6.300: Signal Processing

Short-Time Fourier Transform (STFT)

Windows: Consider the main lobe and sidelobes.

Time Resolution vs. Frequency Resolution

- time resolution: short $w[n] \iff$ wide $W(\Omega)$
- frequency resolution: long $w[n] \iff$ narrow $W(\Omega)$

Analysis: $X[k,m] = \frac{1}{N}\sum_{n=0}^{N-1} x[n+ms]w[n]e^{-jk\frac{2\pi}{N}n}$

Spectrogram: $|X[k,m]|^2$ (magnitude^2 of STFT)

April 9, 2026

Agenda for Recitation

- Windows: Rectangular, triangular, and Hann
- Short-time Fourier transforms and spectrograms
- Examples of spectrograms (as time allows)

What questions do you have from lecture?

Windows

Multiplying $x[n]$ by the window function $w[n]$ corresponds to convolving the DTFT of $x[n]$ with the DTFT of $w[n]$.

Windowing

$$x_w[n] = x[n]w[n] \iff X_w(\Omega) = \frac{1}{2\pi}\big(X * W\big)(\Omega)$$

- time resolution: short $w[n] \iff$ wide $W(\Omega)$
- frequency resolution: long $w[n] \iff$ narrow $W(\Omega)$

There is a **trade-off** between time and frequency resolution. This is Heisenberg's uncertainty principle!

Windows

There are many window functions. SciPy has 25 of them!

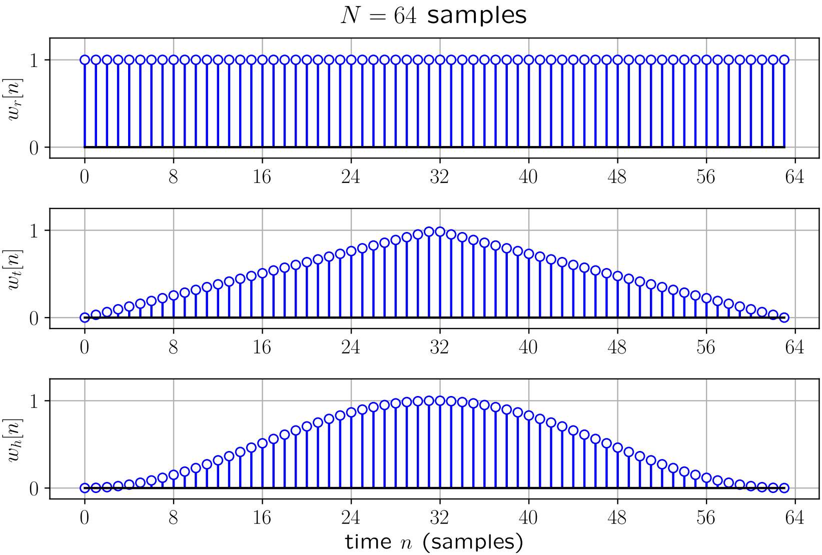
scipy.signal.windows: Bartlett, Bartlett-Hann, Blackman, Blackman-Harris, Bohman, boxcar, cosine, discrete prolate spheroidal sequences, Dolph-Chebyshev, exponential, flat-top, Gaussian, generalized cosine, generalized Gaussian, generalized Hamming, Hamming, Hann, Kaiser, Kaiser-Bessel, Lanczos, Nutall, Parzen, Taylor, triangular, Tukey

Let's examine the DTFT of a few windows.

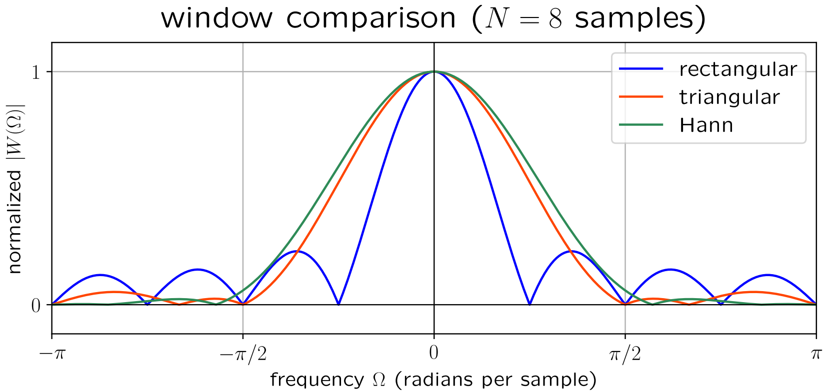
- **rectangular:** narrow mainlobe, high sidelobes
- **triangular:** wider mainlobe, shorter sidelobes
- **Hann:** even wider mainlobe, even shorter sidelobes

Good frequency resolution means "narrow mainlobe."
Problem: Sidelobes may obscure low-power signals.

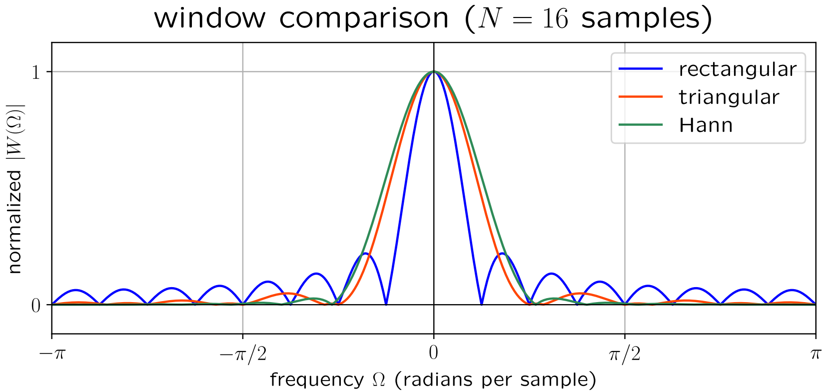
Windows



Windows

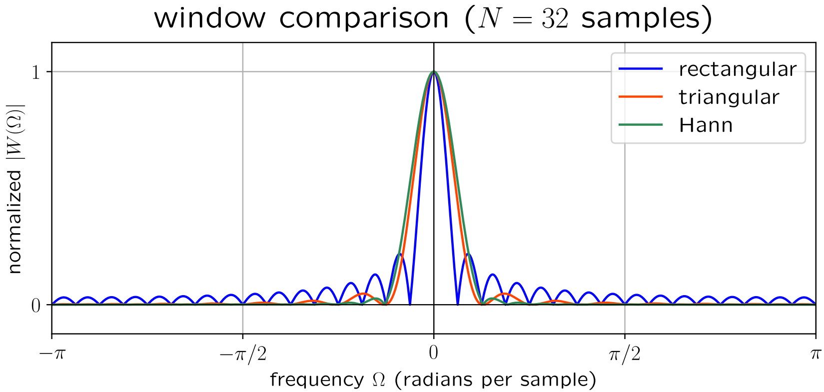

window comparison ($N = 8$ samples)
rectangular
triangular
Hann
normalized $|W(\Omega)|$
1
0
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

Windows



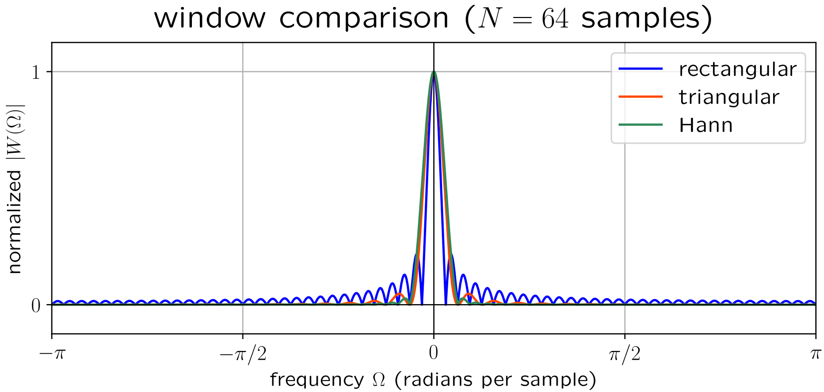
window comparison ($N = 16$ samples)
rectangular
triangular
Hann
normalized $|W(\Omega)|$
1
0
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

Windows



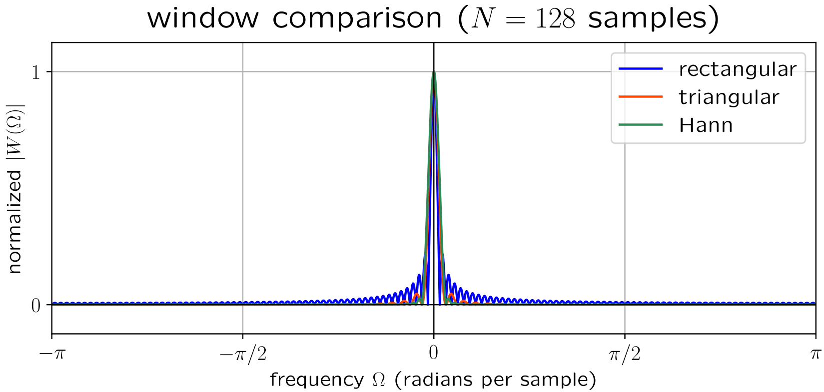
window comparison ($N = 32$ samples)
normalized $|W(\Omega)|$
1
0
rectangular
triangular
Hann
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

Windows



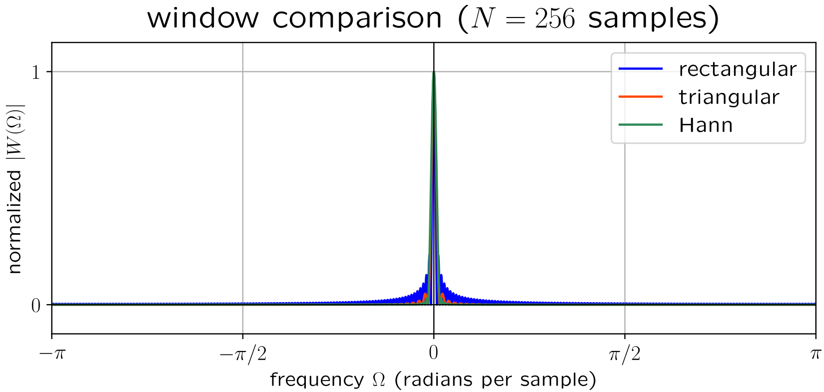
window comparison ($N = 64$ samples)
rectangular
triangular
Hann
1
0
normalized $|W(\Omega)|$
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

Windows

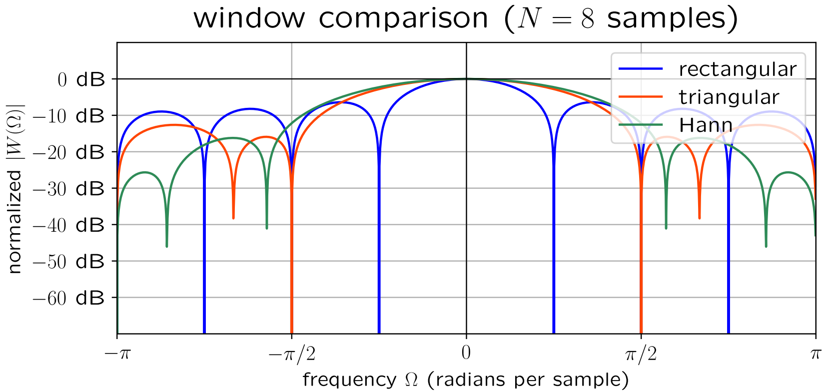


window comparison ($N = 128$ samples)
1
0
normalized $|W(\Omega)|$
rectangular
triangular
Hann
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

Windows


window comparison ($N = 256$ samples)
rectangular
triangular
Hann
normalized $|W(\Omega)|$
1
0
$-\pi$
$-\pi/2$
0
$\pi/2$
π
frequency Ω (radians per sample)

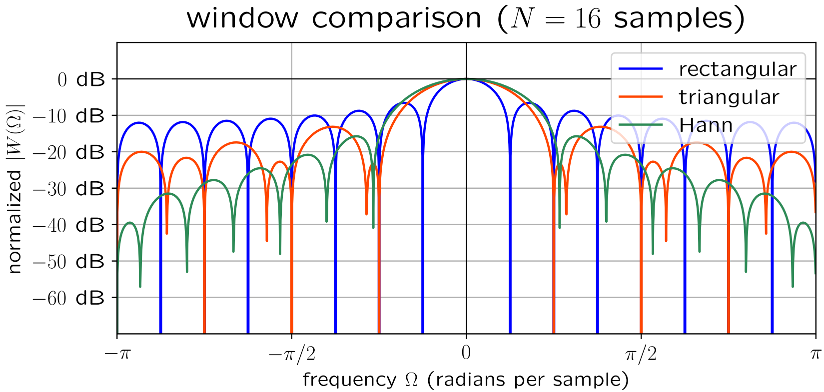
Windows



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

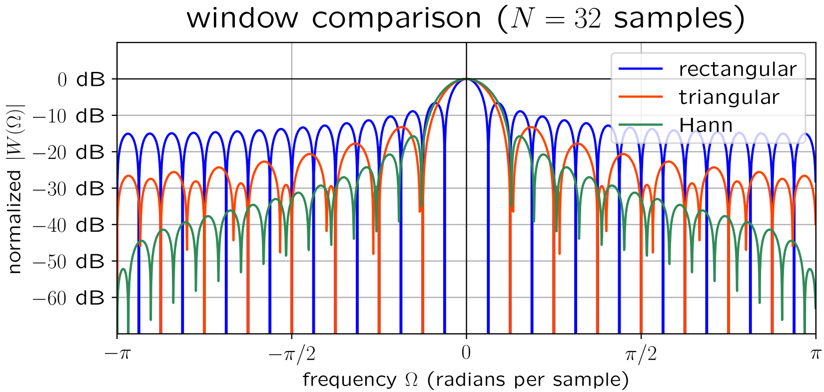
Windows



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

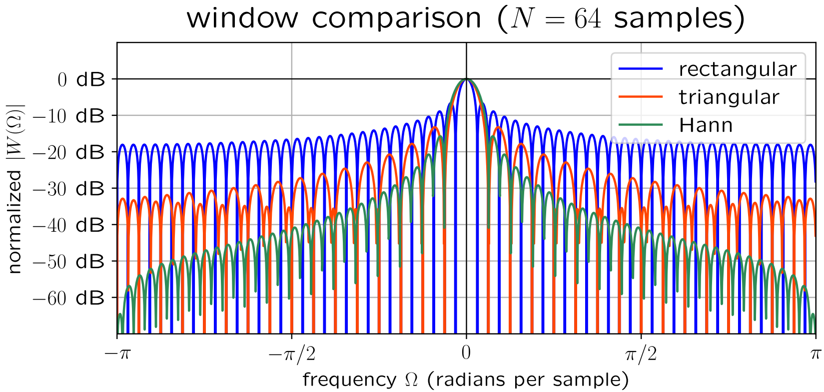
Windows



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

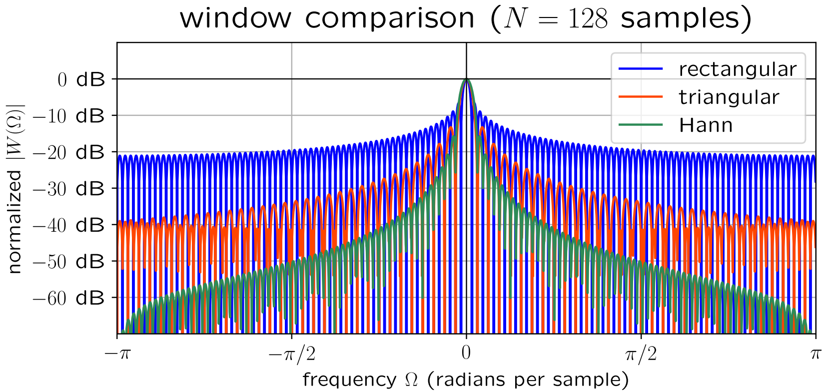
Windows



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

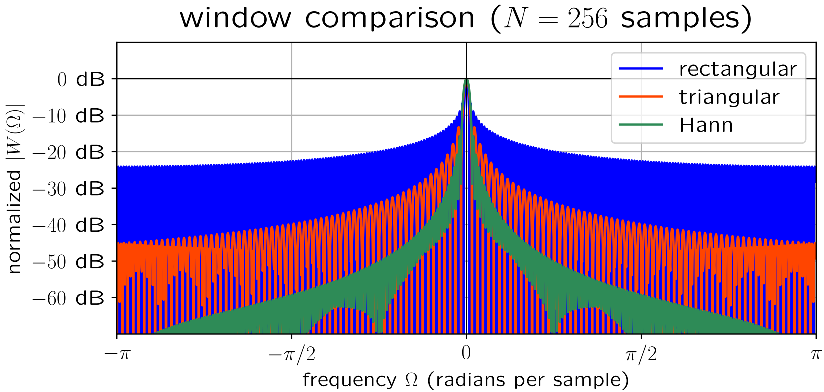
Windows



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

Windows



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

Microsoft Windows

Warning!

Your computer may crash if you use Microsoft Windows.

Short-Time Fourier Transform

STFT: Analyze a signal "block by block." An STFT is a sequence of windowed Fourier transforms.

$$X[k,m] = \frac{1}{N}\sum_{n=0}^{N-1} x[n+ms]w[n]e^{-jk\frac{2\pi}{N}n}$$

Spectrogram: $|X[k,m]|^2$

Parameters:

- type of window: $w[n]$
- DFT size: N
- step size: s

Short-Time Fourier Transform

Parameters:

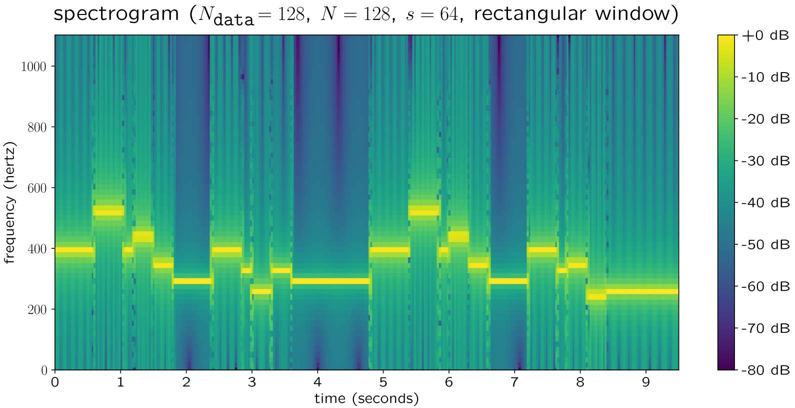
- type of window: $w[n]$
- DFT size: N
- step size: s

Short-Time Fourier Transform

Parameters:

- type of window: $w[n]$
- DFT size: N
- step size: s

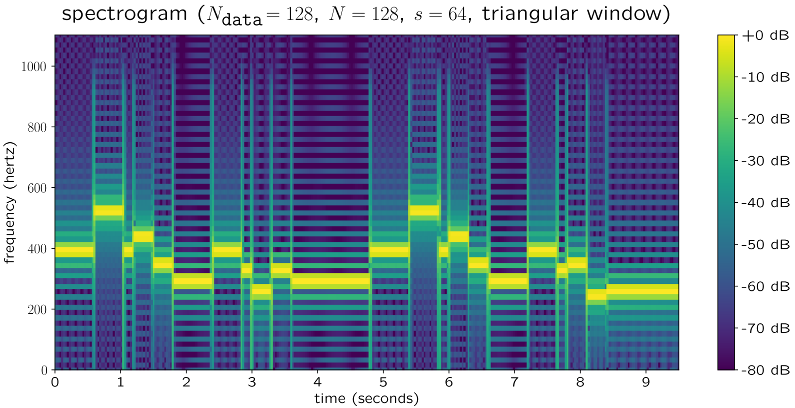
STFT: Windows



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

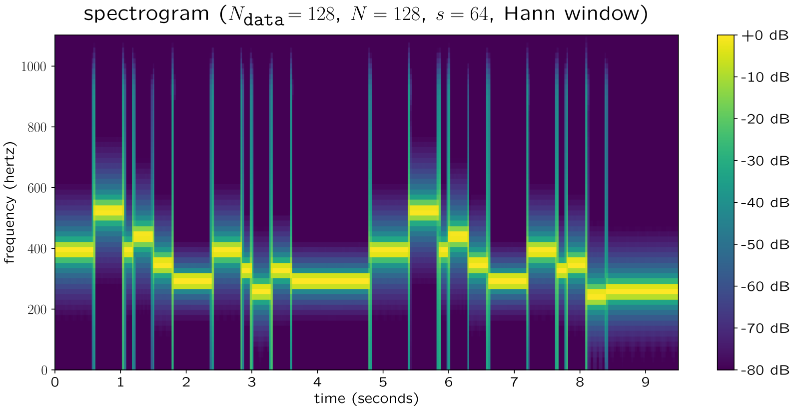
STFT: Windows



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

STFT: Windows



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

Short-Time Fourier Transform

Parameters:

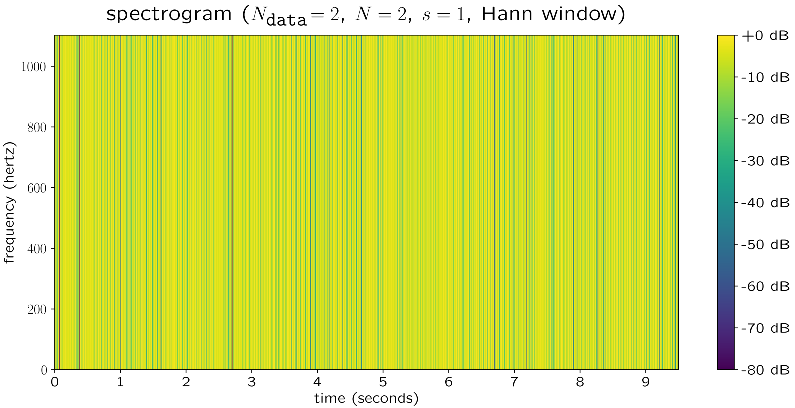
- type of window: $w[n]$
- DFT size: N
- step size: s

Short-Time Fourier Transform

Parameters:

- type of window: $w[n]$
- DFT size: N
- step size: s

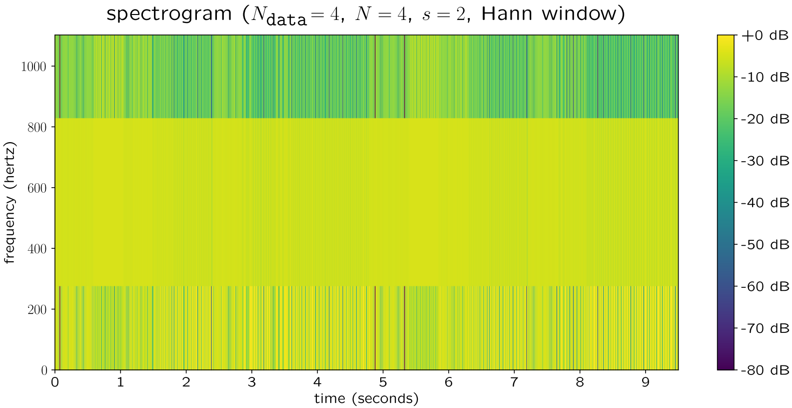
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

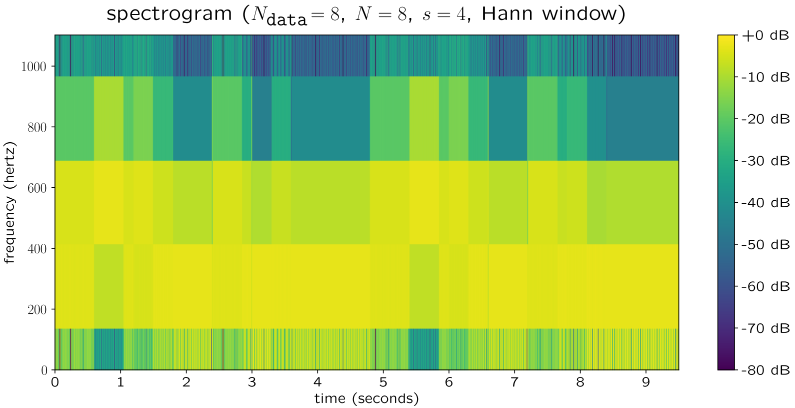
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

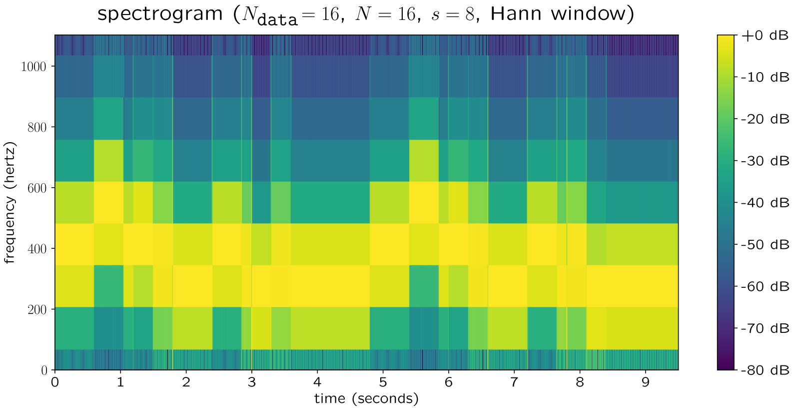
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

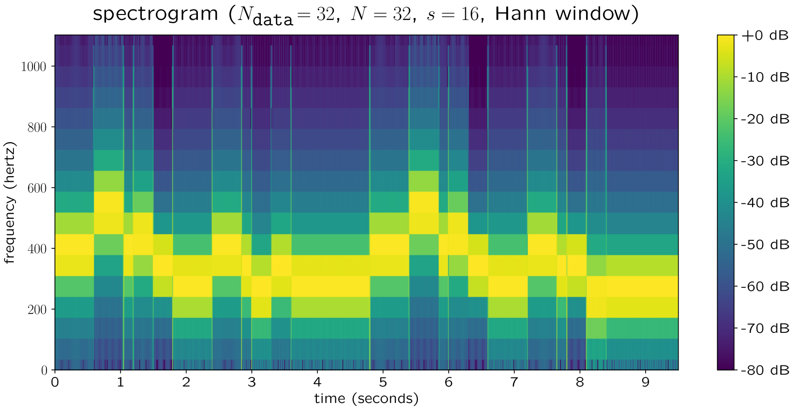
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

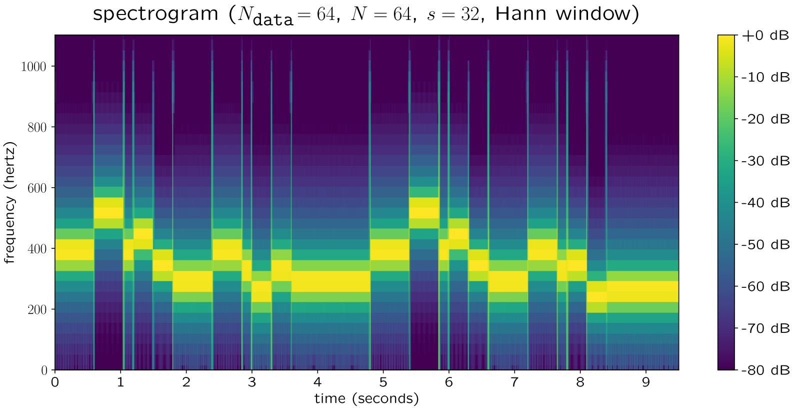
STFT: DFT Size



Decibels: $x_{\mathsf{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

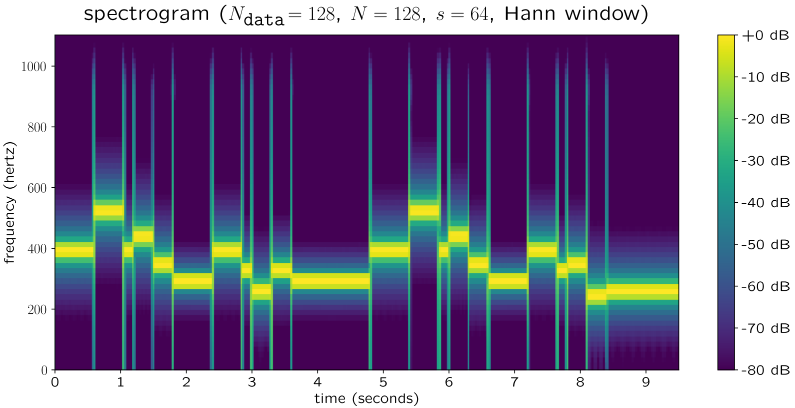
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

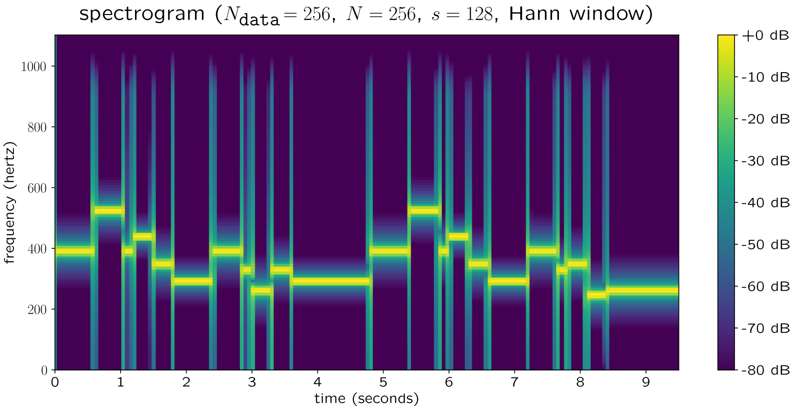
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

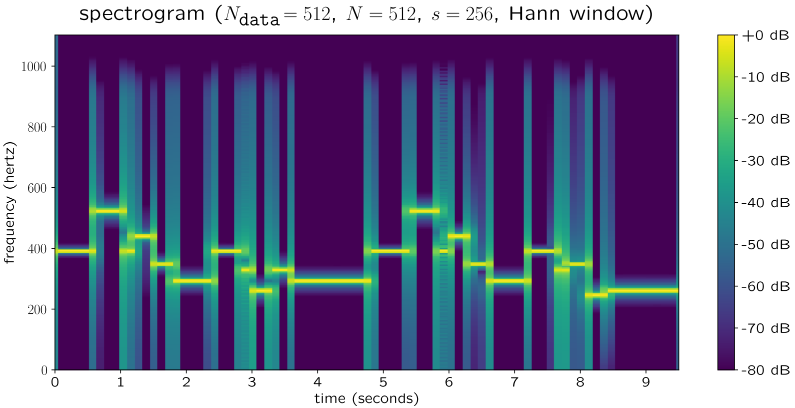
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

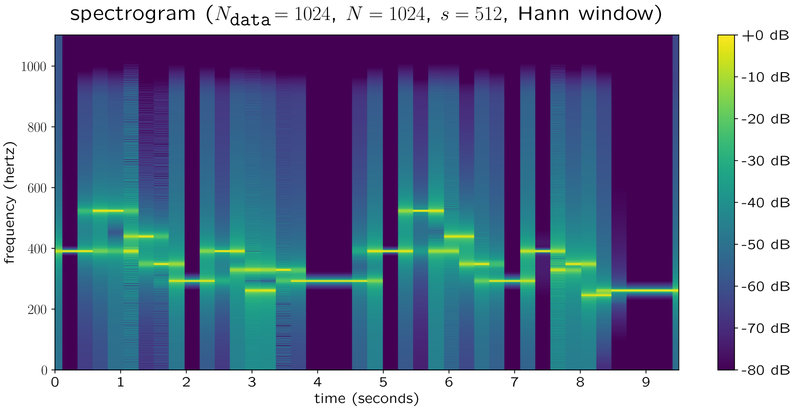
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

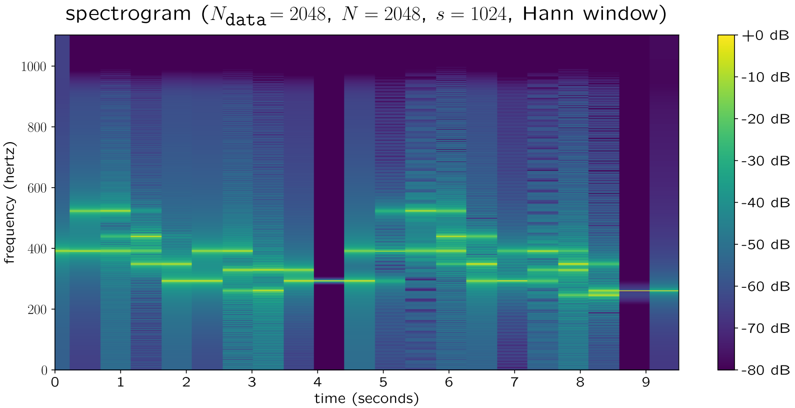
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

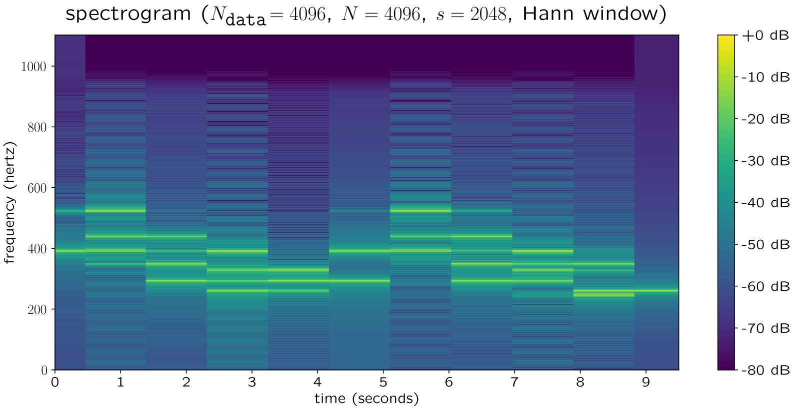
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

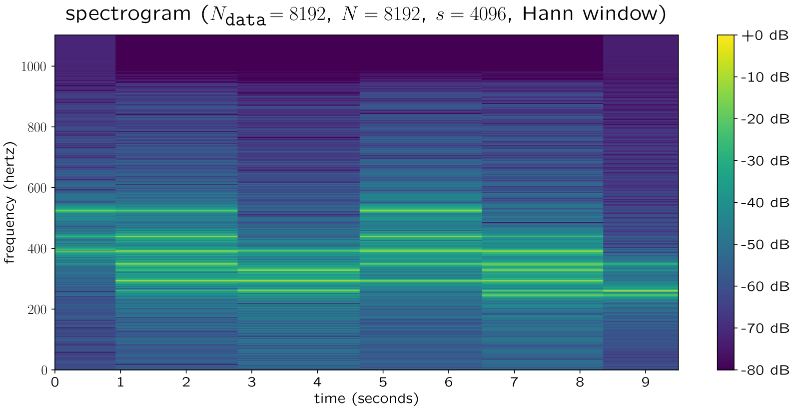
STFT: DFT Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

STFT: DFT Size



Decibels: $x_{\mathsf{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

STFT: Zero-Padding

Remember:

- N_{data} is the length of the data.
- N is the DFT size.

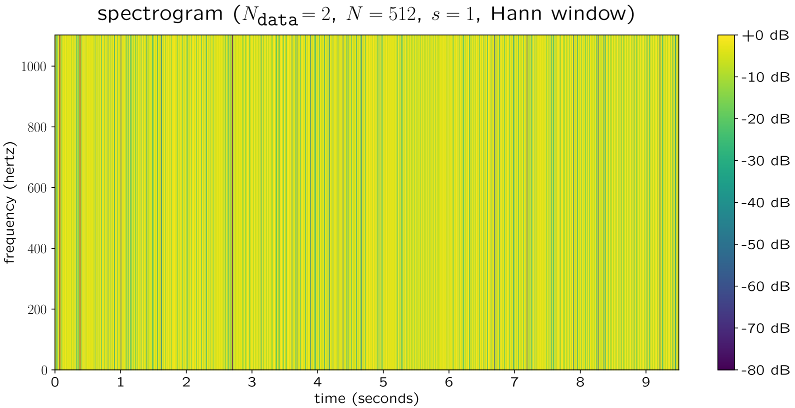
When $N > N_{\text{data}}$, we assume that we are **zero-padding**.

As N increases, we sample the underlying DTFT at more points — but this does not alter $W(\Omega)$, the DTFT of the window function. The width of $W(\Omega)$'s mainlobe ($\sim 1/N_{\text{data}}$) determines frequency resolution.

The more data we have, the larger N_{data} can be. As N_{data} increases, $W(\Omega)$ gets narrower — and the frequency resolution improves.

> You cannot improve frequency resolution by adding "zero information" — only by analyzing more data.

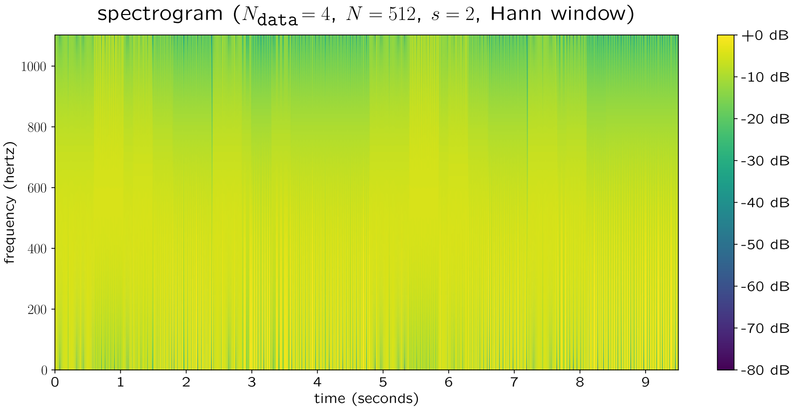
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

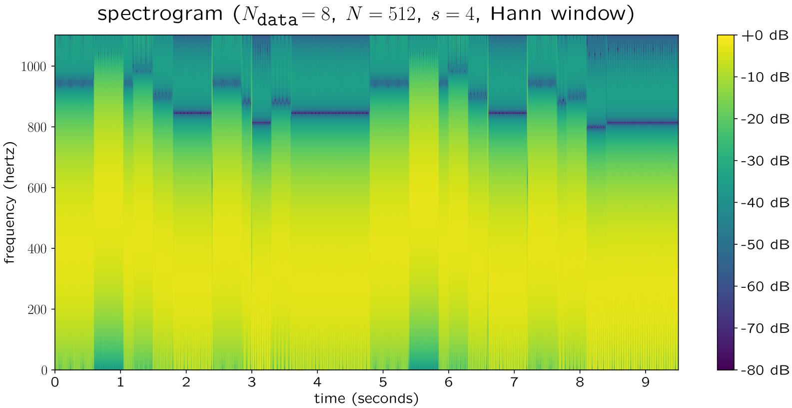
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

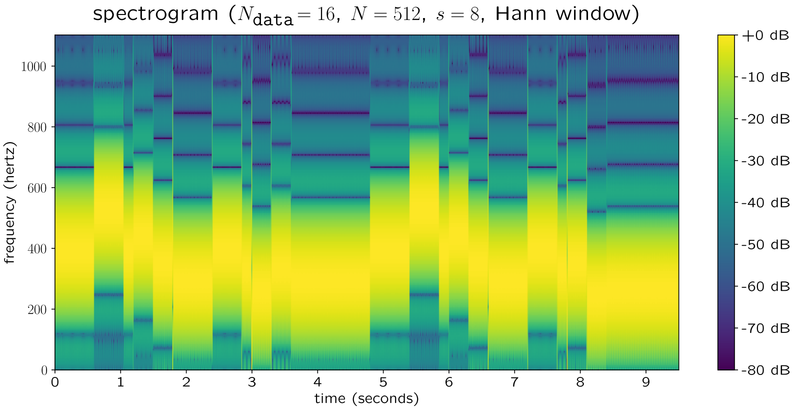
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

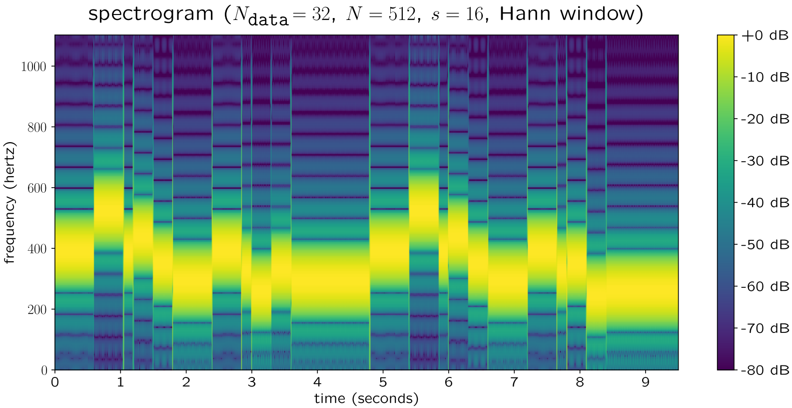
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

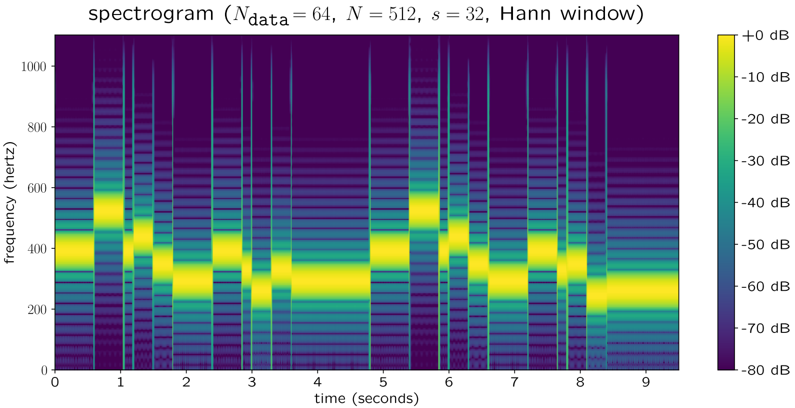
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

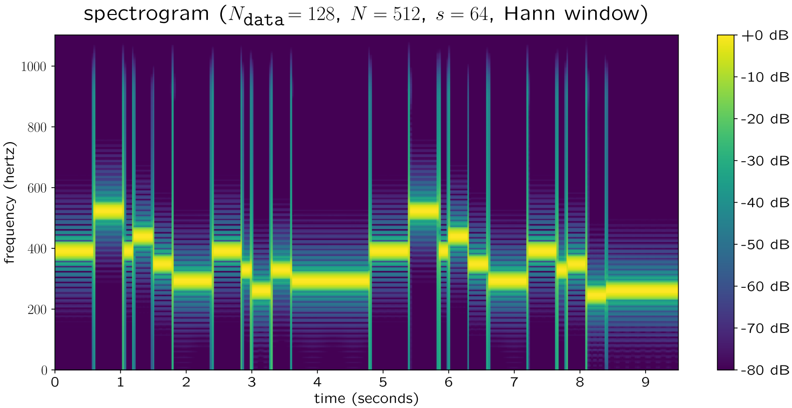
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

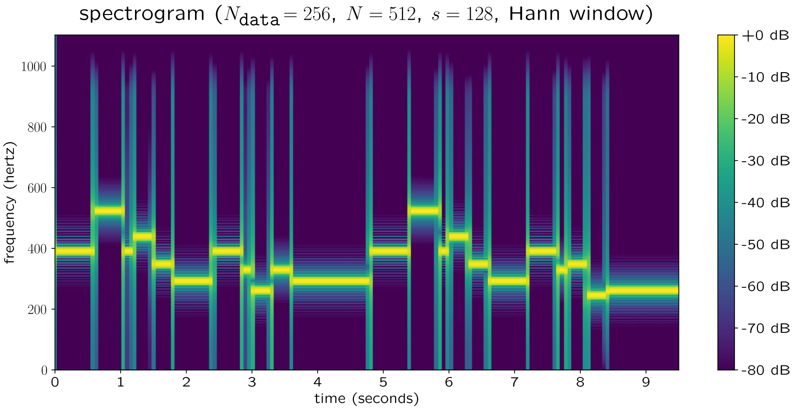
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

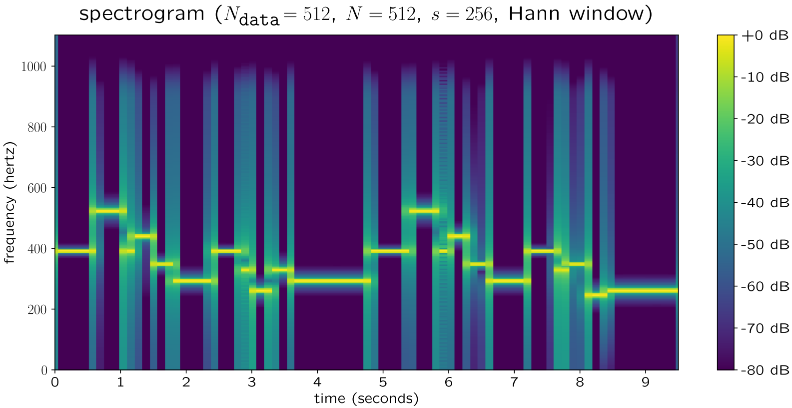
STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

STFT: Zero-Padding



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

Short-Time Fourier Transform

Parameters:

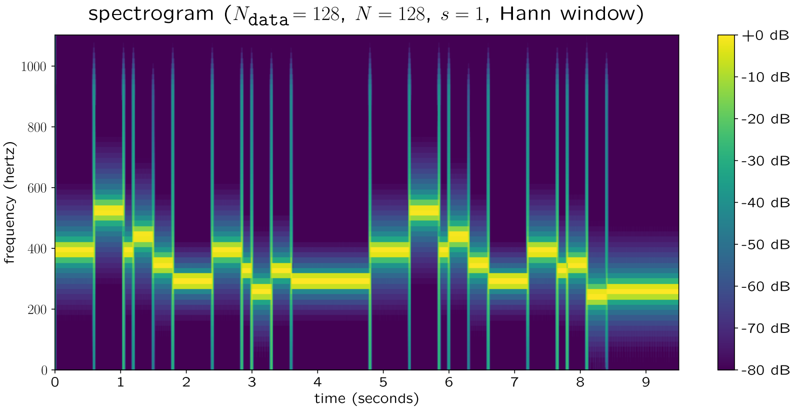
- type of window: $w[n]$
- DFT size: N
- step size: s

Short-Time Fourier Transform

Parameters:

- type of window: $w[n]$
- DFT size: N
- step size: s

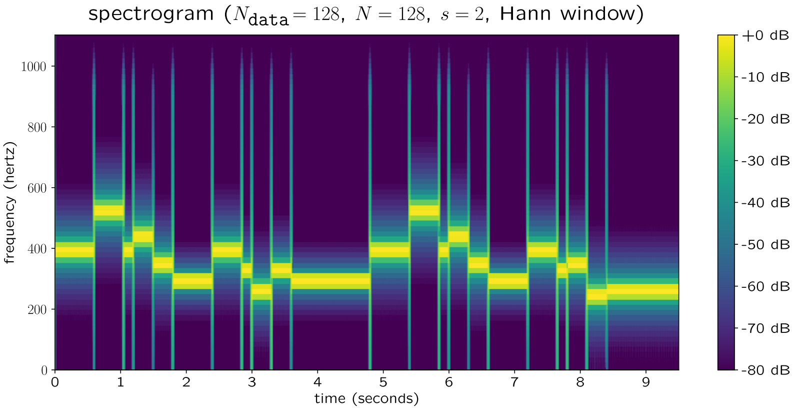
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

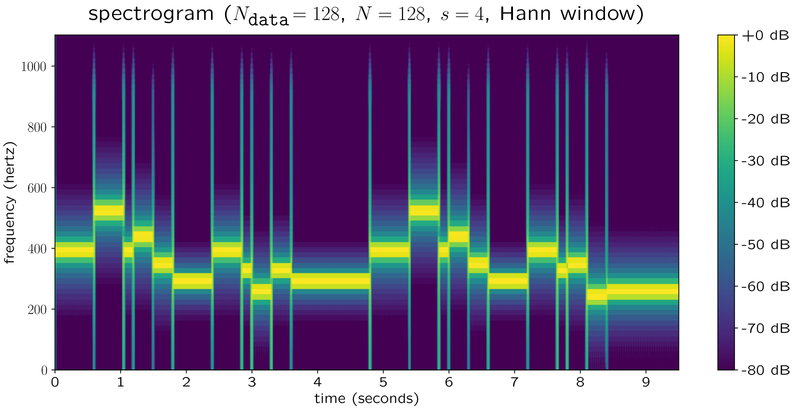
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

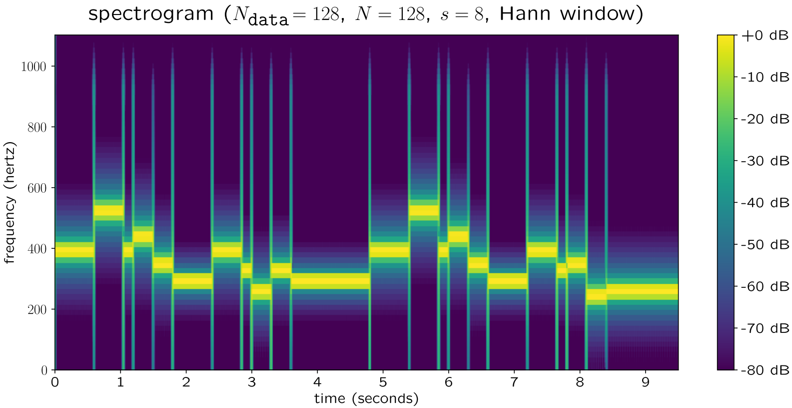
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

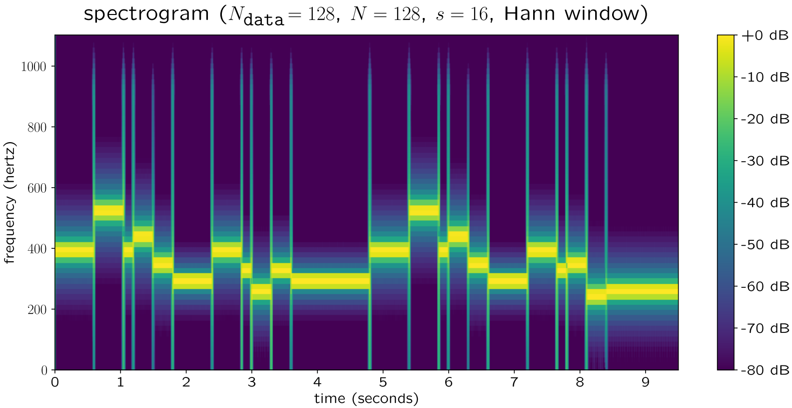
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

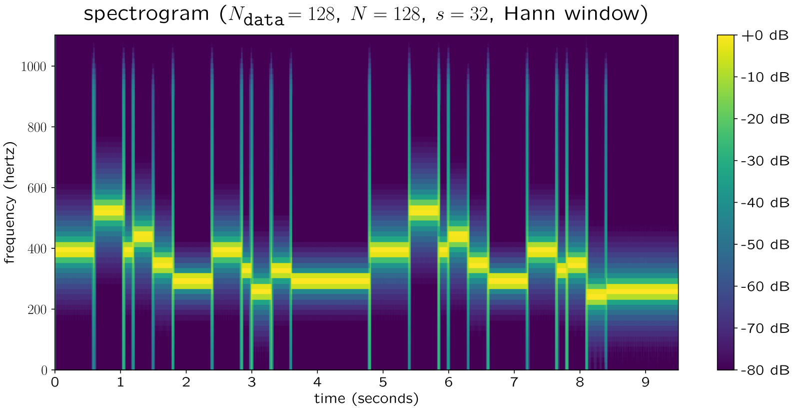
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

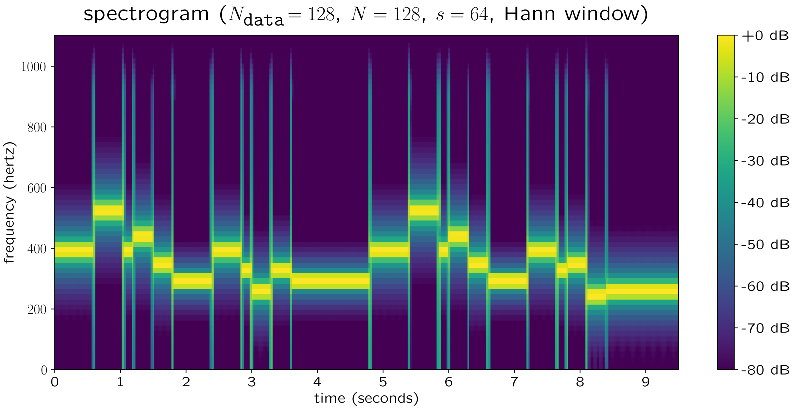
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

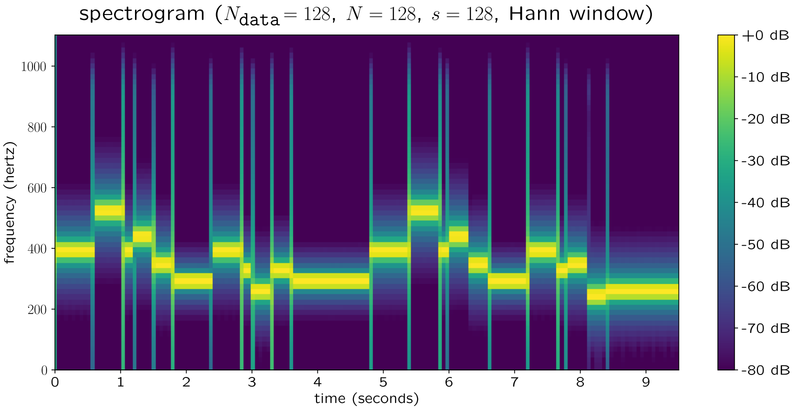
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

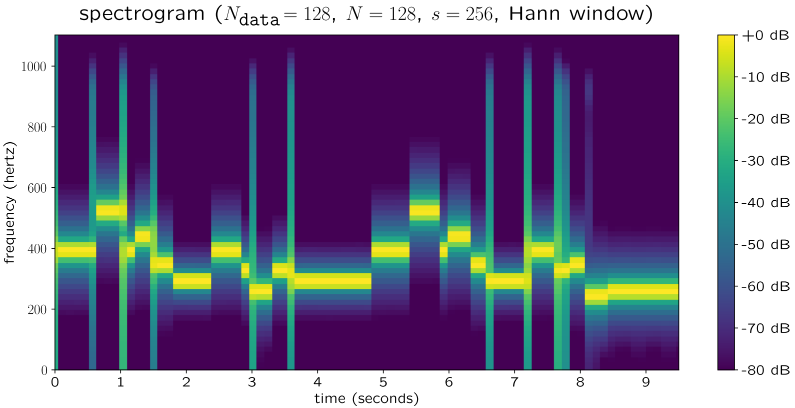
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

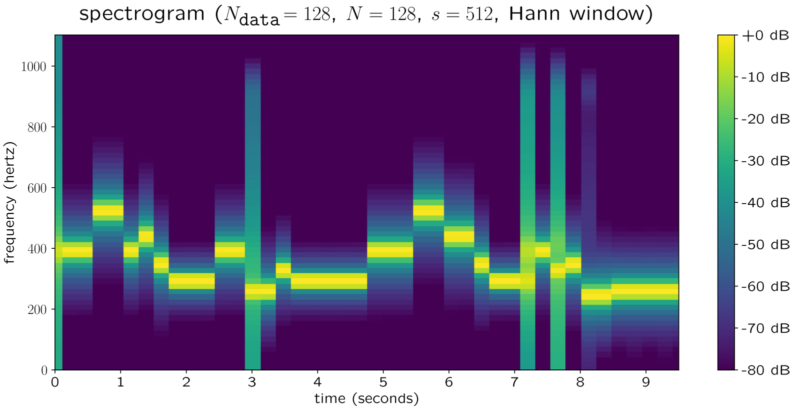
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

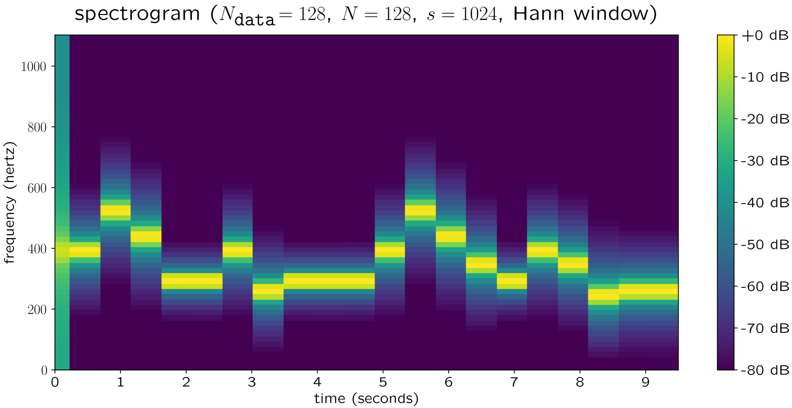
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

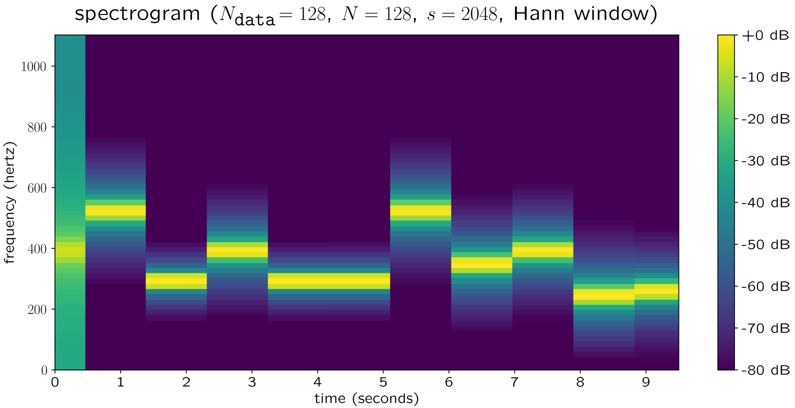
STFT: Step Size



Decibels: $x_{\text{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

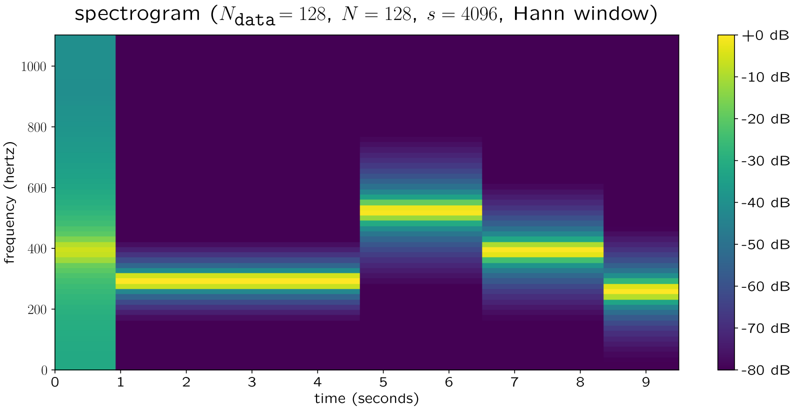
STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

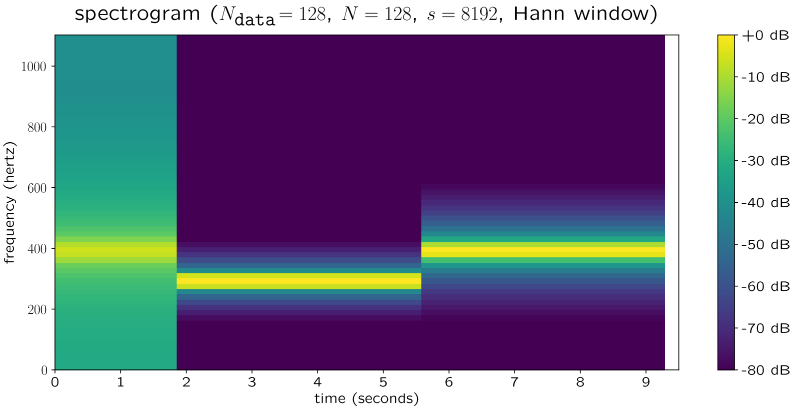
STFT: Step Size



Decibels: $x_{\mathsf{dB}} = 10 \log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

STFT: Step Size



Decibels: $x_{\text{dB}} = 10\log_{10}(x)$

e.g., -10 dB $\iff$ $x = 10^{-1}$ and -20 dB $\iff$ $x = 10^{-2}$

Short-Time Fourier Transform

Parameters:

- type of window: $w[n]$
- DFT size: N
- step size: s

Lessons Learned

A **short-time Fourier transform (STFT)** is a sequence of windowed Fourier transforms. A **spectrogram** displays the (squared) magnitude of an STFT.

Windows: Consider the main lobe and sidelobes.

Time Resolution vs. Frequency Resolution

- time resolution: short $w[n] \iff$ wide $W(\Omega)$
- frequency resolution: long $w[n] \iff$ narrow $W(\Omega)$

Analysis: $X[k,m] = \frac{1}{N}\sum_{n=0}^{N-1} x[n+ms]w[n]e^{-jk\frac{2\pi}{N}n}$

Spectrogram: $|X[k,m]|^2$ (magnitude^2 of STFT)

Question of the Day

Describe what each STFT parameter means.

- type of window: $w[n]$
- DFT size: N
- length of data: N_{data}
- step size: s