Hi all,

Here's what you folks might find interesting today.

Jet Flavor Classification and Substructure

Simplex Demixing: Disentangling Multiple Light-Flavor Jets at Colliders -- Gregorio de la Fuente, **Jesse Thaler**

This paper introduces "simplex demixing," an ML framework that generalizes the quark/gluon jet topic approach to T jet flavors extracted from M data mixtures. The method translates a multi-category classifier output into a bounded geometric object (a simplex) whose vertices correspond to maximally separable flavor categories, requiring minimal assumptions about the underlying mixture composition. The authors validate the approach on synthetic down/up/gluon mixtures and propose a realistic tag-and-probe strategy for dijet events at the LHC. This might be interesting for you, Aritra, because it directly addresses jet tagging with novel ML methodology and pushes the understanding of jet flavor structure beyond the binary quark-vs-gluon paradigm. The identifiability analysis -- showing how flavor abundance and hadron-level information control separability -- is a nice example of using information-theoretic reasoning to understand the physics limits of jet classification. Having Jesse Thaler as an author also signals this is likely to be a well-motivated, physically grounded contribution.

ML for PDF Reconstruction and Uncertainty Quantification

Normalizing Flows to Reconstruct Pseudo-PDFs -- Yamil Cahuana Medrano, Kostas Orginos

This work combines Gaussian Process priors with invertible neural networks (normalizing flows) to learn a posterior distribution over parton distribution functions from limited Ioffe-time data, preserving physical constraints and extrapolation behavior. This might be interesting for you, Jingjing, as it touches on ML uncertainty quantification in a particle physics context -- the posterior-over-functions framing is conceptually adjacent to simulation-based inference and density estimation techniques you work with. The architecture's ability to maintain physical constraints while providing full posterior distributions rather than point estimates is a useful methodological reference, even if the application domain (lattice-extracted PDFs) is somewhat outside your usual LHC focus.

Bayesian Inference in Heavy-Ion Physics

Path-length dependence of parton energy loss across collision systems: a Bayesian analysis -- Fouad A. Majeed et al.

This paper performs a joint Bayesian analysis of charged-particle RAA across four collision systems (O+O through Pb+Pb) using nested sampling for model selection, extracting a path-length scaling exponent consistent with radiative energy loss. You might find this relevant, Jingjing, as it demonstrates Bayesian inference machinery (nested sampling, Bayes-factor model comparison) applied to a particle physics problem, though it is not simulation-based inference in the neural network sense. The systematic exploration of 160 analysis variants and the degeneracy discussion between medium density and path length are methodologically instructive for thinking about identifiability challenges in inference problems more broadly.

A relatively quiet day for your core topics -- the Simplex Demixing paper is the standout. Back tomorrow with more!

P.S. In the most recent submission window: hep-ph: 49 papers hep-ex: 18 papers hep-th: 42 papers