Hello folks,

Here's what you might find interesting today.

Anomaly Detection and Statistical ML at the LHC

NAE, Statistically -- Ranit Das, Jonathan Ostertag-Henning, Tilman Plehn, Lorenz Vogel

This paper tackles a real pain point in ML-based anomaly detection: the lack of a principled statistical interpretation for neural anomaly scores. The authors show that the normalized autoencoder (NAE) ties the anomaly score to a learned likelihood, and they validate this relation on a toy model and on jets using a dual-NAE setup. Crucially, they extend this to a Bayesian NAE that quantifies uncertainty on the learned likelihood itself. This is directly relevant to you, Aritra, given your interest in **anomaly detection at the LHC** and model-agnostic searches -- the probabilistic grounding of the anomaly score is exactly what's needed for deployment. And this might also catch your eye, Jingjing, since the Bayesian treatment of the NAE likelihood touches on **ML uncertainty quantification**, which is central to your interests. Note that Tilman Plehn is among the authors.

Distilling Normalizing Flows for Real-Time Anomaly Detection at the LHC -- Tara P. A. Tahseen, Noah Clarke Hall, Nikolaos Konstantinidis, Paula Martinez Suarez

This one is for you, Aritra. The authors address the latency bottleneck that has prevented normalizing-flow-based **anomaly detection** from running in hardware triggers. They distill the per-event likelihood from a large conditional normalizing flow into lightweight student models (decision trees and quantized neural networks) suitable for FPGA deployment. The conditional flow setup also handles missing input features gracefully. The combination of improved likelihood quality and sub-microsecond inference latency makes this a concrete step toward real-time ML anomaly detection in the trigger chain.

The Search Budget of the BSM Resonance Program -- Theresa Reisch, Tobias Golling

This paper quantifies the trials factor of the ATLAS BSM resonance program -- roughly 7.9 x 10^3 effective independent looks across 104 spectra -- and then asks what a fully combinatorial scan would cost. The key insight for ML-based searches is that when the trials factor cannot be enumerated (as in machine-learning-based anomaly detection), two-stage unblinding becomes essential to guard against spurious signals. This is relevant to you, Aritra, for **model-agnostic searches** and understanding the statistical cost of broad scans. Jingjing, you might also find this interesting for its discussion of statistical safeguards in ML-driven searches, which connects to your interest in rigorous inference methodology.

Jet Substructure and QCD Dynamics

Tracing Vacuum Hadronization with Conserved Currents -- Weiyao Ke, Hai Tao Li, Wanchen Li, Xiaohui Liu, Ding Yu Shao

This paper develops conserved-current observables -- net electric charge, baryon number, and strangeness flows within jets -- that directly probe QCD vacuum hadronization. The authors show these satisfy a Gell-Mann--Nishijima relation at the jet level, and that the net jet baryon number encodes the diquark-to-quark production ratio during hadronization. While there is no ML component, this is relevant to you, Aritra, because it deepens the **physics of jets** at the substructure level and defines flavor-resolved observables that could serve as targets or benchmarks for ML-based jet taggers. Understanding what quantum-number information is carried by hadrons in jets is foundational for designing features that ML models exploit.

The EEC-Hedron: Positivity Bounds on Energy Correlators -- Bianka Mejaj, Ian Moult, Hidde Stoffels, Matthew T. Walters, Yuan Xin

The authors derive positivity constraints on the partial-wave coefficients of energy-energy correlators (EECs), defining a bounded region they call the "EEC-hedron." For CFTs, these constraints translate into bounds on OPE coefficients, demonstrated in the 3d Ising and O(2) theories. This is tangentially relevant to you, Aritra, as EECs are a key **jet substructure** observable at the LHC, and understanding their fundamental theoretical constraints could inform what ML models trained on EECs are actually learning. It is more theoretical than applied, but the connection between positivity bounds and measurable jet observables is worth noting.

Back tomorrow with more!

P.S. In the most recent submission window: hep-ph: 30 papers hep-ex: 10 papers hep-th: 43 papers