## Anatomy of Hadron Collisions - and Future Challenges

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

 SOCIETY


## The Goal

## Use LHC measurements to test hypotheses about Nature

Problem 1: no exact solutions to QFT
$\rightarrow$ Perturbative Approximations


Elementary Fields, Symmetries, Interactions


Problem 2: Confinement
We collide - and observe - hadrons

## The Goal

## Use LHC measurements to test hypotheses about Nature

Problem 1: no exact solutions to QFT
$\rightarrow$ Perturbative Approximations



Problem 2: Confinement
We collide - and observe - hadrons

## Anatomy of LHC collisions

## Physics

Maths
Separation of scales
Factorizations


## Hard Process \& Fixed-Order Corrections



## Infinite-Order Perturbative Corrections



## Perturbative Approaches

## P.T. ~ Calculate the area of a shape ( $\mathrm{d} \sigma$ ) with higher and higher detail

Difference from exact area $\propto \alpha^{n+1}$


Example: Koch Snowflake


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 ( $\mathrm{d} \sigma$ ) with higher and higher detail

Difference from exact area $\propto \alpha^{n+1}$


Example: Koch Snowflake





Massless gauge theories
Scale invariance $\rightarrow$ fractal substructure (+ not hard to build in running coupling)

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

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


## 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:

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

$L \propto \ln \left(p_{\text {土eto }^{2}} / Q_{\text {hard }}^{2}\right)$ - arising from integrals over propagators $\propto \frac{\mathrm{d} p_{\perp}^{2}}{p_{\perp}^{2}} \mathrm{~d} y$
Total loss of predictivity for $p_{\perp \text { veto }} \ll Q_{\text {hard }} \Longrightarrow \alpha_{s} L^{2} \sim 1$.
Reduced precision even for higher veto scales. Logs counteract naive suppression.

The Case for Embedding Fixed-Order Calculations within Showers


## Warmup: NLO + Shower with POWHEG

(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 $\mathrm{PT}_{T}$

Shower then takes over and generates all further emissions


Pseudorapidity of the emitted parton

## Powheg Box - A Subtlety

## Industry Standard: "Powheg Box"

Exploits having its own definition of " $\mathrm{PT}^{\prime \prime}$ $\neq$ shower's definition of $\mathrm{p}_{\mathrm{T}}$
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 arXiv:1303.3922

VBF: Höche et al., SciPost Phys. 12 (2022) 1

## 2. From NLO to NNLO

$\mathrm{MiNNLO}_{\text {ps }}$ 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 $\mathrm{d} \sigma_{X}$ to NNLO $\mathcal{O}\left(\alpha_{s}^{3}\right)$ 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:
$p_{\perp}^{\text {Shower } v s} p_{\perp}^{\text {Powheg }}$ vs $Q_{N N L L}^{\text {resummation }} \&$ differential NNLO spreading
(+ possible efficiency bottleneck: $p_{\perp} \rightarrow 0$ singularity $\times$ Sudakov veto)

What if we could lift that restriction?

## 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

## (arXiv:2108.07133 \& arXiv:2310.18671)

$\Rightarrow$ expect small matching systematics


Need:

(1) Born-Local NNLO $\left(\mathcal{O}\left(\alpha_{s}^{2}\right)\right)$ K-factors: $k_{\mathrm{NNLO}}\left(\Phi_{2}\right)$
(2) NLO $\left(\mathcal{O}\left(\alpha_{s}^{2}\right)\right)$ MECs in the first $2 \rightarrow 3$ shower emission: $k_{\mathrm{NLO}}^{2 \rightarrow 3}\left(\Phi_{3}\right)$
(3) LO $\left(\mathcal{O}\left(\alpha_{s}^{2}\right)\right)$ MECs for next (iterated) $2 \rightarrow 3$ shower emission: $k_{\text {LO }}^{3 \rightarrow 4}\left(\Phi_{4}\right)$


4 Direct $2 \rightarrow 4$ branchings for unordered sector, with LO $\left(\mathcal{O}\left(\alpha_{s}^{2}\right)\right) \mathrm{MECs}: k_{\mathrm{LO}}^{2 \rightarrow 4}\left(\Phi_{4}\right)$

## Preview: VINCIA NNLO+PS for $H \rightarrow b \bar{b}$

"NNLO" Reference $=$ EERAD3 NLO $H \rightarrow b \bar{b} g$
Coloretti, Gehrmann-de Ridder, Preuss, JHEP 06 (2022) 009
NNLO accuracy in $H \rightarrow 2 j$ implies NLO correction in first emission and LO correction in second emission.


So for Thrust, NNLO $H \rightarrow b \bar{b}$ is effectively NLO for $\tau<1 / 3$
LO for $\tau>1 / 3$

VINCIA NNLO+PS: shower as phase-space generator: efficient \& no negative weights!
> Looks ~ $5 \times$ faster than EERAD3 (for equivalent unweighted stats)

+ is matched to shower (add shower resummation without auxiliary input/scales) + can be hadronized
Proof of concepts now done for $Z / H \rightarrow q \bar{q}$; work remains for $p p$ (\& for NnLL accuracy)


## Part II - Nonperturbative Aspects



## Hadronization



## 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


## New Discoveries in Hadronization

## LHC experiments also report very large (factor-10) enhancements in heavyflavour baryon-to-meson ratios at low $\mathrm{p}_{\mathrm{T}}$ !

> Conventional models (eg PYTHIA Monash) $\rightarrow$ constant baryon-to-meson ratio

(Just showing $\Lambda_{c}^{+}$here; same pattern for other heavy-flavour baryons \& also seen by LHCb)


Figure from Altmann \& PZS, String Junctions Revisited, in progress

## 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 $\bar{q}$

"Cornell Potential" fit: $V(r)=-\frac{a}{r}+\kappa r \quad$ with $\kappa \sim 1 \mathrm{GeV} / \mathrm{fm}$

## From Partons to Strings

## Map:



Gluon $=$ kink on string, carrying energy and momentum

Physics then in terms of string worldsheet evolving in spacetime
"Nambu-Goto action" $\Longrightarrow$ Area Law.

## String Breaking

Non-perturbative $g \rightarrow q \bar{q}$

$\Longrightarrow$ 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:


## Schwinger Effect

Non-perturbative creation of $\mathrm{e}^{+} \mathrm{e}^{-}$pairs in a strong external Electric field

Probability from Tunneling Factor

$$
\mathcal{P} \propto \exp \left(\frac{-m^{2}-p_{\perp}^{2}}{\kappa / \pi}\right)
$$

( $\kappa$ is the string tension equivalent)
$\Longrightarrow$ Suppression of $m_{s}^{2} / \kappa$ relative to $m_{u, d}^{2} / \kappa$ (+ occasionally get a "diquark" too $\rightarrow$ baryons)
$\Longrightarrow$ universal (constant) ratios (for constant $m, k$ )

## Beyond the Static Limit

## Regard tension $\kappa$ as an emergent quantity? Not fundamental strings

Cyclonic and Anticyclonic Winds

May depend on (invariant) time $\tau$
E.g., hot strings which cool down Hunt-Smith \& PZS EPJC 80 (2020) 11

## May depend on $\sigma$ (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 SU(3) multiplet:


Confining fields may be reaching higher effective representations than simple $\mathrm{q} \overline{\mathrm{q}}(3)$ ones.

$\rightarrow$ Is "emergent tension" driving strangeness enhancement in pp?

Altmann \& PZS work in progress ...

## What about Baryon Number?

Types of string topologies:

## Open Strings


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

## SU(3) String Junction

Closed Strings


Gluon rings
E.g., $H \rightarrow g g+$ shower Open strings with $N_{C}=3$ endpoints

$$
\Upsilon \rightarrow g g g+\text { shower }
$$


E.g., Baryon-Number violating neutralino decay $\tilde{\chi}^{0} \rightarrow q q q+$ shower

## String Junctions at LHC?

Stochastic sampling of $\operatorname{SU}(3)$ group probabilities (e.g., $3 \otimes 8=15 \oplus 6 \oplus 3$ )
$\Longrightarrow$ Random (re)connections in colour space (weighted by group weights)

Christiansen \& PZS 2015
For example:

Illustration by J. Altmann


Sjöstrand \& PZS 2002; Altmann \& PZS 2024

 restrictions ( $0=$ none)

## 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., MiNNLOps, VinciaNNLO, Geneva)
$\rightarrow$ 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

## Most bremsstrahlung is

driven by divergent propagators $\rightarrow$ simple structure

Mathematically, gauge amplitudes factorize in singular limits

$$
\left.\begin{array}{l}
\underset{\rightarrow \text { Partons ab }}{\rightarrow \text { collinear: }}\left|\mathcal{M}_{F+1}(\ldots, a, b, \ldots)\right|^{2} \xrightarrow{a \| b} g_{s}^{2} \mathcal{C} \frac{P(z)}{2\left(p_{a} \cdot p_{b}\right)}\left|\mathcal{M}_{F}(\ldots, a+b, \ldots)\right|^{2} \\
\qquad P(z)=\text { DGLAP splitting kernels", with } z=E_{a} /\left(E_{a}+E_{b}\right) \\
\underset{\rightarrow \text { soft: }}{\text { Gluon } \mathrm{j}} \mid
\end{array}\left|\mathcal{M}_{F+1}(\ldots, i, j, k \ldots)\right|^{2} \xrightarrow{j_{g} \rightarrow 0} g_{s}^{2} \mathcal{C} \frac{\left(p_{i} \cdot p_{k}\right)}{\left(p_{i} \cdot p_{j}\right)\left(p_{j} \cdot p_{k}\right)}\left|\mathcal{M}_{F}(\ldots, i, k, \ldots)\right|^{2}\right)
$$

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 $g \rightarrow g g$ vertices contain further complicated structures)

- Who ends up confined with whom?


## Colour Tracing

## 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: $N \otimes \bar{N}=\left(N^{2}-1\right) \oplus 1$
Fundamental representation (quarks) $\downarrow$
Antifundamental representation (antiquarks) $\downarrow \quad \uparrow \quad \begin{gathered}\text { Singlet (becomes irrelevant for large } N \text { ) }\end{gathered}$
Thus, for large $N$ ("leading colour"), we can approximate ( $N^{2}-1$ ) $\sim N \otimes \bar{N}$
LC: gluons $\rightarrow$ direct products of colour and anticolour; for SU(3) this is valid to $\sim 1 / N_{C}^{2} \sim 10 \%$ $\Rightarrow$ Rules for colour flow (= colour-space vertices) in MC Event Generators:


$$
g \rightarrow g g
$$

$$
\begin{array}{cc}
9 g \\
\text { cerer }
\end{array} \rightarrow
$$

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

## LC Colour Flow in an ee Collision

MCs: $N_{C} \rightarrow \infty$ limit formalised by letting each "colour line" be represented by a unique Les Houches colour tag ${ }^{\dagger}$ (no interference between different colour lines in this limit)
t: hep-ph/0109068; hep-ph/0609017


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

| $\boldsymbol{\#}$ | id | name | status | mothers | daughters | colours | p_x | p_y | p_z | e | m |  |  |  |
| :--- | :---: | :--- | :--- | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 5 | 23 | (ZO) | -22 | 3 | 4 | 6 | 7 |  | 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 | 8 | 9 | 0 | 101 | 12.368 | -16.523 | -40.655 | 45.594 | 0.000 |
| $\mathbf{8}$ | 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)

## Consider two (uncorrelated) parton systems

NB: much more important in LHC collisions $\rightarrow$ next lecture

Textbook example: $e^{+} e^{-} \rightarrow W^{+} W^{-} \rightarrow$ hadrons


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 $\Uparrow$, 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, $f\left(z, Q_{\mathrm{HAD}}^{2}\right)$


## 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 $\mathrm{n}^{\text {th }}$ "rank" hadron to scale like ~

$$
E_{n} \sim\langle z\rangle(1-\langle z\rangle)^{n-1} E_{0}
$$

[^0]A simple prediction: constant rapidity density of hadrons along string

## Rapidity

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

Recall: expect energy of $\mathbf{n}^{\text {th }}$ "rank" hadron $E_{n} \sim\langle z\rangle(1-\langle z\rangle)^{n-1} E_{0}$

$$
\Longrightarrow y_{n} \sim y_{1}+(n-1) \ln (1-\langle z\rangle)
$$

## Rapidity difference between two adjacent hadrons:

$$
\Delta y=y_{n+1}-y_{n} \sim \ln (1-\langle z\rangle) \quad \longleftarrow \text { Constant, independent of } n\left(\text { and of } E_{0}\right)
$$

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 $E_{q}$ :
$\langle y\rangle_{1} \sim \ln \left(\frac{2\langle z\rangle E_{q}}{\left\langle m_{\perp}\right\rangle}\right)$
$\sim 5$ for $E_{q} \sim 100 \mathrm{GeV},<z>\sim 0.5$, and $\left\langle m_{\perp}\right\rangle \sim 0.5 \mathrm{GeV}$
Changing $E_{q} \Longrightarrow$ logarithmic change in rapidity range:


$$
\left\langle n_{\mathrm{ch}}\right\rangle \approx c_{0}+c_{1} \ln E_{\mathrm{cm}}, \sim \text { Poissonian multiplicity distribution }
$$



## The Rapidity Plateau

## Expect ~ flat Rapidity Plateau along string axis

Estimate of rapidity range for fixed $E_{q}$ :

$$
\langle y\rangle_{1} \sim \ln \left(\frac{2\langle z\rangle E_{q}}{\left\langle m_{\perp}\right\rangle}\right)
$$

$\sim 5$ for $\left.E_{q} \sim 100 \mathrm{GeV},<\mathrm{z}\right\rangle \sim 0.5$, and $\left\langle m_{\perp}\right\rangle \sim 0.5 \mathrm{GeV}$

Changing $E_{q} \Longrightarrow$ logarithmic change in rapidity range:


[^1]

Actual difference is smaller $\longmapsto \sim 0.5$
(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

vs composition of background and bremsstrahlung combinatorics

See also
Larkoski et al., JHEP 1411 (2014) 129
Thaler et al., Les Houches, arXiv:1605.04692

## What do String Junctions do?

Assume Junction Strings have same properties as ordinary ones (u:d:s, Schwinger $\mathrm{p}_{\mathrm{T}}, \ldots$. .
> No new string-fragmentation parameters


## LHCb: also in Bottom

$\Lambda_{b}$ asymmetry
LHCb, JHEP 10 (2021) 060 • arXiv: 2107.09593

$$
A=\frac{\sigma\left(\Lambda_{\mathrm{b}}^{0}\right)-\sigma\left(\bar{\Lambda}_{\mathrm{b}}^{0}\right)}{\sigma\left(\Lambda_{\mathrm{b}}^{0}\right)+\sigma\left(\bar{\Lambda}_{\mathrm{b}}^{0}\right)}
$$

Baseline Expectations: \& ■ $b$ quark combines with the proton beam remnant $\Longrightarrow \Lambda_{b}$ production Not possible for $\bar{\Lambda}_{b}$ (no $\bar{p}$ remnant at LHC)


QCD CR with "string junctions" [Chistiansen \& Sknad JHEP 08 (2015) 003]
Adds large amount of low-pT $\Lambda_{b}$ and $\bar{\Lambda}_{b}$, in equal amounts. Dilutes asymmetry!


[^0]:    ${ }^{*}$ ) more correctly, the $\mathrm{p}_{+}$light-cone momentum coordinate

[^1]:    $\left\langle n_{\mathrm{ch}}\right\rangle \approx c_{0}+c_{1} \ln E_{\mathrm{cm}}, \sim$ Poissonian multiplicity distribution

