Towards Photon Counting at the ALPS II Experiment:
Efficient Background Discrimination for TES-Based Single-Photon Detectors

E. Rivasto[1], J. A. Rubiera Gimeno[2] , K.-S. Isleif[4], F. Januschek[2], A. Lindner[2], M. Meyer[1], G. Othman[3], C. Schwemmbauer[2]

[1]CP3-Origins, University of Southern Denmark, Odense, Denmark
[2]Deutsches Elektronen-Synchrotron DESY, Hamburg, Germany
[3]Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany
[4]Helmut-Schmidt-University, Hamburg, Germany

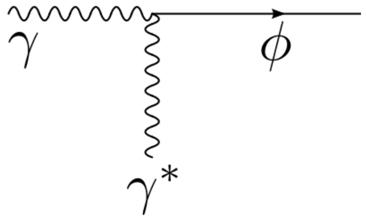
ALPS II @ DESY

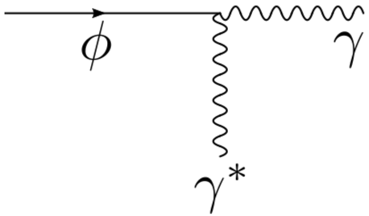
- **A**ny **L**ight **P**article **S**earch **II**
- Laboratory experiment looking for:
  - Weakly Interacting Sub-eV Particles (WISPs)
    - Axions
    - Axion-like particles (ALPs)
- Model independent!
- Located at DESY, Hamburg, Germany (old HERA)
- Started data taking on May 2023 – still improving sensitivity

Motivation: **Axions**

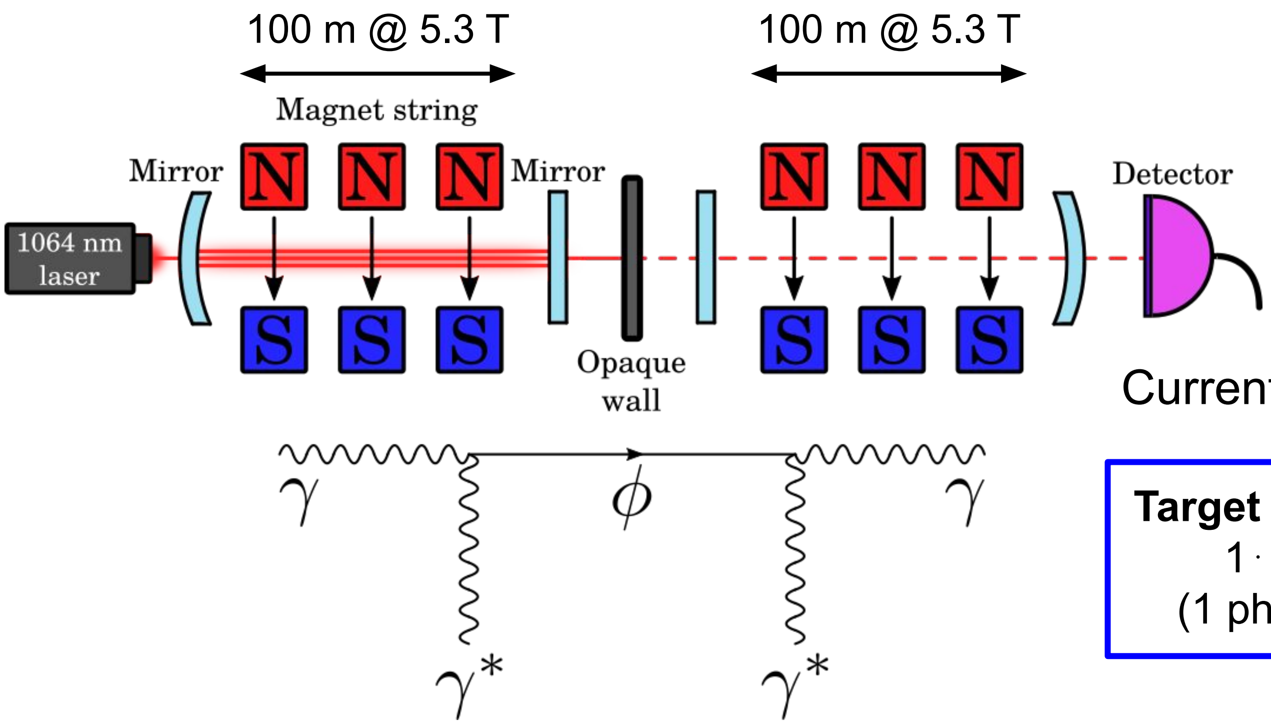
- Solution for the strong CP-problem
  - Peccei & Quinn (1977)
  - Treat θ as dynamical field →Spontaneous relaxation to zero
  - Proposed $U(1)_{PQ}$ symmetry gives rise to the **axion**
  - Weak coupling ($g_{a\gamma\gamma}$) to photons→Candidate for **dark matter**!
- TeV transparency of the universe
- Stellar cooling
  - Limits coupling to $g_{a\gamma\gamma}$~10^{-12} GeV^{-1}

$$\mathcal{L}_{\mathrm{QCD}} \supset \mathcal{L}_{\mathrm{CP-viol.}} = \frac{\alpha_s}{4\pi} \cdot \theta \cdot \mathrm{Tr}(G_{\mu\nu}\tilde{G}_{\mu\nu})$$





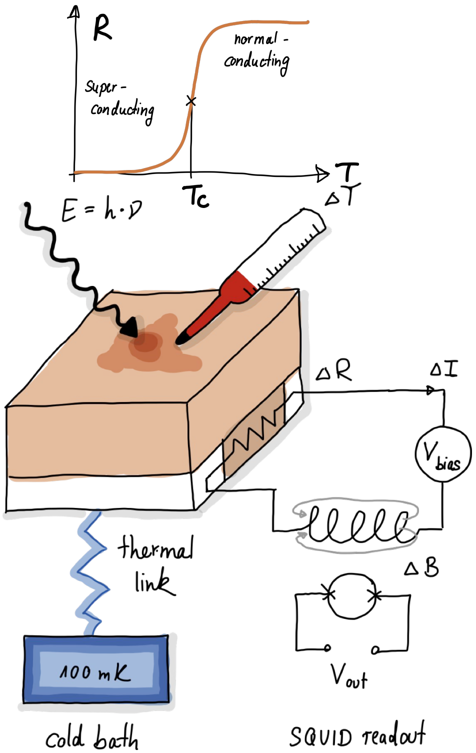
ALPS II – A Light Shining Through a Wall Experiment



Currently: HET

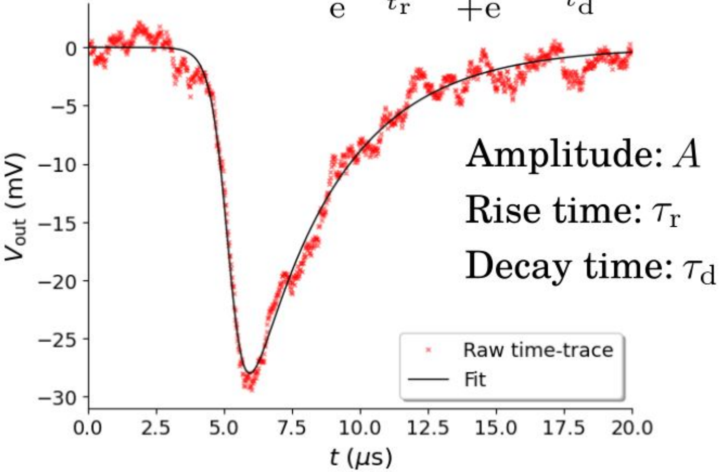
Target sensitivity:
$1 \cdot 10^{-5}$ Hz
(1 photon/day)

Transition Edge Sensors (TES)



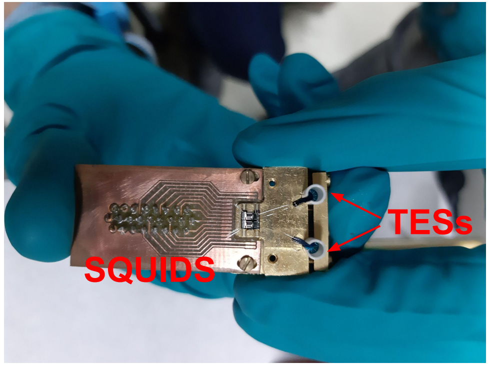
Photon Energy ∝ Pulse Integral

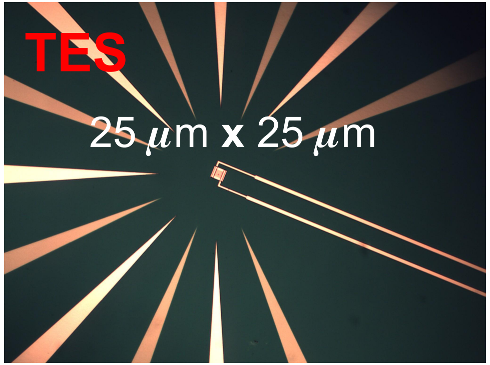
$$V_{\mathrm{out}}(t) = -\frac{2A}{e^{\frac{(t_0-t)}{t_r}} + e^{-\frac{(t_0-t)}{t_d}}}$$



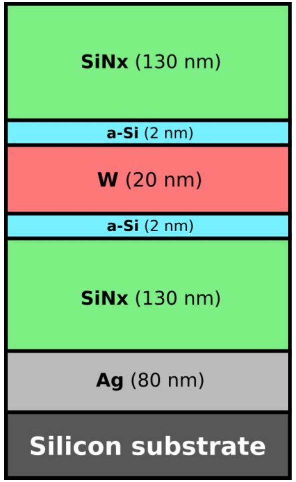
Transition Edge Sensors (TES)

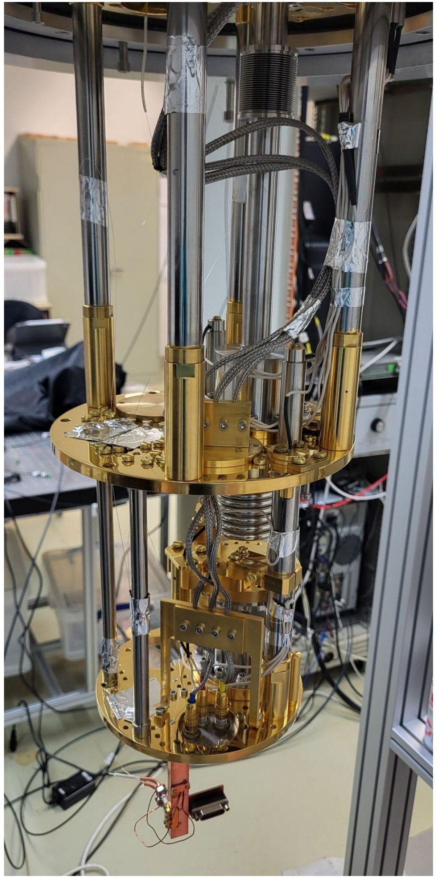
- TESs provided by NIST
  - Based on tungsten ($T_c \approx 140$ mK)
  - Optimized for 1064 nm photons
- Packaging and SQUIDs provided by PTB
- Operated within Bluefors SD dilution refrigerator





TES structure:





Transition Edge Sensors (TES)

- Energy resolution of **5%** (σ) @ 1.165 eV (1064 nm photon)
    - 1.165 eV, σ=0.06 eV
    - 1064 nm, σ= 64 nm
- Detection efficiency **>80%** (further measurement ongoing)

PROCEEDINGS OF SCIENCE

A TES system for ALPS II - Status and Prospects

José Alejandro Rubiera Gimeno,[a,*] Friederike Januschek,[a] Katharina-Sophie Isleif,[a,] Axel Lindner,[a] Manuel Meyer,[b] Gulden Othman,[c] Christina Schwemmbauer[a] and Rikhav Shah[d,***]**

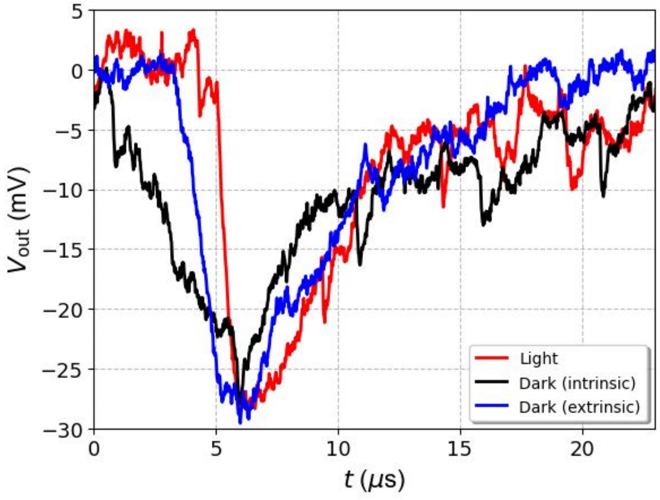
[a]*Deutsches Elektronen-Synchrotron (DESY), Notkestr. 85, 22607 Hamburg, Germany*
[b]*CP3-Origins, University of Southern Denmark, Campusvej 55, 5230 Odense M, Denmark*
[c]*Institut für Experimentalphysik, Universität Hamburg (UHH), Notkestr. 85, 22607 Hamburg, Germany*
[d]*Institute for Physics, Johannes-Gutenberg-Universität (JGU), Staudingerweg 7, 55128 Mainz, Germany*
[**]*now at Helmut-Schmidt-Universität (HSU)*
[***]*now at Universität Hamburg (UHH)*
E-mail: jose.rubiera.gimeno@desy.de, friederike.januschek@desy.de, katharina-sophie.isleif@desy.de, axel.lindner@desy.de, manuel.meyer@desy.de, gulden.othman@desy.de, rikhav.shah@desy.de, christina.schwemmbauer@desy.de

[1] Gimeno, J. A. R., Januschek, F., Isleif, K. S., Lindner, A., Meyer, M., Othman, G., Schwemmbauer, C., & Shah, R. (2024). A TES system for ALPS II - Status and Prospects. Proceedings of Science, 449, Article 567

Background Reduction

GOAL: Distinguish between 1064 nm photon induced **LIGHT** pulses from background induced **DARK** pulses

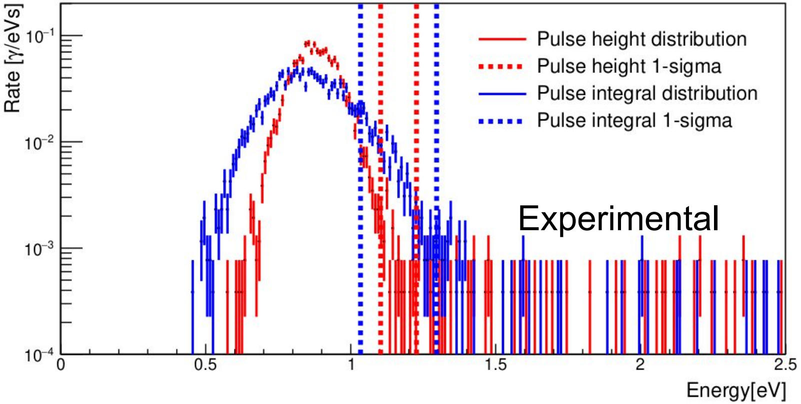
- **Intrinsic background**
  - Radioactive decays
  - Electronic noise
- **Extrinsic background**
  - Black-body radiation



Black-body radiation

- Background limited by TES energy resolution
- Currently minimum rate **$6.9 \cdot 10^{-5}$ Hz**

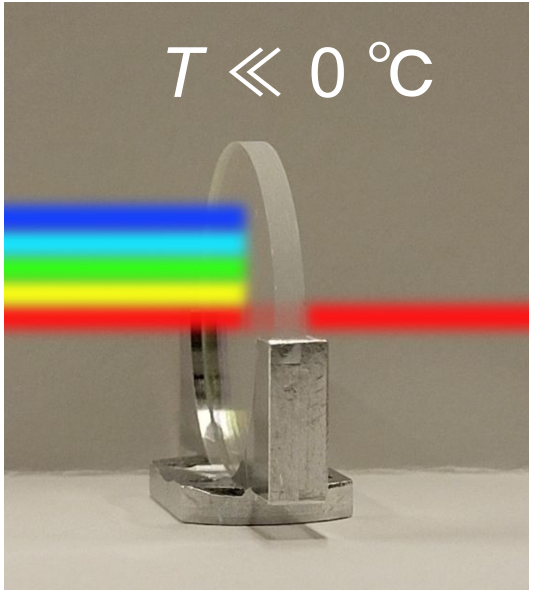
Target sensitivity: $1 \cdot 10^{-5}$ Hz (1 photon/day)



[1] Jose Alejandro Rubiera Gimeno, Dissertation 2024, Optimizing a Transition Edge Sensor detector system for low flux infrared photon measurements at the ALPS II experiment

Cryogenic Optical Filter Bench

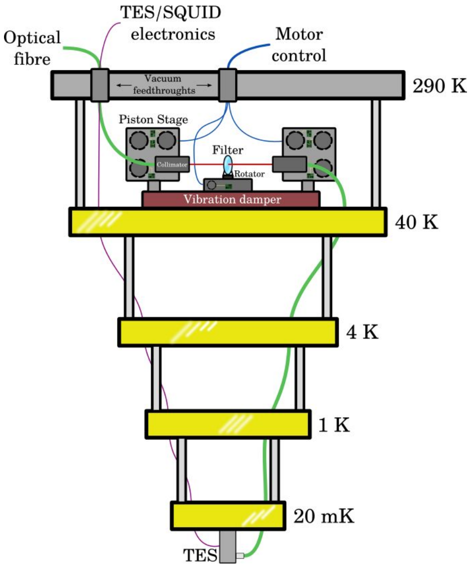
- Ultra Narrow-Band Pass Filter: **1064 nm ± 1 nm**
  - Improve energy resolution! (TES: 1064 nm, σ=64 nm)
  - Remove pileups

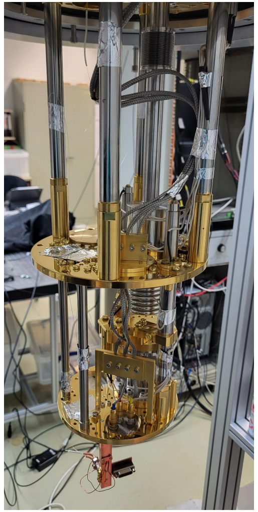


Cryogenic Optical Filter Bench

- Place inside dilution fridge at ~40 K
- Main challenges:
  - Thermal contraction
  - Vibration

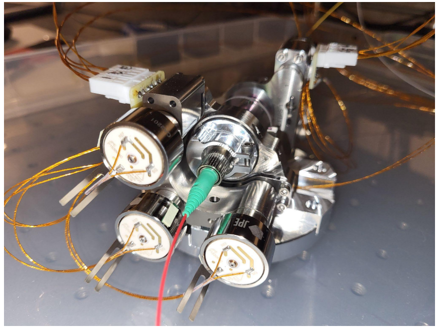
→**Misalignment**

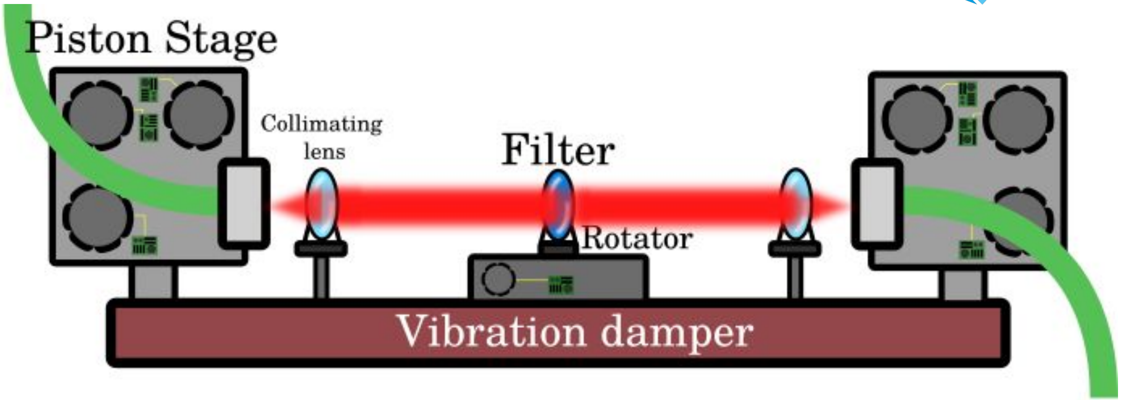


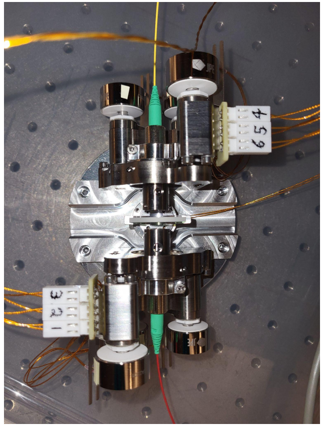


Remotely Controllable Cryogenic Piston Stages

- 3 angles + distance between fibre-end and lens
- Rotator tunes filters transmission window
- Vibrational damper stabilizes the system
- Expecting 70-80% transmission coefficient

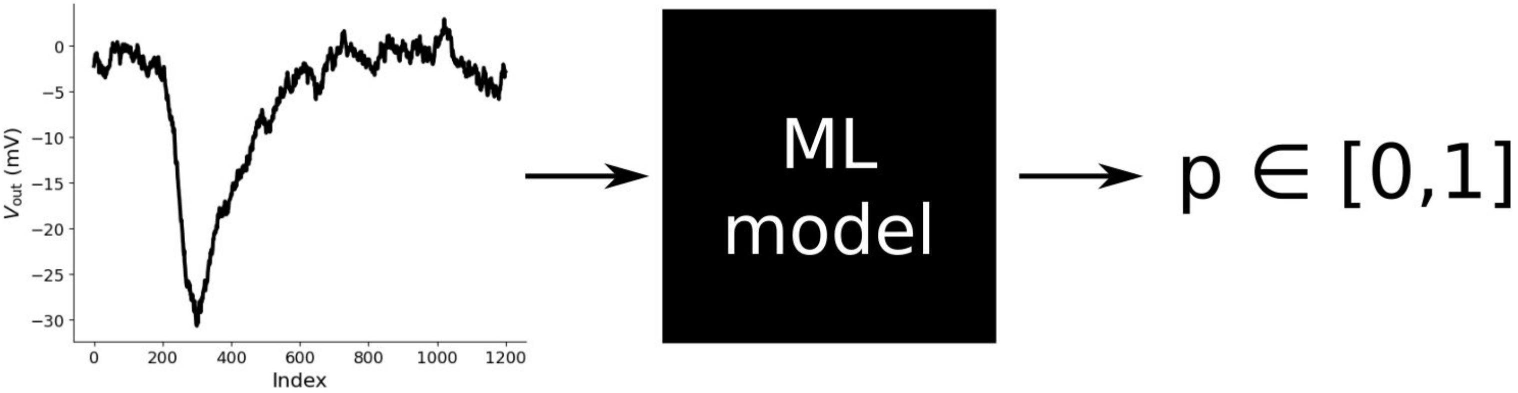






Intrinsic background: Machine Learning

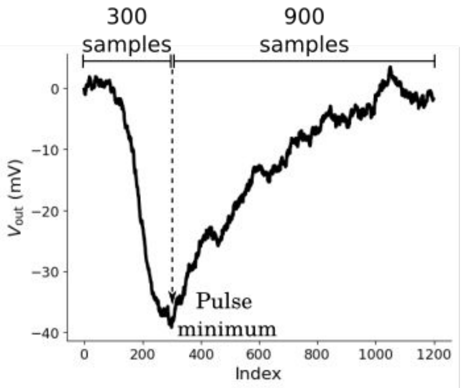
- Binary classification: **LIGHT** vs **DARK**
- **Input** = Time trace
- **Output** = probability that pulse is light
- Convolutional Neural Networks (CNN)



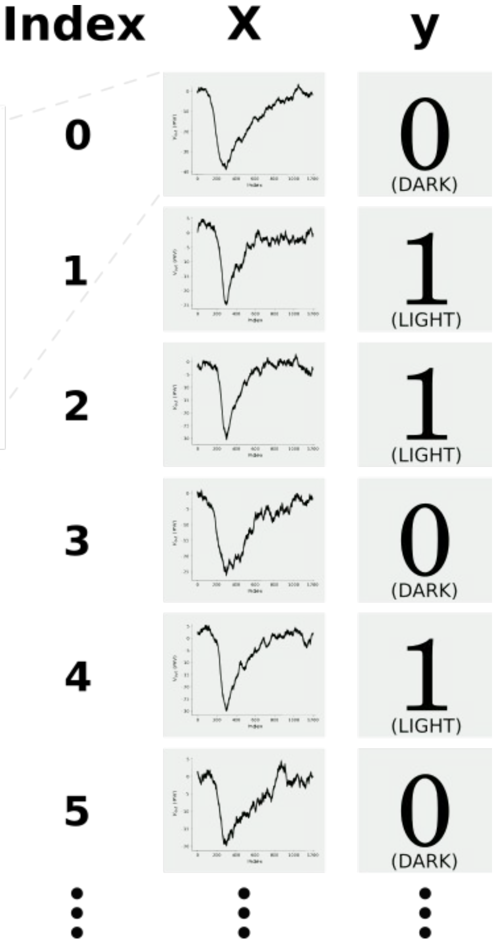
Machine Learning: Dataset

Sampled light-pulses: **25,000**
Sampled dark-pulses: **25,000**

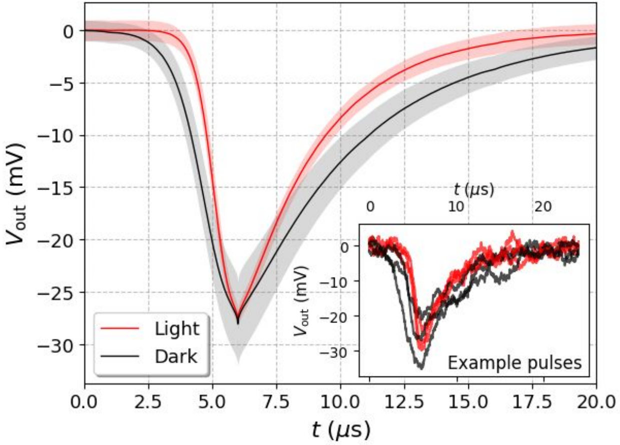
Total dataset size: **50,000**



$= (V_0, V_1, V_2, \ldots, V_{1199})$

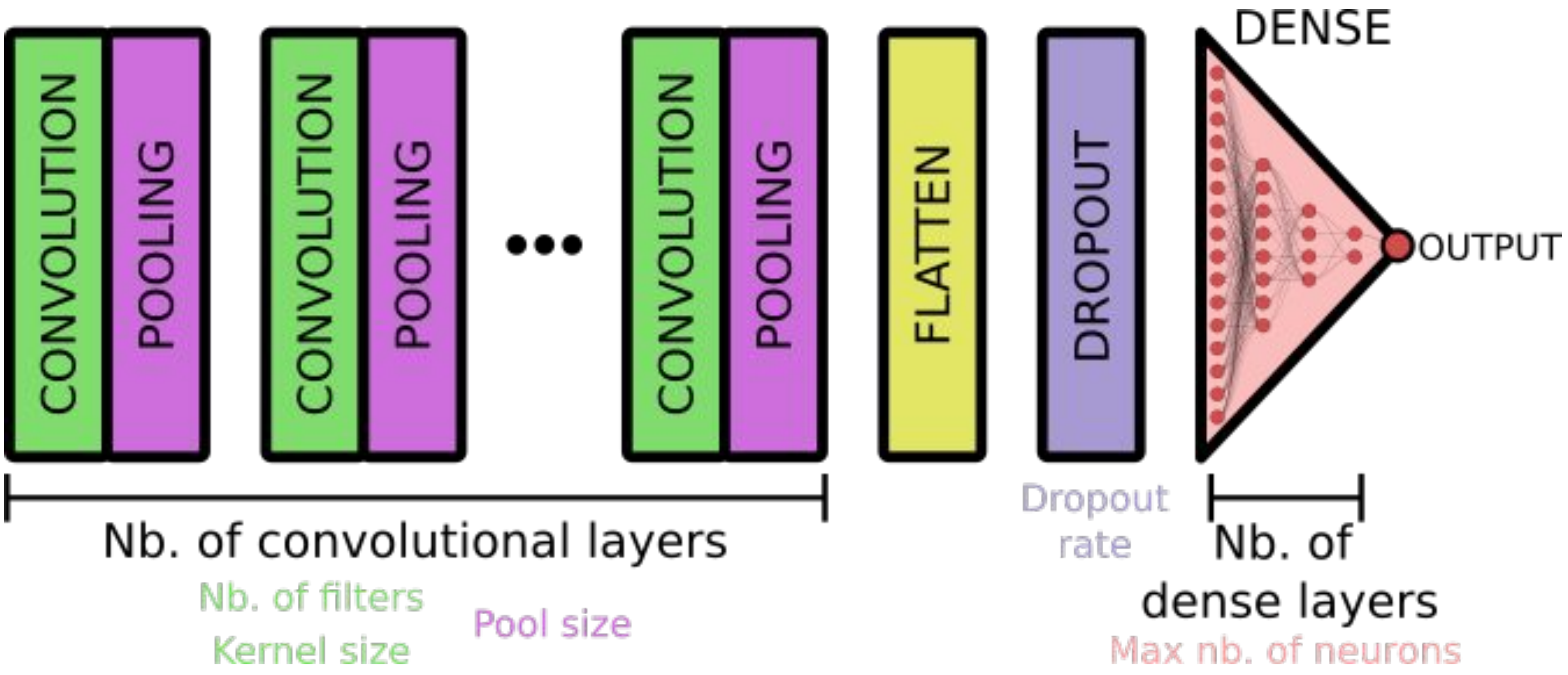


Average simulated pulses



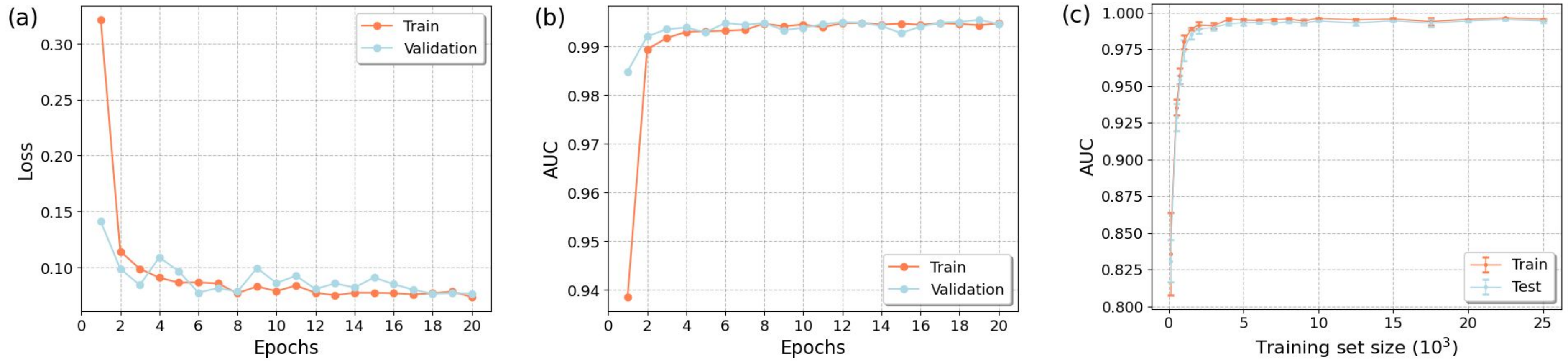
$$V_{out}(t) = -\frac{2A}{e^{\frac{(t_0 - t)}{t_r}} + e^{-\frac{(t_0 - t)}{t_d}}}$$

Convolutional Neural Network (CNN)



Checking for overfitting - Memorization VS Generalization

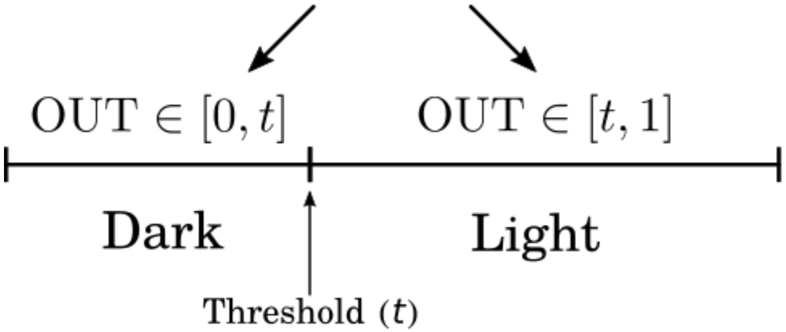
- Very well balanced model
  - Negligible generalization gap
    - No over/underfitting observed
- **AUC=0.995**

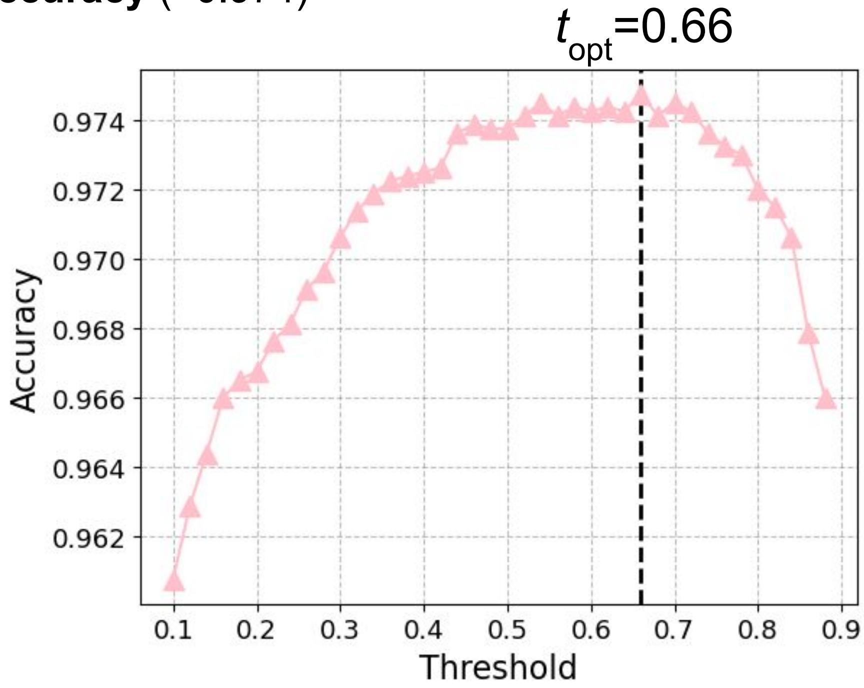


Determining optimal threshold

- Determine optimal threshold based on max **accuracy** (~0.974)

CNN's OUTput





Improvements from previous work

- **Cut-based analysis**
  - Discriminate pulses based on fitting parameters
  - PCA
- **Machine Learning on fitting parameters**
  - Random Forest Classifier (RFC)

Journal of Low Temperature Physics (2022) 209:355–362
https://doi.org/10.1007/s10909-022-02720-0

Check for updates

Characterising a Single-Photon Detector for ALPS II

Rikhav Shah[1] · Katharina-Sophie Isleif[2] · Friederike Januschek[2] · Axel Lindner[2] · Matthias Schott[1]

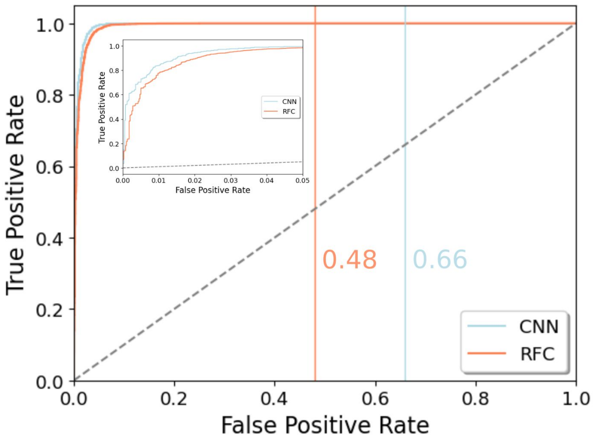
Received: 30 October 2021 / Accepted: 28 March 2022 / Published online: 28 April 2022

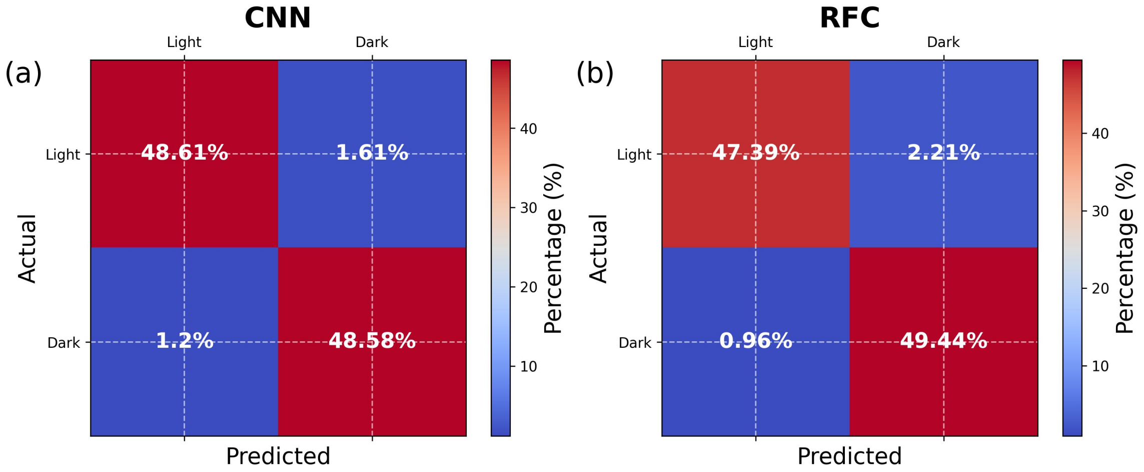
RESEARCH ARTICLE annalen der physik
www.ann-phys.org

A First Application of Machine and Deep Learning for Background Rejection in the ALPS II TES Detector

Manuel Meyer, Katharina Isleif, Friederike Januschek, Axel Lindner, Gulden Othman, José Alejandro Rubiera Gimeno, Christina Schwemmbauer, Matthias Schott, Rikhav Shah, and for the ALPS Collaboration*

Improvements from previous work





CNN: AUC=0.9954 ± 0.0004

RFC: AUC=0.9929 ± 0.0004

	CNN	RFC
Accuracy (%)	97.19	96.83
Precision (%)	97.59	98.01
Recall (%)	96.79	95.55
F_1-score (%)	97.19	96.76

Conclusions

- **ALPS II is looking for axions/ALPS**
- **Target sensitivity: 1 photon/day**
- **Detection system sensitivity improvement by**
    - **TES**
        - **Cryogenic Optical Filter Bench**
            - *Suppression* of extrinsic background (black-body radiation)
            - “Improves the energy resolution of the TES”
            - Work ongoing
        - **Deep learning**
            - Discrimination of intrinsic background (radioactive decays etc.)
            - CNN showed promising results: further testing/development on-going

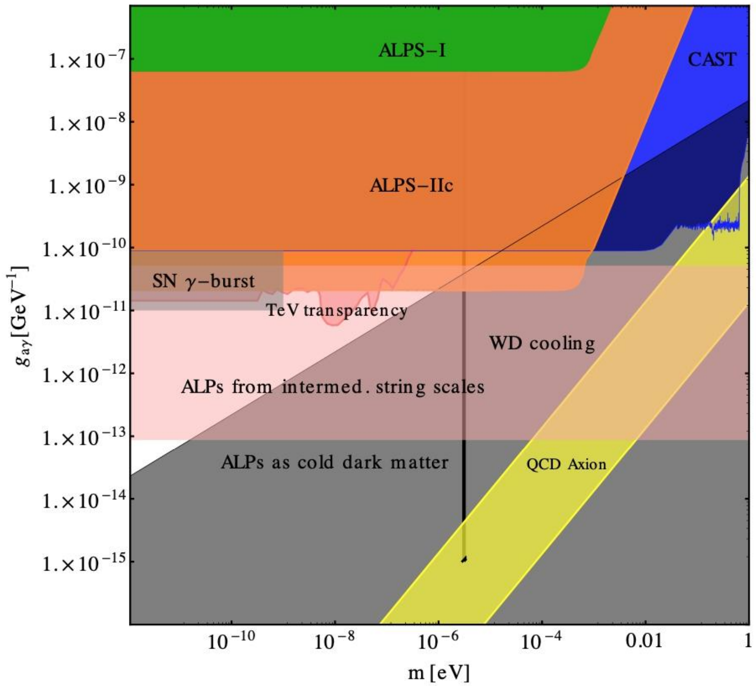
Thank you!

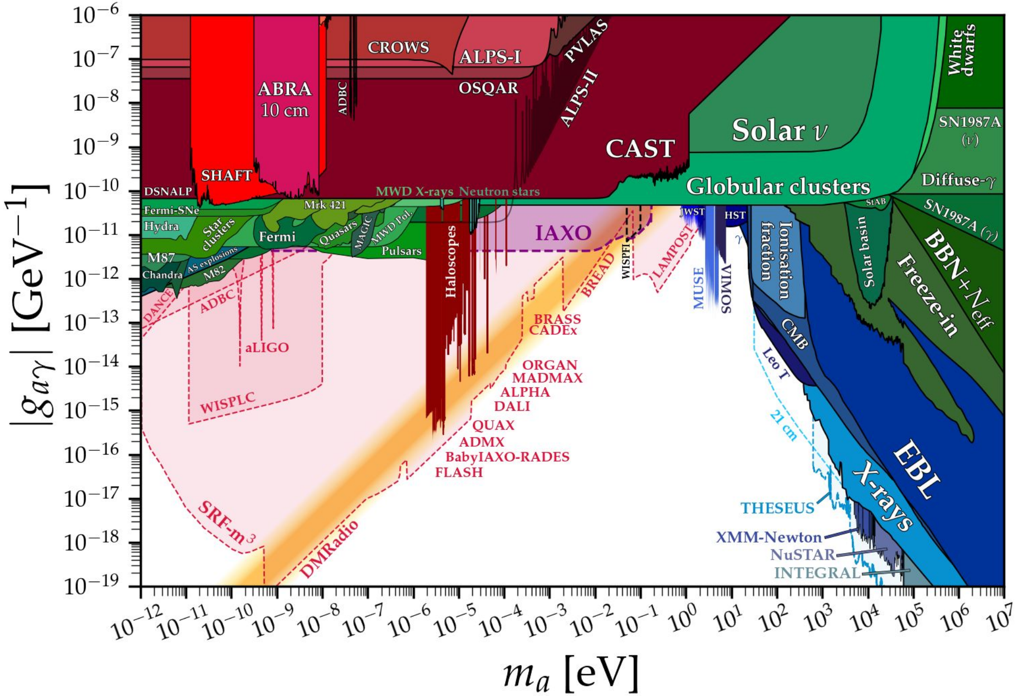
On the behalf of

ALIPS

SUPPORTING SLIDES

ALPS II - Sensitivity Reach

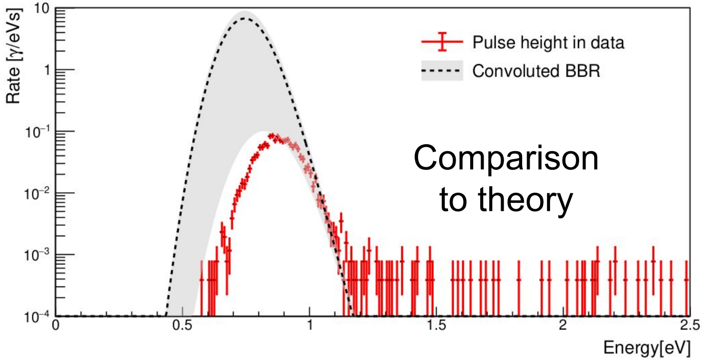


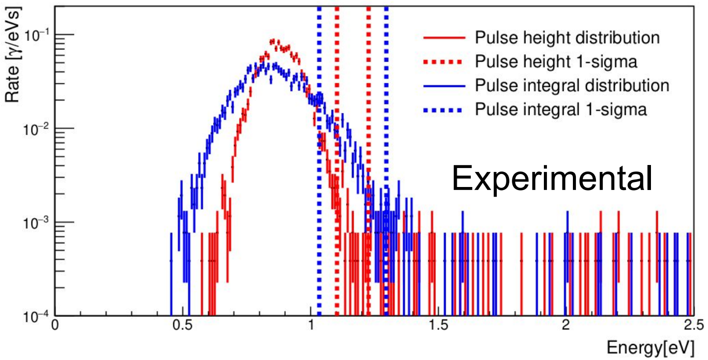


Black-body radiation

- Direct 1064 nm photons + pileups
  - Pileups reduced by
    - Fibre transmission losses
    - Fibre curling
    - TES structure
- Rate highly depends on TES resolution
- Currently minimum rate $6.9 \cdot 10^{-5}$ Hz

Target sensitivity: $1 \cdot 10^{-5}$ Hz (1 photon/day)





[1] Jose Alejandro Rubiera Gimeno, Dissertation 2024, Optimizing a Transition Edge Sensor detector system for low flux infrared photon measurements at the ALPS II experiment

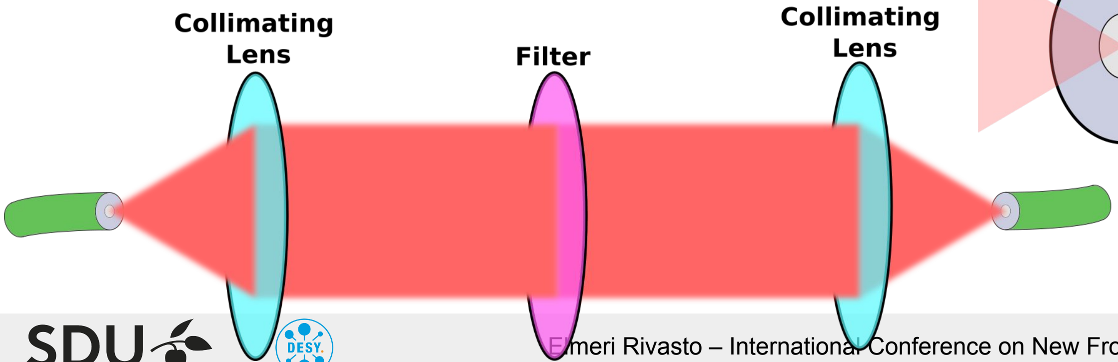
Blackbody background

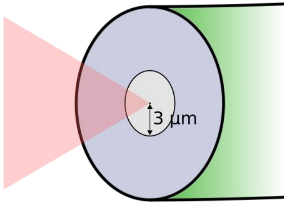
- Experimentally measured blackbody background rates for different TES resolutions:

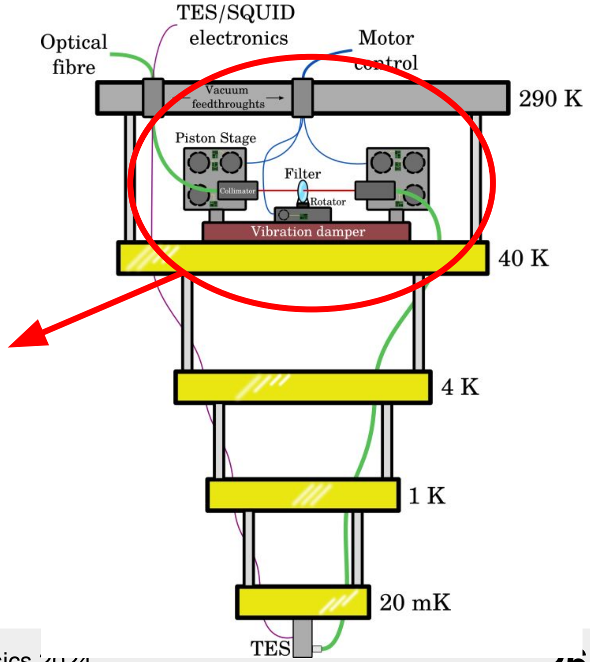
Range (σ)	Analysis efficiency	Rate (I_{Ph})	Rate (h_{FFT})
$-1, 1$	67.2%	$1.7 \cdot 10^{-3}$cps	$1.2 \cdot 10^{-4}$cps
$-2, 2$	93.9%	$5.6 \cdot 10^{-3}$cps	$4.1 \cdot 10^{-4}$cps
$-3, 3$	98.1%	$1.1 \cdot 10^{-2}$cps	$1.5 \cdot 10^{-3}$cps
$0, 3$	49.1%	$4.2 \cdot 10^{-4}$cps	$6.9 \cdot 10^{-5}$cps
$-1, 3$	82.6%	$1.9 \cdot 10^{-3}$cps	$1.6 \cdot 10^{-4}$cps

Challenges: Thermal contraction $\Delta l = \alpha \cdot l \cdot \Delta T$

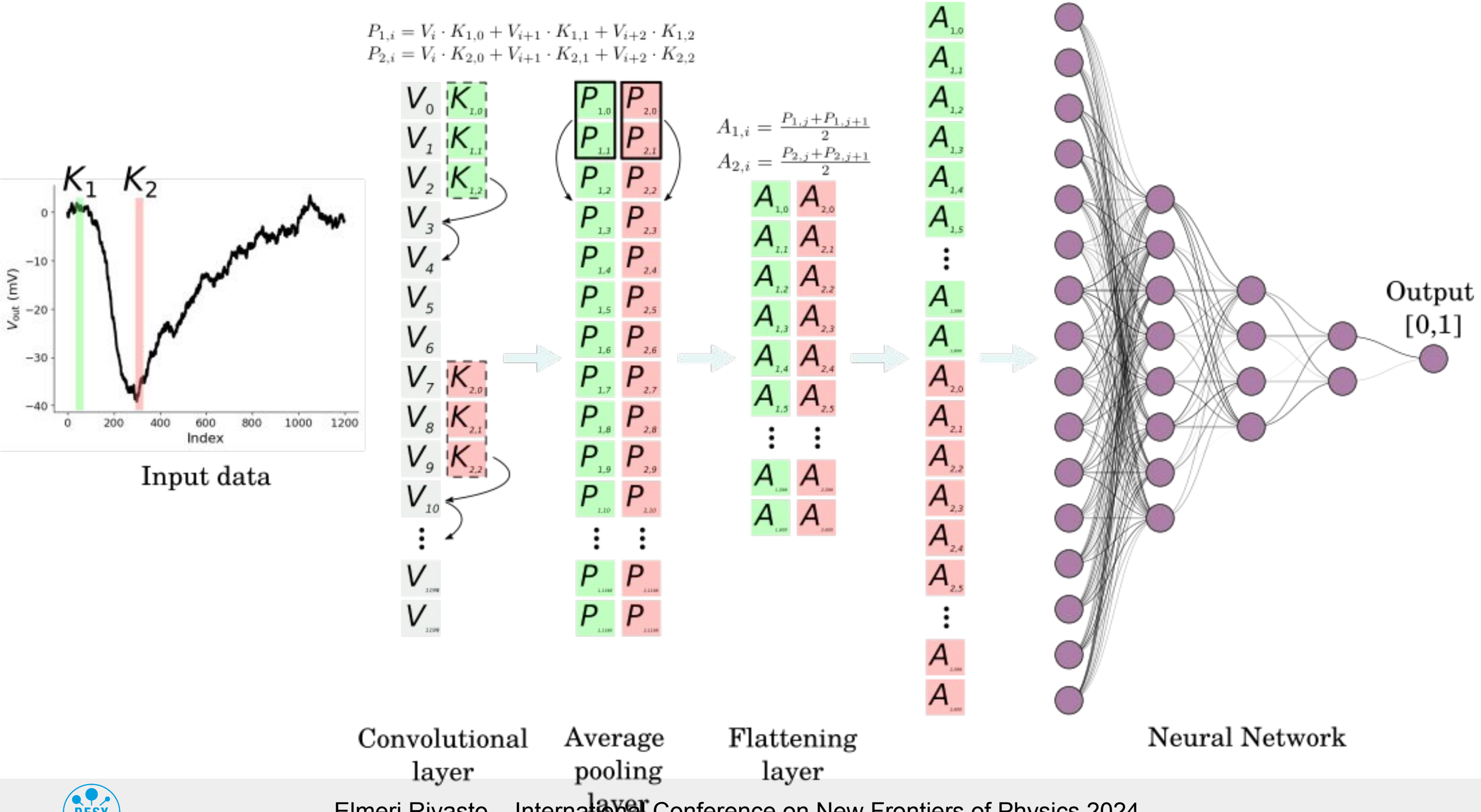
- **Invar** ($Fe_{0.64}Ni_{0.36}$) is a metal with lowest known thermal expansion coefficient
  - Good thermal conductivity
  - $\alpha = 1.2 \cdot 10^{-6}\ K^{-1}$
  - 1 cm long invar structure shrinks 3 µm
- **Realignment required after cooldown!**



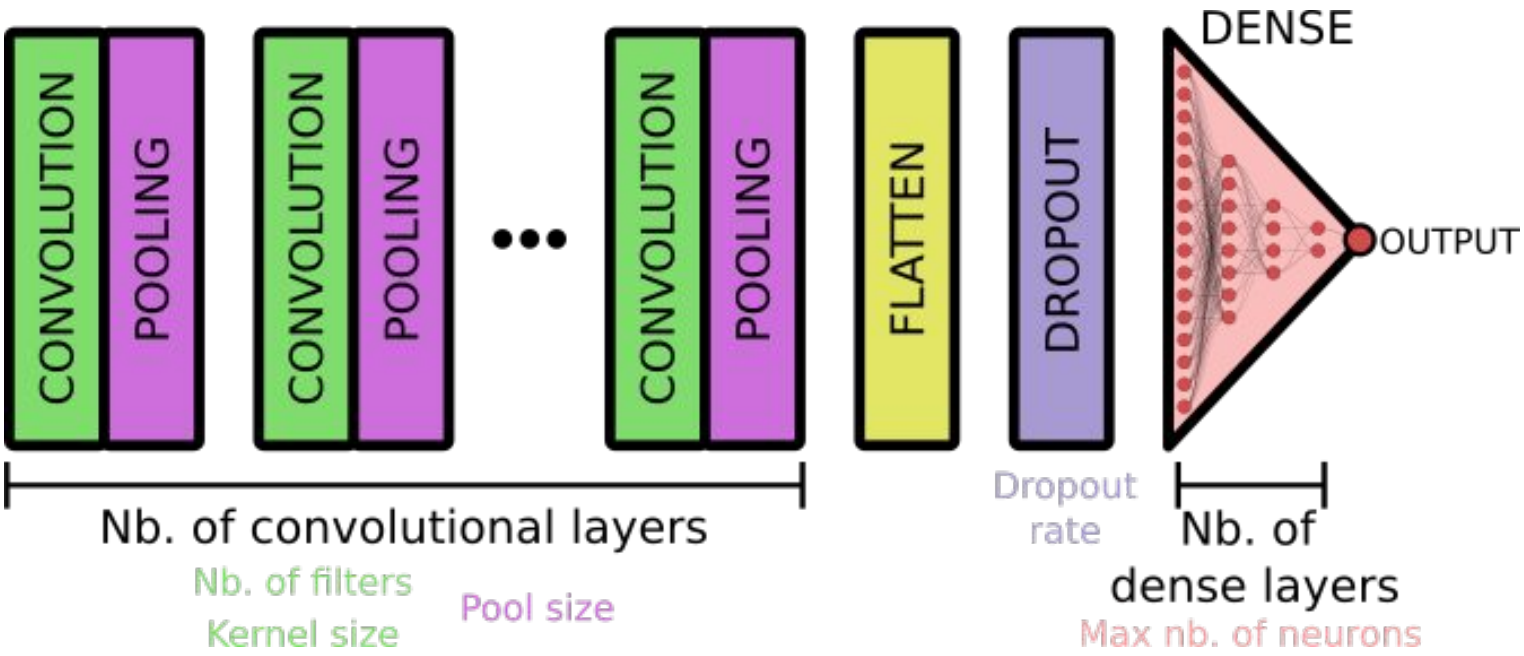




Introduction to CNNs



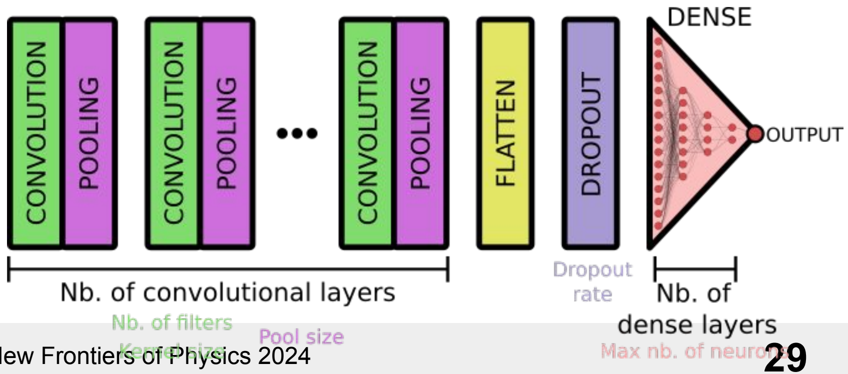
Hyperparameter optimization



Hyperparameter optimization

- Search space size ~10^{10}
- **Random Search**
  - Dataset:
    - Total 40,000 samples
      - 32,000 for training (80-20% train./val.)
      - 8,000 for testing
  - 1000 iterations
- 1 iteration takes ~5 min (~4 days total)
  - Reduced to 1 day via parallel computing!

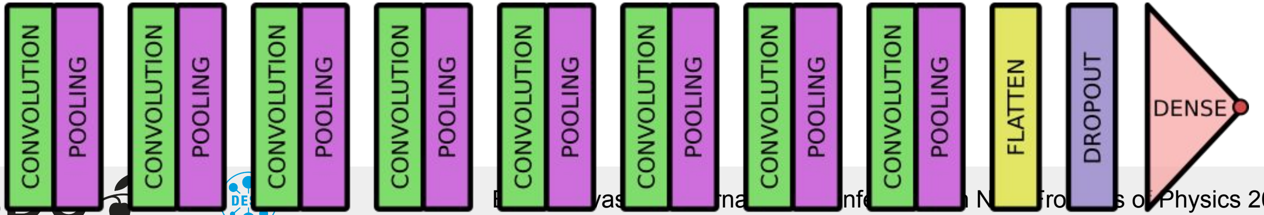
Hyperparameter	Optimization range
Nb. of conv. layers	5–10
Nb. of filters	30–70
Kernel size	5–30
Dropout rate	0–0.2
Nb. of dense layers	1–3
Max nb. of neurons	50–150
Learning rate	10^{-4}–10^{-3}
Epochs	1–30
Batch size	4–128



Optimal parameters:

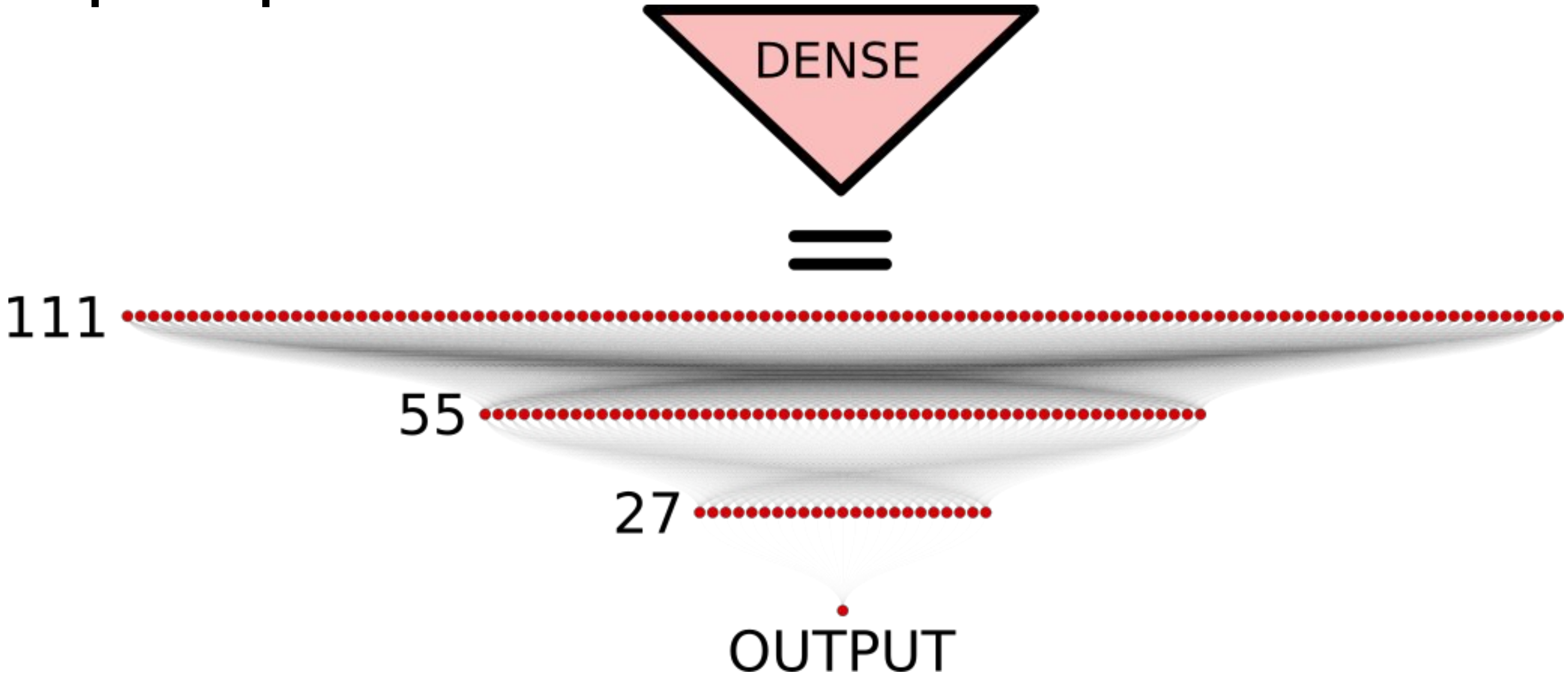
Resulting **AUC=0.9954 ± 0.0004**

Hyperparameter	Optimization range	Found optimum
Nb. of conv. layers	5–10	8
Nb. of filters	30–70	52
Kernel size	5–30	19
Dropout rate	0–0.2	0.122
Nb. of dense layers	1–3	3
Max nb. of neurons	50–150	111
Learning rate	10^{-4}–10^{-3}	10^{-4}
Epochs	1–30	20
Batch size	4–128	126



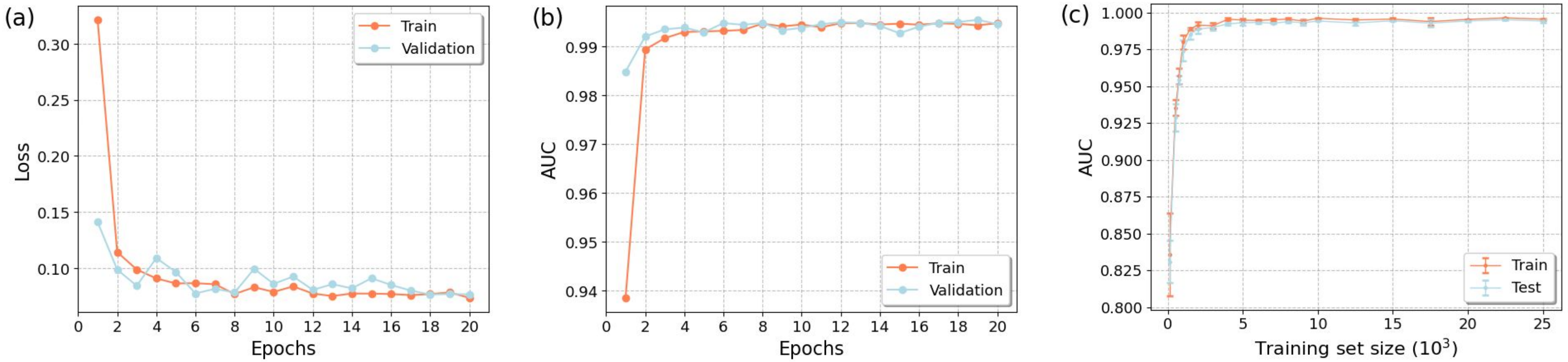

```
Model: "sequential_2"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d_16 (Conv1D)          (None, 1200, 52)          1040

 average_pooling1d_16 (Avera  (None, 600, 52)          0
 gePooling1D)

 conv1d_17 (Conv1D)          (None, 600, 52)           51428

 average_pooling1d_17 (Avera  (None, 300, 52)          0
 gePooling1D)

 conv1d_18 (Conv1D)          (None, 300, 52)           51428

 average_pooling1d_18 (Avera  (None, 150, 52)          0
 gePooling1D)

 conv1d_19 (Conv1D)          (None, 150, 52)           51428

 average_pooling1d_19 (Avera  (None, 75, 52)           0
 gePooling1D)

 conv1d_20 (Conv1D)          (None, 75, 52)            51428

 average_pooling1d_20 (Avera  (None, 37, 52)           0
 gePooling1D)

 conv1d_21 (Conv1D)          (None, 37, 52)            51428

 average_pooling1d_21 (Avera  (None, 18, 52)           0
 gePooling1D)

 conv1d_22 (Conv1D)          (None, 18, 52)            51428

 average_pooling1d_22 (Avera  (None, 9, 52)            0
 gePooling1D)

 conv1d_23 (Conv1D)          (None, 9, 52)             51428

 average_pooling1d_23 (Avera  (None, 4, 52)            0
 gePooling1D)

 flatten_2 (Flatten)         (None, 208)               0

 dropout_2 (Dropout)         (None, 208)               0

 dense_8 (Dense)             (None, 111)               23199

 dense_9 (Dense)             (None, 55)                6160

 dense_10 (Dense)            (None, 27)                1512

 dense_11 (Dense)            (None, 1)                 28

=================================================================
Total params: 391,935
Trainable params: 391,935
Non-trainable params: 0
_________________________________________________________________
```

Optimal parameters:



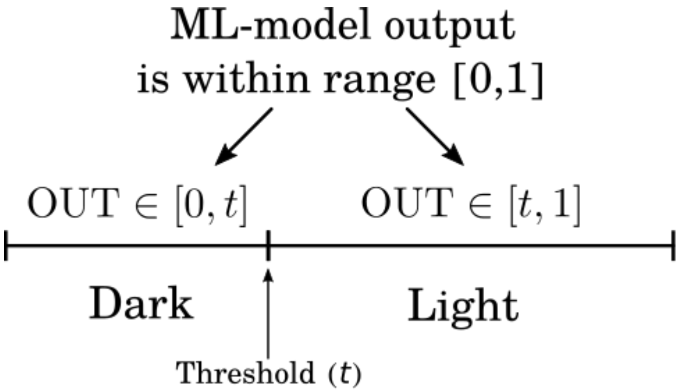
Checking for overfitting - Memorization VS Generalization

- Very well balanced model
    - Negligible generalization gap
        - No over/underfitting observed
- Learning saturates already for training set size of ~5,000



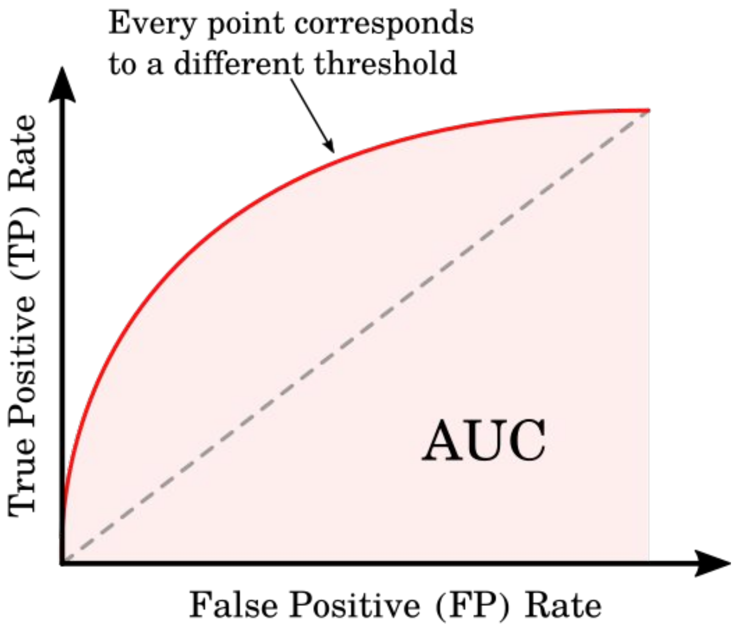
Machine Learning: Metrics

- **A**rea **U**nder the Receiver Operating Characteristics (ROC) **C**urve **(AUC)**



$$\mathrm{TPR} = \frac{\mathrm{TPs}}{\mathrm{TPs}+\mathrm{FNs}}$$

$$\mathrm{FPR} = \frac{\mathrm{FPs}}{\mathrm{FPs}+\mathrm{TPs}}$$



Determining optimal threshold

- Determine optimal threshold based on max **accuracy** (~0.974)