HI JENA
Helmholtz Institute Jena

AI for MMCs

Utilizing Neuronal Networks to Enhance the Application of Metallic-Magnetic Calorimeters for X-ray Spectroscopy

M.O. Herdrich[1,2,3], T. Mattis[4], and Th. Stöhlker[1,2,3]

[1] Helmholtz Institute Jena
[2] IOQ / Friedrich-Schiller-University Jena
[3] GSI Helmholtz-Center for Heavy Ion Research, Darmstadt
[4] Hasso Plattner Institute, Potsdam

4th MODE Workshop, Valencia, 09/24/2024

I Introduction

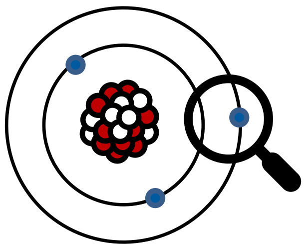
AI for **MMCs**

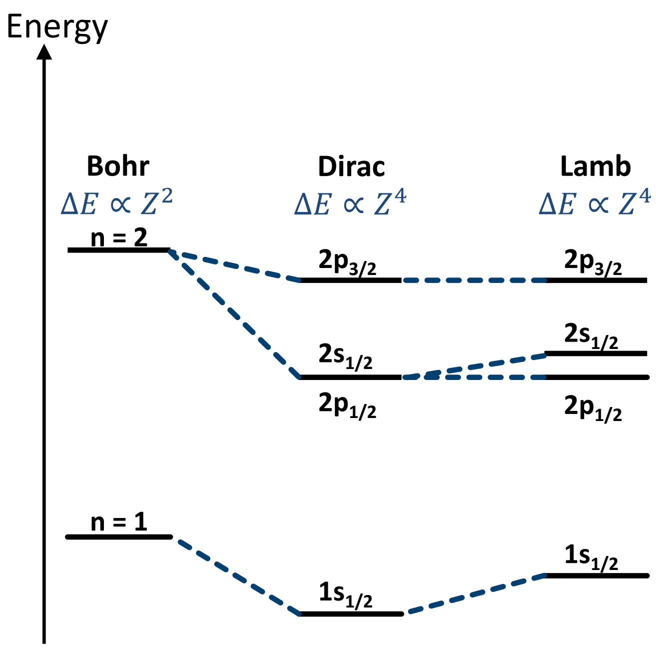
HI JENA
Helmholtz Institute Jena

Motivation: Atomic Physics

Nuclear charge Z
Fine-structure constant α

Atomic Physics





Interaction of electron with…

… nucleus
… other electrons
… **vacuum**

Quantum **E**lectro **D**ynamics (QED)

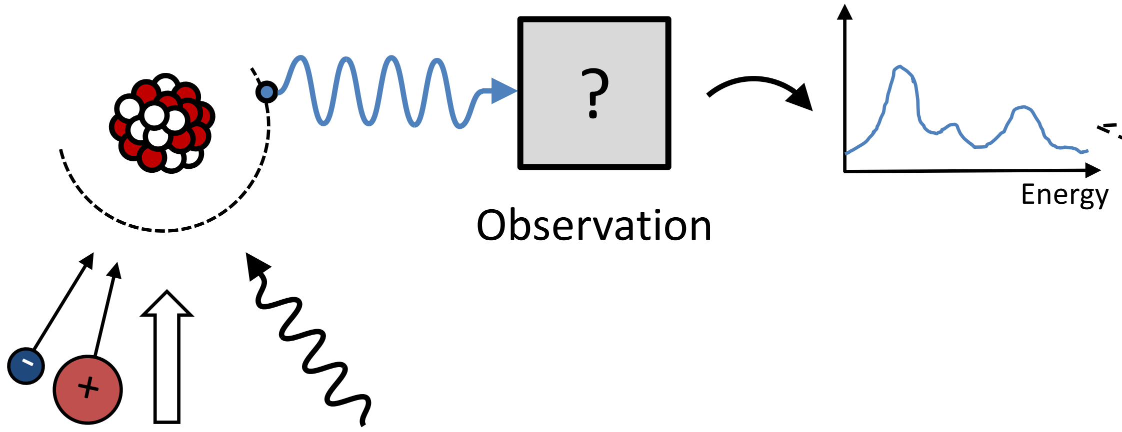
Light elements $Z\alpha \ll 1$: Only perturbative correction and most precisely measured theory

BUT: Heavy elements $Z\alpha \approx 1$: Theory is very complex (all orders) and experiments are sparse

→ Validation required in form of high precision experiments.

HI JENA
Helmholtz Institute Jena

Methods: X-Ray Spectroscopy



Excitation

Heavy ion accelerators like GSI/FAIR

Highly charged heavy ions are routinely available

Electric field strength close to Schwinger-limit

example **U^{91+}** → 1^{st} order QED ~ 100 eV
→ 2^{nd} order QED ~ 1 eV

Important: **Precision** ← Statistics **AND** Resolution

Semiconductor Detector

High quantum efficiency
Broad spectral range (keV-MeV)
BUT: Low resolution (~100 eV)

Crystal spectrometer

Very high resolution (< 1 eV)
BUT: Low efficiency and range

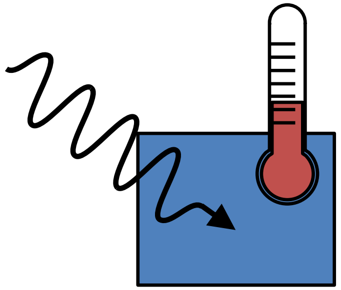
Microcalorimeters

HI JENA
Helmholtz Institute Jena

Technology: Microcalorimeters

Micro-Calorimeter Detectors

$$\Delta E \cdot C^{-1} = \Delta T$$

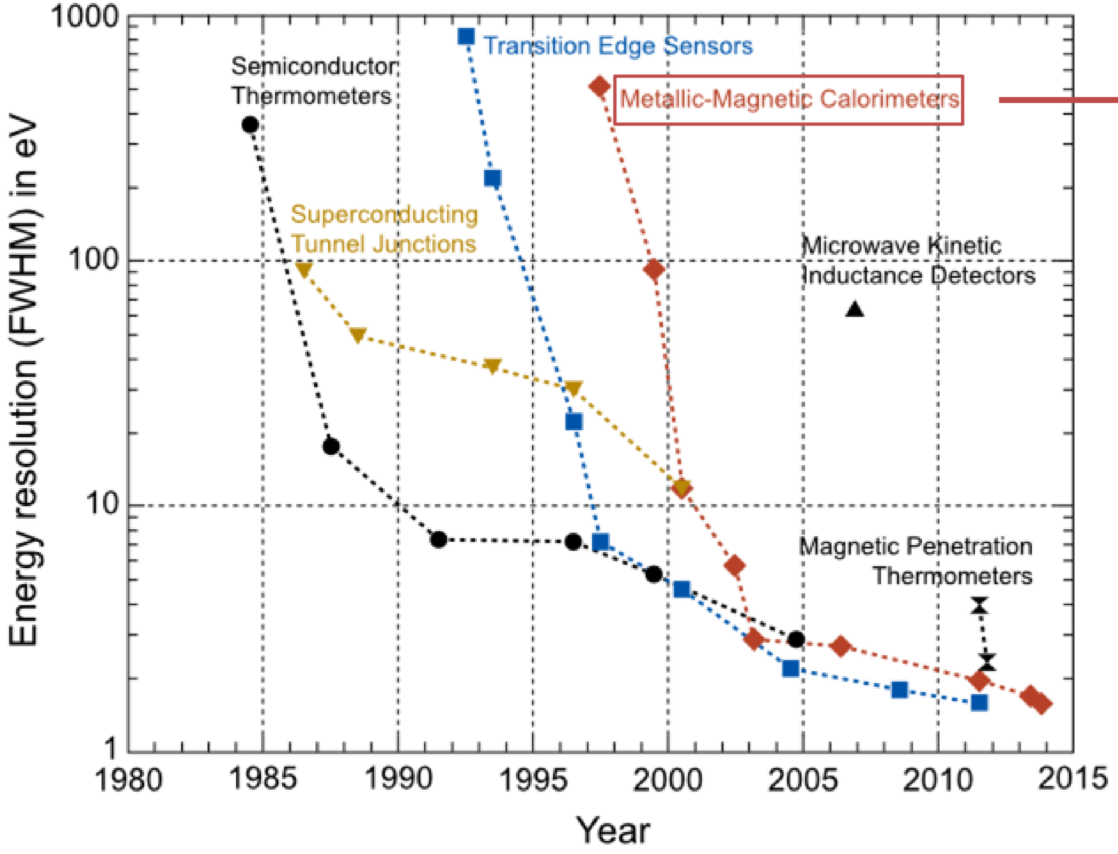


Different Technologies

Phonon excitation (~ 1 meV)
→ Inherently high energy resolution

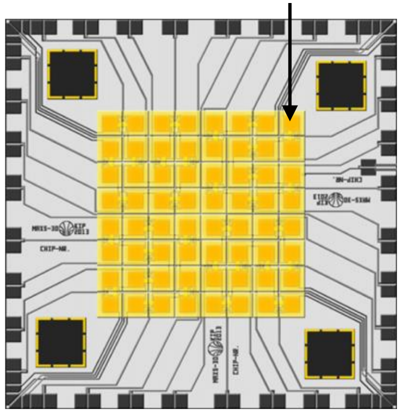
Dedicated photon absorber
→ Broad spectral range
→ High quantum efficiency

Source: S.R. Bundler et al., *Proc. SPIE* **9905**, 99050Q (2016)



MMC

500 µm x 500 µm



maXs-30

Paramagnetic Sensor

⇩

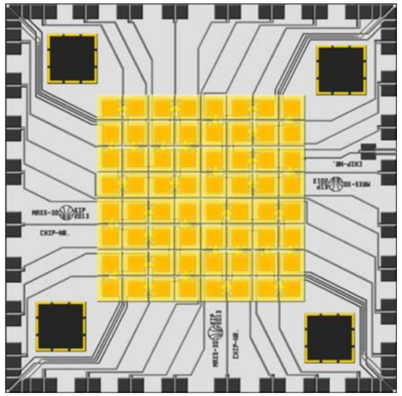
SQUID Magnetometer

HI JENA
Helmholtz Institute Jena

Challenges: How to extract a spectrum?

Fast **signal rise** time up to $\tau_0 \approx 100$ ns
High **energy resolution** $\Delta E_{FWHM} = 1.25$ eV @ 6 keV [1]
Excellent **linearity** $\Delta E / E < 5.9\%$ @ 60 keV

Typical thermal capacity $C \approx 1\,\frac{\mathrm{pJ}}{\mathrm{K}} \rightarrow \frac{\Delta T}{\Delta E} \approx 100\,\frac{\mu K}{keV}$



Operation requires very intricate hardware (cooling < 20 mK, read-out, amplification, …)

High sensitivity leads to susceptibility for external noise (vibrations, external fields, ...)

⇒ Digital pulse shape analysis required to mitigate artifacts

↪ A multitude of hardware settings and numerical parameters require **optimization**

So far: Very labor intensive, partially **manual** process

$\text{Signal} \propto \Delta T = \frac{\Delta E}{C}$ small

$C \propto V = d \cdot A$

efficiency

solid angle

pixelation

Future: Thousands of pixels ⇒ manual optimization per pixel not possible

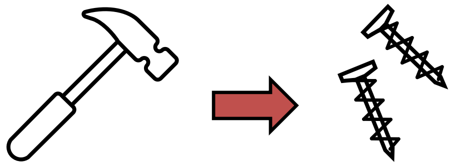
[1] F. Toschi et al., *Phys. Rev. D* **109**, 043035 (2023)

HI JENA
Helmholtz Institute Jena

II Thought Process

AI for MMCs

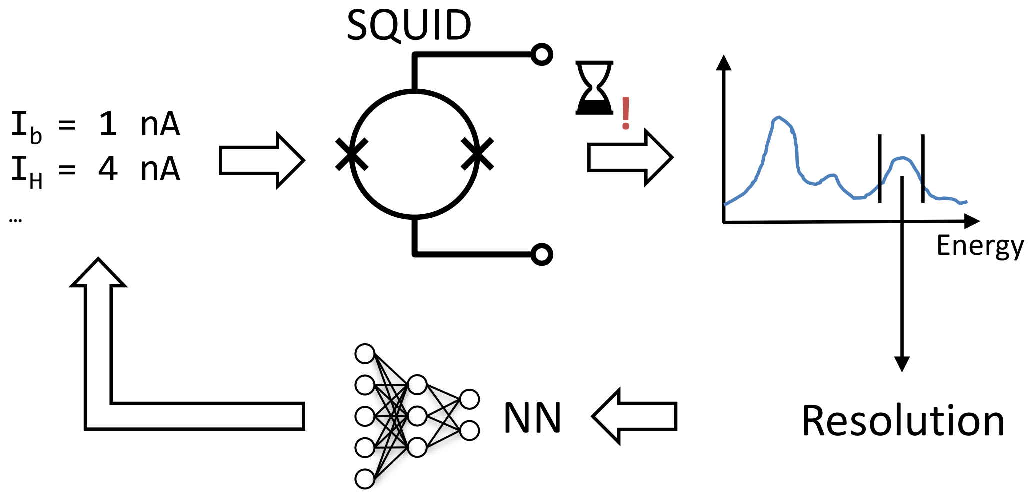
Enhance MMCs using AI?



Where to start?

First idea: Hardware

Amplification/Read-out SQUID tuning
Supervised learning cycle: SQUID parameters vs. performance
+ Could benefit more groups than just us
− Long learning cycle



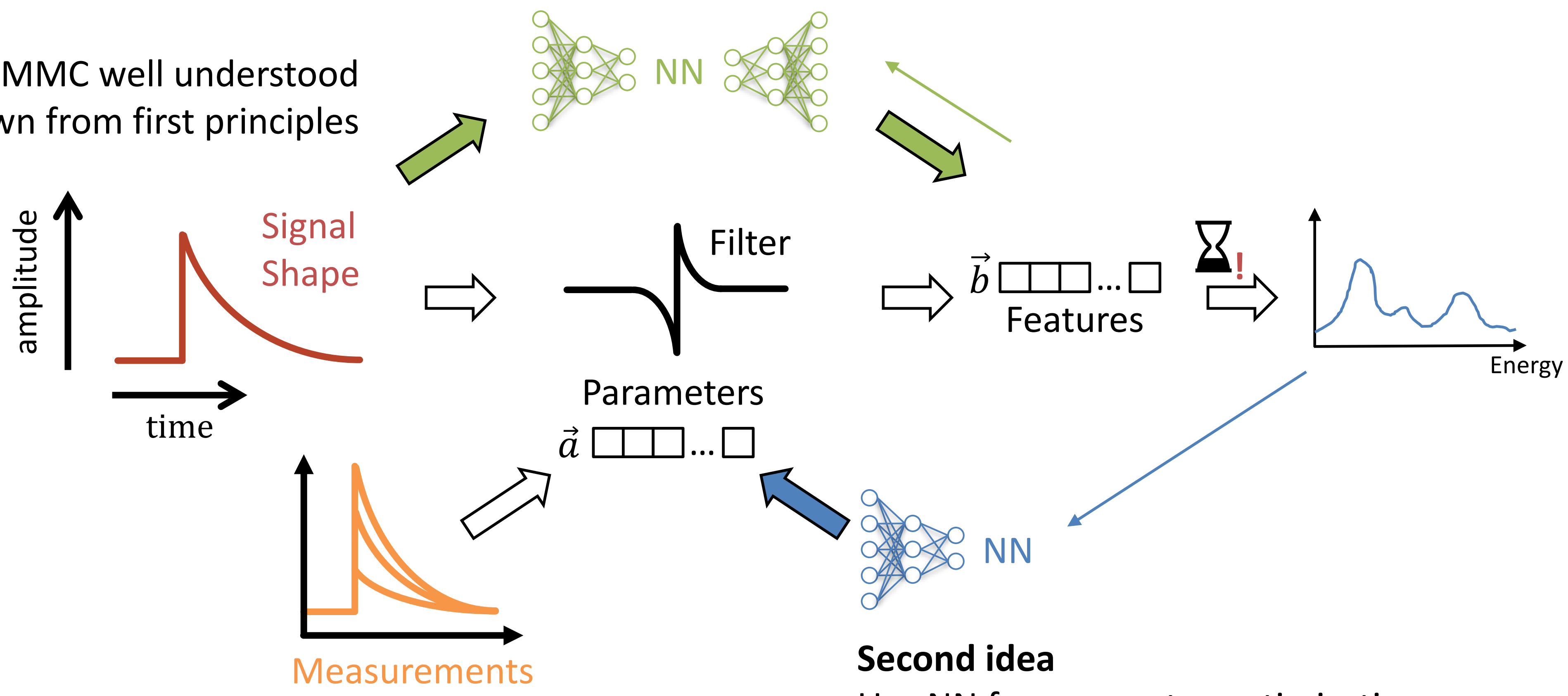
Software

Feature extraction from raw detector signals
- Signal characterization
- Pulse shape analysis
- Artifact correction

Traditional Pulse Shape Analysis

Winning idea: Use NN for pulse shape analysis itself

Thermodynamics of MMC well understood
⇒ Signal is known from first principles



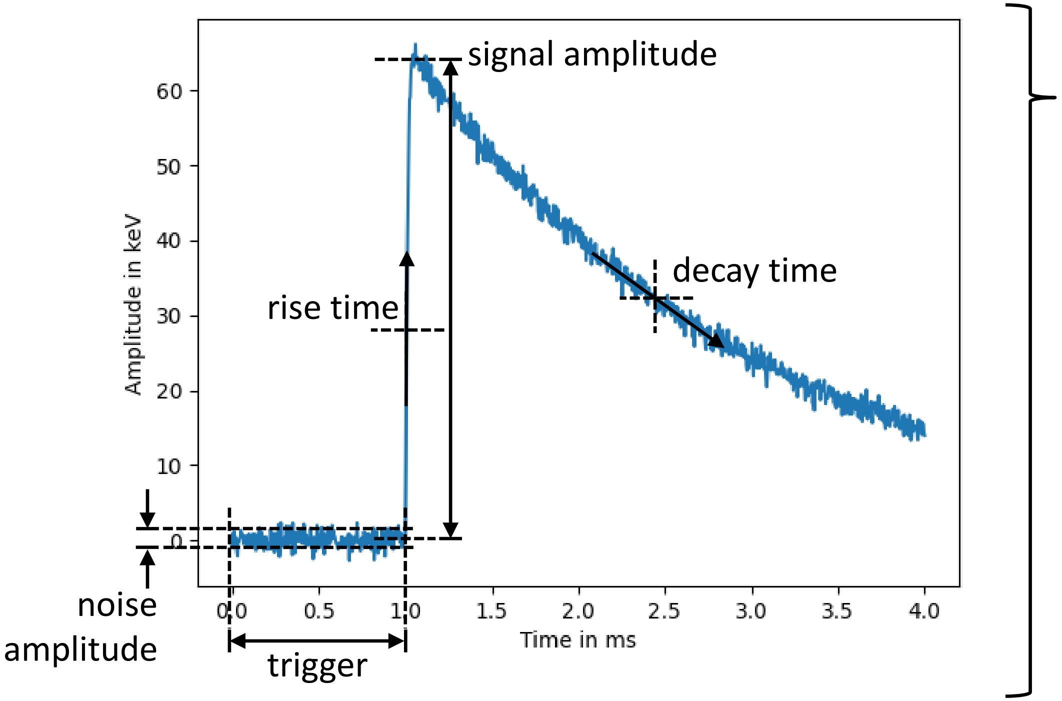
Second idea
Use NN for parameter optimization
Same challenge as for SQUID tuning

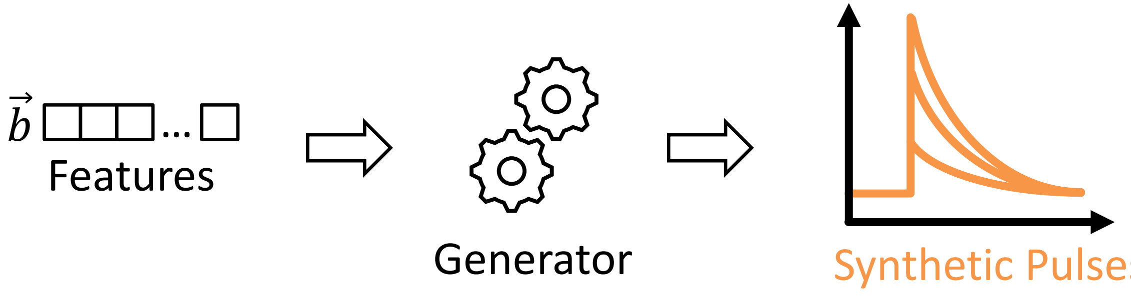
HI JENA
Helmholtz Institute Jena

The Universal Detector

How to label data that you want to find labels for?

Reverse the analysis process!
Utilize, that we know analytically, how our signals should look like.





⇒ Postpone the problem of using real-world data for training.
Evaluate the feasibility of a NN-based approach fast.
Lets us generate …

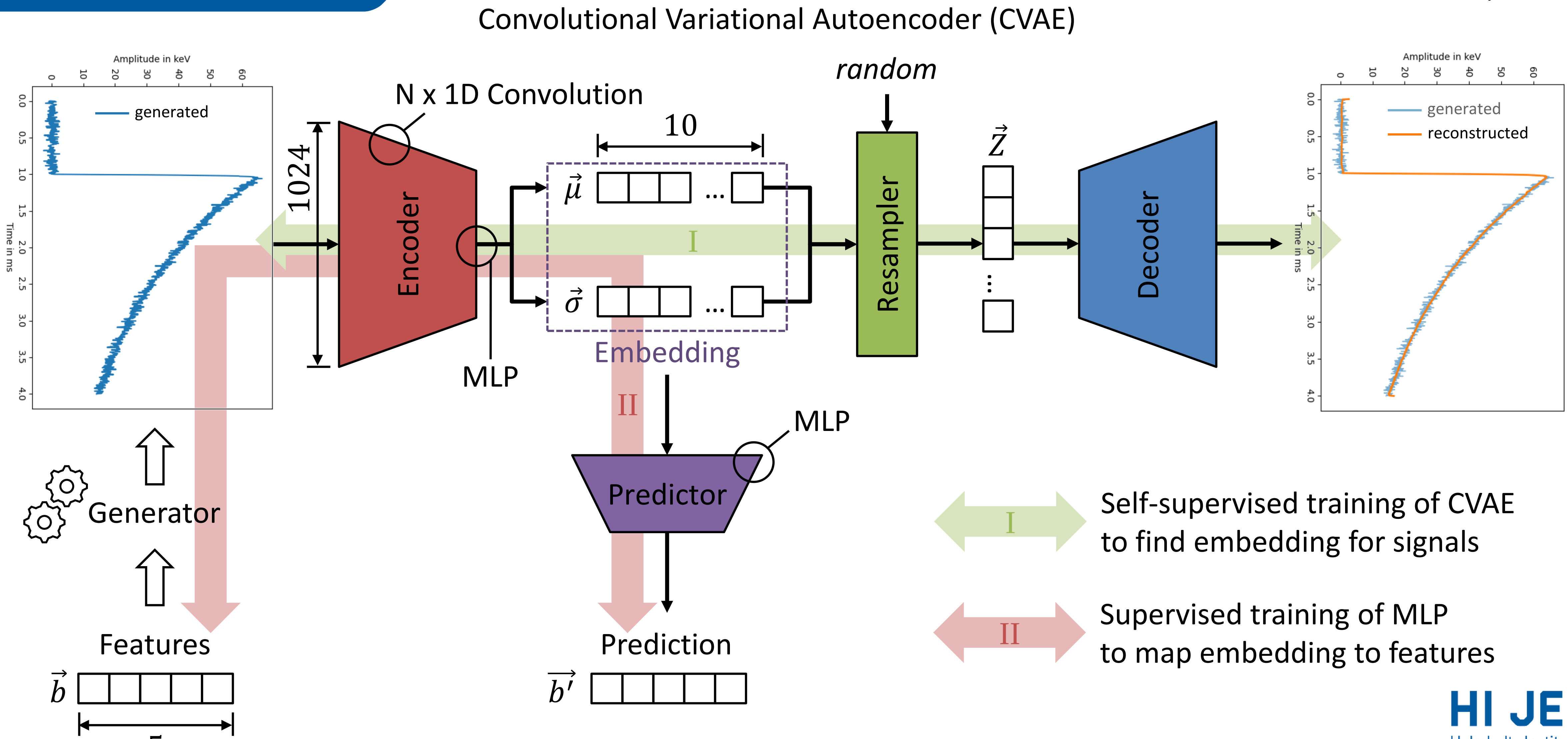
- **well-defined** pulses …
- for **arbitrary** MMCs …
- and as **many** as we need.

III Implementation

AI for MMCs

Improvement: Autoencoder

MLP = Multi-Layer Perceptron



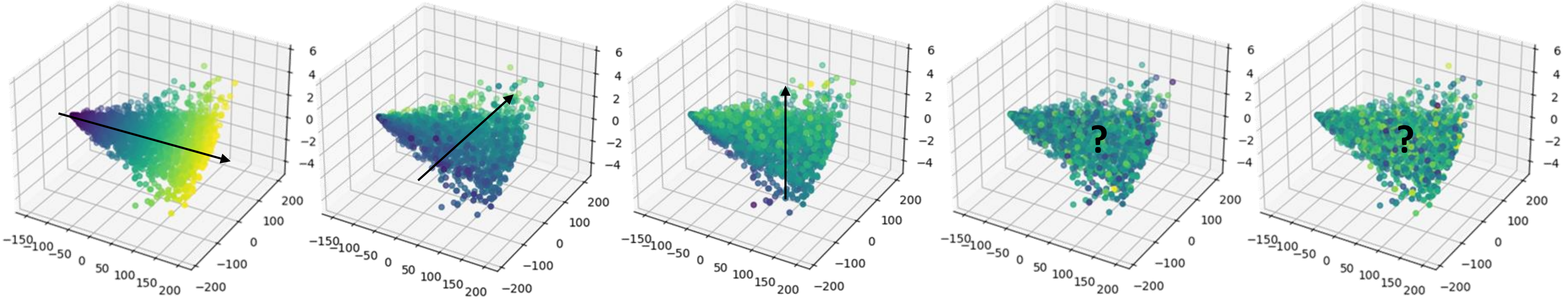
I — Self-supervised training of CVAE to find embedding for signals

II — Supervised training of MLP to map embedding to features

HI JENA
Helmholtz Institute Jena

Latent Space

Principle Component Analysis (PCA) of the embedded signals with n = 3:



Color: Energy | Trigger | Decay Time | Noise Ampl. | Rise Time

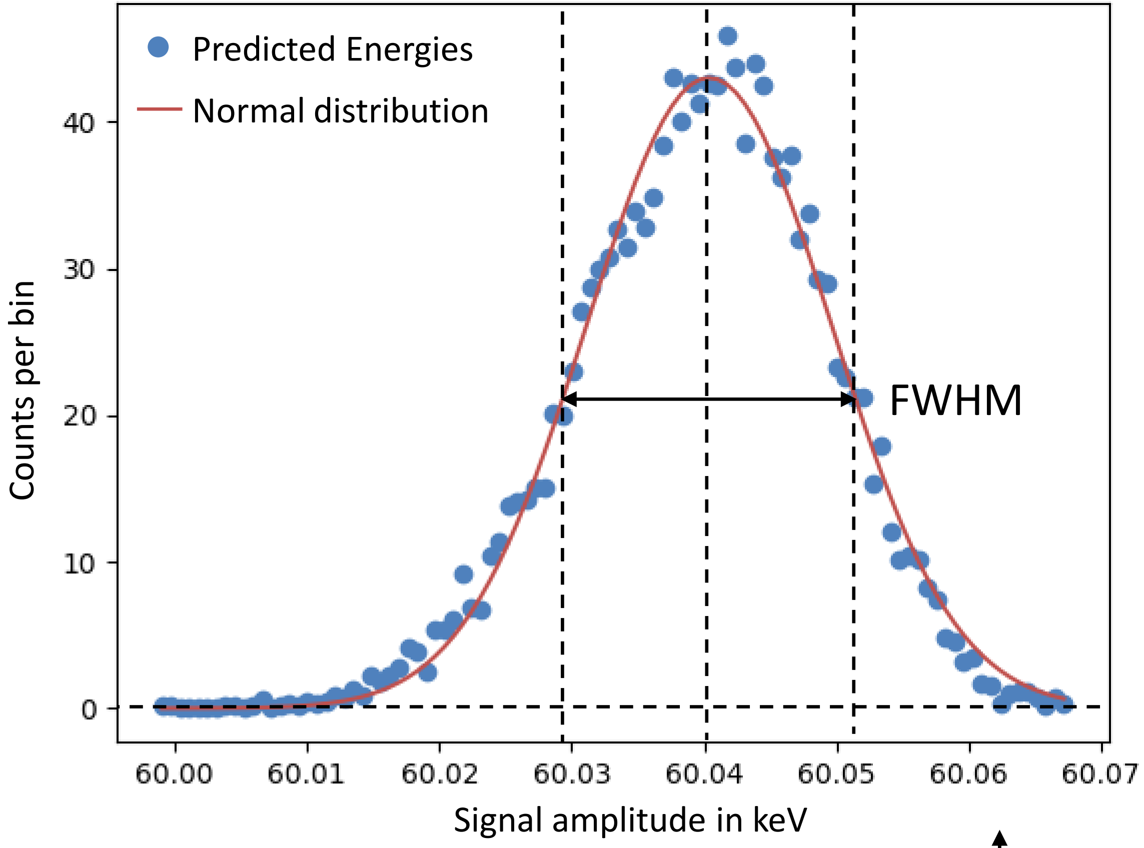
Energy → Largest variance

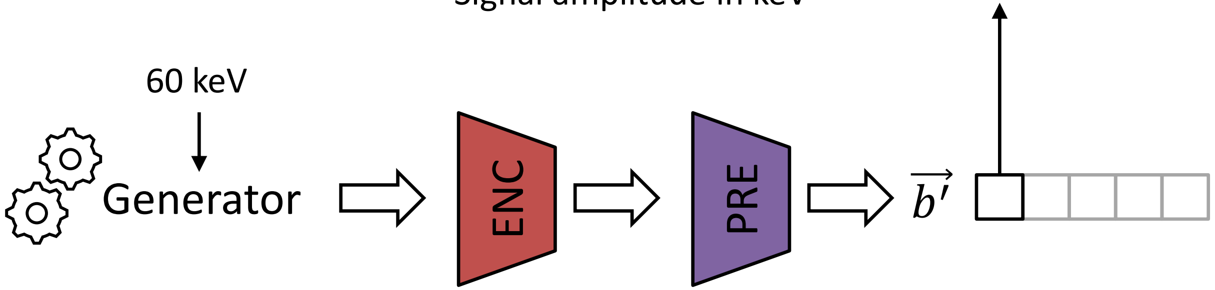
Trigger, Decay Time → Changes shape the most
(also kills FIR performance)

Noise Ampl. → Encoder acts as denoising step

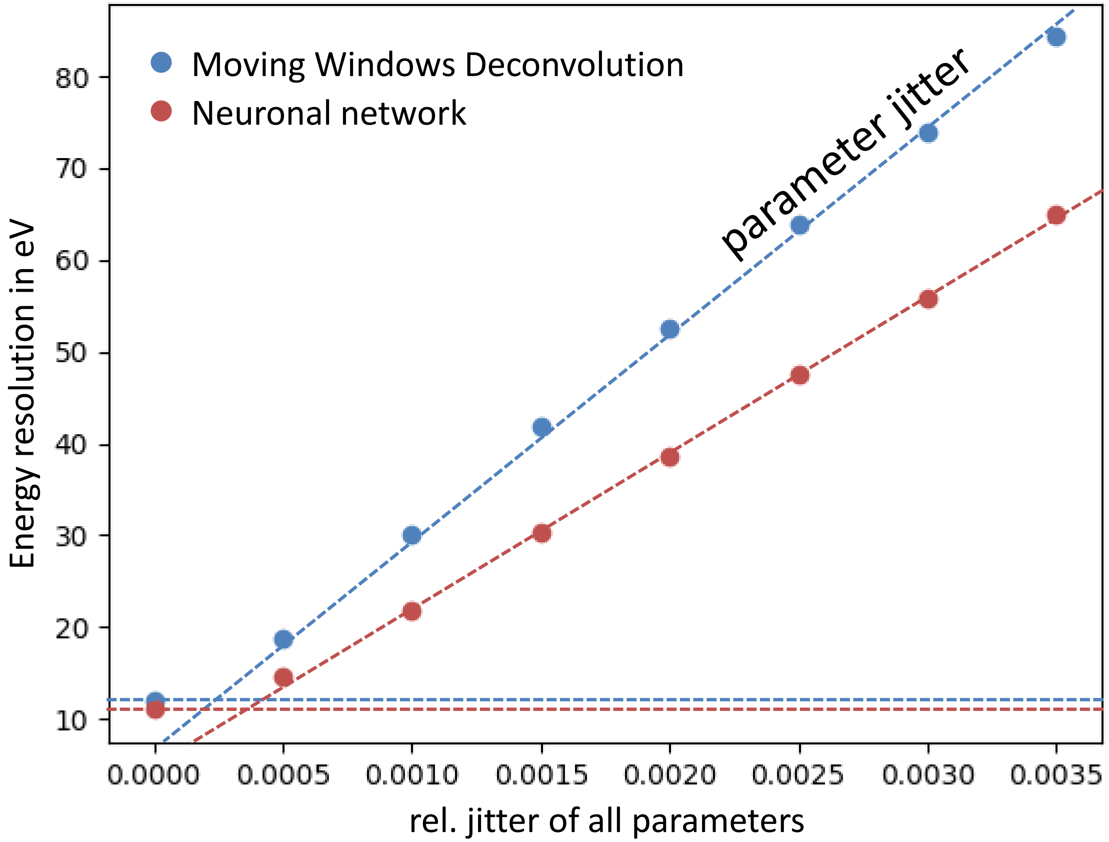
HI JENA
Helmholtz Institute Jena

Results I: Synthetic Data





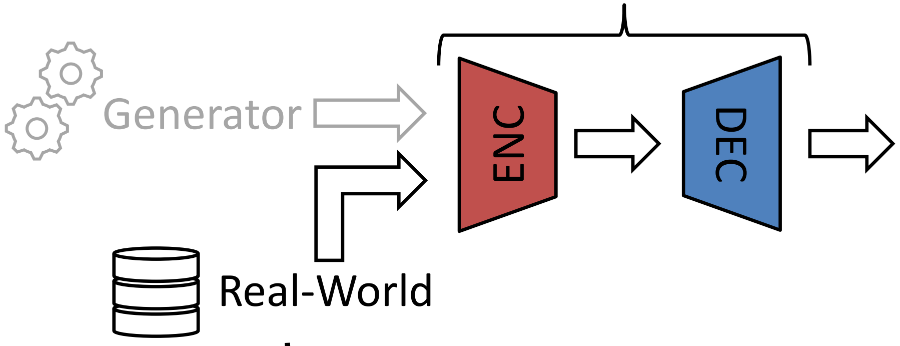
Neuronal network beats traditional filter!



HI JENA
Helmholtz Institute Jena

Real World Data

trained on synthetic data



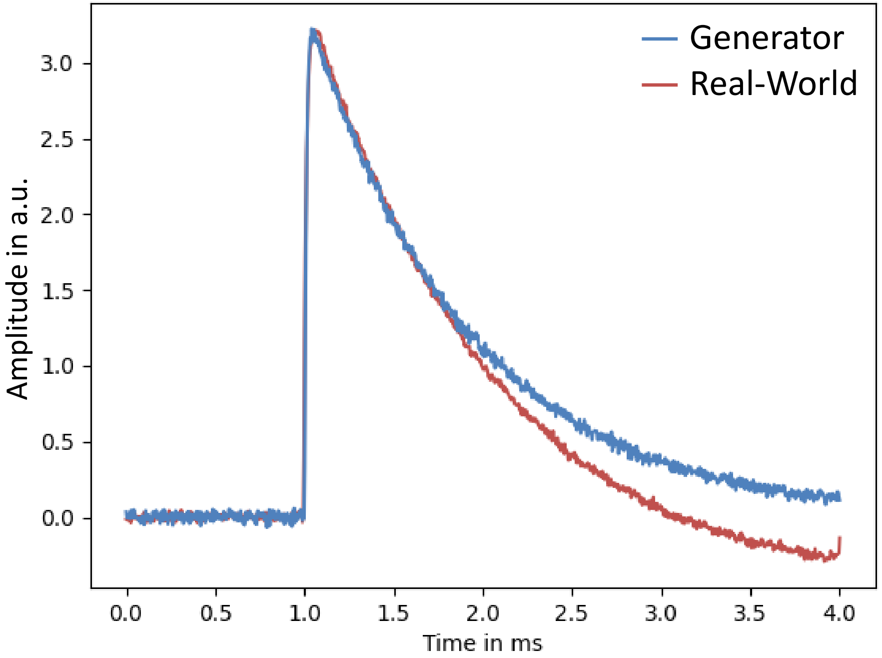
Reconstruction

Mean Squared Error (MSE),
‰ rel. to random signal

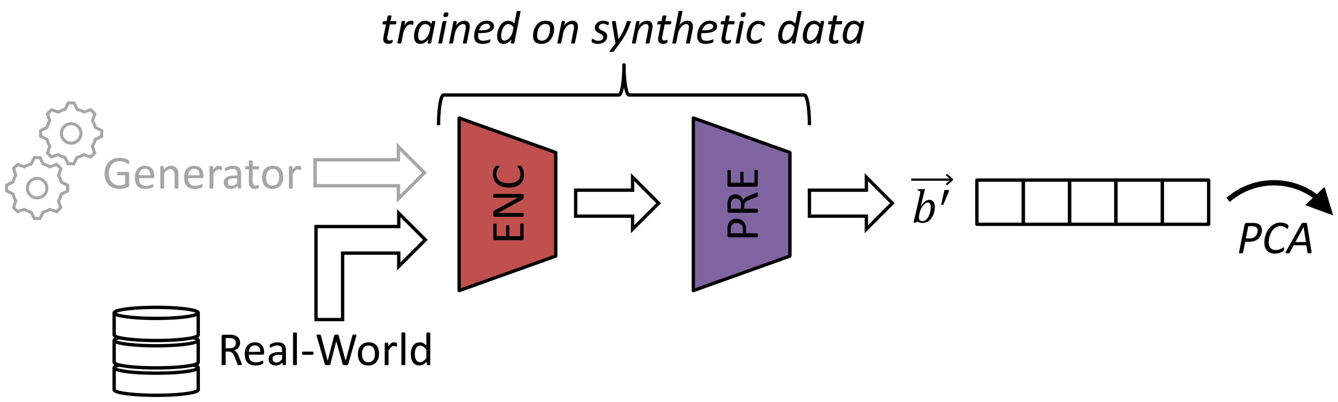
Synthetic	3.94 ± 0.03
Real-World	2.02 ± 0.01

How?

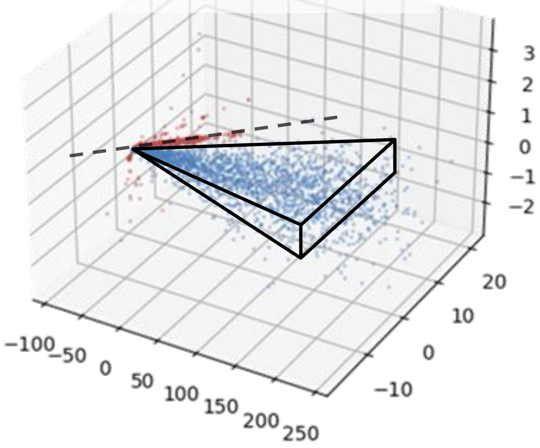
Match parameters as close as possible



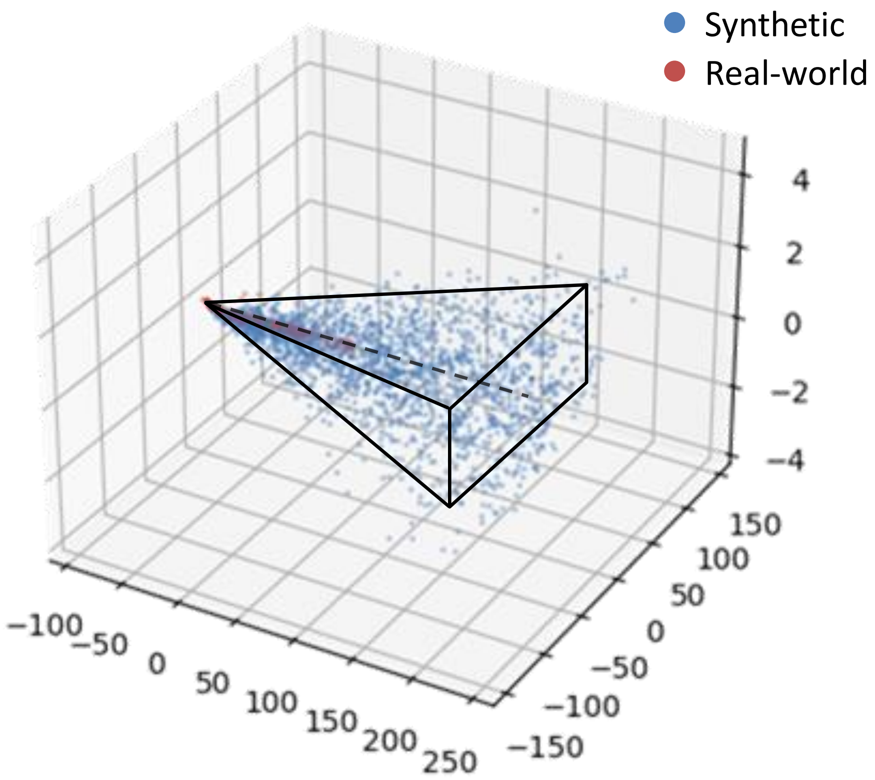
Real World Data Embedding



no trigger jitter

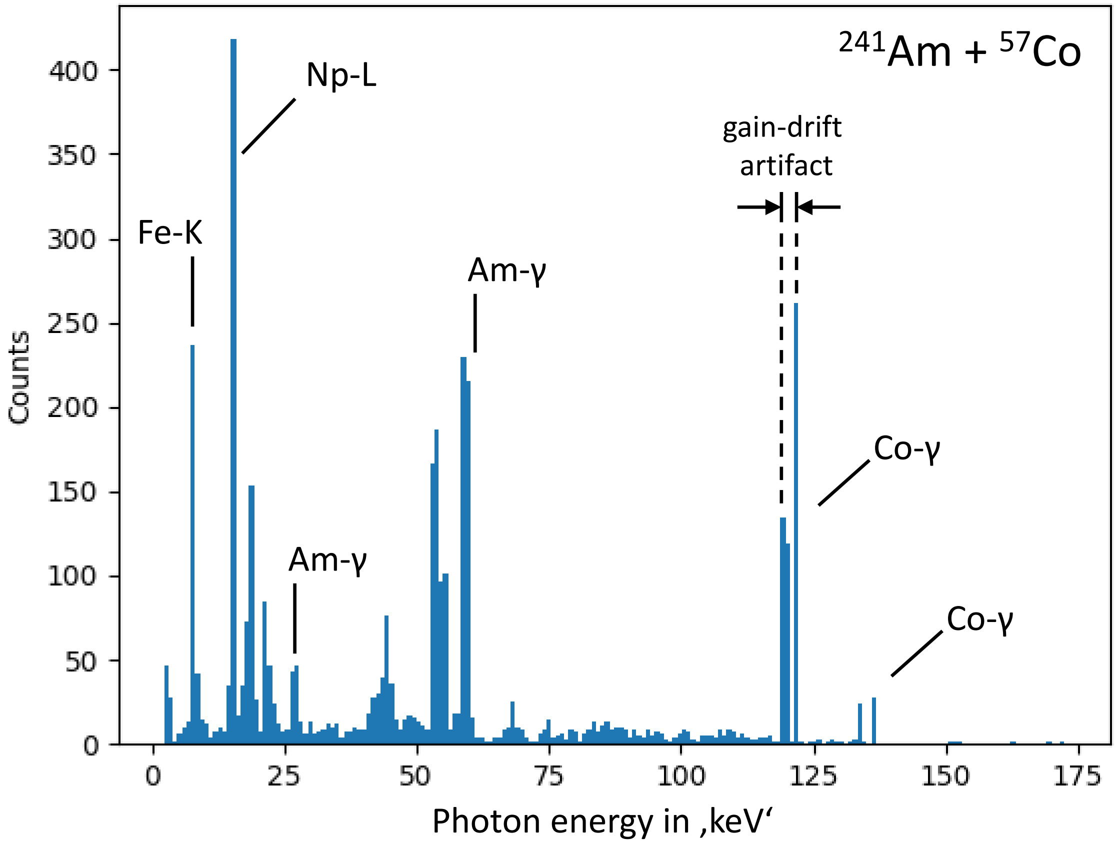


Bonus: Lets us estimate real-world parameters

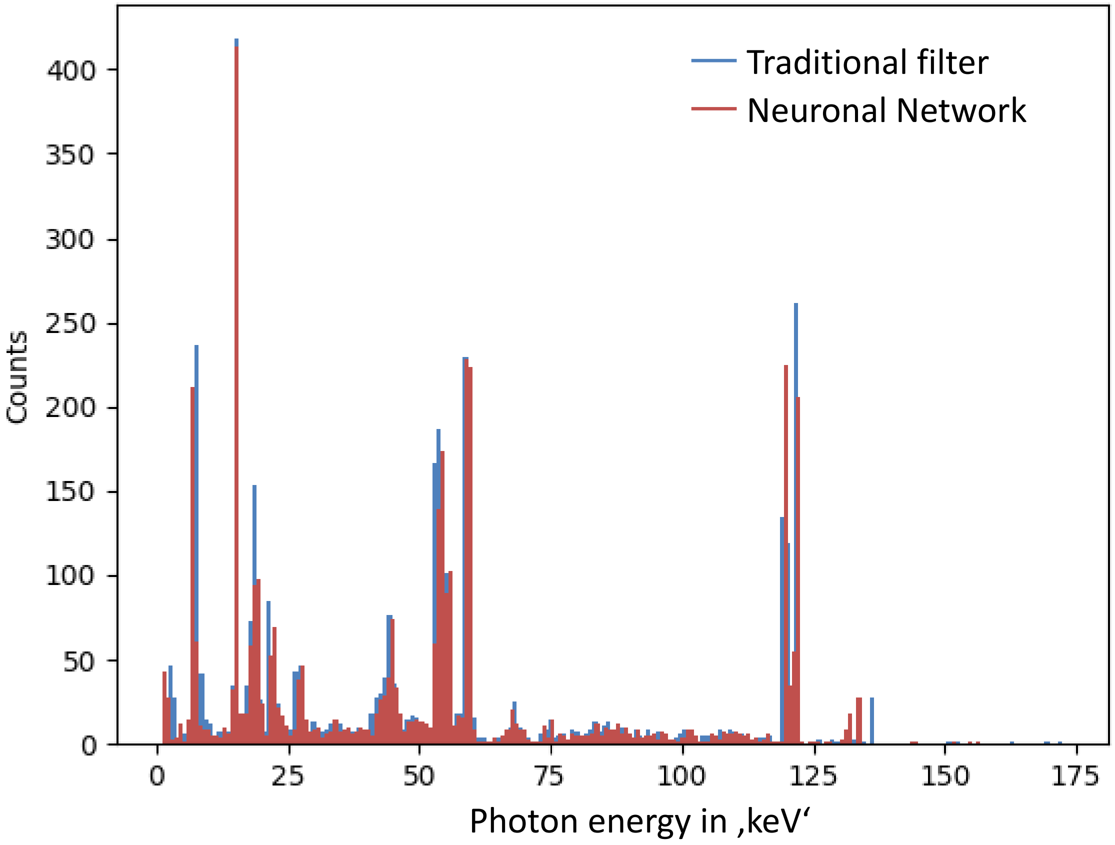


Real-World data embedding is a subvolume of synthetic Data

Results II: Real World Data



Real-world spectrum from NN trained on synthetic data

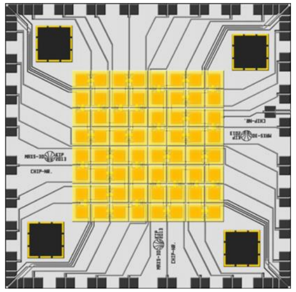


Beam time E138, 2021 at CRYRING@ESR
Single channel, no drift correction
Moving Window Deconvolution

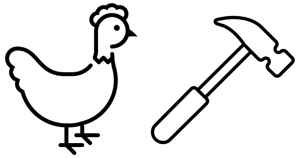
HI JENA
Helmholtz Institute Jena

Conclusion and Outlook

Summary

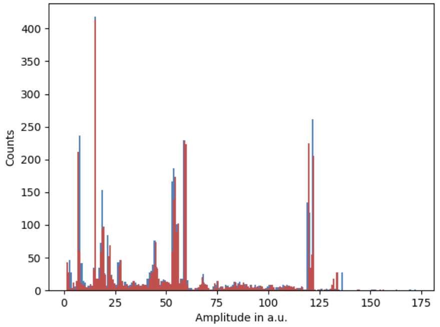


MMCs as promising tool for high precision X-ray spectrometers
Require a lot of knowledge about each individual detector channel
Need to explore new ways to optimize the read-out scheme



Many opportunities to use methods of AI to enhance MMC application
Chosen to start with pulse shape analysis because it’s most accessible
Use synthetic pulses to bootstrap the process
Make the NN learn the general pulse shape

Use CVAE to find embedding of pulse into a lower dimensional latent space
Map compressed pulse information to signal parameters via MLP
Successful benchmark against currently used FIR algorithm
Real-world data subspace of synthetic data embedding
Real-world data spectrum from NN trained on synthetic data

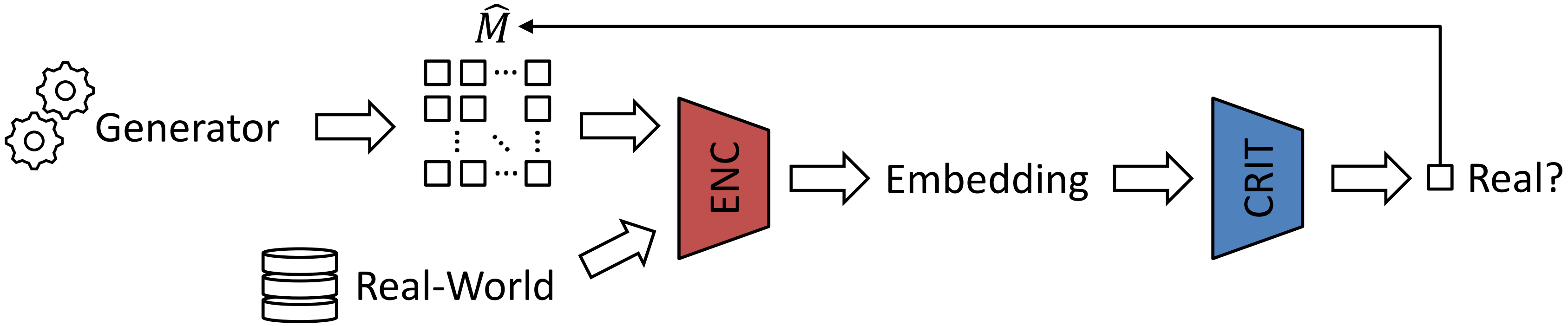


HI JENA
Helmholtz Institute Jena

Future Plans

Near: Improve inclusion of real-world data into training

Use Generative Adversary Network (GAN) to improve generator



Use CVAE+MLP to generate iteratively generate pseudo-labels

Mid: Implement more signal read-out steps

Far: Automated SQUID-tuning with supervised learning

HI JENA
Helmholtz Institute Jena

Thank your for your attention!

Thank you to everyone who has participated and helped
(non-exhaustive list)

HI-J / IOQ
C. Hahn
F. Kröger
G. Weber
O. Forstner
P. Pfäfflein
S. Bernitt
S. Fritzsche
Th. Stöhlker

GSI
A. Gumberidze
A. Kalinin
C. Brandau
G. Vorobjev
M. Lestinsky
N. Petridis
N. Stahlkamp
P.-M. Hillenbrand
S. Trotsenko
T. Gassner
T. Morgenroth
U. Spillmann
Ö. Gülsever

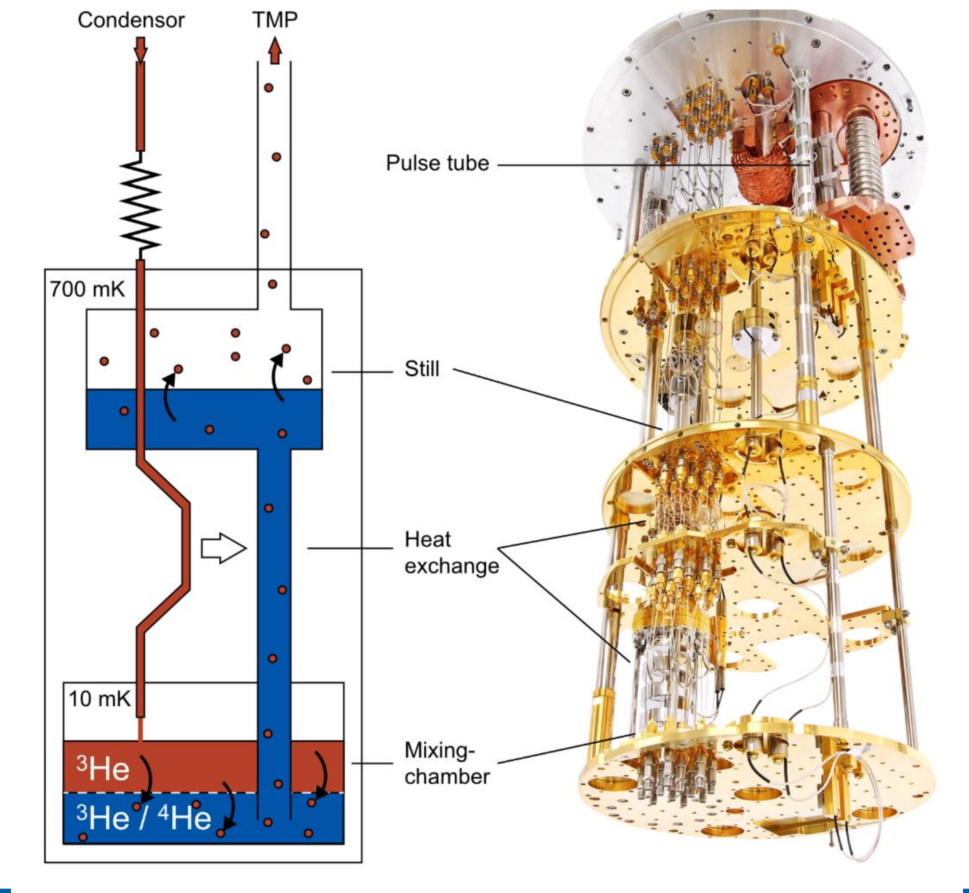
KIP
A. Fleischmann
C. Enss
C. Schötz
D. Hengstler
J. Geist
M. Keller
M. Krantz
S. Allgeier

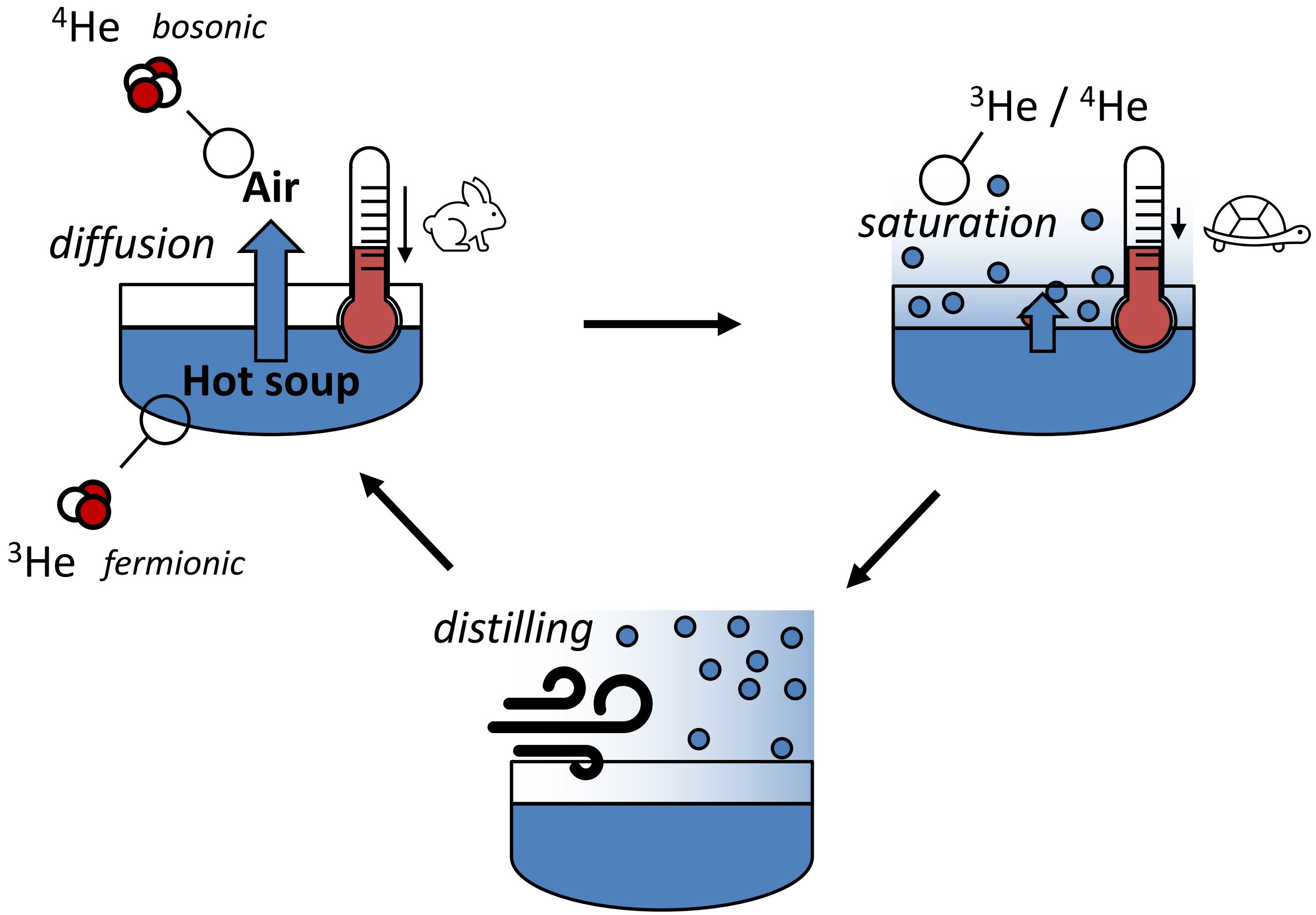
LKB
P. Indelicato

Financial support

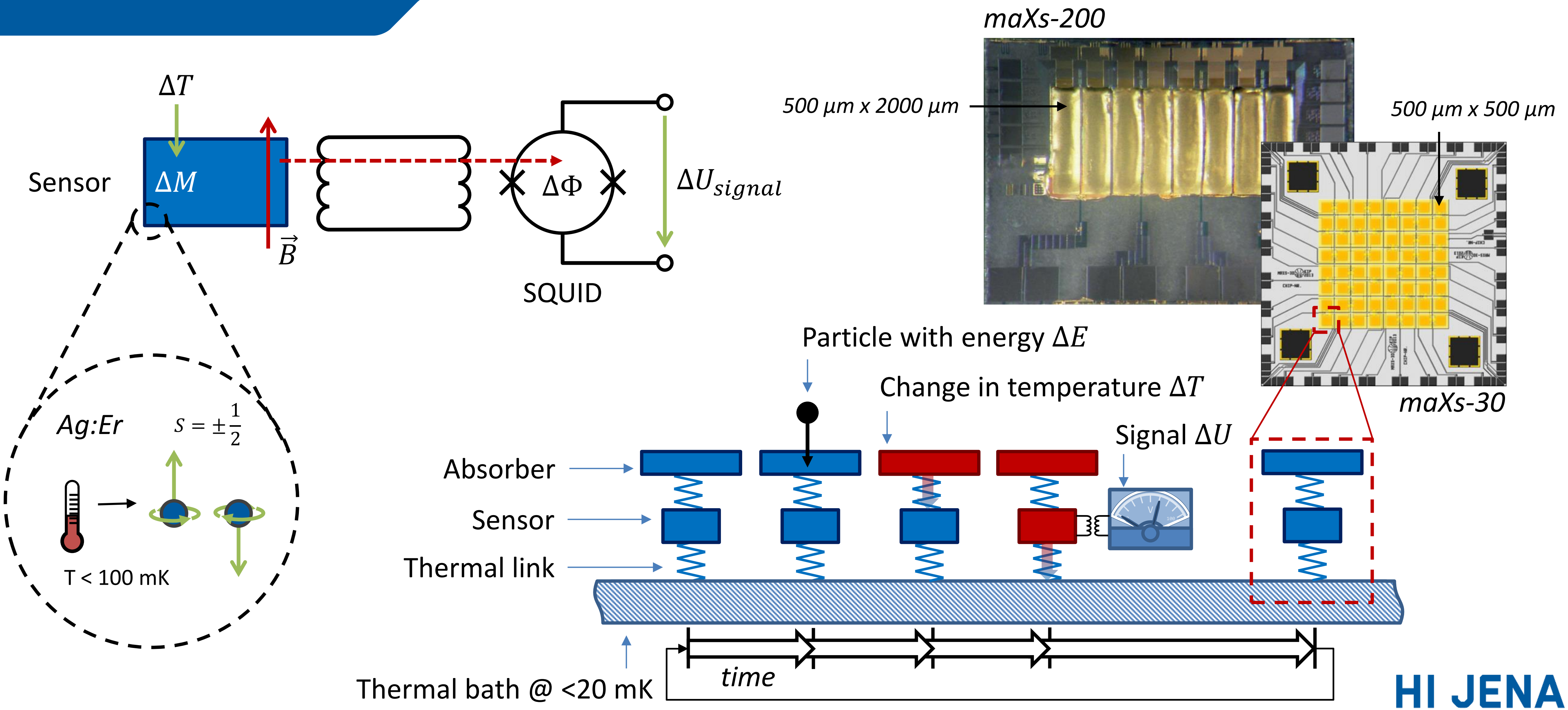
Backup

Dillusion Cryostat



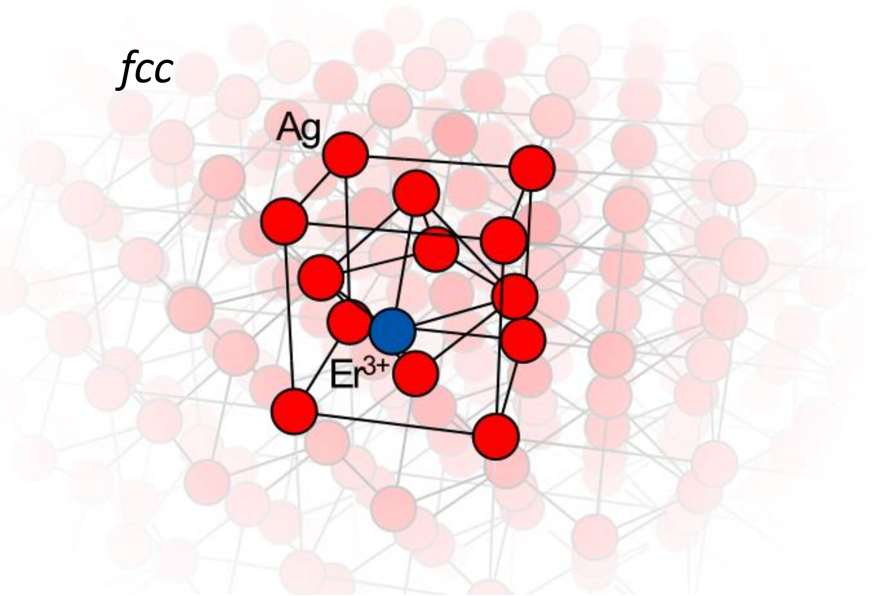


Working Principles



HI JENA
Helmholtz Institute Jena

MMC Sensor



Few 100 ppm of Er in Ag / Au

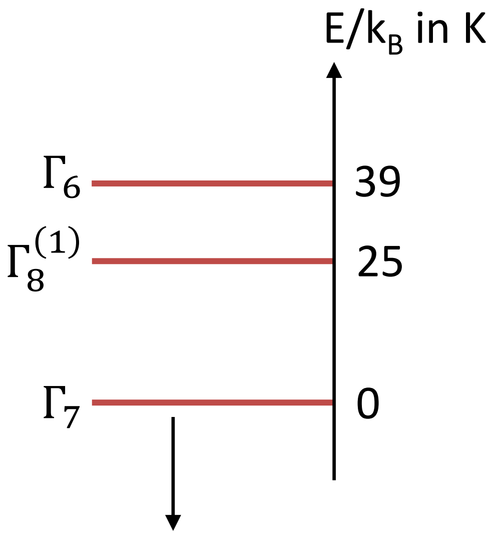
$[\mathrm{Kr}]4\mathrm{d}^{10}4\mathrm{f}^{11}5\mathrm{s}^{2}5\mathrm{p}^{6}$

2 Å, closed

0.6 Å, 3 e^- missing

→magnetic moment

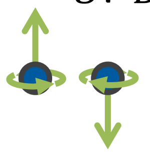
< 100 K: interaction with crystal field splits ground state

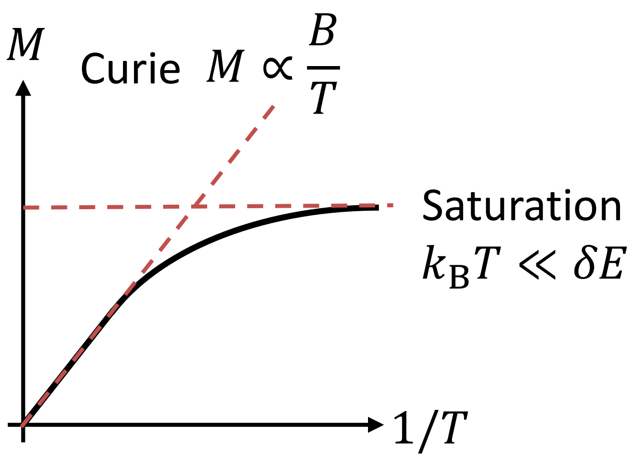


doublet $\tilde{S} = {}^{1}/_{2}$

Zeeman effect

$\delta E = \tilde{g}\mu_{\mathrm{B}}B$





Thermal energy into Zeeman System

Random spin flips out of orientation

$$\Delta M = -\frac{1}{VB}C_{\mathrm{Z}}\Delta T$$

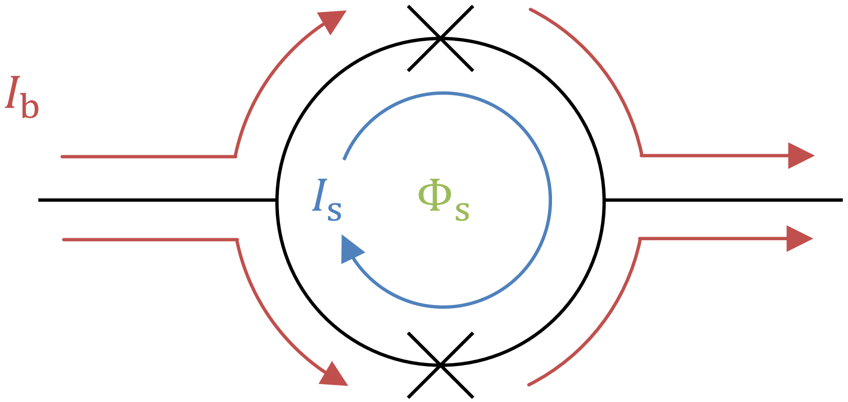
HI JENA
Helmholtz Institute Jena

SQUID

SQUID magnetometer for sensor read-out

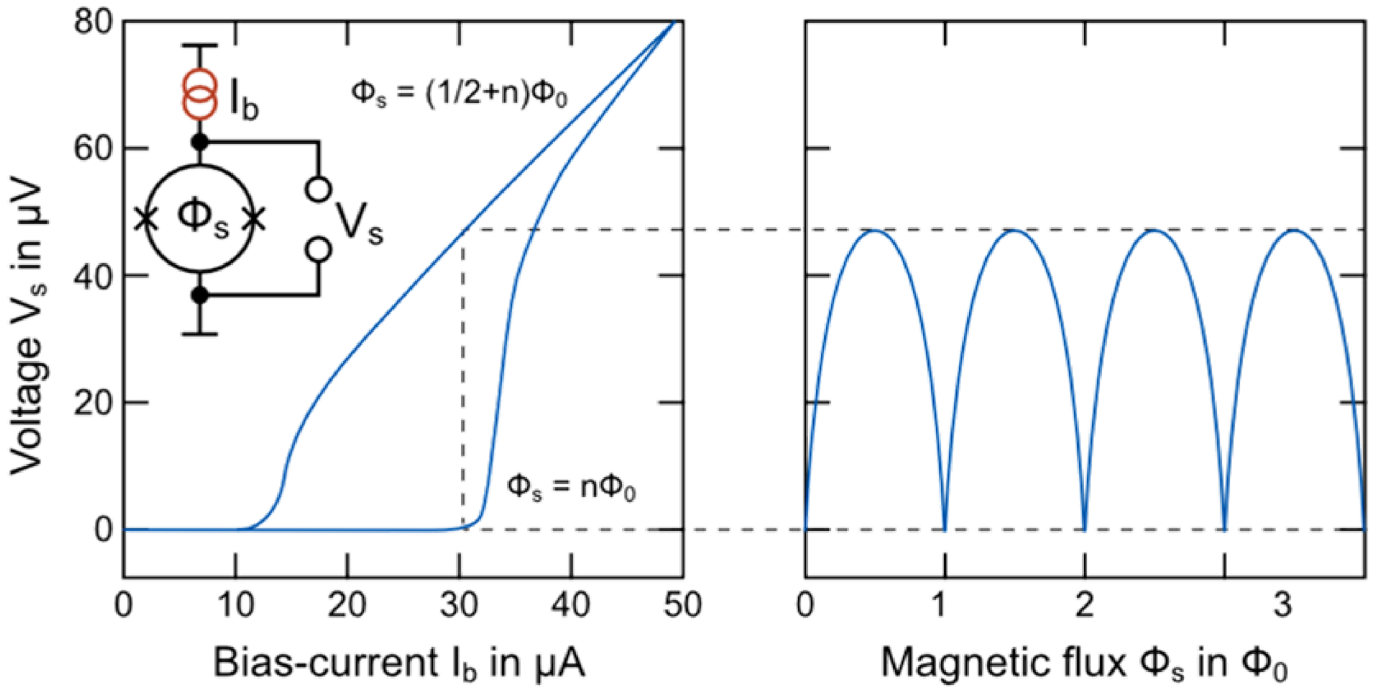
Superconducting **Qu**antum-**I**nterference **D**evice

Sensitive to changes in magnetic flux



$I_J = I_c \sin \Delta\phi$ ⟶ Flux modulates phase

Like optical interferometer



Flux-quantization in closed sc loop → shielding I_s

Overlapping bias I_b and shielding I_s shift critical I_c

Quasi-particle flow leads to U-drop across SQUID

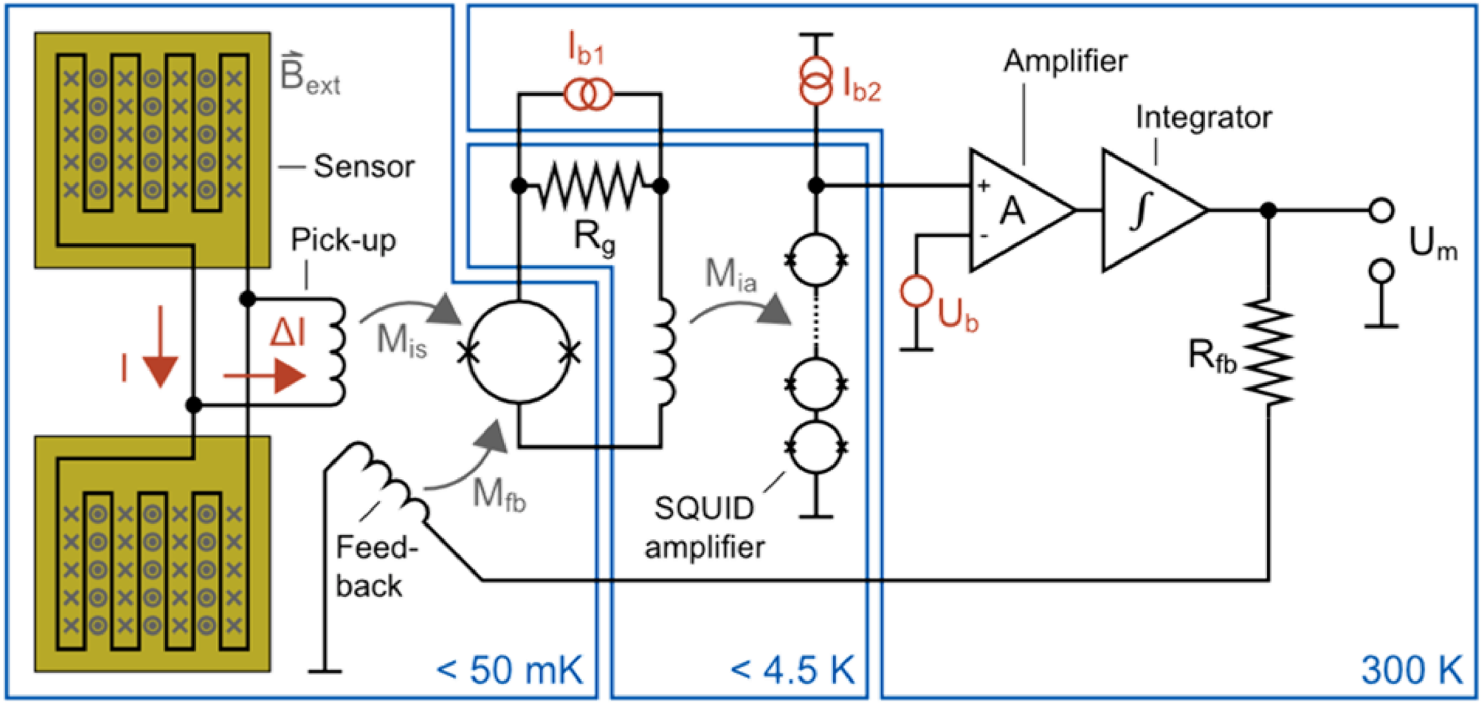
HI JENA
Helmholtz Institute Jena

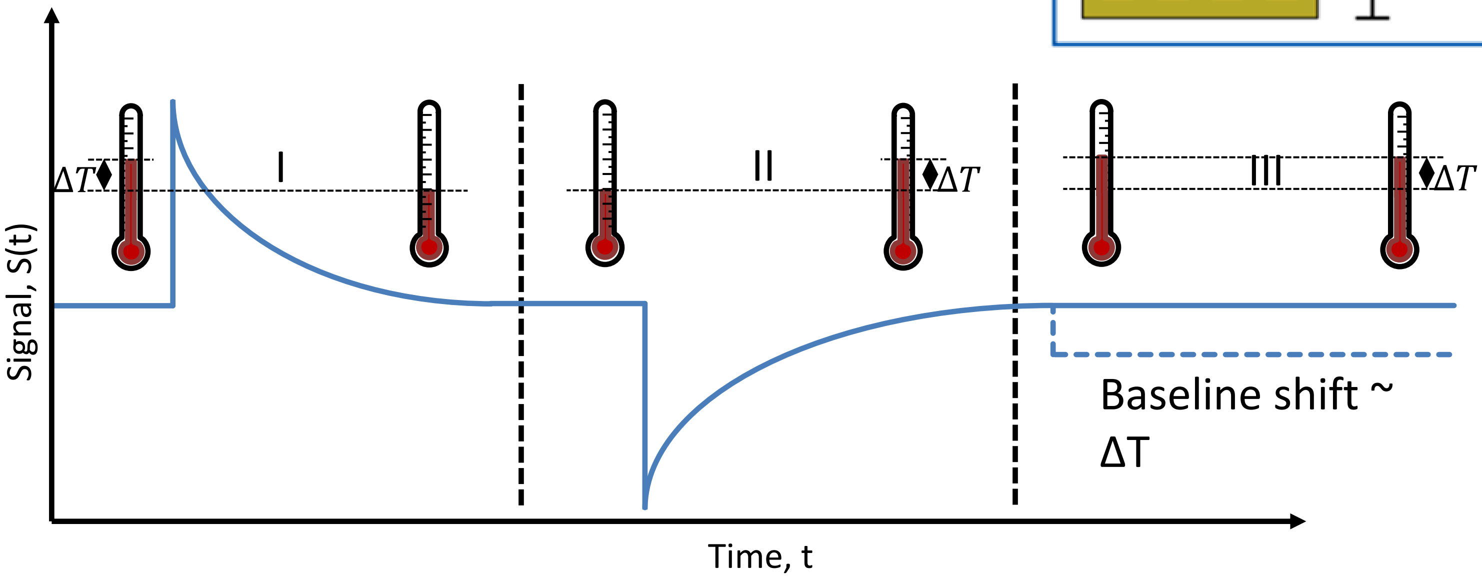
Read-out and Amplification

Gradiometric setup cancels global temperature fluctuation

Imbalance by design creates temperature sensitive pixels

Flux-locked loop (FLL) to stabilize operation point
Cryogenic SQUID based multistage amplifier

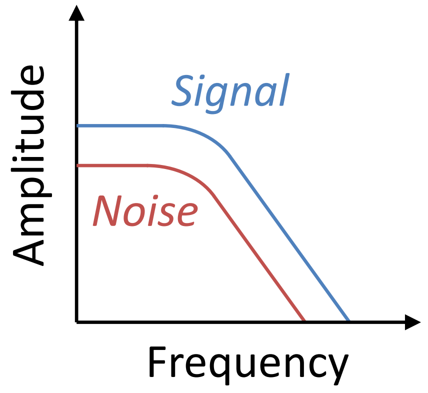




Resolution

Ideal Case: Only one thermal capacity

Noise from thermal fluctions



const. SNR

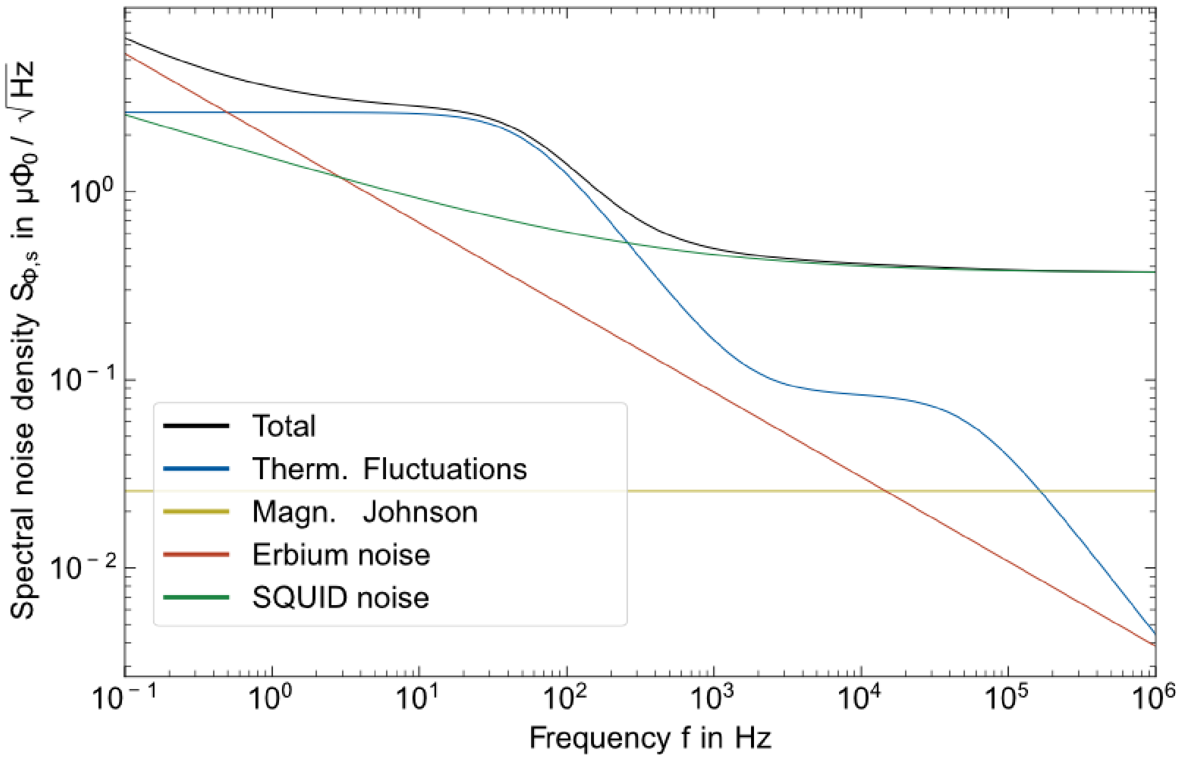
↓

arbitrary resolution

$$\Delta E = \sqrt{4k_B T^2 C} \cdot \left(\frac{1}{\beta(1-\beta)} \frac{\tau_0}{\tau_1} \right)^{-\frac{1}{4}}$$

Reality: Multiple coupled systems

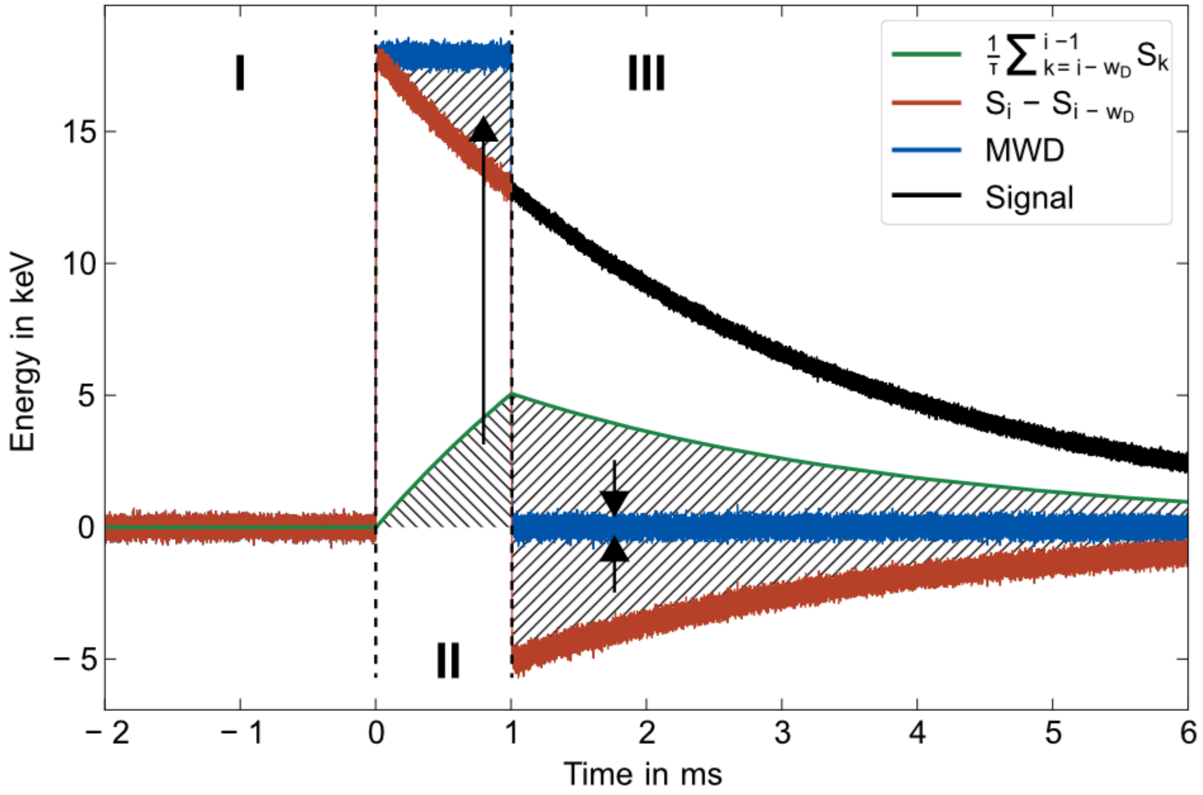
Additional noise sources (environment, read-out, …)

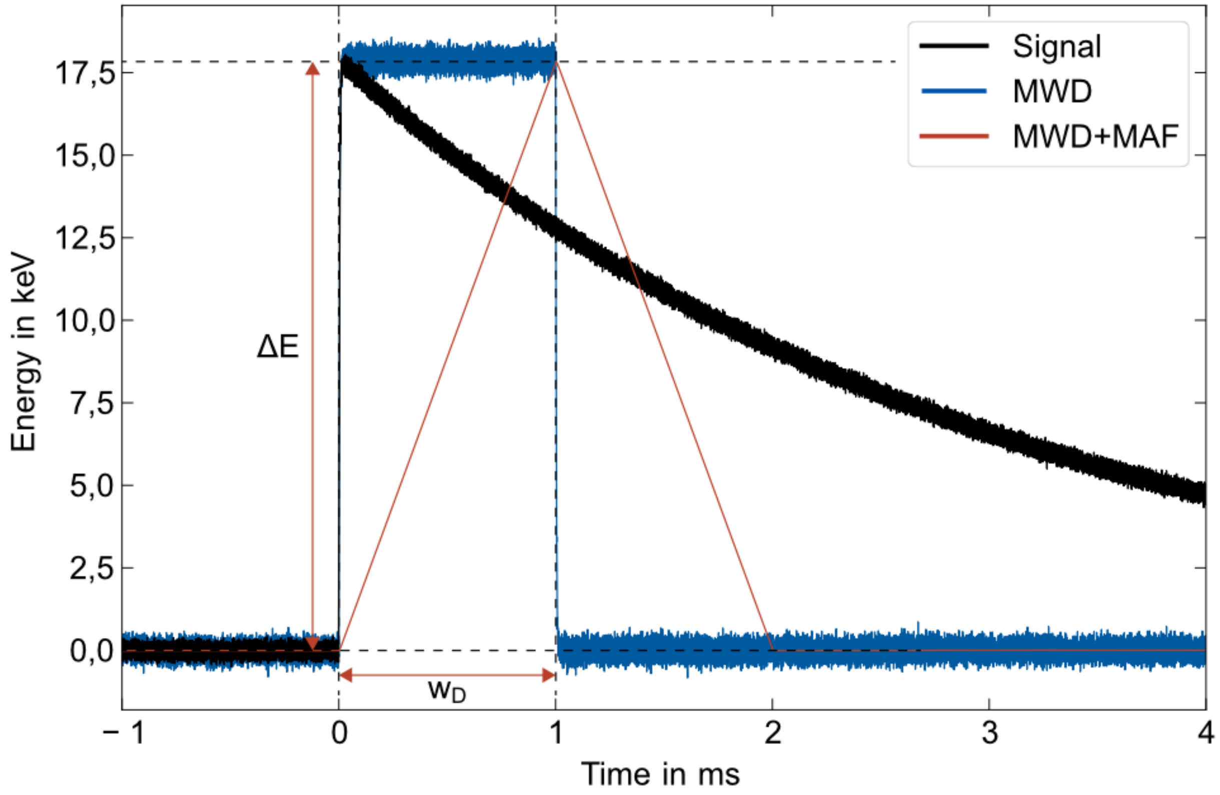


→ (Fundamental) resolution is determined by bandwidth limitations

Can be optimized by component design

FIR-based Signal Processing





Moving Window Deconvolution

$$MWD(S;t_i) \coloneqq S(t_i) - S\left(t_{i-w_D}\right) + \frac{1}{\tau}\sum_{k=i-w_D}^{i-1} S(t_k)$$

Moving Average Filter

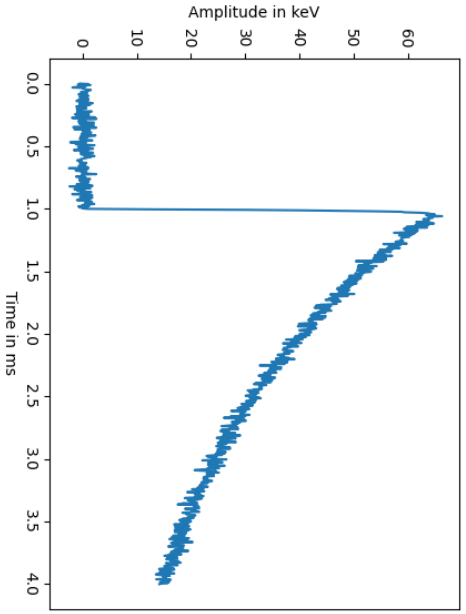
HI JENA
Helmholtz Institute Jena

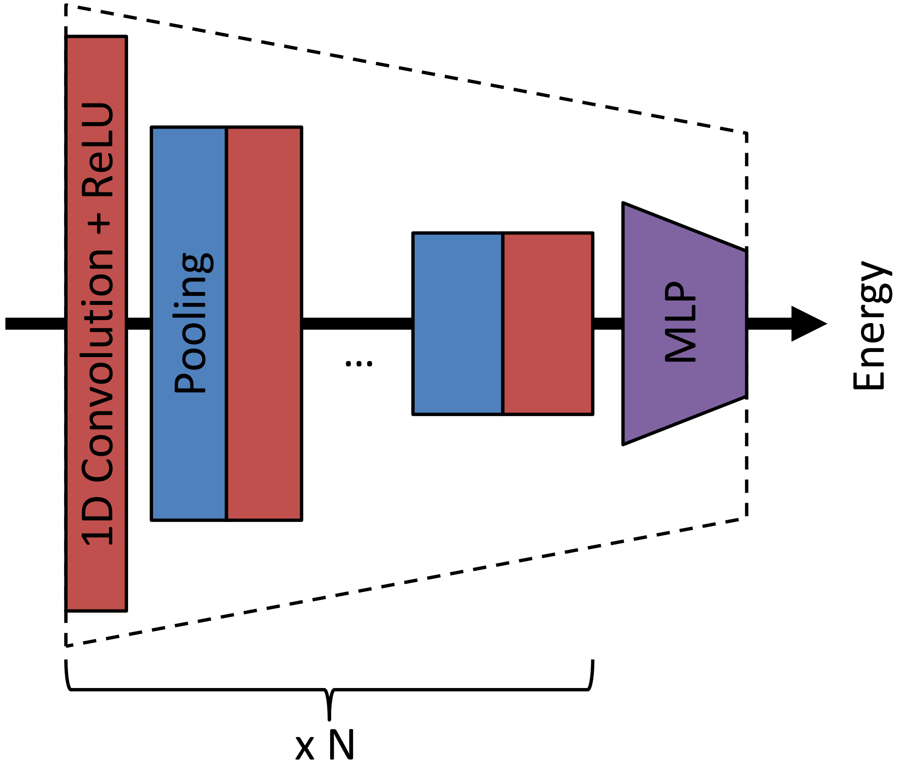
First Attempt: CNN

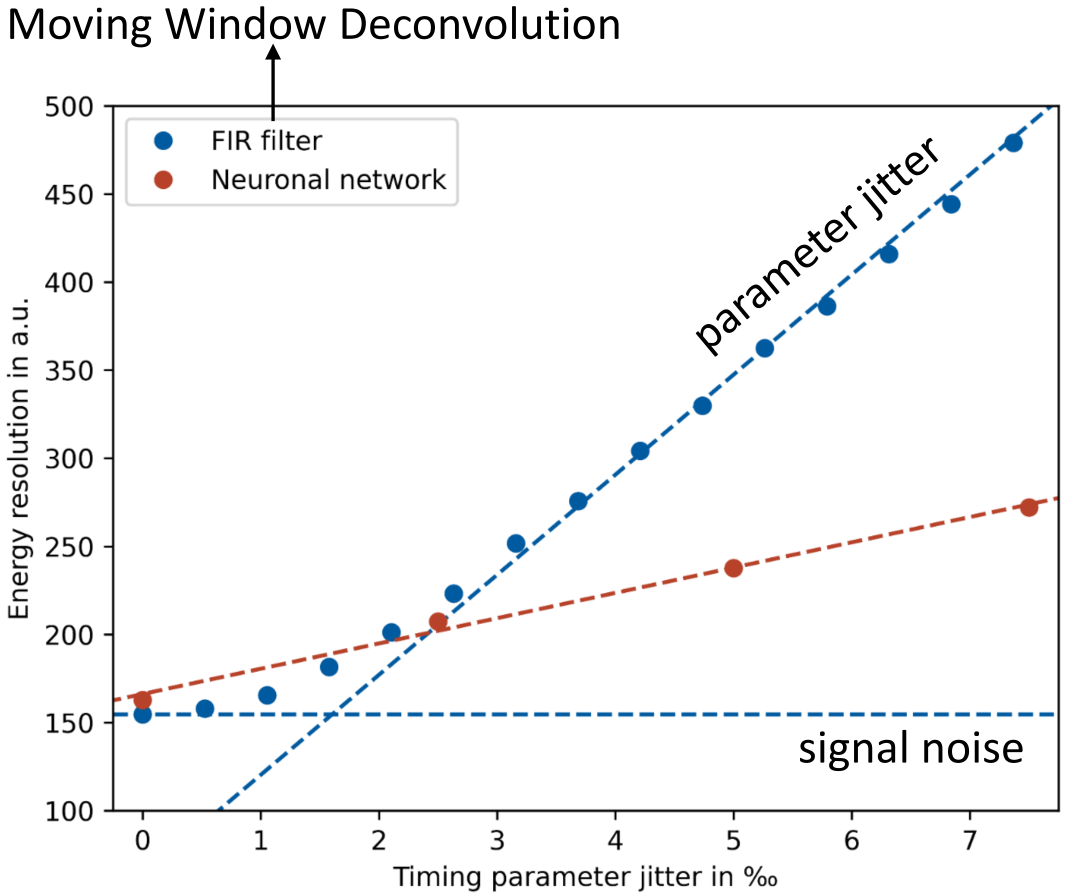
MLP = Multi-Layer Perceptron
FIR = Fininte Impulse Response

Source: S. Kiranyaz et al., *ICASSP* **8360** (2019)

Convolutional Neuronal Network (CNN)







Comparable results to traditional filtering method.

- ➕ Outperforms FIR aproach in scenarios with high parameter jitter.
- ➖ Unstable during training.
- ➖ Difficult to train with real-world data.

HI JENA
Helmholtz Institute Jena