hydro stasis

Separating the signal from the noise:

Signal processing and feature extraction techniques for biological data

Michelle Hoogenhout, PhD
Mathias Ciliberto, PhD

Learning objectives

At the end of the workshop attendees should be able to:

- List 2 use cases
- Able to create spectrogram to visualize data
- Differentiate between frequency and time-domain analysis
- Know when to use frequency and when to use time-domain filters
- Apply at least one frequency and one time-domain filter to the data
- Evaluate whether the filter was successful - e.g. calculate error
- Discuss 2-3 factors to consider in data collection of PPG signals

hydro stasis

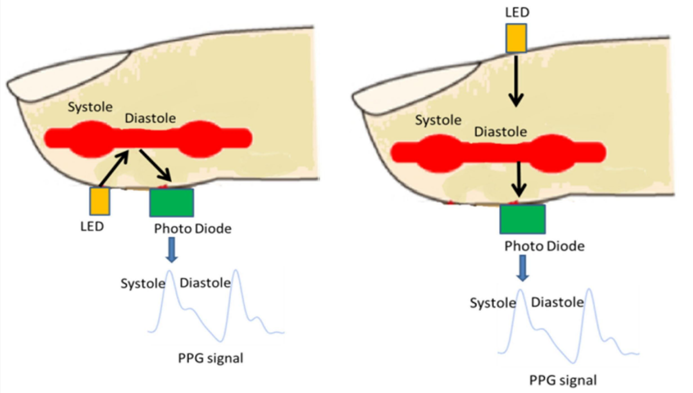
Requirements

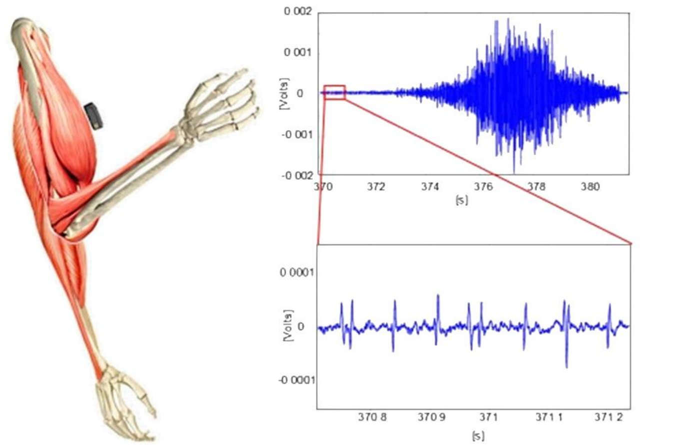
- Will work from Colab Notebook
  - No need to install anything locally
- Assume familiarity with:
  - Python
  - Numpy, Pandas, Scipy packages

hydro stasis

Biological signals

- Photoplethysmogram (PPG)
- Electromyography (EMG)
- Inertial data
  - accelerometer
  - gyroscope
  - magnetometer

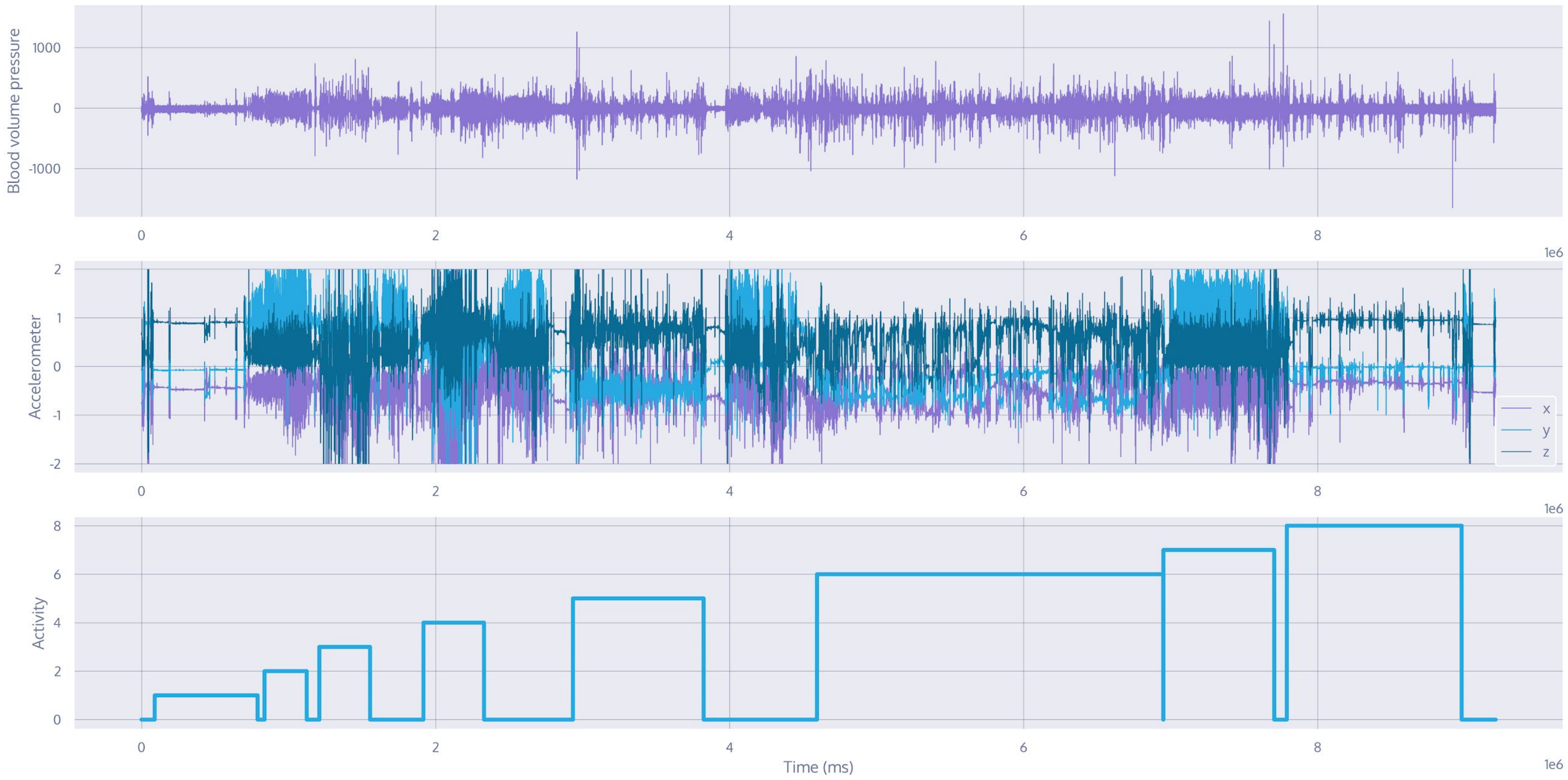




hydro stasis

Visualizing your data

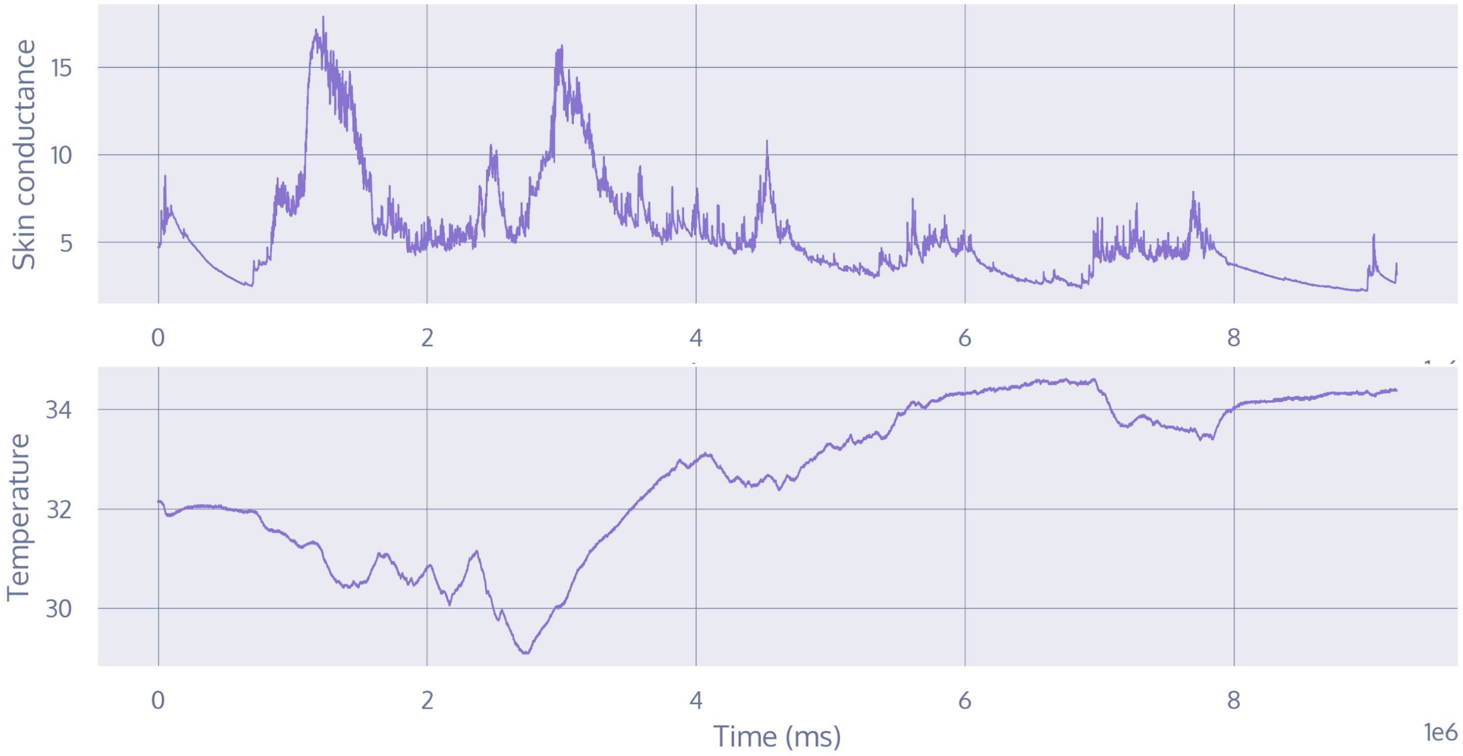
- Inertial
  - Useful for detecting anomalies
- PPG
  - Check for missing data
- Activities (labels)



Time-domain signal example

Interested in amplitude

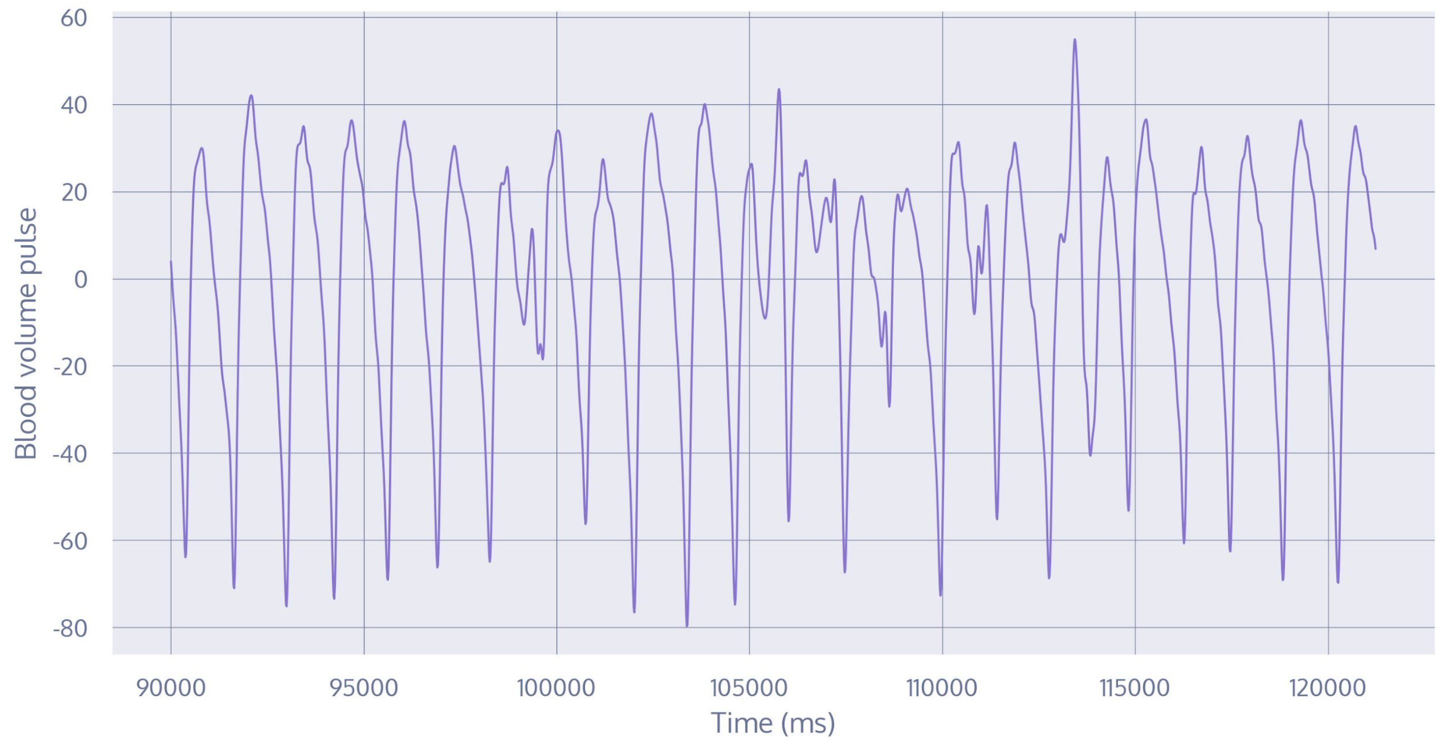
- Skin conductance
- Temperature



Frequency-domain signal example

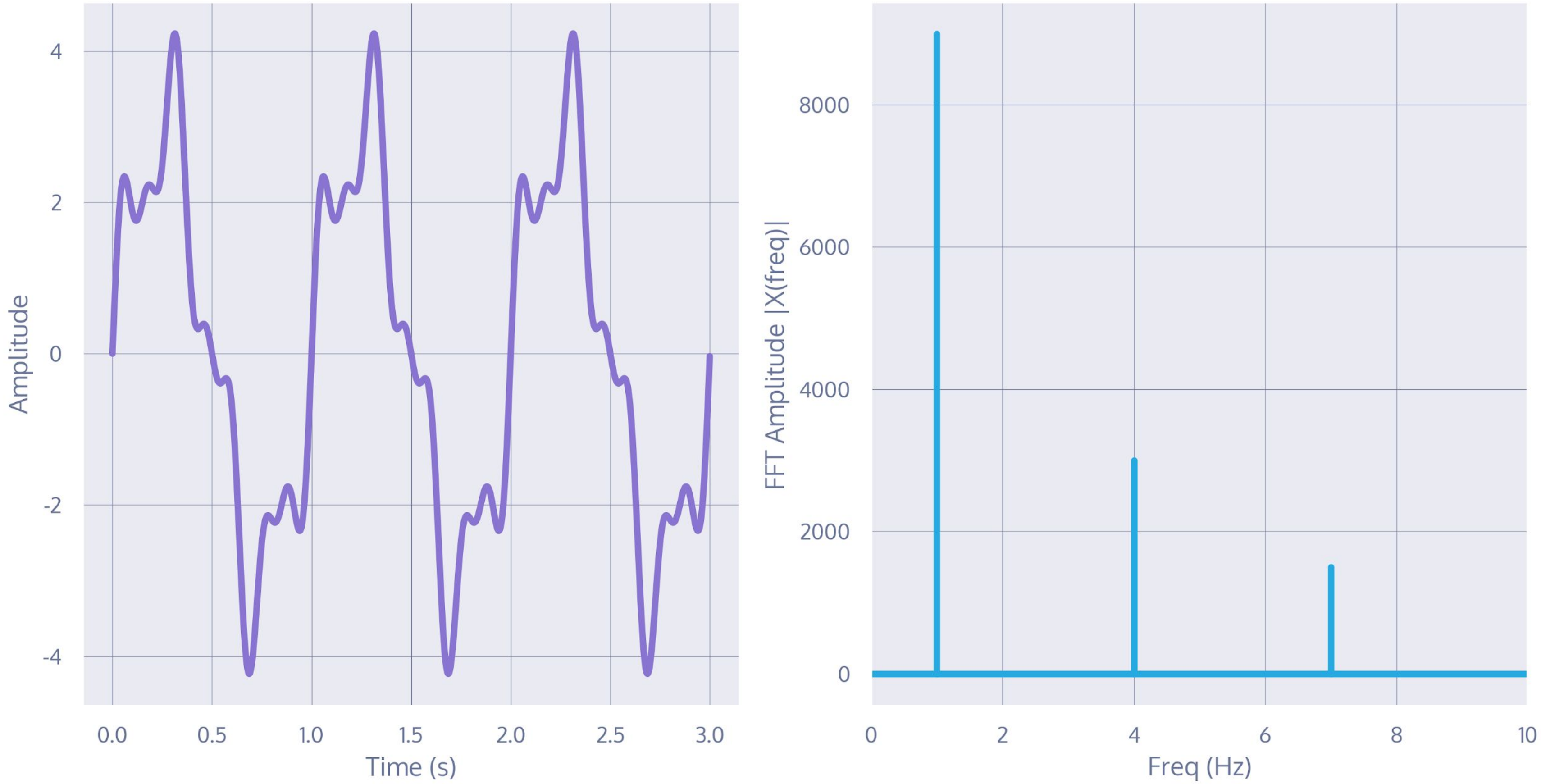
Interested in peaks/repetition:

- Blood volume pressure



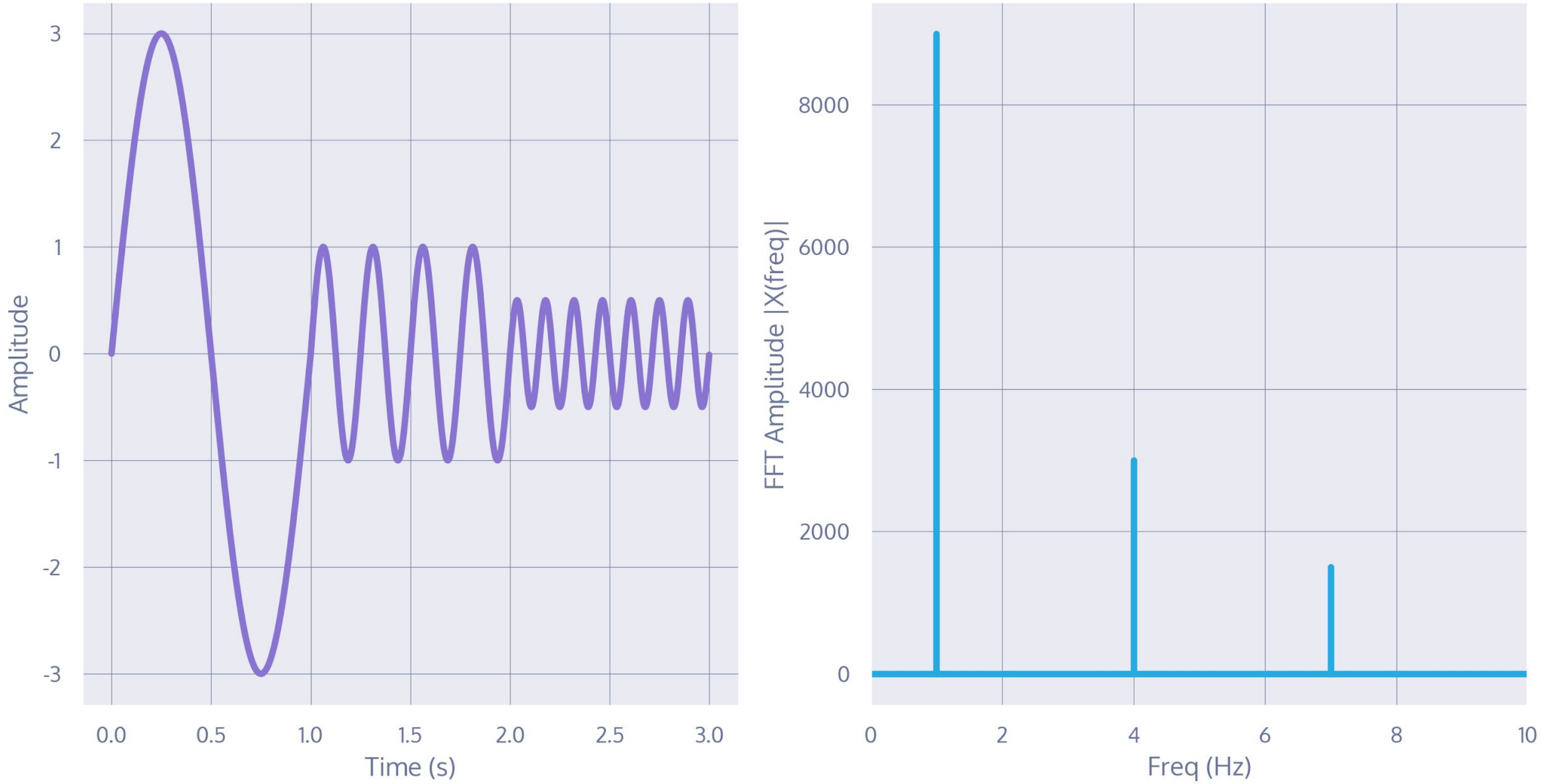
hydro stasis

Time vs Frequency domain

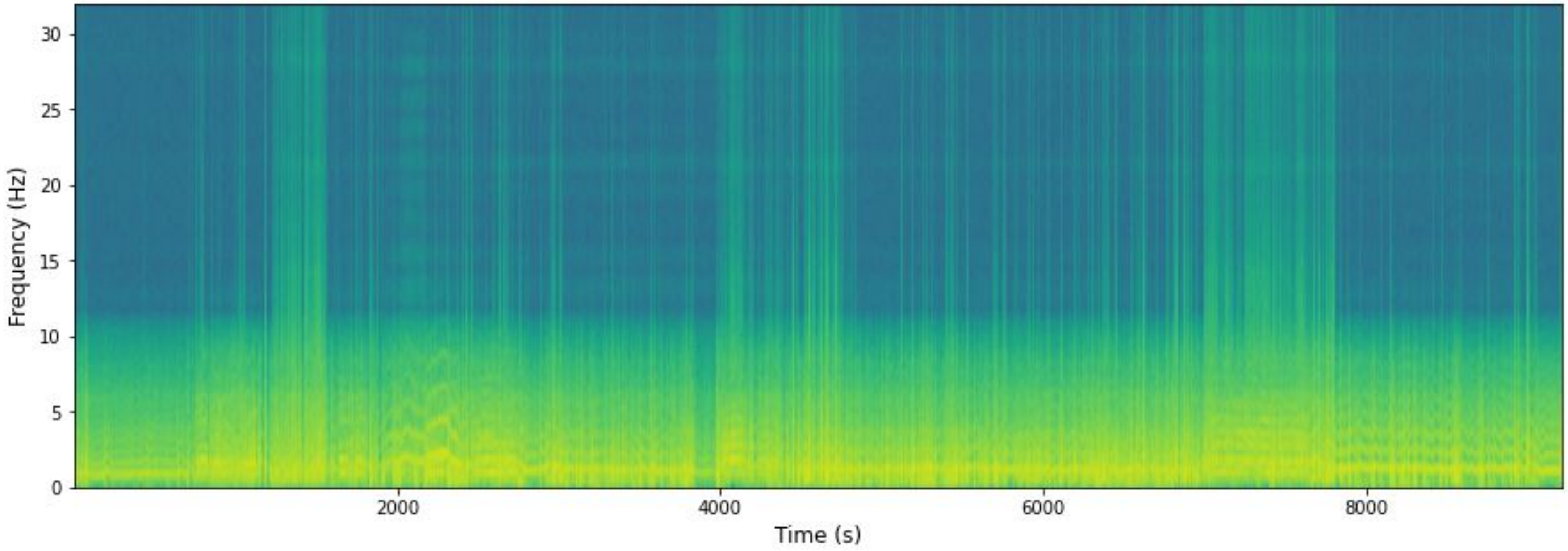


hydro
stasis

Time vs Frequency domain



Spectrograms: Visualize time and frequency at the same time



hydro stasis

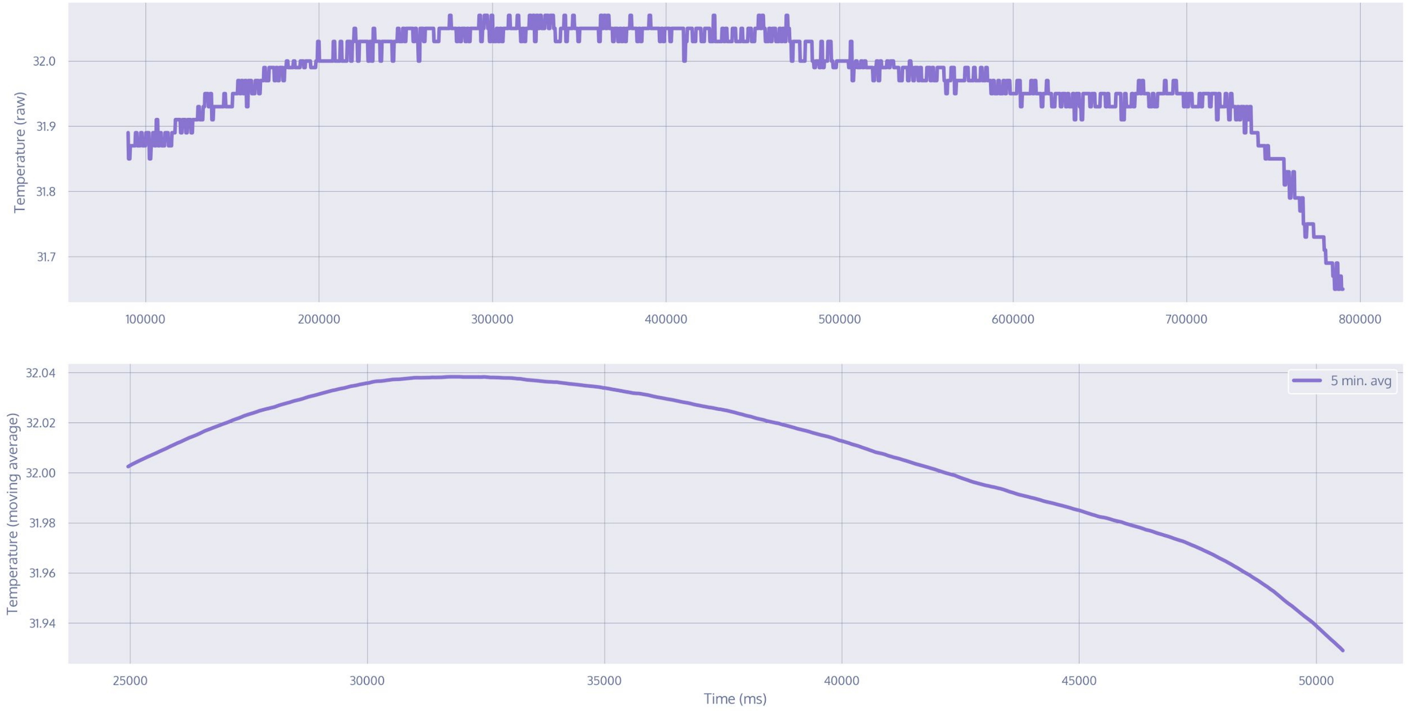
Time-domain filters

Moving average:

$$MA_k = \frac{1}{k} \sum_{i=n-k+1}^{n} p_i$$

k = window length, p= sample

- Good usage:
  - Identify slow-changing process / **trends**
  - Remove small fluctuations

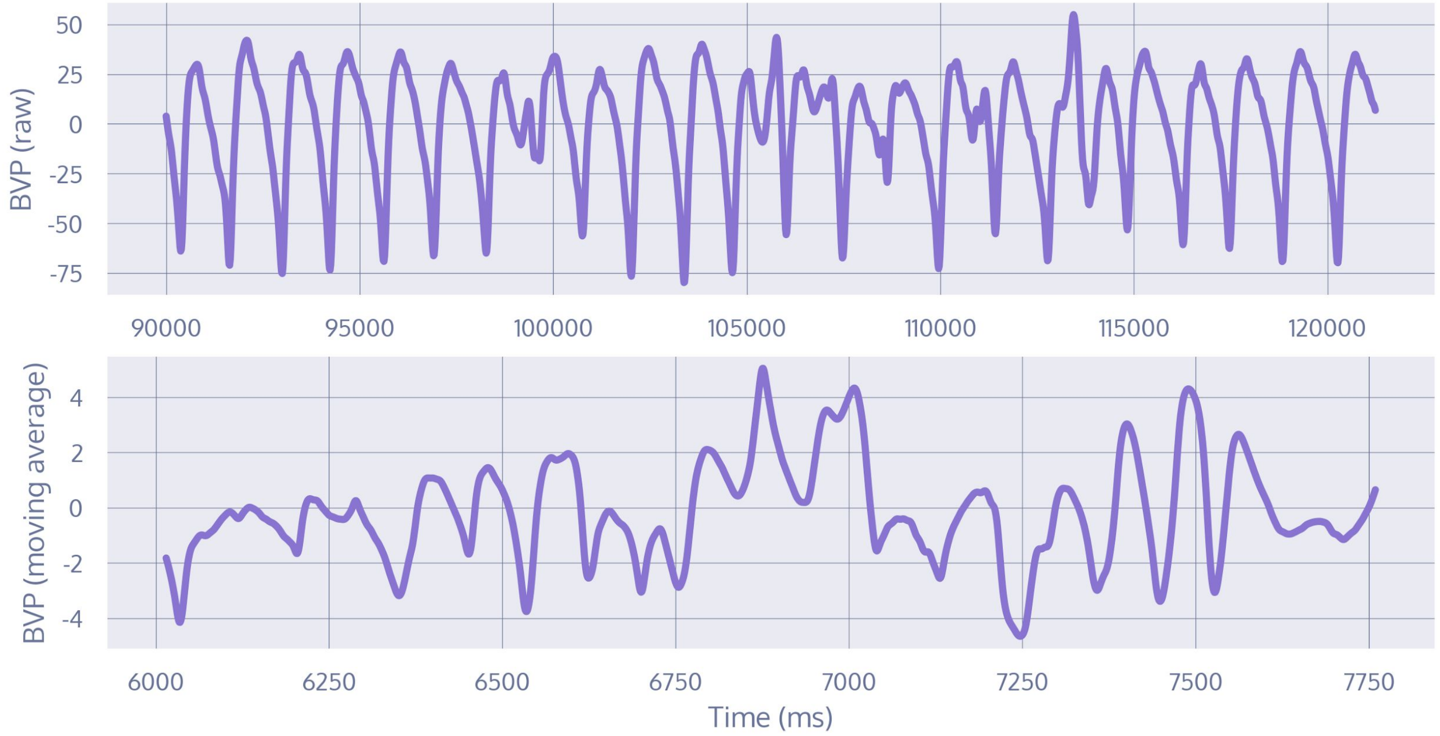


hydro stasis

Time-domain filters

Moving average

- Bad usage
  - Frequency-based signal

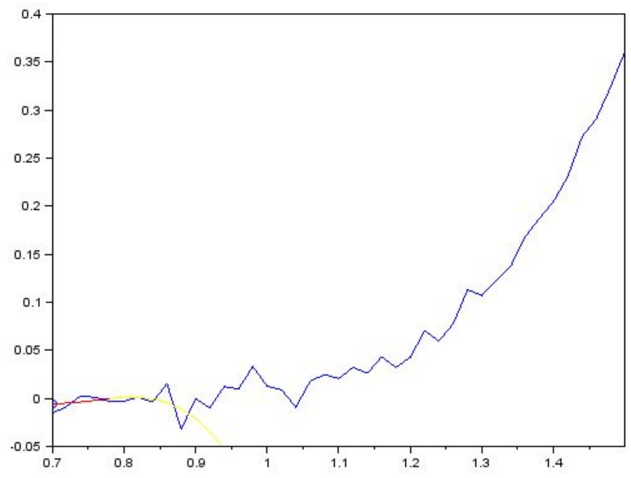


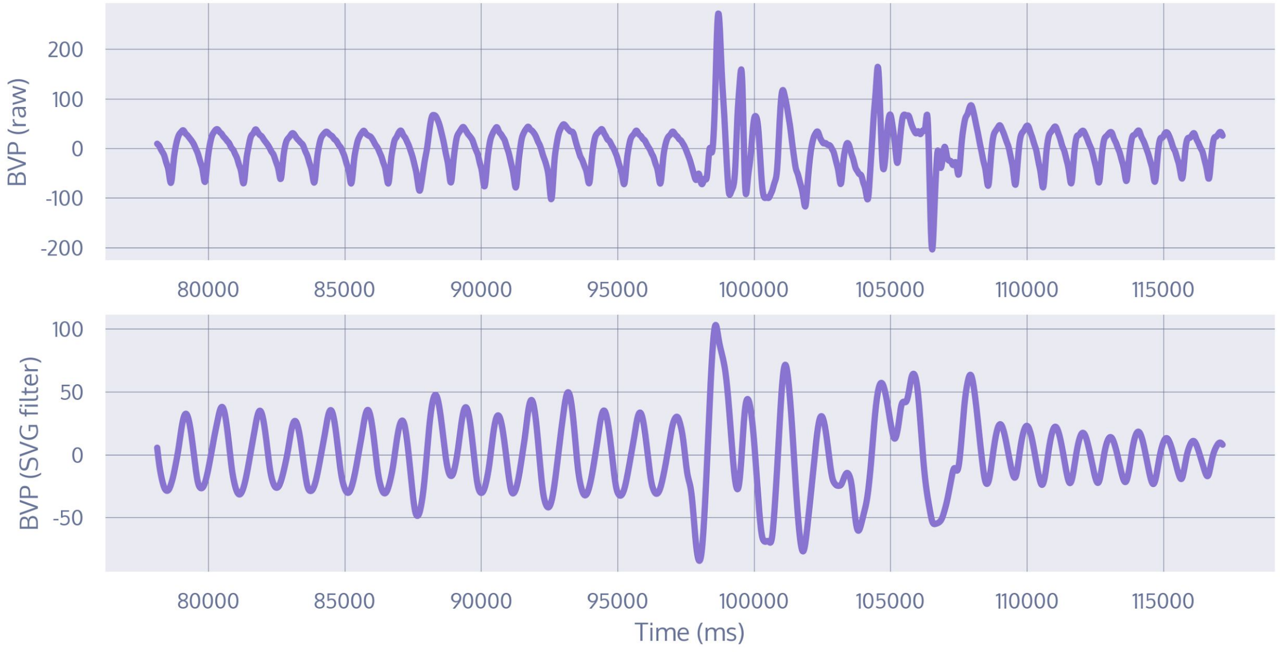
hydro stasis

Time-domain filters

Savitzky-Golay

- Smoothing filter based on convolution
- Good for movements artefact removal (spikes)





hydro stasis

Capturing information in the frequency domain

Fourier transform

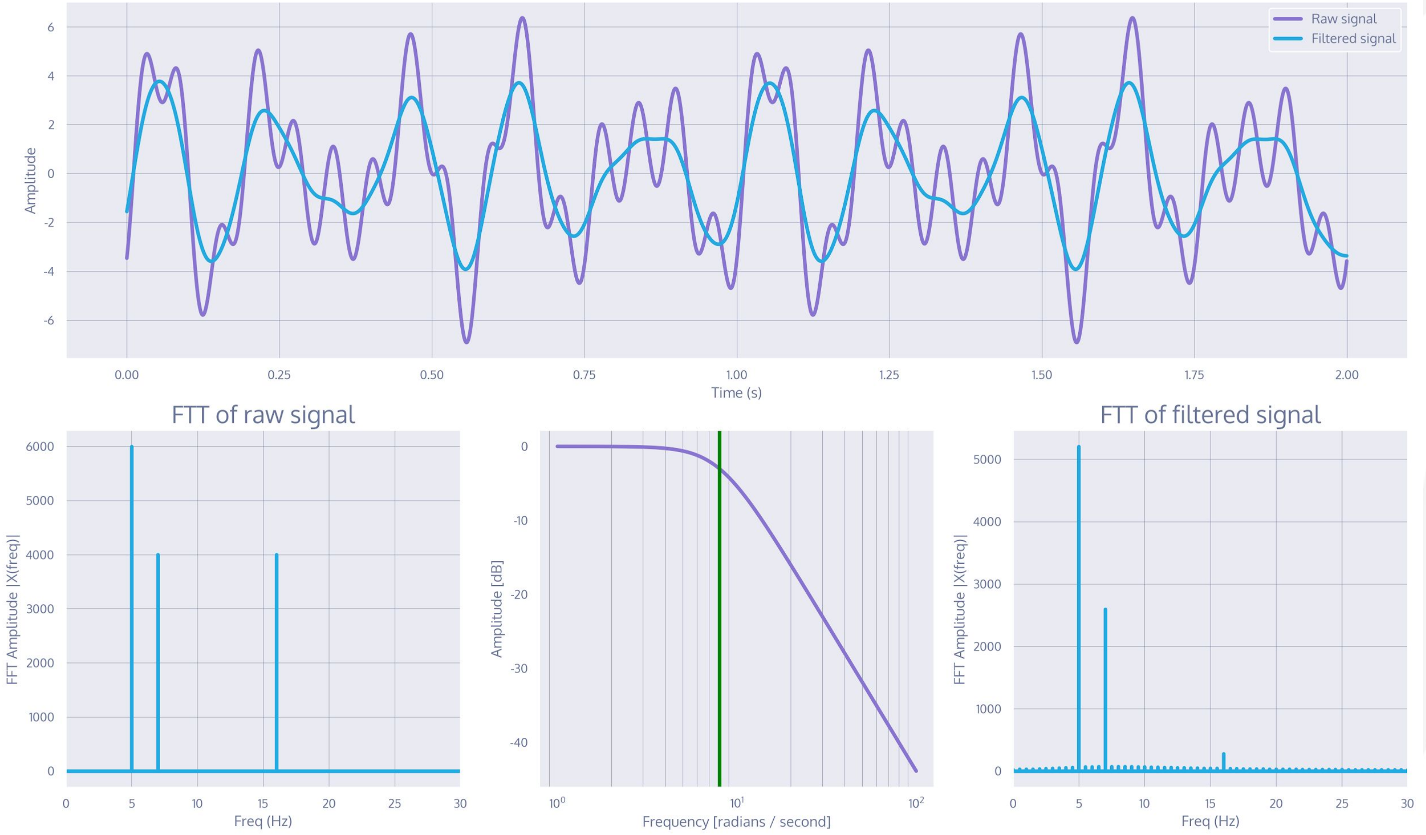
- Useful: What frequencies exist?
- HOWEVER FFT assumes **stationarity**
- Not useful: When do frequencies change?

hydro stasis

Frequency-domain filters

Low-pass filter:

- Remove high frequency noise (i.e. vibrations in inertials sensors data)

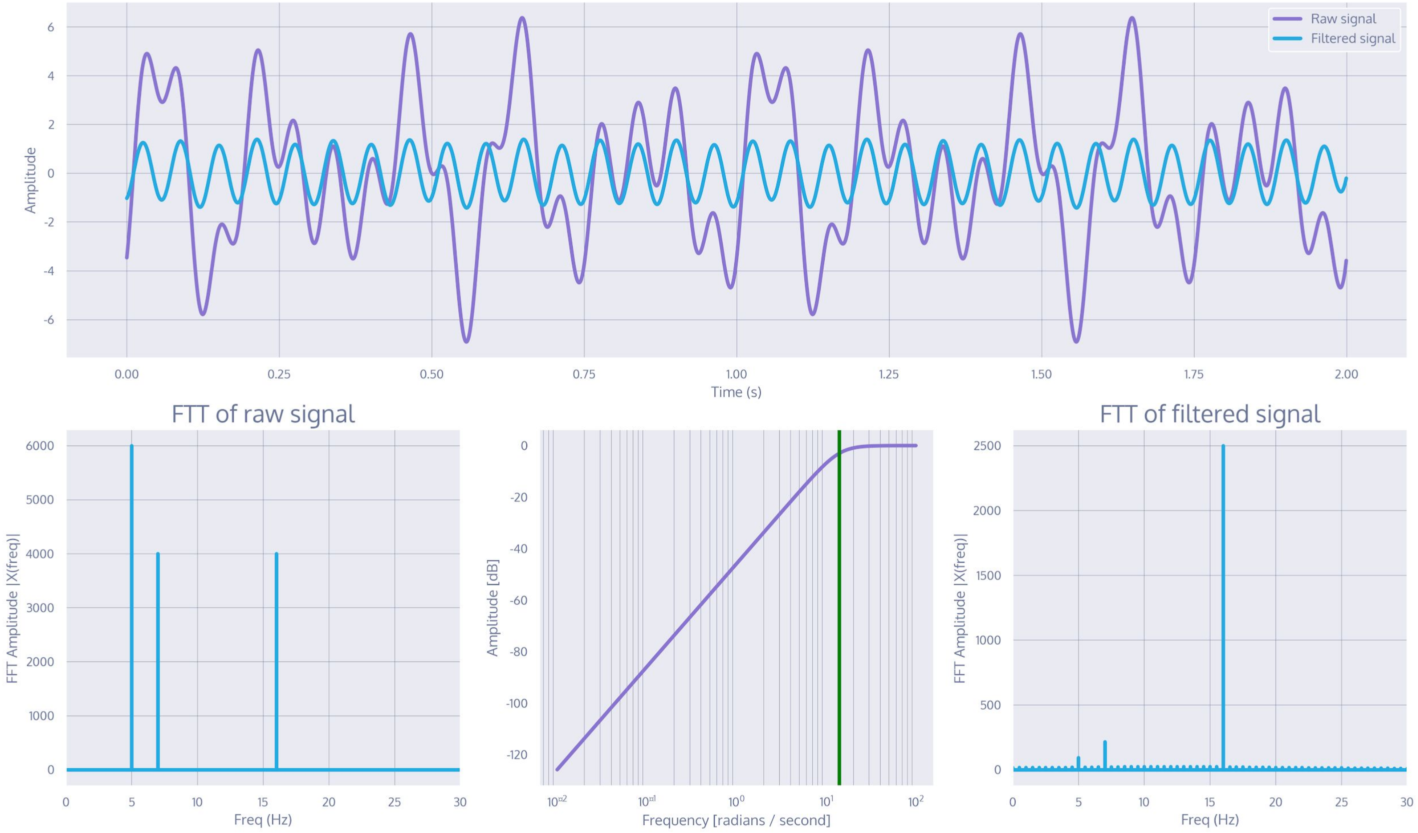


hydro stasis

Frequency-domain filters

High-pass filter:

- Remove low frequency variations (i.e. changes due to temperature variations in optical signals)



Joint time-frequency

Short time Fourier transform

Trade-off between time and frequency resolution

Wavelet functions

- Improves on time/resolution trade-off by having a ‘wavelet’ rather than fixed time window
- Expanded wavelet: Low frequency components, bad time resolution
- Shrunken waveliet: High frequency components, good time resolution

hydro stasis

Check-in

What’s the difference between time and frequency (spectral) analysis?

Can you think of a use case for each that we have not discussed?

Case Study: Detecting heart rate

Goal: Detect a user's heart rate using photoplethysmography (PPG) signals

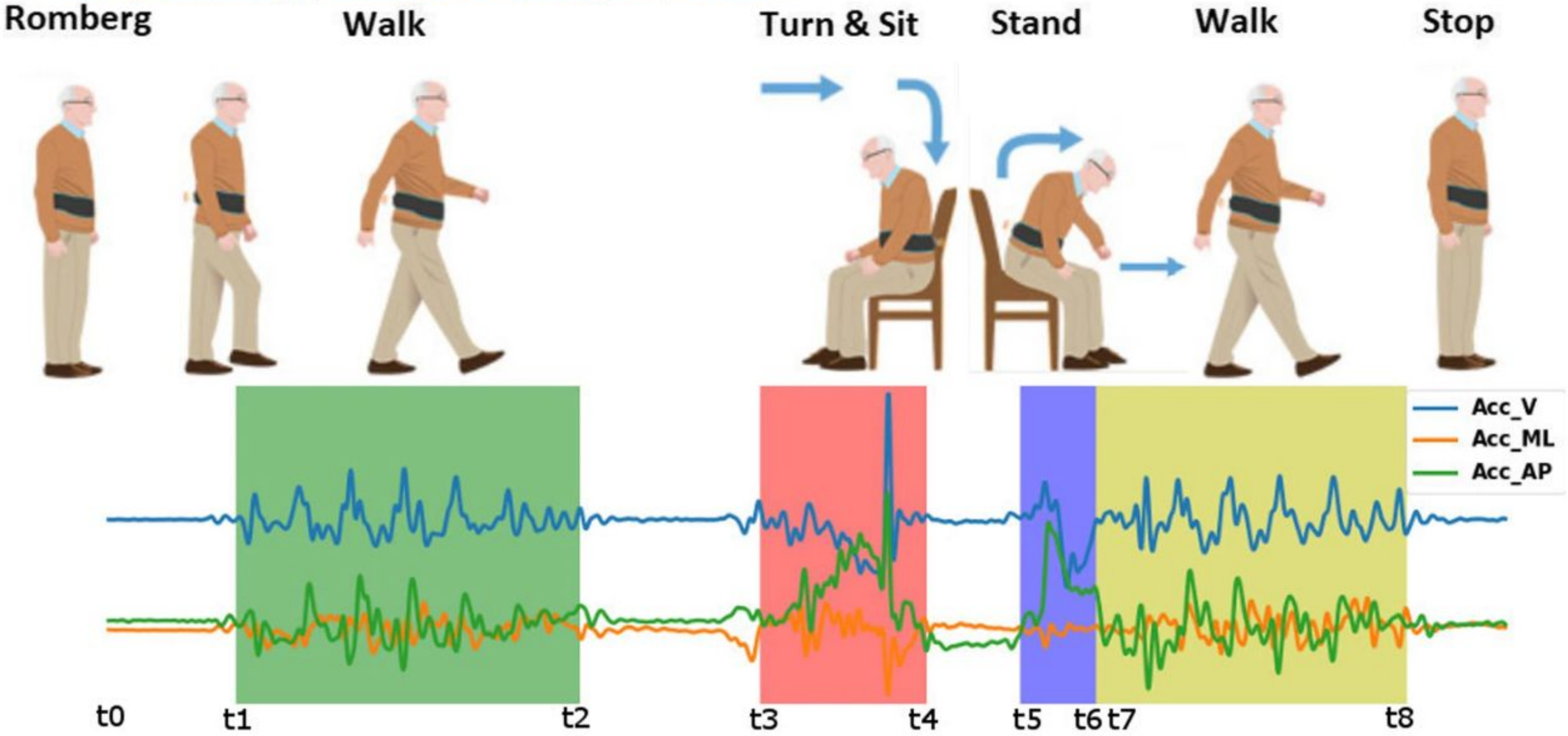
Compare accuracy of heart rate estimates between raw and processed signals under different conditions

[Interactive Colab Notebook session]

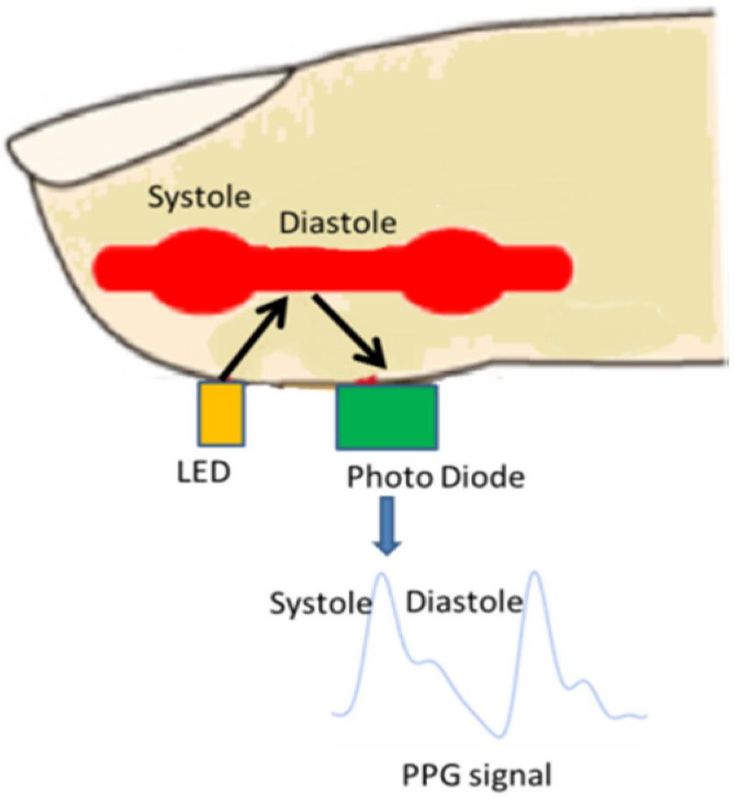
hydro stasis

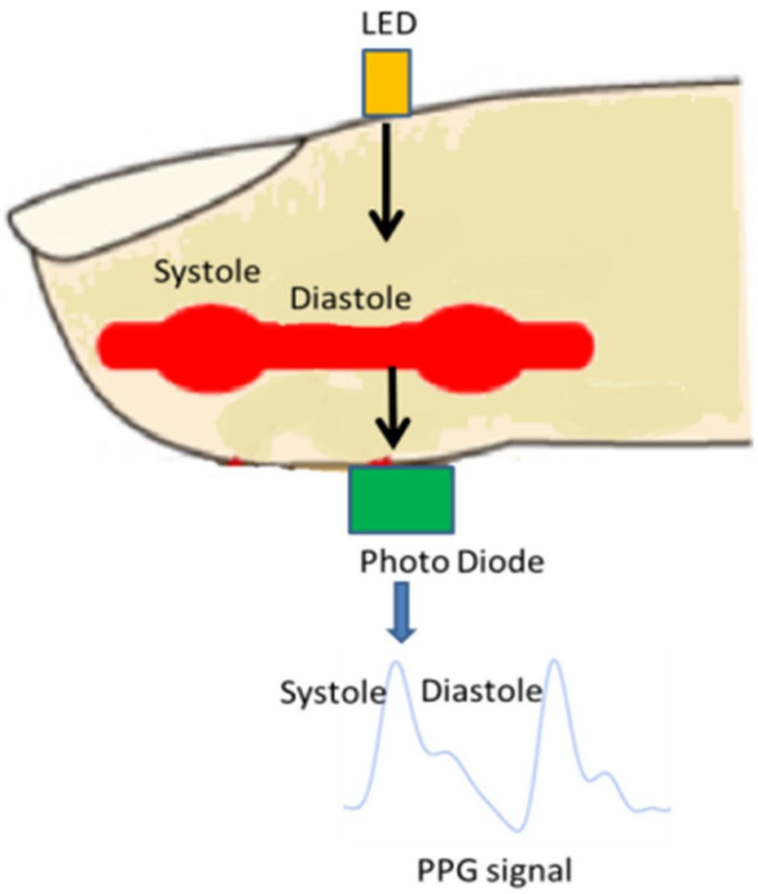
Example accelerometry signal:

From: Mobility assessment in people with Alzheimer disease using smartphone sensors



Example PPG signal:





Factors to consider in data collection and analysis

Easier to solve problems earlier rather than later!

- Know your confounding factors (and controlling for them!)
- Select the right frequency
- Find a way to ensure that your variables of interest are measured in they way you expect
- Finding a good ground truth can be challenging

hydro stasis

Factors to consider in data collection and analysis

What obstacles have you faced in your own analyses?

hydro stasis

Recap: Learning objectives

- List 2 use cases for today's signal processing examples
- Able to create spectrogram to visualize data
- Differentiate between frequency and time-domain analysis
- Know when to use frequency and when to use time-domain filters
- Apply at least one frequency and one time-domain filter to the data
- Evaluate whether the filter was successful - e.g. calculate error
- Discuss 2-3 factors to consider in data collection of PPG signals

hydro stasis

Putting it all together

You can now use spectrograms and filtered signals as input in a variety of statistical learning algorithms

Examples uses:

- spectrograms for detecting activities with CNN
- regression analysis on denoised PPG data to detect blood volume
- wavelet power spectra of skin temperature for detecting covid with SVD
- wavelet denoising of PPG signals to detect heart rate anomalies

hydro stasis

Additional Resources

- Video: Advanced physiological signal processing
- Course: Complete neural signal processing and analysis: Zero to hero

- PPG example analysis: Blood pressure estimation using PPG
- ECG example: Signal processing with MATLAB

- Physics fundamentals: What is the Fourier transform: A visual introduction

hydro stasis