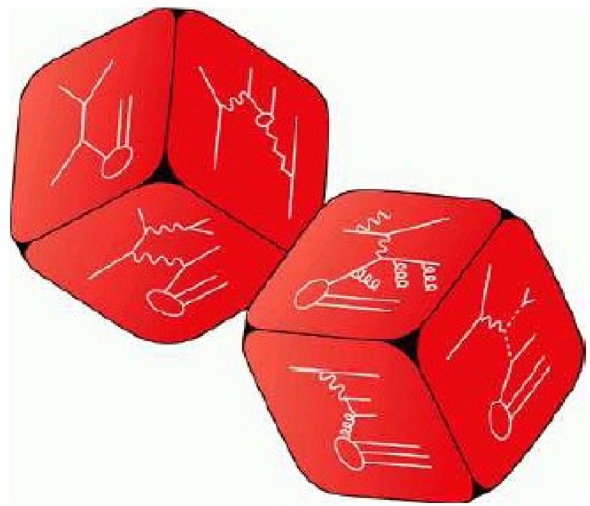
Monte Carlo Generators for Higgs Physics

Guillelmo Gomez-Ceballos
(MIT)
with
Fabio Maltoni, Fabian Stoekli,
Roberto Chierici



Outline

- Monte Carlo generators
- LHC case
- Event generators
- Higgs production processes
- Higgs decays & branching ratios
- Final remarks
- Conclusions

Monte Carlo event generators and HEP

Software tools for simulating real data.
Mandatory tools for high energy physicists. One must be able to:

1. Plan experiments and detectors
    - how should I design my detector on a given collider to be able to see, at 95% CL, a signal of new physics in the channel X with expected cross section Y, with a total integrated luminosity L? What are the backgrounds one should worry about?
2. Making a measurement
    - motivate, validate and tune analyses strategies and cuts
    - extract the expected backgrounds
    - model the expected backgrounds (and subtract them to data, eventually)
    - know the expected signal efficiencies
3. Making a discovery
    - is what I see predicted by simulation?
    - am I seeing a physics effect, or a detector effect?
    - is it something expected? Should I claim something?
    - is it a bug? A bad Monte Carlo tuning instead?…

In all this, one must take care of using those tools which are the best suited (modern, accurate and flexible) for the particular physics (signal and backgrounds) under study

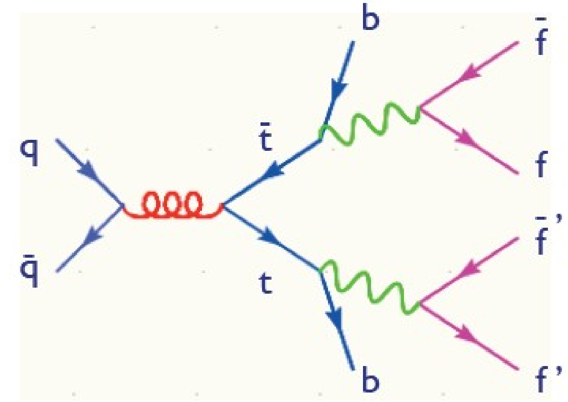
The path towards discoveries

1. Rediscover the known SM

2. Identify excess(es) over SM

3. Identify the nature of BSM:
   from coarse information to measurements of mass spectrum, quantum numbers, couplings

A lesson from the top at the Tevatron



Before the Tevatron the only unknown was the top mass

- though production mechanism was known

1. First mass reconstruction gave a signal consistent with a top production with a mass around 175 GeV/c2

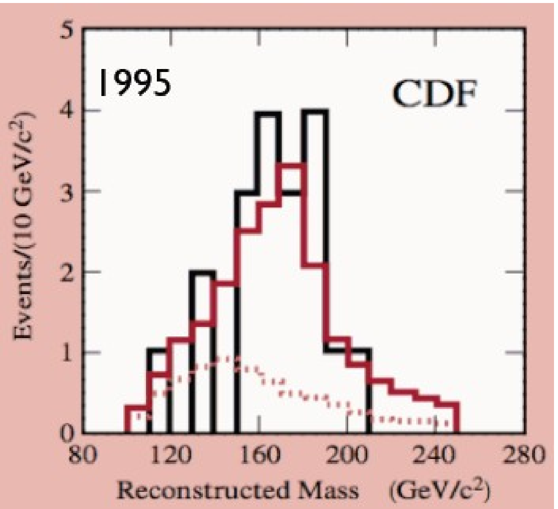
- total QCD cross-section and differential mass distribution
- handful of events!

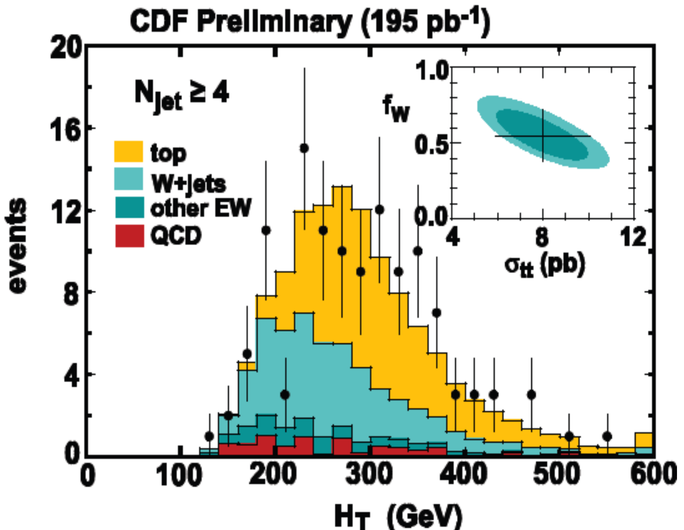
2. Presence of "something" confirmed also in other distributions with more statistics

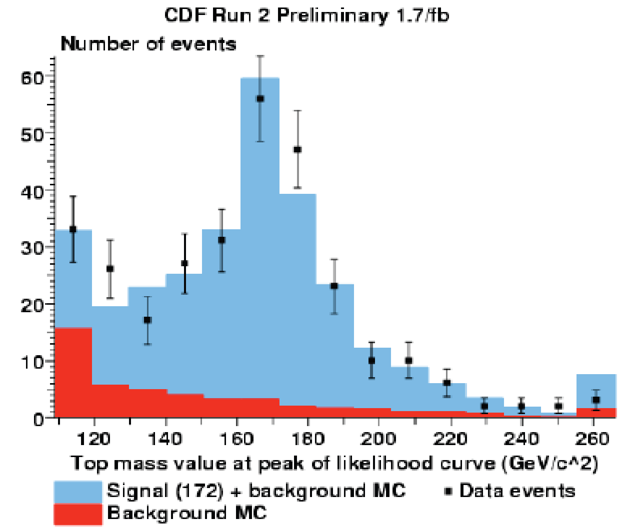
- signal compatible with a top pair production, not explicable otherwise
- consistent picture came out with more data and the possibility of more quantities like its BRs, charge, polarization, and many differential distributions

3. Top studies were established, and Monte Carlos played a role in all steps

- the more sophisticated the study, the more requirement to the Monte Carlo predictions







The principle of event generation

General problem in HEP: given a collision process $ab \to X$, simulate X.
More specifically, make a program able to throw out "events" distributed according to $d\sigma_{ab\to X}/d\Phi(n)$, with Φ phase space in n dimensions.

- an "event" is to be considered as a set of four-momenta representing the final particles of X

$1/\sigma_{ab\to X} * d\sigma_{ab\to X}/d\Phi(n)$ can be seen as an n-dimensional probability density function for X, according to which we want to make our "event generator"

Simple example for illustrating how to proceed: 1D case (Φ==x):

1. throw "events" with "weights"
   - weighted event generation
2. throw, make it an "event" only with frequency and give to it a unitary weight
   - event unweighting
   - events happen with same frequency as in data
   - efficiency is non 1

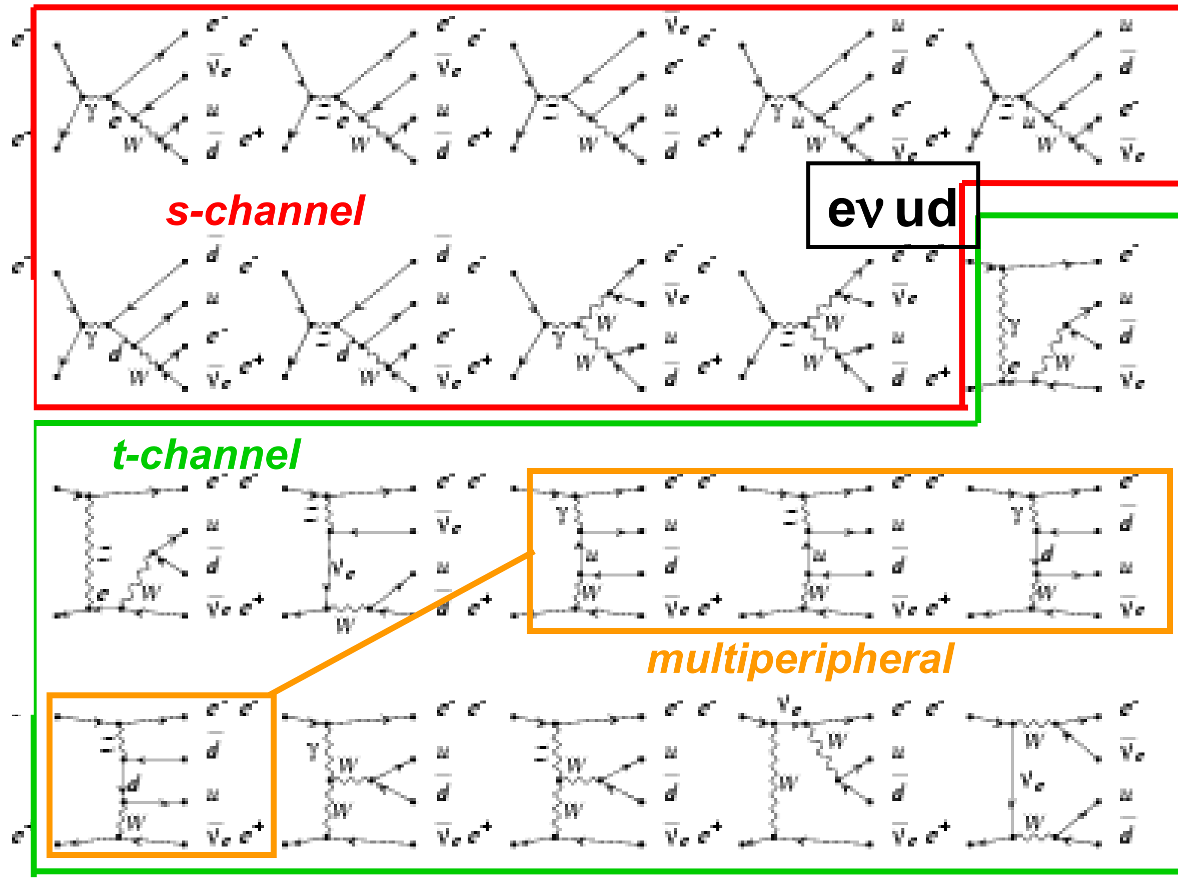
Monte Carlos

The first step is therefore the determination of $\sigma_{ab \rightarrow X}$

$$\sigma = \frac{1}{2s} \int |\mathcal{M}|^2 d\Phi(n)$$

Writing the Matrix Elements M from all contributing Feynman diagrams can be very complex

- It depends on the number of particles in the final state
- It depends on the order of the perturbative calculation



Example at LEP: $e^+e^- \rightarrow e\nu ud$.

- W^+W^- production mixes with single W production and non resonant one
- They all interfere in the calculation of $|M|^2 = |WW+W+\text{non-res}|^2$
- Interferences may be important ! Neglecting some contributions may simply lead you to wrong results

An analogous example at the LHC will turn out to be much more complex

- Need to account for all possible initial states (the proton is a composite object)
- QCD in production complicates a lot with respect to lepton colliders

From partons to hadrons

Perturbative calculations can be performed up to a limited number of particles in the final state

- Factorial growth of Feynman diagrams
- Can be brought to exponential with tricks

n	full Amp	partial Amp	BG
4	4	3	3
5	25	10	10
6	220	36	35
7	2485	133	70
8	34300	501	126
9	559405	1991	210
10	10525900	7335	330
11	224449225	28199	495
12	5348843500	108280	715
	$(2n)!$	3.8^n (using tricks)	n^4 (using recursive relations)

Numerical integrations can become awkward very easily as well

The event generation described so far is essential to bring you predictions of parton level configurations as if they were the final state, but they do not have much to do with what can be seen at the detector level

BUT: in a hadronic machine an event with few hard partons can give origin to events with several hundreds of particles in the final state.
We need to describe/model the mechanisms going from partons to long-lived hadrons to build a usable event generator.

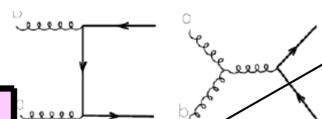
The LHC case

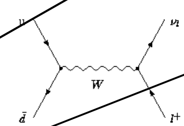
Cross-sections at the LHC

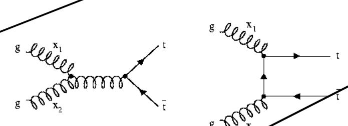
LHC physics = QCD + ϵ

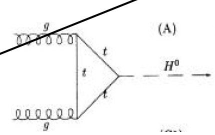
- **Huge statistics @14 TeV:**
  - total of Zs collected at LEP will be produced in ~weeks
  - for Ws it will be ~hours
  - need ~minutes to produce as many tops as produced at the Tevatron

- **There may be more than 10 orders of magnitude between signal (i.e. H) and backgrounds (i.e. QCD)**
  - the "corners" of the phase space is typically what we are interested in

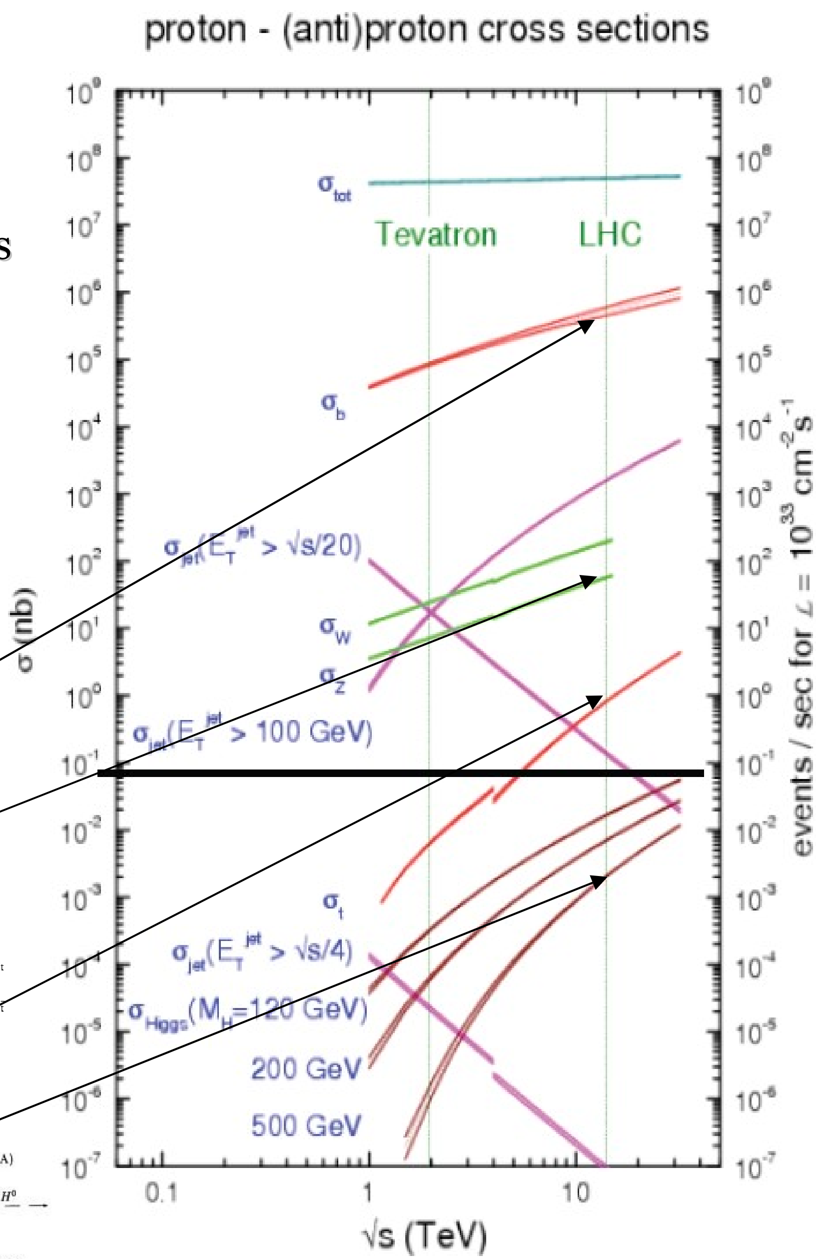








process	Events@ 14TeV/s	Events@ 14TeV/y
bb	10^6	10^{12}
$Z \to ee$	~3	10^7
$W \to e\nu$	~30	10^8
$WW \to e\nu X$	10^{-2}	6×10^3
tt	~2	10^7
H(700 GeV)	2×10^{-3}	10^4



From partons to hadrons

How to transform a parton level event (a few particles) in a full event (order of several hundreds particles) at the LHC?
The full process is sub-divided into pieces, each of which can be treated independently and modelled accordingly:

- This is an approximation, we need to model what we are unable to calculate
- Models, as such, have parameters that we need to adjust to data (tunings)

What are the parts needed, for a p-p collision?

0. Knowledge of the contents of the proton (PDFs)
1. Hard scattering (Perturbative calculations)
2. Radiation (QCD, QED) off the initial or final partons (Parton Shower, ISR and FSR)
3. Description of the underlying event (UE)
4. Fragmentation into hadrons (Hadronisation)
5. Decay of unstable particles

How to get the whole picture right, in terms of average values and fluctuations?

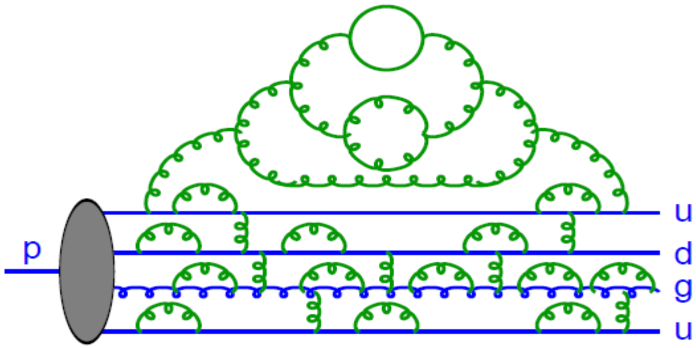
- Make random choices on an event by event basis, as in Nature

How will the final rate/cross-section change?

- $\sigma_{FINAL}=\sigma_{HARD}\times P_{HARD\rightarrow FINAL}$; $P_{HARD\rightarrow FINAL}=P_{ISR}\times P_{FSR}\times P_{UE}\times P_{hadronisation}\times P_{decay}$

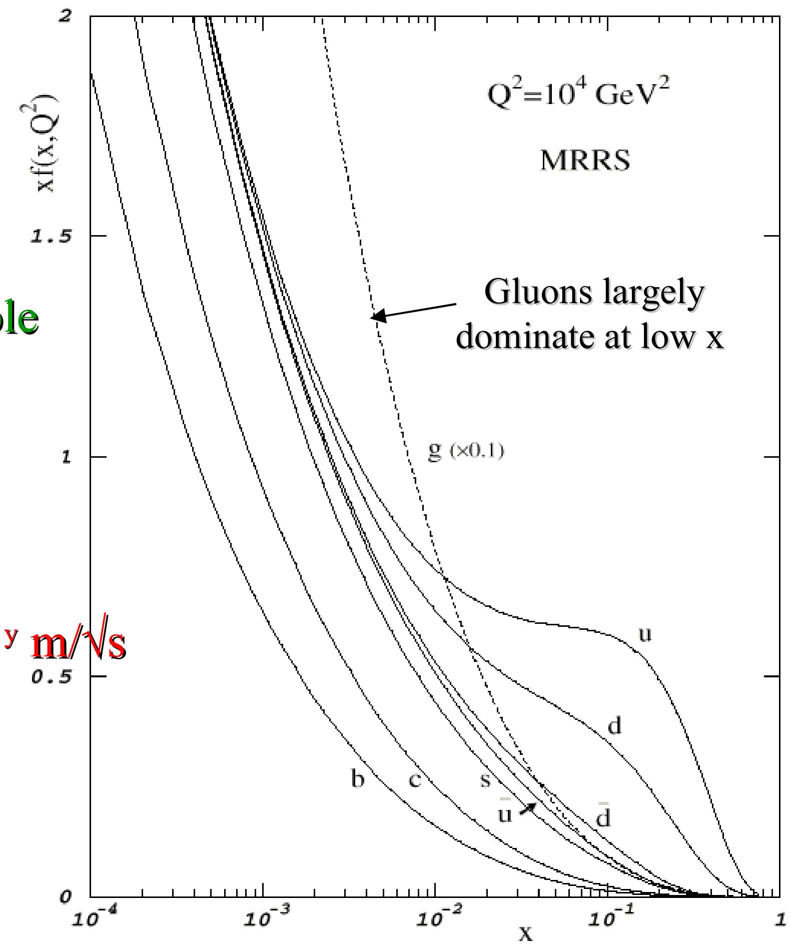
PDFs

Hadrons are composite, with time-dependent structure:



$f_i(x, Q^2) =$ number density of partons i
at momentum fraction x and probing scale Q^2.

$$\sigma_X = \sum_{a,b} \int_0^1 dx_1 dx_2\, f_a(x_1, \mu_F^2) f_b(x_2, \mu_F^2) \times \hat{\sigma}_{ab\to X}\left(x_1, x_2, \alpha_S(\mu_R^2), \frac{Q^2}{\mu_F^2}, \frac{Q^2}{\mu_R^2}\right)$$

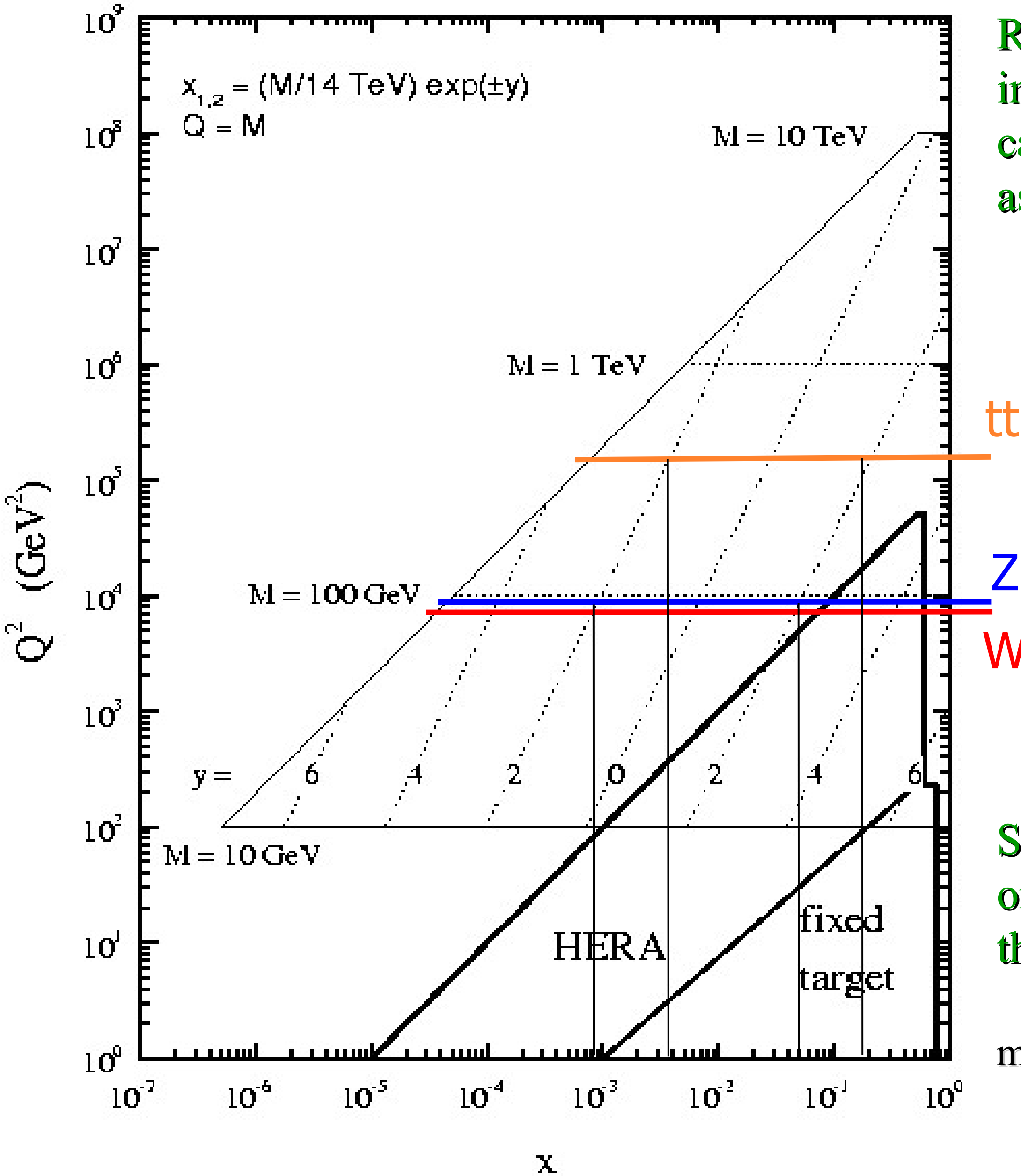


Any cross-section determination must sum over all possible initial states of the proton
The proton density functions are fitted by using heterogeneous collider data

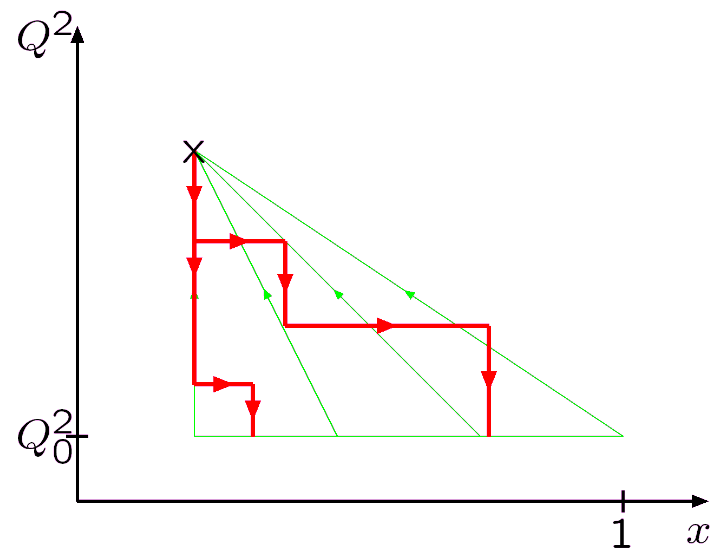
How? For an s-channel process (W, Z, W/ZW/Z, tt)
$m^2=sx_1x_2$ and $y=1/2\ln(E+p_z/E-p_z)=1/2\ln(x_1/x_2) \Rightarrow x_{1/2}= e^{\pm y}\, m/\sqrt{s}$

$$\frac{dN_X}{dy} = \frac{d\sigma_{qq,gg\to X}}{dy} \times L \times pdf_{qq,gg}(x_1, x_2; Q^2)$$

PDFs



Regions where no data is available can be inferred by extrapolation using the so-called DGLAP evolution: PDFs at (x, Q^2) as a function of PDFs at (<x,>Q^2)



Some regions accessible at the LHC will only be fitted by using the LHC data themselves

- especially the gluon PDFs need more data

Parton Showers

Accelerated charged particles radiate in QED, and accelerated coloured particles radiate in QCD (in QCD also gluons radiate)

- Emitted gluons radiate in their turn, and so on → "parton shower"

The main problem of treating QCD/QED radiation is that the real emission rates $q \to qg / e \to e\gamma$ formally diverge when collinear (opening angle $\theta_{qg} \to 0$) or soft ($E_g \to 0$). For QCD also $g \to gg$ is similarly divergent.

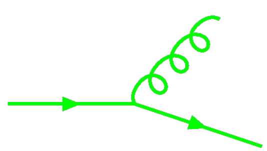
- The divergence is formal, and appears because you touch the very essence of the definition of what a parton is.

  two collinear partons = 1 parton

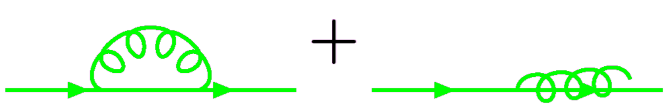
  two partons one of which with zero energy = 1 parton

Need to bound your emission phase space to consider only resolvable emissions. Indeed:

- One can not treat real emissions alone without the introduction of proper phase space cut-offs
- Unresolvable real emission always need to be combined with virtual corrections, bringing in divergencies that exactly cancel those from the real part, making the result finite



Resolvable emission **Finite**

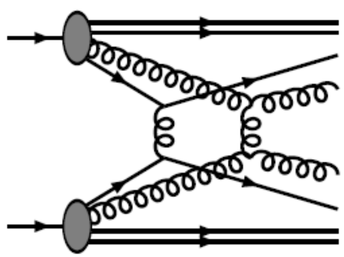


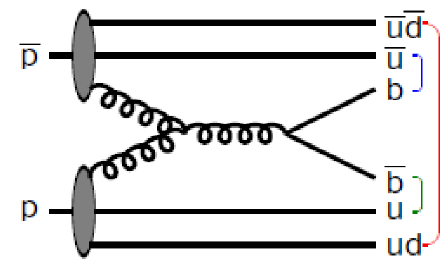
Virtual + Unresolvable emission **Finite**

The underlying event

This is defined as whatever else is in a p-p collision with the exception of the hardest process

UE = Multiple parton-parton interactions (MPI) + Beam remnants, with colour connections

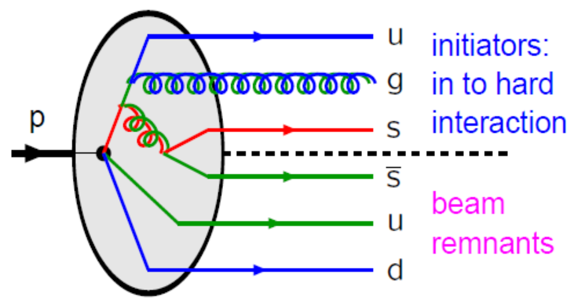




They happen "after" the hard interaction

→ need to rescale the PDFs in order to extract other partons in a correct way

→ primary partons are colour connected to beam remnants, this can have consequences in the hadronization phase



Need to assign:

- correlated flavours
- correlated $x_i = p_{zi}/p_{z\text{tot}}$
- correlated primordial $k_{\perp i}$
- correlated colours
- correlated showers

0) Squeeze range $0 < x < 1$ into $0 < x < 1 - \sum x_i$ (ISR: $i \neq i_{\text{current}}$)

1) Valence quarks: scale down by number already kicked out

2) Introduce companion quark $q/\overline{q}$ to each kicked-out sea quark $\overline{q}/q$, with x based on assumed $g \to q\overline{q}$ splitting

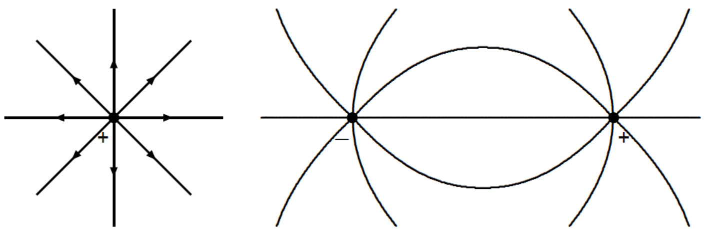
3) Gluon and other sea: rescale for total momentum conservation

Hadronization

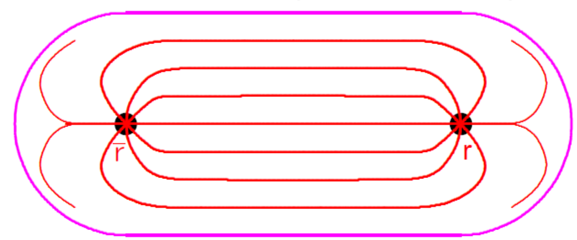
After the showering, at a factorization scale Q, the event is left with a multitude of partons (quarks, gluons, but also leptons and photons)

→ The coloured partons need further evolution down to colourless hadrons

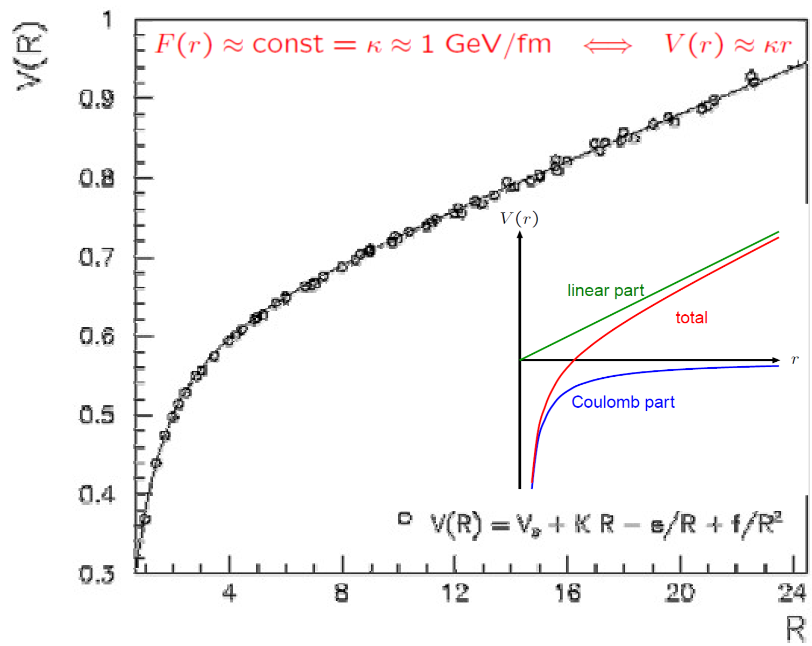
In QED field lines go to infinity, γs do not interact with each other



In QCD field lines are compressed in tube-like regions ("strings")



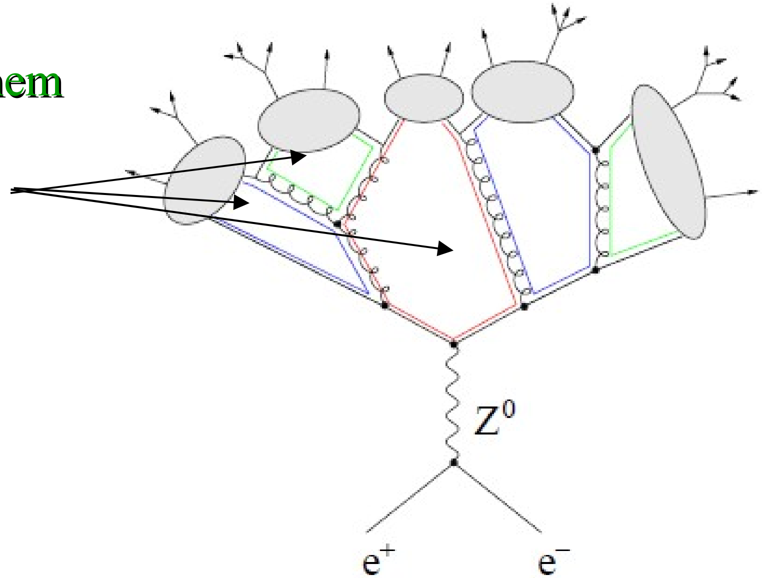
- Two coloured partons are confined linearly, and they cannot freely propagate.
- Over time, two phenomenological models of the hadronization mechanism have survived: the cluster model (implemented in HERWIG) and the lund or string model (implemented in PYTHIA)

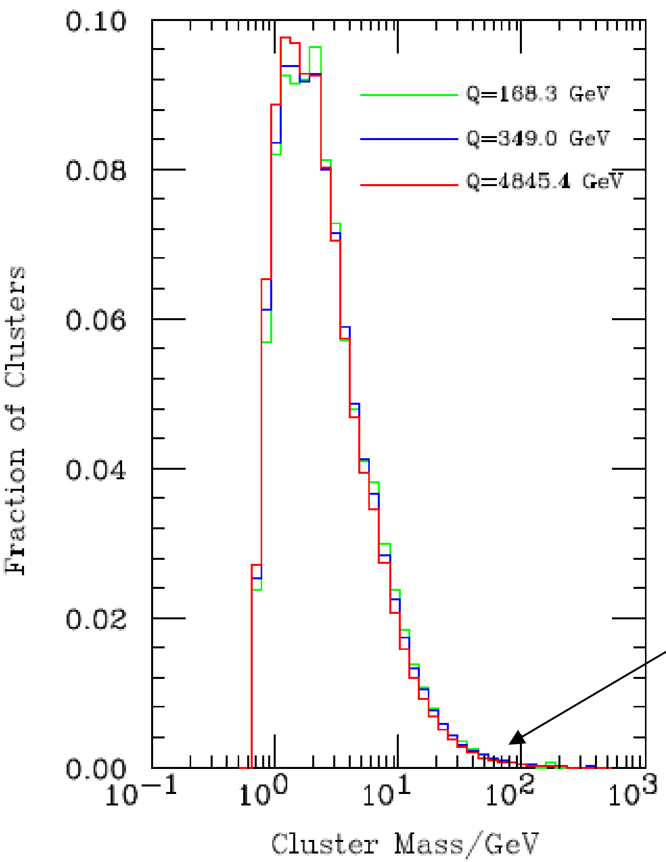


Hadronization: the cluster model

Perturbative evolution of quarks and gluons organizes them into clumps of colour-singlet clusters

- sort of "pre-confinement" with local colour flows
- gluon = colour-anticolour pair





Clusters decay isotropically to two hadrons according to phase space weights (depending on momentum)

- with a forced g→qq branching
- tail to very large mass clusters, need iterative procedure to split big clusters further
- baryon and strangeness production is automatically suppressed

The cutoff separating parton shower and hadronization becomes a fundamental parameter

Hadronization: the string model

PDF PS UE **Hadron** Decay

The confinement string is described by a 1D string with Lorentz invariant formalism

New pairs of quarks are produced via a tunnelling mechanism ($E_{string} \rightarrow q\bar{q}$ pair)

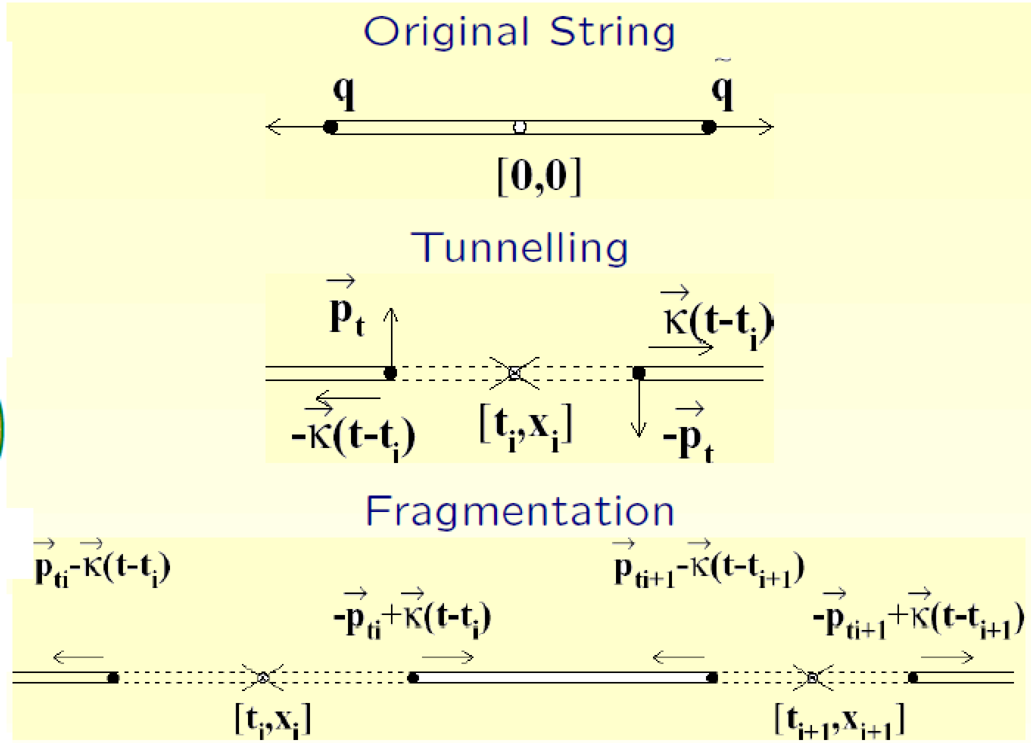
$$\mathcal{P} \propto \exp\left(-\frac{\pi m_{\perp q}^2}{\kappa}\right) = \exp\left(-\frac{\pi p_{\perp q}^2}{\kappa}\right) \exp\left(-\frac{\pi m_q^2}{\kappa}\right)$$

+fragmentation function (with tunable parameters)

Suppression of heavy particles:

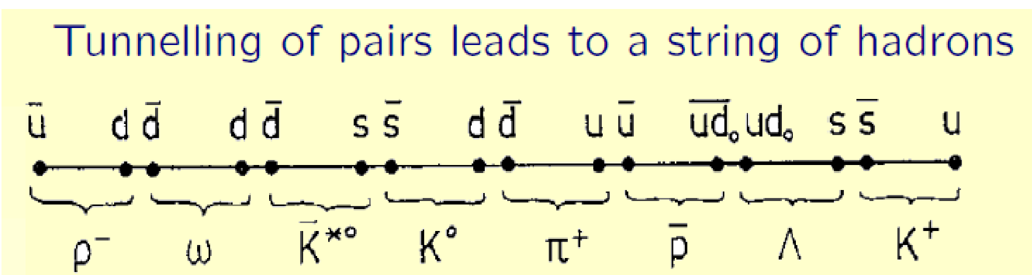
$$u\bar{u} : d\bar{d} : s\bar{s} : c\bar{c} \approx 1 : 1 : 0.3 : 10^{-11}$$

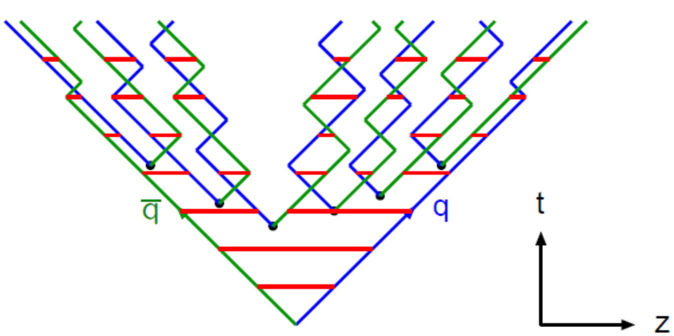
Moderate predictive power for hadron composition (many parameters to be tuned)



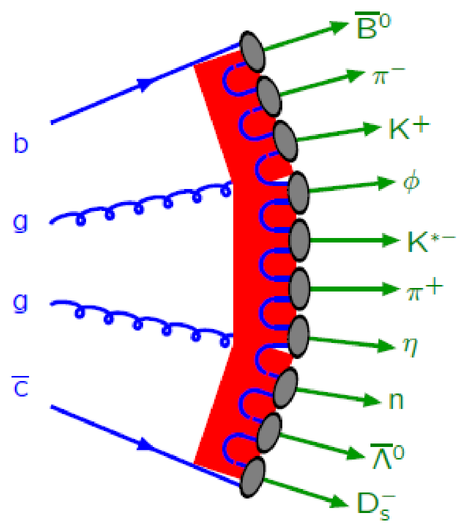
$$E_{had} = \kappa|x_i - x_{i+1}| \quad \vec{p}_{had} = \vec{p}_T + \vec{\kappa}(t_i - t_{i+1})$$

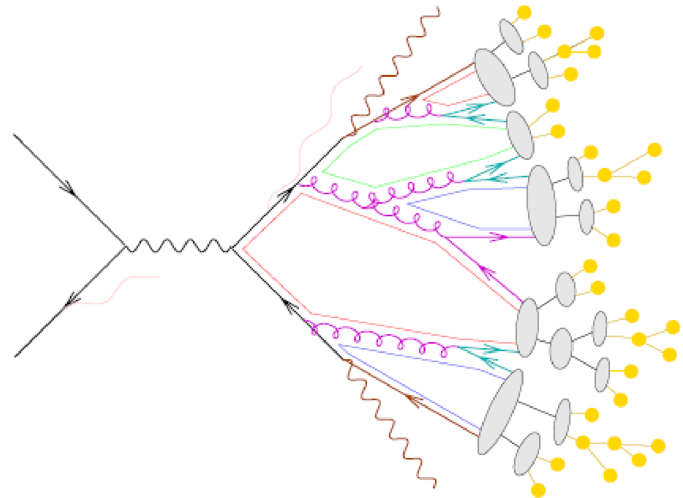
$m_{had}^2 \propto$ area swept out by string





Hadronization overview





program	PYTHIA	HERWIG
model	string	cluster
energy–momentum picture	powerful	simple
	predictive	unpredictive
parameters	few	many
flavour composition	messy	simple
	unpredictive	in-between
parameters	many	few

Parton Shower and Hadronization tunings

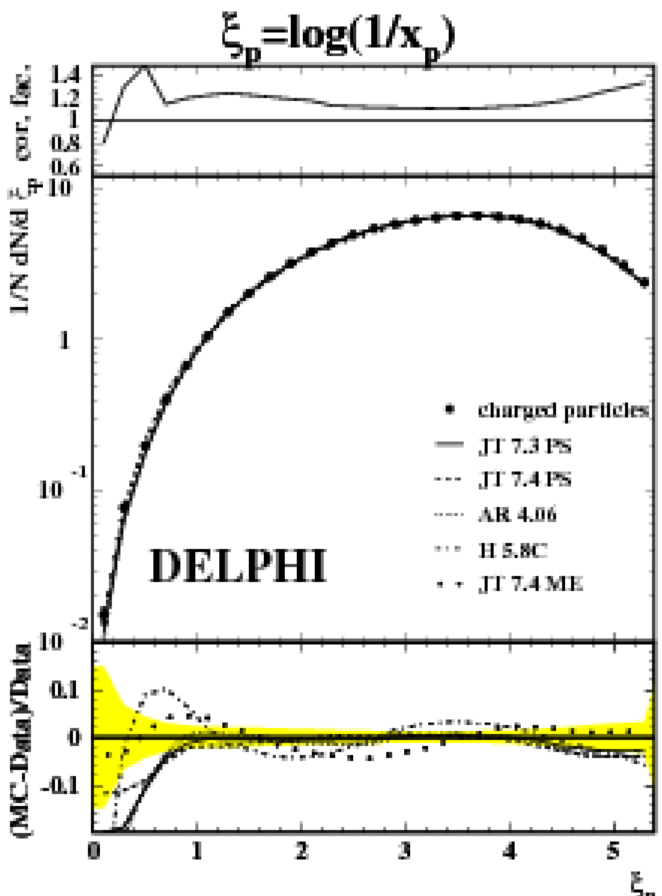
Typically tuned together, they are linked by the factorisation scale (tunable as well!)

Every experiment made its own tunings so far:

- experiments in e+e- collisions: particle densities/multiplicities, event shapes, jet shapes, jet flavour contents
- experiments in hadronic collisions: particle densities/multiplicities, jet shapes

Attempt to put tunings together in the LHC era, using LHC data as well

$$T = \max_{\vec{n}_{Thrust}} \frac{\sum_{i=1}^{N_{particle}} |\vec{p}_i \cdot \vec{n}_{Thrust}|}{\sum_{i=1}^{N_{particle}} |\vec{p}_i|}$$



From a global fit to all observables chosen (accounting for correlations as well) all the needed parameters of a model can be extracted with errors.

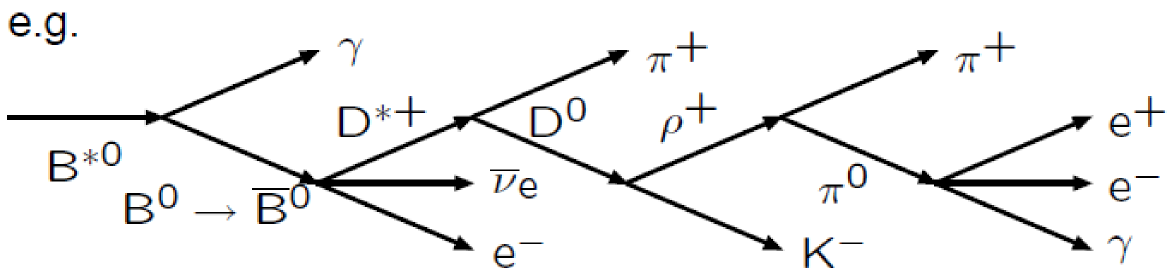
Here the example of PYTHIA

→ At the LHC this effort will need to be repeated before using a model !

Decay

Unspectacular but necessary step: hadron decays.

This is where most of the final particles are produced. Involves hundreds of particle kinds (boring to implement) and thousands of decay modes…



- $B^{*0} \to B^0\gamma$: electromagnetic decay
- $B^0 \to \overline{B}^0$ mixing (weak)
- $\overline{B}^0 \to D^{*+}\overline{\nu}_e e^-$: weak decay, displaced vertex, $|\mathcal{M}|^2 \propto (p_{\overline{B}} p_{\overline{\nu}})(p_e p_{D^*})$
- $D^{*+} \to D^0\pi^+$: strong decay
- $D^0 \to \rho^+ K^-$: weak decay, displaced vertex, ρ mass smeared
- $\rho^+ \to \pi^+\pi^0$: ρ polarized, $|\mathcal{M}|^2 \propto \cos^2\theta$ in ρ rest frame
- $\pi^0 \to e^+e^-\gamma$: Dalitz decay, $m(e^+e^-)$ peaked

Dedicated programs, with special attention to polarization effects:

- EVTGEN: B decays
- TAUOLA: τ decays

Event Generators

Event generators

Multi purpose (PYTHIA, HERWIG, ISAJET…) event generators

- simple LO ME description of 2→2,3 processes, with large coverage of the possible phase space
- PS radiation and hadronization included, output is an hadron level event

Multi purpose N-fermion parton-level generators (HELAC, MadGraph/MadEvent, Sherpa, Whizard+Omega, compHEP…)

- N-fermion ME, with automatic calculations of amplitudes
- can cover all possible final states, but no channel specific optimization: they can require a lot of CPU for complicated final states
- need interface from parton level to final states (PS+hadronization): typically PYTHIA or HERWIG

Dedicated N-fermion parton-level generators (Alpgen, PHANTOM, POWHEG, MC@NLO…)

- optimized for specific processes, not necessarily full coverage of processes
- designed for inclusion of specific features (ie NLO corrections)
- can typically work with pre-made libraries: fast and efficient
- need interface from parton level to final states (PS+hadronization): typically PYTHIA or HERWIG

→ One needs to use the tools that are more suited for the description of the interesting physics signal and backgrounds

Example 1: the PYTHIA process library

No.	Subprocess
Hard QCD processes:	
11	$f_i f_j \to f_i f_j$
12	$f_i \bar{f}_i \to f_k \bar{f}_k$
13	$f_i \bar{f}_i \to gg$
28	$f_i g \to f_i g$
53	$gg \to f_k \bar{f}_k$
68	$gg \to gg$
Soft QCD processes:	
91	elastic scattering
92	single diffraction (XB)
93	single diffraction (AX)
94	double diffraction
95	low-$p_\perp$ production
Open heavy flavour: (also fourth generation)	
81	$f_i \bar{f}_i \to Q_k \bar{Q}_k$
82	$gg \to Q_k \bar{Q}_k$
83	$q_i f_j \to Q_k f_l$
84	$g\gamma \to Q_k \bar{Q}_k$
85	$\gamma\gamma \to F_k \bar{F}_k$
Closed heavy flavour:	
86	$gg \to J/\psi g$
87	$gg \to \chi_{0c} g$
88	$gg \to \chi_{1c} g$
89	$gg \to \chi_{2c} g$
104	$gg \to \chi_{0c}$
105	$gg \to \chi_{2c}$
106	$gg \to J/\psi\gamma$
107	$g\gamma \to J/\psi g$
108	$\gamma\gamma \to J/\psi\gamma$
W/Z production:	
1	$f_i \bar{f}_i \to \gamma^*/Z^0$
2	$f_i \bar{f}_j \to W^\pm$
22	$f_i \bar{f}_i \to Z^0 Z^0$
23	$f_i \bar{f}_j \to Z^0 W^\pm$
25	$f_i \bar{f}_i \to W^+ W^-$
15	$f_i \bar{f}_i \to g Z^0$
16	$f_i \bar{f}_j \to g W^\pm$
30	$f_i g \to f_i Z^0$
31	$f_i g \to f_k W^\pm$
19	$f_i \bar{f}_i \to \gamma Z^0$
20	$f_i \bar{f}_j \to \gamma W^\pm$
35	$f_i \gamma \to f_i Z^0$

No.	Subprocess
36	$f_i \gamma \to f_k W^\pm$
69	$\gamma\gamma \to W^+ W^-$
70	$\gamma W^\pm \to Z^0 W^\pm$
Prompt photons:	
14	$f_i \bar{f}_i \to g\gamma$
18	$f_i \bar{f}_i \to \gamma\gamma$
29	$f_i g \to f_i \gamma$
114	$gg \to \gamma\gamma$
115	$gg \to g\gamma$
Deeply Inel. Scatt.:	
10	$f_i f_j \to f_k f_l$
99	$\gamma^* q \to q$
Photon-induced:	
33	$f_i \gamma \to f_i g$
34	$f_i \gamma \to f_i \gamma$
54	$g\gamma \to f_k \bar{f}_k$
58	$\gamma\gamma \to f_k \bar{f}_k$
131	$f_i \gamma^*_T \to f_i g$
132	$f_i \gamma^*_L \to f_i g$
133	$f_i \gamma^*_T \to f_i \gamma$
134	$f_i \gamma^*_L \to f_i \gamma$
135	$g \gamma^*_T \to f_i \bar{f}_i$
136	$g \gamma^*_L \to f_i \bar{f}_i$
137	$\gamma^*_T \gamma^*_T \to f_i \bar{f}_i$
138	$\gamma^*_T \gamma^*_L \to f_i \bar{f}_i$
139	$\gamma^*_L \gamma^*_T \to f_i \bar{f}_i$
140	$\gamma^*_L \gamma^*_L \to f_i \bar{f}_i$
80	$q_i \gamma \to q_k \pi^\pm$
Light SM Higgs:	
3	$f_i \bar{f}_i \to h^0$
24	$f_i \bar{f}_i \to Z^0 h^0$
26	$f_i \bar{f}_j \to W^\pm h^0$
32	$f_i g \to f_i h^0$
102	$gg \to h^0$
103	$\gamma\gamma \to h^0$
110	$f_i \bar{f}_i \to \gamma h^0$
111	$f_i \bar{f}_i \to g h^0$
112	$f_i g \to f_i h^0$
113	$gg \to g h^0$
121	$gg \to Q_k \bar{Q}_k h^0$
122	$q_i \bar{q}_i \to Q_k \bar{Q}_k h^0$
123	$f_i f_j \to f_i f_j h^0$
124	$f_i f_j \to f_k f_l h^0$

No.	Subprocess
New gauge bosons:	
141	$f_i \bar{f}_i \to \gamma/Z^0/Z'^0$
142	$f_i \bar{f}_j \to W'^+$
144	$f_i \bar{f}_j \to R$
Heavy SM Higgs:	
5	$Z^0 Z^0 \to h^0$
8	$W^+ W^- \to h^0$
71	$Z^0_L Z^0_L \to Z^0_L Z^0_L$
72	$Z^0_L Z^0_L \to W^+_L W^-_L$
73	$Z^0_L W^\pm_L \to Z^0_L W^\pm_L$
76	$W^+_L W^-_L \to Z^0_L Z^0_L$
77	$W^\pm_L W^\pm_L \to W^\pm_L W^\pm_L$
BSM Neutral Higgs:	
151	$f_i \bar{f}_i \to H^0$
152	$gg \to H^0$
153	$\gamma\gamma \to H^0$
171	$f_i \bar{f}_i \to Z^0 H^0$
172	$f_i \bar{f}_j \to W^\pm H^0$
173	$f_i f_j \to f_i f_j H^0$
174	$f_i f_j \to f_k f_l H^0$
181	$gg \to Q_k \bar{Q}_k H^0$
182	$q_i \bar{q}_i \to Q_k \bar{Q}_k H^0$
183	$f_i \bar{f}_i \to g H^0$
184	$f_i g \to f_i H^0$
185	$gg \to g H^0$
156	$f_i \bar{f}_i \to A^0$
157	$gg \to A^0$
158	$\gamma\gamma \to A^0$
176	$f_i \bar{f}_i \to Z^0 A^0$
177	$f_i \bar{f}_j \to W^\pm A^0$
178	$f_i f_j \to f_i f_j A^0$
179	$f_i f_j \to f_k f_l A^0$
186	$gg \to Q_k \bar{Q}_k A^0$
187	$q_i \bar{q}_i \to Q_k \bar{Q}_k A^0$
188	$f_i \bar{f}_i \to g A^0$
189	$f_i g \to f_i A^0$
190	$gg \to g A^0$
Charged Higgs:	
143	$f_i \bar{f}_j \to H^+$
161	$f_i g \to f_k H^+$
401	$gg \to \bar{t} b H^+$
402	$q\bar{q} \to \bar{t} b H^+$

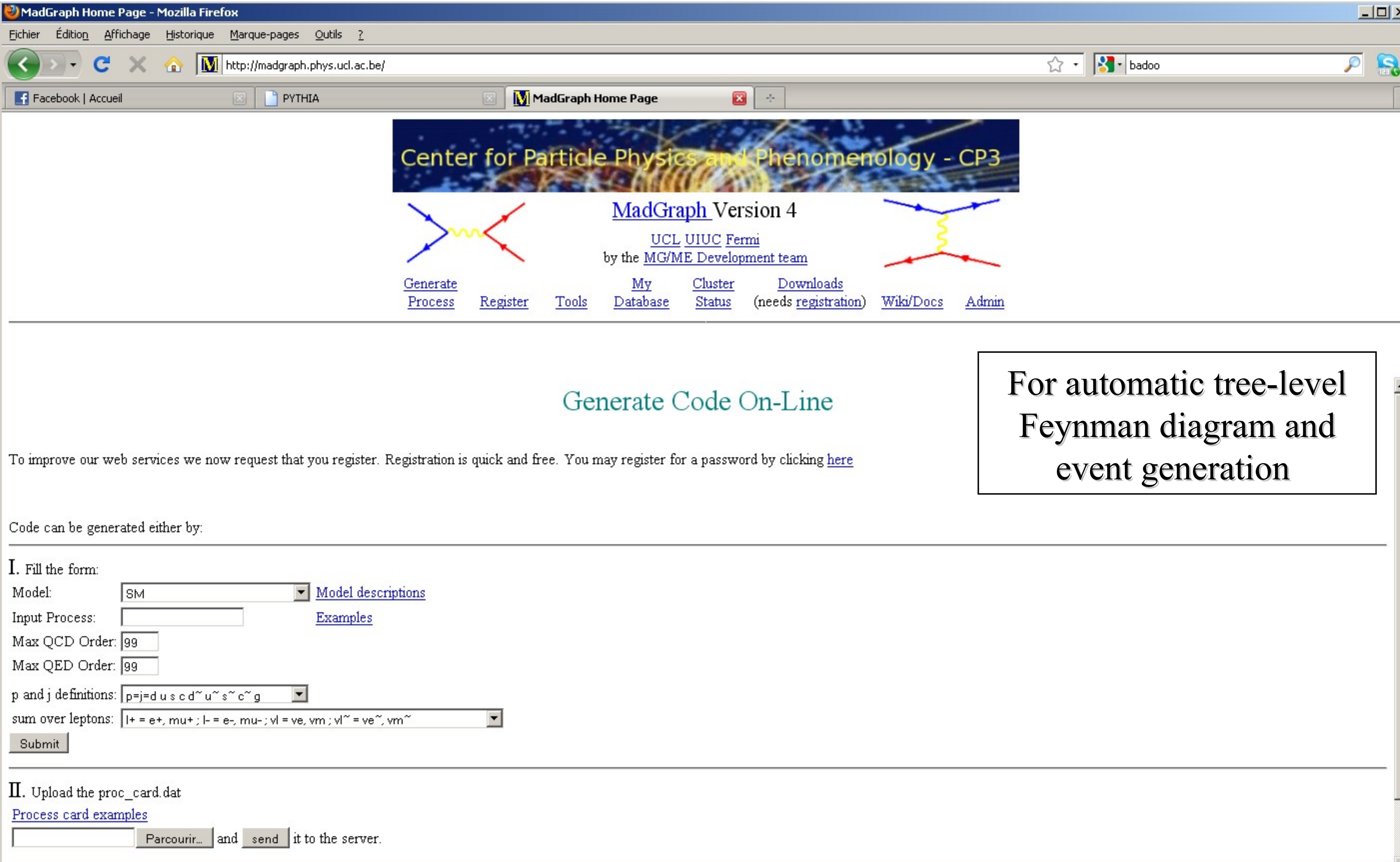
No.	Subprocess
Higgs pairs:	
297	$f_i \bar{f}_j \to H^\pm h^0$
298	$f_i \bar{f}_j \to H^\pm H^0$
299	$f_i \bar{f}_i \to A^0 h^0$
300	$f_i \bar{f}_i \to A^0 H^0$
301	$f_i \bar{f}_i \to H^+ H^-$
Leptoquarks:	
145	$q_i \ell_j \to L_Q$
162	$qg \to \ell L_Q$
163	$gg \to L_Q \bar{L}_Q$
164	$q_i \bar{q}_i \to L_Q \bar{L}_Q$
Technicolor:	
149	$gg \to \eta_{tc}$
191	$f_i \bar{f}_i \to \rho^0_{tc}$
192	$f_i \bar{f}_j \to \rho^+_{tc}$
193	$f_i \bar{f}_i \to \omega^0_{tc}$
194	$f_i \bar{f}_i \to f_k \bar{f}_k$
195	$f_i \bar{f}_j \to f_k \bar{f}_l$
361	$f_i \bar{f}_i \to W^+_L W^-_L$
362	$f_i \bar{f}_i \to W^\pm_L \pi^\mp_{tc}$
363	$f_i \bar{f}_i \to \pi^+_{tc} \pi^-_{tc}$
364	$f_i \bar{f}_i \to \gamma \pi^0_{tc}$
365	$f_i \bar{f}_i \to \gamma \pi'^0_{tc}$
366	$f_i \bar{f}_i \to Z^0 \pi^0_{tc}$
367	$f_i \bar{f}_i \to Z^0 \pi'^0_{tc}$
368	$f_i \bar{f}_i \to W^\pm \pi^\mp_{tc}$
370	$f_i \bar{f}_j \to W^\pm_L Z^0_L$
371	$f_i \bar{f}_j \to W^\pm_L \pi^0_{tc}$
372	$f_i \bar{f}_j \to \pi^\pm_{tc} Z^0_L$
373	$f_i \bar{f}_j \to \pi^\pm_{tc} \pi^0_{tc}$
374	$f_i \bar{f}_j \to \gamma \pi^\pm_{tc}$
375	$f_i \bar{f}_j \to Z^0 \pi^\pm_{tc}$
376	$f_i \bar{f}_j \to W^\pm \pi^0_{tc}$
377	$f_i \bar{f}_j \to W^\pm \pi'^0_{tc}$
381	$q_i q_j \to q_i q_j$
382	$q_i \bar{q}_i \to q_k \bar{q}_k$
383	$q_i \bar{q}_i \to gg$
384	$f_i g \to f_i g$
385	$gg \to q_k \bar{q}_k$
386	$gg \to gg$
387	$f_i \bar{f}_i \to Q_k \bar{Q}_k$
388	$gg \to Q_k \bar{Q}_k$

No.	Subprocess
Compositeness:	
146	$e\gamma \to e^*$
147	$dg \to d^*$
148	$ug \to u^*$
167	$q_i q_j \to d^* q_k$
168	$q_i q_j \to u^* q_k$
169	$q_i \bar{q}_i \to e^\pm e^{*\mp}$
165	$f_i \bar{f}_i (\to \gamma^*/Z^0) \to f_k \bar{f}_k$
166	$f_i \bar{f}_j (\to W^\pm) \to f_k \bar{f}_l$
Extra Dimensions:	
391	$f\bar{f} \to G^*$
392	$gg \to G^*$
393	$q\bar{q} \to gG^*$
394	$qg \to qG^*$
395	$gg \to gG^*$
Left–right symmetry:	
341	$\ell_i \ell_j \to H^{\pm\pm}_L$
342	$\ell_i \ell_j \to H^{\pm\pm}_R$
343	$\ell^\pm_i \gamma \to H^{\pm\pm}_L e^\mp$
344	$\ell^\pm_i \gamma \to H^{\pm\pm}_R e^\mp$
345	$\ell^\pm_i \gamma \to H^{\pm\pm}_L \mu^\mp$
346	$\ell^\pm_i \gamma \to H^{\pm\pm}_R \mu^\mp$
347	$\ell^\pm_i \gamma \to H^{\pm\pm}_L \tau^\mp$
348	$\ell^\pm_i \gamma \to H^{\pm\pm}_R \tau^\mp$
349	$f_i \bar{f}_i \to H^{++}_L H^{--}_L$
350	$f_i \bar{f}_i \to H^{++}_R H^{--}_R$
351	$f_i f_j \to f_k f_l H^{\pm\pm}_L$
352	$f_i f_j \to f_k f_l H^{\pm\pm}_R$
353	$f_i \bar{f}_i \to Z^0_R$
354	$f_i \bar{f}_j \to W^\pm_R$
SUSY:	
201	$f_i \bar{f}_i \to \tilde{e}_L \tilde{e}^*_L$
202	$f_i \bar{f}_i \to \tilde{e}_R \tilde{e}^*_R$
203	$f_i \bar{f}_i \to \tilde{e}_L \tilde{e}^*_R +$
204	$f_i \bar{f}_i \to \tilde{\mu}_L \tilde{\mu}^*_L$
205	$f_i \bar{f}_i \to \tilde{\mu}_R \tilde{\mu}^*_R$
206	$f_i \bar{f}_i \to \tilde{\mu}_L \tilde{\mu}^*_R +$
207	$f_i \bar{f}_i \to \tilde{\tau}_1 \tilde{\tau}^*_1$
208	$f_i \bar{f}_i \to \tilde{\tau}_2 \tilde{\tau}^*_2$
209	$f_i \bar{f}_i \to \tilde{\tau}_1 \tilde{\tau}^*_2 +$

No.	Subprocess
210	$f_i \bar{f}_j \to \tilde{\ell}_L \tilde{\nu}^*_\ell +$
211	$f_i \bar{f}_j \to \tilde{\tau}_1 \tilde{\nu}^*_\tau +$
212	$f_i \bar{f}_j \to \tilde{\tau}_2 \tilde{\nu}^*_\tau +$
213	$f_i \bar{f}_i \to \tilde{\nu}_\ell \tilde{\nu}_\ell^*$
214	$f_i \bar{f}_i \to \tilde{\nu}_\tau \tilde{\nu}^*_\tau$
216	$f_i \bar{f}_i \to \tilde{\chi}_1 \tilde{\chi}_1$
217	$f_i \bar{f}_i \to \tilde{\chi}_2 \tilde{\chi}_2$
218	$f_i \bar{f}_i \to \tilde{\chi}_3 \tilde{\chi}_3$
219	$f_i \bar{f}_i \to \tilde{\chi}_4 \tilde{\chi}_4$
220	$f_i \bar{f}_i \to \tilde{\chi}_1 \tilde{\chi}_2$
221	$f_i \bar{f}_i \to \tilde{\chi}_1 \tilde{\chi}_3$
222	$f_i \bar{f}_i \to \tilde{\chi}_1 \tilde{\chi}_4$
223	$f_i \bar{f}_i \to \tilde{\chi}_2 \tilde{\chi}_3$
224	$f_i \bar{f}_i \to \tilde{\chi}_2 \tilde{\chi}_4$
225	$f_i \bar{f}_i \to \tilde{\chi}_3 \tilde{\chi}_4$
226	$f_i \bar{f}_i \to \tilde{\chi}^\pm_1 \tilde{\chi}^\mp_1$
227	$f_i \bar{f}_i \to \tilde{\chi}^\pm_2 \tilde{\chi}^\mp_2$
228	$f_i \bar{f}_i \to \tilde{\chi}^\pm_1 \tilde{\chi}^\mp_2$
229	$f_i \bar{f}_j \to \tilde{\chi}_1 \tilde{\chi}^\pm_1$
230	$f_i \bar{f}_j \to \tilde{\chi}_2 \tilde{\chi}^\pm_1$
231	$f_i \bar{f}_j \to \tilde{\chi}_3 \tilde{\chi}^\pm_1$
232	$f_i \bar{f}_j \to \tilde{\chi}_4 \tilde{\chi}^\pm_1$
233	$f_i \bar{f}_j \to \tilde{\chi}_1 \tilde{\chi}^\pm_2$
234	$f_i \bar{f}_j \to \tilde{\chi}_2 \tilde{\chi}^\pm_2$
235	$f_i \bar{f}_j \to \tilde{\chi}_3 \tilde{\chi}^\pm_2$
236	$f_i \bar{f}_j \to \tilde{\chi}_4 \tilde{\chi}^\pm_2$
237	$f_i \bar{f}_i \to \tilde{g} \tilde{\chi}_1$
238	$f_i \bar{f}_i \to \tilde{g} \tilde{\chi}_2$
239	$f_i \bar{f}_i \to \tilde{g} \tilde{\chi}_3$
240	$f_i \bar{f}_i \to \tilde{g} \tilde{\chi}_4$
241	$f_i \bar{f}_j \to \tilde{g} \tilde{\chi}^\pm_1$
242	$f_i \bar{f}_j \to \tilde{g} \tilde{\chi}^\pm_2$
243	$f_i \bar{f}_i \to \tilde{g} \tilde{g}$
244	$gg \to \tilde{g} \tilde{g}$
246	$f_i g \to \tilde{q}_{iL} \tilde{\chi}_1$
247	$f_i g \to \tilde{q}_{iR} \tilde{\chi}_1$
248	$f_i g \to \tilde{q}_{iL} \tilde{\chi}_2$
249	$f_i g \to \tilde{q}_{iR} \tilde{\chi}_2$

No.	Subprocess
250	$f_i g \to \tilde{q}_{iL} \tilde{\chi}_3$
251	$f_i g \to \tilde{q}_{iR} \tilde{\chi}_3$
252	$f_i g \to \tilde{q}_{iL} \tilde{\chi}_4$
253	$f_i g \to \tilde{q}_{iR} \tilde{\chi}_4$
254	$f_i g \to \tilde{q}_{jL} \tilde{\chi}^\pm_1$
256	$f_i g \to \tilde{q}_{jL} \tilde{\chi}^\pm_2$
258	$f_i g \to \tilde{q}_{iL} \tilde{g}$
259	$f_i g \to \tilde{q}_{iR} \tilde{g}$
261	$f_i \bar{f}_i \to \tilde{t}_1 \tilde{t}^*_1$
262	$f_i \bar{f}_i \to \tilde{t}_2 \tilde{t}^*_2$
263	$f_i \bar{f}_i \to \tilde{t}_1 \tilde{t}^*_2 +$
264	$gg \to \tilde{t}_1 \tilde{t}^*_1$
265	$gg \to \tilde{t}_2 \tilde{t}^*_2$
271	$f_i f_j \to \tilde{q}_{iL} \tilde{q}_{jL}$
272	$f_i f_j \to \tilde{q}_{iR} \tilde{q}_{jR}$
273	$f_i f_j \to \tilde{q}_{iL} \tilde{q}_{jR} +$
274	$f_i \bar{f}_j \to \tilde{q}_{iL} \tilde{q}^*_{jL}$
275	$f_i \bar{f}_j \to \tilde{q}_{iR} \tilde{q}^*_{jR}$
276	$f_i \bar{f}_j \to \tilde{q}_{iL} \tilde{q}^*_{jR} +$
277	$f_i \bar{f}_i \to \tilde{q}_{jL} \tilde{q}^*_{jL}$
278	$f_i \bar{f}_i \to \tilde{q}_{jR} \tilde{q}^*_{jR}$
279	$gg \to \tilde{q}_{iL} \tilde{q}^*_{iL}$
280	$gg \to \tilde{q}_{iR} \tilde{q}^*_{iR}$
281	$bq_i \to \tilde{b}_1 \tilde{q}_{iL}$
282	$bq_i \to \tilde{b}_2 \tilde{q}_{iR}$
283	$bq_i \to \tilde{b}_1 \tilde{q}_{iR} +$
284	$b\bar{q}_i \to \tilde{b}_1 \tilde{q}^*_{iL}$
285	$b\bar{q}_i \to \tilde{b}_2 \tilde{q}^*_{iR}$
286	$b\bar{q}_i \to \tilde{b}_1 \tilde{q}^*_{iR} +$
287	$f_i \bar{f}_i \to \tilde{b}_1 \tilde{b}^*_1$
288	$f_i \bar{f}_i \to \tilde{b}_2 \tilde{b}^*_2$
289	$gg \to \tilde{b}_1 \tilde{b}^*_1$
290	$gg \to \tilde{b}_2 \tilde{b}^*_2$
291	$bb \to \tilde{b}_1 \tilde{b}_1$
292	$bb \to \tilde{b}_2 \tilde{b}_2$
293	$bb \to \tilde{b}_1 \tilde{b}_2$
294	$bg \to \tilde{b}_1 \tilde{g}$
295	$bg \to \tilde{b}_2 \tilde{g}$
296	$b\bar{b} \to \tilde{b}_1 \tilde{b}^*_2 +$

Example 2: MadGraph automatisation



For automatic tree-level Feynman diagram and event generation

Huge theory community

Developments of modern generators (for the LHC) correspond to the effort of teams of theorists. An incomplete, SM biased, list is here:

	Language	Order of the calculation	Max. number of final particles	Physics signals beyond the SM	Type of program	Site/Documentation
HERWIG	F, C++	LO	3	Yes	Generic library	hepwww.rl.ac.uk/theory/seymour/herwig
PYTHIA	F, C++	LO	3	Yes	Generic library	www.thep.lu.se/~torbjorn/Pythia.html
SHERPA	C++	LO	6	Yes	Generic automatic	www.physik.tu-dresden.de/~krauss/hep
ALPGEN	Fortran	LO	7	No	Generic library	mlm.home.cern.ch/mlm/alpgen
MadGraph	F+web	LO	5	Yes	Generic automatic	madgraph.physics.uiuc.edu
CompHEP	Fortran	LO	4	Yes	Generic automatic	theory.sinp.nsu.ru/comphep
POWHEG	Fortran	NLO	3	No	Many processes	mobydick.mib.infn.it/~nason/POWHEG
MC@NLO	Fortran	NLO		No	Many processes	www.hep.phy.cam.ac.uk/theory/webber/MCatNLO
MCFM	Fortran	NLO		No	Many processes	mcfm.fnal.gov
GR@PPA	Fortran	LO		No	Several processes	atlas.kek.jp/physics/nlo-wg/grappa.html
TopRex	Fortran	LO		No	Dedicated top	cmsdoc.cern.ch/~spitsky/toprex/toprex.html
AcerMC	Fortran	LO		No	A few processes	borut.home.cern.ch/borut
PHANTOM	Fortran	LO	6	No	All SM processes	arxiv.org/PS_cache/arxiv/pdf/0801/0801.3359v2.pdf
PHOX	Fortran	LO		No	A few processes	wwwlapp.in2p3.fr/lapth/PHOX_FAMILY/main.html

Incomplete list !

Typical choices at the LHC (I)

Workhorses are PYTHIA and HERWIG (full events in one go)

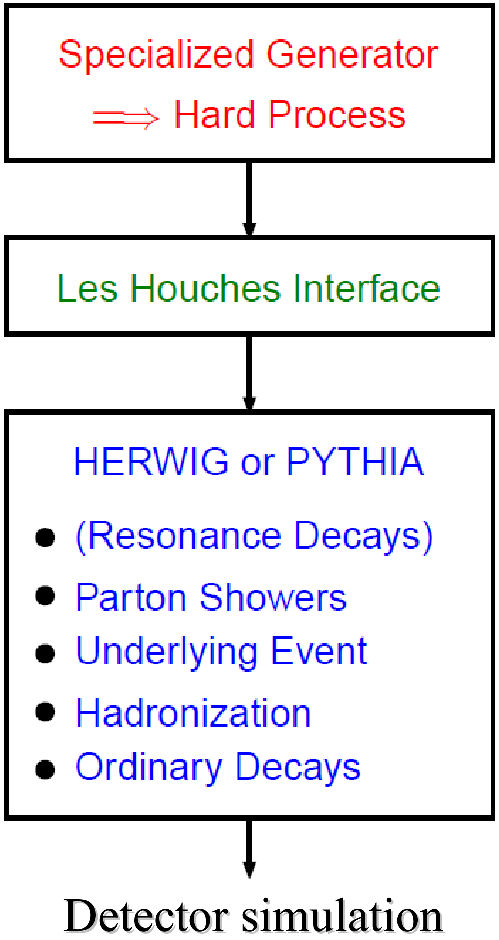
- (soft) QCD, samples for UE tuning and other Monte Carlo tunings
- SM physics for detector calibration and simple backgrounds
- simple new physics processes

Use dedicated codes for the description of high jet multiplicity, like ALPGEN, Sherpa, MadGraph, HELAC, CompHEP

- needed for W+jets, Z/γ+jets, top production, multi boson, and all QCD at high jet multiplicities (see later)
- some can also deal with many new physics signals
- use standard format interface (LHI) for passing events to the further steps in the generation chain

Use dedicated codes for NLO calculations like POWHEG, MC@NLO

- QCD NLO calculation in generators (see later), only for SM
- get shape and normalizations more correct
- useful to get the first QCD emission right, needed for W+jets, Z/γ+jets, top production
- use standard format interface (LHI) for passing events to the further steps in the generation chain



Typical choices at the LHC (II)

Use dedicated codes that are process specific as well, especially for the description of physics beyond the SM (HIGLU, Charybdis, PROSPINO, DIPHOX, …)

- specific features, handy variation of the (specific) new physics parameters
- use standard format interface (LHI) for passing events to the further steps in the generation chain

The physics program at the LHC is very rich, other event generators must be used for specific physics or for the understanding of the detector or calibration purposes

- Diffractive physics (Pomwig, Exhume, EDDE), typically add-on to PYTHIA
- Heavy Ions physics (Hydjet, Pyquen), also add-on to PYTHIA
- Cosmic rays generators
- Generators for beam halo and beam-gas interactions
- Particle "guns" (way to run generic generators like PYTHIA)

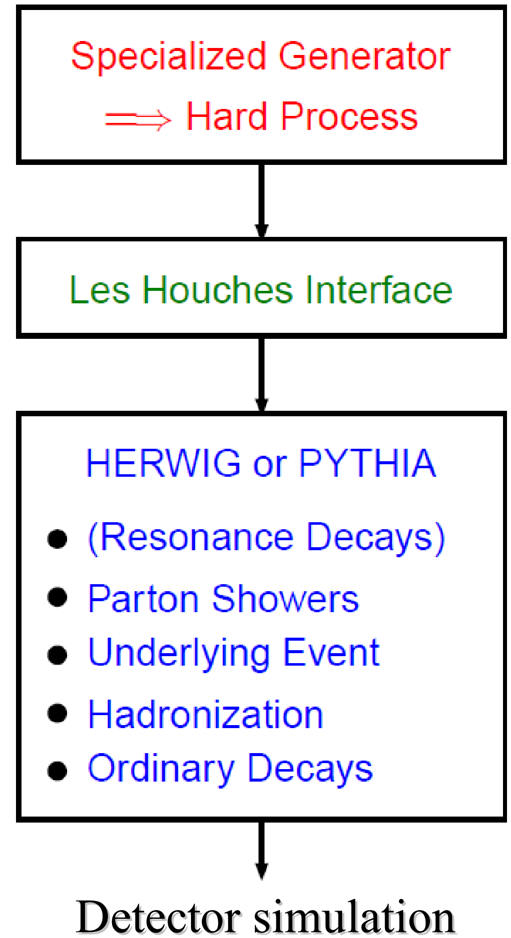
Though PYTHIA and HERWIG can handle the full generation chain, specific tools can be/are typically used for decay handling:

TAUOLA (τ decays) [used when the description of τ decays is relevant]

PHOTOS (QED corrections) [used for inclusion of real QED emissions]

EvtGen (for B hadron decays) [used for precise description of B decays]

In view of the LHC, in the last years there has been enormous progress in the development of ME generators → next slides



Matrix Elements versus Parton Showers

It is essential to understand which techniques are applicable to which kinematic regime, and to make the right choices.

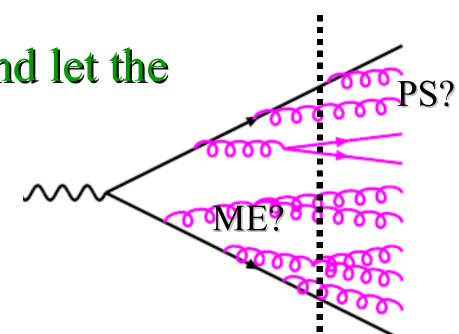
- Parton Shower: infinite serie in α_S keeping only singular terms (collinear approx.):
  - ✥ Excellent at low p_T, with emission at any order, simple interface with hadronization
  - ✥ Large uncertainties away from singular regions
  - ✥ To be used for soft (compared to signal scale) jets
- Fixed order matrix elements: truncated expansion in α_S
  - ✥ Exact expansion in α_S: full helicity structure to the given order
  - ✥ Calculations can easily become very tough
  - ✥ To be used for hard (compared to signal scale) jets

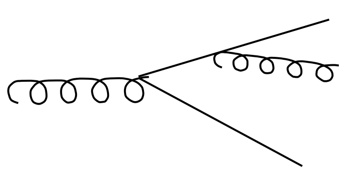
→ High jet multiplicity events are bound to be better described by ME

→ Jet structures and soft QCD are bound to be better described by PS

To which (QCD) order should we stop calculating the matrix elements and let the showering describe extra emissions?

Are there ways to put together the benefits of PS and ME, avoiding to double count processes (ME_N+PS has parts of ME_{N+1})?





This can be either a gluon from PS out of the qq system or generated directly via the qqg ME

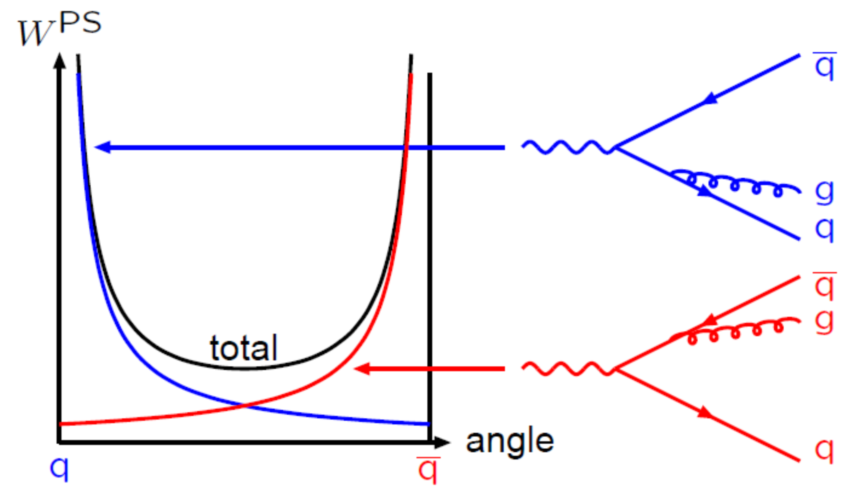
Reweighting PS to ME

This is a simple and convenient way to get the first extra emission more correct without using complex ME tools (technique implemented in general purpose event generators like PYTHIA or HERWIG)

The aim is to cover the full phase space with a smooth transition ME/PS

- just use Parton Shower for process pp→X
- correct it in order to get the right pp→X+g rates and shapes. This reweighting must be done only for the very first branching in the event
- do it for both initial or final showers

$$\overbrace{W^{\mathrm{ME}}}^{\text{wanted}} = \overbrace{W^{\mathrm{PS}}}^{\text{generated}} \overbrace{\frac{W^{\mathrm{ME}}}{W^{\mathrm{PS}}}}^{\text{correction}}$$



Most used multipurpose generators like HERWIG and PYTHIA implement the reweighting for the first gluon emission for some processes only, getting more correct shapes for the first gluon emission

Problems:

- this is an approximation, working only for a few processes
- problem of reweighting in general: you can reweight only populated phase space

If the PS leaves "holes" in the gluon phase space, there is nothing to reweight

Matching ME and PS

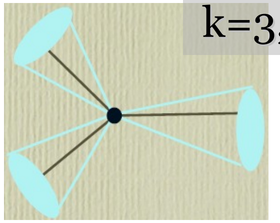
Powerful methods exist to really match PS and ME: CKKW [Catani, Krauss, Kuhn, Webber] (implemented in codes like Sherpa) and MLM [Mangano] (implemented in codes like ALPGEN and MadGraph). Here only the second is discussed

Practical technique to avoid ME and PS double counting in the generation at any order:

1. generate the events at order k in QCD → Npart
2. shower the events (no hadronisation) and cluster them into jets structures → Njets
3. match partons from the ME to the clustered jets (i.e. requiring $\Delta R(\text{parton,jet})<R_{cut}$) → Nmatch
4. if not all partons are matched to jets, discard the event

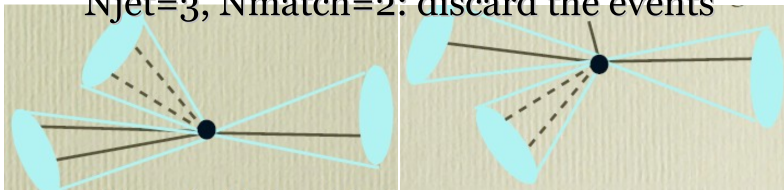
Works independent of the generation procedure

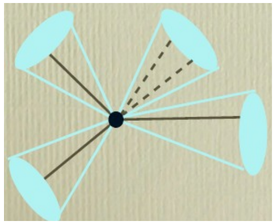
k=3, Npart=3



Njet=3, Nmatch=3: keep the event

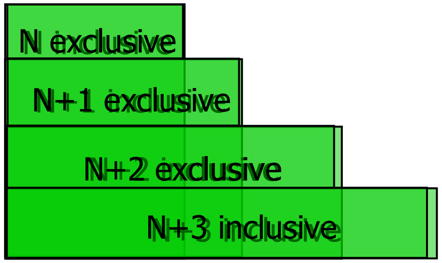
Njet=3, Nmatch=2: discard the events





Njet=4, Nmatch=3: keep the event only for inclusive matching

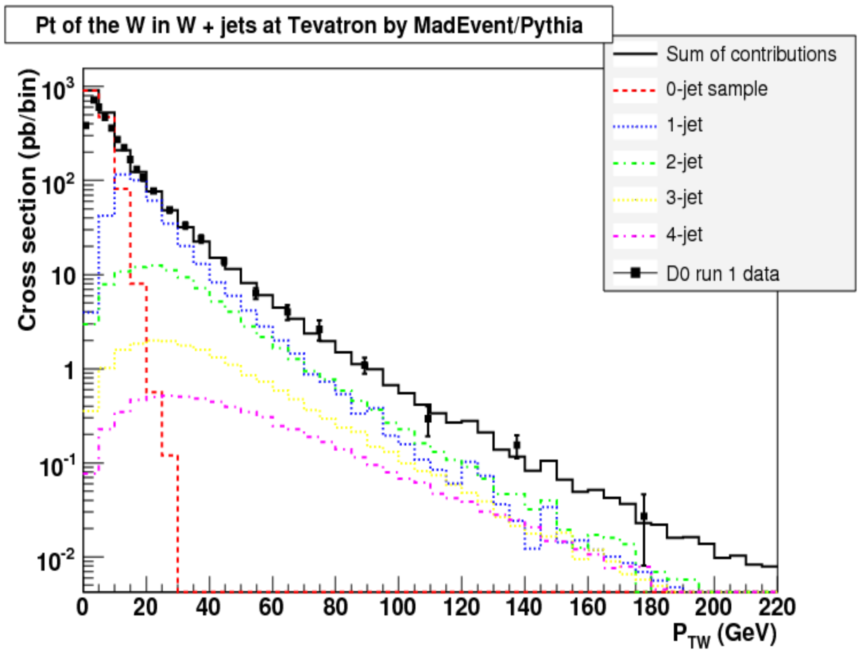
1. ME_N+PS
2. ME_{N+1}+PS
3. ME_{N+2}+PS
4. ME_{N+3}+PS

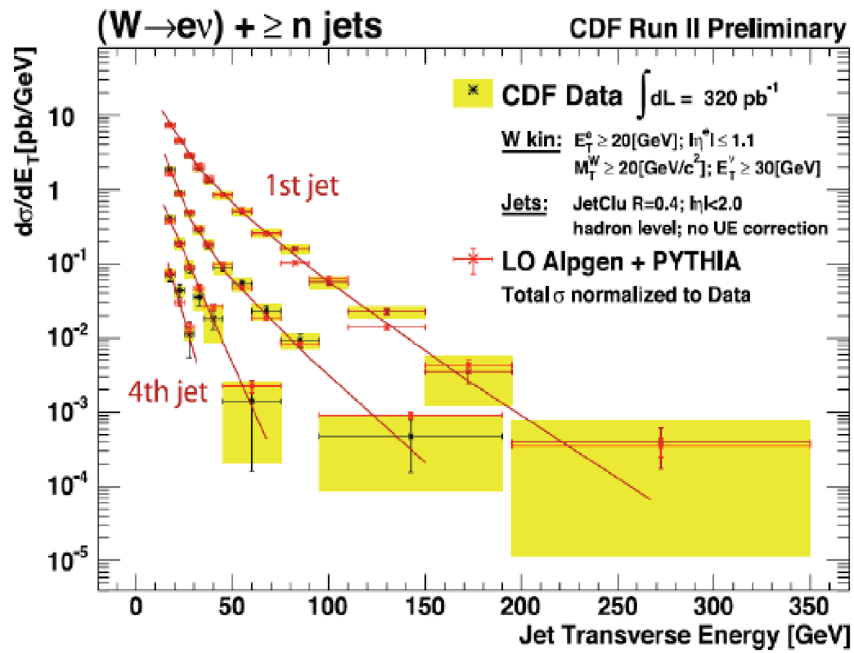


- Can build a ladder of samples including real emissions reliably with matrix elements, but only with real terms (ie the precision of the total cross-section is still LO)
- To keep in mind: PS(today) ≠ PS(yesterday). Tunings need to adapt to the choice of the matching

ME to PS matching at work

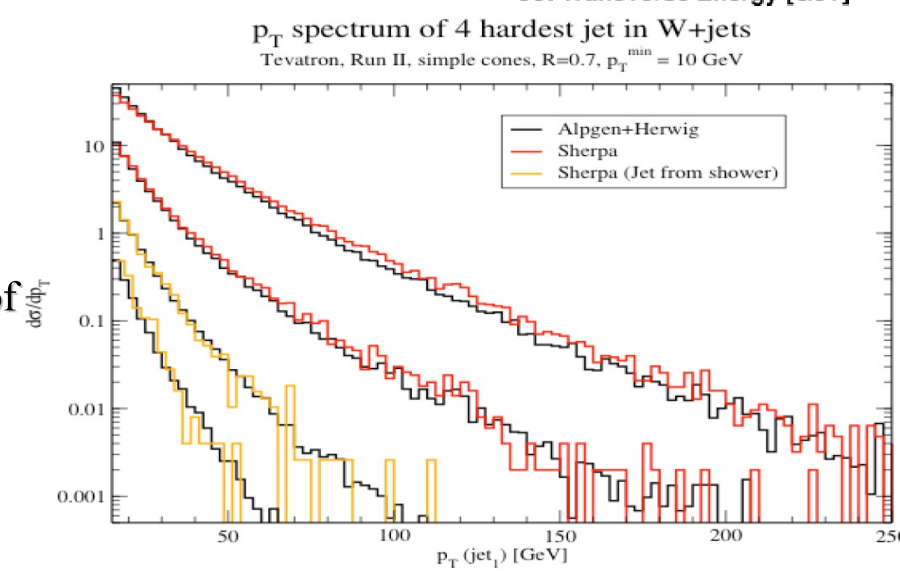
Matching ME and PS is not for fun. It is needed to describe data!



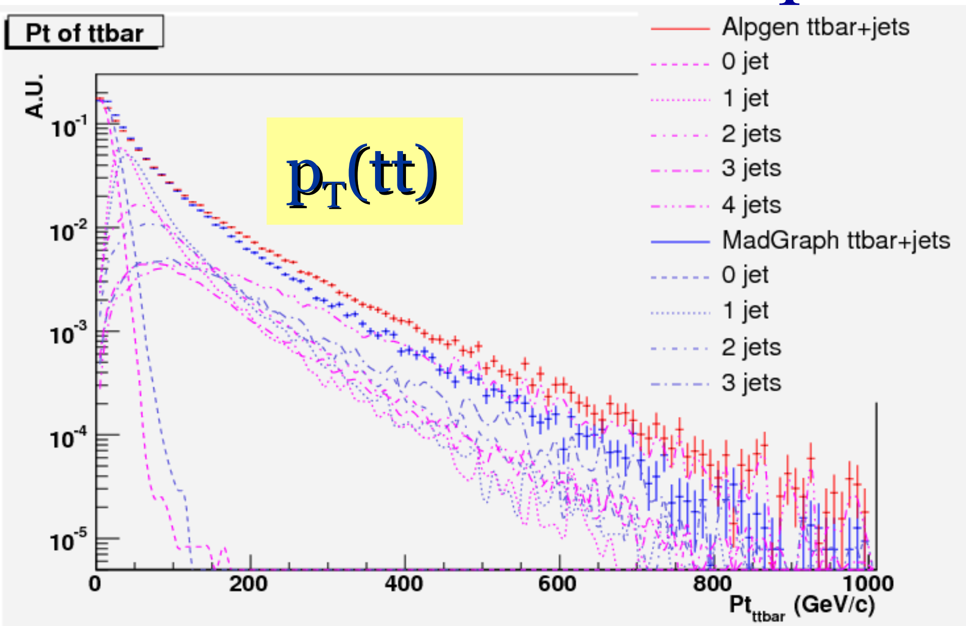


Different codes also give very similar predictions in terms of shapes

- comparison to data more reliable
- understanding of different implementations of the matching (for instance CKKW-MLM)
- assessment of systematic errors



Example: top pair production

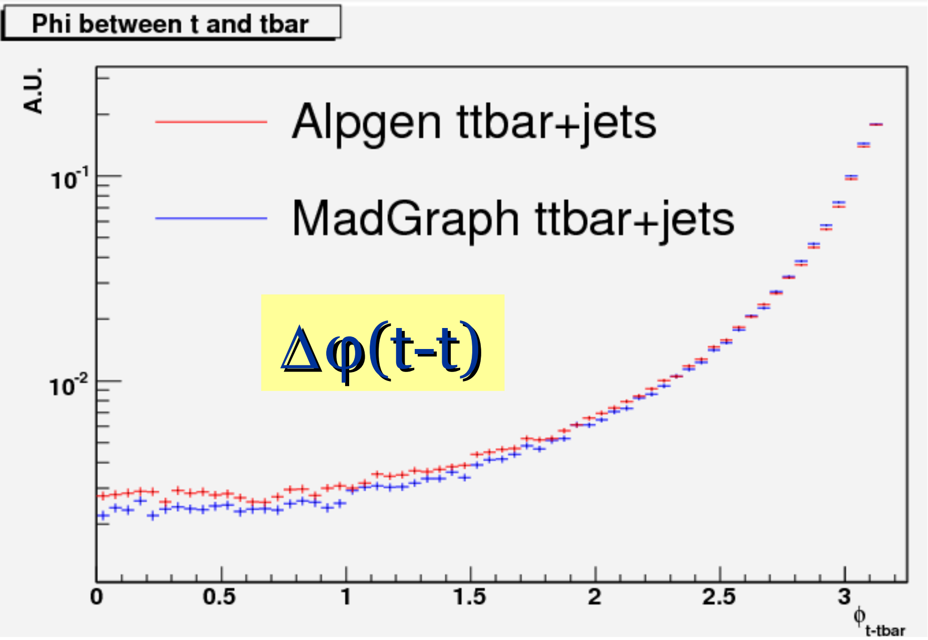


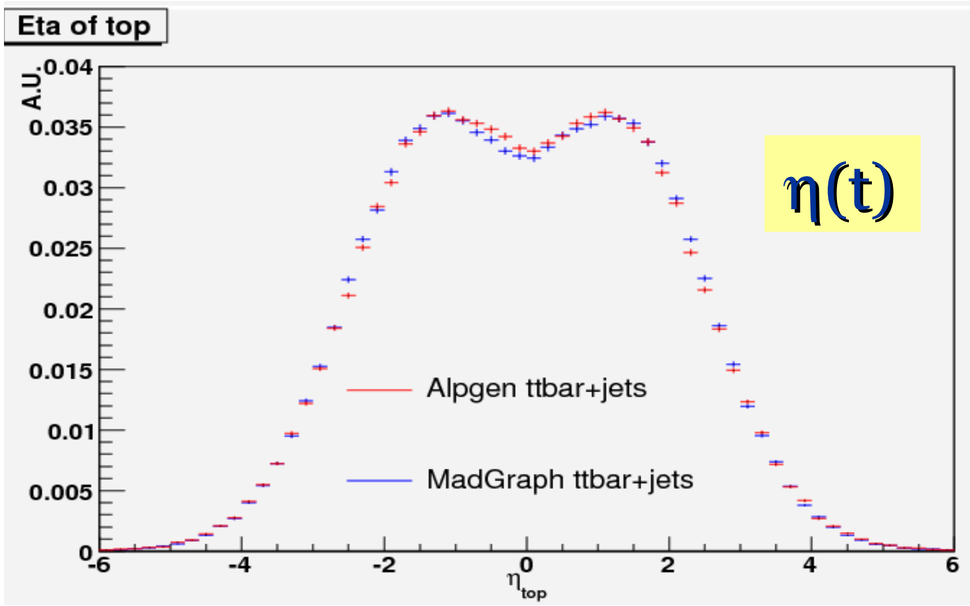
ALPGEN and MadGraph differ by at most 50% on the p_T prediction

Important to understand the residual theory error on the distributions:

- Effect of renormalisation and factorisation scales on the predictions
- Effect of the chosen ME-PS matching scale

Excellent agreement on other variables





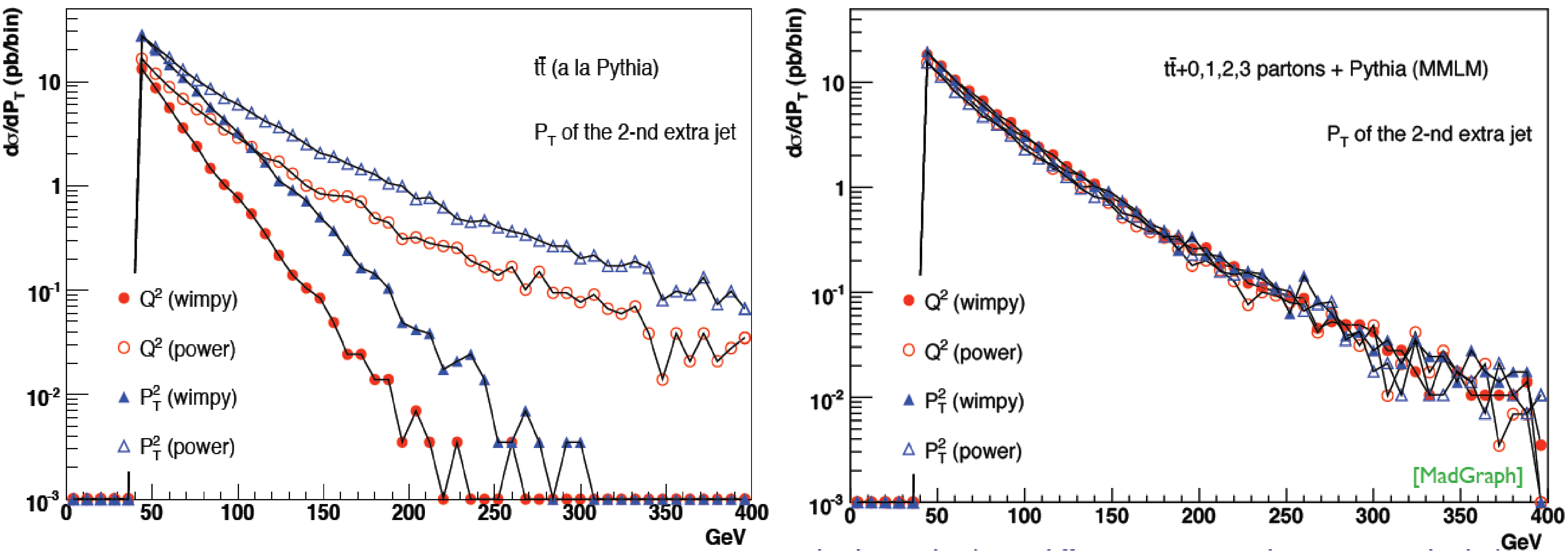
The importance of Matrix Elements tools

A parton shower is by construction an highly tunable tool.

For a matched calculation the effect of tunings in the hard regions are less relevant because this is described by the Matrix Element

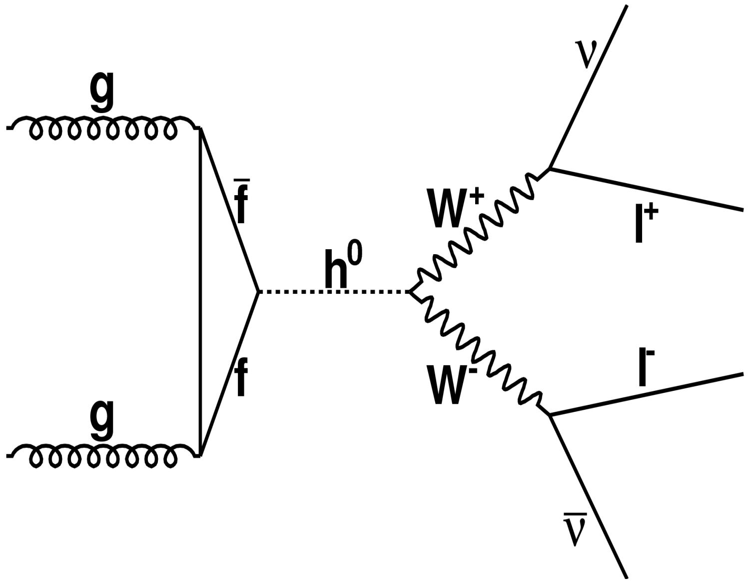
- more predictive power
- less sensitivity to the MC tunings
- systematic errors due to theory/modelling are smaller (should include theory uncertainties of the matching itself)

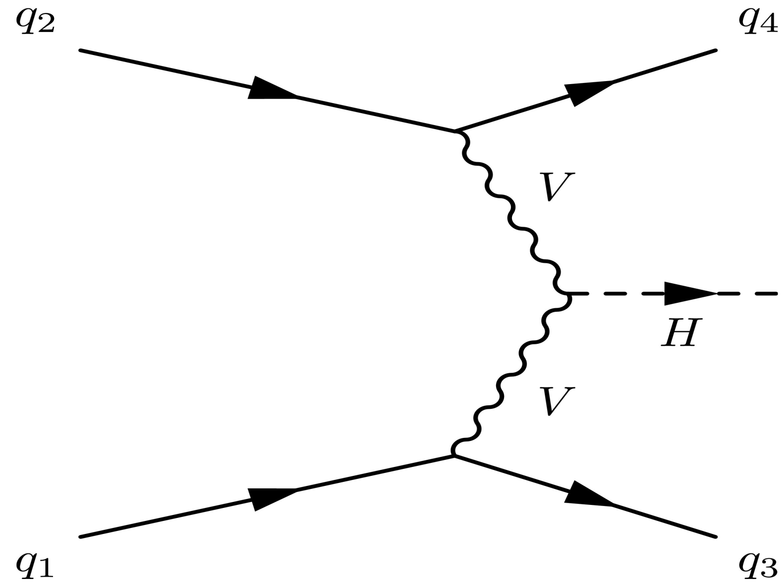
F. Maltoni, top 2008

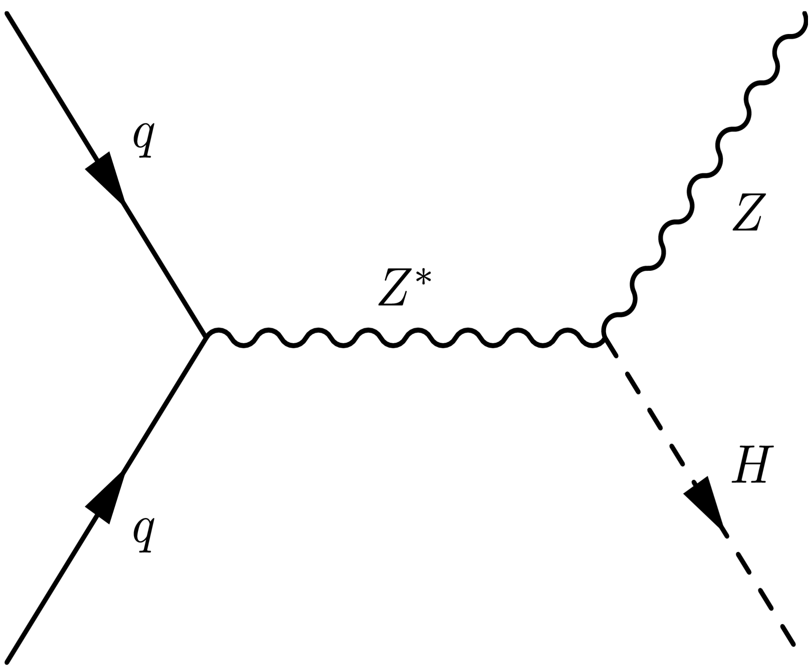


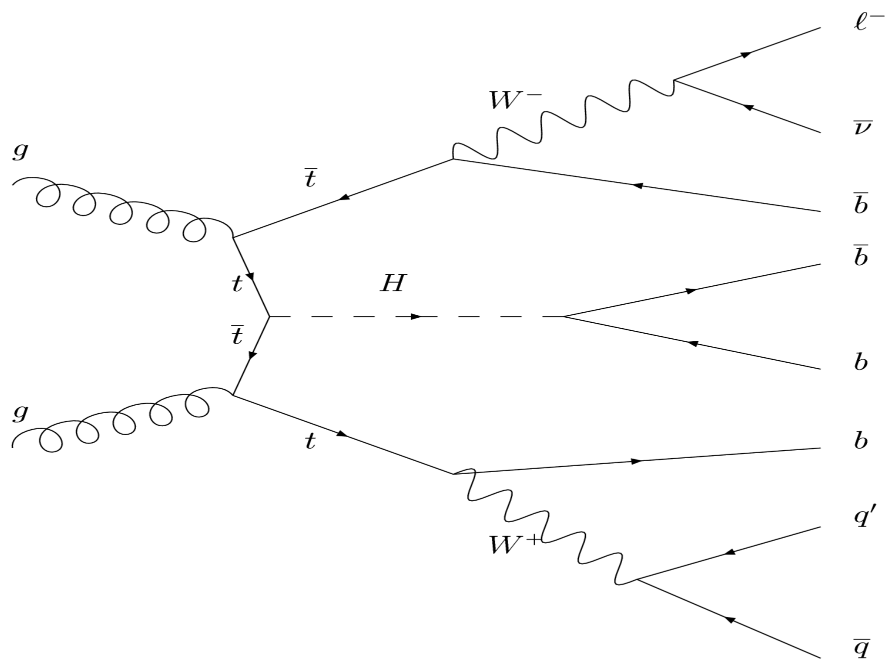
Higgs Production

Production mechanisms









MC Tools

- Fully inclusive : calculates total cross-sections, possibly with some non-realistic cuts

- Parton-level : calculates cross-sections and distributions for any observables

- Fully exclusive : full event simulation down to hadrons...

Presented Numbers

- In the following we present the ‘most up to date’ numbers for the most important SM Higgs production process cross-sections, together with their ‘theoretical uncertainty’
- The tool used to compute them are
  - Everything besides Gluon-Fusion: Dedicated tools by M. Spira
    - http://people.web.psi.ch/spira/proglist.html
  - HggTotal for the Gluon-Fusion, for more detail of the computation, see
    - C. Anastasiou, R. Bougezhal, F. Petriello, JHEP 0904:003, 2009
- The theoretical uncertainties are estimated by
  - Varying the renormalization and factorization scale around a process dependent reference scale μ in the range [μ/2, 2 μ]
  - By scanning over the error PDF sets
- Find the detailed numbers here
  - http://wwweth.cern.ch/HiggsCrossSections

gg→H @ NLO

- MCFM — Campbell & Ellis
- MC@NLO — Bryan & Stefano
- Herwig++ — Hamilton et al.
- POWHEG — Nason et al.

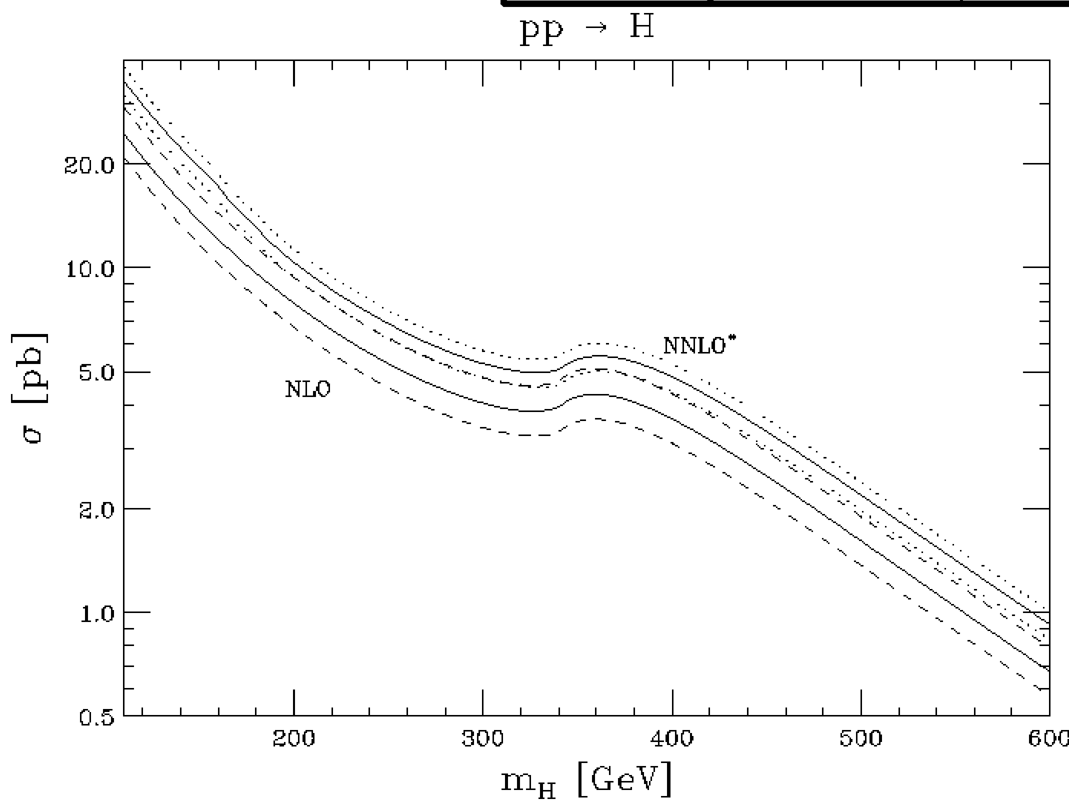
Comments:

Many comparisons exists. All showing that higher order effects are important

Gluon Fusion

Pert. order	NNLO
μ	m_H
PDF set	MSTW2008

m_H [GeV]	σ [pb]	δ (μ) [%]	δ (PDF) [%]
105	3.795×10^{1}	+10.0 -9.20	+1.79 -2.28
300	5.273	+9.16 -9.49	+2.26 -2.53
600	9.236×10^{-1}	+9.06 -9.84	+4.04 -5.17



Vector Boson Fusion

At NLO:

- MCFM Campbell et al.
- VBF@NLO Zeppenfeld et al.

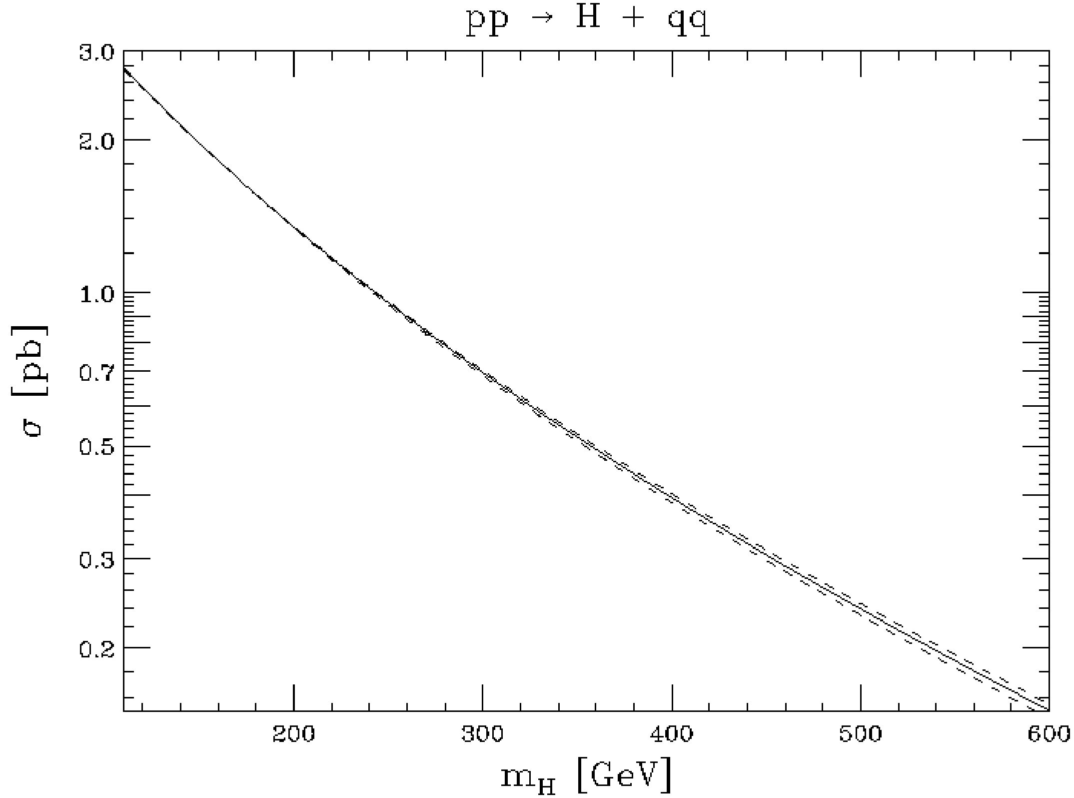
At LO:

- Many tools with matching

Vector Boson Fusion

Pert. order	NLO
μ	m_H
PDF set	CTEQ6M

m_H [GeV]	σ [pb]	δ (μ) [%]	δ (PDF) [%]
105	2.893	-0.71 +0.48	+2.04 -2.05
300	6.973×10^{-1}	+1.06 -1.42	+2.66 -2.59
600	1.507×10^{-1}	+3.07 -3.32	+3.03 -2.98



Heavy quarks + Higgs

Top

At LO everything public and available, even with matching.

At NLO two private codes, one from SPIRA et al. and one from Reina et al.

Cross sections available in FeynHiggs with the effective coupling approximations

Bottom

Much more complicated depending on the scheme

At NLO in the 4F two private codes, one from SPIRA et al. and one from Reina et al.

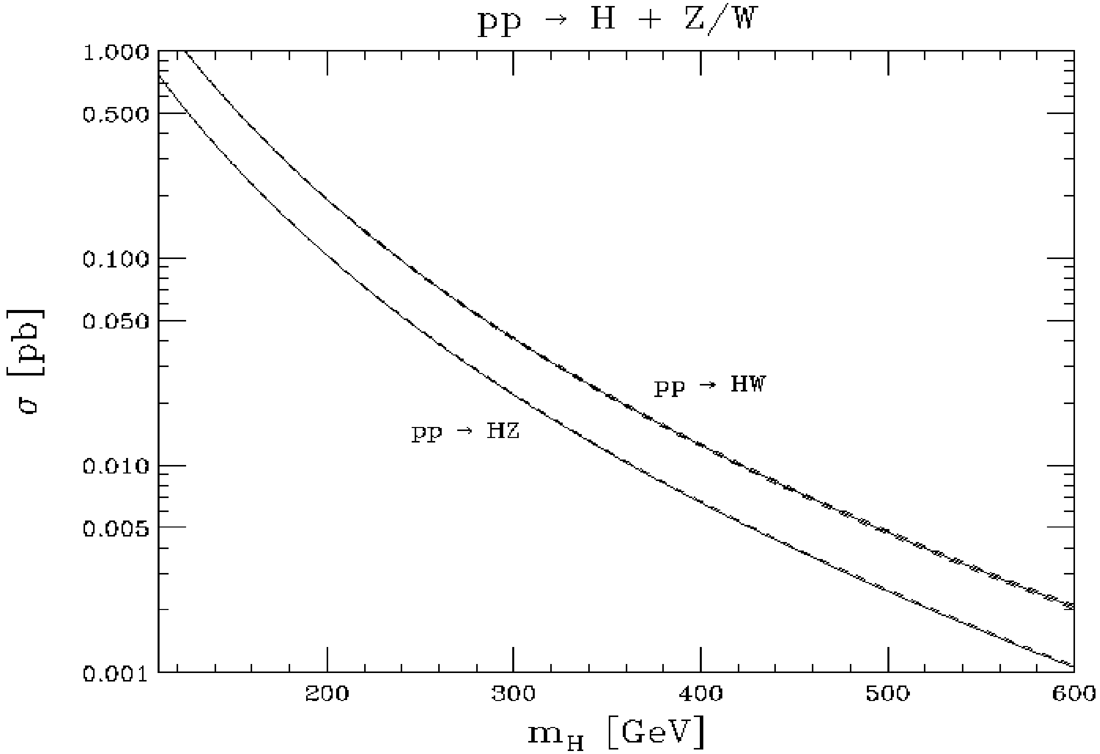
At NNLO one “public” code from Harlander bb@NNLO.

At LO a work still to be done to accomplish solid matching

Association with W/Z Boson

m_H [GeV]	σ (HZ) [pb]	δ (μ) [%]	δ (PDF) [%]	σ (HW) [pb]	δ (μ) [%]	δ (PDF) [%]
105	0.873	+0.37 -0.10	+2.02 -1.90	1.686	+0.37 -0.06	+2.03 -1.89
300	2.176×10^{-2}	+1.58 -1.32	+2.51 -2.51	4.072×10^{-2}	+1.60 -1.34	+2.45 -2.46
600	1.055×10^{-3}	+2.42 -2.40	+3.45 -3.43	2.056×10^{-3}	+2.43 -2.41	+3.43 -3.42

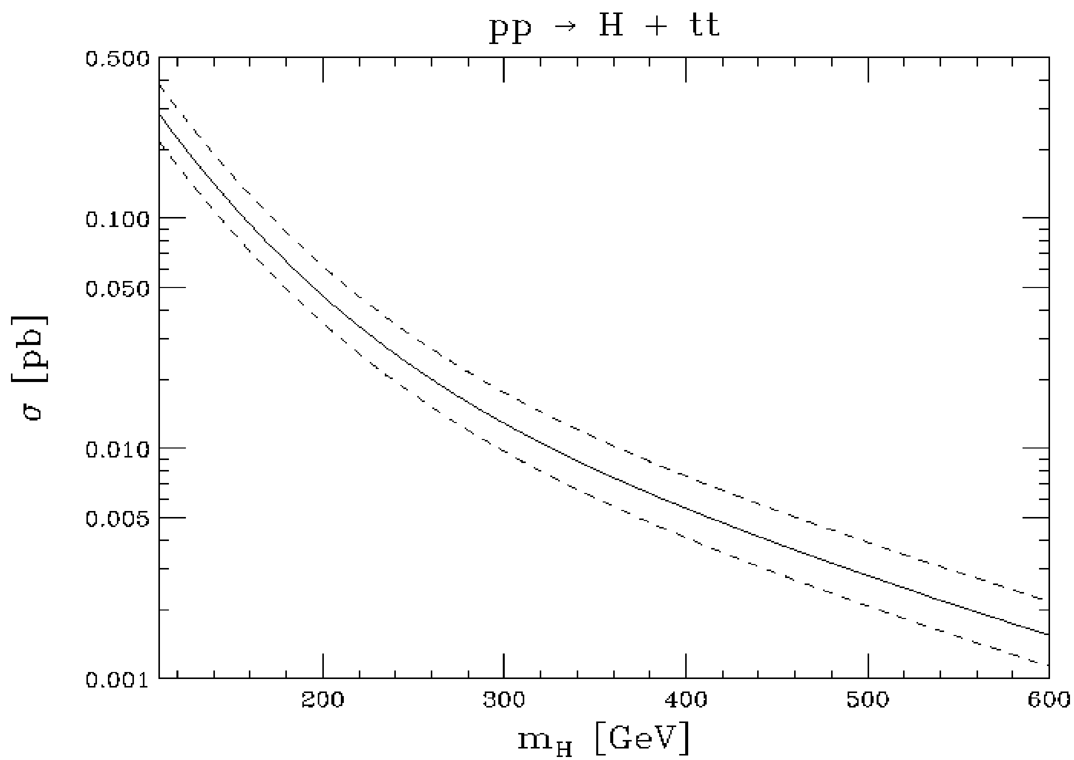
μ	$m_{Z/W}$
PDF set	CTEQ6M



Association with Top-Pair

Pert. order	LO
μ	$(m_H + 2\ m_{top})/2$
PDF set	CTEQ6L1

m_H [GeV]	σ [pb]	δ (μ) [%]
105	3.213×10^{-1}	+34.6 -23.8
300	1.284×10^{-2}	+36.2 -24.5
600	1.542×10^{-3}	+40.5 -26.7



Higgs Decay Modes
&
Branching Ratios

Branching ratios & Higgs mass

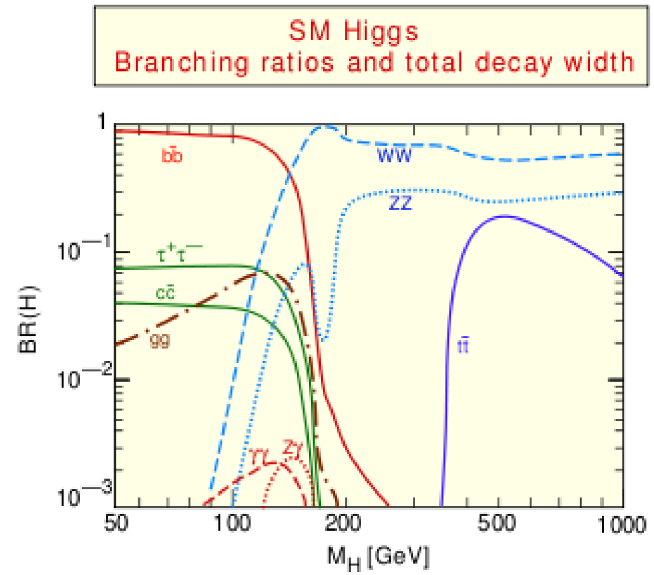
Partial widths at tree level:

- $\Gamma(H \to f\bar{f}) \propto N_c m_f^2 \beta^3 m_H$
- $\Gamma(H \to VV) \propto \delta_V \beta\, m_H^3 (1 - \tau_V + \frac{3}{4}\tau_V^2)$

$\beta^2 = 1 - 4m_f^2/m_H^2,\ \tau_V = 4m_V^2/m_H^2$

$N_c = 3$ for quarks and $N_c = 1$ for leptons

$\delta_W = 2,\ \delta_Z = 1$

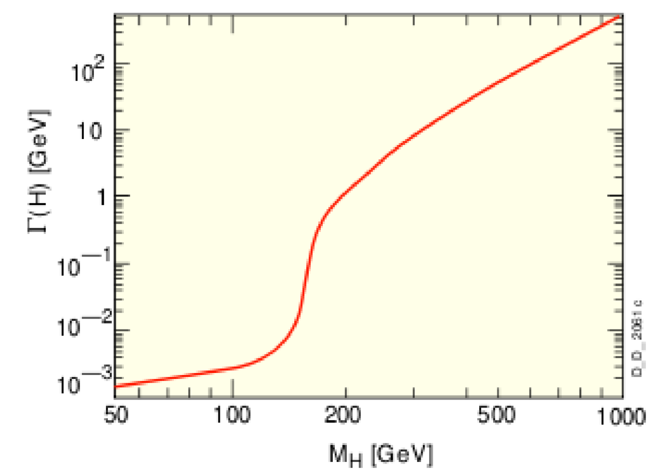


Low mass region:

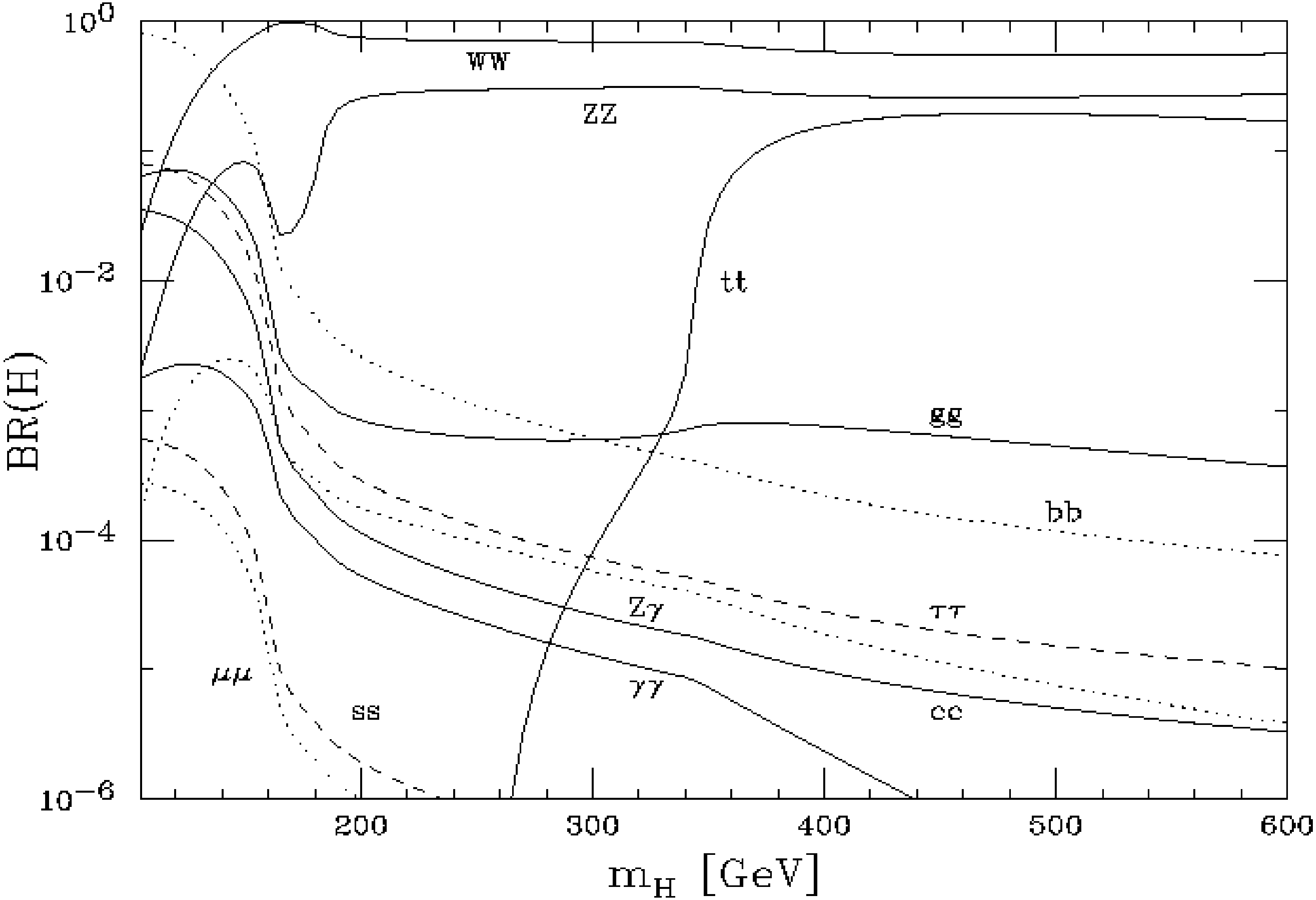
- $H \to b\bar{b}$ dominant

Higgs mass region:

- $\Gamma_{tot} \approx \Gamma(H \to WW) + \Gamma(H \to ZZ)$
- $m_H \approx 1.4\ \text{TeV/c}^2 \to \Gamma_{tot} \approx m_H$



Higgs decay modes



Values computed with HDECAY v.3.4

Shopping list for a coherent MC setup

- **Take care of the SM as correctly (and uniformly!) as possible**

→ Satisfactory ME description with needed corrections. Reweighting is a possibility…
→ Coherent interface to showering, fragmentation, decay
→ Uniform choice for input parameter settings
→ Same cuts/description for signals and backgrounds

- **Diversity wins: add generator redundancy in crucial portions of phase space**

→ One prediction is often not enough
→ NLO, different generators…
→ Different interface to showerings – prepare to tune with data
→ Different settings to study systematics (tunings, scales…)
→ Sensitivity of analyses and reconstruction methods to "theory/modeling" effects

- **Add new physics samples**

→ Main SUSY and BSM points to train analyses

- **Determine tails**

→ Study what tails are most interesting and refine studies there

⇒ **The use of an as much coherent (IPS, PDF, cuts) as possible set-up will ease enormously the tasks of the analyses**

- compare SM and BSM on a similar (equal) footing. Include SM-SM and even SM-BSM interference
- disentangle detector/simulation effects from the physics input to the generation
- speed up things

A common effort

Never underestimate sociological aspects in a large HEP community

Based on the experience at LEP/Tevatron (and actually also at the LHC already), the only winning choice is to build a community working together

1. Theorists and experimentalists
   - we (have) learn(ed) a lot from each other
   - support (online!) on generators setup, tunings, definitions
   - significantly reduce the probability of a misinterpretation of data
   - possibility for theorist to have access to data before an experimental paper comes out

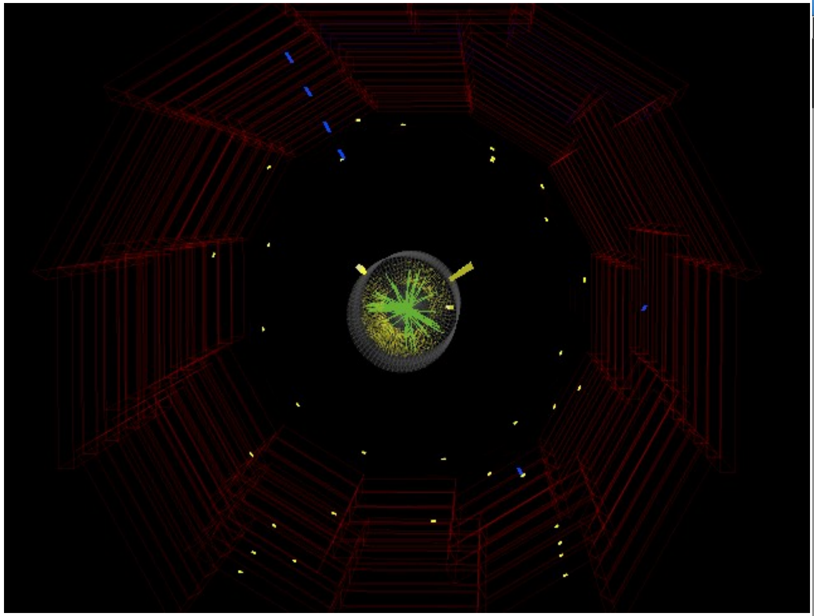
2. Experimentalists and experimentalists
   - different experiments or different analysis should try and work together
   - same process definition (signal for someone, background for some other)
   - coherent setup generation across the phase space
   - combinations (where applicable) need standardizations

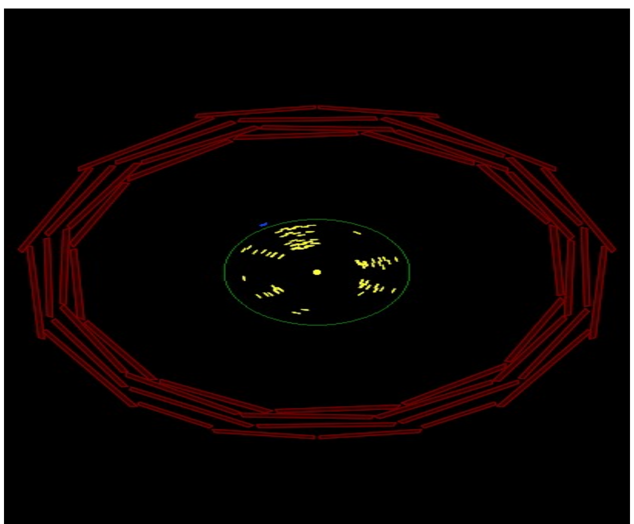
3. Learn from other experiments (i.e. tools, tunings)

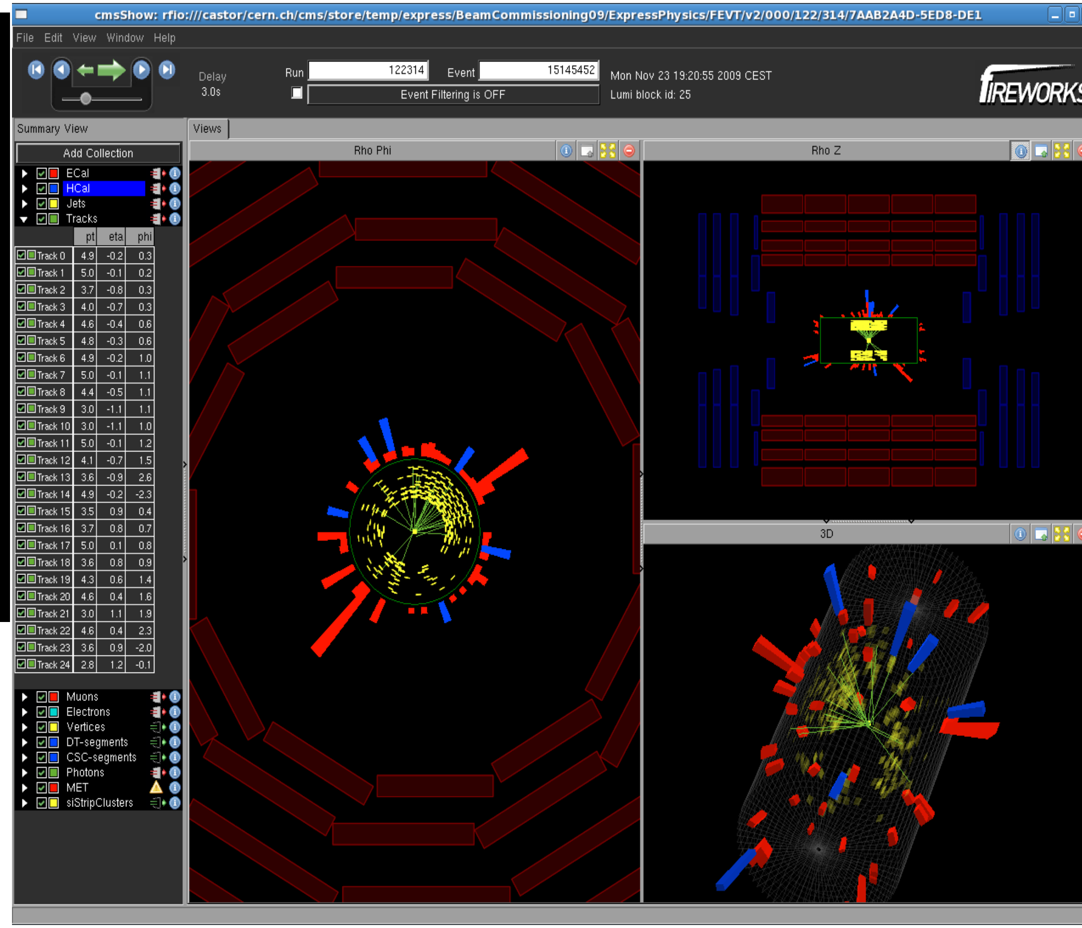
Conclusions

- A lot of tools available
- The highest level of theoretical understanding for cross-sections and observables is actually for the Higgs
- Still room for improvement within the current technical abilities

First collisions!



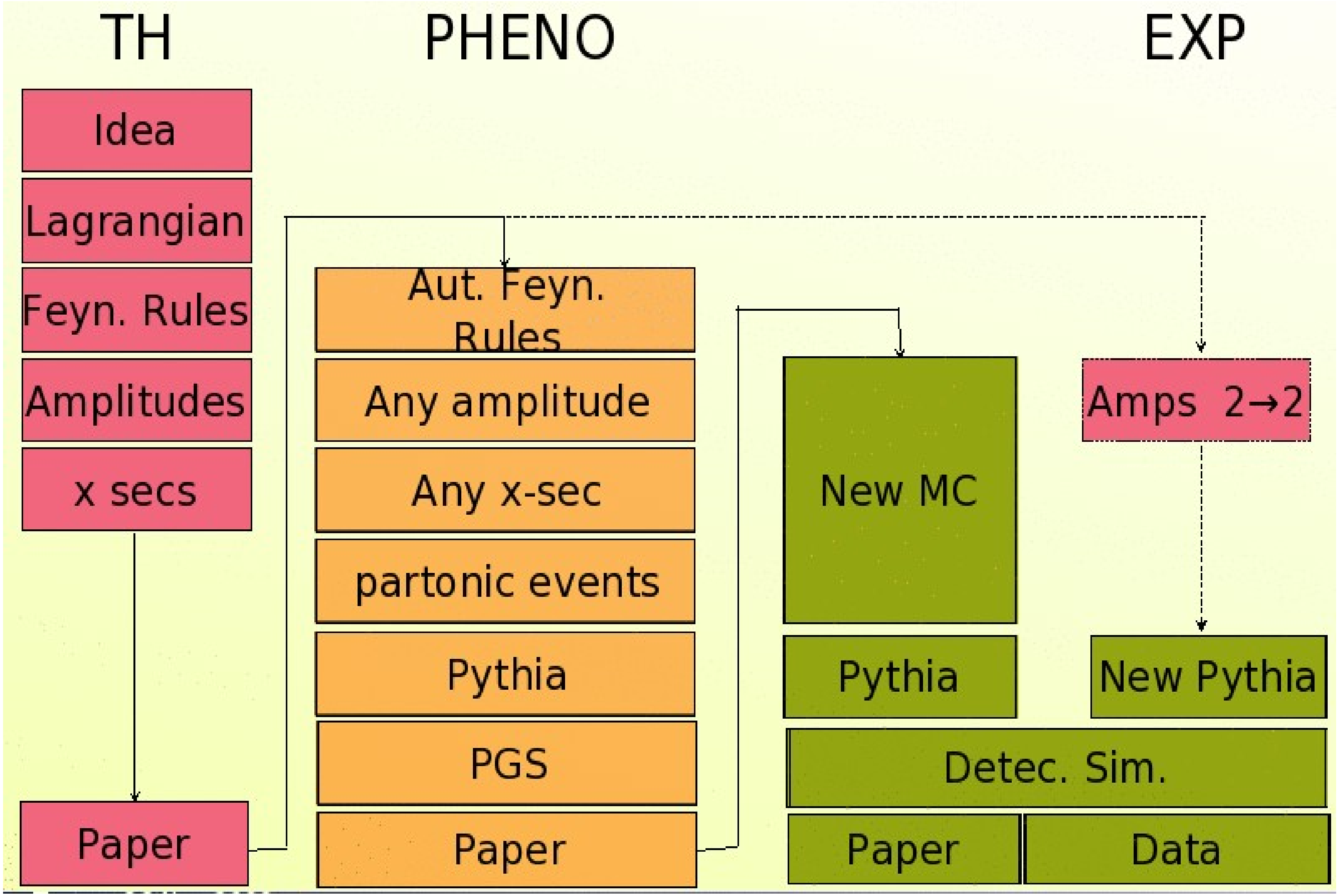




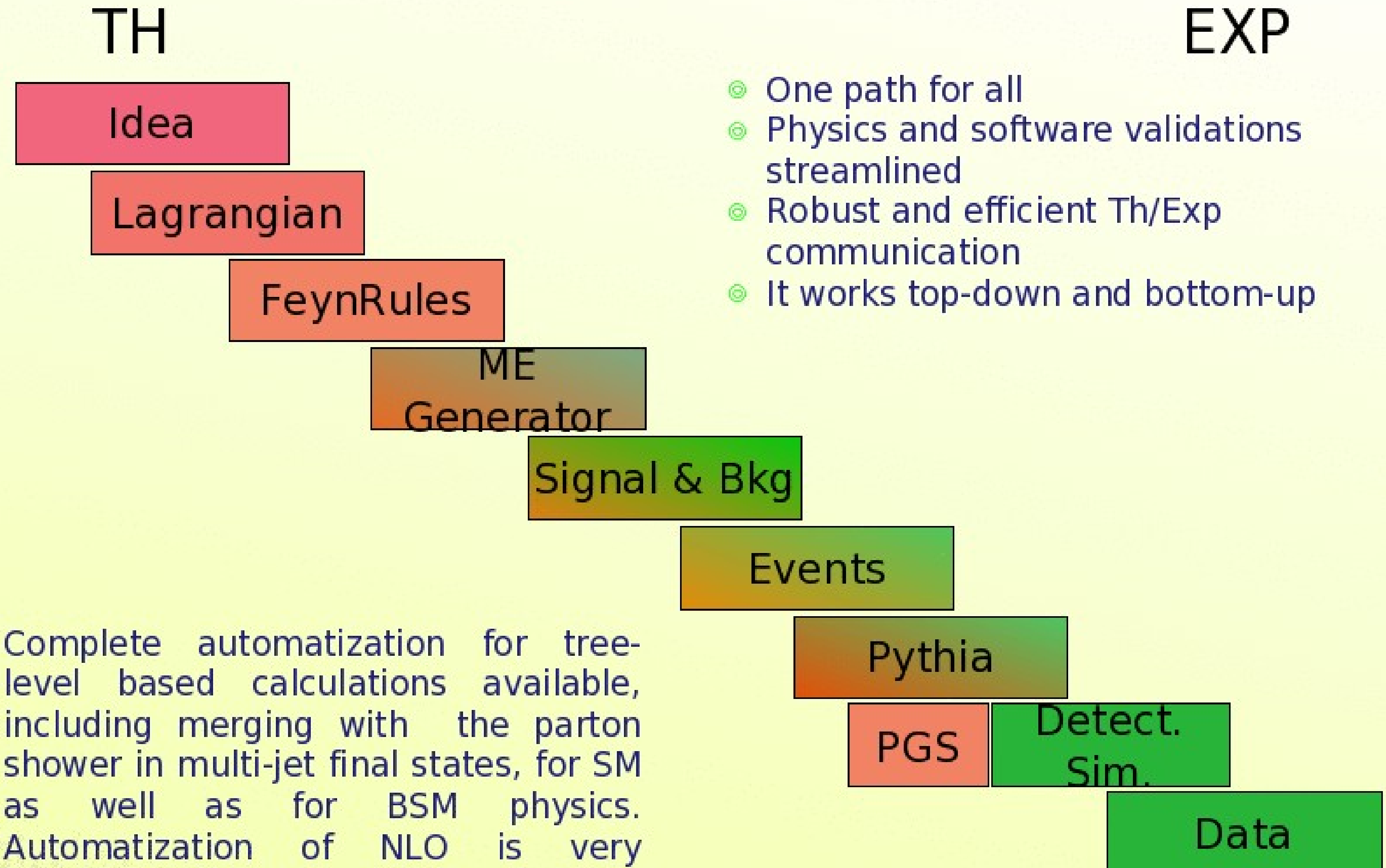
Let's start comparing collision data and Monte Carlo!

Backup slides

A Roadmap (with roadblocks) for BSM @ the LHC



A Roadmap for BSM @ the LHC

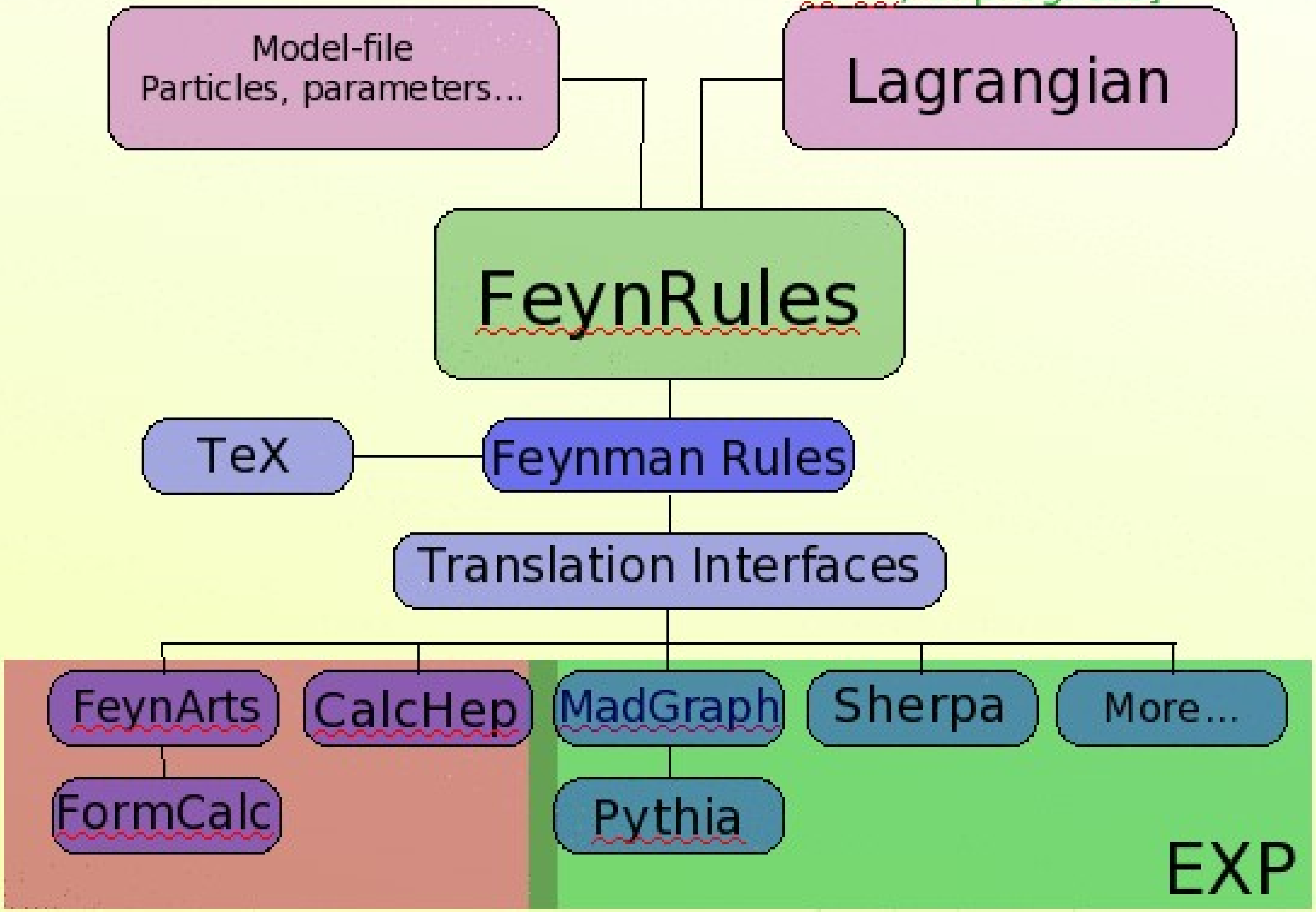


- One path for all
- Physics and software validations streamlined
- Robust and efficient Th/Exp communication
- It works top-down and bottom-up

Complete automatization for tree-level based calculations available, including merging with the parton shower in multi-jet final states, for SM as well as for BSM physics. Automatization of NLO is very

The FeynRules Project

[Christensen, Duhr, 2008; Christensen, Duhr, Fuks, et al., in progress]



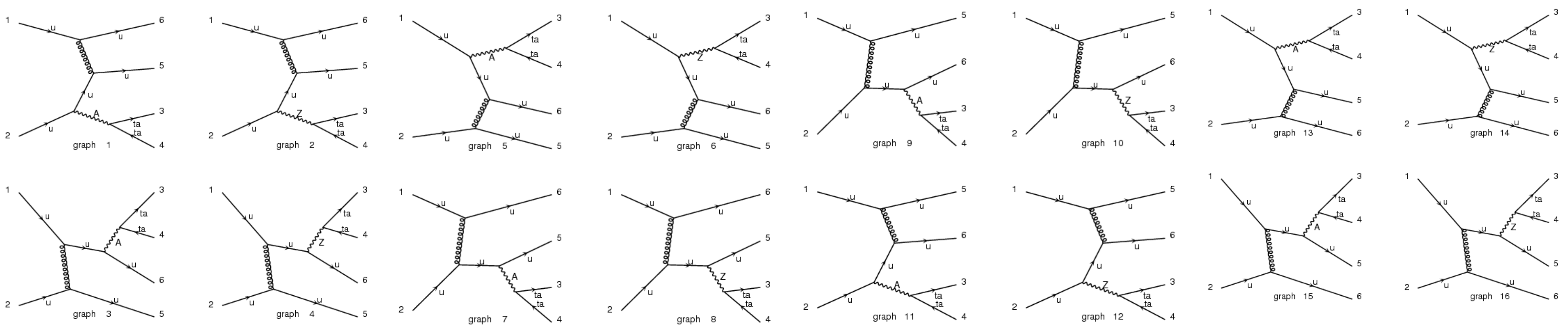
A simple example @LHC: $pp\rightarrow \ell^+\ell^-+2$ jets

Step 0: define the couplings (at leading order) to which you want to perform the calculation:

- QED=2; QCD=2

Step 1: determine all diagrams and subprocesses contributing to the Matrix Elements.

1. Average over initial states: $pp=\Sigma q_i q_j+\Sigma q_i g+gg$ (x 108)
2. Sum over final states: $\ell=e,\mu,\tau$; “jet”=q, g (x 15)
3. For each configuration determine all Feynman diagrams contributing to the subprocess amplitude. Consider for instance the simple case $uu\rightarrow uu\tau\tau$: 16 different diagrams

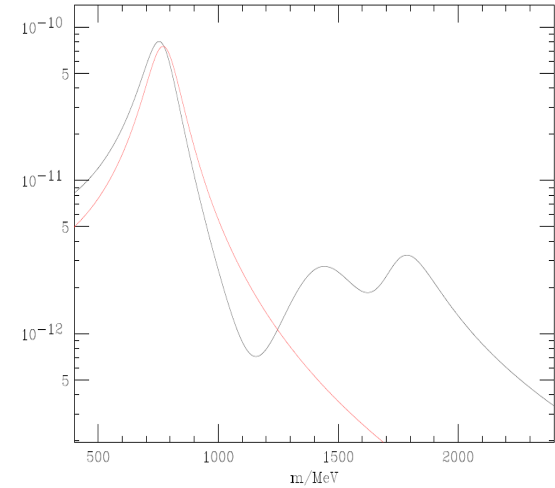


→ Order of thousands Feynman diagrams only for this channel with 4 partons in the final state !

A simple example @LHC: pp→$\ell^+\ell^-$+2 jets

Step 2: calculate the Matrix Elements |M| corresponding to all sub-processes contributing to the observable final state

- This must be done in an automated way, either by using libraries or automated calculations
- The functional form of the final ME will have be very complex, with resonances from internal propagators (for instance the Z in our case)

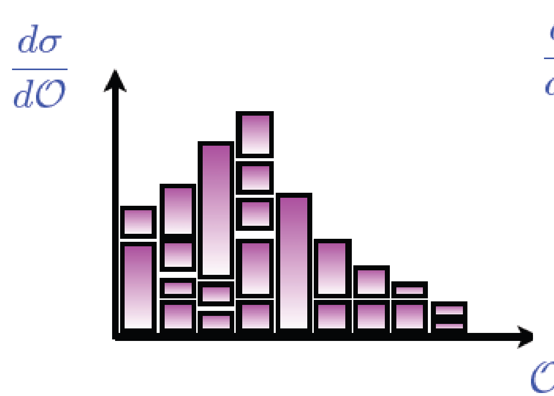


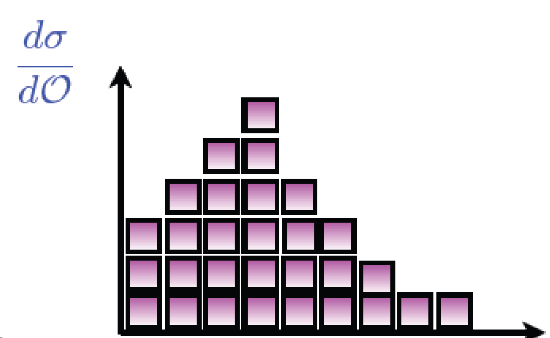
Step 3: integrate the Matrix Elements over the corresponding 4 particles phase space

- Adaptative Monte Carlo integration give the best performance
- In output one can have at the same time weighted "events", each "event" being a point in phase space

Step 4: produce unweighted events

- By using an accept/reject method





Monte Carlos

Next you need to integrate the Matrix Elements over the appropriate phase space:

$$Dim[\Phi(n)] \sim 3n$$

$$\sigma = \frac{1}{2s} \int |\mathcal{M}|^2 d\Phi(n)$$

$$d\Phi_n = \left[\Pi_{i=1}^{n} \frac{d^3 p_i}{(2\pi)^3 (2E_i)} \right] (2\pi)^4 \delta^{(4)}(p_0 - \sum_{i=1}^{n} p_i)$$

This is a 3n dimensional integral (if we know how to write the Matrix Elements) that can be tremendously complex

- it depends on the total number of particles in the final state
- forget pen and paper (ie in general no analytical formula)
- use numerical methods to solve it

The best way to make the integral is to use a Monte Carlo integration (programs like VEGAS)

- fast convergence in many dimensions
- arbitrarily complex integration regions (finite discontinuities not a problem: can include cuts)
- easy error estimations
- can generate events while sampling the phase space! (hence the name Monte Carlo generators)

Monte Carlo integration

The Monte Carlo integration methods locally approximate your function. They work because of the central limit theorem

$$I = \int_{x_1}^{x_2} f(x)\,dx = (x_2 - x_1)\,\langle f(x) \rangle$$

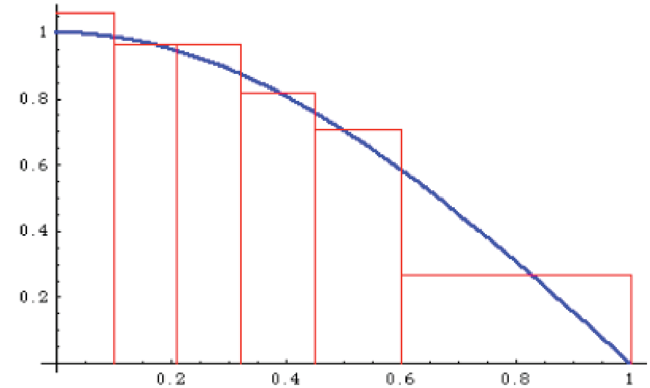
$$I \approx I_N \equiv (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x_i)$$

- error is improved by reducing V_N. Best case for f(x)=c, for which V_N=0.

$$I \approx I_N \pm \sqrt{V_N/N} \qquad V = (x_2 - x_1)\int_{x_1}^{x_2}\left[f(x)\right]^2 dx - \left[\int_{x_1}^{x_2} f(x)\,dx\right]^2$$

Adaptative integration:

1. Sample your function f(x), and determine your integral as an approximated sequence of step functions
2. Try and learn while sampling, throw more points (points=events!) where f(x) is large, i.e. the variance is large
3. Adjust the bin size in such a way that each square has the same area
4. Iterate the procedure



$$p(x) = \frac{1}{N_b \Delta x_i}, \quad x_i - \Delta x_i < x < x_i$$

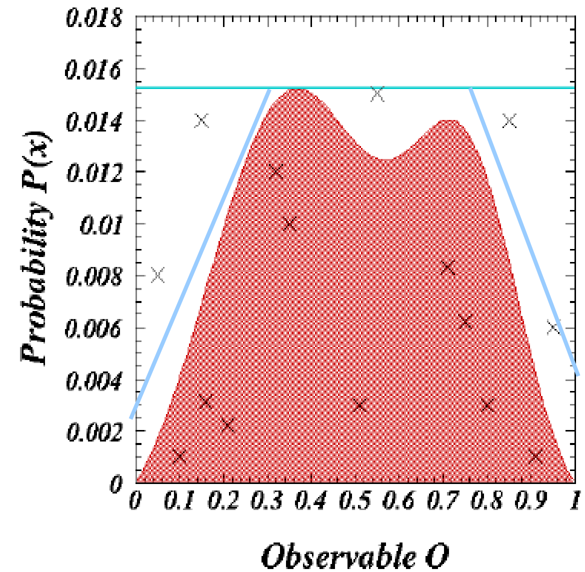
The adaptative integration corresponds to a Jacobian transformation for flattening the phase space

The extension to N dimensions is trivial, the problem is the multidimensional phase space sampling, where complex resonant structures may and do appear (from propagators in the MEs, typically)

Generating events

Most used and effective way to generate evens is via the hit and miss technique.

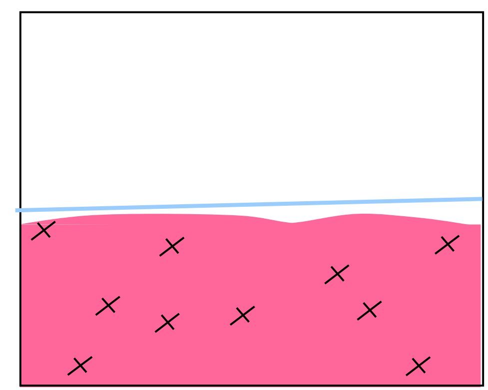
- Throw an x in a range $x=x_{min}+R(x_{max}-x_{min})$
- Find $f(x_{max})$ or an upper value for f(x) in the x range above
- Throw an y in the range 0, $f(x_{max})$
- If $y<f(x)$, keep x, otherwise reject it and start again



This method implies an inherent inefficiency, entirely dependent on the phase space structure.

Simple ways exist to make this method more efficient

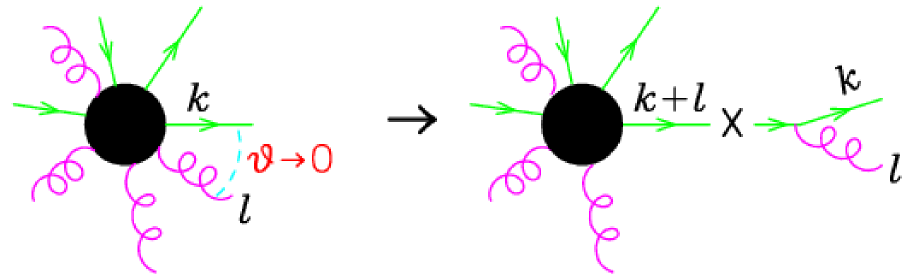
1. Be smart in maximizing f(x)
   - use linear or step functions S such that S>f everywhere
   - use linear sum of known functions always >f, then select the accept-reject condition according to their relative importance in x
2. Find a jacobian transformation J(O) of the variable O for which f(J(O)) is ~flattened up
   - what happens in an adaptative sampling
   - mandatory when phase space contains important peaks



f(J(O))

Parton Showers: basics

Cross sections factorize near collinear limit

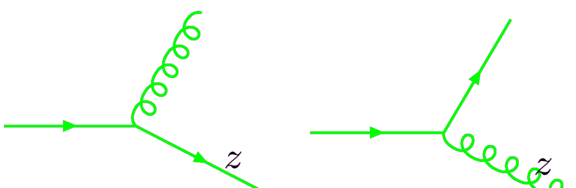


$$d\mathbf{\Phi}_{n+1} = d\mathbf{\Phi}_n \, d\Phi_r \qquad d\Phi_r \div dt\, dz\, d\varphi$$

$$|M_{n+1}|^2 d\mathbf{\Phi}_{n+1} \Longrightarrow |M_n|^2 d\mathbf{\Phi}_n \frac{\alpha_s}{2\pi} \frac{dt}{t} P_{q,qg}(z)\, dz \frac{d\varphi}{2\pi}$$

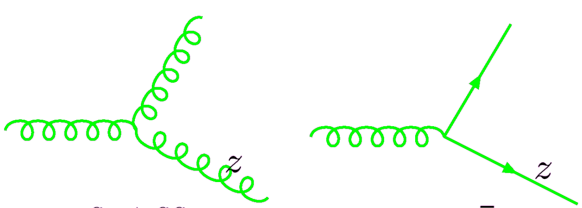
$$\left\{ \begin{array}{ll} \dfrac{dt}{t} \approx \dfrac{d\theta}{\theta} & \text{collinear singularity} \\ \dfrac{dz}{1-z} \approx \dfrac{dE_g}{E_g} & \text{soft singularity} \end{array} \right.$$

Universal DGLAP splitting functions



$q \to qg$ $\quad C_F \frac{1+z^2}{1-z}$

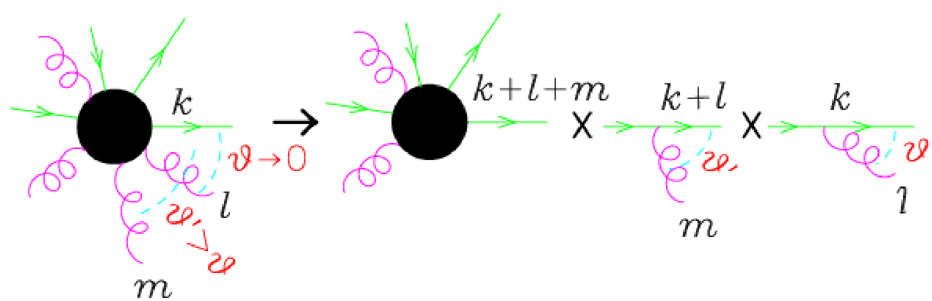
$q \to gq$ $\quad C_F \frac{1+(1-z)^2}{z}$



$g \to gg$ $\quad C_A \frac{z^4+1+(1-z)^4}{z(1-z)}$

$g \to q\bar{q}$ $\quad T_R \left(z^2 + (1-z)^2\right)$

If you have a second collinear emission one can iterate the formula



$\theta' . \theta \longrightarrow 0$ with $\theta' > \theta$

$$\sigma \approx \sigma_0 \, \alpha_s^n \int_{t_0}^{Q^2} \frac{dt_1}{t_1} \int_{t_0}^{t_1} \frac{dt_2}{t_2} \ldots \int_{t_0}^{t_{n-1}} \frac{dt_n}{t_n} \approx \sigma_0 \, \alpha_s^n \frac{1}{n!} \left(\log \frac{Q^2}{t_0}\right)^n$$

$t_0 \approx \Lambda_{QCD}$ is an infrared cutoff : leading-log approximation

Parton Showers: basics

How many time should we branch a line, however?

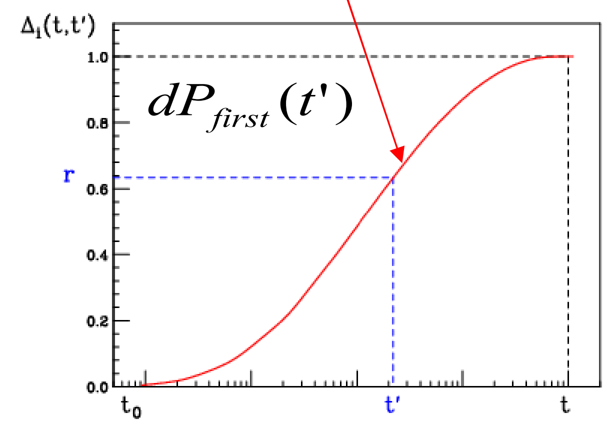
$$dP_{first}(t') = P_{noemis}(t,t')\, dP_{emis}(t'+dt,t')$$

$$dP_{\text{emis}}(t+dt,t) = \frac{dt}{t}\frac{\alpha_s(t)}{2\pi}\int dz\, P_{i,jk}(z)$$

$$dP_{\text{no emis}}(t+dt,t) = 1 - dP_{\text{emis}}(t+dt,t) = 1 - \frac{dt}{t}\frac{\alpha_s(t)}{2\pi}\int dz\, P_{i,jk}(z)$$

$$P_{\text{no emis}}(t_1,t_2) = \lim_{N\to\infty}\prod_{n=1}^{N}\left[1 - \frac{dt}{t_n}\frac{\alpha_s(t_n)}{2\pi}\int dz\, P_{i,jk}(z)\right] = \exp\left\{-\int_{t_2}^{t_1}\frac{dt}{t}\frac{\alpha_s(t)}{2\pi}\int dz\, P_{i,jk}(z)\right\}$$

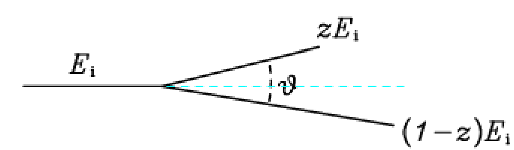
Sudakov form factor



The shower algorithm is then implemented in an iterative way:

- Find your evolution variable t (different choices exist) and your initial scale Q^2
- Branch it the first (and only!) time according to dP_{first}

  1. generate r → determine t'
  2. if $t'<t_0$, stop here (the parton has not radiated)
  3. if $t'>t_0$ generate z, jk with probability $P_{i,jk}(z)$, assigning energies $E_j=zE_i$ and $E_k=(1-z)E_i$ to partons j and k
  4. restart shower from partons j and k, setting the ordering parameter accordingly t=t'
  5. repeat until no further parton can radiate

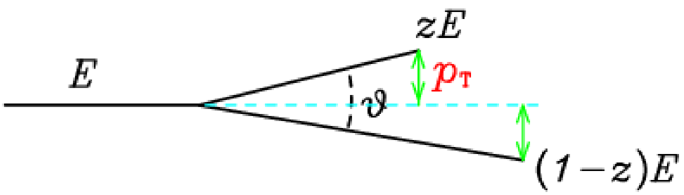


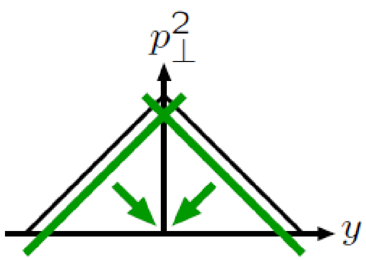
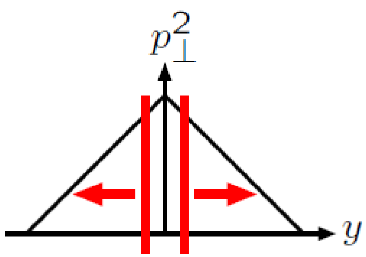
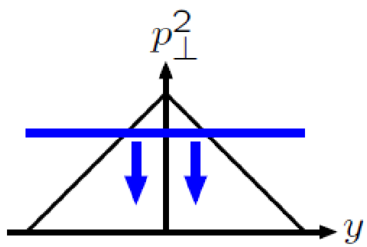
- Q^2 is an upper cutoff for the ordering variable t
- $t_0 \approx \Lambda_{QCD}$ is an infrared cutoff (corresponding to the confinement scale)

Parton Showers: basics

Several possible choices of the ordering variable t:

$$
\begin{array}{llll}
\text{virtuality:} & t & \equiv & E^2 z(1-z)\overbrace{\theta^2}^{2(1-\cos\theta)} \\
p_T^2: & t & \equiv & E^2 z^2 (1-z)^2 \theta^2 \\
\text{angle:} & t & \equiv & E^2 \theta^2
\end{array}
$$

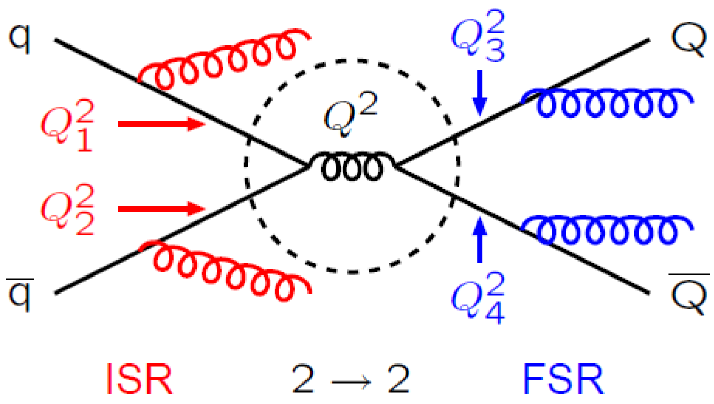


PYTHIA: $Q^2 = m^2$	HERWIG: $Q^2 \sim E^2\theta^2$	ARIADNE: $Q^2 = p_\perp^2$
		
large mass first $\Rightarrow$ "hardness" ordered **coherence brute force** covers phase space ME merging simple	large angle first $\Rightarrow$ **hardness not ordered** coherence inherent **gaps in coverage** **ME merging messy**	large $p_\perp$ first $\Rightarrow$ "hardness" ordered coherence inherent covers phase space ME merging simple

Parton Showers: basics

An illustrated example using the parton shower as implemented in PYTHIA:

$$2 \to n \;=\; (2 \to 2) \oplus \text{ISR} \oplus \text{FSR}$$



FSR = Final-State Rad.;
timelike shower
$Q_i^2 \sim m^2 > 0$ decreasing

ISR = Initial-State Rad.;
spacelike shower
$Q_i^2 \sim -m^2 > 0$ increasing

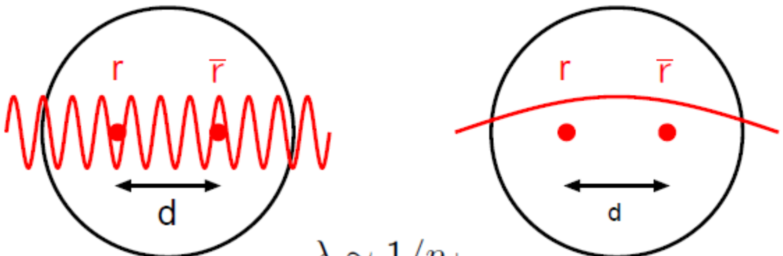
The initial shower is evolved back in 'time' to match the PDF at a scale Q^2 for which a choice of the parton initiator "K" according to $PDF_K(x,Q^2)$ becomes possible (= is picked up)

The final shower evolution is cutoff at some scale Q_0 (factorization scale, tunable parameter, typically of the order of confinement scale ~ 1GeV), after which the event will be left as a set of quarks, leptons and gauge bosons (gluons, γs)
The underlying event will have to be added, then fragmentation takes over

The underlying event: MPI

The 2→2 QCD cross-section is dominated by the t-channel gluon exchange diagram, divergent for low transverse momenta: $d\sigma/dp_\perp^2 \approx 1/p_\perp^4$ for $p_\perp \to 0$

This is unphysical, since perturbative QCD is not valid down to scales comparable to the hadron confinement scale → introduce an empirical cut-off (to be tuned) in the cross-section expression



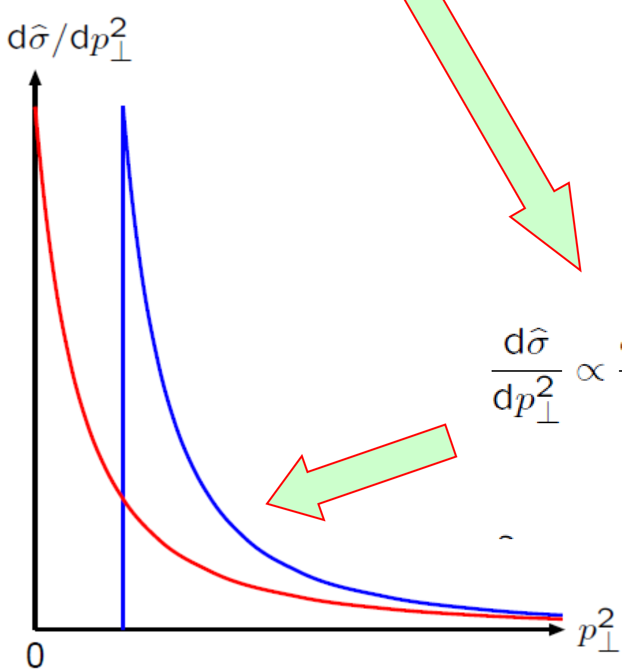
resolved $\lambda \sim 1/p_\perp$ screened

How many interactions will then happen, on average?

- if they were independent they would follow a poissonian statistics
- momentum correlation (E, p conservation) suppresses large numbers of interactions per collision

Rule of thumb for Double Parton Interaction with process A and B:

- σ(A+B) = σ(A)× σ(B)/σ(eff)
- σ(eff) empirically determined: ~10-15 mb at the LHC



$$\frac{d\hat{\sigma}}{dp_\perp^2} \propto \frac{\alpha_s^2(p_\perp^2)}{p_\perp^4} \to \frac{\alpha_s^2(p_\perp^2)}{p_\perp^4}\theta\left(p_\perp - p_{\perp\min}\right) \quad \text{(simpler)}$$

$$\text{or} \to \frac{\alpha_s^2(p_{\perp 0}^2 + p_\perp^2)}{(p_{\perp 0}^2 + p_\perp^2)^2} \quad \text{(more physical)}$$

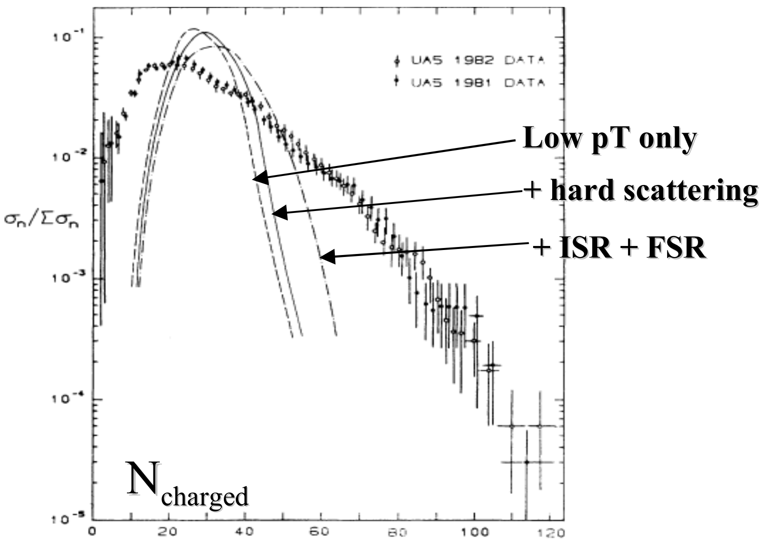
Main parameter in the Monte Carlo tunings !

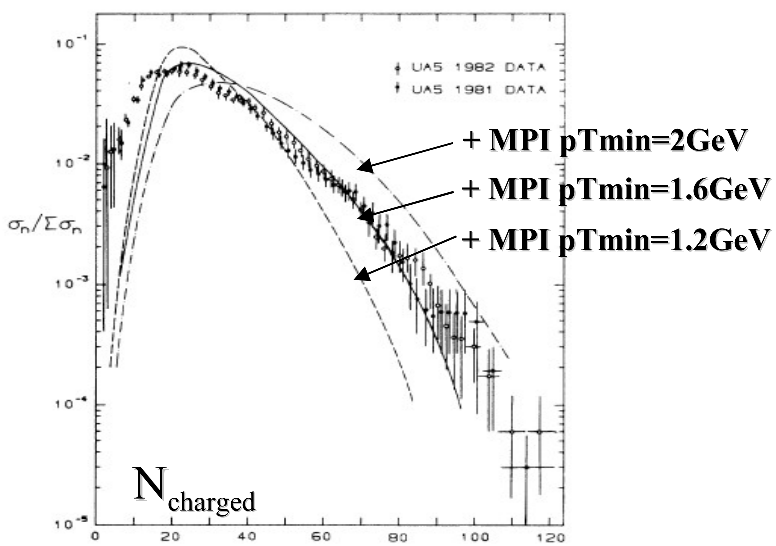
The underlying event: MPI evidence

PDF PS **UE** Hadron Decay

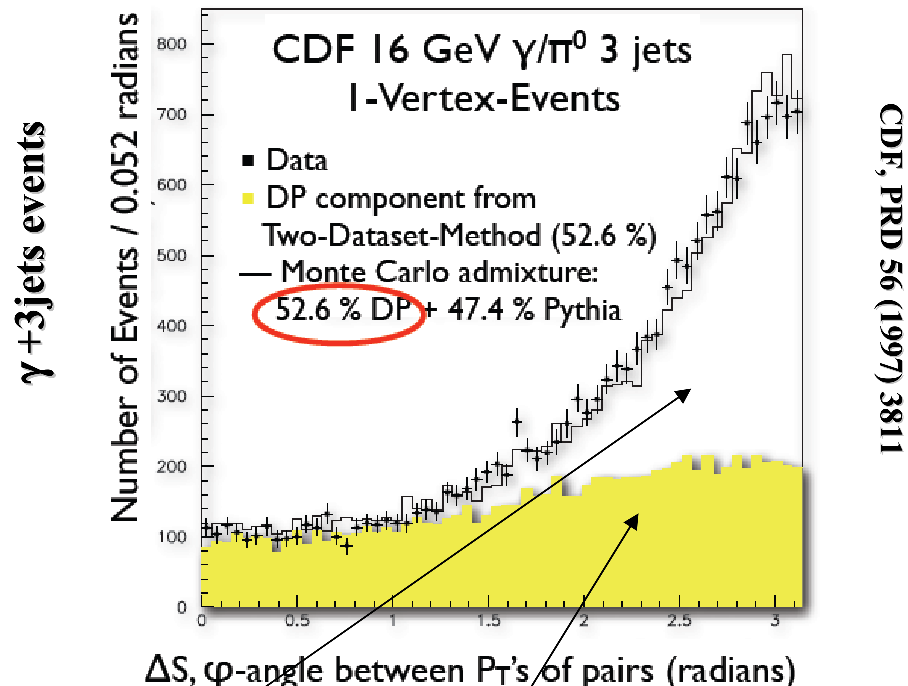
The evidence of MPI is confirmed in other hadron-hadron colliders:

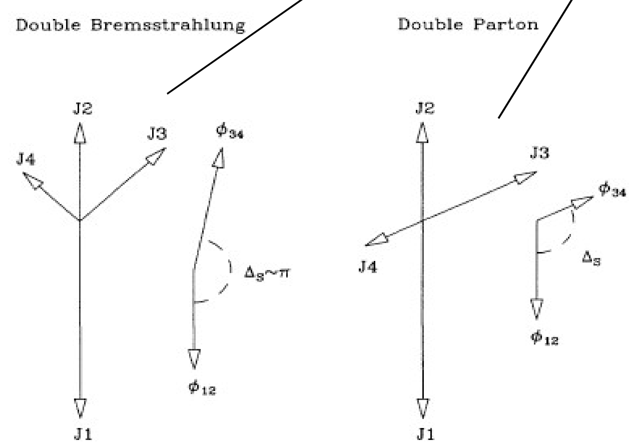
Indirect evidence at UA5 (p-p 540 GeV):





Direct evidence at the Tevatron (p-p 1.8 TeV):





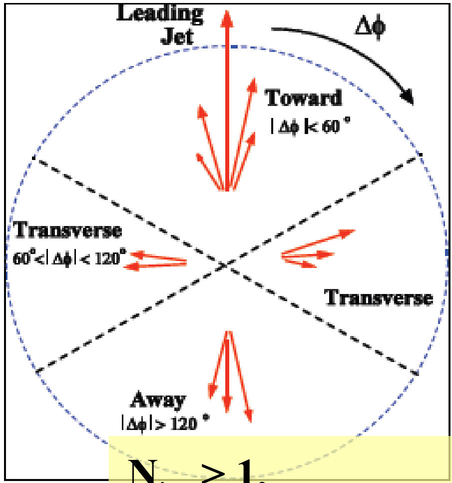
azimuthal angle between the p_T vectors of γ+j and j+j

The underlying event: tunings

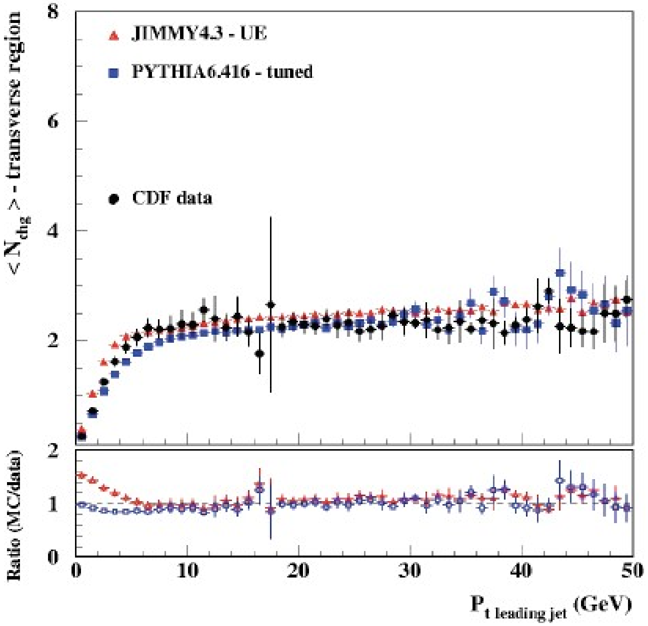
PDF PS **UE** Hadron Decay

Tune on data distributions (differential in η, φ) of $<N_{chg}>$ and $<\Sigma P_T>$ in the region transverse to the leading jet

Use CDF data at 630 GeV and 1.8 TeV (for energy extrapolation)

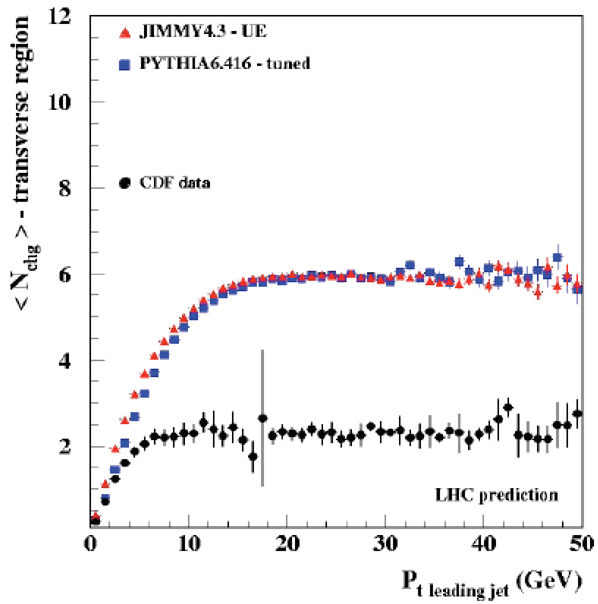


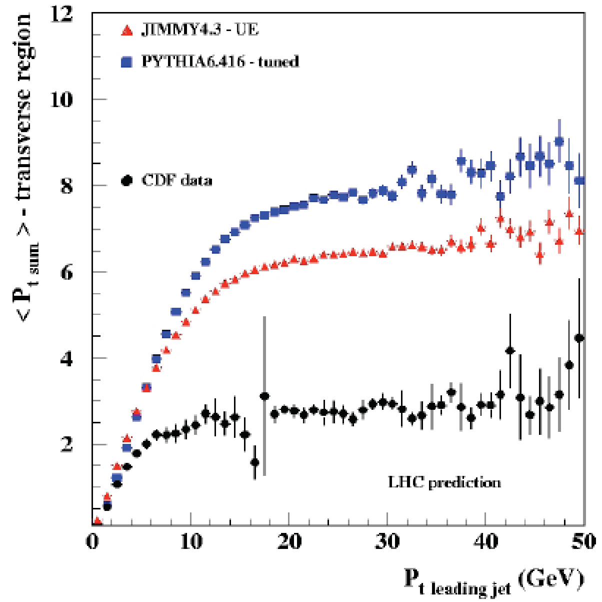
$N_{jets} > 1$,
$|\eta_{jet}| < 2.5$,
$E_T^{jet} > 10$ GeV,
$|\eta_{track}| < 2.5$,



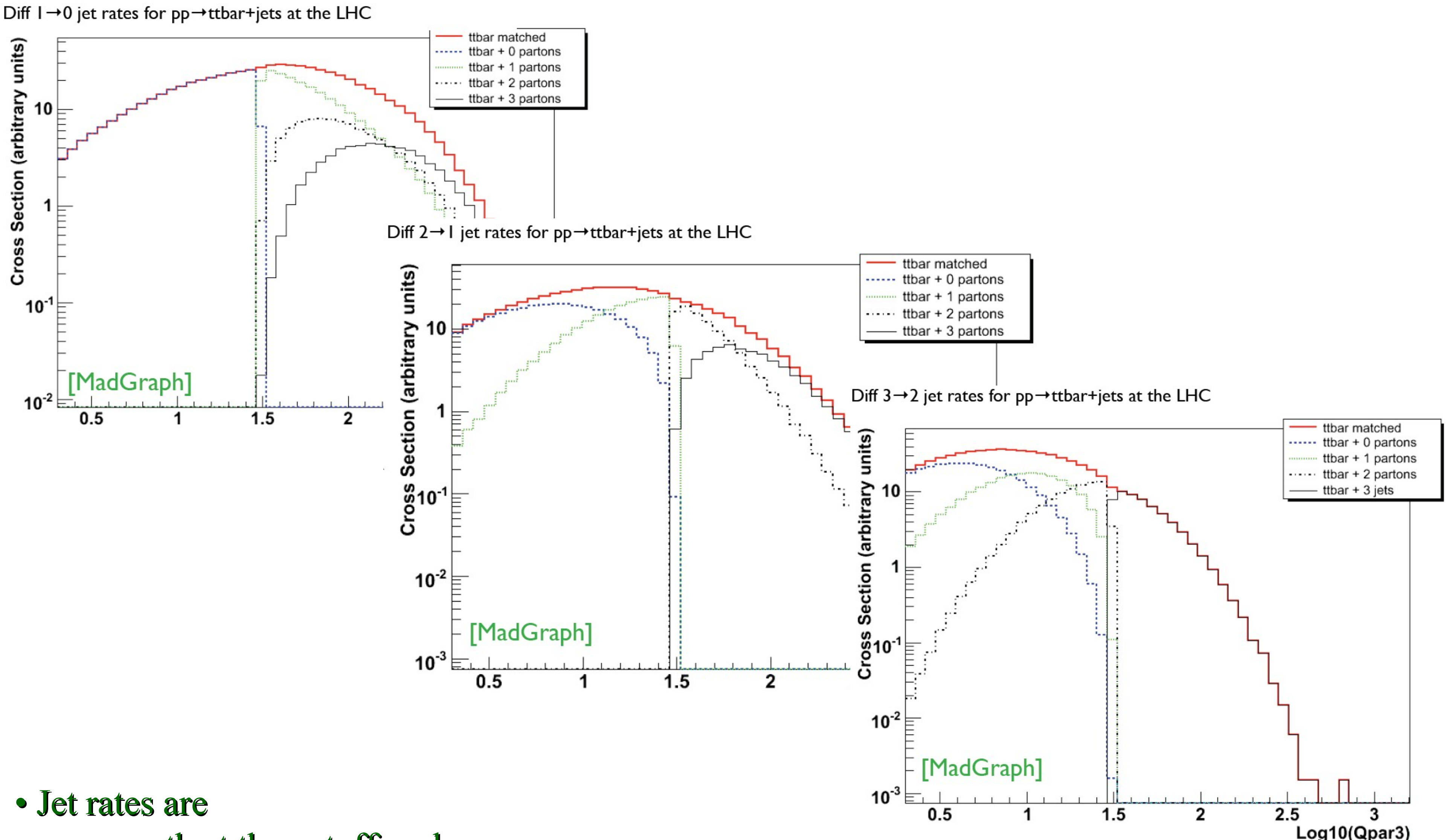
- Leave radiation (IS/FS) as default.
- Tune colour reconnection parameters, matter distribution parameters.
- Main parameter is the p_T cutoff

Extrapolation at LHC energies may give very different results !





Validation of the matching procedure



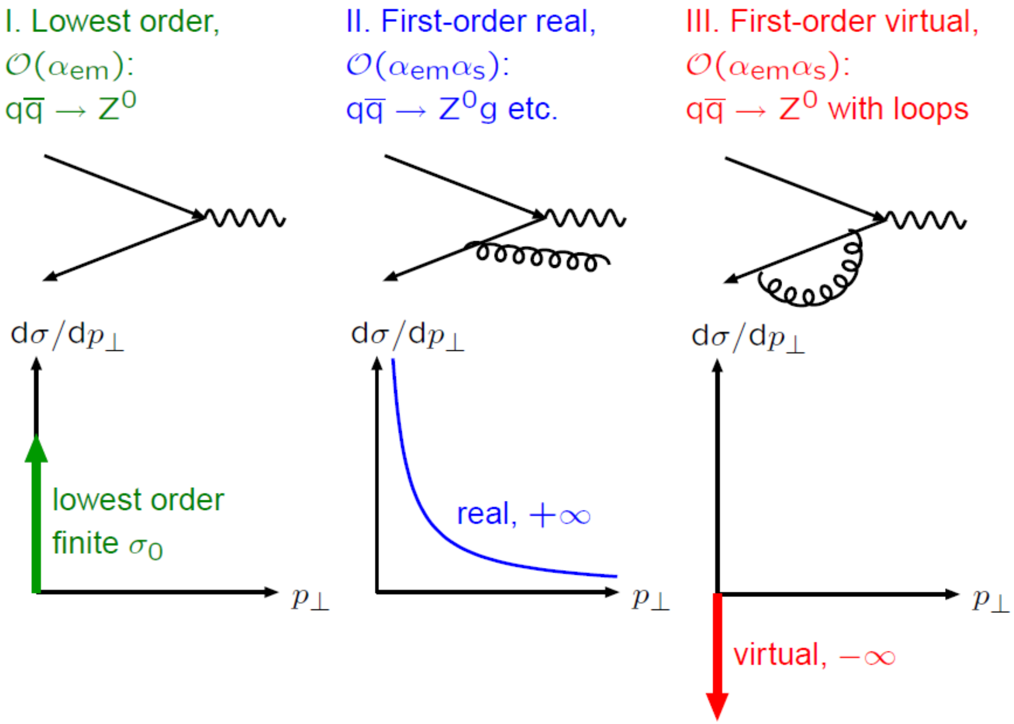
- Jet rates are
  - smooth at the cutoff scale
  - independent upon the cutoff scale (under reasonable variations)

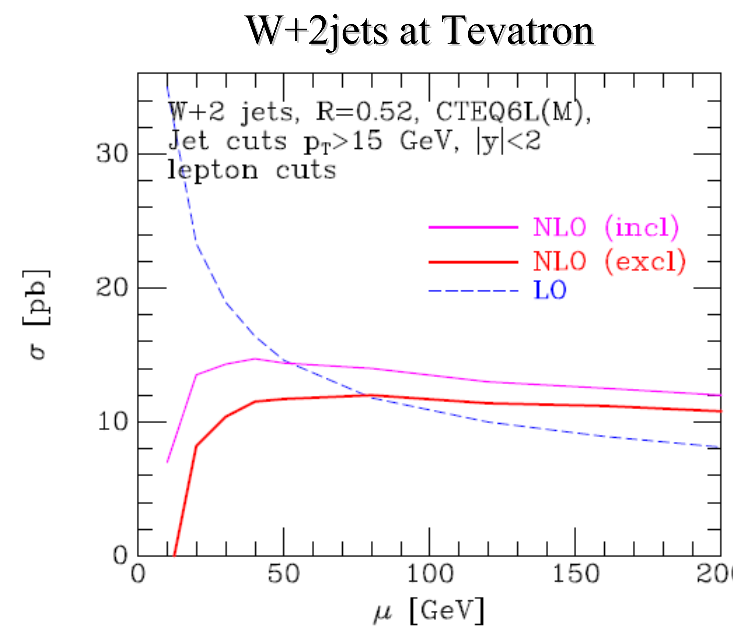
NLO ME

Remember: two separately divergent integrals

$$\sigma_{NLO} = \int_{m+1} d\sigma^R + \int_m d\sigma^V$$

- Must combine them before numerical integration
- Is it possible to include virtual corrections as well in the calculations?





Campbell, Huston, hep-ph/0405276

NLO(k) calculations are superior to LO(k) because they include exactly the virtual corrections. Predictions for (differential) cross-sections are more reliable than LO(k). Parton shower techniques and LO ME only deal with real parton emissions

This allows to predict normalizations and shapes of the distributions more coorectly (NLO(k))

They may be inferior if interested in multiple real emission beyond LO(k) (i.e. LO(k+n))

NLO ME with PS

Aims:

- hard emission described as in NLO computations
- soft and collinear emissions described by parton showers
- the matching between the two regions is everywhere smooth

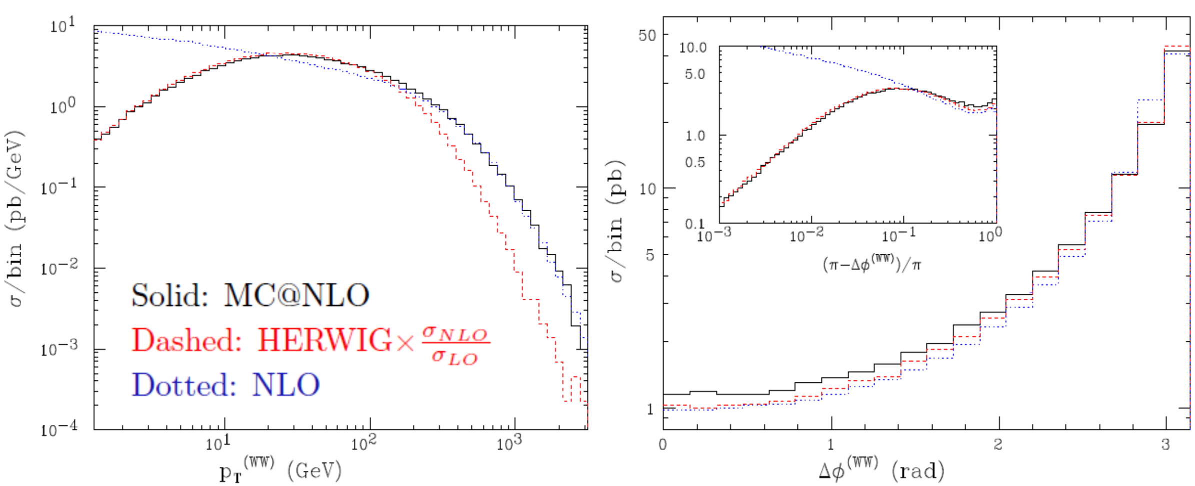
Strategy, in a nutshell:

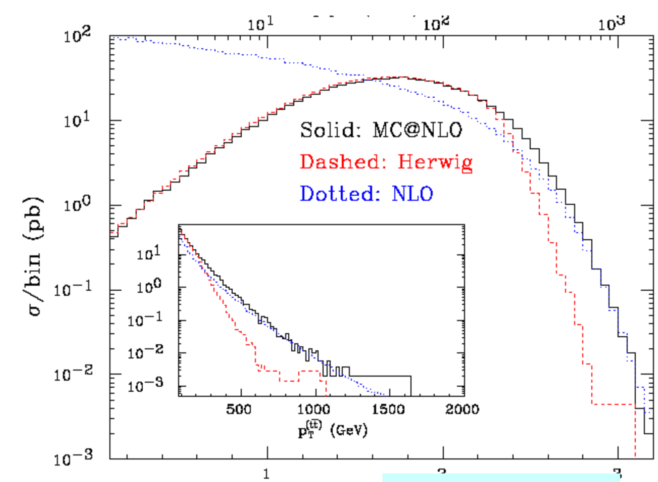
- calculate the NLO ME of an n-body process (real+virtual n-body+ real (n+1)-body)
- calculate analytically how the first shower emission off an n-body phase space populates an (n+1)-body phase space
- subtract the shower expression from the (n+1) ME to get the “bare” (n+1) events
- add showers everywhere

Features:

- normalizations and shapes of the (differential) cross-sections are more precise
- events with negative weights may appear, and need to use them in the analyses

W^+W^- Observables



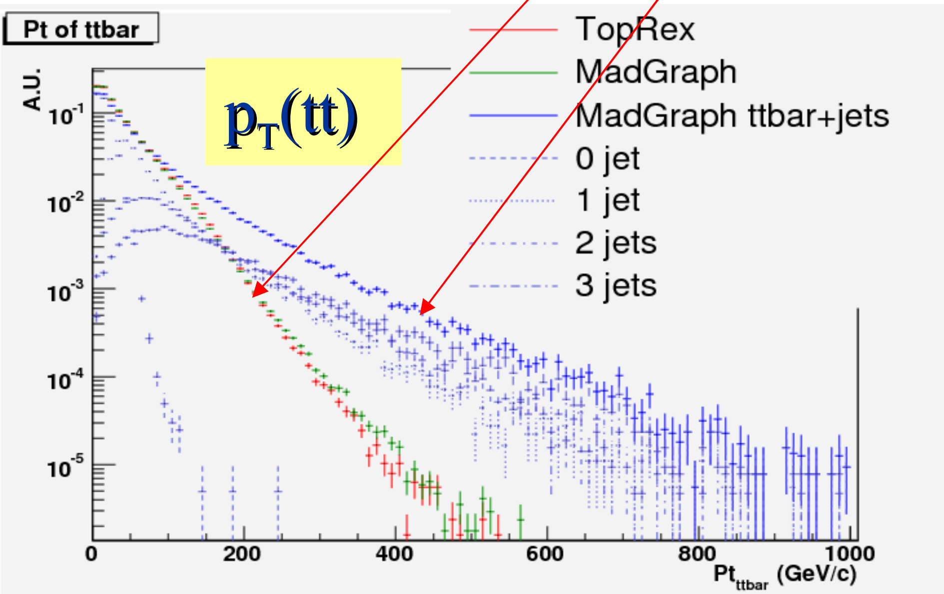


hep-ph/0305252

log(p_T(tt))

Example: top pair production

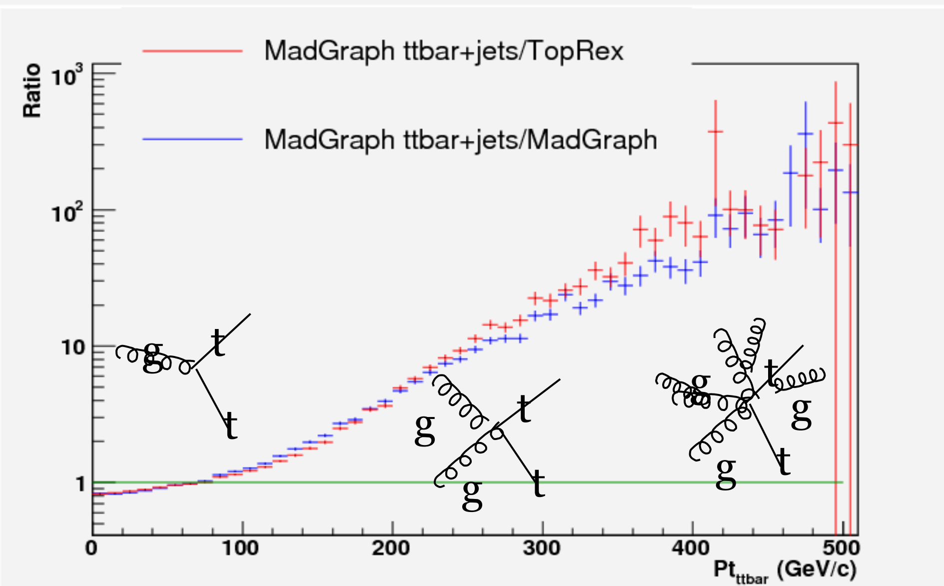
Large differences between PS and ME in transverse variables related to radiation

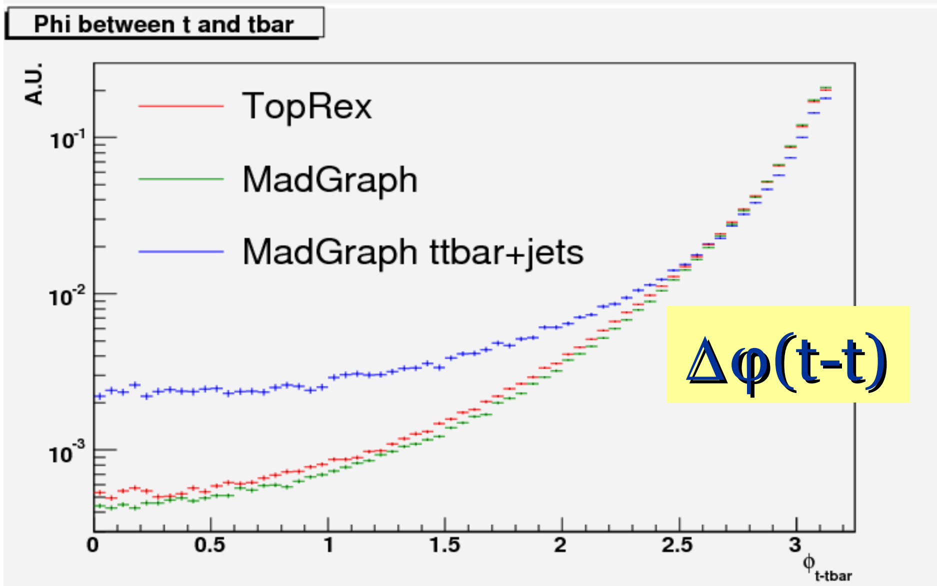


- Large effects at high $p_T(tt)=p_T$(radiation)
- Average $p_T(tt)$~60-70 GeV !
- 40% probability that a tt system recoils against a radiation larger than 50 GeV

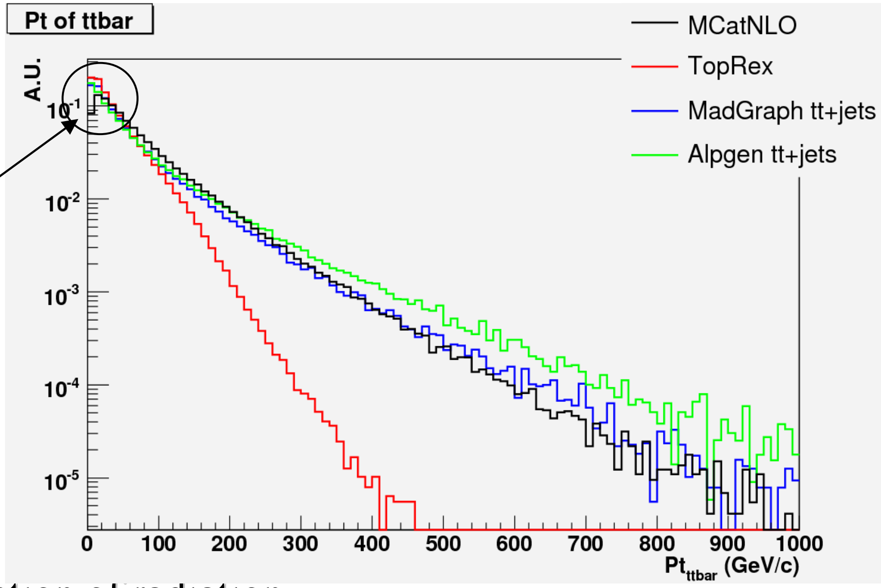
→ effect on reconstruction

→ Mandatory to use the same strategies for physics backgrounds like W/Z+Njets





Example: top pair production



Comparisons to an NLO generator (MC@NLO) show that the hard part of radiation is better described by ME than by PS.

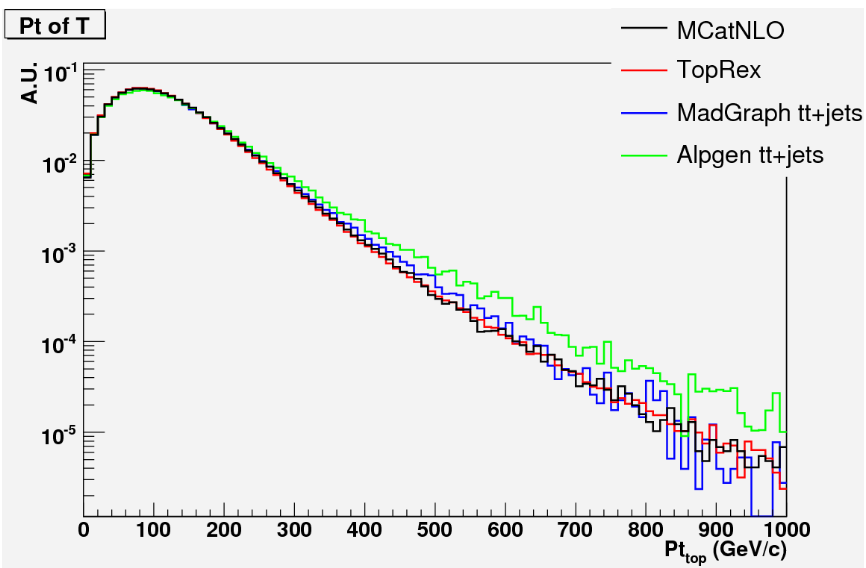
- Non perturbative part treated by HERWIG/JIMMY.

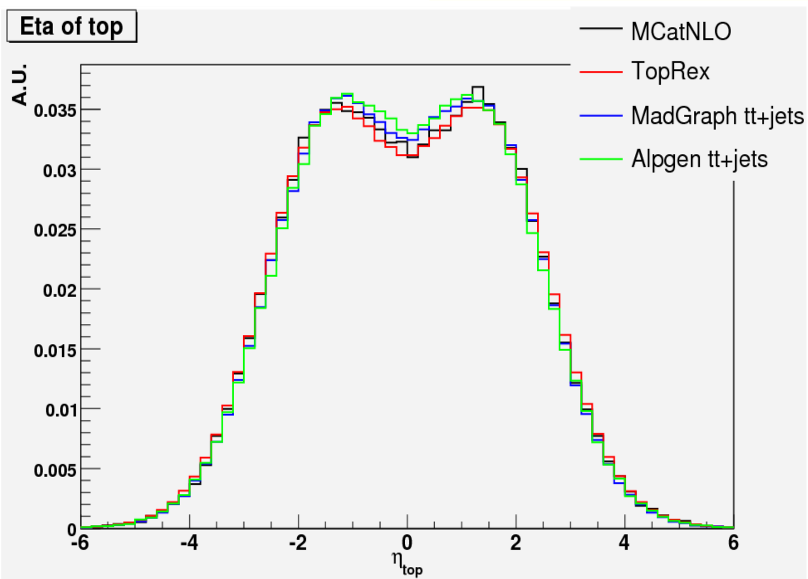
Still a very important step in understanding high p_T radiation and increase our confidence in the process description. Also gives indications on:

- Relative importance of first emission
- Normalization
- Indication of systematic errors associated to the description of radiation.

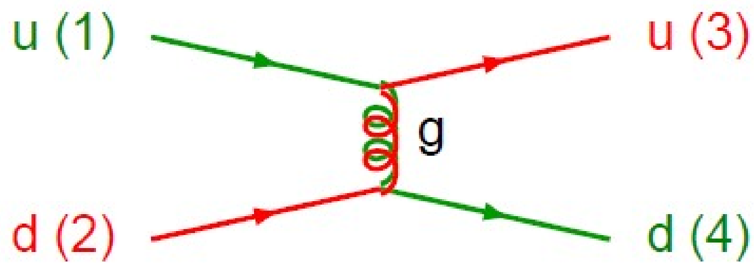
Preliminary generator level study

Essential agreement in the p_T(tt) tail. Good agreement in other distributions.





Kinematics

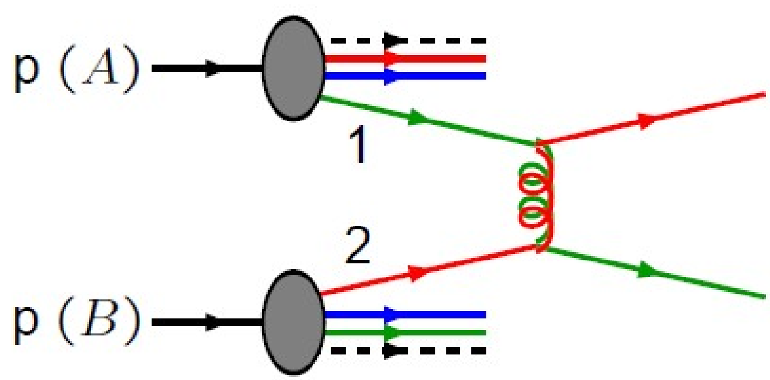


$$\hat{s} = (p_1 + p_2)^2$$

$$\hat{t} = (p_1 - p_3)^2 = -\hat{s}(1 - \cos\hat{\theta})/2$$

$$\hat{u} = (p_1 - p_4)^2 = -\hat{s}(1 + \cos\hat{\theta})/2$$

$$qq' \to qq' \quad : \quad \frac{d\hat{\sigma}}{d\hat{t}} = \frac{\pi}{\hat{s}^2}\frac{4}{9}\alpha_s^2\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} \qquad (\sim \text{Rutherford})$$



$$s = (p_A + p_B)^2$$

$$x_1 \approx E_1/E_A$$

$$x_2 \approx E_2/E_B$$

$$\hat{s} = x_1 x_2 s$$

$$\sigma = \sum_{i,j} \iiint dx_1\, dx_2\, d\hat{t}\, f_i^{(A)}(x_1, Q^2)\, f_j^{(B)}(x_2, Q^2)\, \frac{d\hat{\sigma}_{ij}}{d\hat{t}}$$

Status of calculations