APP Template V1.57 Article id: ps509780 Typesetter: MPS Date received by MPS: 11/02/2015 PE: MAC003876 CE: LE: LE0018

IOP Publishing | Royal Swedish Academy of Sciences Physica Scripta

Phys. Scr. **00** (2015) 000000 (15pp)

Non-commutative tomography and signal processing

R Vilela Mendes[1,2]

Centro de Matemática e Aplicações Fundamentais, Univ. Lisbon

E-mail: rvmendes@fc.ul.pt and rvilela.mendes@gmail.com

Received 11 October 2014, revised 14 January 2015
Accepted for publication 1 February 2015
Published DD MM 2014

Abstract

Non-commutative tomography is a technique originally developed and extensively used by Profs. M. A. Man'ko and V. I. Man'ko in quantum mechanics. Because signal processing deals with operators that, in general, do not commute with time, the same technique has a natural extension to this domain. Here, a review is presented of the theory and some applications of non-commutative tomography for time series as well as some new results on signal processing on graphs.

Keywords: signal processing, transforms, tomograms

1. Introduction

As an alternative to the description of quantum systems by wave functions or density matrices, Wigner [1] introduced in 1932 the function $W(q, p)$, which contains all information on the quantum state and is similar to the classical probability density $f(q, p)$ in phase space. Besides taking negative values, the amplitude of the oscillating cross-terms in the Wigner function may be large in position-momentum regions that carry no physical information. Therefore the Wigner function cannot be interpreted as a probability distribution in phase-space.

However, by using a generalization of the Radon transform [2, 3], a squared amplitude of the projection on the generalized eigenvalues of a linear combination of two non-commutative operators, it was suggested that quantum states might be identified with tomographic probability distributions [4]. This provided for quantum mechanics a description alternative to the one given by wave functions or density matrices. For reviews of this approach to quantum mechanics, refer to [5–7].

The reason why the Wigner function cannot provide a probability distribution in phase-space lies in the fact that position q and momentum p are non-commuting operators. Likewise, the Wigner function for the pair (t, ω), which in signal processing goes under the name of Wigner–Ville transform [8], cannot provide a true probability distribution in the time-frequency plane, because time t does not commute with frequency $\omega = \mathrm{i}\frac{\mathrm{d}}{\mathrm{dt}}$. Hence, serious ambiguity problems may arise in the interpretation of the Wigner–Ville transform, for example, in the analysis of radar signals. The same problem occurs for other time-frequency quasidistributions [9, 10], even for those that are strictly positive [11, 12] but cannot be interpreted as actual probability distributions in the plane of two non-commutative operators.

This situation suggested that also for the pair (t, ω), and in general for signal processing applications, the tomographic point of view might be useful. This program was indeed developed [13–18]. Some of the main results are briefly summarized in section 2. The approach has been quite successful in signal processing applications, in particular to separate the components of complex signals. Two examples of component separation by the non-commutative tomography technique are presented in section 3. Section 4 discusses the application of the tomographic technique when, instead of using either the frequency, the dilation, or another mathematical known operator, the operator itself is constructed from experimental or simulated data. In this way the tomogram becomes sensitive to the very features of the signal that one wants to detect. Finally, the last section contain some recent new results on the application of the technique to signal processing on graphs.

[1] Av. Gama Pinto 2, 1649-003 Lisboa.
[2] Also at Instituto de Plasmas e Fusão Nuclear—IST.

2. Non-commutative tomography and signal processing

A unified framework to characterize linear transforms, quasidistributions, and tomograms was developed in [15]. To fix notation we briefly review it here. Signals $f(t)$ are considered as vectors $|f\rangle$ in a subspace $\mathcal{N}$ of a Hilbert space $\mathcal{H}$ with dual space $\mathcal{N}^*$. Then a family of unitary operators $U(\alpha) = e^{iB(\alpha)}$, α being a label $\{\alpha \in I, I \subset \mathbb{R}^n\}$, is defined on $\mathcal{N}^*$. Using a ket-bra notation we denote $|f\rangle \in \mathcal{N}$ and $\langle f| \in N^*$. In this setting three types of integral transforms are constructed. Let $h \in \mathcal{N}^*$ be a reference vector, and let U be such that the linear span of $\left\{U(\alpha)h \in \mathcal{N}^* : \alpha \in I\right\}$ is dense in $\mathcal{N}^*$. In $\{U(\alpha)h\}$, a complete set of vectors can be chosen to serve as the basis.

1—Linear transforms

$$W_f^{(h)}(\alpha) = \left\langle U(\alpha)h|f\right\rangle \tag{1}$$

2—Quasi-distributions

$$Q_f(\alpha) = \left\langle U(\alpha)f|f\right\rangle \tag{2}$$

3—Tomograms

Given a unitary $U(\alpha) = e^{iB(\alpha)}$, $B(\alpha)$ has the spectral decomposition $B(\alpha) = \int XP(X)dX$. Let

$$P(X) = |X\rangle\langle X|$$

denote the projector on the (generalized) eigenvector $\langle X| \in \mathcal{N}^*$ of $B(\alpha)$. The tomogram is

$$\begin{aligned} M_f^{(B)}(X) &= \langle f|\, P(X)\, |f\rangle \\ &= \langle f|X\rangle\langle X|f\rangle = |\langle X|f\rangle|^2. \end{aligned} \tag{3}$$

The tomogram $M_f^{(B)}(X)$ is the squared amplitude of the projection of the signal $|f\rangle \in \mathcal{N}$ on the eigenvector $\langle X| \in \mathcal{N}^*$ of the operator $B(\alpha)$. Therefore, it is positive. For normalized $|f\rangle$,

$$\langle f|f\rangle = 1$$

the tomogram is normalized

$$\int M_f^{(B)}(X)\, dX = 1. \tag{4}$$

and may be interpreted as a probability distribution on the set of generalized eigenvalues of $B(\alpha)$, that is, as the probability distribution for the random variable X corresponding to the observable defined by the operator $B(\alpha)$.

The tomogram is a homogeneous function

$$M_f^{(B/p)}(X) = |p|\, M_f^{(B)}(pX). \tag{5}$$

Examples:

If $U(\alpha)$ is an unitary operator, generated by $B_F\left(\vec{\alpha}\right) = \alpha_1 t + i\alpha_2 \frac{d}{dt}$, and h is a (generalized) eigenvector of the time-translation operator, the linear transform $W_f^{(h)}(\alpha)$ is the Fourier transform. For the same $B_F\left(\vec{\alpha}\right)$, the quasi-distribution $Q_f(\alpha)$ is the ambiguity function, and the Wigner–Ville transform [1, 8] is the quasi-distribution $Q_f(\alpha)$ for the following B-operator

$$\begin{aligned} B^{(WV)}(\alpha_1, \alpha_2) = &-i2\alpha_1 \frac{d}{dt} - 2\alpha_2 t \\ &+ \frac{\pi\left(t^2 - \frac{d^2}{dt^2} - 1\right)}{2}. \end{aligned} \tag{6}$$

The wavelet transform is $W_f^{(h)}(\alpha)$ for $B_W\left(\vec{\alpha}\right) = \alpha_1 D + i\alpha_2 \frac{d}{dt}$, D being the dilation operator $D = -\frac{1}{2}\left(it\frac{d}{dt} + i\frac{d}{dt}t\right)$. The wavelets $h_{s,\tau}(t)$ are kernel functions generated from a basic wavelet $h(\tau)$ by means of a translation and a rescaling ($-\infty < \tau < \infty$, $s > 0$):

$$h_{s,\tau}(t) = \frac{1}{\sqrt{s}} h\left(\frac{t-\tau}{s}\right) \tag{7}$$

using the operator

$$U^{(A)}(\tau, s) = \exp\left(i\tau\hat{\omega}\right) \exp\left(i\log sD\right), \tag{8}$$

$$h_{s,\tau}(t) = U^{(A)\dagger}(\tau, s)h(t). \tag{9}$$

$\hat{\omega}$ being the operator $i\frac{d}{dt}$.

The Bertrand transform [19, 20] is the quasi-distribution $Q_f(\alpha)$ for B_W. Linear, bilinear, and tomogram transforms are related to one another (see [15]).

2.1. Tomograms: some examples

Tomograms are obtained from projections on the eigenstates of the B operators. These operators may be linear combinations of different (commuting or noncommuting) operators,

$$B = \mu O_1 + \nu O_2$$

meaning that the tomogram explores the signal along lines in the plane $\left(O_1, O_2\right)$. For example, for

$$B(\mu, \nu) = \mu t + \nu\omega = \mu t + i\nu\frac{d}{dt}$$

the tomogram is the expectation value of a projection operator with support on a line in the time–frequency plane

$$X = \mu t + \nu\omega. \tag{10}$$

Therefore, $M_f^{(S)}(X, \mu, \nu)$ is the marginal distribution of the variable X along this line in the time–frequency plane. The line is rotated and rescaled when one changes the parameters μ and ν. In this way, the whole time–frequency plane is sampled, and the tomographic transform contains all the information on the signal. Instead of marginals collected along straight lines on the time–frequency plane, one may use other curves to sample this space [15].

For the tomograms associated to the generators of the conformal group, the B operators are:

Time–frequency

$$B_1 = \mu t + \mathrm{i}\nu\frac{\mathrm{d}}{\mathrm{d}t} \tag{11}$$

Time–scale

$$B_2 = \mu t + \mathrm{i}\nu\left(t\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{2}\right) \tag{12}$$

Frequency–scale

$$B_3 = \mathrm{i}\mu\frac{\mathrm{d}}{\mathrm{d}t} + \mathrm{i}\nu\left(t\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{2}\right) \tag{13}$$

Time-conformal

$$B_4 = \mu t + \mathrm{i}\nu\left(t^2\frac{\mathrm{d}}{\mathrm{d}t} + t\right). \tag{14}$$

The construction of the tomograms reduces to the calculation of the generalized eigenvectors of each one of these B_i operators

$$B_1\psi_1(\mu, \nu, t, X) = X\psi_1(\mu, \nu, t, X)$$

$$\psi_1(\mu, \nu, t, X) = \exp\left[\mathrm{i}\left(\frac{\mu t^2}{2\nu} - \frac{tX}{\nu}\right)\right] \tag{15}$$

with normalization

$$\int \mathrm{d}t\psi_1^*(\mu, \nu, t, X)\psi_1\left(\mu, \nu, t, X'\right) = 2\pi\nu\delta\left(X - X'\right) \tag{16}$$

$$B_2\psi_2(\mu, \nu, t, X) = X\psi_2(\mu, \nu, t, X)$$

$$\psi_2(\mu, \nu, t, X) = \frac{1}{\sqrt{|\,t\,|}}\exp\left[\mathrm{i}\left(\frac{\mu t}{\nu} - \frac{X}{\nu}\log\,|\,t\,|\right)\right] \tag{17}$$

$$\int \mathrm{d}t\psi_2^*(\mu, \nu, t, X)\psi_2\left(\mu, \nu, t, X'\right) = 4\pi\nu\delta\left(X - X'\right) \tag{18}$$

$$B_3\psi_3(\mu, \nu, \omega, X) = X\psi_3(\mu, \nu, \omega, X)$$

$$\psi_3(\mu, \nu, t, X) = \exp\left[-\mathrm{i}\left(\frac{\mu}{\nu}\omega - \frac{X}{\nu}\log\,|\omega|\right)\right] \tag{19}$$

$$\int \mathrm{d}\omega\psi_1^*(\mu, \nu, \omega, X)\psi_1\left(\mu, \nu, \omega, X'\right) = 2\pi\nu\delta\left(X - X'\right) \tag{20}$$

$$B_4\psi_4(\mu, \nu, t, X) = X\psi_4(\mu, \nu, t, X)$$

$$\psi_4(\mu, \nu, t, X) = \frac{1}{|\,t\,|}\exp\left[\mathrm{i}\left(\frac{X}{\nu t} + \frac{\mu}{\nu}\log\,|\,t\,|\right)\right] \tag{21}$$

$$\int \mathrm{d}t\psi_4^*(\mu, \nu, t, s)\psi_4\left(\mu, \nu, t, s'\right) = 2\pi\nu\delta\left(s - s'\right). \tag{22}$$

Then, the tomograms are:

Time–frequency tomogram

$$M_1(\mu, \nu, X) = \frac{1}{2\,\pi\,|\nu|}\left|\int \exp\left[\frac{\mathrm{i}\mu t^2}{2\,\nu} - \frac{\mathrm{i}tX}{\nu}\right]f(t)\mathrm{d}t\right|^2 \tag{23}$$

Time–scale tomogram

$$M_2(\mu, \nu, X) = \frac{1}{2\pi\,|\nu|}\left|\int \mathrm{d}t\,\frac{f(t)}{\sqrt{|t|}}\mathrm{e}^{\left[i\left(\frac{\mu}{\nu}t - \frac{X}{\nu}\log|t|\right)\right]}\right|^2 \tag{24}$$

Frequency–scale tomogram

$$M_3(\mu, \nu, X) = \frac{1}{2\pi\,|\nu|}\left|\int \mathrm{d}\omega\,\frac{\tilde{f}(\omega)}{\sqrt{|\omega|}}\mathrm{e}^{\left[-\mathrm{i}\left(\frac{\mu}{\nu}\omega - \frac{X}{\nu}\log|\omega|\right)\right]}\right|^2 \tag{25}$$

$\tilde{f}(\omega)$ being the Fourier transform of $f(t)$

Time-conformal tomogram

$$M_4(\mu, \nu, X) = \frac{1}{2\pi\,|\nu|}\left|\int \mathrm{d}t\,\frac{f(t)}{|t|}\mathrm{e}^{\left[\mathrm{i}\left(\frac{X}{\nu t} + \frac{\mu}{\nu}\log|t|\right)\right]}\right|^2 \tag{26}$$

The tomograms M_1, M_2, and M_4 interpolate between the (squared) time signal ($\nu = 0$) and its projection on the $\psi_i(\mu, \nu, t, X)$ functions for $\mu = 0$.

For the particular case of the time-frequency tomogram, the projection being on a linear chirp basis, one deals for each fixed X with the modulus squared of a fractional Fourier transform.

In a similar way, tomograms may be constructed for any operator of the general type

$$B_4 = \mu t + \mathrm{i}\nu\left(g(t)\frac{\mathrm{d}}{\mathrm{d}t} + \frac{1}{2}\frac{\mathrm{d}g(t)}{\mathrm{d}t}\right)$$

the generalized eigenvectors being

$$\psi_g(\mu, \nu, t, X) = |\,g(t)\,|^{-1/2} \times \exp\left[\mathrm{i}\left(-\frac{X}{\nu}\int^t\frac{\mathrm{d}s}{g(s)} + \frac{\mu}{\nu}\int^t\frac{s\mathrm{d}s}{g(s)}\right)\right].$$

When dealing with finite-time signals and finite-time tomograms, some normalization modifications are needed. For example, for a time-frequency tomogram, instead of (23), we consider the finite-time tomogram, for a signal defined from t_0 to $t_0 + T$ with $\mu = \cos\theta$ and $\nu = \sin\theta$

$$M_1(\theta, X) = \left|\int_{t_0}^{t_0+T} f^*(t)\psi_{\theta,X}^{(1)}(t)\,\mathrm{d}t\right|^2 = \left|\left\langle f\,|\psi^{(1)}\right\rangle\right|^2 \tag{27}$$

with

$$\psi_{\theta,X}^{(1)}(t) = \frac{1}{\sqrt{T}}\exp\left(\frac{\mathrm{i}\cos\theta}{2\,\sin\theta}t^2 - \frac{\mathrm{i}X}{\sin\theta}t\right). \tag{28}$$

θ is a parameter that interpolates between the time and the frequency operators, running from 0 to $\pi/2$, whereas X is allowed to be any real number. An orthonormalized set of $\psi_{\theta,X}^{(1)}(t)$ vectors is obtained by choosing the sequence

$$X_n = X_0 + \frac{2n\pi}{T}\sin\theta \quad n \in \mathbb{Z}. \tag{29}$$

Likewise, for the finite-time time–scale tomogram $M_2(\mu, \nu, X)$ (24) and the finite-time time-conformal

tomogram $M_4(\mu, \nu, X)$ (26):

$$M_2(\theta, X) = \left| \int_{t_0}^{t_0+T} f^*(t)\psi^{(2)}_{\theta,X}(t)\,\mathrm{d}t \right|^2 = \left| \left\langle f | \psi^{(2)} \right\rangle \right|^2 \tag{30}$$

$$\psi^{(2)}_{\theta,X}(t) = \frac{1}{\sqrt{\log\ |t_0+T| - \log\ |t_0|}} \times \frac{1}{\sqrt{|t|}} \exp\left[\mathrm{i}\left(\frac{\cos\theta}{\sin\theta} t - \frac{X}{\sin\theta} \log\ |t| \right) \right] \tag{31}$$

$$X_n = X_0 + \frac{2n\pi}{\log\ |t_0+T| - \log\ |t_0|} \sin\theta \quad n \in \mathbb{Z} \tag{32}$$

and

$$M_4(\theta, X) = \left| \int_{t_0}^{t_0+T} f^*(t)\psi^{(4)}_{\theta,X}(t)\,\mathrm{d}t \right|^2 = \left| \left\langle f | \psi^{(4)} \right\rangle \right|^2 \tag{33}$$

$$\psi^{(4)}_{\theta,X}(t) = \sqrt{\frac{t_0(t_0+T)}{T}} \frac{1}{|t|} \times \exp \mathrm{i}\left(\frac{\cos\theta}{\sin\theta} \log\ |t| + \frac{X}{t\sin\theta} \right). \tag{34}$$

$$X_n = X_0 + \frac{t_0(t_0+T)}{T} 2\pi n \sin\theta \quad n \in \mathbb{Z}. \tag{35}$$

Tomograms performing a signal analysis on two variables (time and frequency, time and scale, etc.) are more powerful than linear transforms, like Fourier or wavelets. On the other hand, in contrast with the bilinear transforms (Wigner–Ville, Bertrand, etc) and because of their rigorous probabilistic interpretation, they provide a robust and unambiguous characterization of the signals.

3. Component separation and noise suppression

The non-commutative tomography technique has already been used in the identification of biological signals [21], reflectometry in plasma physics [22–25], and astrophysical data [26]. Here I will emphasize the component separation feature.

Most natural and man-made signals are nonstationary and may be thought of as having a multicomponent structure. Bat echolocation, whale sounds, radar, sonar, and many others are examples of this kind of signal. The notion of nonstationarity is easy to define. However, the concept of signal component is not so clearly defined. Because time and frequency descriptions are standard methods of signal analysis, many authors have attempted to base the characterization of signal components on the analysis of the time–frequency plane. There is a large class of time–frequency signal representations (TFR). An important set of such TFRs is Cohen's class [9], obtained by convolutions with the Wigner distribution. Once one particular TFR of the signal is constructed, the search for components may be done by looking for amplitude concentrations in the time–frequency plane. This is the methodology that has been followed by most authors, the notions of instantaneous frequency and instantaneous bandwidth playing an important role in these studies.

An important drawback of the use of TFRs is the fact that they may have negative terms, cross terms, or lack the correct marginal properties in time and frequency. Even if, by the choice of a clever kernel or a smoothing or filtering operation, the TFRs are apparently free from these problems, there is no guarantee that they are free from artifacts that might lead to unwarranted inferences about the signal properties. As stated before, this is a consequence of the basic fact that time (t) and frequency $\left(\omega = \mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}\right)$, being associated to a pair of noncommuting operators, there can never be a joint probability distribution in the time–frequency plane.

Notice that an approach to *component separation* starts from the insight that the notion of *component* depends as much on the observer as on the observed object. That is, when we speak about a component of a signal, we are in fact referring to a particular feature of the signal that we want to emphasize. For example, if time and frequency are the features that interest us, they might indeed be the salient features in the time–frequency plane that should be identified as components. However, if frequency and fractality (scale) interest us, the notion of component and the nature of the decomposition would be completely different. In general, the features of interest correspond to incompatible notions (that is, to noncommuting operators). Therefore, to look for robust characterizations in a joint feature plane is a difficult task, because the noncommutativity of the operators precludes the existence of joint probability densities. Instead, in the tomographic approach, one considers spectral decompositions using the eigenvectors of linear combinations of the operators. The sum of the squares of the signal projections on these eigenvectors having the same norm as the signal, this approach provides an exact probabilistic interpretation.

Once a tomogram for a linear combination of operators O_1 and O_2 $\left(B = \mu O_1 + \nu O_2\right)$ is constructed, what one obtains in the $(X, (\mu, \nu))$ (hyper-) plane is an image of the probability flow from the O_1-description of the signal to the O_2-description, through all the intermediate steps of the linear combination. In contrast with the time–frequency representations, we need not worry about cross-terms or artifacts, thanks to the exact probability interpretation of the tomogram. Then, a *component of the signal* may be defined as any distinct feature (ridge, peak, etc) of the probability distribution in the $(X, (\mu, \nu))$ (hyper-) plane. It is clear that this notion of component is contingent on the choice of the pair $\left(O_1, O_2\right)$.

Here, the method of component separation will be described in some detail for the time–frequency tomogram

case, $M_1(\theta, X)$ (27).

$$M_1(\theta, X) = \left| \left\langle f | \psi^{(1)} \right\rangle \right|^2 \tag{36}$$

with

$$\psi^{(1)}_{\theta,X}(t) = \frac{1}{\sqrt{T}} \exp\left(\frac{\mathrm{i}\cos\theta}{2\sin\theta} t^2 - \frac{\mathrm{i}X}{\sin\theta} t \right). \tag{37}$$

$\theta$is a parameter that interpolates between the time and the frequency operators, running from 0 to $\pi/2$, whereas X is allowed to be any real number. Notice that the $\psi^{(1)}_{\theta,X}$ are generalized eigenfunctions of $B_1(\theta) = t\cos\theta + \mathrm{i}\sin\theta \frac{\mathrm{d}}{\mathrm{d}t}$ for a spectral value X. Therefore $M_1(\theta, X)$ is a (positive) probability distribution as a function of X for each θ. From an abstract point of view, since for different θ the $U(\theta) = \mathrm{e}^{\mathrm{i}B_1(\theta)}$ are unitarily equivalent operators, all the tomograms share the same information. An orthonormalized set of $\psi^{(1)}_{\theta,X}(t)$ vectors is obtained by choosing the sequence (29),

$$\left\langle \psi^{(1)}_{\theta,X_m}(t) | \psi^{(1)}_{\theta,X_n}(t) \right\rangle = \delta_{m,n}. \tag{38}$$

X_0 is freely chosen. (In general, one takes $X_0 = 0$, but it is possible to make other choices, depending on what is more suitable for the signal under study.)

A glance at the shape of the functions (28) shows that the nodes (the zero crossings) t_n of the real (resp. imaginary) part of $\psi^{(1)}_{\theta,X_m}$ are the solutions of

$$\frac{\cos\theta}{2\sin\theta} t_n^2 - \frac{X}{\sin\theta} t_n = 2\pi n \quad (resp.\ \ 2\pi n + \pi/2). \tag{39}$$

It is clear that $\left| t_{n+1} - t_n \right|$ scales as $\sqrt{n}$ and that, for fixed θ, the oscillation length, at a given t, decreases when $|X|$ increases. As a result, the projection of the signal on $\left\{ \psi^{(1)}_{\theta,X_n}(t) \right\}$ will locally explore different scales. On the other hand, changing θ will modify the first term of (39) in such a way that the local time–scale is larger when θ becomes larger, in agreement with the uncertainty principle.

The projections of the signal $f(t)$

$$c^{\theta}_{X_n}(f) = \left\langle f | \psi^{(1)}_{\theta,X_n}(t) \right\rangle \tag{40}$$

are then used for signal processing purposes. In particular, a natural choice for denoising consists in eliminating the $c^{\theta}_{X_n}(s)$ such that

$$\left| c^{\theta}_{X_n}(f) \right|^2 \leqslant \epsilon \tag{41}$$

for some chosen threshold ϵ, the remainder being used to reconstruct a denoised signal. In this case a proper choice of θ is an important issue in the method.

The multi-component analysis is done by selecting subsets $\mathcal{F}_k$ of the X_n and reconstructing partial signals (k-components) by restricting the sum to

$$f_k(t) = \sum_{n \in \mathcal{F}_k} c^{\theta}_{X_n}(f) \psi^{(1)}_{\theta,X_n}(t) \tag{42}$$

for each k. By an appropriate choice of θ, it is possible to use this technique to disentangle the different components of a signal.

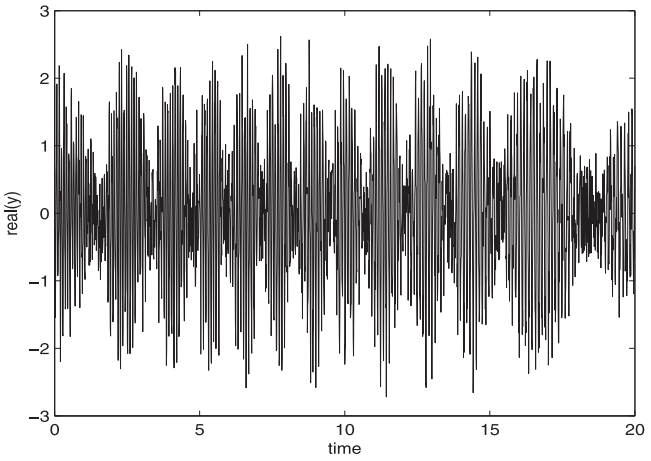

Figure 1. Temporal representation of the real part $\mathcal{R}[y(t)]$ of the simulated data defined by equation (45).

3.1. Examples

Here the general method is applied to two examples: one that shows how two nonlinear chirps, overlapping both in time and frequency, may be correctly separated; the other concerns component separation in reflectometry signals of plasma diagnosis [22, 23].

3.1.1. Separation of a composite chirp signal. Here one analyzes the decomposition into elementary components of a signal which mimics, in a simplified way, the case of an incident plus a reflected wave delayed in time and with an acquired time-dependent change in phase. This would be a typical situation when a signal is sent to probe some environment. In this case the simulated signal $y(t)$ is the sum of an 'incident' chirp $y_0(t)$ and a 'deformed reflected' chirp $y_R(t)$. White noise is added to the signal. The incident chirp is:

$$y_0(t) = \mathrm{e}^{\mathrm{i}\Phi_0(t)} \tag{43}$$

with $\Phi_0(t) = a_0 t^2 + b_0 t$.

The 'instantaneous frequency' of $y_0(t)$ sweeps linearly from 75 rd s^{-1} to 50 rd s^{-1} during $20s$. Its phase derivative is linearly dependent on time: $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_0(t) = 2a_0 t + b_0$. The 'reflected' signal $y_R(t)$ is delayed by $t_R = 3s$ from the incident one and continuously sweeps from 75 rd s^{-1} to 50 rd s^{-1}:

$$y_R(t) = \mathrm{e}^{\mathrm{i}\Phi_R(t)} \tag{44}$$

with $\Phi_R(t) = a_R(t - t_R)^2 + b_R(t - t_R) + 10(t - t_R)^{\frac{3}{2}}$. For $y_R(t)$ the phase derivative $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_R(t)$ is no longer a linear function. The signal $y_R(t)$ is zero during the first $3s$ seconds and ends at $t = 23s$. Finally, the signal to be analyzed is:

$$y(t) = y_0(t) + y_R(t) + b(t). \tag{45}$$

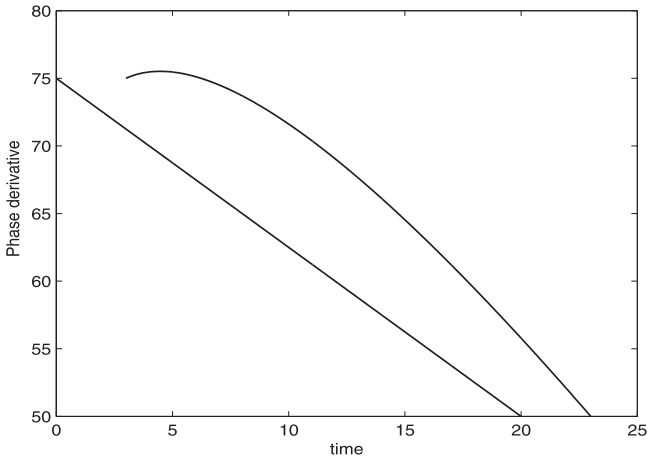

Figure 2. Representation of $F_o = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_0(t)$ and $F_r = \frac{\mathrm{d}}{\mathrm{d}t}\Phi_R(t)$ as a function of time.

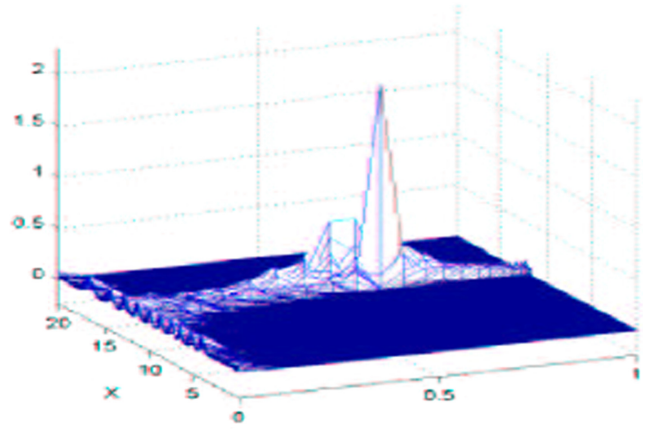

Figure 3. Contour plot of the tomogram for the data defined by equation (45).

The signal-to-noise ratio, $SNR(y, b)$, is 15 dB. The real $\mathcal{R}[y(t)]$ part of this signal is shown in figure 1.

Figure 2 shows $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_0(t)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\Phi_R(t)$ as a function of time. Notice that, except for the three first seconds, there is an almost complete overlap of the 'instantaneous frequency' of the signals $y_0(t)$ and $y_R(t)$.

The tomogram of the first $20s$ of $y(t)$, $M_y(\theta, X) = \left|\langle y|\psi^{(1)}_{\theta,X}(t)\rangle\right|^2$ has a maximum at $\sin(\theta) \approx 0.625$ (figure 3) corresponding to the 'incident' part of the signal that mainly projects in the unique $\psi^{(1)}_{\theta,X}(t)$ that matches $\Phi_0(t)$. One takes the value of $\sin(\theta) \approx 0.625$ to carry out the separation of $y(t)$ in its components. As seen from the tomogram, the signals overlapping both in time and frequency separate in this θ−region. The corresponding spectrum $c^{\theta}_{X_n}(y)$ is shown in figure 4. Based on this spectrum the signal is decomposed into two components. From the first

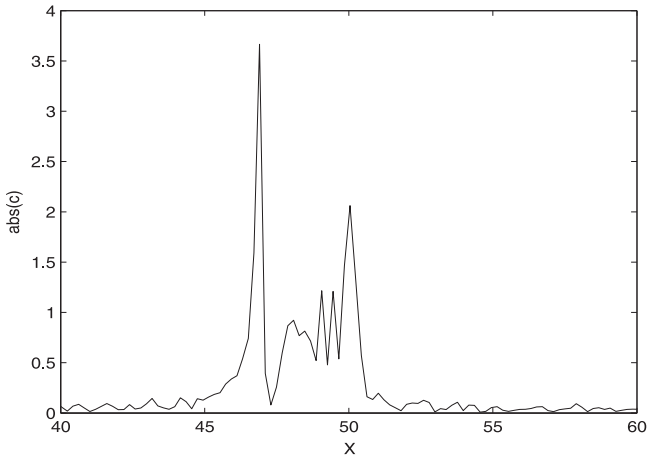

Figure 4. Spectrum of the signal for $\sin(\theta) = 0.625$.

component, the 'incident' chirp $y_0(t)$ is reconstructed by:

$$\tilde{y}_0(t) = \sum_{X_n=45}^{47.25} c^{\theta_0}_{X_n}(y)\psi^{(1)}_{\theta,X_n}(t). \tag{46}$$

The quadratic error, between $\tilde{y}_0(t)$ and $y_0(t)$, $E(y_0, \tilde{y}_0)$, is −9.5 dB.

From the second spectral component, the 'reflected' chirp is given by:

$$\tilde{y}_R(t) = \sum_{X_n=47.5}^{50.5} c^{\theta}_{X_n}(y)\psi^{(1)}_{\theta,X_n}(t). \tag{47}$$

In this case the quadratic error $E(y_R, \tilde{y}_R)$ is −10 dB. This may be compared with a quadratic error $E(y, \tilde{y})$ of −29 dB for the total signal reconstructed from the spectral projection corresponding to $45 < X_n < 50.5$.

A signal that appears completely mixed, both in time and in the frequency spectrum, is thus separated, with good accuracy, into its components. This puts into evidence the convenience of looking for a signal along several different paths in feature space. However, for this analysis to be carried out in a sound way, one should always have a correct probabilistic interpretation of the intensity of the signal as it is projected along each path in feature space. This is exactly what the tomographic analysis provides. I do not know of any other technique fulfilling such requirements.

3.1.2. Component separation in reflectometry data. The reflectometry diagnostic is widely used to determine the electronic density profile in a tokamak. The principle, based upon a radar technique [27], measures the phase of a probing wave reflected by the plasma cut-off layer at a density where the refractive index vanishes. The determination of the density profile can be achieved by continuously sweeping the frequency of the probing wave.

Different techniques are used to measure the density profile on fusion plasmas [28] (phase difference, ultrashort pulses, continuous sweep, etc.) A broadband reflectometer operating in the frequency range 50–75 GHz (V band) [29, 30] and 75–110 GHz (W band) [31] has been developed

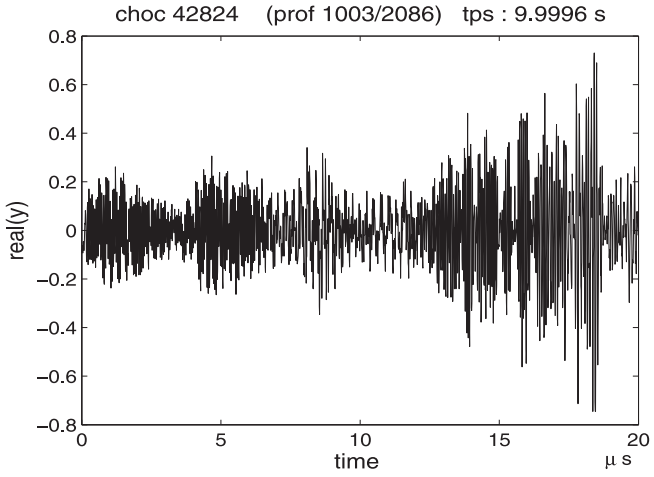

Figure 5. Time representation of the reflectometry signal.

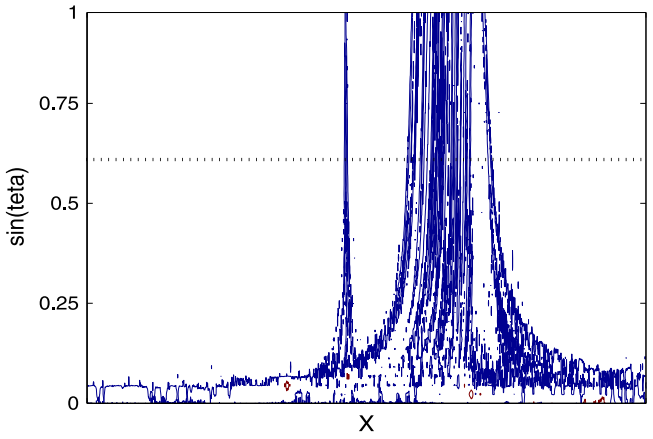

Figure 6. Contour plot of the tomogram of the reflectometry signal ($M_y(X, \theta) \geqslant 0.01$).

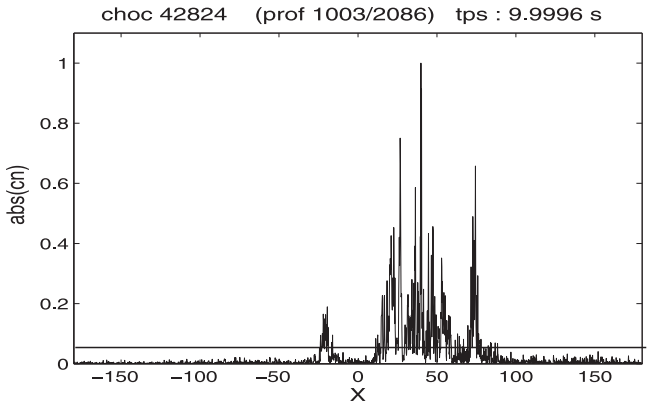

Figure 7. Spectrum $c^{\theta}_{X_n}$ of the reflectometry signal $y(t)$ for $\sin(\theta) = 0.58$.

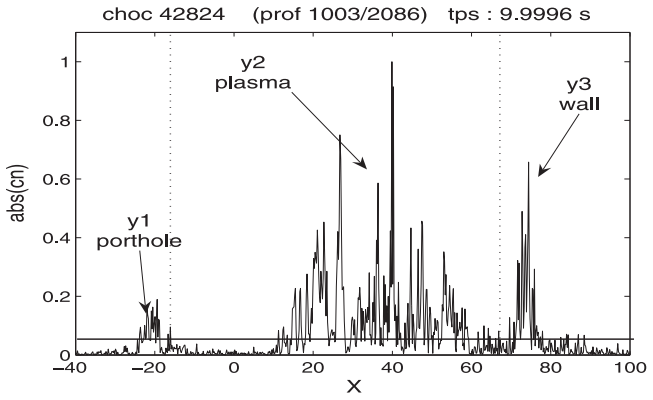

Figure 8. Part of the spectrum $c^{\theta}_{X_n}(y)$ of the reflectometry signal used in the decomposition.

on Tore Supra to measure the electron density profiles at the edge.

The sweep frequency reflectometry system of Tore Supra launches a probing wave on the X-mode polarization in the V band (50–75 GHz). The emitting and receiving antennas are located at about 1.20 m from the plasma edge, outside the vacuum vessel. The reflectometry system operates in burst mode; i.e., the sweeps are performed repeatedly every 25 μs. The duration of one sweep, $E_0(t) = A_0 e^{i\phi(t)}$, is 20 μs, and 5000 chirps are sent during one measurement. During the 20 μs of measurement time, the frequency of the probing wave is continuously varying from 50 GHz to 75 GHz (V band).

The heterodyne reflectometers, with I/Q detection, provide a good signal-to-noise ratio, up to 40 dB. For each sweep, the reflected chirp $E_R(t)$ is mixed with the incident sweep $E_0(t)$, and only the interference term is recorded as in-phase and 90° phase-shifted signals sampled at $T_e = 10^{-8}s$

$$x_1(t) = A_0 A_R(t) \cos(\varphi(t))$$

$$x_2(t) = A_0 A_R(t) \sin(\varphi(t)).$$

For each sweep, the phase $\varphi(t)$ of the reflected signal is represented by

$$y(t) = x_1(t) + i x_2(t) = A(t) e^{i\varphi(t)}. \quad (48)$$

The amplitude of this signal $A(t) = A_0 A_R(t)$ is of low frequency. The real part of one such signal $y(t)$ is shown in figure 5. The labels (choc ...) at the top of figures 5 and 7 and refer to the experiment where this data was acquired.

The contour plot of the tomogram $M_y(X, \theta)$ of the signal is shown in figure 6 where it is possible to see that it carries three main components. The choice of $\sin(\theta) = 0.58$ to perform the decomposition of the signal was done by inspection of this tomogram. The spectrum $c^{\theta}_{X_n}(y)$ of the reflectometry signal for $\sin(\theta) = 0.58$ is shown in figures 7 and 8.

When reconstructing $\tilde{y}(t)$ by:

$$\tilde{y}(t) = \sum_{X_n=-200}^{200} c^{\theta}_{X_n}(y) \psi^{(1)}_{\theta, X_n}(t) \quad (49)$$

the quadratic error $E(y, \tilde{y})$, between the original and the reconstructed signals, is −25 dB.

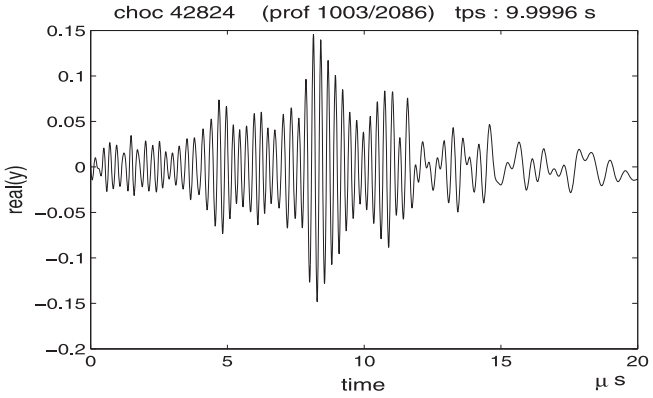

Figure 9. First component of the reflectometry signal corresponding to the reflection on the porthole.

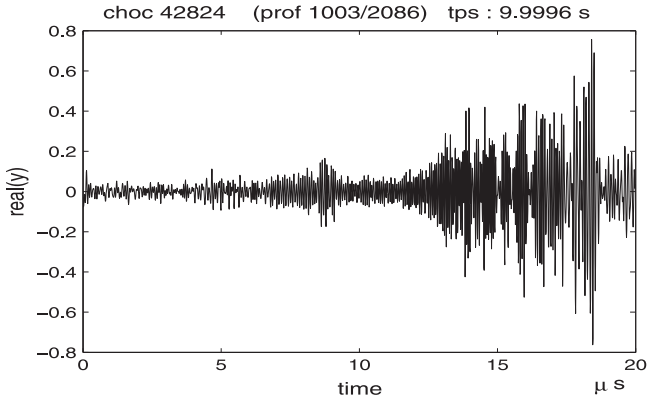

Figure 10. Second component of the reflectometry signal, corresponding to the reflection on the plasma.

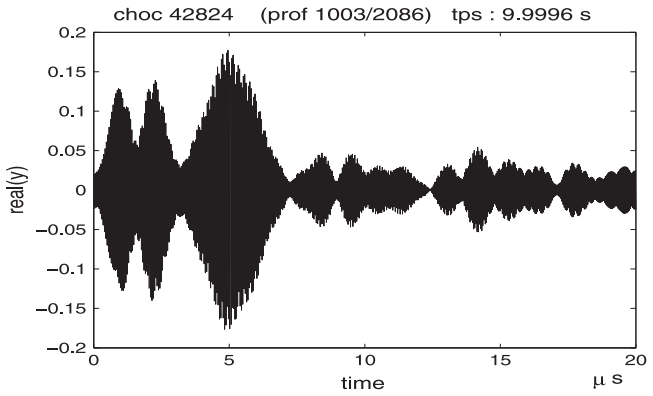

Figure 11. Third component of the reflectometry signal, corresponding to reflection on the wall of the vacuum vessel

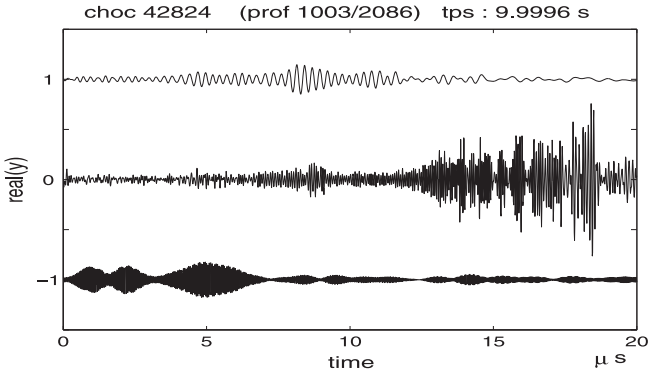

Figure 12. The three components of the reflectometry signal. For visual purposes, the average of $\tilde{y}_1(t)$ is shifted to 1 and the average of $\tilde{y}_3(t)$ to −1.

Decomposition of the reflectometry signal. By taking a threshold equal to $\epsilon = 0.04$, one selects the spectral components corresponding to $\left|c_{X_n}\right| \neq 0$ for $-40 \leqslant X_n \leqslant 100$ (see figure 8). The error between the original and the selected signal is about −18 dB. From this, the spectrum of $y(t)$ splits in three components.

First component the reflection on the porthole. The first component $\tilde{y}_1(t)$ corresponds to $-20 \leqslant x_n \leqslant 0$ and is, therefore, defined as:

$$\tilde{y}_1(t) = \sum_{X_n=-20}^{0} c^{\theta}_{X_n}(y)\psi^{(1)}_{\theta,X_n}(t). \tag{50}$$

It is a low-frequency signal corresponding to the heterodyne product of the probe signal with the reflection on the porthole [31]. It is shown in figure 9.

Second component, the plasma signal. The second component has a Fourier spectrum that fits the expected behavior corresponding to the reflection of the wave inside the plasma of the tokamak [31]. This component, $\tilde{y}_2(t)$, corresponds to $0 \leqslant X_n \leqslant 110$ and is therefore defined as:

$$\tilde{y}_2(t) = \sum_{X_n=0}^{110} c^{\theta}_{X_n}(y)\psi^{(1)}_{\theta,X_n}(t). \tag{51}$$

It is shown in figure 10.

Third component, the first reflection on the wall of the vacuum vessel. The last component corresponds [31] to the first reflection on the wall of the vacuum vessel. This component, $\tilde{y}_3(t)$, corresponds to $110 \leqslant X_n \leqslant 140$ and is therefore defined as:

$$\tilde{y}_3(t) = \sum_{X_n=10}^{140} c^{\theta}_{X_n}(y)\psi^{(1)}_{\theta,X_n}(t). \tag{52}$$

This component is shown in figure 11. Notice that by undertaking a new factorization of this third component it seems possible to separate different successive reflections of the wave.

The three components of the reflectometry signal are presented together on the same plot (figure 12). It is instructive to compare this factorization with the original reflectometry signal (see figure 5).

Before the tomographic technique was developed, the traditional way to separate components in plasma reflectometry was based on the short-time Fourier transform, which provides much poorer results. For detailed comparisons, refer to [22–24].

4. Signal detection with an adapted operator pair

time–frequency tomograms are particularly appropriate to identify the time unfolding of the frequency features of the signals. For example, the component separation success in plasma reflectometry, described before, is to a large extent due to the fact that the plasma is sampled by microwave chirps, and the basis in (28) is exactly a chirp basis. This suggests that, for other types of signals, other types of tomograms should be chosen. It also suggests the best tomographic analysis would be one that mirrors features extracted from the signal itself. This was the motivation to develop a signal-adapted tomogram [26].

In the linear combination $B(\mu, \nu) = \mu t + \nu O$, one chooses an operator O tuned to the signal features that one wants to extract. Then, by looking for particular values in the set (μ, ν) where noise effects might cancel out, we may not only separate the information of weak signals from noise but also obtain reliable information on the temporal structure of the signal. This provides a signal-adapted filtering technique. The construction of an operator suited to particular signals may be done by the same techniques that are used in the bi-orthogonal decomposition [32]. The method for the construction of the adapted operator pair is as follows:

Consider a set of N-dimensional time sequences $\left\{\overrightarrow{x_1}, \cdots, \overrightarrow{x_k}\right\}$, typical of the signal features one wants to detect. From a communication point of view these may be considered as the code words that one wishes to detect in the noisy signal. Form the $k\times N$ matrix $U \in \mathcal{M}_{k\times N}$.

$$U = \begin{pmatrix} x_1(1\Delta t) & x_1(2\Delta t) & \dots & x_1(N\Delta t) \\ \vdots & \vdots & & \vdots \\ x_k(1\Delta t) & x_k(2\Delta t) & \dots & x_k(N\Delta t) \end{pmatrix} \tag{53}$$

with typically $k < N$. Then one constructs the square matrices $A = U^T U \in \mathcal{M}_{N\times N}$ and $B = UU^T \in \mathcal{M}_{k\times k}$. The diagonalization of A provides k non-zero eigenvalues $(\alpha_1, \cdots, \alpha_k)$ and its corresponding orthogonal N-dimensional eigenvectors $(\Phi_1, \cdots, \Phi_k)$, $\Phi_j \in \mathbb{R}^N$. The diagonalization of B would provide the same k eigenvalues and eigenvectors $(\Psi_1, \cdots, \Psi_k)$ with $\Psi_j \in \mathbb{R}^k$. If needed one may obtain, by the Gram–Schmidt method, the remaining $N - k$ eigenvectors to span $\mathbb{R}^N$, which in this context are associated to the eigenvalue zero. The non-zero eigenvalues α_i are the squared non-zero diagonal elements of the singular value decomposition of U.

The linear operator S constructed from the set of typical signals is

$$S = \sum_{i=1}^{k} \alpha_i \Phi_i \Phi_i^t \tag{54}$$

where $S \in \mathcal{M}_{N\times N}$.

For the tomogram, the operator $B(\mu, \nu)$ has the form

$$\begin{aligned} B(\mu, \nu) &= \mu t + \nu S \\ &= \mu \begin{pmatrix} 1\Delta t & & & & \\ & 2\Delta t & & & \\ & & 3\Delta t & & \\ & & & \ddots & \\ & & & & N\Delta t \end{pmatrix} \\ &\quad + \nu \textstyle\sum_{i=1}^{k} \alpha_i \Phi_i \Phi_i^t \end{aligned} \tag{55}$$

with $B \in \mathcal{M}_{N\times N}$.

The eigenvectors of each $B(\mu, \nu)$ are the columns of the matrix that diagonalizes it. Taking the projections of the signal on these eigenvectors, one obtains a tomogram adapted to the operator pair (t, S).

4.1. An example: detection of dust devils

For this example one uses data obtained from the Phoenix Mars Lander [33]. A dust devil is a hot whirlwind generated by a huge temperature contrast between the Martian atmospheric air and the planet surface. Dust devils appear in both temperature and pressure data as sudden drops with a duration between two and three minutes. The upper left panel in figure 13 shows some data from the Phoenix Mars Lander covering a 2000-second interval with a sampling rate of 0.5 Hz. A dust devil is clearly visible at $t \simeq 800\ s$ as a drop in the pressure value.

There have been several efforts to develop systematic methods to detect the effect of dust devils on the Martian atmosphere data. They are based either on checking several ad hoc conditions in the data [33] or on field-programmable gate arrays (FPGAs) [34].

To use the adapted tomographic filtering method for the detection of dust devils, a set of 278 signals was generated. They resemble the shape that a dust devil produces on the data, that is, a sudden drop of about 3% from the baseline, with different durations ranging from 60–80 time units. The upper right panel displays several of these typical signals. Some of the signals have been shifted up or down for representation purposes.

A tomogram is constructed for 20 different values of θ at intervals $\Delta\theta = \pi/40$. A contour plot of the first 999 coefficients of the tomogram is shown in the lower left panel of figure 13.

The coefficient $n = 1000$ corresponds to the largest eigenvalue (and its corresponding eigenvector). This eigenvector contains most of the energy of the signal and is several orders of magnitude larger that any other coefficient, so, for clarity, this coefficient has not been plotted in the tomogram. By direct inspection, one observes that, aside from the coefficient at $n = 1000$, the strongest components concentrate close to $n = 400$. The lower right panel in figure 13 shows the projection on the subspace spanned by the eigenvectors 340–450 and 1000 at $\theta = 19\pi/40$. One sees that the pressure drop produced by the dust devil is very well reconstructed and separated from any other components present in the signal, such as noise or smaller pressure variations.

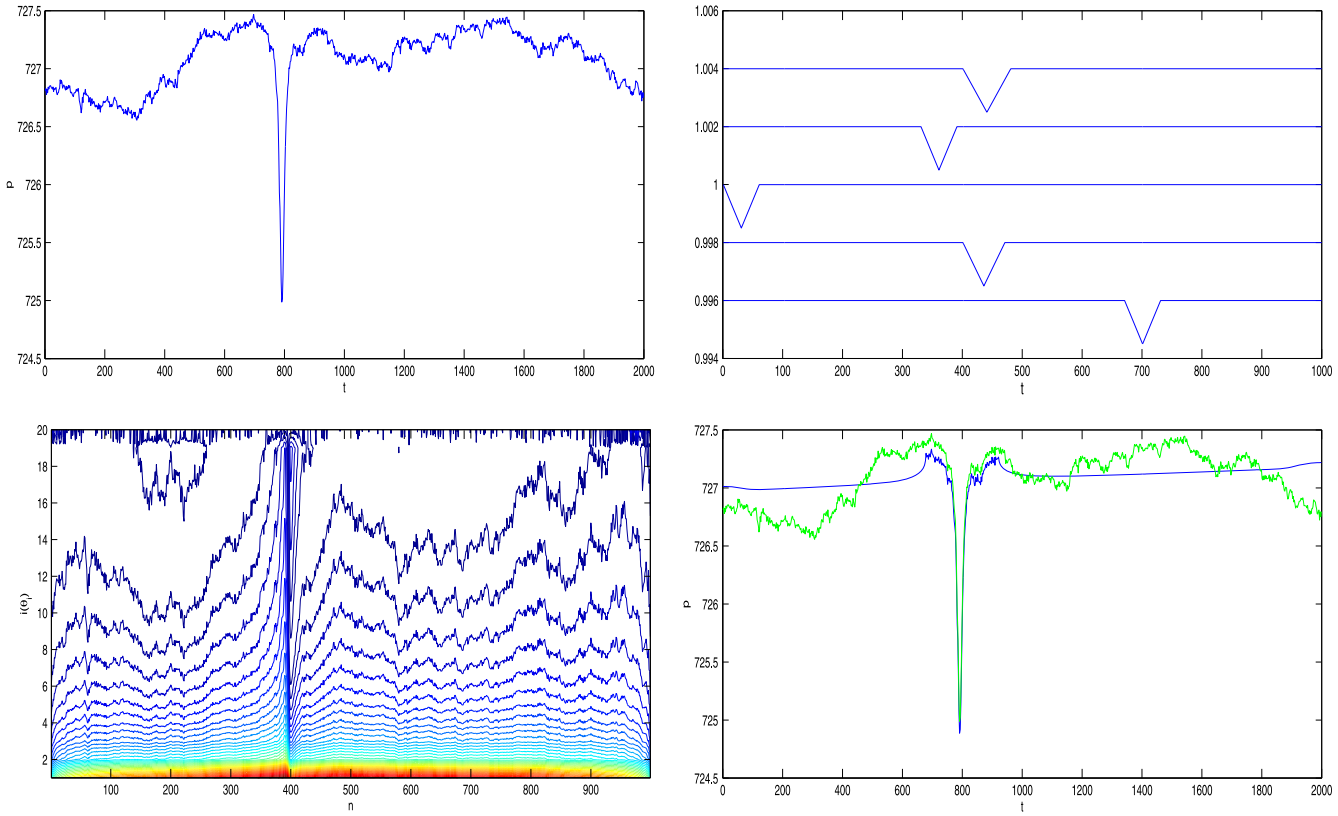

Figure 13. Signal, typical signals, the tomogram (coefficients 1–999) and the projection on the subspace spanned by the eigenvectors 340–450 and 1000 at $\theta = 19\pi/40$.

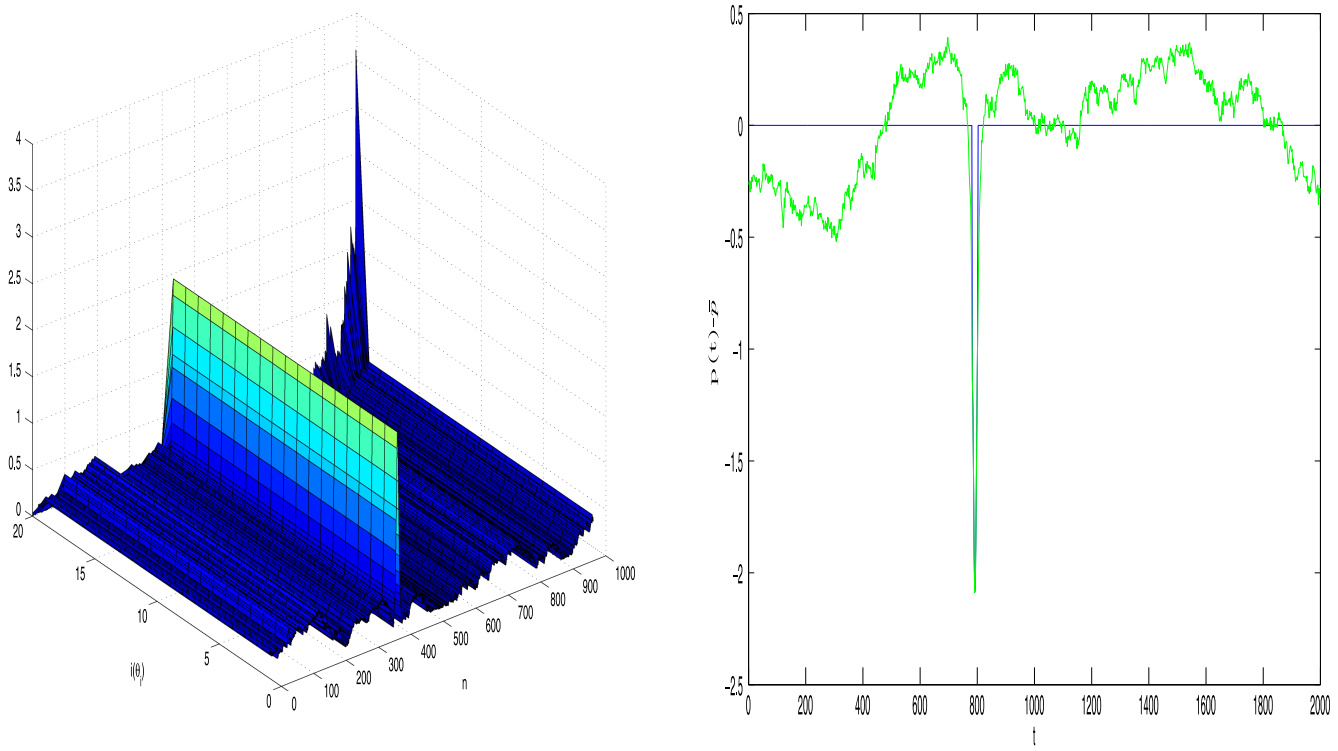

Figure 14. Tomogram for 0 mean typical signals and the projection on the subspace spanned by the eigenvectors 340–450 at $\theta = 19\pi/40$.

As an alternative that avoids the large value of the largest eigenvalue, we may shift both the typical signals and the real data to zero mean signals. In this case there is no eigenvalue much larger than all others. The left panel of figure 14 displays a three-dimensional plot of the tomogram for the 1000 coefficients obtained with zero mean signals. We have also applied a denoising procedure, removing the small coefficients. The right panel in figure 14 shows the projection on the subspace spanned by the eigenvectors 340–450 at $\theta = 19\pi/40$. One sees that the pressure drop produced by the dust devil is completely separated from any other components in the signal.

The signal-adapted tomogram provides not only a multiple-feature extraction capability, as in the tomograms based, for example, on the operators of the conformal group, but it also performs a feature-adapted filtering of the signal.

5. Transforms and tomograms on graphs

Social and economic networks, information networks, power grids, biological networks, etc., generate large sets of raw data from which only a detailed analysis may extract useful information. A first step is the construction of the appropriate signal transforms

From the graph point of view, a time series is a signal on a one-dimensional directed graph with vertices labelled by the times (t_0, t_1, t_2, $\cdots$) and edges connecting t_{k+1} to t_k . That is, the adjacency matrix $\mathbf{A}$ of a time series is, in general

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \end{pmatrix} \tag{56}$$

or, for a time-periodic signal

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{57}$$

Linear signal transforms for a time series are projections on the set of eigenvectors of a linear operator. These operators are not arbitrary but are chosen to extract particular features of the signal that is being analyzed. The Fourier transform looks for periodic features, wavelets for multiscale features, etc. Likewise, useful information from signals on arbitrary graphs may be obtained from projections on sets of vectors associated to suitably chosen linear operators. For the time-periodic signal, it is easy to see that the discrete Fourier transform is the projection on the eigenvectors of the adjacency matrix (57). Therefore, one may generalize the notion of a Fourier transform for graphs as the projection on the eigenvectors (or on the generalized eigenvectors of the Jordan decomposition) of the adjacency matrix. This was the point of view taken by some authors [35–37] to develop a theory of discrete signal processing on graphs. However, this choice is not unique, because, for the time series network, other matrices have the same spectrum, for example, the Laplacian matrix

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

$\mathbf{D}$ being the degree matrix, which for the time series is the identity. Hence, the graph Fourier transform might as well be defined as a projection on the generalized eigenvectors of the Laplacian matrix [38–41]. This operator point of view allows us not only to generalize the notion of transforms but also the notions of filtering and other general linear operations on graph signals.

Let $G = (\mathcal{V}, \mathbf{A})$ be a graph, with $\mathcal{V} = \{v_0, \ldots, v_{N-1}\}$ the set of vertices and $\mathbf{A}$ the weighted adjacency matrix. Each matrix element $\mathbf{A}_{n,m}$ is the weight of a directed edge from v_m to v_n, which can take arbitrary real or complex values. $\mathcal{N}_n = \{m | \mathbf{A}_{n,m} \neq 0\}$ is the neighborhood of v_n, and a graph signal is a map $\mathbf{f} = \{f_n\}$ from the set $\mathcal{V}$ of vertices into the set of complex numbers $\mathbb{C}$, each element f_n being indexed by the vertex v_n.

5.1. Fourier-like transforms

Denote a graph matrix ($\mathbf{A}$, $\mathbf{L}$, or any other) by $\mathbf{M}$. The matrices $\mathbf{M}$ act on the space of graph signals by

$$f_n \to \widetilde{f}_n = \sum_m \mathbf{M}_{n,m} f_m = \sum_{m \in \mathcal{N}_n} \mathbf{M}_{n,m} f_m . \tag{58}$$

When the matrix $\mathbf{M}$ is the adjacency matrix $\mathbf{A}$, this operation generalizes the notion of time shift (when time sequences are looked at as forward-connected graphs).

For many real-world datasets the matrices $\mathbf{M}$ are not diagonalizable. In those cases, to obtain a suitable set of expansion vectors, one may either use the symmetric combinations $\mathbf{M}\mathbf{M}^T$ and $\mathbf{M}^T\mathbf{M}$ to generate an expansion basis or, alternatively, use the block-diagonal Jordan decomposition of $\mathbf{M}$.

$$\mathbf{M} = \mathbf{V}\mathbf{J}\mathbf{V}^{-1} \tag{59}$$

$$\mathbf{J} = \begin{pmatrix} J_{R_{0,0}}(\lambda_0) & & \\ & \ddots & \\ & & J_{R_{M-1,DM-1}}(\lambda_{M-1}) \end{pmatrix} \tag{60}$$

with Jordan blocks associated to the eigenvalues of $\mathbf{M}$

$$J_{r_{m,d}}(\lambda_m) = \begin{pmatrix} \lambda_m & 1 & & \\ & \lambda_m & \ddots & \\ & & \ddots & 1 \\ & & & \lambda_m \end{pmatrix}. \tag{61}$$

The columns of the matrix $\mathbf{V}$, which bring $\mathbf{M}$ to its Jordan normal form, are the eigenvectors

$$(\mathbf{M} - \lambda_m \mathbf{I})\mathbf{v}_{m,d,0} = 0 \tag{62}$$

and the generalized eigenvectors of the Jordan chain

$$(\mathbf{M} - \lambda_m \mathbf{I})\mathbf{v}_{m,d,r} = \mathbf{v}_{m,d,r-1} \tag{63}$$

of $\mathbf{M}$. These vectors may then be used to project the signals

on the graph and, considering the graph signal $\mathbf{f}$ as a column vector, the $\mathbf{M}$-transform is

$$\hat{\mathbf{f}} = V^{-1}\mathbf{f} \tag{64}$$

with inverse transform

$$\mathbf{f} = V\hat{\mathbf{f}}. \tag{65}$$

As stated before, when $\mathbf{M}$ is either the adjacency or the Laplacian matrix, the transforms so obtained correspond to the graph generalization of the Fourier transform, as proposed by several authors. When the matrices are not symmetric, the problem with these transforms lies in the fact that, in general, the set of generalized eigenvectors do not form an orthogonal basis. Therefore it is sometimes more convenient to use $\mathbf{M}\mathbf{M}^T$ and $\mathbf{M}^T\mathbf{M}$ to generate the expansion basis, leading to what one calls the $\mathbf{M}\mathbf{M}^T$—or $\mathbf{M}^T\mathbf{M}$—transform.

5.2. Wavelet-like transforms

The definition of wavelet-like transforms for graphs requires a more elaborate construction. For time series the affine wavelets use, in equation (1), an operator $U(\alpha)$ consisting of the product of a translation and a scale transformation that act on a fixed reference signal (the mother wavelet $h_0(t)$), namely,

$$\begin{aligned} h_{s,a}(t) = U(s, a)h_0(t) &= \mathrm{e}^{\log s\left(t\frac{\mathrm{d}}{\mathrm{d}t}+\frac{1}{2}\right)} \\ &\times \mathrm{e}^{a\frac{\mathrm{d}}{\mathrm{d}t}}h_0(t) = \sqrt{s}\,h_0(st + a). \end{aligned} \tag{66}$$

Translation in the graph is easily generalized, but it is not obvious how to generalize scale transformations. This becomes clearer if one writes the wavelet transform in frequency space,

$$\begin{aligned} f(a, s) &= \int \mathrm{d}t h^*_{s,a}(t)f(t) \\ &= \int \mathrm{d}t\left(\mathrm{e}^{\log s\left(t\frac{\mathrm{d}}{\mathrm{d}t}+\frac{1}{2}\right)}\mathrm{e}^{a\frac{\mathrm{d}}{\mathrm{d}t}}h_0^*(t)\right)f(t) \\ &= \int \mathrm{d}\omega\frac{\mathrm{e}^{-\mathrm{i}\frac{\omega}{s}a}}{\sqrt{s}}\widehat{h_0}^*\left(\frac{\omega}{s}\right)\widehat{f}(\omega) \end{aligned} \tag{67}$$

$\widehat{h_0}$ and $\widehat{f}$ denoting the Fourier transforms of the mother wavelet and of the signal. One sees that the wavelet transform is represented as a sum over the Fourier spectrum Ω with the (frequency) argument of the mother wavelet shifted from ω to $\frac{\omega}{s}$. The mapping $\omega \in \Omega \to \frac{\omega}{s} \in \Omega$ is a one-to-one onto mapping of the Fourier spectrum Ω into itself. Therefore, the natural generalization of the wavelet transform for graphs may be defined as a similar sum, with the spectrum label shift being one of the possible one-to-one onto mappings of the spectrum of the adjacency matrix (or of the Laplacian matrix).

Write the Fourier-like transform on graphs and its inverse as

$$\begin{aligned} \widehat{f}(\eta) &= \sum_i \chi_\eta(i)f(i) \\ f(i) &= \sum_\eta \widehat{f}(\eta)\chi_\eta(i) \end{aligned} \tag{68}$$

where $\chi_\eta(i)$ is an eigenvector of A or L (or a generalized eigenvector or an eigenvector of A^TA or L^TL), and η denotes the spectral label in the spectrum Ω of the matrices. With a localized 'mother wavelet'

$$h^{(k)}(i) = \delta_{k,i} \tag{69}$$

the wavelet-like transform on graphs would be

$$f(a, \tilde{s}) = \sum_\eta \chi_{\tilde{s}(\eta)}(k + a)\widehat{f}(\eta). \tag{70}$$

The mapping $\tilde{s}(\eta)$ is not $\eta \to \frac{\eta}{s}$, because, in general, $\frac{\eta}{s}$ is not in Ω. $\tilde{s}$ is a mapping in the set S of the possible one-to-one onto mappings of Ω, $\tilde{s} \in S$.

The inverse wavelet transform is

$$\widehat{f}(\eta) = \frac{1}{\#S}\sum_{a,\tilde{s}} \chi_{\tilde{s}(\eta)}(a)f(a, \tilde{s}). \tag{71}$$

$\#S$ denotes the cardinality of independent one-to-one onto mappings of Ω.

Hammond, Vandergheynst, and Gribonval [42] have also attempted to generalize the notion of wavelet transform to graph signals. However, instead of the sum with the shifted arguments in the spectrum, their construction corresponds to the introduction of a η-dependent weight on the sum of the 2nd equation in (68), with both the signal component $\widehat{f}(\eta)$ and the eigenvector χ_η associated to the same spectral value η. Therefore, their construction is more in the spirit of a Fourier deformation of the signal rather than of a wavelet transform.

An even more general transform would be

$$f(a, C) = \sum_{\eta,\eta'} C(\eta, \eta')\chi_{\eta'}(a)\widehat{f}(\eta). \tag{72}$$

For comparison with the time series case, this last construction would be similar to the case of 'conformal wavelets' generated by $\mathrm{e}^{\alpha\left(t^2\frac{\mathrm{d}}{\mathrm{d}t}+t\right)}\mathrm{e}^{a\frac{\mathrm{d}}{\mathrm{d}t}}h_0(t)$.

5.3. Graph tomograms

So far signals on graphs have been described either as vectors on vertex space or as projections of these vectors on the generalized eigenvectors of a particular matrix $\mathbf{M}$. Each particular matrix emphasizes a specific topological property of the graph. Tomograms for graphs attempt to obtain information about more than one property by projecting on the generalized eigenvectors of a matrix that interpolates between two distinct matrices $\mathbf{M}_1$ and $\mathbf{M}_2$. This parallels what for a time series is achieved, for example, by the time–frequency tomogram.

In a time series the successive times $t_1, t_2 \cdots t_n$ are the eigenvalues of the time operator t that appears in $B_F(\alpha) = \alpha_1 t + \mathrm{i}\alpha_2 \frac{\mathrm{d}}{\mathrm{d}t}$, the operator that generates the time–frequency tomogram. For a graph, the corresponding notion is the *vertex operator*. For a graph with N vertices the *vertex operator* is

$$\mathbf{T} = \begin{pmatrix} 1 & 0 & 0 & \vdots & 0 \\ 0 & e^{i\frac{2\pi}{N}} & 0 & \vdots & 0 \\ 0 & 0 & e^{i\frac{2\pi}{N}\times 2} & \vdots & 0 \\ \cdots & \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e^{i\frac{2\pi}{N}\times(N-1)} \end{pmatrix} \tag{73}$$

and the vertex signal $\mathbf{f} = \{f_n\}$ corresponds to a projection of $\mathbf{f}$ on the eigenvectors of this operator. Therefore, the construction of a tomogram for graph signals would amount to finding an operator that interpolates between $\mathbf{T}$ and another graph matrix $\mathbf{M}$. A solution would be

$$B_\alpha = (1 - \alpha)\mathbf{T} + \alpha\mathbf{M} \tag{74}$$

with α varying between 0 and 1. The tomogram is obtained by projecting the signal $\mathbf{f}$ on the eigenvectors of B_α. If $\mathbf{M}$ is the adjacency matrix $\mathbf{A}$, this construction, interpolating between $\mathbf{A}$ and the vertex operator $\mathbf{T}$, is, for graphs, the analog of the time–frequency tomogram.

Even if the ordering of the vertices is arbitrary, the vertex operator is always a meaningful entity in the sense that if, for example, the tomogram is used for clustering purposes, it is the $\mathbf{T}$ operator that allows us to identify which vertices belong to each cluster. In addition, more information may be introduced into $\mathbf{T}$ by using, for example, geographical ordering of the nodes or some other property.

Tomograms may also be constructed by using two arbitrary graph matrices $\mathbf{M}_1$ and $\mathbf{M}_2$

$$B_\alpha = (1 - \alpha)\mathbf{M}_1 + \alpha\mathbf{M}_2$$

which may be used to refine the analysis beyond the information obtained from the ($\mathbf{T}$, $\mathbf{A}$)-tomogram

As discussed before, the reason why time and frequency cannot be simultaneously specified is because they correspond to a pair of non-commuting operators. This is the reason why bilinear transforms, like Wigner–Ville, are unreliable, and it is also the main motivation for using tomogram transforms In graphs, the vertex description and the adjacency matrix projection are also incompatible specifications, because in general the $\mathbf{T}$ and $\mathbf{A}$ (or $\mathbf{L}$) matrices do not commute. It is in this sense that, as recently stated [39, 43], there is an uncertainty principle for graphs, that is, a fundamental trade-off between a signal localization on the graph and on its spectral domain.

5.4. Graph tomograms and dynamics

The graph tomogram, as defined above, is appropriate for the study of a static network signal[3]. If during the time evolution the graph structure stays the same, the time series associated to each vertex may simply be projected on the (generalized) eigenvectors, as in the scalar case. However, if the graph itself changes in time, a more general framework must be used.

Consider a graph signal that evolves in (discrete) time. The corresponding graph would be, for each time t, a regular graph, and each one of these graphs is forward-connected to the graph of the subsequent time. A vertex $\nu_n(t)$ at time t connects to the vertex $\nu_n(t+1)$ at time $t+1$. This construction accommodates the possible disappearance of vertices. In that case such vertex $\nu_n(t)$ would not have any forward edges.

The construction of the $\mathbf{M}$-transforms and the graph tomograms will then proceed as before for the global adjacency matrix. To have a feeling for this kind of construction, consider the simple case of a finite-vertex circle graph with N vertices symmetrically connected to nearest-neighbors and forward connected in periodic time with τ time steps. Then, at each time t, the adjacency matrix $\mathbf{A}(t)$ is

$$\mathbf{A}(t) = \begin{pmatrix} 0 & 1 & 0 & 0 & \vdots & 1 \\ 1 & 0 & 1 & 0 & \vdots & 0 \\ 0 & 1 & 0 & 1 & \vdots & 0 \\ 0 & 0 & 1 & 0 & \vdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{pmatrix}. \tag{75}$$

Let, for definiteness and notational simplicity, $N = \tau = 3$. Then the global 9×9 adjacency matrix is

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{pmatrix}. \tag{76}$$

This matrix is

$$\mathbf{A} = \left(\mathbf{I}_3 \otimes \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \right) \oplus \left(\begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \otimes \mathbf{I}_3 \right)$$

with eigenvalues

$$0, \frac{3 \pm \sqrt{5}}{2}, \frac{-3 \pm i\sqrt{3}}{2}, \frac{-5 \pm i\sqrt{3}}{2}, \frac{5 \pm i\sqrt{3}}{2}.$$

The 'Fourier' transform of the dynamical graph signal will be the projection on the corresponding nine-dimensional eigenvectors.

[3] Likewise, the usual time–frequency tomogram may be looked upon as a static description of the whole time history of the system.

For the construction of the tomogram, the vertex operator **T**, as in (73), is

$$\mathbf{T} = \begin{pmatrix} \mathbf{T}^{(3)} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{(3)} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{T}^{(3)} \end{pmatrix}$$

where $\mathbf{T}^{(3)}$ is the 3×3 matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\frac{2\pi}{3}} & 0 \\ 0 & 0 & e^{i\frac{2\pi}{3}\times 2} \end{pmatrix}$.

This general framework, where one takes into account both the network edges and the time links, allows for a unified treatment of both the dynamics over graphs and the dynamics associated to a time-changing topology.

For more details on the tomographic approach to graph signals and some applications, refer to [44].

6. Conclusions

1–Tomograms, a generalization of the Radon transform, first developed for applications in quantum mechanics, are also a powerful tool for the processing of classical signals.

2–As in the case of quantum mechanics, which deals with pairs of non-commuting opertors time and frequency or time and scale, etc, are incompatible features which cannot be simultaneously specified with absolute precision. This is why tomograms, by providing a robust probabilistic interpretation along paths in multi-feature space, are a useful tool in signal processing.

3–The tomographic formulation in signal processing turns out to be a sucessful technique for denoising, component separation, and even, in its signal-adapted form, to extract customer-oriented features of arbitrary signals.

4–The large amount of data that is currently generated in technological and social networks may also benefit from graph signal processing and tomograms on graphs.

Q3 ## References

[1] Wigner E 1932 On the quantum correction for thermodynamic equilibrium *Phys. Rev.* **40** 749–59

[2] Radon J 1917 Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten *Ber. Verh. Sachs. Akad.* **69** 262–77

[3] Deans S R 1983 *The Radon Transform and Some of Its Applications* (New York: Wiley)

[4] Mancini S, Man'ko V I and Tombesi P 1996 Symplectic tomography as classical approach to quantum systems *Phys. Lett.* A **213** 1–6

[5] Man'ko M A and Man'ko V I 2011 Probability Description and Entropy of Classical and Quantum Systems *Found. Phys.* **41** 330–44

[6] Ibort A, Man'ko V I, Marmo G, Simoni A and Ventriglia F 2009 An introduction to the tomographic picture of quantum mechanics *Phys. Scr.* **79** 065013

[7] Man'ko M A, Man'ko V I, Marmo G, Simoni A and Ventriglia F 2013 Introduction to tomography, classical and quantum *Nuovo Cim.* C **36** 163–82

[8] Ville J 1948 Théorie et applications de la notion de signal analytique *Cables Transm.* **2A** 61–74

[9] Cohen L 1966 Generalized phase-space distribution functions *J. Math. Phys.* **7** 781–806

[10] Cohen L 1989 Time–frequency distributions—A review *Proc. IEEE* **77** 941–81

[11] Husimi K 1940 Some formal properties of the density matrix *Proc. Phys. Mat. Soc. Jpn.* **22** 264–314

[12] Kano Y 1965 A new phase-space distribution function in the statistical theory of the electromagnetic field *J. Math. Phys.* **6** 1913–5

[13] Man'ko V I and Vilela Mendes R 1999 Noncommutative time–frequency tomography *Phys. Lett.* A **263** 53–59

[14] Man'ko M A 2000 Quasidistributions, tomography, and fractional Fourier transform in signal analysis *J. Russian Laser Research* **21** 411–37

[15] Man'ko M A, Man'ko V I and Mendes R Vilela 2001 Tomograms and other transforms: a unified view *J. Phys. A: Math. Gen.* **34** 8321–32

[16] Man'ko M A 2002 Electromagnetic Signal Processing and Noncommutative Tomography *J. Russian Laser Research* **23** 433–48

[17] Man'ko M A 2006 Entropy of an analytic signal *J. Russian Laser Research* **27** 405–13

[18] Briolle F, Man'ko V I, Ricaud B and Vilela Mendes R 2012 Non-commutative tomography: A tool for data analysis and signal processing *J. Russian Laser Research* **33** 103–21

[19] Bertrand J and Bertrand P 1992 A class of affine Wigner functions with extended covariance properties *J. Math. Phys.* **33** 2515–27

[20] Goncalvés P and Baraniuk R G 1996 A pseudo-Bertrand distribution for time-scale analysis *IEEE Signal Process. Lett.* **3** 82–84

[21] Aguirre C, Pascual P, Campos D and Serrano E 2011 Single neuron transient activity detection by means of tomography *BMC Neuroscience* **12** 297

[22] Briolle F, Lima R, Man'ko V I and Mendes R Vilela 2009 A tomographic analysis of reflectometry data I: Component factorization *Meas. Sci. Technol.* **20** 105501

[23] Briolle F, Lima R and Vilela Mendes R 2009 A tomographic analysis of reflectometry data: II. The phase derivative *Meas. Sci. Technol.* **20** 105502

[24] Clairet F, Briolle F, Ricaud B and Heuraux S 2011 New signal processing technique for density profile reflectometry on Tore Supra *Rev. Sci. Instrum.* **82** 083502

[25] Ricaud B, Briolle F and Clairet F *Analysis and Separation of Time–Frequency Components in Signals With Chaotic Behavior* (arXiv:1003.0734.) Q1

[26] Aguirre C and Vilela Mendes R 2014 Signal recognition and adapted filtering by non-commutative tomography *IET Signal Process.* **8** 67–75

[27] Hugenholtz C A and Heijnen S H 1991 Pulse radar technique for reflectometry on thermonuclear plasmas *Rev. Sci. Instrum.* **62** 1100–1

[28] Laviron C, Donné A J H, Manso M E and Sanchez J 1996 Reflectometry techniques for density profiles measurements on fusion plasmas *Plasma Phys. Control. Fusion* **38** 905–36

[29] Clairet F, Bottereau C, Chareau J M, Paume M and Sabot R 2001 Edge density profile measurements by X-mode reflectometry on Tore Supra *Plasma Phys. Control. Fusion* **43** 429–42

[30] Clairet F, Sabot R, Bottereau C, Chareau J M, Paume M, Heureaux S, Colin M, Hacquin S and Leclert G 2001 X-mode heterodyne reflectometer for edge density profile measurements on Tore Supra *Rev. Sci. Instrum.* **72** 340–3

[31] Clairet F, Bottereau C, Chareau J M and Sabot R 2003 Advances of the density profile reflectometry on TORE SUPRA *Rev. Sci. Instrum.* **74** 1481–4

[32] Dente J A, Vilela Mendes R, Lambert A and Lima R 1996 The bi-orthogonal decomposition in image processing: Signal analysis and texture segmentation *Signal Process. Image Commun.* **8** 131–48
[33] Smith P H *et al* 2008 Introduction to special section on the phoenix mission: landing site characterization experiments, mission overviews, and expected science *J. Geophys. Res.* **113** E00A18
[34] de Lucas E, Miguel M J, Mozos D and Vázquez L 2012 Martian dust devils detector over FPGA *Geosci. Instrum. Method. Data Syst.* **1** 23–31
[35] Sandryhaila A and Moura J F 2013 Discrete signal processing on graphs *IEEE Trans. on Signal Processing* **61** 1644–56
[36] Sandryhaila A and Moura J F 2014 Discrete signal processing on graphs: frequency analysis *IEEE Trans. on Signal Processing* **62** 3042–54
[37] Sandryhaila A and Moura J F 2014 Big data analysis with signal processing on graphs *IEEE Signal Process. Mag.* **31** 80–90
[38] Ekambaram V N, Fanti G C, Ayazifar B and Ramchandran K 2013 Circulant structures and graph signal processing *20th IEEE Int. Conf, on Image Processing (ICIP)* pp 834–8
[39] Agaskar A and Lu Y M 2012 Uncertainty principles for signals defined on graphs: bounds and characterizations *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)* pp 3493–6
[40] Miller B A, Bliss N T and Wolfe P J 2010 Towards signal processing theory for graphs and non-Euclidean data *Proc. ICASSP* pp 5414–7
[41] Shuman D I *et al* 2013 The emerging field of signal processing on graphs. Extending high-dimensional data analysis to networks and other irregular domains *IEEE Signal Processing Mag.* **30** 83–98
[42] Hammond D K, Vandergheynst P and Gribonval R 2011 Wavelets on graphs via spectral graph theory *Appl. Comput. Harmon. Anal.* **30** 129–50
[43] Agaskar A and Lu Y 2013 A Spectral graph uncertainty principle *IEEE Trans. on Information Theory* **59** 4338–56
[44] Vilela Mendes R, Mendes H C and Araújo T Signal processing on graphs: transforms and tomograms (arXiv:1406.2185.) Q2

QUERY FORM

JOURNAL: Physica Scripta

AUTHOR: R V Mendes

TITLE: Non-commutative tomography and signal processing

ARTICLE ID: ps509780

The layout of this article has not yet been finalized. Therefore this proof may contain columns that are not fully balanced/ matched or overlapping text in inline equations; these issues will be resolved once the final corrections have been incorporated.

Page 14

Q1
Please provide updated details for reference [25] if available.

Page 15

Q2
Please provide updated details for reference [44] if available.

Page 14

Q3
Please check the details for any journal references that do not have a link as they may contain some incorrect information.