Anatomy of Hadron Collisions — and Future Challenges

Peter Z Skands — University of Oxford & Monash University — CERN, March 2024





Australian Government

Australian Research Council



net

The Goal

Use LHC measurements to test hypotheses about Nature

Problem 1: no exact solutions to QFT
→ Perturbative Approximations



The Goal

Use LHC measurements to test hypotheses about Nature

Problem 1: no exact solutions to QFT
→ Perturbative Approximations

> Elementary Fields, Symmetries, Interactions

CONFINEMENT



Problem 2: Confinement We collide — and observe — hadrons

Anatomy of LHC collisions



Hard Process & Fixed-Order Corrections



Infinite-Order Perturbative Corrections



Perturbative Approaches

P.T. ~ Calculate the area of a shape (d σ) with higher and higher detail Difference from exact area $\propto \alpha^{n+1}$



Note: (over)simplified analogy, mainly for IR structure. More at each order than shown here.

Perturbative Approaches

P.T. ~ Calculate the area of a shape (d σ) with higher and higher detail Difference from exact area $\propto \alpha^{n+1}$



Note: (over)simplified analogy, mainly for IR structure. More at each order than shown here.

Perturbation Theory as a Markov Chain

Stochastic differential evolution in "hardness" scale

 $d\sigma$ for generic observable "O", expressed as a Markov chain:

$$\frac{d\sigma}{dO} = \int d\Phi_0 ||M_{Born}|^2 (1 + F_{NLO} + ...) \underbrace{\mathcal{S}(\Phi_0, O)}_{Shower}$$
Fixed-Order Matching Coefficients Shower
$$\xrightarrow{\varphi_0} + 2 + 3 \text{ May 13}$$

$$\mathcal{S}_{+1}(\Phi_n, O) = \underbrace{\overset{Sudakov Factor}{\Delta(\Phi_n, Q_{IR})} \underbrace{\overset{Evaluate O \text{ on } \Phi_n}{\delta(\hat{O}(\Phi_n) - O)}}_{Branching Kernel}$$

$$+ \int d\Phi_{+1} \underbrace{\overset{Sudakov Factor}{\Delta(\Phi_n, Q_{n+1})} \underbrace{\frac{|M_{n+1}|^2}{|M_n|^2}}_{UNITARIY} \mathcal{S}(\Phi_{n+1}, O)}_{UNITARIY}$$
Kernel Ke

Why go beyond **Fixed-Order** perturbation theory?

Fixed-Order calculations most accurate for single-scale problems Effective accuracy reduced for processes/observables with scale hierarchies

Schematic example:

L

NNLO calculation of the rate of events passing a **jet veto**:

$$\begin{array}{c} \underset{F_{0}}{\overset{\text{LO}}{\longrightarrow}} & \underset{\alpha_{s}(L^{2}+L+F_{1})}{\overset{\text{NLO}}{\longrightarrow}} & \underset{\alpha_{s}^{2}(L^{4}+L^{3}+L^{2}+L+F_{2})}{\overset{\text{NNLO}}{\longrightarrow}} \\ L \propto \ln(p_{\perp \text{veto}^{2}}/Q_{\text{hard}}^{2}) & - \text{ arising from integrals over propagators } \propto \frac{\mathrm{d}p_{\perp}^{2}}{p_{\perp}^{2}}\mathrm{d}y \\ \hline \text{Total loss of predictivity for } p_{\perp \text{veto}} \ll Q_{\text{hard}} & \Longrightarrow \alpha_{s}L^{2} \sim 1. \\ \text{Reduced precision even for higher veto scales. Logs counteract naive suppression.} \end{array}$$

The Case for Embedding Fixed-Order Calculations within Showers



Showers		0	1	2	3	
	0	В	PS ₁	PS_2	PS ₃	
sdoo	1	B·∆ _B ¹	$PS_1 \cdot \Delta_1^1$	$PS_2 \cdot \Delta_2^1$	$PS_3 \cdot \Delta_3^1$	
Γo	2	B·∆ _B ²	$PS_1 \cdot \Delta_{1^2}$	$PS_2 \cdot \Delta_2^2$		

Resummation extends domain of validity of perturbative calculations

Showers ➤ Fully exclusive final states

→ can model non-perturbative physics, full-event analyses, fiducial cuts, ...

Target for next generation of MCs: %-level precision @ LHC ⇒ NNLO + NNLL

Not quite there yet — but close ...

Warmup: NLO + Shower with POWHEG

Nason 2004; Fixione, Nason, Oleari 2007

(Just focusing on the real-radiation part)

POWHEG generates hardest emission in a shower-like manner (MECs)

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

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

Sweeping over phase space, from high to low p_{T}

Shower then takes over and generates all further emissions



Powheg Box – A Subtlety

5

gf O Mismatched phase-space regions

Powheg Emission generated with (M⁽⁰⁾{X+1})²

-Phase Space

shower

covered by

Phase Space already Covered by Powheg

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.

E.g., Nason, Oleari <u>arXiv:1303.3922</u> VBF: <u>Höche et al., SciPost Phys. 12 (2022) 1</u>

2. From NLO to NNLO

MiNNLO_{PS} builds on (extends) POWHEG NLO for X + jet [Hamilton et al. 1212.4504.

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: What if we could $p_{\perp}^{\text{Shower}} \text{ vs } p_{\perp}^{\text{Powheg}} \text{ vs } Q_{NNLL}^{\text{resummation}}$ & differential NNLO spreading lift that restriction? (+ possible efficiency bottleneck: $p_{\perp} \rightarrow 0$ singularity \times Sudakov veto)

Monni et al. 1908.069871

 $\sigma_{X+3}(0)$

σ_{X+0}(0)

 $\sigma_{X+0}^{(1)}$

 $\sigma_{\chi + 0}^{(2)}$

0

2

Loops

 $\sigma_{X+1}^{(0)}$

 $\sigma_{X+1}^{(1)}$

 $\sigma_{X+1}(2)$

 $\sigma_{\chi+2}^{(0)}$

 $\sigma_{X+2}^{(1)}$

Legs

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

(arXiv:2108.07133 & arXiv:2310.18671)

QCD seminar

May 13

VINCIA

Continue shower afterwards

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



Need:

1 Born-Local NNLO ($\mathcal{O}(\alpha_s^2)$) K-factors: $k_{\text{NNLO}}(\Phi_2)$

2 NLO ($\mathcal{O}(\alpha_s^2)$) MECs in the first $2 \rightarrow 3$ shower emission: $k_{\text{NLO}}^{2\rightarrow 3}(\Phi_3)$

3 LO ($\mathcal{O}(\alpha_s^2)$) MECs for next (iterated) $2 \rightarrow 3$ shower emission: $k_{\text{LO}}^{3 \rightarrow 4}(\Phi_4)$

• Direct $2 \rightarrow 4$ branchings for unordered sector, with LO ($\mathcal{O}(\alpha_s^2)$) MECs: $k_{LO}^{2 \rightarrow 4}(\Phi_4)$

NNLO

Preview: VINCIA NNLO+PS for $H \rightarrow bb$



0.05



Part II – Nonperturbative Aspects



Hadronization ○ Hard Interaction Hard • Resonance Decays Process MECs, Matching & Merging OCD Final-State Radiation Parton QCD Initial-State Radiation* **Showers** Leeeeeeee Electroweak Radiation ALLER COLLER COLLER Underlying Multiparton Interactions Event Beam Remnants* $\mathrm{d}\hat{\sigma}_0$ $gg \rightarrow tt$ Strings Colour Reconnections Hadronization String Interactions Bose-Einstein & Fermi-Dirac

[Figure from arXiv:2203.11601]

MesonBaryon

Antibaryon

Heavy Flavour

(*: incoming lines are crossed)

Hadron Decays



New Discoveries in Hadronization

What a strange world we live in, said ALICE

Ratios of strange hadrons to pions strongly increase with event activity





Charm hadronization in pp (1):

More charm quarks in baryons in pothan in pothan in the signature of the s



<u>arXiv:2011.06079</u> <u>arXiv:2106.08278</u>

Charm quarks hadronize into baryons 40% of t

\sim 4 times more than in e+e-

(Wi

ll come ba gkI c to these)	$f(\mathbf{c} \rightarrow \mathbf{H}_{\mathbf{c}})[\%]$
\mathbf{D}^0	$39.1 \pm 1.7 (stat)^{+2.5}_{-3.7} (syst)$
\mathbf{D}^+	$17.3 \pm 1.8(\text{stat})^{+1.7}_{-2.1}(\text{syst})$
$\mathrm{D}^+_{\mathrm{s}}$	$7.3 \pm 1.0 (stat)^{+1.9}_{-1.1} (syst)$
Λ_{c}^+	$20.4 \pm 1.3 (stat)^{+1.6}_{-2.2} (syst)$
$\Xi_{\rm c}^0$	$8.0\!\pm\!1.2(\text{stat})^{+2.5}_{-2.4}(\text{syst})$
D^{*+}	$15.5 \pm 1.2(\text{stat})^{+4.1}_{-1.9}(\text{syst})$

Back to Basics – Anatomy of (Linear) Confinement

On lattice, compute potential energy of a colour-singlet $q\bar{q}$ state, as function of the distance, r, between the q and \bar{q}



From Partons to Strings



Gluon = kink on string, carrying energy and momentum

Physics then in terms of string worldsheet evolving in spacetime

"Nambu-Goto action" \implies Area Law.

String Breaking

Non-perturbative $g \rightarrow q\bar{q}$ \implies The strings will "break" Non-perturbative so can't use $P_{g \rightarrow q\bar{q}}(z)$ Our model: Schwinger mechanism Assume const probability per unit worldsheet area:

Meson

time

spatial

separation



J. Schwinger, Phys. Rev. 82 (1951) 664

 $\implies \text{Suppression of } m_s^2/\kappa \text{ relative to } m_{u,d}^2/\kappa \\ (+ \text{ occasionally get a "diquark" too } \text{baryons}) \\ \implies \text{universal (constant) ratios (for constant } m, \kappa)}$

 \vec{g}

Beyond the Static Limit

Regard tension κ as an emergent quantity?

Not fundamental strings

May depend on (invariant) time τ

E.g., hot strings which cool down Hunt-Smith & **PZS** EPJC 80 (2020) 11

May depend on σ (excitations)

Working with E. Carragher & J. March-Russell in Oxford.

May depend on environment (e.g., other strings nearby)

Two approaches (so far) within Lund string-model context:

Colour Ropes [Bierlich, Gustafson, Lönnblad, Tarasov JHEP 03 (2015) 148; + more recent...]

Close-Packing [Fischer & Sjöstrand JHEP 01 (2017) 140; Altmann & PZS in progress ...]





Non-Linear String Dynamics

$MPI \Longrightarrow lots$ of coloured partons scattered into the final states

Count **# of (oriented) flux lines** crossing y = 0 in pp collisions (according to PYTHIA) And classify by SD(3)^PMultiplet:



LE

What about Baryon Number?

Types of string topologies:



Sjöstrand & PZS NPB 659 (2003) 243

String Formation Beyond Leading Colour Christiansen & **PZS** *JHEP* 08 **(2015)** 003

String Junctions at LHC?



Outlook

New insights into perturbation theory at non-trivial orders

NNLO for many hard processes (and N3LO for simple ones)

- Several recent showers achieve NLL for arbitrary (IR safe) observables (e.g., PanScales, Alaric)
 - (& NNLL accuracy not too far, possibly already achieved in evolution variable)
 - (Off-the-shelf coherent ones: at most NLL (?) in observables ~ evolution variable)
- + New ways to combine them (e.g., MiNNLO_{PS}, VinciaNNLO, Geneva)
- → New Paradigm for Perturbative Calculations: NNLO + NNLL matched MCs Expect shift from educated guesses to %-level precision (+ theoretically elegant)

New measurements have challenged naive ideas of hadronization

Appears certain we are seeing effects **beyond the static** $q\bar{q}$ **limit** But *is it* string interactions / junctions? Is it thermal / QGP? Coalescence? ... Last shot hasn't been fired ...

Extra Slides

Parton Showers: Theory

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

Most bremsstrahlung is

driven by divergent propagators → simple structure

Mathematically, gauge amplitudes factorize in singular limits



Partons ab

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

P(z) =**DGLAP splitting kernels**", with $z = E_a/(E_a + E_b)$

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.)

What does it mean that two partons are "colour connected"?

Between which partons should confining potentials form?

E.g., if we have events with lots of quarks and gluons



Complication:

Every quark-gluon vertex contains an SU(3) Gell-Mann matrix in colour space!

(And $g \rightarrow gg$ vertices contain further complicated structures)

> Who ends up confined with whom?



Colour Flow in Event Generators

Event Generators use simplified "colour flow" — to trace colour correlations through hard processes & showers ➤ determine which partons end up "colour connected"

Based on SU(N) group product: $N \otimes \overline{N} = (N^2 - 1) \oplus 1$

 Fundamental representation (quarks)
 Image: Antifundamental representation (antiquarks)
 Image: Adjoint Representation (gluons)

 Antifundamental representation (antiquarks)
 Image: Adjoint Representation (gluons)

Thus, for large N ("leading colour"), we can approximate $(N^2 - 1) \sim N \otimes \overline{N}$

LC: gluons \rightarrow direct products of colour and anticolour; for SU(3) this is valid to ~ $1/N_C^2$ ~ 10% \Rightarrow Rules for colour flow (= colour-space vertices) in MC Event Generators:



(Note: the "colour dipoles" in dipole and antenna showers are also based on these rules)



A corresponding event record from PYTHIA, up to the second gluon emission

#	id	name	status	mothers	daughters	colours	p_x	p_y	p_z	е	m
5	23	(Z0)	-22	3 4	67		0.000	0.000	0.000	91.188	91.188
6	3	(s)	-23	5 0	10 0	101 0	-12.368	16.523	40.655	45.594	0.000
7	-3	(sbar)	-23	5 0	89	0 101	12.368	-16.523	-40.655	45.594	0.000
8	21	(g)	-51	7 0	13 0	103 101	9.243	-9.146	-29.531	32.267	0.000
9	-3	sbar	51	7 0		0 103	3.084	-7.261	-10.973	13.514	0.000
10	3	(s)	-52	6 0	11 12	101 0	-12.327	16.406	40.505	45.406	0.000
11	21	g	-51	10 0		101 102	-2.834	-2.408	1.078	3.872	0.000
12	3	S	51	10 0		102 0	-10.246	17.034	38.106	42.979	0.000
13	21	g	52	8 0		103 101	9.996	-7.366	-28.211	30.823	0.000

Colour Reconnections? (CR)



With a probability of 1/9, both options should be possible (remaining 8/9 allow LC only)

Choose "lowest-energy" one (cf action principle) (assuming genuine quantum superpositions to be rare.)

 \rightarrow small shift in W mass ("string drag") (\rightarrow now important for top quark mass at LHC)

LEP-2: No-CR excluded at 99.5% CL [Phys.Rept. 532 (2013) 119; arXiv:1302.3415]

Measurements consistent with $\sim 1/N_C^2$ expectation but not much detailed information.

The String Fragmentation Function (in momentum space)

Consider a string break \Leftrightarrow , producing a meson M, and a leftover string piece The meson M takes a fraction z of the quark momentum,

Probability distribution in $z \in [0,1]$ parametrised by **Fragmentation Function**, $f(z, Q_{HAD}^2)$



Iterative String Breaks (in momentum space)

Recall: String breaks are causally disconnected \rightarrow May iterate from outside-in

Note: using light-cone momentum coordinates: $p_{+} = E + p_{z}$



On average, expect energy* of nth "rank" hadron to scale like ~

$$E_n \sim \langle z \rangle (1 - \langle z \rangle)^{n-1} E_0$$

*) more correctly, the p+ light-cone momentum coordinate

A simple prediction: constant rapidity density of hadrons along string

Rapidity $y = \frac{1}{2} \ln \left(\frac{E + p_z}{E - p_z} \right) = \frac{1}{2} \ln \left(\frac{(E + p_z)^2}{E^2 - p_z^2} \right) \rightarrow \ln \left(\frac{2E}{m_\perp} \right) \quad \text{(in limit of small } m_\perp = \sqrt{m^2 + p_\perp^2}$ $\ll E$

Recall: expect energy of nth "rank" hadron $E_n \sim \langle z \rangle (1 - \langle z \rangle)^{n-1} E_0$ $\implies y_n \sim y_1 + (n-1) \ln(1 - \langle z \rangle)$

Rapidity difference between two adjacent hadrons:

 $\Delta y = y_{n+1} - y_n \sim \ln(1 - \langle z \rangle) \quad \leftarrow \text{Constant, independent of } n \text{ (and of } E_0\text{)}$

Predicts a flat (uniform) rapidity "plateau" (along the string axis): Also called **"Lightcone scaling";** this is exactly what is observed in practice.

The Rapidity Plateau

Expect ~ flat Rapidity Plateau along string axis

Estimate of rapidity range for fixed E_q :

 $\langle y \rangle_1 \sim \ln\left(\frac{2\langle z \rangle E_q}{\langle m_\perp \rangle}\right)$ ~ 5 for $E_q \sim 100 \text{ GeV}, \langle z \rangle \sim 0.5$, and $\langle m_\perp \rangle \sim 0.5 \text{ GeV}$

Changing $E_q \implies$ logarithmic change in rapidity range:



 $\langle n_{\rm Ch} \rangle \approx c_0 + c_1 \ln E_{\rm Cm}$, \sim Poissonian multiplicity distribution



The Rapidity Plateau

Expect ~ flat Rapidity Plateau along string axis

Estimate of rapidity range for fixed E_q :

$$\langle y \rangle_1 \sim \ln\left(\frac{2\langle z \rangle E_q}{\langle m_\perp \rangle}\right)$$

~ 5 for $E_q \sim 100 \text{ GeV}$, ~ 0.5, and $\langle m_\perp \rangle \sim 0.5 \text{ GeV}$

Changing $E_q \implies$ logarithmic change in rapidity range:



 $\langle n_{\rm Ch} \rangle \approx c_{\rm 0} + c_{\rm 1} \ln E_{\rm Cm}, \sim$ Poissonian multiplicity distribution



Actual difference is smaller ${\color{black}\rightarrowtail}\sim 0.5$

(some energy also goes to increase particle production in the central region, **3-jet events**)

Gluon Kinks: The Signature Feature of the Lund Model



Gluons are connected to two string pieces



What do String Junctions do?

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

► No new string-fragmentation parameters



