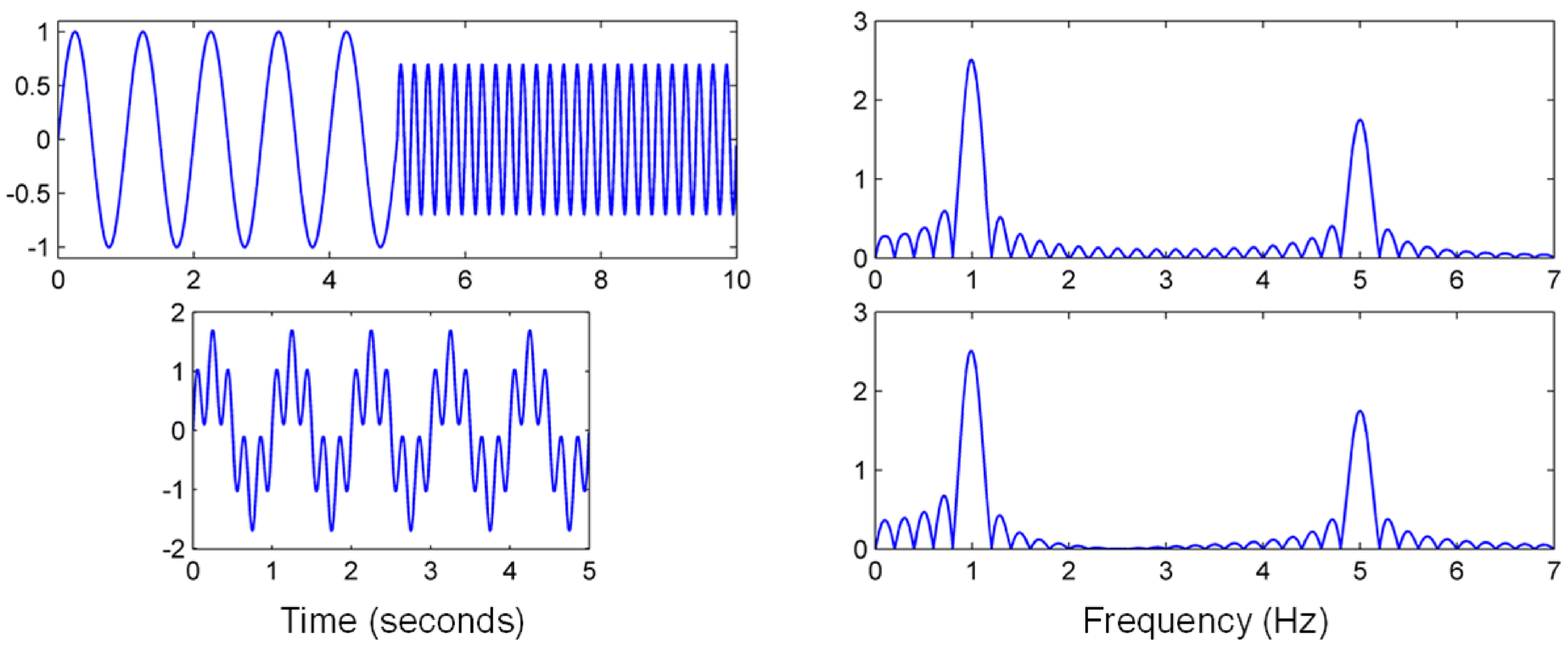
Supplement

HW 2

2017/3/31

Short-time Fourier transform (STFT)

- Fourier analysis assumes that the amplitude/frequency/phase of a signal do not change over time.
- However this is not the case in real world.
- This is why we need STFT to model the temporal dynamics of real world audio signal.


Time (seconds)
Frequency (Hz)

Short-time Fourier transform (STFT)

- Continuous-time STFT
  - Simply, in the continuous-time case, the function to be transformed is multiplied by a window function which is nonzero for only a short period of time.
  - The Fourier transform (a one-dimensional function) of the resulting signal is taken as the window is slid along the time axis, resulting in a two-dimensional representation of the signal.

$$\mathbf{STFT}\{x(t)\}(\tau,\omega) \equiv X(\tau,\omega) = \int_{-\infty}^{\infty} x(t)w(t-\tau)e^{-j\omega t}\,dt$$

w(*t*) is the window function, commonly a Hann window or Gaussian window centered around zero, and *x*(*t*) is the signal to be transformed.

Short-time Fourier transform (STFT)

- Discrete STFT
  - In the discrete time case, the data to be transformed could be broken up into chunks or frames (which usually overlap each other, to reduce artifacts at the boundary).
  - Each chunk is Fourier transformed, and the complex result is added to a matrix, which records magnitude and phase for each point in time and frequency.

$$\mathbf{STFT}\{x[n]\}(m,\omega) \equiv X(m,\omega) = \sum_{n=-\infty}^{\infty} x[n]w[n-m]e^{-j\omega n}$$

likewise, with signal $x[n]$ and window $w[n]$

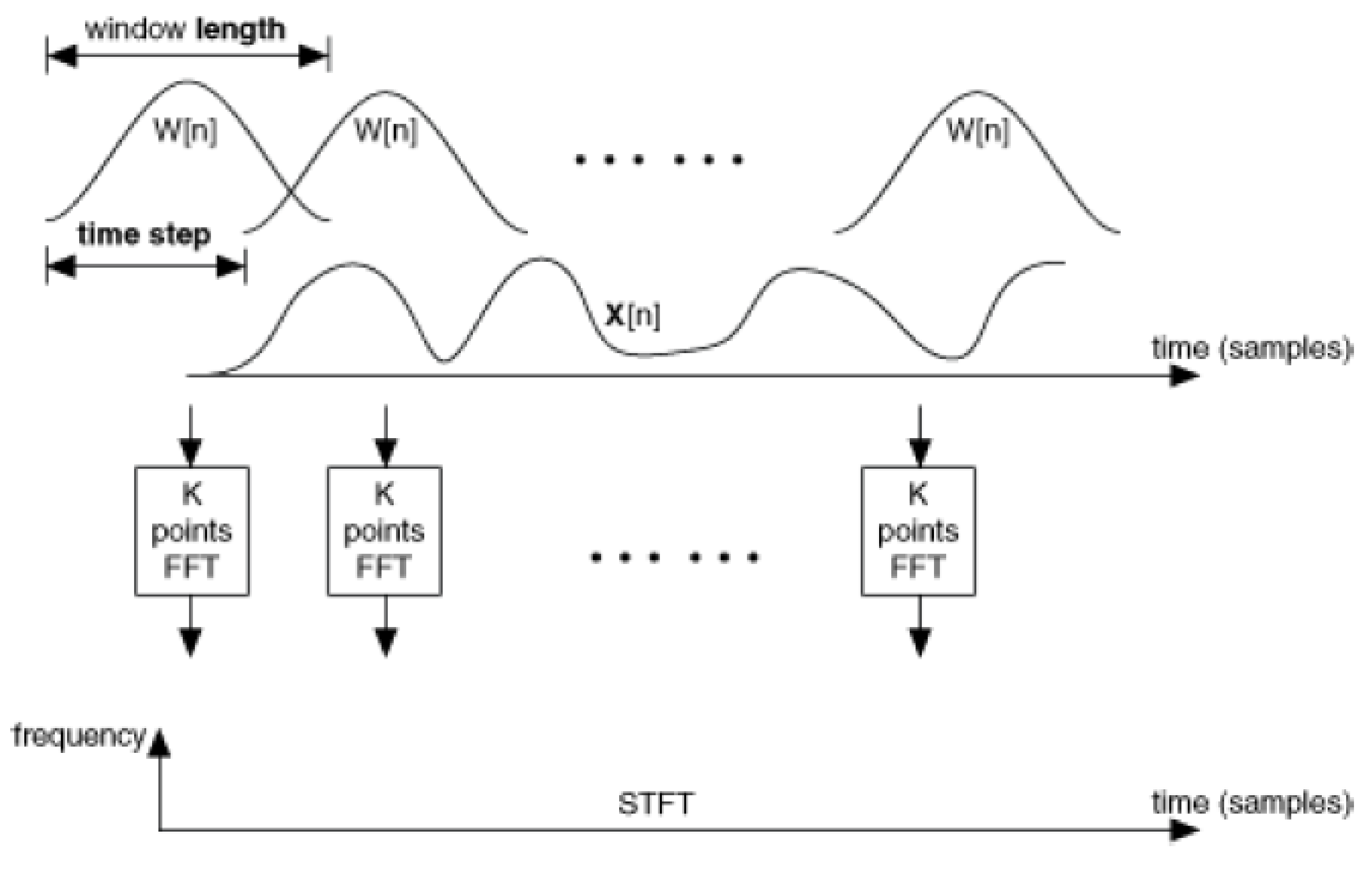
Short-time Fourier transform (STFT)

- Discrete STFT
  - The magnitude squared of the STFT yields the spectrogram of the function:

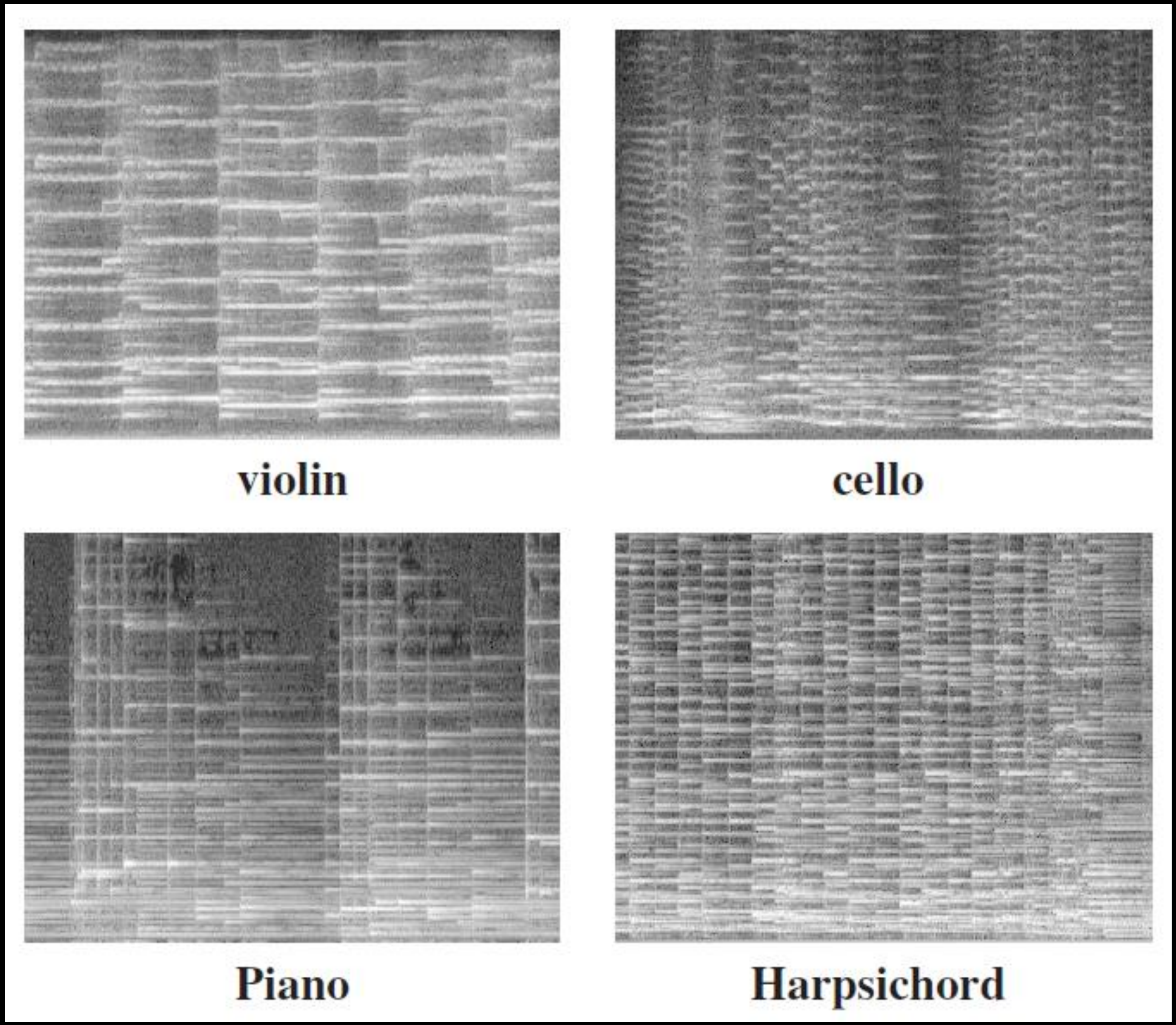
$$\mathrm{spectrogram}\{x(t)\}(\tau,\omega) \equiv |X(\tau,\omega)|^2$$

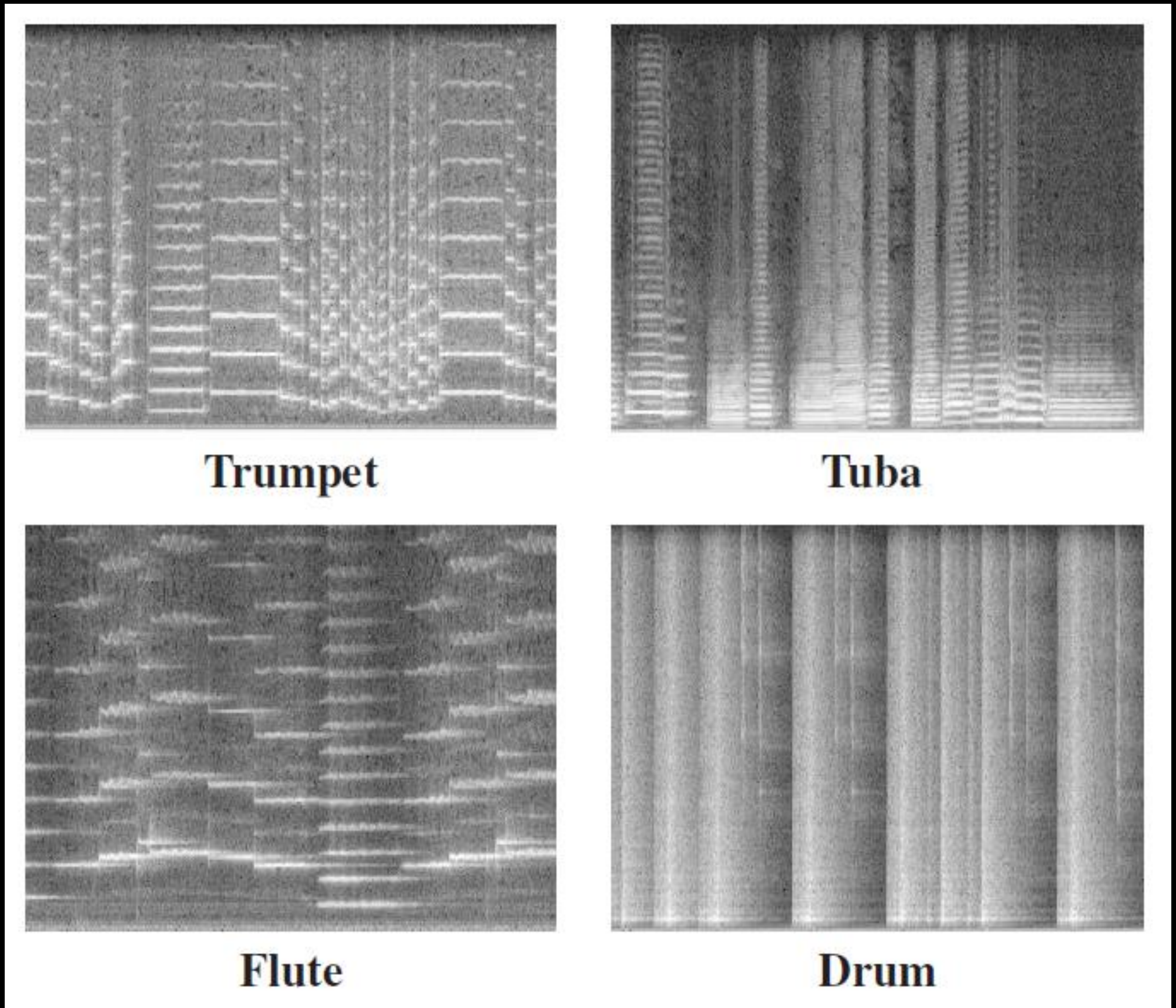
Short-time Fourier transform (STFT)

- Implementation
  1. Slice the signal into frames of segments.
     - Each segment may overlap with its previous segment.
  2. Multiply the short segments by a window function.
  3. Do FFT for each segment.
  4. Aggregate the FFT result of each segment into a matrix S.
  5. The magnitude squared of the matrix S is the spectrogram.



window length
W[n]
W[n]
W[n]
time step
X[n]
time (samples)
K points FFT
K points FFT
K points FFT
frequency
STFT
time (samples)


violin
cello
Piano
Harpsichord


Trumpet
Tuba
Flute
Drum