Hi all,

Here's what you folks might find interesting today.

TOP PICKS

Agentic AI and workflow automation

Articulating Assumptions in AI-Generated Scientific Analyses through Task Decomposition
This is the clearest fit for workflow automation in today's list. The paper proposes a *multi-agent* setup that splits LLM-based analysis into code generation, execution, tracing, validation, and ambiguity checking, rather than treating the whole analysis as one opaque prompt. That matters because the authors are targeting exactly the failure mode that makes AI-generated collider analyses hard to trust: you can run the code, but still not know which assumptions actually drove the result. This might be interesting for you, Jingjing, because it is directly about making agentic analysis pipelines auditable and reproducible for particle-physics tasks, and the ambiguity-rewrite module is especially relevant if you are thinking about automating pieces of MC production or analysis code writing. It is also useful in a broader collider-analysis sense, since they validate on representative collider physics examples rather than staying at the toy-problem level. The novelty here is not "LLMs for physics" in general, but the explicit decomposition into inspectable sub-tasks plus semantic differencing of intended versus implemented quantities.

Di-Higgs and collider phenomenology

ggxy: Fast and flexible NLO QCD corrections to gluon-initiated processes
This one is a strong phenomenology pick because it includes *gg -> HH* with full top-mass dependence at NLO QCD, and packages it in a reusable C++ library rather than just presenting fixed benchmark numbers. That makes it relevant for you, Jingjing, since di-Higgs sensitivity studies often run into the practical bottleneck of needing fast, flexible theory predictions when scanning couplings, masses, or shower-matching choices. The paper also adds *gg -> ZH* and *gg -> ZZ*, so it is useful for signal-background consistency studies in Higgs-related analyses. The Powheg interface is an important detail: it means the tool is meant to be used in realistic event-generation workflows, not just as a formal amplitude result. What is novel here is the combination of analytical approximations for the two-loop virtuals across kinematic regions with exact treatment of the remaining ingredients, giving speed without completely giving up fidelity. There is no explicit ML component, so I would not rank it above ML-native papers in a richer day, but for HH-focused work it is still worth a look.

Jet tagging and LHC ML

Simultaneous efficiency measurements of b- and c-jets in ttbar events from sqrt(s)=13.6 TeV pp collision data collected with the ATLAS detector
This is probably the most relevant experimental paper for you today, Aritra. ATLAS is measuring the performance of the *single-transformer-based GN2 tagger* now used in production, with simultaneous extraction of b-tag efficiency and c-jet mistag rates in semileptonic ttbar events. Even though this is a calibration paper rather than a new architecture paper, it is valuable because it shows how a transformer tagger is behaving in real Run 3 data, including the precision actually achieved as a function of jet pT and tagger output. That kind of deployment-level information matters if you care about where modern jet-tagging architectures are genuinely buying performance versus where systematics and calibration dominate. The use of a kinematic-likelihood reconstruction to map jets to decay products is also a reminder of the analysis machinery needed to make these ML taggers usable in practice. This is not a jet-substructure interpretation paper, but it is still directly connected to the current state of ML-based heavy-flavor tagging at the LHC.

BSM searches with heavy-flavor final states

Searching for charged Higgs bosons in top decays via the t*b channel
This is not an ML paper, but it is collider-relevant enough to be worth flagging because it targets a relatively neglected topology: low-mass charged Higgs decays through the off-shell *H+/- -> t*b *mode. The analysis is built by reinterpreting ATLAS dileptonic measurements in a* ttbar bbbar-*like final state, which makes it potentially useful as a benchmark topology for heavy-flavor-rich searches. This might be interesting for you, Aritra, because the final state overlaps with the kind of complex hadronic environments where tagging, combinatorics, and possibly anomaly-style search strategies can matter. It may also be of some interest to you, Jingjing, from the perspective of reinterpretation methodology and sensitivity studies, though there is no SBI or ML angle in the abstract. A small author boost applies here as* Benjamin Fuks* is on the paper. The main novelty is the focus on an off-shell decay channel below the top threshold that has been largely unconstrained, together with explicit model-independent limits and interpretations in several 2HDM scenarios.

Other papers worth a quick skim

Formal methods and AI reliability in physics

Axioms for physical reasoning: codifying the Seiberg-Witten solution in Lean
This is outside collider ML proper, so not a top-priority read, but it does connect to the trustworthiness of AI-assisted scientific reasoning. The paper's key idea is to separate what is formally proved from what is accepted as explicit physical postulates, so downstream claims can be traced back to named assumptions. That is conceptually adjacent to the task-decomposition paper above, though here the setting is theorem proving and formalization rather than analysis automation. This might be mildly interesting for you, Jingjing, if you are thinking about robust AI-assisted workflows where provenance of assumptions matters. It is more about methodology than direct particle-physics analysis.

Back tomorrow with more!

P.S. In the most recent submission window: hep-ph: 38 papers hep-ex: 15 papers hep-th: 39 papers