Preservation of whiteness in spectral and time-frequency transforms of second order processes

Préservation de la blancheur dans les transformations spectrales et temps-fréquences de processus du second ordre

Roland Badeau

2016D001

Janvier 2016

Département Traitement du Signal et des Images
Groupe AAO : Audio, Acoustique et Ondes

Preservation of whiteness in spectral and time-frequency transforms of second order processes

Préservation de la blancheur dans les transformations spectrales et temps-fréquences de processus du second ordre

Roland Badeau
LTCI, CNRS, Télécom ParisTech, Université Paris-Saclay, 75013, Paris, France
Email : roland.badeau@telecom-paristech.fr

Abstract

In many signal processing applications, recent techniques often rely on the estimation of a probabilistic model. Many times, this model does not focus on the observed data itself, but rather on a spectral or time-frequency transform of this data, such as the discrete Fourier transform (DFT) or the short-time Fourier transform (STFT). A common statistical assumption regarding these transforms is that all spectral or time-frequency bins are uncorrelated. However this assumption is generally inaccurate, either because of the intrinsic properties of the data, or because of the transform itself. In this document, we aim to design transforms from the time domain to the spectral or time-frequency domain, which best fit this statistical assumption. To formulate this idea, we introduce the concept of *preservation of whiteness*, and we characterise the transforms that satisfy this property. We show that several widely used transforms such as the discrete cosine transform (DCT), DFT, modified discrete cosine transform (MDCT), and STFT belong to this class under some conditions.

Index Terms

Second order processes, proper complex processes, spectral transforms, time-frequency transforms, paraunitary filter banks.

Résumé

Dans diverses applications de traitement du signal, les techniques récentes s'appuient souvent sur l'estimation d'un modèle probabiliste. Dans de nombreux cas, le modèle ne représente pas directement les données observées elles-mêmes, mais plutôt une transformée spectrale ou temps-fréquence de ces données, telle que la transformée de Fourier discrète (TFD) ou la transformée de Fourier à court terme (TFCT). Une hypothèse statistique couramment utilisée à propos de ces transformées est que tous les points fréquentiels ou temps-fréquence sont décorrélés. Cependant cette hypothèse est généralement inexacte, soit en raison des propriétés intrinsèques des données, soit en raison de la transformée elle-même. Dans ce document, nous cherchons à définir des transformations depuis le domaine temporel vers le domaine spectral ou temps-fréquence, qui vérifient au mieux cette hypothèse statistique. Pour formaliser cette idée, nous introduisons le concept de *préservation de la blancheur*, et nous caractérisons les transformations qui satisfont cette propriété. Nous montrons que plusieurs transformations couramment utilisées telles que la transformée en cosinus discrète (TCD), la TFD, la transformée en cosinus discrète modifiée (TCDM) et la TFCT appartiennent à cette classe sous certaines conditions.

Mots clés

Processus du second ordre, Processus complexes propres, Transformations spectrales, Transformations temps-fréquences, Bancs de filtres para-unitaires.

I. INTRODUCTION

Spectral and time-frequency (TF) transforms are widely used in many signal processing applications, because they provide a meaningful and often sparse representation of the input signal [1]. Besides, many modern signal processing and machine learning techniques are based on probabilistic modelling and statistical inference [2]. For these reasons, there is much interest today in the probabilistic modelling of TF data. Such models have been proposed in the nonnegative matrix factorisation (NMF) literature for instance, among which NMF with additive Gaussian noise [3], probabilistic latent component analysis (PLCA) [4], NMF as a sum of Poisson components [5], and NMF as a sum of Gaussian components [6]. Usually, it is assumed that all spectral or TF bins are uncorrelated. However, this assumption only holds approximately, for a limited range of random processes (excluding sinusoidal processes for instance). Moreover, spurious correlation artefacts may also be induced by the transform itself, due to spectral and temporal overlap of TF bins. For these reasons, new probabilistic approaches have been proposed in order to model the existing correlation [7], [8], [9], [10], [11], [12], [13]. However these models do not make the difference between the correlation originating from the data and that induced by the transform.

In this work, we investigate whether it is possible to design spectral and TF transforms that do not induce any spurious correlation artefacts, in addition to the correlation inherent to the input process. The proposed approach relies on the idea that if the input process contains no intrinsic TF correlation, then the output TF bins should be uncorrelated. Actually, it turns out that the absence of TF correlation implies that the input process is white. Indeed, in the case of wide sense stationary (WSS) processes, it is well known that the power spectral density (PSD) dynamics is closely related to the temporal autocorrelation [2]. Reciprocally, in the case of uncorrelated processes (whose samples are uncorrelated but may have different variances), temporal power dynamics induces spectral correlation. Therefore a process which has neither temporal nor spectral correlation also has neither temporal nor spectral dynamics, which means that this process is white. In other words, the desired property of the transform is that any white noise is transformed into another white noise. This is what we call the *preservation of whiteness* (PW) property. Besides, in some applications it is required that the transform be invertible, *e.g.* when a signal has to be resynthesized from modified TF data (as in source separation [11], [12], time and pitch scaling [14], and audio inpainting [11], [12], [14]). We will thus focus on transforms which satisfy both conditions of invertibility and PW.

In this document, we investigate this concept using simple mathematics based on linear algebra, and we provide some examples of such transforms. Several cases are distinguished: the input process may be either real or complex-valued, and the spectral or TF output may also be real (such as the DCT of a real input) or complex (such as the DFT of a real or complex input). The document is structured as follows. In Section II, we define real to real, complex to complex, real to complex, and complex to real transforms, we characterise the invertibility of such transforms, and we formally define the PW property. In Section III, we characterise invertible real to real and complex to complex transforms which satisfy the PW property, and we provide a few examples of such spectral and TF transforms. The case of real to complex transforms is similarly addressed in Section IV. Section V provides some additional insights about the whole study. Finally, conclusions are drawn in Section VI.

Notation

The following mathematical notation is used throughout this document:

- $\mathbb{R}$ (resp. $\mathbb{C}, \mathbb{Z}$): set of real (resp. complex, whole) numbers
- $\Re$: real part of a complex number
- $\overline{\boldsymbol{F}}$: complex conjugate of matrix $\boldsymbol{F}$
- $\boldsymbol{F}^T$: transpose of matrix $\boldsymbol{F}$
- $\boldsymbol{F}^H$: conjugate transpose of matrix $\boldsymbol{F}$
- $\boldsymbol{I}$ (resp. $\boldsymbol{0}$): identity (resp. zero) matrix of appropriate dimension
- $\mathbb{E}[.]$: mathematical expectation

II. BASIC DEFINITIONS

Definition 1. *Let $P, N \in \{1, \ldots, +\infty\}$ and $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. Then $\mathrm{M}_{\mathbb{K}}^{P\times N}$ denotes the set of $P \times N$ matrices over field $\mathbb{K}$, such that every row contains a finite number of non-zeros entries if $N = +\infty$.*

Let $\mathbb{K}^N$ denote the set of column vectors of length $N \in \{1, \dots, +\infty\}$ over field $\mathbb{K}$. Then $\mathrm{M}_{\mathbb{K}}^{P\times N}$ is the largest set of matrices such that $\forall \boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{P\times N}$, $\forall \boldsymbol{w} \in \mathbb{K}^N$, the matrix-vector product $\boldsymbol{F}\boldsymbol{w}$ is well-defined (since only finite sums are involved, even when $N = +\infty$).

A. Real to real and complex to complex transforms

Below, we consider linear transforms between $\mathbb{K}^N$ and $\mathbb{K}^P$ (with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), and we characterise their invertibility.

Definition 2 (Linear transform)**.** *Let $P, N \in \{1, \dots, +\infty\}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{P\times N}$. The linear transform defined by matrix $\boldsymbol{F}$ is denoted $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ and transforms any vector $\boldsymbol{w} \in \mathbb{K}^N$ into $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}(\boldsymbol{w}) = \boldsymbol{F}\boldsymbol{w} \in \mathbb{K}^P$.*

The next proposition is a well-known result from linear algebra:

Proposition 1 (Invertible linear transforms)**.** *Let $P, N \in \{1, \dots, +\infty\}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{P\times N}$. Then $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ is invertible if and only if $N \leq P$ and if there is $\boldsymbol{G} \in \mathrm{M}_{\mathbb{K}}^{N\times P}$ such that $\boldsymbol{G}\boldsymbol{F} = \boldsymbol{I}$. If this condition holds, then $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{K}^P\to\mathbb{K}^N}$ is an inverse of $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$.*

B. Real to complex and complex to real transforms

Below, we define transforms between $\mathbb{R}^N$ and $\mathbb{C}^P$, and we characterise their invertibility. For any column vector $\boldsymbol{x} \in \mathbb{C}^P$ where $P \in \{1, \dots, +\infty\}$, we note $\downarrow\boldsymbol{x} = \begin{bmatrix} \boldsymbol{x} \\ \overline{\boldsymbol{x}} \end{bmatrix}$.

Definition 3 (Real to complex transform)**.** *Let $P, N \in \{1, \dots, +\infty\}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{P\times N}$. The real to complex transform defined by matrix $\boldsymbol{F}$ is denoted $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ and transforms any vector $\boldsymbol{w} \in \mathbb{R}^N$ into $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}(\boldsymbol{w}) = \boldsymbol{F}\boldsymbol{w} \in \mathbb{C}^P$. If we note $\downarrow\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F} \\ \overline{\boldsymbol{F}} \end{bmatrix}$ and $\boldsymbol{x} = \mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}(\boldsymbol{w})$, then we can equivalently write $\downarrow\boldsymbol{x} =\downarrow\boldsymbol{F}\boldsymbol{w}$.*

Definition 4 (Complex to real transform)**.** *Let $P, N \in [1, \dots, +\infty]$, and $\boldsymbol{G} \in \mathrm{M}_{\mathbb{C}}^{N\times P}$. The complex to real transform defined by matrix $\boldsymbol{G}$ is denoted $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{C}^P\to\mathbb{R}^N}$ and transforms any vector $\boldsymbol{x} \in \mathbb{C}^P$ into $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{C}^P\to\mathbb{R}^N}(\boldsymbol{x}) = 2\Re(\boldsymbol{G}\boldsymbol{x}) \in \mathbb{R}^N$. If we note $\overrightarrow{\boldsymbol{G}} = [\boldsymbol{G}, \overline{\boldsymbol{G}}]$, then we can equivalently write $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{C}^P\to\mathbb{R}^N}(\boldsymbol{x}) = \overrightarrow{\boldsymbol{G}}\downarrow\boldsymbol{x}$.*

The next proposition is a variant of Proposition 1.

Proposition 2 (Invertible real to complex transforms)**.** *Let $P, N \in \{1, \dots, +\infty\}$ and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{P\times N}$. Then $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ is invertible if and only if $N \leq 2P$ and if there is $\boldsymbol{G} \in \mathrm{M}_{\mathbb{C}}^{N\times P}$ such that $\overrightarrow{\boldsymbol{G}}\downarrow\boldsymbol{F} = \boldsymbol{I}$. If this condition holds, then $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{C}^P\to\mathbb{R}^N}$ is an inverse of $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$.*

Proof. Firstly, it is easy to check that the inverse of $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$, if it exists, is necessarily a complex to real transform as defined in Definition 4. Therefore $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ is invertible if and only if there is $\boldsymbol{G} \in \mathrm{M}_{\mathbb{C}}^{N\times P}$ such that $\forall \boldsymbol{w} \in \mathbb{R}^N$, $\mathcal{T}^{\boldsymbol{G}}_{\mathbb{C}^P\to\mathbb{R}^N}(\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}(\boldsymbol{w})) = \boldsymbol{w} \Leftrightarrow \overrightarrow{\boldsymbol{G}}\downarrow\boldsymbol{F}\boldsymbol{w} = \boldsymbol{w}$, or equivalently $\overrightarrow{\boldsymbol{G}}\downarrow\boldsymbol{F} = \boldsymbol{I}$. The inequality $N \leq 2P$ is a consequence of this matrix equality. □

C. Preservation of whiteness

Let $N \in \{1, \dots, +\infty\}$. A second order process $X = \{X_n\}_{0\leq n<N}$ over field $\mathbb{K}$ is a random process such that $X_n \in \mathbb{K}$ and $\mathbb{E}[|X_n|^2] < +\infty$. Note that all transforms defined in Sections II-A and II-B transform any second order process into another second order process.

Definition 5 (Real white noise)**.** *Let $N \in \{1, \dots, +\infty\}$. A real white noise $W = \{W_n\}_{0\leq n<N}$ of variance $\sigma^2 > 0$ is a real-valued second order process of mean vector $\boldsymbol{\mu}_W = \mathbb{E}[W] = \boldsymbol{0}$ and covariance matrix $\boldsymbol{\Gamma}_W = \mathbb{E}[WW^T] = \sigma^2\boldsymbol{I}$.*

Definition 6 (Proper complex white noise)**.** *Let $N \in \{1, \dots, +\infty\}$. A proper complex white noise $W = \{W_n\}_{0\leq n<N}$ of variance $\sigma^2 > 0$ is a complex-valued second order process of mean vector $\boldsymbol{\mu}_W = \mathbb{E}[W] = \boldsymbol{0}$, covariance matrix $\boldsymbol{\Gamma}_W = \mathbb{E}[WW^H] = \sigma^2\boldsymbol{I}$, and pseudo-covariance matrix $\boldsymbol{\Phi}_W = \mathbb{E}[WW^T] = \boldsymbol{0}$.*

Below, a real white noise will be referred to as a *white noise on* $\mathbb{R}$, and a proper complex white noise as a *white noise on* $\mathbb{C}$.

Definition 7 (Preservation of whiteness)**.** *Let $P, N \in \{1, \ldots, +\infty\}$, $\mathbb{K}_1 = \mathbb{R}$ or $\mathbb{C}$, and $\mathbb{K}_2 = \mathbb{R}$ or $\mathbb{C}$. A transform from $\mathbb{K}_1^N$ to $\mathbb{K}_2^P$ preserves whiteness if and only if the output of any white noise on $\mathbb{K}_1^N$ is a white noise on $\mathbb{K}_2^P$.*

III. Preservation of whiteness in real to real and complex to complex transforms

In this section, we characterise invertible whiteness-preserving transforms between $\mathbb{K}^N$ and $\mathbb{K}^P$ (with $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$), and we provide a few examples of such spectral and TF transforms.

A. General study

Proposition 3 (PW in linear transforms)**.** *Let $P, N \in \{1, \ldots, +\infty\}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{P\times N}$. The linear transform $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ preserves whiteness if and only if $P \leq N$ and $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}$.*

Proof. Let $W = \{W_n\}_{0\leq n<N}$ be a white noise of variance σ^2 over $\mathbb{K}$, and let $X = \mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}(W)$. Then $\boldsymbol{\mu}_X = \boldsymbol{F}\boldsymbol{\mu}_W = \boldsymbol{0}$, and if $\mathbb{K} = \mathbb{C}$, $\boldsymbol{\Phi}_X = \mathbb{E}[XX^T] = \boldsymbol{F}\boldsymbol{\Phi}_W\boldsymbol{F}^T = \boldsymbol{0}$. Therefore X is a white noise of variance σ^2 over $\mathbb{K}$ if and only if $\boldsymbol{\Gamma}_X = \mathbb{E}[XX^H] = \boldsymbol{F}\boldsymbol{\Gamma}_W\boldsymbol{F}^H = \sigma^2\boldsymbol{I}$, which is equivalent to $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}$. □

Proposition 4 (PW in invertible linear transforms)**.** *Let $P, N \in \{1, \ldots, +\infty\}$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{P\times N}$. The linear transform $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ is invertible and preserves whiteness if and only if $N = P$, $\boldsymbol{F}^H \in \mathrm{M}_{\mathbb{K}}^{N\times N}$, and $\boldsymbol{F}$ is a unitary matrix: $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{I}$.*

Proof. According to Proposition 1, $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ is invertible if and only if $N \leq P$ and if there is $\boldsymbol{G} \in \mathrm{M}_{\mathbb{K}}^{N\times P}$ such that $\boldsymbol{G}\boldsymbol{F} = \boldsymbol{I}$. Moreover, according to Proposition 3, $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{K}^N\to\mathbb{K}^P}$ preserves whiteness if and only if $P \leq N$ and $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}$. Both assertions hold if and only if $N = P$ and $\boldsymbol{G}\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{G} = \boldsymbol{F}^H$, which is equivalent to $\boldsymbol{F}^H \in \mathrm{M}_{\mathbb{K}}^{N\times N}$ and $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{F}^H\boldsymbol{F} = \boldsymbol{I}$. □

B. Examples

1) Discrete cosine transform: Let $N = P < +\infty$ and $\mathbb{K} = \mathbb{R}$. The discrete cosine transform (DCT) X of a signal W of length N is a spectral transform defined as $\forall k \in [0 \ldots N-1]$,

$$X_k = \sqrt{\frac{2}{N}} \sum_{n=0}^{N-1} W_n \cos\left(\frac{\pi}{N}(k+\phi)(n+\tau)\right),$$

where $\phi, \tau \in \mathbb{R}$. The entries of the corresponding matrix $\boldsymbol{F} \in \mathrm{M}_{\mathbb{R}}^{N\times N}$ are $F_{kn} = \sqrt{\frac{2}{N}} \cos\left(\frac{\pi}{N}(k+\phi)(n+\tau)\right)$. It can be verified that if $\phi = \tau = \frac{1}{2}$ (which corresponds to the DCT-IV [15]), then $\boldsymbol{F}$ is a unitary matrix. Proposition 4 then proves that $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{R}^N}$ preserves whiteness and that its inverse is defined as $\forall n \in [0 \ldots N-1]$,

$$W_n = \sqrt{\frac{2}{N}} \sum_{k=0}^{N-1} X_k \cos\left(\frac{\pi}{N}\left(k+\frac{1}{2}\right)\left(n+\frac{1}{2}\right)\right).$$

2) Generalised discrete Fourier transform: Let $N = P < +\infty$ and $\mathbb{K} = \mathbb{C}$. The generalised discrete Fourier transform (GDFT) [16] X of a signal W of length N is a spectral transform defined as $\forall k \in [0 \ldots N-1]$,

$$X_k = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} W_n e^{-\frac{i2\pi}{N}(k+\phi)(n+\tau)}. \tag{1}$$

where $\phi, \tau \in \mathbb{R}$. The entries of the corresponding matrix $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{N\times N}$ are $F_{kn} = \frac{1}{\sqrt{N}} e^{-\frac{i2\pi}{N}(k+\phi)(n+\tau)}$. The particular case $\phi = \tau = 0$ corresponds to the regular DFT. It can be verified that $\forall \phi, \tau \in \mathbb{R}$, $\boldsymbol{F}$ is a unitary matrix. Proposition 4 then proves that $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{C}^N\to\mathbb{C}^N}$ preserves whiteness and its inverse is defined as $\forall n \in [0 \ldots N-1]$,

$$W_n = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k e^{+\frac{i2\pi}{N}(k+\phi)(n+\tau)}.$$

3) Paraunitary filter banks: Let $N = P = +\infty$, $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$, and consider a TF transform defined by an infinite matrix $\boldsymbol{F}$ over $\mathbb{K}$, which implements a uniform M-channel filter bank (*uniform* means that the same decimation factor is applied in every subband). According to Definition 1, $\boldsymbol{F} \in \mathrm{M}_{\mathbb{K}}^{N\times N}$ if and only if the filter bank involves only finite impulse response (FIR) analysis filters. Moreover, according to Proposition 1, $\mathcal{T}_{\mathbb{K}^N\to\mathbb{K}^N}^{\boldsymbol{F}}$ is invertible if and only if perfect reconstruction (PR) [1] is provided by means of FIR synthesis filters. In addition, Proposition 4 shows that $\mathcal{T}_{\mathbb{K}^N\to\mathbb{K}^N}^{\boldsymbol{F}}$ preserves whiteness if and only if $\boldsymbol{F}$ is a unitary matrix. Actually, it turns out that the only uniform filter banks which define a unitary transform are critically decimated (CD) paraunitary filter banks [1, sec. 6.4.3]. Moreover, since these filter banks involve matched[2] analysis and synthesis filters, FIR analysis filters lead to FIR synthesis filters. We conclude that the only filter banks which both guarantee PW and PR are CD-FIR paraunitary filter banks.

In this category, two famous examples are worth citing:

- if $\mathbb{K} = \mathbb{R}$: CD-PR-FIR cosine modulated filter banks with matched analysis and synthesis filters [1, sec. 6.6], *e.g.* modified DCT (MDCT) filter banks [17]: $\forall k \in [0 \dots M\text{-}1]$, $\forall t \in \mathbb{Z}$,
$$X_{k,t} = \sum_{n\in\mathbb{Z}} p_n W_{Mt-n} \cos\left(\frac{\pi}{M}\left(k+\frac{1}{2}\right)\left(n+\frac{M+1}{2}\right)\right)$$
where p_n is the impulse response of the *prototype filter*;
- if $\mathbb{K} = \mathbb{C}$: CD-PR-FIR-DFT filter banks involving matched analysis and synthesis filters [1, sec. 6.5]. It is easy to check that if the DFT is replaced by a GDFT as defined in Section III-B2 with any $\phi, \tau \in \mathbb{R}$, then the resulting filter bank is still paraunitary: $\forall k \in [0 \dots M-1]$, $\forall t \in \mathbb{Z}$,
$$X_{k,t} = \sum_{n\in\mathbb{Z}} p_n W_{Mt-n} e^{+\frac{i2\pi}{M}(k+\phi)(n+\tau)}. \tag{2}$$

Whereas the design of the prototype filter is quite flexible in the first example [1, sec. 6.6], the second one is much more constrained [1, sec. 6.5]. Indeed, it corresponds to a generalised STFT (GSTFT) involving contiguous, non-overlapping windows, and the optimal design in terms of stop-band attenuation leads to rectangular windows (*i.e.* in equation (2), $p_{-n} = \frac{1}{\sqrt{M}}$ if $0 \leq n < M$, and 0 otherwise). However, if a more flexible design is desired, it is possible to resort to recursive[3] analysis and synthesis filters [18] [1, sec. 6.5] and to a specific method for designing the prototype filter [19].

Lastly, note that any product of unitary matrices is a unitary matrix. Consequently, it is possible to combine any CD-PR-FIR paraunitary filter banks in order to produce tree-structured non-uniform PR multirate filter banks which preserve whiteness. This approach is very flexible and allows us to design various TF transforms, including wavelet transforms [1, chap. 8].

IV. Preservation of whiteness in real to complex transforms

In this section, we characterise invertible whiteness-preserving transforms from $\mathbb{R}^N$ to $\mathbb{C}^P$, and we provide a few examples of such spectral and TF transforms.

A. General study

Proposition 5 (PW in real to complex transforms)**.** *Let $P, N \in \{1, \dots, +\infty\}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{P\times N}$. Then $\mathcal{T}_{\mathbb{R}^N\to\mathbb{C}^P}^{\boldsymbol{F}}$ preserves whiteness if and only if $2P \leq N$ and $\downarrow\boldsymbol{F}\downarrow\boldsymbol{F}^H = \boldsymbol{I}$.*

Proof. Let $W = \{W_n\}_{0<\leq n<N}$ be a real white noise of variance σ^2, and let $X = \mathcal{T}_{\mathbb{R}^N\to\mathbb{C}^P}^{\boldsymbol{F}}(W)$. Then $\boldsymbol{\mu}_X = \boldsymbol{F}\boldsymbol{\mu}_W = \boldsymbol{0}$, $\boldsymbol{\Phi}_X = \mathbb{E}[XX^T] = \boldsymbol{F}\boldsymbol{\Gamma}_W\boldsymbol{F}^T = \sigma^2\boldsymbol{F}\boldsymbol{F}^T$ and $\boldsymbol{\Gamma}_X = \mathbb{E}[XX^H] = \boldsymbol{F}\boldsymbol{\Gamma}_W\boldsymbol{F}^H = \sigma^2\boldsymbol{F}\boldsymbol{F}^H$. Therefore X is a proper complex white noise of variance σ^2 as defined in Definition 6 if and only if $\boldsymbol{F}\boldsymbol{F}^T = \boldsymbol{0}$ and $\boldsymbol{F}\boldsymbol{F}^H = \boldsymbol{I}$, which is equivalent to $\downarrow\boldsymbol{F}\downarrow\boldsymbol{F}^H = \boldsymbol{I}$. □

[1] In the literature, the term PR generally means that the output signal is a delayed and scaled version of the input. In this document, we use this term in a restricted sense, *i.e.* the output is *equal* to the input.

[2] Two impulse responses h_n and g_n are *matched* when $g_n = \overline{h_{-n}}$.

[3] When recursive filters are employed, matrix $\boldsymbol{F}$ no longer belongs to $\mathrm{M}_{\mathbb{K}}^{N\times N}$. Therefore GDFT filter banks based on stable recursive filters can only be applied to a limited class of second order processes, which still includes white noise and WSS processes.

Proposition 6 (PW in invertible real to complex transforms)**.** *Let $P, N \in \{1, \dots, +\infty\}$, and $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{P\times N}$. Then $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ is invertible and preserves whiteness if and only if $N = 2P$, $\boldsymbol{F}^H \in \mathrm{M}_{\mathbb{C}}^{N\times N}$, and $\downarrow\!\boldsymbol{F}$ is a unitary matrix: $\downarrow\!\boldsymbol{F} \downarrow\!\boldsymbol{F}^H = \downarrow\!\boldsymbol{F}^H \downarrow\!\boldsymbol{F} = \boldsymbol{I}$.*

Proof. According to Proposition 2, $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ is invertible if and only if $N \leq 2P$ and if there is $\boldsymbol{G} \in \mathrm{M}_{\mathbb{C}}^{N\times P}$ such that $\overrightarrow{\boldsymbol{G}} \downarrow\!\boldsymbol{F} = \boldsymbol{I}$. Moreover, according to Proposition 5, $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ preserves whiteness if and only if $2P \leq N$ and $\downarrow\!\boldsymbol{F} \downarrow\!\boldsymbol{F}^H = \boldsymbol{I}$. Both assertions hold if and only if $N = 2P$ and $\overrightarrow{\boldsymbol{G}} \downarrow\!\boldsymbol{F} \downarrow\!\boldsymbol{F}^H = \overrightarrow{\boldsymbol{G}} = \downarrow\!\boldsymbol{F}^H$, which is equivalent to $\boldsymbol{F}^H \in \mathrm{M}_{\mathbb{C}}^{N\times N}$ and $\downarrow\!\boldsymbol{F} \downarrow\!\boldsymbol{F}^H = \downarrow\!\boldsymbol{F}^H \downarrow\!\boldsymbol{F} = \boldsymbol{I}$. □

B. Examples

1) Generalised discrete Fourier transform: Let $P < +\infty$ and $N = 2P$. The GDFT X of a real-valued signal W of length N is defined as in equation (1), but only for $k \in [0 \dots P-1]$. The entries of the corresponding matrix $\boldsymbol{F} \in \mathrm{M}_{\mathbb{C}}^{P\times N}$ are $F_{kn} = \frac{1}{\sqrt{N}} e^{-\frac{i2\pi}{N}(k+\phi)(n+\tau)}$. It can be easily verified that if $\phi = \frac{1}{2}$ (which encompasses the *odd-frequency* DFT [20]), then $\forall \tau \in \mathbb{R}$, $\downarrow\!\boldsymbol{F}$ is a unitary matrix. Proposition 6 then proves that $\mathcal{T}^{\boldsymbol{F}}_{\mathbb{R}^N\to\mathbb{C}^P}$ preserves whiteness and that its inverse is defined as $\forall n \in [0 \dots N-1]$,

$$W_n = \sqrt{\frac{2}{P}} \Re\left(\sum_{k=0}^{P-1} X_k e^{+\frac{i\pi}{P}(k+\frac{1}{2})(n+\tau)}\right).$$

2) GDFT and MDFT filter banks: Let $N = P = +\infty$. As mentioned in Section III-B3, CD-PR-FIR-GDFT filter banks involving matched analysis and synthesis filters are a particular case of paraunitary filter banks. However, if in equation (2) M is even ($M = 2Q$) and $\phi = \frac{1}{2}$ as in Section IV-B1, then the corresponding matrix can be arranged in the form $\downarrow\!\boldsymbol{F}$, where $\boldsymbol{F}$ is defined by keeping only the values of k in $\{0 \dots Q-1\}$, and satisfies all conditions in Proposition 6. Therefore $\boldsymbol{F}$ defines an invertible whiteness-preserving real to complex TF transform.

As mentioned in Section III-B3, the design of CD-PR-FIR-GDFT filter banks involving matched analysis and synthesis filters is highly constrained, since the optimal solution is the GSTFT involving contiguous, non-overlapping rectangular windows. If a more flexible design is desired, it is possible to resort to recursive filters. However, there are other ways of designing uniform invertible real to complex TF transforms which preserve whiteness. One of them consists in using tree-structured CD-PR-FIR filter banks. For instance, it is possible to connect the Q outputs of a Q-channel MDCT filter bank, to Q two-channel GSTFT involving contiguous, non-overlapping rectangular windows. The matrix corresponding to the resulting $M = 2Q$-channel filter bank can be arranged in the form $\downarrow\!\boldsymbol{F}$, where $\boldsymbol{F}$ satisfies all conditions in Proposition 6. Thus it achieves an invertible, whiteness preserving real to complex TF transform. It can be easily verified that this transform corresponds to the MDFT filter bank [21].

V. Discussion

In TF analysis, the conclusion in Section III-B3 was that PR and PW are guaranteed by CD-FIR paraunitary filter banks. However, by focusing only on the PW property, we have lost sight of one important concern: What about the frequency selectivity of the filters? Indeed, paraunitarity does not necessarily imply high frequency selectivity. Actually, frequency (and time) selectivity is important when we deal with processes which slightly deviate from strict whiteness.

A. Wide sense stationary processes

If the input is a WSS process whose PSD has low dynamics and slow variations, then frequency-selective CD-FIR paraunitary filter banks tend to produce an output where TF bins are decorrelated:

- Decorrelation between remote, non-overlapping subbands is due to the low dynamics of the PSD and high attenuation in the stop-band. Indeed, the resulting spectral overlap is negligible, and it is well-known that spectral representations of WSS processes are uncorrelated on non-overlapping frequency intervals [2].
- Decorrelation between adjacent, overlapping subbands, as well as decorrelation over time, is due to the PW property and to the slow variations of the PSD. Indeed, the process can be considered as approximately white in a local area of the spectral domain.

B. Uncorrelated random processes

The case of uncorrelated processes (*i.e.* processes whose samples are uncorrelated but may have different variances) is easily addressed. Indeed, it is obvious that any FIR filter bank produces a TF output where remote, non-overlapping time frames are uncorrelated. If moreover the input is an uncorrelated process whose temporal power variations are slow, then decorrelation between adjacent, overlapping time frames, as well as decorrelation over frequency, is due to the PW property. Indeed, the process can be considered as approximately white in a local area of the time domain.

VI. Conclusions

In this document, we were interested in designing transforms from the time domain to the spectral or TF domain, which introduce as little correlation as possible in the spectral or TF output. In order to properly formulate this idea, we formally defined the concept of preservation of whiteness (PW), that was firstly used in [13] in a restricted framework. We have listed a number of invertible transforms which satisfy this property, such as the unitary DCT and GDFT, and CD-FIR paraunitary filter banks, including MDCT filter banks and some GDFT filter banks. Often, a complex-valued output is desired while the input signal is real-valued. The proposed study gave us some interesting insights about how to design such a transform. In the case of GDFT or GSTFT for instance, the frequency index is shifted by $\frac{1}{2}$ in order to produce a proper complex output distribution. Finally, it is useful to mention that in the particular case of Gaussian processes, the above results can be interpreted in a stronger sense: uncorrelated random variables are independent, a (real or proper complex) white noise is an independent and identically distributed (IID) Gaussian process, and a proper complex random process is a circularly-symmetric complex Gaussian process [22].

Regarding future work, we have shown in Section V that applying CD-FIR paraunitary filter banks, either to WSS processes having a smooth PSD, or to uncorrelated processes having smooth temporal power variations, tends to produce an output where TF bins are decorrelated. However, it would be helpful to mathematically prove and accurately quantify this decorrelation property, and to extend its study to non-stationary processes having a smooth power density over both time and frequency. In other respects, we have noticed in Sections III-B3 and IV-B2 that the design of CD-FIR-GDFT filter banks which guarantee both PR and PW is highly constrained and that some flexibility can be brought in by using stable recursive filters. However such filters can only be applied to a limited class of second order processes (those whose power does not increase too fast as a function of the time index). Therefore an appropriate mathematical framework is needed in order to properly handle recursive filters. Finally, the initial motivation for introducing the PW property was to produce a spectral or TF transform which better fits the statistical decorrelation assumption used in a number of probabilistic frameworks designed for various applications. Therefore some numerical simulations should be carried out, in order to check whether using whiteness-preserving transforms actually improves the performance of existing methods in such applications.

References

[1] Alfred Mertins, *Signal Analysis: Wavelets, Filter Banks, Time-Frequency Transforms and Applications*, John Wiley & Sons, New York, NY, USA, Jan. 1999.

[2] Hisashi Kobayashi, Brian L. Mark, and William Turin, *Probability, Random Processes, and Statistical Analysis*, chapter Spectral representation of random processes and time series, pp. 343–399, Cambridge University Press, Cambridge, UK, 2012.

[3] Mikkel N. Schmidt and Hans Laurberg, "Non-negative matrix factorization with Gaussian process priors," *Computational Intelligence and Neuroscience*, 2008, Article ID 361705, 10 pages.

[4] Paris Smaragdis, "Probabilistic decompositions of spectra for sound separation," in *Blind Speech Separation*, S. Makino, T-W Lee, and H. Sawada, Eds., pp. 365–386. Springer, 2007.

[5] Tuomas Virtanen, Ali Taylan Cemgil, and Simon J. Godsill, "Bayesian extensions to non-negative matrix factorisation for audio signal modelling," in *Proc. 33th IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Las Vegas, Nevada, USA, Apr. 2008, pp. 1825–1828.

[6] Cédric Févotte, Nancy Bertin, and Jean-Louis Durrieu, "Nonnegative matrix factorization with the Itakura-Saito divergence. With application to music analysis," *Neural Computation*, vol. 21, no. 3, pp. 793–830, Mar. 2009.

[7] Alexey Ozerov, Cédric Févotte, and Maurice Charbit, "Factorial scaled hidden Markov model for polyphonic audio representation and source separation," in *Proc. 8th IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, New Paltz, New York, USA, Oct. 2009, pp. 121–124.

[8] Onur Dikmen and Ali Taylan Cemgil, "Gamma Markov random fields for audio source modeling," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 3, pp. 589–601, Mar. 2010.

[9] Gotham J. Mysore, Paris Smaragdis, and Bhiksha Raj, "Non-negative hidden Markov modeling of audio with application to source separation," in *Proc. 9th international Conference on Latent Variable Analysis and Signal Separation (LCA/ICA)*, St. Malo, France, Sept. 2010, 8 pages.

[10] Kazuyoshi Yoshii, Ryota Tomioka, Daichi Mochihashi, and Masataka Goto, "Infinite positive semidefinite tensor factorization for source separation of mixture signals," in *Proc. 30th International Conference on Machine Learning (ICML)*, Atlanta, Georgia, USA, June 2013, pp. 576Ð–584.

[11] Roland Badeau, "High resolution NMF for modeling mixtures of non-stationary signals in the time-frequency domain," Tech. Rep. 2012D004, Telecom ParisTech, Paris, France, July 2012.

[12] Roland Badeau and Mark. D. Plumbley, "Multichannel high resolution NMF for modelling convolutive mixtures of non-stationary signals in the time-frequency domain," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 22, no. 11, Nov. 2014, 11 pages.

[13] Roland Badeau and Mark D. Plumbley, "Probabilistic time-frequency source-filter decomposition of non-stationary signals," in *Proc. 21st European Signal Processing Conference (EUSIPCO)*, Marrakech, Morocco, Sept. 2013, 5 pages.

[14] Ali Taylan Cemgil and Simon J. Godsill, "Probabilistic phase vocoder and its application to interpolation of missing values in audio signals," in *Proc. 13th European Signal Processing Conference (EUSIPCO)*, Antalya, Turkey, Sept. 2005, 4 pages.

[15] Anil K. Jain, "A sinusoidal family of unitary transforms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 4, pp. 356–365, Oct. 1979.

[16] Gincarlo Bongiovanni, Paolo Corsini, and Graziano Frosini, "One-dimensional and two-dimensional generalised discrete Fourier transforms," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 24, no. 1, pp. 97–99, Feb. 1976.

[17] Henrique Malvar, *Signal Processing with Lapped Transforms*, Artech House, Boston, Massachusetts, USA, 1992.

[18] John Håkon Husøy and Tor A. Ramstad, "Application of an efficient parallel IIR filter bank to image subband coding," *Signal Processing*, vol. 20, no. 4, pp. 279–292, Aug. 1990.

[19] Markku Renfors and Tapio Saramäki, "Recursive Nth-band digital filters- Part I: Design and properties," *IEEE Transactions on Circuits and Systems*, vol. 34, no. 1, pp. 24–39, Jan. 1987.

[20] Jean-Louis Vernet, "Real signals fast Fourier transform: Storage capacity and step number reduction by means of an odd discrete Fourier transform," *Proceedings of the IEEE*, vol. 59, no. 10, pp. 1531–1532, Oct. 1971.

[21] Tanja Karp and Norbert J. Fliege, "Modified DFT filter banks with perfect reconstruction," *IEEE Transactions on Circuits and Systems—Part II: Analog and Digital Signal Processing*, vol. 46, no. 11, pp. 1404–1414, Nov. 1999.

[22] Bernard Picinbono, "Second-order complex random vectors and normal distributions," *IEEE Transactions on Signal Processing*, vol. 44, no. 10, pp. 2637–2640, Oct. 1996.

Dépôt légal : 2016 – 1er trimestre
Imprimé à Télécom ParisTech – Paris
ISSN 0751-1345 ENST D (Paris) (France 1983-9999)

Télécom ParisTech

Institut Mines-Télécom - membre de ParisTech

46, rue Barrault - 75634 Paris Cedex 13 - Tél. + 33 (0)1 45 81 77 77 - www.telecom-paristech.fr

Département TSI