## Taming the complex dynamics of scattering events

Pier Monni (CERN)

HET Seminar - BNL, June 2022

## Particle physics at the Large Hadron Collider (LHC)

- LHC about to resume operations:
- Huge boost in experimental precision foreseen (only $\sim 5 \%$ of the total luminosity delivered so far)
- Key open questions to be addressed:
- Establish the Higgs sector
- Broad searches for New Physics (NP)
- Stress test of the Standard Model (SM)




## Broad spectrum searches for NP signatures

- Detailed scan of accessible regions parameter space
- e.g. global EFT fits, dedicated searches \& specific NP models
- test of consistency structure of the theory (op. mixing and correlations)
[Ellis, Madigan, Mimasu, Sanz, You '21]




## Main challenge: controlling the fine structure of collider events



## Vast technological progress (jointly Theory $\otimes$ Experiment)



## E.g. Impressive progress in theoretical calculations

## forma dereopments


event generators

## E.g. Impressive progress in theoretical calculations



## This talk focuses on another crucial aspect: Event Generators

## forma deredoments

andscape of NP modes


## Anatomy of a scattering reaction at the LHC



## Anatomy of a scattering reaction at the LHC

- Short distance (hard)
- scales probed: $\mathrm{O}\left(10^{2}\right)-\mathrm{O}\left(10^{3}\right) \mathrm{GeV}$
- stage sensitive to NP

evolution towards a physical observable state (mainly QCD)

- Long distance (soft)
- transition from $\mathrm{O}\left(10^{2}\right)-\mathrm{O}\left(10^{3}\right) \mathrm{GeV}$ to $\mathrm{O}(1) \mathrm{GeV}$
- hard scattering gets "showered" with soft [and/or collinear] radiation
- Output: what is actually measured


## Event generators simulate all stages of the event formation



- Not a standard theory calculation:
- return events, i.e. particle momenta with a physical probability distribution
- allow the computation of many (~any) observables at once, as opposed to a few of them in perturbative calculations
- deeply different mathematical formulation, difficult to exploit state of the art QFT technology
- Crucial pillar of modern collider physics, e.g. full simulation of experimental analysis, phase-space extrapolation, training of tools (e.g. Machine Learning)


## Strength: Back bone of nearly all LHC analyses

[ATLAS '22]


## Weakness: Inaccuracies are now often the leading systematics

- The improving experimental performance highlights limitations of event generators
- Soon to be the bottleneck of LHC physics programme
- Jet Energy Scale uncertainty ( $\rightarrow$ affecting many measurements)
- ... this is but one example



## Robust training of Machine Learning (ML) algorithms

- ML technology provides a great boost in sensitivity w.r.t. orthodox analysis techniques
- However, this comes often with a dependence on the modelling, i.e. Monte Carlo generator, raising the question of accuracy
- e.g. dependence of 4-pronged tagger on training model \& pseudo-data
- New generation of tools paramount to push this technology in the precision era of LHC



## Extrapolation of experimental measurements

- MC generators used to extrapolate experimental data from fiducial to inclusive phase space (easy comparison with theory and interpretation)
- Inaccuracies may lead to dangerous biases
- e.g. discrepancy in $t \bar{t}$ spin-correlations: new physics or mis-modelling ? (more later)


Parton level $\Delta \phi\left(\Gamma^{+}, \Gamma\right) / \pi[\mathrm{rad} / \pi]$

## The overarching question: Can we do better?

## How do we even define the accuracy of event generators?



- Evolution spans several orders of magnitude in energy scale
- Different perturbation theories needed in different regimes (e.g. fixed-order, logarithmic power counting, subleading power corr.ns)
- We should demand that event generators reproduce these limits correctly
- This talk addresses the two main elements: the hard scattering \& the parton shower


## The parton shower stage

## The parton shower component



- Large hierarchy of scales ( $\mu_{\text {hard }} \gg \mu_{\text {soft }}$ )
- Yet, fully perturbative regime ( $\mu_{\text {soft }} \gg \Lambda_{\mathrm{QCD}}$ )
- Initial conditions for hadronisation
- Several successful public tools:


Herwig


Pythia


Sherpa [also DiRe, Deductor]

## How do they work? [dipole shower case]

- Algorithms based on concepts invented in the mid ' 80 s. Many variants built across the years
e. g. [Sjostrand ' 85 ; Marchesini, Webber ' 88 ; Lonnblad ' 89 ] e.g. $\mathrm{H} \rightarrow \mathrm{bb}$ decay
- Schematically [non-linear evolution]:
-Recursive iteration of $2 \rightarrow 3$ branching probabilities [i.e. LO splitting functions]
- Evolve towards smaller values of a resolution variable [e.g. dipole transverse momentum]
- Kinematic map to restore on-shellness [i.e. recoil scheme]
- Iterate until hadronisation scale is reached


## How do they work? [dipole shower case]

- Algorithms based on concepts invented in the mid ' 80 s. Many variants built across the years
e. g. [Sjostrand ' 85 ; Marchesini, Webber ' 88 ; Lonnblad ' 89 ]
e.g. $\mathrm{H} \rightarrow \mathrm{b} b \mathrm{decay}$
- Schematically [non-linear evolution]:
-Recursive iteration of 2 $\rightarrow 3$ branching probabilities [i.e. LO splitting functions]
- Evolve towards smaller values of a resolution variable [e.g. dipole transverse momentum]
- Kinematic map to restore on-shellness [i.e. recoil scheme]
- Iterate until hadronisation scale is reached



## Energy scale



## How do they work? [dipole shower case]

- Algorithms based on concepts invented in the mid ' 80 s. Many variants built across the years
e. g. [Sjostrand ' ${ }^{5}$; Marchesini, Webber ' ${ }^{8}$; Lonnblad ' ${ }^{8}$ ]
e.g. $\mathrm{H} \rightarrow \mathrm{b} b \mathrm{~b}$ decay


Energy scale


## How do they work? [dipole shower case]

- Algorithms based on concepts invented in the mid ' 80 s. Many variants built across the years
e. g. [Sjostrand '85; Marchesini, Webber '88; Lonnblad '89]
- Schematically [non-linear evolution]:
-Recursive iteration of 2 $\rightarrow 3$ branching probabilities [i.e. LO splitting functions]
- Evolve towards smaller values of a resolution variable [e.g. dipole transverse momentum]
- Kinematic map to restore on-shellness [i.e. recoil scheme]
- Iterate until hadronisation scale is reached
e.g. $H \rightarrow b b$ decay


Energy scale


## How do they work? [dipole shower case]

- Algorithms based on concepts invented in the mid ' 80 s. Many variants built across the years
e. g. [Sjostrand ' ${ }_{5}$; Marchesini, Webber ' ${ }^{88}$; Lonnblad '89]
- Schematically [non-linear evolution]:
-Recursive iteration of $2 \rightarrow 3$ branching probabilities [i.e. LO splitting functions]
- Evolve towards smaller values of a resolution variable [e.g. dipole transverse momentum]
- Kinematic map to restore on-shellness [i.e. recoil scheme]
e.g. $\mathrm{H} \rightarrow \mathrm{b} b \mathrm{decay}$



## What's the logarithmic accuracy of a PS?

- Identify the appropriate QCD perturbative expansion in the multi-scale regime


## Perturbation theory: small coupling, large scale hierarchy [logarithmic counting]

- How can we formulate the concept of accuracy for whole classes of observables at once? e.g. for
- fraction of events passing a jet veto in a rapidity window?
- azimuthal correlation between two sub-jets?
- event shapes?
- ...


## A geometric definition of leading-logarithmic (LL) accuracy

- Radiation phase space conveniently organised in the Lund Plane (LP)
[Anderson, Gustafson, Lonnblad, Pettersson '89]
- $L L \rightarrow$ emissions widely separated in both directions of the LP $\rightarrow \mathcal{O}(50-100 \%)$ uncertainties


Definition used in QCD resummations, e. g.
[Banfi, Salam, Zanderighi '04; Banfi, McAslan, PM, Zanderighi (JHEP 2015)]

## A geometric definition of NLL accuracy

- NLL $\rightarrow$ emissions strongly separated in a single direction of the LP $\rightarrow \mathcal{O}(10 \%)$ uncertainties
e.g. in rapidity at similar transverse momentum

e.g. in transverse momentum at similar rapidities


$$
\text { Line } \longrightarrow \int \mathrm{d} \ln k_{t} \sim \int \mathrm{~d} \eta \sim L / \text { emission }
$$

## A geometric definition of NLL accuracy

- NLL $\rightarrow$ emissions strongly separated in a single direction of the $L P \rightarrow \mathcal{O}(10 \%)$ uncertainties
e.g. in rapidity at similar transverse momentum e.g. in transverse momentum at similar rapidities


## Do existing showers satisfy this?

$$
\text { Line } \longrightarrow \int \mathrm{d} \ln k_{t} \sim \int \mathrm{~d} \eta \sim L / \text { emission }
$$

Definition used in QCD resummations, e. g.

The double-emission matrix element

- Simplest check: probability density for radiating two (soft) gluons
- Compare the result of common showers (e.g. Pythia8, DiRE) to that of a QCD calculation
- Ratio is expected to be $=1$ for NLL showers [Dasgupta, Dreyer, Hamilton, PM, Salam (JHEP 2018)]

Ratio of PS (Py8) radiation pattern to QCD result @ NLL


## The double-emission matrix element

Ratio of PS (Py8) radiation pattern to QCD result @ NLL

- Simplest check: probability density for radiating two (soft) gluons
- Compare the result of common showers (e.g. Pythia8, DiRE) to that of a QCD calculation
- Ratio is expected to be $=1$ for NLL showers [Dasgupta, Dreyer, Hamilton, PM, Salam (JHEP 2018)]


## Common parton showers are

 only LL accurate $\rightarrow$ large uncertainties $O(50-100 \%)$

## Consequences for accuracy: a jet substructure example




- Consider azimuthal distance between two hardest sub-jets
- e.g. Z-boson decay: "quark" jets
- O(60\%) differences with NLL result (large theory uncertainty)


## Consequences for accuracy: a jet substructure example




- Consider azimuthal distance between two hardest sub-jets
- e.g. H-boson decay: "gluon" jets
- unphysical dependence on jet flavour (potential bias for machine learning)


## Formulating NLL parton showers

- Connection between parton showers and perturbative calculations (i.e. resummation) has been an open problem for the past 30 years, i.e. different mathematical language

Anomalous dimensions Renormalisation
Group equations
Power counting


- Mapping of one field into another leads to criteria (e.g. backup) for the building blocks of a PS $\ddagger$
- Methods to create novel algorithms with higher formal accuracy: the PanScales showers
[Dasgupta, Dreyer, Hamilton, PM, Salam, Soyez (PRL 2020)]


## Back to sub-jet's azimuthal correlations



# PanScales showers perfectly agree with NLL, while Pythia/Dire do not 

[Dasgupta, Dreyer, Hamilton, PM, Salam, Soyez (PRL 2020)]

## Repeat the test across several collider observables (e $\mathrm{e}^{-}$- collider case)

[Sjostrand et al. (2020); Hoeche, Prestel (2015)]
Dipole


## A new generation of NLL showers: PanScales



## A new generation of NLL showers: PanScales



## Further developments: towards a full NLL PanScales shower

[Hamilton, Medves, Salam, Scyboz, Soyez (2020)] NLL accuracy tests - Nods method

[van Beekveld, Ferrario Ravasio, Salam, Soto-Ontoso, Soyez (2022)]

Related work on log accuracy in:
[Bewick, Ferrario Ravasio, Richardson, Seymour (2019);
Forshaw, Holguin, Plaetzer (2020); Nagy, Soper (2020)]


\& PanGlobal $(\beta=0) \quad \uparrow$ PanLocal (ant. $\beta=0.5$ )
PanLocal (dip. $\beta=0.5) ~ \square$ Toy shower


## Towards few-percent accuracy: NNLL building blocks

differential collinear fragmentation

[Dasgupta, El-Menoufi (2021); van Beekveld, Dasgupta, ElMenoufi , PM, Salam (in progress)]

NNLL soft (non-g|obal) evolution [a 20 years old problem]


Radiative corrns to hard scattering [preserving PS accuracy]
[Banfi, Dreyer, PM (JHEP 2021 + JHEP 2022)]
Related work by several groups: [Jadach et al. (2015); Li, Skands (2016); Hoeche, Prestel+Krauss+Dulat+Gellersen (2017-2021)]

## The hard scattering

## The hard partonic scattering



## The hard partonic scattering



## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation] e.g. illustration for Higgs+jet production at NLO
(one order less than our target)



## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation] e.g. illustration for Higgs+jet production at NLO


Radiation included by radiative corrections

## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation]
e.g. illustration for Higgs+jet production at NLD



## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation]


> Accuracy broken by double counting across radiation phase space [and virtual corrections]

## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation]

- Not tampering "too much" with the parton shower [i.e. without spoiling its accuracy, so far LL]


## Radiative corrections \& interplay with parton showering

- Computation of radiative corrections to the hard process, while
- Avoiding double-counting with parton shower [PS emits further radiation]


## H

## Explosion of complexity at NNLO [many contributions/configurations, double counting more convoluted]



- Not tampering "too much" with the parton shower [i.e. without spoiling its accuracy, so far LL]

At the same time, we want to keep computational aspects
under control [e.g. fraction of negative weights, stability, ...]

## An LHC example: top-quark pair production

- Main top-quark production mechanism at LHC
[Ferrario Ravasio, Jezo, Nason, Oleari '18]
- Several NP scenarios couple to top quark. Important ingredient of EFT fits
- Inaccuracy of generators already a nuisance




## NNLO event generation

- Variety of methods to handle the production of colourless systems (e.g. EW bosons, Higgs boson)
[Hamilton, Nason, Re, Zanderighi ' ${ }^{13}$; Alioli, Bauer, Berggren, Tackmann, Walsh, Zuberi ' ${ }^{13}$; Hoeche, Li, Prestel '14; PM, Nason, Re, Wiesemann, Zanderighi '19; PM, Re, Wiesemann '20; Campbell, Hoeche, Li, Preuss, Skands '21]
- NNLO event generator for top-pair production has remained a challenge for many years
- Colour charges in initial and final state: involved quantum interference
- Interplay with parton shower highly non trivial
- Many body decays: computationally hard
E.g. first NLO generator for $t \bar{\tau}$ formulated in 2003, it took more than 17 years to achieve NNLO!



## The MiNNLOps method

- Main observation: exploit link between perturbative methods and Monte Carlo language
- Recast NNLO calculation as the first two steps of a parton shower [i.e. radiation ordered in resolution variable, Sudakov factors]
- Fix d.o.f. by matching it to a NNLO perturbative calculation [i.e. resummation properties of $\mathrm{q}_{\mathrm{T}}$ as a resolution variable]
- Advantages:

$\boxed{\square}$ Accurate: Fully differential NNLO QCD
$\triangle$ Fast: Marginal loss in complexity w.r.t. NLO computation
$\boxed{\square}$ Flexible: Possible to tackle complex reactions


## MiNNLOps: NNLO generator for $t \bar{t}$ production

[Mazzitelli, PM, Nason, Re, Wiesemann, Zanderighi (PRL 2021 + JHEP 2022)]

- Validation: verify agreement with perturbative QCD calculations for inclusive observables (i.e. without experimental selection cuts)
- total cross section:

- rapidity distribution of the top pair

> Excellent agreement with pQCD, drastic reduction of theory uncertainties w.r.t. NLO!

## MiNNLOps: broad comparison to experimental data

[Mazzitelli, PM, Nason, Re, Wiesemann, Zanderighi (JHEP 2022)]



## Possible resolution of a long-standing tension in spin correlations?

- Ongoing studies show good theory/data agreement for correlations
[Mazzitelli, PM, Nason, Re, Wiesemann, Zanderighi (JHEP 2022)]



## Conclusions and Outlook

- Modern problems in collider physics demand rethinking the approach to a crucial bridge between theory and experiment: Monte Carlo event generators
- Novel ideas are paving the way to a new generation of tools with a higher and controllable formal accuracy, led by powerful techniques in connection with perturbative QCD:
- New methods to diagnose parton-shower (PS) accuracy and design NLL algorithms
- PS@NLL is today a nearly-solved problem, accessible via public tools in the future. Gearing up for higher orders (NNLL) corrections requires tackling many intriguing conceptual challenges
- Considerable progress in the matching of PS to NNLO calc ${ }^{n}$ for coloured final states. Open problems ahead:
- Consistently preserve higher-order PS accuracy (e.g. matching to PanScales showers)
- First considerations about higher ( $\mathrm{N}^{3} \mathrm{LO}$ ) orders matching have started to emerge


## Backup

## An example: local-recoil dipole showers

- (planar) squared amplitudes built recursively via a Markovian chain of emissions (\& virtuals via unitarity)



## An example: local-recoil dipole showers

- Keep the recoil local, i.e. for each new emission use the map

$$
\text { dipole }\left\{\widetilde{p}_{i}, \widetilde{p}_{j}\right\} \longrightarrow \begin{aligned}
p_{k} & =a_{k} \tilde{p}_{i}+b_{k} \tilde{p}_{j}+k_{\perp} \\
p_{i} & =a_{i} \tilde{p}_{i}+b_{i} \tilde{p}_{j}-f k_{\perp} \\
p_{j} & =a_{j} \tilde{p}_{i}+b_{j} \tilde{p}_{j}-(1-f) k_{\perp}
\end{aligned}
$$

- Typical problem (source of the issues in the heat plot): dipole partitioned in the dipole c.o.m. frame



## The PanLocal algorithm

- Keep the recoil local, i.e. for each new emission use the map

$$
\operatorname{dipole}\left\{\widetilde{p}_{i}, \widetilde{p}_{j}\right\} \longrightarrow \begin{aligned}
p_{k} & =a_{k} \tilde{p}_{i}+b_{k} \tilde{p}_{j}+k_{\perp} \\
p_{i} & =a_{i} \tilde{p}_{i}+b_{i} \tilde{p}_{j}-f k_{\perp} \\
p_{j} & =a_{j} \tilde{p}_{i}+b_{j} \tilde{p}_{j}-(1-f) k_{\perp}
\end{aligned}
$$

- Key element \#1: partitioning ( $\bar{\eta}=0$ ) occurs at equal angles to the dipole ends in the event c.o.m. frame


- In the limit of strong angular ordering and commensurate $\mathrm{k}_{\mathrm{T}}$ 's, $\mathrm{g}_{2}$ takes the recoil from the hard quark



## The PanLocal algorithm

- Keep the recoil local, i.e. for each new emission use the map

$$
\operatorname{dipole}\left\{\widetilde{p}_{i}, \widetilde{p}_{j}\right\} \longrightarrow \begin{aligned}
p_{k} & =a_{k} \tilde{p}_{i}+b_{k} \tilde{p}_{j}+k_{\perp} \\
p_{i} & =a_{i} \tilde{p}_{i}+b_{i} \tilde{p}_{j}-f k_{\perp} \\
p_{j} & =a_{j} \tilde{p}_{i}+b_{j} \tilde{p}_{j}-(1-f) k_{\perp}
\end{aligned}
$$

- Key element \#1: partitioning ( $\bar{\eta}=0$ ) occurs at equal angles to the dipole ends in the event c.o.m. frame


- However, if $\mathrm{g}_{2}$ is produced at larger angles than $g_{1}$, the recoil is still taken from $g_{1}$ in a logarithmic (NLL) region of phase space


## The PanLocal algorithm

- Keep the recoil local, i.e. for each new emission use the map

$$
\operatorname{dipole}\left\{\widetilde{p}_{i}, \widetilde{p}_{j}\right\} \longrightarrow \begin{aligned}
p_{k} & =a_{k} \tilde{p}_{i}+b_{k} \tilde{p}_{j}+k_{\perp} \\
p_{i} & =a_{i} \tilde{p}_{i}+b_{i} \tilde{p}_{j}-f k_{\perp} \\
p_{j} & =a_{j} \tilde{p}_{i}+b_{j} \tilde{p}_{j}-(1-f) k_{\perp}
\end{aligned}
$$

- Key element \#2: modify the evolution variable (instead of dipole $\mathrm{k}_{T}$ )

$$
\begin{aligned}
& k_{t}=\rho v e^{\beta|\bar{\eta}|} \sim v e^{\beta\left|\eta^{w . t . t . ~ e m i t t e r ~}\right|} \\
& \rho=\left(\frac{s_{i} s_{j}}{Q^{2} s_{i \bar{j}}}\right)^{\frac{\beta}{2}}
\end{aligned}
$$

$$
\uparrow^{\ln k_{t}}
$$

## The PanLocal algorithm

- Keep the recoil local, i.e. for each new emission use the map

$$
\operatorname{dipole}\left\{\widetilde{p}_{i}, \widetilde{p}_{j}\right\} \longrightarrow \begin{aligned}
p_{k} & =a_{k} \tilde{p}_{i}+b_{k} \tilde{p}_{j}+k_{\perp} \\
p_{i} & =a_{i} \tilde{p}_{i}+b_{i} \tilde{p}_{j}-f k_{\perp} \\
p_{j} & =a_{j} \tilde{p}_{i}+b_{j} \tilde{p}_{j}-(1-f) k_{\perp}
\end{aligned}
$$

- Key element \#2: modify the evolution variable (instead of dipole $\mathrm{k}_{\mathrm{T}}$ )
- Ordering in v now implies that $\mathrm{k}_{\mathrm{t} 2} \ll \mathrm{k}_{\mathrm{t} 1}$ [i.e. no recoil]
- Interplay of partition $\oplus o r d e r i n g ~ e n s u r e s ~ t h a t ~ t h e ~ r e c o i l ~ i s ~$ always taken from the hard extremities [OK at NLL]


