Uncertainty and Spectrogram Geometry

Patrick Flandrin*

CNRS & École Normale Supérieure de Lyon, France

Erwin Schrödinger Institute, December 2012

* based on joint work with François Auger and Éric Chassande-Mottin

Heisenberg (classical)

Second-order (variance-type) measures for the **individual time and frequency spreadings** of a signal $x(t) \in L^2(\mathbb{R})$ with spectrum $X(\omega)$:

$$\Delta_t^2(x) := \frac{1}{E_x} \int t^2 \, |x(t)|^2 \, dt$$

$$\Delta_\omega^2(X) := \frac{1}{E_x} \int \omega^2 \, |X(\omega)|^2 \, \frac{d\omega}{2\pi}$$

Theorem (Weyl, ’27; Gabor, ’46, ...)

$$\Delta_t(x) \, \Delta_\omega(X) \geq \frac{1}{2},$$

with lower bound attained for Gaussians $x_*(t) = C \, e^{\alpha \, t^2}$, $\alpha < 0$.

Heisenberg (time-frequency)

Second-order (variance-type) measure for the **joint time-frequency spreading** of a signal $x(t) \in L^2(\mathbb{R})$ with energy distribution $C_x(t,\omega;\varphi)$ within Cohen's class:

$$\Delta_{t\omega}(C_x) := \frac{1}{E_x} \iint \left(\frac{t^2}{T^2} + T^2\omega^2 \right) C_x(t,\omega;\varphi)\, dt\, \frac{d\omega}{2\pi}$$

Theorem (Janssen, '91)

$$\begin{aligned} \textit{Wigner} &\Rightarrow \Delta_{t\omega}(W_x) \geq 1 \\ \textit{Spectrogram} &\Rightarrow \Delta_{t\omega}(S_x^h) \geq 2, \end{aligned}$$

with lower bound attained for Gaussian signals and matched Gaussian windows.

time-frequency "covariance"

Definition (Cohen, '95)

$$c(x) := \int t\,|x(t)|^2\,\frac{d}{dt}\arg x(t)\,dt$$

1. **Interpretation**: covariance quantifies the coupling between time and instantaneous frequency

$$c(x) = \langle t\,\omega_x(t)\rangle$$

2. **Intuition**: covariance is zero in case of no coupling

$$c(x) = \langle t\,\omega_x(t)\rangle = \langle t\rangle\,\langle \omega_x(t)\rangle = \langle t\rangle\,\langle\omega\rangle = 0$$

Schrödinger

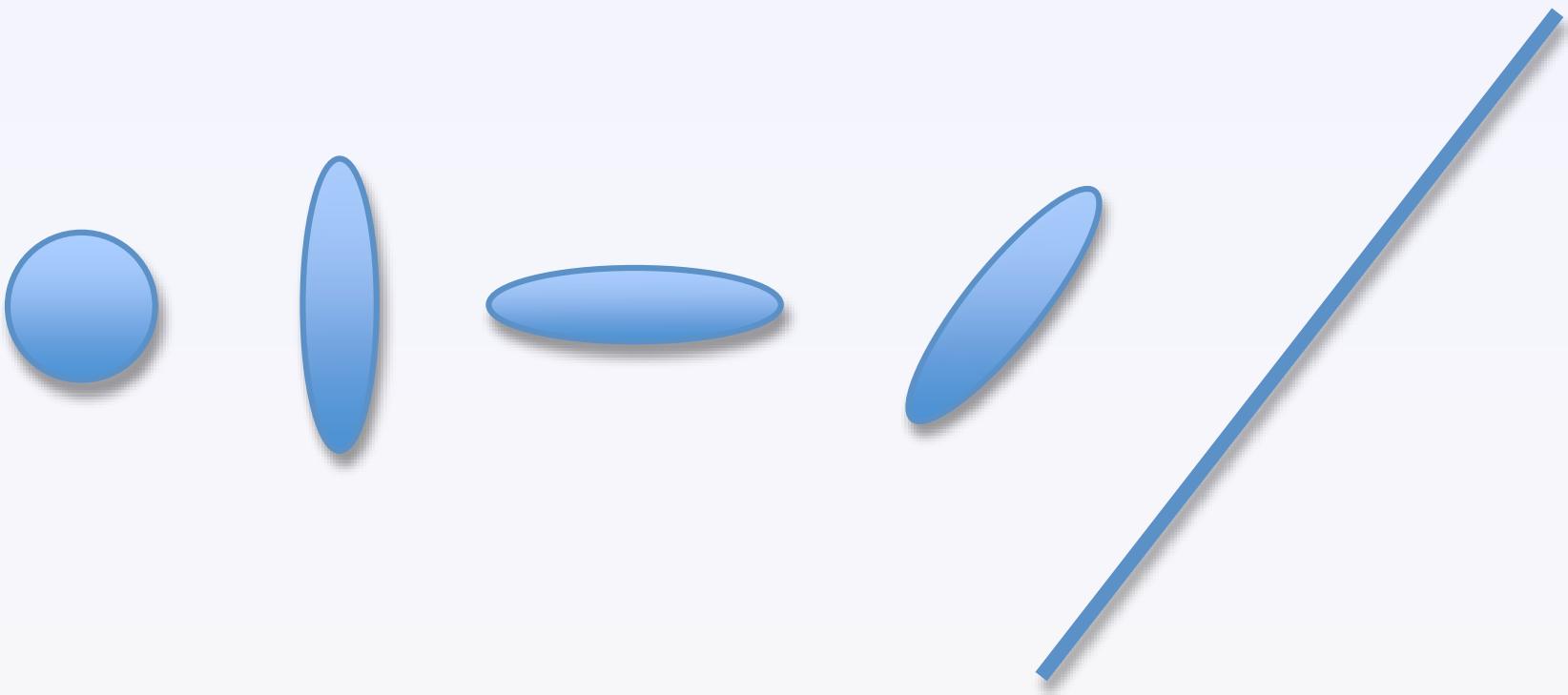
Theorem (Schrödinger, '30)

$$\Delta_t(x)\,\Delta_\omega(X) \geq \frac{1}{2}\sqrt{1+c^2(x)},$$

with lower bound attained for Gaussians $x_*(t) = e^{\alpha\,t^2+\beta\,t+\gamma}$, $Re\{\alpha\} < 0$.

1. **Terminology**: waveforms x_* attaining the lower bound correspond to "squeezed states" in quantum mechanics and "linear chirps" in signal theory
2. **Interpretation**: possibility of localization in the plane beyond pointwise energy concentration

“chirp” localization



how?

1. **Effective localization**: in the Wigner case, we have

$$\lim_{Re\{\alpha\}\to 0_-} W_{x_*}(t,\omega) = \delta\left(\omega - (\beta + 2\,Im\{\alpha\}t)\right)$$

2. **Generalization**: localization on more "arbitrary" curves of the plane by modifying the symmetry rules underlying the Wigner distribution (Gonçalvès and F., '96)
3. ***Caveat***: global localization holds for monocomponent signals only
4. **Way out**: reassignment (Kodera *et al.*, '76)

rationale

Starting with the smoothing relationship

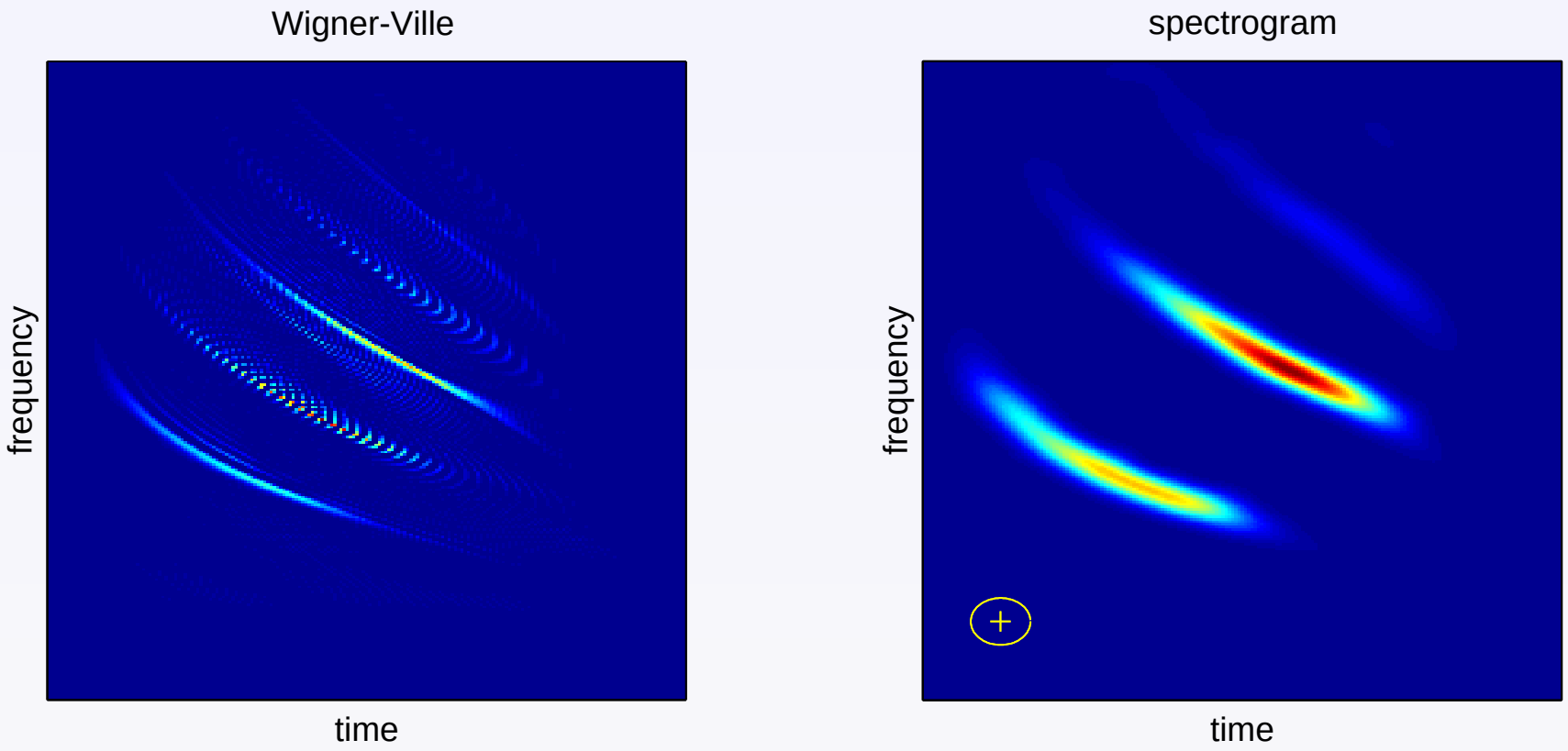
$$S_x^h(t,\omega) = \iint W_x(s,\xi)\, W_h(s-t,\xi-\omega)\, dt\, \frac{d\omega}{2\pi},$$

the key idea is

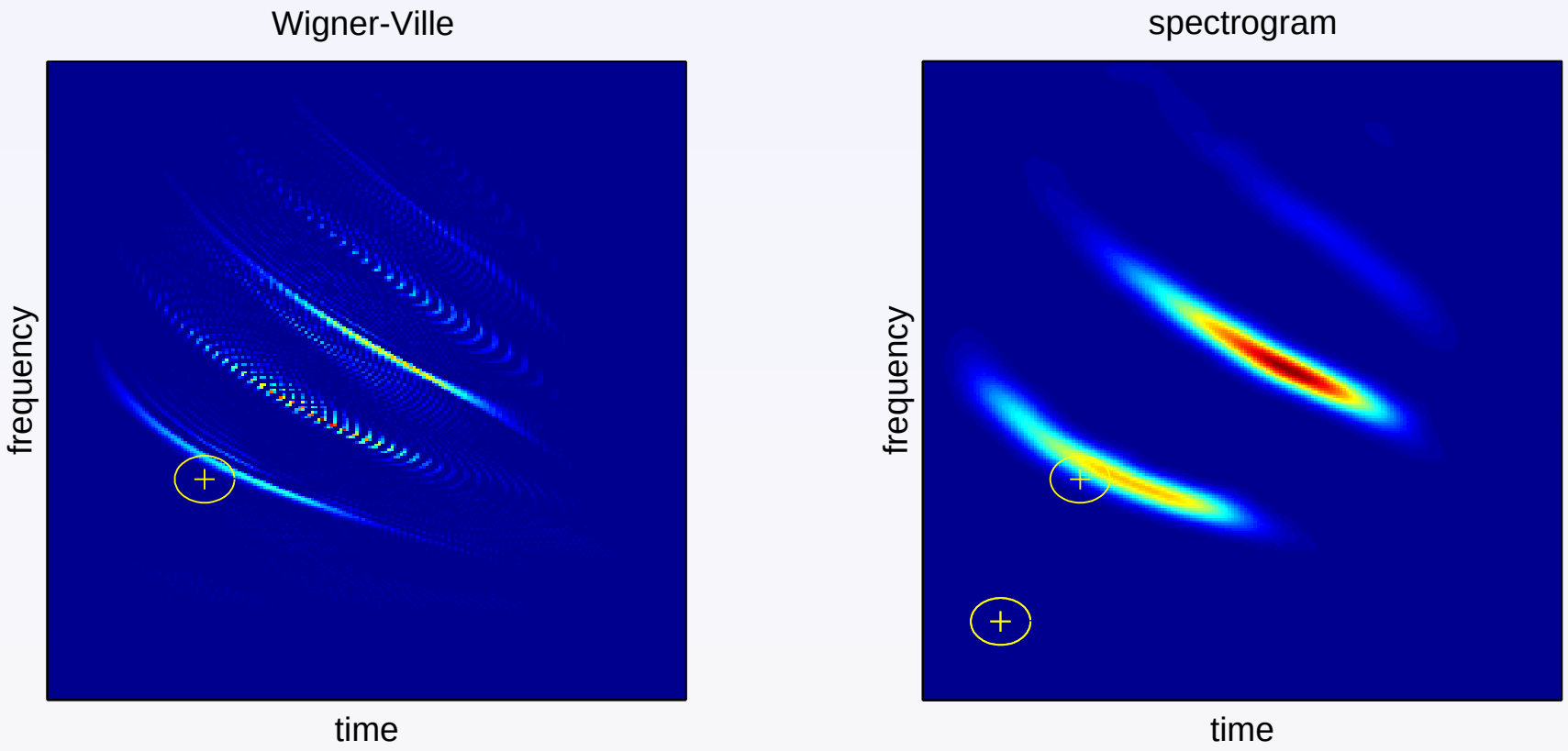
1. to replace the **geometrical** center of the smoothing time-frequency domain by the **center of mass** of the Wigner distribution over this domain, and
2. to **reassign** the value of the smoothed distribution to this local centroïd:

$$\hat{S}_x^h(t,\omega) = \iint S_x^h(\tau,\xi)\delta\left(t-\hat{t}_x(\tau,\xi), \omega-\hat{\omega}_x(\tau,\xi)\right) d\tau \frac{d\xi}{2\pi}$$

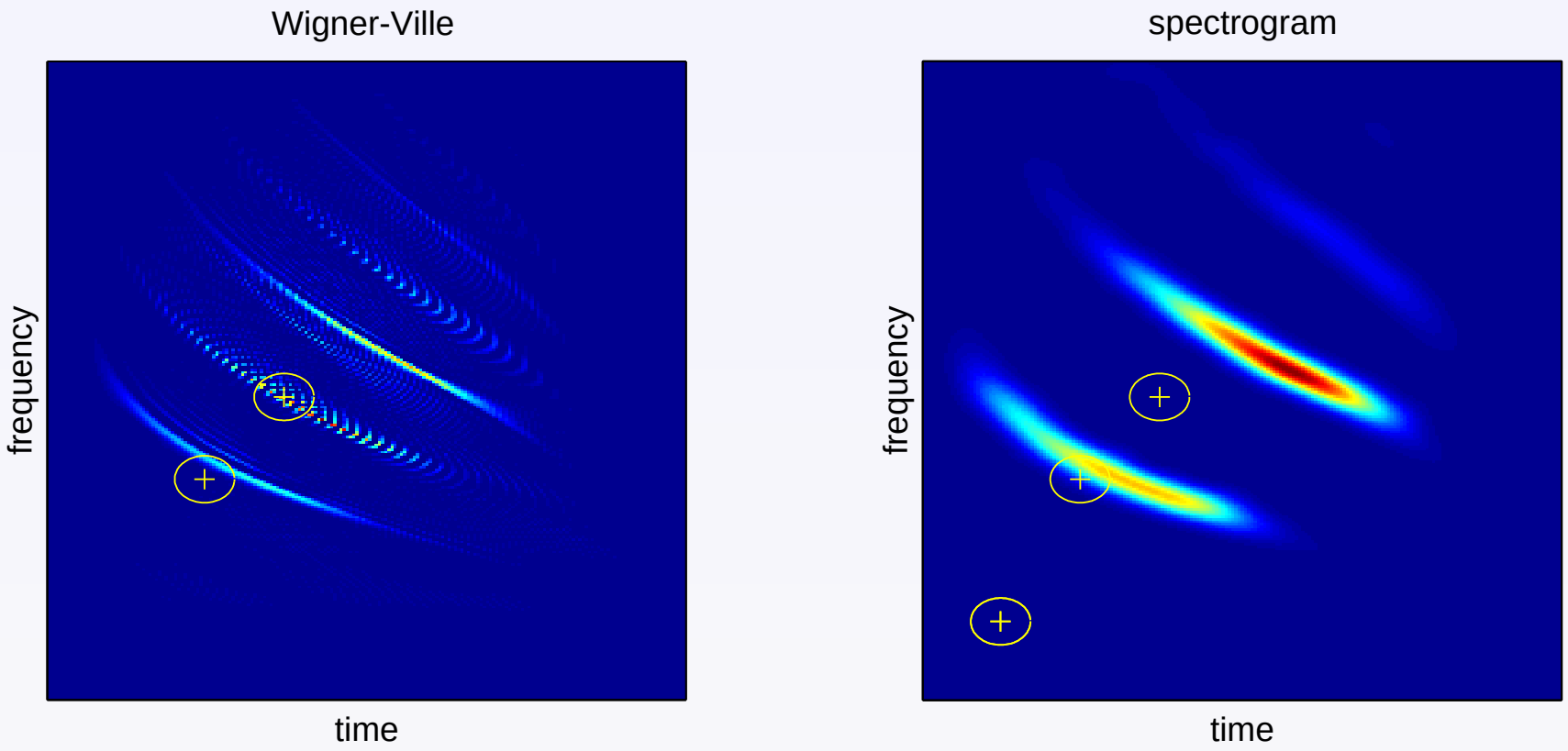
spectrogram = smoothed Wigner



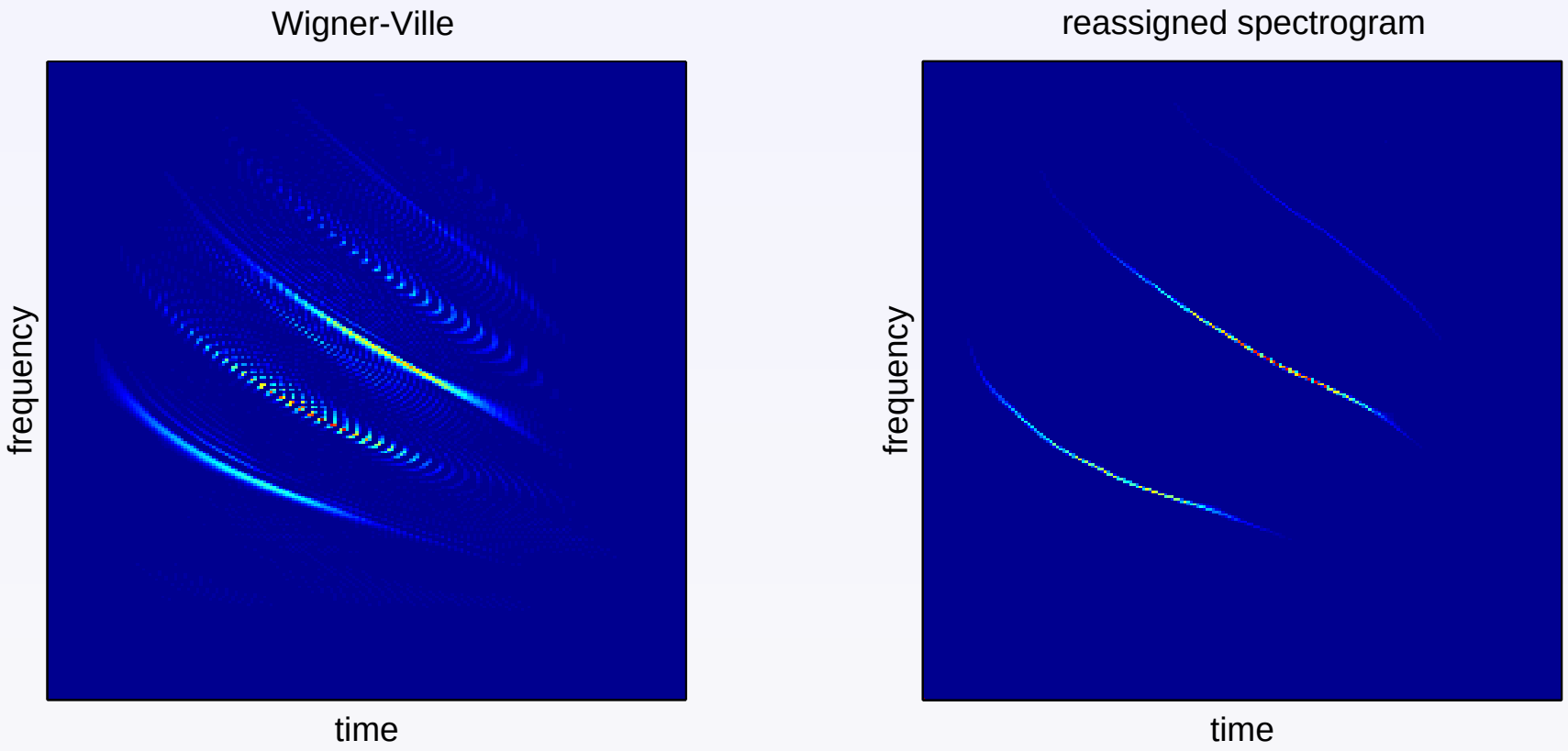
spreading of auto-terms



cancelling of cross-terms



reassignment (Kodera *et al.*, '76, Auger & F., '95)



reassignment in action

1. **Spectrogram**: implicit computation of the local centroïds (Auger & F., '95) via two additional short-time Fourier transforms (STFTs) with the companion windows $(\mathcal{T}h)(t) = t\,h(t)$ and $(\mathcal{D}h)(t) = (dh/dt)(t)$
2. **Gaussian spectrogram**: alternative computation of the local centroïds via partial derivatives of the magnitude of the STFT (Auger, Chassande-Mottin & F., '12)
3. **Beyond spectrograms**: possible generalizations to other smoothings (smoothed pseudo-Wigner-Ville, scalogram, etc.)

linear chirps

- **Framework**: Gaussian window $h(t) = \pi^{-1/4}\, e^{-t^2/2}$
- **Result**: the linear chirp parameterized by

$$\alpha = -\frac{1}{2}\left(\frac{1}{T^2} - i\,a\right) \quad ; \quad \beta = 0$$

is such that

$$\lim_{T\to\infty} \hat{S}_x^h(t,\omega) = \frac{1}{2\pi}\delta(\omega - at)$$

- **Interpretation**: perfect localization

Hermite functions

- **Framework**: Gaussian window $g(t) = \pi^{-1/4}\, e^{-t^2/2}$
- **Hermite functions**:

$$h_k(t) = C_{k-1} \frac{1}{\sqrt{T}} H_{k-1}(\sqrt{2\pi}\, t/T) e^{-\pi (t/T)^2},$$

with

$$H_n(\alpha) = (-1)^n e^{\alpha^2} \frac{d^n}{d\alpha^n} e^{-\alpha^2}, \qquad n = 0, 1$$

- **Radial reassignment**:

$$\frac{\hat{\omega}(t,\omega)}{\hat{t}(t,\omega)} = \frac{\omega}{t}$$

Hermite functions

- **Result** (F., *JFAA* '12):

$$\hat{S}_{h_k}^{(g)}(t,\omega) = \frac{1}{2^{k-1}(k-1)!}\,\hat{V}\left(\sqrt{t^2+\omega^2}\right)\mathbf{1}_{\mathcal{D}}(t,\omega),$$

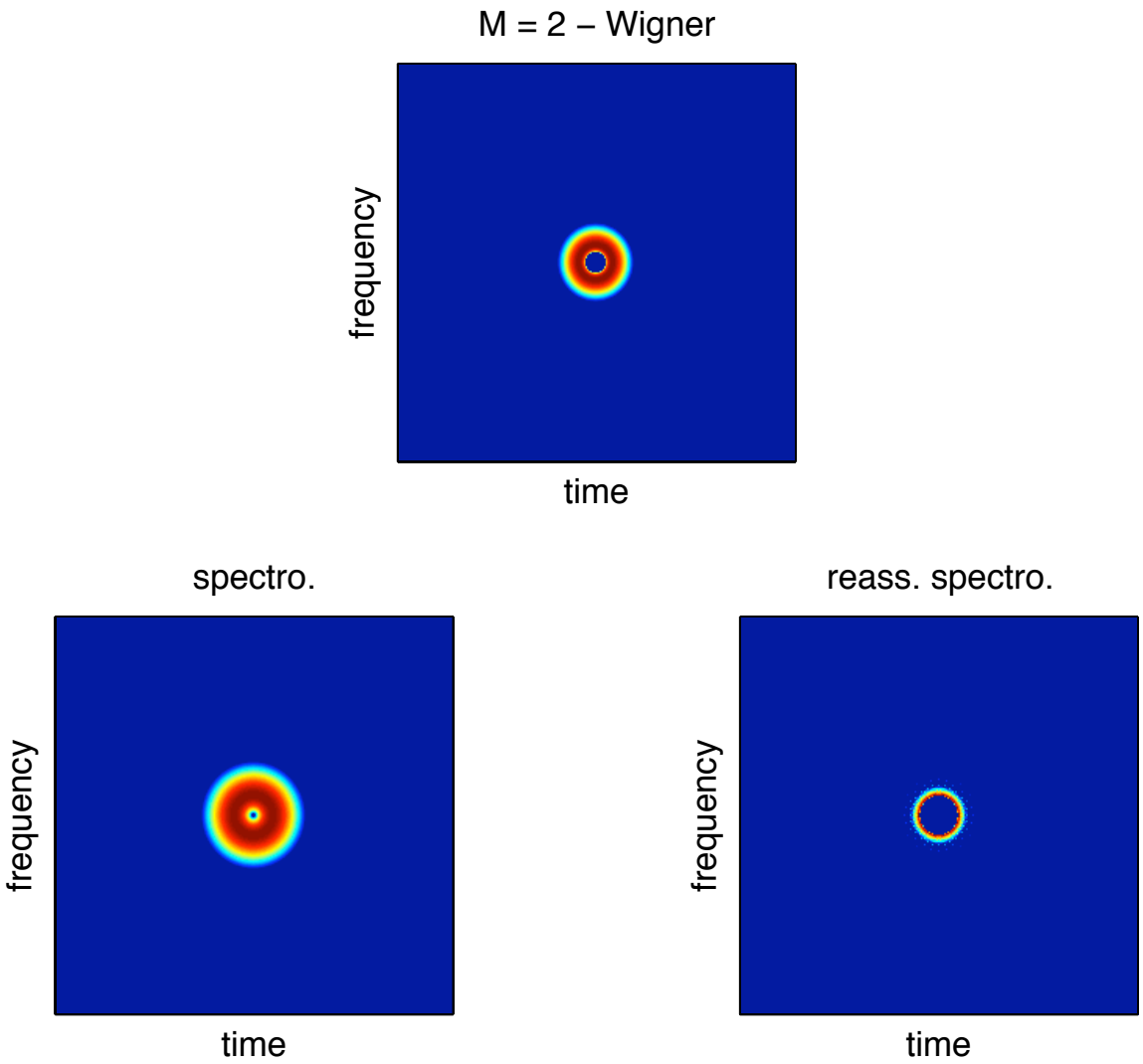
with $\hat{V} \neq \delta$ and:

$$\mathcal{D} = \{(t,\omega) \mid t^2+\omega^2 \geq t_m^2+\omega_m^2 = 2(k-1)\}$$

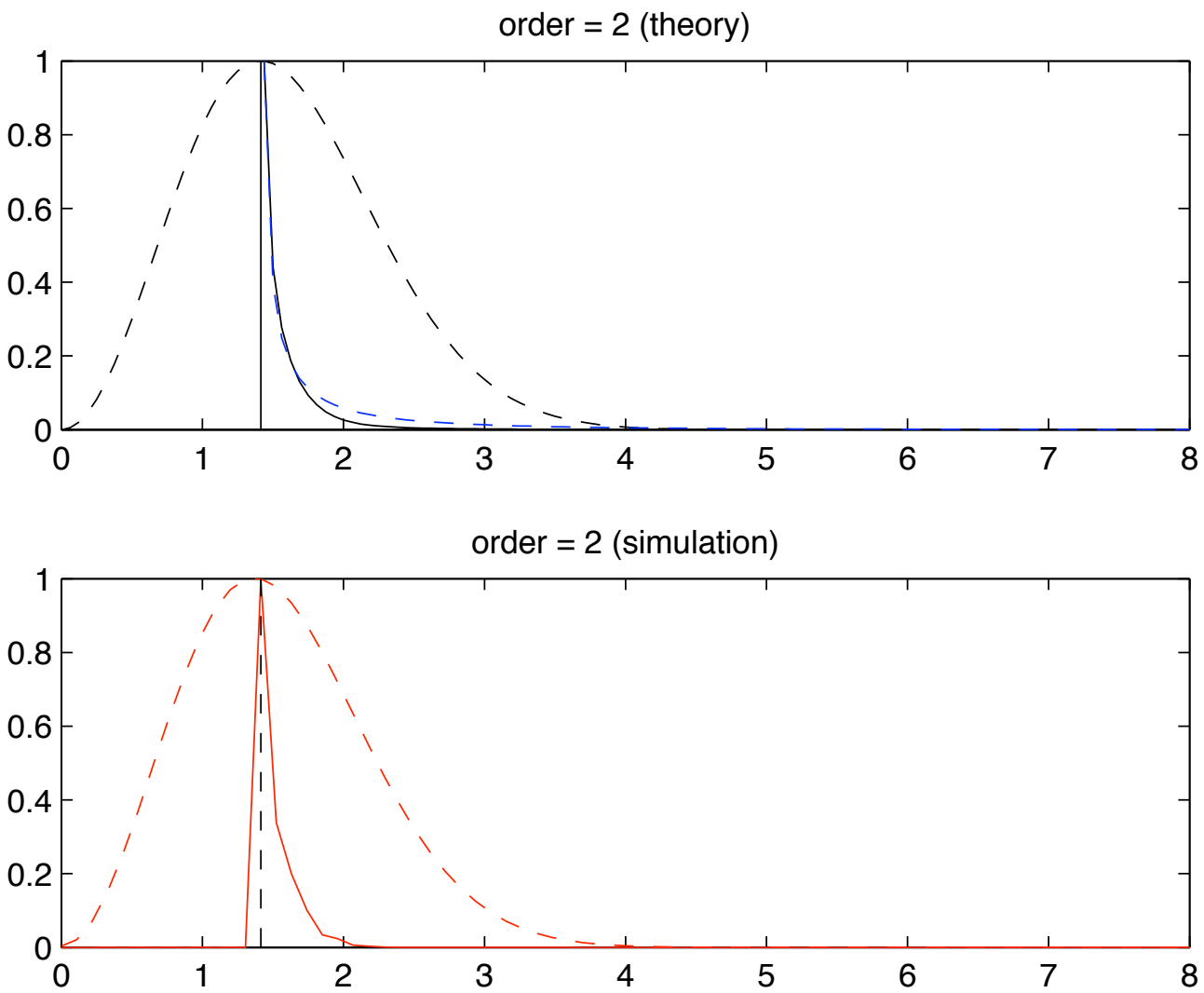
- **Interpretation**: clearance area and no perfect localization
- **Asymptotic localization**:

$$\hat{S}_{h_k}^{(g)}(t,\omega) \overset{k\to\infty}{\longrightarrow} \delta(t^2+\omega^2-2(k-1))$$

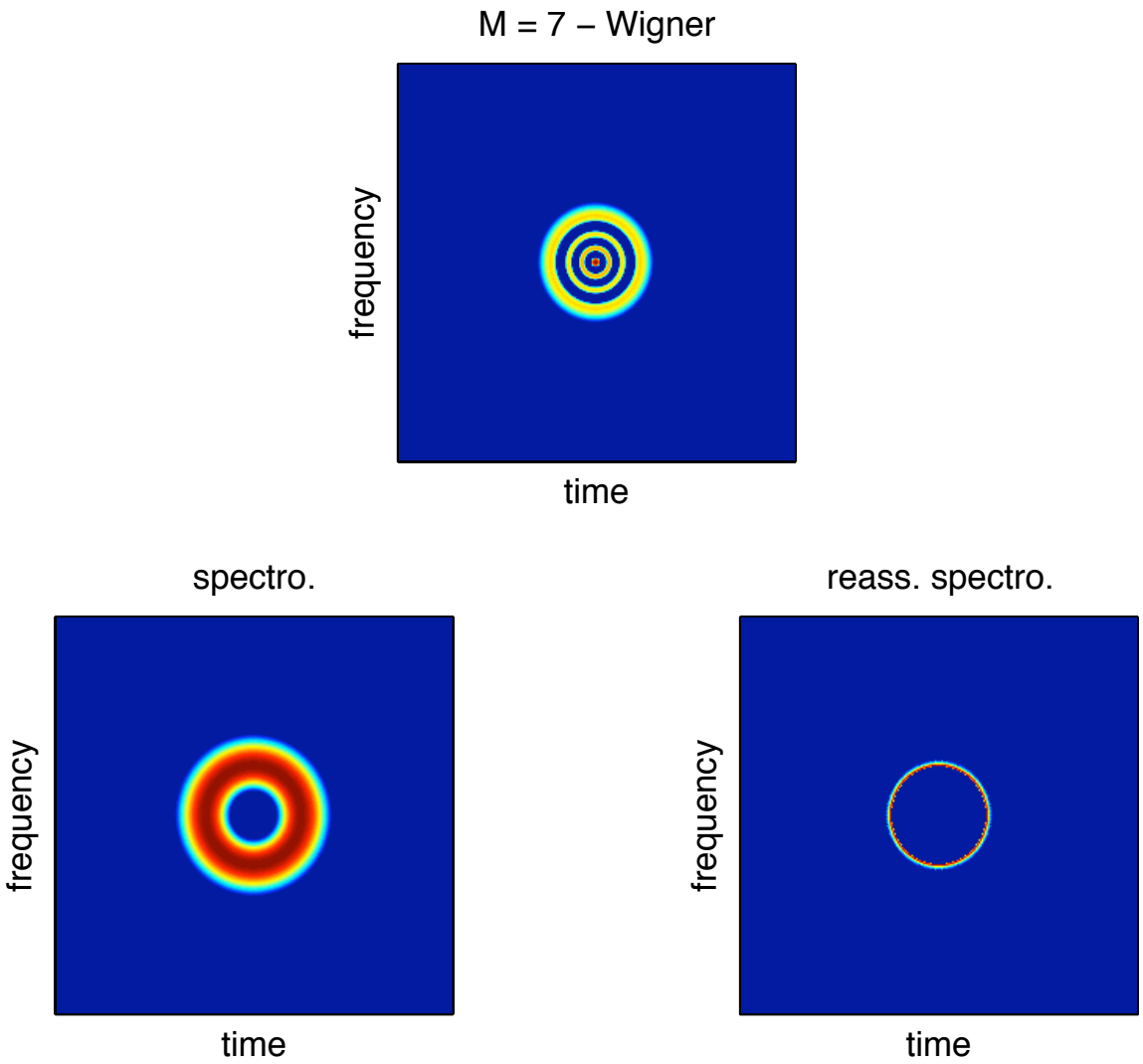
Hermite functions



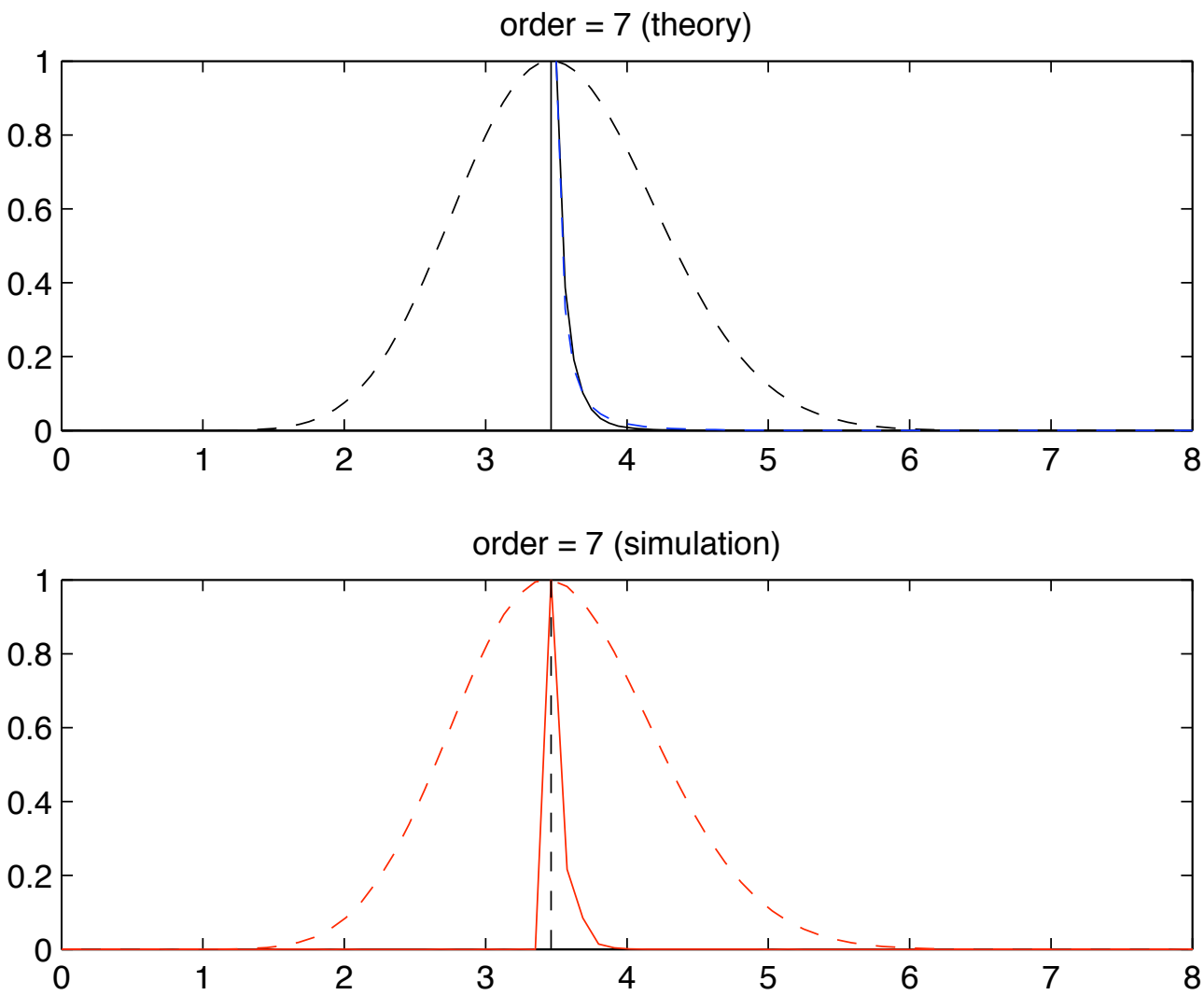
Hermite functions



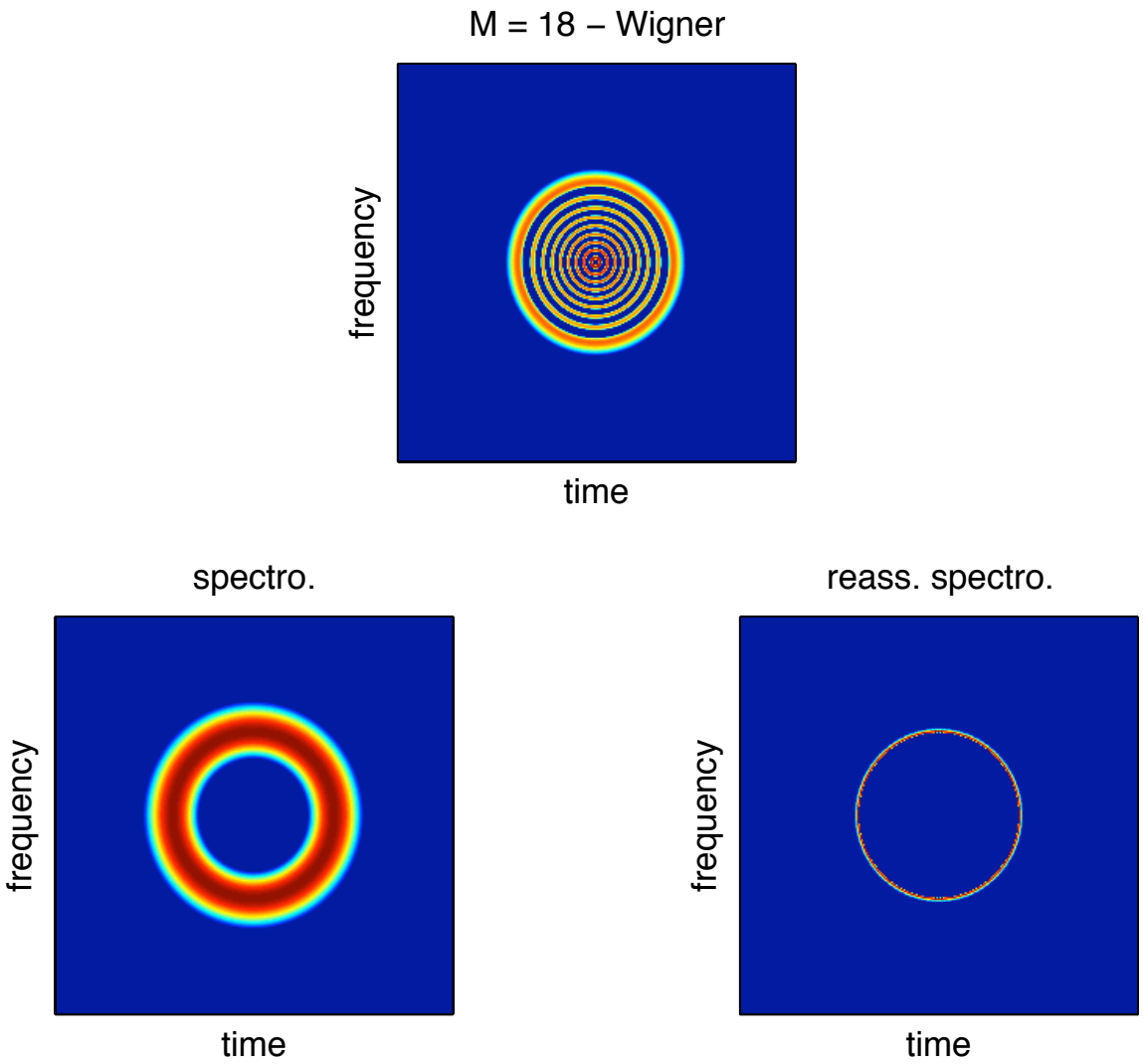
Hermite functions



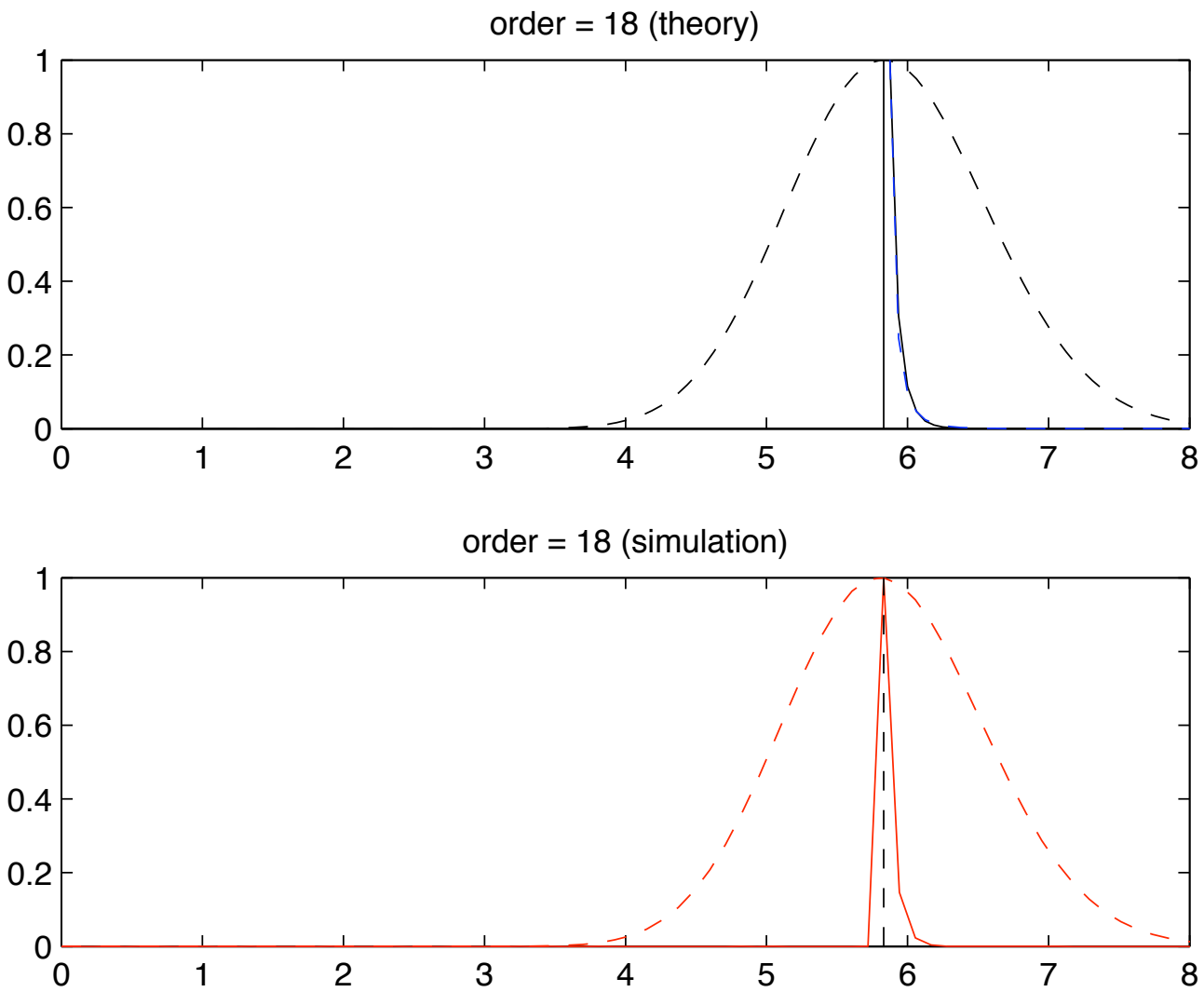
Hermite functions



Hermite functions



Hermite functions



logons

- **Framework**: Gaussian window $h(t) = \pi^{-1/4}\, e^{-t^2/2}$
- **Result**: the Gabor "logon", i.e., the previous linear chirp with $a = 0$, is such that

$$S_x^h(t,\omega) = e^{-\frac{1}{2}(t^2+\omega^2)} \Rightarrow \Delta_{t\omega}(S_x^h) = 2$$

$$W_x(t,\omega) = 2\, e^{-(t^2+\omega^2)} \Rightarrow \Delta_{t\omega}(W_x) = 1$$

$$\hat{S}_x^h(t,\omega) = 4\, e^{-2(t^2+\omega^2)} \Rightarrow \Delta_{t\omega}(\hat{S}_x^h) = \frac{1}{2}$$

- **Interpretation**: Heisenberg defeated?

resolving the paradox

1. **General remark**: as for Fourier, sharp localization not to be confused with resolution (i.e., ability to separate closely spaced components)
2. **Complete picture**: reassignment = squeezed distribution + vector field $\mathbf{r}_x(t,\omega) = (\hat{t}_x(t,\omega) - t, \hat{\omega}_x(t,\omega) - \omega)^t$ such that

$$\mathbf{r}_x(t,\omega) = \frac{1}{2}\nabla \log S_x^h(t,\omega)$$

3. **Interpretation**: basins of attraction
4. **Open question**: geometry of such basins and relationship with Heisenberg?

factorization

- **Bargmann representation**: with a "circular" Gaussian window, the STFT can be expressed as

$$F_x^h(z) = \mathcal{F}_x^h(z)\, e^{-|z|^2/4},$$

where $z = \omega + \jmath t$ and $\mathcal{F}_x^h(z)$ is an entire function of order at most 2

- **Weierstrass-Hadamard factorization**: it follows that

$$\mathcal{F}_x^h(z) = e^{Q(z)} \prod_n (1 - \tilde{z}_n) \exp\left(\tilde{z}_n + \tilde{z}_n^2/2\right),$$

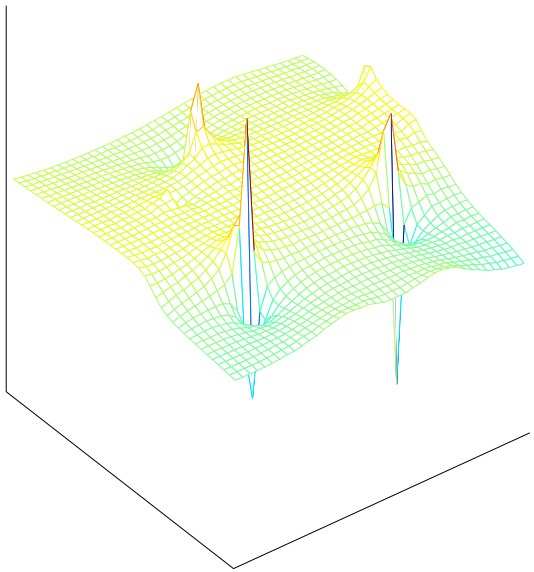
where $Q(z)$ is a quadratic polynomial and $\tilde{z}_n = z/z_n$

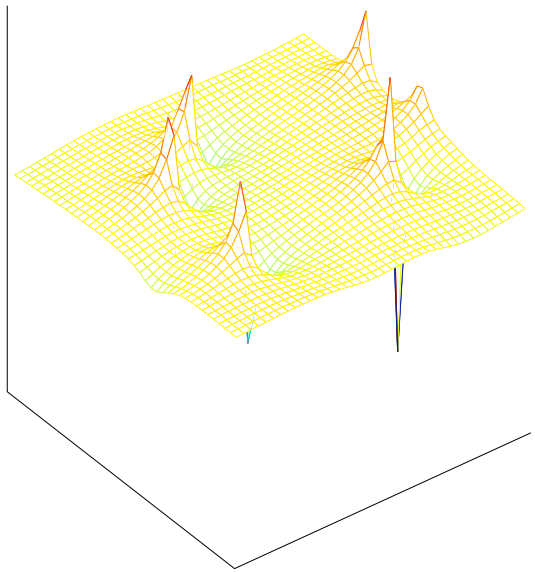
- **Interpretation**: complete characterization by zeroes and, by duality, by local maxima

more on zeroes

Result (Balasz *et al.*, '11)

"Universal" singularity of the phase derivative around zeroes of the STFT





phase-magnitude relationship

In the Gaussian case, simple explanation based on:

Proposition (Auger *et al.*, '12)

Phase and magnitude of a Gaussian STFT form a Cauchy pair, i.e.,

$$\begin{aligned}\frac{\partial \Phi_x^h}{\partial t}(t,\omega) &= \frac{\partial}{\partial \omega} \log \mathcal{M}_x^h(t,\omega) \\ \frac{\partial \Phi_x^h}{\partial \omega}(t,\omega) &= -\frac{\partial}{\partial t} \log \mathcal{M}_x^h(t,\omega),\end{aligned}$$

with $F_x^h(t,\omega) = \mathcal{M}_x^h(t,\omega)\ e^{\jmath \Phi_x^h(t,\omega)}\ e^{-(t^2+\omega^2)/4}$

phase singularity

1. **Magnitude**: local behaviour in the vicinity of a zero (t_0, ω_0)

$$\mathcal{M}_x^h(t,\omega) \propto \left[(\omega-\omega_0)^2 + (t-t_0)^2\right]^{1/2}$$

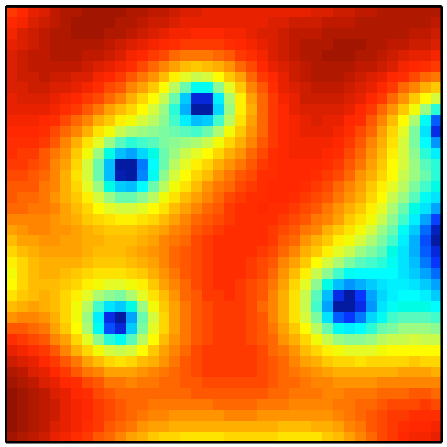
2. **Phase from magnitude**: "universal" hyperbolic divergence of the phase derivative

$$\left.\frac{\partial \Phi_x^h}{\partial t}(t_0,\omega)\right|_{\omega\sim\omega_0} \sim (\omega-\omega_0)^{-1}$$

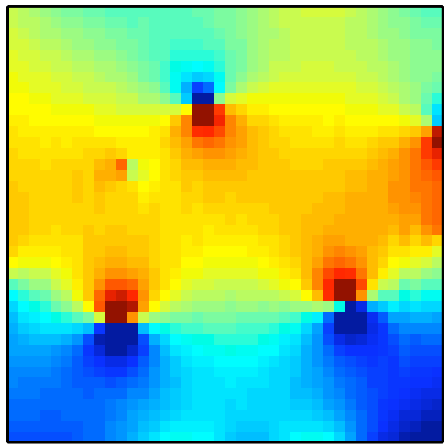
$$\left.\frac{\partial \Phi_x^h}{\partial \omega}(t,\omega_0)\right|_{t\sim t_0} \sim (t_0-t)^{-1}$$

zeroes $\leftrightarrow$ "phase dislocations" (Nye & Berry, '74)

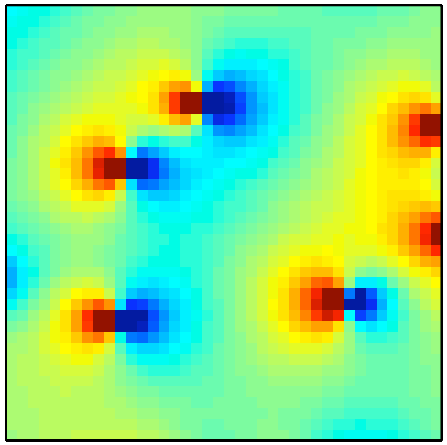
magnitude



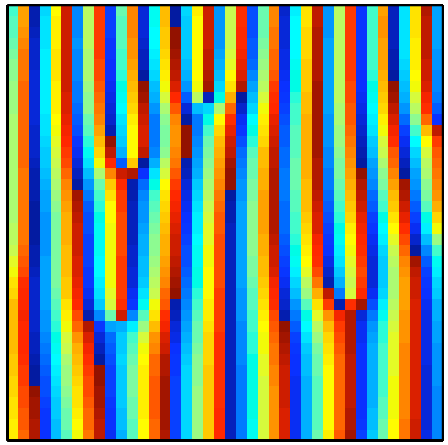
phase gradient in frequency



phase gradient in time



phase



phase dislocation (time domain)

phase dislocation (complex plane)

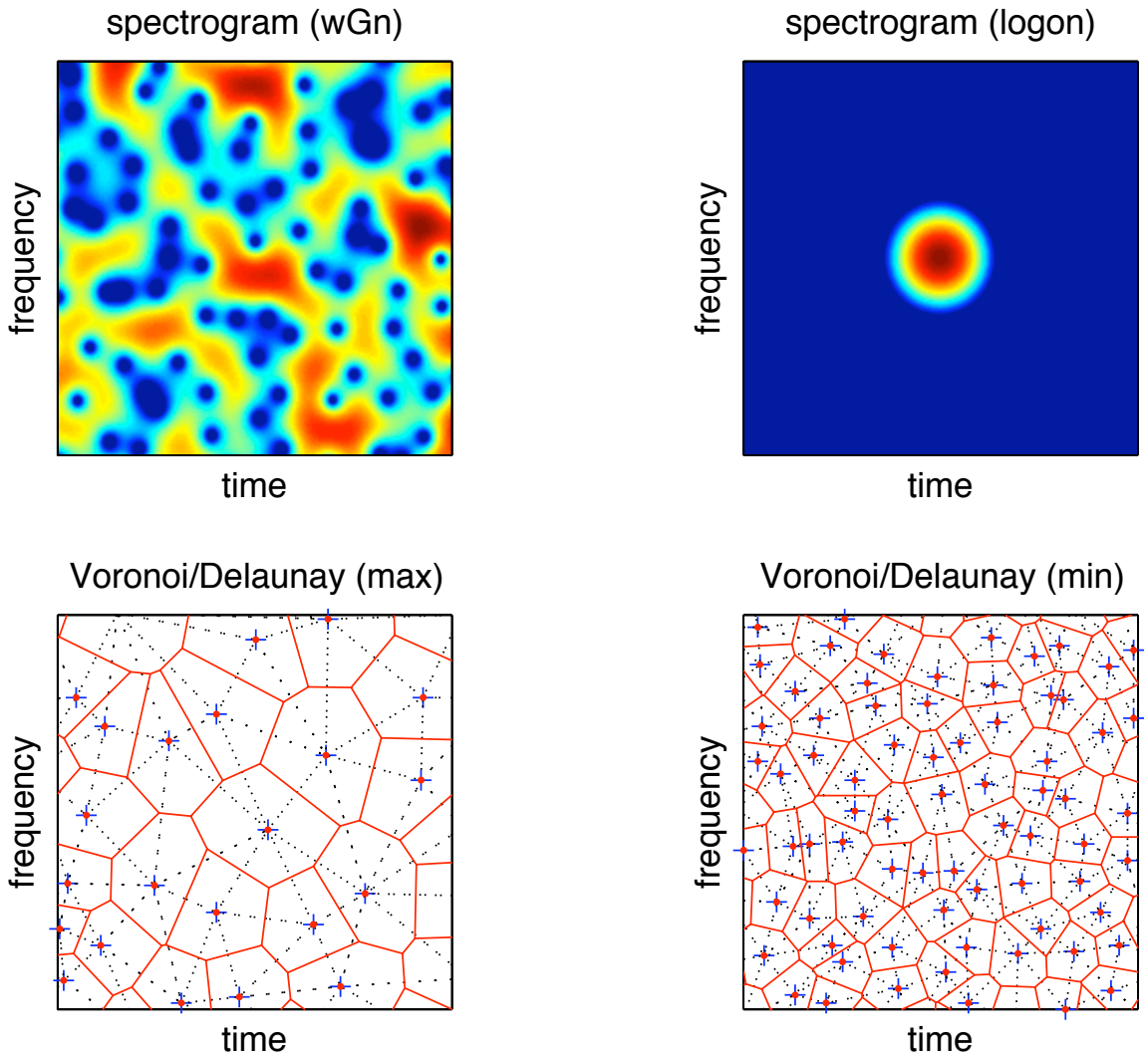
phase dislocation (zoom)

more on extrema

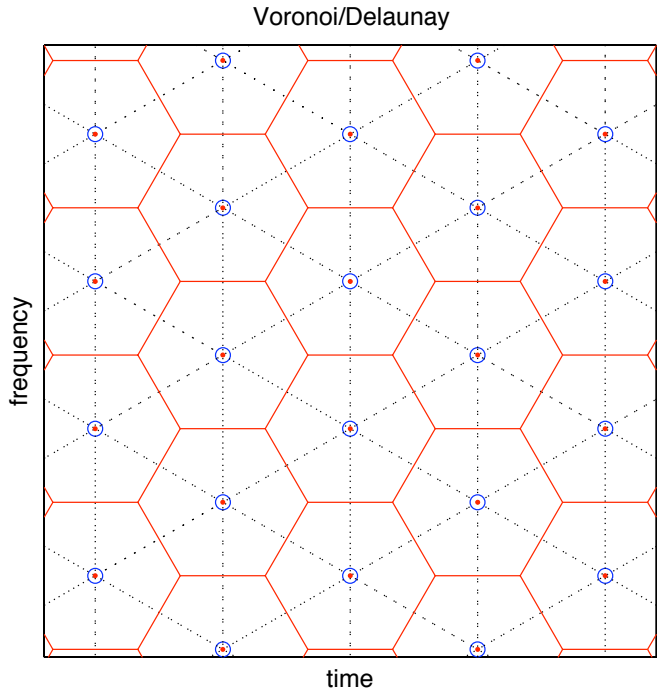
Proposed approach

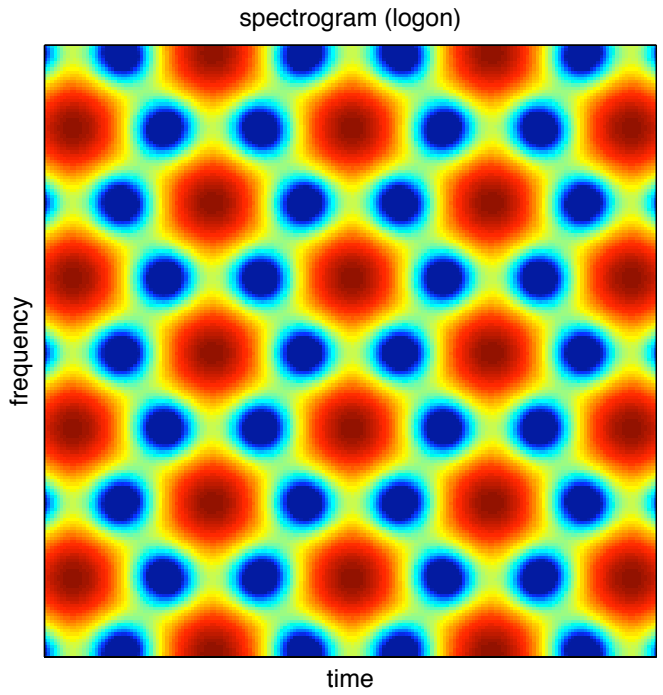
1. **Data**: white Gaussian noise
2. **Time-frequency representation**: Gaussian spectrogram
3. **Characterization**: identification of local extrema (zeroes and maxima) + Voronoi tessellation and Delaunay triangulation
4. **Analysis**: distribution of cell areas, lengths between extrema and heights of local maxima

an example


spectrogram (wGn)
frequency
time
spectrogram (logon)
frequency
time
Voronoi/Delaunay (max)
frequency
time
Voronoi/Delaunay (min)
frequency
time

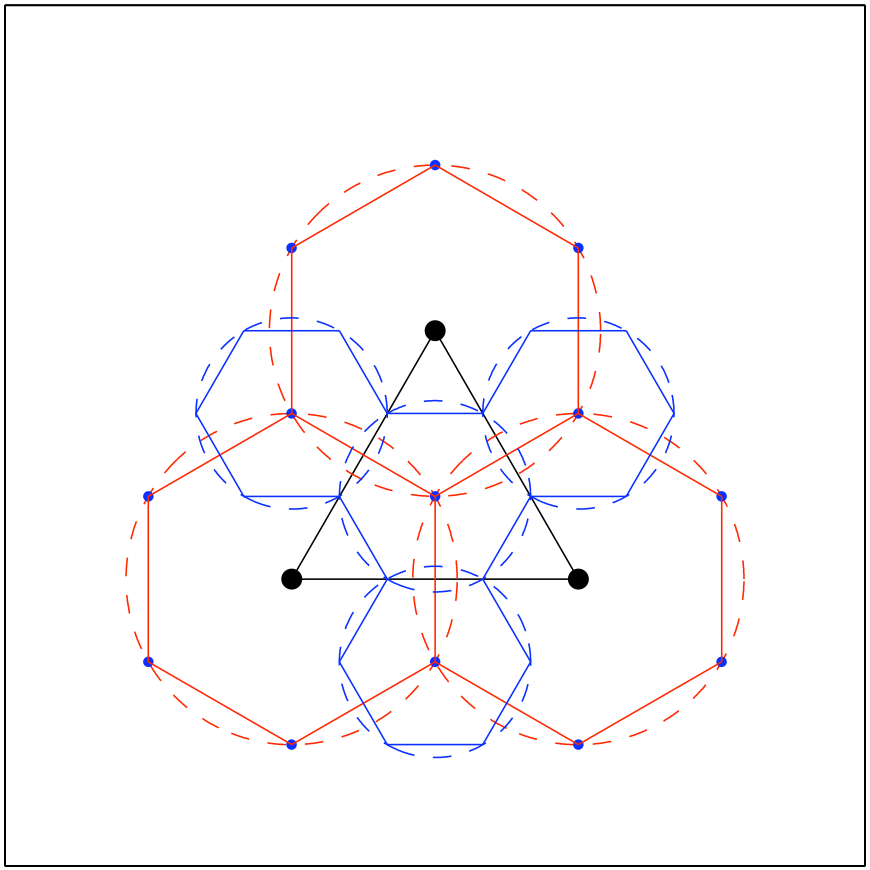
mean arrangement





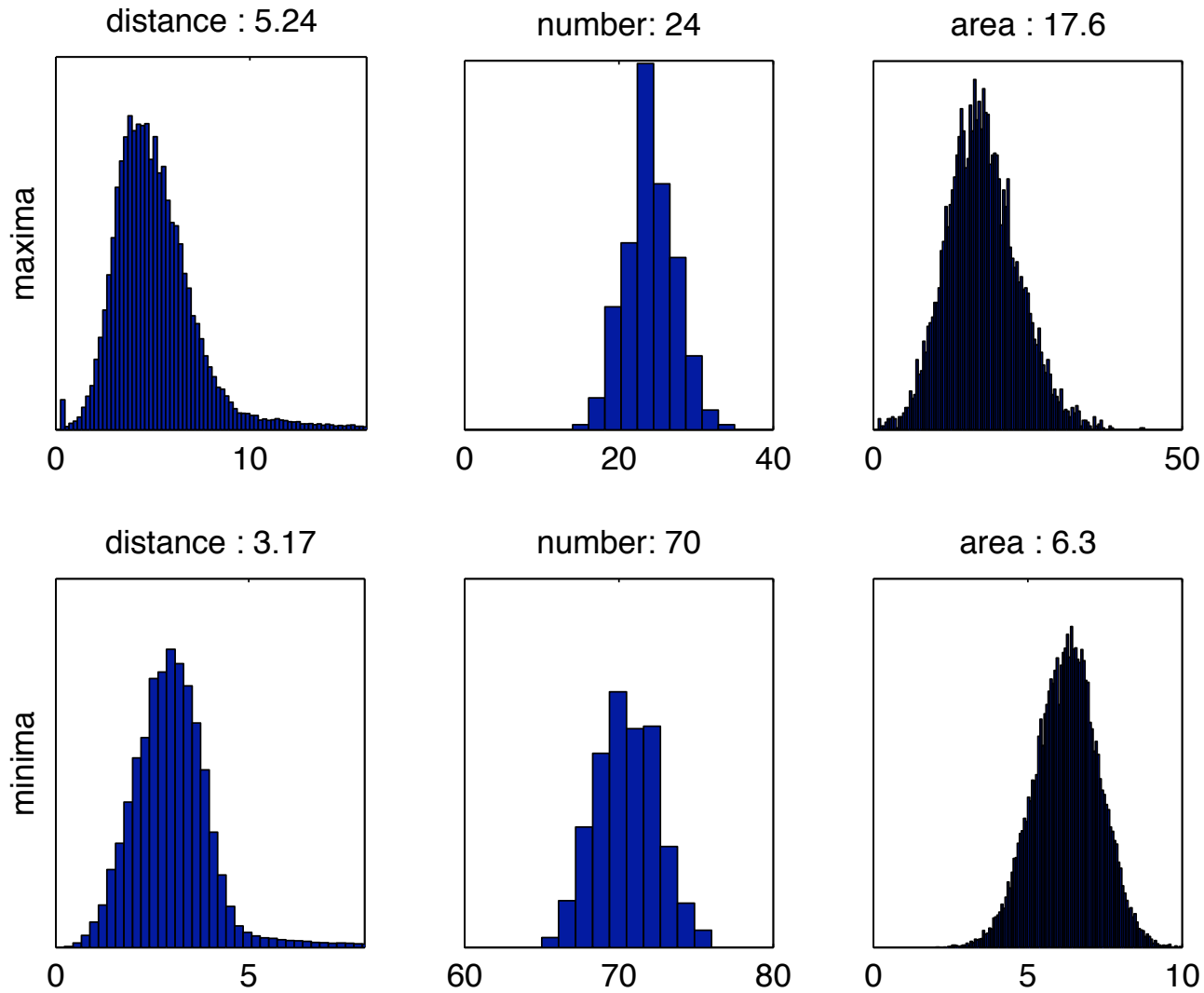
1. **Observation**: average connectivity $\approx 6 \Rightarrow$ tiling with hexagonal cells
2. **Interpretation**: maximum packing of circular patches

predictions

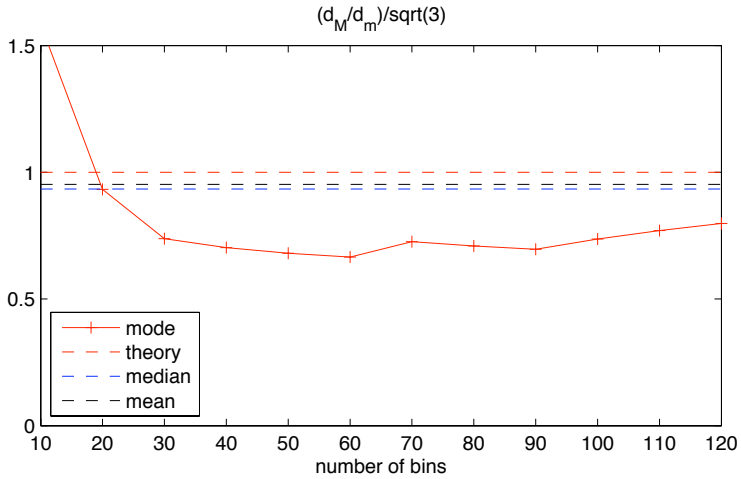


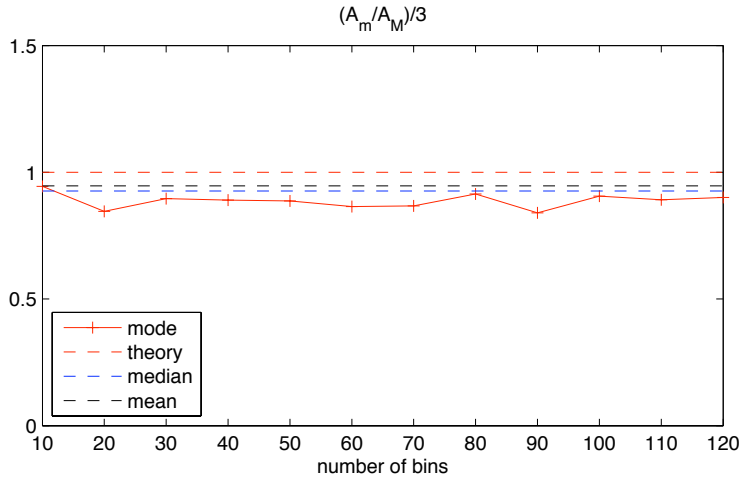
$$D_M/d_m = \sqrt{3}; N_M/N_m = 1/3; A_M/A_m = 3$$

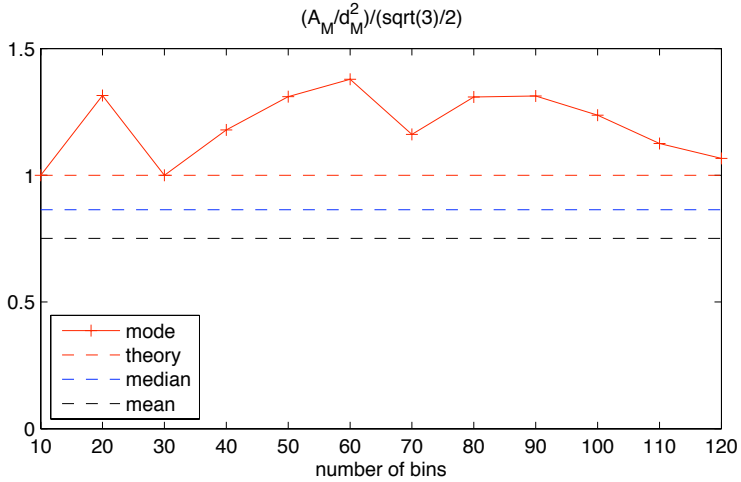
simulation results

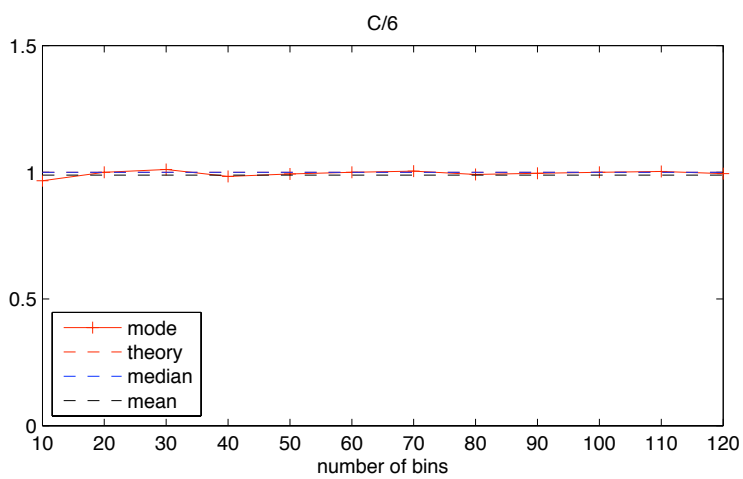


comparison with model









analysis

1. **Distributions**: significant dispersion for lengths and areas
2. **Ratios max/min**: reasonable agreement between experimental results and theoretical predictions
3. **Ranges of values**: if we call "effective domain" of the minimum uncertainty logon the circular domain which encompasses 95% of its energy, its radius and area are equal to ~ 2.6 and 21.8, to be compared to the values $d_M/\sqrt{3} \sim 3$ and $2\pi/(3\sqrt{3})A_M \sim 21.8$ attached to the hexagonal tiling
4. **Interpretation**: tiling cells $\sim$ minimum uncertainty logons

local maxima and Voronoi cells areas (1)

1. **Heights**: well-described by a Gamma distribution
2. **Areas**: *idem* and similar to the heights distribution for a proper renormalization

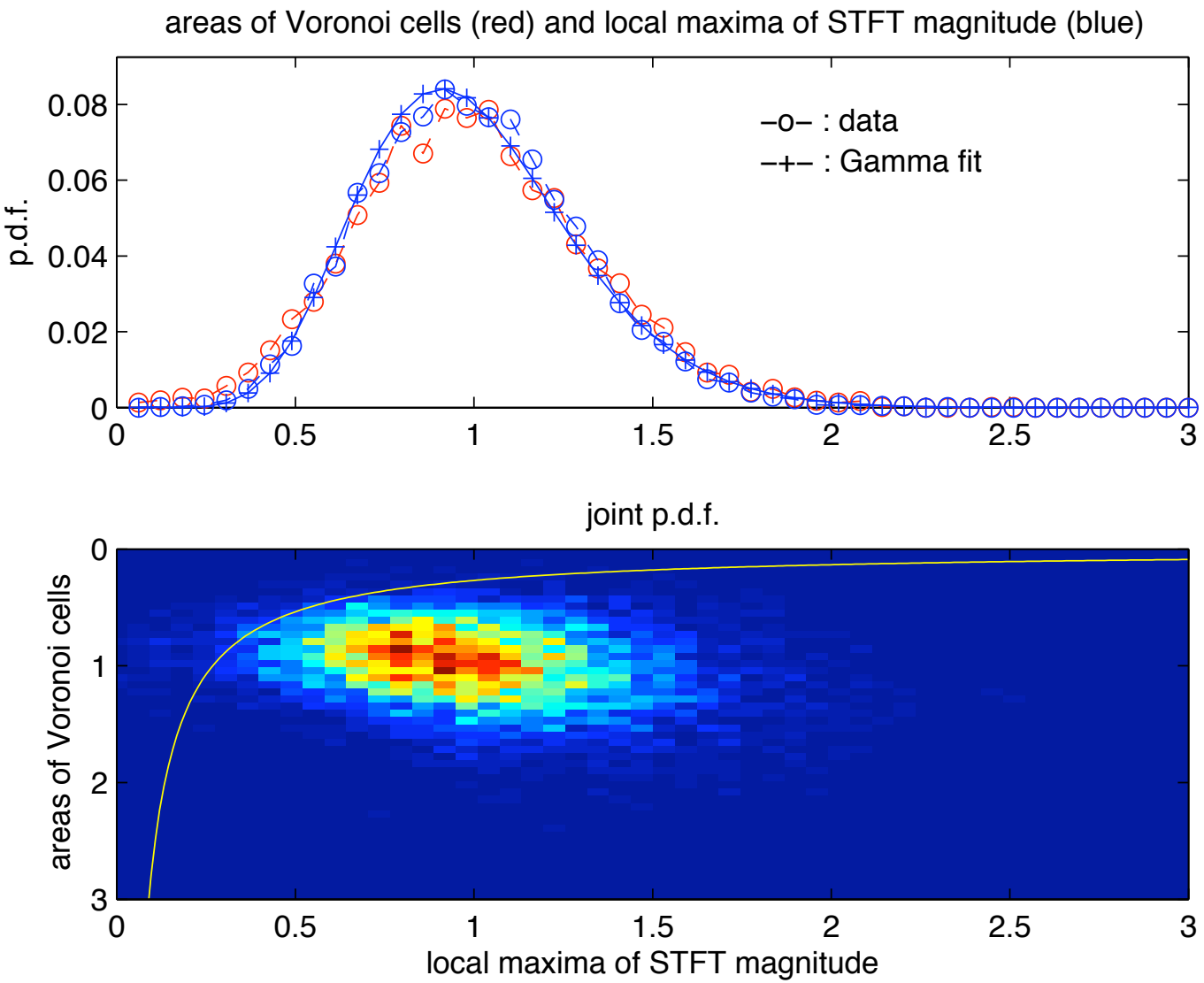
$$2 \leq \iint_A (t^2+\omega^2)\,|F_x^h(t,\omega)|^2 dt\,\frac{d\omega}{2\pi} \leq \left(\frac{|F|_*\,R^2(A)}{2}\right)^2$$

Proposition

The value $|F|_*$ of a local maximum of the STFT magnitude and the area A of the associated Voronoi cell satisfy the uncertainty-type inequality

$$A.|F|_* \geq 3\sqrt{6}$$

local maxima and Voronoi cells areas (2)



towards a random Gabor expansion

Previous results suggest a possible modeling of Gabor spectrograms as

$$S_x^h(t,\omega) = \left| \sum_m \sum_n c_{mn}\, F_h^h(t - t_m, \omega - \omega_n) \right|^2$$

with

1. locations (t_m, ω_n) distributed on some suitable randomized version of a triangular grid
2. magnitudes of the weights c_{mn} Gamma-distributed
3. locations and weights partly correlated

concluding remarks

1. **Spectrogram**: geometry constrained by time-frequency uncertainty
2. **Reassignment**: access to simplified properties, both geometrical and statistical
3. **Gabor case**: complete characterization by zeroes or local maxima?
4. **Modeling**: new ways of sparsifying energy distributions?