QCD Theory Meets Information Theory

Benoît Assi

based on: 2604.00084, 2602.01509, 2501.17219, 2307.00728

IRN Terascale, Apr 21, 2026

Laboratoire
de Physique
des 2 infinis
Irène Joliot-Curie

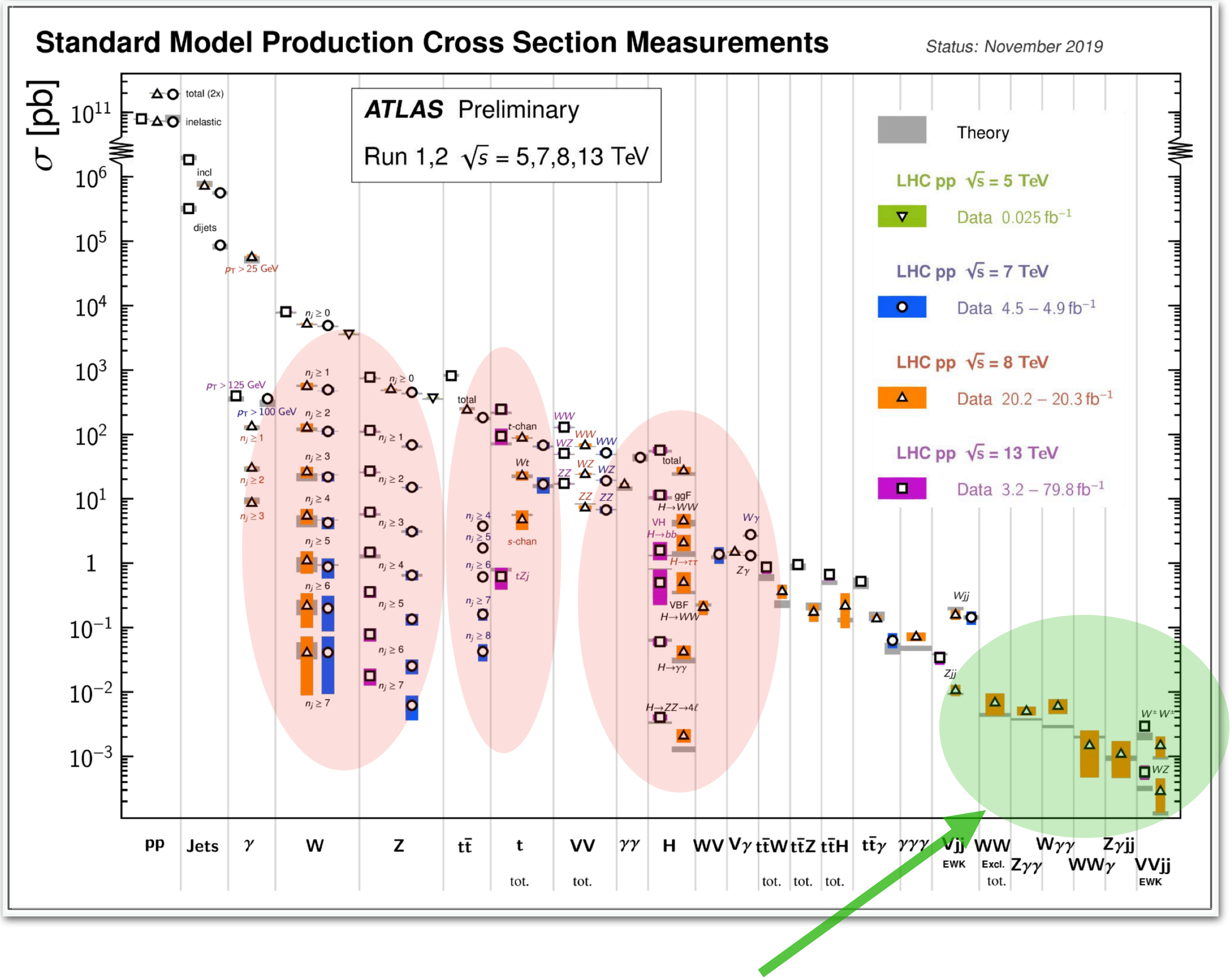
High-Luminosity LHC era

Theory predictions and measurements reaching **incredible level of precision**

Increasing stats means **theory uncertainties** will dominate for many processes and observables

We will need **higher-order precision** SM predictions!

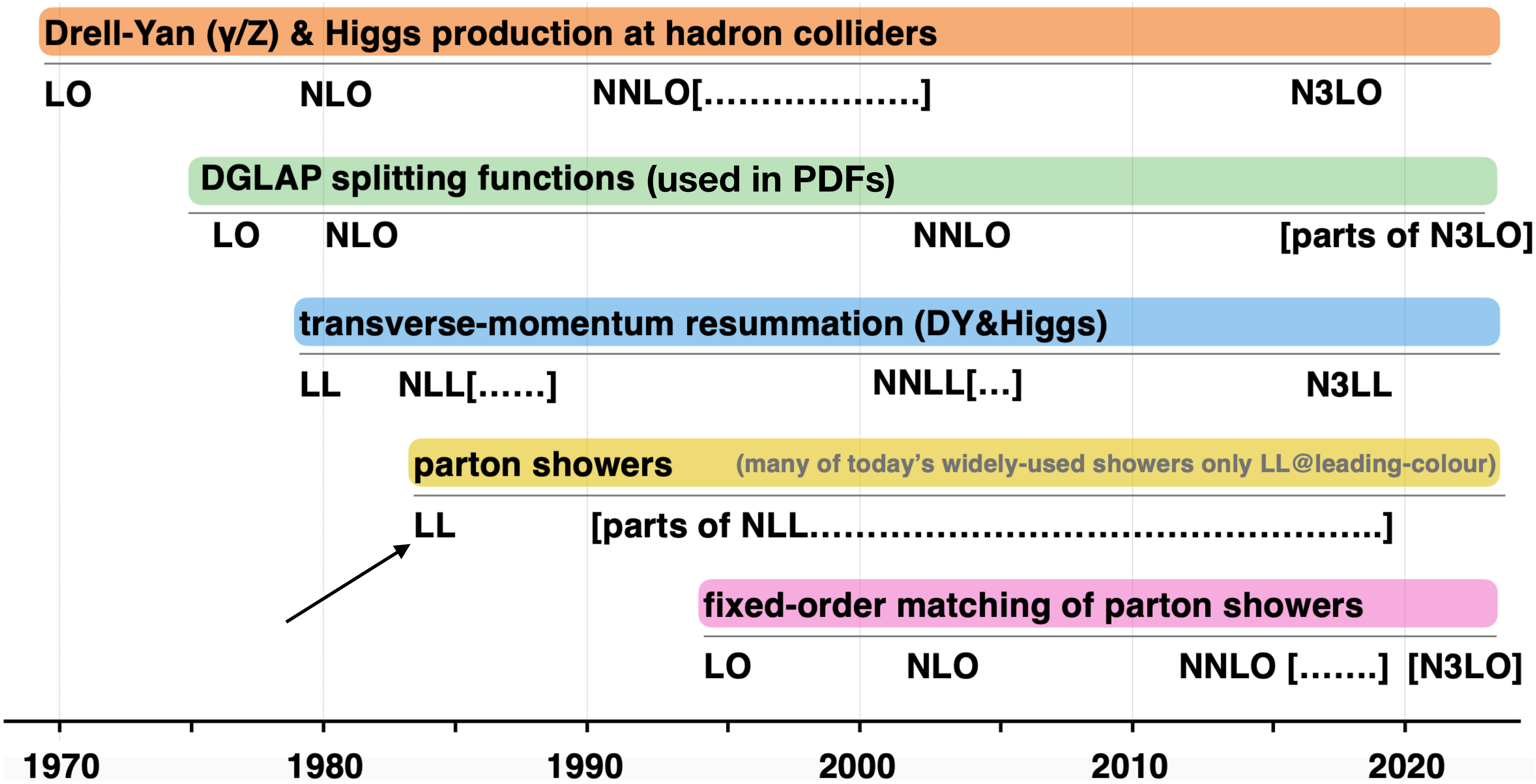
[https://twiki.cern.ch/twiki/bin/view/AtlasPublic/StandardModelPublicResults]



VBS and multi-boson tails — TeV-scale probe of EWSB and gauge self-interactions hidden under **QCD backgrounds**

Quantifying perturbative precision

Current status of (resummed) perturbative precision



Drell-Yan (γ/Z) & Higgs production at hadron colliders

LO NLO NNLO[…………………] N3LO

DGLAP splitting functions (used in PDFs)

LO NLO NNLO [parts of N3LO]

transverse-momentum resummation (DY&Higgs)

LL NLL[……] NNLL[…] N3LL

parton showers (many of today's widely-used showers only LL@leading-colour)

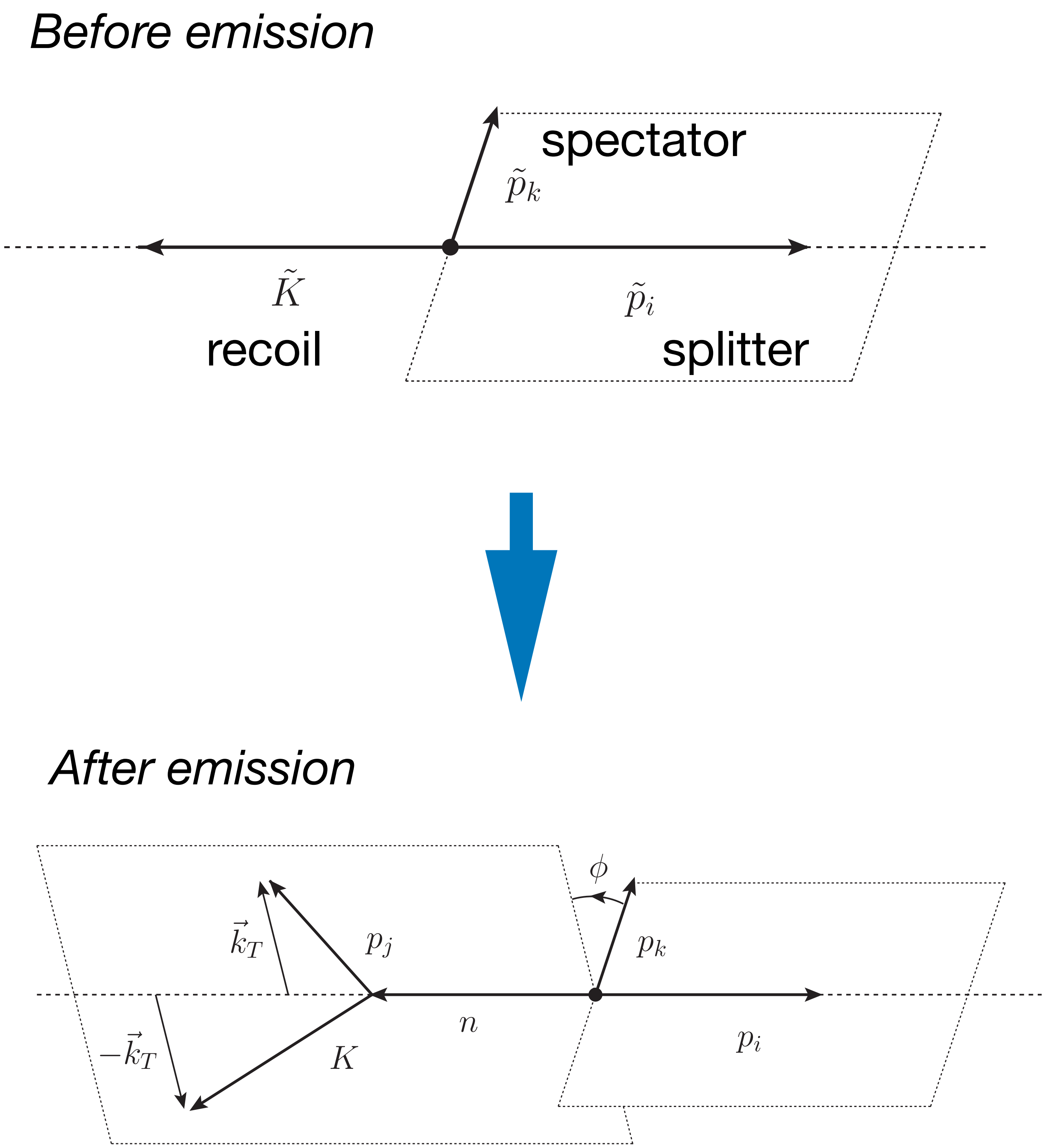
LL [parts of NLL…………………………………………]

fixed-order matching of parton showers

LO NLO NNLO [.......] [N3LO]

G. Salam, Moriond 2023

A provably NLL accurate algorithm in SHERPA


Before emission

After emission

QCD factorisation violated by recoil schemes introducing spurious correlations between emissions at different scales

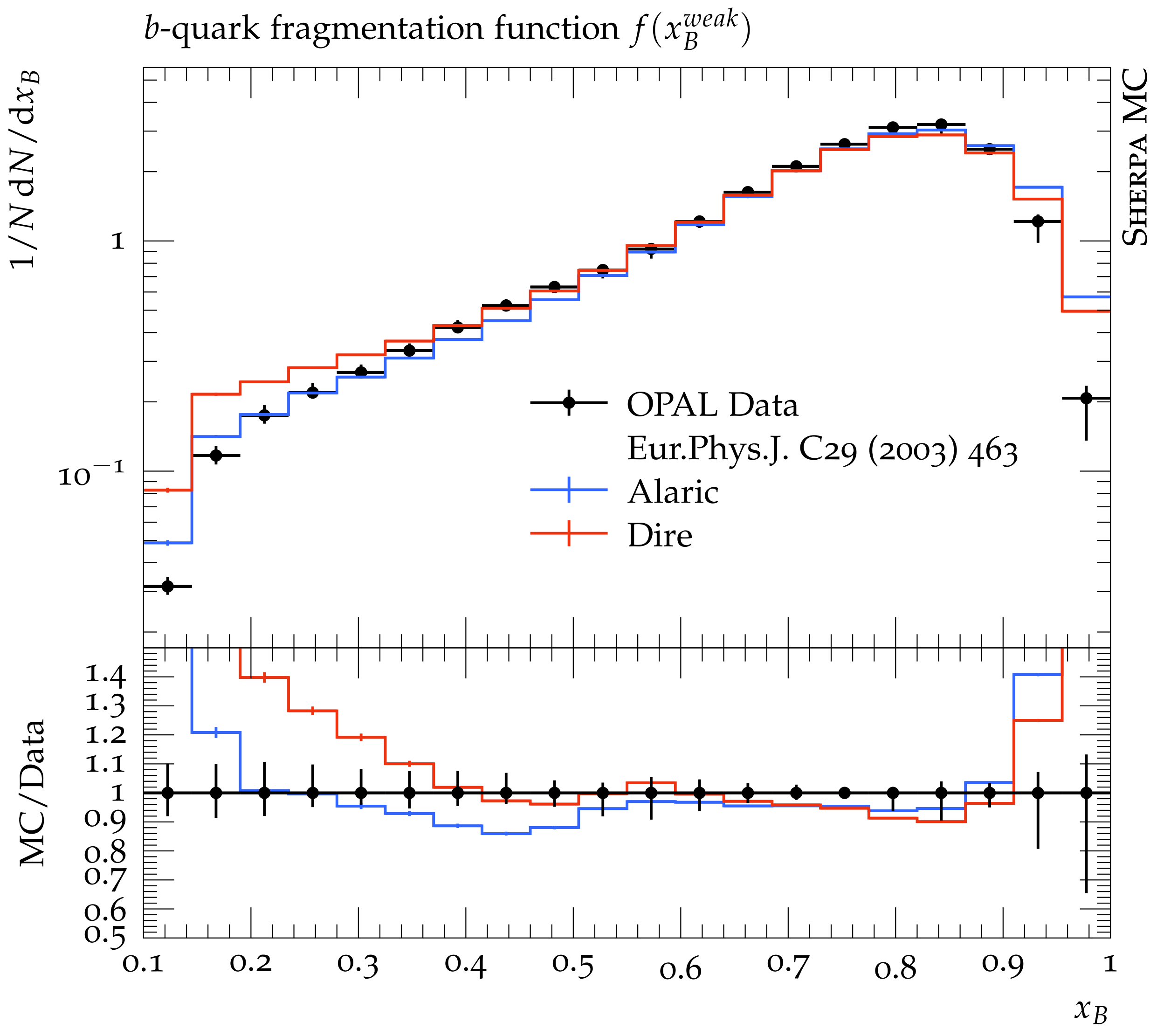
ALARIC: recoil compensated **globally** by neg. sum of multipole momenta $\tilde{K}$

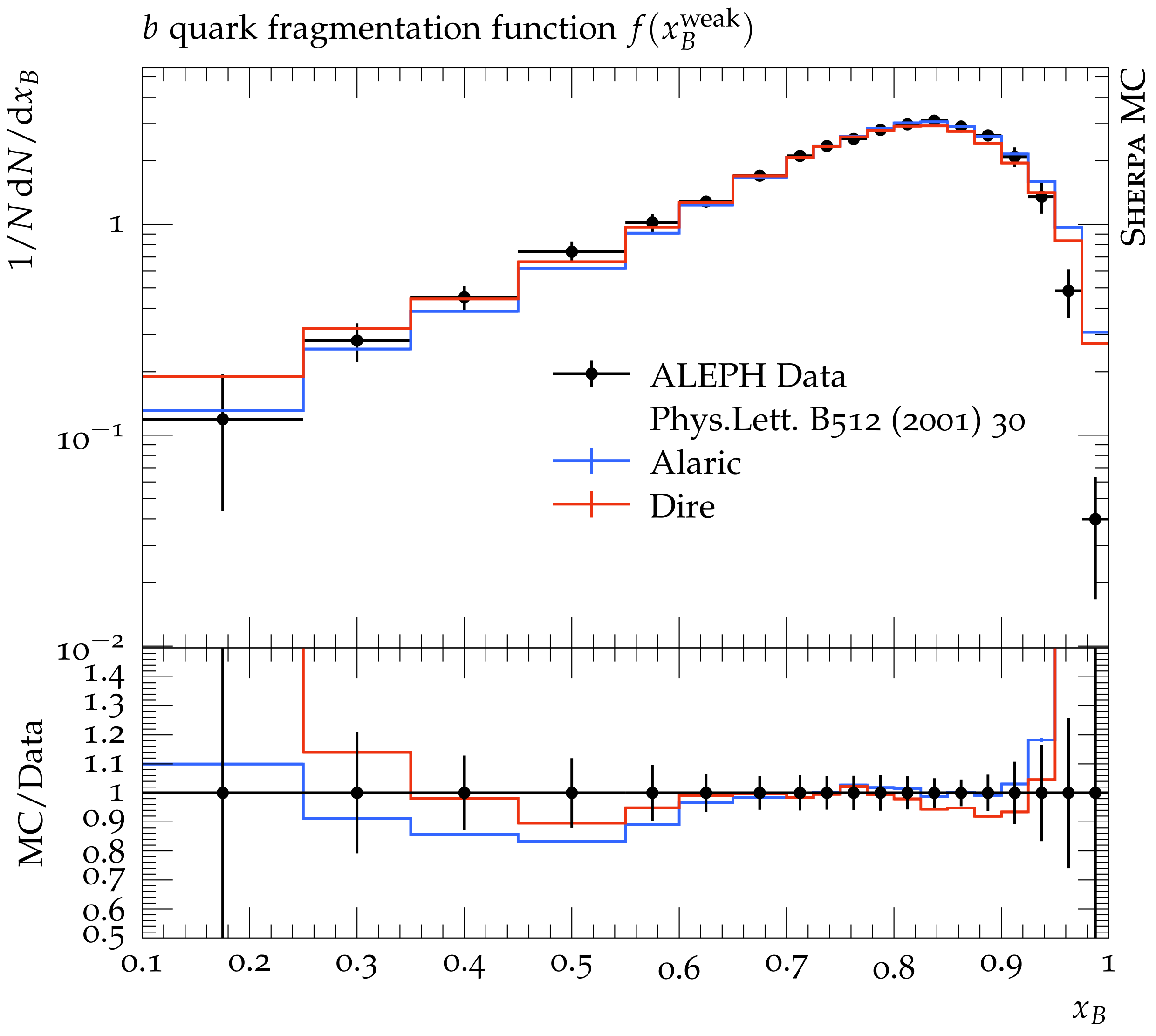
By construction $K^2 \gg k_\perp^2$ and does **not** affect topology of **previous emissions** as its effect scales as $k_\perp^2/K^2$

Uniquely analytically provable for both massless and **massive** partons that this algorithm is NLL accurate

Herren, Höche et al. JHEP 10 (2023)
BA and Höche PRD 109 (2024)

Massive Improvements





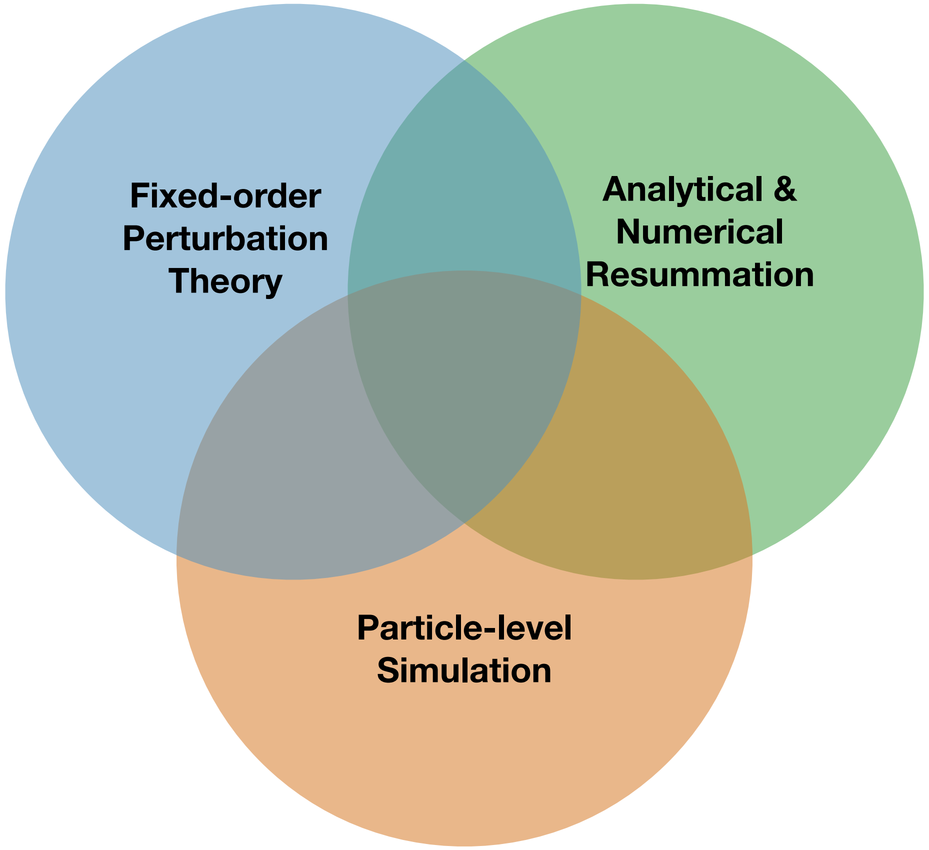
Better agreement for Alaric predictions with experimental data (same hadronization tune)

Can we do this better with ML?

ALARIC is NLO matched in both **massive** and massless case and now **default shower** in SHERPA

BA and Höche PRD 109 (2024)

Can we incorporate higher precision theory information?

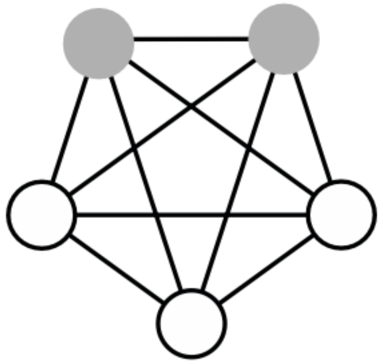


Fixed-order Perturbation Theory

Analytical & Numerical Resummation

Particle-level Simulation

QCD Theory meets Information Theory



Boltzmann machine: Generative model samples according to stat dist

$$P(\mathbf{s}) \sim e^{-E/T}, \quad E = -\sum_{i<j} w_{ij} s_i s_j - \sum_i \theta_i s_i$$

Boltzmann factor: Maximum entropy solution subject to constraint on average energy

$$p(x) = \overset{\text{Prior}}{\overset{\downarrow}{q}}(x) \exp\left[- \overset{\text{Sets Partition Function}}{\overset{\downarrow}{\beta_0}} - \underset{\text{Lagrange Multipler}}{\underset{\uparrow}{\beta_1}} \underset{\text{Constrained Quantity}}{\underset{\uparrow}{f_1}}(x) - \beta_2 f_2(x) - \ldots \right]$$

Sudakov structure: Plucking a random event-shape observable distribution

$$r(\tau)_{\text{LL,f.c.}} = \frac{-2\alpha_s C_F}{\pi} \frac{\ln \tau}{\tau} \underset{\text{Sudakov factor}}{\exp\left[-\frac{\alpha_s C_F}{\pi} \ln^2 \tau \right]}$$

ML practitioner: Sudakov = Boltzmann factor and cusp AD = Lagrange multiplier that enforces constraint on the 2nd log moment of distribution

log moments not previously measured or calculated before in QCD!

BA, Höche, Lee, Thaler PRL 132 (2025)

QCD Theory meets Information Theory

Consider **relative entropy:**

$$-D_{\mathrm{KL}}(P\|Q) = \int dx\, p(x) \log \frac{q(x)}{p(x)} + (\beta_0 - 1)\Big(1 - \int dx\, p(x)\Big)$$

Plus **moment constraints:**

$$+\sum_j \beta_j \Big(c_j - \int dx\, p(x)\, f_j(x)\Big)$$

Extend to all known observables at once**:**

$$c_j[r] = \int d\vec{v}\, r(\vec{v})\, g_j(\vec{v}) \ , \quad d_j[p] = \int d\Phi\, p(\Phi)\, g_j\big(\vec{v}(\Phi)\big)$$

Target moment **Current moment**

Minimized if prior (generator) distribution is modified as:

$$p(\Phi) = q(\Phi)\, w(\Phi) \ , \quad w(\Phi) \equiv \exp\Big[-\sum_j \lambda_j g_j\big(\vec{v}(\Phi)\big)\Big]$$

Note: Efficient procedure as *convex* problem + weight is event-level **afterburner** + weights strictly positive

Simplified thrust calculation

Leading Log at fixed coupling

$$p(\tau) = \frac{-2\alpha_s C_F}{\pi}\frac{\ln\tau}{\tau}\exp\left[-\frac{\alpha_s C_F}{\pi}\ln^2\tau\right]$$

Ordinary Moments

$$\langle\tau\rangle = \frac{2\alpha_s C_F}{\pi} + \mathcal{O}(\alpha_s^2)$$

Mean ⇒ Strong coupling

Characterises *FO information*

Logarithmic Moments

$$\langle\ln\tau\rangle = -\frac{\pi}{2\sqrt{\alpha_s C_F}}$$

Logarithmic Mean ⇒ Sudakov peak

Characterises *Resummed information*

Logarithmic moments of thrust at LEP

Generic Sudakov-log observable v takes the form

$$r(v) = \delta(v) + \sum_{m=1}^{\infty} \sum_{n=1}^{2m-1} k_{mn}^{\mathrm{LP}} \left(\frac{\alpha_s}{4\pi}\right)^m \left[\frac{\ln^n v}{v}\right]_+ + \cdots + \sum_{m=1}^{\infty} \sum_{n=1}^{2m-1} k_{mn}^{\mathrm{N^kLP}} \left(\frac{\alpha_s}{4\pi}\right)^m \frac{\ln^n v}{v} v^{k-1}$$

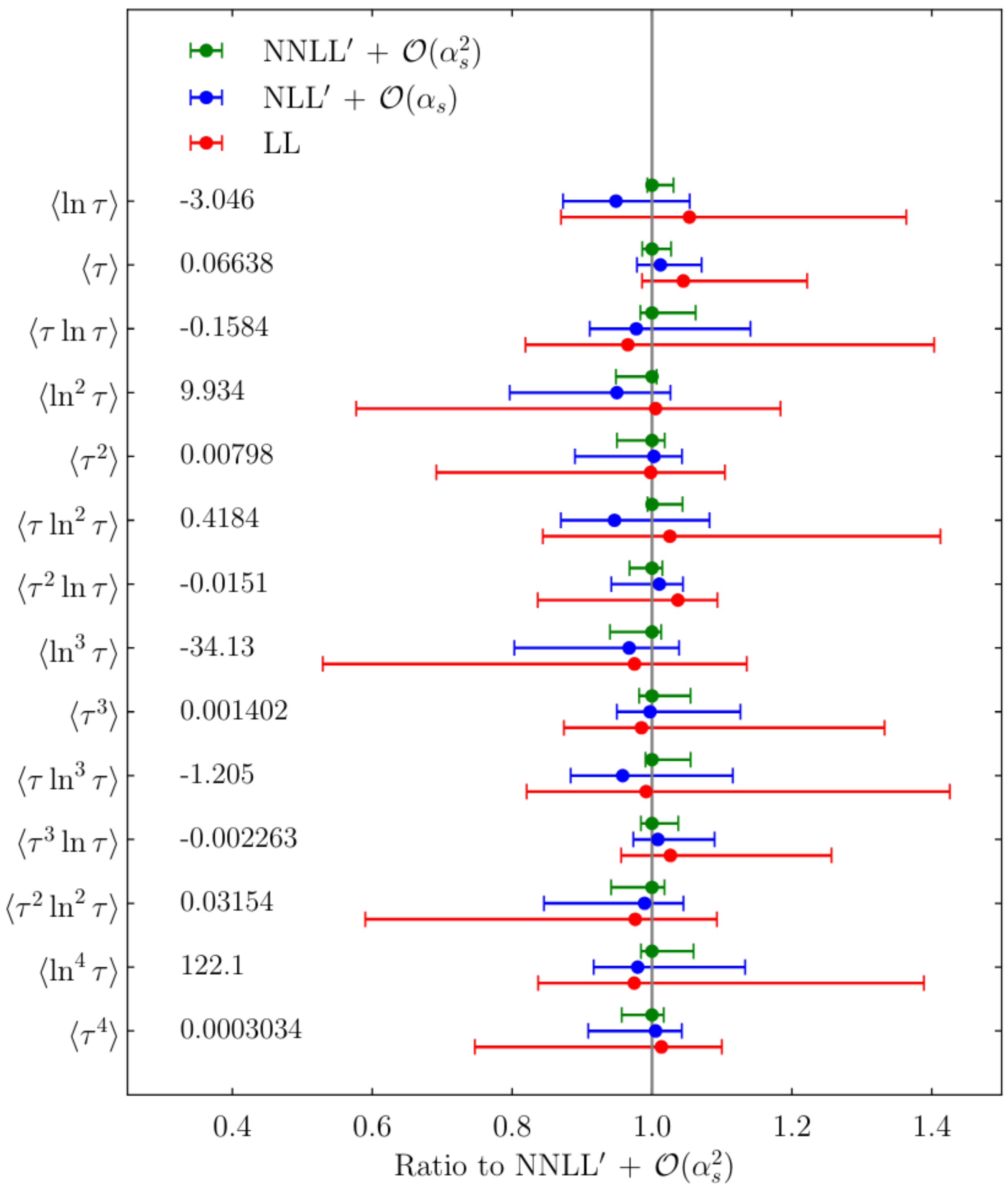
This motivates choosing **basis functions** for thrust of the form

$$g_{mn}(\tau) = \tau^m \ln^n \tau$$

Moments calculable to very **high precision**

Future: Moments of basis functions with support across entire distribution provided alongside precision observable calculations

First QCD calculation of thrust logarithmic moments



BA, Höche, Lee, Thaler, PRL 132 (2025)

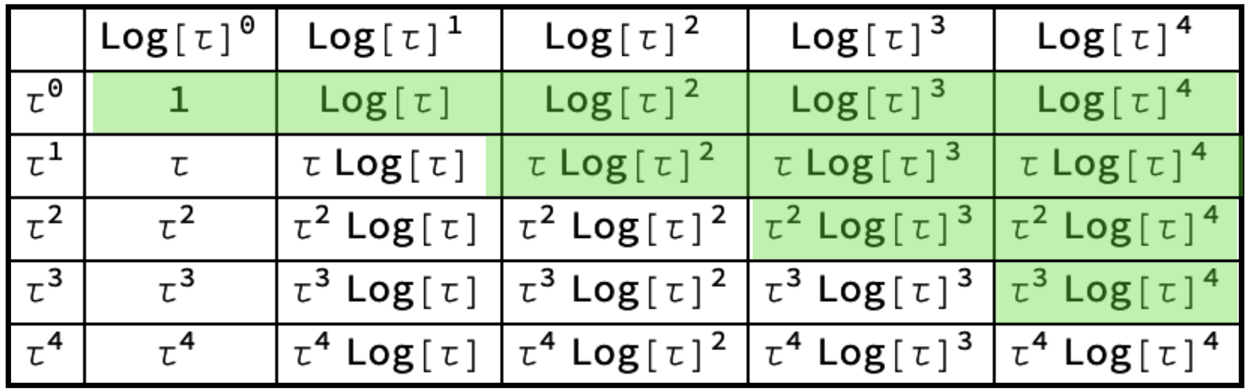
see related moment constructions in Desai, Nachman, Thaler: 2407.11284
and more recently for uncertainty estimation in Gambhir, Mastandrea: 2512.04160

An example of thrust at LEP

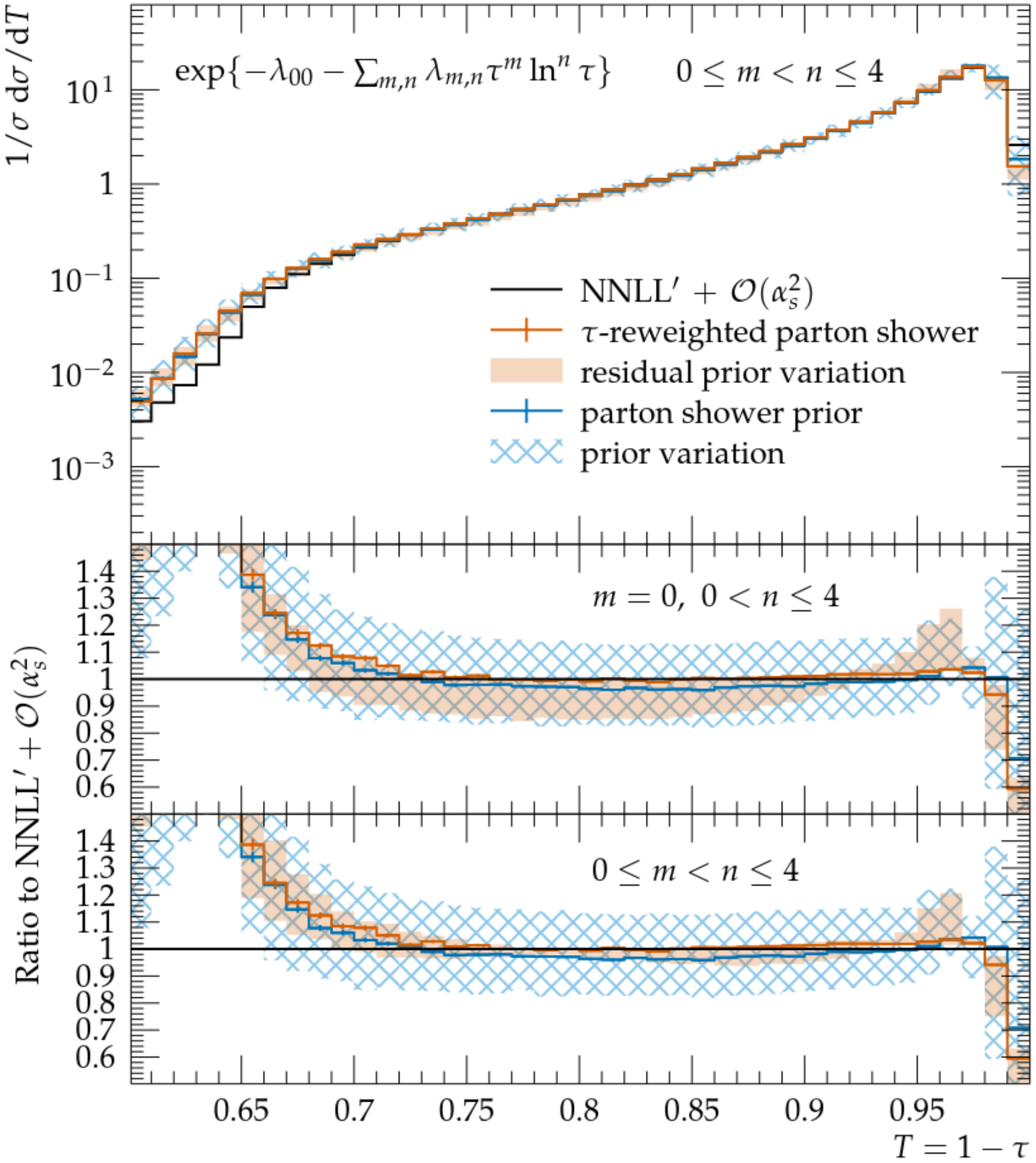
Small subset of mixed log moments sufficient to constrain bulk of thrust distribution

Systematically improvable with more **well-chosen** moments

Choice of moment basis relevant to improve **different physical regions** (2-jet/3-jet region, Sudakov peak, Sudakov shoulder,…) and largely **independent** of MC priors

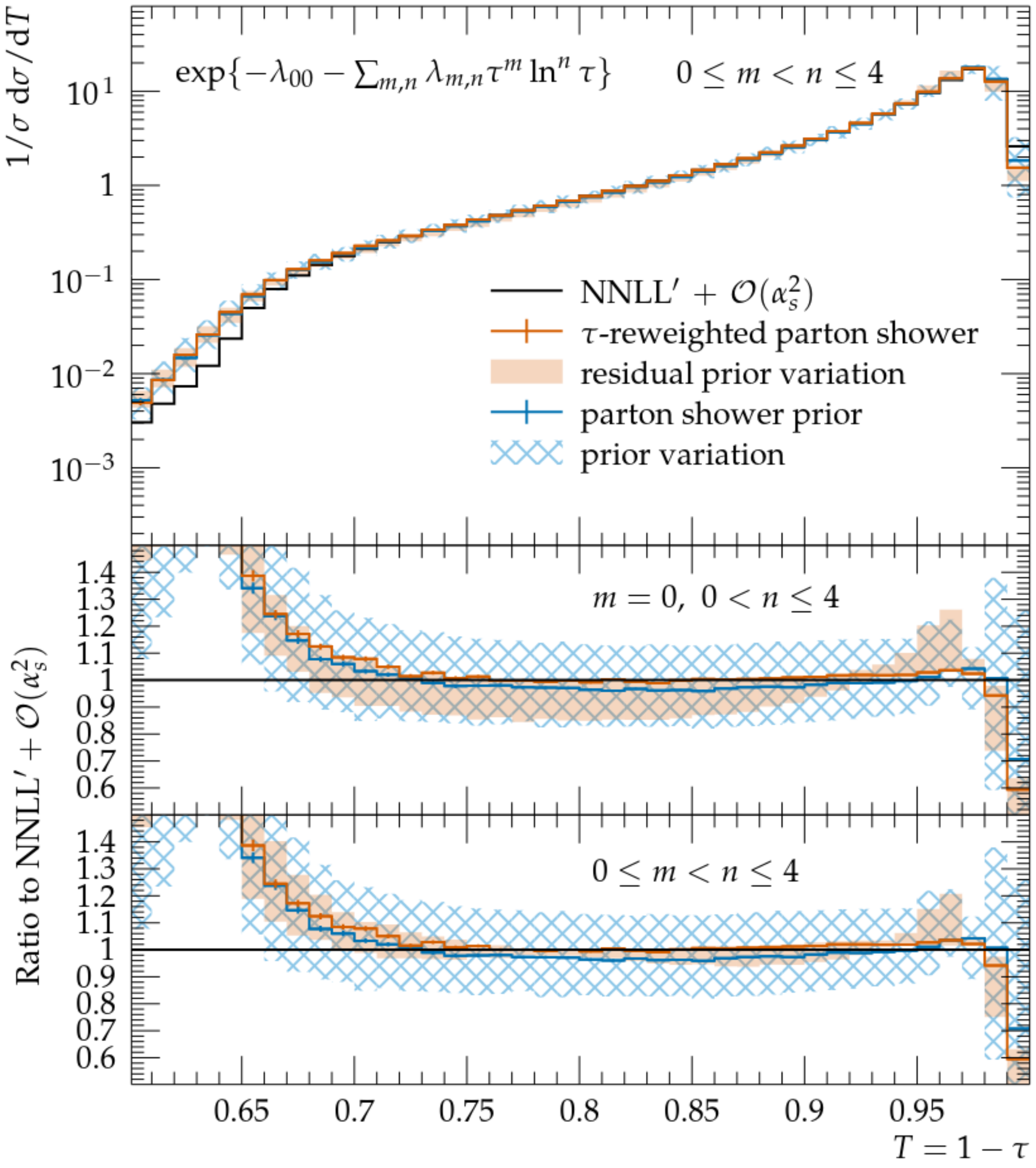


	$\mathrm{Log}[\tau]^0$	$\mathrm{Log}[\tau]^1$	$\mathrm{Log}[\tau]^2$	$\mathrm{Log}[\tau]^3$	$\mathrm{Log}[\tau]^4$
τ^0	🟩 1	🟩 $\mathrm{Log}[\tau]$	🟩 $\mathrm{Log}[\tau]^2$	🟩 $\mathrm{Log}[\tau]^3$	🟩 $\mathrm{Log}[\tau]^4$
τ^1	τ	$\tau\,\mathrm{Log}[\tau]$	🟩 $\tau\,\mathrm{Log}[\tau]^2$	🟩 $\tau\,\mathrm{Log}[\tau]^3$	🟩 $\tau\,\mathrm{Log}[\tau]^4$
τ^2	τ^2	$\tau^2\,\mathrm{Log}[\tau]$	$\tau^2\,\mathrm{Log}[\tau]^2$	🟩 $\tau^2\,\mathrm{Log}[\tau]^3$	🟩 $\tau^2\,\mathrm{Log}[\tau]^4$
τ^3	τ^3	$\tau^3\,\mathrm{Log}[\tau]$	$\tau^3\,\mathrm{Log}[\tau]^2$	$\tau^3\,\mathrm{Log}[\tau]^3$	🟩 $\tau^3\,\mathrm{Log}[\tau]^4$
τ^4	τ^4	$\tau^4\,\mathrm{Log}[\tau]$	$\tau^4\,\mathrm{Log}[\tau]^2$	$\tau^4\,\mathrm{Log}[\tau]^3$	$\tau^4\,\mathrm{Log}[\tau]^4$

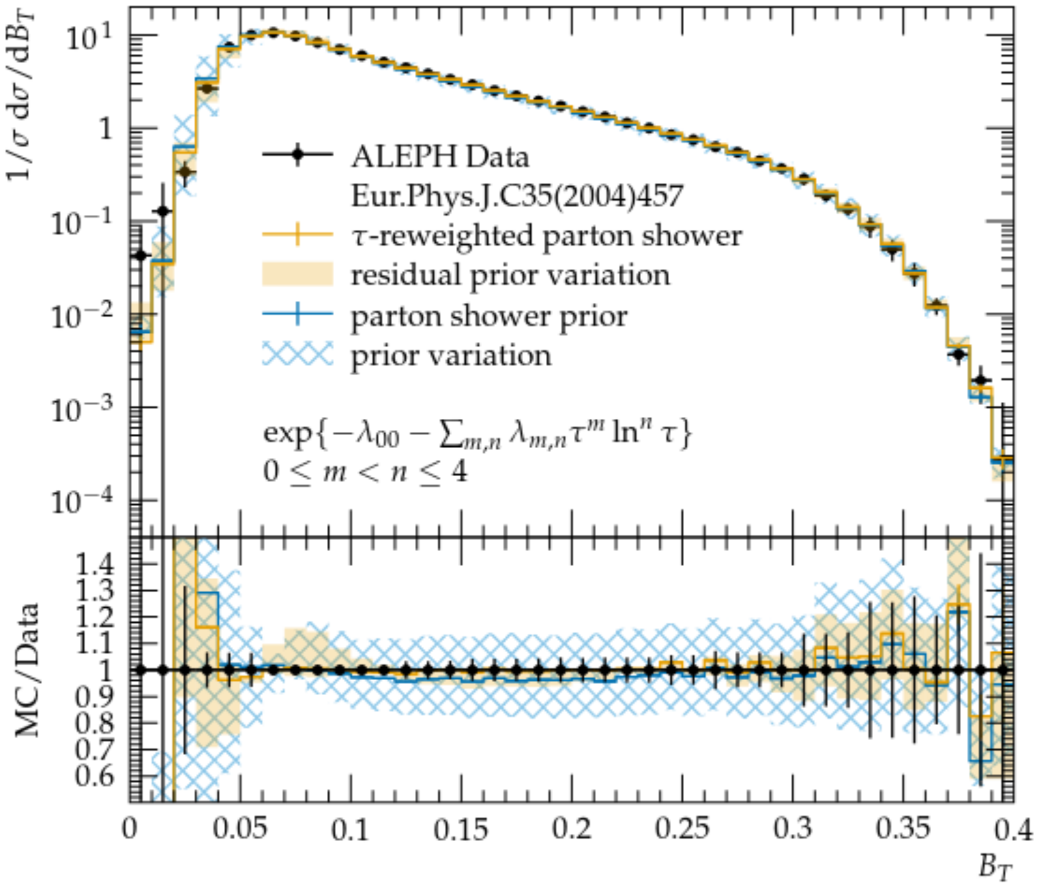


BA, Höche, Lee, Thaler PRL 132 (2025)

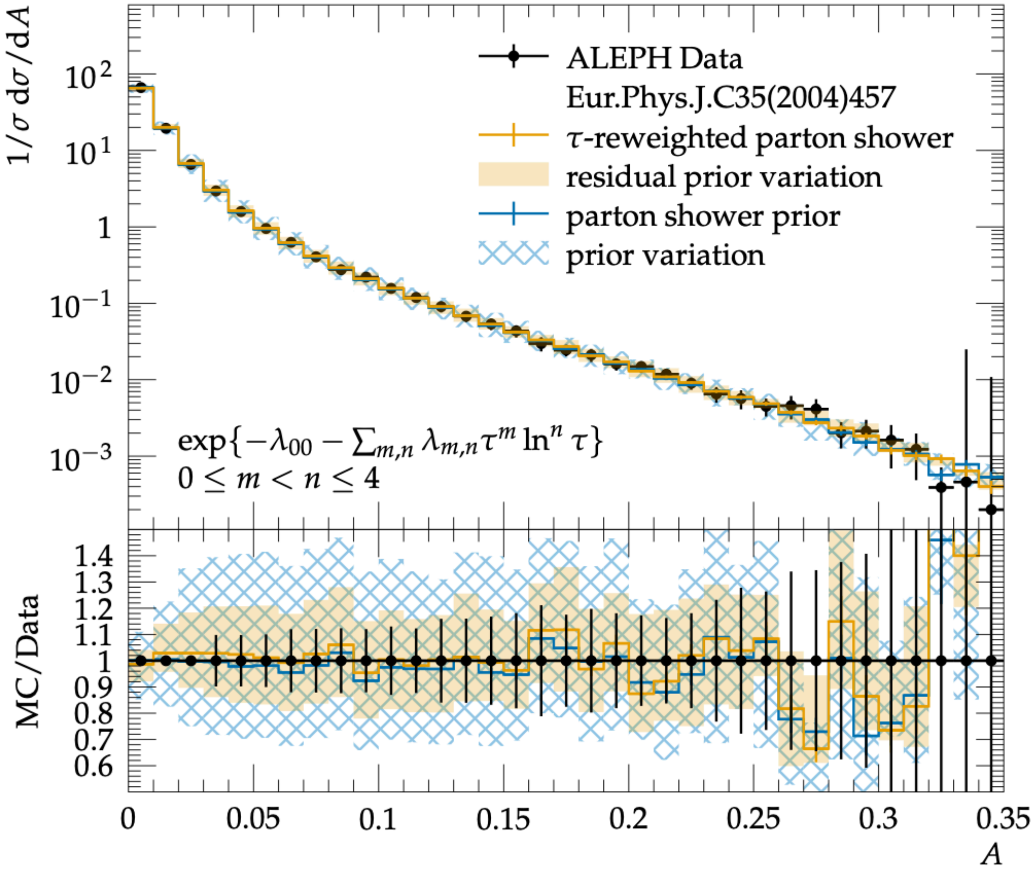
Impact on other observables



Broadening $\simeq$ Thrust



Aplanarity $\not\simeq$ Thrust



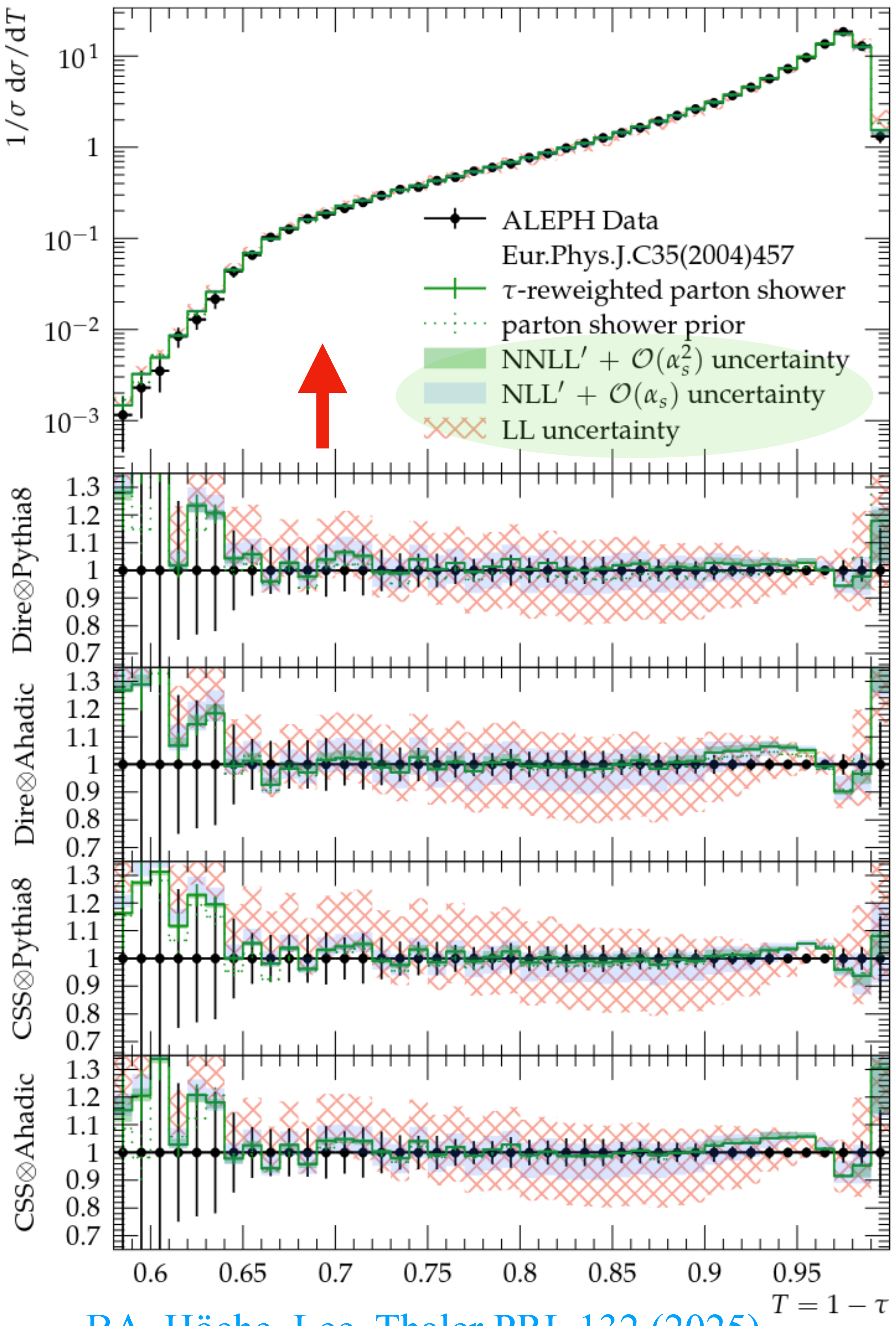
QCD Theory meets Information Theory

Proof-of-Concept with Diverse Priors: Validated with 4 *priors* combining 2 showers (CSShower, Dire) and 2 hadronization models (Pythia8, Ahadic)

Improved Accuracy & Convergence: Uncertainty bands and convergence from LL to $\mathrm{NNLL + NNLO}$ accuracy!

Natural Uncertainty Propagation: Propagate moment uncertainties directly to Lagrange multipliers

Further improvement systematically attainable by including additional moments of **precision observables**



BA, Höche, Lee, Thaler PRL 132 (2025)

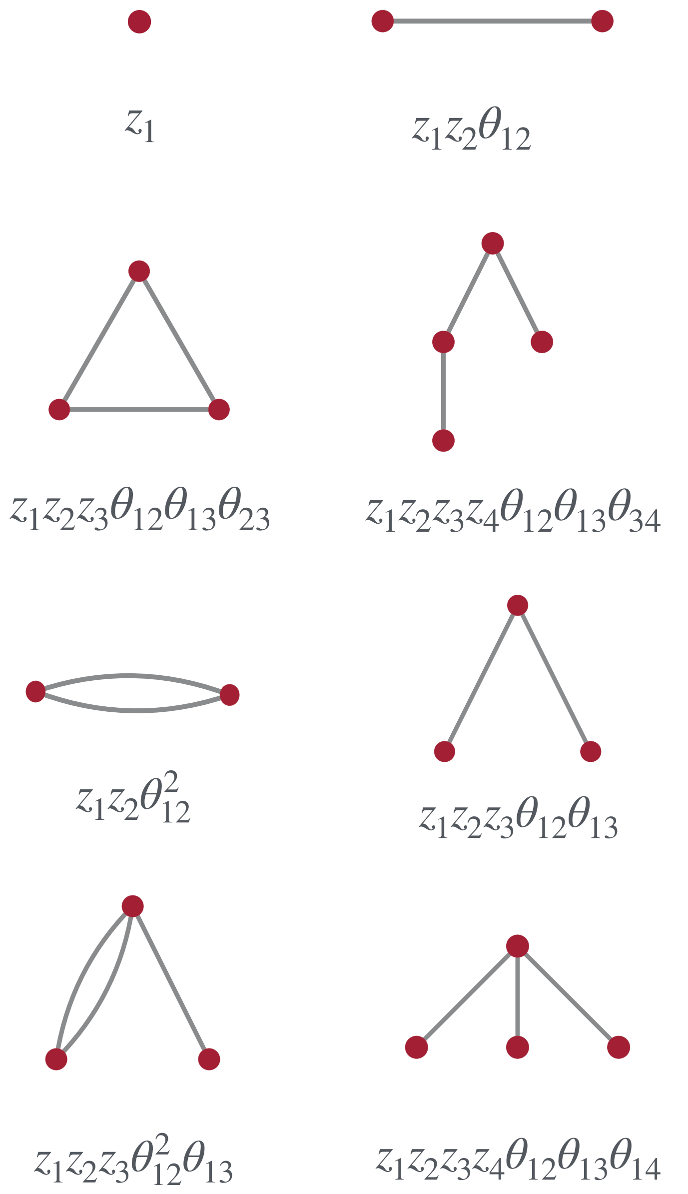
Multiple observables: how many to capture max information?

Basis of observables: Energy flow polynomials

Idea: Employ systematic complete linear basis of IRC-safe observables for multi-observable constraints Komiske, Metodiev, Thaler 1712.07124

$$\boxed{\mathrm{EFP}_G^{(\beta)} \equiv \sum_{i_1,\ldots,i_k \in R} \prod_{v \in V} z_{i_v} \prod_{(a,b) \in E} \theta_{i_a i_b}^{\beta}}$$

Sum over all particle assignments to graph G — graph topology encodes correlation structure



Degree $d = |E|$ → systematic complexity ordering: known low-degree already captures bulk of event structure (over-completeness)

Cal, Thaler, Waalewijn 2205.06818

All IRC-safe observables are (thrust, broadening, C-parameter,…) linear combinations of EFPs

Idea: Constrain mixed moments $\langle \mathrm{EFP}_G^m \ln^n \mathrm{EFP}_G \rangle$ to inject re-summed + FO information

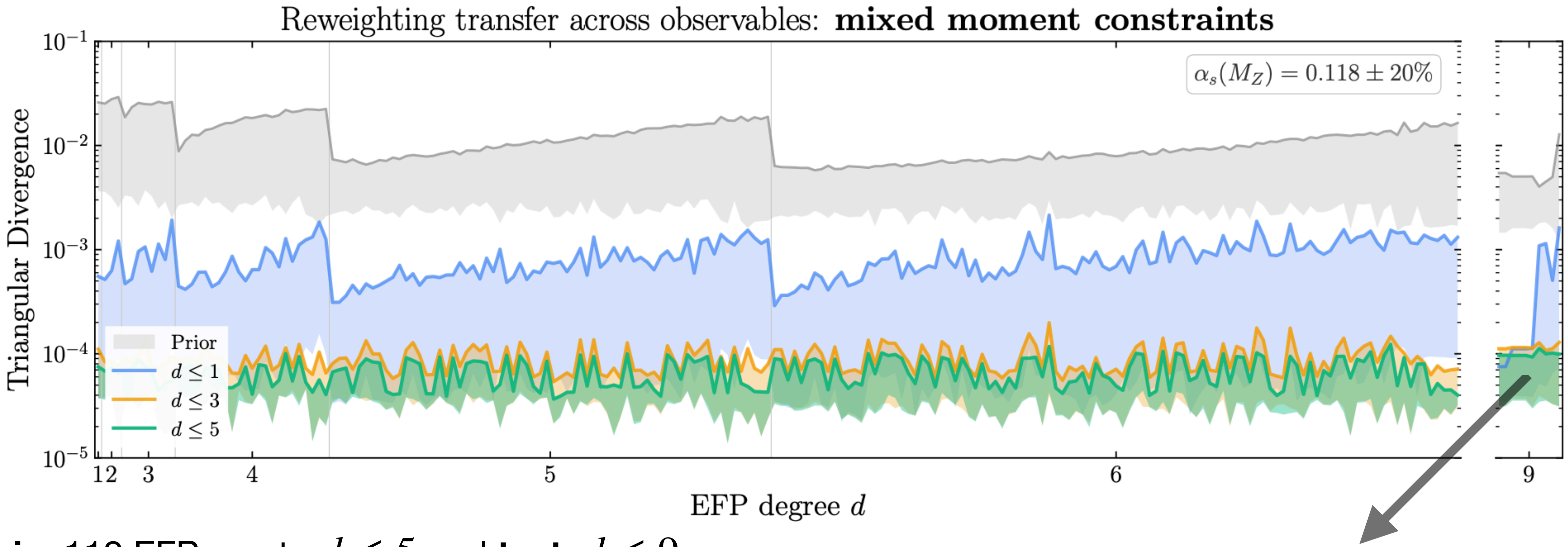
Vertices → energy fractions: $z_i = \dfrac{E_i}{\sum_j E_j}$

Edges → pairwise angles: $\theta_{ij} = \sqrt{2(1 - \cos\angle_{ij})}$

Information saturation

Key question: how many EFP constraints before adding more gives diminishing returns?

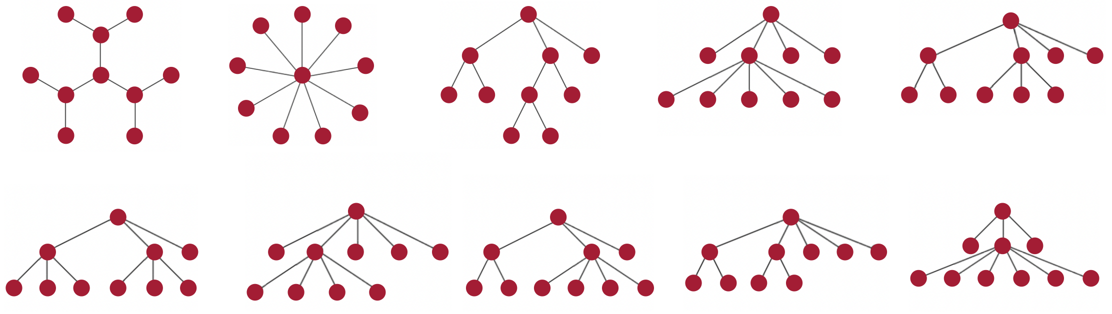
Two-shower setup: Prior is broken CSS shower with stripped non-singular contributions and $g \to q\bar{q}$ channel disabled ⇒ wrong NLL single logs and multiplicity



Train: 112 EFPs up to $d \leq 5$ and **test:** $d \leq 9$

Fix 4-moments each: $\langle \mathrm{EFP}_G \rangle$, $\langle \ln \mathrm{EFP}_G \rangle$, $\langle \ln^2 \mathrm{EFP}_G \rangle$,$\langle \mathrm{EFP}_G \ln \mathrm{EFP}_G \rangle$

Takeaway: Saturation by $d \leq 3$ and robust across $\alpha_s(M_Z) \pm 20\,\%$



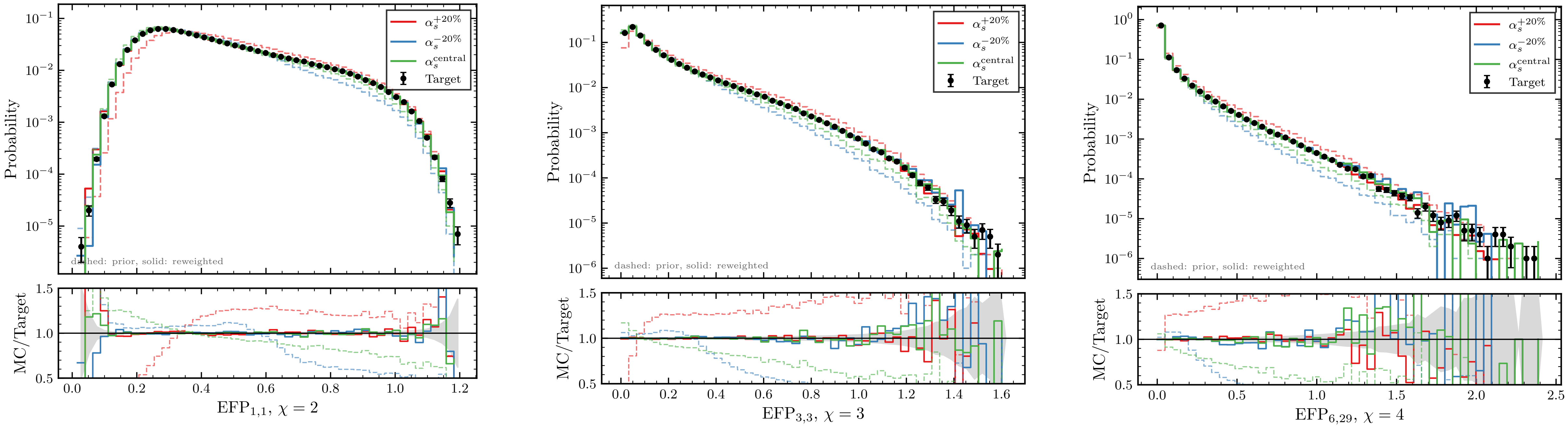
Many degree 9 EFPs: we test the well-defined subset of homeorphically irreducible trees

BA, Lee, Thaler 2604.00084

Information saturation

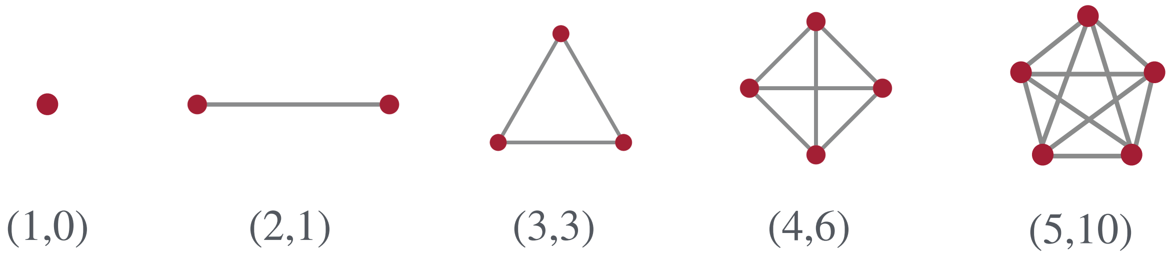
Key question: how many EFP constraints before adding more gives diminishing returns?

Two-shower setup: Prior is broken CSS shower with stripped non-singular contributions and $g \to q\bar{q}$ channel disabled ⇒ wrong NLL single logs and multiplicity



Train: 112 EFPs up to $d \leq 5$ and **test:** $d \leq 9$

Fix 4-moments each: $\langle \mathrm{EFP}_G \rangle$, $\langle \ln \mathrm{EFP}_G \rangle$, $\langle \ln^2 \mathrm{EFP}_G \rangle$,$\langle \mathrm{EFP}_G \ln \mathrm{EFP}_G \rangle$



Higher χ → more particles required → probes increasingly complex multi-particle angular correlations.χ $= 2$: pairwise.χ $= 3$: triangular/planar. χ $= 4$: 3D structure.

Takeaway: Saturation by $d \leq 3$ and robus across $\alpha_s(M_Z) \pm 20\,\%$

BA, Lee, Thaler 2604.00084

Transfer to event shapes

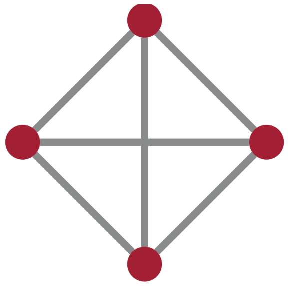
Weights are **event-level** → any observable can be reweighted without retraining
Test: do EFP-trained weights correct standard event shapes never seen during training?

Thrust, sphericity, C-parameter, broadening, N-jettiness — **none** in training set

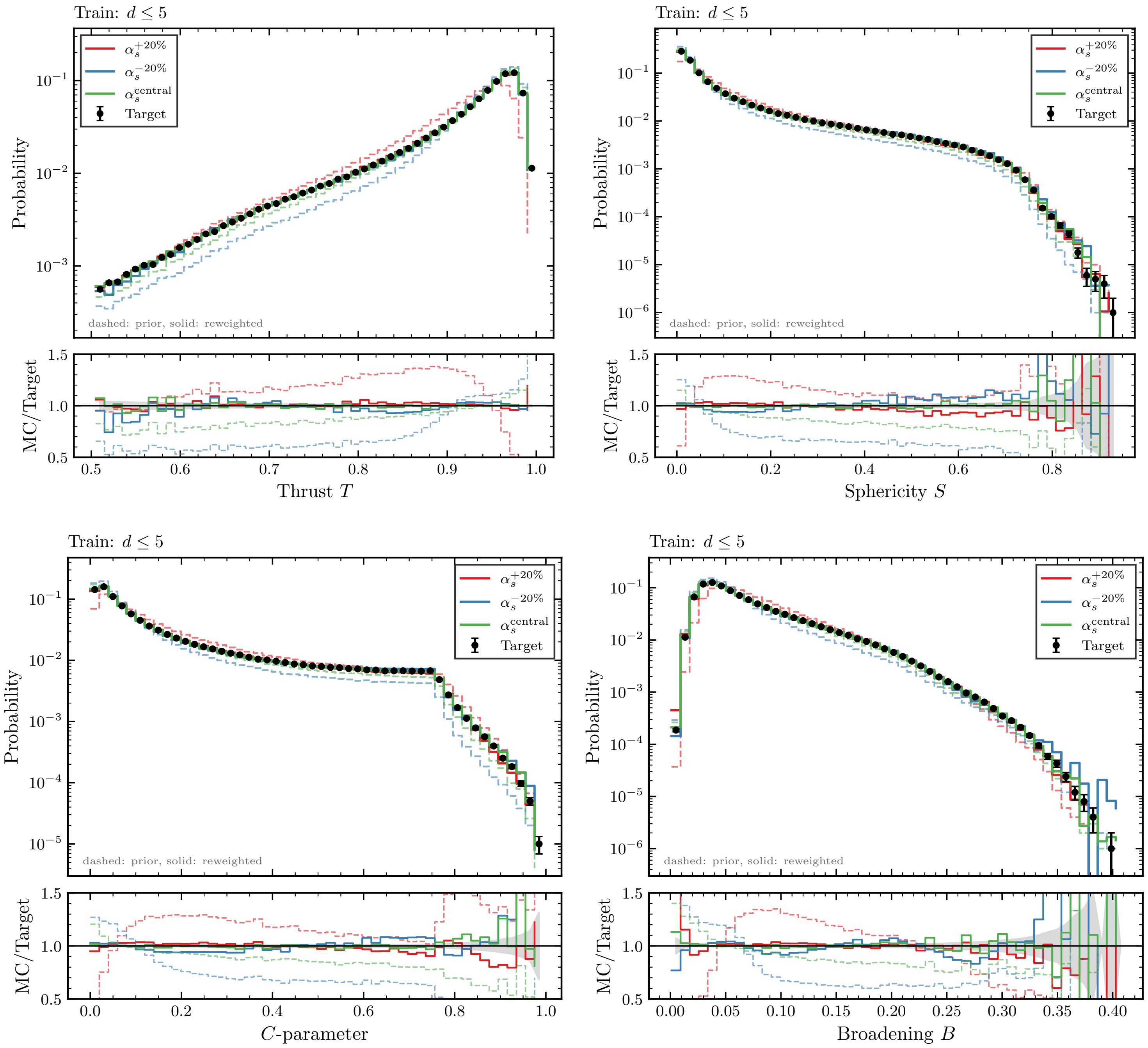
All corrected to near-target agreement across full $\alpha_s(M_Z) \pm 20\,\%$ range

Takeaway: EFPs span IRC-safe observables → constraining EFP moments propagates to event shapes

Exceptions: E.g. aplanarity — sensitive to multi-particle tails, correction poorer for our extreme α_s variations



Lowest order $\chi = 4$ polynomial first arises at $d = 6$



BA, Lee, Thaler 2604.00084

Information theory for efficient optimal observable selection

HDSense: efficient observable ranking

Hadronization models have many parameters θ tuned by correlated observables $\mathcal{O}$ ⇒ need *maximum* constraining power with *minimal* redundancy

Full Fisher information matrix:

$$I_{ab}(\theta) = \mathbb{E}_{p(\mathcal{O}|\theta)}\left[\frac{\partial \ln p(\mathcal{O}\,|\,\theta)}{\partial \theta_a}\frac{\partial \ln p(\mathcal{O}\,|\,\theta)}{\partial \theta_b}\right]$$

Measures how much a set of observables constrains parameters ***but*** *requires full joint distribution*

Idea: derive approximation to full FI using only 1D-observable Fisher $I^{(i)}$

$$\boxed{\mathcal{S}_{\mathrm{HD}}(\mathcal{X}) \;=\; \mathrm{Info}(\mathcal{X})\left[1-\beta\,\mathcal{P}_{\mathrm{overlap}}(\mathcal{X})\right]}$$

Sum of individual constraining powers in traces if observables are uncorrelated

$$\mathrm{Info}(\mathcal{X}) = \sum_{i\in\mathcal{X}} \mathrm{Tr}\, I^{(i)}$$

P_{overlap} **penalizes** pairs with aligned Fisher matrices via Frobenius angle and $\beta \in [0,1]$ ↔ redundancy penalty strength

$$\mathcal{P}_{\mathrm{overlap}}(\mathcal{X}) \;=\; \frac{2}{\sum_{k\in\mathcal{X}} \mathrm{Tr}[I^{(k)}]}\sum_{\substack{i,j\in\mathcal{X}\\ i<j}}\sqrt{\mathrm{Tr}[I^{(i)}]\mathrm{Tr}[I^{(j)}]\cos(\Phi^F_{ij})}$$

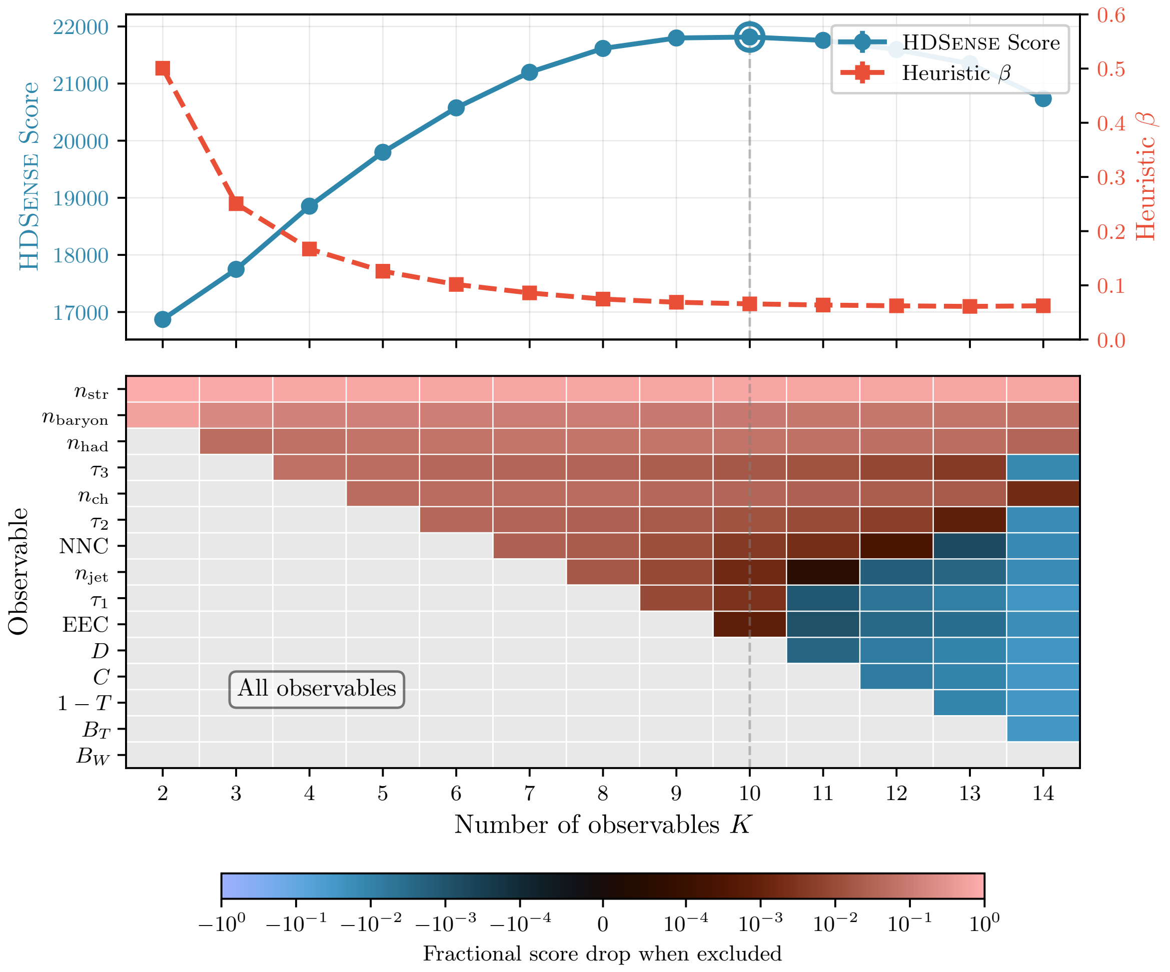
Output: Best K observables = subset with largest S_{HD} score

BA, MLHad 2602.01509

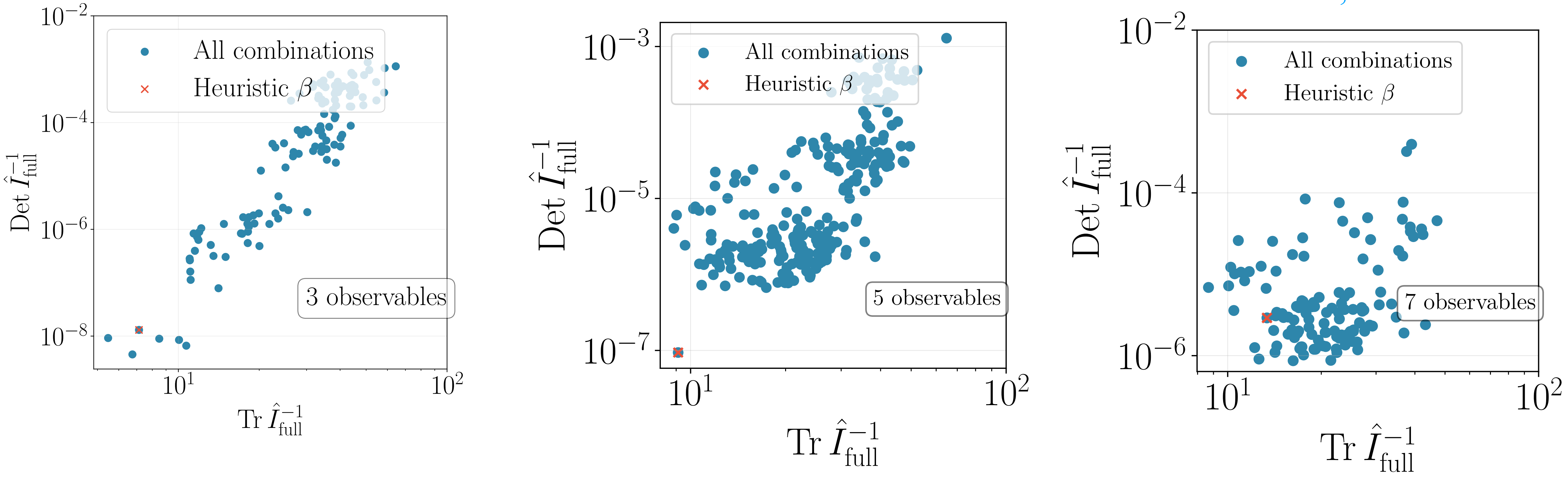
HDSense: Results and comparisons

Apply to rank diverse observables for constraining **5 Lund string parameters** in Pythia, validated against full-likelihood classifier (XGBoost)

- HDSense **efficiently** picks **near-optimal subsets**: 15 Fisher matrices vs. ${}^{15}C_K$ classifier trainings
- Multiplicities rank above event-shapes/correlators — **IRC-unsafe** observables carry more had. info
- n_{bar} and n_{str} irreplaceable: only observables sensitive to ρ (strange suppression) and χ (baryon enhancement)
- **Score saturates** around $K \sim 8$ for this test — beyond is diminishing returns



BA, MLHad 2602.01509



Summary

Directly improved shower accuracy beyond strict LL and various efforts towards MC@NNLO

Shannon & Boltzmann → IT reweighting: injects best theory information with uncertainty via theory-based moment constraints at the event-level

Fisher → HDSense: efficient observable selection for hadronization++ — guides analyses, tuning and training data selection for ML surrogates

Outlook

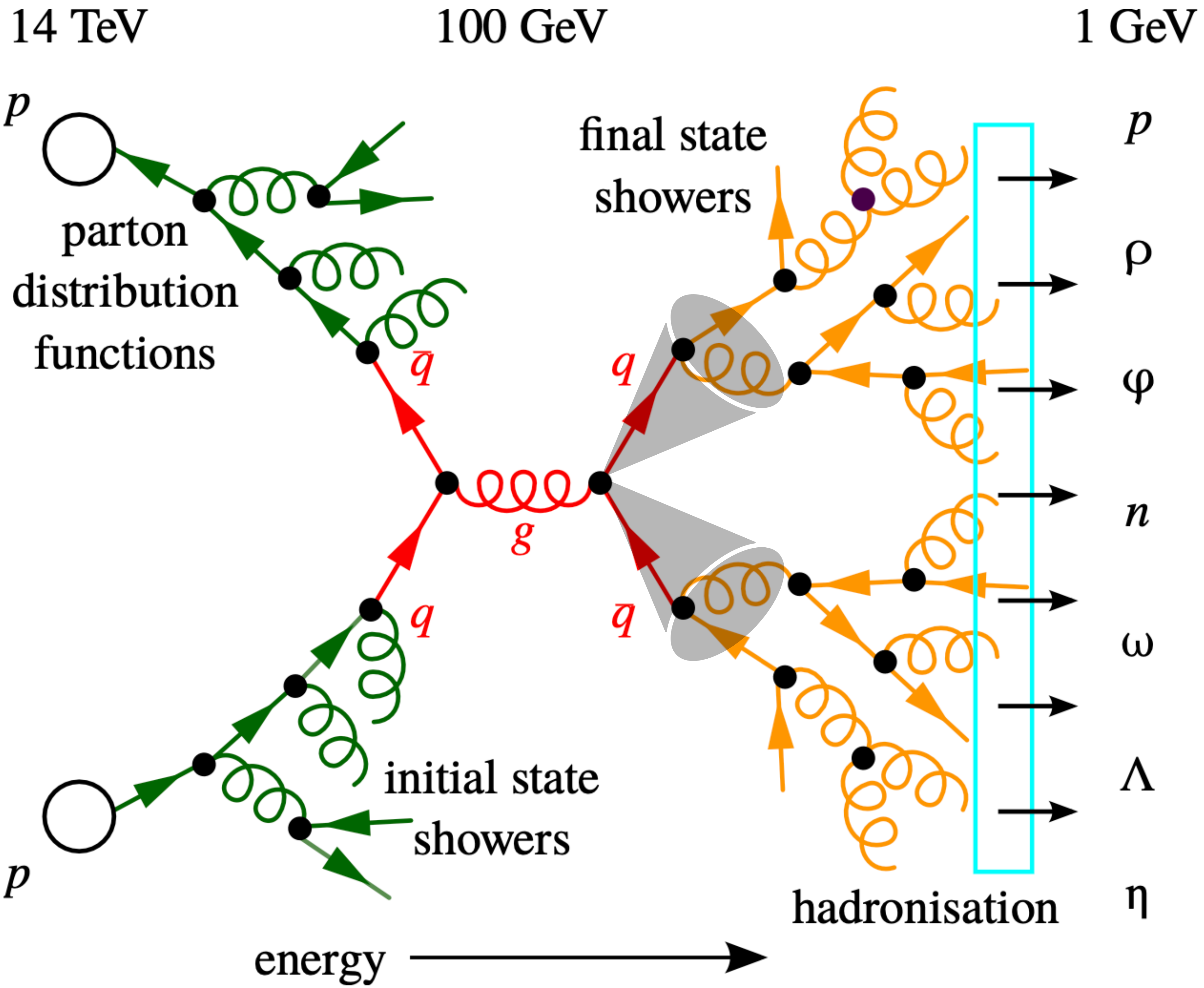
Drell-Yan upgrade ATLAS MC production sample to $\mathrm{N^3LO} + \mathrm{N^4LL}$ + *ab-initio LQCD* via double-diff. q_T and $\Delta\phi$ moment constraints

NNLO accuracy without negative weights: reweight positive-weight LO shower instead of traditional NLO matching

Applying IT-reweighting beyond LEP/LHC → E.g. Neutrino generators

Back-up

Factorizing QCD



Anatomy of a high-energy collision

Standard perturbative QFT expansion

Radiative effects due to **large energy differences** between collision point and detector

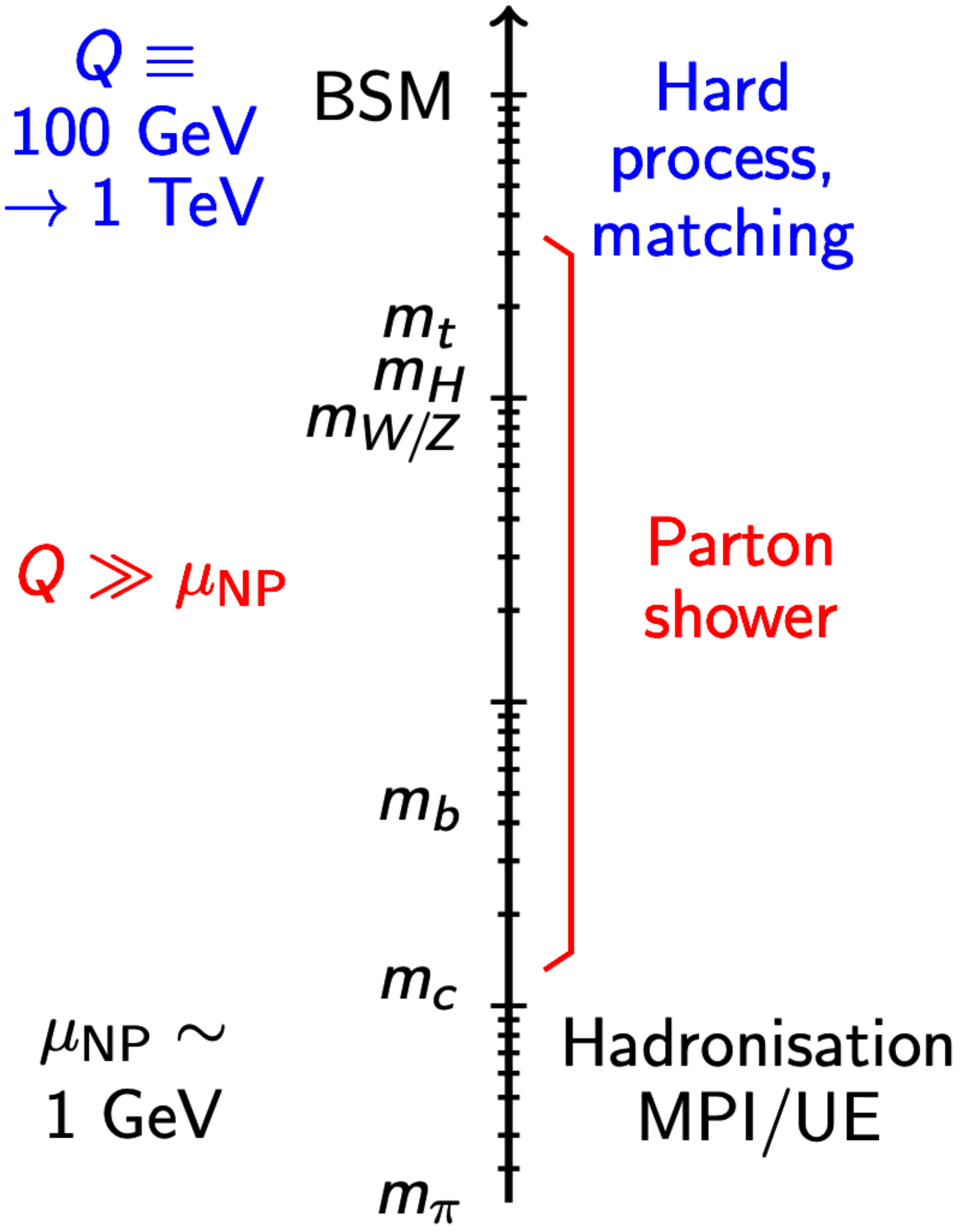
Expect logarithms between disparate scales:

$$\alpha_s \log^2(Q/\mu_{\mathrm{NP}}), \alpha_s \log(Q/\mu_{\mathrm{NP}}), \ldots$$

LL NLL ...

Large logs to **re-sum** analytically or implicitly by shower

Non-perturbative (requires modelling)



Parton Shower

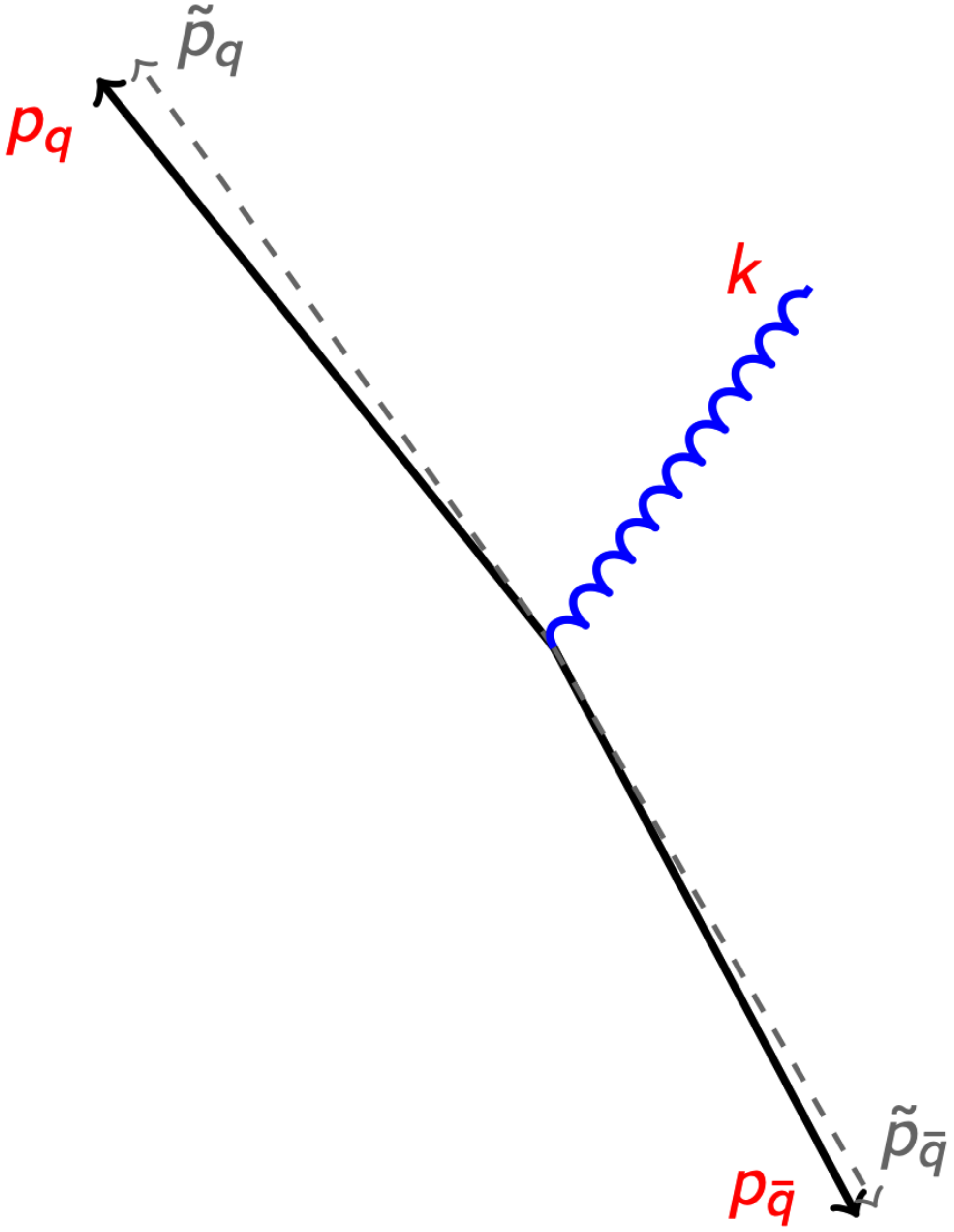
Mechanism: Gluon emission ↔ dipole splitting at large N_c

Branching requires going from a more massive to a less massive state

Energy is taken to boost the mass but should not affect the topology of the event

Parton masses must **stay the same** before and after branching

Careful **momentum mapping** needed from pre- to post-branching



Beyond Leading Accuracy

NLL **accurate** if emissions factorise up to $\mathcal{O}(e^{-\Delta})$ corrections

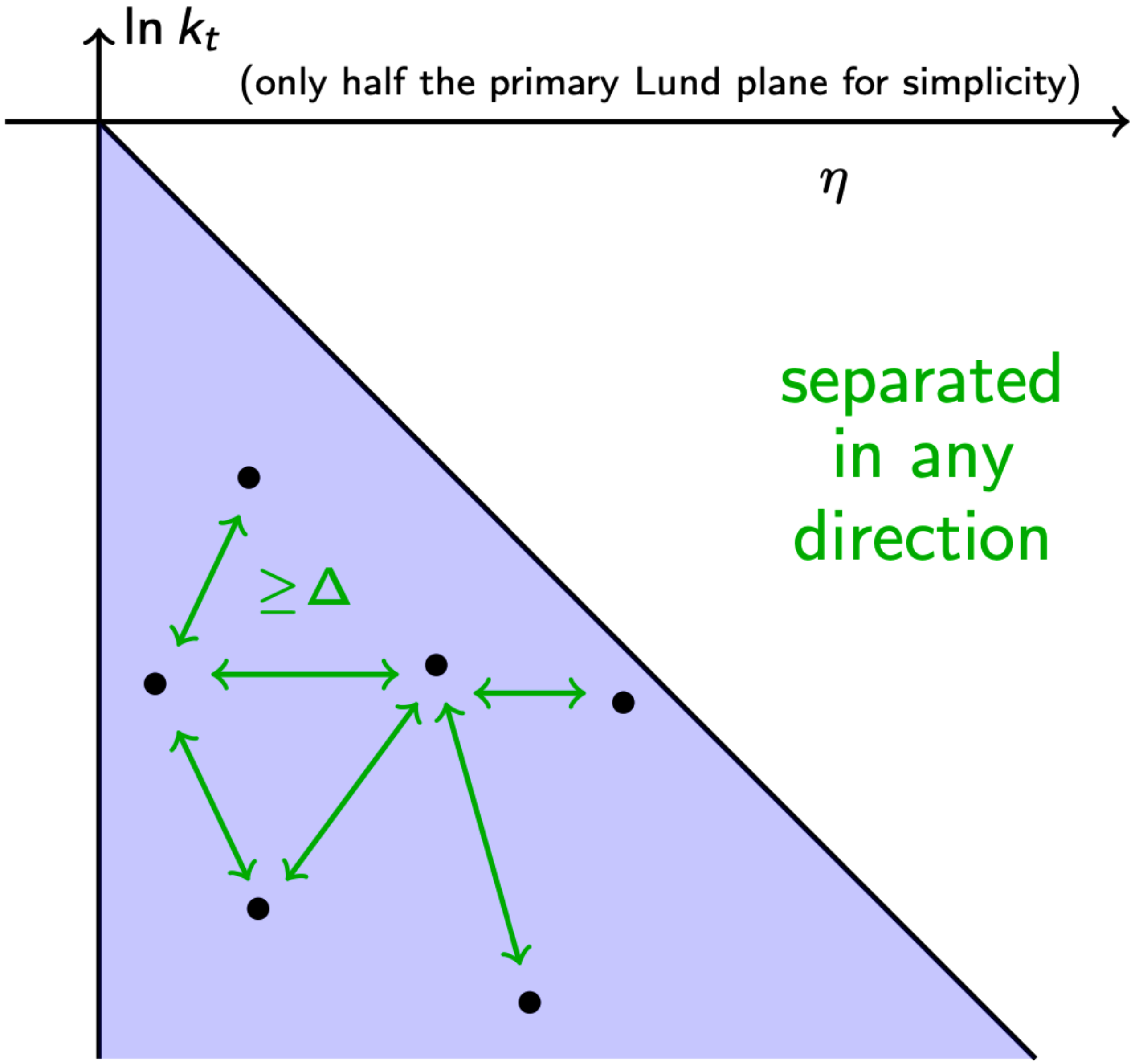
Emissions should be **well-separated** in the lund plane

How to test whether true NLL deviations or subleading effects?

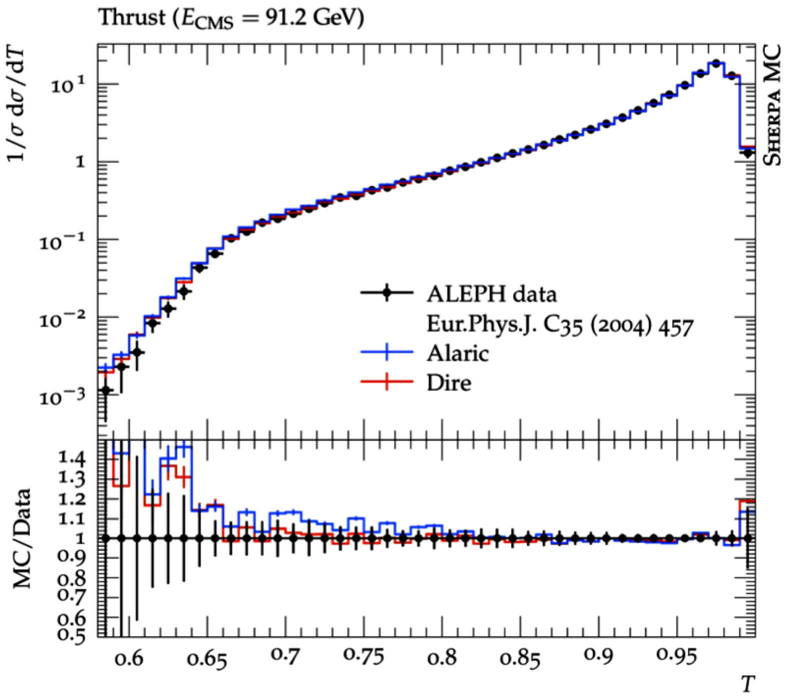
Resummation regime

$\alpha_s \log v \sim 1, \ \alpha_s \ll 1$

Banfi, Salam, Zanderighi JHEP 03 (2005)

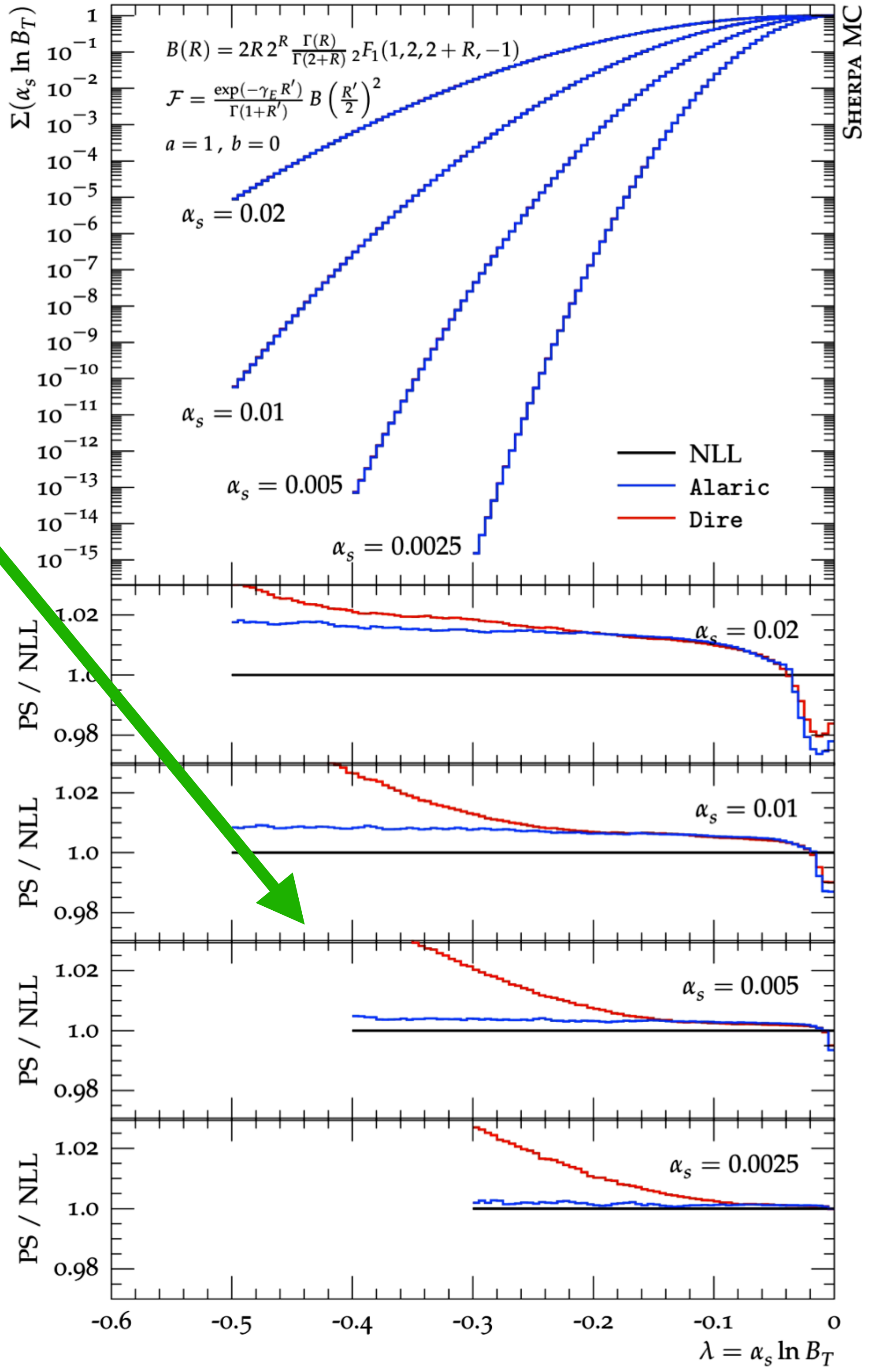


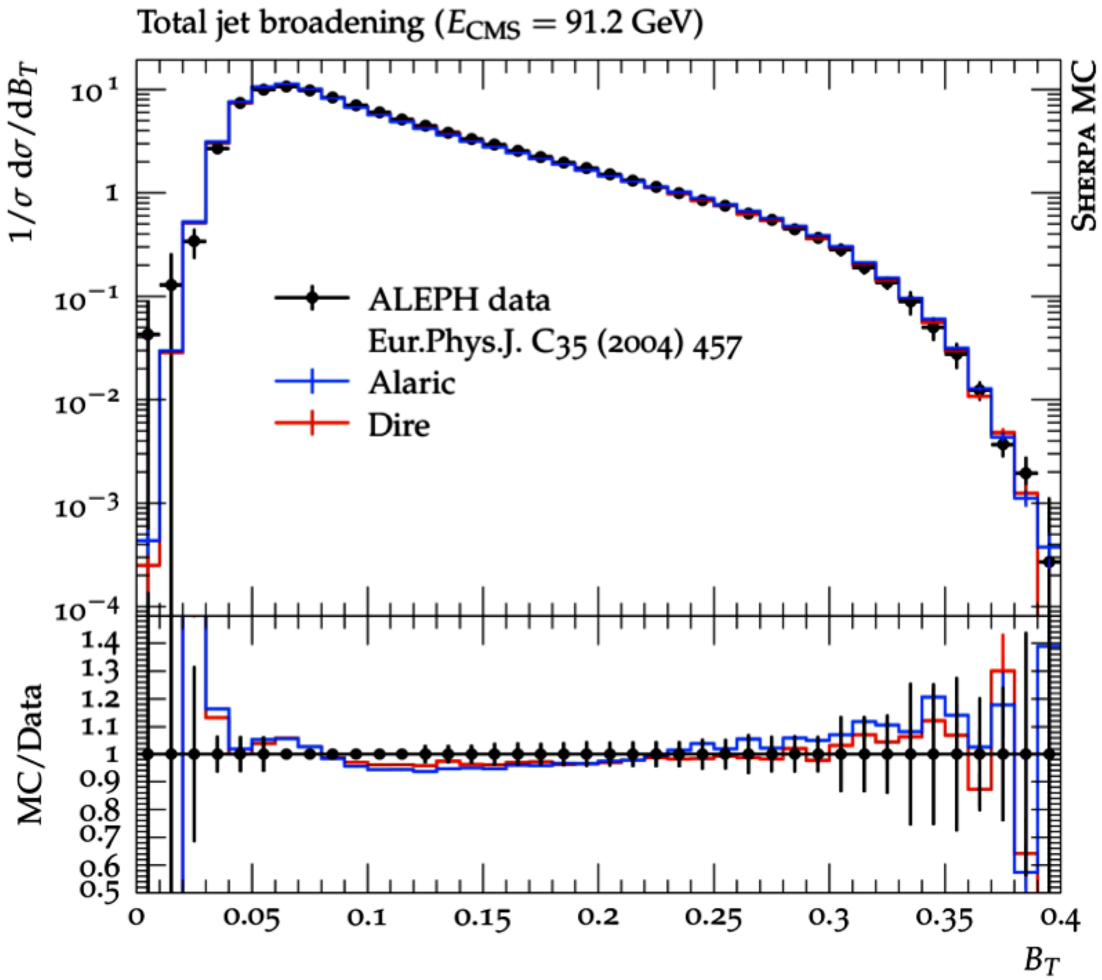
Numerical test and data



For some observables like thrust ***both*** *ALARIC & standard showers NLL accurate*

For others like broadening ***only ALARIC*** *is - sensitive to transverse momentum (soft) recoil effects*



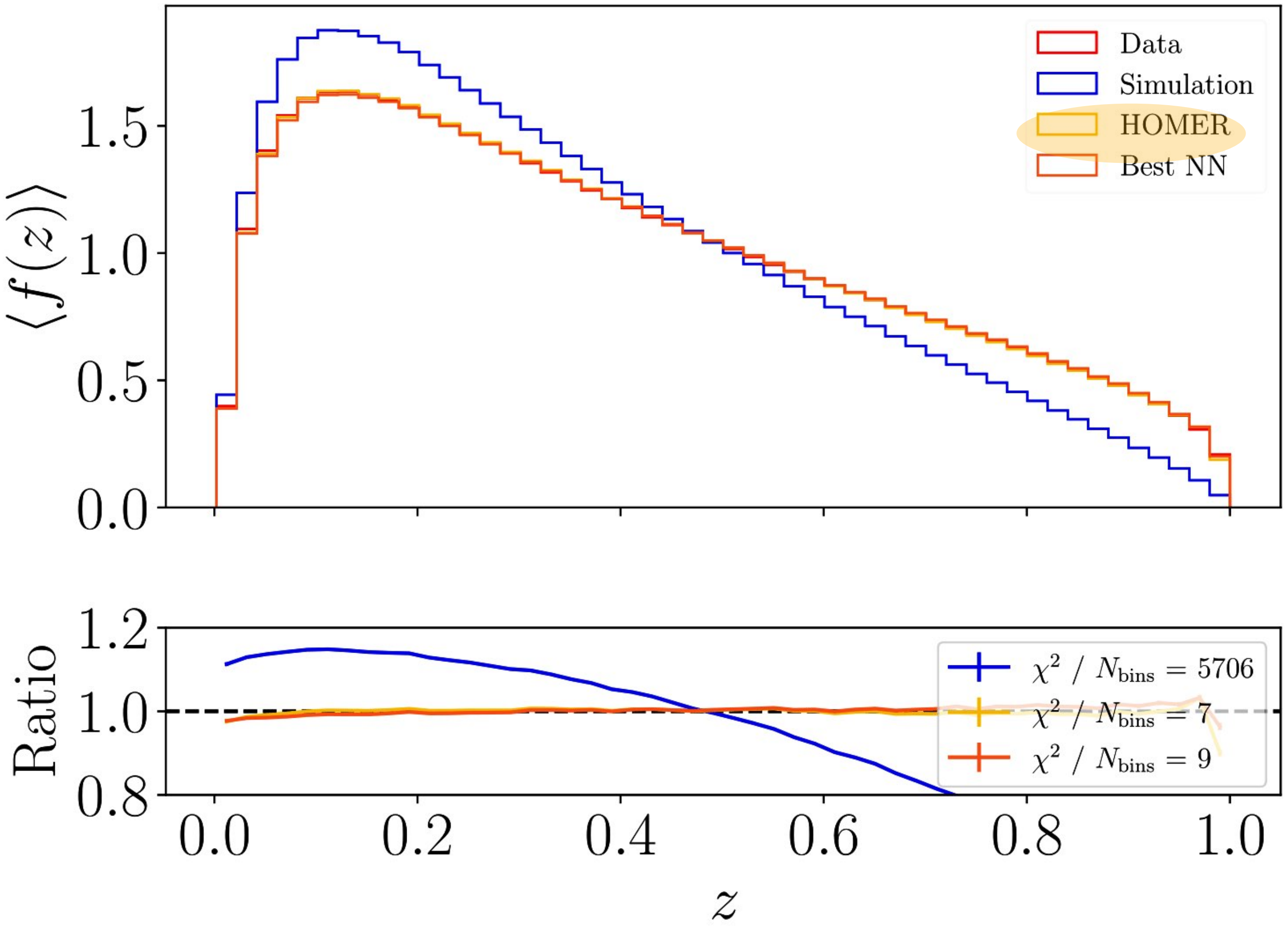


Note: *Perturbative region right - deviations for broadening and a-planarity*

Machine Learning for Hadronization

MLHAD international collaboration of theorists and experimentalists

Big goal: Develop a data-driven, ML-based framework with max theory information for simulating hadronization in event generators (compare NNPDF)



Bierlich et al. 2410.06342

BA et al. 2503.05667

MLHAD

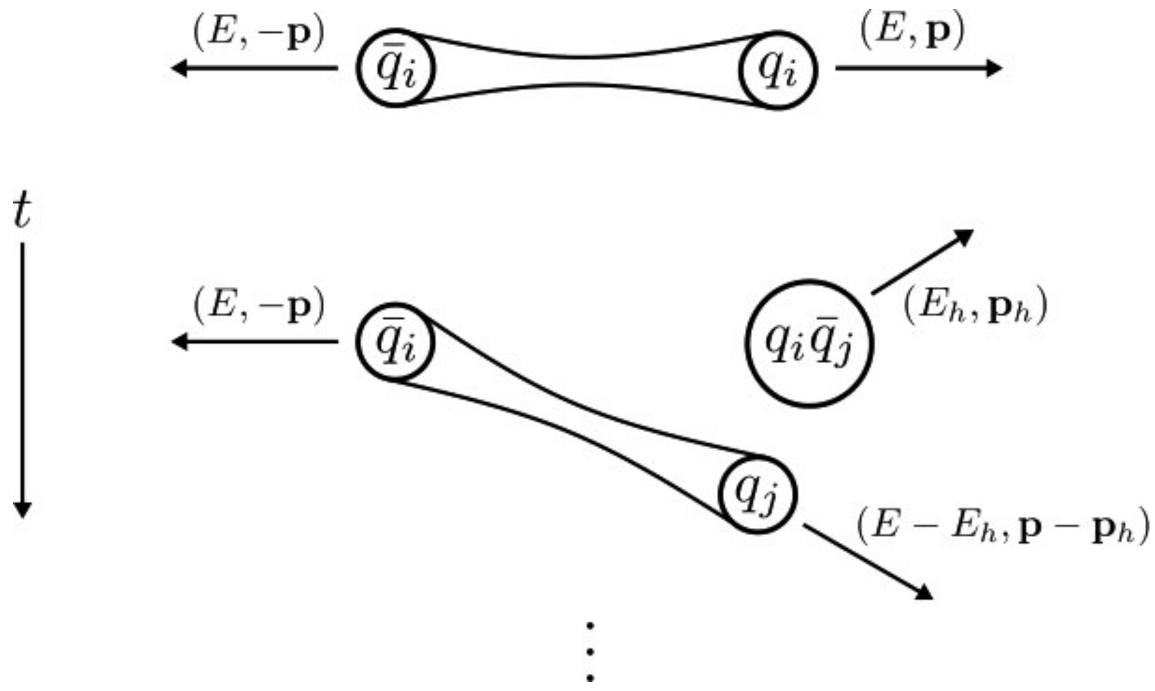
BA, Bierlich, Gambhir, Ilten, Menzo, Mrenna, Szwec, Wilkinson, Zupan

Hadronization empirical models

Current paradigm: Empirical models very successful due to expertise that has been tested against experiments for 40+ years

Lund String model: Explains hadronization kinematics in terms of **LC momentum fraction** z and **transverse momentum of string break** p_T

Hadron emission from string piece defined by **fragmentation function** $f(z)$ → entire hadronization chain reproduced by **iterating**



Pythia Lund String Model: Colored singlets + O(20) parameters → Hadrons

Simplified example from arxiv:2203.04983 .

Machine learn it

Complex problem with **no full model flexible enough** and where **tuning is expensive**

Two groups have recently tackled the subject: **MLHAD** (arxiv:2203.04983, arxiv:2311.09296, arxiv:2410.06342, arxiv:2503.05667, arxiv:2505.00142, arxiv:2508.10090, arxiv:2602.01509) and **HADML** (arxiv:2203.12660, arxiv:2305.17169, arxiv:2312.08453)

Different **generators** (Pythia, Herwig) and different **architectures** (cSWAE,BNF,GNNs, GAN) with different **degrees of implementation**

Flagship Model: HOMER

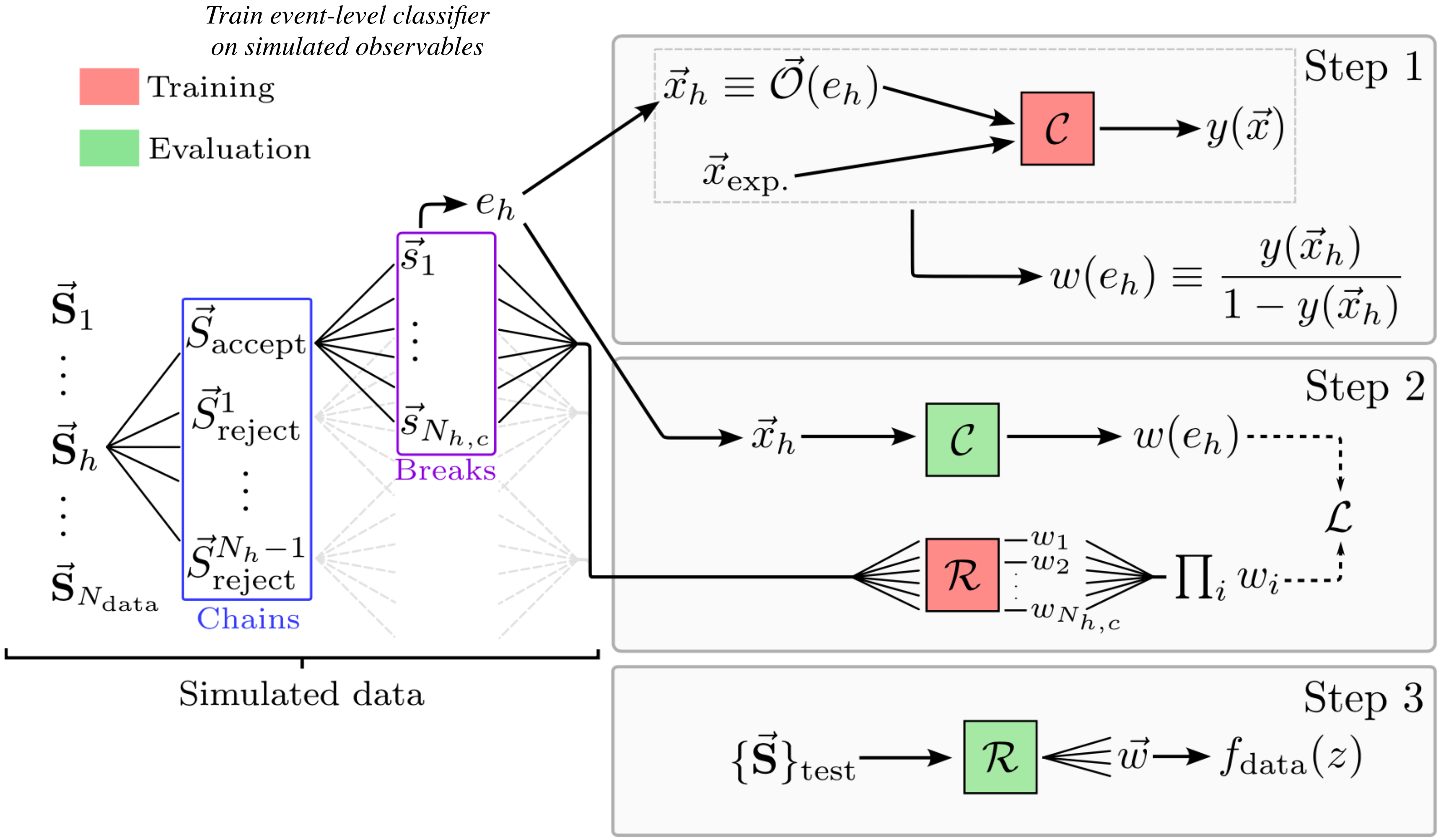
Histories and Observables for Monte-Carlo Event Reweighting: Take **Pythia** as baseline and **reweight**

Three step procedure: Generate **once** and learn the appropriate reweighting function → **New model is Pythia + weights**

1) Learns *which* simulated events look like data

2) Breaks that global information into *local*, per-break reweightings

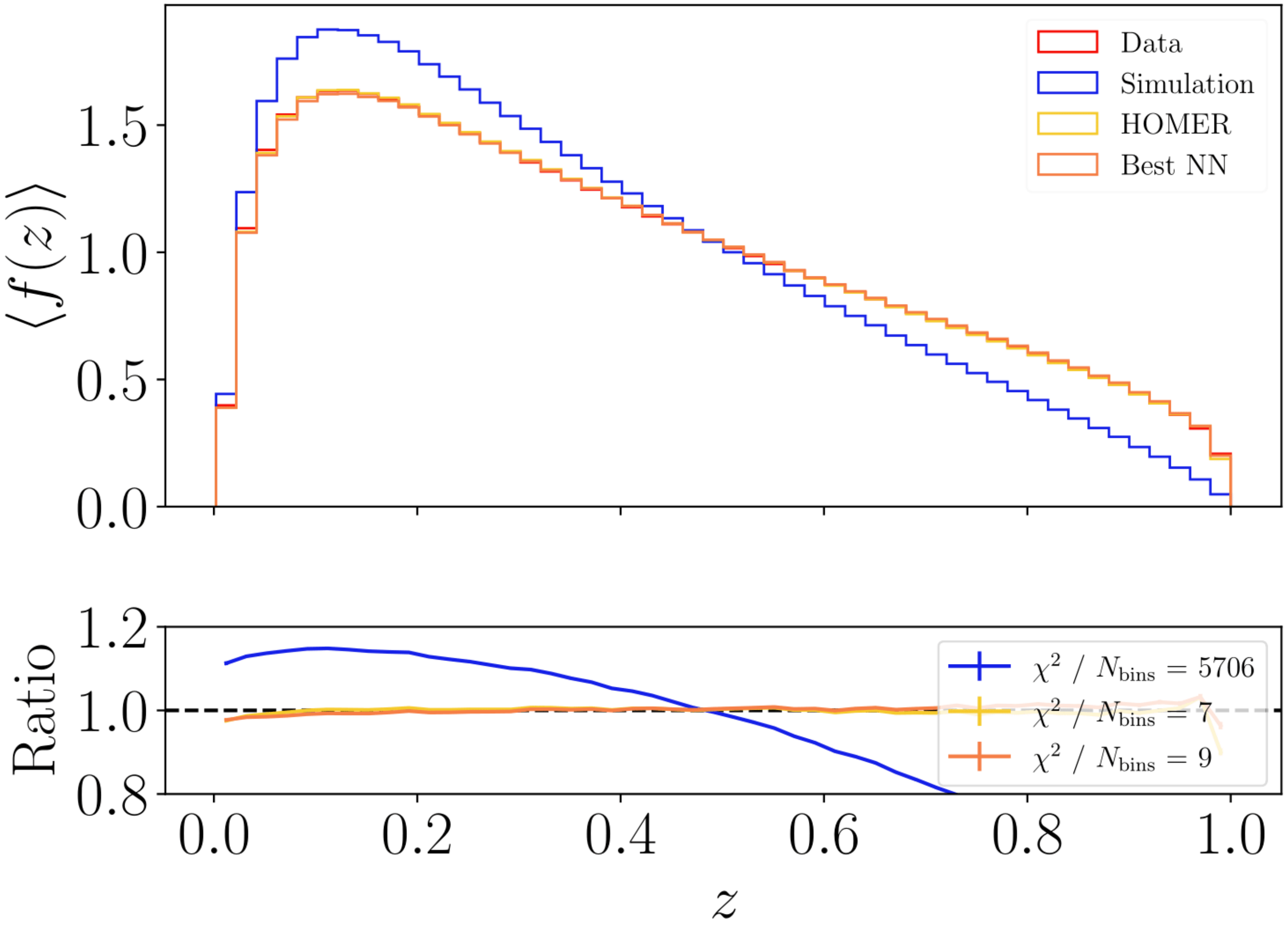
3) Uses those local weights to rebuild any downstream observable in data-driven fashion



New Underlying Model

The weights are translated to a **data-driven fragmentation function** via a Graph Neural Network-based architecture

We see how the model almost **saturates** what we can learn using the chosen NN architecture and is a **great fit to data**



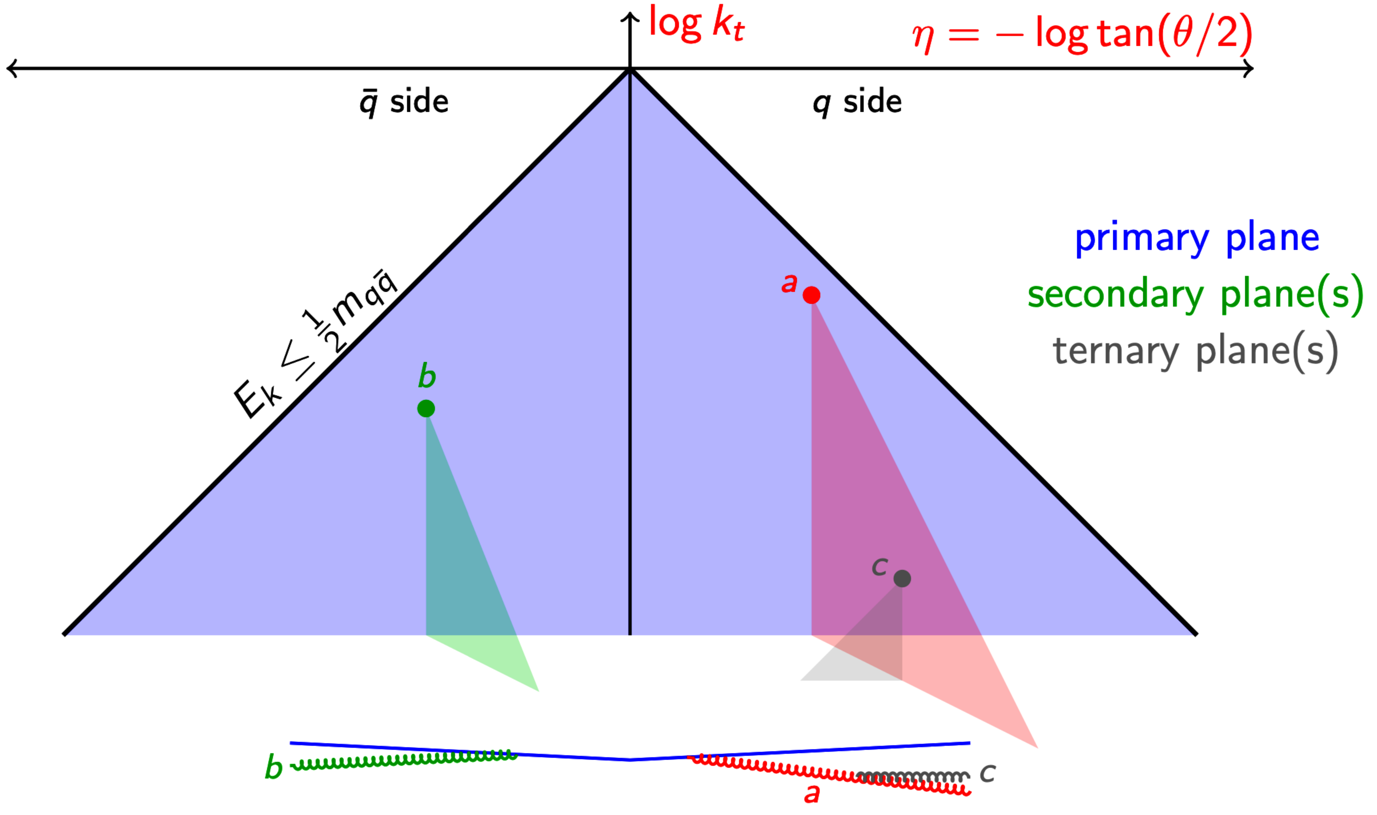
arxiv:2410.06342

arxiv:2503.05667

Beyond Leading Accuracy

Evolution steps:

1) Generate $k_{t,1} < Q$ with **Sudakov probability**
2) Generate η_1 and split dipoles $(q\bar{q}) \to (qg_1) + (g_1\bar{q})$
3) Generate $k_{t,2} < k_{t,1}$ from 2 dipoles
4) Generate η_2 and split dipoles $(g_1\bar{q}) \to (\bar{q}g_1) + (g_1g_2)$
5) Iterate until $k_t = k_{t,\mathrm{cut}}$



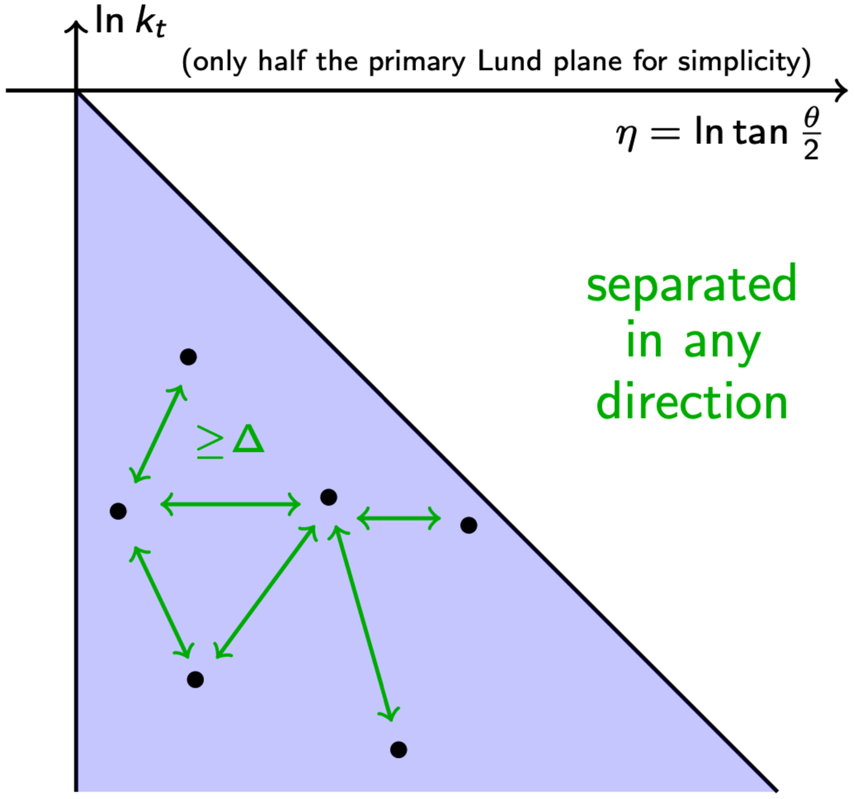
Evolution **resums** large logs at what accuracy?

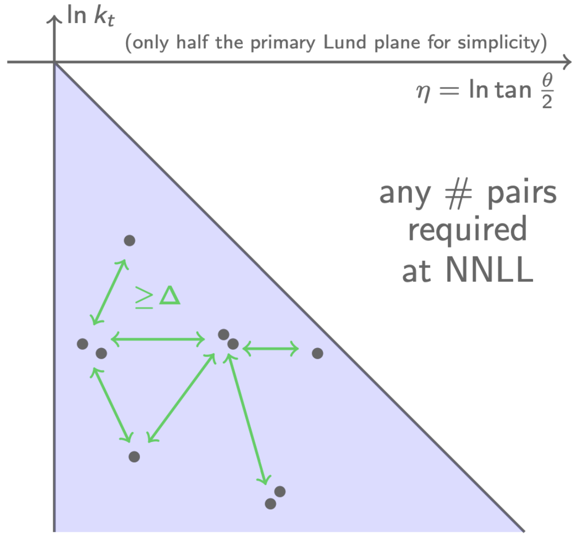
NLL **accurate** if emissions factorise up to $\mathcal{O}(e^{-\Delta})$ corrections

In shower an emission should not be affected by subsequent distant emissions

Test whether true NLL deviations or subleading effects?

Resummation regime: $\alpha_s \log v \sim 1,\ \alpha_s \ll 1$





Banfi,Salam, ’17

An NLL accurate algorithm in SHERPA

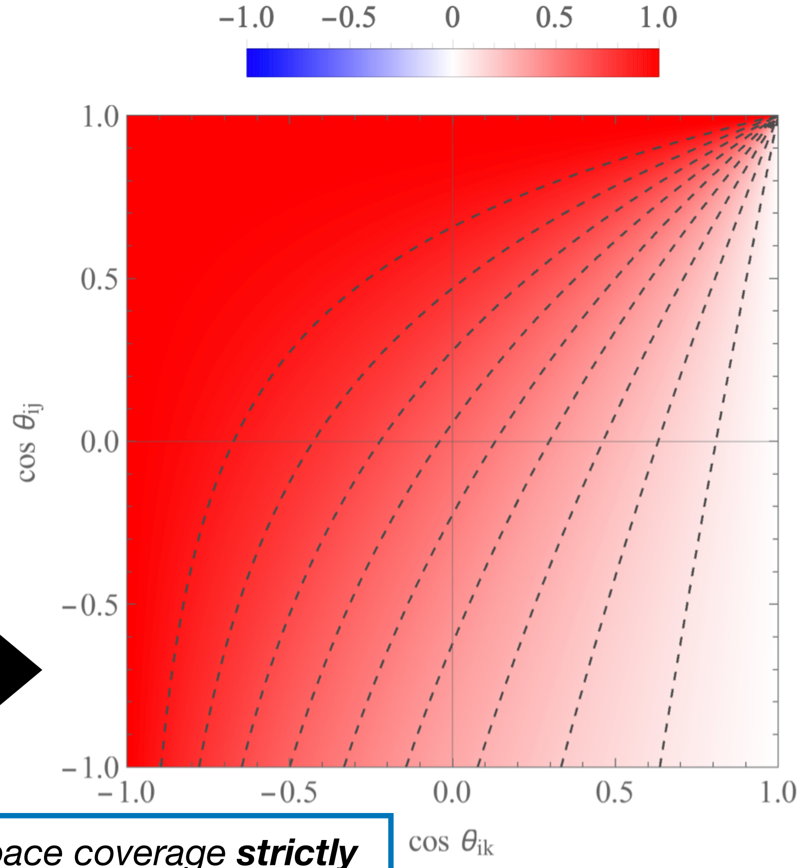
Ingredient 1: Treatment of soft radiation

Matrix element **factorizes** in soft gluon limit Marchesini,Webber ’88

$$|M|^2 \propto \frac{2W_{ik,j}}{E_j^2} = \frac{2}{E_j^2}\frac{(1-\cos\theta_{ik})}{(1-\cos\theta_{ij})(1-\cos\theta_{jk})}$$

Avoid soft double-counting by **partial fractioning**

$$W_{ik,j} = \bar{W}^i_{ik,j} + \bar{W}^k_{ki,j} \text{ with } \bar{W}^i_{ik,j} = \frac{1-\cos\theta_{ik}}{(1-\cos\theta_{ij})(2-\cos\theta_{jk}-\cos\theta_{ij})}$$



*Full phase-space coverage **strictly positive** maintains PD interpretation*

Ingredient 2: Momentum mapping

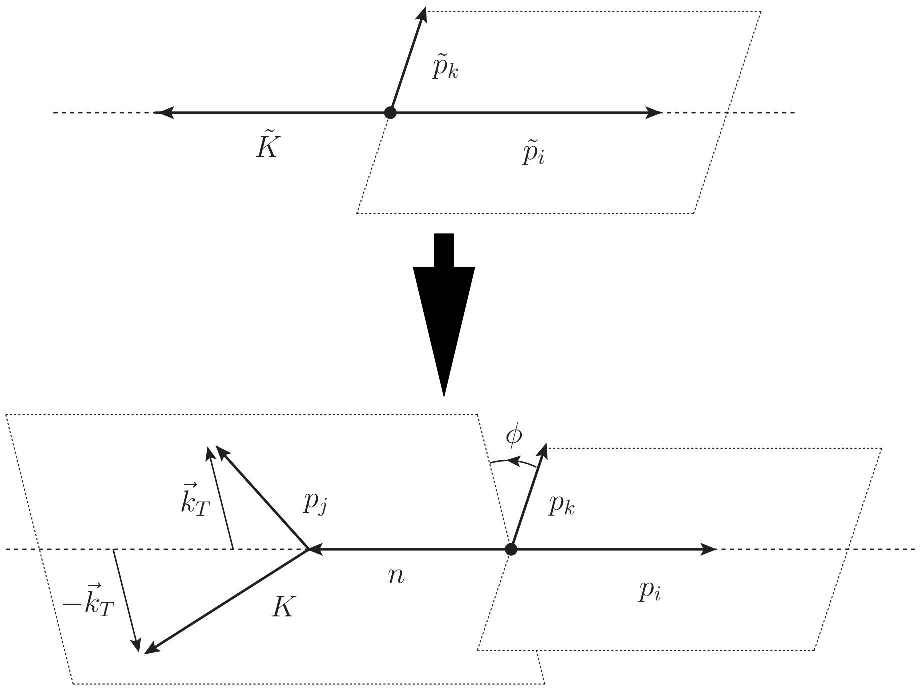
Maintain directions and collinear safety upon emission

$$p_i \xrightarrow{i||j} z\,\tilde{p}_i\,, \qquad p_k \xrightarrow{i||j} \tilde{p}_k\,, \qquad p_j \xrightarrow{i||j} (1-z)\,\tilde{p}_i$$

Recoil compensated **globally** by sum of multipole momenta $\tilde{K}$

$$p_j = (1-z)\,\tilde{p}_i + v\big(\tilde{K} - (1-z+2\kappa)\,\tilde{p}_i\big) + k_\perp, \quad v = \frac{p_i p_j}{p_i\tilde{K}}, \quad \kappa = \frac{\tilde{K}^2}{2\tilde{p}_i\tilde{K}}$$

$$K = \tilde{K} - v\big(\tilde{K} - (1-z+2\kappa)\,\tilde{p}_i\big) - k_\perp$$



Subtraction at NLO — Soft integrals

E.g. Integral with **two massive denominators** given by

$$I_{1,1}^{(2)}(v_{11},v_{12},v_{22}) = \frac{\pi}{\sqrt{v_{12}^2 - v_{11}v_{22}}} \left\{ \log\frac{v_{12}+\sqrt{v_{12}^2-v_{11}v_{22}}}{v_{12}-\sqrt{v_{12}^2-v_{11}v_{22}}} + \epsilon\left(\frac{1}{2}\log^2\frac{v_{11}}{v_{13}^2} - \frac{1}{2}\log^2\frac{v_{22}}{v_{23}^2} \right.\right.$$

$$+2\mathrm{Li}_2\left(1-\frac{v_{13}}{1-\sqrt{1-v_{11}}}\right) + 2\mathrm{Li}_2\left(1-\frac{v_{13}}{1+\sqrt{1-v_{11}}}\right)$$

$$\left.\left.-2\mathrm{Li}_2\left(1-\frac{v_{23}}{1-\sqrt{1-v_{22}}}\right) - 2\mathrm{Li}_2\left(1-\frac{v_{23}}{1+\sqrt{1-v_{22}}}\right)\right) + \mathcal{O}(\epsilon^2)\right\}$$

Massless limit only one simpler integral [arXiv:2208.06057]

$$I_{1,1}^{(1)}(v_{12},v_{12}) = -\frac{\pi}{v_{12}^{1+\epsilon}}\left\{\frac{1}{\epsilon} + \epsilon\,\mathrm{Li}_2\left(1-v_{12}\right) + \mathcal{O}(\epsilon^2)\right\}$$

Recap

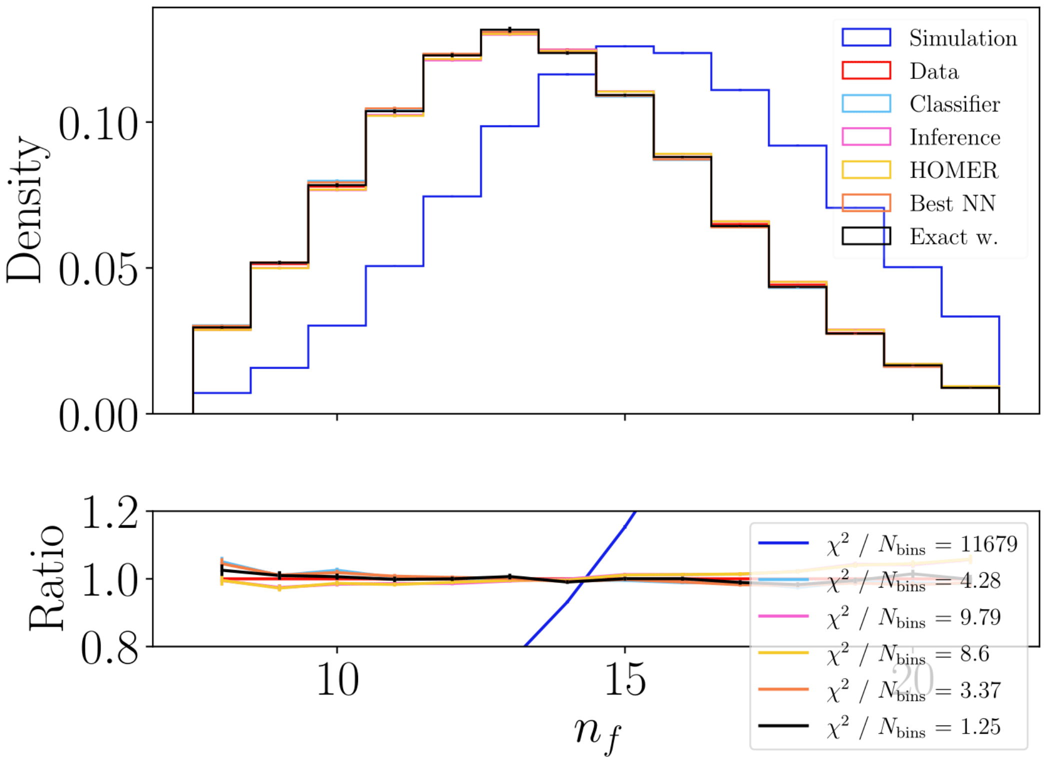
We recover the **underlying fragmentation function** given only **observable quantities**

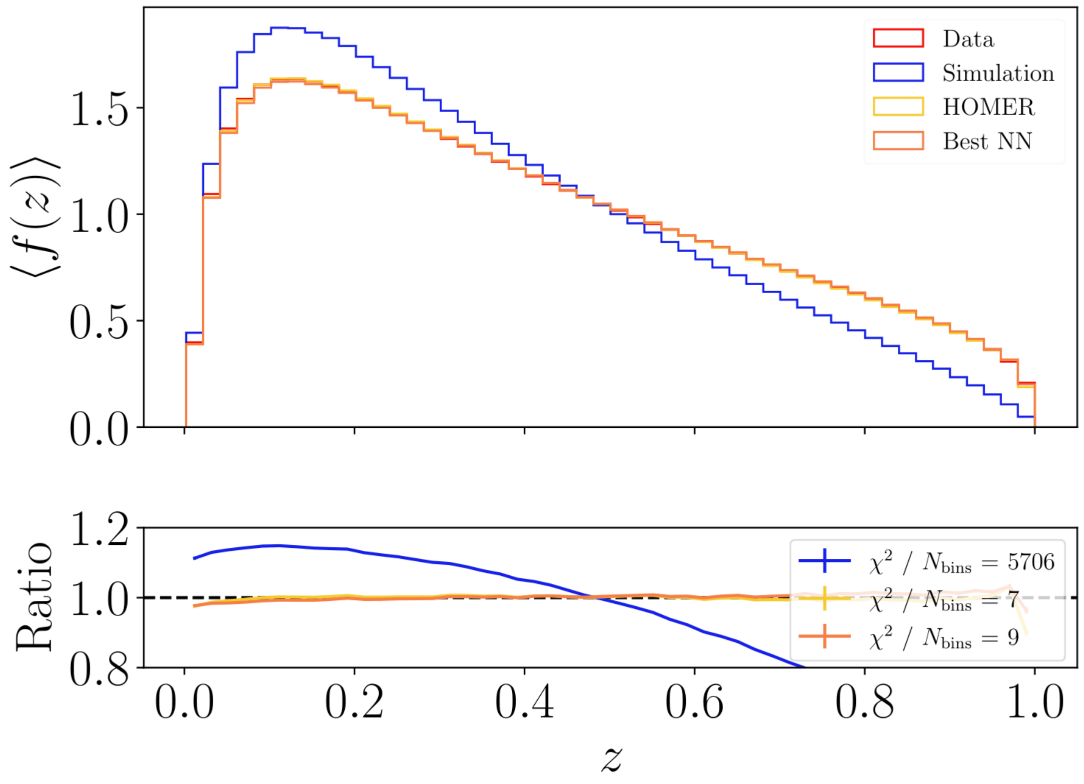
The Lund String model is still used as a **baseline**, properly accounting for the event structure is **essential**

So far **flavour has been ignored,** with only pions considered. Additionally, **gluons make things harder** (see arxiv:2503.05667)

For all three cases, we employ HOMER & performance evaluated both at the measurement level and at the fragmentation level

$P(n_f)$ hadrons from single string frag.





Many QCD subtleties to incorporate and understand

Thrust specific:

- Sudakov Shoulder Resummation
- Three Jet Power Corrections
- Shape Function Parametrization

In general:

- Choice of priors
- Estimating theory uncertainties
- Non-perturbative corrections
- Treatment of mixed-moments
- Convergence of moment expansion

Tackmann: 2411.18606
Lim, Poncelet: 2412.14910
…

Future Directions:

- Multi-differential distributions
- Alternative basis functions
- Inclusive observables (e.g. EEC's)

Application: multi-differential Drell-Yan

Next: DY at hadron colliders — ATLAS MC production samples 35B Sherpa events at 13 TeV → upgrade and deliver precision particle-level events

Double-differential moments

$$\langle r_T^m \ln^n r_T \cdot (\Delta\phi)^p \ln^q \Delta\phi \rangle = \int d^2\vec{q}_T\, dm^2_{\ell\bar{\ell}}\, d\Omega_\ell\, \frac{d^5\sigma}{d^2\vec{q}_T\, dm^2_{\ell\bar{\ell}}\, d\Omega_\ell}\, r_T^m \ln^n r_T \cdot (\Delta\phi)^p \ln^q \Delta\phi$$

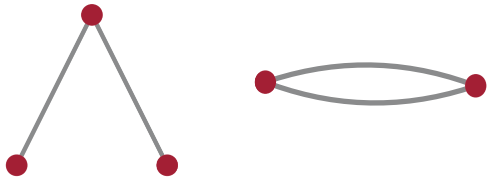
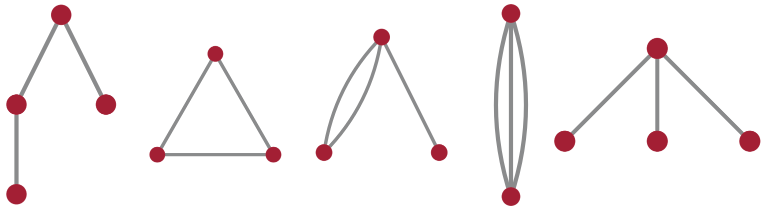
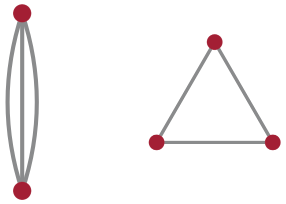
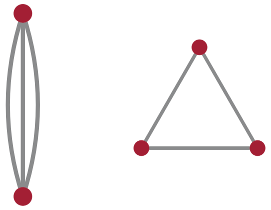
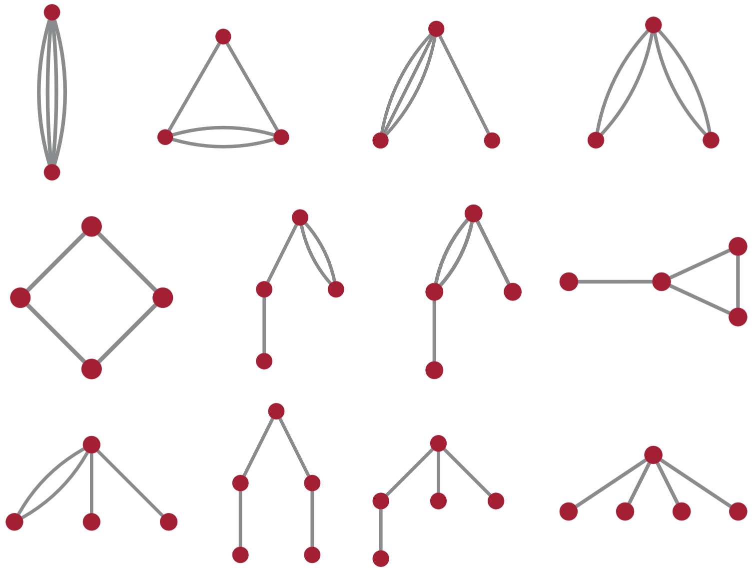
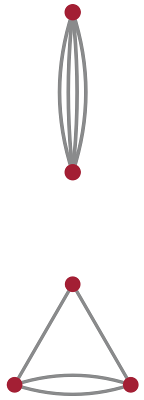
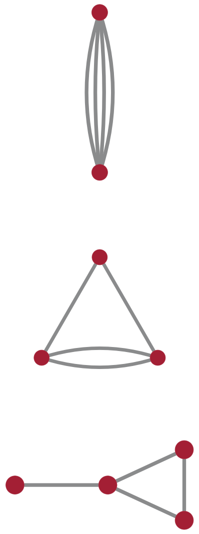
Precision targets: NNLO + $\mathrm{N^4LL}$ matched calculation

Cross-moments $\langle O_1 \cdot O_2 \rangle$ captures r_T and $\Delta\phi$ correlations

Final output: precision-upgraded particle-level events ready for ATLAS analysis chain

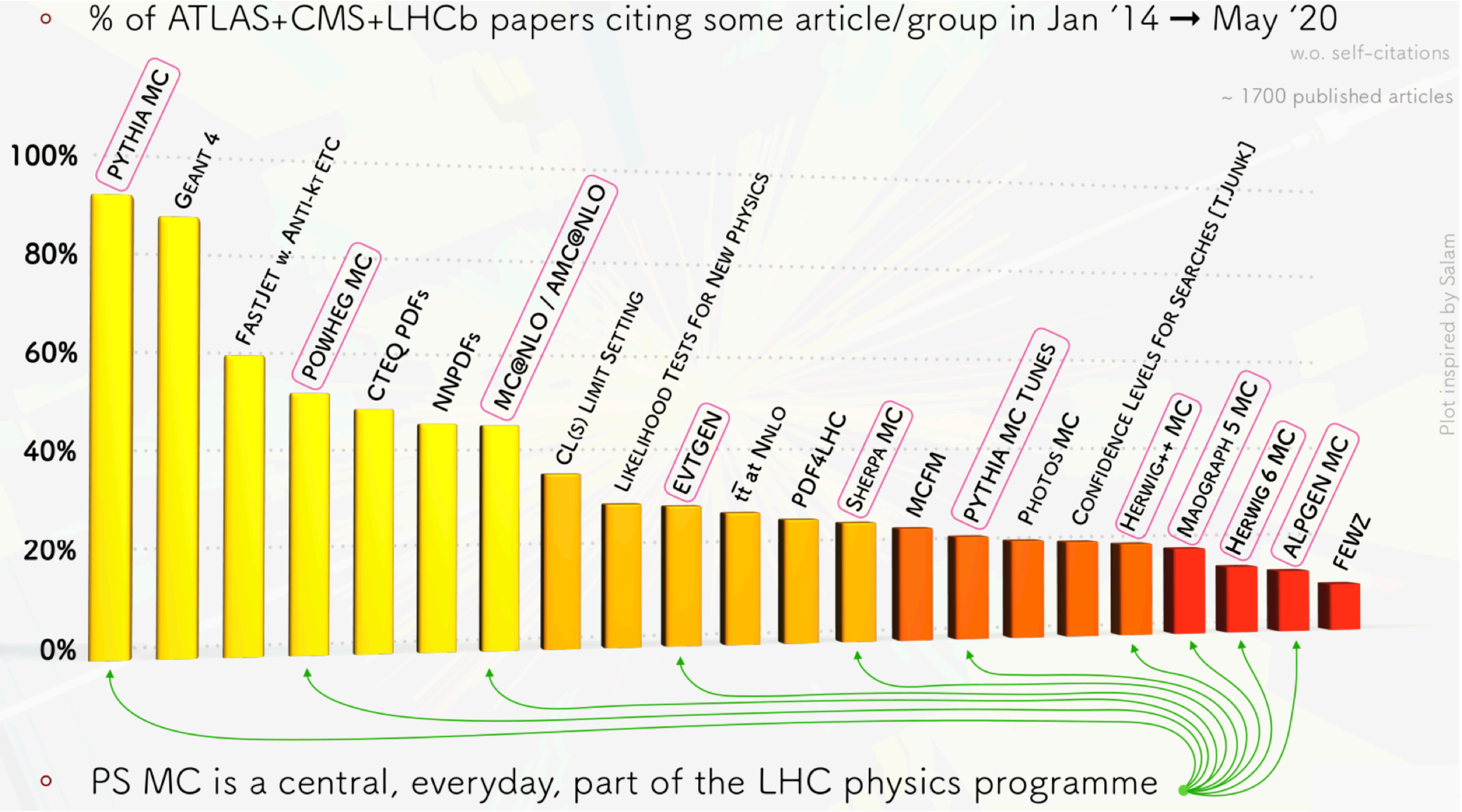
BA, Campbell, Höche, Ju, Thaler et al. *In preparation*

Basis of observables: Energy flow polynomials

Degree	Prime (connected)	Strongly-ordered	1-collinear
$d = 1$			
$d = 2$			
$d = 3$			
$d = 4$			

BA, Lee, Thaler, *In preparation*

MC Generators



- % of ATLAS+CMS+LHCb papers citing some article/group in Jan ’14 → May ‘20

w.o. self-citations

~ 1700 published articles

- PS MC is a central, everyday, part of the LHC physics programme

Keith Hamilton (2021)

Learning the Lund string FF

First instinct: Learn the fragmentation PDF from data

$$p_{\text{Lund}}(x) \to p_{\text{ML}}(x)$$

First step: Show it can be done for "**perfect data**".

Use **simulated first hadronizations** in $e^+e^- \to q\bar{q}$

Checks feasibility + provides an initial model to **correct in data**

We are using simulation to introduce **inductive bias** → Improve over the existing empirical model by first mapping it to a learnable model

Measurement Comparisons

Our simulation and data consist of **full events** — we use different Pythia parameters to generate simulation and pseudo-data

We consider three possibilities for available measurements:

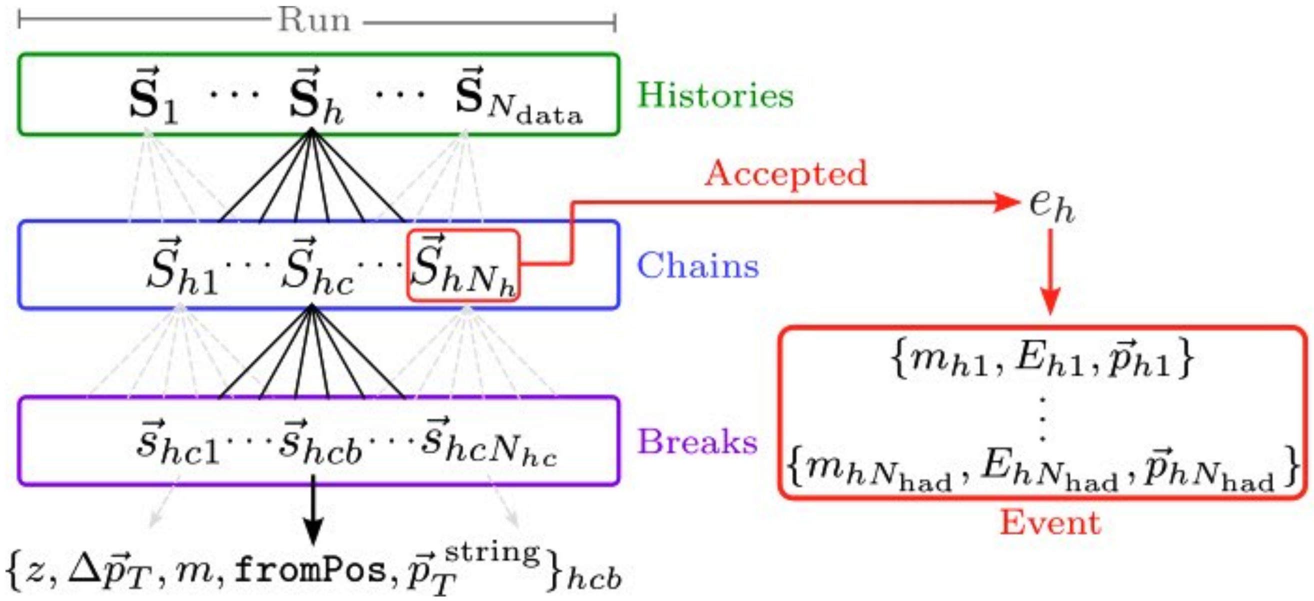
1) High-level observables (thrust, multiplicity,...) available only through 1D histograms (corresponds to available data)
2) High-level observables available on an event-by-event basis
3) The complete set of **particles with associated four-momenta** (the **point cloud** representation).

For all three cases, we employ HOMER & performance evaluated both at the measurement level and at the fragmentation level

Learning from data: HOMER

Histories and Observables for Monte-Carlo Event Reweighting: Take **Pythia** as baseline and **reweight**

We restrict ourselves to the $q\bar{q}$ string emitting pions **and record all information for simulation but** for pseudo-data **only** record observable quantities (no peeking at unobservable internals)



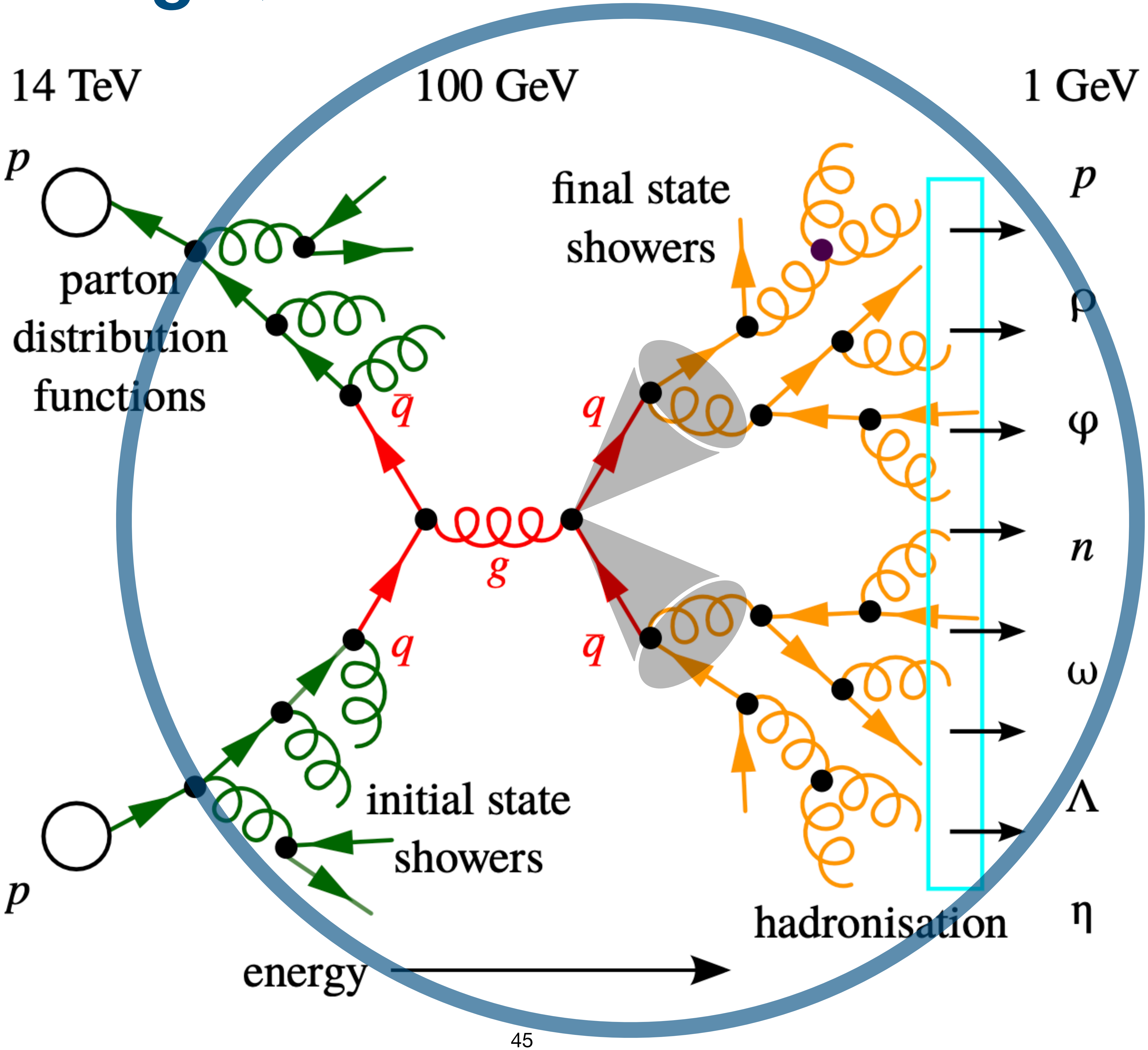
Full sequence of branchings up to perturbative cutoff

ID string segments between partons - several ways to hook up the string pieces so you sample candidate chain topologies

Sample string break configurations, where each s are Lund-model variables for that break

From accepted break configuration you build the full list of hadrons

Factorizing QCD



Shape of an emission

Emission:

$$(\tilde{p}_q, \tilde{p}_{\bar{q}}) \to (p_q k, k p_{\bar{q}})$$

with $k^\mu = z_q \tilde{p}^\mu_q + z_{\bar{q}} \tilde{p}^\mu_{\bar{q}} + k^\mu_t$

Degrees of freedom:

1) Rapidity: $\eta = \frac{1}{2} \log \frac{z_q}{z_{\bar{q}}}$

2) Transverse momentum: k_t

3) Azimuth: ϕ