CERN

Interpretable Anomaly Detection in the LHC Main Dipole Circuit with Non-negative Matrix Factorization

Christoph Obermair, TE-MPE-CB

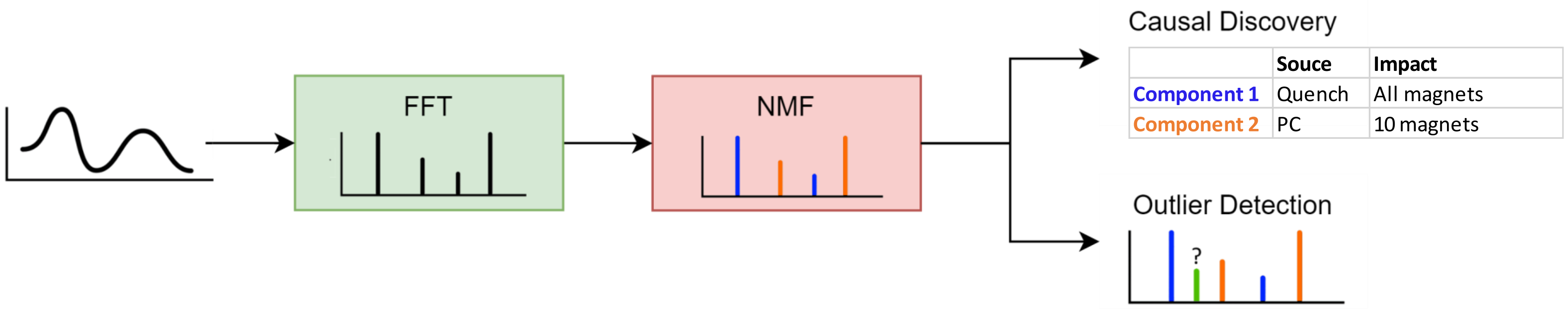
Acknowledgement:
Marvin Janitschke, TE-MPE-PE
Emmanuele Ravaioli, TE-MPE-PE
Mariusz Wozniak, TE-MPE-PE
Zinour Charifoulline, TE-MPE-PE
Lukas Felsberger, TE-MPE-MI
Daniel Wollmann, TE-MPE-CB
Aleksandra Mnich, TE-MPE-CB
Franz Pernkopf, TU-Graz

Outline

Goal: Define and understand **normal** behavior, detect **abnormal** behavior of the RB circuit

Approach:

1. Extract **frequencies** in data → **Fast Fourier transform (FFT)**
2. Group expected frequencies that occur together into **components** → **Non-Negative Matrix Factorization (NMF)**
   a) **Components** help to understand **normal** behavior → **Causal Discovery**
   b) Deviations help to detect **abnormal** behavior → **Outlier detection**



Select data

Signal: U_diode from nQPS in PM

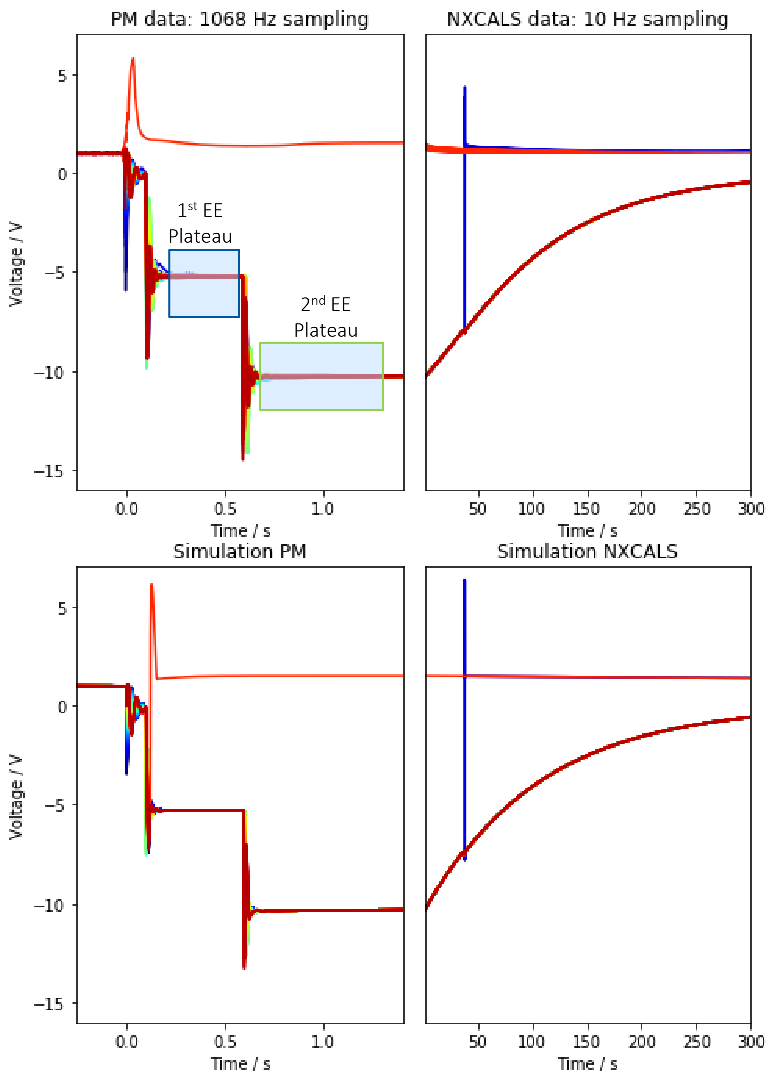
Region: Plateaus after energy extraction

Period: :2018, Quench + Snapshot data

Data size: 731 events x 154 magnets x 400 samples (0.375s)

Selection criterion:

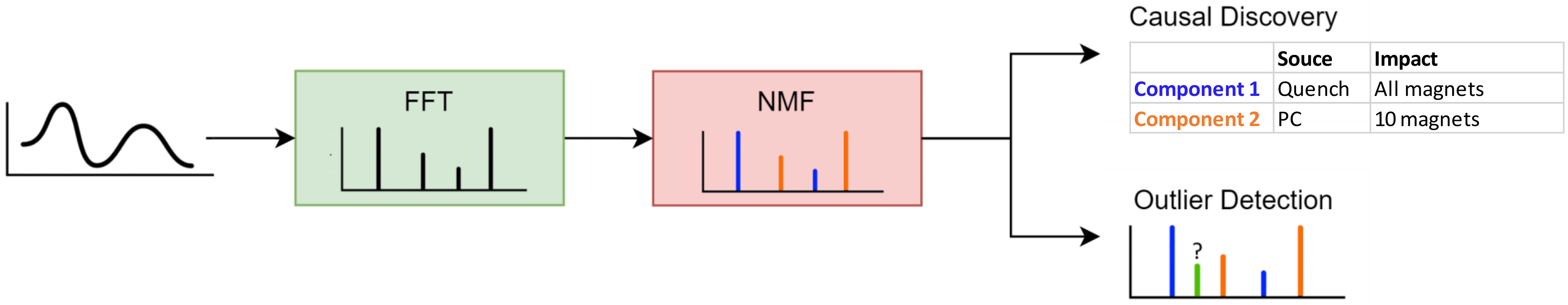
- Unusual behavior (“Wiggle”) in U_diode of RB.A78 B28L8 on 2021-03-28
- PM data available in high resolution (1068Hz)
- EE is similar to transient measurement
- Deviation from data and simulations during plateaus



*simulations done by Marvin Janitschke with STEAM framework

Fast Fourier transform

Extract **frequencies** in data



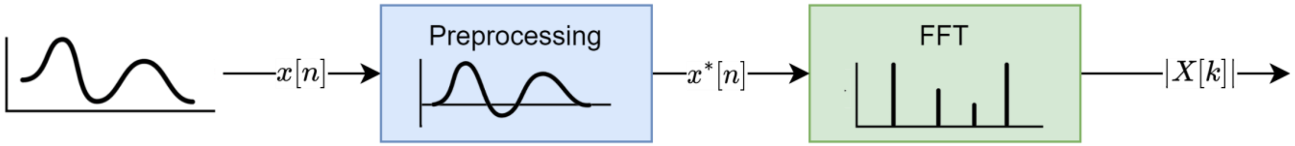
	Souce	Impact
Component 1	Quench	All magnets
Component 2	PC	10 magnets

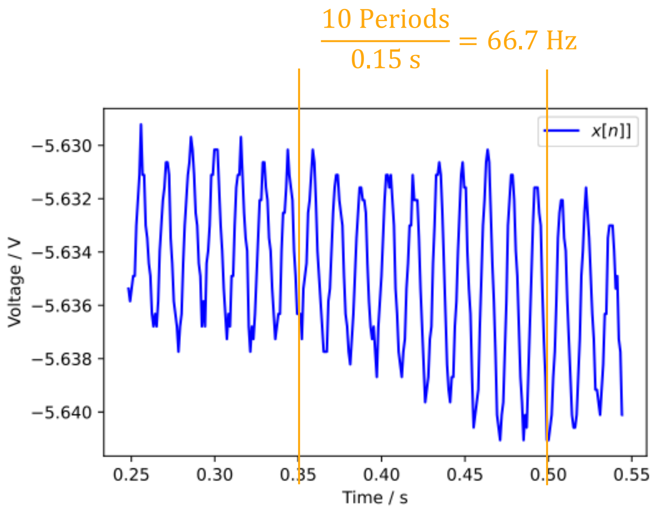
Fast Fourier transform

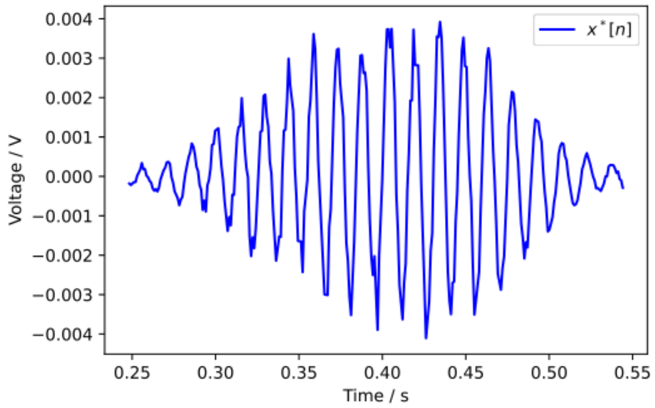
Example signal: B21R3 on 2021-04-18 08:44:17

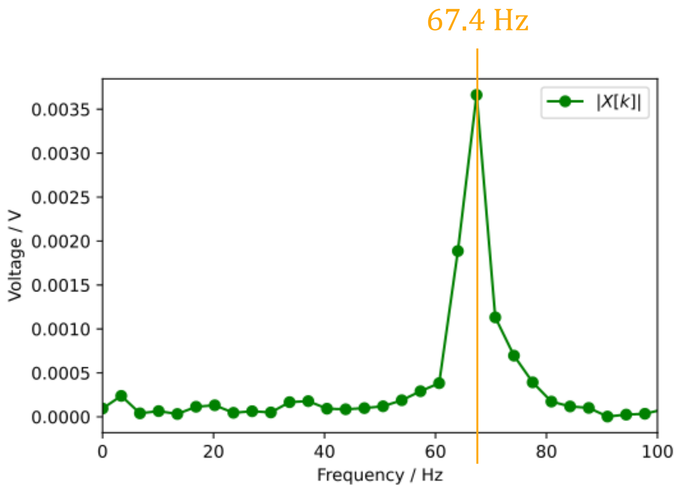
Preprocessing necessary to minimize spectral leakage:

- Subtract offset
- Multiplication with window

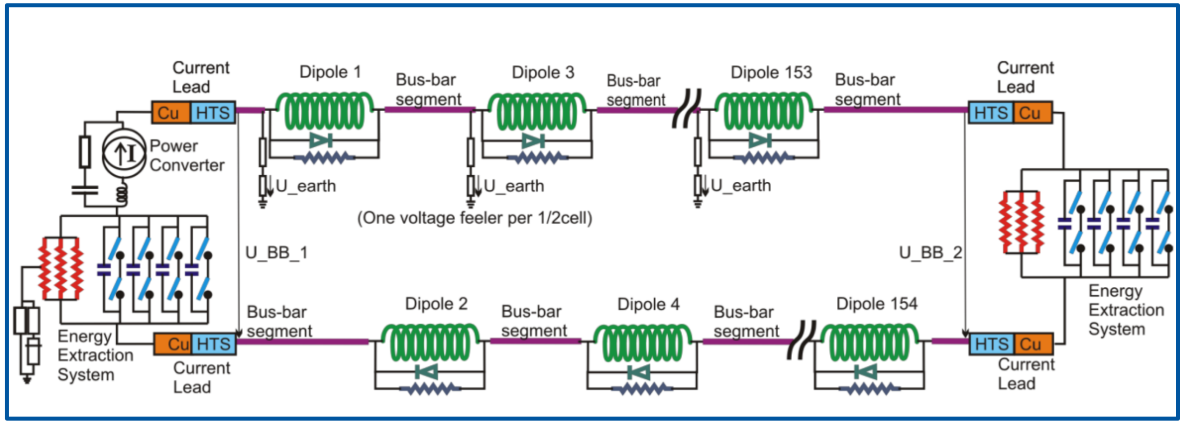


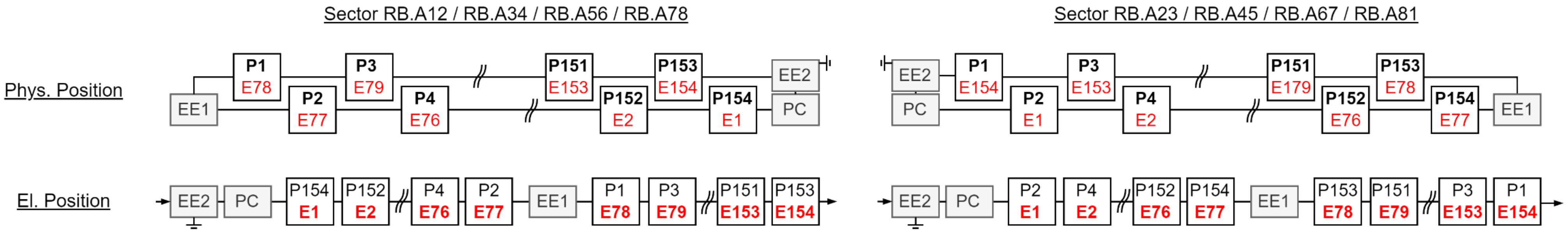




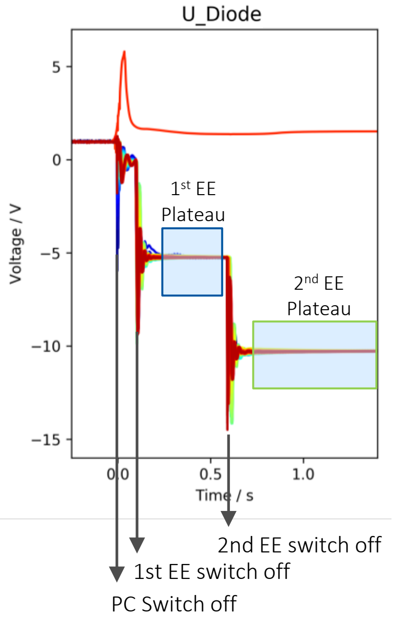


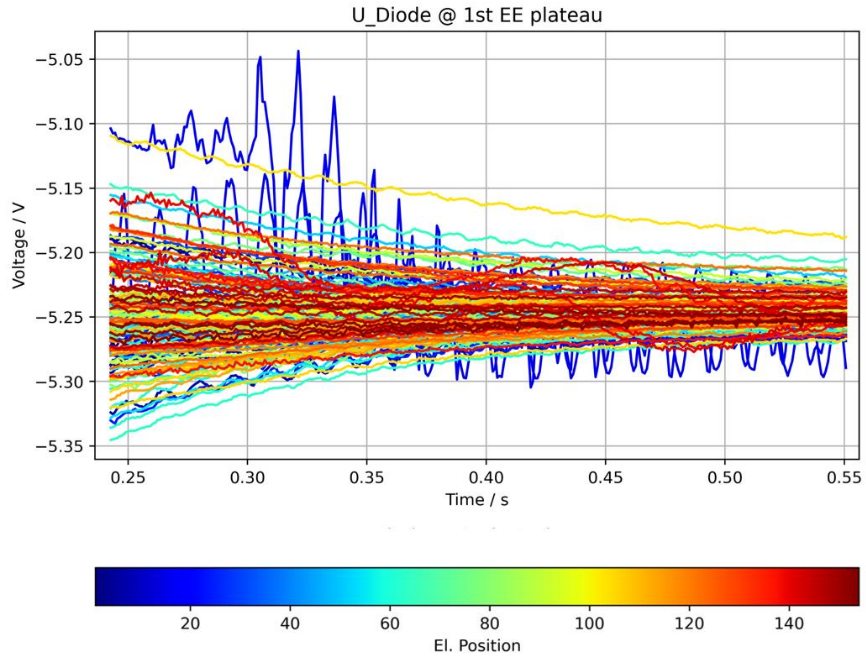
El. Vs Phys. Position





Frequency Position Maps (FPM)

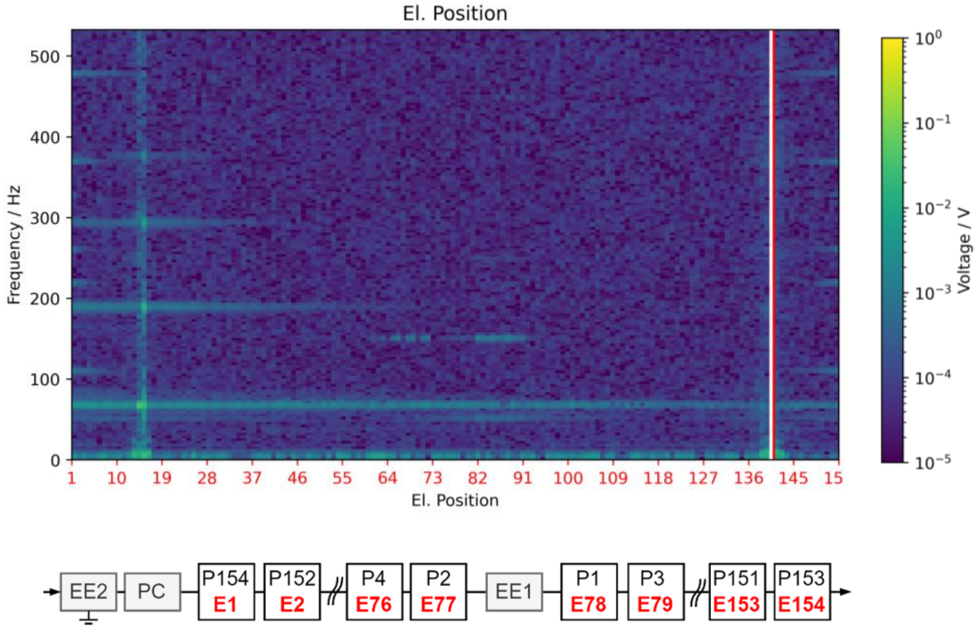




fft

Normal event

FPA identifier: RB_RB.A78_1617170255140000000
Date: 2021-03-31 07:57:35.120000
Max. Current: 11215.0 A

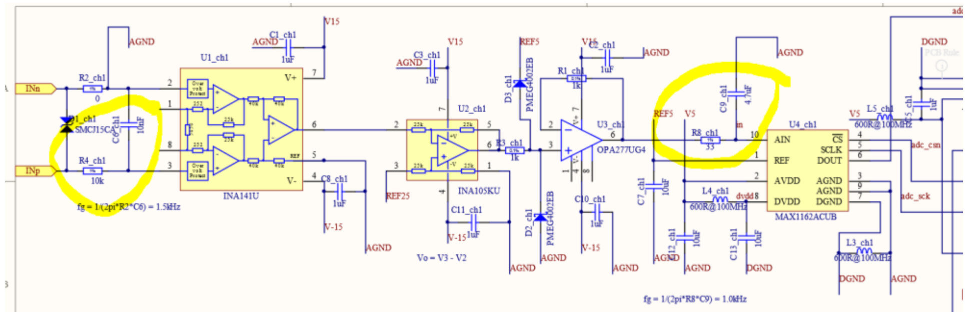


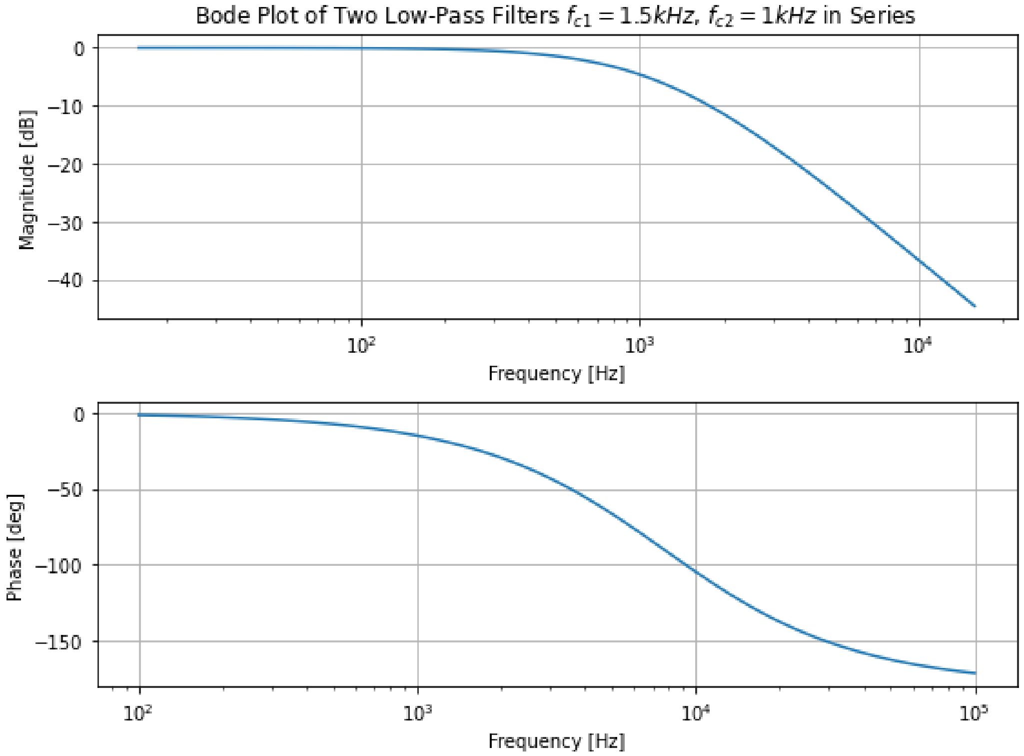
Aliasing

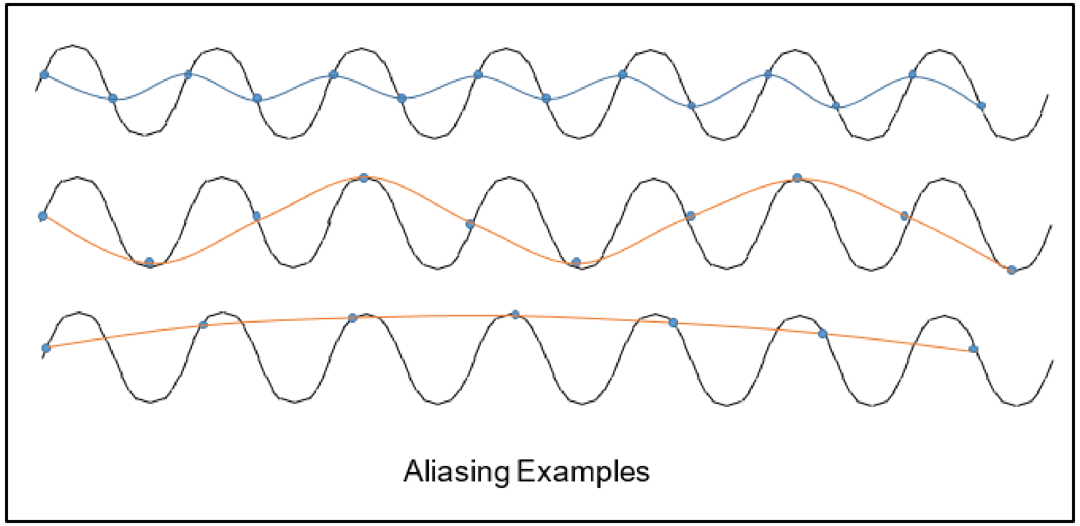
High frequency components could potentially cause **aliasing in results**.

Anti-aliasing filters in the nQPS crates:

- Two 1st order lowpass filters with 1.5 kHz and 1 kHz cutoff frequency*
- Sampling frequency of nQPS crates: 1068 Hz





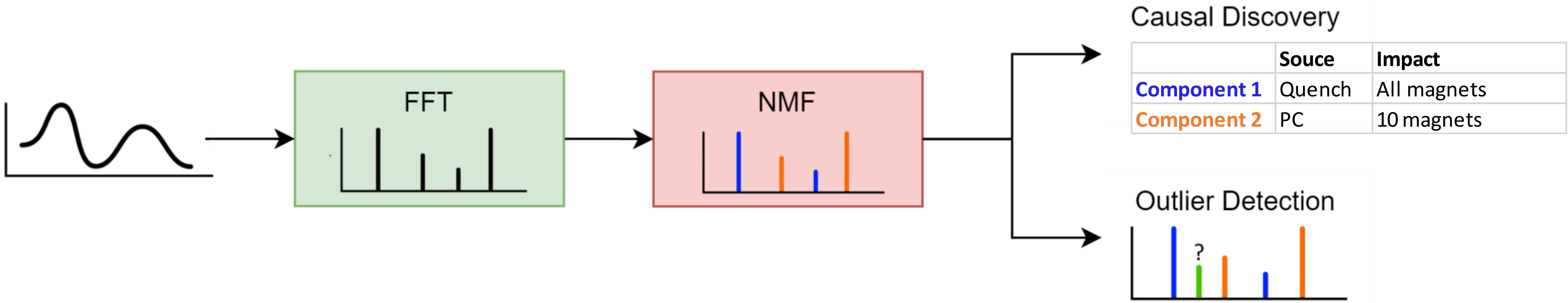


*provided by Jens Steckert

Non-Negative Matrix Factorization (NMF)

Group expected frequencies that occur together into **components**

a) **Components** help to understand **normal** behavior → **Causal Discovery**
b) Deviations help to detect **abnormal** behavior → **Outlier detection**

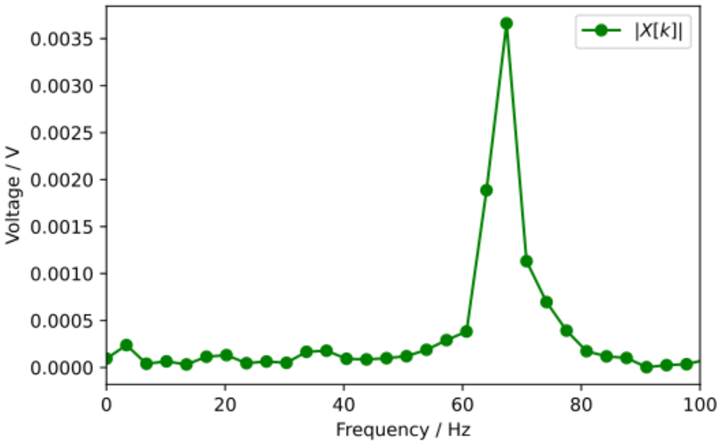


	Souce	**Impact**
Component 1	Quench	All magnets
Component 2	PC	10 magnets

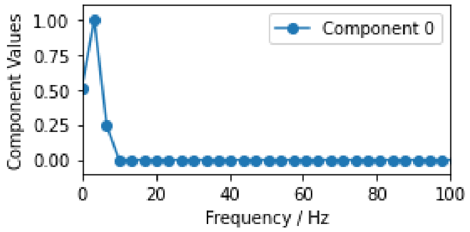
https://www.nature.com/articles/44565
https://proceedings.neurips.cc/paper/2000/file/f9d1152547c0bde01830b7e8bd60024c-Paper.pdf

Method

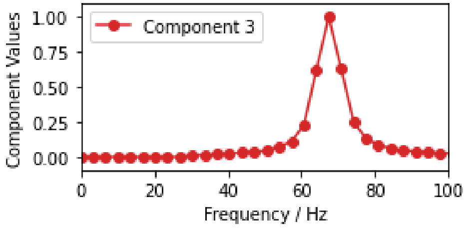
Input		Components	Weights		Reconstruction
$\lvert X[k] \rvert$	$\approx$	$\sum_{k=1}^{r} W_k$	$\cdot \; h_k$	$=$	$\lvert \hat{X}[k] \rvert$



$\approx$

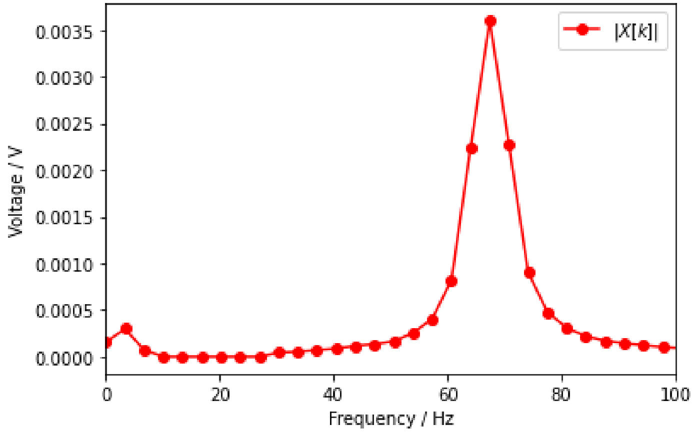


$\cdot$ 0.0003



$\cdot$ 0.0036

$=$

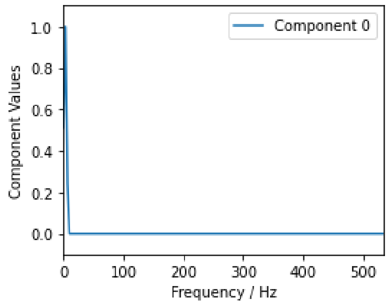
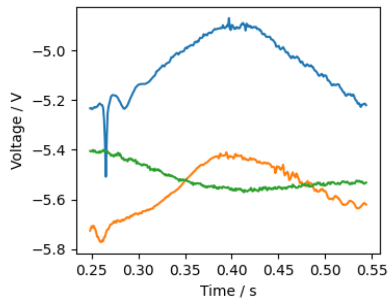
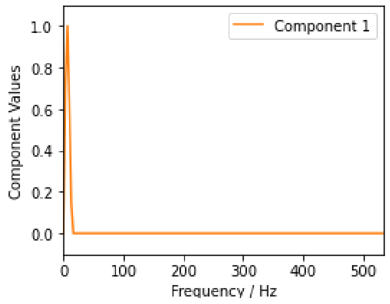
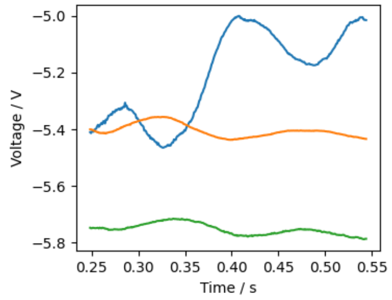
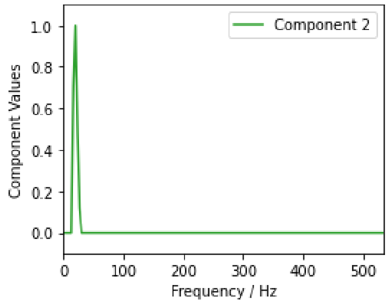
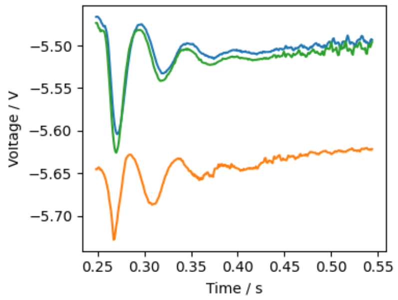


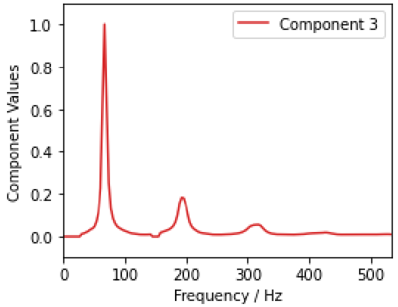
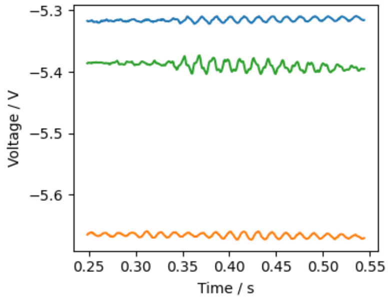
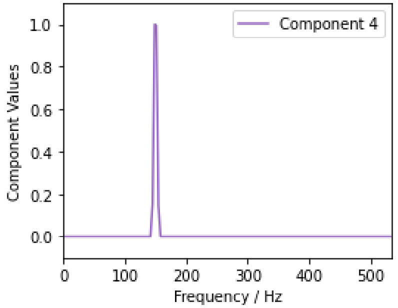
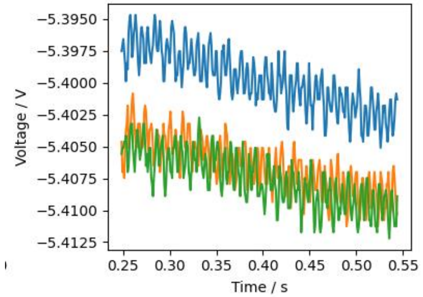
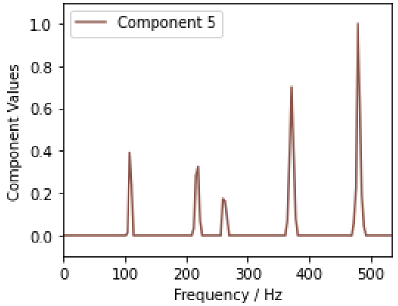
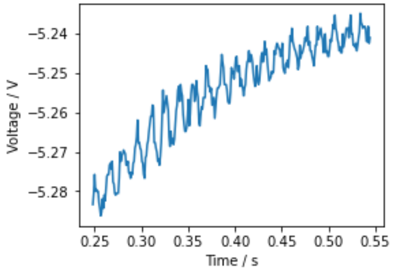
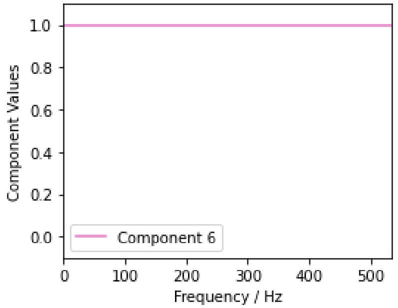
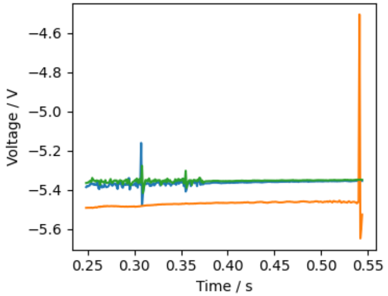
Loss: $\sum_k \left(\lvert X[k] \rvert - \lvert \hat{X}[k] \rvert \right)^2$

How to define W_k & h_k?

1. Manually define r
2. Initialize W_k & h_k randomly
3. Adjust W_k & h_k iterativly until loss over all signals (489 * 154 = 75306) is minimal

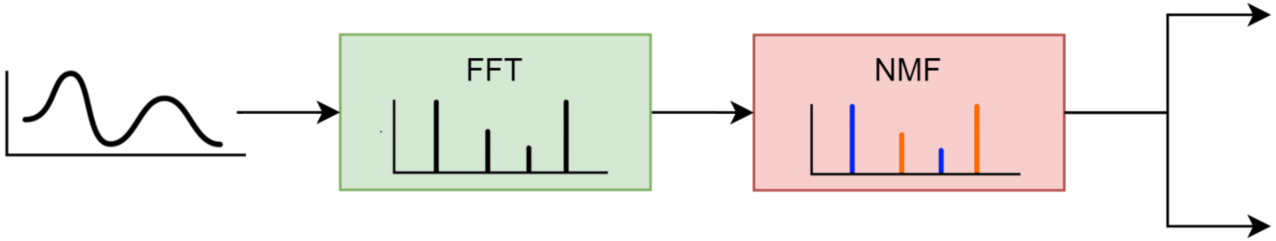
NMF Components

Component Index	Component	3 Event Examples	Dominant Frequencies	Order of magnitude
0*			3 Hz	100mV
1			6 Hz	100mV
2			20 Hz	10mV

Component Index	Component	3 Event Examples	Dominant Frequencies [Hz]	Order of magnitude
3			67 Hz 184 Hz 302 Hz	100mV
4			148 Hz	1mV
5			107 Hz 220 Hz 260 Hz 370 Hz 478 Hz	0.1mV
6			Any	100mV

Causal Discovery

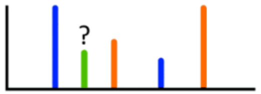
Understand **normal** behavior



Causal Discovery

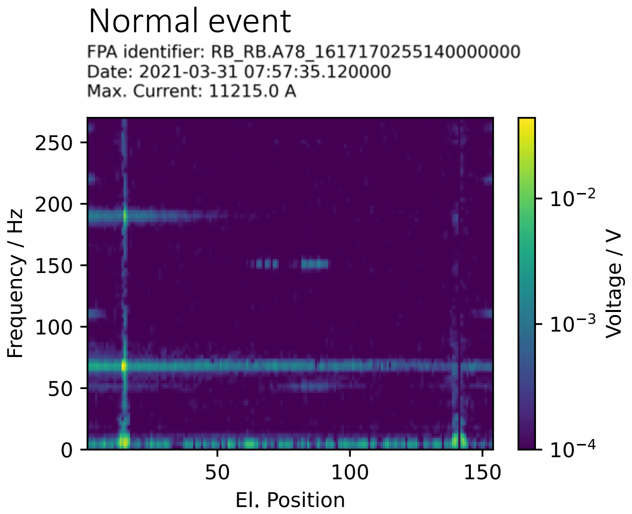
	Souce	Impact
Component 1	Quench	All magnets
Component 2	PC	10 magnets

Outlier Detection

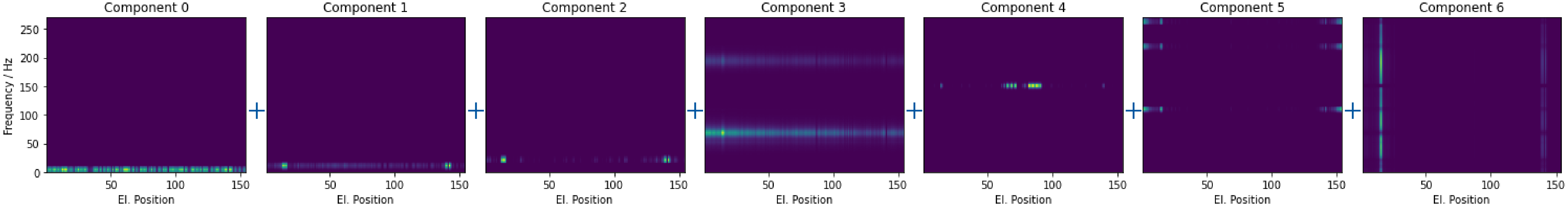


FPM of Components

Frequency **components** along all magnets in event



≈



3Hz	6Hz	20Hz	65Hz	143Hz	478Hz	Any
Electromagnetic perturbation + FFT smiring	Electromagnetic perturbation	Left over of voltage wave traveling along the chain of magnets by magnet inductance	Oscillations caused by quench	Repetitive pattern within QPS crate, aliasing?	Oscillations from the direction of Power converter in RB.A78, aliasing?	Oscillations caused by quench
Source Direction: Quench	Source Direction: Quench	Source Direction: Quench	Source Direction: Quench	Source Direction: Quench	Source Direction: Power converter in RB.A78	Source Direction: Quench
Propagation: Physical	Propagation: Physical	Propagation: Physical	Propagation: Electrical	Propagation: QPS crate position	Propagation: Physical	Propagation: Physical

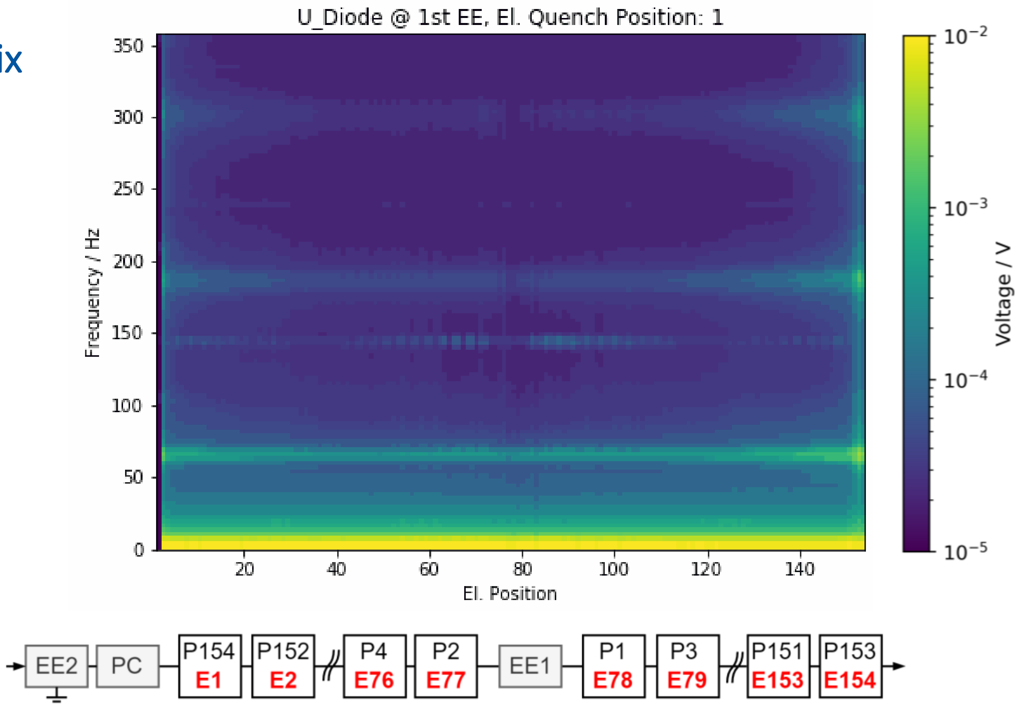
Conclusion Causal Discovery

- Detection of “normal” frequency components with **Non-Negative Matrix Factorization**
- Components allow identification of different frequency **sources**.
- Frequency Position Maps allow to find **propagation** direction

→ Depending on the quench position, a typical FPM would look like this:

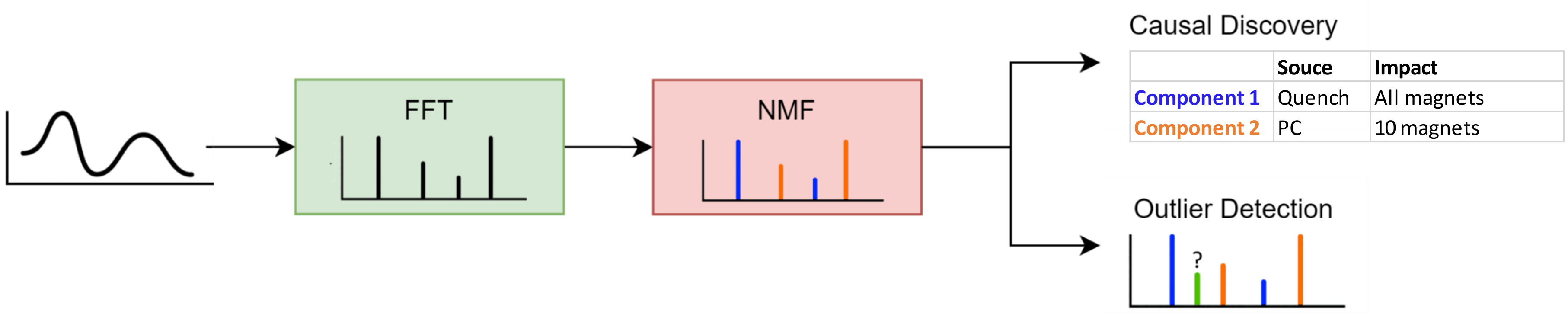
Components with **quench as source**, do not occur in **snapshot** data.

Components in **1st and 2nd EE plateau** are similar.



Outlier Detection

Detect **abnormal** behavior



Approach

Components state frequencies, **expected** to occur

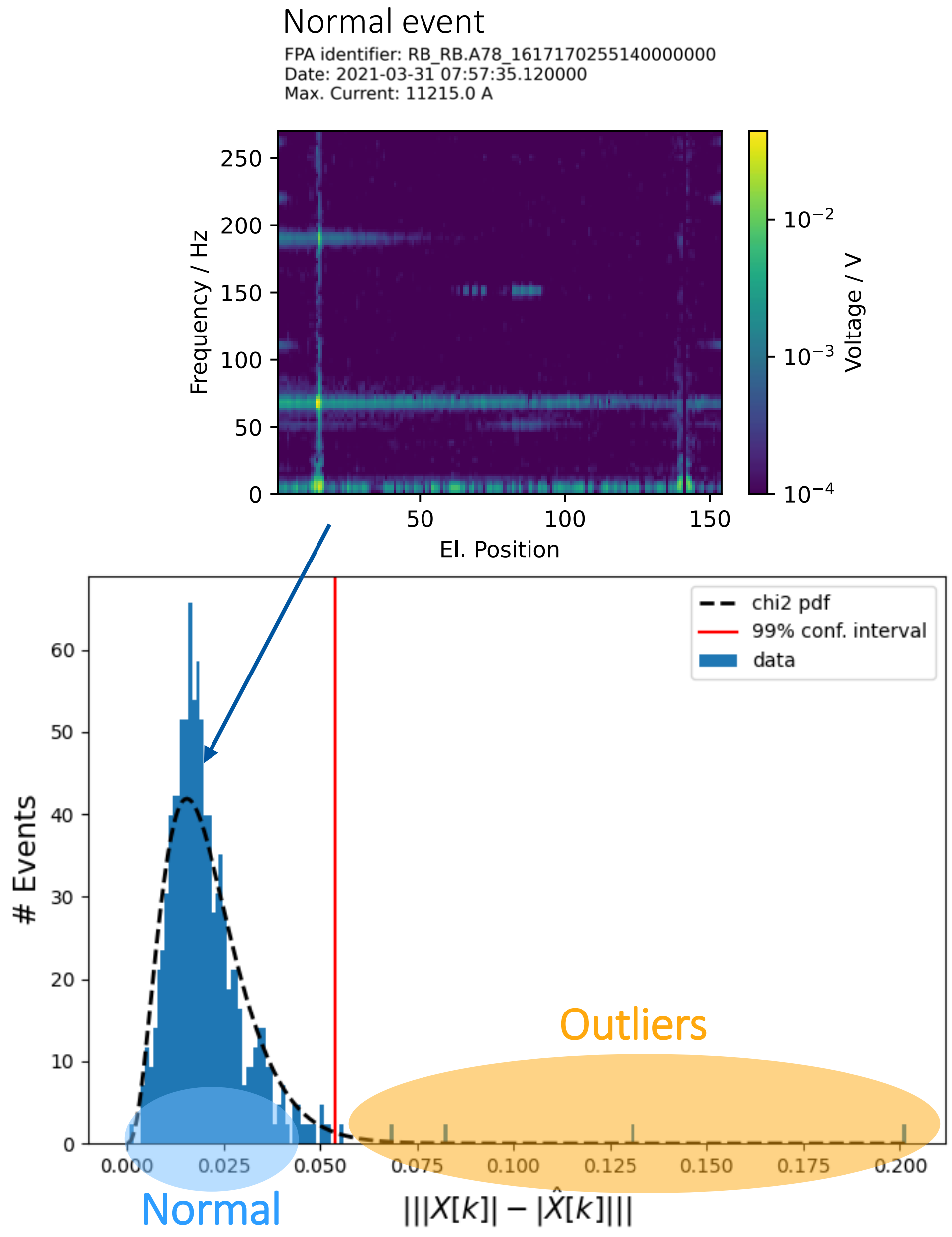
Normal event: Reconstruction with components possible (low loss)

Outlier event: **Unexpected** frequencies occur (high loss)

How to find an outlier:

1. Calculate NMF loss for each event (731)
2. Fit distribution to loss
3. Calculate p value for each event (731)

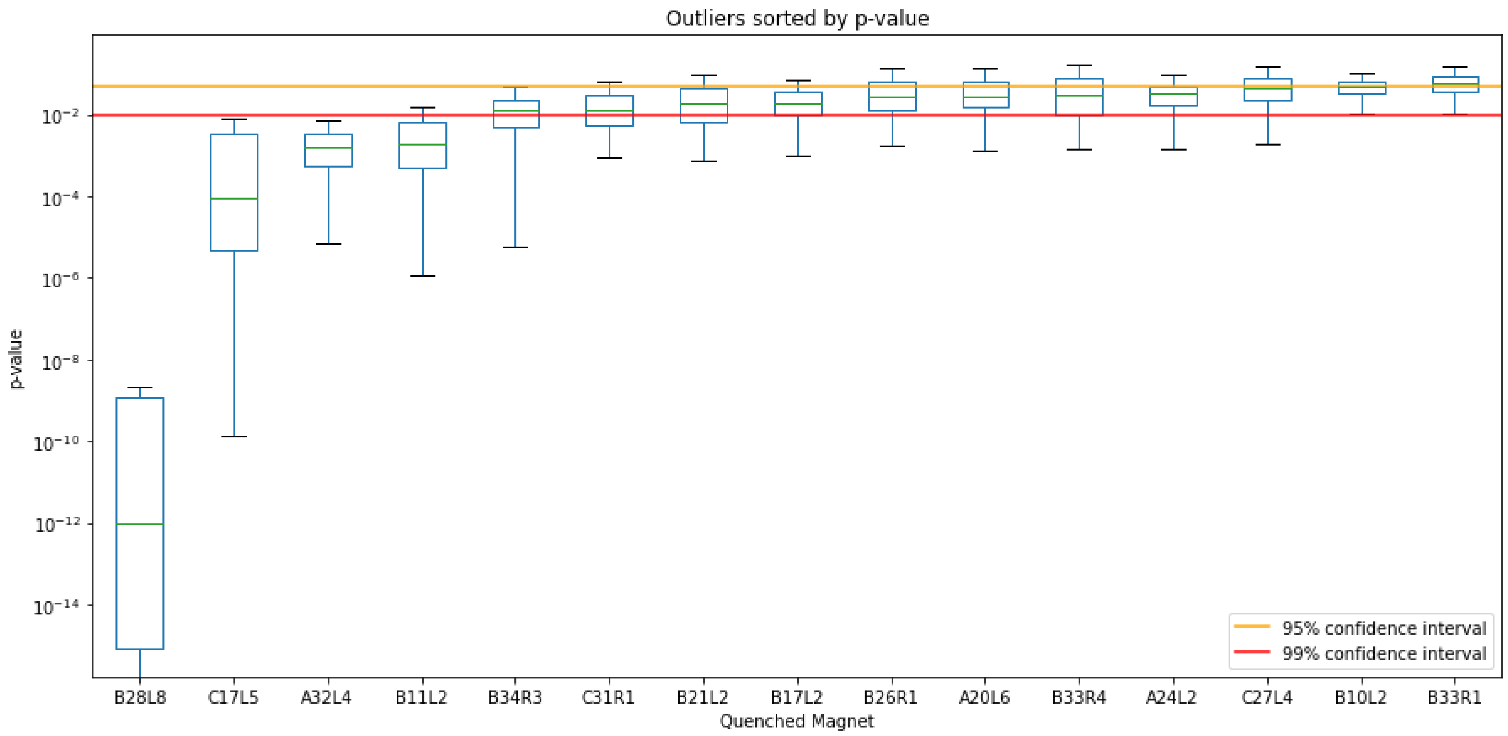
   → probability of obtaining results at least as extreme as the observed

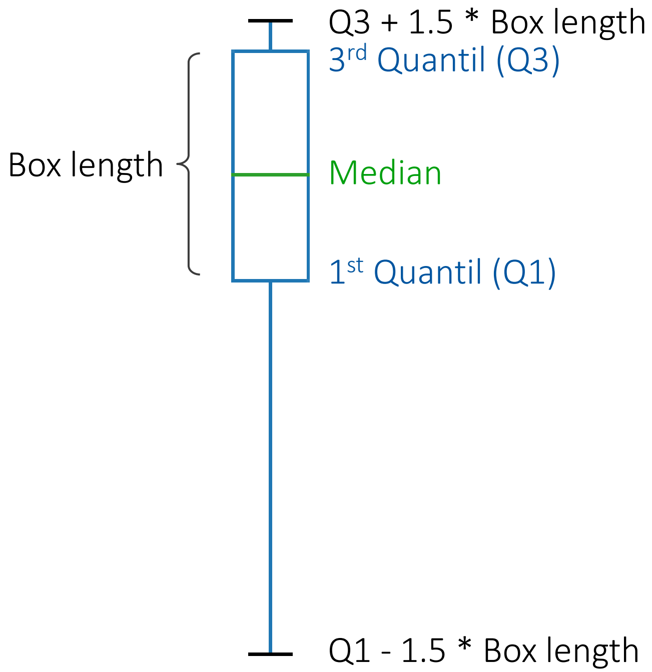


Outliers

Goal: Find outliers robust to assumptions

- Result shows boxplot of 280 different combinations of assumptions
- All outliers occur during 1st EE Plateau → closer in time to quench





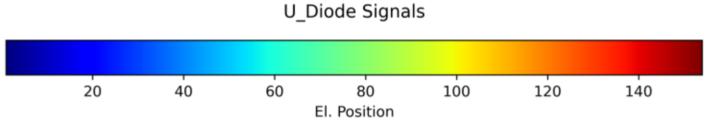
RB.A78 - B28L8 – Intercoil short

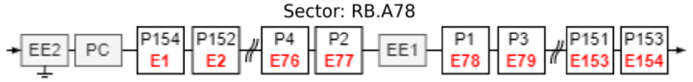
Event with B28L8 quench before:

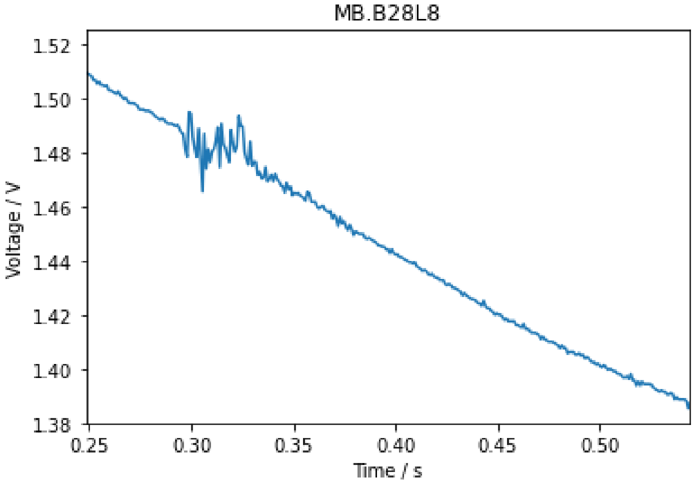
- 2021-03-28: „normal“ at EE plateaus

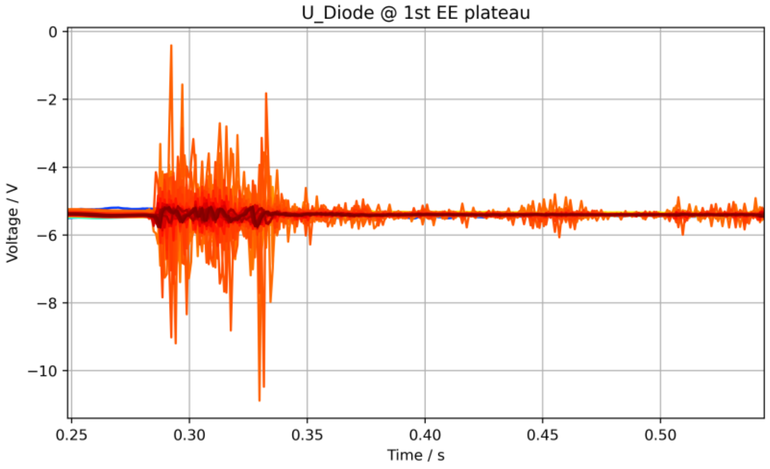
FPA identifier: RB_RB.A78_1619330143440000000
Date: 2021-04-25 07:55:43.418000
Max. Current: 11588.0 A

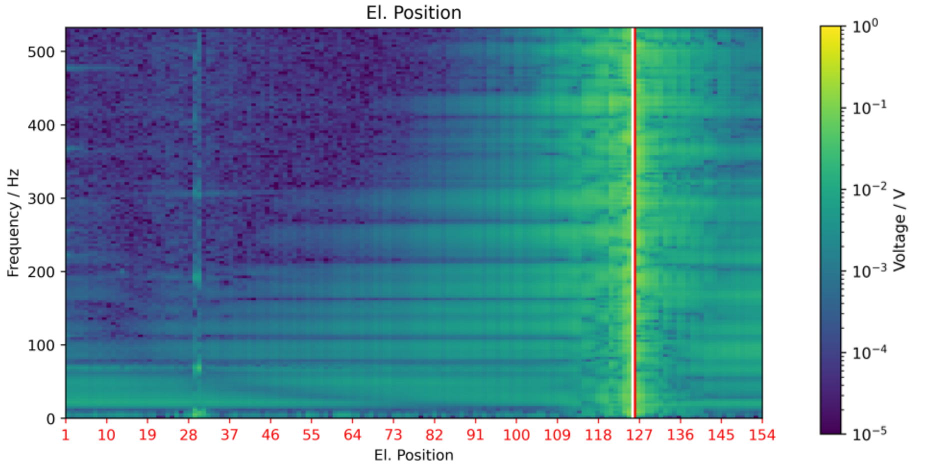
El. Position Primary
Primary quench position: 126
Fast secondary quench: []









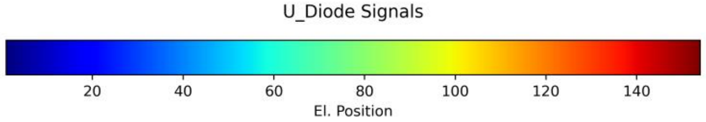


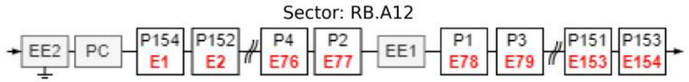
RB.A12 - B11L2

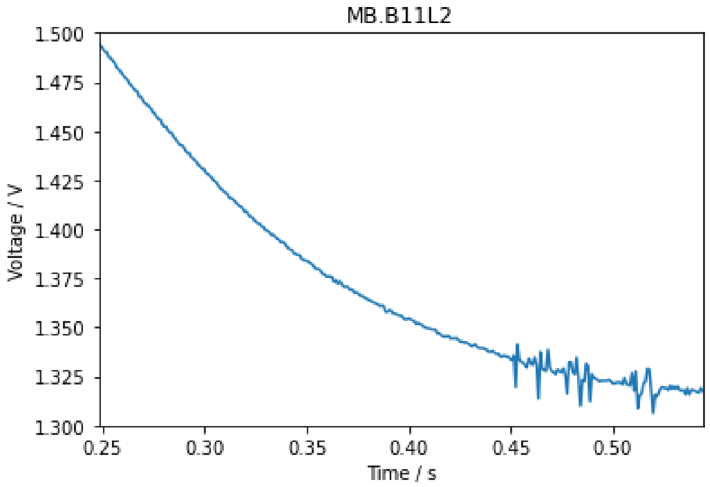
No B11L2 quench before

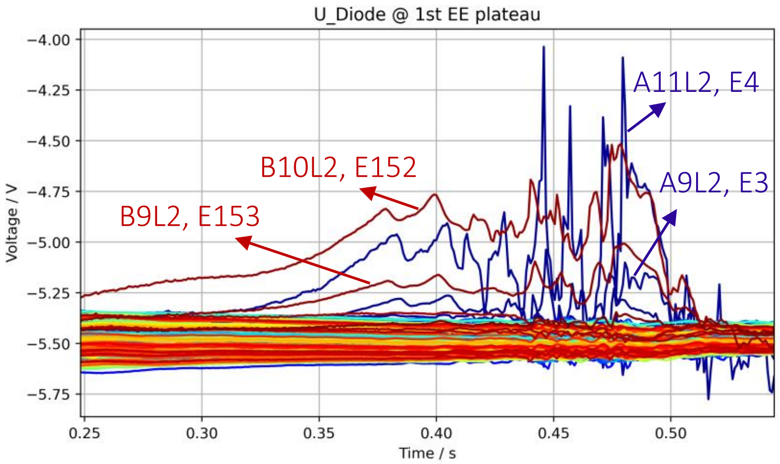
FPA identifier: RB_RB.A12_1621014819920000000
Date: 2021-05-14 19:53:39.901000
Max. Current: 11751.0 A

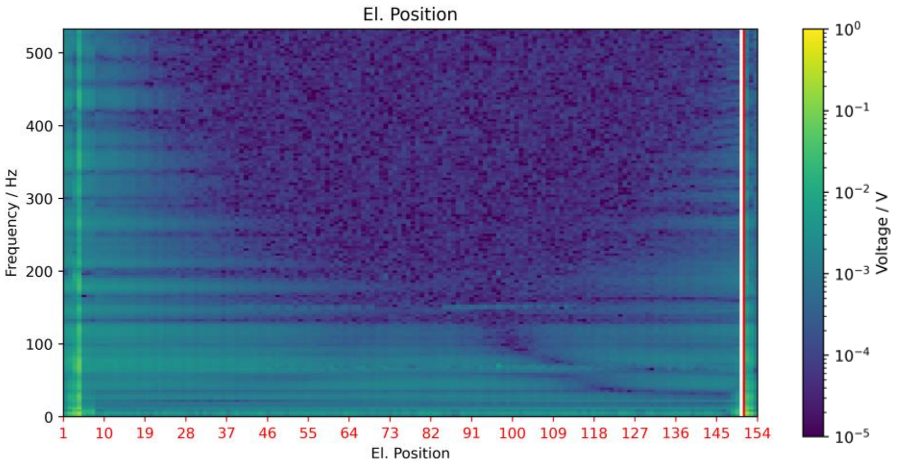
El. Position Primary
Primary quench position: 151
Fast secondary quench: []











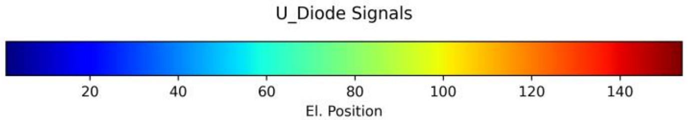
RB.A34 – B28R3

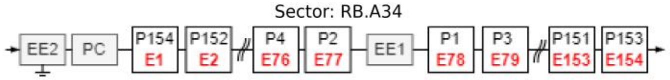
No B28R3 quench before

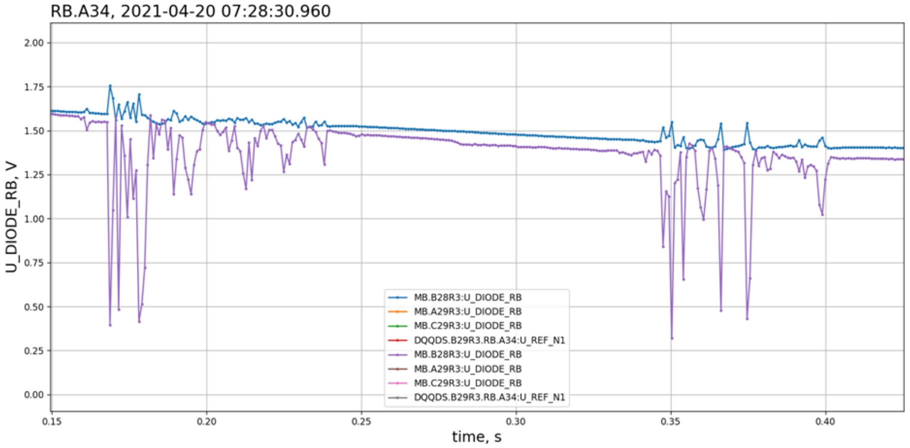
Most likely scenario of noise from simulations:
- Partially emerging resistor, in parallel to diode
- Degraded diode contact?

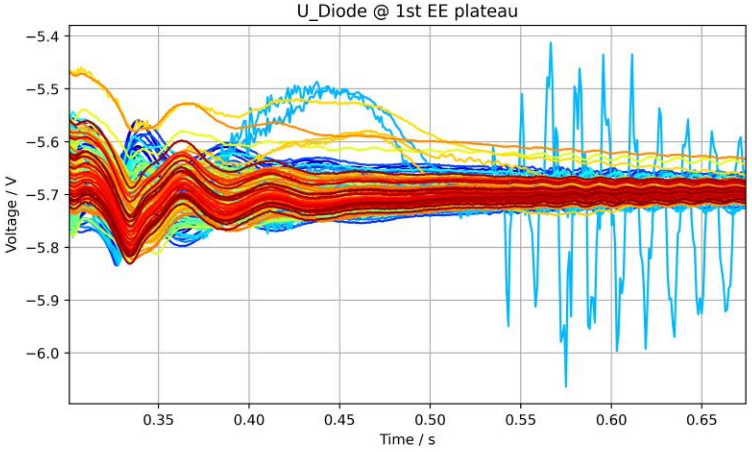
FPA identifier: RB_RB.A34_1618896510960000000
Date: 2021-04-20 07:28:30.924000
Max. Current: 11786.3 A

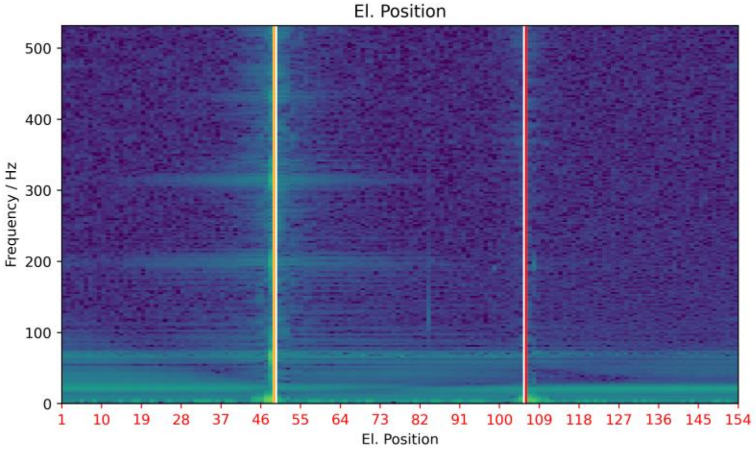
El. Position Primary
Primary quench position: 106
Fast secondary quench: ['49@198ms']











Conclusion Outlier Detection

- Outliers are **events** which cannot be composed out of "normal" frequency components
- Several statistically relevant outliers highlighted
  - All of them are in **1st EE** → closer in time to quench

Ongoing additional **measurements, additional safety measures, possible replacement.**

Next Steps:

- Creation of a MB metadata repository with relevant features from different studies*
  - Goal: **Input for making decisions**
- Modification of **Sigmon** FPA plots planned
- Analysis of further events and signals planned
  - Secondary quenches of U diode
  - U_QS0, U_HDS, I_HDS

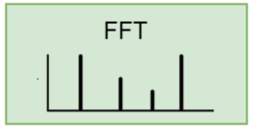
*https://gitlab.cern.ch/machine-protection/mb-feature-classification

CERN

Backup Slides

Fast Fourier transform

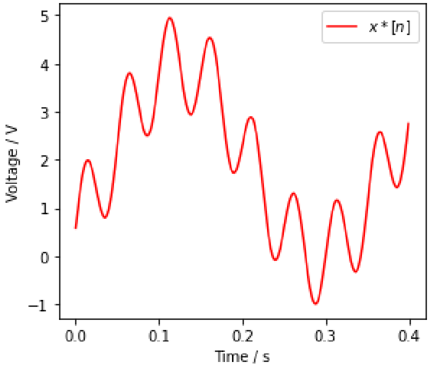
Fast Fourier transform

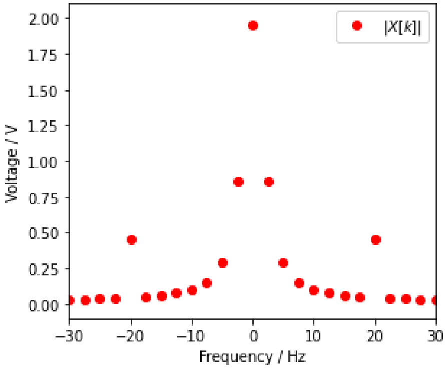


→ The FFT is an algorithm to calculate the discrete Fourier transform (DFT). The DFT is defined as:

$$X[k] = \sum_{n=0}^{N-1} x[n] e^{-i2\pi \frac{nk}{N}}$$

$x[n] = 2\text{V} + 2V * \sin(2\pi\ 3\text{Hz} * n\ - 90°)\ + \sin(2\pi\ 20\text{Hz} * n)$

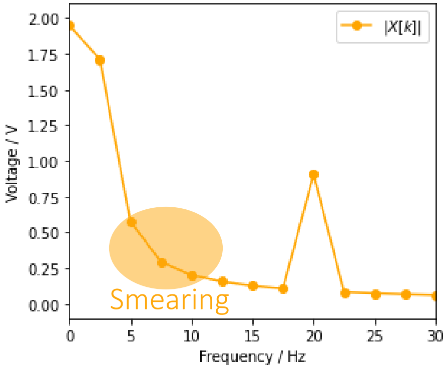




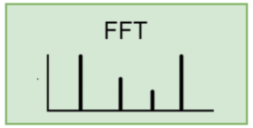
Apply Hilbert transform:
Multiply all but first sample of FFT amplitude with 2

$$z = x + i\text{H}(x)$$

$$\text{where } \text{H}(x) = \begin{cases} -ix, & f > 0 \\ 0, & f = 0 \\ ix, & f < 0 \end{cases}$$

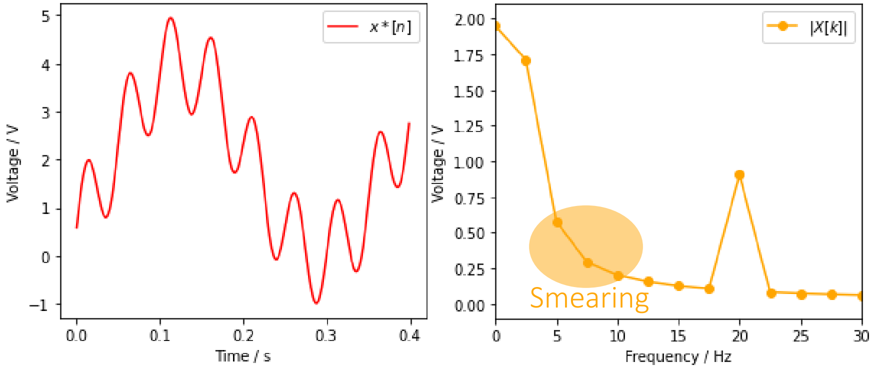


Window data

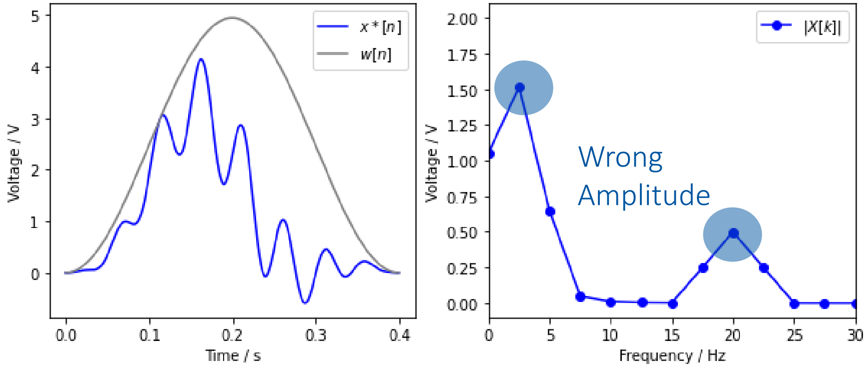


→ Avoid smearing

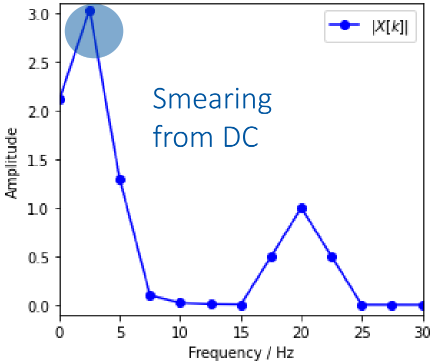
$$x[n] = 2\text{V} + 2V * \sin(2\pi\ 3\text{Hz} * n\ - 90°)\ + \sin(2\pi\ 20\text{Hz} * n)$$

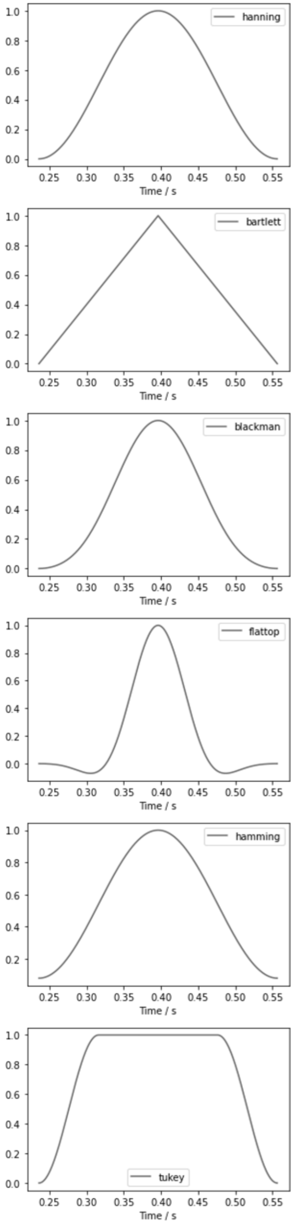


Use Window:
$x^*[n] = x[n] * w[n]$

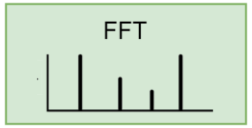


Add Window gain:
Divide $|X[k]|$ with $\frac{1}{N}\sum_n^N w[n]$



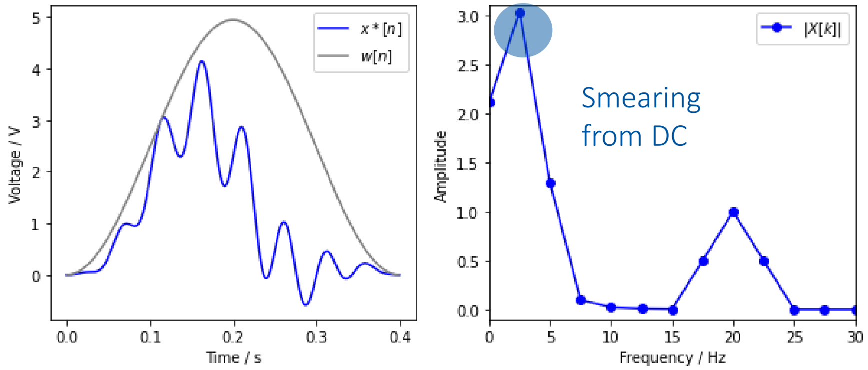


Detrend data

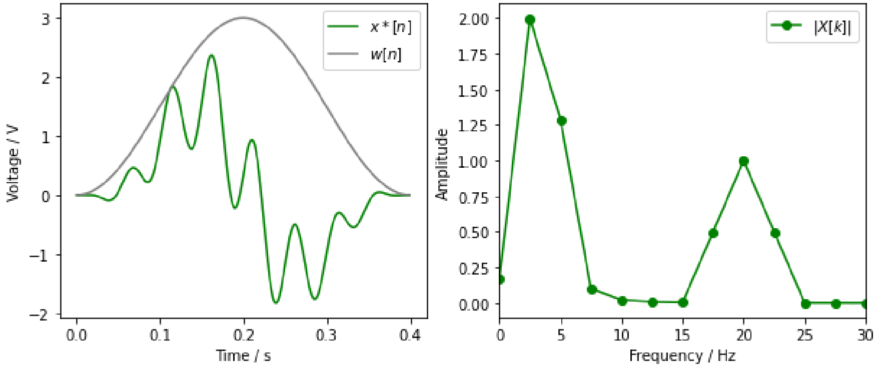


→ Smearing of DC component interferes with low frequency component

$x[n] = 2V + 2V * \sin(2\pi\ 3\text{Hz} * n\ - 90°)\ + \sin(2\pi\ 20\text{Hz} * n)$

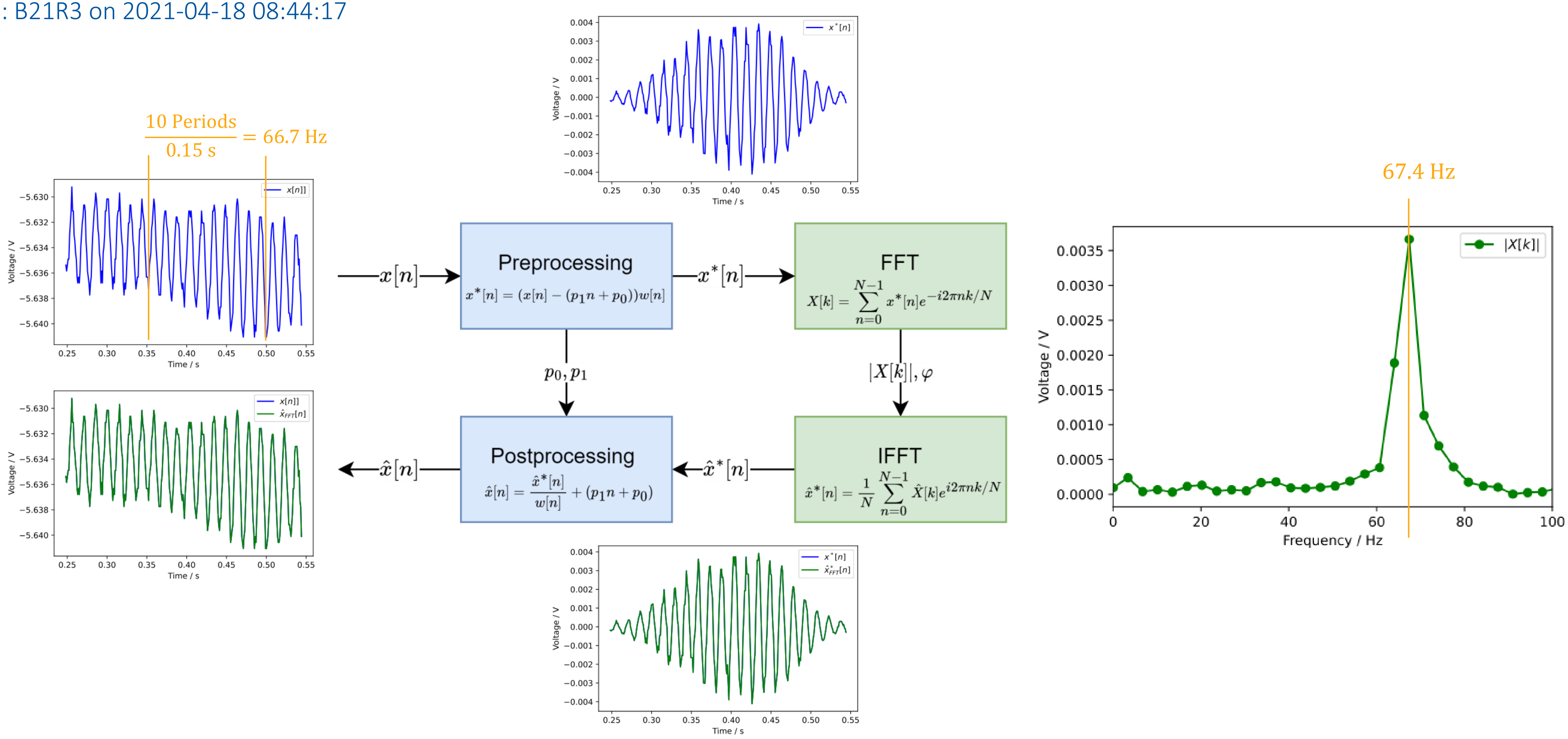


Subtract trend:
$x^*[n] = (x[n] - (p_1 n + p_0)) * w[n]$



Backwards Path

Signal: B21R3 on 2021-04-18 08:44:17



Non-Negative Matrix Factorization (NMF)

https://www.nature.com/articles/44565
https://proceedings.neurips.cc/paper/2000/file/f9d1152547c0bde01830b7e8bd60024c-Paper.pdf

Objective Function

$V \approx WH$

m ... number of i **events * positions** (560 x 154)

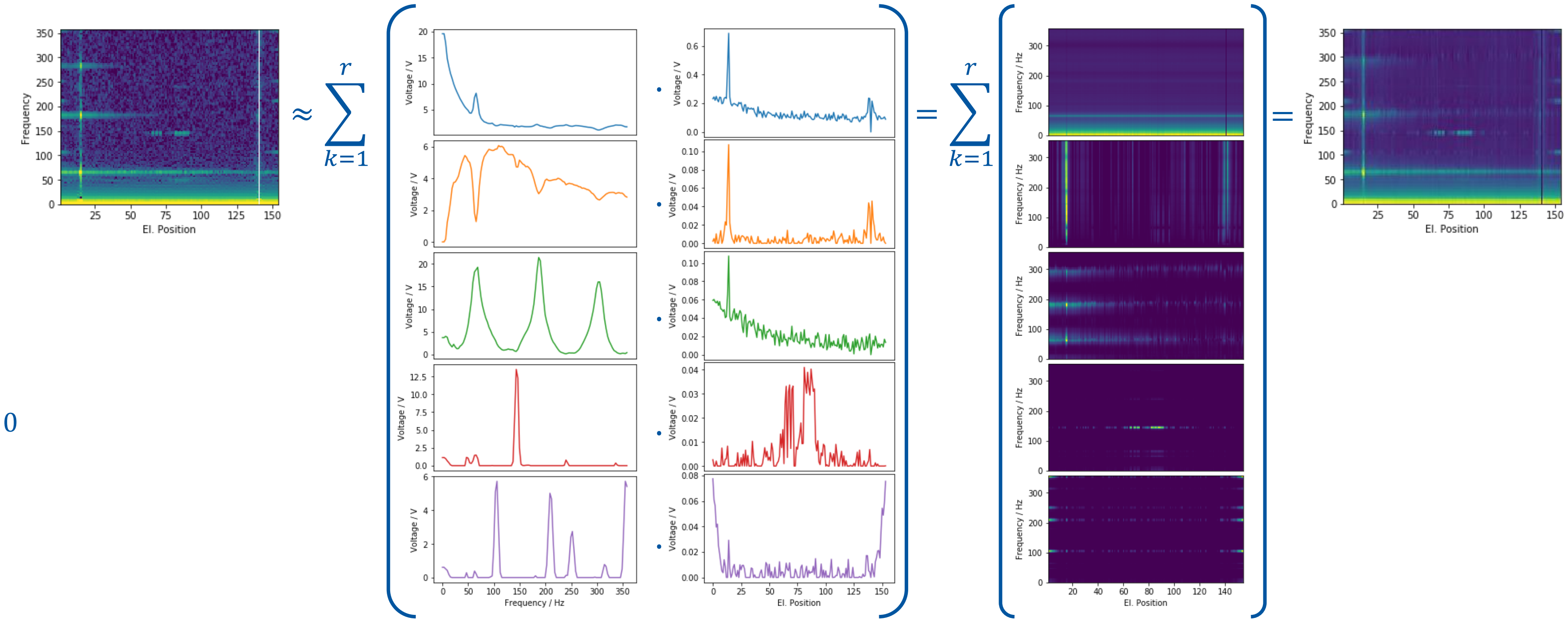
n ... number of **frequencies** (0-360Hz)

r ... number of components (1-5)

$V \in \mathbb{R}_+^{n \times m}$ ... reconstructed event at position

$W \in \mathbb{R}_+^{n \times r}$ ... components

$H \in \mathbb{R}_+^{r \times m}$... presence of components

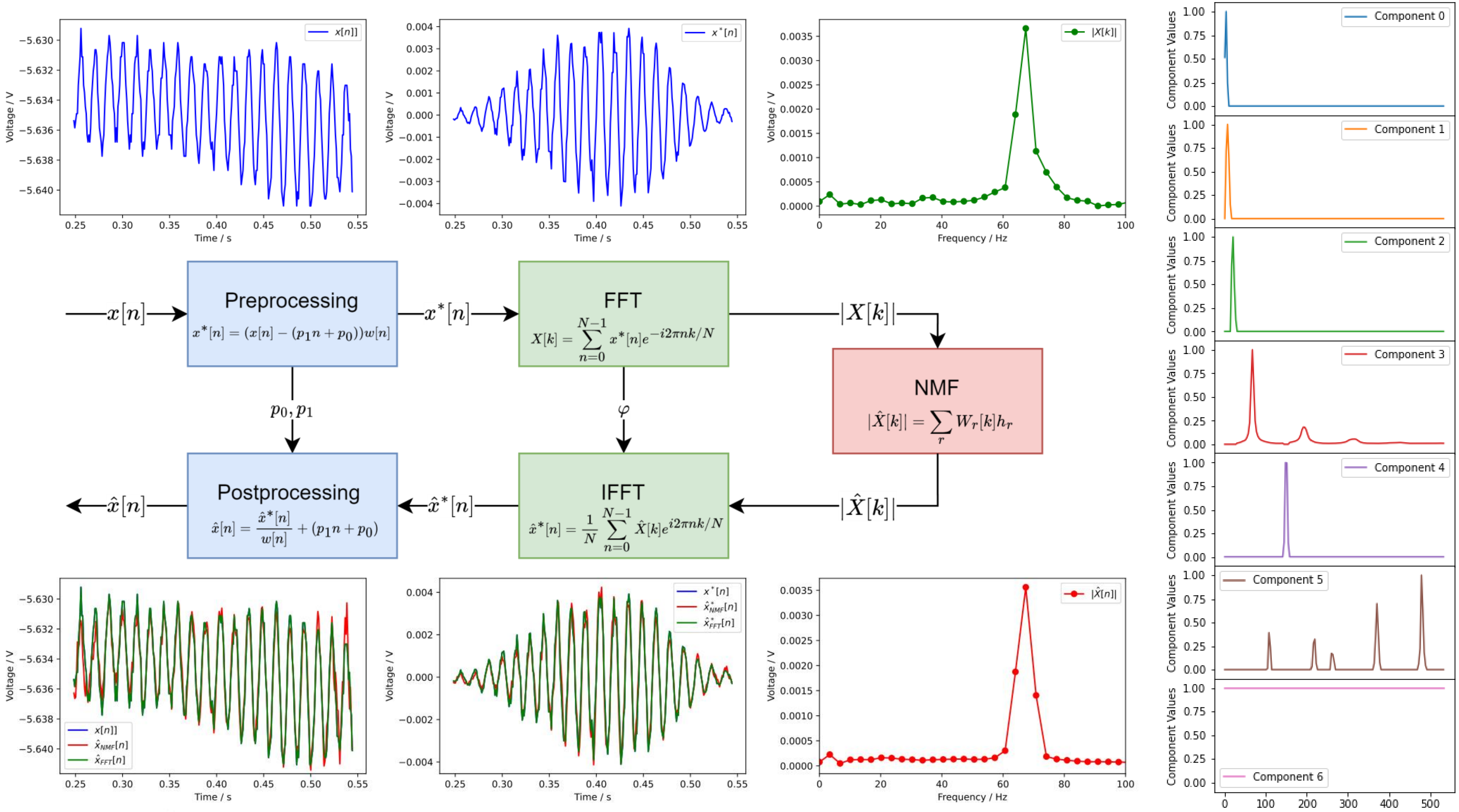


NMF Objective Function:

$$\min_{W,H} f(W,H) \equiv \frac{1}{2}\|V - WH\|_F^2, \text{ s.t. } W, H \geq 0$$

Backwards Path

Signal: B21R3 on 2021-04-18 08:44:17

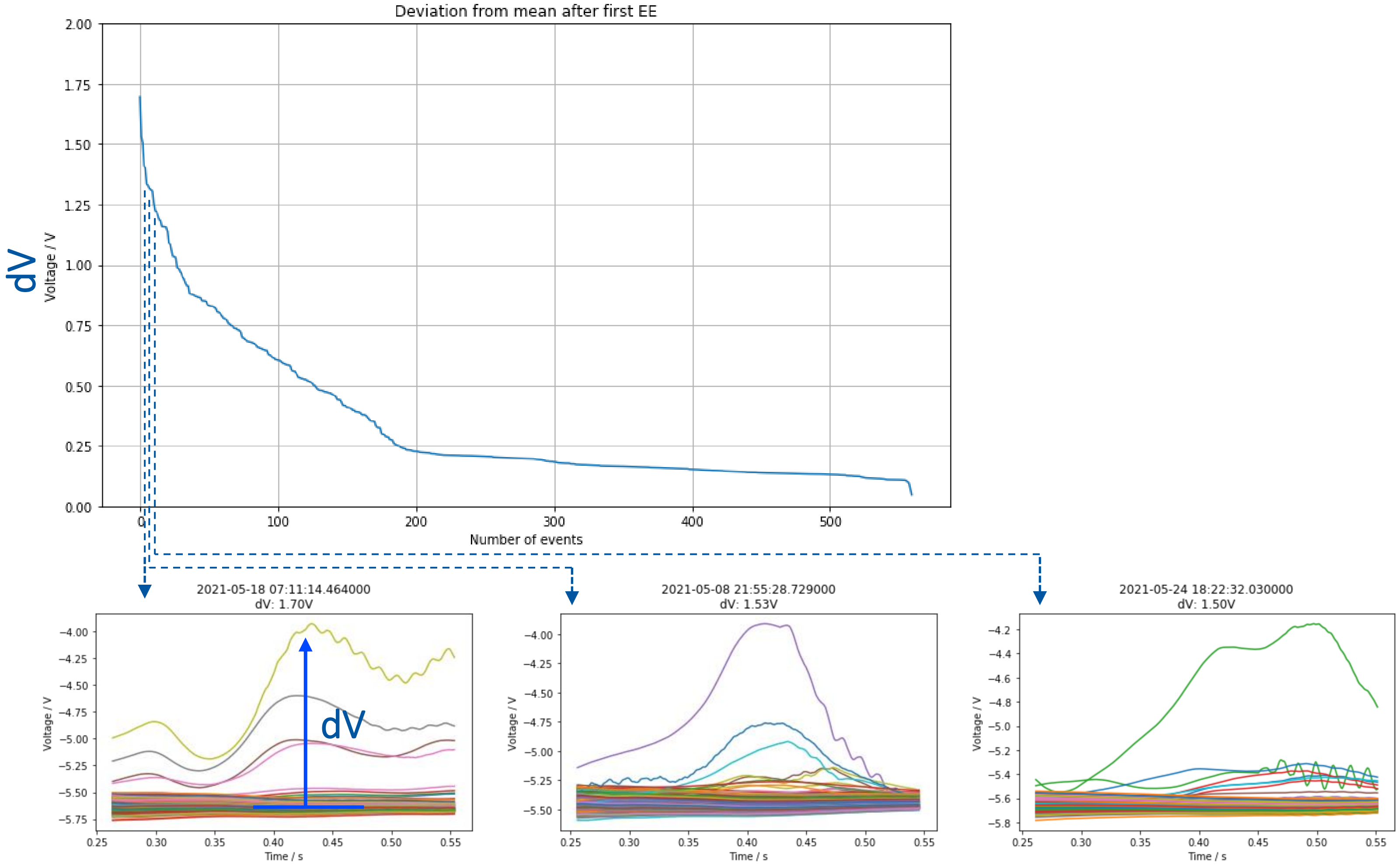


Causal Discovery

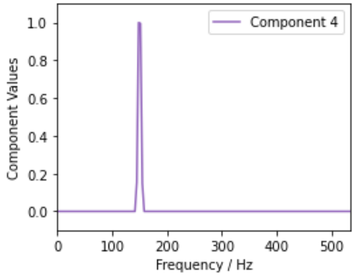
Component 0 & 1: EM Perturbation

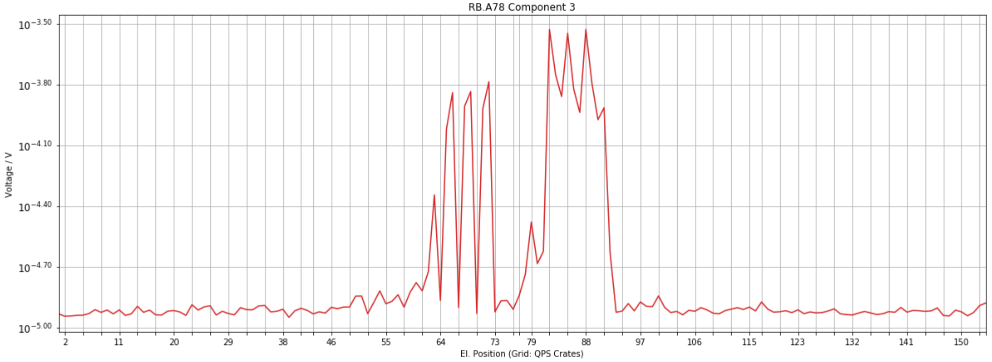
Filter: 20ms moving average

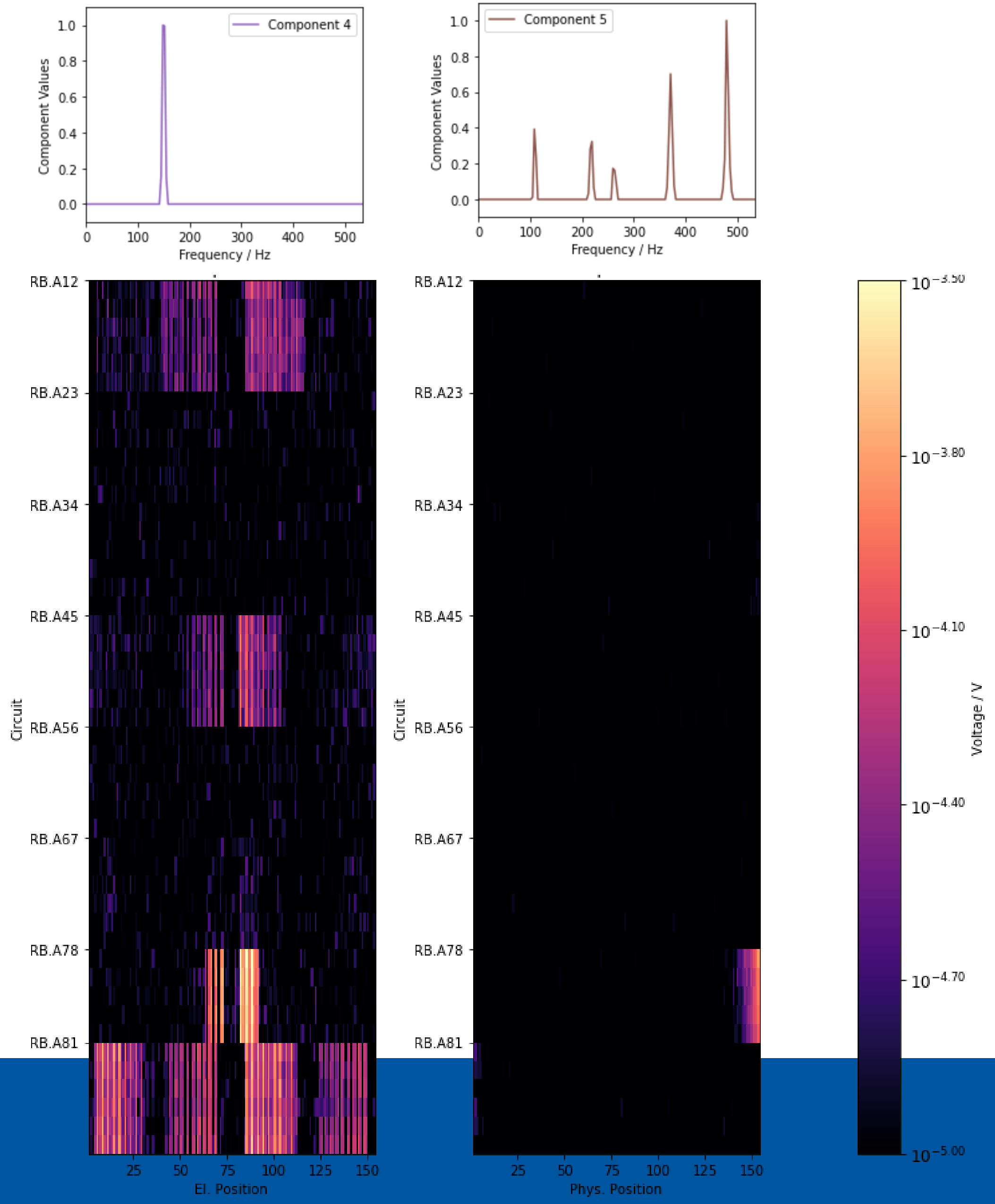
dV is measured similar to sunglass threshold



Component 4 & 5







Outlier Detection

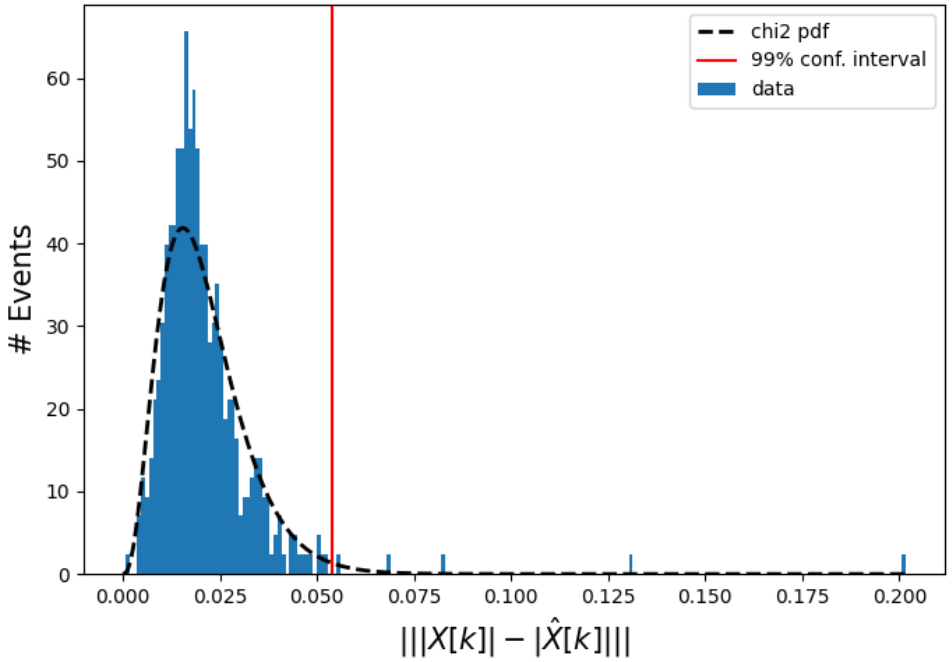
Assumptions

1. **Preprocessing**:
    1. Degree of detrend:
        1. 0 - Offset
        2. 1 - Linear Trend
    2. Window multiplication:
        1. none (best reconstruction, high smearing)
        2. hanning (lowest smearing, no reconstruction)
        3. hamming (low smearing , good reconstruction)
        4. barlett
        5. blackman
        6. flattop (high smearing, accurate amplitude)
        7. tukey
2. **NMF**:
    1. n_components (2-12)
    2. Distance measure*:
        1. Frobenius (Eu)
        2. Kullback-Leibler (KL)

280 x

Assumptions for this plot:

- Linear detrend
- Hamming window
- 4 components
- Frobenius distance



* https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6410389

RB.A34 - A32L4 – Event before