ePIC Barrel HCAL energy calibration using Machine Learning

Churamani Paudel, David Ruth

(New Mexico State University)

08/14/2026

Neutron and Cluster Selection Criterias

Simulation Campaign:26.05.0

Truth Particle

pdgPar == 2112

1.0 <= ePar < 10.0 GeV

abs(etaPar) < 1.0

BHCal cluster quality

abs(etaBHCalCluster) < 1.0

nClustBHCal >= 1

nHitsBHCalCluster >= 2

eBHCalCone03 >= 0.1 GeV,

eScFiCone03>=0.1 GeV

eImageCone03>=0.1 GeV

Truth-to-cluster matching:

DeltaR(truth neutron, BHCal cluster) < 0.30

Truth-association purity:

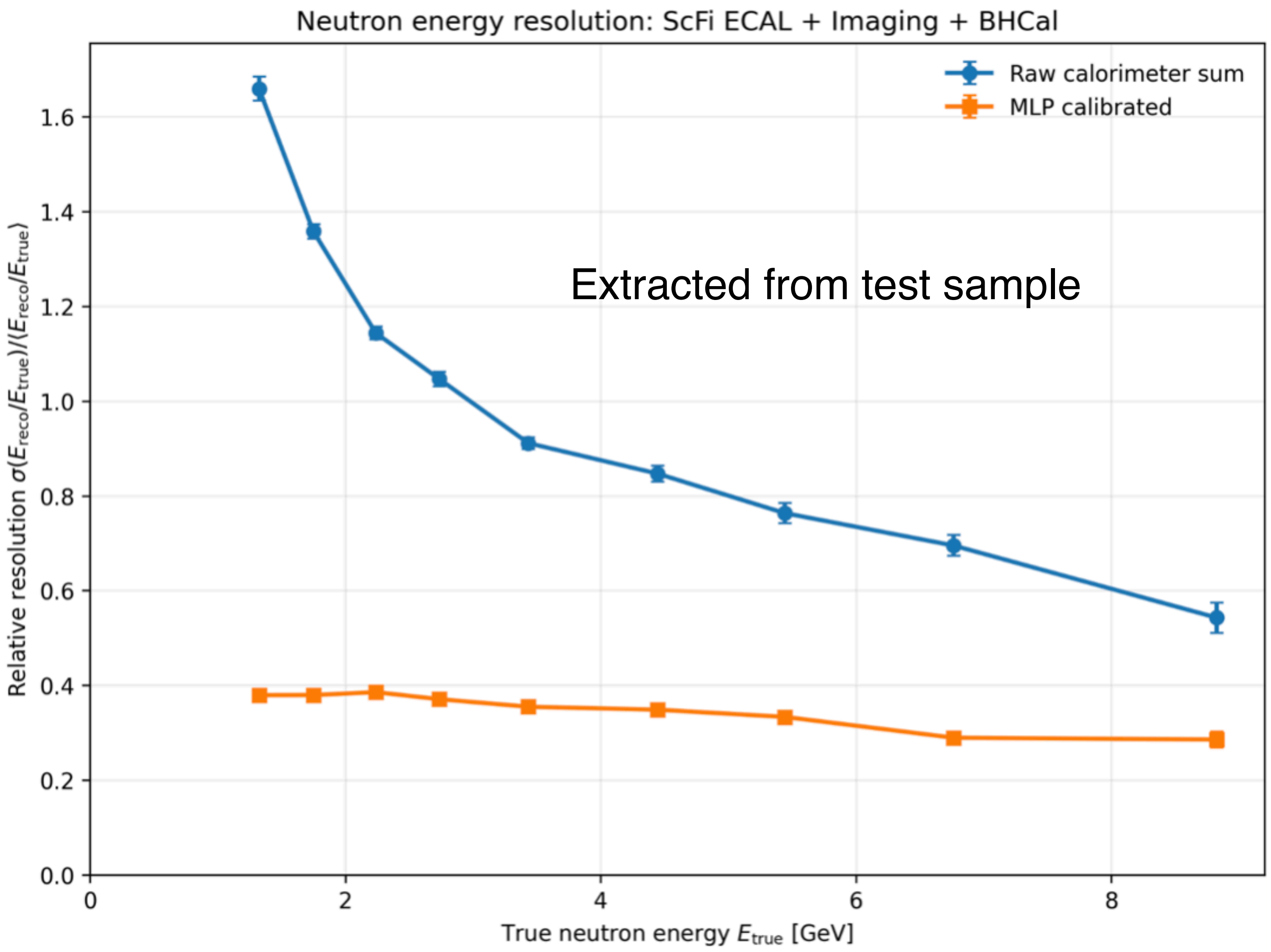
hcalAssocTargetIsLeading == 1
hcalAssocLeadingPDG == 2112
hcalAssocNParticles <= 2

- Select **truth-level neutrons** with PDG code **2112**.
- Require neutron energies between **1 and 10 GeV** and pseudorapidity within **($|\eta| < 1.0$)**.
- Require at least one reconstructed **BHCal cluster** containing **two or more hits**.
- Apply minimum deposited-energy thresholds in the **BHCal, ScFi, and Imaging Calorimeter** systems.
- Match the truth neutron to the BHCal cluster using **($\Delta R < 0.3$)**.
- Retain clusters with a **high-purity truth association** to suppress contamination from unrelated particles.

Motivation: Why is the raw resolution so poor?

WHAT TO SAY

- The full raw ScFi + Imaging + BHCal sum reaches about 166% resolution near 1.3 GeV and remains about 55% near 8.9 GeV.
- The frozen MLP reduces the width to roughly 29–39%, proving that the event information contains useful correction power.
- Goal is not to retrain the model—it is to identify the detector-level cause of the raw width.



Fixed-energy single neutron validation sample

1.0 <= Etrue <= 10.10 GeV

abs(etaTrue) < 1.0

abs(etaBHCalCluster) < 1.0

nClustBHCal >= 1

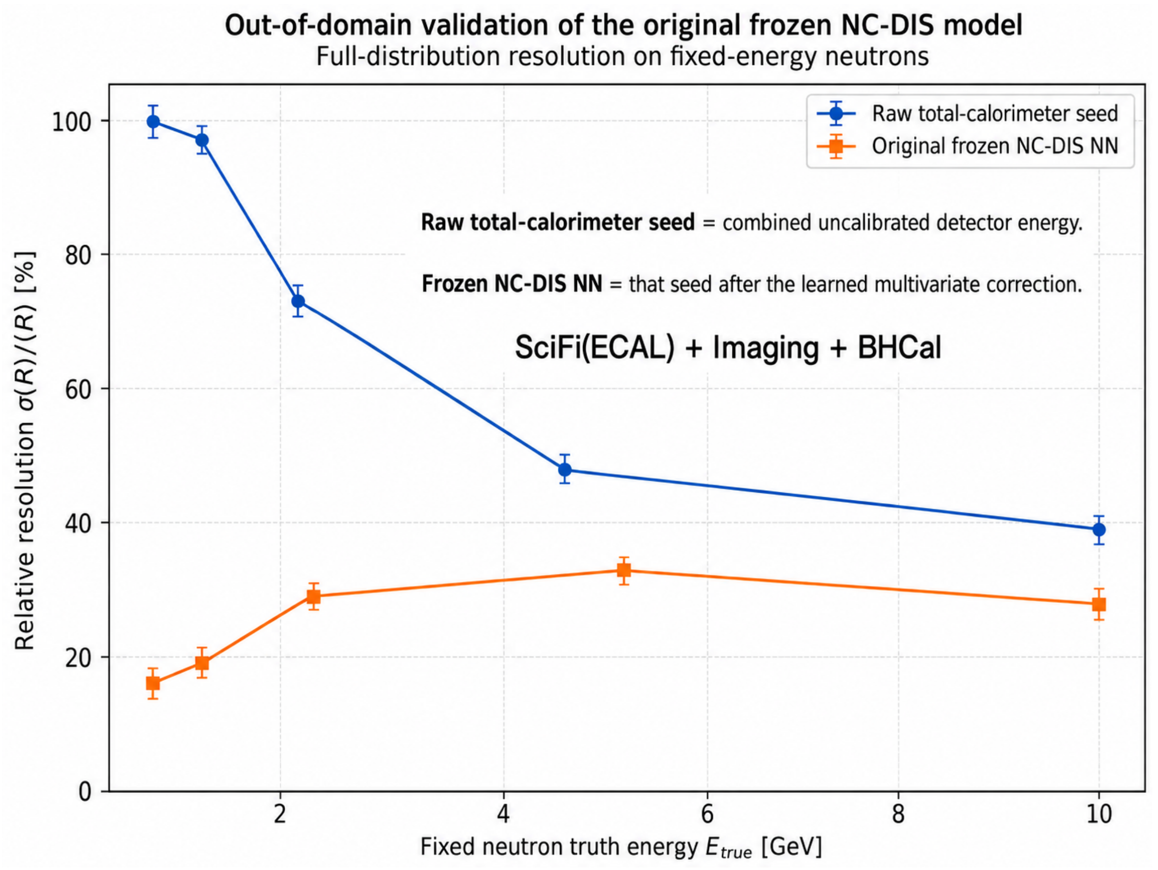
nHitsBHCalCluster >= 2

eBHCalCone03 >= 0.1 GeV

eScFiCone03 >= 0.1 GeV

eImageCone03>=0.1 GeV

DeltaR(truth neutron, BHCal cluster) < 0.30



Raw Energy estimator and role of ML

- Local cone energy in NC-DIS in a complicated DIS-event
- Seed can miss energy outside the cone, also affected by shower overlap and association ambiguity.
- This is what neural-net has to solve.

$$E_{\rm raw}^{\rm NC\,DIS} = E_{\rm ScFi}^{\Delta R<0.30} + E_{\rm Image}^{\Delta R<0.30} + E_{\rm BHCal}^{\Delta R<0.30}$$

$$R_{\rm raw} = \frac{E_{\rm raw}^{\rm NC\,DIS}}{E_{\rm true}}$$

$$E_{NN} = f(x_1, x_2, x_3, x_4, x_5 \ldots\ldots\ldots x_N)$$

$$\frac{\sigma_E}{E} = \frac{\sigma(R)}{\langle R \rangle}$$

What extra variables can indicate:

- Whether energy leaked outside the cone;
- How deeply the shower developed;
- Whether the shower split into several clusters;
- How much energy appeared in different calorimeter layers;
- Whether the matched cluster was contaminated by another particle;
- Whether a low raw energy is truly a low-energy neutron or a poor contained high-energy neutron.

The neural network recovers the low-response population and narrows the response

- The raw ScFi + Imaging + BHCal response is strongly non-Gaussian, with most events far below the ideal response R=1.
- The raw distribution also contains a long high-response tail extending well beyond R=2, producing a very large event-by-event width.
- The frozen neural network shifts the main population toward R=1 and substantially compresses both the low- and high-response tails.
- The NN distribution is much narrower, but it still peaks slightly below unity and retains a residual high-response tail.
- This improvement shows that event-level detector information can distinguish poorly contained events and apply different corrections; a single global scale factor could not produce this shape change.

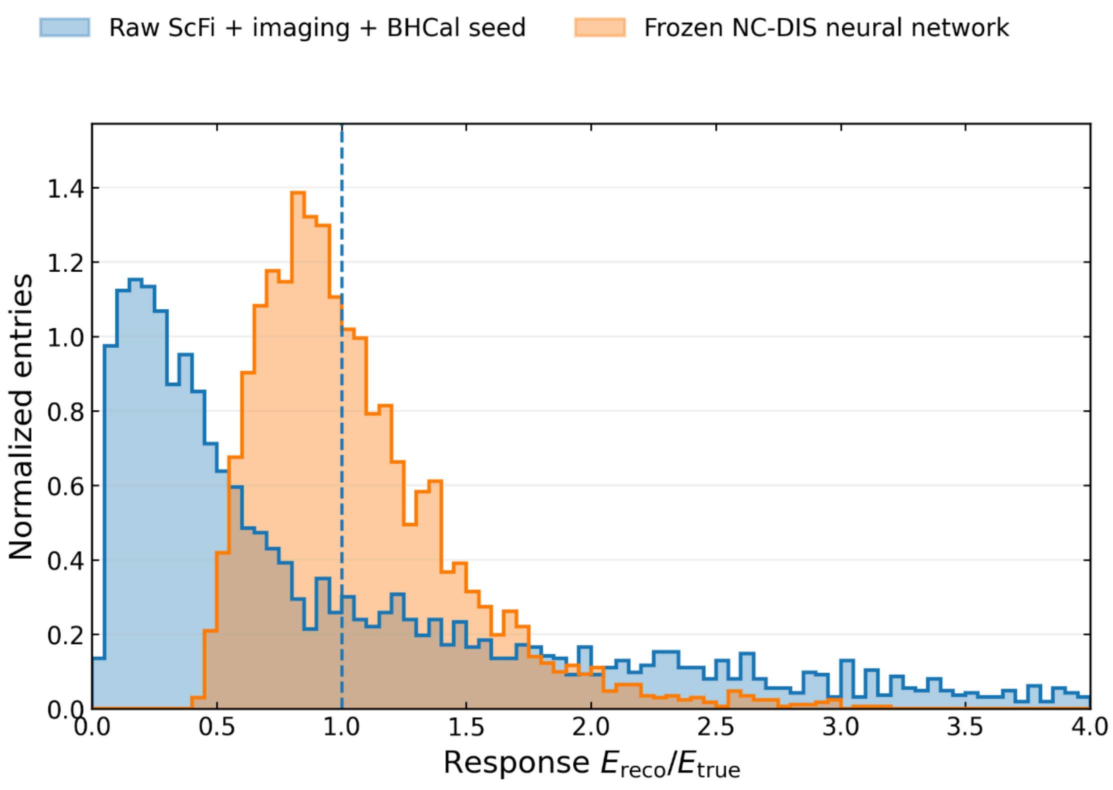


Figure 10: Raw and calibrated response distributions for $2.0 \leq E_{\text{true}} < 2.5$ GeV. The dashed vertical line marks ideal response; all finite events are retained in the numerical moments, while the visible histogram range is $0 < R < 4$. The mean true energy in this bin is 2.238 GeV.

The large raw width comes from strong event-dependent under- and over-response. The neural network improves the resolution by reshaping the distribution, not simply by rescaling its mean.

Analysis Roadmap:From Topology Studies to Truth-Level Diagnostics

WHAT TO SAY

- The study uses 17,344 moderate-clean neutron rows with $\Delta R < 0.30$ and N_assoc ≤ 2.
- Raw response is R = E_seed / E_true with the complete selected distribution included in the resolution.
- We move from reconstructed topology to cone tests, exact SimHits, and finally a truth-axis oracle study.

Input sample

17,344 moderate-clean neutron rows
$\Delta R < 0.30$, N_assoc ≤ 2

Raw estimator

E_seed = E_ScFi + E_Image + E_BHCal
R = E_seed / E_true

Metric

Full-distribution $\sigma(R)/\langle R \rangle$
No Gaussian core fit

Investigation steps

Topology and tails

↓

Cone size, geometry, contamination

↓

Exact target-lineage SimHits

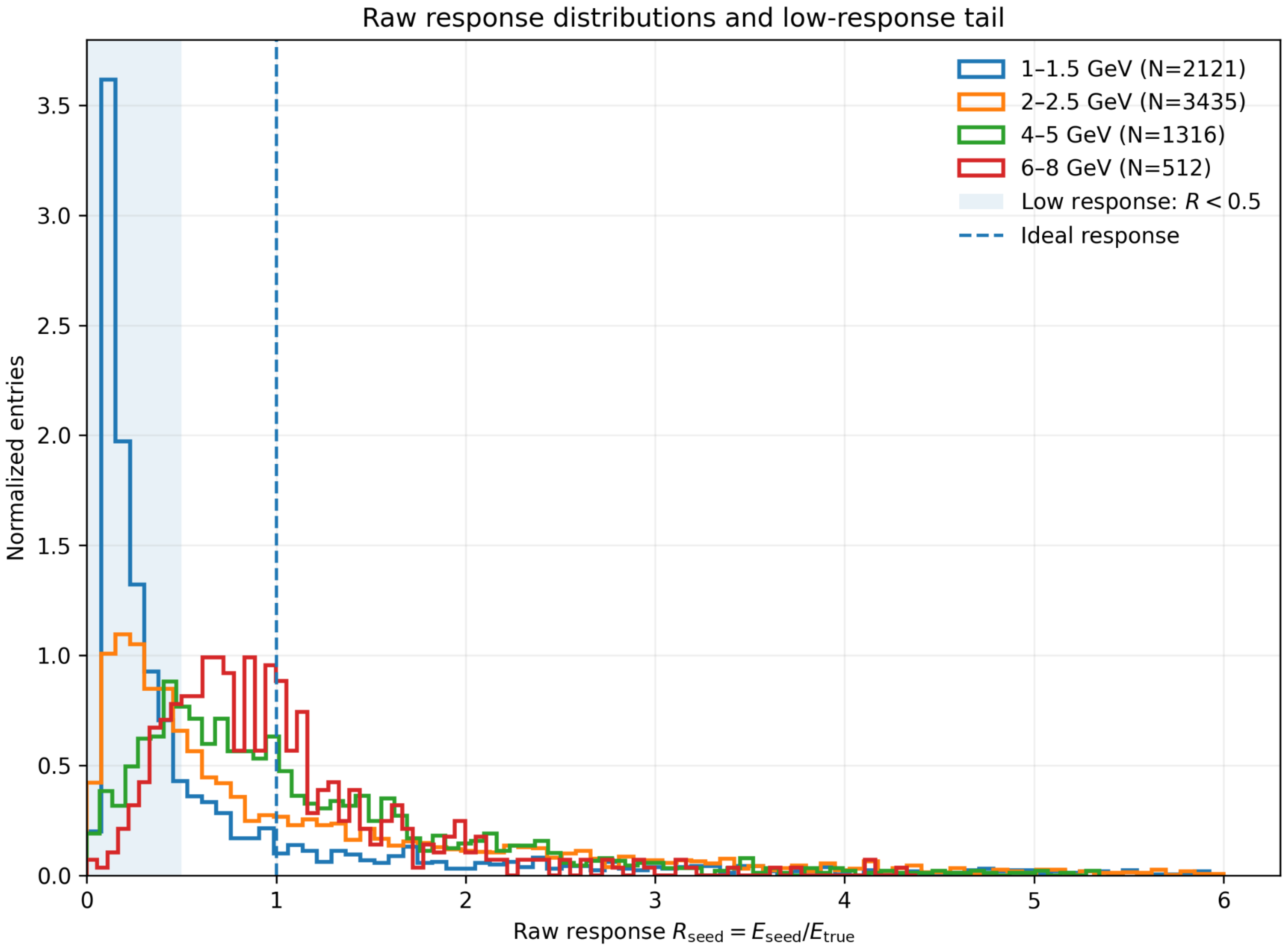
↓

Truth-shower axis oracle test

↓

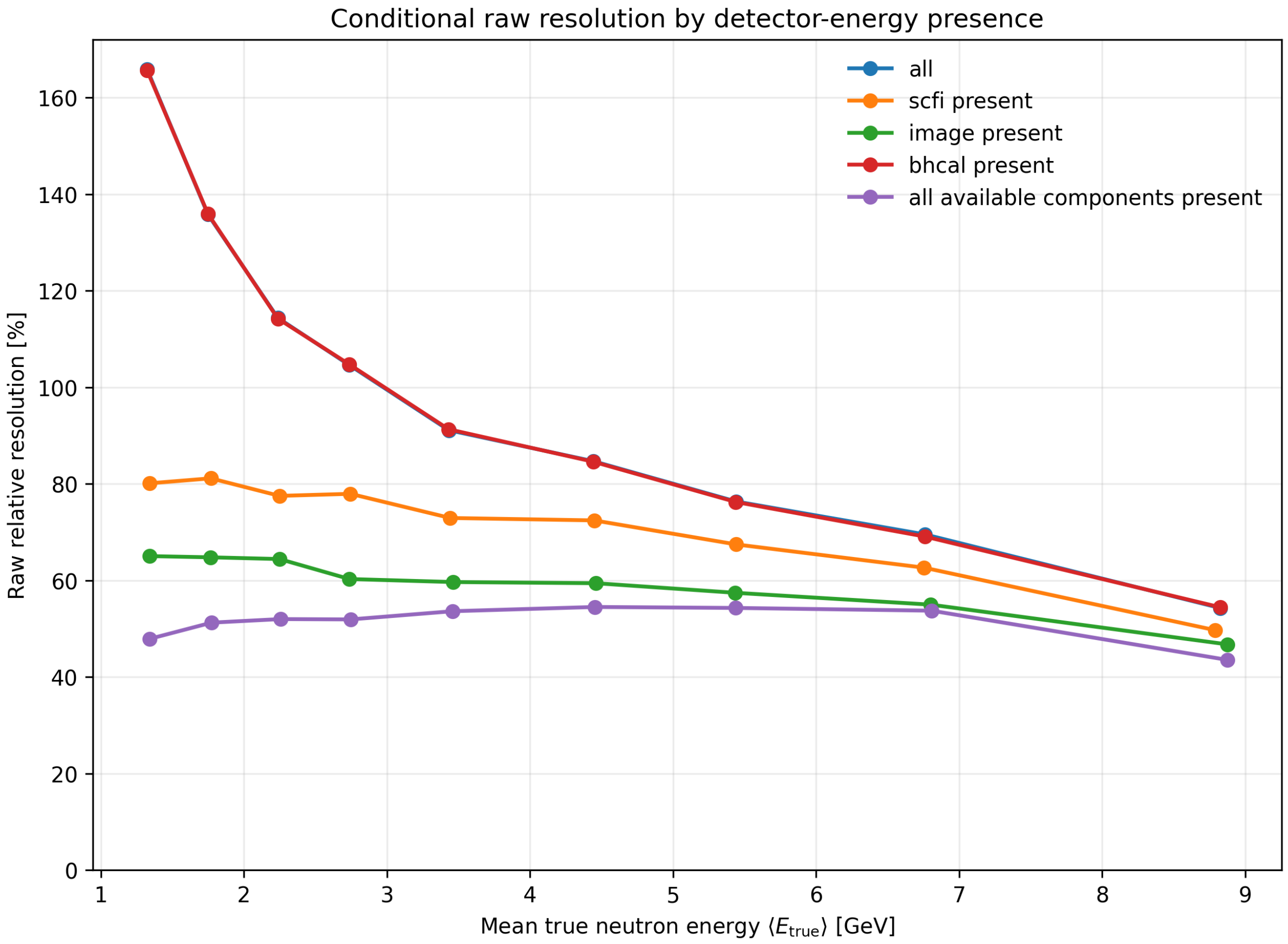
The Raw Response distributions reveal strong non-Gaussian behavior

- Representative energy bins show a large population near very small response and a long high-response tail.
- A single mean or Gaussian-core fit cannot describe the event-by-event response.
- We classify events as low ($R < 0.5$), normal ($0.5 \le R \le 1.5$), or high ($R > 1.5$).



Upstream detector activity identifies better-measured events

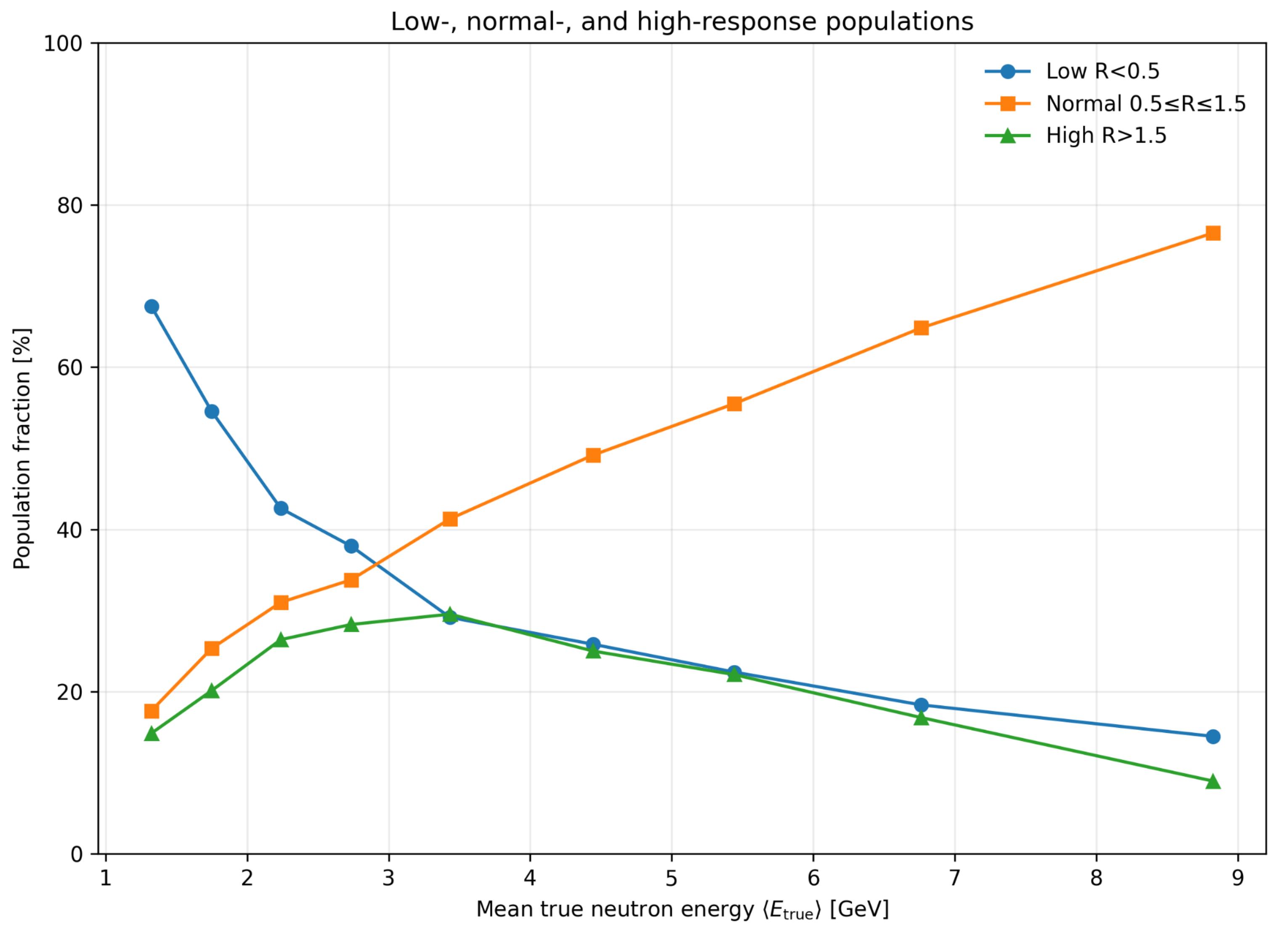
- The inclusive resolution and the **BHCal-present** resolution are nearly identical, showing that BHCal energy is present in almost every selected event and does not isolate the broad-response population.
- Requiring **ScFi activity** reduces the low-energy resolution from about 166% to approximately 80%.
- Requiring **Imaging activity** gives a larger improvement, reducing the low-energy resolution to roughly 65%.
- Events with **all available calorimeter components present** have the narrowest conditional response, approximately 48–55% across most of the energy range.
- The differences become smaller at high energy, where upstream detector activity is more commonly reconstructed.
- These requirements **select** better-response events; they do not themselves recover the missing energy or solve the reconstruction problem.



Energy dependence driven by strong low response population

WHAT TO SAY

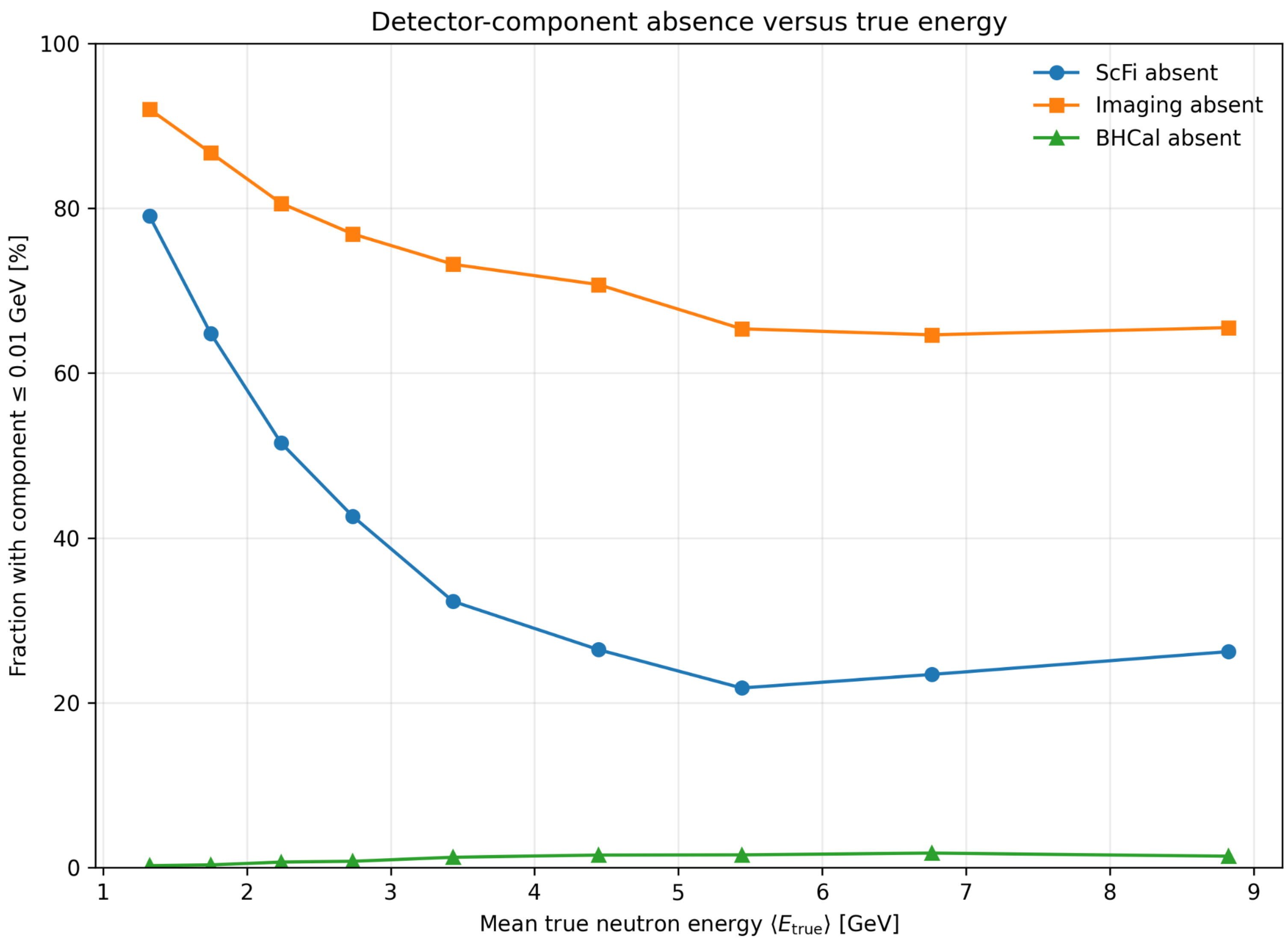
- Near 1.3 GeV, about 68% of events are in the low-response category.
- The low fraction falls steadily with energy, while the normal fraction rises to about 77% near 8.9 GeV.
- The high tail is also substantial, peaking near 30% around 3–4 GeV.



Missing upstream activity explains much of the low tail

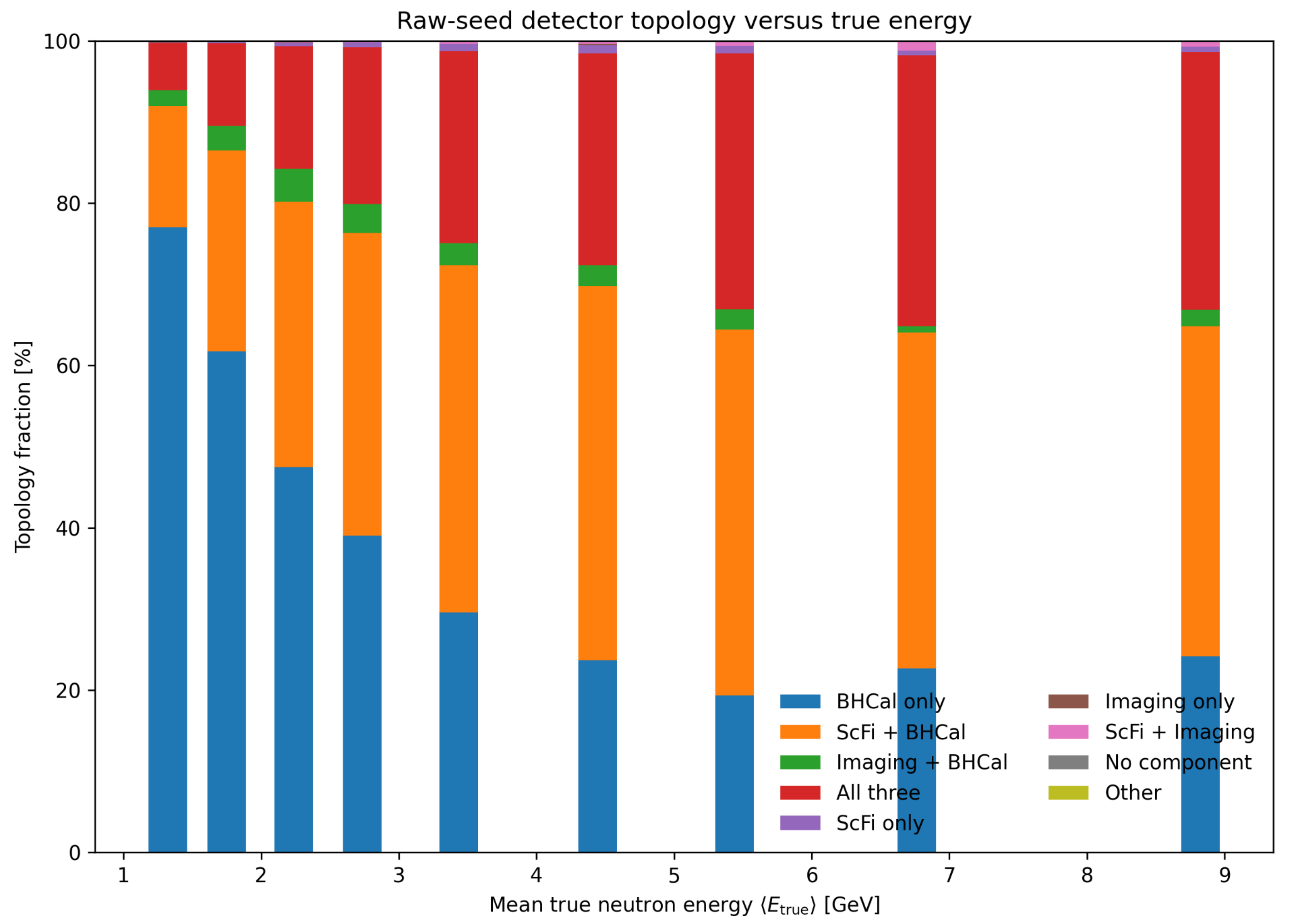
WHAT TO SAY

- BHCal energy is present in almost every selected event, so missing BHCal is not the cause.
- At low energy, ScFi is absent in about 79% of events and Imaging is absent in more than 90%.
- Upstream component presence increases with energy, matching the improvement of the raw resolution.



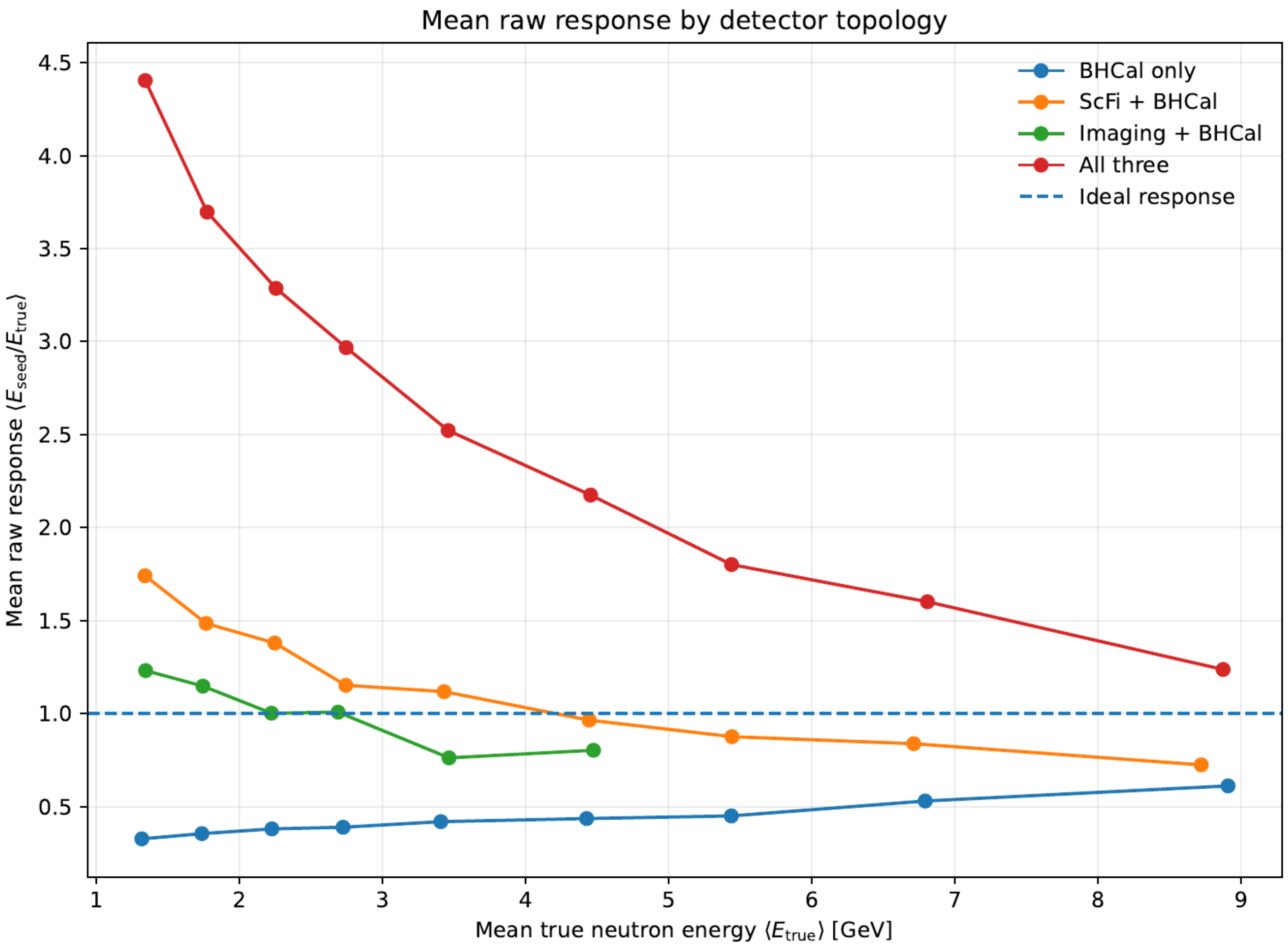
Neutron shower topology evolves with energy

- BHCal-only events dominate the lowest-energy bin at roughly 77%.
- Their fraction drops toward 20–25% as ScFi + BHCal and all-three topologies become more common.
- The inclusive resolution therefore mixes different detector responses in energy-dependent proportions.



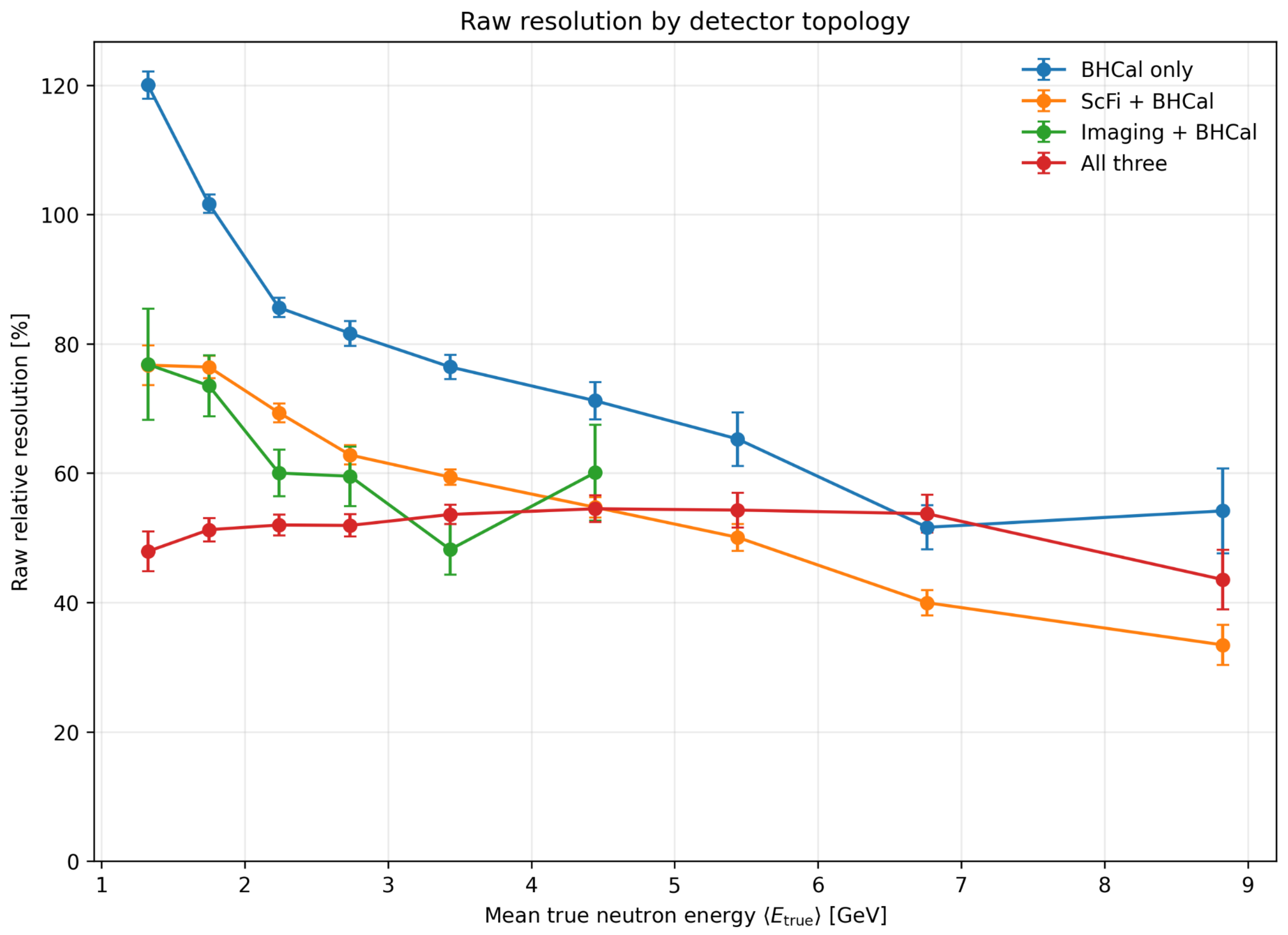
Different topologies exhibit opposite response biases

- BHCal-only events strongly under-respond, with mean response rising only from about 0.3 to 0.6.
- All-three events strongly over-respond at low energy, with mean response near 4.4 at 1.3 GeV.
- ScFi + BHCal transitions from over-response to under-response as energy increases.



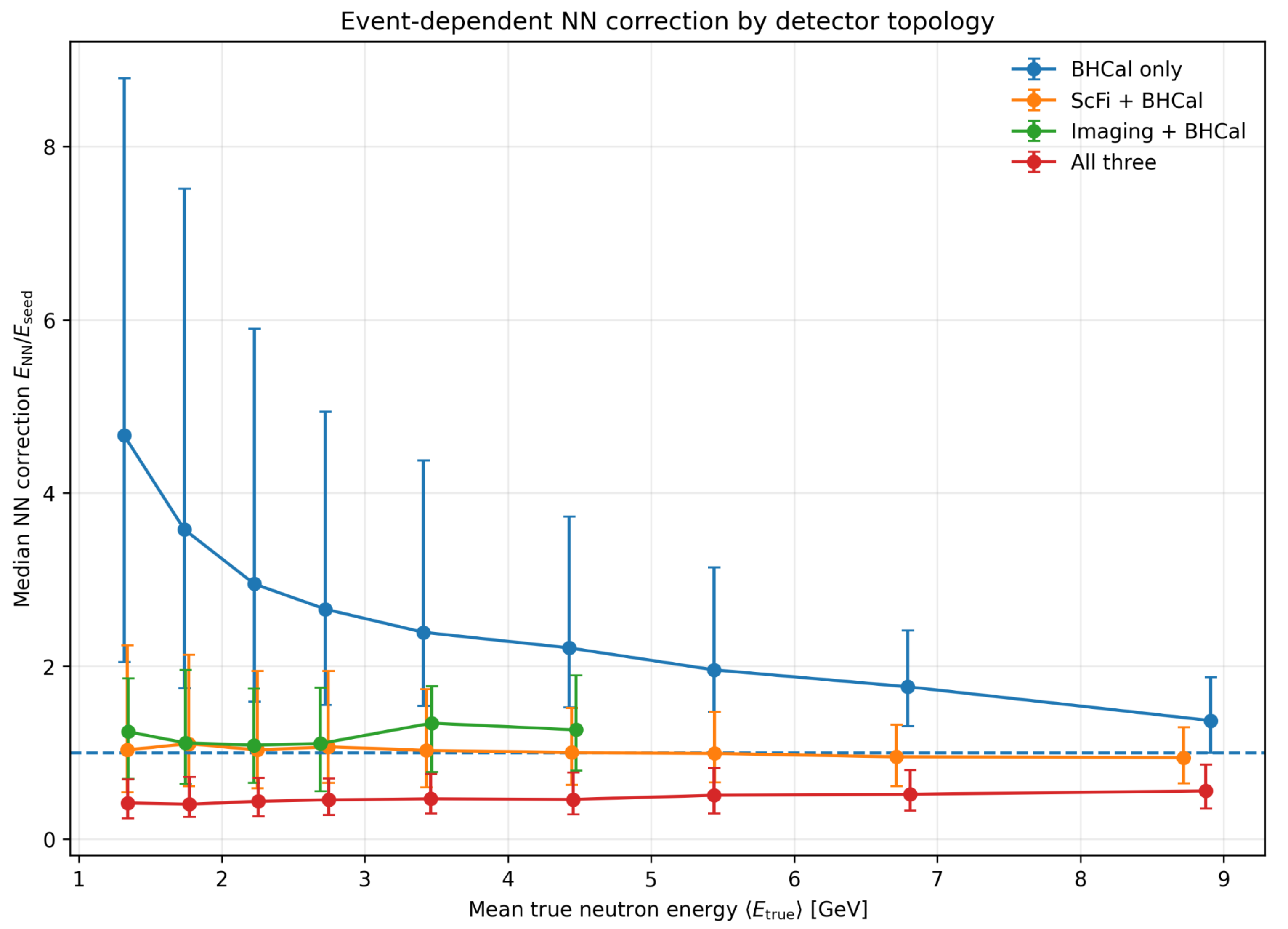
Resolutions and biases must be intrepreted together

- BHCal-only events have the largest relative width: about 120% near 1.3 GeV.
- All-three events have a smaller width near 50%, but their mean response is severely biased high.
- A topology can look “better” in σ/μ while still being badly calibrated.



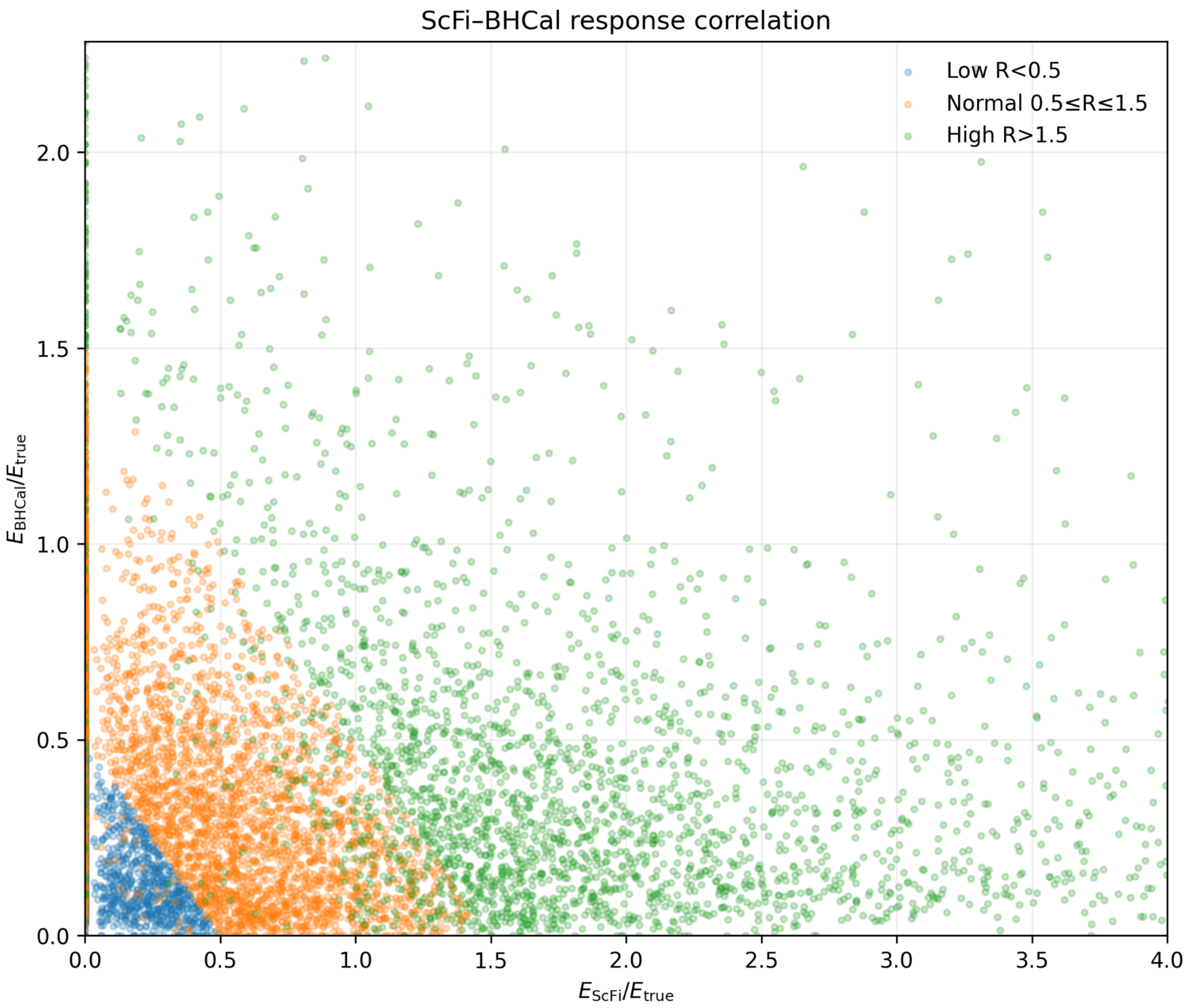
Machine-Learning learns topology dependent corrections

- For BHCal-only events, the median correction factor is very large—about 4.7 at low energy.
- ScFi + BHCal events receive corrections near unity, while all-three events are downscaled to about 0.4–0.6.
- The MLP is learning event-dependent topology and energy corrections rather than one global scale.



Detector Correlations Reveal Origins of Under- and Over-Response

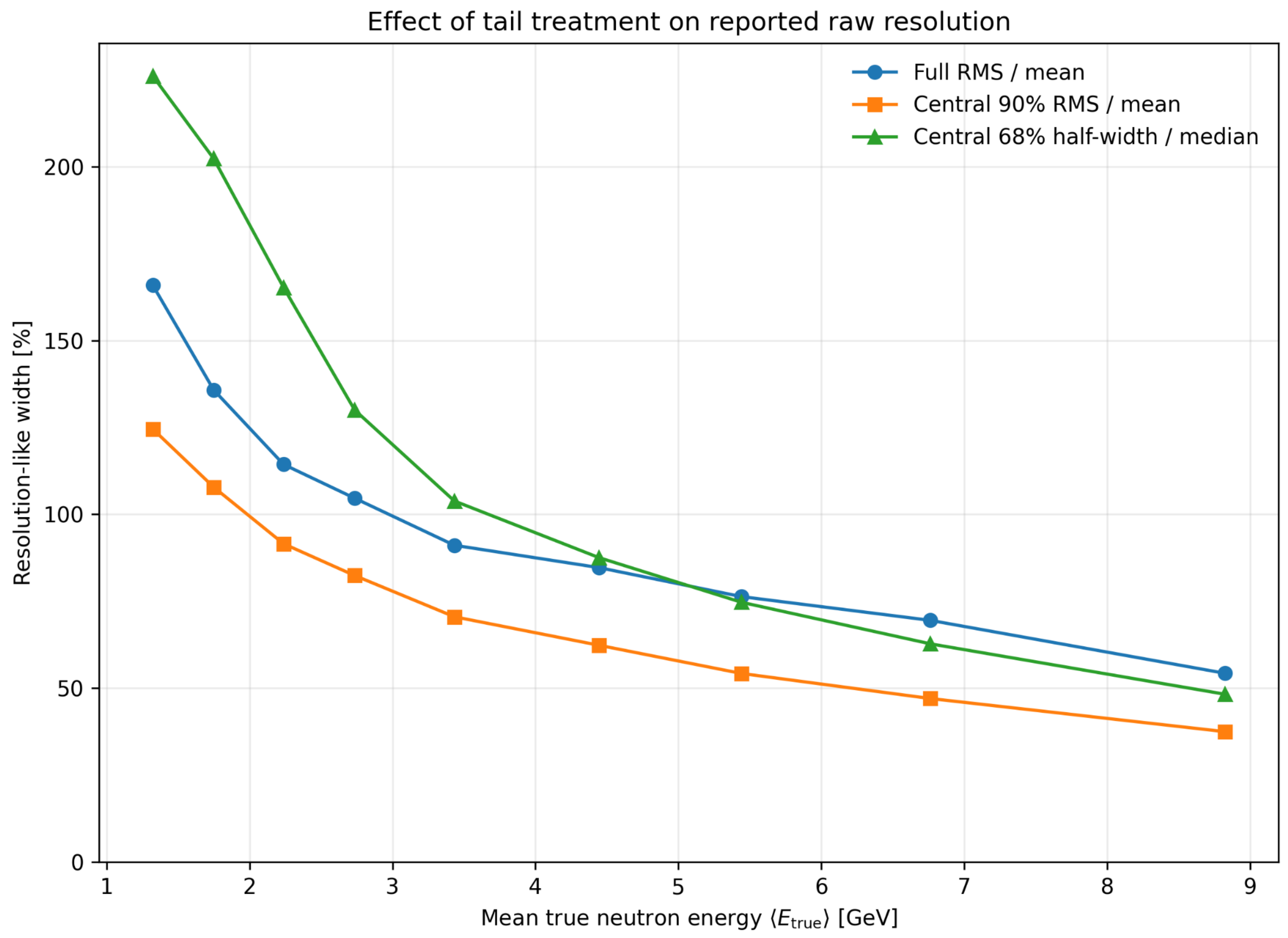
- Low-response events cluster at small ScFi and small BHCal response.
- The high-response population extends mainly toward large ScFi response with a broad BHCal contribution.
- The response is a continuum of detector activity, not a single calibratable line.



The Problem Persists Even After Tail Suppression

WHAT TO SAY

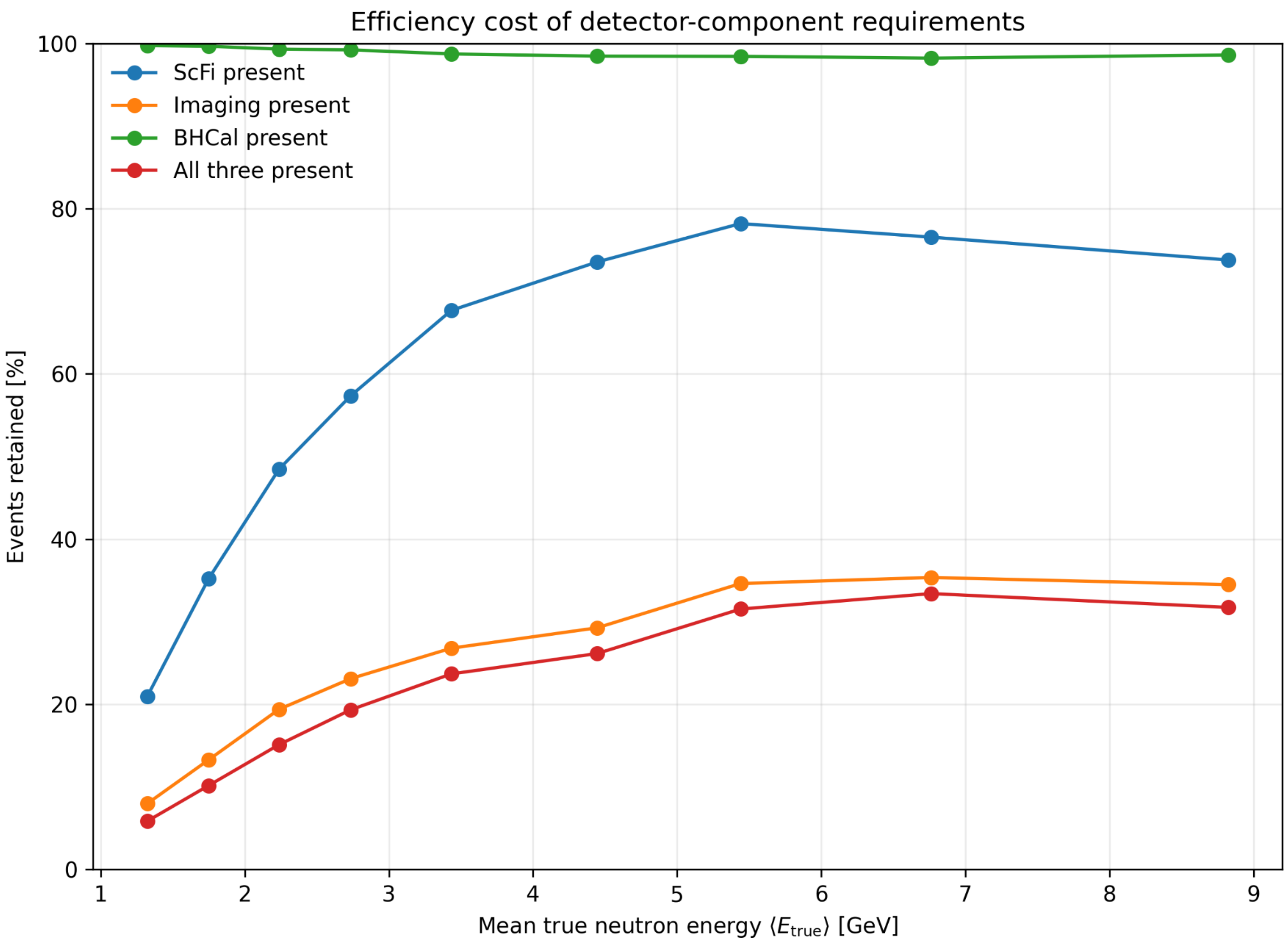
- Removing the outer 10% reduces the reported width, confirming that extreme tails matter.
- Even the central 90% remains very broad, especially below 3 GeV.
- Robust central-width definitions also show a wide non-Gaussian core at low energy.



Resolution Gains from Topology Selection Come at High Efficiency Cost

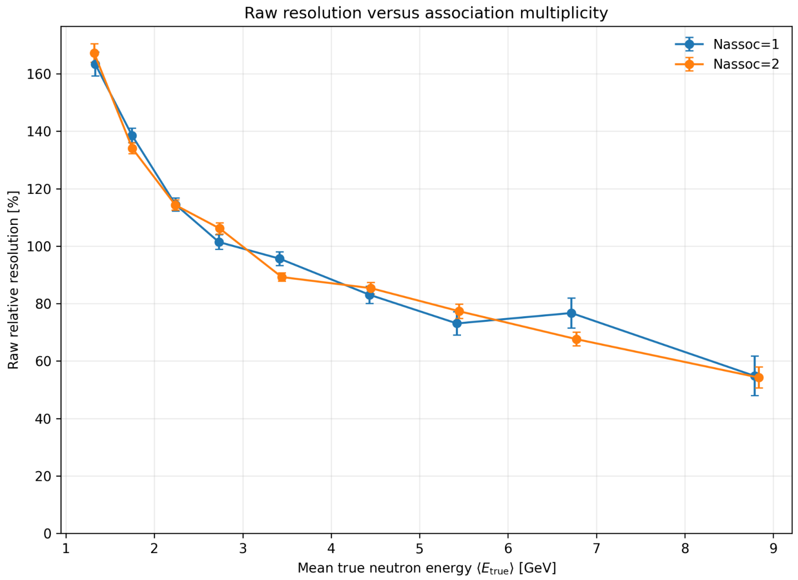
WHAT TO SAY

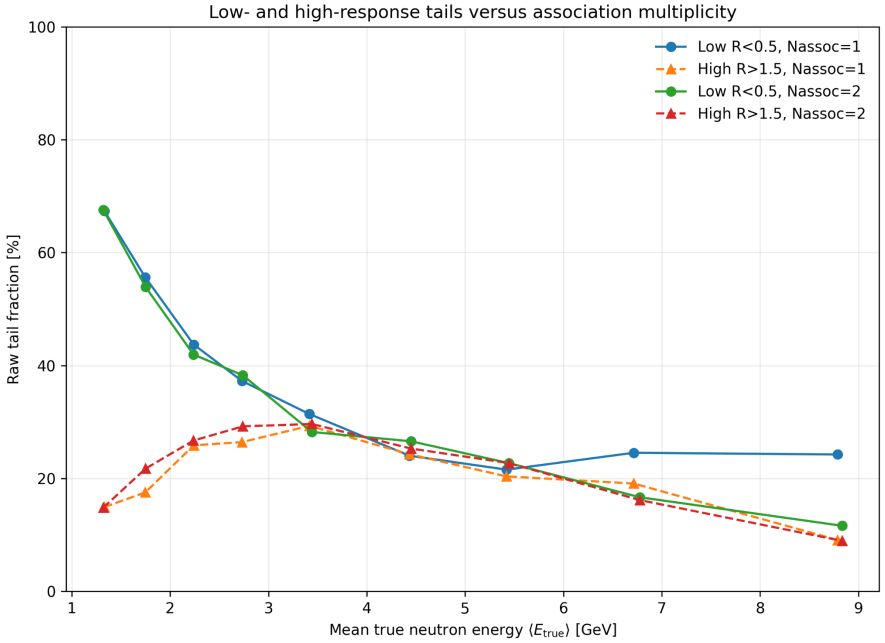
- Requiring ScFi activity retains only about 21% of events near 1.3 GeV.
- Requiring Imaging or all three components keeps only about 6–10% at low energy.
- These cuts improve conditional resolution by discarding the problematic population.



Secondary Associated Particles Are Not the Main Driver

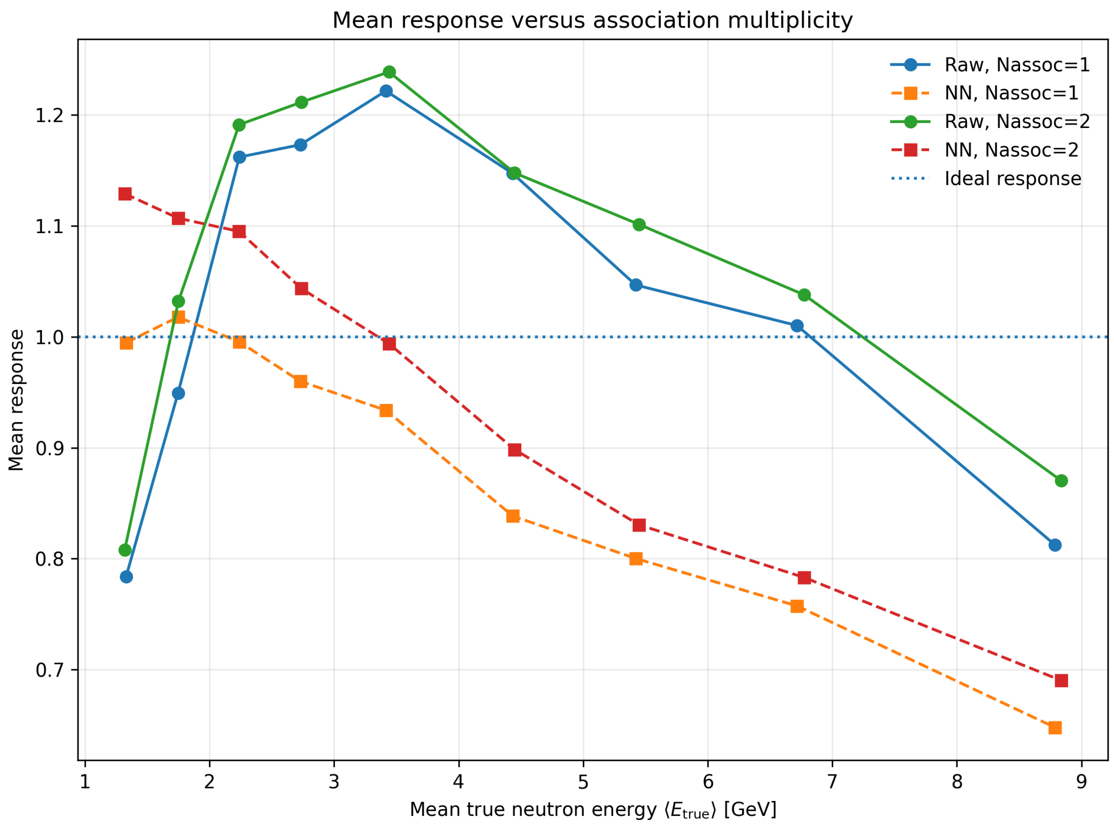
- The resolution curves for one and two associated particles largely overlap.
- Low- and high-tail fractions are also nearly identical over most energies.
- The second associated particle is therefore not required to create the large width.





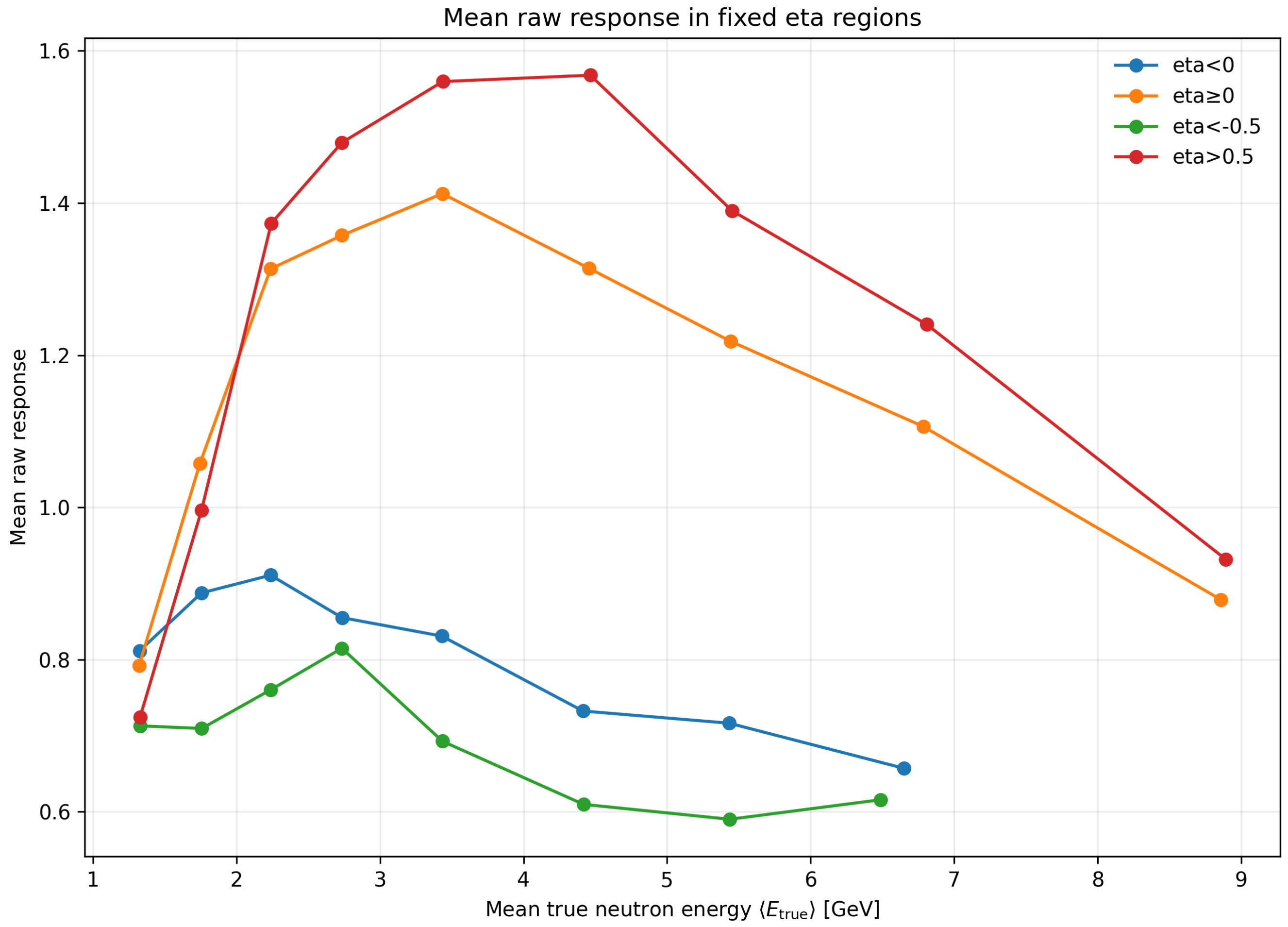
Response Bias Is Largely Independent of Association Multiplicity

- The raw Nassoc =1 and Nassoc =2 samples show nearly the same energy dependence.
- Both samples under-respond at the lowest energy, over-respond around 2–5 GeV, and fall below unity again at high energy.
- Nassoc =2 gives only a modest upward shift relative to Nassoc =1; therefore, the second associated particle does not explain the large raw-response problem.
- The NN response also retains an energy-dependent bias, especially above approximately 3.5 GeV, where both samples increasingly under-respond.



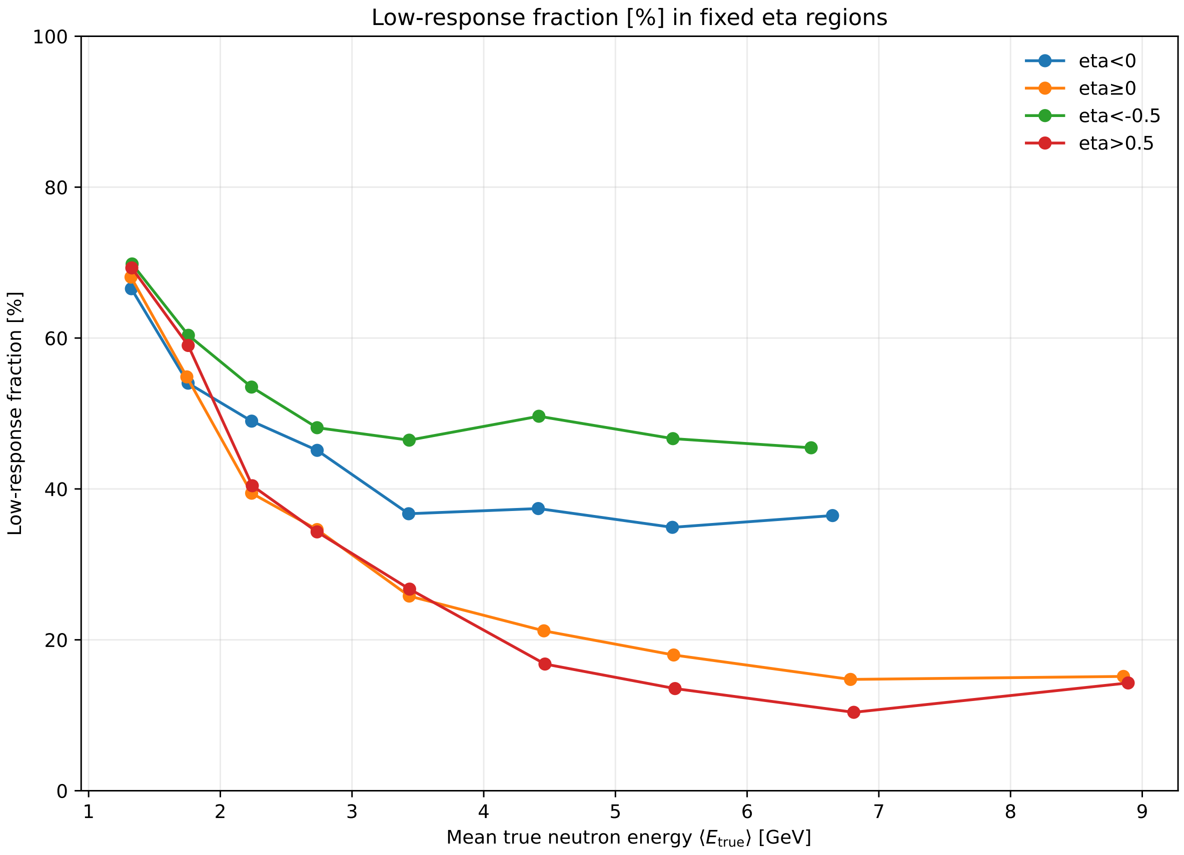
Strong η-Dependent Response Asymmetry

- At the lowest neutron energy, all η regions under-respond, with mean raw response around 0.7–0.8.
- At negative η, the response remains below unity over the full energy range; the strongest under-response occurs for η<–0.5, reaching only about 0.6 at high energy.
- At positive η, the response becomes strongly over-calibrated above approximately 2 GeV, peaking around 1.4–1.6 for Etrue ~3–5 GeV.
- The positive-η over-response decreases at high energy, while the negative-η sample continues to under-respond.
- Therefore, the inclusive raw response combines two opposite detector-region biases rather than one common global calibration offset.



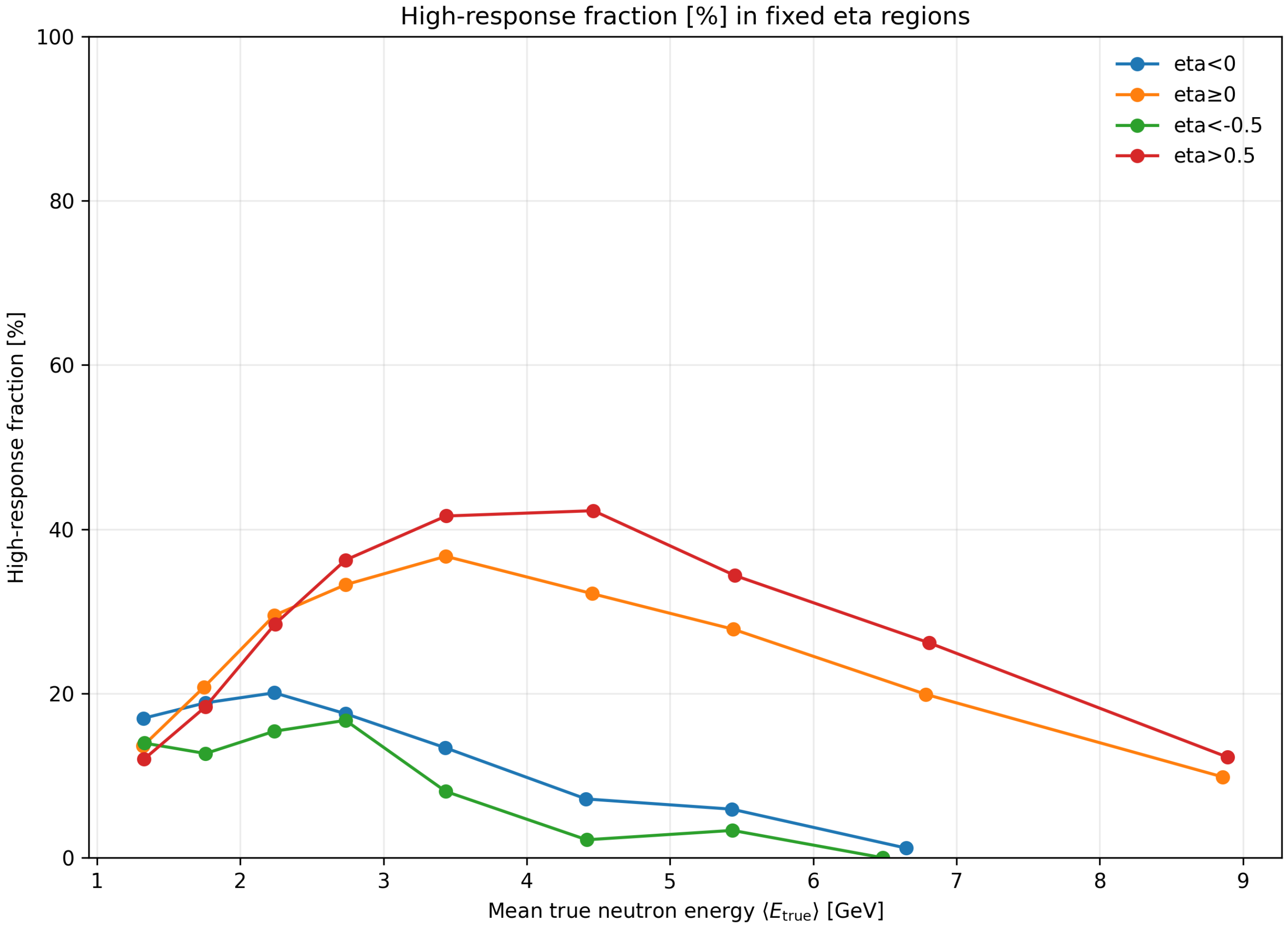
The under-response is concentrated at negative η

- At the lowest neutron energies, the low-response fraction is large in all η regions, reaching approximately 65–70%.
- With increasing energy, the positive-η low-response fraction decreases rapidly, reaching only about 10–20%.
- In the negative-η region, the low-response fraction remains much larger, approximately 35–50% above 3 GeV.
- The strongest persistent under-response occurs for η<–0.5, indicating that the low-energy tail is strongly dependent on detector region.



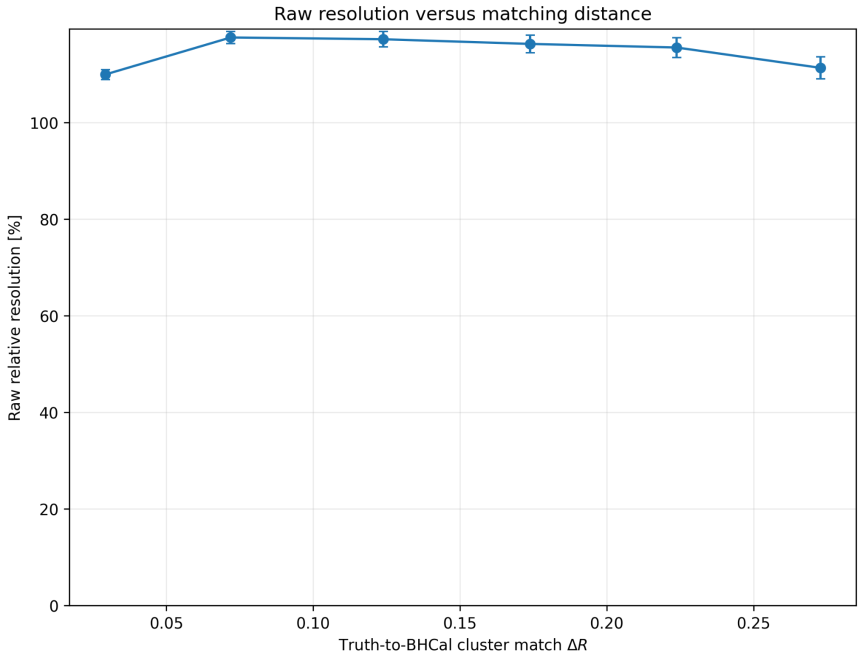
The over-response is concentrated at positive η

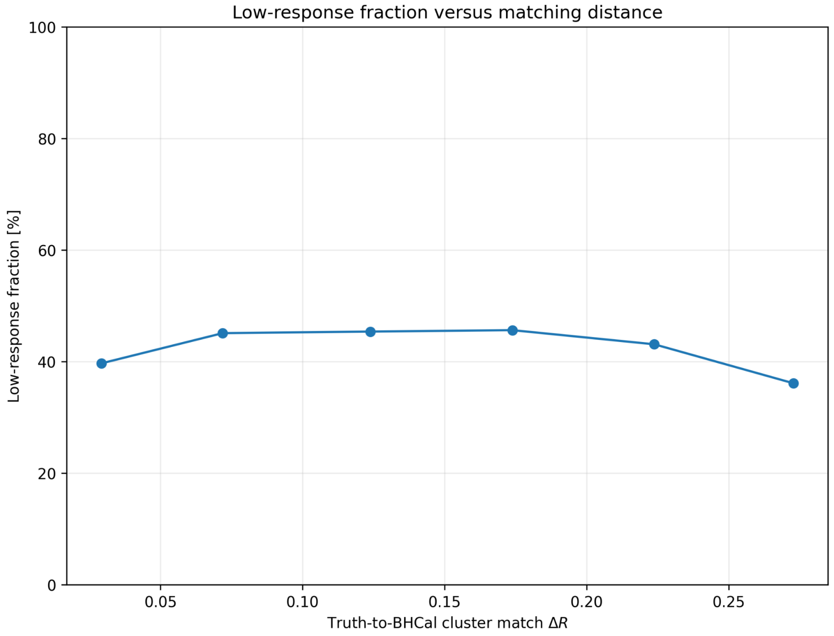
- At the lowest energies, the high-response fractions are similar across the different η regions.
- Above approximately 2 GeV, the positive-η high-response fraction increases strongly, reaching about 35–42% near 3–5 GeV.
- In contrast, the high-response fraction decreases with energy at negative η, approaching zero for η<–0.5 above approximately 6 GeV.
- The opposite behavior of the two tails shows that the inclusive resolution mixes negative-η under-response with positive-η over-response.



Cluster Matching Quality Does Not Explain the Resolution

- The raw resolution remains approximately flat across the tested BHCal matching distance.
- The low-response fraction also stays around 36–46% without a strong monotonic trend.
- Tighter truth-to-BHCal matching therefore does not isolate a good-resolution sample.

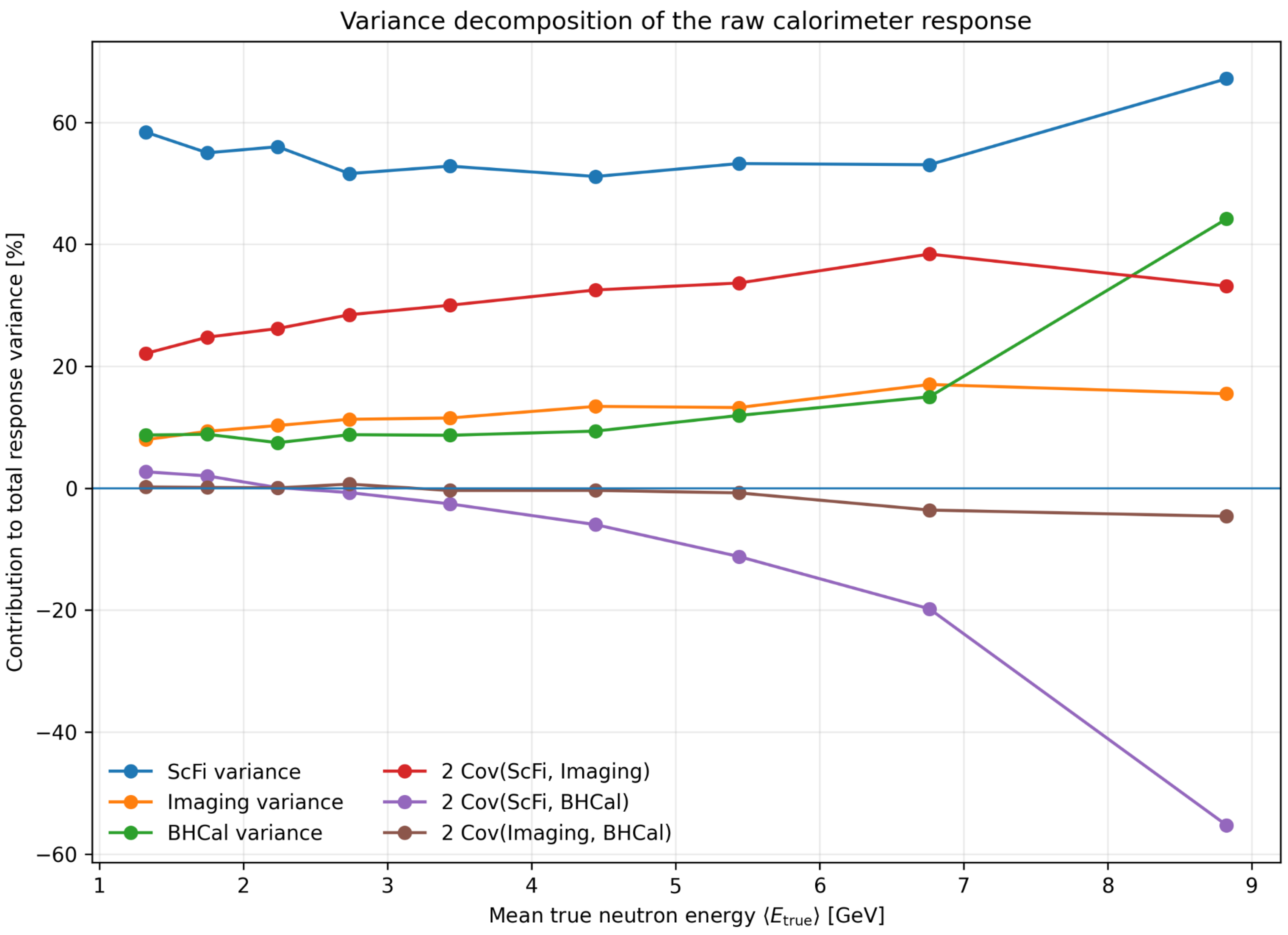




Variance Decomposition Highlights Upstream Fluctuations

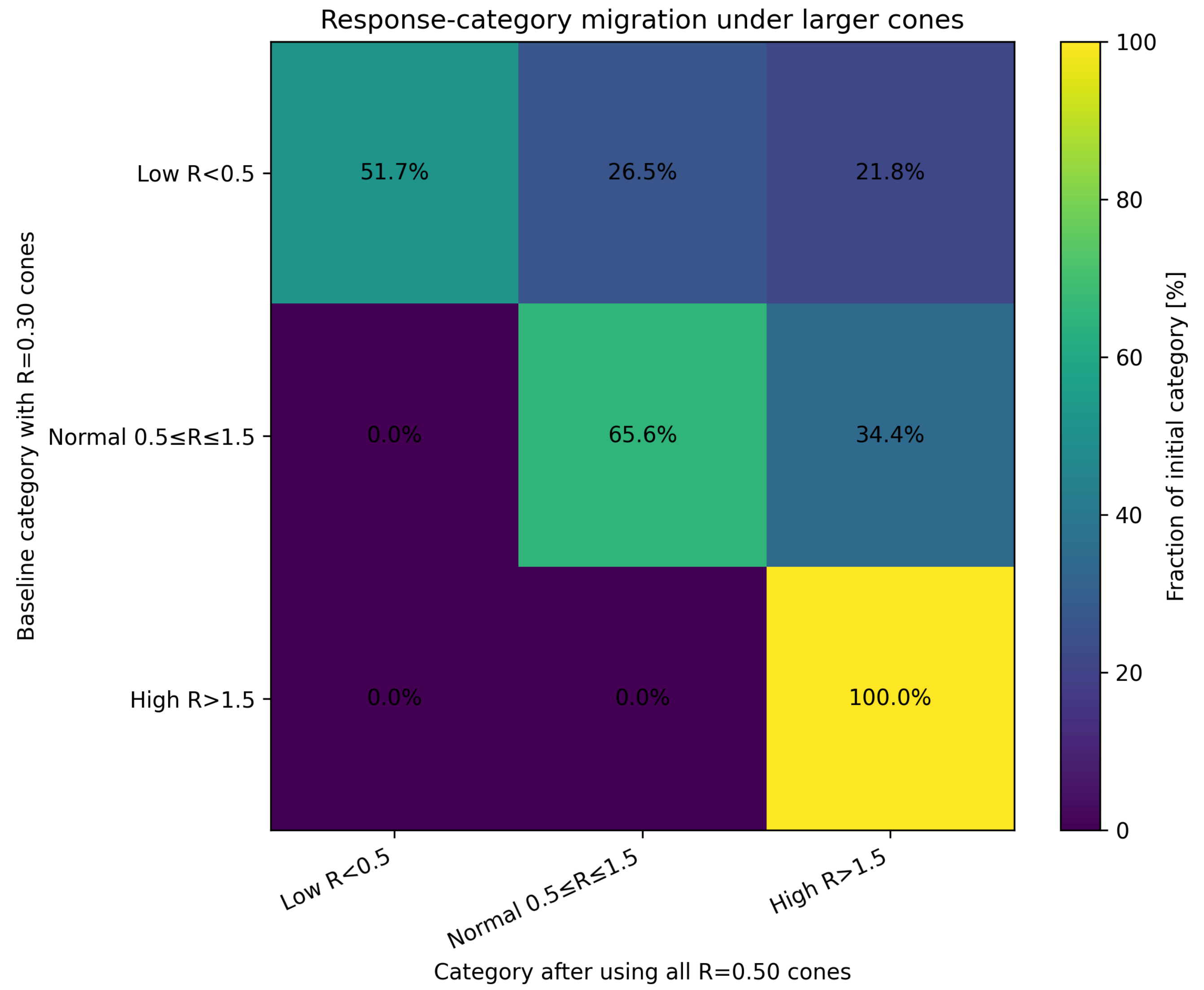
WHAT TO SAY

- ScFi provides the largest positive variance contribution over the full energy range.
- The ScFi–Imaging covariance is also large and positive, linking upstream activity to the high side.
- ScFi–BHCal covariance becomes negative at high energy, showing compensation between detector regions.



Larger Cones Recover Energy but Increase Contamination

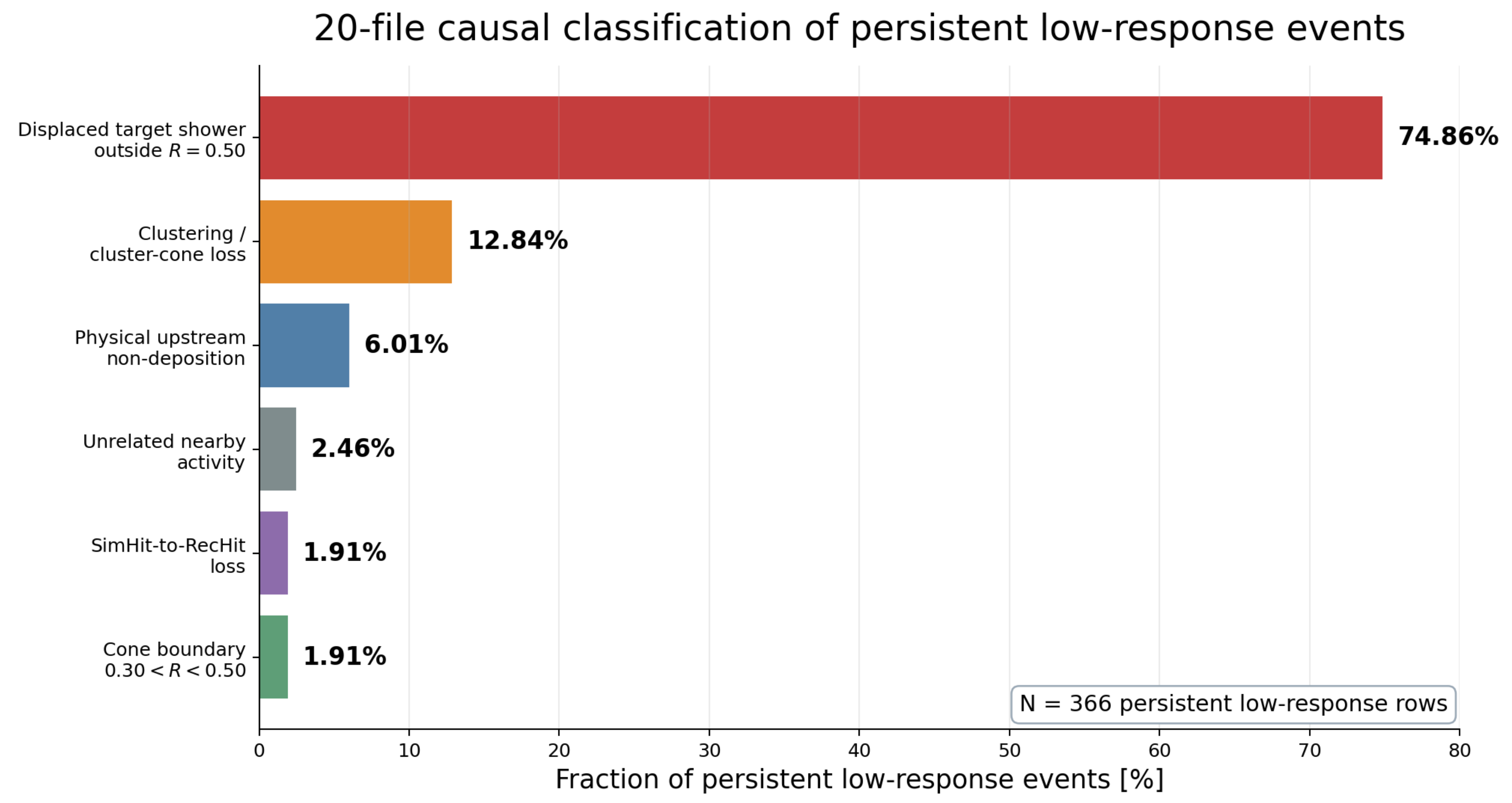
- With all cones increased to R = 0.50, 26.5% of baseline low events become normal.
- Another 21.8% of low events overshoot into the high category.
- Worse, 34.4% of previously normal events become high-response.



Truth-Level SimHits Identify the Primary Low-Tail Mechanism

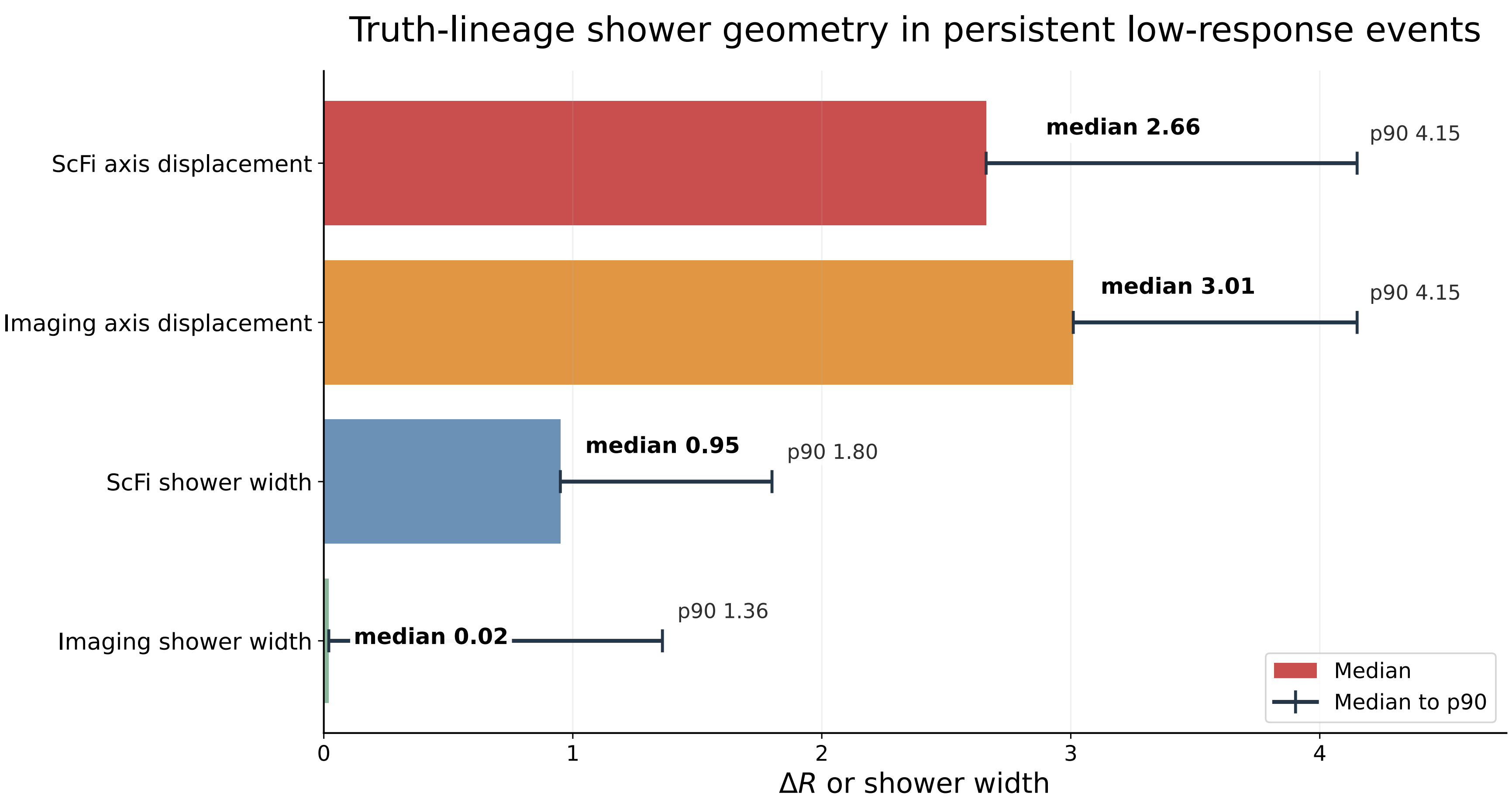
WHAT TO SAY

- The validated 20-file sample contains 366 persistent low-response rows without a parent-axis upstream cluster.
- In 74.9%, target-neutron SimHits exist only outside R = 0.50: the shower is displaced, not absent.
- Direct clustering loss is 12.8%; genuine upstream non-deposition is only 6.0%.



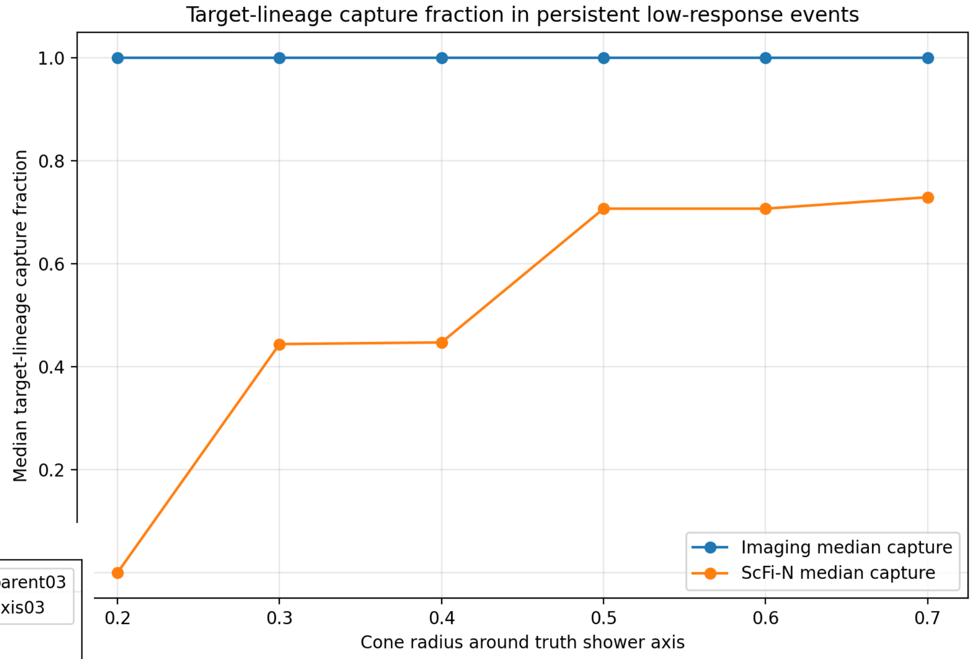
Neutron-Induced Showers Are Frequently Displaced from the Parent Direction

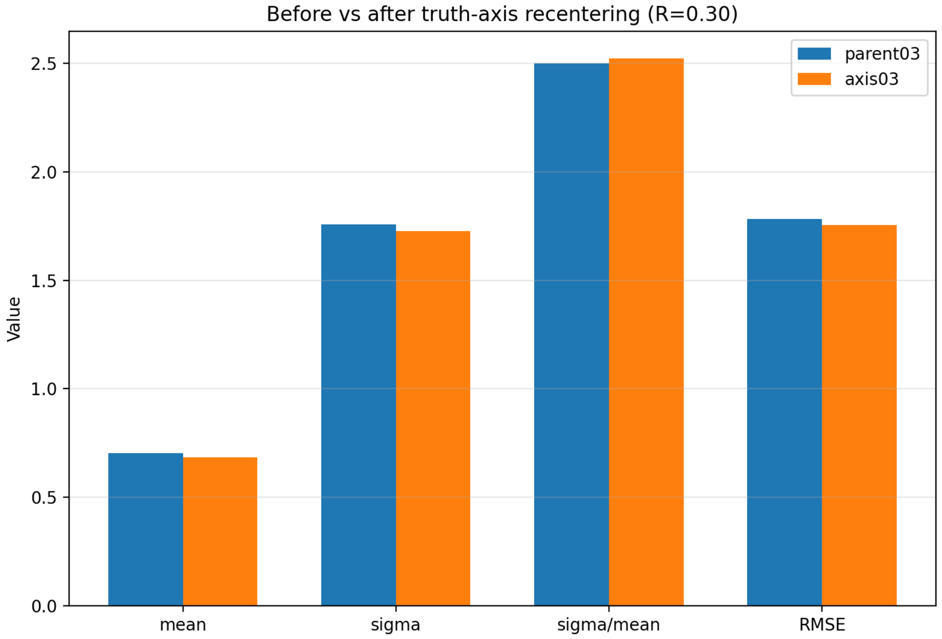
- Persistent low-response events show a **large displacement of the visible shower from the parent-neutron direction**.
- Median truth-lineage axis displacement is ΔR≈2.66 in SciFi and ΔR≈3.01 in Imaging.
- The displacement has a long tail: the **90th percentile reaches ΔR≈4.15** in both detectors.
- The SciFi shower is also relatively broad, with median width ≈0.95, while Imaging is usually more compact but has broad outliers.
- **Conclusion:** a fixed cone around the parent-neutron axis can miss a large fraction of the actual neutron-induced shower activity.



Truth-Axis Recovery Helps, but Not Enough

- Around the truth axis, median ScFi capture rises from about 44% at R = 0.30 to 71% at R = 0.50.
- The raw RMSE improves only from 1.7826 to 1.7547—about 1.6%—and the mean response moves farther below one.
- Miscentering is real, but broad showers, incomplete cluster recovery, and contamination remain.





Conclusions and Future Directions

- The raw neutron response is **strongly non-Gaussian**, with distinct low-response, central, and high-response populations.
- The response depends strongly on **shower topology**: BHCal-only events tend to under-respond, while upstream-rich events can populate the high-response tail.
- Tightening the match, requiring cleaner associations, or increasing the cone size does **not** remove the large raw width.
- Exact SimHit studies show that, for persistent low-response events, about **75% of target-neutron activity lies outside ΔR=0.5** from the parent-neutron direction.
- Local-axis studies show an additional **reconstruction bottleneck**: RecHit energy can be present, while reconstructed cluster energy remains zero in **many events**.

- The large raw neutron-energy resolution is **not caused by a single calibration factor**.
- The dominant problem is that neutron showers are often displaced, fragmented, and delayed relative to the parent-neutron direction.
- A fixed $\Delta R<0.3$ cone around the parent axis therefore misses a significant fraction of the visible shower.
- Even when upstream RecHits are recovered locally, **fragmentation and RecHit-to-cluster losses** can prevent that energy from appearing in reconstructed clusters.
- Simply enlarging the cone is not sufficient because it also collects unrelated activity and worsens the overall response/RMSE.
- The most promising detector-level direction is **BHCal-seeded, timing-aware association of upstream RecHits/components, followed by improved clustering or ML using the recovered information.**

- The poor raw resolution is primarily a **shower-topology and reconstruction problem**—displaced/fragmented neutron activity is missed by the parent-axis cone and is further lost during RecHit-to-cluster reconstruction.