QCD Theory Meets Information Theory
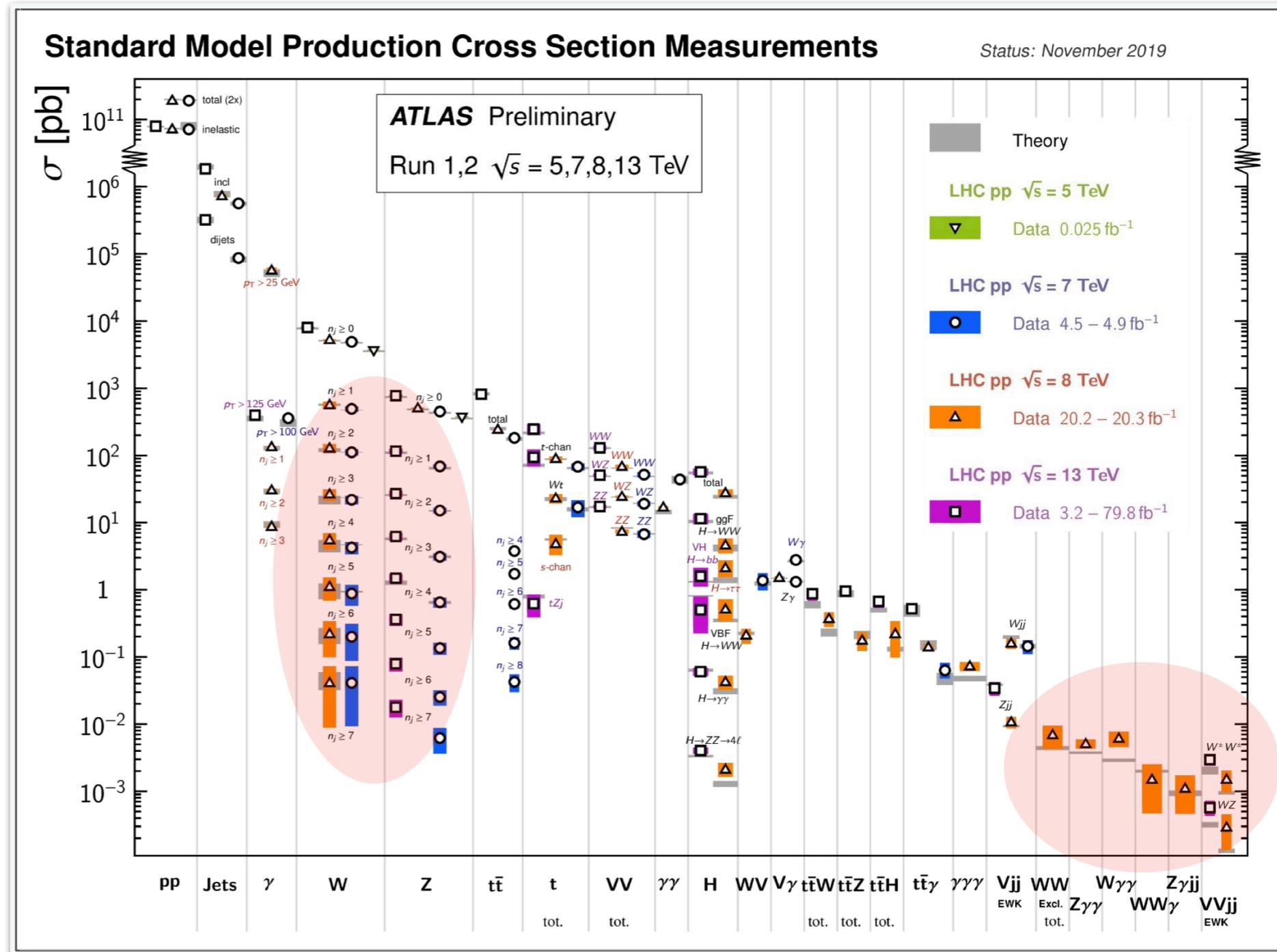
Benoît Assi

RPP, Mar 13, 2026

High-Luminosity LHC era

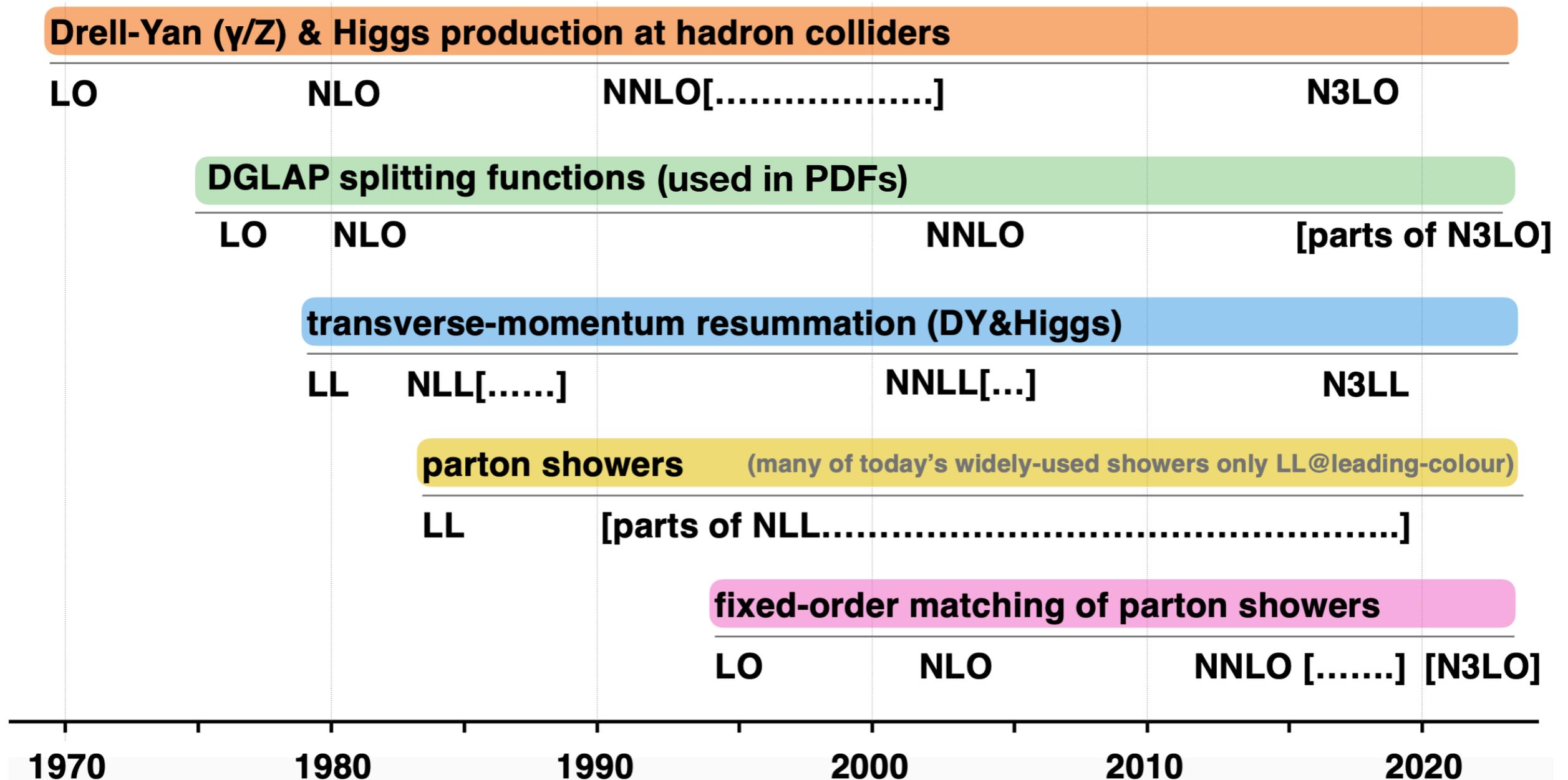Theory predictions and measurements reaching **incredible level of precision**

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

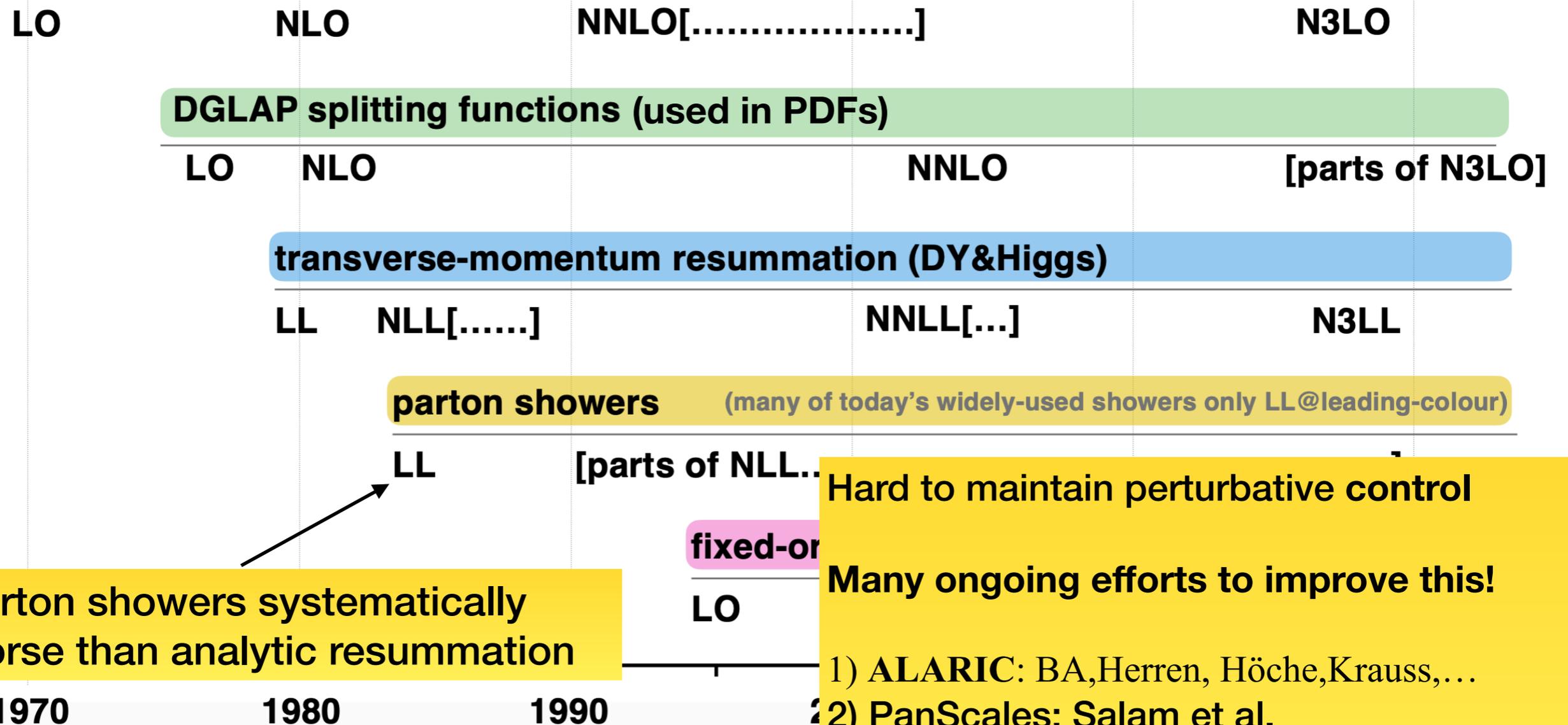We will need **higher precision** QCD predictions!

Quantifying precision

Current status of (resummed) perturbative precision in QCD

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]

1970	1980	1990	2000	2010	2020

Quantifying precision

Current status of (resummed) perturbative precision in QCD

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

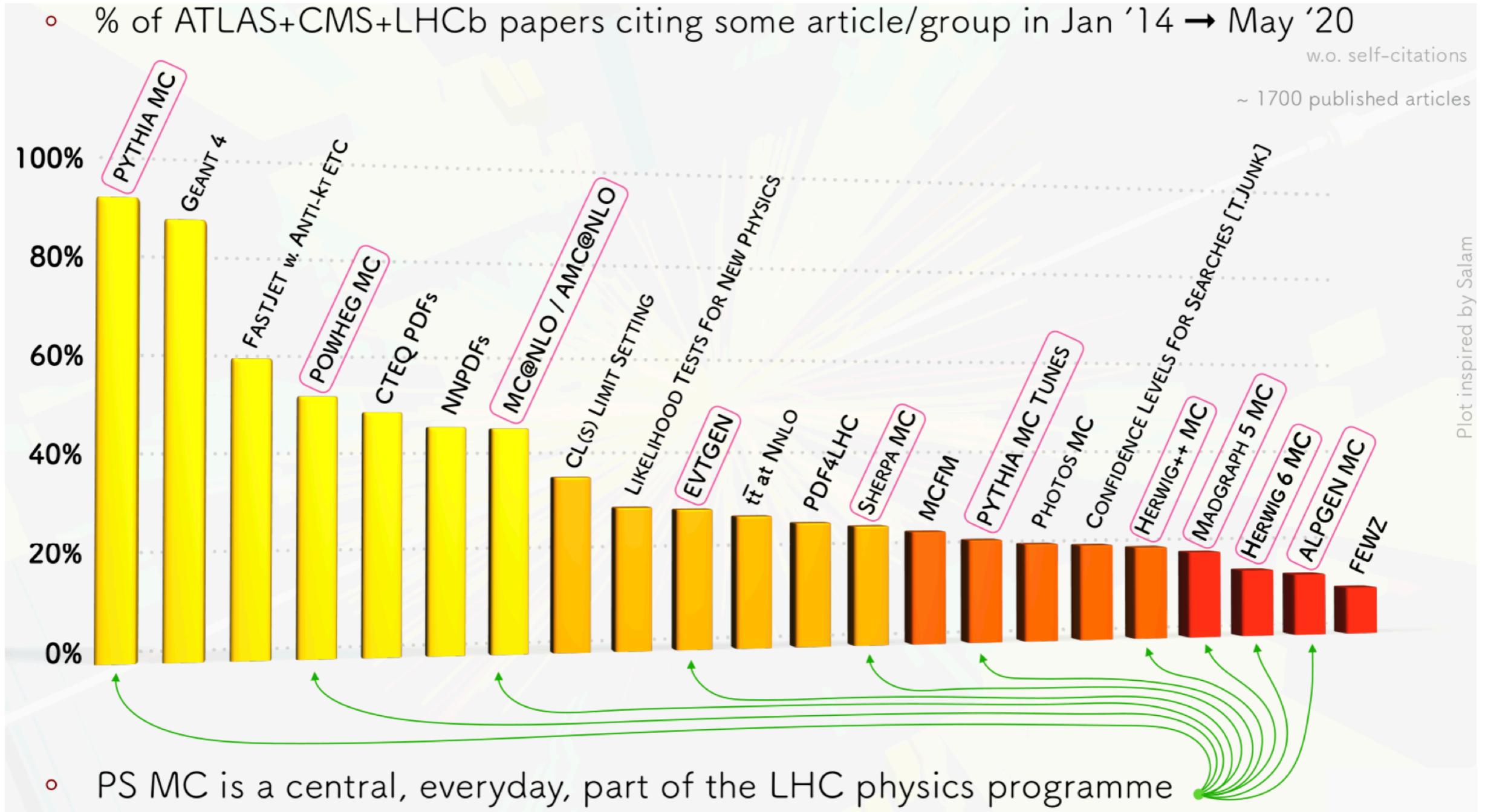| LO	NLO	NNLO[..................]	N3LO

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...]	

Parton showers systematically worse than analytic resummation

fixed-o...

LO

Hard to maintain perturbative **control**

Many ongoing efforts to improve this!

1) **ALARIC**: BA,Herren, Höche,Krauss,…
2) PanScales: Salam et al.
3) Deductor: Nagy et al.
4) Apollo: Hamilton et al.
…

1970	1980	1990	

MC Generators

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

w.o. self–citations

~ 1700 published articles

Plot inspired by Salam

PYTHIA MC · GEANT 4 · FASTJET w. ANTI-kt ETC · POWHEG MC · CTEQ PDFs · NNPDFs · MC@NLO / AMC@NLO · CL(S) LIMIT SETTING · LIKELIHOOD TESTS FOR NEW PHYSICS · EVTGEN · $t\bar{t}$ at NNLO · PDF4LHC · SHERPA MC · MCFM · PYTHIA MC TUNES · PHOTOS MC · CONFIDENCE LEVELS FOR SEARCHES [T.JUNK] · HERWIG++ MC · MADGRAPH 5 MC · HERWIG 6 MC · ALPGEN MC · FEWZ

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

Keith Hamilton (2021)

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:

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

LL   NLL   ...

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

Non-perturbative (requires modelling)

$Q \equiv$ 100 GeV $\rightarrow 1$ TeV

BSM

$Q \gg \mu_{\rm NP}$

$\mu_{\rm NP} \sim$ 1 GeV

m_t
m_H
$m_{W/Z}$

m_b

m_c

m_π

Hard process, matching

Parton shower

Hadronisation MPI/UE
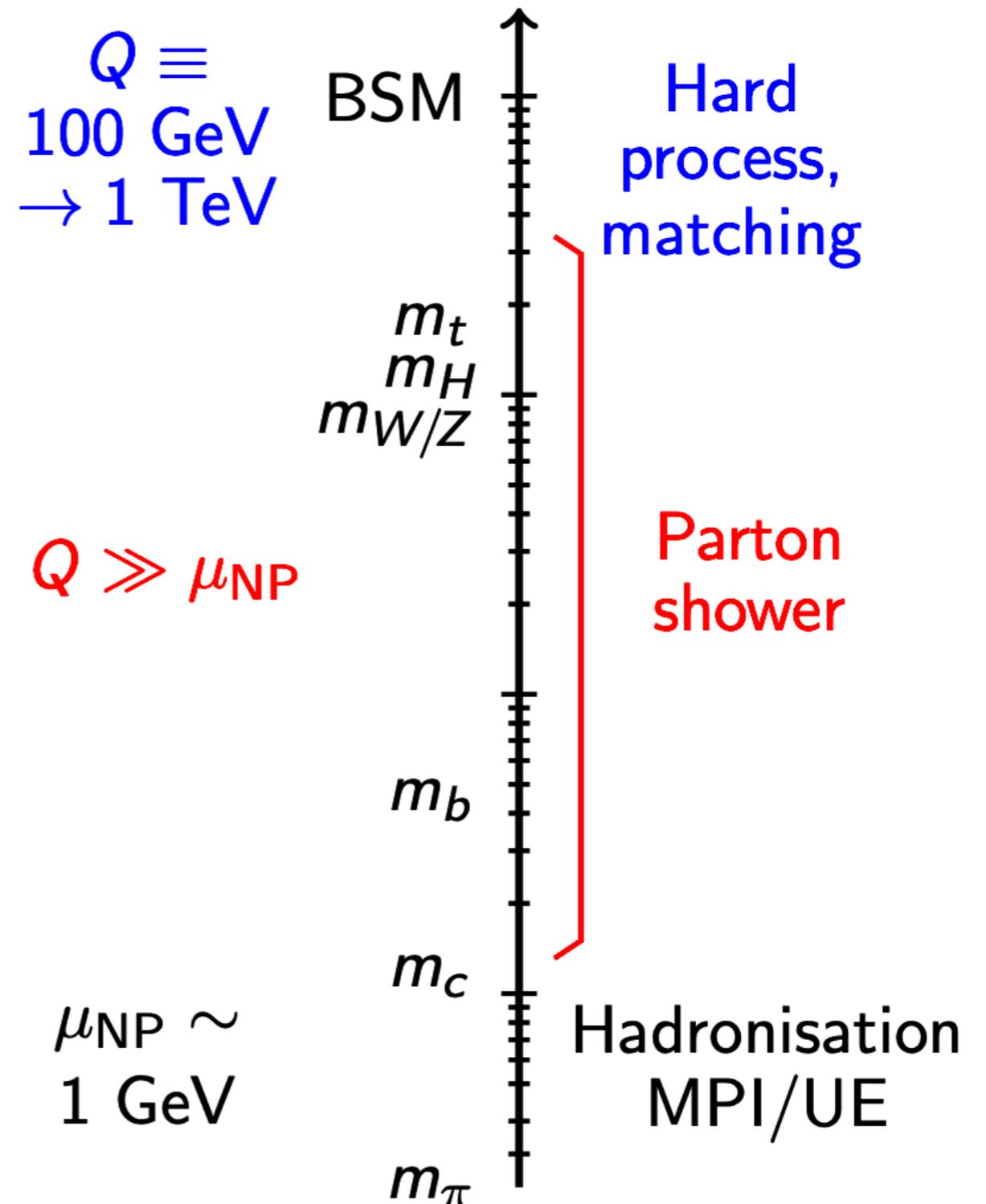
Parton Shower

Mechanism: Gluon emission $\leftrightarrow$ 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
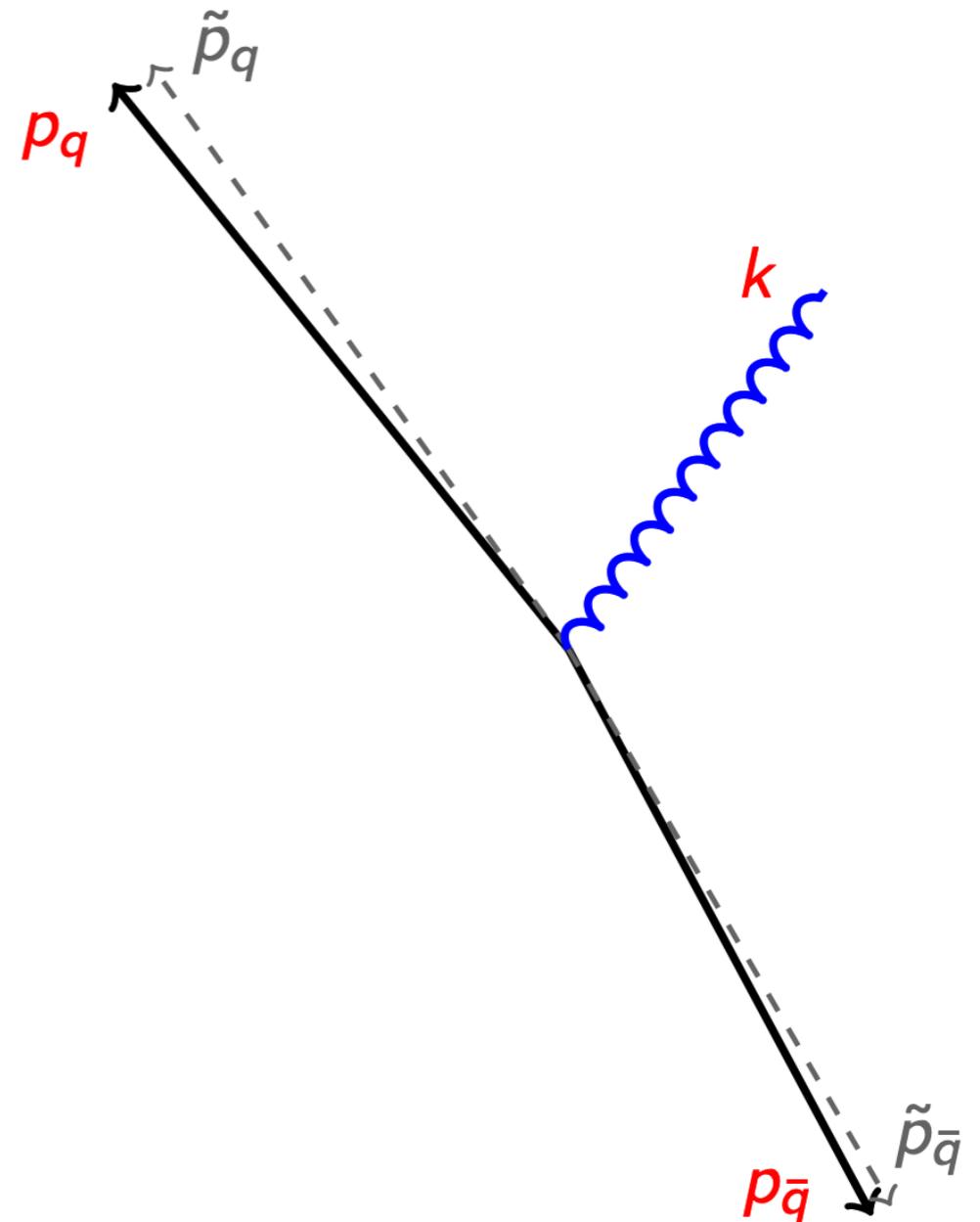
Beyond Leading Accuracy

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)

$\ln k_t$

(only half the primary Lund plane for simplicity)

η

separated in any direction

$\geq \Delta$
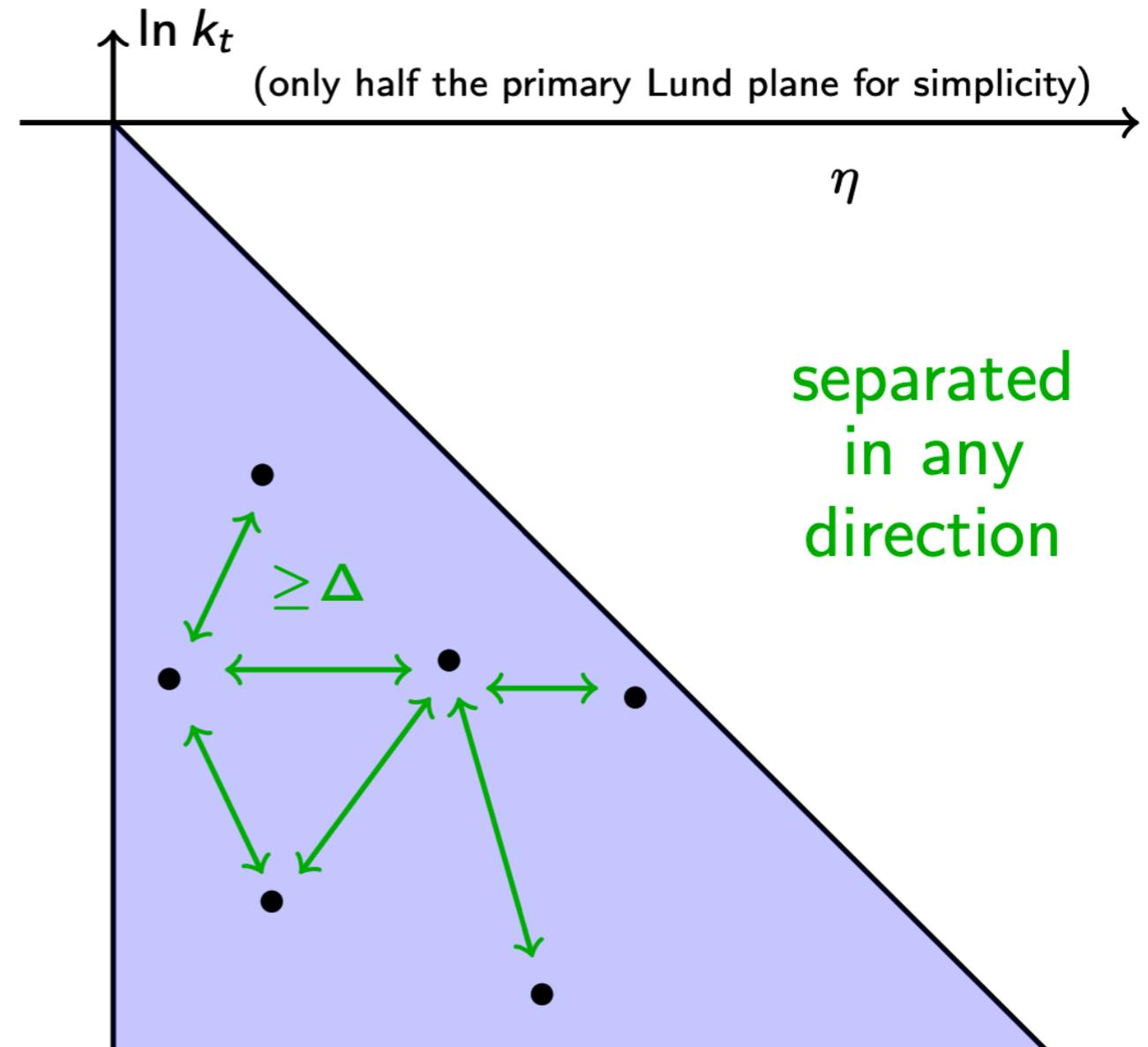
An NLL accurate algorithm in SHERPA

ALARIC momentum mapping:

Recoil compensated **globally** by neg. sum of multipole momenta $\tilde{K}$ including splitter

K receives kick of $\mathcal{O}(k_\perp)$ from the emission that can be understood as a Lorentz TF

Since K chosen such that $K^2 \gg k_\perp^2$ it does **not** affect topology of **previous emissions** in the Lund plane because its effect scales as $k_\perp^2/K^2$

Can analytically **prove** that this algorithm is **NLL accurate** in both massless and **massive** case

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

Before emission
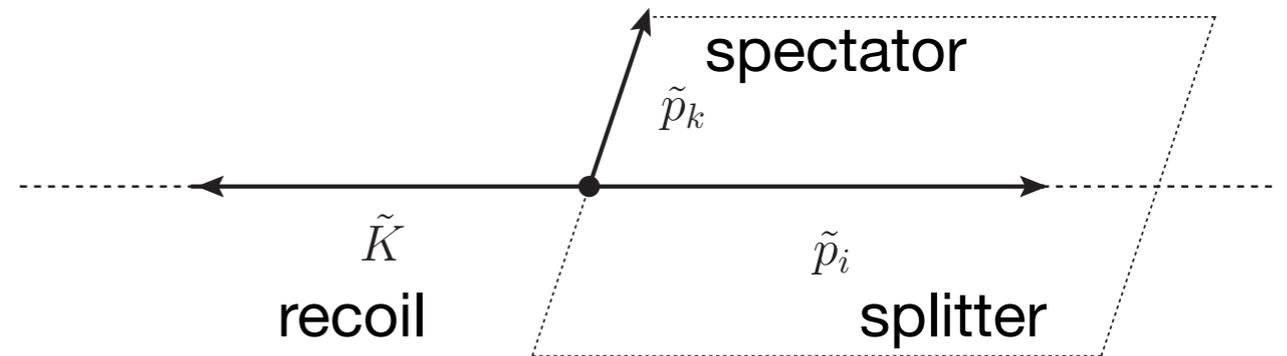
After emission
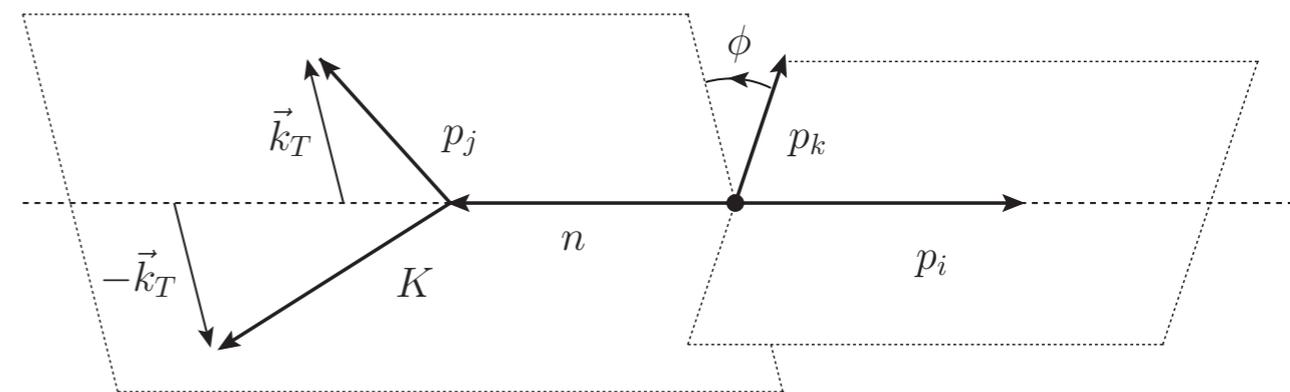
Subtleties of heavy flavour evolution

Both **high-energy** and **threshold** regime require accurate descriptions

Multiple prescriptions - varying success in describing experiments

Norrbin, Sjöstrand; Gieseke, Stephens; Schumann, Krauss; Gehrmann-deRidder, Ritzmann, Skands
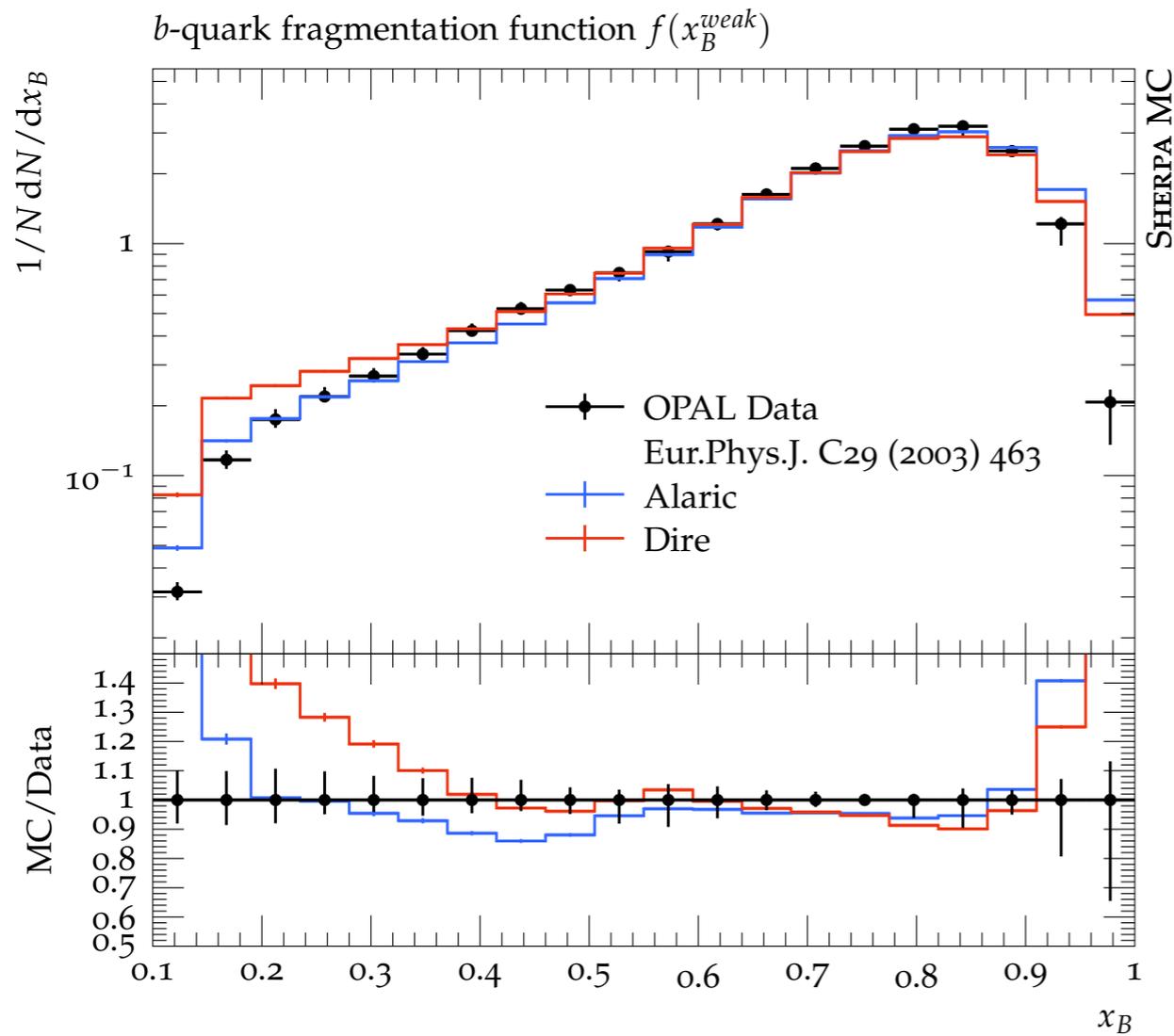
Required for proper description of bottom and charm jet production and quark **fragmentation functions**

Generalized ALARIC for massive flavours: careful analysis of the branching probabilities, matching of soft & quasi-collinear limits, NLO matching (intricate integrals)

BA, Höche PRD 109 (2024)

Norrbin, Sjöstrand hep-ph/0010012
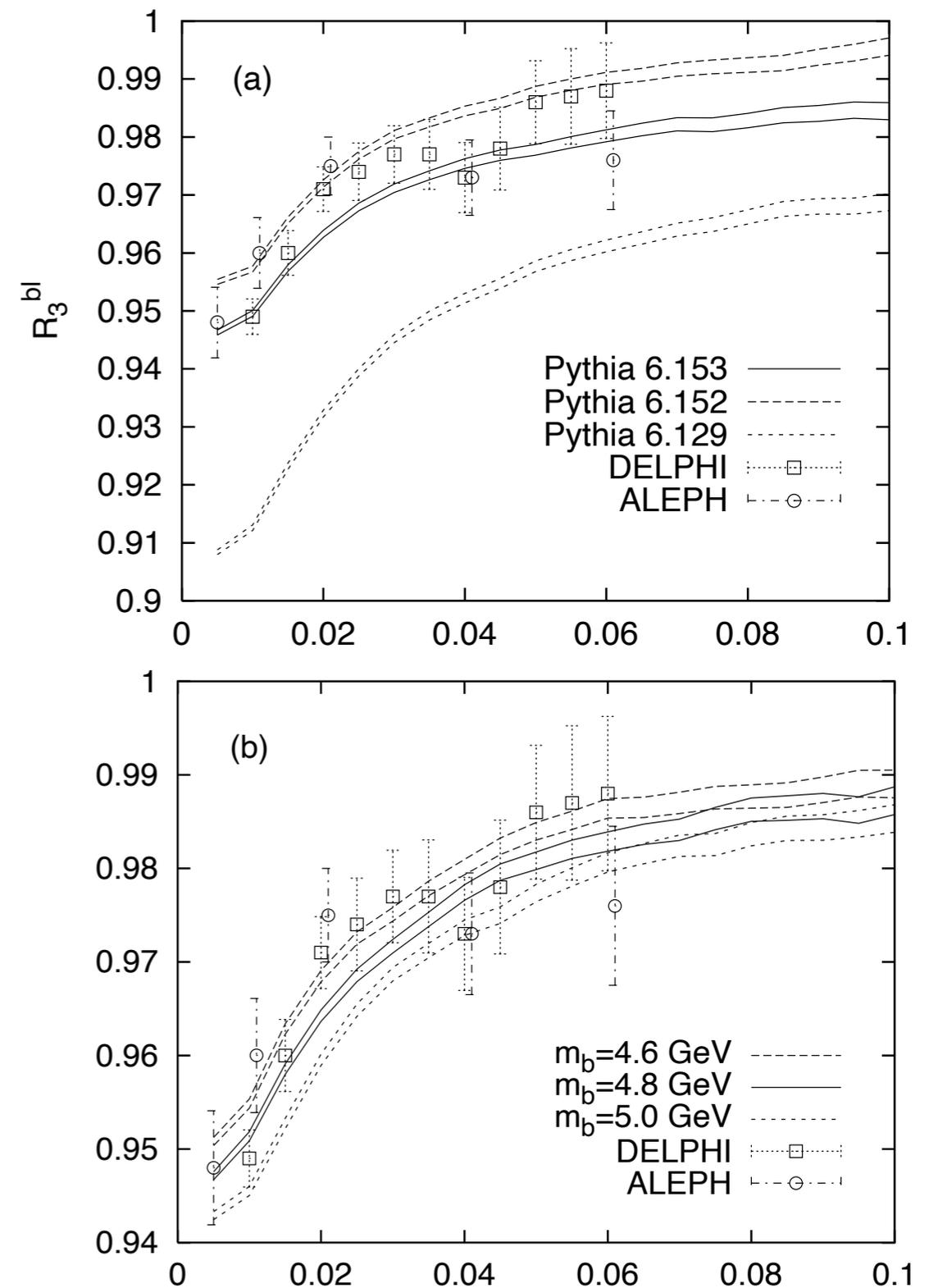
Experimental comparison

b-quark fragmentation function $f(x_B^{weak})$

OPAL Data
Eur.Phys.J. C29 (2003) 463
Alaric
Dire

b quark fragmentation function $f(x_B^{weak})$

ALEPH Data
Phys.Lett. B512 (2001) 30
Alaric
Dire
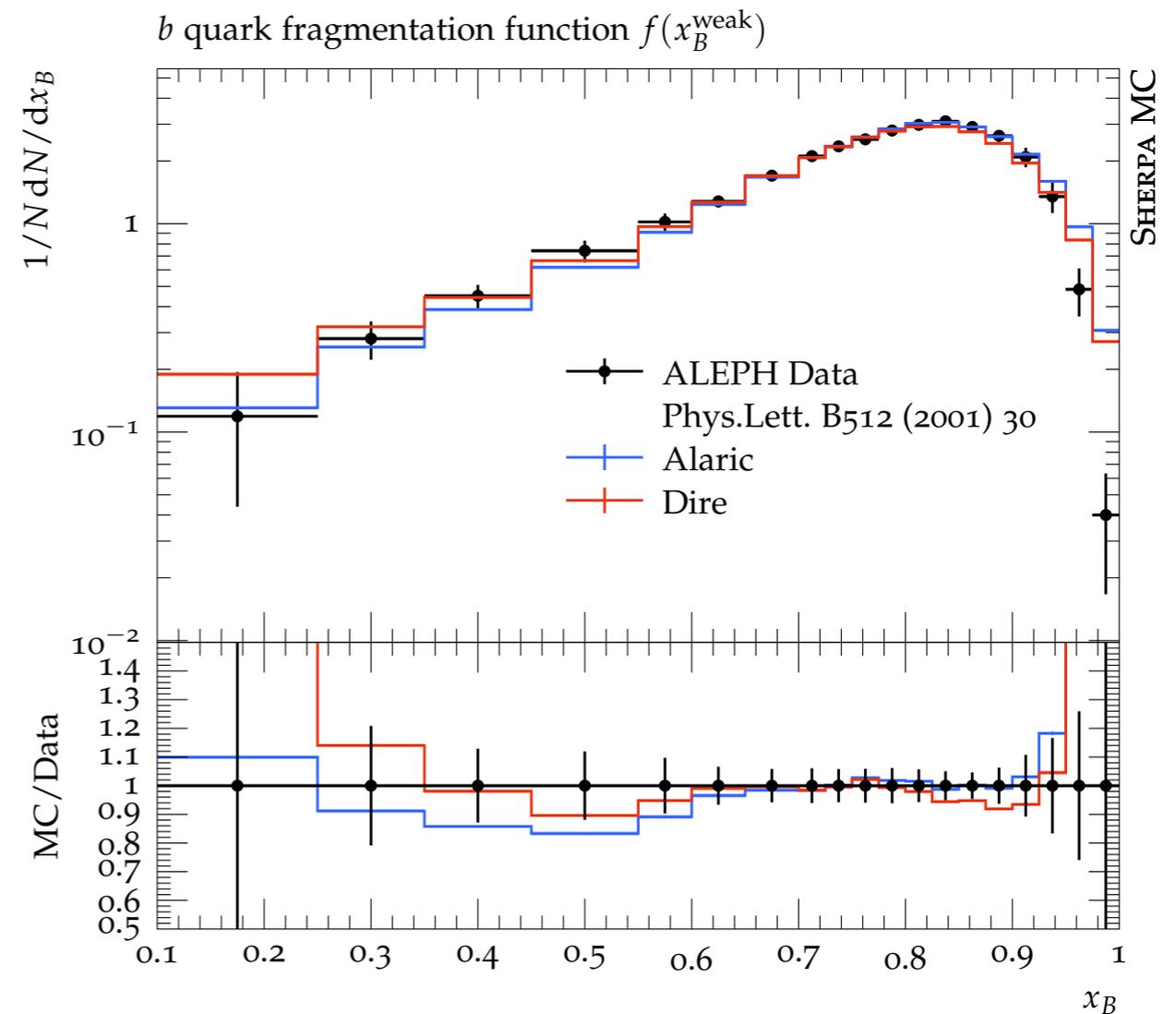
Better agreement for Alaric predictions
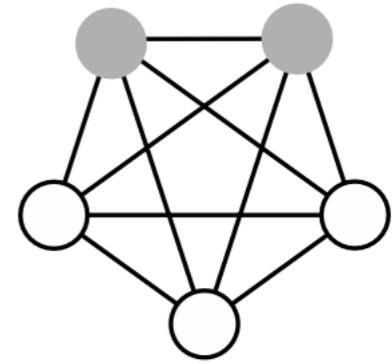with experimental data (same had. tune)

BA, Höche PRD 109 (2024)

*Sins lie here! We'll come back
to this at the end*

11

Can we incorporate higher precision theory information?

Fixed-order Perturbation Theory

Analytical & Numerical Resummation

Particle-level Simulation
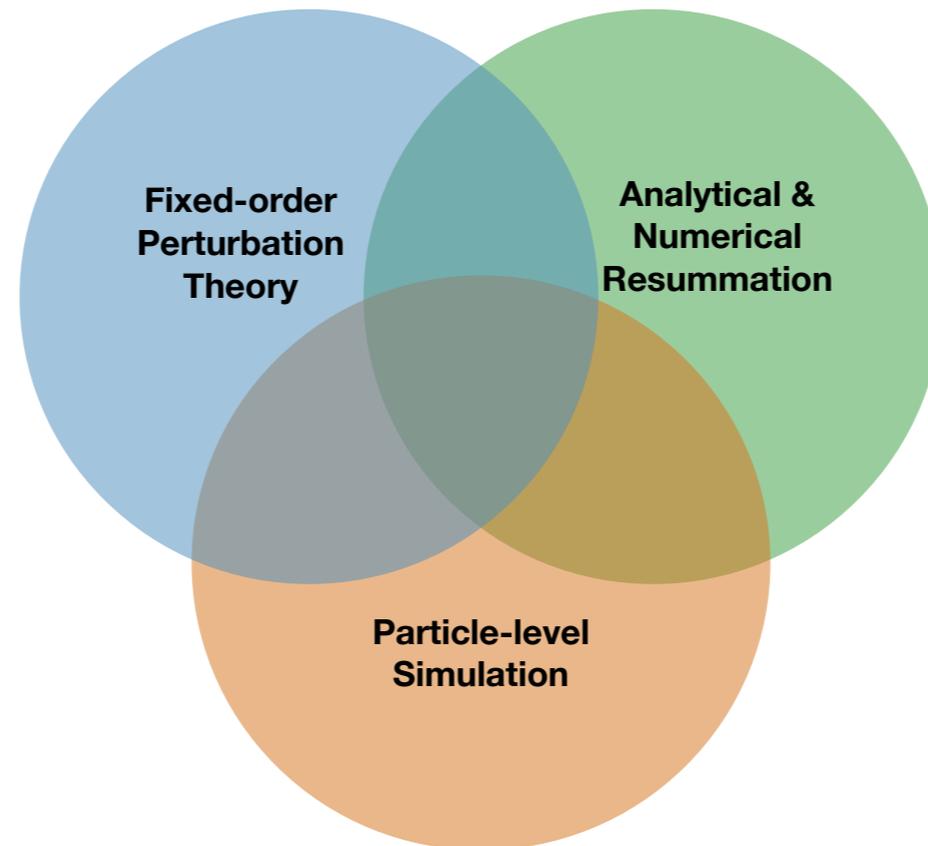
QCD Theory meets Information Theory

Boltzmann machine: Generative model samples according to stat dist

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

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

Prior

Sets Partition Function

$$p(x) = q(x) \exp\left[-\beta_0 - \beta_1 f_1(x) - \beta_2 f_2(x) - \dots\right]$$

Lagrange Multipler   Constrained Quantity

HEP → ? ← ML

Sudakov structure: Plucking a random event-shape observable distribution

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

Sudakov factor

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!

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**:**

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

Target moment **Current moment**

Extremized if prior (generator) distribution is modified as:

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

Loss function given by

$$\mathcal{L} = \sum_j \left(\frac{c_j - d_j}{c_j + d_j}\right)^2$$

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

HEP ⟶ ⟶ ML
?

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 $\Rightarrow$ Strong coupling

Characterises *FO information*

Logarithmic Moments

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

Logarithmic Mean $\Rightarrow$ Sudakov peak

Characterises *Resummed information*

Logarithmic moments of thrust at LEP

Generic Sudakov-log observable v takes the form

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

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
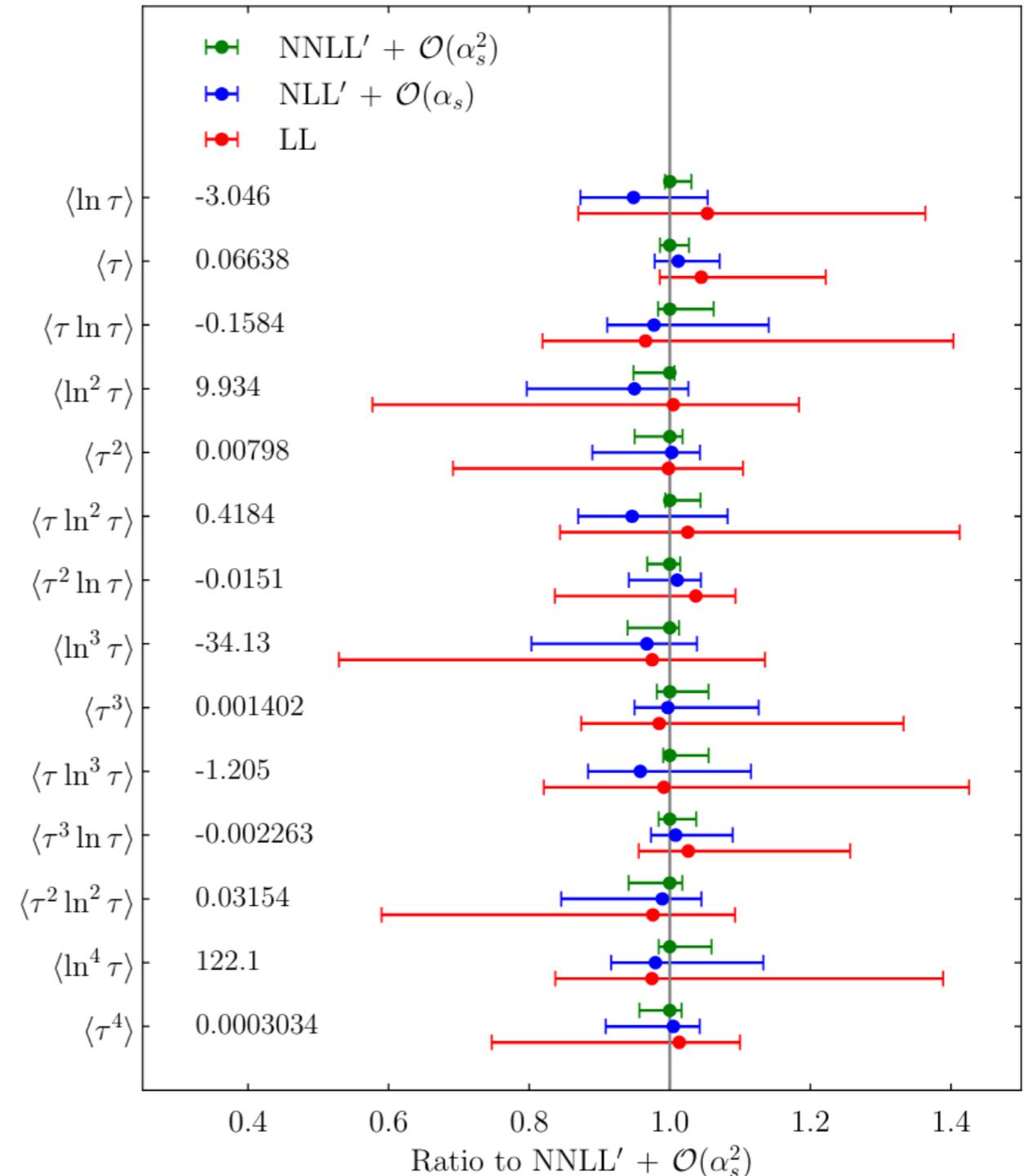
An example of thrust at LEP

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

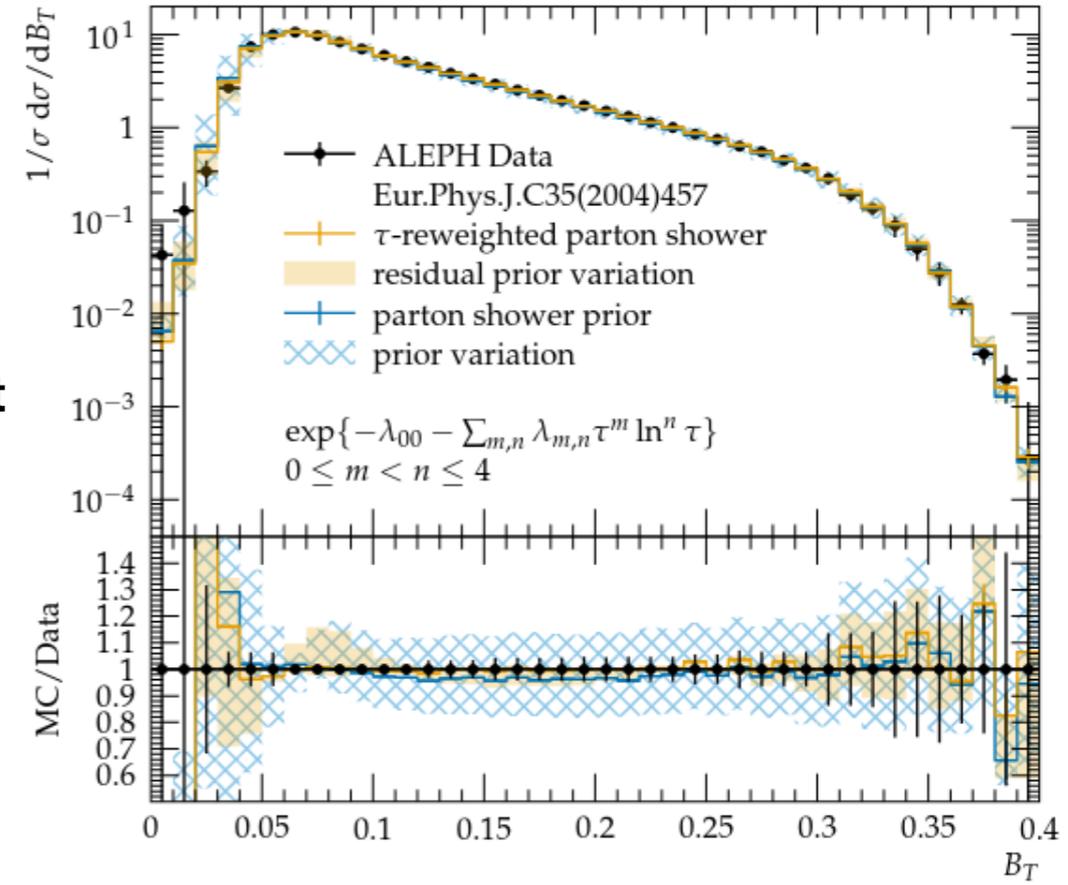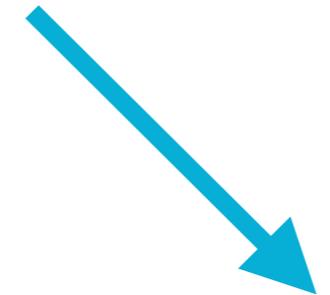|	$Log[\tau]^0$	$Log[\tau]^1$	$Log[\tau]^2$	$Log[\tau]^3$	$Log[\tau]^4$
τ^0	1	$Log[\tau]$	$Log[\tau]^2$	$Log[\tau]^3$	$Log[\tau]^4$
τ^1	τ	$\tau\,Log[\tau]$	$\tau\,Log[\tau]^2$	$\tau\,Log[\tau]^3$	$\tau\,Log[\tau]^4$
τ^2	τ^2	$\tau^2\,Log[\tau]$	$\tau^2\,Log[\tau]^2$	$\tau^2\,Log[\tau]^3$	$\tau^2\,Log[\tau]^4$
τ^3	τ^3	$\tau^3\,Log[\tau]$	$\tau^3\,Log[\tau]^2$	$\tau^3\,Log[\tau]^3$	$\tau^3\,Log[\tau]^4$
τ^4	τ^4	$\tau^4\,Log[\tau]$	$\tau^4\,Log[\tau]^2$	$\tau^4\,Log[\tau]^3$	$\tau^4\,Log[\tau]^4$
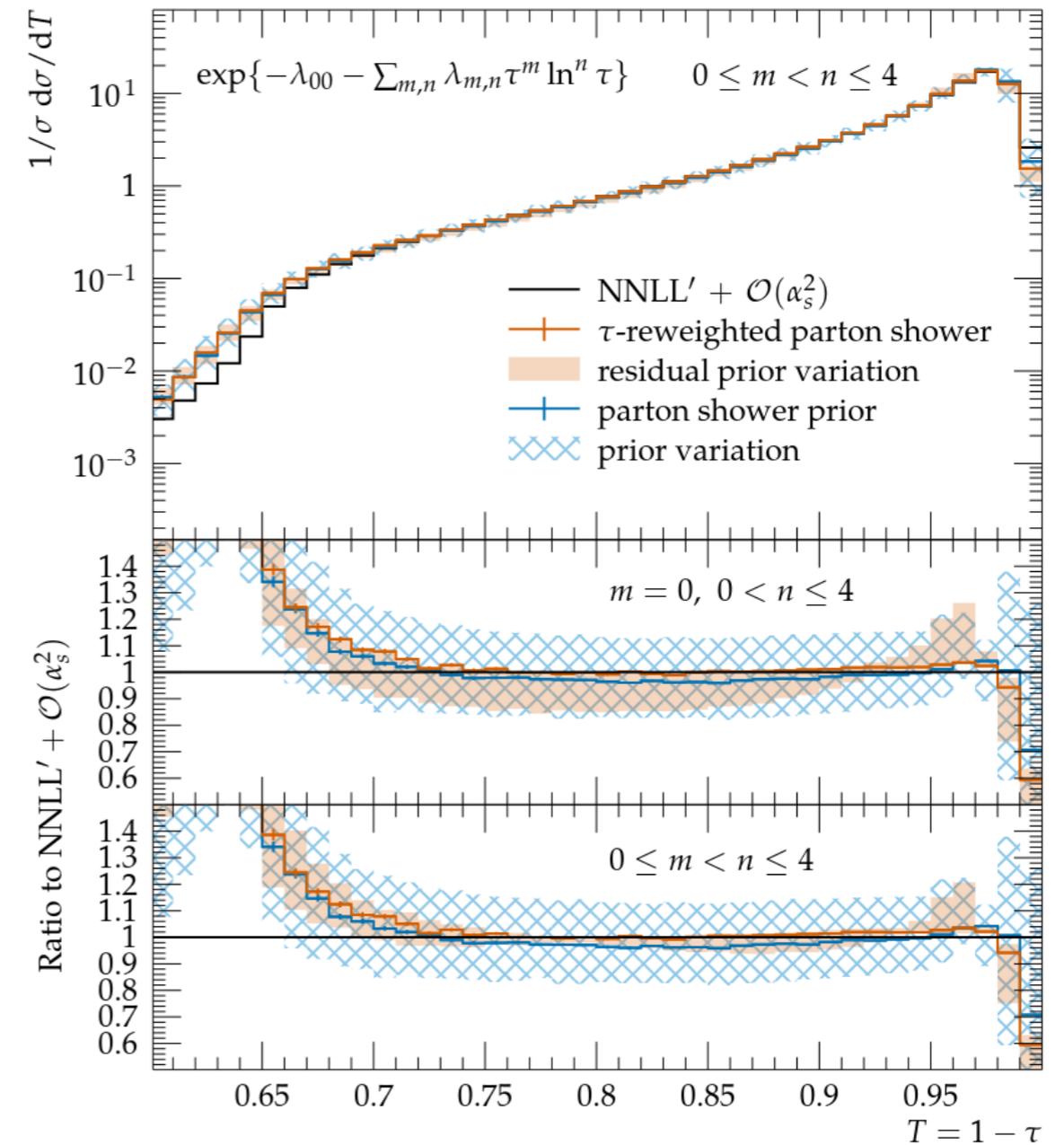
BA, Höche, Lee, Thaler PRL 132 (2025)

Impact on other observables
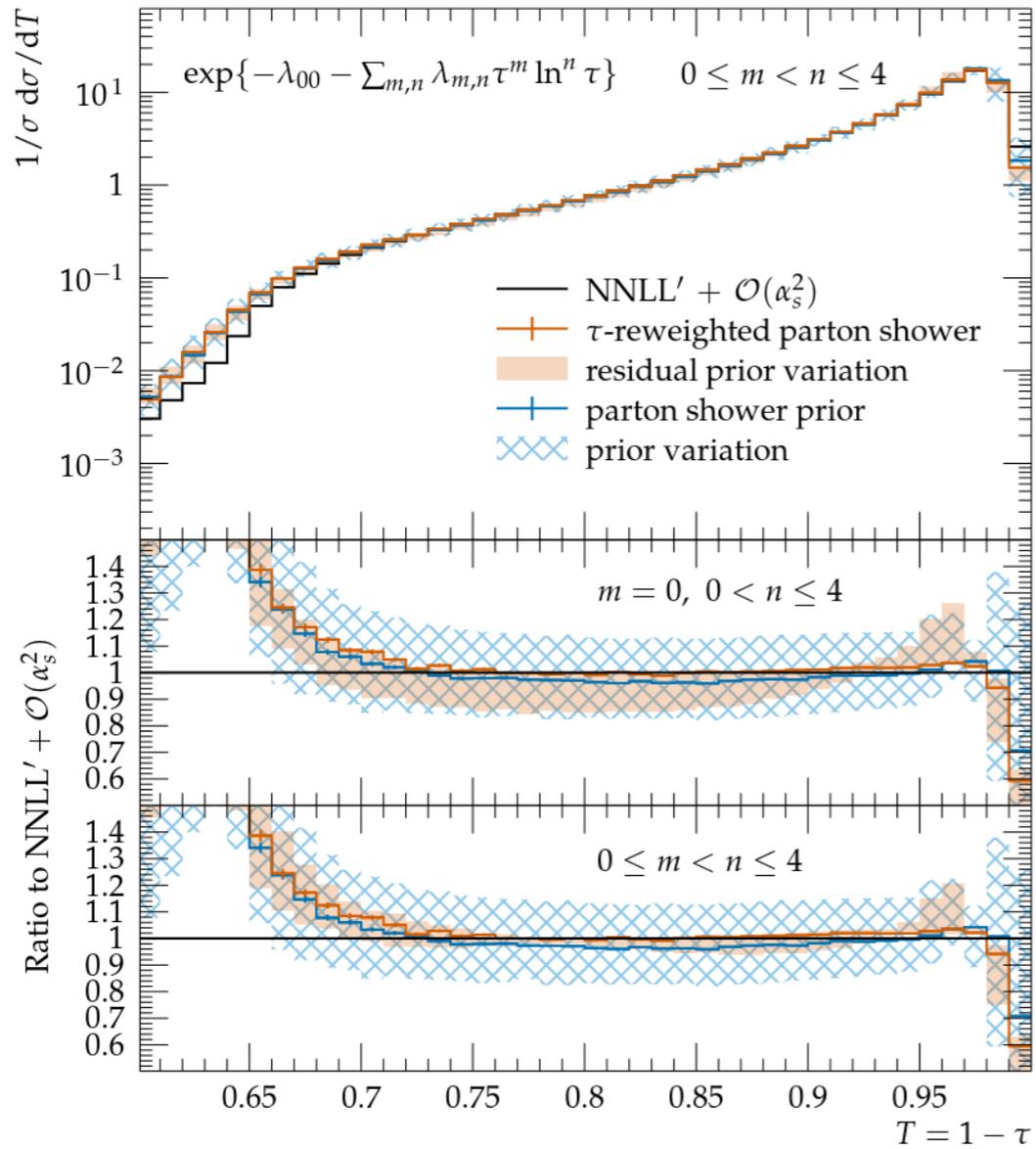
Broadening ≃ Thrust

Aplanarity ≄ Thrust
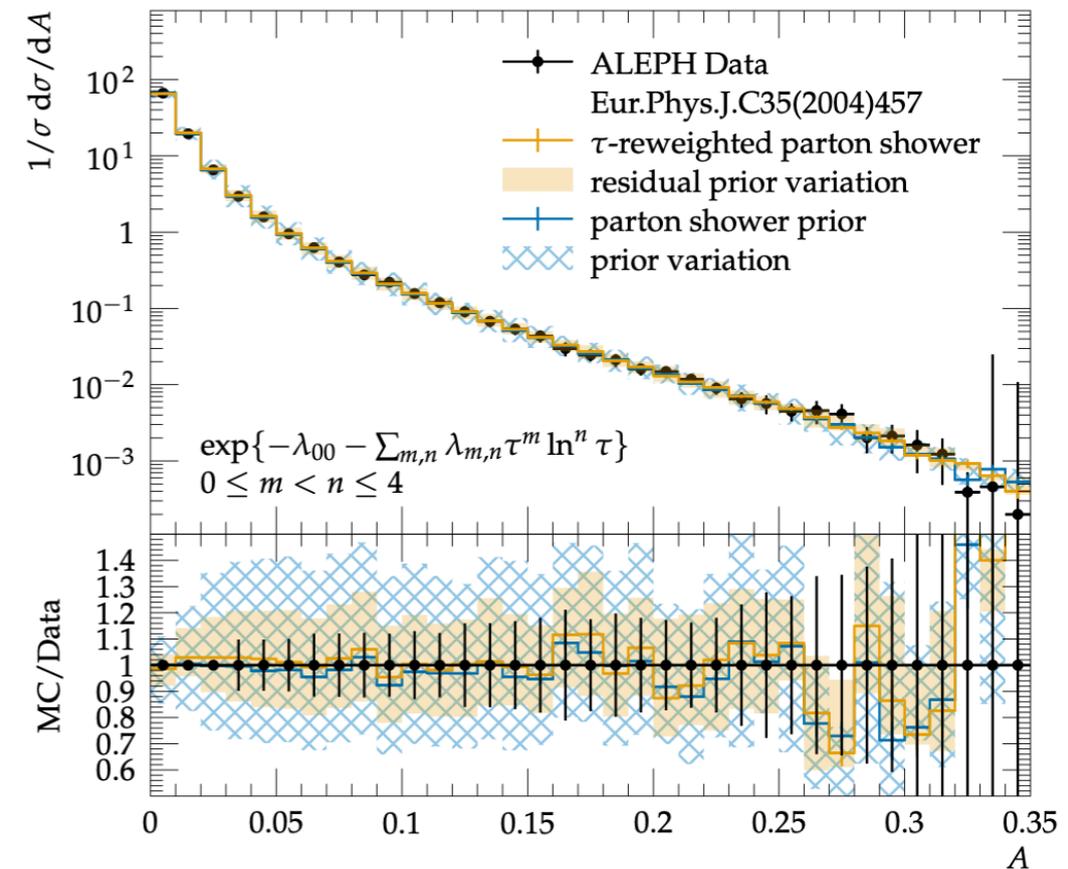
QCD Theory meets Information Theory

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 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)
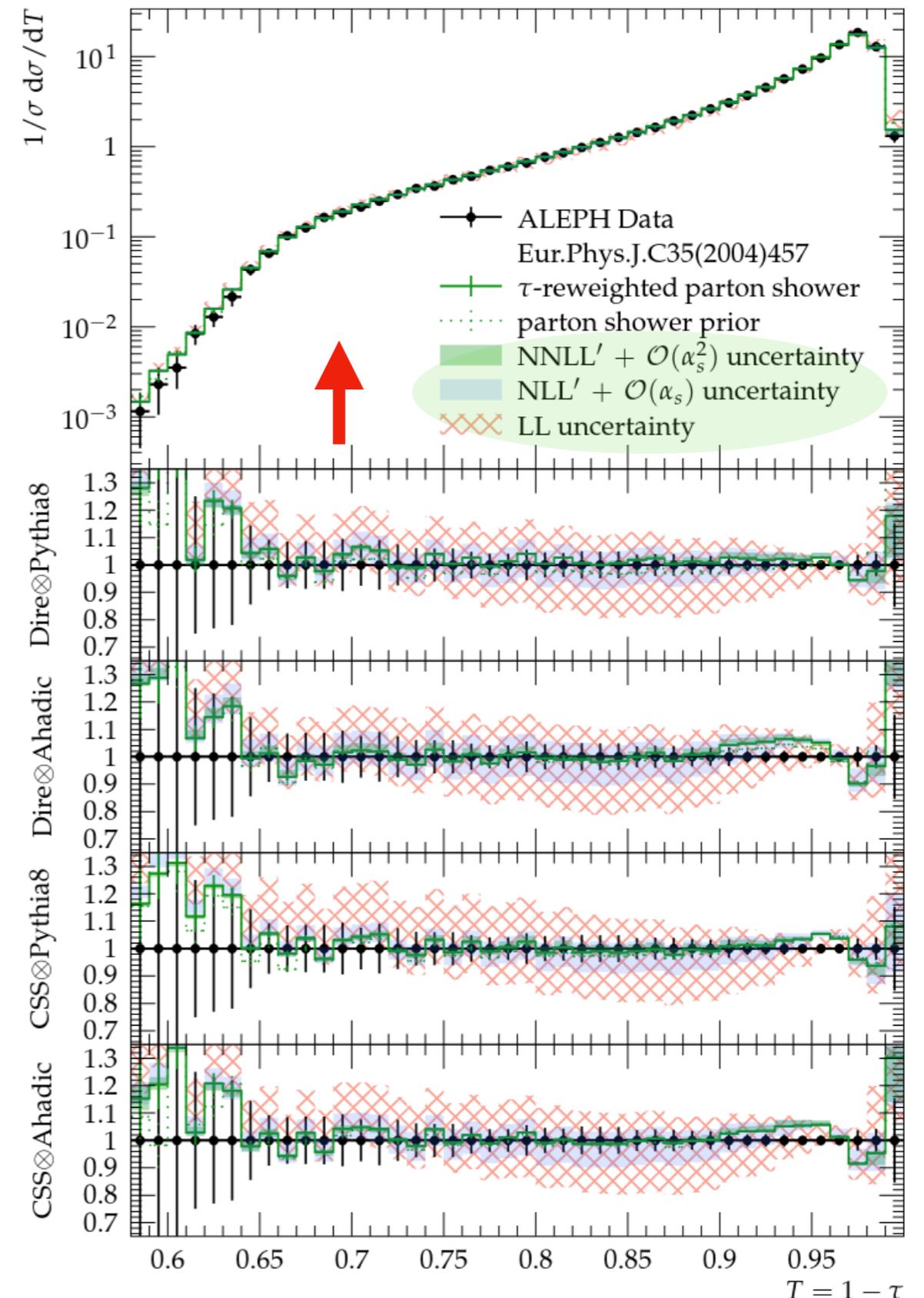
Multiple observables: how many to capture max information?

Basis of observables: Energy flow polynomials

Idea: Need systematic complete linear basis of IRC-safe observables for multi-observable constraints $\Rightarrow$ EFPs   Komiske, Metodiev, Thaler: hep-ph/1712.07124

$$\text{EFP}_G^{(\beta)} \equiv \sum_{i_1,\dots,i_k \in R} \prod_{v \in V} z_{i_v} \prod_{(a,b) \in E} \theta_{i_a i_b}^{\beta}$$
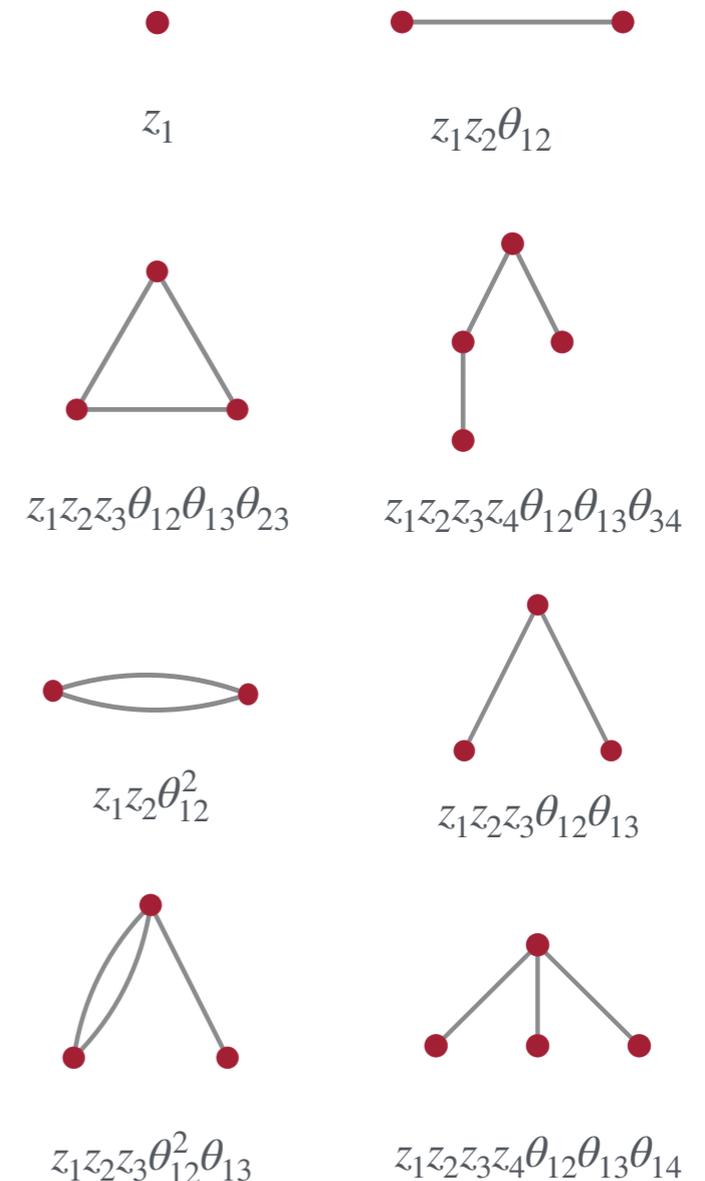
Sum over all particle assignments to graph G — graph topology encodes correlation structure

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

Cal, Thaler, Waalewijn: hep-ph/2205.06818

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

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


z_1

$z_1 z_2 \theta_{12}$

$z_1 z_2 z_3 \theta_{12} \theta_{13} \theta_{23}$

$z_1 z_2 z_3 z_4 \theta_{12} \theta_{13} \theta_{34}$

$z_1 z_2 \theta_{12}^2$

$z_1 z_2 z_3 \theta_{12} \theta_{13}$

$z_1 z_2 z_3 \theta_{12}^2 \theta_{13}$

$z_1 z_2 z_3 z_4 \theta_{12} \theta_{13} \theta_{14}$

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

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

Information saturation

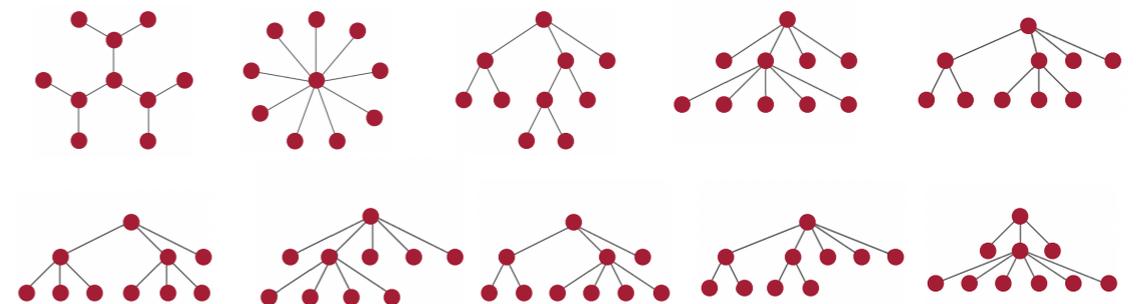**Key question:** how many EFP constraints before adding more gives diminishing returns?

Two-shower setup: Prior is violently broken Sherpa CSS shower collinear singularities removed, $g \to q\bar{q}$ channel disabled $\Rightarrow$ No double logs, wrong multiplicity
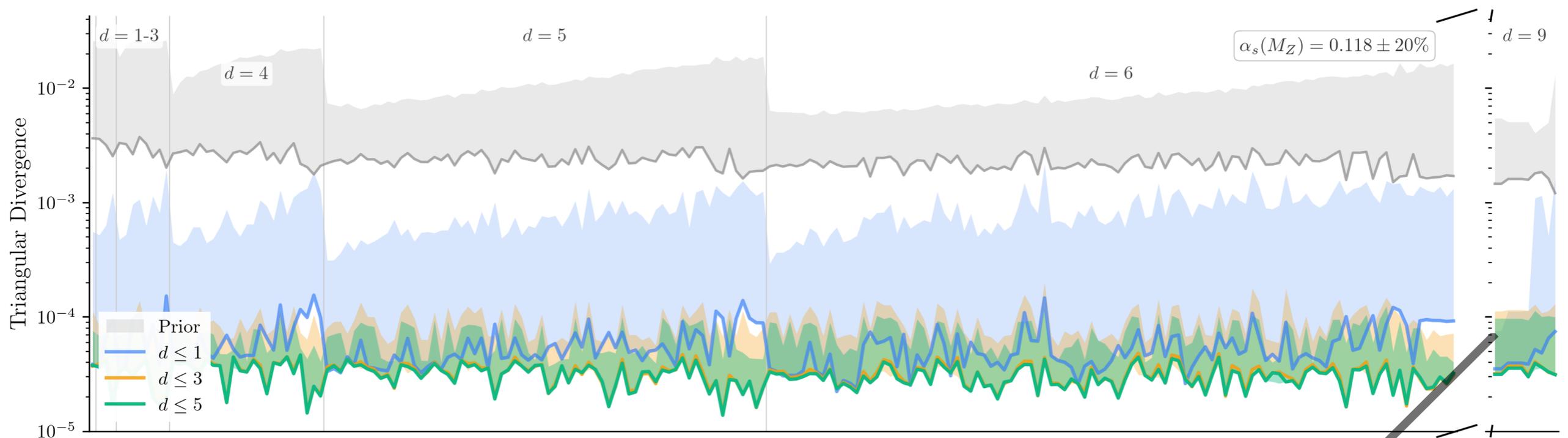
Train: 112 EFPs up to $d \le 5$ and **test:** $d \le 9$

Fix 4-moments each: $\langle \mathrm{EFP}_G \rangle$, $\langle \ln \mathrm{EFP}_G \rangle$, $\langle \ln^2 \mathrm{EFP}_G \rangle$, $\langle \mathrm{EFP}_G \ln \mathrm{EFP}_G \rangle$
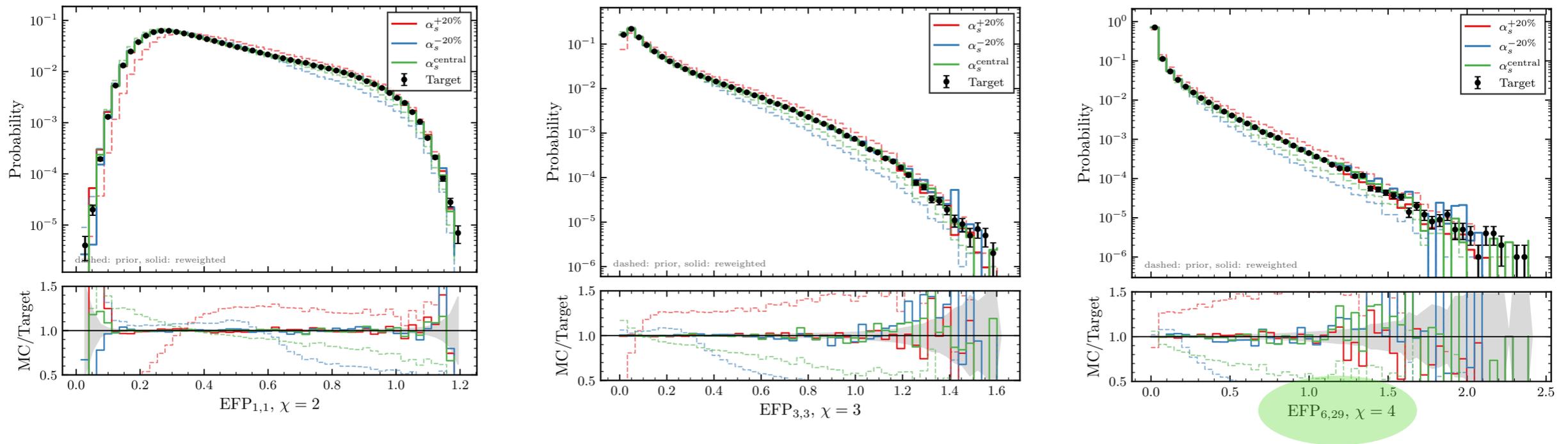
Takeaway: Saturation by $d \le 3$ and robust across $\alpha_s(M_Z) \pm 20\,\%$

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

BA, Lee, Thaler, *In preparation*

Information saturation

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

Two-shower setup: Prior is violently broken Sherpa CSS shower collinear singularities removed, $g \to q\bar{q}$ channel disabled $\Rightarrow$ No double logs, wrong multiplicity

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

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

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

Many degree 9 EFPs: we test a well-defined subset of homeomorphically irreducible trees

BA, Lee, Thaler, *In preparation*
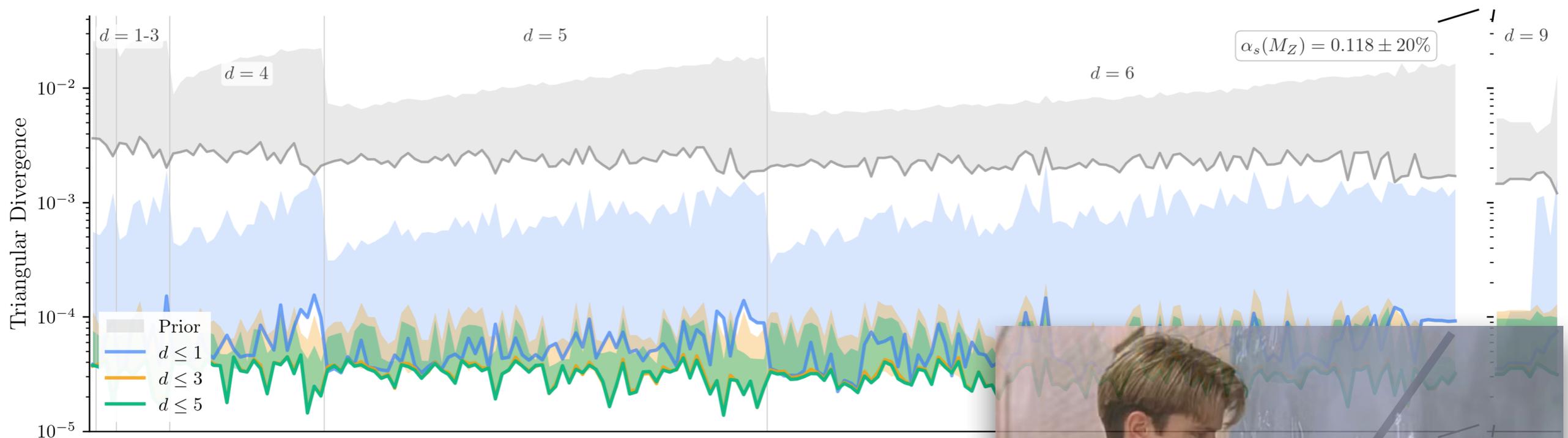
Information saturation

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

Two-shower setup: Prior *violently* broken Sherpa CSS shower, collinear singularities $1/(1-z)$ removed, $g \to q\bar{q}$ channel disabled $\Rightarrow$ No double logs, wrong 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 robus across $\alpha_s(M_Z) \pm 20\%$

$(1,0)$ $(2,1)$ $(3,3)$ $(4,6)$ $(5,10)$
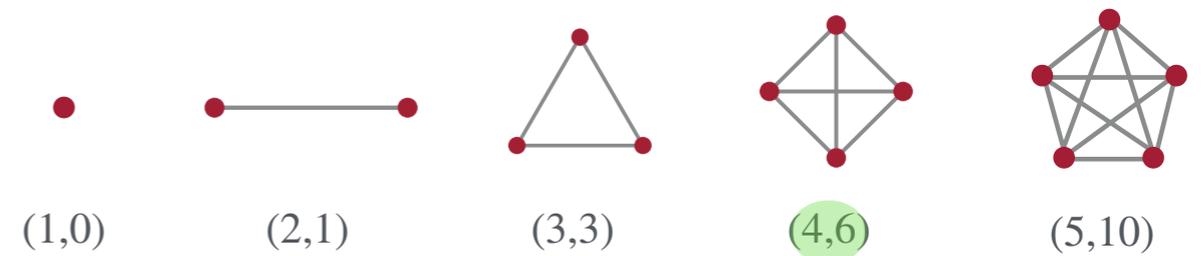
Higher $\chi \to$ more particles required $\to$ probes increasingly complex multi-particle angular correlations. $\chi = 2$: pairwise. $\chi = 3$: triangular/planar. $\chi = 4$: 3D structure.

24

BA, Lee, Thaler, *In preparation*

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 — requires *4 non-coplanar* particles, correction poorer for our extreme α_s variations

Lowest order $\chi = 4$ polynomial first arises at $d = 6$
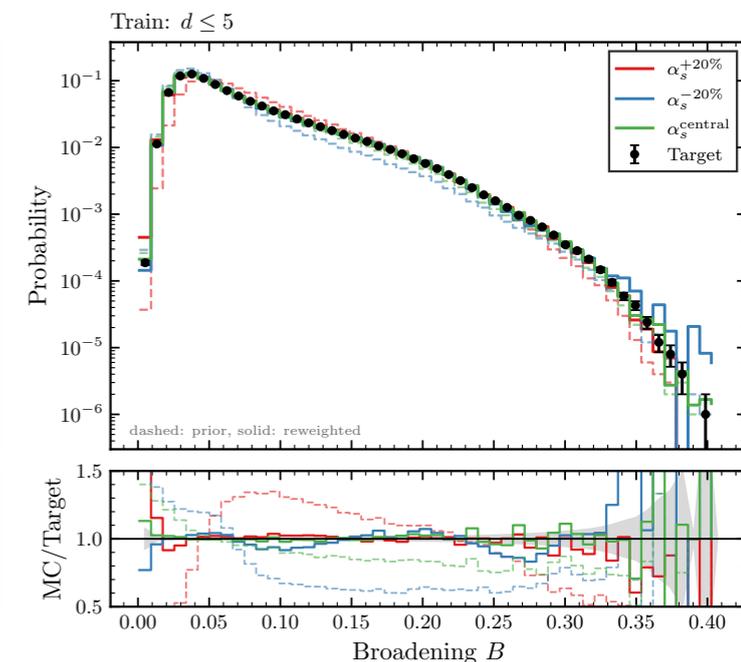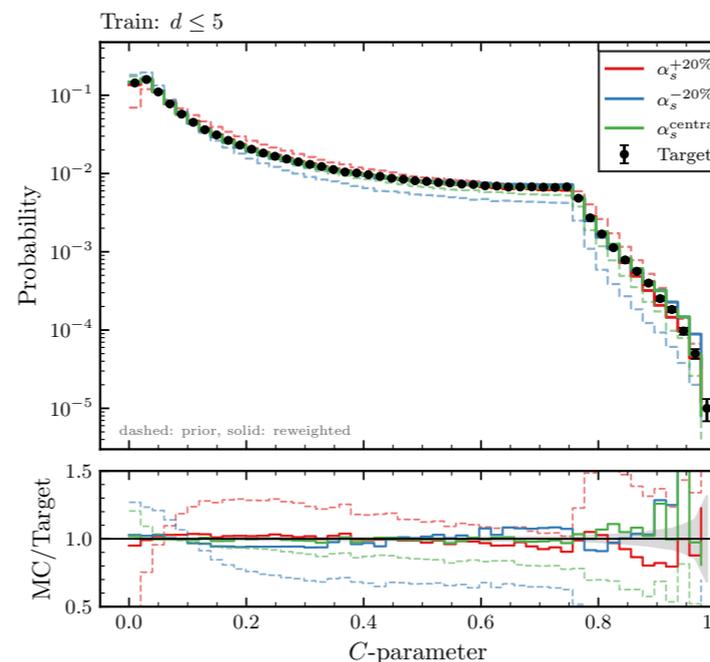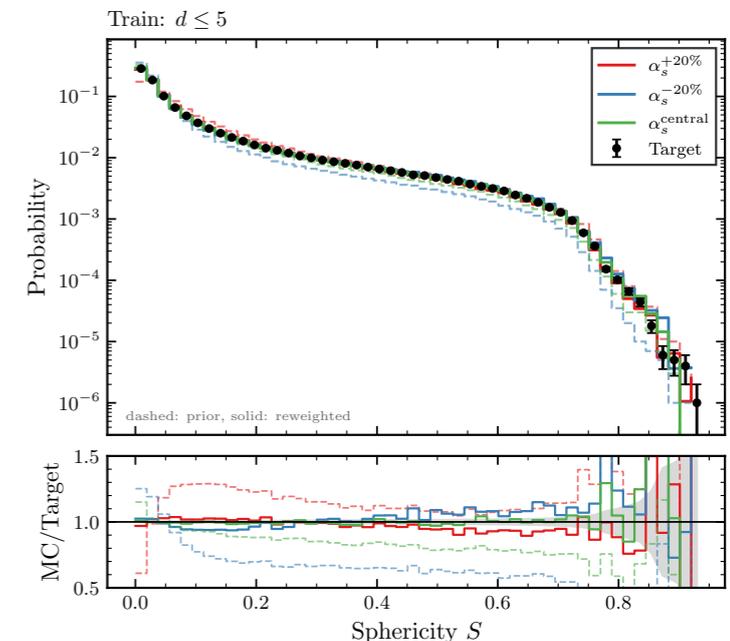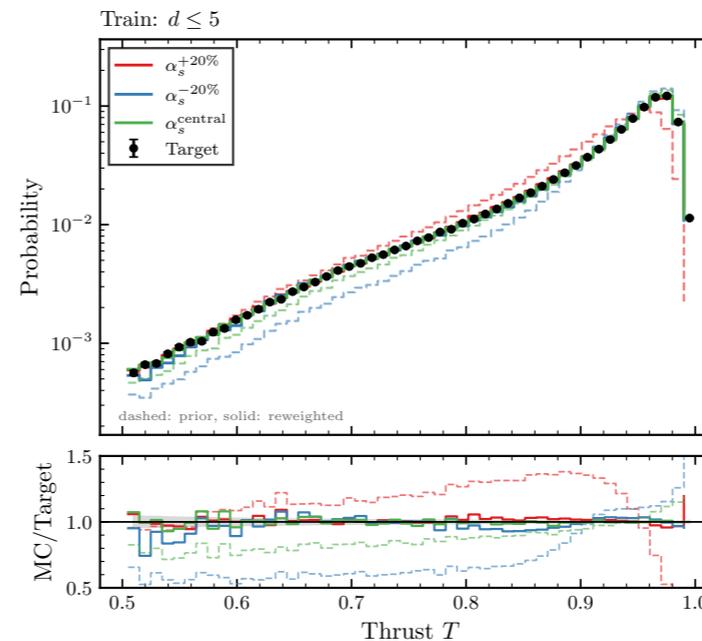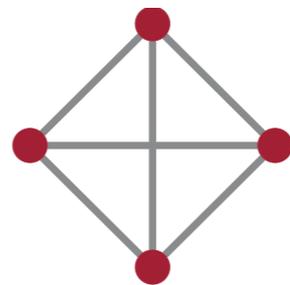
Information theory for efficient optimal observable selection

HDSense: efficient observable ranking

Hadronization models have many parameters θ tuned by correlated observables $\mathcal{O} \Rightarrow$ 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 $I^{(i)}$

Sum of individual constraining powers in traces if observables are uncorrelated

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

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

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

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

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

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
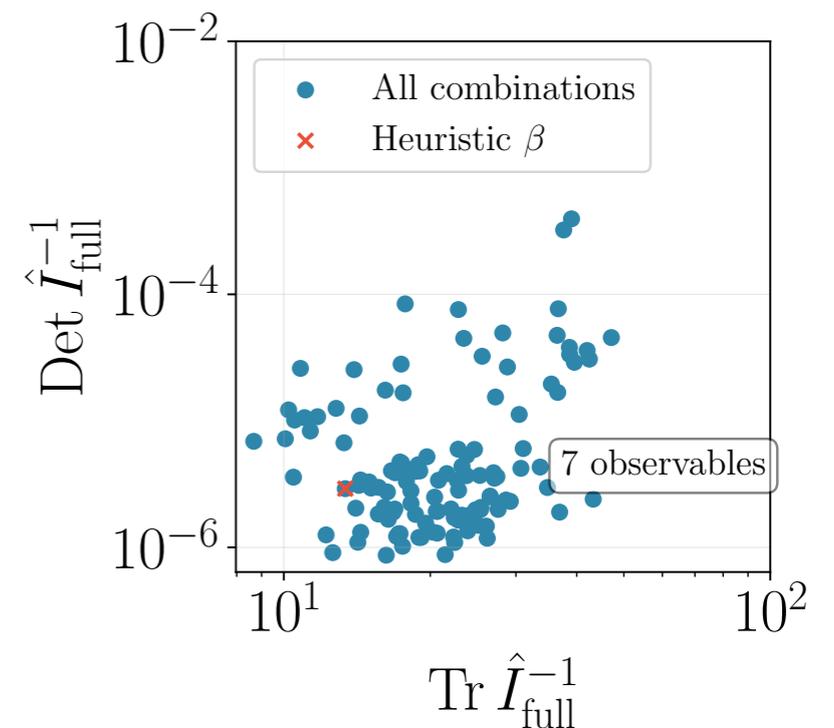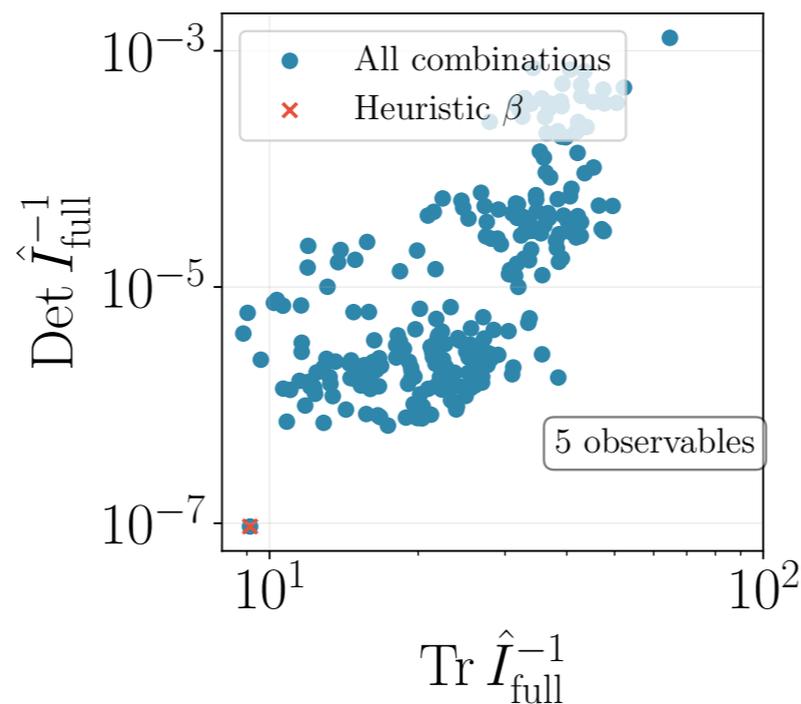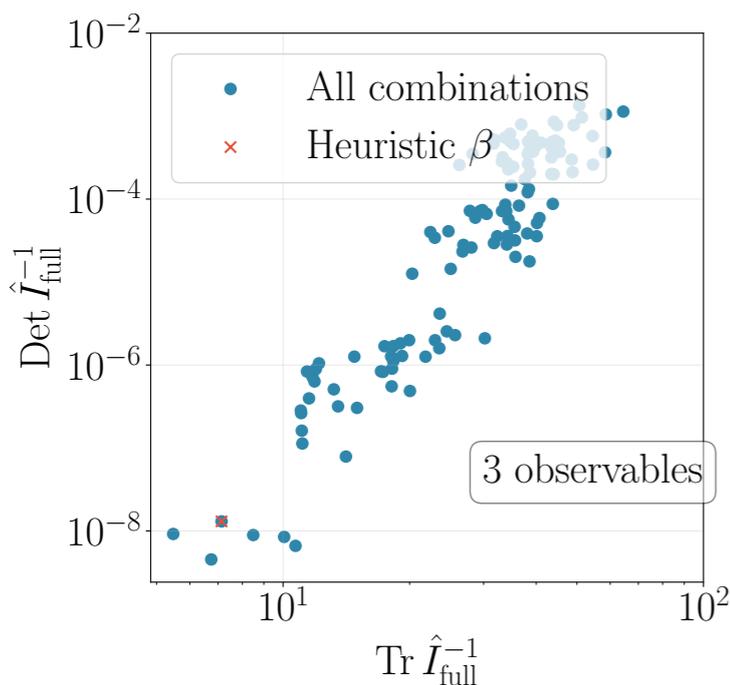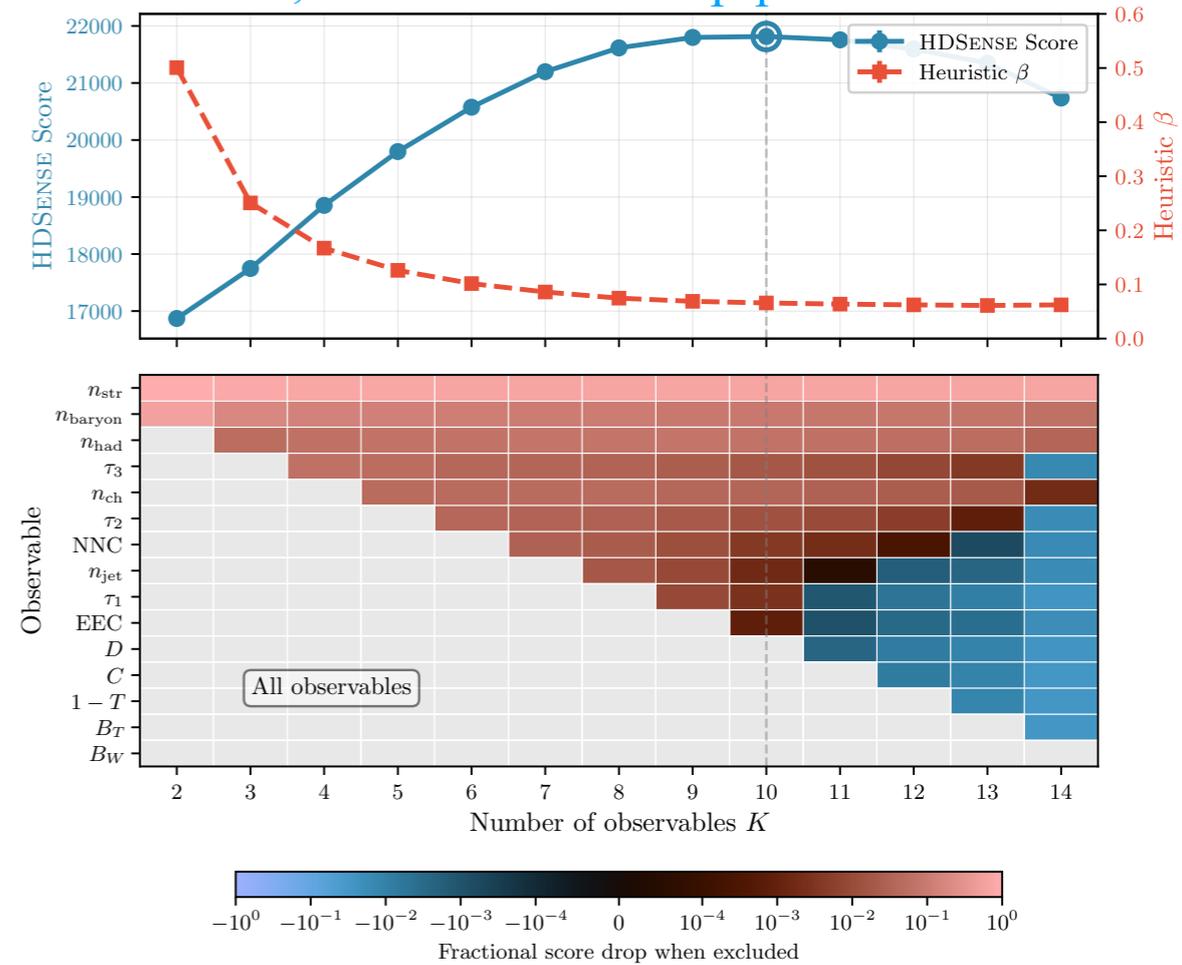
Final Summary

- 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

- Information saturation with linear basis of IRC-safe observables (EFPs): found rapid saturation at lower degrees — moments fix dramatically broken shower

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

Thank you

Back-up

What next?

Drell-Yan upgrade ATLAS MC production sample to $N^3LO + N^4LL$

+ *ab-initio LQCD* via double-diff. q_T and $\Delta\phi$ moment constraints

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

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

BA, Poncelet, Whitehead. *In preparation*

Applying IT-reweighting beyond LEP/LHC $\rightarrow$ E.g. Neutrino generators

Open questions: Systematic uncertainty propagation, optimal moment basis selection, more targeted optimization algorithms
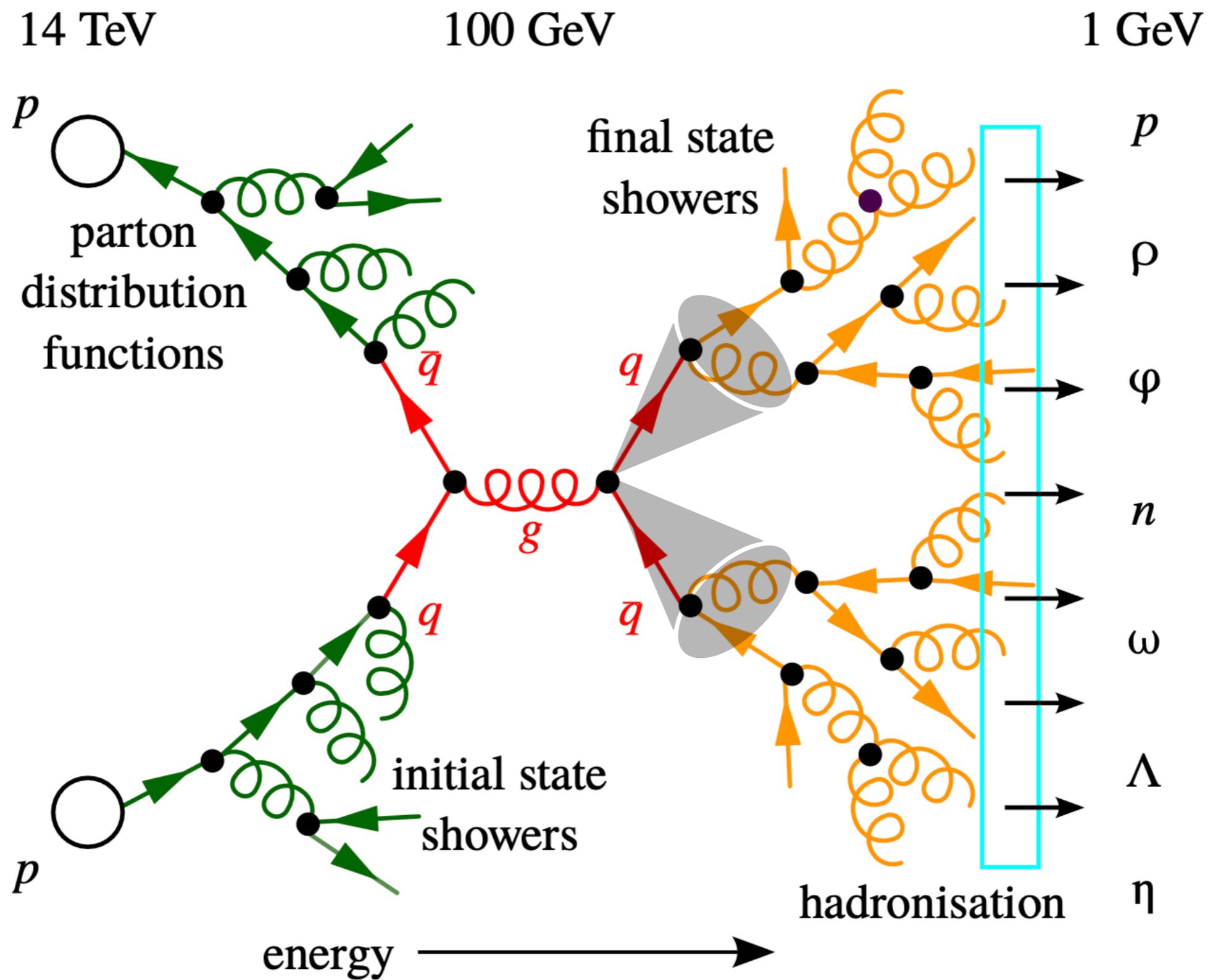
Summary and Higher Orders

$\mathrm{NLL} + \mathrm{NLO}$ matching in both massive and massless case $\Rightarrow$ now **implemented as default** in SHERPA

Towards $\mathrm{MC@NNLO}$ **fully integrated** automated generator - working on **a *shower-aware* subtraction scheme** that matches ALARIC construction + *massive*
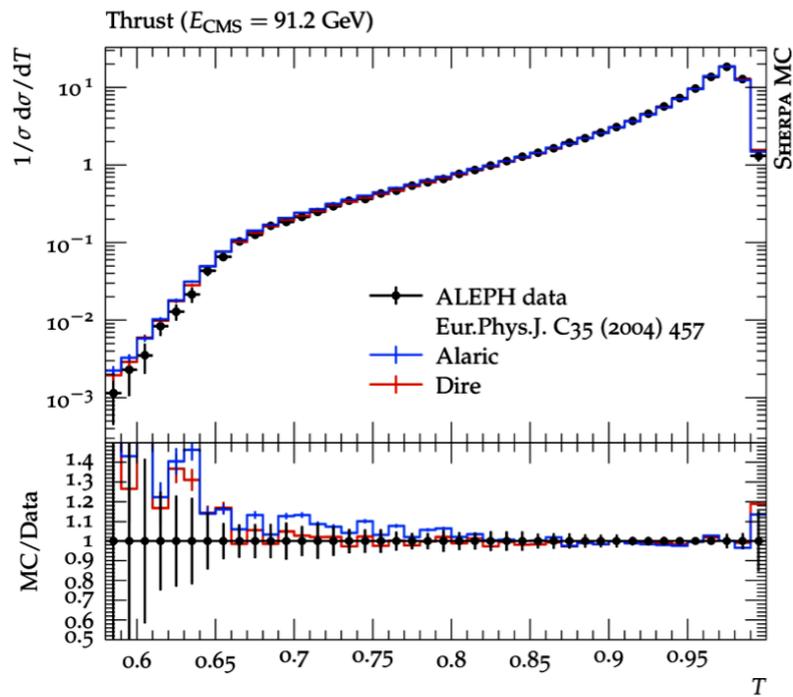
BA, Knobbe, Höche, et al.

Many **more improvements** to ALARIC: spin correlations, beyond leading color, higher multiplicities, complete NNLL, etc.

Factorizing QCD

Numerical test and data

Thrust ($E_{CMS} = 91.2$ GeV)

- ◆ ALEPH data
  Eur.Phys.J. C35 (2004) 457
- Alaric
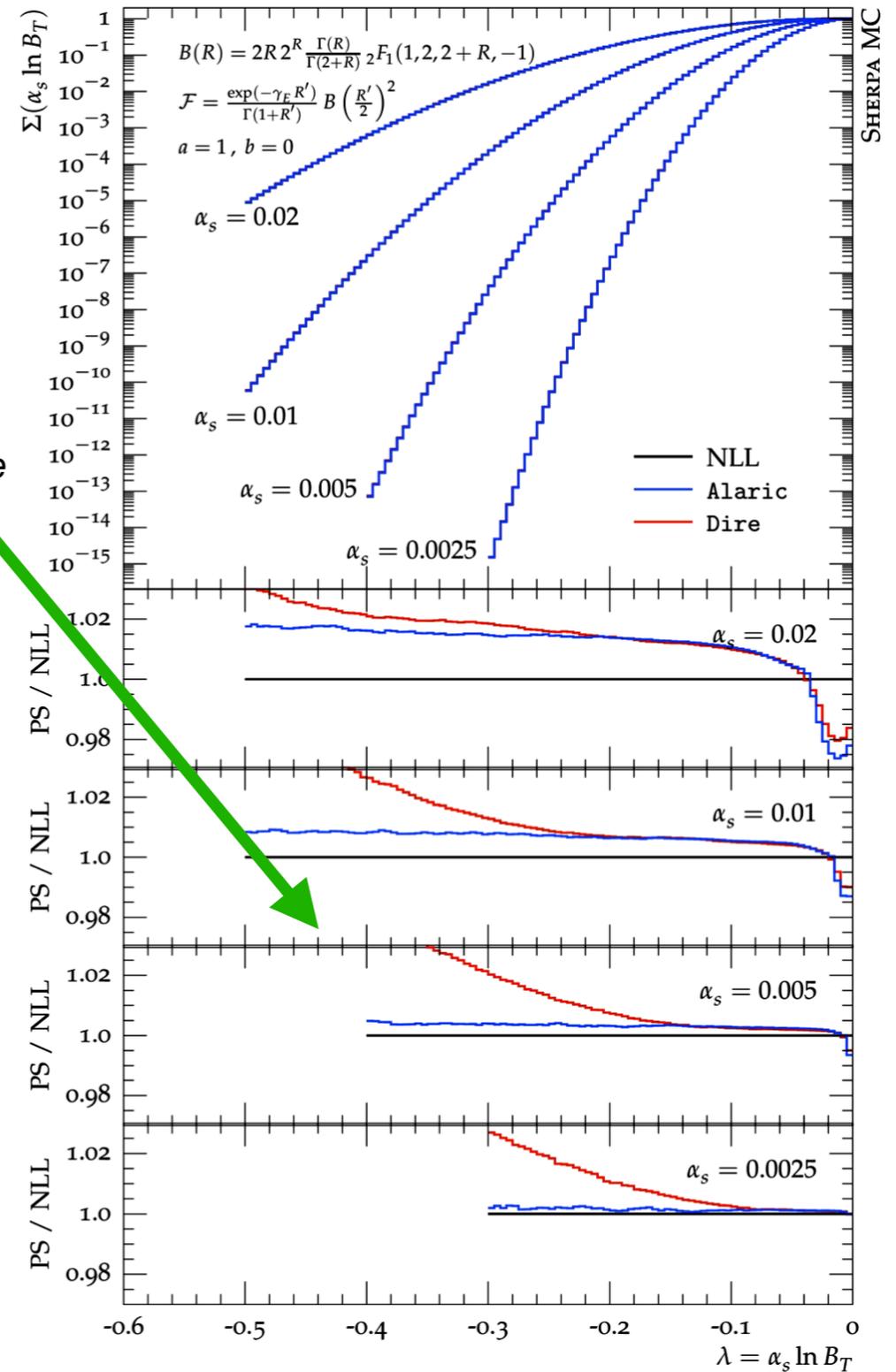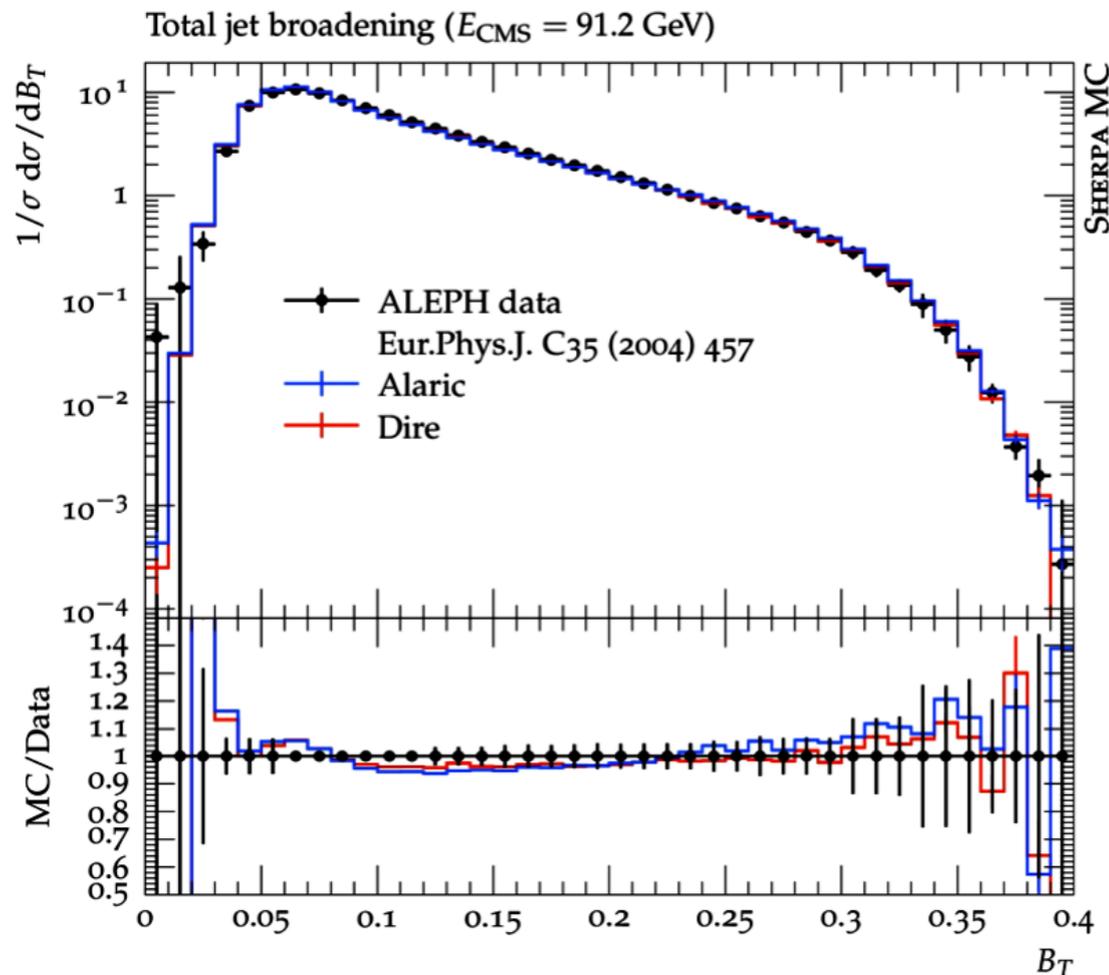- Dire

*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*

$$B(R) = 2R\, 2^R \frac{\Gamma(R)}{\Gamma(2+R)}\, {}_2F_1(1,2,2+R,-1)$$
$$\mathcal{F} = \frac{\exp(-\gamma_E R')}{\Gamma(1+R')}\, B\left(\frac{R'}{2}\right)^2$$
$$a = 1,\ b = 0$$

$\alpha_s = 0.02$
$\alpha_s = 0.01$
$\alpha_s = 0.005$
$\alpha_s = 0.0025$

— NLL
— Alaric
— Dire

$\alpha_s = 0.02$
$\alpha_s = 0.01$
$\alpha_s = 0.005$
$\alpha_s = 0.0025$

$\lambda = \alpha_s \ln B_T$

Total jet broadening ($E_{CMS} = 91.2$ GeV)

- ◆ ALEPH data
  Eur.Phys.J. C35 (2004) 457
- Alaric
- Dire

B_T

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

34

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)
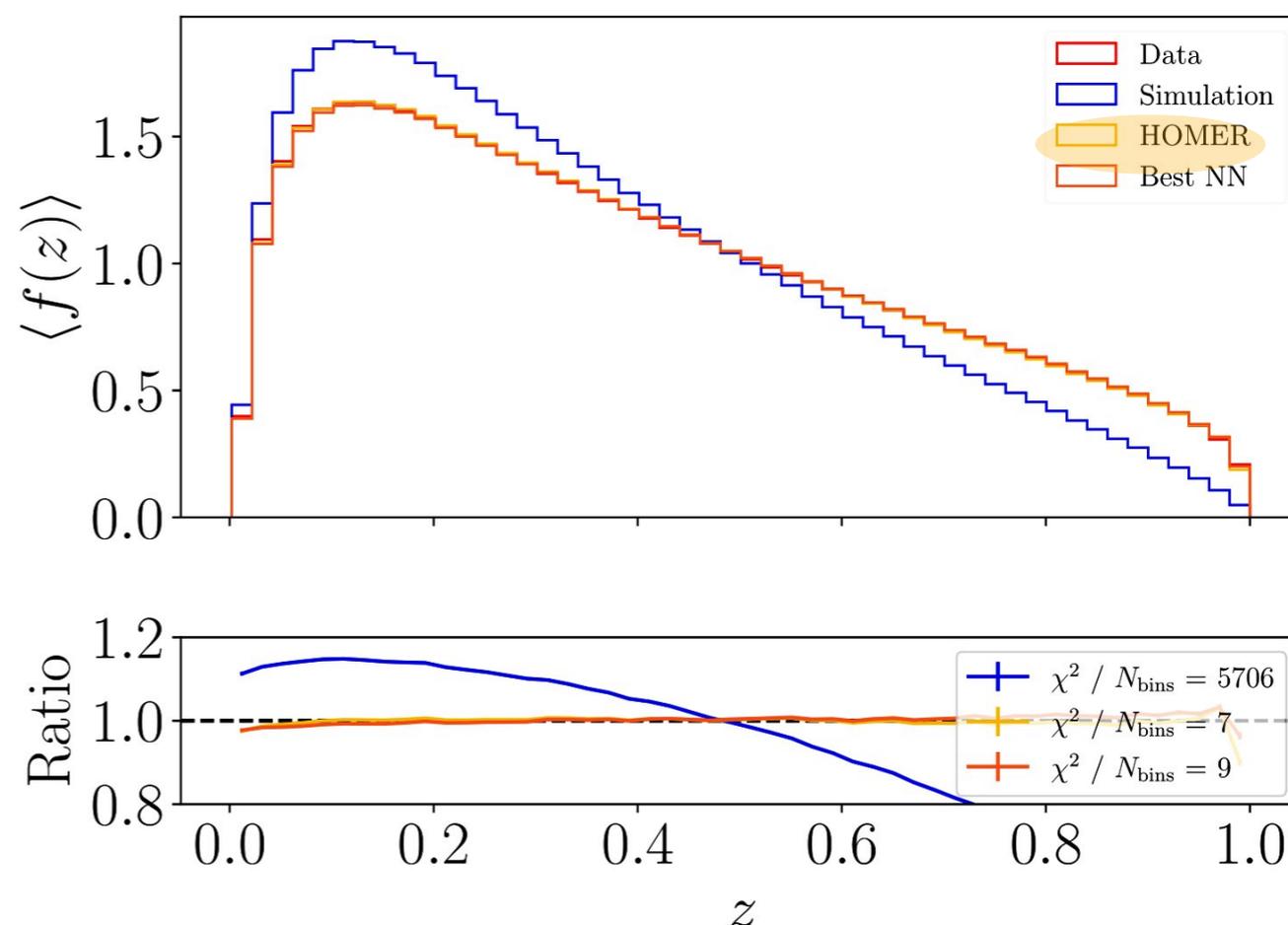


Bierlich et al. 2410.06342

BA et al. 2503.05667

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** $f(z) \rightarrow$ entire hadronization chain reproduced by **iterating**
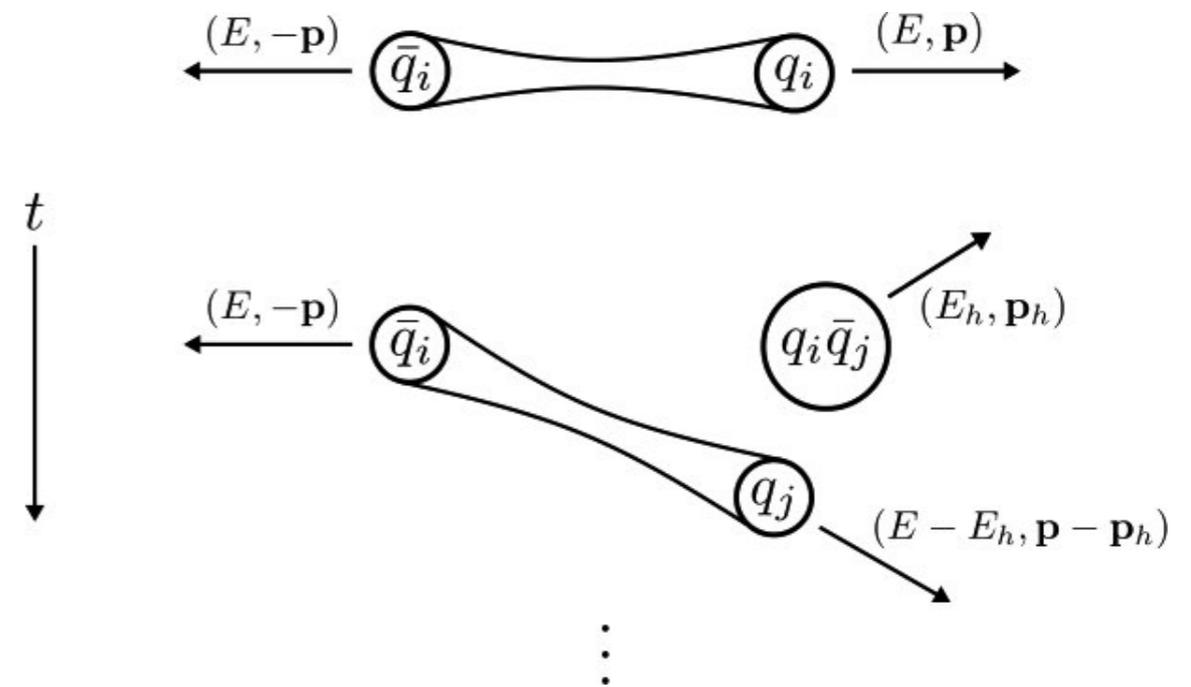
Pythia Lund String Model: Colored singlets + O(20) parameters → Hadrons

Simplified example from <u>arxiv:2203.04983</u> *.*

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
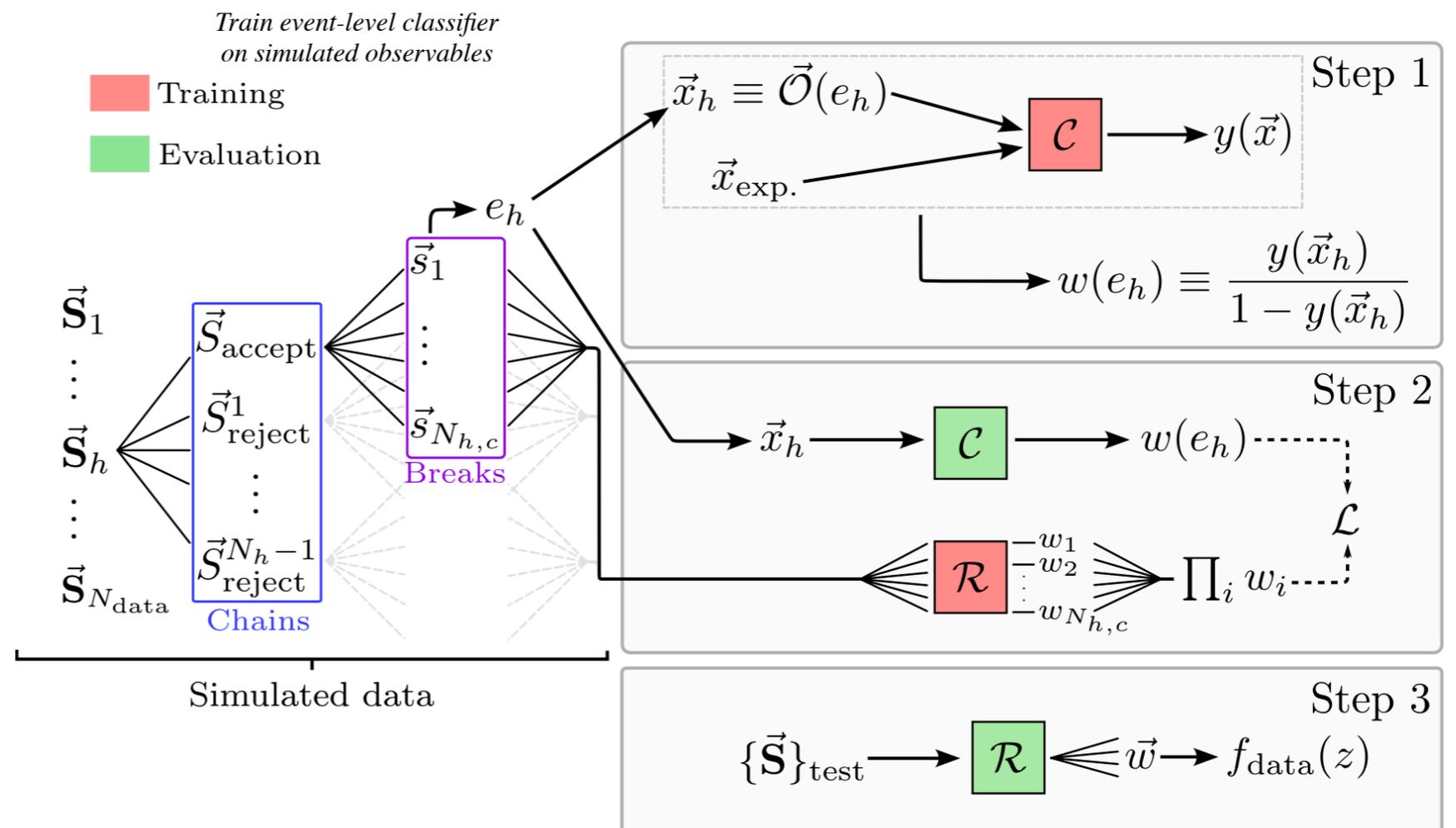
New Underlying Model

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**
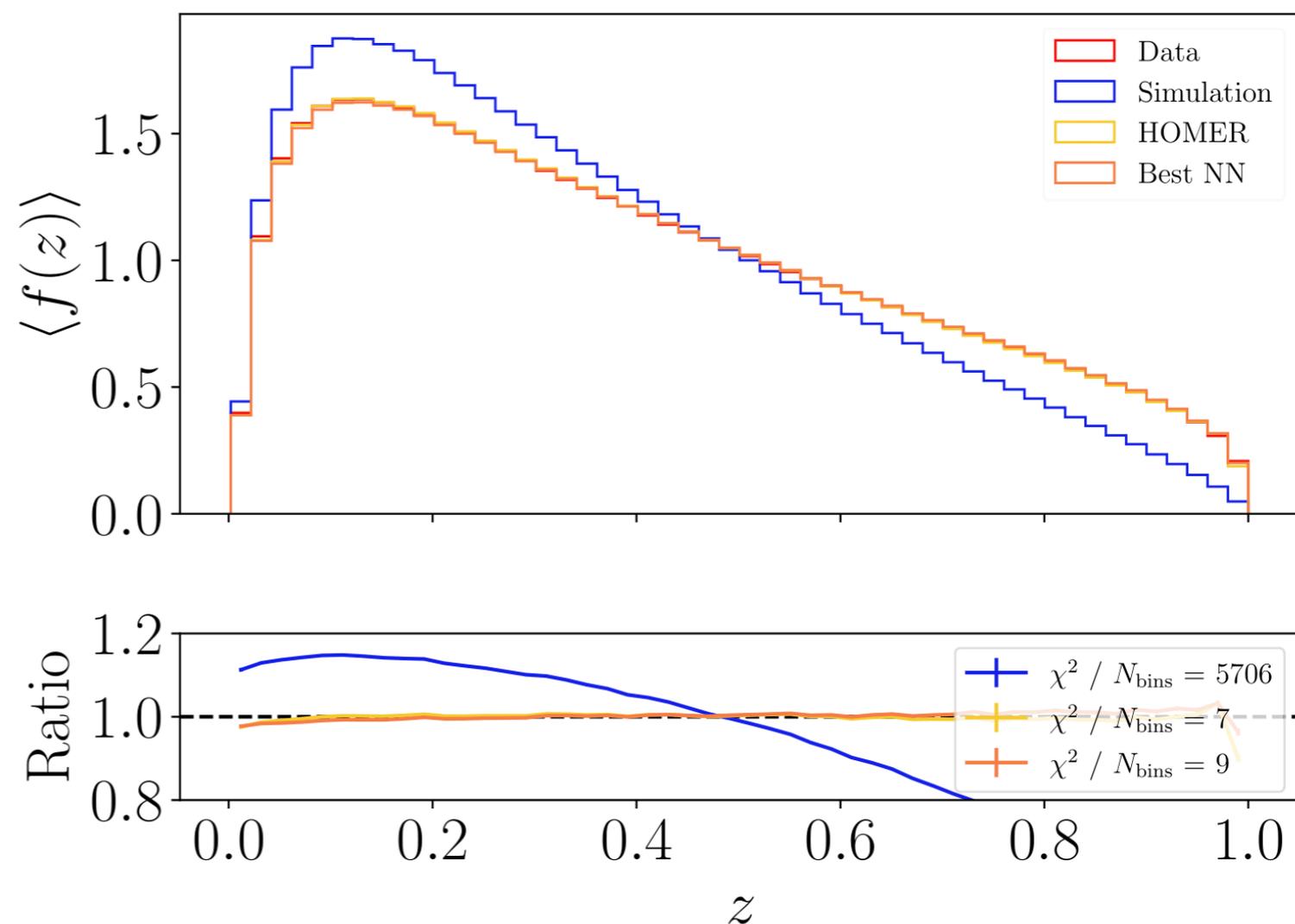
arxiv:2410.06342
arxiv:2503.05667

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
   $(g_1\bar{q}) \to (\bar{q}g_1) + (g_1g_2)$

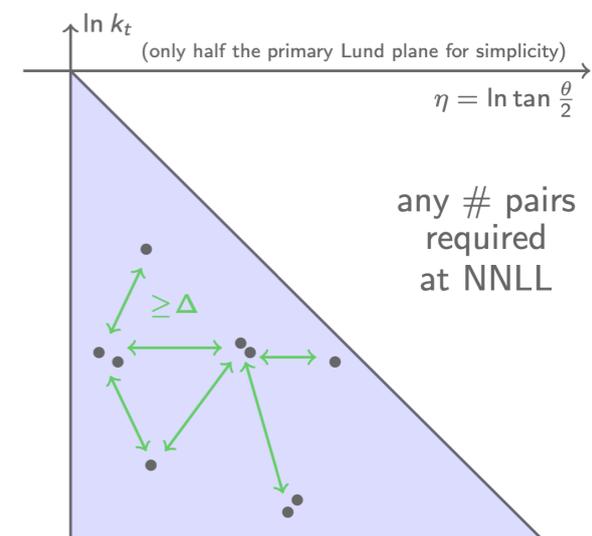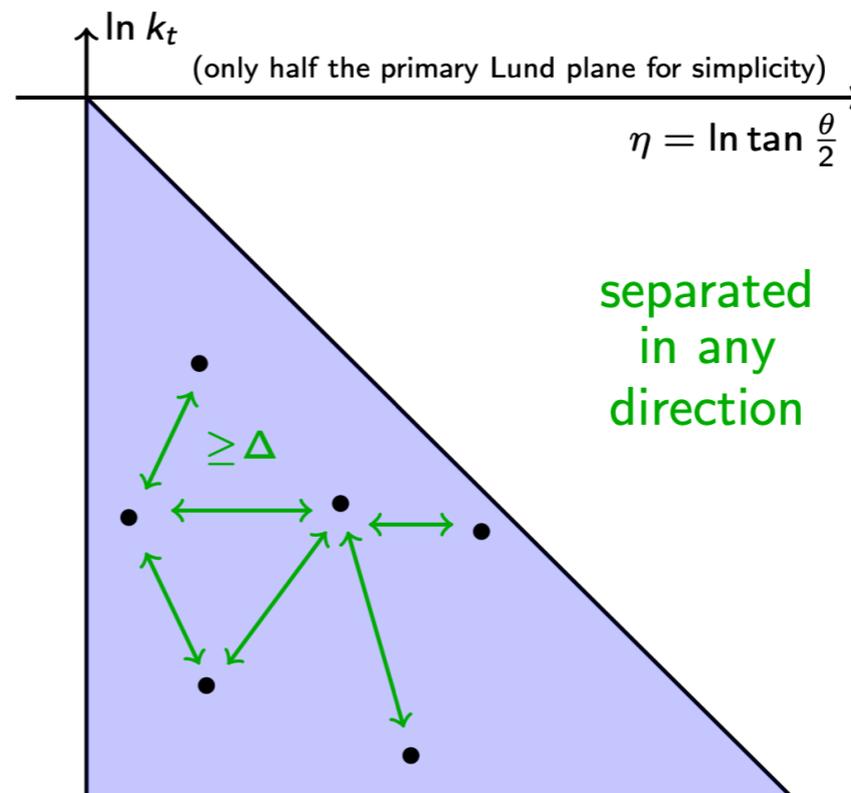5) Iterate until $k_t = k_{t,\text{cut}}$
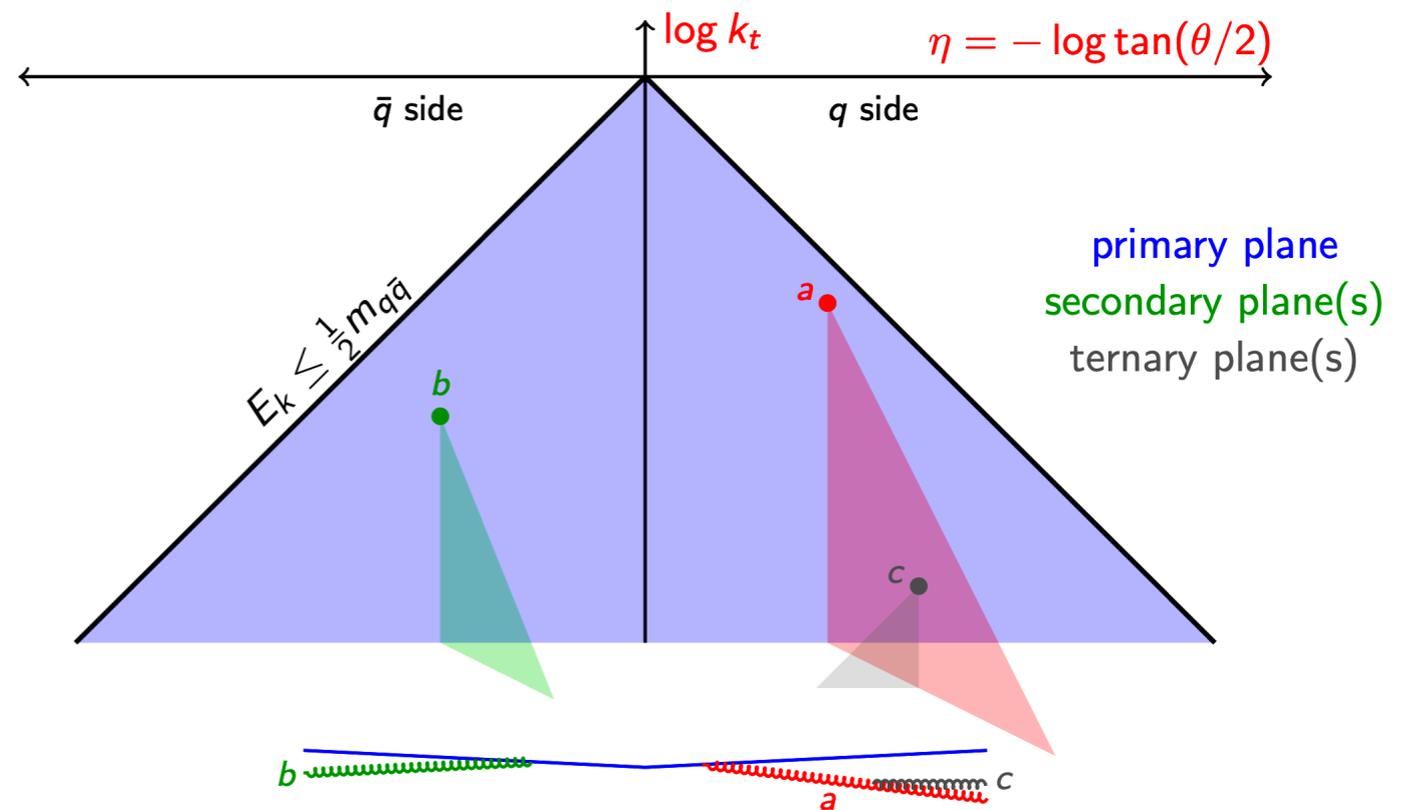
Evolution **resums** large logs at what accuracy?

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

$\log k_t$ $\eta = -\log \tan(\theta/2)$

$\bar{q}$ side q side

primary plane
secondary plane(s)
ternary plane(s)

$E_k \leq \frac{1}{2} m_{q\bar{q}}$

$\ln k_t$ (only half the primary Lund plane for simplicity)

$\eta = \ln \tan \frac{\theta}{2}$

separated in any direction

$\geq \Delta$

$\ln k_t$ (only half the primary Lund plane for simplicity)

$\eta = \ln \tan \frac{\theta}{2}$

any # pairs required at NNLL

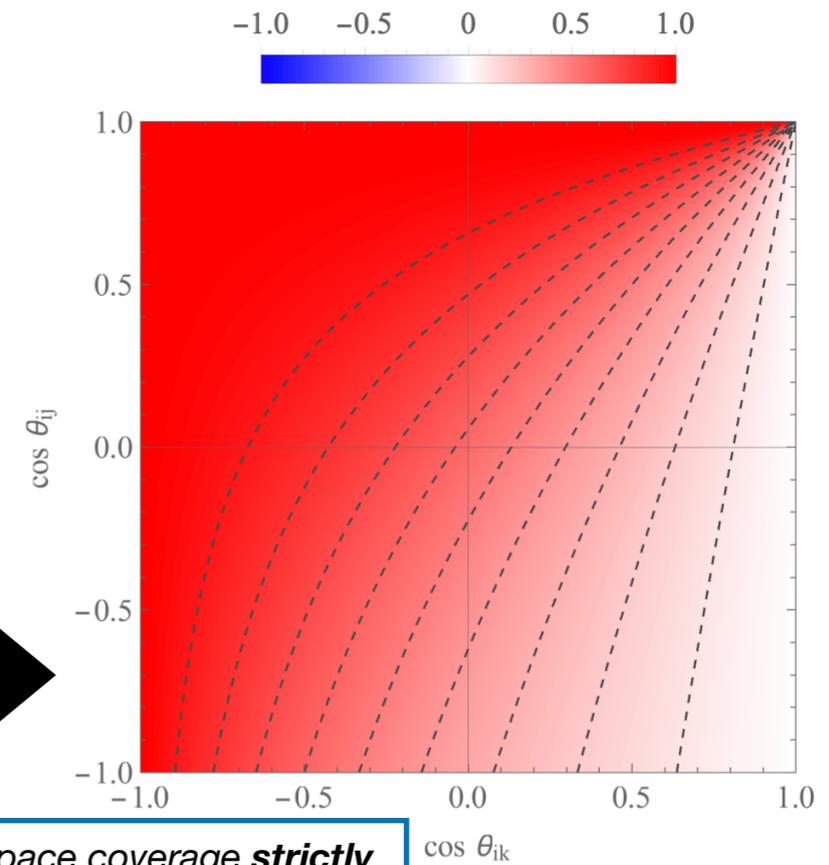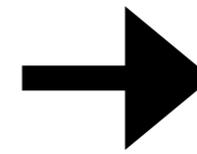$\geq \Delta$

40

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**

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

*Full phase-space coverage **strictly positive** maintains PD interpretation*
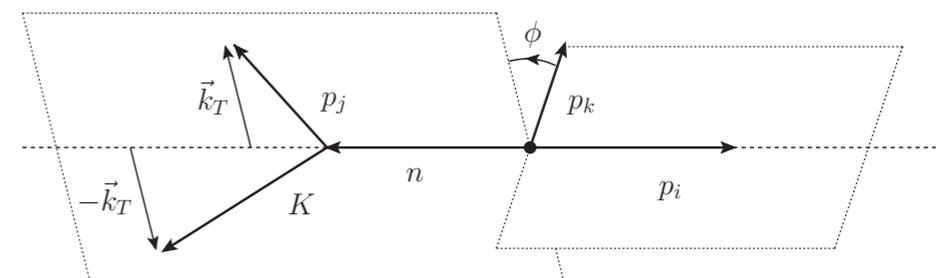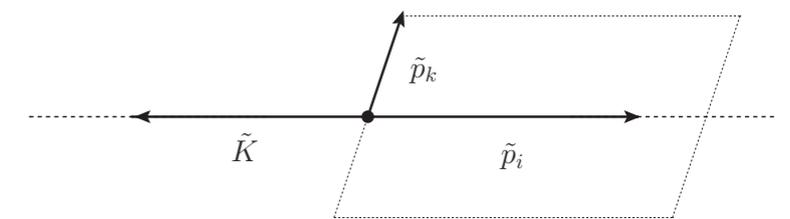
Ingredient 2: Momentum mapping

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

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

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

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

Massless limit only one simpler integral [arXiv:2208.06057]

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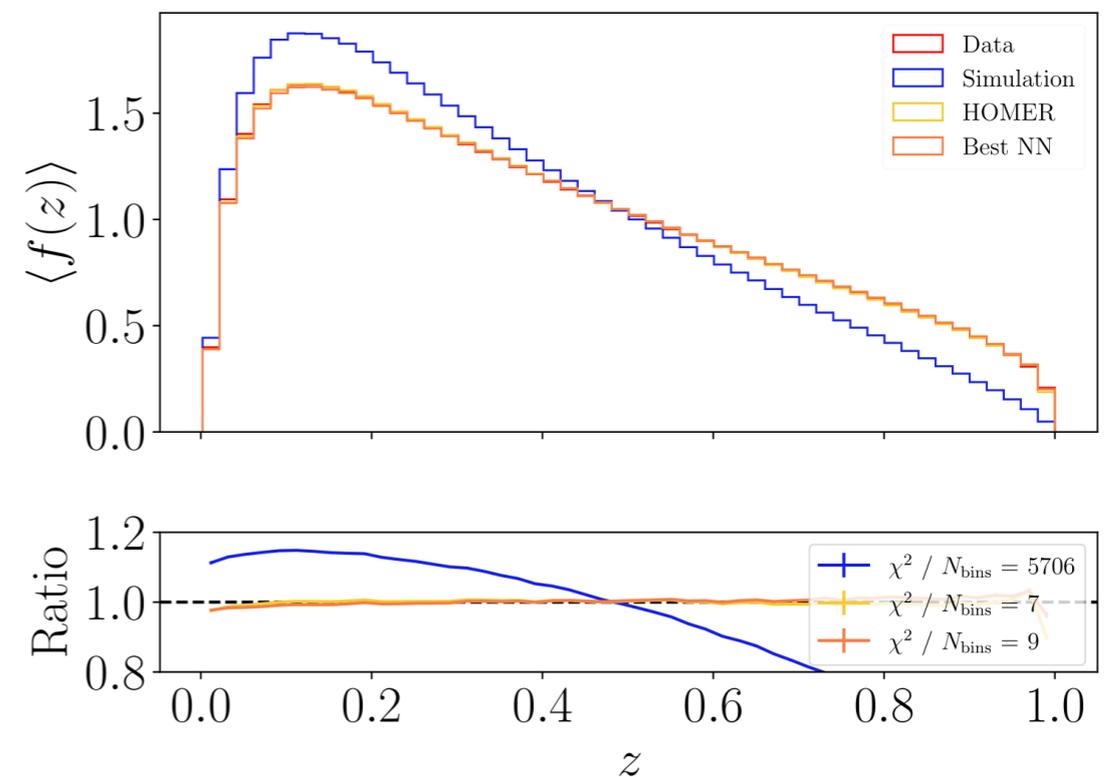
Recap

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   Tackmann: 2411.18606
                Non-perturbative corrections   Lim, Poncelet: 2412.14910
                Treatment of mixed-moments   …
                Convergence of moment expansion

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
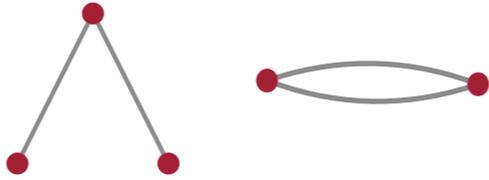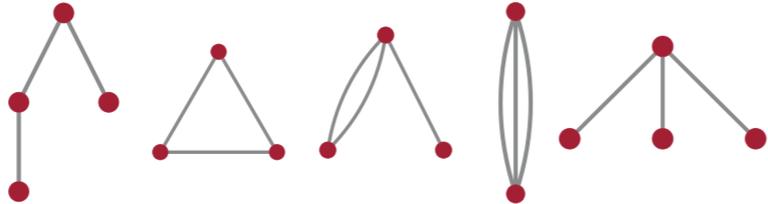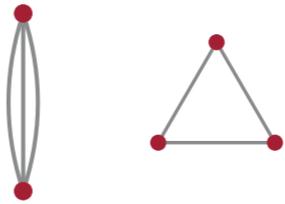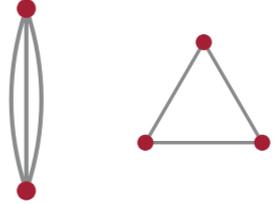
Double-differential moments

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

Precision targets: $\mathrm{NNLO} +$
$\mathrm{N^4LL}$ 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
$d = 1$			
$d = 2$			
$d = 3$			
$d = 4$			
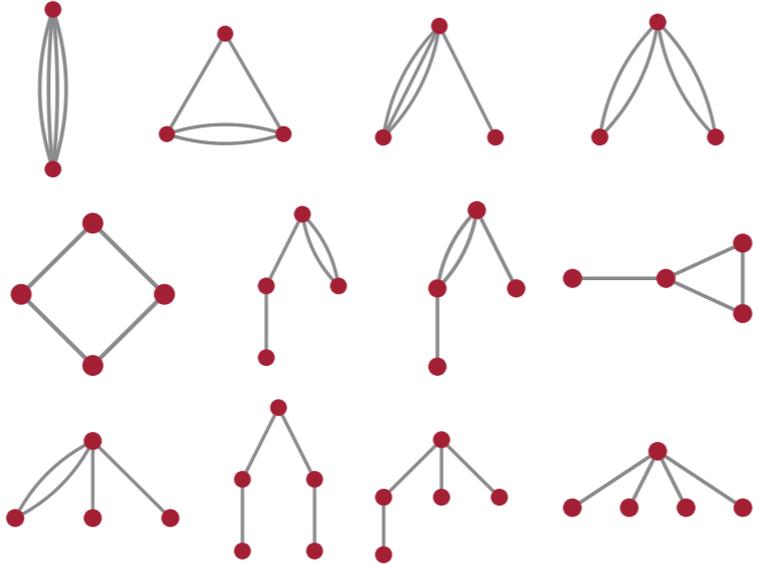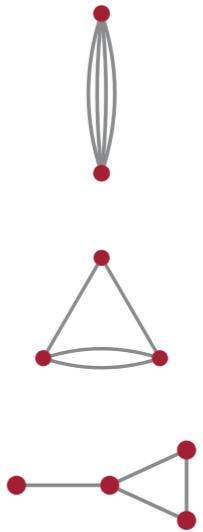
Learning the Lund string FF

First instinct: Learn the fragmentation PDF from data

$$p_{\mathrm{Lund}}(x) \rightarrow p_{\mathrm{ML}}(x)$$

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

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

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

We are using simulation to introduce **inductive bias** $\rightarrow$
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

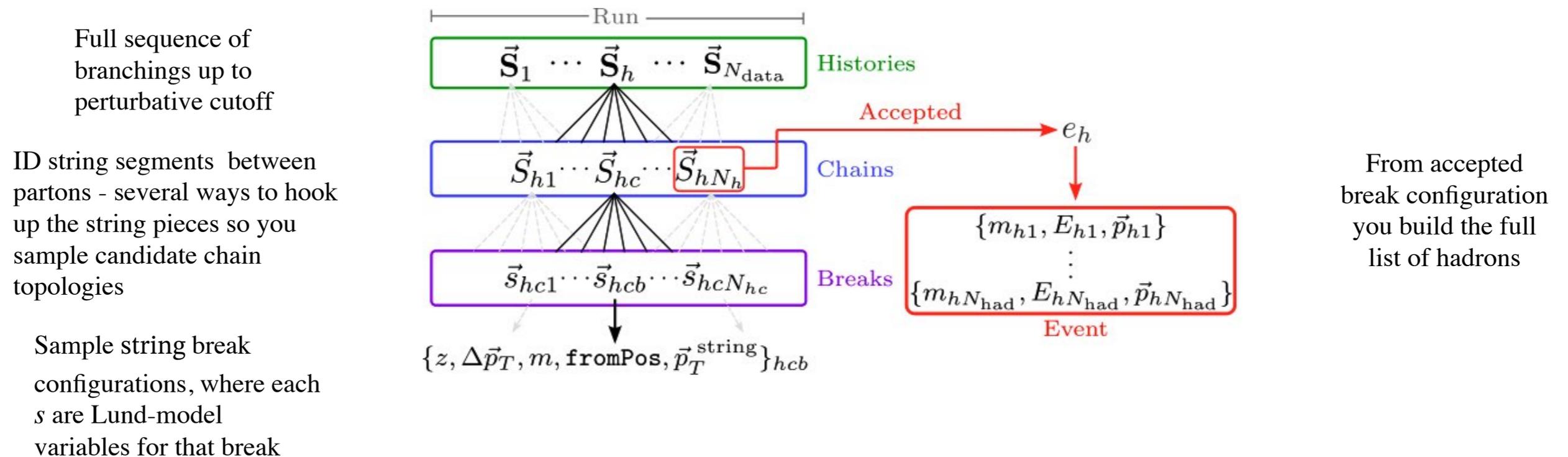We consider three possibilities for available measurements:

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
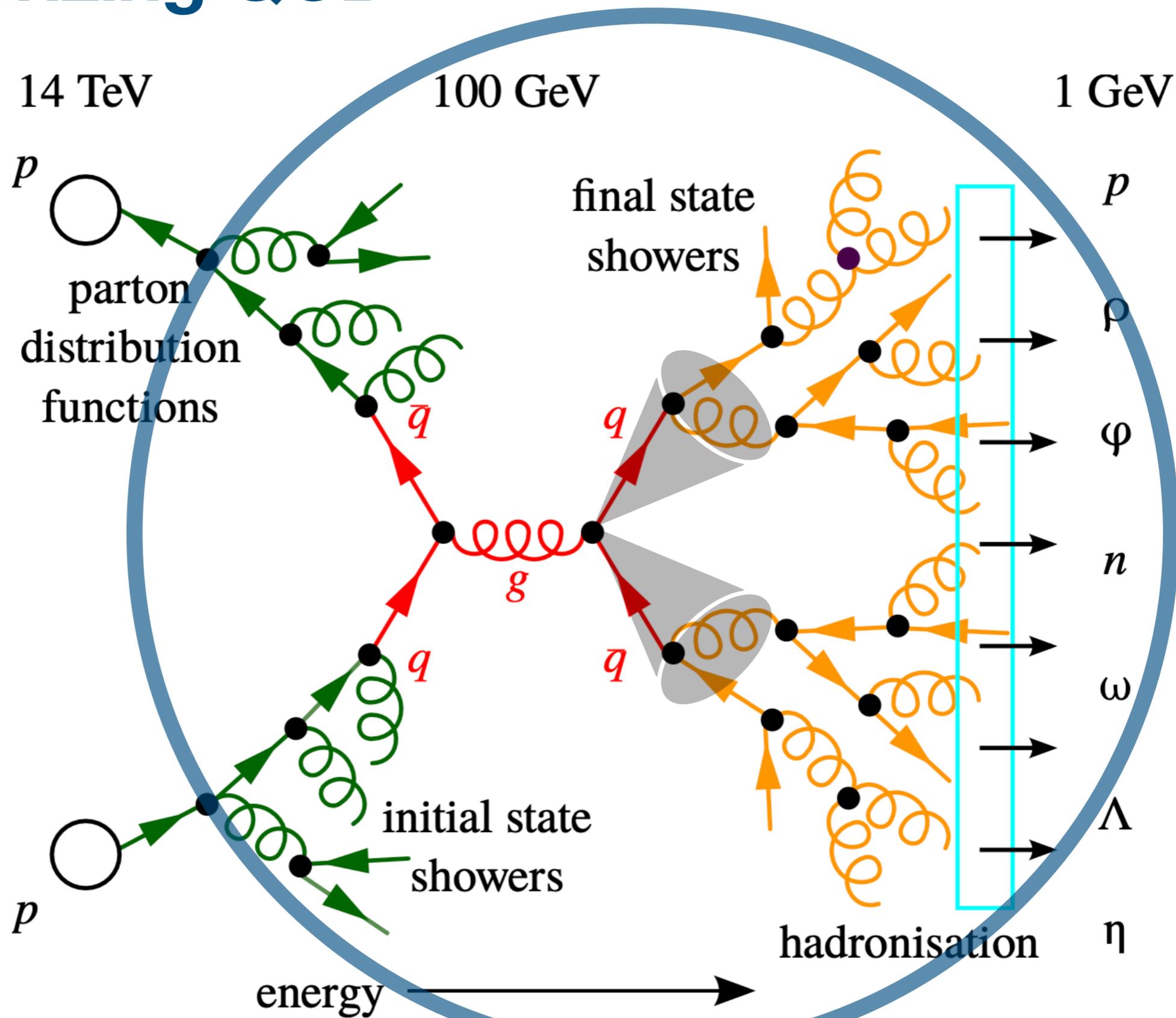
Learning from data: HOMER

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:

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

with $k^\mu = z_q \tilde{p}_q^\mu + z_{\bar{q}} \tilde{p}_{\bar{q}}^\mu + k_t^\mu$

Degrees of freedom:

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

2) Transverse momentum: k_t
3) Azimuth: ϕ

$\log k_t$

$\eta = -\log \tan(\theta/2)$

$\bar{q}$ side q side

Lund plane

$E_k \leq \frac{1}{2} m_{q\bar{q}}$

$\bullet k$

k

q $\bar{q}$