SDSC

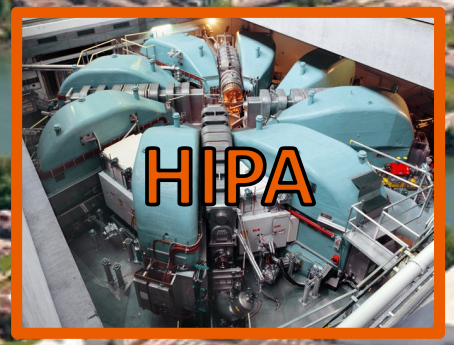
RAPID ELEMENTAL ANALYSIS FROM MUONIC X-RAY SPECTRA

Tomasz Kacprzak (Swiss Data Science Center, PSI), Michael Heiss (LMU, PSI)

with: Alex Amato, Thomas Prokscha, Gianluca Janka, Arndt Remhof, Benjamin Haro (PSI)

PSI Large Research Facilities

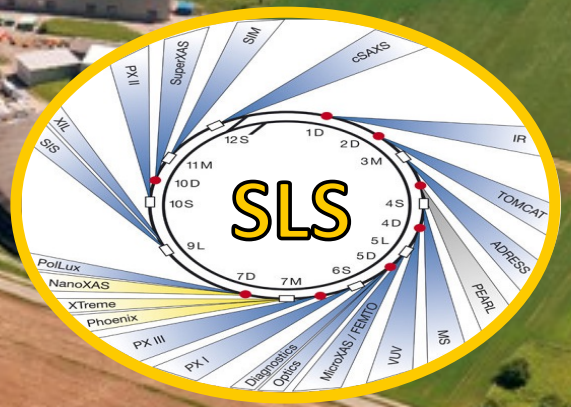
High Intensity
Proton Accelerator



Swiss Spallation
Neutron Source

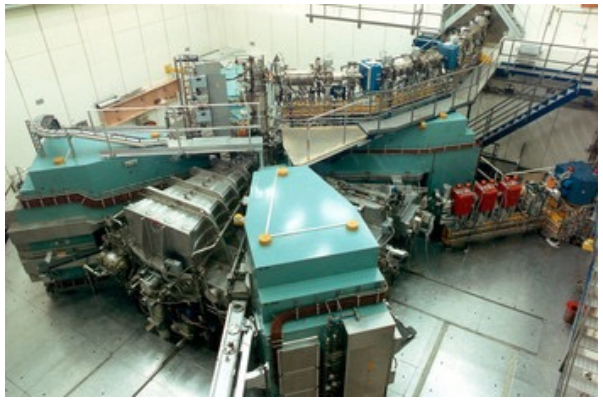
Swiss Free Electron Laser

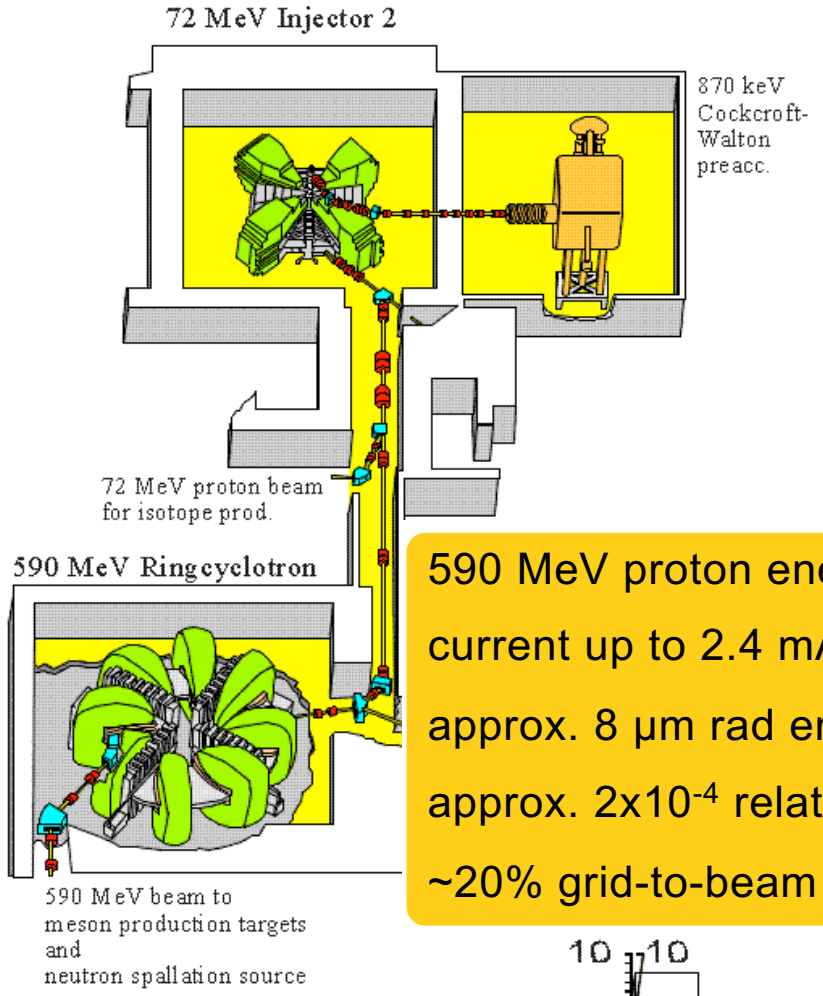
Swiss Muon Source



Swiss Light Source

High Intensity Proton Accelerator Facility (HIPA)





590 MeV proton energy

current up to 2.4 mA (1.4 MW)

approx. 8 µm rad emittance

approx. 2×10^{-4} relative losses

~20% grid-to-beam efficiency

Cyclotron World Record Holder for Beam Power

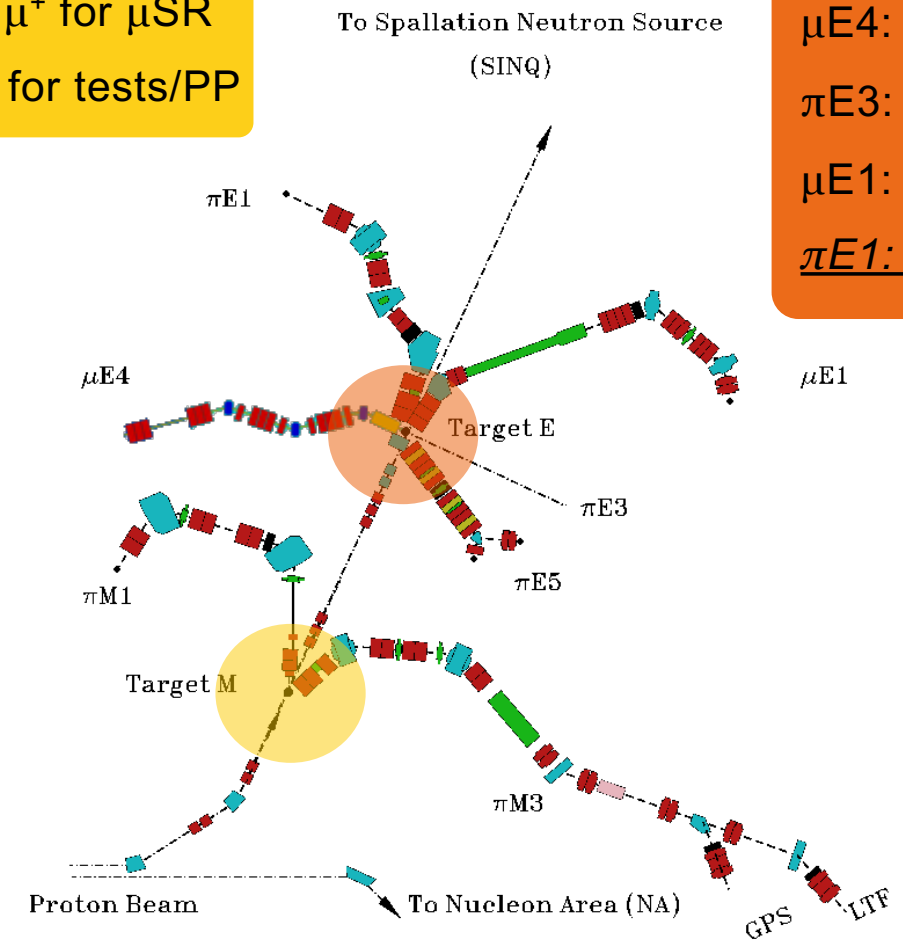
Swiss Muon Source (SμS & CHRISP)

Target M (mince) – 5mm graphite

designed for π production (low rate)

πM3: 10-40 MeV/c (surface) μ^+ for μSR

πM1: 10-300 MeV/c $\pi \rightarrow \mu/e$ for tests/PP



Target E (epaisse) – 40mm graphite

designed for π/μ production (high rate)

πE5: 20-120 MeV/c high rate μ for PP

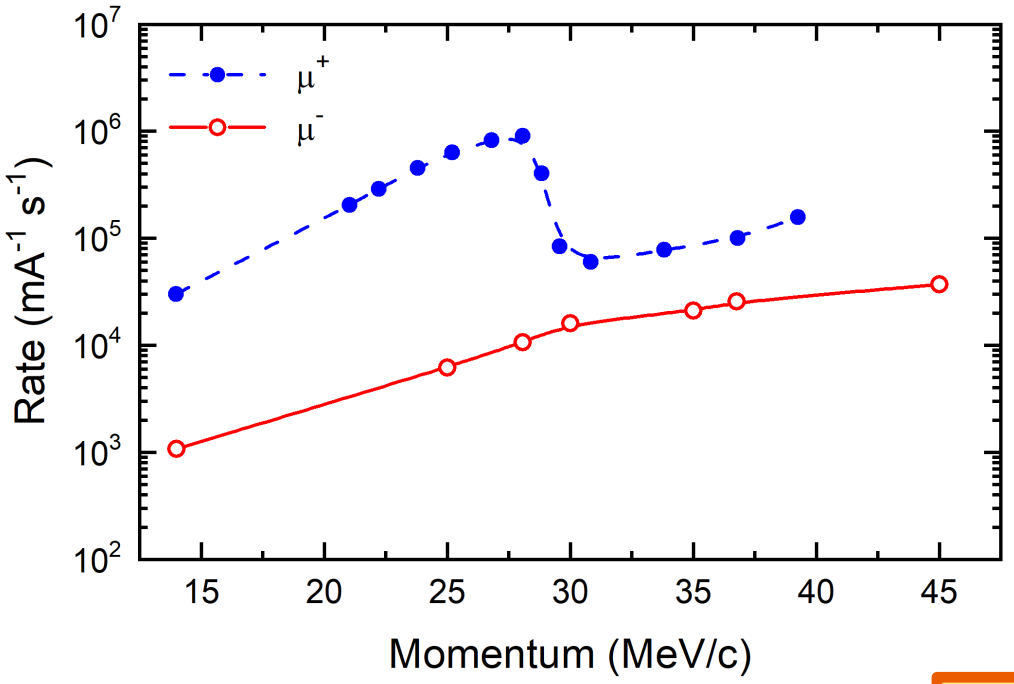
μE4: 10-40 MeV/c μ^+ for LEM – μSR and PP

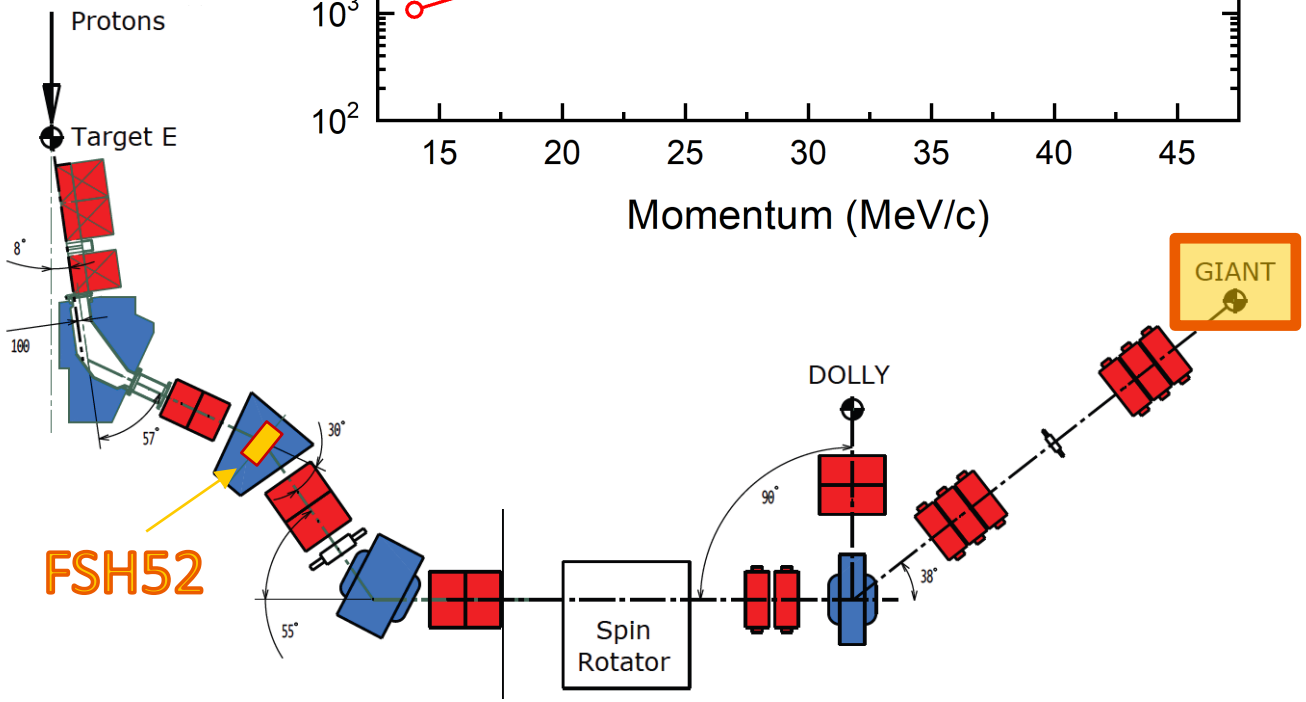
πE3: 10-40 MeV/c (surface) μ^+ for bulk μSR

μE1: 60-120 MeV/c μ for μSR *(and MIXE?)*

πE1: 10-120 MeV/c μ for μSR, MIXE, PP

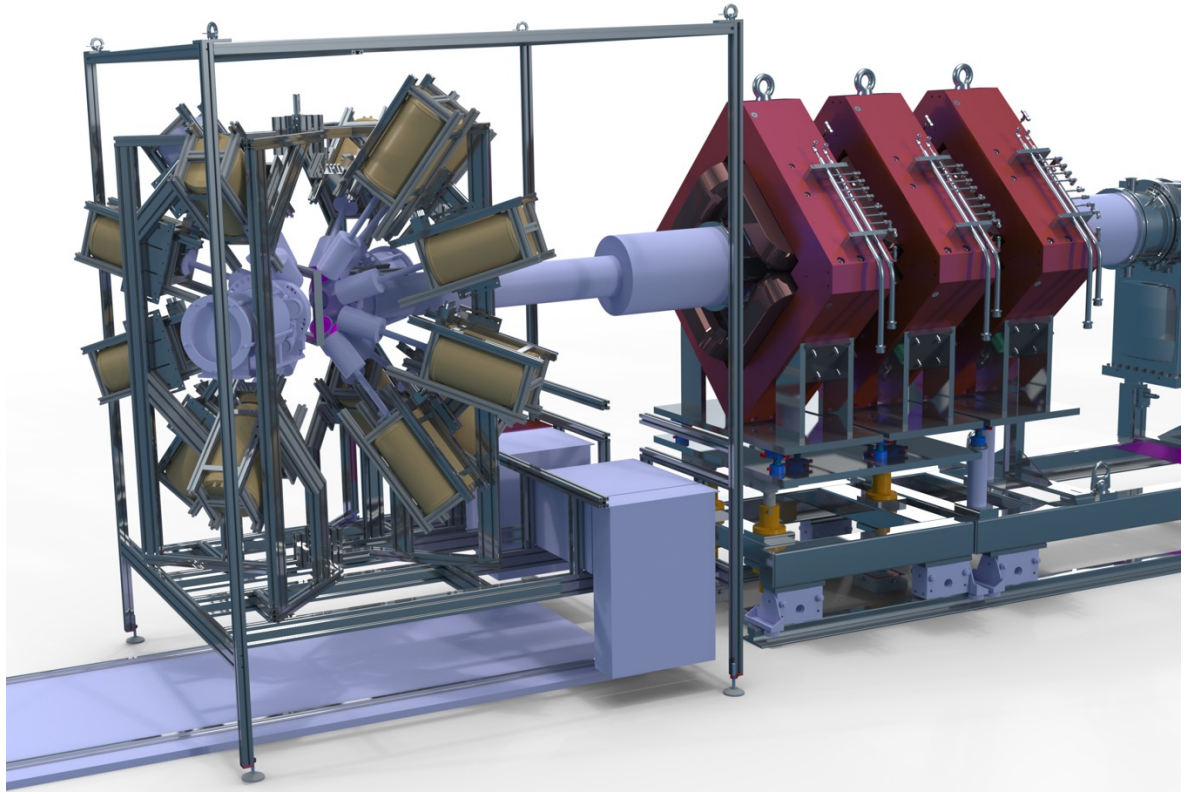
MIXE at beamline πE1

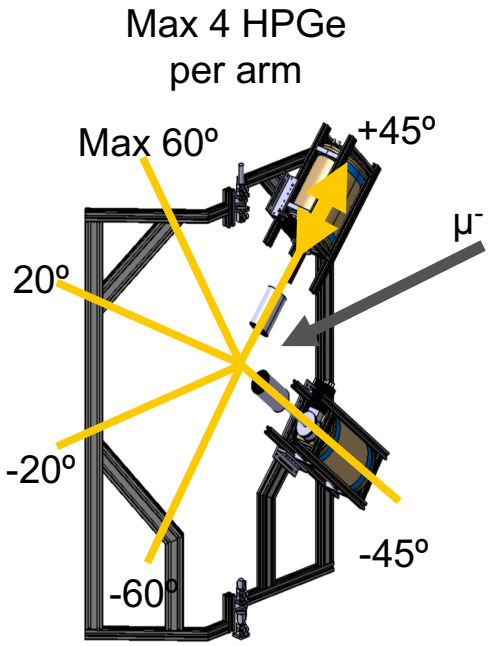


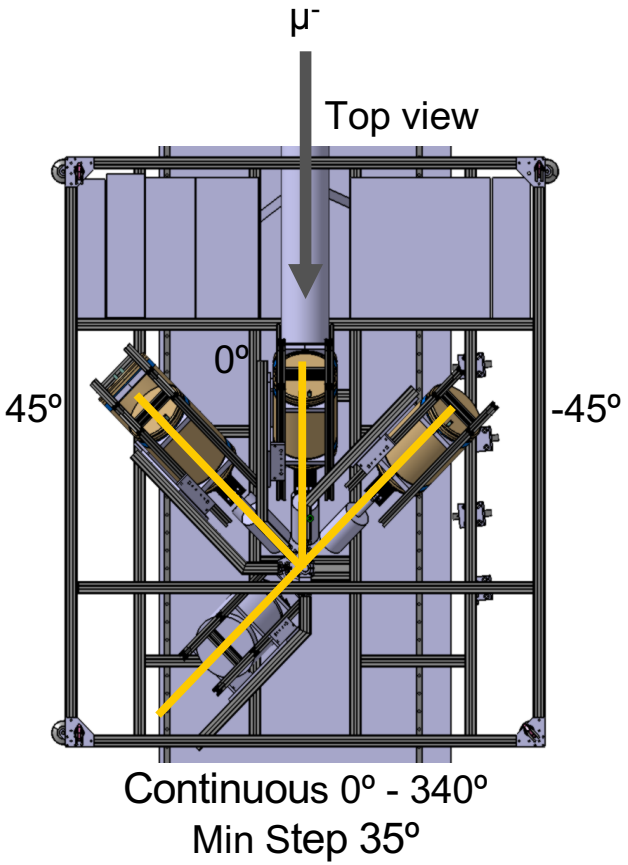


- High intensity π beamline at target E
  - cloud μ^- from π^- decay
- All past MIXE campaigns hosted at πE1.2
  - non-permanent installation
  - approx. 3 weeks beam time per year
- Momentum acceptance
  - selectable via FSH52 slit pair
  - $\Delta p/p \approx 1 - 8$ % FWHM
- Typical μ^- energies: 15-50 MeV/c
- *Rates from 10^3 up to 10^5 μ^-/s on target*
  - close to ideal sampling rate (CW beam!)
  - For the "average" sample, we collect enough statistics within ~1 hour

The GIANT Setup - Hardware Design



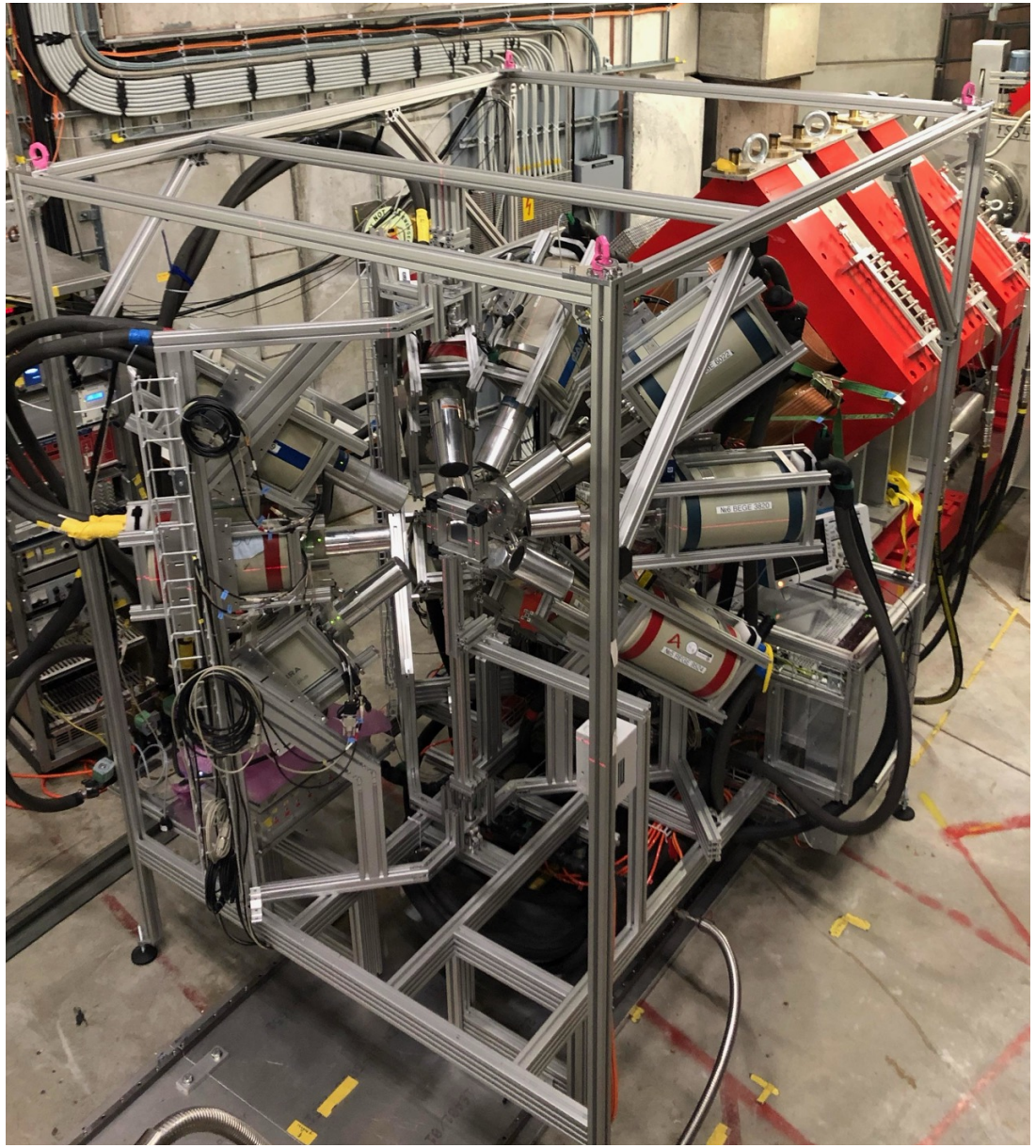




GIANT

German**I**um **A**rray for **N**on-destructive **T**esting

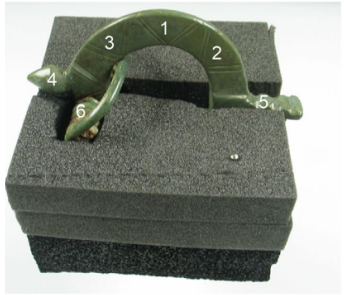
The GIANT Setup - Hardware Reality

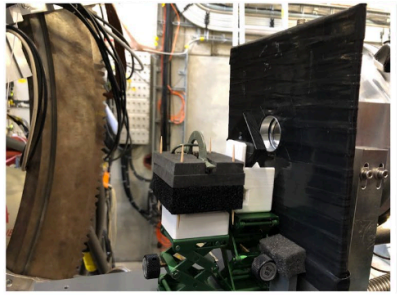


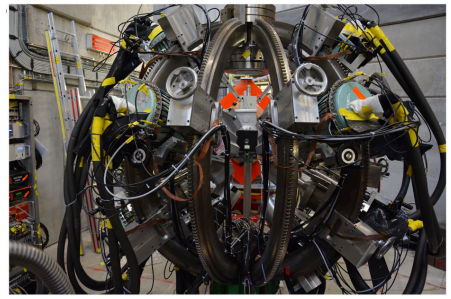
- (Nearly) complete setup
  - 8 freely rotating arms (currently 5)
  - 4 BigMac HPGe per arm
  - up to 30 HPGe detectors
    - currently ~12 detectors
  - shared between multiple experiments
- Reproducible positions and angles
- Setup time approximately 6 hours
- Supervised semi-automatic LN2 refill
- Sample station twin in control room
  - Roughly 5 min sample change
  - *represents approx. 5-10% of the time an "average" sample is measured*

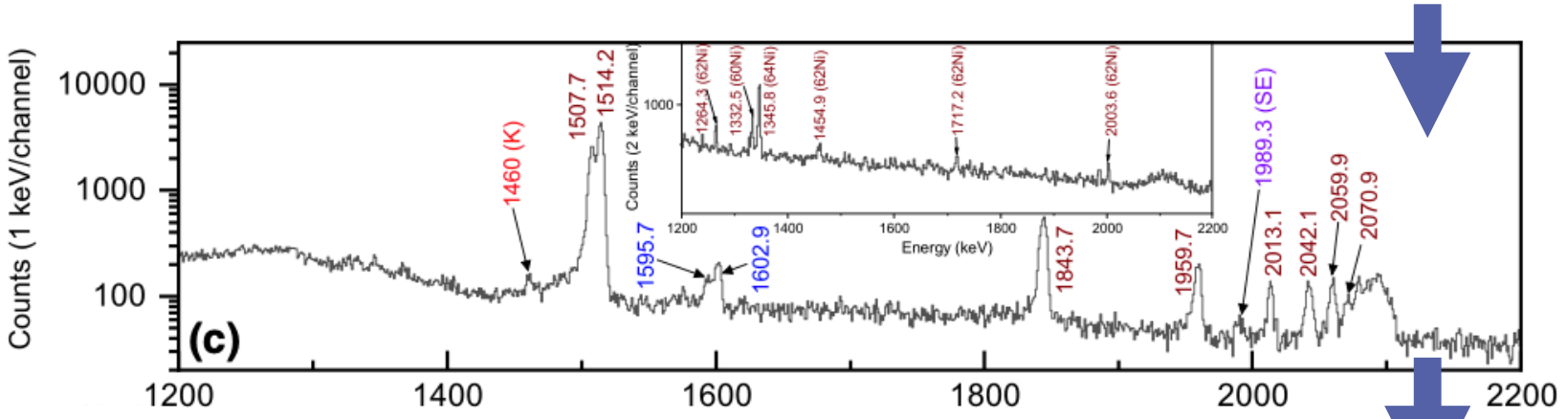
Application to non-destructive elemental analysis

- Goal: automate elemental analysis of samples scanned with Muon beams, with application to GIANT detector at PSI
- Procedure:
  - ➡Obtain muonic X-ray spectra from the GIANT detector
  - ➡Assume systematic errors model (peak spread, count noise level, background level)
  - ➡Use X-ray energies prediction from Mudirac for every element/ isotope
  - ➡Fit the spectra with a model
  - ➡Identify elements present
  - ➡Estimate their proportions inside the sample
- Recent example: analysis of antique knob bow fibula (Bügelknopffibel) from Kaiseraugst/CH by S. Biswas et







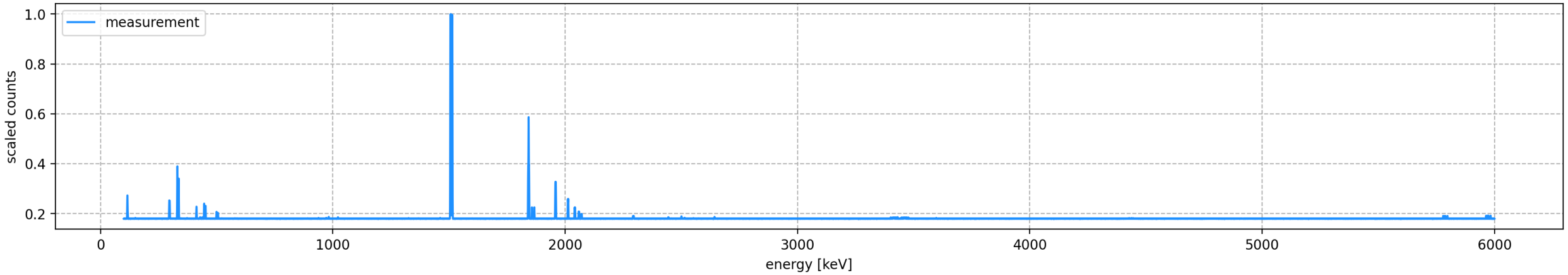


Alloy	Name	Element	wt%	$\frac{wt\%_X}{wt\%_{Cu}}$	at%	$\frac{at\%_X}{at\%_{Cu}}$	$I_{norm} \times 10^4$	$\frac{I_X}{I_{Cu}}$	$\frac{I_{X,\epsilon,cap,br}}{I_{Cu,\epsilon,cap,br}}$
Brass	31X B18 K	Cu	59.37 ± 0.05		60.41 ± 0.06		64.6 ± 0.3		
		Zn	39.41 ± 0.04	0.6638 ± 0.0008	38.97 ± 0.04	0.645 ± 0.001	43.3 ± 0.3	0.670 ± 0.005	0.64 ± 0.08
		Pb	1.018 ± 0.005	0.01715 ± 0.00008	0.318 ± 0.001	0.00526 ± 0.00002	0.30 ± 0.03	0.0047 ± 0.0004	
		others	0.270 ± 0.002		0.300 ± 0.002				
	31X B22 F	Cu	82.47 ± 0.06		82.83 ± 0.10		89.8 ± 0.4		
		Zn	15.92 ± 0.07	0.1930 ± 0.0008	15.54 ± 0.07	0.1876 ± 0.0008	17.1 ± 0.2	0.190 ± 0.003	0.18 ± 0.02
		others	1.599 ± 0.008		1.628 ± 0.008				

Biswas, S., Megatli-Niebel, I., Raselli, L. *et al., https://doi.org/10.1186/s40494-023-00880-0*

The Group-Sparse LASSO method

- We start with a simulated “measured” spectrum
- Assume some peak spread, X-ray counts (for noise model), all of this assumed known
- Background is assumed known and divided out
- We have ~20 previously studied samples from standards, batteries, and cultural artifacts
- X-ray peaks heights assumed known
- Example simulated “measured” spectrum below
- Goal: fit a model to the spectrum with a wide range of elements, and perform element selection simultaneously



Simulation of BCR-691-C Copper alloy

SDSC

Group Sparsity for Modelling of Spectra

- To identify the elements in the sample out of the periodic table, we need to solve a feature selection problem.
- This warrants the use of generalised linear problems with sparsity.
- However, the sparsity should be applied on the level of an element: if the element is selected, all its emission lines are expected to be non-zero.
- This indicates a group-sparse approach, with the simplest method for this task is GroupLASSO.

$$\min_{\beta \geq 0} \|X\beta - y\|_2^2 + \alpha \sum_{g=1}^{G} \|\beta_g\|_2$$

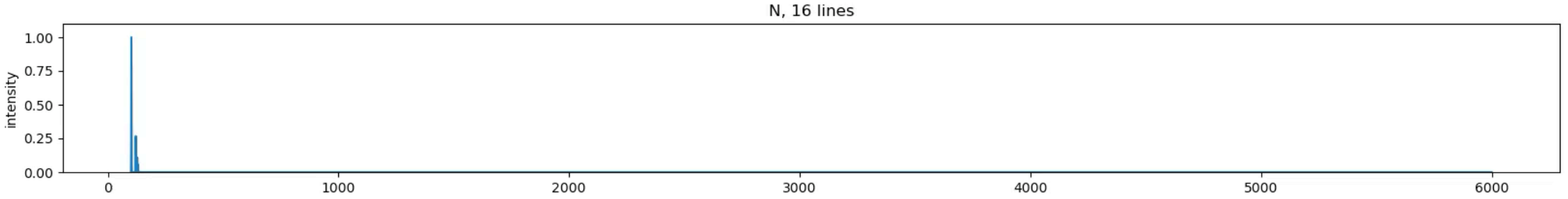
The spectra and features are binned into I bins at energies e_i. Then:

y : measured binned spectrum $y \in \mathbb{R}_+^I$, with elements $y = [y_1, \ldots y_I]^\top$, typically $I \sim 2000$.

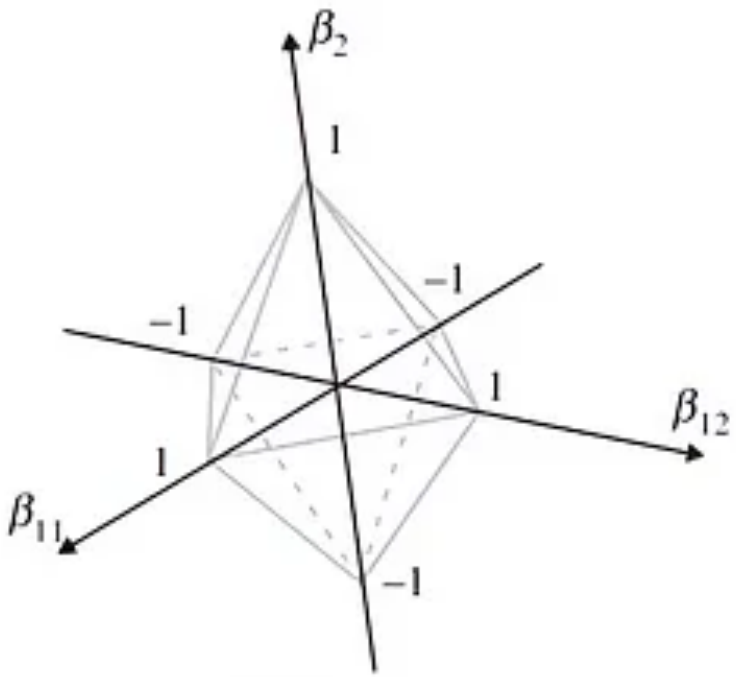
X : feature matrix created from theory energy levels $X \in \mathbb{R}_+^{N \times I}$, $X = [x_1, \ldots, x_N]$, each feature $x_n \in \mathbb{R}^I$ is a spectrum $x_n = [x_{n1}, \ldots, x_{nI}]^\top$, which is a Dirac delta at X-ray emission energy e_n, smoothed with the instrument response (currently Gaussian point spread function, assumed known) $x_{ni} = G(e_i : e_n, \sigma)$. Typically $N \sim 200k$ for a large set of elements.

β : All feature weights $\beta \in \mathbb{R}_+^N$, can be split into groups $\beta = [\beta_1^\top, \ldots, \beta_G^\top]^\top$. Each element group g contains multiple u

Full templates $X_g \beta_g$ for each element:

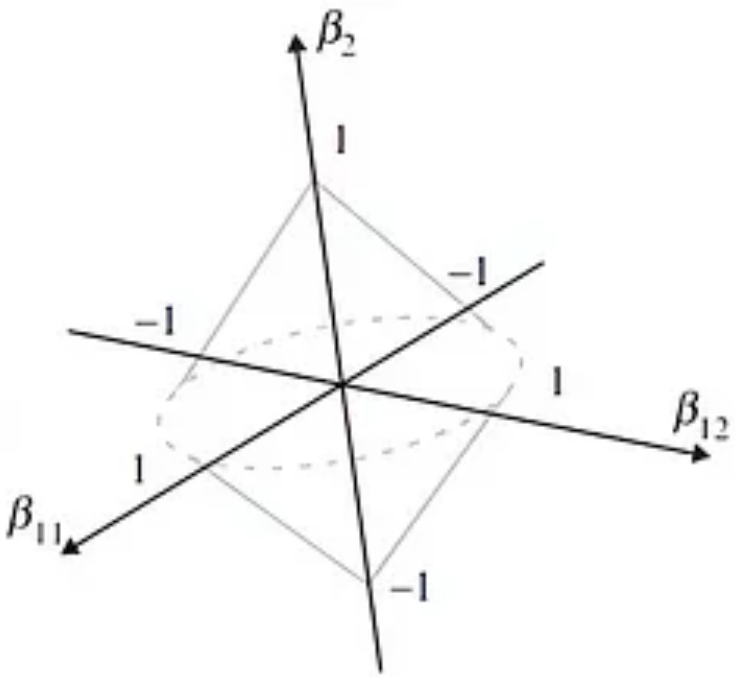


Penalties: GroupLASSO vs LASSO vs Ridge



Lasso

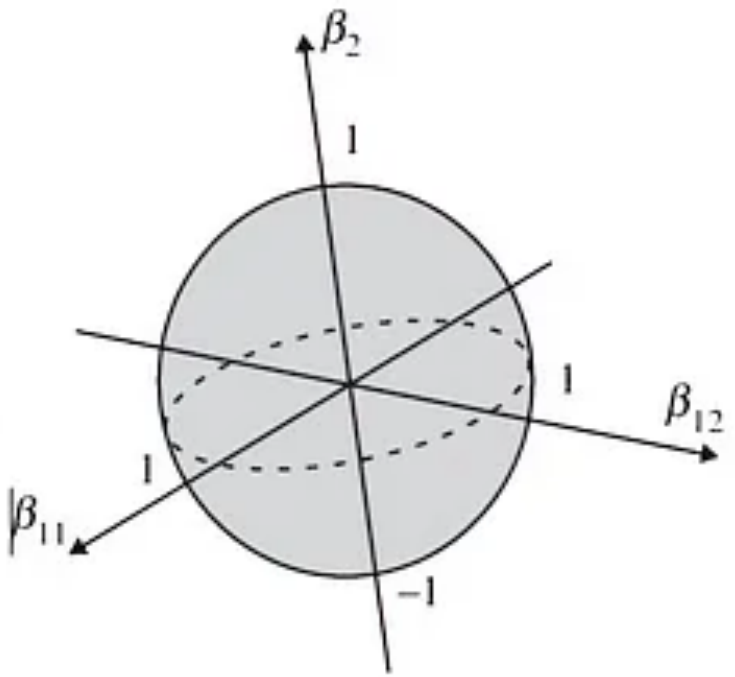
$$\min R(\beta) + \lambda \sum_{i=1}^{p} |\beta_i|$$



Group lasso

$$\min R(\beta) + \lambda \sum_{l=1}^{K} \sqrt{p_l} \left\| \beta^{(l)} \right\|_2$$

groups!



Ridge

$$\min R(\beta) + \lambda \left\| \beta \right\|_2^2$$

SDSC

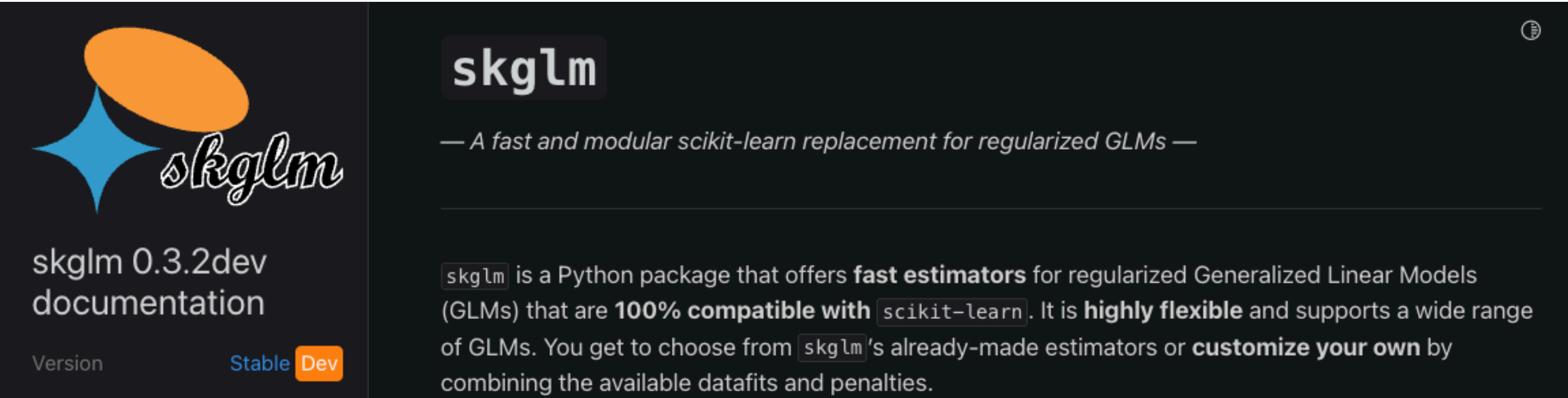
The Solver

Requirements for the GLM fast solver:

- Implemented group sparsity.
- Able to handle sparse X, as it can be big $\sim 200k \times 2k$.
- Include positive bound on feature weights $\beta \geq 0$.
- Fast optimisation with state-of-the-art techniques for sparse problems.
- Very fast results, ~ 1 second per fit

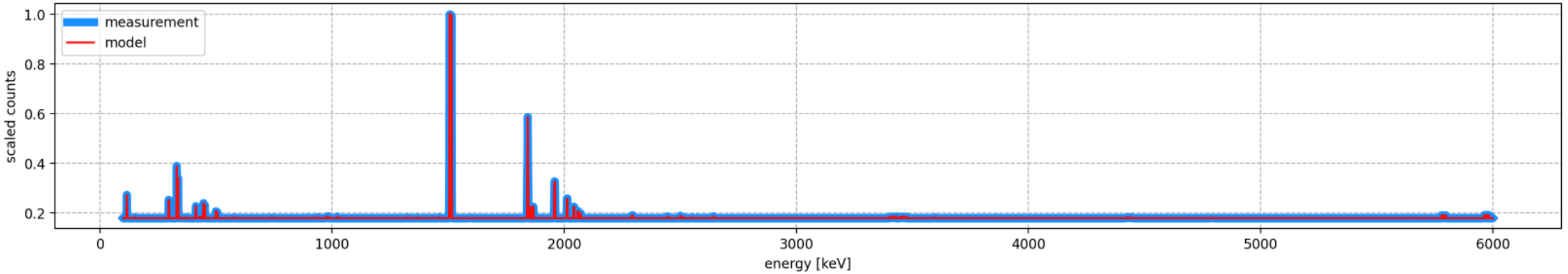
Solution: **SKGLM** scikit-learn replacement for sparse GLMs [3].

- The code is still in alpha, needed to implement additional features. Actively contributing to the development of this package.
- The optimisation engine for GroupLASSO uses block coordinate descent, and converges in ~10 sec for large sparse X matrices [3,4].
- Adding penalty for features within a group, controlled by target Pearson correlation level ρ [5]. This should give the user an option to require β coefficients inside a single element group to vary together.
- Implemented GroupLasso with sparse feature matrices in the SKGLM, merge request in progress.
- Sparsity prior may add biases to fitted weights: re-fit with linear regression using selected elements

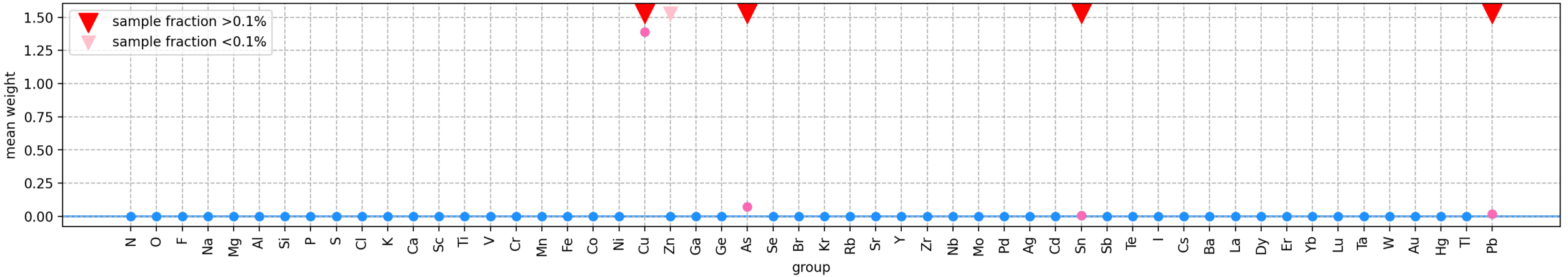


Fitting results for GroupLasso

Fit result:



Average weight per element (all lines), elements present in the sample marked with ▼



Comparison to baseline

Group-sparse solution with :

$$\min_{\beta \geq 0} \|X\beta - y\|_2^2 + \alpha \sum_{g=1}^{G} \|\beta_g\|_2$$

The spectra and features are binned into I bins at energies e_i. Then:

y : measured binned spectrum $y \in \mathbb{R}_+^I$, with elements $y = [y_1, \ldots y_I]^\top$, typically $I \sim 2000$.

X : feature matrix created from theory energy levels $X \in \mathbb{R}_+^{N \times I}$, $X = [x_1, \ldots, x_N]$, each feature $x_n \in \mathbb{R}^I$ is a spectrum $x_n = [x_{n1}, \ldots, x_{nI}]^\top$, which is a Dirac delta at X-ray emission energy e_n, smoothed with the instrument response (currently Gaussian point spread function, assumed known) $x_{ni} = G(e_i : e_n, \sigma)$. Typically $N \sim 200k$ for a large set of elements.

β : All feature weights $\beta \in \mathbb{R}_+^N$, can be split into groups $\beta = [\beta_1^\top, \ldots, \beta_G^\top]^\top$. Each element group g contains multiple lines.

Baseline: single template per element

$$\min_{\beta \geq 0} \|X\beta - y\|_2^2$$

The spectra and features are binned into I bins at energies e_i. Then:

y : measured binned spectrum $y \in \mathbb{R}_+^I$, with elements $y = [y_1, \ldots y_I]^\top$, typically $I \sim 2000$.

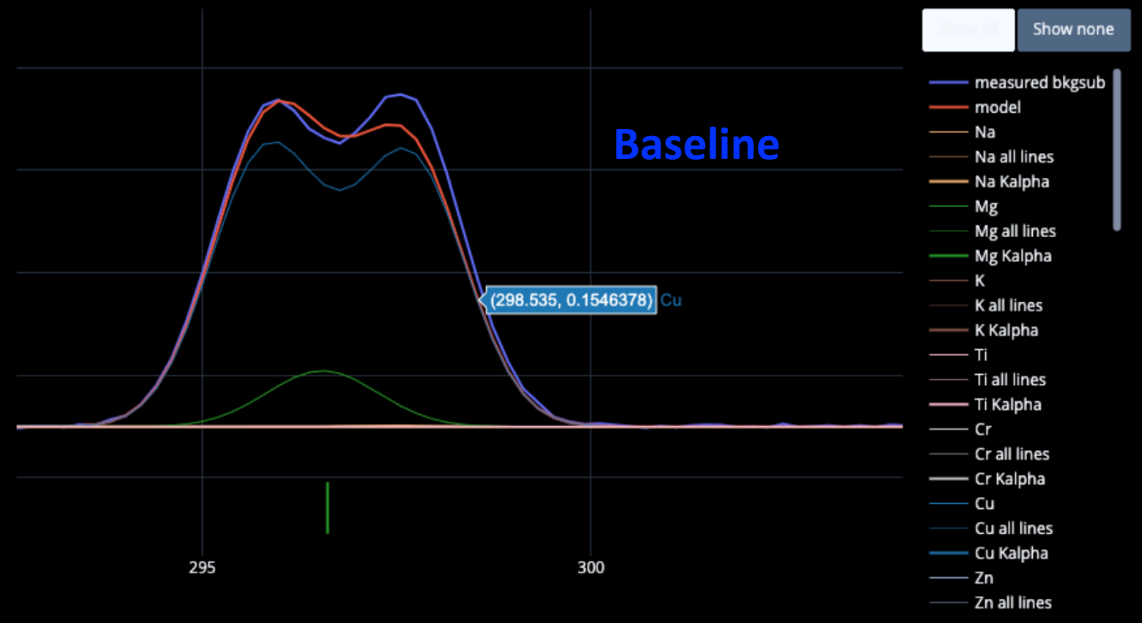
X : feature matrix created from sum of all theory energy levels $X \in \mathbb{R}_+^{G \times I}$, $X = [x_1, \ldots, x_G]$, templates $x_g \in \mathbb{R}^I$ are a spectra $x_g = \sum_{n \in N_g} r_n \cdot x_n^\top$, each of which is a sum of Dirac deltas at X-ray emission energy e_n for a given element g, multiplied by mudirac rate r_n, smoothed with the instrument response (currently Gaussian point spread function, assumed known) $x_n = G(e_i : e_n, \sigma)$. Typically $N \sim 50$ for a large set of elements.

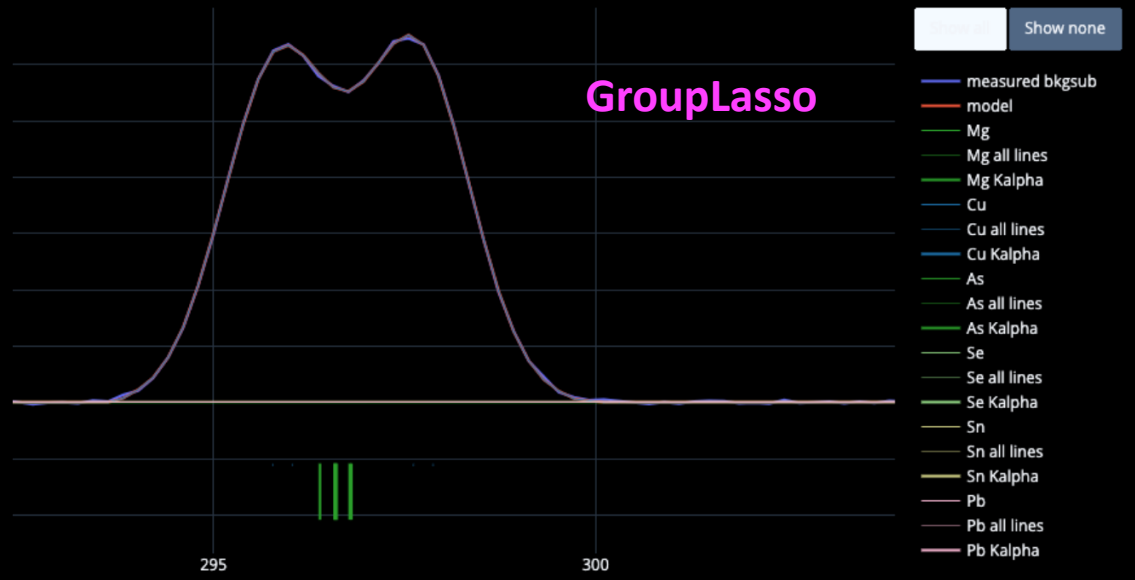
β : All feature weights $\beta \in \mathbb{R}_+^G$

SDSC

Avoiding line confusion

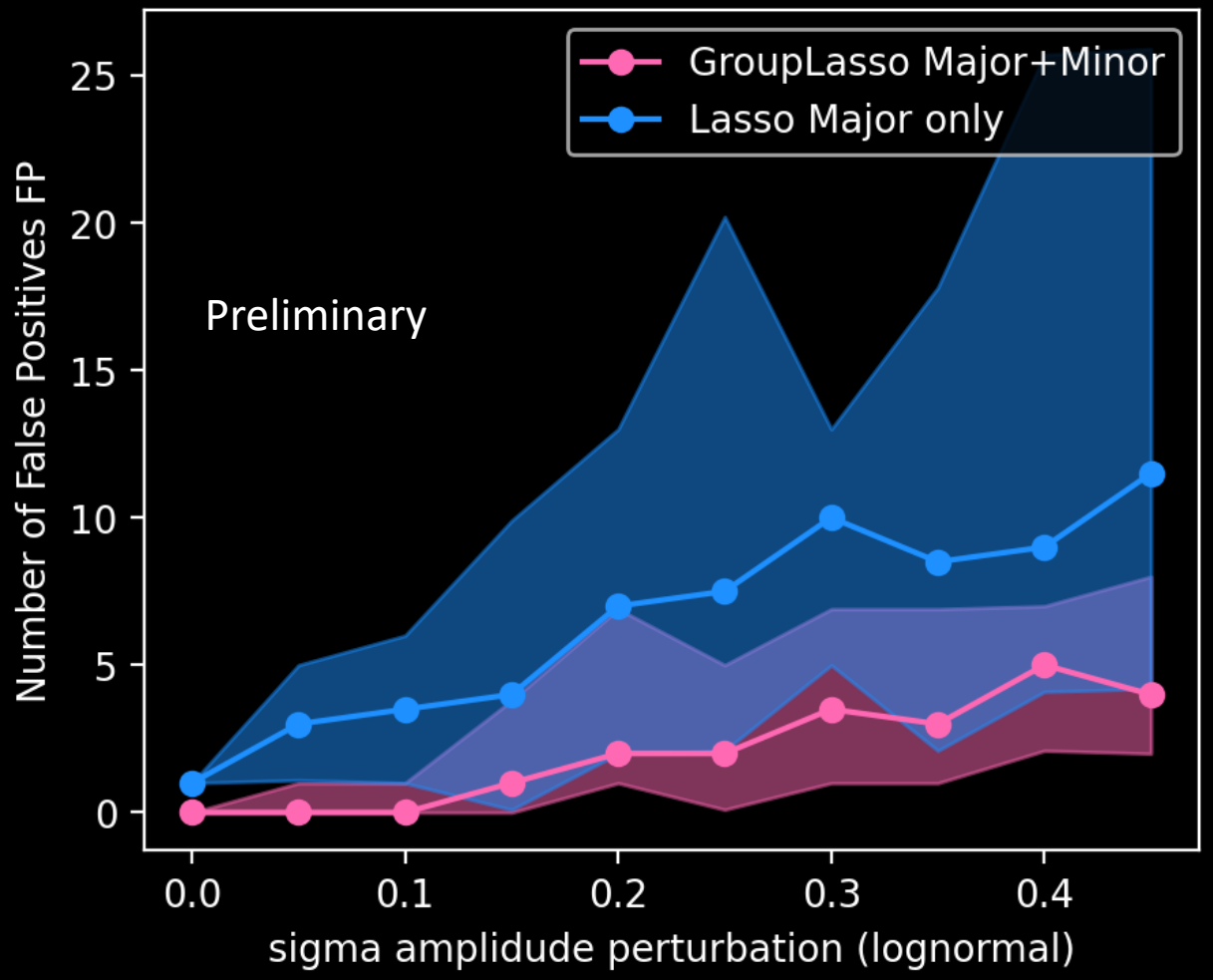
- Problem: when fitting with a large range of elements, there may be high lines (Kα) from trace elements overlapping with low lines (not Kα), from a major abundant element
- Example: Cu alloy sample non-Kα line overlapps with Kα line from Mg, which is not present in the sample, which leads to false positive Mg detection during fitting with baseline method
- Group-sparse LASSO solution avoids this confusion





How much uncertainty on the “rates” can we tolerate?

- Problem: the “rates” for each transition are currently not well known, which can affect the analysis
- Experiment: create a spectrum simulation for a sample using the current rates, and perturb them with some scatter
- Currently using lognormal multiplicative noise model: $\tilde{x} = x \cdot s,\ s \sim \mathrm{lognorm}(0,\sigma)$
- Run 32 random realizations and for each perform a fit with a ~ 57 elements
- Measure the number of false positive element detections as a function of σ
- The true positive rate is usually 100%
- We can currently tolerate up to around 10% of uncertainty on the rates, but results still stable for 20%, still acceptable up to 50%!



Next steps

- Analyse multiple simulated samples
- Give estimates of the proportions of each element in the sample, with quality estimate
- Test the application to rapid spectrum analysis during a live experiment
- Publish the results and code package to the community