



Modeling Absorbers in a Differentiable Muon-Cooling Simulation

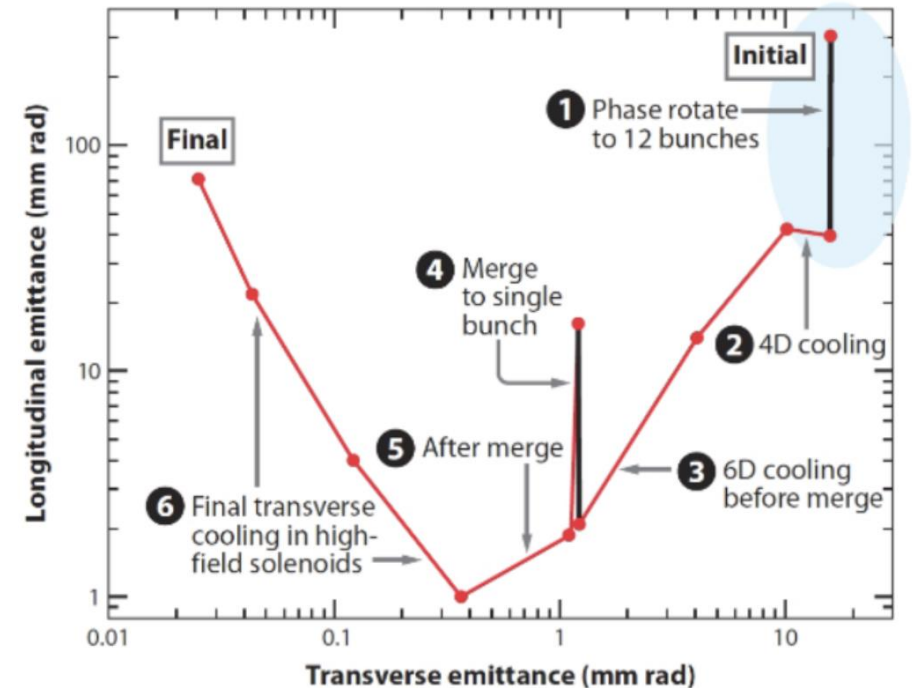
Energy straggling & multiple scattering, built on a JAX beamline stepper

Eric Anderson University of Illinois Chicago

Fermilab · UIC

A muon collider needs a cooling channel

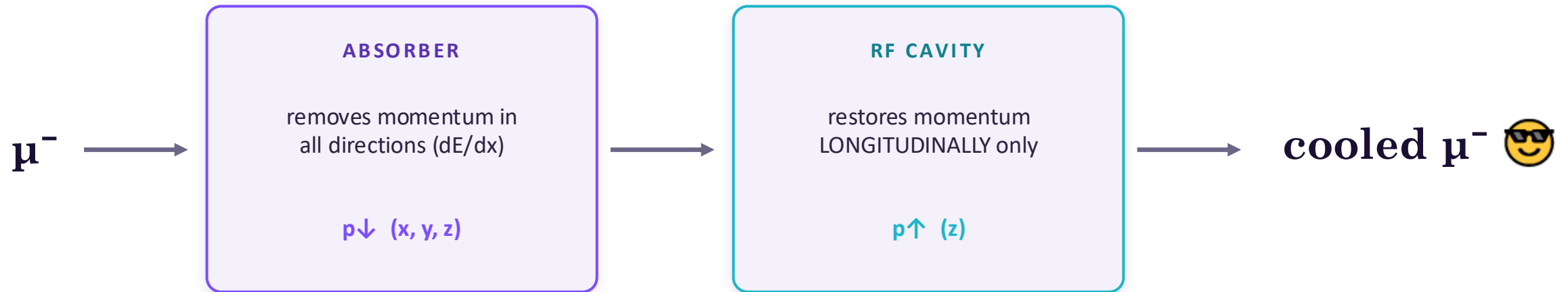
- "A muon collider at FNAL would be cool, but it needs to be cooled." – Nick Smith
- Basic problem: muons aren't not focused into a beam when they're produced.
- But they also have a ~ 2.2 microsecond lifetime, so we have to keep them moving near light speed.
- This makes the cooling channel design one big optimization problem.
- We want to make a cooling channel that will shrink 6D emittance by $\sim 10^5$ over a ~ 1 km lattice.
- 12 stages $\times$ 20–100 cells, each cell with several tunable knobs—fields, geometry, RF.



The simulation problem: an atypical mix of beam optics, material interaction, and collective effects—all differentiable.

HOW IT WORKS

Ionization cooling: repeatedly give and take momentum



Net effect: transverse momentum spread shrinks each pass $\rightarrow$ the beam cools.



But the absorber also heats the beam. Energy-loss straggling and multiple Coulomb scattering spread the beam out. Modelling these two effects is what my work is focused on.

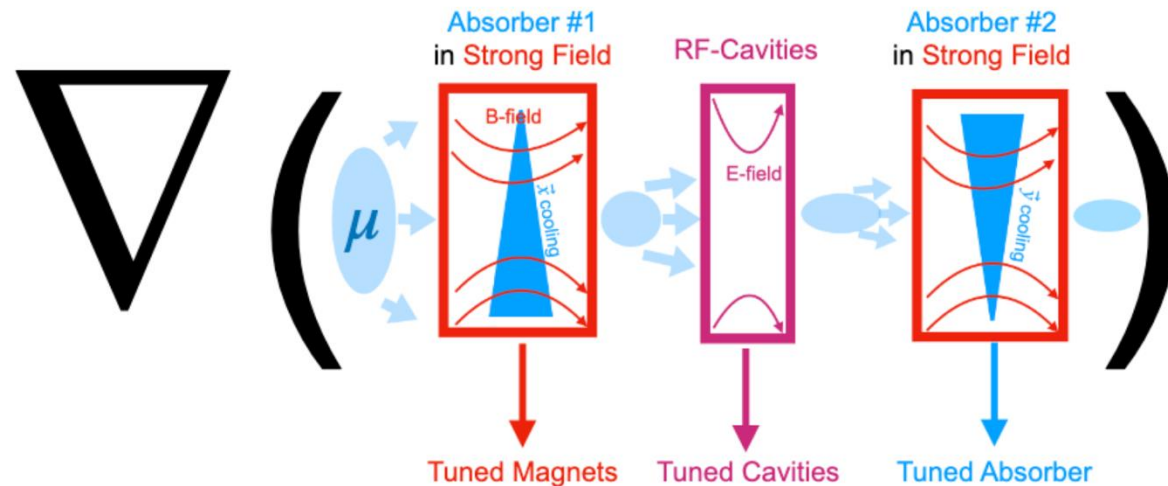
If the simulation is differentiable, gradients tune everything

Hand-optimization

- Tune a few knobs at a time, cell by cell
- Conservative working points
- Improvements don't compose across 100s of cells

Gradient-based optimization

- $\partial(\text{emittance}) / \partial(\text{every parameter})$ at once
- High dimensional optimization, not guess-and-check
- Whole-channel codesign becomes tractable



$$\frac{\partial \epsilon}{\partial \mathbf{B}} = \frac{\partial \epsilon}{\partial A_2} \frac{\partial A_2}{\partial \mathbf{B}} + \frac{\partial \epsilon}{\partial A_2} \frac{\partial A_2}{\partial \text{RF}} \frac{\partial \text{RF}}{\partial A_1} \frac{\partial A_1}{\partial \mathbf{B}} + \dots$$

P. Harris

A ground-up beamline simulator in JAX

“Why not just use GEANT or something?” → Because no existing tool has everything we need:



Optics

beam transport through fields



Material

energy loss & scattering in absorbers



Autodiff

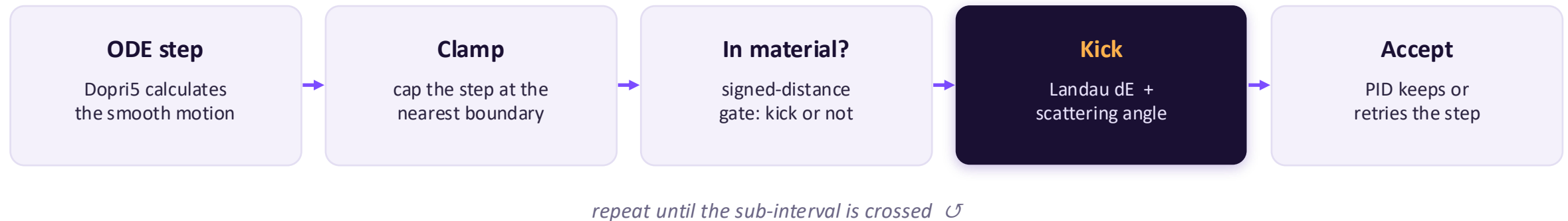
first-class gradients, GPU for free

- **First-class automatic differentiation** — gradients flow through the entire simulation without hand-deriving them
- **GPU/TPU acceleration out of the box** — the same code runs on accelerators with no rewrite
- **Both come "for free" from the framework** — no custom autodiff or GPU-kernel work to build and maintain



How the programming works

We simulate walking the beam through a material by computing effects in small steps. Between materials the muon follows a smooth ODE; inside an absorber each sub-step also draws random kicks.



The kick is always applied

Under jit (a JAX operation) you can't branch. So the angle is `where(in_material,  $\theta$ , 0)`. Outside material it's a zero-angle no-op.

Thickness from geometry

Each sub-step's material thickness is the true 3D path length—so a tilted track automatically sees more material.

Two stochastic effects—the source of the heating

1 Energy straggling

- Energy loss per crossing is random, not fixed
- Highly-skewed Landau distribution
- Characterized by the most-probable loss, not the mean

2 Multiple scattering

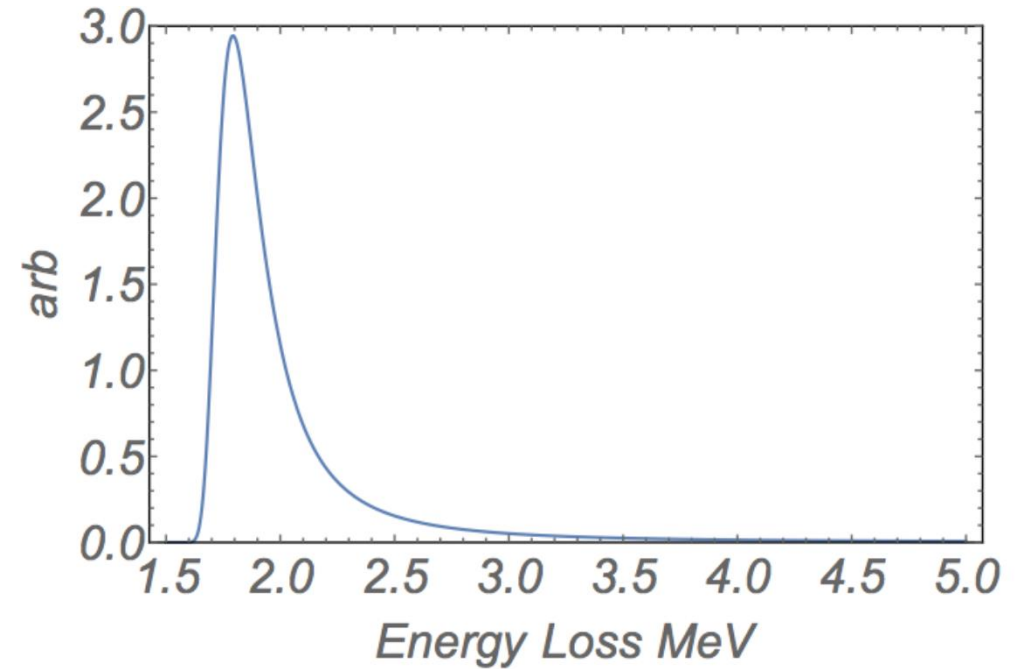
- Many small Coulomb deflections off nuclei
- Gaussian exit angle distribution because of the central limit theorem
- RMS of that distribution calculated from Highland formula

The Landau energy loss spectrum

- The mean is dragged up by rare, hard atomic collisions—experimentally ill-defined.
- The most-probable loss (the peak, Δ_p) is the robust, physical quantity.
- Sampled differentiably from a JAX Landau generator.
- Width scales with ξ ; mode from the Bethe–Landau most probable formula.

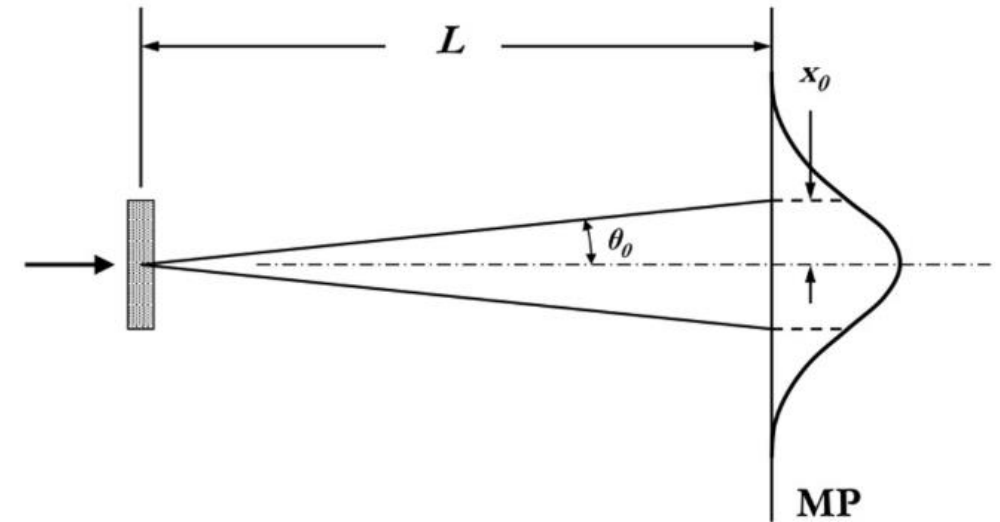
most-probable energy loss

$$\Delta_p = \xi \left[\ln(2mc^2\beta^2\gamma^2/I) + \ln(\xi/I) + j - \beta^2 - \delta \right]$$



Highland scattering, sampled with the right correlations

- Two independent Gaussian deflection planes with the same RMS.
- Angle and lateral offset are correlated ($\rho = \sqrt{3}/2$) and sampled jointly.
- The kick rotates the momentum about its current direction.
- $|p|$ and energy are conserved exactly.

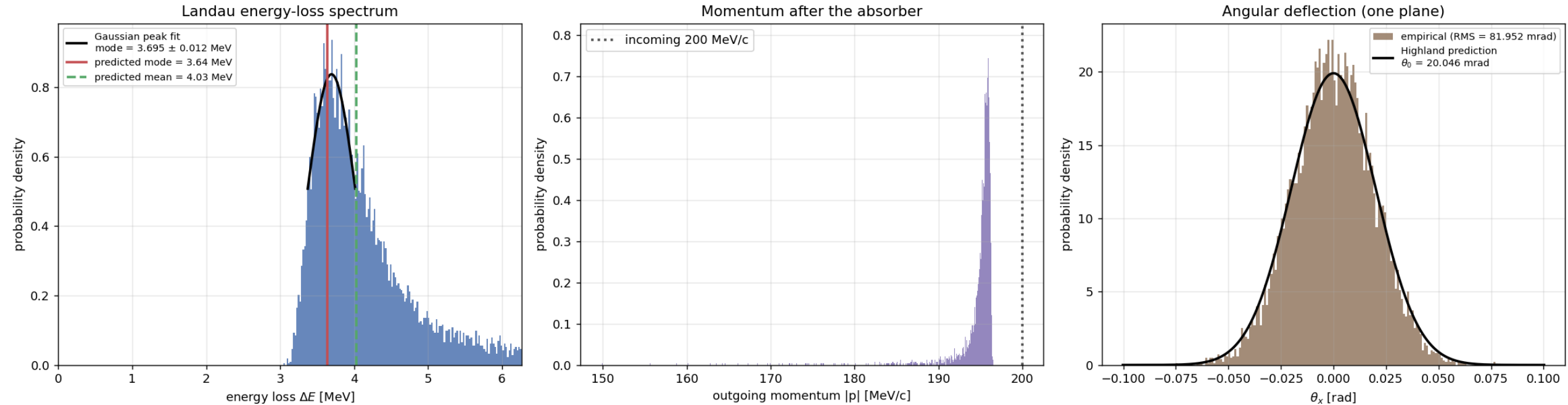


Highland RMS angle

$$\theta_0 = (13.6 \text{ MeV} / \beta c p) \cdot z \cdot \sqrt{(x/X_0) \cdot [1 + 0.038 \ln(x z^2 / X_0 \beta^2)]}$$

A validation test I built

Propagate 10^4 muons (200 MeV/c) through 10 mm of SiO_2 with the real stepper, then check every observable against a closed-form prediction:



● **Landau peak** fitted mode 3.70 MeV = predicted

● **Momentum** 200 $\rightarrow$ ~196 MeV/c, degraded as expected

● **Highland angle** histogram sits on the $\theta_0 = 20.0$ mrad curve

Splitting the crossing underestimates the exit angle

- You're supposed to do a single MCS calculation over the entire length of the absorber.
- We apply many kicks in small substeps through the absorber, and as a result we systematically underestimate the exit angle (PDG warns against this explicitly).
- "So why not just do the calculation how PDG says to?"
 - Because we have to compose MCS calculations with fields and curved trajectories.
 - Real absorbers are wedges, not slabs, so true path length for individual particles varies; the substep handles this naturally.
- Another problem: with substeps, you automatically generate a lateral displacement just from repeated angular displacement + propagation.
- We need to quantify any disagreement that displacement has from the Highland-predicted displacement.

The deficit:

About $[1 - 0.04 \ln (\text{steps})]$

grows with step count

measured

$0.94 \times \theta_0$

bulk θ_x width at ~ 4 sub-steps —
The same as what $1 - 0.04 \ln 4$ predicts

Three potential ways to control the underestimate

1

Telescoping variance

Give each substep a kick of variance $\theta_o^2(x+\Delta x) - \theta_o^2(x)$. The variances telescope to exactly $\theta_o^2(x_total)$, no matter how the PID subdivides.

2

One kick per region

Detect a contiguous material span and apply Highland once over the full thickness. Physically the cleanest, matching PDG's prescription. Not clear how this would square with the substep integrator though.

3

Measure & correct

Accept the deficit, characterize it, and correct downstream. This is where the code sits right now.