8T2: Spectral-based sound transformations (2 of 2)

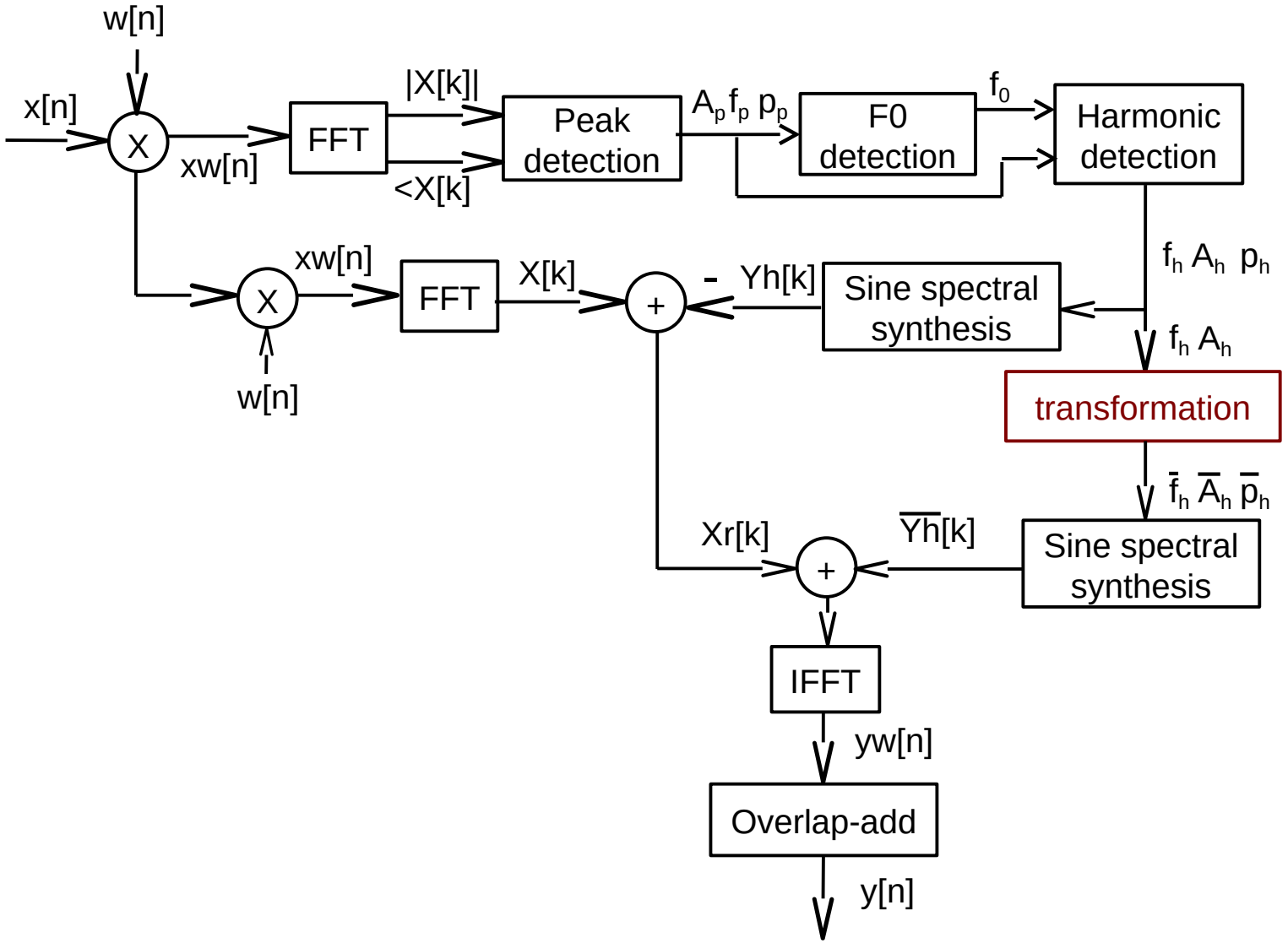
Xavier Serra

Universitat Pompeu Fabra, Barcelona

Index

- Harmonic plus residual model
  - Frequency transformations
- Harmonic plus stochastic model
  - Time scaling
  - Morphing

Harmonic plus residual model



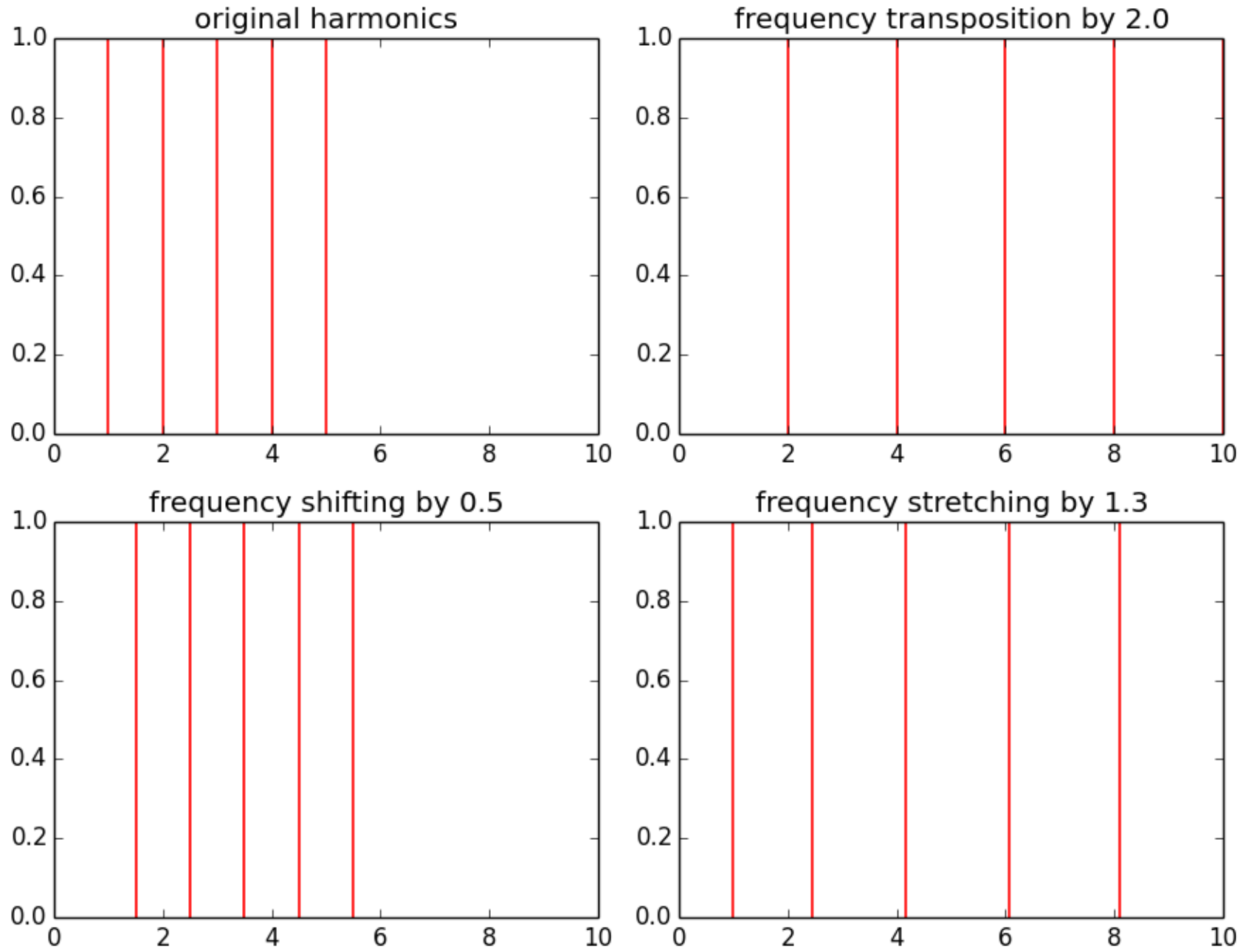
Frequency transformations

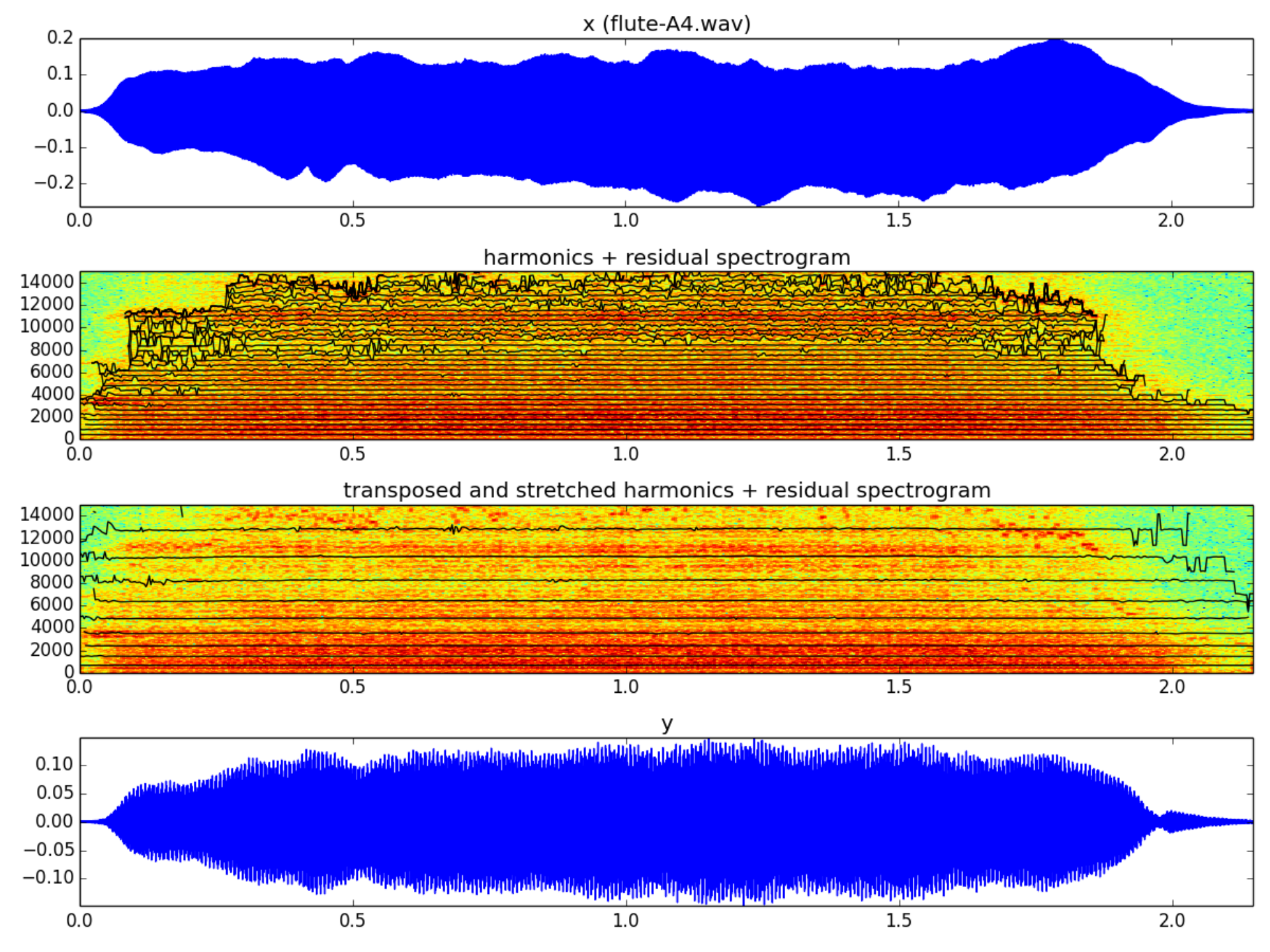
frequency transposition: $\bar{f}_h[l] = sf[l] f_h[l]$

frequency shifting: $\bar{f}_h[l] = sf[l] + f_h[l]$

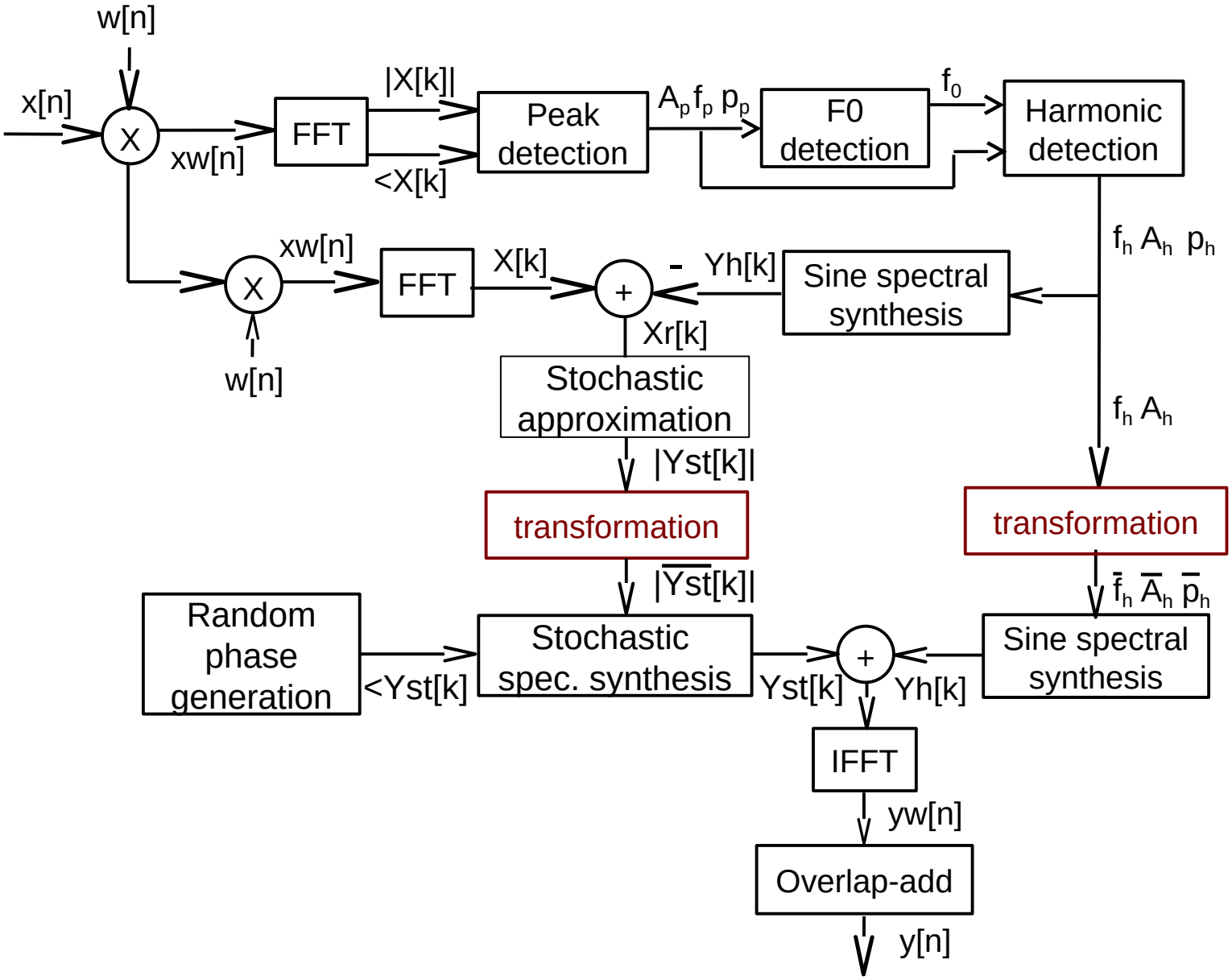
frequency stretching: $\bar{f}_h[l] = (f_h[l]/h) \times h^{sf[l]}$

f_h : input frequency of harmonic h
sf : scaling frequency
$\bar{f}_h$: output frequency of harmonic h


original harmonics
frequency transposition by 2.0
frequency shifting by 0.5
frequency stretching by 1.3
0.0
0.2
0.4
0.6
0.8
1.0
0
2
4
6
8
10


x (flute-A4.wav)
0.2
0.1
0.0
−0.1
−0.2
0.0
0.5
1.0
1.5
2.0
harmonics + residual spectrogram
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0
transposed and stretched harmonics + residual spectrogram
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0
y
0.10
0.05
0.00
−0.05
−0.10
0.0
0.5
1.0
1.5
2.0

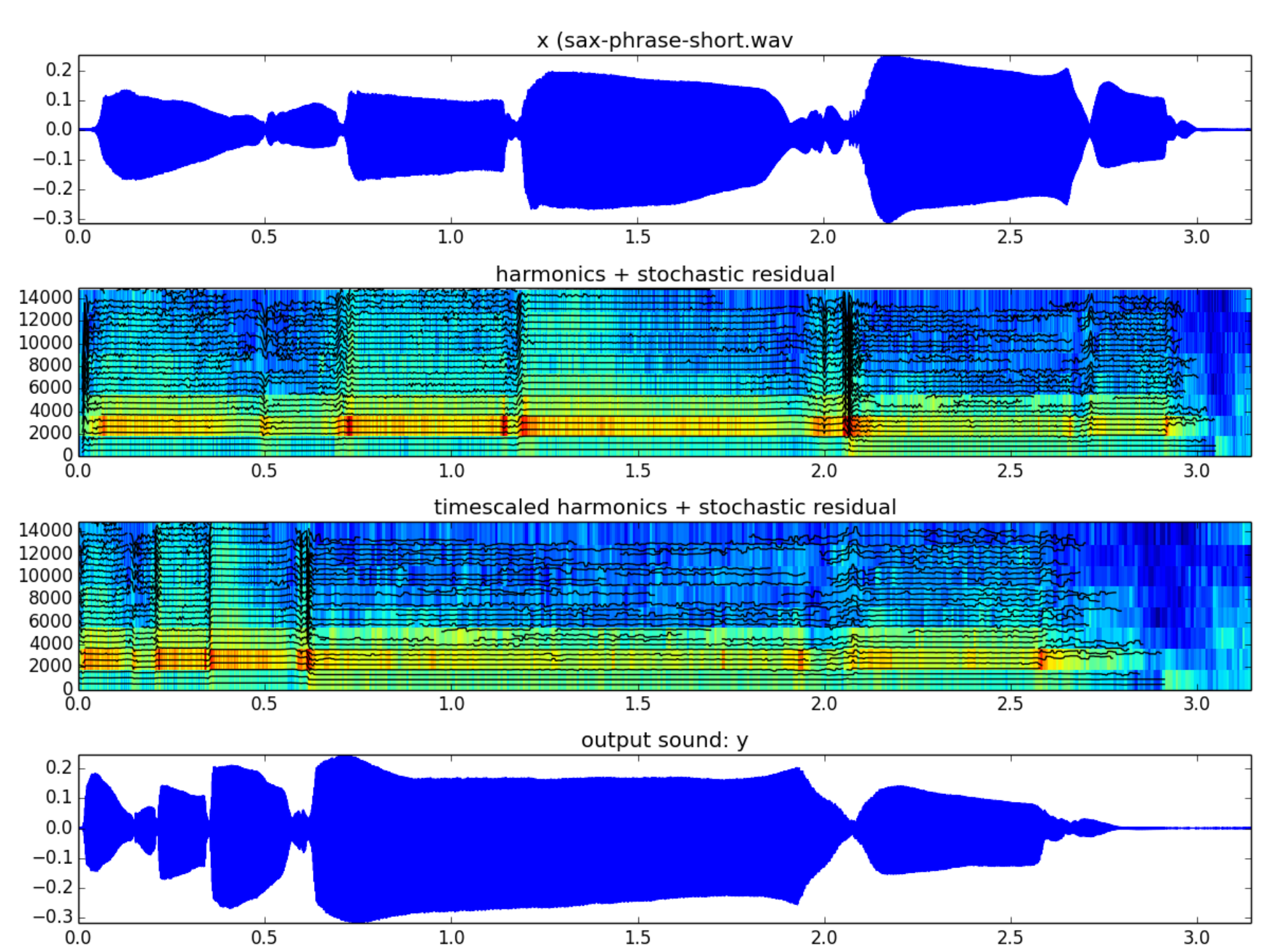
Harmonic plus stochastic model



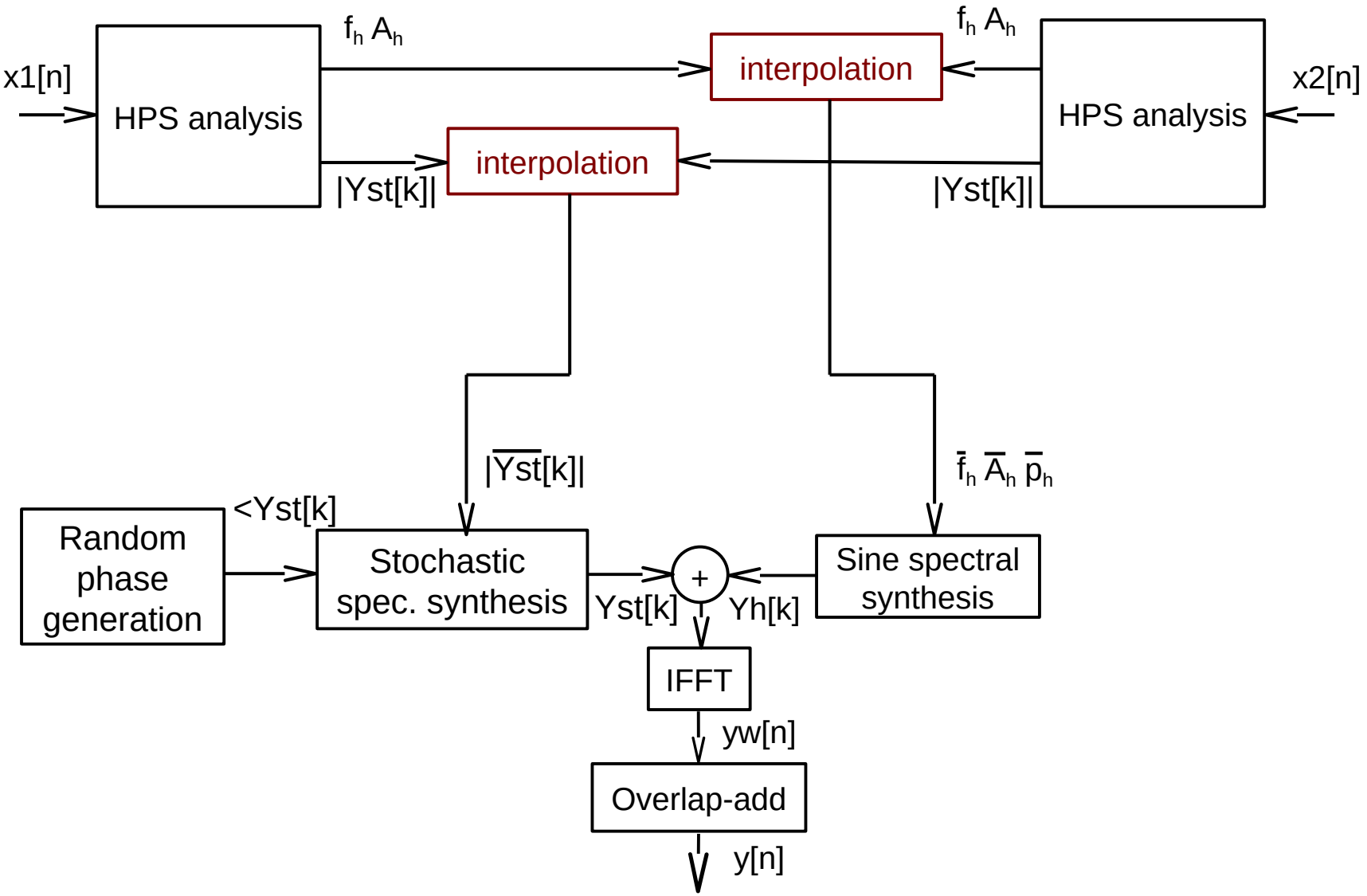
Scaling amplitude, frequency and time

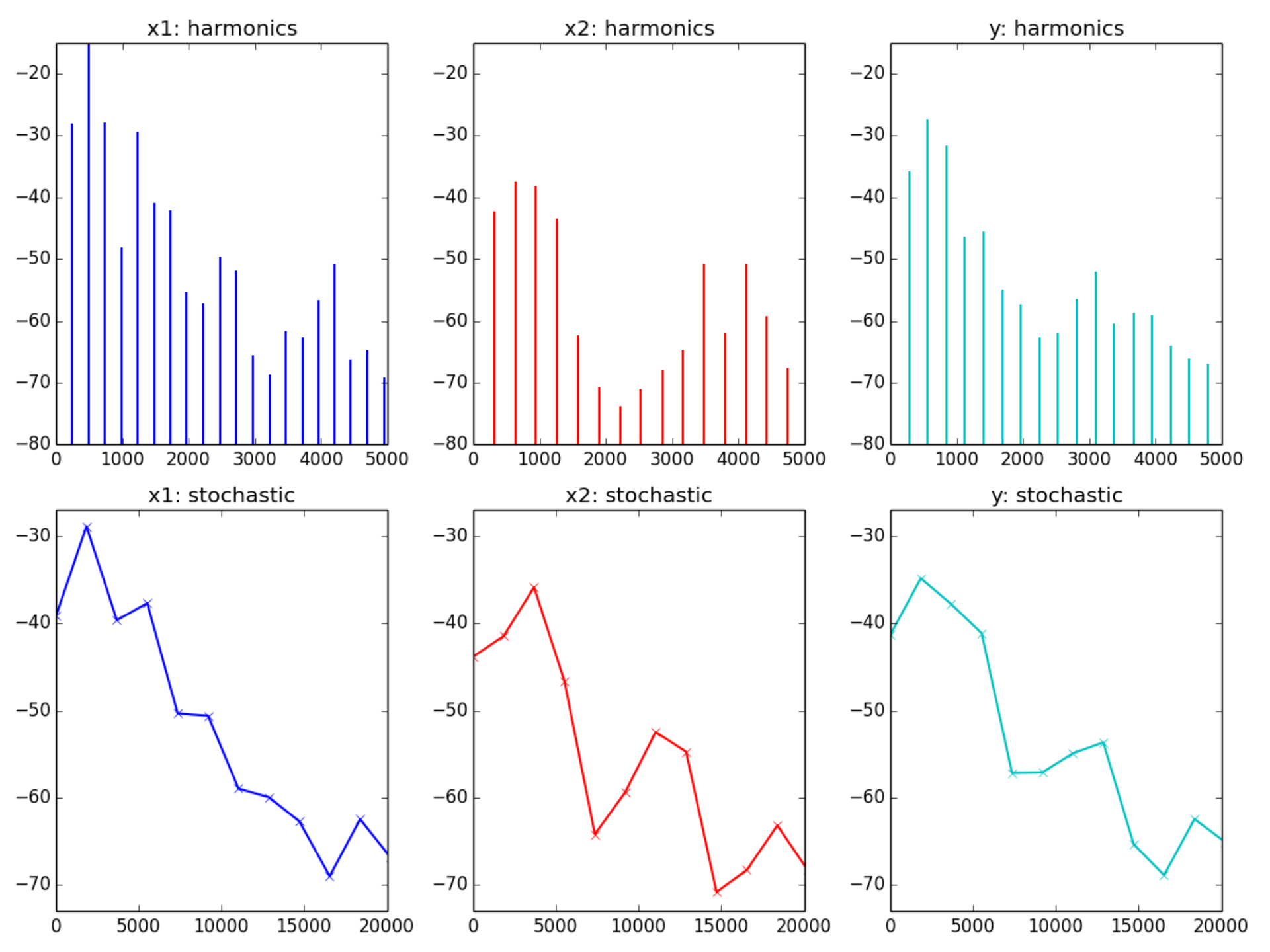
$$\bar{f}_h[q] = sf_h[l] f_t[st_h[l] l]$$
$$\bar{A}_h[q] = sA_h[l] + A_t[st_h[l] l]$$
$$\bar{\phi}_h[q] = \phi_h[q-1] + f_h[q]$$
$$\bar{st}_k[q] = sst_k[l] st_k[st_k[l] l]$$

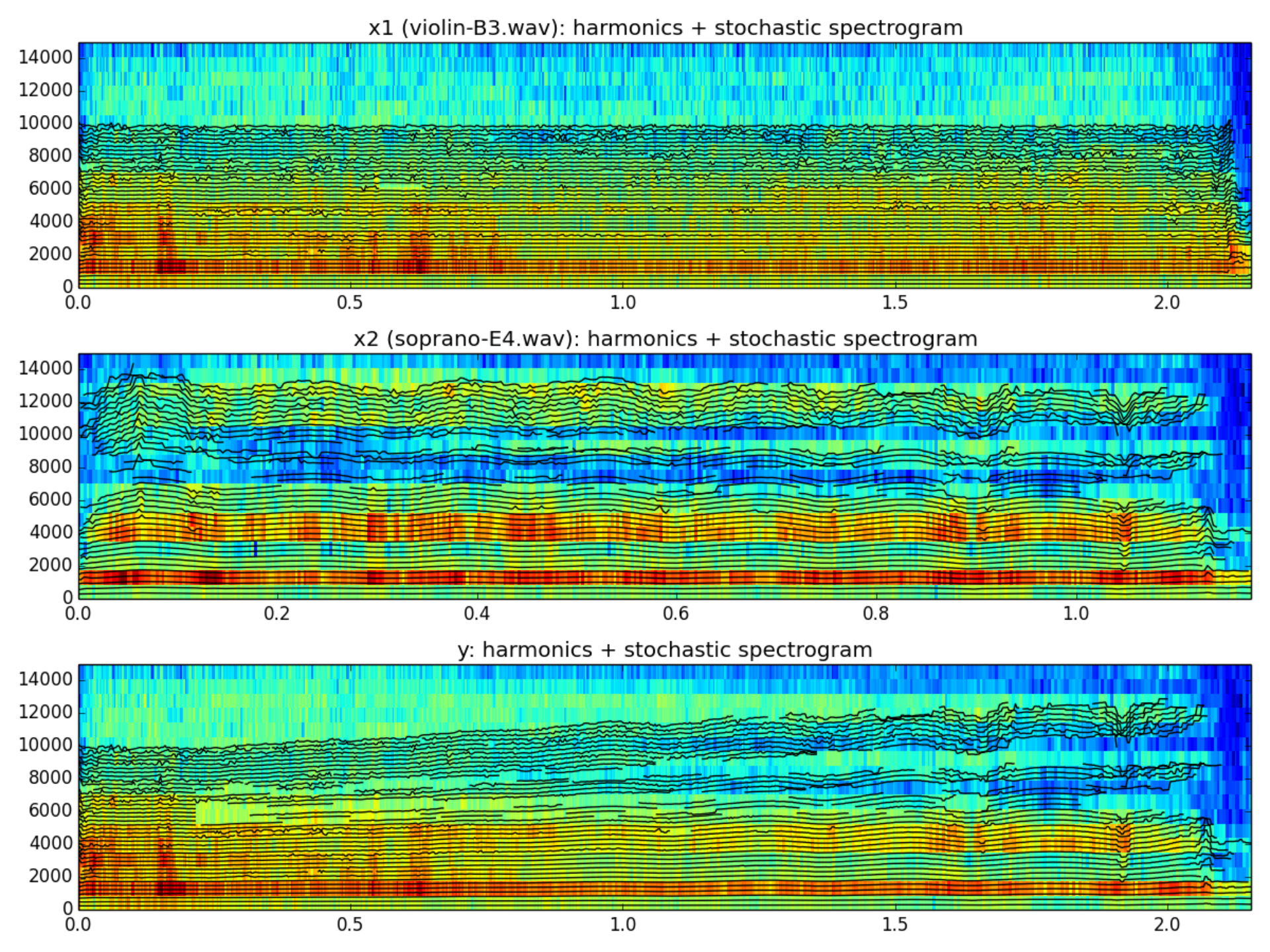
q : output frame ; l : input frame ; h : harmonic
f : input frequency ; A : input amplitude ; st : input stochastic envelope
sf : scaling frequency ; sA : scaling amplitude ; st : scaling time ;
sst : scaling stochastic
$\bar{f}$: output frequency ; $\bar{A}$: output amplitude ; $\bar{\phi}$: output phase ;
$\bar{st}$: output stochastic


x (sax-phrase-short.wav
0.2
0.1
0.0
−0.1
−0.2
−0.3
0.0
0.5
1.0
1.5
2.0
2.5
3.0
harmonics + stochastic residual
14000
12000
10000
8000
6000
4000
2000
0
timescaled harmonics + stochastic residual
output sound: y

Morphing


x1[n]
HPS analysis
f_h A_h
interpolation
f_h A_h
HPS analysis
x2[n]
interpolation
|Yst[k]|
|Yst[k]|
$|\overline{Yst}[k]|$
$\bar{f}_h$ $\overline{A}_h$ $\overline{p}_h$
<Yst[k]
Random phase generation
Stochastic spec. synthesis
+
Sine spectral synthesis
Yst[k]
Yh[k]
IFFT
yw[n]
Overlap-add
y[n]


x1: harmonics
x2: harmonics
y: harmonics
x1: stochastic
x2: stochastic
y: stochastic
−20
−30
−40
−50
−60
−70
−80
0
1000
2000
3000
4000
5000
0
5000
10000
15000
20000


x1 (violin-B3.wav): harmonics + stochastic spectrogram
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0
x2 (soprano-E4.wav): harmonics + stochastic spectrogram
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.2
0.4
0.6
0.8
1.0
y: harmonics + stochastic spectrogram
14000
12000
10000
8000
6000
4000
2000
0
0.0
0.5
1.0
1.5
2.0

References

- More information on this topic from Wikipedia:
  - http://en.wikipedia.org/wiki/Sound_effects
  - http://en.wikipedia.org/wiki/Audio_timescale-pitch_modification
- Sounds: http://www.freesound.org/people/xserra/packs/13038/
- Slides released under CC Attribution-Noncommercial-Share Alike license and code under Affero GPL license; available from https://github.com/MTG/sms-tools

8T2: Spectral-based sound transformations (2 of 2)

Xavier Serra

Universitat Pompeu Fabra, Barcelona