# RICH alignment with Bayesian optimization

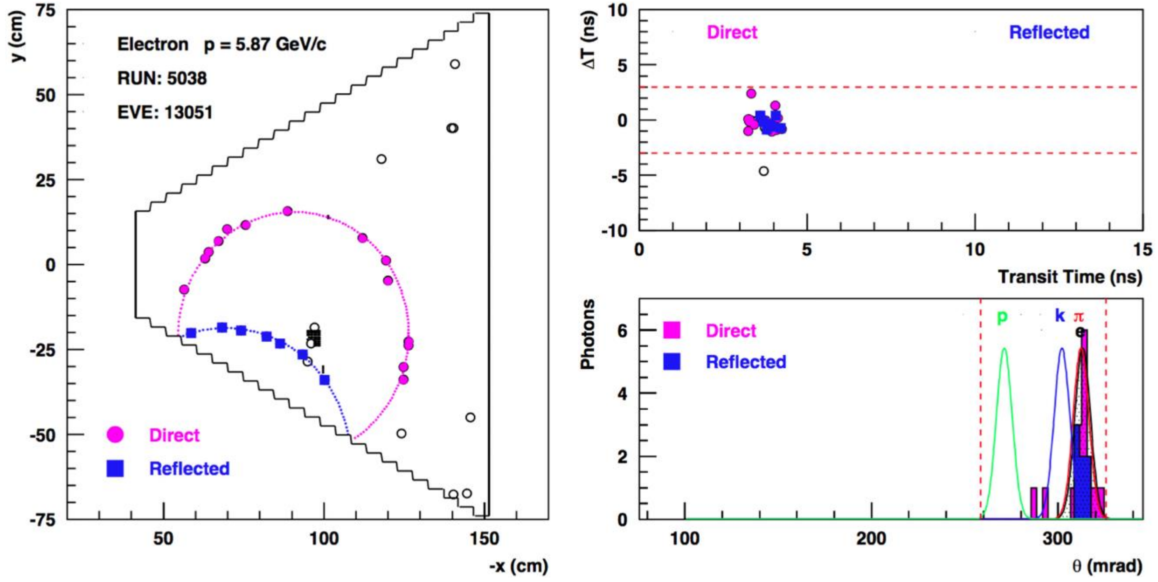
Connor Pecar
Duke University

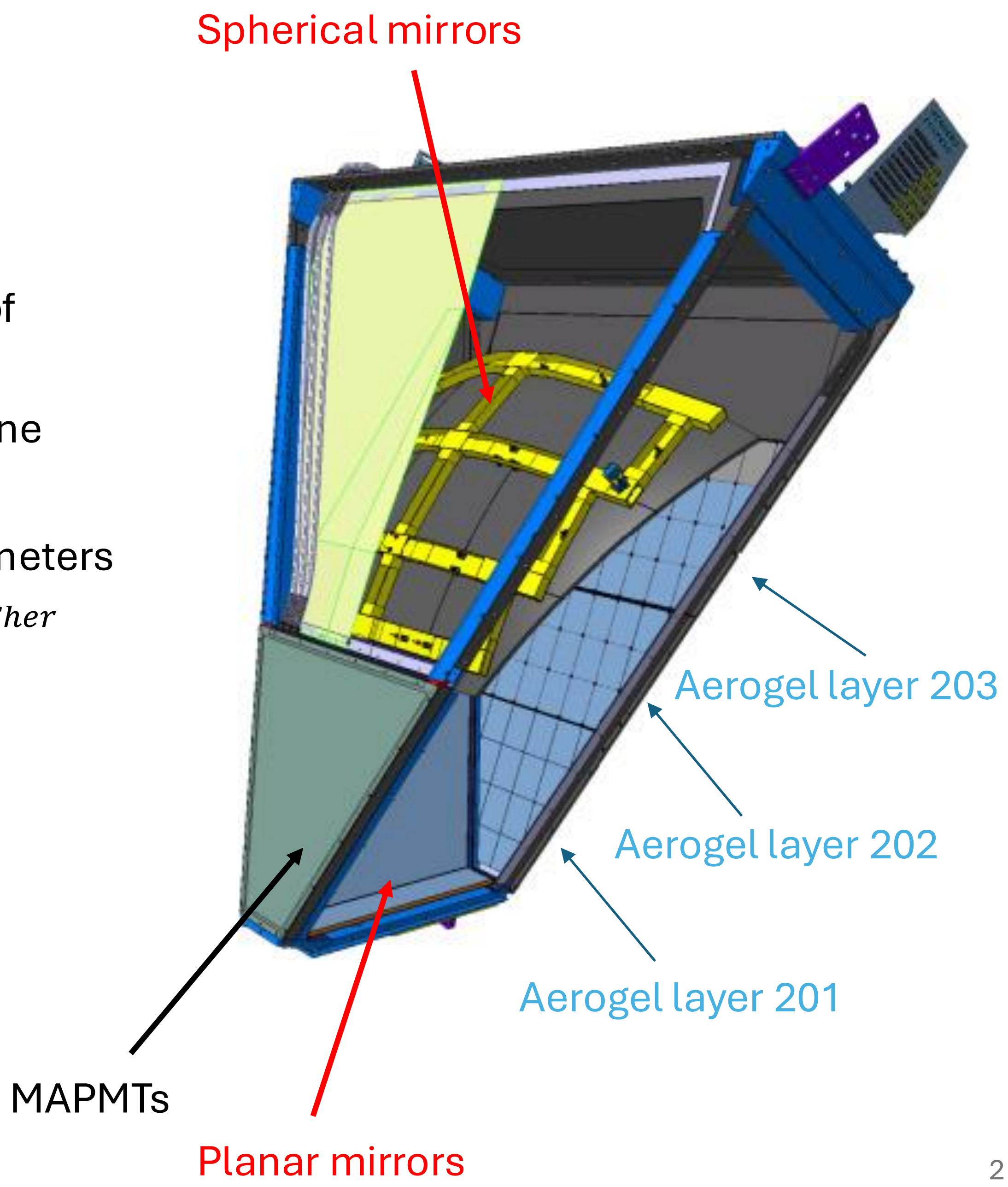
CLAS Collaboration Meeting, March 10th, 2026

Research supported by
DOE Office of Science

# RICH detector and geometry

- Identify charged hadrons using reconstructed Cherenkov angle information
- Two 2cm layers of aerogel tiles (layers 201/202) and one layer of stacked 3cm tiles (layer 203)
- 5 planar + 10 spherical mirrors focusing photons to MAPMT plane
  - Many possible photon reflection paths (photon "topologies")
- **Goal of this work: simultaneously** determine alignment parameters for ALL optical components of the RICH using reconstructed $\theta_{Cher}$

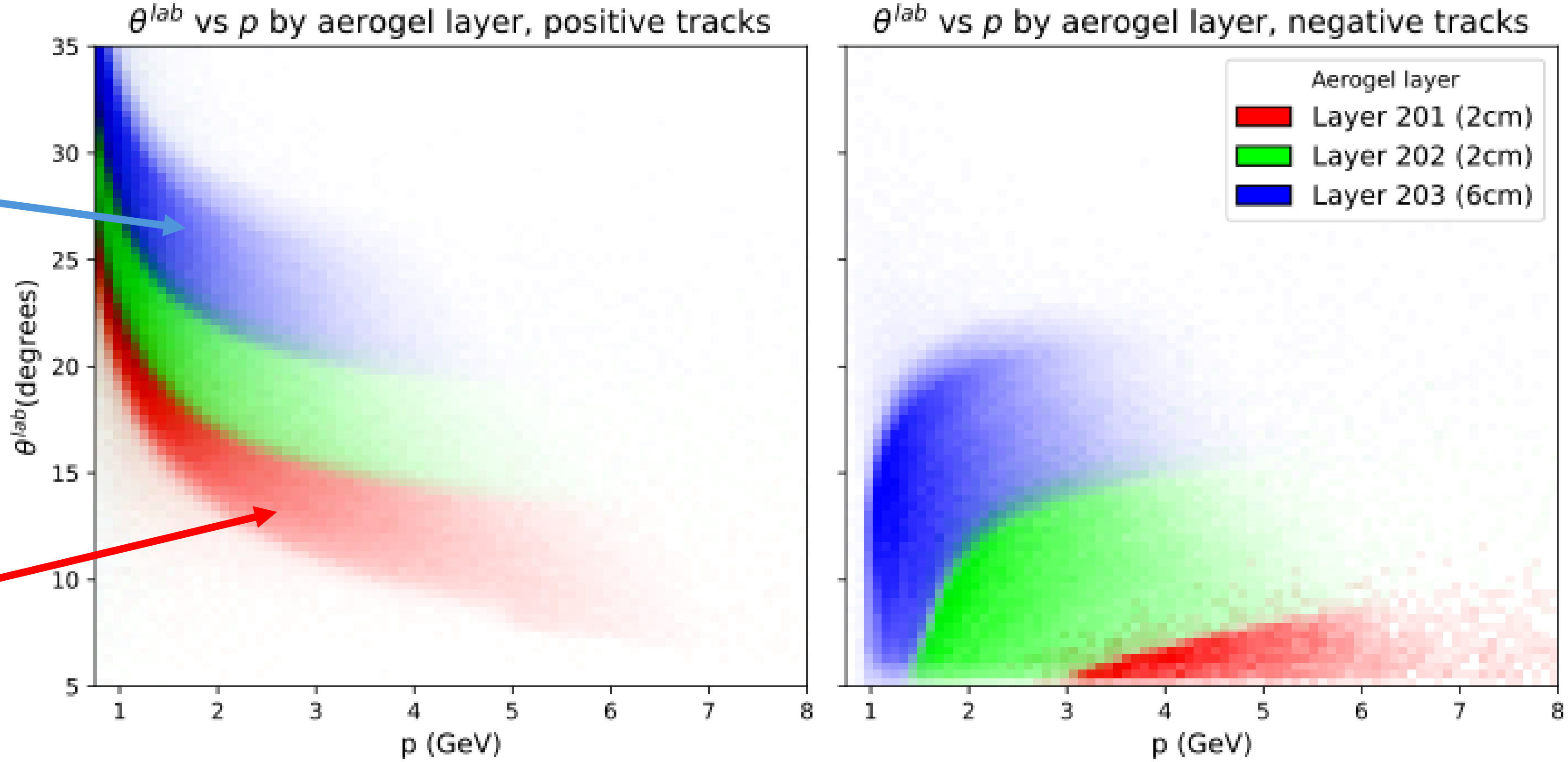


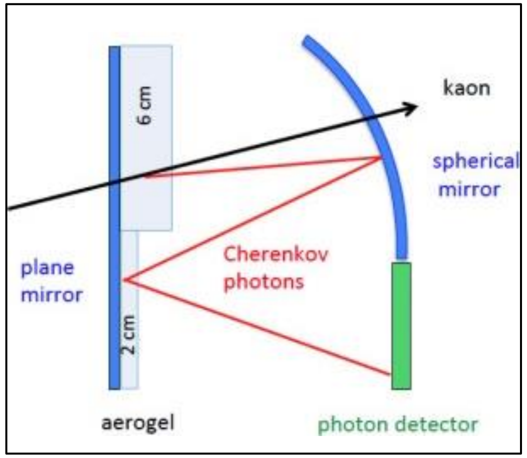



# Acceptance by aerogel layer

Layer 203 primarily 2+ reflection photons

Layer 201 primarily 0 or 1 reflection photons

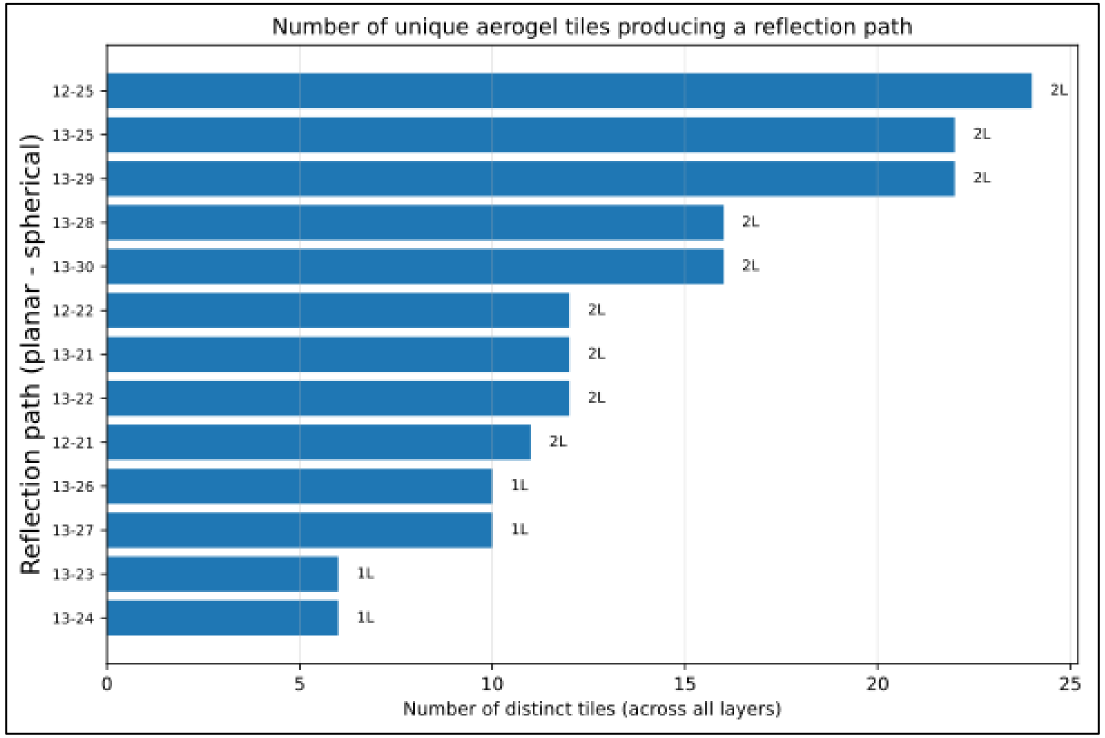



- Layer 203 and upper spherical mirrors previously not successfully aligned
- Critical to large polar angle acceptance

# RICH alignment challenges

- Alignment parameters highly correlated
  - Many possible reflection paths for any given aerogel tile
  - Need to align both the mirrors and the aerogel layers
  - Factorizing into subsets of components not optimal
- $\theta_{cher}$ reconstruction via direct ray tracing → no access to derivatives w.r.t alignment parameters, no a priori knowledge of mirrors hit
- Uncertainty on exact value of $n_{aero}/\theta_{cher}^{expected}$ for each tile, best determined from data
- 114 parameters total considering all DOF
  - Each component with some shifts $(dx, dy, dz)$ and rotations $(d\theta_x, d\theta_y, d\theta_z)$ in RICH reference frame

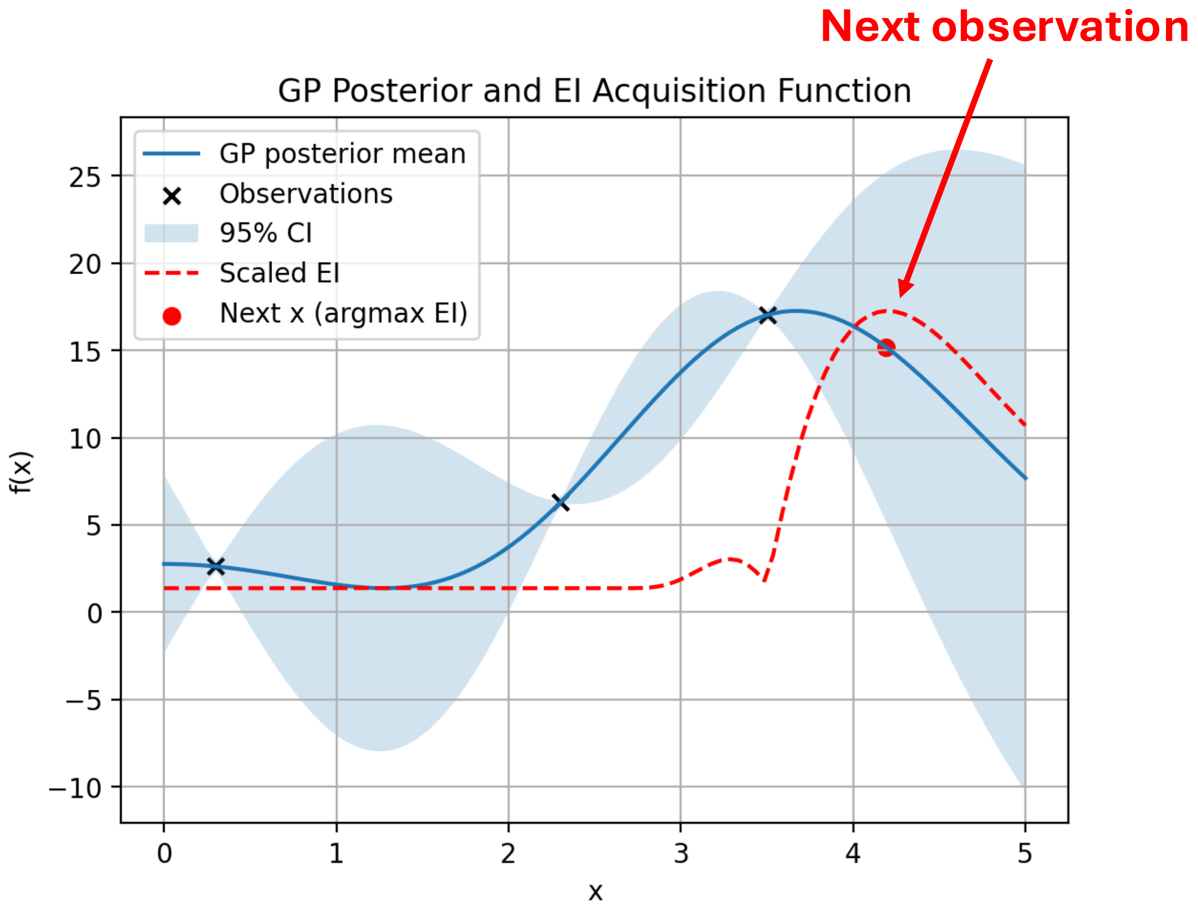



High dimensional, correlated, expensive optimization problem with no gradients → needs a guided and efficient way to search the parameter space

# Chosen tool: Bayesian optimization

**Bayesian optimization (BO)**: sample efficient black-box optimization

- Fit a probabilistic **surrogate model** of the expensive objective function
  - Gaussian process (GP) commonly used
- Use the surrogate posterior distribution to evaluate an **acquisition function**
- Sample from the expensive function (e.g. RICH reconstruction) at the point(s) where the acquisition function is maximized
  - Balances exploration (high uncertainty regions) and exploitation (high performance regions)




$$EI(x) = \sigma(x)\big(z \cdot CDF(z) + PDF(z)\big)$$

$$z = \frac{\mu(x) - f_{max}}{\sigma(x)}$$

# Trust Region Bayesian Optimization (TuRBO)

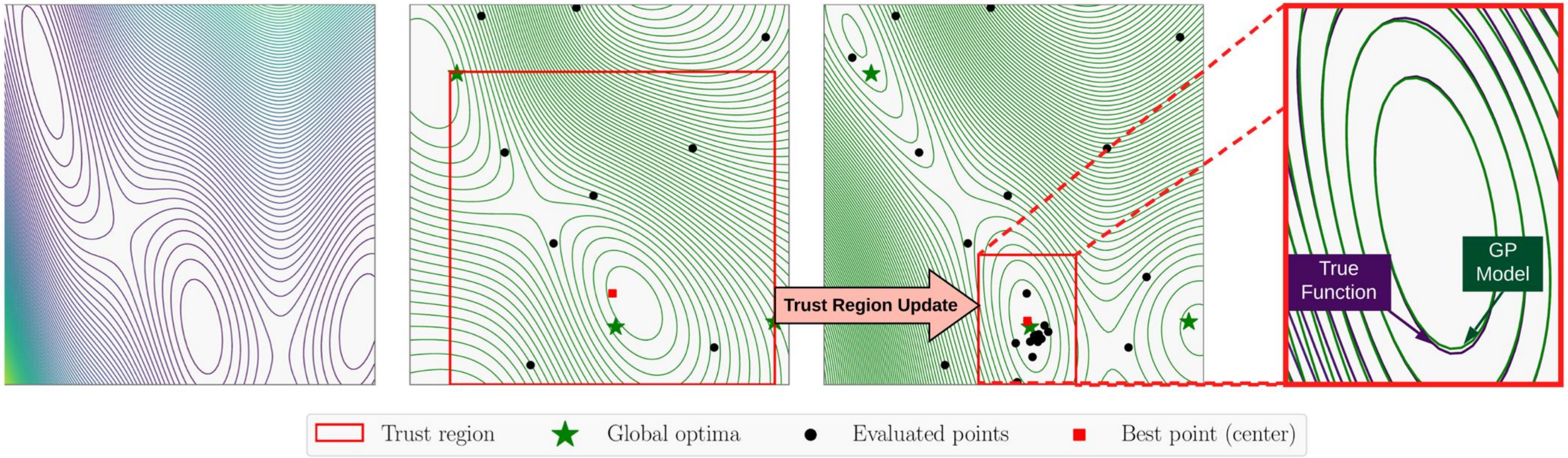



**Figure from arXiv:1910.01739, Branin test function**

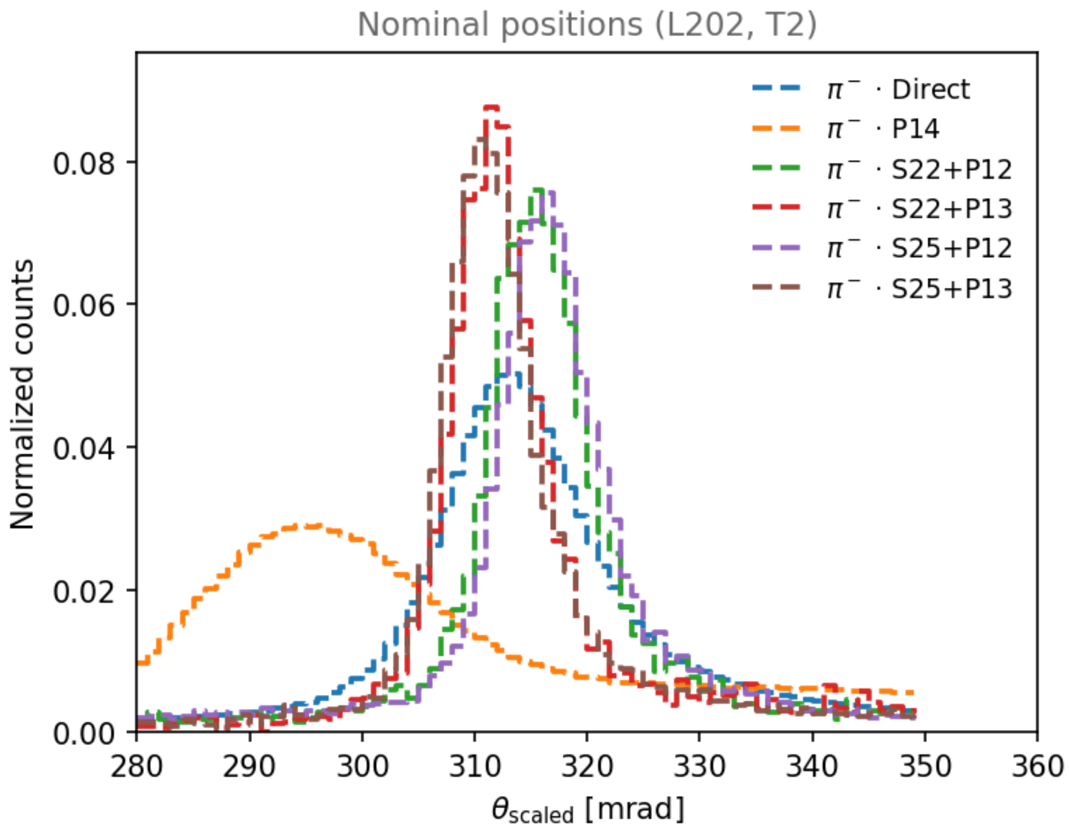
- Utilizing **TuRBO** (arXiv:1910.01739): BO algorithm designed for high-dimensional problems
- Instead of optimizing globally, only **maximizes acquisition function within trust region**
- Trust region (TR): hyperrectangle centered at current best solution
  - Multiple successive improvements: TR expanded (may not be at global optimum)
  - Multiple batches with no improvement: TR contracted around best point
- Currently implemented TuRBO-1 (single TR) for alignment, but (arXiv:1910.01739) also introduced TuRBO-$m$ ($m$ trust regions initialized independently + optimized jointly)

# Objective construction: reflection path agreement

- Photons produced from the same aerogel tile that undergo different reflection paths should agree on the position of the peak of the $\theta_{Cher}$ distribution
- For **each tile**, calculate a spread of the peak positions ($\theta_{peak}$) for the photon topologies produced from that tile,

$$S_i^{var} = \frac{1}{N_i^{topos} \langle IQR^2 \rangle_i} \sum_{j=1}^{N_i^{topos}} (\langle \theta_{peak} \rangle_i - \theta_{peak}^j)^2$$

  - $N_i^{topos}$: number of topologies produced from tile $i$
  - $\langle IQR^2 \rangle_i$ : squared interquartile range for tile $i$ averaged over topologies
  - $\langle \theta_{peak} \rangle_i$: peak position for tile $i$ averaged over topologies
- 2 reflection photons pass through aerogel twice → shift peak position by a small factor estimated from simulation (~3.5mrad)
- **No assumption on exact value of $n_{aero}$**



# Objective construction: resolution

- Then, for **every topology considered**, compute an average of the IQR (width) values over tiles that are well-populated for thaat topology (set $\mathcal{T}$),

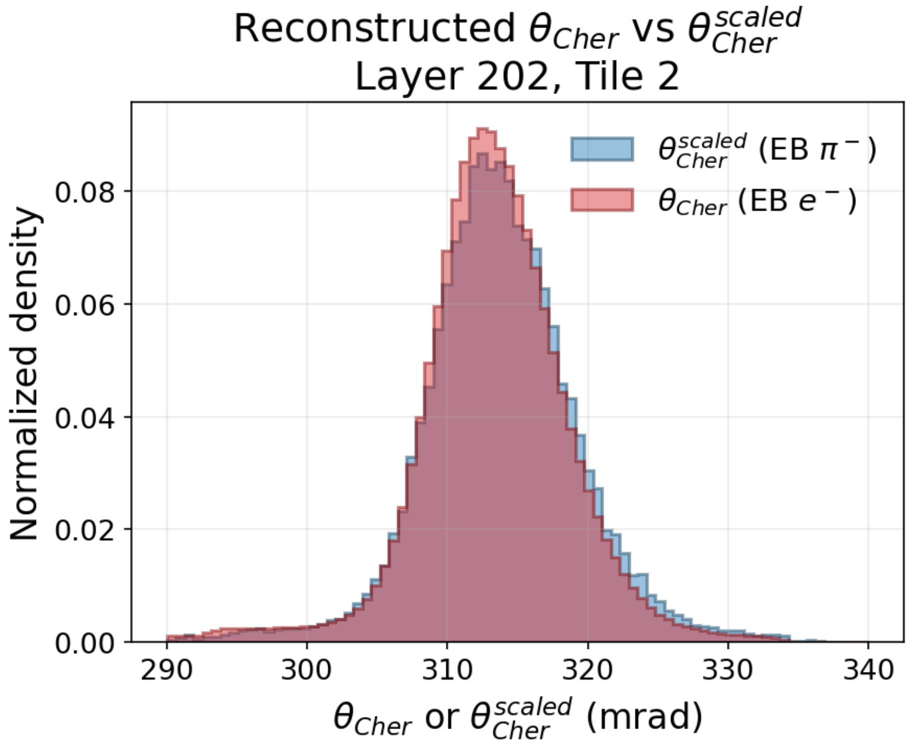
$$R^2_{topology} = \frac{1}{N^{topo}_{tiles}} \sum_{t \in \mathcal{T}} \left( \frac{IQR(\theta^{rec}_t)}{\sigma_{expected}} \right)^2,$$

  - IQR scaled to be equivalent to a Gaussian $\sigma$ ($IQR \sim 1.35\sigma$)
  - Considering only up to 2 reflection topologies
- Then define combined objective as

$$\boldsymbol{M_{align}} = (\sum_{\boldsymbol{i=1}}^{\boldsymbol{N_{tiles}}} \boldsymbol{S_i^{var}}) + \boldsymbol{R^2_{direct}} + \boldsymbol{R^2_{planar,1}} + \cdots$$

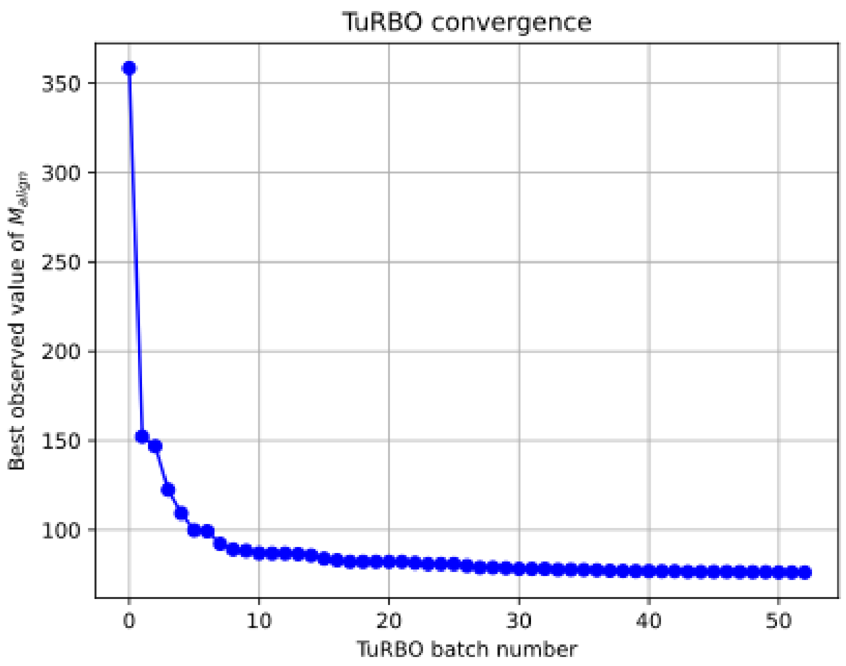
# Objective construction: selected tracks

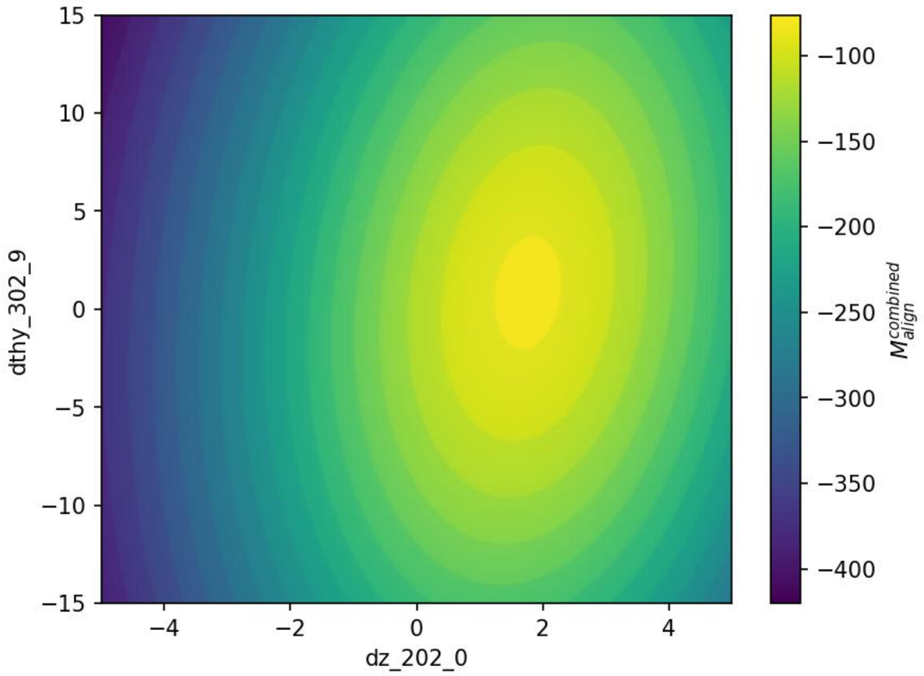
- $e^-$ the "natural" choice for alignment of the detector
- However, outbending $e^-$ do not provide sufficient coverage of planar mirrors, direct photons from aerogel layers 201 and 202
- To make this alignment approach identical for inbending and outbending datasets, **using $\pi^+$ and $\pi^-$ tracks**
  - Scale Cherenkov angle to remove momentum dependence, $\theta_{cher}^{scaled} = \mathrm{acos}(\beta \cos(\theta_{cher}))$
  - $\beta$ from reconstructed momentum and assuming pion mass
- Final objective: $\boldsymbol{M_{align}^{combined} = M_{align}^{\pi^+} + M_{align}^{\pi^-}}$



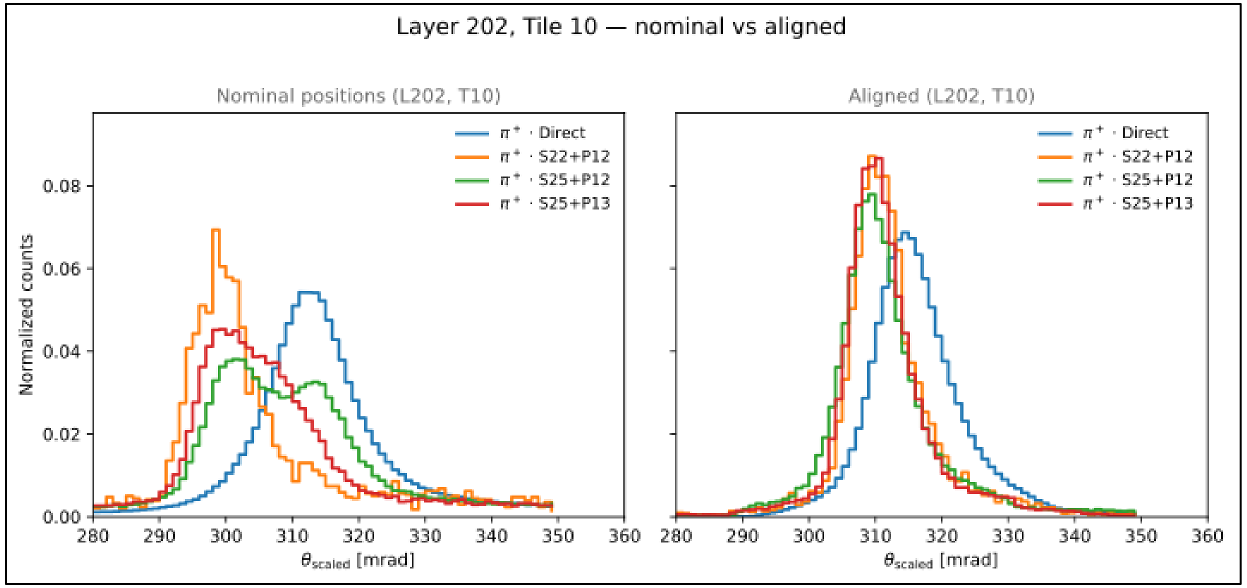
# Alignment workflow

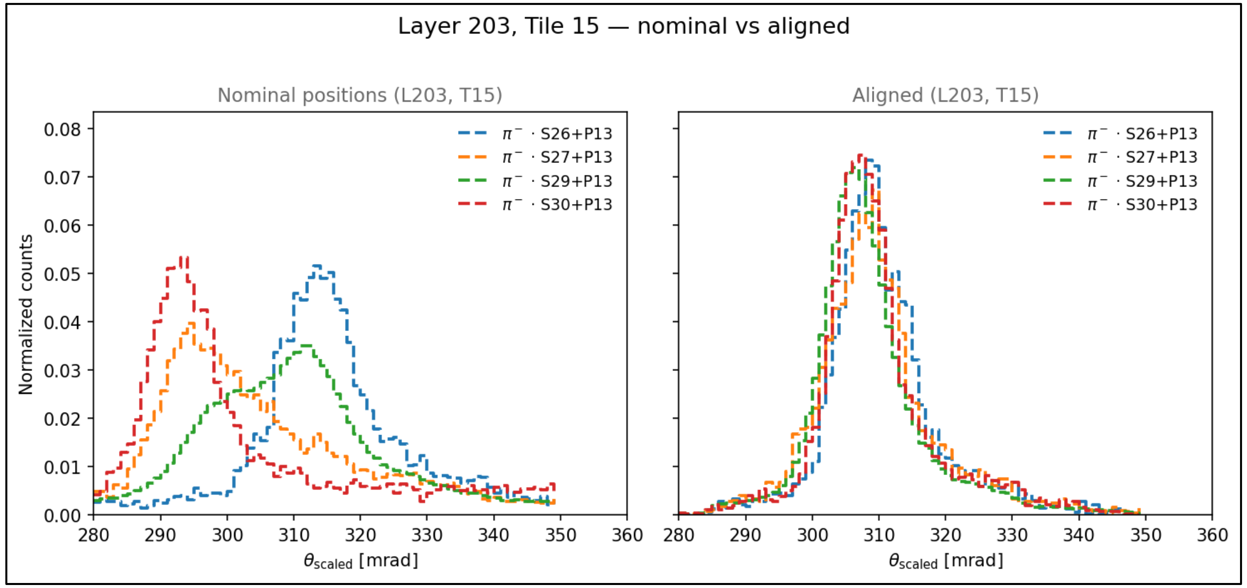
- **Use TuRBO-1 to find alignment parameters that minimize $M_{align}^{combined}$**
- Approach developed as an application of the **AID2E framework** (AI-assisted detector design for the EIC), arxiv:2405.16279
- Reducing the size of the search space by determining only the dominant alignment DOF $dz, d\theta_x, d\theta_y$
  - 54 parameters total determined simultaneously
- ~2000 total trials to convergence (27 trials per batch, run in parallel)
  - 300 pseudorandom initialization trials
  - ~1.5-2 hours per trial (single-threaded re-cooking and analysis)
  - Convergence reached when Trust Region shrinks to set minimum
- Results shown are from the alignment of sector 1 (module 2) for the RGK dataset



# Resulting photon reflection path agreement



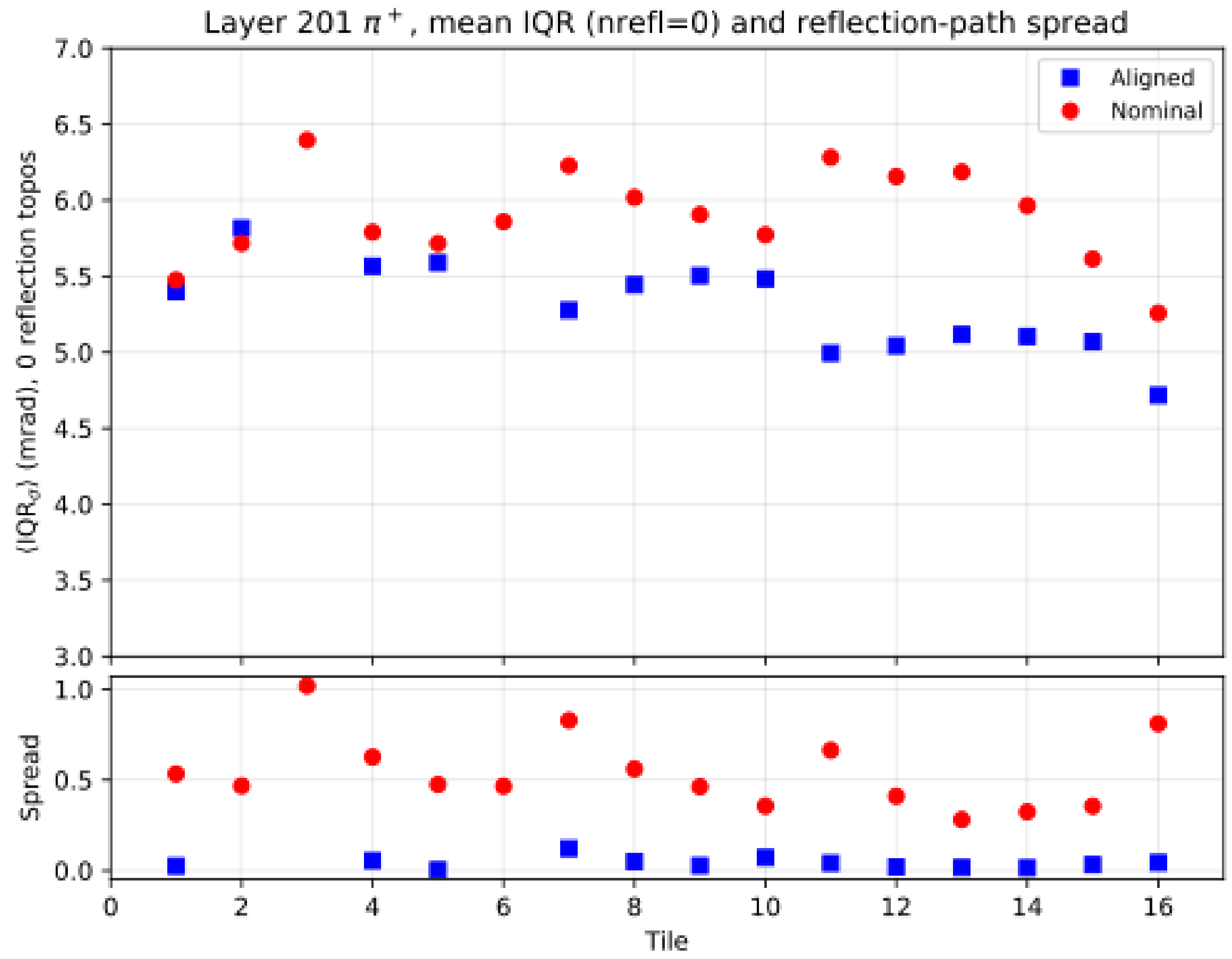

- Comparing scaled Cherenkov angle for pions from different topologies with larger validation dataset
- Simultaneous alignment produces desired agreement between photon reflection paths
  - For both inbending (left) and outbending (right) tracks
- **First alignment of 2*3cm aerogel layer and upper spherical mirrors! (right)**

# Alignment validation, aerogel layer 201

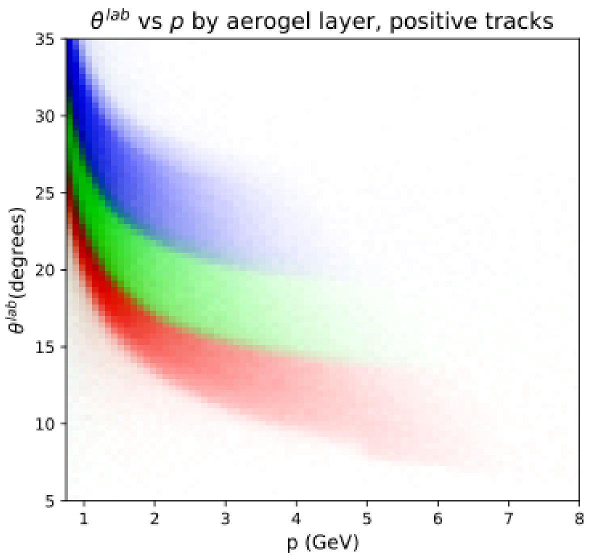
**$\langle IQR \rangle$ for 0 reflection photons**

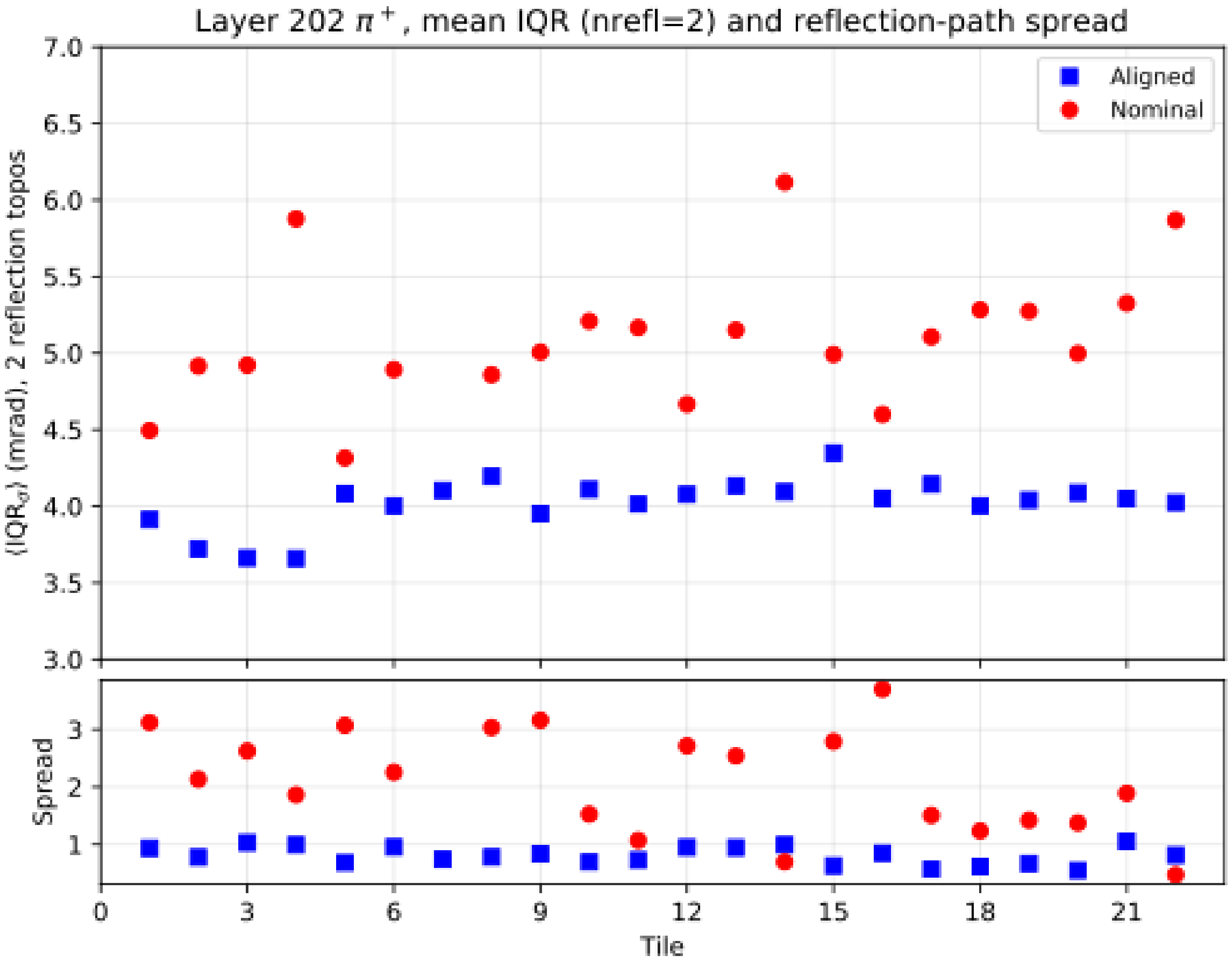
**Topology spread metric**

- $\pi^+$(inbending), first 2cm layer:
  - Good agreement between topologies (0 and 1 reflections from this layer)
  - Improved resolution for direct photons



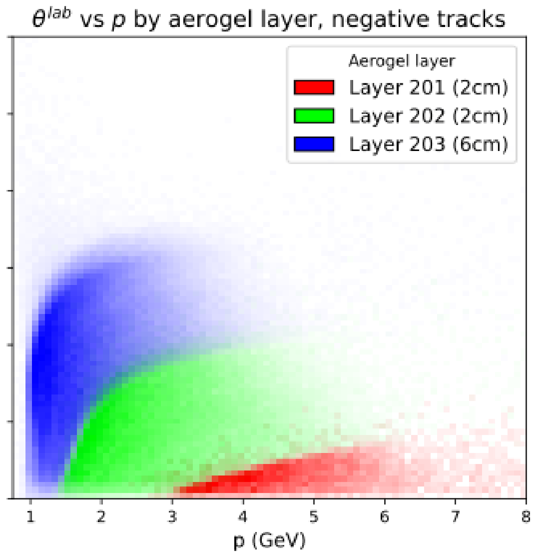
# Alignment validation, aerogel layer 202

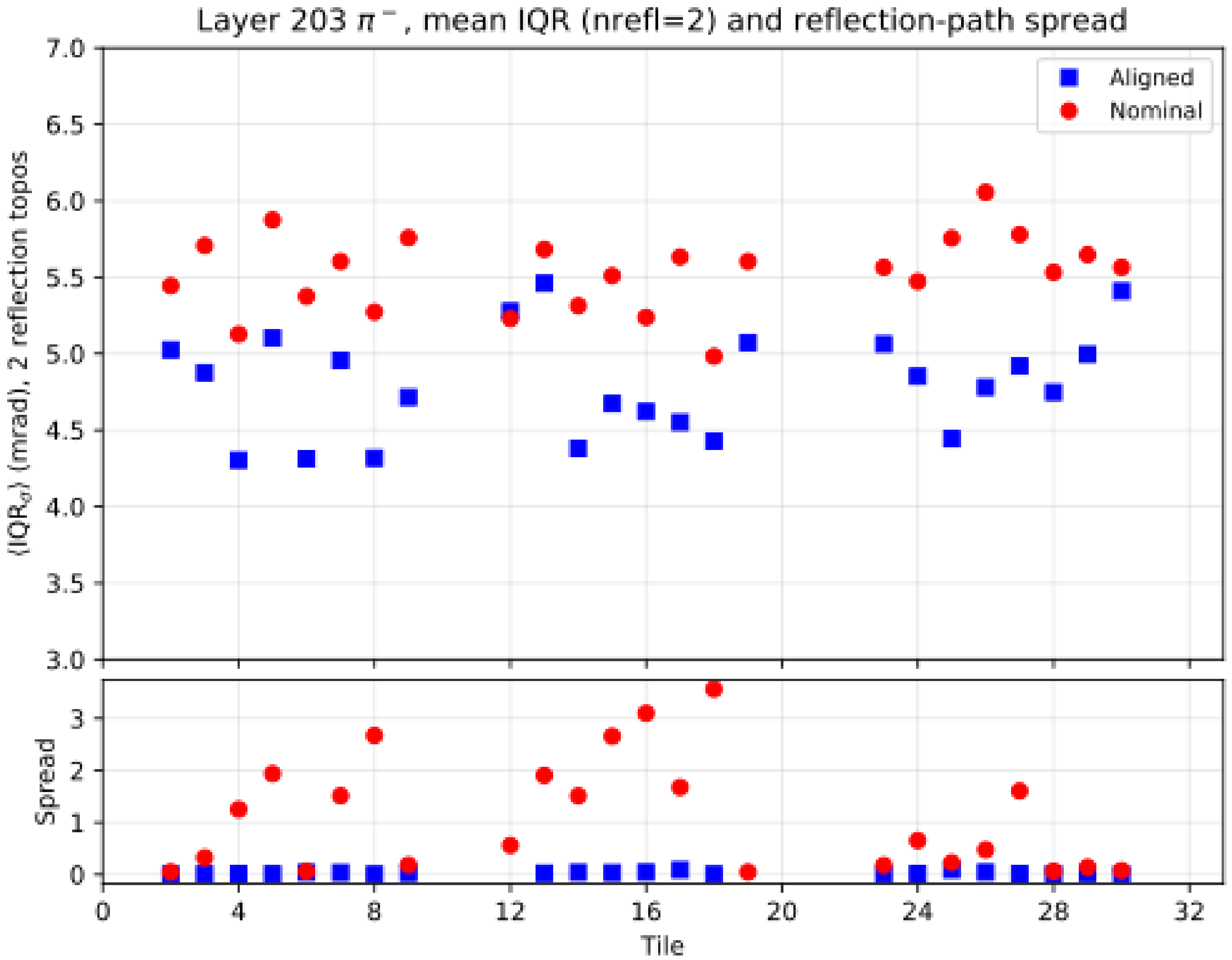



- $\pi^+$(inbending), second 2cm layer:
  - For this layer, have all of 0, 1, and 2+ reflection topologies
  - Performance stable across aerogel tiles, resolution significantly improved (shown 2 reflection resolution)

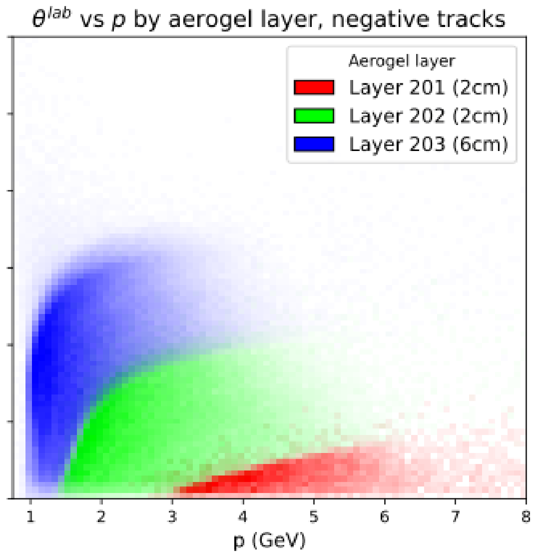
# Alignment validation, aerogel layer 203

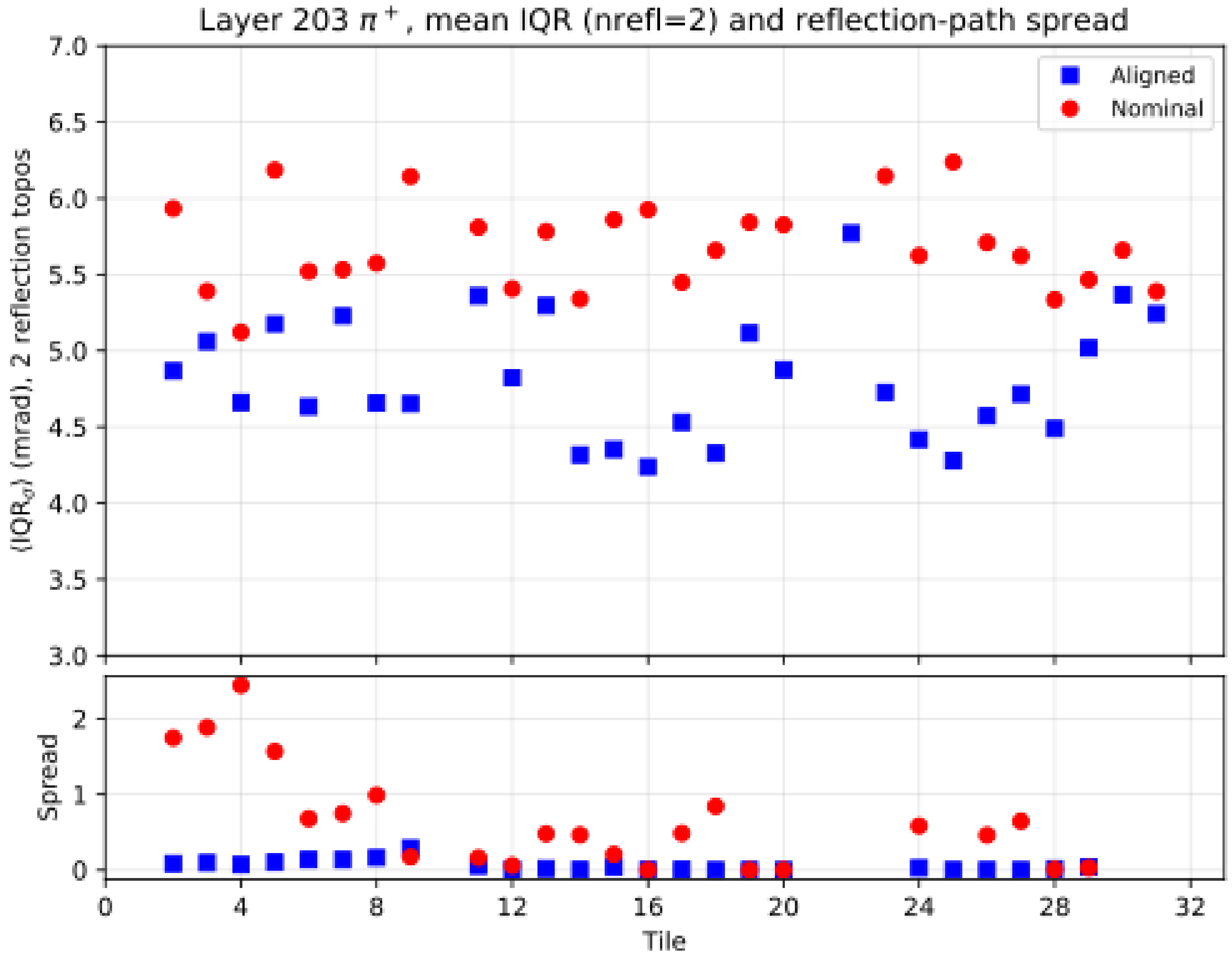



- $\pi^-$(outbending), 6cm layer:
  - For this layer, dominated by 2+ reflection topologies, mirrors correlated
  - First successful alignment of 6cm aerogel layer and upper spherical mirrors!

# Alignment validation, aerogel layer 203




- $\pi^+$(inbending), 6cm layer:
  - For this layer, dominated by 2+ reflection topologies, mirrors correlated
  - First successful alignment of 6cm aerogel layer and upper spherical mirrors!

# Deployment and computing needs

- Approach developed as part of the AID2E framework (AI-assisted Detector Desing for the EIC, arxiv:2405.16279)
  - https://github.com/aid2e/CLAS12-RICH-Align
- Framework currently set up and run on Duke compute cluster with containerized reconstruction
  - Slurm-based batch system, **minimal delay between job submission and start times**
- Each batch of 27 sets of alignment parameters evaluated as independent, single-core jobs
  - Each trial approximately 1-2 hours
- One long-running job needed for trial monitoring, result retrieval, GP fit and acquisition function evaluation
- On average, convergence after about 2000 trials
  - ~4 days with trials evaluated in parallel
- Total compute required: ~2000-4000 CPU hours
- **Currently not feasible to run in full on ifarm**

# Summary

- First simultaneous alignment of all optical components of the RICH has been achieved using TuRBO
  - Now, performance stable for larger polar angle optical components
  - Global alignment now also carried out with this approach as an independent first step
- Some remaining studies to be done (impact of event selection criteria, possible improvements to objective, misalignment in gemc...)
- Alignment successful with $\pi^{\pm}$ for RGK/RGE (figure shown)

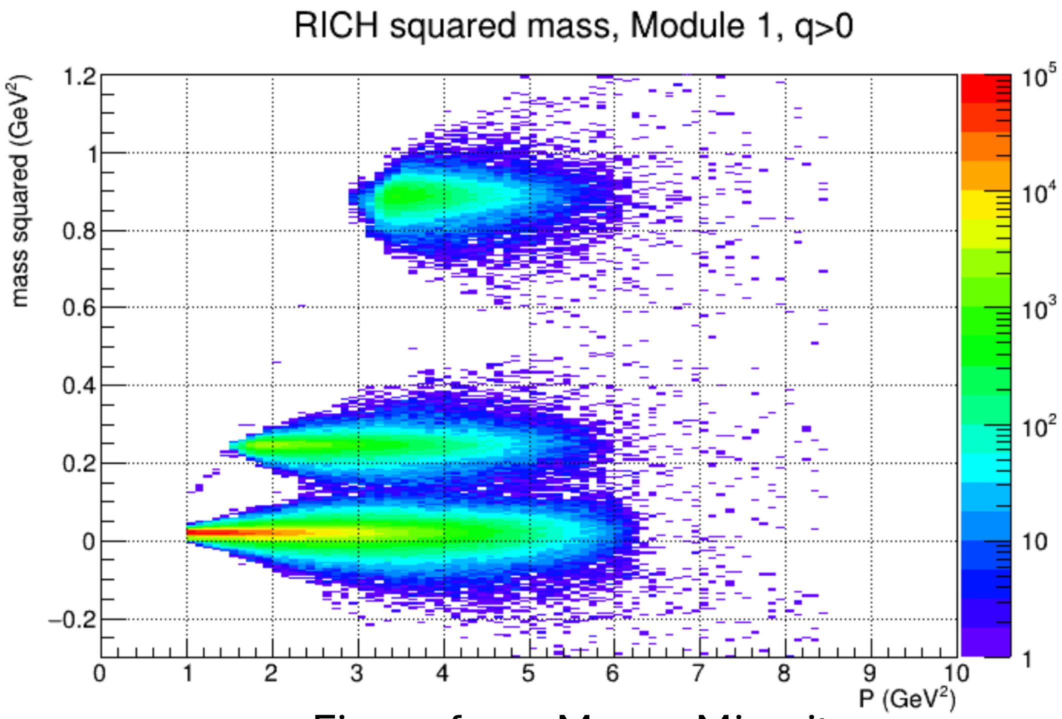



Figure from Marco Mirazita