Universal analytic properties of noise: introducing the *J*-matrix formalism

2009 J. Phys. A: Math. Theor. 42 365202

(http://iopscience.iop.org/1751-8121/42/36/365202)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **42** (2009) 365202 (15pp) doi:10.1088/1751-8113/42/36/365202

Universal analytic properties of noise: introducing the *J*-matrix formalism

Daniel Bessis and Luca Perotti

Department of Physics, Texas Southern University, Houston, TX 77004, USA

Received 30 November 2008, in final form 12 April 2009
Published 19 August 2009
Online at stacks.iop.org/JPhysA/42/365202

Abstract

We propose a new method in the spectral analysis of noisy time-series data for damped oscillators. From the Jacobi three terms recursive relation for the denominators of the Padé approximations built on the well-known *Z*-transform of an infinite time series, we build a Hilbert space operator, a *J*-operator, where each bound state (inside the unit circle in the complex plane) is simply associated with one damped oscillator while the essential spectrum of the *J*-operator, which lies on the unit circle itself, is shown to represent the noise. Signal and noise are thus clearly separated in the complex plane. For a finite time series of length $2N$, the *J*-operator is replaced by a finite order *J*-matrix J_N, having N eigenvalues which are time reversal covariant. Different classes of input noise, such as blank (white and uniform), Gaussian and pink, are discussed in detail, the *J*-matrix formalism allowing us to efficiently calculate hundreds of poles of the *Z*-transform. Evidence of a universal behavior in the final statistical distribution of the associated poles and zeros of the *Z*-transform is shown. In particular, the poles and zeros tend, when the length of the time series goes to infinity, to a uniform angular distribution on the unit circle. Therefore at finite order, the roots of unity in the complex plane appear to be noise attractors. We show that the *Z*-transform presents the exceptional feature of allowing *lossless undersampling* and how to make use of this property. A few basic examples are given to suggest the power of the proposed method.

PACS number: 07.05.Kf

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Spectral analysis of highly noisy time-series data impacts diverse fields such as gravitational wave detection, nuclear magnetic resonance spectroscopy as applied to nuclear waste, brain/breast cancer detection, oil detection and other similar areas of application.

Experimental time series are always affected by the presence of noise. As long as the signal-to-noise ratio is not too poor, several filters are available to denoise the data within the framework of Fourier analysis and its variants. All such techniques, on the other hand, have drawbacks.

Fourier analysis, being linear, treats noise on the same foot as the signal, and therefore is per se unable to distinguish the two. In the case of stationary uncorrelated noise, the Weiner–Khintchine theorem can be used to denoise the data [1], but it is unable to distinguish the signal from correlated or non-stationary noise.

Wavelets denoising methods require some knowledge of the signal to be found in order to avoid smoothing out the signal itself [2].

Prony's method is tailored to analyze signals composed of damped oscillators, but it is known to be unstable in the presence of noise. This is because—at least in its original formulation—it assumes an 'all-pole' system; in other words, it is equivalent to the construction of a Padé approximant (rational approximation) with a constant numerator. Having no zeros it is unable to model noise, as we shall see in section 2.2.3. Modifications have been proposed to stabilize it [3], but, apart from being cumbersome, no effort has been made to classify the poles so as to distinguish the noise ones from the signal ones.

The matched filtering technique, which calculates cross-correlations between noisy detector outputs and reference waveforms from a 'library', is only useful when the waveforms are well predicted theoretically [4].

Still, all the above techniques fail when the (average power) signal-to-noise ratio approaches 1. It is exactly this very high noise case we deal with here, presenting a denoising method based on the analytic properties of the *Z-transform* (generating function) of the data [5].

This paper is organized as follows.

From a given infinite time series, we build—by way of Padé approximations of its Z-transform—a tridiagonal Hilbert space operator, a J-operator, which, for a finite time series of length $2N$, is replaced by a finite order J-matrix J_N, having N time reversal covariant eigenvalues. Such eigenvalues correspond to the poles of the Z-transform itself.

We then introduce a specific class of signals made of an arbitrary number of damped oscillators. The poles of the Z-transform of such signals are known to be *inside the unit circle* [5].

It is, on the other hand, known that a Taylor series with random coefficients has the unit circle as a natural boundary [11].

For a time series of noisy damped oscillators, the spectrum of our J-operator will therefore comprise a discrete spectrum where each bound state (inside the unit circle) is simply associated with one damped oscillator, and an essential spectrum which represents the noise, and lies on the unit circle in the complex plane.

It is known that for a finite time series, the natural boundary due to the noise is replaced by doublets of poles and zeros [13–16] surrounding the vicinity of the unit circle. We numerically study the statistical distribution of such poles and zeros for *purely* noisy time series, considering different input classes, such as blank (white and uniform), Gaussian and even colored noise. We find compelling evidence of universal behavior. In particular, the poles and zeros tend, when the length of the time series goes to infinity, to the roots of unity in the complex plane which appear to be noise attractors.

As the Fourier transform can be seen as the specialization of the Z-transform to the unit circle, the above result, combined with the knowledge that the poles associated with damped oscillators are *inside* the unit circle, explains the poor behavior of the Fourier transform and the like while analyzing damped oscillators spectral data corrupted by heavy noise.

We then prove that the *Z*-transform presents the exceptional feature of allowing *lossless undersampling*. This allows us to use 'interlaced sampling' to obtain multiple time sequences from a single one and thus improve detection through coincidence.

Finally, we give some simple examples to suggest the power of this approach. As is usually done in nuclear magnetic resonance spectroscopy, we consider one or more damped oscillator signals repeatedly observed. By covering the complex plane with a lattice and counting the 'non-zero paired' poles in each cell (that is, poles which have no associated zero in their vicinity), an approximate evaluation of both the detection probability and the false alarm level for a single observation can be obtained. 'Interlaced sampling' can, on the other hand, significantly improve performance even in the case of a single time sequence.

2. Analytic properties of the *Z*-transform

Given a time series $s_0, s_1, s_2, \ldots, s_n, \ldots$, it is a standard practice [5] to associate with it its *Z-transform* defined as

$$Z(z) = \sum_{n \geqslant 0} s_n z^{-n}. \tag{1}$$

Clearly, due to the boundedness of the time-series coefficients, this function is at least analytic outside the unit disk.

2.1. Construction of the tridiagonal Jacobi matrix

We shall now introduce the Hilbert space operator whose resolvent matrix element is precisely the *Z*-transform.

We are going to deal mainly with the $\left[\frac{n-1}{n}\right](z)$ Padé approximant to the *Z*-transform of a given time series [6] and write

$$\left[\frac{n-1}{n}\right](z) = \frac{N_{n-1}(z)}{D_n(z)}, \tag{2}$$

where $D_n(z)$ is *a monic polynomial of degree n* and $N_{n-1}(z)$ *is a polynomial of degree* $n-1$ in z.

It is known [7] that $D_n(z)$ satisfy the Jacobi three terms recursive relation

$$D_{k+1}(z) = [z - A_k]D_k(z) - R_k D_{k-1}(z) \qquad D_{-1}(z) = 1 \qquad D_0(z) = 1, \tag{3}$$

where

$$A_k = -(a_{2k} + a_{2k+1}) \qquad R_k = a_{2k-1}a_{2k} \qquad k \geqslant 1 \quad a_0 = 0, \tag{4}$$

with a_k being the coefficients of the Stieltjes continued fraction expansion of the *Z*-transform.

This relation can be written as an eigenvalue problem by introducing the tridiagonal $(N+1) \times (N+1)$ Jacobi matrix

$$J_N = \begin{bmatrix} A_0 & 1 & 0 & \ldots & 0 & 0 \\ R_1 & A_1 & 1 & 0 & 0 & 0 \\ 0 & R_2 & A_2 & 1 & \ldots & 0 \\ \ldots & & & & 1 & 0 \\ 0 & \ldots & 0 & R_{N-1} & A_{N-1} & 1 \\ 0 & 0 & \ldots & 0 & R_N & A_N \end{bmatrix}. \tag{5}$$

Then, when $D_{N+1}(z) = 0$, the previous recursive relation can be written as

$$J_N V = zV, \tag{6}$$

where V is the column vector

$$V^T = [D_0(z), D_1(z), \ldots, D_N(z),]. \tag{7}$$

We thus see that the zeros of $D_{N+1}(z)$ are the eigenvalues $(z_0, z_1, \ldots, z_N)$ of the tridiagonal Jacobi matrix of order $N+1$. Therefore, the Z-transform appears to be a matrix element of the resolvant of the J-matrix.

We also introduce the norm of the eigenvector:

$$\rho_k^N = \sum_{p=0}^{p=N} |D_p(z_k)|^2. \tag{8}$$

2.2. *The case of finite time series*

2.2.1. Damped oscillating signals Z-transform. One of the most important applications of the *Z-transform* concerns damped oscillating signals, such as those found in nuclear magnetic resonance data [8] and in burst gravitational waves [9, 10].

Let us start by considering the signal in the absence of noise: for a finite ensemble of damped oscillators, the discretized data will read

$$s_k = \sum_p A_p \, \mathrm{e}^{\mathrm{i}\omega_p \frac{k}{N} T} \qquad k = 0, 1, 2, \ldots, N-1 \quad \omega_p = 2\pi f_p + \mathrm{i}\alpha_p, \tag{9}$$

where A_p, f_p and α_p are the amplitude, the frequency and the damping factor of the pth oscillator. T is the recording time and N is the number of data.

We now consider their *Z-transform* (1) and allow both the recording time T and the number of data N to go to infinity, while keeping T/N constant, so that (1) becomes the (formal) Taylor series

$$Z(z) = \sum_{k=0}^{+\infty} s_k z^{-k}. \tag{10}$$

Combining the previous two equations, we get for the *Z-transform* the expression

$$Z(z) = \sum_{k=0}^{+\infty} s_k z^{-k} = \sum_{k=0}^{+\infty} z^{-k} \sum_p A_p \, \mathrm{e}^{\mathrm{i}\omega_p \frac{k}{N} T} = \sum_p A_p \sum_{k=0}^{+\infty} (z^{-1} \, \mathrm{e}^{\mathrm{i}\omega_p \frac{T}{N}})^k = \sum_p \frac{A_p}{1 - z^{-1} \, \mathrm{e}^{\mathrm{i}\omega_p \frac{T}{N}}}. \tag{11}$$

From this we immediately see that, in this noiseless case, the *Z-transform* is a rational fraction in z, the poles of which are

$$z_p = \mathrm{e}^{\mathrm{i}\omega_p \frac{T}{N}}. \tag{12}$$

Note that all poles have to lie strictly inside the unit circle because $\mathrm{Im}\, \omega_p > 0$. The residues are

$$\rho_p = z_p A_p. \tag{13}$$

Therefore, after identifying the poles z_p and residues ρ_p of the *Z-transform*, we get the quantities of interest, which are the frequencies

$$f_p = \frac{N}{T} \frac{\arg z_p}{2\pi}, \tag{14}$$

the damping factors

$$\alpha_p = \frac{N}{T} \log|z_p| \tag{15}$$

and the amplitudes

$$A_p = \frac{\rho_p}{z_p} \tag{16}$$

of the damped oscillators of the signal.

This straightforward calculation shows that for a sum of oscillating damped signals, the *Z*-transform associated with its time series is a sum of poles in the complex plane. The position of each pole is simply linked to the damping factor and the frequency of each of the oscillators. Also, it is important to note that all these poles lie *strictly* inside the unit disk.

2.2.2. Noise Z-transform. Due to the linearity of the *Z-transform*, the *Z-transform* of a noisy time series will be the sum of the *Z-transform* of the signal plus the *Z-transform* of the noise.

It is therefore of the uttermost importance to get a deep insight into the analytic structure of a purely noisy time series.

The following theorem is the key to the answer to this question.

Steinhaus theorem. A Taylor series with random coefficients has, *with probability* 1 *(that is: except for a set of measure zero),* the unit circle as a natural boundary [11].

2.2.3. The case of noisy data. Combining the previous statements, we get the following simple but hitherto unrecognized fundamental result.

The *Z-transform* of a finite number of damped oscillations is, in the presence of noise, an analytic function outside the unit disk, plus a rational function having a finite number of poles strictly inside the unit disk, each of them representing one damped oscillation.

We can state the following fundamental theorem.

Theorem (Bessis–Perotti). *With any noisy time series can be associated a tridiagonal Hilbert space operator, its J-operator, extension when* $N \to \infty$ *of the previously defined* J_N*. With probability* 1 *(that is: except for a set of input data series of measure zero), the spectrum of this operator is made of two parts:*

An essential spectrum with support on the unit circle. This spectrum is associated with the noise (uncorrelated part of the signal). The corresponding eigenfunctions have the infinite norm. At a finite order, when the infinite J-matrix is truncated, it decomposes into the Froissart poles of the Froissart doublets.

A discrete spectrum, made of a finite number of poles inside the unit circle: each pole represents a component of the signal, which is made of a finite number of damped oscillators. The corresponding eigenfunctions have the finite norm.

2.3. The case of finite time series

Up to now, we have been dealing with time series of infinite length, we now come to the more realistic case, where the time series has a finite length *N*.

The problem of constructing the *Z-transform* $Z(z)$ of such finite time series is a standard problem in Mathematics known as the Padé approximation construction [6].

2.3.1. The special case of non-noisy finite time series. If P is the number of damped oscillators, as soon the length of the time series is $N \geqslant 2P+1$, the diagonal Padé approximant P/P gives an elegant solution to the above problem.

2.3.2. The general case. In this case, Froissart [12] has shown that the natural boundary generated by the noise is approximated by doublets of poles and zeros (Froissart doublets—see e.g. [13–16]) surrounding the vicinity of the unit circle. The mean distance of the members of each doublet is of the order of the noise magnitude and *we conjecture that it decreases exponentially with the length of the signal.*

3. Universality properties of the noise part of the signal

Here, we are mainly interested in studying the statistical distribution (both radial and angular) of the Froissart doublets as a function of the input noise statistical properties and the length N of the time series.

The most *unexpected* result is the universality of the distribution: the radial distribution in the purely noisy case is universally Lorentzian, the phase distribution is uniform (approaching the $[N/2]$ roots of unity when N goes to infinity).

Due to this property of *universality,* the only parameter left is the width of the Lorentzian, which depends on the length N of the time series. A conjecture, based on heuristic arguments, about the functional dependence of this width on N will be provided in the following.

3.1. Experimental evidence for universality

We calculated the radial distribution of the diagonal and subdiagonal Padé approximants for time series of up to 600 data, which allow the construction of Padé approximants up to 300/300.

To test universality, we have been studying seven different kinds of noise:

(1) the complex input data have uniform distribution both in modulus and phase;
(2) the complex input data have uniform distribution inside a square domain;
(3) the complex input data have uniform distribution inside a circle domain;
(4) standard pink noise;
(5) the complex input data have Gaussian distributions;
(6) the complex input data are correlated following an autoregressive moving average (ARMA) model;
(7) the input data are real and have a Gaussian distribution.

Typical results for the *single run* radial distribution of the Padé approximant poles in some of the above seven cases are shown in figure 1: *in all cases the behavior is Lorentzian and centered on* 1.

Figure 2 presents instead the phase distributions of the same Froissart poles: *in all cases the poles are uniformly distributed between* 0 *and* 2π, *the deviation being Gaussian*.

The graphs shown refer to time series of 200 data: the case for which we have run the highest number of tests, varying several times the seeds of the random number generators used for each of the seven kinds of noise.

Analogous results apply to the zeros of the Padé approximants, with the difference that the radial distribution is centered *on a point larger than* 1 which approaches unity when N goes to infinity (at present, we have no explanation for this slightly unsymmetrical behavior between the poles and zeros statistical distributions).

We have recently gained evidence that universality extends to the case of noise statistical distributions having infinite second moment and, even, infinite mean. We shall report the details in a forthcoming paper.

For the very special case of Gaussian uniform noise, a mathematical proof of our results can be found in [17].

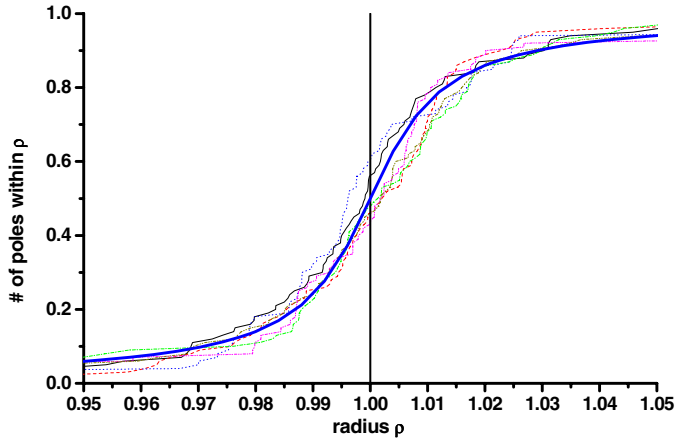

Figure 1. Radial cumulative distribution of the Froissart poles of the [100/100] Padé approximant for different kinds of noise (see the text for a full description). Full line, black: type 1; dash, red: type 2; dot, blue: type 3; dash-dot, green: type 4; dash-dot-dot, magenta and dark yellow: two different cases of type 5. The thick full line is a fitting by a Lorentzian cumulative distribution $((1/\pi)\arctan[(\rho-1)/W]+1/2$, where W is the width of the distribution).

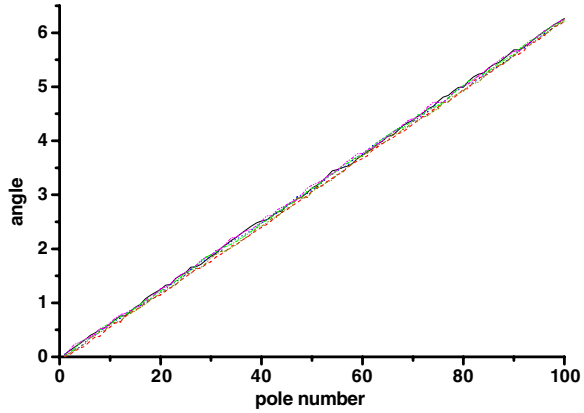

Figure 2. Phase distribution of the Froissart poles of the [100/100] Padé approximant for different kinds of noise (see the text for a full description). Full line, black: type 1; dash, red: type 2; dot, blue: type 3; dash-dot, green: type 4; dash-dot-dot, magenta and dark yellow: two different cases of type 5.

3.2. The width of the universal radial Lorentz distribution

We are left with the problem of finding an expression for the width of the universal radial distribution of both the poles and zeros and for the distance of the center of the radial distribution of the zeros as a function of the time-series length N.

On heuristic arguments, we propose for all three cases the following formula:

$$W \simeq \alpha \frac{\ln(N)}{N}. \tag{17}$$

Figure 3 shows a fit to the above formula of said three quantities. The agreement appears quite good and seems to improve with increasing N.

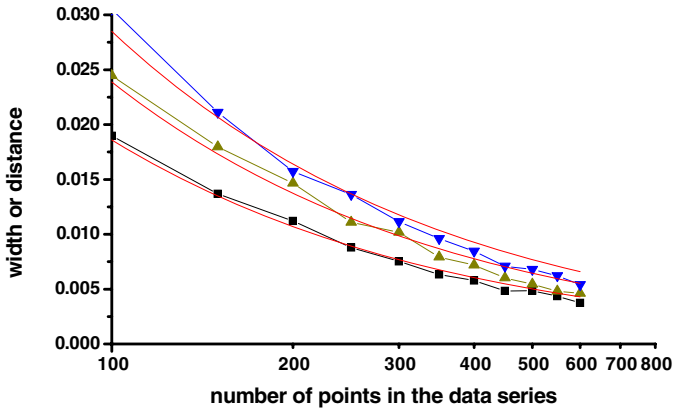

Figure 3. The width of the universal radial distribution of both the poles (squares, black) and zeros (up triangles, dark yellow), and the distance from unity of the center of the radial distribution of the zeros (down triangles, blue) as a function of the time-series length N. Also shown is their fitting to 17. The global multiplicative factor α is, respectively, 0.404 ± 0.005, 0.52 ± 0.01 and 0.62 ± 0.01.

4. First conclusions

Spectral analysis has been dominated by the Fourier transform (and its variants), which is the ideal tool when noise is not too important.

The major drawback of the Fourier transform is its *linearity* which prevents it from distinguishing noise from the signal because the Fourier transform treats the noise on the same foot as the signal, and therefore leaves noise intact.

Furthermore, it is clear that the discrete Fourier transform is nothing but the specialization of the Z-transform to specific points of the complex plane which are the roots of unity. This is dramatic because, as we have shown in our presentation, *the roots of unity are the noise attractors of the complex plane.*

This explains the poor results of data spectral analysis in the presence of heavy noise produced by the FFT.

We therefore suggest to substitute the *Z-transform* for the discrete Fourier transform when analyzing highly corrupted time-series data.

5. Beyond the Fourier analysis: methods for extracting spectra from noisy data

In order to make a clear distinction among the poles of the Z-transform of a noisy time series of length $2N + 2$, and identify the eigenvalues corresponding to the discrete spectrum, we can make use of several properties of the Z-transform. As each method that we are going to present addresses a different property, a comparison of the results of all methods applicable to any specific time series is likely to be the optimal choice.

For some of the methods delineated below, we need to find both poles and zeros of the Z-transform. For the former ones, we compute the $(N+1) \times (N+1)$ Jacobi matrix associated with the denominator of the Z-transform we described in section 2.1 and then diagonalize it. The latter ones can be calculated diagonalizing the $N \times N$ Jacobi matrix for the numerator of the Z-transform Padé approximant, which is simply the denominator Jacobi matrix with the first column and first row removed. The procedure for computing both zeros and poles of the Z-transform therefore consists in the diagonalization of an already tridiagonal matrix;

this makes it computationally extremely efficient and allows us to easily calculate hundreds of poles and zeros.

5.1. Cleaning of Froissart doublets

A first denoising method consists in removing from the complex plane the poles which can be identified as part of Froissart doublets, due to their proximity to zeros of the Z-transform.

This can be done by ordering the poles and zeros in couples in order of increasing distance between the two, and keeping only those whose distance is higher than a given value ϵ (usually $\epsilon \geqslant 0.2$).

5.2. Variational principle

On the other hand, making use of the results of section 2.1, we can consider the Jacobi matrix associated with the denominator of the Z-transform and search for its eigenvalues having *modulus strictly smaller than 1* and for which the norm of the eigenvector has local minima. Such eigenvalues correspond to the frequencies of the signal.

5.3. Stationarity

It is also possible to compare Padé approximants of different orders and look for stable 'non-zero paired' poles inside the unit circle: these will be the signal poles, while the non-stationary poles will be linked to the noise. This method is particularly useful in dealing with poles in the tails of the noise pole distribution.

5.4. Undersampling

Another method makes use of an exceptional feature, never reported up to now, of Padé approximations with respect to undersampling.

Emmanuel Candes [18] has extended the Nyquist–Shannon sampling theorem [19] and shown that it was possible under some circumstances to use bandwidth smaller than those requested by the Nyquist theorem (sufficient condition only but not necessary) without losing information.

It is therefore interesting to look for classes of signals for which the corresponding Z-transforms allow undersampling *without any loss of information. We are going to show that the class of damped oscillators allows lossless undersampling.*

Suppose that in (10), we undersample the time series by taking only s_k with $k = mn$ where m is given fixed, while n runs from zero to infinity to start with. We get the undersampled Z_m-transform

$$Z_m(z) = \sum_{n=0}^{+\infty} s_{mn} z^{-n} \tag{18}$$

or

$$\begin{aligned} Z_m(z) &= \sum_{n=0}^{+\infty} s_{mn} z^{-n} = \sum_{n=0}^{+\infty} z^{-n} \sum_p A_p \mathrm{e}^{\mathrm{i}\omega_p \frac{mn}{N} T} \\ &= \sum_p A_p \sum_{n=0}^{+\infty} (z^{-1}\, \mathrm{e}^{\mathrm{i}m\omega_p \frac{T}{N}})^k \\ &= \sum_p \frac{A_p}{1 - z^{-1}\, \mathrm{e}^{\mathrm{i}m\omega_p \frac{T}{N}}}. \end{aligned} \tag{19}$$

We see that the Padé approximation P/P again gives the exact solution to the undersampled time series. So provided we pick $N = 2P + 1$ equispatiated terms in the time series, we are able to rebuild it exactly.

We point out here that we assume to be dealing with data series such that the periods of the signal oscillations to be reconstructed cover several data points. This means that, as long as m is less than the number of data points per period of each of the signal oscillators, the reconstructed signal frequencies will be given by the mth roots of the poles z_p having the lowest phases.

The above result suggests another technique—which we call 'interlaced sampling'—to improve sensitivity when only a single data time-sequence is available. Taking advantage of the undersampling properties of the Z-transform, we divide the data into m undersampled sequences, the first one comprising the points $1, m + 1, 2m + 1, 3m + 1, \ldots$, the second one comprising the points $2, m + 2, 2m + 2, 3m + 2, \ldots$ and so on. To first order, the ratio of the width of the pole distribution around the unit circle to the distance of the signal pole from it does not change, but instead of a single signal pole, we do now get a cluster of m poles, which are much easier to detect. Preliminary tests have given encouraging results.

6. Application of the method to explicit examples

We shall now proceed to describe some simple examples. The actual procedure we use to reconstruct the signal is very basic, but sufficient for our present aim of showing the potentiality of our proposed system. Plain Fortran was used to write the code implementing the procedure described in the following.

6.1. A first example

As a first example, we consider a damped signal at zero frequency. The chosen damping factor reduces the signal by a factor of 10^{-6} over 300 signal points, resulting in a pole at $(0.952\,38, 0.0)$. Noise is of the type (1) above (the complex input data have uniform distribution both in modulus and phase).

A signal series of 300 points is relatively short. Since the separation of signal and noise is only univocal in the limit of an infinite series, and—moreover—a single pole off the unit circle could be due to the particular noise sequence used, we repeat the sequence several times, as is for example usually done in nuclear magnetic resonance spectroscopy. In the present case, we repeat the sequence 40 times, each time changing the seeds for our noise generator, and search for a *stable* pole inside the unit circle with no zero close to it.

For each of the 40 data sequences we clean most of the Froissart doublets according to the procedure delineated above: we calculate the distance of each pole from all the zeros, arrange the pole-zeros couples in order of increasing distance and keep only the uncoupled pole and the poles whose distance from the coupled zeros is higher than a set value. We find that cut distances between 0.6 and 0.8—which leave us with two or three 'non-zero paired' poles per run—give good results for the range of signal/noise ratios we explored; for higher signal/noise ratios good results can be obtained with much smaller cut distances.

Directly plot of these poles (and their coupled zeros) for all 40 runs, can be misleading to the eye. To better see areas of concentration of poles, we divide the complex plane into a number of square boxes and plot in three dimensions the number of poles minus the number of coupled zeros in each box. For the tests we present, the boxes were of side 0.1.

Figures 4 and 5 show the case of an average power signal/noise ratio (SNR) of approximately -10 db (matched filtering SNR equal to 5.5), using a cut distance of—

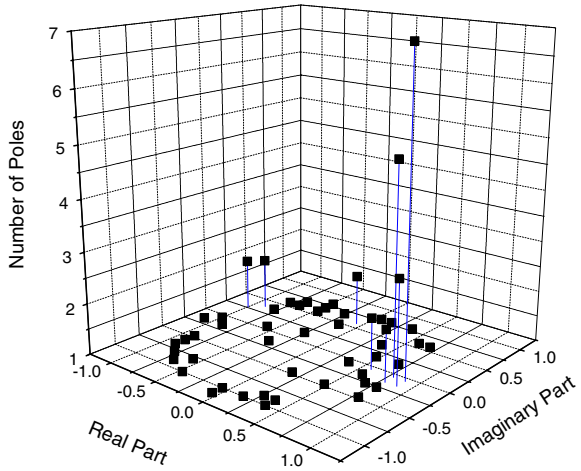

Figure 4. Signal/noise ratio of approximately -10 db. Distribution of the [149/150] Padé approximant poles distant from their coupled zeros more than 0.75. The vertical axis represents the number of poles in each box of side 0.1. The highest peak corresponds to the signal pole at $z = (0.95238, 0.0)$.

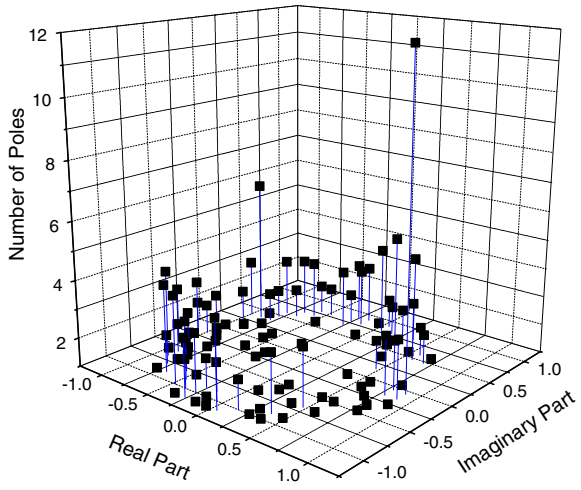

Figure 5. Same as figure 4 but the minimum pole–zero distance has been reduced to 0.2.

respectively—0.75 and 0.2: in both cases the highest peak is at the position of the signal pole. Note that this can be considered a worst case scenario: the signal pole lies close to a 'four corner point'. The reconstructed poles therefore are distributed over four boxes: for a cut distance of 0.75 the four peaks sum to 17, while the height of the highest 'fake' peak is only 2. For a cut distance of 0.2 the four peaks sum to 22, but the height of the highest 'fake' peak is now 6. This suggests that reducing the cut distance this much is not convenient: going from 0.75 to 0.2 improves the detection probability from about 40% to about 50%, but at the same time forces us to increase the false alarm threshold by a factor of 3 (from 2 to 6 out of 40). For the examples that follow we shall therefore show only the results obtained using 0.75 as our cut distance.

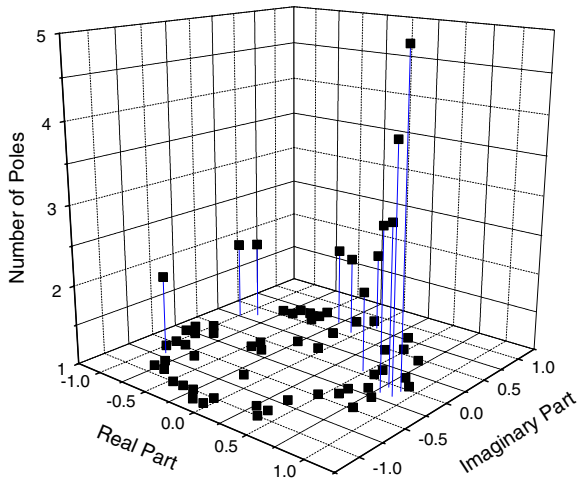

Figure 6. Signal/noise ratio of approximately −15 db. Distribution of the [149/150] Padé approximant poles distant from their coupled zeros more than 0.75. The vertical axis represents the number of poles in each box of side 0.1. The highest peak corresponds to the signal pole at $z = (0.95238, 0.0)$.

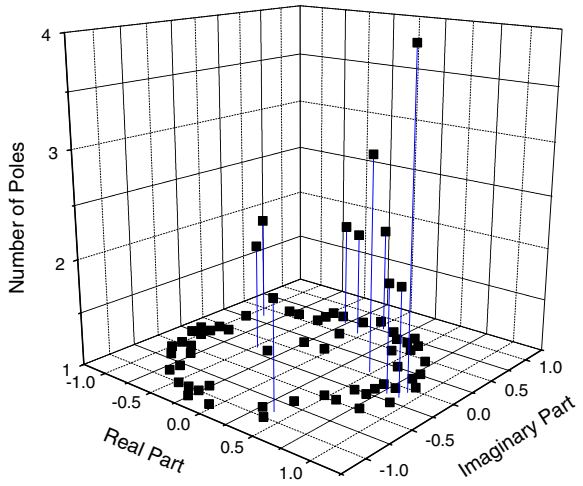

Figure 7. Signal/noise ratio of approximately −18 db. Distribution of the [149/150] Padé approximant poles distant from their coupled zeros more than 0.75. The vertical axis represents the number of poles in each box of side 0.1. The highest peak corresponds to the signal pole at $z = (0.95238, 0.0)$.

Reducing the average signal/noise ratio to approximately −15 db (matched filtering SNR equal to 3.0), the peak is still the highest, but 'fake' peaks far from it become more visible, as can be seen in figure 6. This is mostly due to a higher noise induced spread of the reconstructed poles: if we sum the four signal peaks we get 15, which is not much less than the 17 of the previous case (detection probability is still around 40%), especially considering that the height of the highest 'fake' peak is still 2 which means that we can keep the same false alarm threshold.

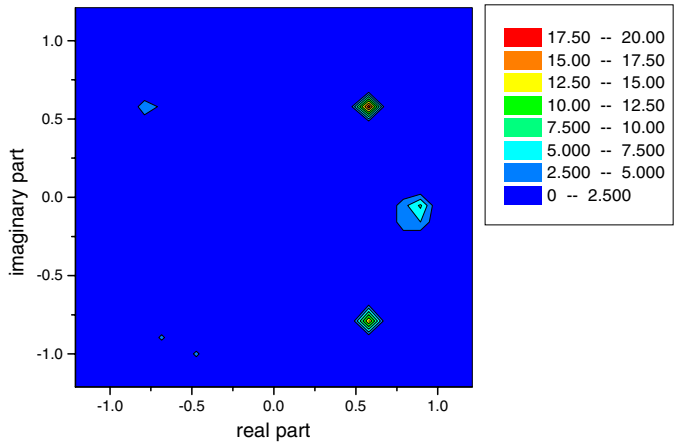

Figure 8. Signal/noise ratio of approximately -15 db for the constant term and of approximately -21 db for the oscillating one. Distribution of the [149/150] Padé approximant poles distant from their coupled zeros more than 0.2. The level curves represent the number of poles in each box of side 0.1. The two highest peaks correspond to the oscillating signal poles at $z = (0.67344, 0.67344)$ and $z = (0.67344, -0.67344)$. The non-oscillating signal pole at $z = (0.95238, 0.0)$ appears to be lower and more spread out.

Figure 7 finally shows the case of an average signal/noise ratio of approximately -18 db (matched filtering SNR equal to 2.2): the peak at the signal pole is still the highest, but peaks far from it (the height of the highest still being 2) are now substantial compared to it and to the sum of the four signal peaks which is now only 9. The false alarm threshold does not need to be changed, but detection probability is now down to 25%.

6.2. A more complicated signal

Adding to the -15 db case an oscillating term, resulting in a couple of complex conjugate poles in $(0.67344, 0.67344)$ and $(0.67344, -0.67344)$ respectively, and having an average signal/noise ratio of approximately -21 db (matched filtering SNR equal to 1.5), gives us figure 8. The oscillating signal, even if weaker, appears to give peaks much sharper than the peak given by the non-oscillating signal, but only because its poles are far from box borders: if we sum the four peaks corresponding to the non-oscillating signal we get 19, the same as the highest oscillating signal peak (the other one is only 15, the asymmetry is due to the noise being complex). Note that the height of the highest 'fake' peak is now 4, which suggests the need for a higher false alarm threshold.

7. General conclusion

There is still plenty of room for improvement: of the methods outlined in section 5 we have used, in the examples given above, only the cleaning of Froissart doublets and the knowledge that the signal poles must be inside the unit circle.

Moreover, we are still not using all the information available. For example, we did not use the residues of the poles, from which the signal amplitude can be obtained, and we have not yet performed a complete optimization over the various parameters used in the analysis. In particular, there is the need for an extensive mapping of the spread of the reconstructed poles as a function of noise level and Padé approximant order. Our last example also suggests

that there might be a connection between the number of signals and the fake alarm threshold, which needs to be investigated, too.

Having said this, we can nonetheless conclude the following.

In our method, the Z-transform appears as an extension of the discrete Fourier transform to complex values of the frequency: in the complex plane, the discrete Fourier transform is the restriction of the Z-transform on values of z located at the roots of unity. However, noise is not uniformly distributed in the complex z-plane and, unfortunately, the roots of unity are noise attractors as shown at the beginning of this paper. Therefore, the Fourier transform is not a good choice when the signal is embedded in high noise as it is the case in many circumstances. The analytic treatment of the noise we propose, distinguishes, in a drastic way, the signal from the noise by their totally different analytic properties. The fact that the noise presents an analytic characterization that is universal, independent of its statistical properties, (the most novel result of our paper) has as consequence a much greater independence of the signal identification on the noise level. In principle, in the limit of very long signals made of data having a large number of digits, the signal extraction becomes independent of the signal-to-noise ratio.

Moreover, the examples we have given show that the Z-transform analysis promises *not only in theory* to be a powerful tool especially for *the frequently encountered case* of *damped signals in a highly noisy environment*, because it gives us a sure signature for the signal by radially separating it in the complex plane from the noise.

Acknowledgments

We thank Professor Marcel Froissart, from College de France, for discussions and suggestions. We thank Professor Carlos Handy, Head of the Physics Department at Texas Southern University, for his support. We thank Professor Bernhard Beckermann of Lille University, France for explaining to us his method of computing Padé approximations avoiding divisions. Special thanks to Professor Mario Diaz, Director of the Center for Gravitational Waves at the University of Texas at Brownsville: without his constant support this work would have never been possible. Supported by subaward CREST to the Center for Gravitational Waves, Texas University at Brownsville, Texas, USA.

References

[1] See, for example Couch L W II 2001 *Digital and Analog Communications Systems* 6th edn (Englewood Cliffs, NJ: Prentice-Hall) pp 406–9

[2] See, e.g., Donoho D L and Johnstone I M 1995 *J. Am. Stat. Assoc.* **90** 1200

[3] See, for example, Osborne M R and Smyth G K 1995 A modified Prony algorithm for fitting sums of exponential functions *SIAM J. Sci. Comput.* **16** 119–38

[4] Wainstein L A and Zubakov V D 1962 *Extraction of Signals from Noise* (Englewood Cliffs, NJ: Prentice-Hall)
Owen B J 1996 *Phys. Rev.* D **53** 6749

[5] Proakis J G and Manolakis D G 2007 *Digital Signal Processing* 4th edn (Upper Saddle River, NJ: Pearson) chapter 3

[6] Baker G A 1975 *Essentials of Pade Approximants* (New York: Academic)

[7] Baker G A and Graves-Morris P 1981 *Padé Approximants, Encyclopedia of Mathematics and its Applications* (Reading, MA: Addison-Wesley)

[8] Belkić D and Belkić K 2006 In vivo magnetic resonance spectroscopy by the fast Padé transform *Phys. Med. Biol.* **51** 1049–75

[9] Echeverria F 1989 *Phys. Rev.* D **40** 3194
Finn L S 1992 *Phys. Rev.* D **46** 5236
Creighton J D E 1999 *Phys. Rev.* D **60** 022001

[10] Goggin L M (for the LIGO Scientific Collaboration) 2006 *Class. Quantum Grav.* **23** S709–S713
[11] Steinhaus H 1929 Uber die Wahrscheinlichkeit dafuer dass der Konvergenzkreis einer Potenzreihe ihre natuerliche Grenze ist *Math. Z.* **31** 408–16
[12] Froissart M 1987 private communication
[13] Gilewicz J and Truong-Van 1988 Froissart doublets in Padé approximants and noise *Constructive Theory of Functions 1987* (Sofia: Bulgarian Academy of Sciences) pp 145–51
[14] Fournier J-D, Mantica G, Mezincescu A and Bessis D 1993 Universal statistical behavior of the complex zeros of Wiener transfer functions *Europhys. Lett.* **22** 325–31
[15] Fournier J-D, Mantica G, Mezincescu A and Bessis D 1995 Statistical properties of the zeros of the transfer functions in signal processing *Chaos and Diffusion in Hamiltonian Systems* ed D Benest and C Froeschle (Gif-sur-Yvette: Editions Frontières)
[16] Bessis D 1996 Padé approximations in noise filtering *International Congress on Computational and Applied Mathematics J. Comput. Appl. Math.* **66** 85–8
[17] Barone P 2005 *J. Approx. Theory* **132** 224–40
[18] Candes E J, Romberg J and Tao T 2004 Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information *IEEE Trans. Inf. Theory* **52** 489–509
[19] Shannon C E 1998 Communication in the presence of noise *Proc. IEEE* **86** 447–57