Biweekly report

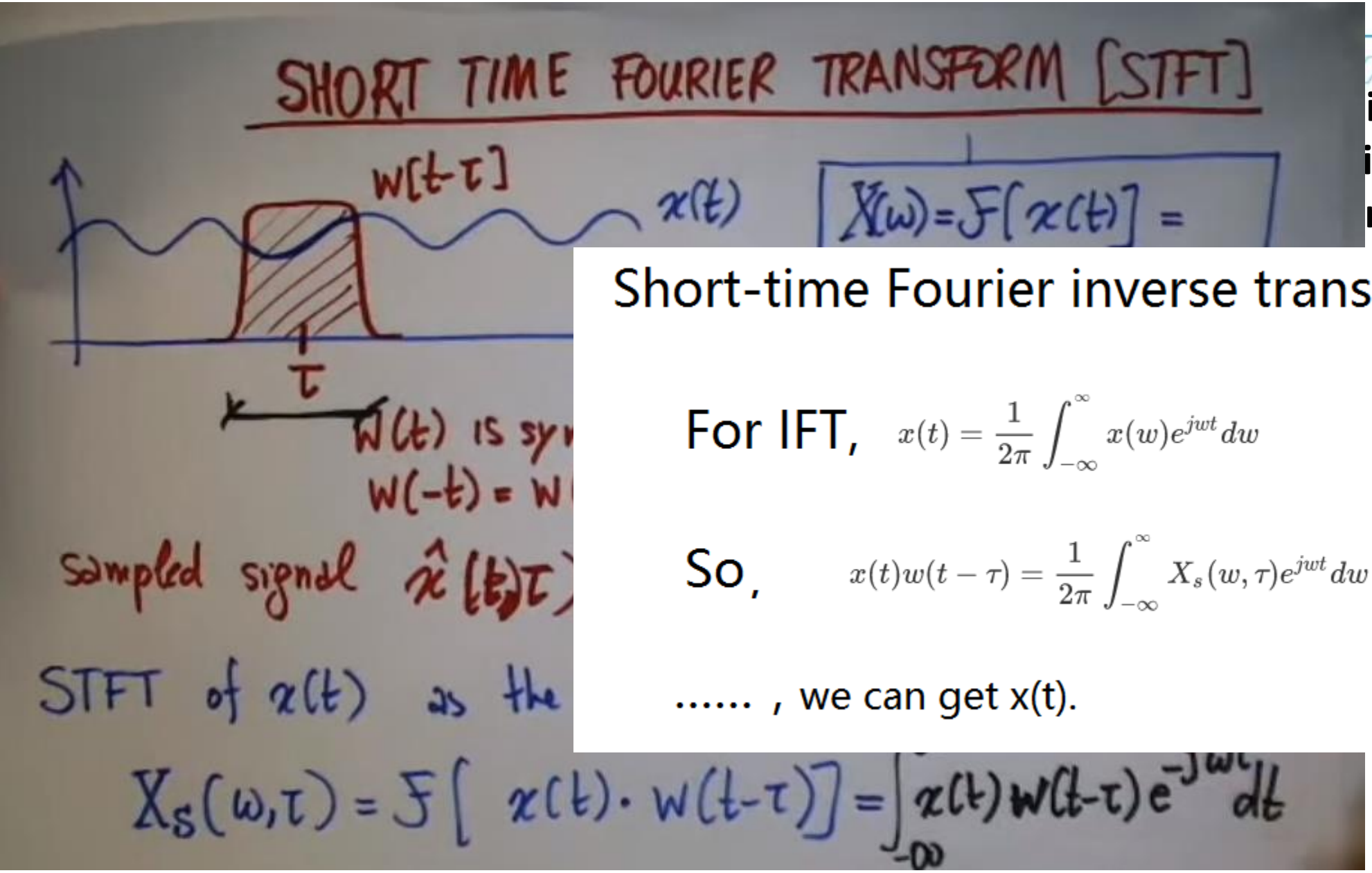
STFT, Spectrogram

Reporter：Xiaofei Kang.

July 11, 2017

CONTENTS

STFT, ISTFT



iding window Fourier
ing multiplication, for
rier transform

Short-time Fourier inverse transform (ISTFT)

For IFT, $x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(w)e^{jwt}dw$

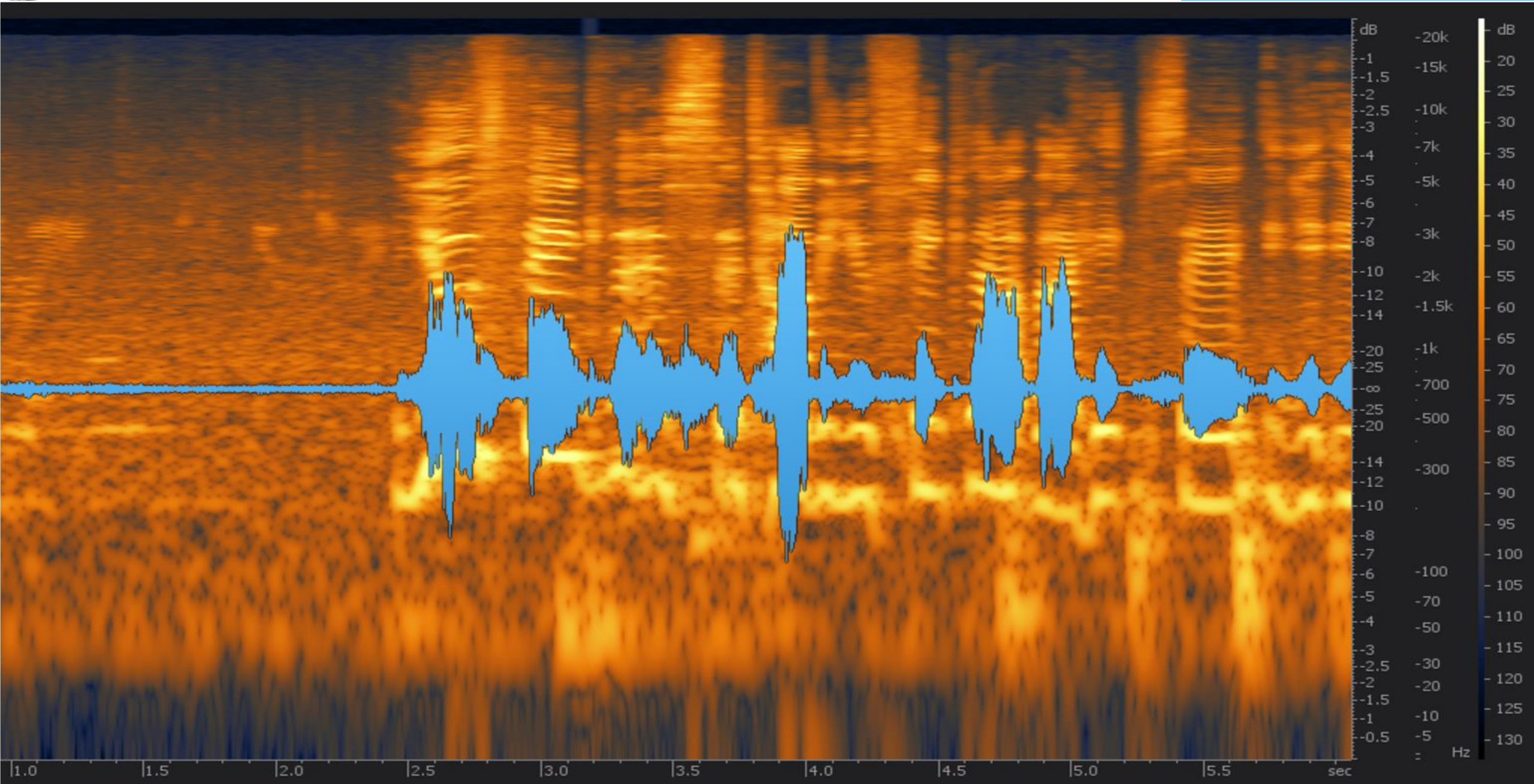
So, $x(t)w(t-\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X_s(w,\tau)e^{jwt}dw$

..... , we can get x(t).

CONTENTS

Spectrogram

CSLT



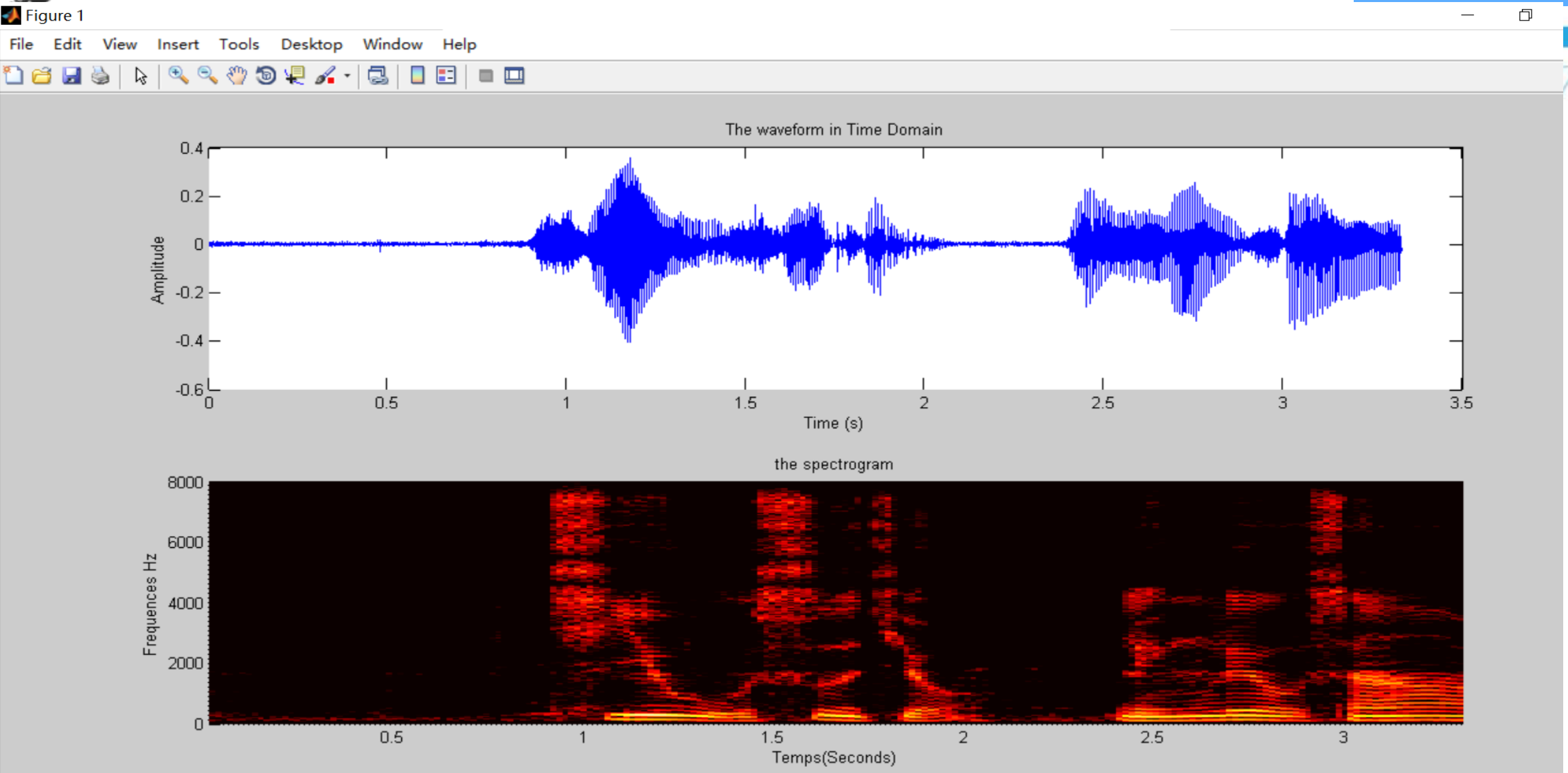
Time-frequency map ： For x (t) to do short-term Fourier transform, you can do the time-frequency analysis of the whole signal, you can get the signal time-frequency distribution map ：

$$X(t,f)=\int_{-\infty}^{\infty} w(t-\tau)x(\tau)\ e^{-j2\pi f\tau}d\tau$$

Spectrogram：a three-dimensional spectrum, indicating the voice spectrum with time changes in the graphics, the vertical axis is the frequency, the horizontal axis is the time, the coordinate point value for the voice data energy.

$$SP(t,f)=|X(t,f)|^2=X(t,f)X^*(t,f)$$

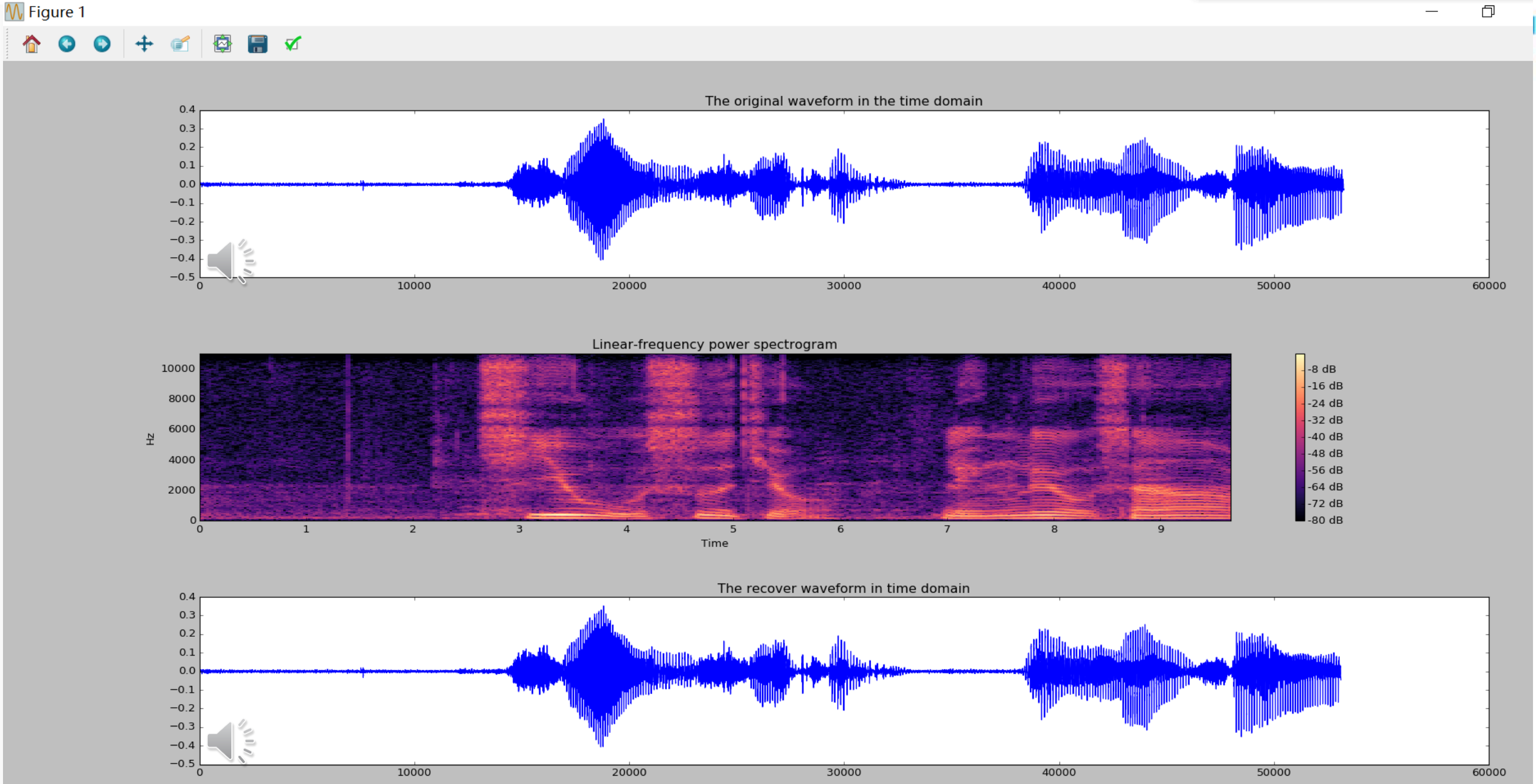
2 Matlab:Waveform->Spectrogram



NB:"Spectrogram" is just a name for a heat map of the magnitude of the STFT, and magnitude alone is not enough to reconstruct a signal.

3 Python:Waveform->Sectrogram->Waveform

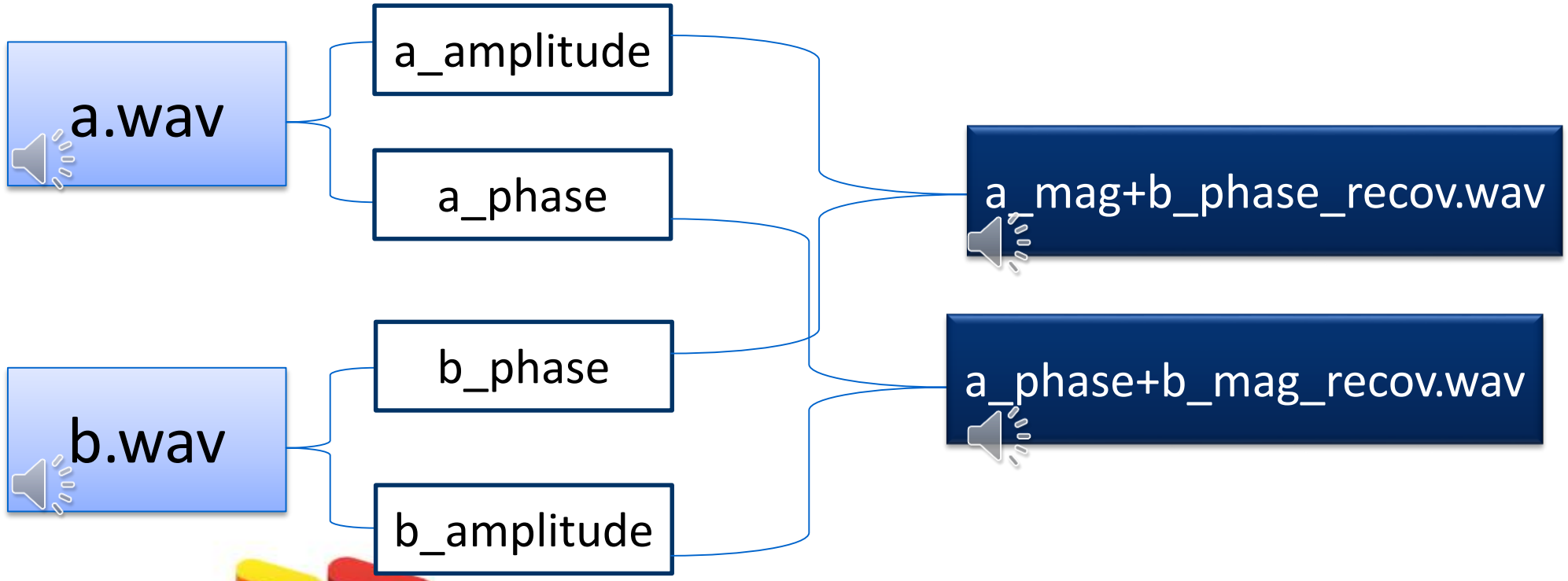
CSLT



CONTENTS

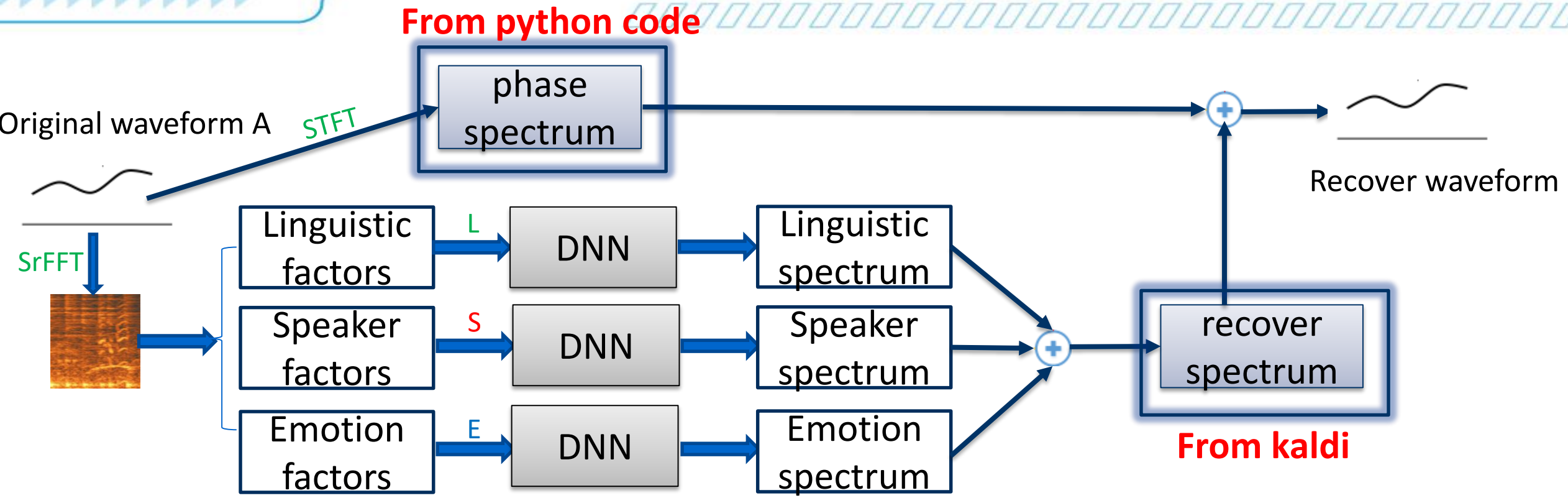
Experiment1—audio mixing

CSLT



2 Experimen2—Goals

CSLT



1、Speaker spectrum(S)→ waveform、Linguistic spectrum (L) → waveform
Emotion spectrum (E) → waveform

2、 Recovery: S1+ L1+ E1→ waveform、
Combination: S2+ L1+ E1→ waveform、 S1+ L2+ E1→ waveform、
S1+ L1+ E2→ waveform……

For Python：

$x(t)$

STFT

$$X(t,f) = |X(t,f)|e^{j\varphi(t,f)} = a + bi$$

$$lnX(t,f) = ln|X(t,f)| + j\varphi(t,f)$$

$$There, |X(t,f)| = \sqrt{a^2 + b^2}$$

$$logmag = \ln|X(t,f)|$$

$$phase = \varphi(t,f)$$

For kaldi：

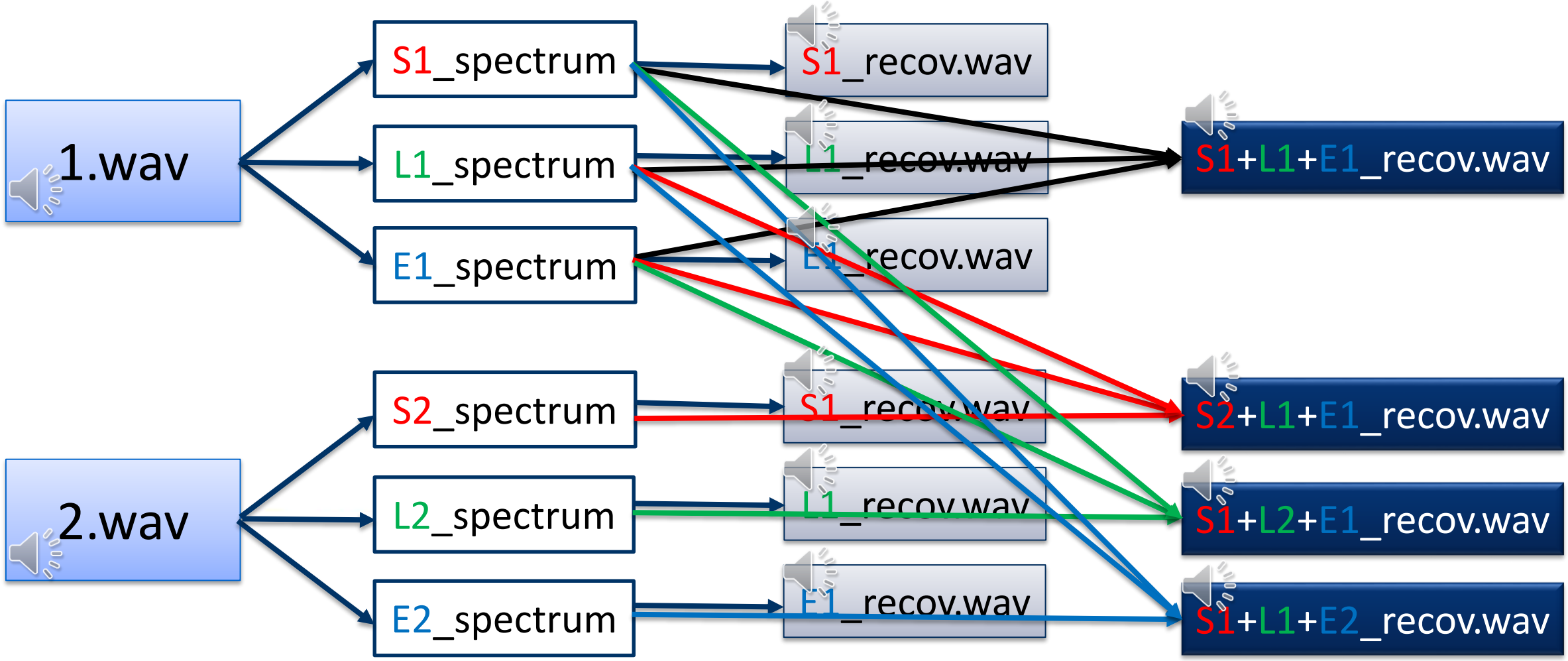
Given data: $1n(a^2 + b^2)$

And, $1n(a^2 + b^2) = 2ln|X(t,f)|$

So, $1n(a^2 + b^2) = 2logmag$

Experiment2 - Speech conversion

CSLT



ShortComings

1、 Python method is different from kaldi method.

- Recover waveform all in the kaldi.
- Recover waveform all in the python.


THANK YOU