

# 8T1: Spectral-based sound transformations (1 of 2)

***Xavier Serra***

Universitat Pompeu Fabra, Barcelona

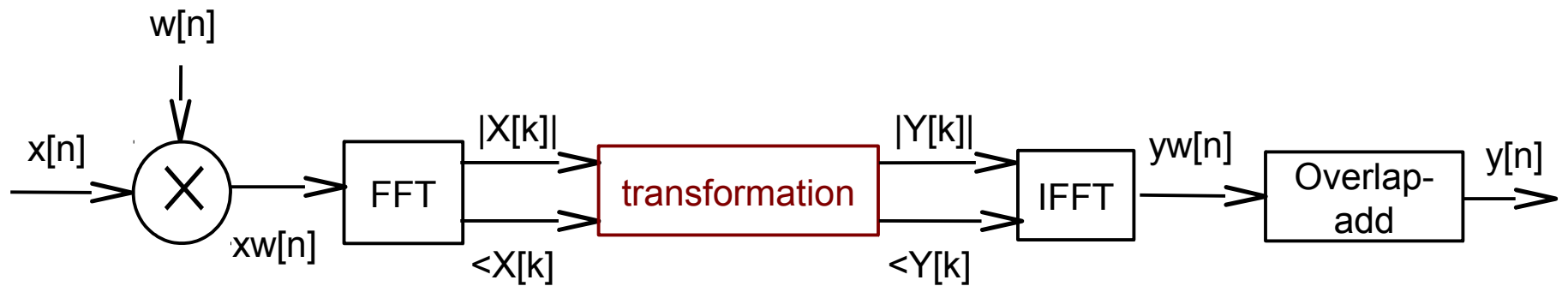
&

Stanford University

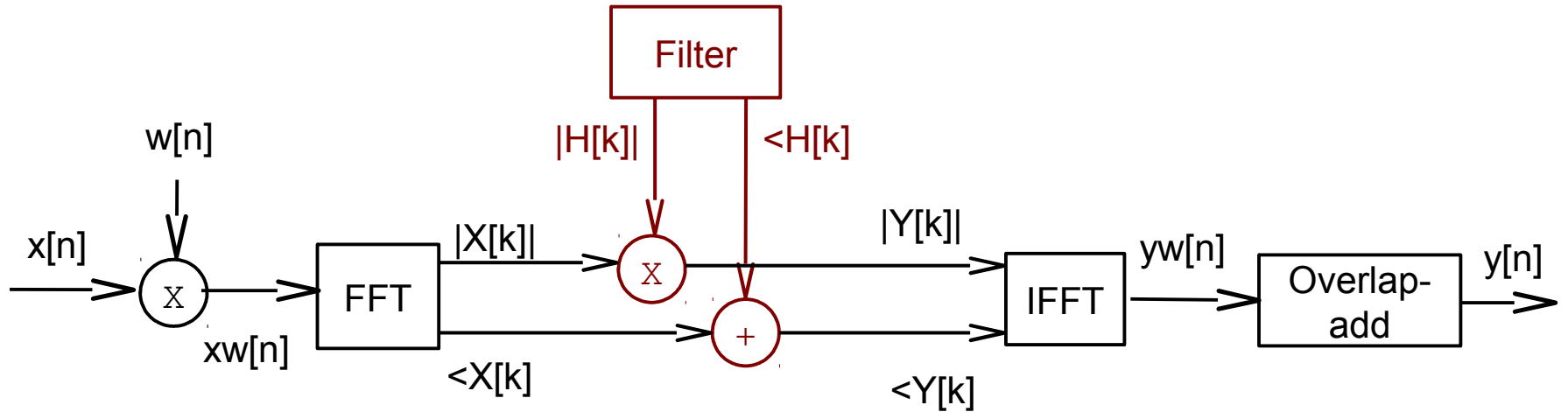
# Index

- Short-time Fourier transform
  - Filtering
  - Morphing
- Sinusoidal model
  - Frequency scaling
  - Time scaling

# Short-time Fourier transform

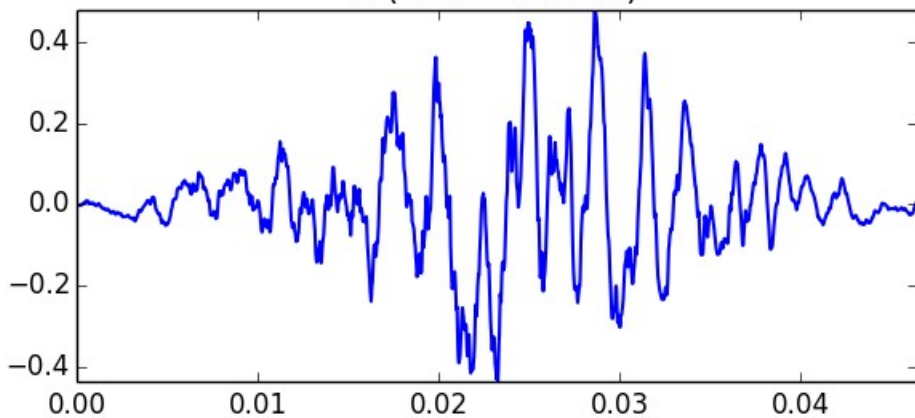


# Filtering with STFT

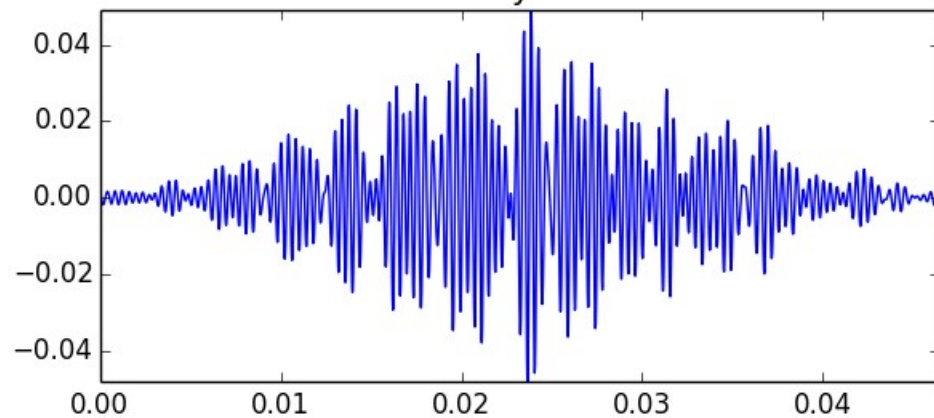


$$Y_l[k] = |H[k]| |X_l[k]| e^{j(\angle H[k] + \angle X_l[k])}$$

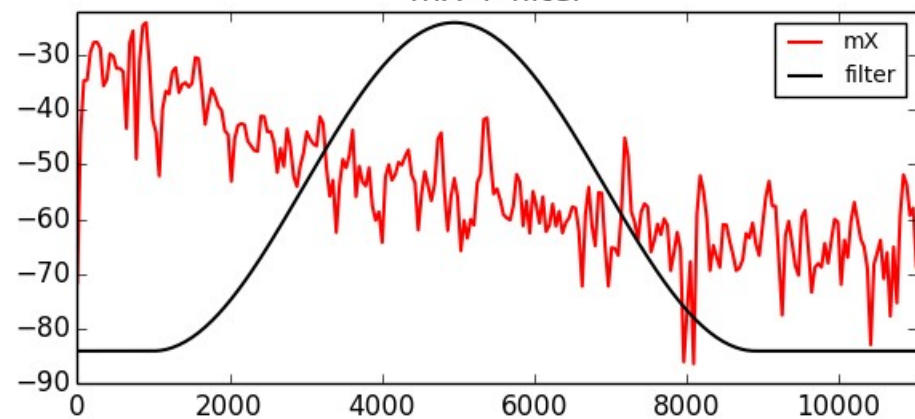
x (orchestra.wav)



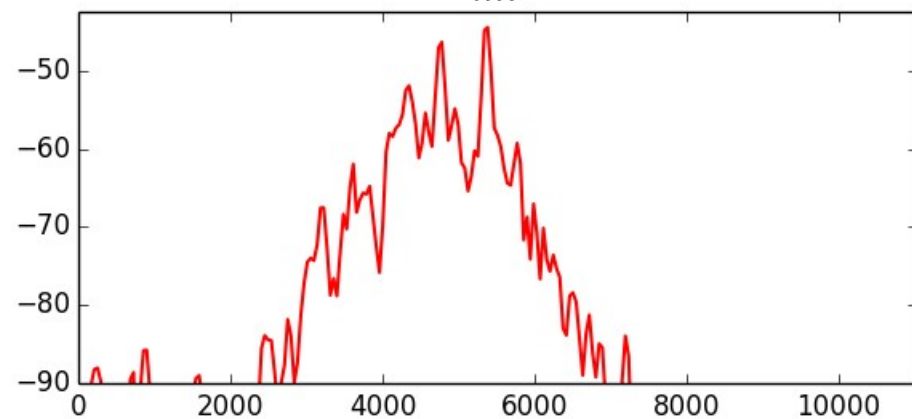
y



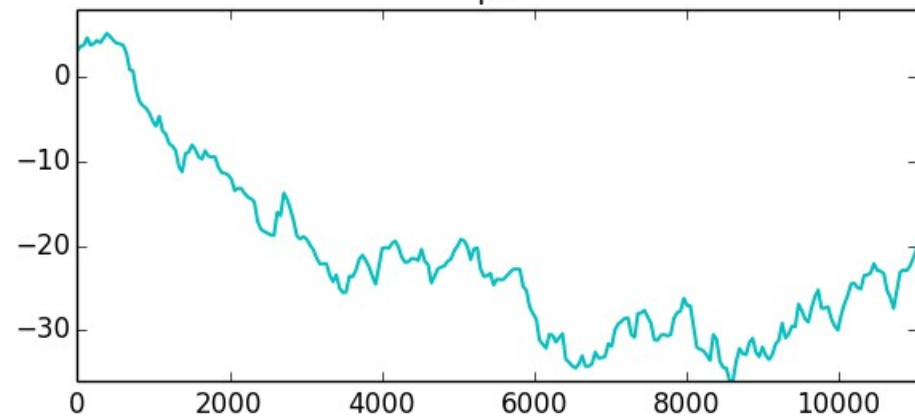
mX + filter



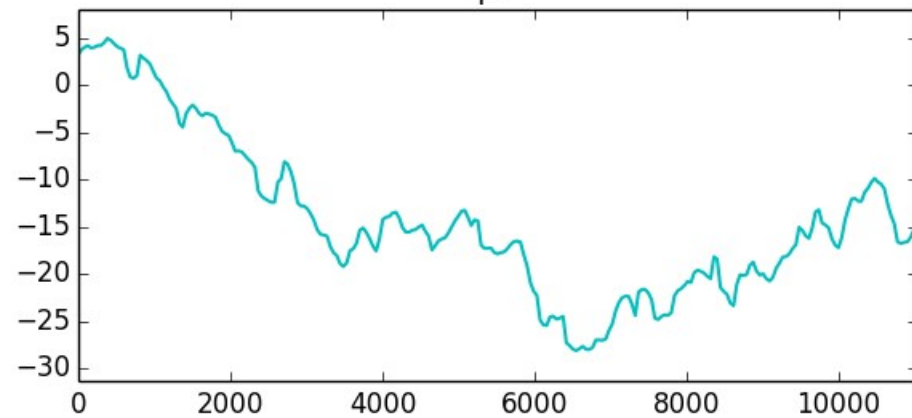
mY



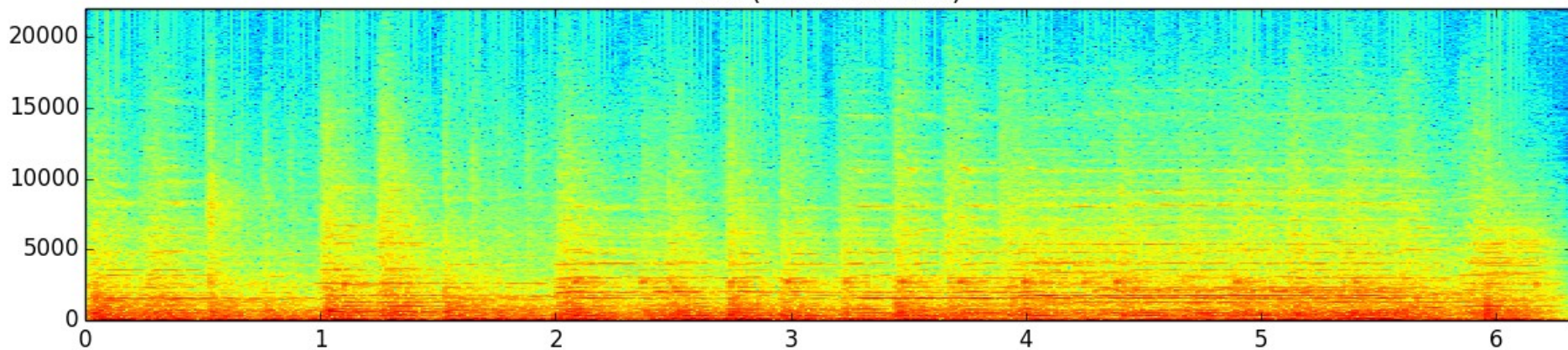
pX



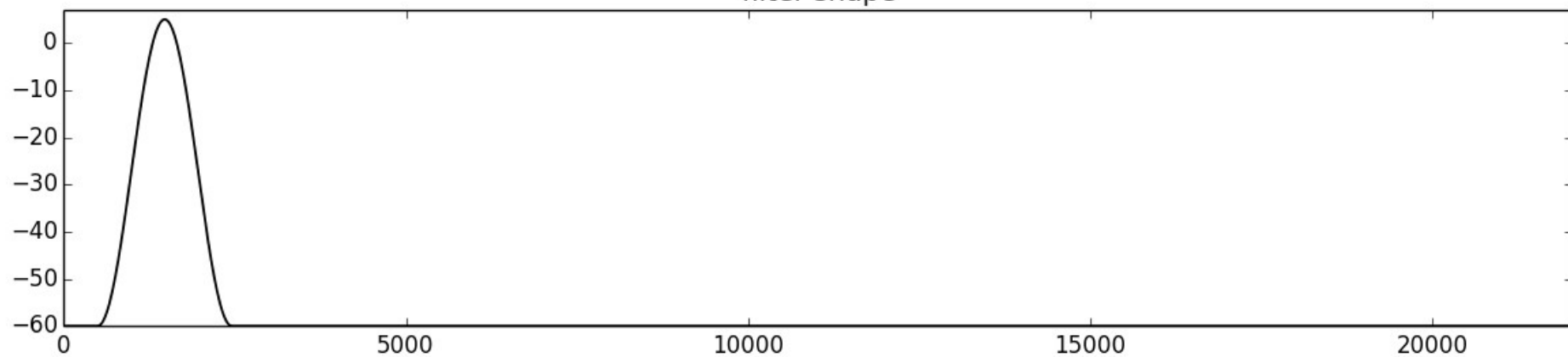
pY



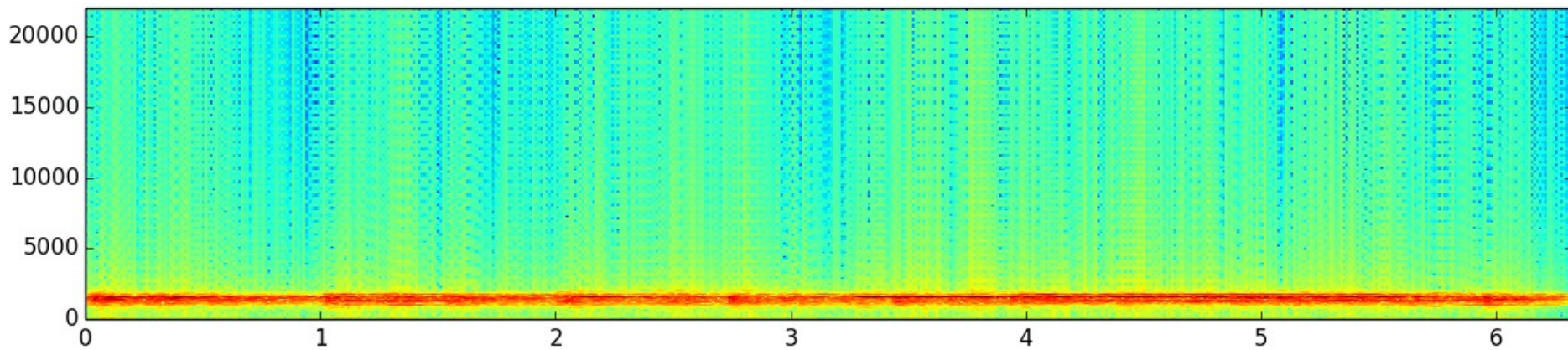
mX (orchestra.wav)



filter shape

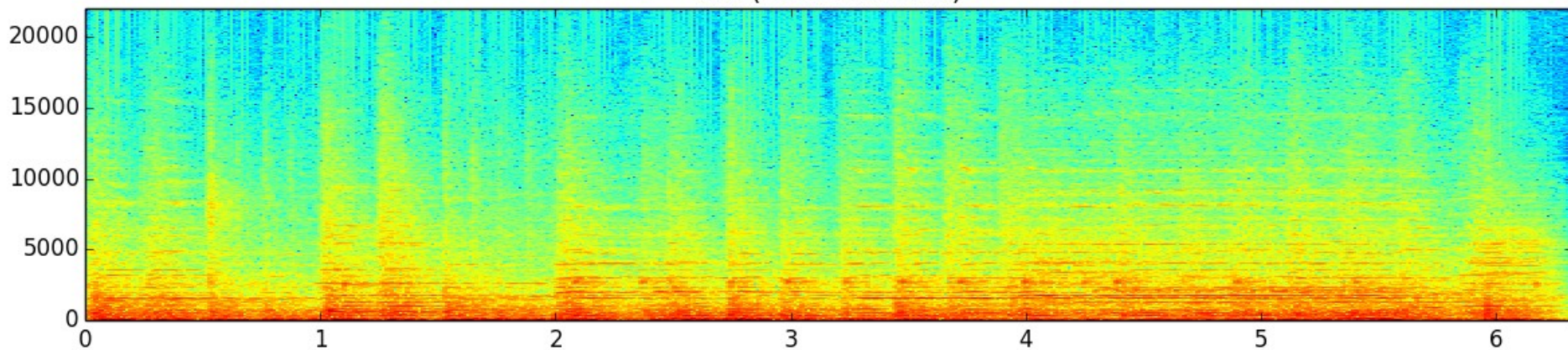


mY

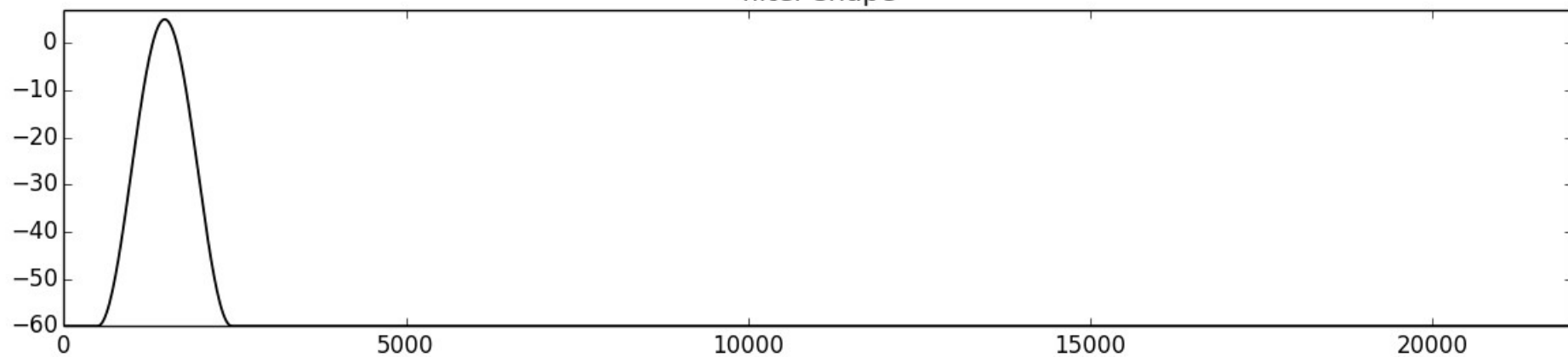




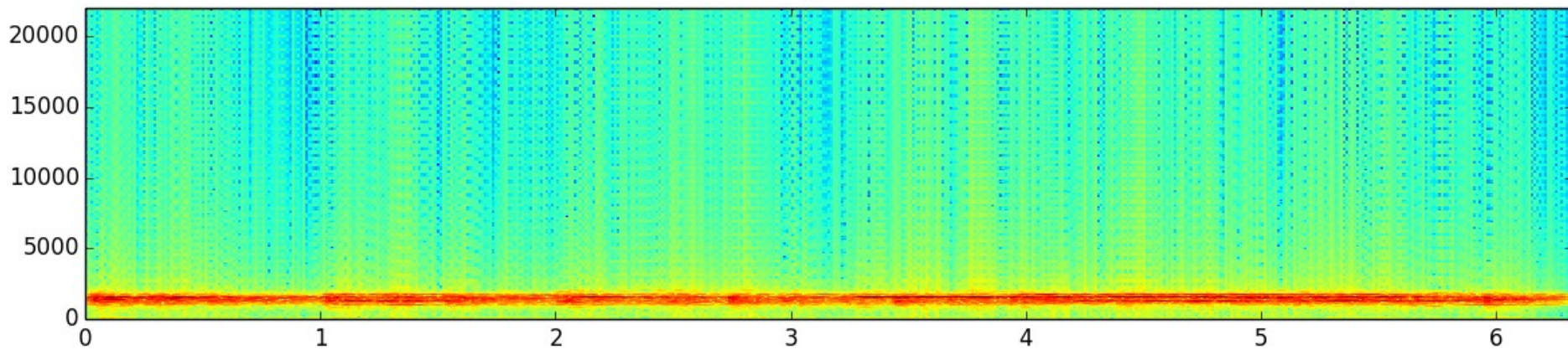
mX (orchestra.wav)



filter shape

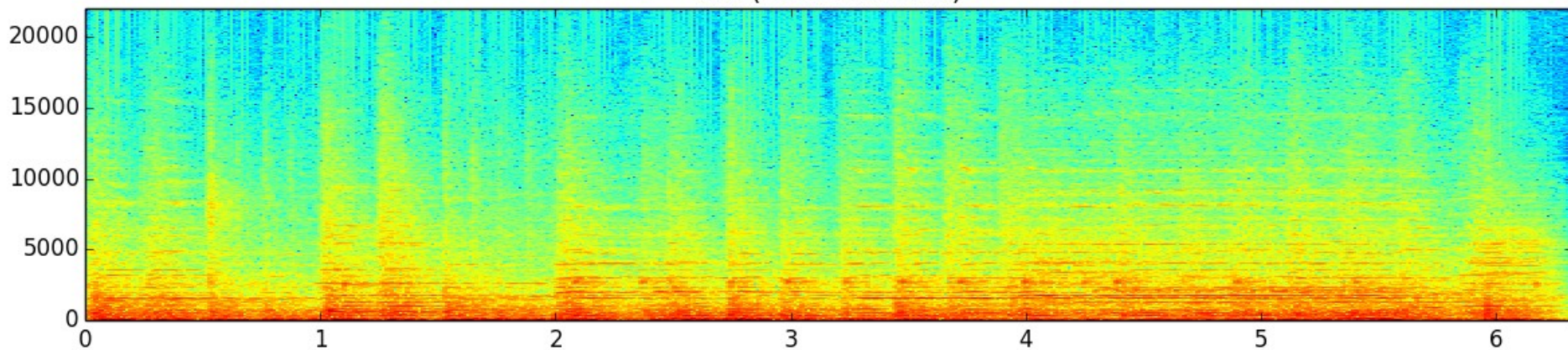


mY

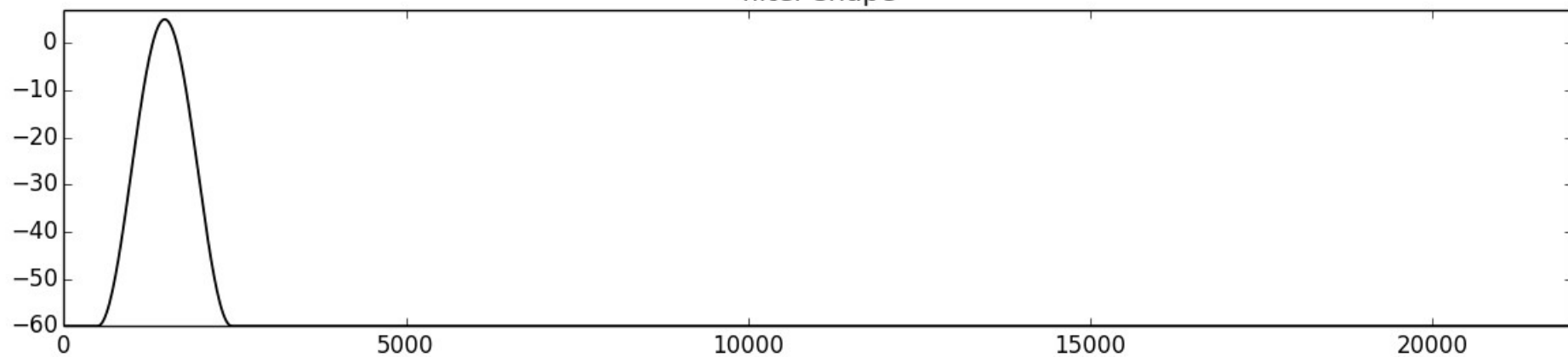




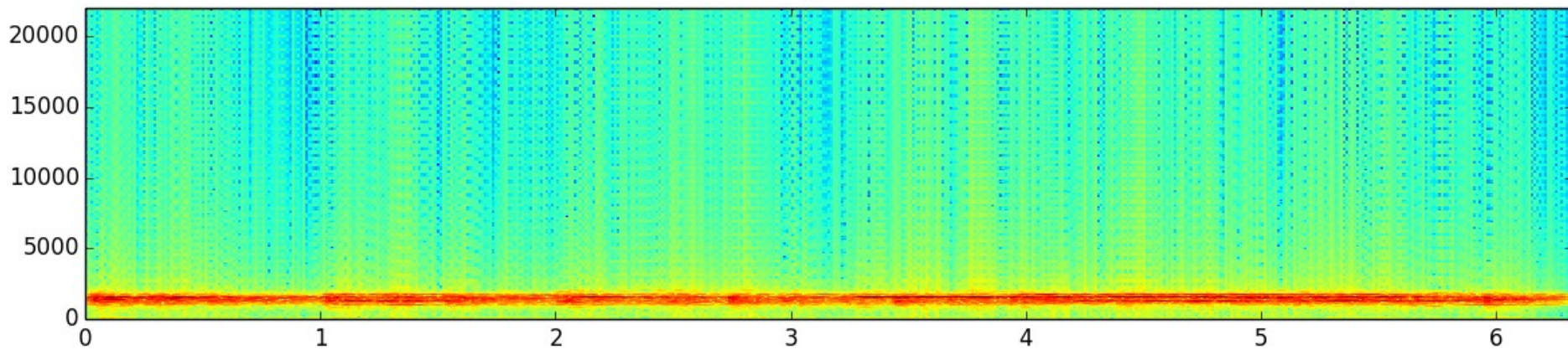
mX (orchestra.wav)



filter shape

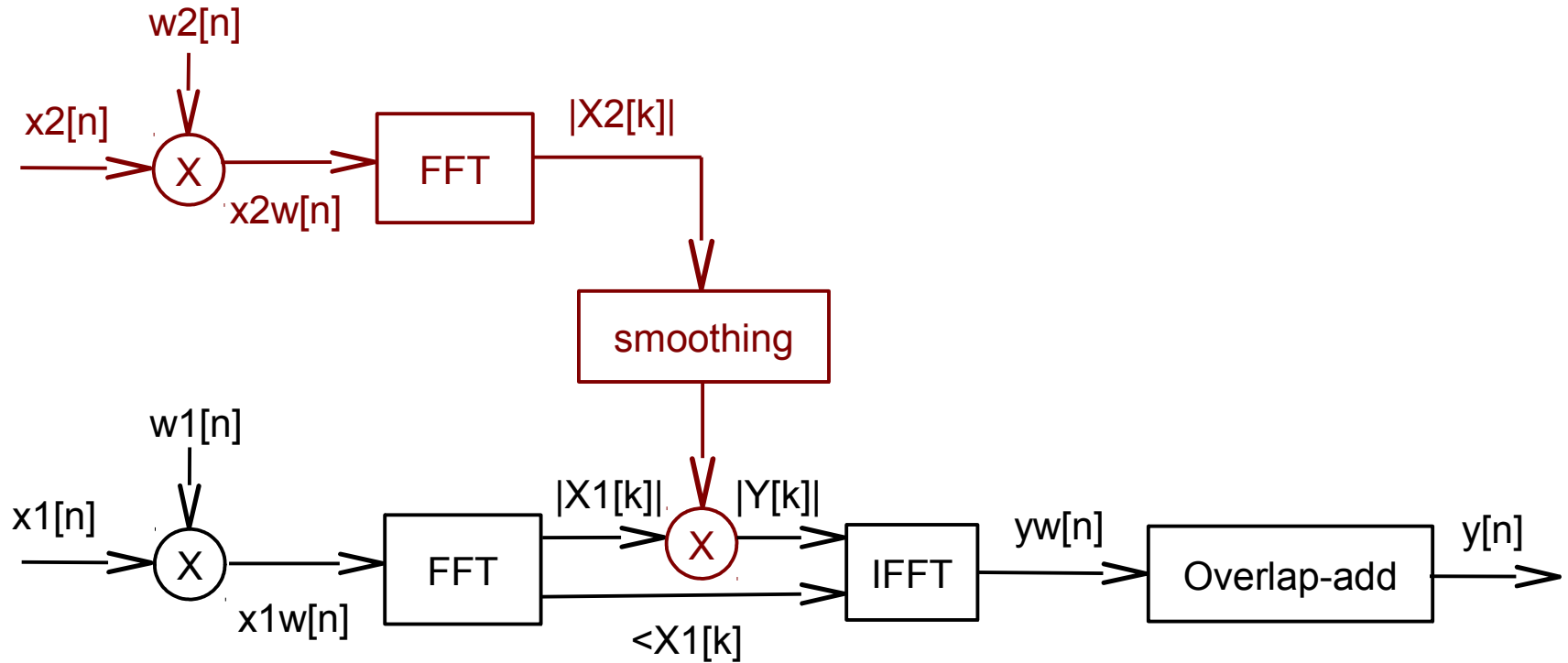


mY



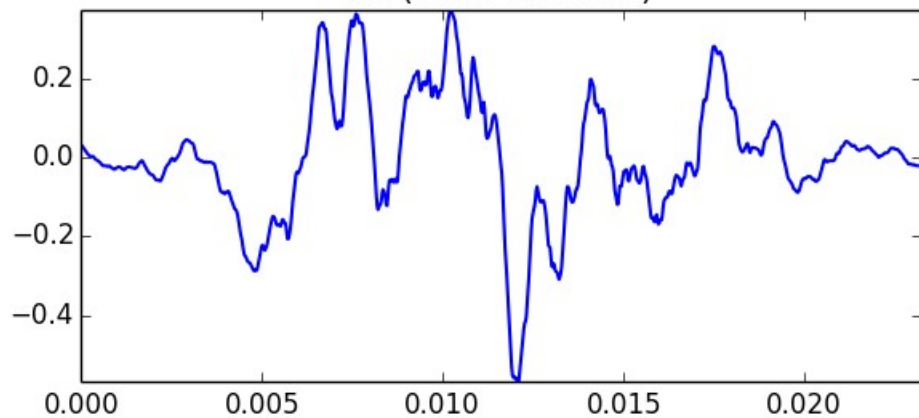


# Morphing with STFT

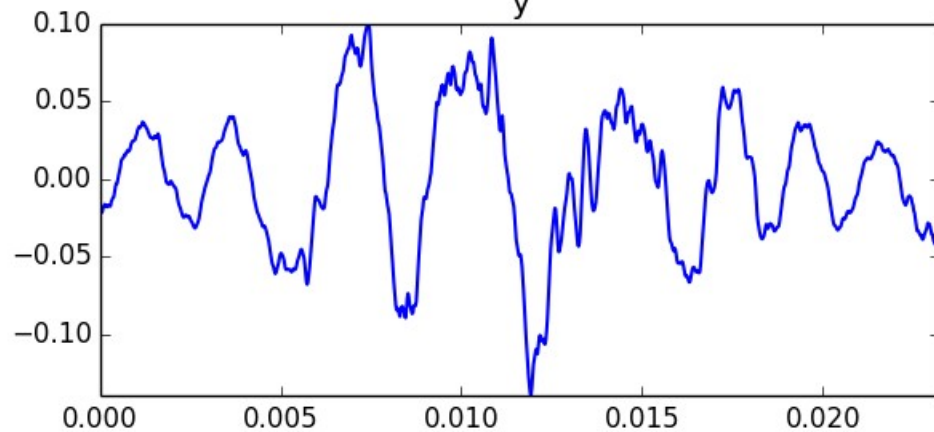


$$Y_l[k] = |X_{2_l}[k]| |X_{1_l}[k]| e^{j \angle X_{1_l}[k]}$$

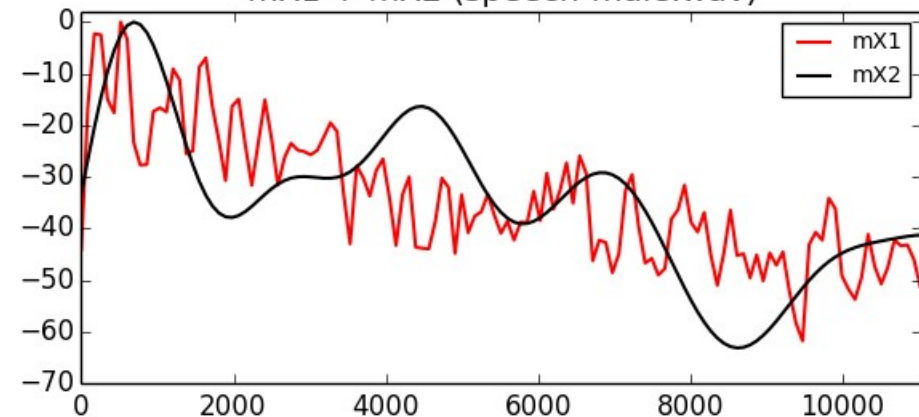
x1 (orchestra.wav)



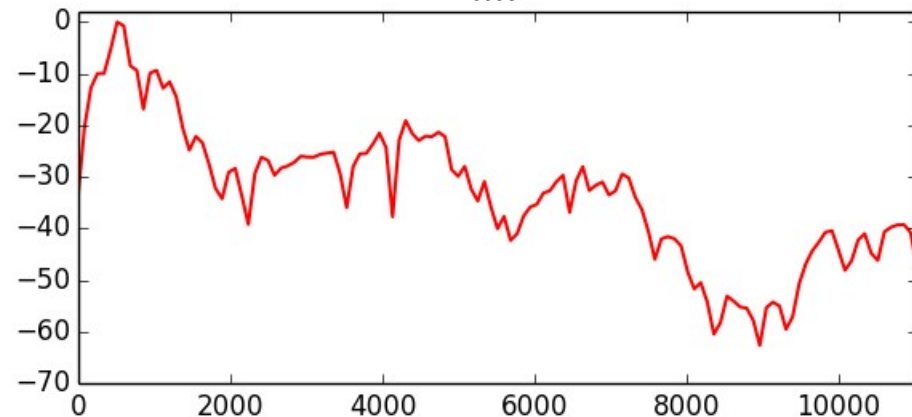
y



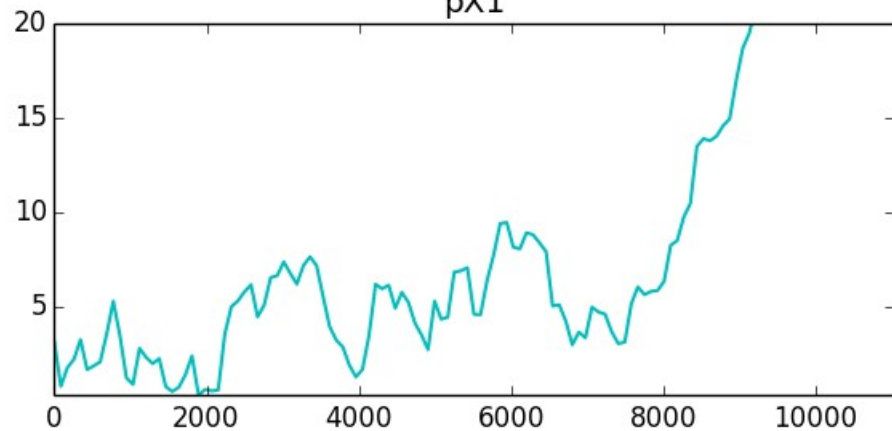
mX1 + mX2 (speech-male.wav)



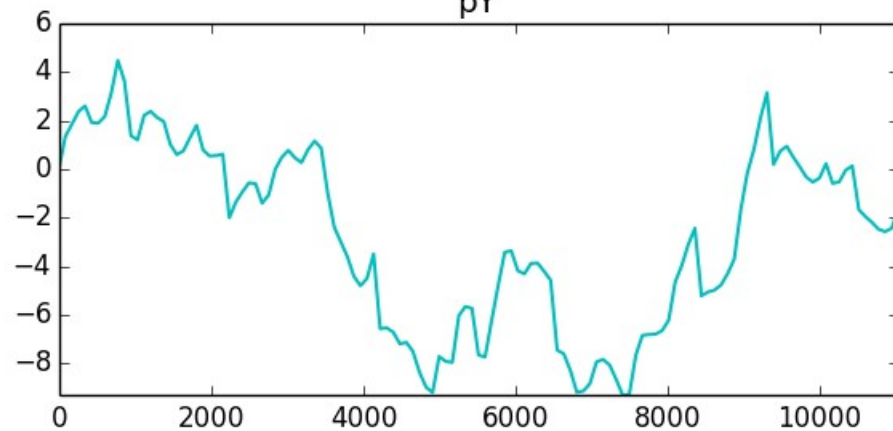
mY



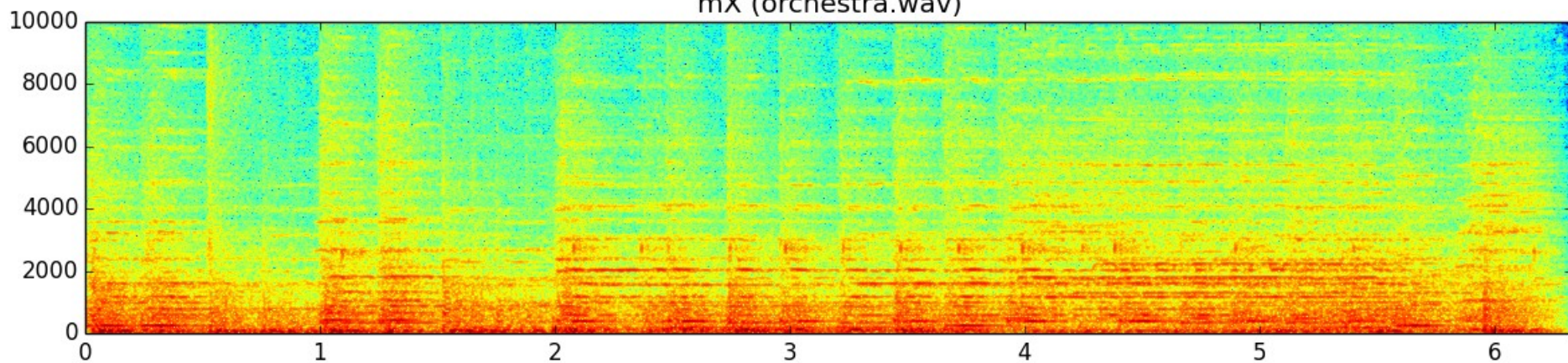
pX1



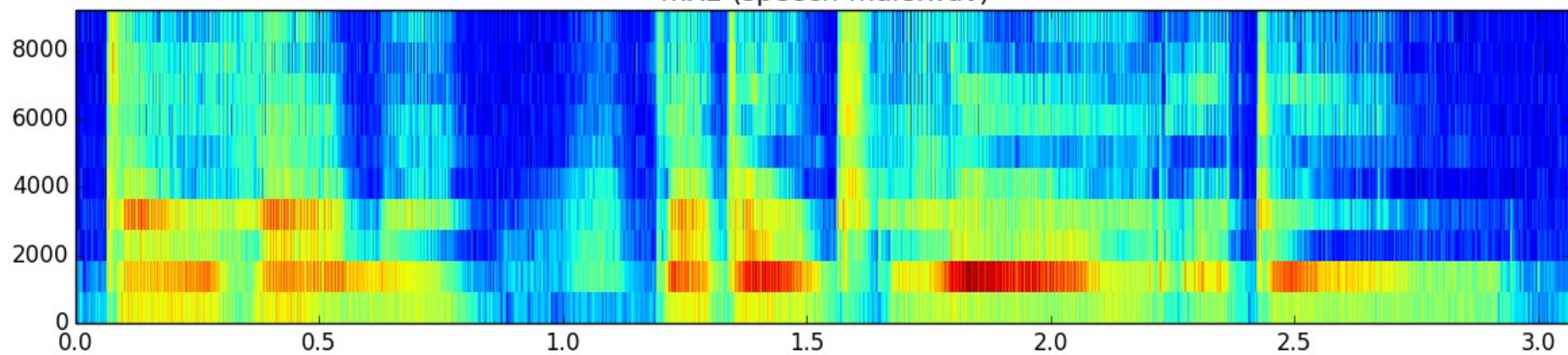
pY



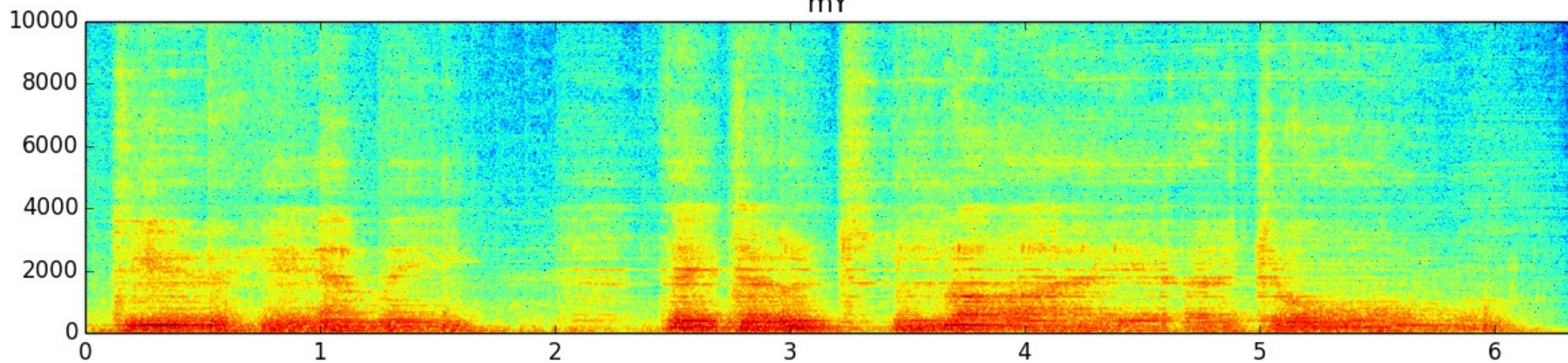
mX (orchestra.wav)



mX2 (speech-male.wav)

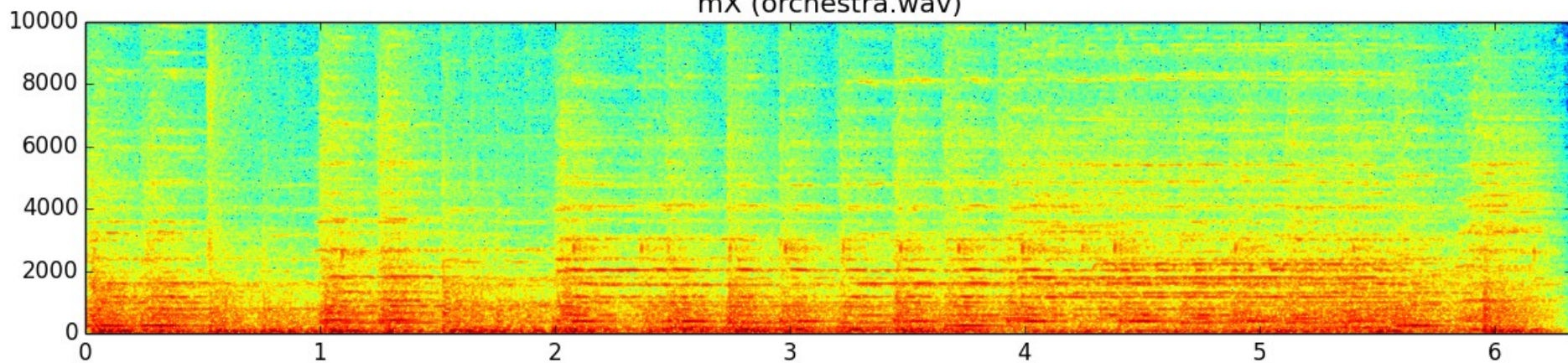


mY

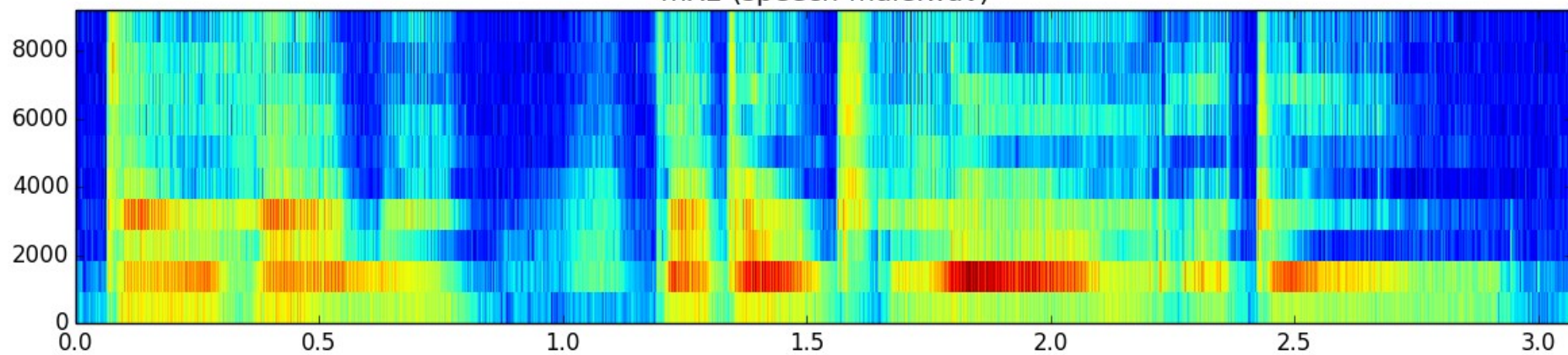




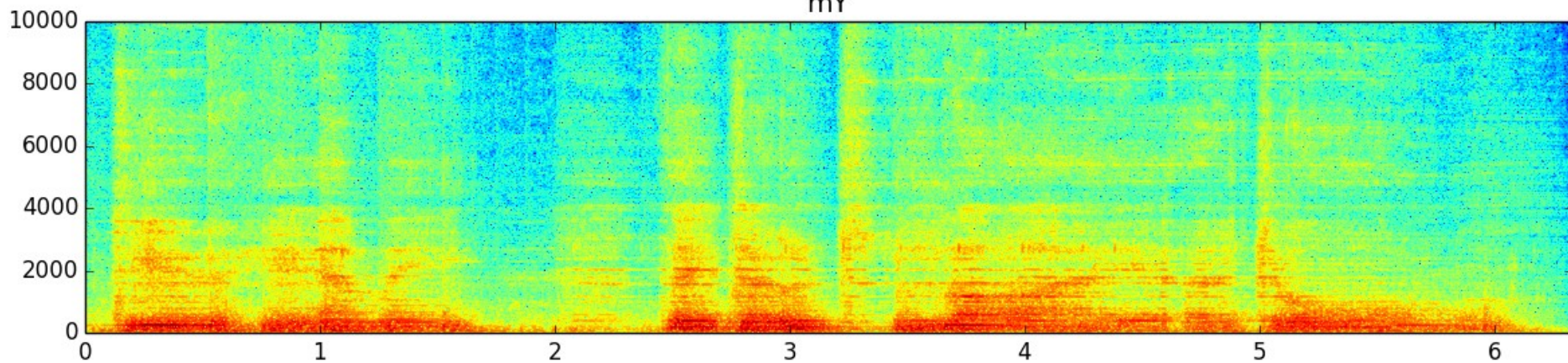
mX (orchestra.wav)



mX2 (speech-male.wav)

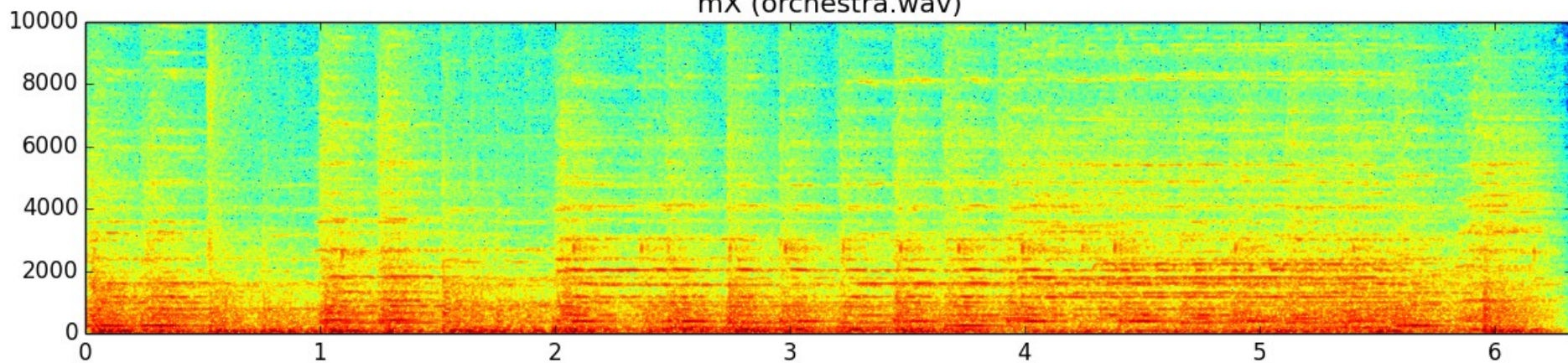


mY

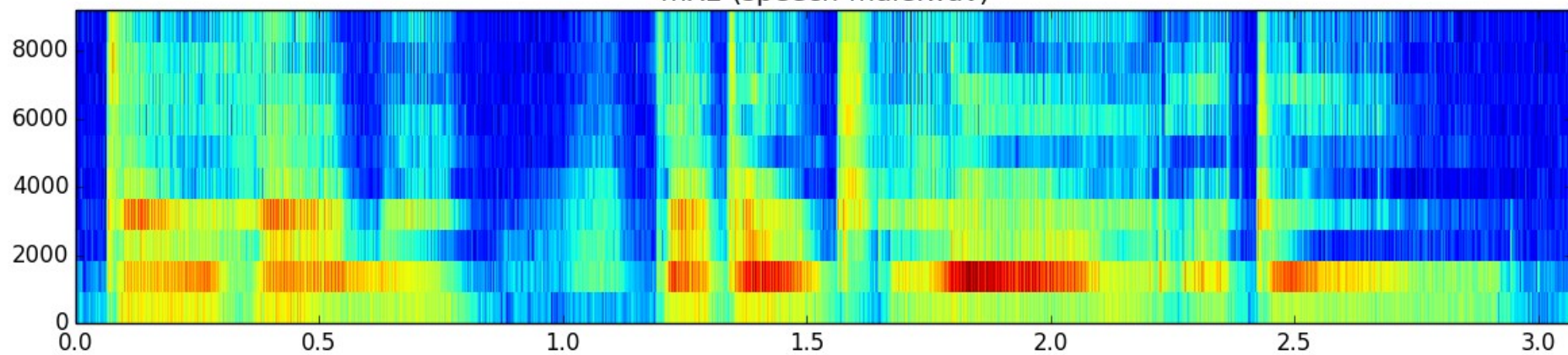




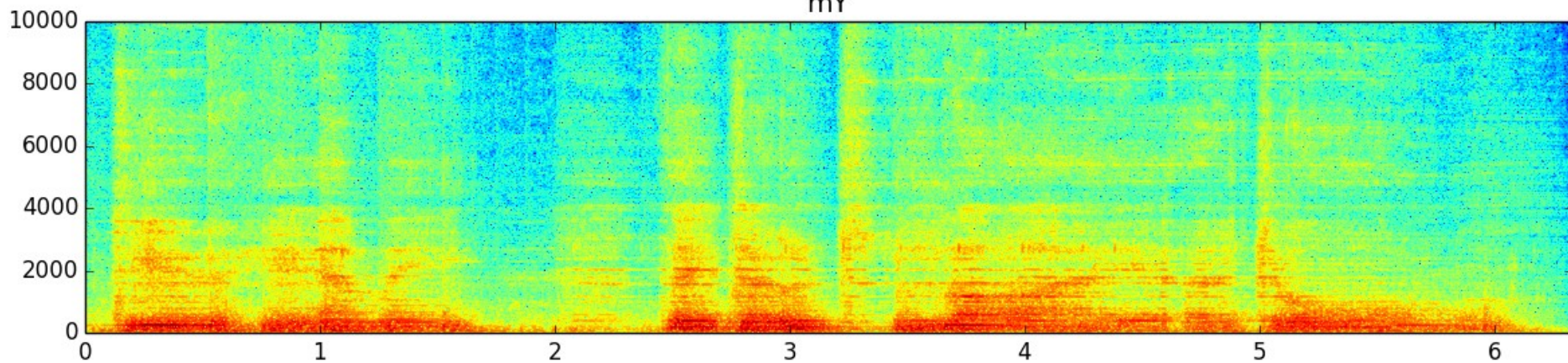
mX (orchestra.wav)



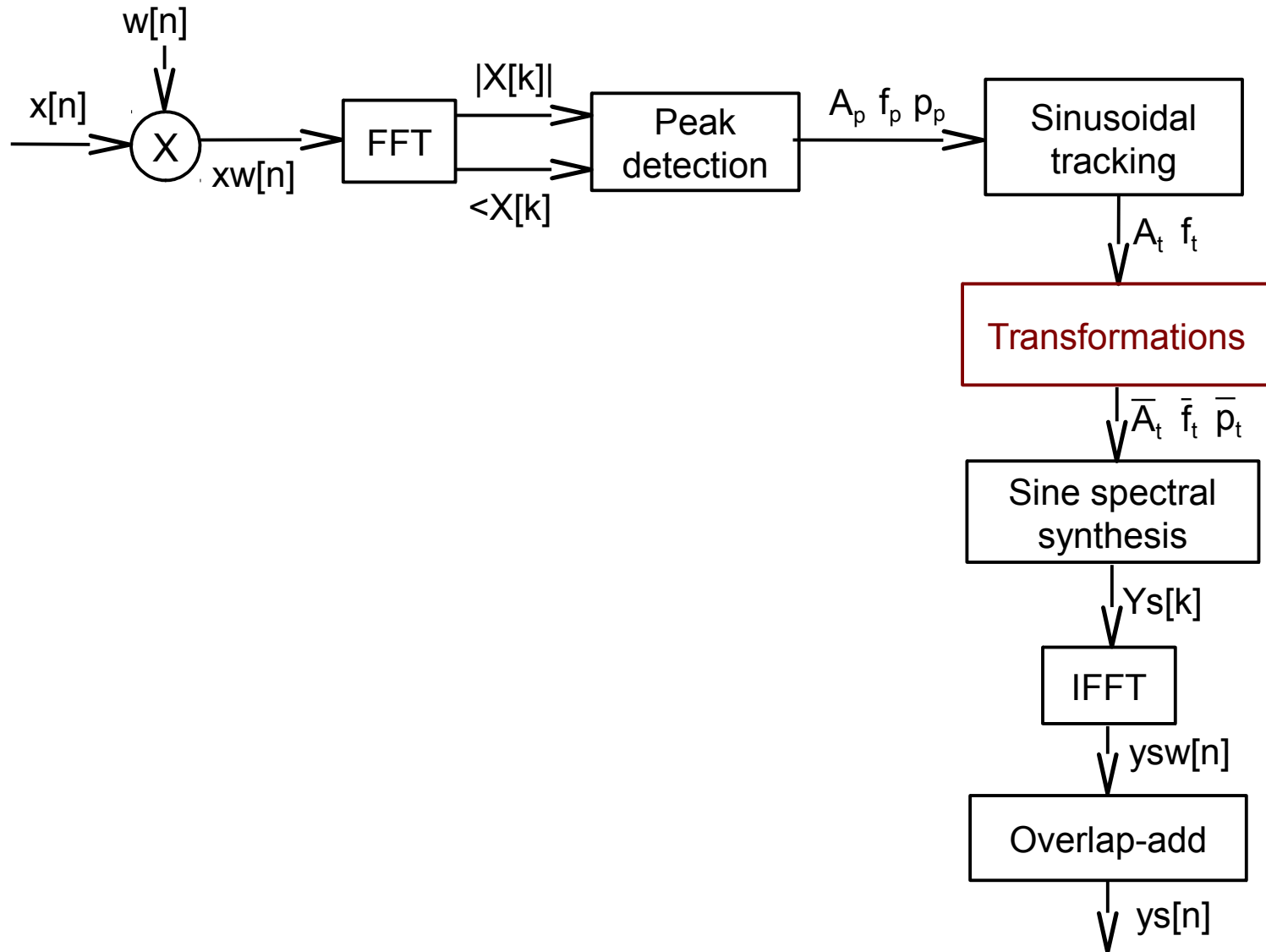
mX2 (speech-male.wav)



mY



# Sinusoidal model



# Scaling frequency, amplitude and time

$$\bar{f}_t[q] = sf_t[l] f_t[st_t[l]l]$$

$$\bar{A}_t[q] = sA_t[l] + A_t[st_t[l]l]$$

$$\bar{\varphi}_t[q] = \varphi_t[q-1] + f_t[q]$$

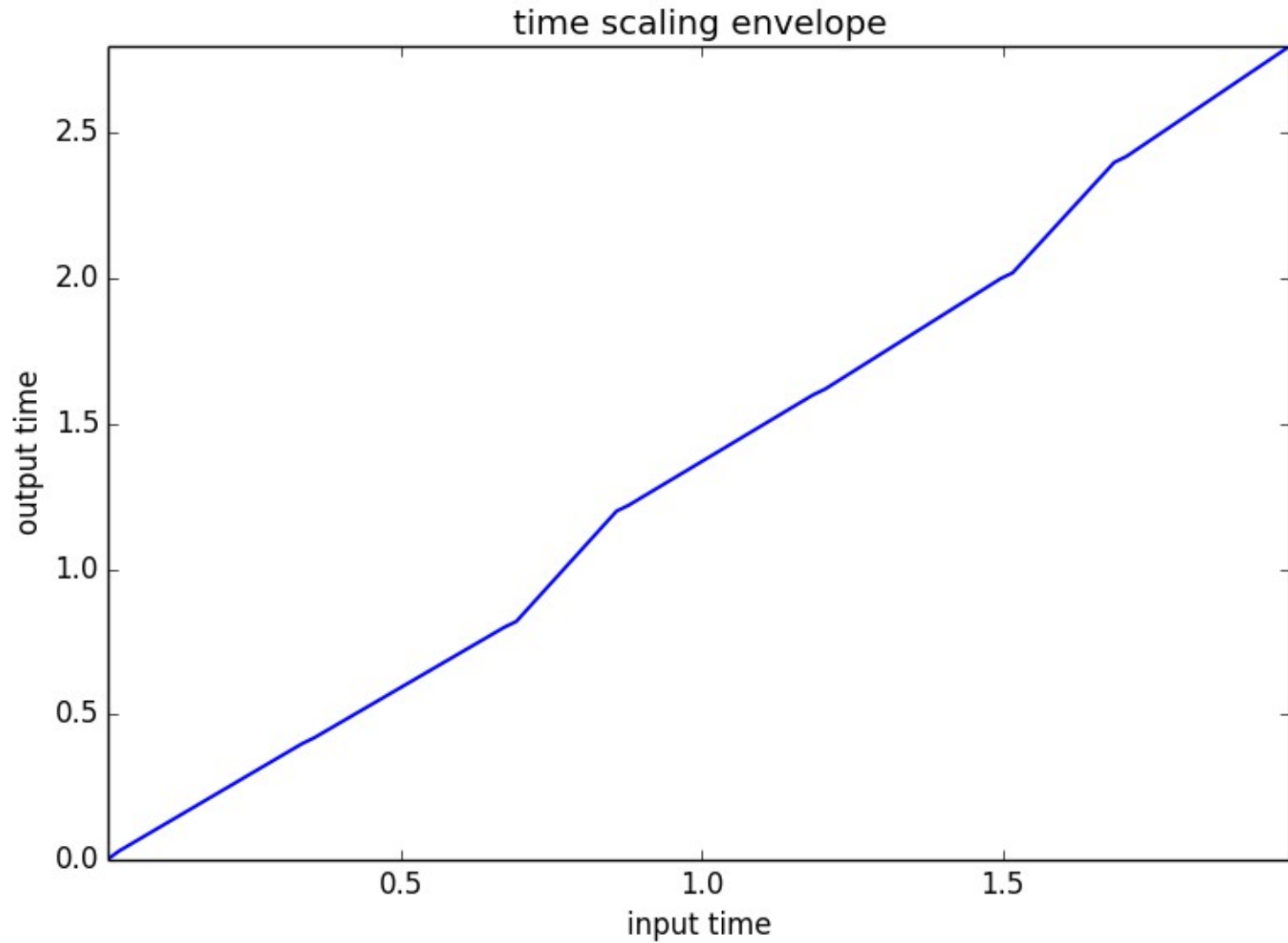
$q$ : output frame index ;  $l$ : input frame index ;  $t$ : sinusoidal track index

$f$ : input frequency in Hz ;  $A$ : input amplitude in dB

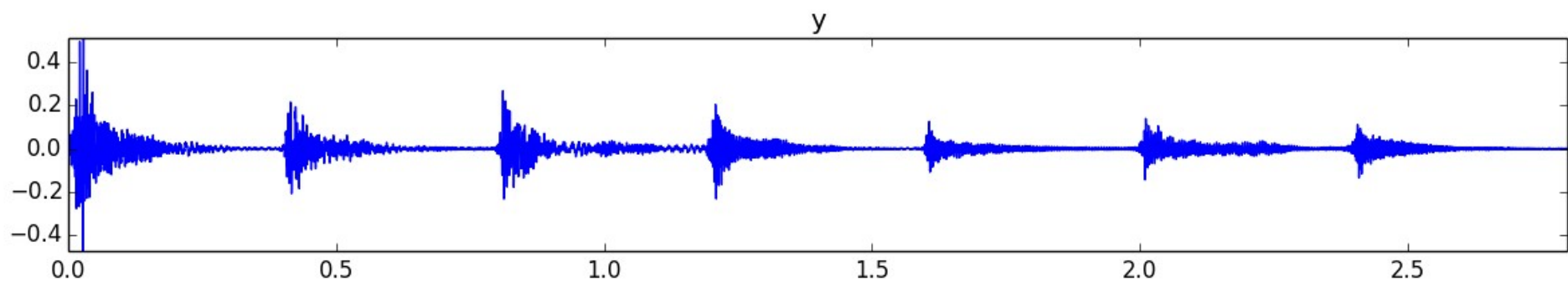
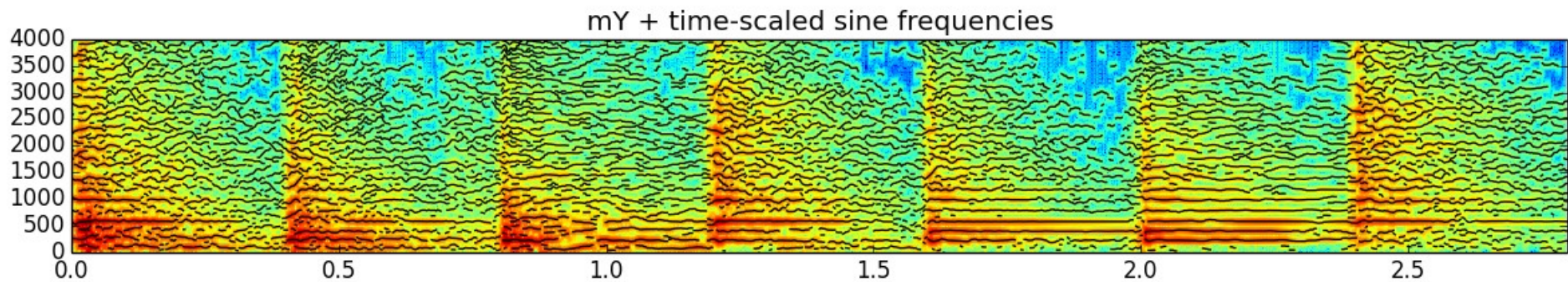
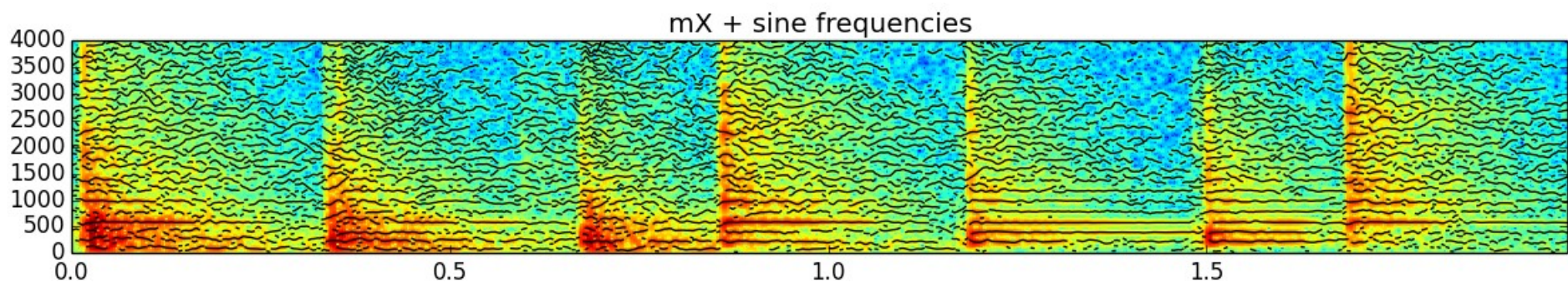
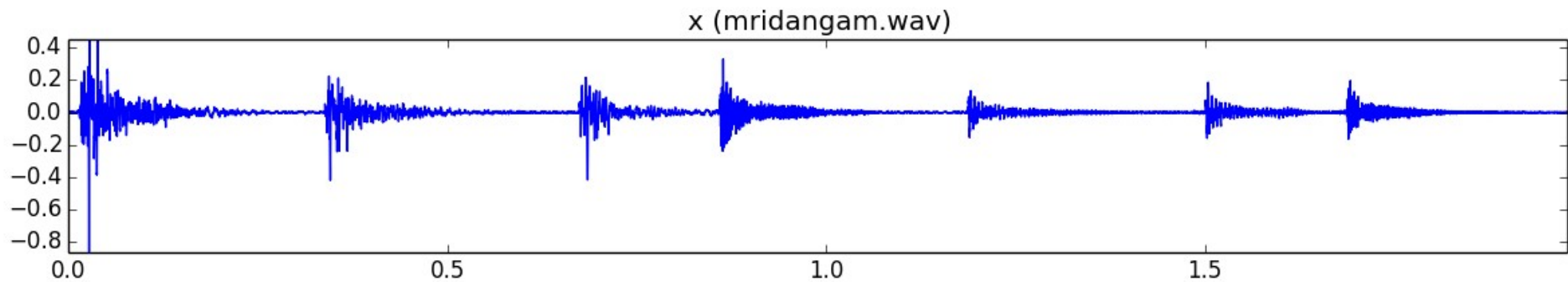
$sf$ : scaling frequency ;  $sA$ : scaling amplitude ;  $st$ : scaling time

$\bar{f}$ : output frequency ;  $\bar{A}$ : output amplitude ;  $\bar{\varphi}$ : output phase

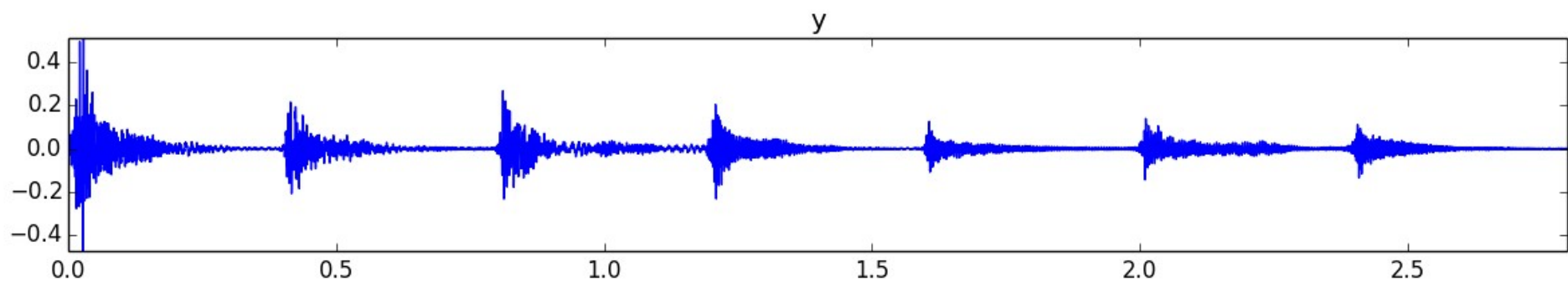
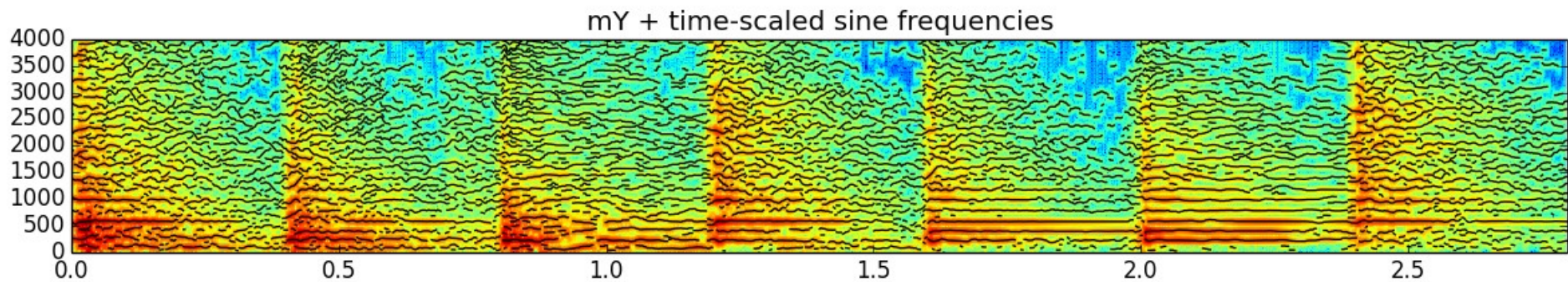
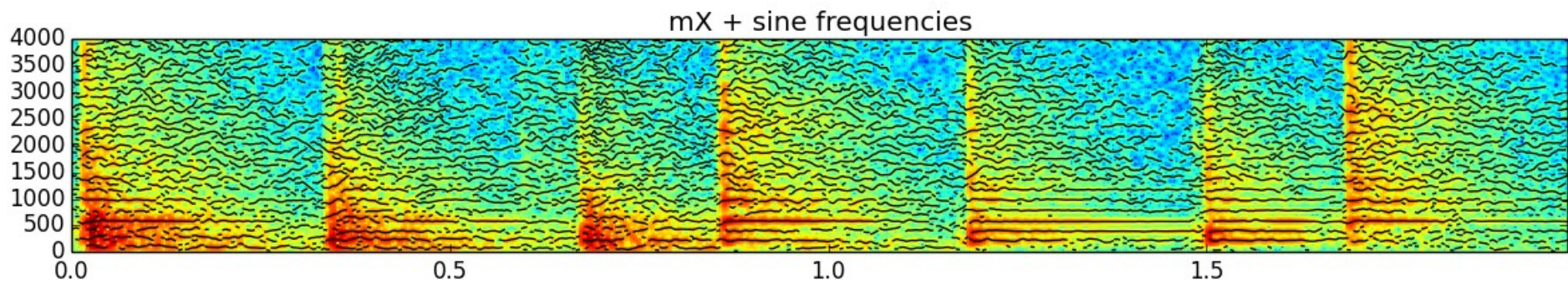
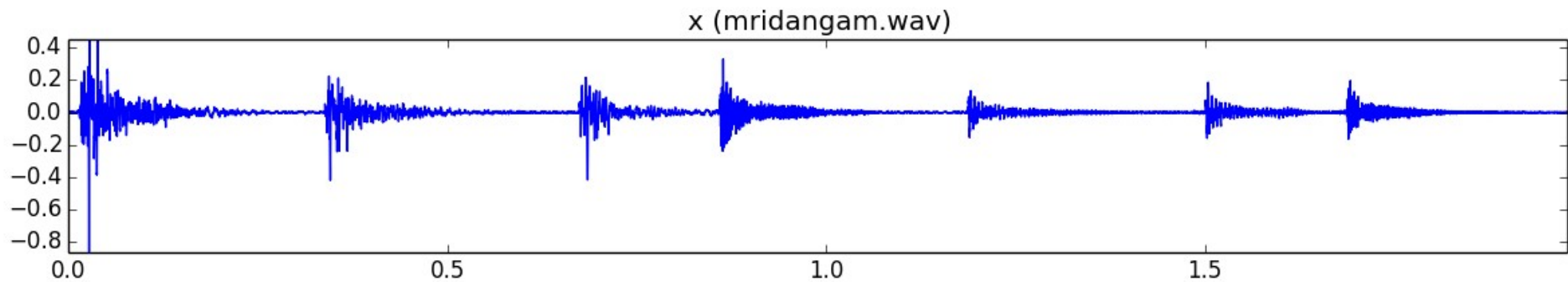
# Time scaling



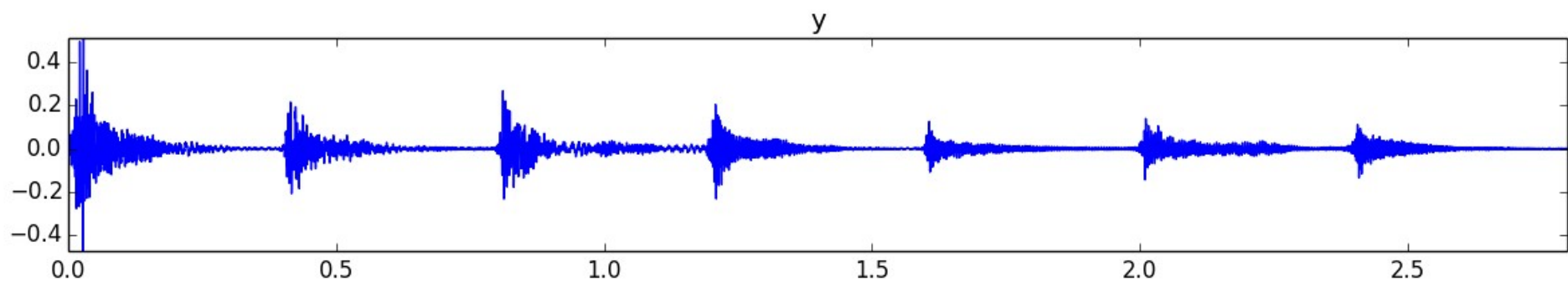
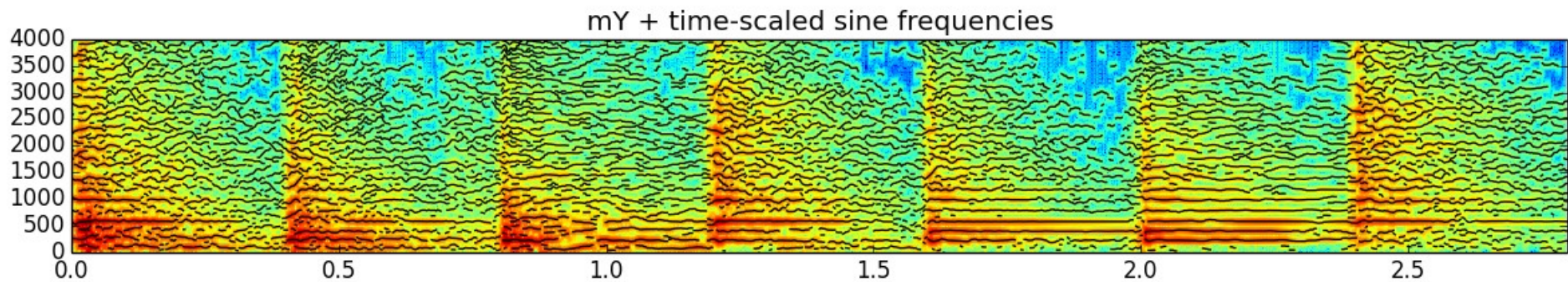
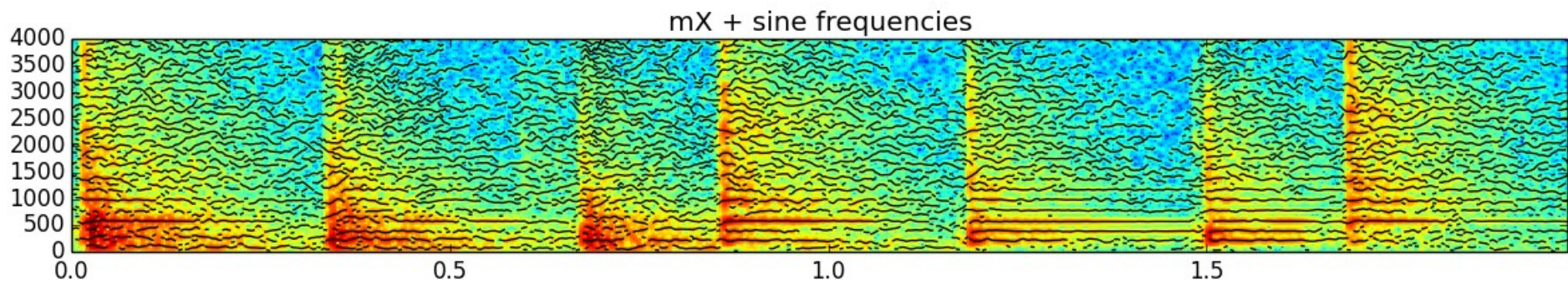
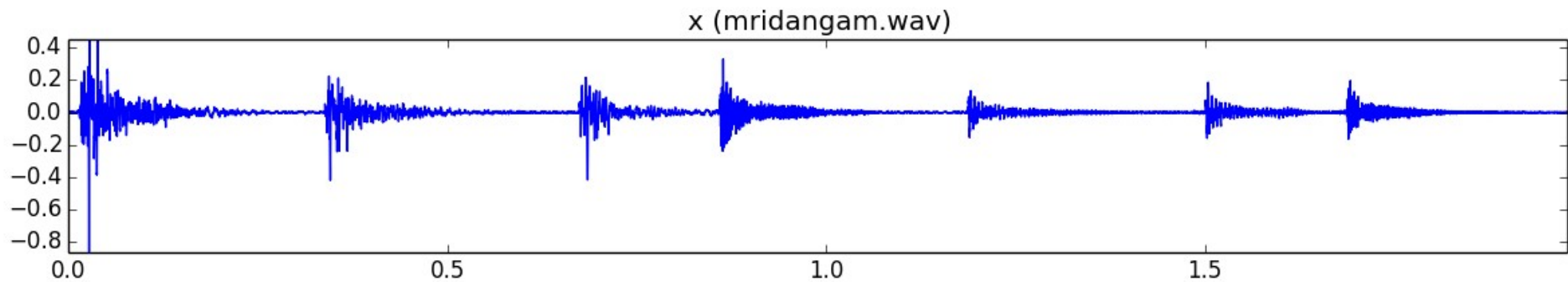




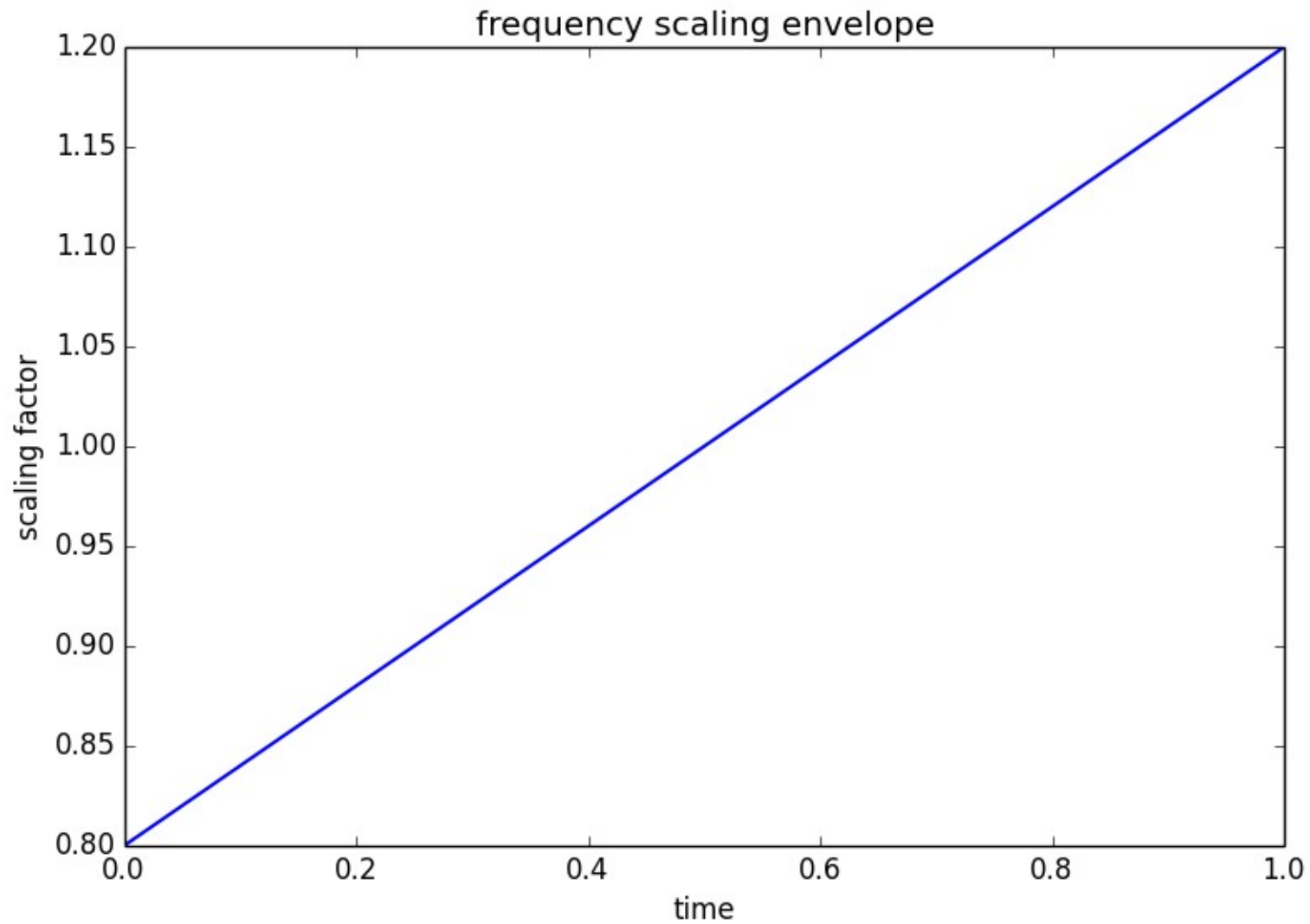




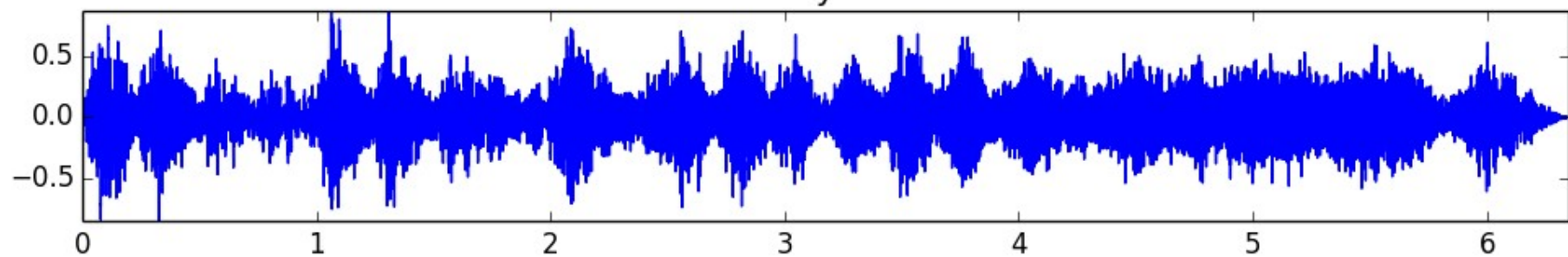
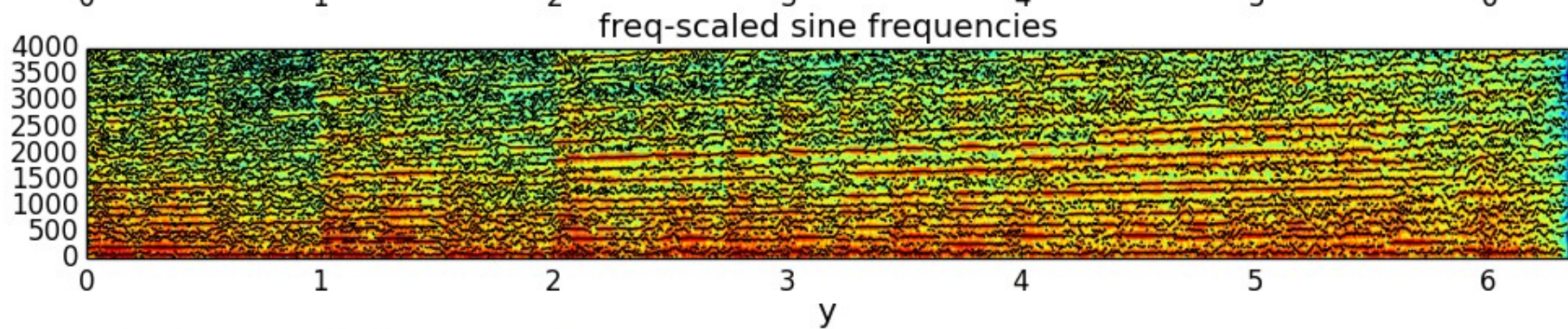
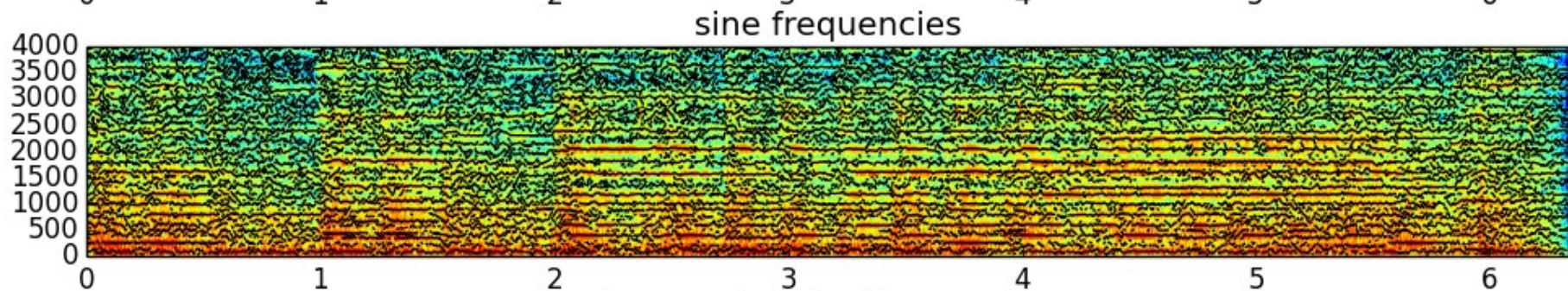
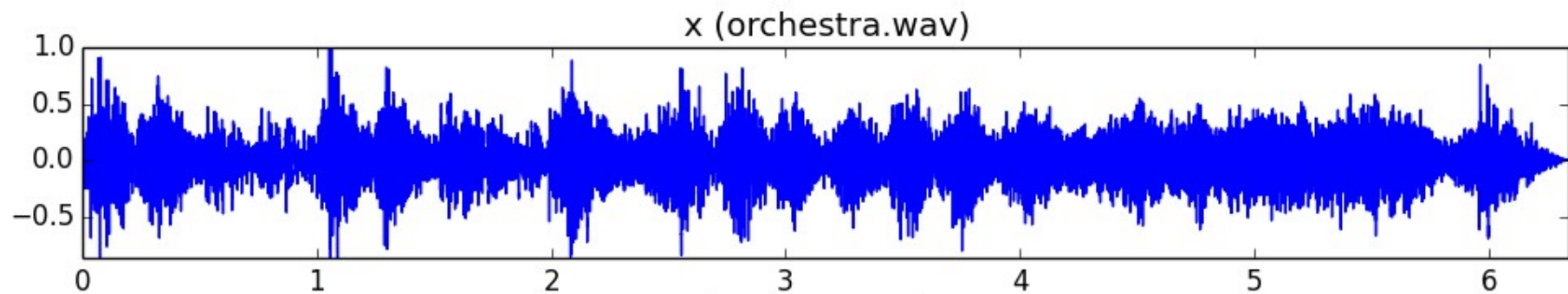




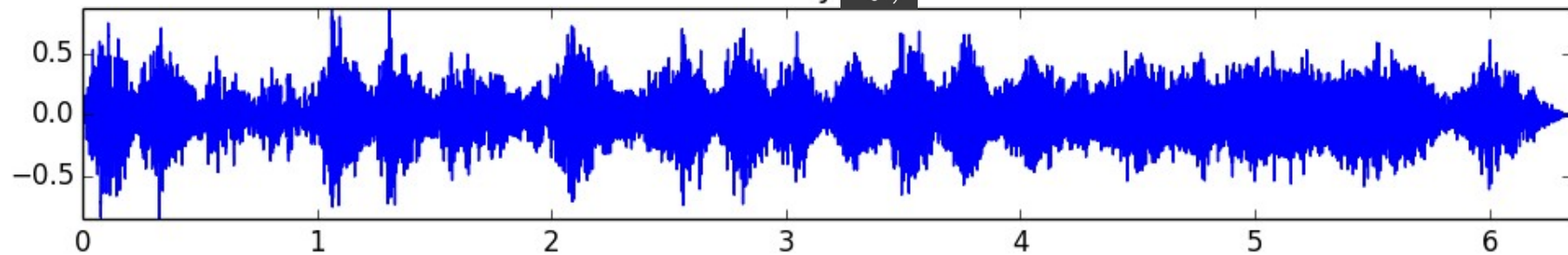
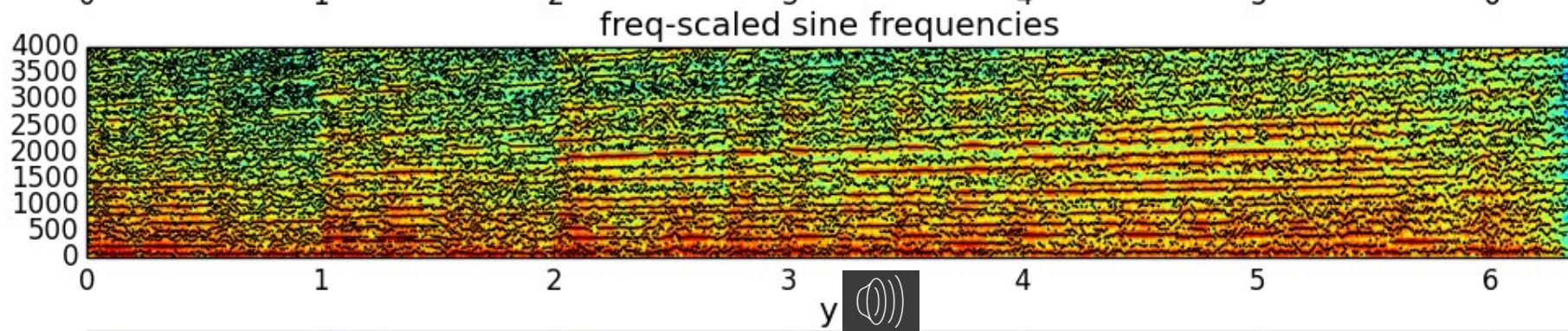
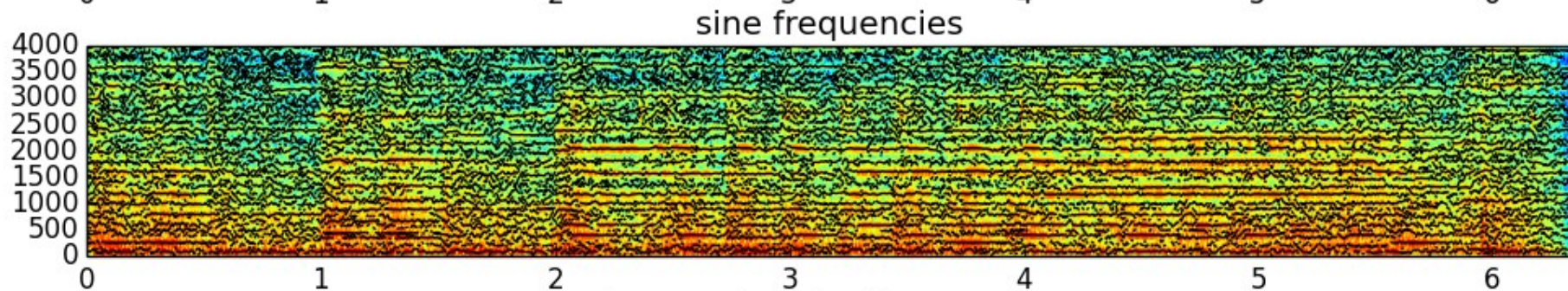
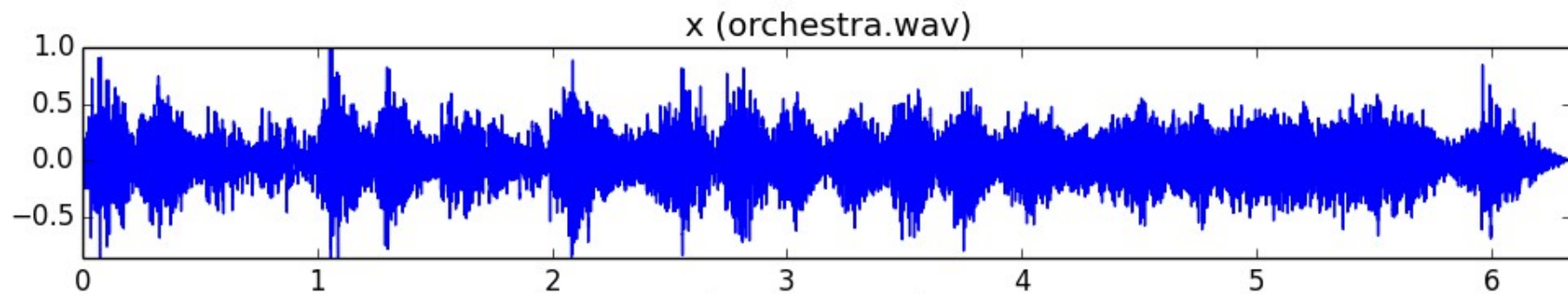
# Frequency scaling











# References

- More information on this topic from Wikipedia:
  - [http://en.wikipedia.org/wiki/Sound\\_effects](http://en.wikipedia.org/wiki/Sound_effects)
  - [http://en.wikipedia.org/wiki/Equalization\\_filter](http://en.wikipedia.org/wiki/Equalization_filter)
  - [http://en.wikipedia.org/wiki/Audio\\_timescale-pitch\\_modification](http://en.wikipedia.org/wiki/Audio_timescale-pitch_modification)
- Sounds: <http://www.freesound.org/people/xserra/packs/13038/>
- The slides and code are released using the CC Attribution-Noncommercial-Share Alike license or the Affero GPL license and available from <https://github.com/MTG/sms-tools>

# 8T1: Spectral-based sound transformations (1 of 2)

***Xavier Serra***

Universitat Pompeu Fabra, Barcelona

&

Stanford University