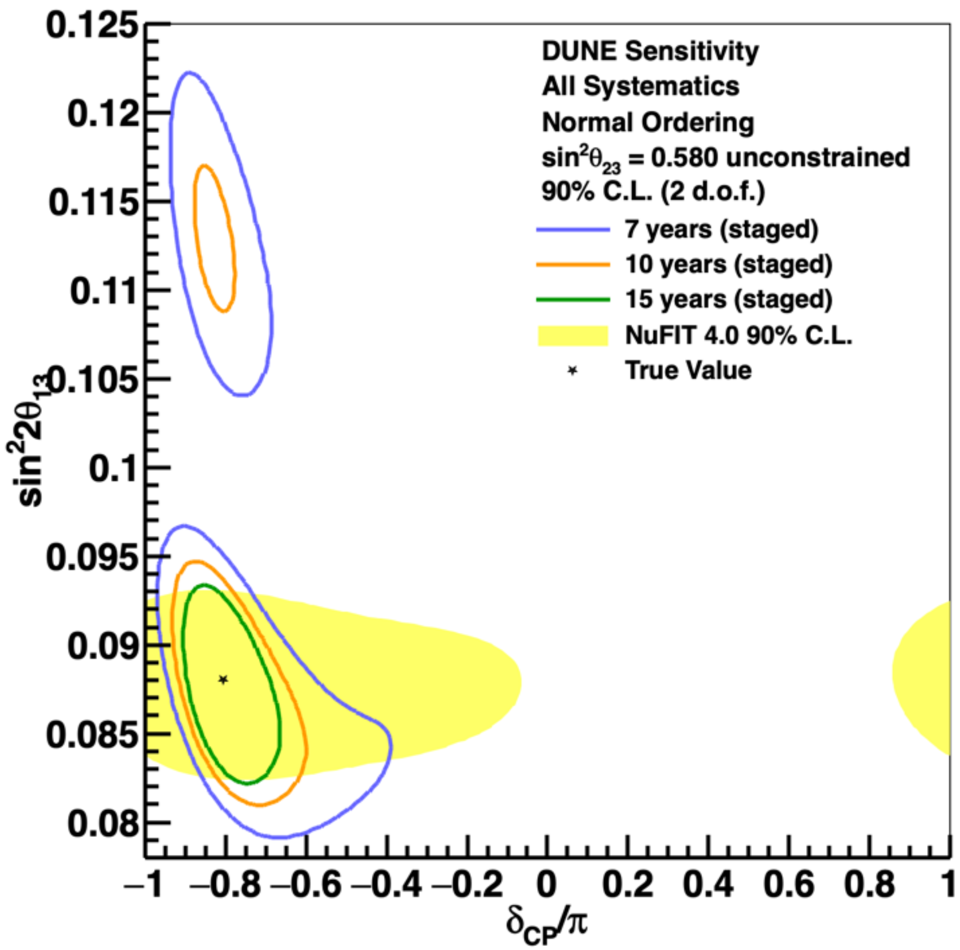
Simultaneous high-dimensional calibration with differentiable simulations for a liquid argon time projection chamber

Yifan Chen
SLAC, Stanford University

The fourth MODE workshop
Valencia, Spain
23 September 2024

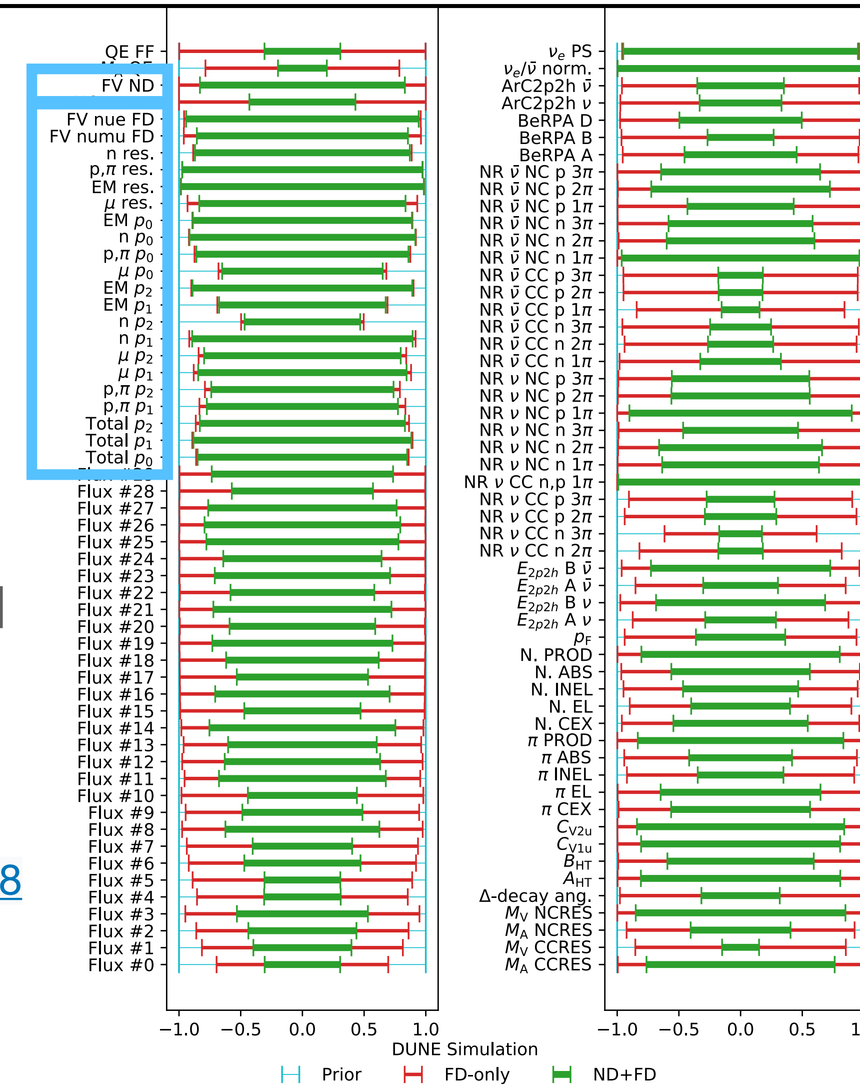
LArTPC's for Accelerator Neutrino Experiments

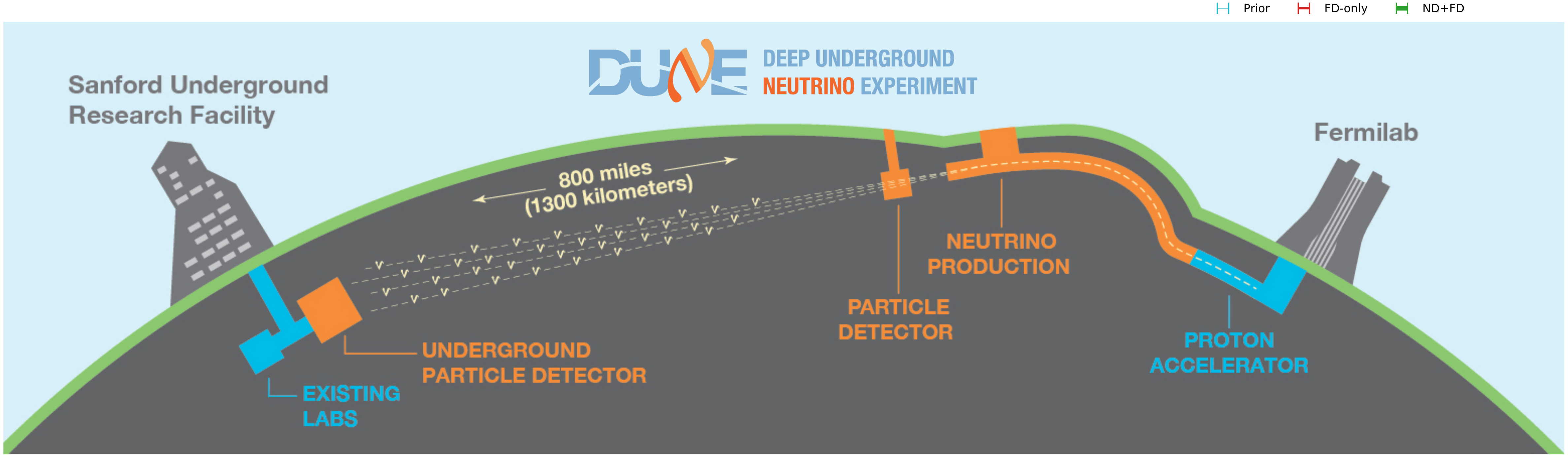


Systematics of detector modeling

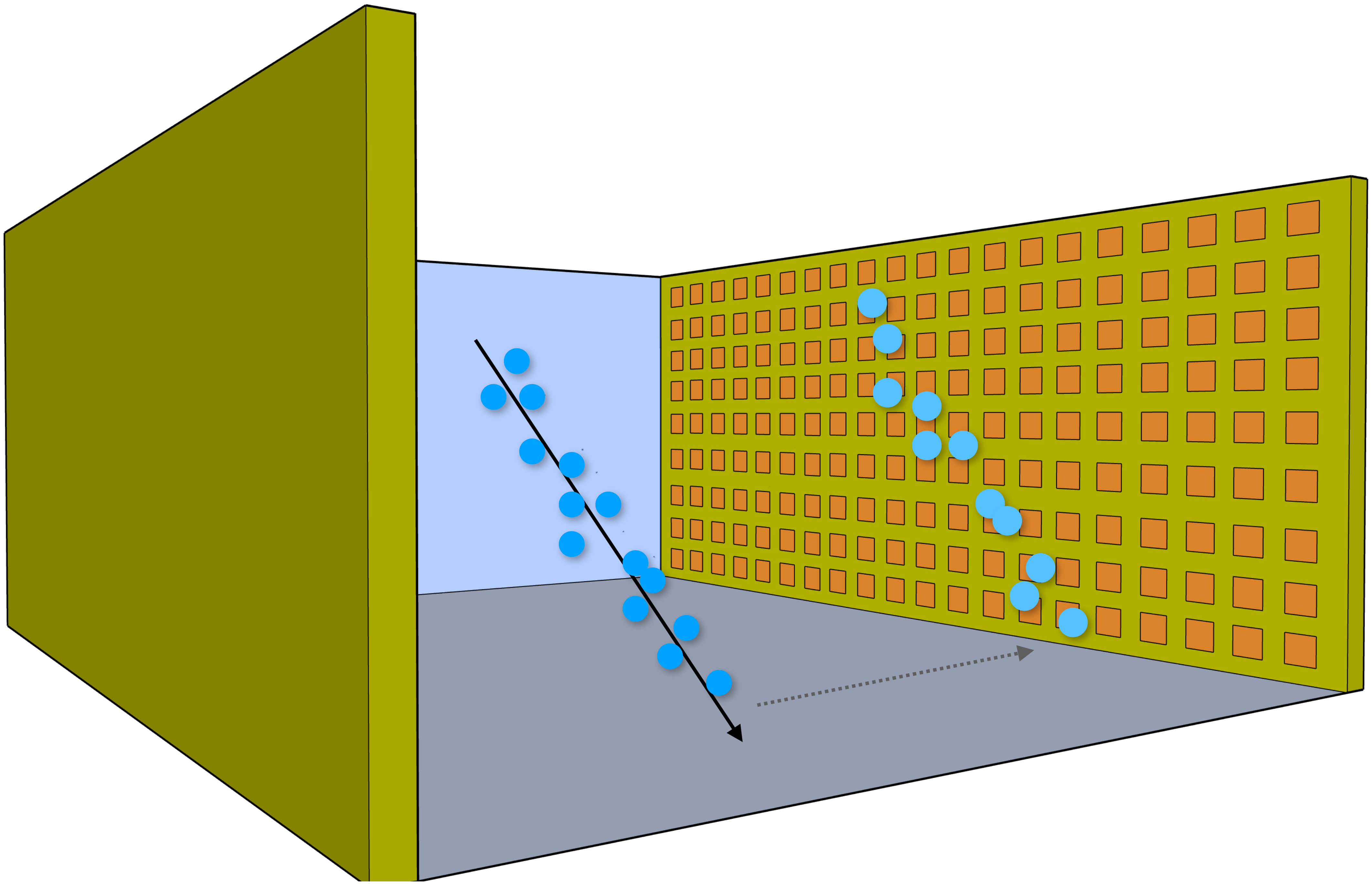
Well-understood detector modeling and calibration are vital

[Eur.Phys.J.C 80 (2020) 10, 978]



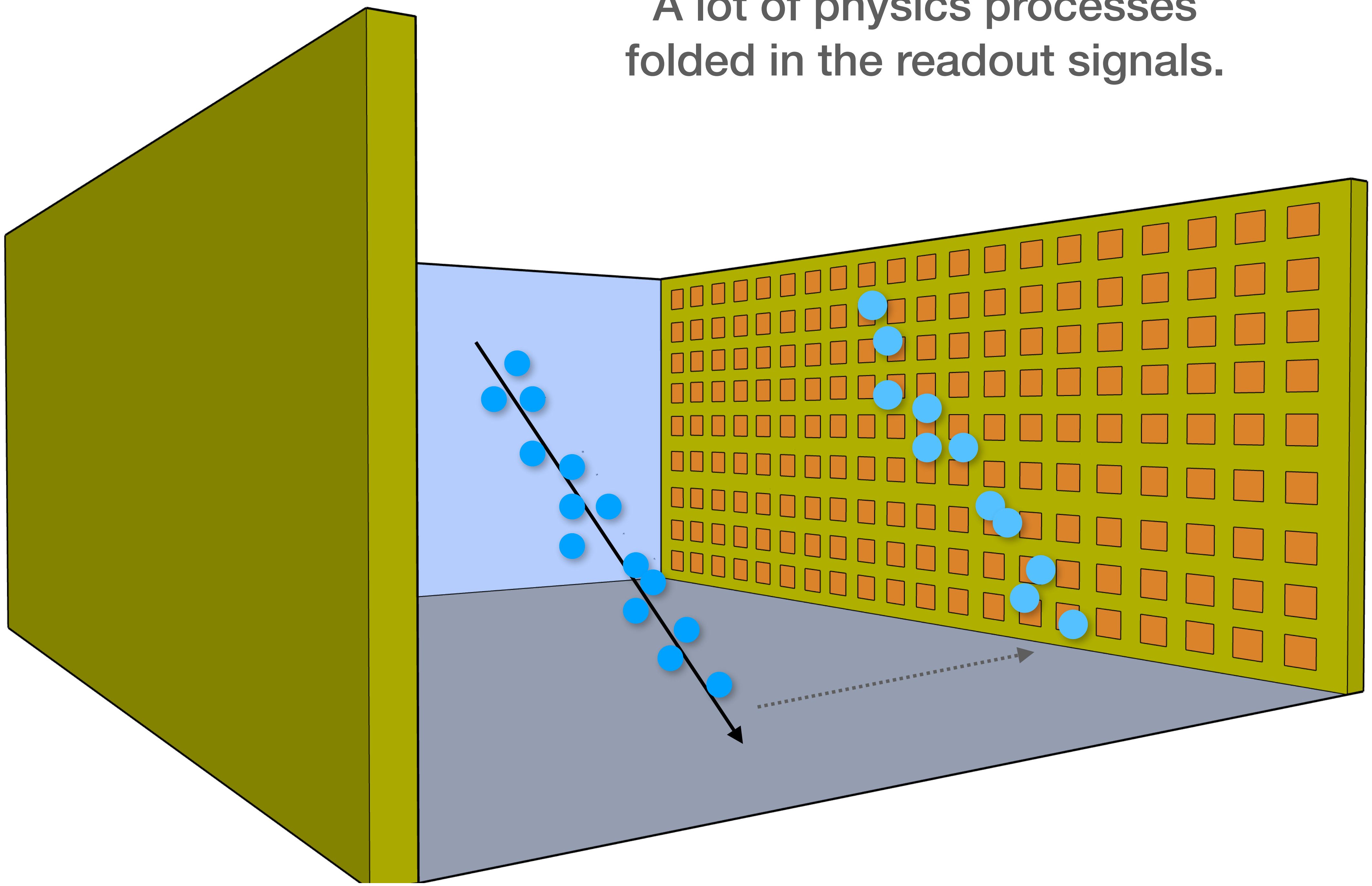


Liquid Argon Time Projection Chamber (LArTPC)



Liquid Argon Time Projection Chamber (LArTPC)

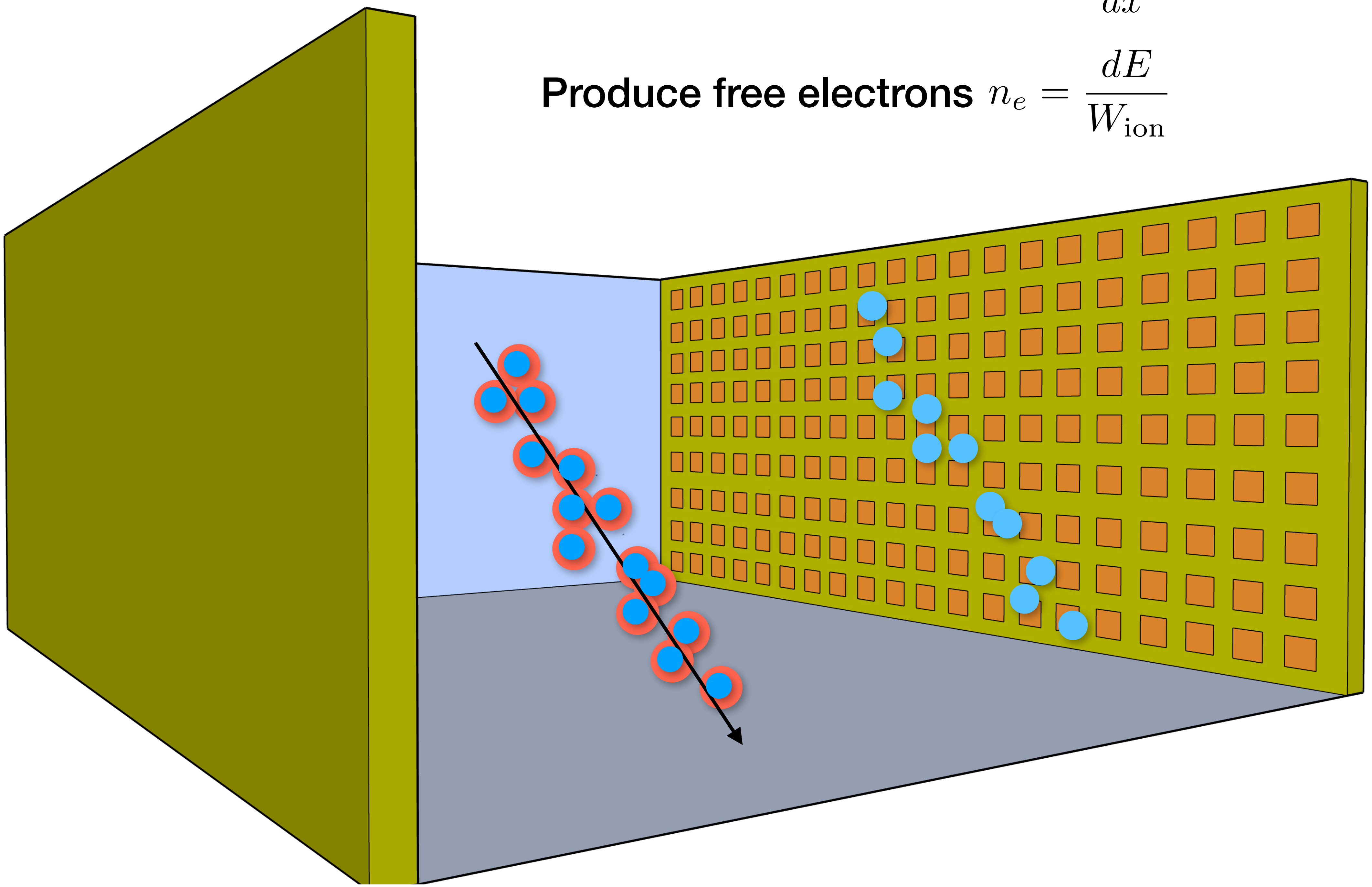
A lot of physics processes
folded in the readout signals.



LArTPC: Charge Production

Particle "segments" deposit $\frac{dE}{dx}$

Produce free electrons $n_e = \frac{dE}{W_{\mathrm{ion}}}$



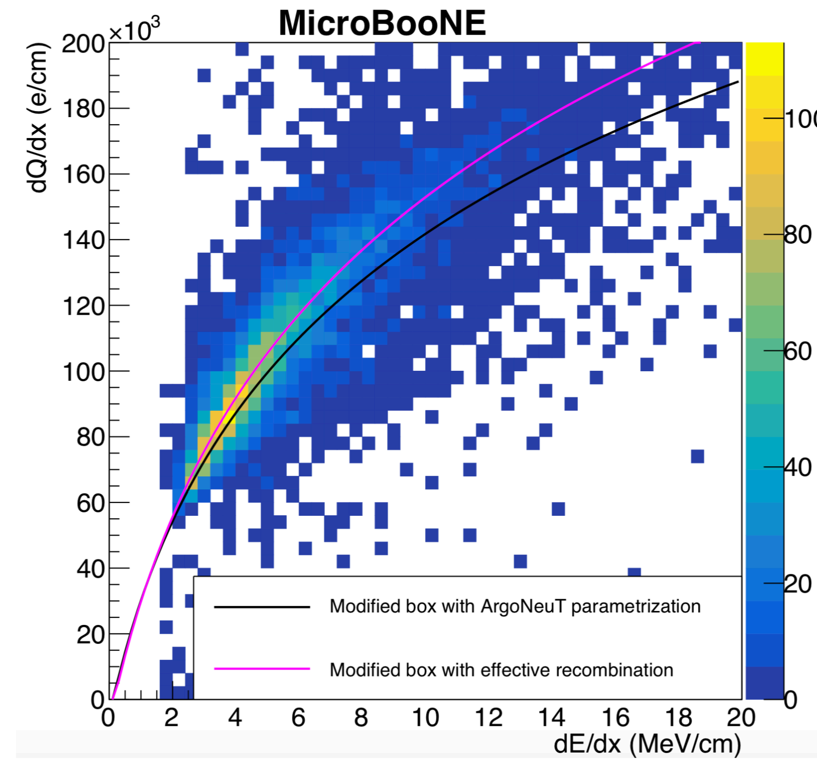
LArTPC: Charge Recombination

Number of electrons after recombination

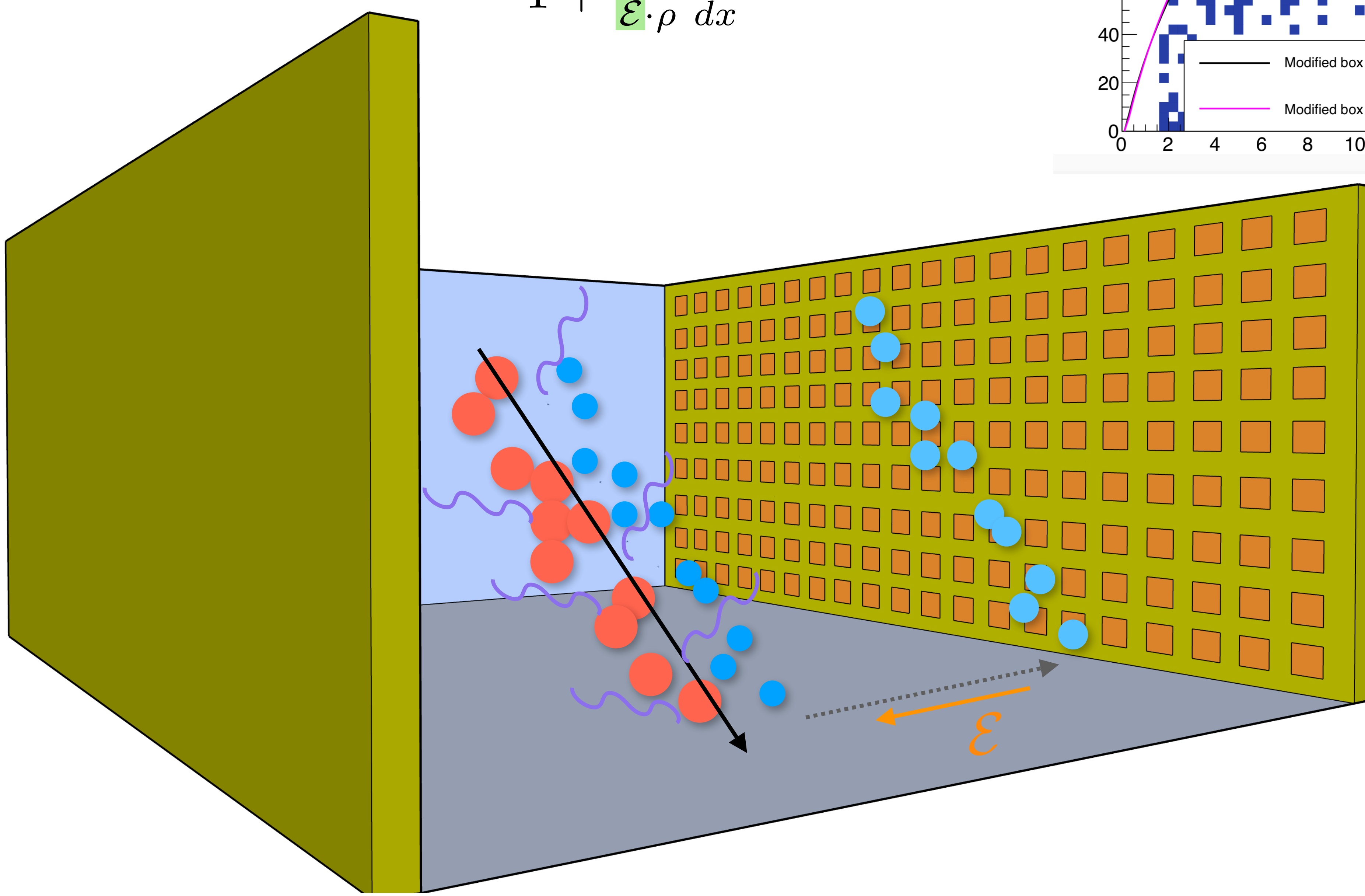
$$n_e^{\text{recomb}} = \alpha^{recomb} \cdot n_e$$

Recombination Birks model

$$\alpha^{recomb} = \frac{\boxed{A_{\mathrm{B}}}}{1 + \frac{\boxed{k_{\mathrm{B}}}}{\boxed{\mathcal{E}}\cdot\rho}\frac{dE}{dx}}$$



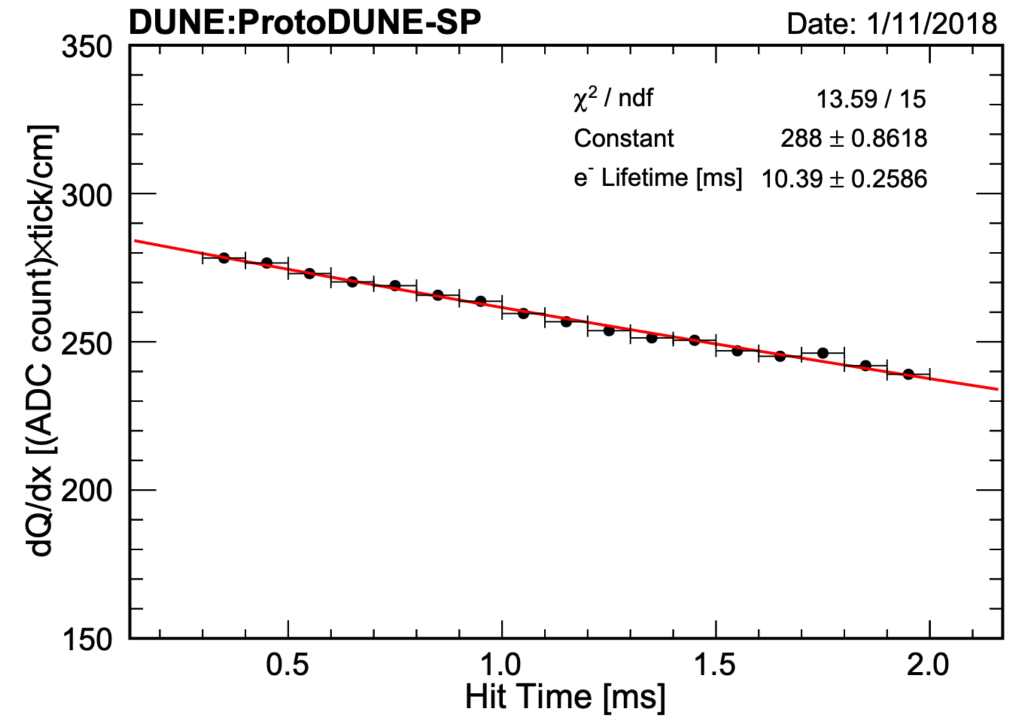
JINST 15, P03022 (2020)



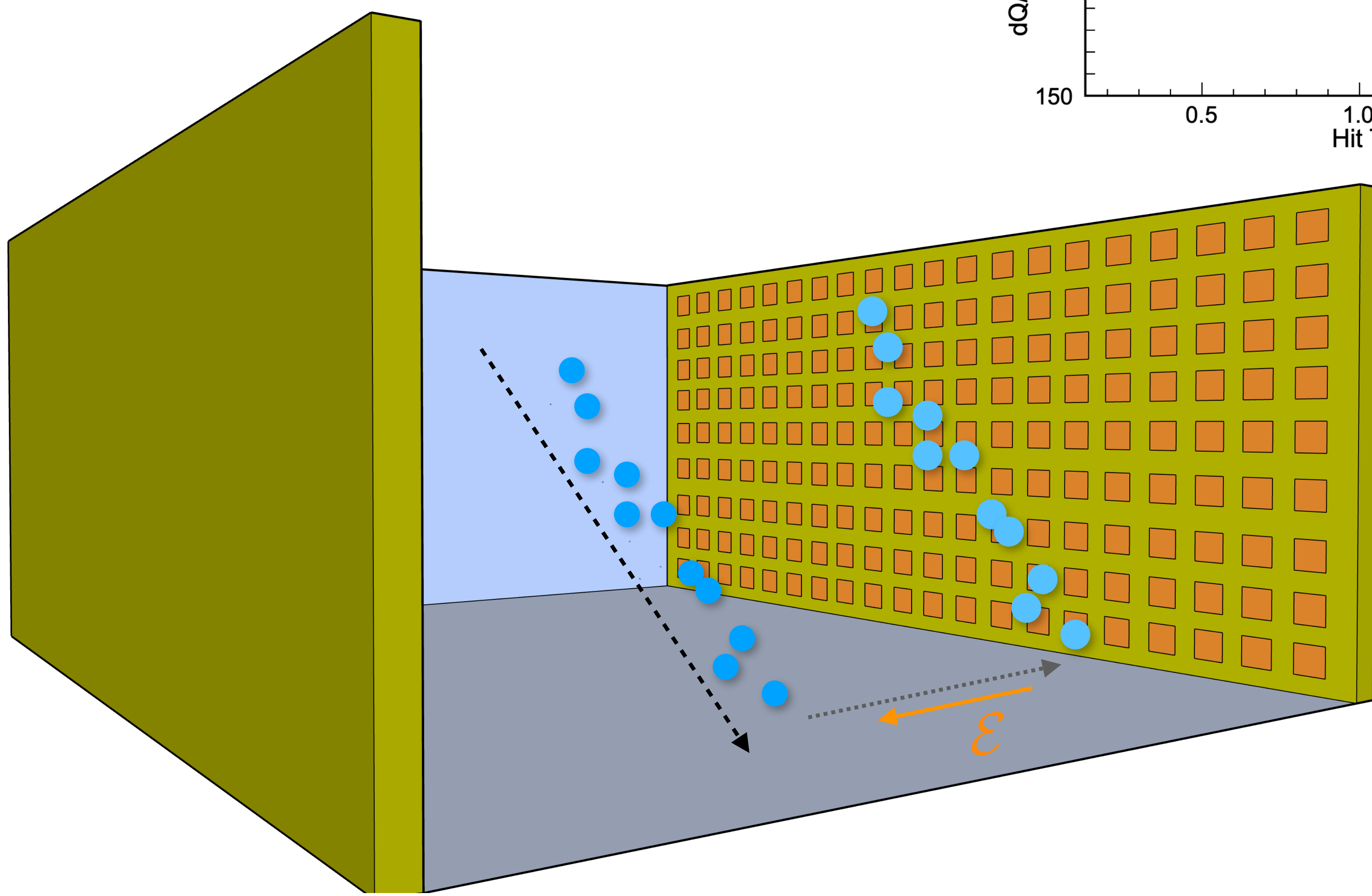
LArTPC: Charge Attenuation

Number of electrons survived charge attenuation

$$n_e^{\text{att}} = n_e^{\text{recomb}} \cdot e^{-\frac{t_{\text{drift}}}{\tau}}$$



JINST 15 (2020) 12, P12004

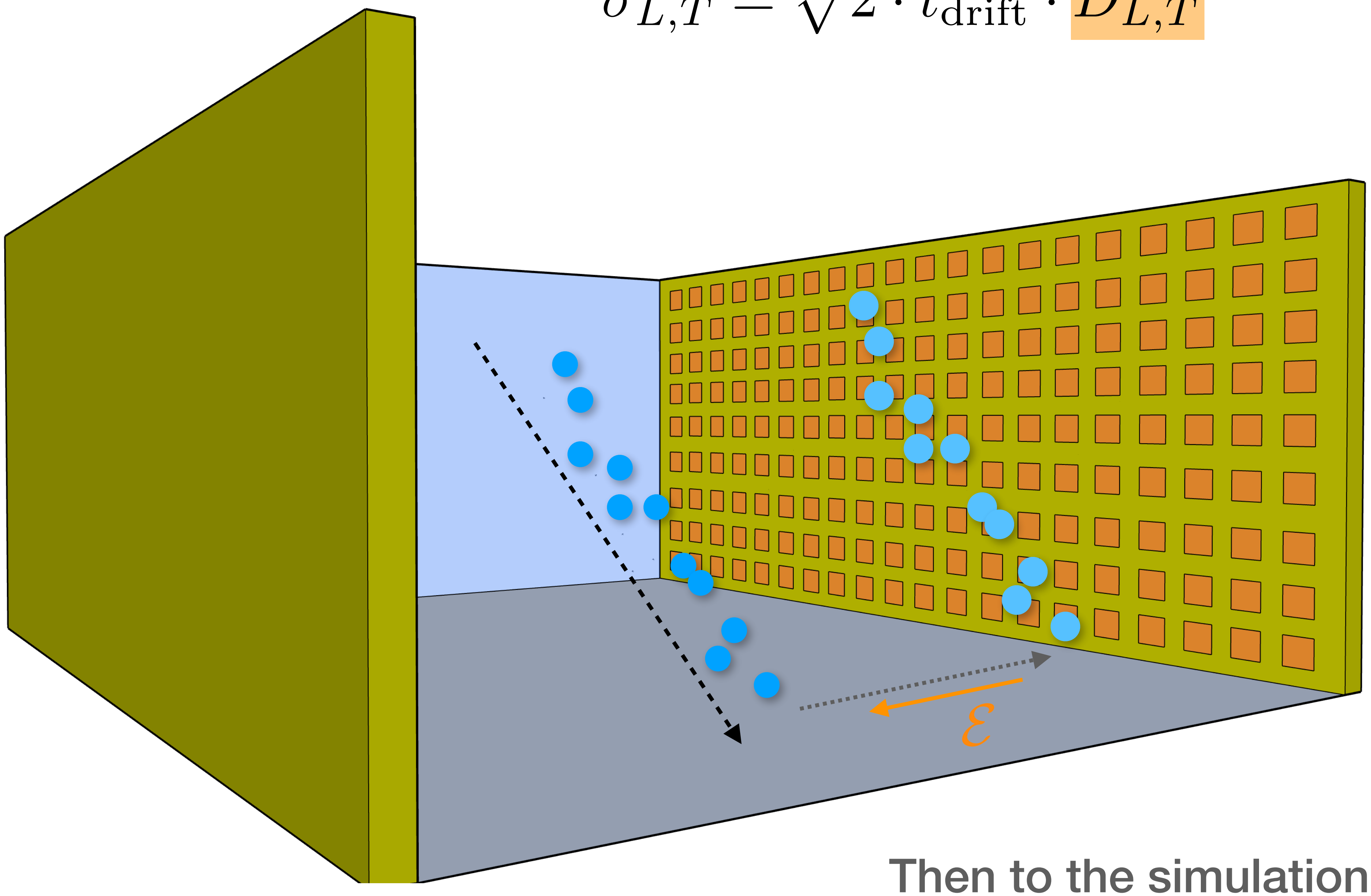


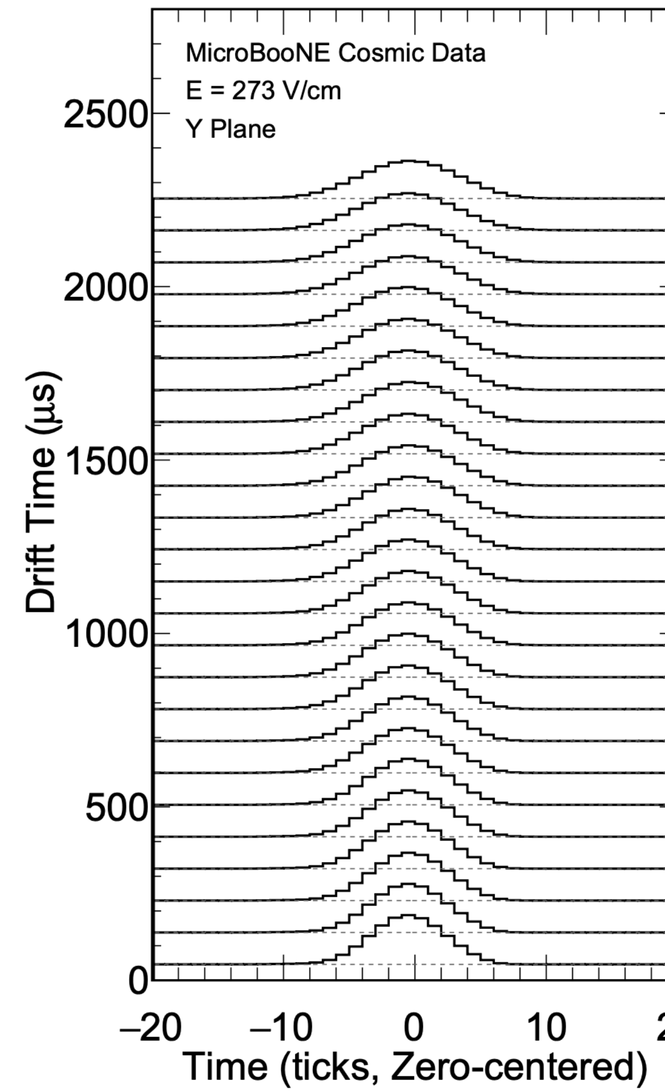
LArTPC: Charge Drift and Diffusion

$$v_{\mathrm{drift}} = \mu \cdot \mathcal{E}$$

$$\mathrm{drift\ position} = v_{\mathrm{drift}} \cdot t_{\mathrm{drift}}$$

$$\sigma_{L,T} = \sqrt{2 \cdot t_{\mathrm{drift}} \cdot D_{L,T}}$$



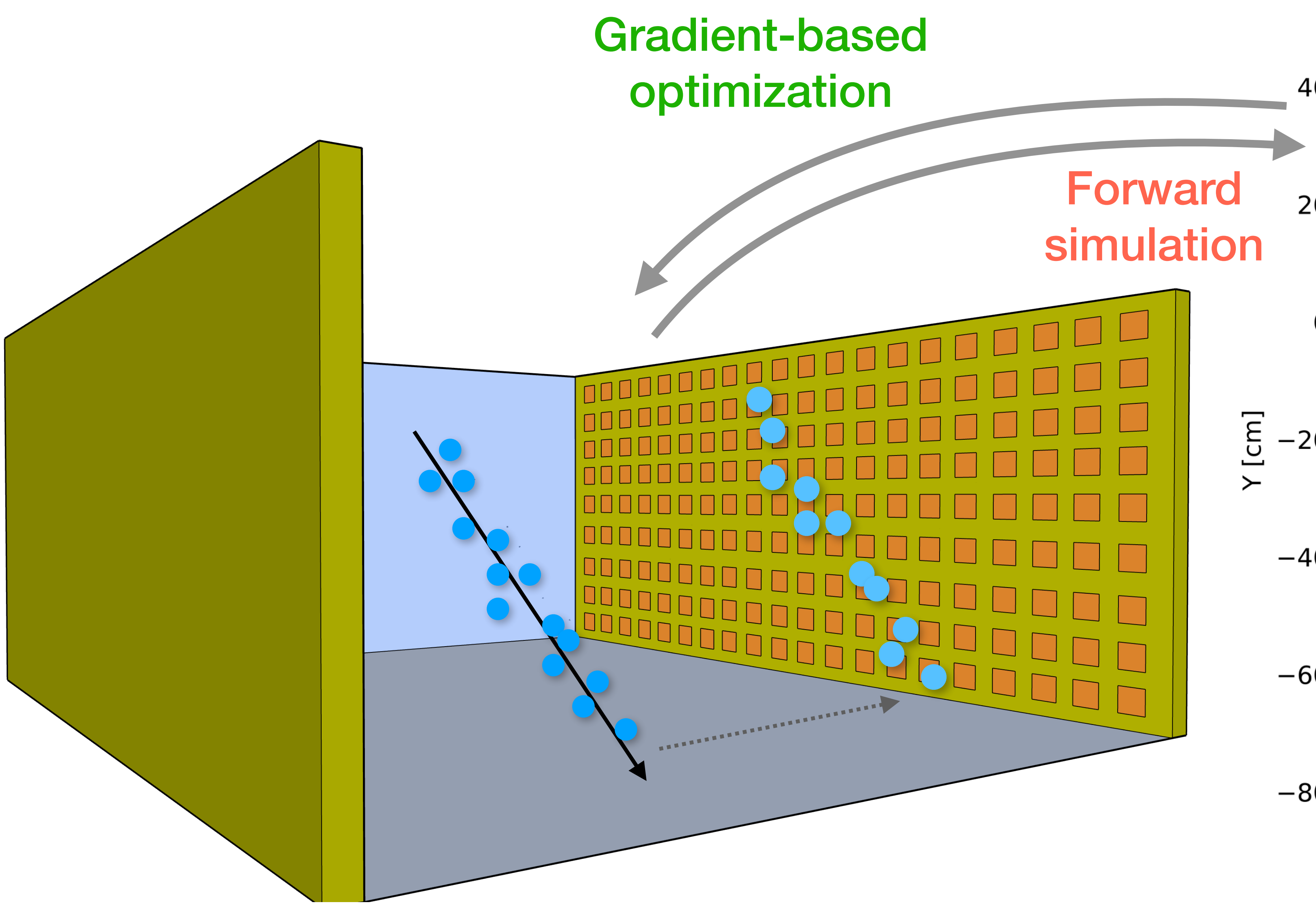


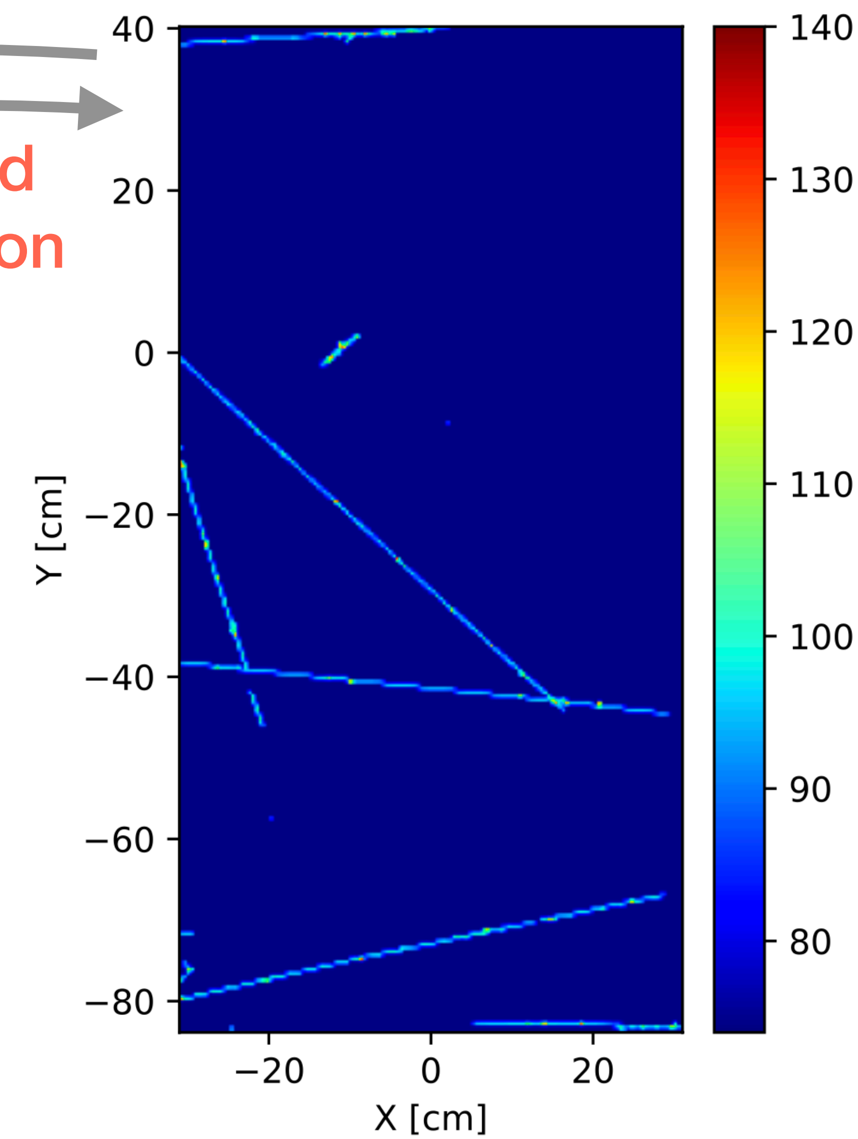
JINST 16 (2021) 09, P09025

Then to the simulation of electronic responses

Simultaneous High-dimensional Calibration

- Challenging for conventional calibration methods
- Development of a differentiable simulation for high-dimensional calibration
  - Simultaneous optimization for multiple model parameters
  - Straightforward application of the calibration
  - Improve simulation fidelity





Differentiable *larnd-sim*

Reference *larnd-sim*
JINST 18 (2023) 04, P04034

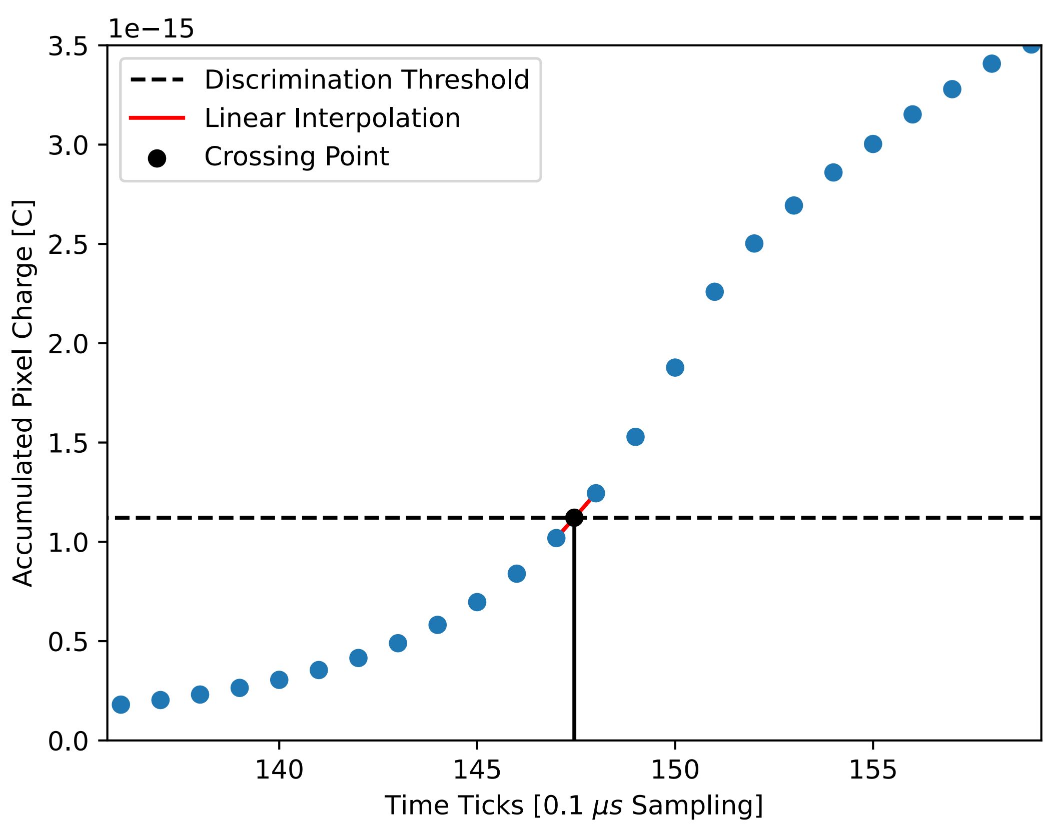
Differentiable *larnd-sim*
Mach. Learn.: Sci. Technol. 5 025012

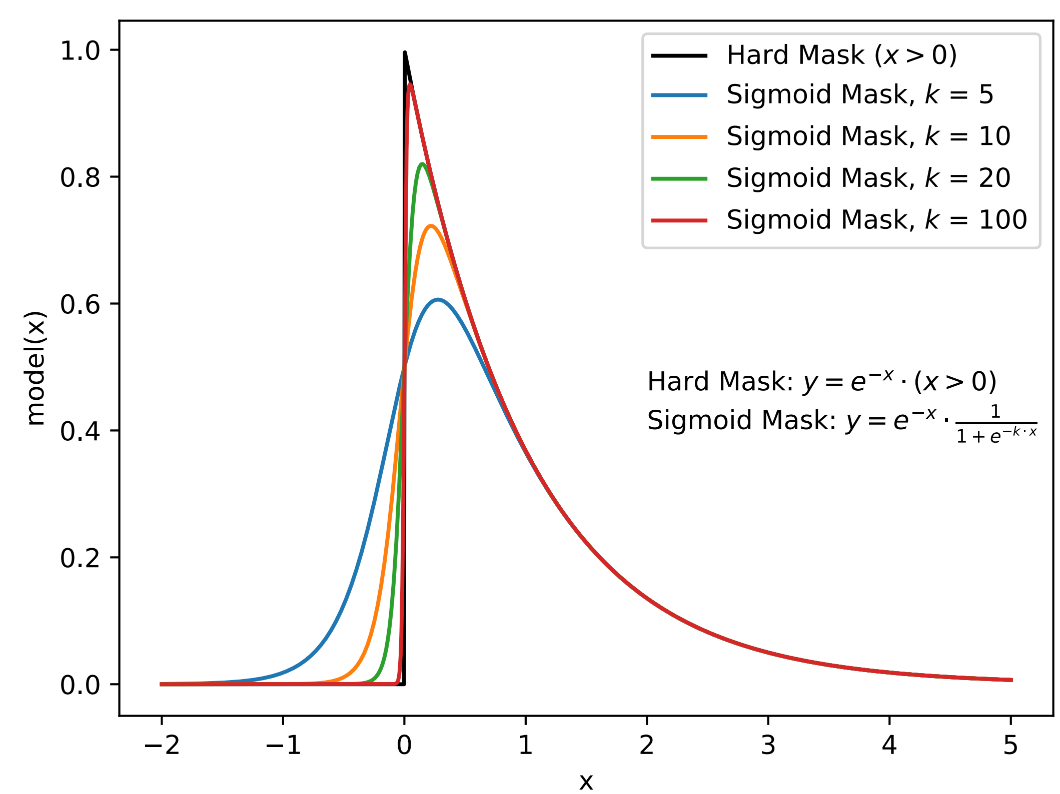
- *larnd-sim* is a detector simulator applicable for DUNE near detector LArTPC and its prototypes
- Rewrote using EagerPy (agnostic to automatic differentiation backend)
- Chose PyTorch for this demonstration

Differentiable relaxation

Important for parameter related operations

- Integer operations → floating point
- Discrete sampling with threshold → interpolation
- Conditional operation (hard mask) → sigmoid threshold



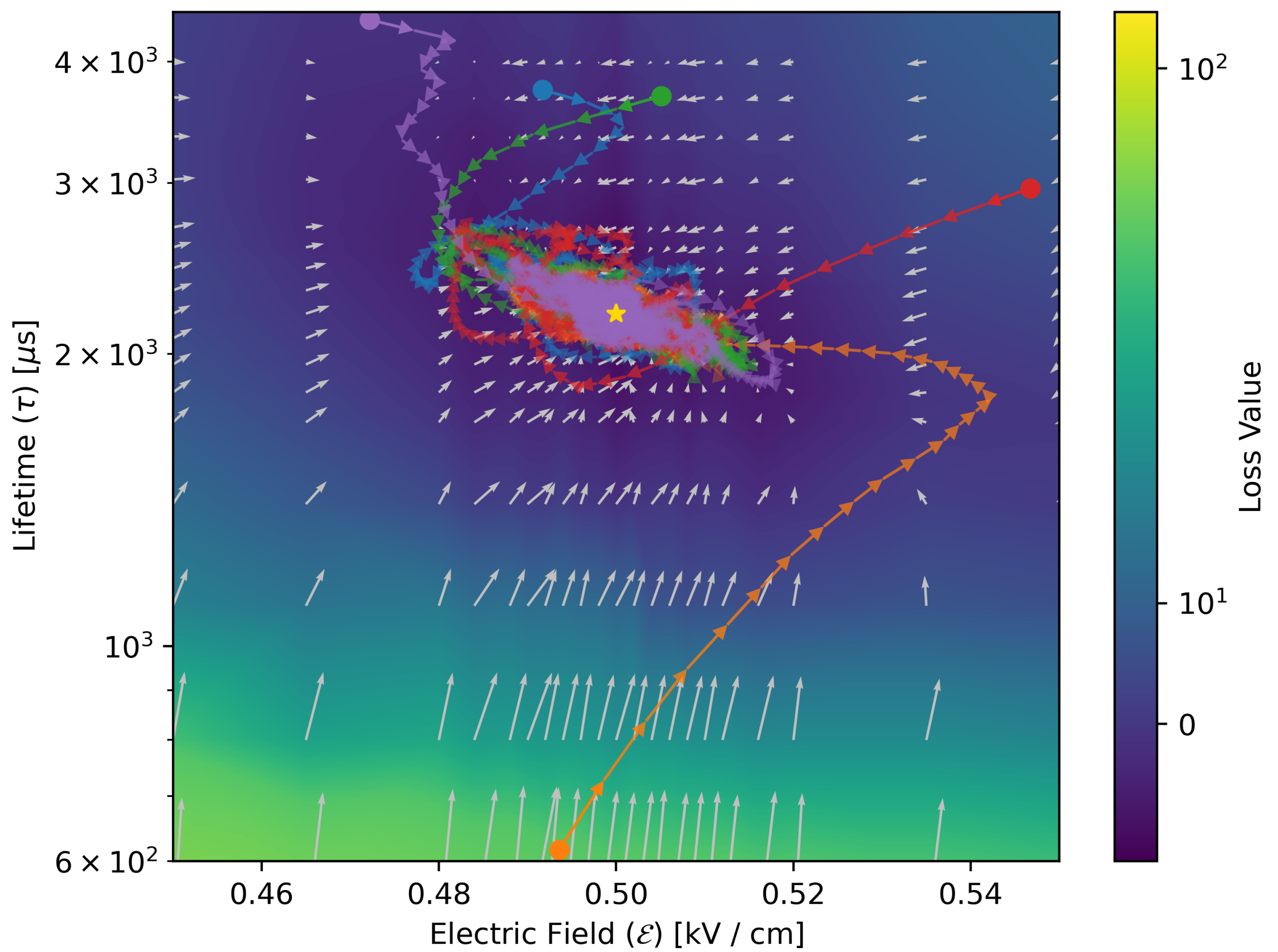


Calibration: Fit the Model Parameters

Input particle segments (position and energy deposition): χ
Model parameters: θ
Differentiable simulation: $f(\chi, \theta)$
Target data: F_{target}

1. Choose the initial parameter values θ_0
2. Run the forward simulation $f(\chi, \theta_0)$
3. Compare the simulation output and the target data with a loss function $\mathcal{L}(f(\chi, \theta_0), F_{\text{target}})$
4. Calculate gradients for the parameters $\nabla_\theta \mathcal{L}(f(\chi, \theta_0), F_{\text{target}})$
5. Update parameter values $\theta_0 \to \theta_i$ based on the gradients

   Iterate step 2. to 5.



For gradient descent, the parameter update takes form of

$$\theta_{i+1} = \theta_i - \eta \cdot \nabla_\theta \mathcal{L}(f(\chi, \theta_i), F_{\text{target}})$$

We use Adam for the optimiser

About the Fit

Parameter [Units]	Nominal Value	Range
A_B	0.8	[0.78, 0.88]
k_B $[kV.g/cm^3/MeV]$	0.0486	[0.04, 0.07]
$\mathcal{E}$ $[kV/cm]$	0.5	[0.45, 0.55]
τ $[\mu s]$	2200	[500, 5000]
D_L $[cm^2/\mu s]$	4×10^{-6}	$[2 \times 10^{-6}, 9 \times 10^{-6}]$
D_T $[cm^2/\mu s]$	8.8×10^{-6}	$[4 \times 10^{-6}, 14 \times 10^{-6}]$

- Normalize the parameters with their nominal values for gradient calculation
- Gradient clips on normalized gradient
- Use the same learning rate for all parameters
- Exponential learning rate decay
- Use Soft Dynamic Time Warping (Soft DTW) for the loss
- Recover the parameter values for the forward simulation

Loss Function: Soft DTW

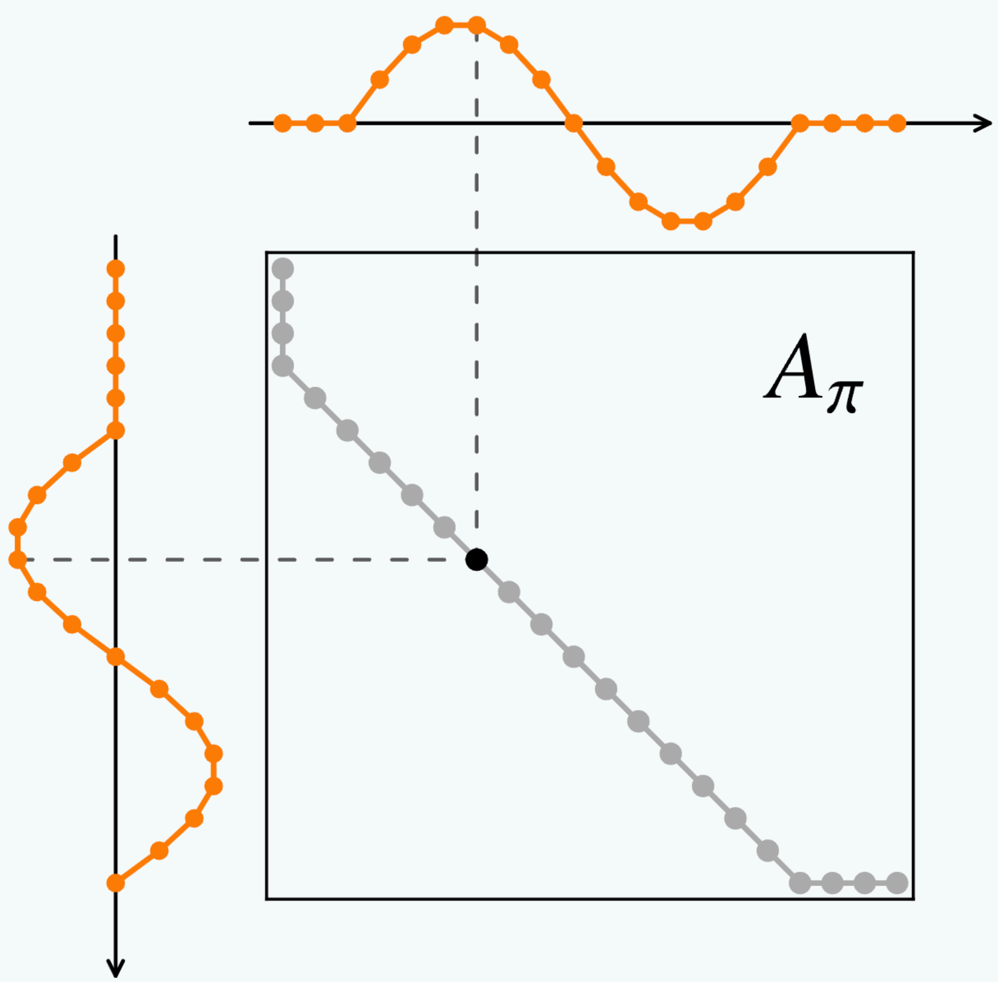
The choice of the loss function is important to the fit performance.

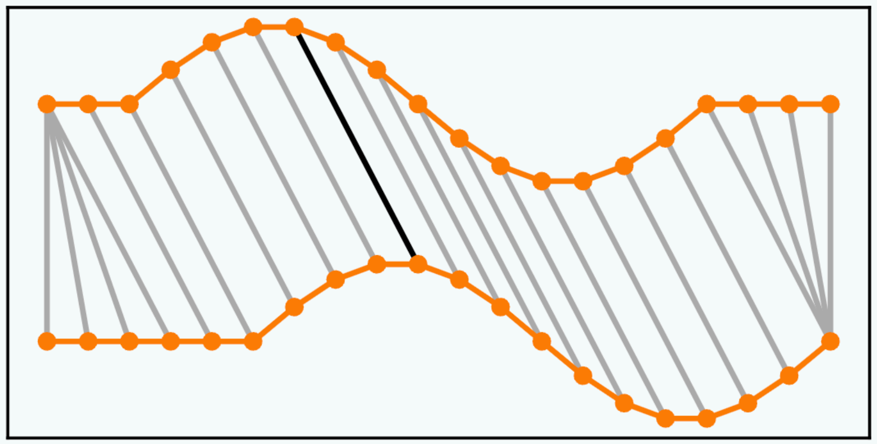
- High discriminating power of the best parameter values
- Differentiable

Challenges for sparse data

- Potentially different length of the simulation output and the target data
- Obscure correspondence between hits from the two sets

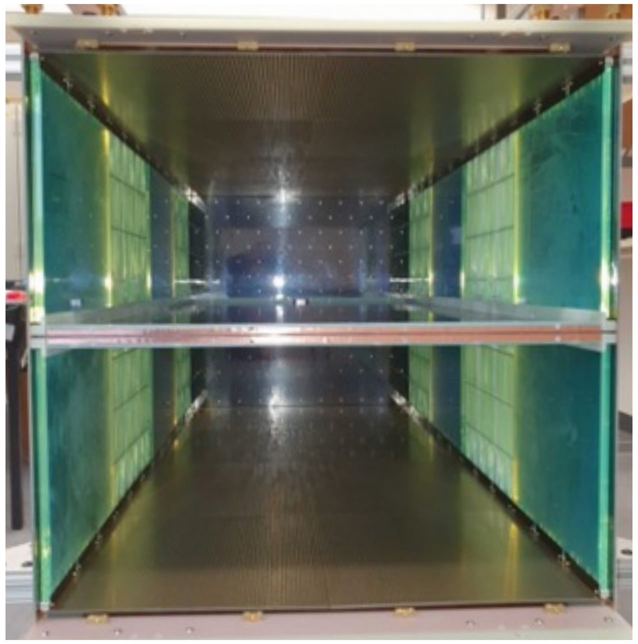
Dynamic Time Warping (DTW) addresses the alignment challenges.
Soft DTW is a differentiable version of it.



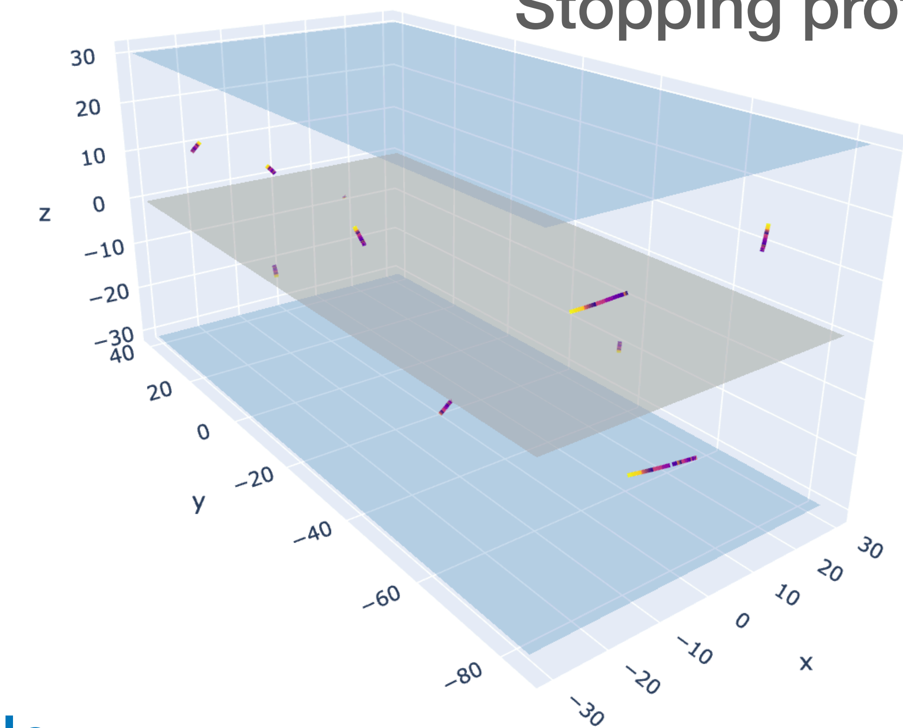


Romain Tavenard
Post: [DTW]
Post: [Soft DTW]

Samples, Selection and Mini-batches



2403.03212 [physics.ins-det]

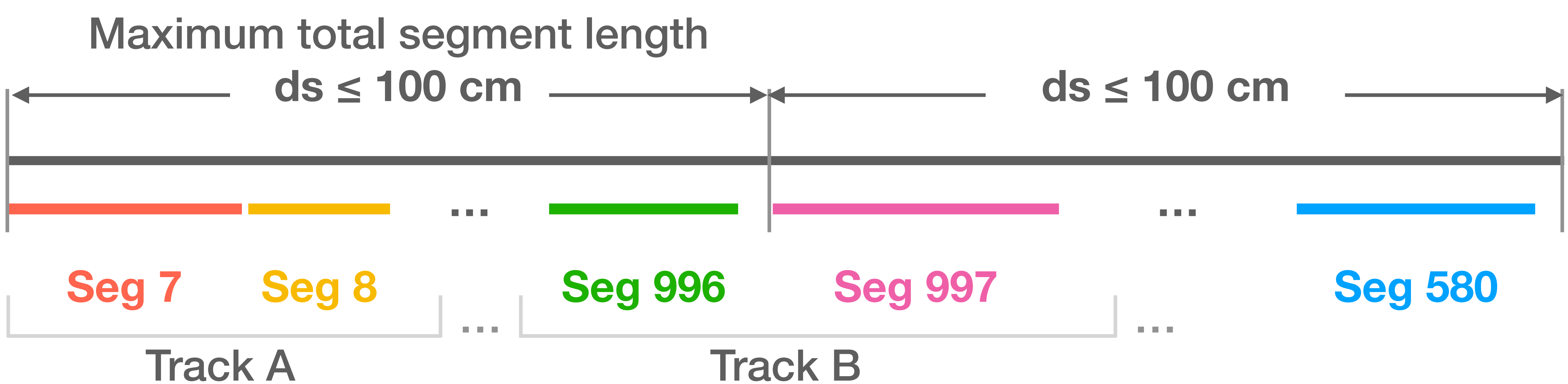


Data sample:

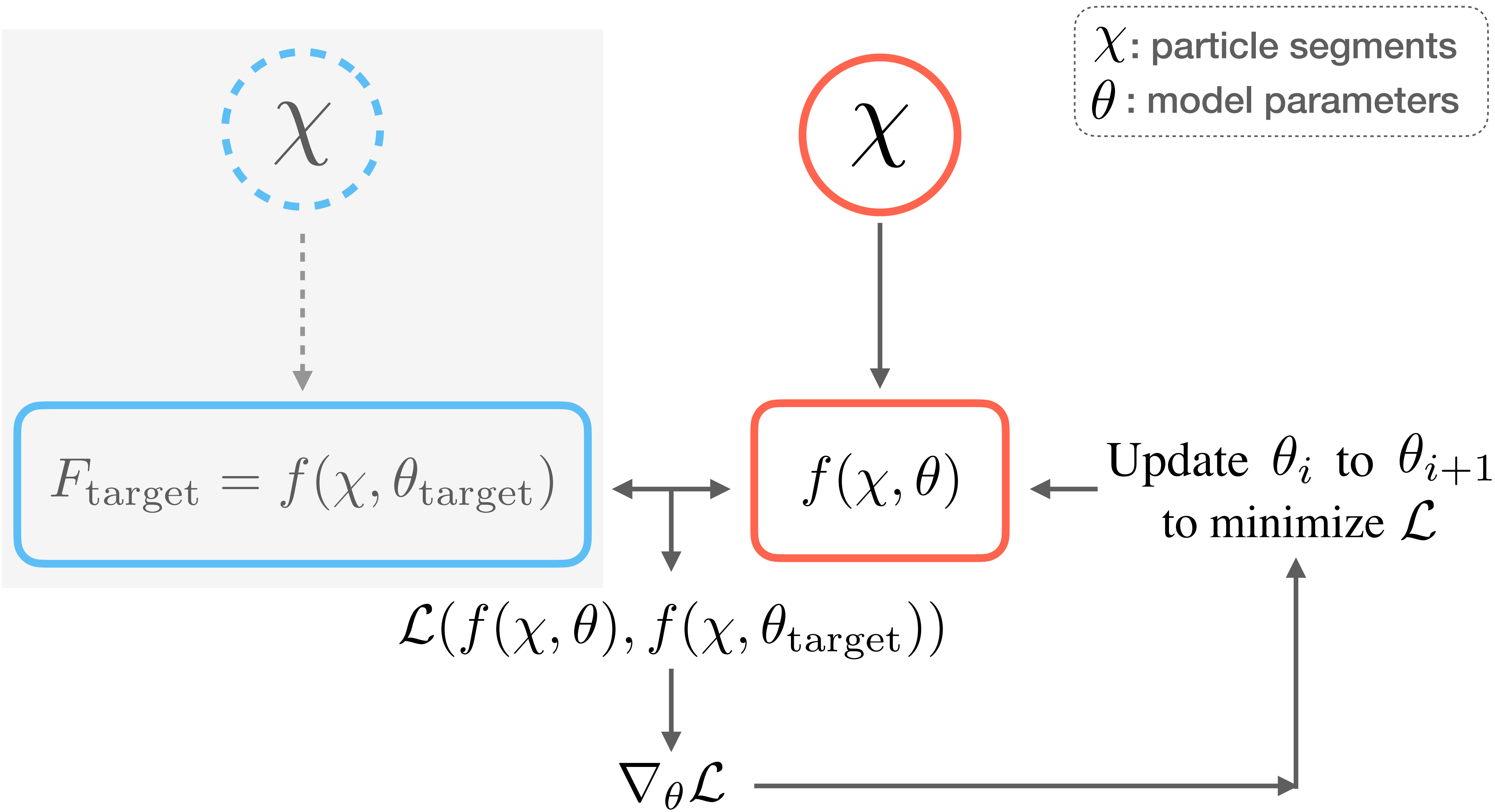
5000 events of ~10 protons with momentum of 200-500 MeV/c.

Selection:

- abs(track segment z) in [15, 28] cm (the half near the anodes)
- Track angle with respect to the drift axis larger than 15°



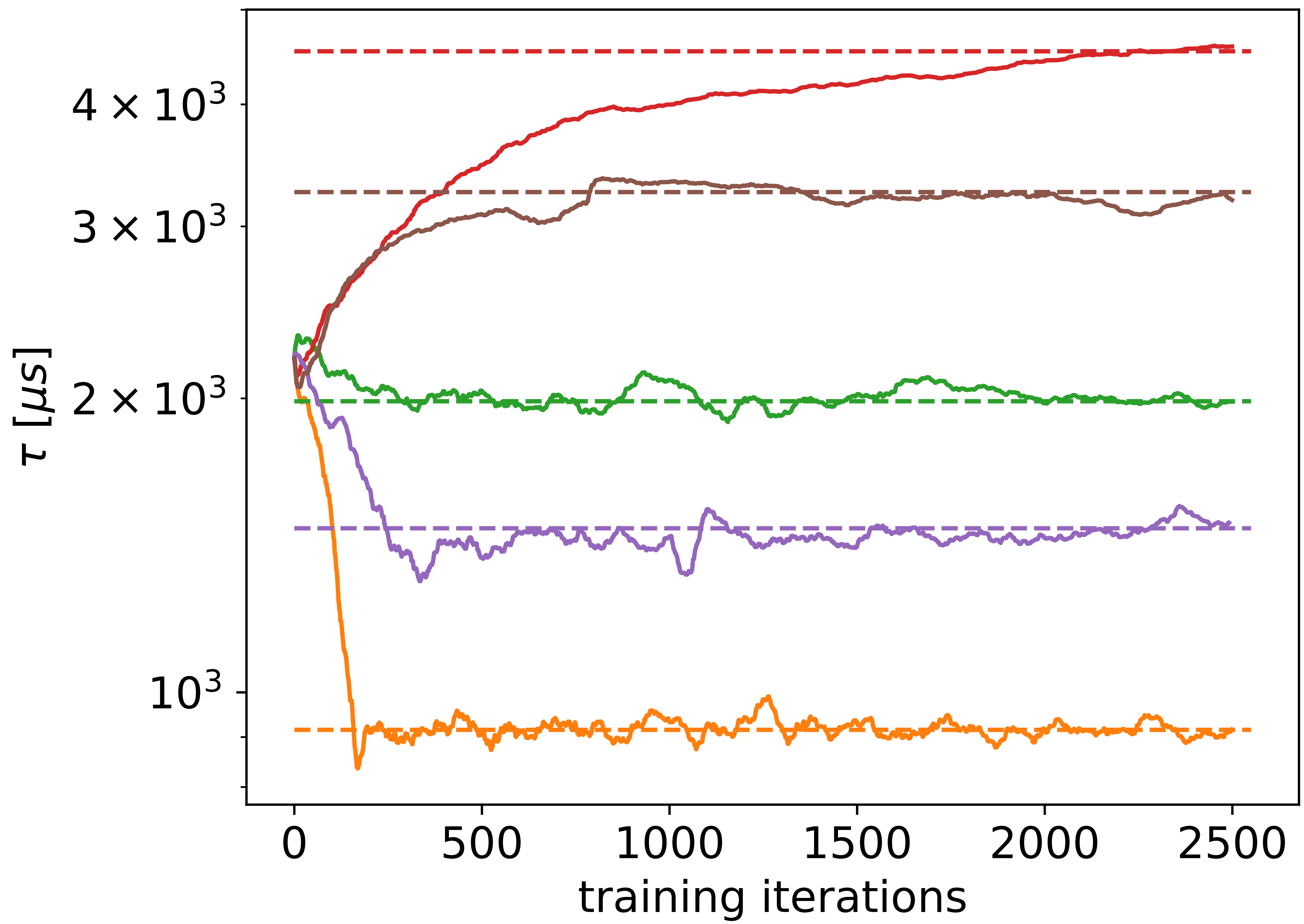
Closure Test of the Fit



Evaluate if θ converge to θ_{target}

Fit Result

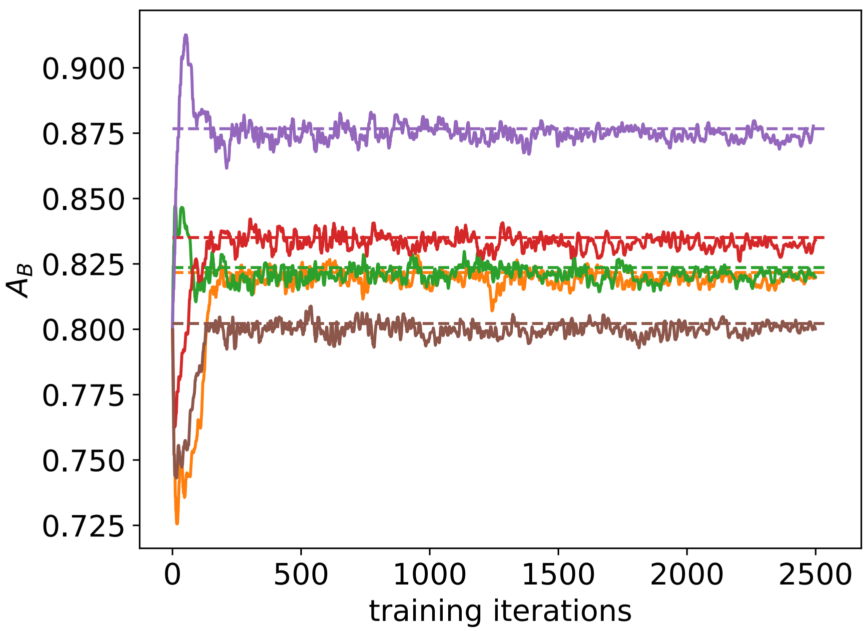
5 fits with different targets in 6D phase space.
The fits use 100 cm mini-batch.



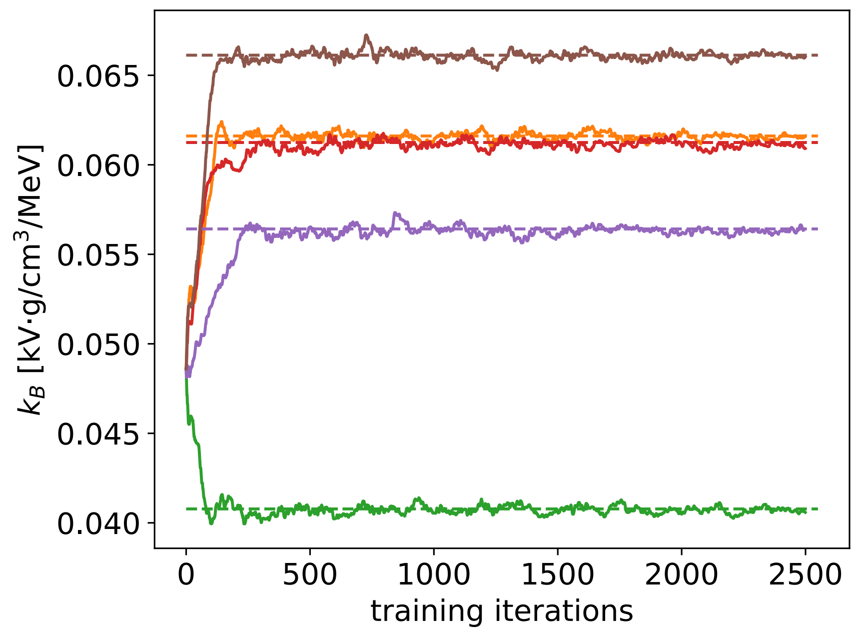
A Full Picture of the Fit Result

5 fits with different targets in 6D phase space.
The fits use 100 cm mini-batch.

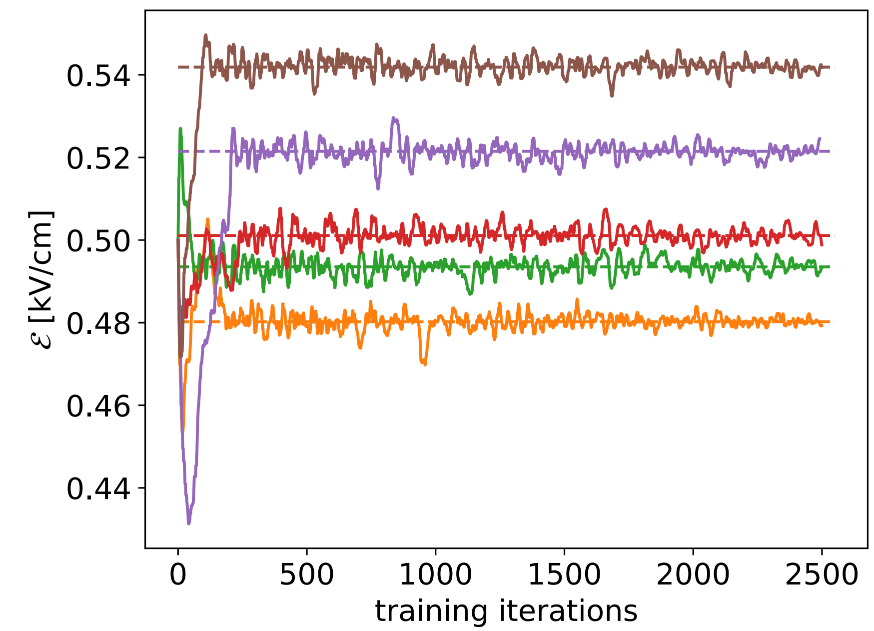
Recombination model A_B



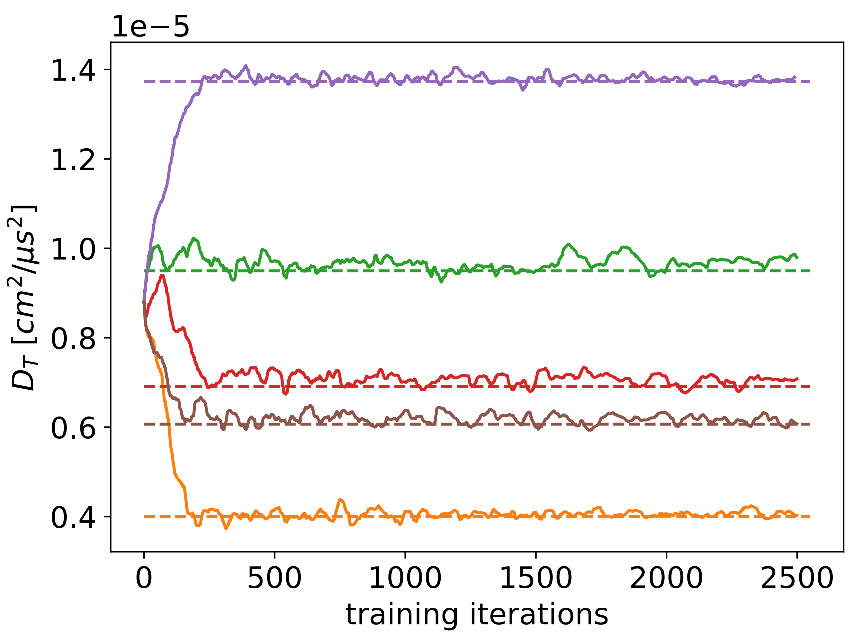
Recombination model k_B



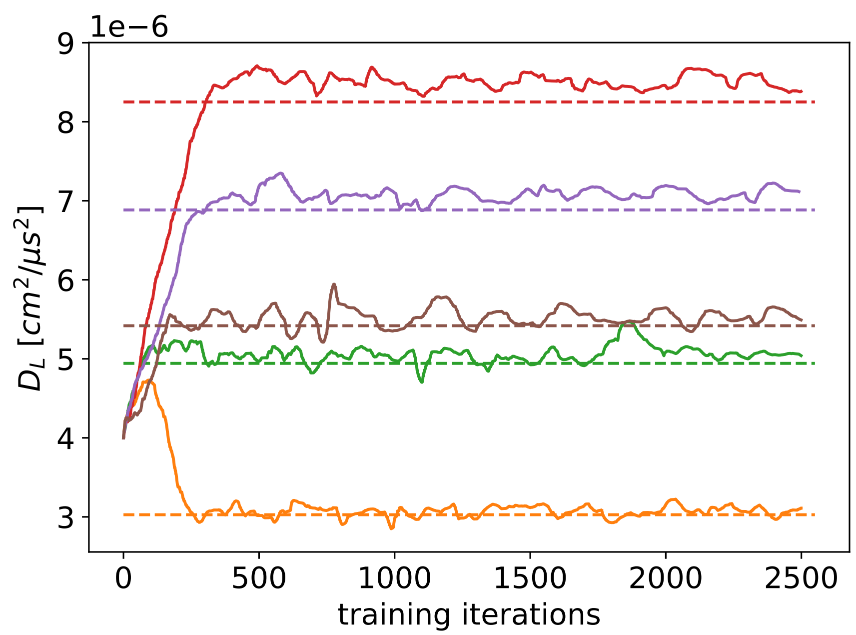
Electric field



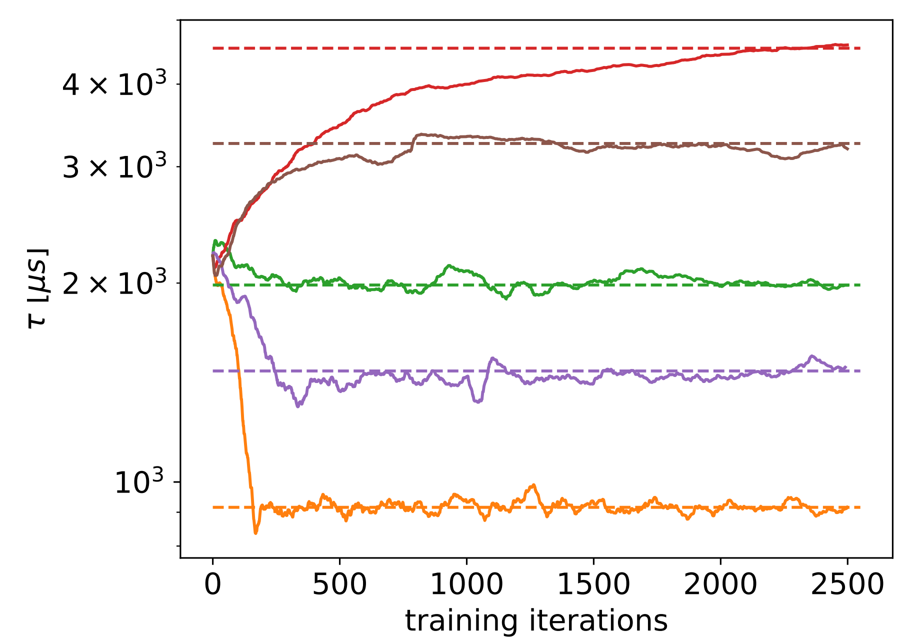
Transverse diffusion coefficient



Longitudinal diffusion coefficient

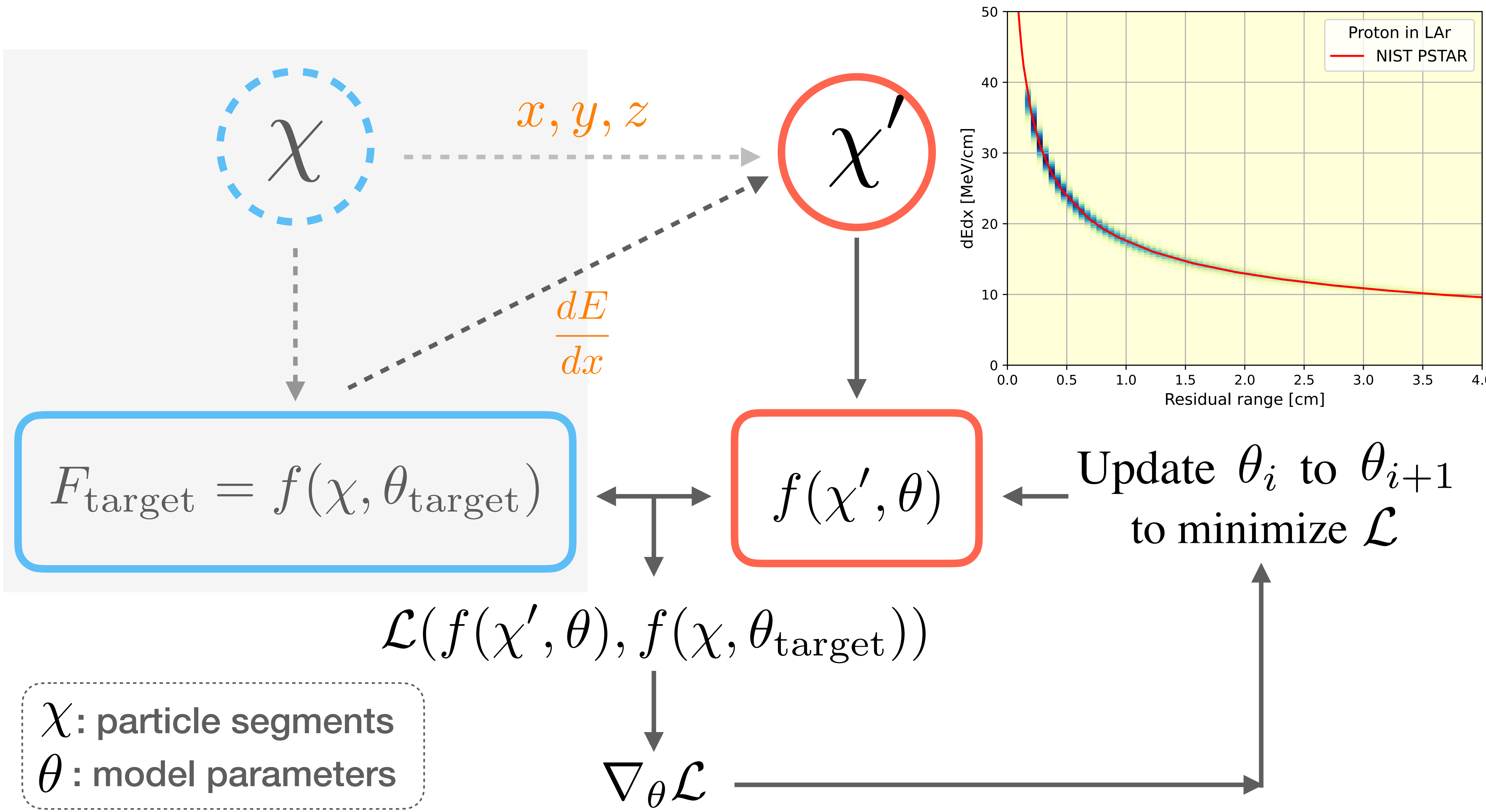


Electron lifetime



Data Application: From the Readout to the Simulation Input

Updated closure test



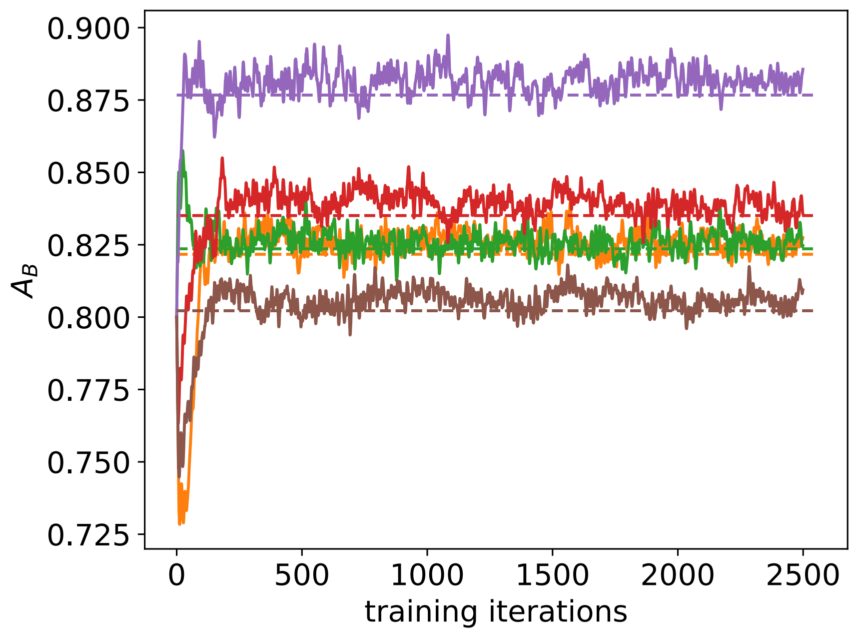
Evaluate if θ converge to θ_{target}

Fit Result

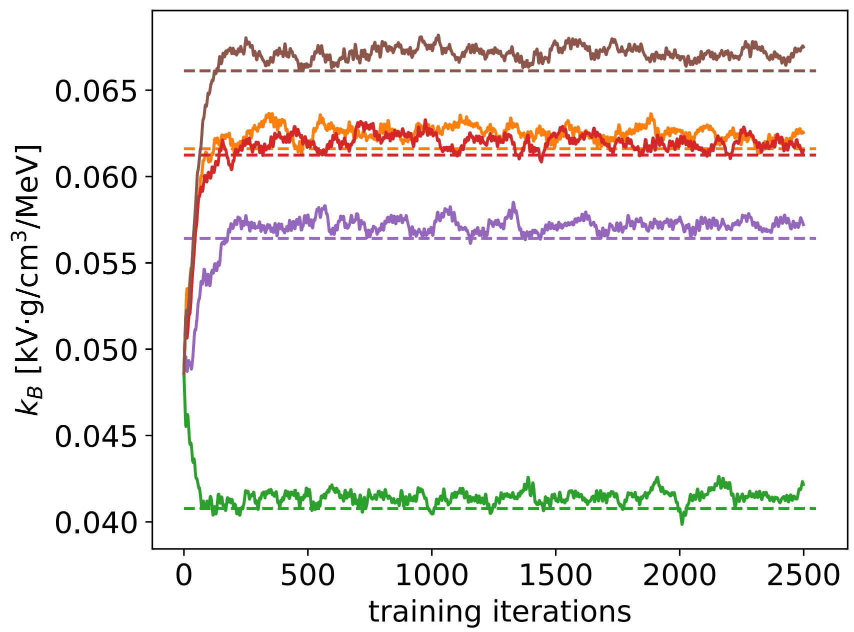
Work in progress

5 fits with different targets in 6D phase space.
The fits use 100 cm mini-batch.

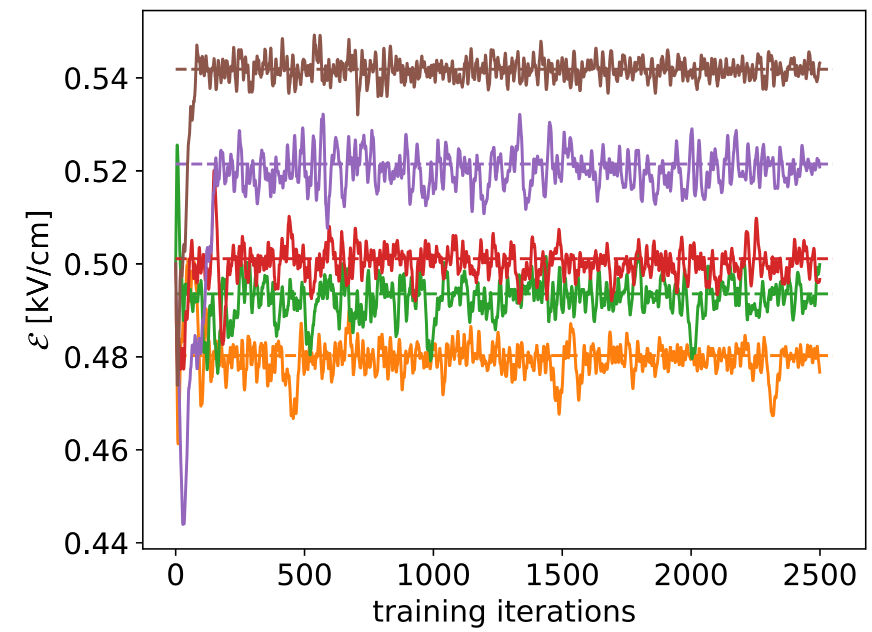
Recombination model A_B



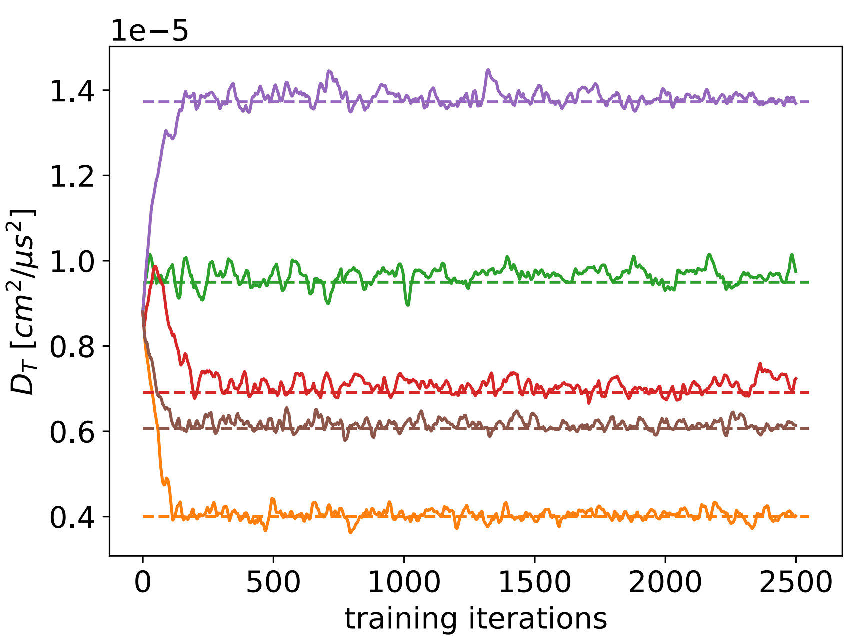
Recombination model k_B



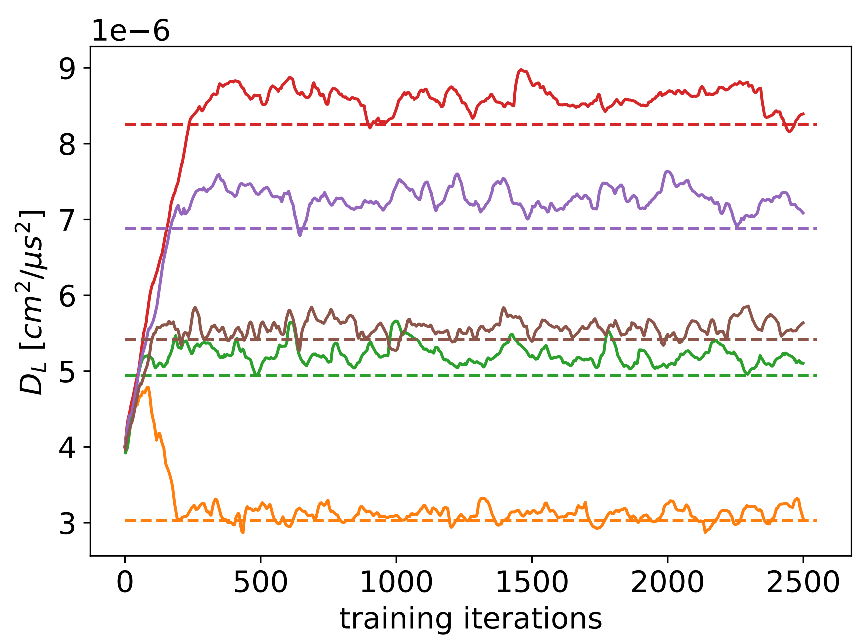
Electric field



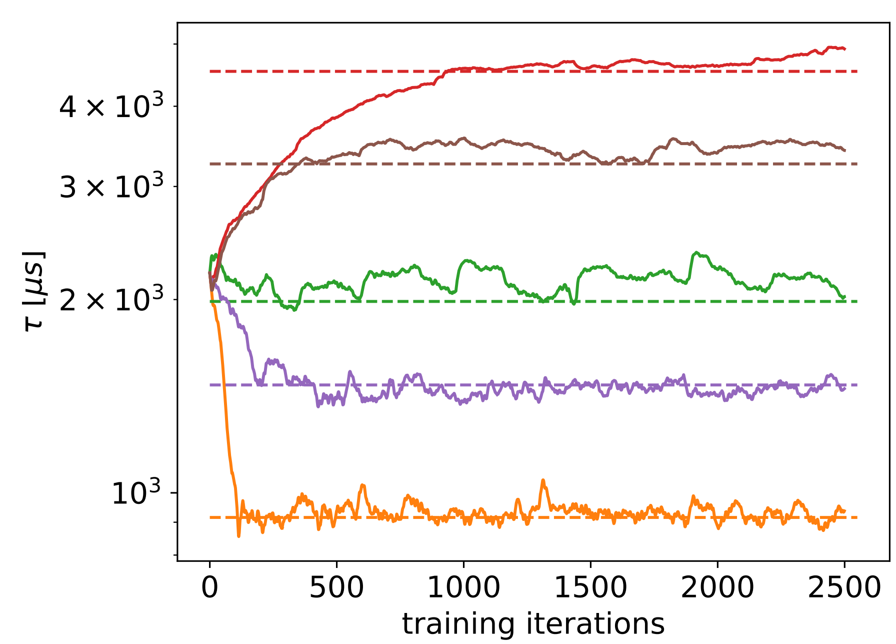
Transverse diffusion coefficient



Longitudinal diffusion coefficient

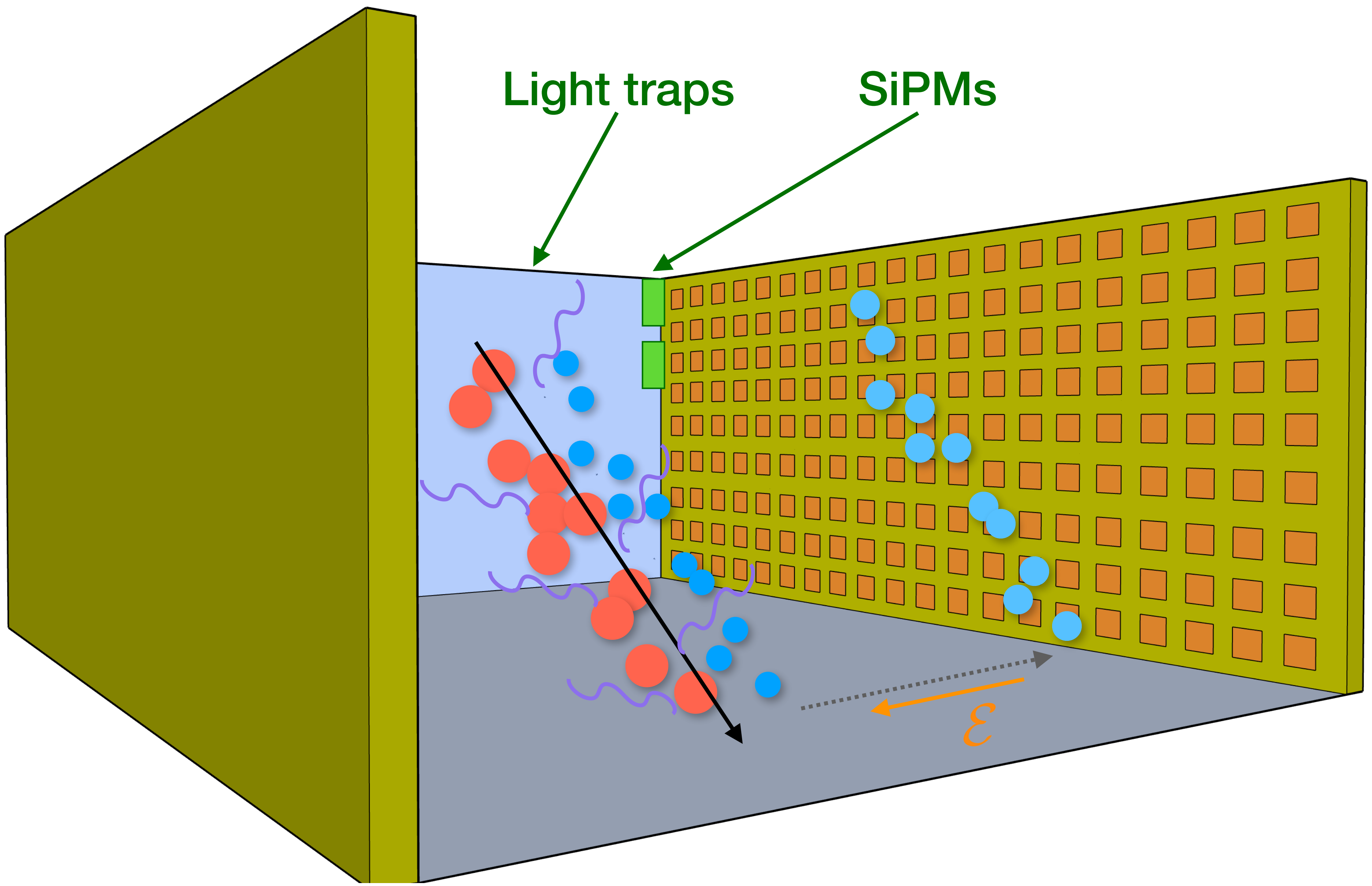


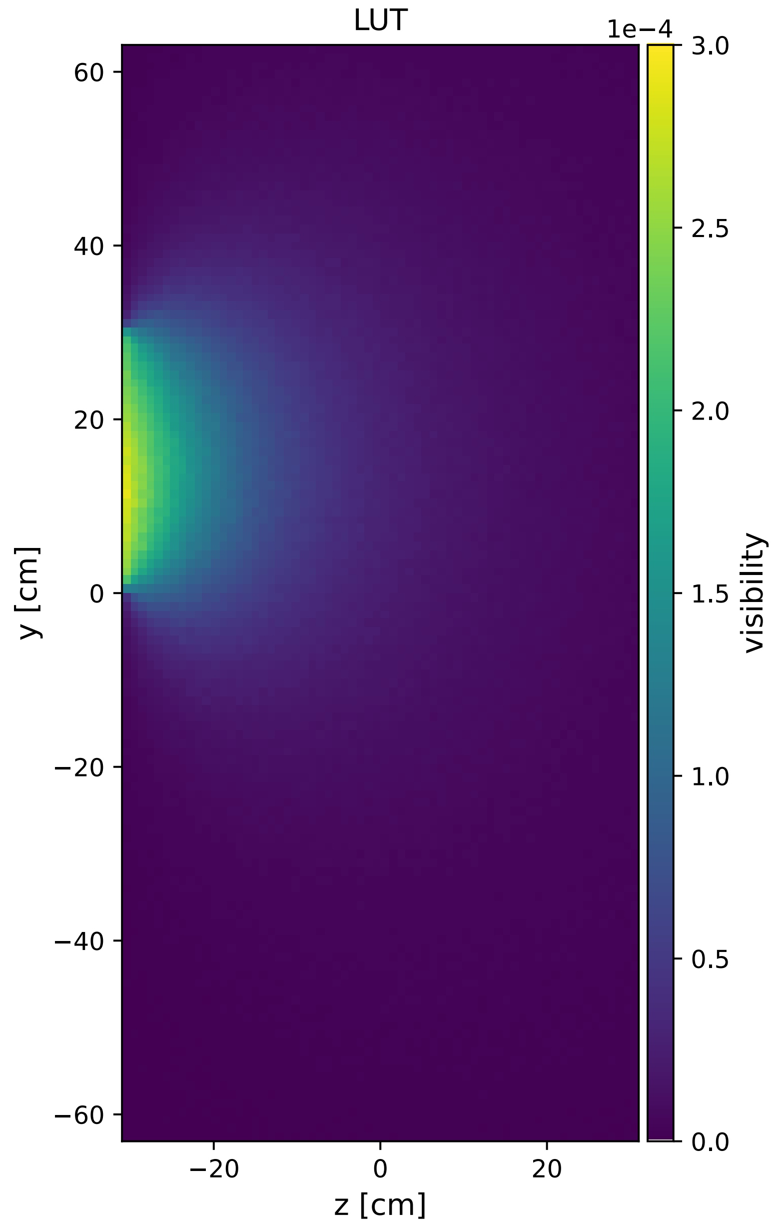
Electron lifetime

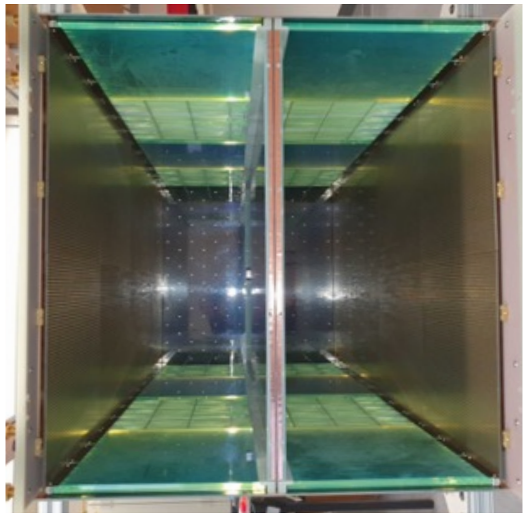


LArTPC: Light propagation

- Number of photons being produced (anti-correlated with charge survival probability per recombination)
- Standalone Geant4 simulation for photon observation probability at the readout and photon arrival time (photon library concluded in a loop-up table, LUT)
- Liquid argon scintillation singlet and triplet decays
- Readout electronic response

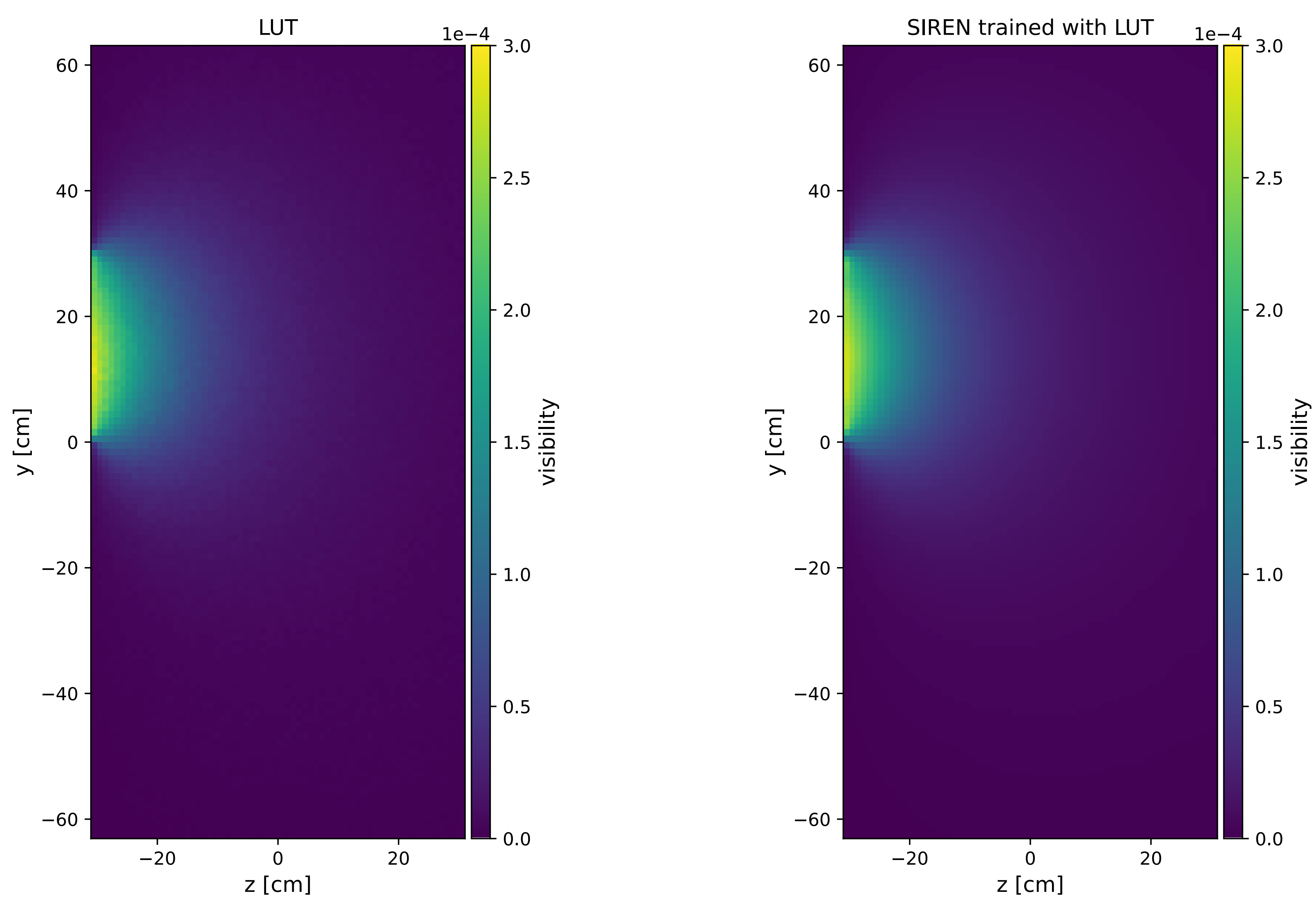






Photon Library with a Differentiable SIREN

- Conventional photon library: LUT that's discrete, sampled, not scalable with respect to the detector size, spatial resolution and readout channels
- Use SIREN (Sinusoidal representation network) for a continuous, differentiable photon library (surrogate model, with O(100-1000) times fewer parameters)
- LUT can be sensitive to the Poisson fluctuations depending on the initial Geant4 simulation statistics
- SIREN models better representing the underlying distribution than the LUT's
- SIREN trained with the LUT using weighted L2 loss of voxel visibilities

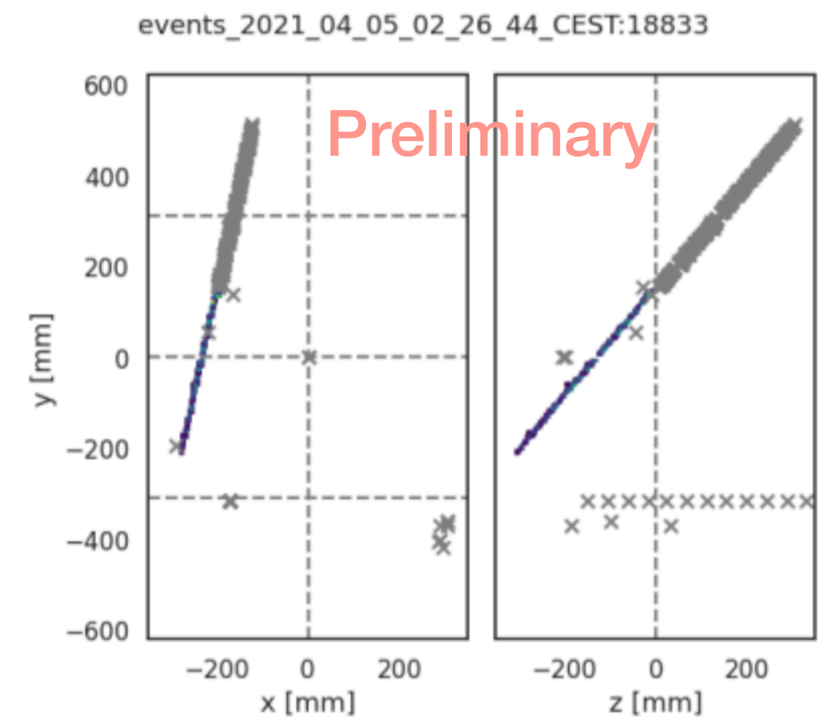


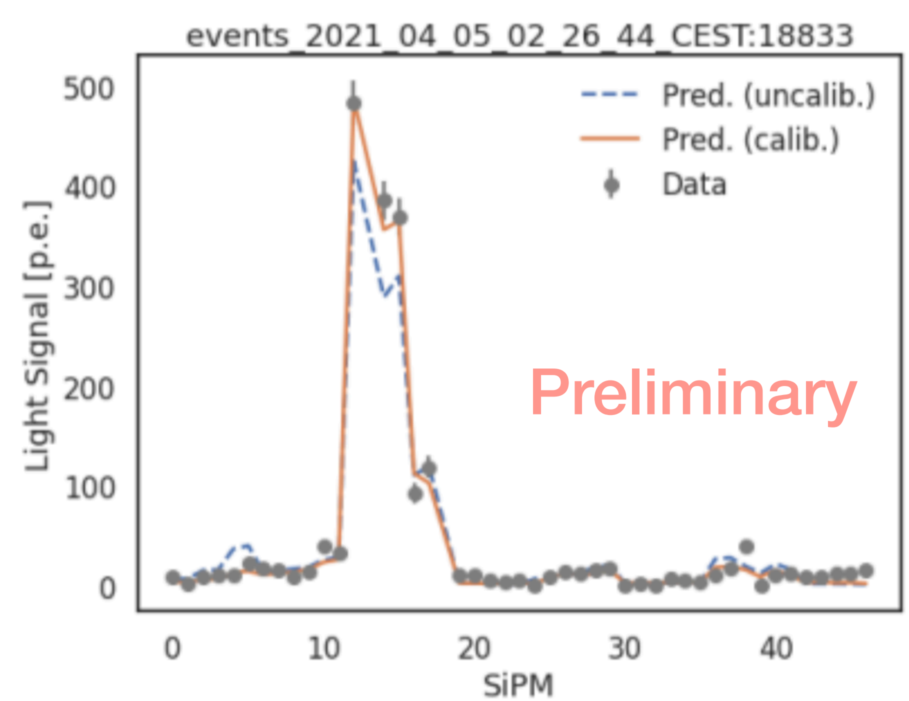
Optimise Light Modelling with SIREN

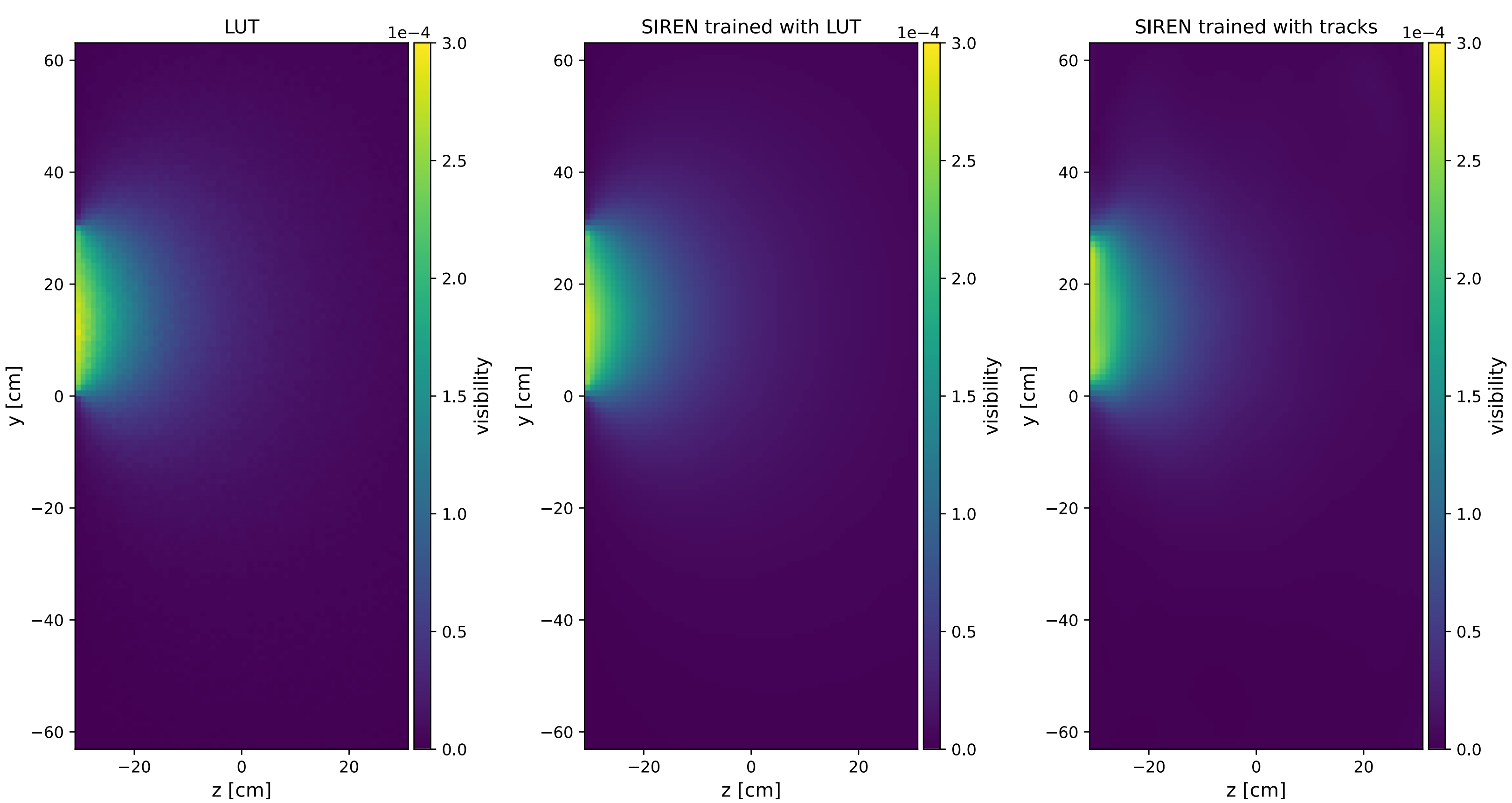
- Train SIREN directly with data (e.g cosmic muon tracks) using PoissonNLLLoss for the total photoelectrons in the light readout channels
- No need of the initial Geant4 simulation
- An in-situ photon library, not susceptible to potential mis-modeling in the light simulation (Rayleigh scattering, efficiency of the wavelength shifter etc.)
- Inclusive improvement of the modeling
- Can be trained with a small track sample

Demonstration with a different detector setup

[arXiv: 2211.01505](https://arxiv.org/abs/2211.01505)



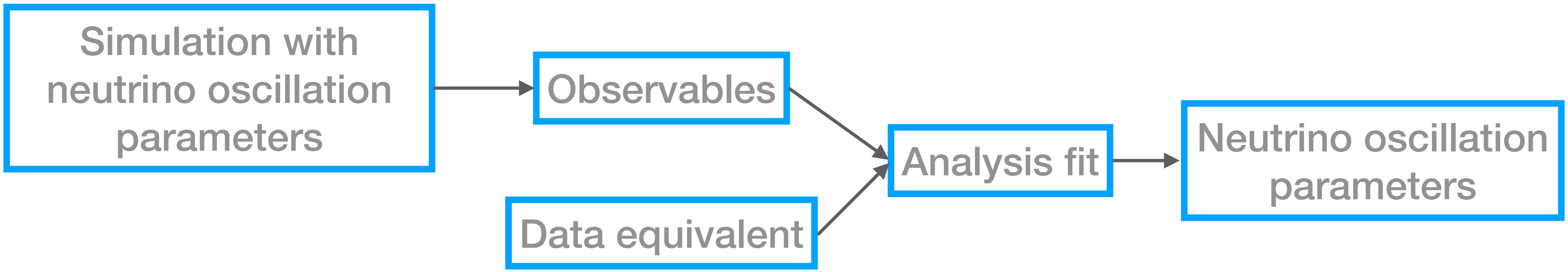


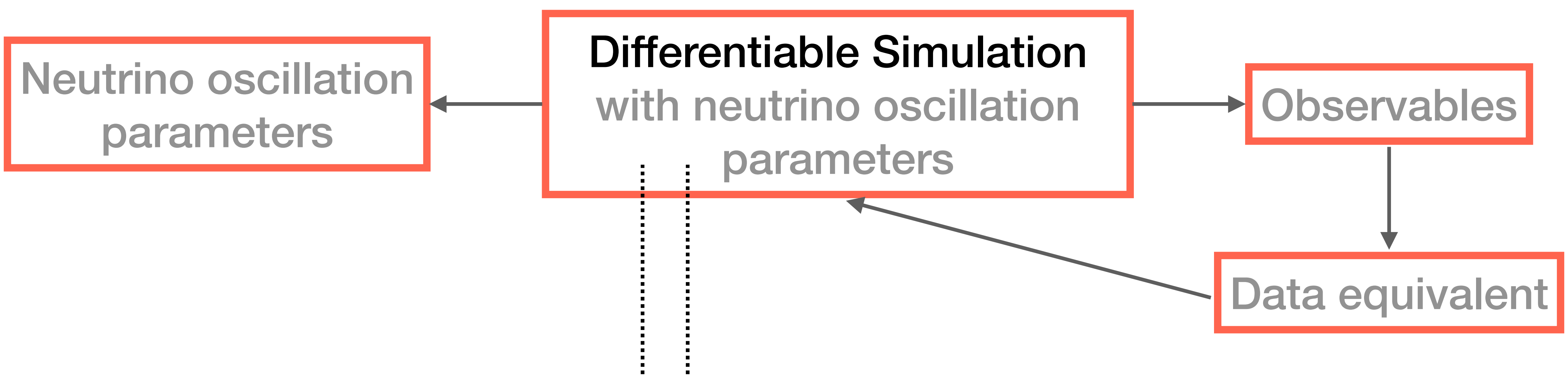


Summary and Next Steps

- Demonstration of simultaneous high-dimensional calibration with a differentiable simulation for both the charge and light signals in LArTPCs to enhance the simulation fidelity.
- Charge with differentiable larnd-sim:
  - A successful closure test for a simultaneous calibration fit of 6 detector parameters.
  - Development towards data application.
  - Work in progress to only use the target reconstructed information for the simulation input.
  - Establish a criteria for fit convergence
  - Speed up
- Light with SIREN:
  - Incorporate the photon arrival time
  - Update loss to be time sensitive
  - Optimise SIREN together with other explicit light modeling parts
- Aim to combine the optimisation of the charge and light modeling in future

Why Differentiable Simulation?





Differentiable detector simulation
High dimensional calibration
Improve the fidelity of the simulation

The Example Simulator: *larnd-sim*

Reference *larnd-sim*

https://github.com/DUNE/larnd-sim

JINST 18 (2023) 04, P04034

Differentiable *larnd-sim*

https://github.com/ynashed/larnd-sim

Mach. Learn.: Sci. Technol. 5 025012

- Applicable for DUNE near detector LArTPC and its prototypes
- Using Numba (a JIT compiler) with CUDA kernels (highly parallelized)

Track segments ($\frac{dE}{dx}$)

Birks' model (A_B, k_B), Electric field (ε) → **Quenching**

$$\alpha_{\text{recomb}} = \frac{A_B}{1 + \frac{k_B}{\varepsilon \cdot \rho} \frac{dE}{dx}}$$

$$n_{\text{elec}} = \frac{dE}{W_{\text{ion}}}$$

$$n_{\text{elec}}^{\text{recomb}} = \alpha_{\text{recomb}} \cdot n_{\text{elec}}$$

Number of free electrons for drifting

Drift velocity (v_{drift}), Lifetime (τ), Transverse/ longitudinal diffusion (D_T, D_L) → **Drifting**

$$v_{\text{drift}} = \mu \cdot \varepsilon$$

$$\sigma_{L,T} = \sqrt{2 \cdot t_{\text{drift}} \cdot D_{L,T}}$$

$$n_{\text{elec}}^{\text{final}} = n_{\text{elec}}^{\text{recomb}} \cdot e^{-t_{\text{drift}}/\tau}$$

Electron distribution at readout

Current

Current induced on each pixel

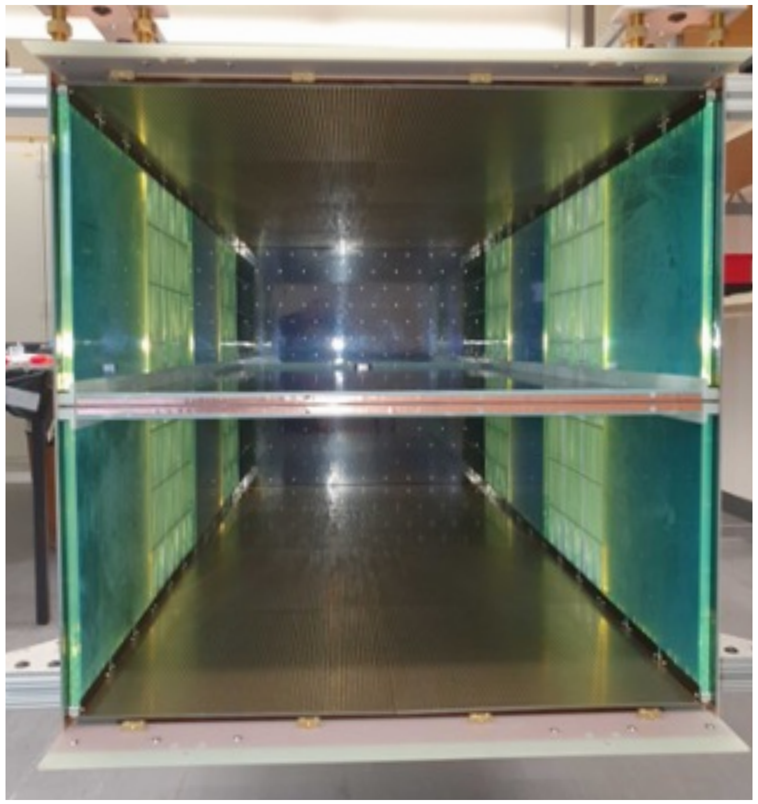
Electronics

Signal digitized here

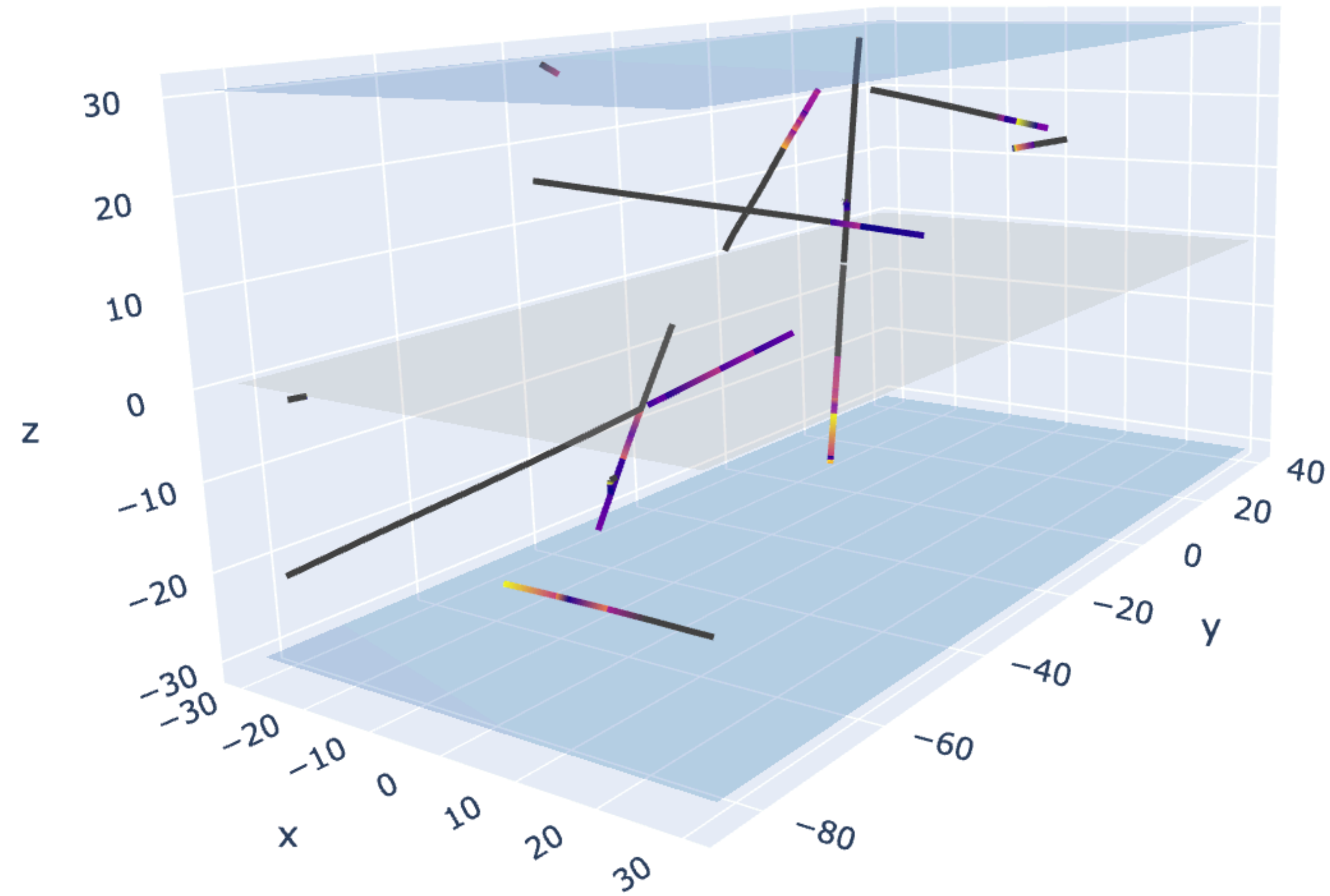
ADC counts per pixel/time

Samples and Selection

DUNE



Default sample   Particle segments



Default sample:

100 events of ~10 muons with 1 GeV kinetic energy (K.E.)

Alternative sample:

100 events of ~10 mixed particles (muons, charged pions, protons) with [0.1, 2] GeV K.E.

Selection:

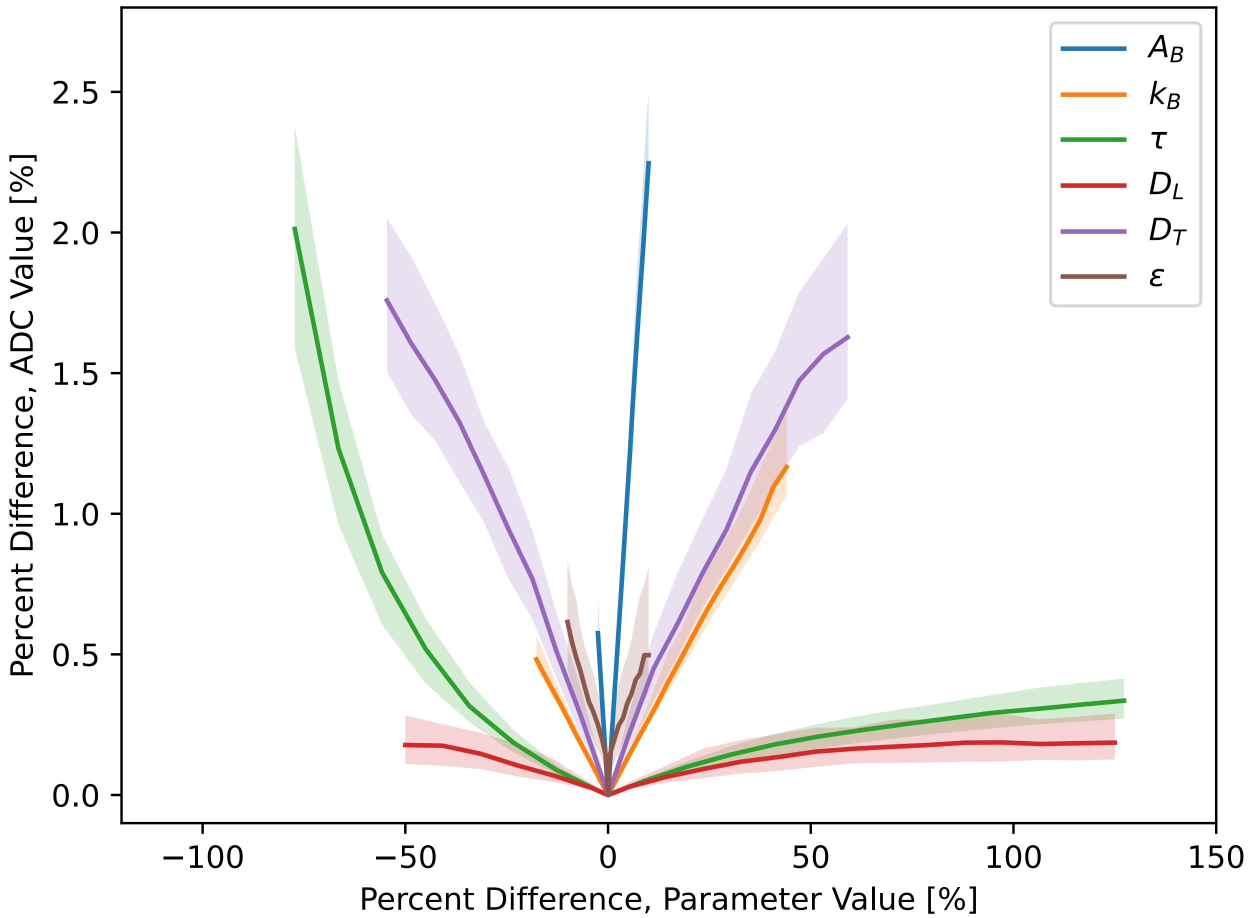
- Track length > 2 cm
- abs(track segment z) in [15, 28] cm (near the anodes)
- Track angle wrt. z larger than 15°

Challenging 1: Electronic Noises

The electronic noises are an example of stochasticity in the simulation.
It increases batch-to-batch variations.
Let noise be in data.
Disable the noise in the forward simulation to reduce stochasticity.

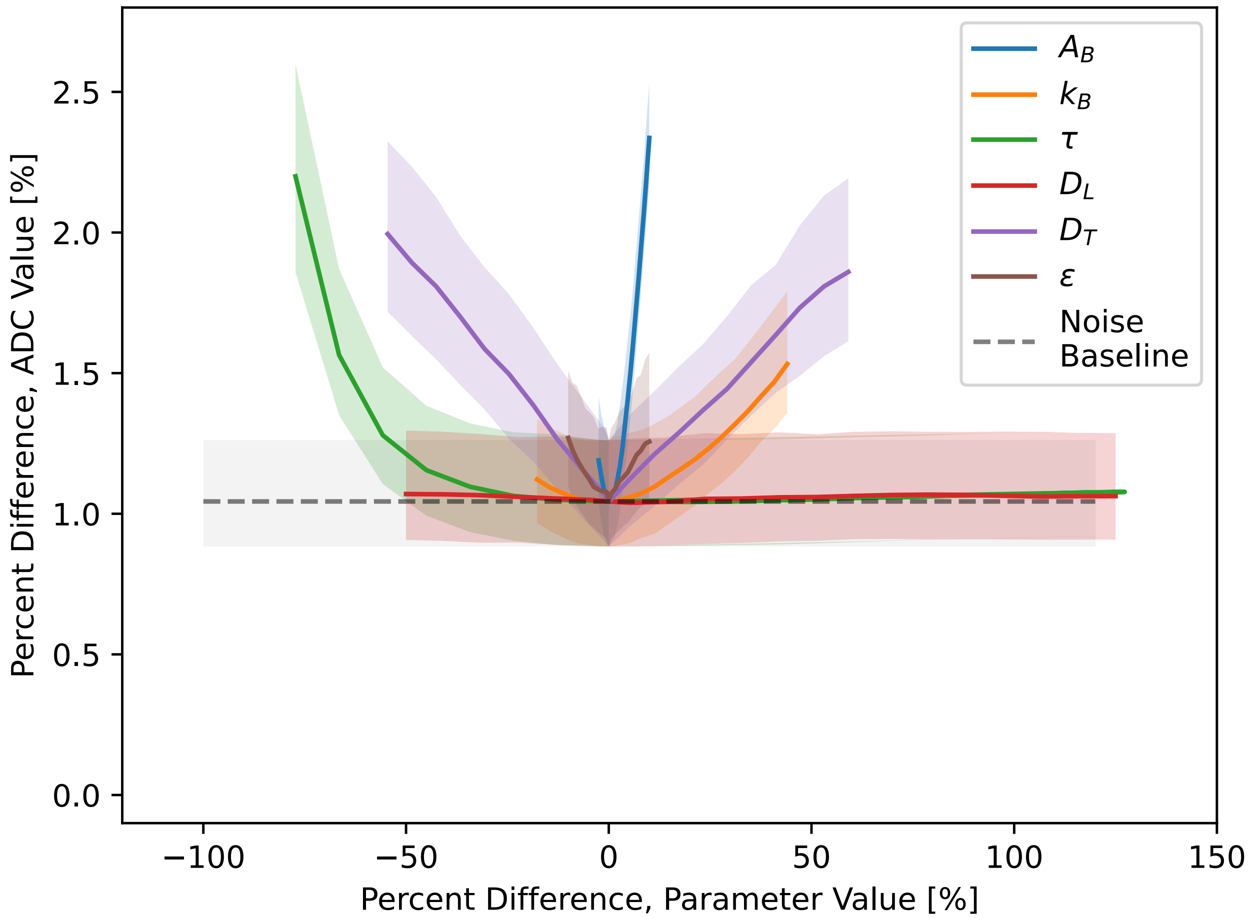
Target without noise

mostly MIP like muons

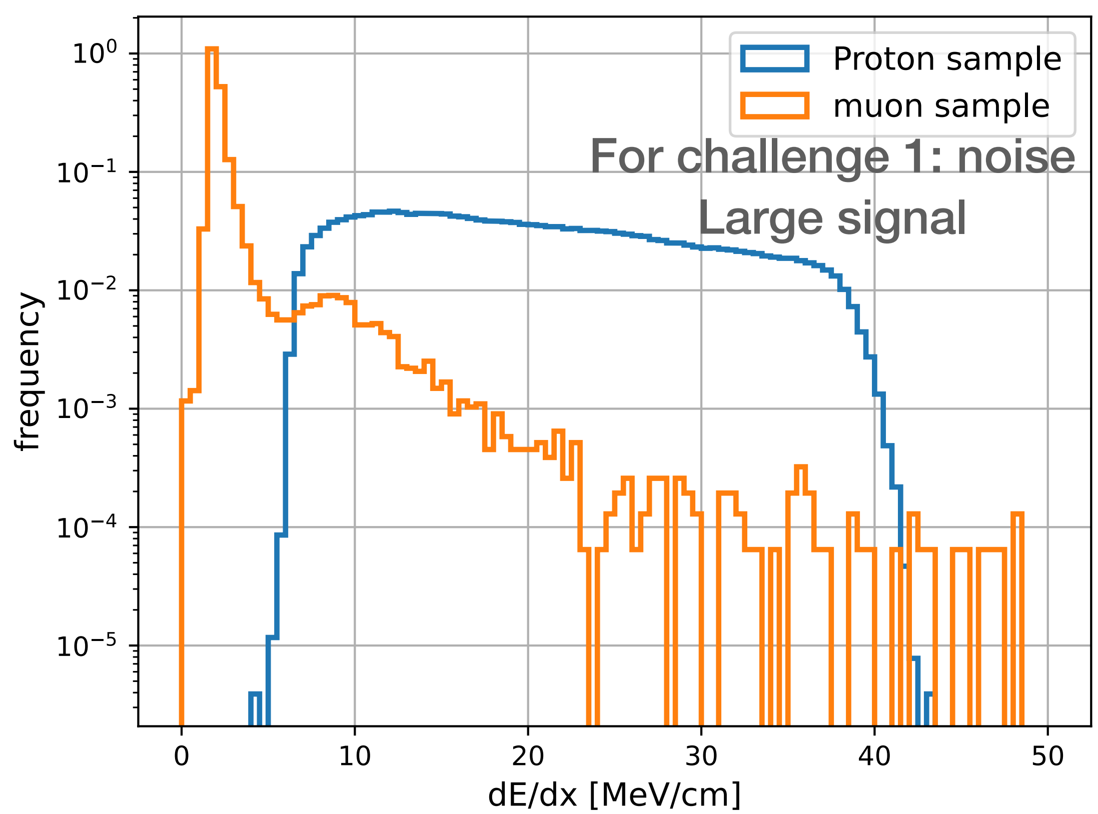


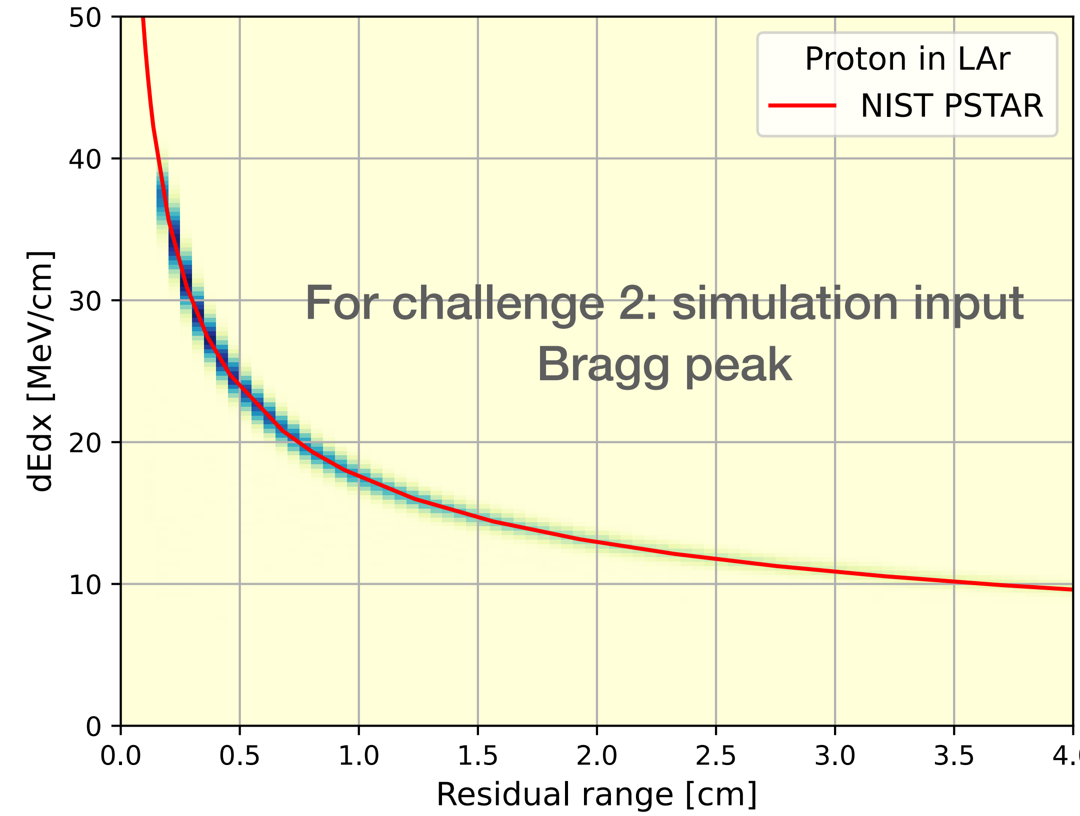
Target with noise

Mostly MIP like muons

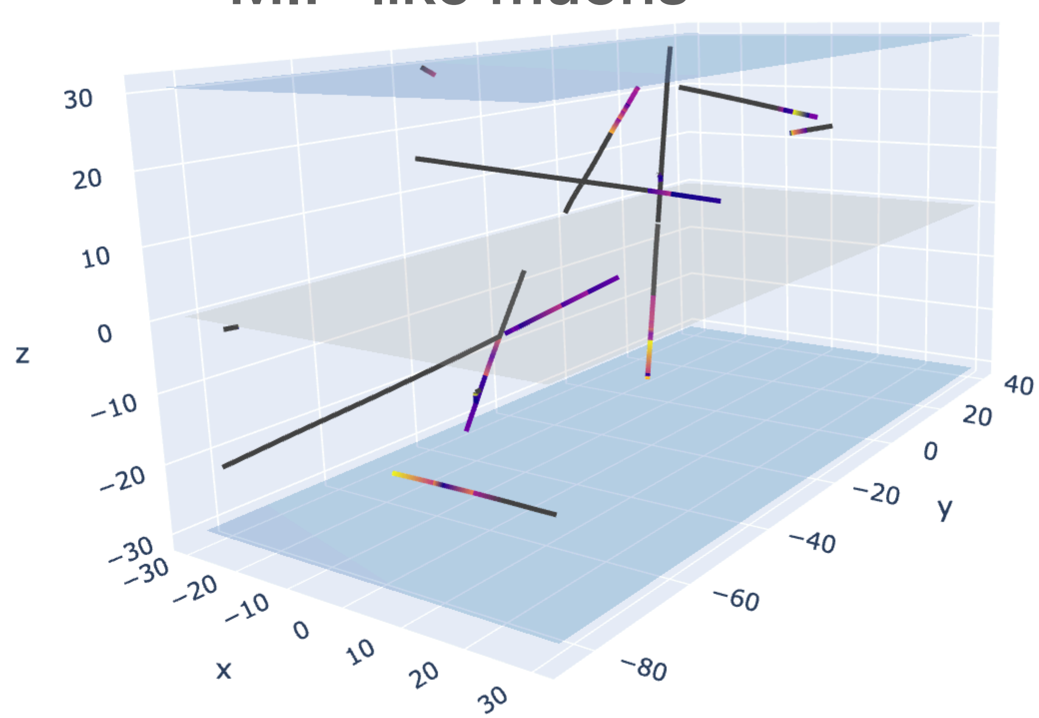


A Sample of Stopping Protons

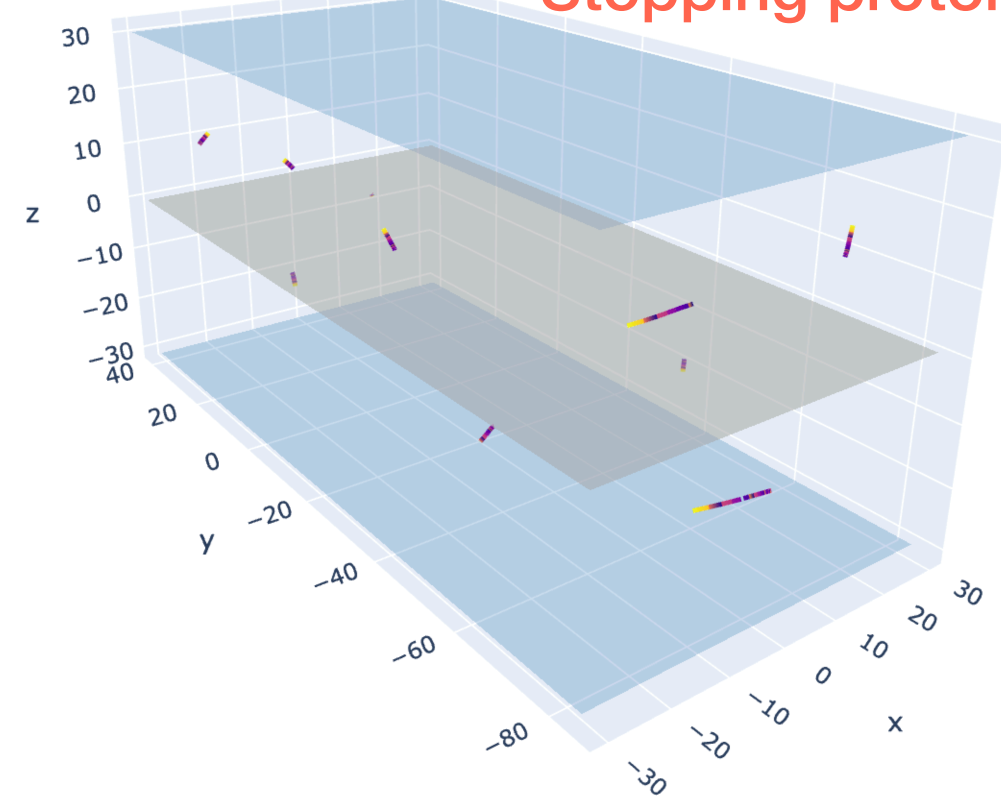




MIP-like muons

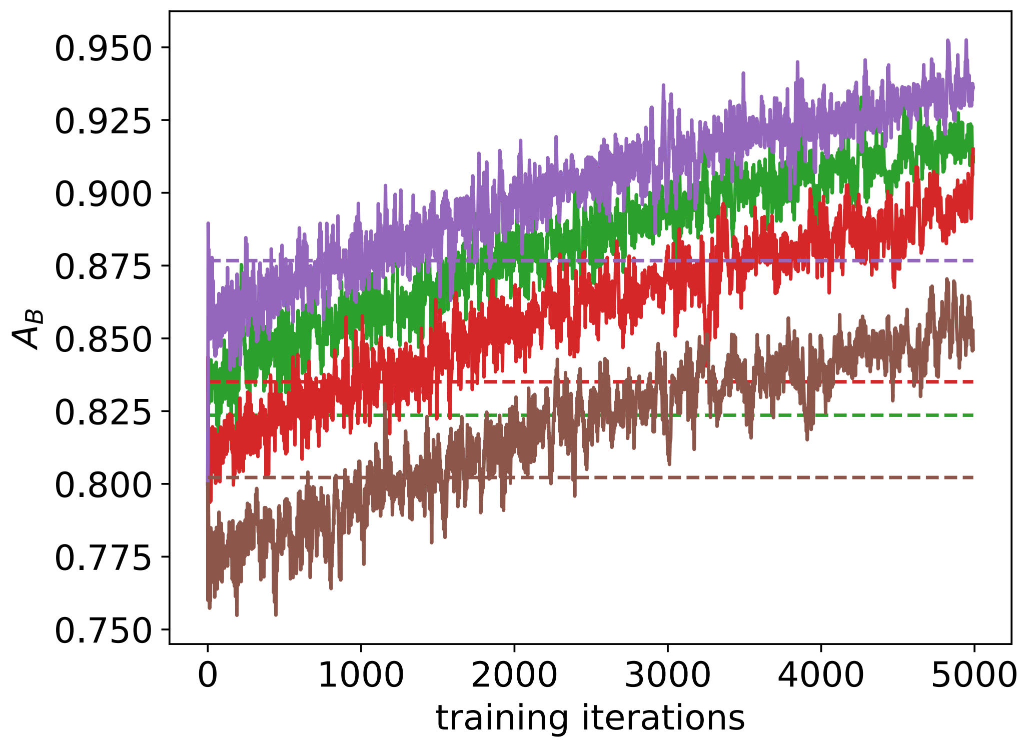


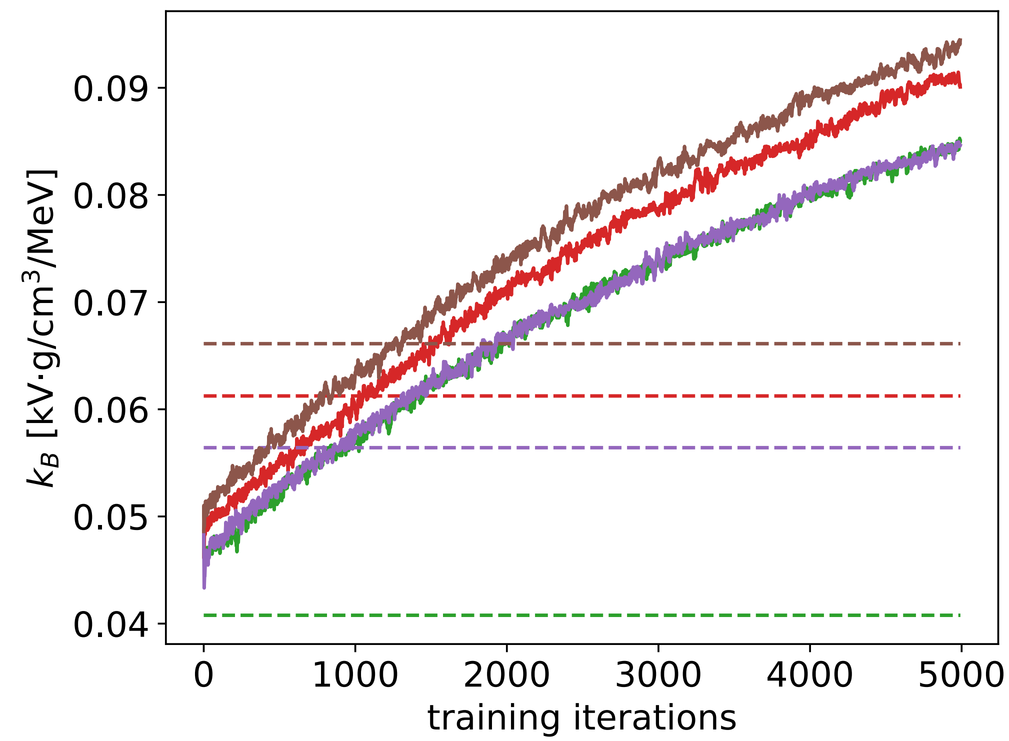
Stopping protons



Simulation Input: MIP Assumption?

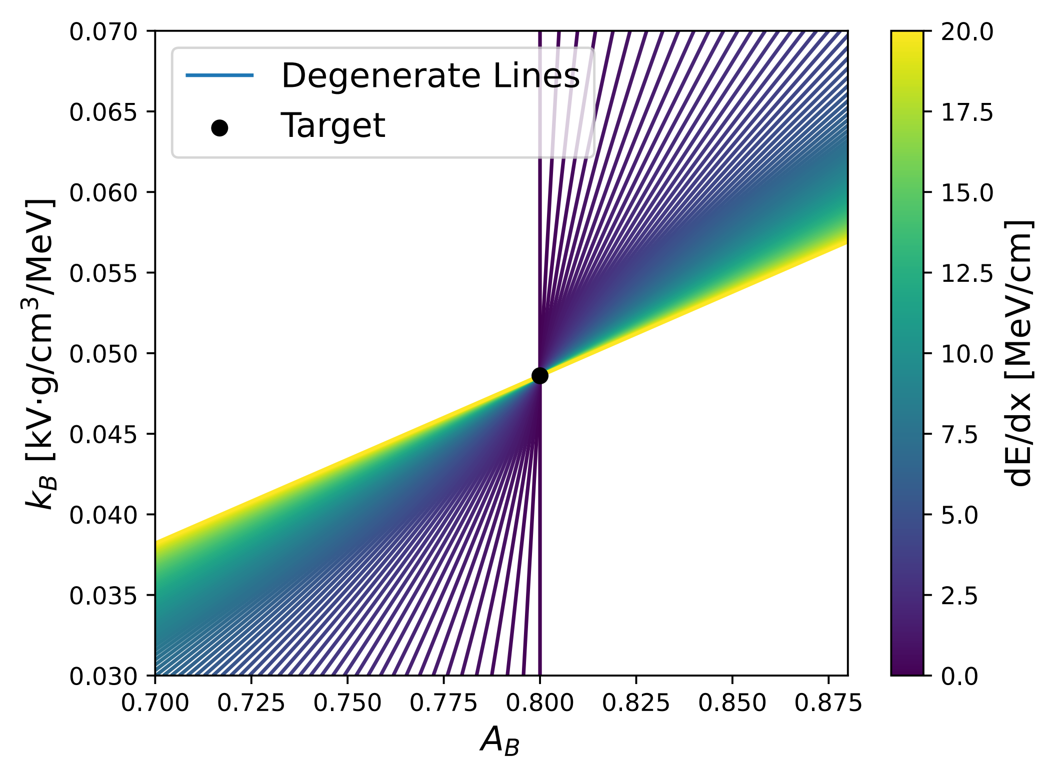
Set dE/dx to 2 MeV/cm (mostly MIP muon) in the simulation input.
Other parameters have converged to the corresponding targets.





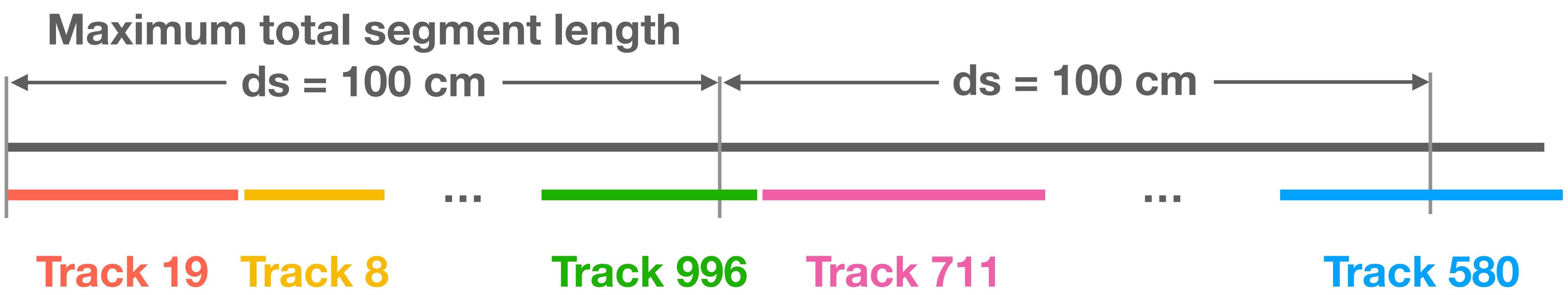
Electron recombination
Birks model

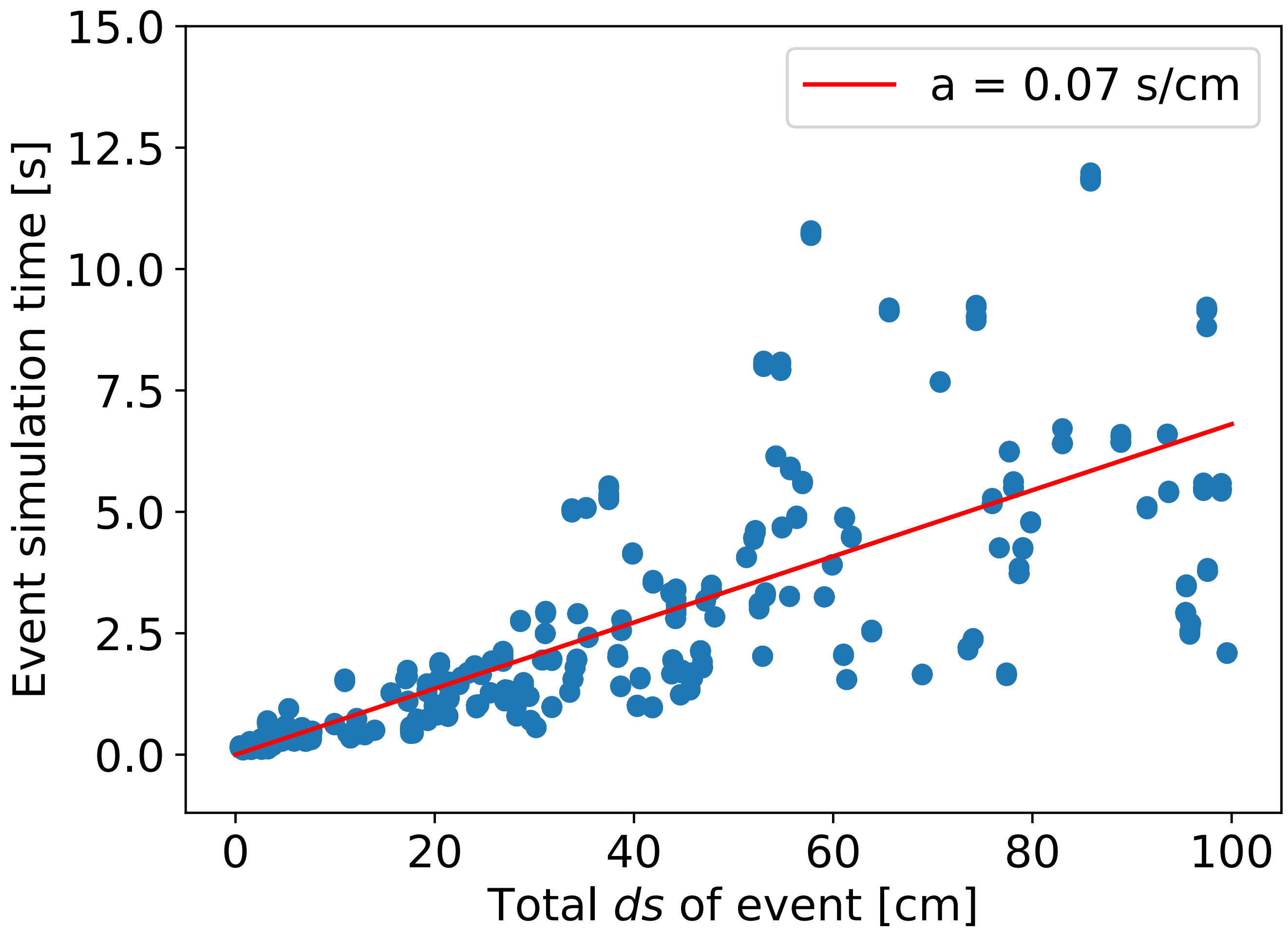
$$A_B = \alpha^*_{\text{recomb}} \cdot \left(1 + \frac{k_B}{\mathcal{E} \cdot \rho} \frac{dE}{dx}\right)$$



Mini-batching

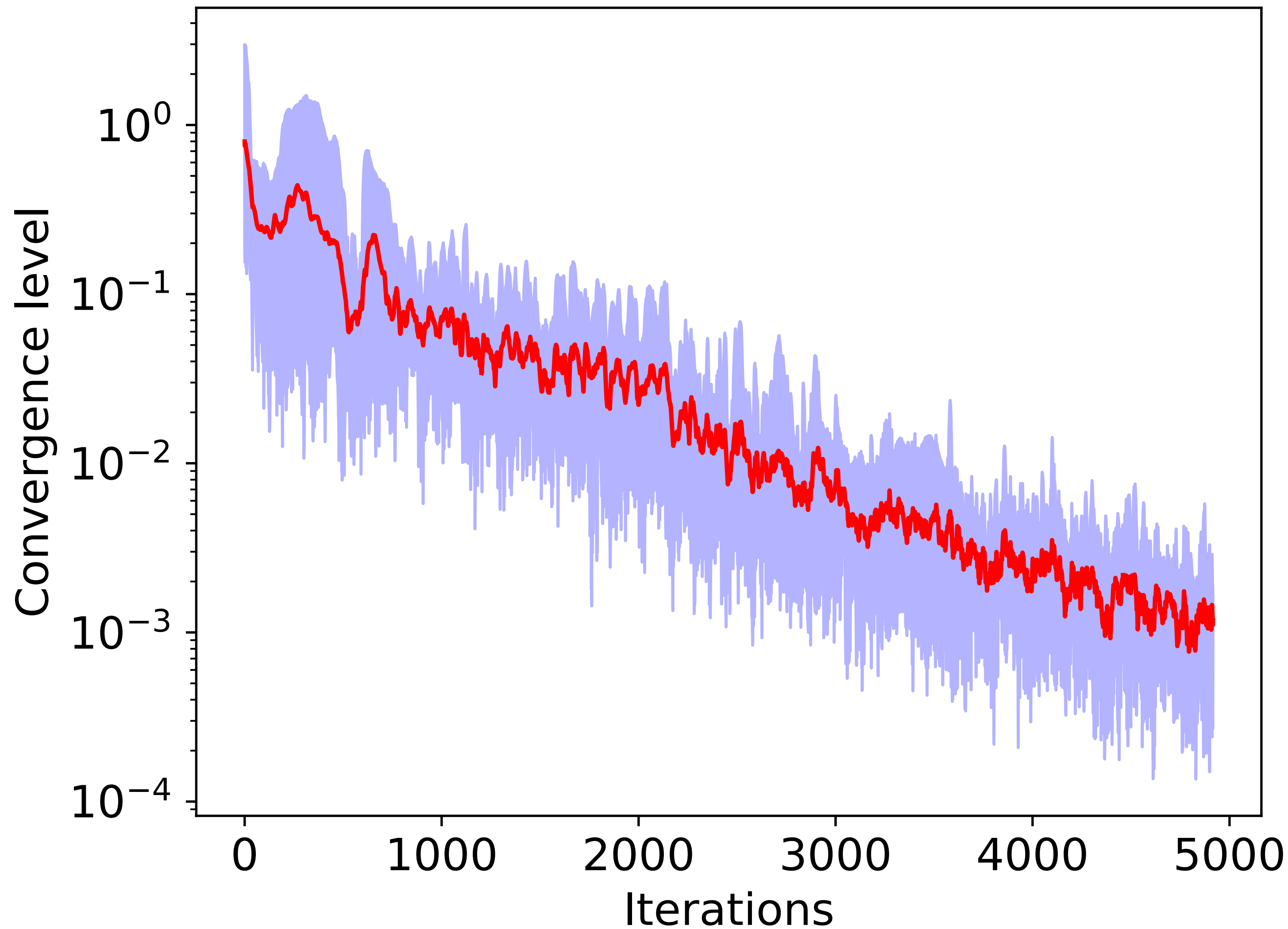
Each fitting iteration runs on a mini-batch



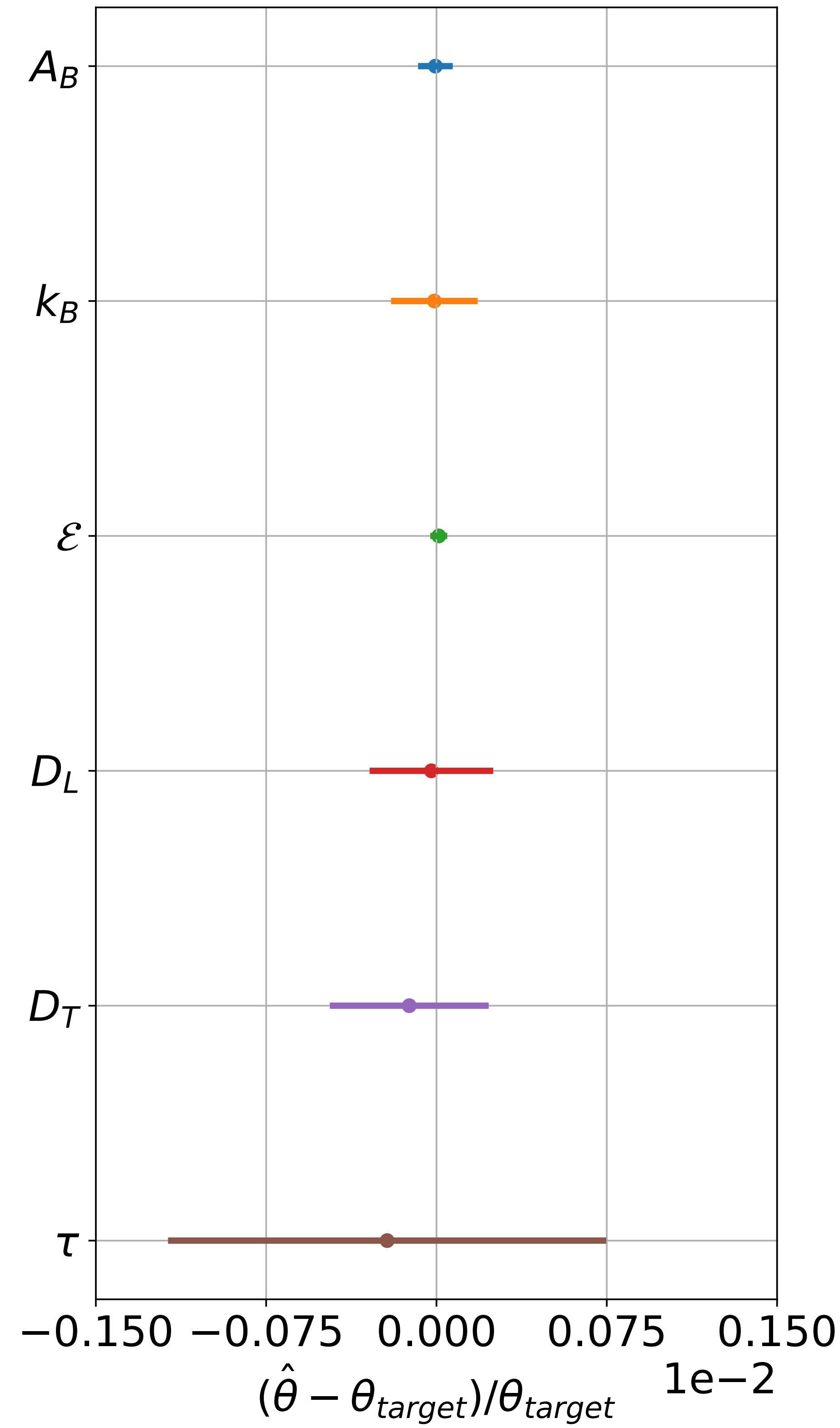


Convergence level

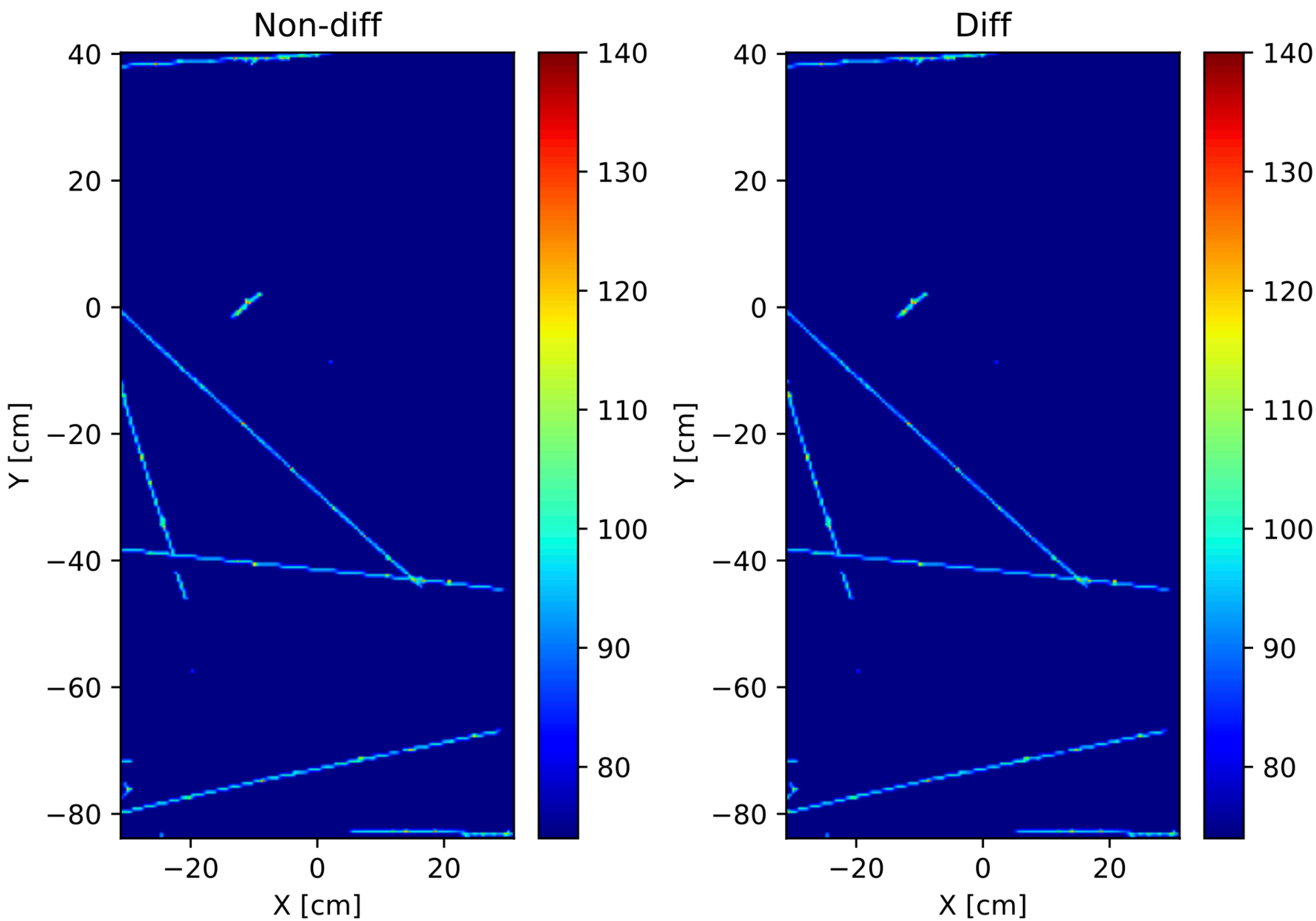
Convergence evolution



Final convergence



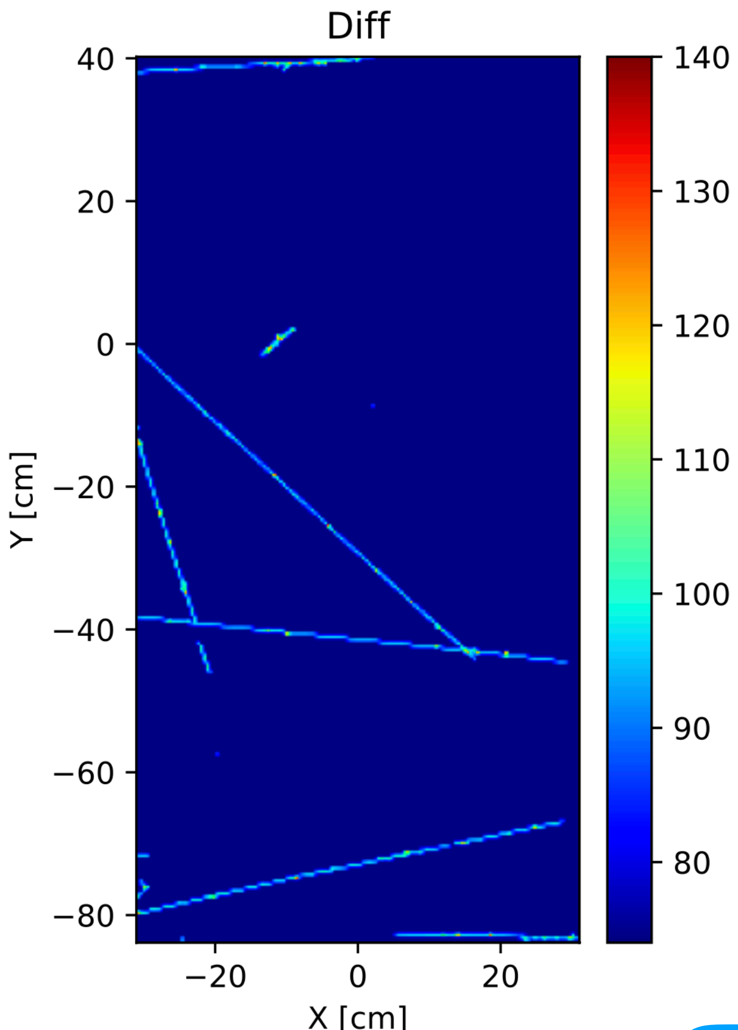
The Differentiable and the Reference *larnd-sim*

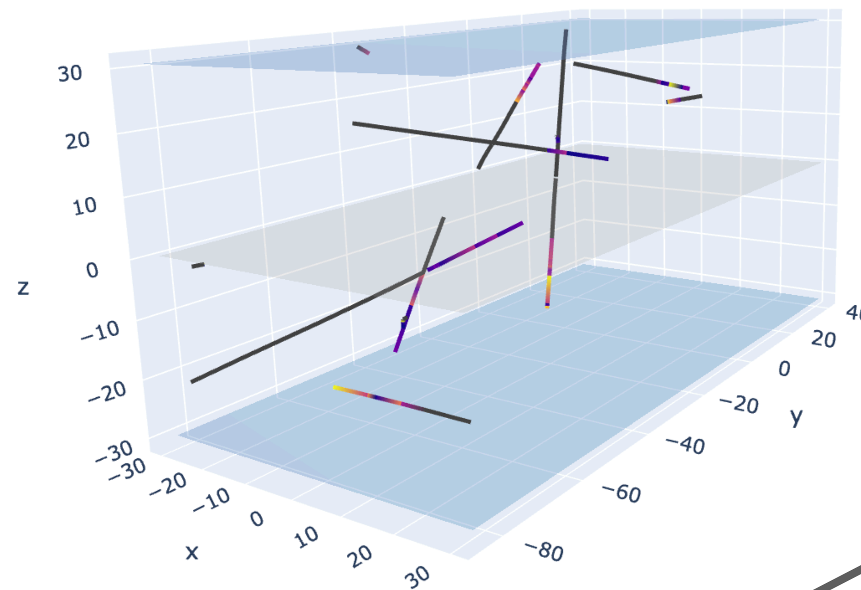


Average deviation: 0.04 ADC counts per active pixel
Far below the typical noise level of a few ADC counts

Future Application: Inverse Solver

Map from detector readout to physics quantity





$$\chi$$

$$\nabla_\chi \mathcal{L}$$

$$F_{\text{target}}$$

$$f(\chi, \theta)$$

$$\mathcal{L}(f(\chi, \theta), F_{\text{target}})$$

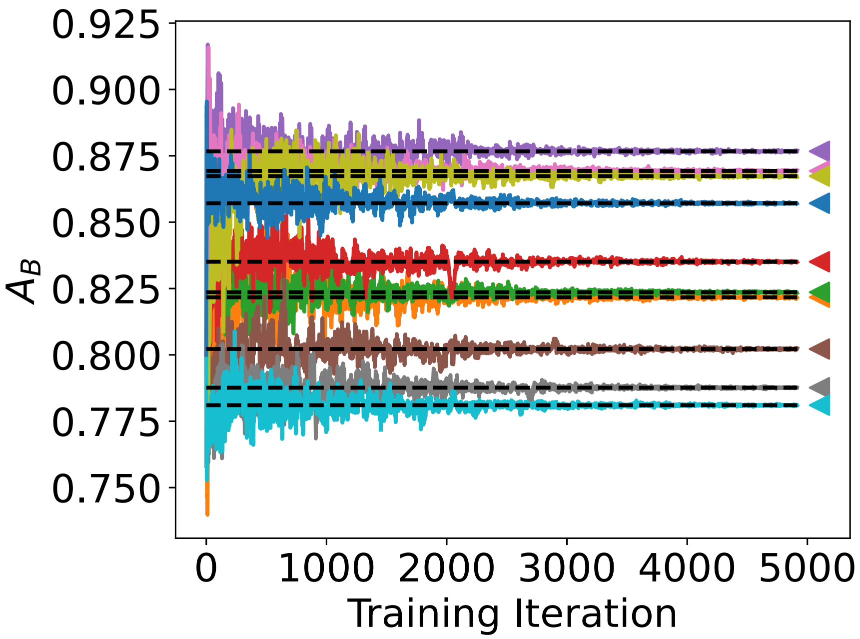
$$\nabla_\chi \mathcal{L}$$

Simultaneous Fit Result

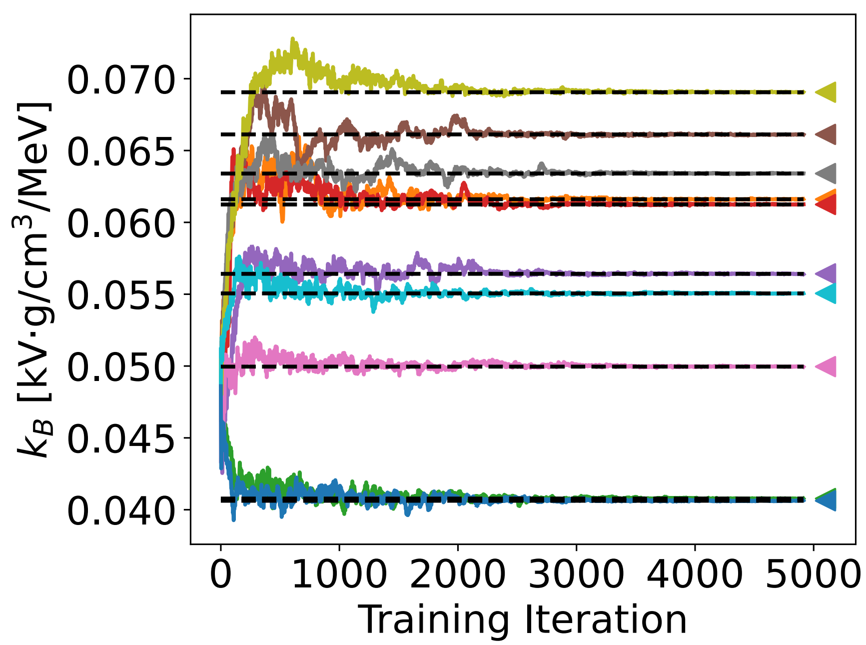
Mach. Learn.: Sci. Technol. 5 025012

10 fits with different targets in 6D phase space
All 6 parameters of interest converge to the target values

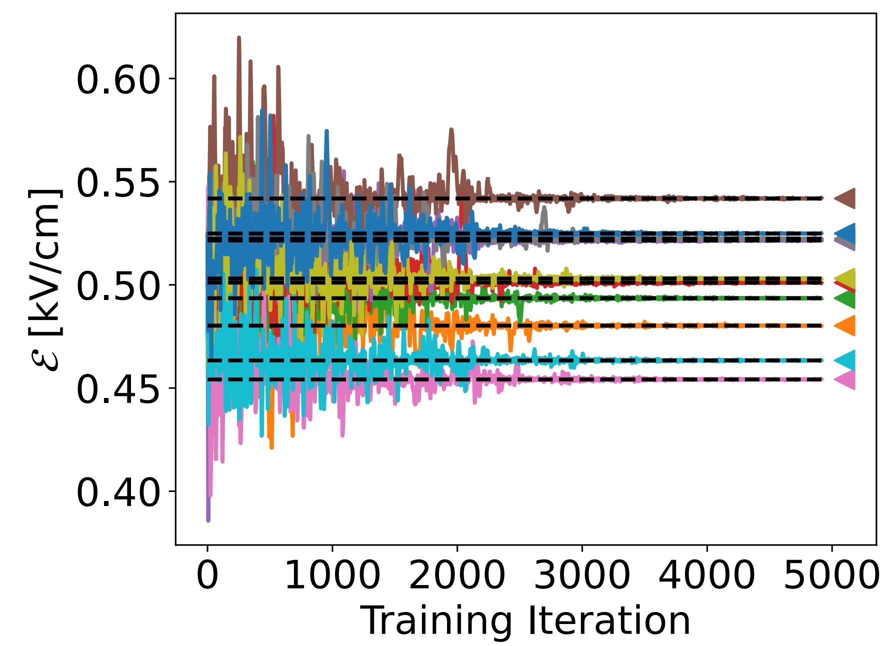
Recombination model A_B



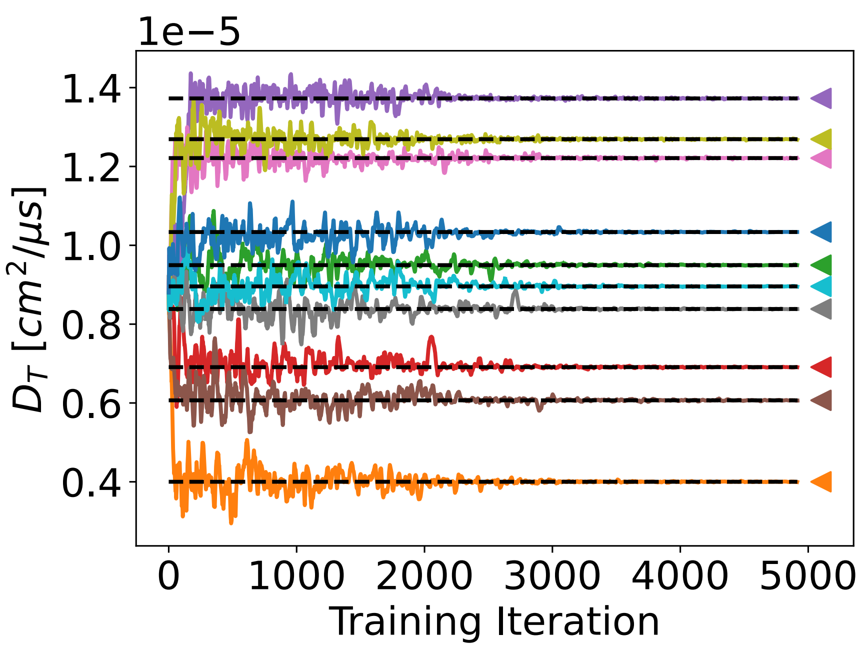
Recombination model k_B



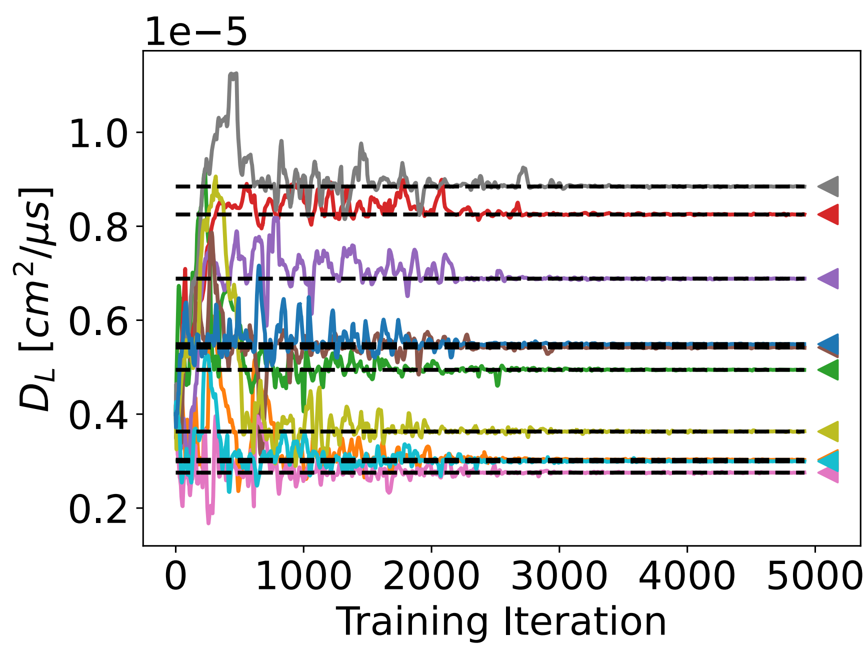
Electric field



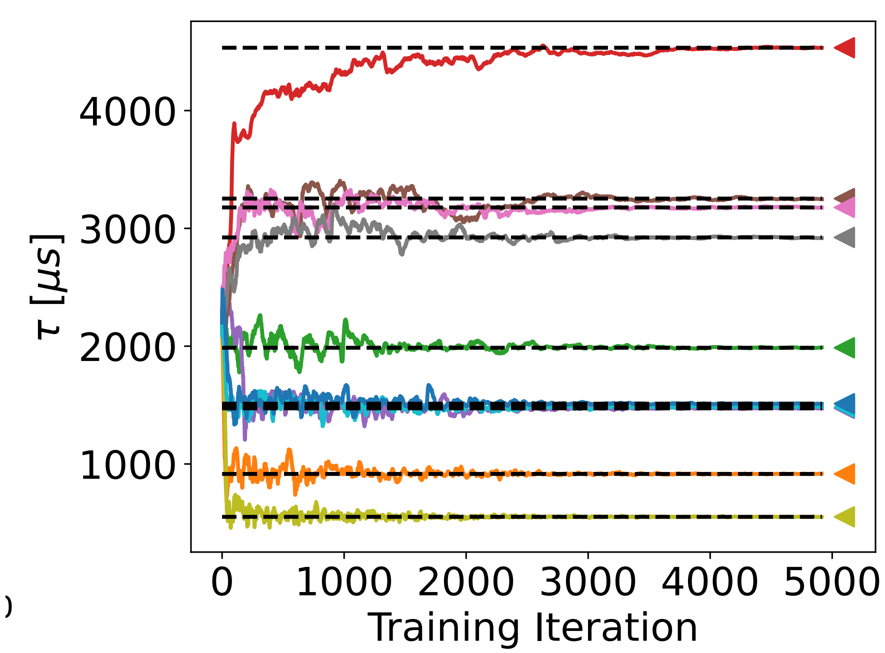
Transverse diffusion coefficient



Longitudinal diffusion coefficient



Electron lifetime



Result with the Alternative Sample

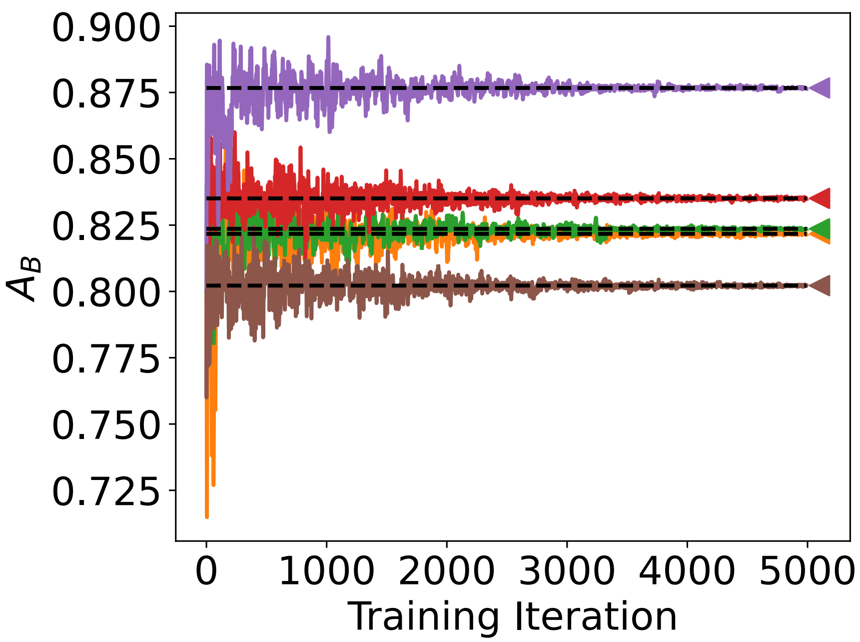
5 different targets in 6D phase space

All 6 parameters of interest converge to the target values

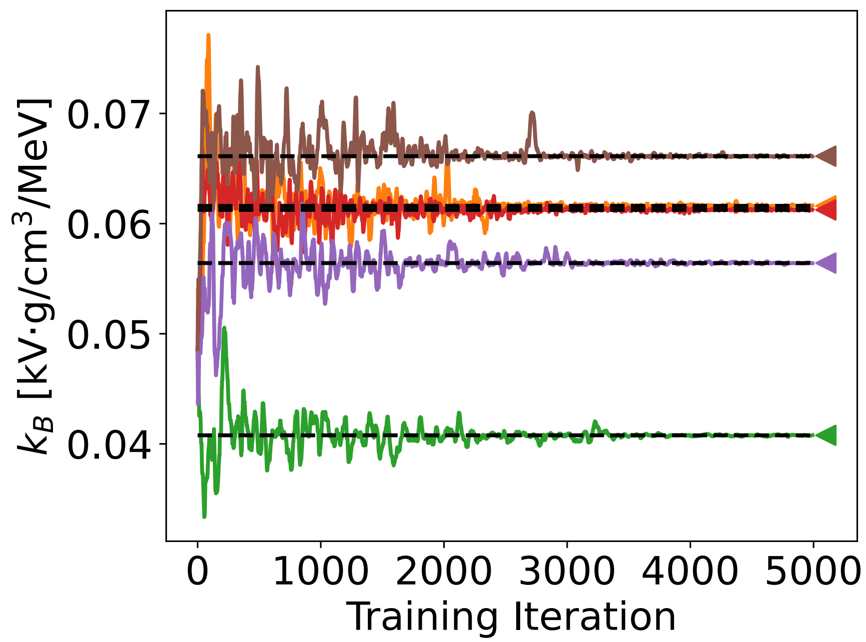
Relaxed requirement on Inputs

(Muons, charged pions and protons of [0.1, 2] GeV kinetic energy)

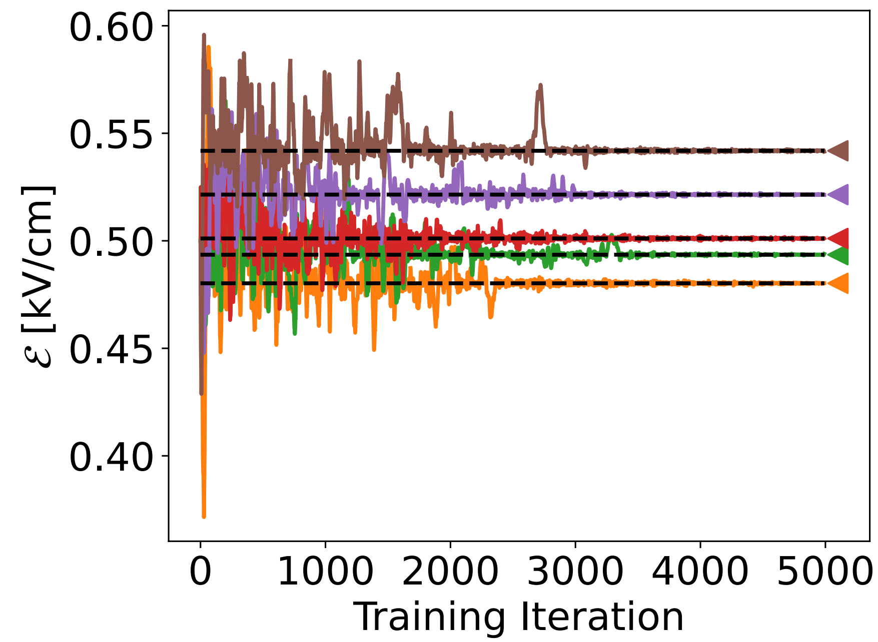
Recombination model A_B



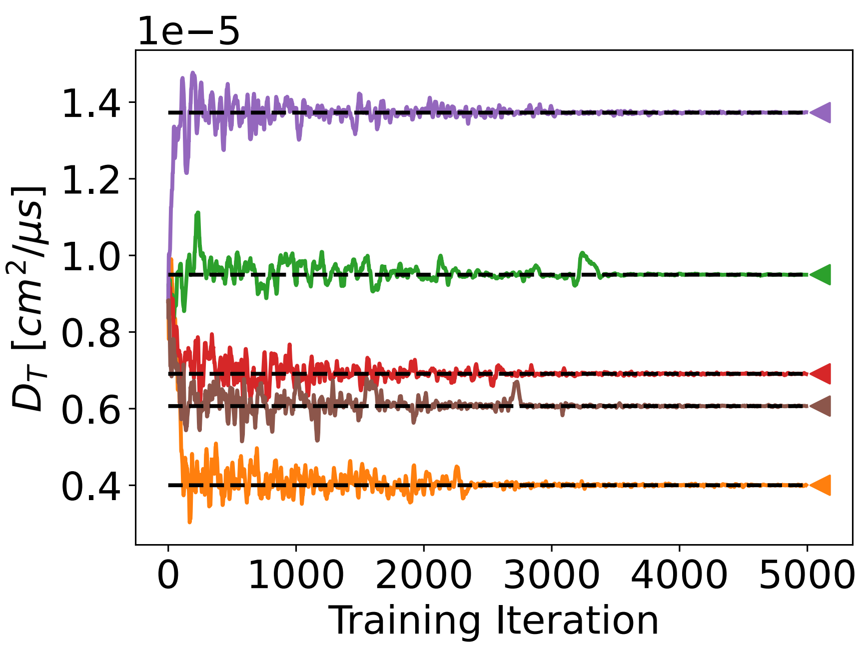
Recombination model k_B



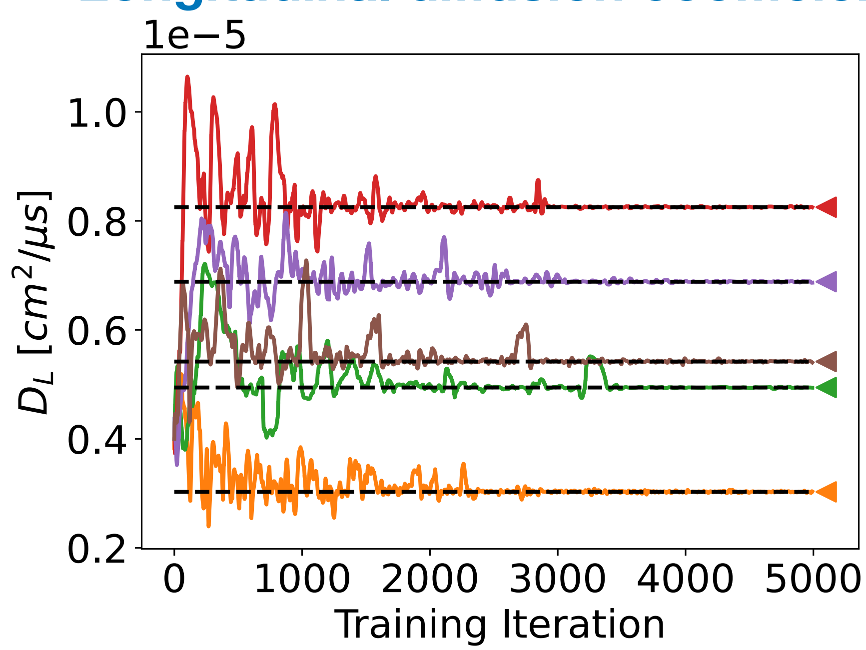
Electric field



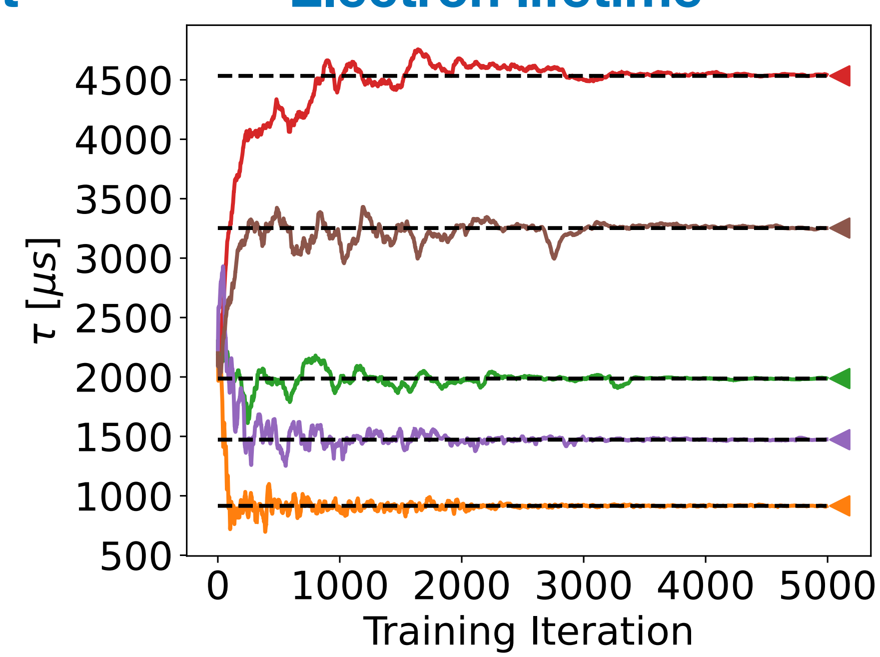
Transverse diffusion coefficient



Longitudinal diffusion coefficient

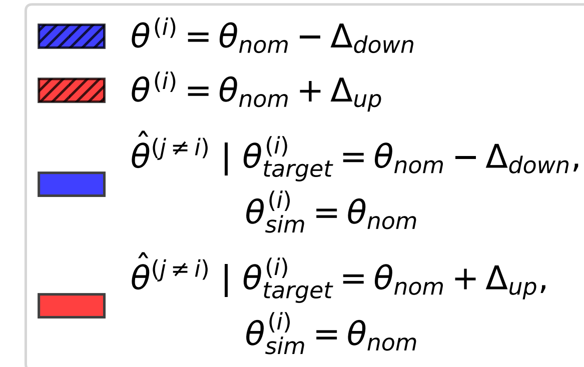


Electron lifetime



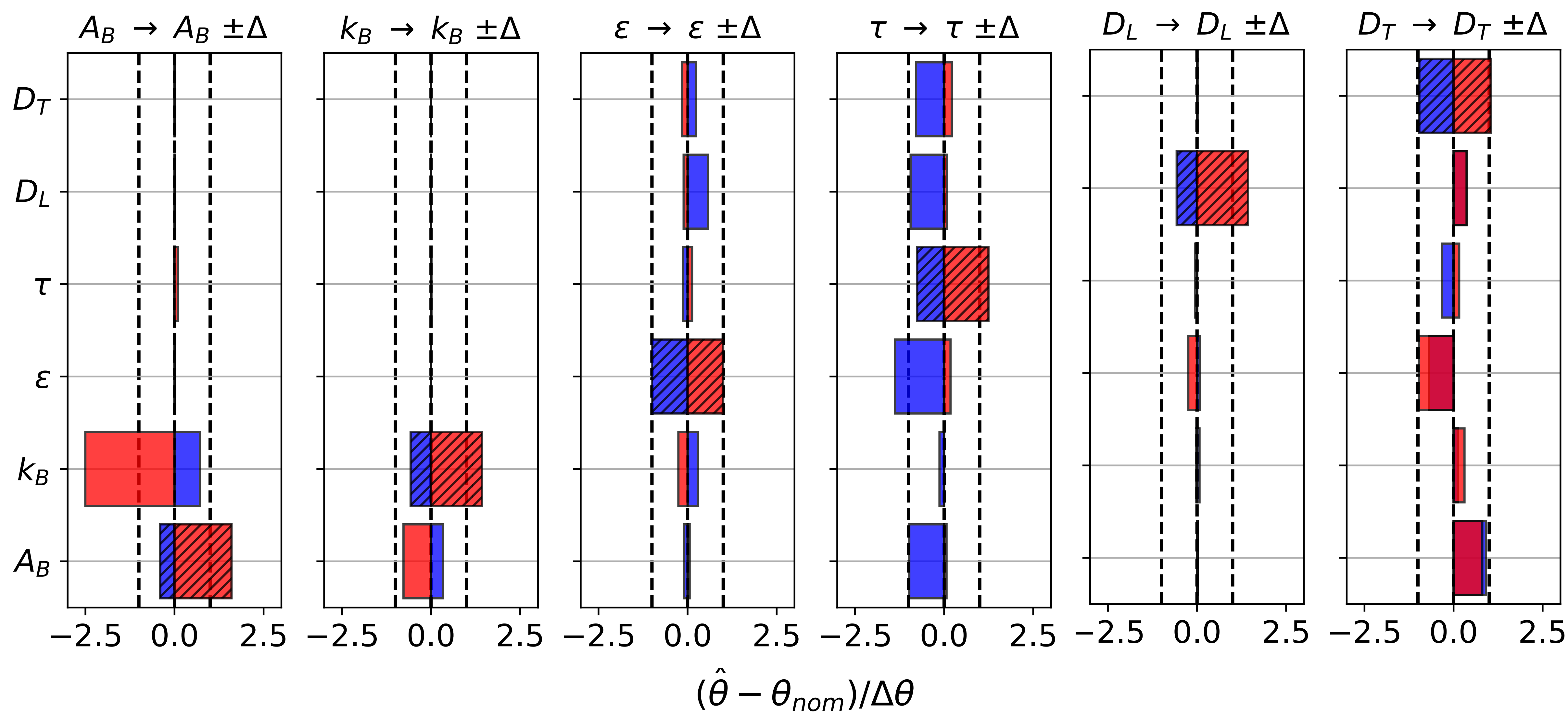
No Fully Independent Parameters

Parameter [Units]	Nominal Value	Range
A_B	0.8	[0.78, 0.88]
k_B $[kV.g/cm^3/MeV]$	0.0486	[0.04, 0.07]
$\mathcal{E}$ $[kV/cm]$	0.5	[0.45, 0.55]
τ $[\mu s]$	2200	[500, 5000]
D_L $[cm^2/\mu s]$	4×10^{-6}	$[2 \times 10^{-6}, 9 \times 10^{-6}]$
D_T $[cm^2/\mu s]$	8.8×10^{-6}	$[4 \times 10^{-6}, 14 \times 10^{-6}]$



$$\Delta\theta = \frac{\Delta_{up} + \Delta_{down}}{2}$$

$\hat{\theta}$ = Fitted Value

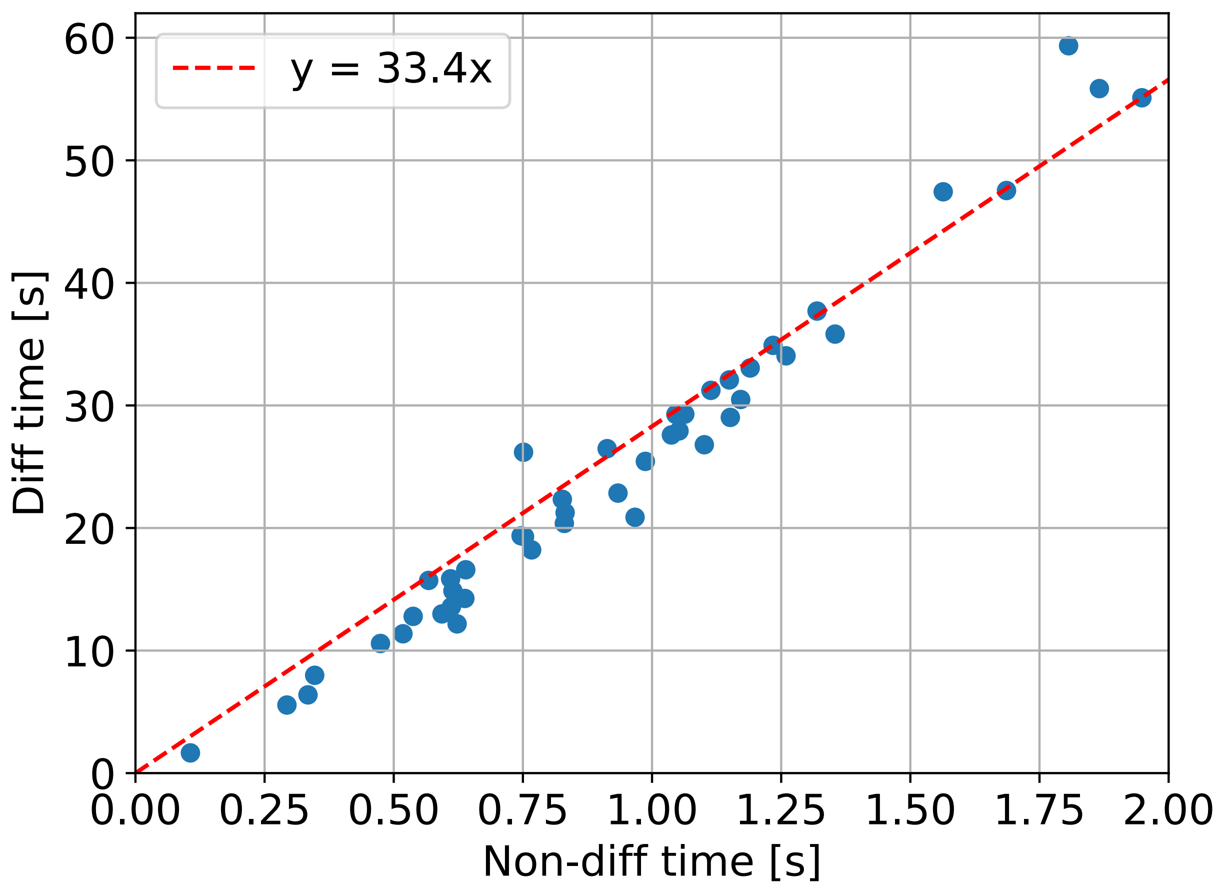


Computation Time

Modular nature of the detector — highly parallelizable

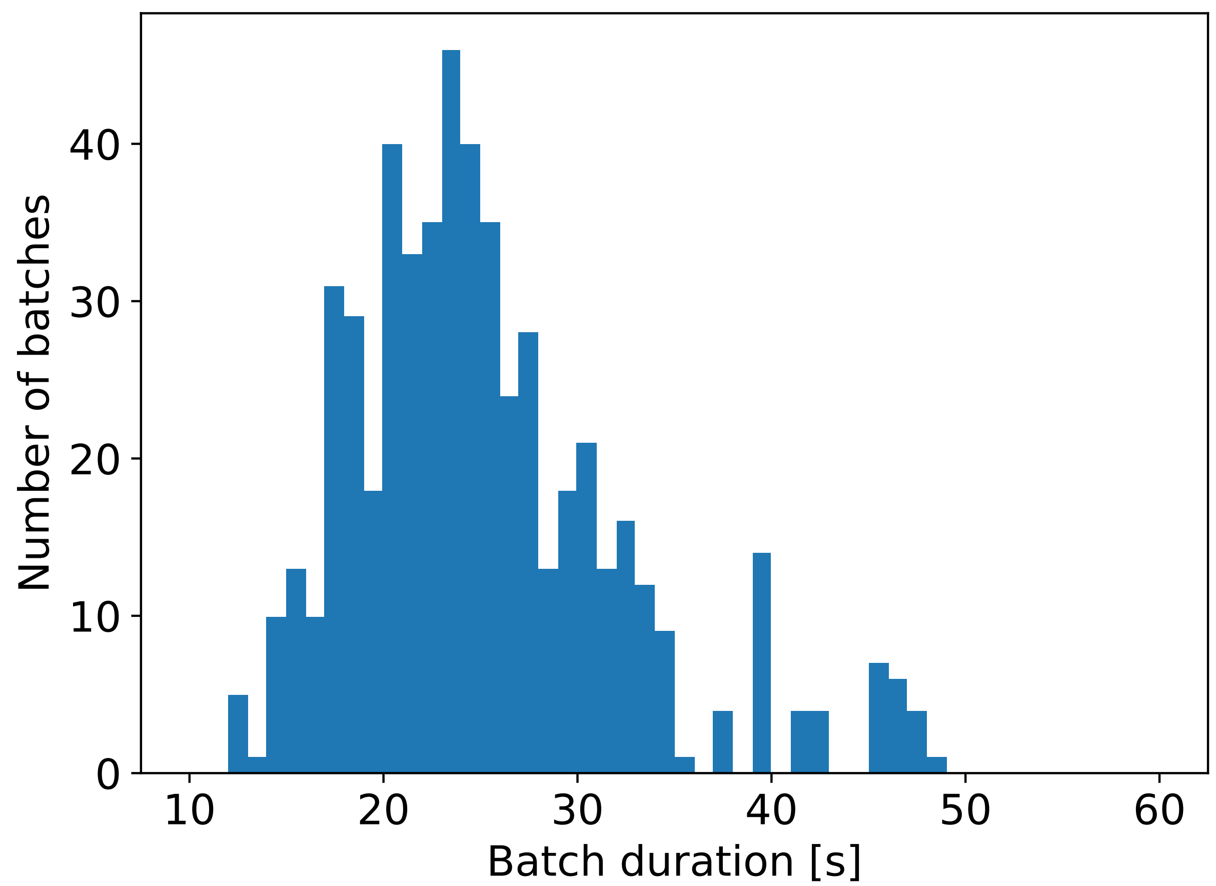
Working on translate the diff-sim using JAX to speed up the simulation

Simulation time per event



Forward simulation

Time per min-batch iteration



Forward simulation + gradient calculation

Memory Usage

Peak memory usage in track current calculation

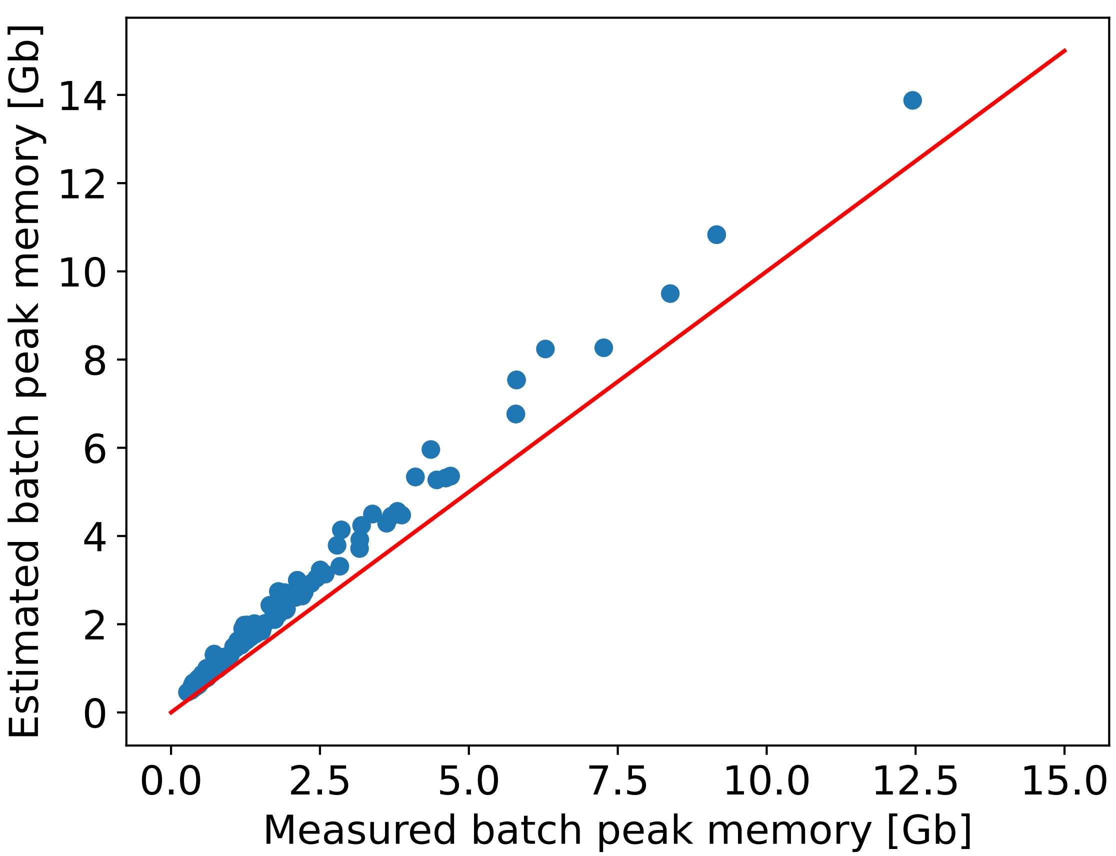
loop   loop

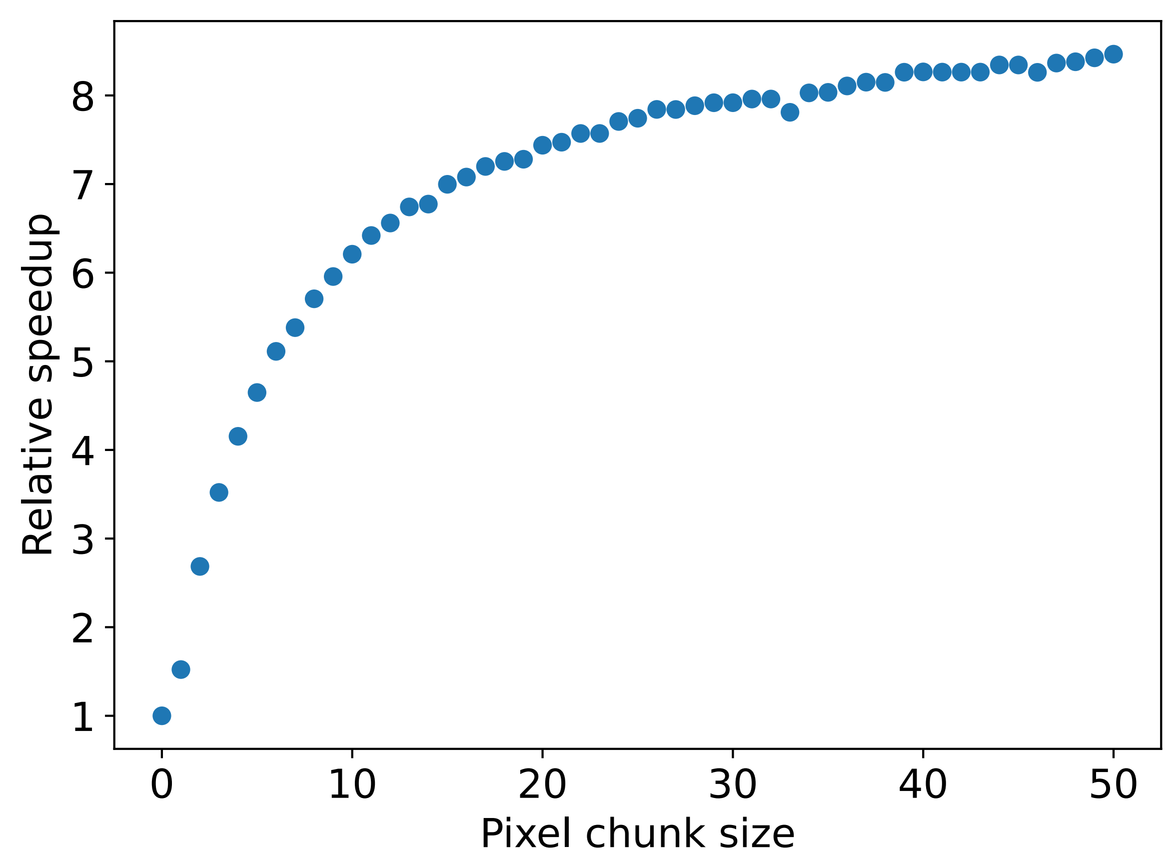
$$\mathcal{M} = N_{\text{segments}} \times N_{T_f} \times N_{T_0} \times N_x \times N_y \times N_{\text{pixels}}$$

chunk size 1   adjustable chunk size

Trade-off between memory and computation time

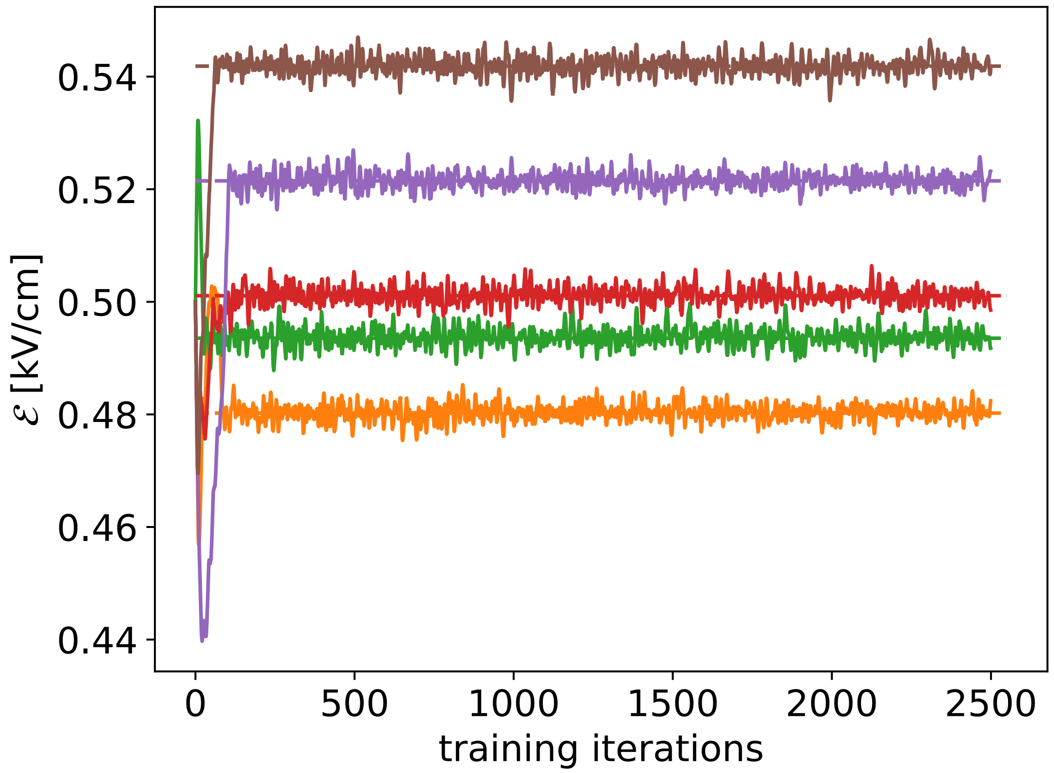
- Shrink the tensor dimension by loops
- “Chunking” operation to reduce loop iterations
- Gradient checkpoint



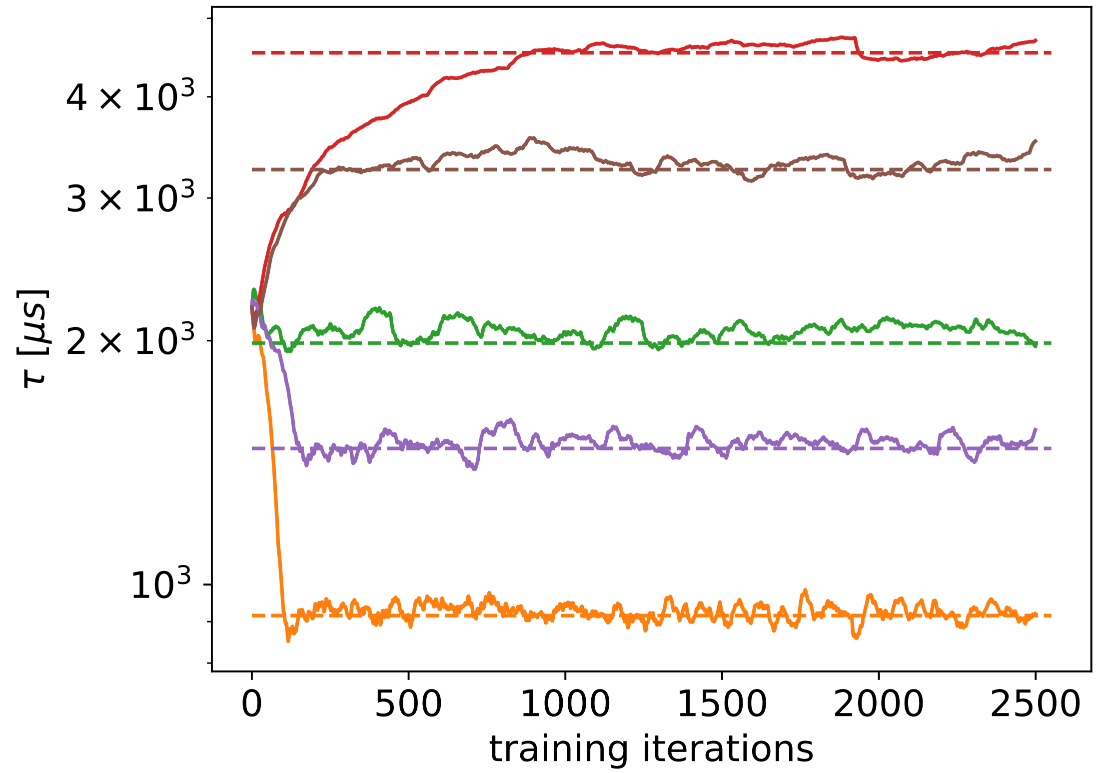


Fits with “Half Fake Data”

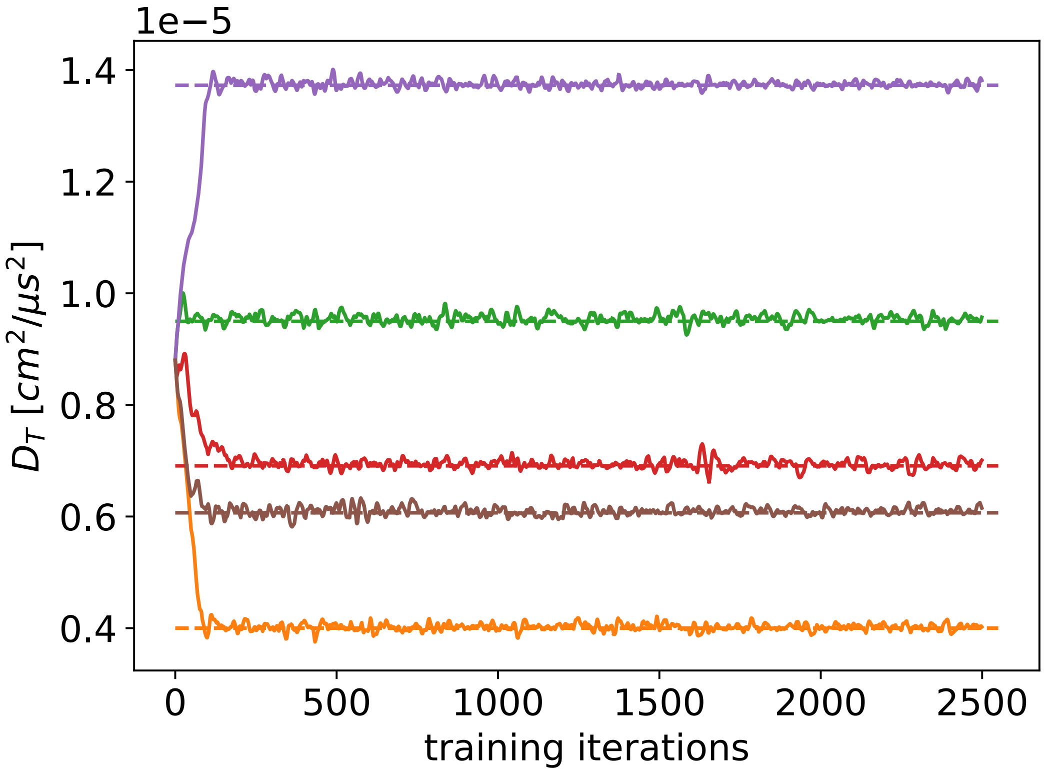
Electric field



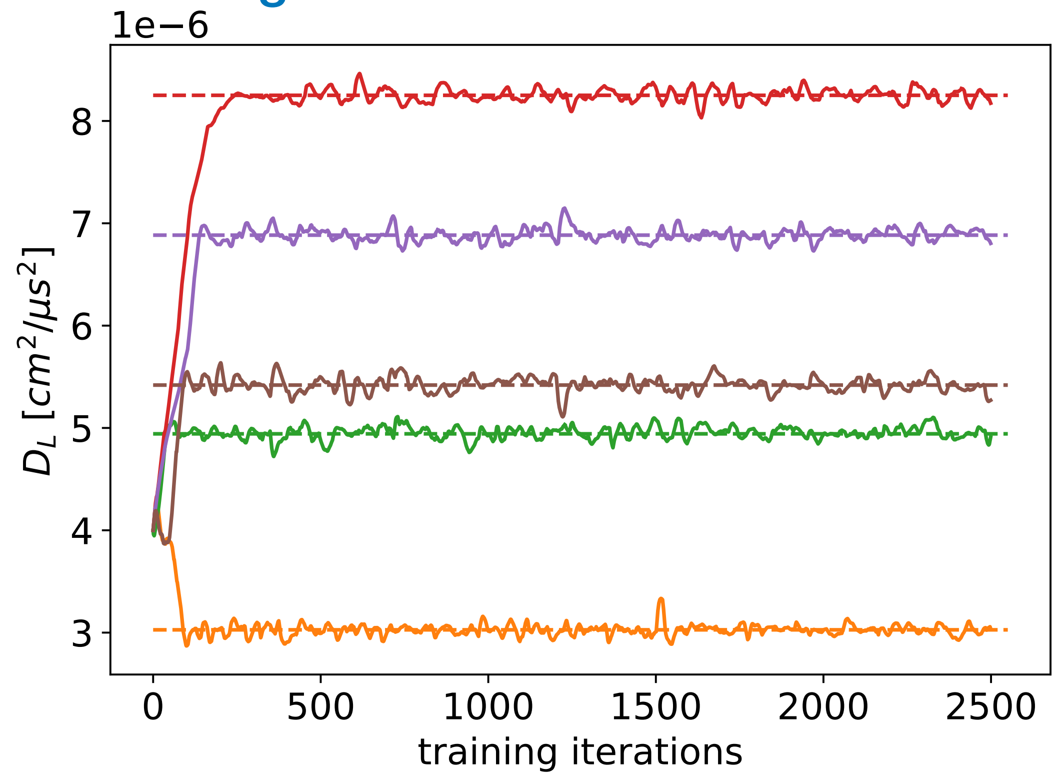
Electron lifetime



Transverse diffusion coefficient



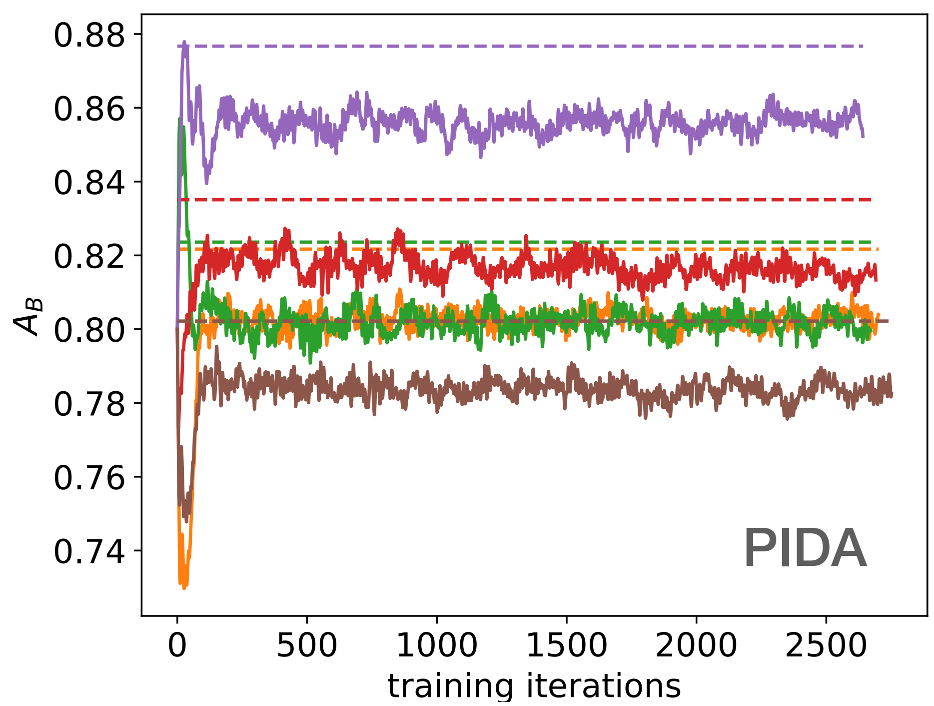
Longitudinal diffusion coefficient

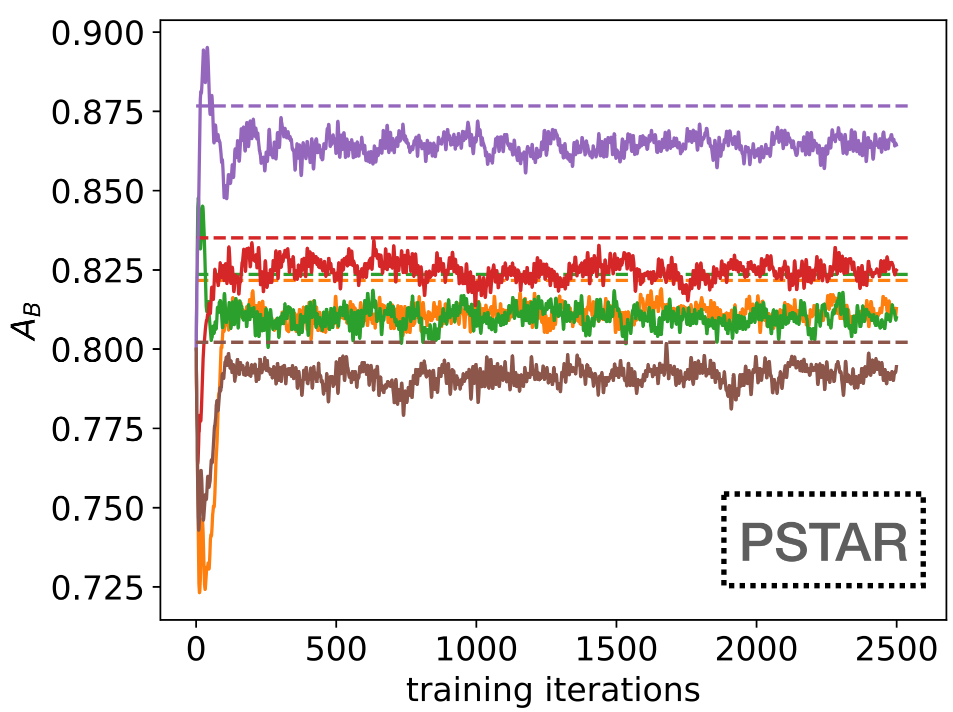


Fits with “Half Fake Data”

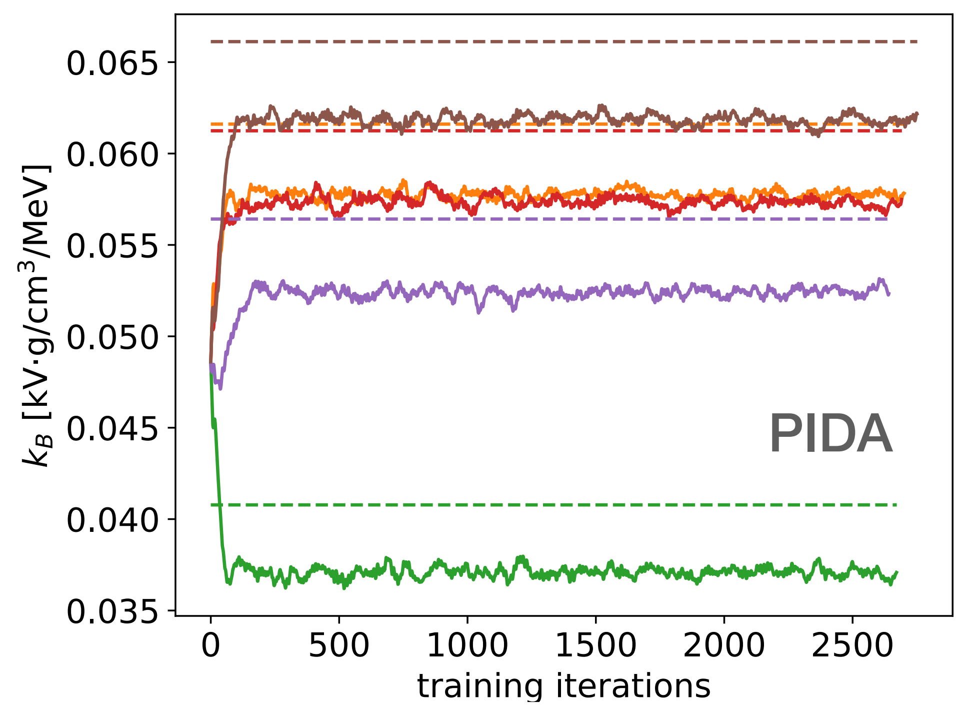
Recombination parameters A_B and k_B converge to values with offsets to the targets. They are sensitive to the absolute dE/dx values, while the other parameters can reach the optimal with the relative calorimetric information and position of the charge deposition

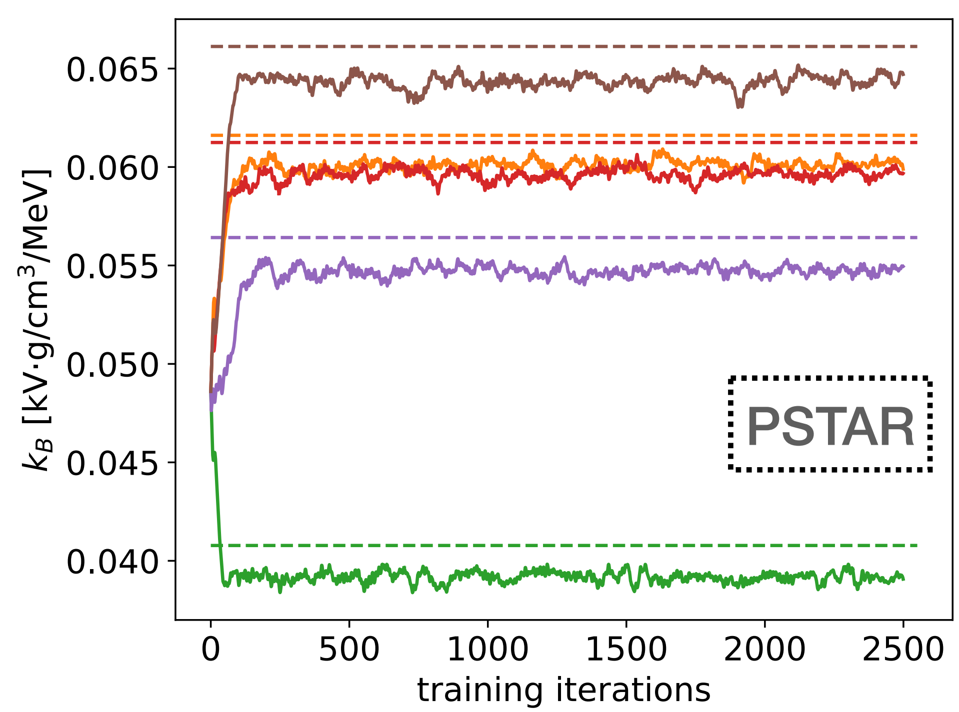
Recombination model A_B





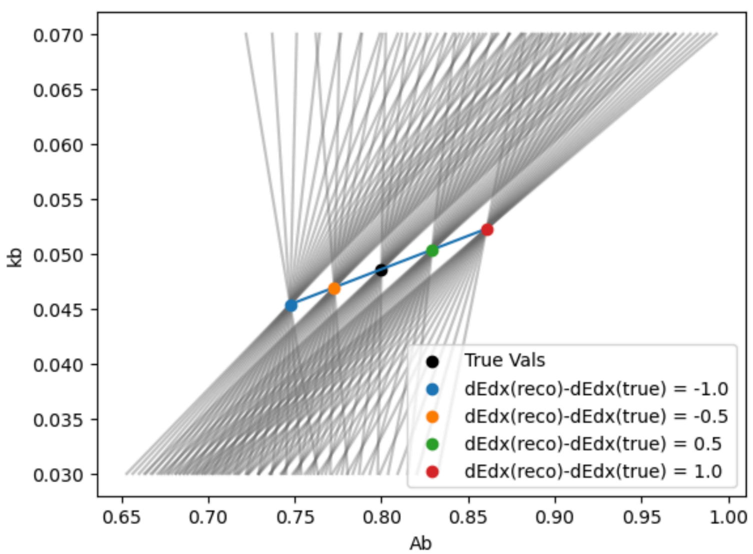
Recombination model k_B





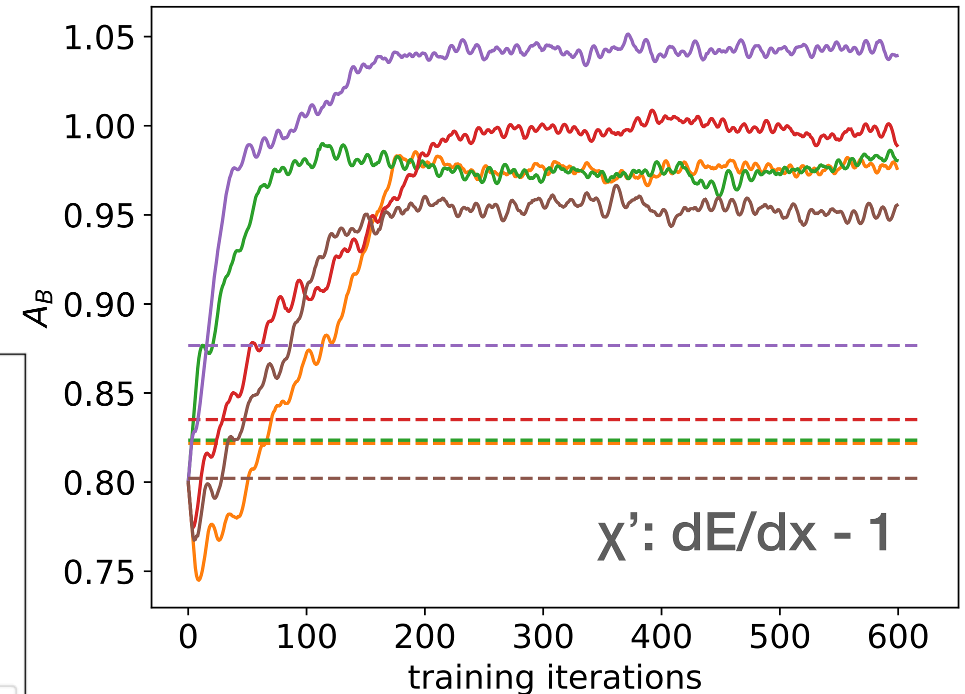
Impact of dE/dx Biases

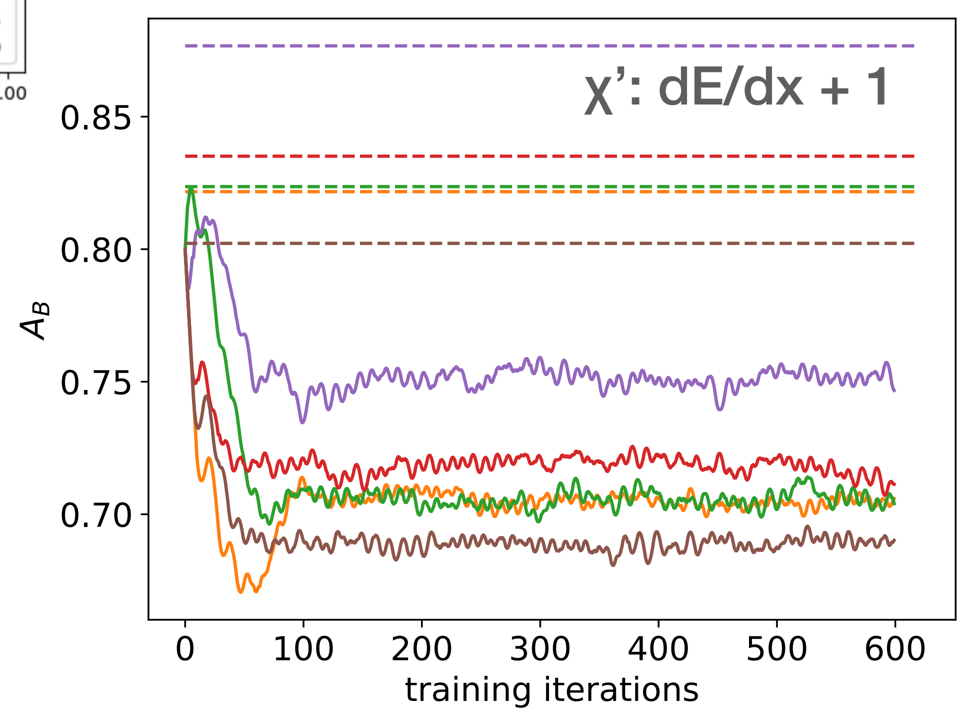
Opposite trend to our calculation 😲



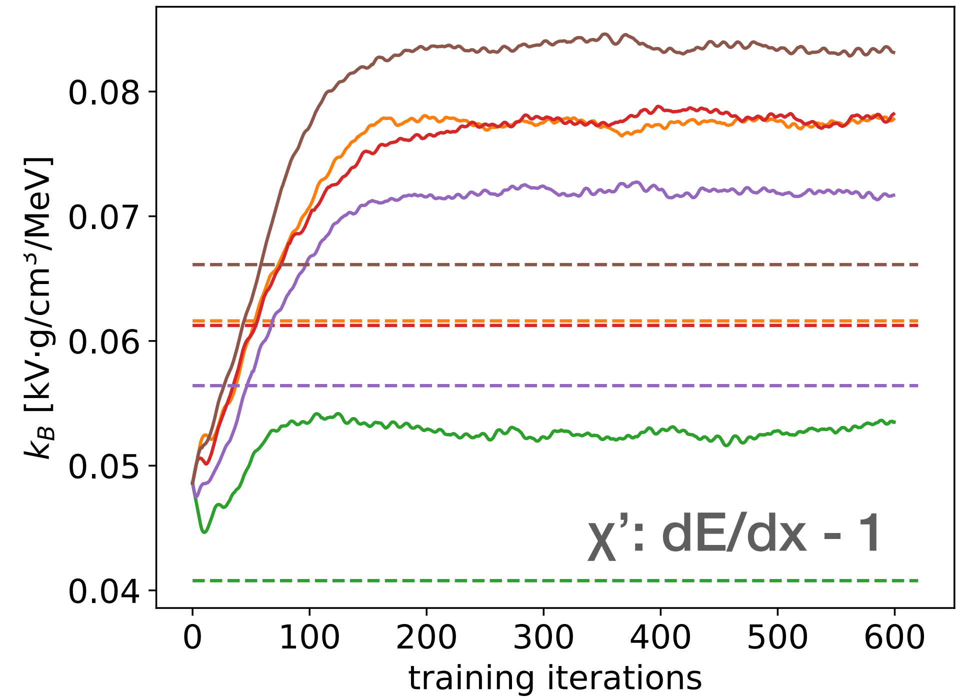
Given that in a batch, each segment dE/dx could be different and mis-modelled differently. The crossings of A_B and k_B are further smeared.

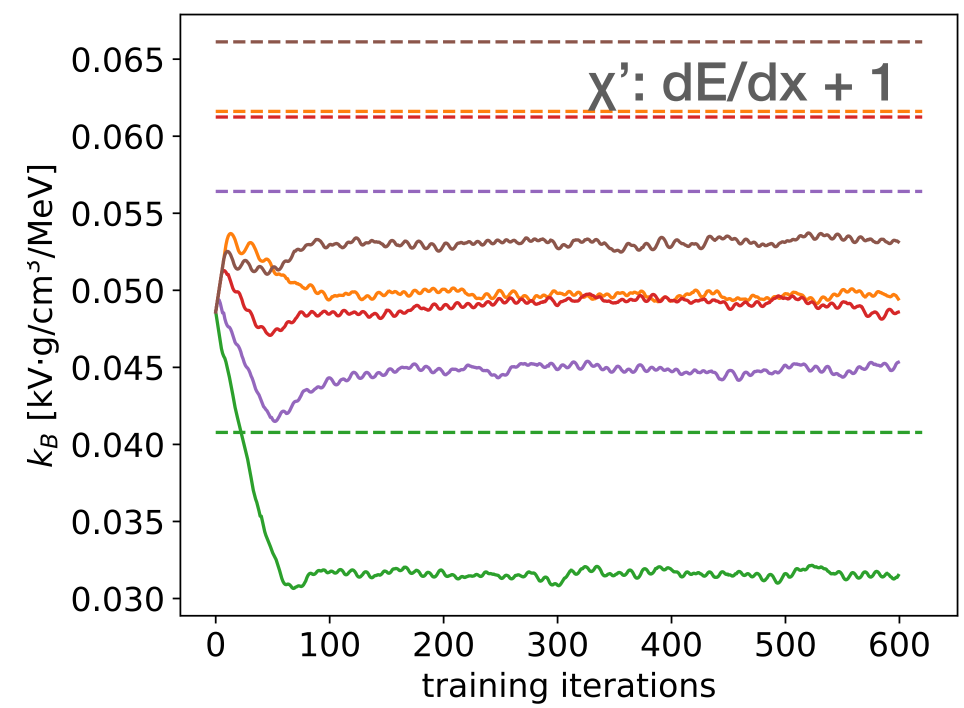
Recombination model A_B





Recombination model k_B

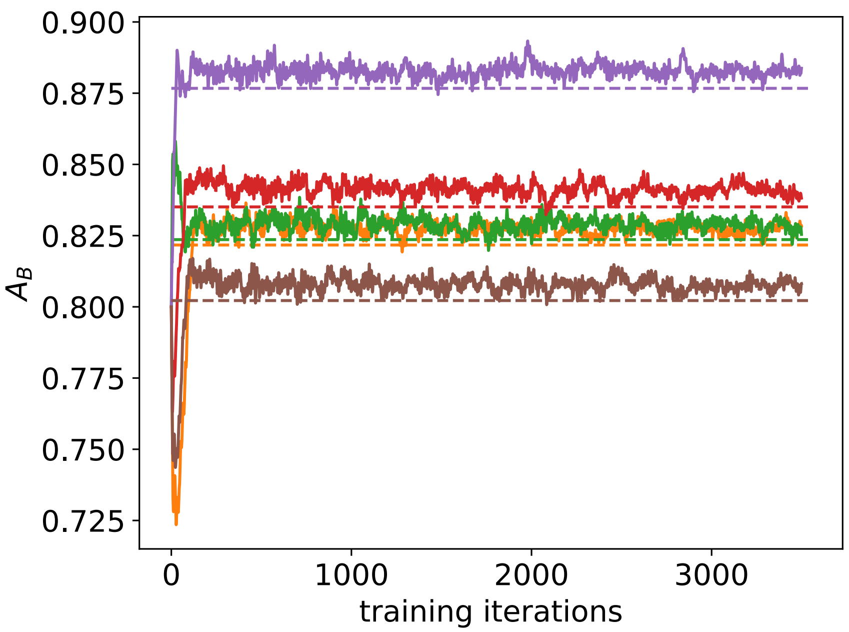




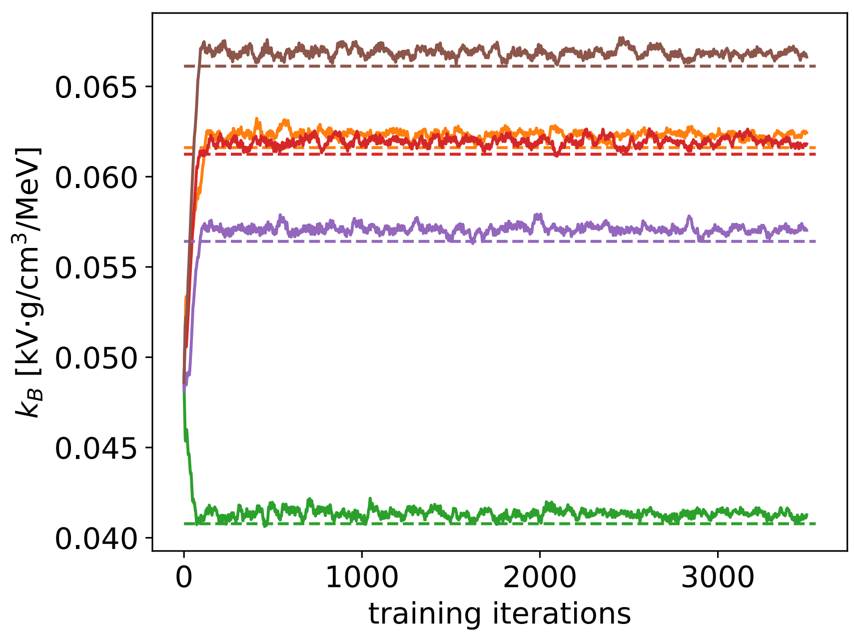
Fits with “Half Fake Data” Excluding the Very End

5 fits with different targets in 6D phase space.
The fits use 100 cm mini-batch.
Small offsets in A_B and k_B (room for improvement).

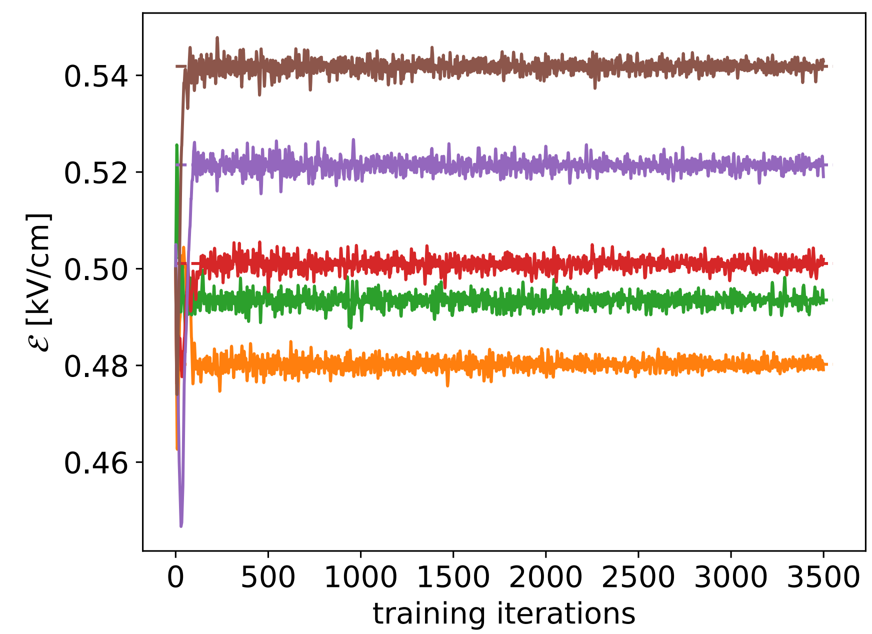
Recombination model A_B



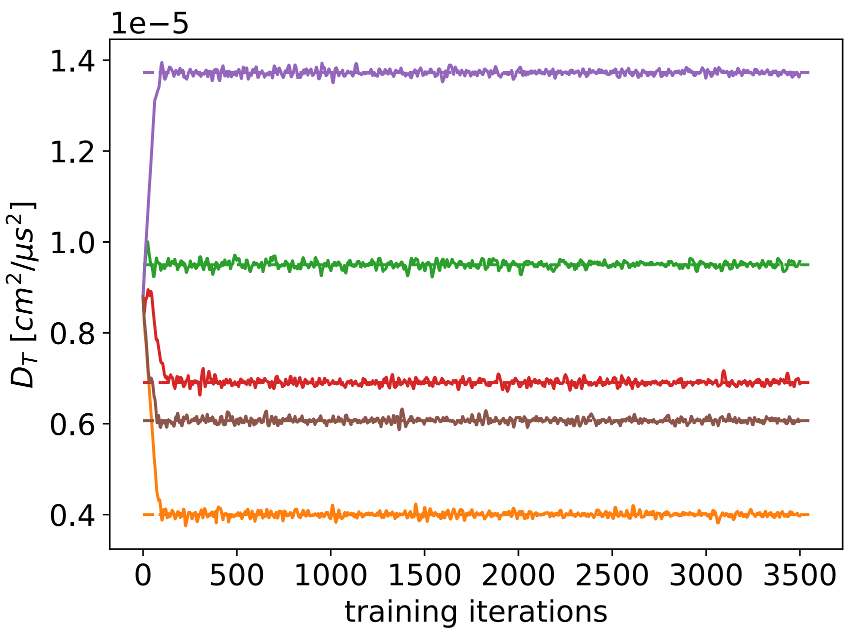
Recombination model k_B



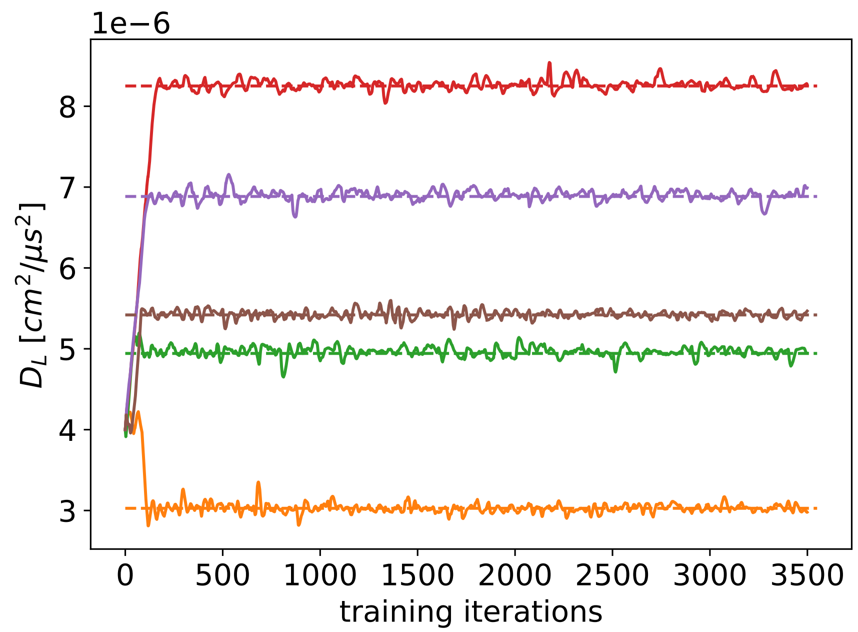
Electric field



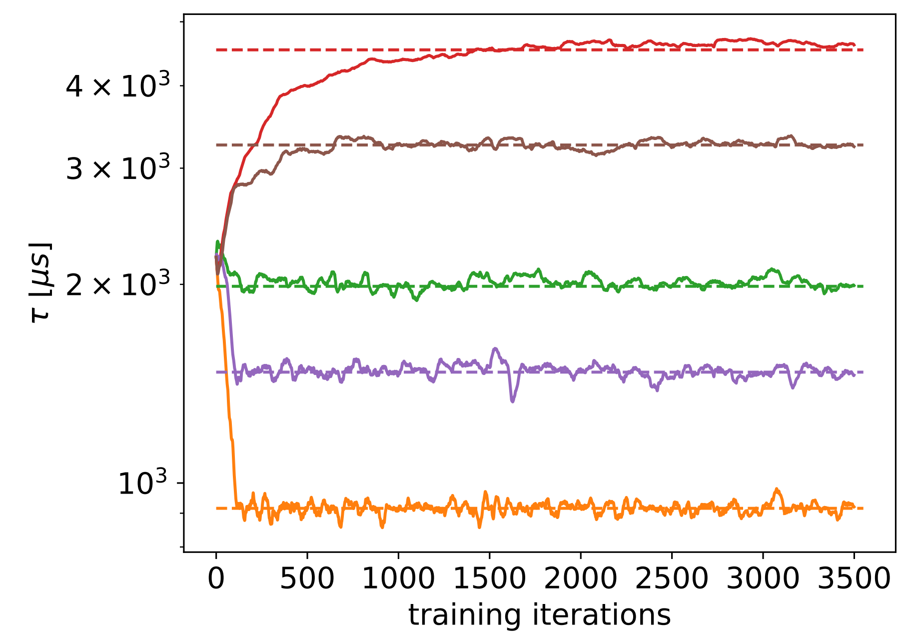
Transverse diffusion coefficient



Longitudinal diffusion coefficient

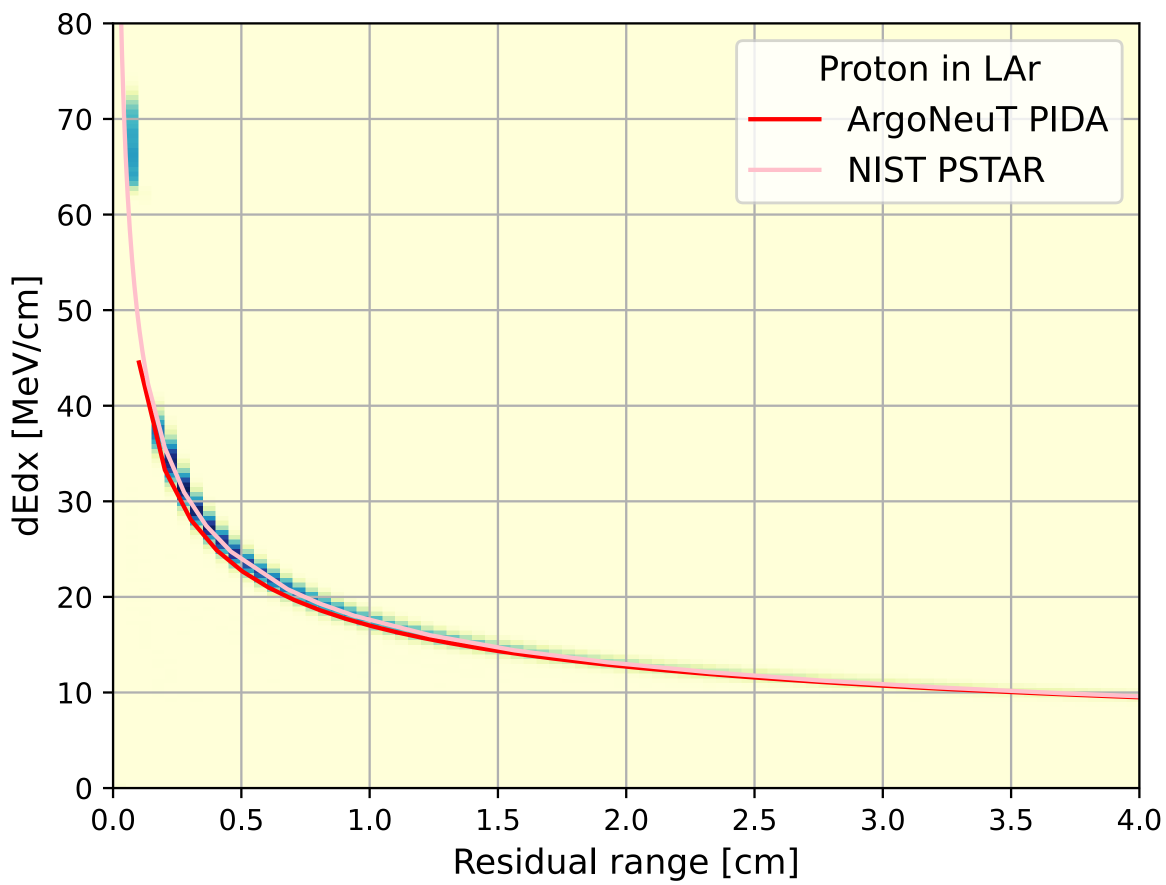


Electron lifetime

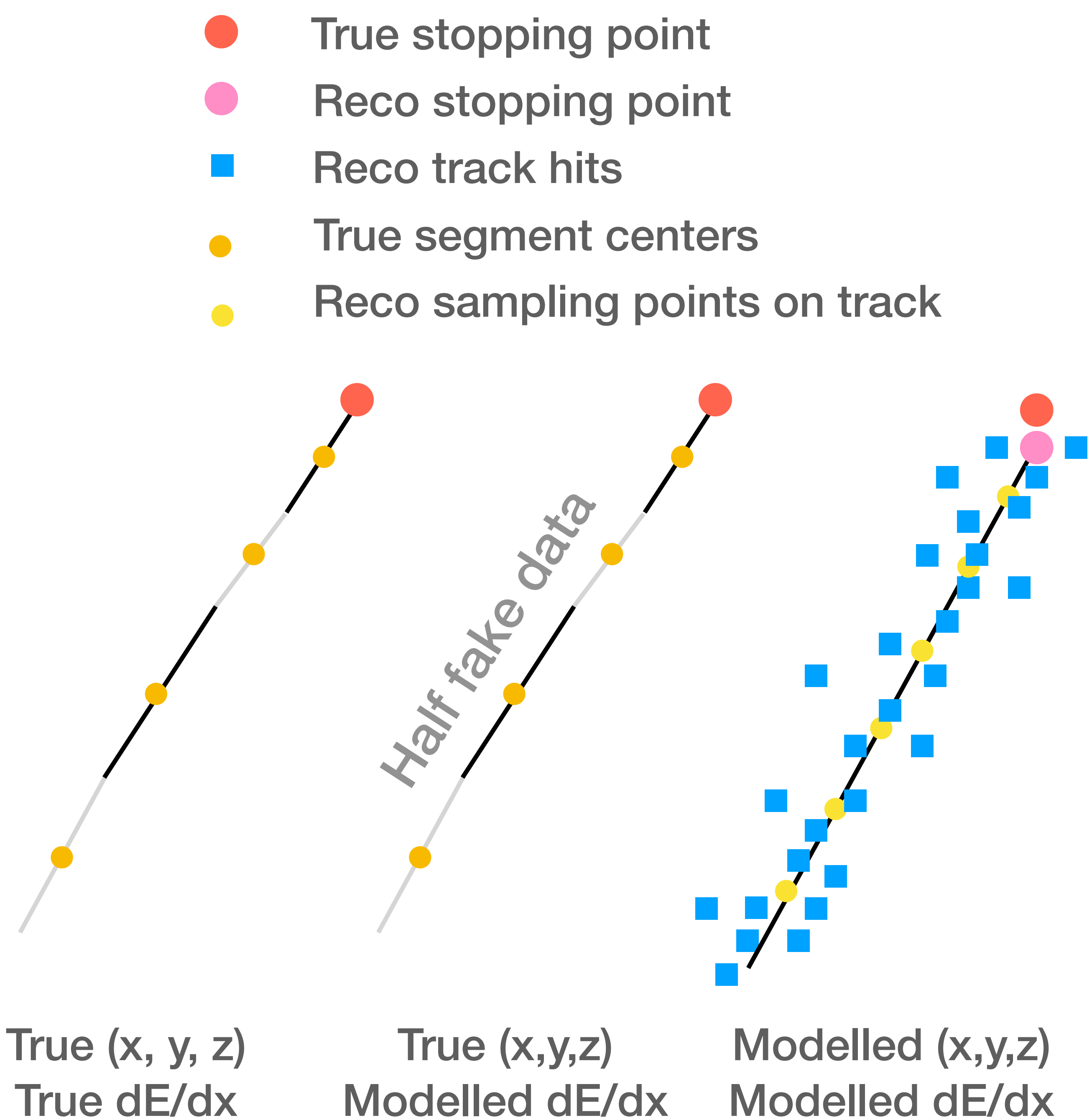


Simulation Input: Residual Range

Segments ("short" lines to represent particle trajectory) with position and dE/dx (dE).
Using a stopping proton sample.
Model dE/dx from the residual range.



Proton could re-scatter which would smear the reco residual range.

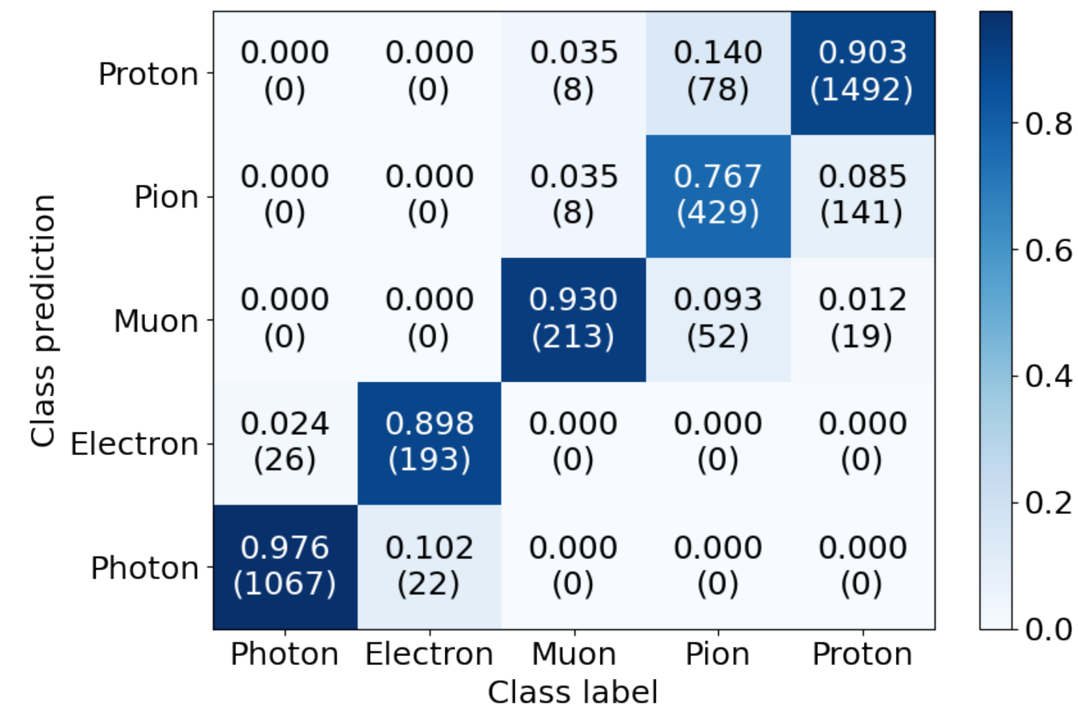


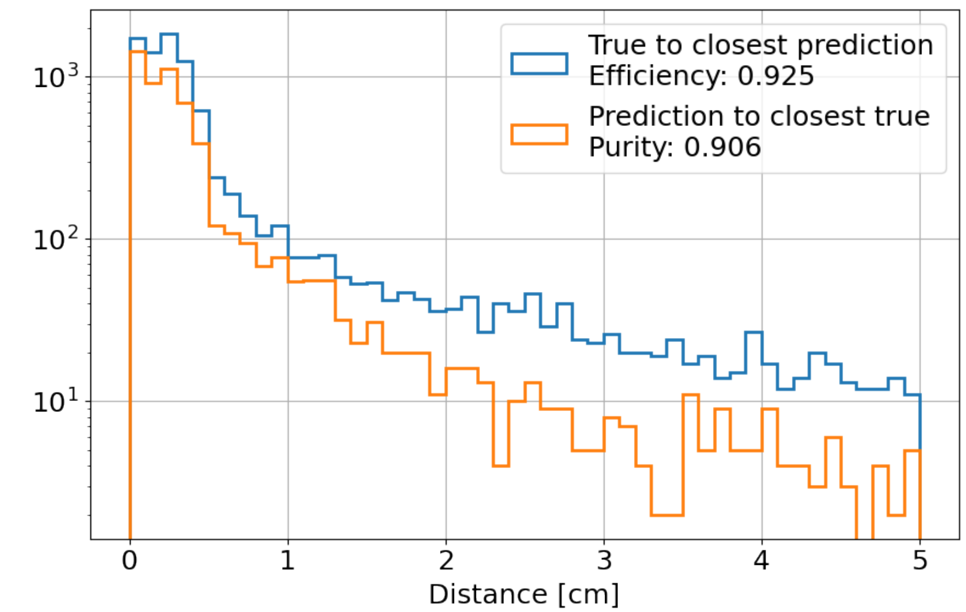
Towards Data Application

Calibrate the simulation with data!

Francois D.

- Fit with reconstructed proton position
- Proton identification and start/end recognition
- Fit with selected and reconstructed stopping protons (Fake data!)
- JAX version for a faster differentiable simulator
- Fit with real data! (Data preparation etc.)
- Produce metrics to express the fitting convergence quantitatively
- Extract uncertainties of the fitted parameters





Towards Data Application

arXiv: 2211.01505