Information-theoretic and quantum-informed approaches for jets at the LHC

Aritra Bal [1,2] , Michael Binder[3], Benedikt Maier[4], Markus Klute[3], Michael Spannowsky[1,2]

ML4Jets 2026, Vienna

1. Institute of Theoretical Physics (ITP), KIT Karlsruhe
2. Institute of Quantum Materials and Technologies (IQMT), KIT Karlsruhe
3. Institute of Experimental Particle Physics (ETP), KIT Karlsruhe
4. Imperial College, London

Karlsruhe Institute of Technology

Motivation

- Jets → Several high energy particles (and many *soft* ones)
  - What particular combination is important?
  - To what order do we model correlations?
  - Might have implications for jet tagging

- Information Theory could help:
  - Jet substructure variables
  - Higher order machine learning architectures (such as hypergraphs)
  - Fisher Information Matrix and higher order cumulant tensors

Can we connect them?

Karlsruhe Institute of Technology

Jet Substructure Observables

For a jet constituent i:

$$z_i = \frac{E_i}{\sum_j E_j} \qquad \text{or} \qquad z_i = \frac{p_{T,i}}{\sum_j p_{T,j}},$$

Energy Correlator Function (ECF):

$$e_n^{(\beta)}(x) = \sum_{1 \leq i_1 \leq \cdots \leq i_N \leq N} \left(\prod_{r=1}^{n} z_{i_r} \right) \left(\prod_{r<s} \theta_{i_r i_s}^{\beta} \right)$$

And likewise the EEC:

$$\mathrm{EEC}_x(\chi) = \sum_{i<j} 2 z_i z_j \, \delta\big(\cos\chi - \cos\chi_{ij} \big)$$

Aritra Bal 18 September 2026 aritra.bal@kit.edu

Higher Order Statistics: Beyond 2nd order

Set of jet observables →joint fluctuations

Exponential probability distribution

$$p_\theta(x) = h(x)\exp[\theta^a \phi_a(x) - \psi(\theta)]$$

Second Order: Fisher Information

$$\mathcal{C}_{ab} = \langle \Delta\mathcal{O}_a \Delta\mathcal{O}_b \rangle$$

$$\mathcal{I}^{(2)}_{ab}(\theta) \quad = \mathrm{Cov}_\theta(\phi_a, \phi_b) = \langle s_a s_b \rangle_\theta$$

Pairwise fluctuations of observables

Higher Orders: Cumulant Tensors

$$\mathcal{T}_{abc} = \langle \Delta\mathcal{O}_a \Delta\mathcal{O}_b \Delta\mathcal{O}_c \rangle$$

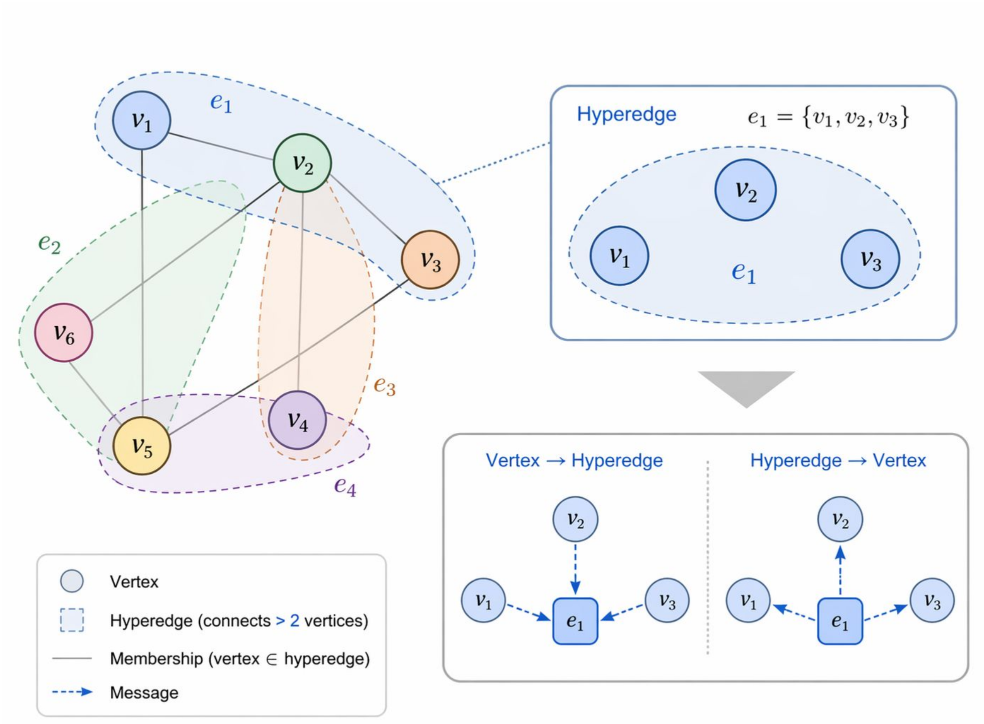
$$\mathcal{I}^{(3)}_{abc}(\theta) \quad = \kappa_\theta(\phi_a \phi_b \phi_c) = \langle s_a s_b s_c \rangle_\theta$$

But are all interactions pairwise-reducible?

Local KL Expansion of a distribution: quadratic response + cubic correction

Karlsruhe Institute of Technology

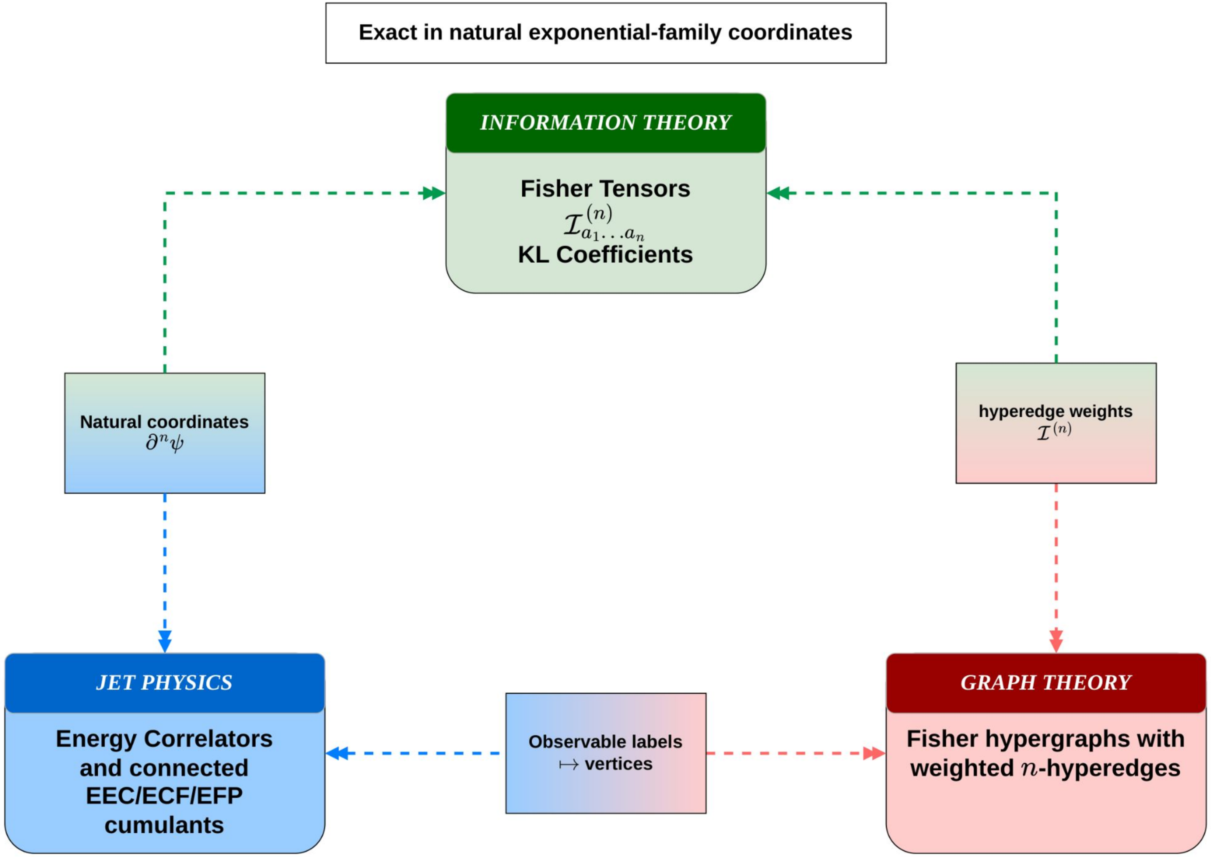
Hypergraphs



Ordinary graph: edges
Hypergraph: ***hyper*-edges**

Message passing → collates information along a hyper-edge

Introducing a Triality Relation



- Higher order Fisher tensors
- Energy Correlation Functions (ECFs) or Energy-Energy Correlators (EECs)
- n>2 hypergraphs

Connecting the Edges of the Triality

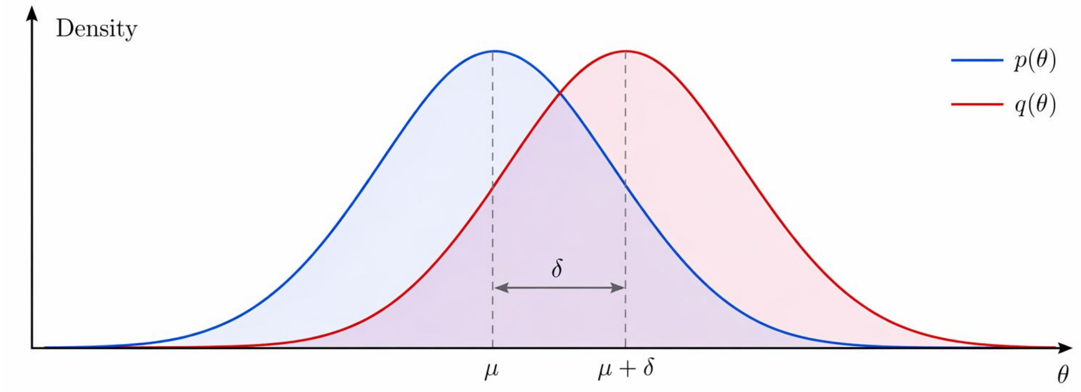
- Model jet data as exponential probability distribution
  - Observables → jet EECs + ECFs
  - Background sample: at parameter point θ
  - Local deformation: θ + δ

For an exponential-family model built from energy correlators, the tensor $\mathcal{I}^{(n)}_{a_1 \cdots a_n}$ has three equivalent interpretations:

1. *it is the coefficient of $\delta^{a_1} \cdots \delta^{a_n}/n!$ in the exact KL expansion*
2. *it is the connected cumulant of the correlator observables,*
3. *it is the signed weight of the n-hyperedge $\{a_1, \ldots, a_n\}$ in the Fisher hypergraph.*

Karlsruhe Institute of Technology

Toy Demonstration: tracking KL divergence



KL Divergence (from p to q)

$$D_{\mathrm{KL}}\,(p \,\|\, q) = \int_{\Theta} p(\theta) \log \frac{p(\theta)}{q(\theta)}\, d\theta$$

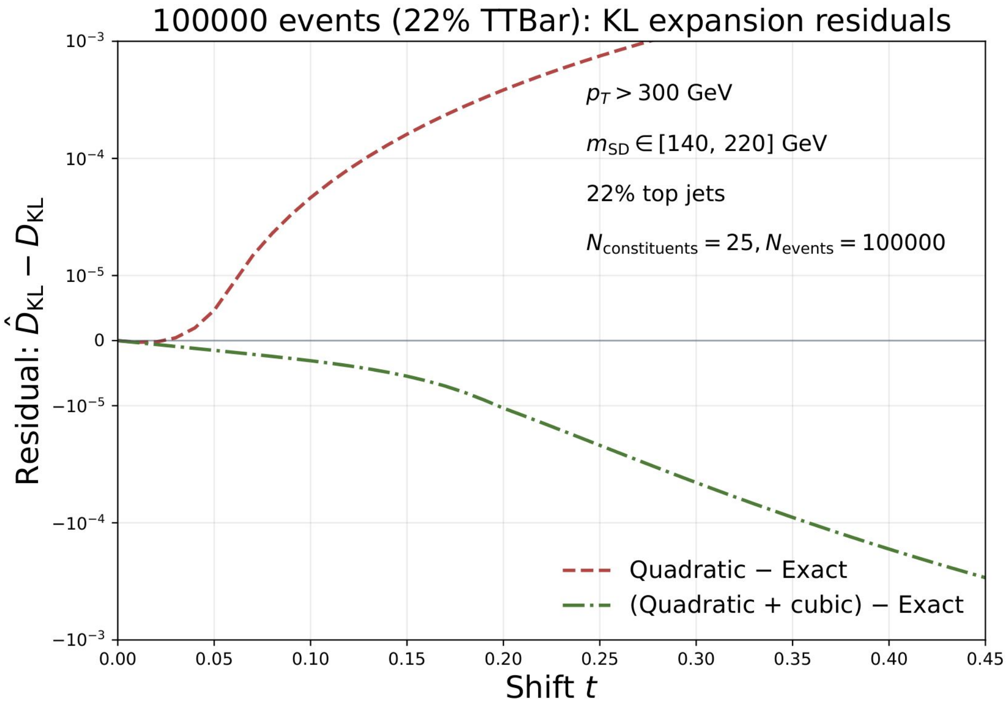
$$q(\theta) = p(\theta + \delta)$$

$$\int_{\Theta} p(\theta)\, d\theta = 1, \quad \int_{\Theta} q(\theta)\, d\theta = 1$$

$$D_{\mathrm{KL}}(p \,\|\, q) = \sum_{n=1}^{\infty} \frac{1}{n!} \mathcal{I}_{i_1 \ldots i_n}\, \delta^{i_1} \cdots \delta^{i_n}$$

- Two distributions: signal and background
- Irreducible 3-correlation information cannot be modeled at n=2

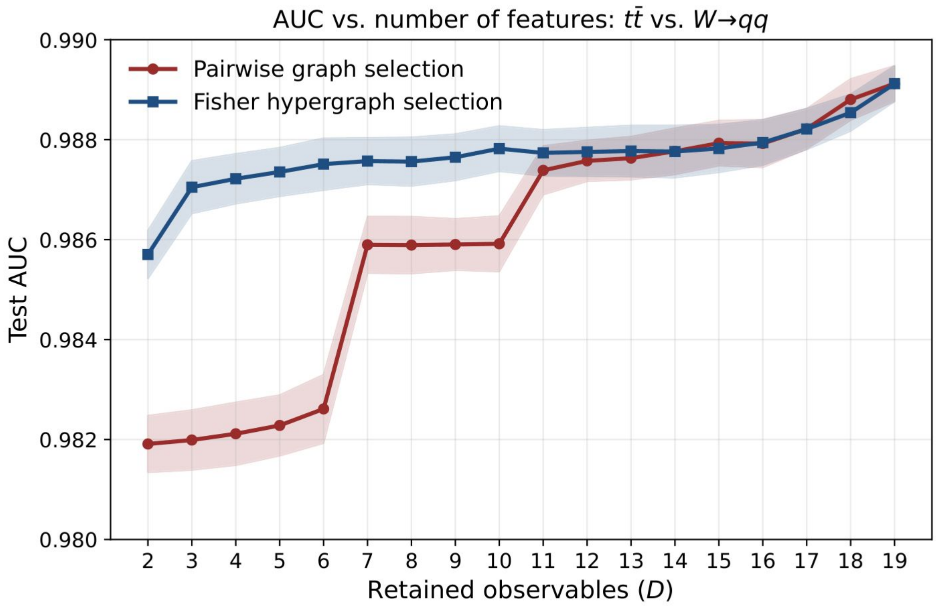
Toy Demonstration: tracking KL divergence



- Use observable basis of size 4
- Dataset with 2 classes of jets → KL expansion along class mean direction
- Third order terms built from the EEC/ECF basis reduces error by upto 29x at larger deformations
- Minimal demonstration: but … can this additional structure **actually** help?

Observable Compression

- Real analyses → hundreds of observables (binned EECs, ECFs, EFPs, etc)
- Can we find the best subset with minimal effort?
- Sample dataset: ttbar (3-pronged) vs W (2-pronged)
- Large basis of 18 observables
  - Start from 2 (common) observables
  - Design 2nd (3rd) order selectors using the respective tensors
  - Greedy optimization: maximise classifier score at each selection step



Karlsruhe Institute of Technology

Physics-informed Hypergraph Learning

- Graph-based ML widely used in HEP already
- Low-capacity application: graph vs hypergraph
- Probe sensitivity to EFT parameters via a small classifier:
  - Top jet sample (3 prongs)
  - Soft 4th prong deformation sample
- Network structure kept simple: graph/hypergraph structure described by corresponding Laplacian, followed by single MLP readout
  - Truncate to top 150 entries at second order, and top 200 entries at third order

Classifier performance

$AUC^{(2)} = 0.8184 \pm 0039$
$AUC^{(3)} = 0.8258 \pm 0029$

6000 labeled events,

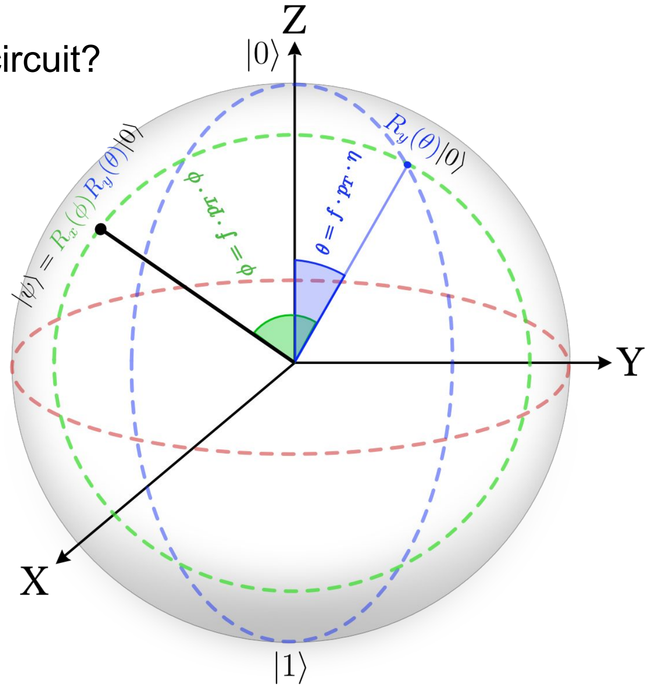
Read the preprint here

ONWARDS TO QUANTUM ..

Quantum Machine Learning: Why?

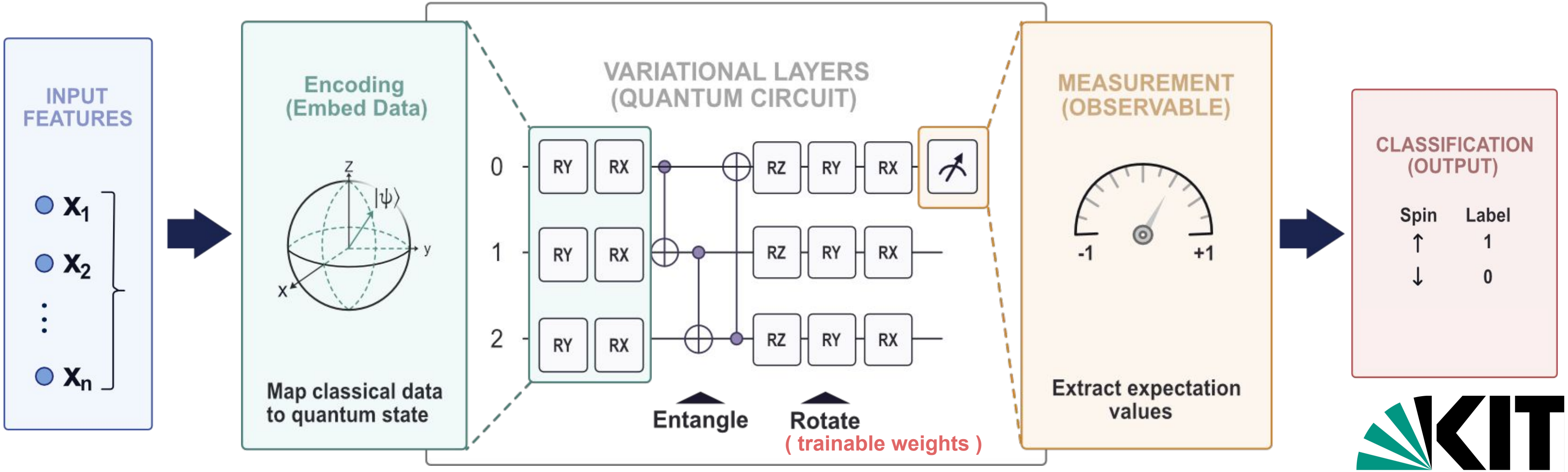
- Is there (similar) useful **geometric information** in a quantum circuit?
- Goal: model complex correlations in high-dimensional data
- Superposition/entanglement →no classical equivalents

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \;\; \Rightarrow \;\; |\psi\rangle = \cos\left(\frac{\theta}{2}\right)|0\rangle + e^{-i\phi}\sin\left(\frac{\theta}{2}\right)|1\rangle$$



Variational Quantum Circuits as Neural Networks

- Circuit acts as a (**Quantum**) Neural Network
- Data encoding → Entanglement → Parameterized unitaries → Measurement



The One Particle - One Qubit (1P1Q) Mapping

PHYSICAL REVIEW D **112**, 076004 (2025)

One particle - one qubit: Particle physics data encoding for quantum machine learning

Aritra Bal* and Markus Klute
Institute of Experimental Particle Physics (ETP), Karlsruhe Institute of Technology, Karlsruhe, Germany

Benedikt Maier,† Melik Oughton, and Eric Pezone
Imperial-X and Department of Physics, Imperial College Of Science, Technology And Medicine, London, United Kingdom

Michael Spannowsky‡
Institute for Particle Physics Phenomenology (IPPP), Durham University, Durham, United Kingdom

(Received 2 March 2025; accepted 20 August 2025; published 6 October 2025)

We introduce 1P1Q, a novel quantum data encoding scheme for high-energy physics (HEP), where each particle is assigned to an individual qubit, enabling direct representation of collision events on quantum circuits without classical compression. We demonstrate the effectiveness of 1P1Q in quantum machine learning (QML) through two applications: a quantum autoencoder (QAE) for unsupervised anomaly detection and a variational quantum circuit (VQC) for supervised classification of top quark jets. Our results show that the QAE successfully distinguishes signal jets from background quantum chromodynamics (QCD) jets, achieving superior performance compared to a classical autoencoder while utilizing significantly fewer trainable parameters. Similarly, the VQC achieves competitive classification performance, approaching state-of-the-art classical models despite its minimal computational complexity. Furthermore, we validate the QAE on real experimental data from the CMS detector, establishing the robustness of quantum algorithms in practical HEP applications. These results demonstrate that 1P1Q provides an effective and scalable quantum encoding strategy, offering new opportunities for applying quantum computing algorithms in collider data analysis.

DOI: 10.1103/l8y2-87vq

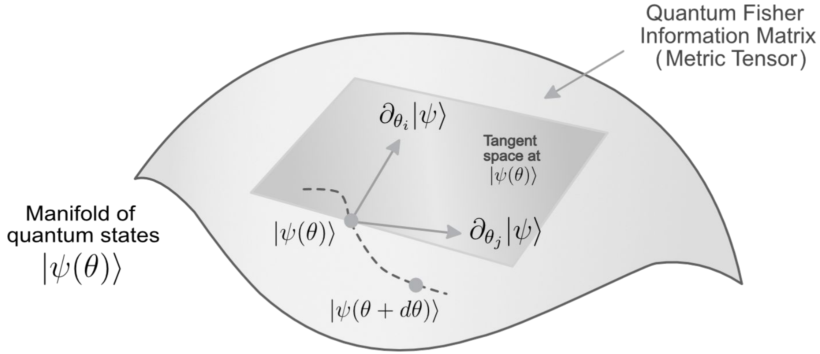
- Flexible encoding for jet constituent kinematics
  - Maps one particle → one qubit
  - 3 trainable parameters per qubit

- First demonstration of QML on (CMS) Open Data

Read the paper here

Karlsruhe Institute of Technology

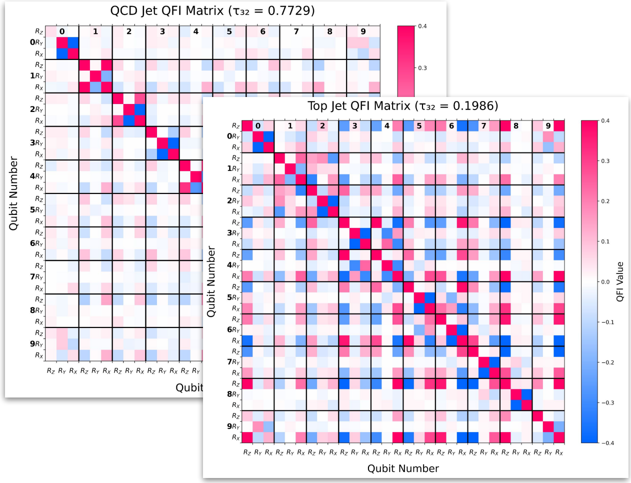
Quantum Geometry



$$\boxed{ds^2 = d\theta^T F(\theta)\, d\theta}$$

$$F_{ij} = 4\,\mathrm{Re}\left[\langle \partial_i \psi \mid \partial_j \psi \rangle - \langle \partial_i \psi \mid \psi \rangle \langle \psi \mid \partial_j \psi \rangle\right]$$

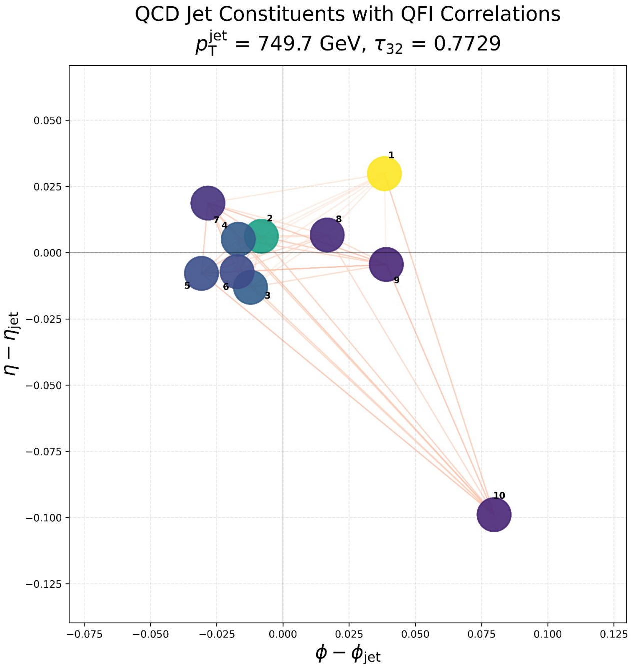
- Quantum Fisher Information Matrix (QFIM)
- Operating under a 1 particle → 1 qubit mapping
- Sensitivity of the final (entangled) state to individual parameters (or qubits)

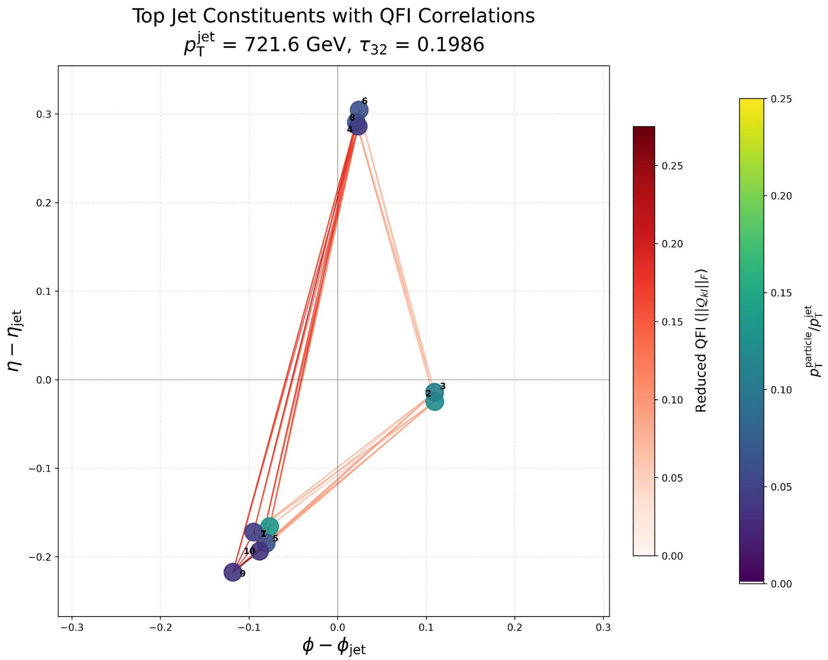


- QCD: off-diagonals suppressed
- Top: more populated QFIM

Karlsruhe Institute of Technology

Interpreting the QFI



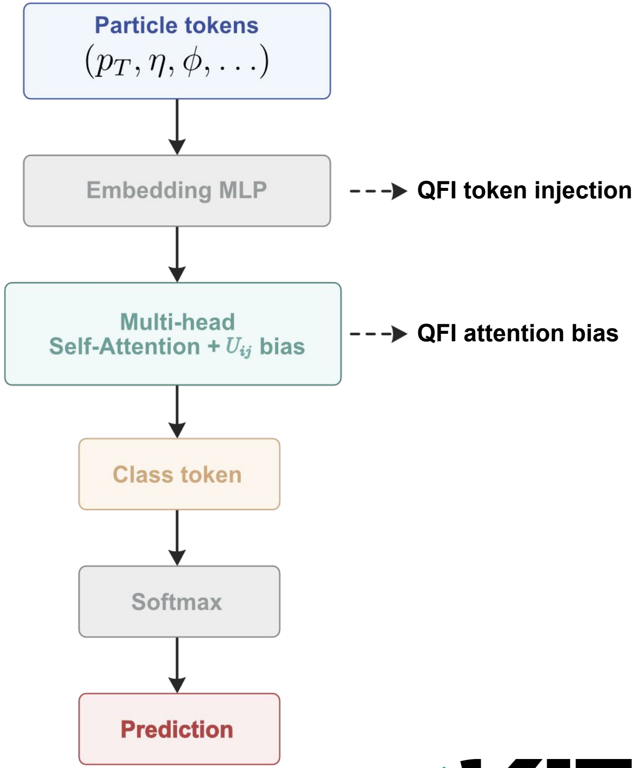


- Large QFI components → connect constituents of different *subjets*
- 3 parameters per qubit → 1 particle per qubit → One 3x3 block per qubit (particle) pair interaction

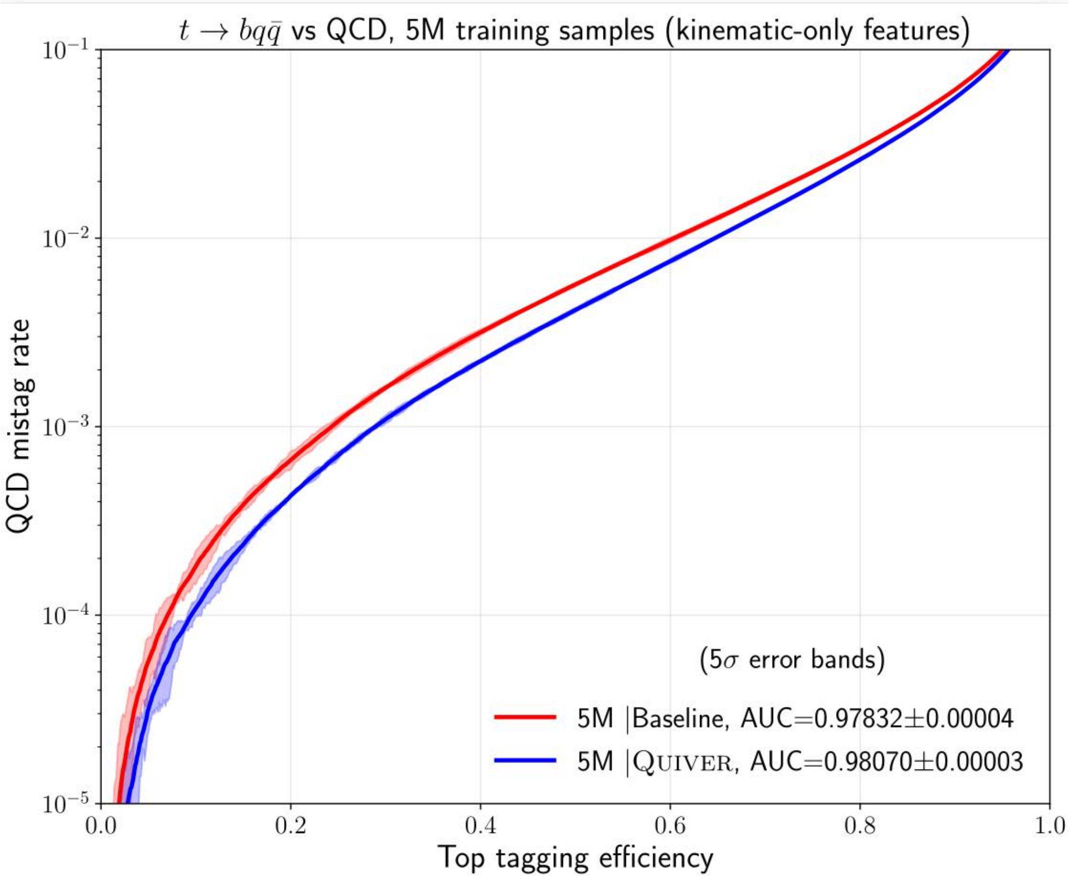
Karlsruhe Institute of Technology

The Particle Transformer (ParT)

- **State-of-the-art jet classifier**
    - ~2.2M trainable parameters
    - Transformer architecture on point clouds
    - **Benchmarks:**
        - Top Tagging
        - H→c$\bar{c}$
- **Baseline** architecture → use for QFIM augmentation
- QFIM → represent as additional (embedded) tokens and concatenate
    - Outperforms alternate augmentation schemes
- Applies to both a restricted kinematic input set and the full ParT feature set



Results



Karlsruhe Institute of Technology

Results

Features	Metric	Method	$N = 0.1\,\mathrm{M}$	$N = 0.5\,\mathrm{M}$	$N = 5\,\mathrm{M}$
Kinematic	AUC	ParT	0.97140 ± 0.00038	0.97629 ± 0.00015	0.97832 ± 0.00004
		QUIVER	$\mathbf{0.97368} \pm 0.00013$	$\mathbf{0.97848} \pm 0.00045$	$\mathbf{0.98070} \pm 0.00003$
	$1/\epsilon_B$	ParT	107 ± 2	146 ± 3	176 ± 1
		QUIVER	130 ± 1	191 ± 9	240 ± 1
Full	AUC	ParT	0.98875 ± 0.00008	0.99080 ± 0.00017	0.99235 ± 0.00003
		QUIVER	$\mathbf{0.98893} \pm 0.00005$	$\mathbf{0.99095} \pm 0.00003$	$\mathbf{0.99244} \pm 0.00003$
	$1/\epsilon_B$	ParT	570 ± 13	921 ± 13	1306 ± 8
		QUIVER	590 ± 7	951 ± 17	1362 ± 28

Karlsruhe Institute of Technology

Conclusion and Outlook

- Classical and quantum geometric quantities can inform the design and improve performance of jet taggers

- Quantum information → potentially complementary
  - Learns better representations
  - Improves performance of SOTA models like Particle Transformer

- Quantum-enhanced improvements → not necessarily limited to HEP

Read the preprint here

Presented at **ICML 2026,**
AI4Physics Workshop

Karlsruhe Institute of Technology

THANK YOU

Karlsruhe Institute of Technology