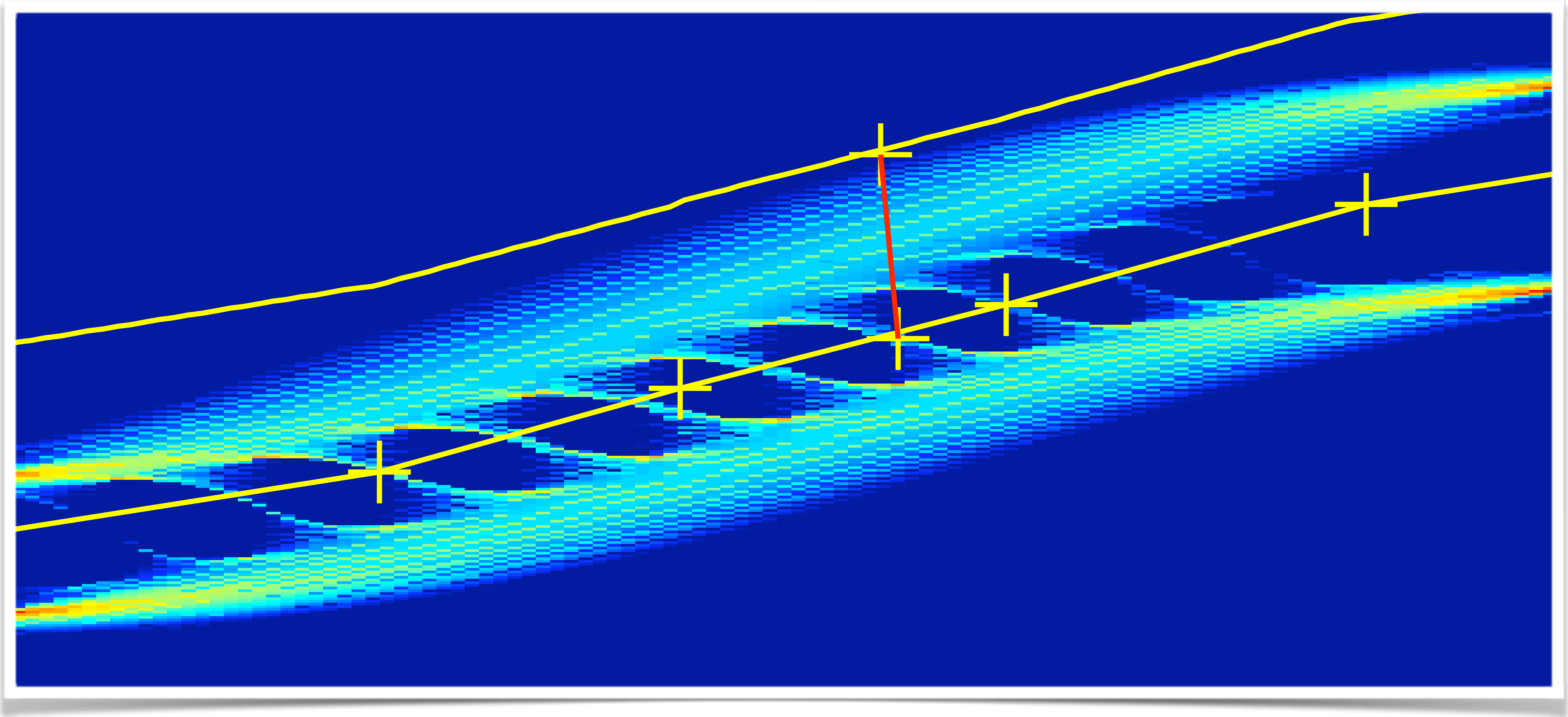
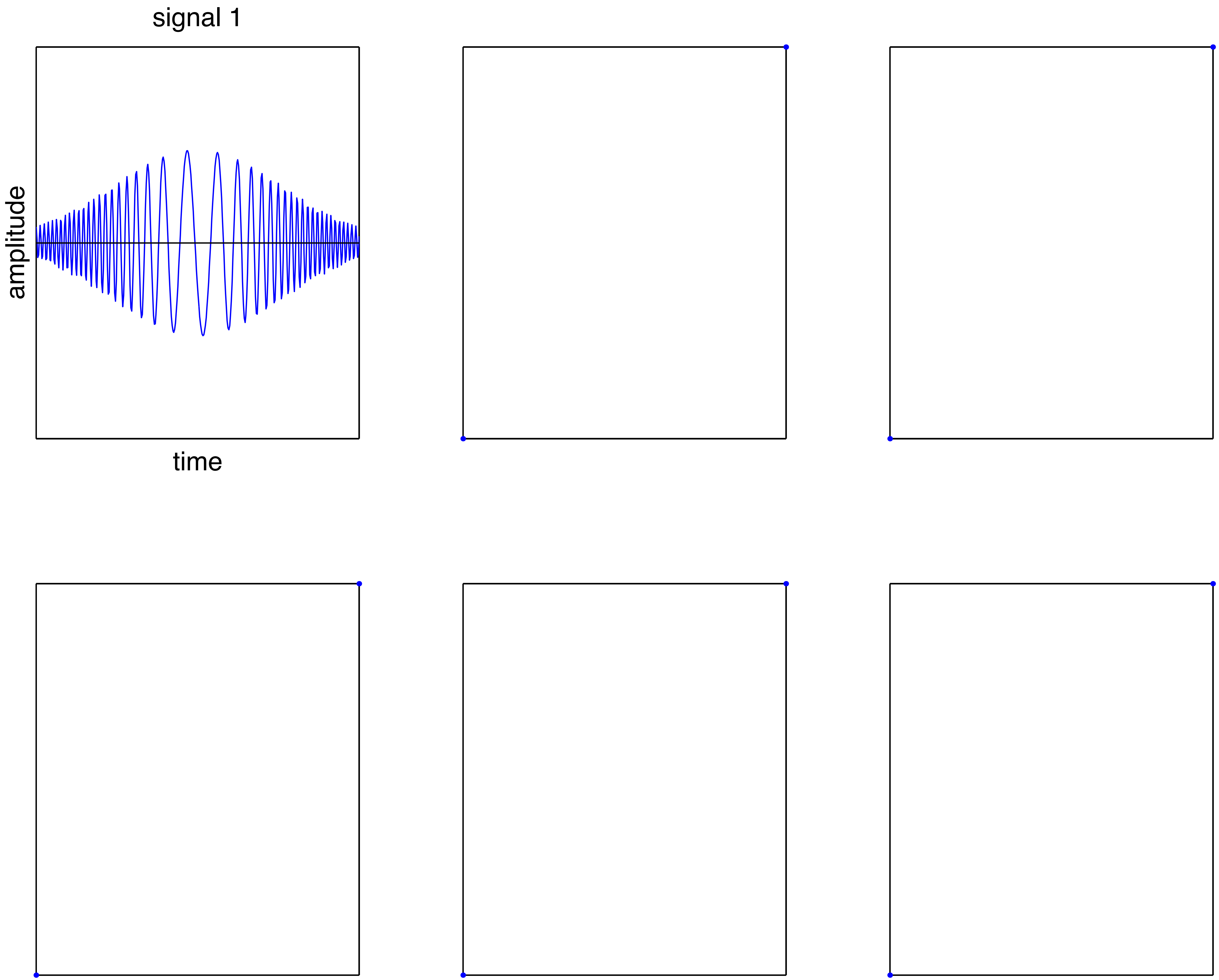
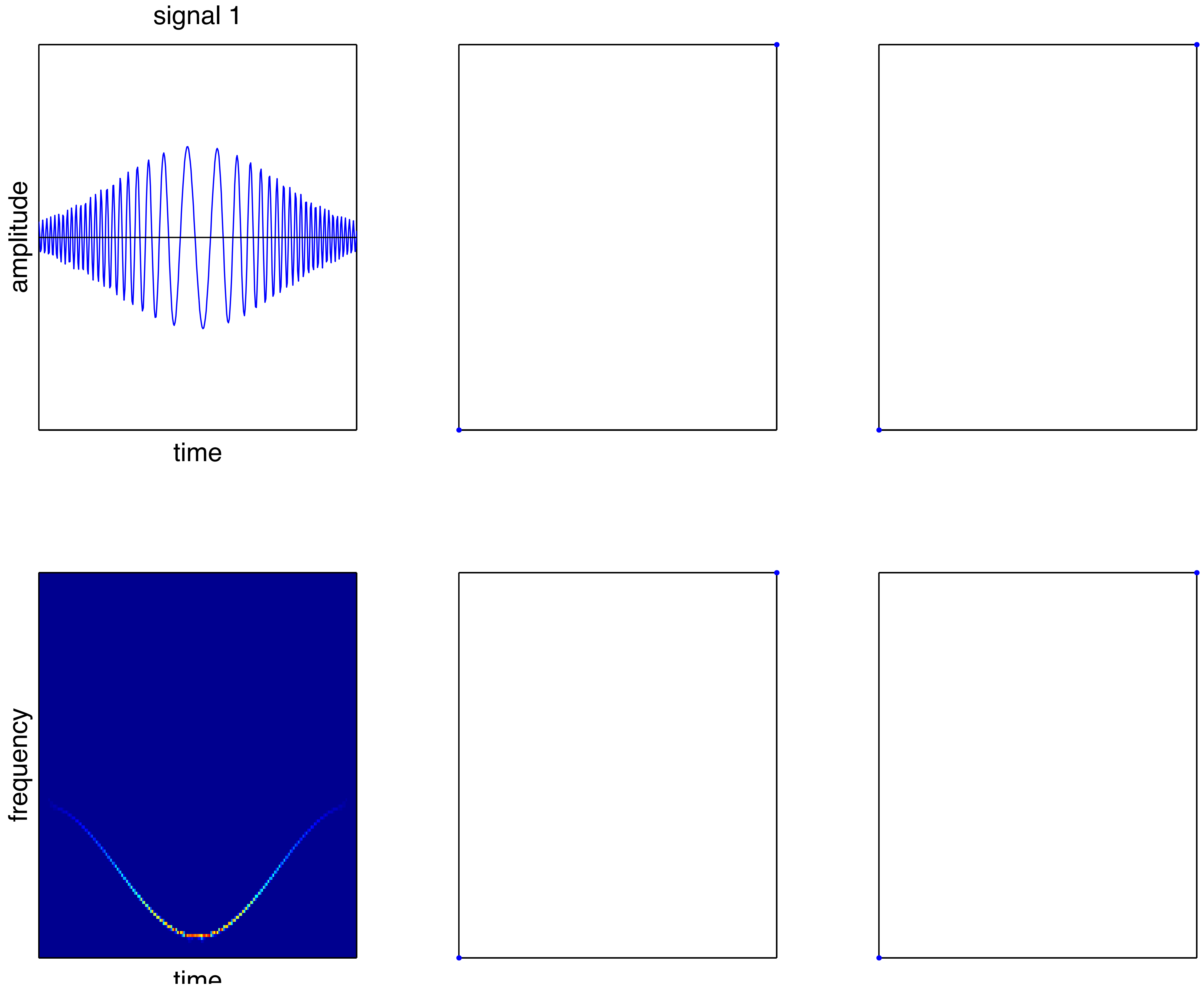


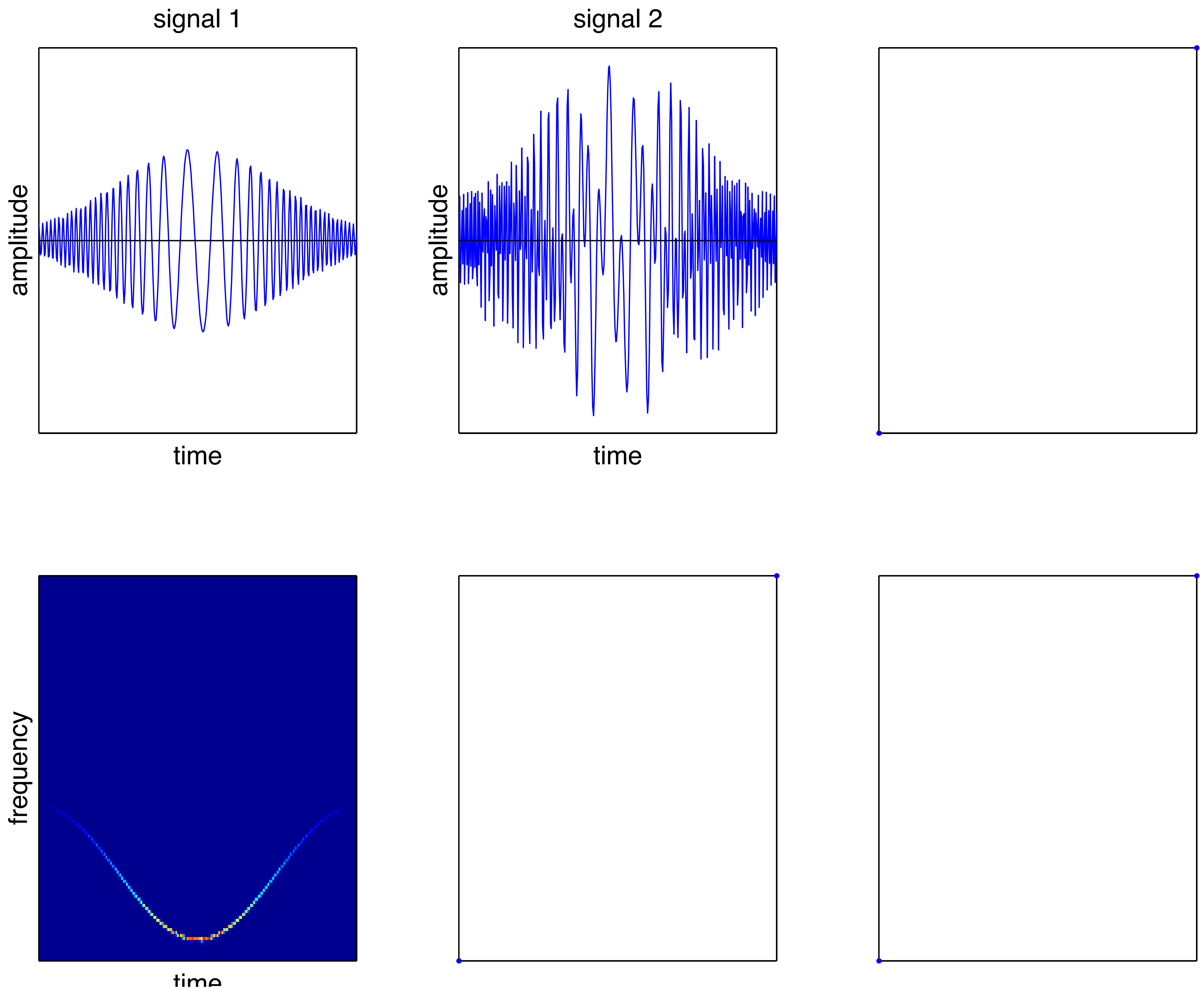
Data-driven time-frequency analyses

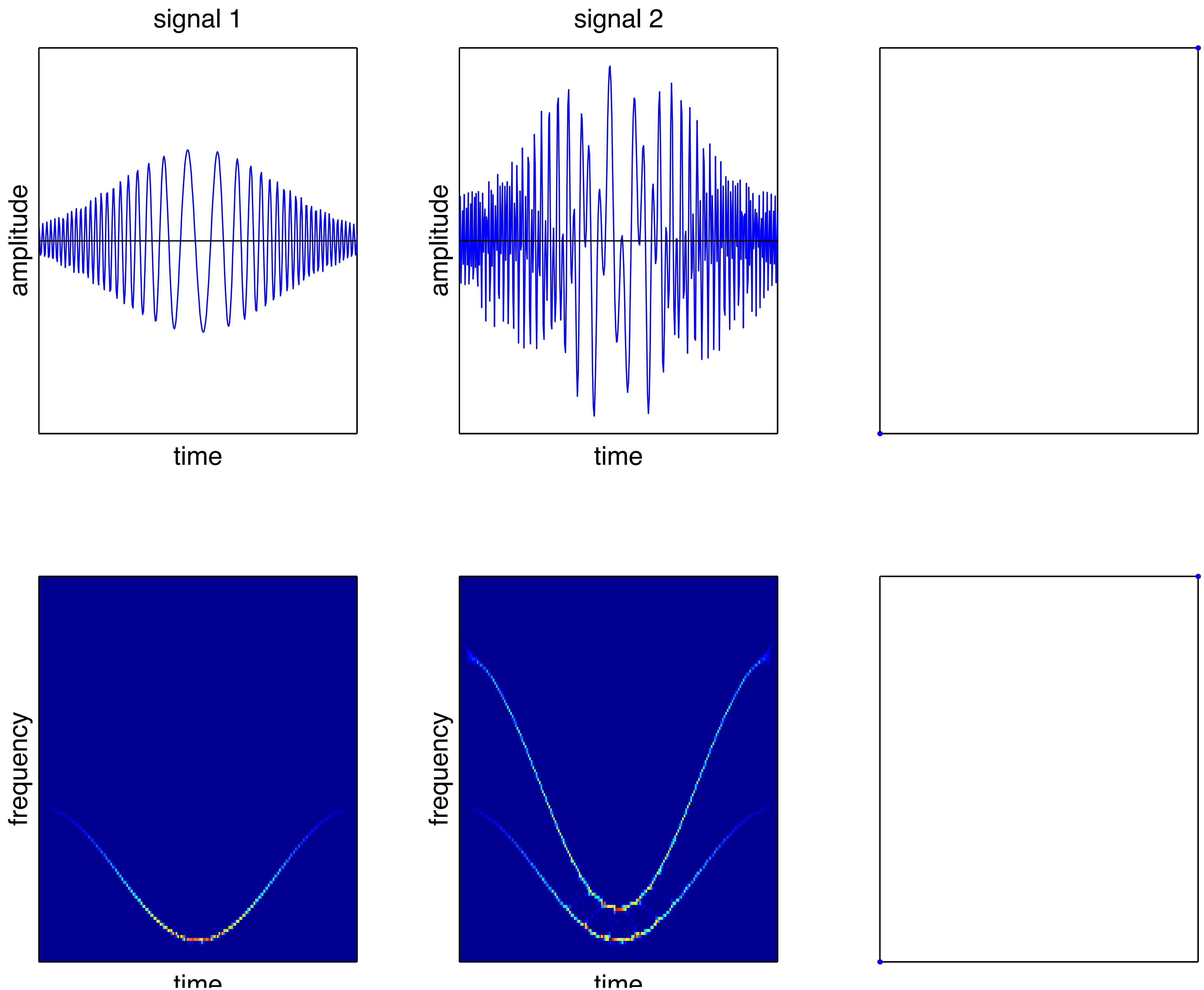
Patrick Flandrin

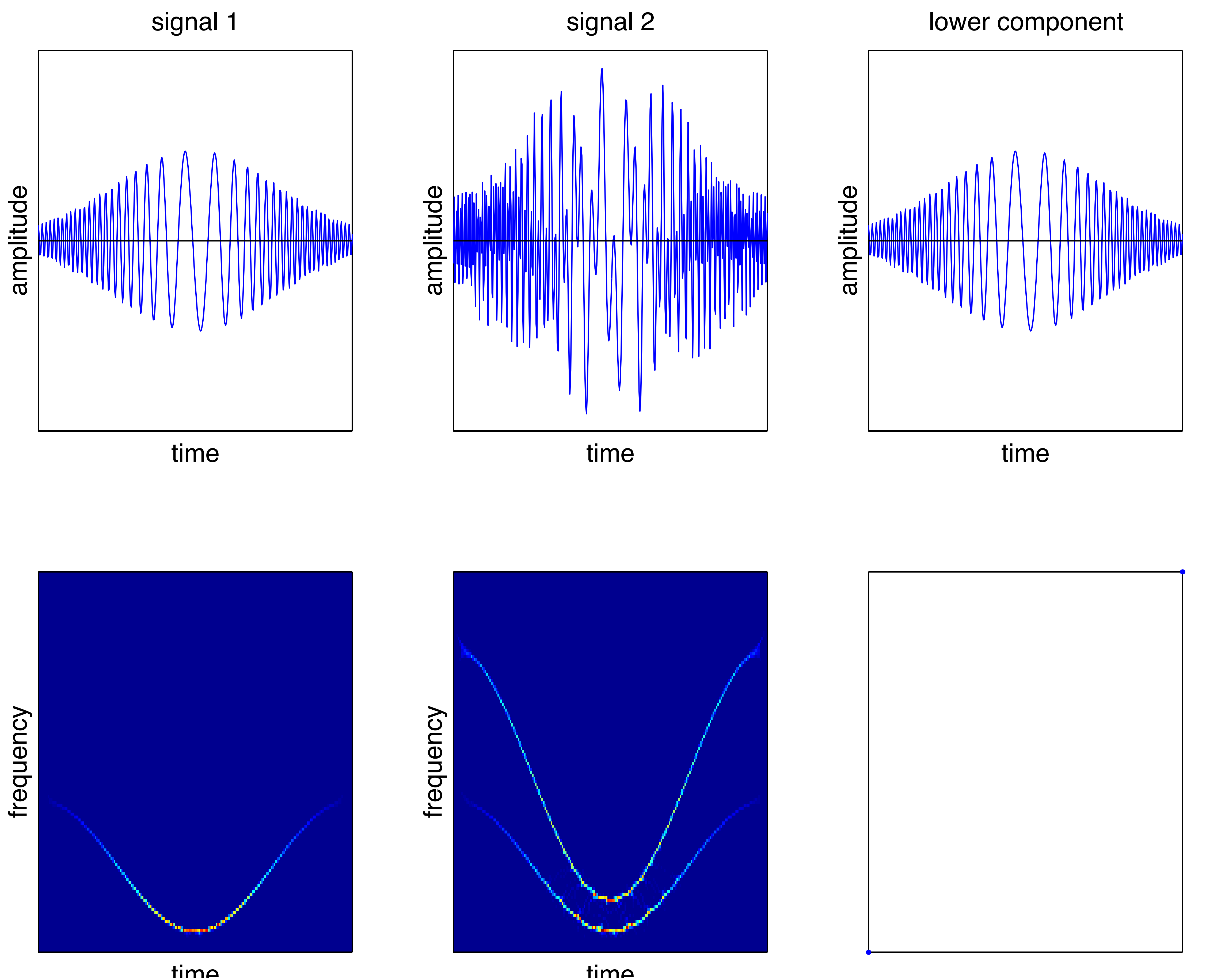
ENS DE LYON

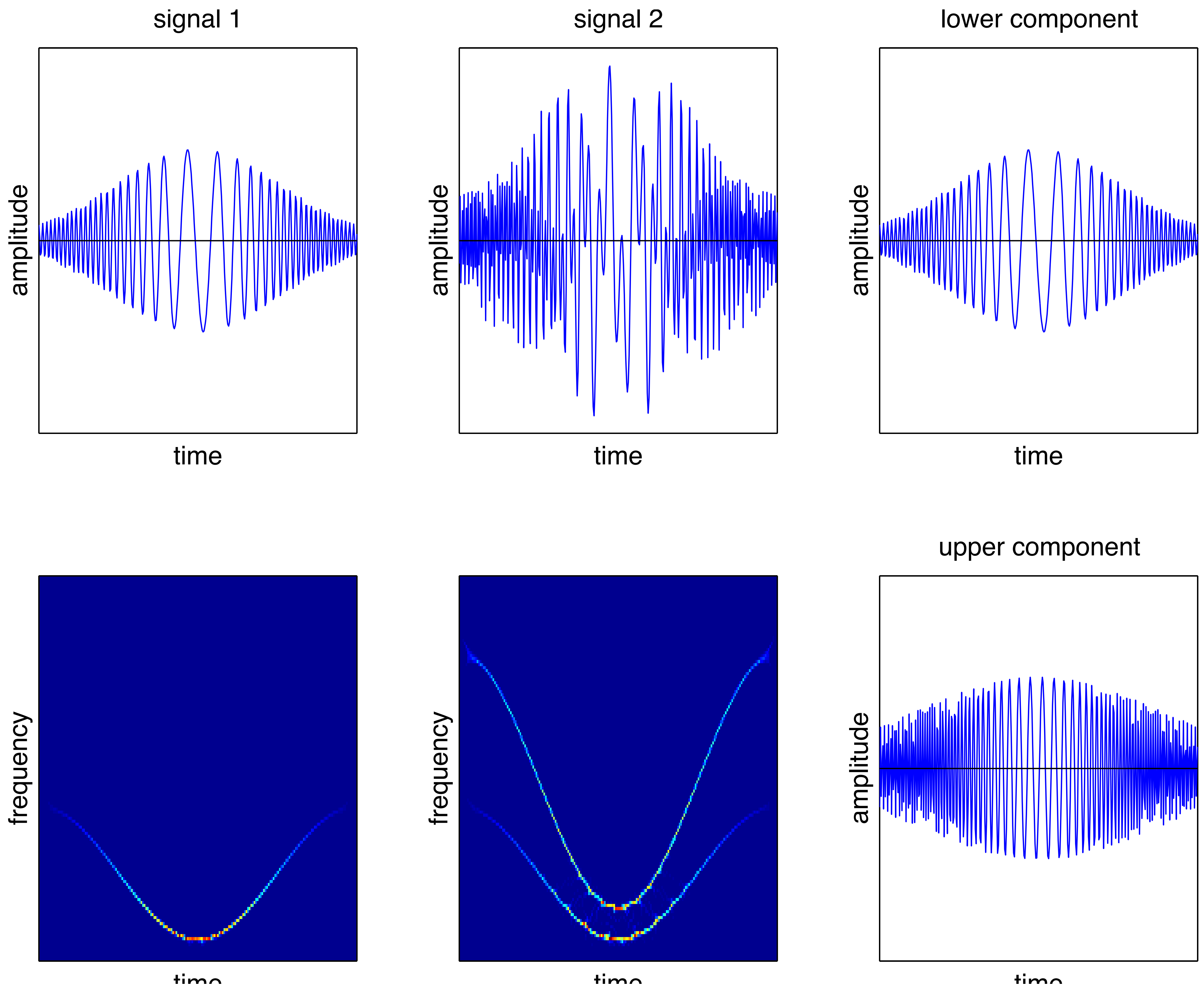

signal 1
amplitude
time

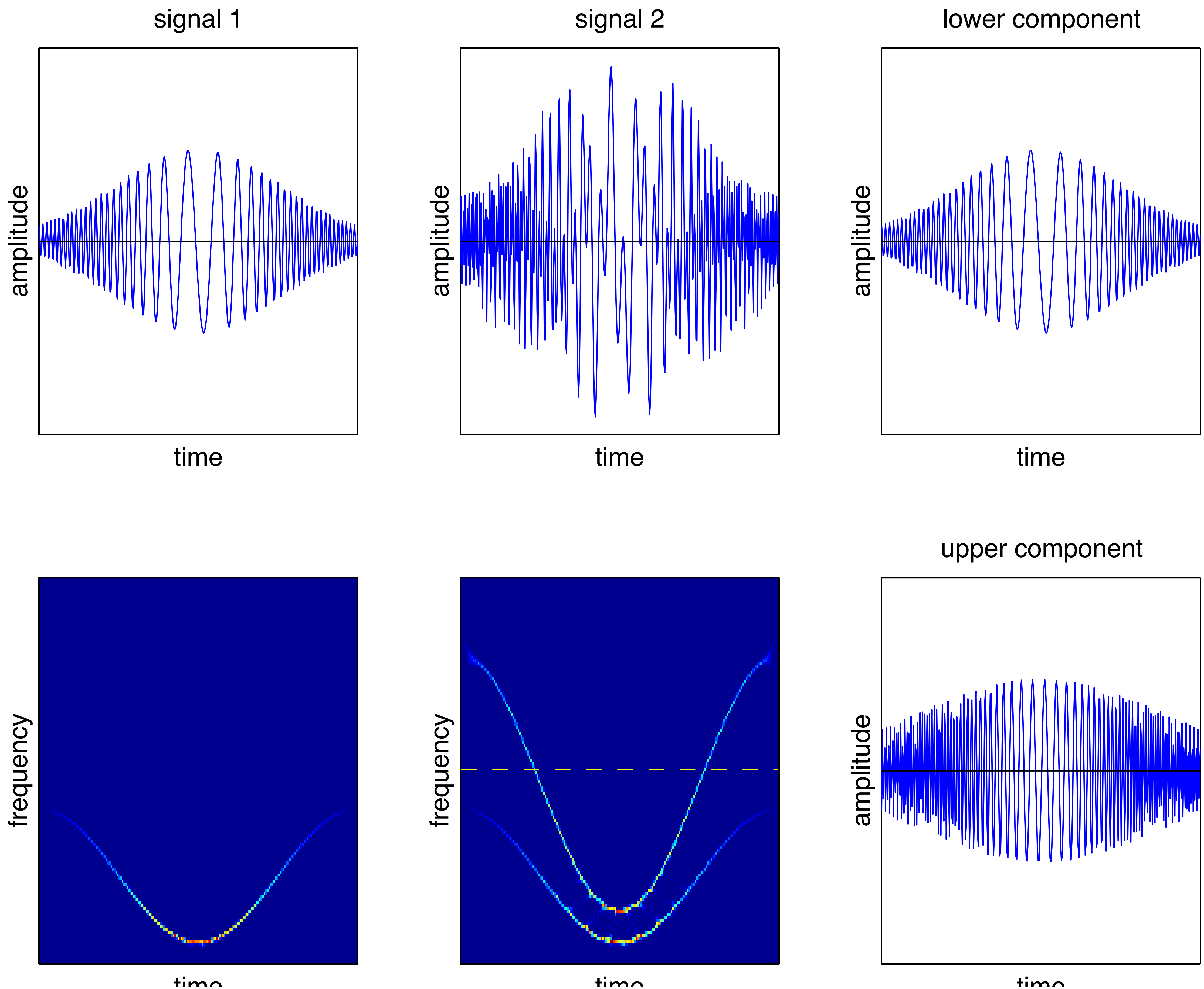

signal 1
amplitude
time
frequency
time

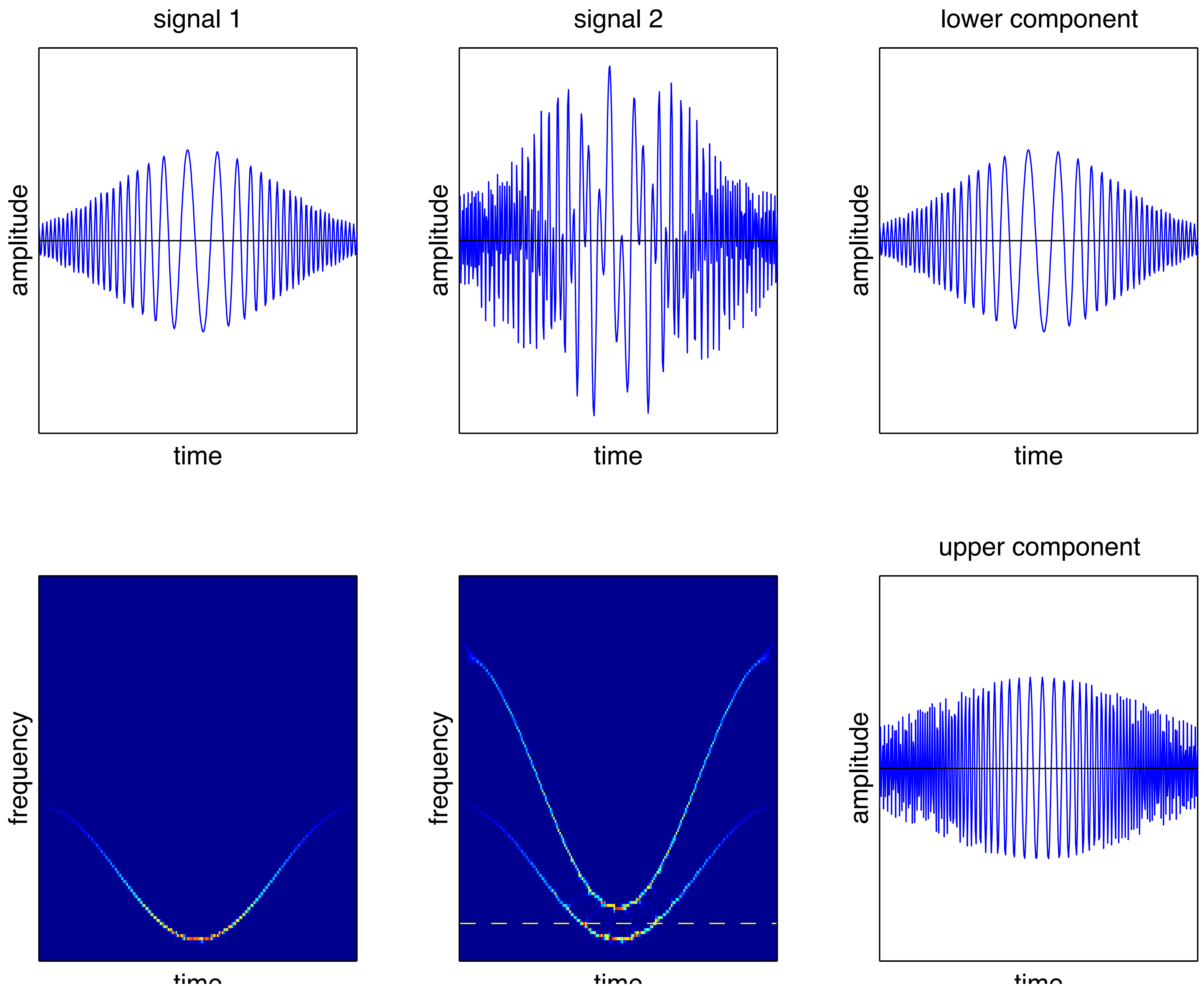

signal 1
signal 2
amplitude
time
amplitude
time
frequency
time


signal 1
signal 2
amplitude
time
amplitude
time
frequency
time
frequency
time

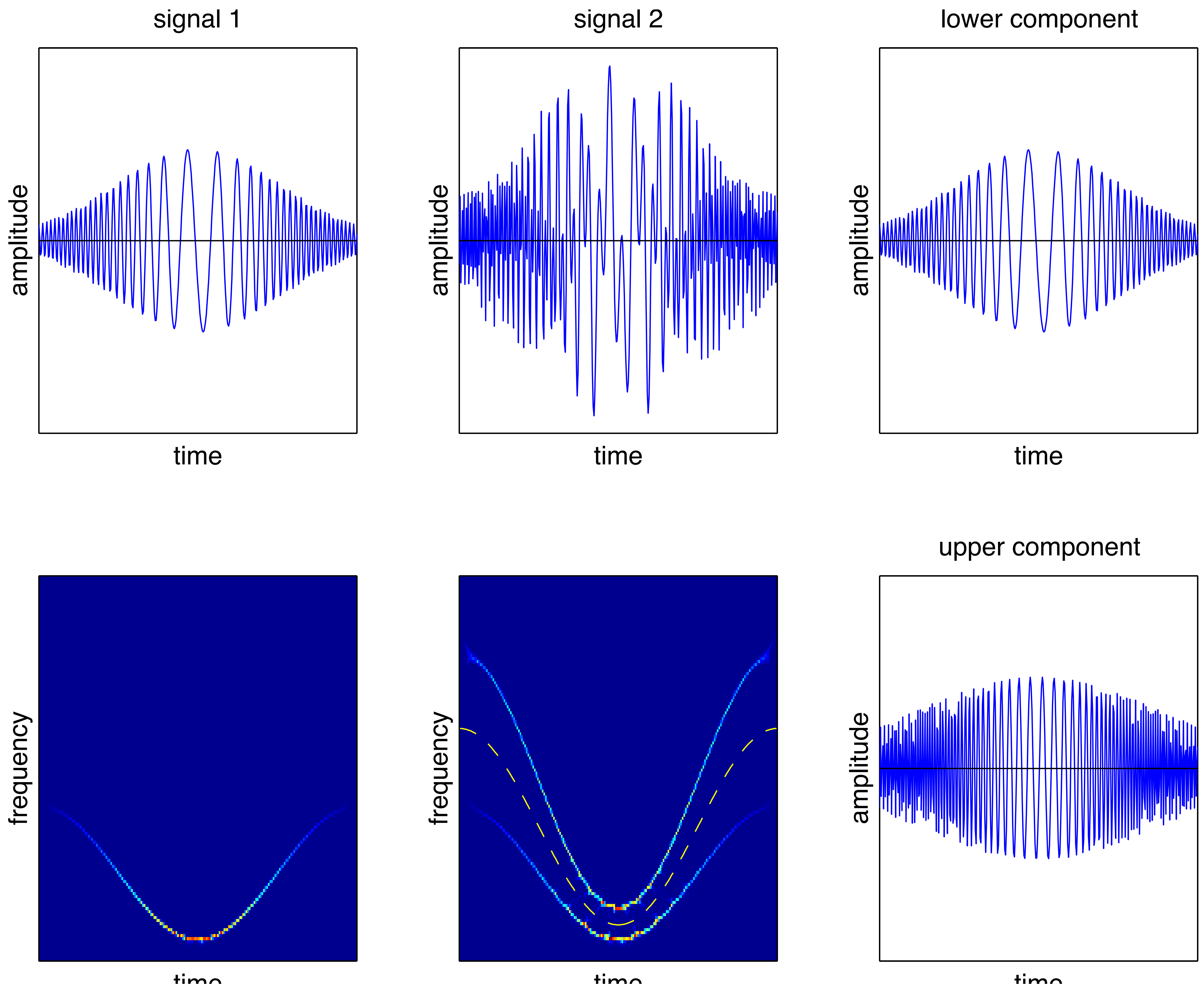

signal 1
signal 2
lower component
amplitude
time
amplitude
time
amplitude
time
frequency
time
frequency
time


signal 1
signal 2
lower component
amplitude
time
amplitude
time
amplitude
time
upper component
frequency
frequency
amplitude


signal 1
amplitude
time
signal 2
amplitude
time
lower component
amplitude
time
frequency
time
frequency
time
upper component
amplitude
time



signal 1
signal 2
lower component
upper component
amplitude
time
frequency


signal 1
signal 2
lower component
amplitude
time
upper component
frequency

Time-frequency

- Limitations of Fourier analysis in **nonstationary/nonlinear** situations
- **Many research efforts** (theory and/or applications) since 30 years
- Most approaches aimed at making time-dependent
  - *the **Fourier transform** itself (linear signal expansions: Short-Time Fourier Transform, wavelets, etc.)*
  - *the associated **spectrum analysis** (quadratic energy/power distributions: spectrogram, Wigner-Ville, Cohen's class, etc.)*
- Mostly **pre-determined transforms**, yet sometimes with some data-adaptive variations

The data revolution

- Data—rare and expensive up to a recent past—are now
  - ***abundant*** *(« Big Data »)*
  - ***multiform*** *(nD, hyperspectral, multimodal, graphs, etc.)*
  - ***cheap***
- New **opportunities** (learning, classification)
- New **challenges** (complexity, variability)

Data-driven analyses

- Two consequences of the « **data deluge** »
  - *need to face* ***extended variabilities***
  - ***no « universal » method*** *expected to be equally efficient in any context*
- Recent move towards **data-driven** methods so as to
  - ***soften*** *the rigidity of pre-determined transforms/models*
  - ***tailor*** *analyses to individual specificities*
- ***What about time-frequency?***

Outline of the talk

- Revisit of « classical » time-frequency analyses from a signal-dependent perspective
- Special focus on 3 data-driven techniques (with examples)

    - *reassignment*

    - *synchrosqueezing*

    - *Empirical Mode Decomposition*

- Concluding remarks
- Useful links

Wigner-Ville as a data-adaptive STFT

- **Short-Time Fourier Transform** (STFT)

$$F_x^{(h)}(t, f) = \int x(s)\, \overline{h(s-t)}\, e^{-i2\pi f s}\, ds$$

- « **Matched filter** » principle:

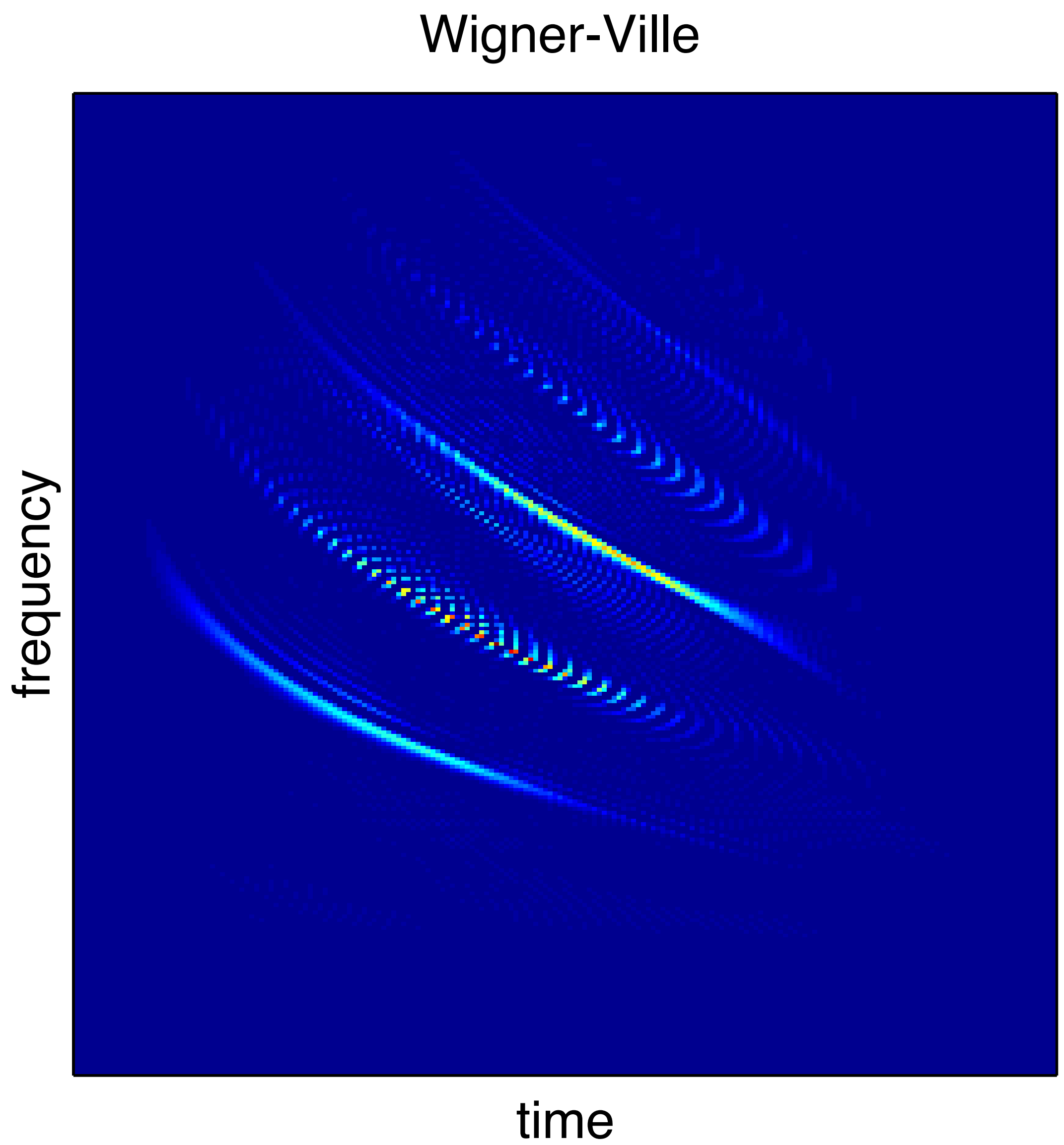
$$h(t) = x_-(t) := x(-t) \Rightarrow F_x^{(x_-)}(t, f) = W_x(t/2, f/2)/2,$$

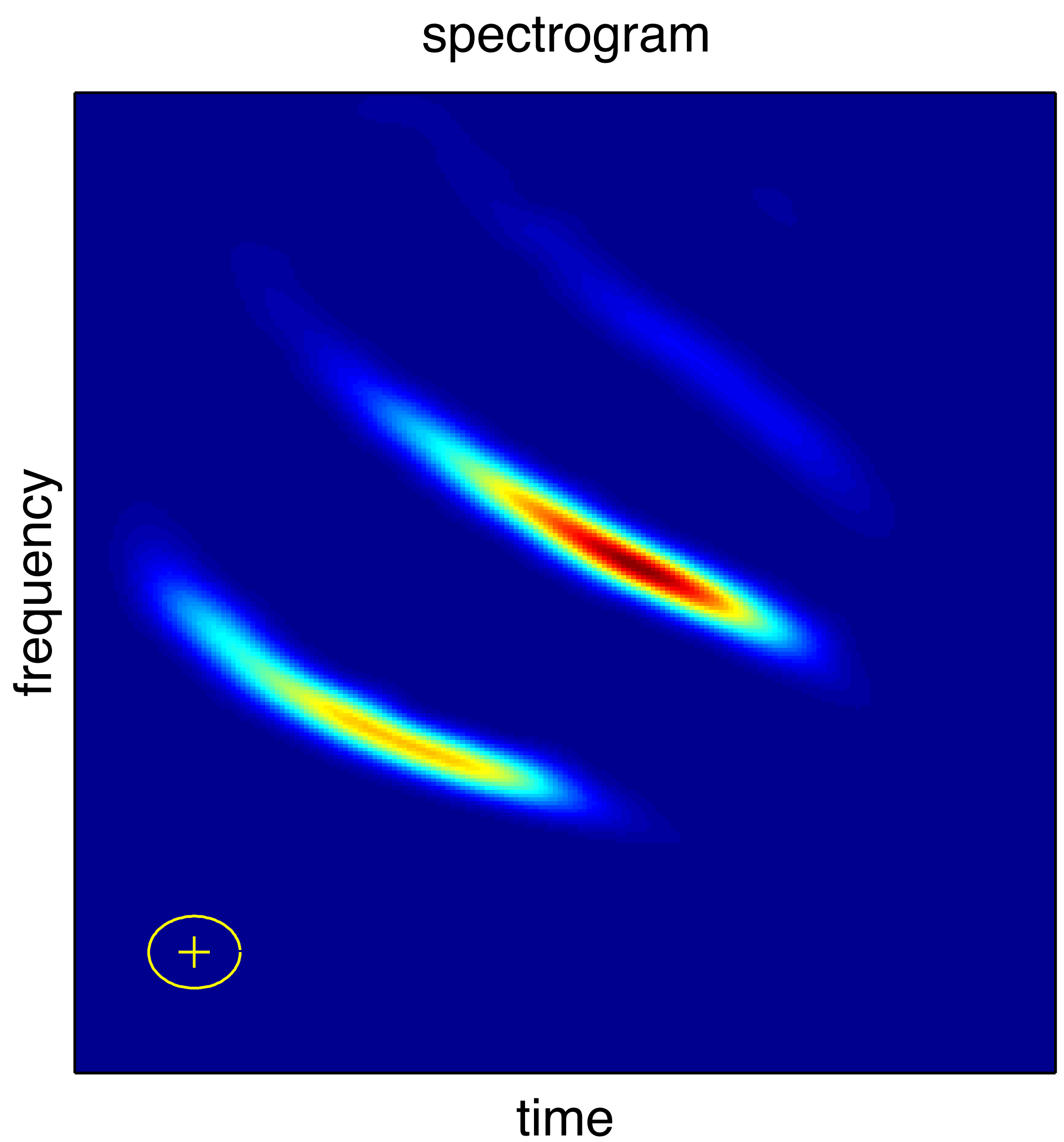
with

$$W_x(t, f) := \int x(t + \tau/2)\, \overline{x(t - \tau/2)}\, e^{-i2\pi f \tau}\, d\tau$$

the **Wigner-Ville distribution**

Wigner-Ville vs. STFT / spectrogram





Wigner-Ville – pros and cons

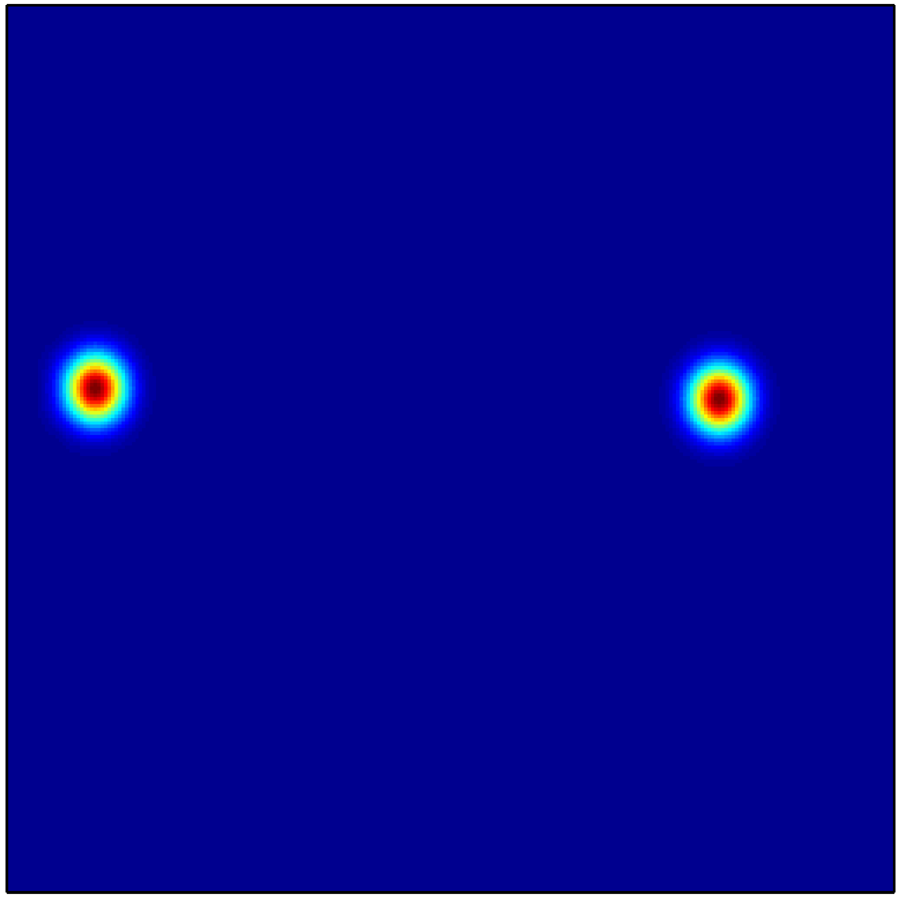
- (P) **Perfect localization** for monocomponent linear chirps
- (C) **Interference terms** for multicomponent signals

Both (P) and (C) result from the very same

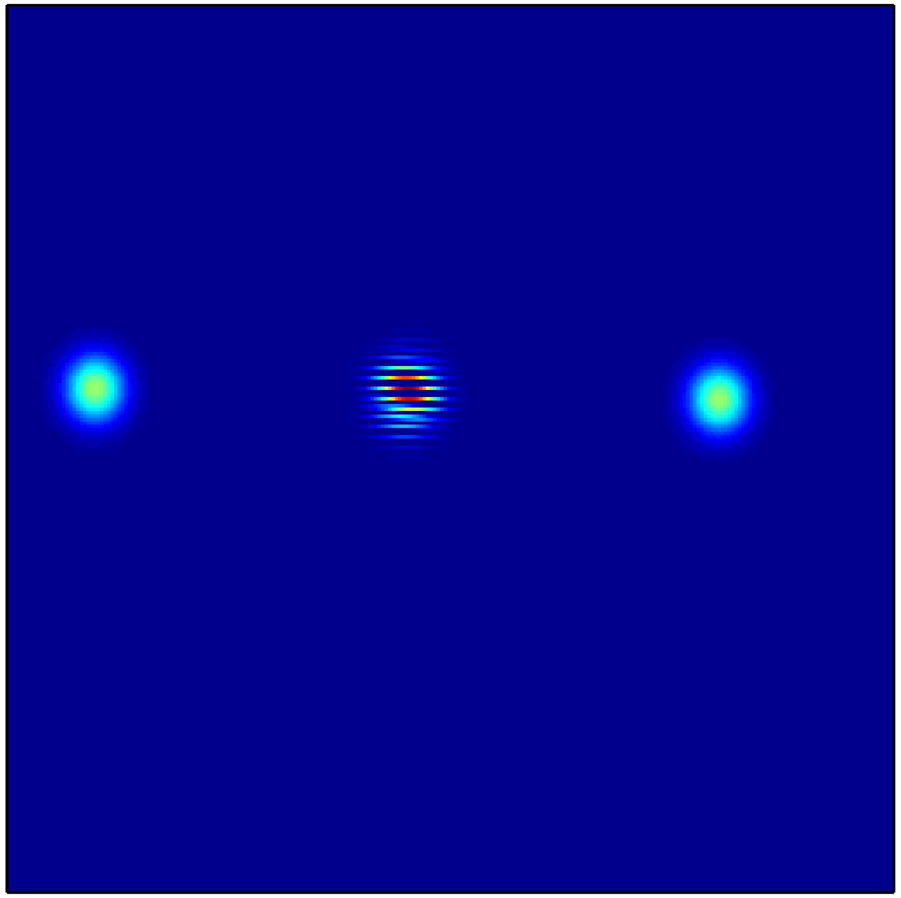
quadratic superposition principle

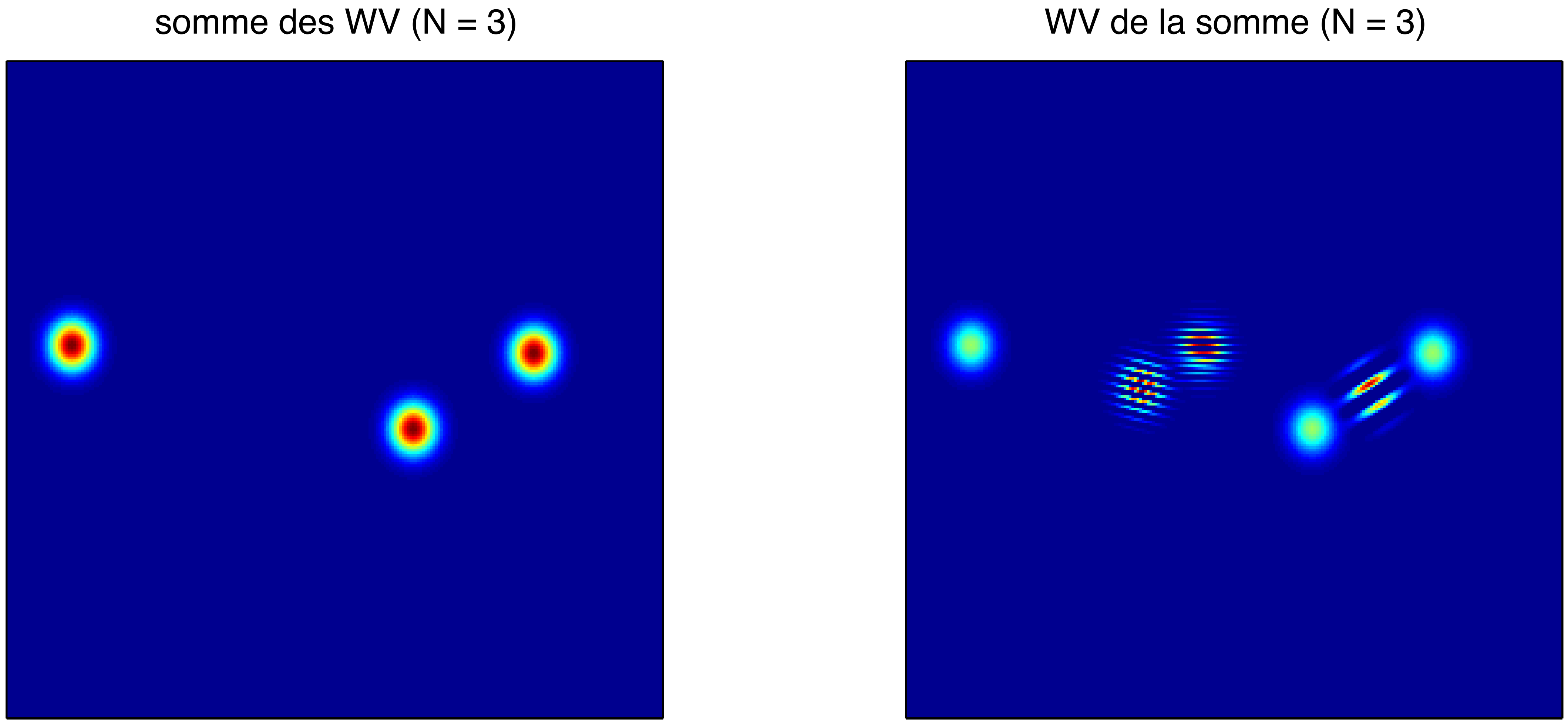
- New type of **time-frequency trade-off**, complementary to Heisenberg's

somme des WV (N = 2)

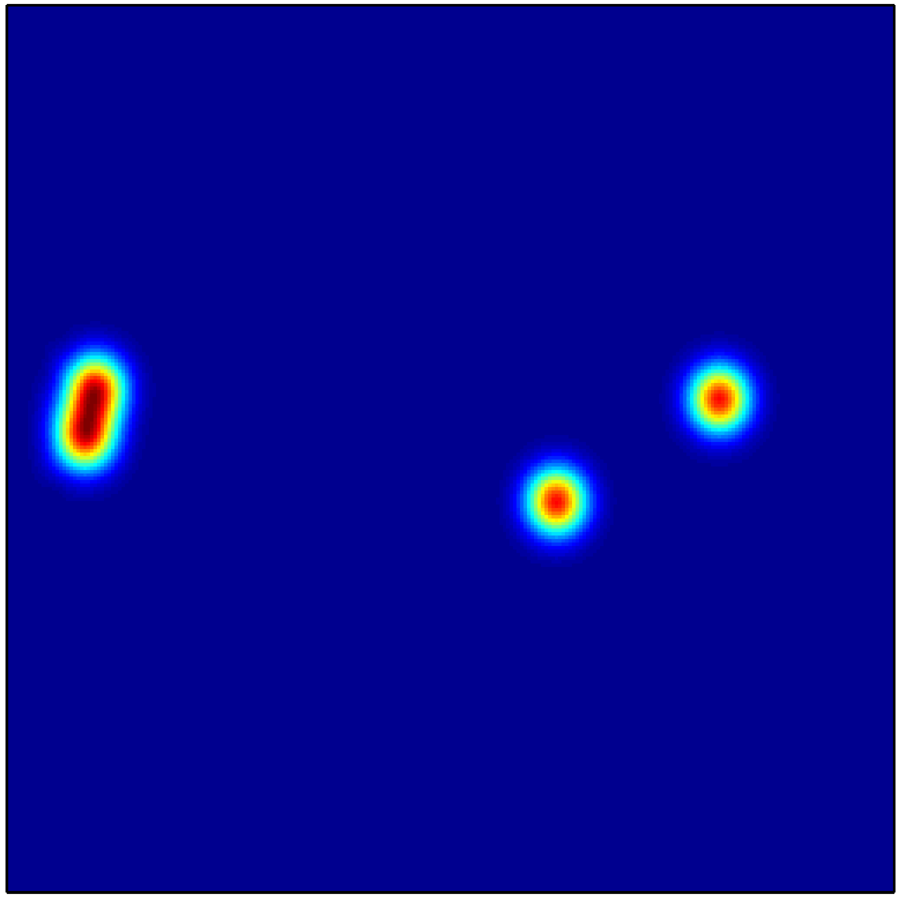


WV de la somme (N = 2)

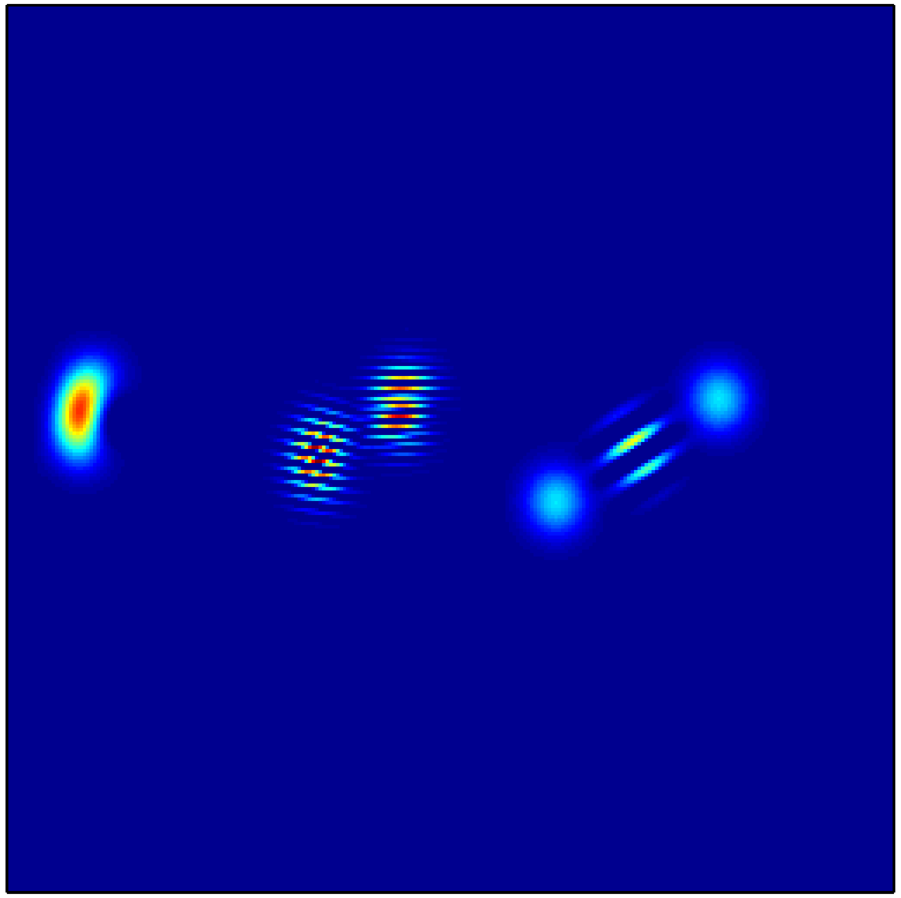


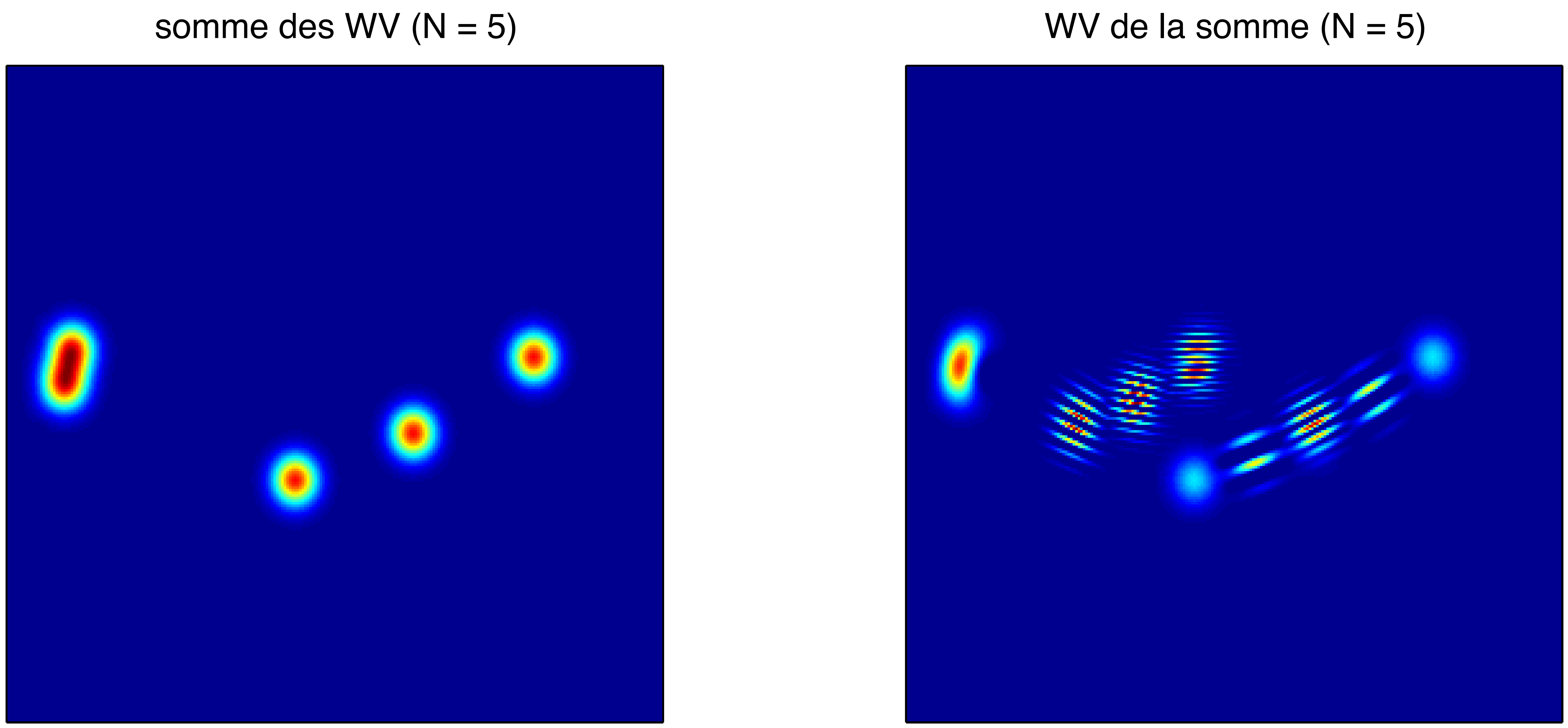

somme des WV (N = 3)
WV de la somme (N = 3)

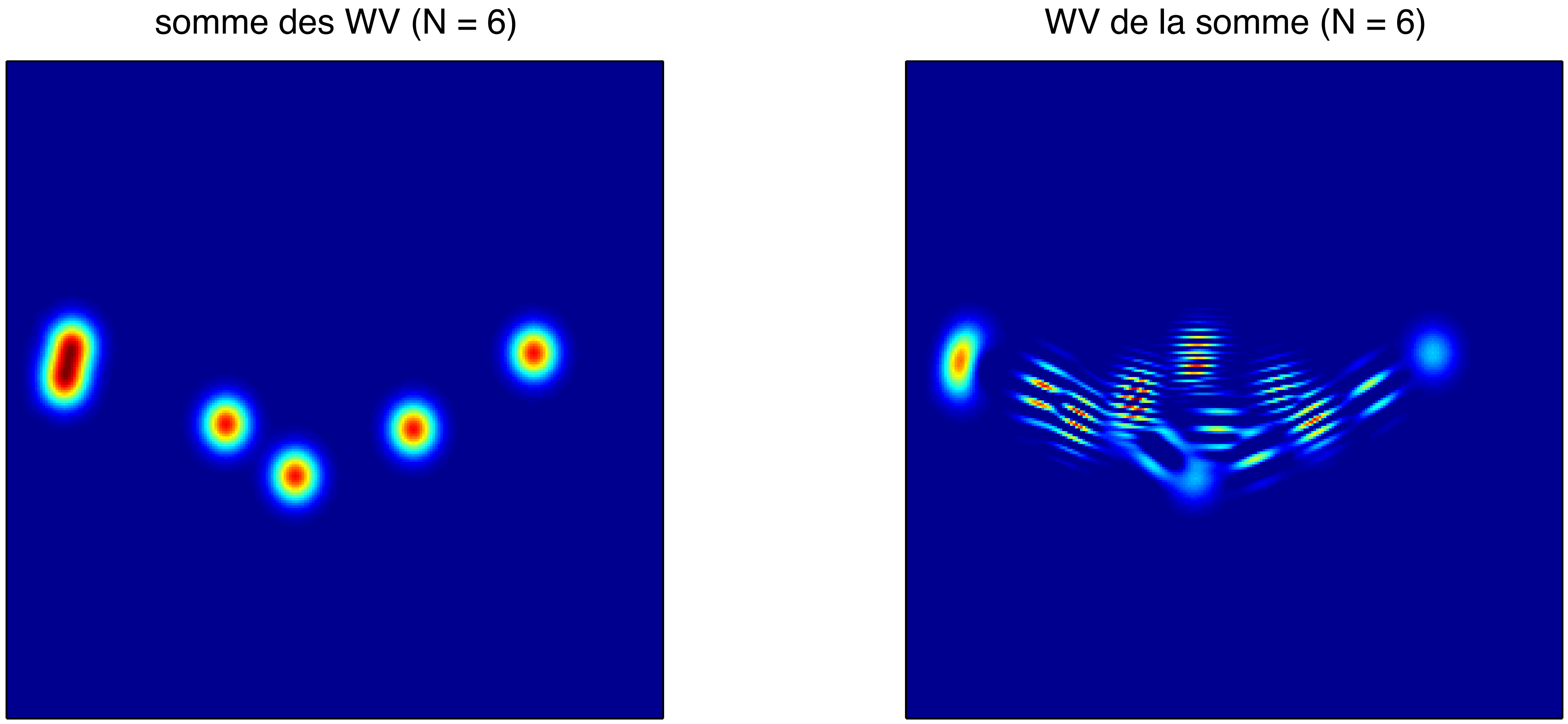
somme des WV (N = 4)

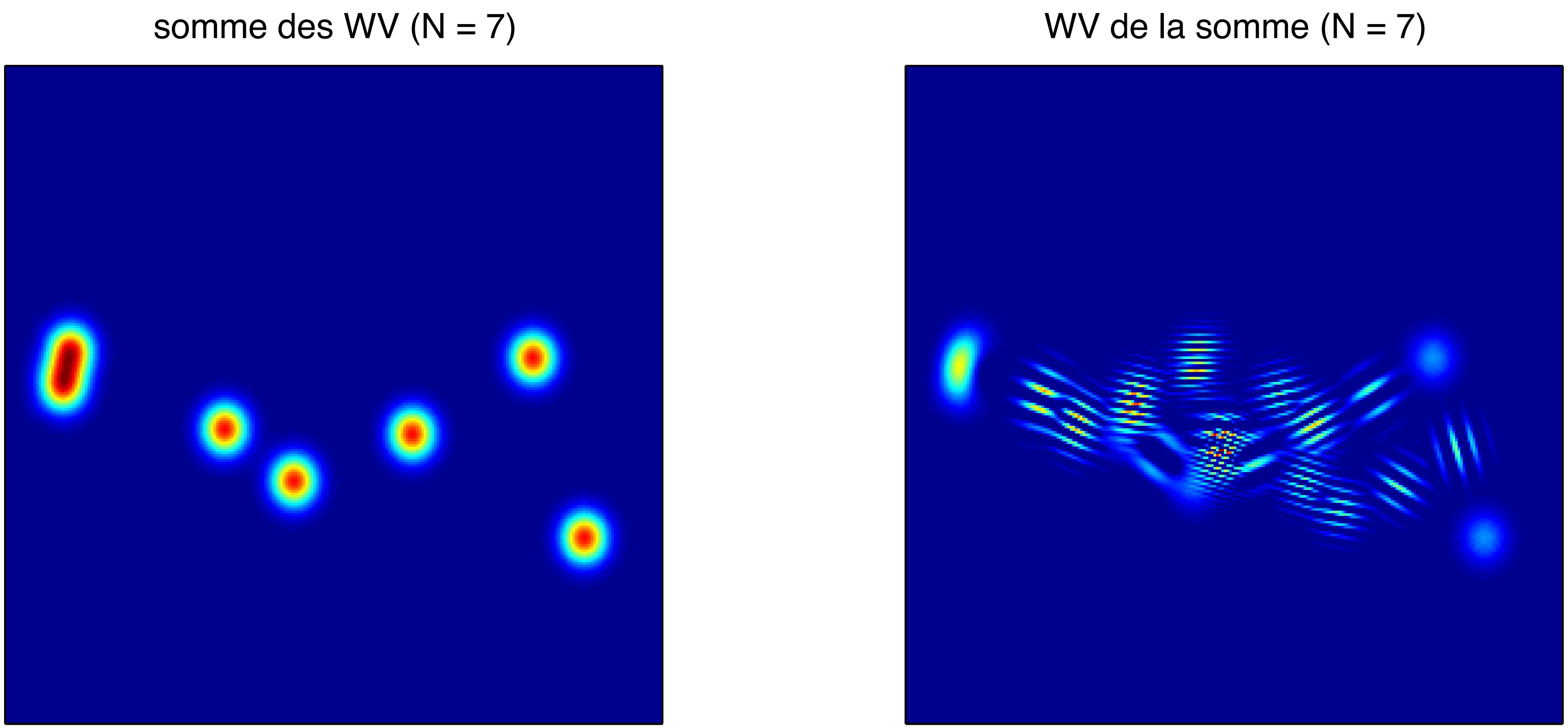


WV de la somme (N = 4)

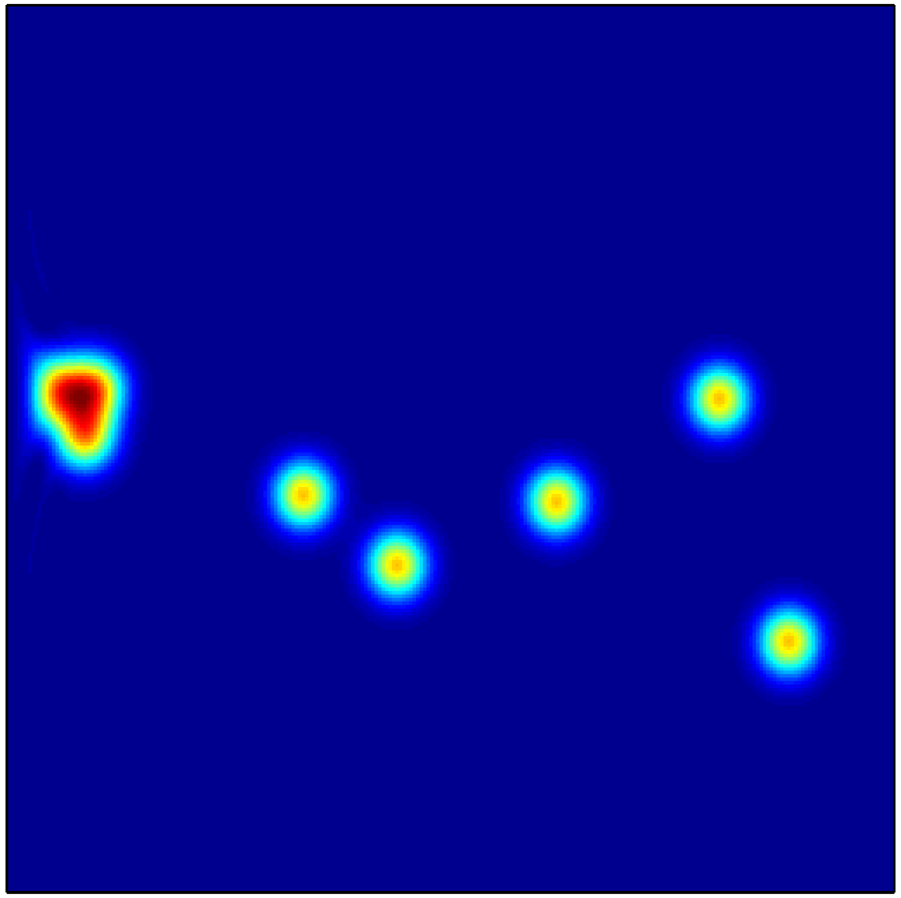



somme des WV (N = 5)
WV de la somme (N = 5)

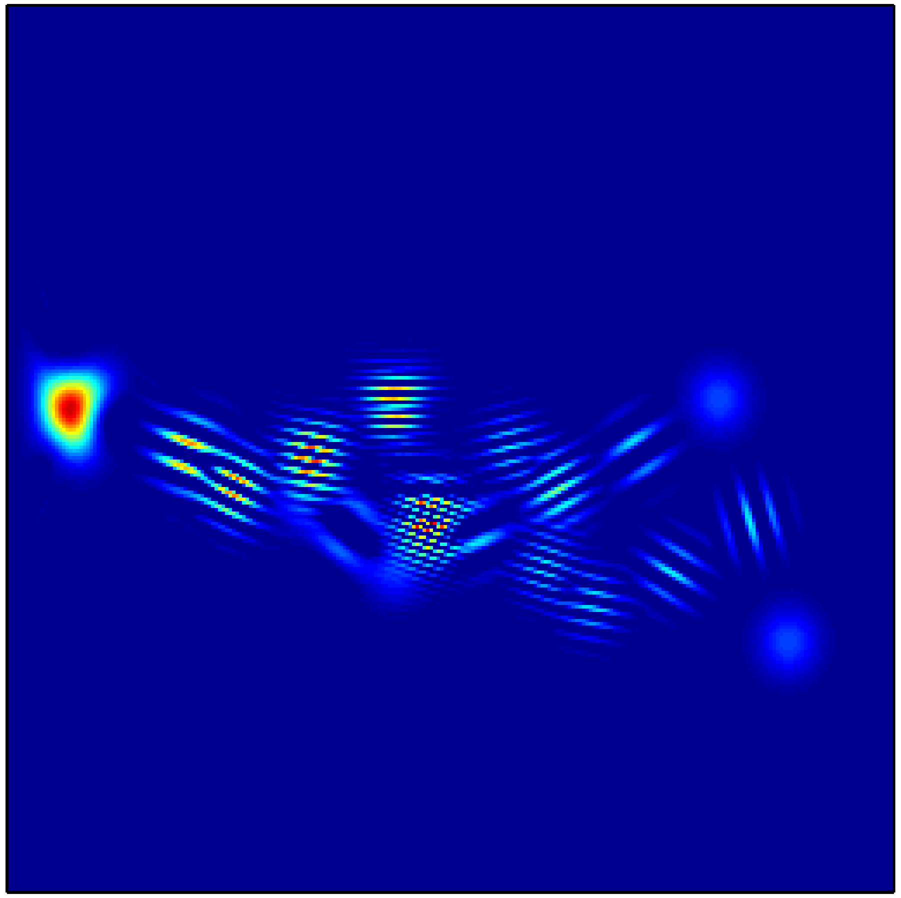

somme des WV (N = 6)
WV de la somme (N = 6)


somme des WV (N = 7)
WV de la somme (N = 7)

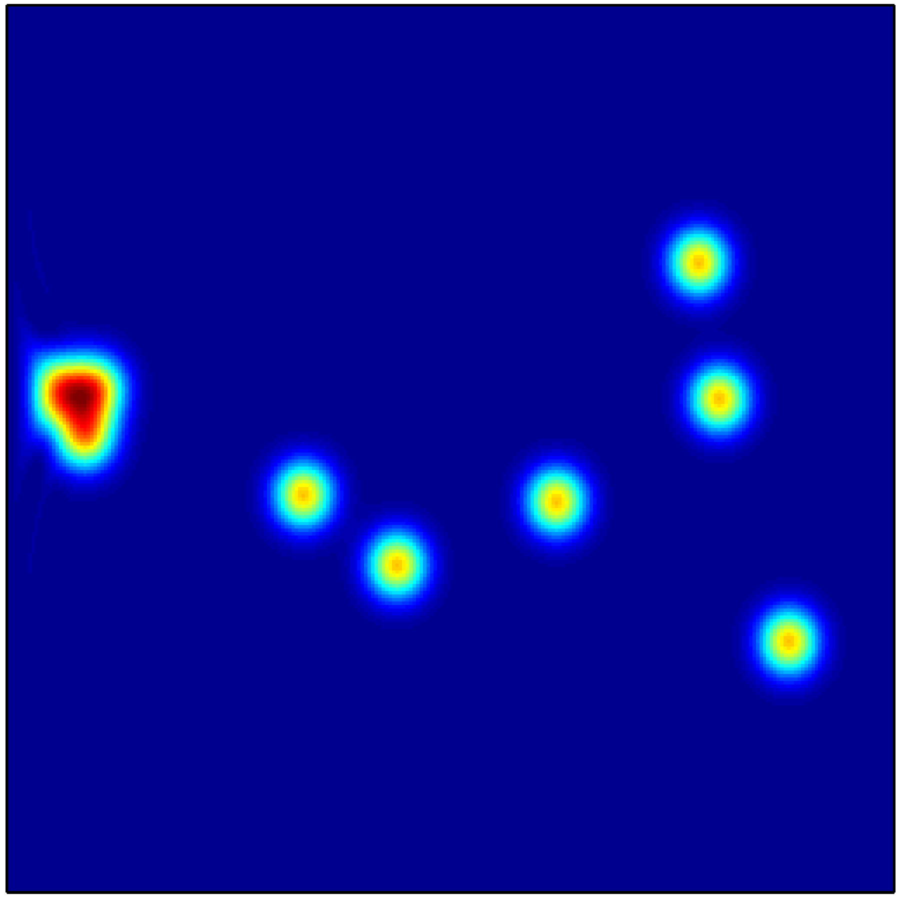
somme des WV (N = 8)



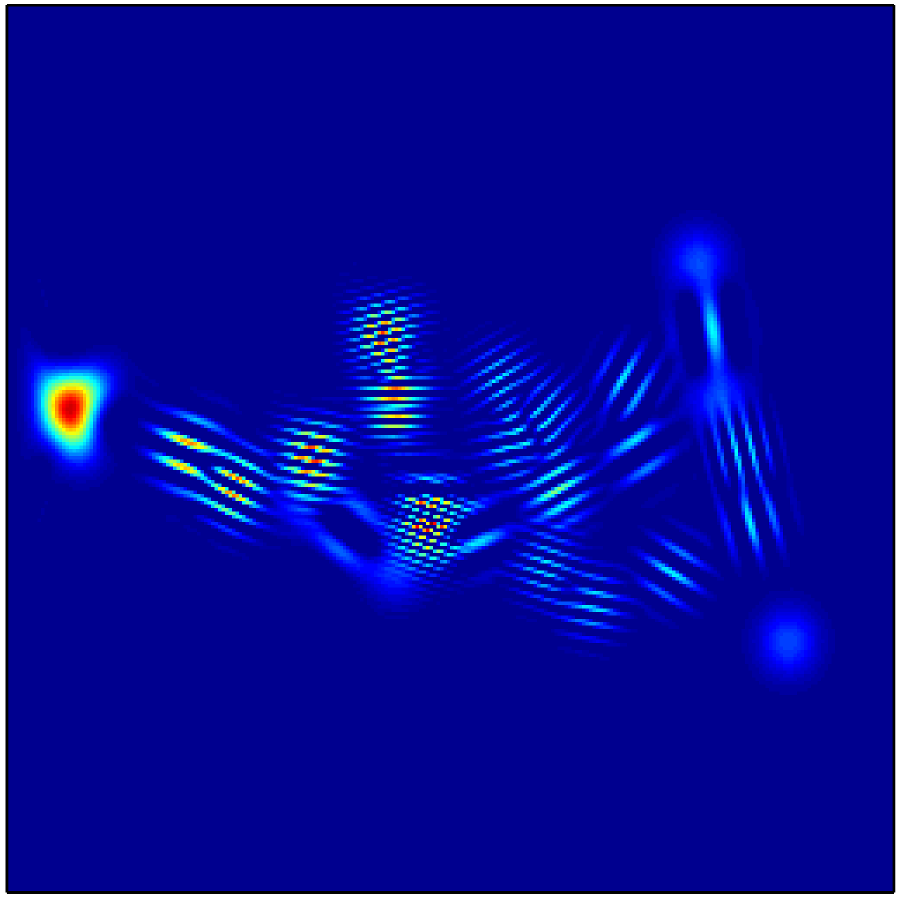
WV de la somme (N = 8)

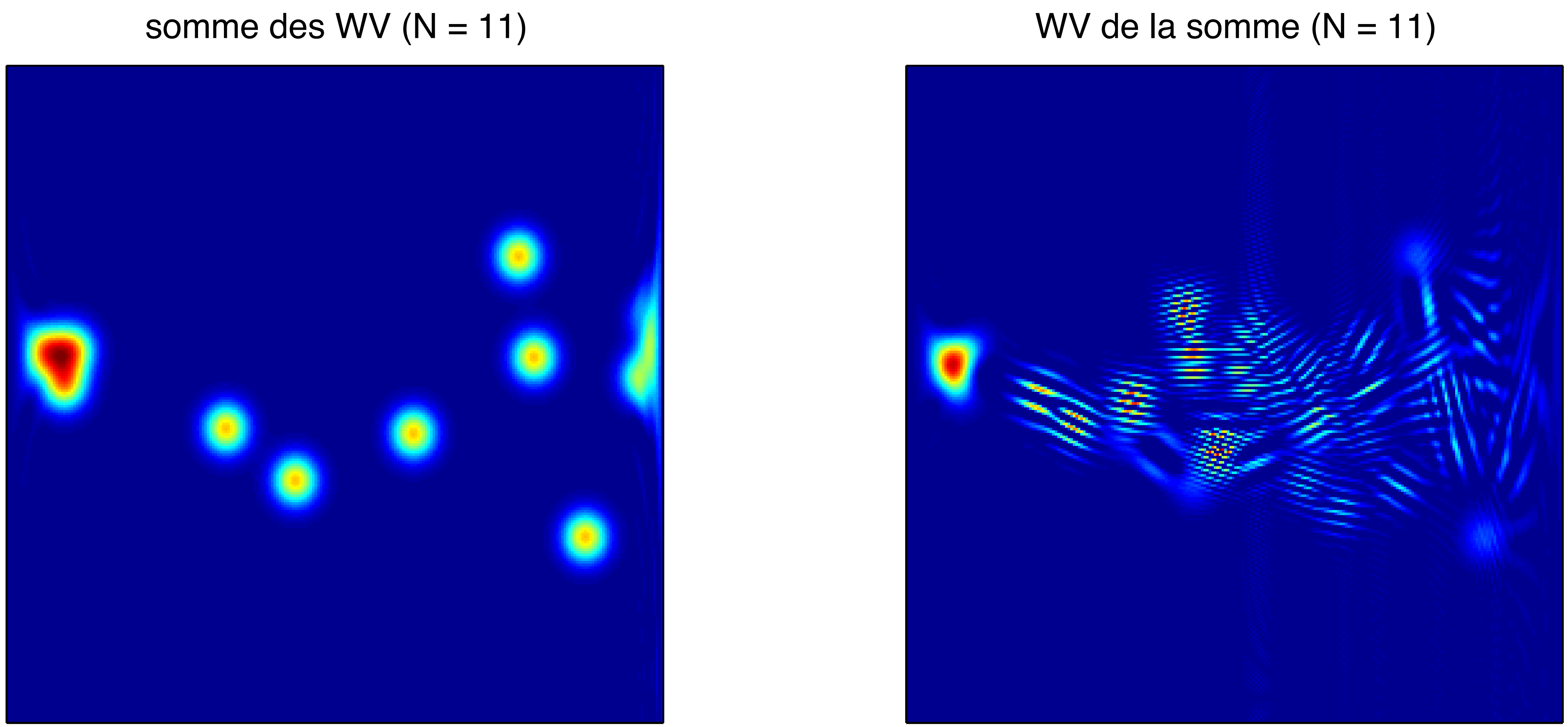


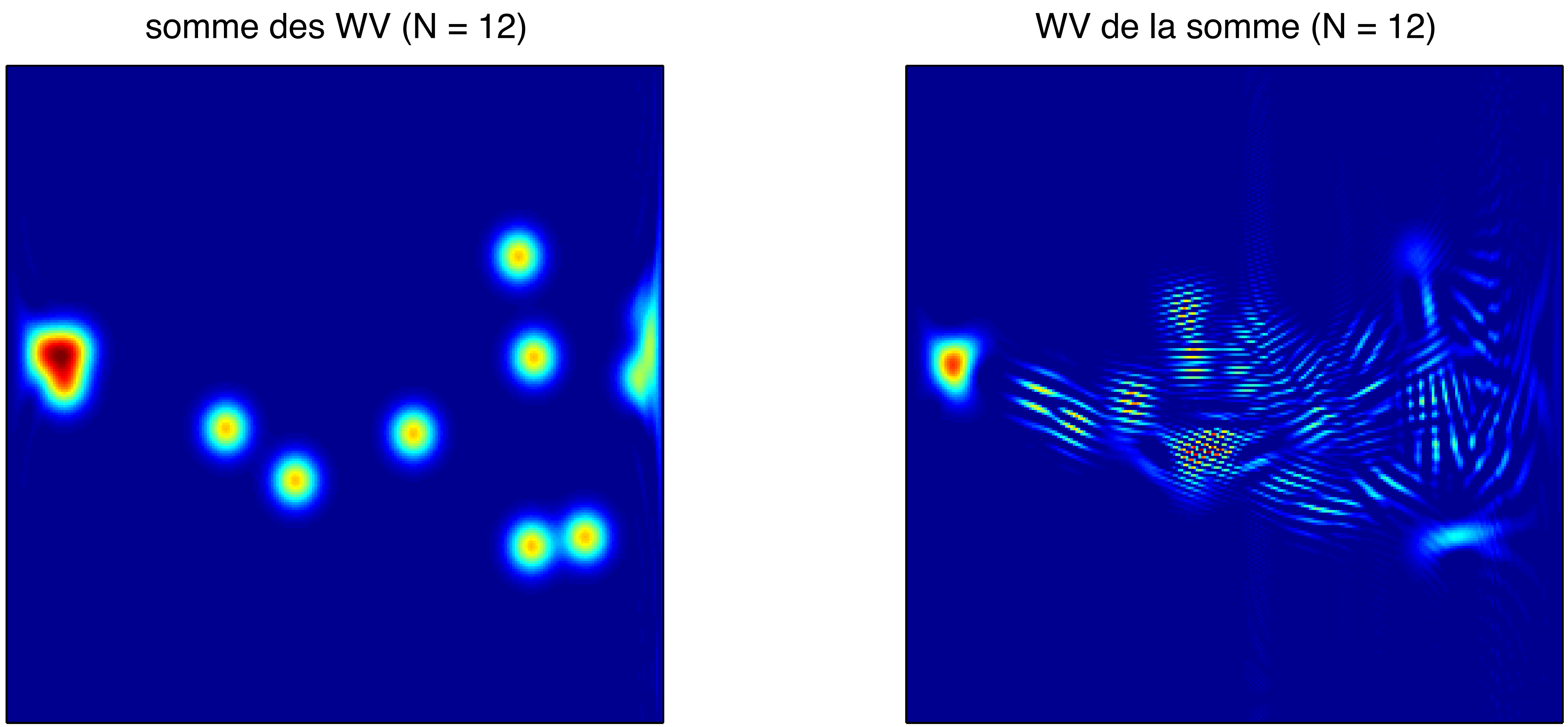
somme des WV (N = 9)

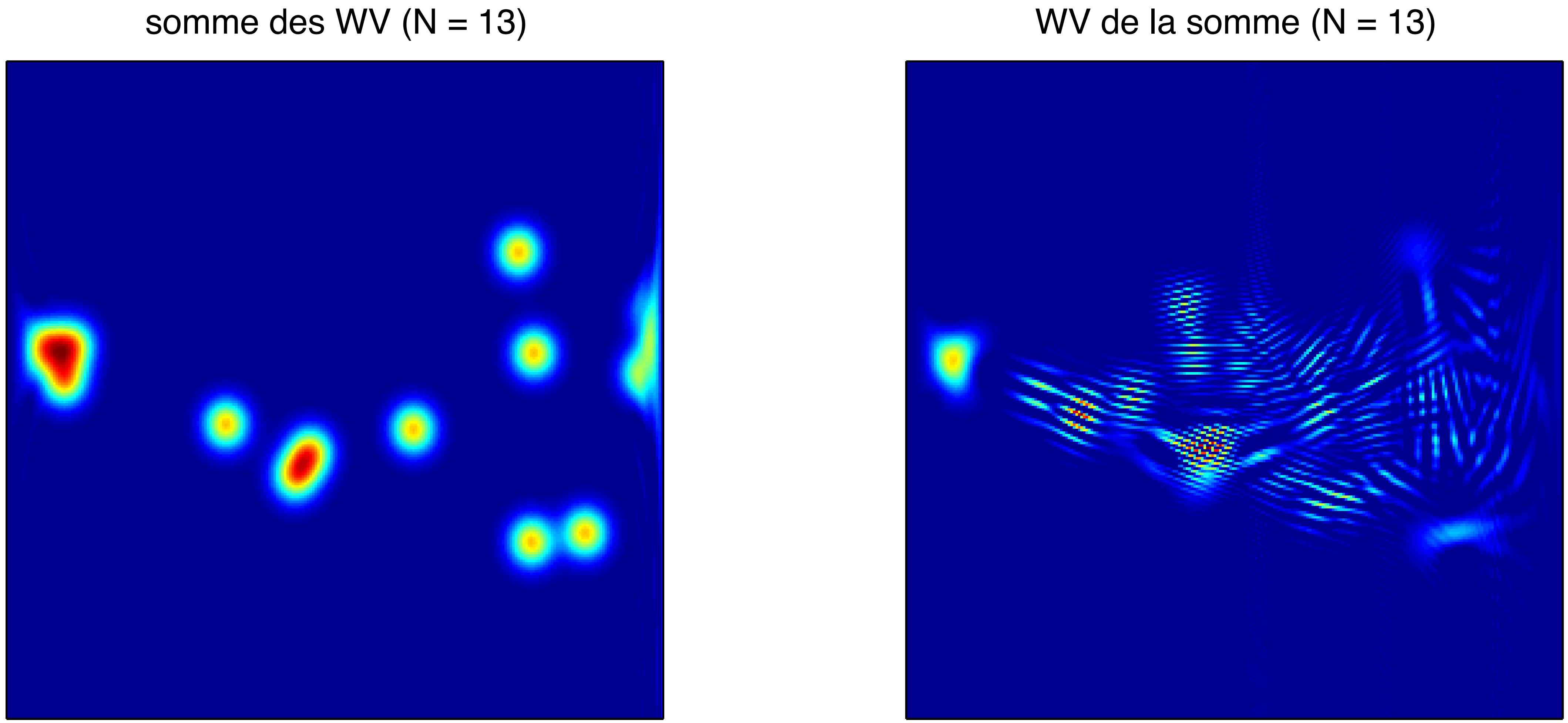


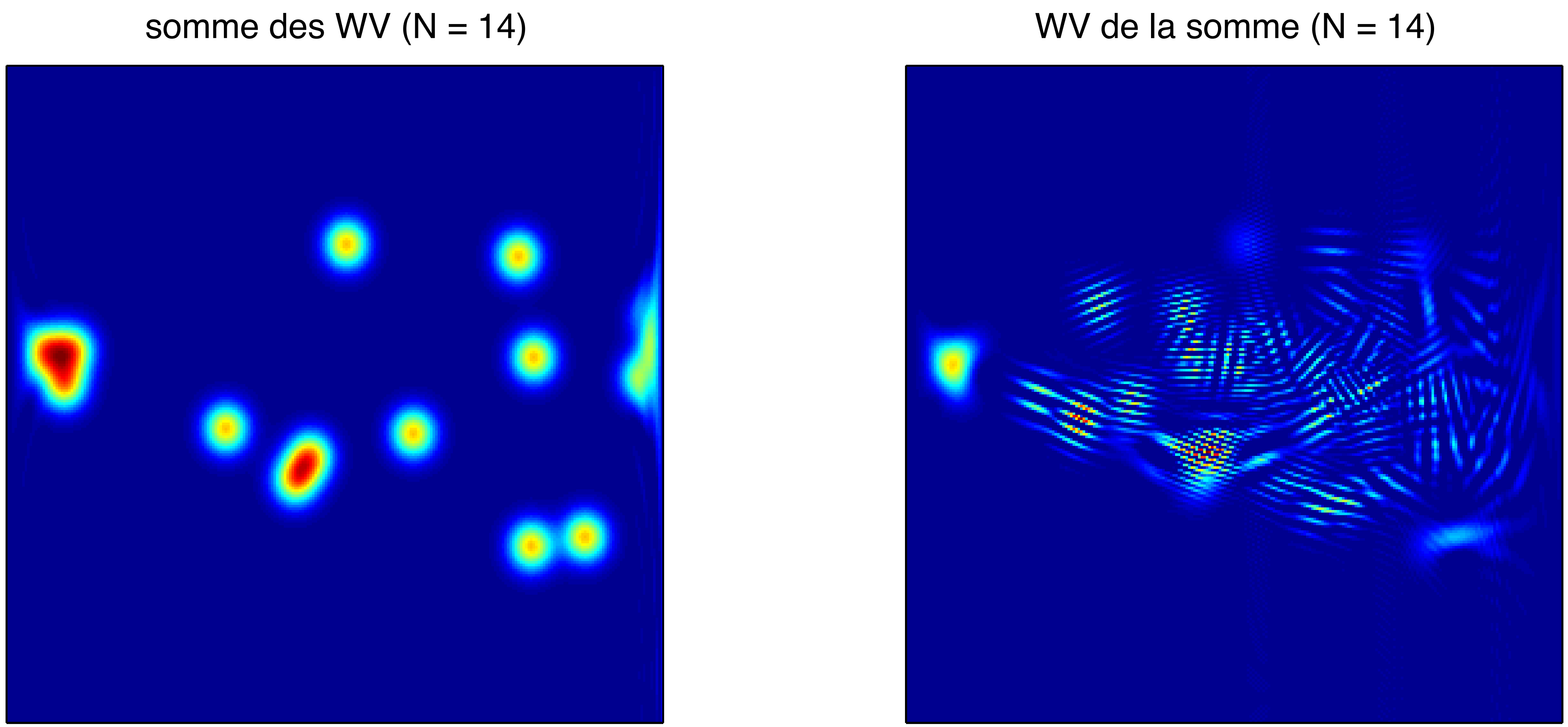
WV de la somme (N = 9)

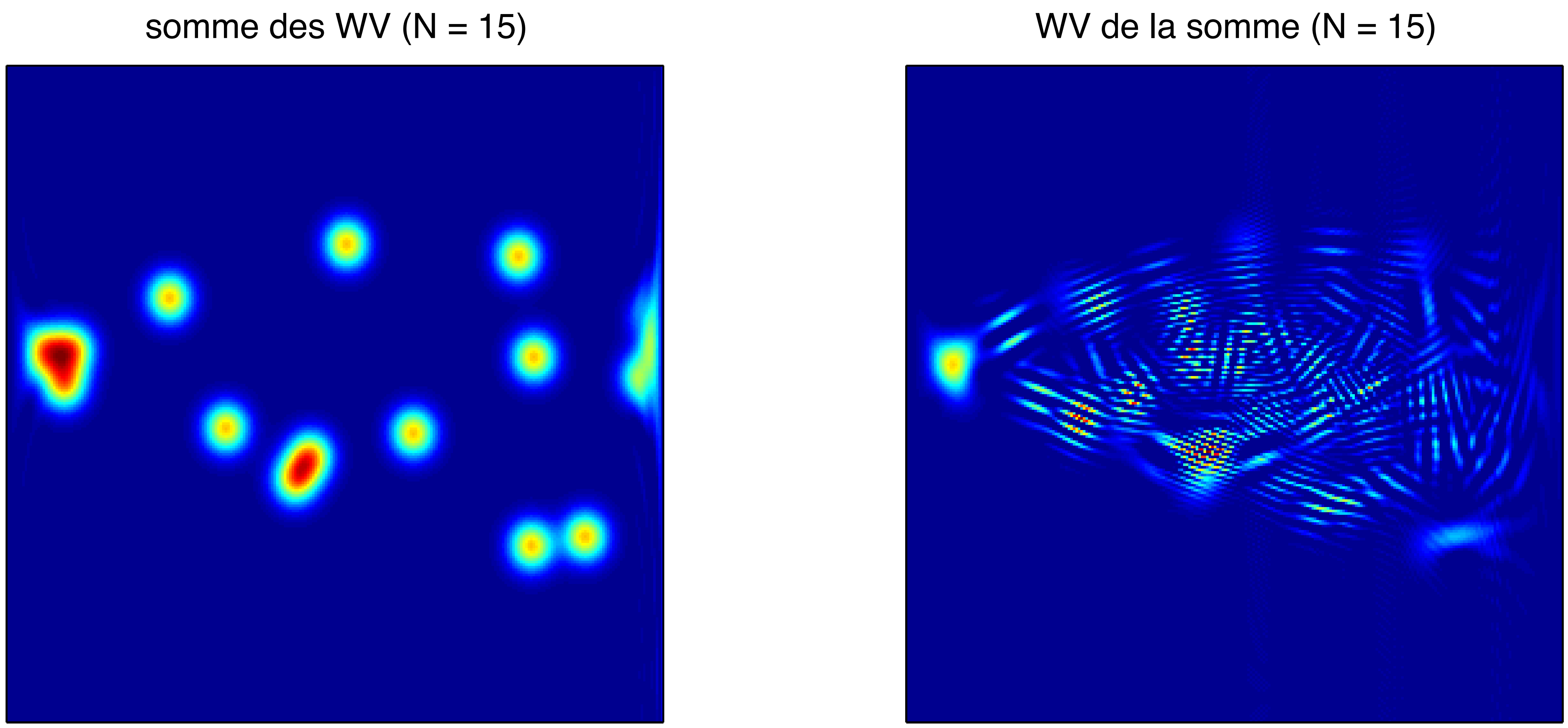



somme des WV (N = 11)
WV de la somme (N = 11)

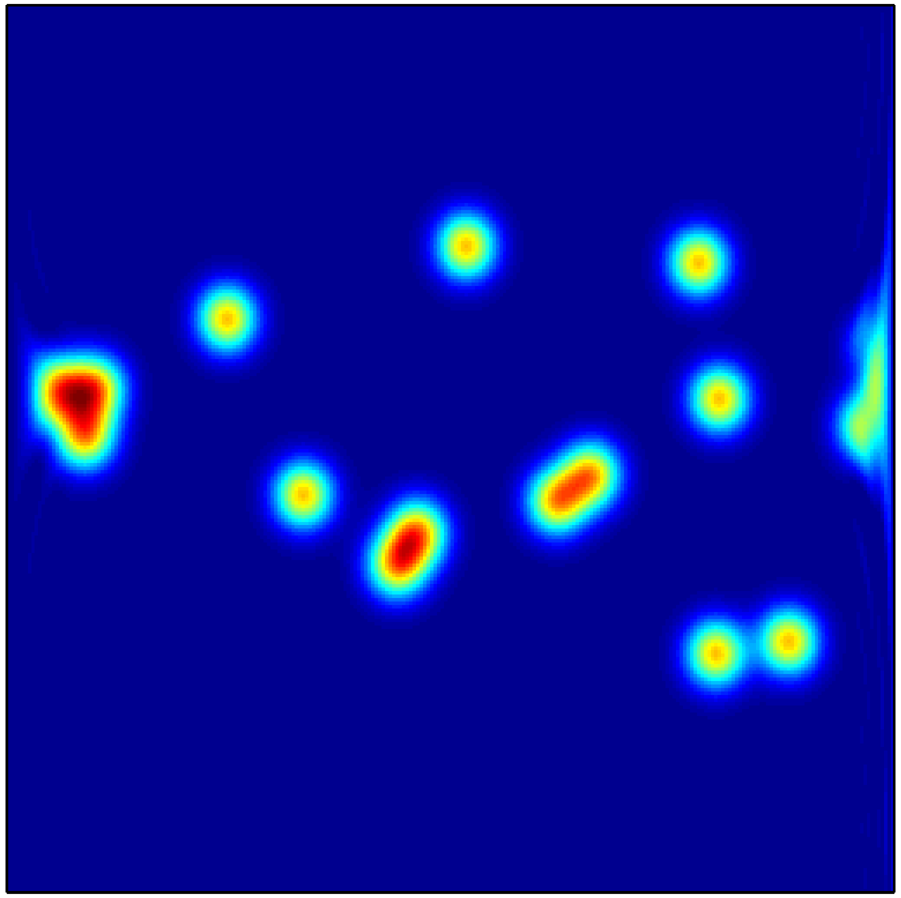

somme des WV (N = 12)
WV de la somme (N = 12)


somme des WV (N = 13)
WV de la somme (N = 13)

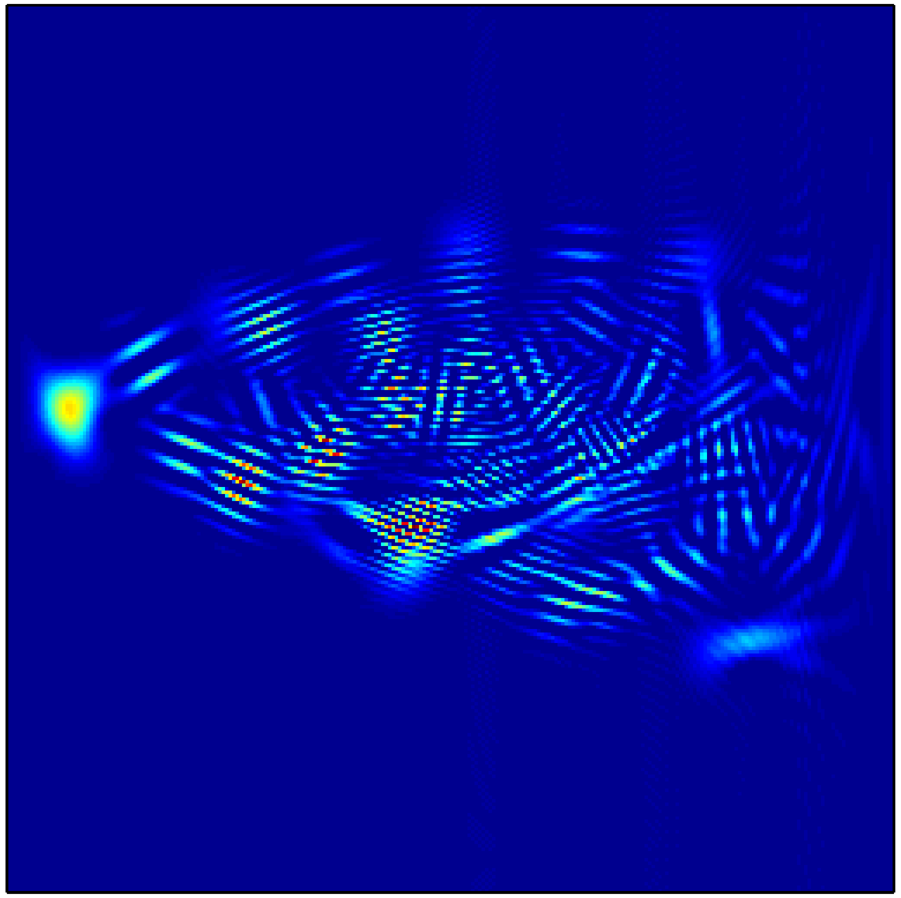

somme des WV (N = 14)
WV de la somme (N = 14)

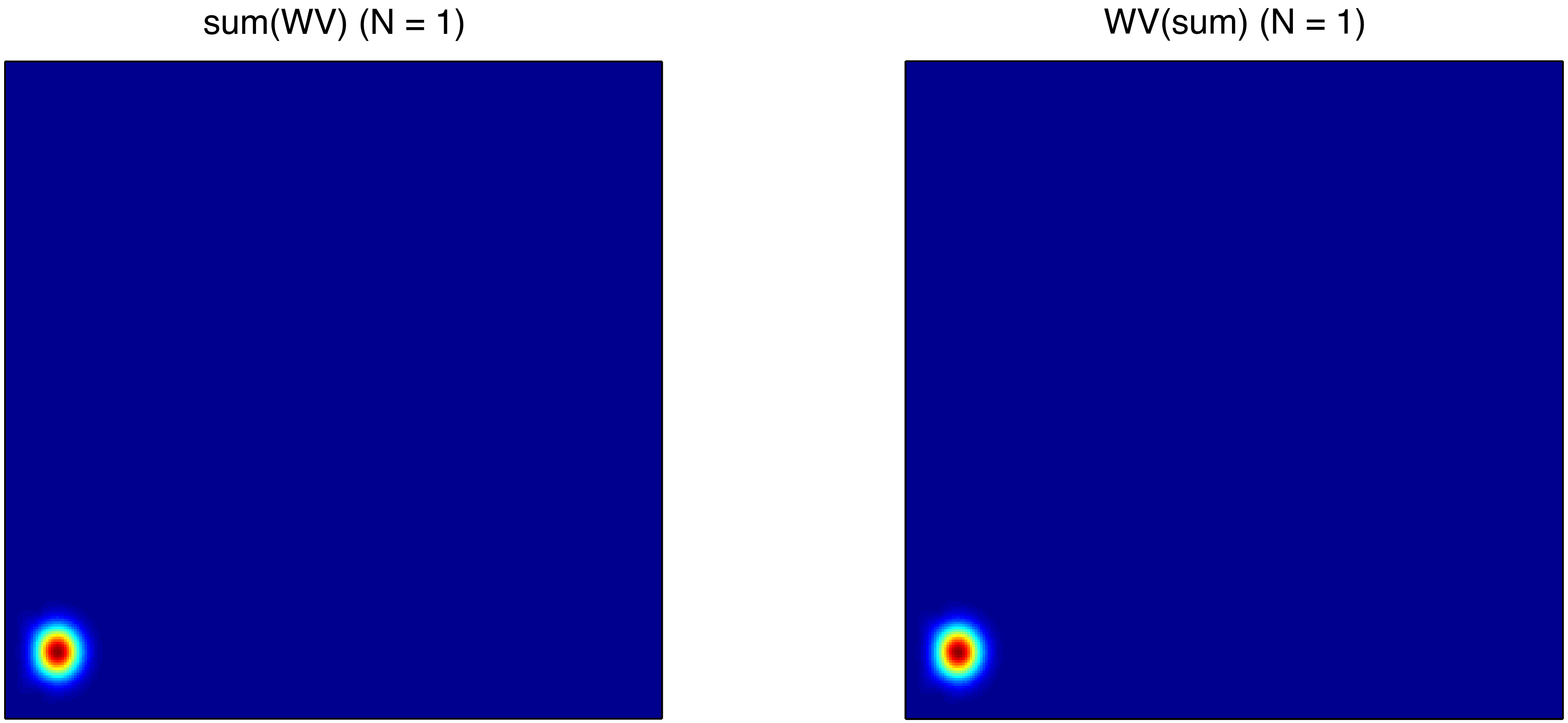

somme des WV (N = 15)
WV de la somme (N = 15)

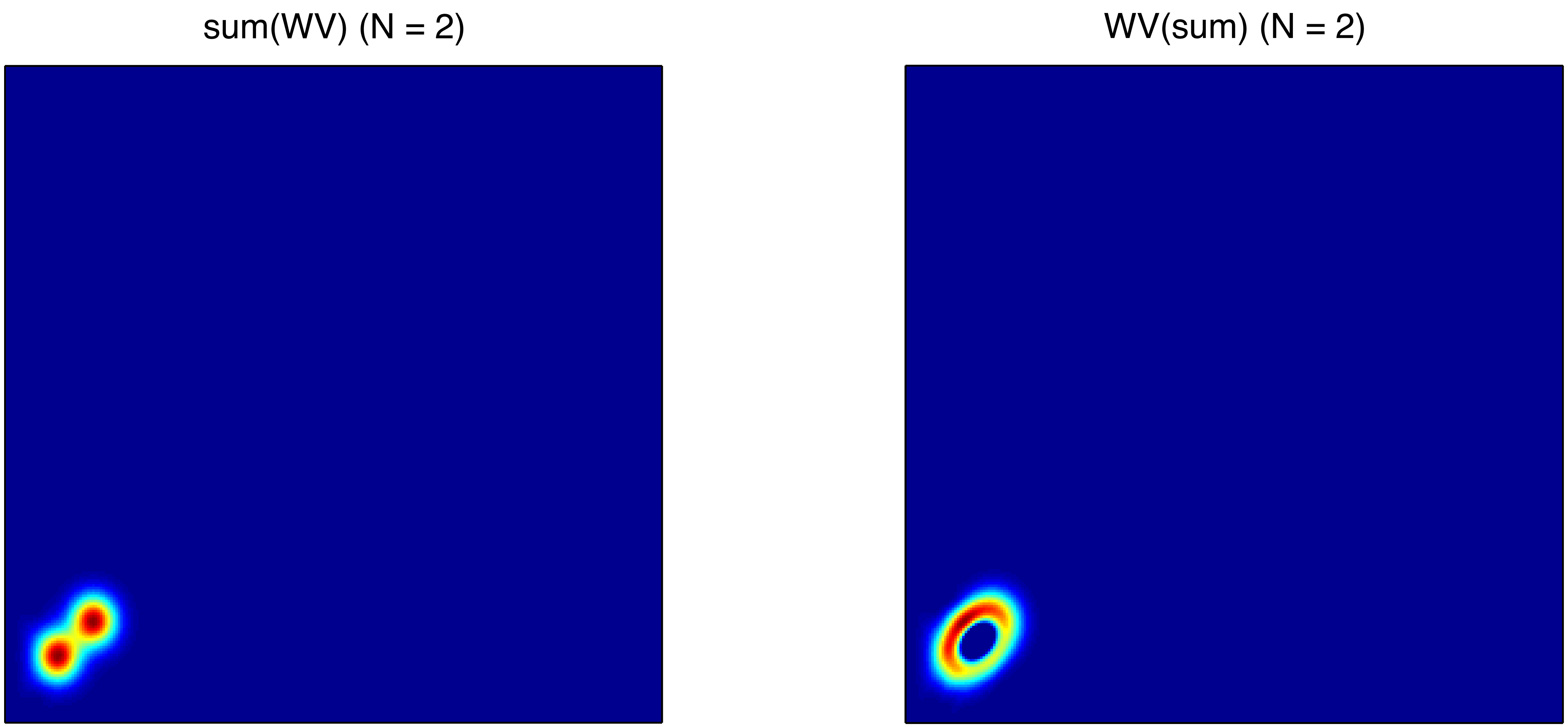
somme des WV (N = 16)



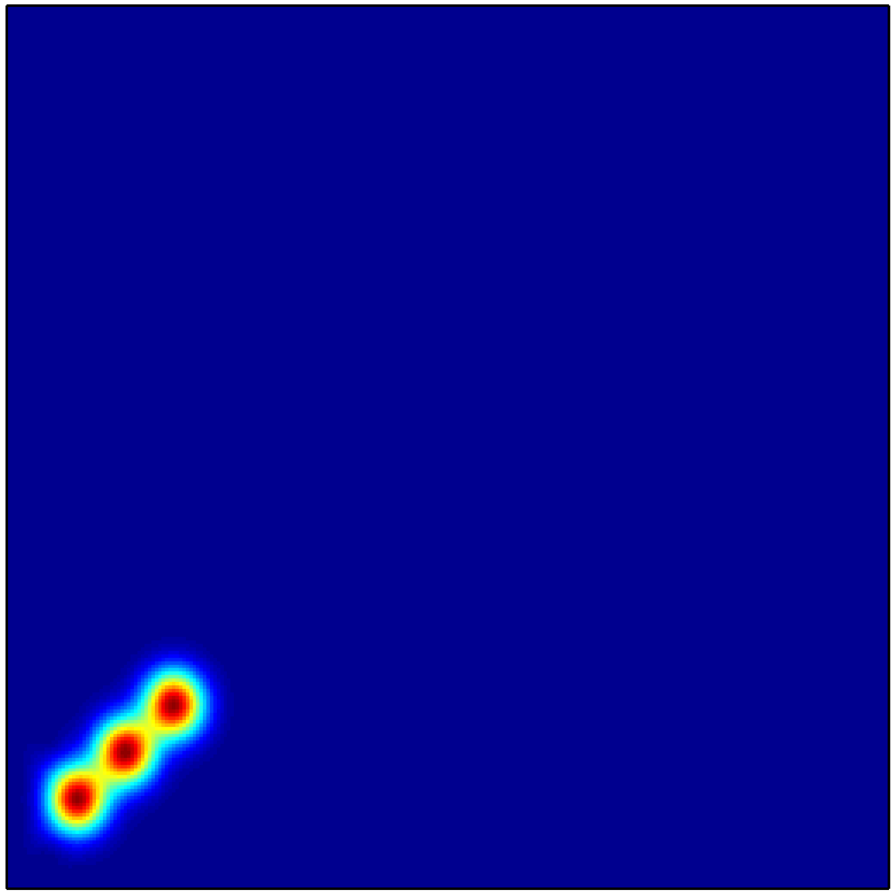
WV de la somme (N = 16)



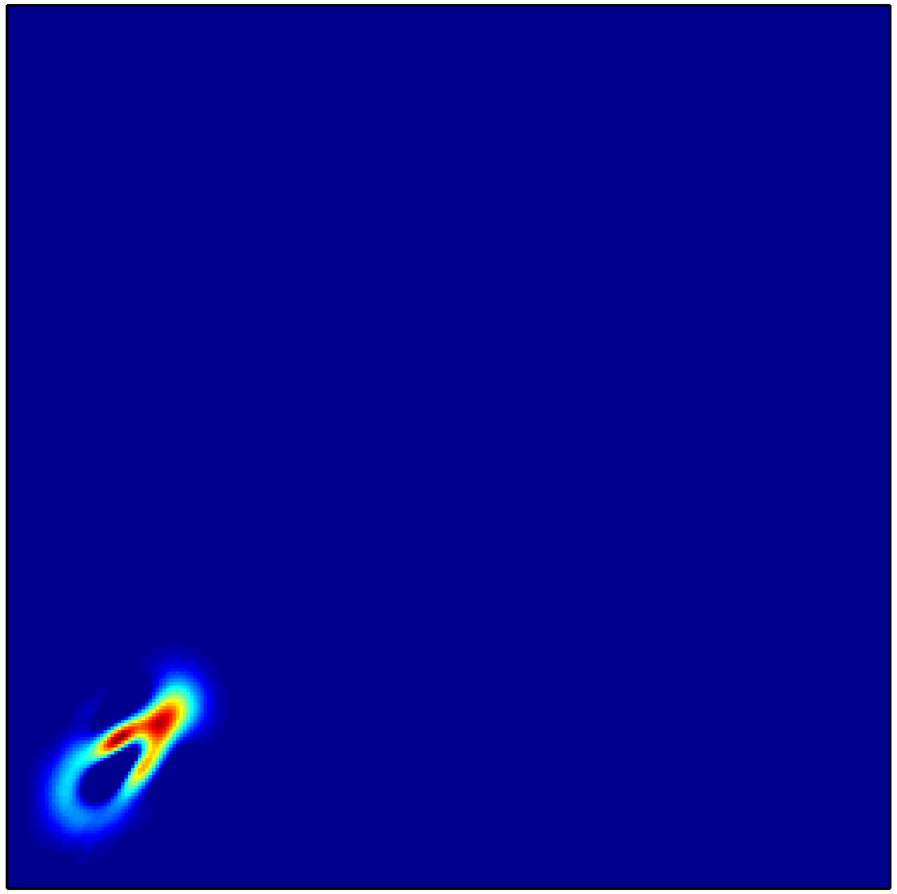

sum(WV) (N = 1)
WV(sum) (N = 1)

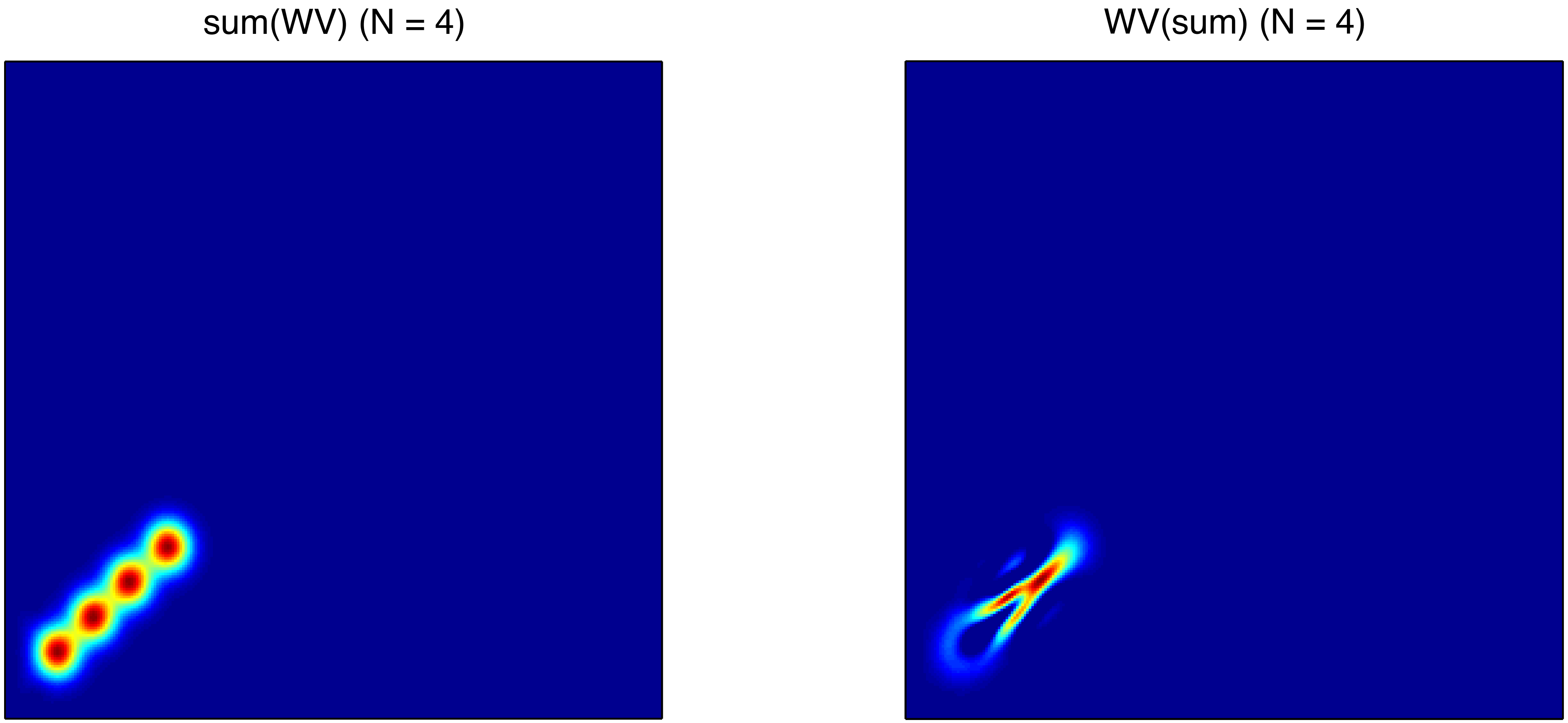

sum(WV) (N = 2)
WV(sum) (N = 2)

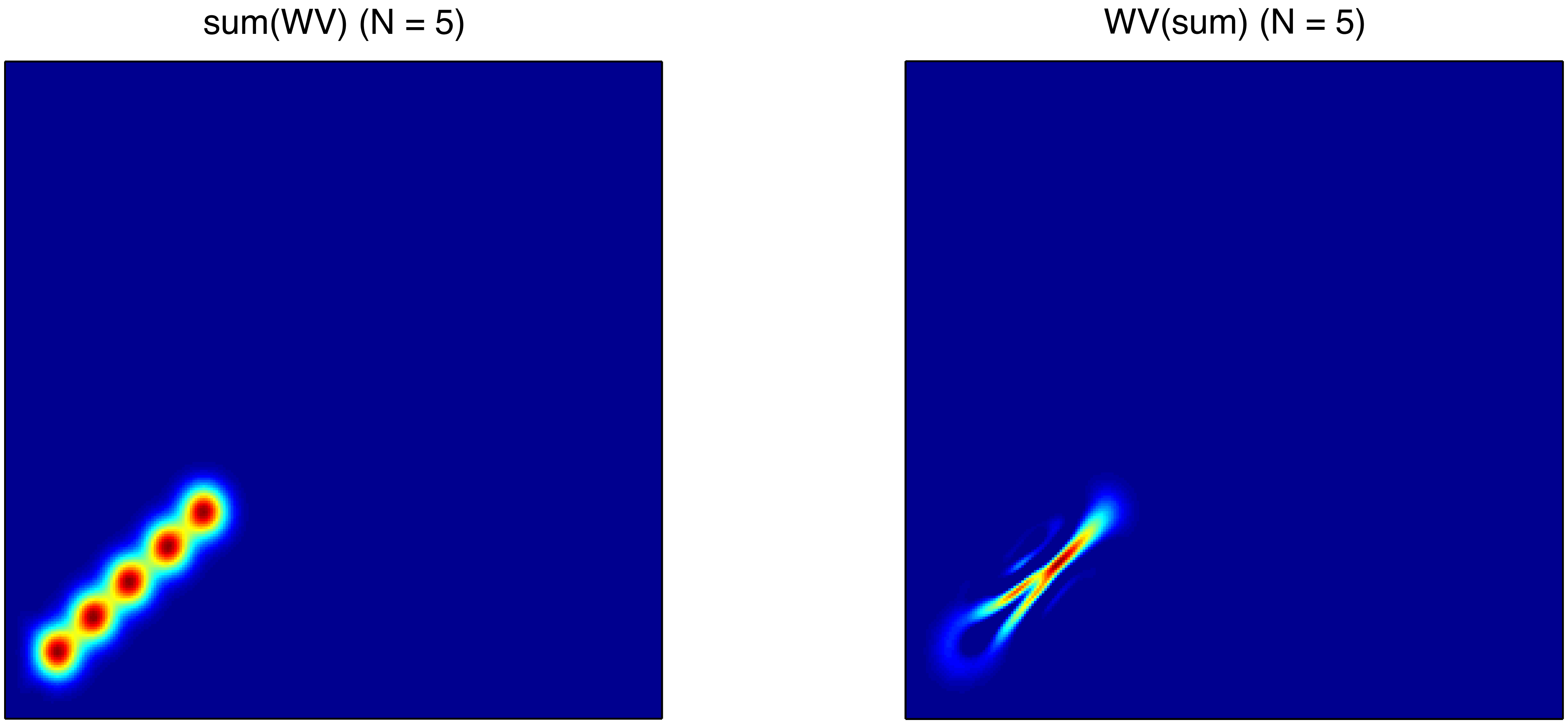
sum(WV) (N = 3)

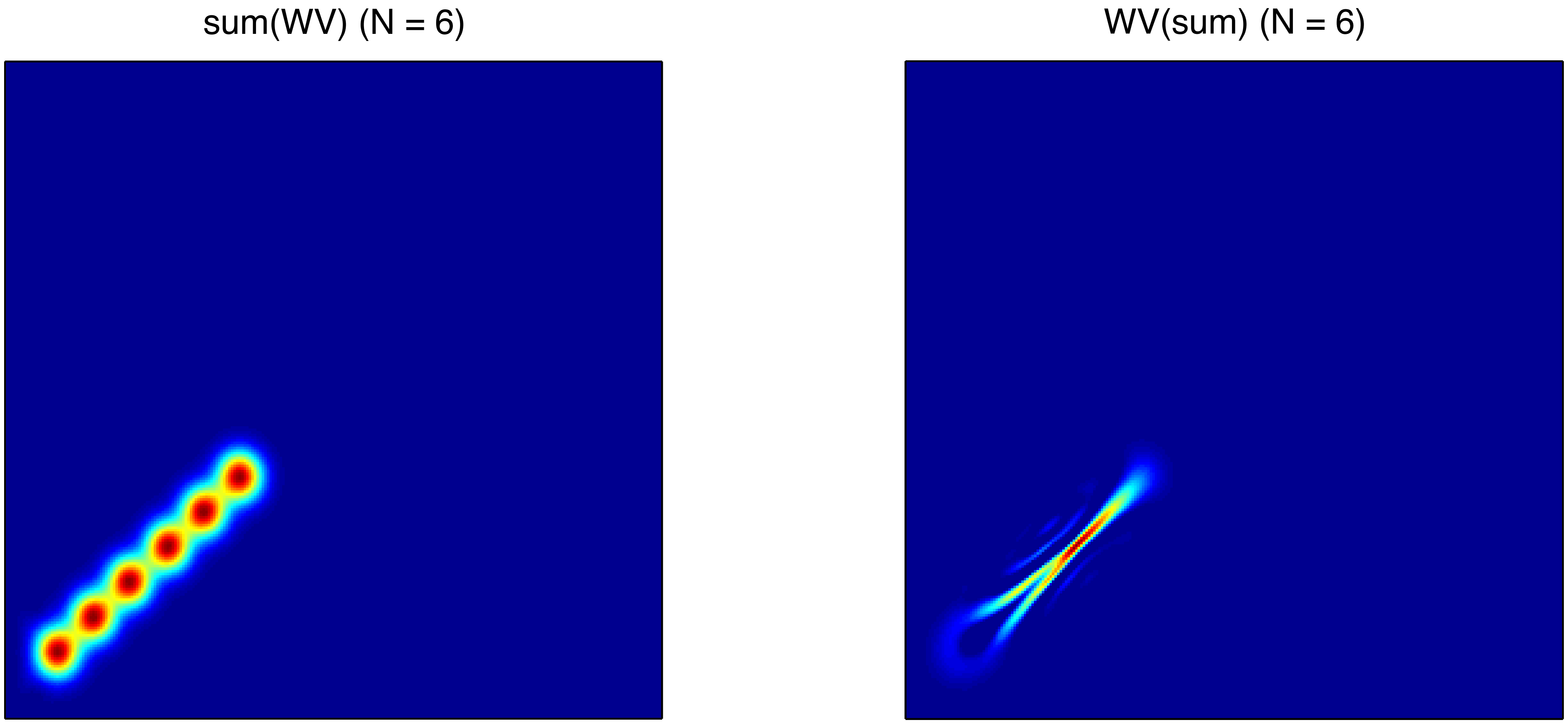


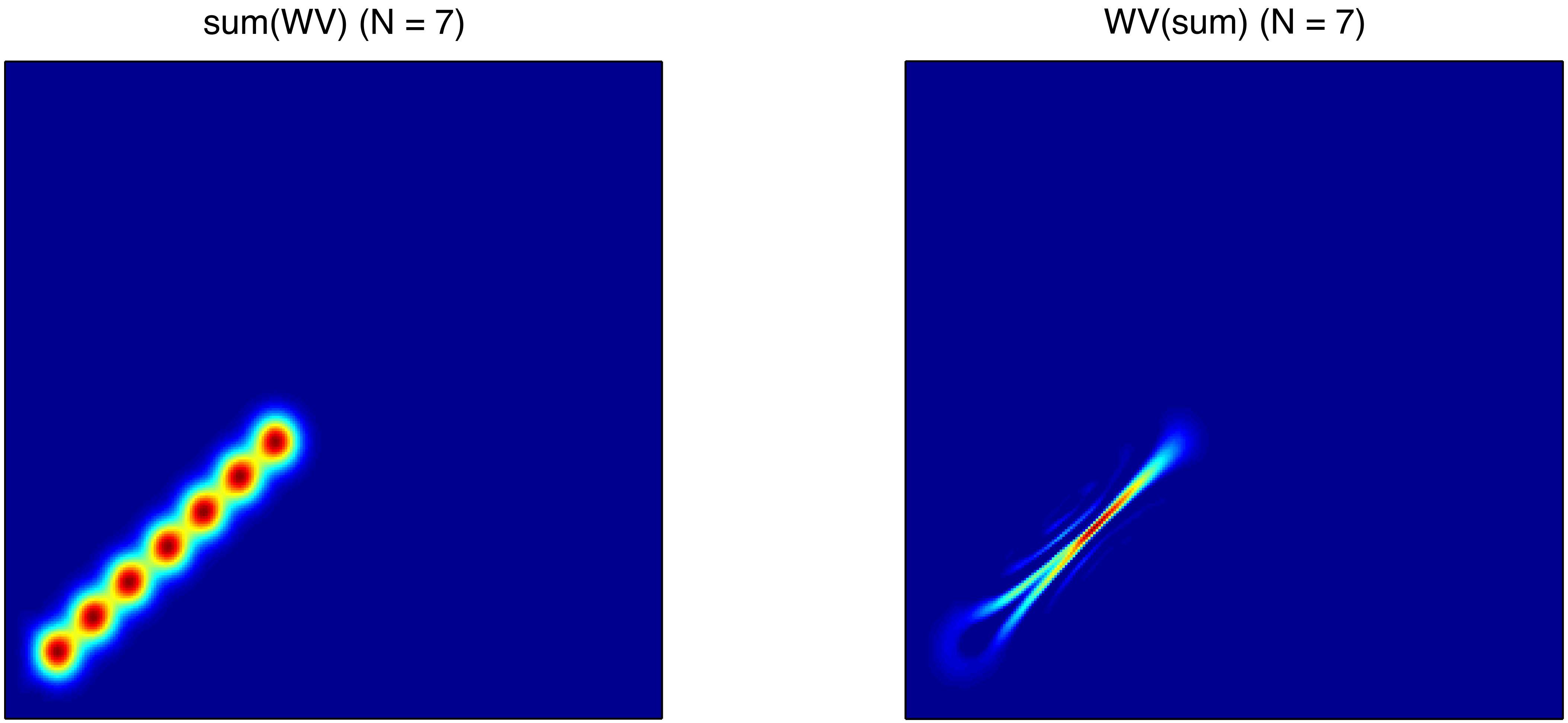
WV(sum) (N = 3)



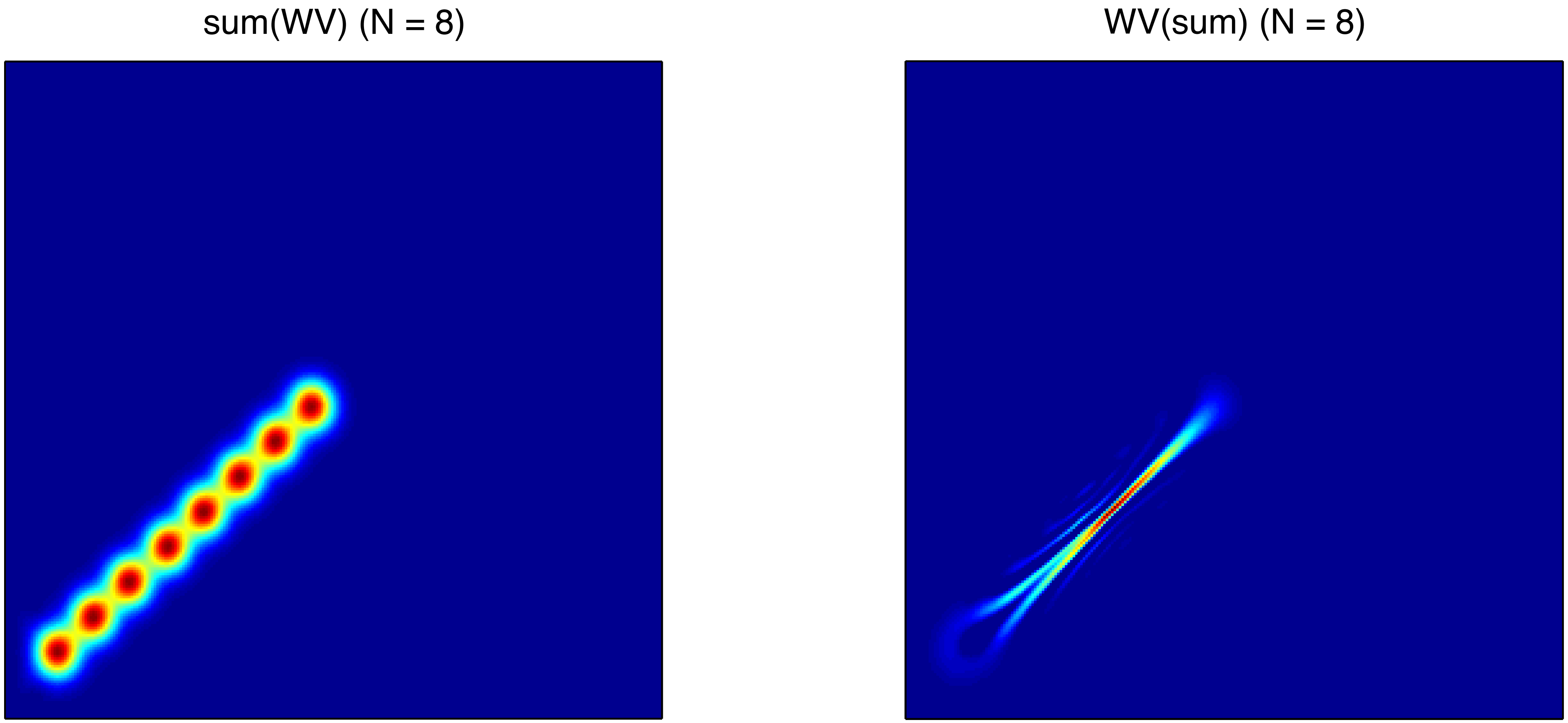

sum(WV) (N = 4)
WV(sum) (N = 4)

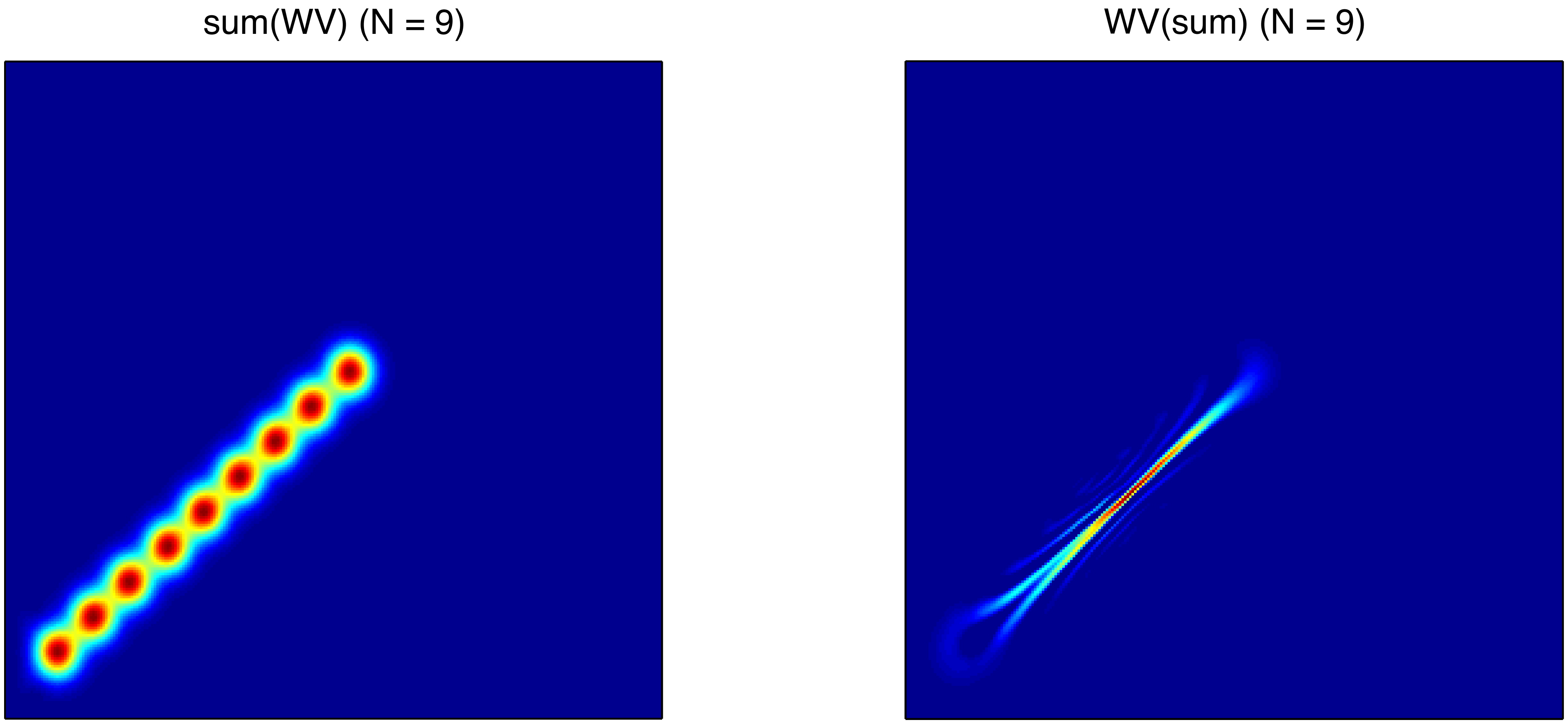

sum(WV) (N = 5)
WV(sum) (N = 5)

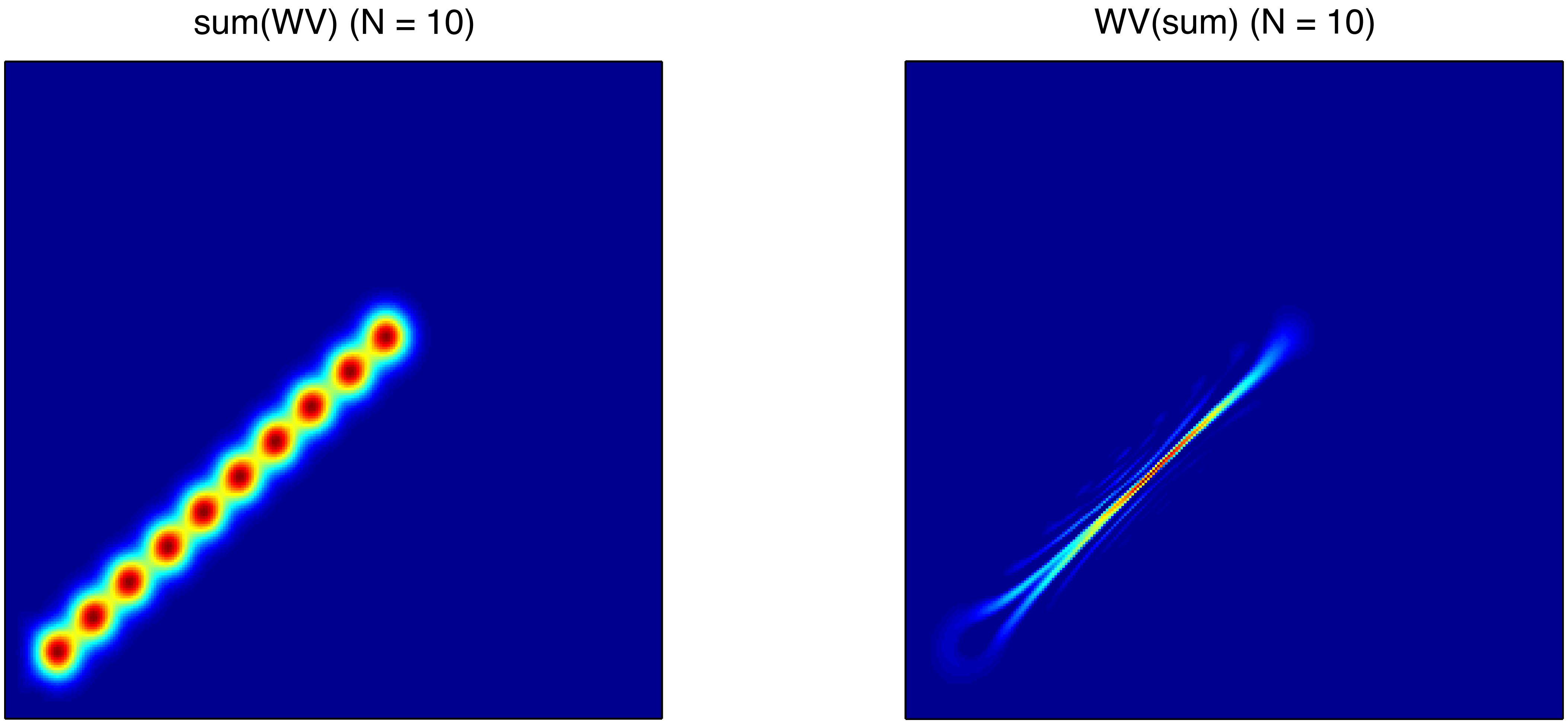

sum(WV) (N = 6)
WV(sum) (N = 6)

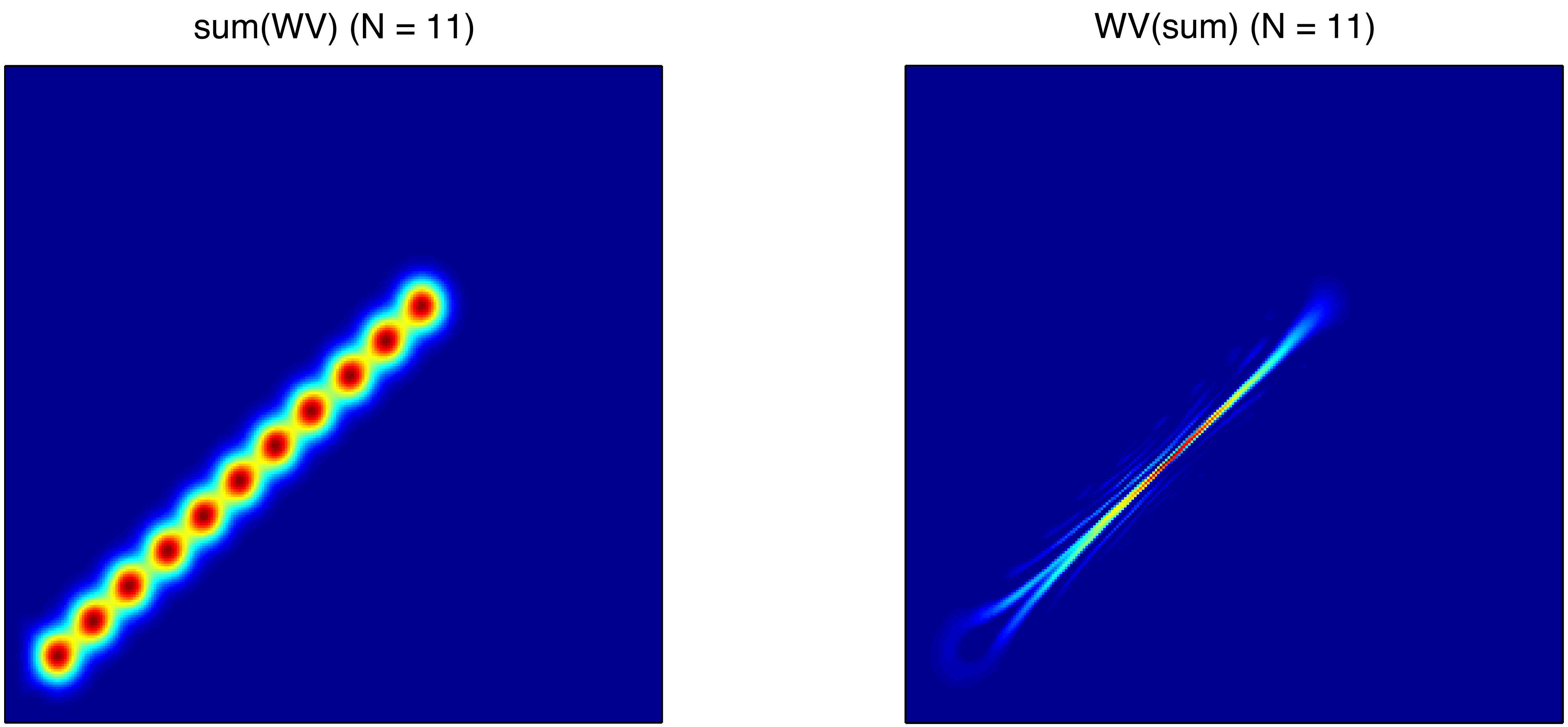


sum(WV) (N = 7)
WV(sum) (N = 7)

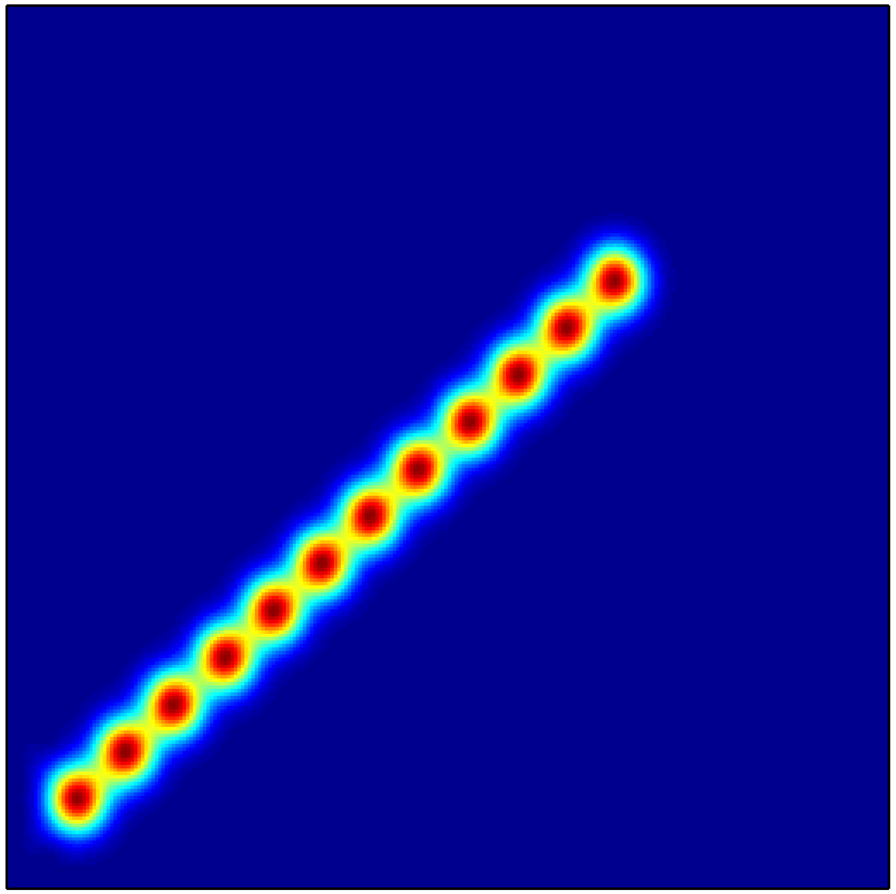

sum(WV) (N = 8)
WV(sum) (N = 8)


sum(WV) (N = 9)
WV(sum) (N = 9)

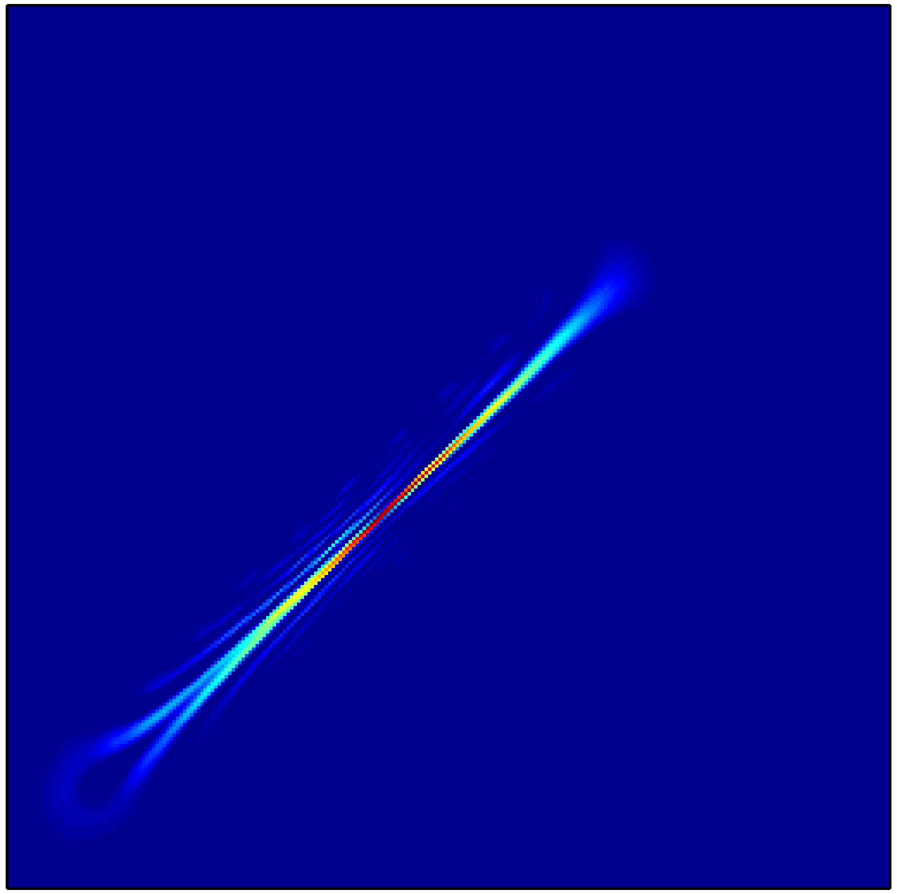

sum(WV) (N = 10)
WV(sum) (N = 10)


sum(WV) (N = 11)
WV(sum) (N = 11)

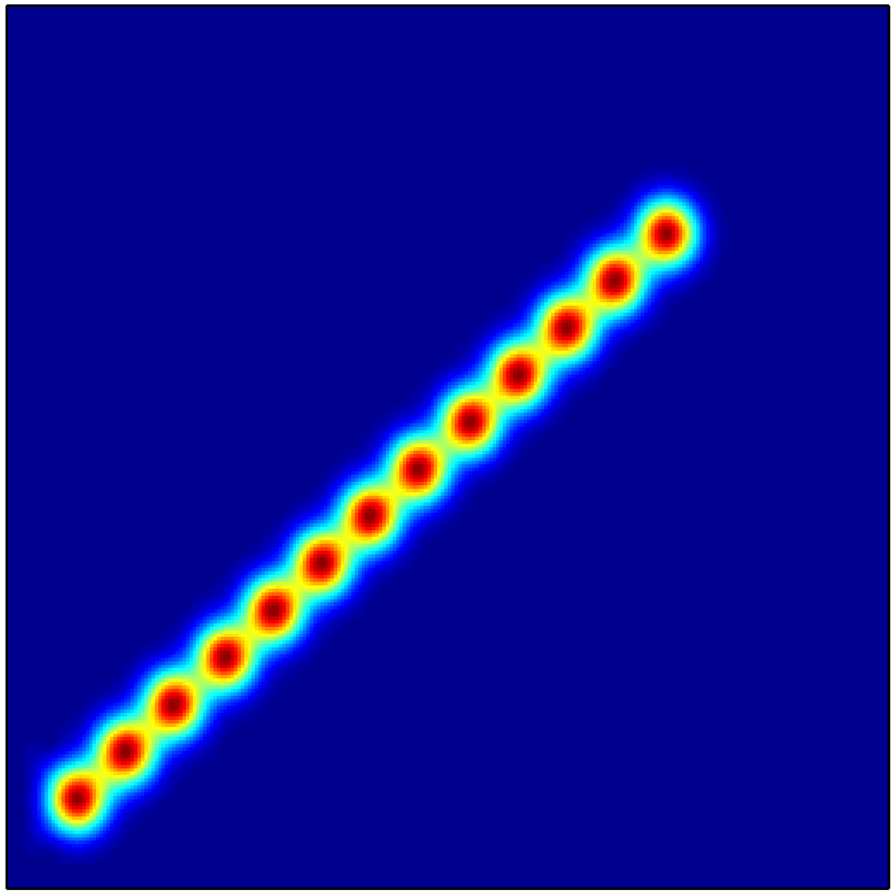
sum(WV) (N = 12)



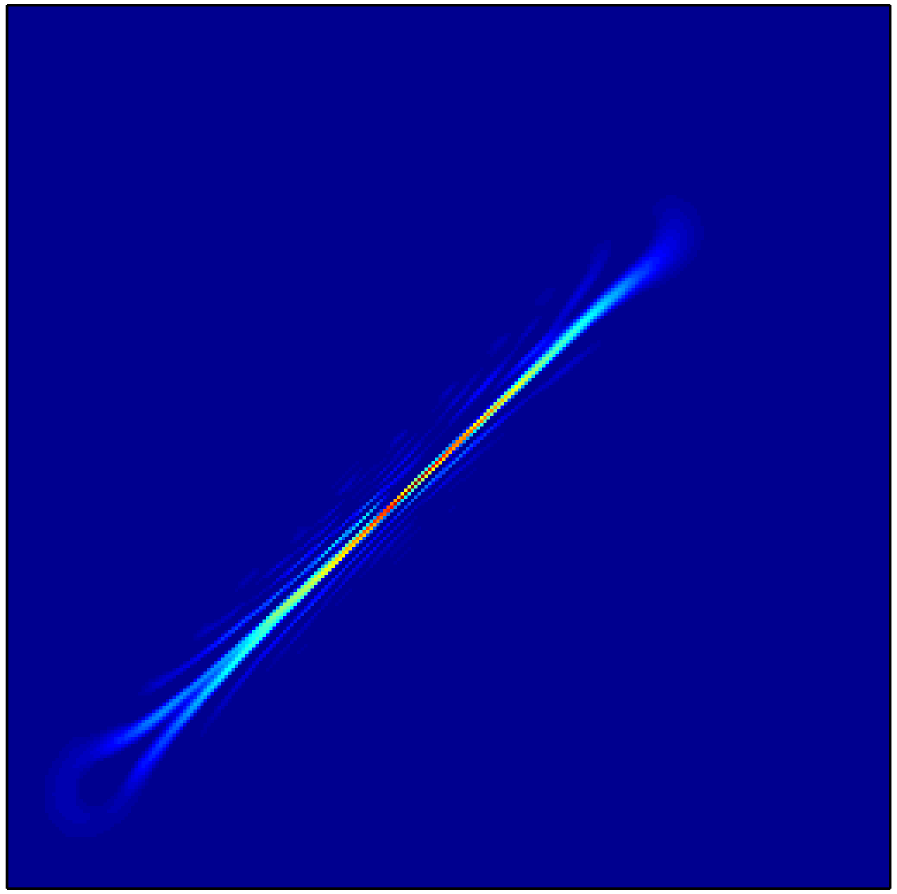
WV(sum) (N = 12)

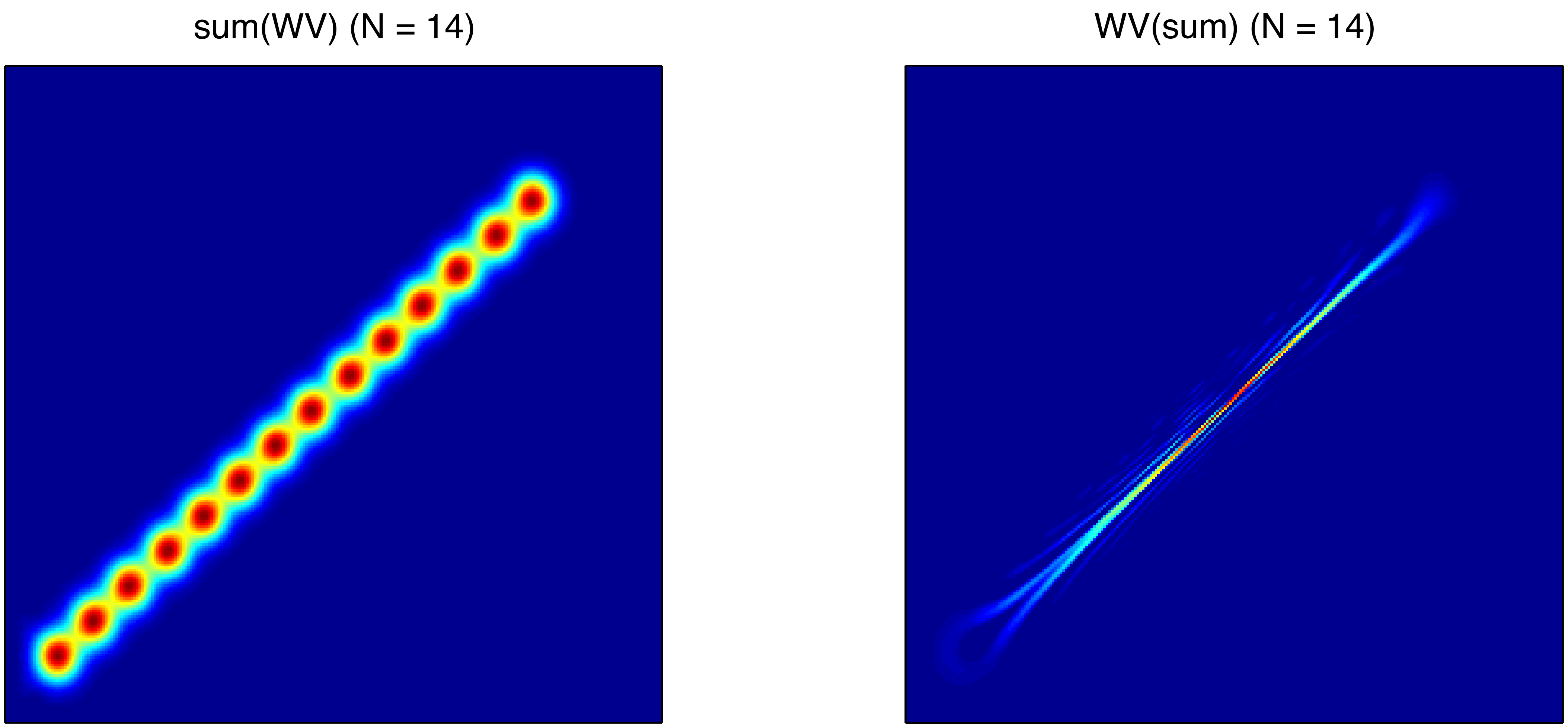


sum(WV) (N = 13)

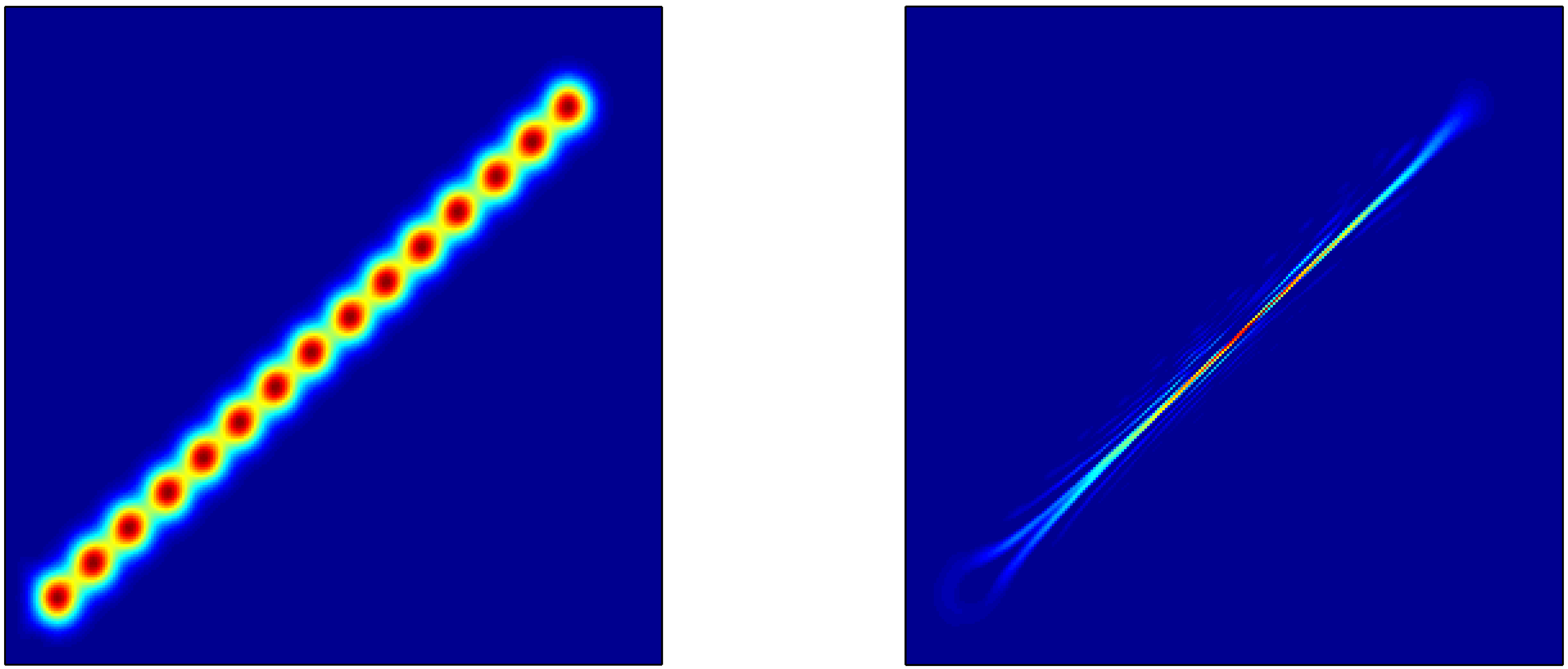


WV(sum) (N = 13)

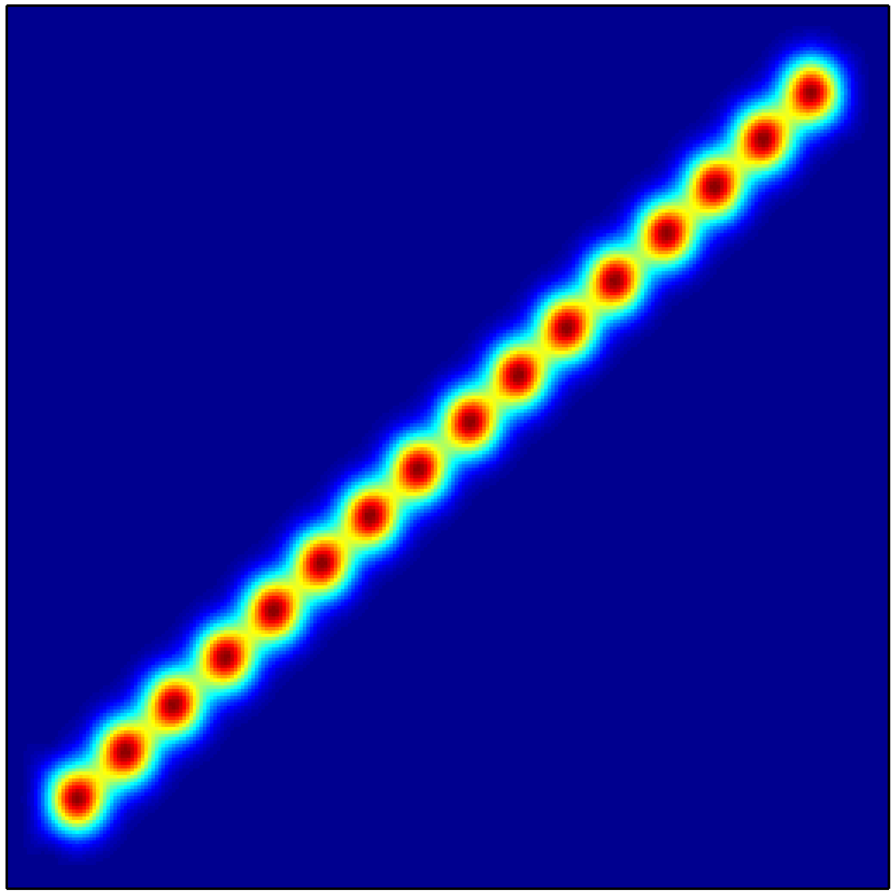



sum(WV) (N = 14)
WV(sum) (N = 14)

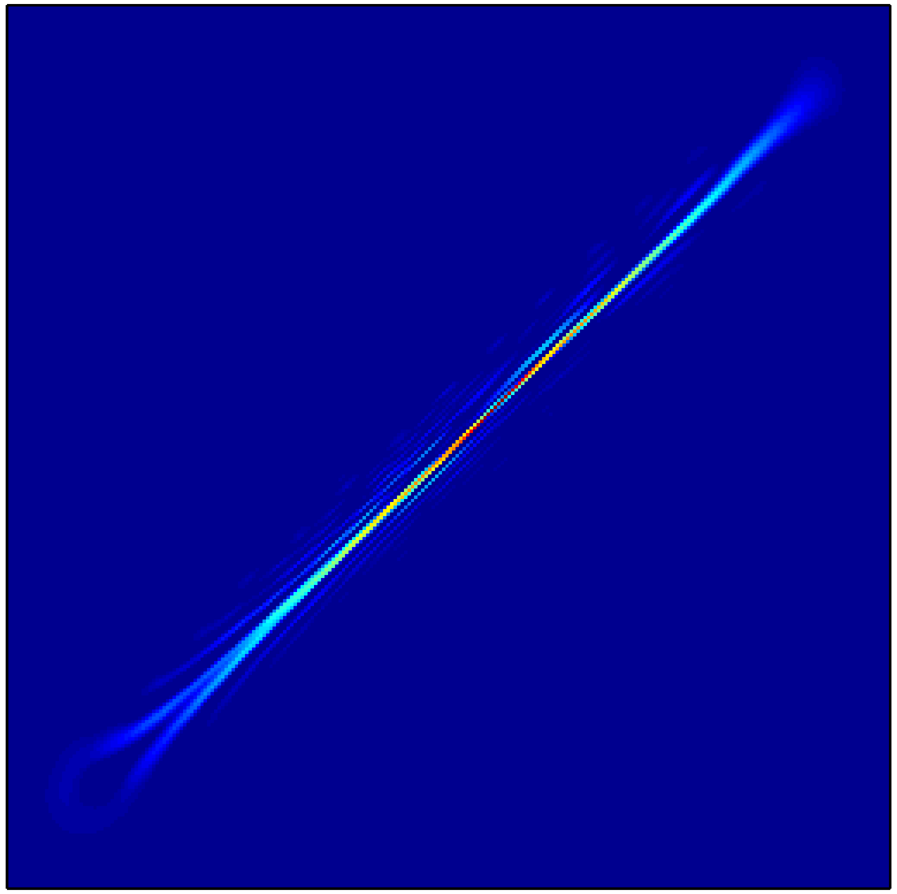
sum(WV) (N = 15) WV(sum) (N = 15)



sum(WV) (N = 16)



WV(sum) (N = 16)



Revisiting the spectrogram (1)

- Classically, « spectrogram = **squared STFT** »

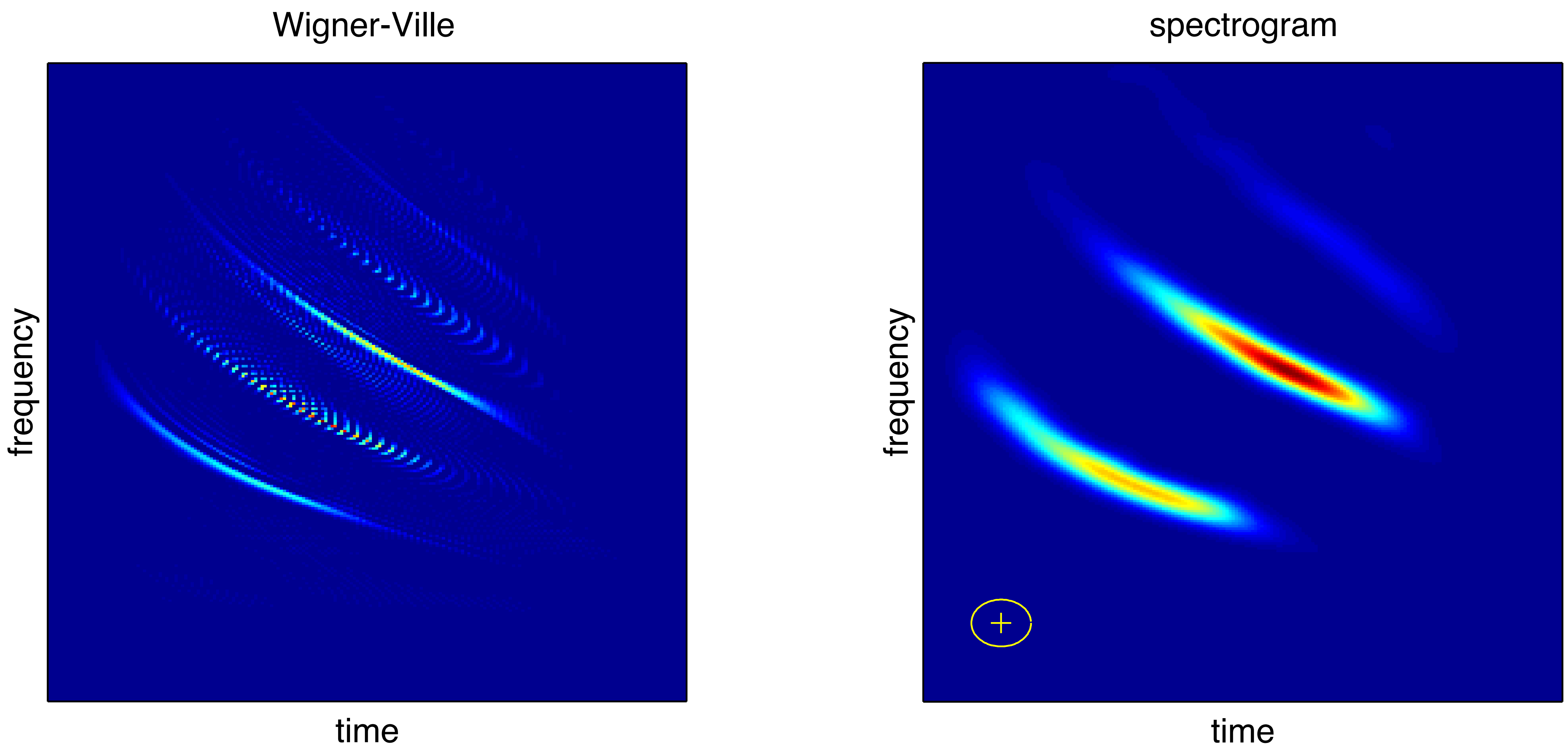
$$S_x^{(h)}(t,f) = \left| F_x^{(h)}(t,f) \right|^2$$

- As an alternative, « spectrogram = **smoothed Wigner-Ville** »

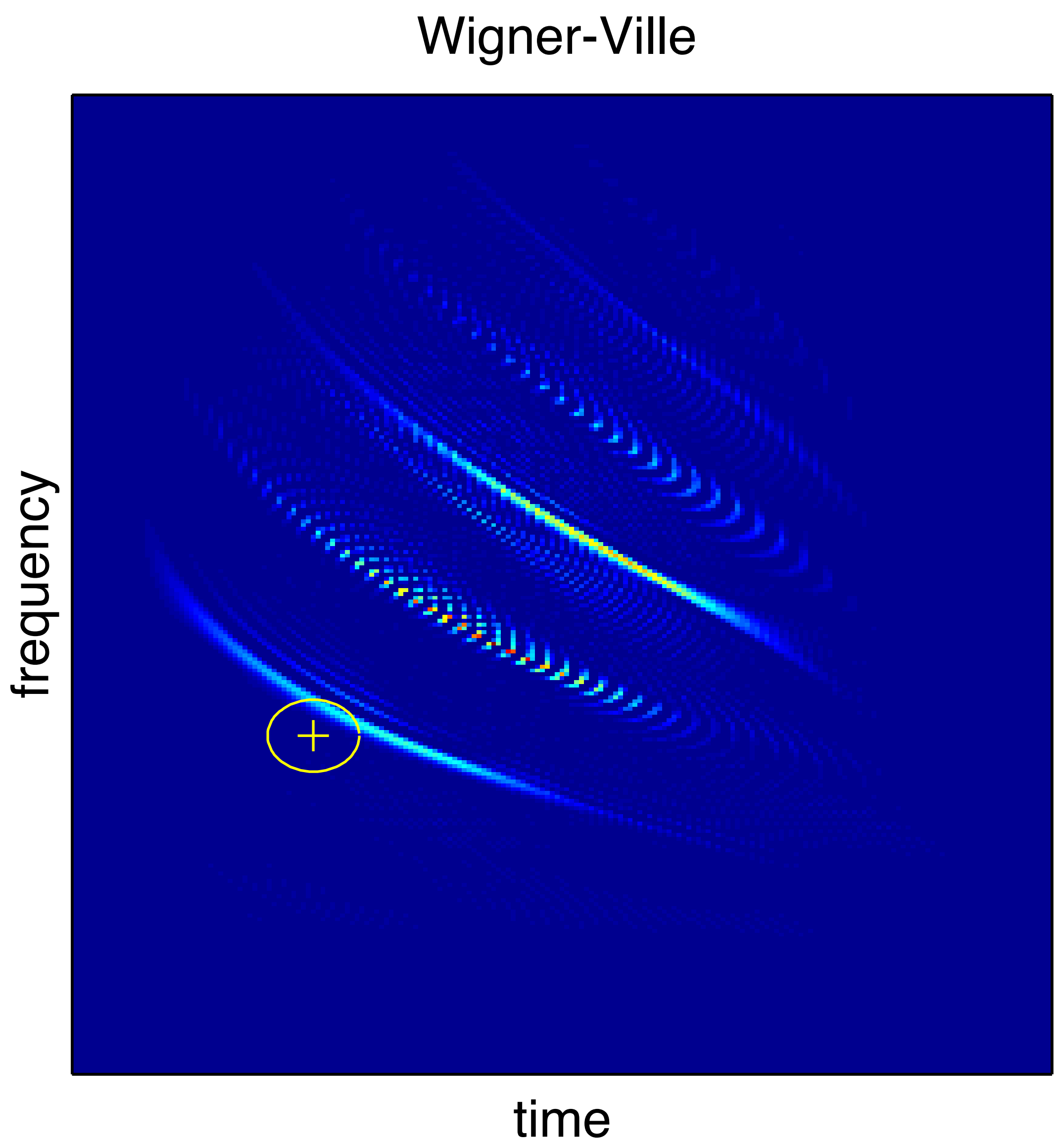
$$S_x^{(h)}(t,f) = \iint W_x(s,\xi)\, W_h(s-t,\xi-f)\, ds\, d\xi$$

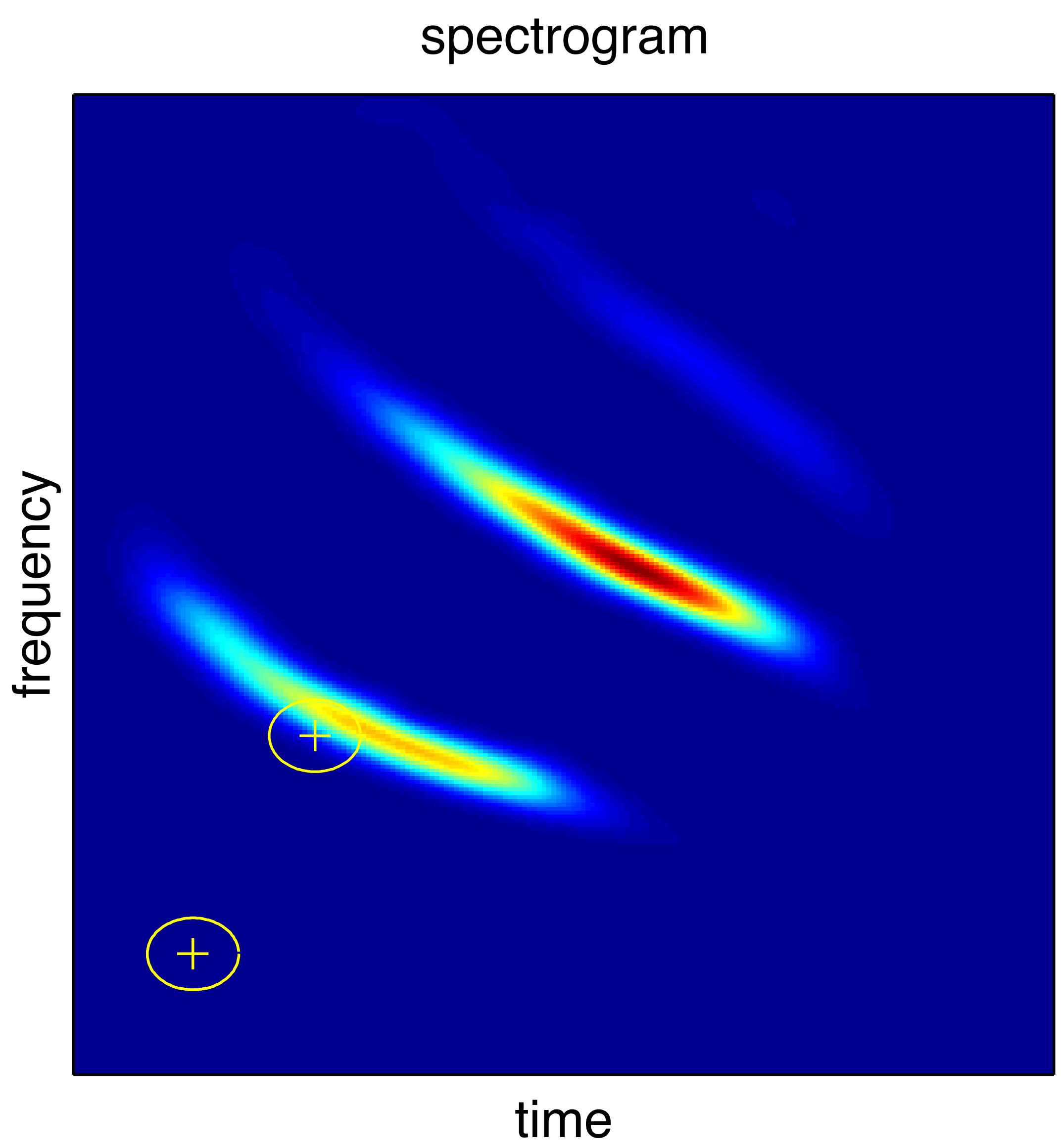
- Mathematical equivalence, but **different physical interpretations**

« Heisenberg » smoothing

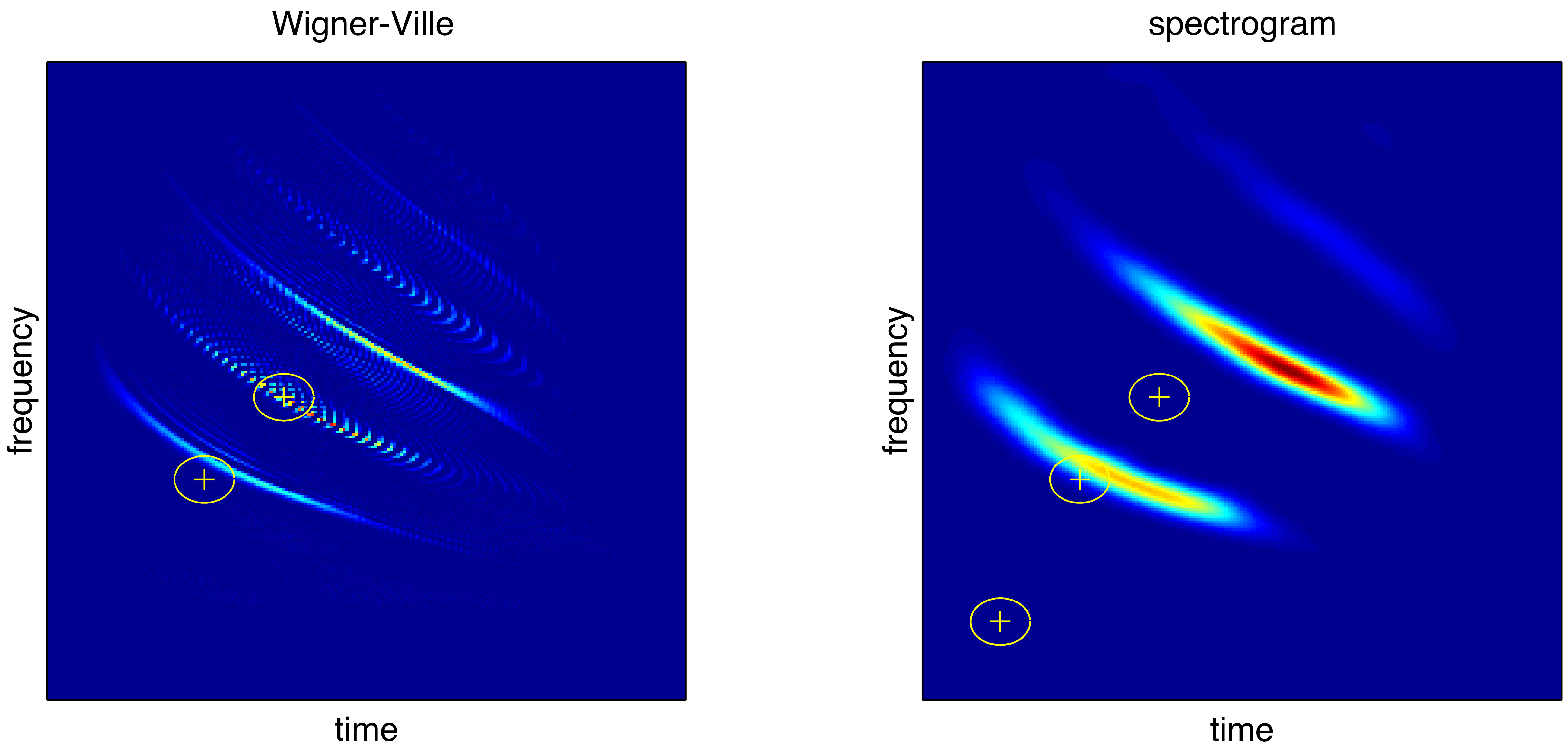


Smearing of signal components :-(





Smoothing out of interferences :-)



Revisiting the spectrogram (2)

$$S_x^{(h)}(t,f) = \iint W_x(s,\xi)\, W_h(s-t,\xi-f)\, ds\, d\xi$$

- Smoothing = **summing up** local contributions
- A mechanical analogy:
  - *replace a distribution of mass within a domain by* ***one number*** *(the total mass)*
  - *unless uniform distribution,* ***no meaning*** *for the geometrical center of the domain*
  - *summarized information best attached to the* ***center of mass***

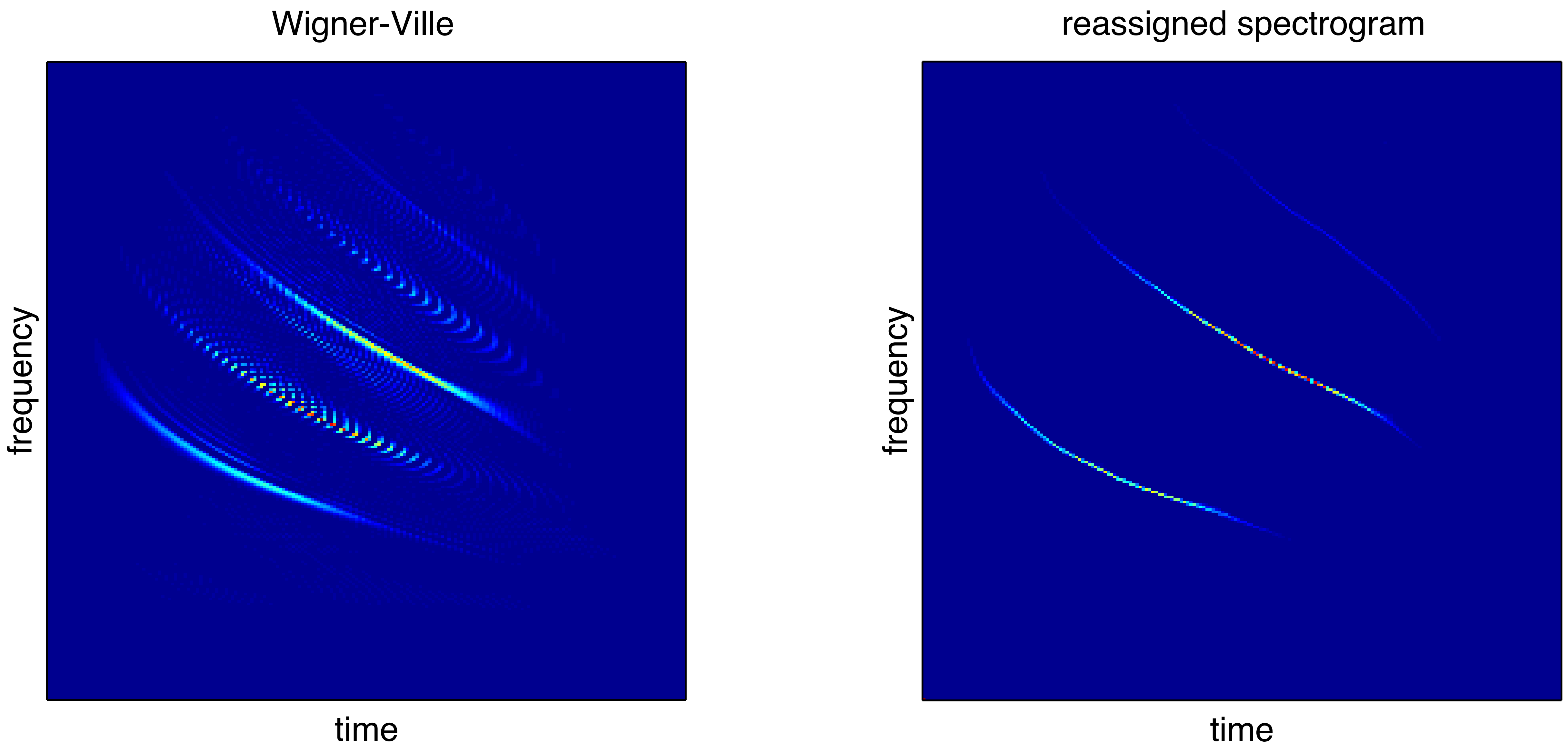
Reassignment principle

1. **Compute** STFT and spectrogram
2. **Identify** local centroids
3. **Reassign** value to this location

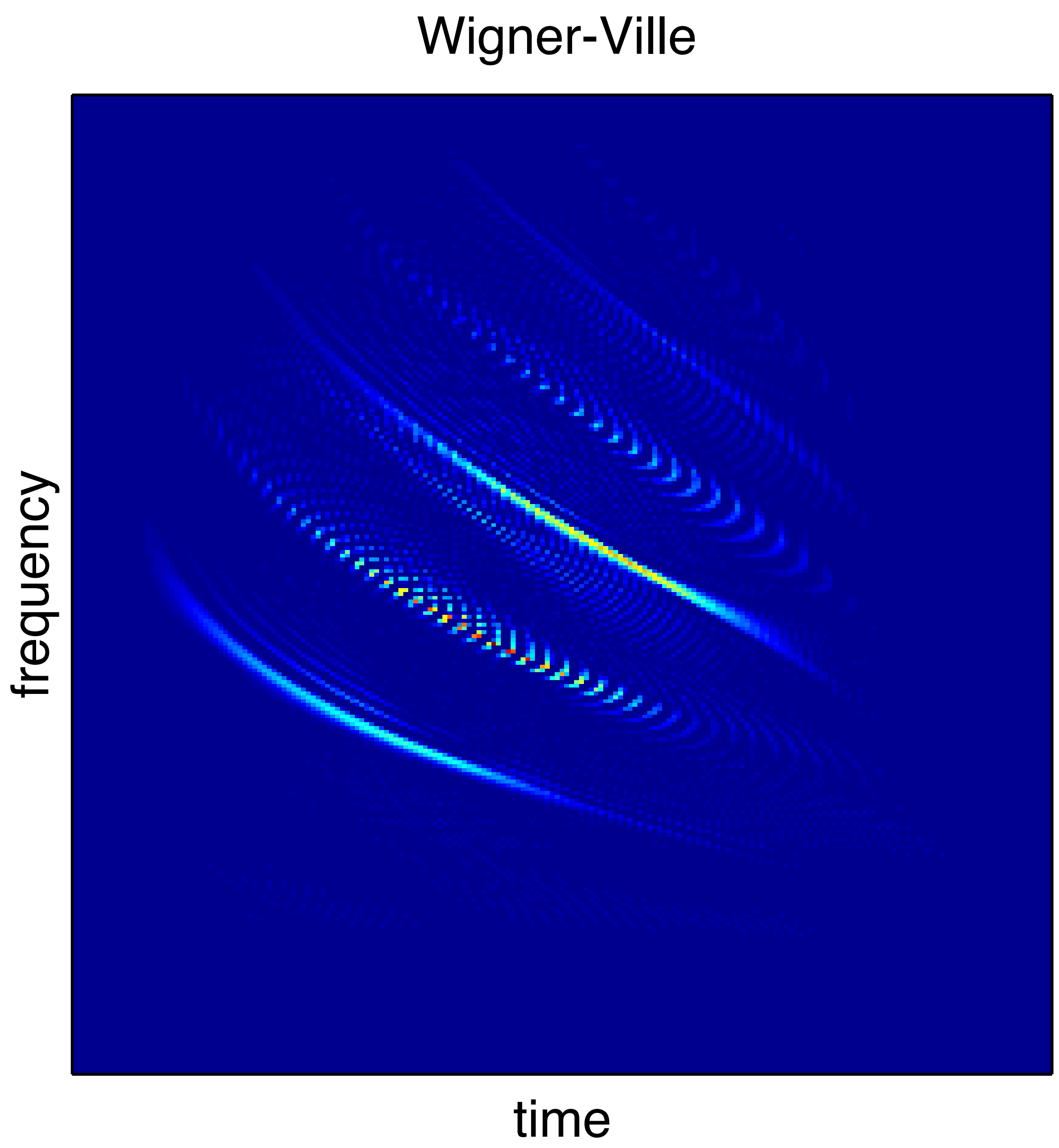
$$S_x^{(h)}(t,f) \mapsto \iint S_x^{(h)}(s,\xi)\, \delta\left(t - \hat{t}_x(s,\xi), f - \hat{f}_x(s,\xi)\right)\, ds\, d\xi$$

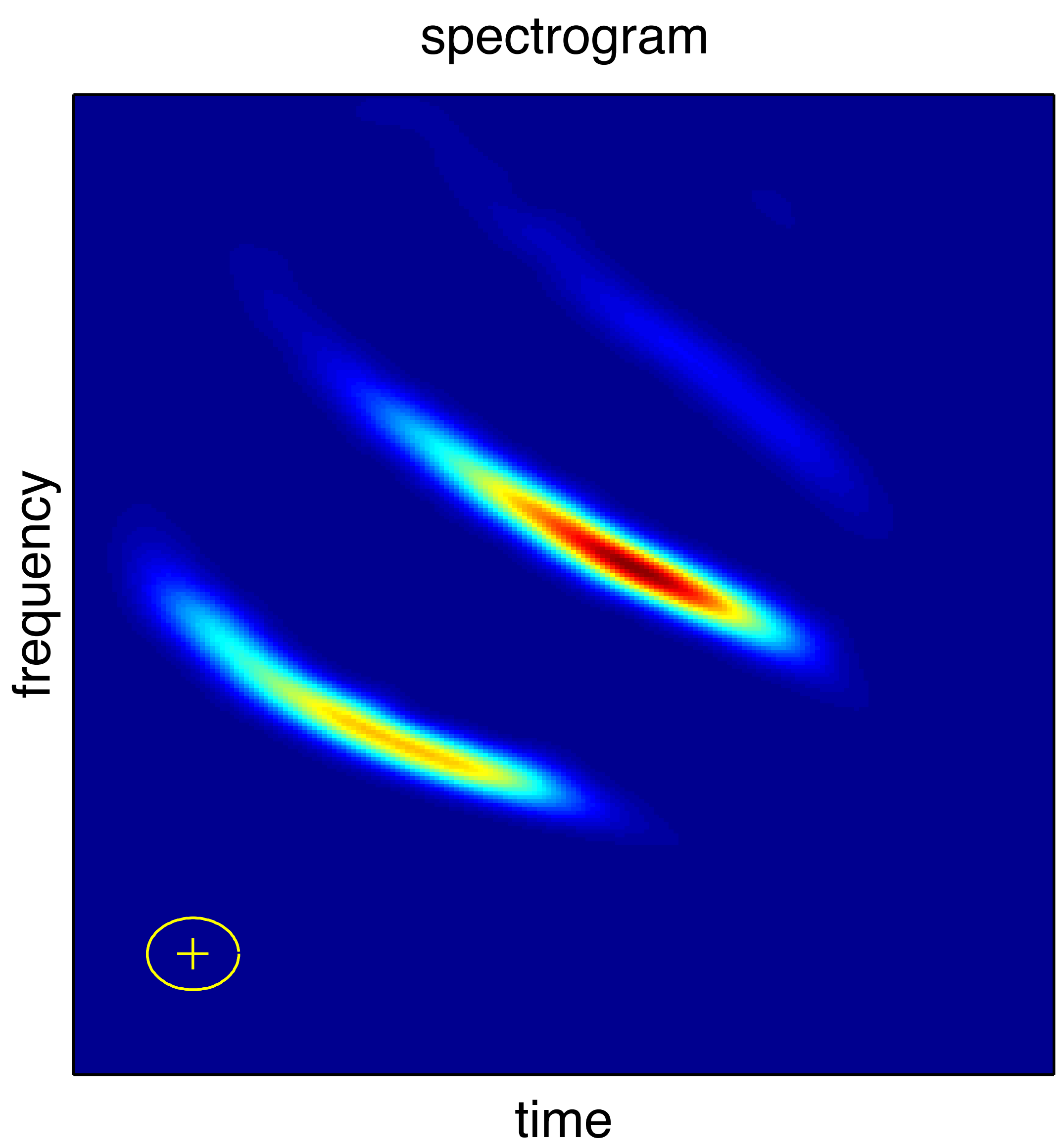
Can be done either from **phases** (Kodera et al., '76) or by **combining** STFTs with suitably chosen windows (Auger & F., 94)

Localization + cross-terms reduction :-))



Localization + cross-terms reduction :-))





Synchrosqueezing

- Reassigning in both time and frequency makes **reconstruction difficult**

- « **Synchrosqueezing** » (Daubechies *et al.*, ’94) as a variant, based on

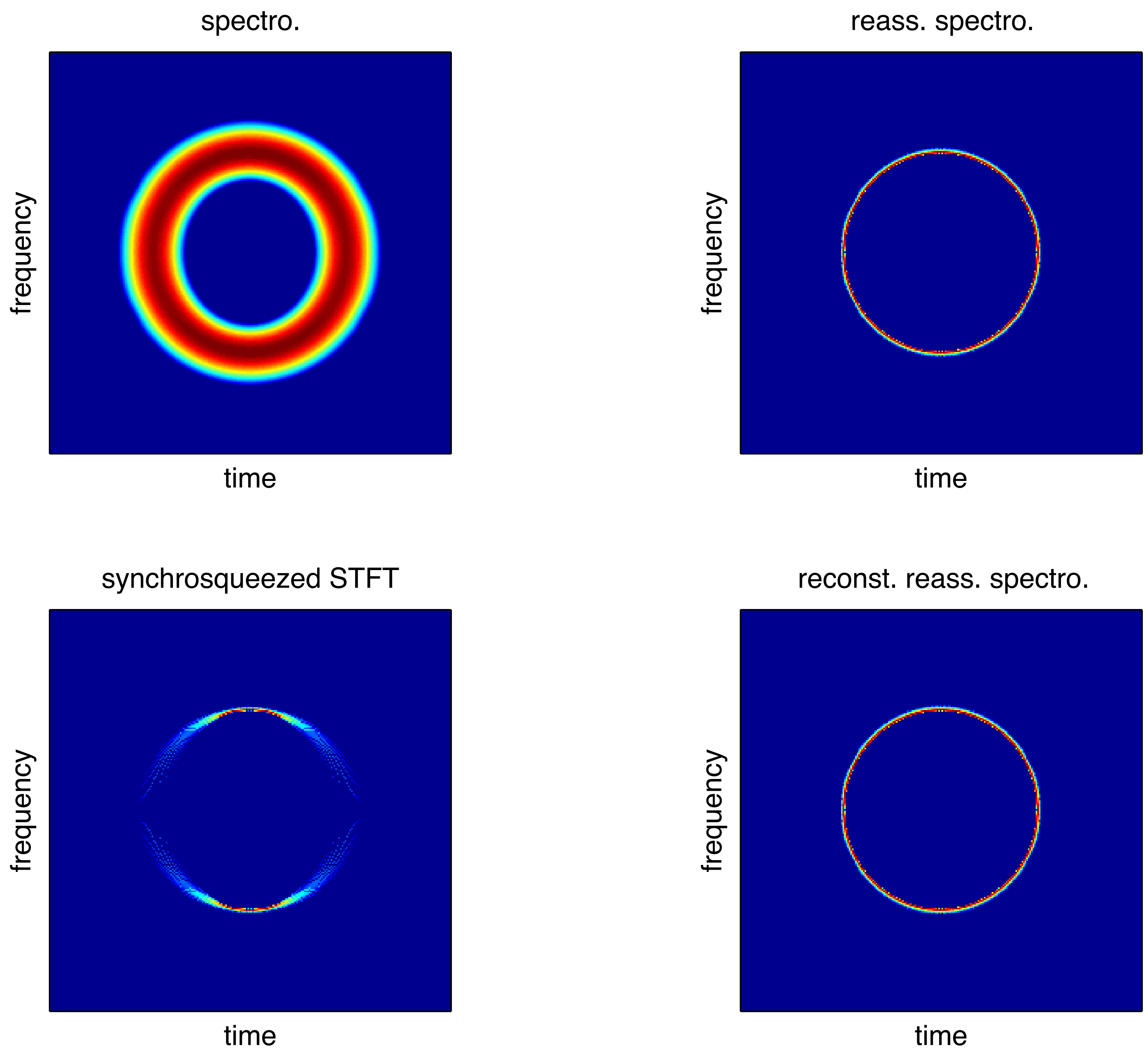
  - *frequency-only reassignment*

  - *applied to **wavelet transform** or **STFT**, e.g.,*

$$F_x^{(h)}(t,f) \mapsto \hat{F}_x^{(h)}(t,f) = \int F_x^{(h)}(t,\xi)\, \delta\left(f - \hat{f}_x(\xi)\right)\, d\xi$$

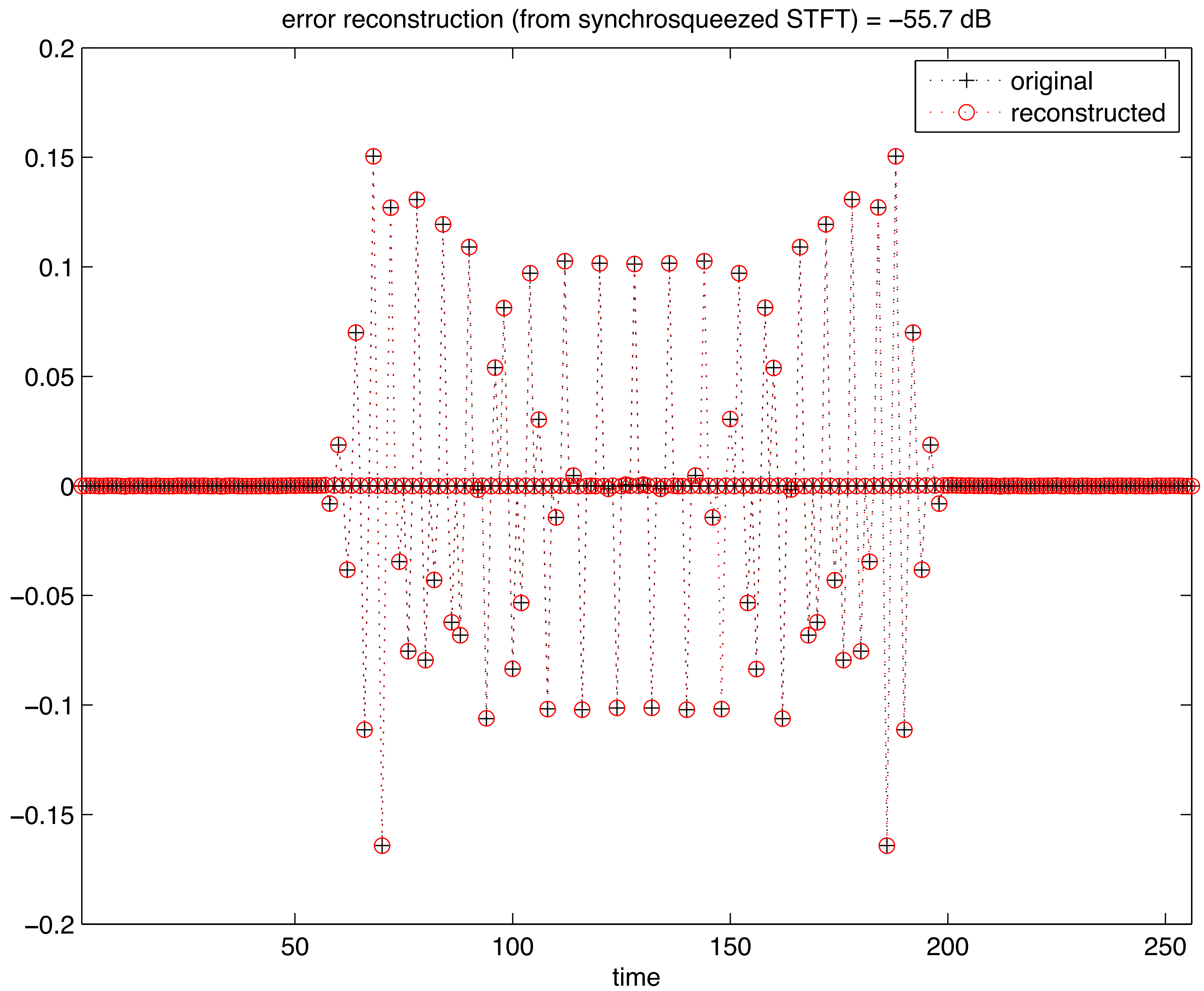
$$x(t) = \int_\Omega \hat{F}_x^{(h)}(t,\xi)\, d\xi$$

Example 1



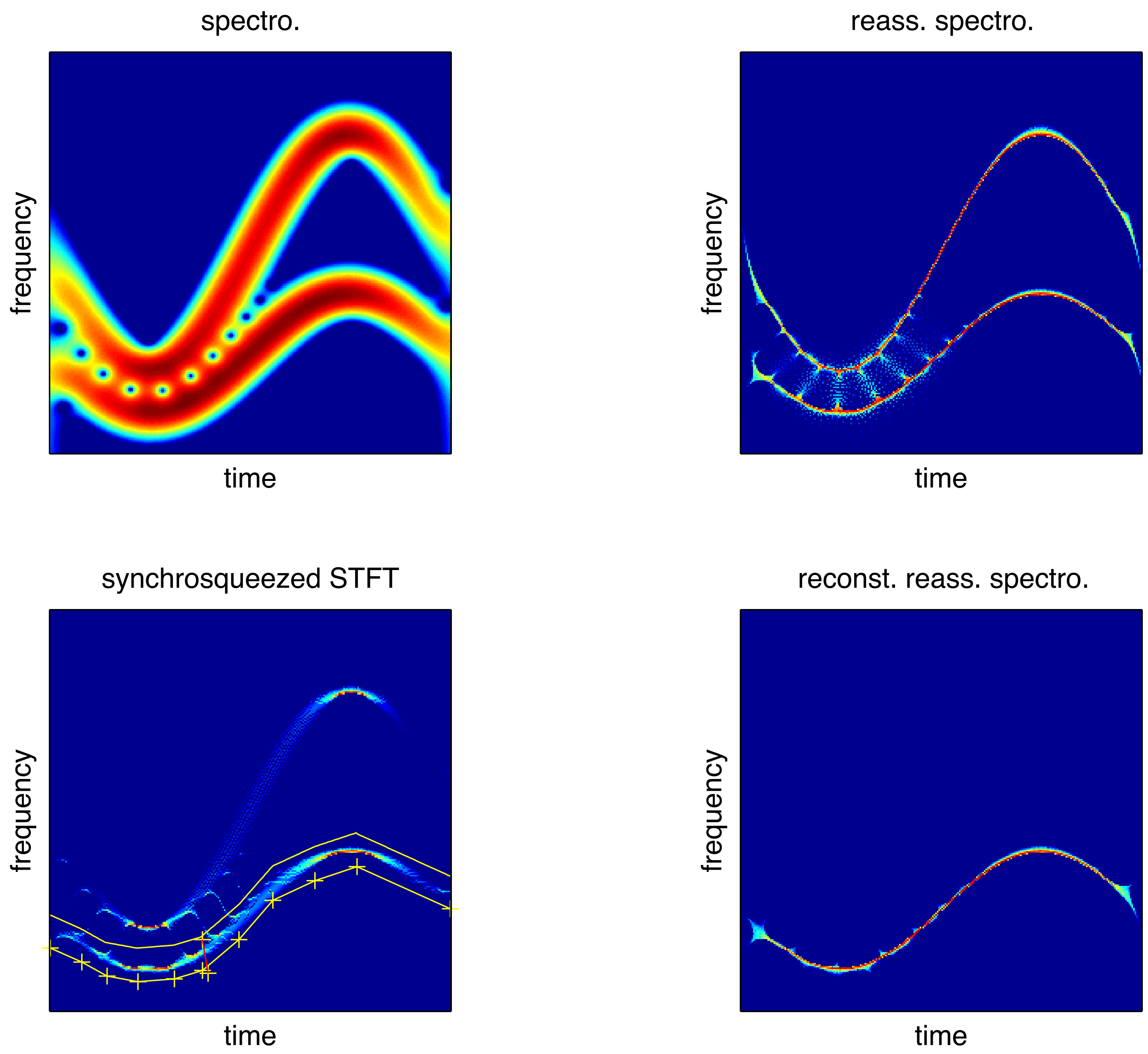
spectro.
frequency
time
reass. spectro.
frequency
time
synchrosqueezed STFT
frequency
time
reconst. reass. spectro.
frequency
time

Example 1



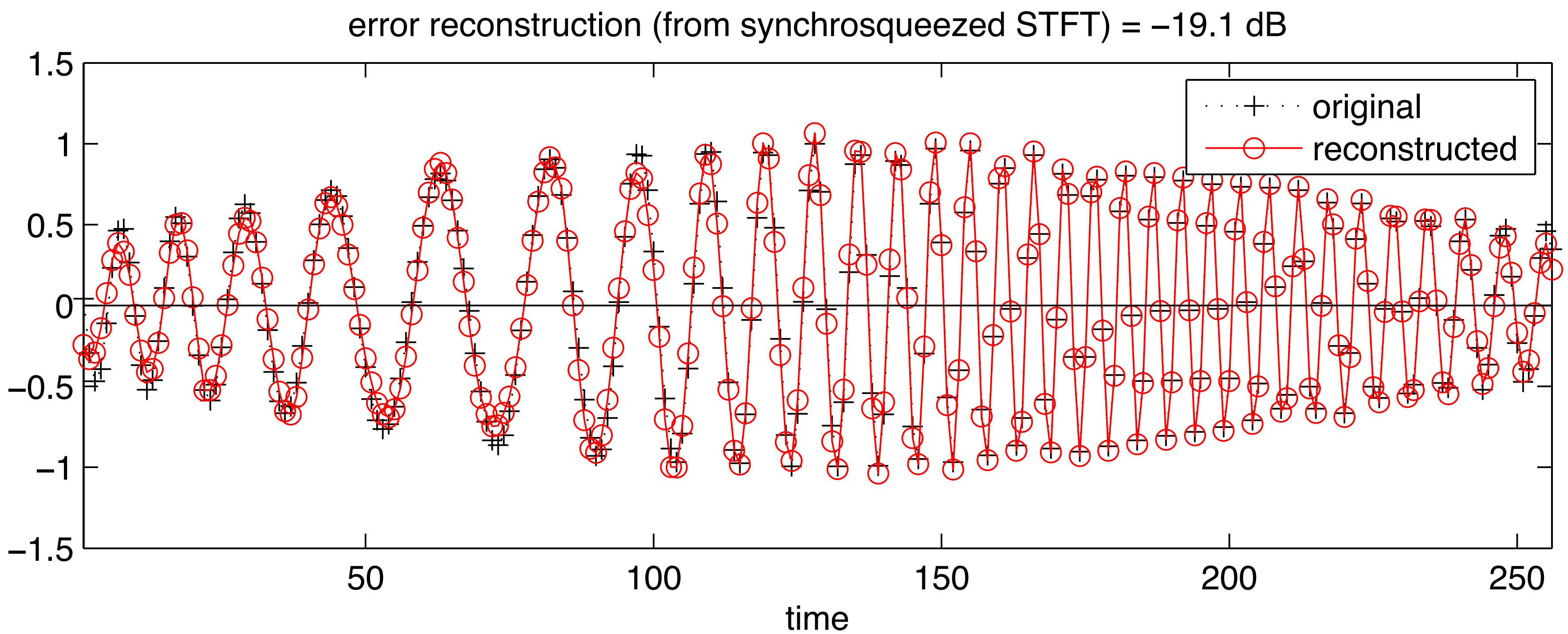
error reconstruction (from synchrosqueezed STFT) = −55.7 dB
original
reconstructed
0.2
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
−0.2
50
100
150
200
250
time

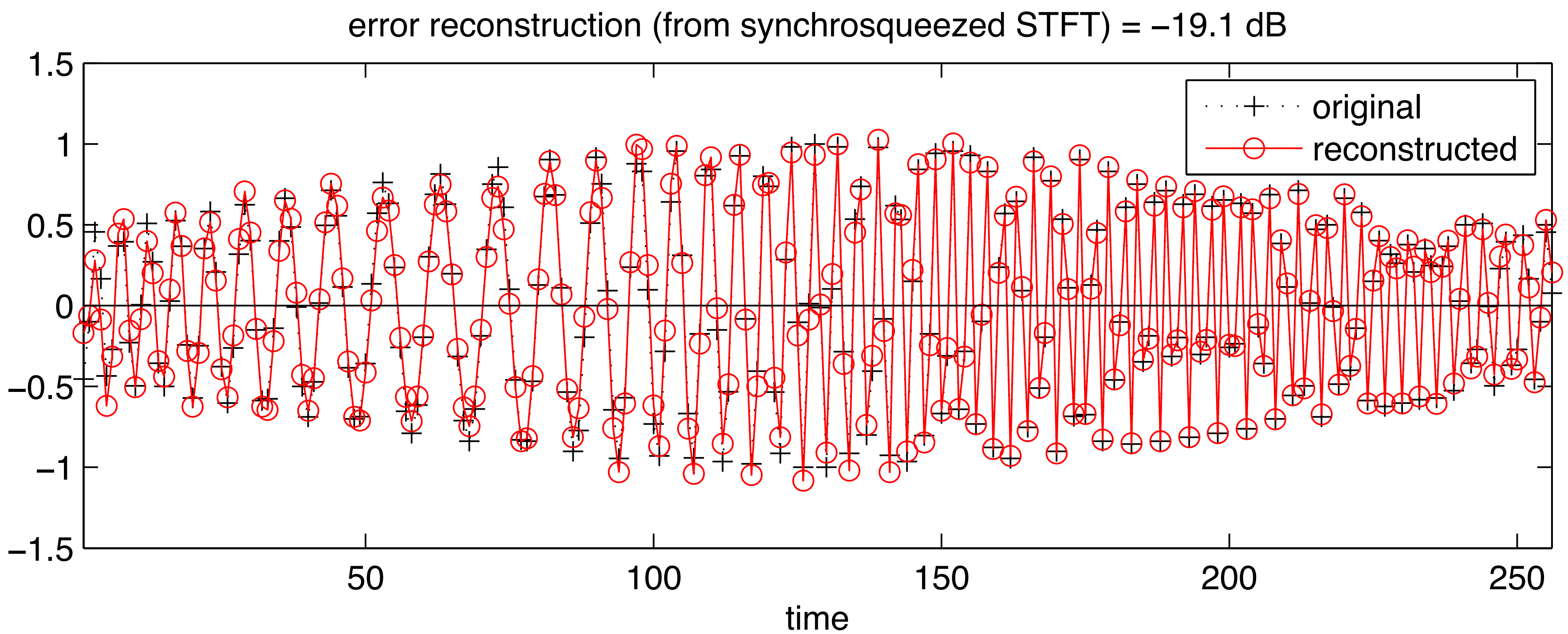
Example 2



spectro.
frequency
time
reass. spectro.
frequency
time
synchrosqueezed STFT
frequency
time
reconst. reass. spectro.
frequency
time

Example 2


error reconstruction (from synchrosqueezed STFT) = −19.1 dB
original
reconstructed
1.5
1
0.5
0
−0.5
−1
−1.5
50
100
150
200
250
time


error reconstruction (from synchrosqueezed STFT) = −19.1 dB
original
reconstructed
1.5
1
0.5
0
−0.5
−1
−1.5
50
100
150
200
250
time

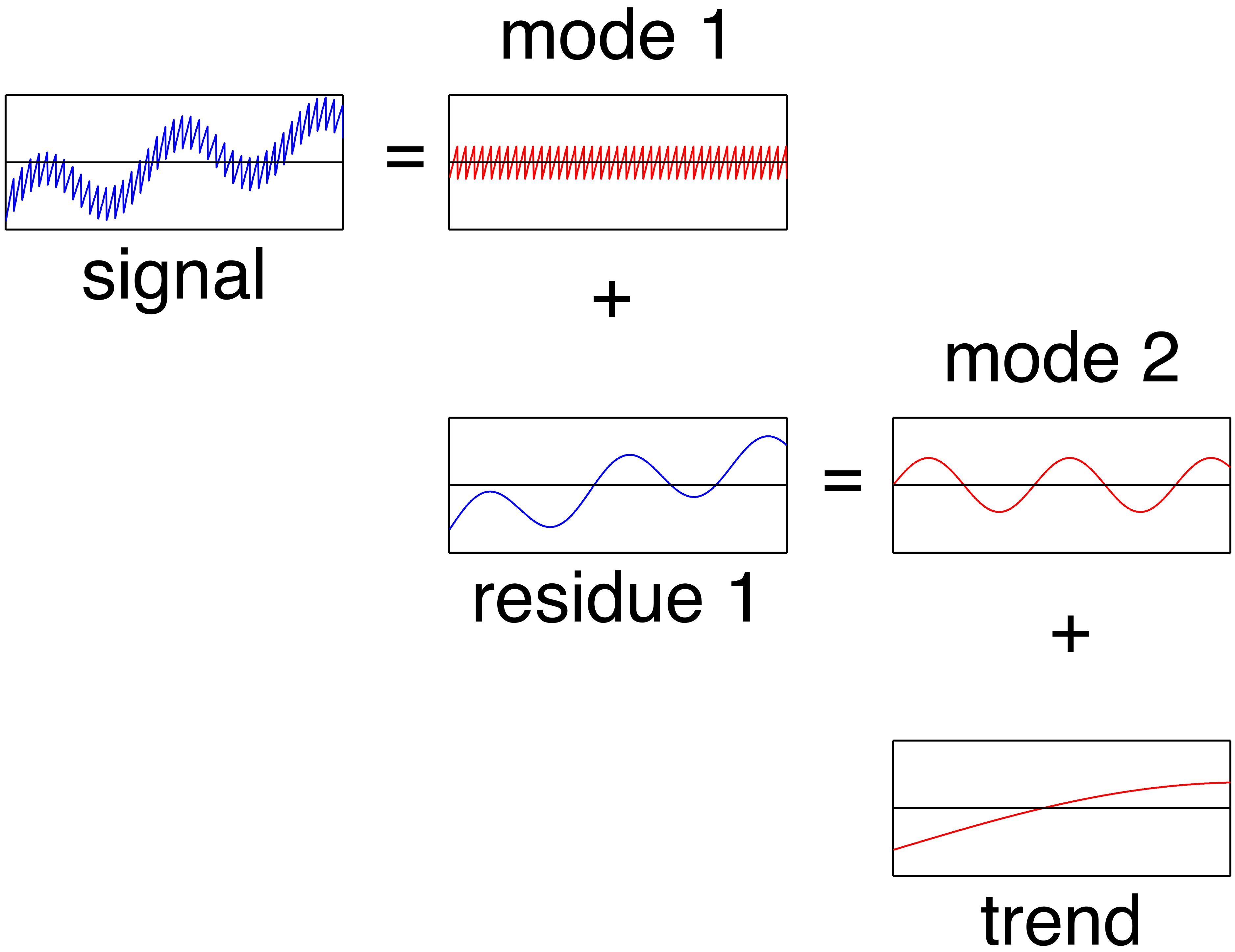
Empirical Mode Decomposition

- Goal:
  - ***disentangle*** *multicomponent « AM-FM-like » signals*
  - *accommodate for* ***nonstationary and/or nonlinear*** *oscillations*
- Approach (Huang *et al.*, '98):
  - ***local***
  - ***model-free***
  - ***no explicit transform***

EMD rationale

signal = fast oscillation + slow oscillation + iteration

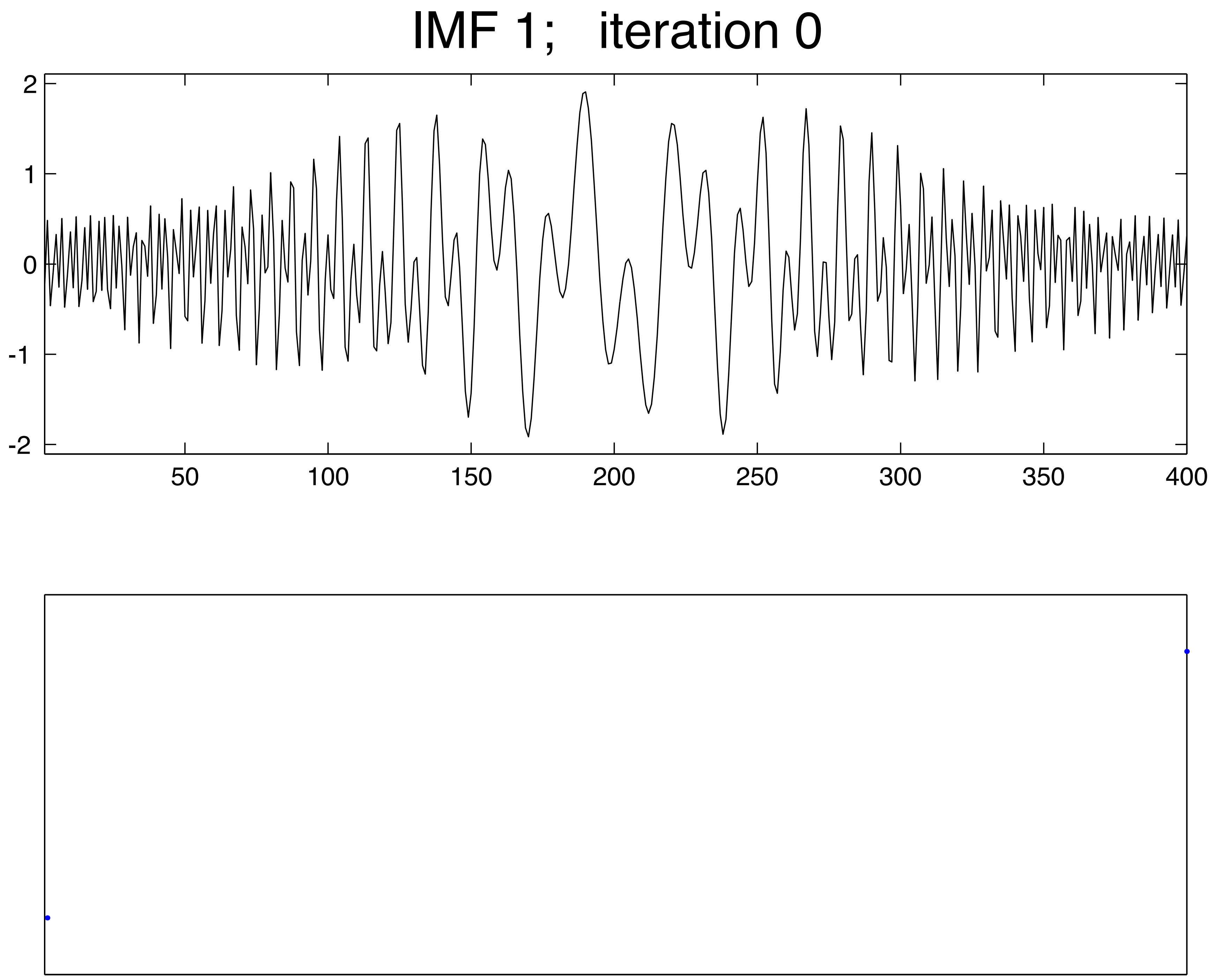
- **Hierarchical** decomposition
- In the spirit of **wavelets**, but with « fast » vs. « slow » separation
    - *not determined by fixed « high-pass » vs. « low-pass » filters*
    - *fully **data-driven***
    - *controlled by **local extrema***


mode 1
=
signal
+
mode 2
=
residue 1
+
trend

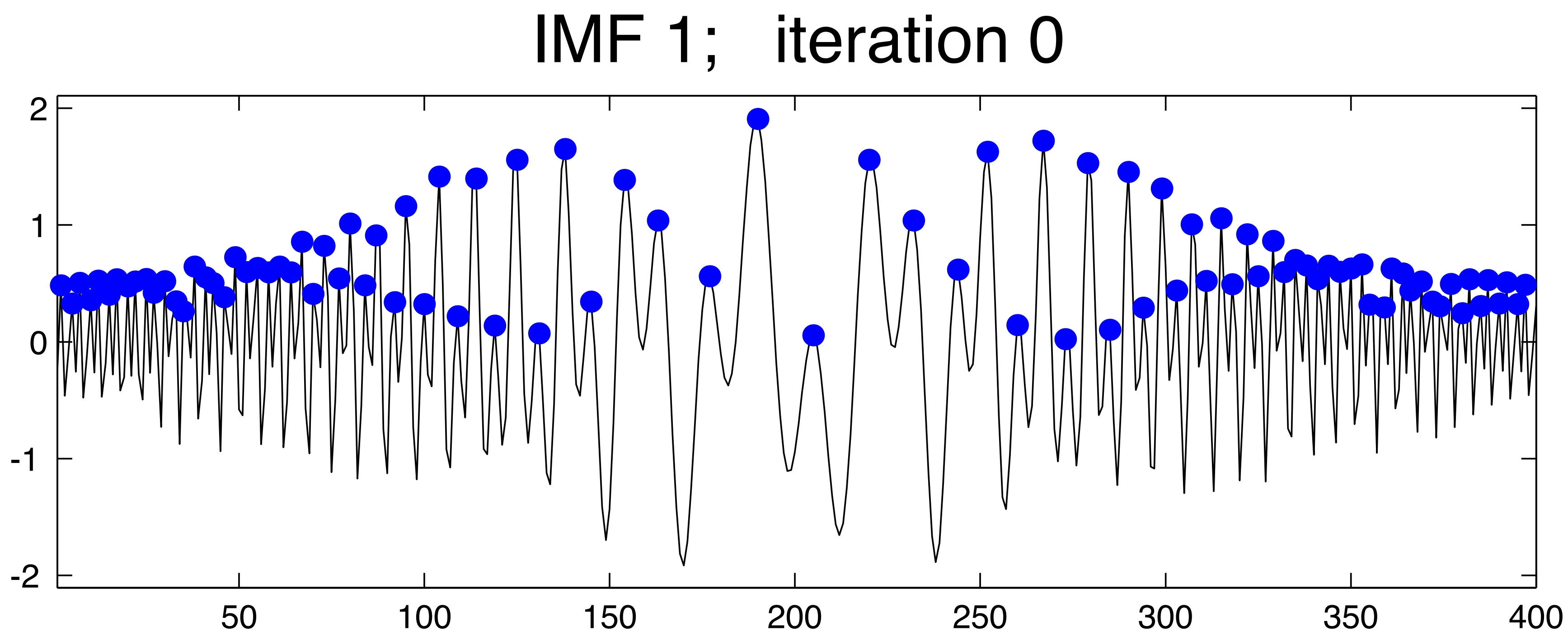
EMD as a data-driven algorithm

1. identify **local extrema** (maxima and minima) in the signal
2. form upper and lower **envelopes** by interpolation (cubic splines)
   - *subtract the mean envelope from the signal*
   - *iterate until **mean = 0***
3. subtract the so-obtained « **Intrinsic Mode Function** » (IMF) from the signal
4. **iterate** on the residual

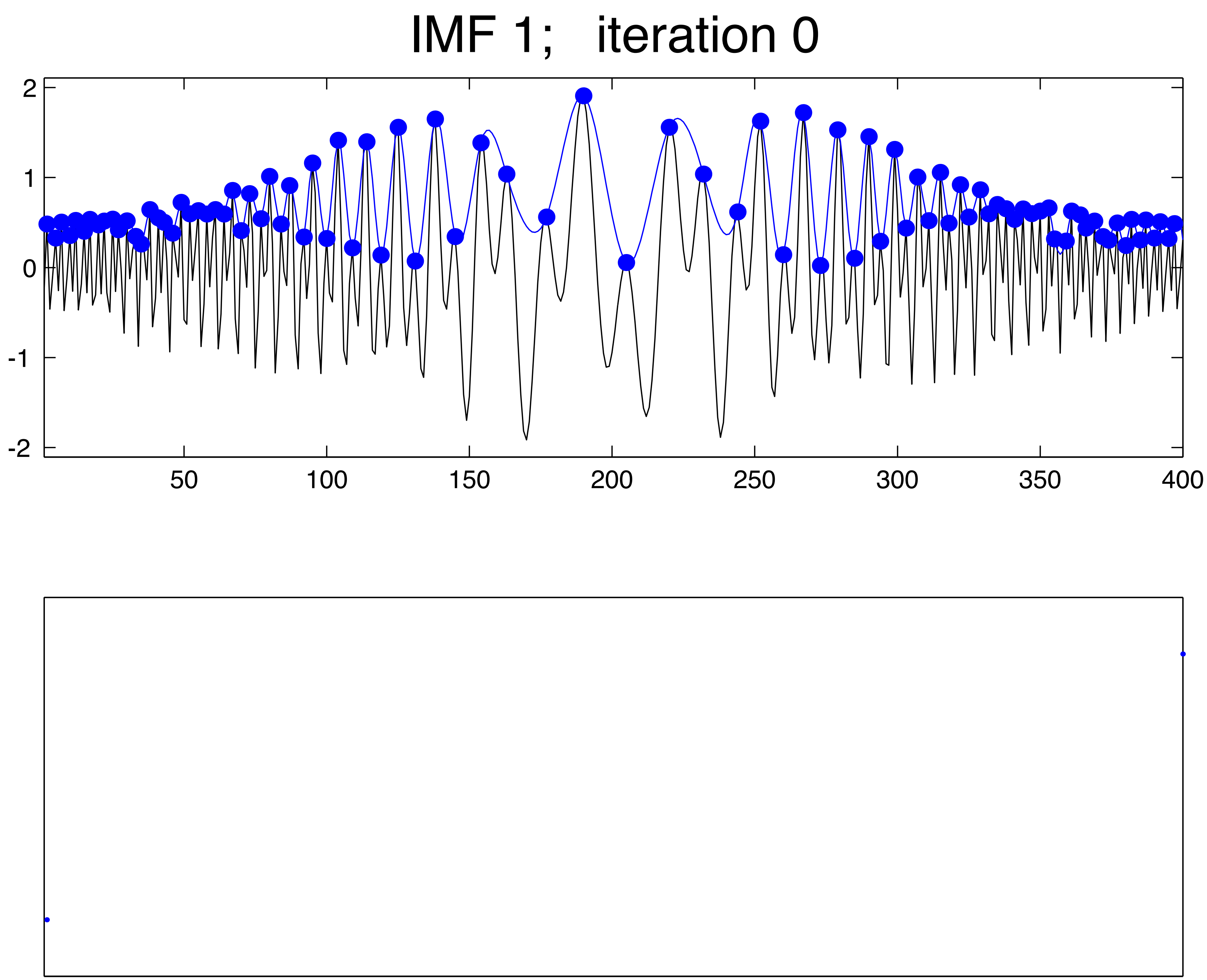
initialization: input = signal



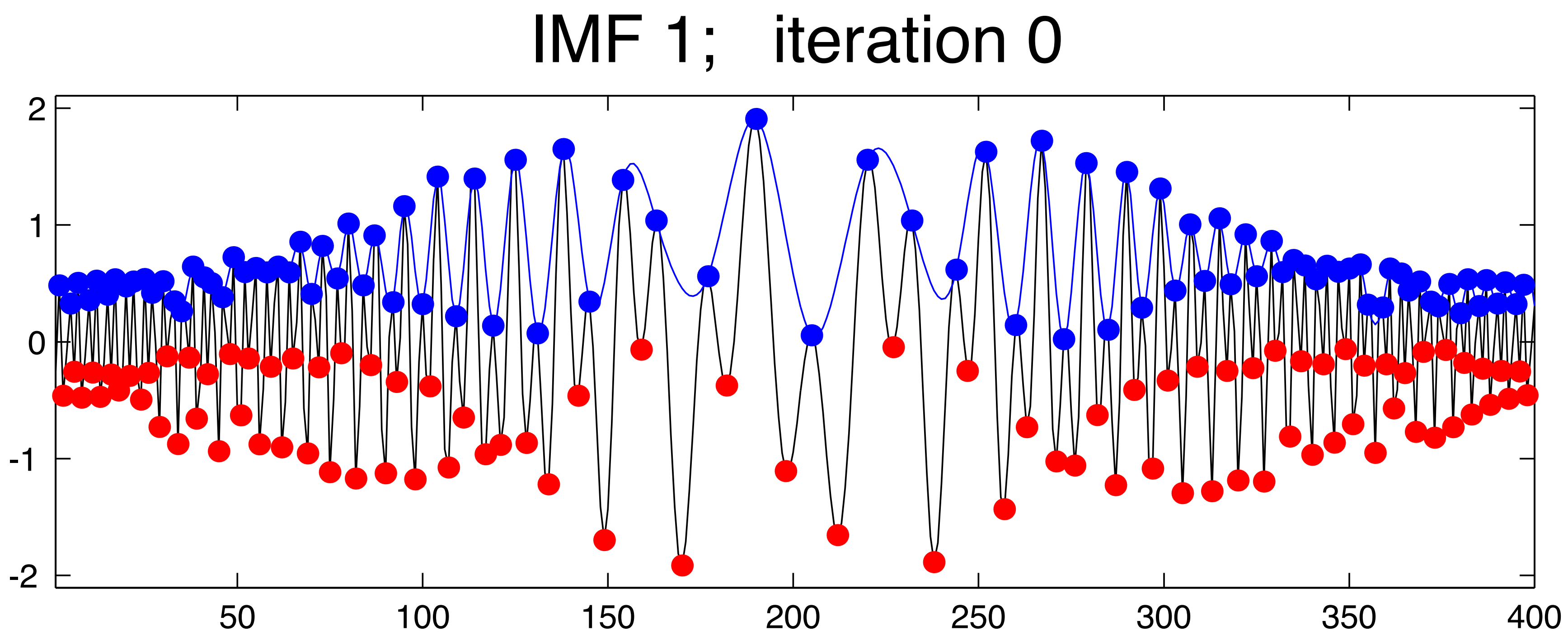
local maxima

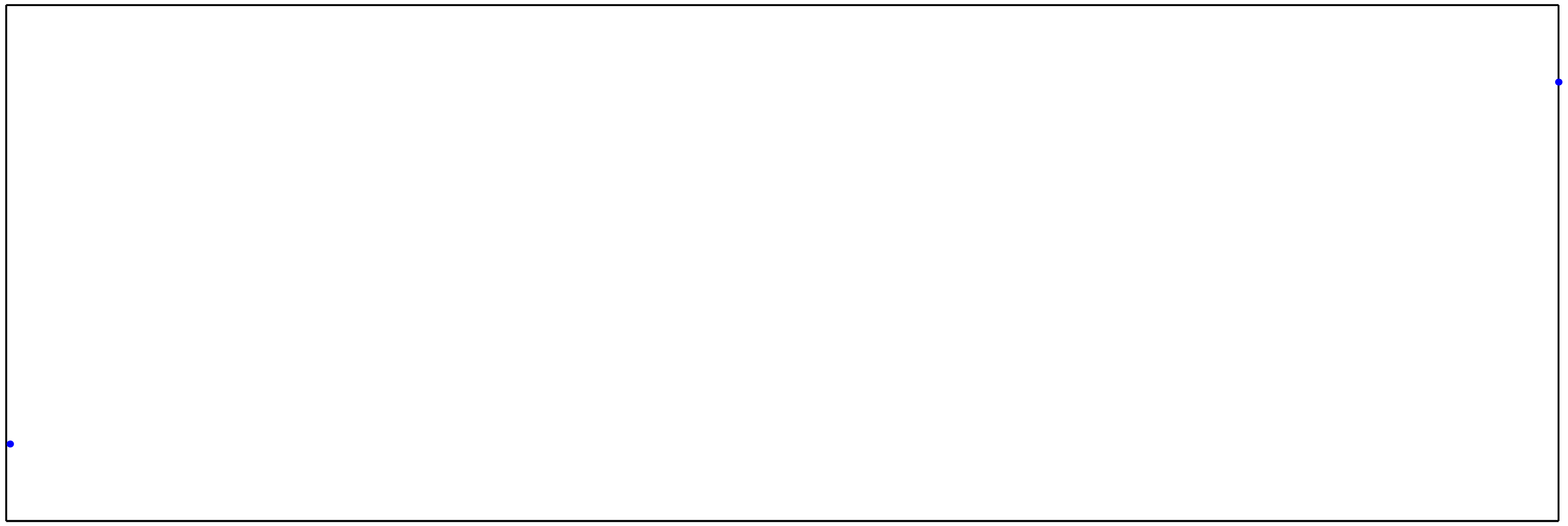


upper envelope

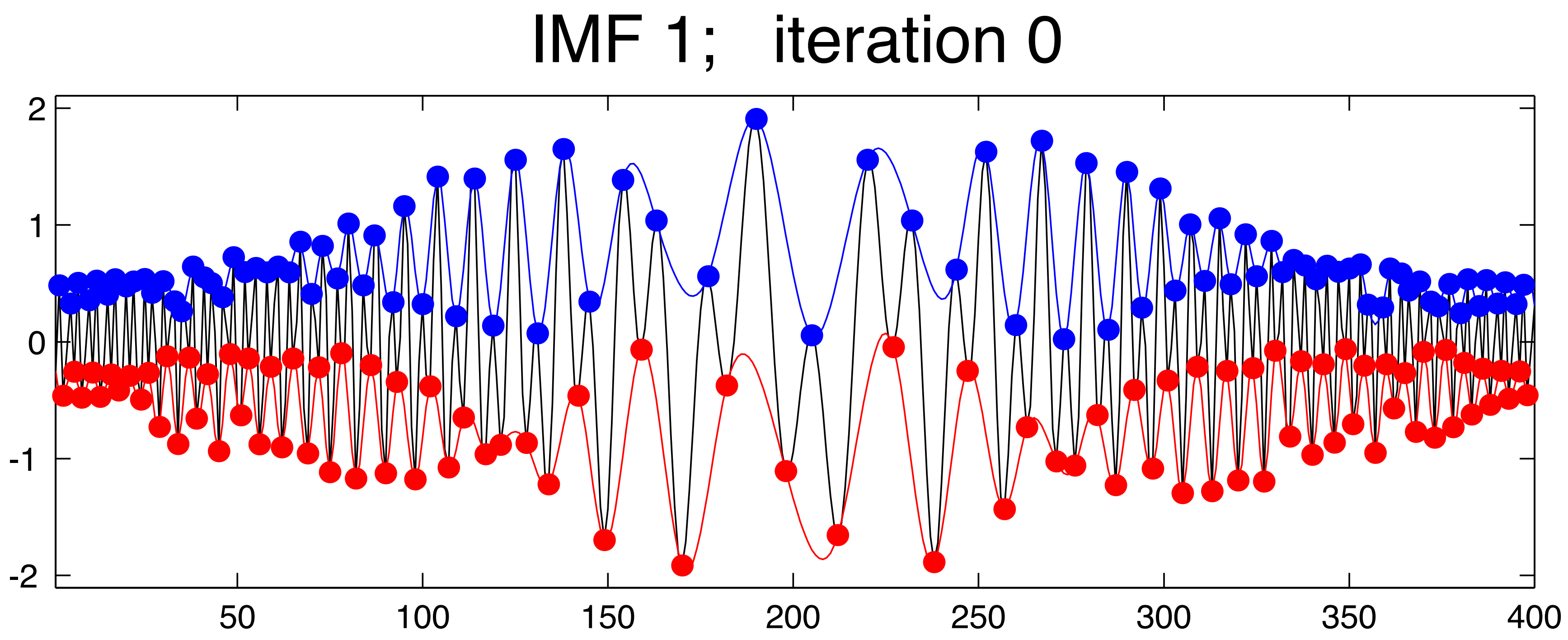


local minima

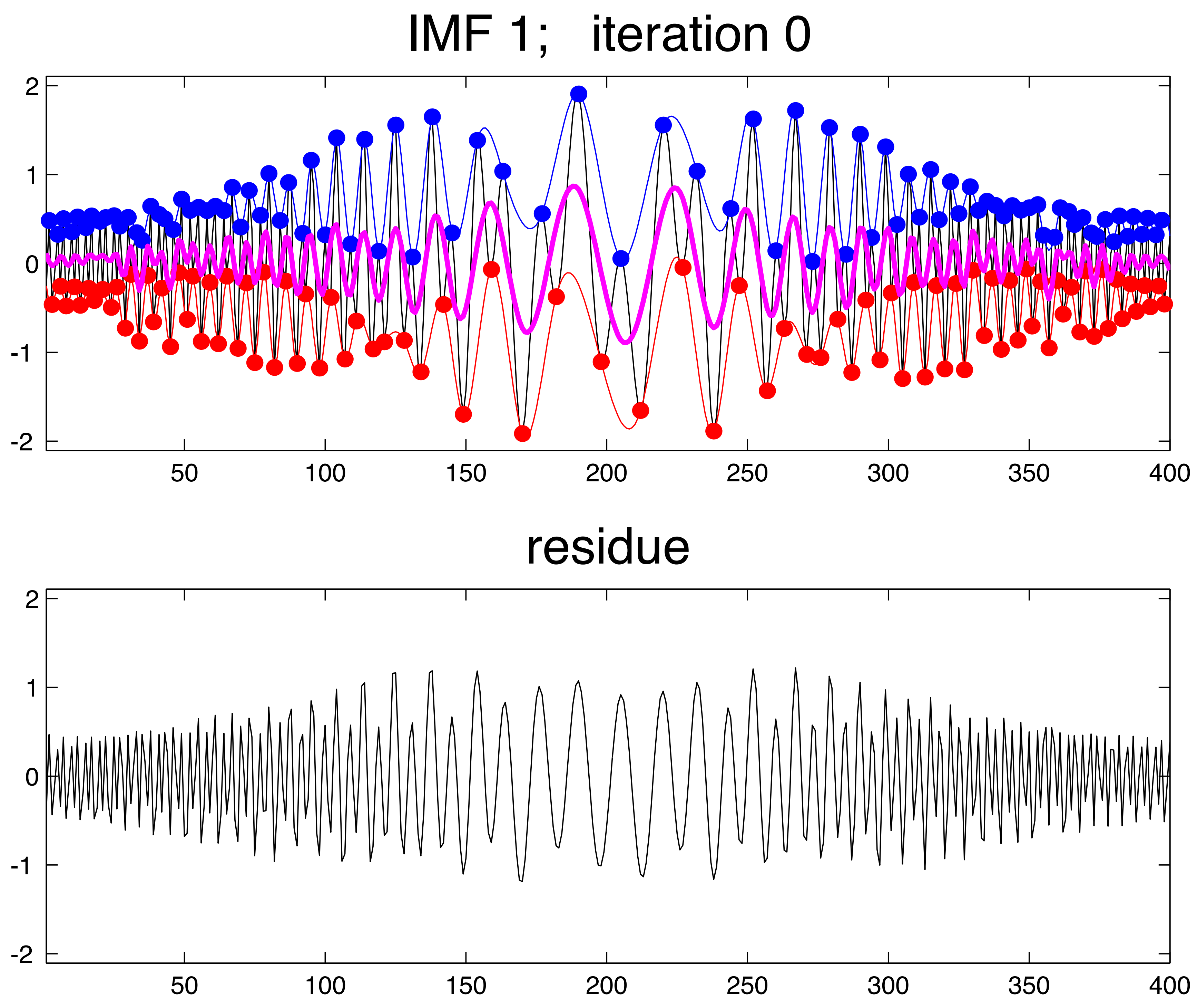




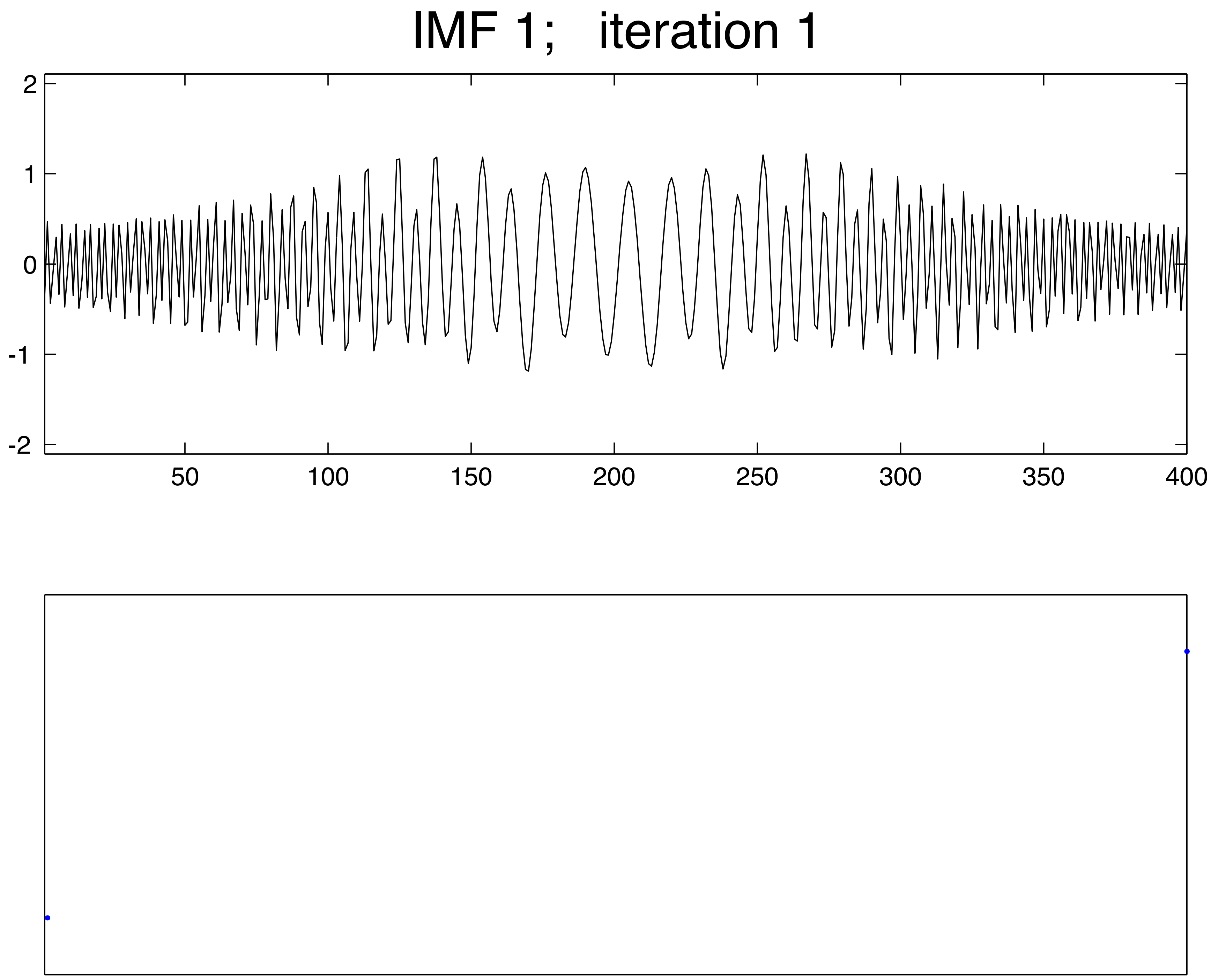
lower envelope



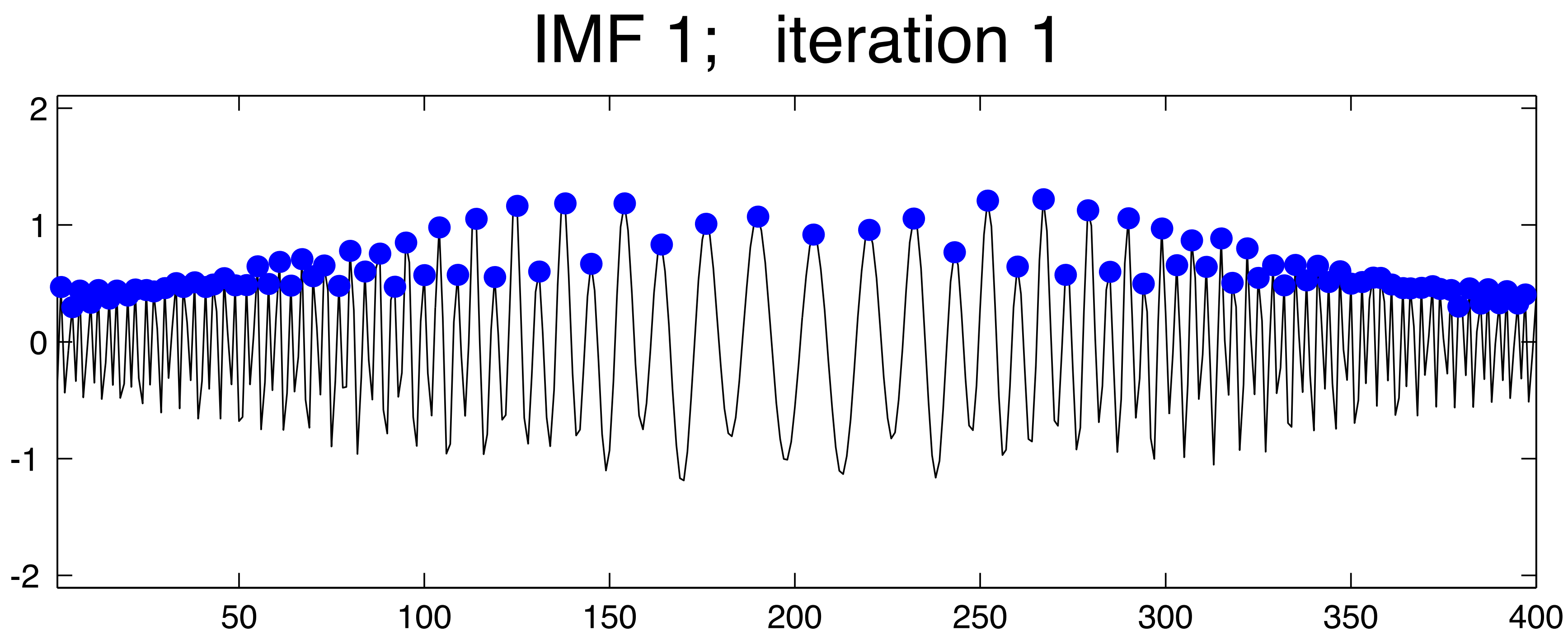
proto-mode 1 = signal – mean envelope



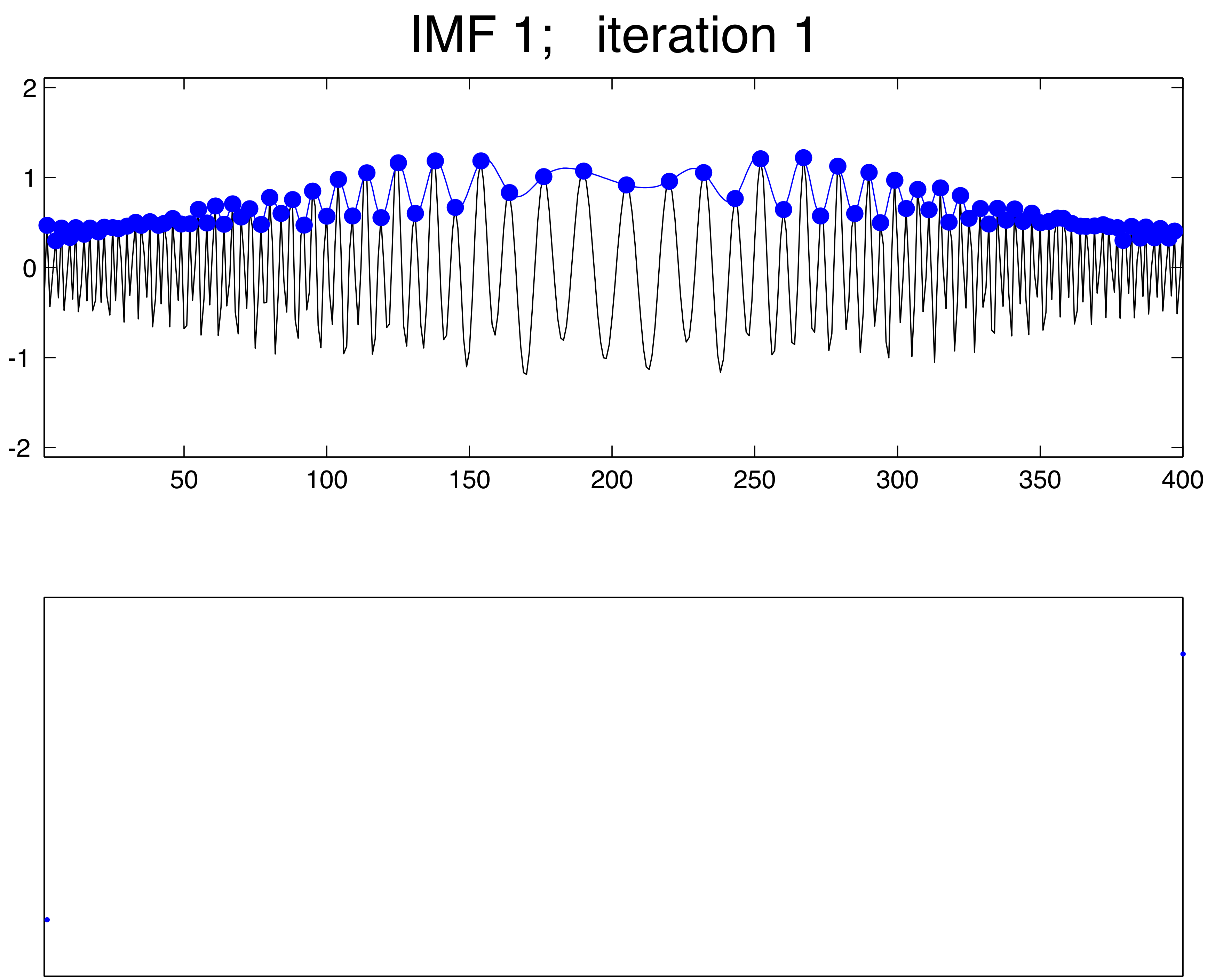
sifting: input = proto-mode 1



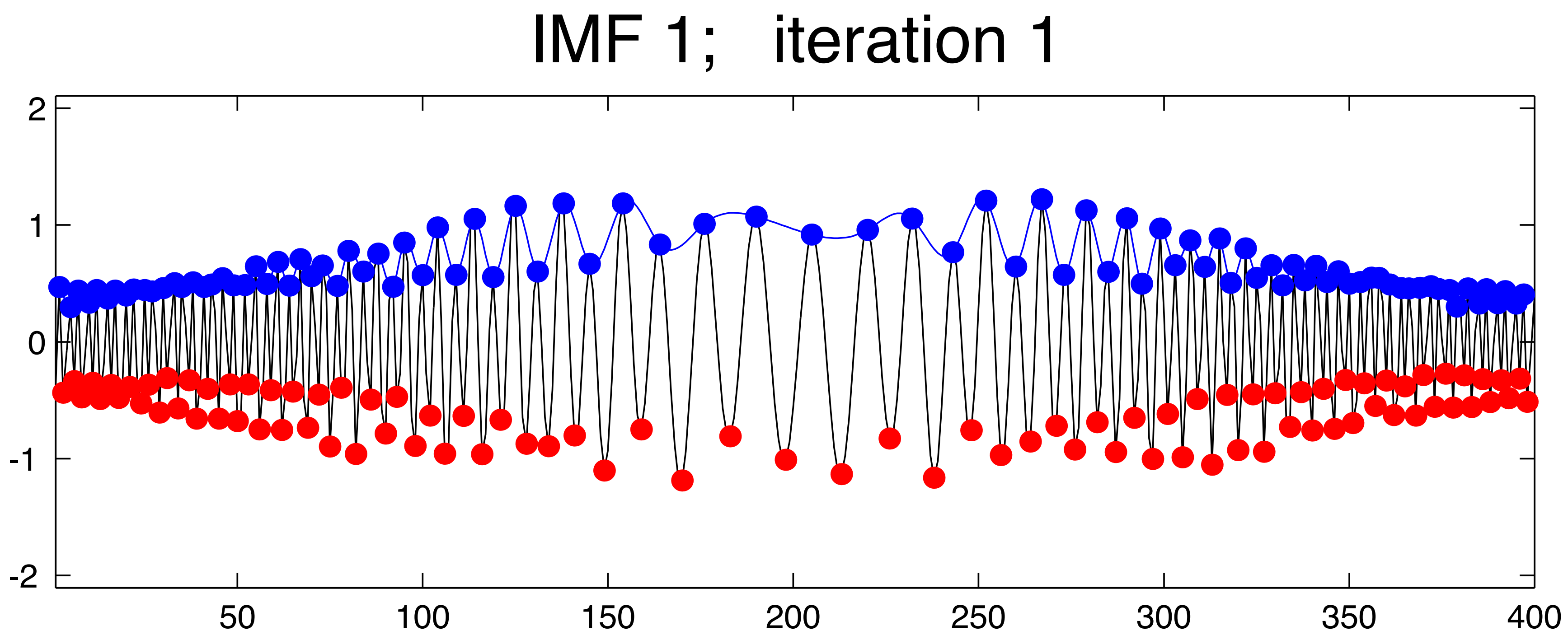
local maxima

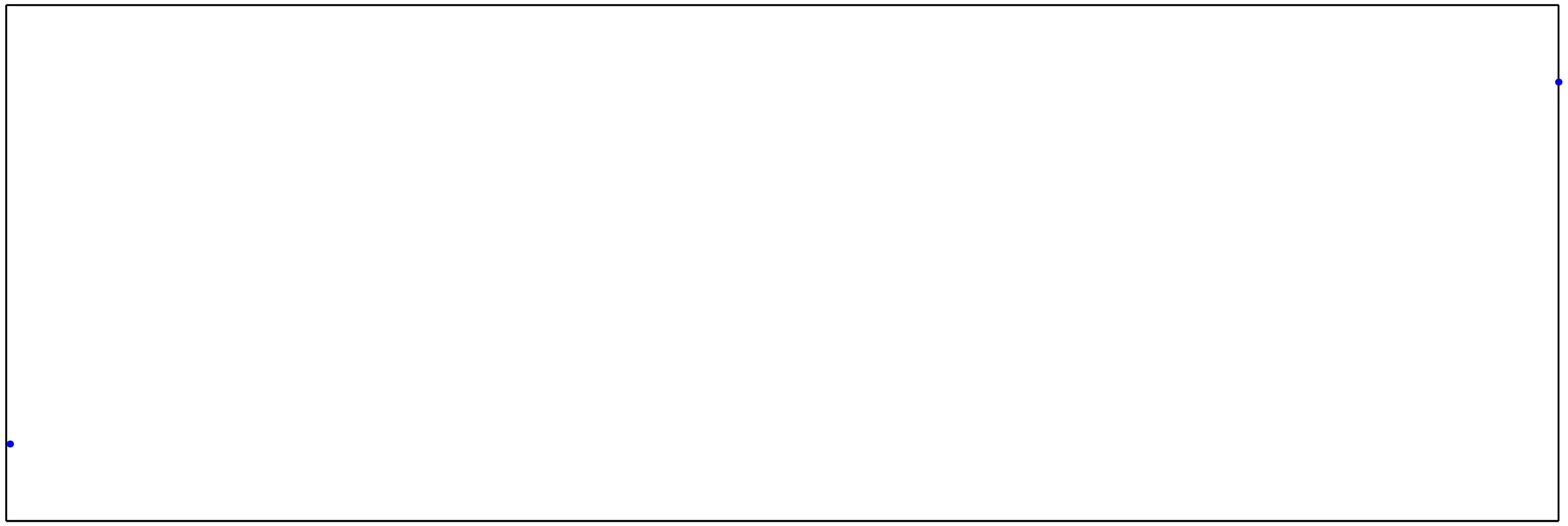


upper envelope

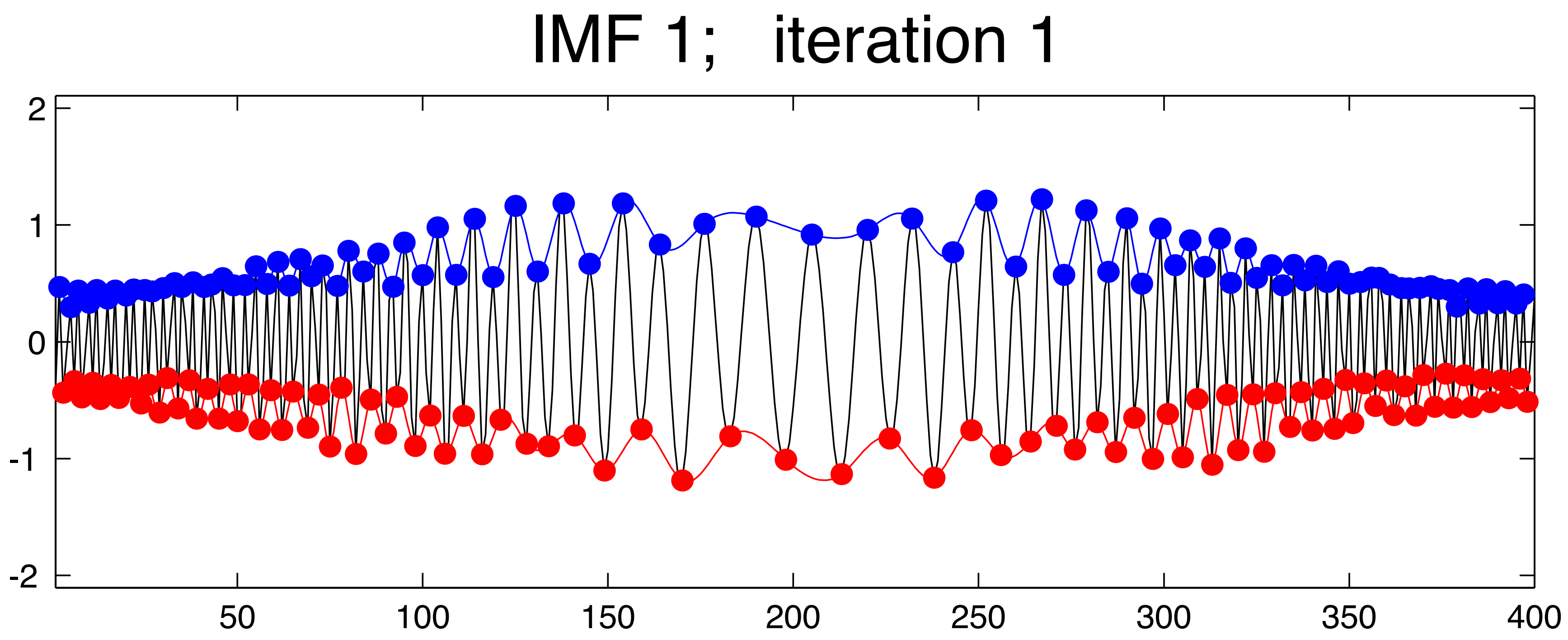


local minima

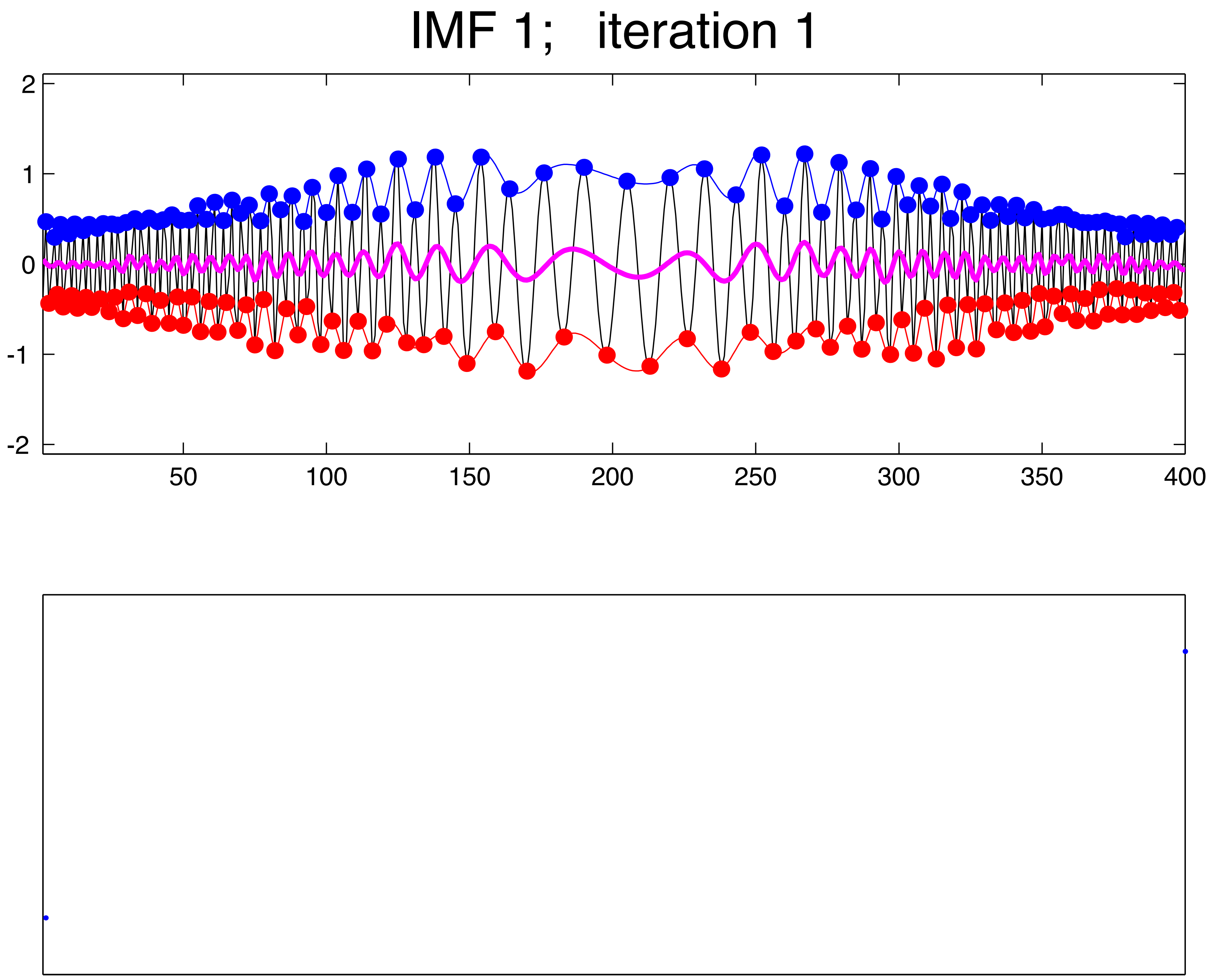




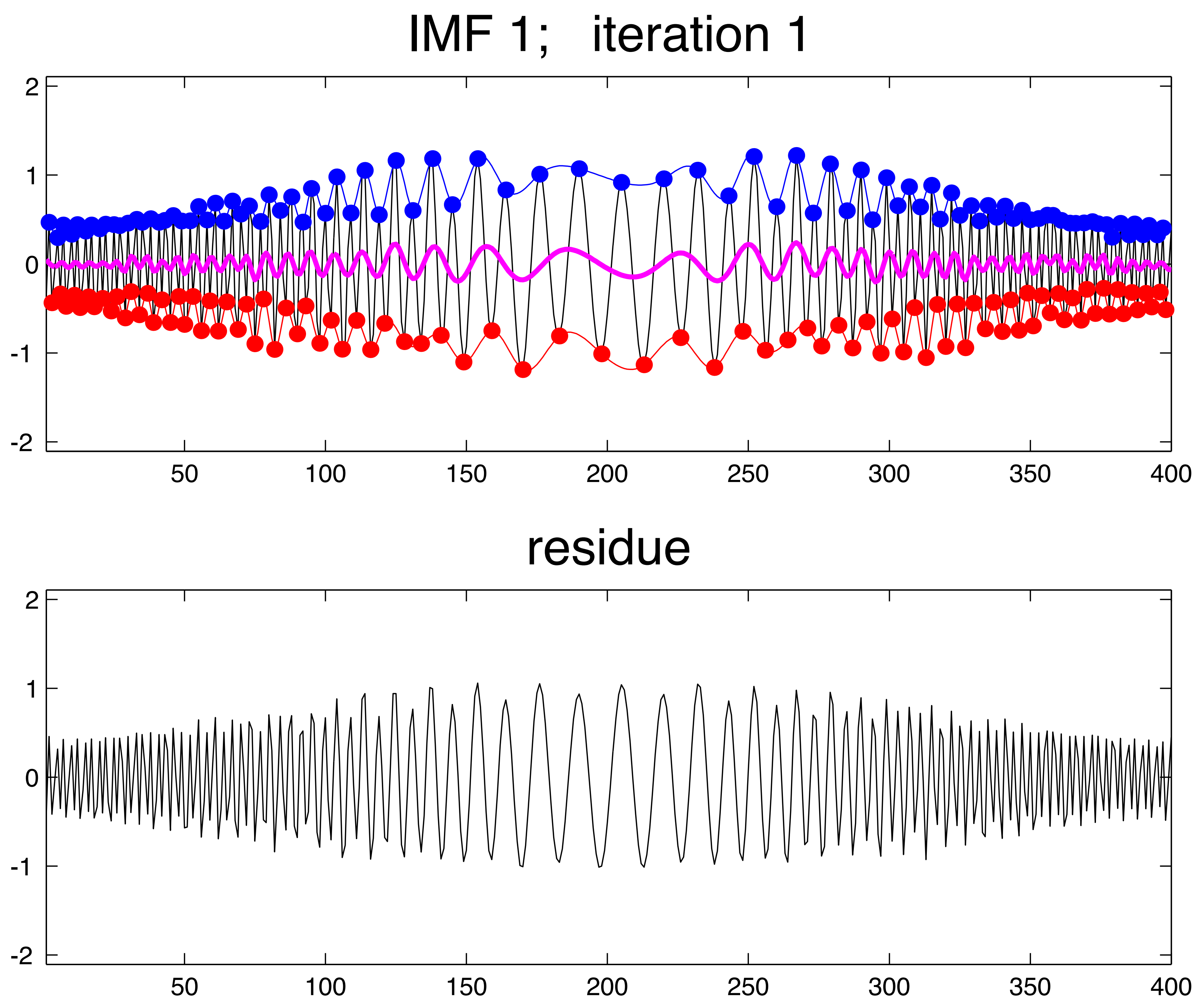
lower envelope



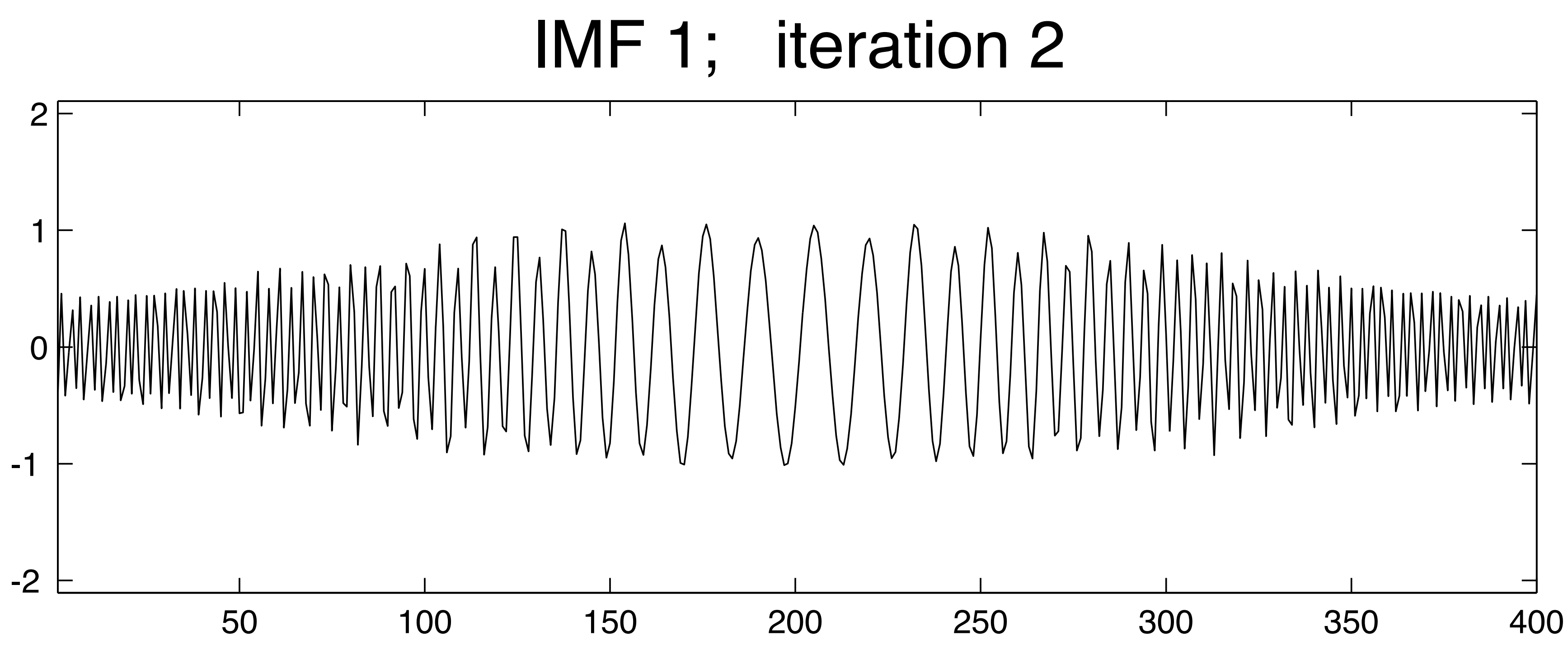
mean envelope



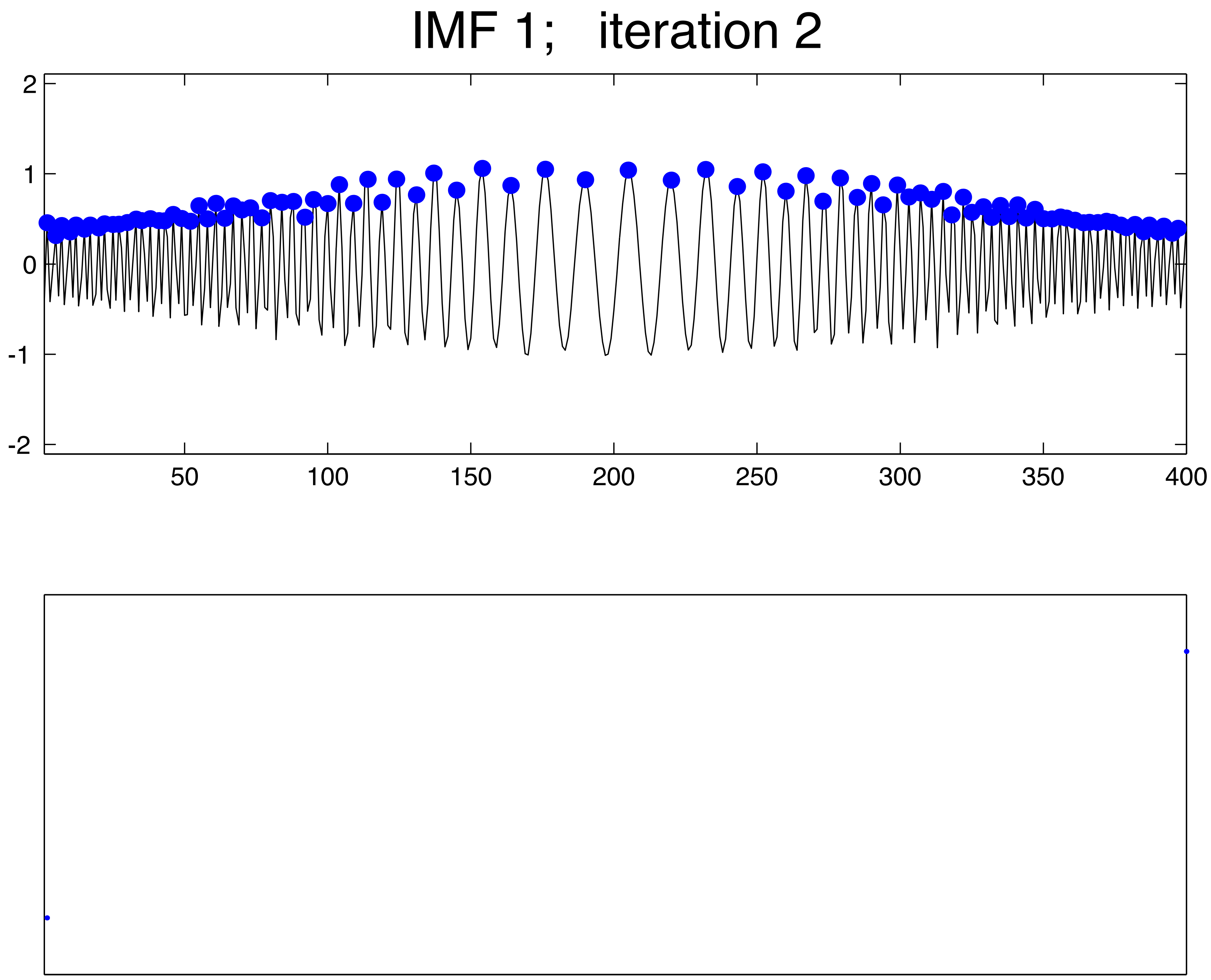
proto-mode 2 = proto-mode 1 – mean envelope



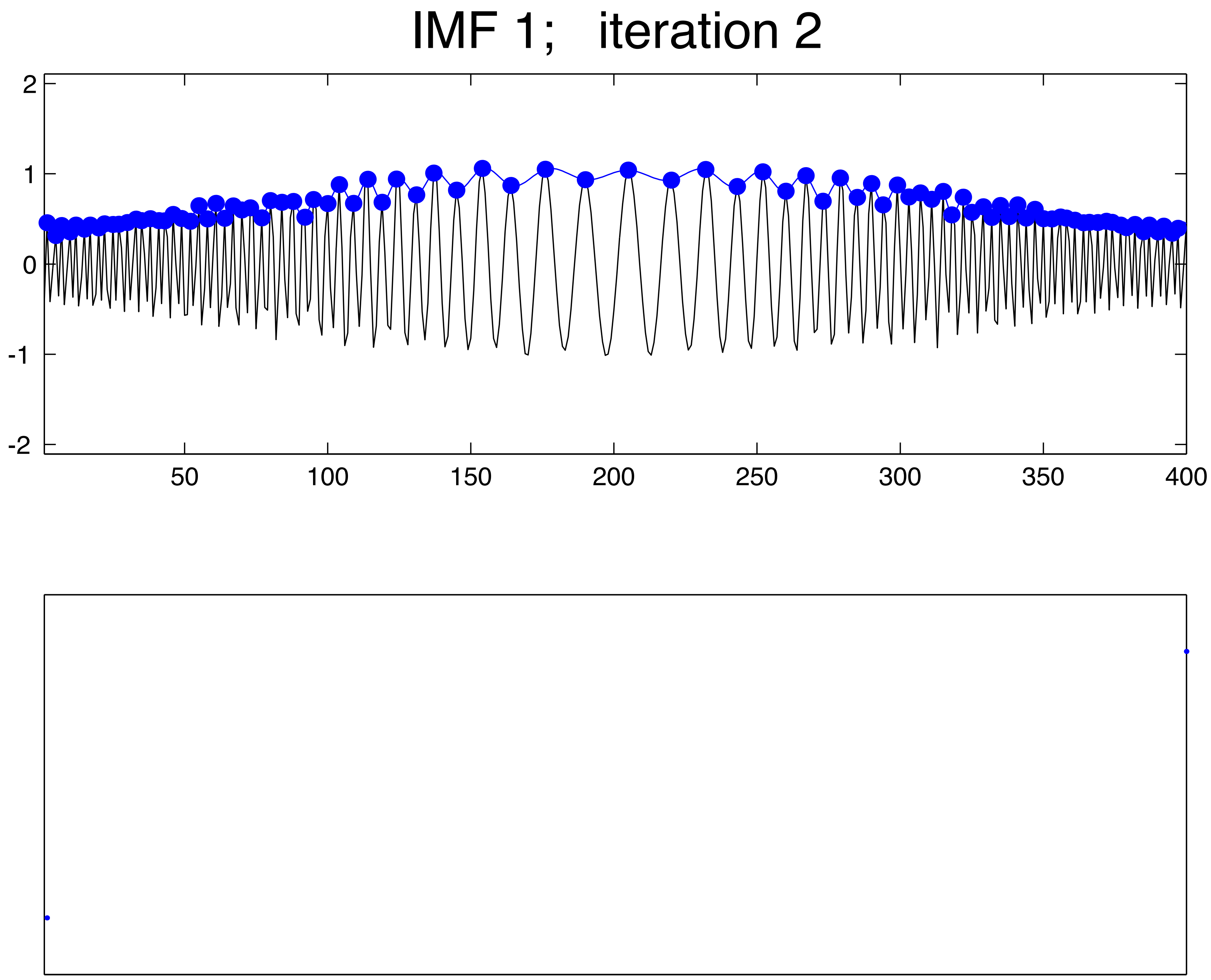
input = proto-mode 2



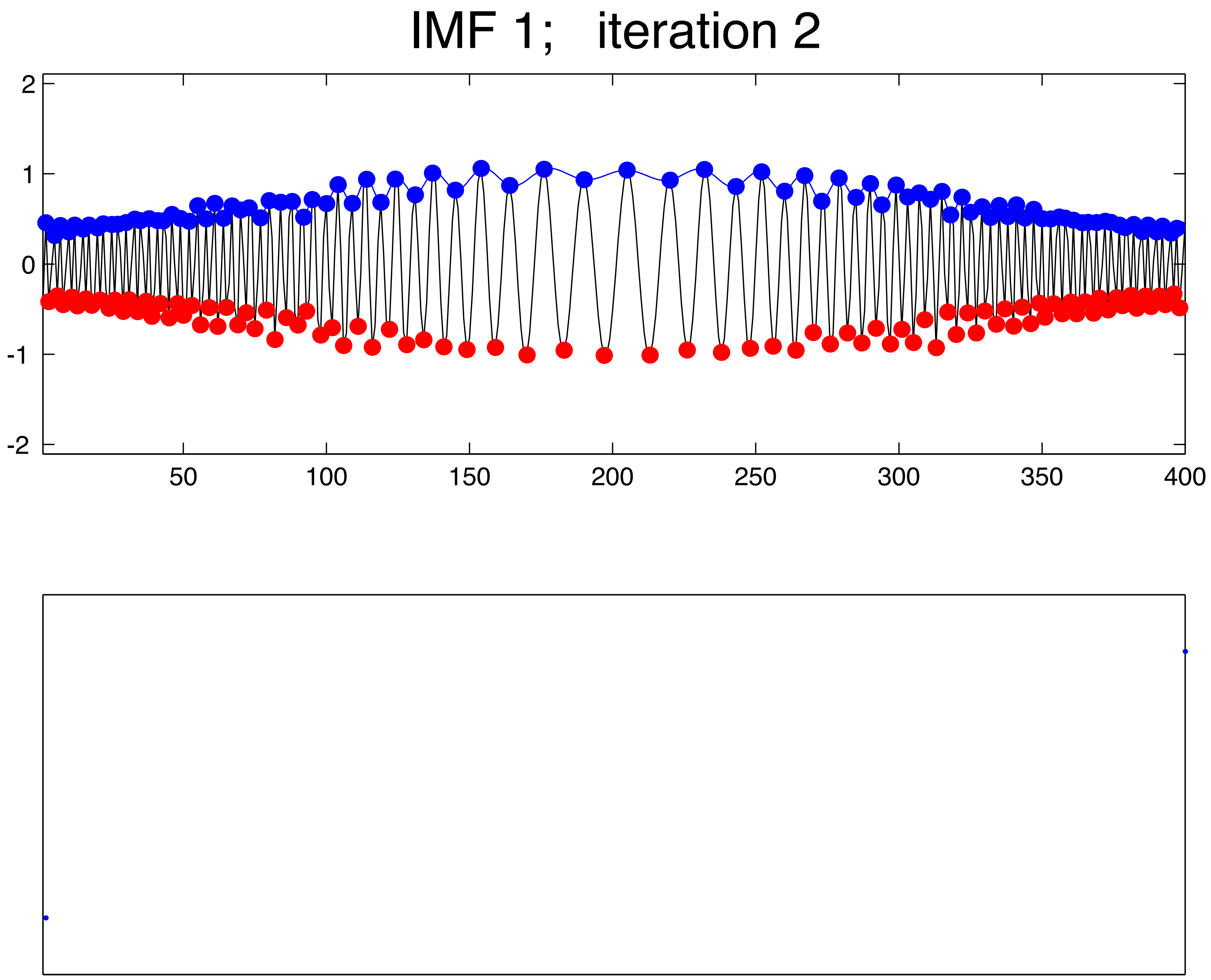
local maxima



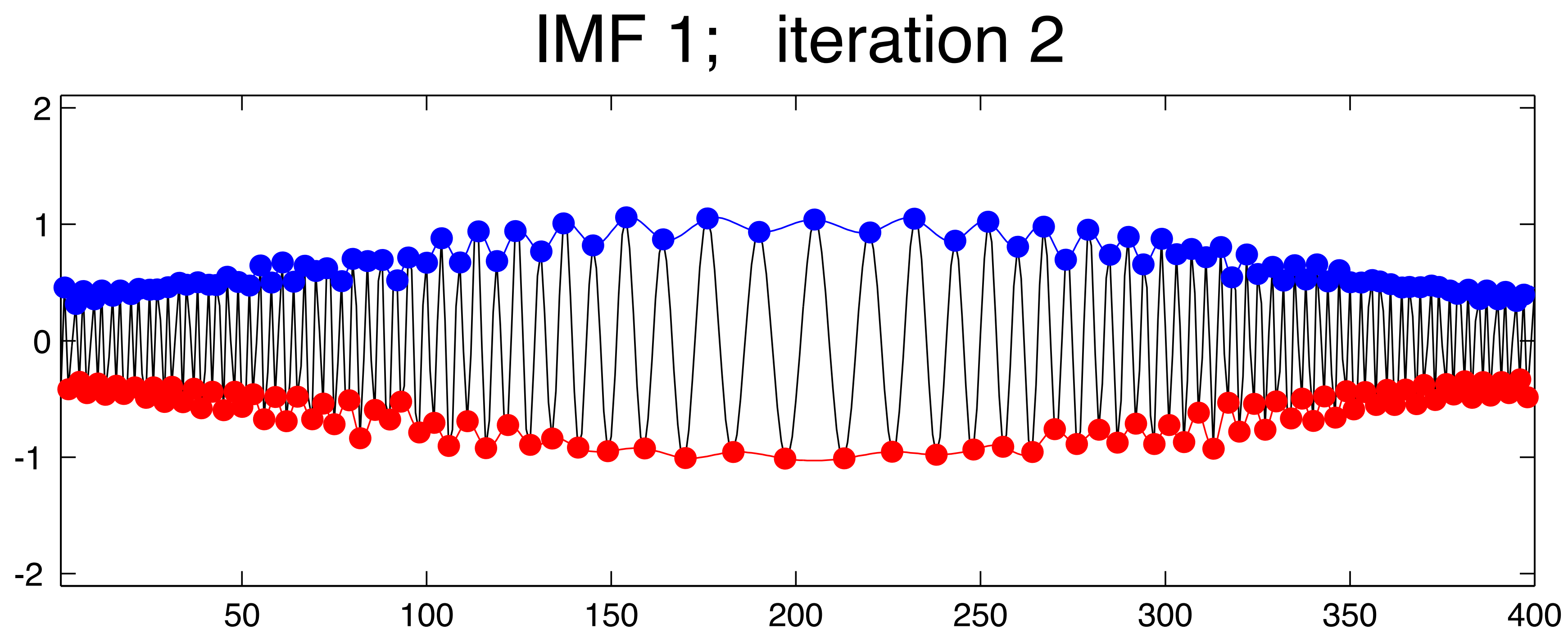
upper envelope

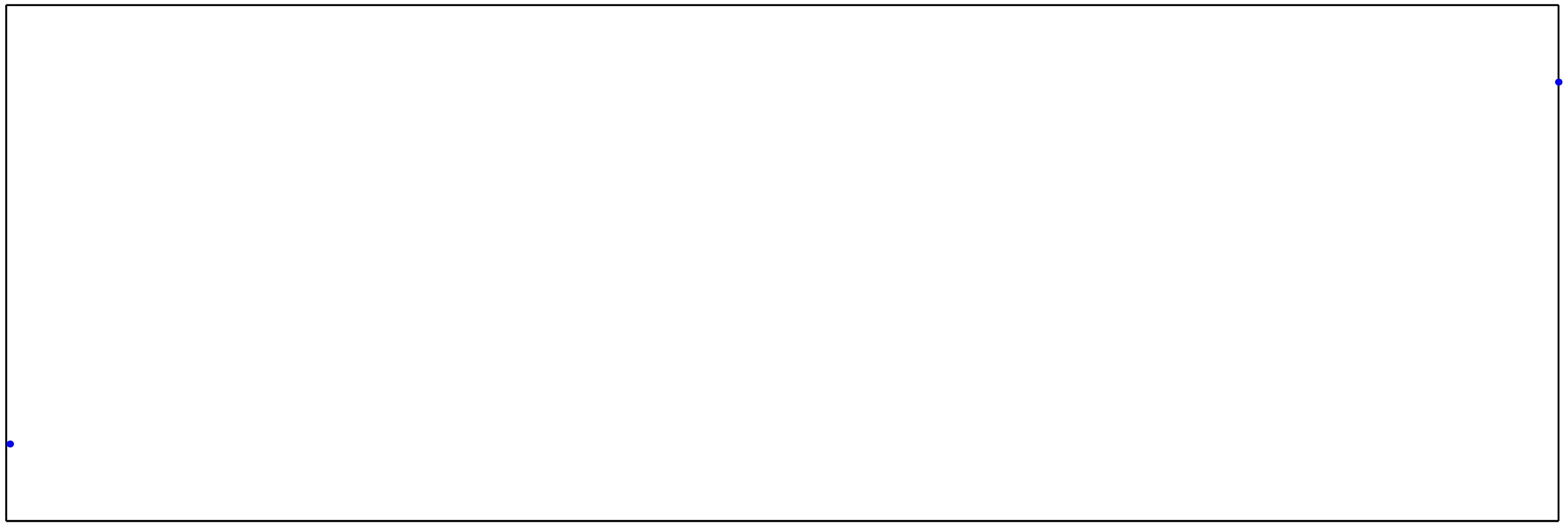


local minima

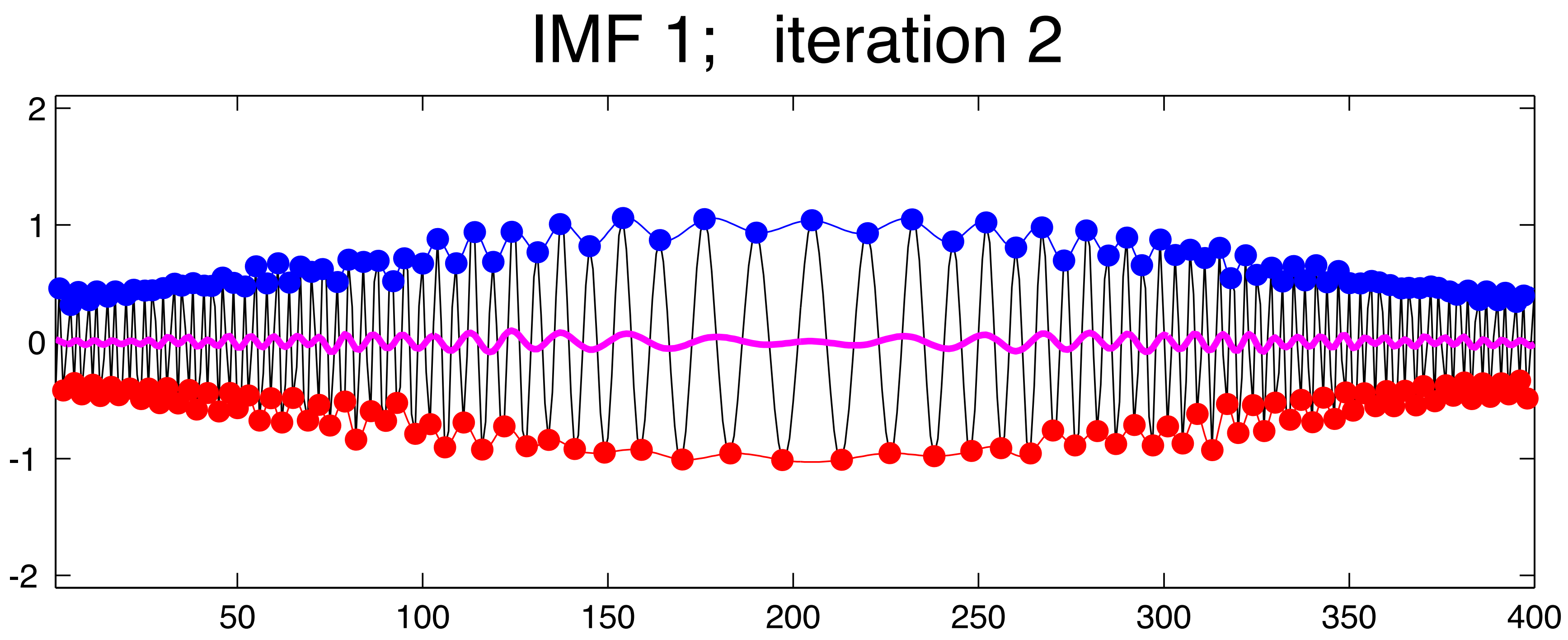


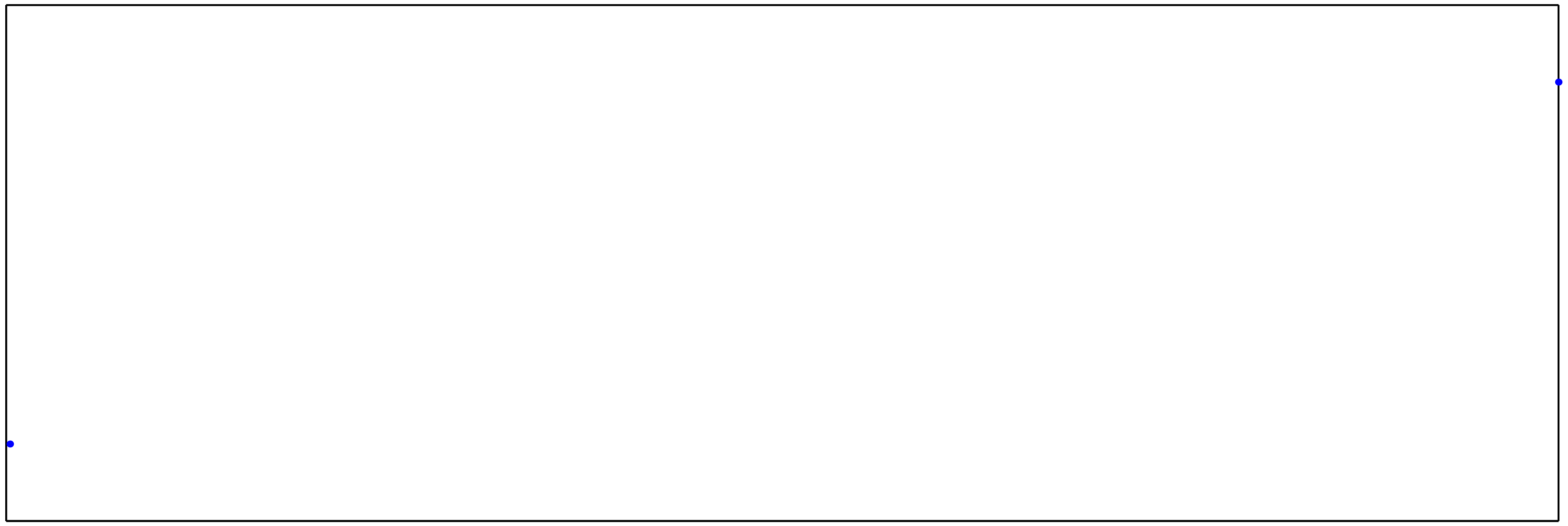
lower envelope



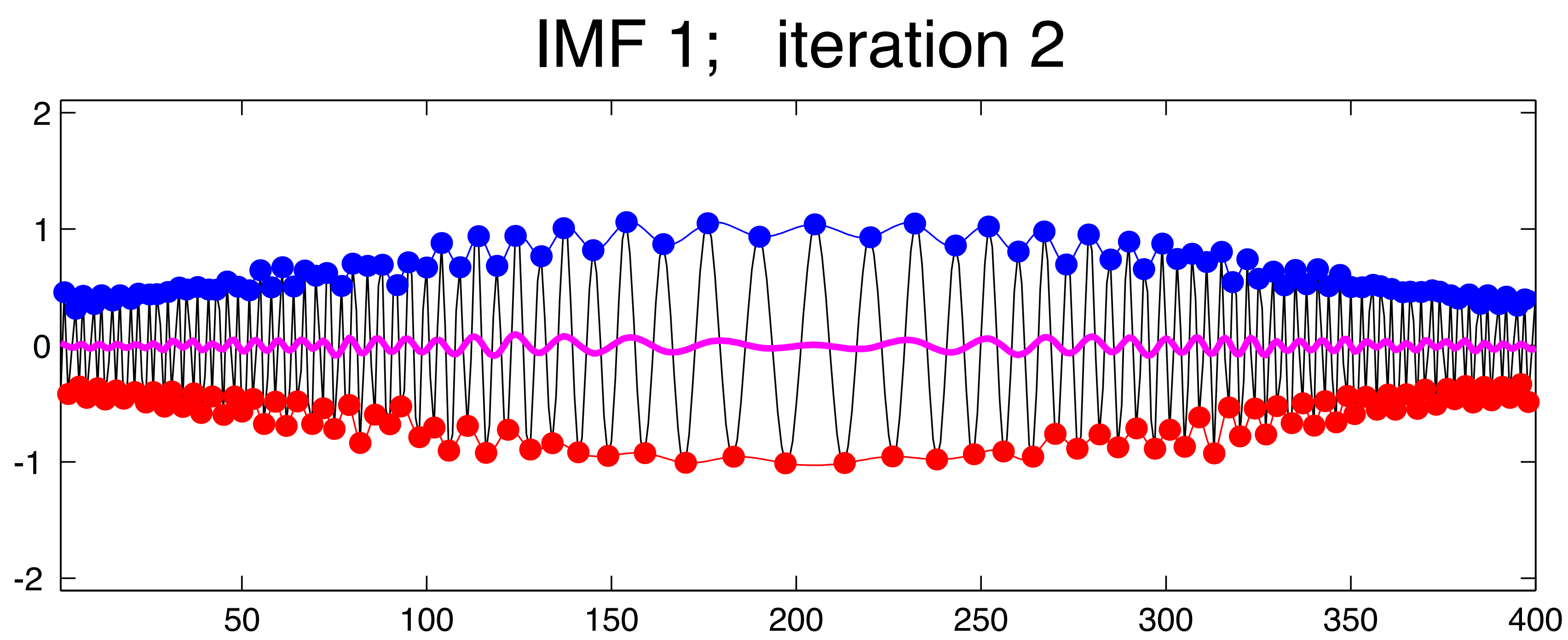


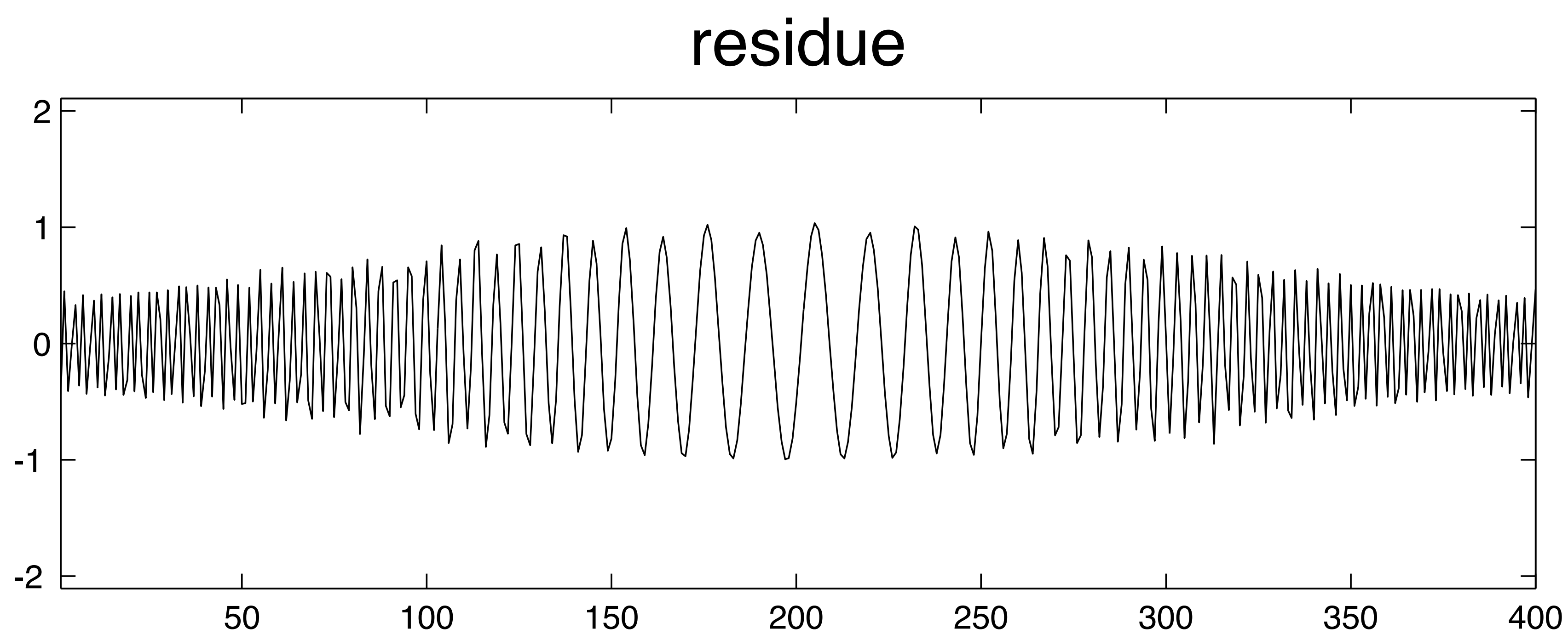
mean envelope



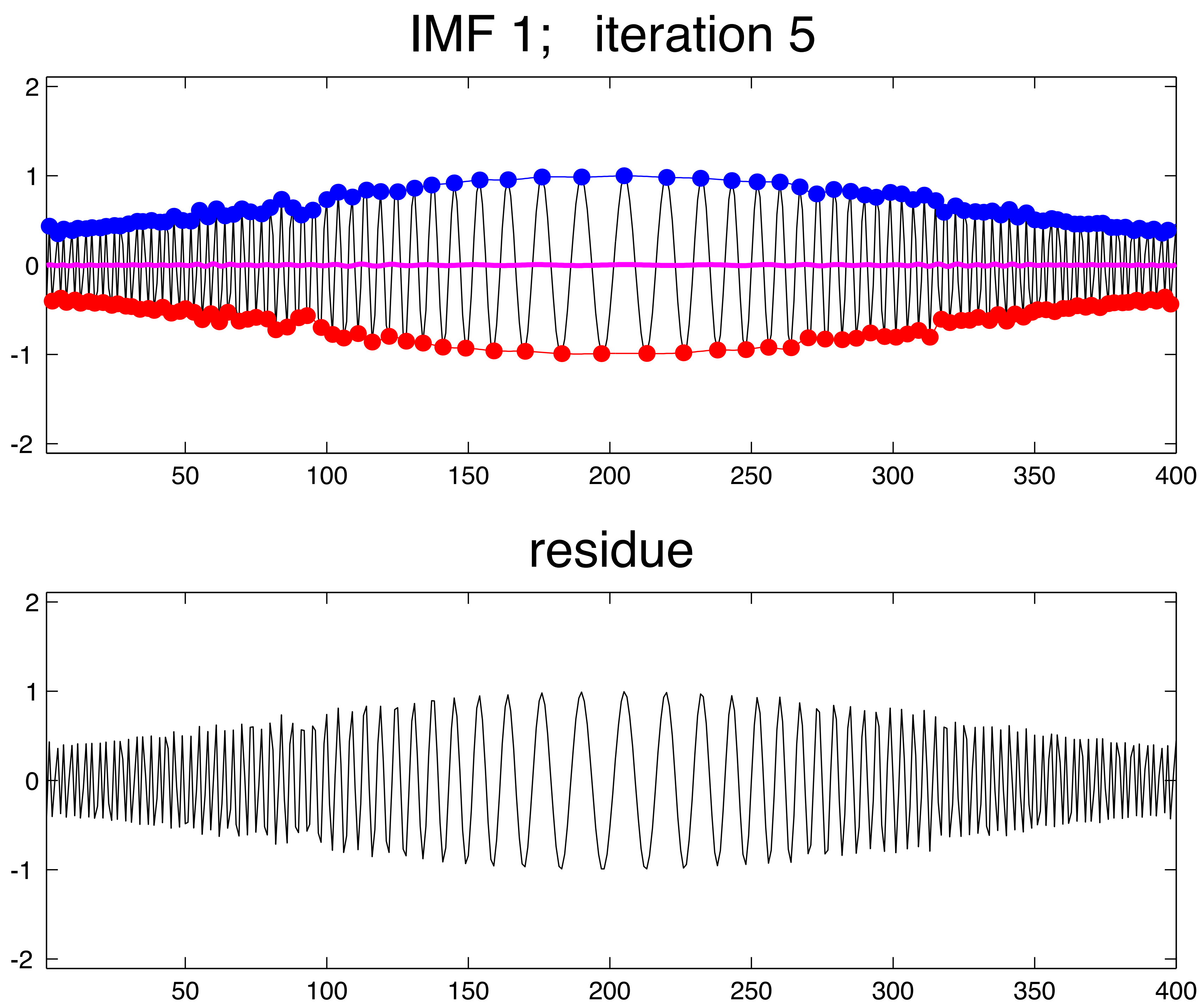


proto-mode 3 = proto-mode 2 – mean envelope



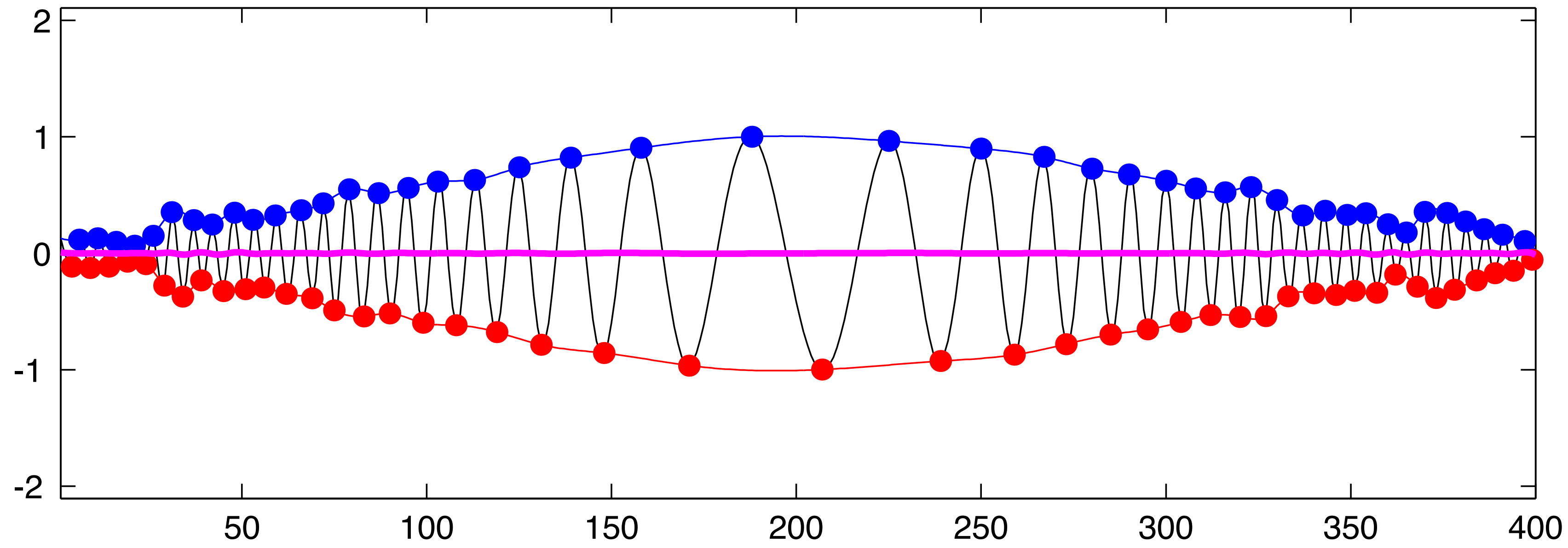


iteration until « zero » mean envelope

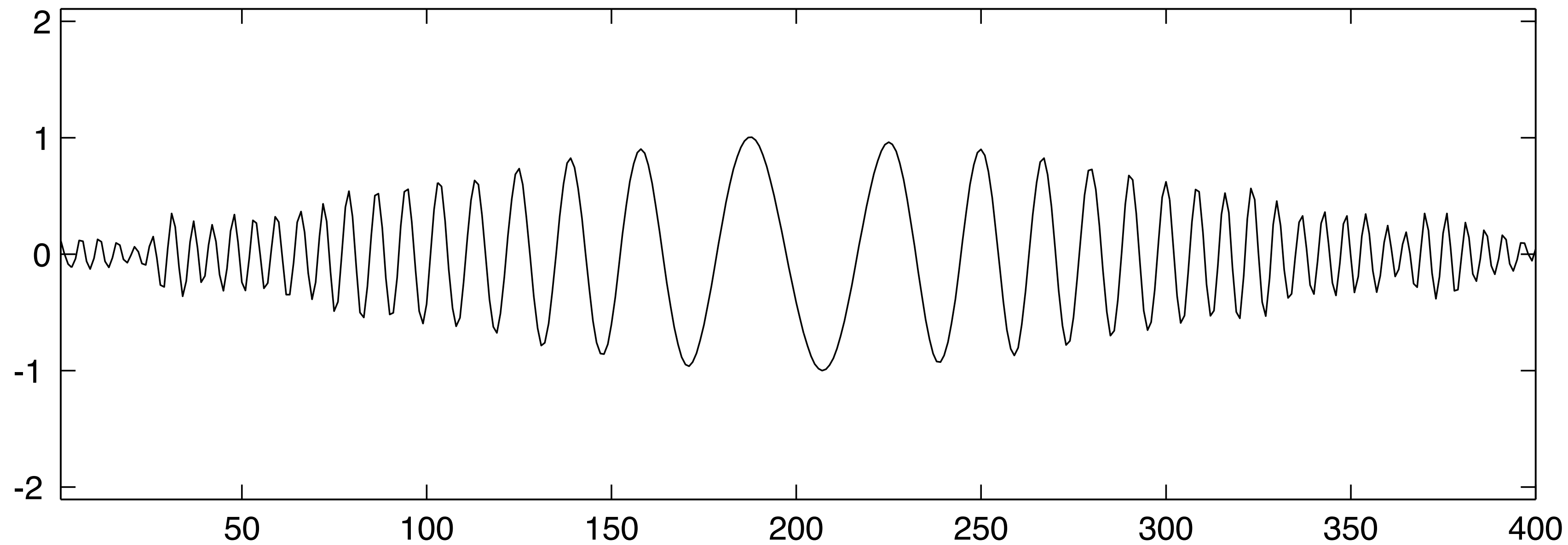


input = signal – mode

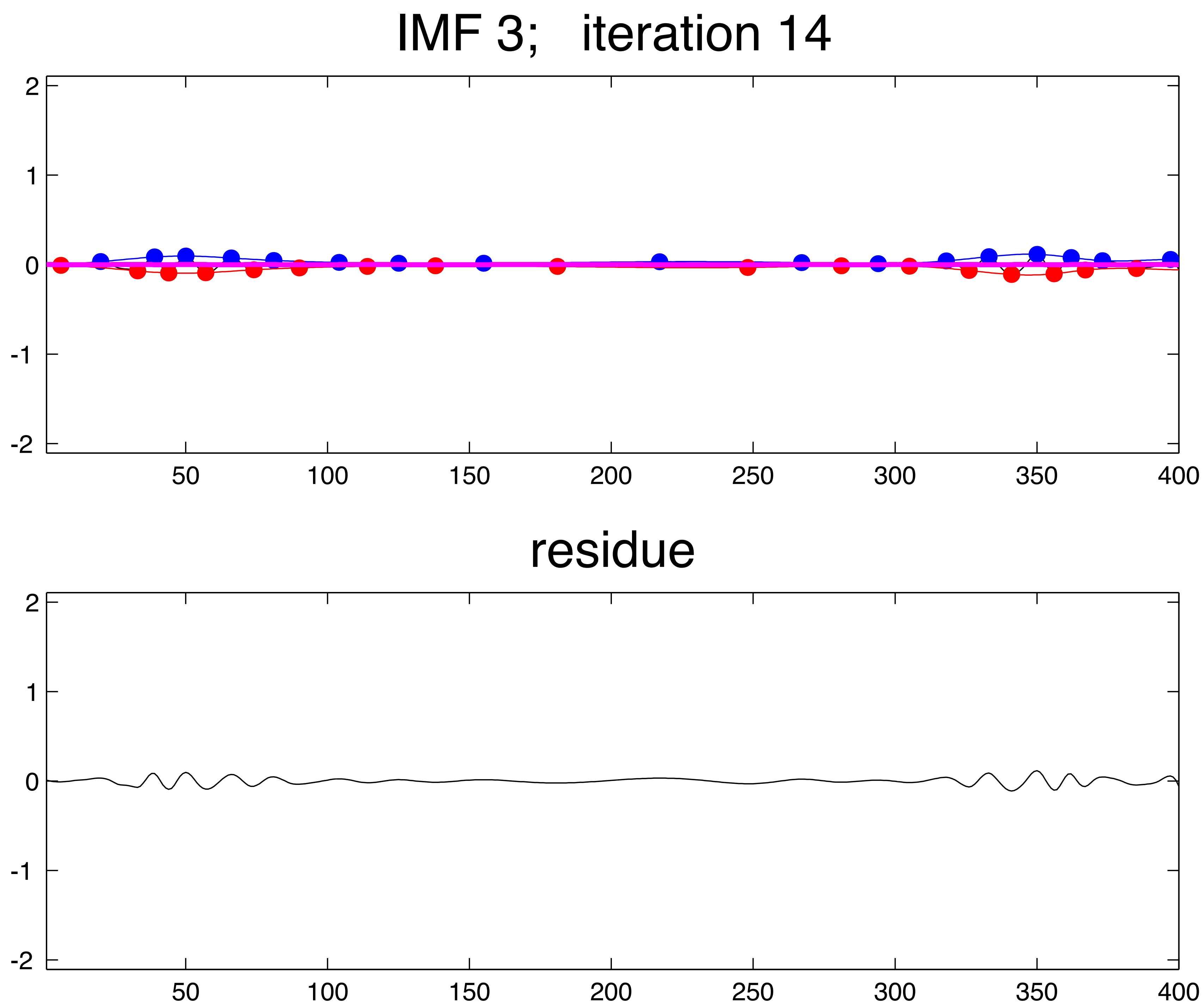
IMF 2;   iteration 2



residue

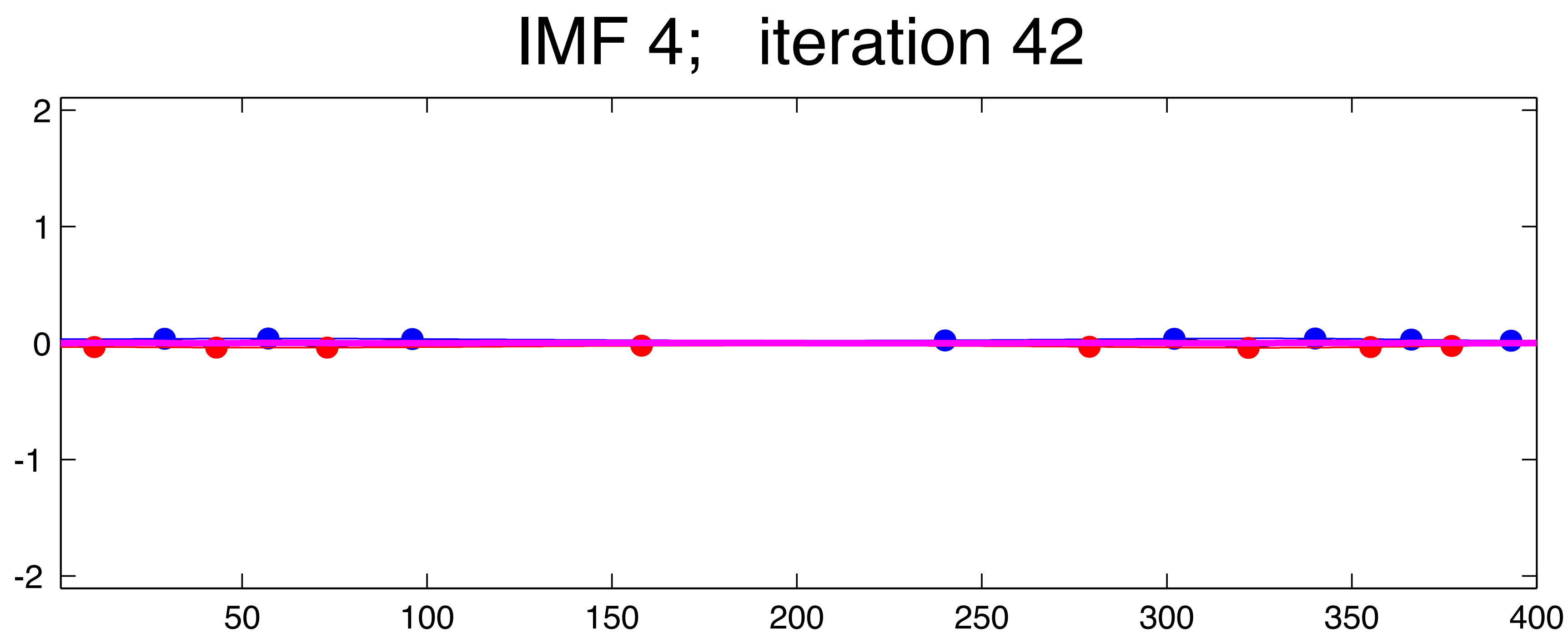


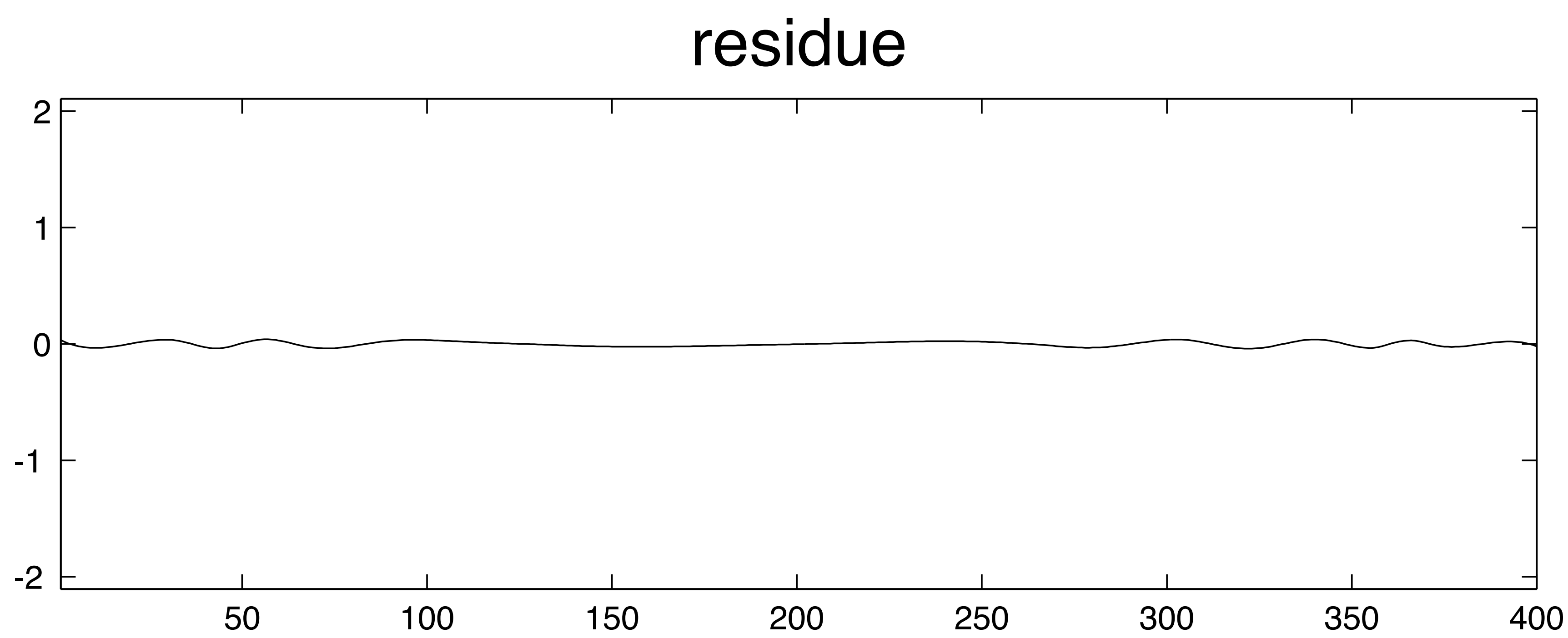
and again…



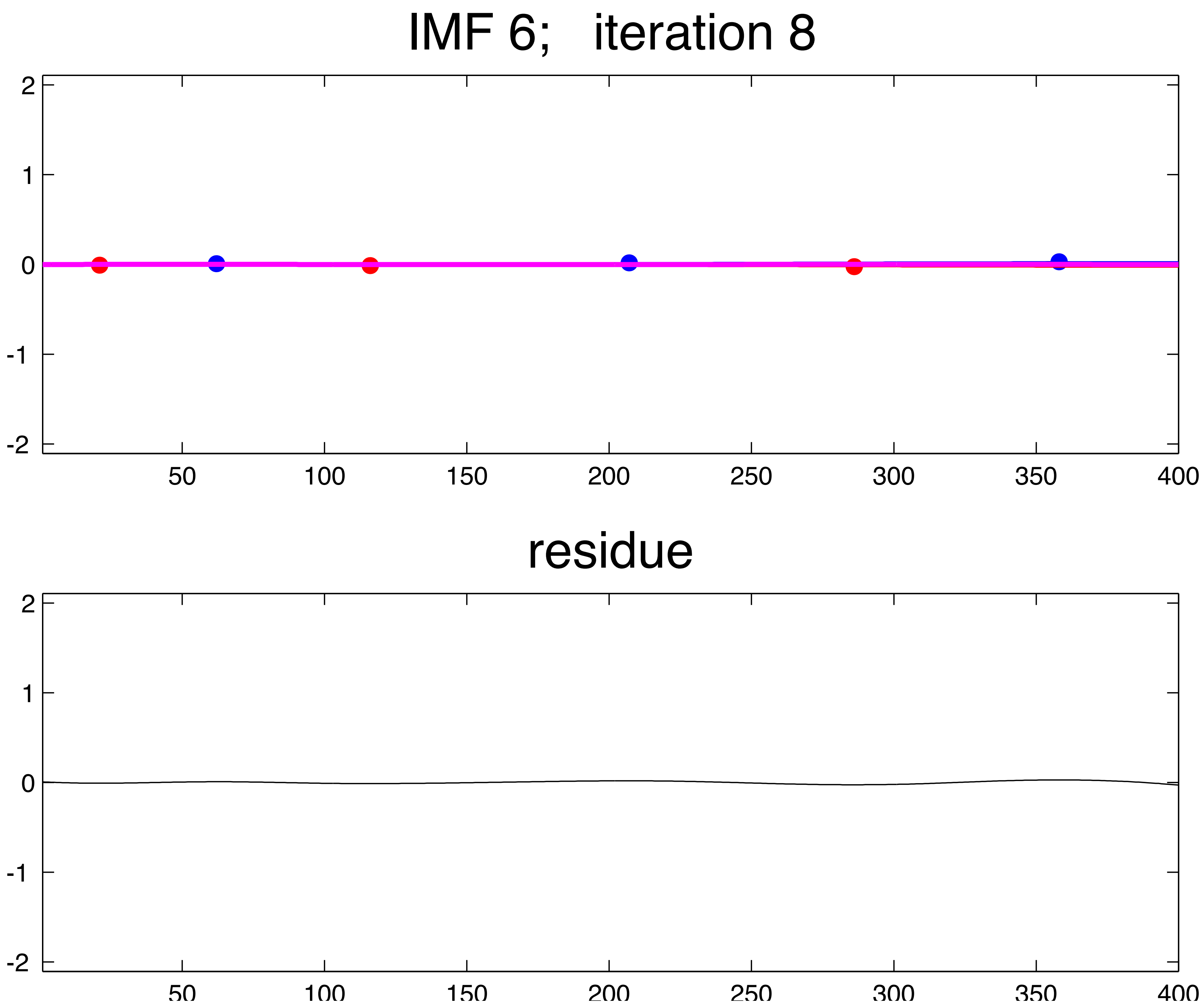
IMF 3; iteration 14
2
1
0
-1
-2
50
100
150
200
250
300
350
400
residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400

and again…


IMF 4; iteration 42
2
1
0
-1
-2
50
100
150
200
250
300
350
400

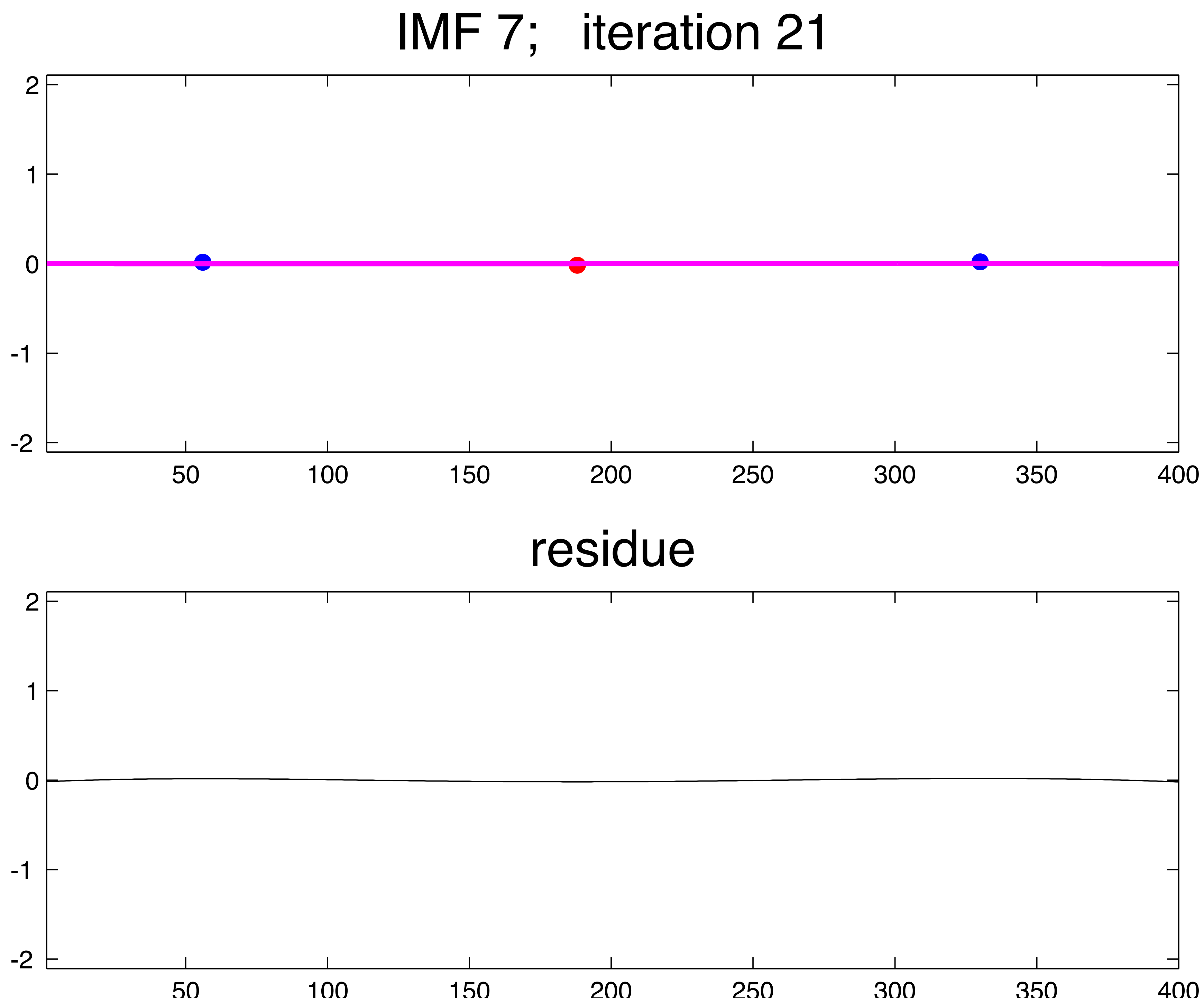

residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400

and again...

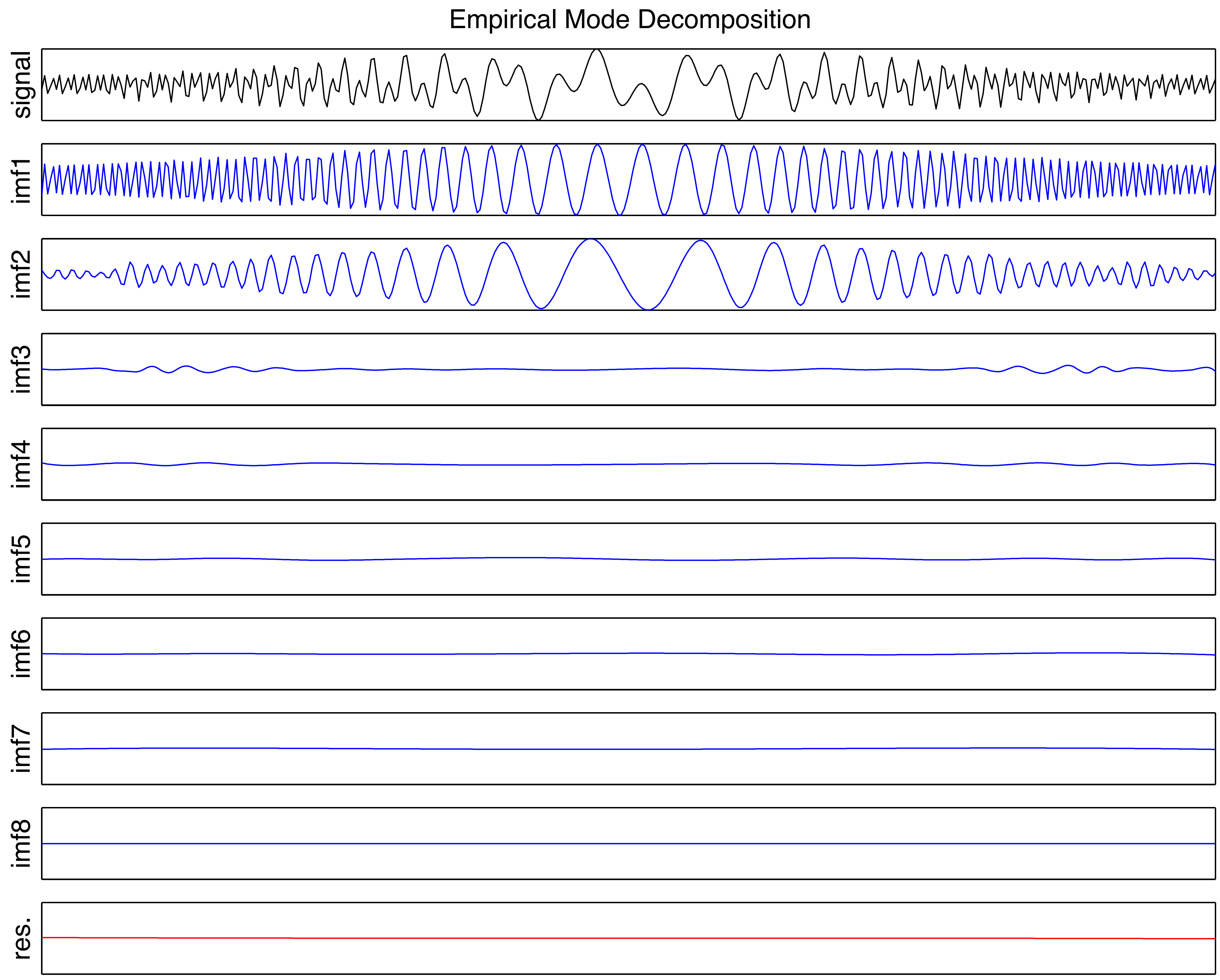


IMF 6; iteration 8
2
1
0
-1
-2
50
100
150
200
250
300
350
400
residue
2
1
0
-1
-2
50
100
150
200
250
300
350
400

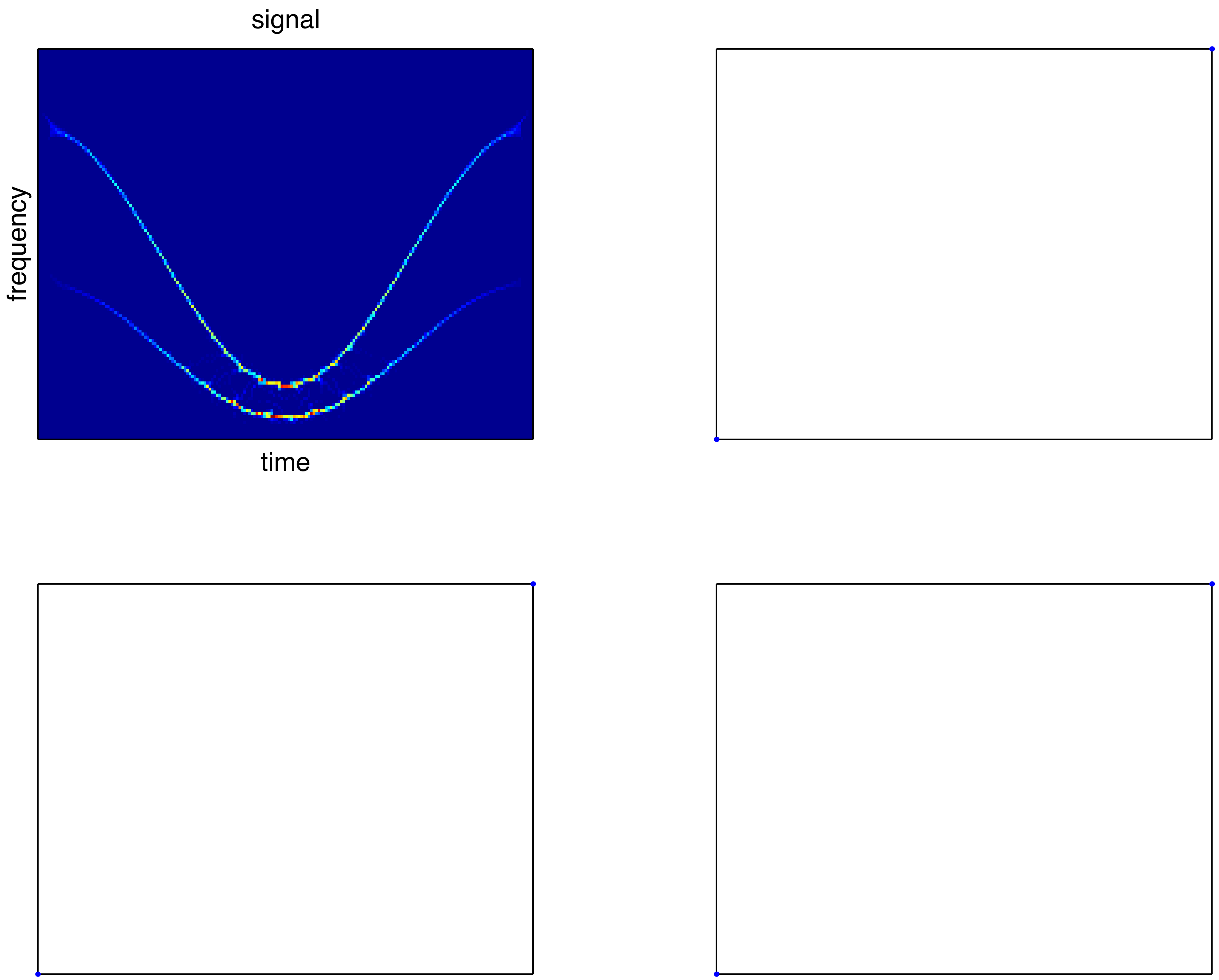
… until no extrema anymore



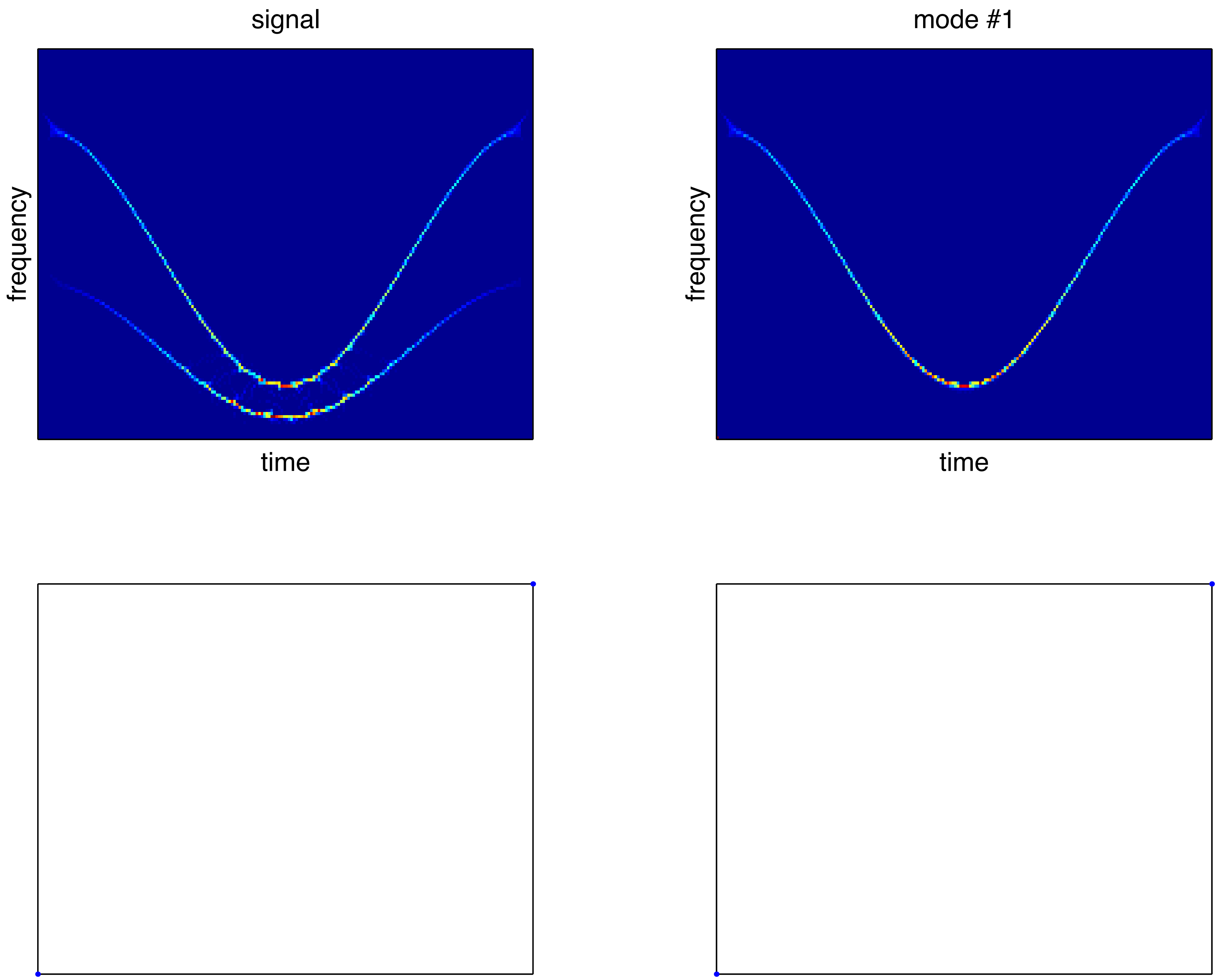
signal = « Intrinsic Mode Functions » + residual



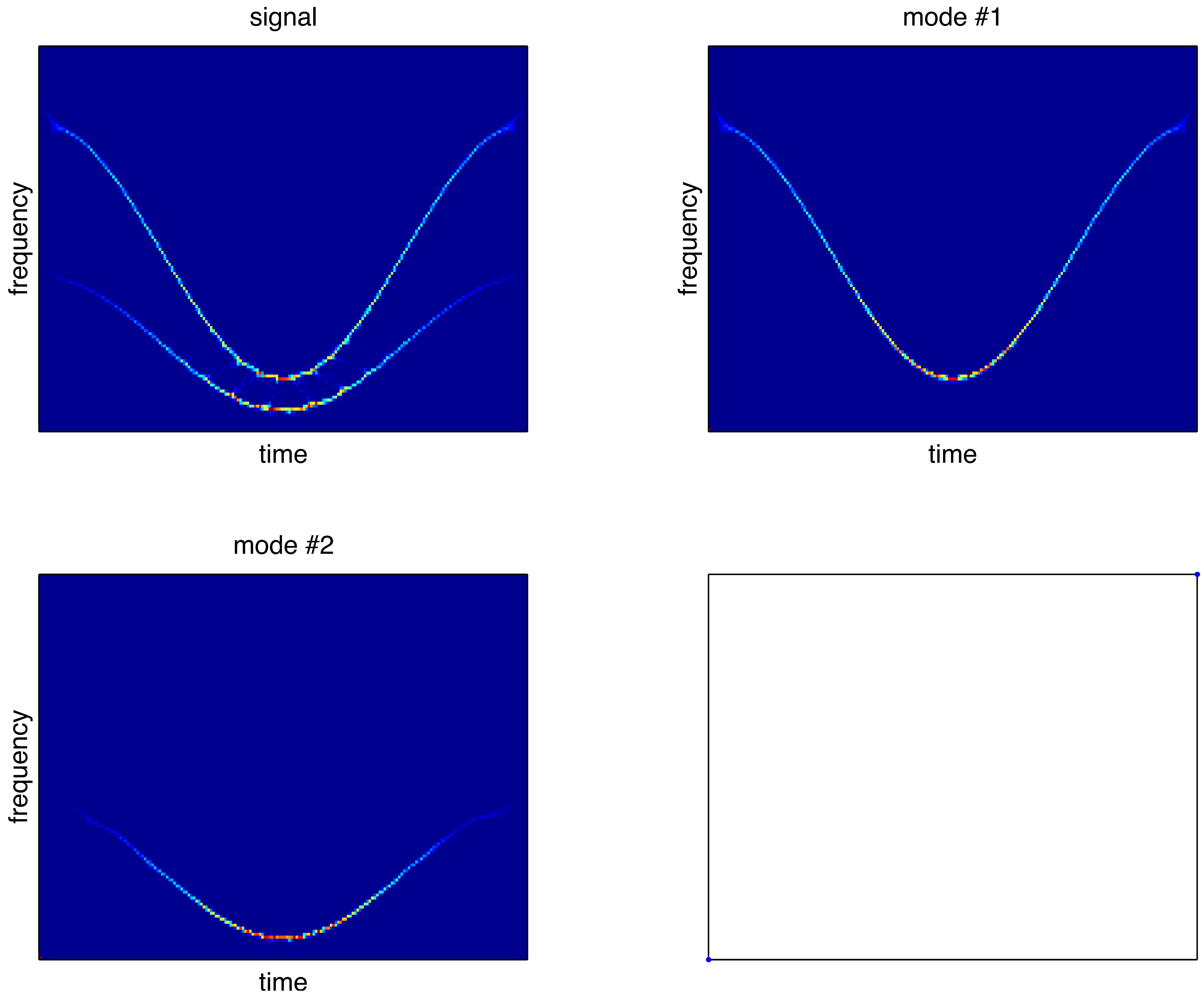
signal



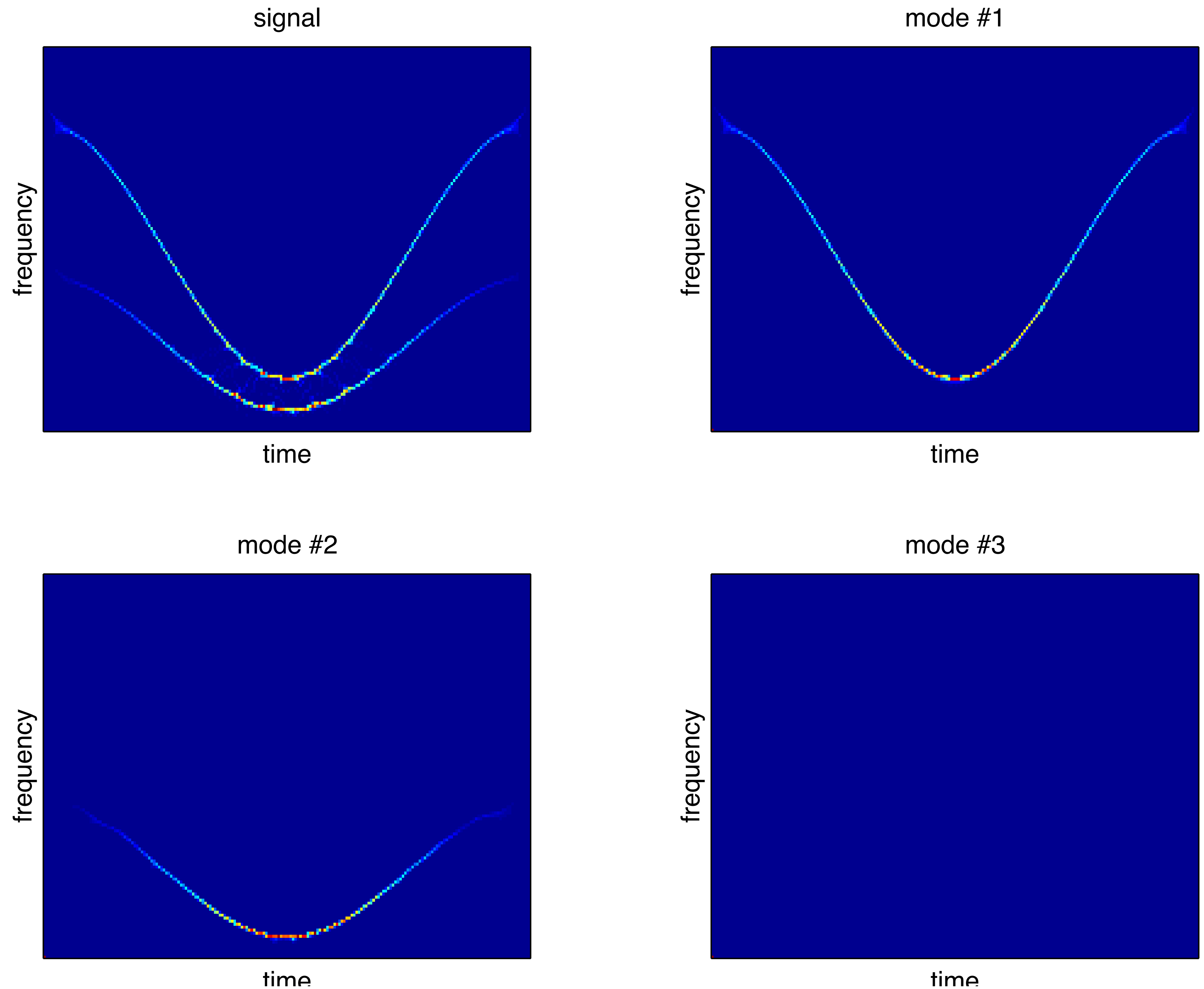
signal = IMF1 + ...



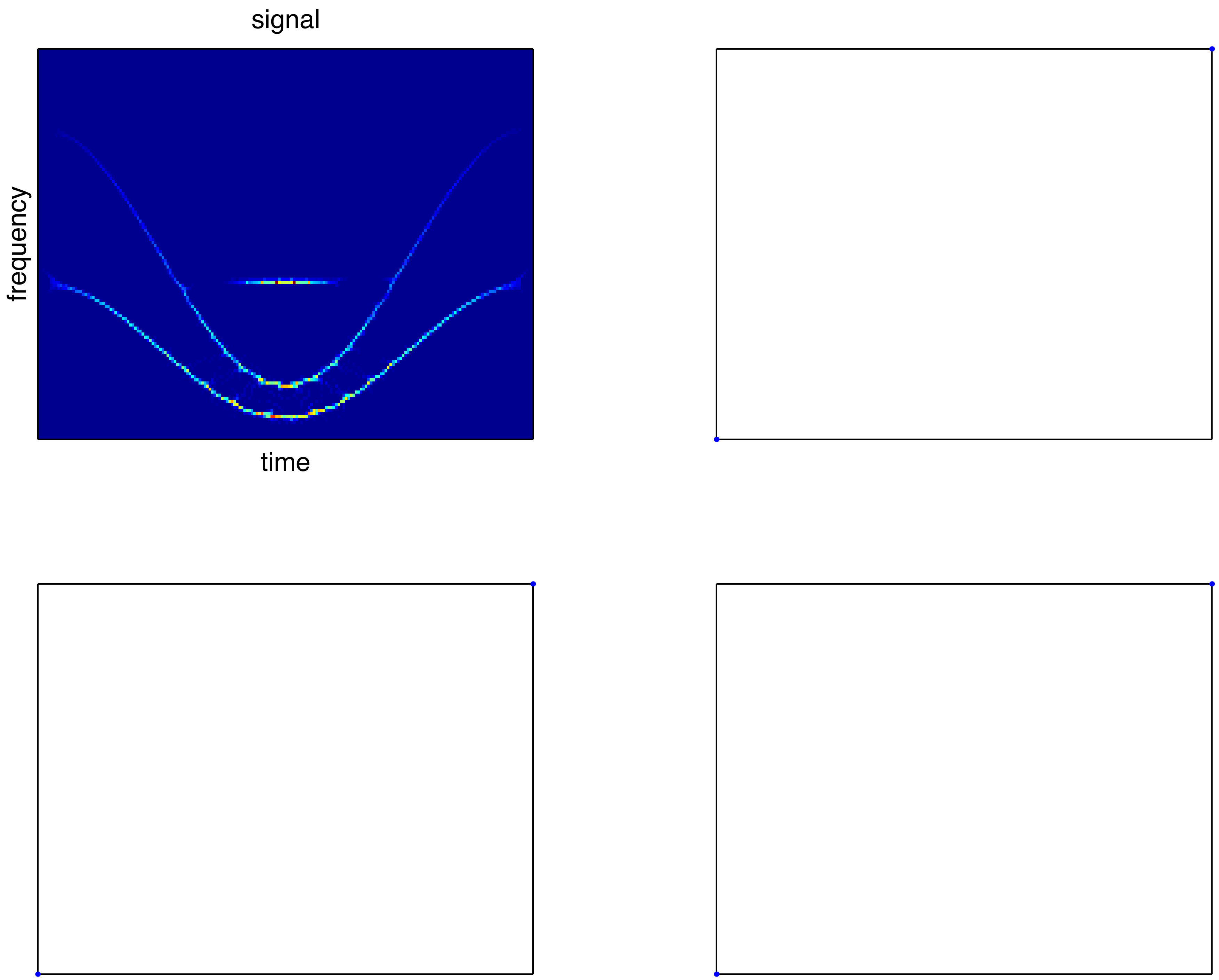
signal = IMF1 + IMF 2 + ...



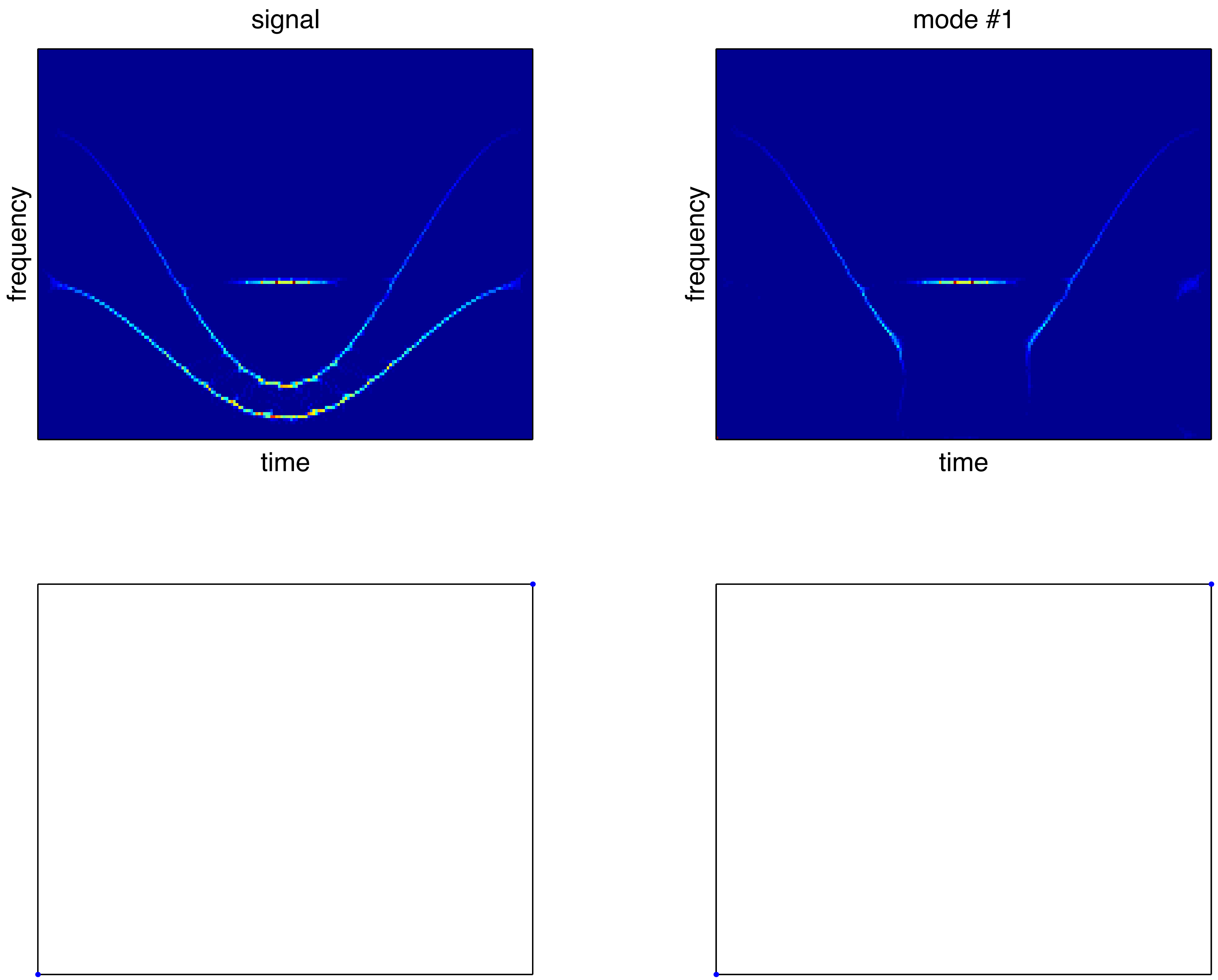
signal = IMF 1 + IMF 2 + residual



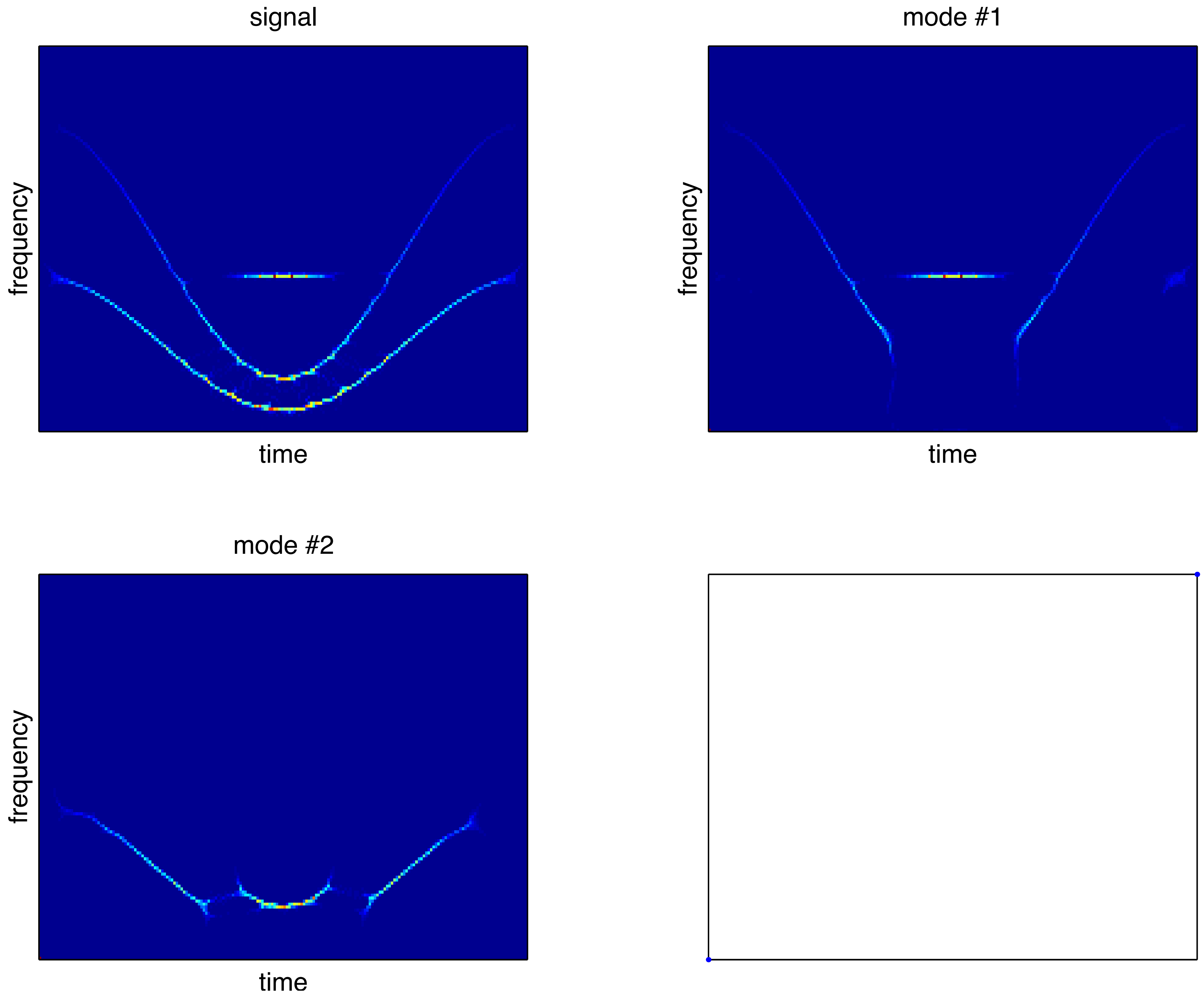
signal



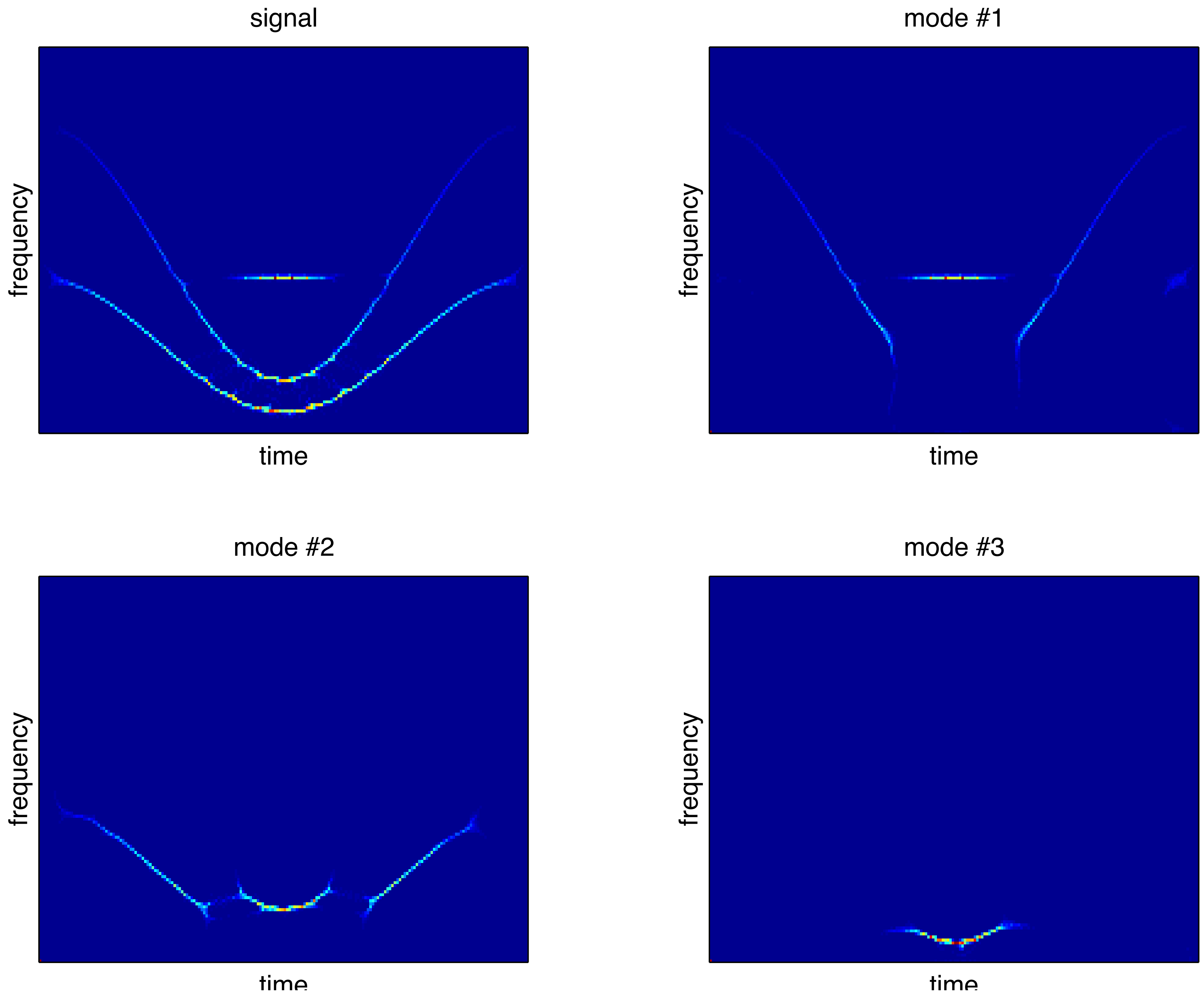
signal = IMF 1 + ...



signal = IMF 1 + IMF 2 + …



signal = IMF 1 + IMF 2 + residual



Interpreting EMD

- Decomposition = **output** of an algorithm (iterative, with an inner loop)
- Detailed performance analysis **unreachable**
- Gain insight from generic **test-situations**:
  - *disentangling* ***tones***
  - *decomposing* ***noise***
- Approach expected to pave the way for **variants**, **processing** and and **effective uses** on real data

1 or 2 frequencies?

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\left(\frac{\omega_1 + \omega_2}{2} t\right)\cos\left(\frac{\omega_1 - \omega_2}{2} t\right)$$

- **Mathematical equivalence**, with different physical interpretations

1 or 2 frequencies?

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\left(\frac{\omega_1 + \omega_2}{2} t\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t\right)$$

- Mathematical equivalence, with different physical interpretations

*2 distinct, **constant amplitude** tones*

1 or 2 frequencies?

$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\left(\frac{\omega_1 + \omega_2}{2} t\right) \cos\left(\frac{\omega_1 - \omega_2}{2} t\right)$$

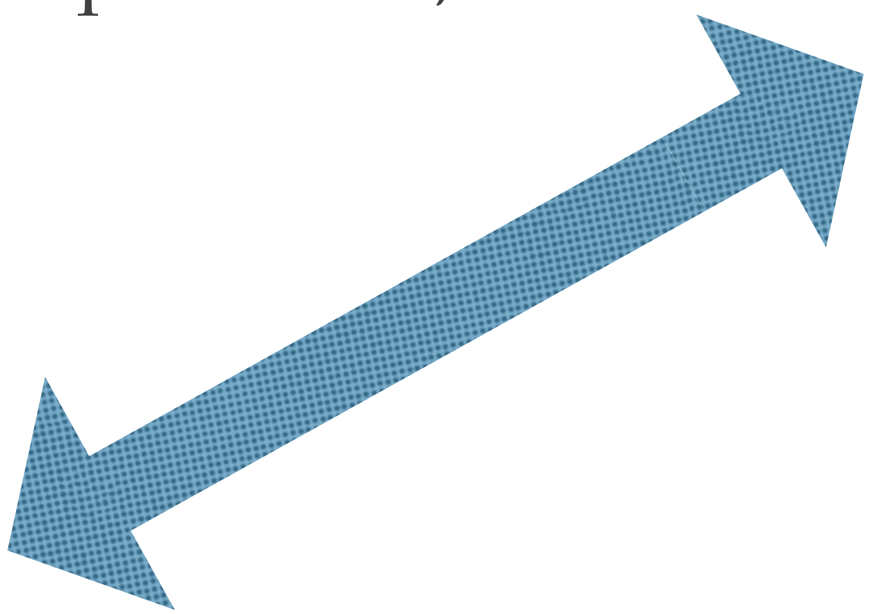
- Mathematical equivalence, with different physical interpretations

*1 single, **amplitude modulated** tone (« beating effect »)*

1 or 2 frequencies?

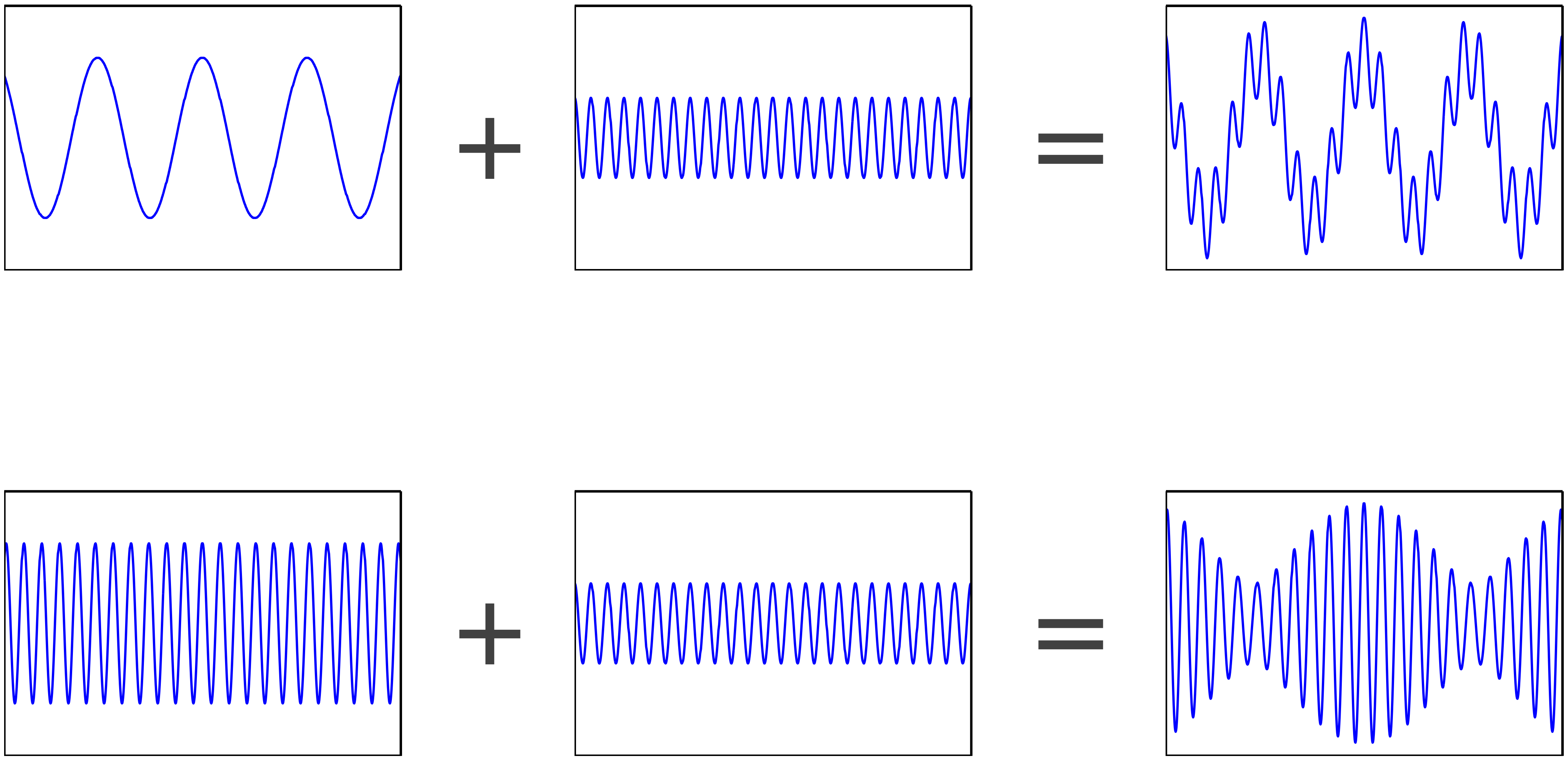
$$\cos(\omega_1 t) + \cos(\omega_2 t) = 2\cos\left(\frac{\omega_1 + \omega_2}{2} t\right)\cos\left(\frac{\omega_1 - \omega_2}{2} t\right)$$

- Mathematical equivalence, with **different physical interpretations**

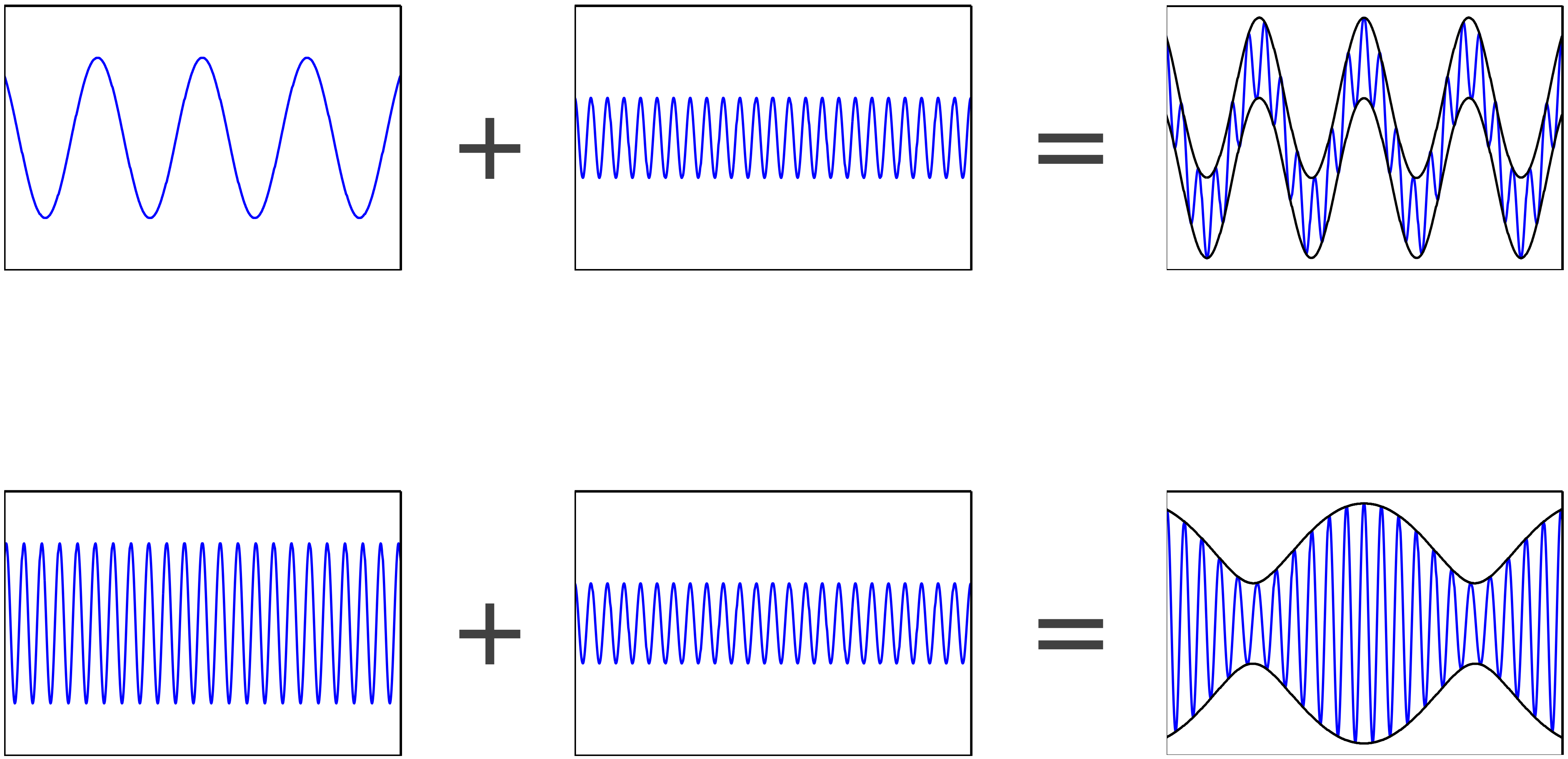


- **Context-based** preference

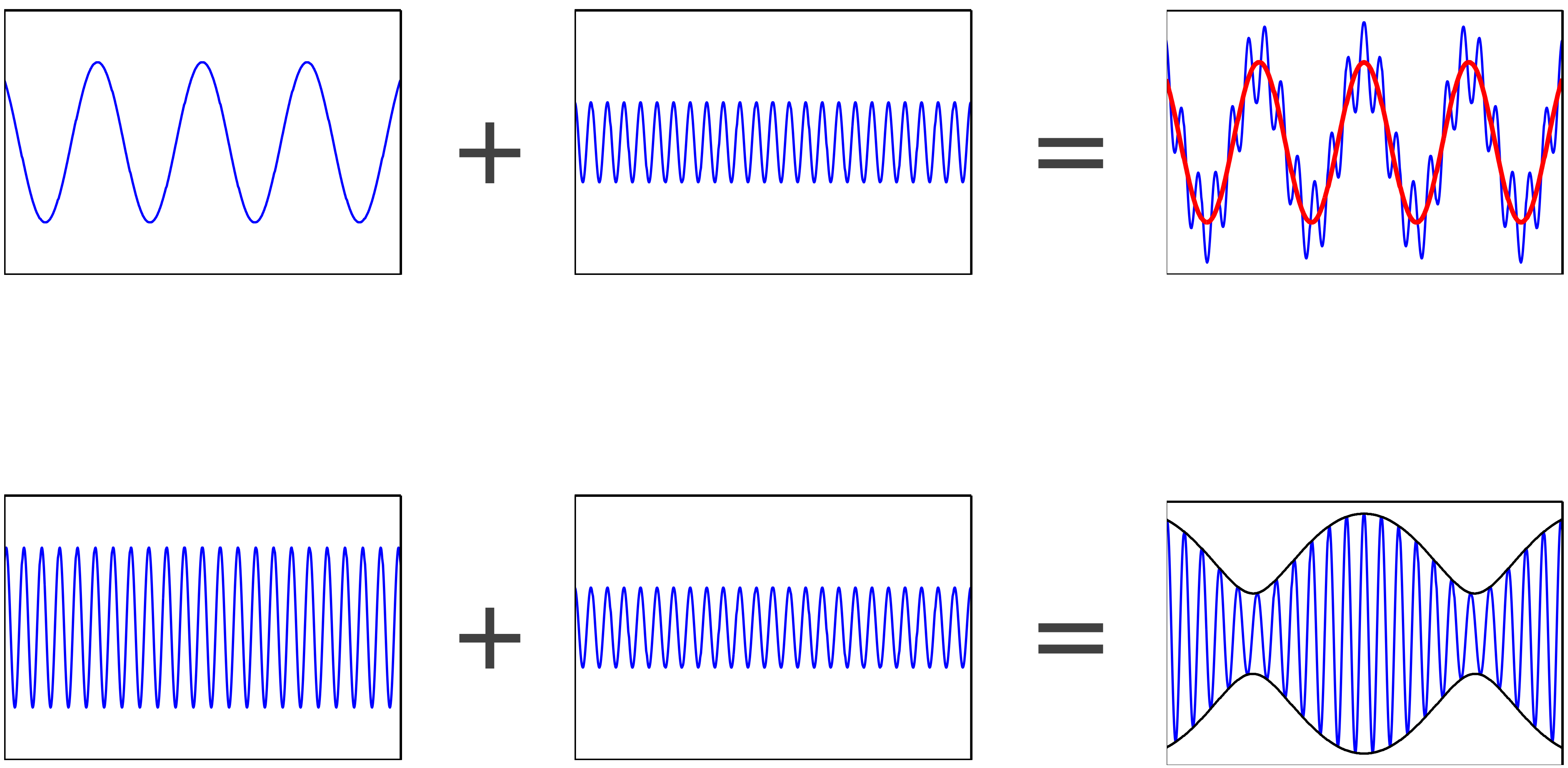
1 or 2 frequencies?



1 or 2 frequencies?

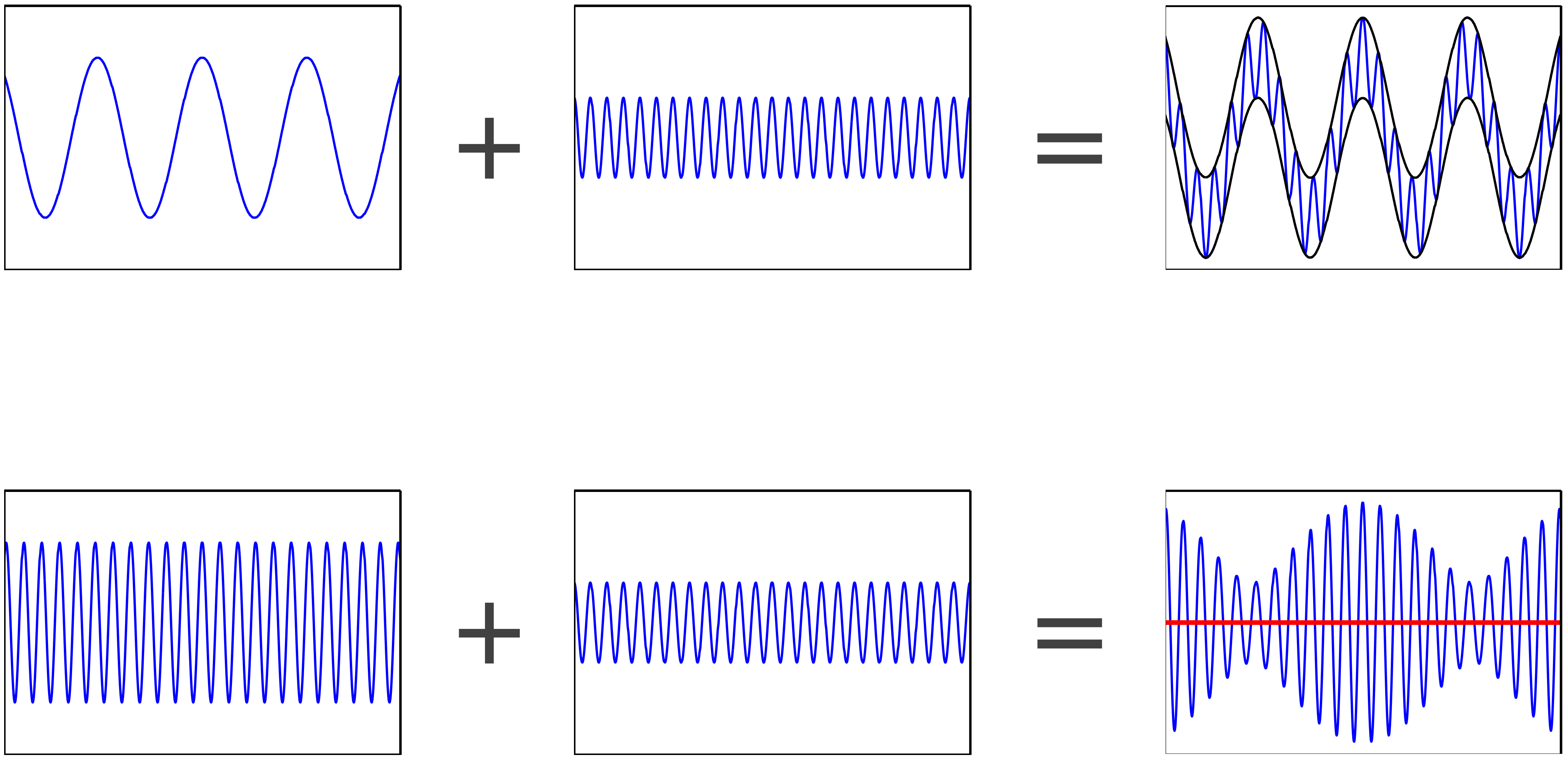


1 or 2 frequencies?



2

1 or 2 frequencies?



1

1 or 2 frequencies?

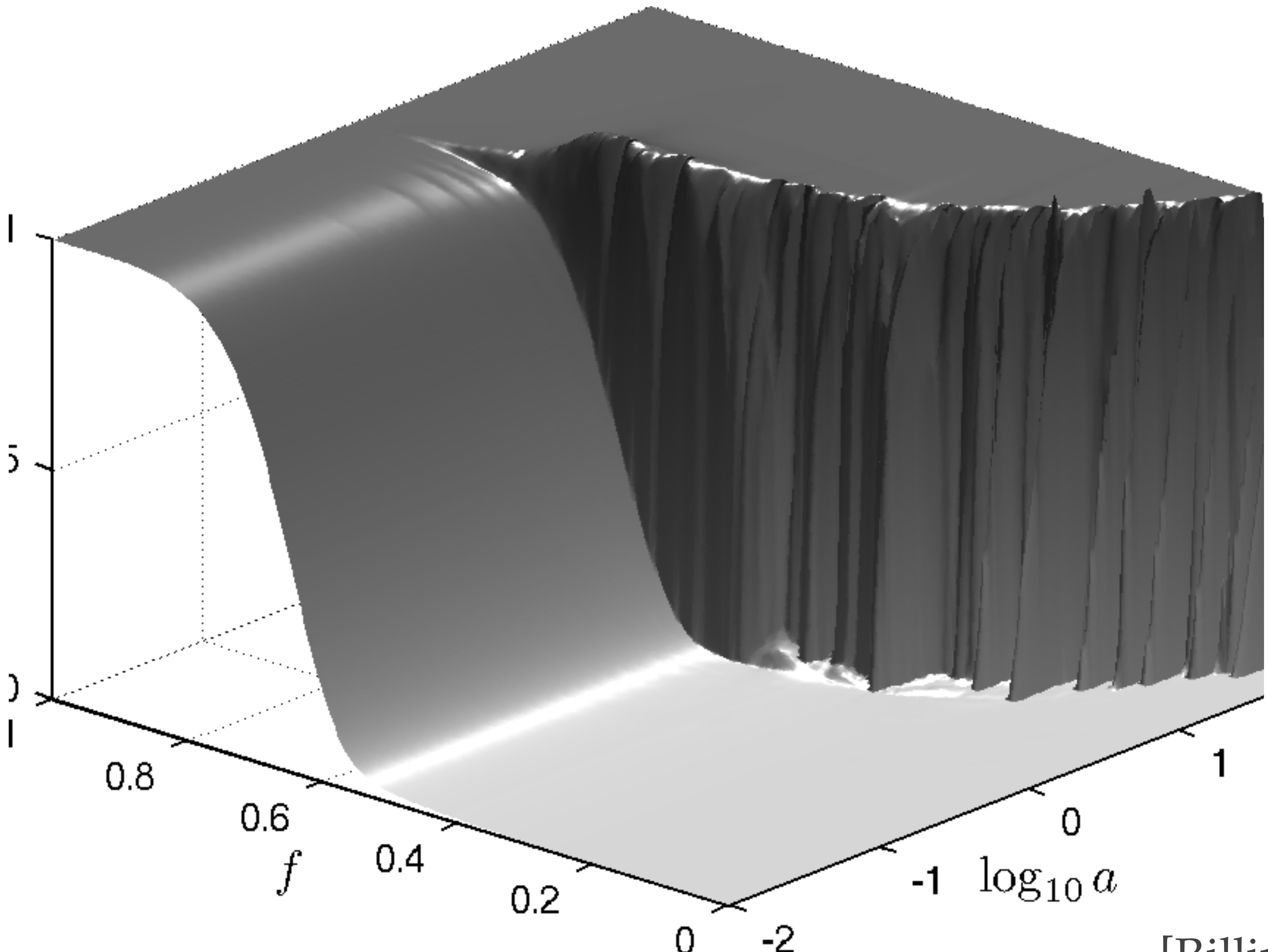
- **Model**: $$x(t) = \underbrace{a_1 \cos(2\pi f_1 t)}_{x_1(t)} + \underbrace{a_2 \cos(2\pi f_2 t + \varphi)}_{x_2(t)}, \quad f_1 > f_2$$

- **Analysis**: if separation, the 1st IMF should be equal to the highest frequency component

- **Criterion** (= 0 if separation):

$$c\left(\frac{a_2}{a_1}, \frac{f_2}{f_1}, \varphi\right) = \frac{\|\mathrm{IMF}_1(t) - x_1(t)\|_{\ell_2}}{\|x_2(t)\|_{\ell_2}}$$

1 or 2 frequencies?



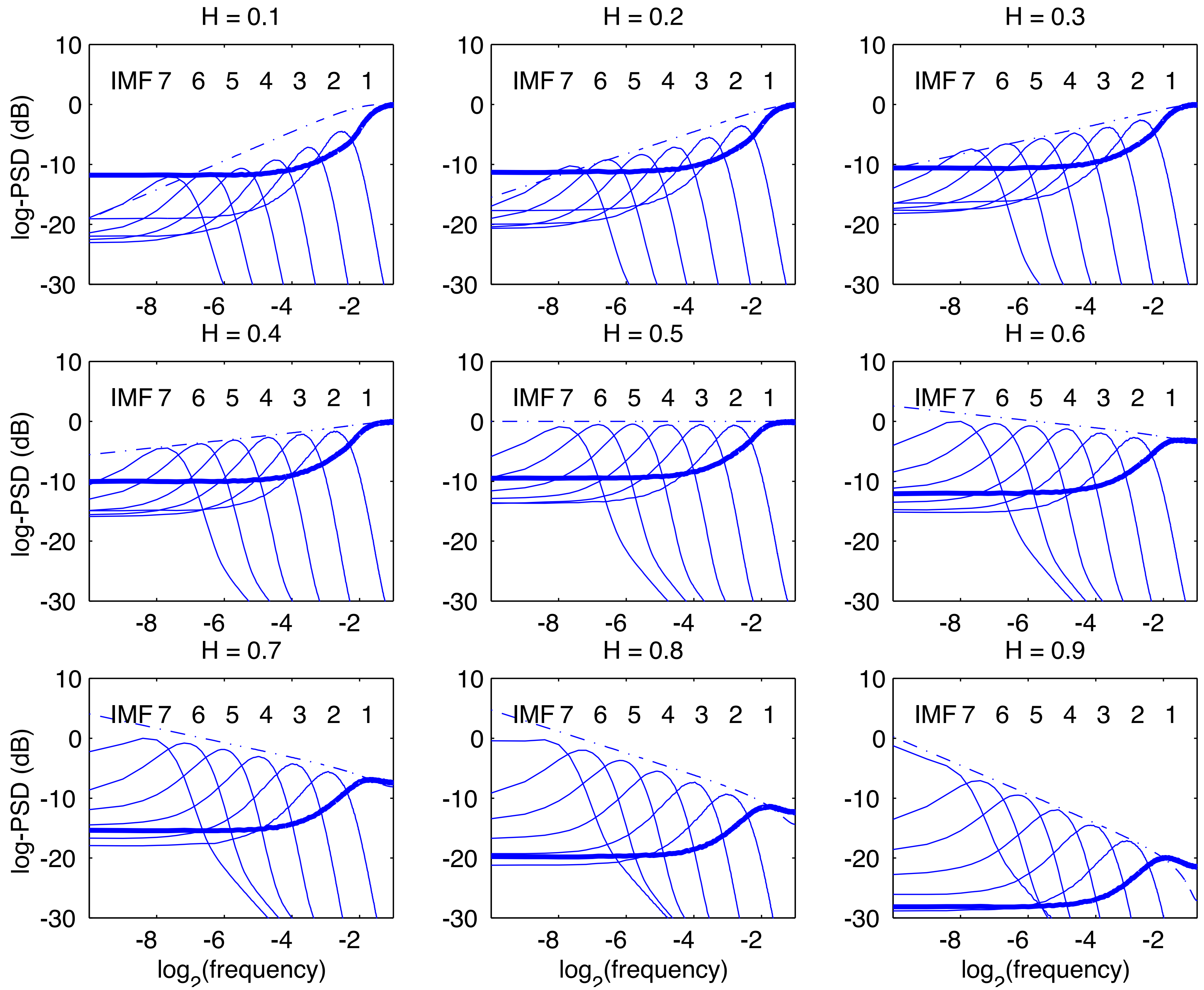
[Rilling & F., ’10]

Multiresolution

- **Stochastic frequency approach**: decomposition and spectrum analysis, mode by mode, of a wideband noise.
- **Model**: Fractional Gaussian noise (extension of white Gaussian noise, with short/long range dependencies)
- **Result** [F., Gonçalvès & Rilling, '03]: « Spontaneous » emergence of a quasi-dyadic, self-similar, filterbank structure

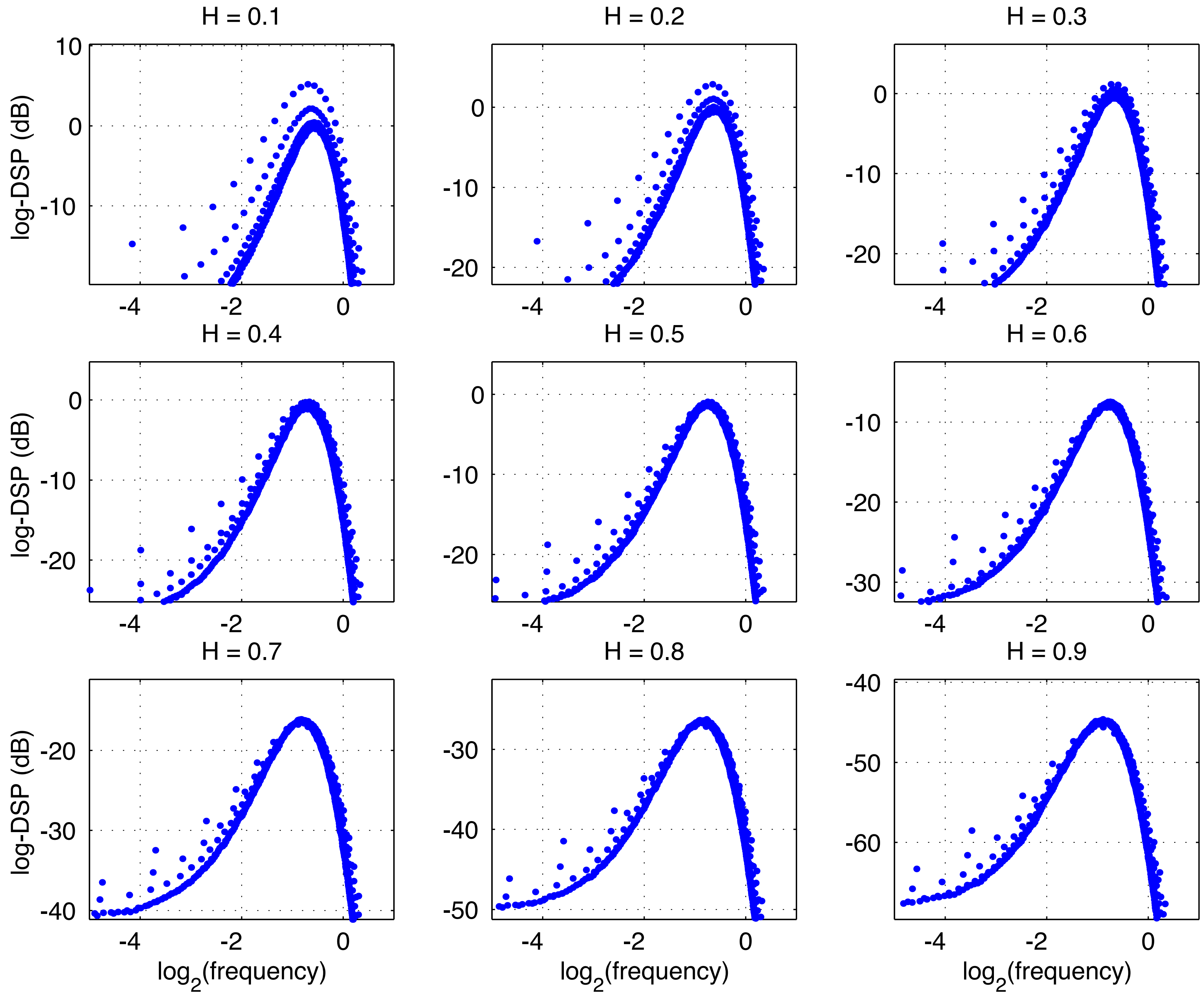
$$\mathcal{S}_{k',H}(f) \sim 2^{(2H-1)(k'-k)}\,\mathcal{S}_{k,H}(2^{k'-k}f)$$

Multiresolution



[F., Rilling & Gonçalvès, ’03]

Multiresolution

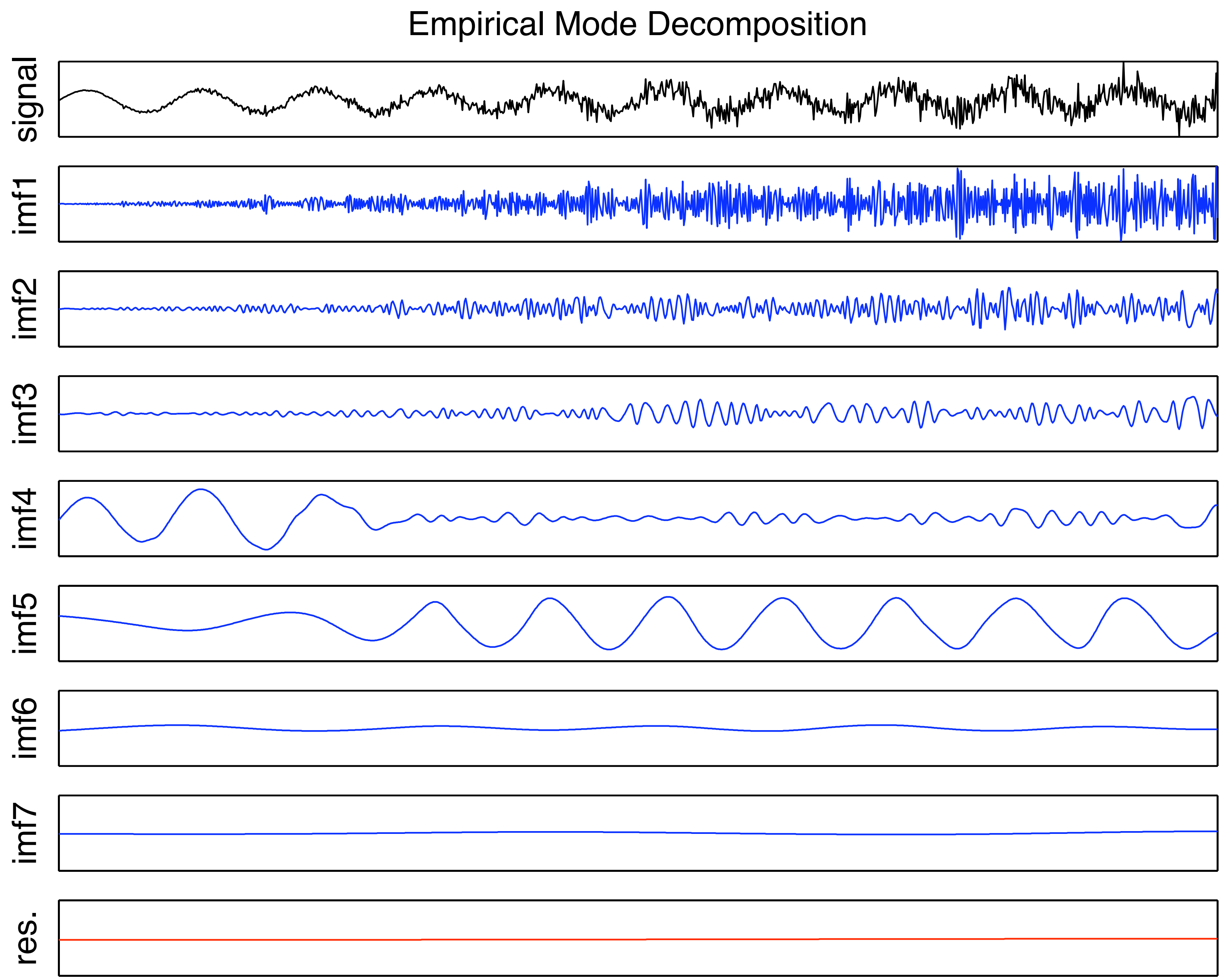


[F., Rilling & Gonçalvès, ’03]

Some variations on the theme

- EMD as a **general principle**
- Many **variations**, **extensions** and **uses** over the last 15 years
- Elements on
  - ***Ensemble EMD***: *noise-assisted variant to prevent « mode mixing »*
  - ***Multivariate extensions***
  - ***Optimization-based*** *alternatives*
  - *EMD-based* ***denoising/detrending***

« Mode mixing »


Empirical Mode Decomposition
signal
imf1
imf2
imf3
imf4
imf5
imf6
imf7
res.

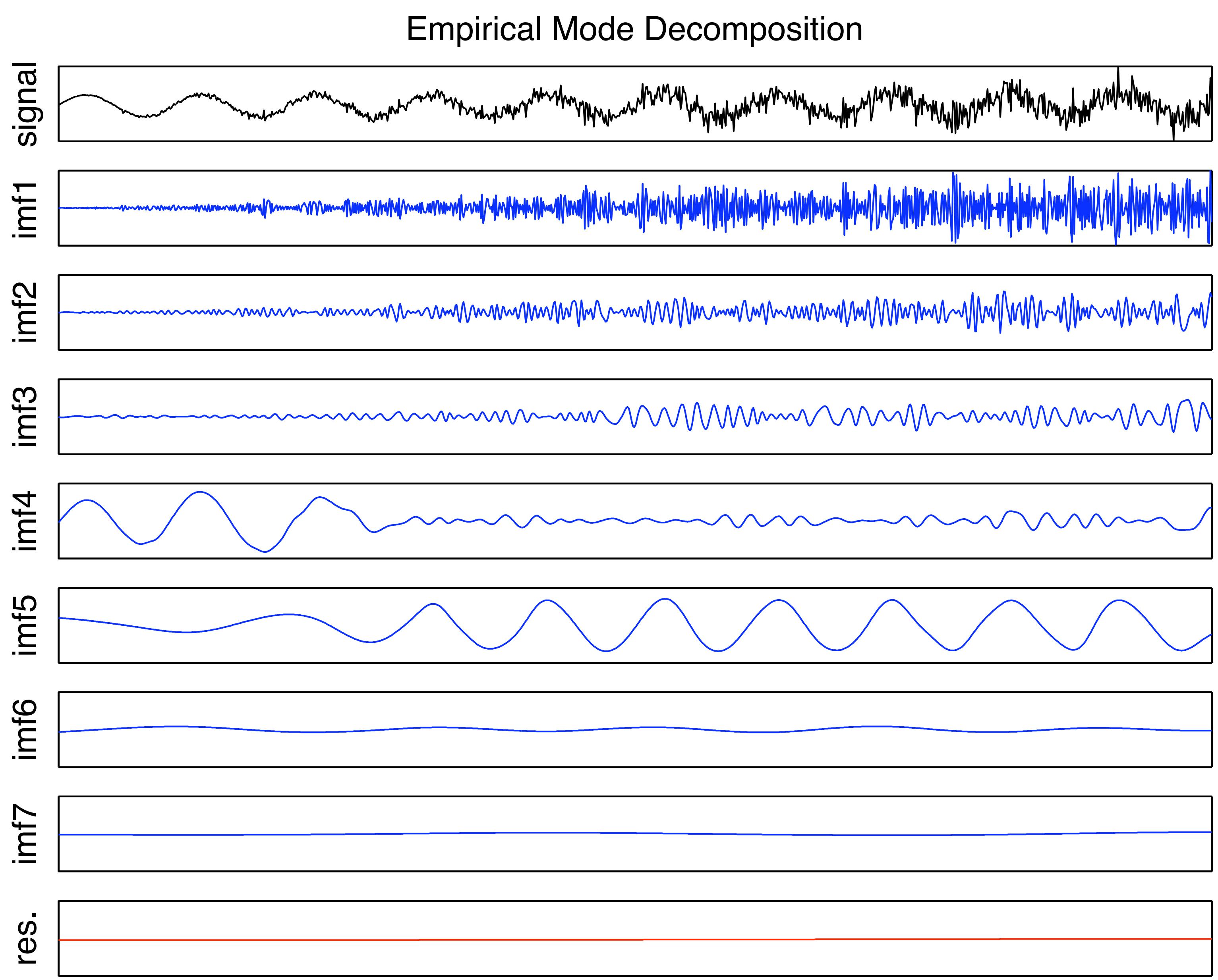
Ensemble EMD

- **Rationale**: noise-assisted variant to prevent « mode mixing »
- **Implementation** [Wu & Huang, '09]

  1. ***add*** *some controlled* ***noise*** *to data*
  2. ***compute*** *EMD*
  3. ***iterate*** *on a number of realizations and* ***average***

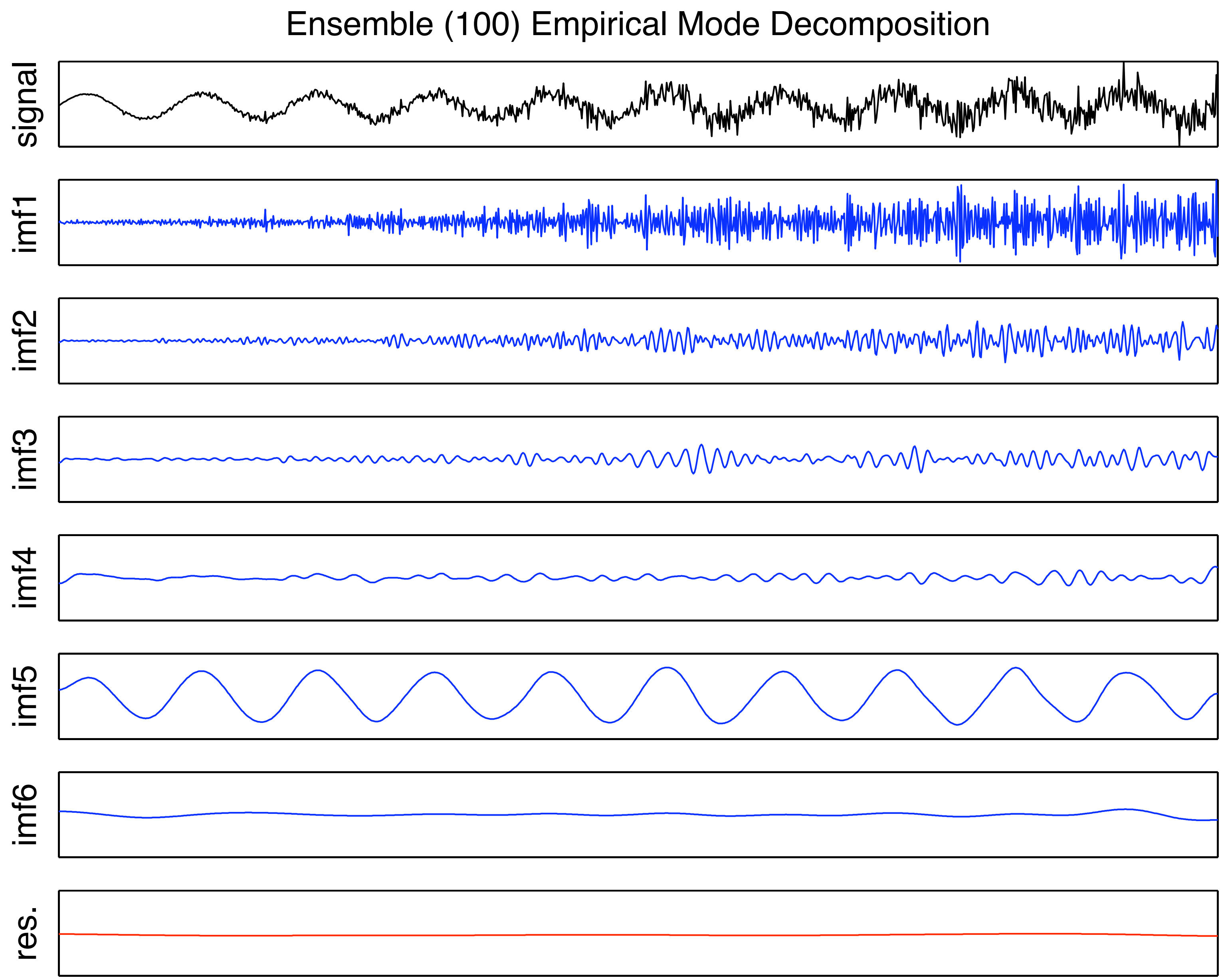
- **Improvement**: CEEMDAN (« Complete EEMD with Adaptive Noise ») [Torres *et al.*, '11]

« Mode mixing »



Empirical Mode Decomposition
signal
imf1
imf2
imf3
imf4
imf5
imf6
imf7
res.

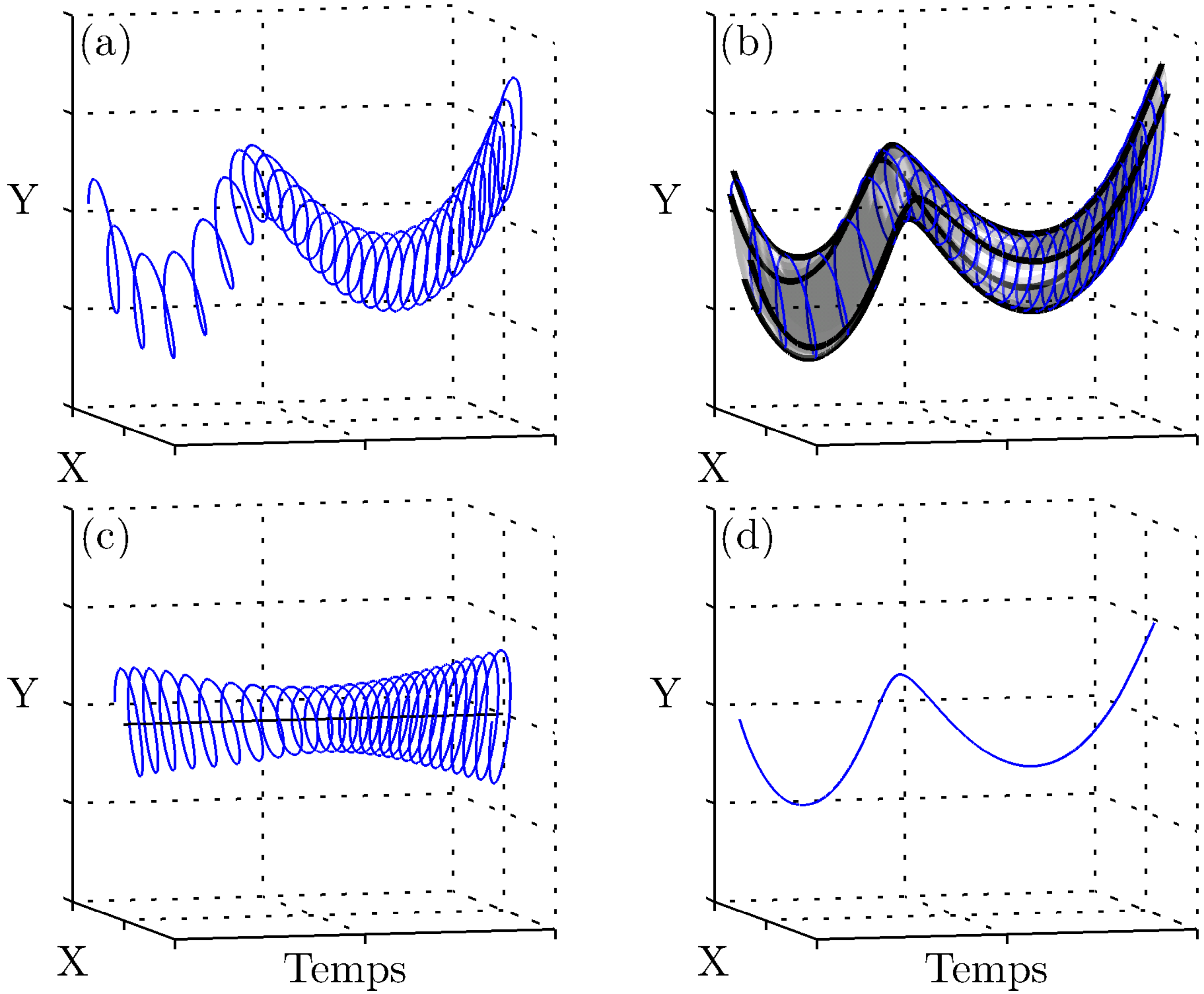
« Mode un-mixing »



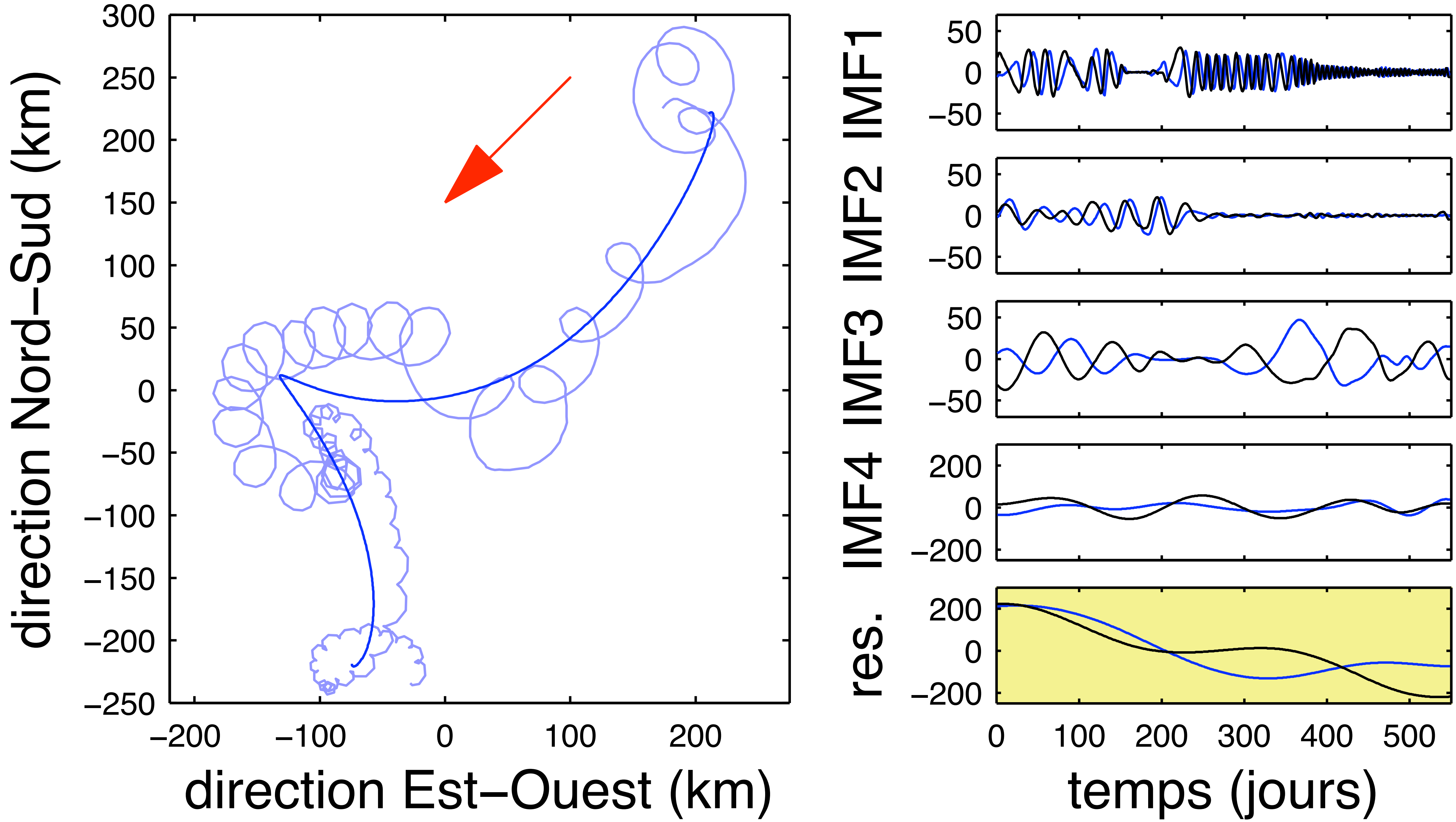
Multivariate EMD

- **Purpose**: deal in some **joint** and **coherent** way with multiple components rather than component-wise
- **Rationale**: consider « oscillations » in higher dimensions
- **Implementation**
  - ***Bivariate*** *setting [Rilling et al., '07]*
    1. *switch from oscillations to* ***rotations***
    2. *replace envelopes by* ***tubes***
    3. ***project*** *and apply the* ***usual EMD machinery***
  - ***Multivariate*** *setting [Rehman & Mandic, '10]*

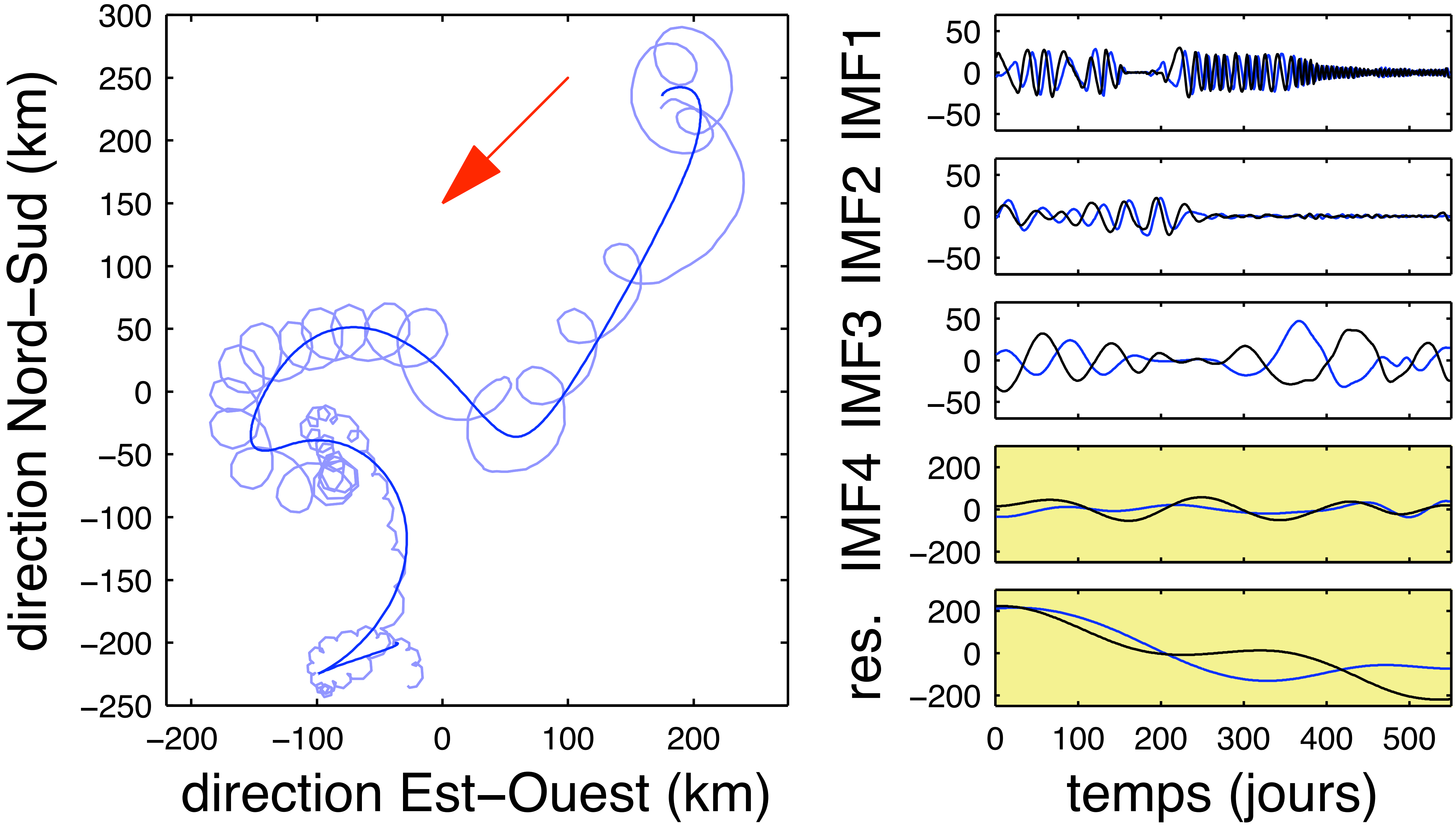
Bivariate EMD – Principle



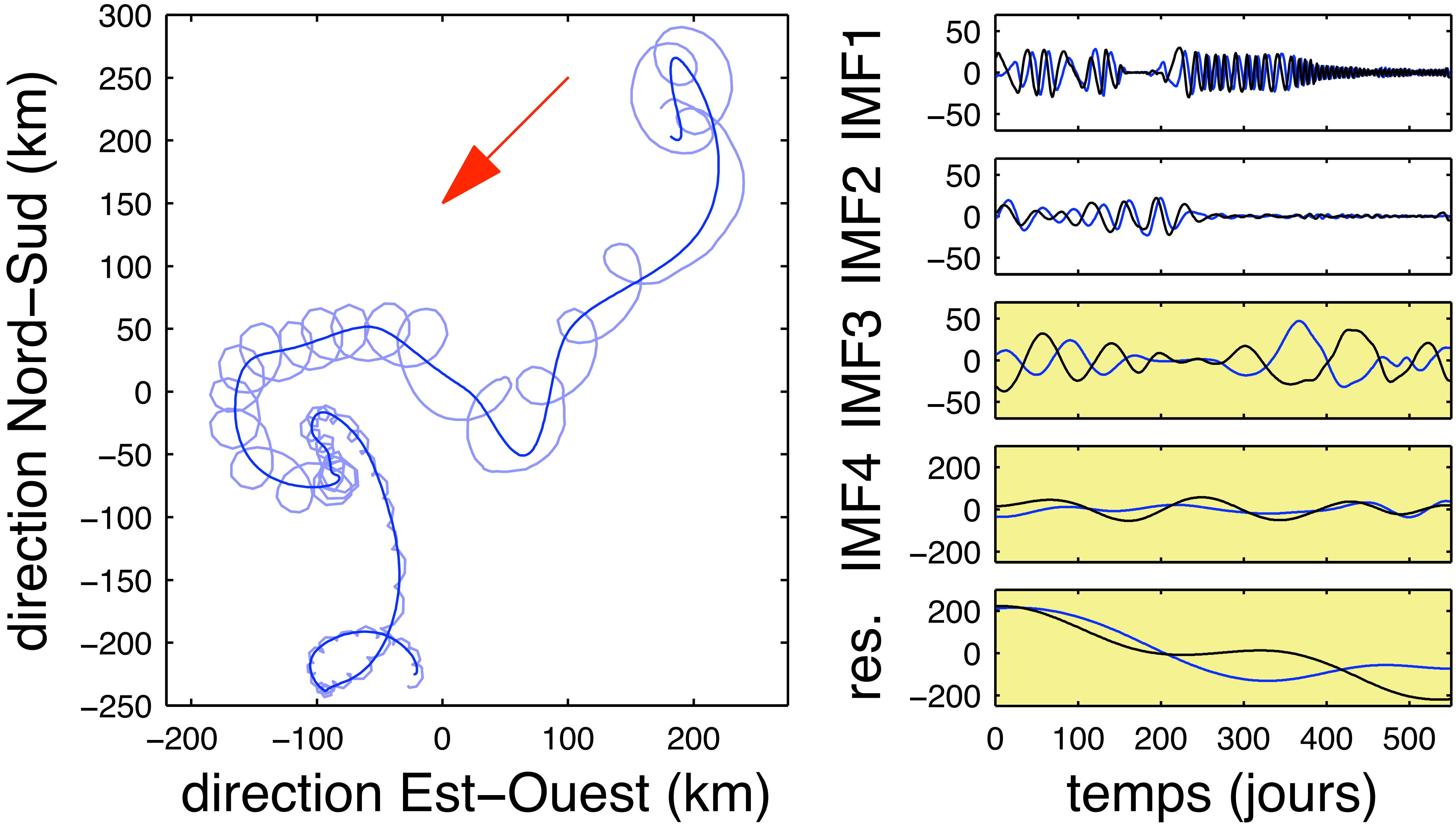
Bivariate EMD – Example



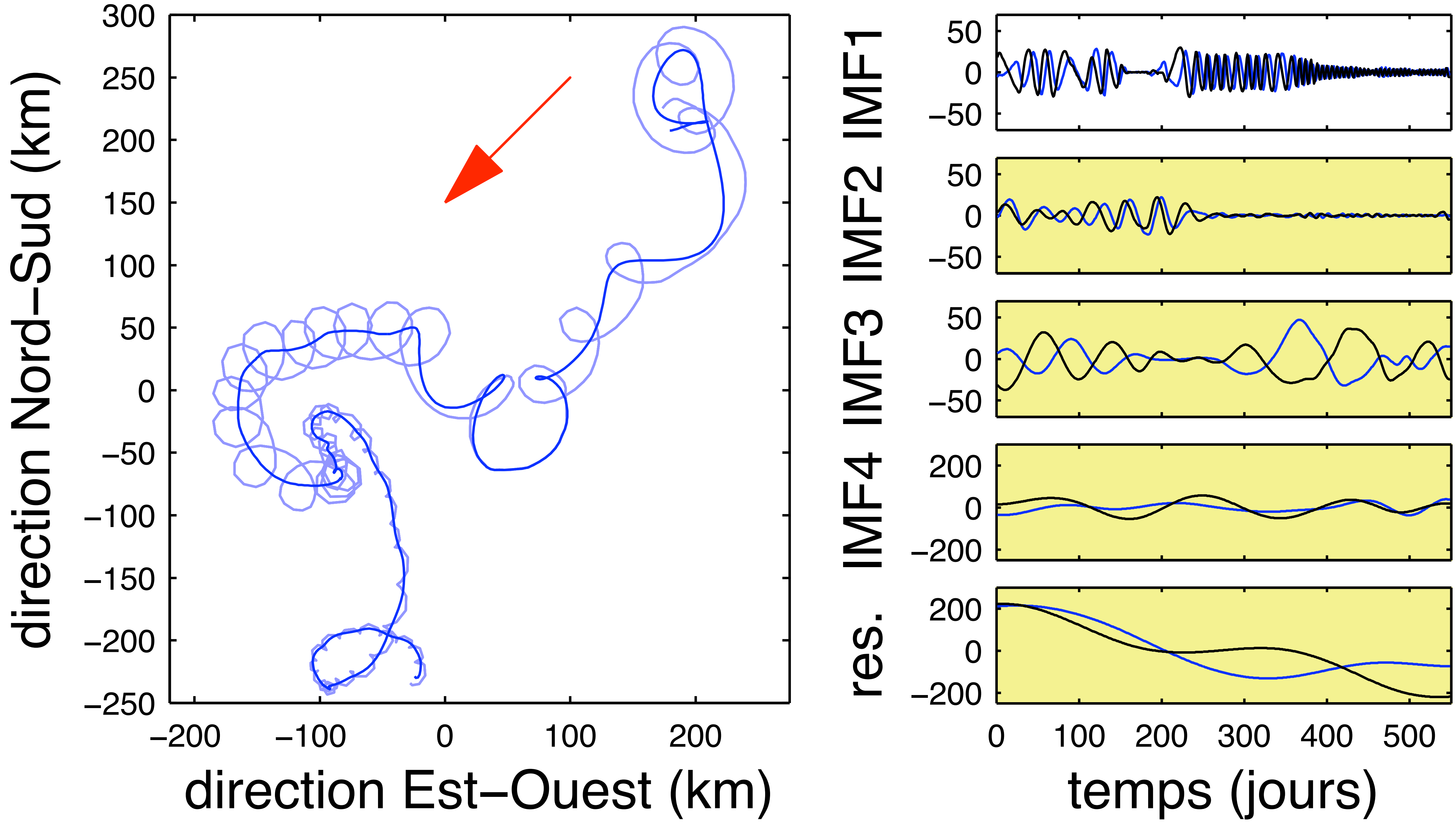
Bivariate EMD – Example



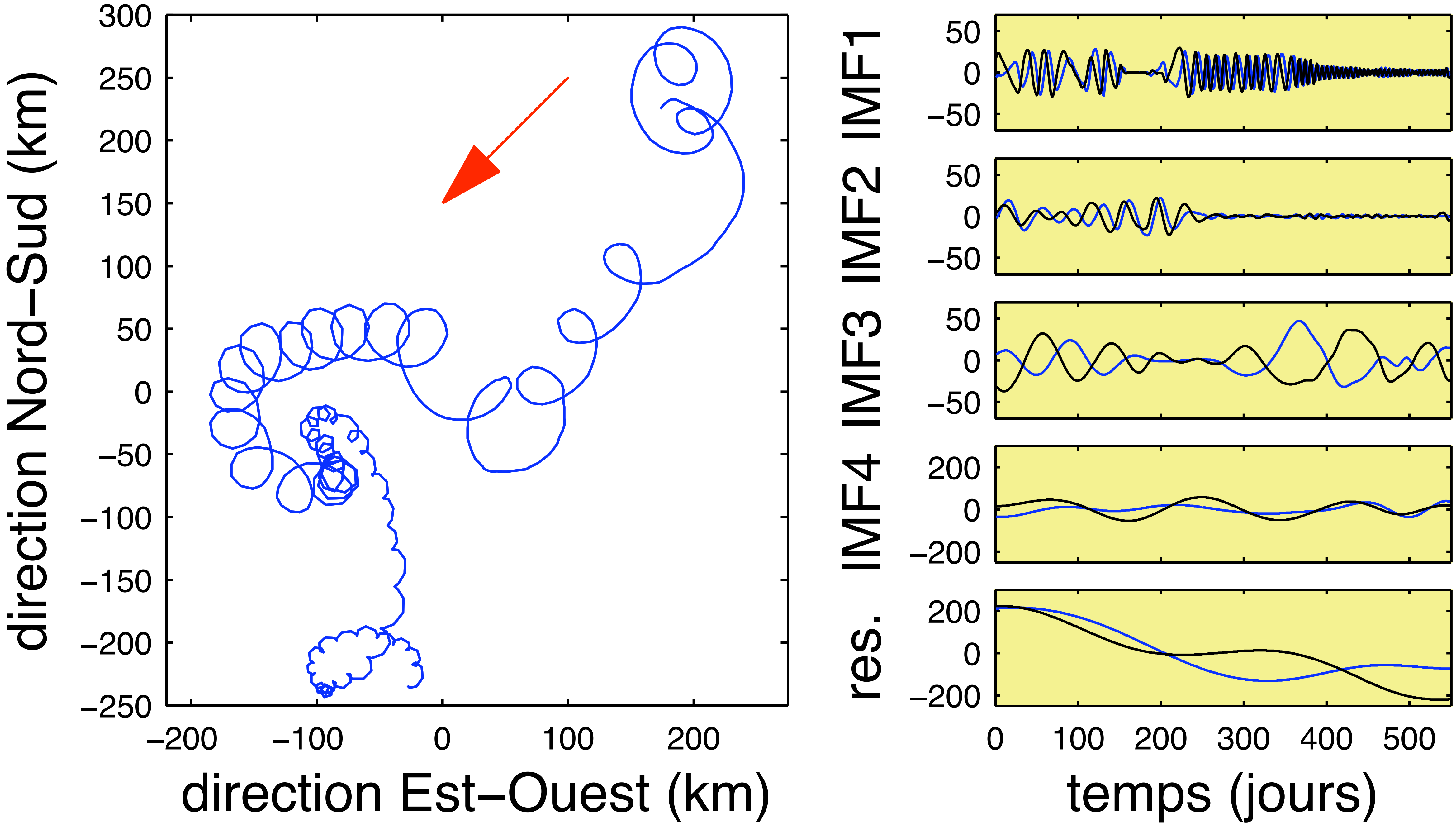
Bivariate EMD – Example



Bivariate EMD – Example



Bivariate EMD – Example



Optimization-based EMD

- **Rationale**
  - ***formalize*** *the distinction between IMF and residual*
  - *replace sifting by* ***optimized*** *criteria*
- **Different approaches**
  - *model-based [Hou & Shi, '11-'13]*
  - *model free [Oberlin et al., '12][Pustelnik et al., '12-'14]*
- **Extension to images** [Schmitt et al., '14]

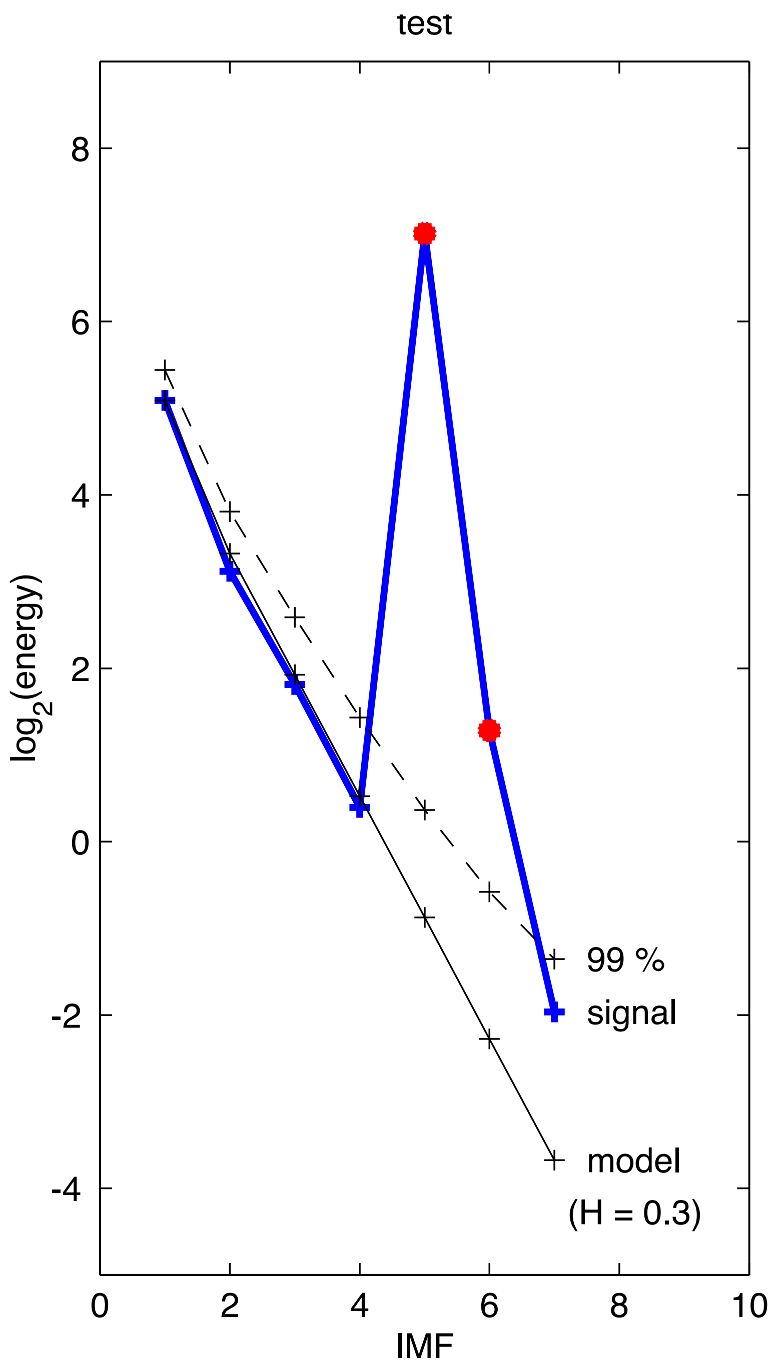
IMF selection

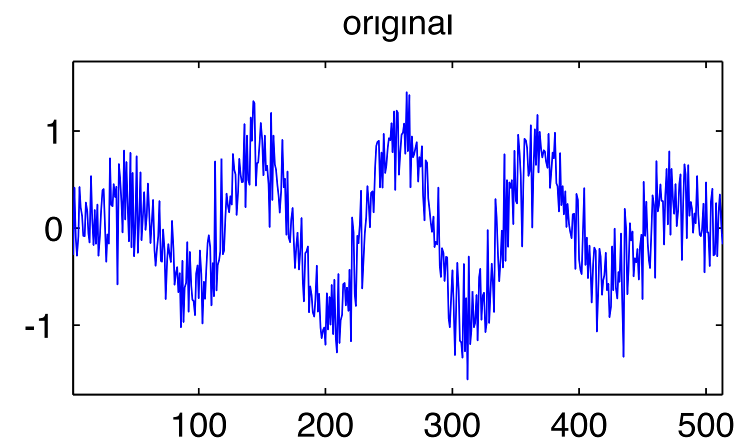
$$x(t) = \sum_{k \in \mathcal{K}} d_k(t) + \sum_{k \notin \mathcal{K}} d_k(t) + a_K(t)$$

with $\mathcal{K}$ some subset of the K IMFs

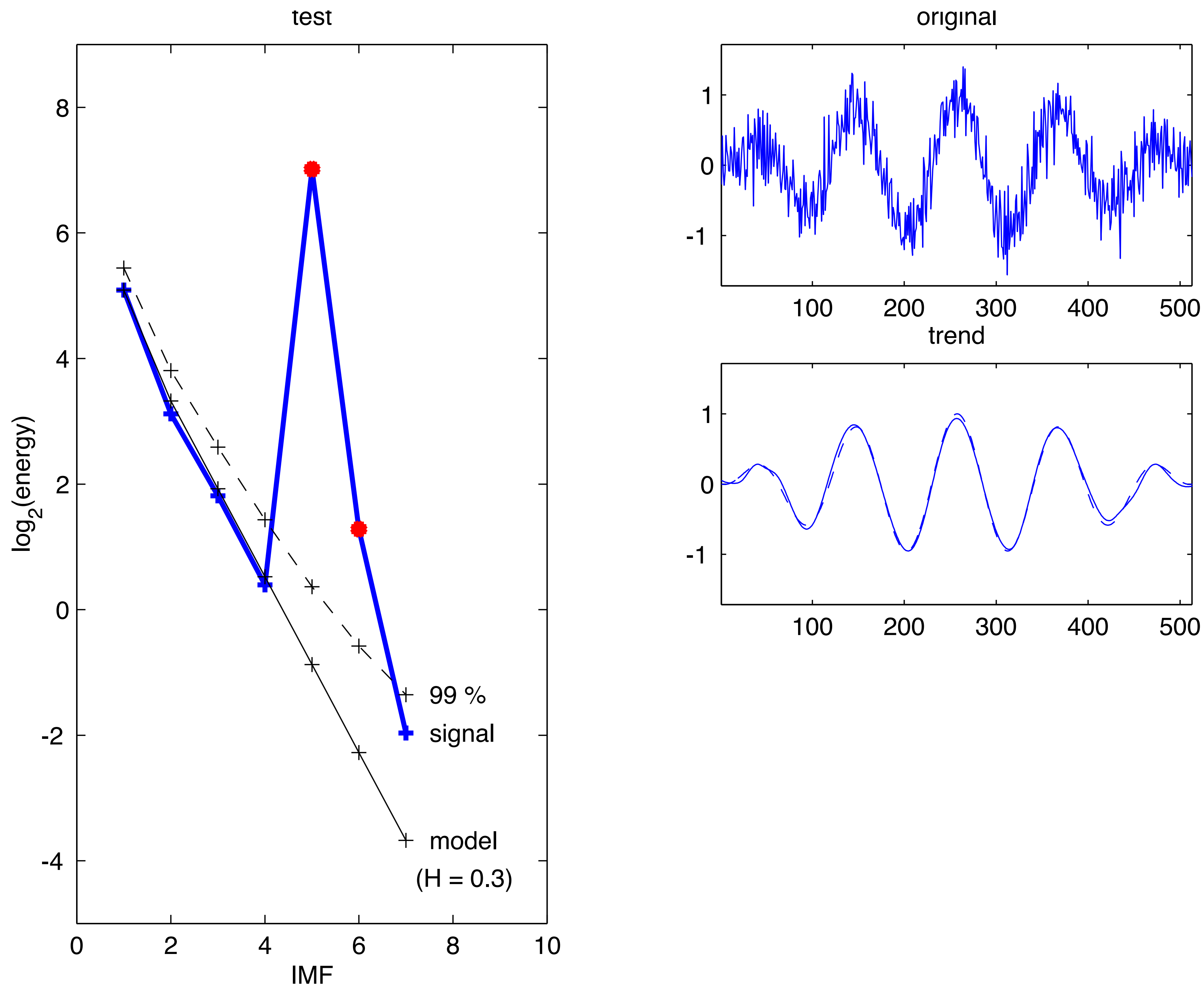
- Different situations:
  - $\mathcal{K}$ = set of distinct, non necessarily contiguous k's : mode **selection/removal**
  - $\mathcal{K} = \{1, \ldots k^*\}$: **denoising/detrending**

IMF selection/removal

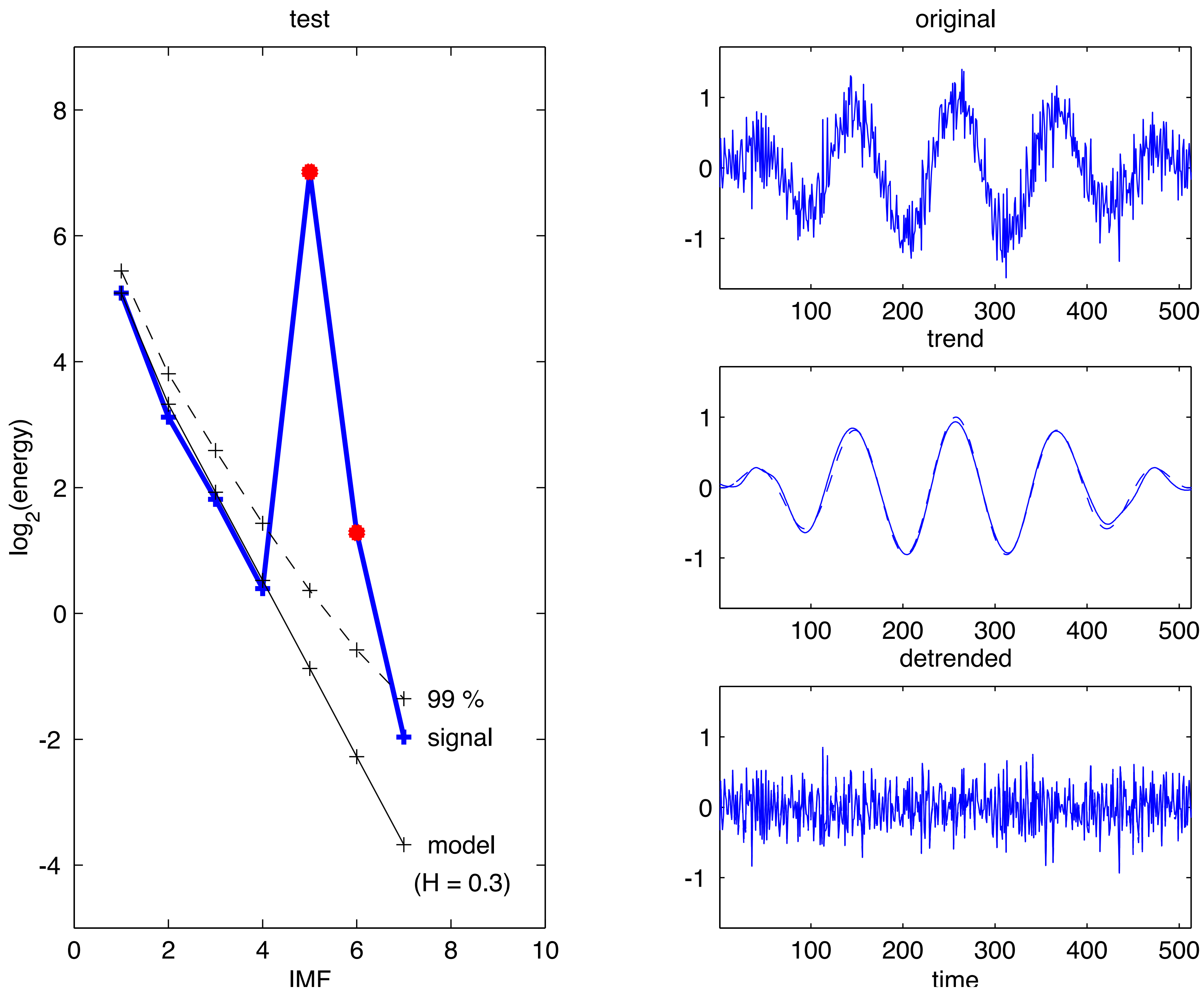




IMF selection/removal



IMF selection/removal

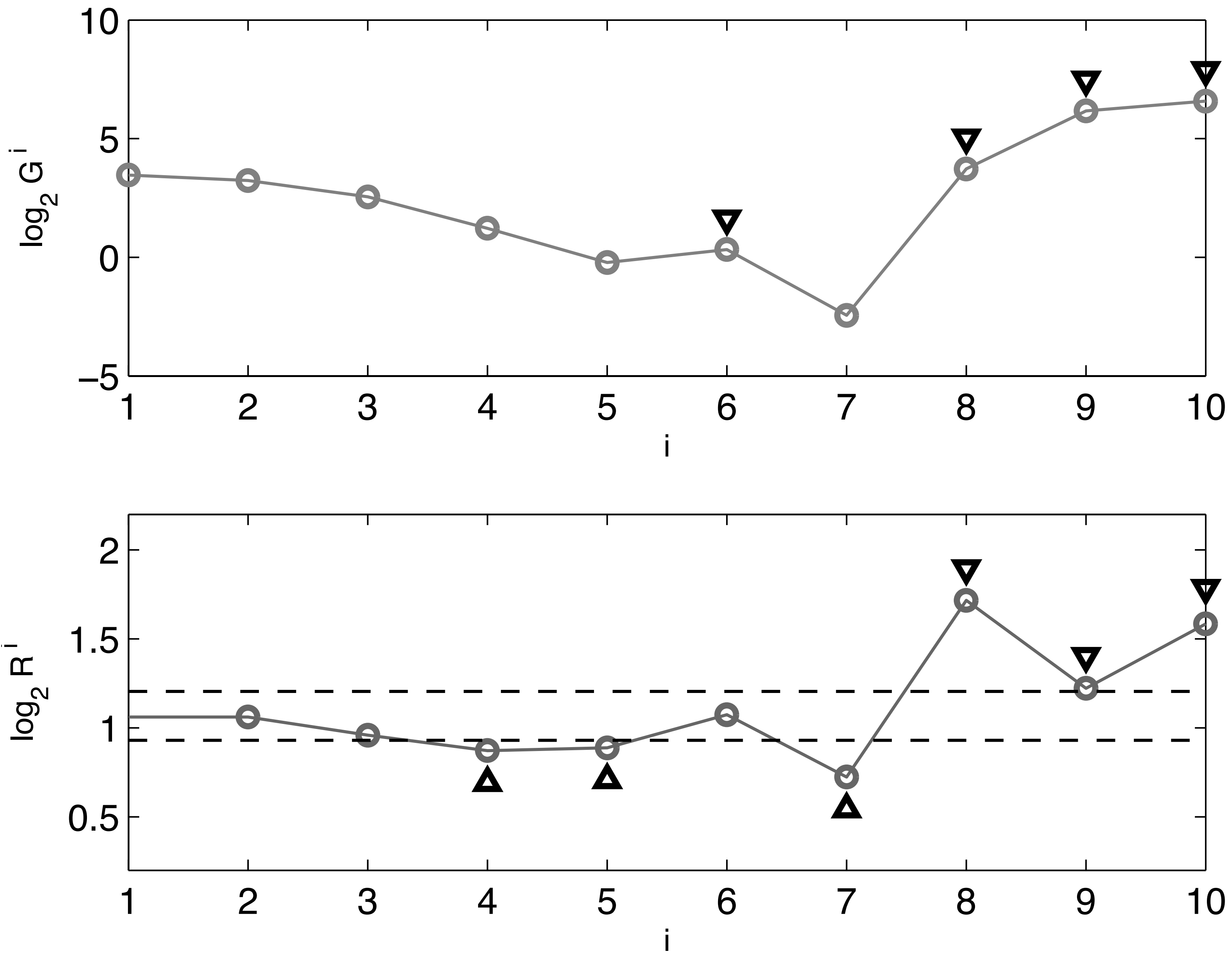


Denoising/detrending

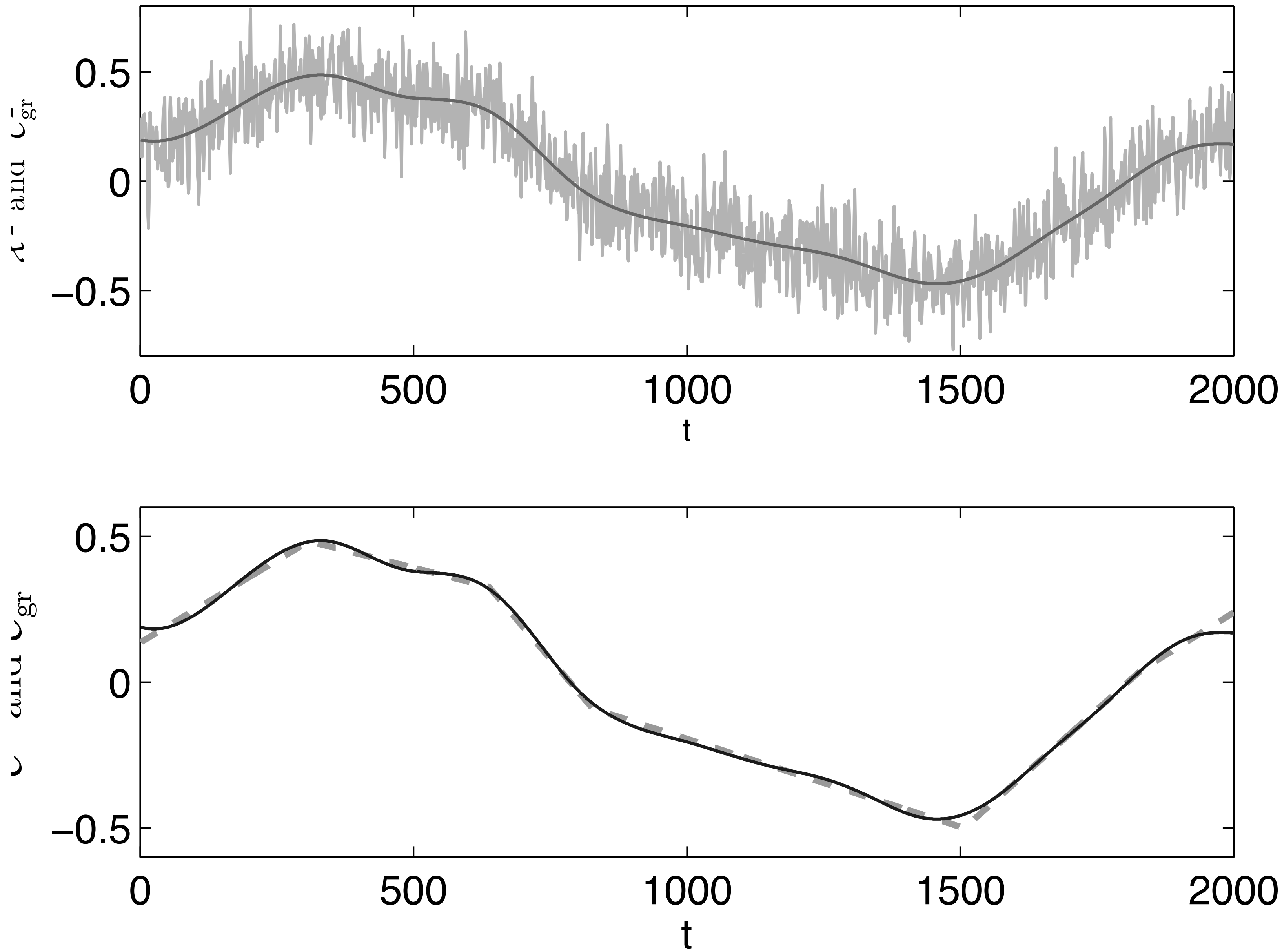
- **Definition**: A « **trend** » is a loosely defined object, e.g., a « long-term change in the mean » [Chatfield, '96]
- As opposed to « **fluctuations** », an EMD-based definition of a trend may correspond to (some of) the **last IMF(s)**
- A possible **strategy** [Moghtaderi *et al.*, '11] for selecting those relevant IMFs combine **ratios** of
  - *zero-crossings*
  - *energy*

  between successive modes

Denoising/detrending



Denoising/detrending



Some concluding remarks

- **Nonstationarities** (+ nonlinearities): adapted **time-frequency** methods
- **Variability** + model-free: **data-driven** techniques
- 3 distinct features to consider
  - ***mathematical*** *setting*
  - ***physical*** *interpretation*
  - ***algorithmic*** *efficiency*
- **Context-driven** selection of methods

More

- **Two recent references**
  - F. Auger *et al.*, « An overview of time-frequency reassignment and synchrosqueezing, » *IEEE Signal Proc. Mag.*, 30(6):32-41, 2013
  - N.E. Huang and S.P. Shen (*eds*.), *Hilbert-Huang Transform and Its Applications: 2nd Edition*, World Scientific, 2014
- **(P)reprints & freewares**

`http://perso.ens-lyon.fr/patrick.flandrin`