ENHANCING SDD ENERGY RESPONSE WITH ML AND DIFFERENTIAL PROGRAMMING

Fabrizio Napolitano for the VIP Collaboration

INFN
Istituto Nazionale
di Fisica Nucleare

John
Templeton
Foundation

CENTRO
FERMI
Enrico Fermi

MUSEO
STORICO DELLA FISICA
E
CENTRO
STUDI E RICERCHE
ENRICO FERMI

SMI

FWF
Der Wissenschaftsfonds.

TEQ

FQXi
FOUNDATIONAL QUESTIONS INSTITUTE

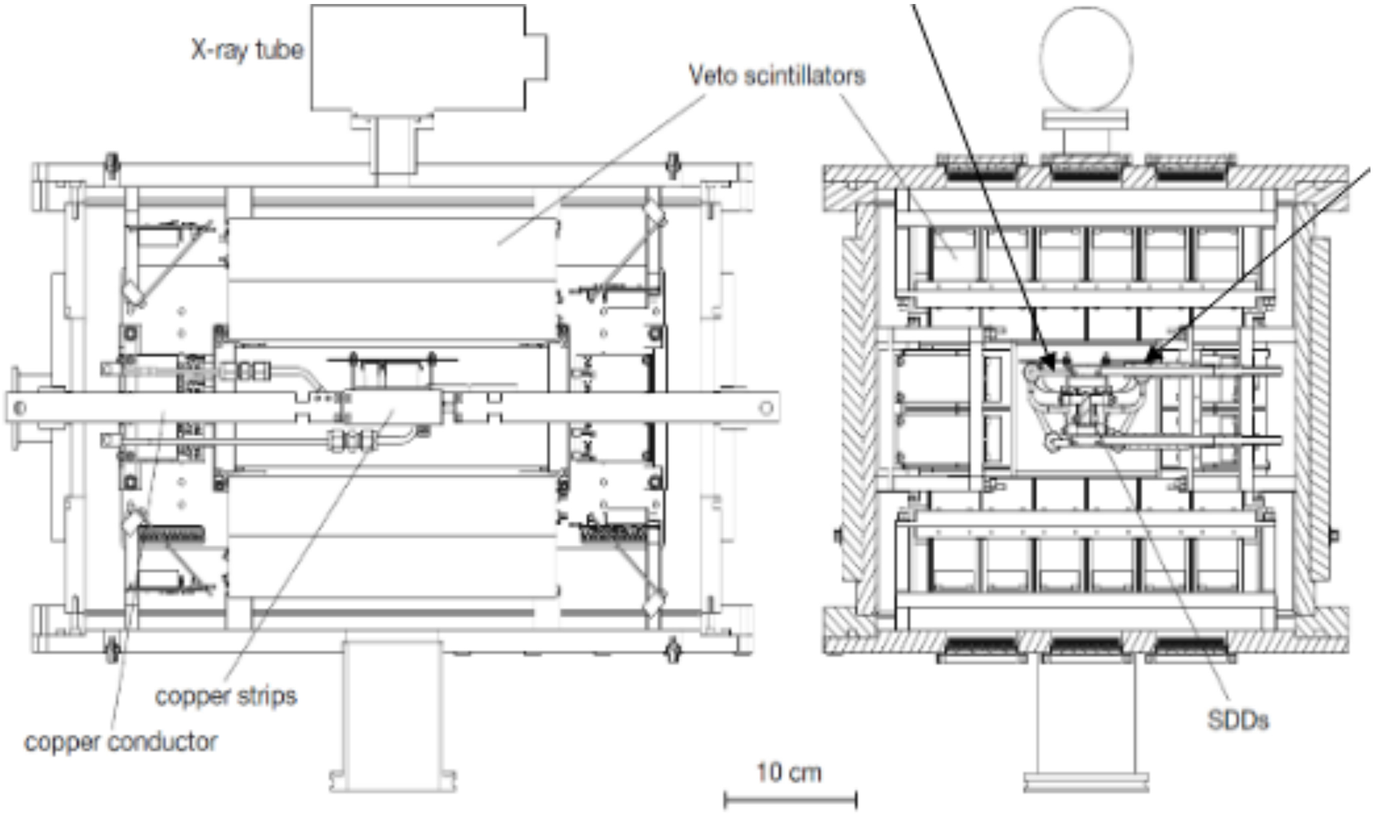
fabrizio.napolitano@lnf.infn.it

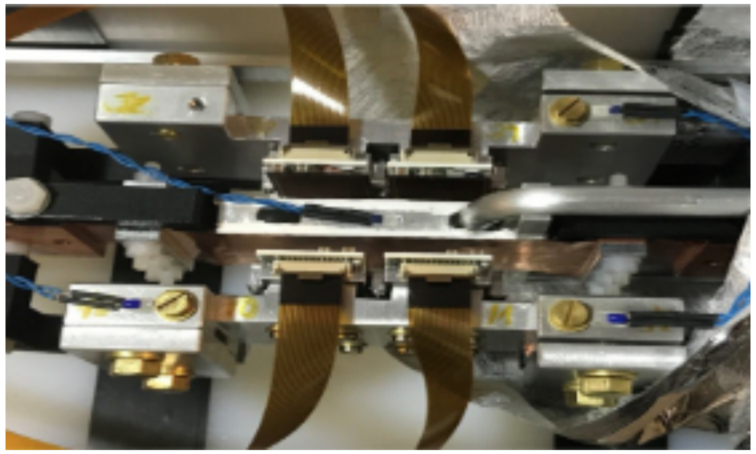
HPXM2023

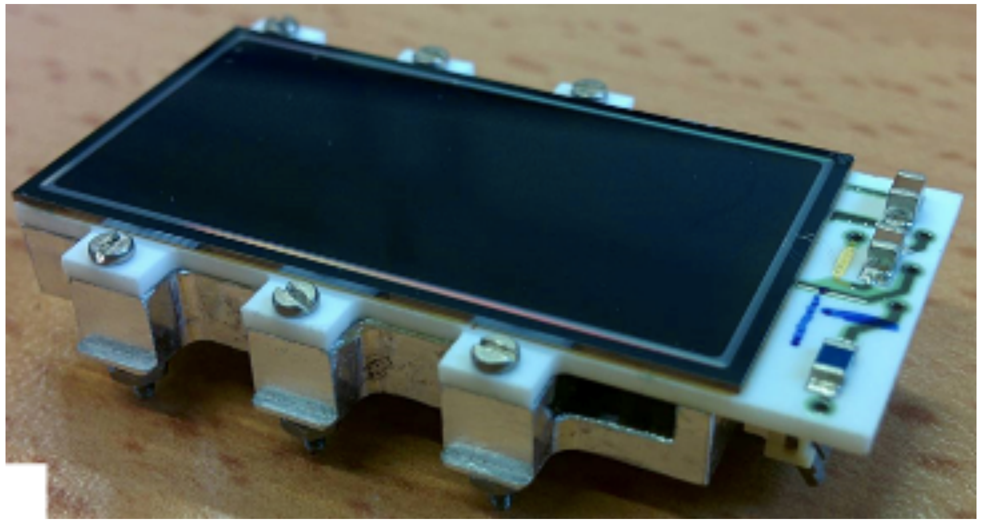
June 19th to 23rd at Laboratori Nazionali di Frascati (LNF - INFN), Italy

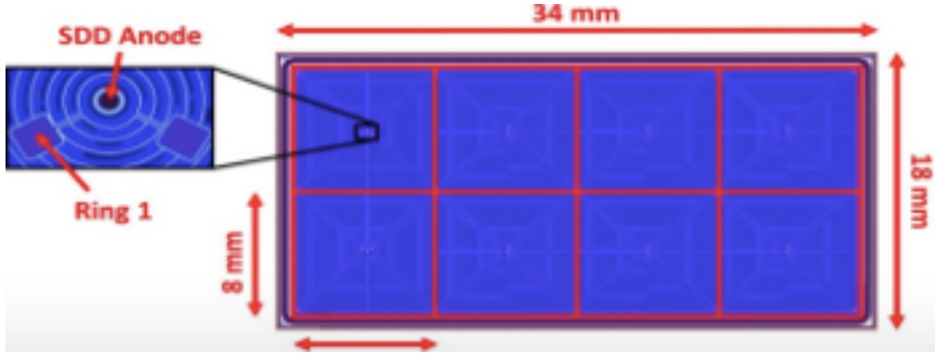
The VIP-2 Experiment

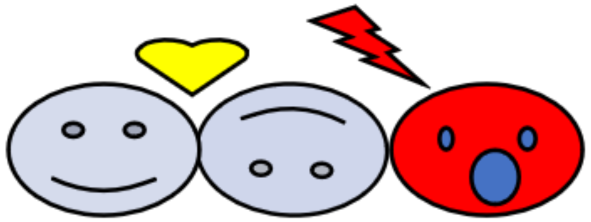
*Silicon Drift Detectors (**SDDs**) higher resolution (190 eV FWHM at 8.0 →keV), faster (triggerable) detectors. 4 arrays of 2 x 4 SDDs 8mm x 8mm each, liquid argon closed circuit cooling 170 °C*

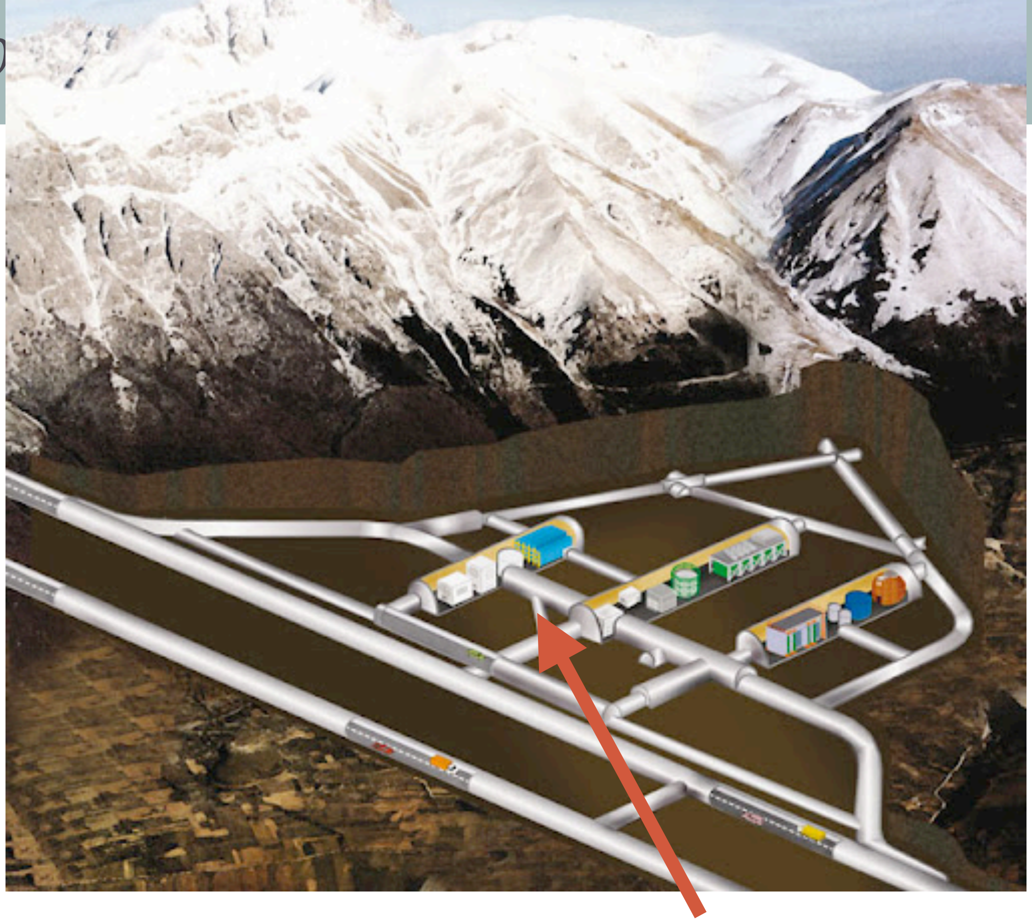






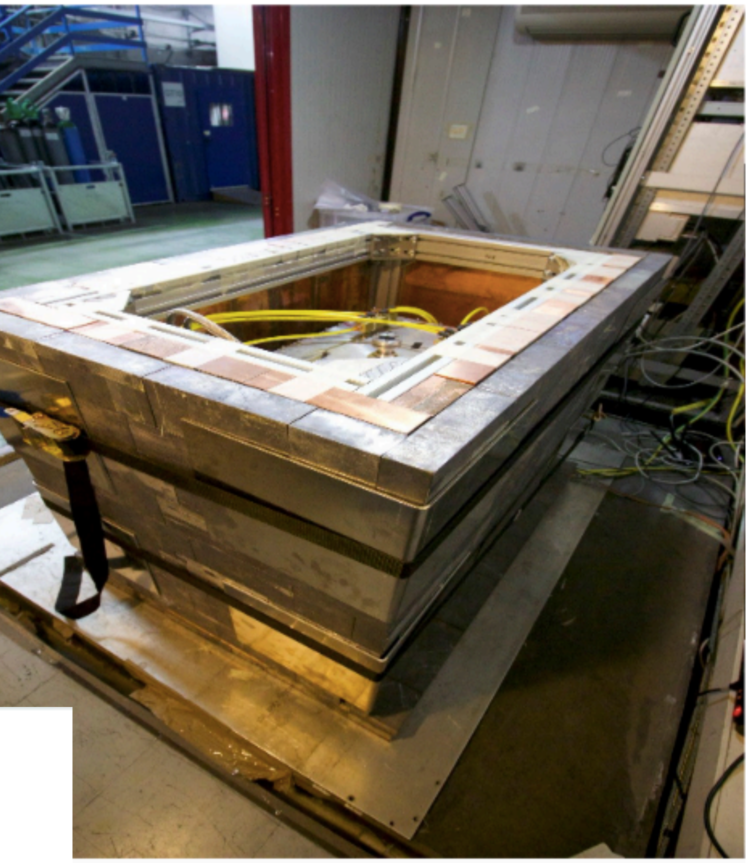




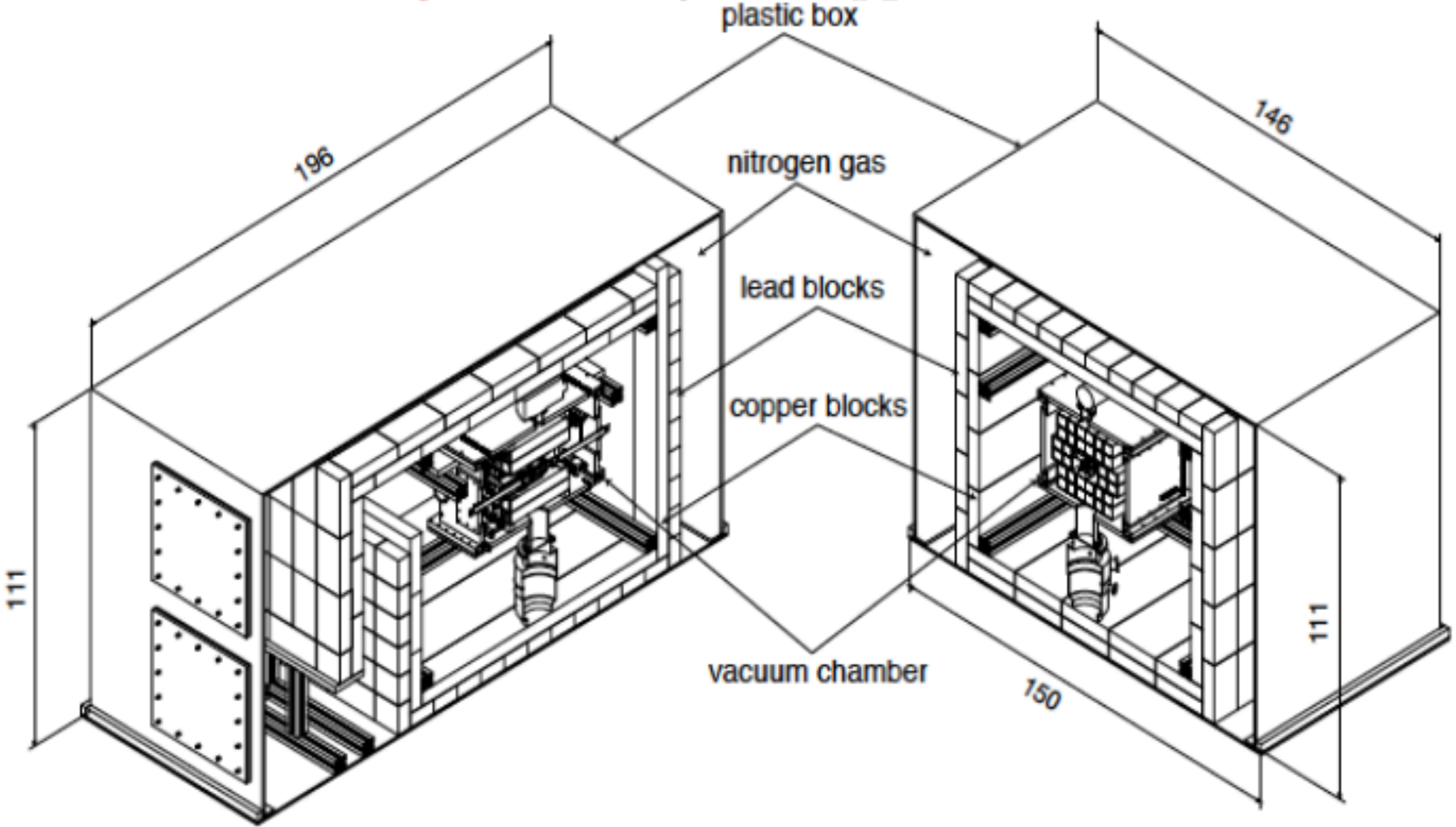


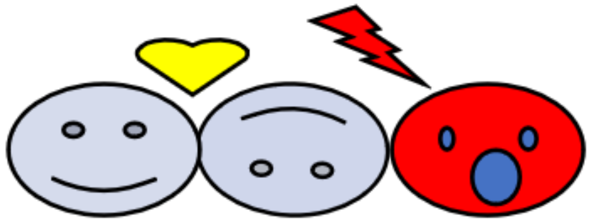
1400 m rock coverage

Upgrade concluded in April 2019:



Passive scielding → two layers, copper inside lead outside





1400 m rock coverage

Upgrade concluded in April 2019:

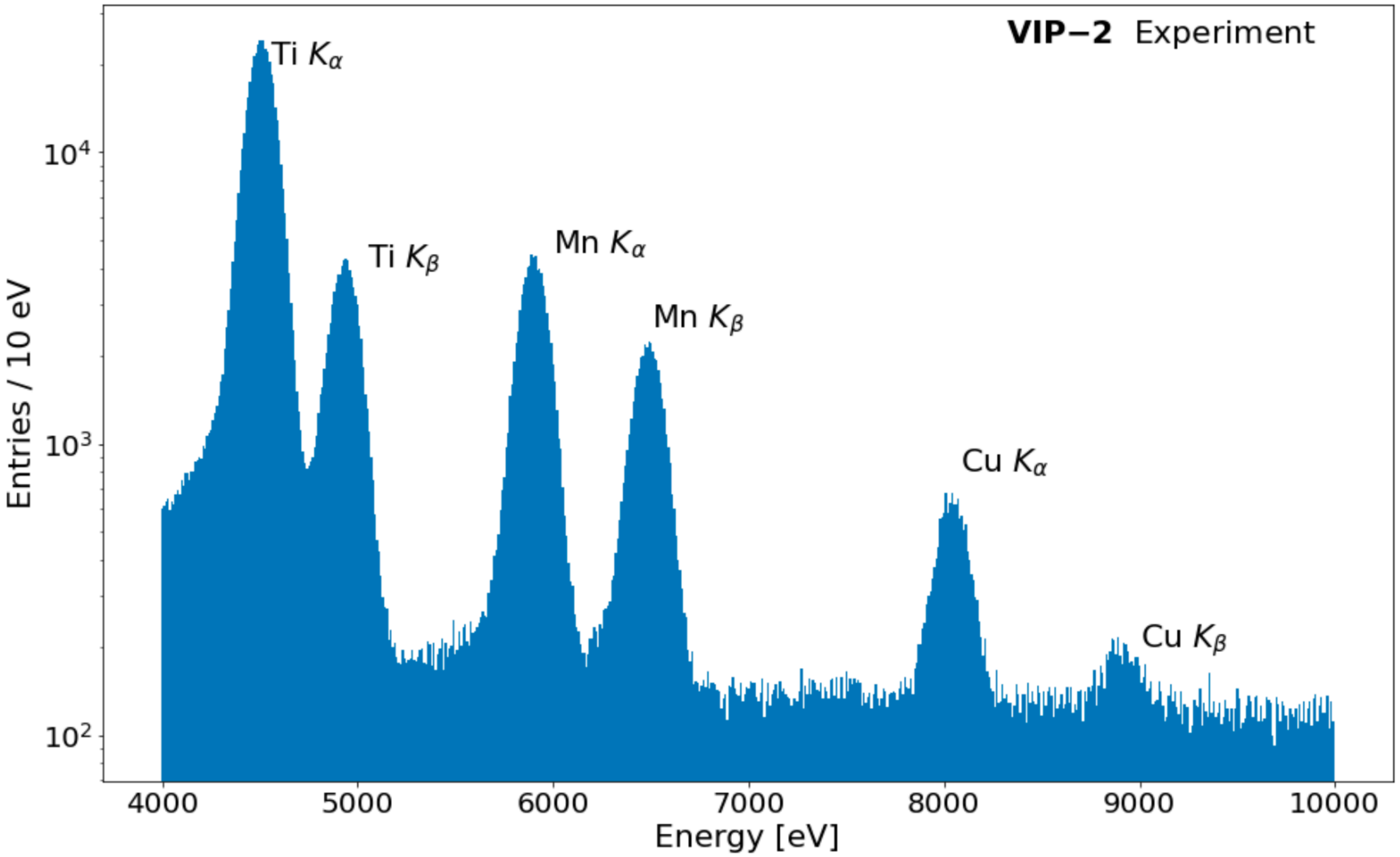
Ideal place for precision X-ray spectroscopy

Searching for Pauli Exclusion Principle Violations

Quantum Gravity

… and more: see next talks!

Passive shielding → two layers, copper inside lead outside



Calibrated spectrum of 4 SDD arrays.

Not easy to calibrate because:

- *Copper line at orders of magnitude smaller than Ti and Mn*
- *Tiny distortions of FEE*

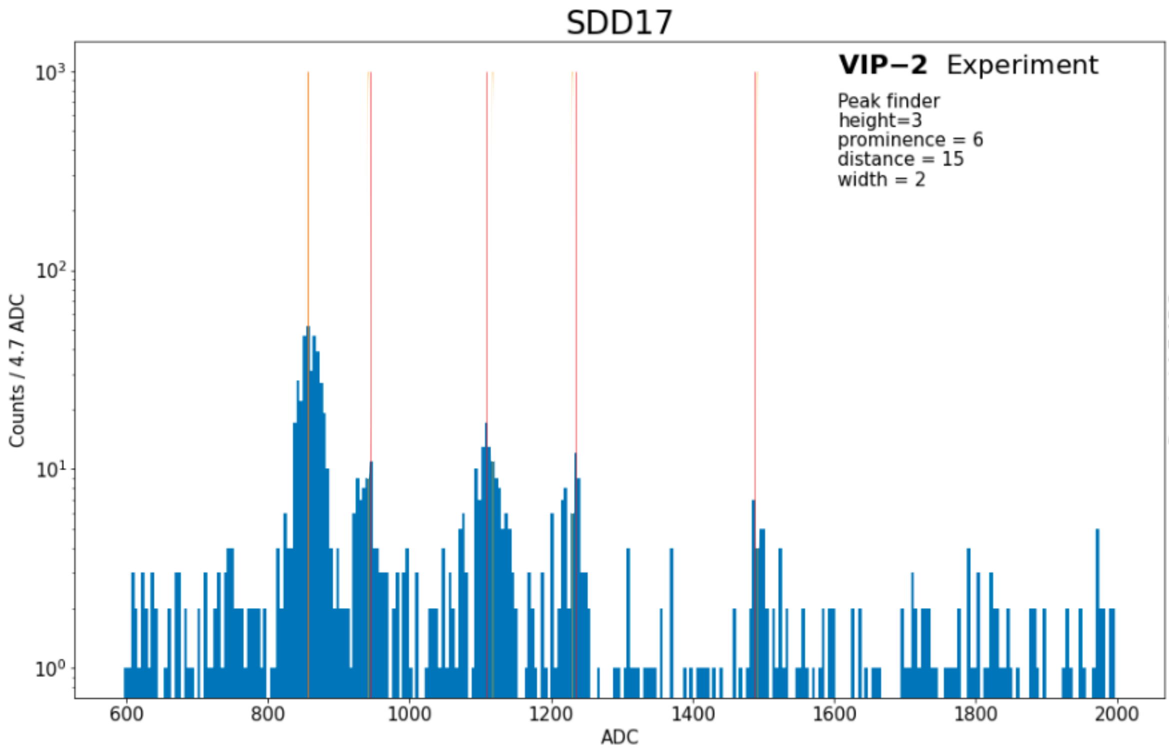
Calibration can be done in big or small batches

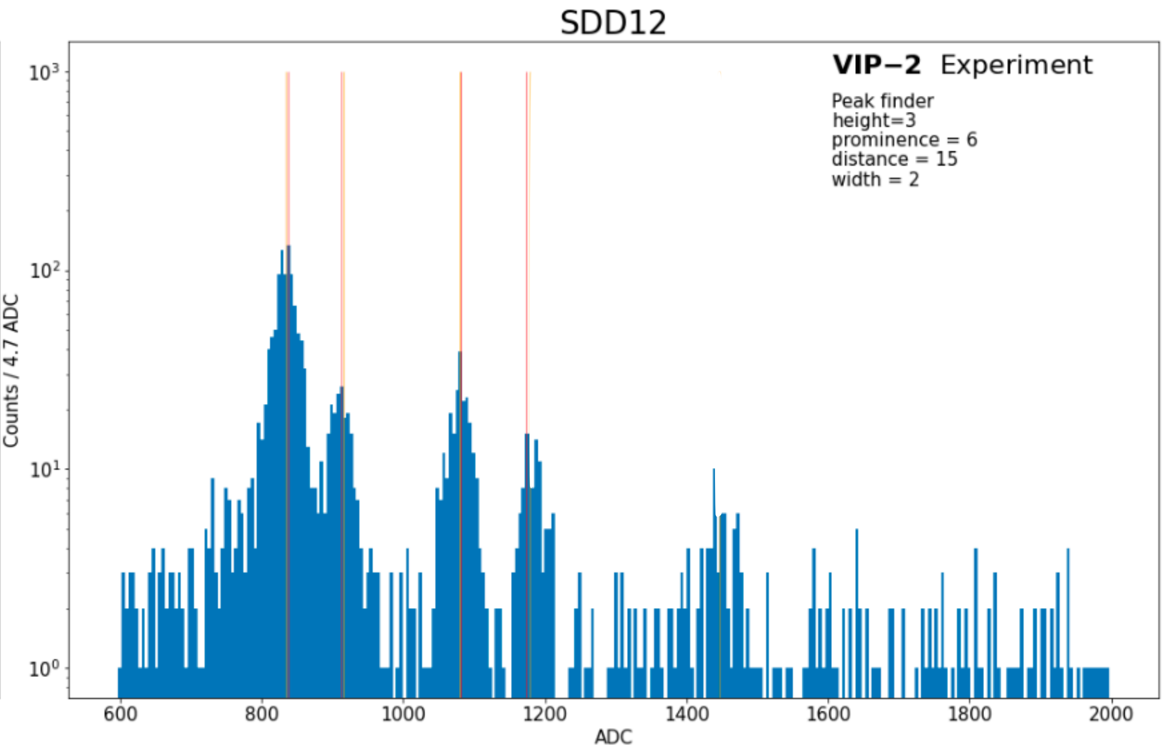
Big batches

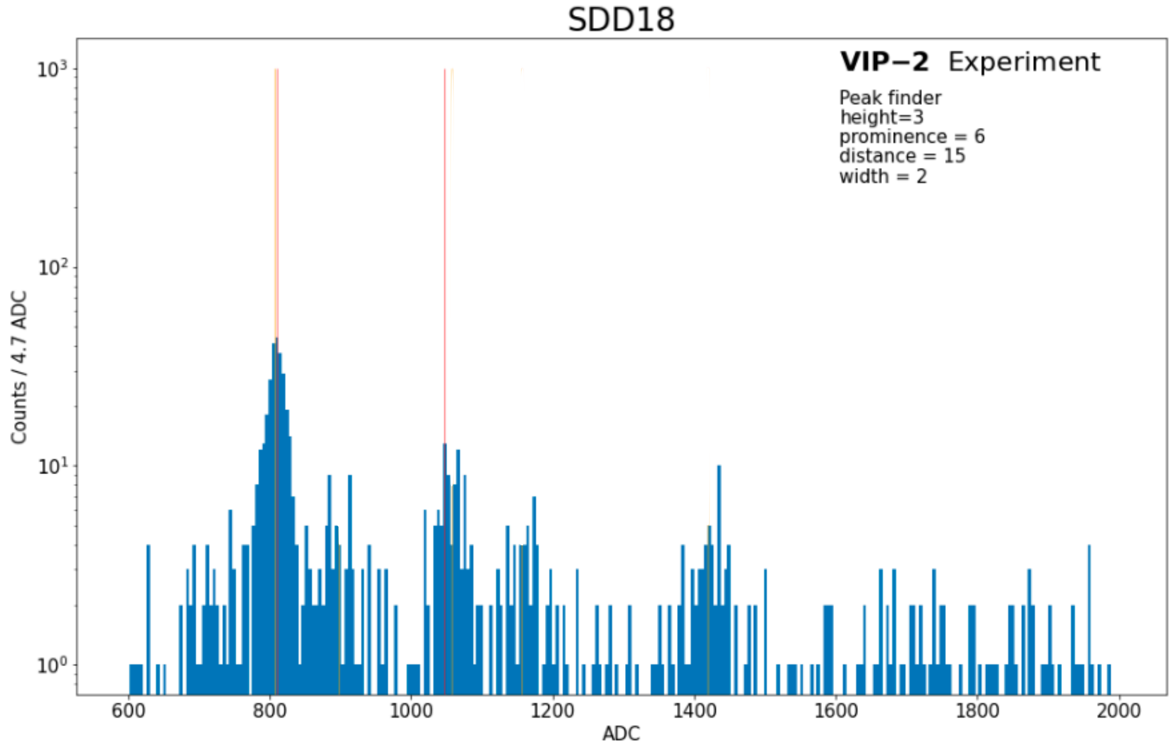
Can determine better the Copper position but cannot capture fluctuations

Small batches

Can capture fluctuations but cannot determine the Copper position well



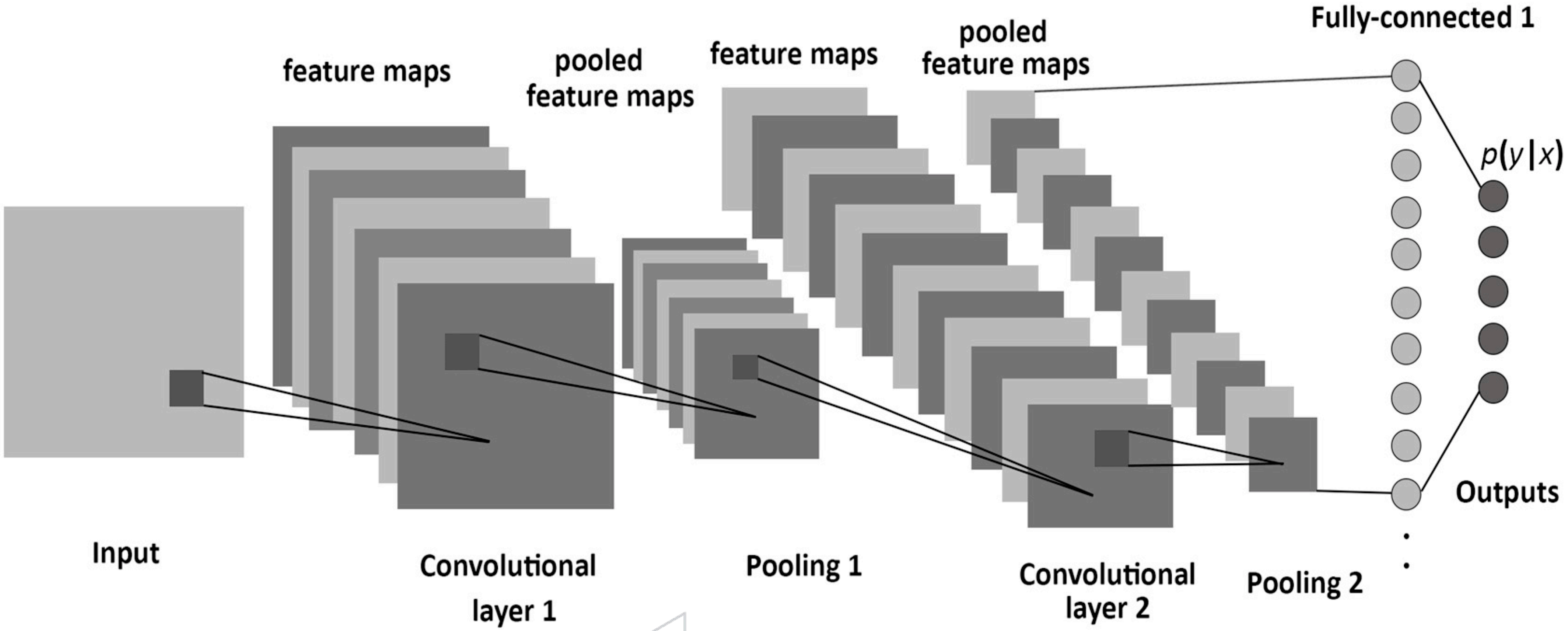


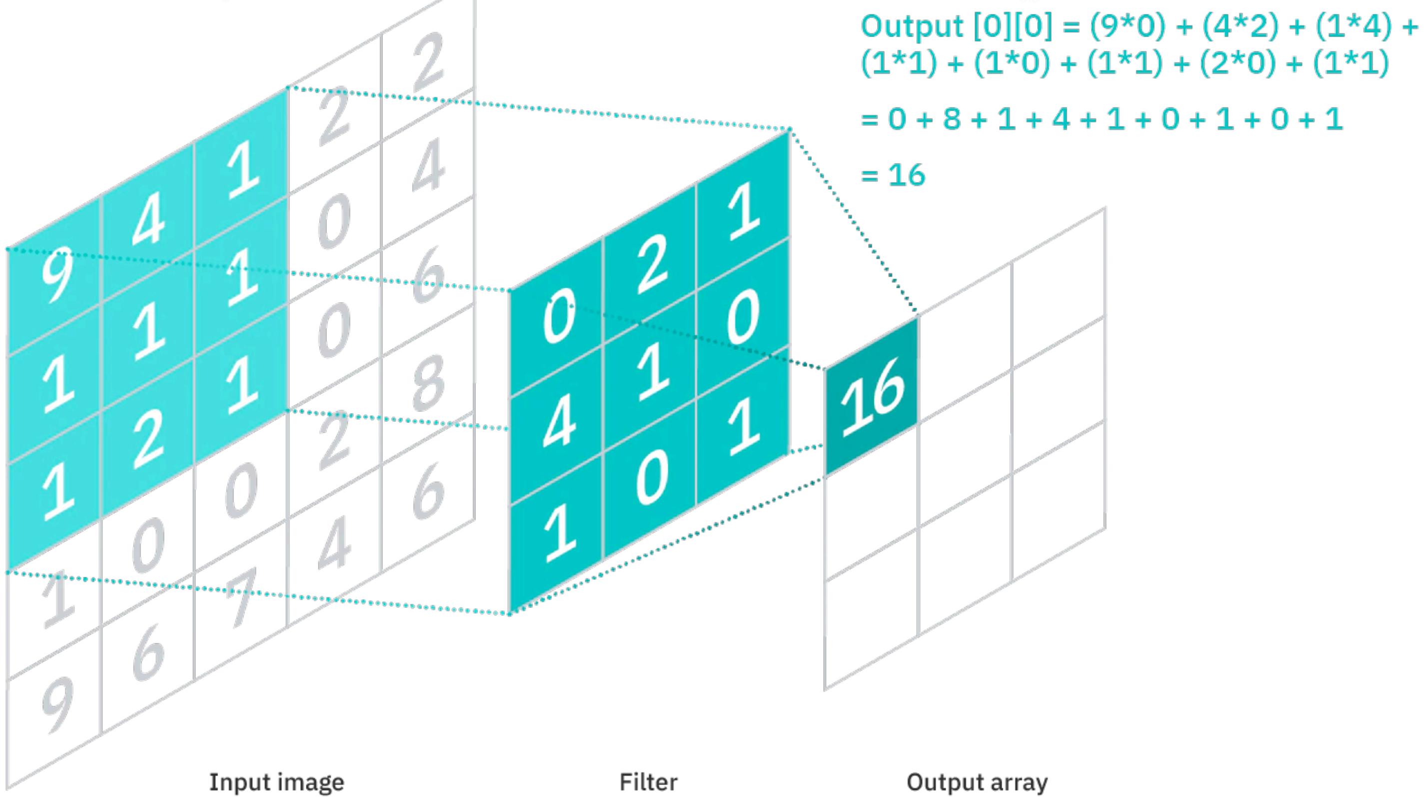


Statistical fluctuations at low counts can make the use of peak finder algorithm tricky to setup needs constant care calibrated to be resilient algo params need to be tuned

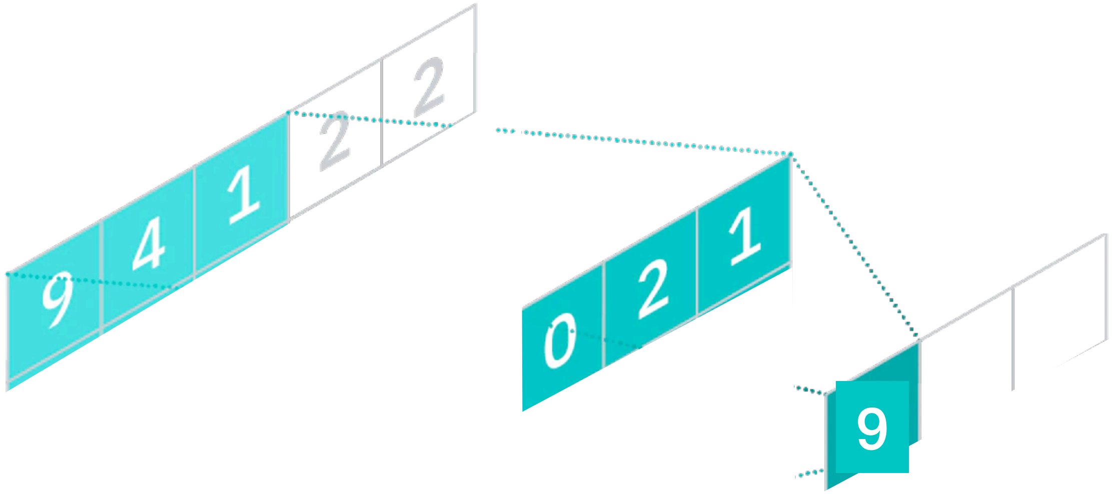
Use two step approach - 1st: convolutional neural network as peak finder

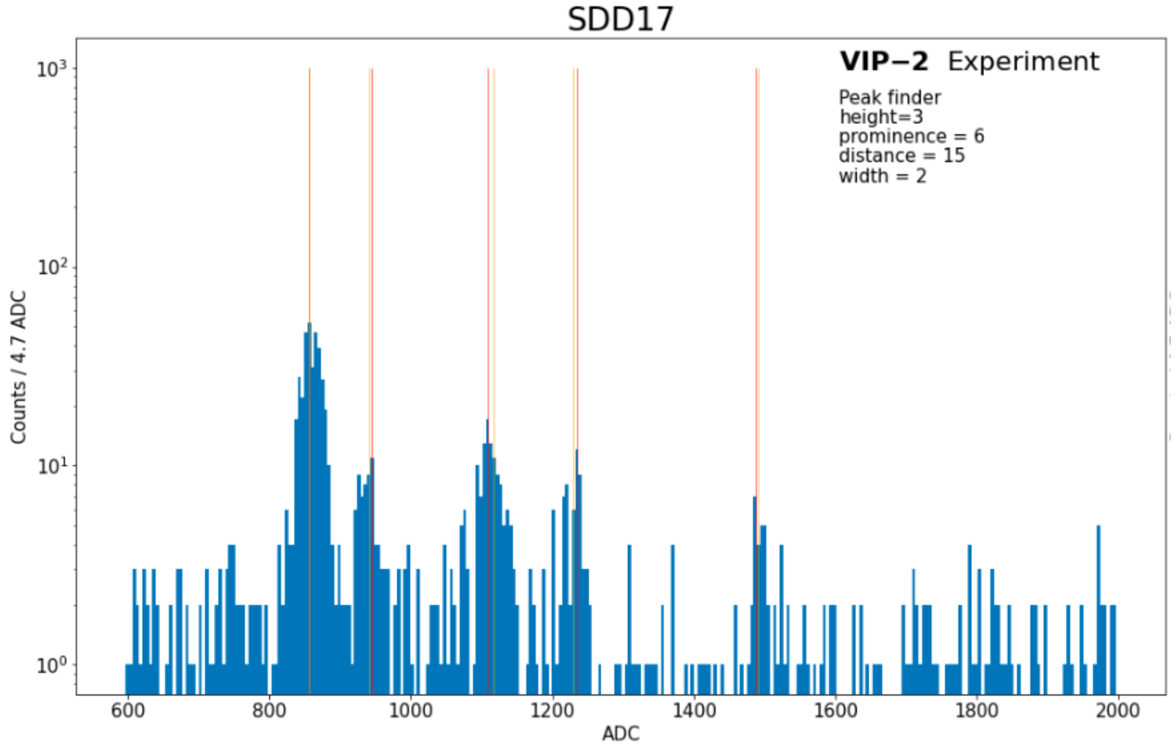
https://www.mdpi.com/1099-4300/19/6/242

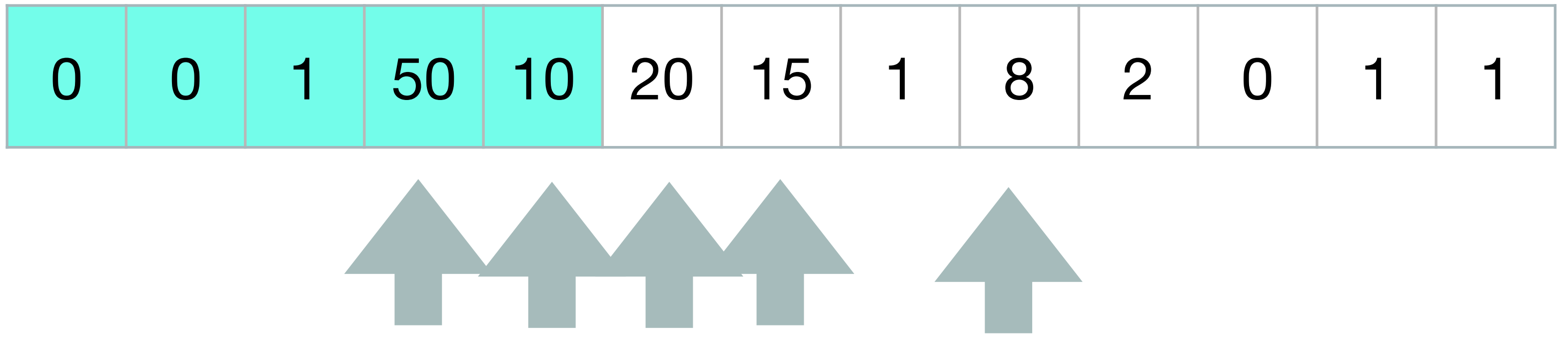




Use two step approach - 1st: convolutional neural network as peak finder

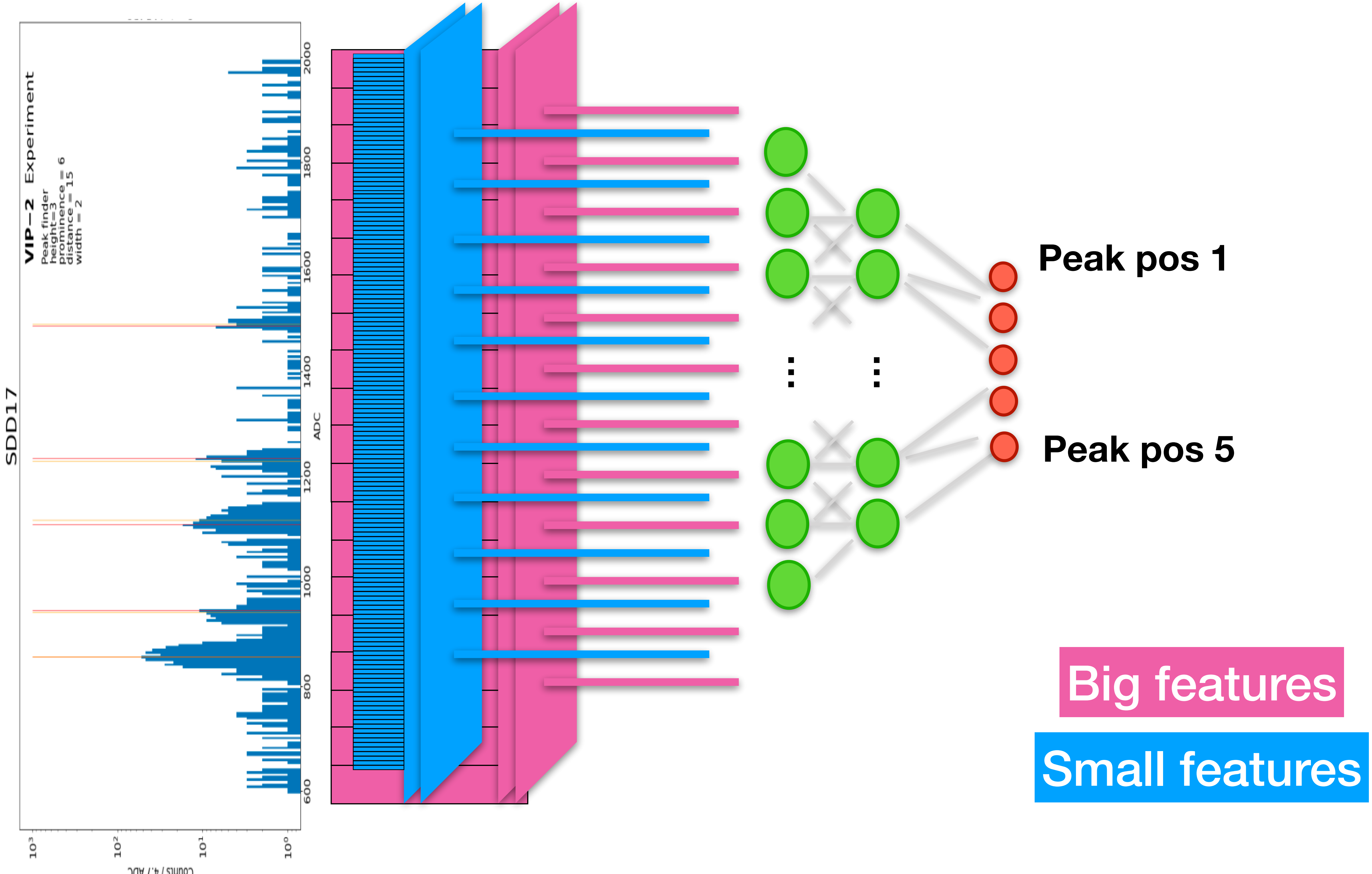






0	0	1	50	10	20	15	1	8	2	0	1	1

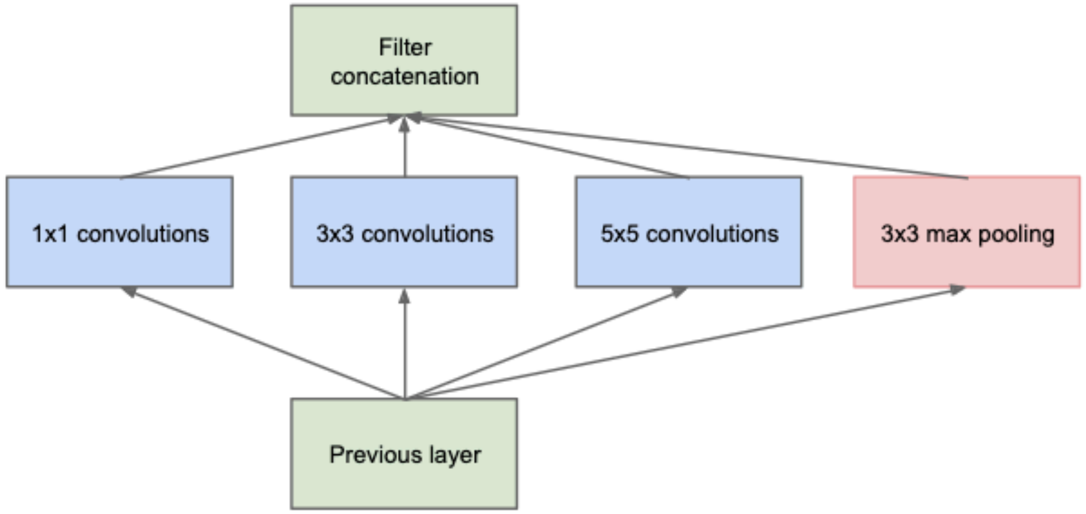
Use two step approach - 1st: convolutional neural network as peak finder



Use two step approach - 1st: convolutional neural network as peak finder

https://arxiv.org/pdf/1409.4842.pdf

The "Inception" architecture



Filter concatenation

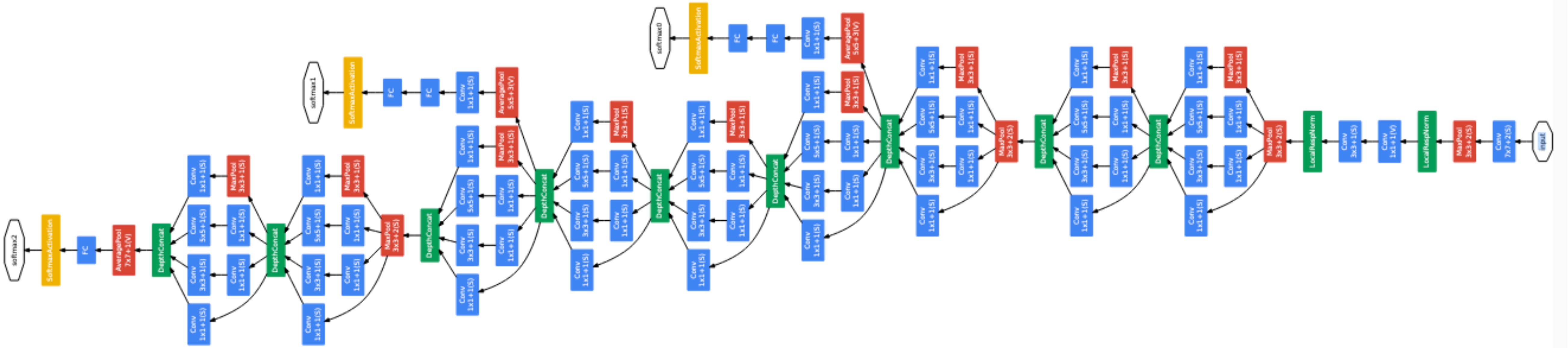
1x1 convolutions

3x3 convolutions

5x5 convolutions

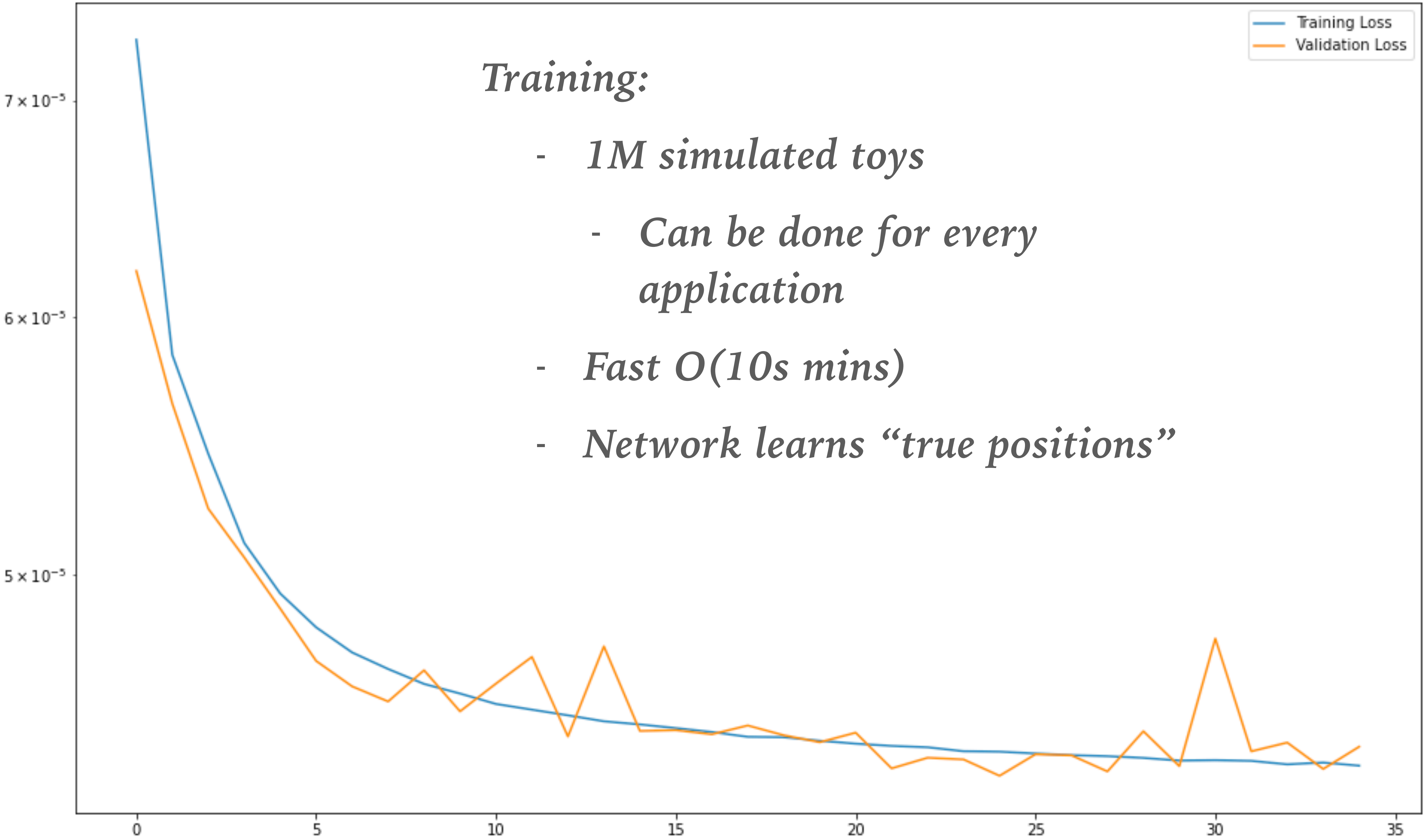
3x3 max pooling

Previous layer



GoogLeNet network

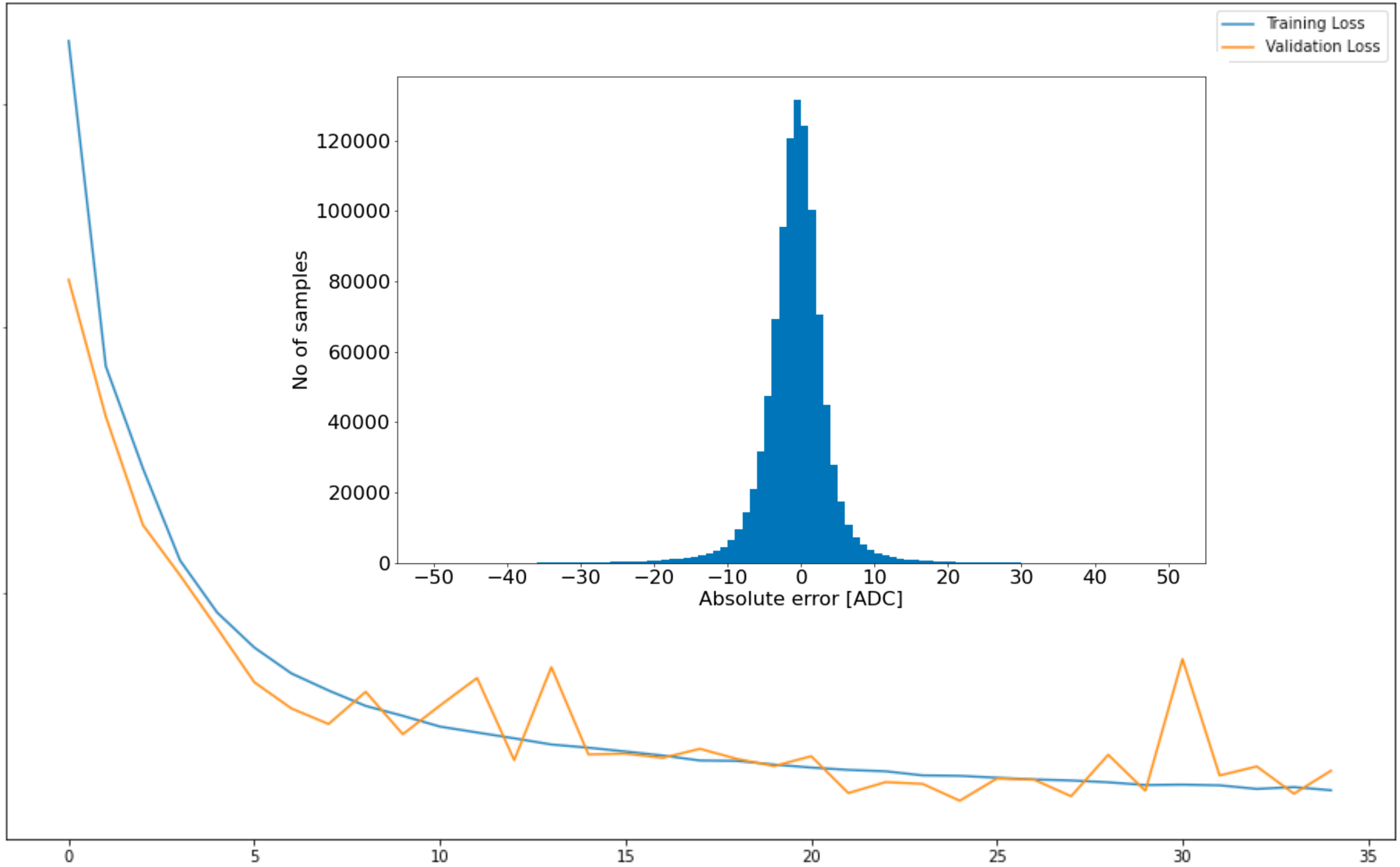
Use two step approach - 1st: convolutional neural network as peak finder



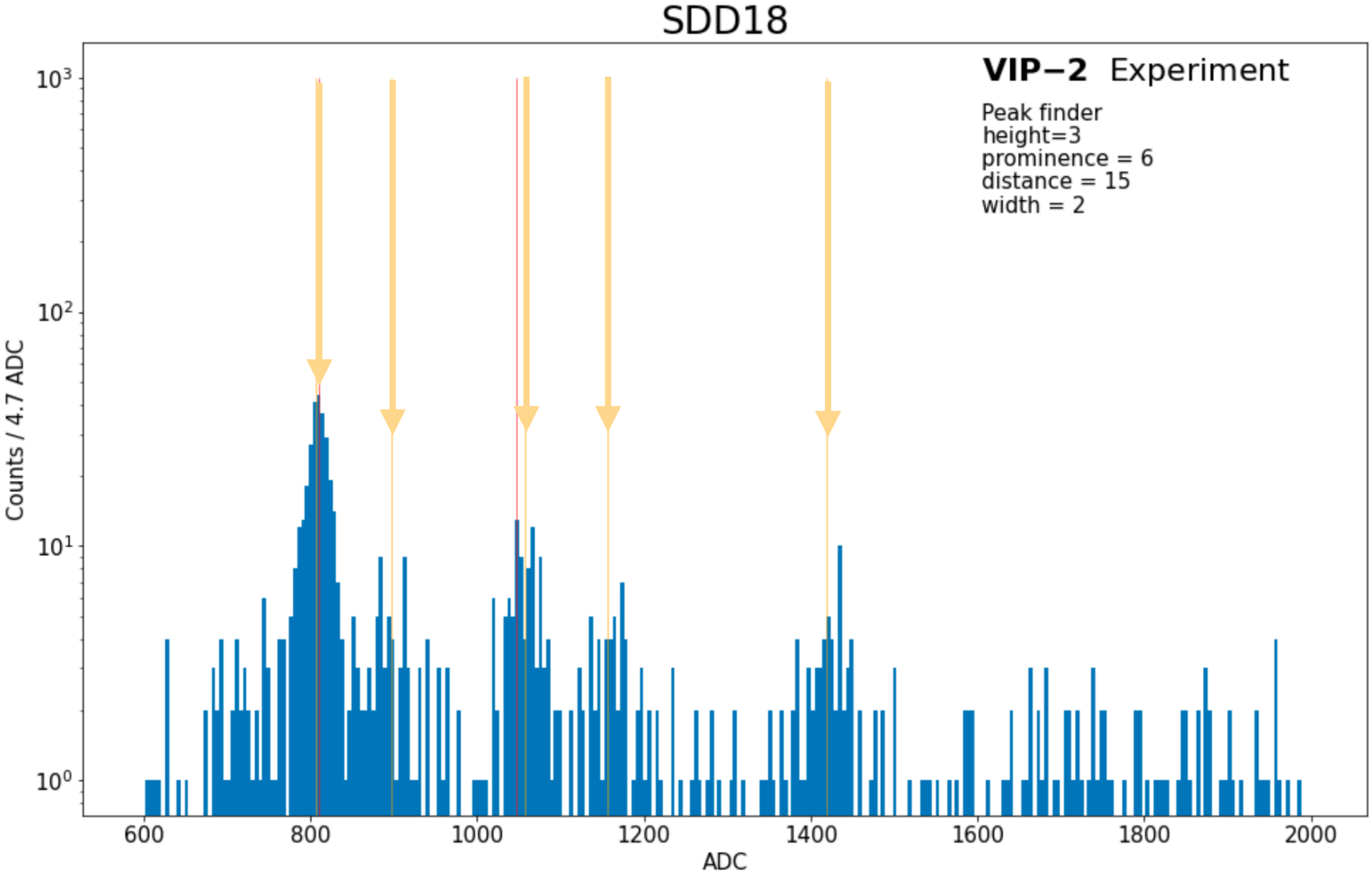
Training:

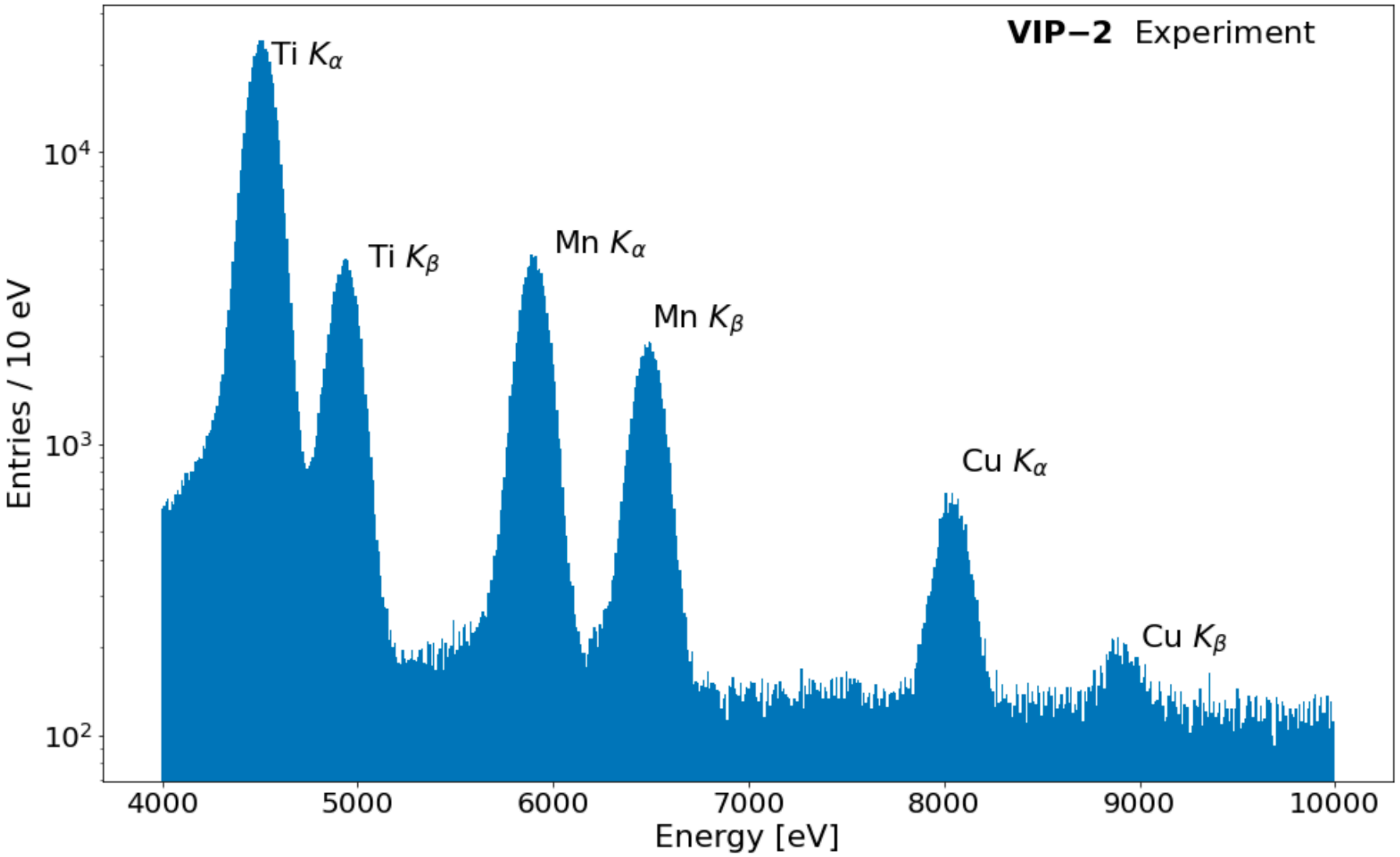
- *1M simulated toys*
    - *Can be done for every application*
- *Fast O(10s mins)*
- *Network learns “true positions”*

Use two step approach - 1st: convolutional neural network as peak finder



Use two step approach - 1st: convolutional neural network as peak finder





Calibrated spectrum of 4 SDD arrays.

Not easy to calibrate because:

- *Copper line at orders of magnitude smaller than Ti and Mn*
- *Tiny distortions of FEE*

Calibration can be done in big or small batches

Big batches

Can determine better the Copper position <u>but</u> cannot capture fluctuations

<u>Small batches</u>

Can capture fluctuations <u>but</u> cannot determine the Copper position well

Use two step approach - 2nd: to the limit with differentiable programming

Machine Learning

~~THE KING~~ IS DEAD

FEBRUARY

THE HALIFAX CHRONICLE-HERALD

The Voice of the Maritimes — The Choice of the Maritimes

WEATHER FORECAST
HALIFAX REGION
Cloudy with widely scattered showers. High today 40 degrees.

VOLUME 4, NUMBER 32
HALIFAX, CANADA, WEDNESDAY, FEBRUARY 6, 1952
PRICE FIVE CENTS—20 PAGES

-- LONG LIVE THE ~~QUEEN~~ --

Differentiable Programming

Use two step approach - 2nd: to the limit with differentiable programming

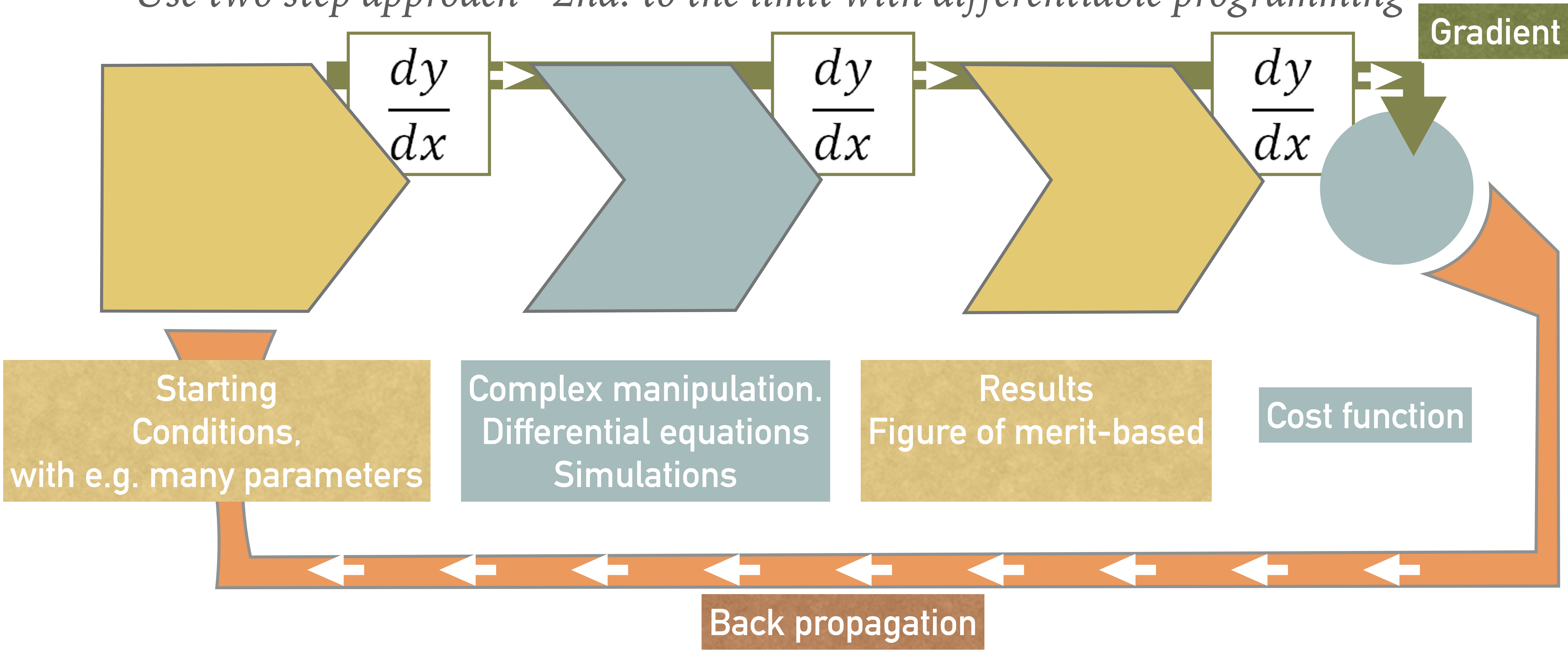
Starting
Conditions,
with e.g. many parameters

Complex manipulation.
Differential equations
Simulations

Results
Figure of merit-based

Cost function

Use two step approach - 2nd: to the limit with differentiable programming



Starting
Conditions,
with e.g. many parameters

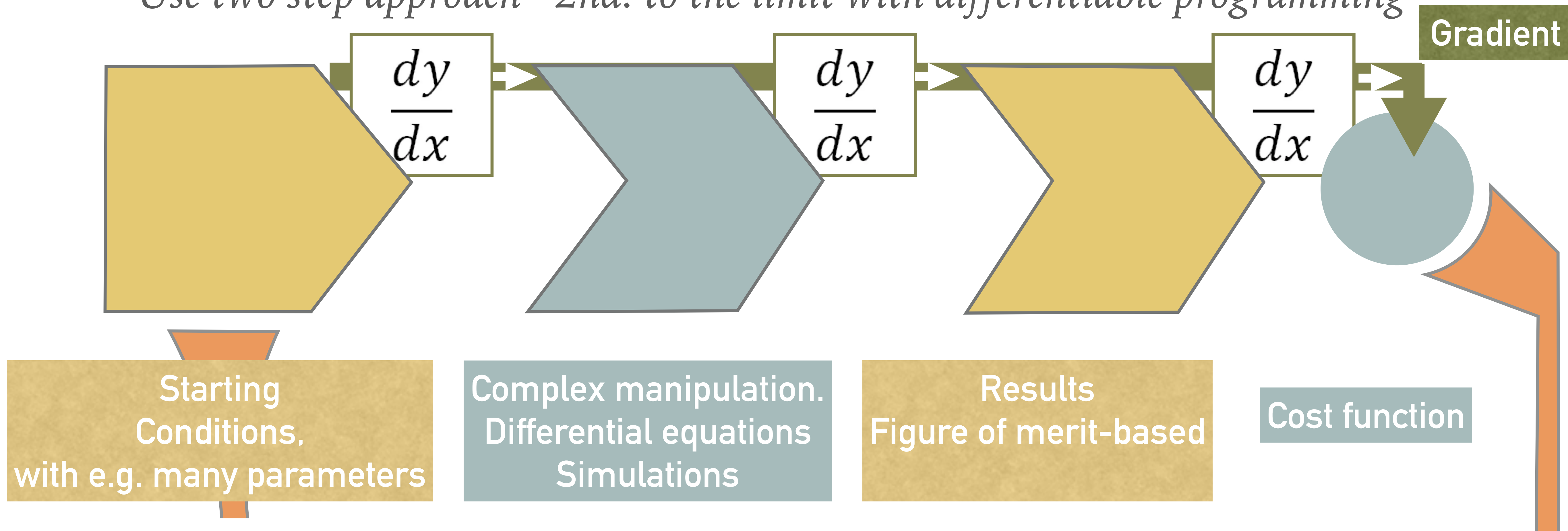
Complex manipulation.
Differential equations
Simulations

Results
Figure of merit-based

Cost function

Back propagation

Use two step approach - 2nd: to the limit with differentiable programming



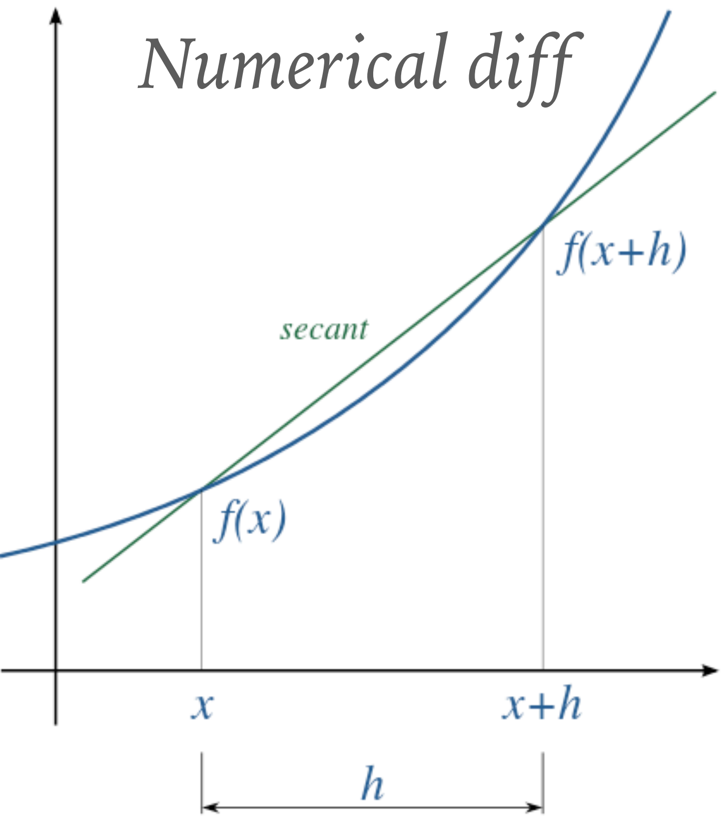
Starting
Conditions,
with e.g. many parameters

Complex manipulation.
Differential equations
Simulations

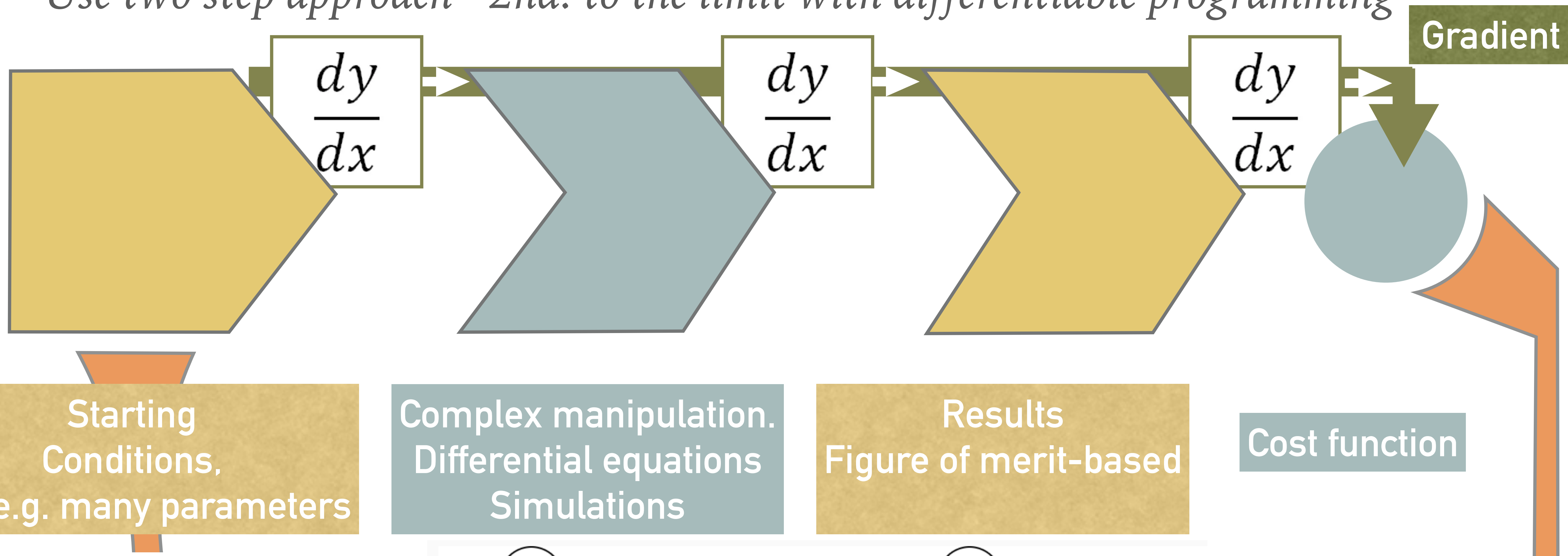
Results
Figure of merit-based

Cost function

Back propagation



Use two step approach - 2nd: to the limit with differentiable programming

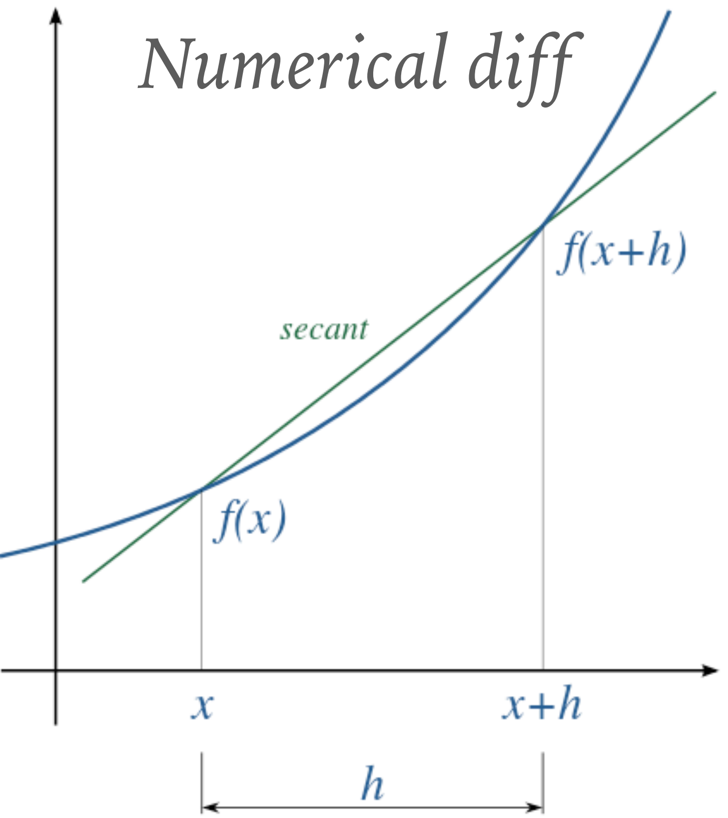


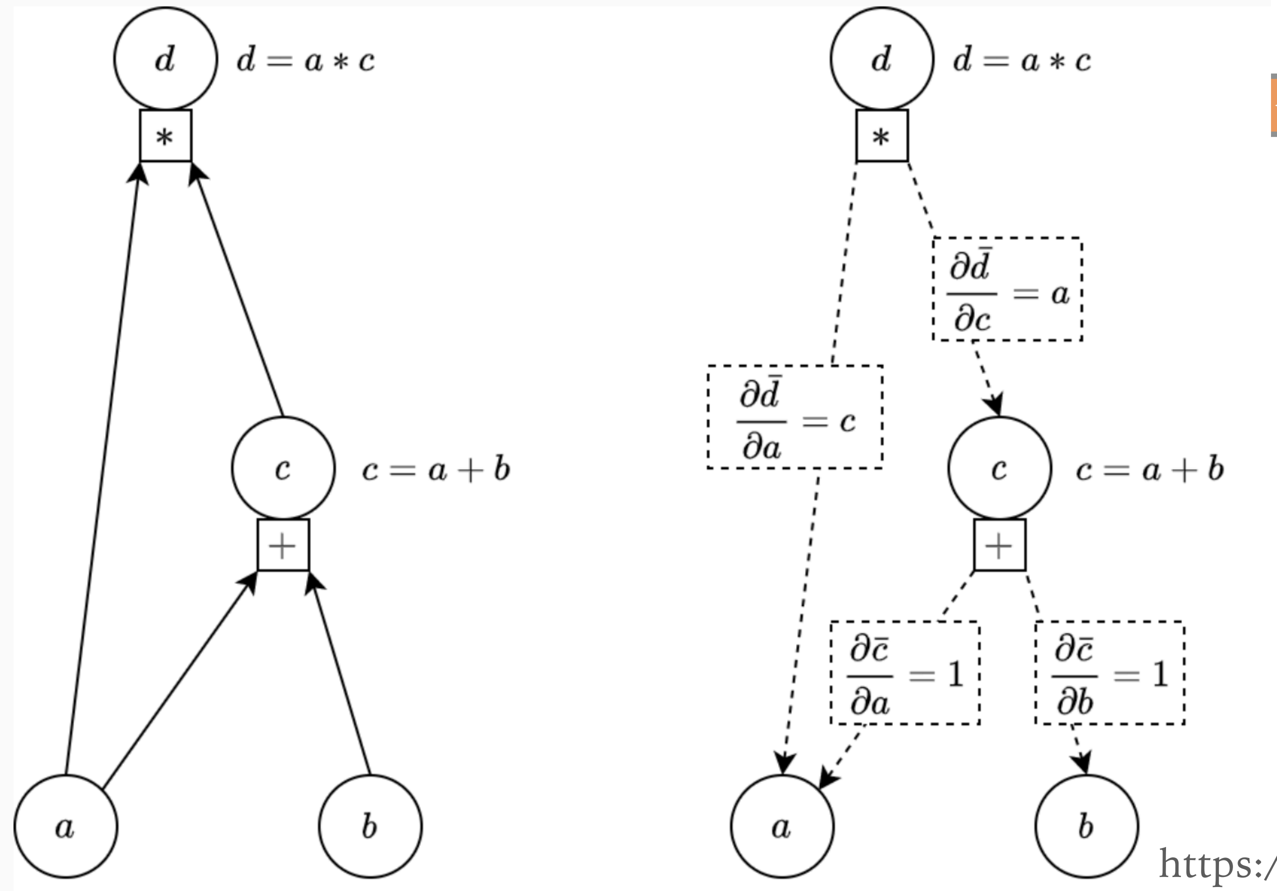
Starting Conditions, with e.g. many parameters

Complex manipulation. Differential equations Simulations

Results Figure of merit-based

Cost function





$$\frac{\partial d}{\partial a} = \frac{\partial \bar{d}}{\partial a} + \frac{\partial \bar{d}}{\partial c} * \frac{\partial \bar{c}}{\partial a}$$

$$\frac{\partial d}{\partial a} = c + a * 1$$

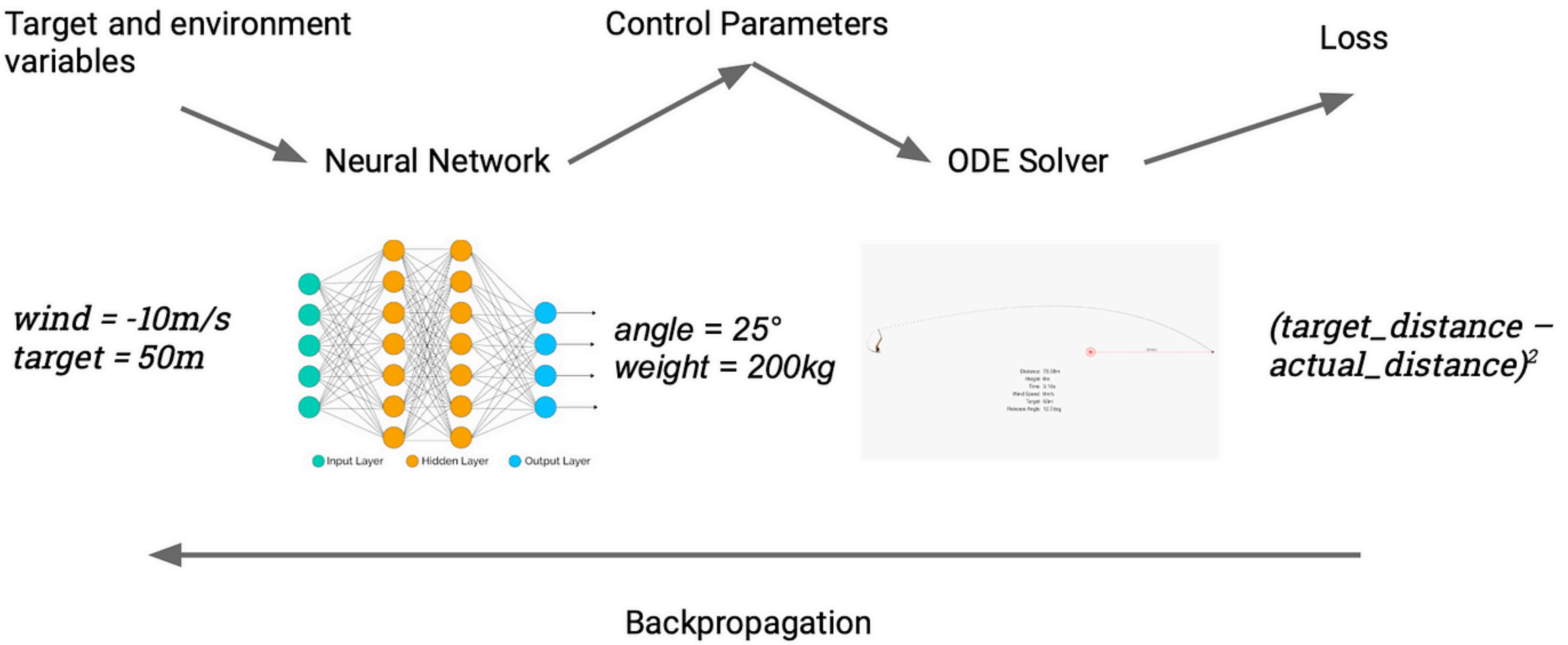
$$\frac{\partial d}{\partial a} = a + b + a$$

$$\frac{\partial d}{\partial a} = 2a + b$$

https://sidsite.com/posts/autodiff/

Use two step approach - 2nd: to the limit with differentiable programming

https://github.com/JuliaComputing/ODSC2019


Target and environment variables

Neural Network

Control Parameters

ODE Solver

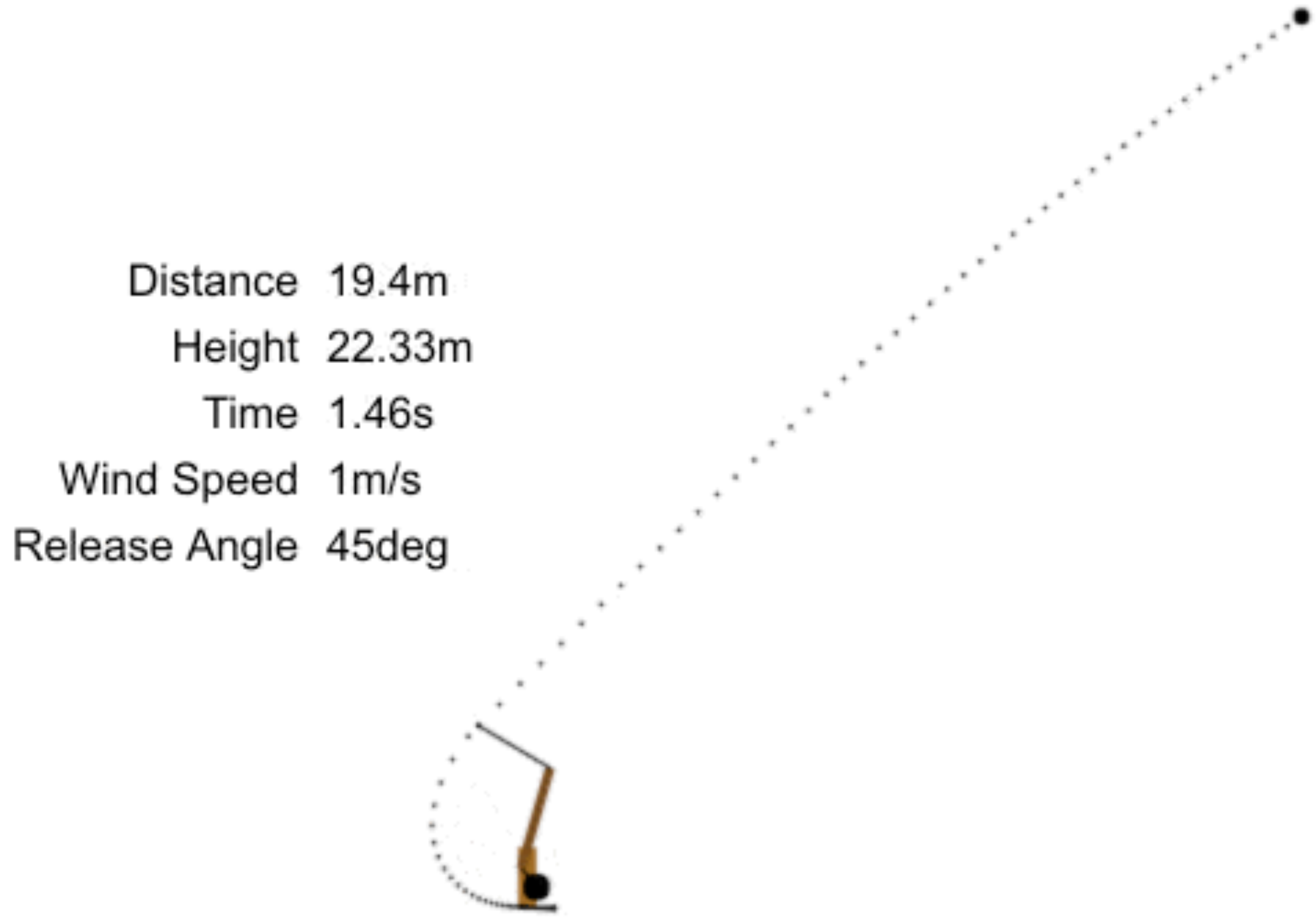
Loss

wind = -10m/s
target = 50m

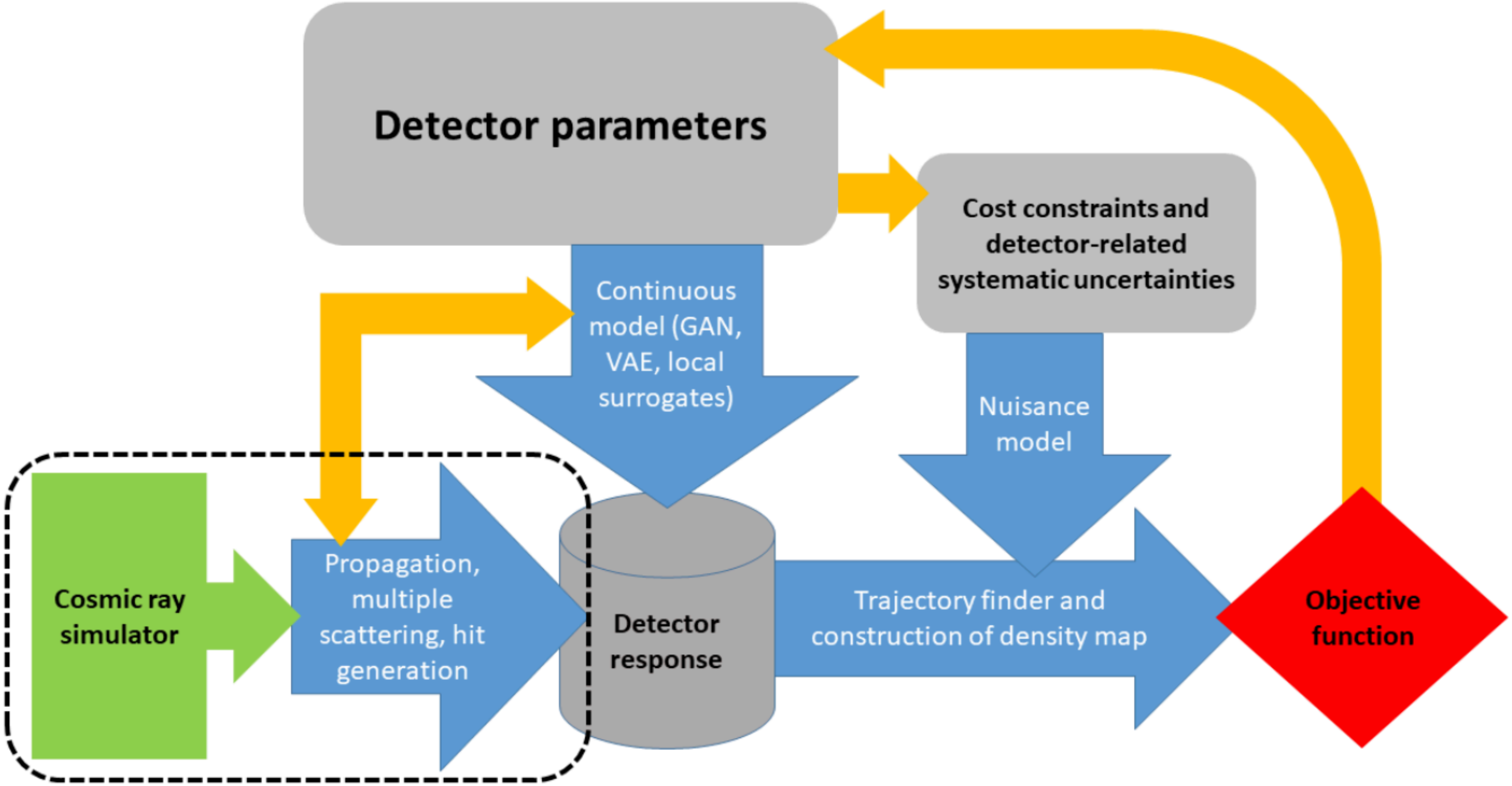
angle = 25°
weight = 200kg

(target_distance – actual_distance)²

Backpropagation



Use two step approach - 2nd: to the limit with differentiable programming

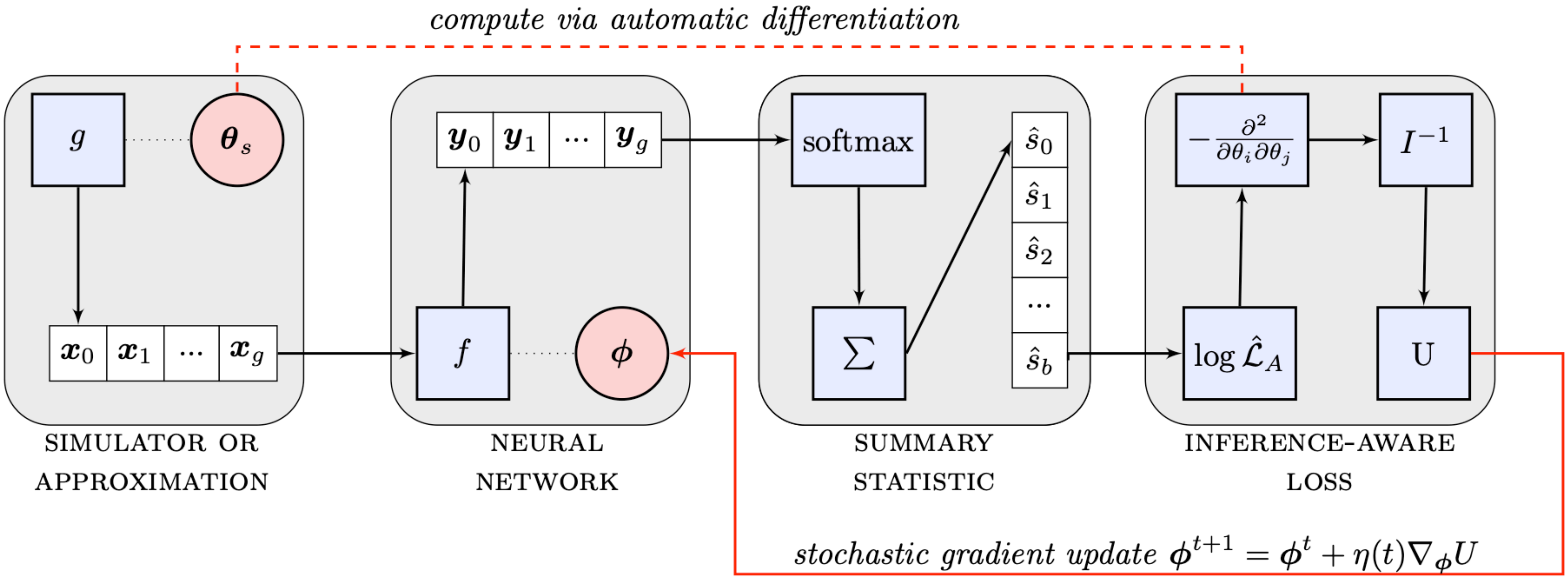


MODE
Machine-learning Optimized Design of Experiments

Optimization of detector design and operation

https://inspirehep.net/files/8f6784eac6a7f06c184cfc50757de656

Use two step approach - 2nd: to the limit with differentiable programming



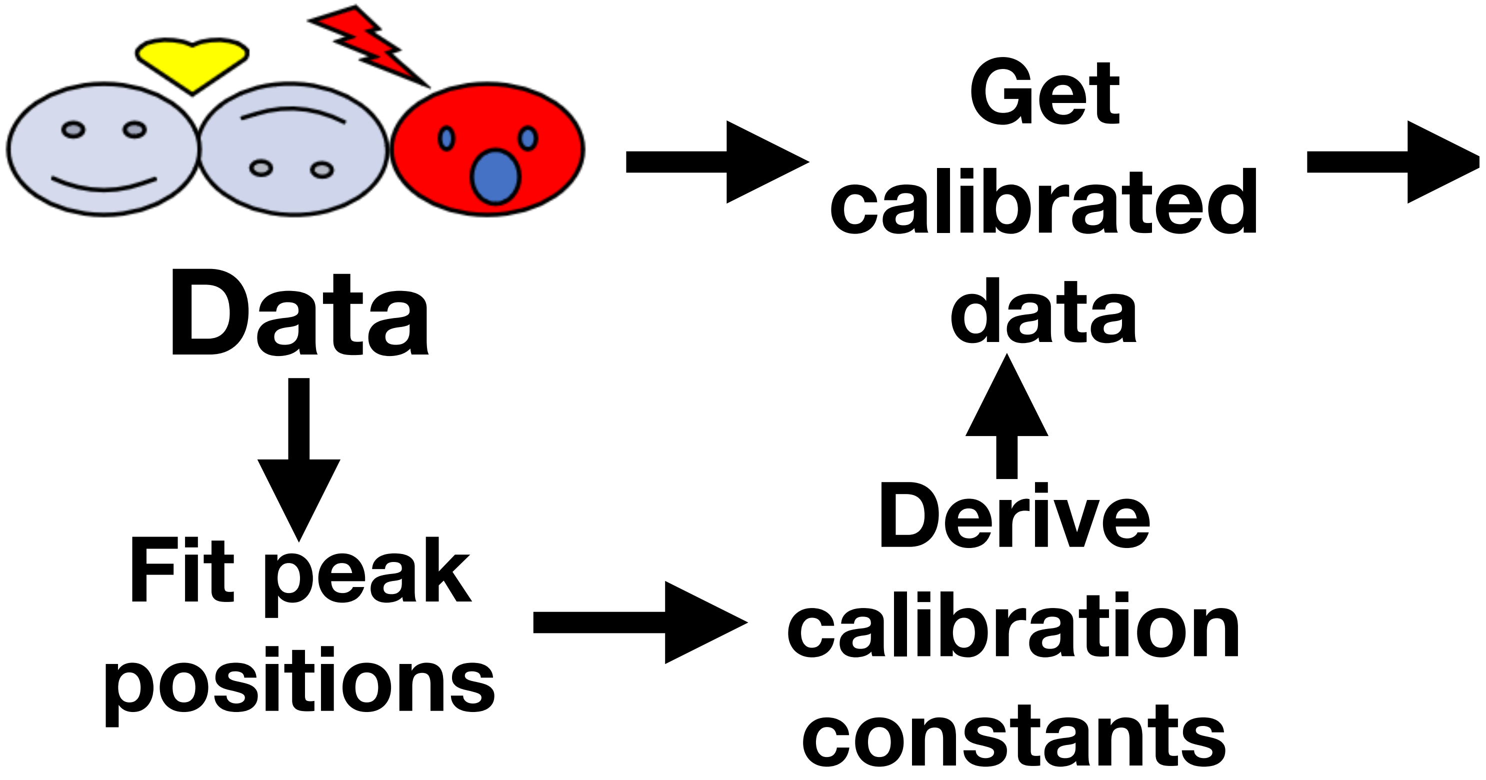
Sketch of the INFERNO algorithm. Batches from a simulator are passed through a neural network and a differentiable summary statistic is constructed that allows to calculate the variance of the POI. The parameters of the network are then updated by stochastic gradient descent.

https://cds.cern.ch/record/2807001/plots

Use two step approach - 2nd: to the limit with differentiable programming

Idea: use automatic differentiation to compute gradients of functions

Use gradients to find global optima

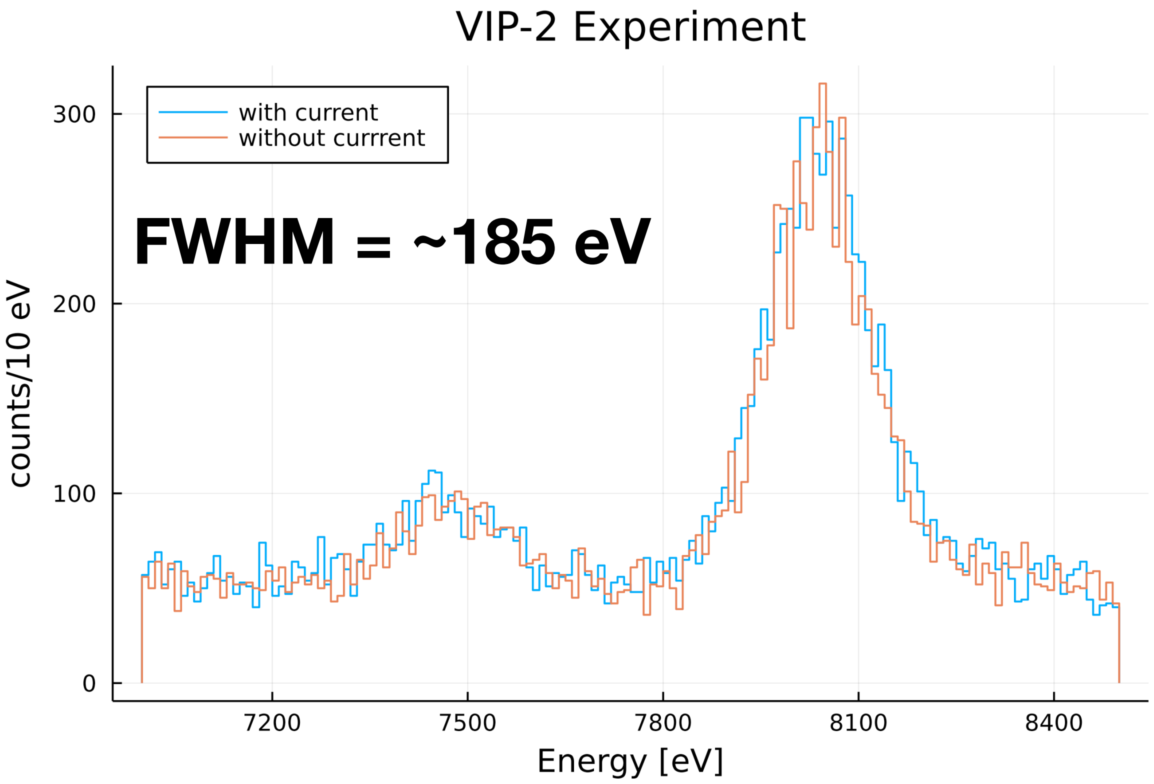


Data

Get calibrated data

Fit peak positions

Derive calibration constants

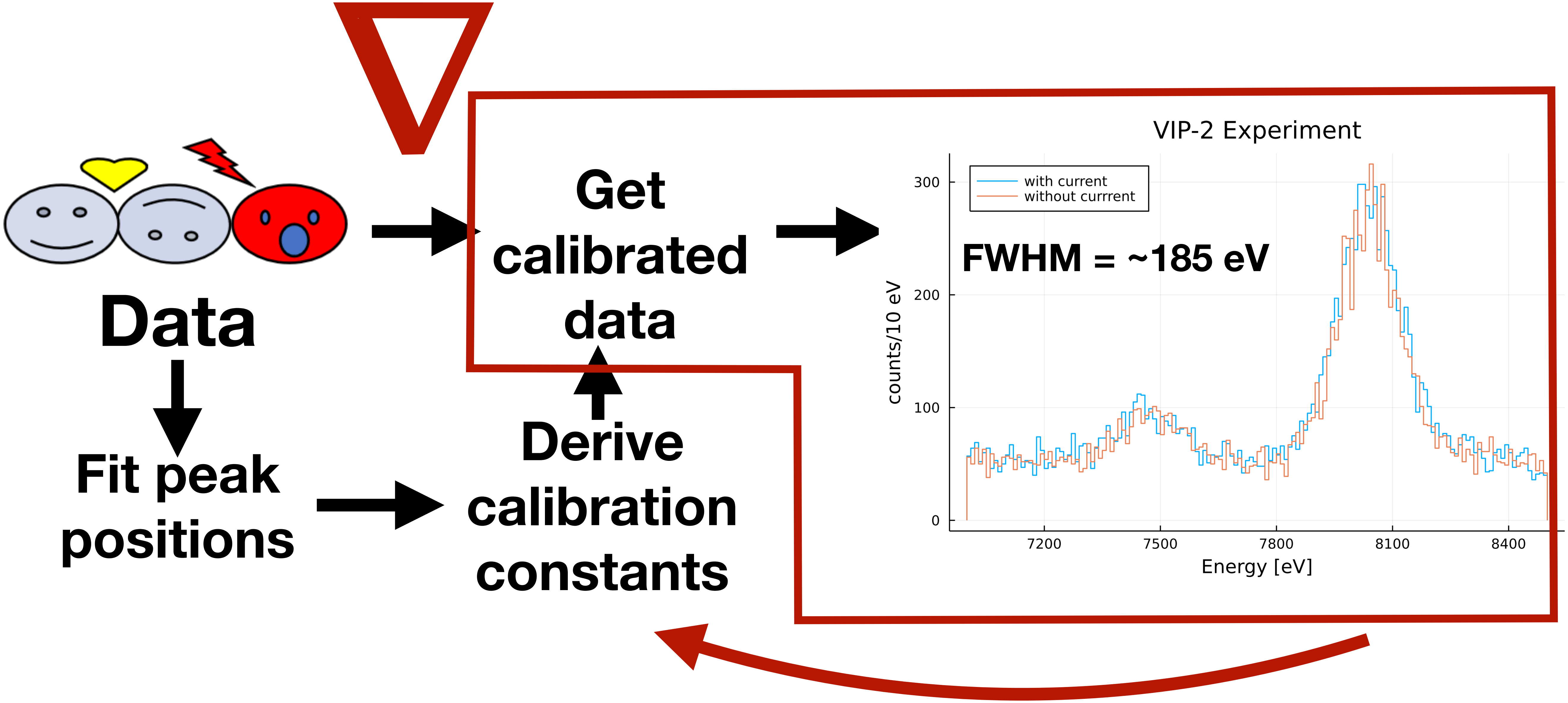


Our Calibration Flow

Use two step approach - 2nd: to the limit with differentiable programming

Idea: use automatic differentiation to compute gradients of functions

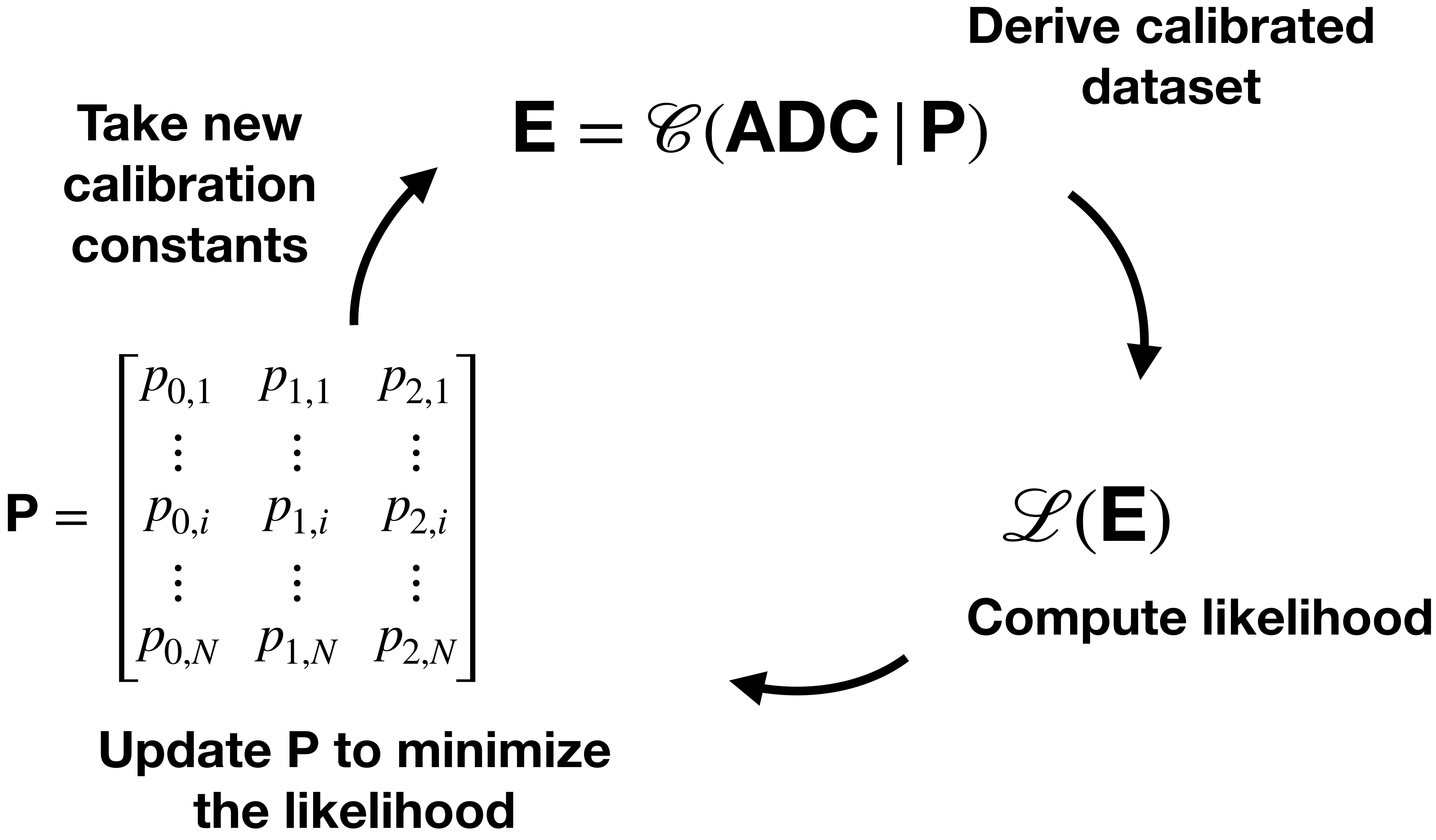
Use gradients to find global optima



Following the gradient, change the constants to enhance FWHM

Use two step approach - 2nd: to the limit with differentiable programming

Loss function for the gradient: log likelihood



Derive calibrated dataset

$$\mathbf{E} = \mathscr{C}(\mathbf{ADC} \,|\, \mathbf{P})$$

Compute likelihood

$$\mathscr{L}(\mathbf{E})$$

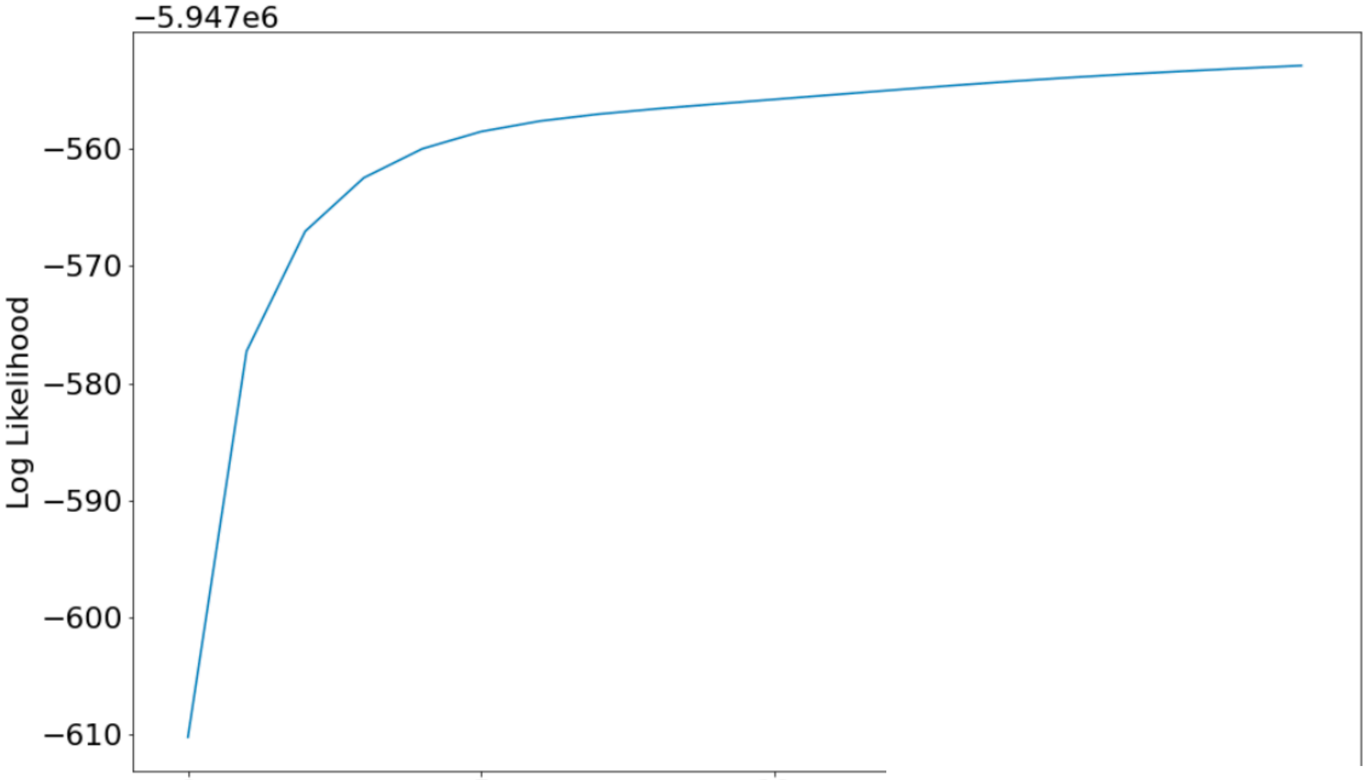
Update P to minimize the likelihood

$$\mathbf{P} = \begin{bmatrix} p_{0,1} & p_{1,1} & p_{2,1} \\ \vdots & \vdots & \vdots \\ p_{0,i} & p_{1,i} & p_{2,i} \\ \vdots & \vdots & \vdots \\ p_{0,N} & p_{1,N} & p_{2,N} \end{bmatrix}$$

Take new calibration constants

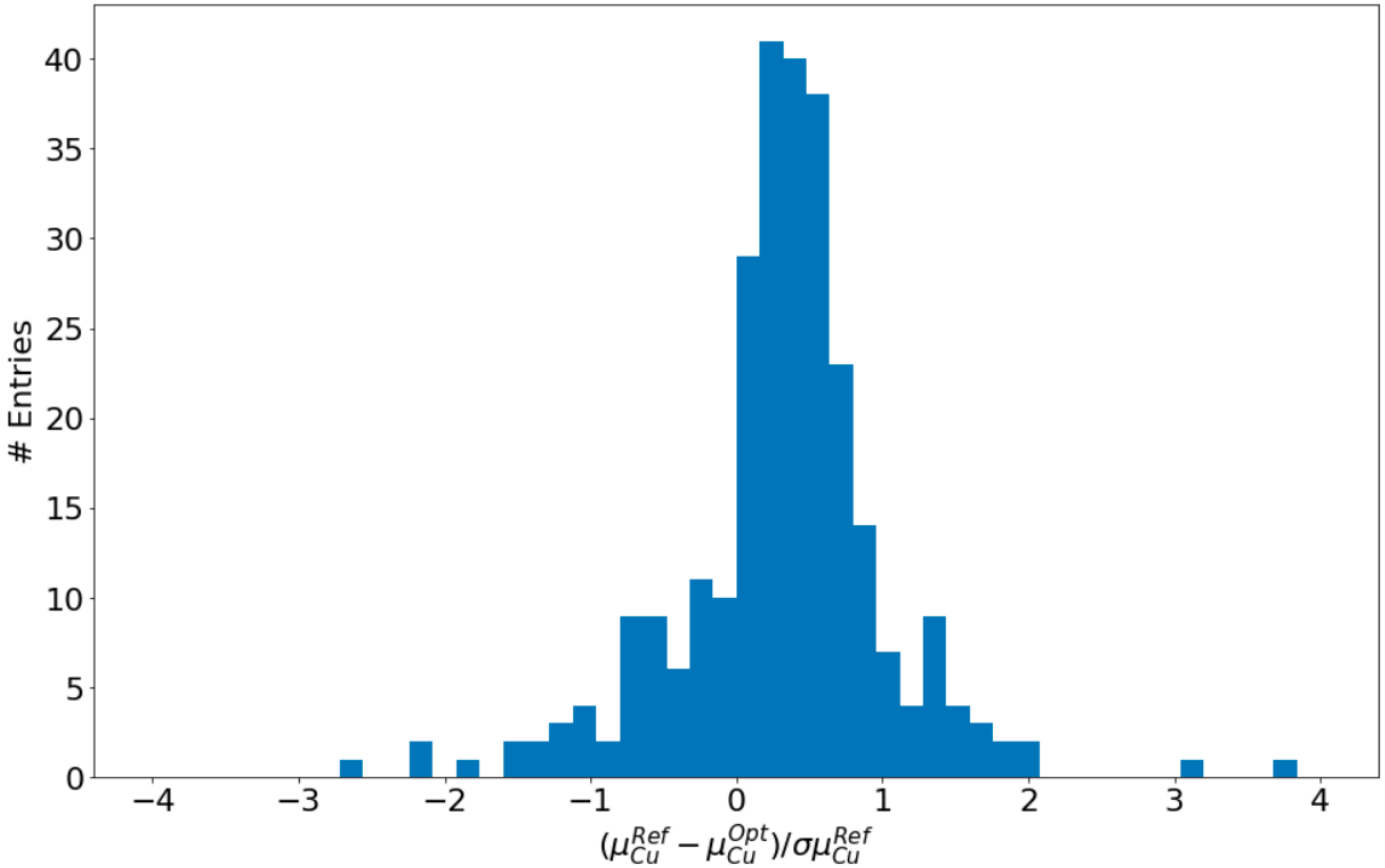
$$\mathcal{L}(\mathbf{ADC}, \mathbf{P}) = \prod_{i} \prod_{j \in i} \Big(\mathrm{Gauss}(\mathcal{C}(\mathrm{ADC}_{i,j}, P_i) - \mu_{Cu_{K_{\alpha 1}}}, \sigma) + \mathrm{Gauss}(\mathcal{C}(\mathrm{ADC}_{i,j}, P_i) - \mu_{Cu_{K_{\alpha 2}}}, \sigma) \Big)$$

Use two step approach - 2nd: to the limit with differentiable programming



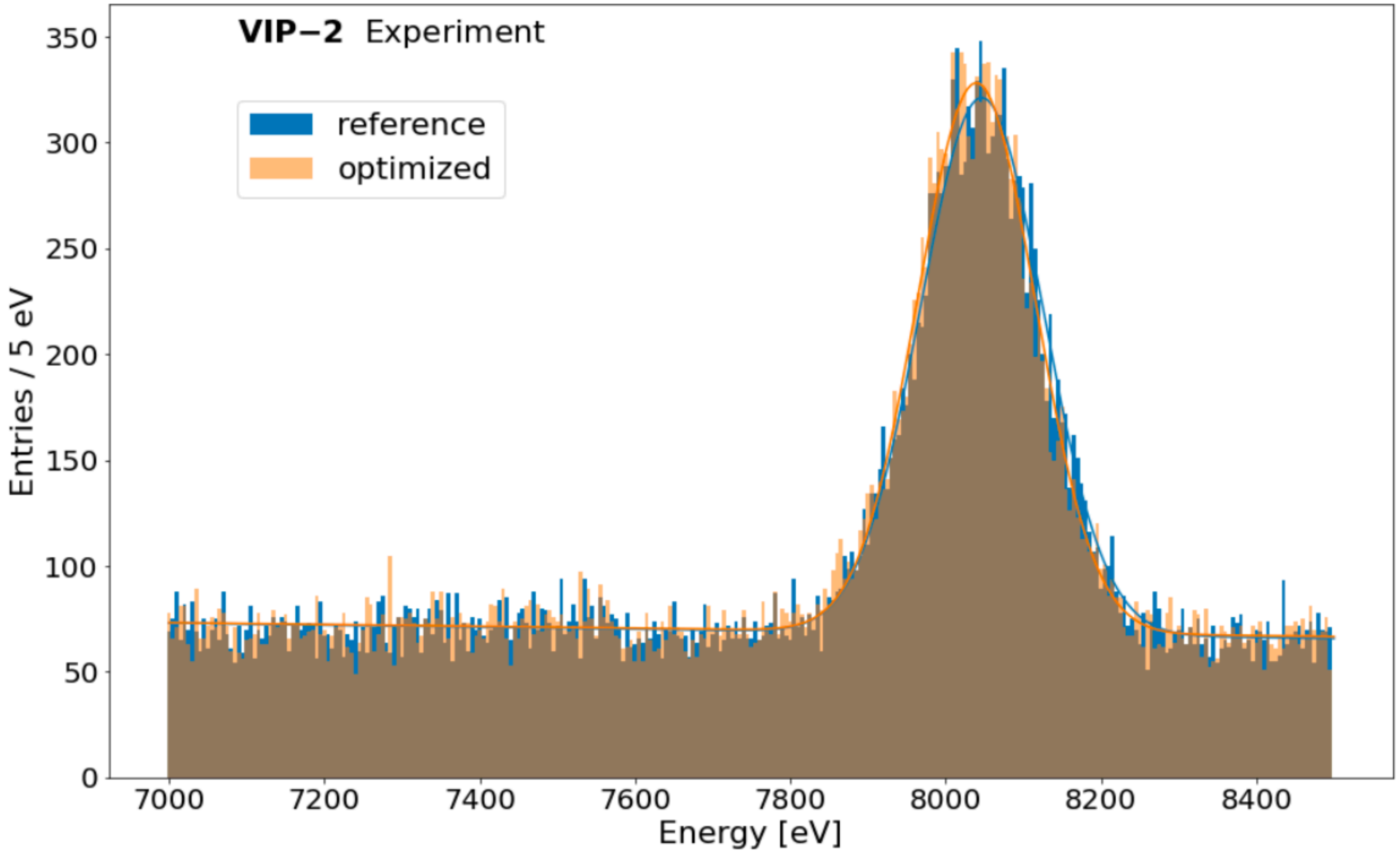
Gradient descended into plateau

Most of the calibration parameters do not change much



Results

https://arxiv.org/abs/2305.17153
Submitted to Meas. Sci. Tech.



	Position [eV]	FWHM [eV]	χ^2/ndf
Reference	8050 ± 1	185 ± 2	1.64
Optimized	8048 ± 1	176 ± 2	1.25

$$f(x, A, \mu, \sigma) = A \times \frac{51}{100} \times \mathrm{Gauss}(x-\mu-20, \sigma) + T_2(x) + A \times \mathrm{Gauss}(x-\mu, \sigma) + T_1(x) + m \times x + C$$

$$T_i(x) = \frac{A_i}{2\beta\sigma} \times e^{\frac{x-\nu}{\beta\sigma} \frac{1}{2\beta^2}} \times erfc\left(\frac{x-\nu}{\sqrt{2\pi}} + \frac{1}{\sqrt{2}\beta}\right)$$

Conclusions

- *VIP-2 experiment at LNGS uses X-ray spectroscopy as tool for precision tests QM and Standard Model*
- *Showed a ML and DP way to increase the detector resolution*
  - *Below 180 eV FWHM at Cu for the first time in VIP-2*
  - *The method can correct for miscalibration*
  - *Closer compatibility with the model*
  - *Higher discovery significance of forbidden transitions*

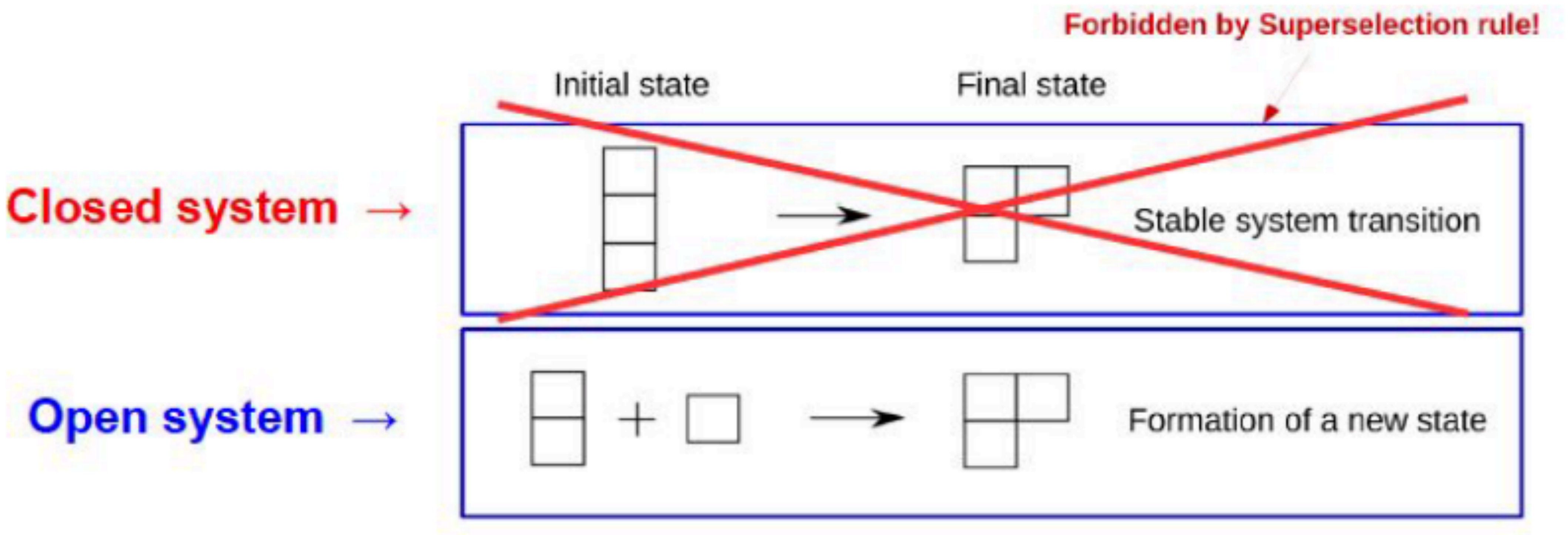
Thank you for your attention!

Questions?

Messiah-Greenberg super-selection rule:

Superposition of states with different symmetry are not allowed →

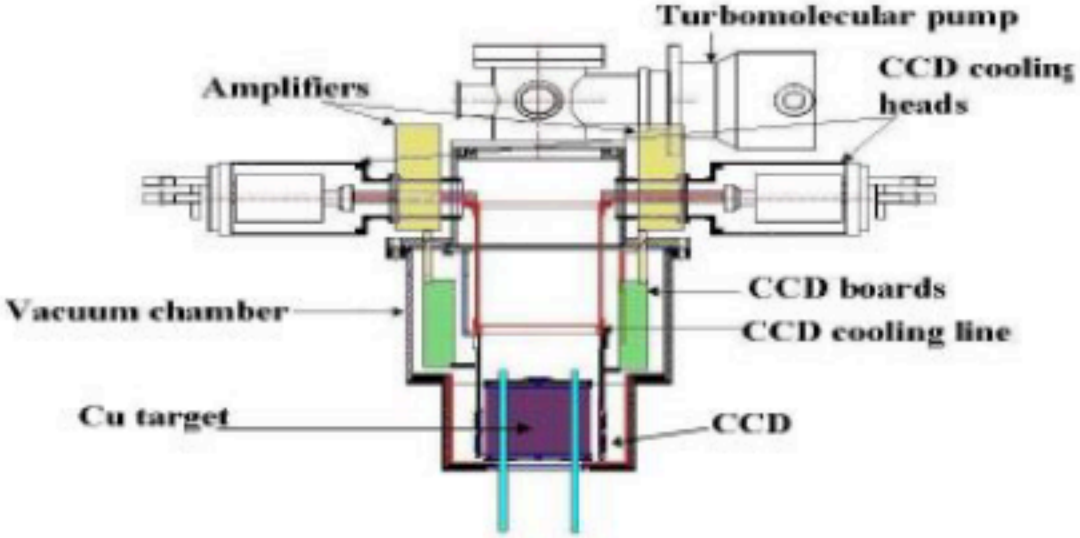
Transition probability between two symmetry states is ZERO

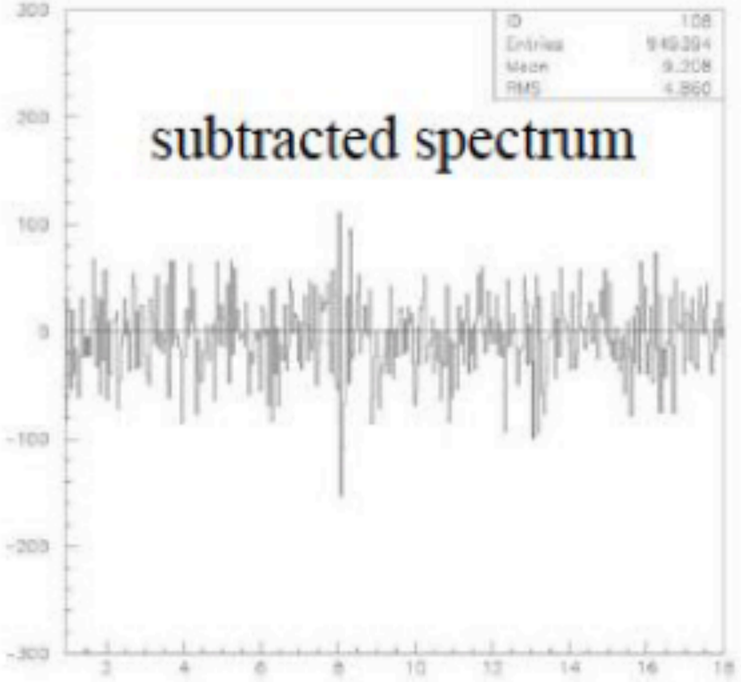


VIP-2 Experiment: best limits on PEP violation of an elementary particle respecting the Messiah-Greenberg super-selection rule

From VIP to VIP-2

a) **copper ultrapure cylindrical foil**
b) **surrounded by 16 Charge Coupled Devices (CCD) res. at 8 keV 320 eV (FWHM)**
c) **inside a vacuum chamber: CCDs cooled to 168K by a cryogenic system**
d) **amplifiers + read out ADC boards.**





$$\beta^2/2 \leq 4.7\times10^{-29}$$

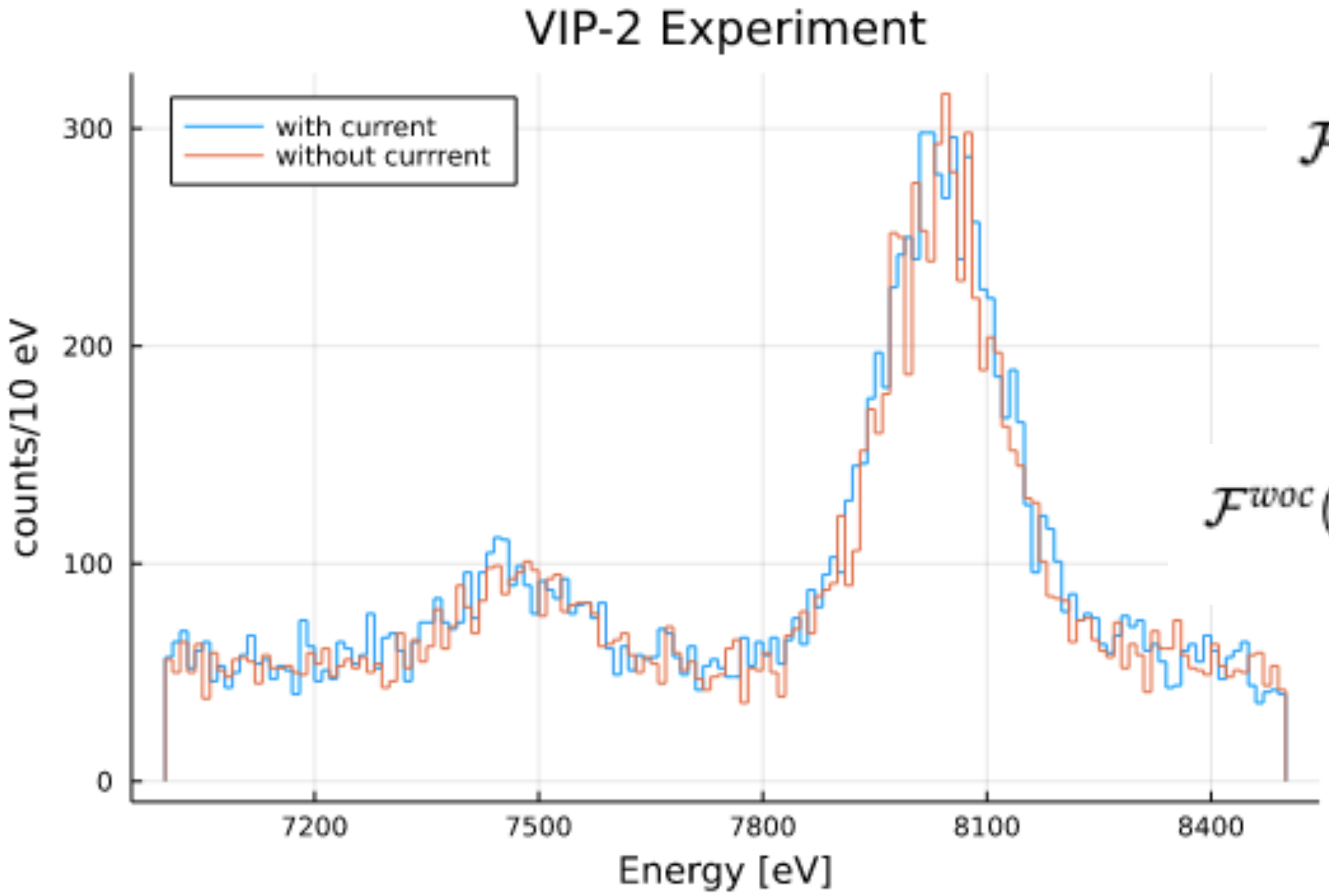
improved the limit obtained by Ramberg & Snow by a factor ~ 400

(Foundation of Physics 41 (2011) 282+ other papers)

GOAL OF VIP-2 : improve the VIP result of 2 orders of magnitude

Symmetry **2022**, *14*(5), 893;
https://doi.org/10.3390/sym14050893

Six months of data taking



Description spectrum with current

$$\mathcal{F}^{wc}(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}) = y_1 \times Ni(\theta_1,\theta_2) + y_2 \times Cu(\theta_3,\theta_4) + \\ + y_3 \times \mathrm{pol}_1(\theta_5) + \mathcal{S} \times PEPV(\theta_4)$$

Description spectrum without current

$$\mathcal{F}^{woc}(\boldsymbol{\theta},\boldsymbol{y}) = y_1 \times Ni(\theta_1,\theta_2) + y_2 \times Cu(\theta_3,\theta_4) + y_3 \times \mathrm{pol}_1(\theta_5)$$

Likelihood

$$\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}) = \mathrm{Poiss}(\mathcal{D}^{wc}|\mathcal{F}^{wc}(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})) \times \mathrm{Poiss}(\mathcal{D}^{woc}|\mathcal{F}^{woc}(\boldsymbol{\theta},\boldsymbol{y} \times \mathcal{R}))$$

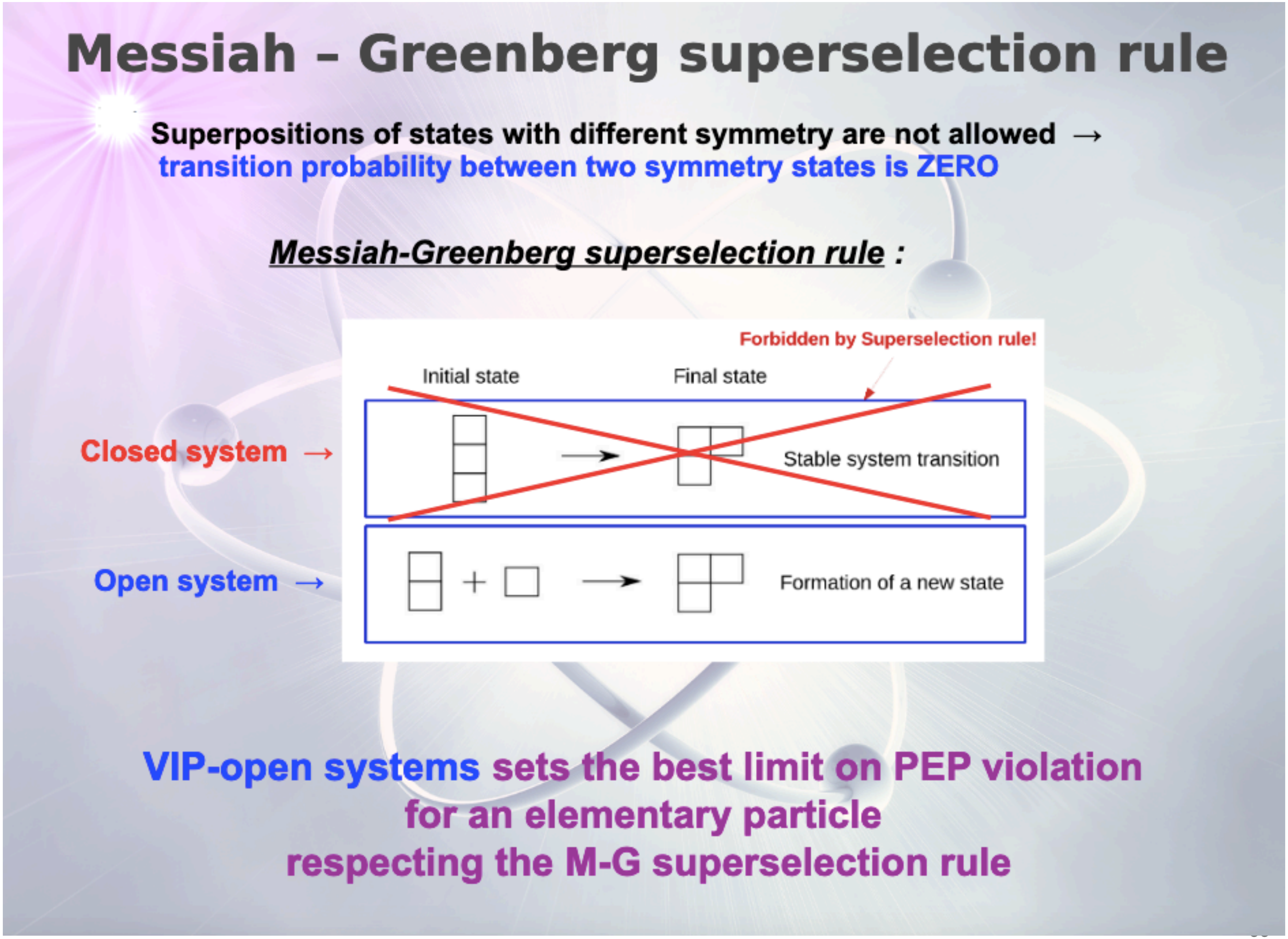
Bayesian

$$p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}|\mathcal{D}^{wc},\mathcal{D}^{woc}) = \\ = \frac{\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})}{\int d\boldsymbol{\theta}d\boldsymbol{y}\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})}.$$

Frequentist

$$t_{\mathcal{S}} = -2\ln\Lambda(\mathcal{S}) = -2\ln\frac{\mathcal{L}(\hat{\hat{\boldsymbol{\theta}}},\hat{\hat{\boldsymbol{y}}},\mathcal{S})}{\mathcal{L}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{y}},\hat{\mathcal{S}})}$$

$$p_{\mathcal{S}} = \int_{t_{obs}}^{\infty} f(t_{\mathcal{S}}|\mathcal{S})dt_{\mathcal{S}}$$



Messiah - Greenberg superselection rule

Superpositions of states with different symmetry are not allowed →
transition probability between two symmetry states is ZERO

***Messiah-Greenberg superselection rule* :**

Forbidden by Superselection rule!

Initial state Final state

Closed system → Stable system transition

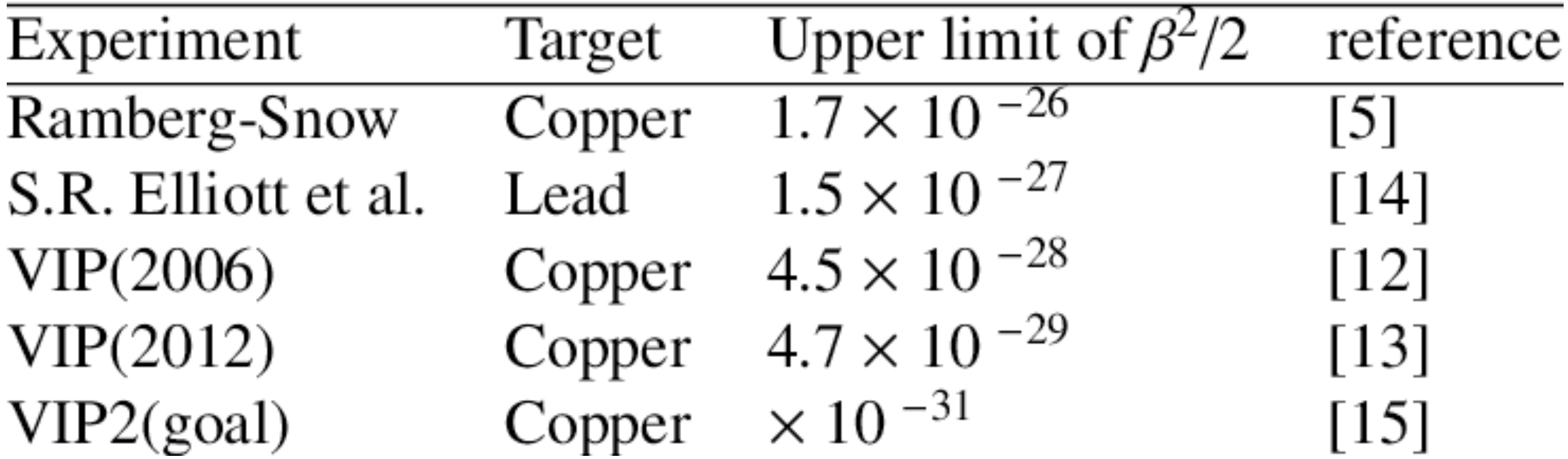
Open system → Formation of a new state

VIP-open systems sets the best limit on PEP violation for an elementary particle respecting the M-G superselection rule

VIP-2 experiment goal

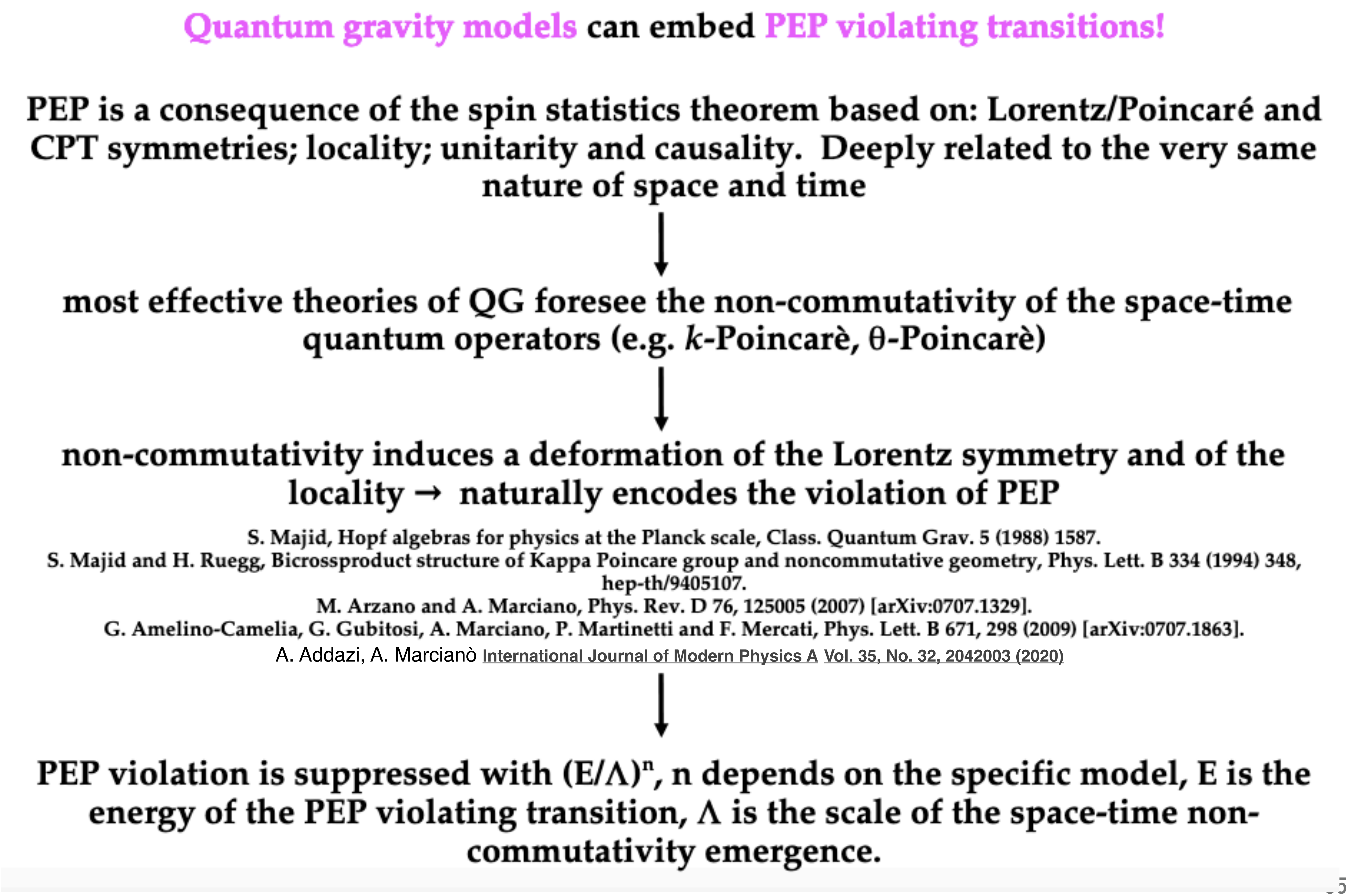
(Upper limit not using Close Encounters (CE) treatment)

As reference for past experiments



Experiment	Target	Upper limit of $\beta^2/2$	reference
Ramberg-Snow	Copper	1.7×10^{-26}	[5]
S.R. Elliott et al.	Lead	1.5×10^{-27}	[14]
VIP(2006)	Copper	4.5×10^{-28}	[12]
VIP(2012)	Copper	4.7×10^{-29}	[13]
VIP2(goal)	Copper	$\times 10^{-31}$	[15]

New paradigm for VIP-2



Quantum gravity models can embed PEP violating transitions!

PEP is a consequence of the spin statistics theorem based on: Lorentz/Poincaré and CPT symmetries; locality; unitarity and causality. Deeply related to the very same nature of space and time

↓

most effective theories of QG foresee the non-commutativity of the space-time quantum operators (e.g. *k*-Poincarè, θ-Poincarè)

↓

non-commutativity induces a deformation of the Lorentz symmetry and of the locality → naturally encodes the violation of PEP

S. Majid, Hopf algebras for physics at the Planck scale, Class. Quantum Grav. 5 (1988) 1587.
S. Majid and H. Ruegg, Bicrossproduct structure of Kappa Poincare group and noncommutative geometry, Phys. Lett. B 334 (1994) 348, hep-th/9405107.
M. Arzano and A. Marciano, Phys. Rev. D 76, 125005 (2007) [arXiv:0707.1329].
G. Amelino-Camelia, G. Gubitosi, A. Marciano, P. Martinetti and F. Mercati, Phys. Lett. B 671, 298 (2009) [arXiv:0707.1863].
A. Addazi, A. Marcianò International Journal of Modern Physics A Vol. 35, No. 32, 2042003 (2020)

↓

PEP violation is suppressed with $(E/\Lambda)^n$, n depends on the specific model, E is the energy of the PEP violating transition, Λ is the scale of the space-time non-commutativity emergence.

How to model PEP violations

- *Ignatiev & Kuzmin model: Fermi oscillator with a third state*

(Ignatiev, A.Y., Kuzmin, V. , Quarks '86: Proceedings of the 229 Seminar, Tbilisi, USSR, 1517 April 1986)

$$\begin{array}{ll} a^{+}|0\rangle = |1\rangle & a|0\rangle = 0 \\ a^{+}|1\rangle = \beta|2\rangle & a|1\rangle = |0\rangle \\ a^{+}|2\rangle = 0 & a|2\rangle = \beta|1\rangle \end{array}$$

β quantifies the degree of violation in the transition

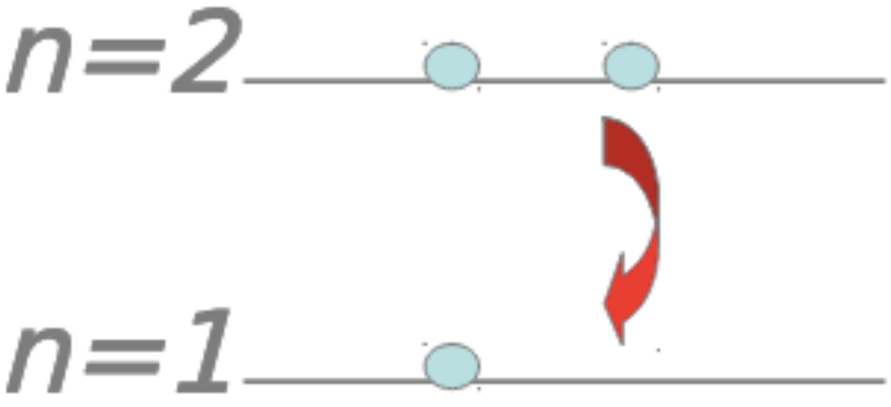
- *Greenberg & Mohapatra: Local Quantum Field Theory, q parameter deforms anticommutators [Phys. Rev. Lett. 1987,59,2507]:*

$$a_k a^{+}_l - q a^{+}_l a_k = \delta_{k,l}$$

- *Rahal & Campa: global wave function of the electrons not exactly antisymmetric,*

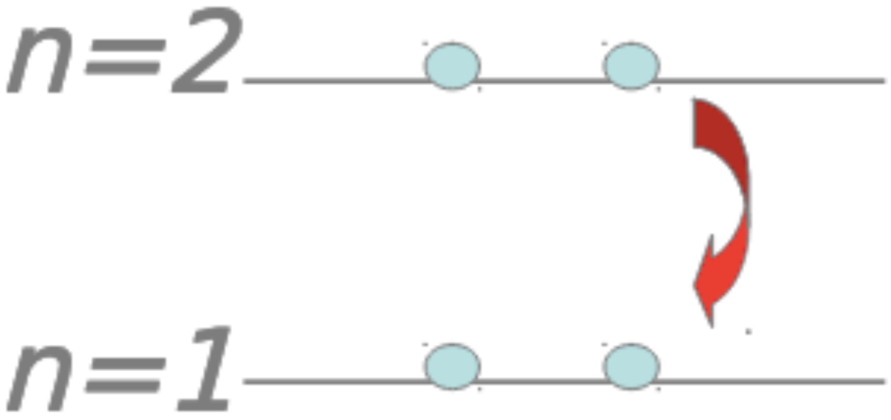
  PEP holds as long as the number of wrongly entangled pairs is small

Search for anomalous X-ray transitions performed by electrons introduced in a target trough a DC current (open system)



Normal 2p → 1s transition

~ 8.05 keV in Cu



2p → 1s transition violating Pauli principle

~ 7.7 keV in Cu

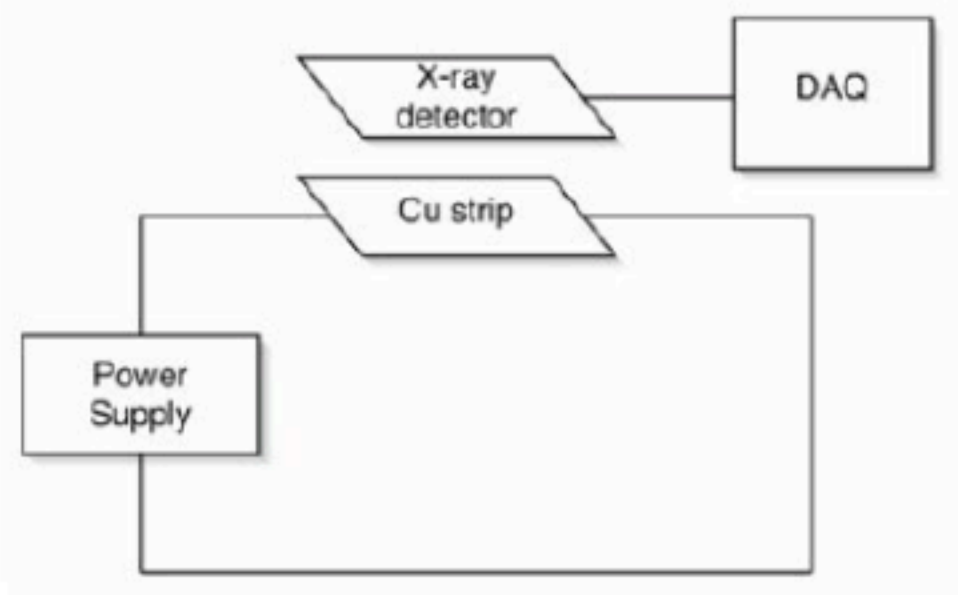
Paul Indelicato (Ecole Normale Supérieure et Université Pierre et Marie Curie)

Multiconfiguration Dirac-Fock approach

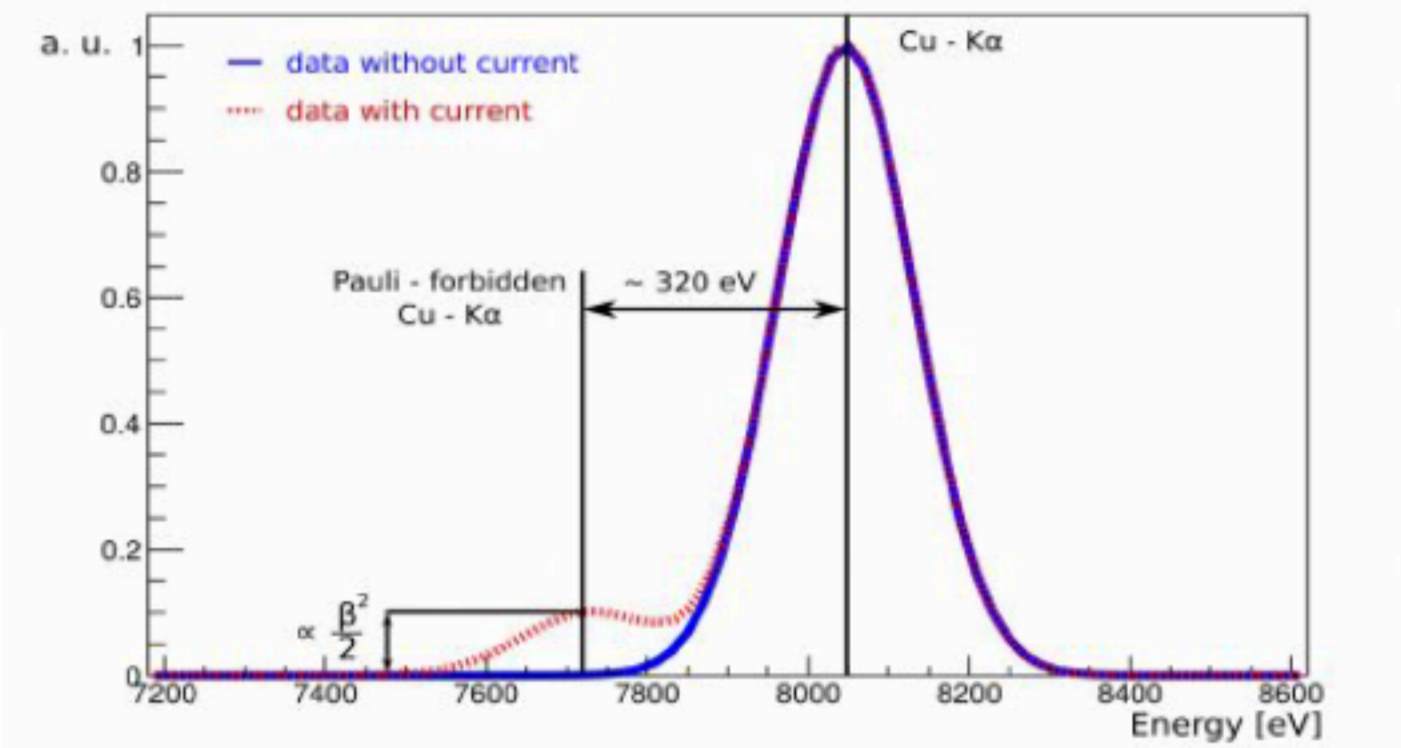
Accounts for the shielding of the two inner electrons

Greenberg, O. W. & Mohapatra, R. N., Phys Rev Lett 59, (1987).
E. Ramberg and G. A. Snow, Phys Lett B 238, 438-441(1990)

Search for anomalous electronic transitions in Cu
induced by a circulating current
introduced electrons interact with the valence electrons
search transition from 2p to 1s already filled by 2 electrons
alternated to X-ray background measurements without current

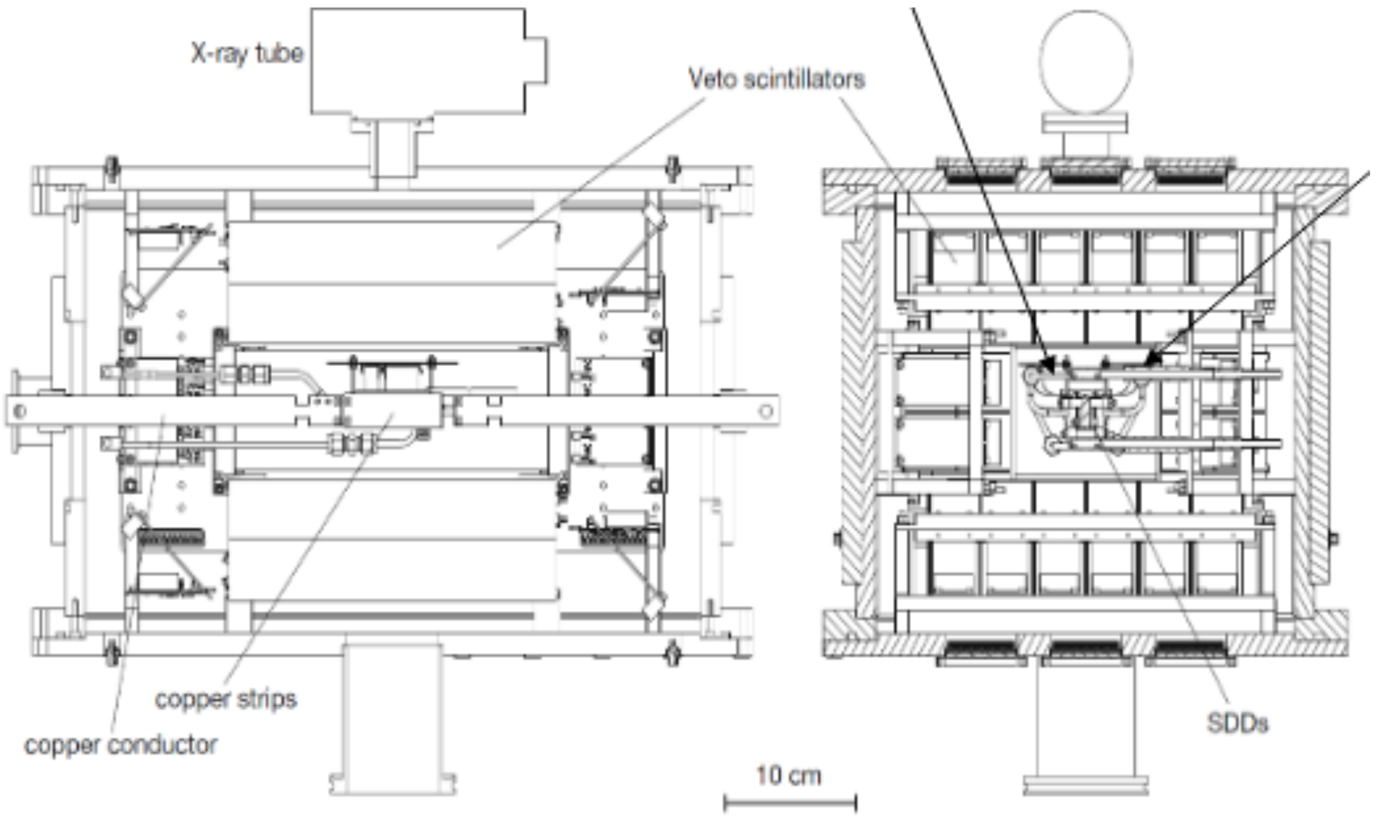


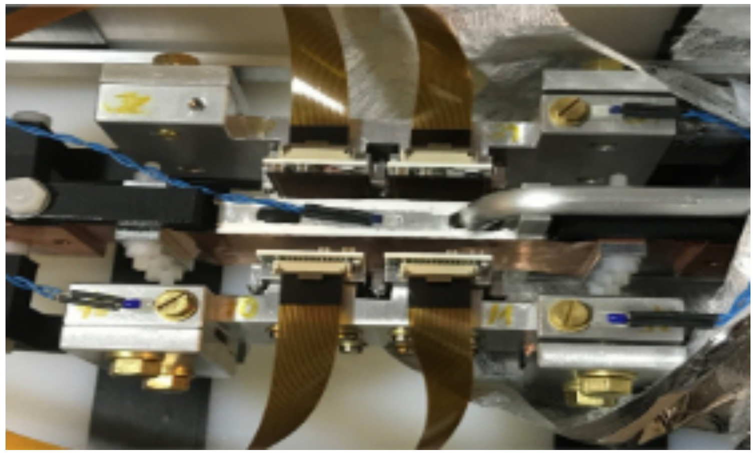
PEP Violation Signal:

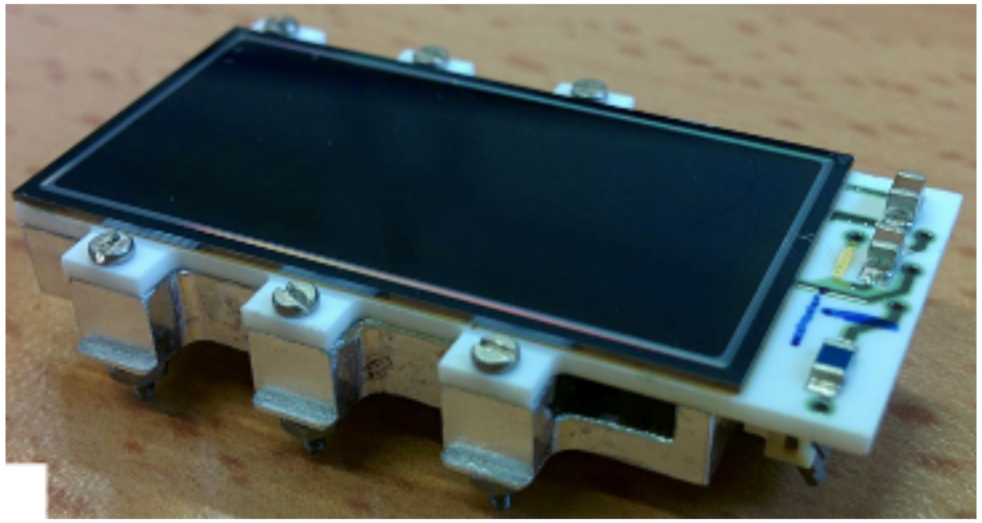


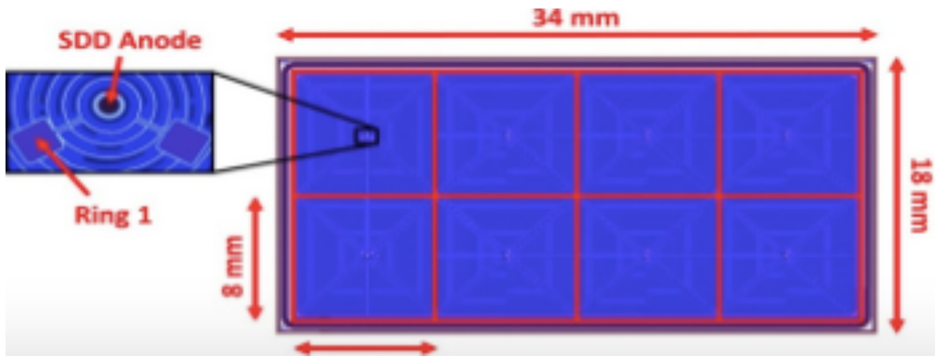
The VIP-2 Experiment

*Silicon Drift Detectors (**SDDs**) higher resolution (190 eV FWHM at 8.0 →keV), faster (triggerable) detectors. 4 arrays of 2 x 4 SDDs 8mm x 8mm each, liquid argon closed circuit cooling 170 °C*









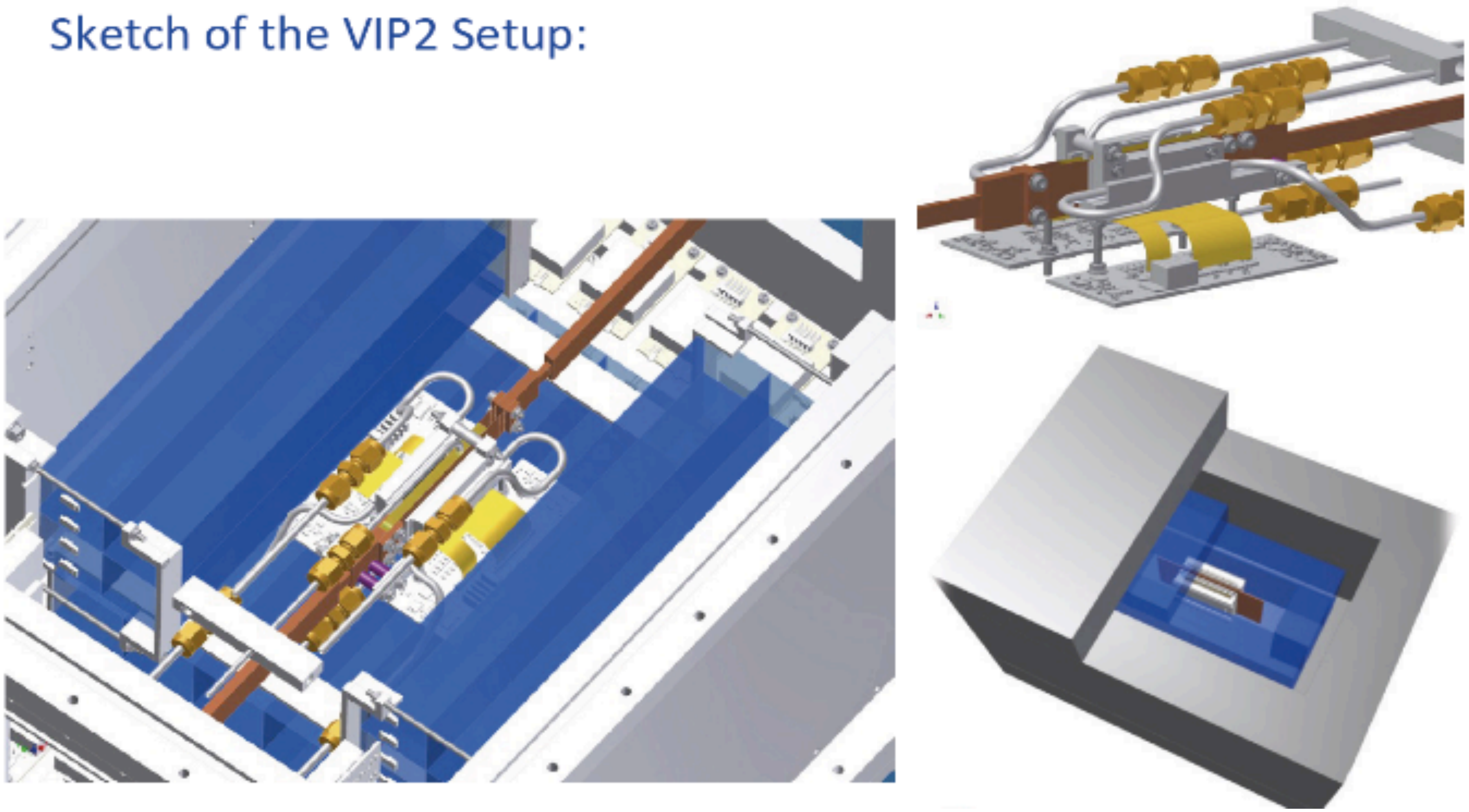
The VIP-2 Experiment

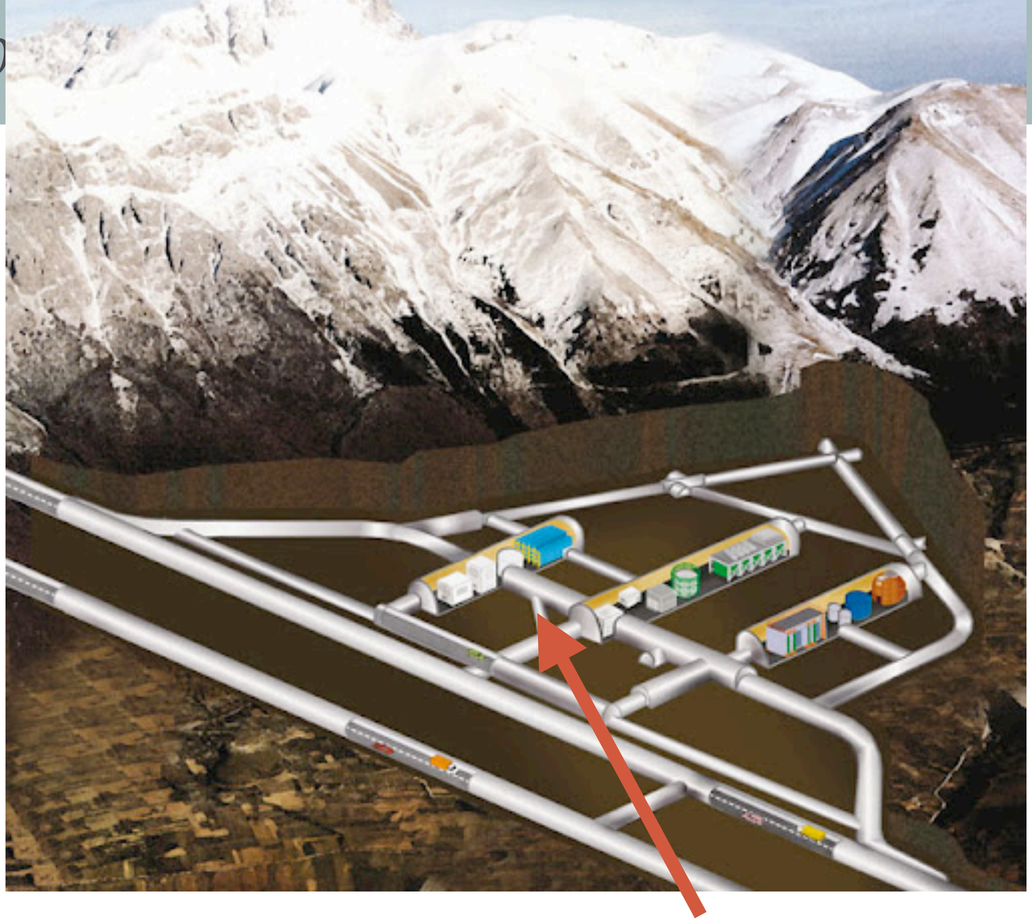
*2 strip shaped Cu targets (25 um x 7 cm x 2 cm) more compact target → higher acceptance, thinner → **higher efficiency***
DC current supply to Cu bars

*Cu strips cooled by a closed Fryka chiller circuit →**higher current***

(100 A) @ 20 °C of Cu target implies 1 °K heating in SDDs

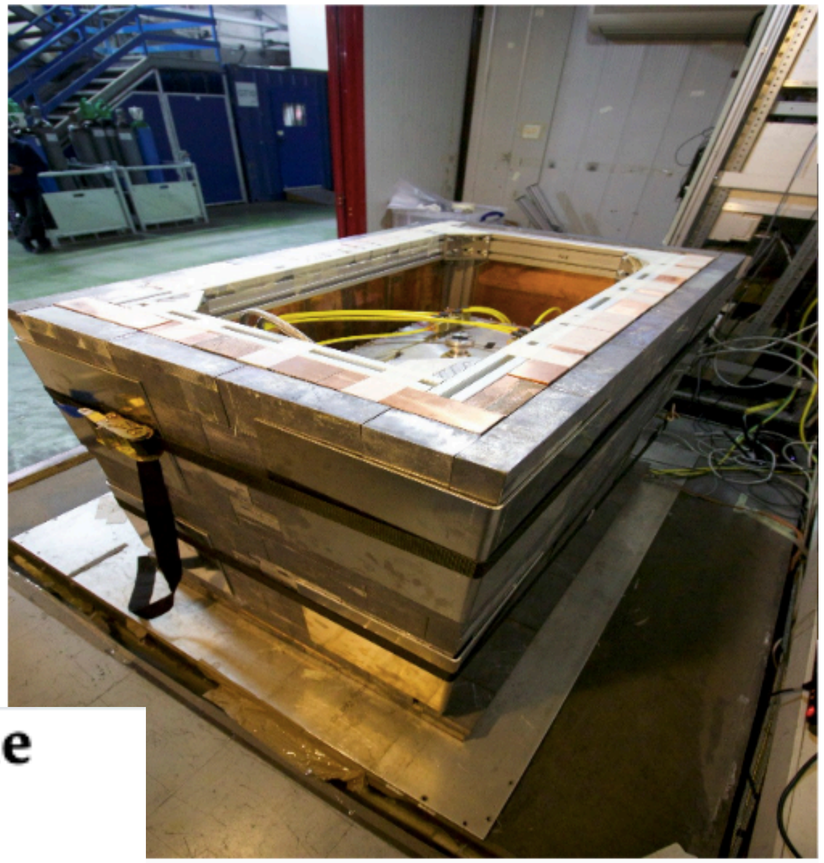
Sketch of the VIP2 Setup:



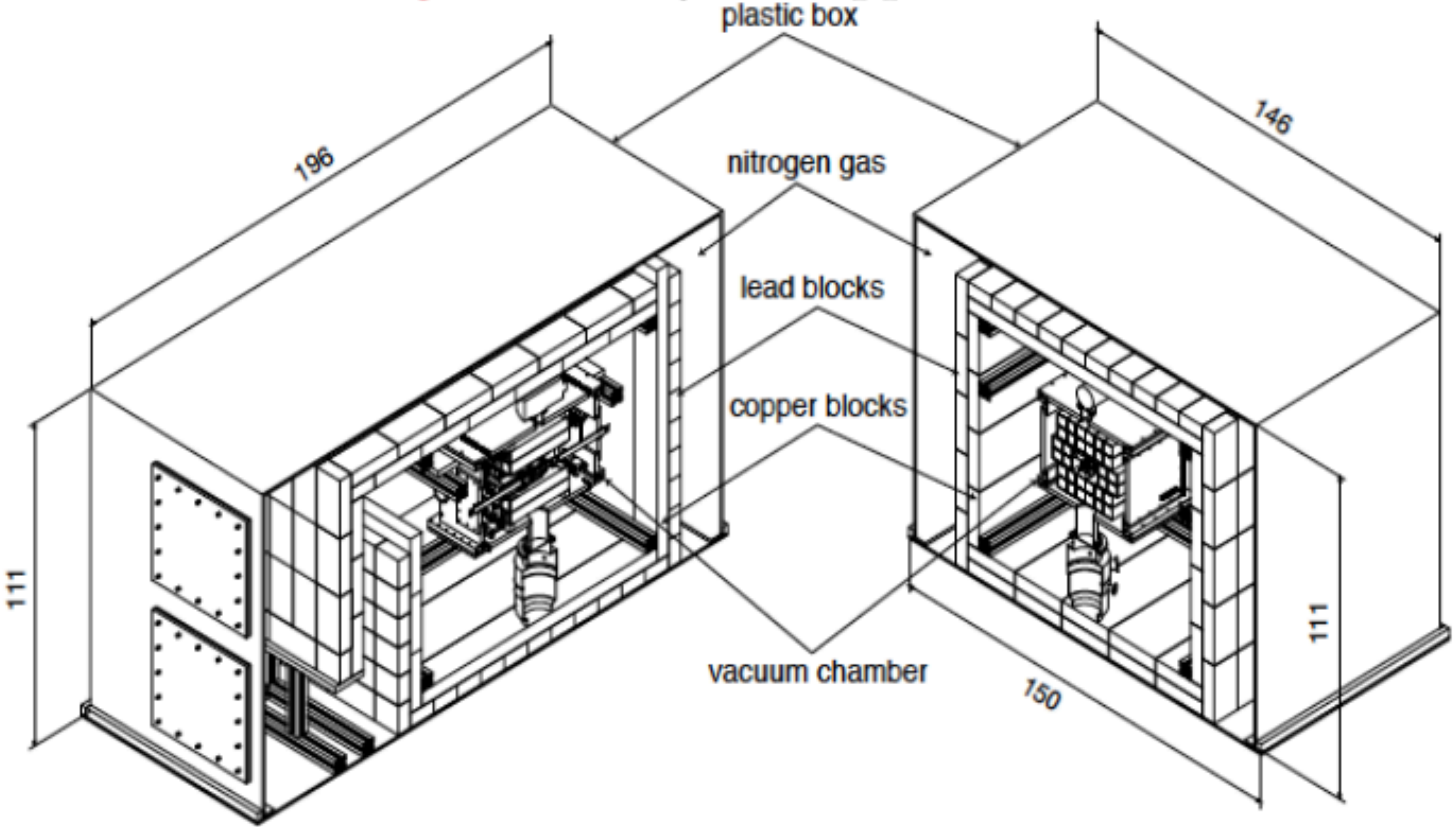


1400 m rock coverage

Upgrade concluded in April 2019:

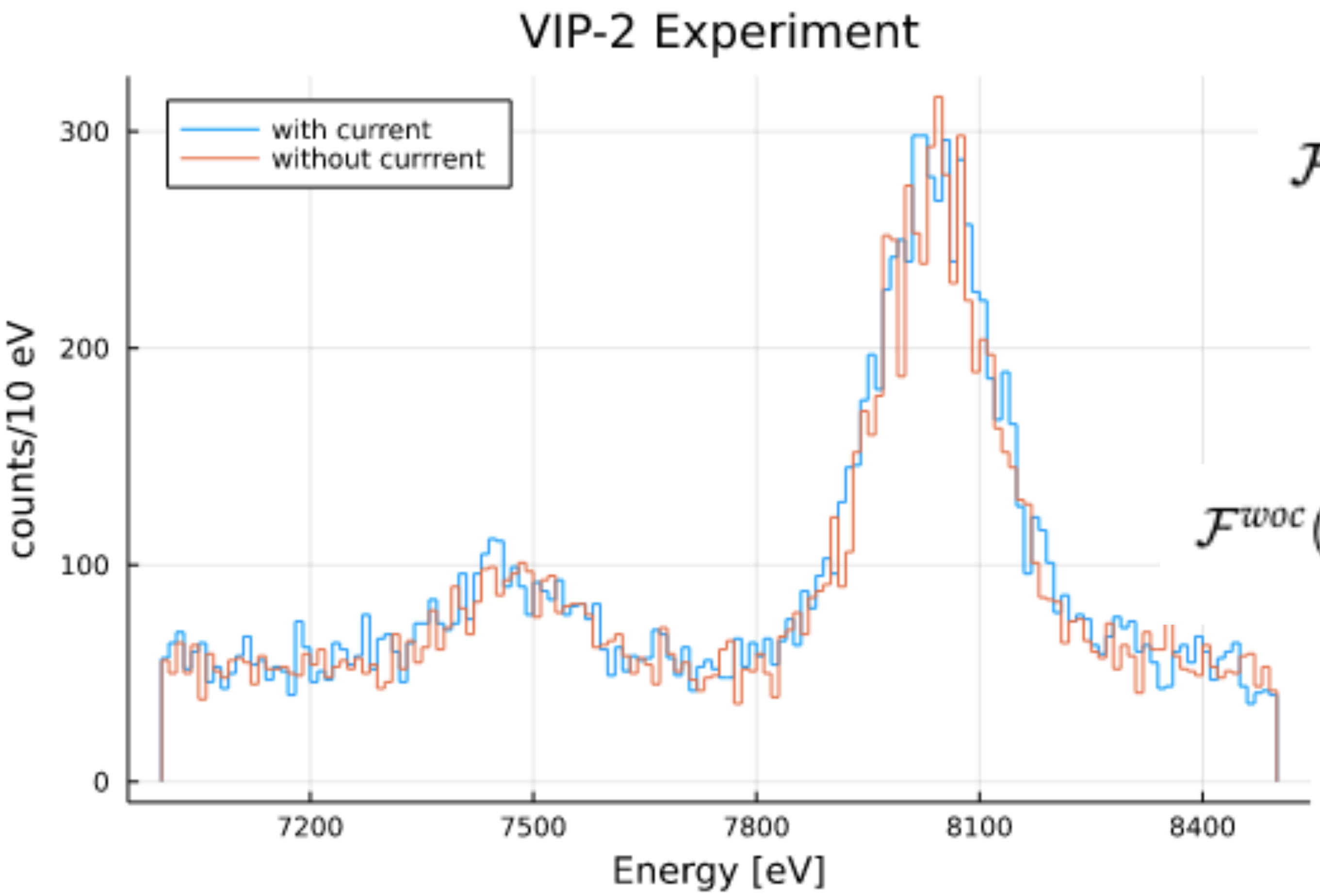


Passive scielding → **two layers, copper inside lead outside**



Symmetry **2022**, *14*(5), 893;
https://doi.org/10.3390/sym14050893

Six months of data taking



Description spectrum with current

$$\mathcal{F}^{wc}(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}) = y_1 \times Ni(\theta_1,\theta_2) + y_2 \times Cu(\theta_3,\theta_4) + \\ + y_3 \times \mathrm{pol}_1(\theta_5) + \mathcal{S} \times PEPV(\theta_4)$$

Description spectrum without current

$$\mathcal{F}^{woc}(\boldsymbol{\theta},\boldsymbol{y}) = y_1 \times Ni(\theta_1,\theta_2) + y_2 \times Cu(\theta_3,\theta_4) + y_3 \times \mathrm{pol}_1(\theta_5)$$

Likelihood

$$\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}) = \mathrm{Poiss}(\mathcal{D}^{wc}|\mathcal{F}^{wc}(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})) \times \mathrm{Poiss}(\mathcal{D}^{woc}|\mathcal{F}^{woc}(\boldsymbol{\theta},\boldsymbol{y}\times\mathcal{R}))$$

Bayesian

$$p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S}|\mathcal{D}^{wc},\mathcal{D}^{woc}) = \\ = \frac{\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})}{\int d\boldsymbol{\theta}d\boldsymbol{y}\mathcal{L}(\mathcal{D}^{wc},\mathcal{D}^{woc}|\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})p(\boldsymbol{\theta},\boldsymbol{y},\mathcal{S})}.$$

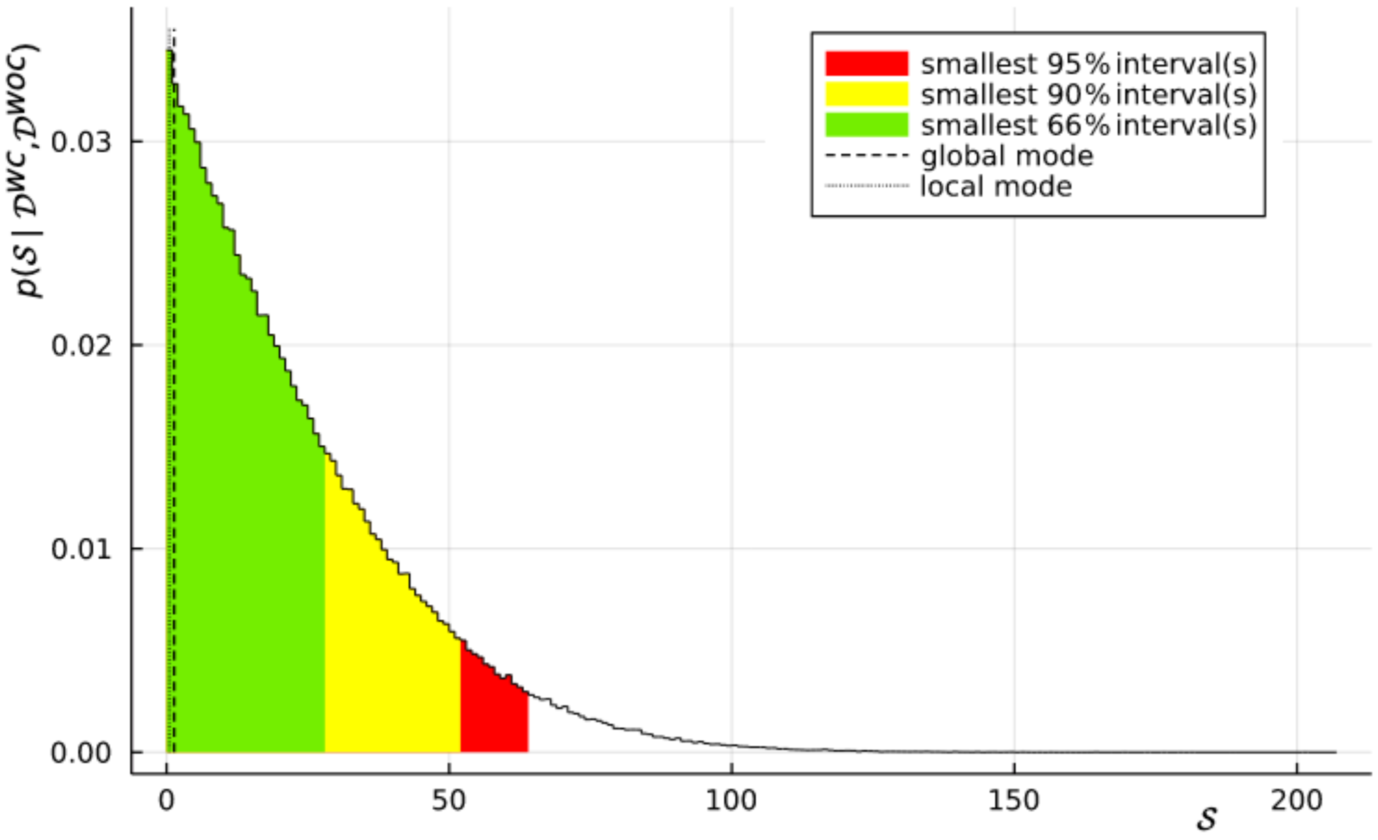
Frequentist

$$t_{\mathcal{S}} = -2\ln\Lambda(\mathcal{S}) = -2\ln\frac{\mathcal{L}(\hat{\hat{\boldsymbol{\theta}}},\hat{\hat{\boldsymbol{y}}},\mathcal{S})}{\mathcal{L}(\hat{\boldsymbol{\theta}},\hat{\boldsymbol{y}},\hat{\mathcal{S}})}$$

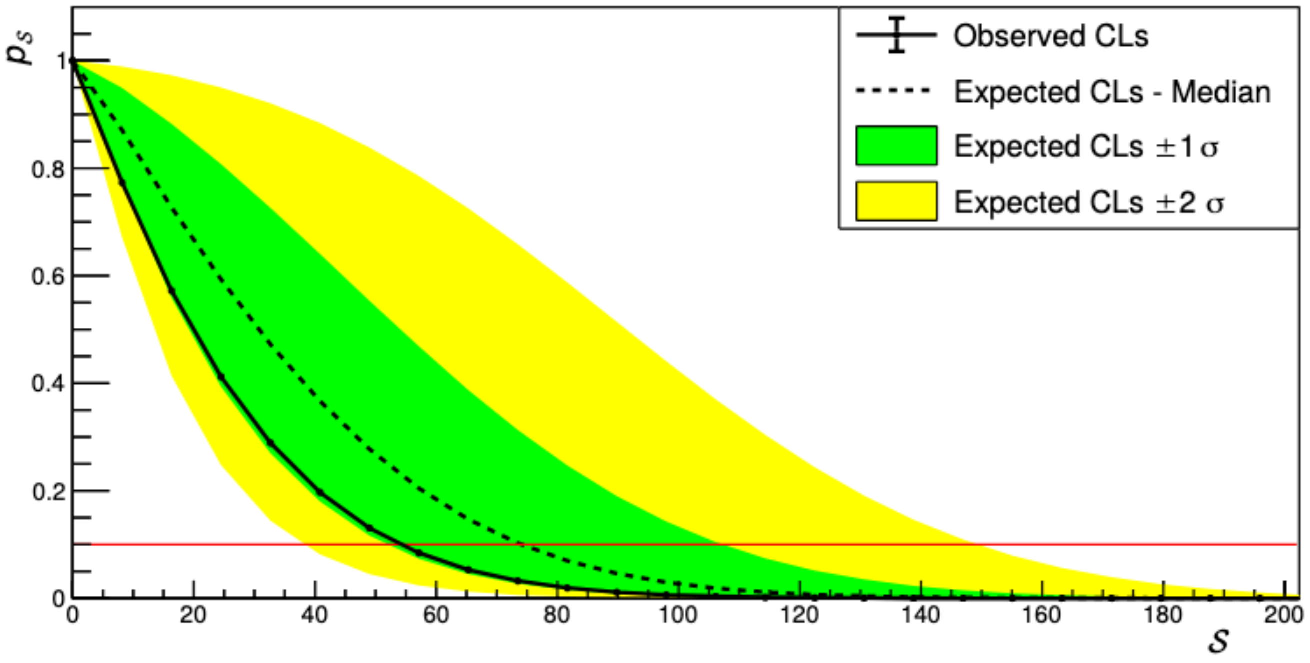
$$p_{\mathcal{S}} = \int_{t_{obs}}^{\infty} f(t_{\mathcal{S}}|\mathcal{S})dt_{\mathcal{S}}$$

Symmetry **2022**, *14*(5), 893;
https://doi.org/10.3390/sym14050893

Bayesian



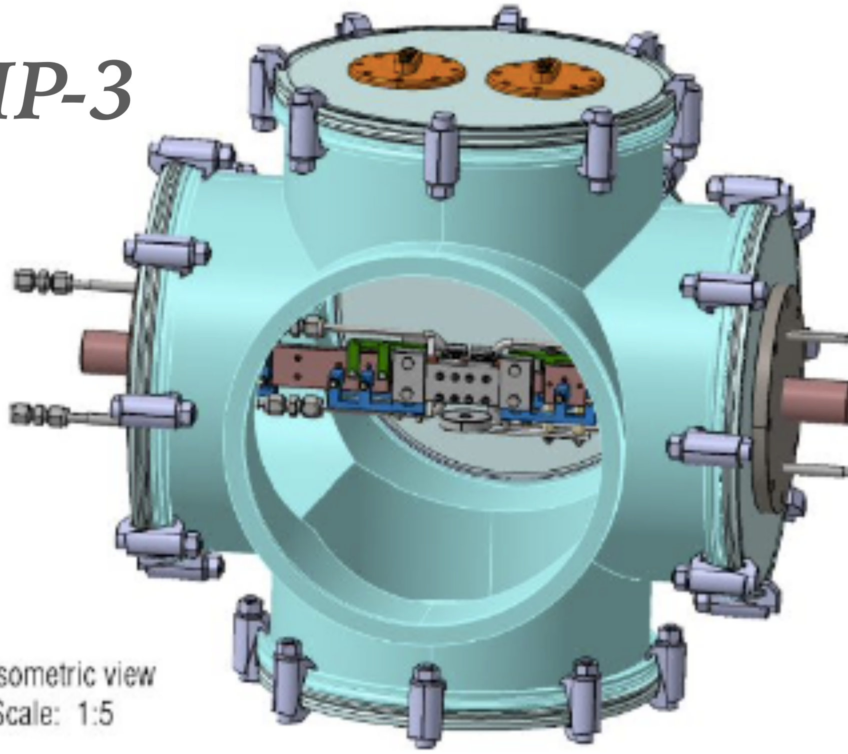
Frequentist



$$\beta^2/2 \leq 8.6 \times 10^{-31} \text{ (Bayesian)}, \qquad \beta^2/2 \leq 8.9 \times 10^{-31} \text{ (CL}_s\text{)}$$

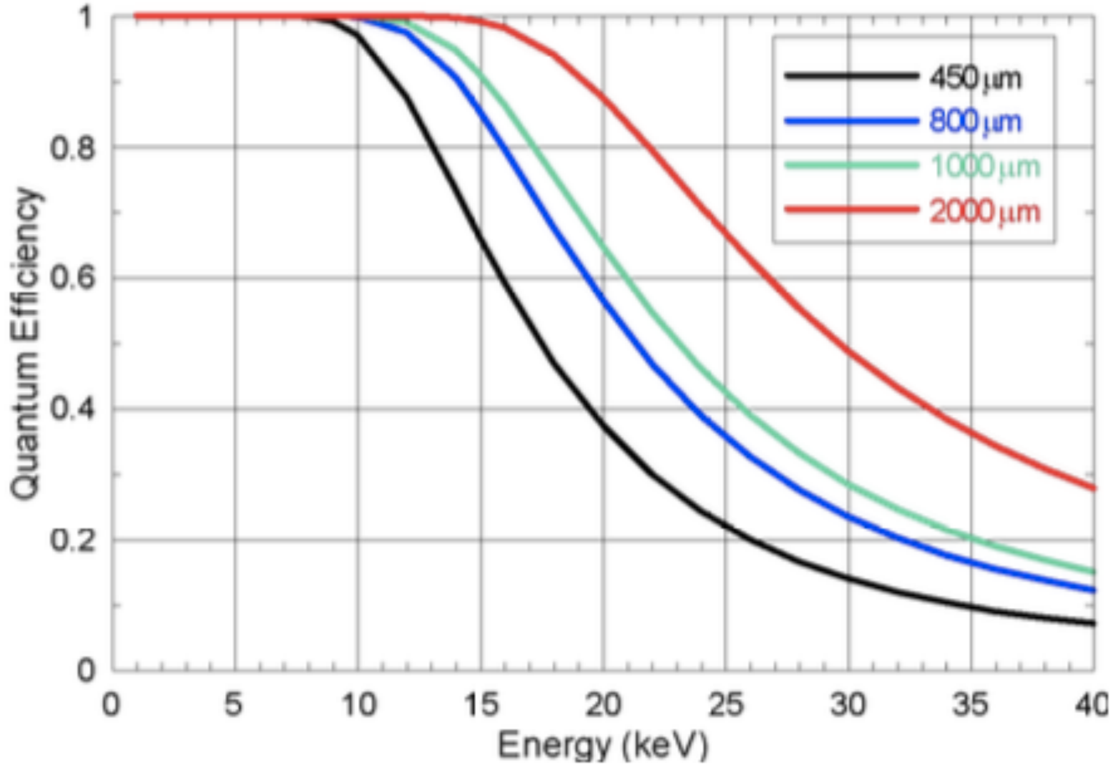
New article in preparation with all the available statistics!

VIP-2 experimental upgrade: VIP-3



- ***new vacuum chamber,*** *increase the number of SDD detectors, increase the geometrical efficiency, higher current up to* ***400 A***
- *New thermal contact between cold finger and SDDs*
- *New target cooling system*

- ***Higher quantum efficiency needed for the SDDs at higher Z: use 1 mm thick SDDs, allowing to scan e.g. Ag, Sn and Pd***



- ***2x4 SDDs, 8x8 mm² each, in production with FBK & politecnico di Milano***