Studies on detector optimisation through end-to-end surrogate models including discrete parameters

24.9.2024

Karlsruher Institut für Technologie

Nikhil Kota, Jan Kieseler, Kylian Schmidt (KIT)

Institut für Experimentelle Teilchenphysik

Studies in collaboration with:

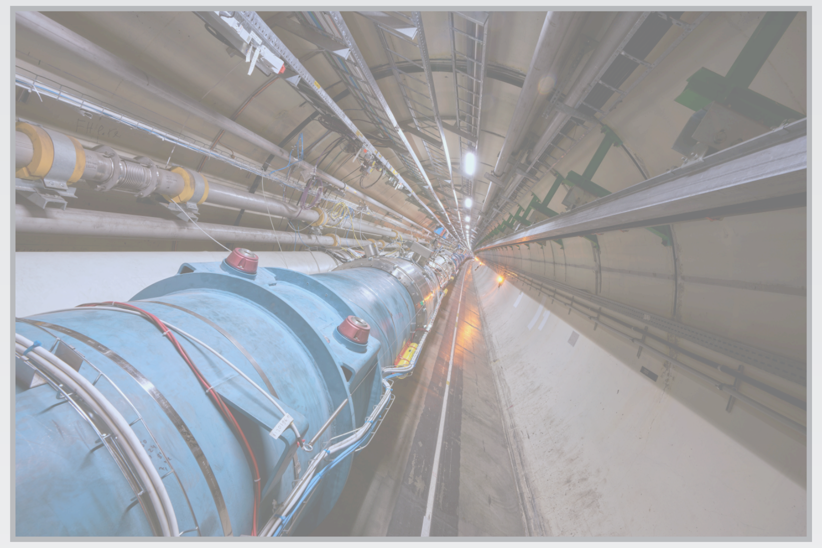
Tommaso Dorigo (INFN Padova and LTU), Enrico Lupi (INFN Padova and University of Padova), Xuan-Tung Nguyen (INFN Padova and RPTU), Joseph Willmore (INFN Padova), Max Aehle, Nicholas R. Gauger, Alexander Schilling Tobias Kortus Ralf Keidel (RPTU), Fredrik Sandin (LTU)

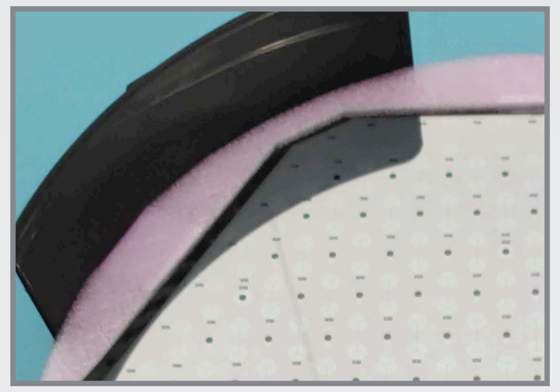
MODE

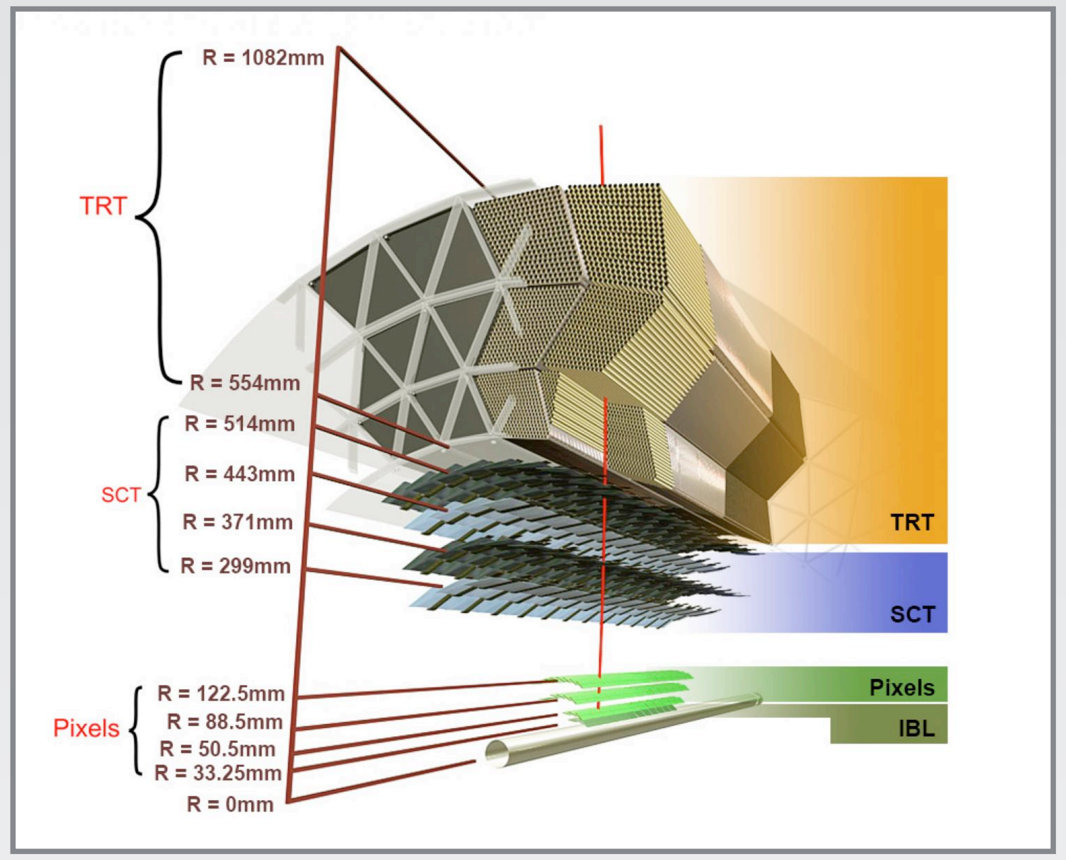
and

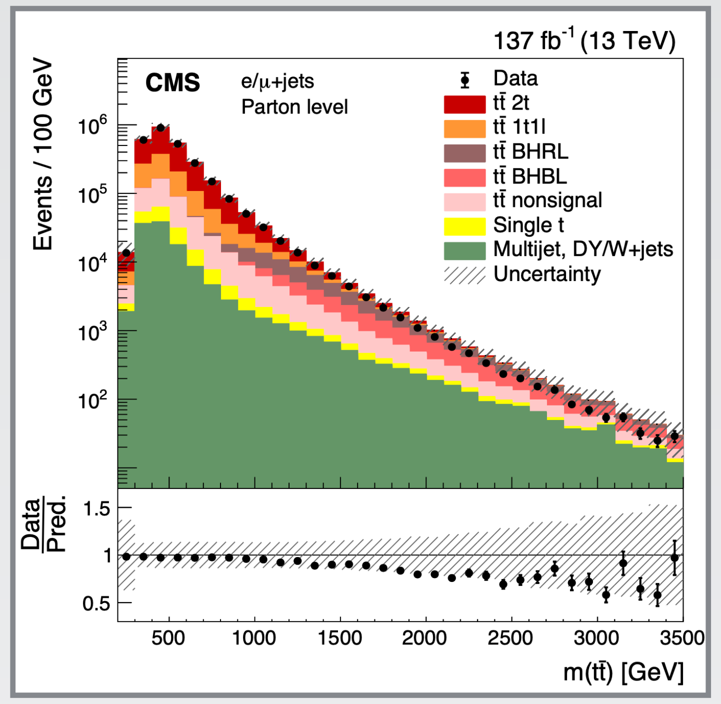
Francois Fleuret, Tobias Golling, Stephen Mulligan, Atul Sinha, Kinga Wozniak (U. Geneva)

The usual chain







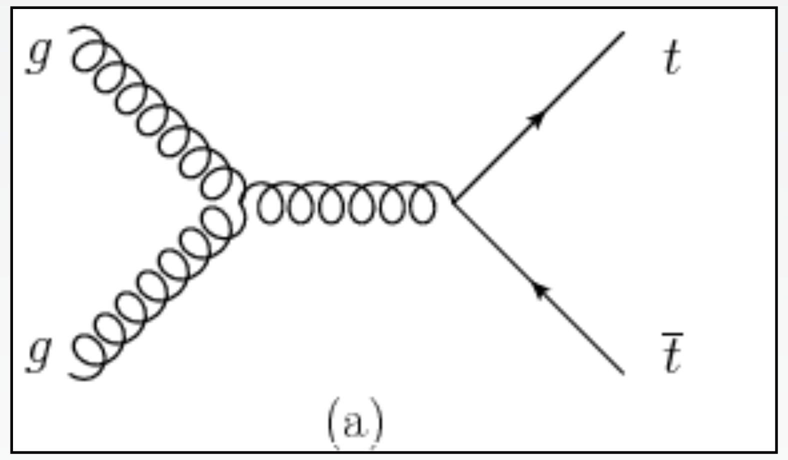


Collider / "Space"

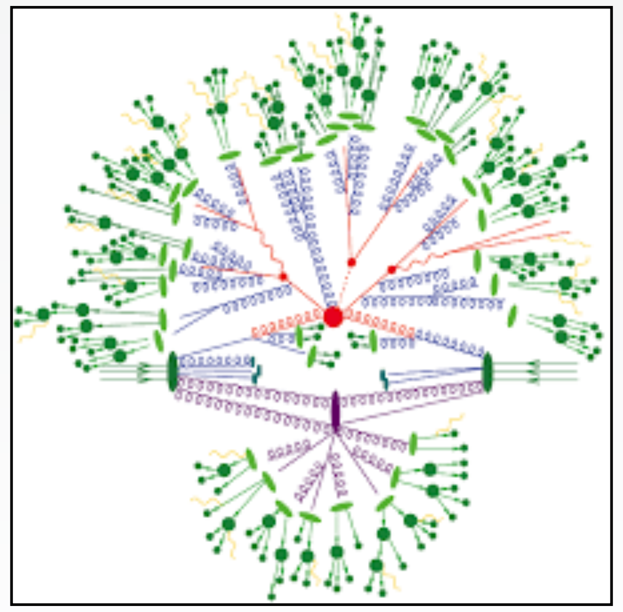
Sensors Electronics Geometry

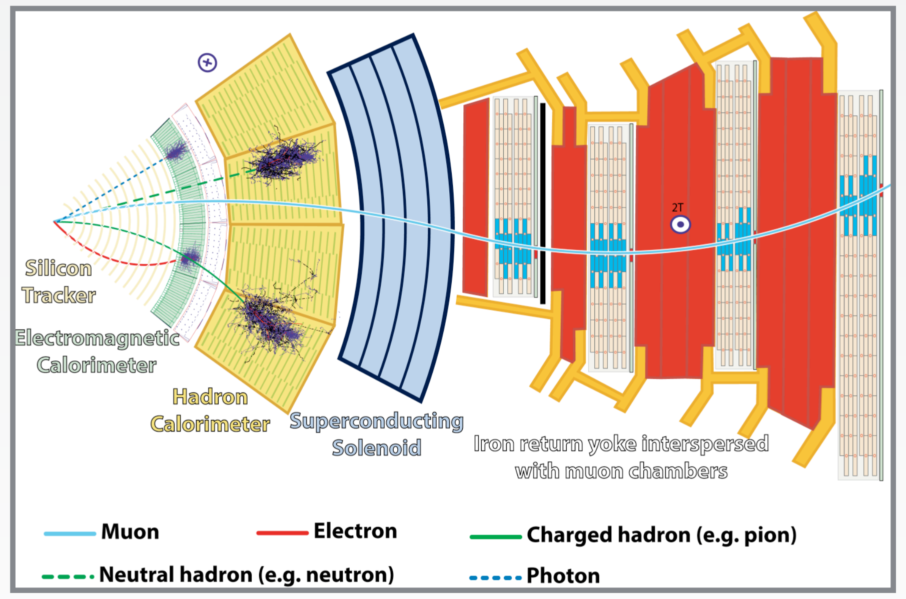
Recon-struction

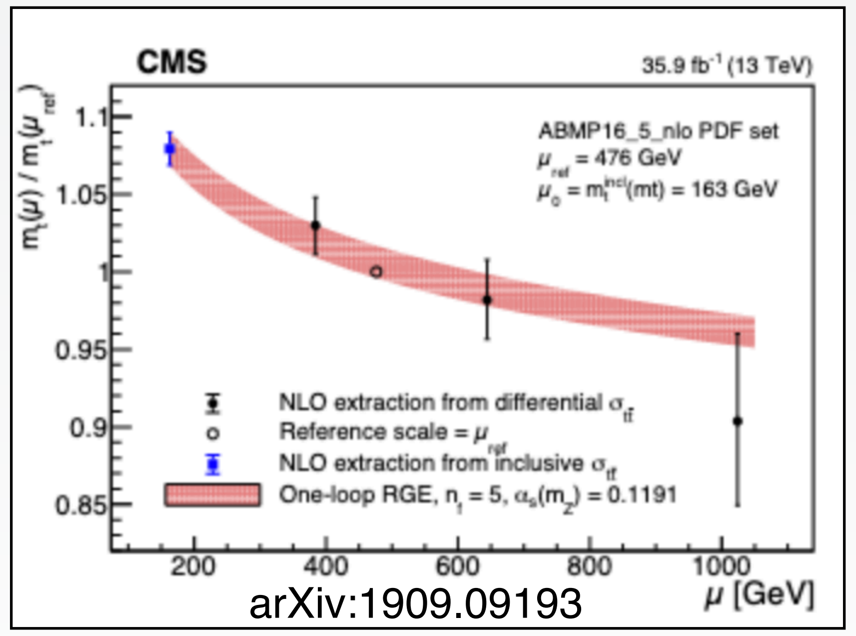
Analysis



Physics

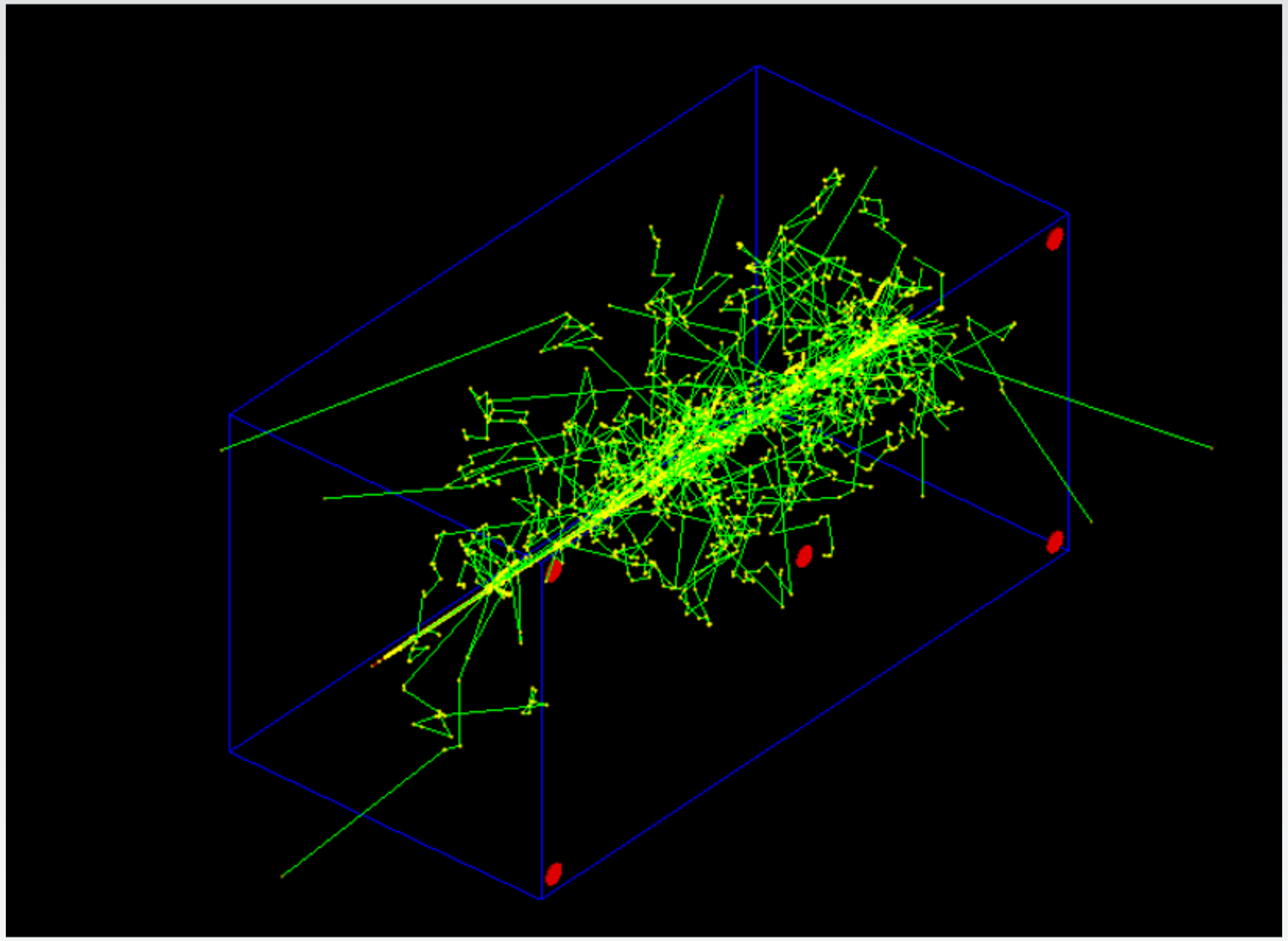






What happened here?

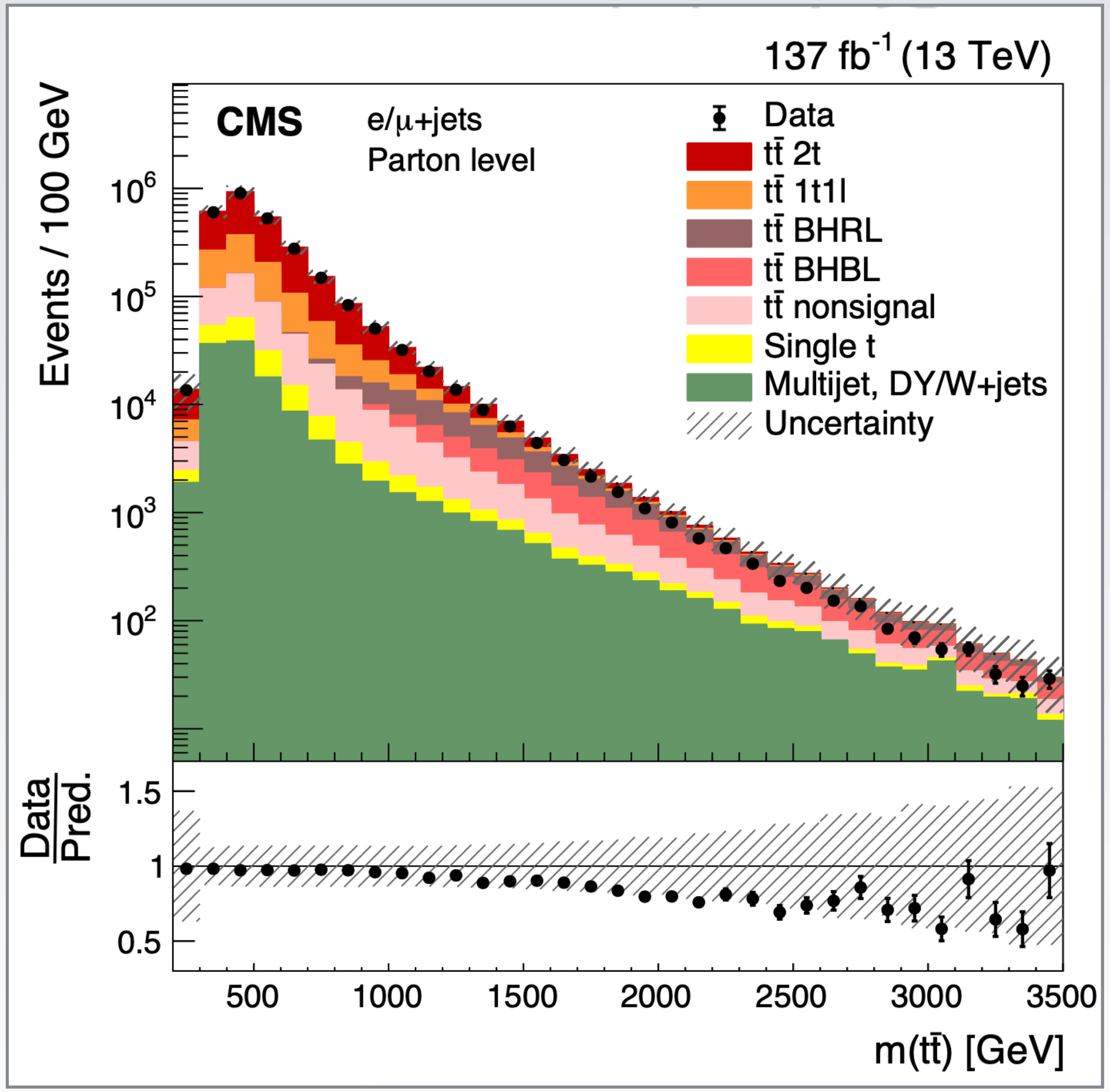
Our simulation is fantastic



A particle being stopped in dense material

- High fidelity simulation of particles interacting with matter
- Carefully validated

- Validity also spans orders of magnitude
- Not* differentiable



http://arxiv.org/abs/2108.02803

Complexity of the Problem

Collider / "Space"

Sensors
Electronics
Geometry

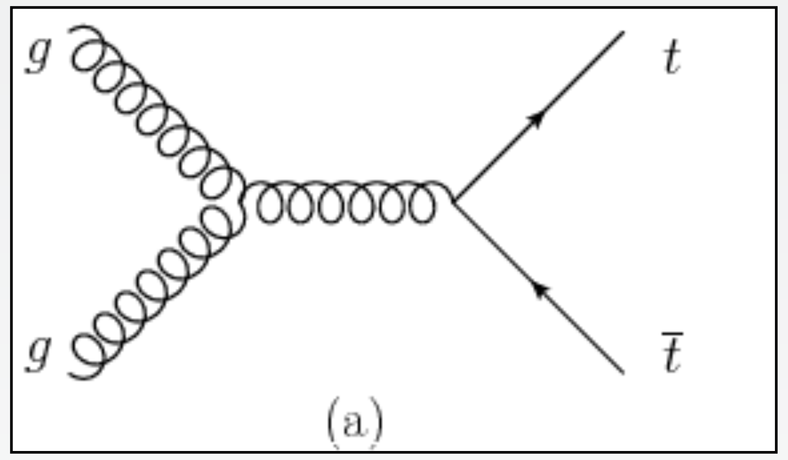
Recon-
struction

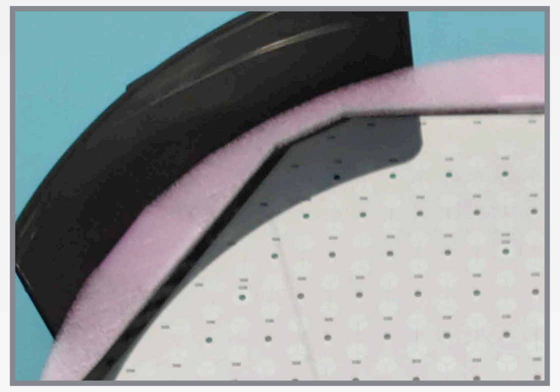
Analysis

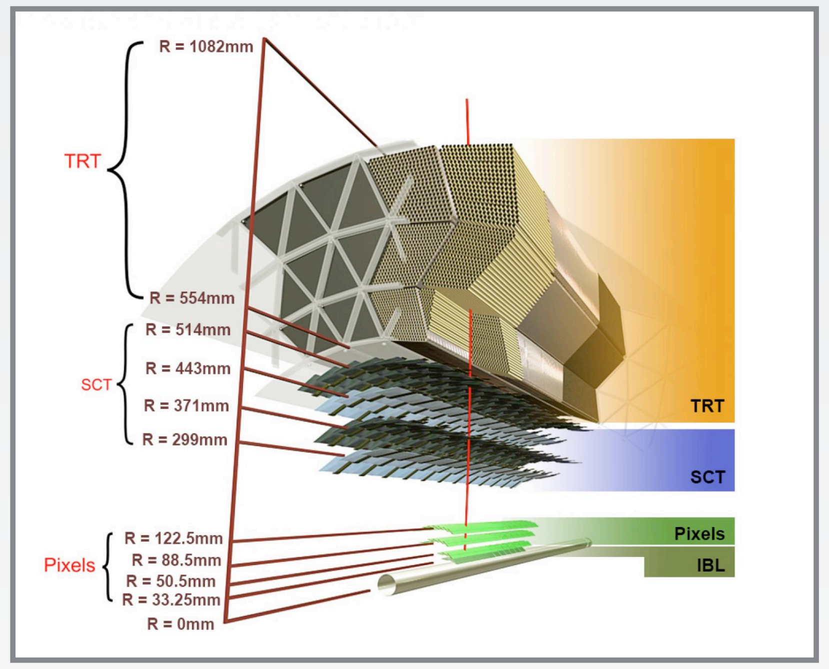
A few quantities

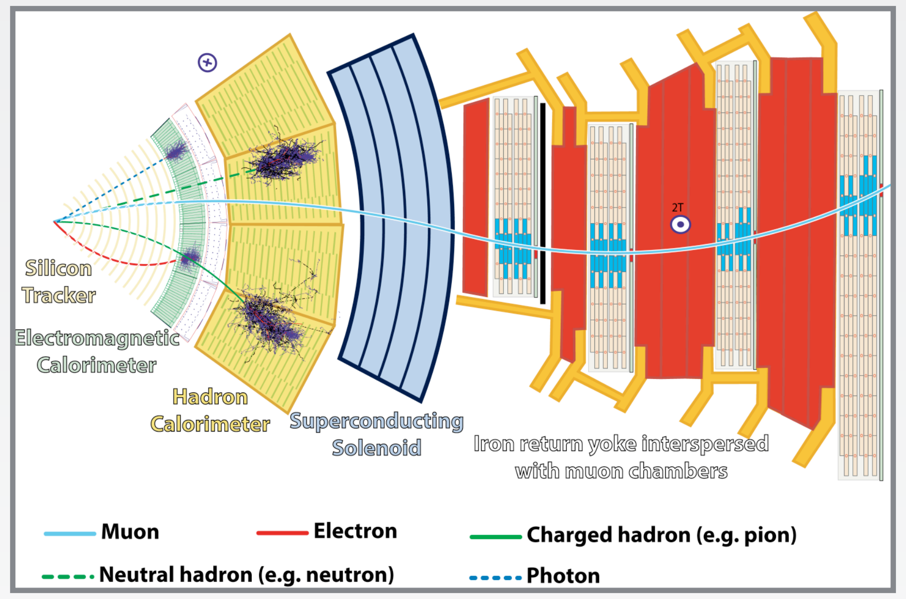
Many pixels, detector elements, hits

A few quantities

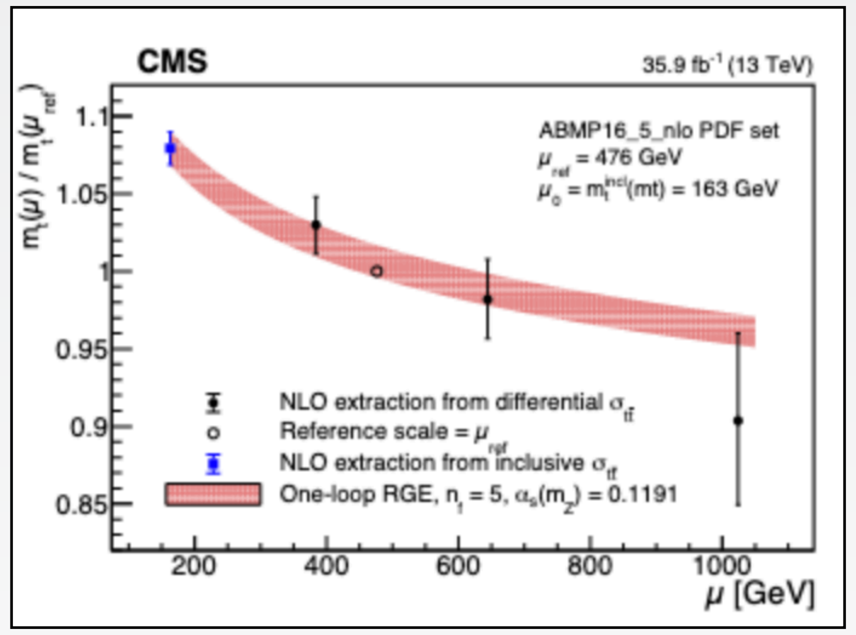




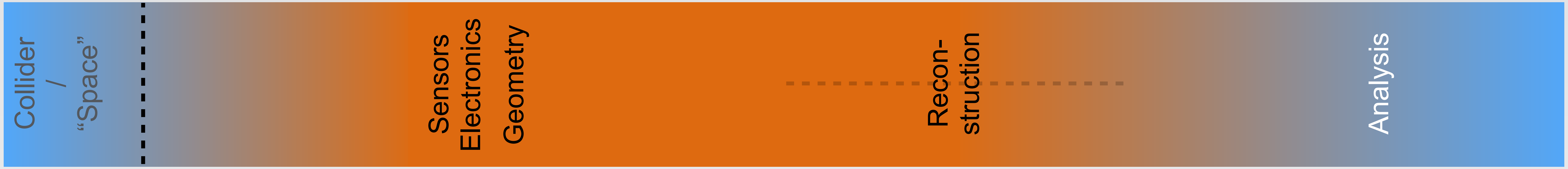




arXiv:1909.09193



Make Wider Steps?

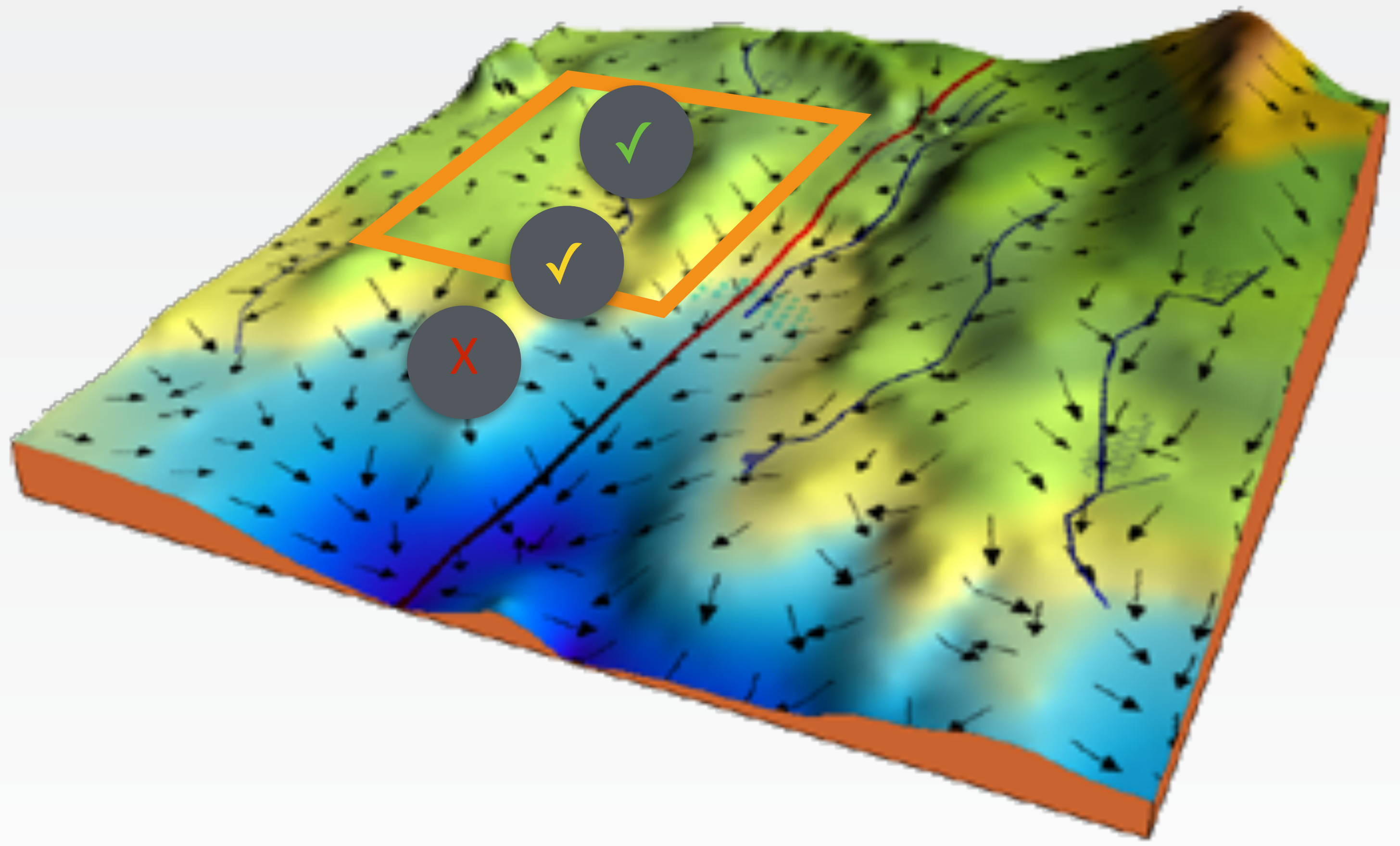


A few quantities

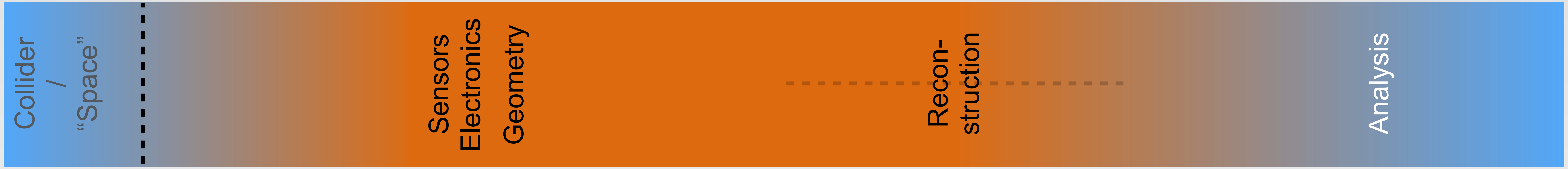
Many pixels, detector elements, hits

A few quantities

Very complex surrogate model



Make Wider Steps?



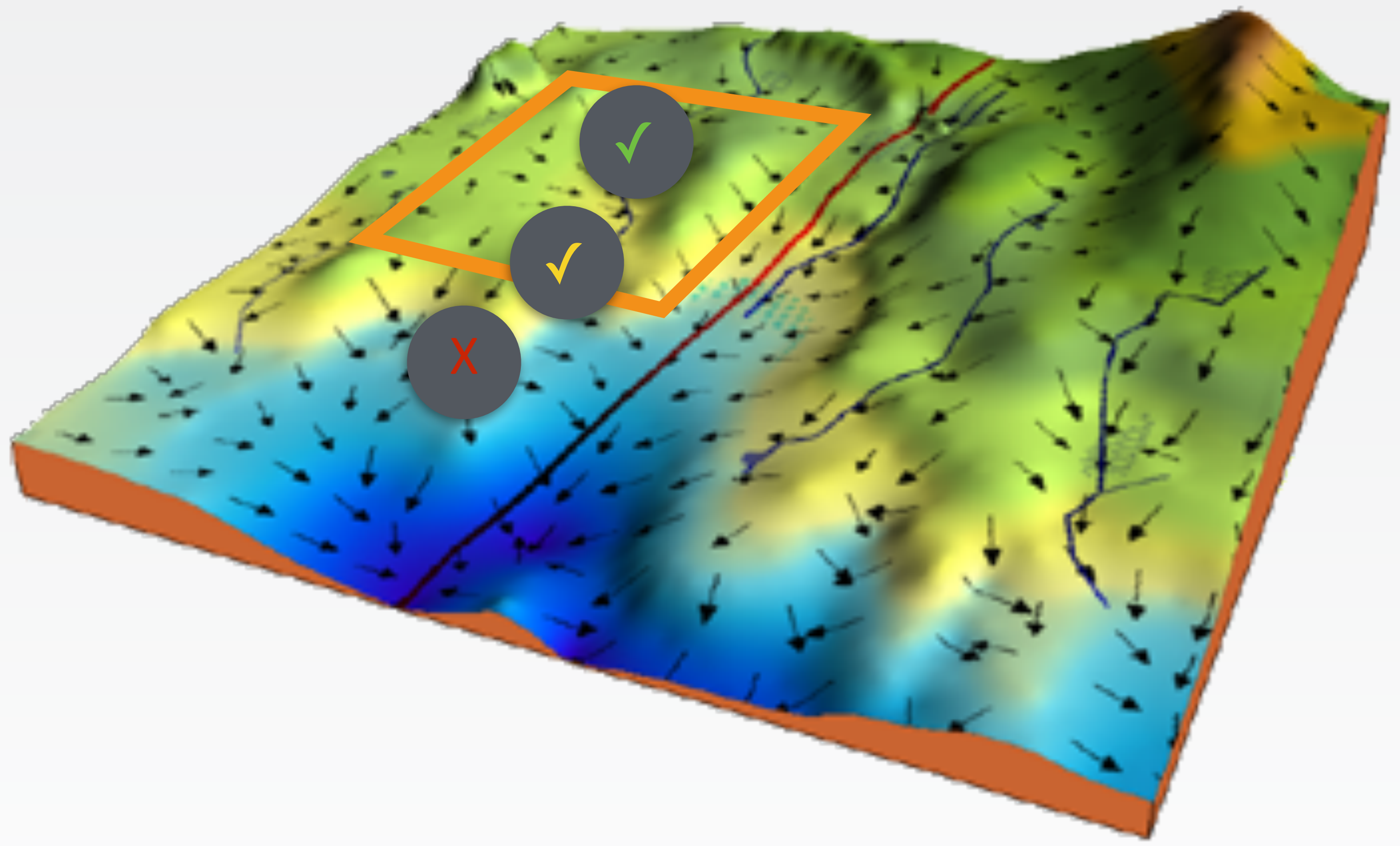
A few quantities

Many pixels, detector elements, hits

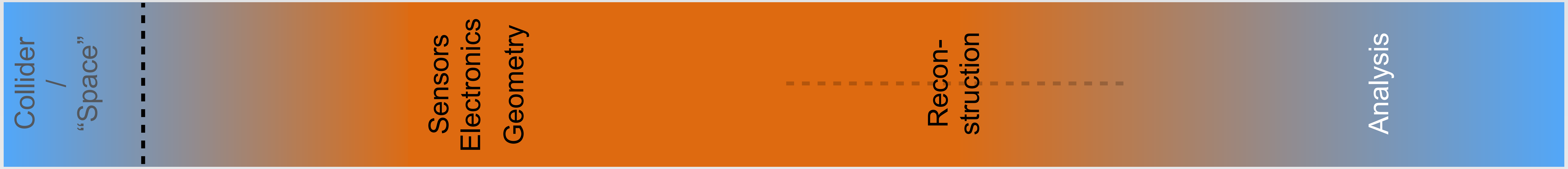
A few quantities

Very complex surrogate model

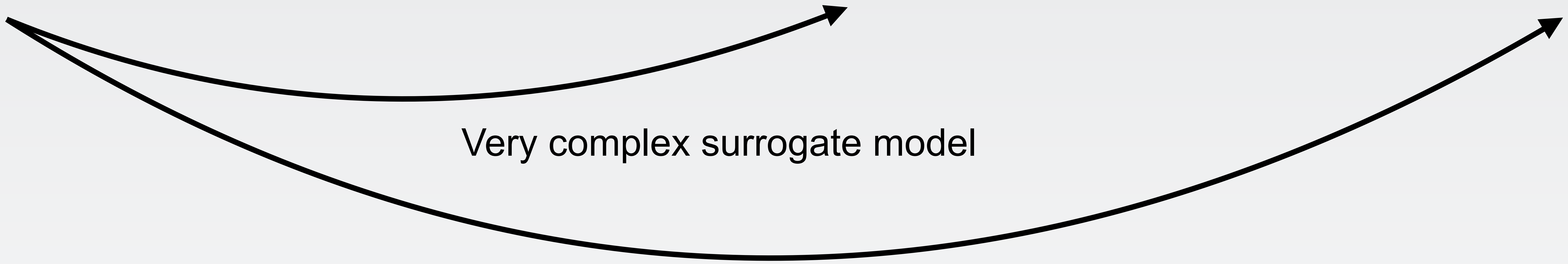
✓ Much simpler surrogate model



Make Wider Steps?

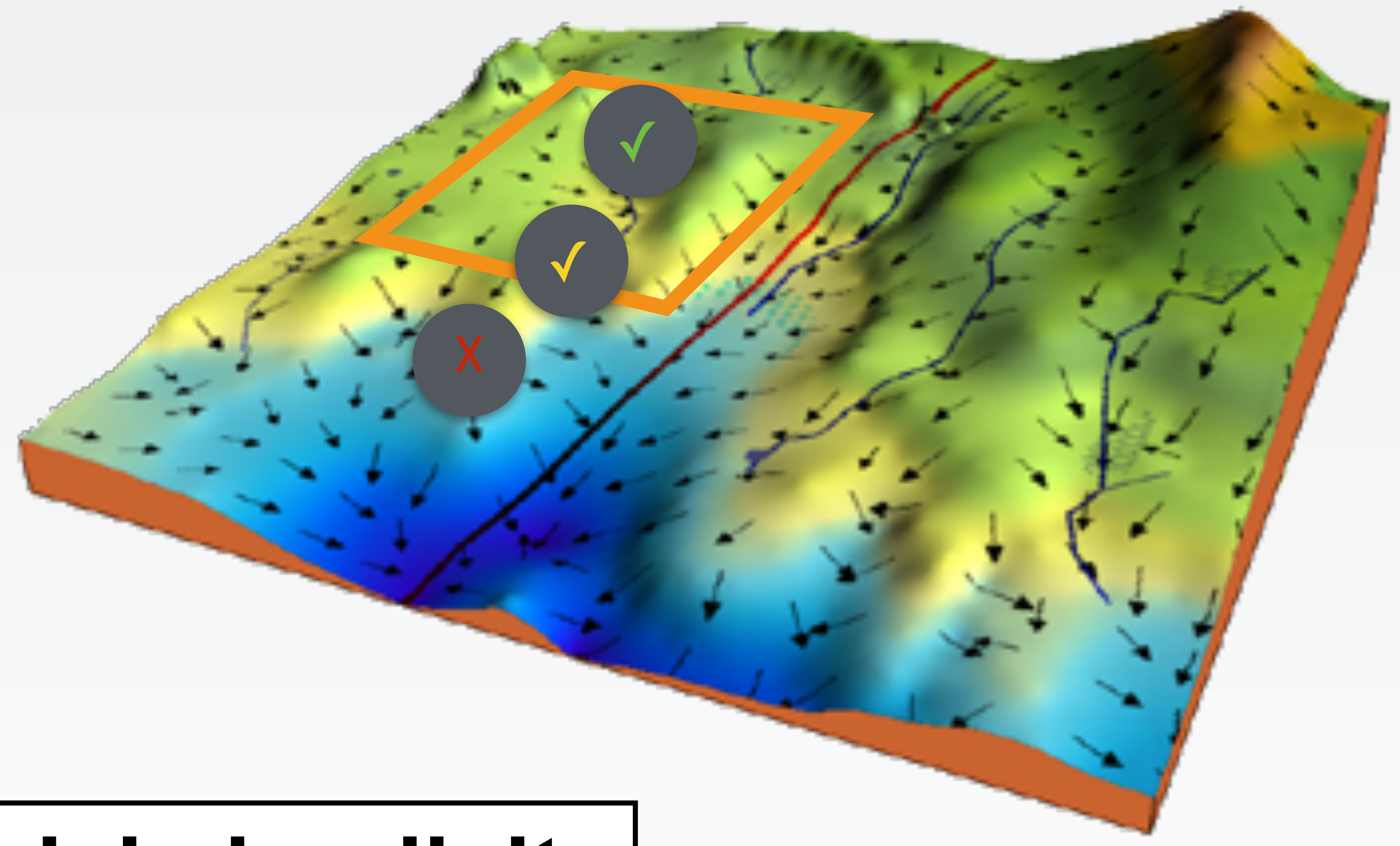


A few quantities | Many pixels, detector elements, hits | A few quantities



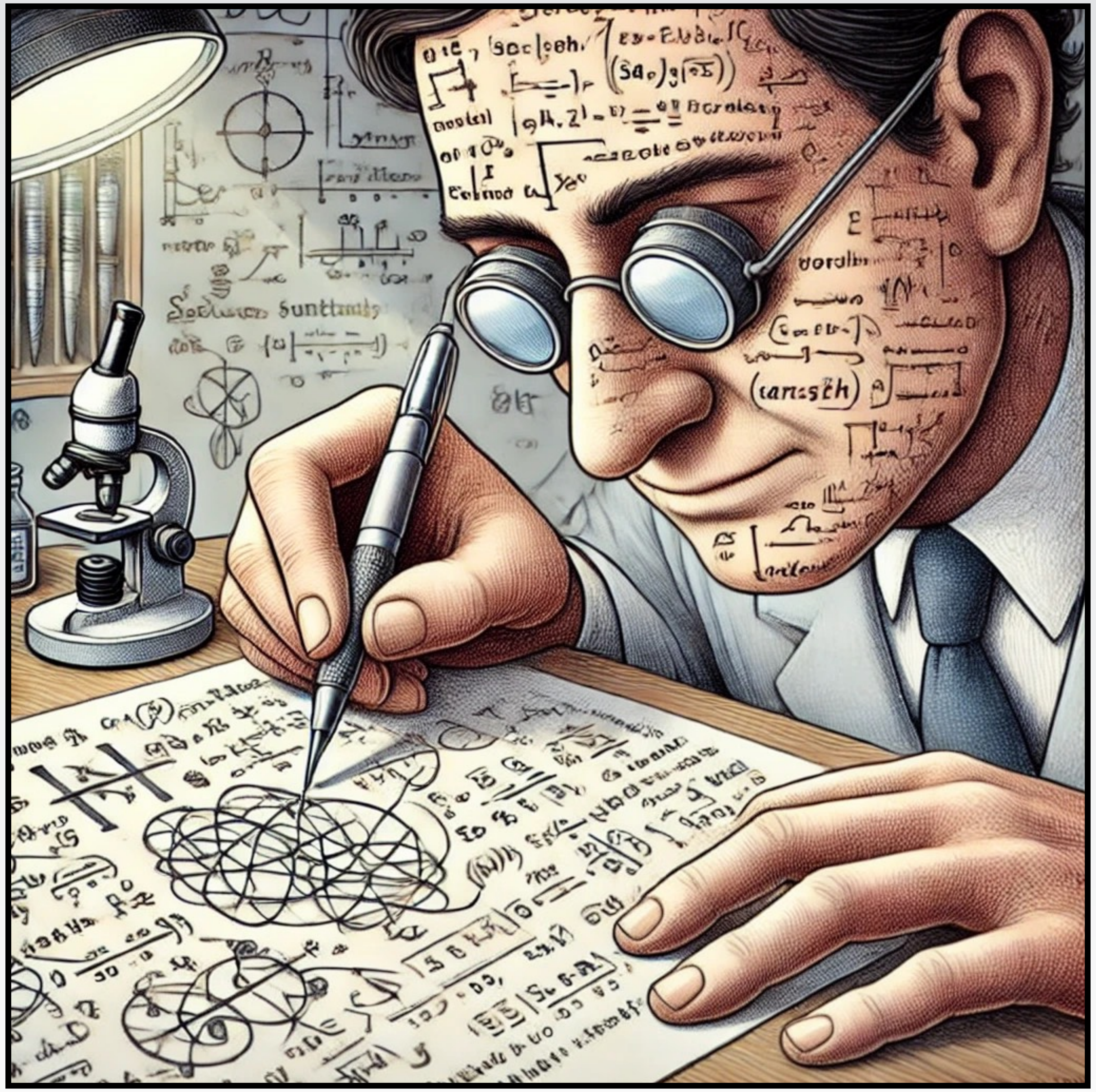
Very complex surrogate model

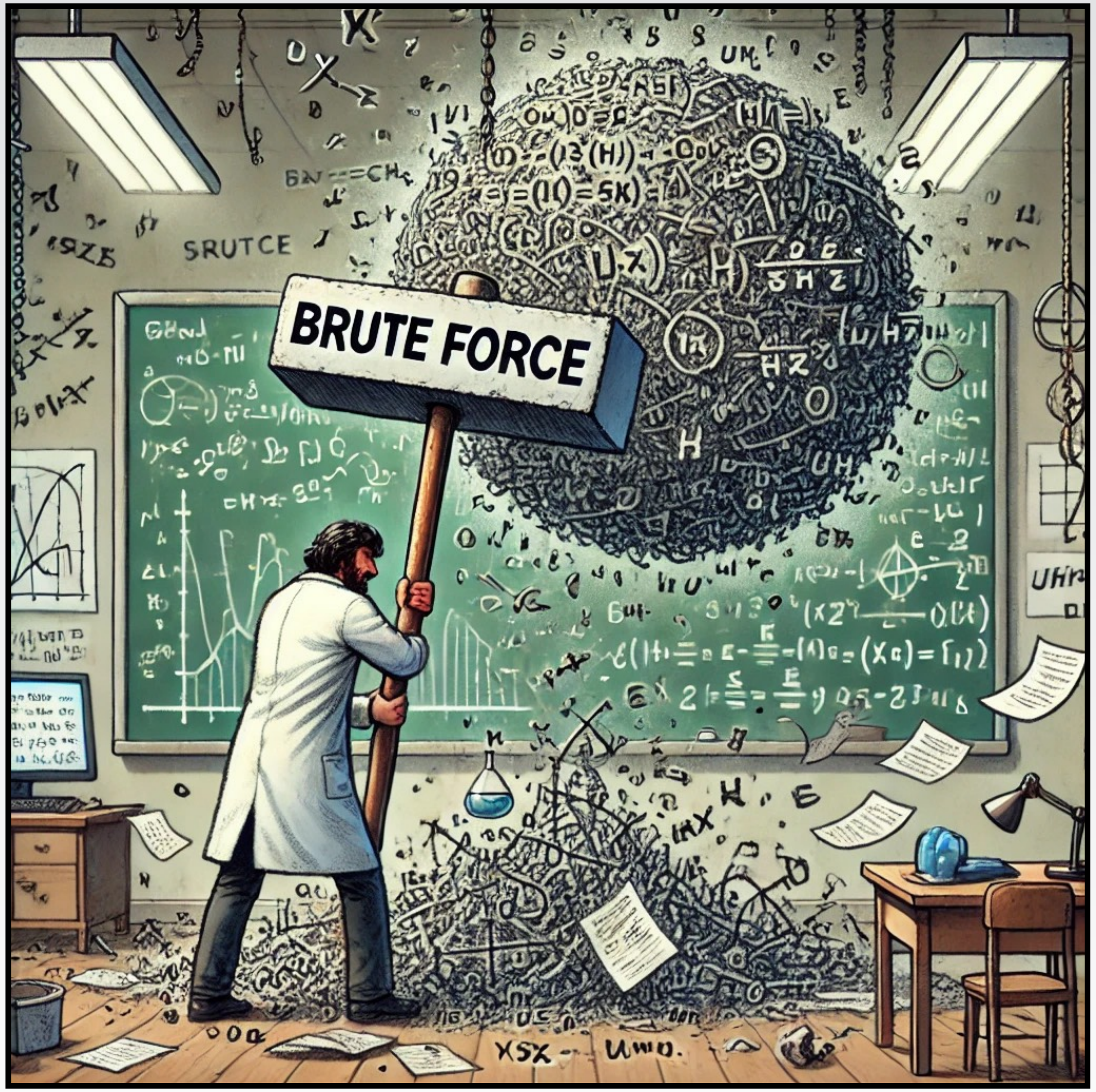
✓ Much simpler surrogate model



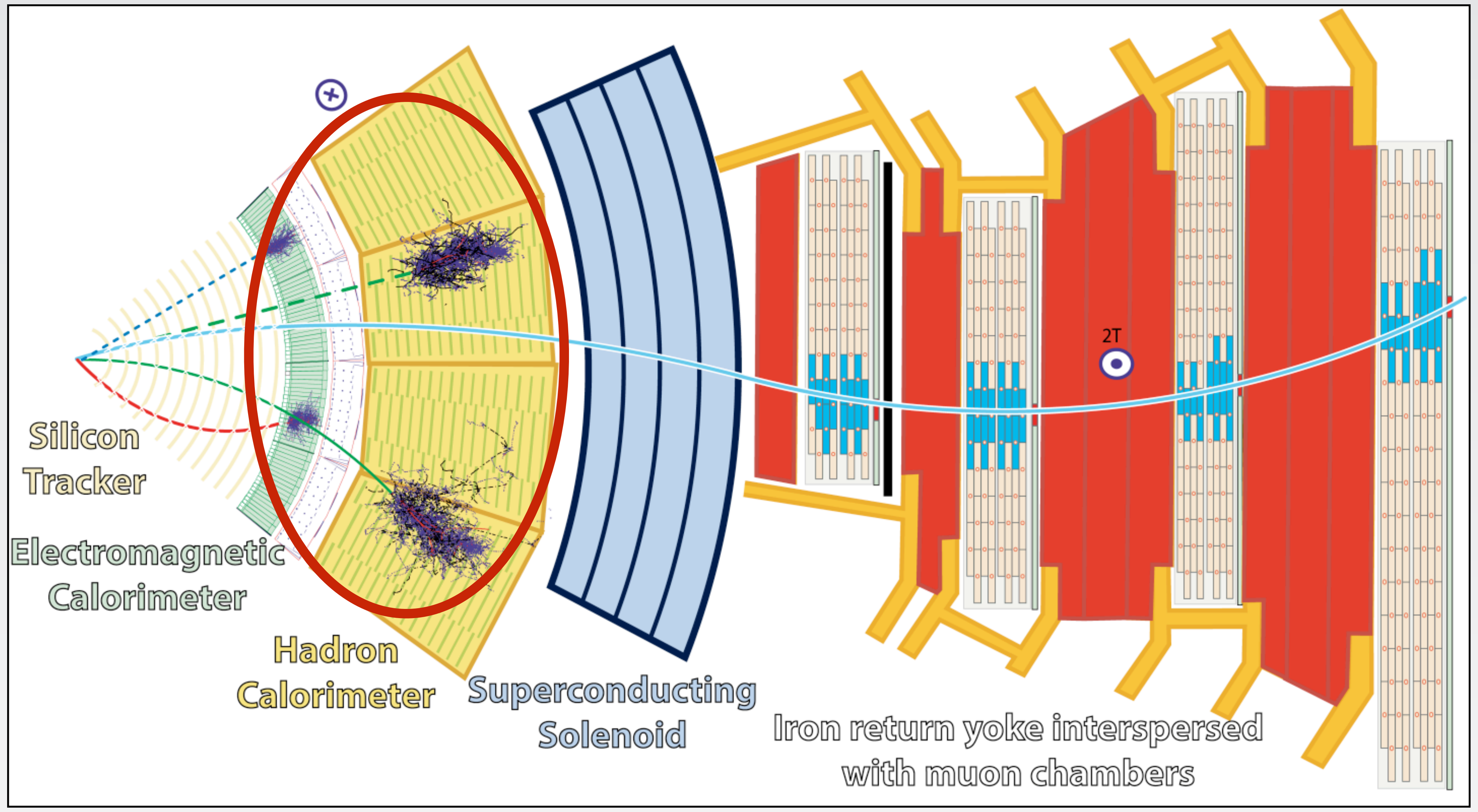
Trade off between computing resources and surrogate model simplicity

The type of approach

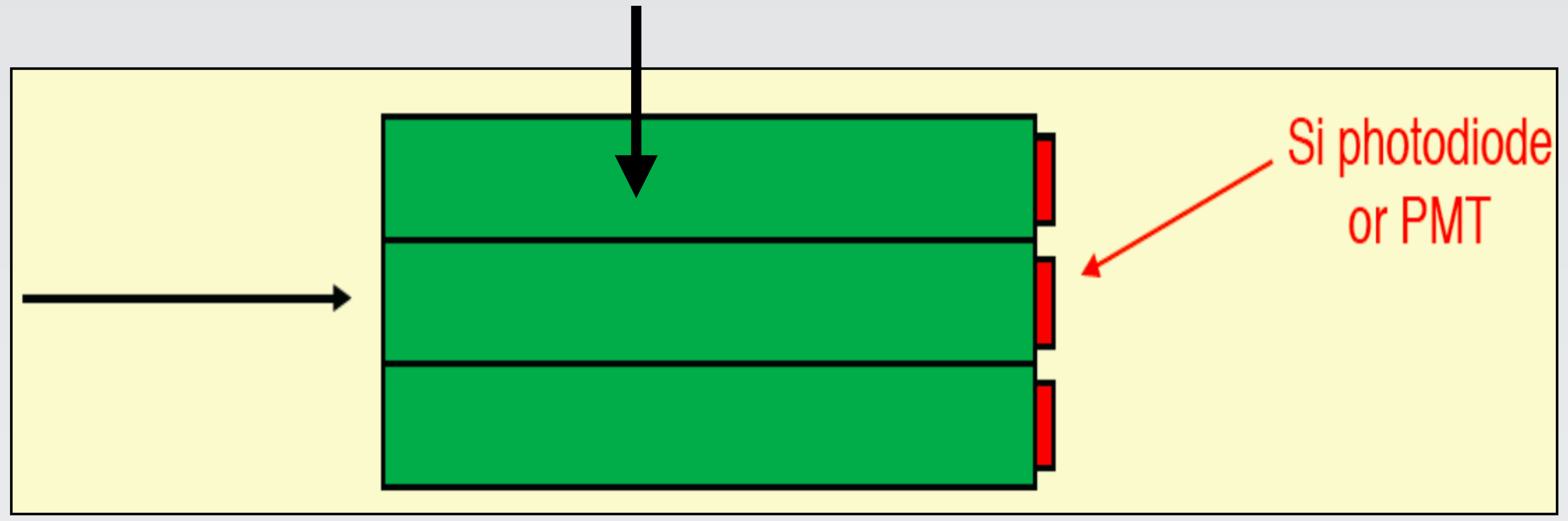




Calorimeters

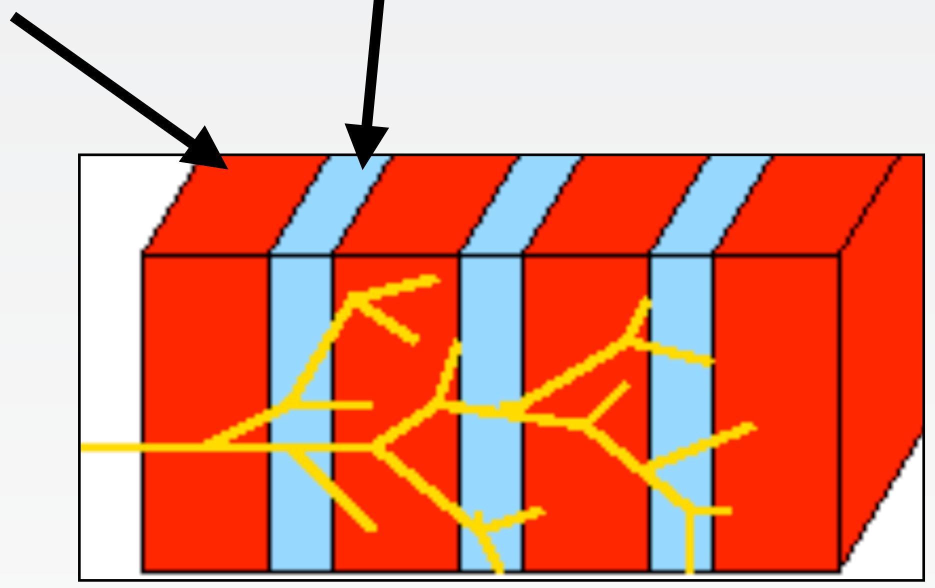


Absorb/stop and detect



Better physics performance

Absorb/stop
Detect (sensors)



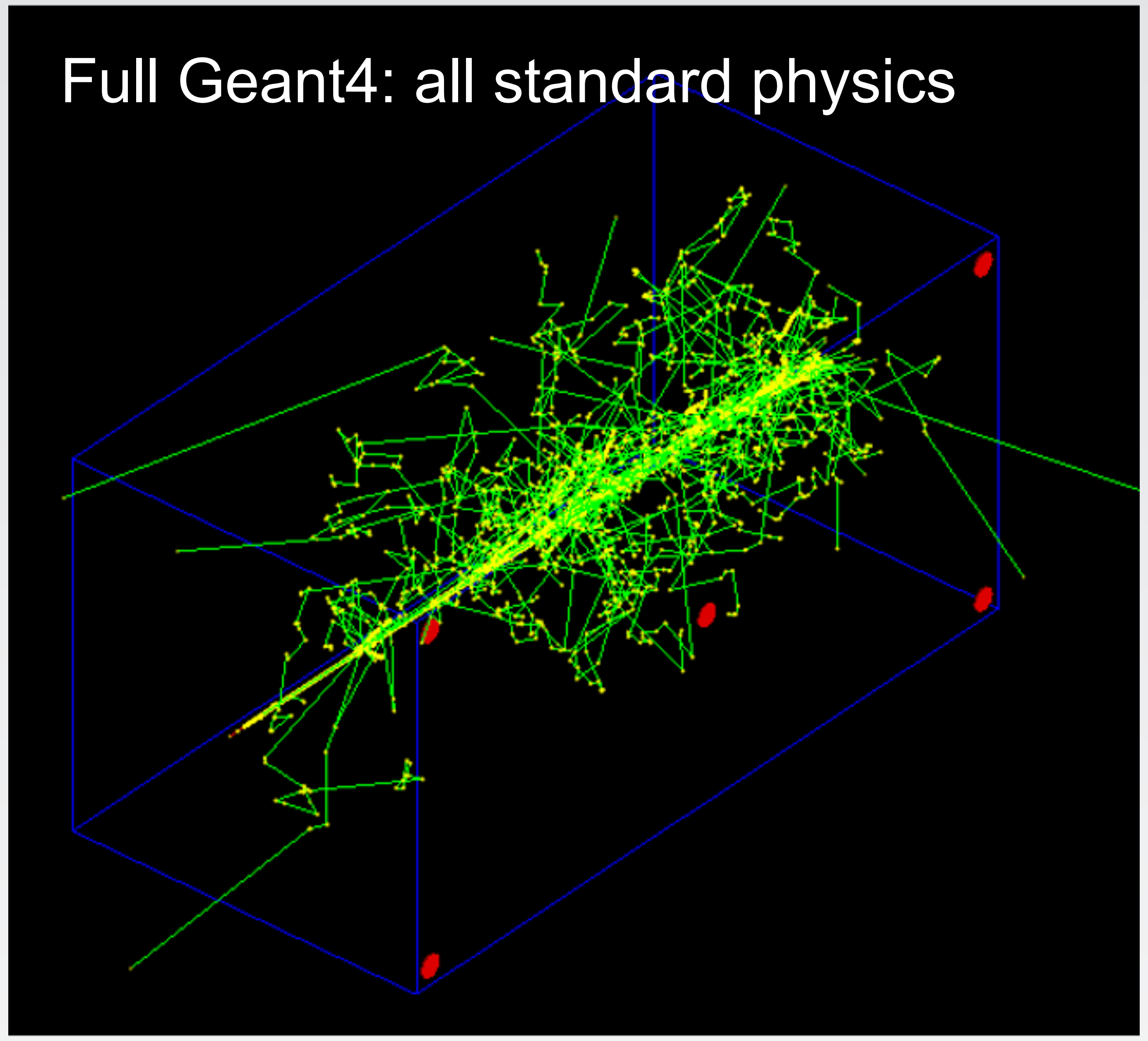
Much cheaper

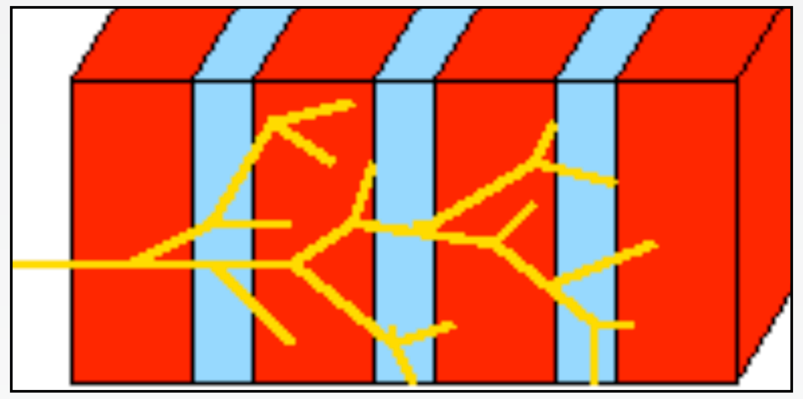
- Complex showers
- So far designs relatively simple
- Good place to invest in systematic gradient-based optimisation

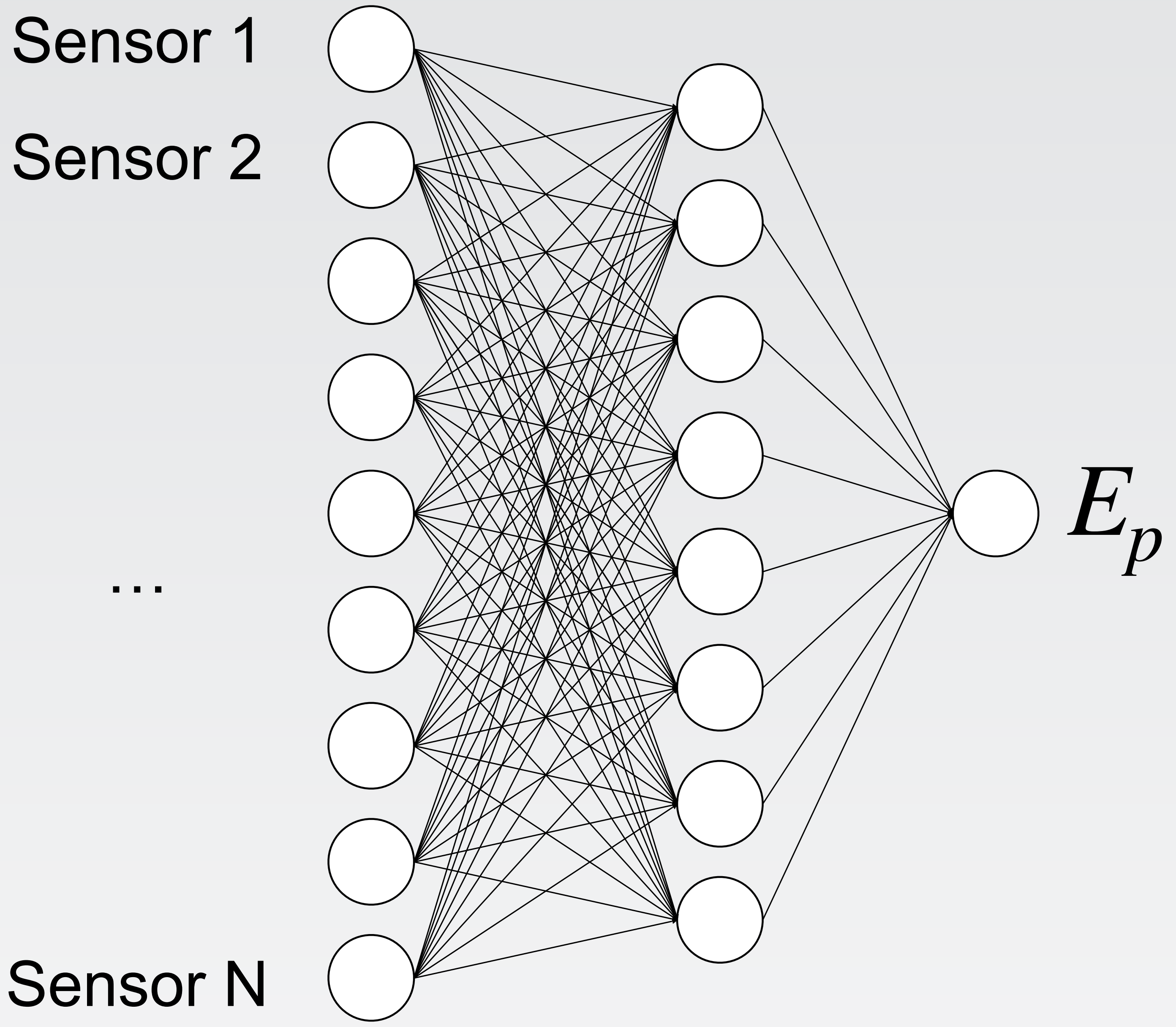
The simulation and reconstruction

```
double p = /* diff'able */;
if( rng->flat() < p ){
  // ... do something ...
}
```

E_t







E_t : e.g. single value

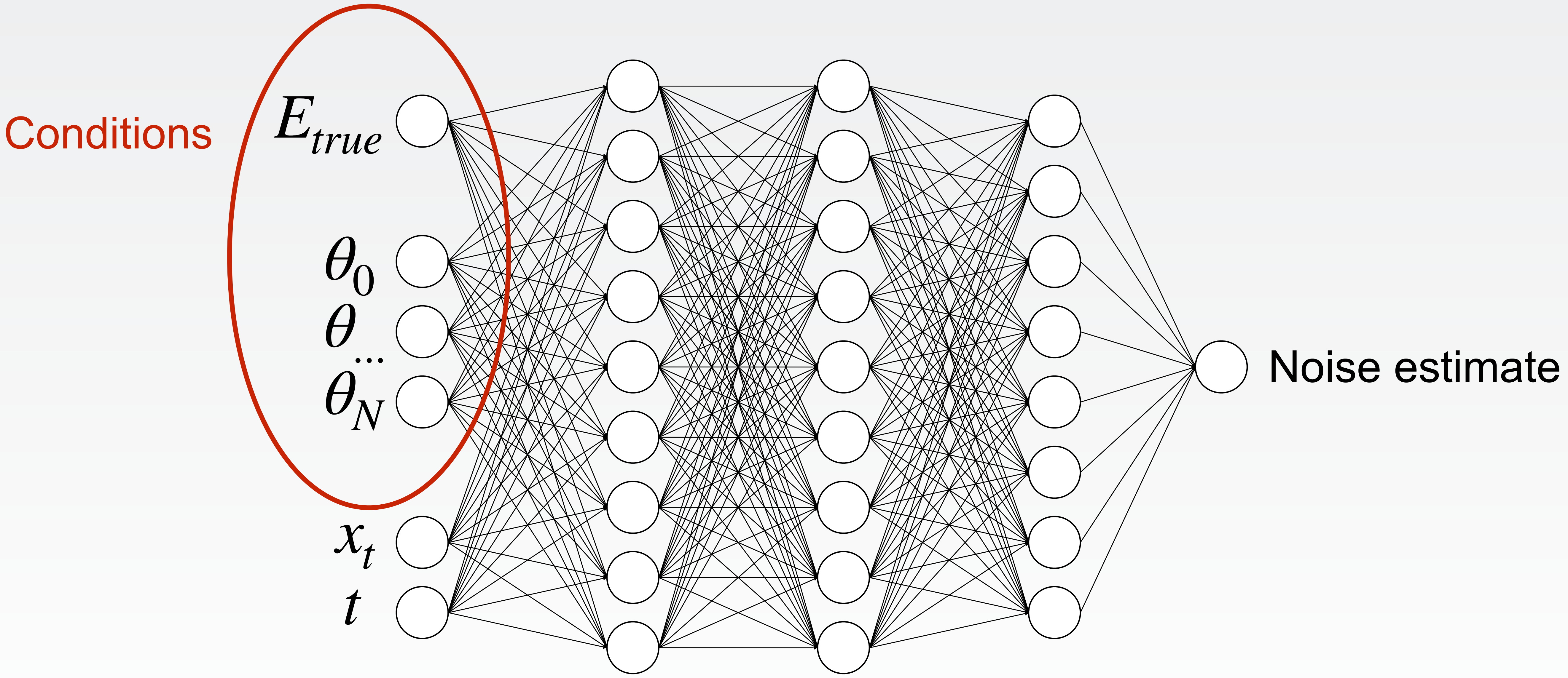
E_p/E_t : approx. Gaussian distribution (narrow = better)

A diffusion model as a surrogate

https://scholar.harvard.edu/binxuw/classes/machine-learning-scratch/materials/foundation-diffusion-generative-models

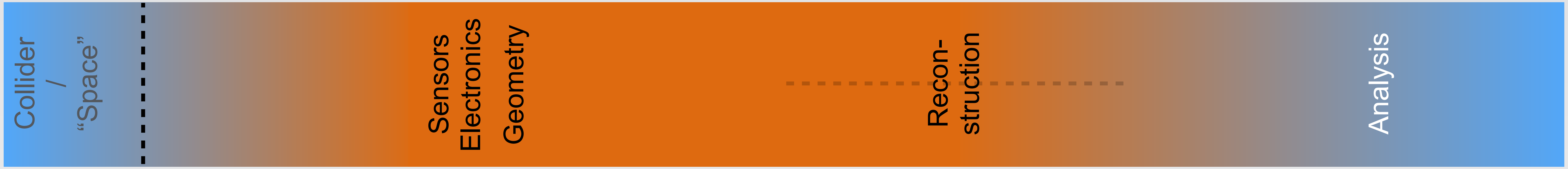
Adding conditioning

https://scholar.harvard.edu/binxuw/classes/machine-learning-scratch/materials/foundation-diffusion-generative-models



- Sampling over time steps ($O(100)$) in principle slow
- Very low-dimensional problem here
- Fast enough and **very easy to train**

A pipeline



Single particles
(Photons and hadrons)
E = [1,20] GeV

Full Geant4 Simulation (θ)
- Layer material*
- Layer thicknesses

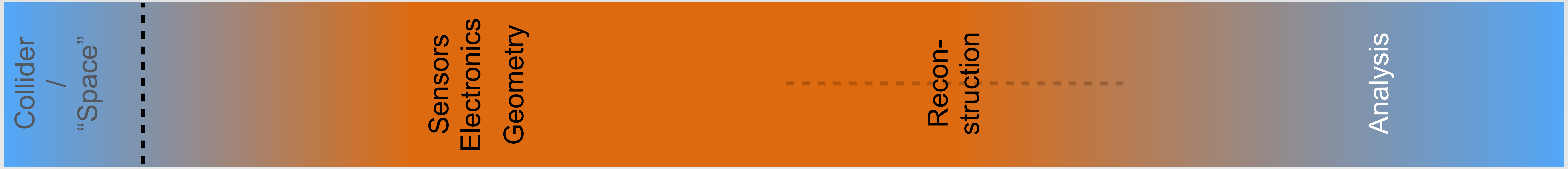
Simple DNN reconstruction
- Has access to θ

Objective:

$$L = \frac{(E_p - E_t)^2}{E_t}$$

Simple diffusion model (DNN)
- One input (noise) and one final output (E)
- Conditioned on θ, E_t

Training



Single particles
(Photons and hadrons)
E = [1,20] GeV

Full Geant4 Simulation ($\mathcal{N}(\theta, \sigma)$)
- Layer material*
- Layer thicknesses

Simple DNN reconstruction
- Has access to $\hat{\theta}$

Objective:

$$L = \frac{(E_p - E_t)^2}{E_t}$$

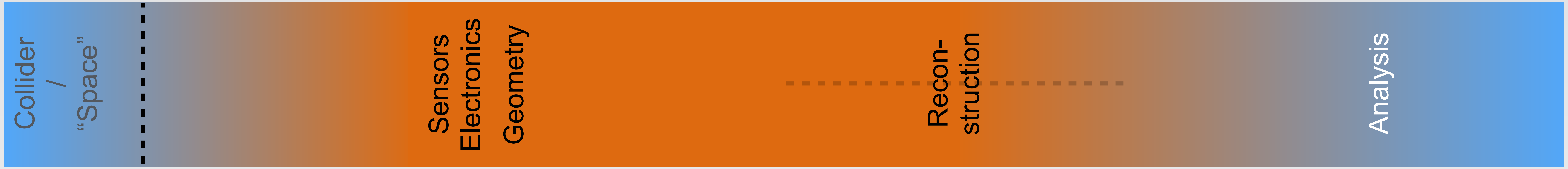
1.

2.

Simple diffusion model (DNN)
- One input (noise) and one final output (E)
- Conditioned on $\hat{\theta}$, E_t

Sample specific $\hat{\theta}$
σ: 'reasonable' values
Local surrogate

Optimisation



Single particles (Photons and hadrons) E = [1,20] GeV	Full Geant4 Simulation ($\mathcal{N}(\theta, \sigma)$) • Layer material* • Layer thicknesses	Simple DNN reconstruction • Has access to $\hat{\theta}$	Objective: resolution $L = \frac{(E_p - E_t)^2}{E_t}$

+ constraints (length, cost)

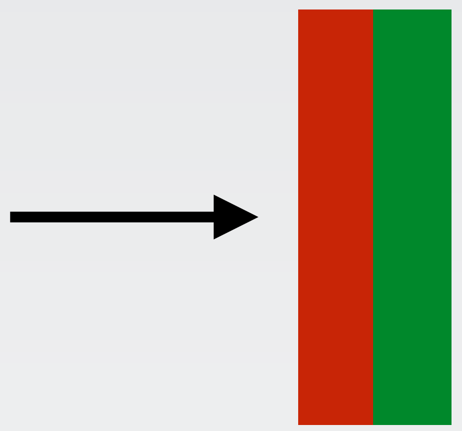
Simple diffusion model (DNN)
- One input (noise) and one final output (E)
- Conditioned on $\hat{\theta}$, E_t

3. Optimise $\hat{\theta}$ with SGD (Adam)
Stop if $\theta^s - \theta > \sigma$

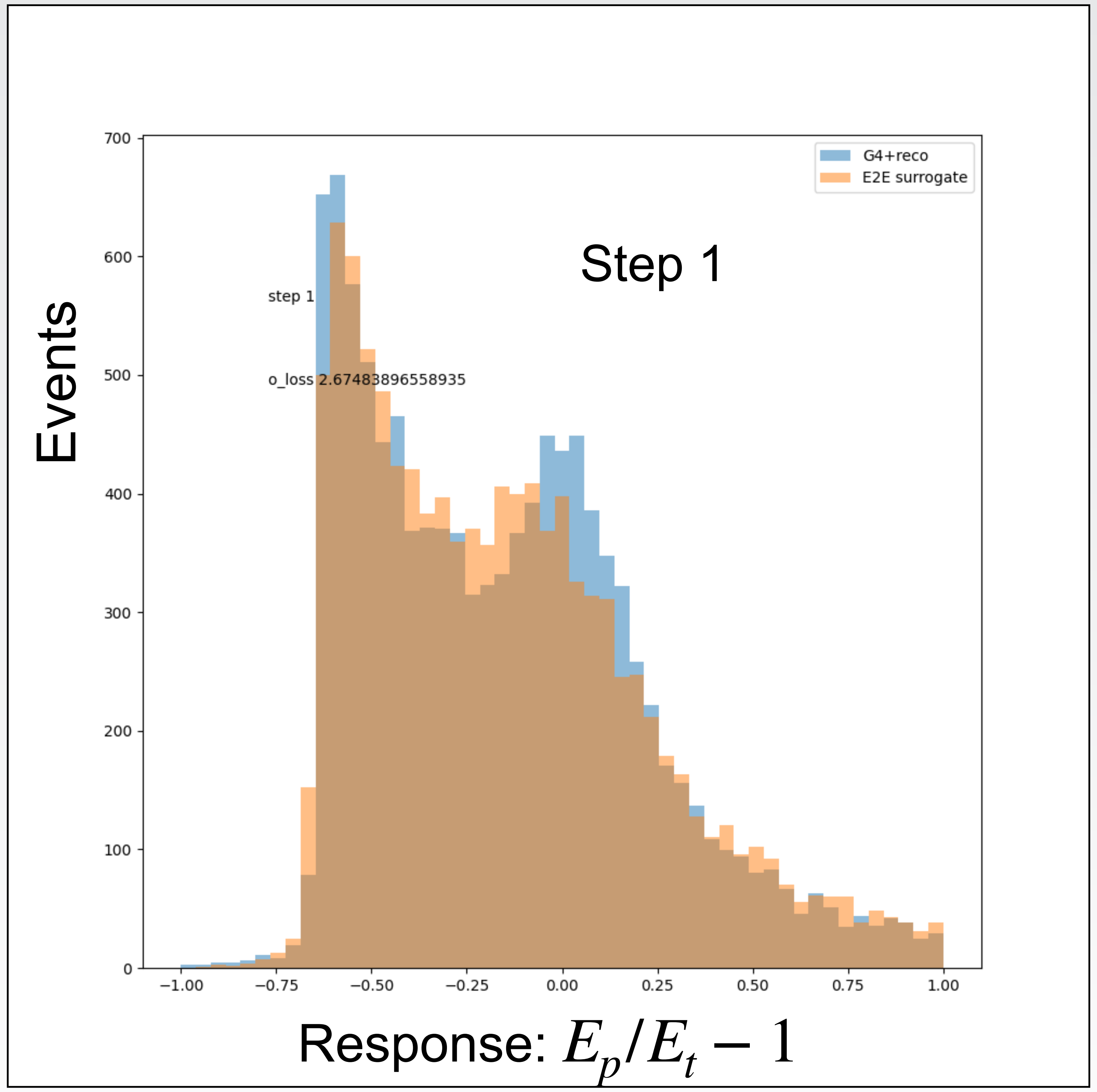
- Input normalisations are kept fixed after first training iteration
- Reconstruction and diffusion models are ***refined*** on the new data, not trained from scratch
  - Transfer-learning significantly reduces the amount of simulation resources needed *

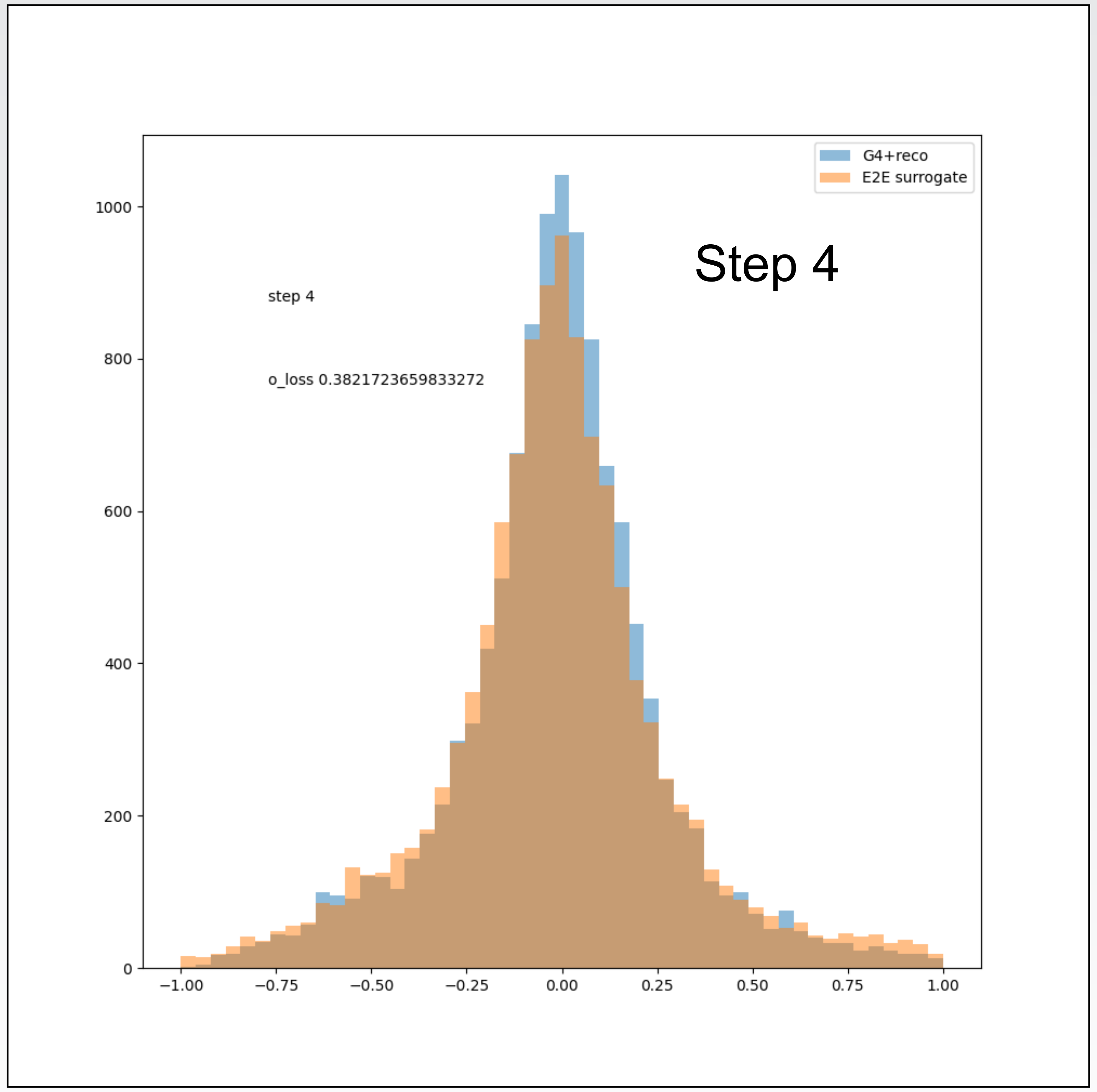
Sanity check

- Photons: 1-20 GeV, uniform
- Absorber Pb: 1cm
- Active PbWO4 : 1cm
- Constraint: depth of 25 cm

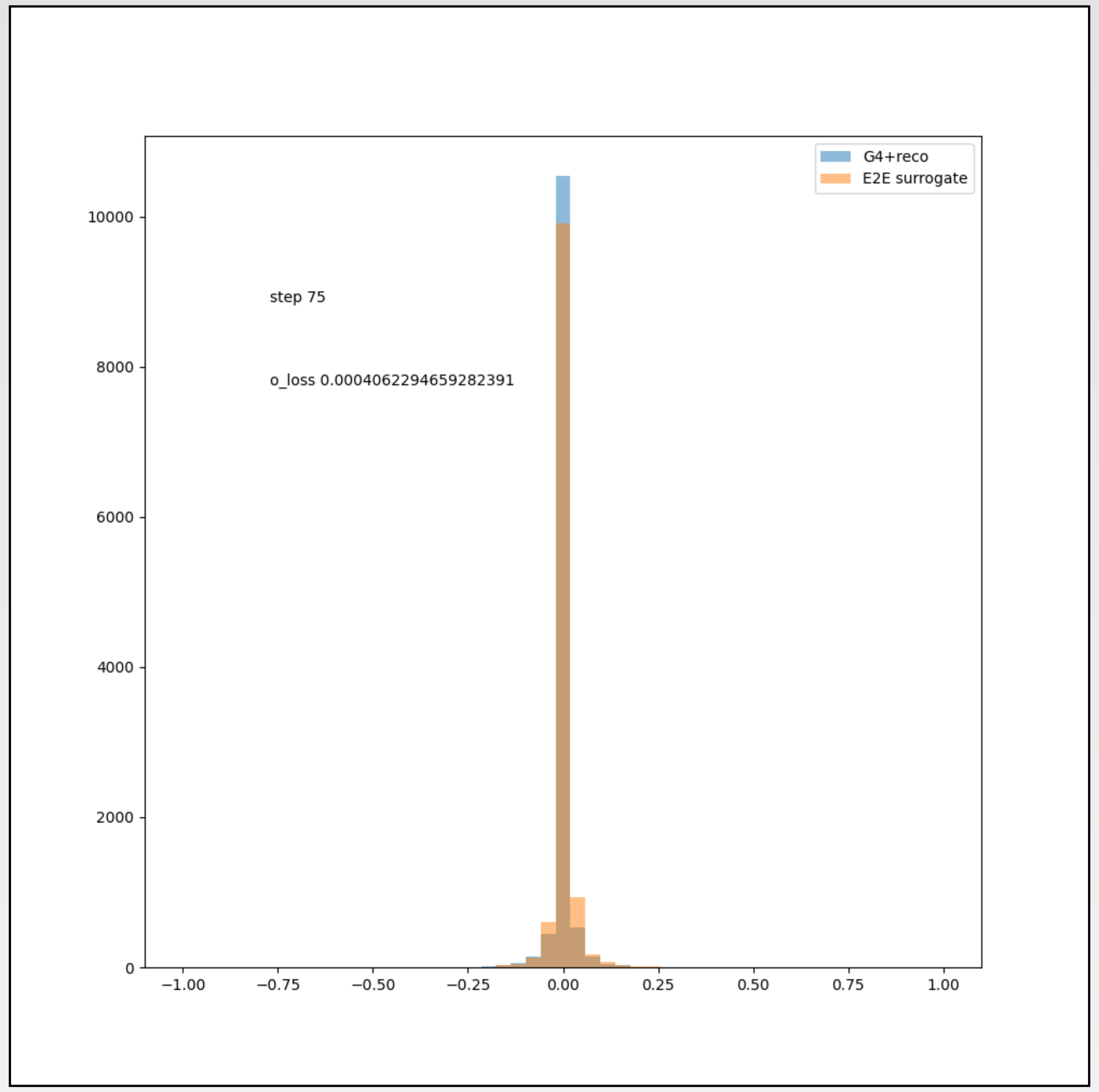


Expected best solution



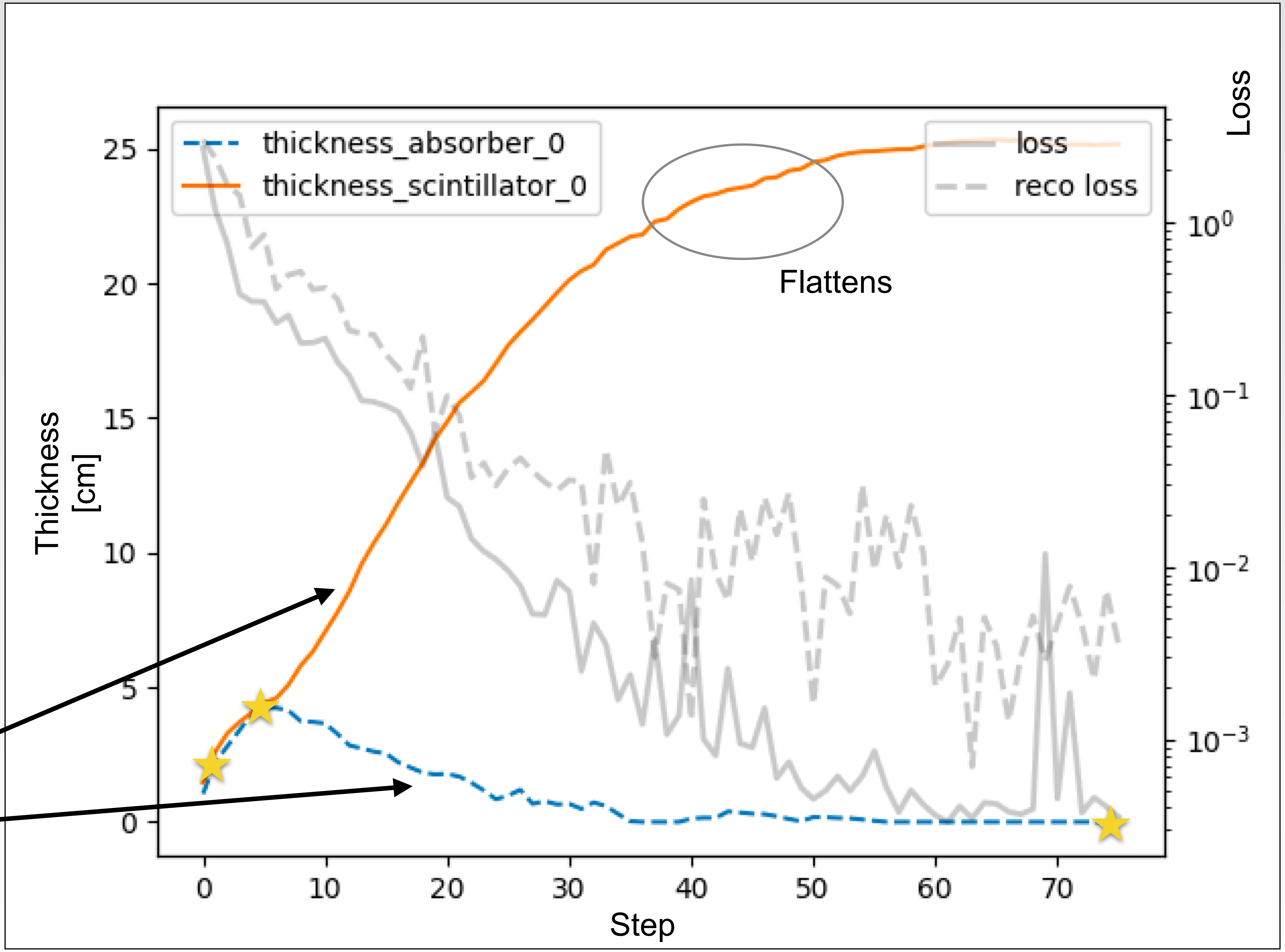


It works



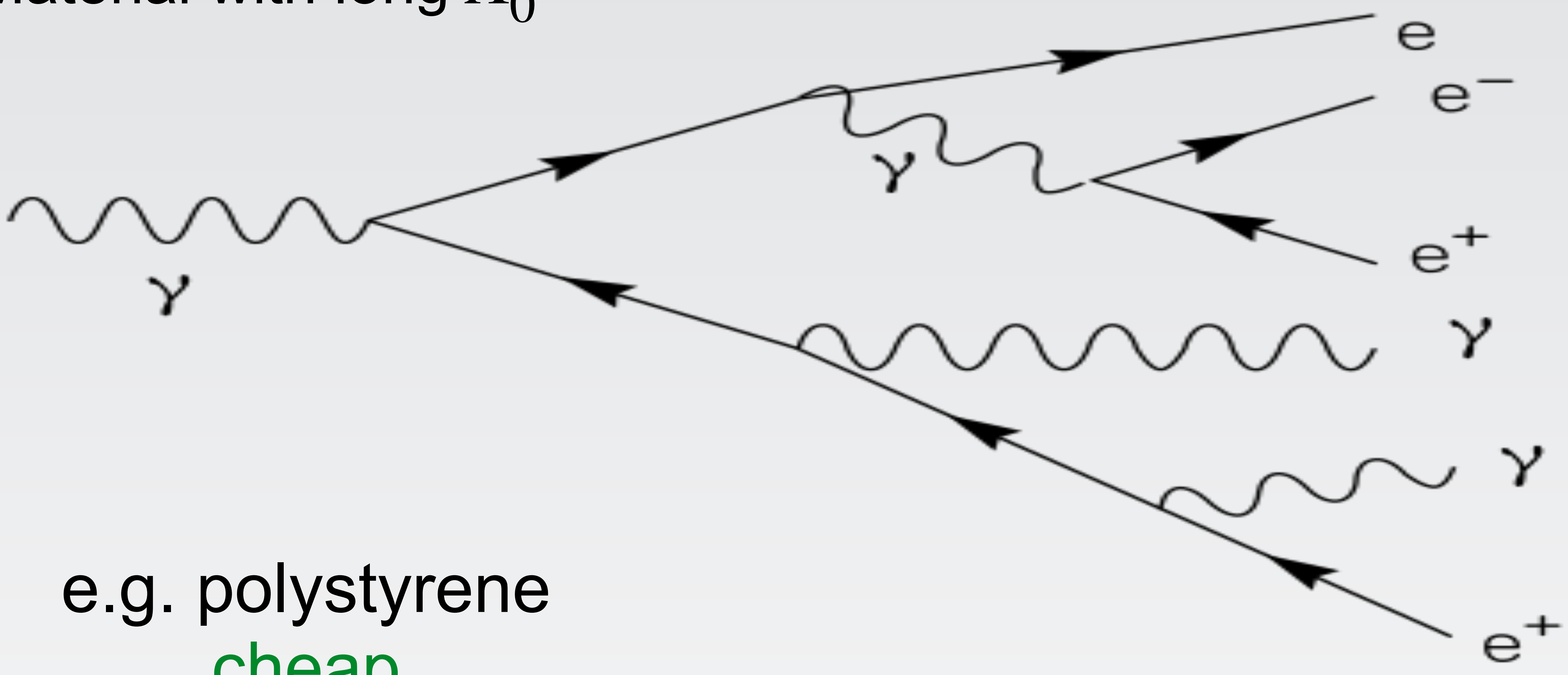
Final configuration

0 cm 25 cm

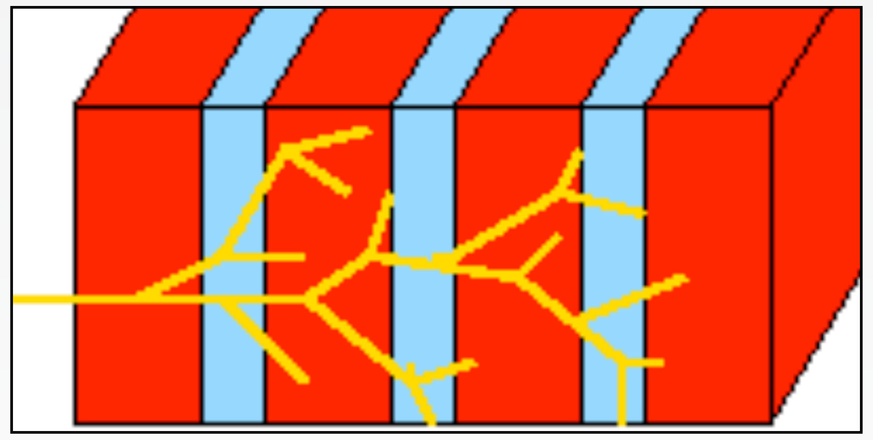


A look at discrete parameters: Material

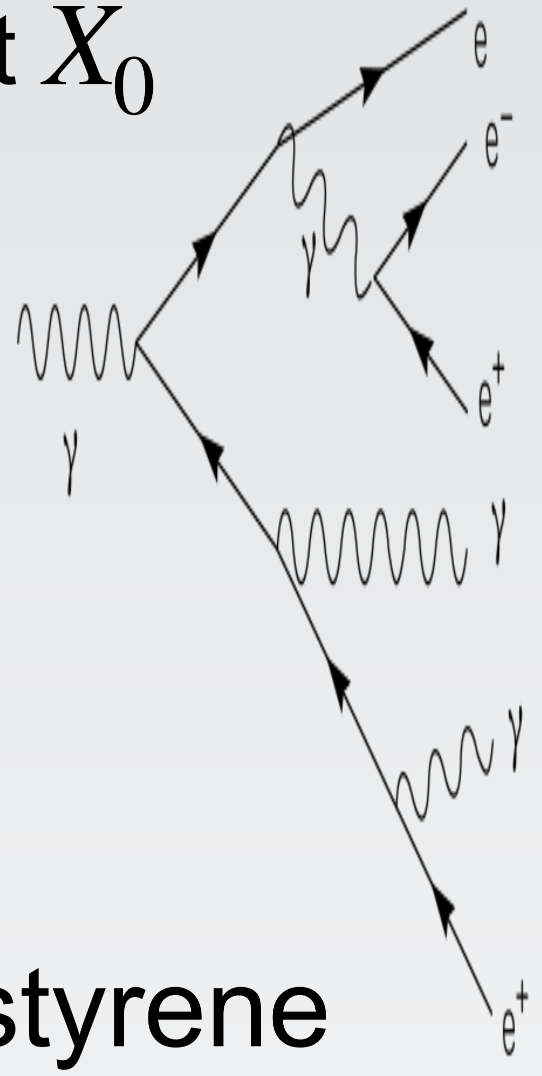
Material with long X_0



e.g. polystyrene
cheap
low resolution

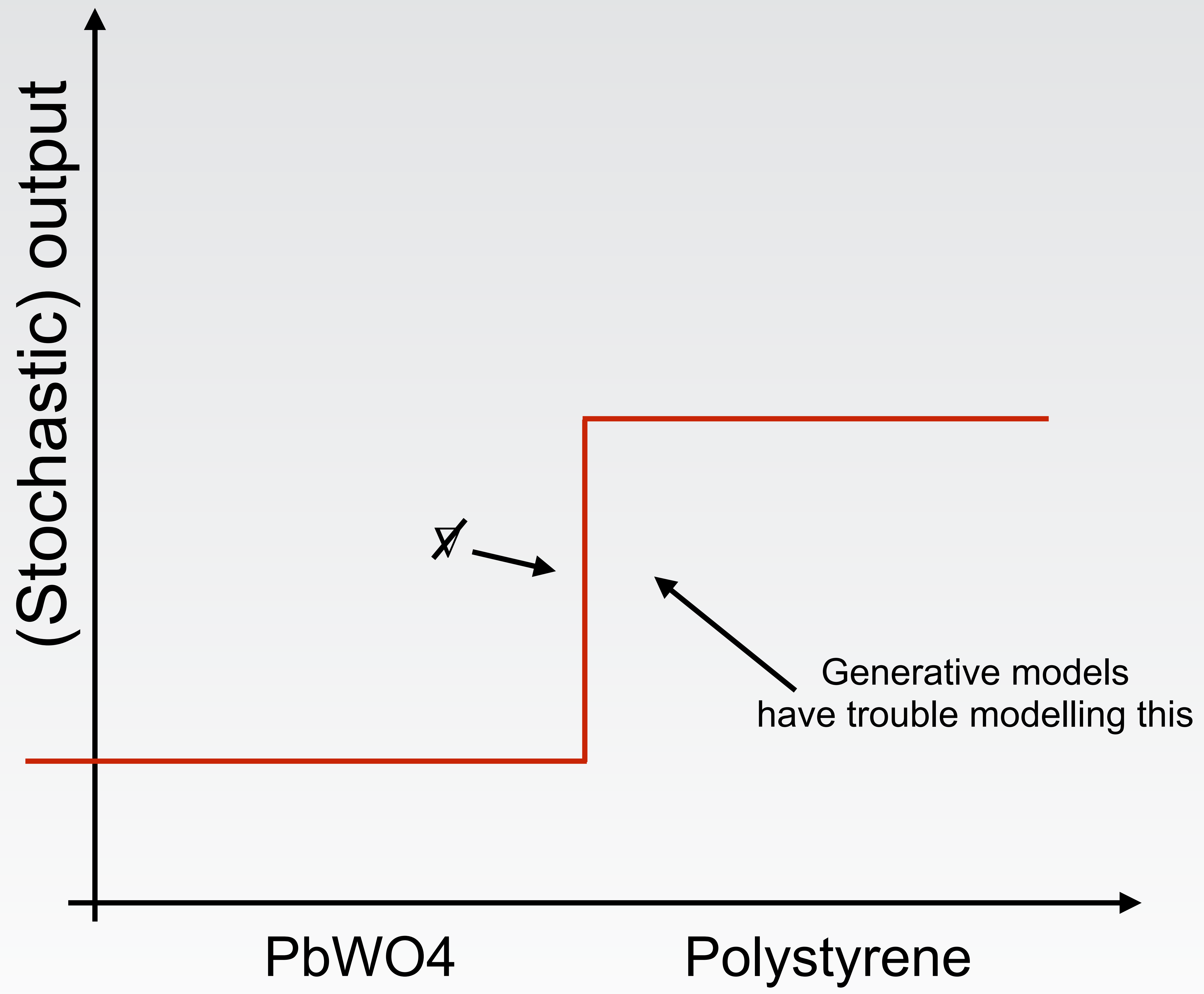


Material with short X_0

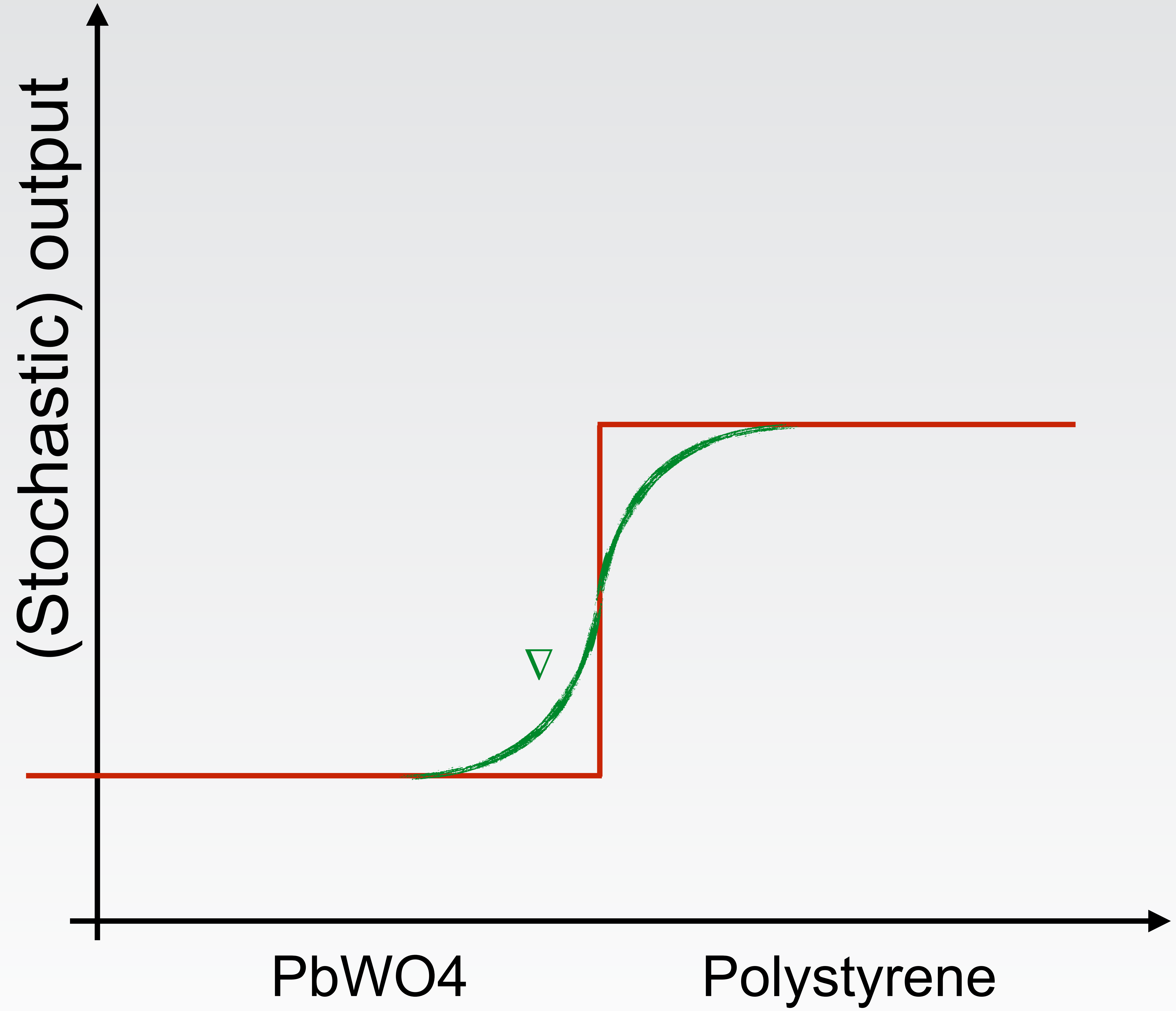


e.g. polystyrene
expensive
high resolution

The problem with discrete parameters

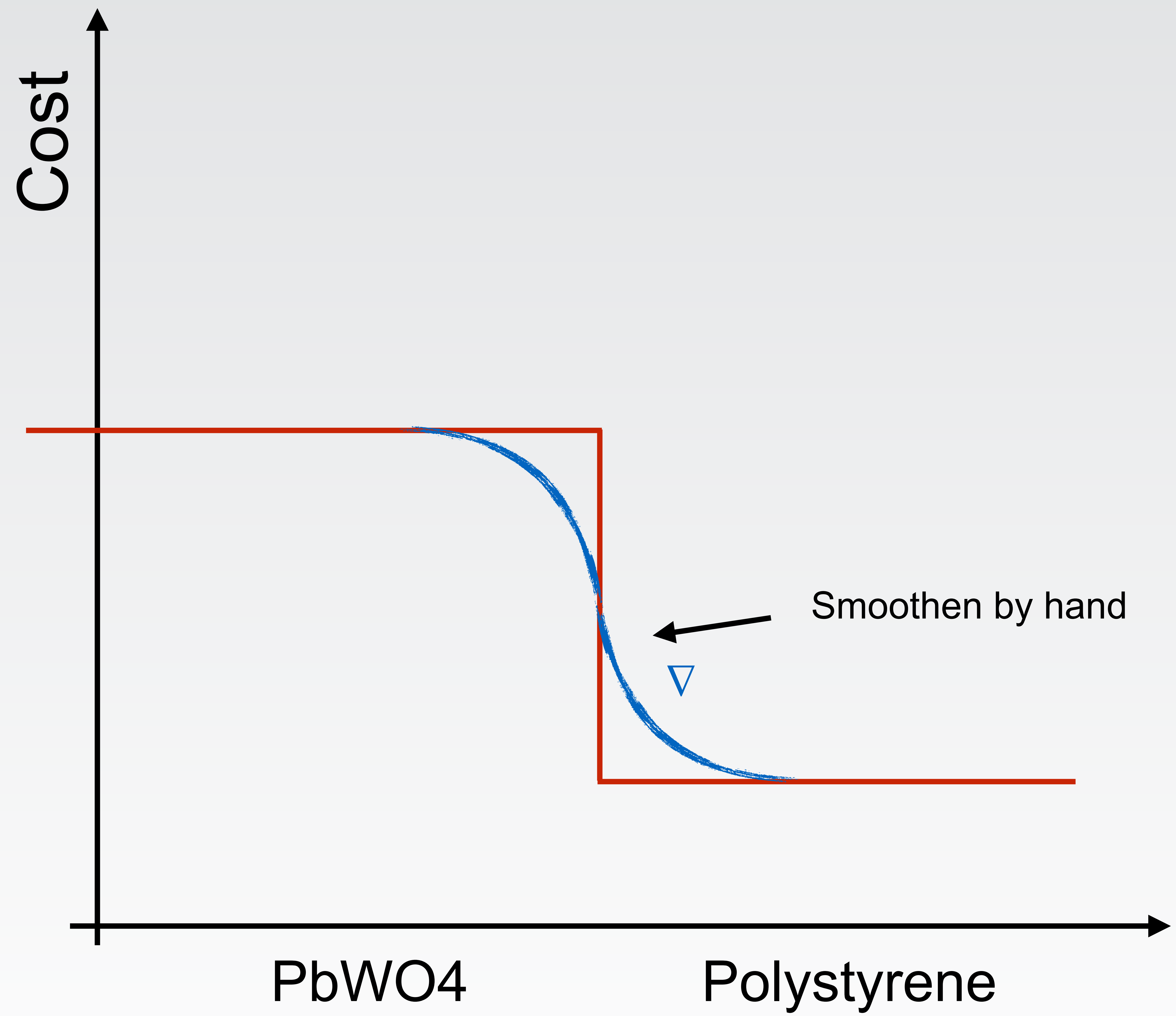


The output as a function of the discrete parameter



- We can use the strong tendency of DNNs to “interpolate everything” to effectively model (gradients on) discrete parameters

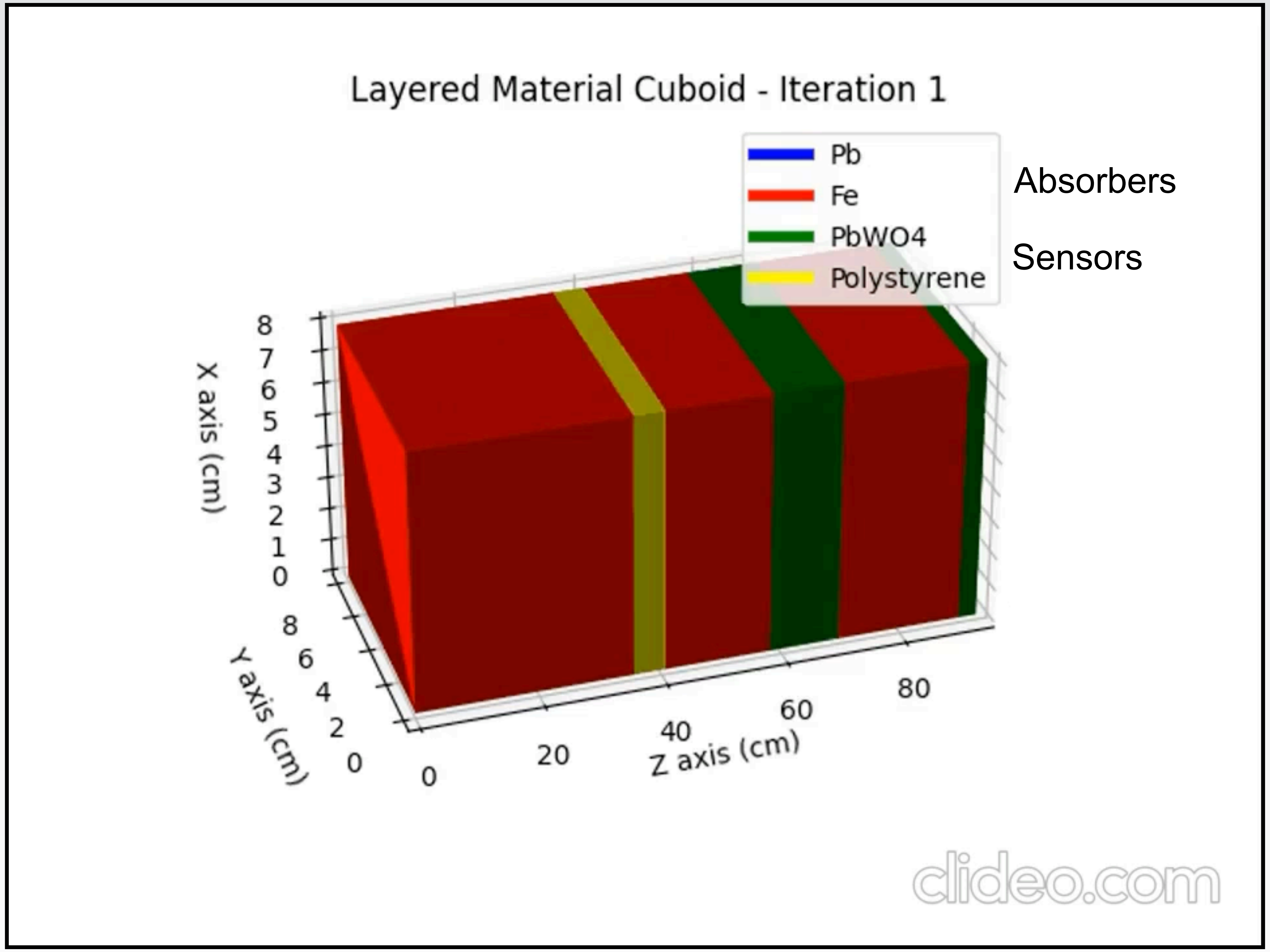
Cost in objective function



- Implement also the material cost into the objective function using a smoother dependence

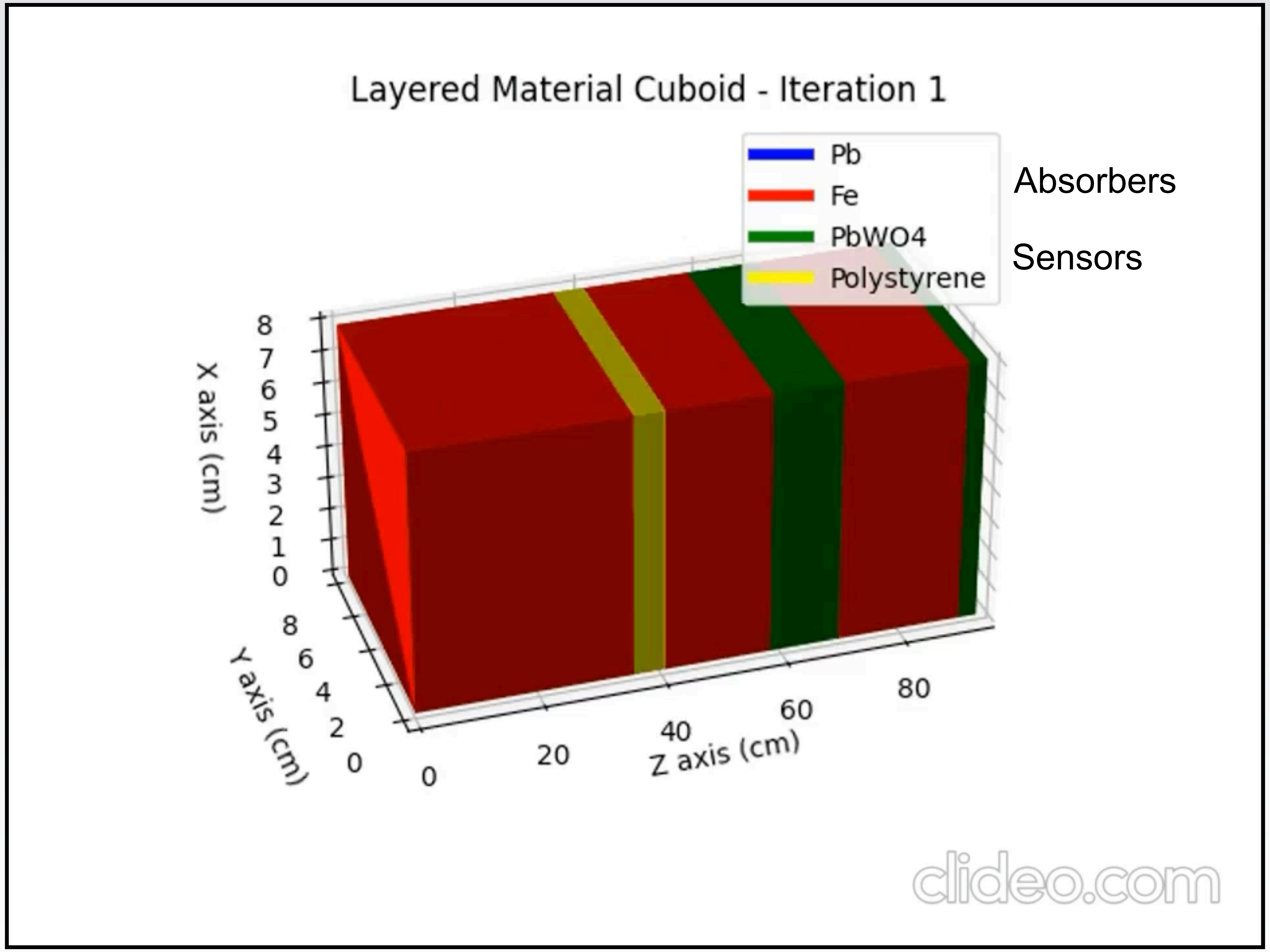
Does it work?

- 3 x (absorber + sensor)
- Different particles (electromagnetic **and hadronic**)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
  - e.g. no photon will actually reach the first sensor



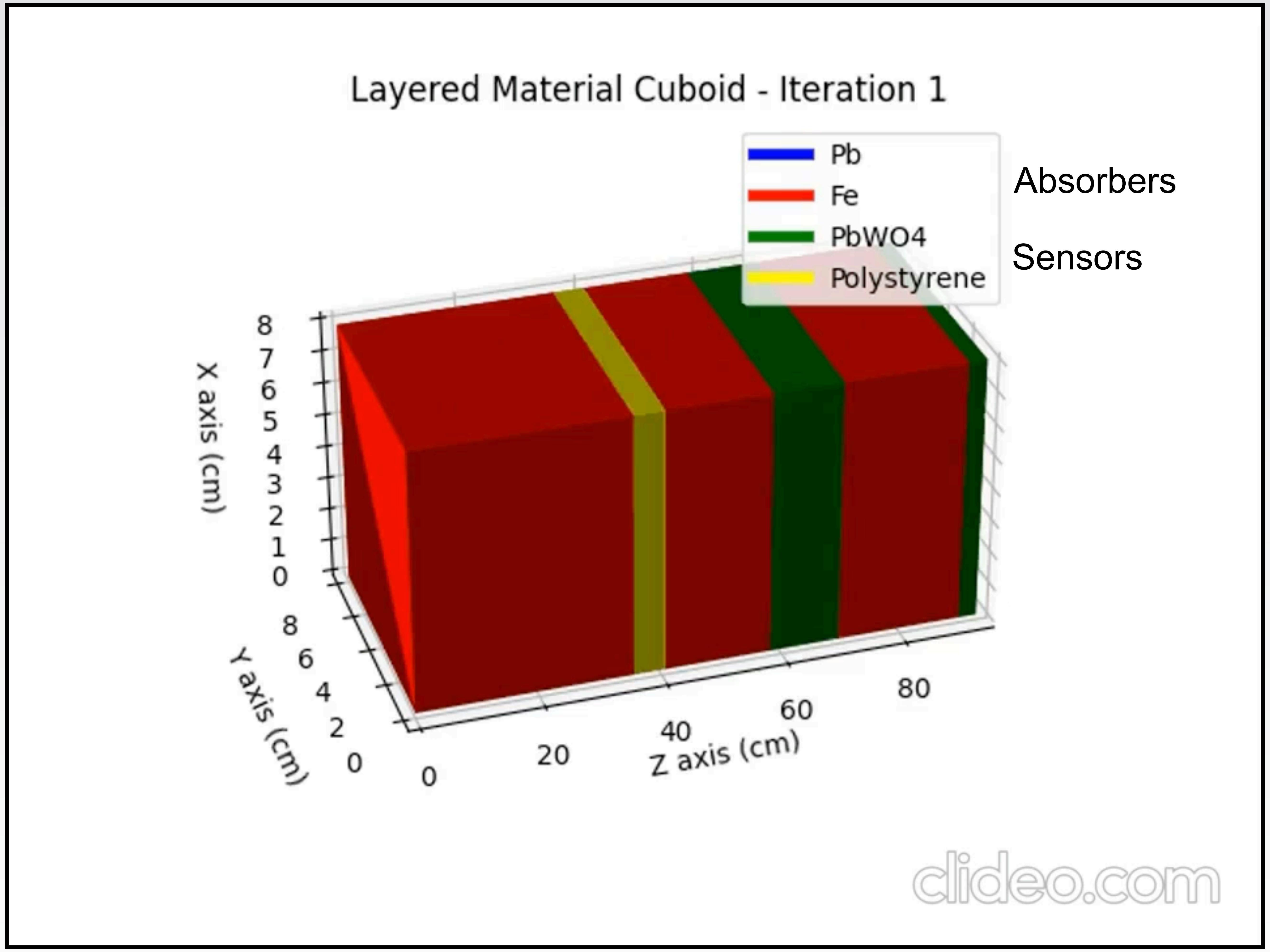
Does it work?

- 3 x (absorber + sensor)
- Different particles (electromagnetic **and hadronic**)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
  - e.g. no photon will actually reach the first sensor



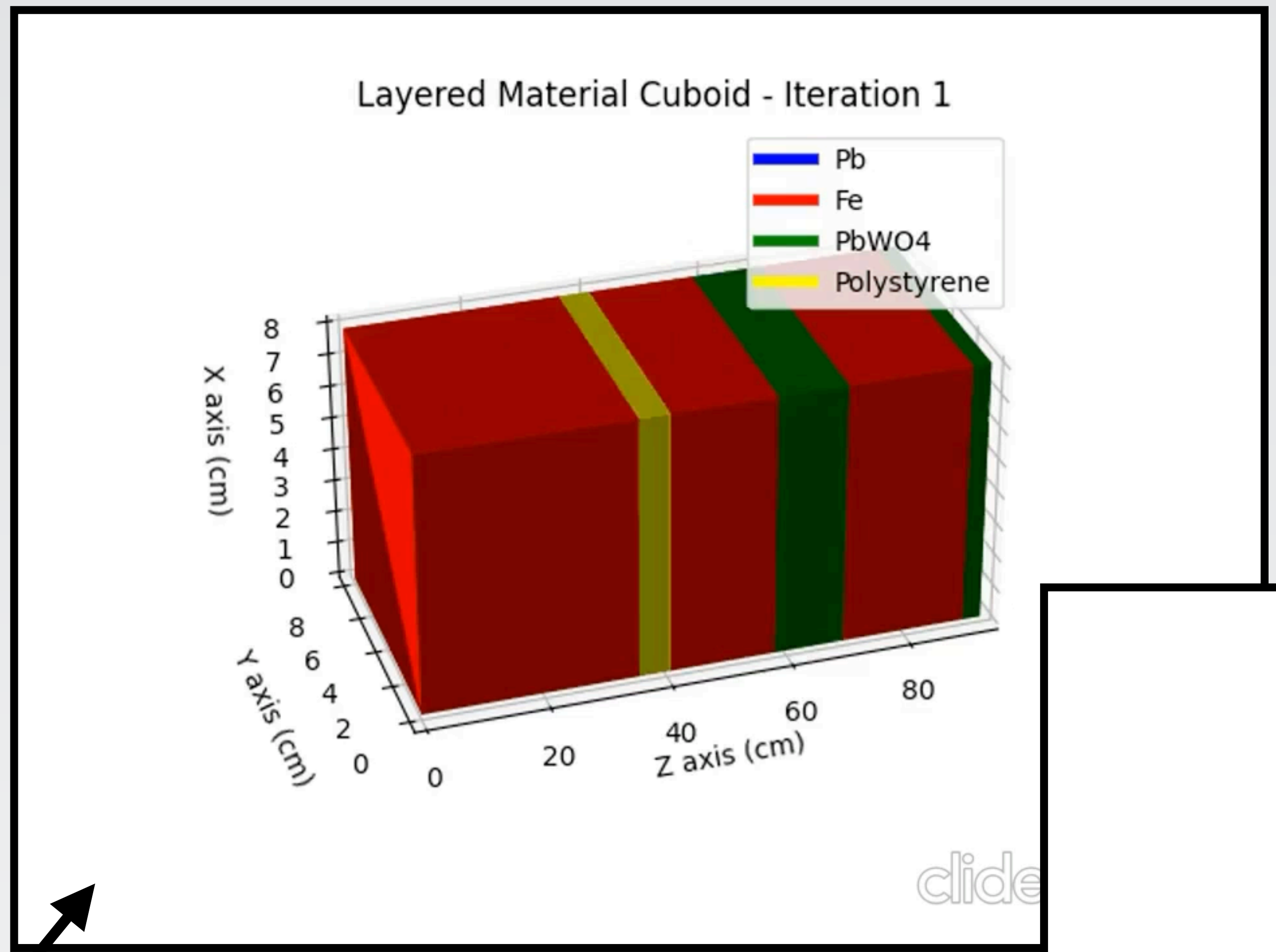
Does it work?

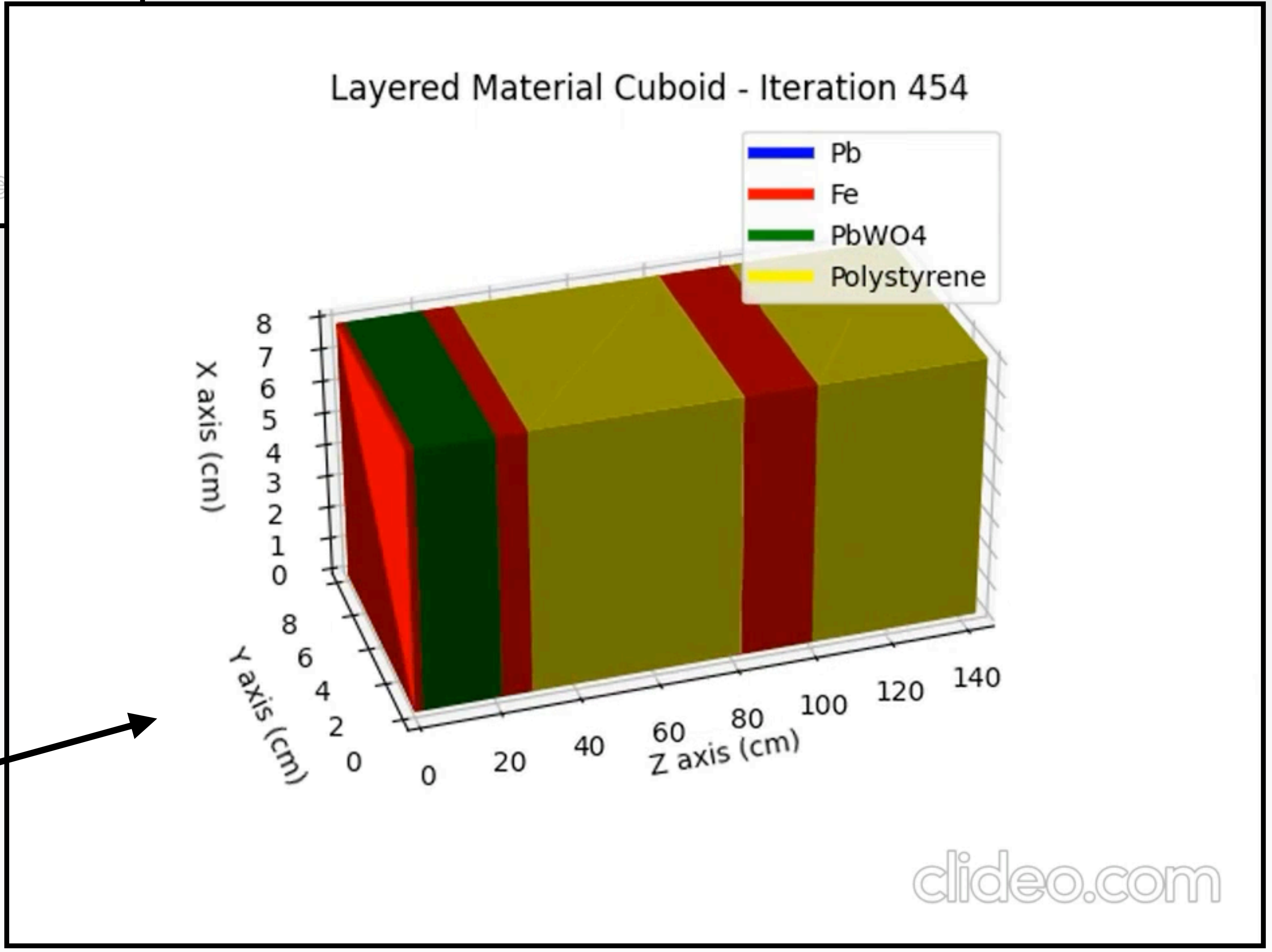
- 3 x (absorber + sensor)
- Different particles (electromagnetic **and hadronic**)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
  - e.g. no photon will actually reach the first sensor



Yes.

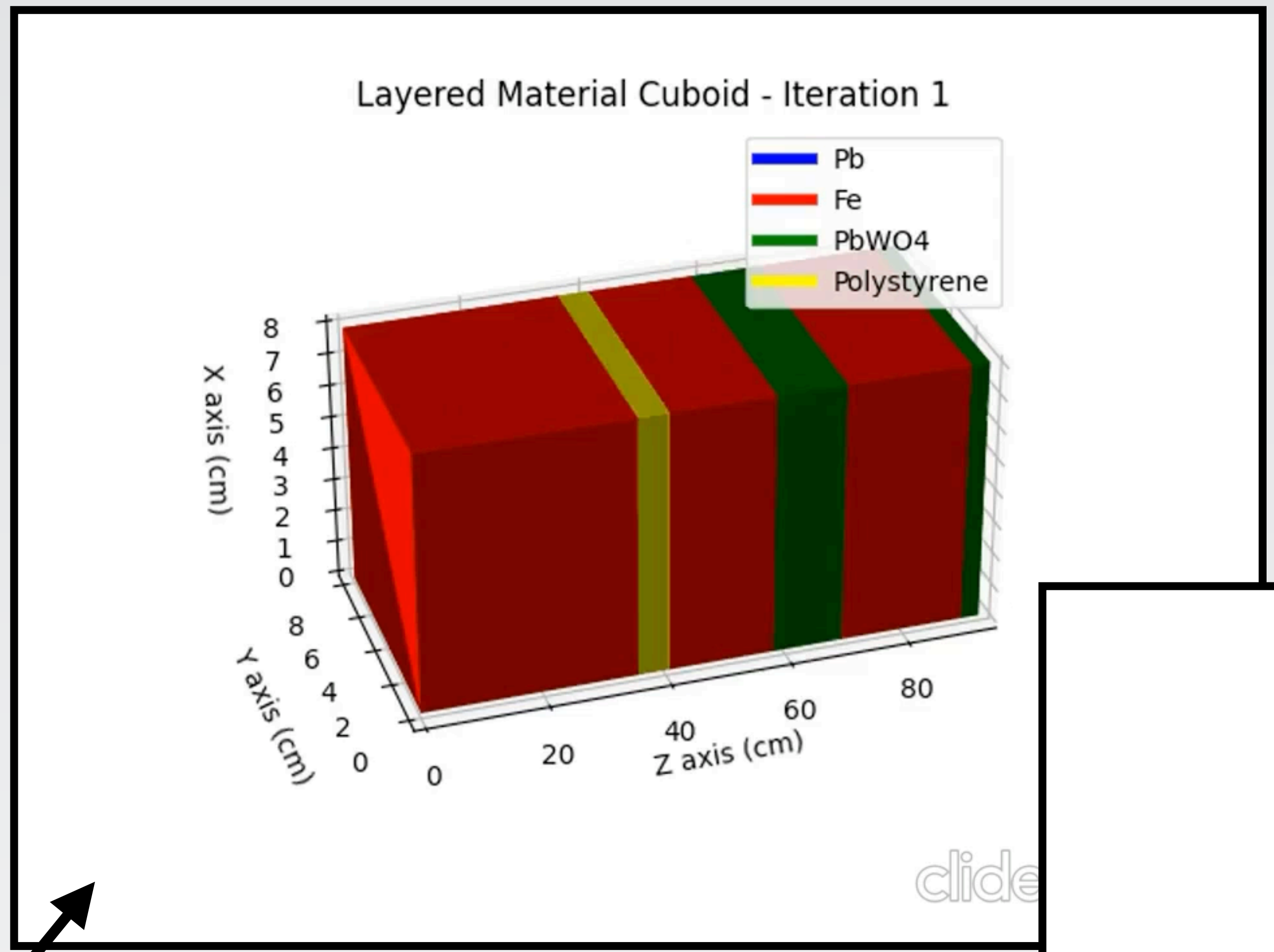
- Different particles (electromagnetic and hadronic)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
- After a few hundred iterations: a ***very reasonable*** configuration with significantly better performance (and close to real-world detector designs)

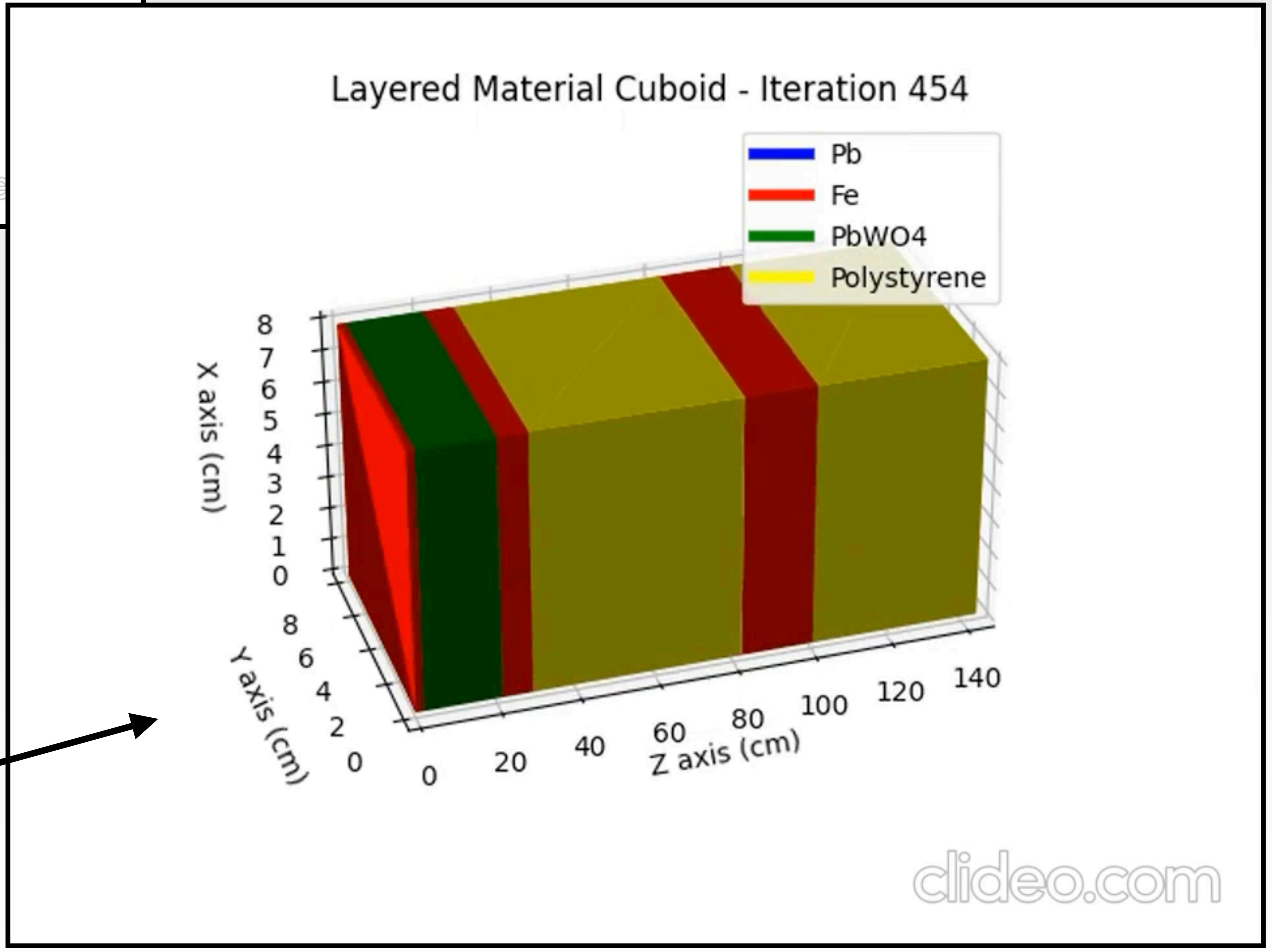




Yes.

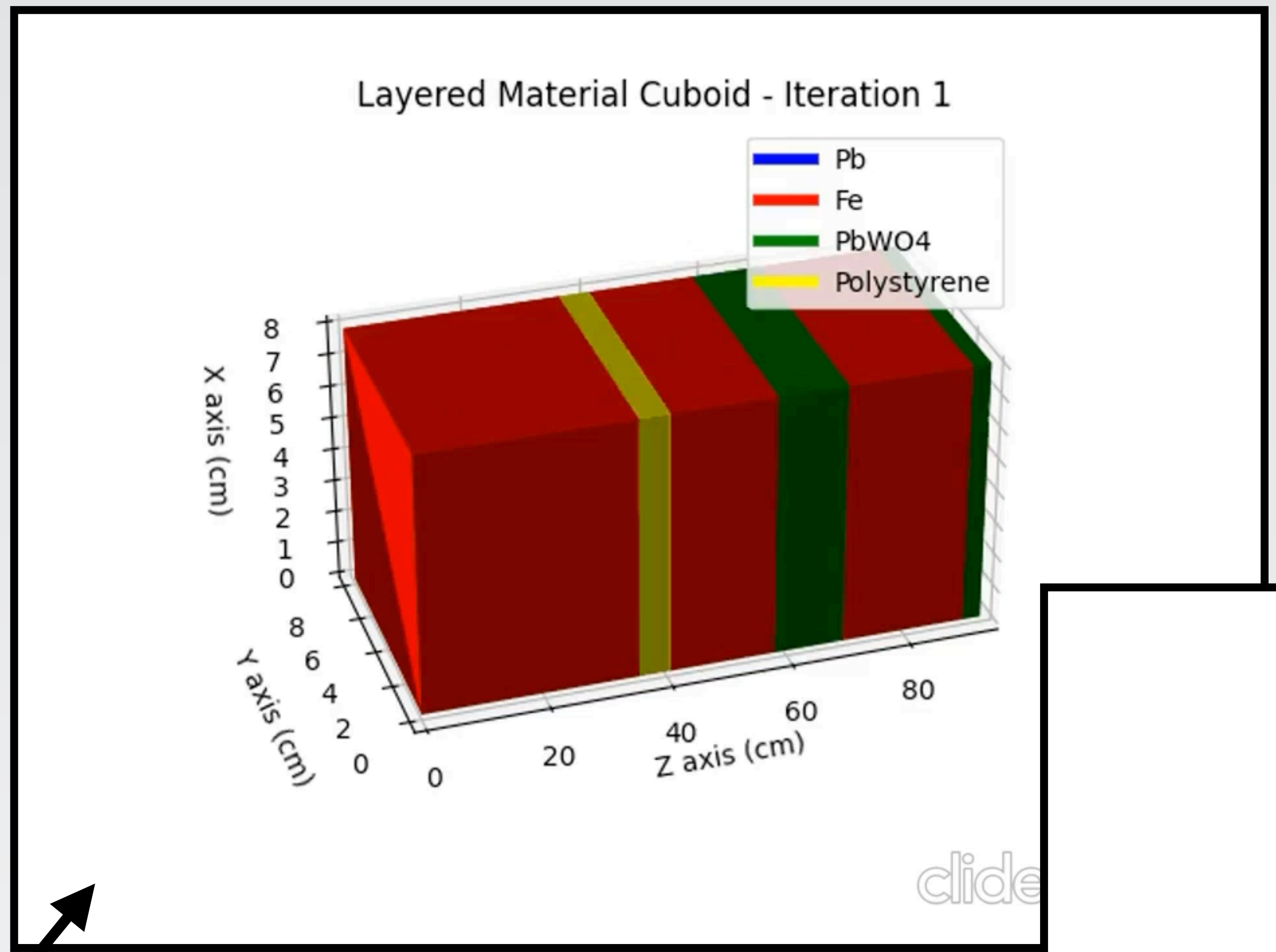
- Different particles (electromagnetic and hadronic)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
- After a few hundred iterations: a ***very reasonable*** configuration with significantly better performance (and close to real-world detector designs)

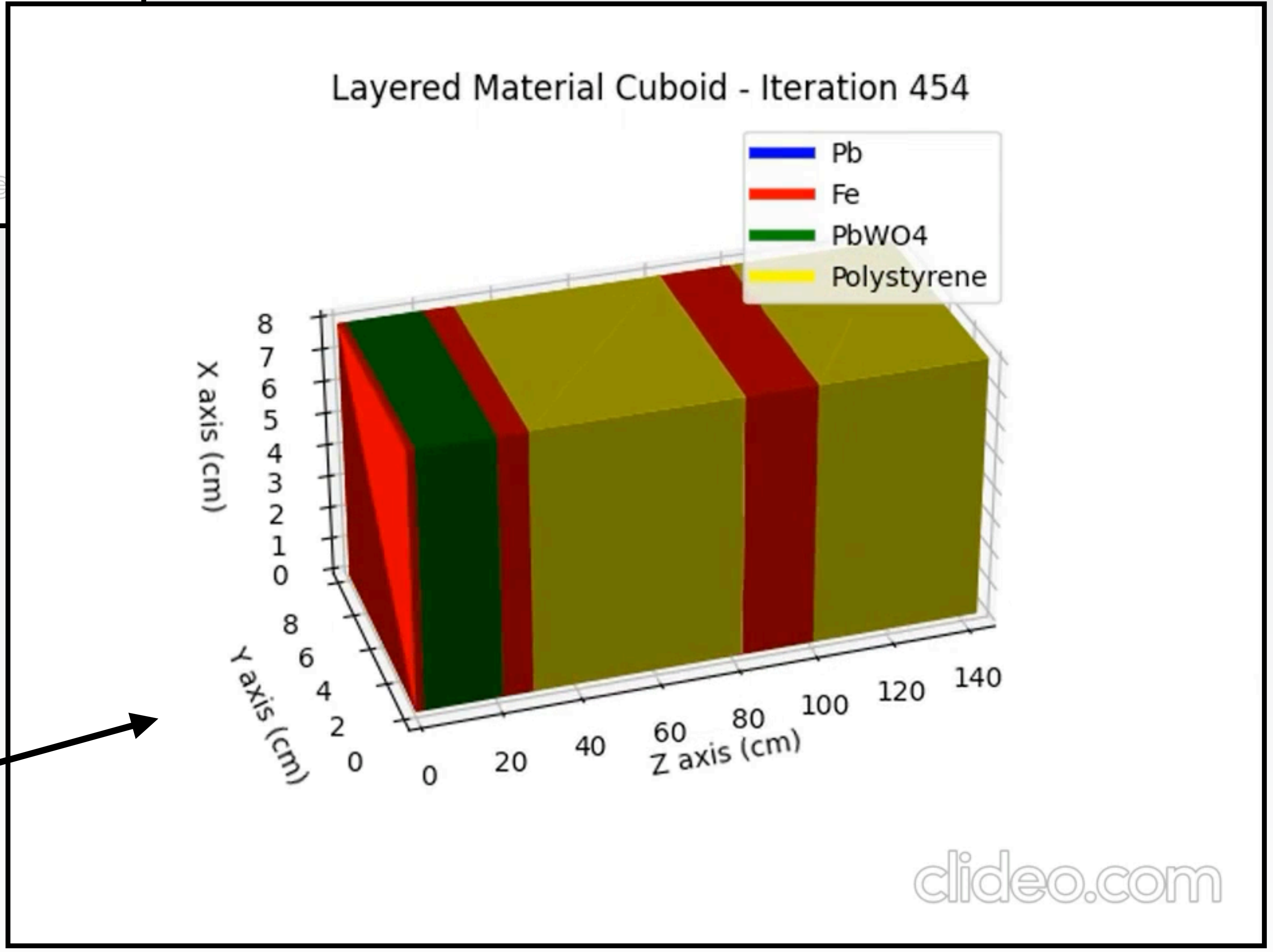




Yes.

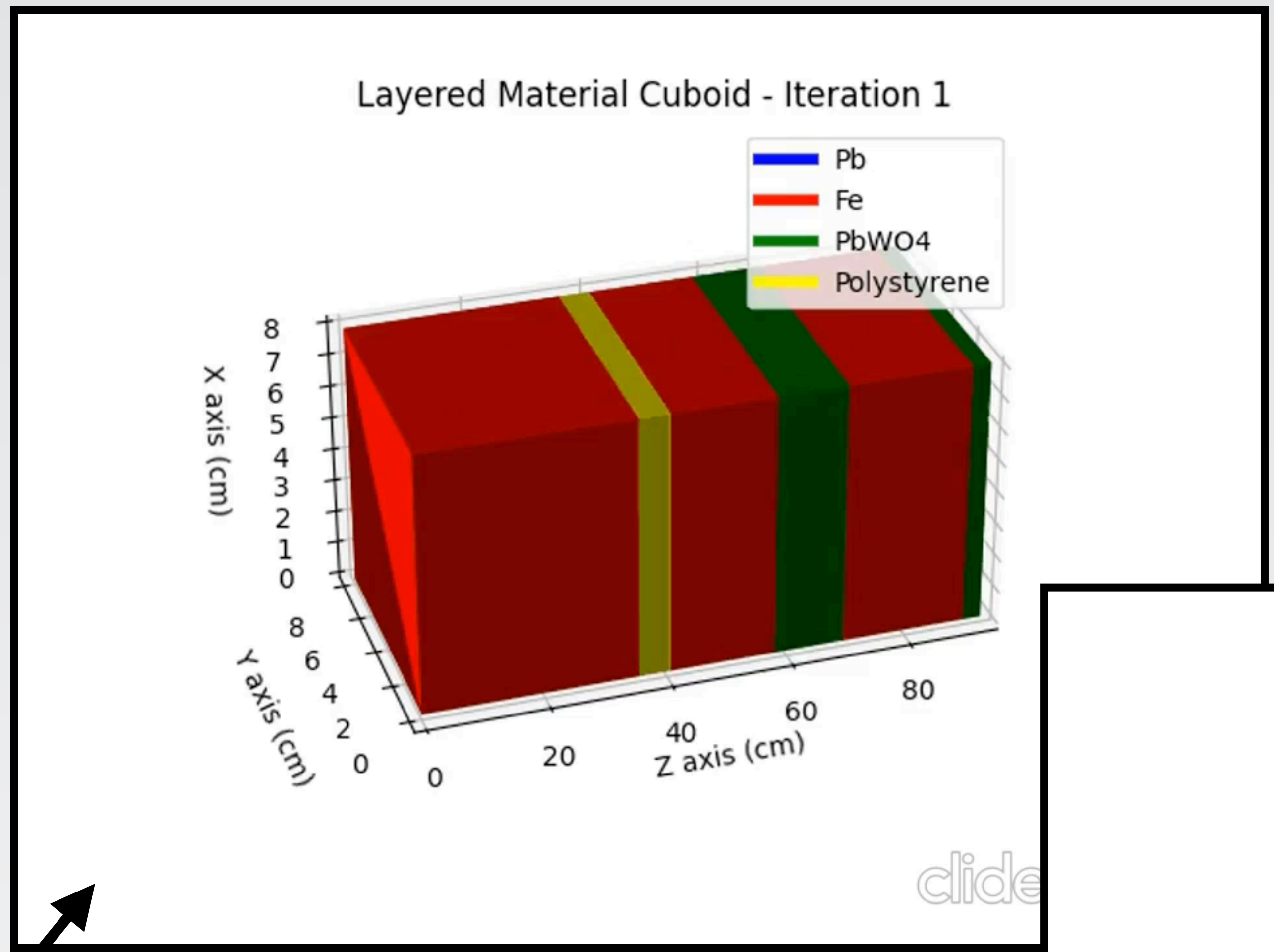
- Different particles (electromagnetic and hadronic)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
- After a few hundred iterations: a ***very reasonable*** configuration with significantly better performance (and close to real-world detector designs)

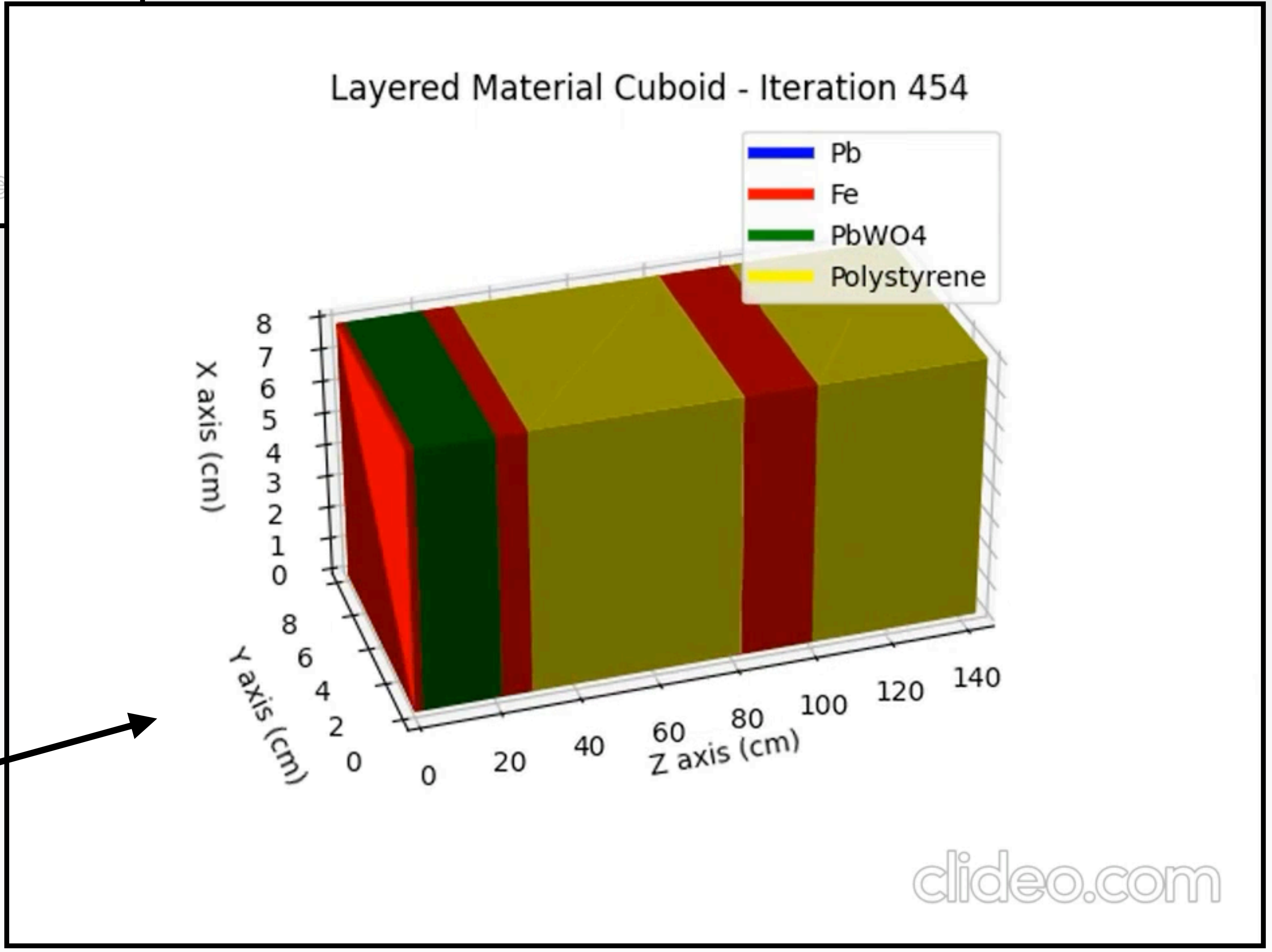




Yes.

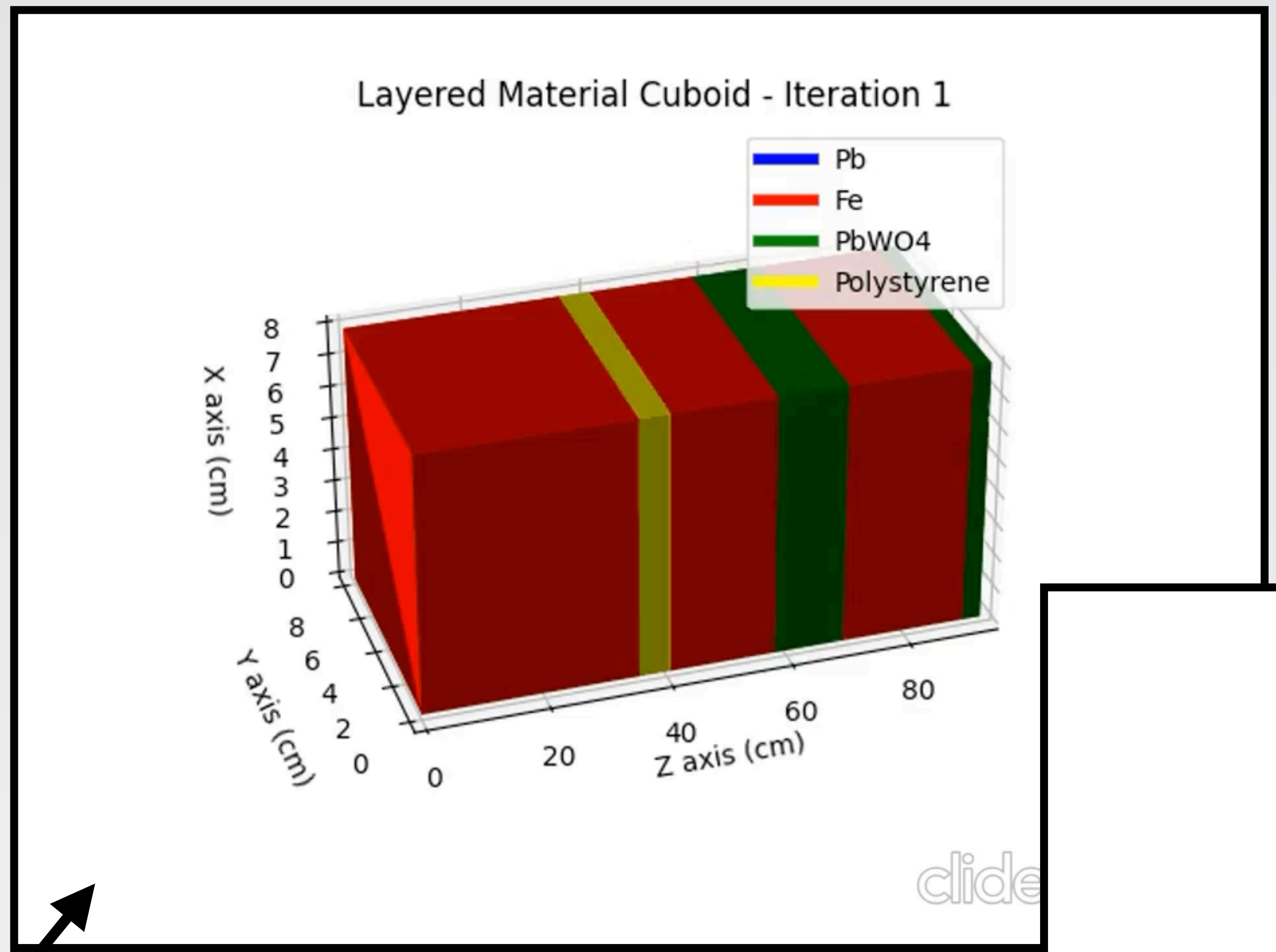
- Different particles (electromagnetic and hadronic)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
- After a few hundred iterations: a ***very reasonable*** configuration with significantly better performance (and close to real-world detector designs)

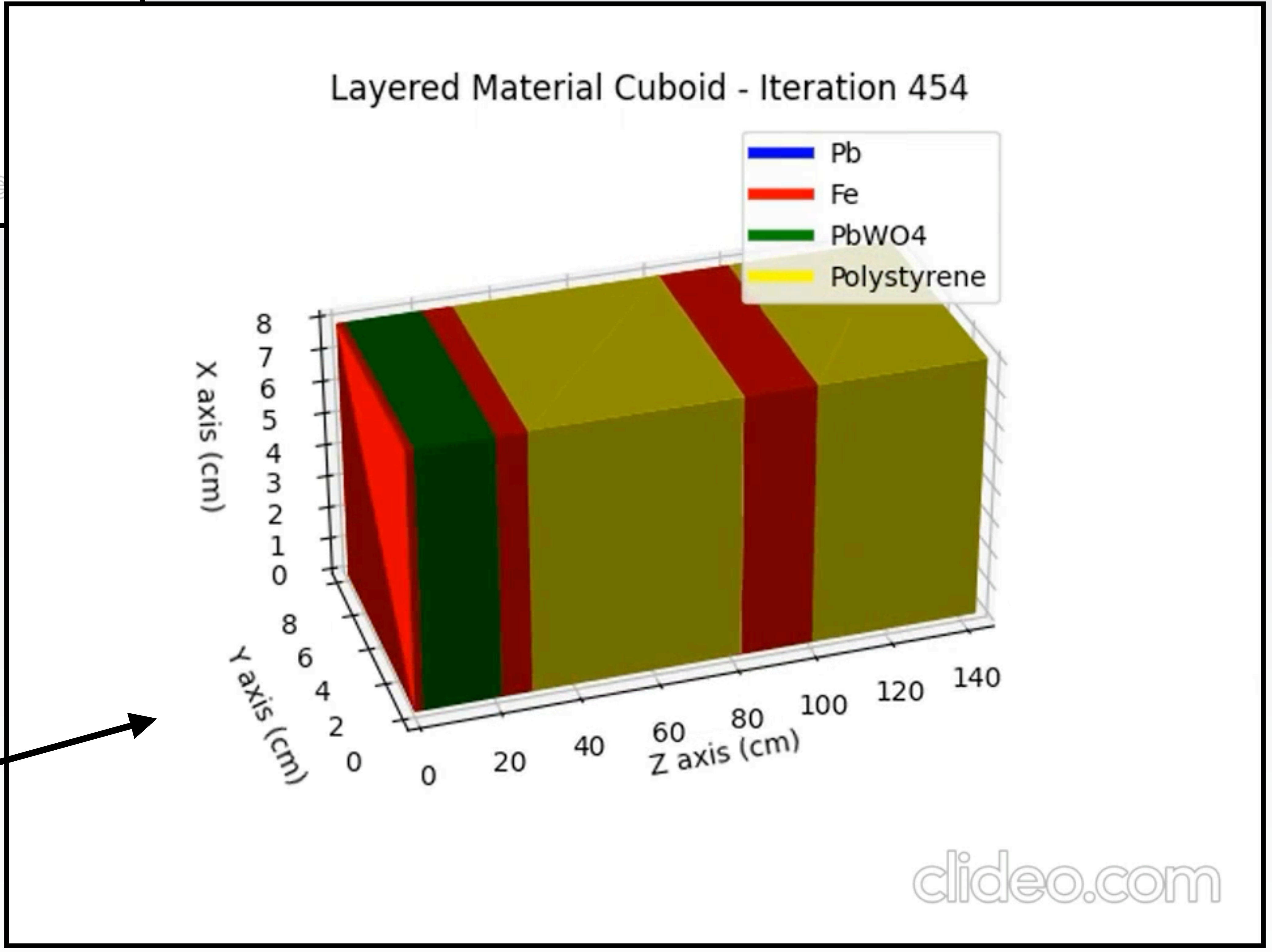




Yes.

- Different particles (electromagnetic and hadronic)
  - Short showers
  - Deep showers
- All with energies between 1 and 20 GeV
- Material cost <50k CHF
- Length <180 cm
- Start with a ***horrible*** configuration
- After a few hundred iterations: a ***very reasonable*** configuration with significantly better performance (and close to real-world detector designs)





Summary

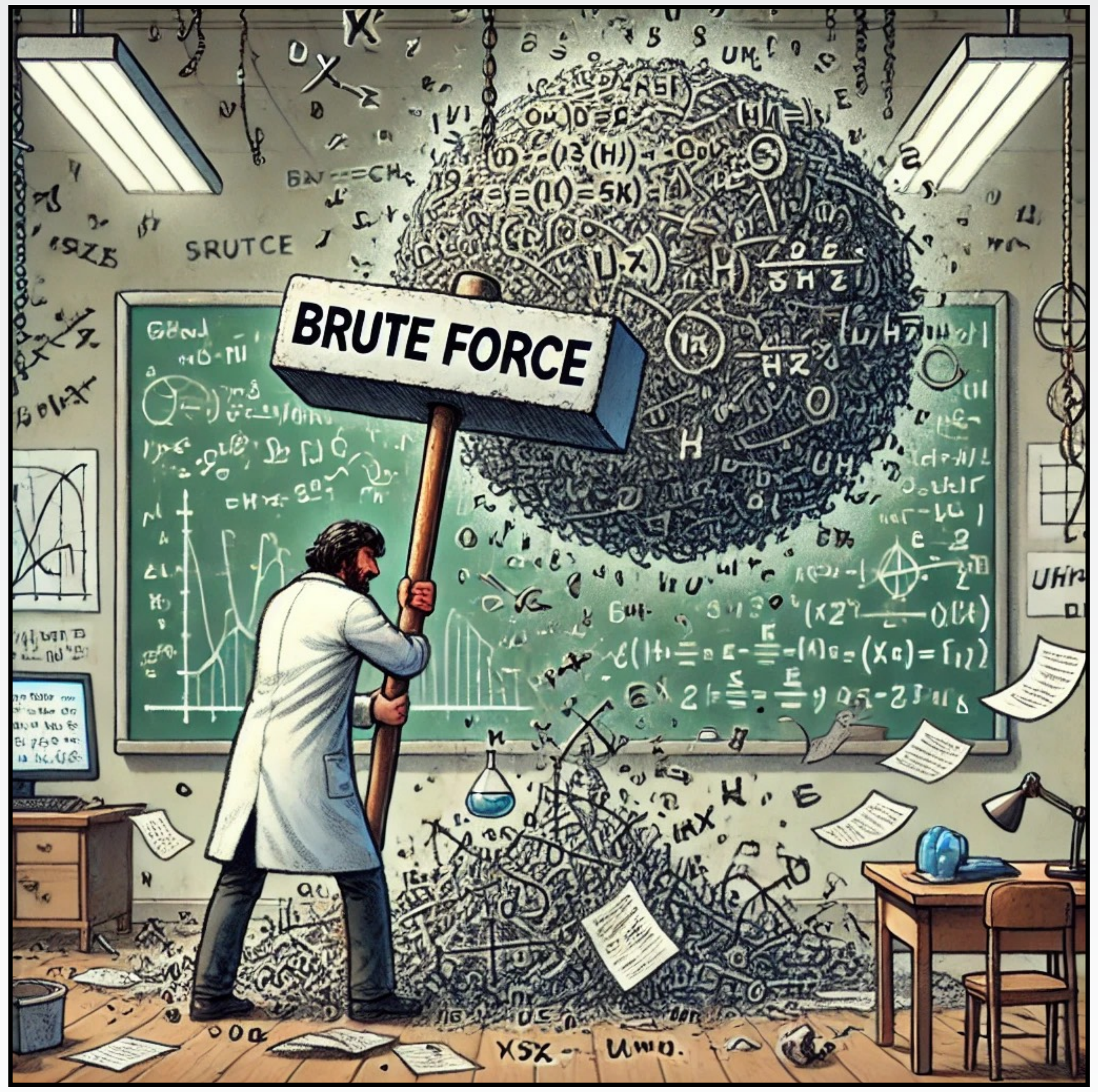
An end-to-end surrogate avoids creating a differentiable model for the complex, high dimensional intermediate detector-level state

It smoothes out non-differentiable operations in the simulation (and reconstruction)

In the (simplified) setting of calorimeters, it seems to work well

The dependence on discrete parameters can be modelled in an effective manner

More detailed studies ongoing - stay tuned



I have omitted many physics and technical details - please feel free to ask