Monte Carlo tools for hadronic collisions

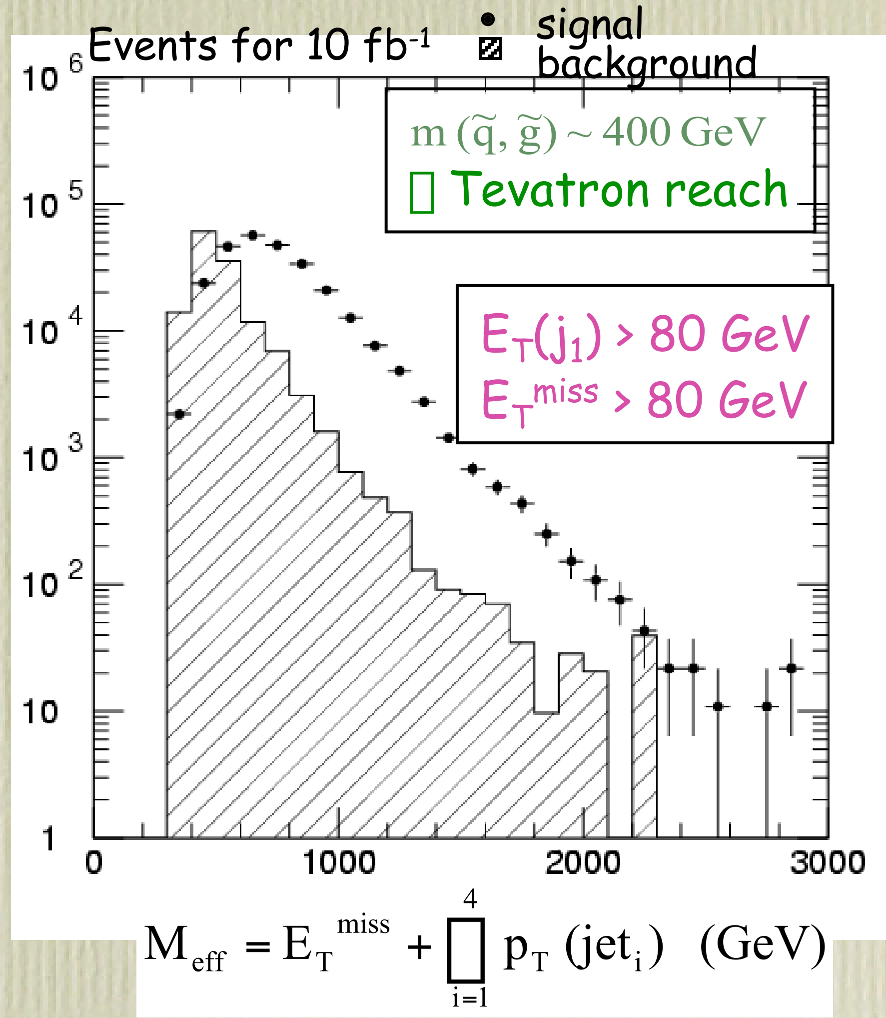
KITP Collider Physics Conference,
UCSB, Jan 12-16 2004

Michelangelo Mangano
TH Division
CERN, Geneva, Switzerland

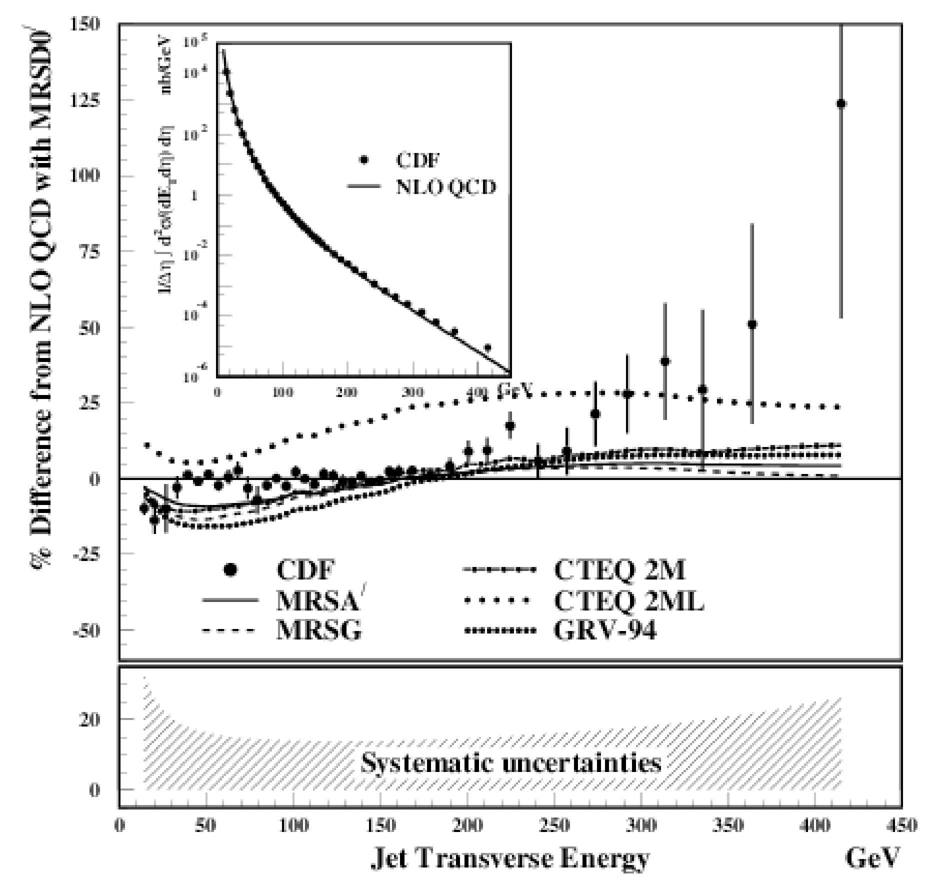
It is often said (and written) that we need better MC's to allow the discovery of new physics in future experiments

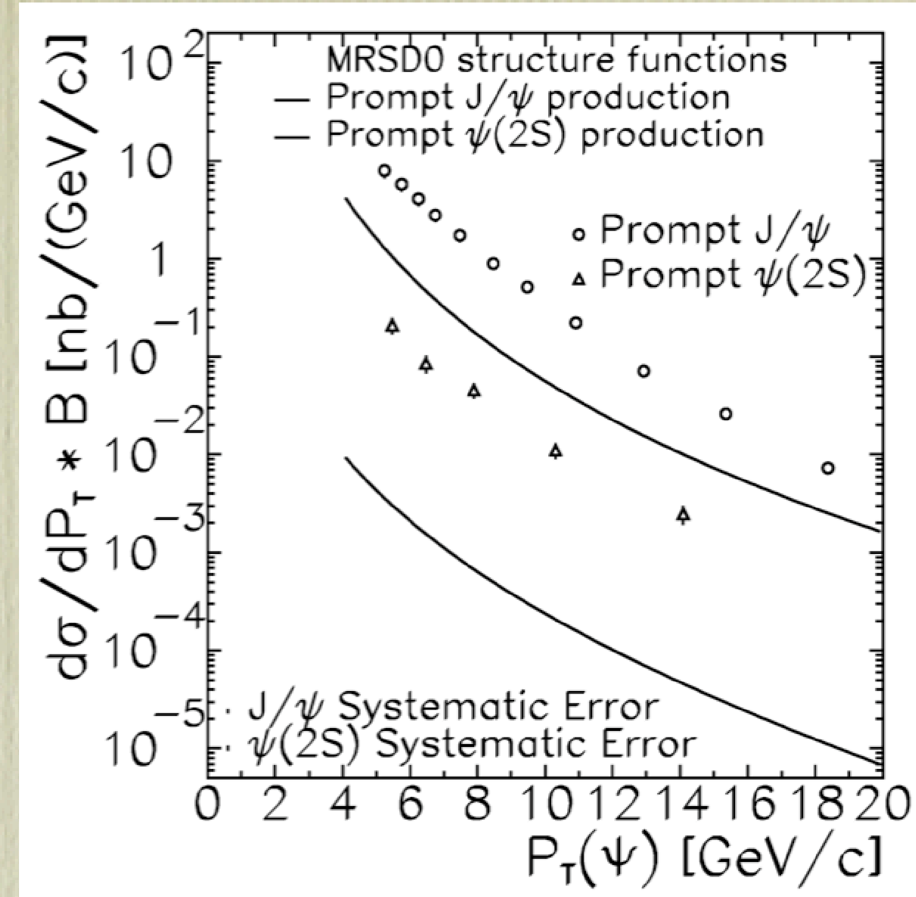
To my knowledge, no discovery in HEP has ever been been obtained by comparing data to a MC: the data always spoke for themselves

This **might** (need to) change with the LHC



Examples of good reasons not to trust a MC to claim a discovery:





Reality is often more complicated than a MC can predict. Unfortunately we find this out only when data and MC disagree. A discovery claim in this case is based on two assumptions being true at the same time: that a new phenomenon exists, and that our MC is right.

While we can debate whether MCs will ever be discovery tools, one point is beyond doubt:

MC's are essential to measure the properties of the (possibly new) objects being studied:

masses and cross sections

This requires control over the complete behavior of both signals (typically easier) and backgrounds (typically harder)

A good MC should be able to describe the data, having enough knobs to be tuned allowing proper fits (⇒ Ian's remarks yesterday)

A better MC should do so by just using first principles, rather than ad hoc models, to provide a clear relation between input parameters (physical constants) and observables

A good experimentalist should identify the best observables to tune the MC and improve their quality

A better experimentalist, in addition to being good, will work as much as possible without a MC, using it only as an auxiliary tool to extrapolate the knowledge obtained from control samples to the observable being studied

As a result of insufficient MC validation studies in hadronic collisions, I **do not** think we **have today a solid understanding of what the theoretical uncertainty is** for many important measurements that will be possible at the Tevatron and at the LHC

$\Delta^{th}(m_{top})$??

$\Delta^{th}(m_W)$??

$\Delta^{th}(\sin^2\theta_W)$??

$\Delta^{th}(\sigma_W)$??

....

Improvement of our tools, via **theoretical developments** and via strategies for the **validation of the theoretical systematics** is a crucial duty of our community

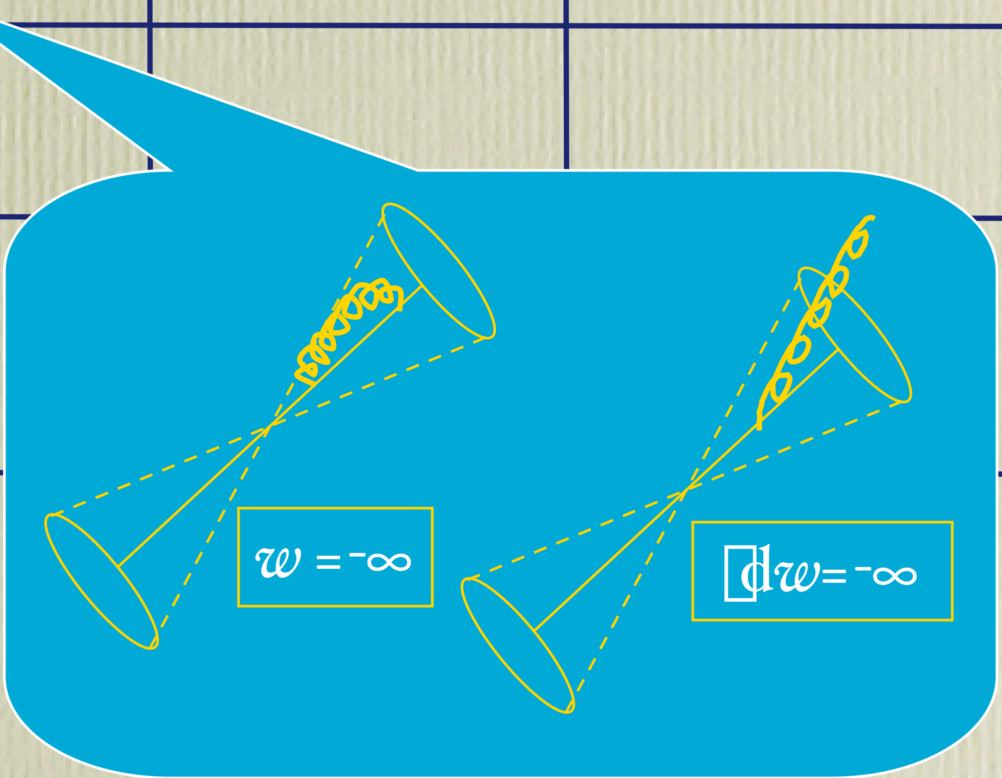
Three complementary approaches

	ME MC's	X-sect evaluators	Shower MC's
Final state description	Hard partons jets. Describes geometry, correlations, etc	Limited access to final state structure	Full information available at the hadron level
Higher order effects: loop corrections	Hard to implement, require introduction of negative probabilities	Straighforward to implement, when available	Included as vertex corrections (Sudakov FF's)

Kosower's talk

Three complementary approaches

	ME MC's	X-sect evaluators	Shower MC's
Final state description	Hard partons jets. Describes geometry, correlations, etc	Limited access to final state structure	Full information available at the hadron level
Higher order effects: loop corrections	Hard to implement, require introduction of negative probabilities	Straighforward to implement, when available	Included as vertex corrections (Sudakov FF's)



Three complementary approaches

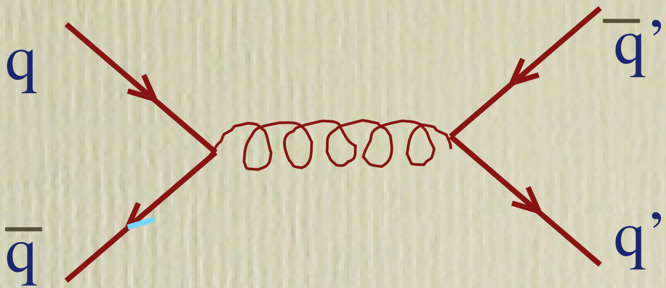
	ME MC's	X-sect evaluators	Shower MC's
Final state description	Hard partons jets. Describes geometry, correlations, etc	Limited access to final state structure	Full information available at the hadron level
Higher order effects: loop corrections	Hard to implement, require introduction of negative probabilities	Straighforward to implement, when available	Included as vertex corrections (Sudakov FF's)
Higher order effects: hard emissions	Included, up to high orders (multijets)	Straighforward to implement, when available	Approximate, incomplete phase space at large angle
Resummation of large logs	??	Possible, when available	Unitary implementation (i.e. correct shapes, but not total rates)

Sterman's talk

Huston's talk

2’ guide to shower MC’s

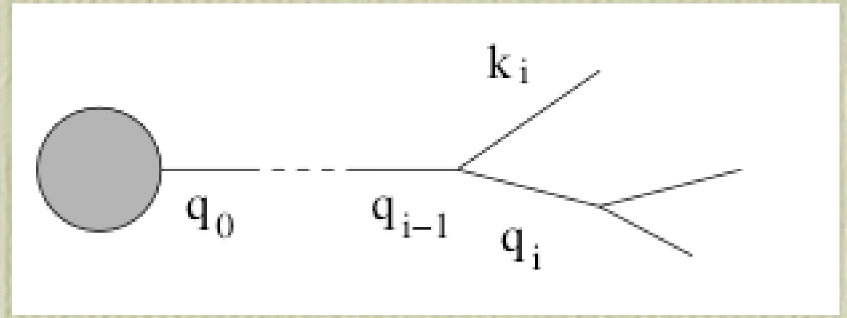
- After the generation of a given parton-level configuration (typically LO, 2 1 or 2 2) , each possible IS and FS parton-level evolution (***shower***) is generated, with probability defined by the shower algorithm (**unitary evolution**).



- ***Algorithm***: numerical, Markov-like evolution, implementing within a given appoximation scheme the QCD dynamics:
  - branching probabilities:
    - selection of evolution variables ⇒

Choice of shower-evolution variables

Q^2, z , α( )



$Q^2=$
q^2
k_i q_i
$q_T^{\ 2}$
...

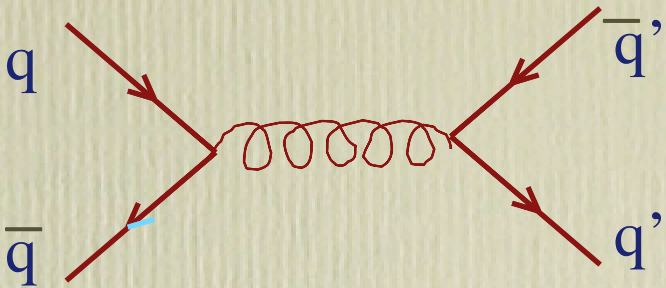
$z \approx p_i/p_{i-1}$, p=
E
$q^{//}$
$q^{//}+E$
...

$^2=$ f(z,Q)

While at LL all choices of evolution variables and of scale for αs are equivalent, more intelligent choices can lead to improved description of NLL effects and allow a more accurate and easy -to-implement inclusion of angular-ordering constraints and mass effects, as well as to a better merging of multijet ME’s with the shower

New work appeared recently identifying new, improved, evolution variables. Catani, Dittmaier&Trocsany,
Herwig++, Sherpa, Sjöstrand

2' guide to shower MC's

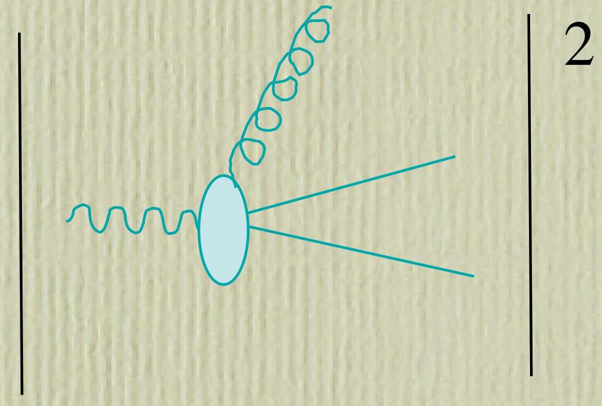


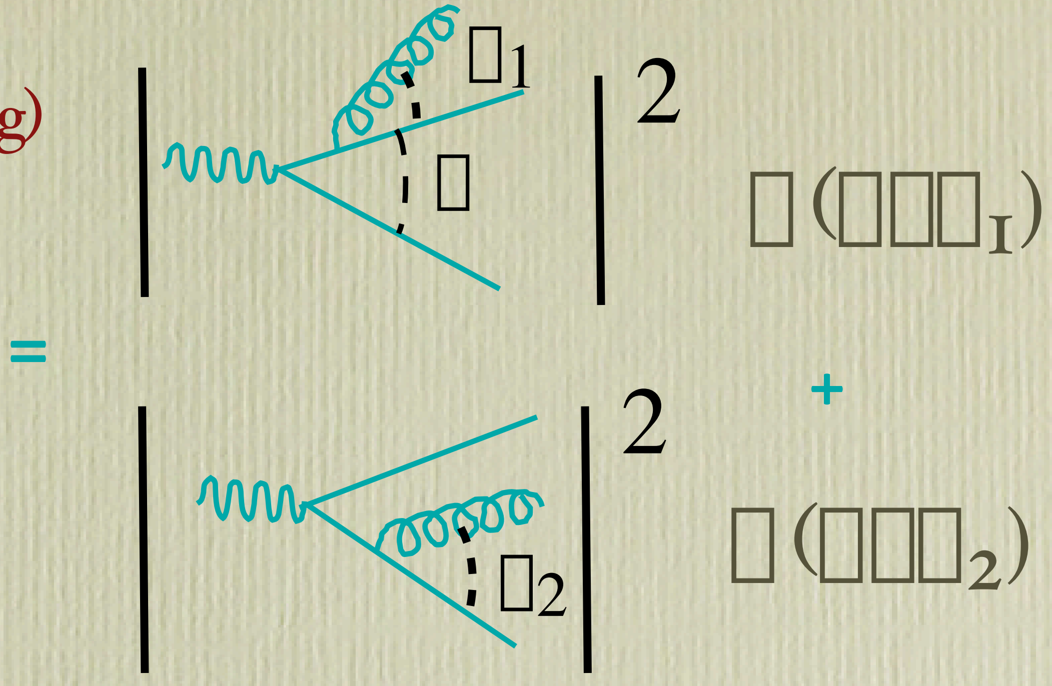
- After the generation of a given parton-level configuration (typically LO, 2 1 or 2 2) , each possible IS and FS parton-level evolution (***shower***) is generated, with probability defined by the shower algorithm (**unitary evolution**).

- ***Algorithm***: numerical, Markov-like evolution, implementing within a given appoximation scheme the QCD dynamics:
  - branching probabilities:
    - selection of evolution variables
    - implementation of quantum coherence ⇒

Solution

(a.k.a. angular ordering)





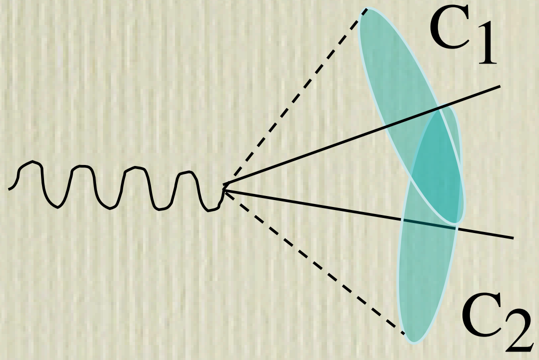
Limitations:

no emission outside $C_1 \cup C_2$:

- lack of hard, large-angle emission
- poor description of multijet events

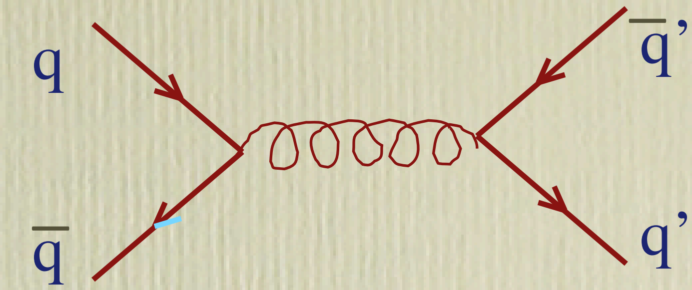
incoherent emission inside $C_1 \cap C_2$:

- loss of accuracy for intrajet radiation



2' guide to shower MC's

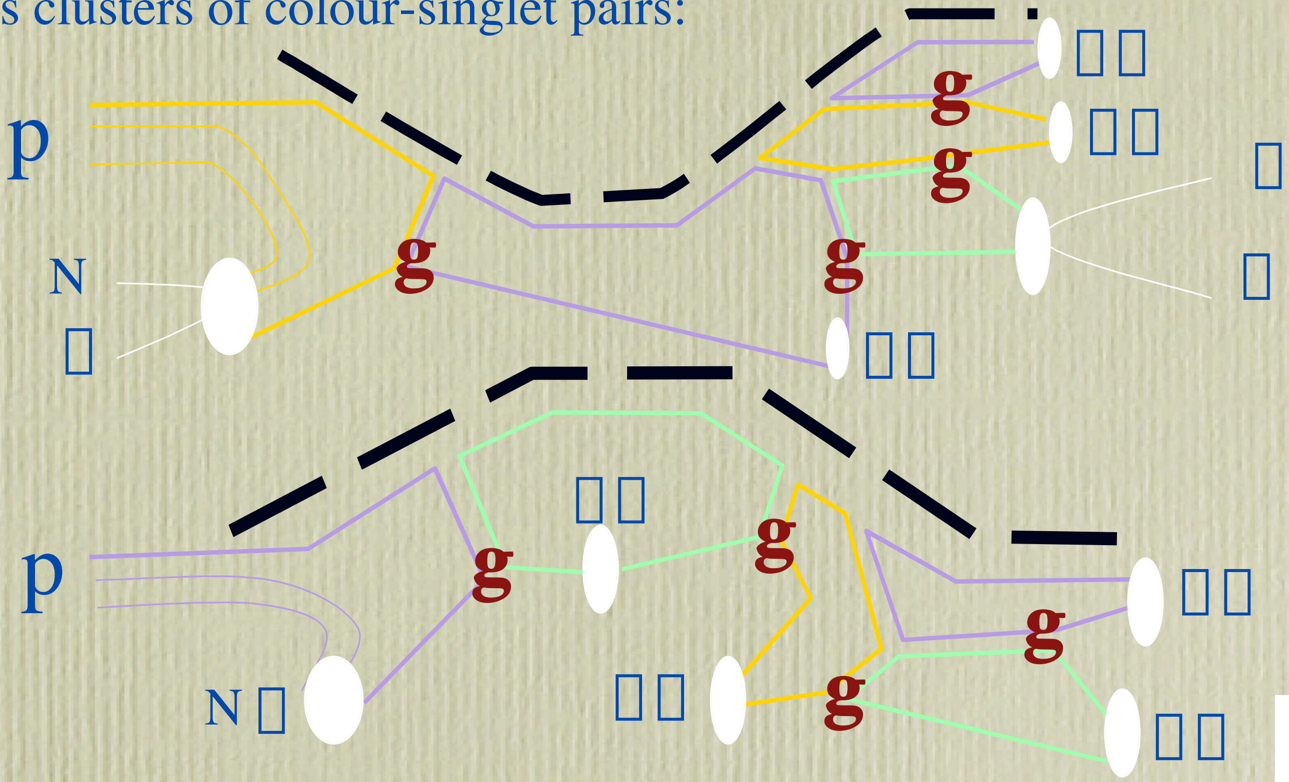
- After the generation of a given parton-level configuration (typically LO, 2 1 or 2 2) , each possible IS and FS parton-level evolution (***shower***) is generated, with probability defined by the shower algorithm (**unitary evolution**).



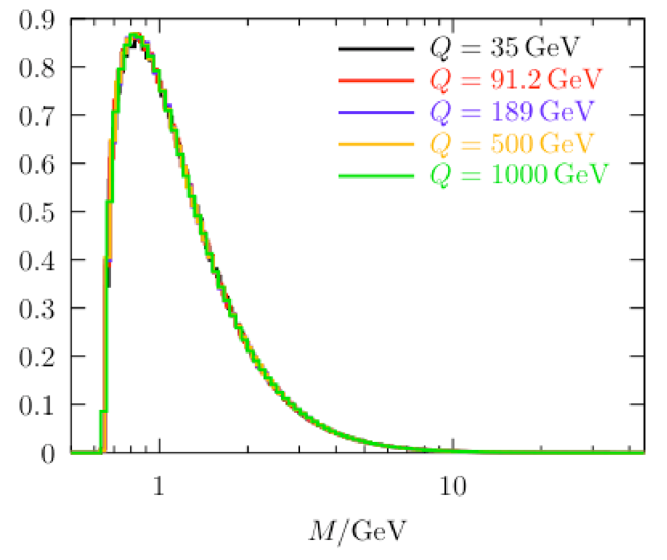
- ***Algorithm***: numerical, Markov-like evolution, implementing within a given appoximation scheme the QCD dynamics:
    - branching probabilities:
        - selection of evolution variables
        - implementaiton of quantum coherence
    - infrared cutoff scheme
    - hadronization model ⇒

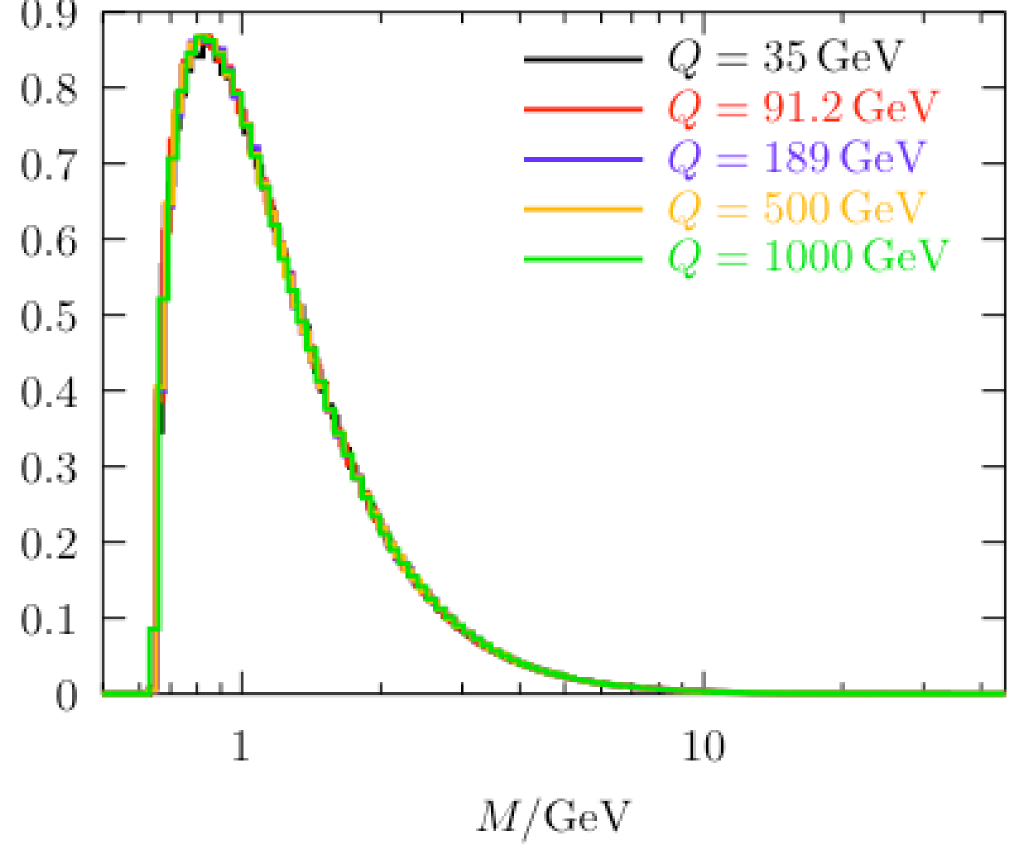
Hadronization

At the end of the perturbative evolution, the final state consists of quarks and gluons, forming, as a result of angular-ordering, low-mass clusters of colour-singlet pairs:



Thanks to the cluster pre-confinement, hadronization is local and independent of the nature of the primary hard process, as well as of the details of how hadronization acts on different clusters.
Models for hadronization can then be tuned on e+e- data at a given energy, and applied elsewhere





The existence of high-mass clusters, however rare, is unavoidable, due to IR cutoff which leads to a non-zero probability that no emission takes place. This is particularly true for evolution of massive quarks (as in, e.g. Z→bb or cc). Prescriptions have to be defined to deal with the "evolution" of these clusters. **This has an impact on the z→1 behaviour of fragmentation functions.**

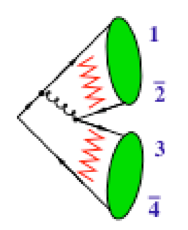
Phenomenologically, this leads to uncertainties, for example, in the background rates for H→γγ (jet→γ).

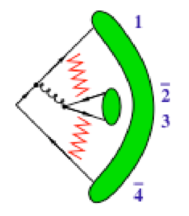
New cluster model

(Winter, Krauss, Soff,
hep-ph/0311085)

implementing:

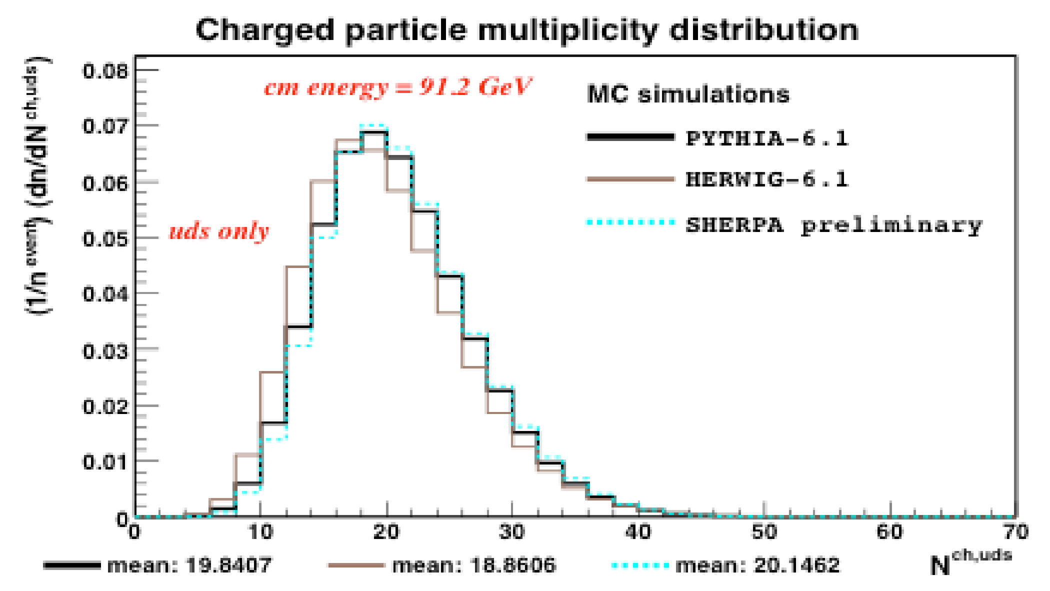
- colour reconnections ($1/N^2$ effects),
- flavour-dependent cluster evolution
- z-dependent non-perturbative gluon splitting

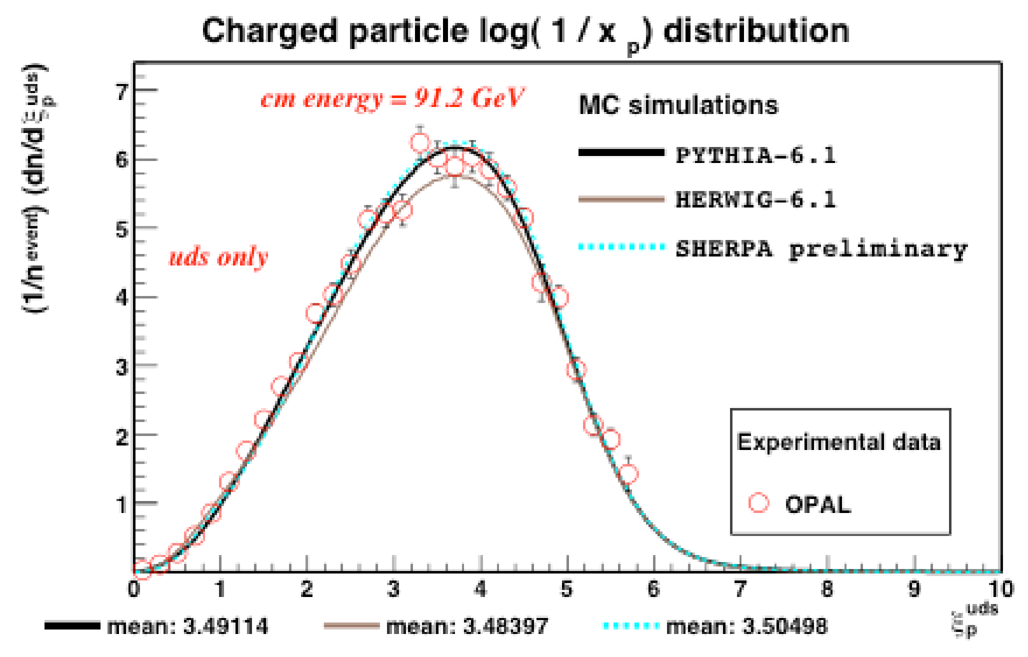


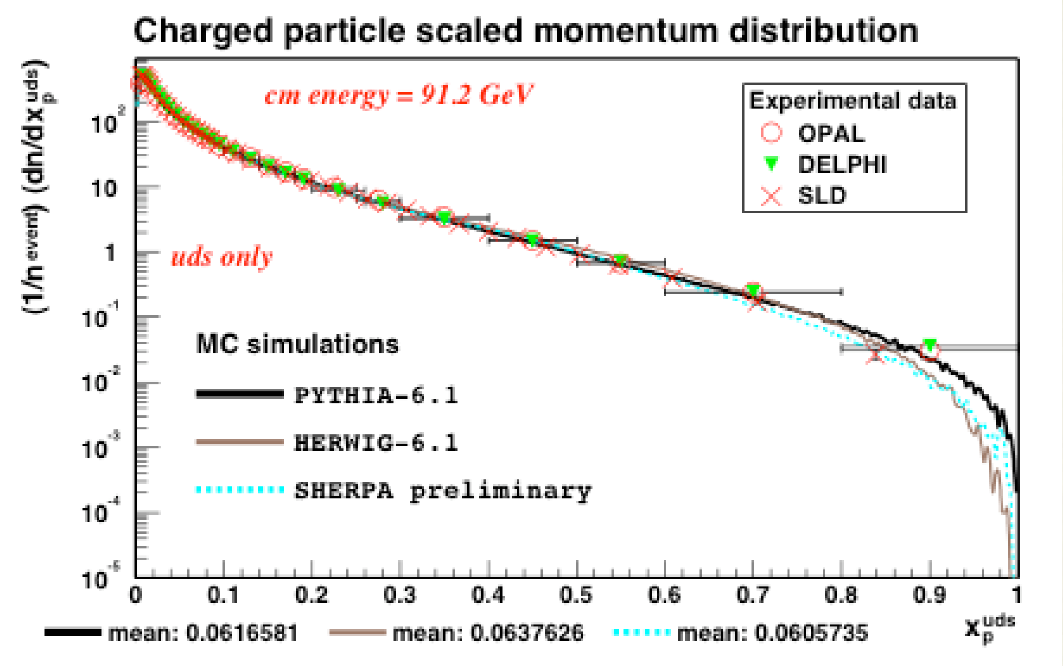


Leads to:

- lower cluster masses
- better description of z→1 region
- better description of <Nch>







Recent progress in MC-related tools

- New tools to calculate ME's for high multiplicity multijet final states (Alpgen, MacEvent, **2002**)

Codes available for:

- W/Z/gamma + N jets (N≤6)
- W/Z/gamma + Q Qbar + N jets (N≤4)
- Q Qbar + N jets (N≤4)
- Q Qbar Q' Q'bar + N jets (N≤2)
- Q Qbar H + N jets (N≤3)
- nW + mZ + kH + N jets (n+m+k+N ≤8, N≤2)
- N jets (N≤8)

ALPGEN: MLM, Moretti, Piccinini, Pittau, Polosa
MADGRAPH: Maltoni, Stelzer
CompHEP: Boos etal
VECBOS: Giele et al
NJETS: Giele et al
Kleiss, Papadopoulos
......

Example of complexity of the calculations, for gg-> N gluons:

Njets	2	3	4	5	6	7	8
# diag's	4	25	220	2485	34300	5x10^5	10^7

For each process, flavour state and colour flow (leading 1/Nc) are calculated on an event-by-event basis, to allow QCD-coherent shower evolution

Recent progress in MC-related tools

- New tools to calculate ME's for high multiplicity multijet final states (Alpgen, MadEvent, **2002**)
- New NLO parton-level event generators ⇨ **Campbell talk**
- NLO matrix elements in shower MC's (Dobbs (**2001**), Grace (**2002**), MC@NLO, (**2003**) ⇨ **Frixione talk**

On the role of NLO, NNLO, ...

- (N)NLO calculations are essential to extract reliable estimates of the total production rates
- It is highly non-trivial, however, to establish an accurate connection between what is calculated and what is observed.
- QCD physics at LEP taught us that the concept of IR and collinear safety, while essential to justify the use of fixed-order perturbative calculations, does not guarantee the accuracy of such calculations.
- The impact of power corrections, as well as of the resummation of large logs, is crucial for a faithful description of the data. This is true even at high-Q
- **A balance between perturbative accuracy and realism in the description of the physical observables (e.g. in the description of the structure of an experimental jet) is mandatory**

NLO results are available today for most processes of interest. The technique by Frixione and Webber allows their consistent merging with shower MC's. **Extension to NNLO is far** from being even just theoretically formulated, let alone numerically implemented.

Example: accuracy in the extraction of the W cross-section

- NNLO total X-sections known, residual theory uncertainty ~few%.
- MC necessary to evaluate acceptance, before the comparison with the inclusive calculation.
- New calculations available for the W differential distributions (Anastasiou, Dixon, Melnikov, Petriello, hep-ph/0312266, PRL 91 (03) 182002)
- Effects other than NNLO can have however an effect on acceptance more important than the NLO-NNLO difference. Keeping them into account in a NLO event generator could be more important than having the full NNLO EvGgen

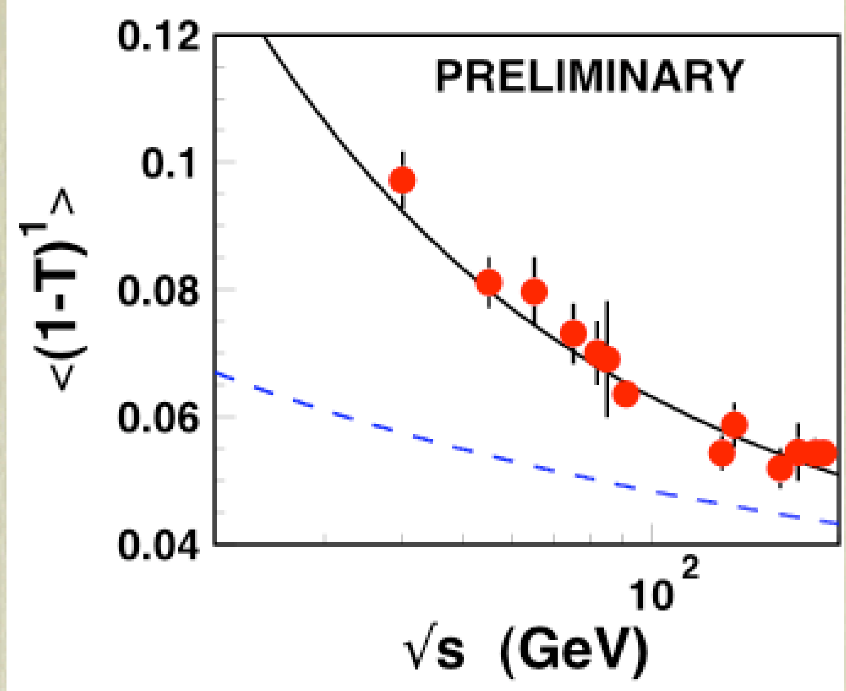
Acceptance for lepton with p_T>20 GeV and |η|<2.5, using different parameters or approximations: easily ~5% differences. Larger effects at the Tevatron, due to more limited acceptance

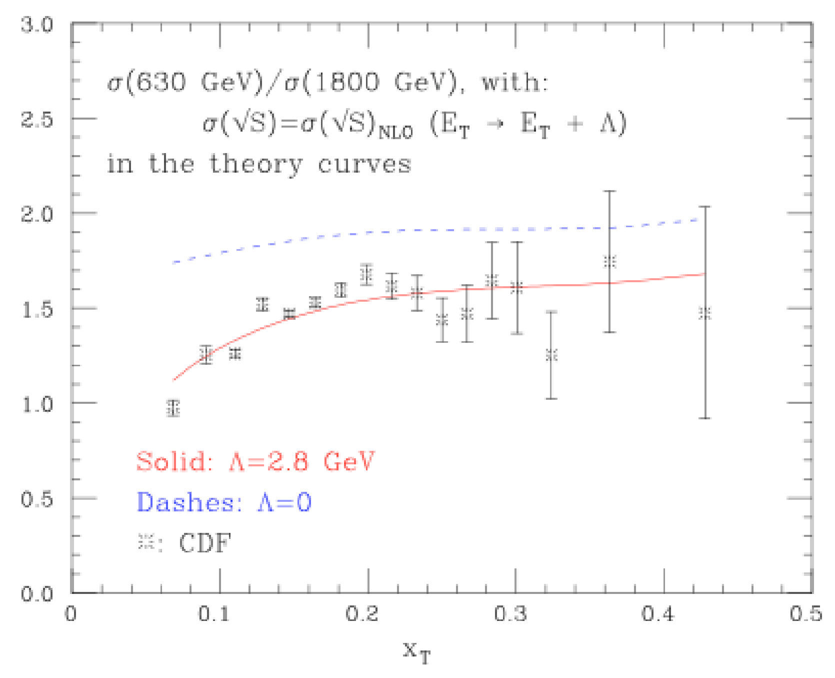
LO	LO, Γ_W=0	LO, no spin corr's	LO, PDF= CTEQ6.19
0.4890(2)	0.4971(2)	0.5259(2)	0.5245(2)

Effects induced by **ISR, hadronization**, etc need to be evaluated (MLM&Frixione, in progress)

Power corrections

- NLL description of `jet shapes’, and inclusion of power corrections (see LEP):
  - formalism established and tested with great success at LEP, where it provides an essential tool for the high-accuracy determination of αs
  - essential to extend the formalism to hadronic collisions, to exploit the lever arm in Q in the measurement of αs. Effects can be significant event at large ET, due to the rapidly falling spectra





Example: B cross section in PT

$$\frac{d\square(B)}{dp_T} = \int \frac{dz}{z} \frac{d\square(b)}{d\hat{p}_T} f(b \to B; z) \; , \qquad \hat{p}_T = p_T/z$$

known in PT to NLO+NLL resummation of collinear logs

fit to e+e- data, under the assumption of factorization and universality.

There is **no control** in PT over the **corrections to factorization and universality**. Not even the definition of the fragmentation variable z is uniquely fixed.

Corrections to total production rates will be small, of order $(\Lambda_{QCD}/m_b)^2 \approx 1\%$

Corrections to cross sections with cuts can be significantly larger!

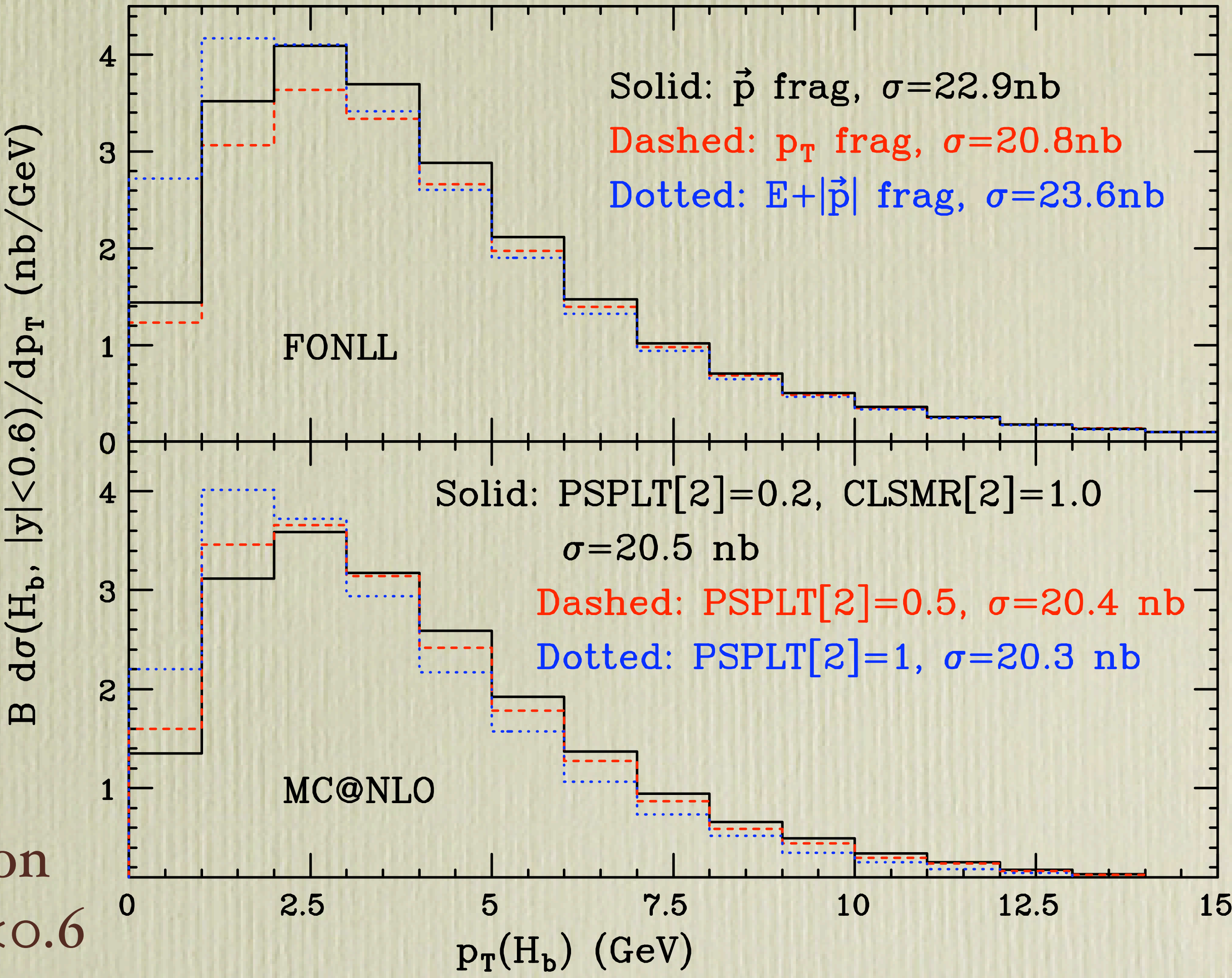
CDF measured recently σ(B) in the domain p_T(B)>0, |y|<0.6

Significant dependence of the NLO+NLL prediction on the definition of fragmentation variable:

$p_T(B) = z\, p_T(b)$, $y_B = y_b$

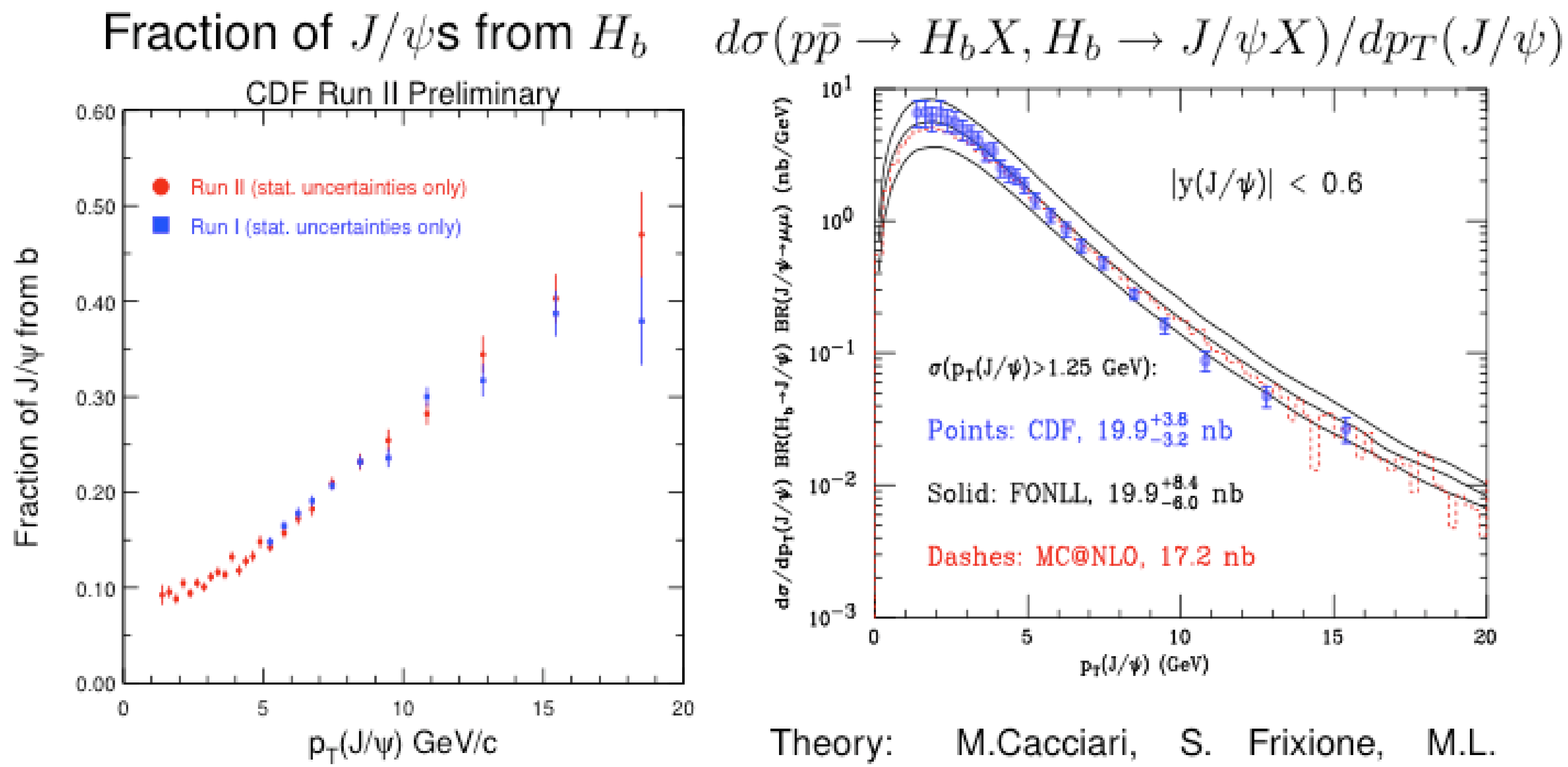
$\vec{p}(B) = z\, \vec{p}(b)$

$E+|\vec{p}|\,(B) = z\,(E+|\vec{p}|)\,(b)$



The shower MC prediction for the total rate within |y|<0.6 is instead very stable

Incidentally, in case you are interested here is the comparison between the latest CDF data at 1.96TeV and NLO+NLL QCD:



Fraction of J/ψs from H_b

$d\sigma(p\bar{p} \to H_b X, H_b \to J/\psi X)/dp_T(J/\psi)$

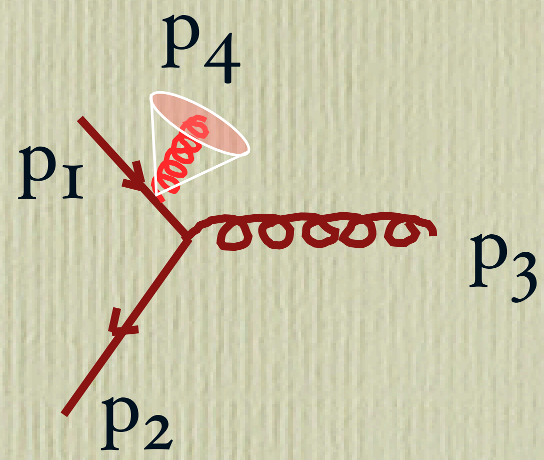
Theory: M.Cacciari, S. Frixione, M.L. Mangano, P. Nason. G. Ridolfi hep-ph/0312132

Mari Bishai (CDF) FNAL Wine&Cheese, Dec 2003

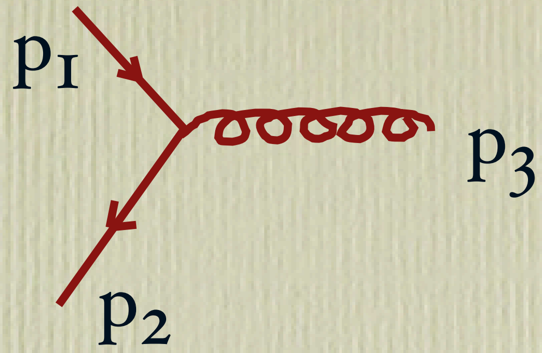
Recent progress in MC-related tools

- New tools to calculate ME's for high multiplicity multijet final states (Alpgen, MadEvent, **2002**)
- New NLO parton-level event generators ➪ **Campbell talk**
- NLO matrix elements in shower MC's (Dobbs (**2001**), Grace (**2002**), MC@NLO, (**2003**) ➪ **Frixione talk**
- New techniques for merging of multijet ME's and shower MC's (Catani, Krauss, Kuhn, Webber (**2001**), Lönnblad (**2002**), Mrenna&Richardson (**2003**))

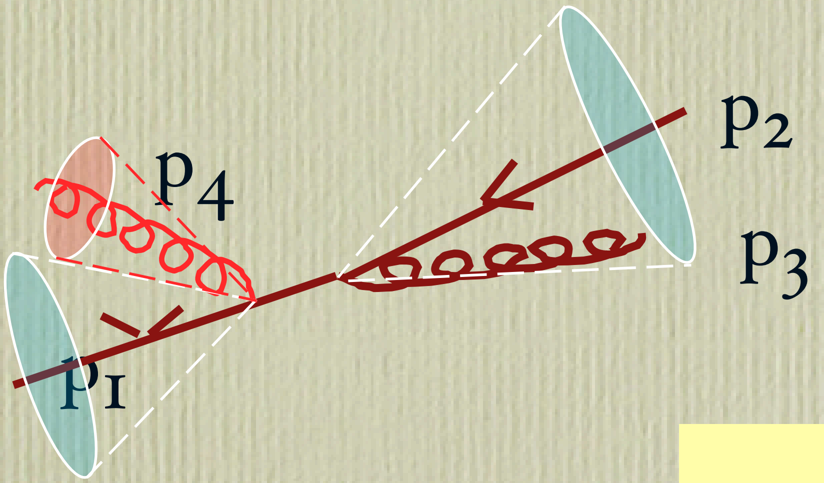
The problem: Leading vs subleading accuracy and double counting



is of $\mathcal{O}(\alpha_s)$ relative to the LO process



unless:

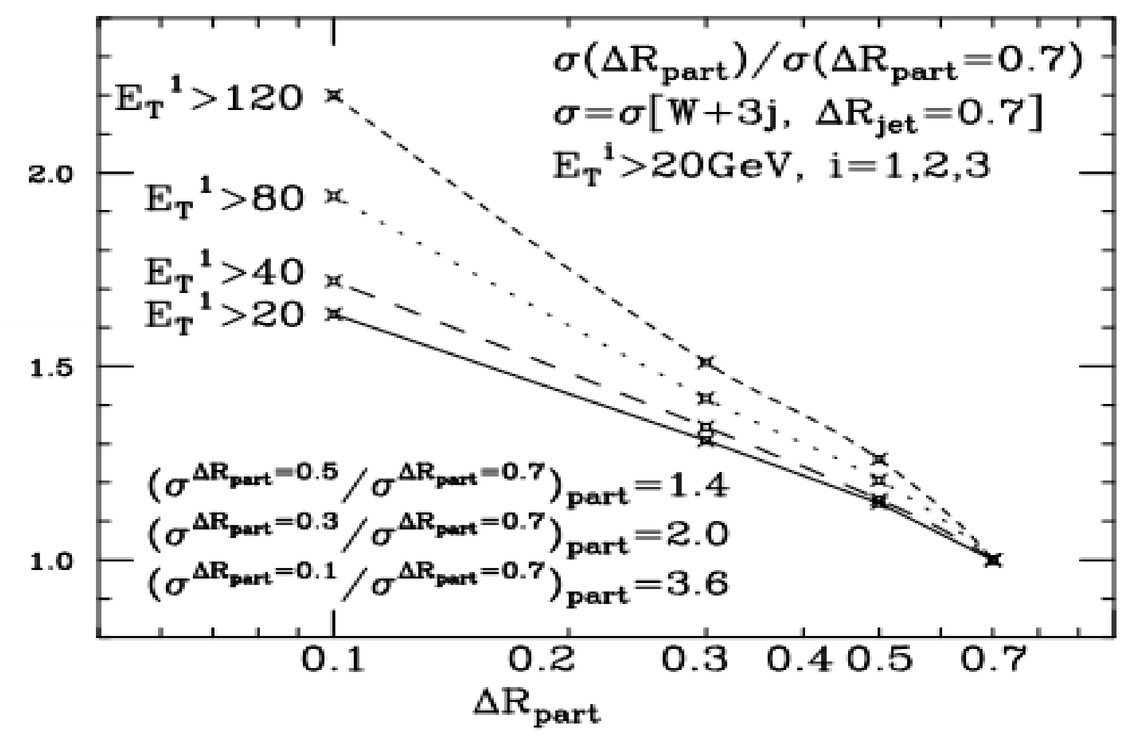


which gives a contribution to $\sigma_{3\text{-jet}}$ of order

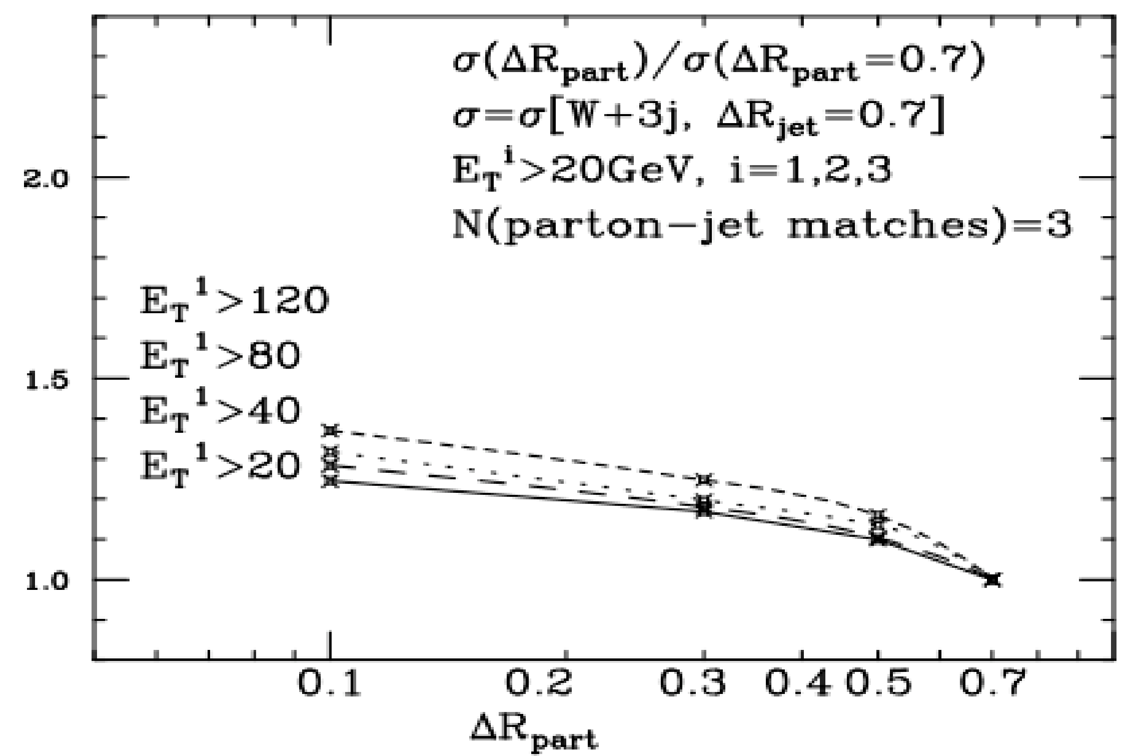
$$\alpha_s \log\frac{(p_2+p_3)^2}{E_{T\,jet}^2} \sim \alpha_s\left(\log\frac{p_T^{\max}}{p_T^{\min}} + \log\frac{1}{\Delta R}\right)$$

Double counting is sub-leading only if ΔR and $p_T^{\max}/p_T^{\min}$ are not too large

How large is too large? Cut-generation dependence, example



Dependence on generation ΔR in the spectrum of ΔR =0.7 jets



Same, after applying a (one among many possible) parton-jet matching requirement

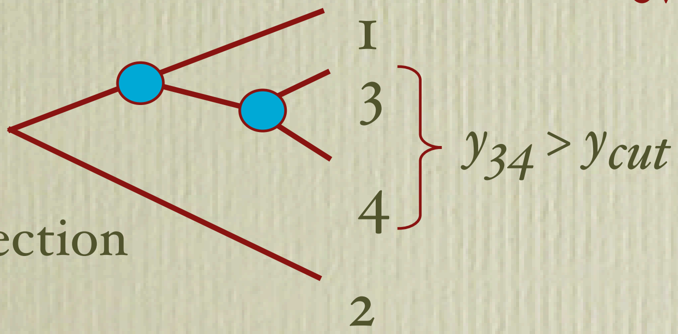
Progress towards solution: vetoed showers
(CKKM: Catani, Krauss, Kuhn, Webber)

- Generate samples of different jet multiplicities according to exact tree-level ME's, with N_{jet} defined using a k_{perp} algorithm

$$y_{ij} = \frac{2\min\left\{E_i^2, E_j^2\right\}\left(1 - \cos\theta_{ij}\right)}{s} \geq y_{cut} = \frac{Q_{cut}^2}{s}$$

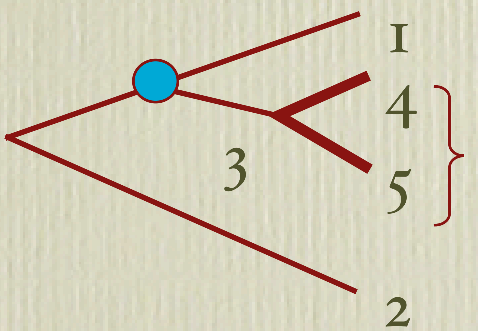
- Reweight the matrix elements by vertex Sudakov form factors, assuming jet clustering sequence defines the colour flow
- Remove double counting by vetoing shower histories (i.e. y_{ij} sequences already generated by the matrix elements)
- Fully successfull for e^+e^- collisions, being extended to hadronic collisions (Richardson, Krauss, Mrenna, Alpgen)

From the sample of 4-hard-parton events



● : Sudakov correction

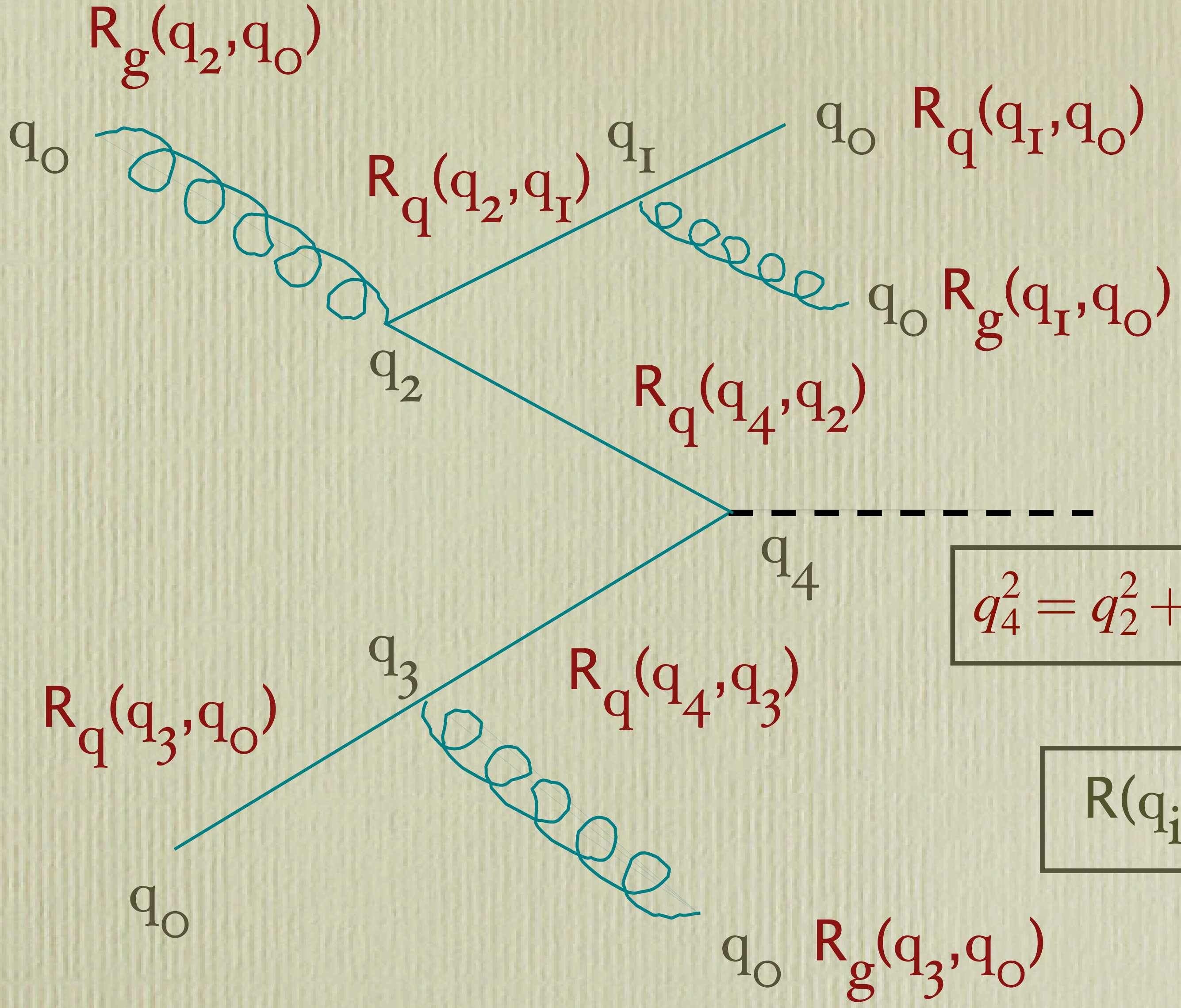
From the sample of 3-hard-parton events



(splitting rejected if $y_{45} > y_{cut}$)

CKKW prescription in a nutshell

- Generate samples of N-jet configurations, defined by the $k_\perp$ algorithm, with a resolution parameter k_0
- Since all N-jets have to be resolved w.r.t. the beam, $k_0 = p_T^{min}$. No cut on η can be set, however
- Cluster the partons using the $k_\perp$ algorithm, allowing only for physical branchings in the tree
- Reevaluate α_s at each vertex of the tree, using $k_\perp$ as a scale
- For each line in the tree, associate a Sudakov weight giving the probability that no emission takes place along this line
- Samples of different N-jet multiplicity can now be put together, and evolved through the vetoed shower

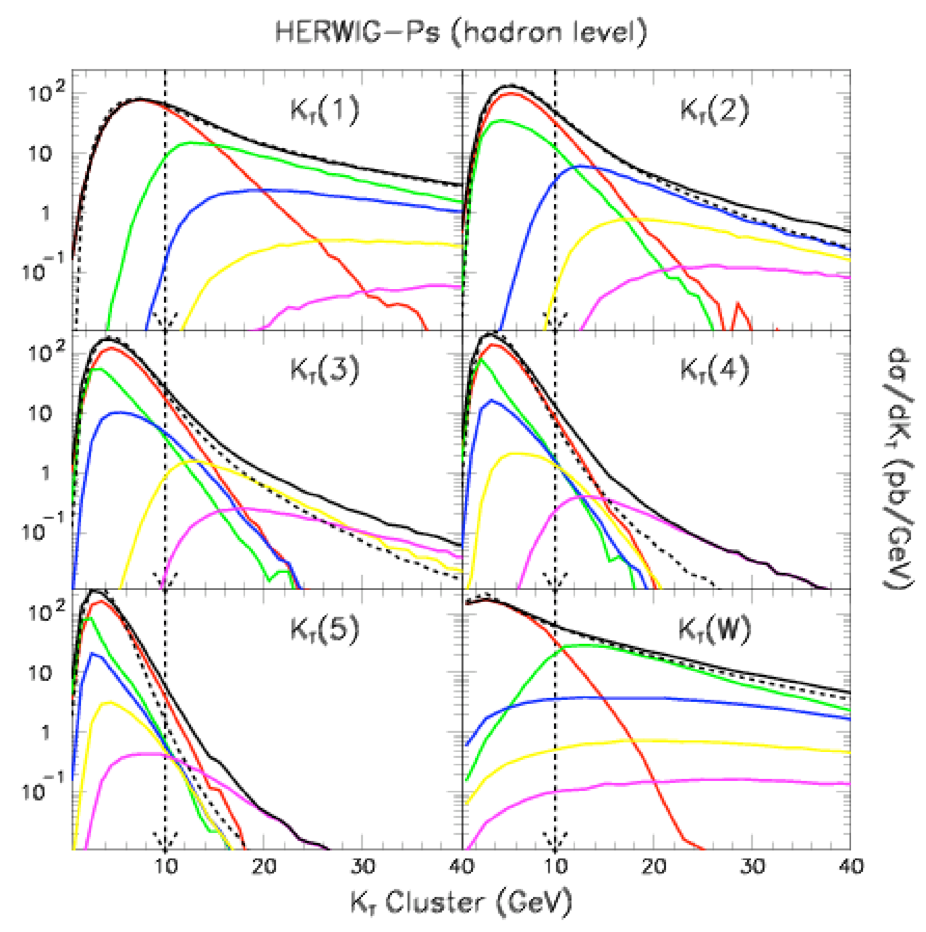


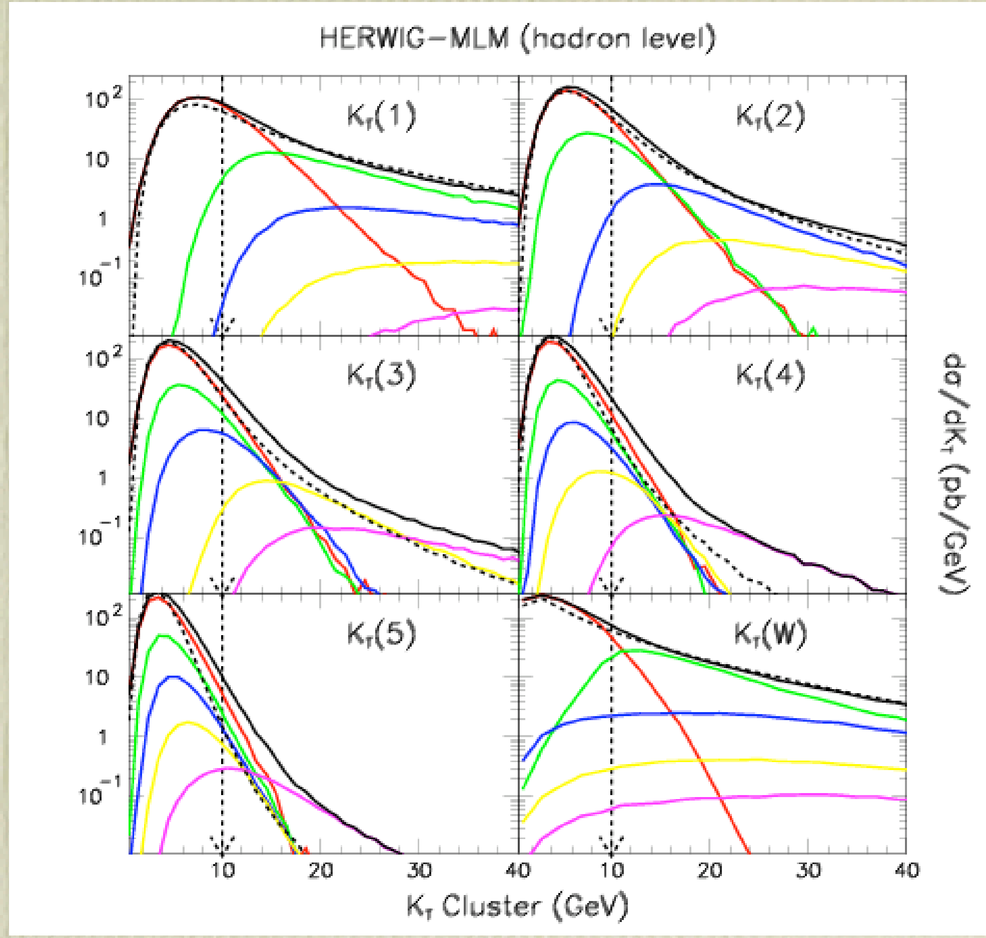
$$\boxed{q_4^2 = q_2^2 + q_3^2 + m_W^2 - min(q_2^2, q_3^2, m_W^2)}$$

$$\boxed{R(q_i,q_j)=\Delta(q_i,q_o)/\Delta(q_j,q_o)}$$

$$w = \prod_1^3 \frac{\alpha_s(q_i)}{\alpha_s(q_0)} \times \prod R(q_i, q_j)$$

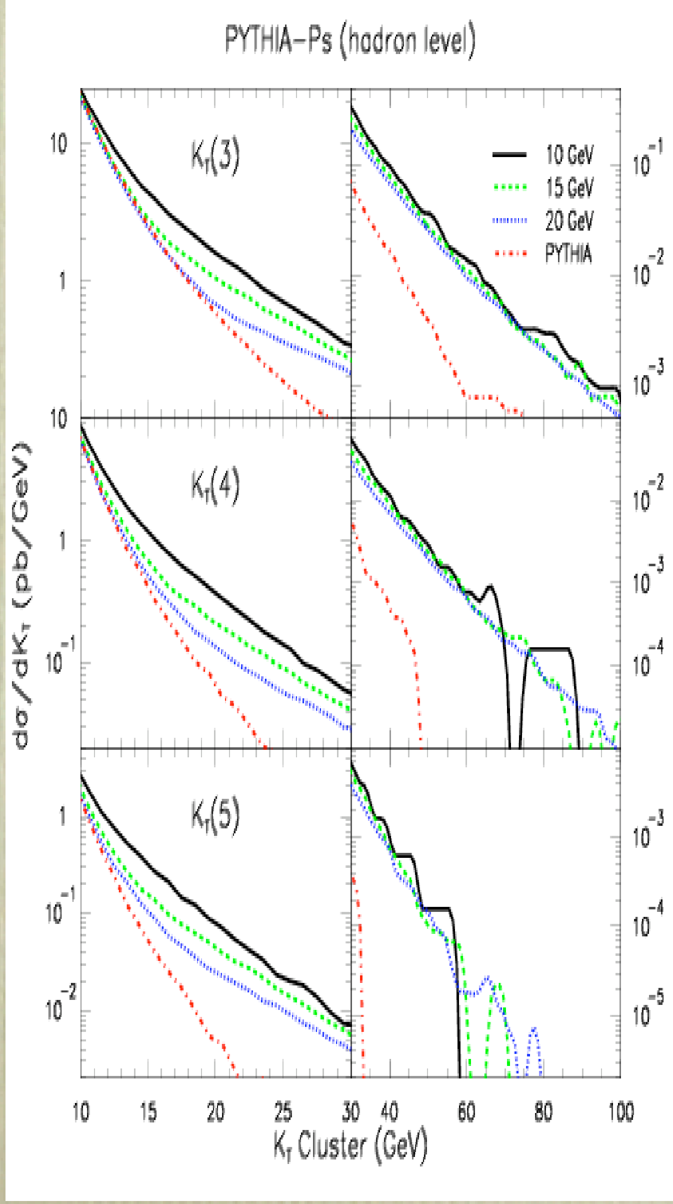
Study of CKKW implementation in hadronic collisions, W+multijets, Mrenna&Richardson hep-ph/0312274

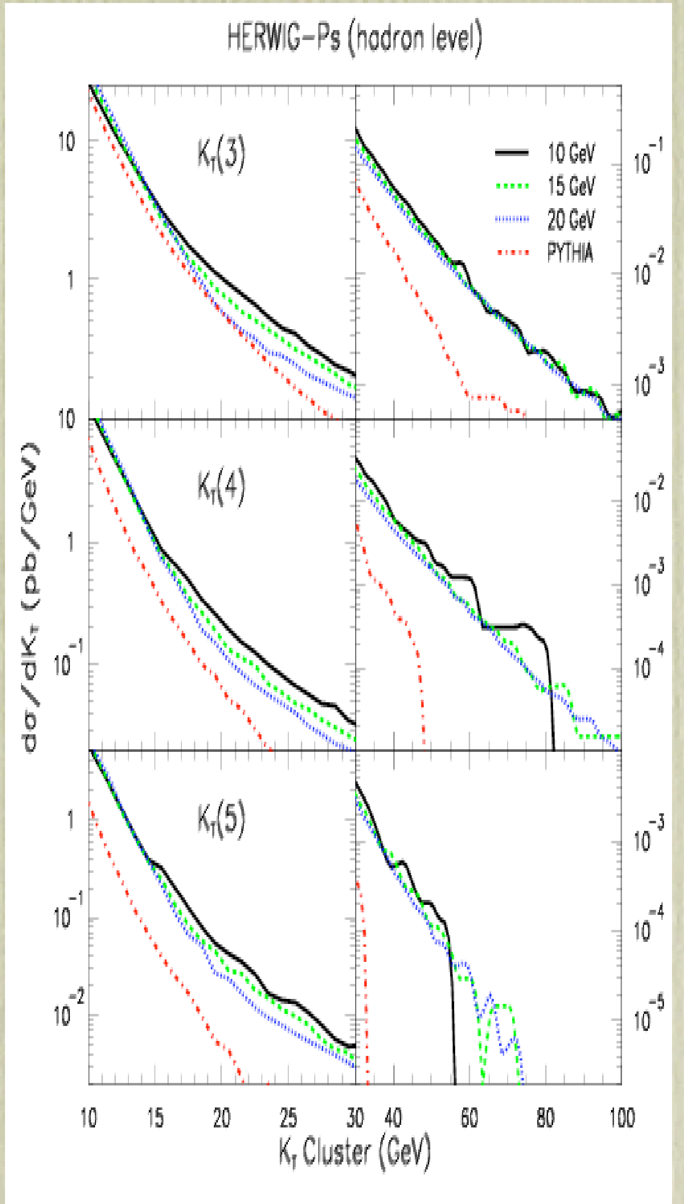


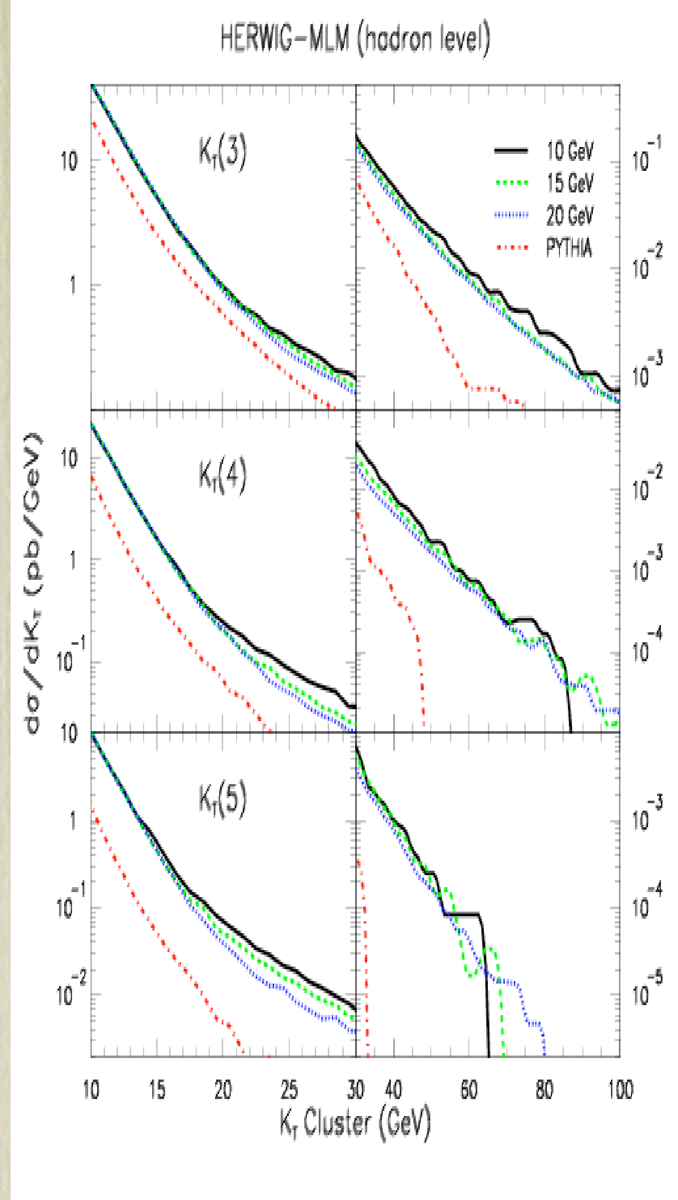


$k_T(n) \approx p_T(\text{n-th jet})$

Cut dependence for three different merging prescriptions: OK, but not perfect!

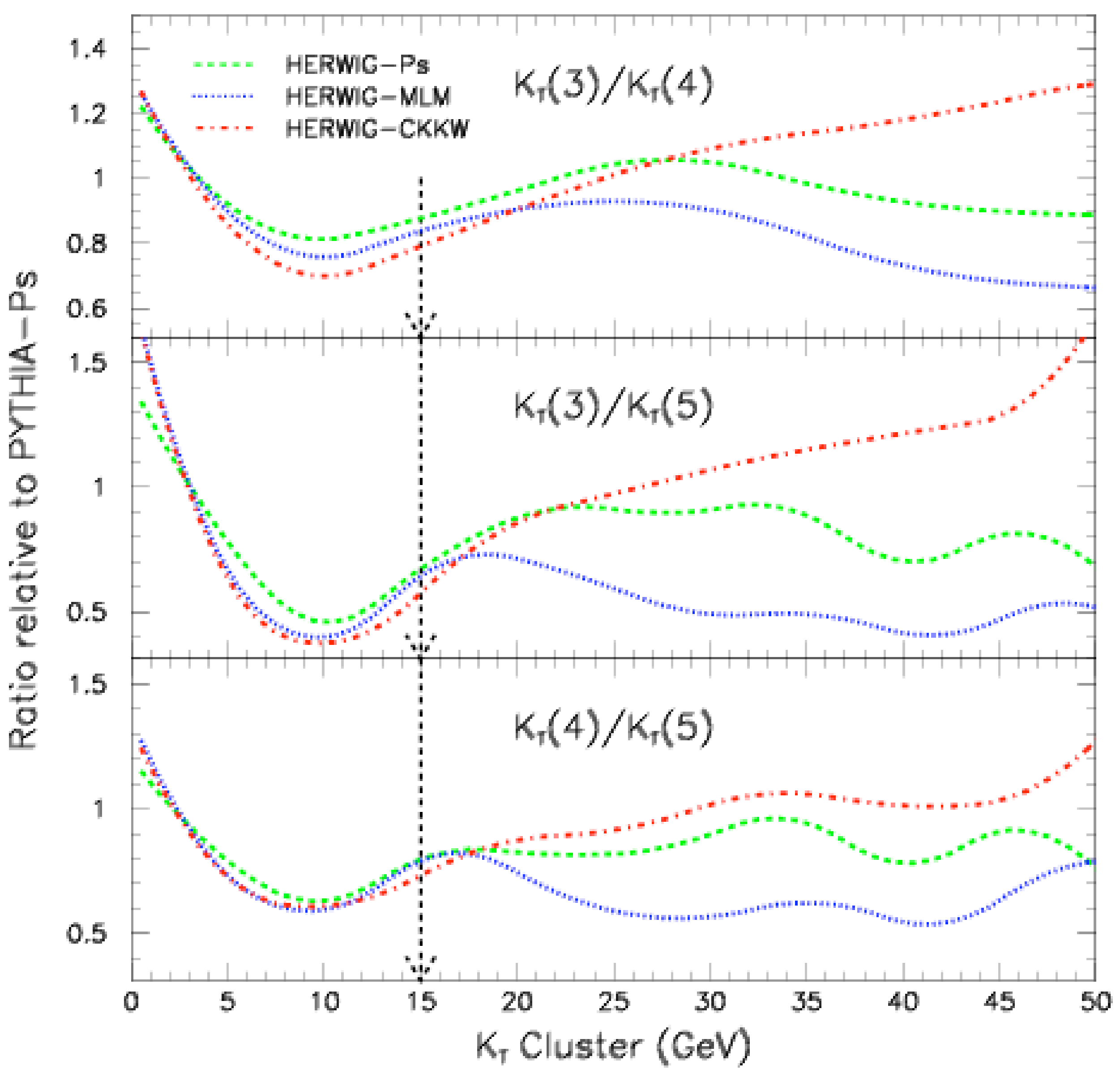






Three different prescriptions, normalized to a reference one

Merging-systematics of the order of ±30% for x-section ratios!



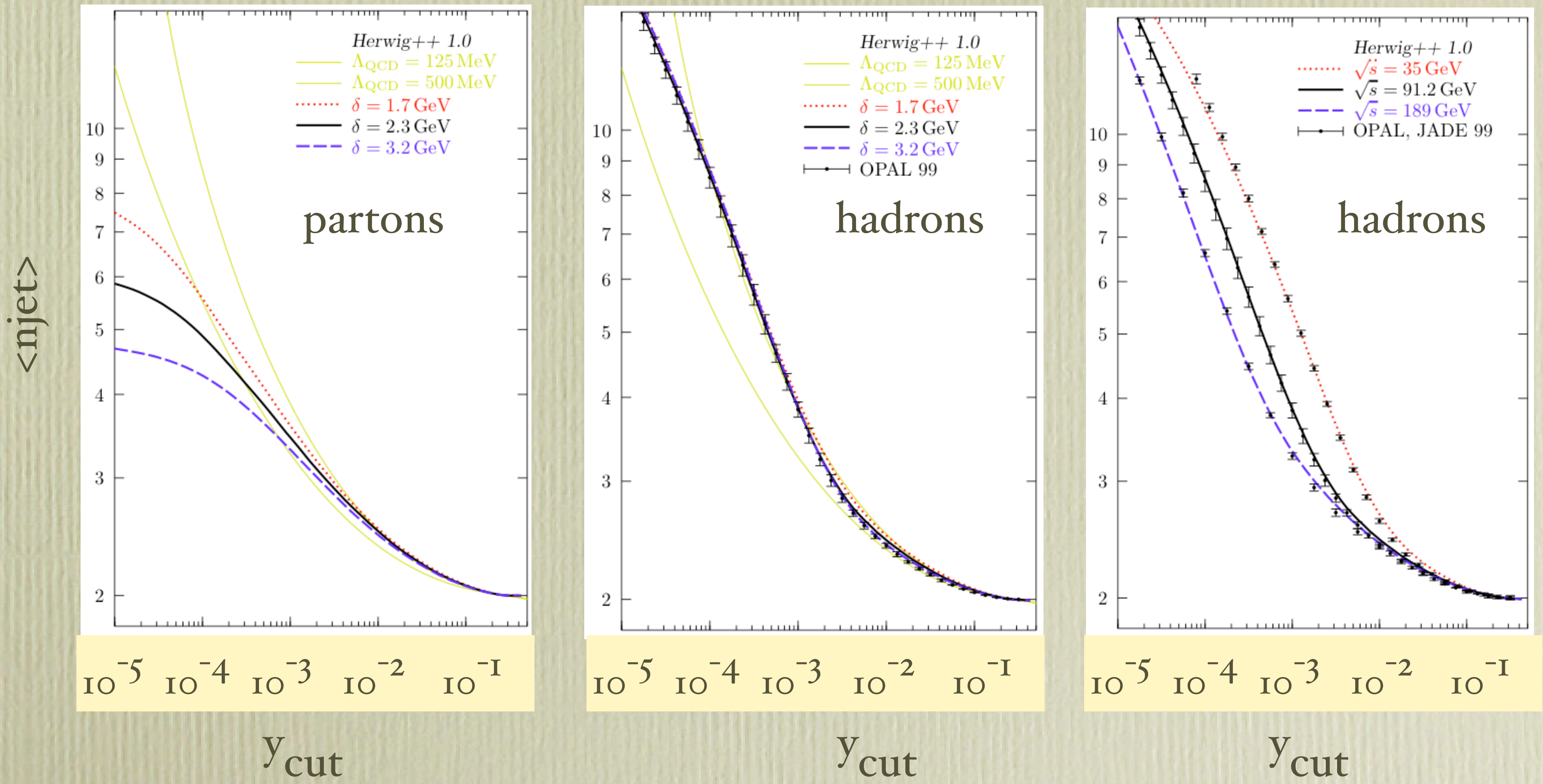
Recent progress in MC-related tools

- New tools to calculate ME's for high multiplicity multijet final states (Alpgen, MadEvent, **2002**)
- New NLO parton-level event generators ⇨ **Campbell talk**
- NLO matrix elements in shower MC's (Dobbs (**2001**), Grace (**2002**), MC@NLO, (**2003**) ⇨ **Frixione talk**
- New techniques for merging of multijet ME's and shower MC's (Catani, Krauss, Kuhn, Webber (**2001**), Lönnblad (**2002**), Mrenna&Richardson (**2003**))
- New shower MC codes (**Sherpa**: Gleisber, Höche,Krauss, Schälicke, Schumann, Winter, **2003**), with new:
    - shower algorithms
    - hadronization schemes
- New incarnation of old MC codes. Pythia/Herwig=>C++ (**2003**) with
    - new features, better QCD, better hadronization

Examples of results from Herwig++ (e+e-)

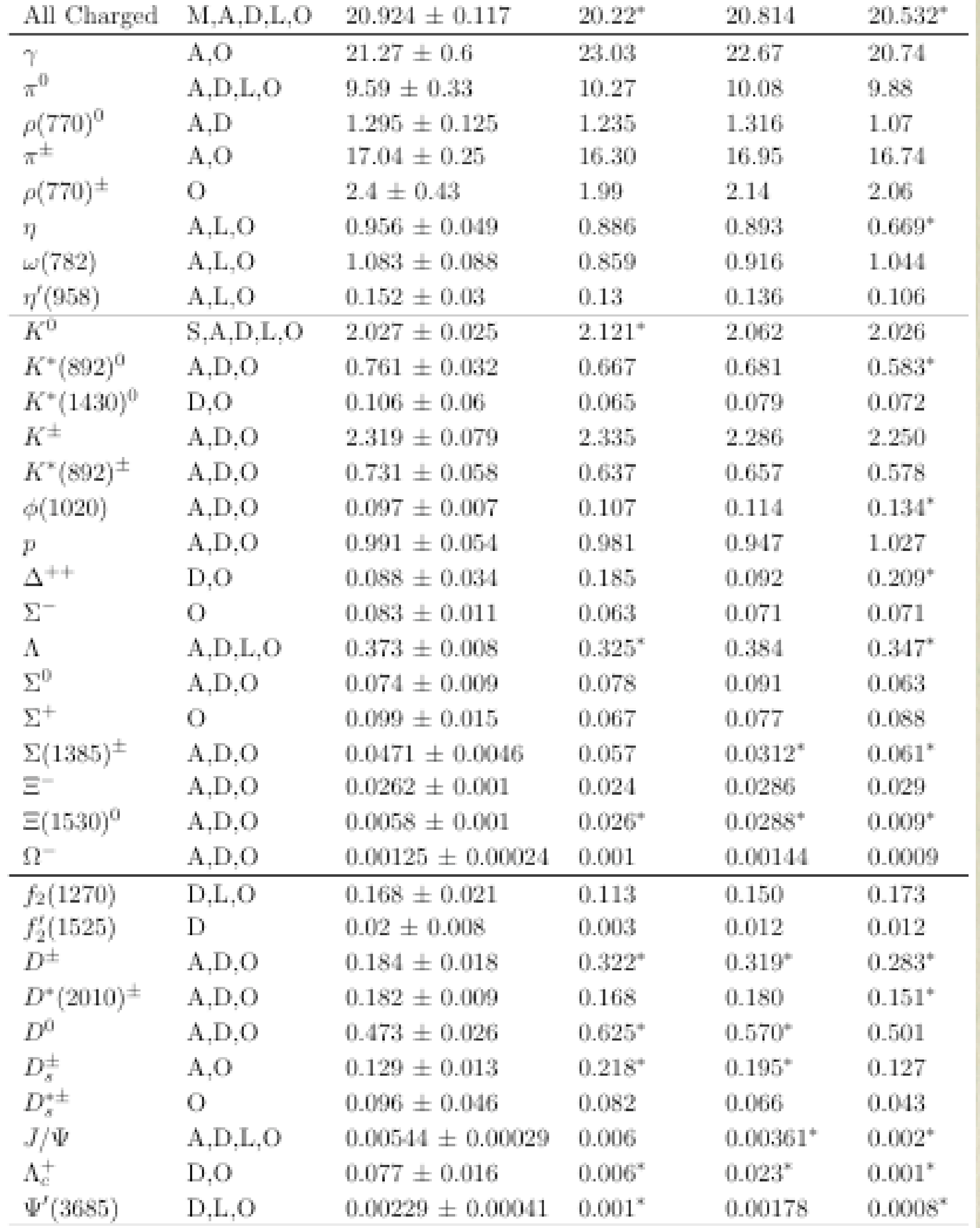
Gieseke, Ribon, Seymour, Stephens, Webber, hep-ph/0311208

Jet multiplicities:



Hadron-level results are rather independent of the IR cutoff (δ) ⇒ consistent merging of the PT↔nPT phases

Particle multiplicities:



Particle	Experiment	Measured	Old Model	Herwig++	Fortran
All Charged	M,A,D,L,O	20.924 ± 0.117	20.22*	20.814	20.532*
γ	A,O	21.27 ± 0.6	23.03	22.67	20.74
π^0	A,D,L,O	9.59 ± 0.33	10.27	10.08	9.88
$\rho(770)^0$	A,D	1.295 ± 0.125	1.235	1.316	1.07
$\pi^\pm$	A,O	17.04 ± 0.25	16.30	16.95	16.74
$\rho(770)^\pm$	O	2.4 ± 0.43	1.99	2.14	2.06
η	A,L,O	0.956 ± 0.049	0.886	0.893	0.669*
$\omega(782)$	A,L,O	1.083 ± 0.088	0.859	0.916	1.044
$\eta'(958)$	A,L,O	0.152 ± 0.03	0.13	0.136	0.106
K^0	S,A,D,L,O	2.027 ± 0.025	2.121*	2.062	2.026
$K^*(892)^0$	A,D,O	0.761 ± 0.032	0.667	0.681	0.583*
$K^*(1430)^0$	D,O	0.106 ± 0.06	0.065	0.079	0.072
$K^\pm$	A,D,O	2.319 ± 0.079	2.335	2.286	2.250
$K^*(892)^\pm$	A,D,O	0.731 ± 0.058	0.637	0.657	0.578
$\phi(1020)$	A,D,O	0.097 ± 0.007	0.107	0.114	0.134*
p	A,D,O	0.991 ± 0.054	0.981	0.947	1.027
Δ^{++}	D,O	0.088 ± 0.034	0.185	0.092	0.209*
Σ^-	O	0.083 ± 0.011	0.063	0.071	0.071
Λ	A,D,L,O	0.373 ± 0.008	0.325*	0.384	0.347*
Σ^0	A,D,O	0.074 ± 0.009	0.078	0.091	0.063
Σ^+	O	0.099 ± 0.015	0.067	0.077	0.088
$\Sigma(1385)^\pm$	A,D,O	0.0471 ± 0.0046	0.057	0.0312*	0.061*
Ξ^-	A,D,O	0.0262 ± 0.001	0.024	0.0286	0.029
$\Xi(1530)^0$	A,D,O	0.0058 ± 0.001	0.026*	0.0288*	0.009*
Ω^-	A,D,O	0.00125 ± 0.00024	0.001	0.00144	0.0009
$f_2(1270)$	D,L,O	0.168 ± 0.021	0.113	0.150	0.173
$f_2'(1525)$	D	0.02 ± 0.008	0.003	0.012	0.012
$D^\pm$	A,D,O	0.184 ± 0.018	0.322*	0.319*	0.283*
$D^*(2010)^\pm$	A,D,O	0.182 ± 0.009	0.168	0.180	0.151*
D^0	A,D,O	0.473 ± 0.026	0.625*	0.570*	0.501
$D_s^\pm$	A,O	0.129 ± 0.013	0.218*	0.195*	0.127
$D_s^{*\pm}$	O	0.096 ± 0.046	0.082	0.066	0.043
J/Ψ	A,D,L,O	0.00544 ± 0.00029	0.006	0.00361*	0.002*
Λ_c^+	D,O	0.077 ± 0.016	0.006*	0.023*	0.001*
$\Psi'(3685)$	D,L,O	0.00229 ± 0.00041	0.001*	0.00178	0.0008*

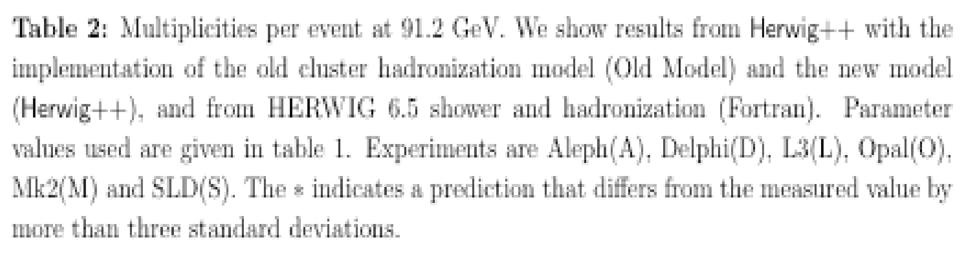
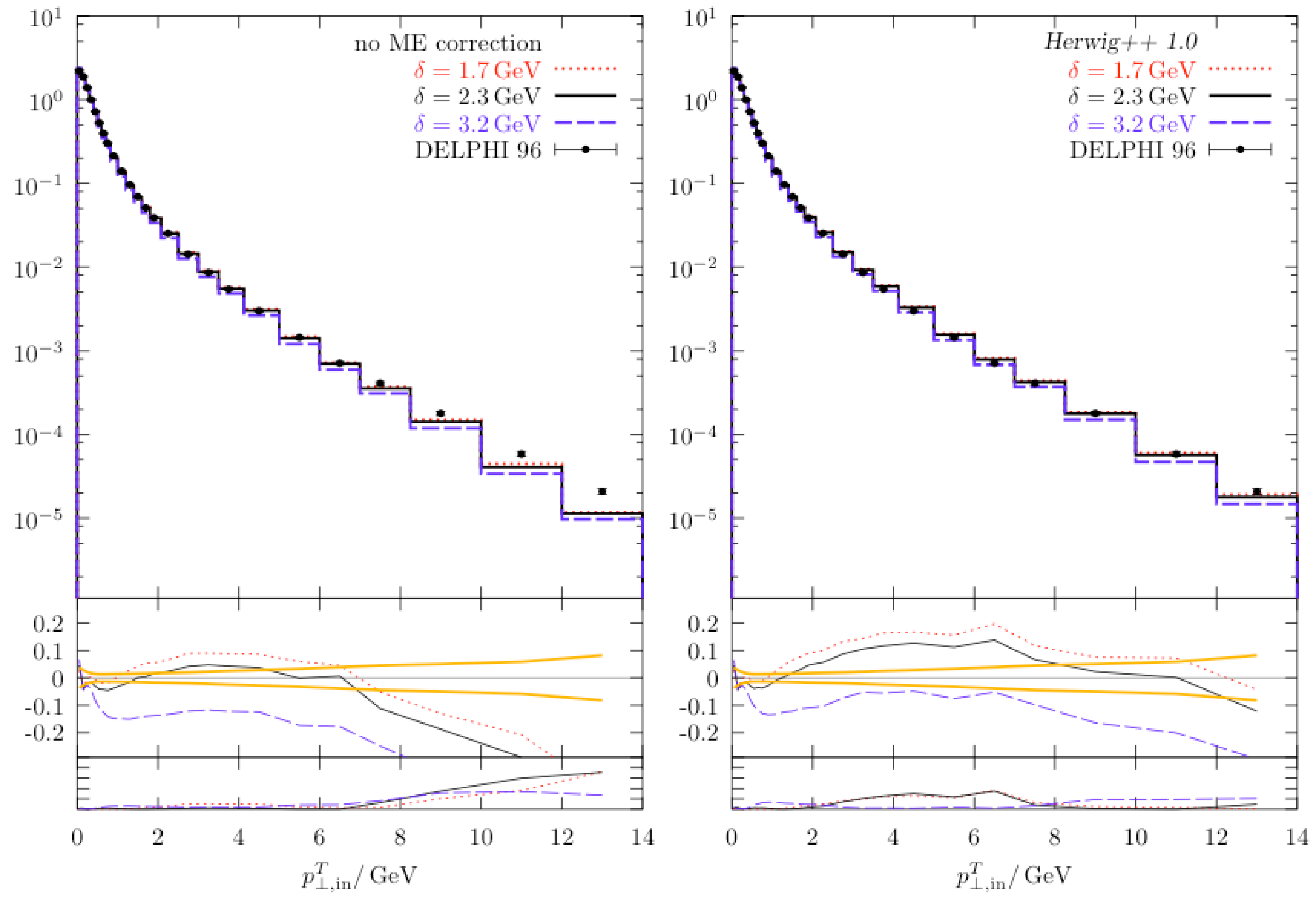


Table 2: Multiplicities per event at 91.2 GeV. We show results from Herwig++ with the implementation of the old cluster hadronization model (Old Model) and the new model (Herwig++), and from HERWIG 6.5 shower and hadronization (Fortran). Parameter values used are given in table 1. Experiments are Aleph(A), Delphi(D), L3(L), Opal(O), Mk2(M) and SLD(S). The $*$ indicates a prediction that differs from the measured value by more than three standard deviations.

Transverse momenta w.r.t. thrust axis:



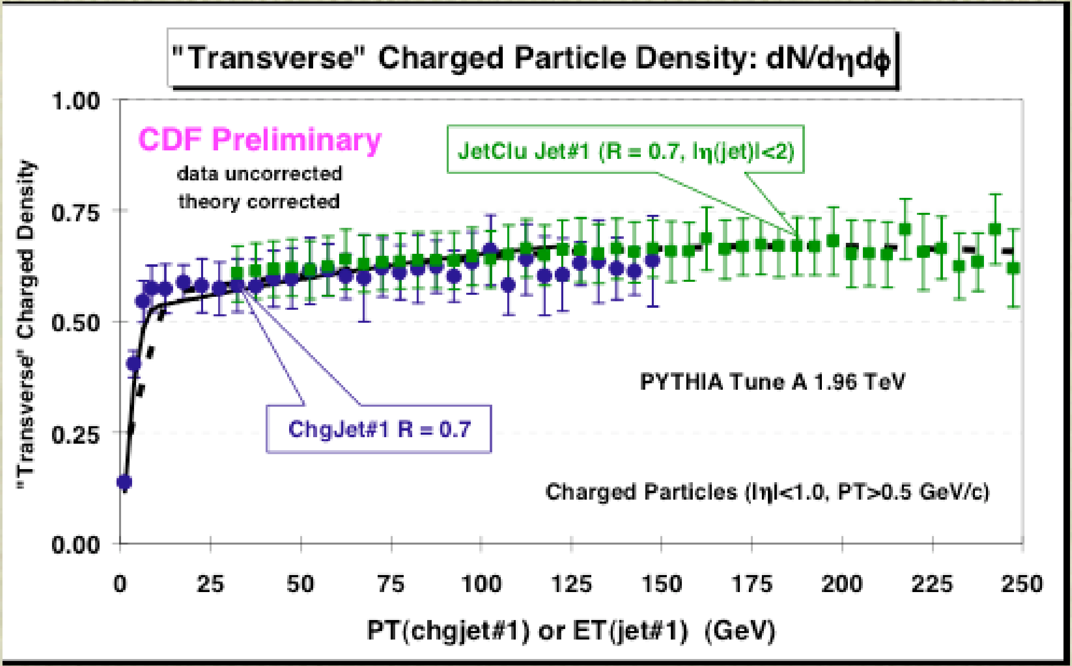
Improvement in the shower algorithm reduces the impact of Matrix Element corrections:

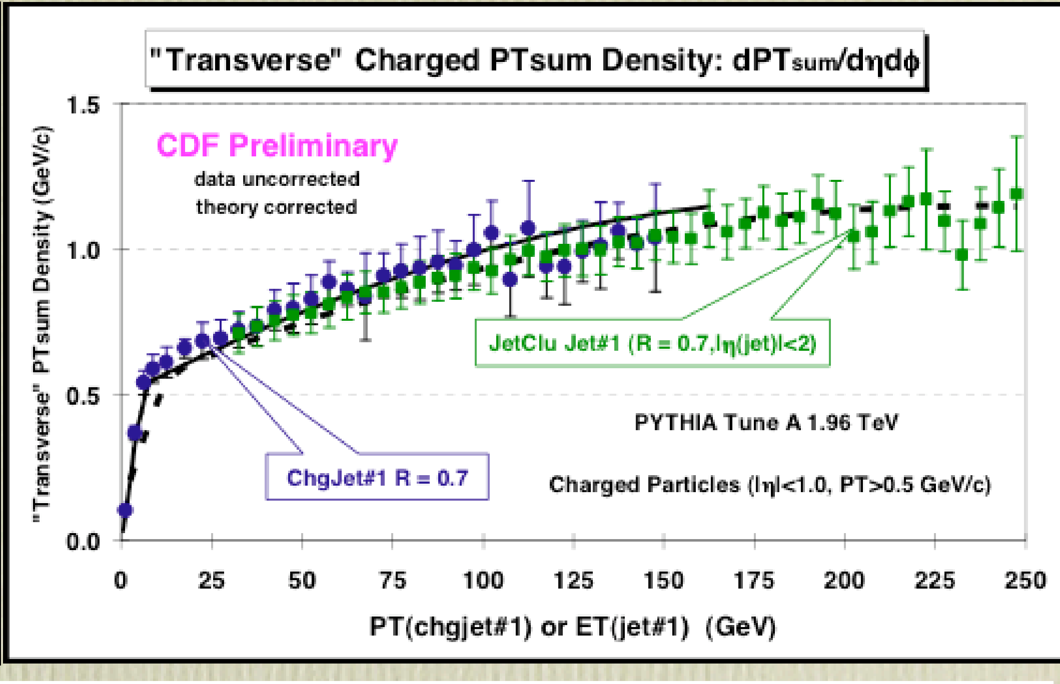
=> expect improvement in the description of higher jet multiplicities

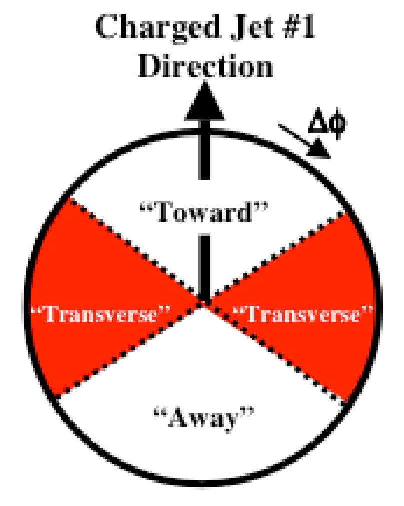
Recent progress in MC-related tools

- New tools to calculate ME's for high multiplicity multijet final states (Alpgen, MadEvent, **2002**)
- New NLO parton-level event generators ➪ **Campbell talk**
- NLO matrix elements in shower MC's (Dobbs (**2001**), Grace (**2002**), MC@NLO, (**2003**) ➪ **Frixione talk**
- New techniques for merging of multijet ME's and shower MC's (Catani, Krauss, Kuhn, Webber (**2001**), Lönnblad (**2002**), Mrenna&Richardson (**2003**))
- New shower MC codes (**Sherpa**: Gleisber, Höche,Krauss, Schälicke, Schumann, Winter, **2003**), with new:
  - shower algorithms
  - hadronization schemes
- New incarnation of old MC codes. Pythia/Herwig=>C++ (**2003**) with
  - new features, better QCD, better hadronization
- Data from Tevatron to study and model the underlying event (R.Field-CDF, **2002**). New models (Skands & Sjöstrand, **2003**)

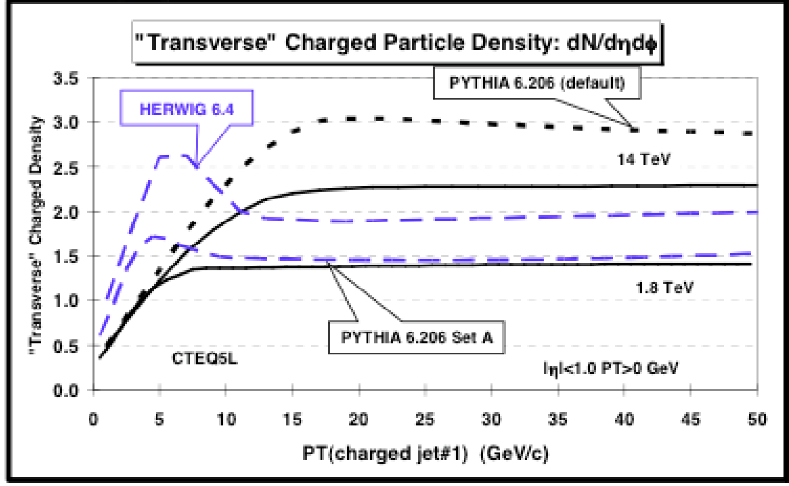
MC UE tuning with CDF data (R.Field, CDF)

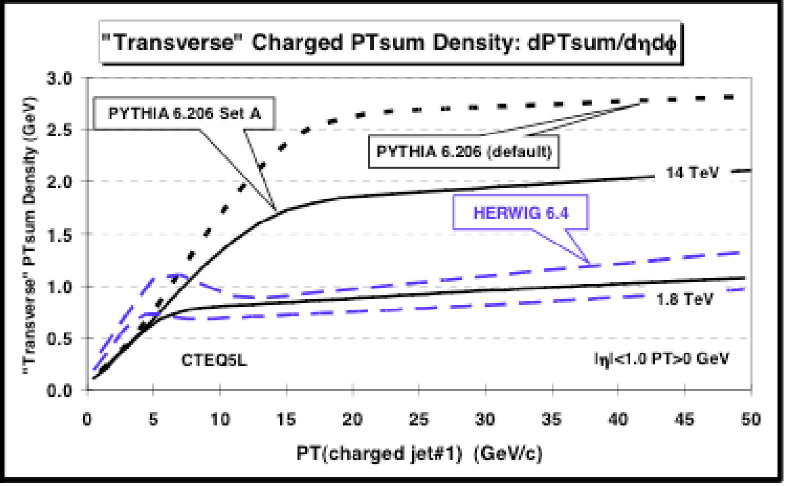






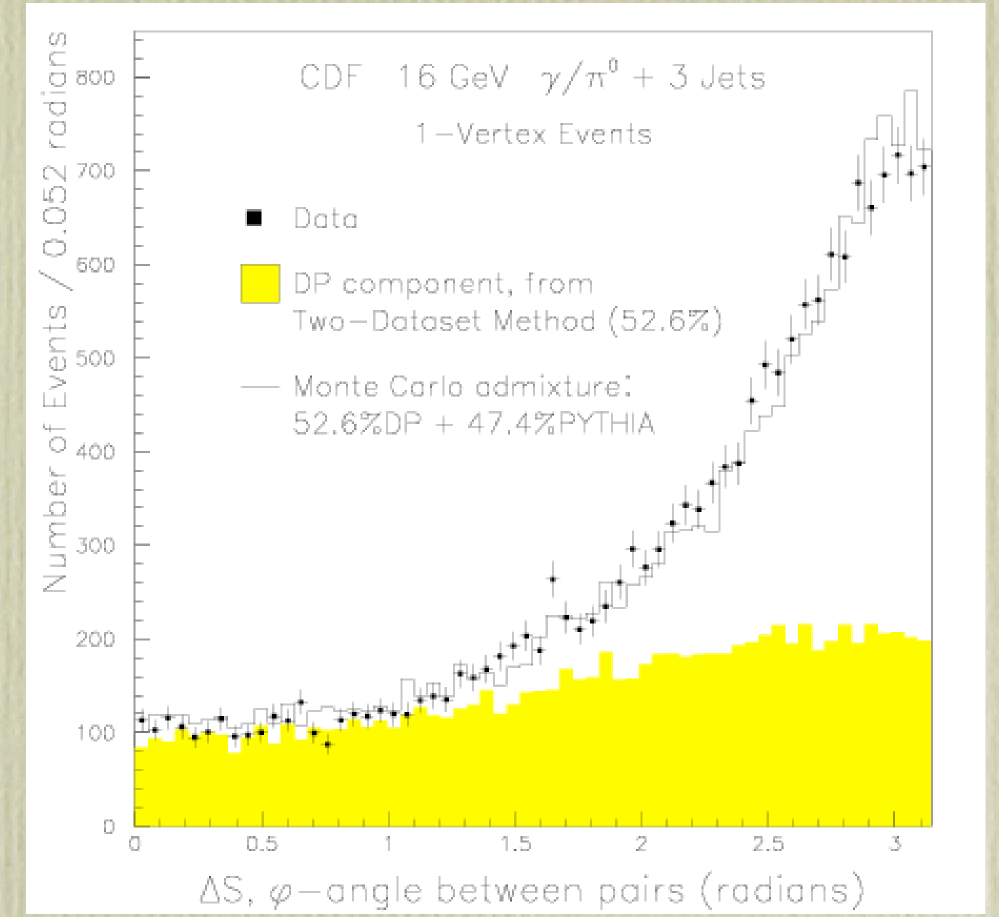
Tuned PYTHIA (Set A)
LHC Predictions





Direct evidence for multiparton collisions

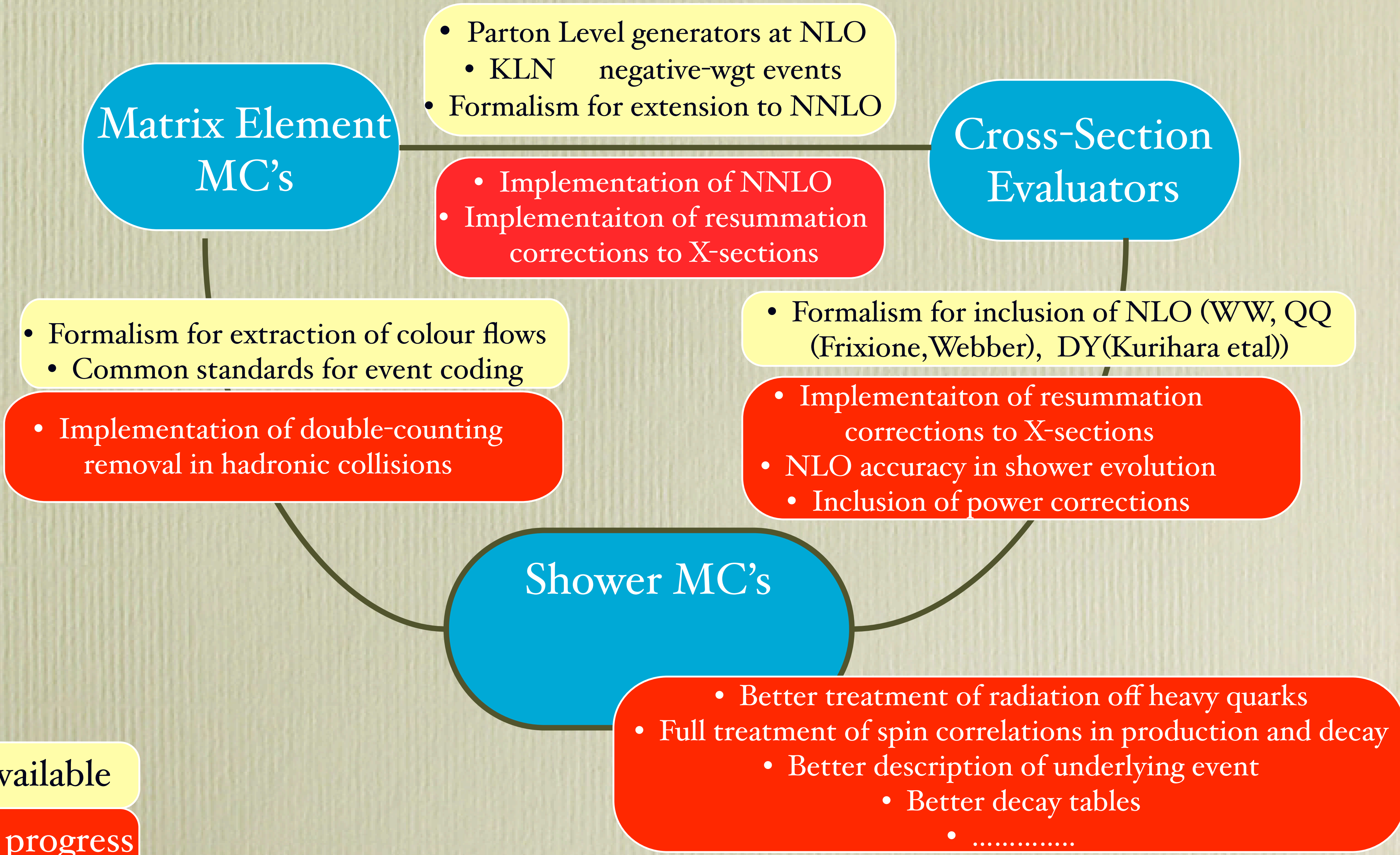
CDF, run I,
γ+3jet events



Since $\sigma_{tot}=\sigma_{jet}$(Et>few GeV), each individual collision at the LHC will lead to multiple hard scatterings

Need concrete models to describe correlations in multiparton density distributions. Recent developments include **momentum, flavour and colour correlations** among partons contributing to the multiple interactions (Skands&Sjöstrand, hep-ph/0310315)

M(*ontecarlo*) o(*f*) E(*verything*)



- Parton Level generators at NLO
- KLN negative-wgt events
- Formalism for extension to NNLO

Matrix Element MC's

Cross-Section Evaluators

- Implementation of NNLO
- Implementaiton of resummation corrections to X-sections

- Formalism for extraction of colour flows
- Common standards for event coding

- Implementation of double-counting removal in hadronic collisions

- Formalism for inclusion of NLO (WW, QQ (Frixione,Webber), DY(Kurihara etal))

- Implementaiton of resummation corrections to X-sections
- NLO accuracy in shower evolution
- Inclusion of power corrections

Shower MC's

- Better treatment of radiation off heavy quarks
- Full treatment of spin correlations in production and decay
- Better description of underlying event
- Better decay tables
- ....

available

in progress

Final remarks

- A lot of progress has taken place in the recent years, but 30 yrs after QCD, still a lot of work to be done to achieve a satisfactory description of all high-Q^2 processes accessible at LHC
- most of the key conceptual difficulties have been recently, or are being, solved, and their implementation into concrete MC schemes should be achievable in the next 5 years
- with the level of accuracy reached today with NLO and NNLO **parton level** calculations, attention needs to be shifted to the impact of violations from the naive factorization assumptions. Shower MC's, especially MC@NLO, provide an excellent tool to explore the effects of hadronization and "explicit resummation"
- forthcoming data from Tevatron and HERA will help improving our tools, but the final test will need real LHC data (FYI, a year-long Workshop sponsored by CERN-DESY will start on March 26-27 dedicated to the interplay of HERA/LHC)
- **there is plenty of room for creative and rewarding work for young phenomenologists!**