CERN

Department of Theoretical Physics

The path to NNLL accurate parton showers

Alexander Karlberg

Joint INFN-UNIMI-UNIMIB Pheno Seminar

Based on

JHEP 03 (2023) 224 [K. Hamilton, AK, G. P. Salam, L. Scyboz, R. Verheyen] Phys.Rev.Lett. 131 (2023) 16 [S. Ferrario Ravasio, K. Hamilton, AK, G. P. Salam, L. Scyboz, G. Soyez] 2406.02661 [eid. + M. v. Beekveld, M. Dasgupta, B. K. El-Menoufi, J. Helliwell, P. F. Monni, A. Soto-Ontoso]

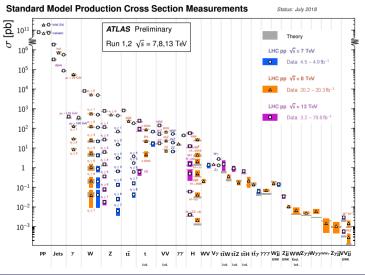
 $^+$

using analytic understanding developed in

JHEP 01 (2019) 083 [A. Banfi, B. K. El-Menoufi, P. F. Monni] JHEP 12 (2021) 158 [M. Dasgupta, B. K. El-Menoufi] JHEP 05 (2024) 09 [eid. + M. v. Beekveld, J. Helliwell, P. F. Monni]

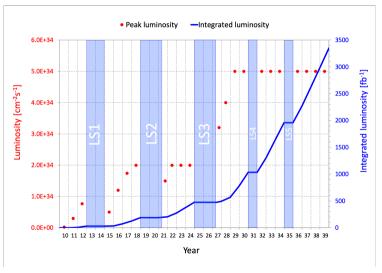
27th June 2024 Slide 1/42

The precision era of the LHC



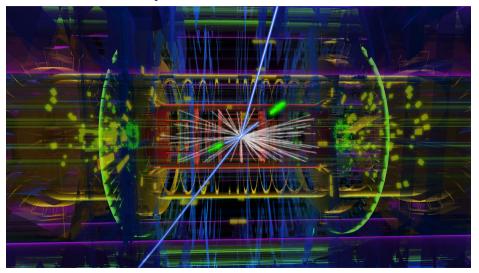


The precision era of the LHC





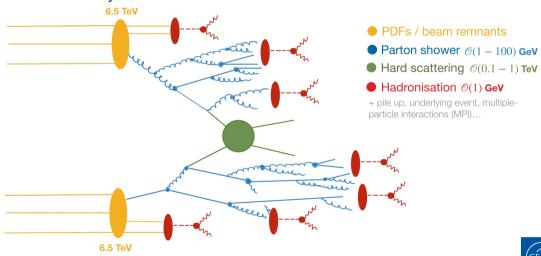
The LHC: A messy environment





Slide 4/42 — Alexander Karlberg — NNLL parton showers

Anatomy of an LHC collision





courtesy M. van Beekveld

The ubiquitous Parton Shower



Pythia 8

Torbiörn Siöstrand (Lund U., Dept. Theor. Phys.), Stefan Ask (Cambridge U.), Jesper R. Christiansen (Lund U., Dept. Theor. Phys.), Richard Corke (Lund U., Dept. Theor. Phys.),

Published in: Comput. Phys.Commun. 191 (2015) 159-177 + e-Print: 1410.3012 [hep-ph] 2 DOI ☐ cite

Herwig 7



Sherpa

<i>1</i> /1	Herwig++ Physics and Manual #1 Event generation wi	th SHERPA 1.1 #1
fan Ask (Cambridge U.), Jesper R. orke (Lund U., Dept. Theor. Phys.), 4)	Grellscheid (Durham U., IPPP), K. Hamilton (Louvain U.) et al. (Mar, 2008) Schonherr (Dresden, Tec	fan. Hoeche (Zurich U.), F. Krauss (Durham U., IPPP), M. ch. U.), S. Schumann (Edinburgh U.) et al. (Nov, 2008)
9-177 · e-Print: 1410.3012 [hep-ph]		2009) 007 • e-Print: 0811.4622 [hep-ph]
→ 4,050 citations	D pdf ∂ links ∂ DOI ⊡ cite 3 2,644 citations	∂ DOI ⊆ cite

Parton Showers enter one way or another in almost 95% of all ATLAS and CMS analyses. Collider physics would not be the same without them.

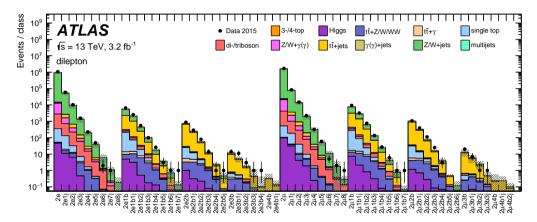


An introduction to PYTHIA 8 2

Ch ndt

Nishita Desai (U. Heidelberg, ITP) et al. (Oct 11, 2014)

The ubiquitous Parton Shower



ATLAS [1807.07447]



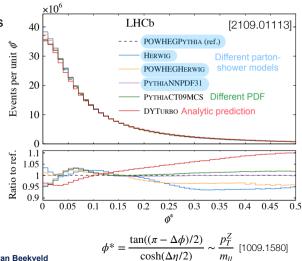
But differences matter...

Consider measurement of W boson mass

Measurements of p_T^Z in $Z/\gamma^* \rightarrow l^+l^-$ decays used to validate the MC predictions for p_T^W

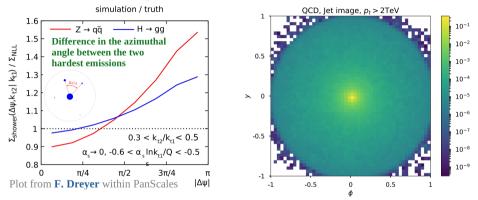
The envelope of shifts in m_W originating from differences in these shower predictions is the dominant theory uncertainty (11 MeV)

$$m_W = 80354 \pm 23_{\text{stat}} \pm 10_{\text{exp}} \pm 17_{\text{theory}} \pm 9_{\text{PDF}} \text{ MeV}$$



Melissa van Beekveld

Machine learning and jet sub-structure



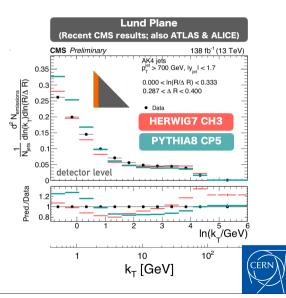
de Oliveira, Kagan, Mackey, Nachmann, Schwartzman [1511.05190]

Machine learning might learn un-physical "features" from MC \rightarrow can significantly impact the potential of new physics searches.



Lund Plane measurements

- Despite common showers doing an amazing job at the LHC, there are still places where big differences are seen
- In particular as we zoom into very differential phase space regions of jets, these differences can easily reach 10-30%
- The region shown here is particularly sensitive to soft emissions
- This is a region where some of the developments discussed later are relevant
- See also CMS [2312.16343]



selected collider-QCD accuracy milestones

	DCL	P splitting function							
	DGL	-		ons					
	LO	NLO				NNLO	[parts o	of N3LO	
		transverse-momentum resummation (DY&Higgs)							
		LL	NLL[]			NNLL[]	N3LL		
			parton she	owers	(many of	today's widely-used s	howers only LL@leading	J-colour)	
			LL	[parts	of NLL]		
					fixed-ord	der matching of	parton showers		
					LO	NLO	NNLO []		

selected collider-QCD accuracy milestones

					LO	NLO	NNLO []	[N3LO]
					fixed-ord	er matching of p	arton showers	
			LL	[parts	of NLL]	NLL NN
th	is t	alk	parton sh	owers	(many of t	oday's widely-used sh	owers only LL@leading	-colour) pa
		LL	NLL[]			NNLL[]	N3LL	
		trans	verse-mom	entum	resummat	ion (DY&Higgs)		
	LO	NLC)			NNLO	[parts o	of N3LO]
	DGL	AP spli	itting funct	ions				
0		NLO		NNLC)[]	N3LO	

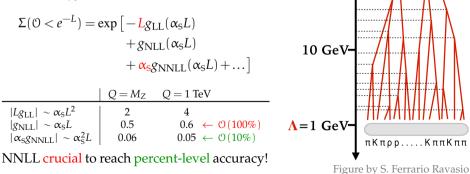
Why are we talking about logarithmic accuracy?

Q=1 TeV -

100 GeV-

Parton showers evolve hard states $Q \sim \sqrt{\hat{s}}$ down to the scale where hadronisation takes place $\Lambda \sim 1$ GeV

This evolution generates logarithms of the form $L \sim \ln \frac{Q}{\Lambda} \gg 1$, $(g_X(\alpha_S L) \sim \alpha_S L)$

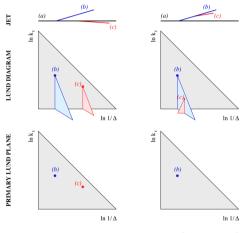


Current status of parton showers

- The most widely-used event generators at the LHC, Pythia, Herwig, and Sherpa, are all limited to LL (some exceptions where NLL can be reached, cf. Bewick, Ferrario Ravasio, Richardson, Seymour [1904.11866])
- Although there has been significant progress in improving the hard matrix elements of event generators with NNLO matching and NLO multi-jet merging, the logarithmic accuracy has been limited to LL for a very long time
- For this reason, there has been a concerted effort in taking parton showers from LL \rightarrow NLL in the last couple of years
- This has been achieved by several groups including PanScales [1805.09327], [2002.11114], [2011.10054], [2103.16526], [2111.01161], [2205.02237], [2207.09467], [2305.08645], [2312.13275], ALARIC Herren, Höche, Krauss, Reichelt, Schoenherr [2208.06057], [2404.14360], APOLLO Preuss [2403.19452], DEDUCTOR Nagy, Soper [2011.04773], and Forshaw-Holguin-Plätzer [2003.06400]
- $\rightarrow\,$ Very recently we have taken significant steps towards general NNLL (focus of this talk)



The Lund Plane

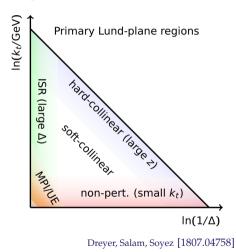


Dreyer, Salam, Soyez [1807.04758]

- To better make the connections between parton showers and their logarithmic accuracy we need to introduce the Lund Plane:
- Cluster the event with the Cambridge/Aachen algorithm, producing an angular ordered clustering sequence.
- Decluster the last clustering and record the transverse momentum and the opening angle of the declustering (plus other kinematics).
- Iterate along the hardest branch after each declustering to produce the primary Lund Plane.
- Following the softer branch produces the secondary, tertiary, etc Lund Plane.
- One can impose cuts easily on the declusterings (e.g. that they satisfy $z > z_{cut}$)



Logarithms in the Lund Plane



• The emission probability in the Lund Plane is then

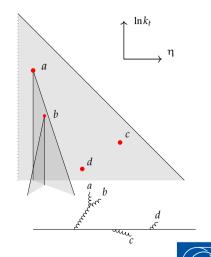
 $\mathrm{d}\rho \sim \alpha_{\mathrm{S}} \mathrm{d}\ln k_T \mathrm{d}\ln \theta$

- Hence emissions that are well-separated in both directions are associated with double logarithms of the form $\alpha_n^s L^{2n}$
- Emissions separated along one direction are associated with single logarithms of the form αⁿ_SLⁿ
- Emissions that are close in the Lund Plane are associated with a factor αⁿ_S
- We are now ready to state the PanScales NLL criteria for Parton Showers



NLL showers in a nutshell

- A necessary condition for a shower to be NLL is that it correctly describes configurations where all emissions are well-separated in a Lund plane Dasgupta, Dreyer, Hamilton, Monni, Salam [1805.09327]
- A core principle in this picture is that two emissions that are well-separated, should not influence each other (e.g. emission *d* cannot change the kinematics of *c*)^{*a*}.
- This principle is violated by most standard dipole-showers, due to the way the recoil is distributed after an emission First observed by Andersson, Gustafson, Sjogren '92
- For NLL 2-loop running coupling in the CMW scheme is also required
- For full NLL one also needs to include spin-correlations and sub-leading colour corrections



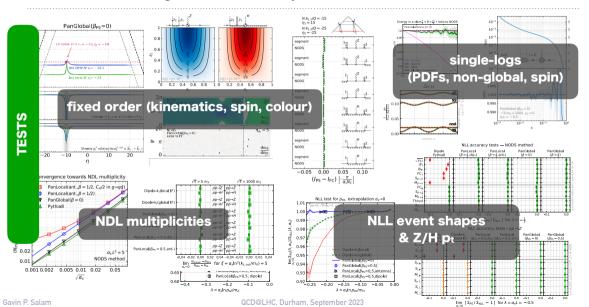


^{*a*}Spin-correlations are an exception in this context as they introduce long-range azimuthal correlations at NLL. Collinear spin understood in angular ordered showers for decades due to work of Collins '88 and Knowles '88. Extension to dipole showers studied in Richardson, Webster [1807.01955]. Both collinear and soft spin-correlations are included in PanScales at NLL.

PanLocal	PanGlobal	Colour	Spin
$k_t \sqrt{\theta}$ ordered	k_t or $k_t \sqrt{\theta}$ ordered	nested ordered double soft	for correct azimuthal
Recoil ⊥: local +: local –: local	Recoil ⊥: global +: local -: local	(NODS) Designed to ensure LL are	structure in collinear and soft→collinear
Dipole partition event CoM	Dipole partition event CoM	full colour (also gets many NLL at full colour)	[Collins-Knowles extended to soft sector]
[2002.11114]; pp (w/spin rario Ravasio, Salam, Soto-Ontoso,	Hamilton, Monni, Salam, Soyez h+colour): van Beekveld, Fer- Soyez, Verheyen [2205.02237]; + 09467]; DIS+VBF: van Beekveld,	Hamilton, Medves, Salam, Scyboz, Soyez [2011.10054]	AK, Salam, Scyboz, Verheyen [2103.16526], eid. + Hamilton [2111.01161]



a selection of the logarithmic accuracy tests



14

Oxford



Gavin Salam



Jack Helliwell



Silvia Zanoli

NIKHEF





Mrinal Dasgupta



Monash



Basem El-Menoufi



Ludo Scyboz



Gregory Soyez

CERN



AK



Silvia Ferrario Ravasio

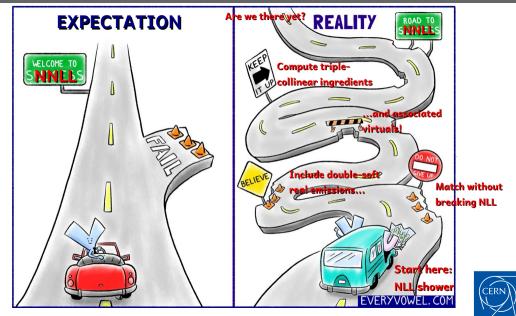




Pier Monni

Alba Soto-Ontoso

PanScales current members A project to bring logarithmic understanding and accuracy to parton showers



Analytic structure beyond NLL

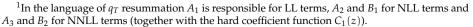
Taking an event shape, 0, to be less than some value $e^{-|L|}$ we have at NNLL (focusing for now on e^+e^- only)

$$\Sigma(\mathcal{O} < e^{-|L|}) = (1 + \alpha_{\rm s}C_1 + \dots) \exp\left[\frac{1}{\alpha_{\rm s}}g_1(\alpha_{\rm s}L) + g_2(\alpha_{\rm s}L) + \alpha_{\rm s}g_3(\alpha_{\rm s}L) + \dots\right]$$
(1)

where g_1 accounts for LL terms, g_2 for NLL terms, and g_3 and C_1 for NNLL terms¹. Whereas an analytic resummation in principle retains only the terms that are put in (i.e. g_1 and g_2 at NLL) the shower will instead generate spurious higher order terms

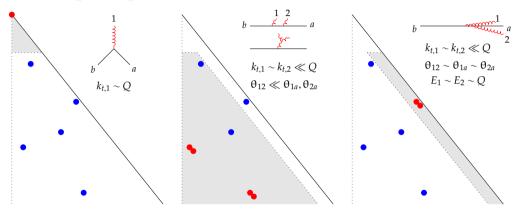
$$\Sigma(\mathcal{O} < e^{-|L|}) = \left(1 + \alpha_{\rm s} \tilde{\mathbf{C}}_1 + \dots\right) \exp\left[\frac{1}{\alpha_{\rm s}} g_1(\alpha_{\rm s} L) + g_2(\alpha_{\rm s} L) + \alpha_{\rm s} \tilde{\mathbf{g}}_3(\alpha_{\rm s} L) + \dots\right]$$
(2)

When thinking about going beyond NLL we need to address two things: 1) what are the necessary analytic ingredients from resummation and 2) how do we compensate the NNLL terms already present in the shower?



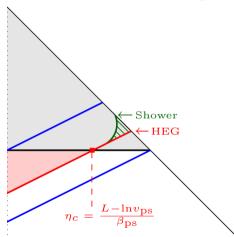


Lund plane picture



 $\begin{array}{ll} \mbox{hard matching} \rightarrow & \mbox{double-soft} \rightarrow & \mbox{triple-collinear} \rightarrow \\ \mbox{α_{s} correct for first emission } & \mbox{get any pair of soft commen-} & \mbox{account for genuine $2 \rightarrow 4$} \\ \mbox{-surate energy/angle right } & \mbox{collinear splittings} \end{array}$

Match without breaking NLL

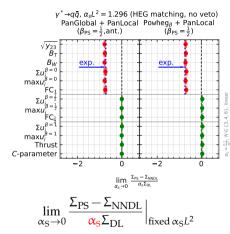


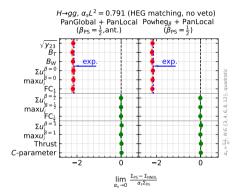
- We have so far explored the two-body decays γ → qq̄ and h → gg @ NLO
- For matching schemes that rely on the shower to generate the first emission (such as MC@NLO, KrkNLO, and MAcNLOPS) the matching works more or less out of the box.
- For POWHEG style matchings (including MiNNLO and GENEVA) log accuracy is more subtle to maintain.
- Main concern related to kinematic mismatch between shower and hardest emission generator (in general they are only guaranteed to agree in the soft-collinear region). This issue has been studied in the past Corke, Sjöstrand [1003.2384] but logarithmic understanding is new.



Hamilton, AK, Salam, Scyboz, Verheyen [2301.09645]

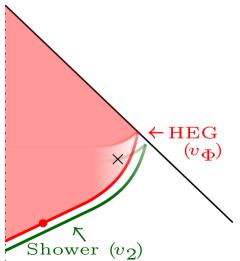
HEG without a veto is not NNDL ($\alpha_s^n L^{2n-2}$) accurate





Without a veto NLL accurate showers fail our NNDL ($\alpha_s^n L^{2n-2}$) event shape tests. The failure is O(1), and hence phenomenologically relevant. The dashed blue line indicates the analytically calculated expected value.

Further subtleties



- Even when the contours are fully aligned there are issues associated with how dipole showers partition the $g \rightarrow gg(q\bar{q})$ splitting function.
- In PanScales we use

$$\frac{1}{2!}P_{gg}^{\mathrm{asym}}(\zeta) = C_A \left[\frac{1+\zeta^3}{1-\zeta} + (2\zeta-1)w_{gg} \right],$$

such that $P_{gg}^{\text{asym}}(\zeta) + P_{gg}^{\text{asym}}(1-\zeta) = 2P_{gg}(\zeta)$

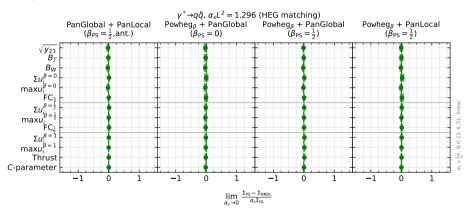
- This partitioning takes place to isolate the two soft divergences in the splitting function (ζ → 0 and ζ → 1), but there is some freedom in how one handles the non-singular part.
- The HEG needs to partition in exactly the same way. Not clear how easy this is in general, in particular in the soft-large angle region.



MILAN SEMINAR

Hamilton, AK, Salam, Scyboz, Verheyen [2301.09645]

Proper HEG achieves NNDL ($\alpha_s^n L^{2n-2}$) accuracy



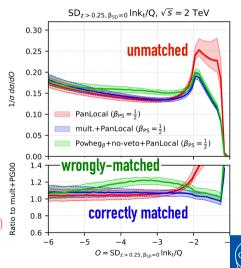
This can be achieved through a standard kinematic veto, as long as shower partioning matches the exact matrix element. A veto however complicates the inclusion of double-soft emissions, since it effectively alters the second emission, complicating the path to further logarithmic enhancement.



Phenomenological impact

- Contour mismatch by area αΔ leads to breaking of NLL and exponentiation
- Correct matching on the other hand augments the shower from NLL to NLL+NNDL for event shapes.
- Impact of NLL breaking terms vary for SoftDrop they have a big impact due to the single-logarithmic nature of the observable. In particular the breaking manifests as terms with super-leading logs

$$\partial_L \Sigma_{\rm SD}(L) = \bar{\alpha} c \, e^{\bar{\alpha} c L - \bar{\alpha} \Delta} - 2 \bar{\alpha} L e^{-\bar{\alpha} L^2} (1 - e^{-\bar{\alpha} \Delta})$$

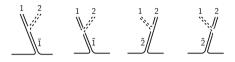


Include double-soft real emissions

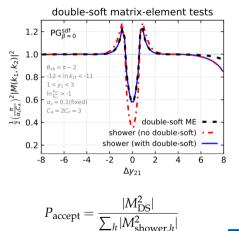
- NLO matching is a necessary ingredient for going beyond NLL, but to some extent NLO matching is a solved problem
- Until recently the inclusion of double-soft emissions in an NLL shower was still an open question
- To get them right we must ensure that any pair of soft emissions with commensurate energy and angles should be produced with the correct ME
- Any additional soft radiation off that pair must also come with the correct ME
- In addition must get the single-soft emission rate right at NLO (CMW-scheme)
- This should achieve NNDL accuracy for multiplicities, i.e. terms $\alpha_s^n L^{2n}$, $\alpha_s^n L^{2n-1}$ and $\alpha_s^n L^{2n-2}$
- and next-to-single-log (NSL) accuracy for non-global logarithms, for instance the energy in a rapidity slice, $\alpha_s^n L^n$ and $\alpha_s^n L^{n-1}$ (albeit only at leading- N_c for now)



The double-soft ME

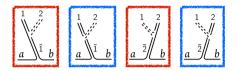


- For now we have focused on PanGlobal
- Any two-emission configuration in a dipole-shower comes with up to four histories (for PanLocal this would in fact be eight)
- We accept any such configuration with the true ME divided by the shower's effective double-soft ME summed over all histories that could have lead to that configuration.

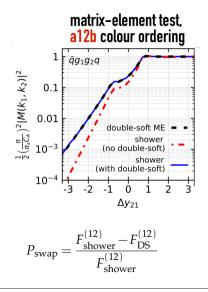




Correcting the colour-ordering



- We have two distinct colour orderings *a*12*b* and *a*21*b*
- We need to get the relative fractions *F*⁽¹²⁾ and *F*⁽²¹⁾ right in order to ensure that any further emissions are also correct.
- In practice we accept a colour ordering if the shower generates too little of it, and swap them if the shower generates too much (and similarly for $q\bar{q}$ vs gg branchings).



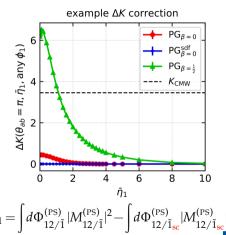


...and associated virtuals!

• The PanScales showers have correct soft emission intensity at NLO in the softcollinear (sc) region due to the use of the CMW-coupling

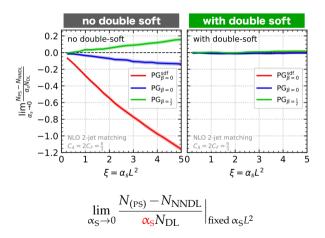
$$\alpha_{\rm s} \rightarrow \alpha_{\rm s} + \alpha_{\rm s}^2 K_1/2\pi$$

- This in general is not enough the get to soft wide-angle region right and we need to add a ΔK_1 which depends on the rapidity of the single soft emission
- This is related to the fact, that the shower $\bar{\eta}_1$ organises its phase space in such a way, that the rapidity of soft pair, y_{12} , does not $\Delta K_1 = \int d\Phi_{12/\tilde{1}}^{(PS)} |M_{12/\tilde{1}}^{(PS)}|^2 \int d\Phi_{12/\tilde{1}_{sc}}^{(PS)} |M_{12/\tilde{1}_{sc}}^{(PS)}|^2$.





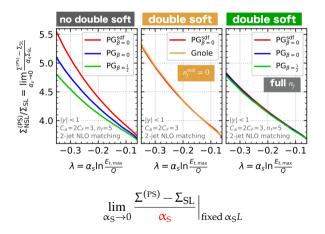
Lund Multiplicities at NNDL ($\alpha_s^n L^{2n-2}$)



- Reference NNDL analytic result from Medves, Soto-Ontoso, Soyez [2205.02861]
- We take α_s → 0 limit to isolate NNDL terms. This is significantly more challenging than at NDL due to presence of 1/α_s in denominator.
- Showers without double-soft corrections show clear differences from reference (and each other).
- Adding the double-soft corrections brings NNDL agreement.



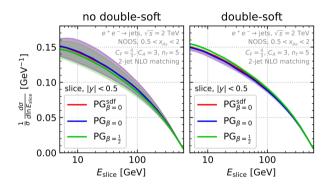
Energy in a slice at NSL ($\alpha_s^n L^{n-1}$)



- Reference NSL from Gnole Banfi, Dreyer, Monni [2111.02413] (see also Becher, Schalch, Xu [2307.02283]).
- We did this test semi-blind: only compared to Gnole after we had agreement between the three PanGlobal variants.
- We have NSL agreement with Gnole (using n_f^{real} = 0) and agreement between all showers with full-n_f dependence (first calculation of this kind as a by-product!)



What about pheno?



- We studied energy flow between two hard (1 TeV) jets as a preliminary pheno case
- The three PanGlobal variants are remarkably close without double-soft corrections, but have large uncertainties
- With double-soft corrections we see a small shift in central values but a significant reduction in uncertainties.



Compute triple-collinear ingredients

- Double-soft corrections are not by themselves enough to reach NNLL accuracy for event shapes. We also need triple-collinear ingredients (cf. Dasgupta, El-Menoufi [2109.07496], eid. + van Beekveld, Helliwell, Monni [2307.15734], eid. + AK [2402.05170] for work in this direction)
- However, it turns out that with the inclusion of real double-soft emissions, only the Sudakov form factor needs to be modified to reach NNLL for event shapes, i.e. we do not need the fullly differential triple-collinear structure
- Taking

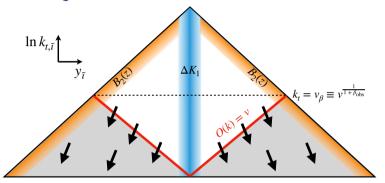
$$\alpha_{\text{eff}} = \alpha_{\text{s}} \left[1 + \frac{\alpha_{\text{s}}}{2\pi} \left(K_1 + \Delta K_1(y) + \frac{B_2(z)}{4\pi^2} \right) + \frac{\alpha_{\text{s}}^2}{4\pi^2} K_2 \right]$$

there are two pieces missing - B_2 which is of triple-collinear origin [2109.07496], [2307.15734] and K_2 (A_3) which is known Banfi, El-Menoufi, Monni [1807.11487], Catani, De Florian, Grazzini [1904.10365]

• NB: NLL showers generate spurious \tilde{B}_2 and $\tilde{K}_2 \rightarrow$ must be compensated



An intuitive picture



Imagine an emission, $\tilde{1}$, sitting anywhere right at the observable boundary (red line). The key observation is that whenever the shower splits $\tilde{1} \rightarrow 12$, the kinematic variables $(y_{12}, k_{t,12}, z_{12})$ of the resulting pair, do not agree with that of the parent $(y_{\bar{1}}, k_{t,\bar{1}}, z_{\bar{1}})$. Since the Sudakov was computed assuming conserved kinematics of $\tilde{1}$, and the observable is computed with the actual kinematics of (12), we have generated a mismatch. We can compute these drifts!

Relation between shower and resummation ingredients

It is fairly straightforward to see that at NNLL we only depend on ΔK_1 and B_2 through their respective integrals

$$\Delta K_1^{\text{int}} \equiv \int_{-\infty}^{\infty} dy \,\Delta K_1(y) \,, \, B_2^{\text{int}} \equiv \int_0^1 dz \frac{P_{gq}(z)}{2C_F} B_2(z).$$

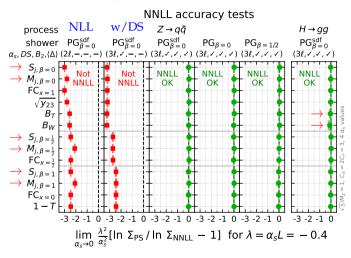
These (and K_2) can be related to the drifts in y ($\langle \Delta_y \rangle$), $\ln z$ ($\langle \Delta_{\ln z} \rangle$), and $\ln k_t$ ($\langle \Delta_{\ln k_t} \rangle$) and analytical resummation through

$$\Delta K_1^{\mathrm{int,PS}} = 2 \langle \Delta_y \rangle, \quad B_2^{\mathrm{int,PS}} = B_2^{\mathrm{int,NLO}} - \langle \Delta_{\ln z} \rangle, \quad K_2^{\mathrm{PS}} = K_2^{\mathrm{resum}} - 4\beta_0 \langle \Delta_{\ln k_t} \rangle.$$

Using these relations and taking $B_2^{\text{int,NLO}}$ from [2109.07496], [2307.15734] and K_2^{resum} from [1807.11487] one can prove that our showers are NNLL accurate for event-shape observables.



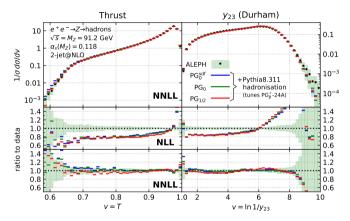
Are we there yet?



- →: New analytic results, not available in literature van Beekveld, Buonocore, El-Menoufi, Ferrario Ravasio, Monin, Soto-Ontoso, Soyez [in preparation]
- With no NNLL improvements, the coefficient of NNLL difference is significant, O(2 - 3), indicating importance of getting NNLL right
- With the inclusion of double-soft, observables with the same β_{obs} align but do still not agree with the analytics
- After inclusion of shifts and B₂ and K₂ we have perfect agreement



Not far now...



Long-standing tension between LEP data and Pythia8 unless using an anomalously large value of $\alpha_{\rm S}(M_Z) = 0.137$ Skands, Carrazza, Rojo [1404.5630] (also present for PanScales showers)

Inclusion of NNLL brings large corrections with respect to NLL

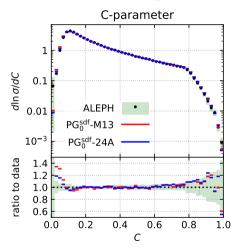
Agreement with data achieved without anomalously large value of $\alpha_{\rm S}$

Beware: no 3j@NLO which is known to be relevant in the hard regions

Residual uncertainties still need to be worked out



What about tuning?



Improved agreement with data across a large range of event shapes

Tuning here still rough

→ We start from the Monash tune (see ref. above) but fix $\alpha_{\rm S}(M_Z) =$ 0.118 (M13)

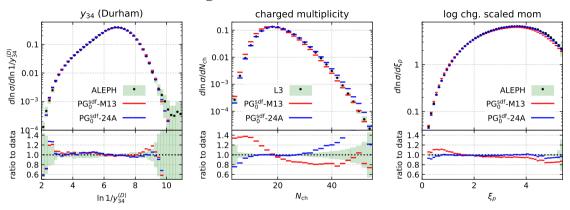
For our NLL showers this is the tune we use

For the NNLL showers we tune a number of parameters in the string model semi-automatically (24A)

Full tuning exercise still to be done!



What about tuning?



Impact of tune very minor on infrared safe observables, even those that are only NLL accurate

Impact on unsafe observables much larger, bringing good agreement with ALEPH data.



Conclusions and outlook

- As the experiments at the LHC record more and more data, it will become increasingly more important to improve on the accuracy of event generators
- NLL accurate showers have now been established by several groups
- First steps towards general NNLL accuracy was taken recently with the inclusion of double-soft corrections in the PanGlobal family of showers
- With these corrections we have reached NNDL accuracy for multiplicity and NSL accuracy for non-global observables
- The next natural step is to get NNLL right for event shapes
- This can be achieved using known ingredients from resummation together with an understanding of how the showers differ from analytic resummation through mainly recoil
- This we have achieved very recently
- The associated NNLL code has been made public in a the 0.2 release of the PanScales code
- Naturally we now are thinking about how to bring these advances to hadron-collisions
- For full general NNLL the shower needs to also correctly reproduce triple-collinear kinematics (e.g. for fragmentation functions)
- Work in that direction is also ongoing

