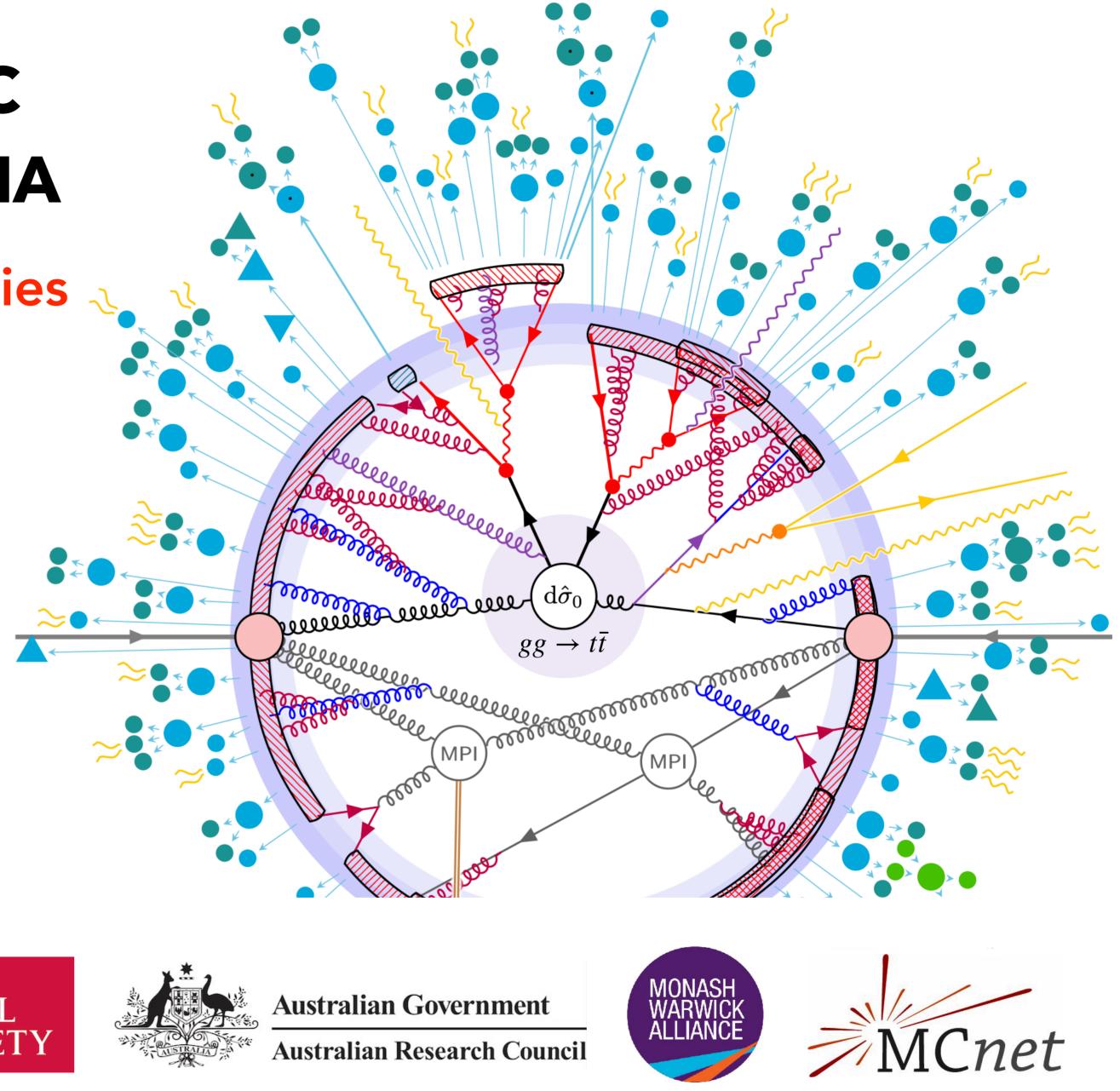
Modelling of LHC Collisions in PYTHIA

— Physics & Uncertainties

Peter Z Skands

University of Oxford & Monash University

Oxford, April 2024





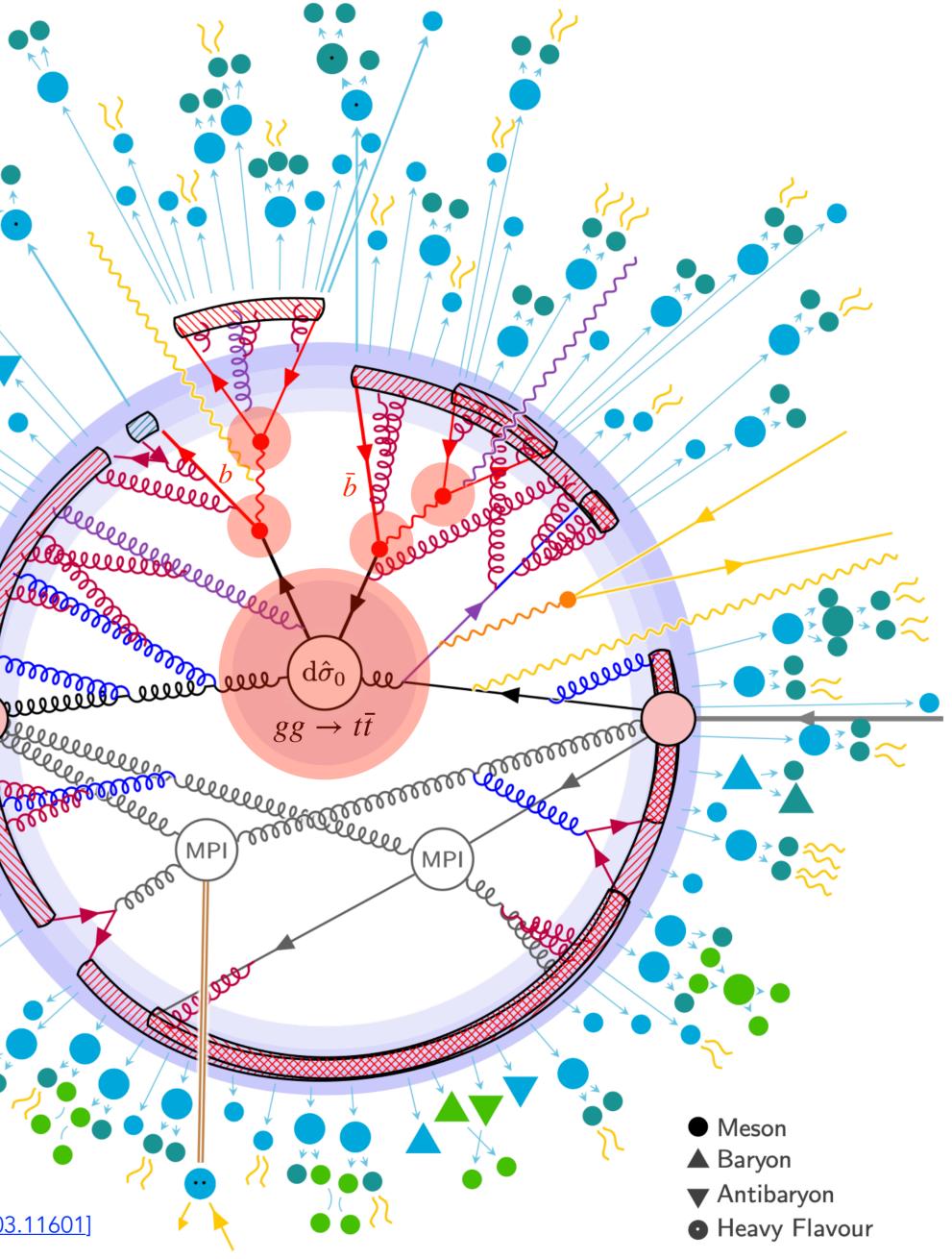




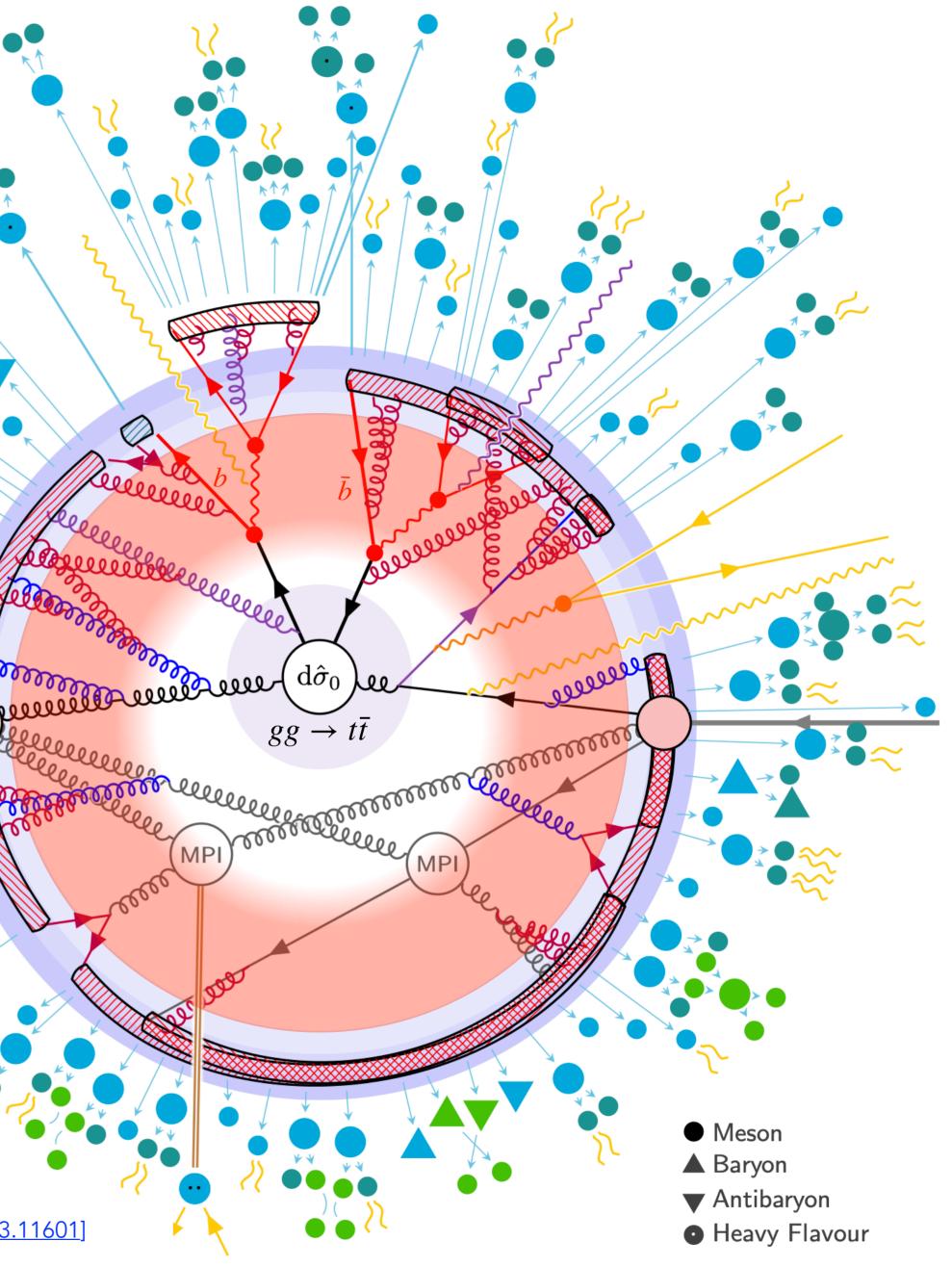




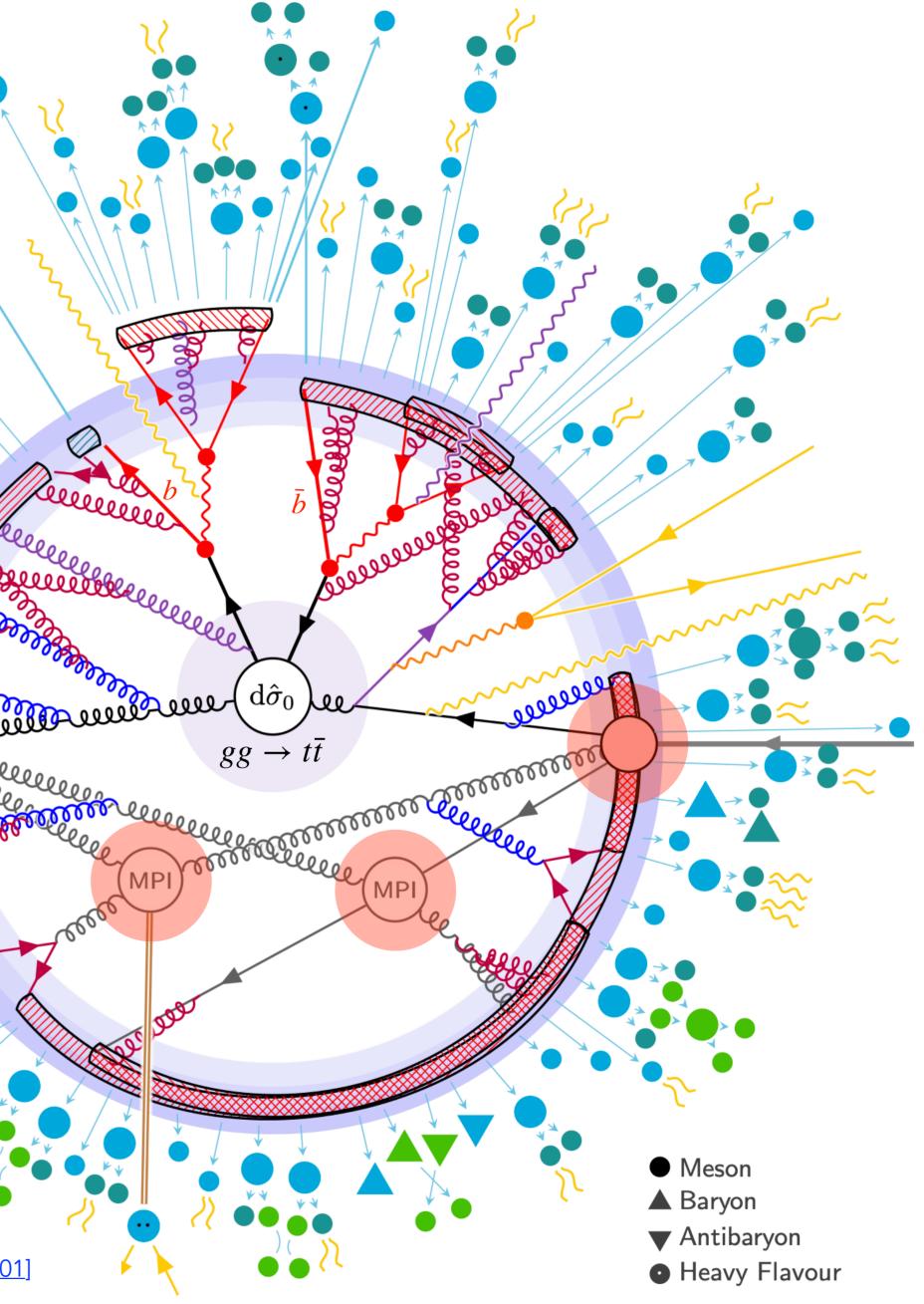
- Hard Process
- O Hard Interaction
- Resonance Decays
- MECs, Matching & Merging



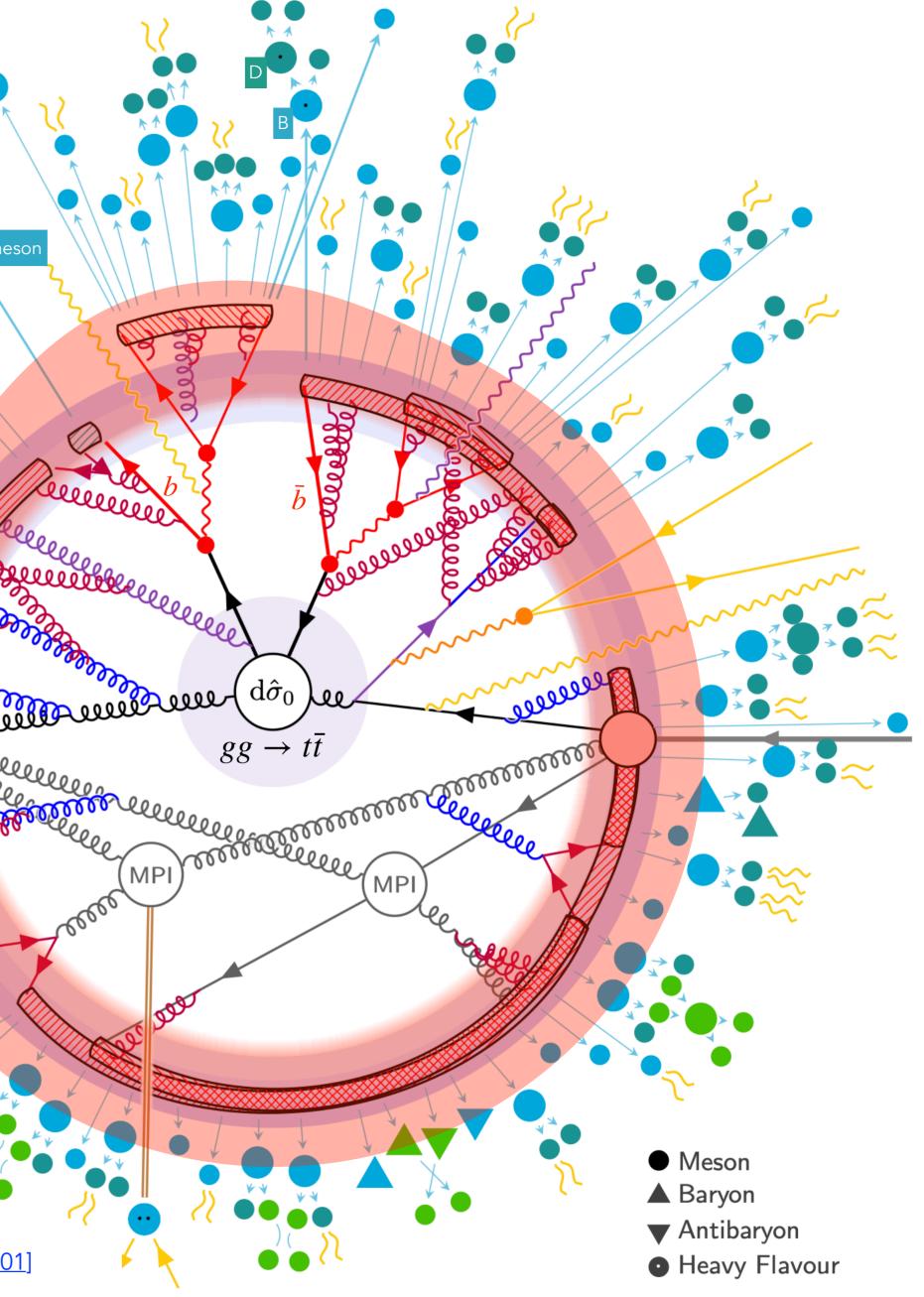
Hard Process	 Hard Interaction Resonance Decays MECs, Matching & Merging 	
Parton Showers	 QCD Final-State Radiation QCD Initial-State Radiation* Electroweak Radiation 	



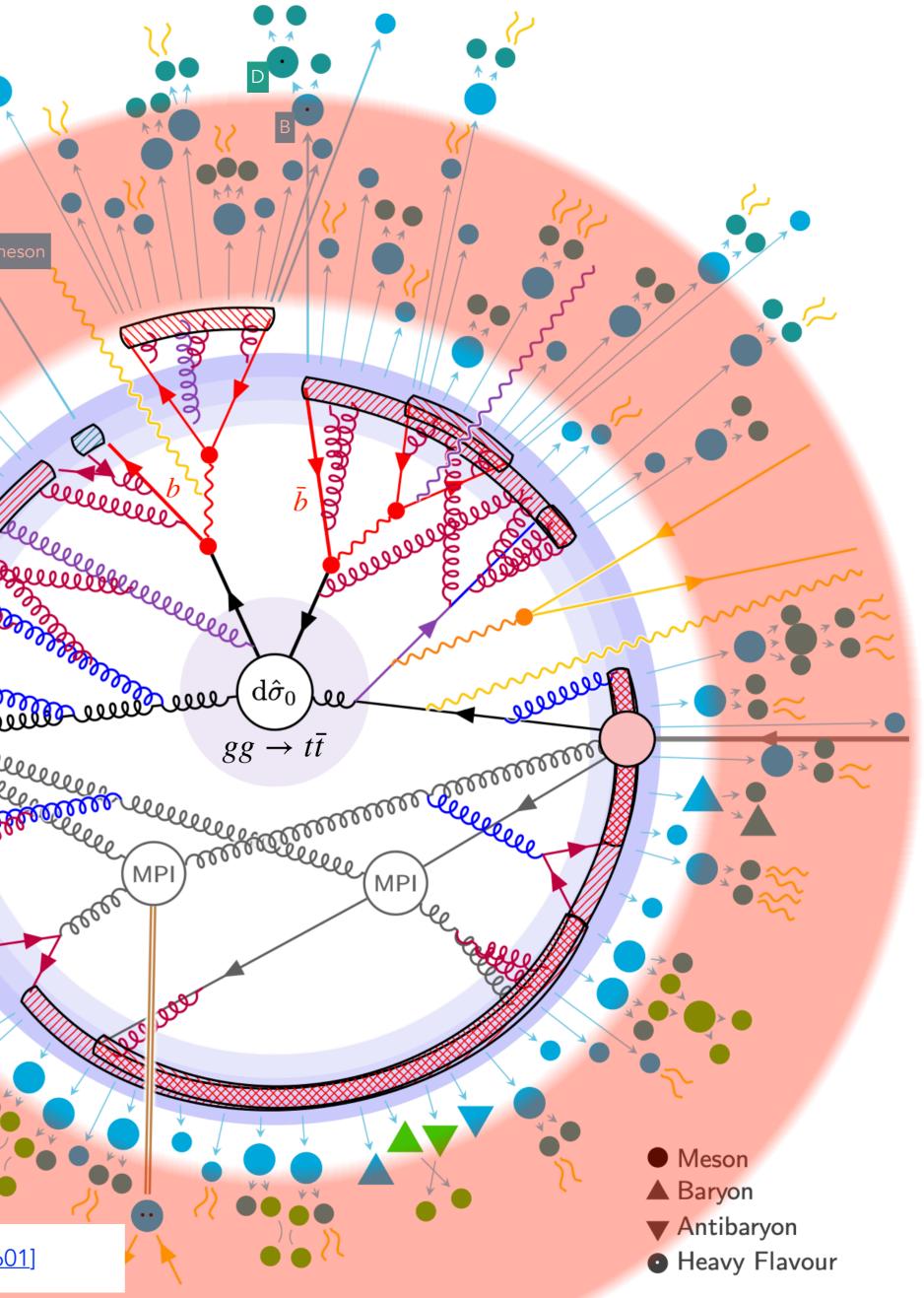
Hard Process	 Hard Interaction Resonance Decays MECs, Matching & Merging 	
Parton Showers	 QCD Final-State Radiation QCD Initial-State Radiation* Electroweak Radiation 	
Underlying Event	 Multiparton Interactions Beam Remnants* 	



Hard Process	O Hard Interaction	D meson	
	Resonance Decays		B mes
	MECs, Matching & Merging		
	<u></u>		
Parton Showers	QCD Final-State Radiation		
	QCD Initial-State Radiation*		
	Electroweak Radiation		
			peop
Underlying	Multiparton Interactions		elee
Event	Beam Remnants*		
	-		00000
Hadronization	🔯 Strings		feee
	Colour Reconnections		REFE
	String Interactions		2000
	Bose-Einstein & Fermi-Dirac		



Hard Process	 Hard Interaction Resonance Decays 	Br				
	MECs, Matching & Merging					
Parton Showers	QCD Final-State Radiation					
	QCD Initial-State Radiation*					
	Electroweak Radiation					
		1/h				
Underlying Event	O Multiparton Interactions	fee				
	Beam Remnants*	1				
		2000				
	Strings	See.				
	Colour Reconnections	1980				
Hadronization	String Interactions	fee.				
	Bose-Einstein & Fermi-Dirac					
	Primary Hadrons					
Hadron (& т) Decays	Secondary Hadrons					
	Hadronic Reinteractions	رہے				



Recent Studies

Focus on SM precision environments \leftrightarrow BSM backgrounds

- 1. NLO Matching Systematics with POWHEG-Box (examples: VBF, tt) 2. From NLO to NNLO (examples: tī, V, H, VH, VV, ...)
- 3. The computational bottleneck in **ME merging** (example: V+jets)
- 4. New Discoveries in Hadronization (examples: HF baryons, JES)

NB: want to address/explain state of the art & systematics in real contexts \rightarrow a bit theory heavy

1. NLO + Shower with POWHEG

(Just focusing on the real-radiation part)

POWHEG generates the first (hardest) emission with $|M_{X+1}|^2$

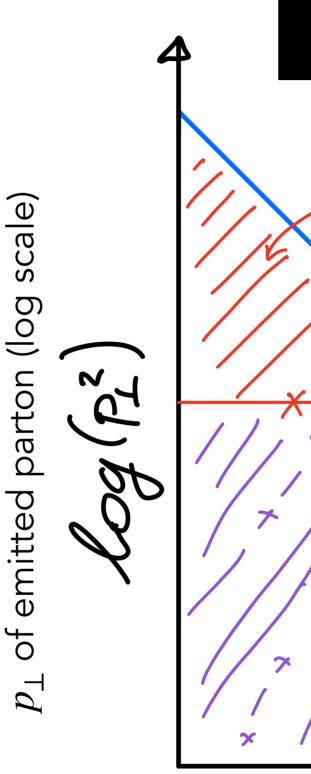
It does so in a shower-like manner: sweeping over phase space, from high to low p_T

Matrix-Element Corrections (MECs) [Bengtsson & Sjöstrand 1987 + ...]

+ NLO Born Normalization [Nason 2004; Fixione, Nason, Oleari 2007]

Shower then takes over and generates all further emissions

Using soft/collinear approximations



Pseudorapidity of the emitted parton

Nason 2004; Fixione, Nason, Oleari 2007

Generic emission phase space

Phase Space already Covered by Powhea Powheg Emission generated with $[M_{X+1}^{(0)}]^2$ -Phase Space Covered by shower

Powheg Box – A Subtlety

Industry Standard: "Powheg Box"

- Exploits having its own definition of "p_T"
- \neq shower's definition of p_T
- Breaks clean matching

Solution: Vetoed Showers

- (+ truncated showers)
- Works very well for simple cases

Induces an uncertainty/ambiguity

Purely associated with the matching scheme (not physical)

Can be important for complex / multi-scale processes.



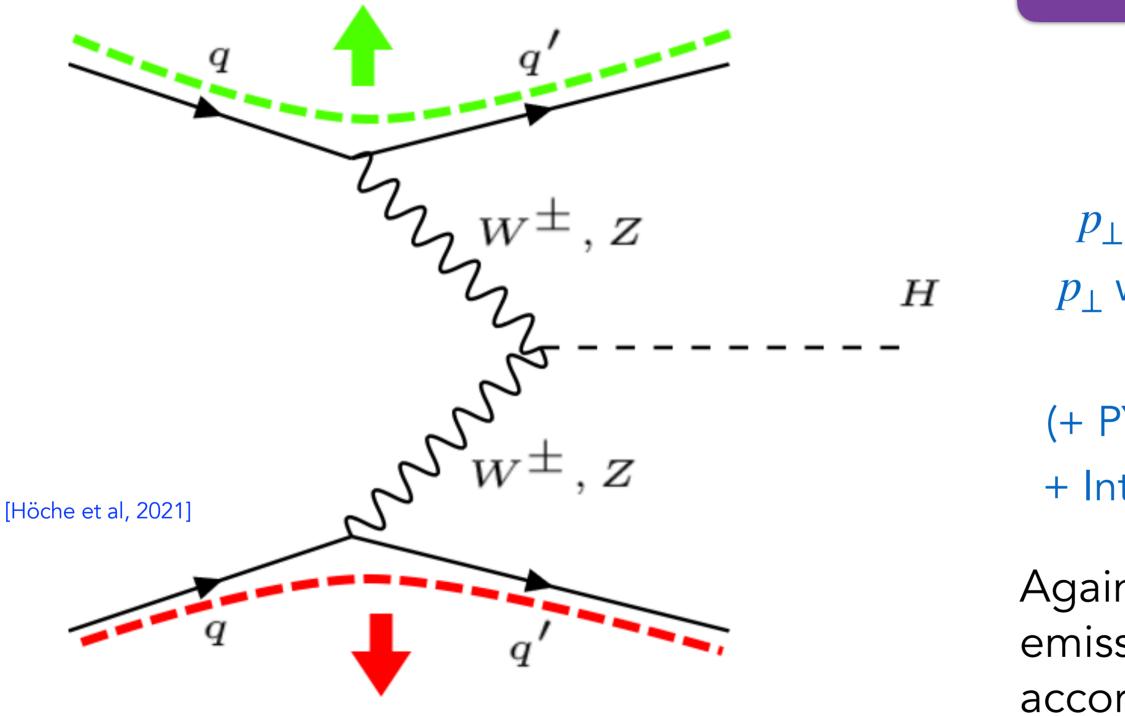


[Alioli et al, 2010]

Mismatched phase-space regions Phase Space already Covered by Powheg Powheg Emission generated with [M⁽⁰⁾_{X+1}]² -Phase Space Covered by shower

A More Complex Process

Vector boson fusion, $qq \rightarrow q'q'H$



Note: similar concerns for any process with coloured partons in the final state at Born level $t\bar{t} \ (\& t \rightarrow bW), V/H + jet(s), dijets, trijets, ...$

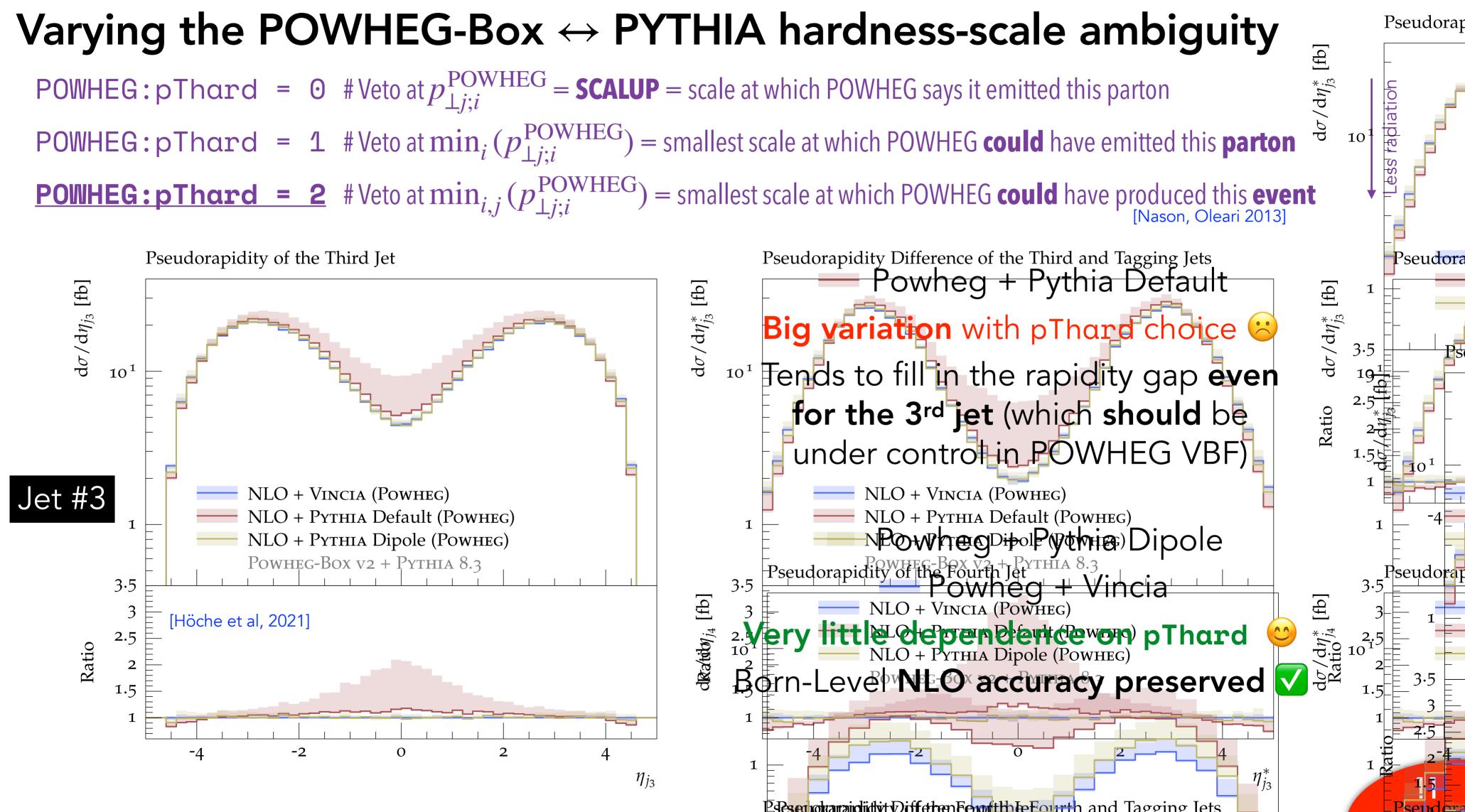
Multiple emitters → several overlapping phase spaces

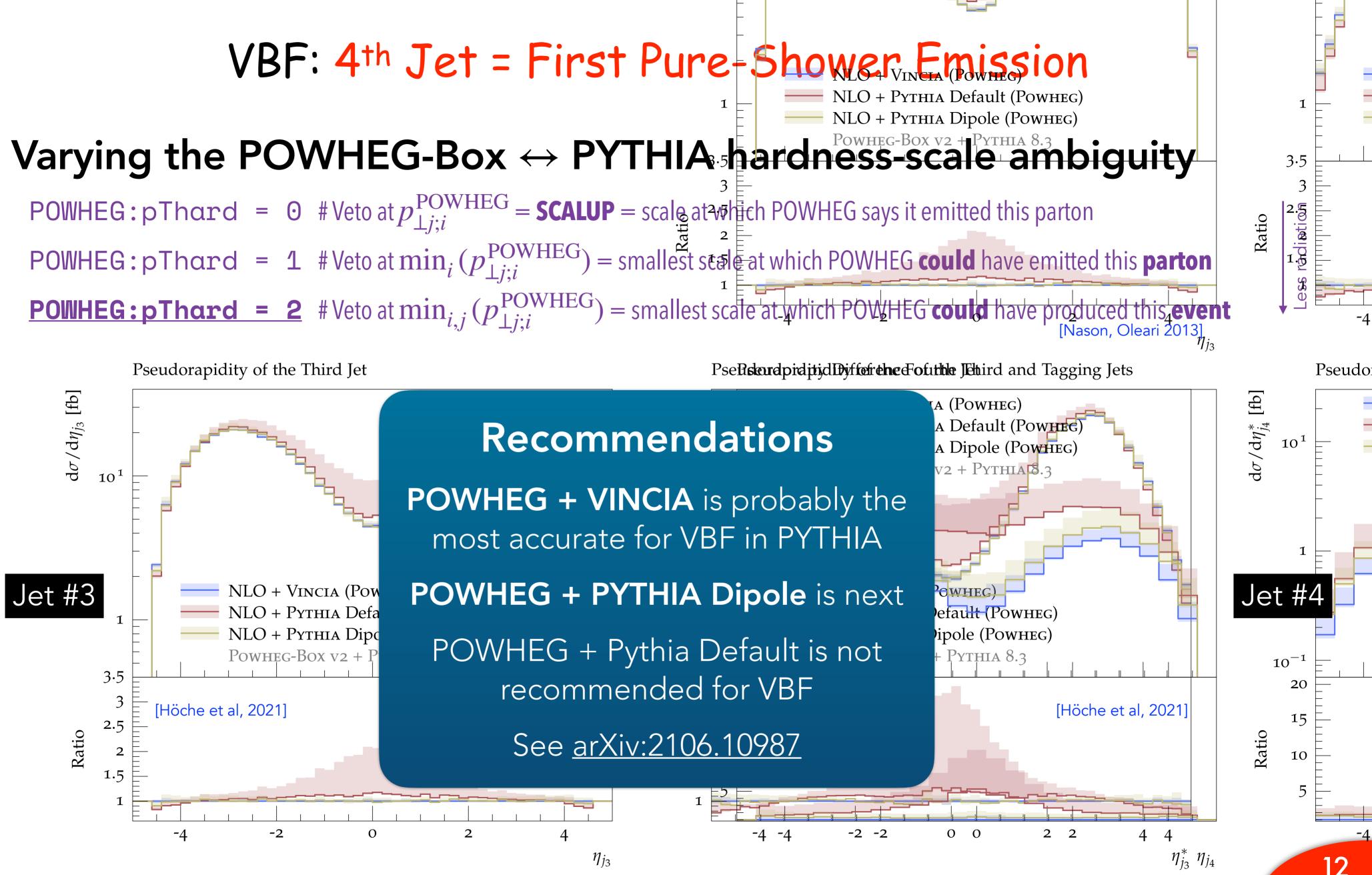
And many possible p_T definitions:

 p_{\perp} with respect to the beam p_{\perp} with respect to the final-state q' partons p_{\perp} with respect to either of the (q_*q') dipoles p_{\perp} with respect to the H? (+ PYTHIA defines a problematic (q'q') dipole) + Interpolations/combinations of the above ...

Again, POWHEG-Box generates the first emission, which it judges to be the "hardest" according to its own p_T definition

POWHEG-Box Matching Systematics





Pseudorapidity Difference of the Fourth and Tagging Lets

2. From NLO to NNLO

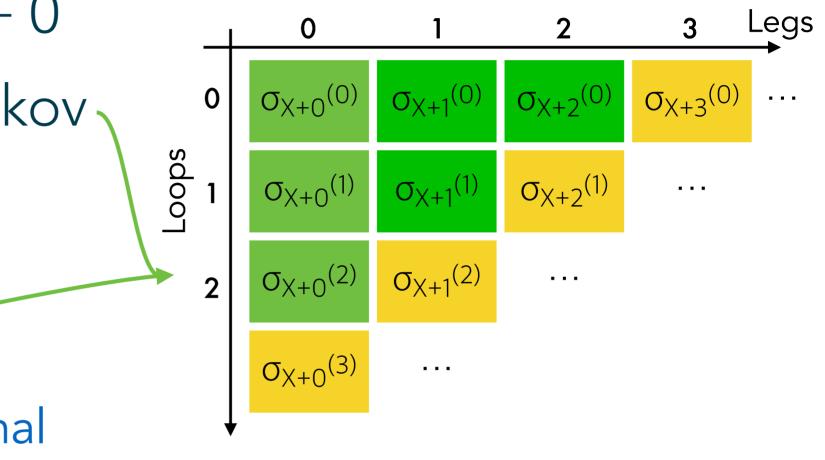
MiNNLO_{PS} builds on (extends) POWHEG NLO for X + jet

- Allow the first jet to approach $p_{\perp} \rightarrow 0 \sim X + 0$
- Tame divergence with analytic (NNLL) Sudakov
 - (introduces additional hardness scale = resummation scale)
- **Normalize** inclusive $d\sigma_X$ to NNLO
 - ($\mathcal{O}(\alpha_s^3)$) ambiguity on how to "spread" the additional contributions in phase space.)

~ Best you can do with current off-the-shelf parton showers

Is approximate; introduces some ambiguities: $p_{\perp}^{\text{Shower}} \operatorname{vs} p_{\perp}^{\text{Powheg}} \operatorname{vs} Q_{NNLL}^{\text{resummation}}$ & differential NNLO spreading (+ possible efficiency bottleneck: $p_{\perp} \rightarrow 0$ singularity × Sudakov veto)

[Hamilton et al. 1212.4504] Monni et al. 1908.069871



What if we could lift that restriction?

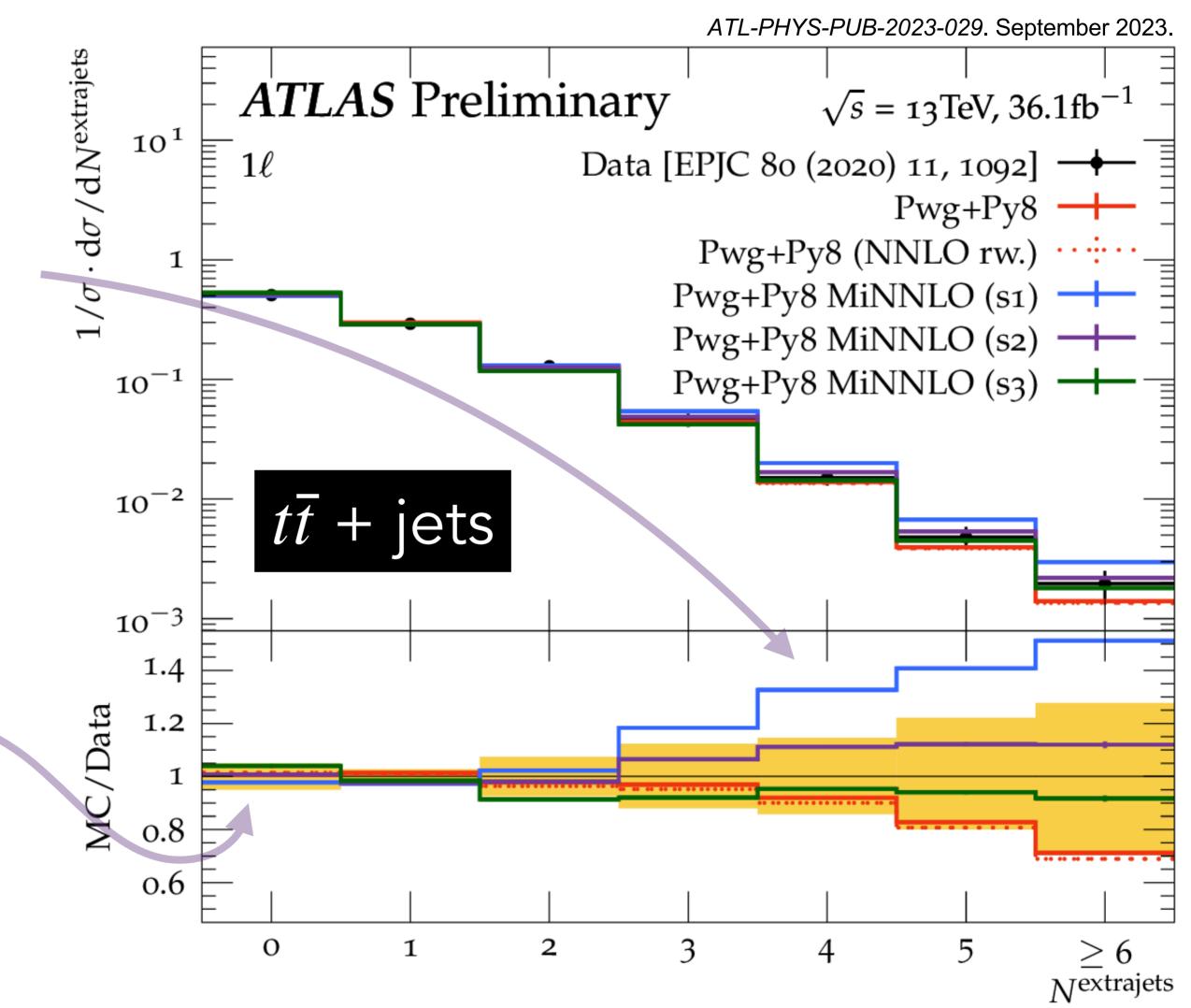
MINNLOPS inherits some issues from POWHEG-Box

Large dependence on pThard scale

Big variations in predictions for further jets

Calculation "anchored" in NLO for X+jet

⇒ Fairly big variations for Born-level (0-jet) observable.



Recommendations to Users of these Calculations

MiNNLO_{PS} is an approximate matching scheme ~ Best you can do with current off-the-shelf parton showers! But: does not "match" shower to NNLO point by point in phase space (Impossible to do so with LL showers.) Does not (always) do vetoed showers (This can in principle be done.) Depends on several auxiliary scales (Intrinsic to scheme. Physical observables should not depend on them \rightarrow vary!)

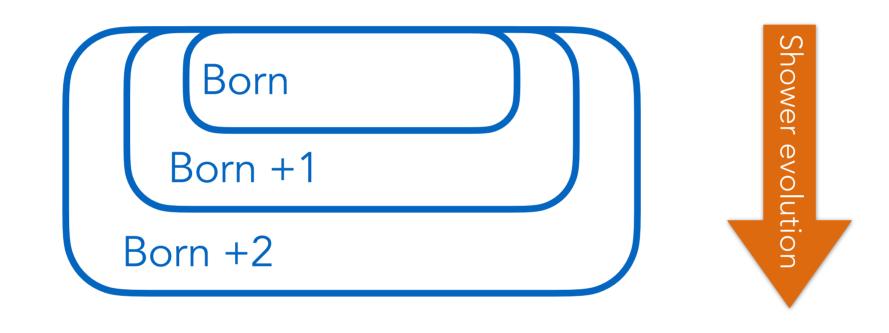
Do comprehensive variations to estimate scheme uncertainties Subsequent shower not fully guaranteed to preserve accuracy (Also applies to POWHEG + showers)

Towards True* NNLO Matching

*In the sense of the fixed-order and shower calculations matching each other point by point in each phase space

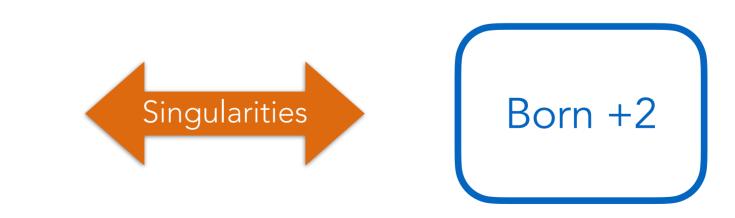
Idea: Use (nested) Shower Markov Chain as NNLO Phase-Space Generator

- Harnesses the power of showers as efficient phase-space generators for QCD
 - Pre-weighted with the (leading) QCD singular structures = soft/collinear poles



Different from conventional Fixed-Order phase-space generation (eg VEGAS)



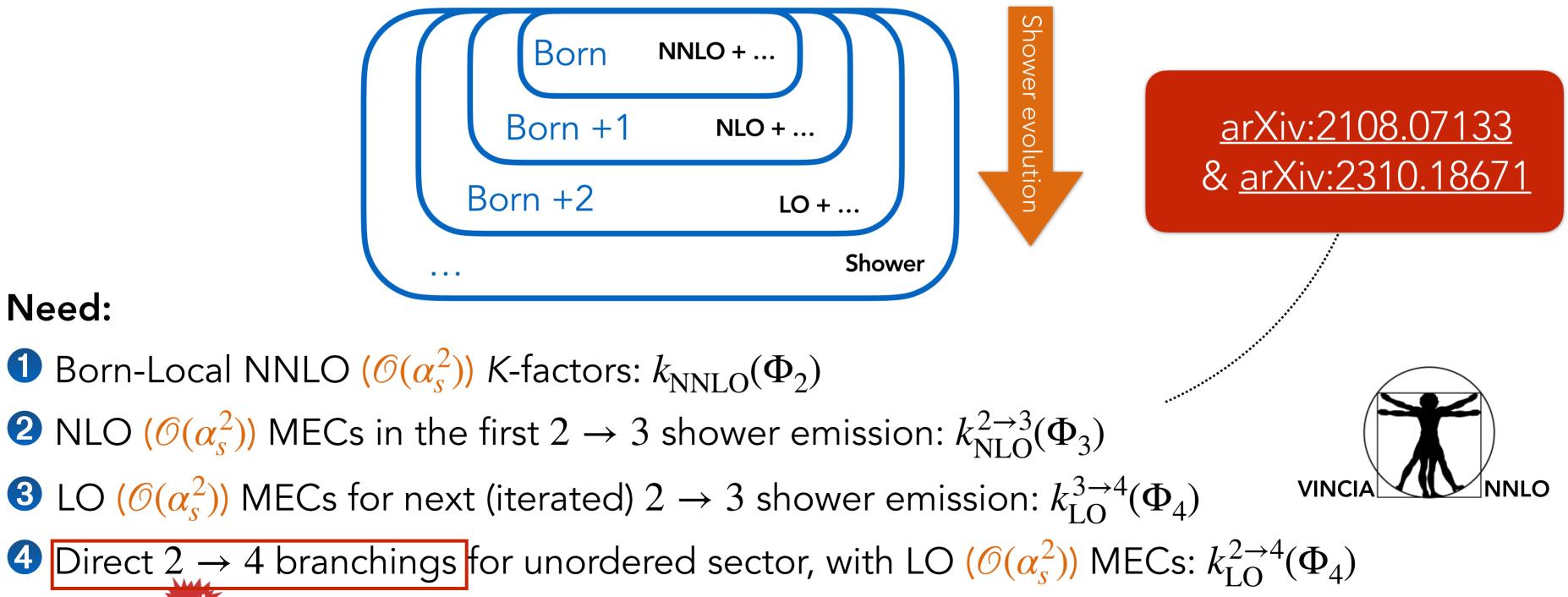


Towards True* NNLO Matching

*In the sense of the fixed-order and shower calculations matching each other point by point in each phase space

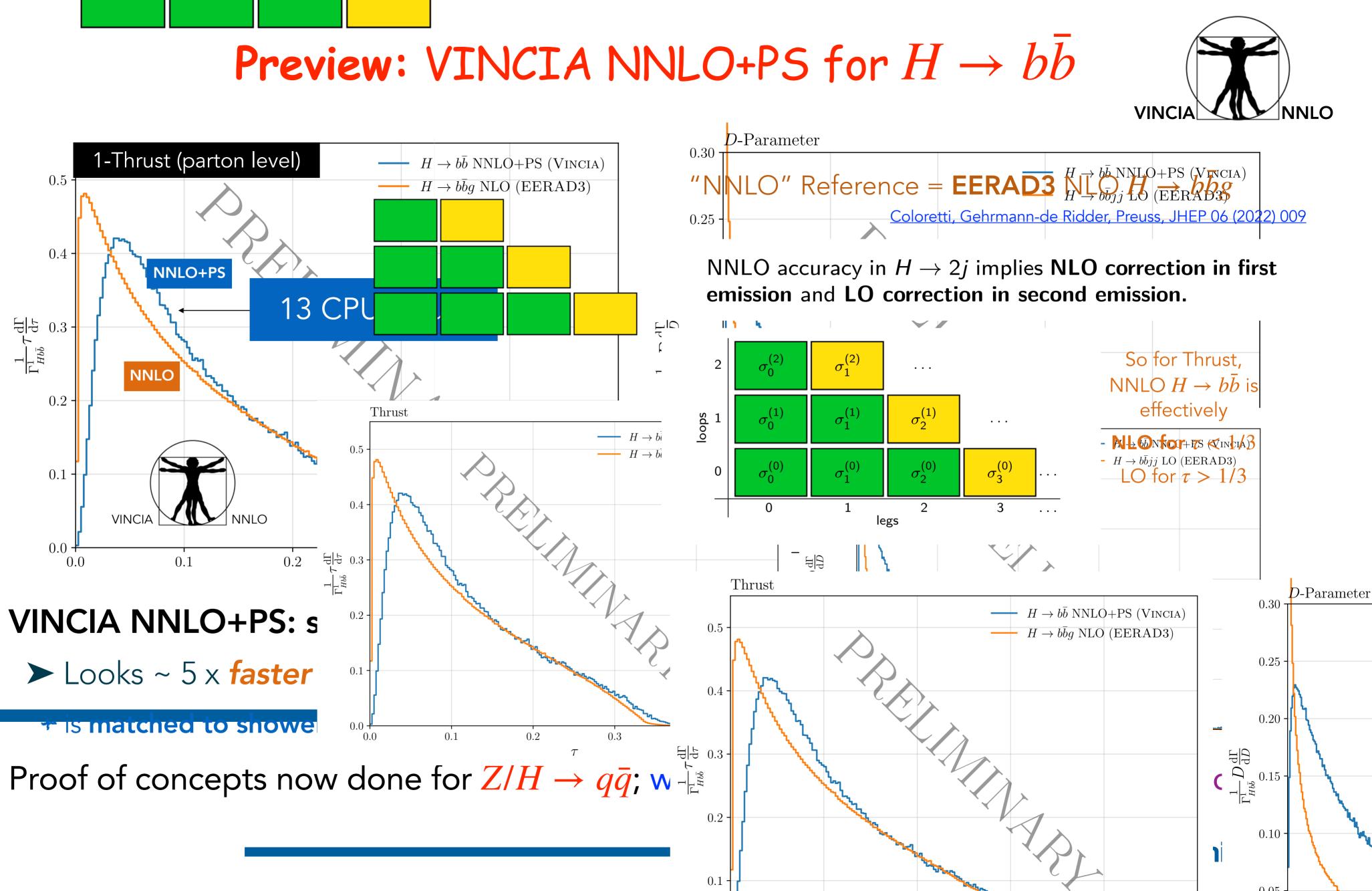
Continue shower afterwards

- No auxiliary / unphysical scales
 - \Rightarrow expect small matching systematics



Need:

- **1** Born-Local NNLO ($\mathcal{O}(\alpha_s^2)$) K-factors: $k_{NNLO}(\Phi_2)$
- 2 NLO ($\mathcal{O}(\alpha_s^2)$) MECs in the first $2 \rightarrow 3$ shower emission: $k_{\rm NLO}^{2\rightarrow 3}(\Phi_3)$



Recent Studies

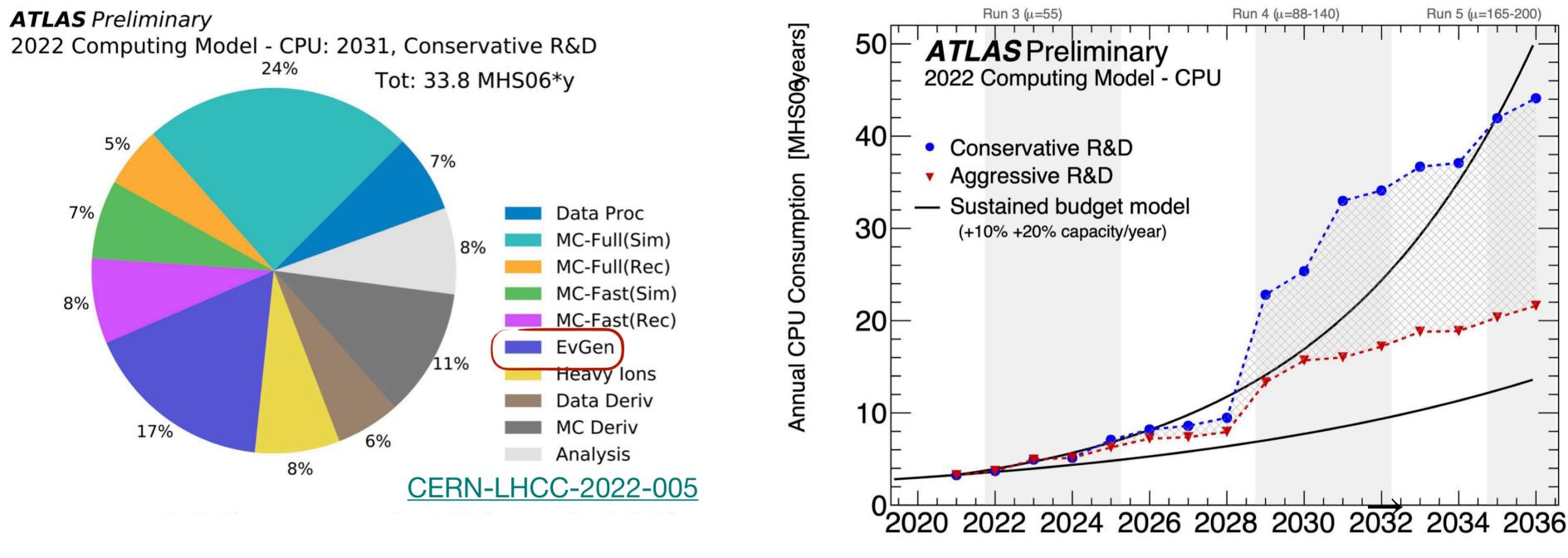
Focus on SM precision environments ↔ BSM backgrounds

- **NLO** Matching Systematics with POWHEG-Box (examples: VBF, tt) 1.
- 2. From NLO to NNLO (examples: tī, V, H, VH, VV, ...)
- 3. The computational bottleneck in **ME merging** (example: V+jets)
- 4. New Discoveries in Hadronization (examples: HF baryons, JES)

NB: want to address/explain state of the art & systematics in real contexts \rightarrow a bit theory heavy

The Computational Bottleneck in ME Merging

Something about CPU resources



Largest fraction of EvGen CPU time is taken by generation of multi-leg MC predictions

namely, multijet merged Sherpa V+jets

Year

Matrix-Element Merging — The Complexity Bottleneck

For CKKW-L style merging: (incl UMEPS, NL3, UNLOPS, ...)

Need to take all contributing "shower histories" into account.

In conventional parton showers (Pythia, Herwig, Sherpa, ...)

Each phase-space point receives contributions from many possible branching "histories" (aka "clusterings")

of histories grows ~ # of Feynman Diagrams, faster than factorial

Number of Histories for n Branchings							
Starting from a single $qar q$ pair	$\mid n=1$	n = 2	n = 3	n = 4	n = 5	n = 6	n = 7
CS Dipole	2	8	48	384	3840	46080	645120

Bottleneck for merging at high multiplicities (+ high code complexity)



Sector Showers (without maths)

VINCIA's shower is unique in being a "Sector Shower" Partition N-gluon Phase Space into N "sectors" (using step functions). Each sector \leftrightarrow one specific gluon being the "softest" in the event Inside each sector, only one kernel contributes (the most singular one)! Sector Kernel = the eikonal for the soft gluon and its collinear DGLAP limits for z > 0.5. Unique properties: shower operator is bijective and is a true Markov chain

The crucial aspect:

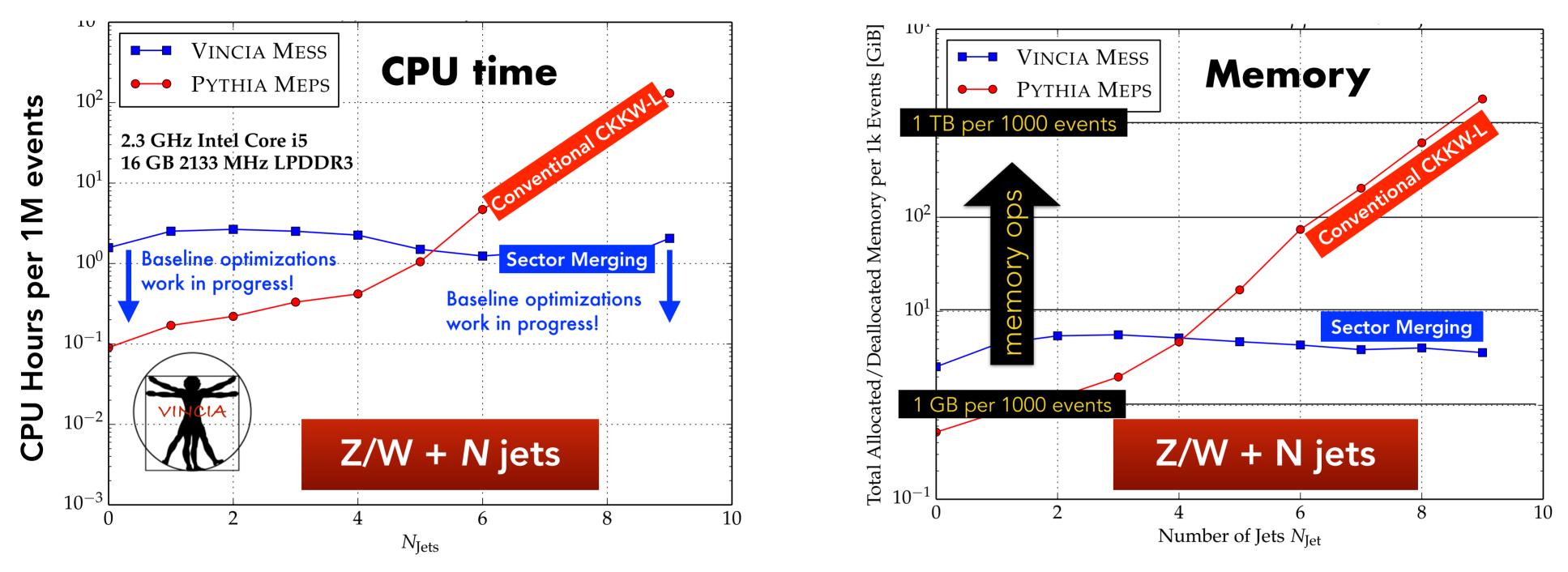
Only a **single history** contributes to each phase-space point !

\implies Factorial growth of number of histories reduced to constant! (And the number of sectors only grows linearly with the number of gluons) $(g \rightarrow q\bar{q} \rightarrow q\bar{q} \rightarrow leftover factorial in number of same-flavour quarks; not a big problem)$

- **PS** & Villarejo JHEP 11 (2011) 150 Brooks, Preuss, **PS** JHEP 07 (2020) 032

Sectorized CKKW-L Merging publicly available from Pythia 8.306

Brooks & Preuss, "Efficient multi-jet merging with the VINCIA sector shower", arXiv:2008.09468



Demonstrated constant scaling with multiplicity. Extensions now pursued:

Optimisations of baseline algorithm

Sectorized iterated tree-level ME corrections (demonstrated in PS & Villarejo arXiv:1109.3608)

Sectorized multi-leg merging at NLO (active research grants, with C. Preuss, Wuppertal)

Recent Studies

Focus on SM precision environments ↔ BSM backgrounds

- **NLO** Matching Systematics with POWHEG-Box (examples: VBF, tt) 1.
- 2. From NLO to NNLO (examples: tt, V, H, VH, VV, ...)
- 3. The computational bottleneck in **ME merging** (example: V+jets)

4. New Discoveries in Hadronization (examples: HF baryons, JES)

NB: want to address/explain state of the art & systematics in real contexts \rightarrow a bit theory heavy

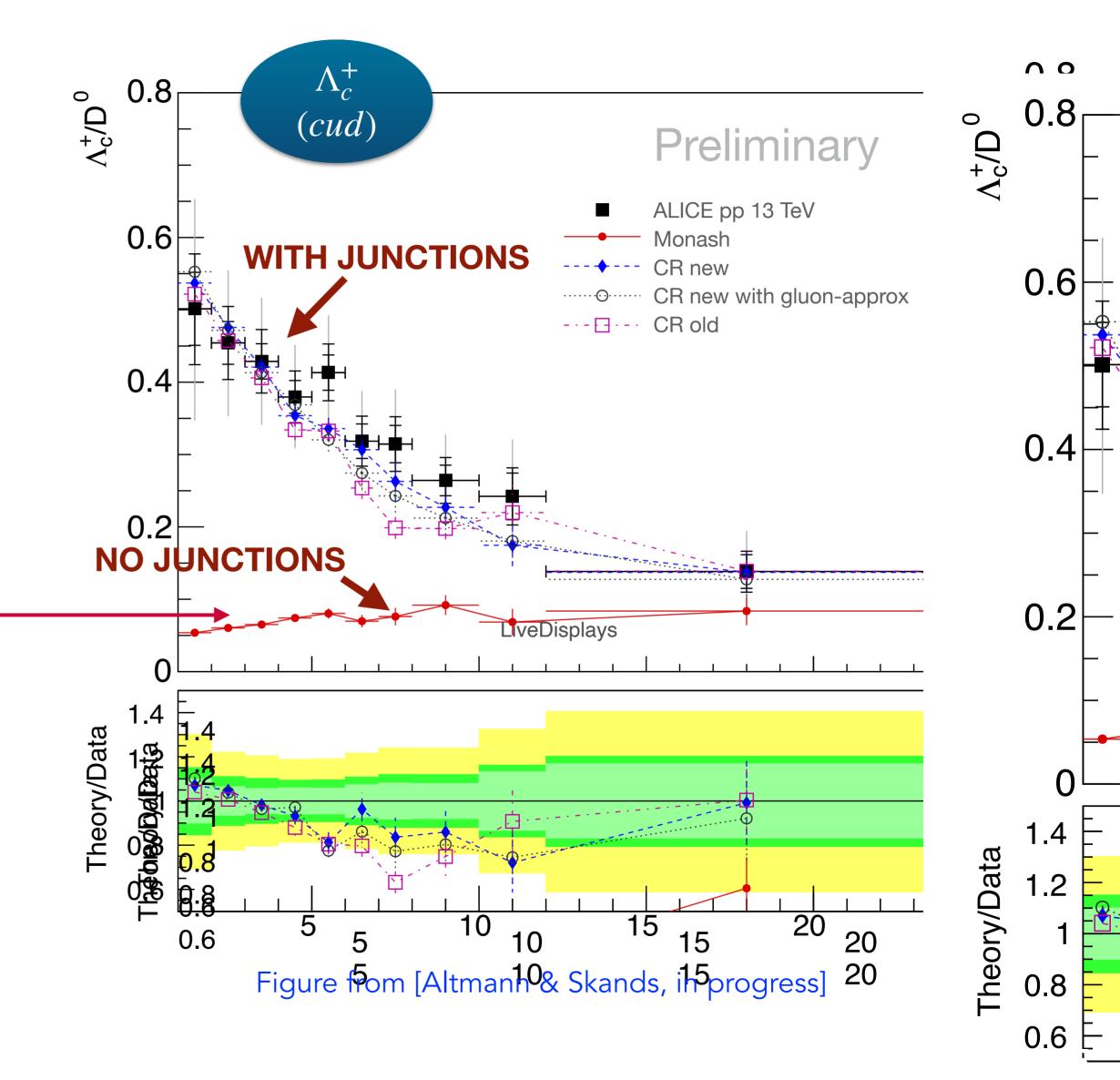


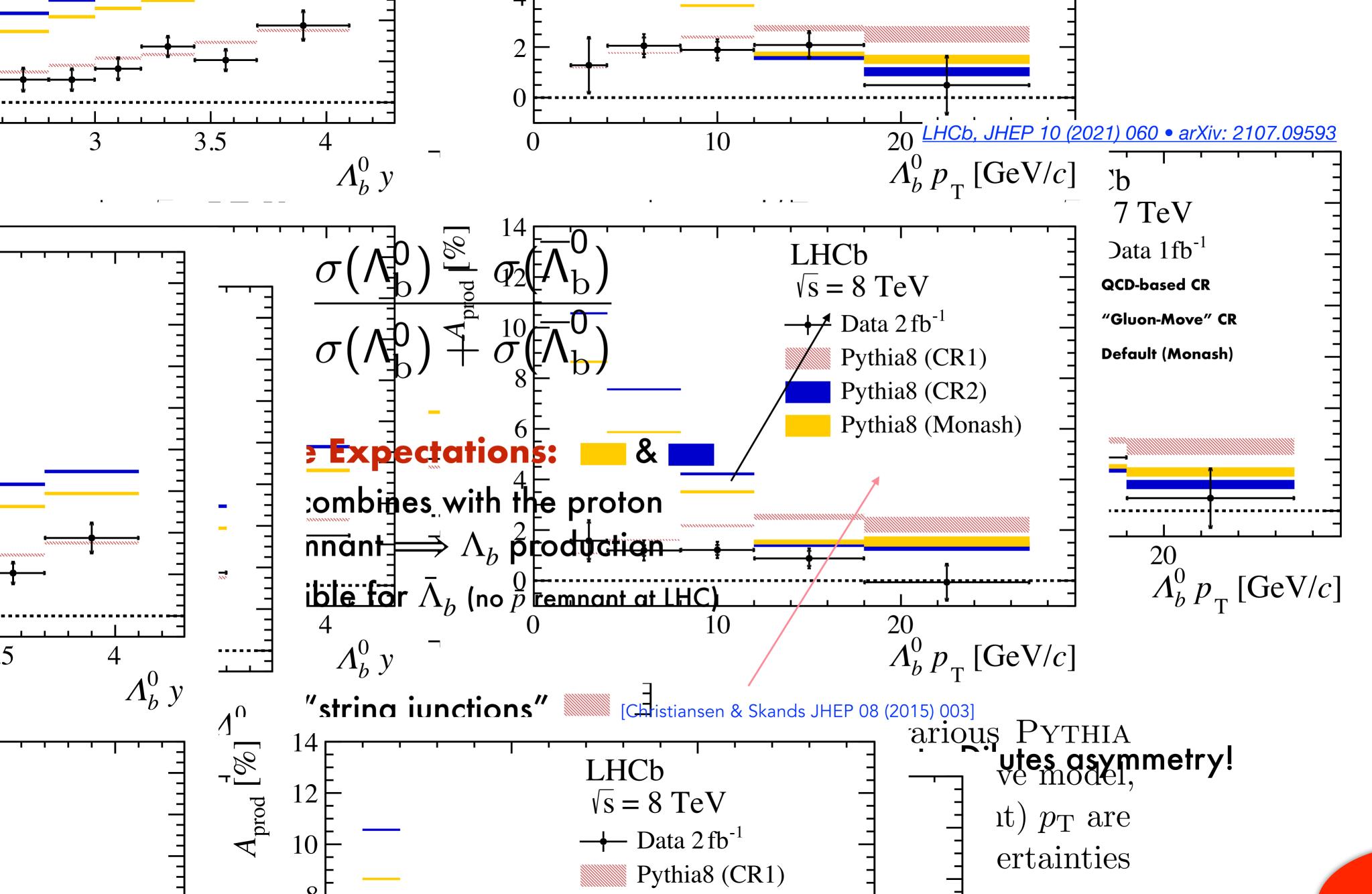
New Discoveries in Hadronization

LHC experiments report very large (factor-10) enhancements in heavyflavour baryon-to-meson ratios at low p_T!

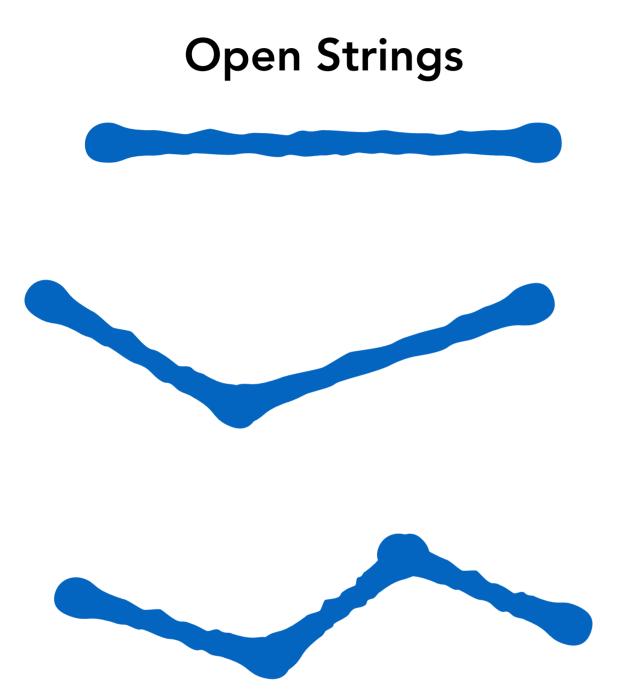
Not predicted by default Pythia (Monash)

Very exciting!

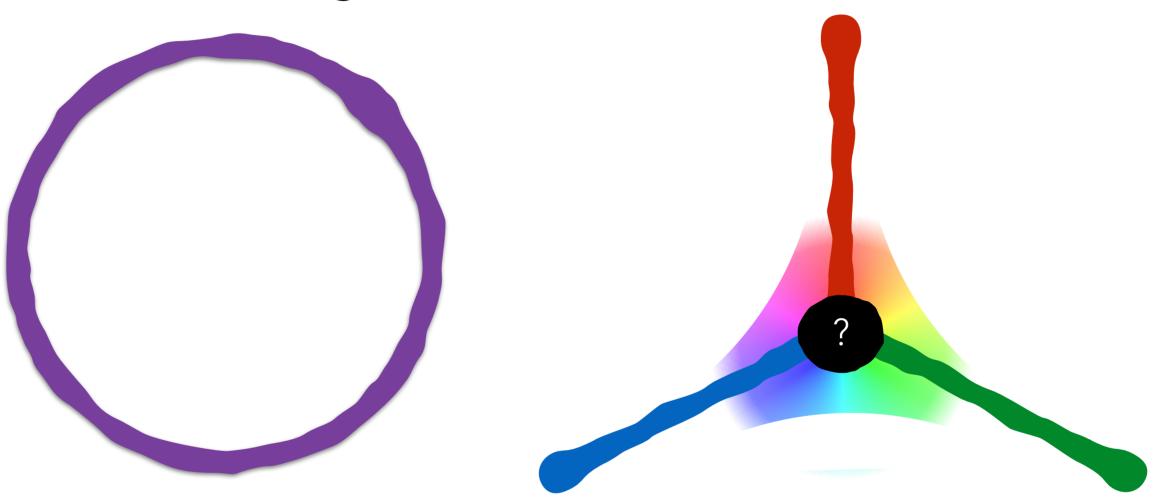




What are "String Junctions"?



Closed Strings



 $q\bar{q}$ strings (with gluon kinks) E.g., $Z \rightarrow q\bar{q}$ + shower $H \rightarrow b\bar{b} + \text{shower}$

Gluon rings E.g., $H \rightarrow gg$ + shower Open strings with $N_C = 3$ endpoints $\Upsilon \rightarrow ggg + shower$ E.g., Baryon-Number violating neutralino decay $\tilde{\chi}^0 \rightarrow qqq$ + shower



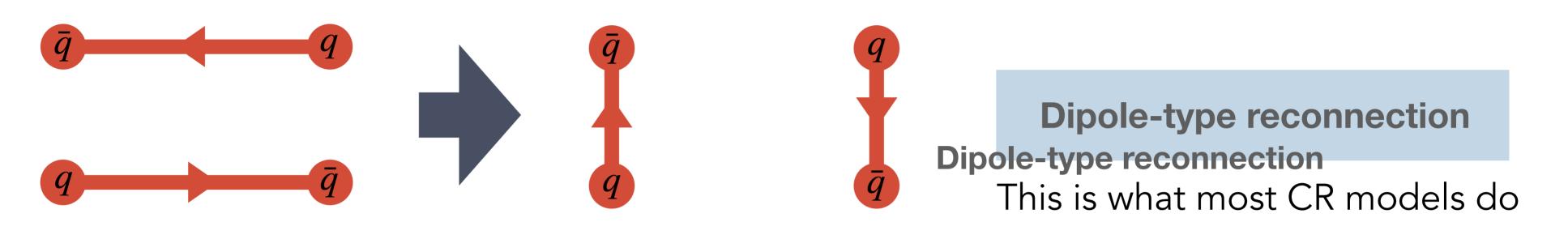
SU(3) String Junction

27

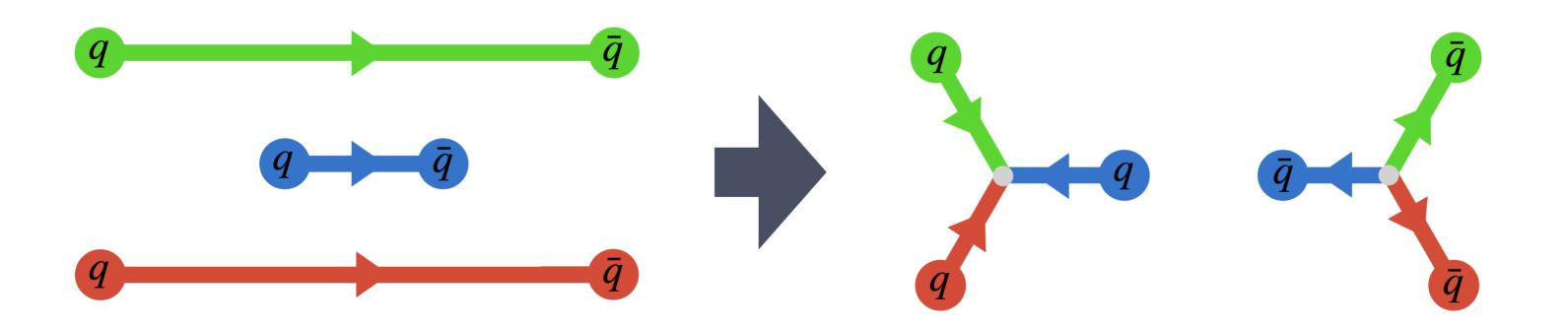
How do Colour Reconnections Create String Junctions?

Stochastically restores colour-space ambiguities according to SU(3) algebra

> Allows for reconnections to minimise string lengths



What about the red-green-blue colour singlet state?



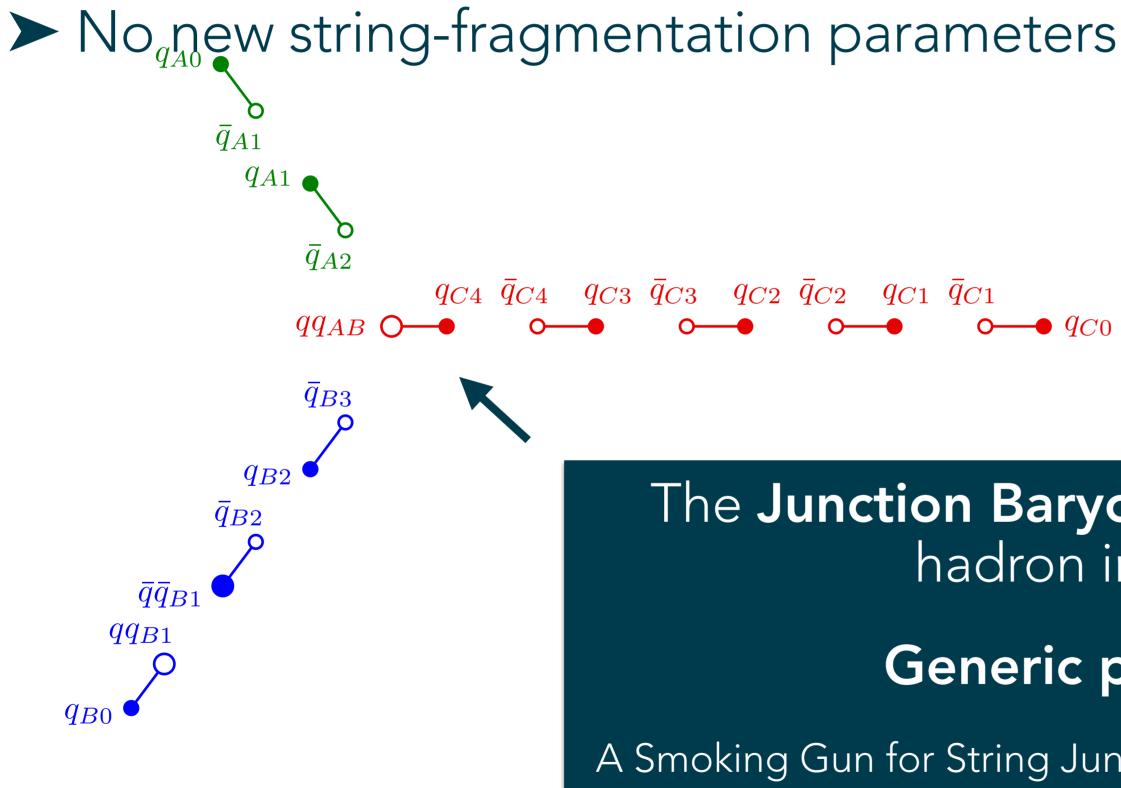
[Christiansen & **PS**] JHEP 08 (2015) 003]



Junctions!

What do String Junctions do?

Assume Junction Strings have same properties as ordinary ones (u:d:s, Schwinger p_T, etc)



[Sjöstrand & **PS**, <u>NPB 659 (2003) 243</u>] [+ J. Altmann & **PS**, in progress]

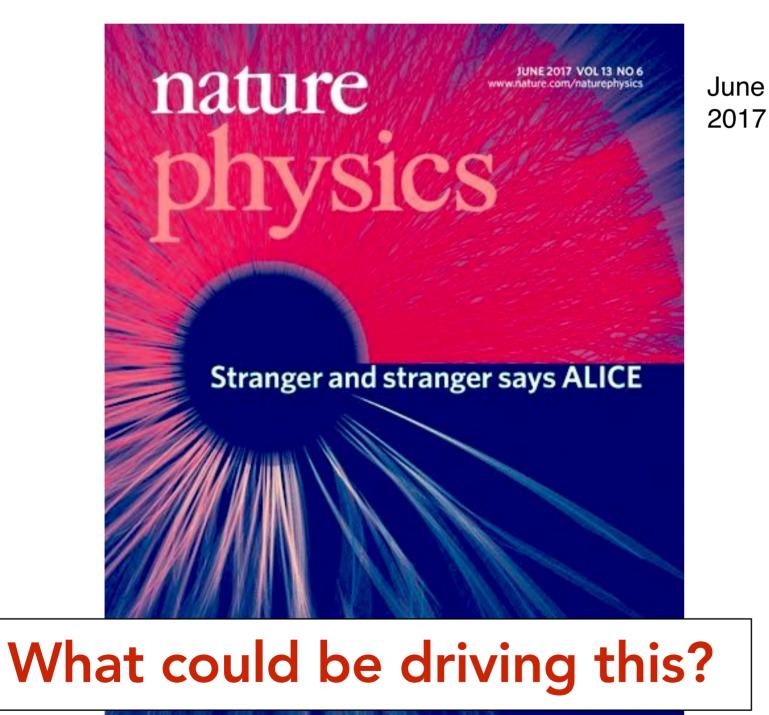
The Junction Baryon is the most "subleading" hadron in all three "jets".

Generic prediction: low p_T

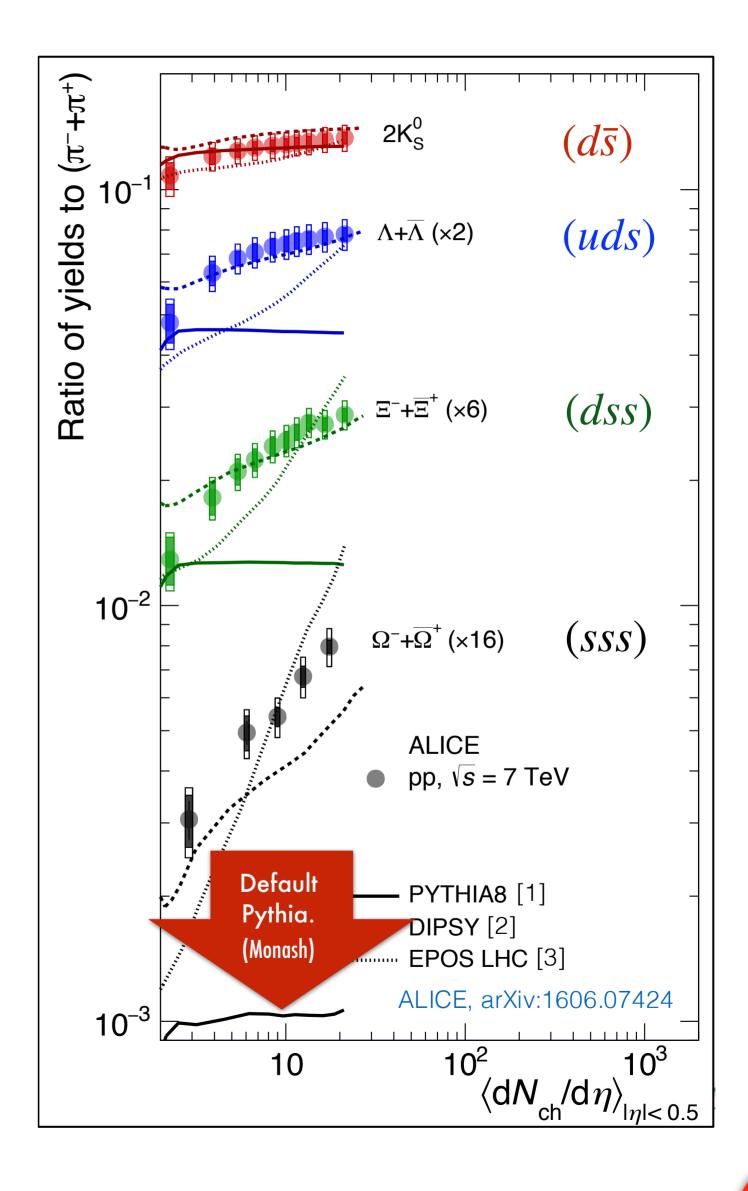
A Smoking Gun for String Junctions: **Baryon enhancements at low p**

What a strange world we live in, said Alice

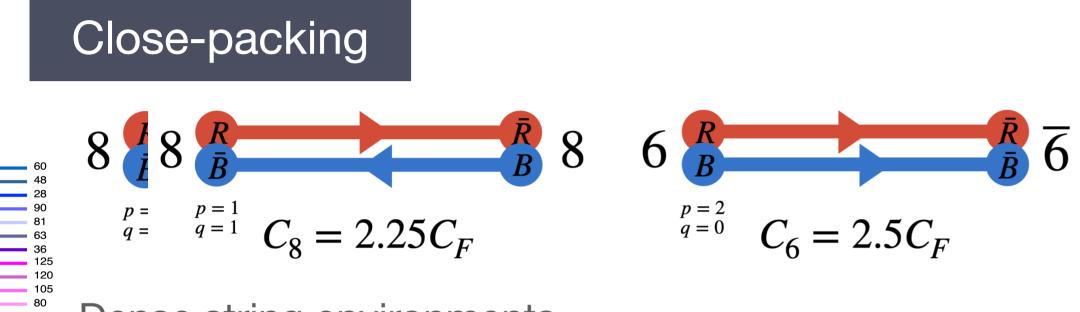
We also know ratios of strange hadrons to pions strongly increase with event activity

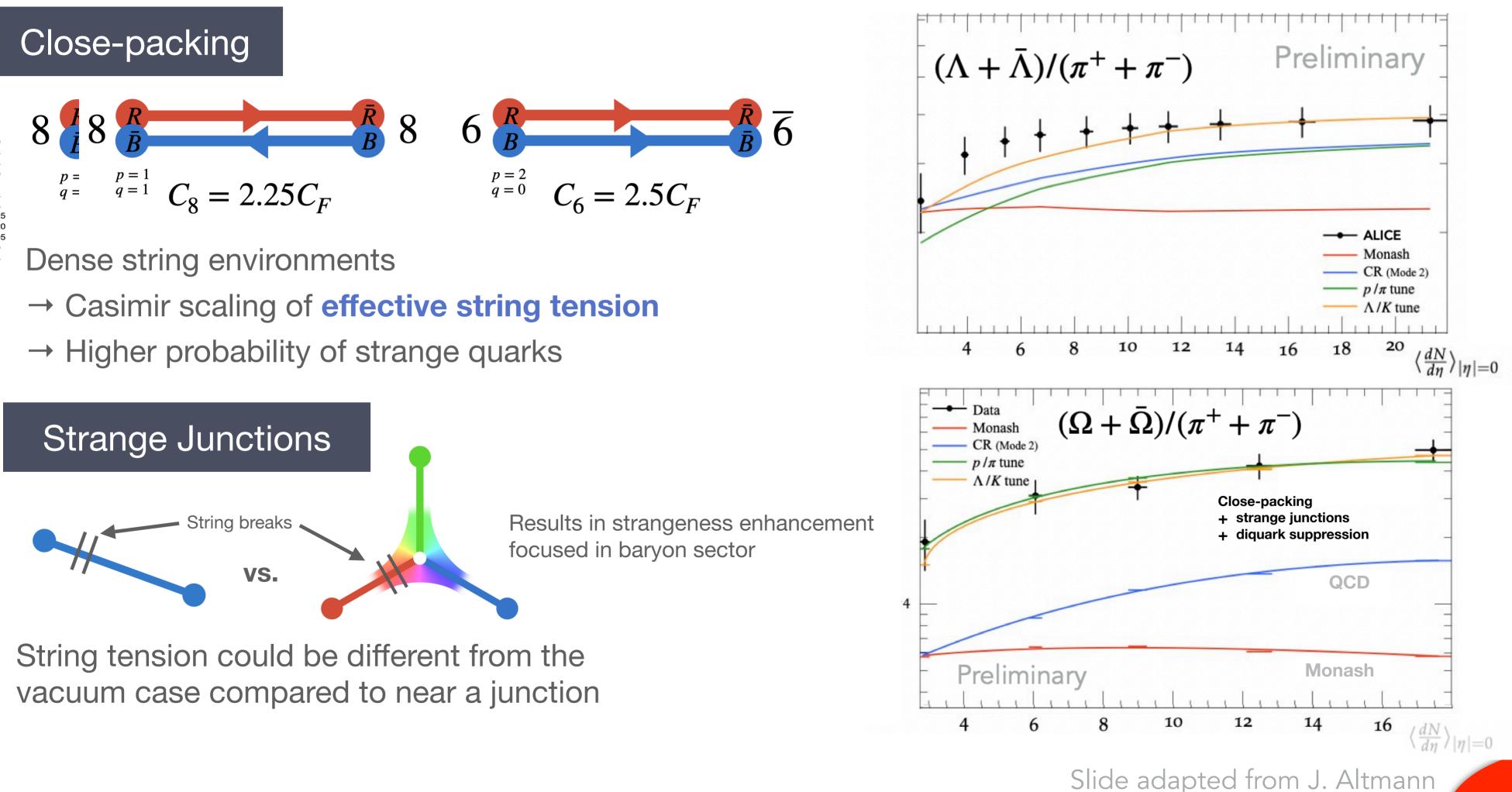


TOPOLOGICAL PHOTONICS Optical Weyl points and Fermi arcs



FIn Progress: Strangeness Enhancement from Close-Packing S Enhancement Idea: each string exists in an effective background produced by the others





Particle Composition: Impact on Jet Energy Scale



ATLAS PUB Note

ATL-PHYS-PUB-2022-021

29th April 2022



Dependence of the Jet Energy Scale on the Particle Content of Hadronic Jets in the ATLAS Detector Simulation

The dependence of the ATLAS jet energy measurement on the modelling in Monte Carlo simulations of the particle types and spectra within jets is investigated. It is found that the hadronic jet response, i.e. the ratio of the reconstructed jet energy to the true jet energy, varies by $\sim 1-2\%$ depending on the hadronisation model used in the simulation. This effect is mainly due to differences in the average energy carried by kaons and baryons in the jet. Model differences observed for jets initiated by *quarks* or *gluons* produced in the hard scattering process are dominated by the differences in these hadron energy fractions indicating that measurements of the hadron content of jets and improved tuning of hadronization models can result in an improvement in the precision of the knowledge of the ATLAS jet energy scale.

- Variation largest for gluon jets For $E_T = [30, 100, 200]$ GeV Max JES variation = [3%, 2%, 1.2%]
- Fraction of jet E_T carried by baryons (and kaons) varies significantly
 - Reweighting to force similar baryon and kaon fractions
 - Max variation → [1.2%, 0.8%, 0.5%]
 - Significant potential for improved Jet Energy Scale uncertainties!
- Motivates Careful Models & Careful Constraints
 - Interplay with advanced UE models
 - In-situ constraints from LHC data
 - Revisit comparisons to LEP data

Summary & Outlook

State of the art for perturbation theory: NNLO (\rightarrow N3LO)

Showers + hadronization mandatory for collider studies (+ resummation extends range)

Now: can use off-the-shelf showers with MiNNLO_{PS}

Based on POWHEG-Box + Analytical Resummation + NNLO normalisation Approximate method; depends on several auxiliary unphysical scales \rightarrow can exhibit large variations

Work in progress: VinciaNNLO

Based on nested shower-like phase-space generation with second-order MECs True NNLO matching -> Expect small matching systematics So far only worked out for colour-singlet decays. (Also developing extensions towards NLL, NNLL showers ...)

Beautiful Strings

New discoveries at LHC on particle composition, esp. **baryons and strangeness**

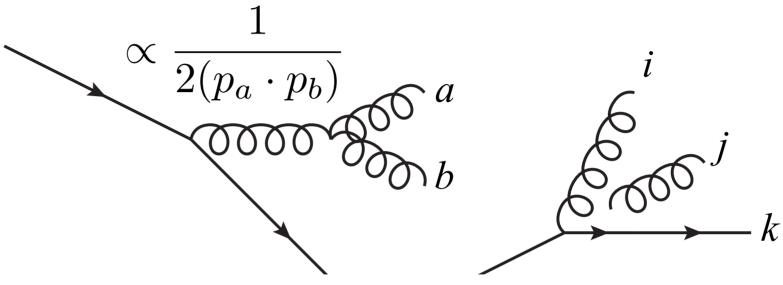
New research grant with LHCb (Warwick) focusing on strings with b-quark endpoints And QED corrections in B decays



Extra Slides

Parton Showers: Theory

Most bremsstrahlung is driven by divergent **propagators** \rightarrow simple structure



Mathematically, gauge amplitudes factorize in singular limits

Partons ab \rightarrow collinear: $|\mathcal{M}_{F+1}(\ldots, a, b, \ldots)|^2 \xrightarrow{a||b} g_s^2 \mathcal{C} \frac{P(a)}{2(p_a)}$

P(z) = DGLAP splitting kernels'

Gluon j

$$\rightarrow$$
 soft: $|\mathcal{M}_{F+1}(\ldots,i,j,k\ldots)|^2 \xrightarrow{j_g \to 0} g_s^2 \mathcal{C} \frac{(p_i \cdot p_k)}{(p_i \cdot p_j)(p_j \cdot p_k)} |\mathcal{M}_F(\ldots,i,k,\ldots)|^2$

Coherence \rightarrow Parton j really emitted by (i,k) "dipole" or "antenna" (eikonal factors)

These are the **building blocks of parton showers** (DGLAP, dipole, antenna, ...) (+ running coupling, unitarity, and explicit energy-momentum conservation.)

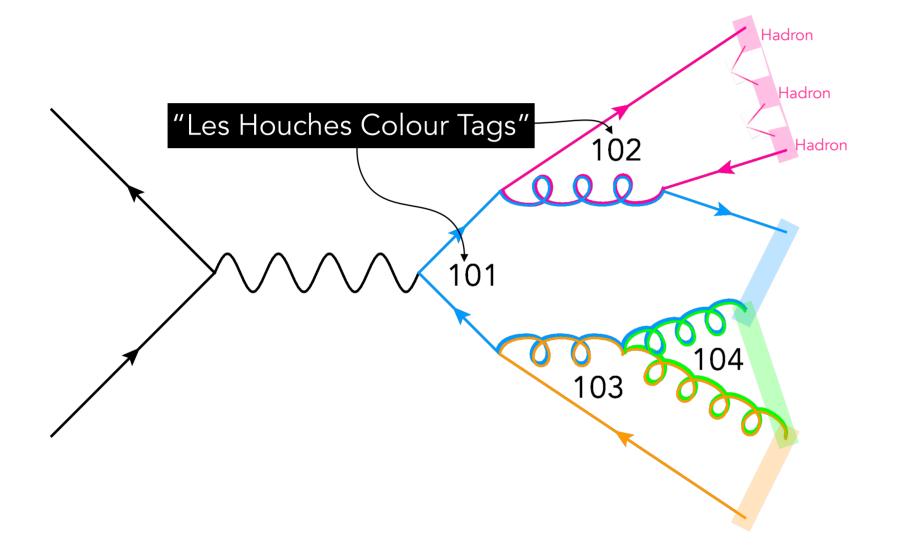
see e.g PS, Introduction to QCD, TASI 2012, arXiv:1207.2389

$$\frac{P(z)}{a \cdot p_b} |\mathcal{M}_F(\dots, a+b,\dots)|^2$$

', with
$$z = E_a / (E_a + E_b)$$

Confinement in PYTHIA: The Lund String Model

Simplified (leading-N_C) "colour flow" → determine between which partons to set up confining potentials

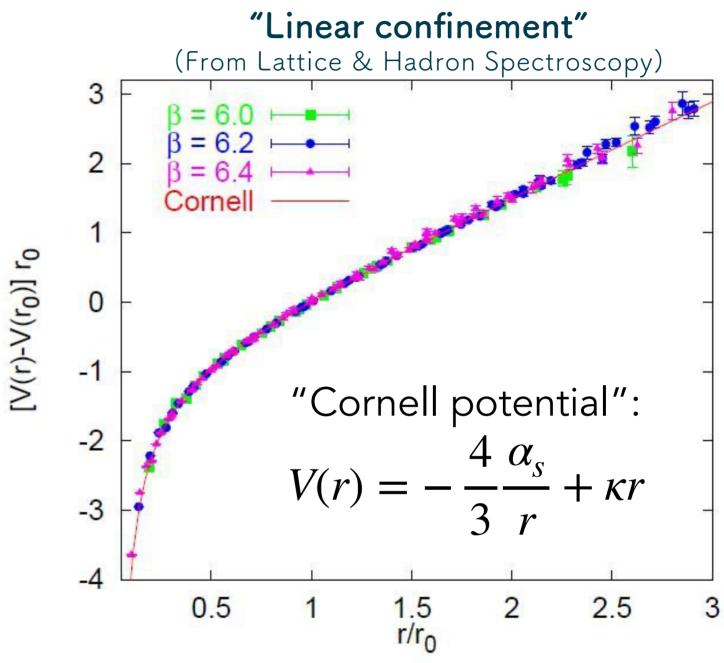


Map from Partons to Strings:

Quarks string endpoints; gluons transverse "kinks"

System then evolves as a string world sheet

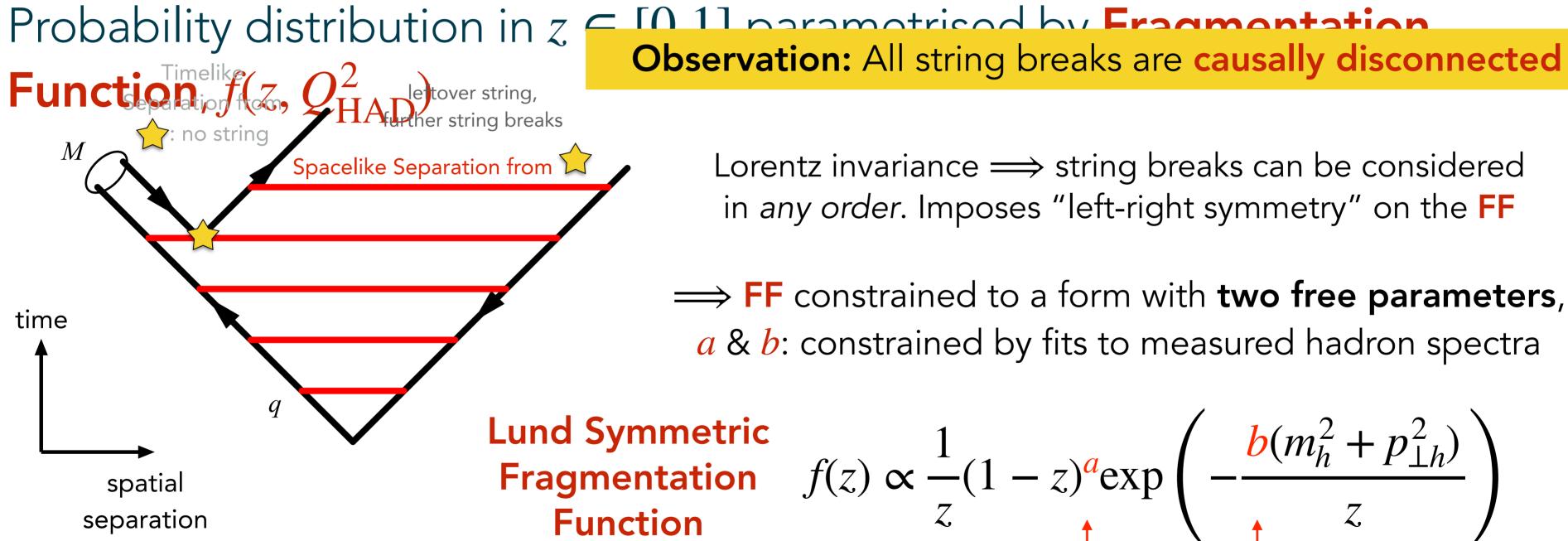
+ String breaks via spontaneous $q\bar{q}$ pair creation ("Schwinger mechanism") \rightarrow hadrons



The String Fragmentation Function

Consider a string break d_{2} , producing a meson M, and a leftover string piece

The meson M takes a fraction z of the quark momentum,



Lorentz invariance \implies string breaks can be considered in any order. Imposes "left-right symmetry" on the FF

 \implies **FF** constrained to a form with **two free parameters**, *a* & *b*: constrained by fits to measured hadron spectra

Lund Symmetric Fragmentation $f(z) \propto \frac{1}{z}(1-z)^{a} \exp\left(-\frac{b(m_{h}^{2}+p_{\perp h}^{2})}{t}\right)$ **Function**

Supresses high-z hadrons

Supresses low-z hadrons



Automated Hadronization Uncertainties

Problem:

Given a colour-singlet system that (randomly) broke up into a specific set of hadrons:

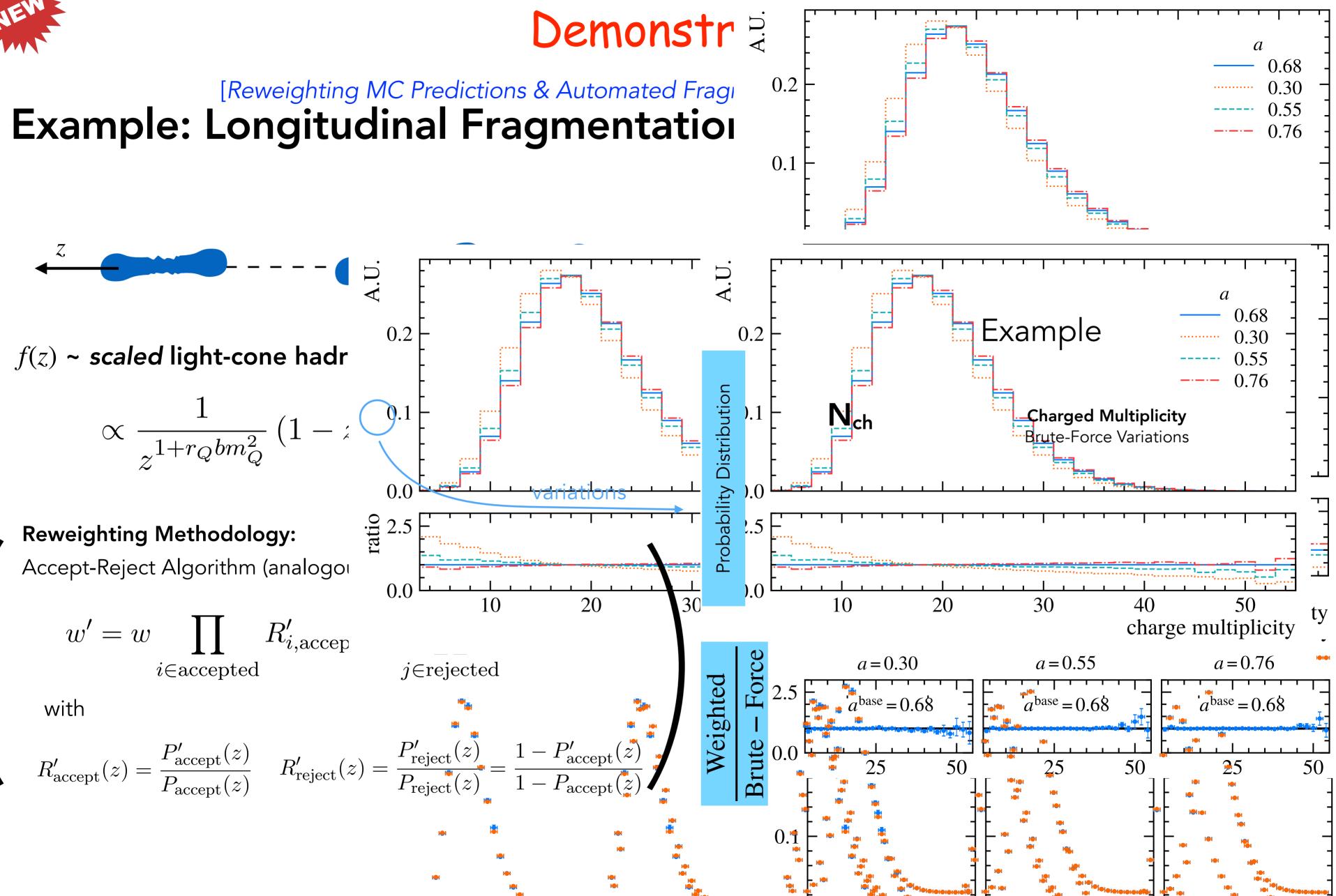


- What is the **relative probability** that same system would have resulted, if the fragmentation parameters had been different? Would this particular final state become more likely (w' > 1)? Or less likely (w' < 1)Crucially: maintaining unitarity \implies inclusive cross section remains unchanged!
- August 2023: Bierlich, Ilten, Menzo, Mrenna, Szewc, Wilkinson, Youssef, Zupan [Reweighting MC Predictions & Automated Fragmentation Variations in Pythia 8, 2308.13459]
 - Method is general; demonstrated on variations of the 7 main parameters governing longitudinal and transverse fragmentation functions in PYTHIA 8
- https://gitlab.com/uchep/mlhad-weights-validation





[Reweighting MC Predictions & Automated Fragi



39

A Brief History of MPI in PYTHIA

