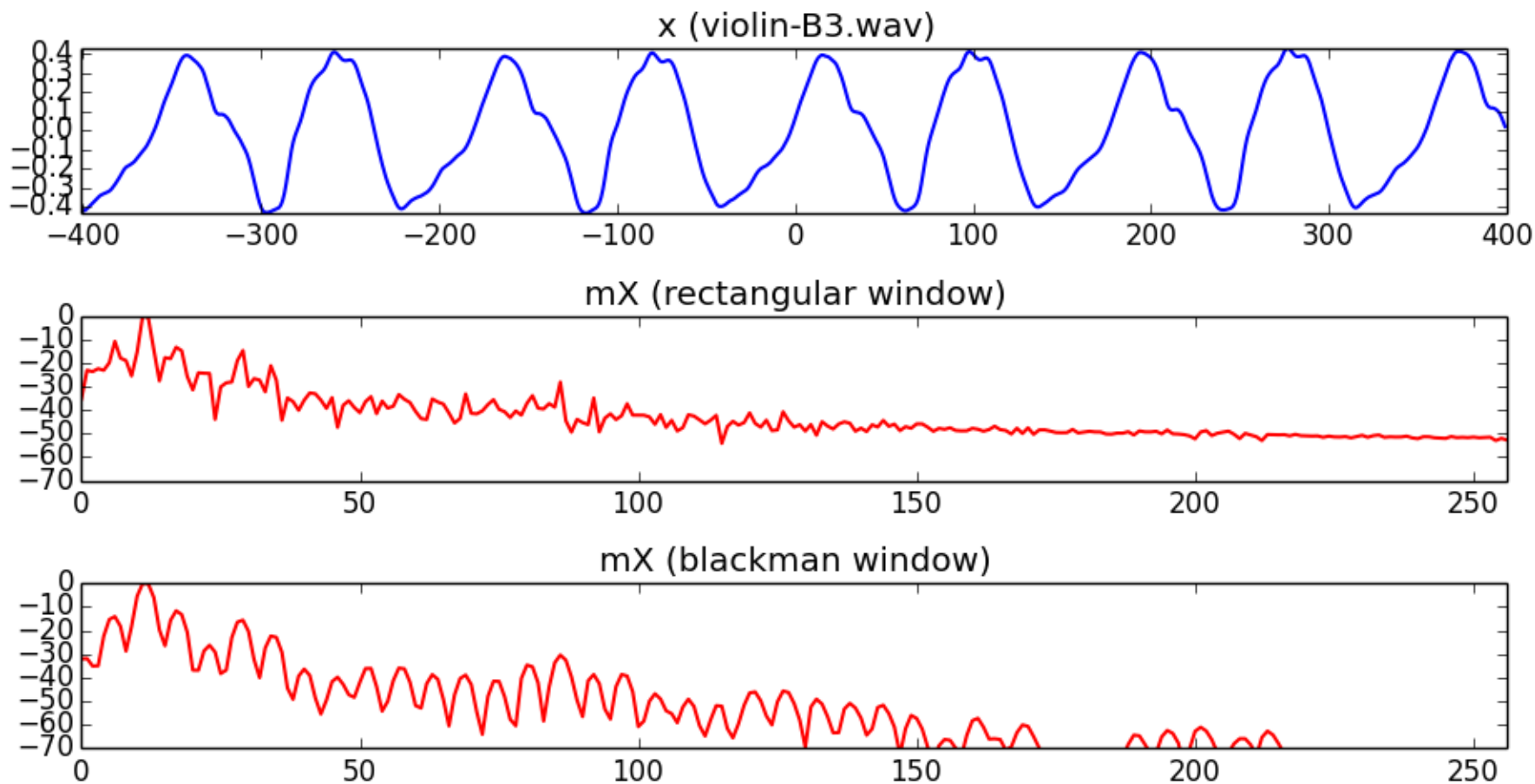# 4T2: The Short-Time Fourier Transform (2 of 2)

Xavier Serra

Universitat Pompeu Fabra, Barcelona

Index
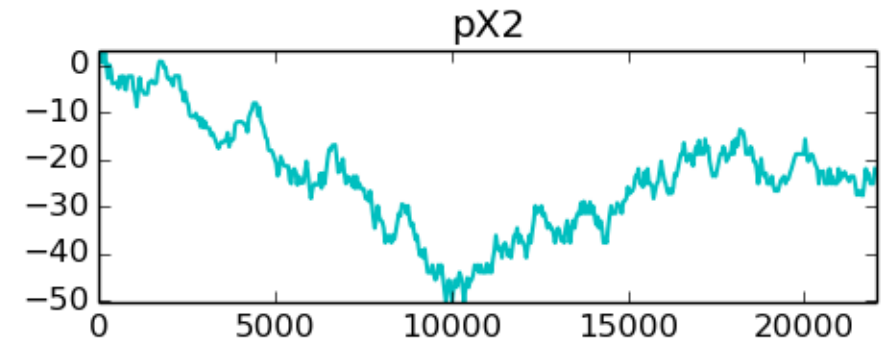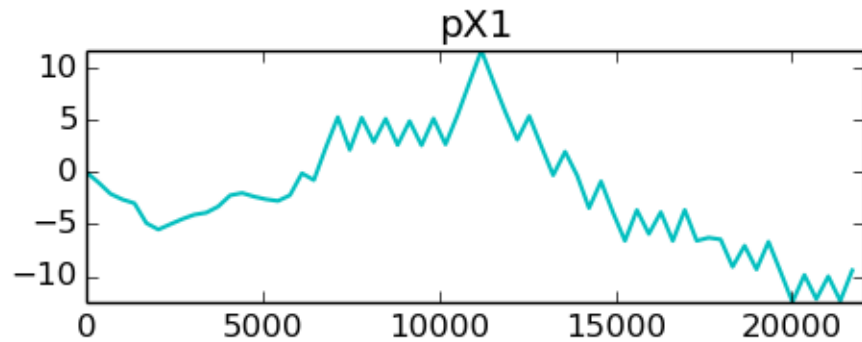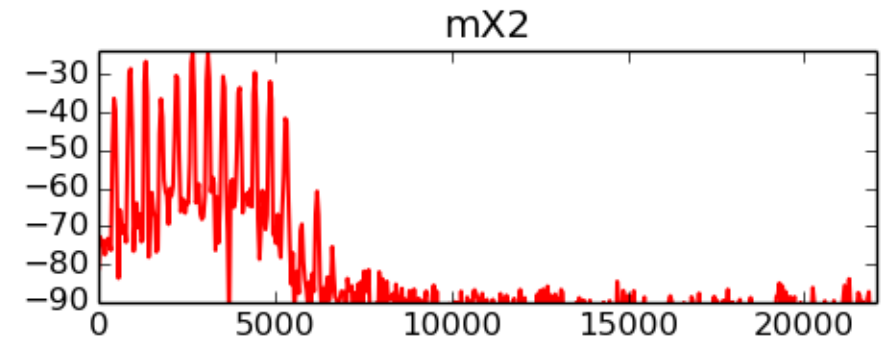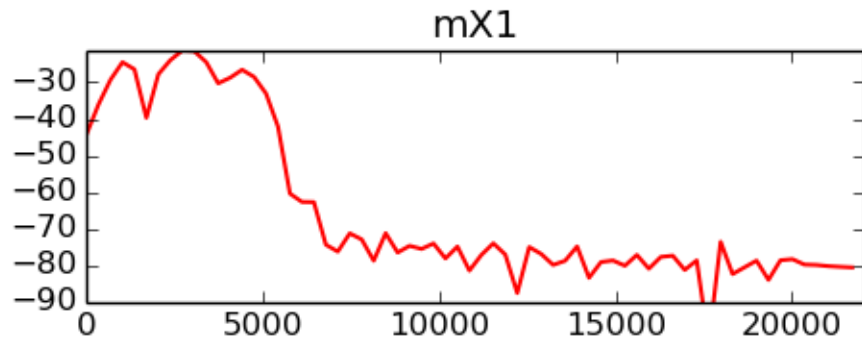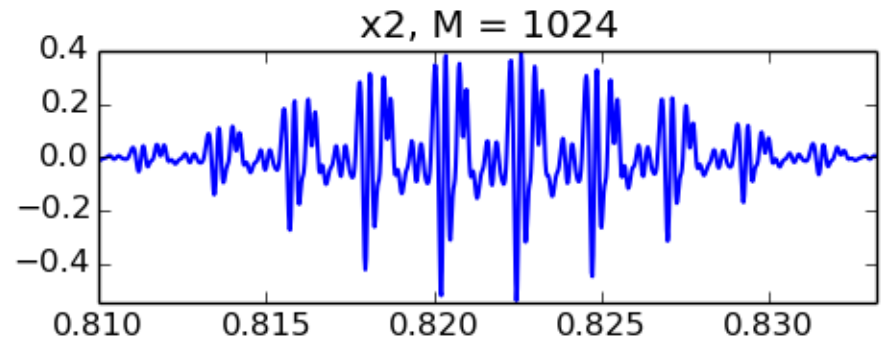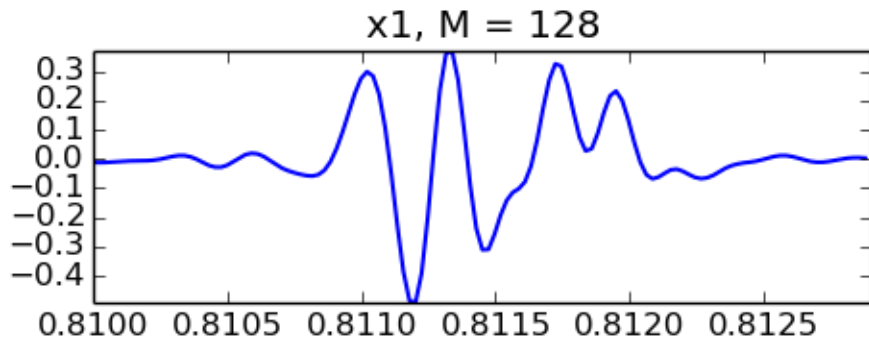
- STFT and analysis window

- Window size

- FFT size

- Hop size

- Time-frequency compromise

- Inverse STFT

- STFT system

STFT and analysis window

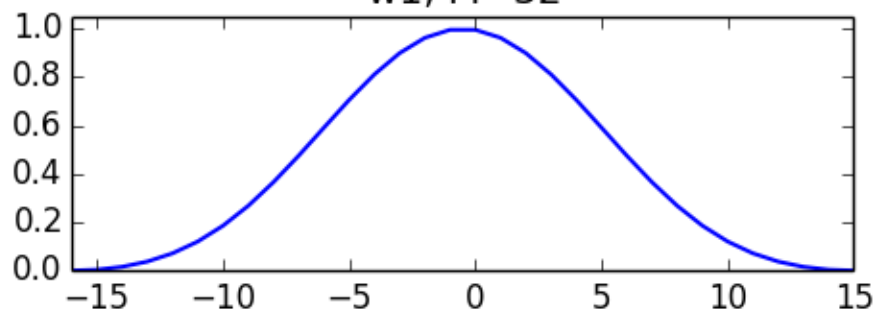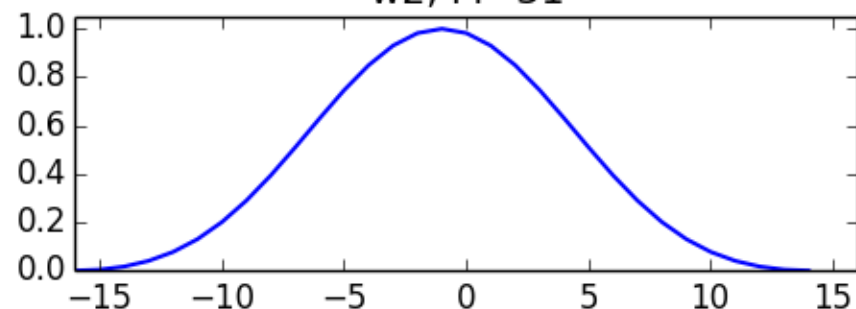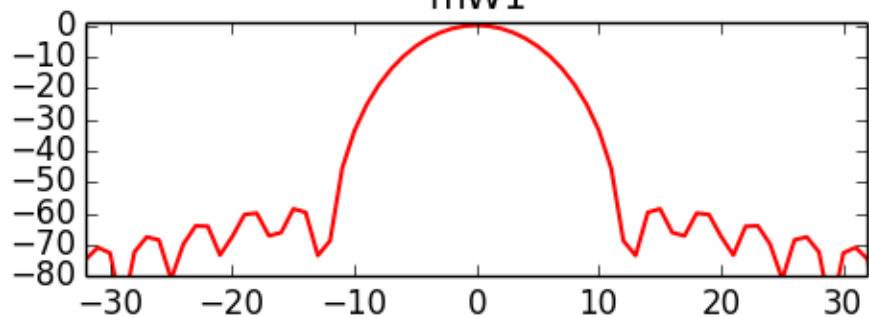$$X_l[k]= \sum_{n=-N/2}^{N/2-1} w[n]x[n+lH]e^{-j2\pi kn/N} \qquad l=0,1,\ldots,$$

x (violin-B3.wav)

mX (rectangular window)

mX (blackman window)

Window size

Even-odd size window

FFT size

x (oboe-A4.wav), M = 512

mX, N = 512

mX, N = 2048

Hop size $A_w[n] = \sum_{l=0}^{L-1} w[n-lH] = c$

Blackman, M=201, H=100

Blackman, M=201, H=50
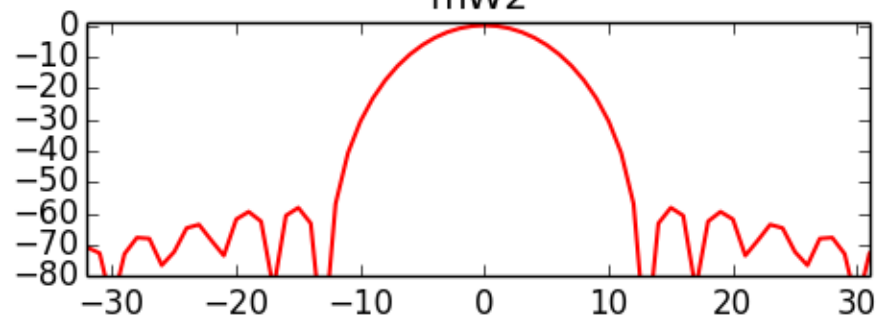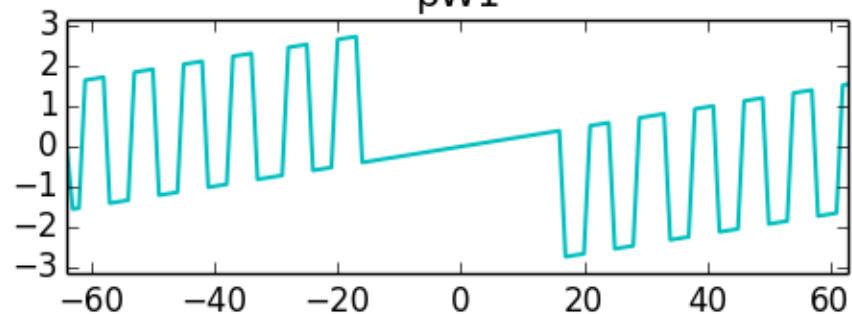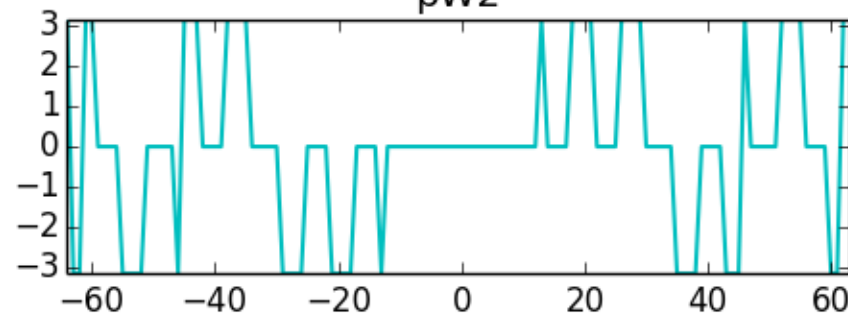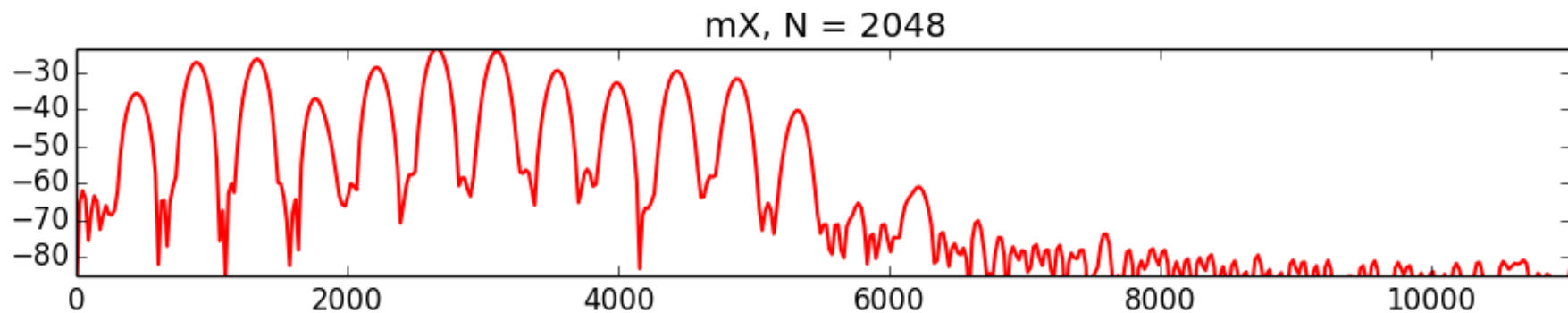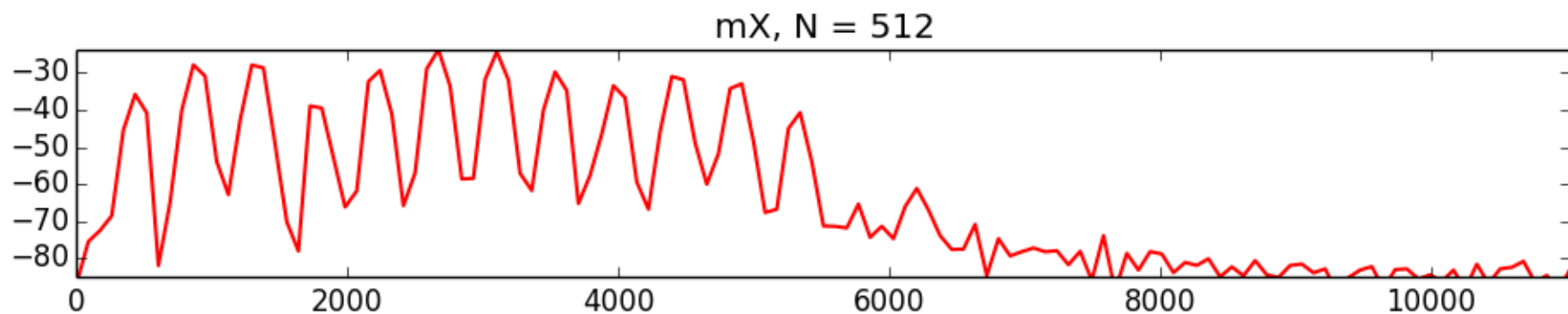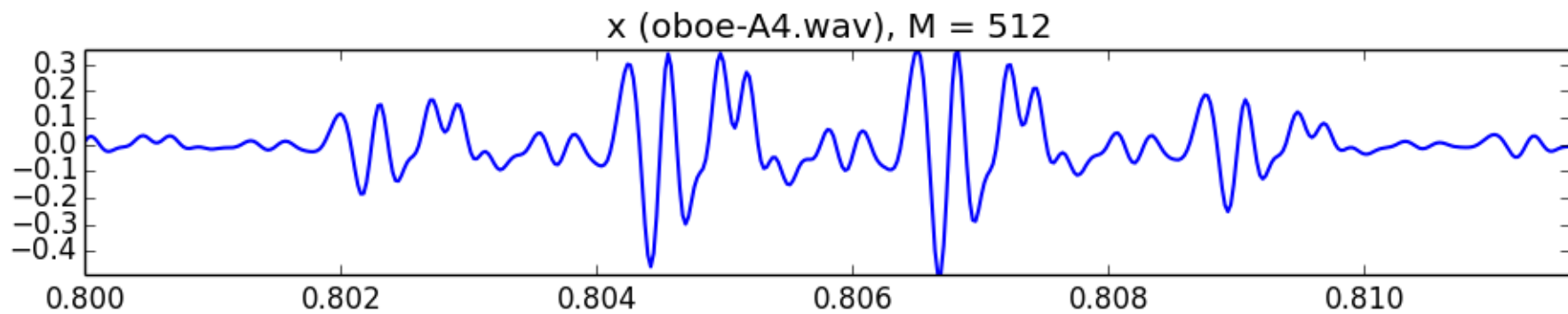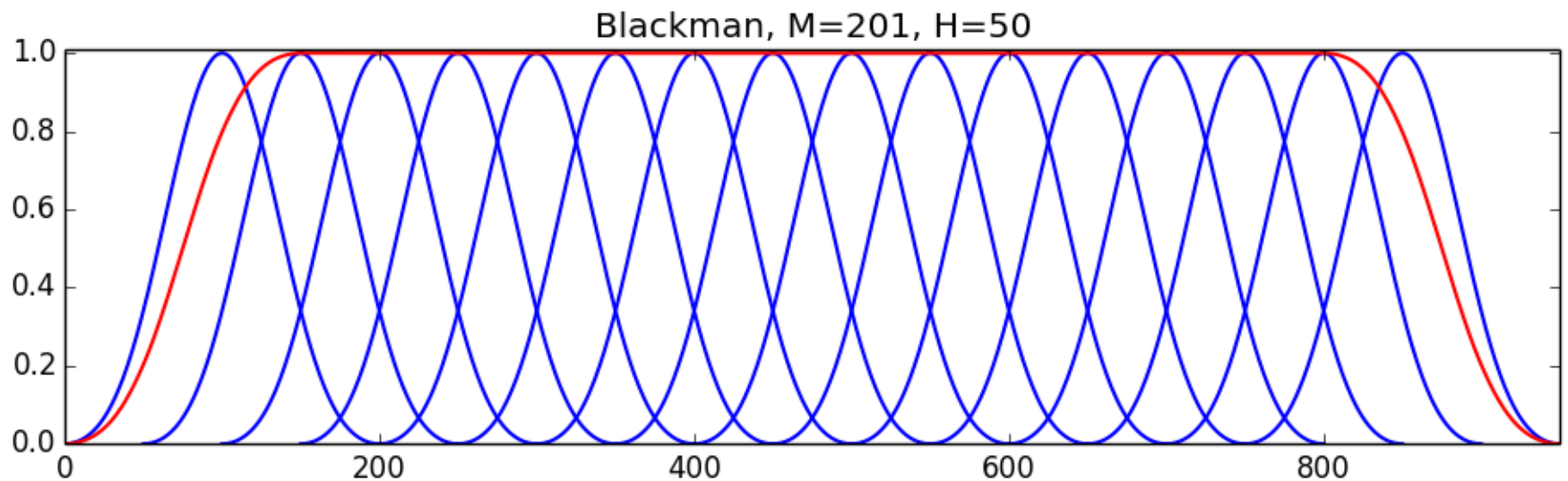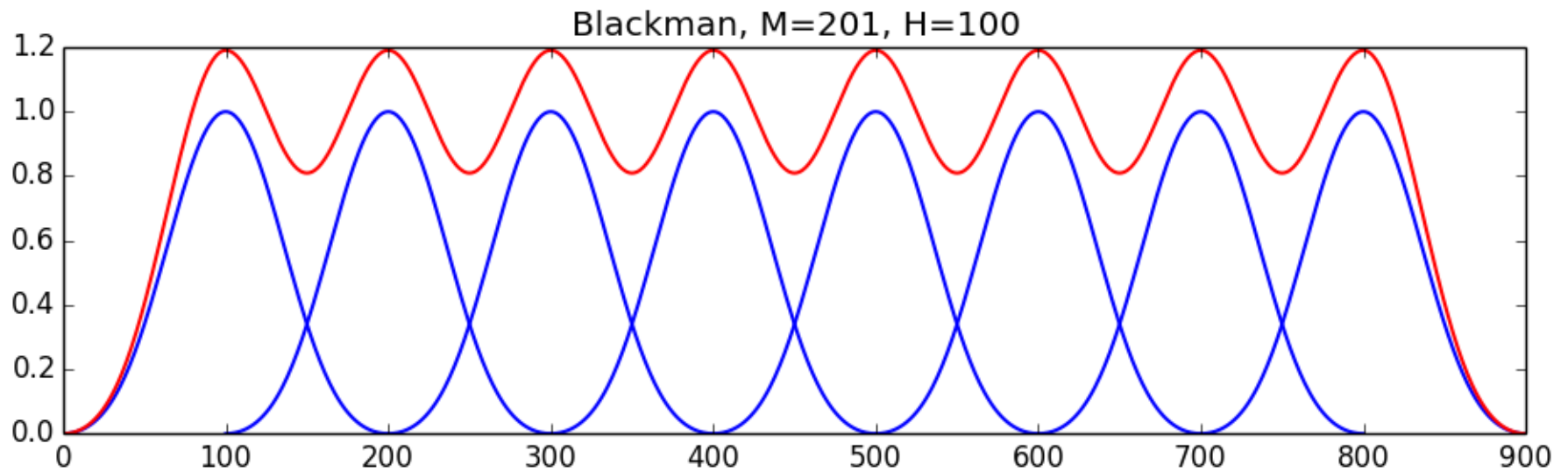
Time-frequency compromise

mX (piano.wav), M=256, N=256, H=128
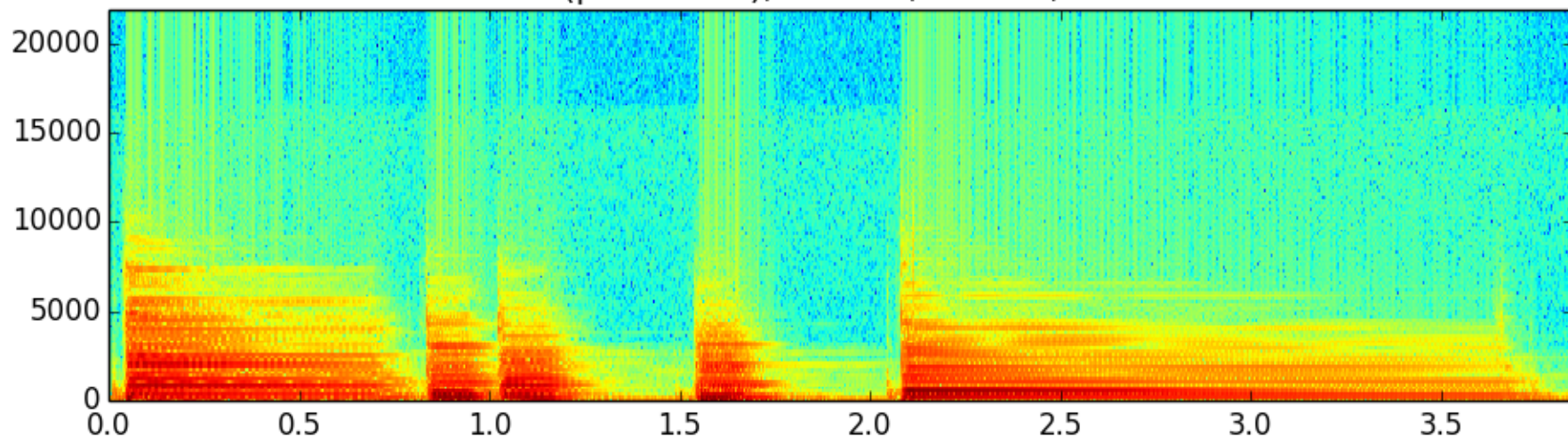
mX (piano.wav), M=1024, N=1024, H=128

Amplitude and phase spectrogram

mX (piano.wav), M=1001, N=1024, H=256
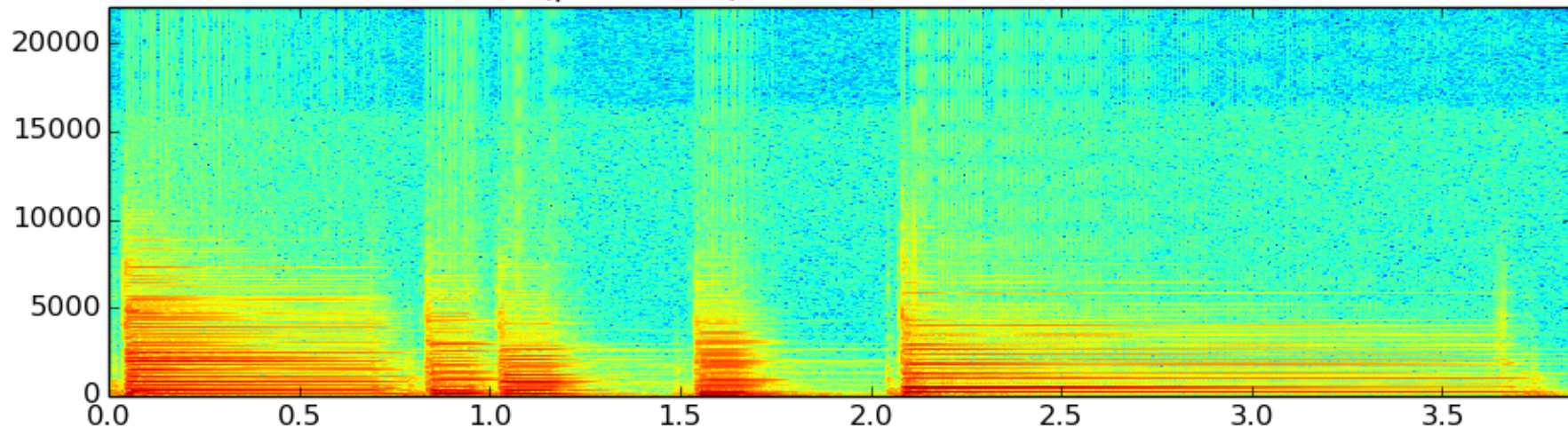
pX derivative (piano.wav), M=1001, N=1024, H=256
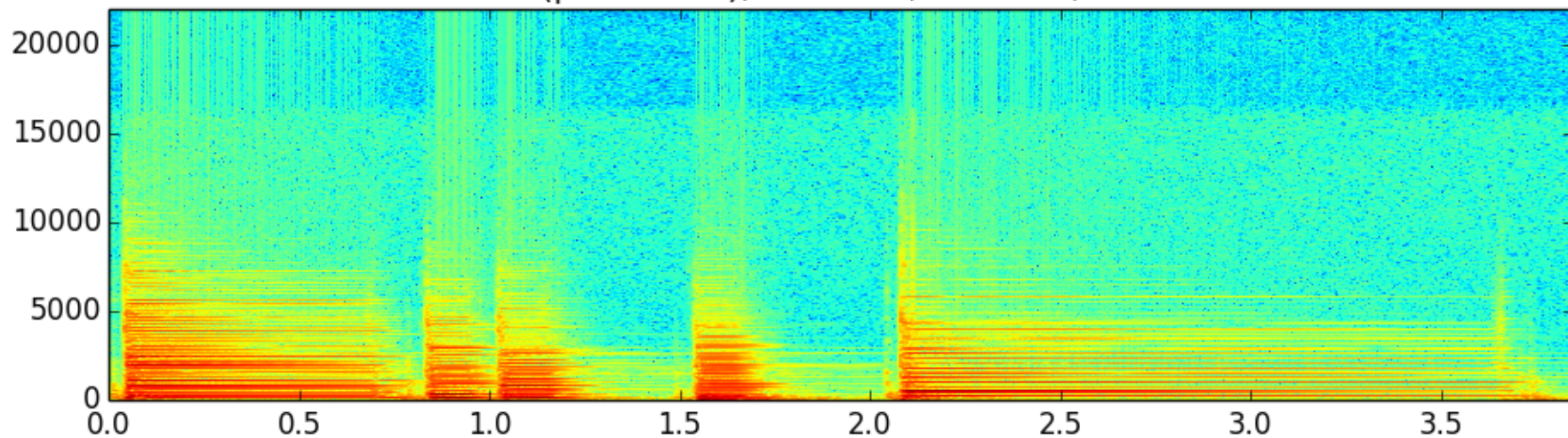
Inverse STFT

$$y[n] = \sum_{l=0}^{L-1} Shift_{lH,n} [\frac{1}{N} \sum_{k=-N/2}^{N/2-1} X_l[k] e^{j2\pi kn/N}]$$

each output frame is:

$$yw_l[n] = x(n+lH) w[n]$$

and the output sound is:

$$y[n] = \sum_{l=0}^{L-1} yw_l[n] = x[n] \sum_{l=0}^{L-1} w[n-lH]$$

$$yw_l[n] = w[n]x[n+lH] \qquad l = 0,1,\ldots,$$

STFT system
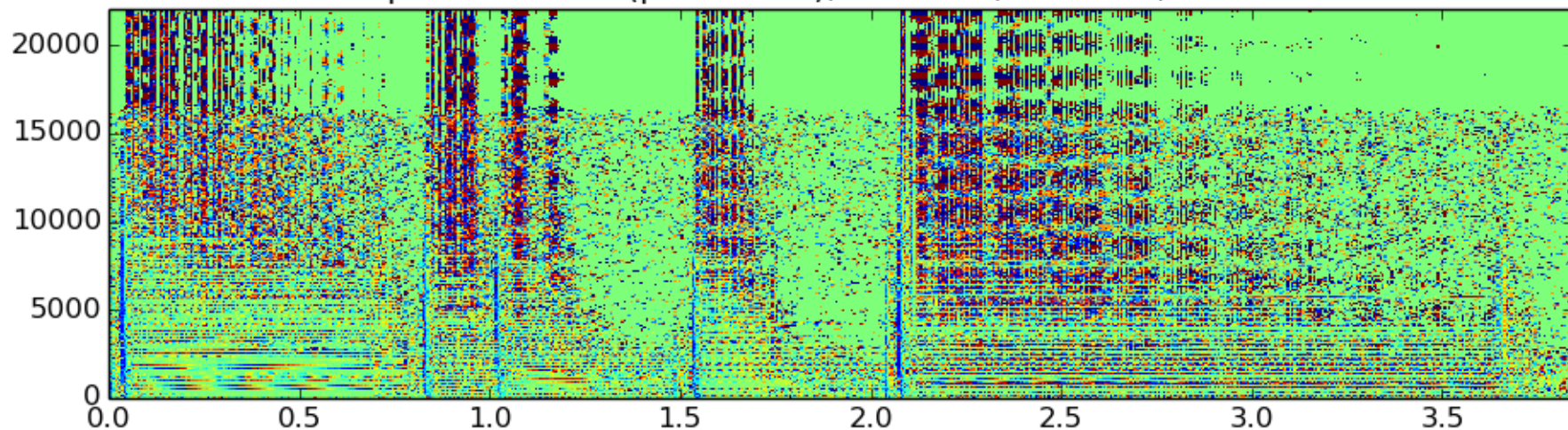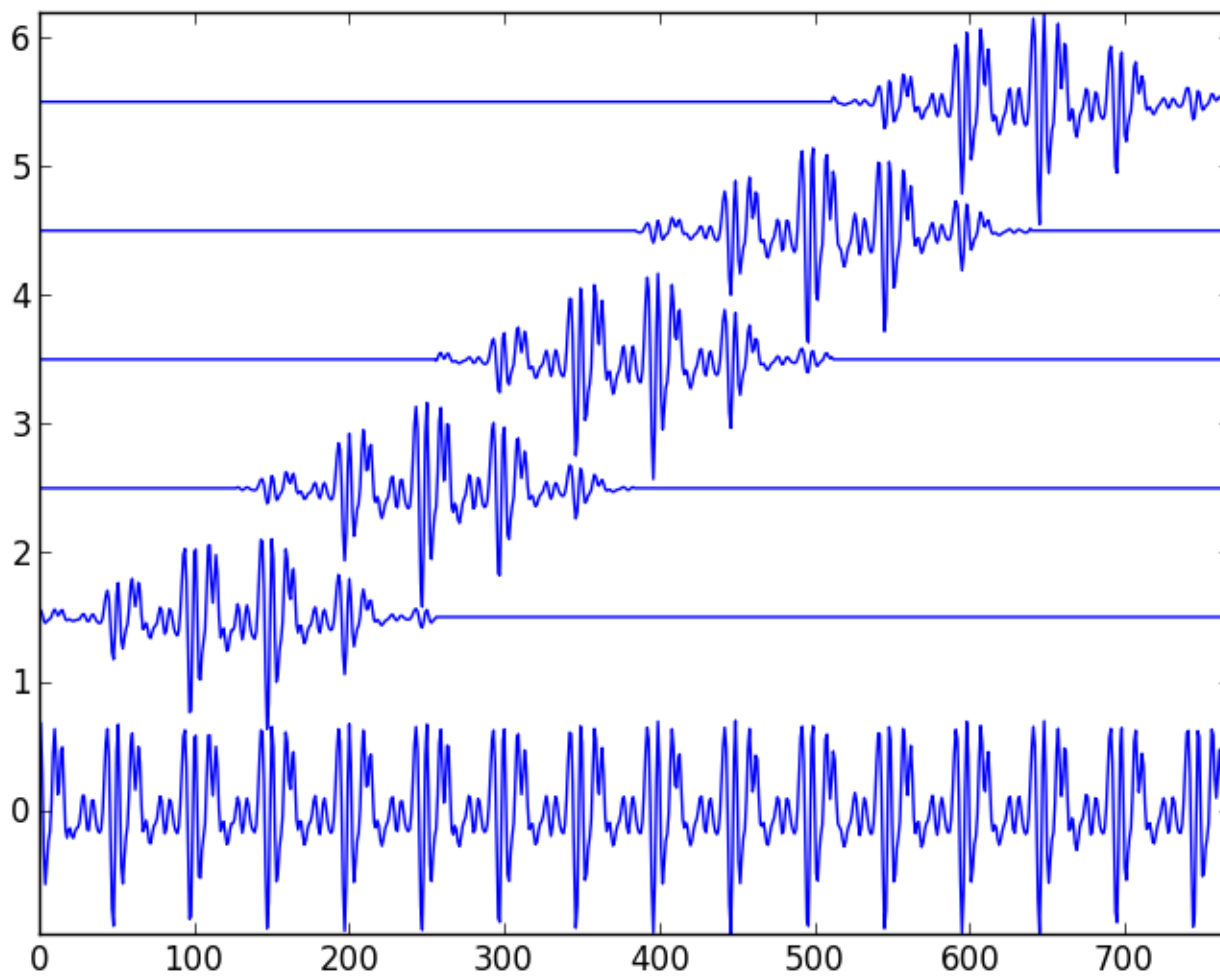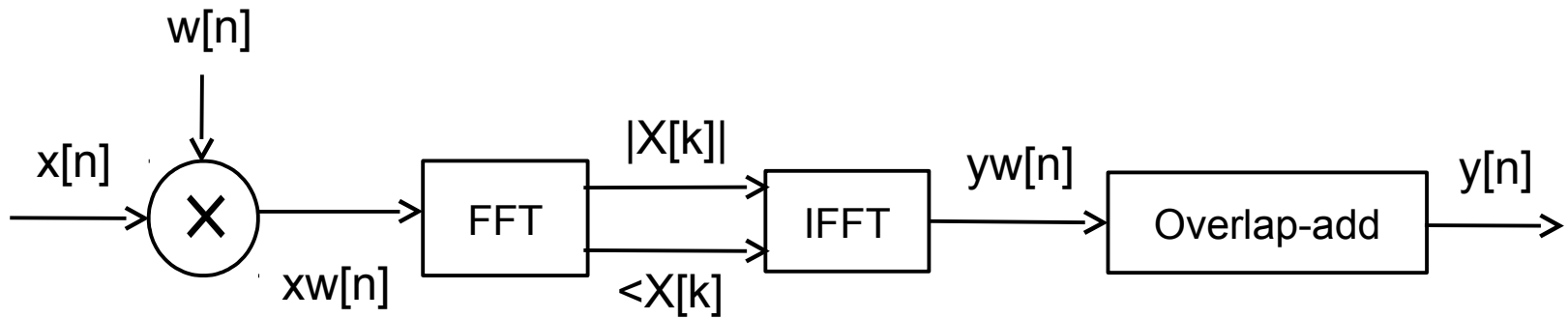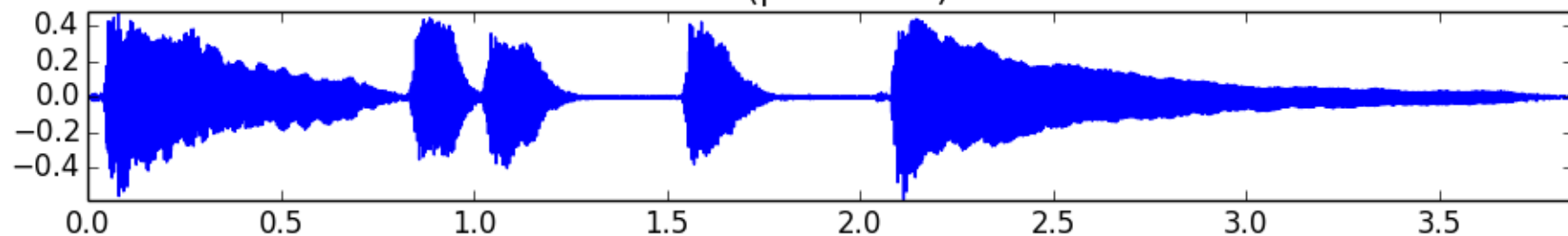

w[n]

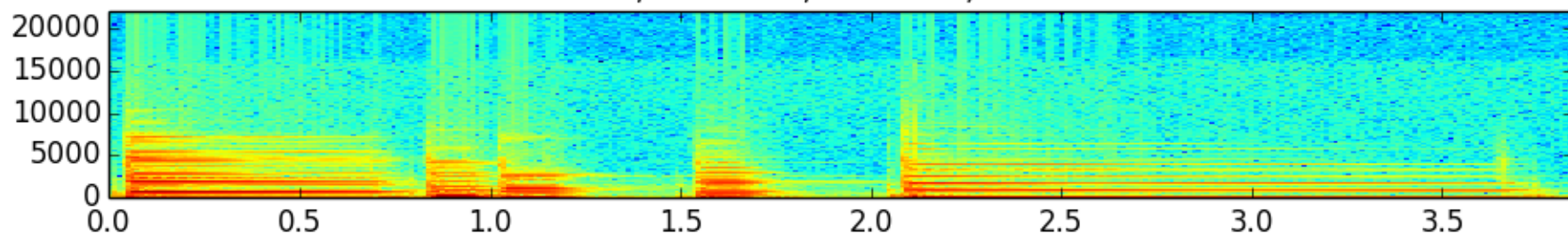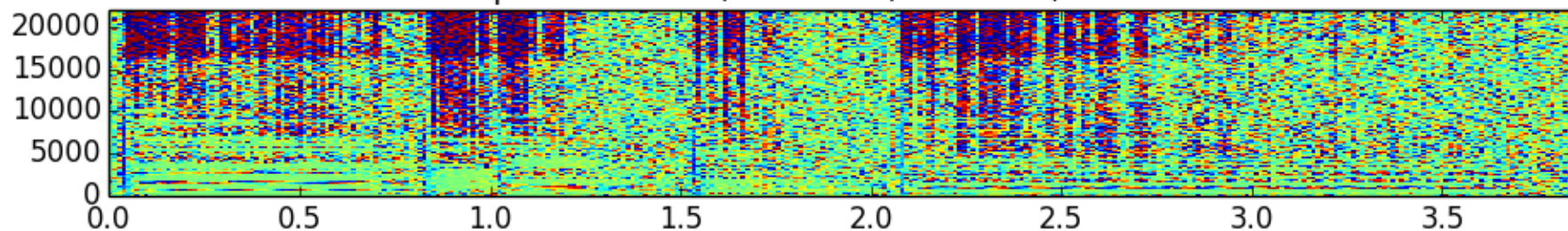x[n] → ⊗ → xw[n] → | FFT | → |X[k]| / <X[k] → | IFFT | → yw[n] → | Overlap-add | → y[n]

x (piano.wav)
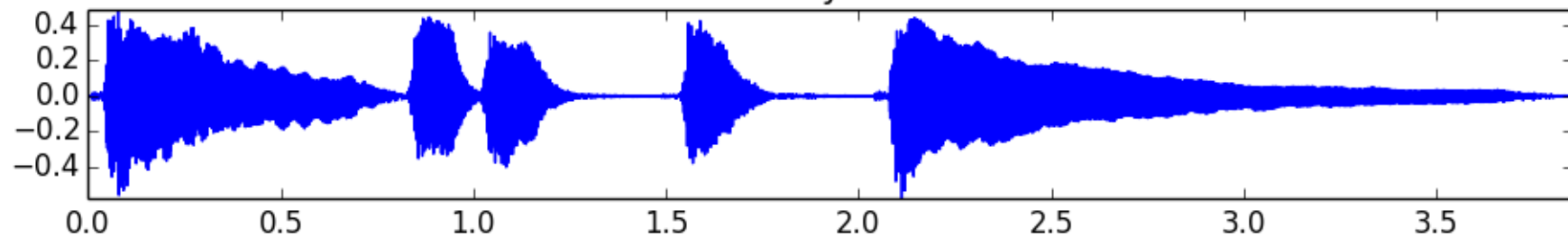
mX, M=1024, N=1024, H=512

pX derivative, M=1024, N=1024, H=512

y

References and credits

- More information in:
  https://en.wikipedia.org/wiki/STFT
  https://en.wikipedia.org/wiki/Window_function
  http://en.wikipedia.org/wiki/Spectrogram

- Reference on the STFT by Julius O. Smith:
  https://ccrma.stanford.edu/~jos/sasp/

- Sounds from:
  http://www.freesound.org/people/xserra/packs/13038/

- Slides released under CC Attribution-Noncommercial-Share Alike license and code under Affero GPL license. All available from https://github.com/MTG/sms-tools

4T2: The Short-Time Fourier Transform (2 of 2)

Xavier Serra

Universitat Pompeu Fabra, Barcelona