A METHOD FOR OPTIMIZING THE AMBIGUITY FUNCTION CONCENTRATION

H. G. Feichtinger[1], D. Onchis-Moaca[1], B. Ricaud[2], B. Torrésani[2], C. Wiesmeyr[1]

[1]NuHAG - University of Vienna
Nordbergstraße 15
1090 Wien, Austria.

[2]LATP, Aix-Marseille Univ./CNRS
UMR7353, 39 rue F. Joliot-Curie,
13453 Marseille cedex 13, France.

ABSTRACT

In the context of signal analysis and transformation in the time-frequency (TF) domain, controlling the shape of a waveform in this domain is an important issue. Depending on the application, a notion of optimal function may be defined through the properties of the ambiguity function. We present an iterative method for providing such optimal functions under a general concentration constraint of the the ambiguity function. At each iteration, it follows a variational approach which maximizes the ambiguity localization via a user-defined weight function F. Under certain assumptions on this latter function, it converges to a waveform which is optimal according to the localization criterion defined by F.

1. INTRODUCTION

In the context of time-frequency analysis (such as short time Fourier analysis or Gabor analysis), it is commonly agreed that best results are obtained when one uses window functions that are well localized in the time-frequency domain. In the mathematically idealized setting of continuous time, infinite support signals, i.e. functions on the whole real line $\mathbb{R}$, the optimally localized waveforms in the TF domain are well known to be Gaussians [1]. The latter are the minimizers of the variance uncertainty inequality (Heisenberg uncertainty principle) as well as the Hirschman entropic inequality and the Lieb time-frequency inequalities [2]. However, in other contexts, of importance for particular applications, the optimal function is not known; the localization criterion does not rely on such measures and the definition of optimality can be different. For example, in the discrete, finite time setting, the variance is not defined and the entropy point of view gives a "picket fence" signal as a minimizer. The latter is indeed optimally concentrated, has smaller entropy than the periodized discrete Gaussian for example, but is definitely not well localized. In addition, according to the components/information one wants to retrieve

or visualize inside a signal, there exist more appropriate families of windows (Hann, Hamming, prolate spheroidal, etc...), optimal in a different sense.

In the framework of the short time Fourier transform (STFT), the ambiguity function (which also characterizes the reproducing kernel of the STFT), provides a sensible measure of the spreading of the analysis window in the time-frequency domain. This function plays an important role in the estimation of optimal localization properties in many different areas e.g. for operator approximation by Gabor multipliers [3] or for RADAR and coding applications [4, 5, 6]. Depending on the application, the search for an optimal function emphasizes some particular property, e.g. support properties, peakiness, fast decay away from some reference point. Hence the ambiguity function must satisfy a set of constraints being stated by the user as well as its own inherent constraints.

We present here a method for designing optimally localized windows in the TF plane, based on the maximization of some measure of the ambiguity function concentration. Under general hypotheses and shape constraints, the proposed algorithm is able to provide an optimal function as the solution of a variational problem.

2. TIME-FREQUENCY CONCENTRATION: CONTINUOUS TIME

2.1. The ambiguity function

We first consider the case of continuous time, infinitely supported signals, and denote by $\langle \cdot, \cdot \rangle$ the L^2 inner product. Let $\psi \in L^2(\mathbb{R})$ be a finite energy waveform, and denote by ψ_z, with $z = (\tau, \xi)$ a time-frequency shifted copy of ψ with time and frequency shifts τ and ξ:

$$\psi_z(t) = e^{2i\pi\xi t}\psi(t - \tau).$$

The ψ_z will be termed Gabor atoms. Let us introduce the STFT $\mathcal{V}_\psi f$ of $f \in L^2(\mathbb{R})$, and the ambiguity function $\mathcal{A}_\psi$ of ψ:

$$\mathcal{V}_\psi f(z) = \langle f, \psi_z \rangle, \qquad \mathcal{A}_\psi(z) = \mathcal{V}_\psi \psi(z). \quad (1)$$

The ambiguity function for an arbitrary $\psi \in L^2(\mathbb{R})$ is bounded, continuous, decays at infinity, satisfies

$\|\mathcal{A}_\psi\|_2 = \|\psi\|_2^2$ and is the kernel of a positive semi-definite operator. Furthermore its modulus satisfies the symmetry relation $|\mathcal{A}(z)| = |\mathcal{A}(-z)|$, attains its maximum value at the origin and satisfies specific concentration constraint [7]. This shows that not every function defined on the TF plane is the ambiguity function of a window. Therefore the design of windows such that their TF picture satisfies arbitrary properties is in general impossible. Still, one can try to find waveforms whose ambiguity function matches as closely as possible such prescribed localization properties (see e.g. [8], [9]). The current paper considers a new localization criterion in the ambiguity domain, and proposes corresponding new optimization algorithms.

2.2. Concentration measures

Classical families of concentration measures are provided by L^p-(quasi)-norms, for $p \geq 0$

$$I_\psi(p) = \|\mathcal{A}_\psi\|_p^p = \int_{\mathbb{R}^2} |\langle \psi, \psi_z \rangle|^p dz, \qquad (2)$$

which are closely related to Rényi entropies (for details see e.g. [10]), and the Shannon entropy measure:

$$S(\mathcal{A}_\psi) = -\int_{\mathbb{R}^2} |\mathcal{A}_\psi(z)|^2 \log(|\mathcal{A}_\psi(z)|^2)\, dz = I'_\psi(2)\,. \tag{3}$$

For $p < 2$, maximizing the L^p-norm (under unit L^2 norm constraint) favors spreading while for $p > 2$ this increases sparsity. Small values for the Shannon entropy is a sign of sparsity. In the continuous case, Lieb [2] gives the following bounds on the L^p-norm of the ambiguity function:

$$\|\mathcal{A}_\psi\|_p^p \leq (2/p)\|\psi\|_2^{2p}, \quad p > 2 \tag{4}$$

with a reversed inequality for $p < 2$. These bounds are sharp and all yield Gaussian functions (up to translations, dilations and rotations in the TF plane) as minimizers. Lieb also provides a lower bound for the entropy, with again the same family of Gaussians as minimizers

$$S(\mathcal{A}_\psi) \geq 1\,, \tag{5}$$

These cases are examples of a family of problems for which our method provides optimizers.

2.3. Optimization principle

We describe a general scheme for optimizing prescribed time-frequency localization properties. The rationale is to set the problem as a variational problem, and optimize a weighted L^2-norm of the ambiguity function. The problem to solve reads

$$\psi_{\text{opt}} = \arg\max_{\psi:\|\psi\|=1} \int F(|\mathcal{A}_\psi(z)|, z)\, |\mathcal{A}_\psi(z)|^2\, dz\,, \tag{6}$$

where the *density function* $F : \mathbb{R}^+ \times \mathbb{R}^2 \to \mathbb{R}^+$ is chosen so as to enforce some specific localization properties. We shall consider here the two following specific cases, but the approach is not limited to these.

1. Find a ψ which optimizes the sparsity of $\mathcal{A}_\psi$, we choose F of the form

$$F(|\mathcal{A}_\psi(z)|, z) = |\mathcal{A}_\psi(z)|^{p-2} \quad (p > 2)$$

and (6) becomes equivalent to finding the function which maximize $\|\mathcal{A}_\psi\|_p^p$. This will lead to the Gaussian here but our method clearly extends to more general settings than $L^2(\mathbb{R})$. We shall see in the next section that this also leads to discrete versions of the Gaussian function.

2. The other extreme case is to choose weight function of the form

$$F(u, z) = F(0, z)\,,$$

i.e. a fixed weight in the ambiguity plane. In this case, the optimization (6) is expected to produce waveforms ψ whose ambiguity function is concentrated in TF regions where F takes large values. Given the constraints on ambiguity functions, we limit ourselves to weight functions concentrated around the origin and satisfying the symmetry property $F(-z) = F(z)$. When F is an indicator of a ball the solution of the problem is the Gaussian function [11].

Optimizing Eq. (6) does not lead to a closed-form solution. Alternatively, we shall search for approximate solutions, by recursive quadratizations.

Suppose a guess ϕ is available (fixed function in $L^2(\mathbb{R})$), then we can maximize the following quantity (cost function) with respect to ψ:

$$\Gamma_\lambda[\psi|\phi] = \int F(|\mathcal{A}_\phi(z)|, z)|\mathcal{V}_\phi\psi(z)|^2 dz - \lambda(\|\psi\|^2 - 1) \tag{7}$$

where the constraint of unit norm was added via a Lagrange multiplier λ. This is equivalent to constraining the STFT of ψ (with window ϕ) to be localized in regions where $F(|\mathcal{A}_\phi(\cdot)|, \cdot)$ is concentrated. Let us denote by ψ_1 the function maximizing (7) and $\Gamma_\lambda^{(1)}$ the maximum. The next step is to replace ϕ by ψ_1 and solve the variational problem again. This yields a new optimal function ψ_2, with maximum $\Gamma_\lambda^{(2)}$. Iterating this procedure leads to a (possibly local) maximum for Γ_λ. For special weight functions we have the following result:

Proposition 1. *For any function F bounded, positive and even with respect to z, maximizing* (7) *is equivalent to solving the following eigenvalue problem*

$$\int_{\mathbb{R}^2} F(|\mathcal{A}_\phi(z)|, z)\mathcal{V}_\phi\psi(z)\phi_z\, dz = \lambda\psi\,, \tag{8}$$

where λ is the largest eigenvalue. Furthermore, the sequence $\Gamma_\lambda^{(n)}$ converges to a (local) maximum for weight functions F independent of ϕ.

Sketch of proof: Using the classical method for the calculus of variations, i.e. equating the functional derivative of Γ_λ with respect to ψ to zero yields the linear equation (8). The cost function Γ_λ converges for F real, positive, even, bounded and independent of ϕ since Γ_λ is bounded, and at rank n

$$\Gamma_\lambda[\psi_n|\psi_{n-1}] = \Gamma_\lambda[\psi_{n-1}|\psi_n] \le \Gamma_\lambda[\psi_{n+1}|\psi_n]. \square$$

Remark 1. a). *The convergence of $\Gamma_\lambda^{(n)}$ does not necessarily imply the convergence of the sequence $\psi^{(n)}$. Conditions on F for the convergence of $\psi^{(n)}$ to ψ_{opt} can be derived; this is the object of current research and will be developed elsewhere.* **b).** *When it exists, the limiting function ψ_{opt} satisfies Eq. (8), with $\phi = \psi_{\text{opt}}$ and hence maximizes Γ_λ with maximal value $\Gamma_\lambda[\psi_{\text{opt}}|\psi_{\text{opt}}]$. This turns out to match the expression of Eq. (6), and hence it is at least a local solution of the initial optimization problem.*

Remark 2. a). *When ϕ is the standard Gaussian and $F(|\mathcal{A}_\phi(\tau,\xi)|, (\tau,\xi)) = F(0, \tau^2 + \xi^2)$ is radially symmetric, Daubechies showed in [12] that the solution of (8) is ϕ. Hence it is a function which gives a (local) maximum for (6). Moreover, she showed that the eigenvectors of the STFT-multiplier are the Hermite functions.* **b).** *In our iterative process, the diagonalization of the operator obtained in the last iteration yields an orthonormal basis which can be seen as a set of generalized Hermite functions.*

3. ALGORITHM IN DISCRETE TIME

3.1. Discrete time considerations

In terms of localization, the finite discrete case is not just a discretization of the continuous one. Replacing integrals by finite sums yields a completely different world. For example, for signals of length N analogues of the Lieb inequalities can be obtained.

Proposition 2. *Assume $\psi \in \mathbb{C}^N$ is such that $\|\psi\|_2 = 1$. Then, assuming $p < 2$,*

$$\|\mathcal{A}_\psi\|_p \ge N^{\frac{1}{p} - \frac{1}{2}}, \qquad (9)$$
$$S(\mathcal{A}_\psi) \ge \log(N). \qquad (10)$$

The family of "picket fence" signals, translated and modulated copies of the following periodic series of Kronecker deltas:

$$\omega(t) = \frac{1}{\sqrt{b}} \sum_{n=1}^{b} \delta(t - an), \quad ab = N,$$

saturates the inequality.

Sketch of the proof: The proof is directly adapted from the proof of Lieb's inequality for the continuous case [2]. The ambiguity function is expressed as a finite Fourier transform, its L^p-norm can be bounded using the Hausdorff-Young inequality. The resulting bound takes the form of a finite periodic convolution product which is itself bounded using Young's convolution inequality. Finally, the entropic inequality is obtained by a standard limiting argument. $\square$

Note that the minimizer is not unique. Picket fence signals turn out to optimize many finite domain uncertainty principles (see e.g. [1]). This is quite different to the continuous case described in Sec. 2.2. In particular, picket fences are concentrated but not localized: they are not the most relevant windows for applications.

In this discrete setting, it is interesting to find an equivalent of the Gaussian function. One may define it as the optimal function of the following discretized version of (6):

$$\psi_{\text{opt}} = \arg\max_{\psi: \|\psi\|=1} \sum_z F(|\mathcal{A}_\psi(z)|, z) |\mathcal{V}_\psi \psi(z)|^2, \quad (11)$$

where $F(|\mathcal{A}_\psi(z)|, z) = |\mathcal{V}_\psi \psi(z)|^{p-2}$ for any $p > 2$. This is made possible by the proposed algorithm.

3.2. Algorithm

As in the continuous case recursive quadratization, as described in Eq. (7) and below can be used to make the problem treatable. It leads to diagonalizing the STFT multiplier matrix: $\sum_z F(|\mathcal{A}_\phi(z)|, z)\langle \phi_z, \cdot\rangle \phi_z$. All the given arguments for convergence apply in the discrete case also. Given any density function F, an initial window ϕ has to be specified to start with at the beginning. The steps are summarised in Algorithm 1. The stopping criterion is based on the L^2-norm difference between functions issued from two successive iterations.

Algorithm 1 compute the optimal window ψ_{opt}

   **Input:** ϕ - initial window, F - weight, eps.
   Individual Steps:
1:   $\psi_0 = 0, \psi_1 = \phi$
2:   **while** $\|\psi_1 - \psi_0\| > eps$ **do**
3:      $\psi_0 \leftarrow \psi_1$
4:      Compute the STFT-multiplier with weight F and window ψ_0
5:      Compute ψ_1 as the eigenvector to the largest eigenvalue
6:      Update F if necessary
7:   **end while**
8:   $\psi_{\text{opt}} = \psi_1$

3.3. Results and convergence issues

Fixed weight ($F(0, z)$): For a radial symmetric weight function F as stated in Remark 2.a), the iterative proce-

dure presented in Algorithm 1 converges to a Gaussian-like function. In Fig. 1 a plot of F (top right) and of the convergence rate (top left) is presented. Notice that it linearly decreases in log scale. The optimal function is plotted (middle) together with an example of another eigenvector of the operator given at the last iteration (bottom). We may define the discrete Hermite functions to be the eigenfunctions of such an operator, as pointed out in Remark 2.
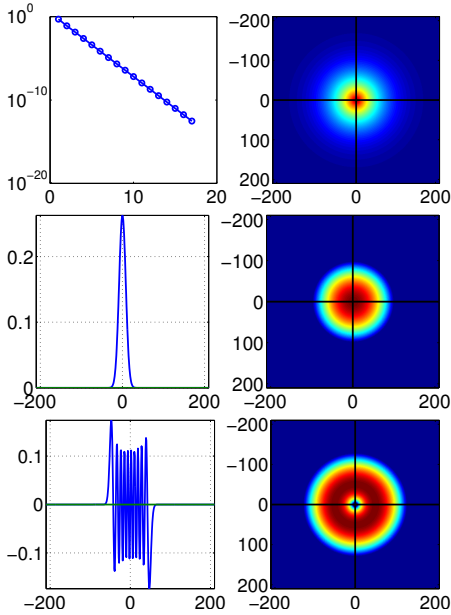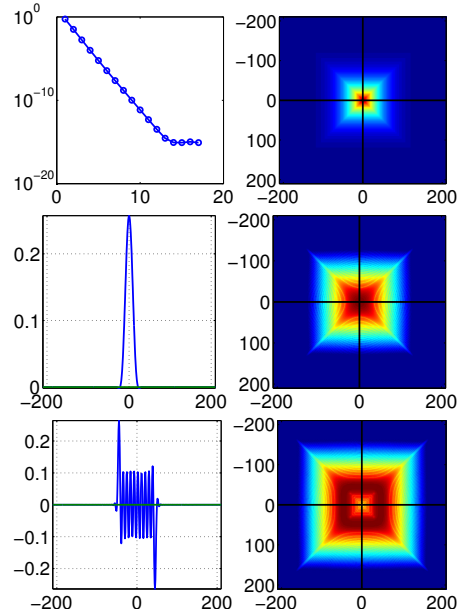

Fig. 1. Results in the case of a radial TF mask: $F(|\mathcal{A}_\psi(z)|, z) = F(0, z)$. Top left: $\|\psi_{n+1} - \psi_n\|_2$, log scale. Top right: TF mask. Middle: optimal function in time (left) and in the TF plane through its ambiguity function (right). Bottom: an eigenvector of the last STFT-multiplier, the "generalized" Hermite function H_{20} in time (left) and in the TF plane through its ambiguity function (right). The signal length is $N = 420$.

Weight functions which are constant along squares centered at the origin are another interesting example, see Fig. 2. Algorithm 1 leads to a square-shaped optimal function. Notice the generalized Hermite function which forms a square ring in the TF plane.

As stated in Remark 1.a), the sequence of $\psi^{(n)}$ may not converge although Γ does; this is the case when F has the shape of a cross in the TF plane for example. The algorithm is "hesitating" between convergence to one of the cross branch or the other.

Weight depending on $|\mathcal{A}_\psi|$ $(F(|\mathcal{A}_\psi(z)|))$**:** The main difference to the previous case is that the weight function evolves at each iteration and that the result depends on the initial window. The maximum of the objective function Γ_λ is very degenerate (there are at least $2N$ solutions: N Kronecker deltas and N sine waves) and


Fig. 2. Results in the case of a square-shaped TF mask: $F(|\mathcal{A}_\psi(z)|, z) = F(0, z)$. Top left: $\|\psi_{n+1} - \psi_n\|_2$, log scale. Top right: TF mask. Middle: optimal function in time (left) and in the TF plane through its ambiguity function (right). Bottom: an eigenvector of the last STFT-multiplier, the "generalized" Hermite function H_{20} in time (left) and in the TF plane through its ambiguity function (right). The signal length is $N = 420$.

there are many local minima. The algorithm seems to converge to the maximum closest to the initialization. For example, starting with a widespread signal (e.g. white noise) generally leads to a picket fence, or more simply a constant function (a Kronecker delta in the Fourier domain). The solution for a mildly localized starting window is a Gaussian. In Fig. 3 we display two examples corresponding to these two situations. The value of both cost functions at each iteration has been plotted on the top graph: it shows that the Gaussian-like solution gives only a local maximum for Γ_λ. For these two examples, the respective relative entropy $S = S(\mathcal{A}_{\psi_{\text{opt}}})/\log(N)$ yields: $S = 1$ for the picket fence and $S = 1.16$ for the Gaussian like function, showing again that the Gaussian is only a local maximum for (11), while the picket fence signals are global minimizers. Last remark: the convergence speed increases with the value of p, up to a certain level; beyond that the algorithm starts to behave unstable.

4. CONCLUSION

We have proposed a novel approach for designing waveforms with controlled localization in the ambiguity plane. Our algorithm is able to provide optimally concentrated functions for a wide range of
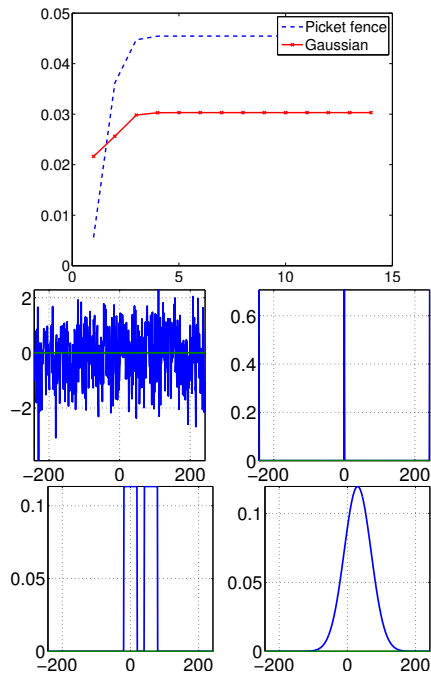
Fig. 3. Results for the optimization of the L^p norm of the ambiguity function ($p = 3$). Top: Cost function leading to the picket fence (dashed) and to the Gaussian-like function (crossed). Middle: initial delocalized function (left) and final function, i.e. two-picket fence (right). Bottom: initial localized function and final one, the Gaussian-like function.

constraints/applications via a flexible weight function. It covers the intuitive spreading measure associated to L^p-norms as well as the design of concentration masks in the TF plane. We recover the optimal functions given in the literature for particular choices of weight functions. Furthermore, by choosing a weight function F, one can define an application-tailored measure of localization and obtain the optimal window accordingly. For example a square-shaped window in the TF plane may be an interesting window for tiling the TF domain, when using a square lattice for the Gabor transform. It should be less redundant than the round-shaped Gaussian atoms.

The investigation of conditions of F which guarantee convergence is the object of ongoing work, together with the analysis of more general expressions for F.

This procedure may also be of interest for other problems. Firstly, when diagonalizing the STFT-multiplier matrix with respect to the limiting window and the chosen weight F, one obtains a whole orthogonal basis, one of them being the optimal function. The other vectors have original shapes in the TF plane, looking like deformed rings around the origin. This could be seen as a generalization of the Hermite functions basis.

5. REFERENCES

[1] D. L. Donoho and P. B. Stark, "Uncertainty principles and signal recovery," *SIAM J. Appl. Math.*, vol. 49, no. 3, pp. 906–931, June 1989.

[2] Elliott H. Lieb, "Integral bounds for radar ambiguity functions and wigner distributions," *J. Math. Phys.*, vol. 31, pp. 594–599, 1990.

[3] M. Doerfler and B. Torrésani, "Representation of operators in the time-frequency domain and generalized Gabor multipliers," *J. Fourier Anal. and Appl.*, vol. 16, no. 2, pp. 261–293, 2010.

[4] M. A. Herman and T. Strohmer, "High-resolution radar via compressed sensing," *IEEE Trans. Sig. Proc.*, vol. 57, no. 6, pp. 2275–2284, 2009.

[5] W.O. Alltop, "Complex sequences with low periodic correlations," *IEEE Trans. Info. Th.*, vol. 26, no. 3, pp. 350–354, May 1980.

[6] J. P. Costas, "A study of a class of detection waveforms having nearly ideal range-doppler ambiguity properties," *Proc. of the IEEE*, vol. 72, pp. 996–1009, 1984.

[7] B. Demange, "Uncertainty principles for the ambiguity function," *J. London Math. Soc.*, vol. 72, no. 3, pp. 717–730, 2005.

[8] S. M. Sussman, "Least-squares synthesis of radar ambiguity functions," *IRE Transactions on Information Theory*, vol. 8, no. 3, pp. 246–254, 1962.

[9] Franz Hlawatsch and W. Krattenthaler, "Bilinear signal synthesis," *IEEE Trans. on Sig. Proc.*, vol. 40, no. 2, pp. 352–363, 1992.

[10] R. Baraniuk, P. Flandrin, A. J.E.M. Janssen, and O. Michel, "Measuring time frequency information content using the renyi entropies," *IEEE Trans. Inf. Theory*, vol. 47, no. 4, pp. 1391–1409, 2001.

[11] Patrick Flandrin, "Maximum signal energy concentration in a time-frequency domain," *ICASSP-88*, pp. 2176–2179, 1988.

[12] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach," *IEEE Trans. Info. Th.*, vol. 34, no. 4, pp. 605–612, July 1988.