Inequalities in Mellin-Fourier signal analysis

Patrick Flandrin

Ecole Normale Supérieure de Lyon
Laboratoire de Physique (Ura 1325 Cnrs)
46 allée d'Italie, 69364 Lyon Cedex 07, FRANCE
Tel. : +33 4 72 72 81 60   Fax : +33 4 72 72 80 80
E-mail : `flandrin@physique.ens-lyon.fr`

Abstract

A specific form of the Mellin transform, referred to as the "scale transform," is known to be a natural complement to the Fourier transform for wideband analytic signals. In this paper, limitations for the simultaneous localization of scale transform pairs are investigated. A number of inequalities are established and discussed, based on various measures of spread (Heisenberg-type inequalities for variance-like measures and Hirschman-type inequalities for entropy). The same issue of maximally concentrating a signal in both scale and frequency domains is also addressed via spread measures which are applied directly to joint scale-frequency distributions. A simple way of obtaining inequalities for Altes-type distributions is pointed out, new results pertaining to the unitary Bertrand distribution are established, as well as a new form of uncertainty relation for the wavelet transform.

1 Introduction

Signals are generally considered as functions of time t or frequency f, both (mutually exclusive) descriptions being connected by a Fourier transformation. More precisely, well-behaved signals $x(t), t \in I\!R$, (e.g., finite energy signals $x(t) \in L^2(I\!R)$) admit classically the Fourier representation

$$x(t) := \int_{-\infty}^{+\infty} X(f) \, e^{i2\pi ft} \, df, \tag{1}$$

where $X(f), f \in I\!R$, stands for their frequency spectrum[1].

Nevertheless, and although it is the one that is most commonly used, the Fourier transform is supplemented in some circumstances by other types of transforms, among which

---

[1]Throughout the paper, we will adopt the convention of using capital letters for spectra and the corresponding lower case symbols for their inverse Fourier transform (time-domain signals).

one can single out the Mellin transform. Given a function $F(v), v \in \mathbb{R}_+$, the most general form of its Mellin transform is [8]

$$\left(\mathcal{M}F\right)(\sigma) := \int_0^{+\infty} F(v)\, v^{\sigma-1}\, dv,$$

with σ a complex variable. Here, we will be interested in a somewhat more restricted definition that proves well-adapted to the analysis of analytic signals, i.e., signals whose frequency spectrum $X(f)$ vanishes over the real half-line of negative frequencies. Following [8], we can consider signals in $L^2(\mathbb{R}_+, f^{2r+1}\, df)$ and define the quantity

$$\underline{X}^{(r)}(s) := \int_0^{+\infty} X(f) f^{i2\pi s+r} df,$$

with $s \in \mathbb{R}$. It is easy to verify that

$$\underline{X}^{(r)}(s) = \left(\mathcal{M}X\right)\left((r+1) + i2\pi s\right),$$

thus emphasizing the fact that, in the form $\underline{X}^{(r)}$ above, we are primarily interested in looking at the Mellin transform $\mathcal{M}X$ as a function of only the imaginary part of its output variable σ, the real part being considered as fixed. Since this real part turns out to be a parameter rather than a variable, a further simplification can be achieved by fixing r to $-\frac{1}{2}$, an unrestrictive constraint that, at this point, just amounts to multiplying $X(f)$ by $f^{r+\frac{1}{2}}$ and thus dealing with signals whose spectrum is in $L^2(\mathbb{R}_+, df)$. Doing so, we end up with a restricted version of the Mellin transform that exactly coincides with what is referred to as a "scale transform" in [10].

We will therefore adopt thereafter the following

Definition 1 *The scale transform $\underline{X}(s)$ of $X(f) \in L^2(\mathbb{R}_+, df)$ is*

$$\underline{X}(s) := \int_0^{+\infty} X(f) f^{i2\pi s-\frac{1}{2}} df. \tag{2}$$

We can immediately remark that, according to Definition 1, the scale transform (2) can be equivalently expressed as

$$\underline{X}(s) = \int_{-\infty}^{+\infty} \widetilde{X}(u)\, e^{i2\pi su} du, \tag{3}$$

with

$$\widetilde{X}(u) := X(e^u) e^{u/2}. \tag{4}$$

Rewritten this way, the scale transform $\underline{X}(s), s \in \mathbb{R}$, of a given signal $X(f), f \in \mathbb{R}_+$, appears therefore as an ordinary Fourier transform applied to some suitably warped version $\widetilde{X}(u)$, $u \in \mathbb{R}$, of $X(f)$.

The Mellin transform is known to be invertible and, considering the simplified form of the scale transform given above, the spectrum of a signal can be simply recovered from it as

$$X(f) = \int_{-\infty}^{+\infty} \underline{X}(s) f^{-i2\pi s - \frac{1}{2}} \, ds. \tag{5}$$

The transform is furthermore isometric from $L^2(\mathbb{R}_+, df)$ to $L^2(\mathbb{R}, ds)$ [8]:

$$\int_{-\infty}^{+\infty} |\underline{X}(s)|^2 \, ds = \int_0^{+\infty} |X(f)|^2 \, df, \tag{6}$$

both members of the above equality defining a quantity referred to in the following as the energy of the signal.

The importance of both Mellin and scale transforms in the context of signal analysis comes primarily from the fact that they are naturally adapted to scale changes. In fact, if a spectrum is transformed according to $X(f) \mapsto X_\alpha(f) := \alpha^{-\frac{1}{2}} X(f/\alpha)$, its inverse Fourier transform $x(t)$ (i.e., the corresponding signal in the time domain) is dilated or compressed according to $x(t) \mapsto x_\alpha(t) = \alpha^{\frac{1}{2}} x(\alpha t)$, $\alpha > 0$, whereas its scale transform is only affected by a phase factor: $\underline{X}(s) \mapsto \underline{X}_\alpha(s) = \underline{X}(s) \alpha^{-i2\pi s}$, so that its scale energy density remains unchanged: $|\underline{X}_\alpha(s)|^2 = |\underline{X}(s)|^2$.

Different and complementary interpretations of the dimensionless variable s can be provided. A first one is obtained by considering (5) from a "building blocks" point of view, the complete spectrum $X(f)$ turning out to result from a (suitably weighted) linear superposition of elementary waveforms of the form

$$H_s(f) := f^{-i2\pi s - \frac{1}{2}} U(f), \tag{7}$$

3

with $U(f)$ the unit step function. Such waveforms have a non-constant group delay which reads

$$
\begin{aligned}
t_{H_s}(f) \quad &:= \quad -\frac{1}{2\pi} \frac{\partial}{\partial f} \arg H_s(f) \\
&= \quad -\frac{1}{2\pi} \frac{\partial}{\partial f} \left(-2\pi s \log f\right) \\
&= \quad \frac{s}{f},
\end{aligned}
\tag{8}
$$

thus allowing us to interpret s as an *hyperbolic chirp rate*.

Within this interpretation, the dimensionless parameter s can be viewed as resulting from a time $\times$ frequency product, which allows for its companion interpretation in terms of *scale*. The reason is that the operator $\mathcal{S}$ associated to $s = tf$ (in the sense of, e.g., Weyl's association rule [10, 17, 18]) is

$$
\mathcal{S} = \frac{\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T}}{2},
$$

where $\mathcal{T}$ and $\mathcal{F}$ stand for the usual time and frequency operators, defined, respectively, by $(\mathcal{T}x)(t) = t\,x(t)$ and $(\mathcal{F}x)(t) = (-i/2\pi)(dx/dt)(t)$. It thus follows that, for any $\alpha > 0$, we have

$$
\left(\alpha^{i2\pi\mathcal{S}} X\right)(f) = \alpha^{-\frac{1}{2}} X\left(\frac{f}{\alpha}\right),
\tag{9}
$$

making of $\mathcal{S}$ the infinitesimal generator of the scaling operator whose eigenfunctions are precisely the hyperbolic chirps $H_s(f)$ defined in (7).

Since the scale transform plays an important role for analytic signals, it is natural to consider in which respects scale transform pairs $\{X(f), \underline{X}(s)\}$ may be faced with joint limitations, as it is known for Fourier transform pairs. This will be addressed in the following from a number of different perspectives. Section 2 will consider scale transform pairs in place of Fourier transform pairs, and basic limitations will be established and discussed, based on various measures of spread (Heisenberg-type inequalities for variance-like measures and Hirschman-type inequalities for entropy). The same issue of maximally concentrating a signal in both scale and frequency domains will then be re-addressed in Section 3 via spread measures of the same type, but applied directly to a joint scale-frequency distribution. A

collection of results pertaining to Altes and Bertrand distributions will be established in this direction, as well as a new form of uncertainty relation for the wavelet transform.

While being essentially of a theoretical nature, all of the results discussed here should prove useful for a number of signal processing applications. One can mention for instance the optimal design of wideband signals, the performance evaluation of wideband systems or the characterization of elementary waveforms in signal decompositions, to name but a few.

2 Inequalities for scale transforms

2.1 Variance inequalities

If we consider (normalized) energy distributions—such as those given in (6)—in analogy with probability distribution functions, a first natural way of measuring how much they are localized amounts to evaluating their spread around some central value. To make things more precise, we will adopt the following conventions and notations:

Definition 2 *Given an integrable density $\rho(v)$ defined on $\mathbb{R}$, its energy $E(\rho)$, its arithmetic mean $m_a(\rho)$ and its arithmetic variance $V_a(\rho)$ are defined by*

$$E(\rho) := \int_{-\infty}^{+\infty} \rho(v)\, dv, \tag{10}$$

$$m_a(\rho) := \frac{1}{E(\rho)} \int_{-\infty}^{+\infty} v\, \rho(v)\, dv, \tag{11}$$

$$V_a(\rho) := \frac{1}{E(\rho)} \int_{-\infty}^{+\infty} (v - m_a(\rho))^2\, \rho(v)\, dv, \tag{12}$$

respectively.

As far as variances of a Fourier transform pair are concerned, a classical result in Fourier analysis is the Heisenberg-Gabor inequality, according to which a time signal $x(t)$ and its frequency spectrum $X(f)$ are such that [19]:

$$V_a\left(|X|^2\right) V_a\left(|x|^2\right) \geq \frac{1}{16\pi^2}, \tag{13}$$

with equality if and only if $X(f)$ (and, hence, $x(t)$) is a Gaussian, i.e., is of the form:

$$X(f) = K \exp\left(-a(f - \log b)^2 + i(cf + d)\right), \tag{14}$$

5

with $a, b \in I\!\!R_+$ and $c, d, K \in I\!\!R$.

In the space of analytic signals, Gaussians are clearly not allowed, with the consequence that the lower bound cannot be attained. One can therefore ask for Heisenberg-Gabor-type inequalities that would apply specifically to analytic signals, and in particular when considering the scale transform in place of the Fourier transform. A first result in this direction is given by the following (proofs can be found in [8, Sect. 11.3.1.2] or [10, Sect. 18.9] and will not be repeated here, the original result going back to [23])

Proposition 1 *An analytic signal $X(f)$ and its scale transform $\underline{X}(s)$ have arithmetic variances such that*

$$V_a\left(|X|^2\right) V_a\left(|\underline{X}|^2\right) \geq \frac{m_a^2(|X|^2)}{16\pi^2}, \tag{15}$$

with equality if and only if $X(f)$ is a "Klauder wavelet" of the form

$$X(f) = K \, \exp\left(a \log f - bf + i(c \log f + d)\right) U(f), \tag{16}$$

with $a > -\frac{1}{2}$, $b \in I\!\!R_+$ and $c, d, K \in I\!\!R$.

2.2 Modified variance inequalities

Proposition 1 provides indeed a lower bound for a joint concentration of a signal in both frequency and scale, but the use of the arithmetic variance in (15) can be questioned. In fact, in the case of analytic signals, the use of an ordinary (arithmetic) variance introduces some unnecessary weight on negative frequencies, for which no energy contributions exist (by construction, analytic signals vanish on the real half-line of negative frequencies). It furthermore defines a measure of spread that is not scale-invariant, whereas scale-invariance would be expected to be a key feature in this context. We can therefore consider a more general definition of variance, adapted to signals whose spectrum is in $I\!\!R_+$ only, as follows

Definition 3 *Given an integrable density $\rho(v)$ defined on $I\!\!R_+$, its energy $E_+(\rho)$ and its generalized variance $V(\rho)$ are defined by*

$$E_+(\rho) := \int_0^{+\infty} \rho(v) \, dv,$$

$$V(\rho) := \inf_m \frac{1}{E_+(\rho)} \int_0^{+\infty} (\varphi(v) - \varphi(m))^2 \, \rho(v) \, dv, \tag{17}$$

respectively, with $\varphi : \mathbb{R}_+ \to \mathbb{R}$ some monotonic function.

Differentiating (17) with respect to m, we get readily that the natural choice for the corresponding "mean" of the density ρ is

$$m = m(\rho) := \varphi^{-1}\left(\frac{1}{E_+(\rho)} \int_0^{+\infty} \varphi(v) \, \rho(v) \, dv\right), \tag{18}$$

which has the form of a "quasi-arithmetic" generalized mean, in the sense of [20, Chapter III], the ordinary (arithmetic) mean being attached to the specific choice $\varphi(v) = v$.

If we now impose a scale-invariance requirement on (17), we are led to the following

Proposition 2 *Given an integrable density $\rho(v)$ defined on $\mathbb{R}_+$, the only generalized variance (in the sense of Definition 3) that is scale-invariant, i.e., such that $m(\rho_\alpha) = m(\rho)/\alpha$ and $V(\rho_\alpha) = V(\rho)$ for any $\rho_\alpha(v) := \alpha \, \rho(\alpha v), \alpha > 0$, is characterized by $\varphi(v) = \log v$, up to an affine transformation.*

Proof — The scaled distribution $\rho_\alpha(v)$ is such that $E_+(\rho_\alpha) = E_+(\rho)$ and its generalized mean reads

$$m(\rho_\alpha) = \varphi^{-1}\left(\frac{1}{E_+(\rho)} \int_0^{+\infty} \varphi\left(\frac{v}{\alpha}\right) \rho(v) \, dv\right).$$

For the scale-invariance requirement on variance to make sense, the mean has itself to be scale-covariant, i.e., such that $m(\rho_\alpha) = m(\rho)/\alpha$. Such a condition of homogeneity on the mean is known [20, pages 68 and 150] to restrict solutions to $\varphi(v) = \log v$ or $\varphi(v) = v^r, r \neq 0$, up to an affine transformation. Assuming this to hold, we can write

$$
\begin{aligned}
V(\rho_\alpha) &= \frac{1}{E_+(\rho_\alpha)} \int_0^{+\infty} (\varphi(v) - \varphi(m(\rho_\alpha))^2 \, \rho_\alpha(v) \, dv \\
&= \frac{1}{E_+(\rho)} \int_0^{+\infty} \left(\varphi\left(\frac{v}{\alpha}\right) - \varphi\left(\frac{m(\rho)}{\alpha}\right)\right)^2 \rho(v) \, dv,
\end{aligned}
$$

and it follows that the scale invariance $V(\rho_\alpha) = V(\rho)$ requires that

$$\psi(v, \alpha) := \varphi\left(\frac{v}{\alpha}\right) - \varphi(v)$$

be independent of v, for any α. Differentiating $\psi(v, \alpha)$ with respect to v, we must then have

$$\frac{1}{\alpha}\varphi'\left(\frac{v}{\alpha}\right) = \varphi'(v)$$

for any v and any α. It thus follows, by fixing $\alpha = v$, that we must have $\varphi(v) = \log v$ (up to an affine transformation) and the proof is complete.

Ignoring affine transformations, the choice $\varphi(v) = \log v$ in Definition 3 appears as the most relevant one for analytic signals, and we will adopt the following

Definition 4 *Given an integrable density $\rho(v)$, defined on $\mathbb{R}_+$ and of energy $E_+(\rho)$, its geometric mean $m_g(\rho)$ and its geometric variance $V_g(\rho)$ are defined by*

$$m_g(\rho) := \exp\left(\frac{1}{E_+(\rho)}\int_0^{+\infty} \log v\, \rho(v)\, dv,\right) \tag{19}$$

$$V_a(\rho) := \frac{1}{E_+(\rho)}\int_0^{+\infty} \log^2\left(\frac{v}{m_g(\rho)}\right)\rho(v)\, dv, \tag{20}$$

respectively.

This definition easily leads us to the following

Proposition 3 *An analytic signal $X(f)$ and its scale transform $\underline{X}(s)$ have, respectively, geometric and arithmetic variances such that*

$$V_g\left(|X|^2\right)V_a\left(|\underline{X}|^2\right) \geq \frac{1}{16\pi^2}, \tag{21}$$

with equality if and only if $X(f)$ is an "Altes wavelet" of the form

$$X(f) = K\exp\left(-\frac{1}{2}\log f - a\log^2(f/b) + i(c\log f + d)\right)U(f), \tag{22}$$

with $a, b \in \mathbb{R}_+$ and $c, d, K \in \mathbb{R}$.

Proof — Starting from the remark that $\left\{\underline{X}(s), \widetilde{X}(u)\right\}$ form a Fourier transform pair (see eq.(3)), we can readily apply to it the usual Heisenberg-Gabor inequality (13) under the form:

$$V_a\left(|\underline{X}|^2\right)V_a\left(|\widetilde{X}|^2\right) \geq \frac{1}{16\pi^2}. \tag{23}$$

A straightforward calculation shows that

$$m_a\left(|\widetilde{X}|^2\right) = \log m_g\left(|X|^2\right)$$

and

$$V_a\left(|\widetilde{X}|^2\right) = V_g\left(|X|^2\right),$$

and we obtain immediately the inequality (21).

Since equality in (23) is achieved if and only $\widetilde{X}(u)$ (and, hence, $\underline{X}(s)$) is a Gaussian as in (14), signals guaranteeing that the lower bound of (21) is attained are then simply derived from (14) by using the equivalence $X(f) = f^{-\frac{1}{2}}\widetilde{X}(\log f)U(f)$, thus leading to the form (22).

Remark — The family of waveforms (22) has been introduced by R.A. Altes [1] in a context of active sonar, such as encountered in bat echolocation. "Altes wavelets" are, in fact, solutions of the problem of "Doppler tolerance," which consists of estimating at best (i.e., with no bias and maximum signal-to-noise ratio) a time delay in the presence of some unknown Doppler shift.

Having introduced the geometrical variance as a spread measure that is well-adapted to signals whose frequency spectrum vanishes on the real half-line of negative frequencies, we can think of making use of it, not only for establishing scale-frequency inequalities, but also for having a new look at time-frequency inequalities. For doing so, we will need the following

Definition 5 *Given an integrable density $\rho(v)$, defined on $\mathbb{R}_+$ and of energy $E_+(\rho)$, its harmonic mean $m_h(\rho)$ is defined by*

$$m_h(\rho) := \left(\frac{1}{E_+(\rho)}\int_0^{+\infty}\frac{1}{v}\,\rho(v)\,dv\right)^{-1}. \tag{24}$$

With this definition, which is nothing but the special case of (18) attached to $\varphi(v) = 1/v$, we can establish the following

Proposition 4 *An analytic signal $X(f)$ and its Fourier transform $x(t)$ have, respectively, geometric and arithmetic variances such that*

$$V_g\left(|X|^2\right) V_a\left(|x|^2\right) \geq \frac{1}{16\pi^2 m_h^2(|X|^2)}, \tag{25}$$

under the assumption that $m_a(|x|^2) = 0$, and that $|X(f)|^2$ and $|X(f)|^2 \log f$ both vanish when f goes to zero and infinity.

Proof — Let I be the integral defined as

$$I := \int_0^{+\infty} \log\left(\frac{f}{m_g(|X|^2)}\right) \overline{X(f)}\, X'(f)\, df. \tag{26}$$

Assuming that $|X(f)|^2$ and $|X(f)|^2 \log f$ both vanish when f goes to zero and infinity, an integration by parts leads to

$$\mathrm{Re}\,\{I\} = -\frac{1}{2} \int_0^{+\infty} \frac{1}{f}\, |X(f)|^2\, df$$

or, using Definition 5,

$$\mathrm{Re}\,\{I\} = -\frac{E_+(|X|^2)}{2\, m_h(|X|^2)},$$

and it follows that

$$|I|^2 \geq (\mathrm{Re}\,\{I\})^2 = \frac{E_+^2(|X|^2)}{4\, m_h^2(|X|^2)}. \tag{27}$$

Going back to (26) and using the fact that the time-domain density $|x(t)|^2$ is supposed to be zero-mean, application of the Cauchy-Schwarz inequality gives

$$
\begin{aligned}
|I|^2 &\leq \int_0^{+\infty} \log^2\left(\frac{f}{m_g(|X|^2)}\right) |X(f)|^2\, df \times \int_0^{+\infty} |X'(f)|^2\, df \\
&= 4\pi^2\, E_+^2\left(|X|^2\right) V_g\left(|X|^2\right) V_a\left(|x|^2\right),
\end{aligned}
$$

whence the claimed inequality, when combining this result with (27).

Remark — The bound in (25) cannot be expected to be tight, since conditions for equality in the Cauchy-Schwarz inequality would lead to waveforms such that

$$|X(f)|^2 = K\, \exp\left(-a\left(f \log(f/b) - f\right)\right) U(f),$$

with $a, b, K \in I\!\!R_+$, thus violating the assumed boundary conditions at $f = 0$.

2.3 Entropy inequalities

Entropy summarizes the complete information about a distribution, taking into account moments of all orders. This contrasts with variance-type measures, which only consider the two first moments and, except in the Gaussian case, do not exhaust all of the information available from a distribution. Furthermore, entropy providing a global measure of "peakiness" for a distribution, it may serve as a basis for deriving inequalities relative to the energy distribution of a signal and of its spectrum.

In order to simplify derivations and to cope more easily with conventional definitions of entropy, we will assume throughout this Section that all the considered signals are of unit-energy, i.e., that $E(|X|^2) = 1$, and we will adopt the following

Definition 6 *Given an integrable density $\rho(v)$ defined on $\mathbb{R}$, its Shannon entropy $H(\rho)$ is defined by*

$$H(\rho) := -\int_{-\infty}^{+\infty} \rho(v) \log \rho(v) \, dv. \tag{28}$$

In the case where $\rho(v)$ is defined on $\mathbb{R}_+$, we will make use of the notation:

$$H_+(\rho) := -\int_{0}^{+\infty} \rho(v) \log \rho(v) \, dv. \tag{29}$$

A classical result of information theory, first stated by Shannon, is that entropy and (arithmetic) variance satisfy the inequality [2, Theorem 8.3.3]

$$H(\rho) \leq \frac{1}{2} \log \left(2\pi e \, V_a(\rho) \right), \tag{30}$$

with equality for Gaussian distributions and only for them.

Another important result—as it has been initially conjectured by Hirschman [21] and later proved by Beckner [5]—is that

$$H\left(|X|^2\right) + H\left(|x|^2\right) \geq 1 - \log 2, \tag{31}$$

with equality if and only if X (and, hence, x) is a Gaussian of the form (14).

Results along similar lines can be obtained for analytic signals and their scale transforms. They are given in the following Lemma and Proposition.

Lemma 1 *The Shannon entropy and the geometric variance of a unit-energy analytic signal* $X(f)$ *satisfy*

$$H_+\left(|X|^2\right) \leq \frac{1}{2}\log\left(2\pi e\, V_g\left(|X|^2\right)\right) + \log m_g\left(|X|^2\right), \tag{32}$$

with equality if and only if $X(f)$ *is an Altes wavelet of the form (22).*

Proof — According to the model (22), the unit energy analytic signal $A(f)$ having $m_g(|A|^2)$ for geometric mean and $V_g(|A|^2)$ for geometric variance reads :

$$A(f) = \left(2\pi V_g\left(|A|^2\right)\right)^{-1/4} f^{-1/2} \exp\left(-\frac{1}{4V_g(|A|^2)}\log^2\left(f/m_g\left(|A|^2\right)\right)\right) U(f).$$

If $X(f)$ is a given unit-energy analytic signal, we know that [2]

$$\begin{aligned}
H_+\left(|X|^2\right) &= -\int_0^{+\infty} |X(f)|^2 \log |X(f)|^2 df \\
&\leq -\int_0^{+\infty} |X(f)|^2 \log |Y(f)|^2 df
\end{aligned}$$

for any unit-energy analytic signal $Y(f)$, with equality if and only if $X(f) = Y(f)$. If we now let $A(f)$ be such that $m_g(|X|^2) = m_g(|A|^2)$ and $V_g(|X|^2) = V_g(|A|^2)$, it suffices to set $Y(f) = A(f)$ to obtain the claimed result.

Given the above Definition 6 and Lemma 1, we can establish the following (Hirschman-type) inequality :

Proposition 5 *A unit-energy analytic signal* $X(f)$ *and its Mellin transform* $\underline{X}(s)$ *have Shannon entropies such that*

$$H_+\left(|X|^2\right) + H\left(|\underline{X}|^2\right) \geq 1 - \log 2 + \log m_g\left(|X|^2\right), \tag{33}$$

with equality if and only if $X(f)$ *is of the form (22).*

Proof — We know from Beckner's sharp form of the classical Hausdorff-Young inequality [5] that any function $Y(x) \in L^p(I\!R, dx), 1 \leq p \leq 2$, has a Fourier transform $\widehat{Y}(\xi)$ such that

$$\|\widehat{Y}\|_q \leq C_{p,q}\|Y\|_p, \tag{34}$$

with $1/p + 1/q = 1$, $C_{p,q} = \left(p^{1/p}/q^{1/q}\right)^{1/2}$ and where the L^p-norm is defined by

$$\|Y\|_p = \left(\int_{-\infty}^{+\infty} |Y(x)|^p dx\right)^{1/p}.$$

We can therefore apply the inequality (34) to the pair $\left\{\widehat{Y}, Y\right\} = \left\{\underline{X}, \widetilde{X}\right\}$ to obtain

$$\|\underline{X}\|_q^q \leq C_{q/(q-1),q}^q \|\widetilde{X}\|_{q/(q-1)}^{q-1}. \tag{35}$$

Remarking that $\|\underline{X}\|_2 = \|\widetilde{X}\|_2 = \|X\|_2 = 1$ and $C_{2,2} = 1$, we can make use of the following argument, outlined in [19, Lemma 5.5]: if two functions F and G are such that $F(x) \leq G(x), x \geq a$, and $F(a) = G(a)$, then their derivatives satisfy $F'(a) \leq G'(a)$. Differentiating both sides of the inequality (35) with respect to q and evaluating them at $q = 2$, we get therefore that

$$\left.\frac{\partial}{\partial q}\|\underline{X}\|_q^q\right|_{q=2} \leq \left.\frac{\partial}{\partial q}\left[\frac{q^{\frac{q}{2}-1}}{(q-1)^{\frac{q-1}{2}}}\left(\int_0^{+\infty} |X(f)|^{\frac{q}{q-1}} f^{\frac{1}{2}\frac{q}{q-1}-1} df\right)^{q-1}\right]\right|_{q=2},$$

thus leading to the claimed result, after some manipulations.

In order to prove that signals of the form (22) are indeed minimizers for the entropy inequality (33), we can combine (30), (32) and (33) to obtain :

$$\begin{aligned}
\frac{1}{2}\log\left[4\pi^2 e^2\, V_a\left(|\underline{X}|^2\right) V_g\left(|X|^2\right) m_g^2\left(|X|^2\right)\right] &\geq H_+\left(|X|^2\right) + H\left(|\underline{X}|^2\right) \\
&\geq 1 - \log 2 + \log m_g\left(|X|^2\right) \\
&= \frac{1}{2}\log\left(\frac{e^2}{4} m_g^2\left(|X|^2\right)\right).
\end{aligned}$$

Comparing the most left and most right terms of this sequence of inequalities, we thus recover the inequality (21), whose lower bound has been shown in Proposition 3 to be tight for signals of the form (22), and the proof is complete.

Remark — The proof above has been given for a sake of completeness, but a short-cut could have been used, assuming the inequality (31) to be known. In fact, proceeding as in the proof of Proposition 3, it is easy to establish that

$$H\left(|\widetilde{X}|^2\right) = H_+\left(|X|^2\right) - \log m_g\left(|X|^2\right).$$

It then suffices to apply Hirschman's inequality (31) to the pair $\left\{\underline{X}(s), \widetilde{X}(u)\right\}$ for getting (33).

2.4 Narrowband limit

We can remark that, depending on the choice we make for measuring the frequency spread, we may end up with different "most concentrated" signals which can be equally considered as "optimal" in the sense that the lower bound of the corresponding inequality is attained. A further observation is however that all the inequalities considered so far (and, hence, the associated "optimal" signals) tend to become identical in the limit of narrowband signals.

More precisely, it is easy to show that if a density $\rho(v)$ is mainly supported over an interval $[m_a(\rho) - \varepsilon/2, m_a(\rho) + \varepsilon/2]$, with $\varepsilon \ll m_a(\rho)$, we then have

$$
\begin{aligned}
\int_0^{+\infty} \log v \, \rho(v) \, dv \quad &\sim \quad \int_{-\varepsilon/2}^{+\varepsilon/2} \log\left(m_a(\rho) + \xi\right) \rho\left(m_a(\rho) + \xi\right) d\xi \\
&\sim \quad \int_{-\varepsilon/2}^{+\varepsilon/2} \left(\log m_a(\rho) + \frac{\xi}{m_a(\rho)}\right) \rho\left(m_a(\rho) + \xi\right) d\xi \\
&\sim \quad \log m_a(\rho) \, E(\rho)
\end{aligned}
$$

and, thus, $m_g(\rho) \sim m_a(\rho)$. Using similar approximations, we can derive in the same way that $m_h(\rho) \sim m_a(\rho)$ and that

$$
V_g(\rho) \sim \frac{1}{m^2(\rho)} \, V_a(\rho),
$$

with $m(\rho) := m_a(\rho)$, $m_g(\rho)$ or $m_h(\rho)$, indifferently.

It then follows that, in the narrowband limit, inequality (25) naturally reduces to the classical Heisenberg-Gabor inequality (13), and that (21) reduces to the standard scale-frequency inequality (15). Furthermore, in the same narrowband limit, we have

$$
\begin{aligned}
|\underline{X}(s)|^2 \quad &\sim \quad \left|\int_{-\varepsilon/2}^{+\varepsilon/2} X(m(\rho) + \xi) \left(m(\rho) + \xi\right)^{i2\pi s - \frac{1}{2}} d\xi\right|^2 \\
&\sim \quad \frac{1}{m(\rho)} \left|x\left(\frac{s}{m(\rho)}\right)\right|^2,
\end{aligned}
\tag{36}
$$

so that

$$
V_a\left(|\underline{X}|^2\right) \quad \sim \quad m_a^2\left(|x|^2\right) V_a\left(|x|^2\right)
$$

14

$$\sim \quad m_g^2\left(|x|^2\right) V_a\left(|x|^2\right),$$

and both scale-frequency inequalities (21) and (15) reduce to the standard Heisenberg-Gabor inequality (13).

As far as entropy is concerned, the approximation (36) guarantees that

$$H\left(|\underline{X}|^2\right) \sim H\left(|x|^2\right) + \log m\left(|X|^2\right),$$

so that the modified Hirschman-type inequality (33) reduces to the classical one (31).

An interpretation of this narrowband limit can be given on the time-frequency plane. Localization of a signal in frequency amounts to concentrating most of its energy within a strip of the plane, whereas—from (33)—localization in scale is associated to a time-frequency domain limited by two hyperbolae. It thus follows that the specificity of scale analysis mostly concerns wideband signals, while narrowband signals only face the usual localization limitations of Fourier analysis on "rectangular" domains. A graphical illustration of this behaviour is given in Figure 1.

A companion illustration of this property is given in Figure 2, where Klauder and Altes wavelets are both evidenced to become asymptotically undistinguishable from Morlet wavelets (i.e., Gaussian modulated wavelets) in the narrowband limit.

3 Inequalities on the scale-frequency plane

Limitations for a simultaneous localization of a signal in both scale and frequency were considered so far from the point of view of *individual* energy densities in each variable. Figure 1 suggests however a complementary approach, which would consist of an evaluation of how much localized a *joint* distribution of scale and frequency could be, a situation that is of course reminiscent of similar approaches that may have been used in time-frequency and/or time-scale analysis (see, e.g., [12, 13, 14, 15, 17, 22]). To make things precise, some basics of scale-frequency distributions have first to be recalled.

3.1 Joint distributions of scale and frequency

General procedures for obtaining joint distributions of arbitrary variables (associated with non-commuting operators) have been proposed and discussed, leading generally to *classes* of solutions [3, 9, 10, 17]. As far as scale and frequency are concerned, one way of obtaining a family of admissible distributions is to make use of the time-frequency interpretation illustrated in Figure 1, and to impose suitable covariance requirements. We have seen in (9) that the exponential of the scale operator acts on the spectrum of a signal as a dilation operator whose eigenfunctions are the hyperbolic chirps (7). It is therefore natural to require that scale-frequency distributions $\check{P}_X(s, f)$ adapted to scale-changing operations be such that

$$\check{P}_{X_{\alpha,\sigma}}(s, f) = \check{P}_X\left(s - \sigma, \frac{f}{\alpha}\right),$$

when the analyzed signal is transformed according to

$$X(f) \mapsto X_{\alpha,\sigma}(f) := \alpha^{-\frac{1}{2}} X\left(\frac{f}{\alpha}\right) H_\sigma(f),$$

with $H_\sigma(f)$ defined as in (7).

Assuming that $\check{P}_X$ is sesquilinear in X—so as to have a simple energetic interpretation for the distributions—, such a covariance requirement with respect to dilations and hyperbolic shifts results in the scale-frequency version (i.e., with the identification $s = tf$) of the so-called "hyperbolic class" [9]. We will adopt in the following the[2]

Definition 7 *The hyperbolic class of an analytic signal $X(f)$ is given by all scale-frequency distributions of the form*

$$\check{P}_X(s, f) = \int_{-\infty}^{+\infty} \int_0^{+\infty} \Phi\left(s - \sigma, \log\frac{f}{\xi}\right) \check{Q}_X(\sigma, \xi) \, d\sigma \, \frac{d\xi}{\xi}, \tag{37}$$

with $\Phi(s, f)$ some arbitrary kernel and

$$\check{Q}_X(s, f) := f \int_{-\infty}^{+\infty} X\left(f \, e^{u/2}\right) \overline{X\left(f \, e^{-u/2}\right)} \, e^{i2\pi su} \, du \tag{38}$$

---

[2]Distributions of the hyperbolic class are usually considered as time-frequency distributions. In order to put a specific emphasis on their scale-frequency interpretation, we will make use of the following convention for the notation: a time-frequency distribution will be written as $P_X(t, f)$, and its scale-frequency counterpart as $\check{P}_X(s, f)$, with the equivalences $P_X(t, f) = \check{P}_X(tf, f)$ and $\check{P}_X(s, f) = P_X(s/f, f)$.

the *"Altes distribution."*

As for the scale transform, which—according to (4)—can be expressed in the form of a Fourier transform applied to a suitably warped version of the signal, the Altes distribution can be written as

$$\check{Q}_X(s, f) = W_{\widetilde{X}}(s, \log f),$$

where $W_X(t, f)$ is the usual *Wigner distribution*[3] [9, 10, 17]

$$W_X(t, f) := \int X\left(f + \frac{\xi}{2}\right) \overline{X\left(f - \frac{\xi}{2}\right)} e^{i2\pi\xi t} \, d\xi.$$

More generally, all members of the hyperbolic class (37) can be expressed as

$$\check{P}_X(s, f) = C_{\widetilde{X}}(s, \log f), \tag{39}$$

where

$$C_X(t, f) := \int_{-\infty}^{+\infty} \int \Phi\left(t - \tau, f - \xi\right) W_X(\tau, \xi) \, d\tau \, d\xi \tag{40}$$

is a distribution of the so-called "Cohen class" [9, 10, 17], i.e., the class of all sesquilinear time-frequency distributions which are covariant with respect to shifts in time and frequency.

The hyperbolic class (37) and the Cohen class (40) are equivalent by warping, but they have no intersection, since none of the covariances of the Cohen class (shifts in both time and frequency) is shared by the hyperbolic class. Nevertheless, there exist other classes of distributions which may have an intersection with either the hyperbolic class or the Cohen class. This is especially the case of the "affine class" [9] which, by construction, is covariant with respect to shifts in time and dilations. A key distribution of the affine class that also belongs to the hyperbolic class is the (unitary) "Bertrand distribution" [7], whose (scale-frequency) definition reads

$$\check{B}_X(s, f) := f \int_{-\infty}^{+\infty} \sqrt{\lambda(u)\,\lambda(-u)}\, X\left(f\,\lambda(u)\right) \overline{X\left(f\,\lambda(-u)\right)} \, e^{i2\pi su} \, du, \tag{41}$$

with

$$\lambda(u) := \frac{u}{1 - e^{-u}}. \tag{42}$$

---

[3]In the general case, integration is performed from $-\infty$ to $+\infty$ but, for analytic signals, the integration limits are actually $-2f$ and $+2f$.

Even if they cannot be expressed as scale-frequency distributions *stricto sensu*, other members of the affine class may be of interest. The most prominent example is the "scalogram"—the squared modulus of a wavelet transform—whose (time-frequency) definition is [9, 17]

$$|T_X(t,f)|^2 := \frac{f_0}{f} \left| \int_0^{+\infty} X(\xi) \, \overline{\Psi\left(\frac{f_0}{f}\xi\right)} \, e^{i2\pi\xi t} \, d\xi \right|^2, \tag{43}$$

where $\Psi(f)$ is the frequency spectrum of the analyzing wavelet and f_0 a (non-zero) reference frequency, related to some "mean" of the bandpass density $|\Psi|^2$. Given (43), a (formal) scale-frequency version of the scalogram will be defined in the following as

$$\left|\check{T}_X(s,f)\right|^2 := \left|T_X\left(\frac{s}{f}, f\right)\right|^2. \tag{44}$$

3.2 Variance-type inequalities

Adopting the point of view of characterizing an analytic signal by means of a scale-frequency distribution, we can address the problem of its joint localization in both scale and frequency *via* joint spread measures defined on the half-plane. We will adopt the following definitions:

Definition 8 *Given a scale-frequency distribution $R(s,f)$, defined on $\mathbb{R} \times \mathbb{R}_+$ and integrable with respect to the measure $ds\,df/f$, its energy $E(R)$ and its arithmetic-arithmetic variance $V_{aa}(R)$ are defined by*

$$E(R) := \int_{-\infty}^{+\infty} \int_0^{+\infty} R(s,f)\,ds\,\frac{df}{f},$$

$$V_{aa}(R) := \frac{1}{E(R)} \int_{-\infty}^{+\infty} \int_0^{+\infty} \left[(s - m_a(R^{(s)}))^2 + (f - m_a(R^{(f)}))^2\right] R(s,f)\,ds\,\frac{df}{f}, \tag{45}$$

respectively, with $m_a(\rho)$ as in (11), and the marginal distributions $R^{(s)}(s)$ and $R^{(f)}(f)$ defined by

$$R^{(s)}(s) := \int_0^{+\infty} R(s,f)\,\frac{df}{f}$$

and

$$R^{(f)}(f) := \frac{1}{f} \int_{-\infty}^{+\infty} R(s,f)\,ds.$$

Definition 9 *Given a scale-frequency distribution $R(s, f)$, defined on $\mathbb{R} \times \mathbb{R}_+$ and integrable with respect to the measure $ds\, df/f$, its arithmetic-geometric variance $V_{ag}(R)$ is defined by*

$$V_{ag}(R) := \frac{1}{E(R)} \int_{-\infty}^{+\infty} \int_0^{+\infty} \left[(s - m_a(R^{(s)}))^2 + \log^2\left(\frac{f}{m_g(R^{(f)})}\right) \right] R(s, f)\, ds\, \frac{df}{f}, \quad (46)$$

with $m_g(\rho)$ as in (19).

Based on those definitions, it immediately follows that marginal distributions play a key role for making of the joint spread measure an equivalent substitute to the individual spread measures used previously. More precisely, we have (proof is trivial and will be omitted)

Lemma 2 *Given an analytic signal $X(f)$, and any of its scale-frequency distributions $\check{P}_X(s, f)$ with "correct" marginals, i.e., such that $\check{P}_X^{(s)}(s) = |\underline{X}(s)|^2$ and $\check{P}_X^{(f)}(f) = |X(f)|^2$, we have*

$$V_{aa}(\check{P}_X) = V_a\left(|\underline{X}|^2\right) + V_a\left(|X|^2\right) \tag{47}$$

and

$$V_{ag}(\check{P}_X) = V_a\left(|\underline{X}|^2\right) + V_g\left(|X|^2\right). \tag{48}$$

It follows from this lemma that joint spread measures of joint scale-frequency distributions with correct marginals merely re-express the inequalities obtained previously for a scale transform pair. This is made explicit by the following

Proposition 6 *Let $\check{P}_X(s, f)$ be a scale-frequency distribution with correct marginals, then we have*

$$V_{aa}(\check{P}_X) \geq \frac{m_a(|X|^2)}{2\pi}, \tag{49}$$

with equality if and only $X(f)$ is a Klauder wavelet of the form (16). In a similar way,

$$V_{ag}(\check{P}_X) \geq \frac{1}{2\pi}, \tag{50}$$

with equality if and only $X(f)$ is an Altes wavelet of the form (22)

Proof — Starting from (47), we can write

$$
\begin{aligned}
V_{aa}(\check{P}_X) &= V_a\big(|\underline{X}|^2\big) + V_a\big(|X|^2\big) \\
&= \left(\sqrt{V_a(|\underline{X}|^2)} - \sqrt{V_a(|X|^2)}\right)^2 + 2\sqrt{V_a(|\underline{X}|^2)\, V_a(|X|^2)} \\
&\geq 2\sqrt{V_a(|\underline{X}|^2)\, V_a(|X|^2)} \\
&\geq 2\frac{m_a(|X|^2)}{4\pi},
\end{aligned}
$$

the last inequality stemming from (15), and (49) follows. Similar reasoning, starting from (48) but using (21), would lead to (50). Both lower bounds are attained as previously.

The most simple examples of distributions with correct marginals within the hyperbolic class are the Altes distribution (38) and the unitary Bertrand distribution (41) but, more generally, all distributions of the form (37) with $\phi(0,1) = \phi(1,0) = 1$ (with $\phi(\xi, \tau)$ the two-dimensional Fourier transform of $\Phi(s, f)$) are solutions as well [9]. All these distributions are faced with the same limitations on their joint localization.

In the specific case of the unitary Bertrand distribution, inequality (50) complements the inequality (49), which had already been proven (in a slightly different form) in [16, Prop. 17]. Moreover, thanks to the warping equivalence mentioned in (39), we have that

$$
\int_{-\infty}^{+\infty}\int_{0}^{+\infty} G(s, \log f)\, \left|\check{P}_X(s, f)\right|^p\, ds\,\frac{df}{f} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} G(t, f)\, \left|C_{\widetilde{X}}(t, f)\right|^p\, dt\, df
$$

for any function G and any positive p (provided that the above integrals exist), so that many other inequalities can be derived from more standard time-frequency inequalities within the Cohen class, such as, e.g., those considered in [16, Sect. 3] or [22]. This will not be followed up here. We will rather concentrate on a more specific issue concerning wavelet transforms and scalograms.

3.3   An uncertainty relation for the wavelet transform

Numerous comments can be found in the literature about localization properties of the wavelet transform but, in most cases, it turns out that the considered localization (in

the time and frequency domains) concerns much more the wavelet used in computing the transform than the resulting transform itself. A few results can however be found (e.g. in [11], [13] or [25, 26]), that address explicitly the problem of the joint localization of a wavelet transform in the plane, and we will propose here a new result in that direction.

Being in fact primarily interested in energy localization, the analysis will be based on (quadratic) scalograms rather than on (linear) wavelet transforms. Scalograms (43) are affine time-frequency distributions and, as such, they may be expressed in terms of the unitary Bertrand distribution, which itself belongs to the affine class. More precisely, if we consider distributions (41) as affine time-*frequency* functions by using the aforementioned notation $B_X(t, f) := \check{B}_X(tf, f)$, unitarity (and real-valuedness) of the Bertrand distribution guarantees that

$$\left| \int_0^{+\infty} X(\xi) \, \overline{Y(\xi)} \, d\xi \right|^2 = \int_{-\infty}^{+\infty} \int_0^{+\infty} B_X(\tau, \xi) \, B_Y(\tau, \xi) \, d\tau \, d\xi \qquad (51)$$

for any $X(f)$ and $Y(f) \in L^2(\mathbb{R}_+, df)$. If we now let $Y(f)$ be some time-shifted and scaled version of a wavelet spectrum $\Psi(f)$ according to

$$Y(\xi) = \Psi_{t,f}(\xi) := \sqrt{\frac{f_0}{f}} \, \Psi\left(\frac{f_0}{f} \, \xi \right) e^{-i2\pi\xi t},$$

the left-hand side of (51) exactly coincides with the scalogram defined in (43). Moreover, we readily see from the definition (41) that

$$B_{\Psi_{t,f}}(\tau, \xi) = B_\Psi\left(\frac{f}{f_0} \, (\tau - t), \frac{f_0}{f} \, \xi \right),$$

so that, after inclusion in (51), we end up with the "regularization" formula [6]

$$|T_X(t, f)|^2 = \int_{-\infty}^{+\infty} \int_0^{+\infty} B_X(\tau, \xi) \, B_\Psi\left(\frac{f}{f_0} \, (\tau - t), \frac{f_0}{f} \, \xi \right) d\tau \, d\xi. \qquad (52)$$

The interpretation of this formula is that scalograms can be viewed as (affine) smoothed versions of Bertrand distributions, exactly as spectrograms are (standard) smoothed versions of Wigner distributions [9, 10, 17].

Owing to their definition, scalograms have no "correct" marginals and Lemma 2 cannot be applied to them. Furthermore, since they are not members of the hyperbolic class, they

cannot be expected to be covariant with respect to hyperbolic shifts. As a consequence, we will have—in order to make use of measures such as those defined in (45) or (46), explicitly constructed on $s = tf$—to restrict things further and deal only with those signals whose group delay is zero, so that the mean of their scale density be itself zero: such a property is in fact a corollary to the following [10]

Lemma 3 *Let $X(f)$ be any analytic signal of group delay $t_X(f)$. Assuming that $f\,|X(f)|^2$ vanishes when f goes to zero and infinity, we have*

$$m_a\left(|\underline{X}|^2\right) = m_a\left(t_X|X|^2\right).$$

Proof — We can write

$$
\begin{aligned}
m_a\left(|\underline{X}|^2\right) &= \int_{-\infty}^{+\infty} s\,|\underline{X}(s)|^2\,ds \\
&= \int_{-\infty}^{+\infty} s\left|\int_0^{+\infty} X(f)\,f^{i2\pi s-\frac{1}{2}}\,df\right|^2\,ds \\
&= \frac{1}{i2\pi}\int_0^{+\infty}\int_0^{+\infty} X(\xi)\,\overline{X(\zeta)}\,\delta'\left(\log\frac{\xi}{\zeta}\right)\frac{d\xi\,d\zeta}{\sqrt{\xi\zeta}} \\
&= \frac{i}{2\pi}\left(\int_0^{+\infty} f\,\overline{X(f)}\,X'(f)\,df + \frac{1}{2}\,E_+\left(|X|^2\right)\right).
\end{aligned}
$$

Assuming that $f\,|X(f)|^2$ vanishes when f goes to zero and infinity, an integration by parts shows that

$$\mathrm{Re}\left\{\int_0^{+\infty} f\,\overline{X(f)}\,X'(f)\,df\right\} = -\frac{1}{2}\,E_+\left(|X|^2\right),$$

whence

$$m_a\left(|\underline{X}|^2\right) = -\frac{1}{2\pi}\,\mathrm{Im}\left\{\int_0^{+\infty} f\,\overline{X(f)}\,X'(f)\,df\right\}.$$

With the notation $X(f) := |X(f)|\,\exp\{i\Phi_X(f)\}$, we may write

$$\mathrm{Im}\left\{\int_0^{+\infty} f\,\overline{X(f)}\,X'(f)\,df\right\} = \int_0^{+\infty} f\,\Phi'(f)\,|X(f)|^2\,df,$$

from which we get the claimed result, thanks to the definition

$$t_X(f) = -\frac{1}{2\pi}\,\Phi'(f)$$

22

of the group delay of $X(f)$.

Based on this material, we can now prove a new form of "uncertainty relation" for wavelet transforms. We will need one more definition:

Definition 10 *Given an integrable density $\rho(v)$, defined on $I\!R_+$ and of energy $E_+(\rho)$, its inverse quadratic mean $m_{iq}(\rho)$ is defined by*

$$m_{iq}(\rho) := \left(\frac{1}{E_+(\rho)} \int_0^{+\infty} \frac{1}{v^2} \, \rho(v) \, dv \right)^{-\frac{1}{2}}. \tag{53}$$

Proposition 7 *Let $X(f)$ be any analytic signal whose group delay is zero, and let $\Psi(f)$ be any wavelet of reference frequency $f_0 := m_g(|\Psi_0|^2)$ with $\Psi_0(f) := f^{-1/2}\Psi(f)$. Assuming that $\Psi(f) \in L^2(I\!R_+, f^{-(n+1)} \, df)$ for $n = 0, 1$ and 2, the corresponding scalogram $|\check{T}_X(s,f)|^2$ is such that*

$$V_{ag}\left(|\check{T}_X|^2 \right) \geq \frac{D(\Psi)}{2\pi}, \tag{54}$$

with

$$D(\Psi) := m_g\left(|\Psi_0|^2 \right) \left(\frac{1}{m_{iq}(|\Psi_0|^2)} + \frac{1}{m_h(|\Psi_0|^2)} \right) \geq 2, \tag{55}$$

and V_{ag}, m_g, m_{iq} and m_h as in (46), (53), (19) and (24), respectively.

Proof — Decomposing the arithmetic-geometric variance (46) in two terms and using the simplified notation

$$m := m_g\left(\frac{1}{f} \int_0^{+\infty} |\check{T}_X(s,f)|^2 \, ds \right),$$

we first get from the definition (43) of the scalogram that

$$
\begin{aligned}
I_2 \; &:= \; \frac{1}{E\left(|\check{T}_X|^2\right)} \int_{-\infty}^{+\infty} \int_0^{+\infty} \log^2\left(\frac{f}{m}\right) |\check{T}_X(s,f)|^2 \, ds \, \frac{df}{f} \\
&= \; \frac{1}{E\left(|\check{T}_X|^2\right)} \int_{-\infty}^{+\infty} \int_0^{+\infty} \log^2 f \, |\check{T}_X(s,f)|^2 \, ds \, \frac{df}{f} - \log^2 m \\
&= \; \frac{f_0}{E\left(|\check{T}_X|^2\right)} \int_0^{+\infty} \int_0^{+\infty} \log^2\left(\frac{f_0 \xi}{\zeta}\right) |X(\xi)|^2 \, |X(\zeta)|^2 \, d\xi \, \frac{d\zeta}{\zeta} - \log^2 m.
\end{aligned}
$$

23

Expanding the logarithm in the above integral, noting that $m = m_g(|X|^2)$ and that

$$E\left(|\check{T}_X|^2\right) = f_0\, C_0(\Psi)\, E_+\left(|X|^2\right), \tag{56}$$

with

$$C_n(\Psi) := \int_0^{+\infty} |\Psi(f)|^2\, \frac{df}{f^{n+1}}, \tag{57}$$

we obtain after some calculations that

$$I_2 = V_g\left(|X|^2\right) + V_g\left(|\Psi_0|^2\right), \tag{58}$$

with V_g as in (20).

Considering then the first term of (46), and assuming that the group delay of $X(f)$ is zero, Lemma 3 guarantees that it suffices to consider the quantity

$$
\begin{aligned}
I_1 &:= \int_{-\infty}^{+\infty} \int_0^{+\infty} s^2\, |\check{T}_X(s,f)|^2\, ds\, \frac{df}{f} \\
&= \int_{-\infty}^{+\infty} \int_0^{+\infty} t^2 f^2\, |T_X(t,f)|^2\, dt\, df.
\end{aligned}
\tag{59}
$$

The evaluation of this integral is better achieved by using the identity (51). Plugging (51) into (59) and making the appropriate changes of variables, we find that

$$
\begin{aligned}
I_1 =\ & f_0^3 \Bigg[\int_{-\infty}^{+\infty} \int_0^{+\infty} t^2 f^2\, B_X(t,f)\, dt\, df \times \int_{-\infty}^{+\infty} \int_0^{+\infty} B_\Psi(t,f)\, dt\, \frac{df}{f^3} \\
& -2 \int_{-\infty}^{+\infty} \int_0^{+\infty} tf\, B_X(t,f)\, dt\, df \times \int_{-\infty}^{+\infty} \int_0^{+\infty} t\, B_\Psi(t,f)\, dt\, \frac{df}{f^2} \\
& + \int_{-\infty}^{+\infty} \int_0^{+\infty} B_X(t,f)\, dt\, df \times \int_{-\infty}^{+\infty} \int_0^{+\infty} t^2\, B_\Psi(t,f)\, dt\, \frac{df}{f} \Bigg].
\end{aligned}
\tag{60}
$$

Remarking that

$$\int_{-\infty}^{+\infty} \int_0^{+\infty} tf\, B_X(t,f)\, dt\, df = m_a\left(|\underline{X}|^2\right)\, E_+\left(|X|^2\right),$$

we obtain from Lemma 3 and the assumption of zero-mean for the group delay that the middle term in (60) vanishes, and an explicit evaluation of the other terms leads to

$$I_1 = f_0^3\, E_+\left(|X|^2\right) \left[V_a\left(|\underline{X}|^2\right) C_2(\Psi) + \int_{-\infty}^{+\infty} \int_0^{+\infty} t^2\, B_\Psi(t,f)\, dt\, \frac{df}{f} \right]. \tag{61}$$

The evaluation of the remaining integral requires some more effort and can be carried out as follows. One can write

$$
\begin{aligned}
J &:= \int_{-\infty}^{+\infty} \int_0^{+\infty} t^2 \, B_\Psi(t,f) \, dt \, \frac{df}{f} \\
&= \int_0^{+\infty} \left(\int_{-\infty}^{+\infty} s^2 \, \check{B}_\Psi(s,f) \, ds \right) \frac{df}{f^4} \\
&= -\frac{1}{4\pi^2} \int_0^{+\infty} \frac{d^2}{du^2} \left[\mu(u) \, \Psi(f \, \lambda(u)) \, \overline{\Psi(f \, \lambda(-u))} \right]_{u=0} \frac{df}{f^3},
\end{aligned}
$$

with $\lambda(u)$ defined as in (42) and $\mu(u) := \sqrt{\lambda(u)\lambda(-u)}$.

Differentiating twice the bracketed term with respect to u, evaluating the obtained derivative in $u = 0$ by using the identities $\lambda'(0) = 1/2$, $\lambda''(0) = 1/6$ and $\mu''(0) = -1/12$, and integrating finally in f, we obtain the intermediate result:

$$
J = \frac{1}{4\pi^2} \, C_2(\Psi) \left[\frac{C_0(\Psi')}{C_2(\Psi)} - \frac{3}{4} \right]. \tag{62}
$$

This can be cast in a more interesting form using the modified wavelet $\Psi_0(f)$. From its definition, we have

$$
|\Psi'(f)|^2 = \frac{1}{4} \frac{|\Psi_0(f)|^2}{f} + f \, |\Psi_0'(f)|^2 + \mathrm{Re} \left\{ \Psi_0(f) \, \overline{\Psi_0'(f)} \right\},
$$

so that, after an integration by parts, we obtain

$$
\frac{C_0(\Psi')}{C_2(\Psi)} - \frac{3}{4} = \frac{1}{C_2(\Psi)} \int_0^{+\infty} |\Psi_0'(f)|^2 \, df = 4\pi^2 \frac{C_0(\Psi)}{C_2(\Psi)} V_a\big(|\psi_0|^2\big).
$$

Plugging this result into (62), normalizing the resulting contribution (61) and combining with (58), we finally obtain

$$
V_{ag}\big(|\check{T}_X|^2\big) = f_0^2 \frac{C_2(\Psi)}{C_0(\Psi)} V_a\big(|\underline{X}|^2\big) + V_g\big(|X|^2\big) + f_0^2 \, V_a\big(|\psi_0|^2\big) + V_g\big(|\Psi_0|^2\big). \tag{63}
$$

Proceeding as in the proof of Proposition 6, it is easy to show that the right-hand side of (63) is bounded below as

$$
V_{ag}\big(|\check{T}_X|^2\big) \geq 2 \sqrt{f_0^2 \frac{C_2(\Psi)}{C_0(\Psi)}} \sqrt{V_a(|\underline{X}|^2) \, V_g(|X|^2)} + 2 \sqrt{f_0^2 \, V_a(|\psi_0|^2) \, V_g(|\Psi_0|^2)}. \tag{64}
$$

Considering integrals $C_n(\Psi)$ of the type (57) and making explicit the definition of the reference frequency f_0 as the geometric mean of the density $|\Psi_0|^2$, we have

$$\sqrt{f_0^2 \, \frac{C_2(\Psi)}{C_0(\Psi)}} = \frac{m_g(|\Psi_0|^2)}{m_{iq}(|\Psi_0|^2)}.$$

As far as the second term of the sum in (64) is concerned, one may write

$$\sqrt{f_0^2 \, V_a(|\psi_0|^2) \, V_g(|\Psi_0|^2)} = \frac{m_g(|\Psi_0|^2)}{m_h(|\Psi_0|^2)} \sqrt{m_h^2(|\Psi_0|^2) \, V_a(|\psi_0|^2) \, V_g(|\Psi_0|^2)},$$

and (64) becomes

$$V_{ag}\Big(|\check{T}_X|^2\Big) \geq$$
$$2 \left[\frac{m_g(|\Psi_0|^2)}{m_{iq}(|\Psi_0|^2)} \sqrt{V_a(|\underline{X}|^2) \, V_g(|X|^2)} + \frac{m_g(|\Psi_0|^2)}{m_h(|\Psi_0|^2)} \sqrt{m_h^2(|\Psi_0|^2) \, V_a(|\psi_0|^2) \, V_g(|\Psi_0|^2)} \right].$$

According to the results of Propositions 3 and 4, both square roots in the above expression are each bounded below by $1/4\pi$, thus leading to (54).

Using the standard sequence of inequalities [20, Chapter VI]

$$m_{iq}\Big(|\Psi_0|^2\Big) \leq m_h\Big(|\Psi_0|^2\Big) \leq m_g\Big(|\Psi_0|^2\Big)$$

which hold for the different means involved in the definition of $D(\Psi)$, we finally get that

$$D(\Psi) = m_g\Big(|\Psi_0|^2\Big) \left(\frac{1}{m_{iq}(|\Psi_0|^2)} + \frac{1}{m_h(|\Psi_0|^2)} \right) \geq 2 \, \frac{m_g(|\Psi_0|^2)}{m_h(|\Psi_0|^2)} \geq 2,$$

and this completes the proof.

Remark 1 — Proposition 7 gives a lower bound for the (arithmetic-geometric) variance of scalograms on the plane, but it fails to characterize signal-wavelet pairs for which the lower bound would be attained. Given a fixed wavelet, one can however go back to the inequality (64) and consider it as an affine function (with positive weights) of the spread measure $\sqrt{V_a(|\underline{X}|^2) \, V_g(|X|^2)}$. We know from Proposition 3 that the minimum of this quantity is attained for $X(f)$ of the form (22), and the same applies to the joint measure $V_{ag}\Big(|\check{T}_X|^2\Big)$.

Remark 2 — We know that all of the different means involved in $D(\Psi)$ become identical in the limit of narrowband wavelets. In this limit, wavelet analysis tends however to be

equivalent to Fourier analysis and one can observe that, in this case, the asymptotic lower bound $1/\pi$ is identical to the one which controls the lowest (arithmetic-arithmetic) variance attainable by spectrograms when using the spread measure (45) [16].

3.4 An entropy inequality for the unitary Bertrand distribution

In a way similar to what has been done previously for scale transform pairs, variance-type inequalities can be supplemented by entropy inequalities in the case of joint scale-frequency distributions. Again, thanks to the warping equivalence (39), many results can be readily inferred from entropy inequalities relative to Wigner distributions and their smoothed versions within the Cohen class [4, 24]. This will not be detailed further.

We will here only consider the specific example of the unitary Bertrand distribution which, while being a member of the hyperbolic class, cannot be treated in a straightforward manner on the basis of the results given in [4] (this is so because the extensions of Lieb's results given in [4] only concern those members of the Cohen class such that their kernel function is in $L^1(I\!\!R \times I\!\!R_+)$, which is definitely not the case for the kernel function used for expressing the unitary Bertrand distribution as a "smoothed" Altes distribution [9]).

We will first give a definition and prove two lemmata.

Definition 11 *Given a positive scale-frequency density $R(s, f)$, integrable and normalized to one over $I\!\!R \times I\!\!R_+$, its Shannon entropy is defined by*

$$H_1(R) := - \int_{-\infty}^{+\infty} \int_0^{+\infty} R(s, f) \log R(s, f) \, ds \, \frac{df}{f}. \tag{65}$$

Lemma 4 *The unitary Bertrand distribution of an analytic signal $X(f)$ is bounded according to*

$$\|\check{B}\|_\infty \le 2 \, E_+ \left(|X|^2 \right). \tag{66}$$

Proof — Starting from the definition (41) of the unitary Bertrand distribution and using the Cauchy-Schwarz inequality, we can write

$$|\check{B}_X(s, f)|^2 = \left| \int_{-\infty}^{+\infty} \sqrt{f \, \lambda(u)} \, X(f \, \lambda(u)) \sqrt{f \, \lambda(-u)} \, \overline{X(f \, \lambda(-u))} \, e^{i2\pi su} \, du \right|^2$$

$$\leq \ \left(\int_{-\infty}^{+\infty} f\,\lambda(u)\,|X(f\,\lambda(u))|^2\,du\right)^2,$$

so that we get, after a change of variables, the first majoration

$$|\check{B}_X(s,f)| \leq \int_0^{+\infty} |X(\xi)|^2\, g\!\left(\frac{\xi}{f}\right) d\xi, \tag{67}$$

with

$$g(x) := \frac{x}{\lambda'(\lambda^{-1}(x))}. \tag{68}$$

Another way of obtaining a majoration consists of writing

$$
\begin{aligned}
|\check{B}_X(s,f)|^2 \ &= \ \left|\int_{-\infty}^{+\infty} \sqrt{f\,\lambda(-u)}\,X(f\,\lambda(u))\,\sqrt{f\,\lambda(u)}\,\overline{X(f\,\lambda(-u))}\,e^{i2\pi su}\,du\right|^2 \\
&\leq \ \left(\int_{-\infty}^{+\infty} f\,\lambda(-u)\,|X(f\,\lambda(u))|^2\,du\right)^2,
\end{aligned}
$$

so that, using the relation [7]

$$\lambda(u) = e^u\,\lambda(-u),$$

we end up this time with the second majoration

$$|\check{B}_X(s,f)| \leq \int_0^{+\infty} |X(\xi)|^2\, h\!\left(\frac{\xi}{f}\right) d\xi, \tag{69}$$

with

$$h(x) := g(x)\,e^{-\lambda^{-1}(x)}. \tag{70}$$

In order to study the behaviour of $g(x)$ and $h(x)$ (which are both plotted in Figure 3), we can first rewrite $g(x)$ as

$$g(x) = \frac{\lambda(\lambda^{-1}(x))}{\lambda'(\lambda^{-1}(x))}$$

and remark that, $\lambda : I\!\!R \to I\!\!R_+$ being a strictly increasing function [7], it suffices to restrict the study to the function

$$G(u) := \frac{\lambda'(u)}{\lambda(u)} = \frac{1}{g(\lambda(u))}. \tag{71}$$

This auxiliary function $G(u)$ is nothing but the logarithmic derivative of $\lambda(u)$ and, together with the explicit form (42), we get

$$G(u) = \frac{1}{u} - \frac{1}{e^u - 1} \tag{72}$$

28

and

$$G'(u) = \frac{1}{4 \, \sinh^2(u/2)} - \frac{1}{u^2} \le 0, \tag{73}$$

and it thus follows from (71) that $g(x)$ is an increasing function.

Considering then the function $h(x)$ defined in (70), it is again enough to restrict the study to the auxiliary function

$$H(u) := \frac{1}{G(u)} \, e^{-u},$$

whose derivative expresses as

$$H'(u) = -\frac{e^{-u}}{G^2(u)} \left[G'(u) + G(u) \right].$$

According to the explicit expressions (72) and (73), we have that $H'(u) \le 0$ and, as a consequence, $h(x)$ is a decreasing function. Given the definition (70), both functions intersect for $x = \lambda(0) = 1$, taking on the value $g(1) = h(1) = 1/\lambda'(0) = 2$. It follows from (67) and (69) that

$$\|\check{B}_X\|_\infty \le \sup_f \int_0^{+\infty} |X(\xi)|^2 \, \kappa\left(\frac{\xi}{f}\right) \, d\xi, \tag{74}$$

with

$$\kappa(x) := \min \{g(x), h(x)\} \le 2,$$

whence the claimed result.

Remark — According to (66), the unitary Bertrand distribution turns out to be bounded in exactly the same way as the Wigner distribution. In the narrowband limit, the integration in (74) is in fact restricted to $\xi \sim f$. It follows that $\kappa(\xi/f) \sim 2$, and the factor 2 in the upper bound is in accordance with the interpretation of the Wigner distribution as the limit form of the unitary Bertrand distributions in the case of narrowband signals [7]. In more general situations, Figure 3 suggests however that the majoration $\kappa(x) \le 2$ for all x's is a crude one and that sharper bounds should be attainable. In fact, while (66) gives a global bound, supposed to be valid for all frequencies, (74) is indeed frequency-dependent, and it follows from the form of $\kappa(x)$ that, for example,

$$\lim_{f \to 0} \sup_f |\check{B}_X(s, f)| = 0,$$

provided that $X(f) \in L^2(I\!R_+, df)$.

Based on the above Lemma 4, we can establish the more general result:

Lemma 5 *For any $p \geq 2$, the L^p-norm of the unitary Bertrand distribution of an analytic signal $X(f)$ is bounded according to*

$$\|B_X\|_p \leq 2^{1-\frac{2}{p}} E_+\left(|X|^2\right). \tag{75}$$

Proof[4] — Given any function $G(t, f) \in L^r(I\!R \times I\!R_+) \cap L^s(I\!R \times I\!R_+)$, with $1 \leq r \leq p \leq s < \infty$, we can use Hölder's inequality to show that

$$\|G\|_p \leq \|G\|_r^t \|G\|_s^u, \tag{76}$$

with

$$t = \frac{1 - s/p}{1 - s/r}, \qquad u = \frac{1 - r/p}{1 - r/s}.$$

This follows from the fact that p can be uniquely written as $p = \alpha r + (1 - \alpha)s$, with $\alpha \in [0, 1]$, so that Hölder's inequality gives

$$\begin{aligned}
\int_0^{+\infty} \int_{-\infty}^{+\infty} |G(t, f)|^p \, dt \, df &= \int_0^{+\infty} \int_{-\infty}^{+\infty} |G(t, f)|^{\alpha r} |G(t, f)|^{(1-\alpha)s} \, dt \, df \\
&\leq \left(\int_0^{+\infty} \int_{-\infty}^{+\infty} |G(t, f)|^r \, dt \, df\right)^{\alpha} \left(\int_0^{+\infty} \int_{-\infty}^{+\infty} |G(t, f)|^s \, dt \, df\right)^{1-\alpha},
\end{aligned}$$

whence the inequality (76).

For $r = 2$, $s \to \infty$ and $r \leq p \leq s$, we get

$$\|G\|_p \leq \|G\|_2^{\frac{2}{p}} \|G\|_\infty^{1-\frac{2}{p}}.$$

Applying this result to $G(t, f) = B_X(t, f)$ thus leads to the claimed inequality (75), knowing that $\|B_X\|_2^2 = E_+\left(|X|^2\right)$ and that $\|B_X\|_\infty \leq 2 E_+\left(|X|^2\right)$, from Lemma 4.

Based on Lemma 5, a bound can be obtained for the Shannon entropy of the unitary Bertrand distribution. This is given by the following

---

[4]Thanks to Guido Janssen for having suggested a simplification of the original proof and proposed the one reproduced here.

Proposition 8 *The squared unitary Bertrand distribution of a unit-energy analytic signal* $X(f)$ *has a Shannon entropy such that*

$$H_1\left(|\check{B}_X|^2\right) \geq \log \frac{1}{4}. \tag{77}$$

Proof — Thanks to the unitarity condition (51), the squared Bertrand distribution of a unit-energy signal integrates to one over the half-plane, so that we can consider its Shannon entropy as defined in (65).

Using (75) and proceeding as in the proof of Proposition 5, we can consider it as a function of p, differentiate both members with respect to p and evaluate at $p = 2$. Doing so, we get

$$
\begin{aligned}
\frac{\partial}{\partial p} \|B_X\|_p^p \Big|_{p=2} &= \int_{-\infty}^{+\infty} \int_0^{+\infty} |B_X(t,f)|^2 \log |B_X(t,f)| \, dt \, df \\
&= \int_{-\infty}^{+\infty} \int_0^{+\infty} |\check{B}_X(s,f)|^2 \log |\check{B}_X(s,f)| \, ds \, \frac{df}{f} \\
&= -\frac{1}{2} H_1\left(|\check{B}_X|^2\right) \\
&\leq \frac{\partial}{\partial p} \left(2^{p-2}\right)\Big|_{p=2} \\
&= \log 2,
\end{aligned}
$$

whence the claimed result.

Remark — Lemma 5 gives a partial (and weaker) counterpart to Lieb's results given in [24]. In fact, we know that in the narrowband limit, the Bertrand distribution tends to become identical to the Wigner distribution, for which Lieb's inequality [24]

$$\|W_X\|_p \leq \frac{2^{1-\frac{1}{p}}}{p^{\frac{1}{p}}} E_+\left(|X|^2\right)$$

guarantees that

$$H_1\left(|\check{W}_X|^2\right) \geq 1 + \log \frac{1}{4}.$$

4 Conclusion

Investigating limitations in scale-frequency signal analysis, we have considered specific measures of spread that proved well-adapted to the problem, and we have established a number of inequalities that can be attached to them. Those inequalities can be classified in two main groups. Some of them are a direct consequence of a warping equivalence with classical Fourier and time-frequency analyses. While interesting in giving explicit bounds, they mostly correspond to a distorted perspective of standard situations. More interestingly, other inequalities have been derived, that call for specific approaches. Those include a new form of uncertainty relation for the wavelet transform, as well as an entropy inequality for the Bertrand distribution. Some of the results presented here are still partial and preliminary, especially with respect to the sharpness of the bounds. Nevertheless, they are expected to give new insights in scale-frequency analysis, and in particular in its relation with time-frequency analysis, special emphasis having been put on narrowband limits for which specificities of scale-frequency analysis tend to fade out.

References

[1] Altes, R.A. (1976). Sonar for generalized target description and its similarity to animal echolocation systems. *J. Acoust. Soc. Am.* **59**, 97–105.

[2] Ash, R.B. (1965). *Information Theory*. Dover, New York, NY.

[3] Baraniuk, R.G. and Jones, D.L. (1995). Unitary equivalence: a new twist on signal processing. *IEEE Trans. on Signal Proc.* **43** (10), 2269–2282.

[4] Baraniuk, R.G., Flandrin, P., Janssen, A.J.E.M. and Michel, O. (1998). Measuring time-frequency information content using the Rényi entropies. Preprint available from `http://www-dsp.rice.edu/publications/pub/info98.ps.Z`.

[5] Beckner, W. (1975). Inequalities in Fourier analysis. *Ann. of Math.* **102**, 159–182.

[6] Bertrand, J. and Bertrand, P. (1988). Time-frequency representations of broad-band signals. *IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP–88*, New York, NY, 2196–2199.

[7] Bertrand, J. and Bertrand, P. (1992). A class of affine Wigner functions with extended covariance properties. *J. Math. Phys.* **33** (7), 2515–2527.

[8] Bertrand, J., Bertrand, P. and Ovarlez, J.P. (1996). The Mellin transform. *The Transforms and Applications Handbook* (A.D. Poularikas, ed.). CRC Press, Boca Raton, FL, 829–885.

[9] Boudreaux-Bartels, G.F. (1996). Mixed time-frequency signal transformations. *The Transforms and Applications Handbook* (A.D. Poularikas, ed.). CRC Press, Boca Raton, FL, 887–962.

[10] Cohen, L. (1995). *Time-Frequency Analysis*. Prentice-Hall, Englewoods Cliffs, NJ.

[11] Dahlke, S. and Maass, P. (1995). The affine uncertainty principle in one and two dimensions. *Computers Math. Applic.* **30** (3–6), 293–305.

[12] Daubechies, I. (1988). Time-frequency localization operators: a geometric phase-space approach. *IEEE Trans. on Info. Theory* **34**, 605–612.

[13] Daubechies, I. and Paul, Th. (1988). Time-frequency localization operators: a geometric phase-space approach — II. The use of dilations. *Inverse Problems* **4**, 661–680.

[14] de Bruijn, N.G. (1967). Uncertainty principles in Fourier analysis. *Inequalities* (O. Shisha, ed.). Academic Press, New York, NY, 57–71.

[15] Flandrin, P. (1988). Maximum signal energy concentration in a time-frequency domain. *IEEE Int. Conf. on Acoust., Speech and Signal Proc. ICASSP–88*, New York, NY, 2176–2179.

[16] Flandrin, P. (1998). Separability, positivity and minimum uncertainty in time-frequency energy distributions. *J. Math. Phys.* **39** (3), 4016–4040.

[17] Flandrin, P. (1998). *Time-Frequency/Time-Scale Analysis.* Academic Press, San Diego, CA.

[18] Folland, G.B. (1989). *Harmonic Analysis in Phase Space.* Princeton University Press, Princeton, NJ.

[19] Folland, G.B. and Sitaram, A. (1997). The uncertainty principle: a mathematical survey. *J. Fourier Anal. Appl.* **3** (3), 207–238.

[20] Hardy, G., Littlewood, J.E. and Pólya, G. (1934). *Inequalities.* Cambridge University Press, Cambridge.

[21] Hirschman, I.I. (1957). A note on entropy. *Amer. J. Math.* **79**, 152–156.

[22] Janseen, A.J.E.M. (1982). On the locus and spread of pseudo-density functions in the time-frequency plane. *Philips J. Res.* **37**, 79–110.

[23] Klauder, J.R. (1980). Path integrals for affine variables. *Functional Integration: Theory and Applications* (J.P. Antoine and E. Tirapegui, eds.). Plenum Press, New York, NY, 101–119.

[24] Lieb, E. (1990). Integral bounds for radar ambiguity functions and the Wigner distribution. *J. Math. Phys.* **31** (3), 594–599.

[25] Wilczok, E. (1997). Zur Funktionalanalysis der Wavelet– und der Gabor–transformation. PhD Thesis, Univ. Erlangen Nürnberg, Germany.

[26] Wilczok, E. (1998). New uncertainty principles for the continuous Gabor transform and the continuous wavelet transform. Preprint available from `http://www-m6.mathematik.tu-muenchen.de/~elke/Papers/uncertain.ps`.
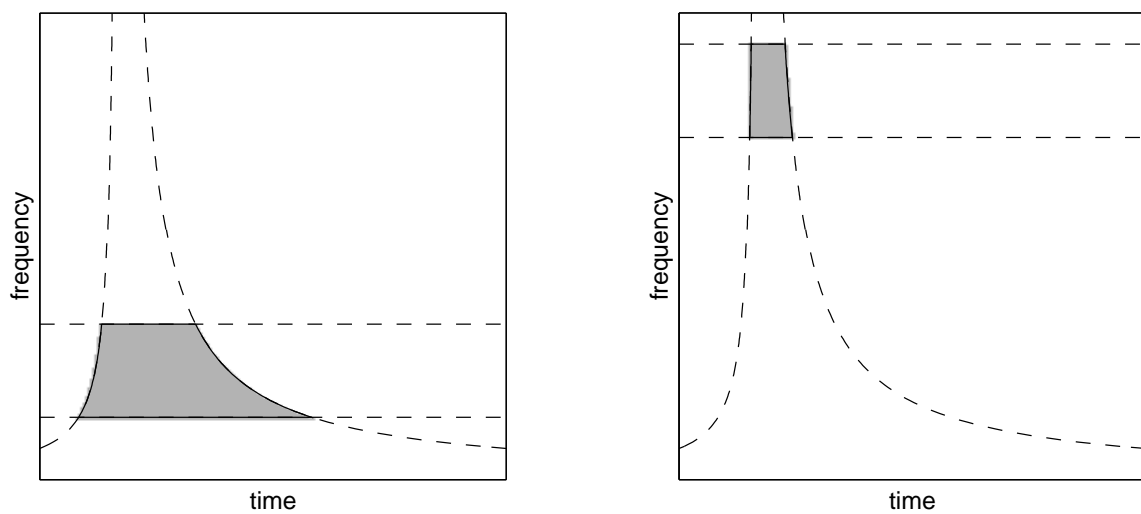
Figure 1: *Scale-frequency localization on the time-frequency plane — The narrowband limit.* Localization of a signal in frequency (respectively, in scale) amounts to concentrating most of its energy within a strip of the plane (respectively, a domain limited by two hyperbolae). The gray shaded areas represent (symbolically) time-frequency domains of simultaneous localization in scale and frequency, for both wideband (left) and narrowband (right) signals. In the narrowband limit, classical "rectangular" domains of Fourier analysis are recovered.
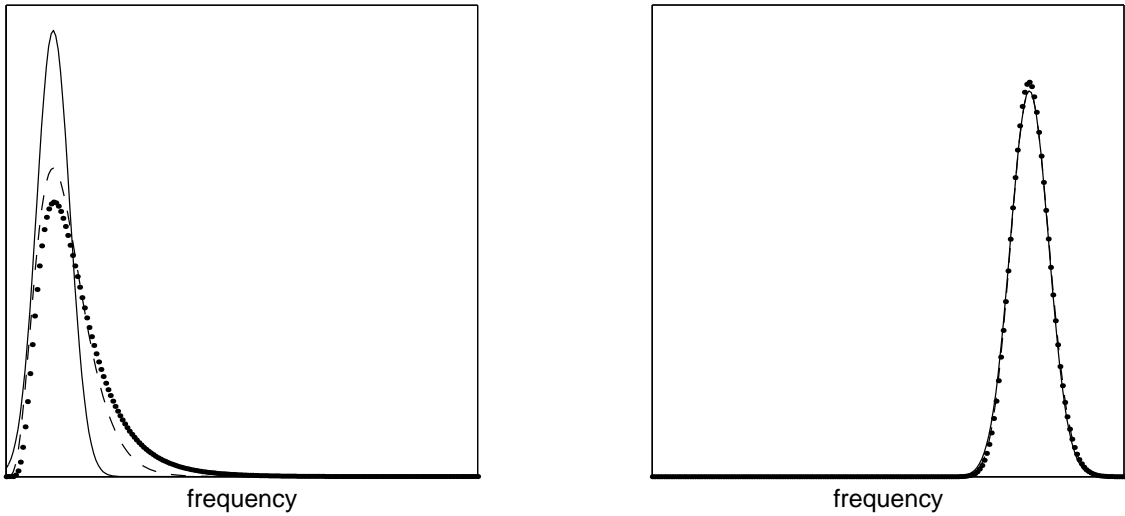
Figure 2: *Most localized signals in scale and frequency — The narrowband limit.* Depending on whether frequency localization is evaluated according to arithmetic or geometric measures of spread, signals which are most localized in both scale and frequency are either Klauder wavelets (dashed lines) or Altes wavelets (dotted lines), respectively. In the narrowband limit, they both become undistinguishable from Morlet wavelets (solid line). In the left diagram (wideband case), the Morlet wavelet correspond to (14) with $a = 2$ and $\log b = 1$, the Klauder wavelet to (16) with $a = 2$ and $b = 2$, and the Altes wavelet to (22) with $a = 1$ and $b = 1.3$. In the right diagram (narrowbandband case), the Morlet wavelet correspond to (14) with $a = 1.5$ and $\log b = 8$, the Klauder wavelet to (16) with $a = 192$ and $b = 24$, and the Altes wavelet to (22) with $a = 100$ and $b = 8$. In all cases, energy densities are plotted, normalized to be unit energy.
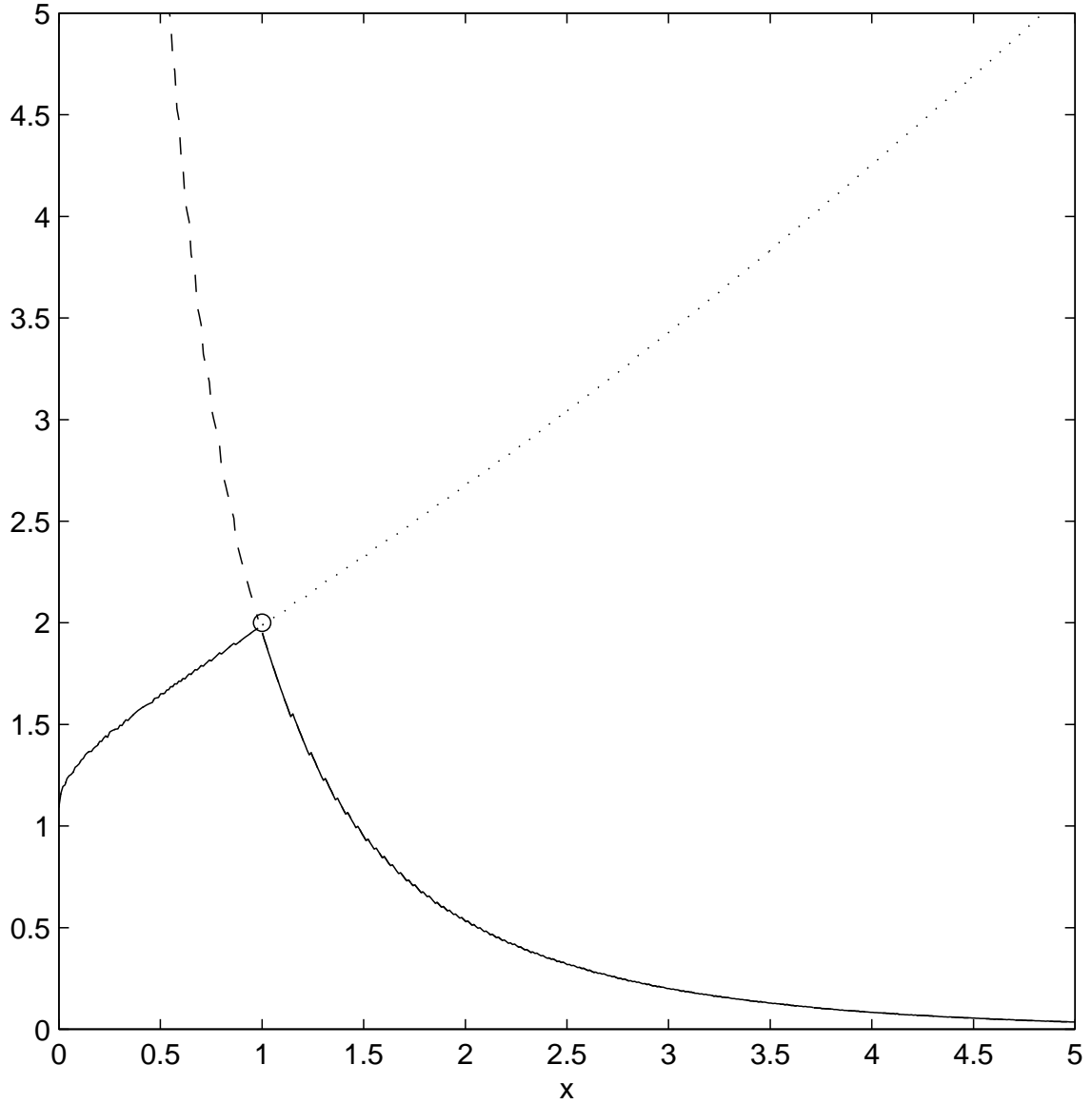
Figure 3: *Upper bound for the unitary Bertrand distribution.* Evaluating an upper bound for absolute values of the unitary Bertrand distribution involves the two functions $g(x)$ (dotted line) and $h(x)$ (dashed line), defined in (68) and (70), respectively. Evaluation of the bound is actually based on the auxiliary function $\kappa(x) = \min\{g(x), h(x)\}$ (superimposed solid line), which attains its maximum at the (circled) point $(1, 2)$, thus allowing to guarantee that the modulus of a Bertrand distribution never exceeds twice the energy of the signal on which it is constructed (see text).