Hey folks,

Here's what you might find interesting in today's arXiv batch. Fair warning: it's a relatively thin day for ML-in-HEP, but a few papers stand out.

Top Picks

Differentiable Simulation and Inference

[ADoNIS: A Differentiable generatOr of Neutrino Interaction Samples] by Cesar Jesus-Valls.

This might be interesting for you, Jingjing. ADoNIS is a fully differentiable neutrino interaction event generator that propagates exact gradients through the nuclear ground state, hard-scattering amplitudes, and the stochastic intranuclear cascade -- all following the physics choices of ACHILLES. The key novelty is that differentiability is achieved without sacrificing physical fidelity, and the resulting Jacobian directly exposes which measurements and phase-space regions constrain each parameter. This is directly relevant to simulation-based inference: differentiable simulators enable gradient-based posterior estimation and likelihood-free methods in a way that reweighting or finite-difference approaches fundamentally cannot. The authors also show applications to frequentist and Bayesian inference as well as unfolding through the detector response, hitting multiple topics on your list.

ML Interpretability for Particle Physics

[Finding and using interpretable latents in a neutrino foundation model with sparse autoencoders] by Raphael Bonnet-Guerrini, Johann Ioannou-Nikolaides, Inar Timiryasov, and Vincenzo Piuri.

This is interesting for you, Jingjing, but you might also find it relevant, Aritra. The authors apply sparse-autoencoder-based mechanistic interpretability to a neutrino foundation model pretrained on IceCube data and fine-tuned for direction reconstruction. They identify a validated "atlas" of physical concepts in the model's latent representation using a strict protocol of held-out tests, nuisance controls, and replication across independent dictionary trainings. The striking result is that the direction head barely uses this atlas -- but an uncertainty head trained on the same representation depends causally on quality and brightness features from it, improving the median angular resolution from 20.2 degrees to 3.2 degrees at 20% selection efficiency. This is a first application of mechanistic interpretability tooling to particle physics and touches on ML uncertainty quantification in a novel way, Jingjing. For you, Aritra, it is a concrete example of probing the internal physics knowledge of a trained model using information-theoretic tools, which connects to your interest in understanding what ML models learn.

Jet Substructure and QCD

[Color Coherence and the Soft Structure of QCD Jets in Vacuum and the QGP] by Paul Caucal and Yacine Mehtar-Tani.

This might be interesting for you, Aritra, in the context of understanding the physics of jets -- though there is no ML component here. The paper formulates a theoretical framework for the evolution of QCD jets in vacuum and in the quark-gluon plasma via resummation of large energy logarithms, showing that near-threshold jet observables can be expressed in terms of Wilson-line correlators obeying BMS evolution. The key insight is that soft radiation resolves the internal color structure of the jet, producing a hierarchy of nonlinear evolution equations governing color coherence and decoherent energy loss. In the large-Nc limit the evolution is closely related to the Balitsky-Kovchegov equation, connecting saturation physics to jet quenching. This is the kind of jet-substructure physics that underpins the observables your ML architectures operate on, and understanding these coherence effects is relevant for designing features or inductive biases for jet tagging.

Unfolding and Detector Effects

A Weighting Method for Incorporating Mass Resolution Effects in Amplitude Analysis by Benhou Xiang, Wenqian Zheng, Hongxun Yang, Xiaolin Kang, and Shuangshi Fang.

This could be relevant to you, Jingjing, for its connection to unfolding and detector response handling. The method constructs event-by-event weights from truth-reconstruction correspondences in fully simulated MC samples to map theoretical amplitudes from true phase space to the smeared observational space, avoiding explicit multidimensional convolution or on-the-fly detector simulation. It naturally handles nonuniform resolution across the Dalitz plane and can be extended to momentum and angular variables. While this is not ML-based, the weighting approach is conceptually adjacent to density ratio estimation and could inspire ML-based unfolding or deconvolution methods that similarly exploit truth-reconstruction pairings.

Higgs Pair Production Phenomenology

Pair production of h in the U(1)_XSSM by Yue-Tong Liu, Shu-Min Zhao, Meng-Zi Cao, Shuang Di, Rong-Zhi Sun, and Xing-Xing Dong.

Noting this for you, Jingjing, as it touches on HH phenomenology at the LHC -- though without an ML component. The paper studies lightest neutral Higgs pair production via gluon fusion at 14 TeV in a U(1) extension of the MSSM, analyzing parameter dependence of the cross section and identifying the gauge couplings as the most sensitive parameters. The model yields sizable new physics corrections under current experimental constraints. It is a useful phenomenology reference for HH sensitivity studies, but the lack of any ML methodology means it is only tangentially relevant to your core interests.

That's the lot for today. The pickings are slim on the ML side -- no jet tagging, anomaly detection, Lorentz-equivariant, or agentic AI papers in this batch. The sparse-autoencoder paper and ADoNIS are the standouts worth your time.

Back tomorrow with more.

P.S. In the most recent submission window: hep-ph: 31 papers hep-ex: 14 papers hep-th: 28 papers