LEARNING FROM FREQUENCY DOMAIN DATA

Representation Learning from Frequency Domain Data

Reminders and Announcements

- Look for HW2 on the class webpage (will be out by 11:59pm).
- Your 2-page project proposal is due (by email to me, as a PDF attachment) by 2/26.
- For your presentations, presentation guidance was posted on Piazza
  - Search for Piazza note @45
- Concluding remarks on Debate #1 were posted on Piazza
  - Search for Piazza note @48
- Debate #2 will occur in class today.

Do not edit
How to change the design

Join at slido.com
#3175898

ⓘ The Slido app must be installed on every computer you're presenting from

slido

Do not edit
How to change the design

In IoT contexts, a transformer-based encoder architecture will generally outperform smaller and simpler neural networks (that feature a combination of convolutional and recurrent layers) at creating representations suitable for downstream classification tasks.

Today’s Debate:

In IoT contexts, a transformer-based encoder architecture will generally outperform smaller and simpler neural networks (that feature a combination of convolutional and recurrent layers) at creating representations suitable for downstream classification tasks.

Part I: Group #7 Presents “Pro” Arguments

Jingjie He

Kevin Lin

Ricky Wang

Part II: Group #8 Presents “Anti” Arguments

McKenzy Heavlin

Maciek Hryciuk

Part III: Moderated Open Discussion: All are welcome to join in…

Do not edit
How to change the design

Join at slido.com
#3175898

ⓘ The Slido app must be installed on every computer you're presenting from

slido

In IoT contexts, a transformer-based encoder architecture will generally outperform smaller and simpler neural networks (that feature a combination of convolutional and recurrent layers) at creating representations suitable for downstream classification tasks.

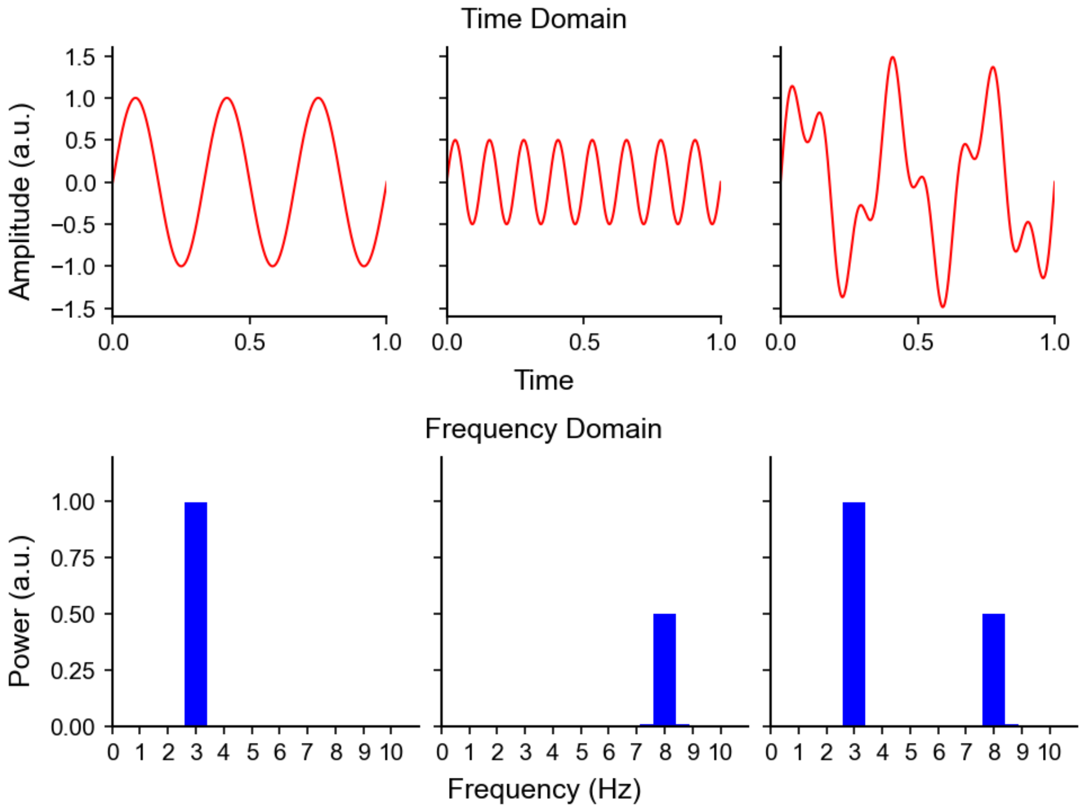
Today's Goal

Understand key intuitions about frequency domain signals that allow you to design better representation learning solutions for data in the frequency domain (e.g., better encoder architectures, better pretext/training tasks, etc)

What Is the Frequency Domain?

Reminder:

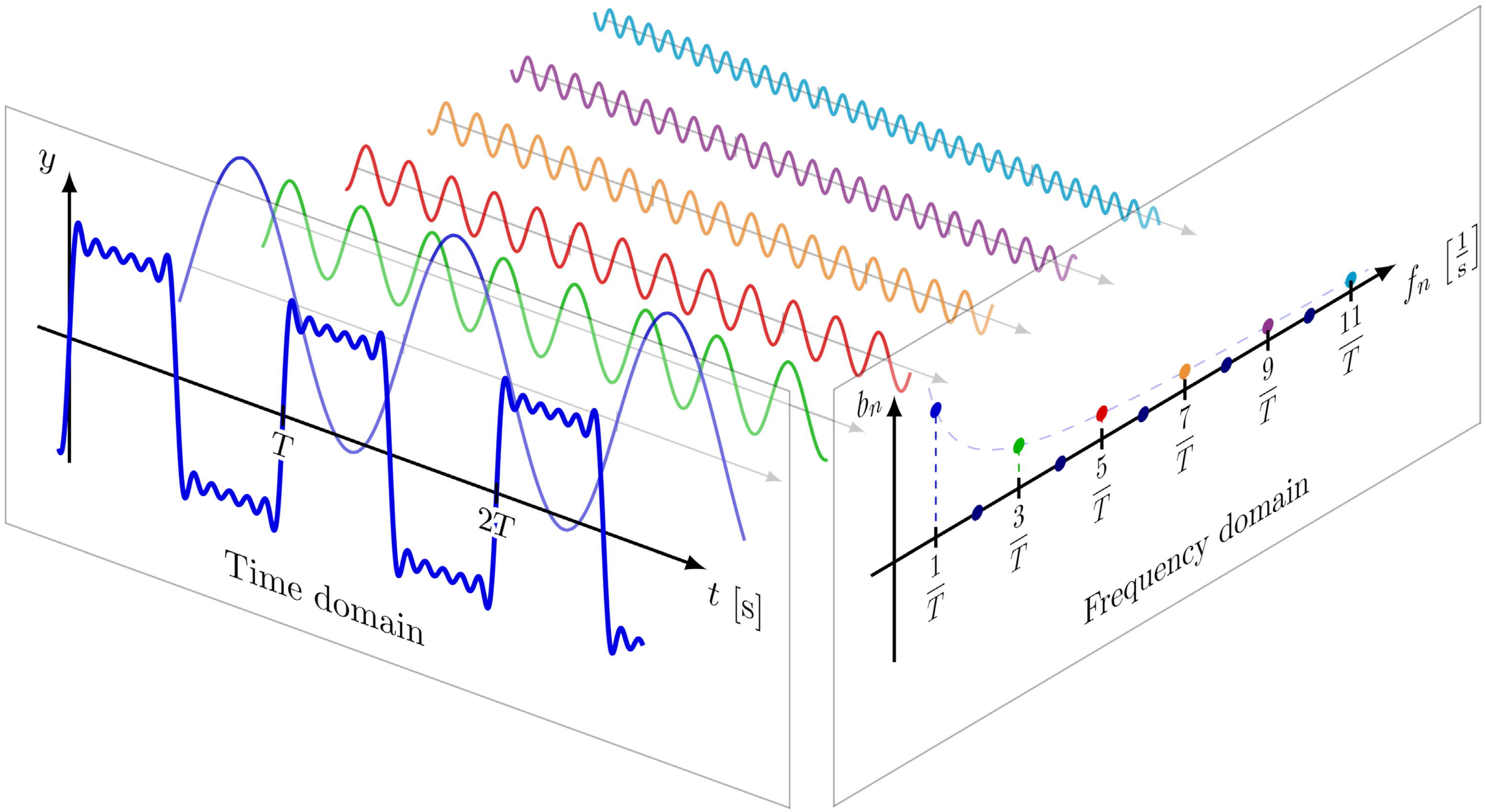
- Any periodic function of time (a time-domain signal) can be represented as a summation of sinusoids of different amplitudes and frequencies.
- The frequency domain representation of a time-domain signal plots the amplitudes of its component sinusoids versus their frequencies.



https://dibsmethodsmeetings.github.io/fourier-transforms/

What Is the Frequency Domain?

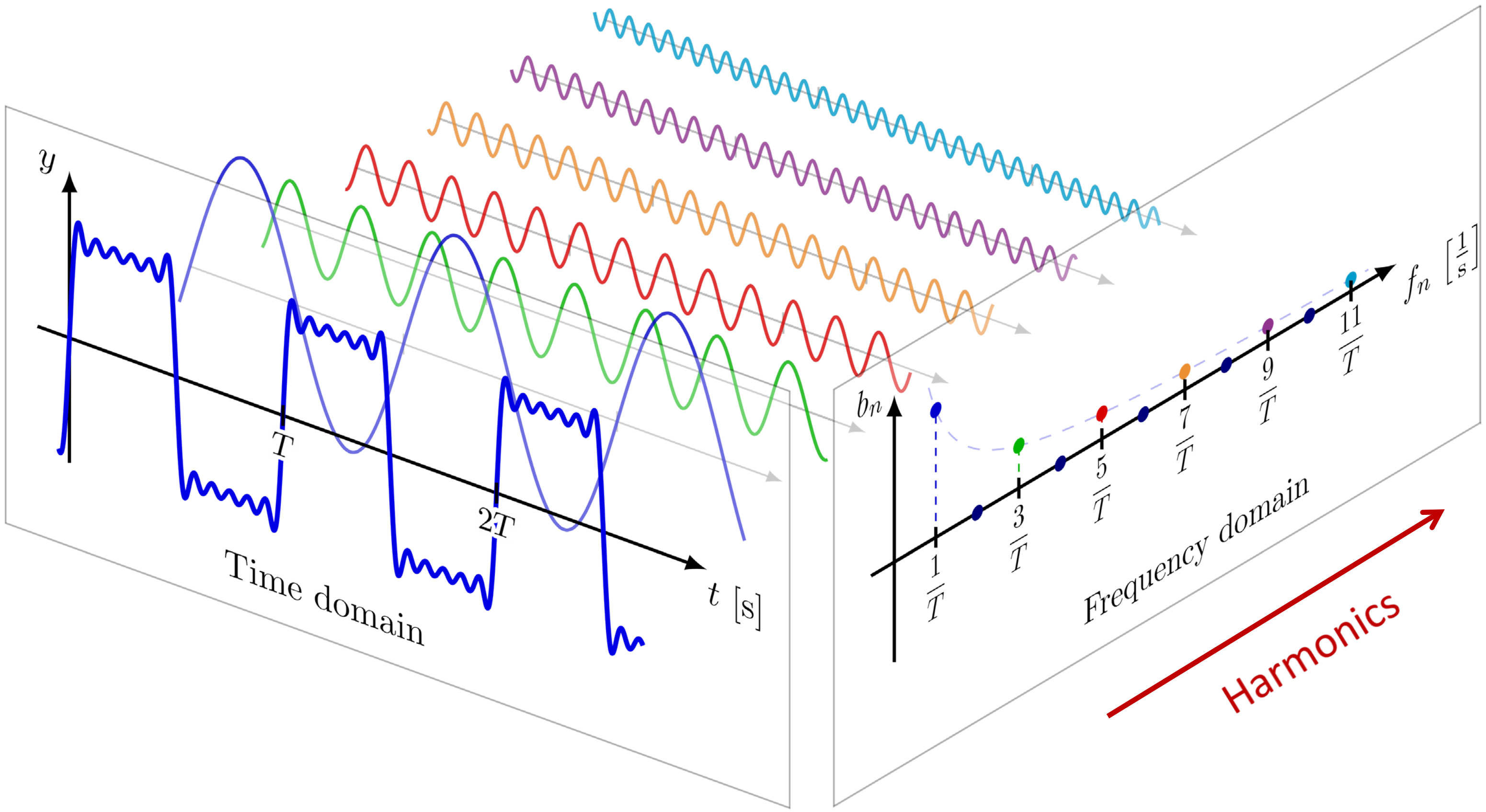
Intuition: Higher-frequency component sinusoids are needed to paint finer detail of the time-domain signal, whereas lower-frequency component sinusoids account for smoother components of the original signal



https://dibsmethodsmeetings.github.io/fourier-transforms/

Harmonics

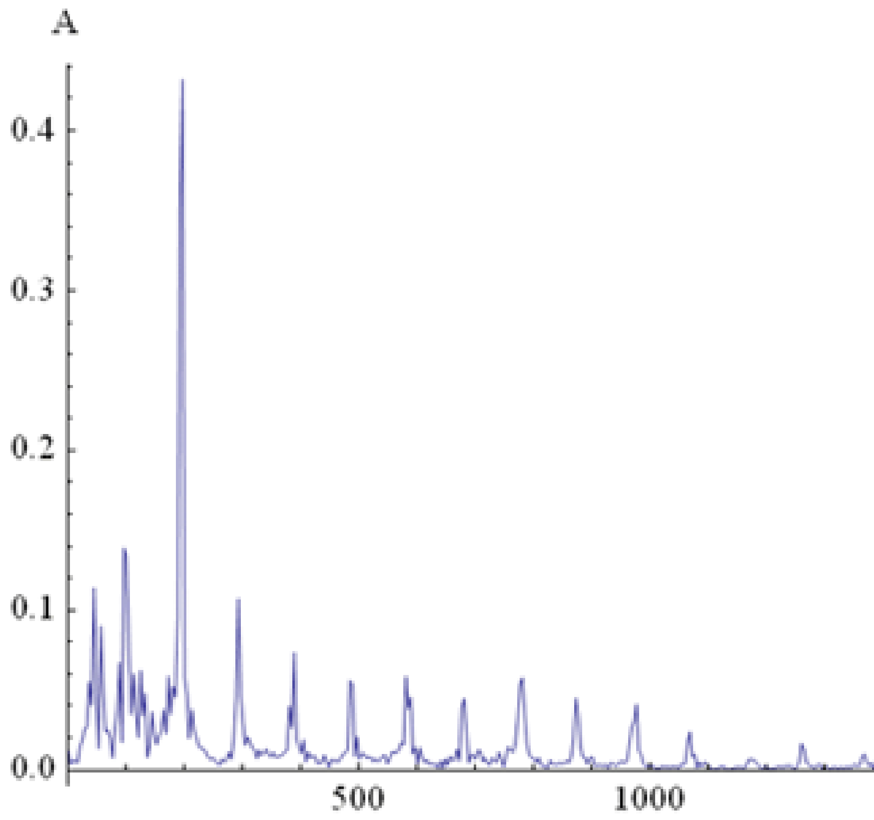
A periodic signal of period T (i.e., frequency $f = 1/T$) in the time domain is decomposed into sinusoids whose frequencies are f and its multiples ($2f$, $3f$, $4f$, …). They are called the base frequency and harmonics, respectively.



https://dibsmethodsmeetings.github.io/fourier-transforms/

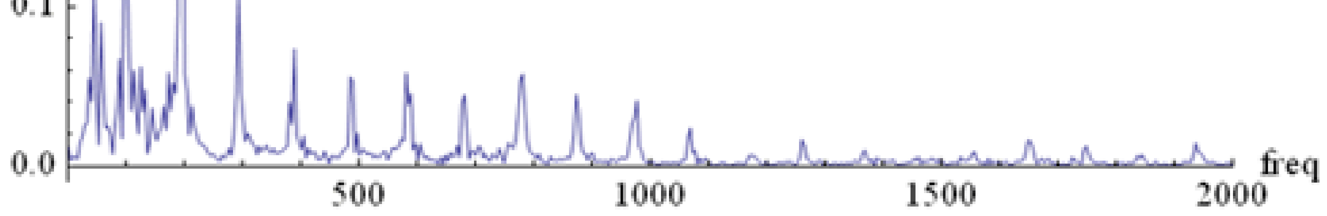
Harmonics
Example: A Cello Note

Example: A music note (e.g., of a string instrument such as a cello) is generated by string vibrations at a particular frequency. Since the vibration is a periodic movement, but is not exactly sinusoidal, it gives rise to several harmonics that are multiples of the base frequency of the vibration.



Harmonics
Example: A Cello Note

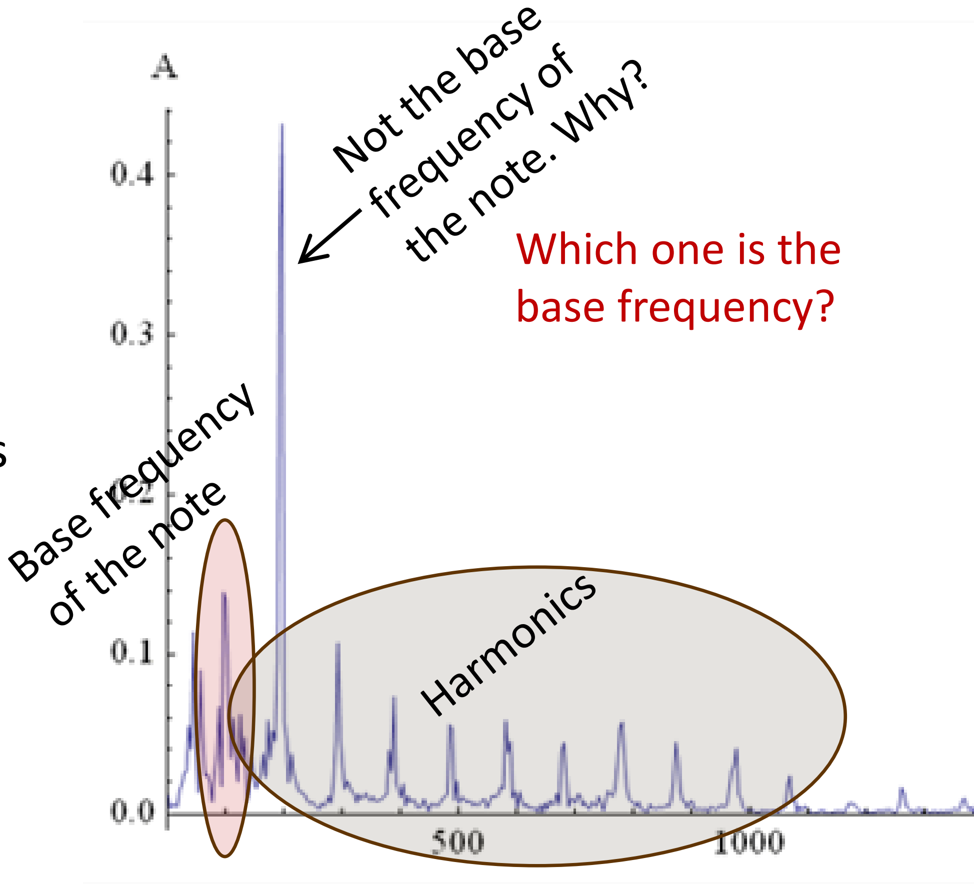
Example: A music note (e.g., of a string instrument such as a cello) is generated by string vibrations at a particular frequency. Since the vibration is a periodic movement, but is not exactly sinusoidal, it gives rise to several harmonics that are multiples of the base frequency of the vibration.



Which one is the base frequency?

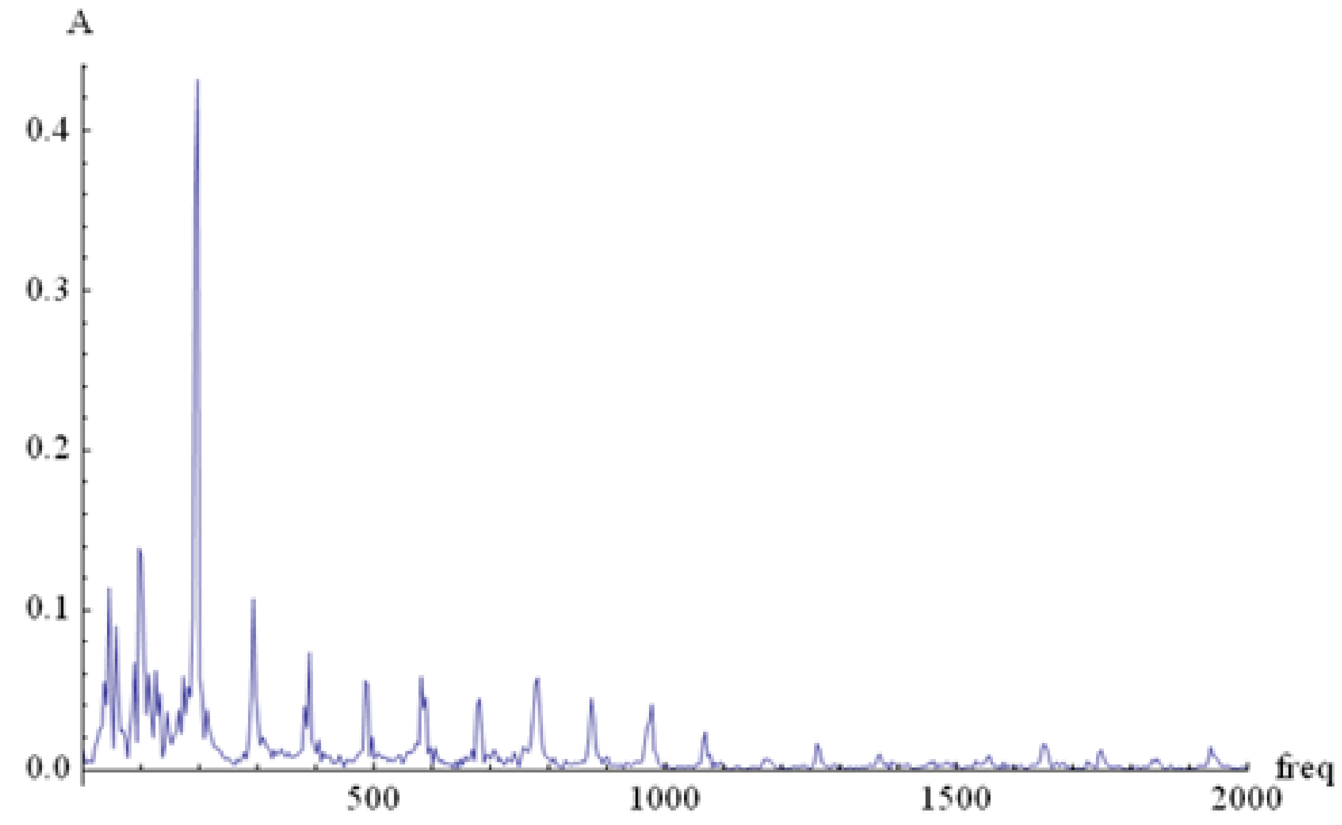
Harmonics
Example: A Cello Note

Example: A music note (e.g., of a string instrument such as a cello) is generated by string vibrations at a particular frequency. Since the vibration is a periodic movement, but is not exactly sinusoidal, it gives rise to several harmonics that are multiples of the base frequency of the vibration.



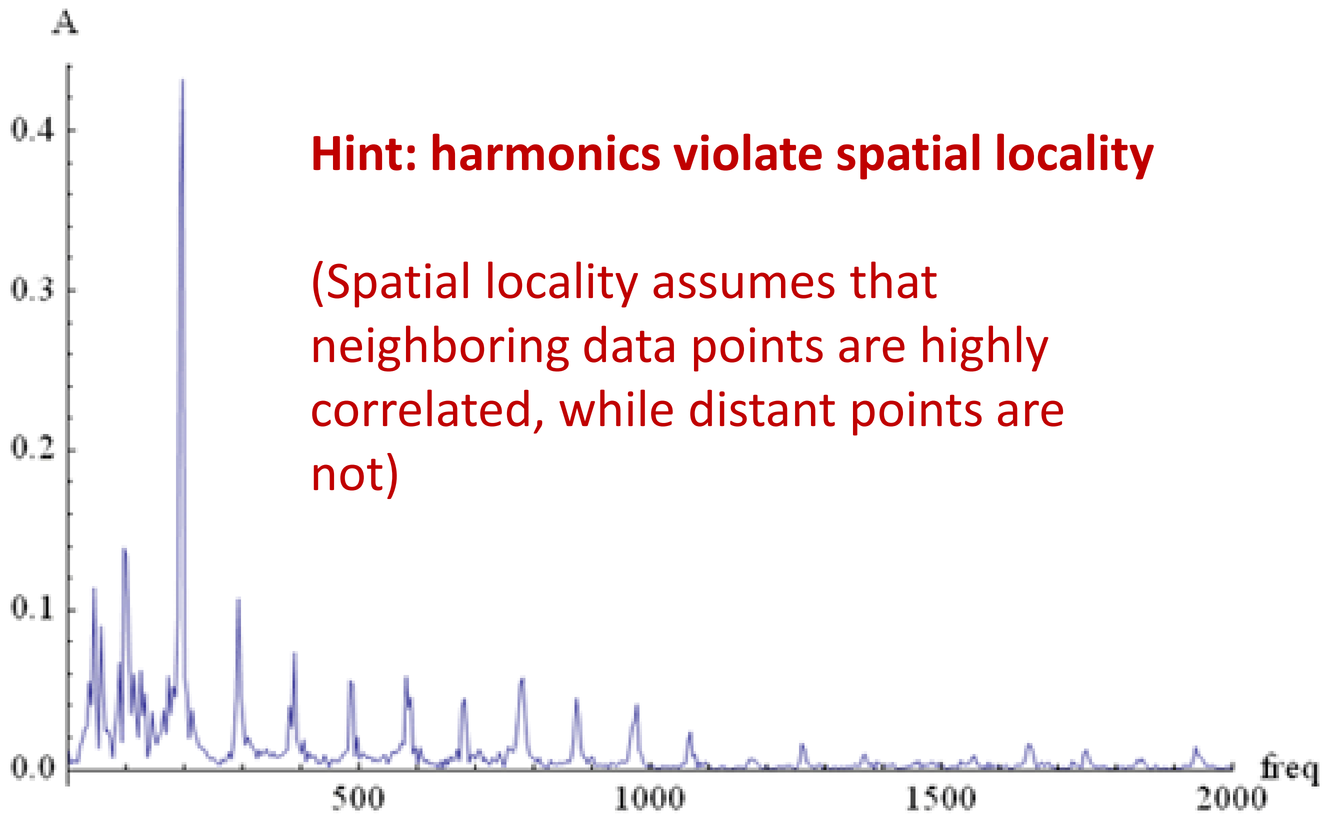
Harmonics
Implications on Representation Learning

Design question: What does the existence of harmonics imply in terms of encoder architecture and similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?



Harmonics
Implications on Representation Learning

Design question: What does the existence of harmonics imply in terms of encoder architecture and similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?

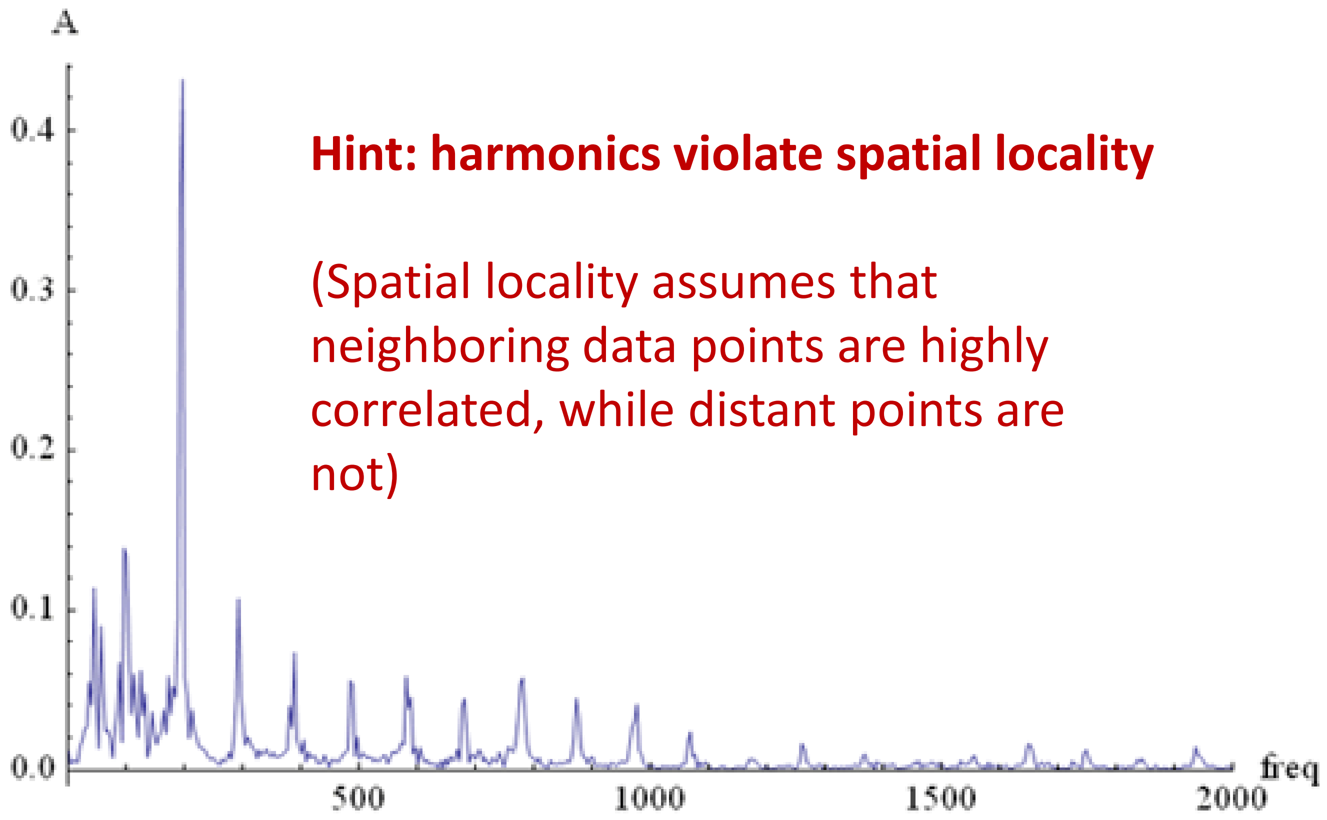


Hint: harmonics violate spatial locality

(Spatial locality assumes that neighboring data points are highly correlated, while distant points are not)

Harmonics
Implications on Representation Learning

Design question: What does the existence of harmonics imply in terms of encoder architecture and similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?



Hint: harmonics violate spatial locality

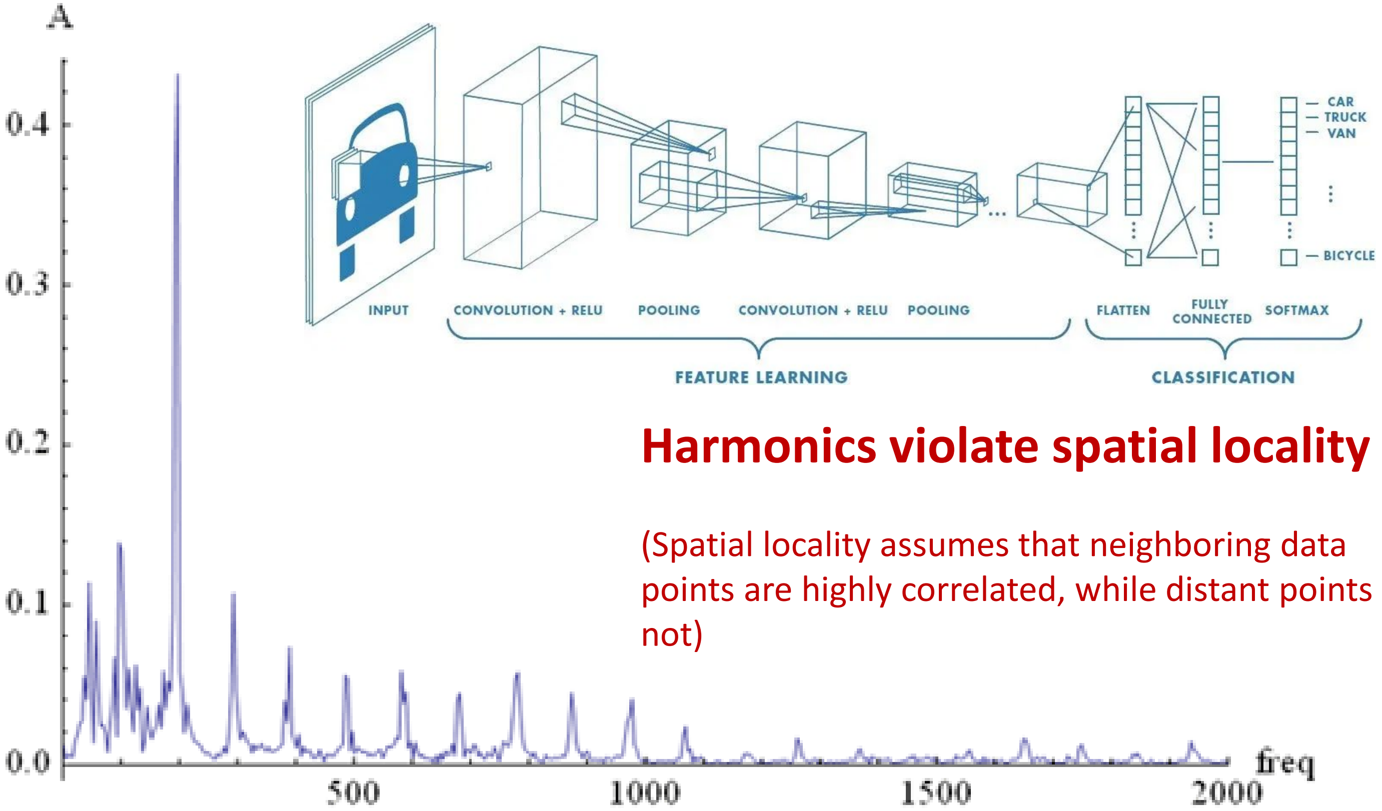
(Spatial locality assumes that neighboring data points are highly correlated, while distant points are not)

Need new self-attention concepts that are better at capturing harmonic patterns (not just local patterns)

Harmonics
Implications on Representation Learning

Would convolutional neural networks make good encoders of frequency domain data?

Design question: What does the existence of harmonics imply in terms of encoder architecture and similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?



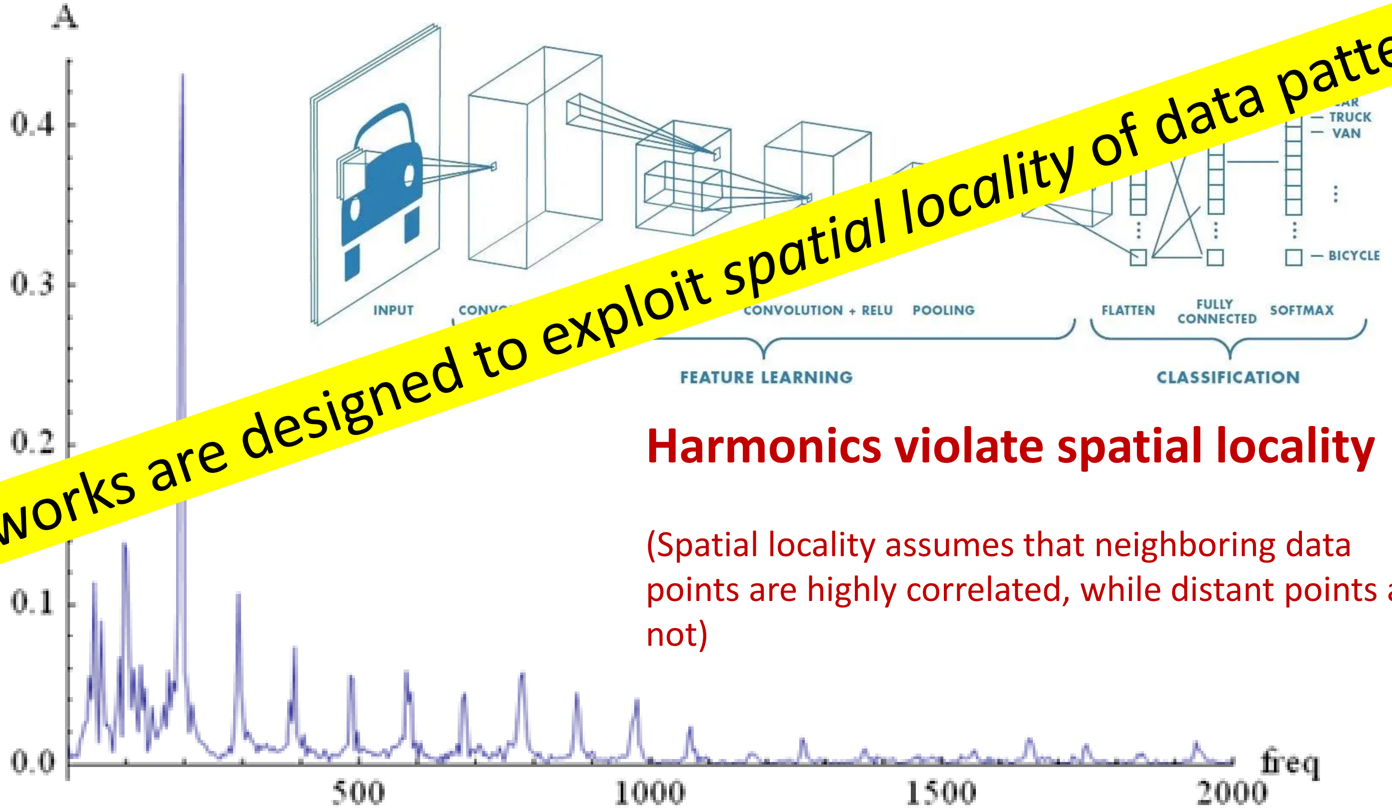
Harmonics violate spatial locality

(Spatial locality assumes that neighboring data points are highly correlated, while distant points are not)

Harmonics
Implications on Representation Learning

Would convolutional neural networks make good encoders of frequency domain data? → **Not really**

Design question: What does the existence of harmonics imply in terms of encoder architecture and similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive lea

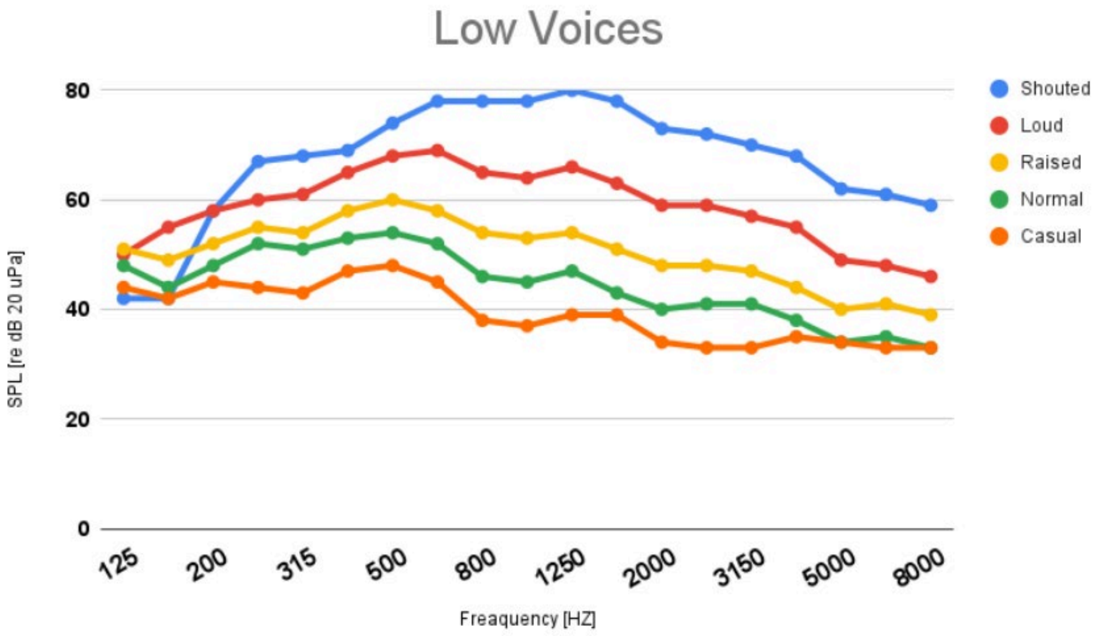


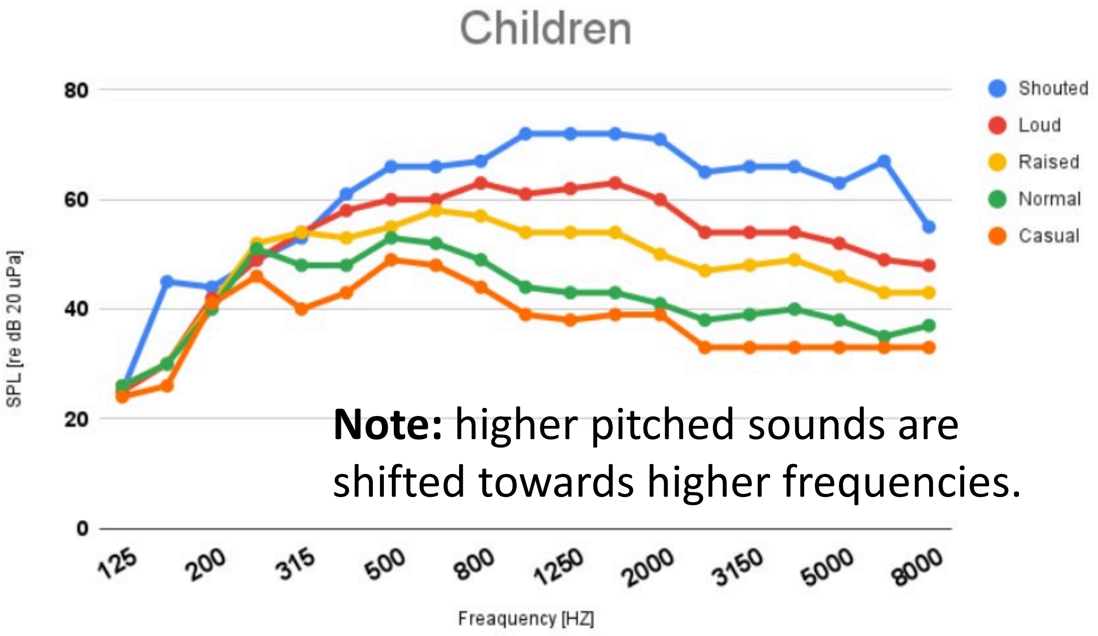
Hint: Convolutional neural networks are designed to exploit spatial locality of data patterns

Harmonics violate spatial locality

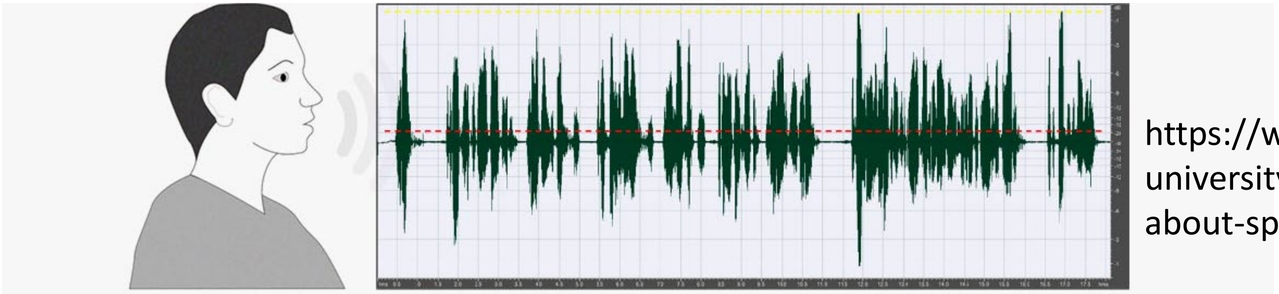
(Spatial locality assumes that neighboring data points are highly correlated, while distant points are not)

More Complex Waveform Example: Human Voices in the Frequency Domain





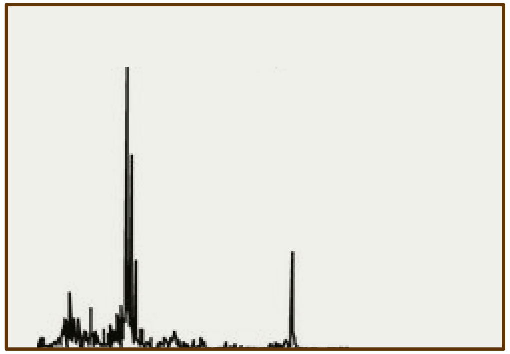
Note: higher pitched sounds are shifted towards higher frequencies.

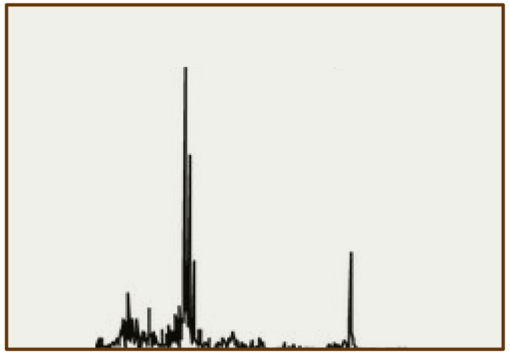


https://www.dpamicrophones.com/mic-university/background-knowledge/facts-about-speech-intelligibility/

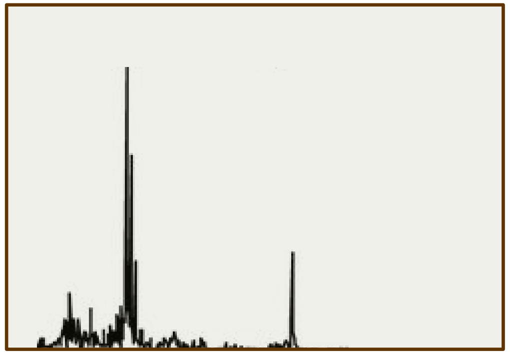
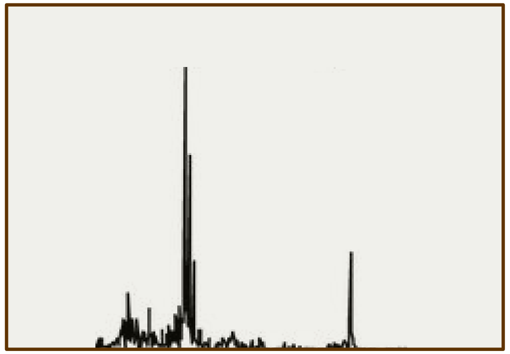
Implication on Similarity Metrics in the Frequency Domain

Similar?





Implication on Similarity Metrics in the Frequency Domain

Similar?

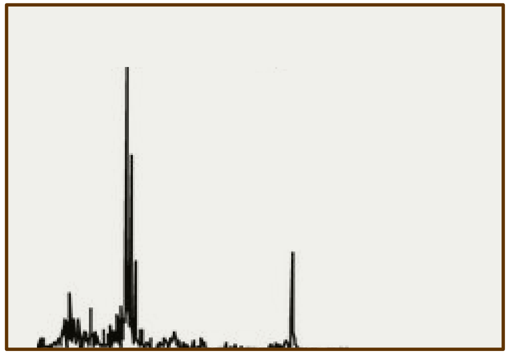
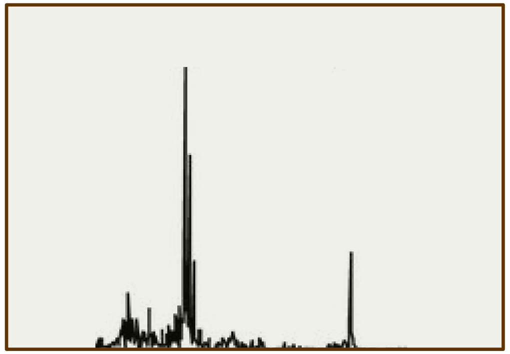
YES

Similar?

Implication on Similarity Metrics in the Frequency Domain


Similar?

YES

Similar?

NO

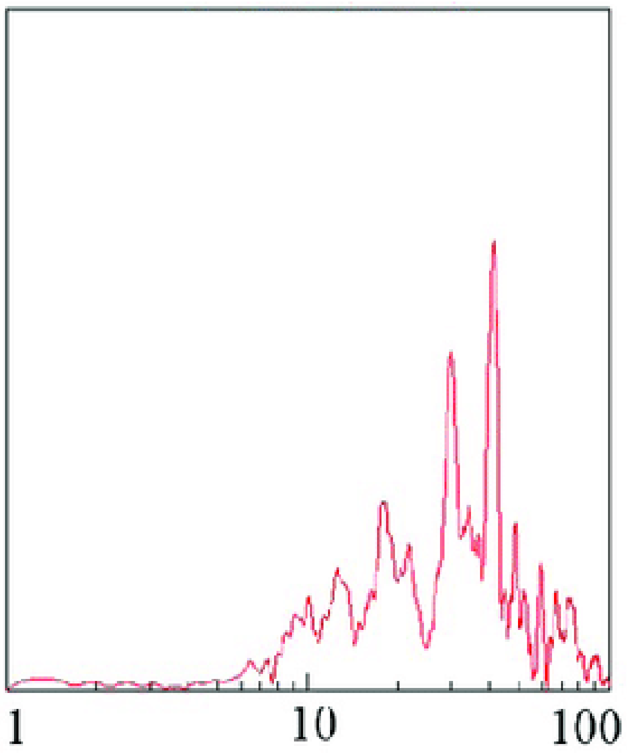
In the frequency domain, signal semantics do **not** exhibit location invariance!

Recall how children's voices looked like adult voices that are shifted right in the frequency domain, but the sounds are very different.

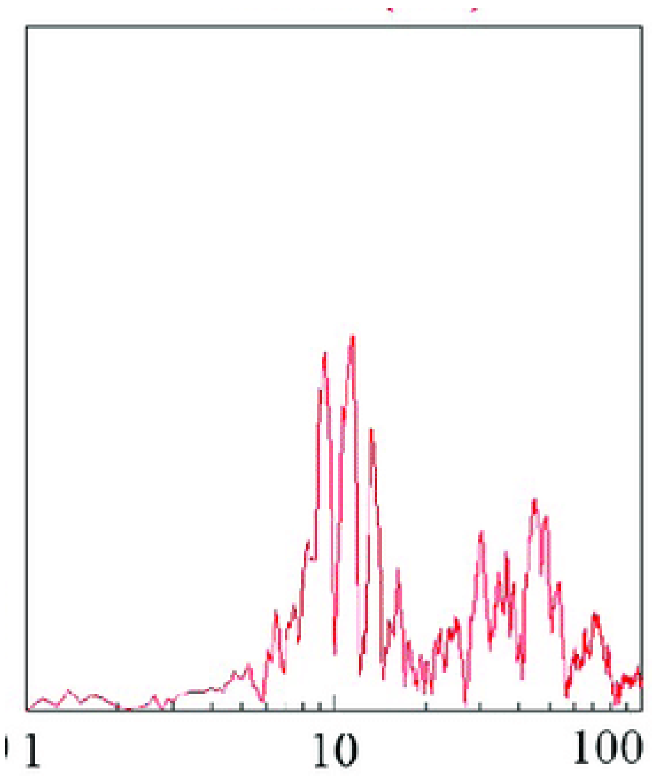
Contrastive learning needs different similarity metrics

"Decoding" Frequency Domain Signals Intuitions: Which Car Is Going Faster?

If the Fourier Transforms shown to the right are of cars of the same make and model going at different speeds. Which car is going faster, Car A or Car B?



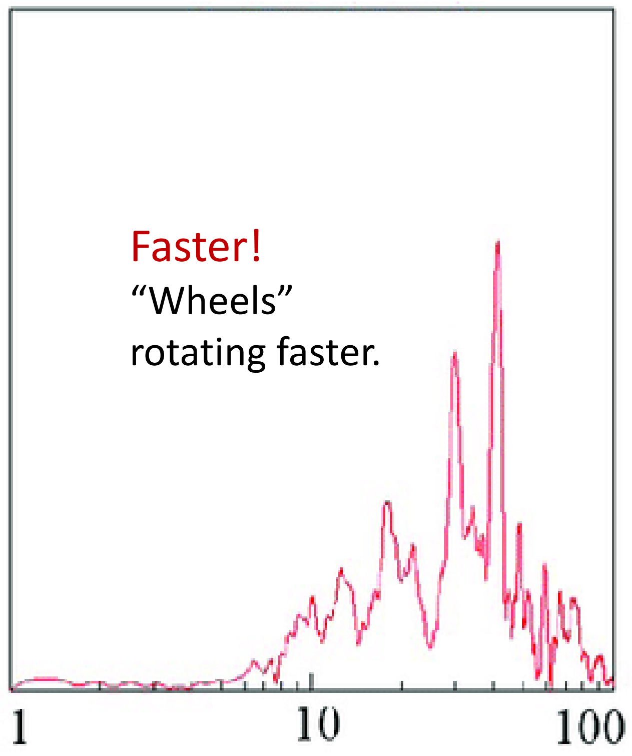
Car A



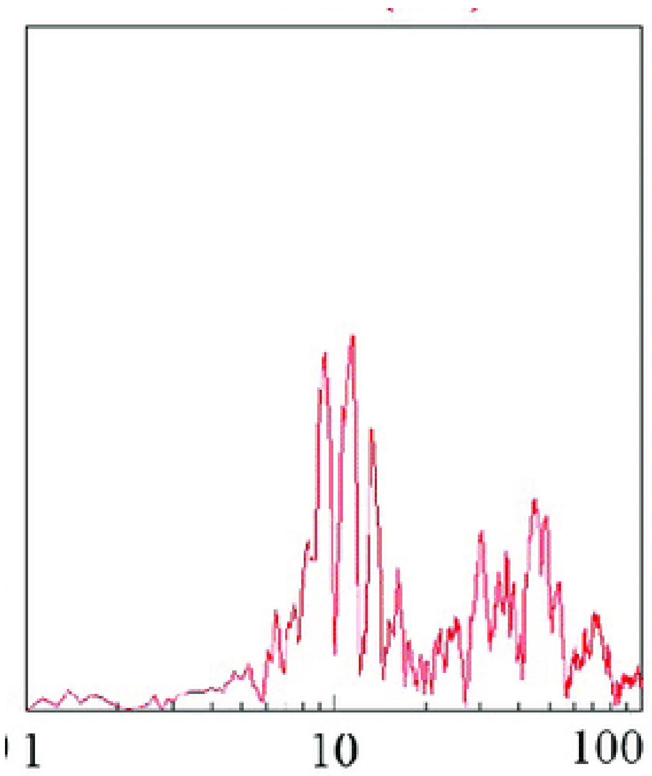
Car B

"Decoding" Frequency Domain Signals Intuitions: Which Car Is Going Faster?

If the Fourier Transforms shown to the right are of cars of the same make and model going at different speeds. Which car is going faster, Car A or Car B?



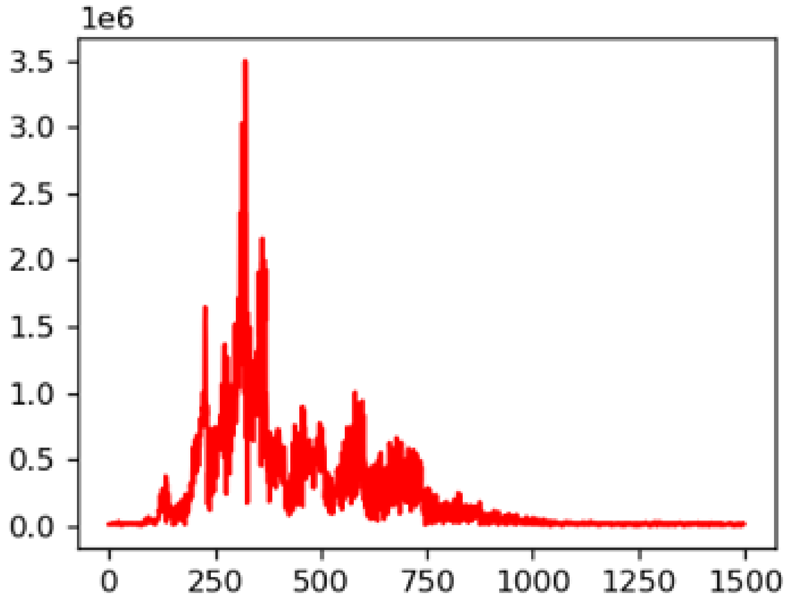
Car A



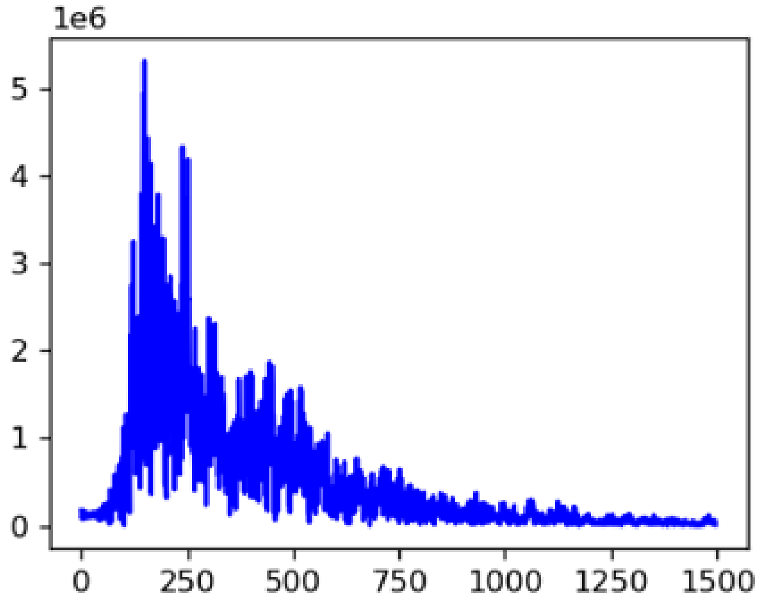
Car B

"Decoding" Frequency Domain Signals Intuitions: Which Car Is Bigger?

Alternatively, if the Fourier Transforms shown to the right are of two cars moving roughly the same speed, which car is likely bigger, Car A or Car B?



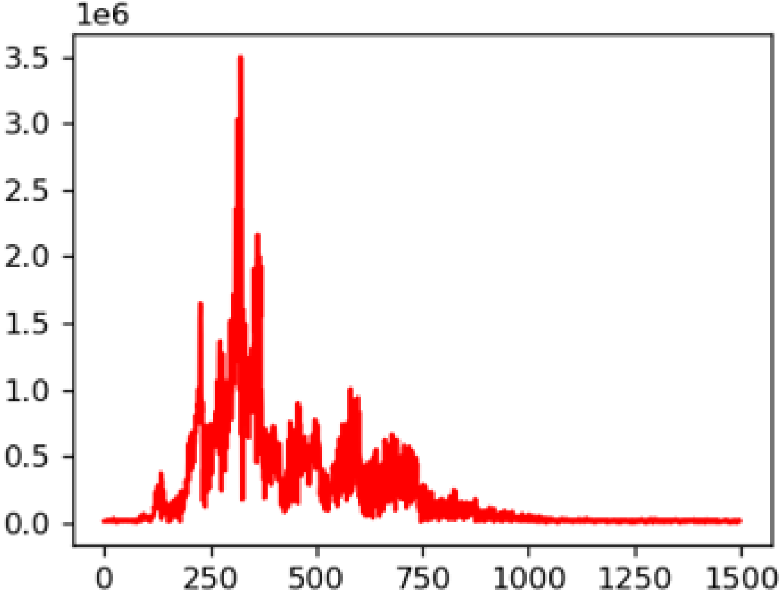
Car A



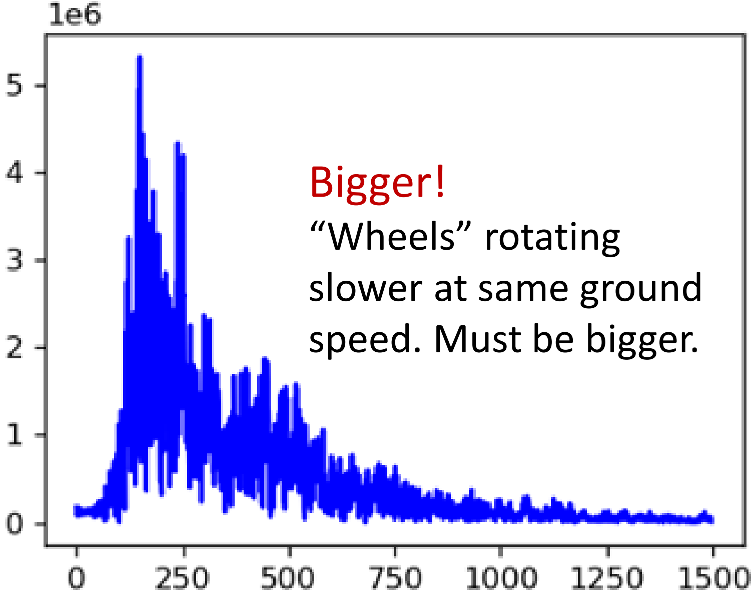
Car B

"Decoding" Frequency Domain Signals Intuitions: Which Car Is Bigger?

Alternatively, if the Fourier Transforms shown to the right are of two cars moving roughly the same speed, which car is likely bigger, Car A or Car B?



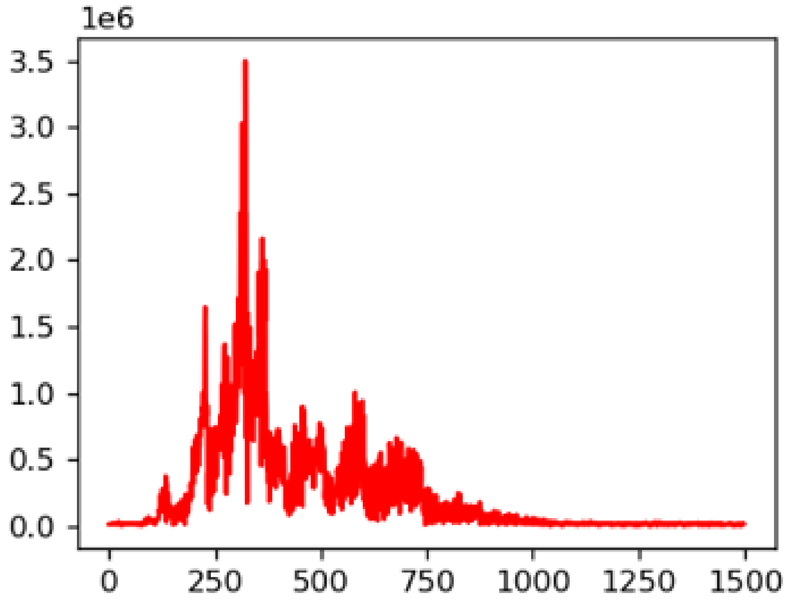
Car A



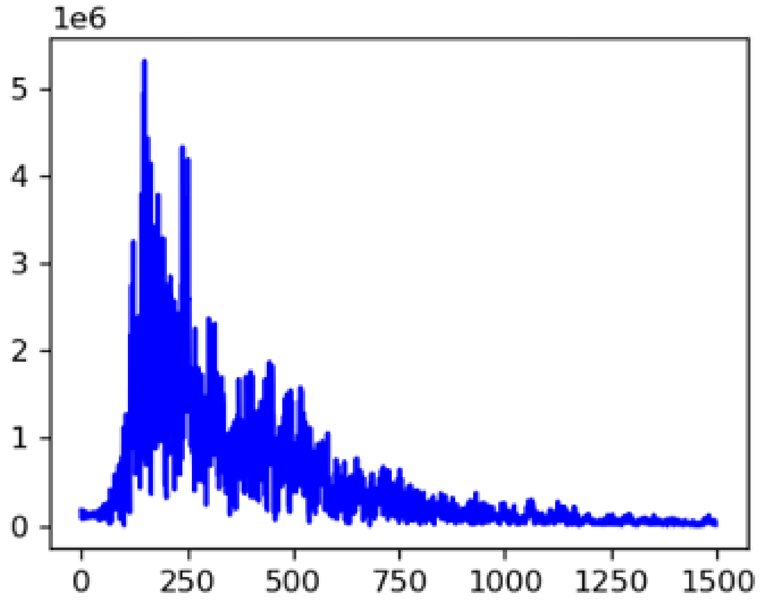
Car B

Intuitions: Which Car Has a Better Suspension System?

Alternatively, if the Fourier Transforms shown to the right are of two cars moving roughly the same speed, which car has a better suspension system, Car A or Car B?



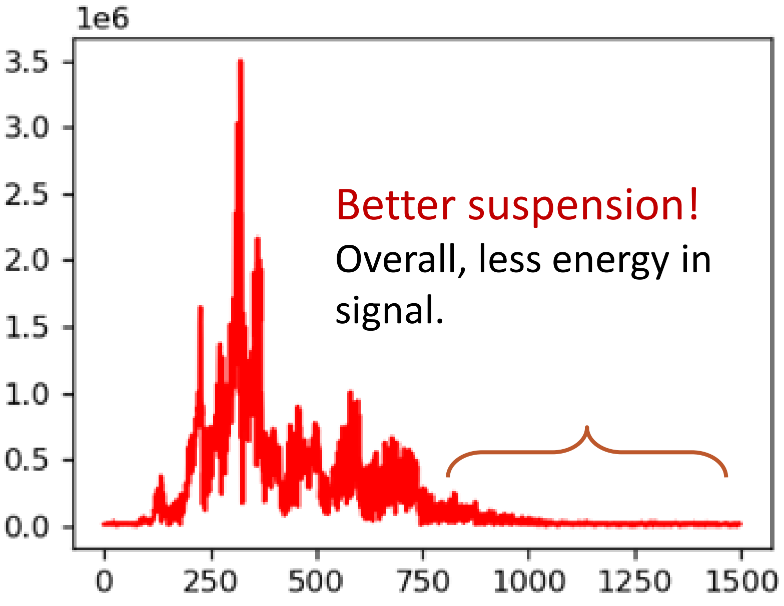
Car A



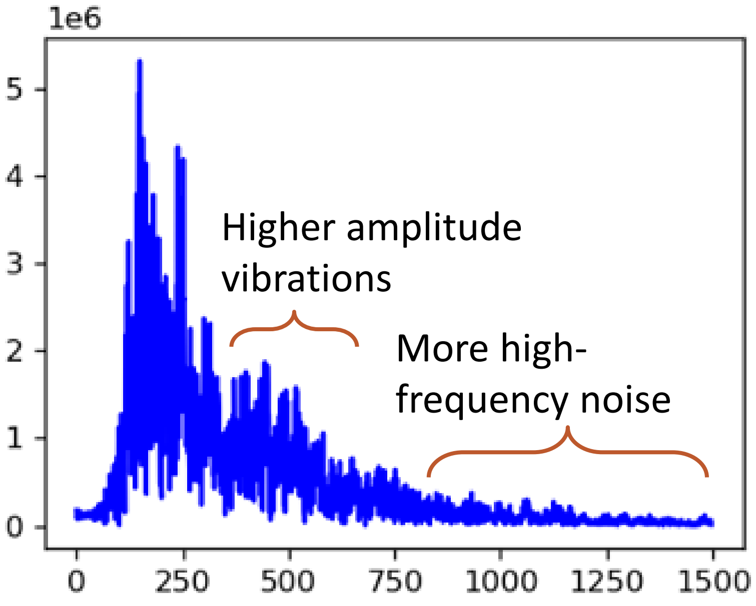
Car B

Intuitions: Which Car Has a Better Suspension System?

Alternatively, if the Fourier Transforms shown to the right are of two cars moving roughly the same speed, which car has a better suspension system, Car A or Car B?



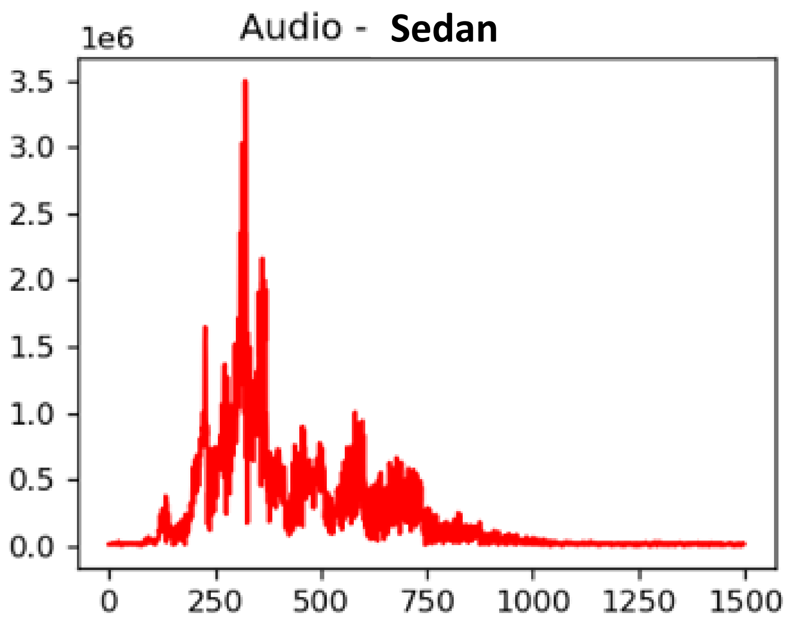
Car A



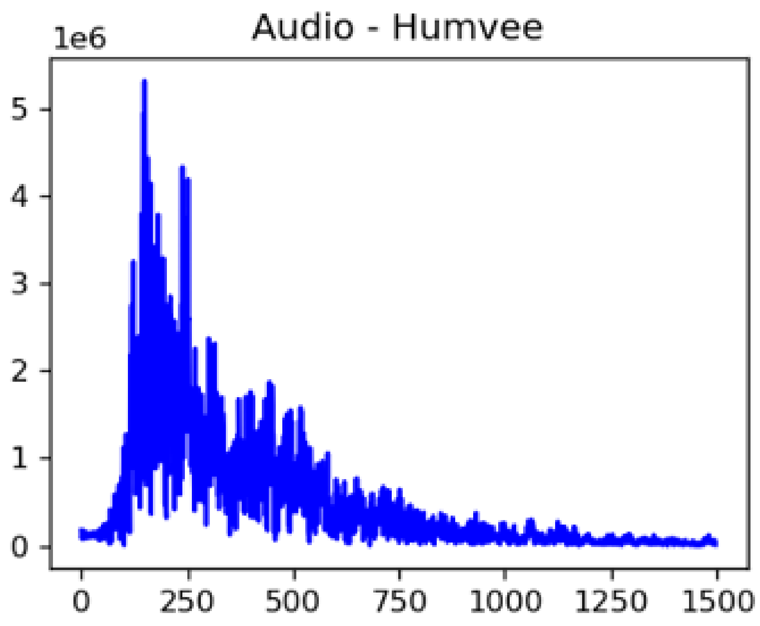
Car B

Intuitions: Which Car Has a Better Suspension System?

Alternatively, if the Fourier Transforms shown to the right are of two cars moving roughly the same speed, which car is likely bigger, Car A or Car B?



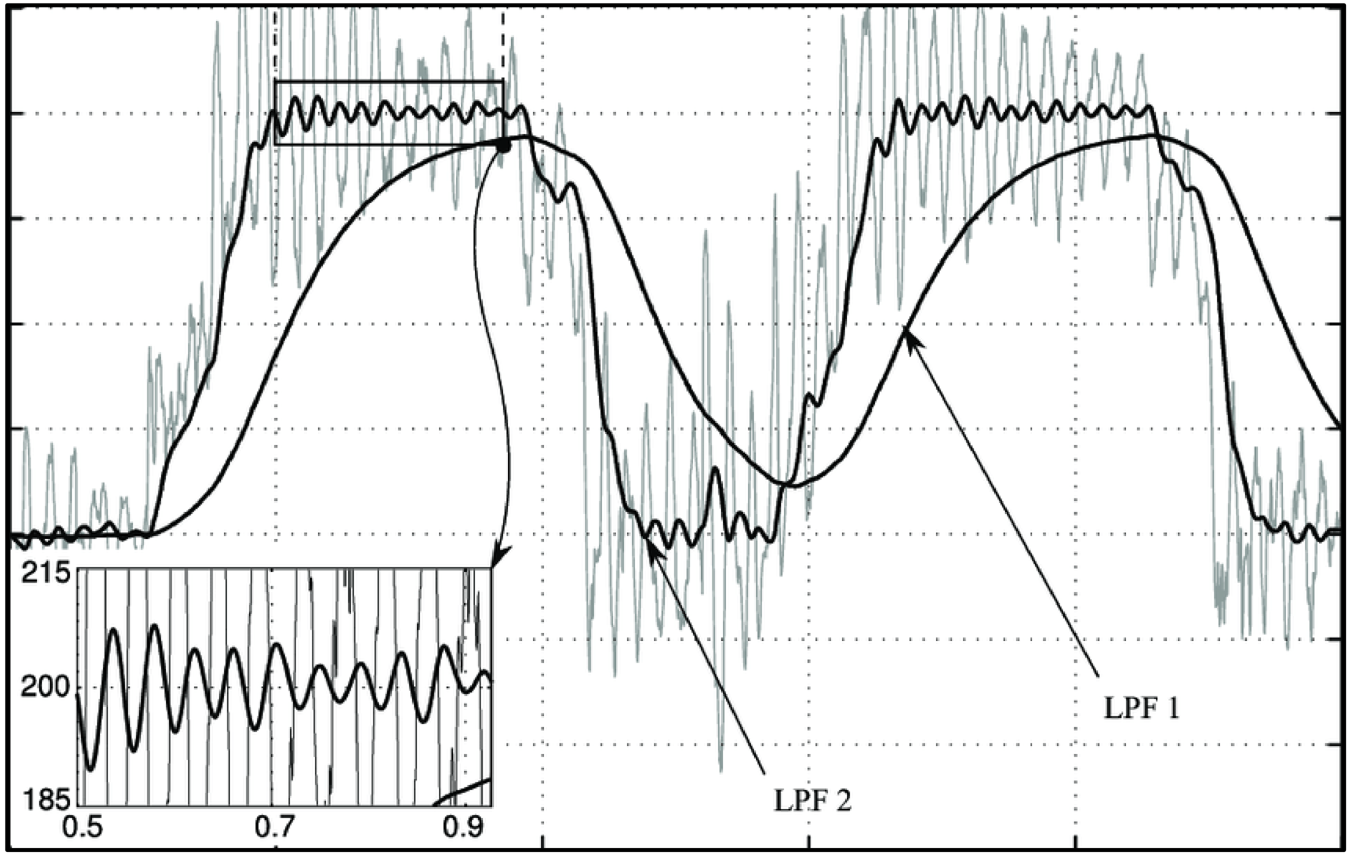
Car A



Car B

Low Pass Filters

Low pass filters let lower frequency components through, but attenuate (reduce) high-frequency components. Most physical systems have inertia (which means they resist higher frequency changes more), so they act as low pass filters



https://www.researchgate.net/figure/The-fundamental-limitation-of-linear-time-invariant-low-pass-filtering-the-length-of-a_fig5_260804992

Note: Most Transmission Media Act as Low Pass Filters for Transmitted Signals

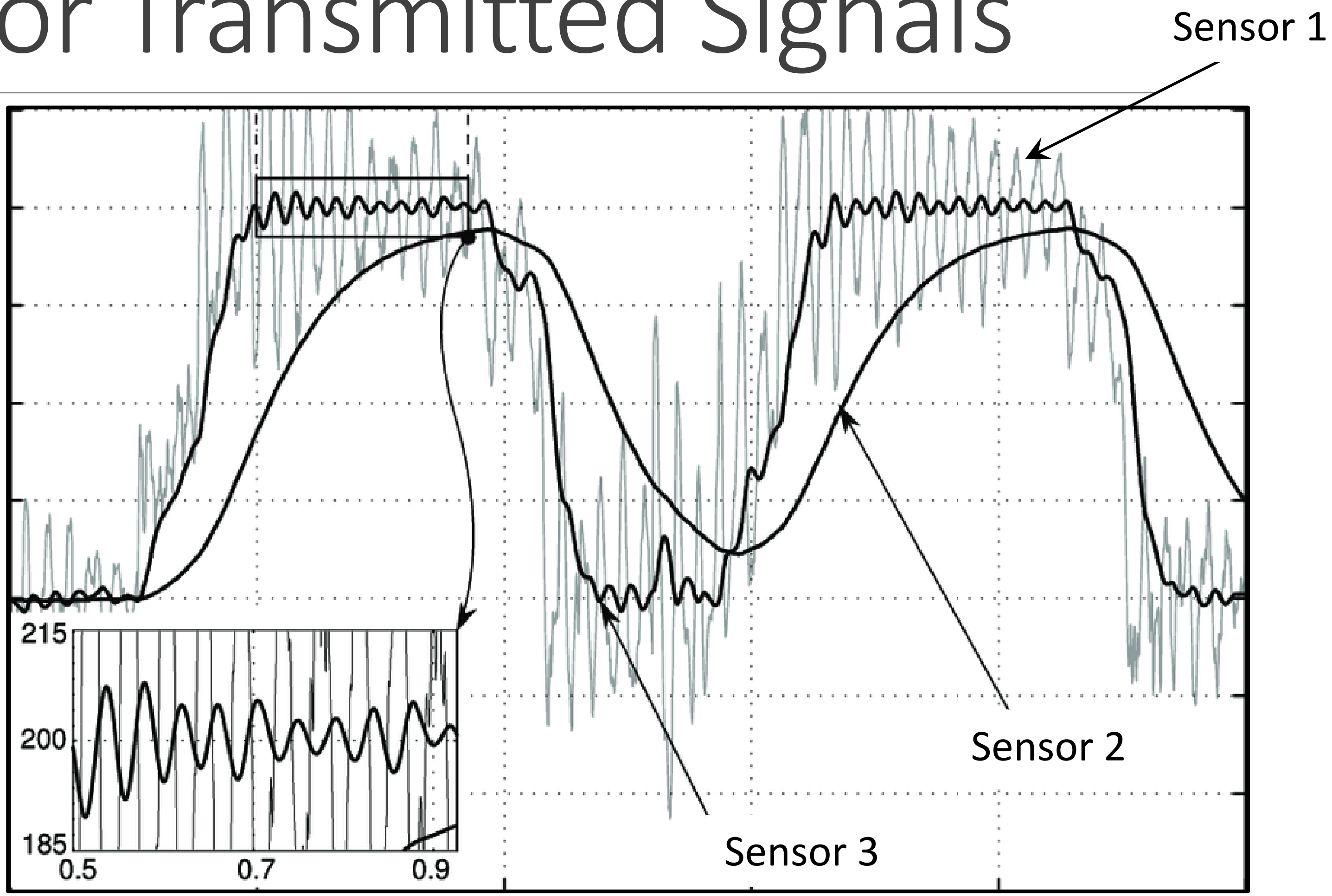
Why?

Note: Most Transmission Media Act as Low Pass Filters for Transmitted Signals

Why? Because of inertia (that resists faster changes causing more attenuation at higher frequencies)

Note: Most Transmission Media Act as Low Pass Filters for Transmitted Signals

Question: If the three sensor waveforms shown in picture were seismic footprints of the same car measured at a distance of 4 meters, 100 meters, and 900 meters from the vehicle, which sensor is at which distance?



https://www.researchgate.net/figure/The-fundamental-limitation-of-linear-time-invariant-low-pass-filtering-the-length-of-a_fig5_260804992

What Is the Frequency Domain?

Reminder:

- Any periodic function of time (a time-domain signal) can be represented as a summation of sinusoids of different amplitudes and frequencies.
- The frequency domain representation of a time-domain signal plots the amplitudes of its component sinusoids versus their frequencies.

What if the time-domain signal is not periodic?

What Is the Frequency Domain?

Reminder:

- Any periodic function of time (a time-domain signal) can be represented as a summation of sinusoids of different amplitudes and frequencies.
- The frequency domain representation of a time-domain signal plots the amplitudes of its component sinusoids versus their frequencies.

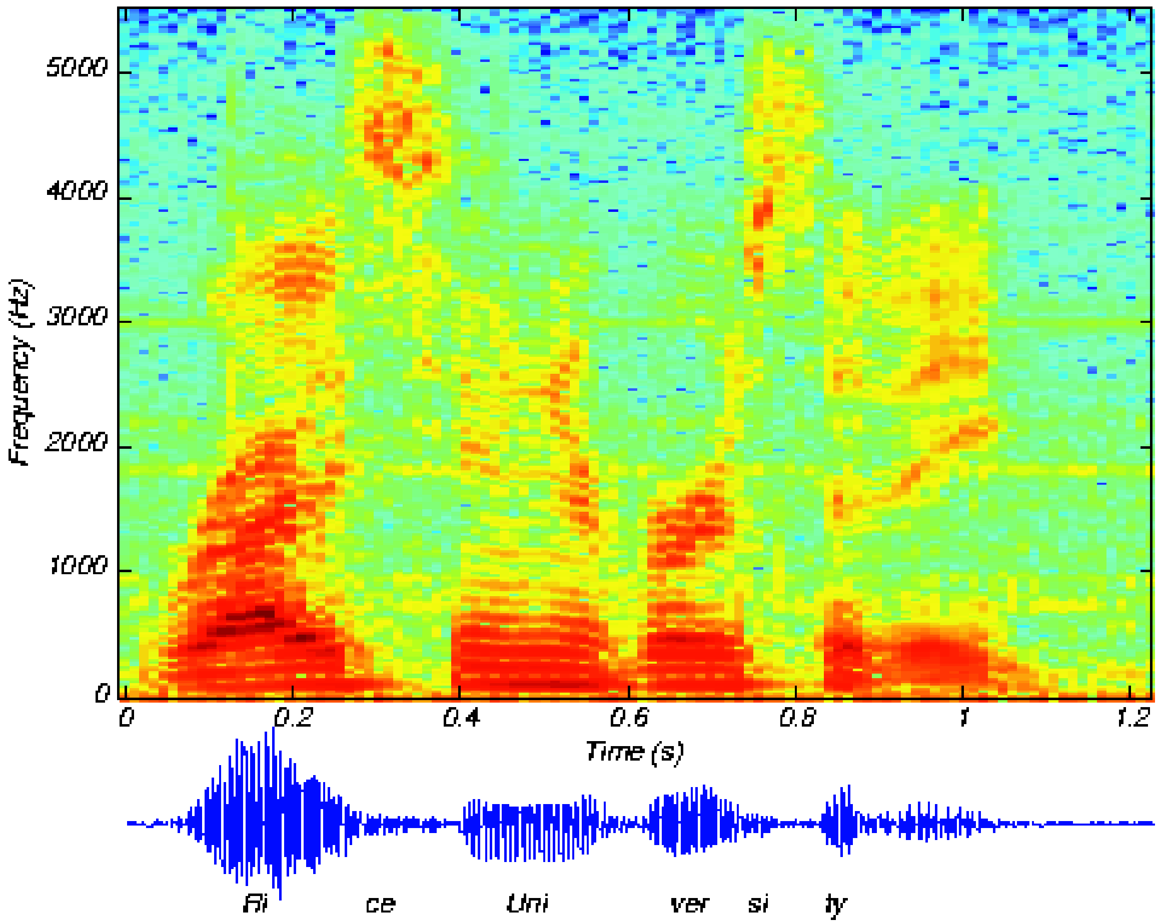
What if the time-domain signal is not periodic?

Cut it into chunks of length, say, T. Transform each chunk to the frequency domain as if it were a periodic signal of period T.

The frequency domain representation will thus differ for every window T.

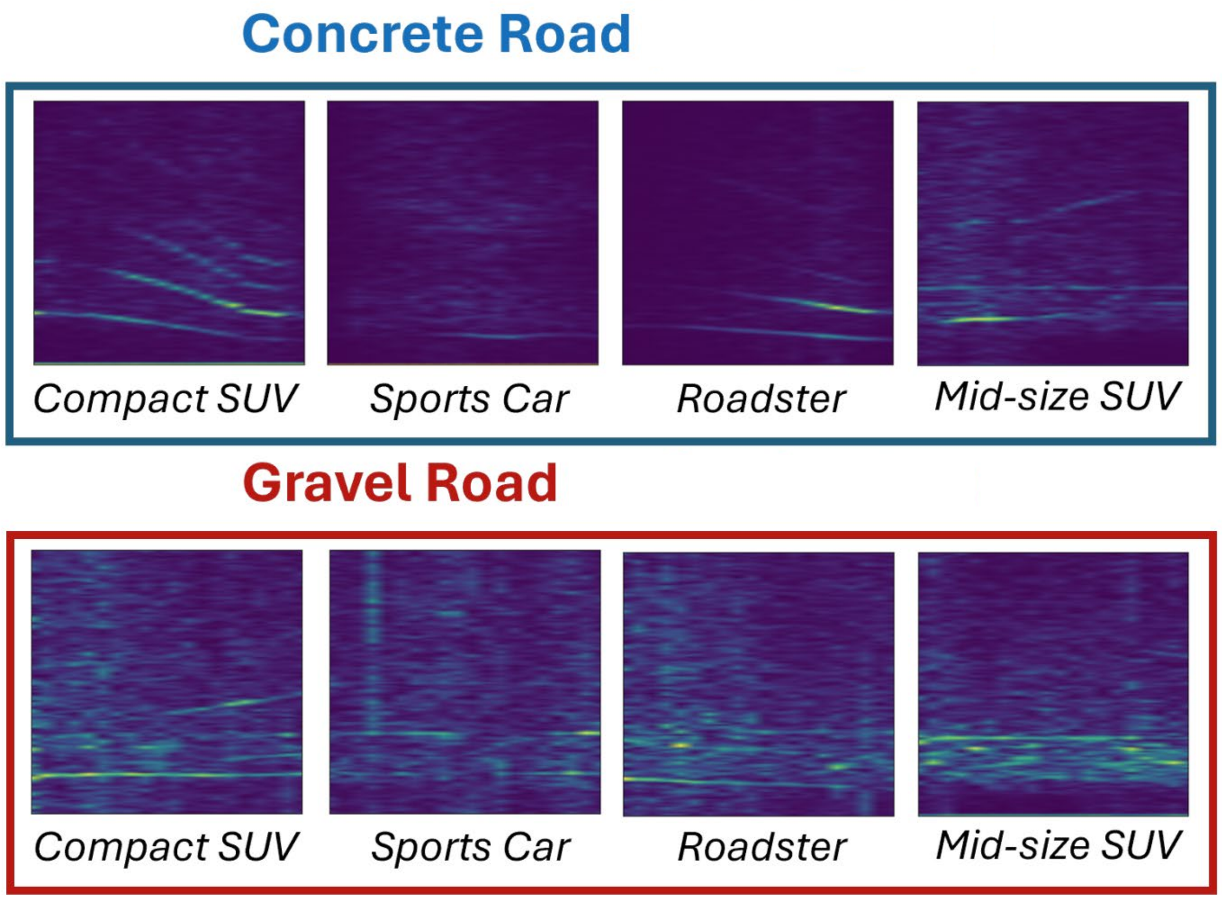
Spectrograms (A Speech Example)

Plot signal energy (in the frequency domain) as a heat map for different frequencies (on the vertical axis) and time slots (on the horizontal axis)



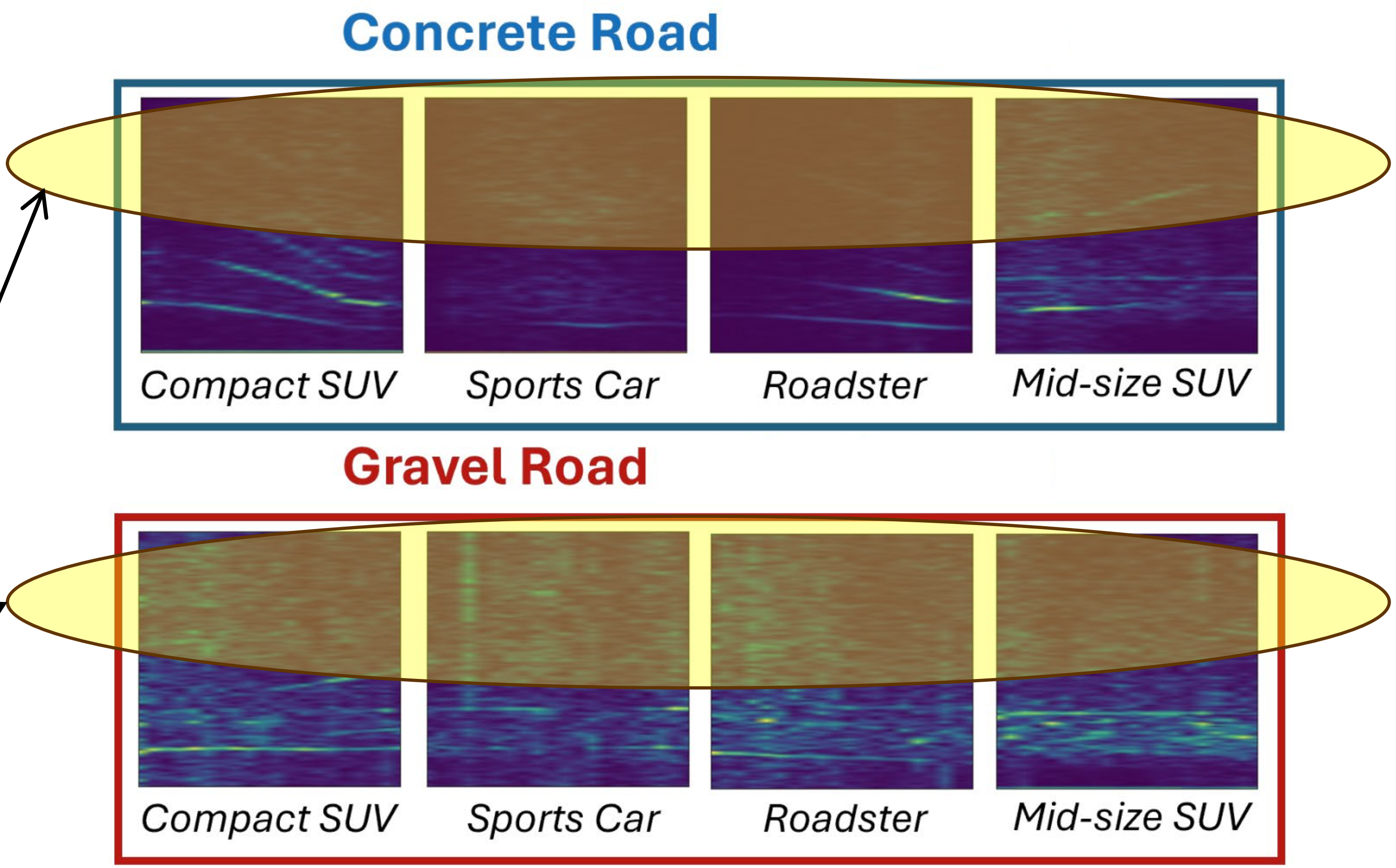
Spectrograms
Example: Seismic Signature of a Vehicle

Plot signal energy (in the frequency domain) as a heat map for different frequencies (on the vertical axis) and time slots (on the horizontal axis)



Spectrograms
Example: Seismic Signature of a Vehicle

Plot signal energy (in the frequency domain) as a heat map for different frequencies (on the vertical axis) and time slots (on the horizontal axis)



Why are there more significant high-frequency components on gravel compared to concrete?

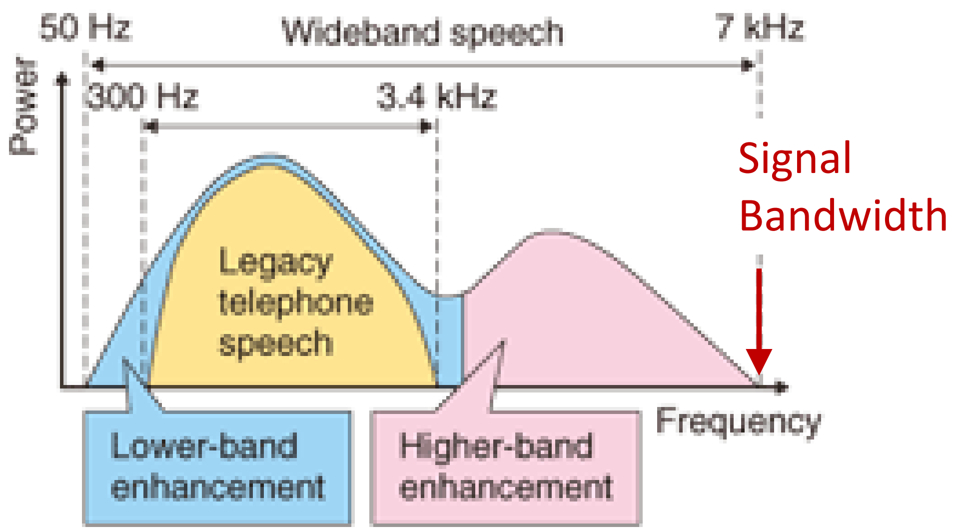
Spectrograms
Example: Seismic Signature of a Vehicle

Representation learning challenge: How to disentangle the effect of signal “sources” (vehicles, persons, tanks, …) from the effect of “propagation media” (types of terrain, multipath reflections, etc)

Signal Bandwidth and Nyquist Frequency

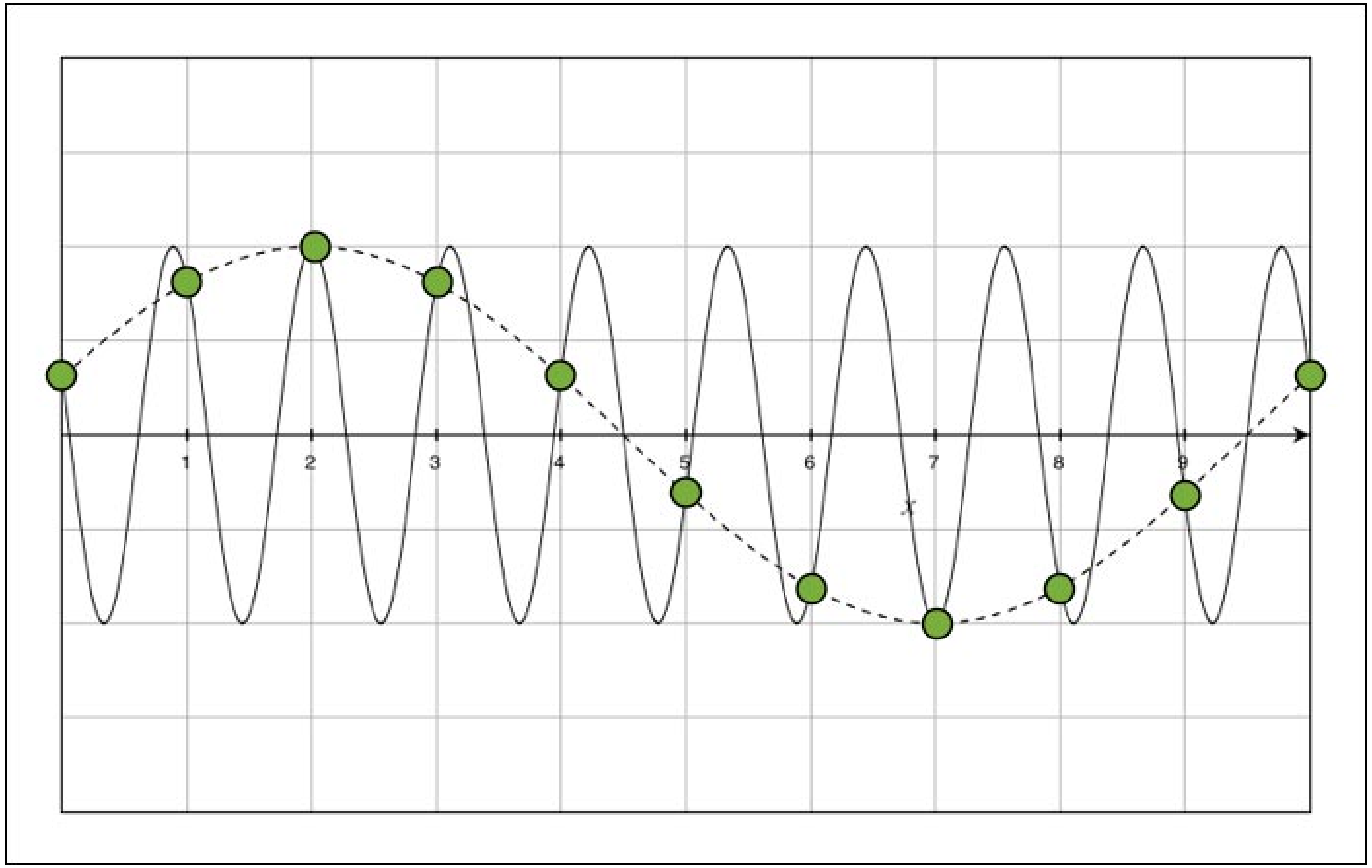
Bandwidth: The signal bandwidth represents the maximum frequency for which the signal carries non-trivial energy in the frequency domain.

Nyquist Frequency: In order to correctly reconstruct the original signal from a sampled version, the sampling frequency should be at least twice the signal bandwidth (this minimum sampling frequency needed for correct reconstruction is called the Nyquist frequency)



Under-sampling and Aliasing

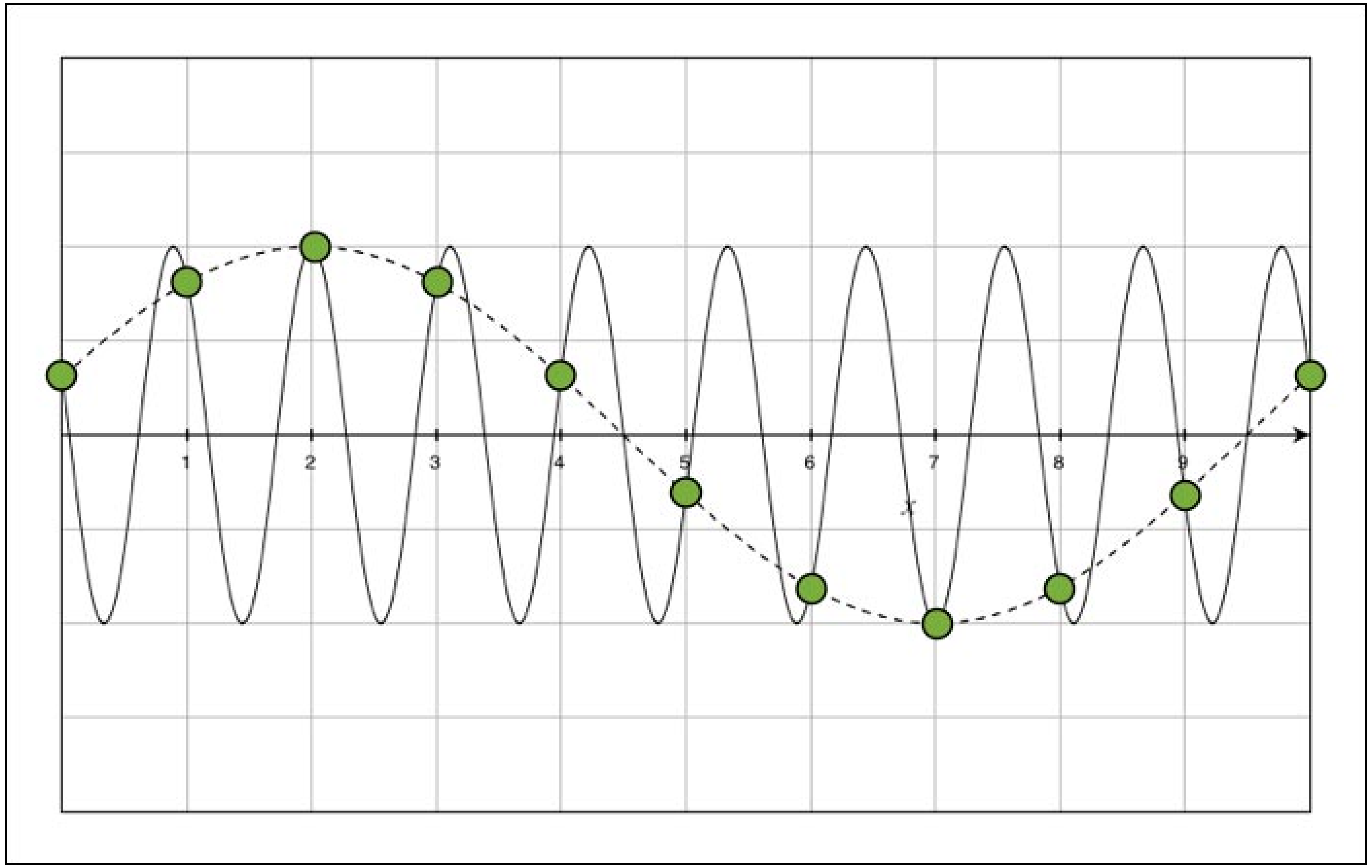
Sampling below the Nyquist frequency causes a shift in the frequency component (the sampled signal looks as if it has a lower frequency). This shift is called Aliasing.



Under-sampling and Aliasing

Sampling below the Nyquist frequency causes a shift in the frequency component (the sampled signal looks as if it has a lower frequency). This shift is called Aliasing.

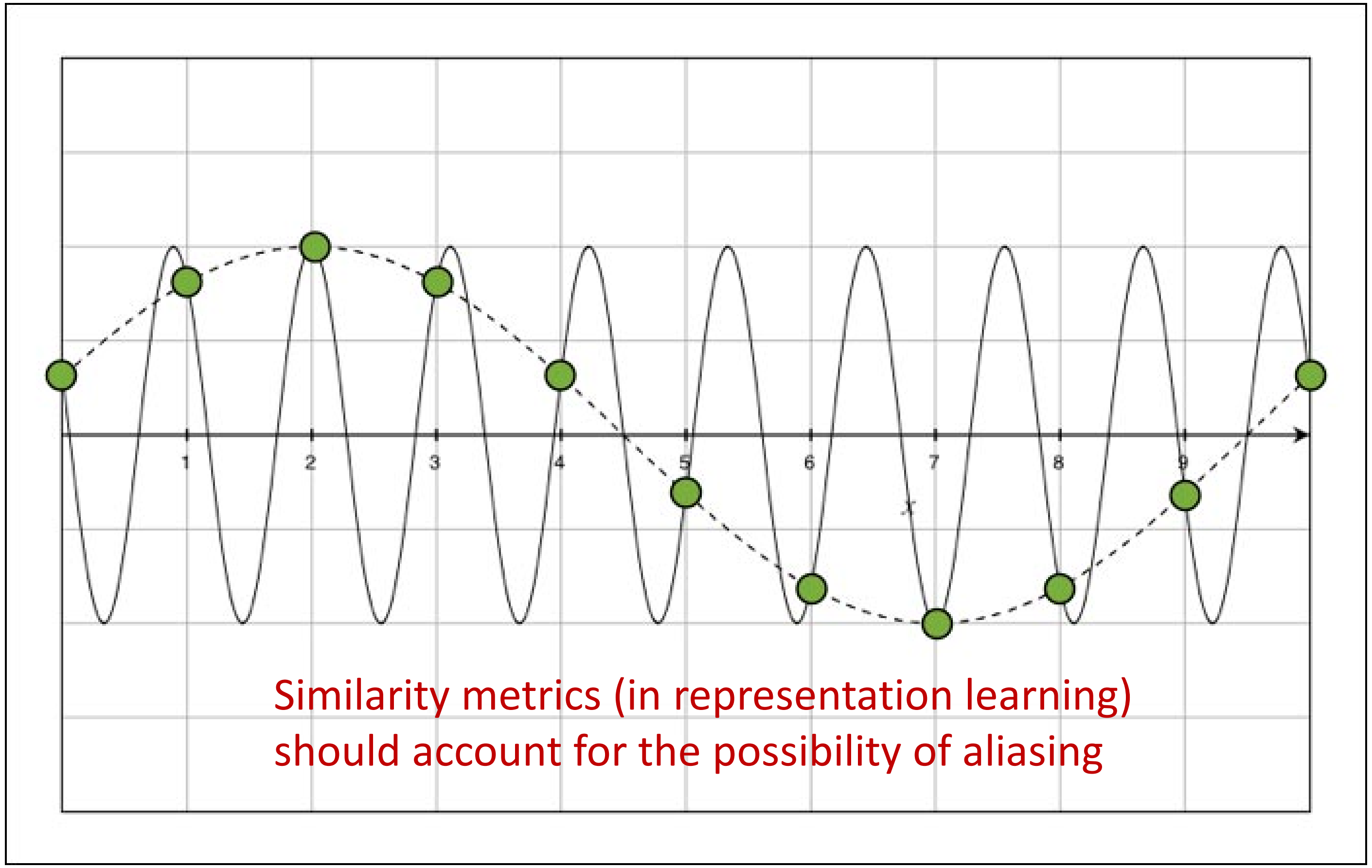
Note: What does aliasing imply in terms of similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?



Under-sampling and Aliasing

Sampling below the Nyquist frequency causes a shift in the frequency component (the sampled signal looks as if it has a lower frequency). This shift is called Aliasing.

Note: What does aliasing imply in terms of similarity notions between signals (and signal augmentations) when using a representation learning approach like contrastive learning?



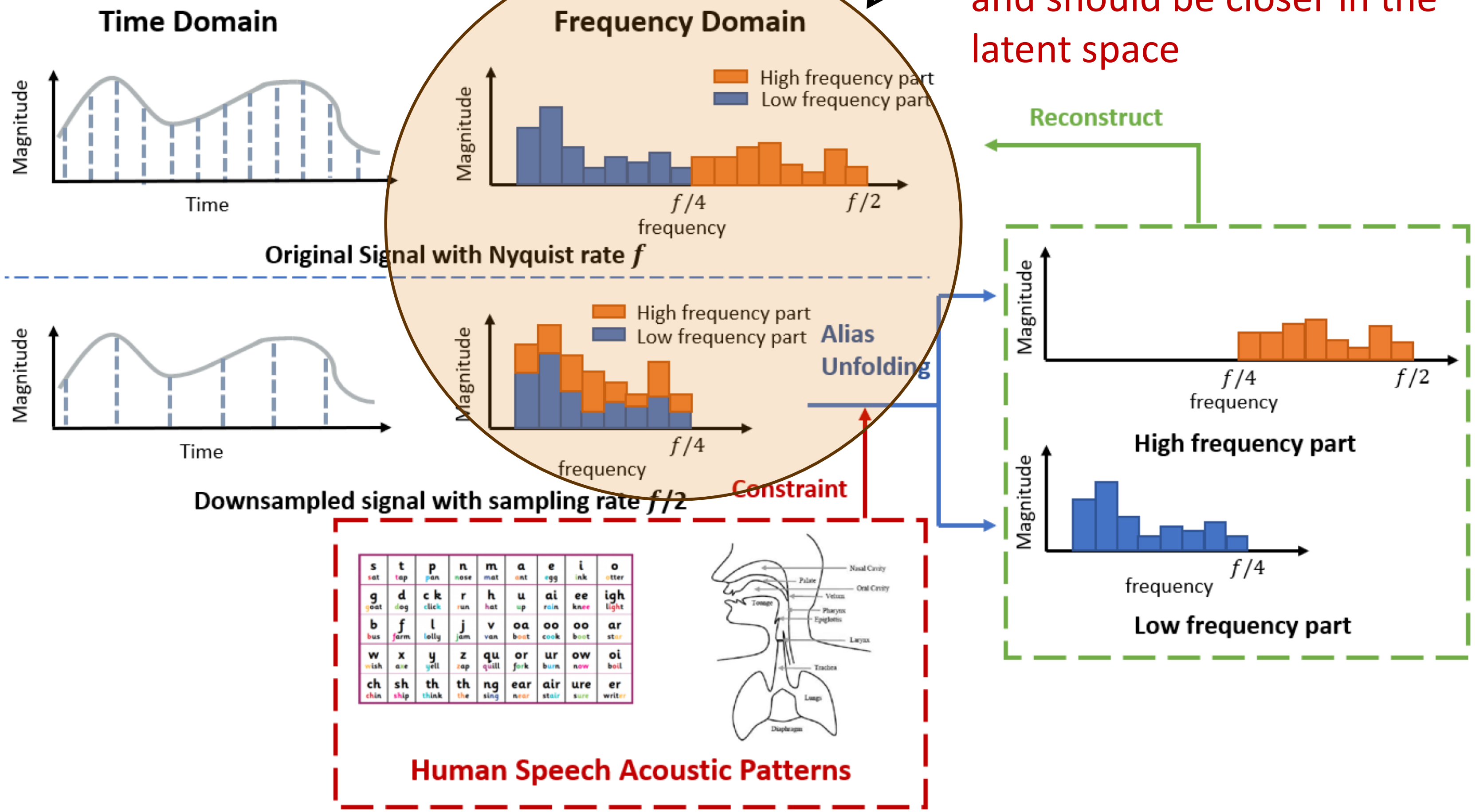
Similarity metrics (in representation learning) should account for the possibility of aliasing

Aliasing and Similarity

Thought Experiment:
Can a properly trained neural network undo aliasing in order to reconstruct signals sampled at significantly below their Nyquist rate?

Example:
Sound reconstruction from accelerometer data

Representation learning insight: In principle, signals before and after aliasing are "semantically similar" and should be closer in the latent space



Spectrogram Feature Representation Learning: Lessons Learned

Lesson #1: For purposes of representation learning (e.g., contrastive learning, auto-encoding, etc) can one treat spectrograms as images? If not, why not?

Spectrogram Feature Representation Learning: Lessons Learned

Lesson #2: How can we inject application bias into representation learning of frequency domain data to best leverage its unique properties?