QCD vs. Monte Carlo Event Generators

http://www.terascale.de/mc

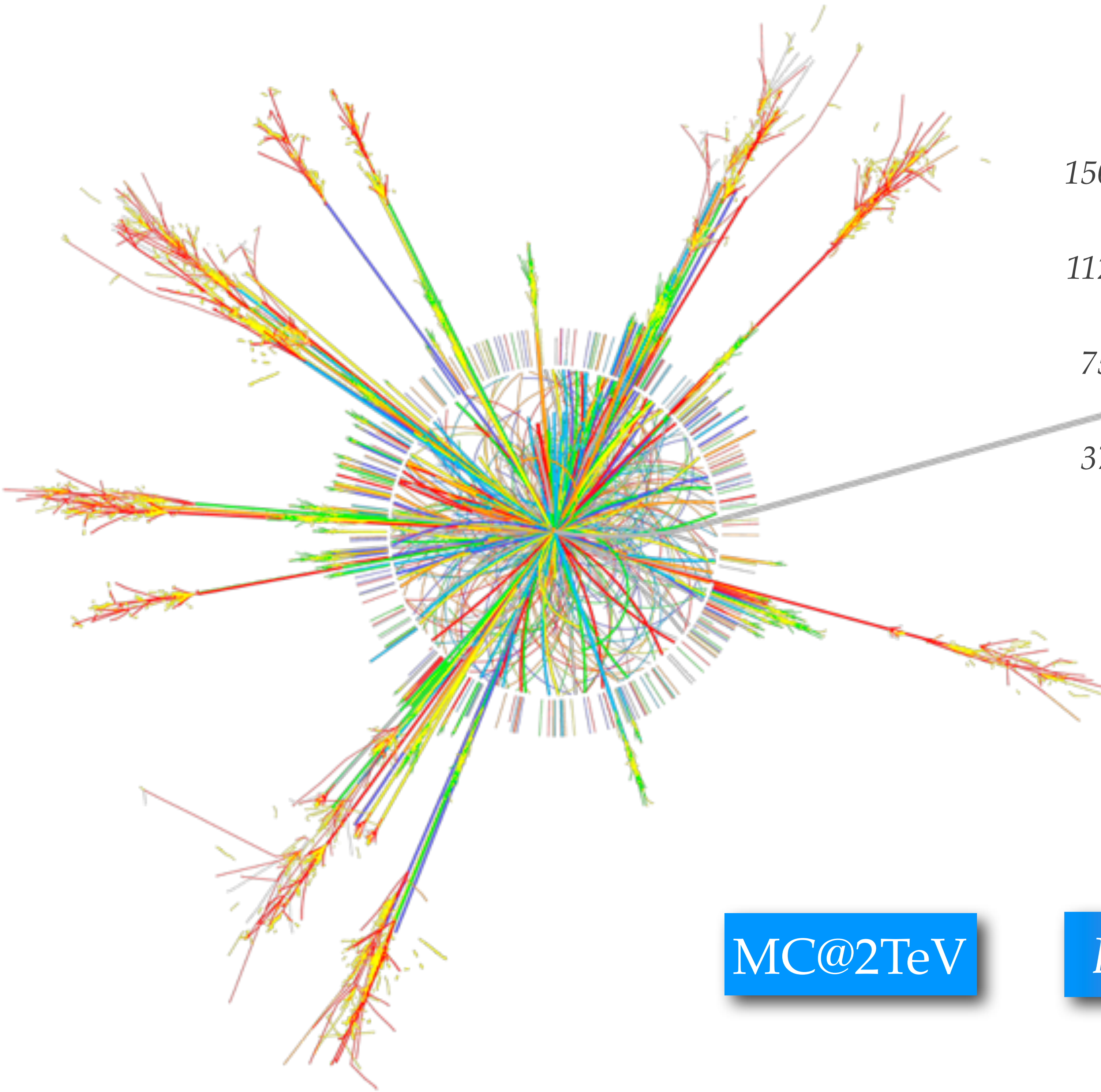
Zoltán Nagy
DESY, Terascale Analysis Center

in collaboration with Dave Soper

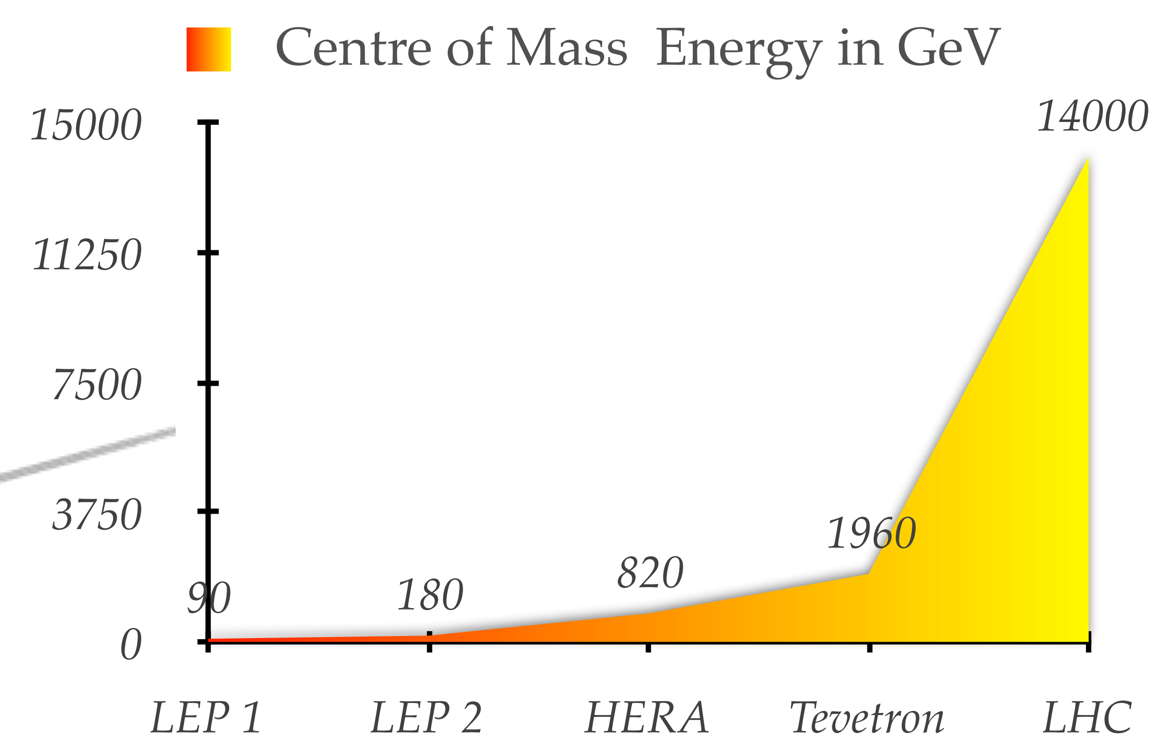
-- MC Meeting -- -- December 4, 2009,Karlsruhe ---

Introduction

The LHC is almost running and we will have to deal with the data soon.



Picture: ATLAS simulation



MC@2TeV → *Extrapolation* → MC@14TeV

1996 Tevatron Era

In 1996 both CDF and D0 found big discrepancies between the data and the theory in the one-jet inclusive cross section. Some people were thinking about new physics. That time we didn't have systematic error estimates for the PDFs. The next dialog might have happened between an experimentalist (EXP) and a self respecting theorist (SRT):

EXP: "We JUST need A PDF that can describe the data."

SRT: "WRONG! You have to make predictions in perturbative QCD and understand how much variation is allowed by the data and theory in the gluon distribution."

EXP: "But we estimate the uncertainty by using CTEQ and MRST."

SRT: "How can you estimate and control this uncertainty systematically if you don't have the tool that can systematically consider error of the PDF functions"

EXP: .....

And the theorist was right. There wasn't new physics. It turned out the gluon distribution wasn't well constrained by the DIS data.

D. Soper, ICHEP-98, **Nucl.Phys.Proc.Suppl.**54A:97-101

2009 LHC Era

The LHC is running and recently CMS have seen the first collisions. It seems to us that we are ready to discover new physics. Just like in 1996 the experimentalists and the theorists have very similar discussions:

EXP: "We JUST need A PROGRAM that can describe the data."

SRT: "WRONG, WRONG, WRONG! You need a program that can make predictions in perturbative QCD."

EXP: "But we estimate the uncertainties by running PYTHIA, HERWIG and their different versions."

SRT: "How can you estimate and control these uncertainties systematically if you don't have the tool that can make systematically improvable predictions?"

EXP: .....

We are looking forward to the data...

Parton Shower

The result and the derivation strongly depends on the shower algorithm, so it is useful to stick at one. My choice an shower algorithm with quantum interference.

Z.N, D.E. Soper: JHEP 0709:114,2007; JHEP 0803:030,2008; JHEP 0807:025,2008

Now, the shower equation is

$$\frac{d}{dt}\big|\rho(t)\big) = [\mathcal{H}_{\mathrm{I}}(t) - \mathcal{V}(t)]\,\big|\rho(t)\big)$$

- Fully exclusive and systematical formulation of the parton shower
- Quantum interferences are considered properly
    - Color evolution
    - Spin correlations
- Full control over the kinematics
    - Mapping based on exact phase space factorization
    - Ordering in virtuality (this is the most natural ordering variable)

QCD vs. MC

Shower cross section

$$\sigma^S[F] = \big(F\big|\rho\big)$$

- It is an all order but approximated calculation
- Based on soft and collinear factorization of the amplitudes
- Usually more approximation considered (e.g: large Nc,...)
- Implemented in general purpose MC programs (HERWIG, PHYTIA,...)
- *Sums up large logarithms*

QCD cross section

$$\sigma^{QCD}[F]$$

- It is an all order but approximated calculation
- Based on soft and collinear factorization of the amplitudes
- Precise in color
- Case-by-case rather elaborate calculation
- *Sums up large logarithms, correctly*

⇨ *Let us compare them!*

QCD vs. MC

Shower cross section

$$\sigma^S[F] = (F|\rho)$$

- It is an all order but approximated calculation
- Based on soft and c… factorization of the…
- Usually more appr… considered (e.g: lar…
- Implemented in general purpose MC programs (Herwig, Phytia,...)
- *Sums up large logarithms*

QCD cross section

$$\sigma^{QCD}[F]$$

- It is an all order but approximated
- …and collinear …of the amplitudes
- …r
- …rather elaborate calculation
- *Sums up large logarithms, correctly*

Herwig has been tested for

- e+e-: thrust, C-parameter, Durham jet rates, jet mass distribution, ...
- DIS, DY: large x

Let us compare them!

QCD vs. Parton Shower

Recent paper by Marchesini and Dokshitzer indicates that the color dipole based showers are not consistent with the parton evolution picture. They studied the quark energy distribution.

This has been checked both analytically and numerically and the shower is consistent with the DGALP equation.

Z.N, D.E. Soper: JHEP 0905:088,2009;

P. Skands, S. Weinzierl: arXiv:0903:2150

From shower equation

$$\frac{d}{dt}\big(x,q\big|\mathcal{U}(t,t')\big|M_2\big) = \big(x,q\big|[\mathcal{H}_I(t)-\mathcal{V}(t)]\,\mathcal{U}(t,t')\big|M_2\big)$$

No approximation and assumptions.
Only algebraic manipulations.

to DGLAP

$$\frac{d}{dt}D_q(t,t',x) = \int_x^1 \frac{dz}{z} P_{qq}(z) D_q(t,t',x/z) + \mathcal{O}(e^{-t})$$

QCD: Drell-Yan pT distribution

The NLL expression of the pT distribution was obtained using the renormalization group technique and the result is

$$\frac{d\sigma}{d\boldsymbol{p}_\perp dY} \approx \int \frac{d^2\boldsymbol{b}}{(2\pi)^2}\, e^{i\boldsymbol{b}\cdot\boldsymbol{p}_\perp}$$

$$\times \sum_{a,b} \int_{x_a}^{1} \frac{d\eta_a}{\eta_a} \int_{x_b}^{1} \frac{d\eta_b}{\eta_b}\, f_{a/A}(\eta_a, C^2/\boldsymbol{b}^2)\; f_{b/B}(\eta_b, C^2/\boldsymbol{b}^2)$$

$$\times \exp\left(-\int_{C^2/\boldsymbol{b}^2}^{M^2} \frac{d\boldsymbol{k}_\perp^2}{\boldsymbol{k}_\perp^2}\left[A(\alpha_s(\boldsymbol{k}_\perp^2))\log\left(\frac{M^2}{\boldsymbol{k}_\perp^2}\right) + B(\alpha_s(\boldsymbol{k}_\perp^2))\right]\right)$$

$$\times \sum_{a',b'} H^{(0)}_{a'b'}\; C_{a'a}\left(\frac{x_a}{\eta_a}, \alpha_s\left(\frac{C^2}{\boldsymbol{b}^2}\right)\right) C_{b'b}\left(\frac{x_b}{\eta_b}, \alpha_s\left(\frac{C^2}{\boldsymbol{b}^2}\right)\right) .$$

$$x_A = \sqrt{\frac{M^2}{s}}\, e^{Y}$$

$$x_B = \sqrt{\frac{M^2}{s}}\, e^{-Y}$$

$$C = 2e^{-\gamma_E}$$

where

$$A(\alpha_s) = 2\,C_F\,\frac{\alpha_s}{2\pi} + 2\,C_F\left\{C_A\left[\frac{67}{18} - \frac{\pi^2}{6}\right] - \frac{5\,n_f}{9}\right\}\left(\frac{\alpha_s}{2\pi}\right)^2 + \cdots \quad ,$$

$$B(\alpha_s) = -4\,\frac{\alpha_s}{2\pi} + \left[-\frac{197}{3} + \frac{34 n_f}{9} + \frac{20\pi^2}{3} - \frac{8 n_f \pi^2}{27} + \frac{8\zeta(3)}{3}\right]\left(\frac{\alpha_s}{2\pi}\right)^2 + \cdots \quad ,$$

$$C_{a'a}(z, \alpha_s) = \delta_{a'a}\delta(1-z) + \frac{\alpha_s}{2\pi}\left[\delta_{a'a}\left\{\frac{4}{3}(1-z) + \frac{2}{3}\delta(1-z)\left(\pi^2 - 8\right)\right\} + \delta_{ag}\, z(1-z)\right]$$

MC: Drell-Yan process

Now, the shower equation is

$$\frac{d}{dt}\big(\hat{\boldsymbol{p}}, Y\big|\rho(t)\big) = \big(\hat{\boldsymbol{p}}, Y\big|\mathcal{H}_{\rm I}(t) - \mathcal{V}(t)\big|\rho(t)\big)$$

After some harmless approximations, algebraic manipulations and about 2 months of hard work the result is

$$\frac{d\sigma}{d\boldsymbol{p}_\perp\, dY} = \int \frac{d\boldsymbol{b}}{(2\pi)^2}\, e^{{\rm i}\boldsymbol{p}_\perp\cdot\boldsymbol{b}} \exp\left\{-C_{\rm F}\int_{C^2/\boldsymbol{b}^2}^{M^2} \frac{d\boldsymbol{k}^2}{\boldsymbol{k}^2}\,\frac{\alpha_{\rm s}(\lambda\boldsymbol{k}^2)}{\pi}\left[\log\frac{M^2}{\boldsymbol{k}^2} - \frac{3}{2}\right]\right\}$$
$$\times \sum_{a,b} H^{(0)}_{a,b}\, f_{a/A}\left(x_{\rm A}, \frac{C^2}{\boldsymbol{b}^2}\right)\, f_{b/B}\left(x_{\rm A}, \frac{C^2}{\boldsymbol{b}^2}\right)$$

With the support of the *standard DGLAP equation* for the PDFs:

$$\mu_F^2 \frac{d}{d\mu_F^2} f_{a/{\rm A}}(x, \mu_F^2) = \sum_{\hat{a}} \int_0^1 \frac{dz}{z}\, \frac{\alpha_{\rm s}(\mu_{\rm F}^2)}{2\pi} P_{\hat{a},a}(z)\, f_{\hat{a}/{\rm A}}(x/z, \mu_F^2)$$

MC: Drell-Yan process

The result is strongly depends on the choice of the *argument of the α_s in the shower*:

$$\frac{\alpha_s(\lambda \boldsymbol{k}_\perp^2)}{2\pi} = \frac{\alpha_s(\boldsymbol{k}_\perp^2)}{2\pi} - 2\beta_1 \log(\lambda) \left(\frac{\alpha_s(\boldsymbol{k}_\perp^2)}{2\pi}\right)^2 + \mathcal{O}(\alpha_s^3)$$

Using scaled transverse momentum for the argument of strong coupling with

$$\lambda = \exp\left(-\frac{C_A\left[67 - 3\pi^2\right] - 15 n_f}{3\,(33 - 2\,n_f)}\right)$$

we can reproduce the QCD cross section at NLL level

$$A^{MC}(\alpha_s) = 2\,C_F\,\frac{\alpha_s}{2\pi} + 2\,C_F\,\left\{C_A\left[\frac{67}{18} - \frac{\pi^2}{6}\right] - \frac{5\,n_f}{9}\right\}\left(\frac{\alpha_s}{2\pi}\right)^2 + \cdots \quad ,$$

$$B^{MC}(\alpha_s) = -4\,\frac{\alpha_s}{2\pi} + \cdots \quad ,$$

$$C_{a'a}^{MC}(z, \alpha_s) = \delta_{a'a}\delta(1 - z) + \cdots$$

Other choices

- ✗ The shower based on Catani-Seymour factorization fails to reproduce the analytic answer. *The shower result doesn't exponentiate in b-space.* Bad choice of the momentum mapping.

- ✗ One can fix this bad mapping but that can fix only the LL logs. NLL resummation fails due to the bad choice of the soft gluon distribution function.

$$A_{lk}(q) = \frac{\hat{p}_k \cdot q}{(\hat{p}_k + \hat{p}_l) \cdot q}$$

- ✗ What happens when we don't change anything (splitting function, mapping, correct interference terms, ...), but we use the *transverse momentum as evolution variable* (like in PHYTIA,...).

  - ➡ The result is *correct at LL level* but very likely that *it fails at NLL level*. Lack of angular ordering.
  - ➡ PHYTIA also uses (as far I can see) the partitioning function above.

- ✓ What happens when we don't change anything (splitting function, mapping, correct interference terms, ...), but we use the *emission angle as evolution variable* (like in HERWIG,...).

  - ➡ This gives the *right answer at NLL level.* HERWIG looks OK, but the momentum mapping make trouble at NLL level.

Summary

- It is important to test parton shower against resummed QCD calculation.
- This can help us to treat it more systematically.
- ✓ Our parton shower can sum up the pT logs at NLL level
- ✗ Unfortunately this algorithm hasn't been implemented,yet.
- ✗ *Don't use modified LO PDFs (LO* & LO**) and don't produce such creatures!*
- Need more work on testing parton showers systematically against known QCD results.
- The summation of the large logs in the parton shower algorithms is very sensitive even for the smallest change which formally looks OK.
- Need more work on color evolution, spin correlations, non-global effect, higher order corrections, ..., more theory work required.
- In principle shower *has a chance to* sum up all the *LL and the LO NLL* contributions.

  "has a chance to" ≠ "does"
- Shower is only an "exponentiated LO" (*one can call it to eLO*) calculation.

Summary

- Non of the available algorithms is good for every observables.
- E.g.: Herwig is for a certain class of observables.
  - What is this class?
  - What are the constraints and limitations?
- Can we have a general and common shower framework?
- What would be the minimal requirements?
- ......

Conclusion from the PDF School

Conclusion from the PDF School

- I hope you have a different view on Monte Carlo programs after this talk. These programs are much more than just empirical tools. In principle they can predict cross section but you have to know their limitations.

Conclusion from the PDF School

- I hope you have a different view on Monte Carlo programs after this talk. These programs are much more than just empirical tools. In principle they can predict cross section but you have to know their limitations.

- Experimentalists usually say: "*We need just a program that can describe the data.*"

Conclusion from the PDF School

- I hope you have a different view on Monte Carlo programs after this talk. These programs are much more than just empirical tools. In principle they can predict cross section but you have to know their limitations.

- Experimentalists usually say: "*We need just a program that can describe the data.*"

- This is wrong, you should say: "*We need a program that can make (systematically improvable) predictions in perturbative QCD.*"

Conclusion from the PDF School

- I hope you have a different view on Monte Carlo programs after this talk. These programs are much more than just empirical tools. In principle they can predict cross section but you have to know their limitations.

- Experimentalists usually say: "*We need just a program that can describe the data.*"

- This is wrong, you should say: "*We need a program that can make (systematically improvable) predictions in perturbative QCD.*"

- E.g.: Would you trust in PDF function which has a modified DGLAP equation in order to have better agreement to a certain set of data?

Conclusion from the PDF School

- I hope you have a different view on Monte Carlo programs after this talk. These programs are much more than just empirical tools. In principle they can predict cross section but you have to know their limitations.
- Experimentalists usually say: "*We need just a program that can describe the data.*"
- This is wrong, you should say: "*We need a program that can make (systematically improvable) predictions in perturbative QCD.*"
- E.g.: Would you trust in PDF function which has a modified DGLAP equation in order to have better agreement to a certain set of data?

Please use the Monte Carlos wisely!