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 ~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)





Current integrated luminosity

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)





Main challenge: controlling the fine structure of collider events

Candidate $H \rightarrow \gamma \gamma$ event at the LHC



Vast technological progress (jointly Theory \otimes **Experiment)**











event generators



non-perturbative (QCD) corrections

(...)

$$\tilde{\alpha}_s(\mu^2) \equiv \tilde{\alpha}_s(0) + \int_0^\infty \mathrm{d}m^2 \frac{\mu^2}{m^2 + \mu^2} \frac{\mathrm{d}\alpha_{\mathrm{eff}}(m^2)}{\mathrm{d}m^2}$$

landscape of NP models

novel strategies (e.g. ML, new observables)



convolutional neural network



- - - ,

E.g. Impressive progress in theoretical calculations

formal developments

landsc





event generators



non-perturbative (QCD) corrections

Understanding

of QFTs

 $\tilde{\alpha}_s(\mu^2) \equiv \tilde{\alpha}_s(0) + \int_0^\infty \mathrm{d}m^2 \frac{\mu^2}{m^2 + \mu^2} \frac{\mathrm{d}\alpha_{\mathrm{eff}}(m^2)}{\mathrm{d}m^2}$

pe of NP models



perturbative methods

Experiment & pheno

novel strategies (e.g. ML, new observables)



convolutional neural netwo







E.g. Impressive progress in theoretical calculations



[Chen, Gehrmann, Glover, Huss, PM, Re, Rottoli, Torrielli '22]





/GeV

[bp

 $d\sigma$



This talk focuses on another crucial aspect: Event Generators

formal developments

landsc



perturb

event generators



Understanding of QFTs

non-perturbative (QCD) corrections

 $\tilde{\alpha}_s(\mu^2) \equiv \tilde{\alpha}_s(0) + \int_0^\infty \mathrm{d}m^2 \frac{\mu^2}{m^2 + \mu^2} \frac{\mathrm{d}\alpha_{\mathrm{eff}}(\mu^2)}{\mathrm{d}m^2} \frac{$

pe of NP models

ative methods



Experiment & pheno

novel strategies (e.g. ML, new observables)



convolutional neural netwo







Anatomy of a scattering reaction at the LHC



- Short distance (hard)
 - scales probed: O(10²)-O(10³) GeV
- stage sensitive to NP



Anatomy of a scattering reaction at the LHC



- Short distance (hard)
- scales probed: O(10²)-O(10³) GeV
- stage sensitive to NP



evolution towards a physical observable state (mainly QCD)

- Long distance (soft)
- transition from $O(10^2)-O(10^3)$ GeV to O(1) 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



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



- 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)





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 ($\mu_{hard} \gg \mu_{soft}$)
- Yet, fully perturbative regime ($\mu_{\rm soft} \gg \Lambda_{\rm QCD}$)
- Initial conditions for hadronisation
- Several successful public tools:



[also DiRe, Deductor]





Algorithms based on concepts invented in the mid '80s. Many variants built across the years

- Schematically [non-linear evolution]:
- -Recursive iteration of 2→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







Algorithms based on concepts invented in the mid '80s. Many variants built across the years

- Schematically [non-linear evolution]:
- -Recursive iteration of 2→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





Algorithms based on concepts invented in the mid '80s. Many variants built across the years

- Schematically [non-linear evolution]:
- -Recursive iteration of 2→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





Algorithms based on concepts invented in the mid '80s. Many variants built across the years

- Schematically [non-linear evolution]:
- -Recursive iteration of 2→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





Algorithms based on concepts invented in the mid '80s. Many variants built across the years

- Schematically [non-linear evolution]:
- -Recursive iteration of 2→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





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)
- LL \rightarrow emissions widely separated in both directions of the LP $\rightarrow O(50 100\%)$ uncertainties



Definition used in QCD resummations, e.g.

[Banfi, Salam, Zanderighi '04; Banfi, McAslan, PM, Zanderighi (JHEP 2015)]

[Anderson, Gustafson, Lonnblad, Pettersson '89]



A geometric definition of NLL accuracy

• NLL \rightarrow emissions strongly separated in a single direction of the LP $\rightarrow O(10\%)$ uncertainties

e.g. in rapidity at similar transverse momentum



Definition used in QCD resummations, e.g.

[Banfi, Salam, Zanderighi '04; Banfi, McAslan, PM, Zanderighi (JHEP 2015)]

e.g. in transverse momentum at similar rapidities





A geometric definition of NLL accuracy

• NLL \rightarrow emissions strongly separated in a single direction of the LP $\rightarrow O(10\%)$ uncertainties



Definition used in QCD resummations, e.g.

[Banfi, Salam, Zanderighi '04; Banfi, McAslan, PM, Zanderighi (JHEP 2015)]

e.g. in rapidity at <u>similar transverse momentum</u> e.g. in transverse momentum at similar rapidities **<u>Criterion</u>: a parton shower is expected to** reproduce correctly all these limits at once [+ consistent treatment of virtual corrections] Do existing showers satisfy this? \boldsymbol{q} g_2

Line $\longrightarrow \int d\ln k_t \sim \int d\eta \sim L/emission$



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)]













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)]

Common parton showers are only LL accurate \rightarrow large uncertainties $\mathcal{O}(50 - 100\%)$













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

 \bigcirc problem for the past 30 years, i.e. different mathematical language



- Mapping of one field into another leads to criteria (e.g. backup) for the building blocks of a PS[‡]
- Methods to create novel algorithms with higher formal accuracy: the PanScales showers

[‡]QCD resummation provides guidelines, more than one architecture is possible

Connection between parton showers and perturbative calculations (i.e. resummation) has been an open

[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+e- collider case)



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

Plot: relative deviations from exact NLL [taken in the relevant kinematic limit]





A new generation of NLL showers: PanScales



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



A new generation of NLL showers: PanScales



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

Further developments: towards a full NLL PanScales shower

Related work on log accuracy in:

Forshaw, Holguin, Plaetzer (2020); Nagy, Soper (2020)]

[Karlberg, Salam, Scyboz, Verheyen (2021)] EEEC

Towards few-percent accuracy: NNLL building blocks

differential collinear fragmentation

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

Radiative corr^{ns} to hard scattering [preserving PS accuracy]

NNLL soft (non-global) evolution [a 20 years old problem]

[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)] 36

The hard scattering

The hard partonic scattering

 QCD well described by the radiation of a fixed number of partons (quarks & gluons) [and corresponding virtual corrections]

Perturbation theory: [fixed-order] $\alpha_s(M_Z) \sim 0.118 \longrightarrow \begin{cases} NLO & \sim 10\% \\ NNLO & \sim 1\% \\ \dots \end{cases}$

The hard partonic scattering

 QCD well described by the radiation of a fixed number of partons (quarks & gluons) [and corresponding virtual corrections]

Perturbation theory: [fixed-order] $\alpha_s(M_Z) \sim 0.118 \longrightarrow \begin{cases} NLO & \sim 10\% \\ NNLO & \sim 1\% \end{cases}$

Target: (At least!) Next-to-next-to-leading ordercorrections with respect to Born approximation:up to two extra emissions

- 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)

- 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

- 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

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

NLO solutions: [Frixione, Webber '02; Nason '04; Jadach et al. '15; Nason, Salam '22]

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

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

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

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

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

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]

An LHC example: top-quark pair production

- Main top-quark production mechanism at LHC
 - Several NP scenarios couple to top quark. Important ingredient of EFT fits
- Inaccuracy of generators already a nuisance

NNLO event generation

- NNLO event generator for top-pair production has remained a challenge for many years \bigcirc
 - 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 $t\bar{t}$ formulated in 2003, it took more than 17 years to achieve NNLO!

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]

The MiNNLO_{PS} 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 q_T as a resolution variable]
- Advantages:
 - **Accurate:** Fully differential NNLO QCD
 - **Fast:** Marginal loss in complexity w.r.t. NLO computation
 - **Flexible**: Possible to tackle complex reactions

[PM, Nason, Re, Wiesemann, Zanderighi (JHEP 2019); PM, Re, Wiesemann (EPJC 2020)]

MiNNLO_{PS}: NNLO generator for *tt* **production**

- experimental selection cuts)
 - total cross section:

rapidity distribution of the top pair

MiNNLO_{PS}: broad comparison to experimental data

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

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

Conclusions and Outlook

- and experiment: Monte Carlo event generators
- led by powerful techniques in connection with perturbative QCD:
- New methods to diagnose parton-shower (PS) accuracy and design NLL algorithms
- orders (NNLL) corrections requires tackling many intriguing conceptual challenges
- Consistently preserve higher-order PS accuracy (e.g. matching to PanScales showers)
- First considerations about higher ($N^{3}LO$) orders matching have started to emerge

• Modern problems in collider physics demand rethinking the approach to a crucial bridge between theory

Novel ideas are paving the way to a new generation of tools with a higher and controllable formal accuracy.

- PS@NLL is today a nearly-solved problem, accessible via public tools in the future. Gearing up for higher

• Considerable progress in the matching of PS to NNLO calcⁿ for coloured final states. Open problems ahead:

Backup

An example: local-recoil dipole showers

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

te with n particles
$$S_n$$
 to S_{n+1}

$$\int d\bar{\eta} \frac{d\phi}{2\pi} \frac{\alpha_s(k_t) + K\alpha_s^2(k_t)}{\pi}$$

$$\tilde{j}_k(a_k) + g(-\bar{\eta})b_k P_{\tilde{j} \to jk}(b_k)],$$
L0 splitting function

Dipole partitioning (rapidity within the dipole)

Recoil assigned according to a map $S_n \rightarrow S_{n+1}$

An example: local-recoil dipole showers

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

dipole
$$\{\widetilde{p}_i, \widetilde{p}_j\}$$

• Typical problem (source of the issues in the heat plot): dipole partitioned in the dipole c.o.m. frame e.g. Pythia8 / DiRE $\ln k_t$

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp,$$

$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp,$$

$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1 - f) k_\perp$$

In the limit of strong angular ordering and commensurate k_T 's, g_2 can still take the recoil from g₁

→ i.e. violation of locality in the LP

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

dipole $\{\widetilde{p}_i, \widetilde{p}_j\}$ —

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

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

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp ,$$

$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp ,$$

$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1 - f) k_\perp$$

In the limit of strong angular ordering and commensurate k_T 's, g_2 takes the recoil from the hard quark

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

dipole $\{\widetilde{p}_i, \widetilde{p}_j\}$

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

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

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp,$$

$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp,$$

$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1 - f) k_\perp$$

 However, if g₂ is produced at larger angles than g_1 , the recoil is still taken from g_1 in a logarithmic (NLL) region of phase space

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

dipole
$$\{\widetilde{p}_i, \widetilde{p}_j\}$$

• Key element #2: modify the evolution variable (instead of dipole k_T)

$$k_{t} = \rho v e^{\beta |\bar{\eta}|} \sim v e^{\beta |\eta^{\text{w.r.t. emitter}}|}$$
$$\rho = \left(\frac{s_{\tilde{\imath}} s_{\tilde{\jmath}}}{Q^{2} s_{\tilde{\imath}\tilde{\jmath}}}\right)^{\frac{\beta}{2}}$$

 k_T ordering corresponds to $\beta=0$

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

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp,$$

$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp,$$

$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1 - f) k_\perp$$

 g_1

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

dipole $\{\widetilde{p}_i, \widetilde{p}_j\}$ —

• Key element #2: modify the evolution variable (instead of dipole k_T)

- Ordering in v now implies that k_{t2} << k_{t1} [i.e. no recoil]
- Interplay of partition
 ordering ensures that the recoil is always taken from the hard extremities [OK at NLL]

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

$$p_k = a_k \tilde{p}_i + b_k \tilde{p}_j + k_\perp,$$

$$p_i = a_i \tilde{p}_i + b_i \tilde{p}_j - f k_\perp,$$

$$p_j = a_j \tilde{p}_i + b_j \tilde{p}_j - (1 - f) k_\perp$$

