C2^P2 Meeting at UIC

From Likelihoods to Language Models: AI's Evolving Role in High-Energy Physics

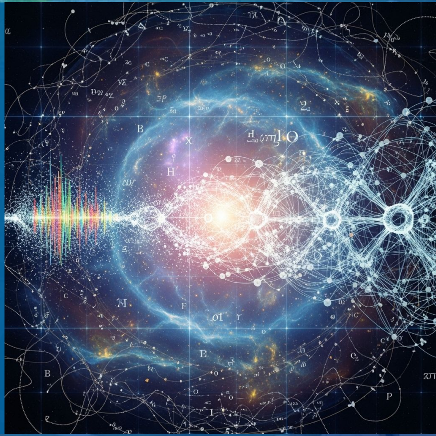
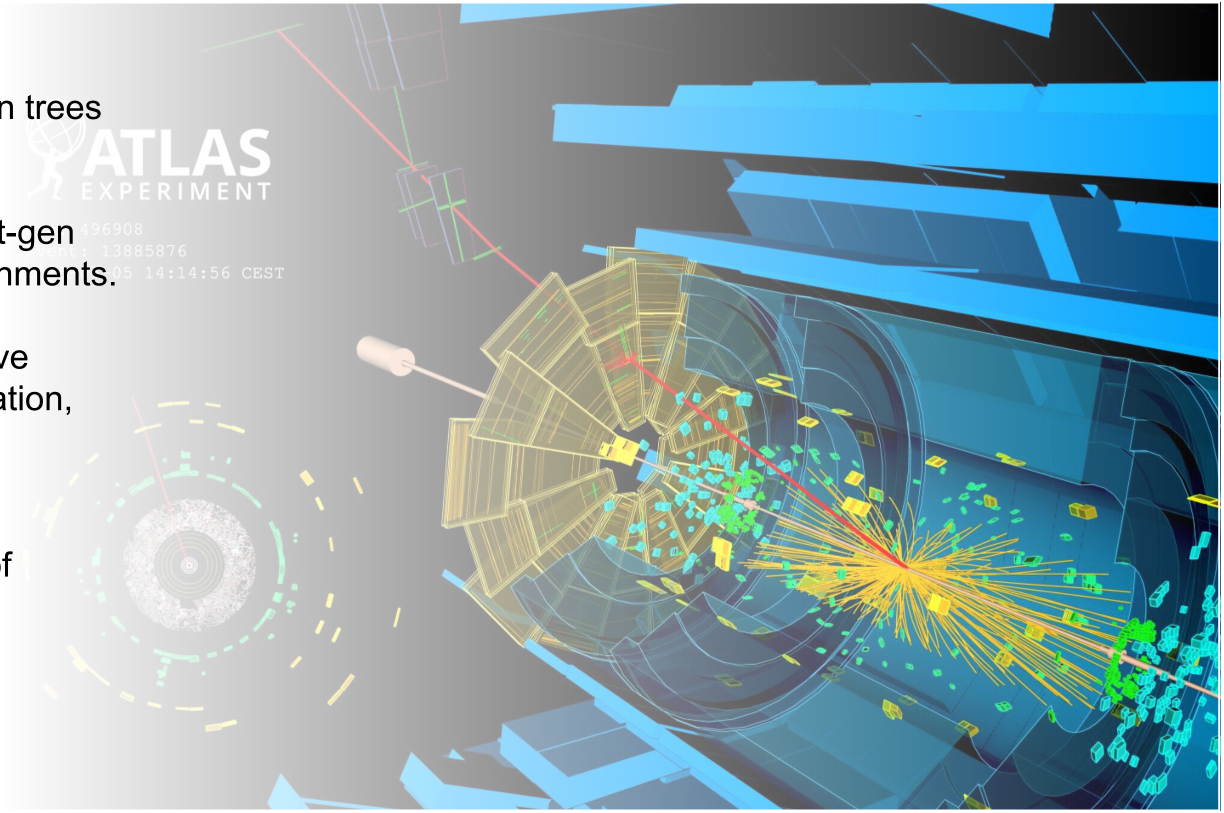


Imagen-4

J. Taylor Childers
Deputy Technical Directory, ALCF-4
Computational Scientist

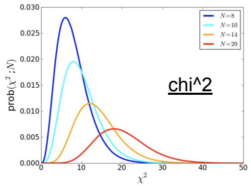
From Likelihoods to Language Models: AI’s Evolving Role in High-Energy Physics

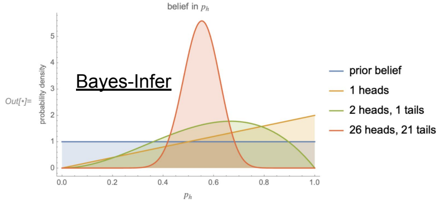
- **HEP’s long history with advanced analysis methods**
  From likelihood fits to boosted decision trees & random forests.
- **Why AI now?**
  Data deluge from LHC/HL-LHC & next-gen experiments; complex detector environments.
- **The leap to modern AI**
  Deep learning, transformers, generative models—tackling classification, simulation, and anomaly detection.
- **Impact**
  Improving physics reach, speeding simulations, and enabling new kinds of searches.

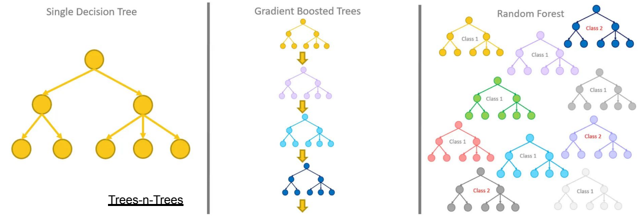


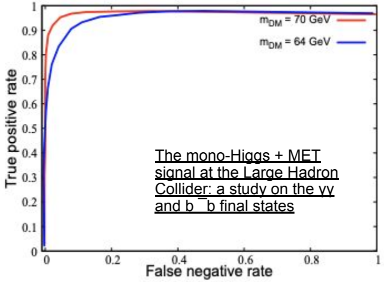
A Legacy of Statistical Learning in HEP

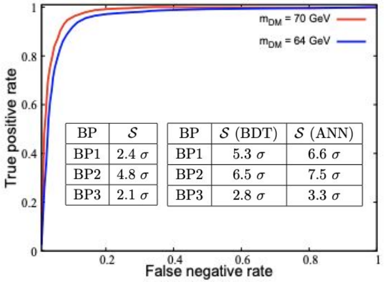
- **Analytical methods have always been central to discovery**
  — Likelihood fits, chi-squared minimization, Bayesian inference.
- **Adoption of early machine learning in the 2000s–2010s**
  — Boosted Decision Trees (BDTs) in Higgs searches (e.g., ATLAS, CMS).
  — Random Forests for particle ID and event classification.
- **Strengths and limits**
  — Highly interpretable, good for small datasets.
  — Struggle with complex correlations in high-dimensional feature spaces.
- **The setup for today**
  — Increasing detector complexity and HL-LHC data rates demand models with higher capacity and automation.



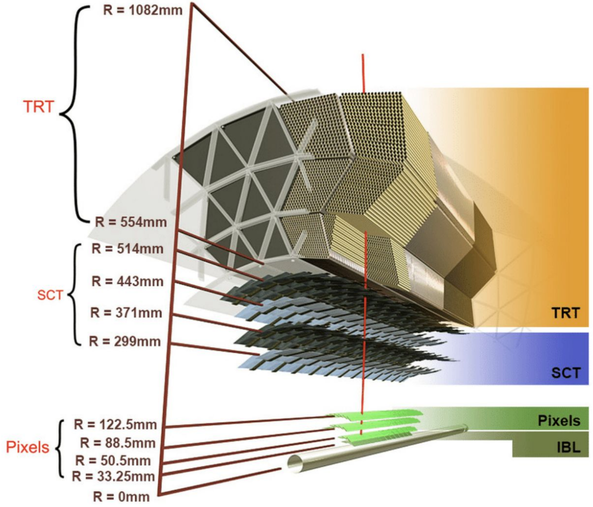




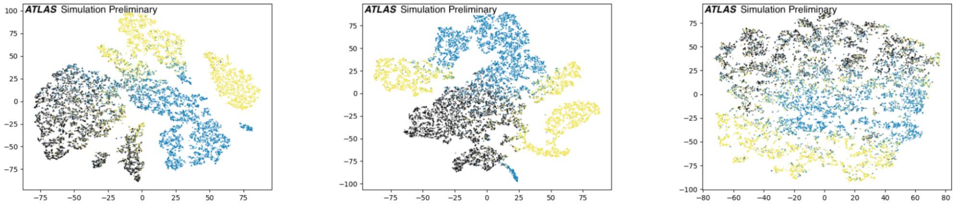




Modern AI for Object Identification



- **From handcrafted features to end-to-end learning**
  — Deep networks learn directly from detector hits or reconstructed objects.
- **Example architectures in use**
  — CNNs — for calorimeter image-based classification.
  — Graph Neural Networks (GNNs) — model particles or detector hits as nodes with physics-motivated edges.
  — Point-Cloud Networks — process particle-flow candidates or tracker hits directly.
- **Impact**
  — Higher tagging accuracy for b-jets, τ-jets, and boosted objects.
  — Resilience to pile-up and detector effects.
- **Key Examples**
  — ATLAS Point-Cloud Segmentation for particle flow object reconstruction (ATLAS note).
  — LArTPC hit-based topology classification with quantum machine learning and symmetry (arXiv:2503.12655).
  — A Comparison of Deep Learning Models for Proton Background Rejection with the AMS Electromagnetic Calorimeter (arxiv:2402.16285).

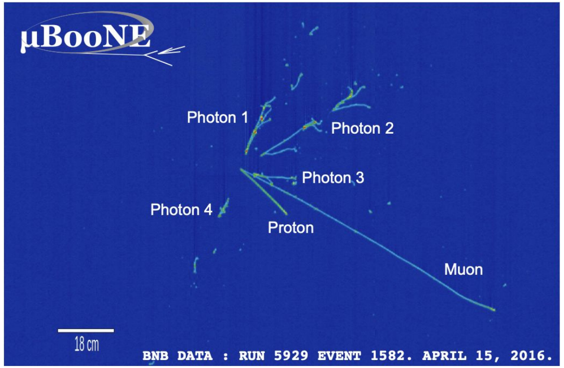


(a) PointNet++ (b) DGCNN (c) GravNet

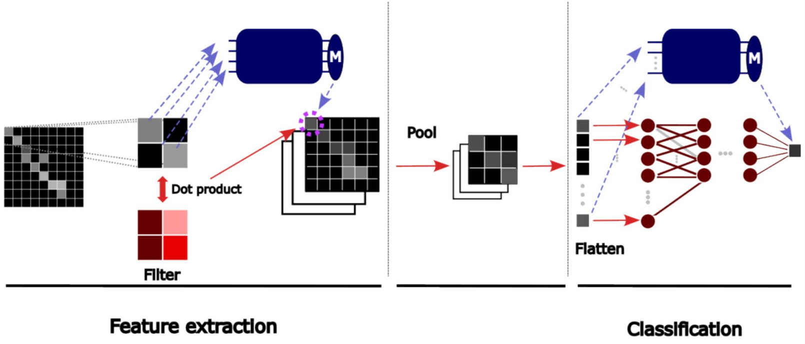
jet (black), electron (yellow), or background (blue)

	mIoU	Jet IoU	Electron IoU	Background IoU
PointNet++	0.776 ± 0.009	0.842 ± 0.007	0.61 ± 0.01	0.882 ± 0.007
GravNet	0.60 ± 0.02	0.74 ± 0.01	0.32 ± 0.02	0.75 ± 0.02
GarNet	0.43 ± 0.02	0.45 ± 0.04	0.13 ± 0.02	0.70 ± 0.02
DGCNN	**0.826 ± 0.005**	**0.885 ± 0.002**	**0.69 ± 0.0.01**	**0.904 ± 0.002**

Modern AI for Object Identification

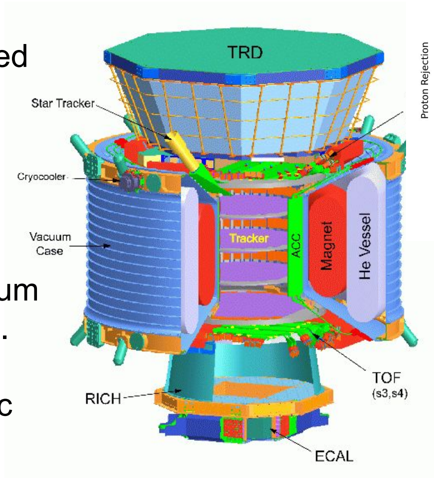


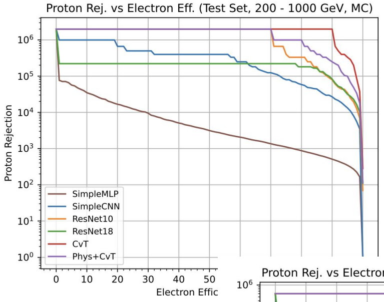
- **From handcrafted features to end-to-end learning**
  — Deep networks learn directly from detector hits or reconstructed objects.
- **Example architectures in use**
  — CNNs — for calorimeter image-based classification.
  — Graph Neural Networks (GNNs) — model particles or detector hits as nodes with physics-motivated edges.
  — Point-Cloud Networks — process particle-flow candidates or tracker hits directly.
- **Impact**
  — Higher tagging accuracy for b-jets, τ-jets, and boosted objects.
  — Resilience to pile-up and detector effects.
- **Key Examples**
  — ATLAS Point-Cloud Segmentation for particle flow object reconstruction (ATLAS note).
  — LArTPC hit-based topology classification with quantum machine learning and symmetry (arXiv:2503.12655).
  — A Comparison of Deep Learning Models for Proton Background Rejection with the AMS Electromagnetic Calorimeter (arxiv:2402.16285).

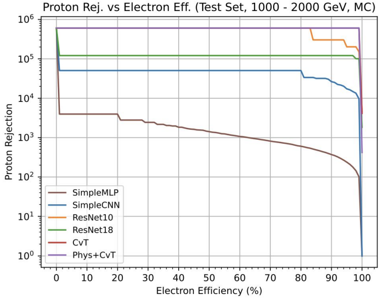


Modern AI for Object Identification

- **From handcrafted features to end-to-end learning**
  — Deep networks learn directly from detector hits or reconstructed objects.
- **Example architectures in use**
  — CNNs — for calorimeter image-based classification.
  — Graph Neural Networks (GNNs) — model particles or detector hits as nodes with physics-motivated edges.
  — Point-Cloud Networks — process particle-flow candidates or tracker hits directly.
- **Impact**
  — Higher tagging accuracy for b-jets, τ-jets, and boosted objects.
  — Resilience to pile-up and detector effects.
- **Key Examples**
  — ATLAS Point-Cloud Segmentation for particle flow object reconstruction (ATLAS note).
  — LArTPC hit-based topology classification with quantum machine learning and symmetry (arXiv:2503.12655).
  — A Comparison of Deep Learning Models for Proton Background Rejection with the AMS Electromagnetic Calorimeter (arxiv:2402.16285).

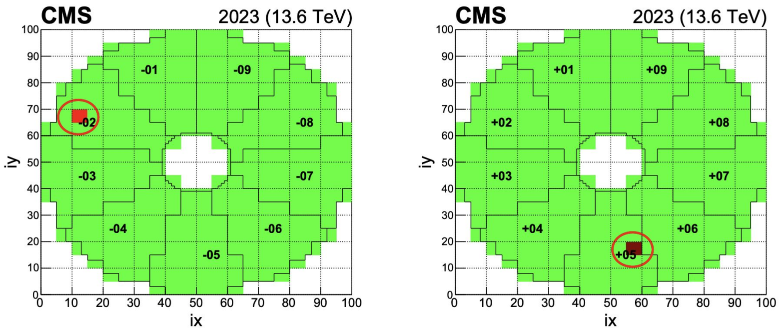


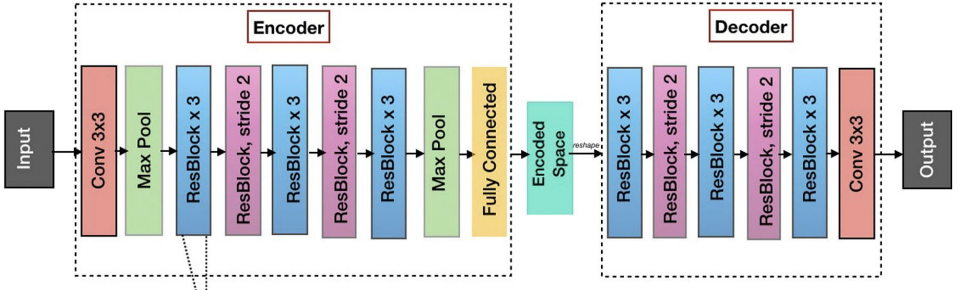


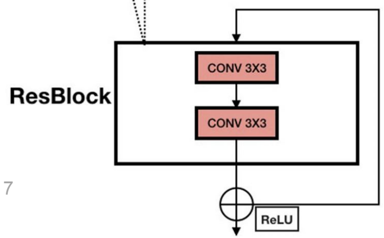


AI for Data Quality & Anomaly Detection

- **Data Quality Monitoring (DQM)**
  - — Detects detector malfunctions or calibration drifts in real-time.
  - — AI models learn normal detector behavior and flag deviations.
- **Anomaly Detection for Physics Searches**
  - — Trains on Standard Model “background” data only.
  - — Flags events with unusual topology, kinematics, or detector signatures.
- **Key Techniques**
  - — Autoencoders, Variational Autoencoders, Density Estimation, Normalizing Flows.
- **Impact**
  - — Early detection of detector issues → reduced downtime.
  - — Potential to uncover rare or unexpected physics signals without explicit search models.



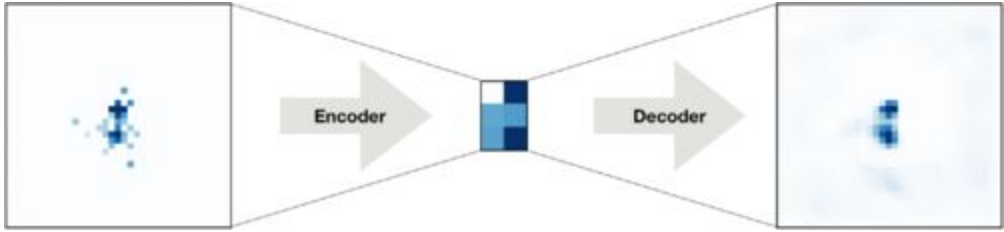


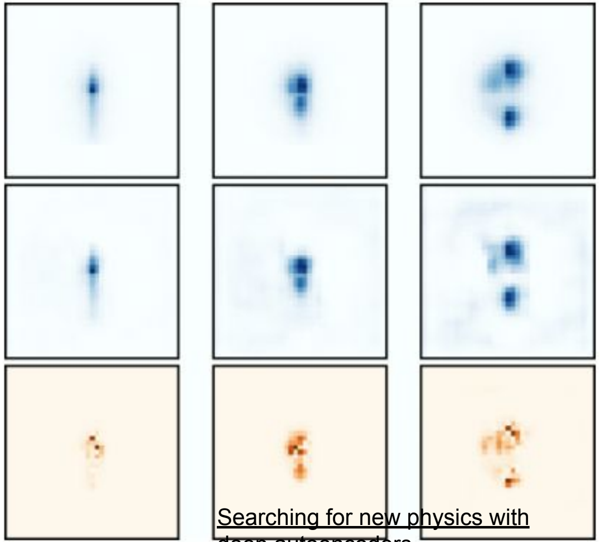


[Autoencoder-Based Anomaly Detection System for Online Data Quality Monitoring of the CMS Electromagnetic Calorimeter]

AI for Data Quality & Anomaly Detection

- **Data Quality Monitoring (DQM)**
  - Detects detector malfunctions or calibration drifts in real-time.
  - AI models learn normal detector behavior and flag deviations.
- **Anomaly Detection for Physics Searches**
  - Trains on Standard Model “background” data only.
  - Flags events with unusual topology, kinematics, or detector signatures.
- **Key Techniques**
  - Autoencoders, Variational Autoencoders, Density Estimation, Normalizing Flows.
- **Impact**
  - Early detection of detector issues → reduced downtime.
  - Potential to uncover rare or unexpected physics signals without explicit search models.



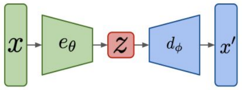
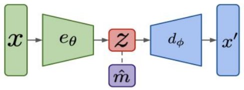
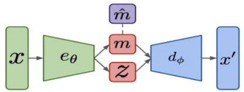
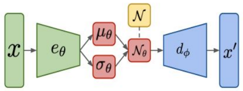
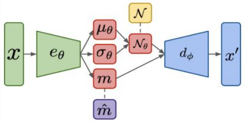


Each panel represents the average of 100,000 jet images. Pixel intensity corresponds to the total p_T in each pixel. Upper row: original sample. Middle row: after reconstruction. Lower row: pixelwise squared error. Left column: QCD jets. Middle column: top jets. Right column: g-jets.

[Searching for new physics with deep autoencoders]

AI for Probabilistic Modeling & Physics Interpretation

- **From data to theory parameters**
  - Neural networks can learn physics PDFs directly from collider data.
- **PDFdecoder (PRD 111.014028)**
  - Encoder–decoder architecture in Mellin space.
  - Latent space dimensions correspond to physically interpretable modes.
- **Advantages over traditional fits**
  - Flexible function approximation.
  - Naturally incorporates correlations between parameters
  - Allows uncertainty quantification via Bayesian or ensemble methods.
- **Beyond PDFs**
  - Similar approaches for detector response functions, cross-section unfolding, and global fits.

Name	Diagram	Loss	Recreates PDFs	Tractable Latent	Free Latent Dimension	Moment Constraint
AE		$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2$	✓	✗	✓	✗
AE-CL		$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2 + \|z - \hat{m}\|_2^2$	✓	✓	✓	✓
AE-WC		$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2 + \|m - \hat{m}\|_2^2$	✓	✗	✓	✓
VAE		$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2 + KL(\mathcal{N}(\mu_\theta, \sigma_\theta \| \mathcal{N}(0,1))$	✓	✓	✓	✗
VAIM		$\mathcal{L} = \|x - d_\phi(e_\theta(x))\|_2^2 + \|m - \hat{m}\|_2^2 + KL(\mathcal{N}(\mu_\theta, \sigma_\theta \| \mathcal{N}(0,1))$	✓	✓	✓	✓

Learning PDFs through interpretable latent representations in Mellin space

Fast Simulation with Generative AI

- **Motivation**

— Geant4 simulations are accurate but extremely slow — a major bottleneck at HL-LHC.

- **Generative Models**

— GANs (Generative Adversarial Networks).

— VAEs (Variational Autoencoders).

— Normalizing Flows and Diffusion Models for stable, high-fidelity generation.

- **Benchmark: CaloChallenge 2022**

— Community-wide comparison of generative AI methods for calorimeter simulation.

— Metrics: shower shape, energy resolution, computing throughput.

- **Impact**

— 100–10,000× speedup over Geant4 for certain workflows.

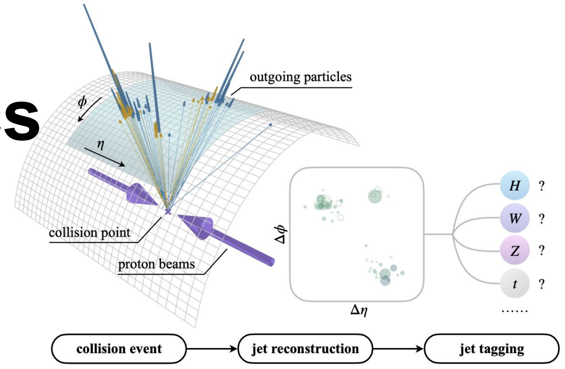
— Enables large-scale MC production, rapid detector studies, and real-time inference applications.

Table 1: Models submitted to the CaloChallange.

Approach	Model	Code	Dataset $1-\gamma$	Dataset $1-\pi$	Dataset 2	Dataset 3	Section
GAN	`CaloShowerGAN` [21]	[22]	✓	✓			3.1
	`MDMA` [23, 24]	[25]			✓	✓	3.2
	`BoloGAN` [26]	[27]	✓	✓			3.3
	`DeepTree` [28, 29]	[30]			✓		3.4
NF	`L2LFlows` [31, 32]	[33]			✓	✓	4.1
	`CaloFlow` [34, 35]	[36, 37]	✓	✓	✓	✓	4.2
	`CaloINN` [38]	[39]	✓	✓	✓		4.3
	`SuperCalo` [40]	[41]			✓		4.4
	`CaloPointFlow` [42]	[43]			✓	✓	4.5
Diffusion	`CaloDiffusion` [44]	[45]	✓	✓	✓	✓	5.1
	`CaloClouds` [46, 47]	[48, 49]				✓	5.2
	`CaloScore` [50, 51]	[52, 53]	✓		✓	✓	5.3
	`CaloGraph` [54]	[55]	✓	✓			5.4
	`CaloDiT` [56]	[57]			✓		5.5
VAE	`Calo-VQ` [58]	[59]	✓	✓	✓	✓	6.1
	`CaloMan` [60]	[61]	✓	✓			6.2
	`DNNCaloSim` [62, 63]	[64]		✓			6.3
	`Geant4-Transformer` [65]	[66]				✓	6.4
	`CaloVAE+INN` [38]	[39]	✓	✓	✓	✓	6.5
	`CaloLatent` [67]	[68]			✓		6.6
CFM	`CaloDREAM` [69]	[70]			✓	✓	7.1
	`CaloForest` [71]	[72]	✓	✓			7.2

CaloChallenge 2022: A Community Challenge for Fast Calorimeter Simulation

Transformers in High-Energy Physics



Particle Transformer for Jet Tagging

- **Why Transformers?**

— Designed for sequence modeling — natural fit for particle lists, detector hits, and even time-series data.

— Handle variable-length, unordered inputs and capture global correlations through self-attention.

— Scalable: performance improves with model size and training data availability.

- **Applications in HEP**

— Jet Tagging — Particle Transformer (ParT) on JETCLASS dataset shows SOTA classification (arXiv:2202.03772).

— Event Classification — Transformer architectures for multi-object event-level tasks, outperforming CNNs and GNNs on physics benchmarks (PRD 109.096035).

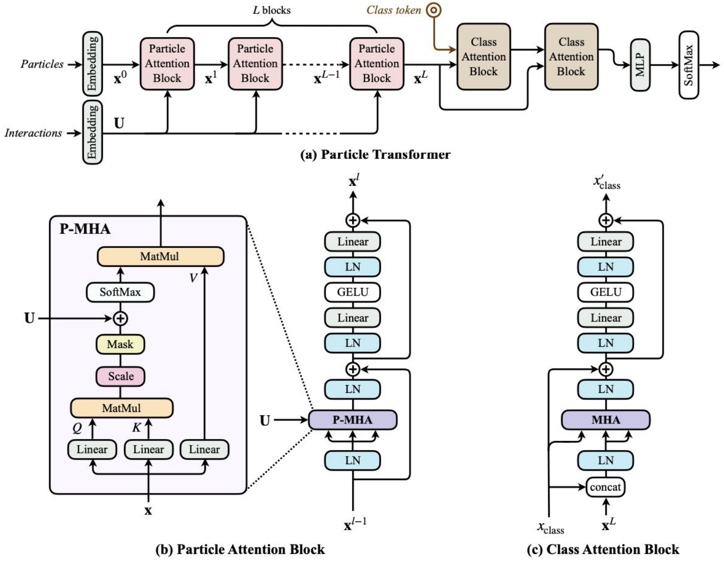
— Physics Object Identification — HL-LHC-ready attention-based particle flow and reconstruction algorithms (arXiv:2507.17807).

- **Advantages Over Previous Architectures**

— Captures long-range correlations beyond local neighborhoods (unlike CNNs or GNNs).

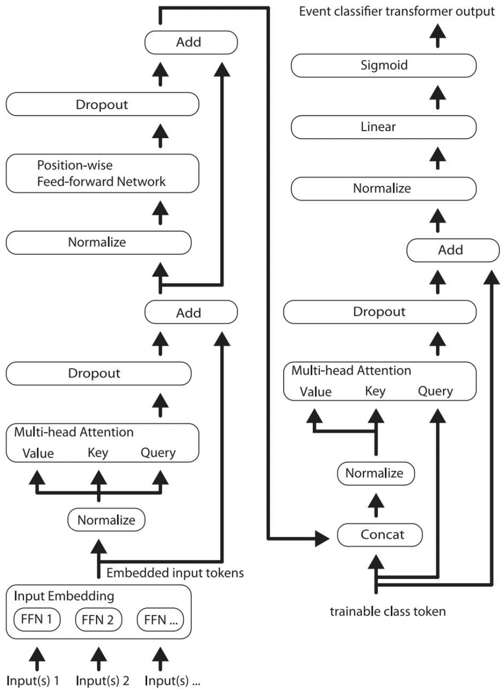
— Flexible for multi-modal data (tracking + calorimeter + timing layers).

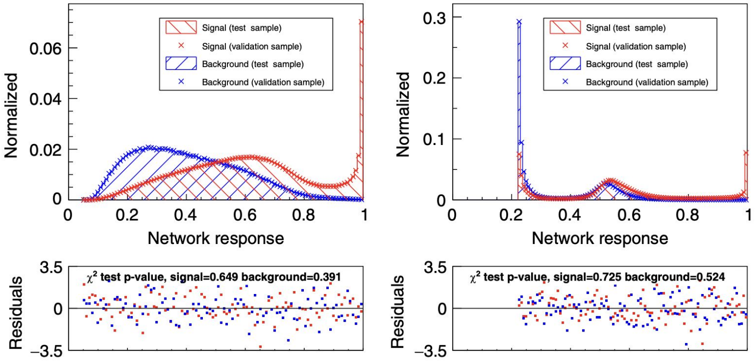
— Works well in low-level and high-level feature spaces.



Transformers in High-Energy Physics

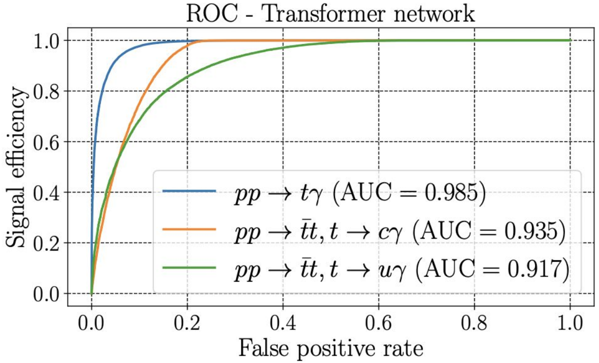
- **Why Transformers?**
  - Designed for sequence modeling — natural fit for particle lists, detector hits, and even time-series data.
  - Handle variable-length, unordered inputs and capture global correlations through self-attention.
  - Scalable: performance improves with model size and training data availability.
- **Applications in HEP**
  - Jet Tagging — Particle Transformer (ParT) on JETCLASS dataset shows SOTA classification (arXiv:2202.03772).
  - Event Classification — Transformer architectures for multi-object event-level tasks, outperforming CNNs and GNNs on physics benchmarks (PRD 109.096035).
  - Physics Object Identification — HL-LHC-ready attention-based particle flow and reconstruction algorithms (arXiv:2507.17807).
- **Advantages Over Previous Architectures**
  - Captures long-range correlations beyond local neighborhoods (unlike CNNs or GNNs).
  - Flexible for multi-modal data (tracking + calorimeter + timing layers).
  - Works well in low-level and high-level feature spaces.

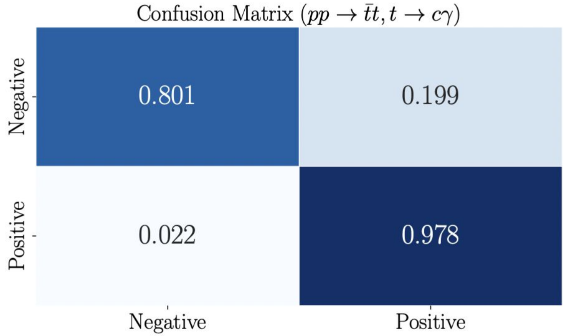




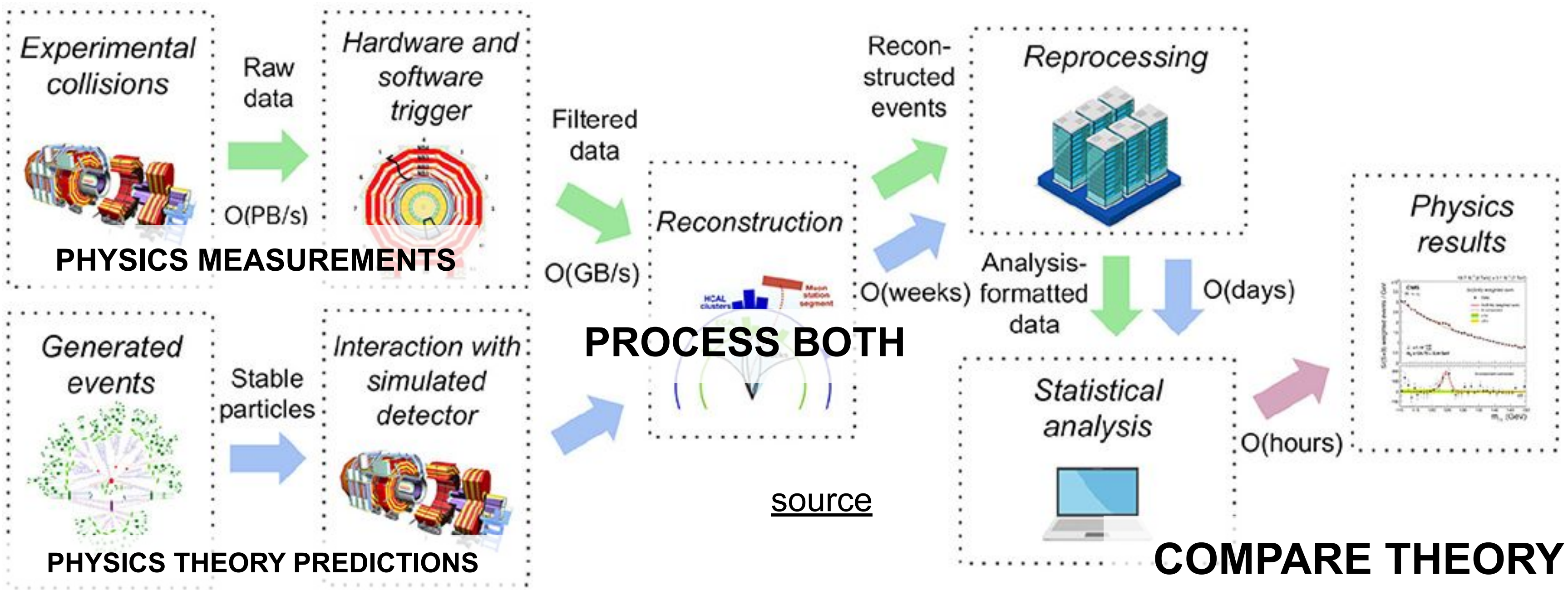
Transformers in High-Energy Physics

- **Why Transformers?**
  - — Designed for sequence modeling — natural fit for particle lists, detector hits, and even time-series data.
  - — Handle variable-length, unordered inputs and capture global correlations through self-attention.
  - — Scalable: performance improves with model size and training data availability.
- **Applications in HEP**
  - — Jet Tagging — Particle Transformer (ParT) on JETCLASS dataset shows SOTA classification (arXiv:2202.03772).
  - — Event Classification — Transformer architectures for multi-object event-level tasks, outperforming CNNs and GNNs on physics benchmarks (PRD 109.096035).
  - — Physics Object Identification — HL-LHC-ready attention-based particle flow and reconstruction algorithms (arXiv:2507.17807).
- **Advantages Over Previous Architectures**
  - — Captures long-range correlations beyond local neighborhoods (unlike CNNs or GNNs).
  - — Flexible for multi-modal data (tracking + calorimeter + timing layers).
  - — Works well in low-level and high-level feature spaces.





Toward Real-Time Physics Fits with AI

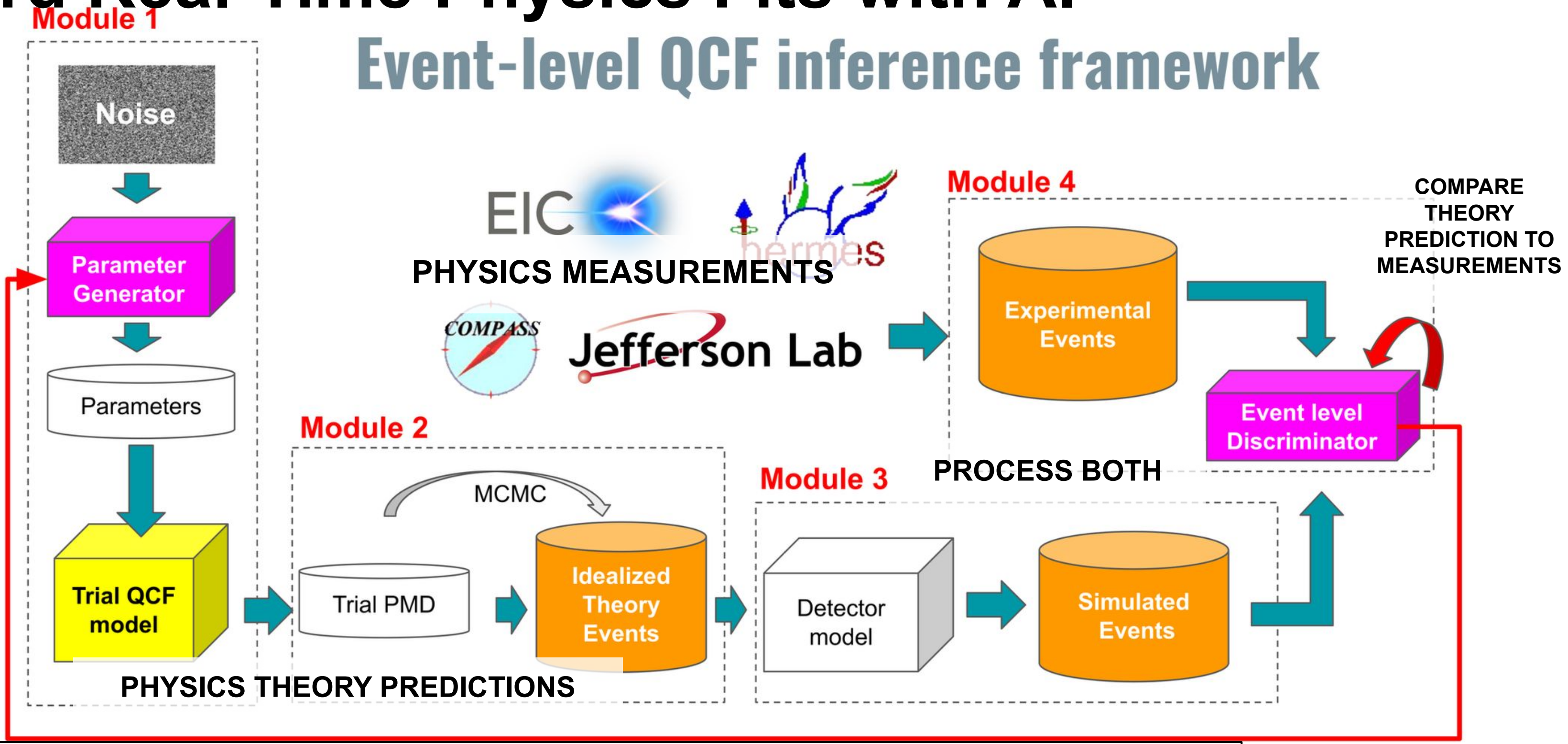


source

COMPARE THEORY PREDICTION TO MEASUREMENTS

- **The Traditional Model in HEP**
  - Experiments run for months/years, producing massive datasets.
  - Full analysis performed after data collection is complete.
  - Theory fits and interpretations only finalized long after running ends.

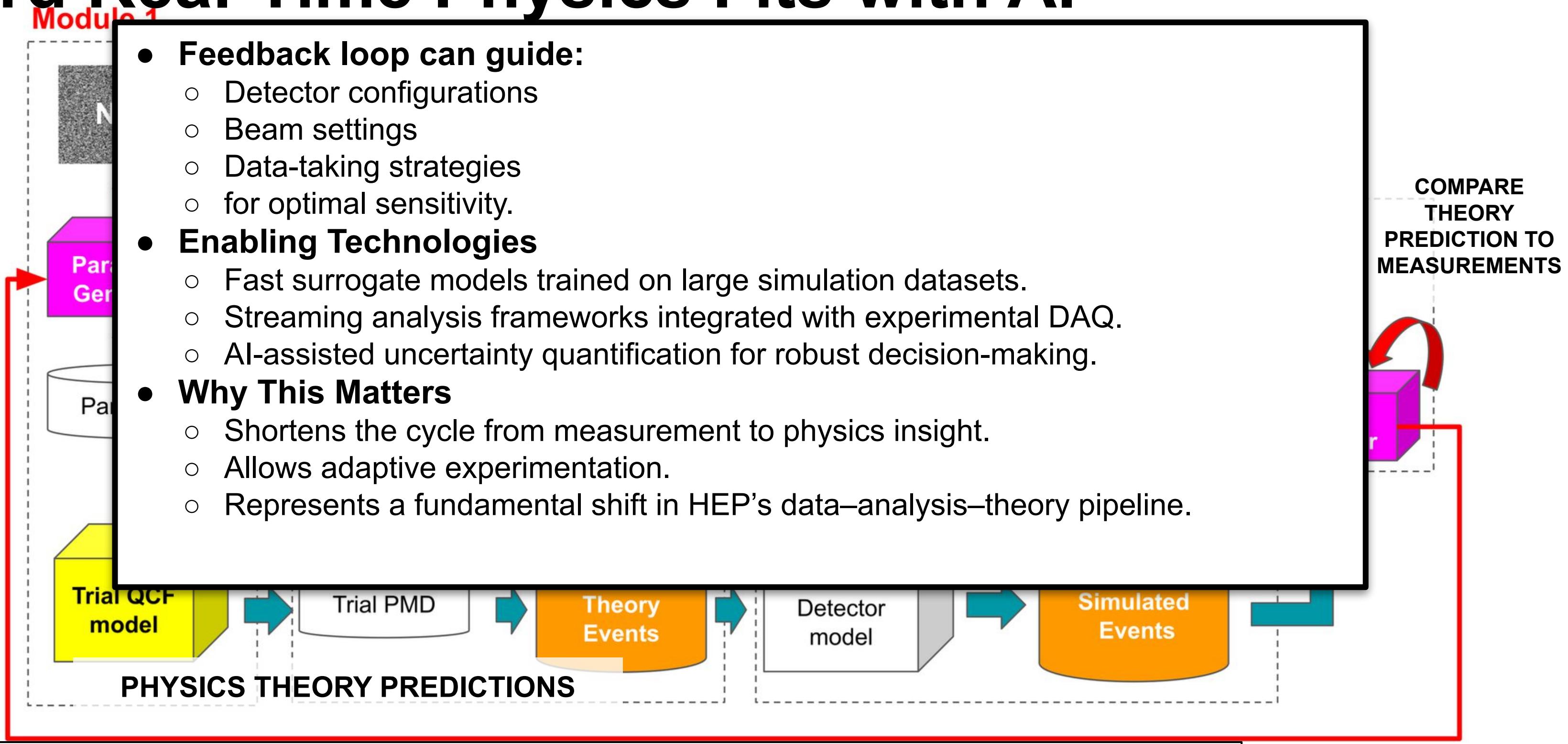
Toward Real-Time Physics Fits with AI



SAGIPS

- **A New Paradigm: Real-Time Fitting**
  - SciDAC project for the EIC explores AI-enabled parameter estimation during data collection.
  - AI models ingest live experimental data, update theory fit parameters in near real-time.

Toward Real-Time Physics Fits with AI



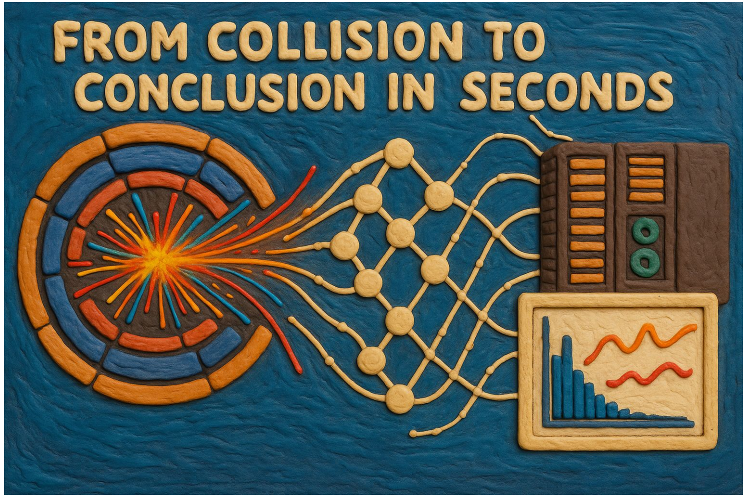
- **Feedback loop can guide:**
  - Detector configurations
  - Beam settings
  - Data-taking strategies
  - for optimal sensitivity.
- **Enabling Technologies**
  - Fast surrogate models trained on large simulation datasets.
  - Streaming analysis frameworks integrated with experimental DAQ.
  - AI-assisted uncertainty quantification for robust decision-making.
- **Why This Matters**
  - Shortens the cycle from measurement to physics insight.
  - Allows adaptive experimentation.
  - Represents a fundamental shift in HEP’s data–analysis–theory pipeline.

SAGIPS

- **A New Paradigm: Real-Time Fitting**
  - SciDAC project for the EIC explores AI-enabled parameter estimation during data collection.
  - AI models ingest live experimental data, update theory fit parameters in near real-time.

Takeaways & Outlook

- **HEP has always been data-driven** — AI is the next step in a long tradition of statistical innovation.
- **Modern AI methods** (CNNs, GNNs, Transformers, Generative Models) are already improving classification, simulation, and anomaly detection.
- **Integration with theory is deepening** — probabilistic models, real-time fits, and adaptive experiments are emerging.
- **The future is faster, smarter, and more adaptive** — AI will shorten the path from data to discovery.



GPT-5