Overview of Monte Carlo Generators

John Campbell, Fermilab

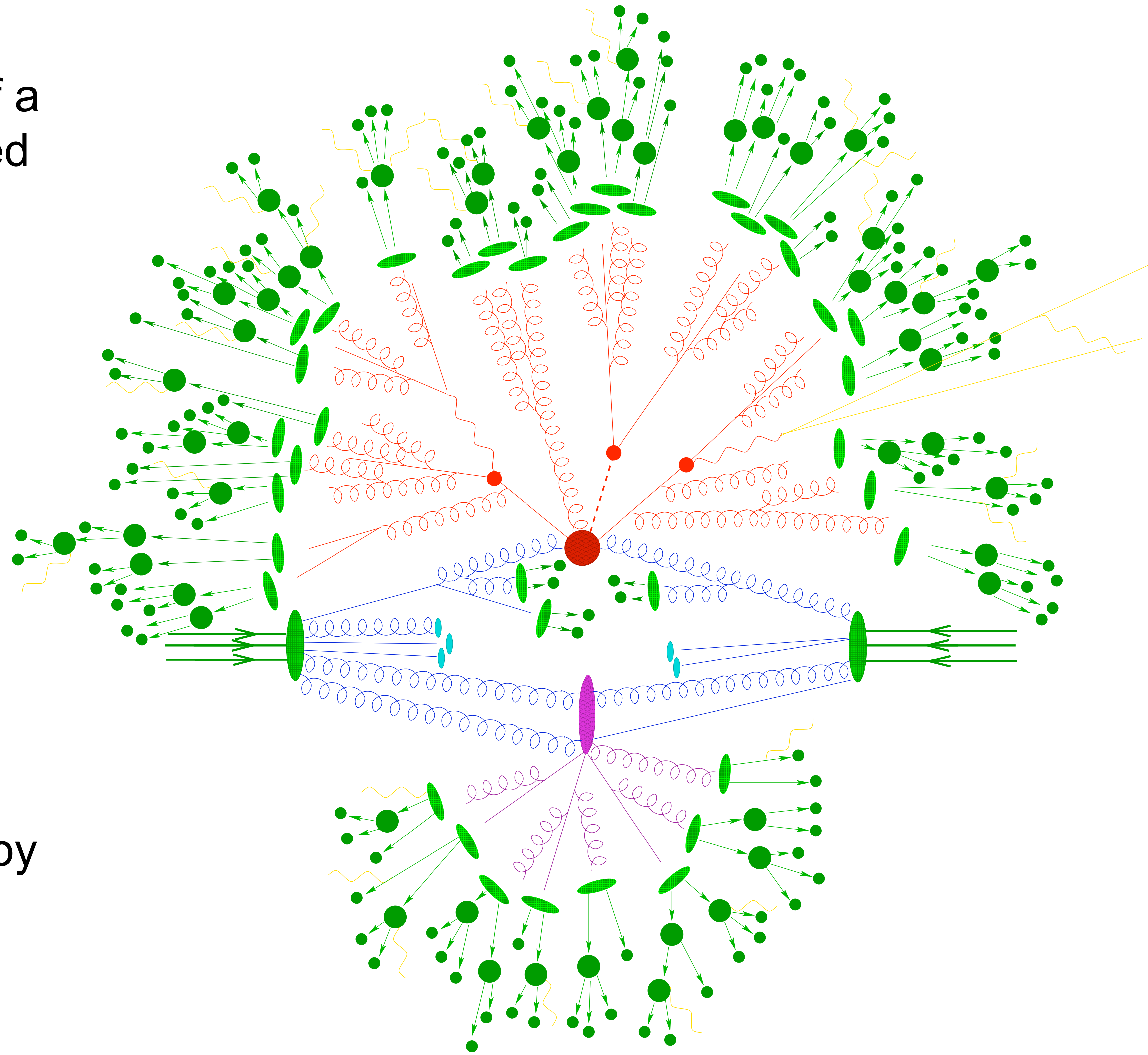
Monte Carlo overview: history

- Theoretical description of a sample of events recorded by an experiment.
- Distinction between:
- flexible event generators that provide an exclusive **description of the full final state**

  e.g. Pythia, HERWIG

- specialized **parton level predictions** that can be systematically improved by including higher orders in perturbation theory

  e.g. MCFM, BlackHat



More recently ...

- Difference significantly blurred by an understanding of how to combine exact fixed order results with the event generator capability of a parton shower

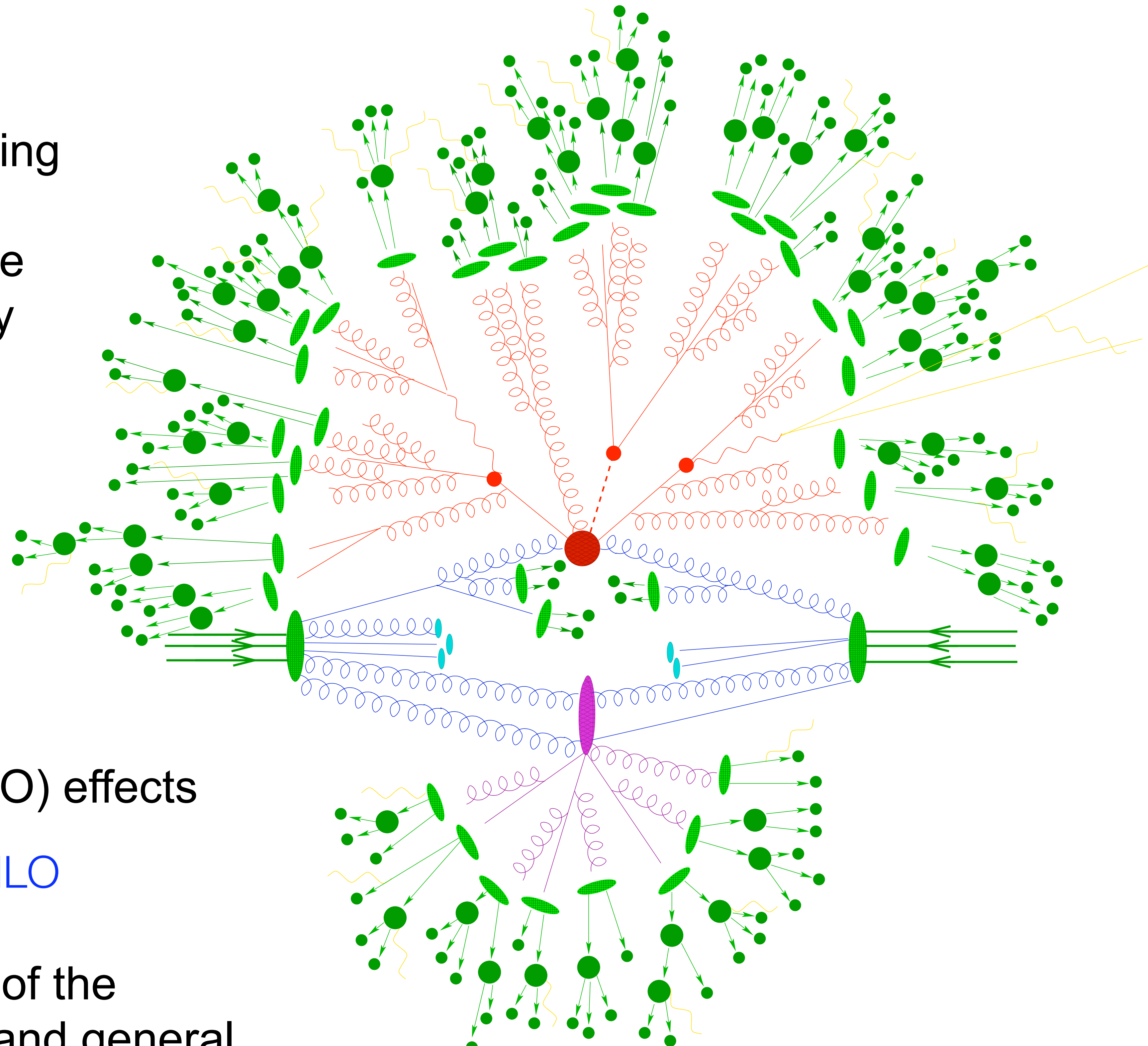
- Better treatment of hard radiation via merging of samples

  e.g. ALPGEN, SHERPA

  Systematic inclusion of next-to-leading order (NLO) effects

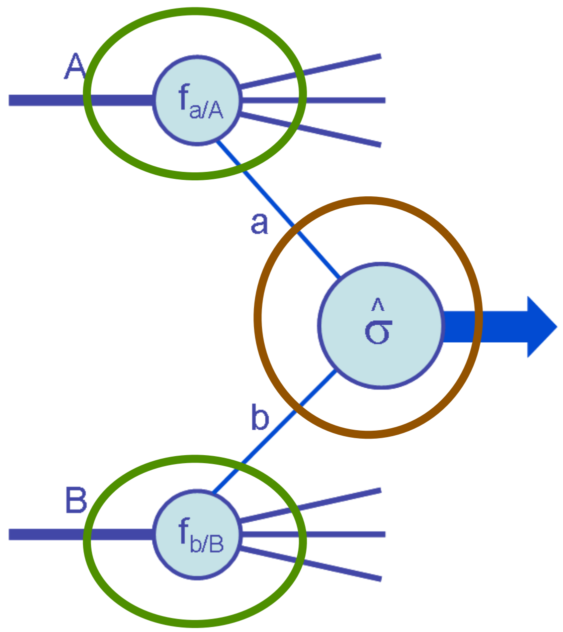
  e.g. POWHEG, (a)MC@NLO

- Will try to describe some of the theoretical underpinning and general features of the above.



Factorization

- A theoretical description of the process relies on **factorization** of the problem into long- and short-distance components:



universal parton distribution functions

$$\sigma_{AB} = \int dx_a dx_b \, f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \, \hat{\sigma}_{ab\to X}$$

hard scattering matrix element

- Sensible description in theory only if **scale Q^2 is hard**, i.e. production of a massive object or a jet with large transverse momentum.
- Asymptotic freedom then allows a perturbative expansion of all ingredients that appear in the calculation, e.g.

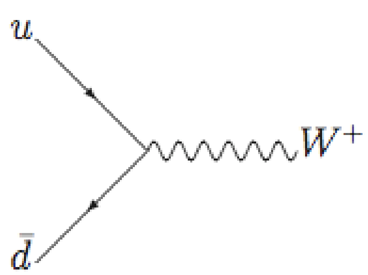
$$\hat{\sigma}_{ab\to X} = \hat{\sigma}^{(0)}_{ab\to X} + \alpha_s(Q^2)\hat{\sigma}^{(1)}_{ab\to X} + \alpha_s^2(Q^2)\hat{\sigma}^{(2)}_{ab\to X} + \ldots$$

Leading order

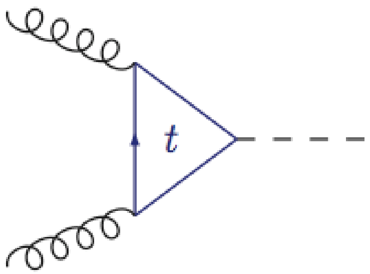
- Simplest picture: **leading order parton level prediction**.

$$\sigma_{AB} = \int dx_a dx_b \; f_{a/A}(x_a, Q^2) f_{b/B}(x_b, Q^2) \; \hat{\sigma}^{(0)}_{ab \to X}$$

1. Identify the leading-order partonic process that contributes to the hard interaction producing X
    - each jet is replaced by a quark or gluon ("**local parton-hadron duality**")
2. Calculate the corresponding matrix elements.



usually a tree diagram, e.g. Drell-Yan



... but not always, e.g. Higgs from gluon fusion

3. Combine with appropriate combinations of pdfs for initial-state partons *a*,*b*.
4. Perform a numerical integration over the energy fractions x_a, x_b and the phase-space for the final state X.

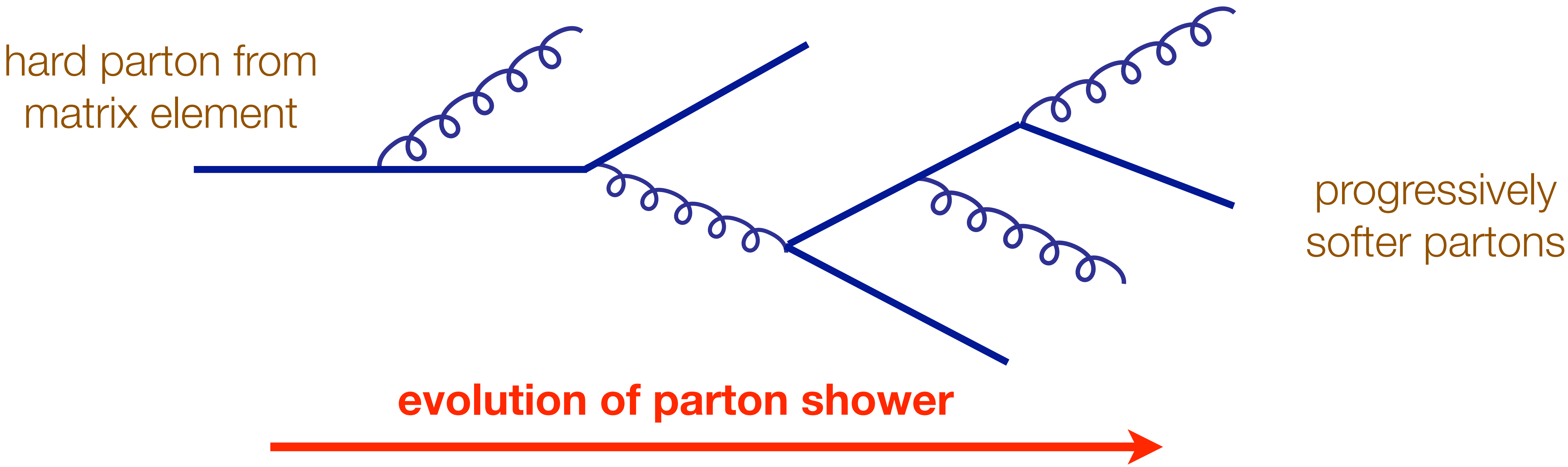
Tools for LO calculations

- Practically a solved problem for over a decade - many suitable tools available.
- Computing power can still be an issue.
- This is mostly because the number of Feynman diagrams entering the amplitude calculation **grows factorially** with the number of external particles.
- hence smart (**recursive**) methods to generate matrix elements.

ALPGEN	M. L. Mangano et al. http://alpgen.web.cern.ch/alpgen/
AMEGIC++	F. Krauss et al. http://projects.hepforge.org/sherpa/dokuwiki/doku.php
CompHEP	E. Boos et al. http://comphep.sinp.msu.ru/
HELAC	C. Papadopoulos, M. Worek http://helac-phegas.web.cern.ch/helac-phegas/helac-phegas.html
MadGraph	F. Maltoni, T. Stelzer http://madgraph.hep.uiuc.it/

Parton showers

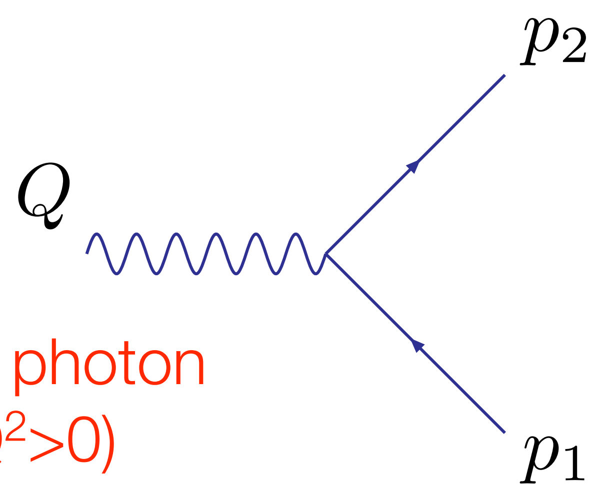
- A parton shower is a way of simulating the radiation of additional quarks and gluons from the hard partons included in the matrix element.

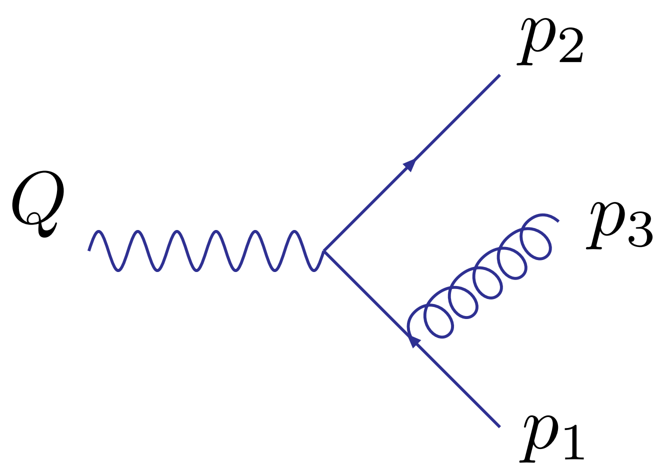


- This radiation is important for describing more detailed properties of events, e.g. the structure of a jet.
- Eventually, evolution produces partons that are are soft (~GeV) and that must be arranged into hadrons (**hadronization**) $\rightarrow$ true event generator.

Underlying theory of parton showers

- The construction of a parton shower is based on another type of factorization: of cross sections in **soft** and **collinear** limits.
- Easiest way to see the behavior is in the matrix elements.





(2 diagrams)

$$|\mathcal{M}_{\gamma^* \bar{q} q}|^2 = 4 N_c e_q^2 Q^2 \qquad |\mathcal{M}_{\gamma^* \bar{q} q g}|^2 = 8 N_c C_F e_q^2 g_s^2 \left(\frac{(2p_1.p_3)^2 + (2p_2.p_3)^2 + 2Q^2(2p_1.p_2)}{4\, p_1.p_3\, p_2.p_3} \right)$$

- In the limit that quark 2 and gluon 3 are collinear, $p_2 = zP \; , \quad p_3 = (1-z)P$ there is a remarkable factorization:

$$|\mathcal{M}_{\gamma^* \bar{q} q g}|^2 \xrightarrow{\text{coll.}} \frac{2 g_s^2}{2 p_2 . p_3} |\mathcal{M}_{\gamma^* \bar{q} q}|^2 P_{qq}(z) \qquad P_{qq}(z) = C_F \left(\frac{1+z^2}{1-z} \right)$$

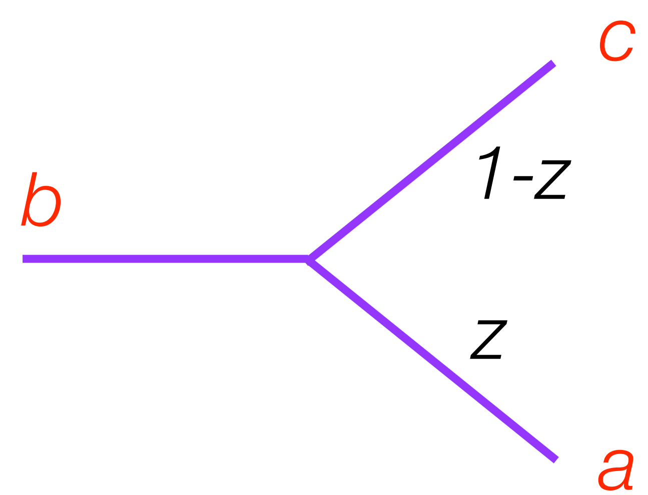
splitting function

Universal soft and collinear factorization

- The important feature is that this behaviour is **universal**, i.e. it applies to the appropriate collinear limits in **all processes involving QCD radiation**.
- They are a feature of the QCD interactions themselves.

$$|\mathcal{M}_{ac\ldots}|^2 \xrightarrow{a,\,c\text{ coll.}} \frac{2g_s^2}{2p_a.p_c}|\mathcal{M}_{b\ldots}|^2 P_{ab}(z)$$

collinear singularity



$$P_{qq}(z) = C_F\left(\frac{1+z^2}{1-z}\right)$$

additional soft singularity as $z \to 1$

$$P_{gg}(z) = 2N_c\left(\frac{z^2+(1-z)^2+z^2(1-z)^2}{z(1-z)}\right)$$

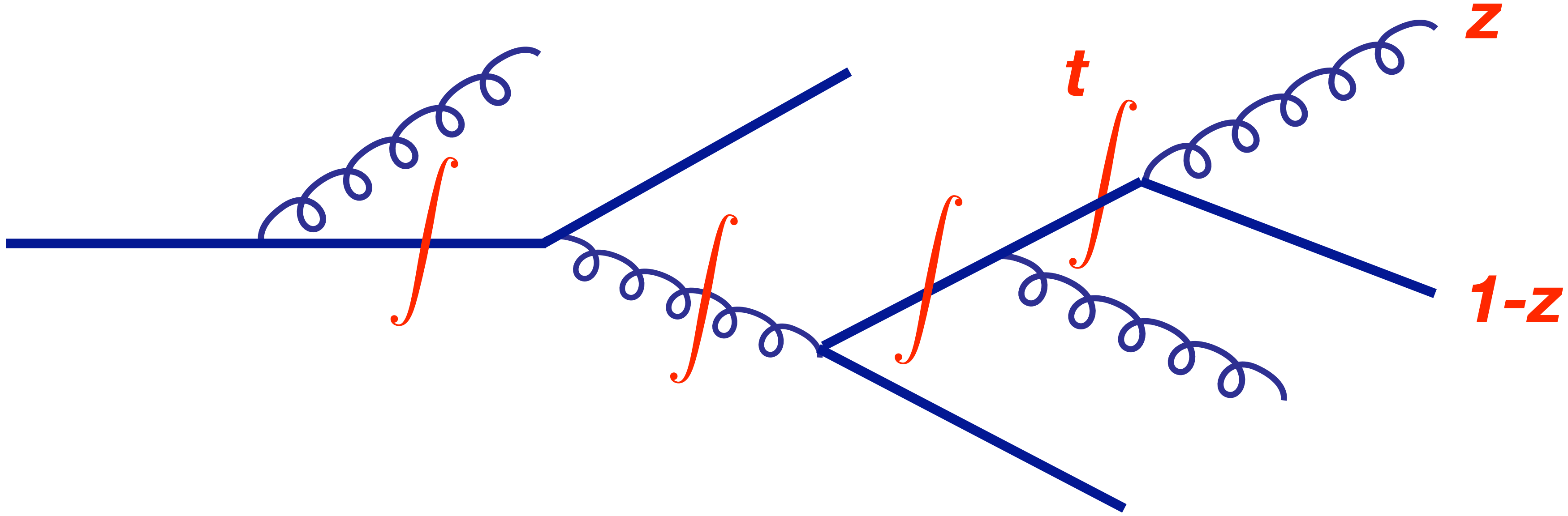
soft for $z \to 0$, $z \to 1$

$$P_{qg}(z) = T_R\left(z^2+(1-z)^2\right)$$

Parton showers

- Factorization extends to the phase space too and hence to the level of cross sections:

$$d\sigma_{n+1} = d\sigma_n \left(\frac{\alpha_s}{2\pi}\right) \frac{dt}{t} P_{ab}(z)\, dz$$



- In this equation the virtuality, ***t*** is the evolution variable; other choices are **angular ordered** and **p_T ordered**.

$$\frac{dt}{t} = \frac{d\theta^2}{\theta^2} = \frac{dp_T^2}{p_T^2}$$

- Here we have considered timelike branching (all particles are outgoing, *t>0*).
  - extension to the spacelike case (radiation on an incoming line) is similar.
- This is the principle upon which all **parton shower** simulations are based.

Sudakov logarithms

- Solution of evolution equation in terms of **Sudakov form factor**, corresponding to probability of no resolvable emission
- Resolvable means not arbitrarily soft: remove singularities as z→0 and z→1:

$$\frac{t'}{t_0} < z < 1 - \frac{t'}{t_0}$$

- Probability of **no resolvable branchings** from a quark:

$$\Delta_q(t) = \exp\left[-\int_{t_0}^{t} \frac{dt'}{t'} \int_{t_0/t'}^{1-t_0/t'} dz \left(\frac{\alpha_s}{2\pi}\right) P_{qq}(z)\right]$$

$$\sim \exp\left[-C_F \left(\frac{\alpha_s}{2\pi}\right) \int_{t_0}^{t} \frac{dt'}{t'} \int_{t_0/t'}^{1-t_0/t'} \frac{dz}{1-z}\right]$$

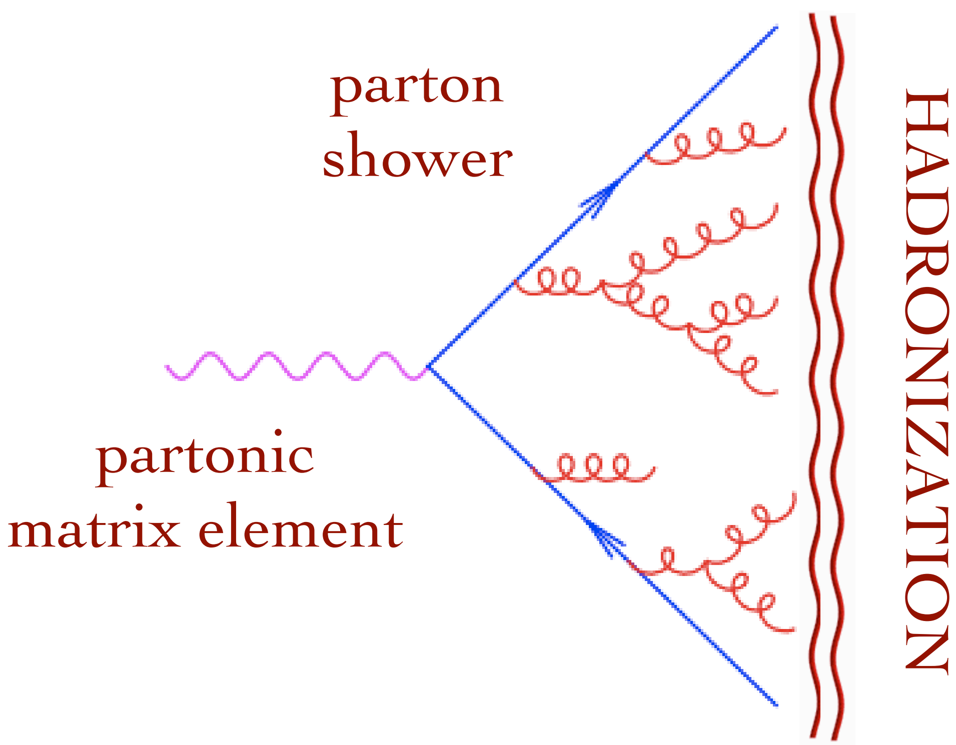
$$\sim \exp\left[-C_F \left(\frac{\alpha_s}{2\pi}\right) \log^2 \frac{t}{t_0}\right]$$

leading log parton shower

- **Exponentiation** sums all terms with greatest number of logs per power of α_s

Hadronization

- At very small scales of *t* perturbation theory is no longer valid
  - no further branching beyond ~ GeV.
- All partons produced in the shower are showered further, until same condition.



- Once this point is reached, no more perturbative evolution possible.
- Partons should be interpreted as hadrons according to a **hadronization model**.
  - examples: **string model** (Pythia), **cluster model** (Herwig, Sherpa).

- Most importantly: these are phenomenological models.
- They require inputs that cannot be predicted from the QCD Lagrangian ab initio and must therefore be tuned by comparison with data (mostly LEP).

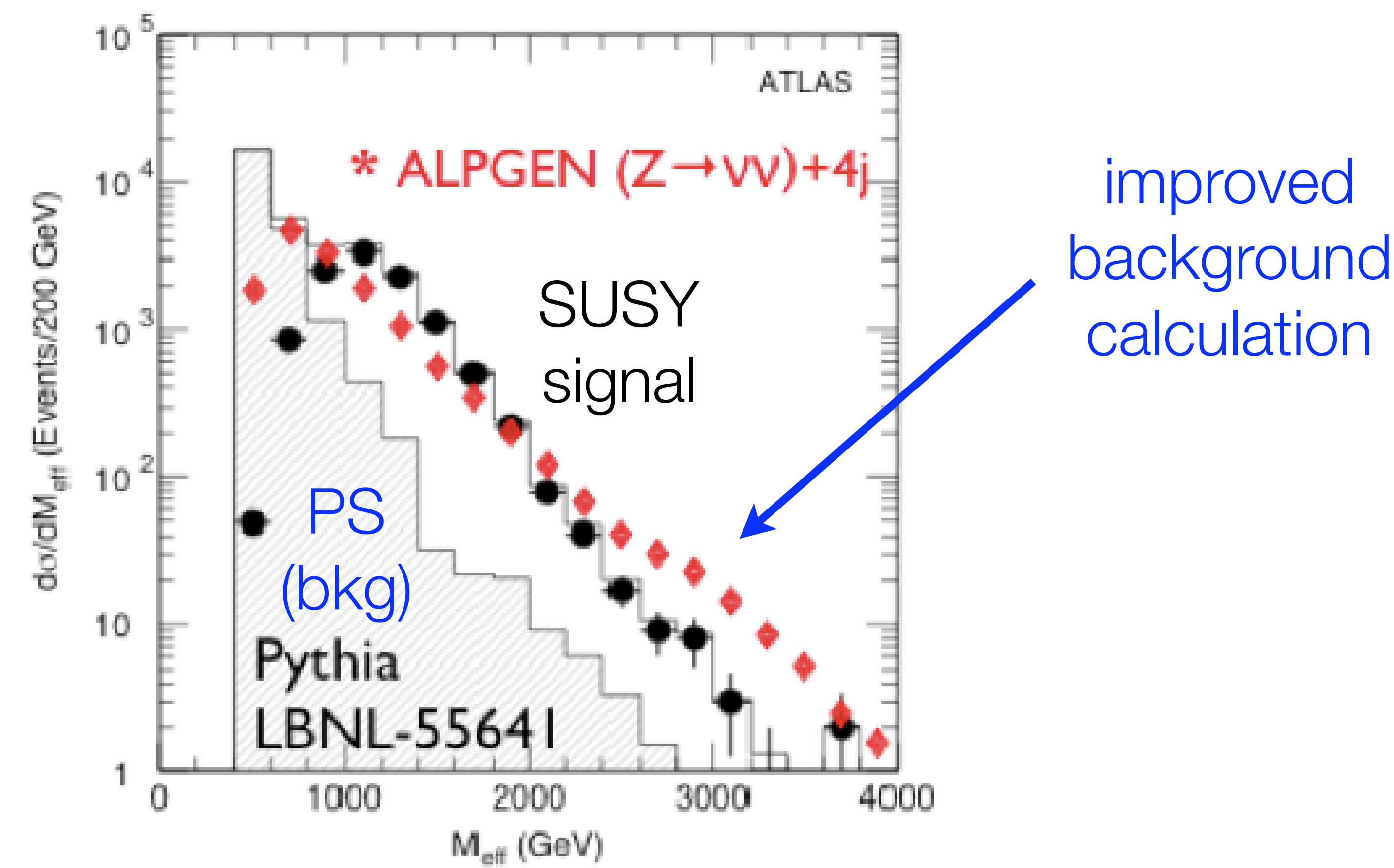
Popular parton shower programs

PYTHIA	T. Sjöstrand et al. http://home.thep.lu.se/~torbjorn/Pythia.html
HERWIG	G. Corcella et al. http://hepwww.rl.ac.uk/theory/seymour/herwig/
HERWIG++	S. Gieseke et al. http://projects.hepforge.org/herwig/
SHERPA	F. Krauss et al. http://projects.hepforge.org/sherpa/dokuwiki/doku.php
ISAJET	H. Baer et al. http://www.nhn.ou.edu/~isajet/

Warnings

- By construction, a parton shower is correct only for successive branchings that are collinear or soft (i.e. only leading logs, or next-to-leading logs with care).

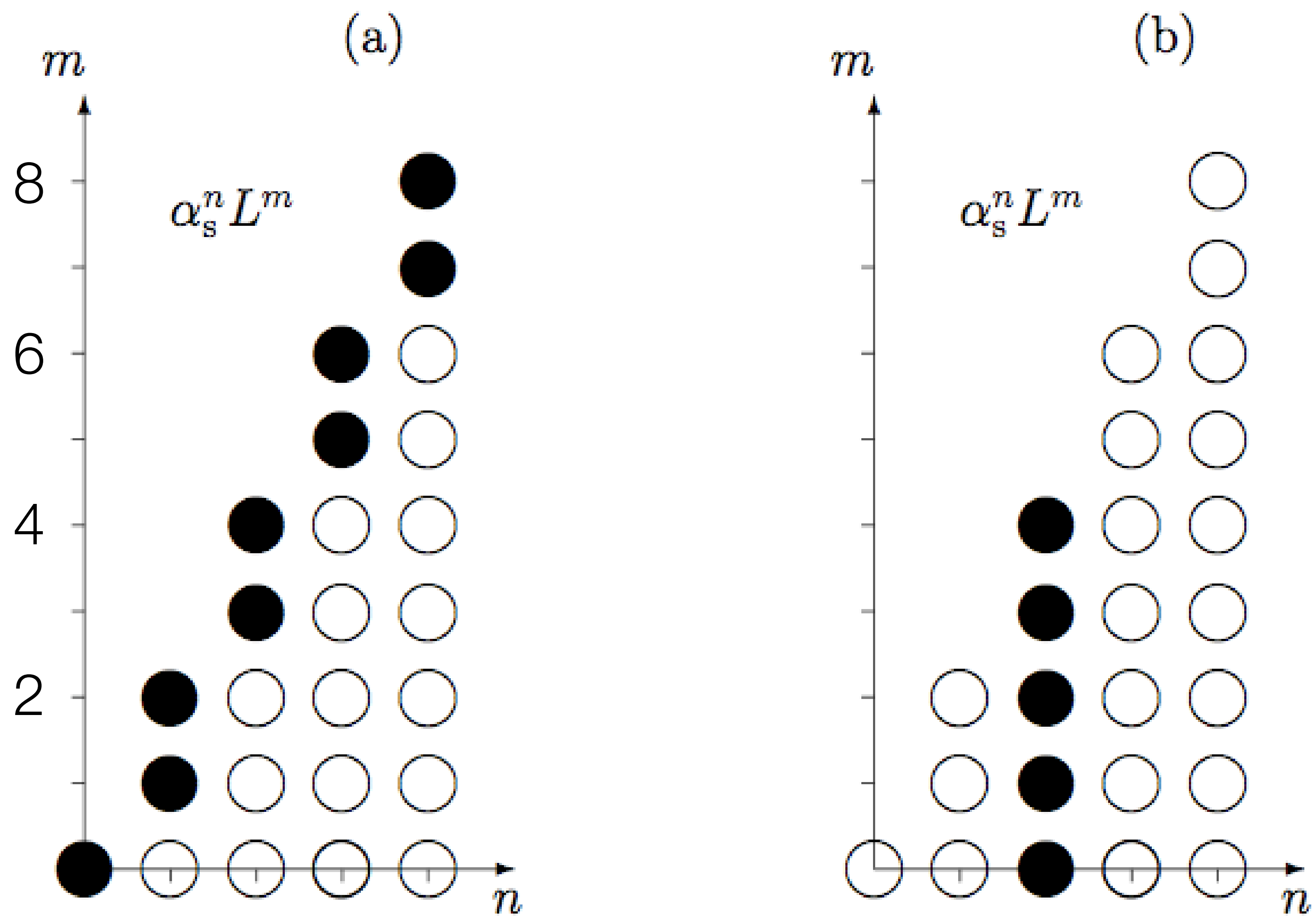
- **Must take care** when describing final states in which there is either manifestly multiple hard radiation, or its effects might be important.
    - example: simulation of background to a SUSY search in the ATLAS TDR.



$$M_{\text{eff}} = \Sigma_i \left|p_{T(i)}\right| + \not{E}_T$$

Parton shower extensions

- As simplest example, consider Drell-Yan process: $pp \to Z \ (+n \text{ jets})$
  (one power of α_s per jet)



accuracy of NLL parton shower

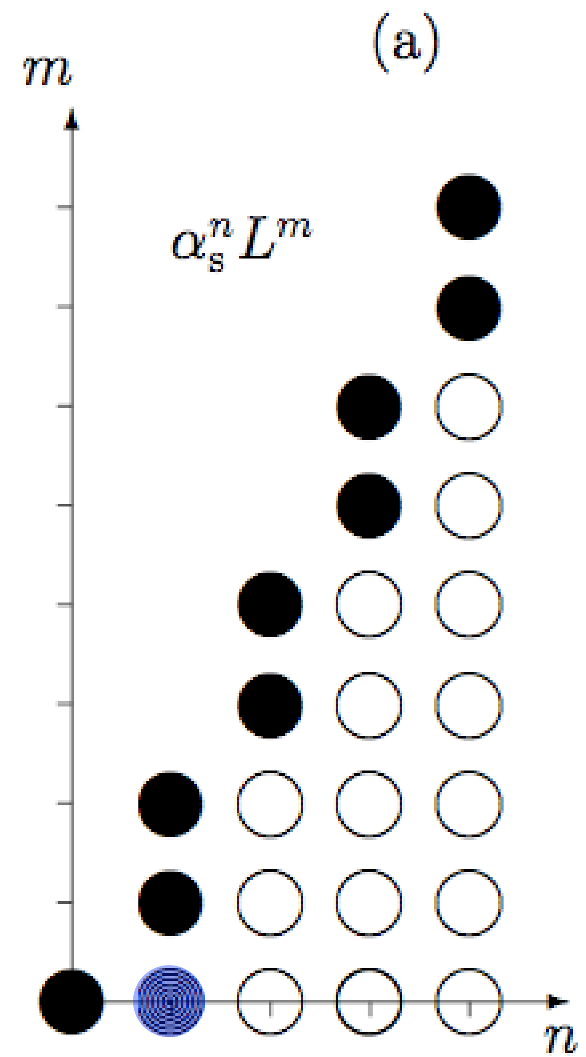
accuracy of tree-level Z+2 jet calculation

c.f. earlier, leading log: $\alpha_s^n L^{2n}$

How can parton shower recover more of fixed-order accuracy?

Tree-level matching

- Use exact matrix elements for the hardest emission from the parton shower instead of approximate form.
- Captures one extra term in the expansion
  - does not account for all corrections
  - real radiation is taken into account but not virtual (loop diagram) contributions
- Hence shape improved for observables dominated by low-multiplicity emission, but overall normalization same as before.

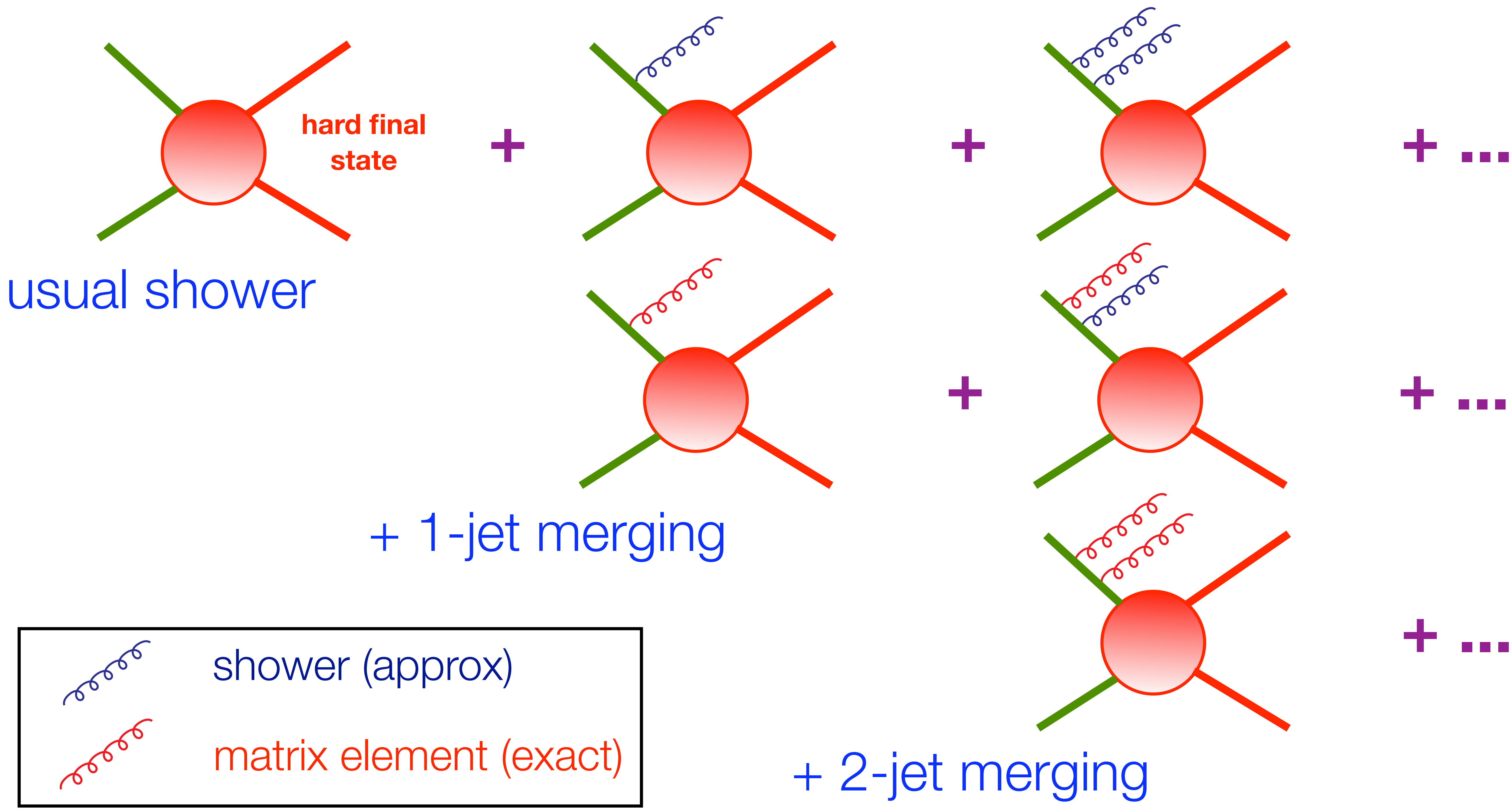


Mostly historical, not a feature of modern generators

PS + one matched emission

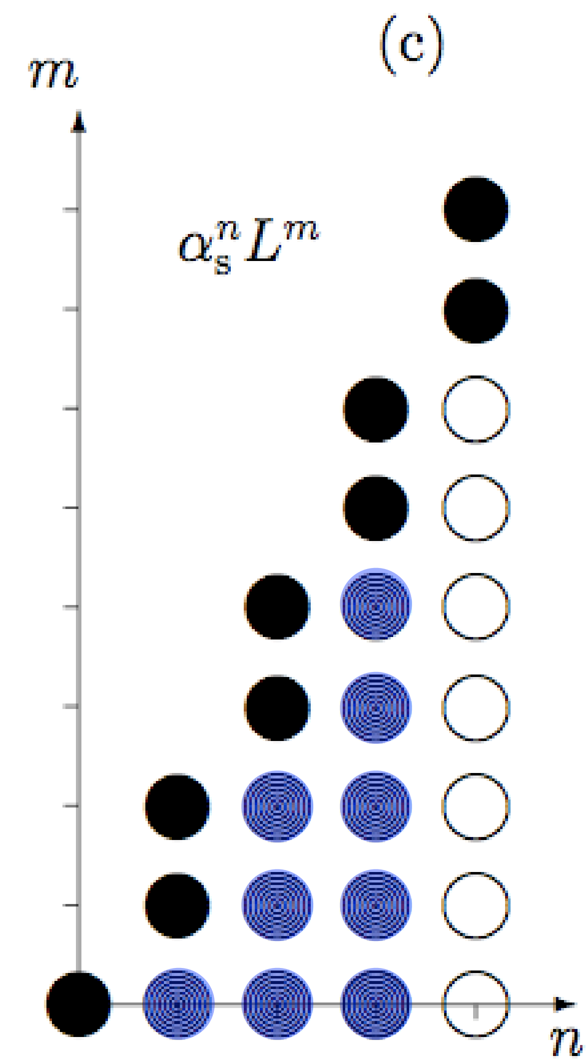
Multi-jet merging in pictures

- Merging: **include more exact matrix elements as initial hard scatters**, with merging scale determining transition from approximate to exact MEs.



Multi-jet merging

- Perform matrix element corrections to multiple emissions
- Introduces an unphysical merging scale in order to perform corrections for each jet multiplicity
  - again, **impact only on shapes** of relevant distributions - for observables up to the number of jet samples merged
  - again, no possible improvement in rate cross section **remains a leading order estimate**
- Various techniques for combining samples without overcounting in shower:
  - CKKW (Catani, Krauss, Kuhn, Webber) CKKW-L (Lönnblad)
    SHERPA
  - MLM (Mangano)
    ALPGEN



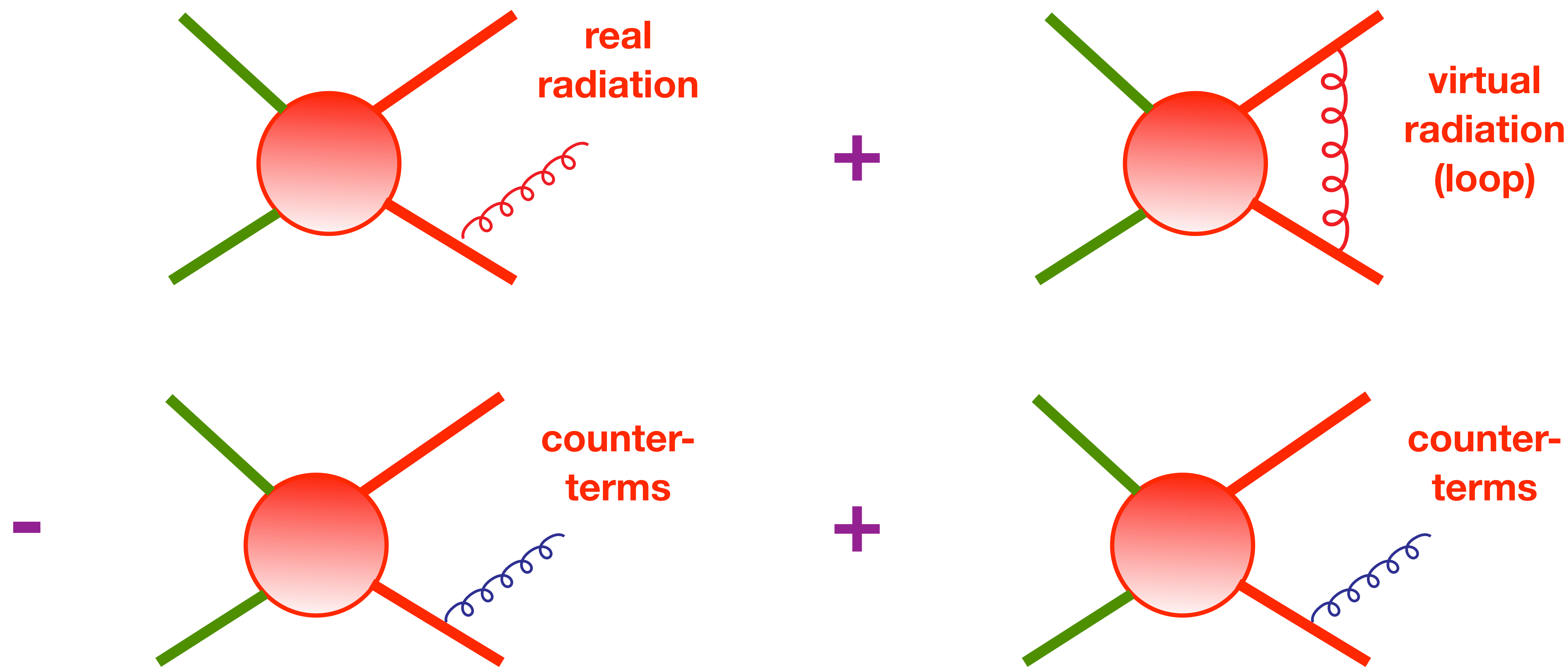
PS + multi-jet merging for up to 3 jets

Higher orders

- Systematic method of improving the perturbative prediction
  - at each order of calculation, result is formally independent of the choice made for the **renormalization scale** (used in strong coupling) and **factorization scale** (used to evaluate pdfs)
  - uncertainty is usually estimated by examining residual sensitivity to these scales; typically reduced as more orders are considered
  - scales should be motivated by physics; range of variation is a subject for debate and resulting uncertainties certainly not Gaussian!
- Approximate hierarchy:
  - LO: ballpark estimate (~ within a factor of two)
  - NLO: first serious estimate (~ 10% accuracy)
  - NNLO: precision (~ few percent), serious estimate of uncertainty

General structure of NLO calculation

- Two divergent contributions, only finite together (**Kinoshita-Lee-Nauenberg**)



soft/collinear singularities cancelled numerically

singularities cancelled analytically

- In general: many “counter-events” for each real radiation event.
- Naive “event generator” produces mixed sample of parton configurations with large positive and negative weights; **not suitable for usual analysis**.

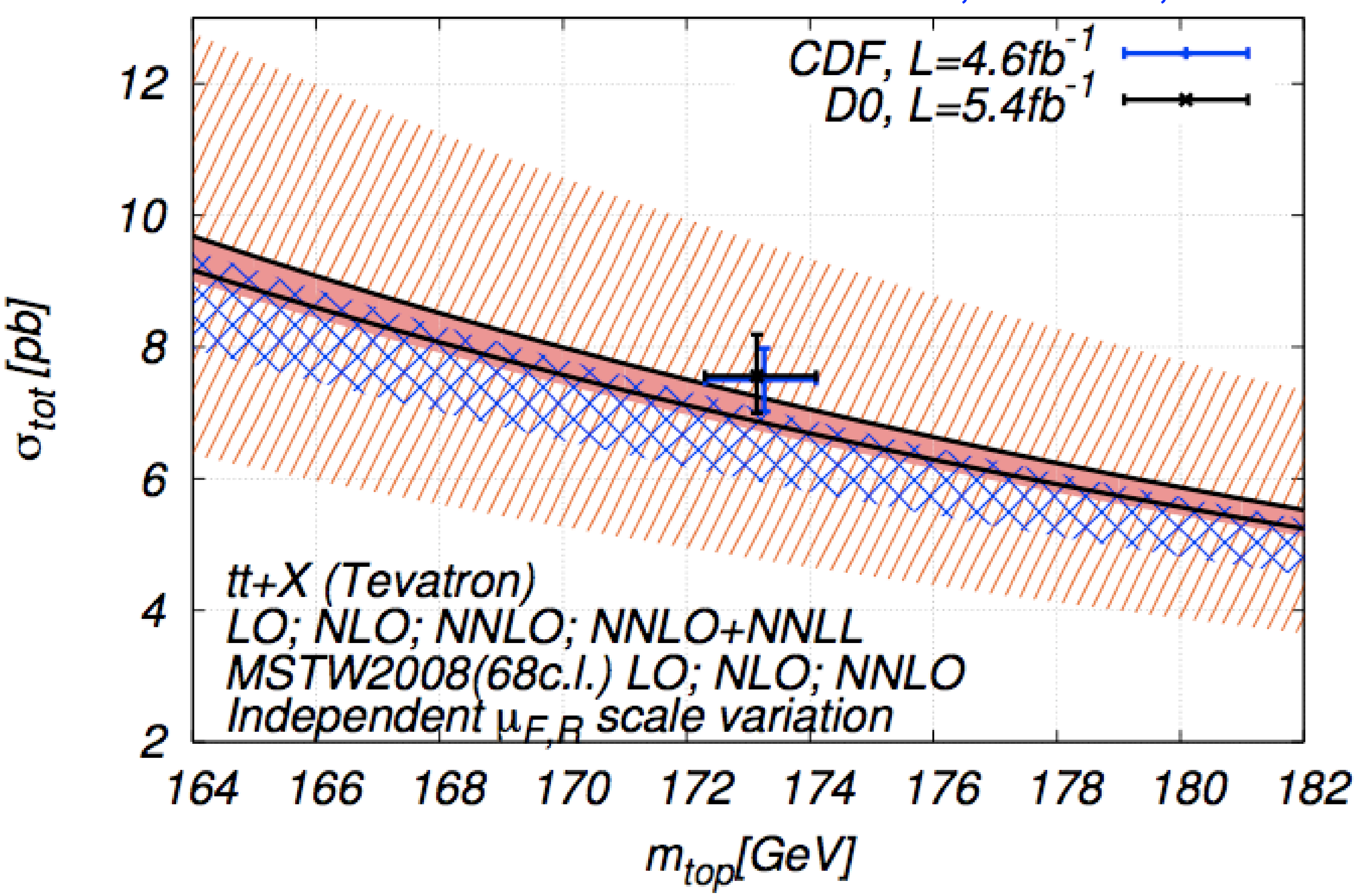
Parton-level NLO

- Automated 1-loop approaches
  - **great flexibility** and range of applications; heavy use of numerical methods; may be slow
  - MadLoop: mostly limited by CPU time — Alwall et al
  - HELAC-NLO: e.g. tt+2 jets, tttt — Bevilacqua et al
  - GoSam: e.g. WW+2 jets, H+3 jets — Cullen et al
- Specialized codes
  - less flexibility; may contain more sophisticated treatments and **much faster**
    - MCFM: W/Z/H+2 jets, Wbb, Zbb, diboson, top processes — JC, Ellis, et al
    - VBFNLO: double and triple boson production, VBF — Arnold et al
    - Rocket/MCFM+: W+3 jets, WW+up to 2 jets — Melia et al
    - BlackHat: 4 jets, W/Z + up to 5 jets, photon+jets — Bern et al

NNLO progress

- Even more diagrams and singularities to consider at NNLO
    - beginning of era of **widespread availability of NNLO**: W, Z, H (for a while), diphoton, top pairs, Higgs+jet, jet production (in the last couple of years)

Barnreuther, Czakon, Mitov



decreasing error bands:

LO

NLO

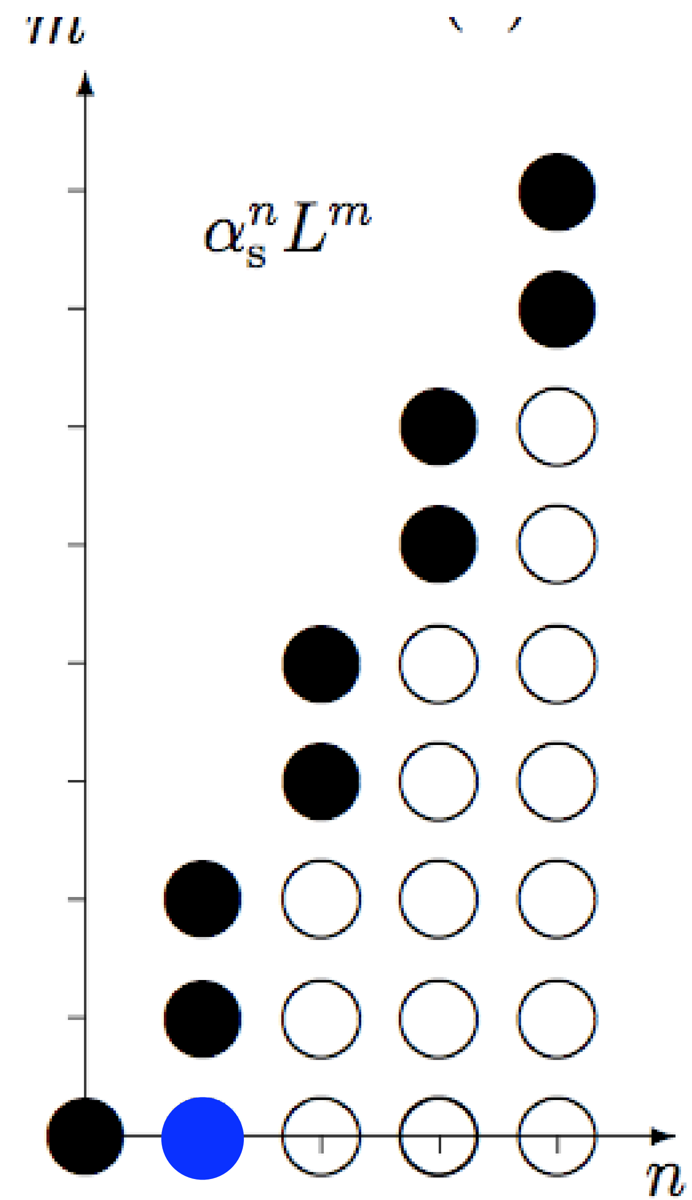
NNLO

NNLO + NNLL

(resummation)

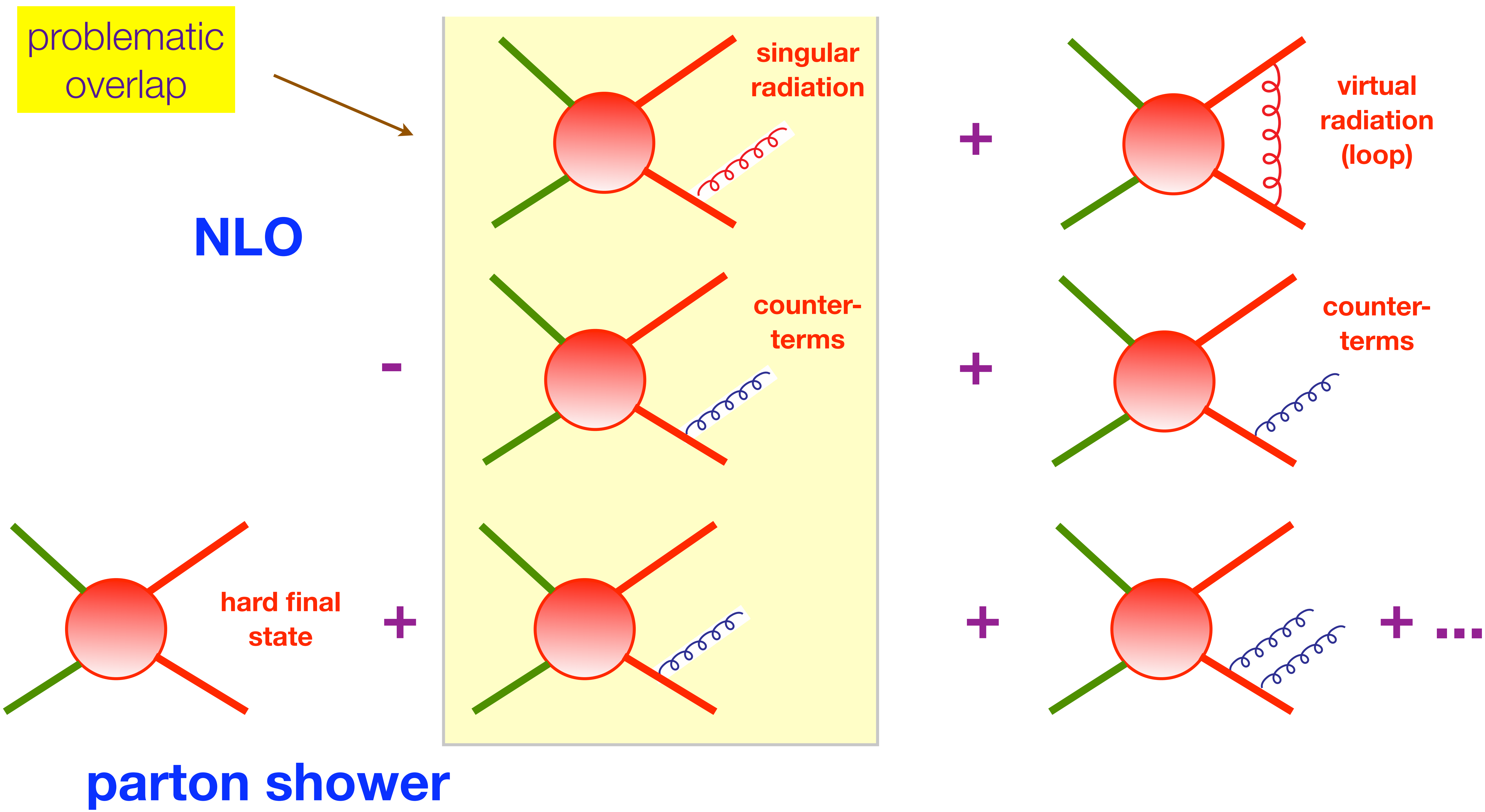
NLO + parton shower

- Many NLO parton-level predictions available - but most come without parton shower benefits.
  - do not produce unweighted events
  - hard to correct for detector effects
- Obvious problem:
  - NLO already includes one extra parton emission.
  - the hard part of this can be matched as before.
  - the soft/collinear part contains singularities that must be accounted for in the usual way
- **Technical challenge now overcome**.



PS + NLO inclusive

NLO + parton shower overlap



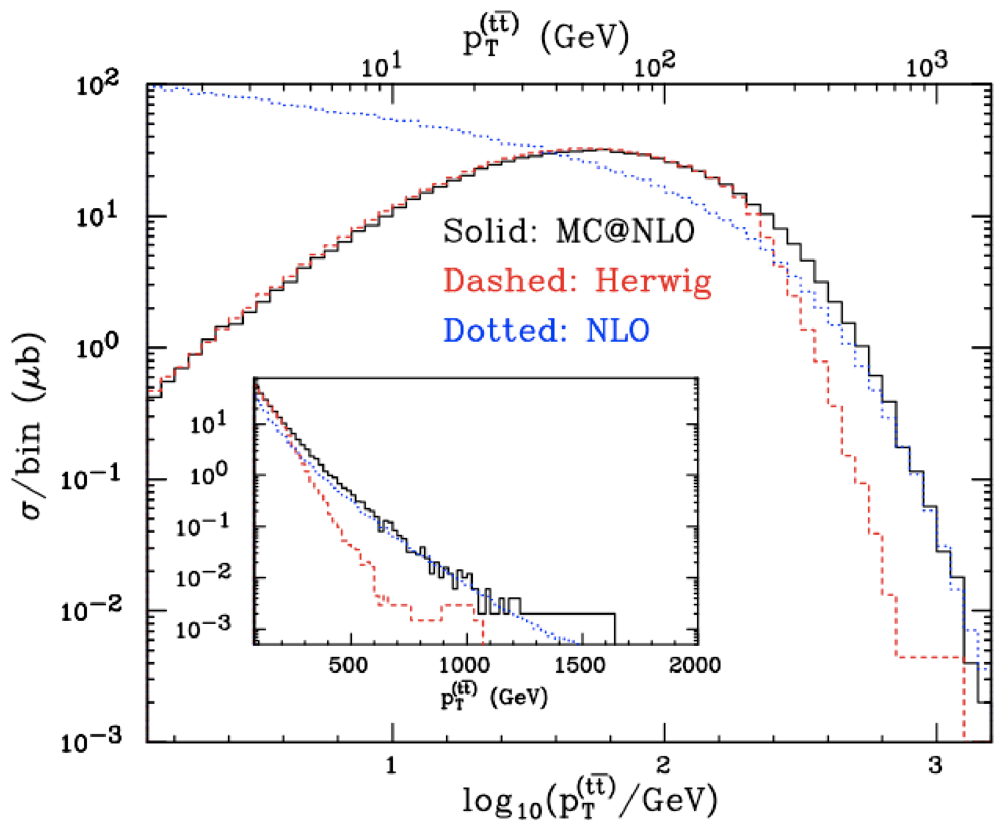
Solution

- Two main methods for dealing with the overlap: POWHEG and MC@NLO.
  - although formally of the same accuracy, differences do occur (see Nason and Webber, arXiv:1202.1251)

- POWHEG (Frixione, Nason, Oleari)
  - "local K-factor", real exponentiated in Sudakov, **positive-weight generator**
  - many processes implemented in POWHEG-BOX; same overall framework for each process, but many contributions from different theorists

- MC@NLO (Frixione, Webber)
  - usual Sudakov factor but **possibility of negative weights**
  - procedure has been automated in aMC@NLO (Alwall et al)
  - access to NLO+parton shower predictions at unprecedented level

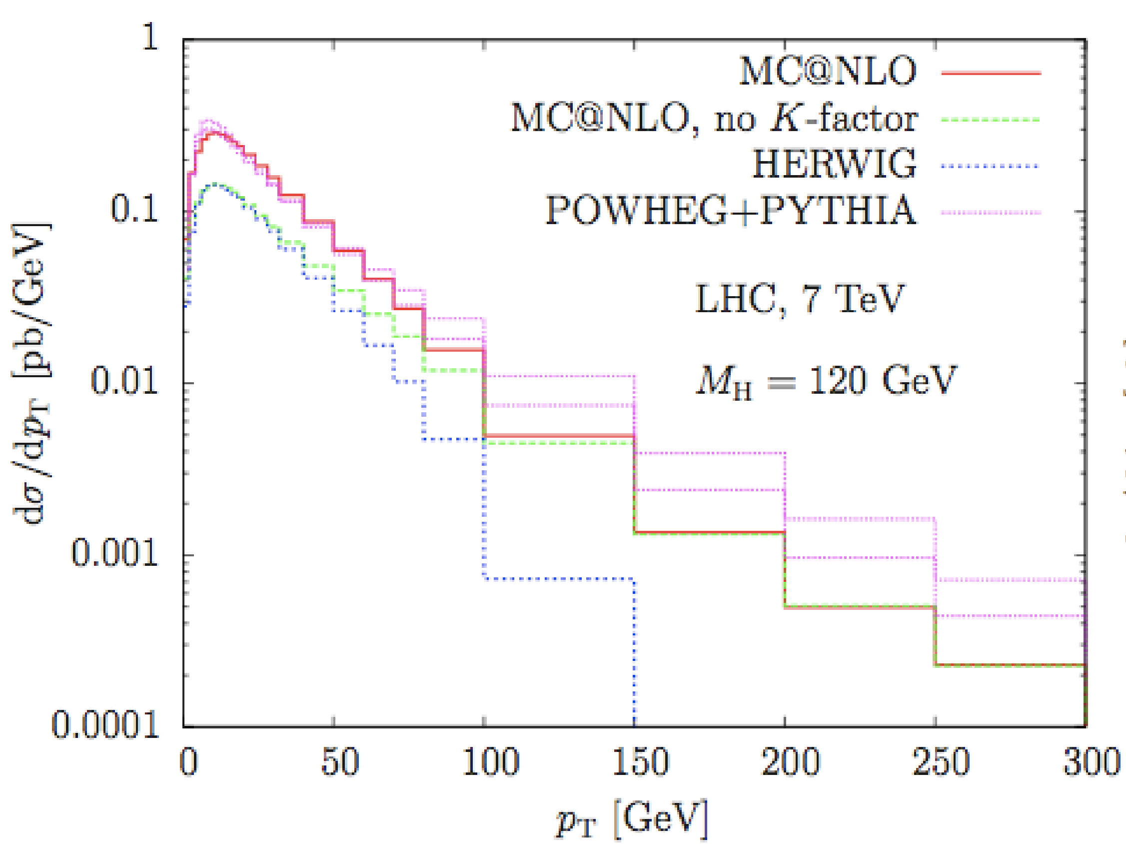
NLO+PS in action

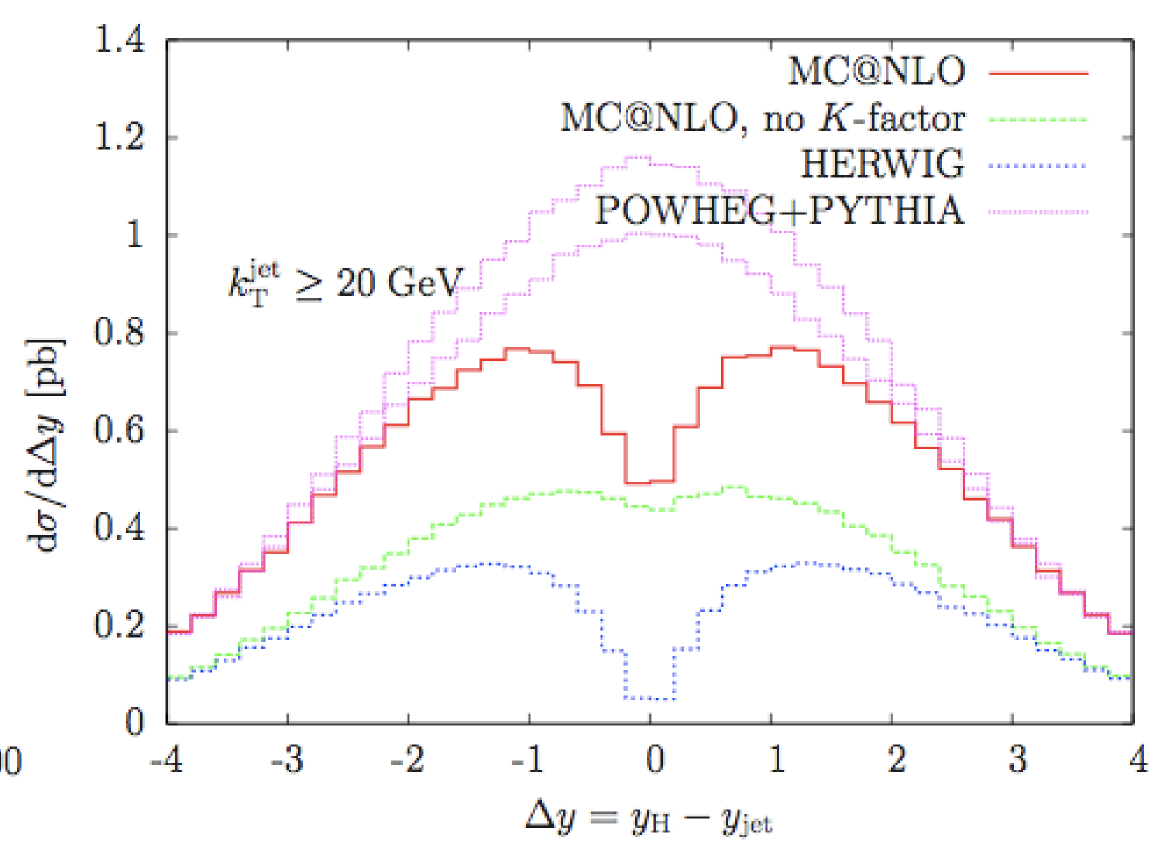


transverse momentum of top pairs, from MC@NLO

Frixione and Webber (2003)

differences in regions where NNLO important

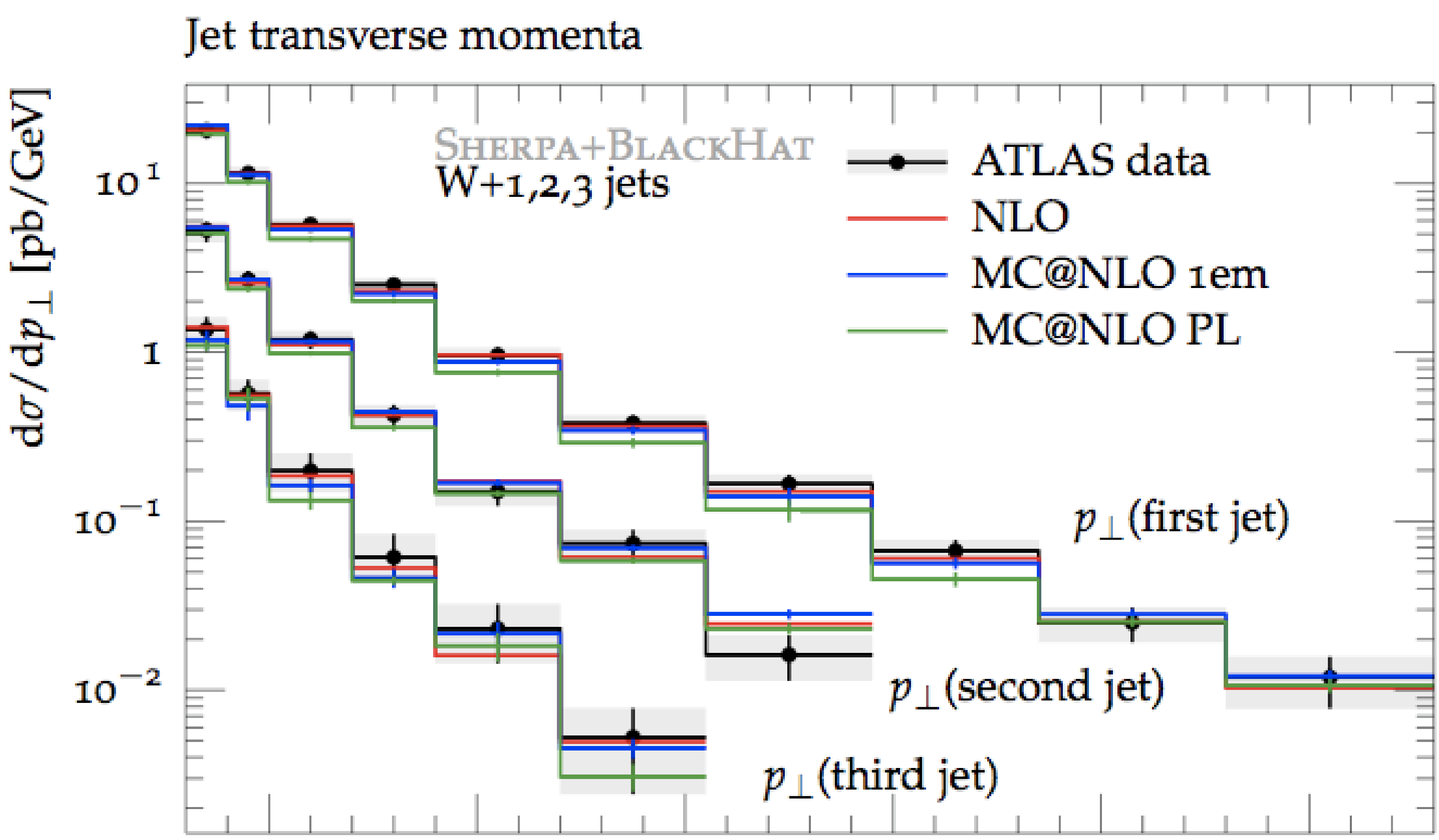




SHERPA

- POWHEG and MC@NLO methods also used in **SHERPA**.
- Recent application to W+1,2,3 jets.

Höche, Krauss, Schönherr, Siegert



- Smooth interpolation between POWHEG, MC@NLO procedures.

Beyond MC@NLO and POWHEG

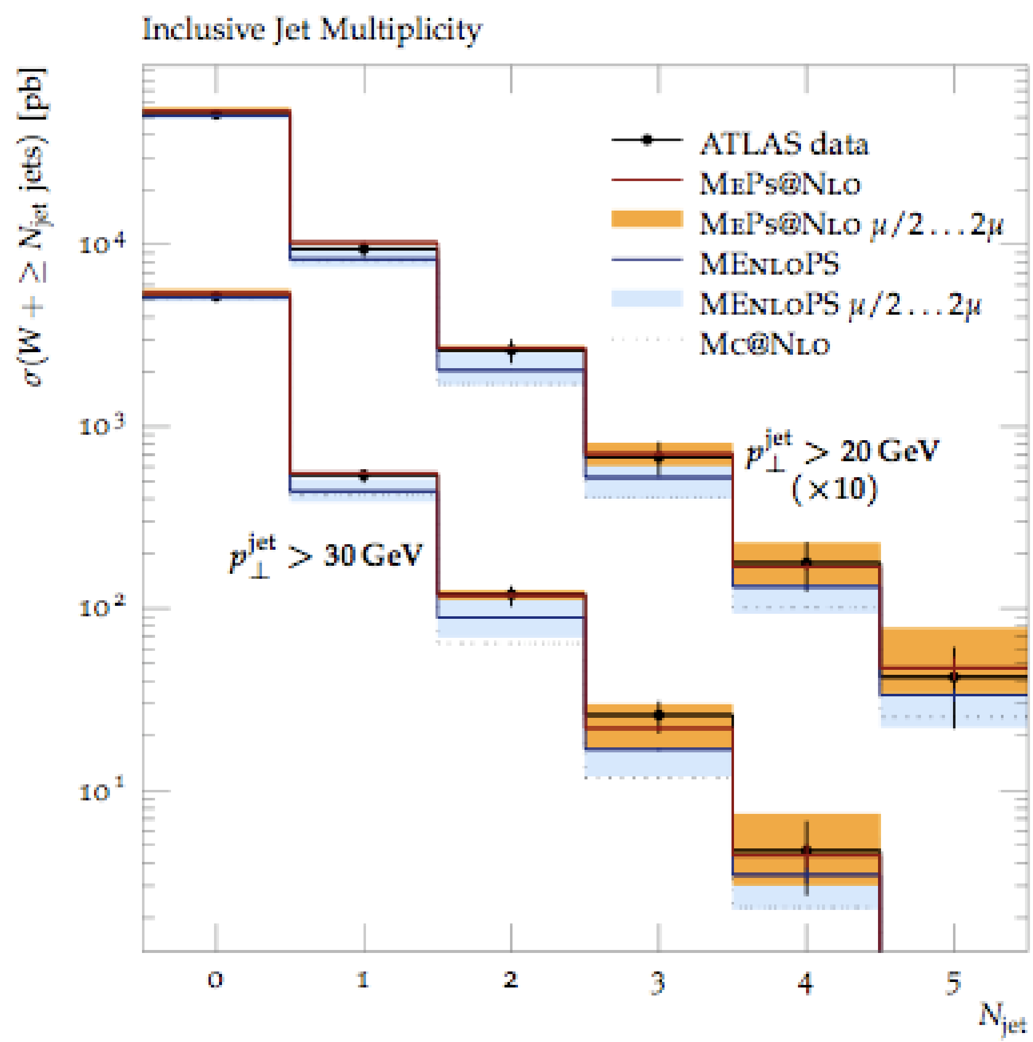
- NLO inclusive cross section, exact matrix elements for further jets.
    - “MENLOPS” and “ME&TS”.

      Hamilton, Nason;
      Höche, Krauss, Schönherr, Siegert

- **NLO precision for each jet emission**: merging samples that are each matched to correct NLO
    - “MEPS@NLO”

      Höche, Krauss, Schönherr, Siegert

- Many recent developments and ongoing intense activity.



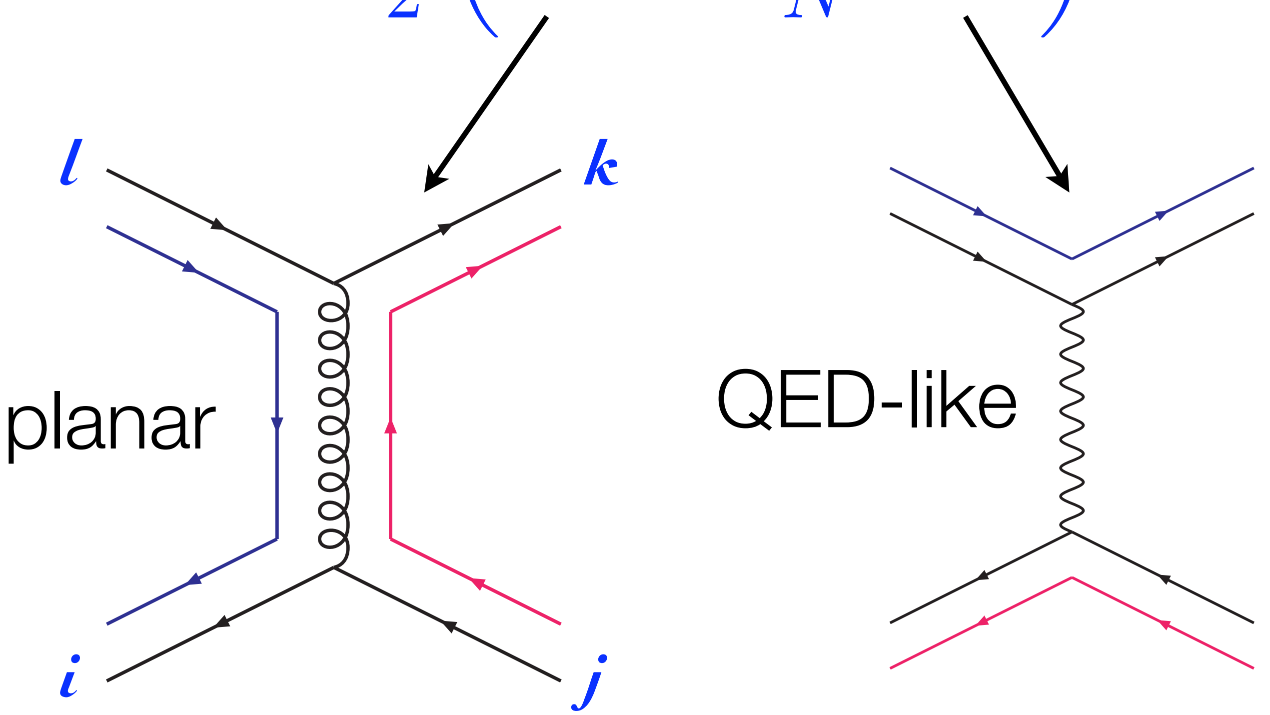
MENLOPS: W+0 jet NLO, W+1,2,3,4 LO

MEPS@NLO: W+0,1,2 jets NLO, W+3,4 LO

Crystal ball

- **Next frontier**: parton shower + NNLO?
  - problem: NNLO calculations very slow and numerically delicate
  - will get there eventually, but cannot reasonably expect much progress in the next few years
- **Orthogonal direction**: improvements in the parton shower evolution
  - example: treatment of color, where parton showers usually work in the "leading color approximation"
  - also called "planar"; gluon = (quark, antiquark) in terms of color.
  - will eventually need this level of precision too

$$t^A_{ij} t^A_{kl} = \frac{1}{2}\left(\delta_{il}\delta_{jk} - \frac{1}{N}\delta_{ij}\delta_{kl}\right)$$



Summary

- At the parton level:
    - NLO calculations standard, mostly available in an automated form
    - more complicated NLO calculations available in standalone code
    - NNLO calculations becoming available for an array of very important cases
- Modern parton showers come in many flavors:
    - capability for multi-jet merged samples (MLM, CKKW)
    - NLO matched to first emission (MC@NLO, POWHEG)
    - NLO matched to first emission, multi-jet merged (ME&TS, MENLOPS)
    - NLO matched for many jets (MEPS@NLO)
- Availability and maturity of predictions in that order