An optimally concentrated Gabor transform for localized time-frequency components

Benjamin Ricaud, Guillaume Stempfel, Bruno Torresani, Christoph Wiesmeyr, Hélène Lachambre, Darian Onchis

HAL Id: hal-00904813
https://hal.science/hal-00904813

Submitted on 23 Mar 2016

An optimally concentrated Gabor transform for localized time-frequency components*

Benjamin Ricaud   Guillaume Stempfel   Bruno Torrésani   Christoph Wiesmeyr

Hélène Lachambre   Darian Onchis

November 13, 2013

Abstract

Gabor analysis is one of the most common instances of time-frequency signal analysis. Choosing a suitable window for the Gabor transform of a signal is often a challenge for practical applications, in particular in audio signal processing. Many time-frequency (TF) patterns of different shapes may be present in a signal and they can not all be sparsely represented in the same spectrogram. We propose several algorithms, which provide optimal windows for a user-selected TF pattern with respect to different concentration criteria. We base our optimization algorithm on l^p-norms as measure of TF spreading. For a given number of sampling points in the TF plane we also propose optimal lattices to be used with the obtained windows. We illustrate the potentiality of the method on selected numerical examples.

1   Introduction

The Gabor transform is widely used for the analysis of non-stationary signals, in particular in audio signal processing. In contrast to the classical Fourier Transform it allows for detection and estimation of localized time-frequency (TF) components or time-evolving frequency components. It is also used for modifying or filtering the signal in a limited TF region (through Gabor multipliers). The Gabor transformation coefficients are obtained by testing the signal against time frequency shifted copies of a window function. The TF representation hence depends crucially on the choice of this predefined window and on the lattice giving the coordinates of the TF shifts.

This freedom of choice however rises the important question of which window and lattice are the best for some given signal. Indeed, this choice is crucial as the precision of the Gabor transform in time and frequency is limited by the uncertainty principle and it is not possible to have arbitrarily good time and frequency resolution at the same time. It is in particular the decay of the chosen window on the time and frequency side that allows balancing time and frequency precision of the corresponding Gabor transform. It often requires a delicate tuning from the user based on a visual appreciation of the spectrogram. Moreover, in some cases changing the size of the window is not sufficient for obtaining sharply concentrated information in the TF plane. Typical examples, where one needs to control more than the width of the window are signals that contain chirps, i.e. frequency components with time-varying frequency.

Although often neglected, the choice of the Gabor lattice is an important issue. It does not directly concern precision in time and frequency but more the computational effort needed to process the representation of the signal and its modification. A dense lattice with many sampling points implies a high redundancy of the representation and will therefore require a high computational cost and a heavy load for the computer memory. On the other hand, if one samples too coarsely, firstly the TF representation will be coarse, preventing the detection of fine details. Secondly, if one chooses a bad sampling strategy, the Gabor transform may not even be invertible or it may result in an unstable inversion process.

The question of the adaptation of the window to a signal of interest had been considered in [11], where it has been shown that quite often a global adaptation does not yield sensible results. A typical example is a musical signal which contains harmonic sounds from string or wind instruments and impact sounds from percussive instruments.
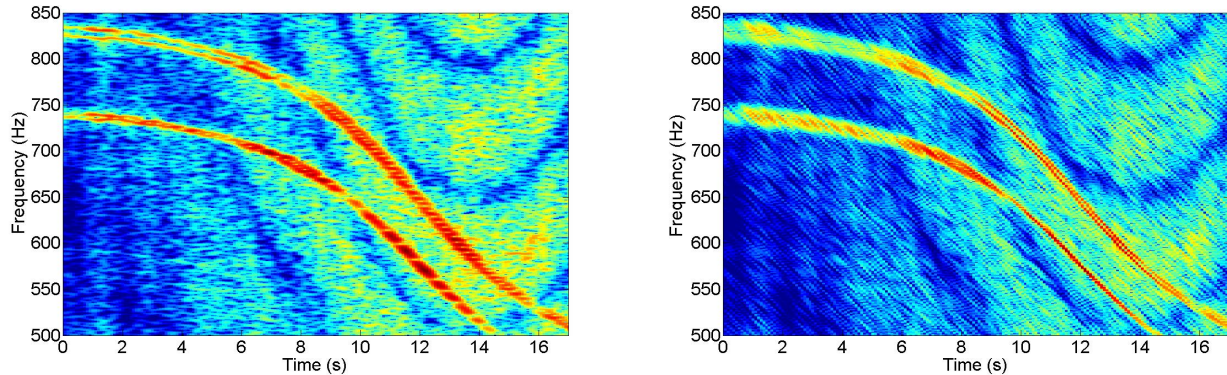
Figure 1: Audio recording of an aircraft passing above the recording device (doppler effect). Left: spectrogram computed with a Gaussian window. Right: spectrogram computed with the optimal window. The optimization reveals the fine structure (i.e. 2 separated frequency components) in the TF domain where the optimization is performed, but kind of blurs the picture in other regions (namely the first 4 seconds) where another window would be better.

The associated TF patterns can be described as lines parallel to the time axis for constant frequencies and lines parallel to the frequency axis for percussive events. It is hence impossible to find a window (and lattice) suitable for the analysis of the entire signal. However it is often useful to be able to represent one particular pattern with high precision inside the signal even if the cost is a worse representation in other TF regions. Motivations are either for determining the characteristics of this TF shape, for highlighting it among other TF signatures often nearby located or for helping a systematic detection of every occurrence of this special pattern inside the signal.

In the following we propose an algorithm for obtaining a Gabor window and lattice adapted to a particular TF pattern via the solution of an optimization problem. According to our definition, a window is adapted when it concentrates the pattern optimally, in other words when the pattern becomes the sparsest possible in the TF plane. Our procedure hence minimizes the sparsity in the TF plane of a chosen component embedded in the whole signal. The idea is similar to the one of Jaillet-Torrésani [11]. However, our optimization scheme is non-parametric and optimizes over all possible functions (windows) in the signal space. In addition their sparsity measure (entropy) is a particular case of our more general ℓ^p-norm measure. These quantities can be related via Rényi entropies [13]. The present work has been inspired by [5] in which an algorithm is proposed for building the maximally concentrated window fitting a given shape in the TF (Ambiguity) plane. The optimization approach using ℓ^p-norms is similar however there are many differences: in the motivation, in the algorithm and the more detailed theoretical study. In addition, our method is numerically faster as it relies on an iterative gradient method instead of the diagonalization of a large matrix (Gabor multiplier).

To illustrate the idea, two spectrograms of the same signal using different Gabor windows are shown in Fig. 1. On the left, the spectrogram has been computed with a standard window (Gaussian function). One can notice two frequency components evolving in time (thick red lines). The region, where we seek to have an optimal representation, is chosen to be one of the main components at a recording time of around 12 seconds and has a length of around 1 second. The spectrogram using the optimal window is shown on the right: the frequency components get thinner (sparser) between 10 and 16 seconds recording time. Moreover, as the structure is resolved in more details, it reveals that the components are more complex than single lines.

In addition to the optimal window our method also provides an adapted lattice, suited for the optimal window. It is constructed to give an efficient placement of the Gabor atoms in the TF plane. For a fixed redundancy, the purpose is to provide the best tiling of the TF plane in order to represent as accurately as possible the energy distribution of the signal, with the given window. A comprehensive study of the Gabor transform for general lattices can be found in [18]. For example, a standard Gaussian window has a round-shaped ambiguity function and the optimal lattice would intuitively be hexagonal (like the problem of orange packing). As in [18], we have chosen three tuning parameters to adapt the lattice: the time and frequency step sizes (as in standard lattices) and the shearing. Later we will give explicit sampling strategies for a class of windows called *generalized Gaussians*, as this class in particularly adapted to the representation of audio signals.

2 Preliminaries

Here we will outline the theory of discrete Gabor transforms, for a more complete picture of the theory and algorithms we refer to [6]. An in-depth investigation of the computational complexities of the finite discrete Gabor transform can be found in [14, 18]. The continuous theory is described in [9, 12].

To be able to define the Gabor transform we make use of the *translation* and *modulation* operators denoted by T and M respectively. For $x, \xi \in \{0, \ldots, N-1\}$ we define them as

$$T_x g(t) = g(t-x) \text{ and} \tag{1}$$

$$M_\xi g(t) = e^{2\pi i \xi t/N} g(t), \tag{2}$$

for any $g \in \mathbb{C}^N$. At this point we note that all the indices are to be interpreted modulo N throughout the rest of this contribution (periodic boundary conditions).

With this notation we define the Gabor family with respect to some sampling set $\Lambda \subseteq \mathbb{Z}_N^2$ and some window function g as

$$\mathcal{G}(g, \Lambda) = \{M_\xi T_x g : (x, \xi) \in \Lambda\}. \tag{3}$$

The set $\mathbb{Z}_N^2$ is called *time-frequency plane*, as every point corresponds to a *time-frequency shift* of the window g. The Gabor transform $\mathcal{V}_g f$ of a given signal f with respect to g and sampling set Λ is given by its transform coefficients:

$$\{\mathcal{V}_g f(x, \xi) = \langle f, M_\xi T_x g \rangle : (x, \xi) \in \Lambda\}, \tag{4}$$

where $\langle \cdot, \cdot \rangle$ stands for the scalar product in $\mathbb{C}^N$.

It is possible to perfectly reconstruct the signal from its coefficients if and only if $\mathcal{G}(g, \Lambda)$ is a frame [3], i.e. when there exist two bounds $A, B > 0$ such that for all f in the signal space

$$A\|f\|_{\ell^2(\mathbb{C}^N)}^2 \leq \|\mathcal{V}_g f\|_{\ell^2(\Lambda)}^2 \leq B\|f\|_{\ell^2(\mathbb{C}^N)}^2.$$

The frame property of the family $\mathcal{G}(g, \Lambda)$ is equivalent to the existence of a (possibly non-unique) dual window g_d, such that the inverse transform is given by

$$f = \sum_{(x,\xi) \in \Lambda} \langle f, M_\xi T_x g \rangle M_\xi T_x g_d. \tag{5}$$

We call a frame tight if the window is dual to itself.

In the definition of frames we have been overly careful with denoting norms, throughout the rest of the paper we will denote the p-norm by $\|\cdot\|_p$, as the underlying space will always be clear.

The space of time-frequency centered windows $\mathcal{C}$. In the following, we will use p-norms as measure of optimality. Due to the property that for any $\xi, x \in \Lambda$ and $g \in \ell^2$, $\|M_\xi T_x g\|_p = \|g\|_p$, any time-frequency shifted copy of a solution will be a solution. To avoid this ambiguity, we introduce the set of time-frequency centered windows $\mathcal{C}$ in which the optimization will be processed. To define this set we first define what is a TF centered window. Let us introduce the mean value in time of $g \in \mathbb{C}^N$:

$$\mu_t(g) = \arg\left(\sum_{n=0}^{N-1} e^{2i\pi \frac{n}{N}} |g(n)|^2\right), \tag{6}$$

which is the definition of the mean value for functions on periodic domains (von Mises' mean). The mean value in frequency $\mu_f(g)$ is obtained by replacing g by its Fourier transform in (6). The TF centered version of g reads

$$M_{-\mu_f} T_{-\mu_t} g.$$

The set $\mathcal{C}$ of TF centered functions is

$$\mathcal{C} = \left\{g \in \ell^2(\mathbb{C}^N), \mu_t(g) = \mu_f(g) = 0\right\}.$$

3 Optimal window

For a given Gabor lattice, we call a window optimal if it maximally concentrates the energy of the signal in the TF plane. Given a window with unit energy we use the ℓ^p-norm with $p > 2$ (resp. $p < 2$) as a measure of spread. Maximizing ($p > 2$) or minimizing ($p < 2$) this norm leads to sparse representations (see Ricaud-Torrésani [13]). In the following we will choose $p > 2$. In this case the quantity $\|\mathcal{V}_g f\|_p^p$ is differentiable with respect to g and we can use a standard gradient method to find the optimum (here the maximum).

The first optimization problem reads as

$$g_{\text{opt}} = \underset{g \in \mathcal{C}, \|g\|_2 = 1}{\text{Argmax}} \|\mathcal{V}_g f\|_p^p = \underset{g \in \mathcal{C}, \|g\|_2 = 1}{\text{Argmax}} \sum_{\xi, x \in \Lambda} |\langle M_\xi T_x g, f \rangle|^p, \tag{7}$$

with $p > 2$. Under this latter condition, the p-norm is a measure of sparsity and maximizing this quantity leads to a sparser TF representation of f. The solution g_{opt} of (7) (whenever it exists) is the optimal window. However, the set of functions g such $\|g\|_2 = 1$ is not concave. Hence the above formulation does not lead to a concave problem and convergence issues have to be investigated. In addition, any time frequency shifted copy $M_\xi T_x g_{\text{opt}}$ of the optimal window is also an optimal window. This is why we restrict the space of windows to the ones in $\mathcal{C}$, centered in time and frequency. We also apply a TF centering to the function f, see Sec. 3.4 for the signal preprocessing.

It is also interesting to look at the first order term when p is close to 2 and $\|f\|_2 = 1$: it yields the entropic optimization problem (c.f. Remark 4),

$$\underset{g \in \mathcal{C}, \|g\|_2 = 1}{\text{Argmax}} \sum_{\xi, x \in \Lambda} -|\langle M_\xi T_x g, f \rangle|^2 \ln |\langle M_\xi T_x g, f \rangle|^2, \tag{8}$$

which also maximizes the sparsity and has been used in Jaillet-Torresani [11] for the optimization of a parametric window. We may also generalize their approach in the following way. We introduce the parametrized window (chirped and dilated Gaussian-like window),

$$g_{\sigma,s}(t) = \sqrt[4]{\frac{2}{N\sigma}} e^{-\pi \frac{t^2}{N\sigma} + i\pi s t^2 \frac{N+1}{N}}, \tag{9}$$

where $\sigma > 0$ is the width in time and $s \in \mathbb{R}$ is the chirping parameter. The number of samples of the signal is N. The optimization problem becomes:

$$\underset{\sigma, s}{\text{Argmax}} \|\mathcal{V}_{g_{\sigma,s}} f\|_p^p. \tag{10}$$

The projection on the non-concave set disappears but still the cost function may not be concave in general.

Eventually, another optimization problem which is concave is the following:

$$\underset{g}{\text{Argmax}} \|\mathcal{V}_g f\|_p^p - \lambda \|h - g\|_2^2, \qquad p > 2, \tag{11}$$

where h is for example a standard Gaussian function centered in time and frequency. An elliptic Gaussian or any other well-concentrated function may also do the trick. The parameter $\lambda > 0$ allows one to tune the shape of the optimal window. This formulation as well as (10) are of interest for practical applications: one often wants to obtain a window which is not too eccentric. Indeed, in most applications, TF localization is as important as sparsity. Localization in the TF plane is the reason why one chooses the Gabor transform to analyze a signal. If this property is lost the Gabor transform coefficients are of little interest. For example an optimal window whose TF transform is very elongated in one direction of the plane won't be a good optimal window in practice as the localization in that direction will be poor (although the localization in the perpendicular direction is extremely good). The interpretation of the spectrogram will be difficult and filtering based on localized TF regions will not be efficient. In (11) the localization is constrained by the choice of λ and in (10) it is constrained by limits on the values σ can take (upper and lower bound).

Before going to the details of the window optimization, let us point out some important remarks.

Remark 1. *The optimization is done on a fixed Gabor lattice which may not be the ideal one for the detection of the TF pattern. Hence, after optimizing the window the next step of our optimization process is to adapt a new lattice $\Lambda^{(1)}$ to the optimal window g_{opt}. This new sampling strategy depends on the shape of the ambiguity function of the window. It seeks to tile the TF plane optimally with such atoms, a detailed explanation can be found in Section 4.*

Remark 2. *We enforce sparsity of the Gabor representation from an analysis point of view. Therefore it is not necessary to compute a dual Gabor window or an inverse Gabor transform during the optimization process.*

3.1 The non-concave, non-parametric case

Problem (7) is non concave but satisfies the hypotheses stated in [1, Sec. 5], whenever p is a rational number. To prove that, let us first introduce $\widetilde{g} = (\mathsf{Re}(g), \mathsf{Im}(g)) \in \mathbb{R}^{2N}$ and rewrite (7) as:

$$\underset{\widetilde{g} \in \mathbb{R}^{2N}}{\operatorname{Argmin}} \mathcal{F}(\widetilde{g}) + i_C(\widetilde{g}), \tag{12}$$

where

$$\mathcal{F}(\widetilde{g}) = - \sum_{\xi, x \in \Lambda} \left(\sqrt{\mathsf{Re}\left(\langle g, T_x M_\xi f \rangle \right)^2 + \mathsf{Im}\left(\langle g, T_x M_\xi f \rangle \right)^2} \right)^p, \qquad p > 2,$$

and i_C is the indicator function of the closed set $C = \{g : \|g\|_2 = 1\}$. Introducing

$$(K_{\xi, x}(\widetilde{g}, f))^2 = \mathsf{Re}\left(\langle g, T_x M_\xi f \rangle \right)^2 + \mathsf{Im}\left(\langle g, T_x M_\xi f \rangle \right)^2,$$

and replacing g and $T_x M_\xi f$ by $\mathsf{Re}(g) + i\mathsf{Im}(g)$ and $\mathsf{Re}(T_x M_\xi f) + i\mathsf{Im}(T_x M_\xi f)$ in the above expression shows that $K_{\xi, x}(\cdot, f)$ is a semialgebraic function. The right hand side is a polynomial in the elements of vector $\widetilde{g}$.

Then $\mathcal{F} = - \sum_{\xi, x \in \Lambda} (K_{\xi, x}(\cdot, f))^p$ and for any integer p, the real function $\mathcal{F}$ is a polynomial in the elements $\{K_{\xi, x}\}$. Hence $\mathcal{F}$ is semialgebraic which implies that it is a lower semicontinuous Kurdyka-Lojasiewicz function (see [1] for more details). It can be generalized to any p rational by writing $p = c/d$ and replacing $K_{\xi, x}$ by $K_{\xi, x}^d$. Concerning the lower bound on the cost function, $\mathcal{F}$ is strictly negative and $\mathcal{F} = 0$ (maximum) if and only if $g = 0$ since $\{T_x M_\xi f\}$ is a frame. The constraint $\|g\|_2 = 1$ (and $\{T_x M_\xi f\}$ is a frame) implies that the function $\mathcal{F} + i_C$ is bounded from below. Moreover $\mathcal{F}$ is differentiable and has a L-Lipschitz continuous gradient. To see that one has to differentiate with respect to $\mathsf{Re}(g)$ and $\mathsf{Im}(g)$. It is more convenient to differentiate (in the complex sense) with respect to $\overline{g} \in \mathbb{C}^N$ and use Cauchy-Riemann equations, this is what we do in the following. We have (see below for details of the calculation):

$$\nabla_{\overline{g}} \|\mathcal{V}_g f\|_p^p = \frac{p}{2} G_{f, \Lambda, \Gamma}(g), \tag{13}$$

where $G_{f, \Lambda, \Gamma}$ denotes the following time-frequency masking operator:

$$G_{f, \Lambda, \Gamma}(g) = \sum_{\xi, x \in \Lambda} \Gamma_g(x, \xi) \langle g, M_\xi T_x f \rangle M_\xi T_x f. \tag{14}$$

with mask $\Gamma_g(x, \xi) = |\langle M_\xi T_x f, g \rangle|^{\frac{p}{2} - 1}$. In [1] the authors show that under these conditions the forward-backward algorithm converges to a critical point of $\mathcal{F} + i_C$. This procedure iteratively computes

$$g_{i+1/2} = g_i + \gamma \nabla_{\overline{g}} \|\mathcal{V}_{g_i}(h)\|_p^p, \tag{15}$$

where $\gamma \in (0, L)$ and then projects the current iterate to the ℓ^2-sphere by

$$g_{i+1} = \frac{g_{i+1/2}}{\|g_{i+1/2}\|_2}.$$

Remark 3. *Let us first treat the particular case of $p = \infty$. We have the following result:*

$$\lim_{p \to \infty} \left(\sum_{\xi, x} |\langle M_\xi T_x g, f \rangle|^p \right)^{1/p} = \max_{\xi, x} |\langle M_\xi T_x g, f \rangle|.$$

The maximum of the scalar product is attained when $M_\xi T_x g = f$. The optimal window is hence equal to the signal (up to a translation and modulation).

The convergence results are summarized in the following theorem.

Theorem 1. *The forward-backward algorithm associated to the optimization problem (7) converges to a critical point g_{opt} which depends on p in the following manner: for any finite $p > 2$, g_{opt} is the limit of g_i when i tends to infinity of the gradient ascent (15) with $\nabla_{\overline{g}} \|\mathcal{V}_g(h)\|_p^p = \frac{p}{2} G_{h, \Lambda, \Gamma}(g)$ where $\Gamma_g(x, \xi) = |\langle M_\xi T_x g, f \rangle|^{\frac{p}{2} - 1}$.*

Proof. Convergence issues have been treated in the beginning of this section. we concentrate on the derivation of the gradient. Introduce $A(g, \overline{g}) = |\langle M_\xi T_x g, f \rangle|^2$, its derivative with respect to $\overline{g}$ (taking g as independent of $\overline{g}$, see [16] and refs. therein) is:

$$\nabla_{\overline{g}} A(g, \overline{g}) = \langle M_\xi T_x g, f \rangle T_x M_\xi f.$$

One can write $(p > 2)$

$$\nabla_{\overline{g}} \sum_{\xi, x} |\langle M_\xi T_x g, f \rangle|^p = \sum_{\xi, x} \nabla_{\overline{g}} A^{p/2}(g, \overline{g}) = \frac{p}{2} \sum_{\xi, x} A^{\frac{p}{2}-1}(g, \overline{g}) \nabla_{\overline{g}} A(g, \overline{g}).$$

$\square$

Remark 4 (Near $p = 2$, the Shannon entropy is not far)**.** *Introduce $\eta = 2 - p$. One can write $A^\eta = \exp(\eta \ln A) = 1 + \eta \ln A + \eta^2 \ln^2 A + \cdots$. When p is close to 2 the ℓ^p-norm is:*

$$\sum_{\xi, x} |\langle M_\xi T_x g, f \rangle|^p = \sum_{\xi, x} |\langle M_\xi T_x g, f \rangle|^2 |\langle M_\xi T_x g, f \rangle|^\eta$$

$$= 1 + \frac{\eta}{2} \sum_{\xi, x} |\langle M_\xi T_x g, f \rangle|^2 \ln |\langle M_\xi T_x g, f \rangle|^2 + \mathcal{O}(\eta^2)$$

For small η the second term is the leading one in the optimization process. The expression of its gradient is $\nabla_{\overline{g}}(A \ln A) = (1 + \ln A) \nabla_{\overline{g}} A$. Note that if $\{T_x M_\xi f\}$ is a tight frame, $\sum_{\xi, x} \nabla_{\overline{g}} A(g, \overline{g}) = C g$, ($C > 0$ being a constant depending on the frame). The leading term in the gradient will be

$$\frac{\eta}{2} \left(C g + G_{h, \Lambda, \Gamma}(g) \right),$$

where the operator $G_{h, \Lambda, \Gamma}(g)$ has the following mask:

$$\Gamma_g(x, \xi) = \ln |\langle M_\xi T_x g, f \rangle|^2.$$

3.2 The non-concave parametric case

In practice the Gabor transform is often used to visually estimate the behavior of some TF component. The analysis window g must not be too exotic in order to keep the intuitive interpretation of the transform coefficients. A first requirement is that the window should be intrinsically concentrated in time and frequency, a second one is that it should be localized in only one connected region of time and frequency. A third requirement is that the shape of the window should be close to an ellipsis in the TF plane. It would allow the window to have a privileged direction in the TF plane but not only restricted to either time or frequency. A sensible way to obtain such a function is to do the optimization in the subset of chirped Gaussians.

Let us introduce the subclass $\mathcal{G}$ of chirped Gaussians in $\mathbb{C}^N$ (or generalized Gaussians). A function in this class is of the form $M_\xi T_x \phi_{\sigma, s}$, where $\phi_{\sigma, s} = g_{\sigma, s}$ of (9), for all possible values of the dilation parameter $\sigma > 0$ and chirping parameter $s \in \mathbb{R}$. Note that $\|M_\xi T_x \phi_{\sigma, s}\|_2 = 1$. We want to maximize the following quantity over σ and s:

$$\|\mathcal{V}_{\phi_{\sigma, s}} f\|_p^p, \qquad p > 2.$$

As in the previous section the problem is not concave, it depends on f in a non trivial way even after factoring out the constant phase indeterminacy. Actually it is interesting to notice that if f is the sum of two Gaussians with same mean but with two different variances, the function $|\langle \phi_{\cdot, s=0}, f \rangle| : \sigma \to \mathbb{R}_+$ has two maxima (hence non-concave).

If we assume that f is such that $|\langle \phi_{\cdot, \cdot}, f \rangle|$ is bi-concave (concave in each variable separately) standard optimization algorithms will converge but may converge to a local maximum [8]. In the following we will add this property as a condition on our function f, which in practice covers many TF patterns. One has to keep in mind that there are also many examples of f which do not lead to concavity or bi-concavity. As a few examples we can list:

- the above example of a sum of 2 Gaussians with different spreading,

- a sum of 2 localized functions with different chirping values, forming a "cross" pattern in the TF plane,

- broad band signals.

The convergence of the optimization process is not guaranteed for these patterns.

Using the results of the previous section, the gradient is given by

$$\frac{\partial \|\mathcal{V}_{\phi_{\sigma,s}} f\|^p}{\partial \sigma} = p \ \mathsf{Re}[\langle G_{f,M}\phi_{\sigma,s}, \frac{\partial \phi_{\sigma,s}}{\partial \sigma}\rangle], \tag{16}$$

$$\frac{\partial \|\mathcal{V}_{\phi_{\sigma,s}} f\|^p}{\partial s} = p \ \mathsf{Re}[\langle G_{f,M}\phi_{\sigma,s}, \frac{\partial \phi_{\sigma,s}}{\partial s}\rangle], \tag{17}$$

where $G_{f,\Lambda,\Gamma}$ is the operator defined in (14) with window f and mask $\Gamma = |\langle M_\xi T_x f, \phi_{\sigma,s}\rangle|^{p-2}$. One can see it by writing for $p > 2$ even, any window g and function f:

$$\sum_{\xi,x} |\langle M_\xi T_x g, f\rangle|^p = \sum_{\xi,x} \langle M_\xi T_x g, f\rangle^{p/2} \langle f, M_\xi T_x g\rangle^{p/2}$$

and

$$\frac{\partial}{\partial s}\langle M_\xi T_x \phi_{\sigma,s}, f\rangle^{p/2} = \frac{p}{2}\langle \frac{\partial}{\partial s}\phi_{\sigma,s}, T_{-x}M_{-\xi}f\rangle\langle \phi_{\sigma,s}, T_{-x}M_{-\xi}f\rangle^{p/2-1}.$$

At each step of the gradient method, one computes:

$$\sigma_{n+1} = \sigma_n + \alpha p \ \mathsf{Re}[\langle G_{f,\Lambda,\Gamma^{(n)}}\phi_{\sigma,s}^{(n)}, \frac{\partial \phi_{\sigma,s}^{(n)}}{\partial \sigma}\rangle] \tag{18}$$

$$s_{n+1} = s_n + \alpha p \ \mathsf{Re}[\langle G_{f,\Lambda,\Gamma^{(n)}}\phi_{\sigma,s}^{(n)}, \frac{\partial \phi_{\sigma,s}^{(n)}}{\partial s}\rangle] \tag{19}$$

and $\phi_{\sigma,s}^{(n+1)} = \phi_{\sigma_{n+1},s_{n+1}}$.

The derivatives of the function $\phi_{\sigma,s}$ with respect to σ and s read:

$$\frac{\partial \phi_{\sigma,s}}{\partial \sigma}(t) = \left(\frac{\pi t^2}{\sigma^2 N} - \frac{1}{4\sigma}\right)\phi_{\sigma,s}(t), \qquad \frac{\partial \phi_{\sigma,s}}{\partial s}(t) = \frac{i\pi}{N}t^2\phi_{\sigma,s}(t).$$

Remark 5 (Control over the parameters). *At each step we can also control the value of the parameters σ and s. Additional constraints may be added in the loop. For example if one wants an optimal window with a limited spread in one direction, the value of σ may be forced to stay within some predefined bounds.*

Remark 6 (Gradient step and the BFGS optimization scheme). *Experiments show that it is difficult to find an appropriate gradient step to ensure the convergence of the algorithm (large steps lead to divergence, small steps to extremely slow convergence). Even an adaptive step is not satisfactory because of the shape of the function. Thus, we decided to resort to a second-order-like method. Since the computation of the second derivatives is hard, we used a BFGS (Broyden-Fletcher-Goldfarb-Shanno, see e.g. [7]) technique. Basically, BFGS, like other Quasi-Newton procedures, is an iterative algorithm which is an approximation of Newton's method, without the need to compute the inverse of the hessian matrix. This hessian matrix is estimated using the first-order gradient computed for each iteration. The larger the number of iterations is, the more accurate the hessian estimation will be. A line search is performed in order to find an acceptable step size (actually an α value which satisfies Wolfe conditions). The algorithm iterates until the gradient norm becomes small enough, which indicates that a local optimum has been found. Experimentally, under this form, the algorithm converges rather quickly. The BFGS procedure is designed for tackling concave or convex problems. Even though our objective function was not concave, it did not appear to be a problem in the computations. In our simulations, if the starting point of the optimization was reasonably chosen (typically a Gaussian time-support around half a second, with a chirping parameter equal to zero), the algorithm always converged to an acceptable solution. However, because of the non-concavity of the function, we cannot guarantee the quality of obtained solutions theoretically.*

3.3   The concave case

The optimization problem (11) is concave and the solution may be obtained using a gradient ascent. Starting from any initial g_0 the iterative algorithm computes at step $i + 1$:

$$g_{i+1} = (1 - \gamma\lambda)g_i + \gamma\lambda h + \gamma\nabla_{\overline{g_i}}\|\mathcal{V}_{g_i} f\|_p^p, \tag{20}$$

where $\gamma > 0$ is the gradient step. The expression of the gradient $\nabla_{\overline{g_i}}\|\mathcal{V}_{g_i} f\|_p^p$ depends on p and may be found in Th. 1.

Preliminary tests were not convincing in term of signal representation compared to the other cases. Moreover choosing γ and λ is delicate. That is why we concentrated more on the non-convave parametric problem which was more promising. We present it only in order to show that it is possible to find a concave optimization problem for the window optimization.

3.4 Localizing a pattern of interest in the TF plane (data preprocessing)

The preprocessing of the data before optimization consists of the 3 following steps.

Localization. Let us suppose that there is a localized TF pattern of interest embedded in the signal f. A Gabor transform with standard lattice and standard Gaussian window can be used to get a first crude estimate of its TF location and spreading. In order to optimize the Gabor window with respect to this particular pattern we may proceed as follow.

In the first step we extract this pattern by multiplying all the TF coefficient not in the area of interest with zeros and reconstruct (compute the inverse Gabor transform). The resulting signal, lets call it h, contains only the pattern (or more precisely a good approximation of it) to be represented in sparser way.

TF centering. By translating and modulating h we can center the selected component at the origin of the TF plane. Since the windows belong to the set of TF centered functions, doing so avoids unnecessary TF shifts in the optimization process. It also allows data reduction (see next step). The centering is performed according to the formula given in Sec. 2.

Data reduction. To further simplify the problem one can reduce the number of samples of h, which is at this point still equal to the number of samples of the original function. Often, it is possible to neglect part of the samples in time if the region of interest has a small time width. If the pattern covers only a small portion of the frequencies it is possible to downsample h without deleting information.

4 Adapted lattice

The choice of a Gabor lattice may be as important as the choice of a window. In this section we describe firstly what is the set of all possible Gabor lattices. Secondly we present what is a good lattice for a given window, in the sense that it optimally packs the windows in the TF plane (for a fixed transform redundancy). Thirdly, we show how to compute the Gabor transform with a non-standard lattice as fast as the one having a standard lattice.

In the discrete setting with signal length N any Gabor lattice Λ can be expressed by

$$\begin{pmatrix} a & 0 \\ s & b \end{pmatrix} \cdot \mathbb{Z}_N^2, \tag{21}$$

for some uniquely determined a, b which must divide N and $s = k\, b/\gcd(L/a, b)$ for some integer number k [10]. The time frequency plane is given by $\mathbb{Z}_N^2$, which is at the same time the finest (and densest) possible lattice. This means that we can generate all possible lattices by choosing a, b and s satisfying the conditions above. The redundancy of a lattice of the form (21) is given by $R = L/(ab)$, regardless of the parameter s.

As we have seen previously, dilated and chirped Gaussians play a central role in our optimization scheme. They possess good localization properties, are easy to control and allow for a good interpretation of the spectrogram. In order to choose the optimal lattice it seems sensible to seek to minimize the condition number of the resulting Gabor system, i.e. the ratio between the upper and lower frame bound B/A. Geometrical arguments have led to the conclusion that hexagonal patterns are the most appropriate for standard Gaussian functions in the continuous case [17].

For windows which are dilated and chirped versions of the standard Gaussian, the adapted lattice (which has the same properties as the hexagonal one for the round Gaussians) may be computed knowing the dilation and chirping parameters. We refer to [9, 4] for the continuous case and to [18, 10] for the discrete setting. It is shown in [9, Ex. 9.4.1] that two Gabor transforms with two different lattices are equivalent provided their respective windows are related by some dilation and chirping; the values of which are directly proportional to the lattice step and shearing values.

Under the assumption that hexagonal sampling is optimal for the standard Gaussian, a careful analysis of the connection between continuous and discrete Gabor transforms yields the theoretical optimal sampling points for a window of the form $\phi_{\sigma,s}$ defined in (9):

$$\Lambda = \sqrt{\frac{N}{R}} \begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix} \begin{pmatrix} \sqrt{\sigma} & 0 \\ 0 & 1/\sqrt{\sigma} \end{pmatrix} \begin{pmatrix} \frac{\sqrt[4]{3}}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt[4]{3}\sqrt{2}} & \frac{\sqrt{2}}{\sqrt[4]{3}} \end{pmatrix} \cdot \mathbb{Z}_N^2. \tag{22}$$

The 3 successive matrices correspond, from left to right, to the shearing, the dilation and the hexagonal shaping of the $N \times N$ square lattice $\mathbb{Z}_N^2$. Several entries of these matrices are irrational numbers which will of course not satisfy the requirements we mentioned above. Therefore we make a first approximation by choosing feasible a, b and s of Λ which are closest to the true values.

In the following we show in more details how the shearing operation is related to chirping in the signal domain. Moreover, we demonstrate how *a good approximation* of the Gabor transformation with a sheared lattice may be computed using the regular Gabor transform. Since we intend to compute optimal windows using non-regular lattices in Sec. 5, it is important to have a fast computation of the non-separable Gabor transform.

In the discrete setting, due to periodic conditions, the chirp functions are restricted to the ones which revolve around the TF plane s times for $s \in \mathbb{N}_0$. Let U_s denote the chirping operator on a signal $g \in \mathbb{C}^N$:

$$U_s g(n) = e^{i\pi s n^2 \frac{N+1}{N}} g(n), \tag{23}$$

and let Λ, Λ' be two Gabor lattices related by a shearing of s:

$$x' = x, \qquad \xi' = \xi + xs \qquad \forall (x', \xi') \in \Lambda', \ (x, \xi) \in \Lambda. \tag{24}$$

For an integer value of s one has [18]

$$\langle f, M_{\xi'} T_{x'} g \rangle = e^{is\pi x^2 \frac{N+1}{N}} \langle U_{-s} f, M_\xi T_x U_{-s} g \rangle. \tag{25}$$

To prove the equation above one needs to interchange chirping and translation. This is only possible for integer chirping parameters (periodic boundary conditions).

Under the same restriction, the operator $G_{g,\Lambda',\Gamma}$ involved in the optimization procedure[1], which is associated to a sheared lattice may be computed using:

$$
\begin{aligned}
G_{g,\Lambda',\Gamma}(f) &= \sum_{x',\xi'} \Gamma(x',\xi') \langle f, M_{\xi'} T_{x'} g \rangle M_{\xi'} T_{x'} g \\
&= \sum_{x',\xi'} \Gamma(x',\xi') \langle U_{-s}^* U_{-s} f, M_{\xi'} T_{x'} U_{-s}^* U_{-s} g \rangle U_{-s}^* U_{-s} M_{\xi'} T_{x'} g \\
&= U_{-s}^* \sum_{x,\xi} \Gamma(x, \xi + sx) \langle U_{-s} f, M_\xi T_x U_{-s} g \rangle M_\xi T_x U_{-s} g,
\end{aligned}
$$

where

$$\Gamma(x', \xi') = |\langle f, M_{\xi'} T_{x'} g \rangle|^{\frac{p}{2}-1} = |\langle U_{-s} f, M_\xi T_x U_{-s} g \rangle|^{\frac{p}{2}-1}.$$

The phase factor disappears in the expression of the operator. As a consequence, $G_{g,\Lambda',\Gamma}(f)$ may be calculated with a standard lattice Λ but with a chirped version of f and g.

As mentioned above, for the finite discrete setting $s \in \mathbb{N}$ is necessary to make sure that the chirp does an integer number of revolutions around the TF plane. The restriction comes from the periodic structure of the discrete signal domain. However, for concentrated and TF centered signal and window function the boundary terms of the TF representation become negligible, as we will show in the following. Hence we can act as if we were without any periodicity requirements and choose whatever $s \in \mathbb{R}$ is needed.

Let f and g be well localized signals around the origin ($t = N/2$). In the following we will show that the energy of the TF representation $\mathcal{V}_g f$ is negligible around the boundaries. Assume that for all $t \notin [N/2 - N/8, N/2 + N/8]$, $|f(t)| \leq \varepsilon$ and $|g(t)| \leq \varepsilon$, where ε is small. Translating the window by x gives $T_x g$ centered around $t = N/2 + x$ and as a consequence the overlap between f and $T_x g$ is decreased. Then for $x \notin [-N/4, N/4]$ and for all ξ,

$$
\begin{aligned}
|\langle f, M_\xi T_x g \rangle| = |\sum_{t=0}^{3N/8} \varepsilon M_\xi T_x g(t) &+ \sum_{t=3N/8}^{5N/8} f(t)\varepsilon + \sum_{t=5N/8}^{N-1} \varepsilon M_\xi T_x g(t)| \\
&\leq \varepsilon \left(\|f\|_1 + \|M_\xi T_x g\|_1 \right) = \varepsilon \left(\|f\|_1 + \|g\|_1 \right) \\
&\leq \varepsilon \sqrt{N} \left(\|f\|_2 + \|g\|_2 \right).
\end{aligned}
\tag{26}
$$

The same results hold along the frequency axis, it can be shown by replacing f and g by their Fourier transform. As a consequence, away from the center of the TF plane the energy of the signal representation may be neglected as well as its influence on the optimal window calculation.

---

[1] This operator is a Gabor multiplier if the frame is tight.

Concerning dilations σ, the same property of concentration around the origin allows us to dilate as in the continuous case, without taking care of what could happen at the boundaries. The value of σ is directly given as the output of the window optimization and the lattice parameters (a, b) are replaced by $(\sqrt{\sigma}a, b/\sqrt{\sigma})$ plus an extra term coming from the redundancy (see Eq. (22)).

Remark 7. *If the window is given by the unconstrained optimization problems (7) or (11), the lattice parameters may be fixed by the following process. Using formula (9), a set of dilated and chirped (periodized or truncated) Gaussian is constructed. The scalar product between the optimal window and the elements of the set is computed. The maximal value obtained for the scalar product gives the dilated and chirped Gaussian closest to the optimal window. The parameters of this function (dilation σ, chirping s) are then used to construct the adapted lattice. Of course, the shape of the optimal window may be far from an ellipse in the TF plane, especially when using (7) and one has to be aware of this limitation.*

5   Alternating optimal window and adapted lattice

We point out that the computation of the optimal window is done with a fixed given lattice. The output of the previous step (section 4) gives a lattice adapted to the optimal window. But this window is optimal (in terms of concentration of the TF representation of f) for a different lattice. It is interesting to see now what is the optimal window for the new lattice. Ultimately, we want to carry out these steps iteratively, as described below.

1. Choose an initial Gabor lattice,

2. Find the optimal window g,

3. Compute the lattice adapted to g,

4. Set this lattice as the Gabor lattice and loop to 2.

Connected to the above procedure there are some rather delicate convergence questions. Even though we will not provide theoretical justifications for doing this iterative loop of window and lattice optimization, we are able to back it up with numerical experiments. Some of the experiments carried out numerically are described and visualized in Section 6.

6   Applications

In this section we illustrate the method(s) with a few numerical examples. We focus on the main properties and results of the optimization procedures. A comprehensive description of their outputs for applications is not within the scope of the present study. A detailed analysis of their applications for audio signal analysis is planned in the near future.

All the numerical applications have been computed using MATLAB and the LTFAT toolbox [15]. This toolbox contains functions for time-frequency analysis and synthesis. In particular we have used the discrete Gabor transform and functions designed for the construction of Gabor multipliers. It also contains a Gabor transform for non-separable (sheared) lattices as well as a function for displaying the spectrogram for these non conventional lattices.

6.1   Some examples of optimal windows for the standard lattice

Non-parametrized vs. parametrized. We start by comparing the optimal windows obtained by minimizing the non-parametric case (7) and the parametric case (10). The signal is a real valued quadratic chirp i.e. a bent pattern in the positive frequencies part of the TF plane and its symmetric counterpart in the negative ones. As a region of interest we pick the positive frequencies in a time window between 0.4 and 0.7 seconds. The results are plotted in Figure 2. The two optimal windows are significantly different. In the non-parametric case the window is allowed to have non-local or disconnected components. This freedom permits a better optimal window in terms of energy concentration in the TF plane. However, this window is not appropriate in terms of interpretation of the spectrogram as it introduces artificial components: a sort of shadow chirp component appears (see the spectrograms of the full signal on Fig. 2). The parametrized Gaussian is local and does not allow for these kind of "artefacts". Even if the optimization procedure giving this latter window is restricted to a small subspace, the improvement on the signal representation is noticeable. It is even comparable to the one computed with the non-parametric optimal
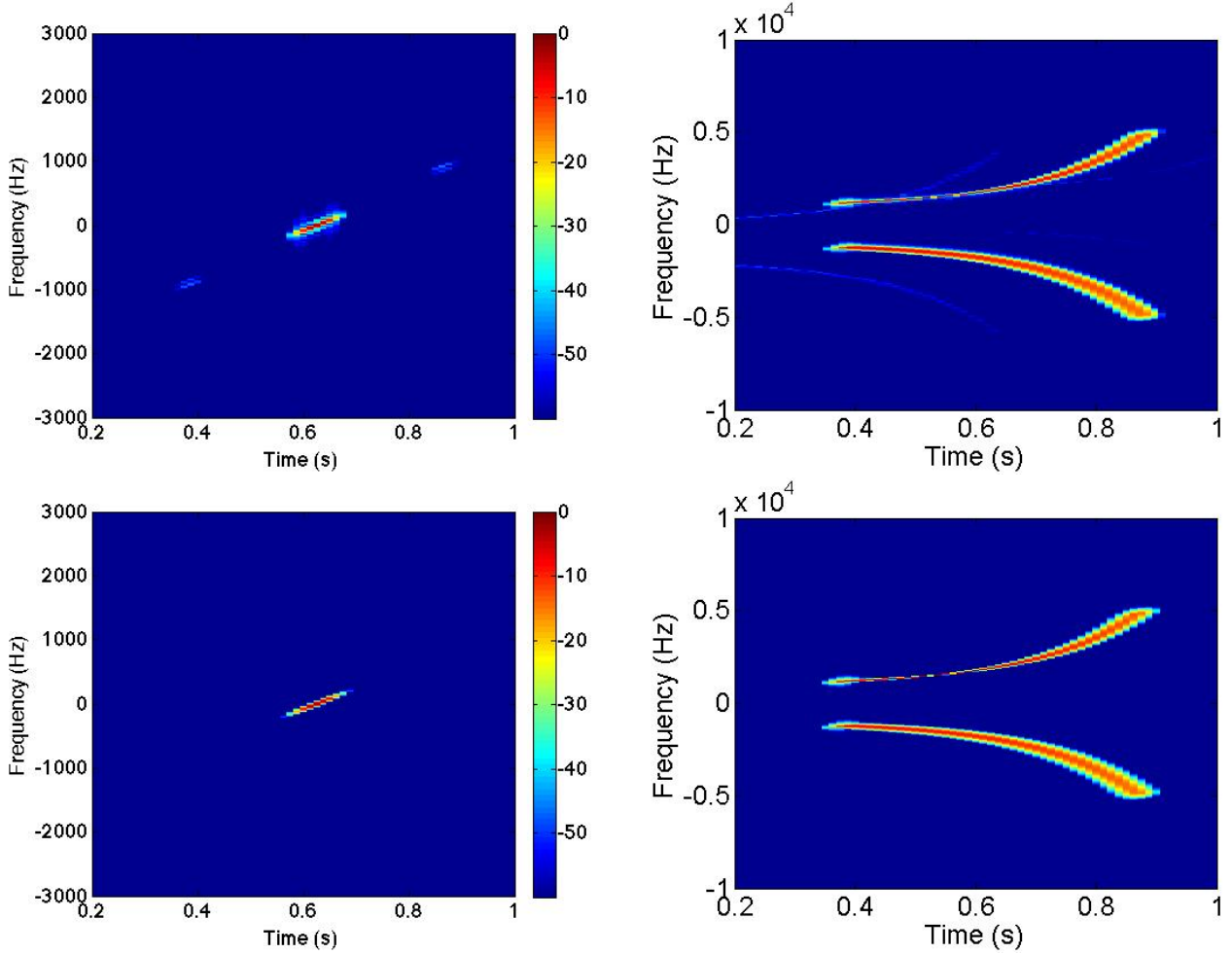
Figure 2: Left: Modulus of the optimal windows ambiguity function (top: non-parametric case, bottom: parametric case), log scale. Right: Spectrograms of a synthetic sound composed of a quadratic chirp using the two respective optimal windows on the left. The optimization has been done with respect to the part of the signal between time 0.4s and 0.7s and frequencies 0 to 5000 Hertz.

window. The region chosen for the optimization is a thin line in the TF plane in both cases. Let us notice that the windows are optimal for a selected part of signal, which is the reason for the badly resolved chirps outside of the user selected region of interest.

Let us remark that several tests were performed using large values of p in the non-parametric case. As expected, the optimal window tends to match the extracted signal when p increases (see remark 3).

Accurate representation of a localized part. The next test reveals the main asset of the method for time-frequency analysis. A better concentration in the TF plane implies a better separation of close components. This is illustrated on Fig. 3. On the upper left image the spectrogram of a function f has been computed with a standard (Gaussian) window. One can see that the time-frequency oscillating pattern is made of several nearby components, which seem to merge repeatedly over time. In fact the signal is composed of a sum of three very close frequency components with constant amplitude. These components have instantaneous frequencies f_i, with $i = 1, 2, 3$, of the type:

$$f_i = \cos(b \cos ct + a_i t + \theta), \tag{27}$$

for some constants a_i, b, c, θ. The modulation is the same for all components so that they stay equally separated over time. The Gaussian window is not adapted for discriminating the three components around regions where the frequency is highly varying as for example around time 1.8 seconds. As a consequence, the components appear as one thick line in the TF representation. Our window optimization procedure allows to find an appropriate window for this part of the spectrogram. We have extracted the part of the signal between 1.6 and 2s and have run the optimal window procedure. The result is shown on the upper right figure. This time the 3 components are separated
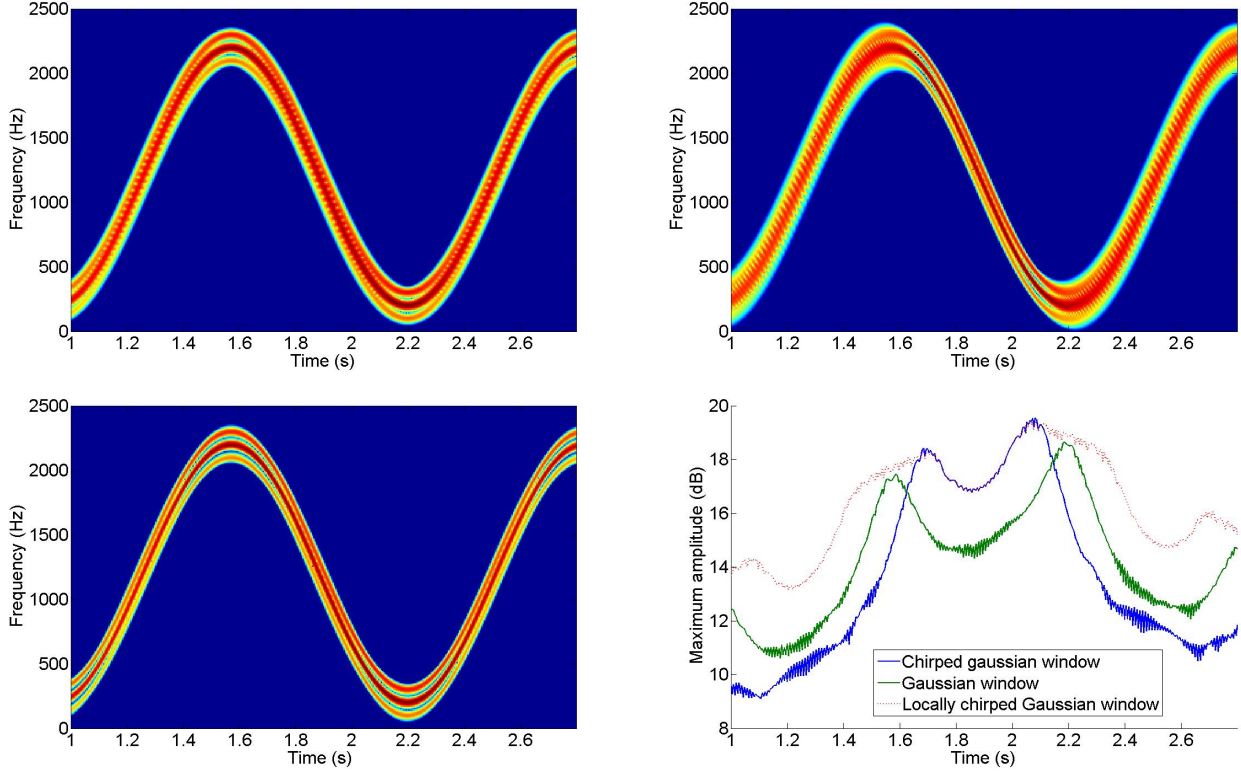
11

Figure 3: Spectrogram of an oscillating chirp in the TF plane with three different windows. Top left: Gaussian window. Top right: Optimal window for time interval $[1.6, 2]$. Bottom left: Optimal window evolving with time. Bottom right: Evolution in time of the maximum over frequencies of the spectrogram coefficients, for the 3 windows.

around the region $t = 1.8s$. However, the optimal window is of course optimal for only a part of the signal. One can see that this window is unable to separate the components well in the area where the Gaussian window succeeded. The bottom left figure could be the ideal representation for this type of signal. The signal has been cut in different successive pieces for which different optimal windows have been computed. For each part the Gabor transform has been computed with its optimal window. The total spectrogram is the sum of the spectrograms of the different parts. The bottom right figure shows the evolution of the energy maximum over time:

$$ m(t) = \max_{\xi} \mathcal{V}_g f(\xi, t), $$

for the different windows g. Since the windows are normalized, a higher maximum means a sparser representation.

6.2   Combining optimal window and optimal lattice

Given the optimal window, the computation of the adapted lattice gives the most appropriate time and frequency steps. These are two key parameters for an accurate representation. Fig 4 illustrates this point. Even if the window is optimal, an inappropriate lattice will destroy the accuracy of the representation.

The whole optimization process containing the alternative window and lattice optimization (Sec. 5) has been tested in different cases. The preliminary results show that the algorithm converges in many cases. However it appears that the redundancy parameter has an influence on the convergence. Choosing $R = 15$ makes the algorithm converge rapidly in 4-5 iterations. For lower redundancies, the risk of alternating infinitely between several optimal windows and lattices increases.

Tests show that there are often only minor improvements of the representation sparsity when comparing representations using the optimal window/lattice at convergence and the optimal window/lattice given after the first iteration, see Fig. 5. On this figure, the spectrograms using these two optimal windows are almost identical. One
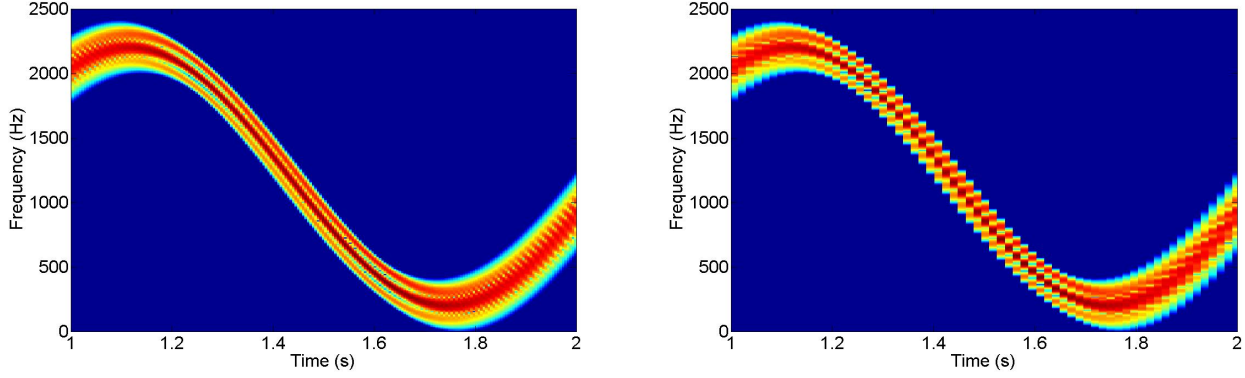
Figure 4: Spectrograms of a signal with the same optimal window but for two different lattices having the same redundancy R. Left: Optimal time and frequency steps. Right: inappropriate lattice.
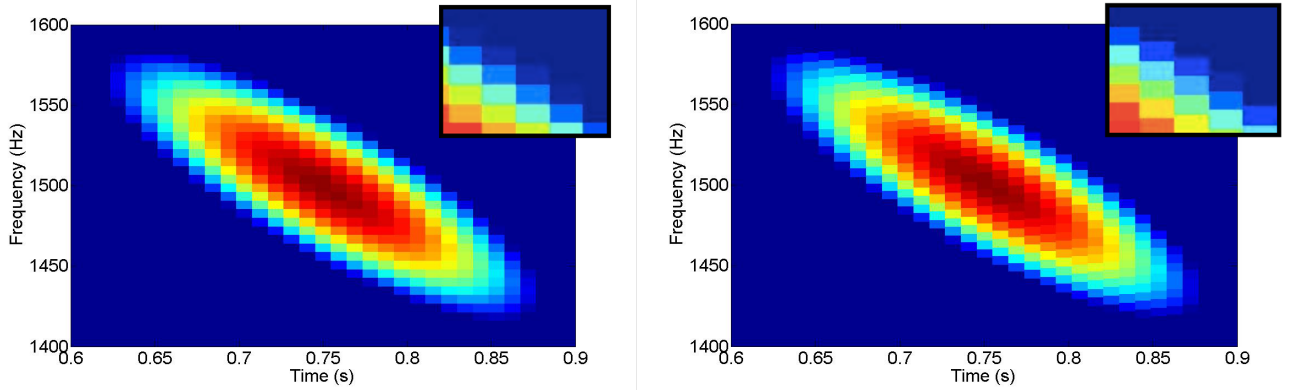

Figure 5: Spectrograms of a chirp signal with two different windows and lattices. Left: with the optimal window associated to the optimal lattice obtained after the first iteration of the algorithm of Sec. 5. Right: with the optimal window and lattice after convergence of the algorithm of Sec. 5. In the top right corner of each picture a box magnifies the details of the TF tiling. Each pixel (rectangle) corresponds to the magnitude of one Gabor coefficient. The pixels are not aligned in the same manner (the shear values are different).

difference is in the position of the Gabor coefficients (see the zoom boxes); they are shifted along the frequency axis by a value varying at each time step, according to the shear parameter. In this case, the sparsity of the representation is not improved by iterating the optimal window/lattice procedure. However, the gain of using the optimal lattice resides in the denser packing in the TF plane of the analysis window. This is an important point for the synthesis aspect. By giving an optimal window and lattice, the optimization process of Sec. 5 may further improve the packing, this is to be tested in future work. Eventually, one must also notice that the TF pattern on which the optimization is made is rather simple: a short-time chirp signal. The results might be different for other types of patterns (to be investigated).

7   Conclusion and perspectives

The work presented here shows promising results and several directions for further research. The 2-parameter optimization scheme, although restricted, proved to be the best choice for the analysis of real world signals. We noticed that it always converged to an appropriate optimal window for the patterns we were interested in. We must keep in mind that this might not always be the case; it could reach a bad local optimum for unusual and highly complex patterns in the TF plane. Another advantage over non-constraint problem is that the optimization of only two parameters allows a fast computation.

The optimal lattice procedure is found to be highly valuable as it adds extra power to the optimization both for the visualization (adapted lattice steps) and the synthesis (improving the condition number). It appears to be

indispensable. In addition we provide a way for the fast computation of this adapted lattice (even though it is an approximation of it). This is an asset for practical applications.

It is particularly interesting to notice the relation between dilation and chirping of a window and dilation and shearing of a lattice. As seen in Sec. 4, all possible Gabor lattices can be described by 3 parameters: dilation, shearing and redundancy. One can associate to any of these lattices a window parametrized by a chirp and a dilation which will be as optimally packed as would be round Gaussians with an hexagonal lattice. This points out the fact that the set of dilated and chirped version of any well-concentrated window is the natural and most general set of windows for the framework of the standard Gabor transform.

We demonstrated that the alternative lattice/window optimization can be done. This is only a preliminary work which serves as a proof of concept. We have seen that this principle is interesting for the automatic analysis of engineering signals. Future work on this topic must include a study of the convergence of the algorithm as well as an analysis of its relevance beyond a few iterations.

From a practical point of view, we found highly interesting results using a window whose shape evolves in time (c.f. Fig. 3). As a motivating example, it allows to follow and separate close frequency patterns over time. The results shown on the figures were obtained by selecting manually successive time intervals and optimizing a window on each. This process deserves to be made automatic. However, a non-stationary window breaks the underlying structure of the Gabor transform. The setting of non-stationary Gabor transforms [2] is likely to be an appropriate setting to closer investigate windows evolving in time. This gives a promising direction for future research.

References

[1] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized gauss-seidel methods. *math. Program.,*, Ser. A(137):91–129, 2013.

[2] P. Balazs, M. Dörfler, F. Jaillet, N. Holighaus, and G. A. Velasco. Theory, implementation and applications of nonstationary Gabor frames. *J. Comput. Appl. Math.*, 236(6):1481–1496, 2011.

[3] O. Christensen. *Frames and Bases. An Introductory Course.* Applied and Numerical Harmonic Analysis. Basel Birkhäuser, 2008.

[4] M. A. de Gosson and D. Onchis. Multivariate symplectic Gabor frames with Gaussian windows. *preprint*, 2012.

[5] H. G. Feichtinger, D. Onchis-Moaca, B. Ricaud, B. Torrésani, and C. Wiesmeyr. A method for optimizing the ambiguity function concentration. *Proceedings of Eusipco 2012*, 2012.

[6] H. G. Feichtinger and T. Strohmer. *Gabor Analysis and Algorithms. Theory and Applications.* Birkhäuser, Boston, 1998.

[7] D. Goldfarb. A Family of Variable-Metric Methods Derived by Variational Means. *Mathematics of Computation*, 24(109):23–26, Jan. 1970.

[8] J. Gorski, F. Pfeuffer, and K. Klamroth. Biconvex sets and optimization with biconvex functions: a survey and extensions. *Mathematical Methods of Operations Research*, 66(3):373–407, 2007.

[9] K. Gröchenig. *Foundations of Time-Frequency Analysis.* Appl. Numer. Harmon. Anal. Birkhäuser Boston, Boston, MA, 2001.

[10] M. Hampejs, N. Holighaus, L. Tóth, and C. Wiesmeyr. On the subgroups of the group $Z_m \times Z_n$. *ArXiv e-prints*, nov 2012.

[11] F. Jaillet and B. Torrésani. Time-frequency jigsaw puzzle: adaptive multiwindow and multilayered gabor expansions. *International Journal of Wavelets Multiresolution and Information Processing*, 5(2):293 – 315, 2007.

[12] S. Mallat. *A wavelet tour of signal processing: The sparse way.* Academic Press, Third edition, 2009.

[13] B. Ricaud and B. Torrésani. A survey of uncertainty principles and some signal processing applications. *Submitted, available at http://arxiv.org/abs/1211.5914*, 2013.

[14] P. Sondergaard. Efficient algorithms for the discrete Gabor transform with a long FIR window. *J. Fourier Anal. Appl.*, 18(3):456–470, 2012.

[15] P. L. Søndergaard, B. Torrésani, and P. Balazs. The Linear Time Frequency Analysis Toolbox. *International Journal of Wavelets, Multiresolution Analysis and Information Processing*, 10(4), 2012.

[16] L. Sorber, M. van Barel, and L. de Lathauwer. Unconstrained optimization of real functions in complex variables. *SIAM J. Optim.*, 22(3):879–898, 2012.

[17] T. Strohmer and S. Beaver. Optimal OFDM system design for time-frequency dispersive channels. *IEEE Trans. Comm.*, 51(7):1111–1122, July 2003.

[18] C. Wiesmeyr, N. Holighaus, and P. Sondergaard. Efficient algorithms for discrete Gabor transforms on a nonseparable lattice. *IEEE Trans. Signal Process.*, 61(20):5131 – 5142, 2013.