Time-Frequency Feature Extraction via Synchrosqueezing Transform and Its Application to Data Sonification

Alex Berrian, Jordan Leung & Naoki Saito

Dept. of Mathematics; Dept. of Statistics
University of California, Davis

JSIAM Annual Meeting
Kanazawa, Japan
September 11, 2015

Outline

Acknowledgment

- Xiang Cui (formerly Nanjing Univ., currently Univ. Washington)
- Geoff Schladow (UC Davis Tahoe Environmental Research Center)
- Hau-Tieng Wu (Univ. Toronto)
- Support from Office of Naval Research grant: ONR N00014-12-1-0177
- Support from National Science Foundation grants: DMS-1148643; DMS-1418779

Outline

Outline

Amplitude-phase decomposition

Let $f:\mathbb{R}\to\mathbb{C}$ be a signal and assume f can be written as a finite sum of (unknown) *amplitude-phase components (modes)*

$$f(t)=\sum_{k=1}^{K}A_k(t)\exp(2\pi\mathrm{i}\phi_k(t)),$$

where the $A_k(t)$ represent *instantaneous amplitudes* (IAs) and the $\phi_k(t)$ represent *instantaneous phases*. Denote $f_k(t):=A_k(t)\exp(2\pi\mathrm{i}\phi_k(t))$.

- We call the decomposition $f=\sum_{k=1}^{K}f_k$ an *amplitude-phase decomposition*.
- The $\phi_k'(t)$ are called *instantaneous frequencies* (IFs).
- For definiteness of this decomposition, we impose certain requirements on A_k and ϕ_k' (to be discussed later).

Motivation for amplitude-phase decomposition

What kinds of signals can be modeled using an amplitude-phase decomposition?

- Audio signals (music, speech, bat echolocation calls...)
- Medical/physiological signals (EEG, ECG, sEMG...)
- Mechanical signals (vibration, force signals; for machine condition monitoring and testing...)
- Radar/sonar signals

These signals have *time-varying oscillatory characteristics*.

General problem: determining f_k

General problem: We want to express

$$f(t) = \sum_{k=1}^{K} f_k(t) = \sum_{k=1}^{K} A_k(t) \exp(2\pi i \phi_k(t)).$$

How do we determine the amplitude-phase components f_k given that only f is known?

Starting point: Use the *short-time Fourier transform* to visualize A_k and ϕ_k'.

Outline

Short-time Fourier transform (STFT)

Given a window function $g \in L^2(\mathbb{R})$ centered at 0, the *short-time Fourier transform* (STFT) of a signal f is defined by

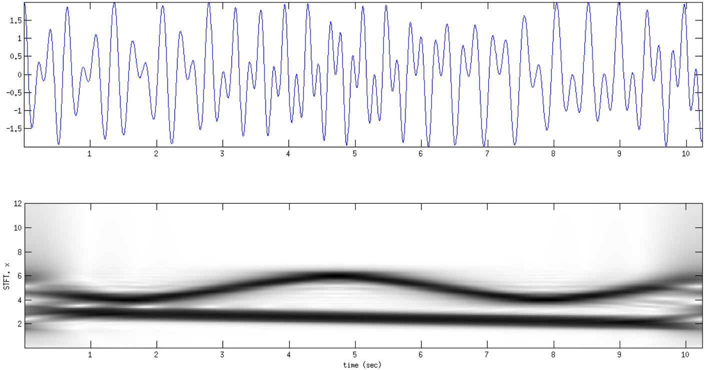
$$V_g f(t,\xi) := \int_{\mathbb{R}} f(x)\overline{g(x-t)}\mathrm{e}^{-2\pi \mathrm{i}\xi(x-t)}\,\mathrm{d}x.$$

If $\hat{f}, \hat{g} \in L^1(\mathbb{R})$ then we can *reconstruct* f from the STFT via the inversion formula

$$f(t) = \frac{1}{\overline{g(0)}} \int_{\mathbb{R}} V_g f(t,\xi)\,\mathrm{d}\xi,$$

provided that g is continuous in a neighborhood of 0 and $g(0) \neq 0$.

STFT example (spectrogram)



First plot: $f(t) := f_1(t) + f_2(t)$ where
$f_1(t) := [1 + 0.2\cos(t+1)] \cdot \cos(2\pi(3t - t^2/20))$ and
$f_2(t) := \sqrt{0.8 + t/10}\cos(2\pi(5t + \cos(t)))$ for $t \in [0, 10]$, 2000 samples.

Second plot: STFT of f, using window function g given by
$\hat{g}(\eta) := \exp\left([(\eta/\sigma)^2 - 1]^{-1}\right)$ for $|\eta| \le \sigma$, $\hat{g} = 0$ elsewhere, $\sigma = 0.8$.

How to retrieve f_k using STFT?

Due to the *uncertainty principle*, the STFT visualization of f is blurry. So the exact IAs and IFs are obscured.

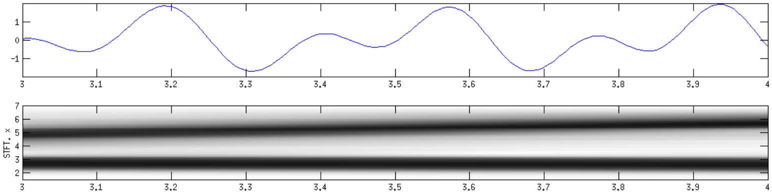
Solution: A post-processing method known as the *Synchrosqueezing transform* (SST)! [Daubechies & Maes 1996]

Outline

Outline

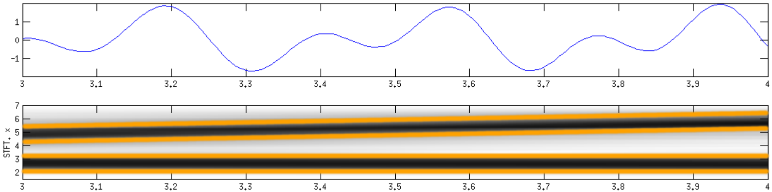
SST in pictures: frequency reassignment

Step 1: For each t, *reassign* the frequency locations η of *sufficiently large* STFT coefficients $V_g f(t,\eta)$ to an IF estimate $\xi_f(t,\eta)$, yielding the SST $S_{\gamma,f}(t,\xi)$.



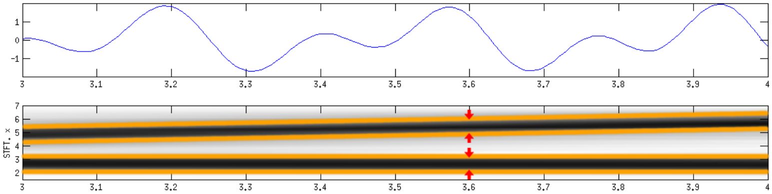
SST in pictures: frequency reassignment

Step 1: For each t, *reassign* the frequency locations η of *sufficiently large* STFT coefficients $V_g f(t,\eta)$ to an IF estimate $\xi_f(t,\eta)$, yielding the SST $S_{\gamma,f}(t,\xi)$.



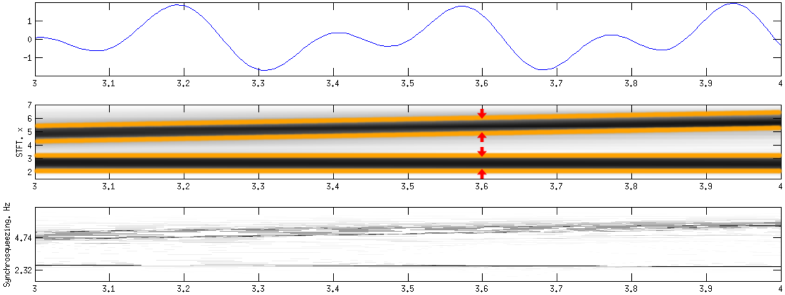
SST in pictures: frequency reassignment

Step 1: For each t, *reassign* the frequency locations η of *sufficiently large* STFT coefficients $V_g f(t,\eta)$ to an IF estimate $\xi_f(t,\eta)$, yielding the SST $S_{\gamma,f}(t,\xi)$.



SST in pictures: frequency reassignment

Step 1: For each t, *reassign* the frequency locations η of *sufficiently large* STFT coefficients $V_g f(t,\eta)$ to an IF estimate $\xi_f(t,\eta)$, yielding the SST $S_{\gamma,f}(t,\xi)$.

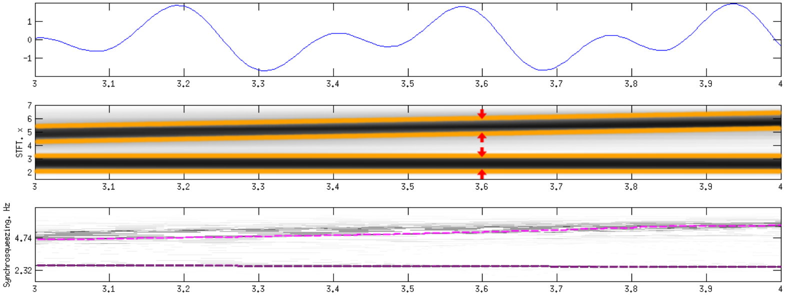


Define the SST by $S_{f,\gamma}(t,\xi) := \int_{A_{\gamma,f}(t)} V_g f(t,\eta)\delta\left(\xi - \xi_f(t,\eta)\right) \mathrm{d}\eta$, where

$A_{\gamma,f}(t) := \{\eta \in \mathbb{R}_+ : |V_g f(t,\eta)| > \gamma\}$, $\gamma > 0$, and $\xi_f(t,\xi) := \dfrac{\partial_t \left[V_g f(t,\xi)\right]}{2\pi \mathrm{i} V_g f(t,\xi)}$.

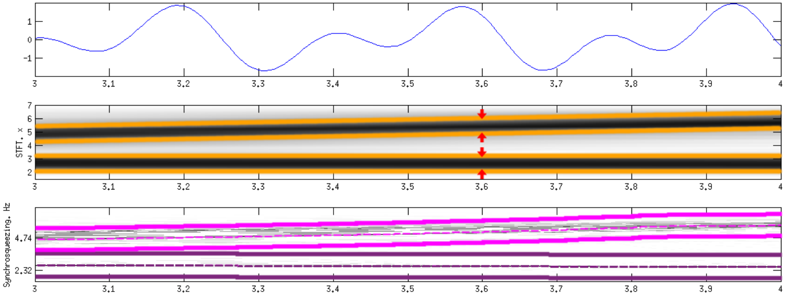
SST in pictures: reconstruction

Step 2: For each t, *integrate* the SST $S_{\gamma,f}(t,\xi)$
over a frequency band around an estimate of the curve $\phi_k'(t)$ to reconstruct $f_k(t)$.



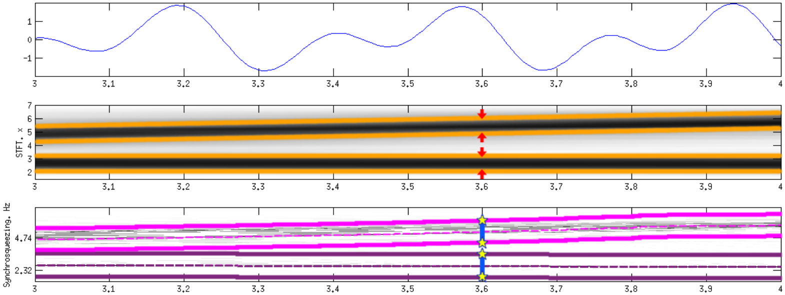
SST in pictures: reconstruction

Step 2: For each t, *integrate* the SST $S_{\gamma,f}(t,\xi)$ over a frequency band around an estimate of the curve $\phi_k'(t)$ to reconstruct $f_k(t)$.



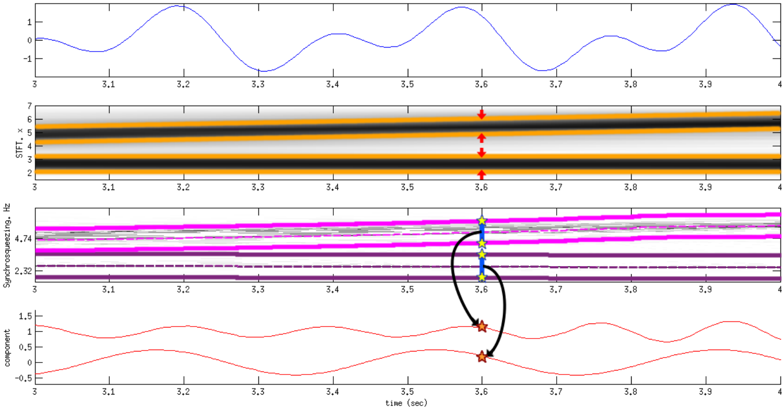
SST in pictures: reconstruction

Step 2: For each t, *integrate* the SST $S_{\gamma,f}(t,\xi)$ over a frequency band around an estimate of the curve $\phi_k'(t)$ to reconstruct $f_k(t)$.



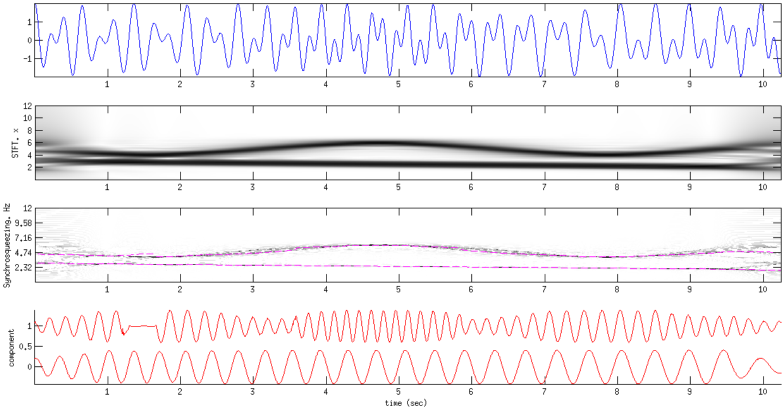
SST in pictures: reconstruction

Step 2: For each t, *integrate* the SST $S_{\gamma,f}(t,\xi)$ over a frequency band around an estimate of the curve $\phi_k'(t)$ to reconstruct $f_k(t)$.



Reconstruction formula: $f_k(t) \approx \int_{\{\xi\ :\ |\xi - \phi_k'(t)| < \gamma\}} \frac{1}{g(0)} \cdot S_{f,\gamma}(t,\xi)\, d\xi.$

SST in pictures: whole signal



Outline

Applications of SST

- Speaker identification from speech signal [Daubechies & Maes 1996]
- Fault diagnosis in planetary gearboxes for wind turbines [Feng, Chang & Liang 2015]
- Extracting heart-rate variability from electrocardiogram (ECG) signal [Daubechies, Lu & Wu 2011]
- Detecting stages of sleep from respiratory signal [Wu 2013]
- Quantifying the effect of solar radiation on a key paleoclimate change on Earth [Thakur et al. 2013]
- ...

Outline

SST-STFT: Assumptions on signal and window

Definition: We say that $f \in \mathscr{B}_{\epsilon,d}^{\mathrm{STFT}}$ if we can write $f = \sum_{k=1}^{K} f_k = \sum_{k=1}^{K} A_k \exp(2\pi i \phi_k)$ for some $K \in \mathbb{Z}^+$ and if moreover there exist $\epsilon, d > 0$ such that for each $k \in \{1, \dots, K\}$,

- A_k and ϕ_k' are *bounded and sufficiently smooth*: $A_k \in C^1(\mathbb{R}) \cap L^\infty(\mathbb{R})$, $\phi_k \in C^2(\mathbb{R})$, $\phi_k' \in L^\infty(\mathbb{R})$, $\inf_{t\in\mathbb{R}} A_k(t) > 0$ and $\inf_{t\in\mathbb{R}} \phi_k'(t) > 0$;
- A_k is *weakly-modulated*: $\forall\, t \in \mathbb{R}$, $|A_k'(t)| \le \epsilon |\phi_k'(t)|$;
- ϕ_k' is *slowly-varying*: $\forall\, t \in \mathbb{R}$, $|\phi_k''(t)| \le \epsilon |\phi_k'(t)|$;
- ϕ_k' is *well-separated from the other IFs*: if $k \ge 2$, $t \in \mathbb{R}$ we have $\phi_k'(t) - \phi_{k-1}'(t) > d$.

Definition: We say that $g \in \mathscr{W}_d^{\mathrm{STFT}}$ if $g \in \mathscr{S}(\mathbb{R})$ (the Schwartz space), $g(0) \neq 0$, and $\mathrm{supp}(\hat{g}) \subset [-d/2, d/2]$.

SST-STFT: Theorem

The following theorem is due to Thakur & Wu (2011) and Oberlin, Meignen & Perrier (2014):

Theorem: Suppose there exist $\epsilon, d > 0$ such that $f = \sum_{k=1}^{K} f_k \in \mathscr{B}_{\epsilon,d}^{\mathsf{STFT}}$ and $g \in \mathscr{W}_d^{\mathsf{STFT}}$. Let $\tilde{\epsilon} = \epsilon^{1/3}$. Then, provided that $\tilde{\epsilon}$ is sufficiently small, the following results hold:

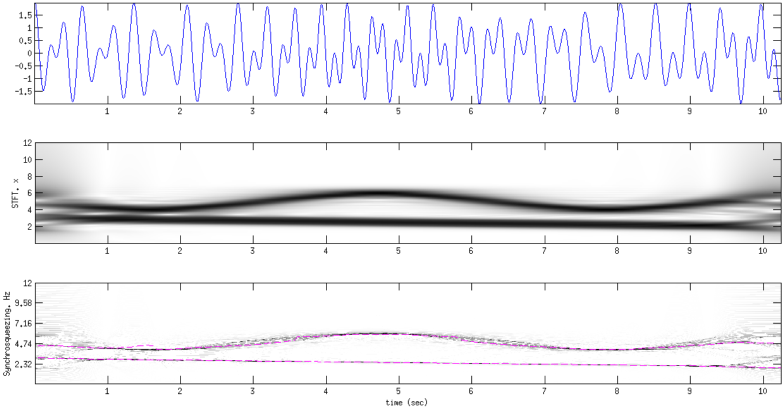
- **(Concentration of STFT around IF curve)**
  $|V_g f(t,\xi)| > \tilde{\epsilon}$ only when there exists $k \in \{1,\ldots,K\}$ such that $(t,\xi) \in Z_k := \{(t,\xi)\ :\ |\xi - \phi_k'(t)| < d/2\}$.
- **(Closeness of reassigned frequency ξ_f to IF)**
  For all $k \in \{1,\ldots,K\}$ and all $(t,\xi) \in Z_k$ such that $|V_g f(t,\xi)| > \tilde{\epsilon}$, we have $|\xi_f(t,\xi) - \phi_k'(t)| \leq \tilde{\epsilon}$.
- **(Accuracy of reconstruction)**
  For every $k \in \{1,\ldots,K\}$ there is a constant $C > 0$ such that for all times $t \in \mathbb{R}$,

$$\left| \int_{\{\xi\ :\ |\xi - \phi_k'(t)| < \tilde{\epsilon}\}} \frac{1}{\overline{g(0)}} \cdot S_{f,\tilde{\epsilon}}(t,\xi)\, \mathrm{d}\xi - f_k(t) \right| \leq C\tilde{\epsilon}.$$

Outline

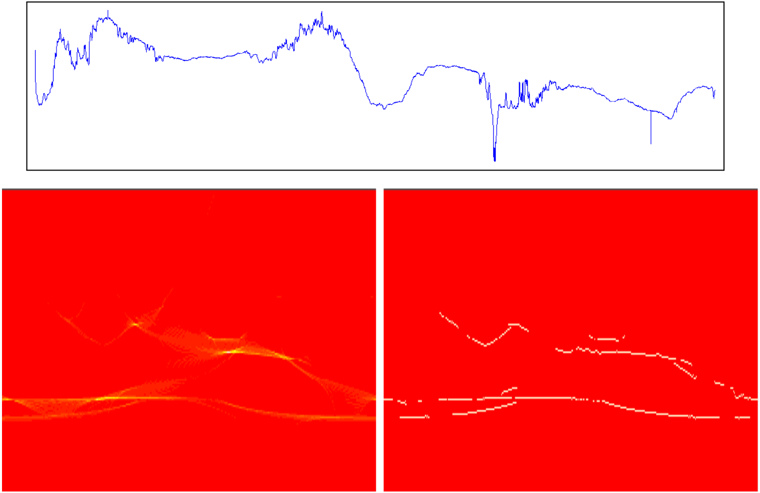
Default ridge extraction method

Default method: Find maximum of a functional measuring the time-frequency energy and smoothness of curves using a greedy algorithm.



Crazy climbers method

Crazy climbers method: [Carmona, Hwang & Torrésani 1999]
Markov chain Monte Carlo approach to detect ridges.



Outline

Open problems

1. How to *increase time resolution* and still provide a reconstruction formula for the modes f_k?
2. Can we address the case of *strongly modulated and/or discontinuous* IAs A_k?
3. How to handle *crossing* IFs?

Quilted STFT to address open problems 1 & 2

1. ***Quilted Gabor Transform* (QGT) for improved time resolution:**
   Generalization of STFT, permitting for different analysis functions in different regions of interest *(quilt patches)* in time-frequency plane. Theoretically guaranteed to allow for reconstruction [Dörfler 2011].
2. **QGT for strongly modulated/discontinuous IAs:**
   QGT captures sudden amplitude changes more accurately with improved time resolution.

On open problem 3, we will not address in this presentation.

Outline

Outline

Quilted short-time Fourier transform

First, we define a two-parameter family of *quilted window functions* $h_{t,\xi}(x)$.

- For each $(t,\xi) \in \mathbb{R} \times \mathbb{R}^+$, $h_{t,\xi}$ is a function in $L^2(\mathbb{R})$ which is *centered at 0*.
- To guarantee SST accuracy, the $h_{t,\xi}$ will satisfy certain requirements which we give later.

Then we define the *quilted short-time Fourier transform* (QSTFT) of a signal f by

$$V_g^Q f(t,\xi) := \int_{\mathbb{R}} f(x)\overline{h_{t,\xi}(x-t)}\mathrm{e}^{-2\pi\mathrm{i}\xi(x-t)}\,\mathrm{d}x.$$

Outline

SST-QSTFT: Continuous formulation

We define the *Synchrosqueezing transform based on the quilted short-time Fourier transform* (SST-QSTFT) of a signal f, with tolerance $\gamma > 0$, by

$$S^{Q}_{f,\gamma}(t,\xi) := \int_{A^{Q}_{\gamma,f}(t)} V^{Q}_{g} f(t,\eta)\delta\left(\xi - \xi^{Q}_{f}(t,\eta)\right) \mathrm{d}\eta,$$

where $A^{Q}_{\gamma,f}(t) := \{\eta \in \mathbb{R}_{+} : |V^{Q}_{g} f(t,\eta)| > \gamma\}$, and where

$$\xi^{Q}_{f}(t,\xi) := \frac{\partial_t \left[V^{Q}_{g} f(t,\xi)\right]}{2\pi \mathrm{i} V^{Q}_{g} f(t,\xi)}$$

is a function which gives an approximation to IF.

Definition: Given $\epsilon, d > 0$, we say that the two-parameter family $\{h_{t,\xi}\}_{(t,\xi)\in\mathbb{R}\times\mathbb{R}^+}$ is of the class $\mathscr{W}_{d,\epsilon}^{\text{QSTFT}}$ if:

- For each $(t,\xi)\in\mathbb{R}\times\mathbb{R}^+$ we have $h_{t,\xi}\in\mathscr{W}_d^{\text{STFT}}$.
- For each $t\in\mathbb{R}$ and $k\in\{1,\ldots,K\}$, there is some constant $a_{t,k}\in\mathbb{R}$ such that $h_{t,\xi}\equiv a_{t,k}$ for all ξ in the frequency band $\{\xi : |\xi-\phi_k'(t)| < d/2\}$.
- The integrals $I_m(t,\xi) := \int_{\mathbb{R}} |u|^m |h_{t,\xi}(u)|\,\mathrm{d}u$ and $J_m(t) := \int_{\mathbb{R}} |u|^m |h_{t,\xi}'(u)|\,\mathrm{d}u$ satisfy $\sup_{t,\xi\in\mathbb{R}\times\mathbb{R}^+} \{I_m(t,\xi), J_m(t,\xi)\} < \infty$ for all $m\in\{0,1,2\}$.
- Defining $h(x,t,\xi) := h_{t,\xi}(x)$, we have that for all $(t,\xi)\in\mathbb{R}\times\mathbb{R}^+$, $\int_{\mathbb{R}} |\partial_t h(u,t,\xi)|\,\mathrm{d}u < \infty$.

The third condition ensures the closeness of ξ_f^Q to ϕ_k'. The fourth condition bounds the amount of variation that the window family $h_{t,\xi}$ can exhibit over t.

SST-QSTFT

Theorem (B., 2015): Let $\epsilon > 0$, $\tilde{\epsilon} = \epsilon^{1/3}$, $d > 0$. Assume that $\{h_{t,\xi}\}_{(t,\xi)\in\mathbb{R}\times\mathbb{R}^+}$ is of the class $\mathscr{W}_{d,\epsilon}^{\text{QSTFT}}$. Suppose that $f = \sum_{k=1}^{K} f_k \in \mathscr{B}_{\epsilon,d}^{\text{STFT}}$. Then, if ϵ is sufficiently small, the following results hold:

- **(Concentration of QSTFT around IF curve)** $|V_g^Q f(t,\xi)| > \tilde{\epsilon}$ only when there is a $k \in \{1,\ldots,K\}$ such that $(t,\xi) \in Z_k := \{(t,\xi) \,:\, |\xi - \phi_k'(t)| < d/2\}$.
- **(Closeness of reassigned frequency ξ_f^Q to IF)** For all $k \in \{1,\ldots,K\}$ and all $(t,\xi) \in Z_k$ such that $|V_g^Q f(t,\xi)| > \tilde{\epsilon}$, we have $|\xi_f^Q(t,\xi) - \phi_k'(t)| \le \tilde{\epsilon}$.
- **(Accuracy of reconstruction)** For every $k \in \{1,\ldots,K\}$ there is a constant C_k such that for all times $t \in \mathbb{R}$,

$$\left| \int_{\{\xi \,:\, |\xi - \phi_k'(t)| < \tilde{\epsilon}\}} \frac{1}{\overline{h_{t,\xi}(0)}} \cdot S_{f,\gamma}^Q(t,\xi)\, \mathrm{d}\xi - f_k(t) \right| \le C_k \tilde{\epsilon}.$$

Numerical example 1: Synthetic test signal

We try the SST-QSTFT on a challenging synthetic test signal $f(t) := f_1(t) + f_2(t)$ for $t \in [0,1]$, where

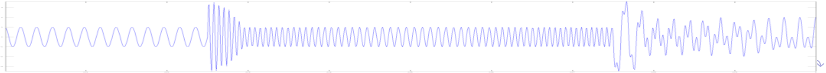
$$f_1(t) := \begin{cases} A_1(t)\cos(2\pi \cdot 51 t) & \text{if } t \in [0,.25) \cup (.75,1] \\ 0 & \text{if } t \in [.25,.75], \end{cases}$$

$$f_2(t) := \begin{cases} 0 & \text{if } t \in [0,.25) \\ A_2(t)\cos(2\pi \cdot (131 t - \frac{30}{4\pi}\sin(4\pi t))) & \text{if } t \in [.25,1], \end{cases}$$

and where $A_1(t)$ and $A_2(t)$ are strongly-modulated, discontinuous amplitude functions of the form

$$A_k(t) := \begin{cases} 1 & \text{if } t \in S_k \\ 7 \cdot \left(\exp\left[\left(\left(\frac{t-t_k}{.05}\right)^2 - 1\right)^{-1}\right] + 1\right) & \text{if } t \in O_k, \end{cases}$$

where $O_1 = [.75,.8]$, $O_2 = [.25,3]$ are the intervals of strong amplitude modulation, $t_1 = .75$, $t_2 = .25$ are the onset times, and $S_1 = [0,.25] \cup (.8,1]$ and $S_2 = (.3,1]$ are the interval of stable, constant amplitude 1.



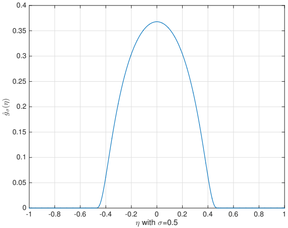
Numerical example 1: Quilted window function choice

Next, we use a *spectral flux function* [Dixon 2006], which measures temporal change in magnitude in each frequency bin and sums them within some range of frequency band, to estimate the onset and offset times of signal components. Let $\tilde{t}_1 \approx .743$ and $\tilde{t}_2 \approx .243$ be the estimates of the actual onset times $t_1 = .75$ and $t_2 = .25$ from before. Then, our window $h_{t,\xi}$ has the form

$$h_{t,\xi}(x) := \begin{cases} g_{60}(x) & \text{for } t \in [\tilde{t}_2 - \nu, \tilde{t}_2 + \nu],\ \xi \in [90, 240], \\ g_{45}(x) & \text{for } t \in [\tilde{t}_1 - \nu, \tilde{t}_1 + \nu],\ \xi \in [0, 100], \\ g_{30}(x) & \text{elsewhere}, \end{cases}$$

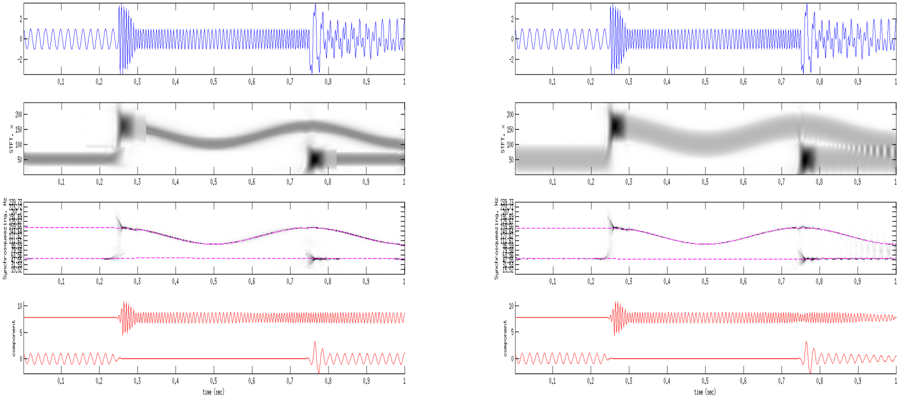
where $\nu > 0$ is a user-prescribed parameter, and for generic $\sigma > 0$ we define g_σ to be the Fourier-side bump function given by

$$\widehat{g_\sigma}(\eta) := \begin{cases} \exp\left[\left(\left(\eta/\sigma\right)^2 - 1\right)^{-1}\right] & \text{if } |\eta| < \sigma \\ 0 & \text{if } |\eta| \geq \sigma. \end{cases}$$



We remark that for each fixed t, $h_{t,\xi}$ is constant in ξ over the frequency bands $\{\xi : |\xi - \phi_k'(t)| < 10\}$, $k = 1, 2$.

Numerical example 1 (continued)

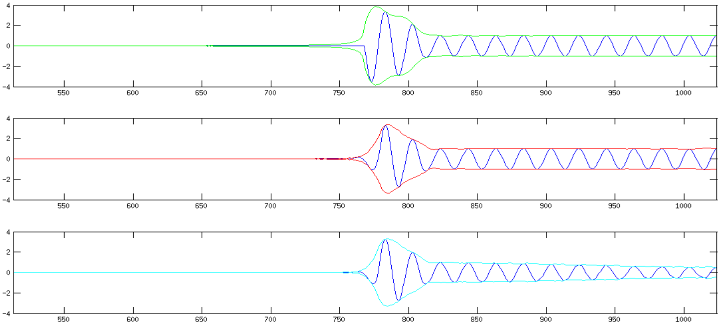


Left: SST-QSTFT of f, using window g_{60} on the upper component at the time of its onset, g_{45} on the lower component at the time of its second onset, and g_{30} elsewhere.

Right: SST-STFT of f, using g_{60} as the window.

For this example, we sample f at sampling rate $f_s = 1024$ and take $\nu = 80/f_s$.

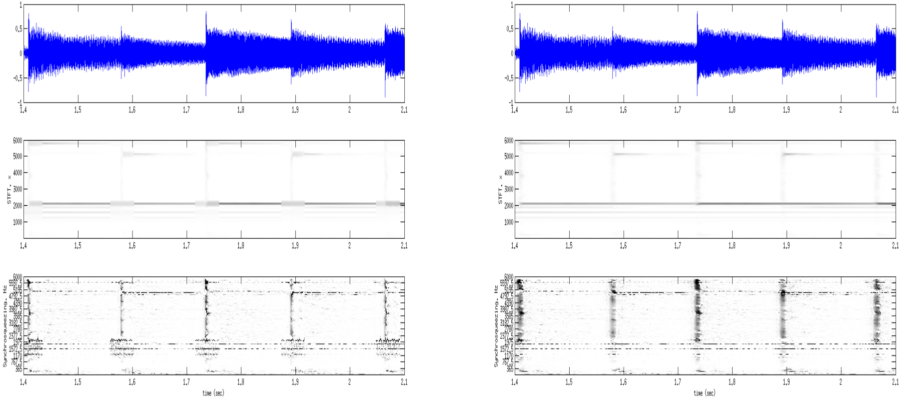
Numerical example 1 (continued)



Transform	Relative ℓ^2 reconstruction error	Relative ℓ^1 reconstruction error
SST-QSTFT	.3195	.1406
SST-STFT	.3626	.2760

Comparison of relative ℓ^2- and ℓ^1-norm errors in the envelope of the reconstructed signal from samples 513 to 1024 (the second half of the signal) in the first component $f_1(t)$ of the signal $f(t)$.

Numerical example 2 (onset detection)



Glockenspiel signal, analyzed by SST-QSTFT and SST-STFT.

Left: SST-QSTFT, using window g_{250} around estimated onset times and g_{100} elsewhere. Right: SST-STFT, using window g_{100}.

Outline

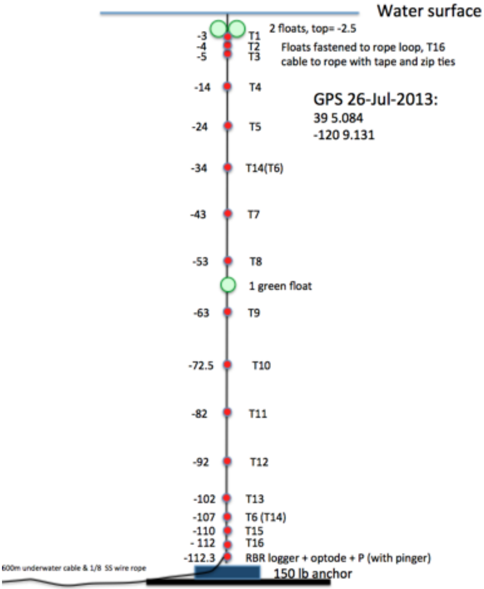
Outline

Lake Tahoe temperature dataset: description

In this part, we work with a dataset of lake temperatures, described as follows:

- 16 thermometers placed at a location in Lake Tahoe, each at different depths.
- Temperature measurements taken every 30 seconds, for approximately 70 days. $202150 \cdot 16$ data points in total.

Lake Tahoe temperature dataset: description



Outline

Sonification: Motivation

- **Goal:** Convey the separate short-term oscillatory and long-term trend information from each of the temperature signals, while still enabling their simultaneous "reading."
- **Obstacle:** Visualization of all this information may be difficult to read.
- **Idea:** Using the power of our auditory system, one has some hope of "hearing" all the information together.

This leads to the idea of *sonification*: the translation of data into sound!

Sonification with SST: Main ideas

- SST can be used to extract the oscillatory characteristics from the data, in the form of IF curves.
- Oscillatory characteristics represent diverse events occurring in the fluid flow of the lake.
- **Musical sonification:** Since music signals share similar oscillatory characteristics to those in our data, it is natural to consider a musical model of these IF curves.

Hence, we propose to model this Lake Tahoe temperature dataset as a musical composition using SST.

Outline

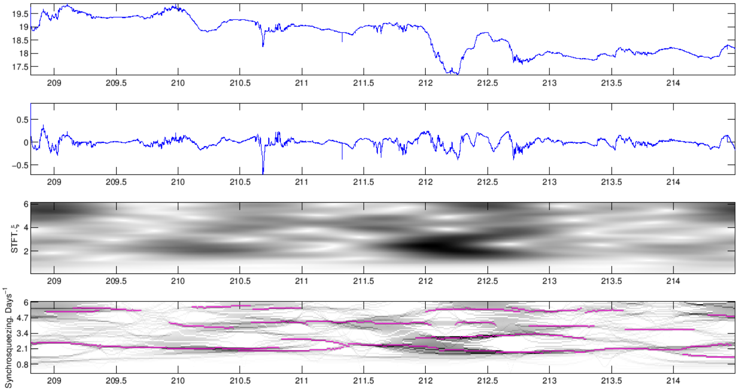
Sonification with SST: Algorithm

The algorithm to sonify the temperature dataset proceeds as follows:

1. Assign an instrument to each thermometer.
   - Lower thermometers are assigned lower-pitched instruments; higher thermometers are assigned higher-pitched instruments.
2. Extract IF curves ϕ'_k from the data using SST.
3. Map the IF curves to pitches in a musical scale (major, minor, Dorian, etc.) in an audible range, using the open-source software JythonMusic.
4. Use a LOESS (locally weighted polynomial regression) method to extract each signal's trend (seasonality).
5. Map the trend to MIDI volume values.
6. Export the information as MIDI data, using JythonMusic.

Outline

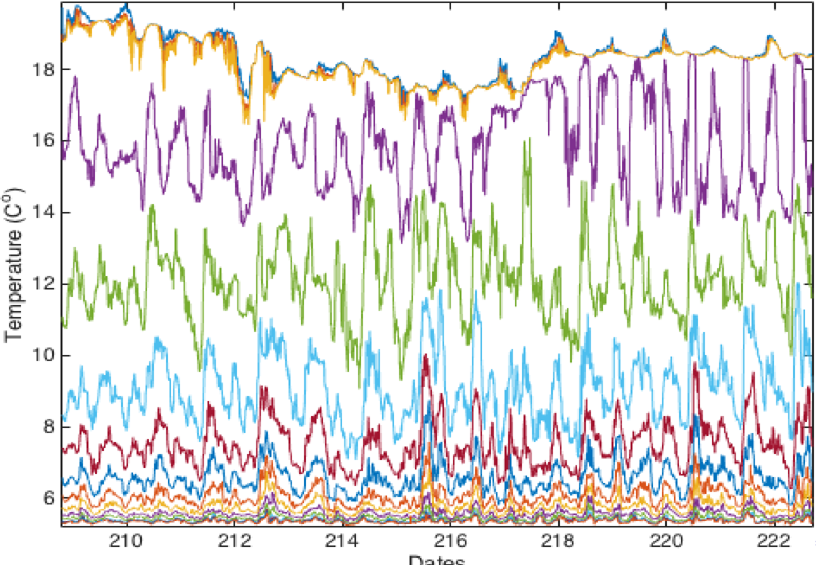
Steps of SST-based sonification



From top to bottom: the temperature measurement `T1`; its detrended version; the magnitude of the STFT; SST-STFT with the extracted IF curves overlaid in magenta.

Performance of SST-based sonification

About 14 days amount of data after subsampling every 5th time sample:



Summary

- SST is quite useful for revealing and extracting modes in nonstationary multicomponent signals for various applications.
- In particular, SST-QSTFT can provide improved time resolution and handle signals having strongly-modulated and/or discontinuous IAs.
- SST could extract IA/IF curves from the Lake Tahoe temperature measurements, which allowed us to *sonify* them nicely.
- Yet, the current versions of SST have difficulty to handle *crossing* IF curves, which often occur in practice. Investigate *Chirplet transforms*?
- Also, it is important to consider how to (semi-)automatically set all the parameters in SST-QSTFT .
- As for the Lake Tahoe dataset, we will examine the *isothermal displacement* data generated from those temperature measurements.

Thank you!