Exploring Glitchland - II

T-F Glitch Entomology

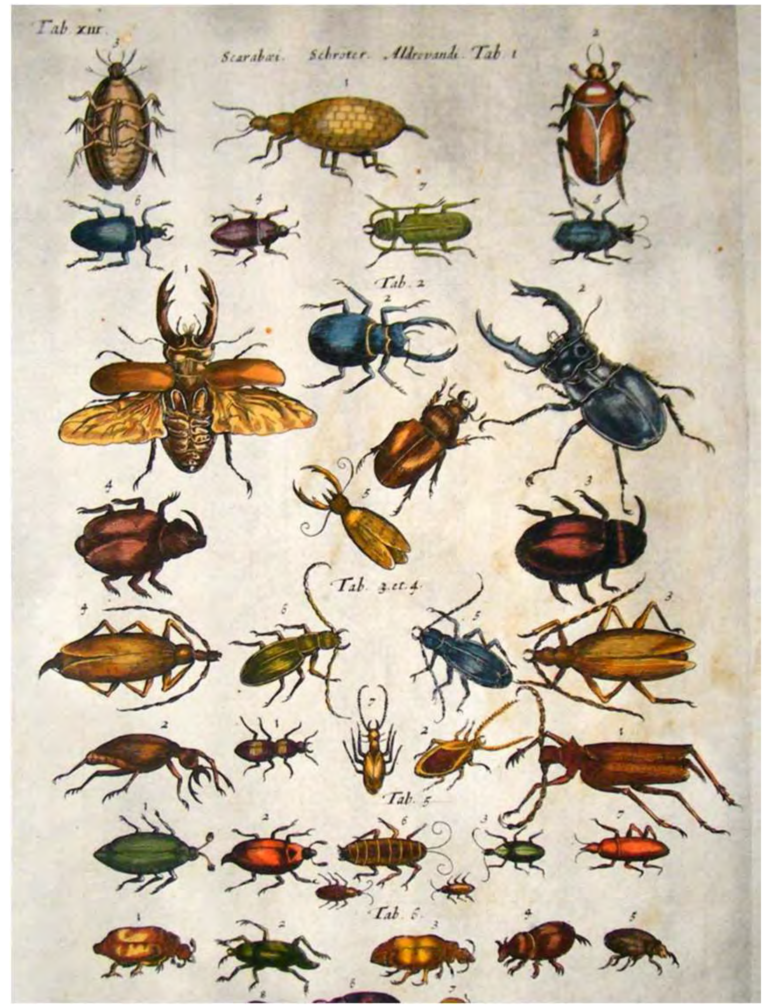
Innocenzo M. Pinto

University of Sannio at Benevento and INFN

Institute of Cosmic Ray Research, University of Tokyo, May 23, 2014

Entomology

The study and under-standing (Greek: λογία) of the insects (Greek: ἔντομοι), i.e., literally "entities which consist of (several) parts"



Glitch Entomologist Roadmap

- Validate a quick reliable procedure for extracting training sets of "clean" glitches from (long) time series . **Key problem: robustness** .

- Investigate "fine structure" of glitches (look for evidences of the fact that, in general, glitches are "multi-component") ➡ **skeleton representations .**

- Investigate glitch (linear) decompositions using a *pre-determined* set of elementary "glitches". Simplest choice: SG-atoms ➡ **Matching - Pursuit representation of glitches using a Gabor-atom dictionary** .

- Investigate glitch (linear) decompositions using an **adaptively - determined** set of elementary glitches ➡ **Basis - Pursuit, K-means-SVD dictionary construction**

This Talk

Time Frequency Tools

Resolution duality

The Wigner Ville Transform

Getting Rid of WV Artifacts

Compressed Coding and Skeletons

CCS vs QT

Atomic Decompositions

The “Dark Energy” problem

Hilbert-Huang Transform

Specific Issues

Time-Frequency Tools

Natural choice for analyzing non-stationary (transient) signals;

All limited by Gabor *resolution bound* (frequency · time uncert-ainty) - to different extents, e.g.,

Short-Time Fourier Transform (*worst* of them all in terms of TF resolution);

Non-uniform TF-tiling linear transforms, including wavelet (e.g., Kleine-Welle, L. Blackburn, LIGO-T060221 etc), and constant-Q (e.g., Q-pipeline, S. Chatterji, LIGO-G060044 etc) transforms;

Wigner-Ville transform, *best* (in energy-preserving transform class) in terms of resolution ...
... but, being bi-linear, exhibiting signal-signal and signal-noise *intermodulation artifacts* which may affect its readability ...

Gabor Atomic Decomposition;

Hilbert – Huang "Empirical Mode" Decomposition

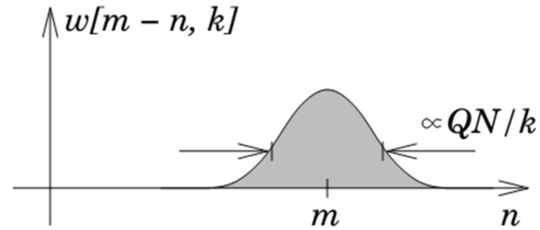
The Q-Transform

Project signal into time-shifted, time-windowed-sinusoids , whose time-widths are *inversely proportional* to their center frequencies.

frequency

$$X_Q[m, k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nk/N} w[m-n, k]$$

time



Efficient computation in terms of D(W)FT,

$$\tilde{X}[l] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi nl/N}, \quad \tilde{W}[l] = \sum_{n=0}^{N-1} w[n, k] e^{-i2\pi nl/N}$$

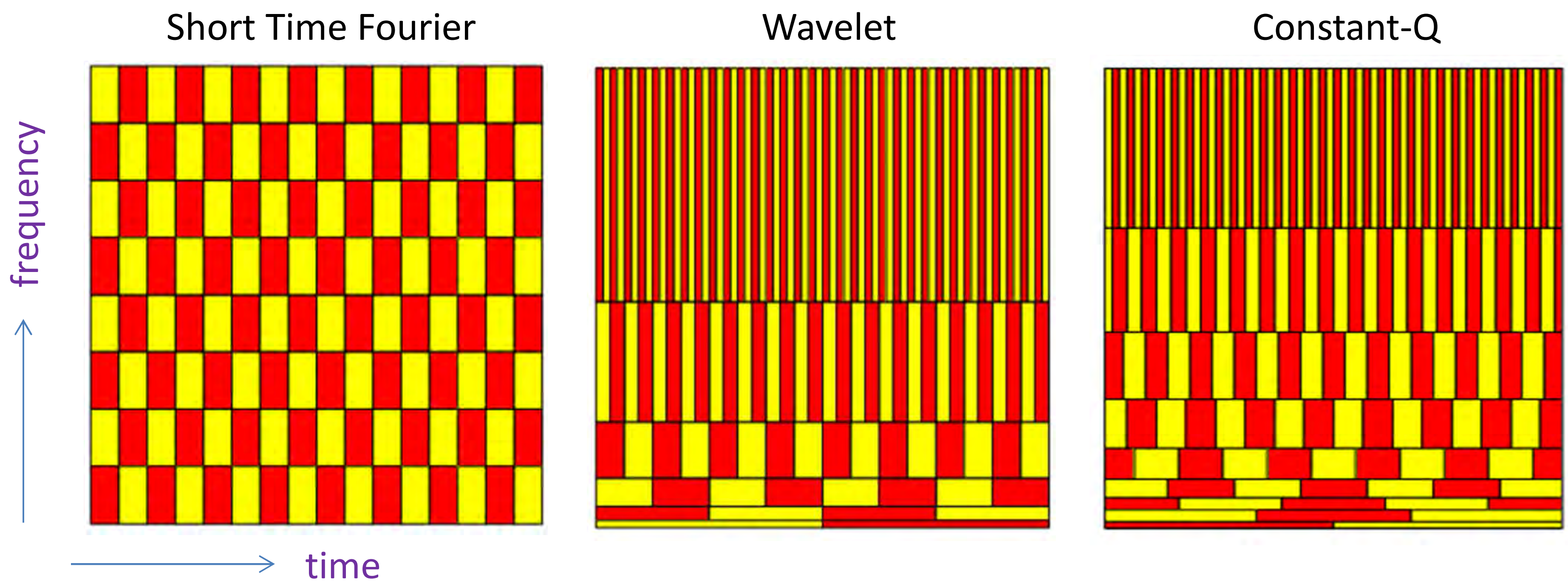
$$\longrightarrow \quad X_Q[m, k] = \sum_{l=0}^{N-1} \tilde{X}[l+k] \tilde{W}[l, k] e^{-i2\pi ml/N}$$

[J.C. Brown, "Calculation of a constant Q spectral transform," JASA 89 (1991) 425]

[J. C. Brown and M. S. Puckette, "An Efficient Algorithm for the Calculation of a Constant Q Transform," JASA 92 (1992) 2698.]

[S. Chatterji et al., "Multiresolution Techniques for the Detection of Gravitational - Wave Bursts," CQG 21 (2004) S1809]

TF Plane Tilings



QT can be regarded as template bank (in the whitened signal manifold), whose templates are sinusoidal-Gaussians of different central time, central frequency, and quality factor Q. Typical Q values are in the range (4, 64); typical center frequencies from 10^2 to 10^3 Hz.

QT stands at the core of the (widely adopted) **Qscan**, **QOnline** and **Omega** pipelines.

[R. Khan and S. Chatterji, "Enhancing the Capabilities of LIGO Time-Frequency Plane Searches Through Clustering," CQG 26 (2009) 155009]

This Talk

Time Frequency Tools

Resolution duality

The Wigner Ville Transform

Getting Rid of WV Artifacts

Compressed Coding and Skeletons

CCS vs QT

Atomic Decompositions

The “Dark Energy” problem

Hilbert-Huang Transform

Specific Issues

The Wigner-Ville Transform

$$W_x(t,f) = \int_{-\infty}^{+\infty} \tilde{x}\left(t+\frac{\tau}{2}\right)\tilde{x}^*\left(t-\frac{\tau}{2}\right)e^{-i2\pi f\tau}d\tau$$

where: $\tilde{x}(t) = x(t) + i\,\mathscr{H}[x(t)]$

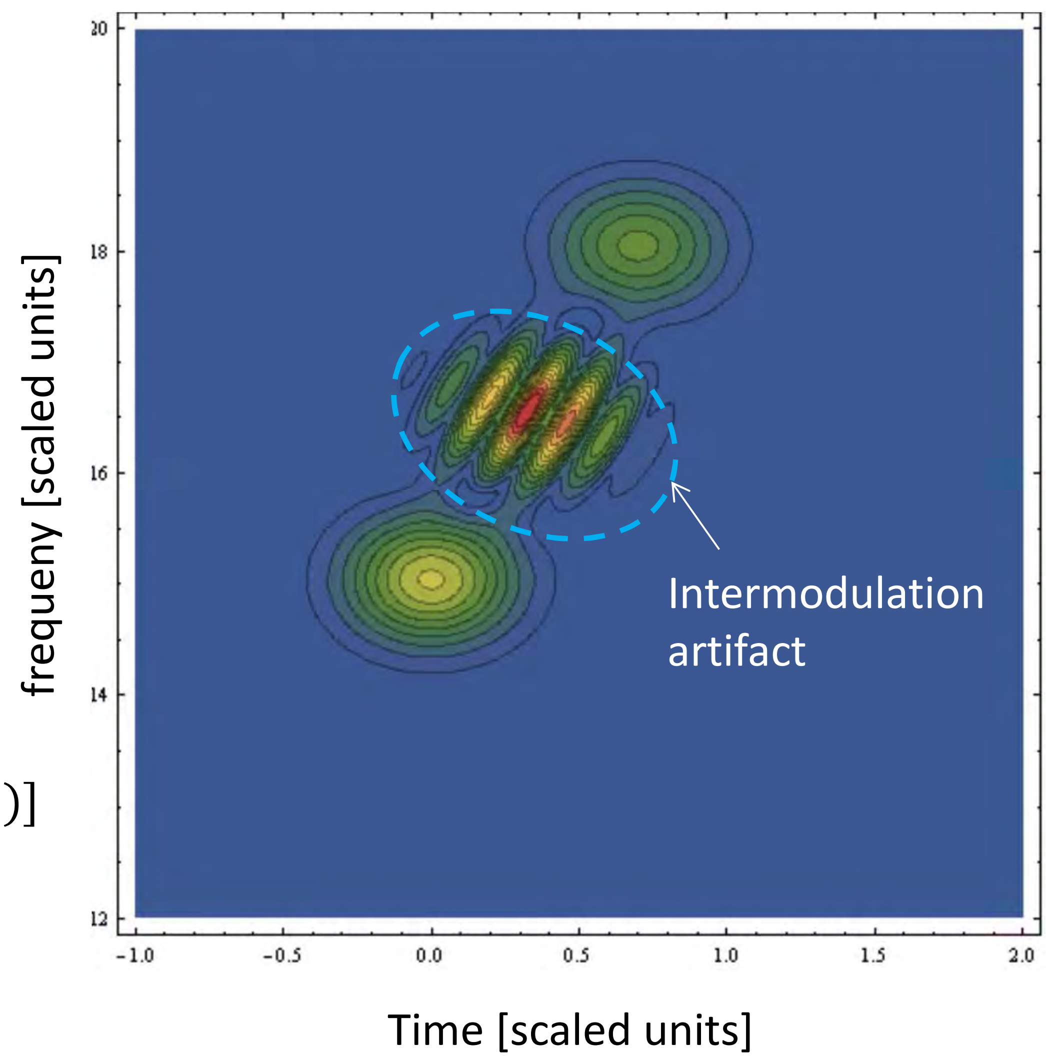
analytic mate of $x(t)$ **Hilbert transform**

WV intermodulation artifacts ***do not*** appear ***only*** for two special signals:

- *linear* chirps, $f = f_0 + \beta t$;
- Gabor (SG) atoms

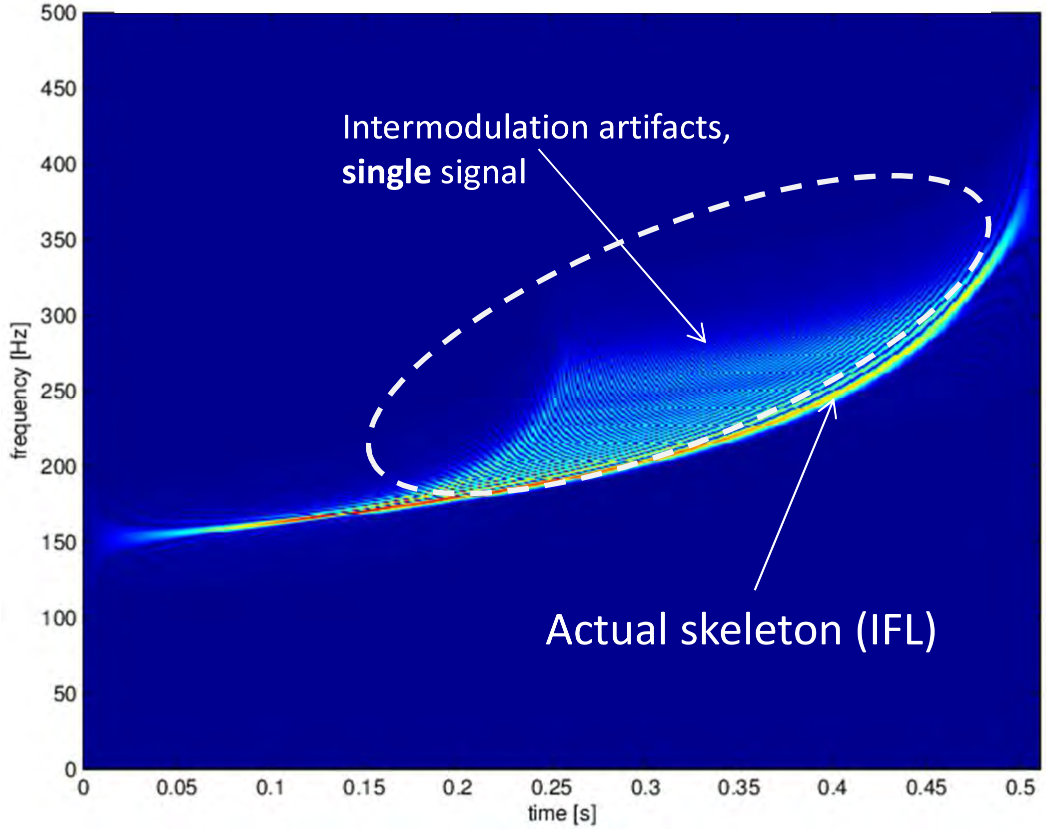
$$s(t) = C\,exp\left[\left(\frac{t-t_0}{\delta t}\right)^2\right]cos[\Omega(t-t_0)]$$

WV of 2-atom Gabor molecule



The Wigner-Ville Transform, contd.

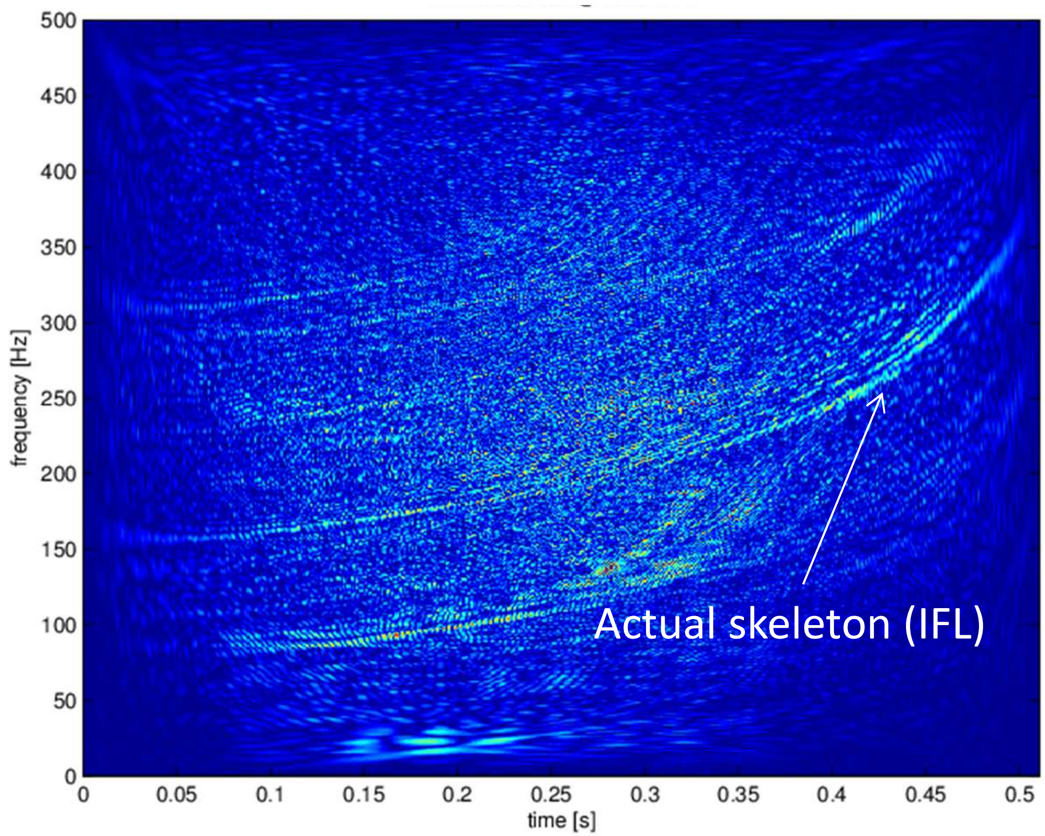
WV of 0-PN Chirp, $f = f(0)(1 - \frac{t}{t_c})^{-3/8}$



Think of actual instantaneous frequency as a superposition (chain) of time-gated (short) consecutive linear chirps. ..

The Wigner-Ville Transform, contd.

WV of 0-PN Chirp, $f = f(0)(1 - \frac{t}{t_c})^{-3/8}$ in AWGN



Intermodulation between signal and noise may badly blur the TF representation .
In this specific case, the actual skeleton is barely visible …

Getting Rid of WV Artifacts

Smoothing by use of an ambiguity-function (AF) adapted kernel [R. Baraniuk and D. Jones, Signal Proc. 32 (1993) 263] - Husimi / Choi-Williams can be seen as special cases; **Suppress intermodulation terms (which map far from the origin in the AF plane) *at the expense of a worse TF resolution***

Re-assigment (re-squeezing) [P. Flandrin et al., IEEE T-SP 43 (1995) 1968] **re-allocate value of smoothed WV at each TF point P to *barycenter* of the WV smoothed by kernel at P**

Retrieving the TF skeleton of the WV from a *reduced cardinality* AF-data subset, using a ***sparsity constraint*** [P. Flandrin and P. Borgnat, IEEE T-SP58 (2010) 2974]

This Talk

Time Frequency Tools

Resolution duality

The Wigner Ville Transform

Getting Rid of WV Artifacts

Compressed Coding and Skeletons

CCS vs QT

Atomic Decompositions

The “Dark Energy” problem

Hilbert-Huang Transform

Specific Issues

CC Skeleton - Algorithm

Compute Wigner-Ville TF-representation of data *W(t,f)* ;

Compute related ambiguity function (2D Fourier transform of WV) , call it *A(t,f)* ;

Construct Beraniuk-Jones optimal smoothing kernel for *A(t,f)* ;

Identify Beraniuk-Jones optimal smoothing kernel level curves;

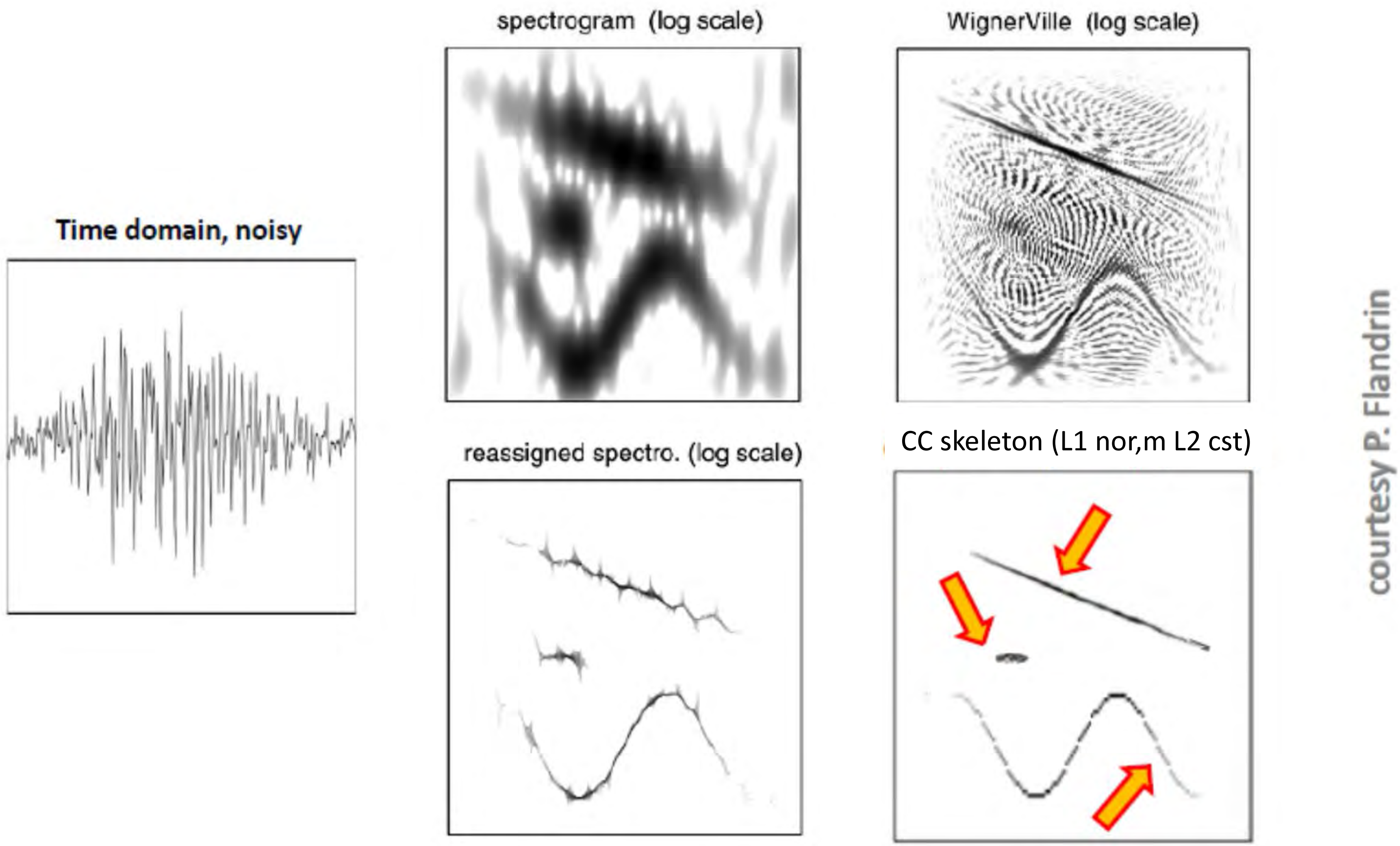
Select the contour encircling $\sim N$ samples (*Heisenberg cardinality constraint*);

Find the TF skeleton $\Sigma(t,f)$ by solving the constrained optimization problem (sparse synthesis problem)

$$\min_{\Sigma} \|\Sigma\|_0 \text{ subject to } \|\mathfrak{F}(\Sigma) - A\|_2 \leq \varepsilon$$

Tuning: N and ε key params

CC Skeletons – Seeing is Believing



CC Skeletons for GW Hunters

Compressed Coding for Time-Frequency Gravitational Wave Data Analysis

Paolo Addesso[1], Maurizio Longo[1], Stefano Marano [1], Vincenzo Matta[1], Maria Principe[2] and Innocenzo M Pinto[2]

[1] Dept. of Electrical and Computer Engineering and Applied Mathematics, University of Salerno, Italy, [2]WavesGroup, University of Sannio at Benevento, Italy, INFN and LVC

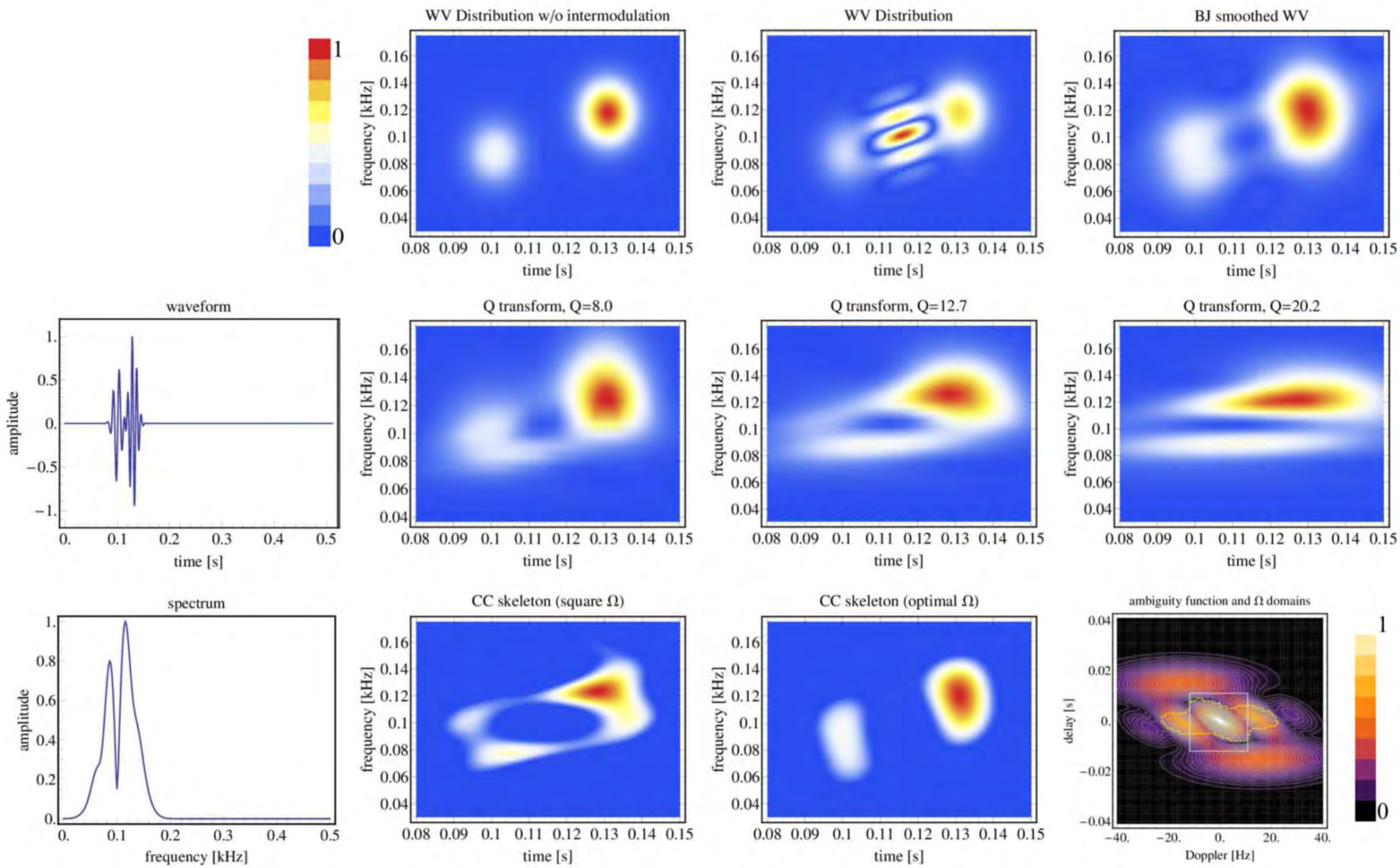
E-mail: pinto@sa.infn.it

PACS number: 04.80.Nn, 07.05.Kf, 95.55.Sz

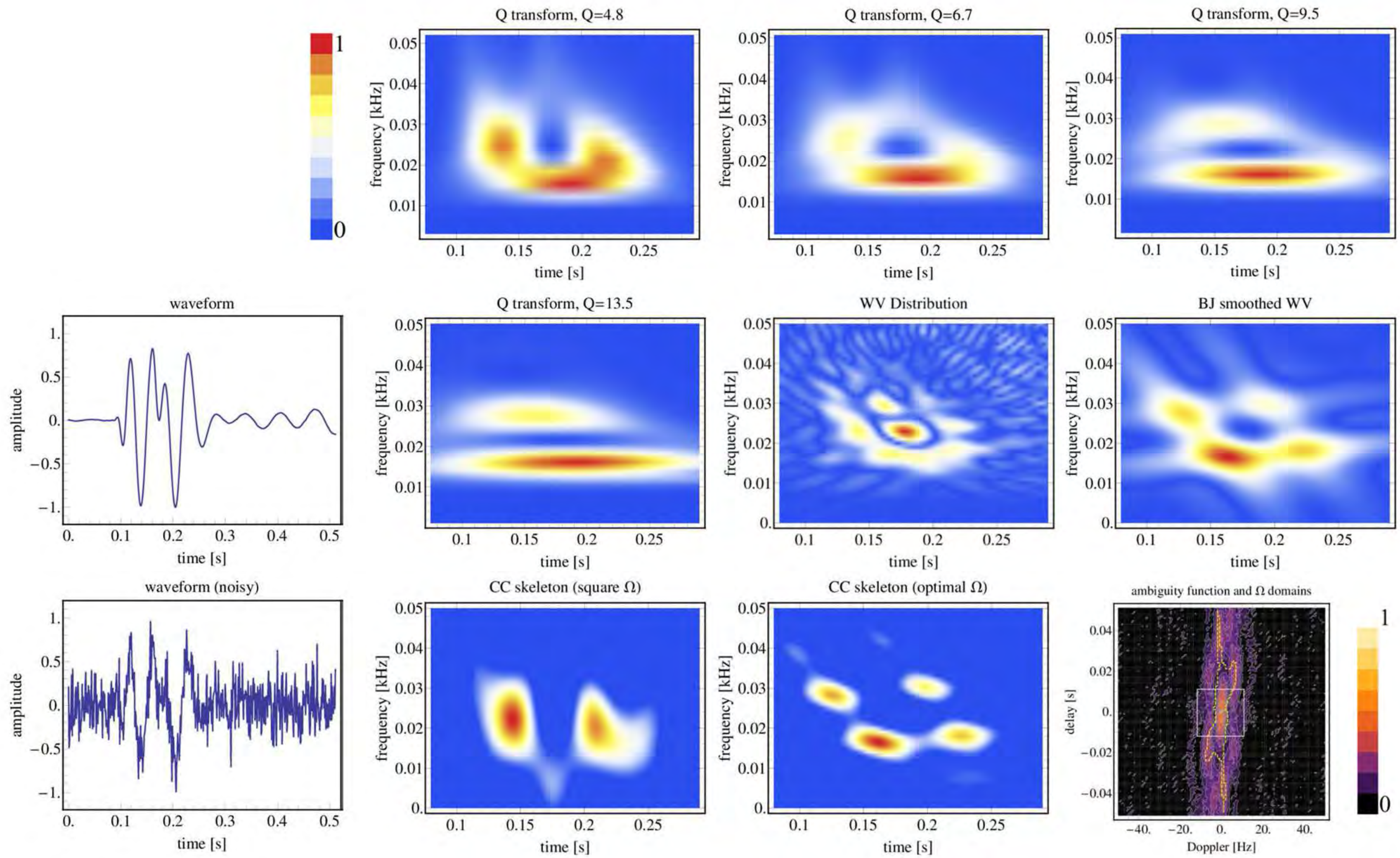
Abstract. The potential of sparse time frequency representations (compressed coding) for gravitational wave data analysis is illustrated, by comparison with existing methods, as regards i) shedding light on the fine-structure of noise glitches, in preparation of their classification, and ii) boosting the performance of waveform consistency tests in the network detection of unmodeled transient gravitational wave signals co-existing with unmodeled noise glitches.

[LIGO-P1200170]

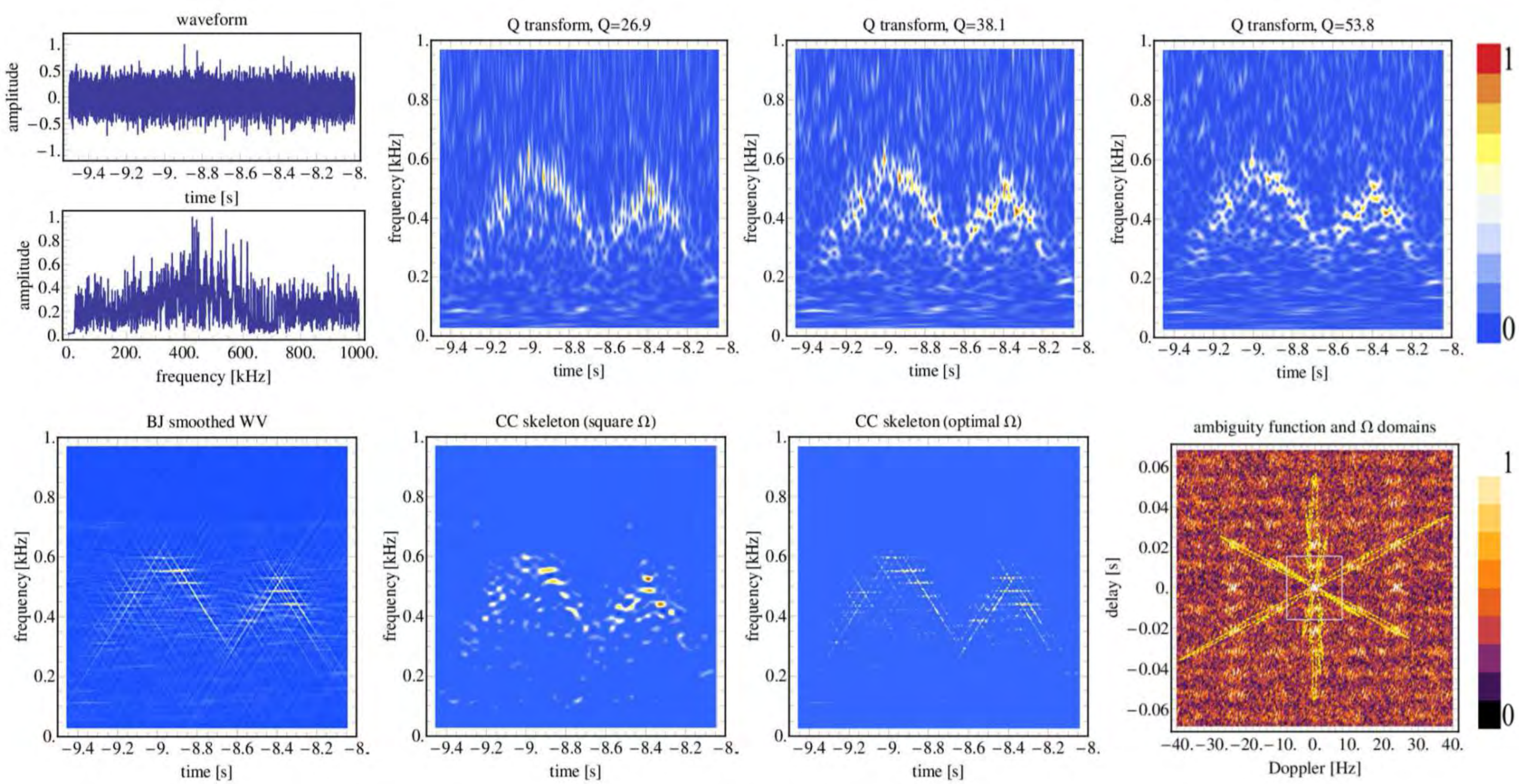
A Gabor «Molecule»



A LIGO Glitch



A GEO 600 Glitch



This Talk

Time Frequency Tools

Resolution duality

The Wigner Ville Transform

Getting Rid of WV Artifacts

Compressed Coding and Skeletons

CCS vs QT

Atomic Decompositions

The “Dark Energy” problem

Hilbert-Huang Transform

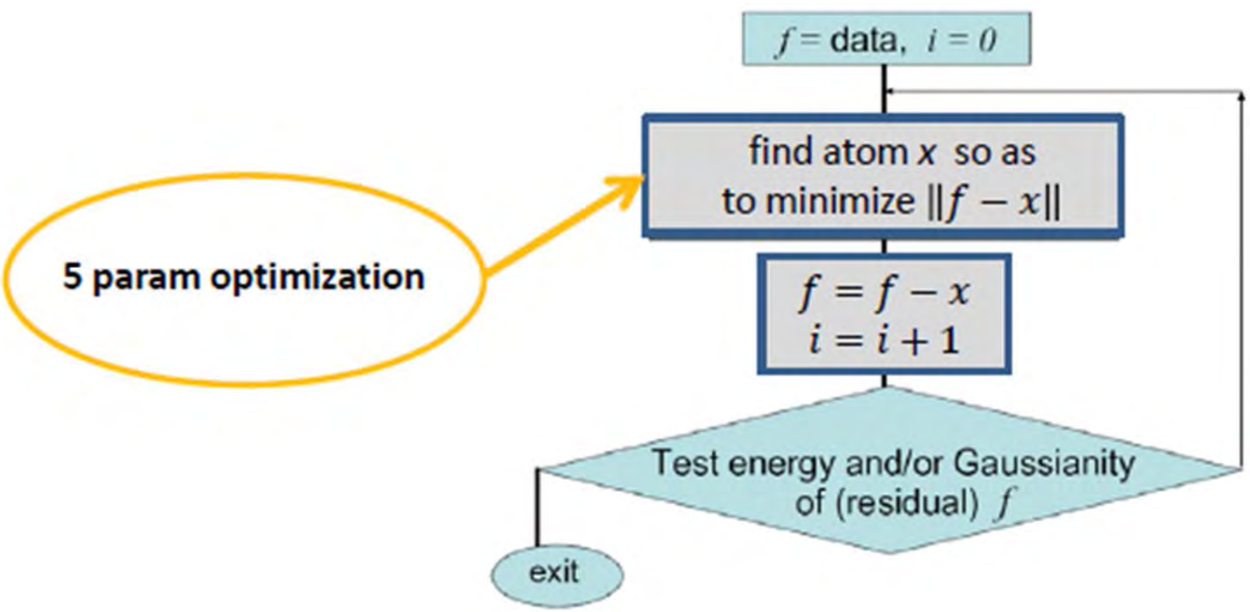
Specific Issues

Atomic Decomposition

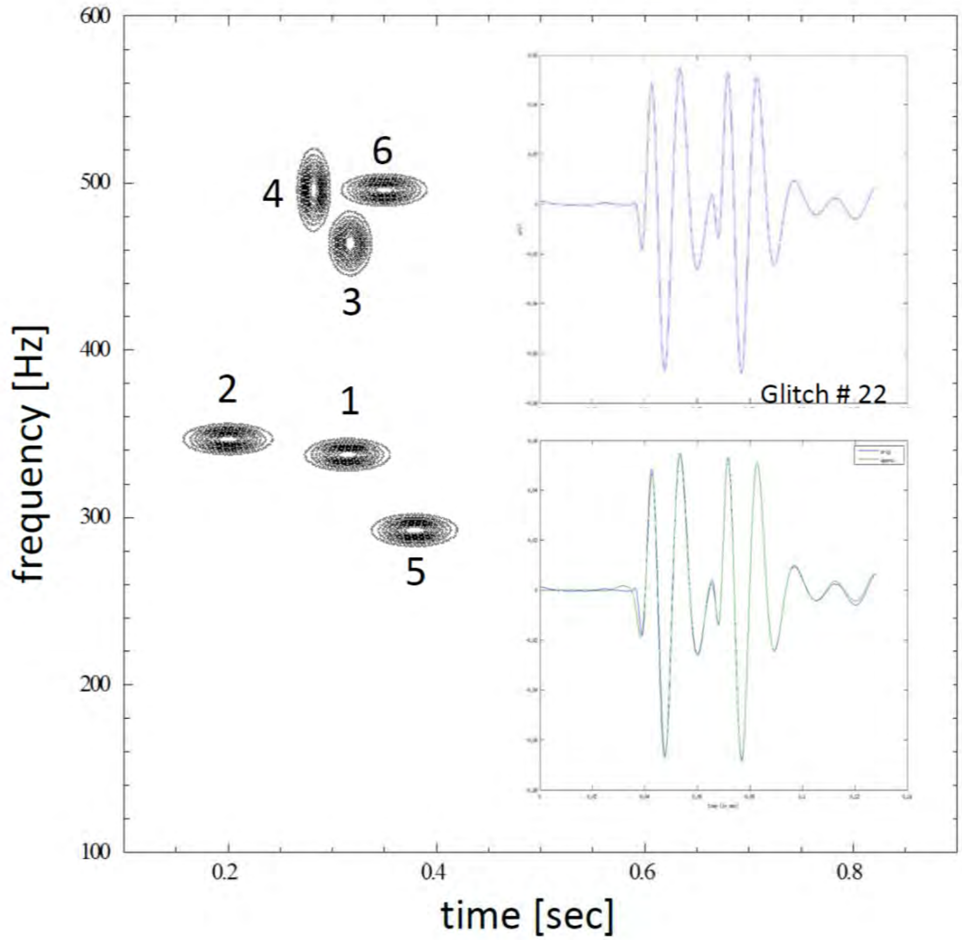
Gabor (Sine Gaussian) atom: 5-parameter waveform **dictionary**

$$x(t \,|\, A_0, t_0, f_0, \tau_0, \psi_0) = A_0 exp\left(-\frac{(t-t_0)}{{\tau_0}^2}\right) Cos(2\pi f_0 t + \psi_0)$$

Matching Pursuit (MP) approximation (*greedy* algorithm)

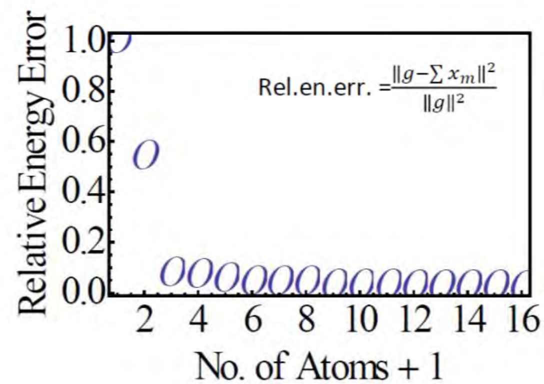


Atomic Decomposition, contd.



m	A_m	t_m	f_m	τ_m	ψ_m	err_{m-1}
1	0.339411	314.525	337.51	36.9824	-0.778158	1.
2	0.339664	199.664	346.794	38.7676	-0.688176	0.536083
3	0.033311	317.117	463.991	18.3848	-0.008061	0.071473
4	0.059282	282.063	495.989	14.1777	-2.50435	0.067004
5	0.055749	378.865	292.419	37.0195	1.49288	0.052851
6	0.010988	349.75	495.989	37.2266	2.08993	0.040335
7	0.031284	225.506	201.535	72.75	-1.54849	0.039849
8	0.052121	236.	495.989	44.0488	-3.75086	0.035908
9	0.010925	286.563	495.989	-7.29688	-1.49696	0.024968
10	0.004872	304.848	319.056	77.9707	-0.390494	0.024487
11	0.029298	507.117	286.594	62.916	7.90958	0.024391
12	0.006489	443.998	463.971	91.6738	5.85081	0.020935
13	0.002232	329.318	208.774	3.45313	-0.084683	0.020765
14	0.001885	336.752	310.541	-4.35156	0.628395	0.020745
15	0.007372	275.125	467.546	4.39648	-2.59703	0.020731
16	0.061547	162.752	463.991	20.0586	-6.72294	0.020512

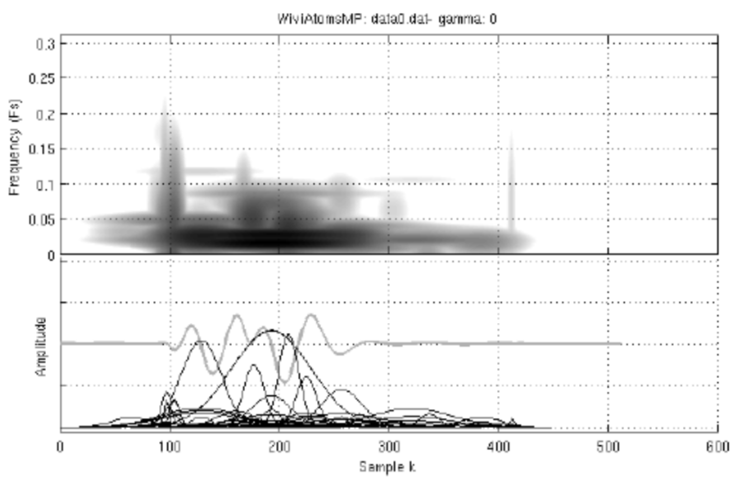
t_m, τ_m [ms]



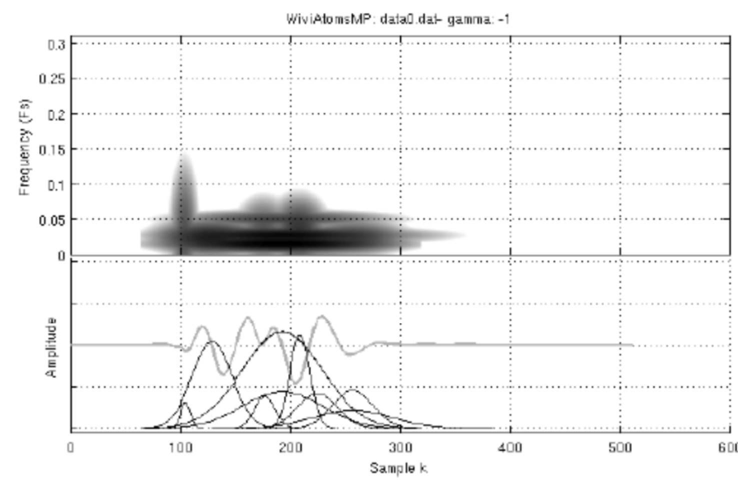
The «Dark Energy» Problem

Naïve atomic decomposition shows atoms which are strikingly outside the time support of the glitch (dark-energy atoms) . Must include a penalty functional in the greedy optimization algorithm

naive



smart



This Talk

Time Frequency Tools

Resolution duality

The Wigner Ville Transform

Getting Rid of WV Artifacts

Compressed Coding and Skeletons

CCS vs QT

Atomic Decompositions

The “Dark Energy” problem

Hilbert-Huang Transform

Specific Issues

The Hilbert Huang Transform

Step 1 - "decompose" signal into Huang's ***intrinsic modes***

$$s(t) = \sum_{k=1}^{N} a_k(t) cos[\phi_k\ (t)]$$

No sound theoretical basis.
Just an algorithm ("sifting").
NASA hold several related patents.

signal has at least one max & one min (otherwise must be differentiated first);

"time scale" defined by time lapse between extrema;

at any point, mean value of the (local) max-envelope and (local) min-envelope is zero for each IMF;

In each IMF the number of min, max and zero crossings does not differ by more than 1.

Step 1 - compute Hilbert – transform ($\mathcal{H}$) of Huang modes ⟶ ***(sparse) TF coding***

$$A_k(t)\exp[i\Phi_k(t)] = a_k(t)cos[\phi_k\ (t)] + i\ \mathcal{H}\ \{ a_k(t)cos[\phi_k\ (t)] \}$$

$$\longrightarrow \quad \omega_k = \frac{d\Phi_k}{dt} \qquad \textit{(instantaneous frequencies)}$$

$$A_k(t) \qquad \textit{(instantaneous amplitudes)}$$

$s(t)$ is "encoded" in the TF plane as a set of TF arcs (ridges), whose levels are determined by the related instantaneous amplitudes

HHT Specific Issues

i) The IMF decomposition is not unique:, there is an infinite number of IMF choices which will reproduce the same s(t), differing in the way the non-stationarity is embedded in the AM, FM or both. To avoid this indetermination we need to constrain the definition of a signal *component* e.g., as a *connected* TF region with *"significant"* energy (Bedrosian).

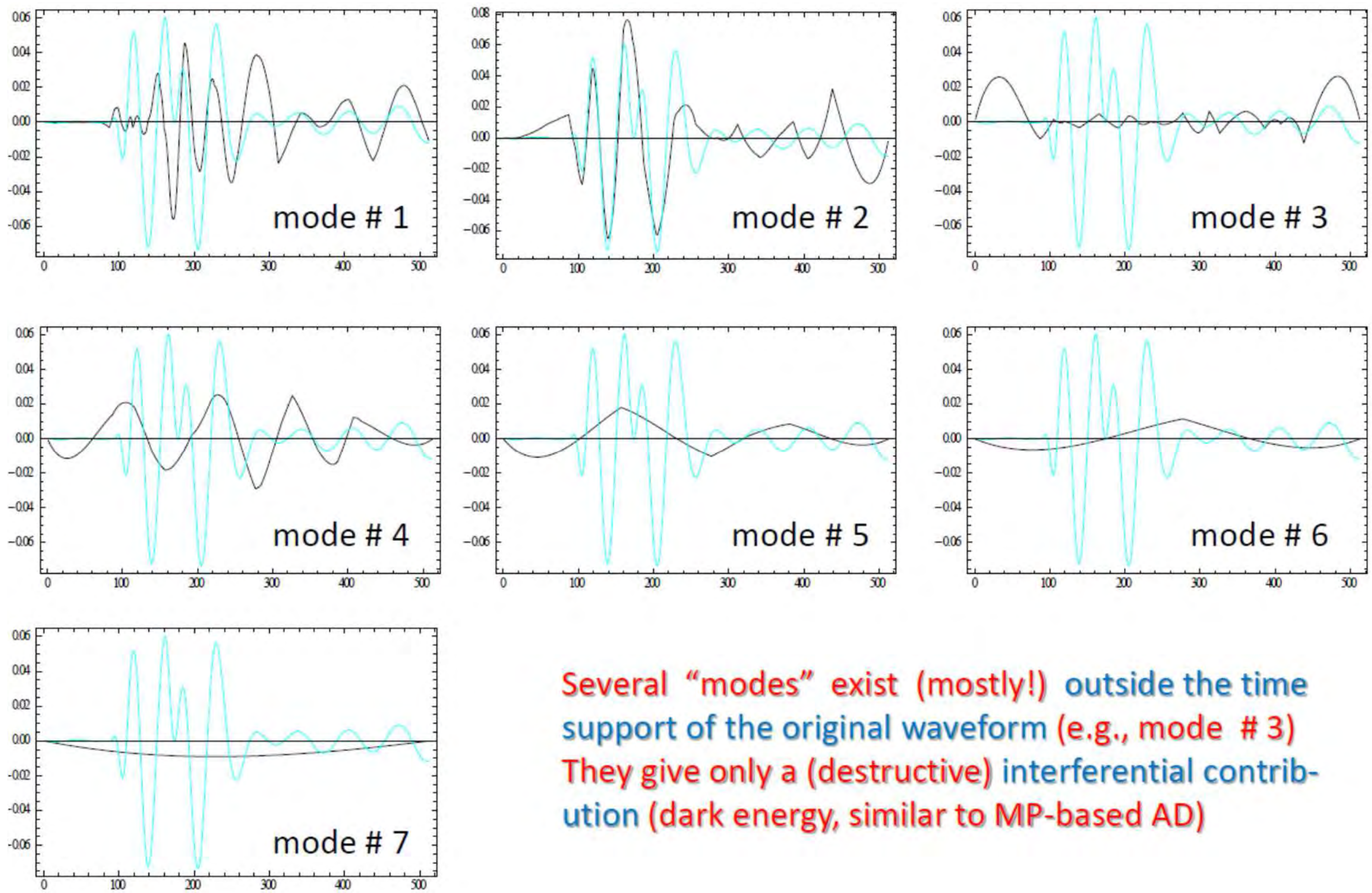
ii) The EMD algorithm implies time-continuity of each IMF. As a result, *distinct* signal components may be *grouped together* even if their time supports do not originally overlap.

iii) The claim frequently made in the GW Literature that "the HH method achieves *infinite frequency resolution*" is unsubstantial: evidence (Flandrin) shows that two (pure) tones ***will not*** be solved by EMD unless their freq-uencies differ by a ***confusion band***- depending on the tone amplitudes ratio.

iv) The numerical differentiation used to retrieve the instantaneous frequency is quite sensitive to noise. Boashash and coworkers compared the STFT, WD, RGK-smoothed WV, BD, HHT in the presence of noise for a number of represent-ative waveforms. They found that the RGK-smoothed WV and the HHT gave the best and worst performance, respectively.

HHT «Dark Energy» Problem

Huang empirical modes
(original waveform in cyan)



Several "modes" exist (mostly!) outside the time support of the original waveform (e.g., mode # 3)
They give only a (destructive) interferential contribution (dark energy, similar to MP-based AD)

Conclusions

TF tools are a natural choice to investigate the fine structure of glitches, and hopefully improving our understanding about their origin, and our skill in getting rid of them.

The Gabor duality bound limits TF resolution.

No Holy Grail TF technique exists: all techniques have their own strenghts and pitfalls.

CC-based TF skeletons produce readable TF representations of complex multi-component wave-forms, which are free from intermodulation artifacts and yield as much time-frequency resolution as possible *uniformly* throughout the TF plane.

Thanks for Your Kind Attention

ありがとうございました

More on Sparse Representations and CS applications in Physics at

http://dsp.rice.edu/cs

The Q-Transform, contd.

- The squared mag of the Q transform coefficients X are $\chi^2 -$ distributed with 2 d.o.f. ;
- Define a normalized TF-tile energy as $Z = \frac{|X|^2}{\langle |X|^2 \rangle_{(\text{same } f_0, Q)}}$;
- This is related to the "SNR" of the "corresponding" SG-correlator by $SNR = \sqrt{2Z}$;
- In the H_0 hypothesis (noise only) Z is exp-distributed, $prob(Z > \zeta) = \exp(-\zeta)$, which allows to set a threshold to identify candidate "triggers" ; the QT can be accordingly used to identify significant transients in
  - The GW Channel (GW bursts and glitches)
  - Auxiliary detector channels (glitches)
  - Environmental monitoring data (environment excitations)
- Performance improves if significant ("hot") tiles are *clustered* on the basis of their "proximity" in the TF plane [R. Khan and S. Chatterji, "Enhancing the Capabilities of LIGO Time-Frequency Plane Searches Through Clustering," CQG 26 (2009) 155009].