Tools and Monte-Carlos

for the

Old and New Physics

Fawzi BOUDJEMA LAPTH-Annecy, France

Great Idea: A New Physics Model

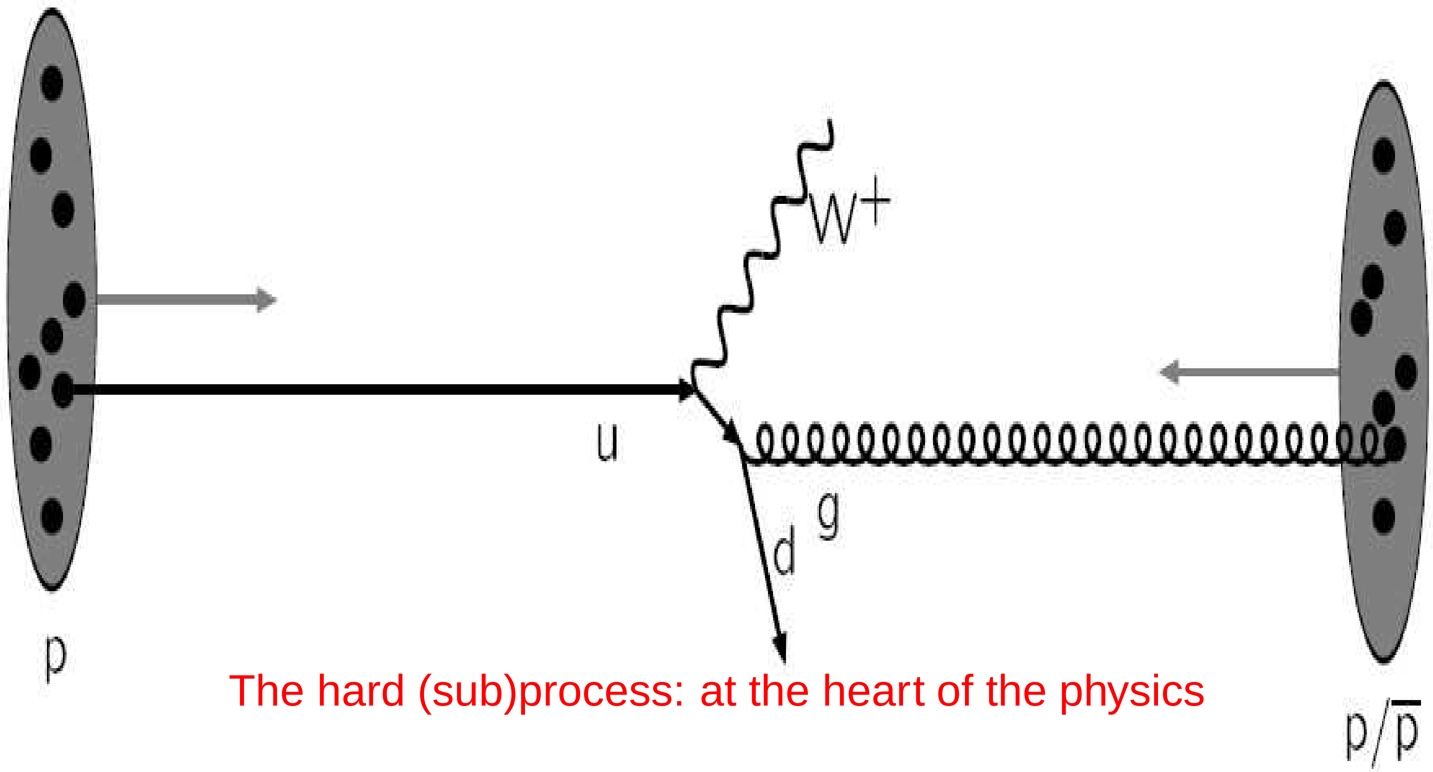
FINAL AIM

Nobel Prize if LHC validates!

Movie: The structure of an event

Incoming beams: partons densities

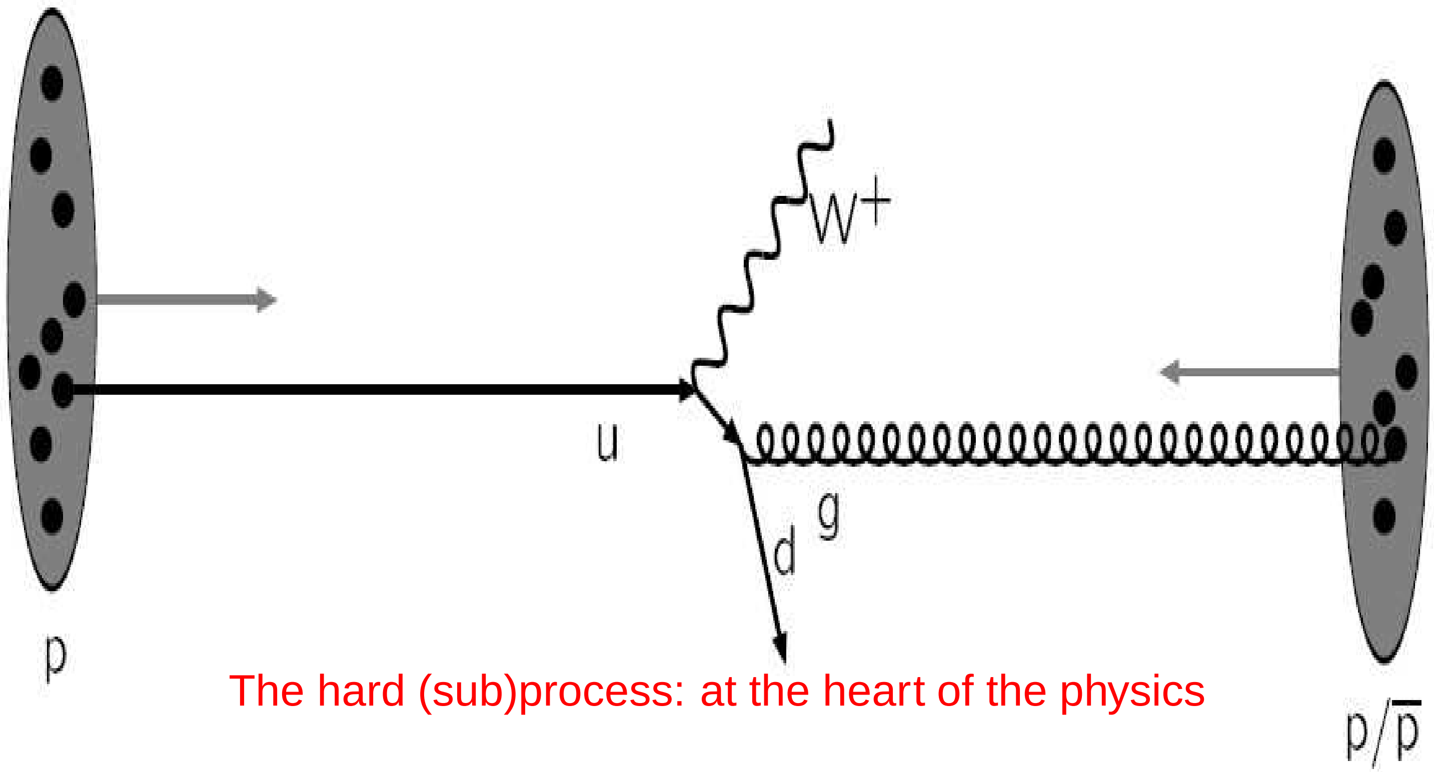
Movie: The structure of an event



The hard (sub)process: at the heart of the physics

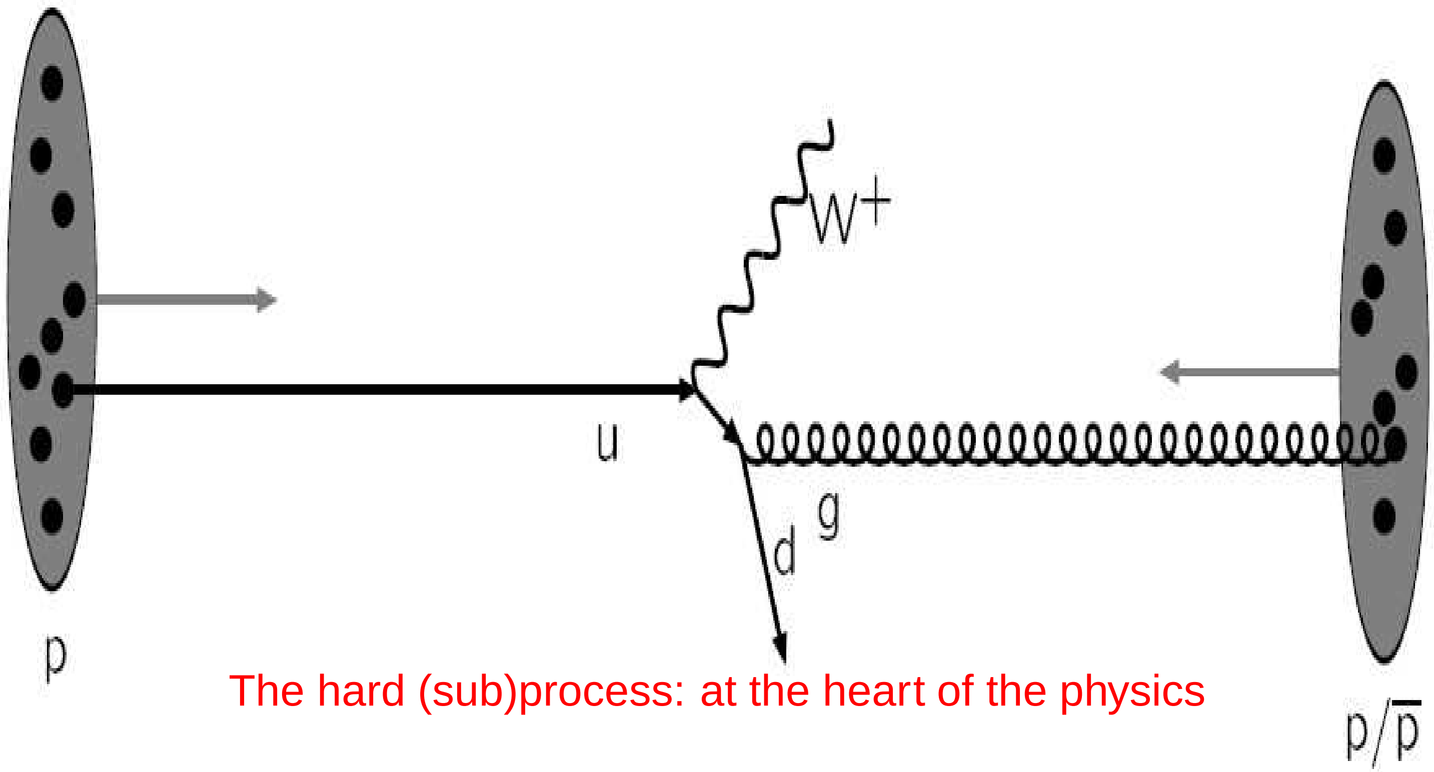
- Hard process is well understood and well described: relies on a firm perturbative framework.

Movie: The structure of an event



The hard (sub)process: at the heart of the physics

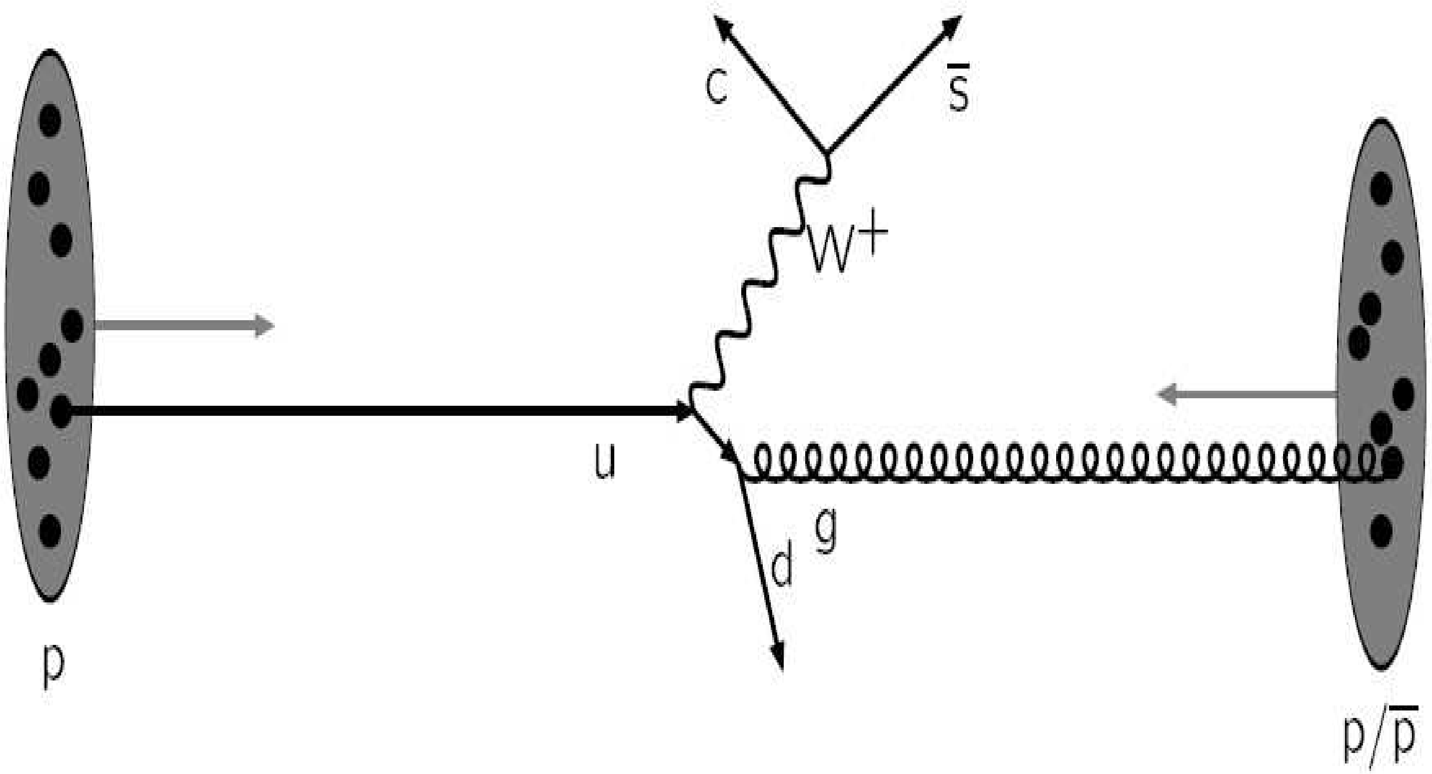
- Hard process is well understood and well described: relies on a firm perturbative framework.
- described by Matrix Elements (ME)
  This does not mean that it is very well calculated



The hard (sub)process: at the heart of the physics

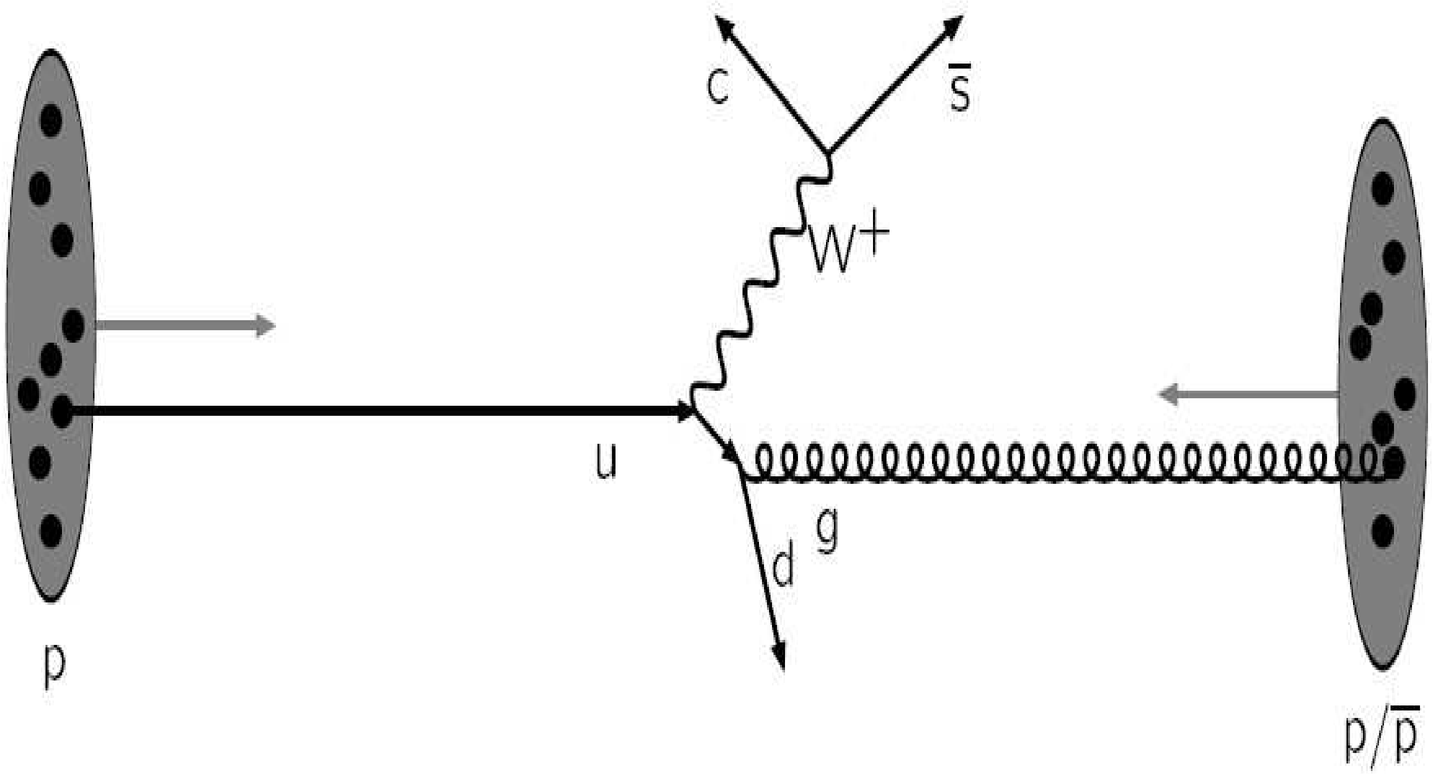
- Hard process is well understood and well described: relies on a firm perturbative framework.
- described by Matrix Elements (ME)
  This does not mean that it is very well calculated
- issue of higher order (NLO), most calculations only LO say

Movie: The structure of an event



Decays of resonances: correlated with hard process

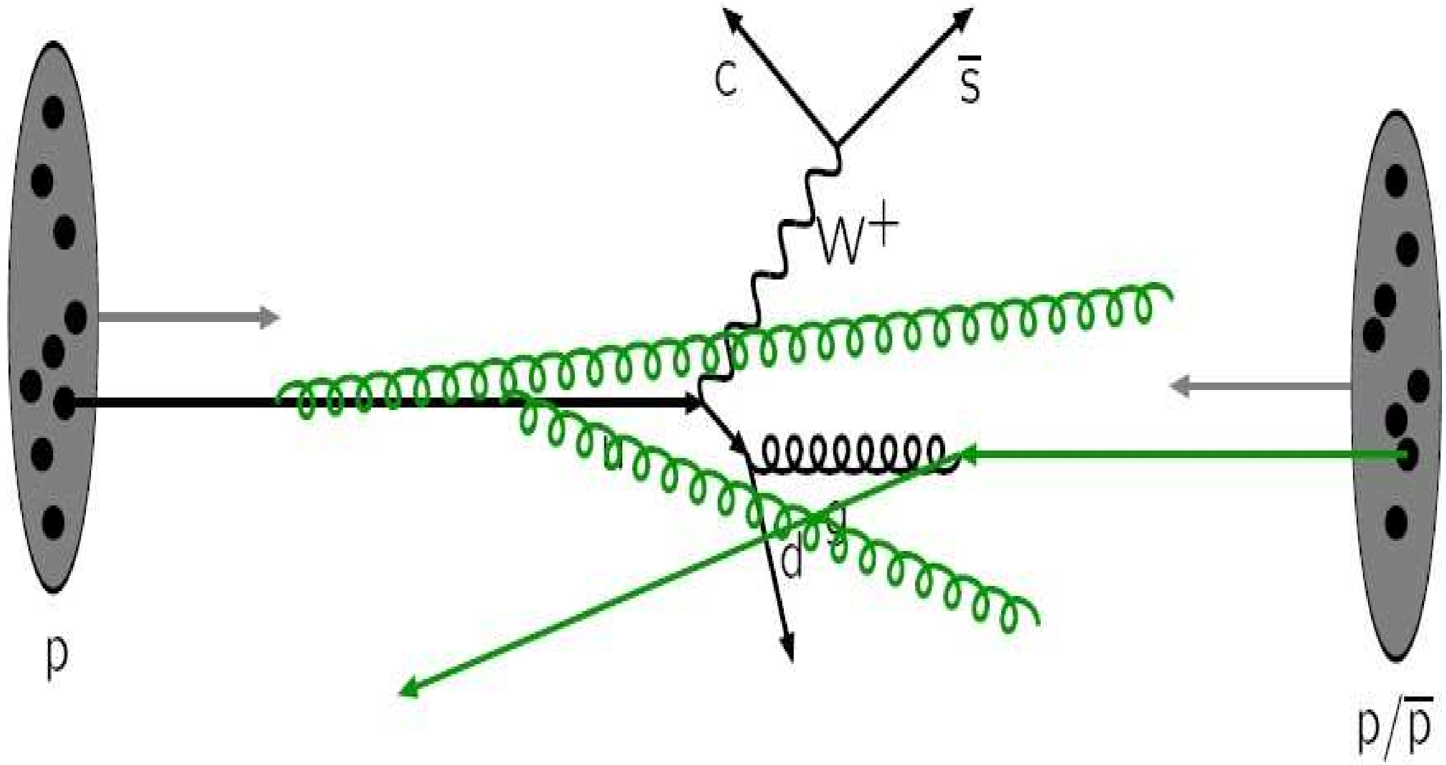
- Approximation: W on-shell



Decays of resonances: correlated with hard process

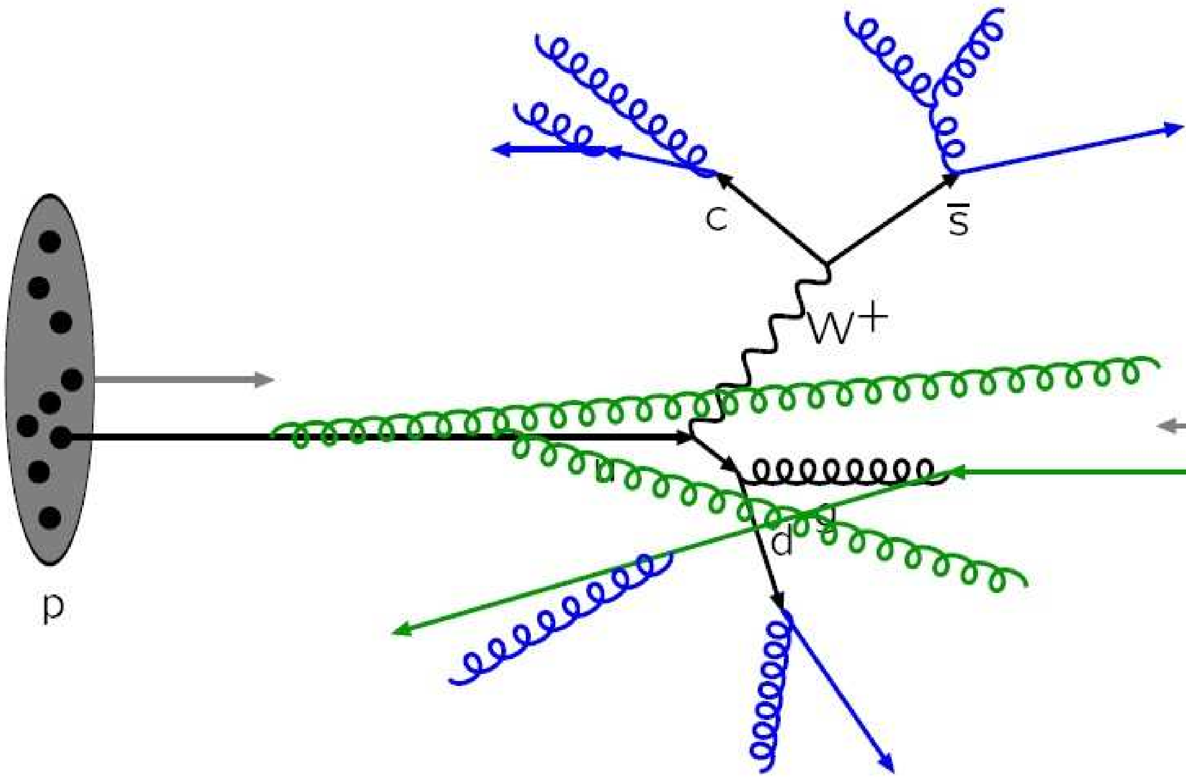
- Approximation: W on-shell
- Spin effect in decays?

Movie: The structure of an event



ISR: Initial State Radiation

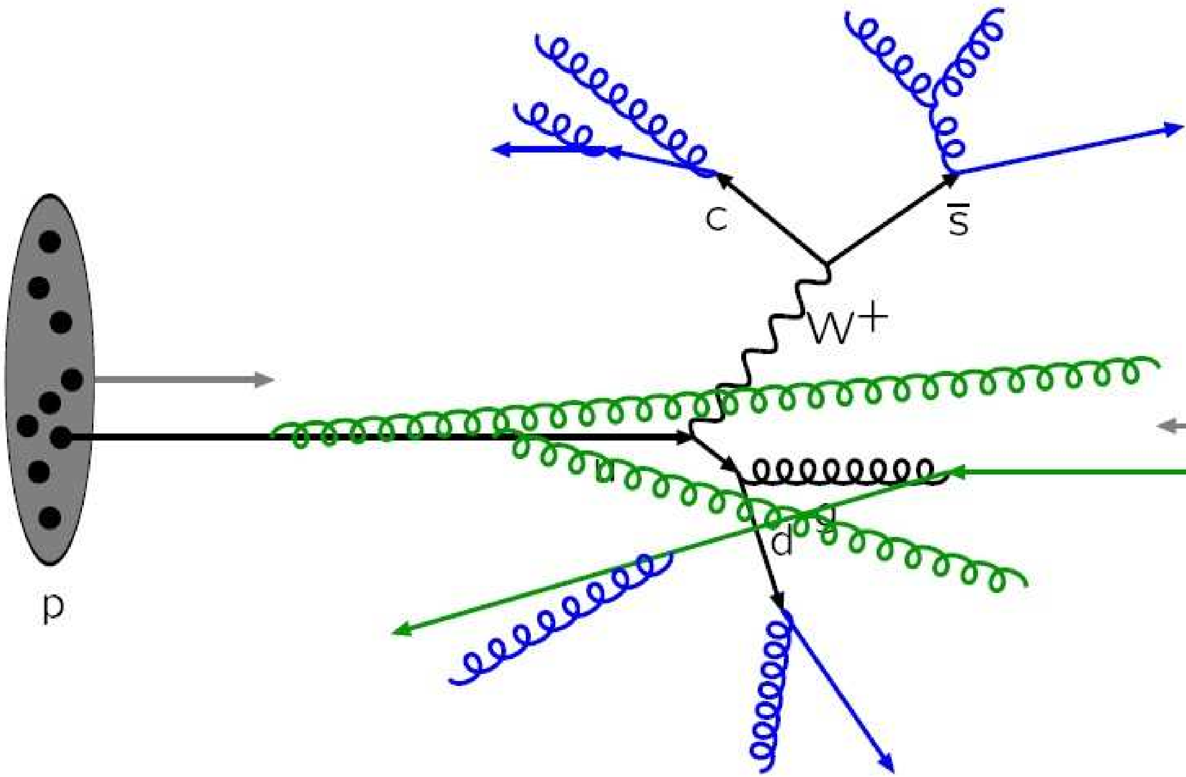
Space-like parton showers (PS)



FSR: Final State Radiation

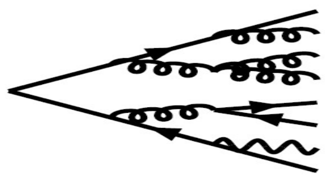
time-like parton showers (PS)

Movie: The structure of an event



FSR: Final State Radiation

time-like parton showers (PS)



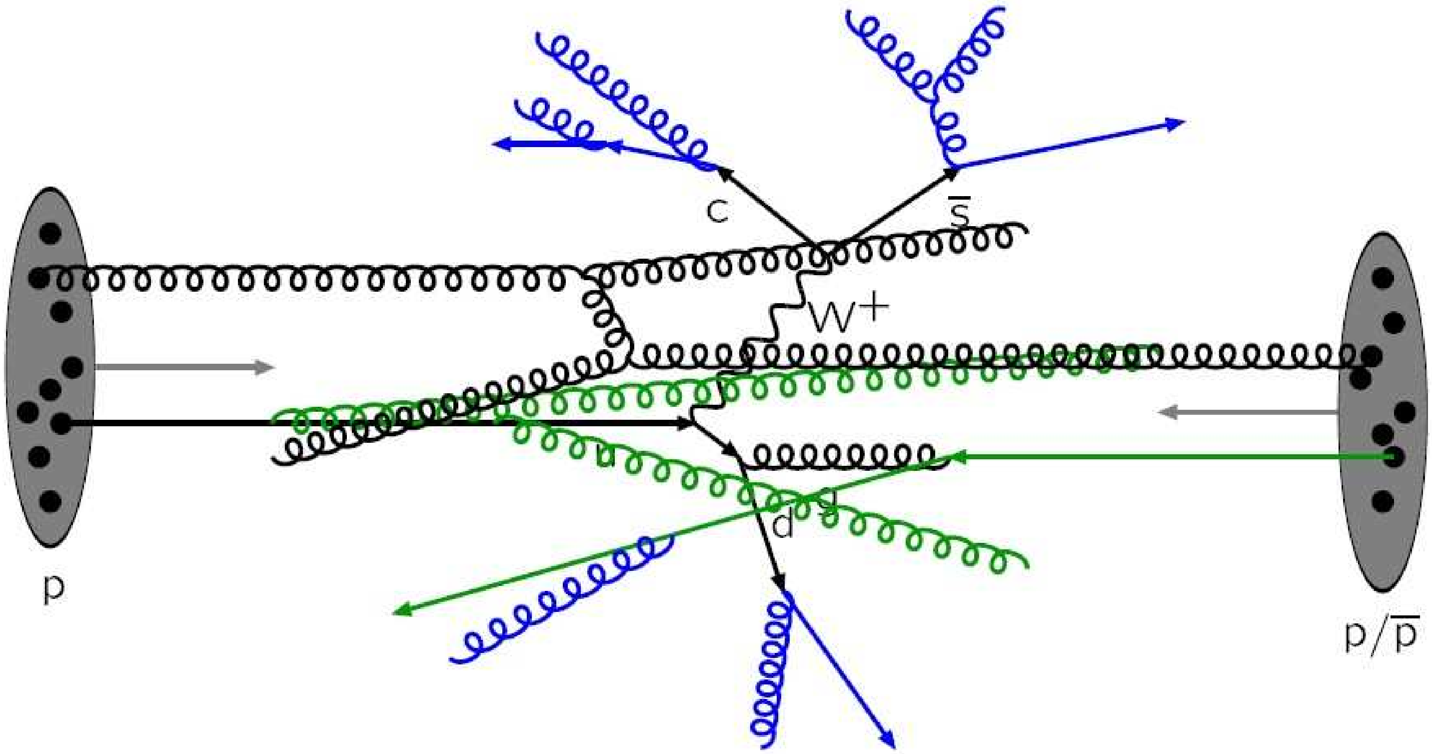
$q \longrightarrow qg$

$g \longrightarrow gg$

$g \longrightarrow q\bar{q}$

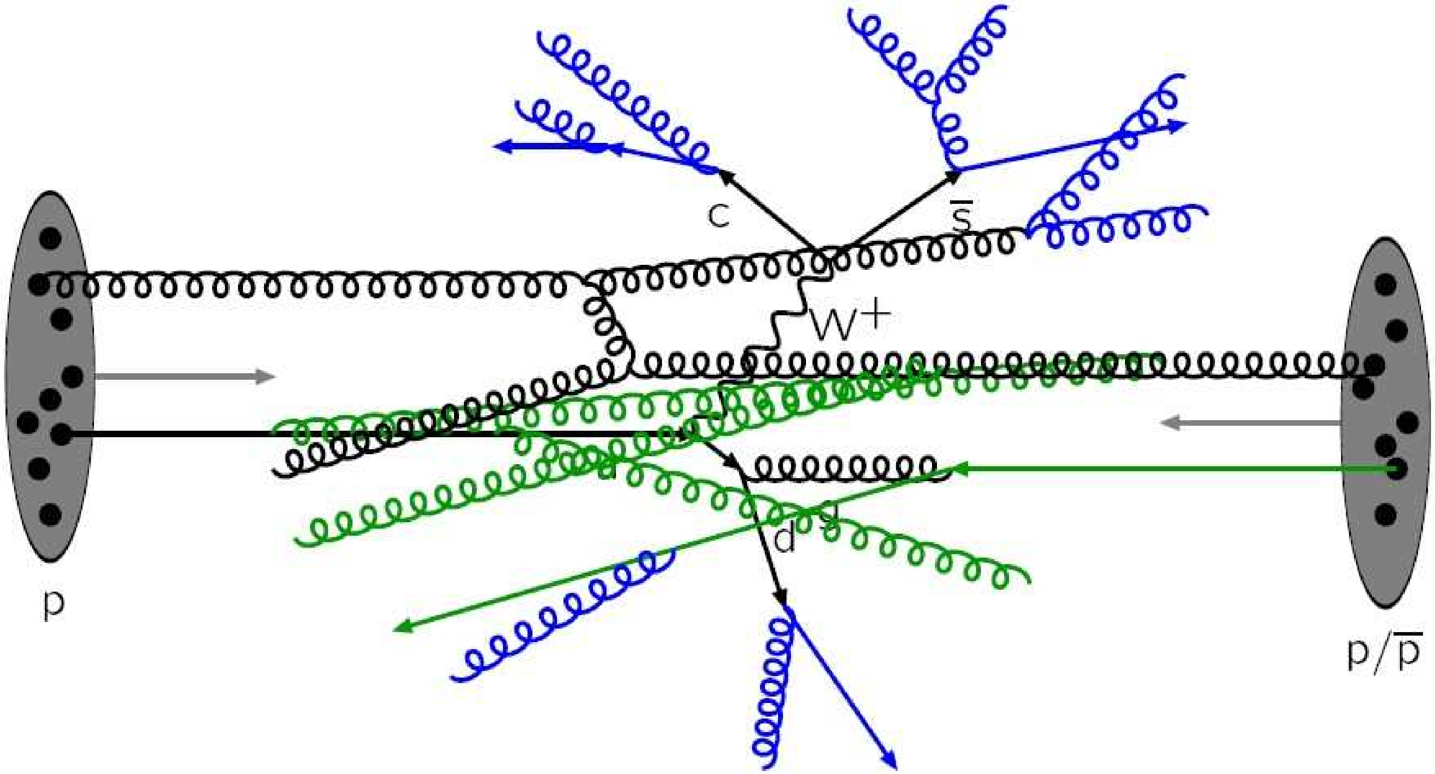
$q \longrightarrow q\gamma$

Movie: The structure of an event



Multiple parton-parton interactions (MPI)

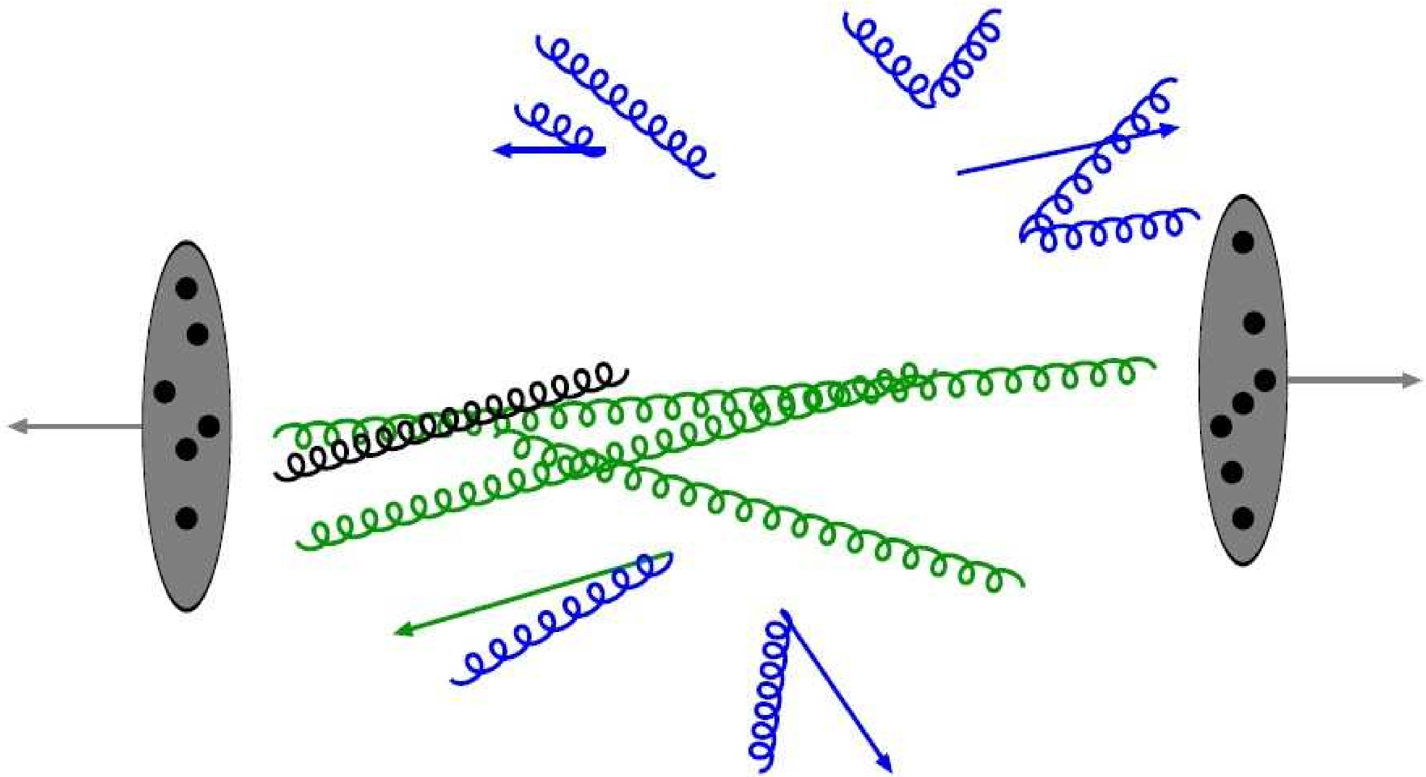
The muck



MPI with ISR and FSR !

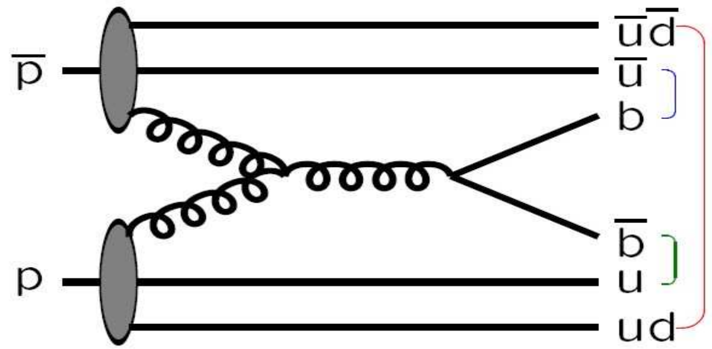
The muck

Movie: The structure of an event

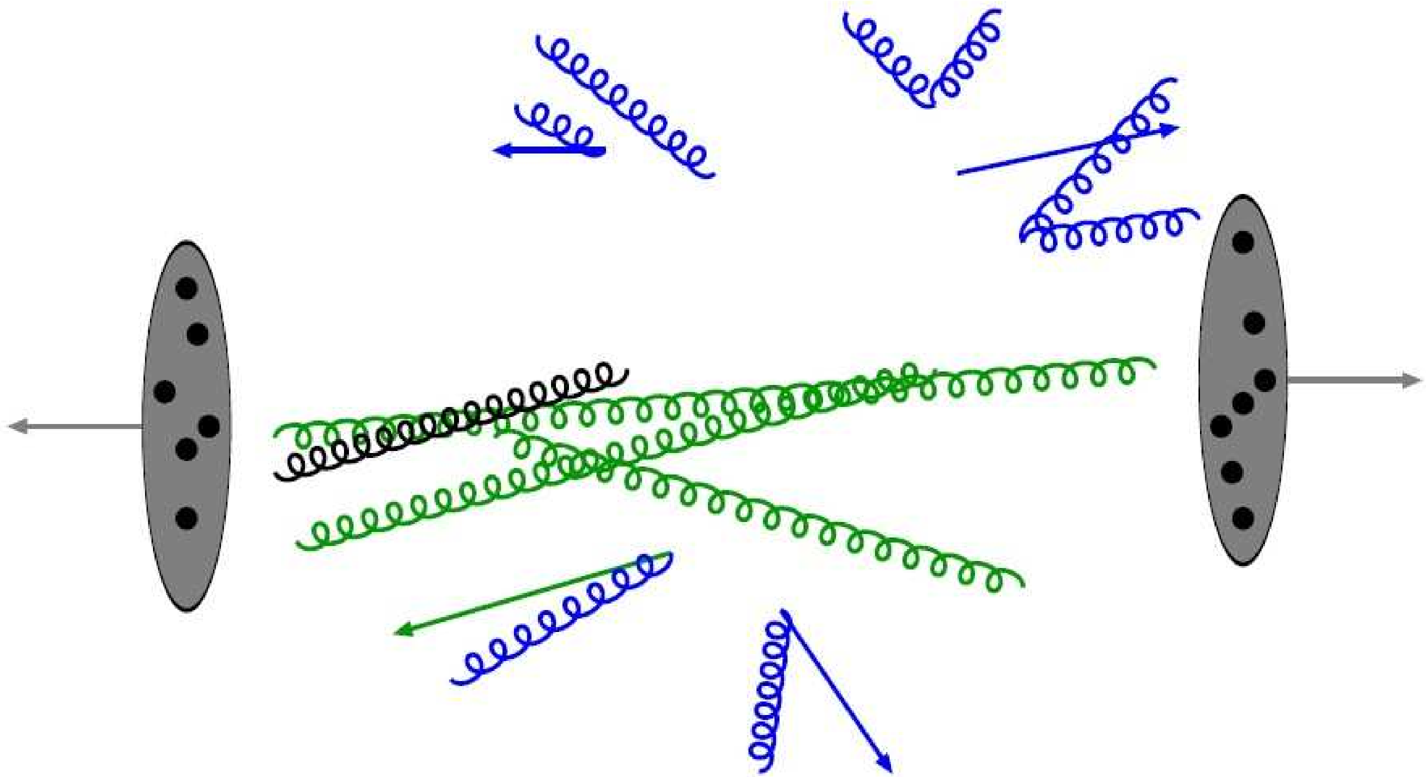


Beam remnants and other outgoing partons !

The muck

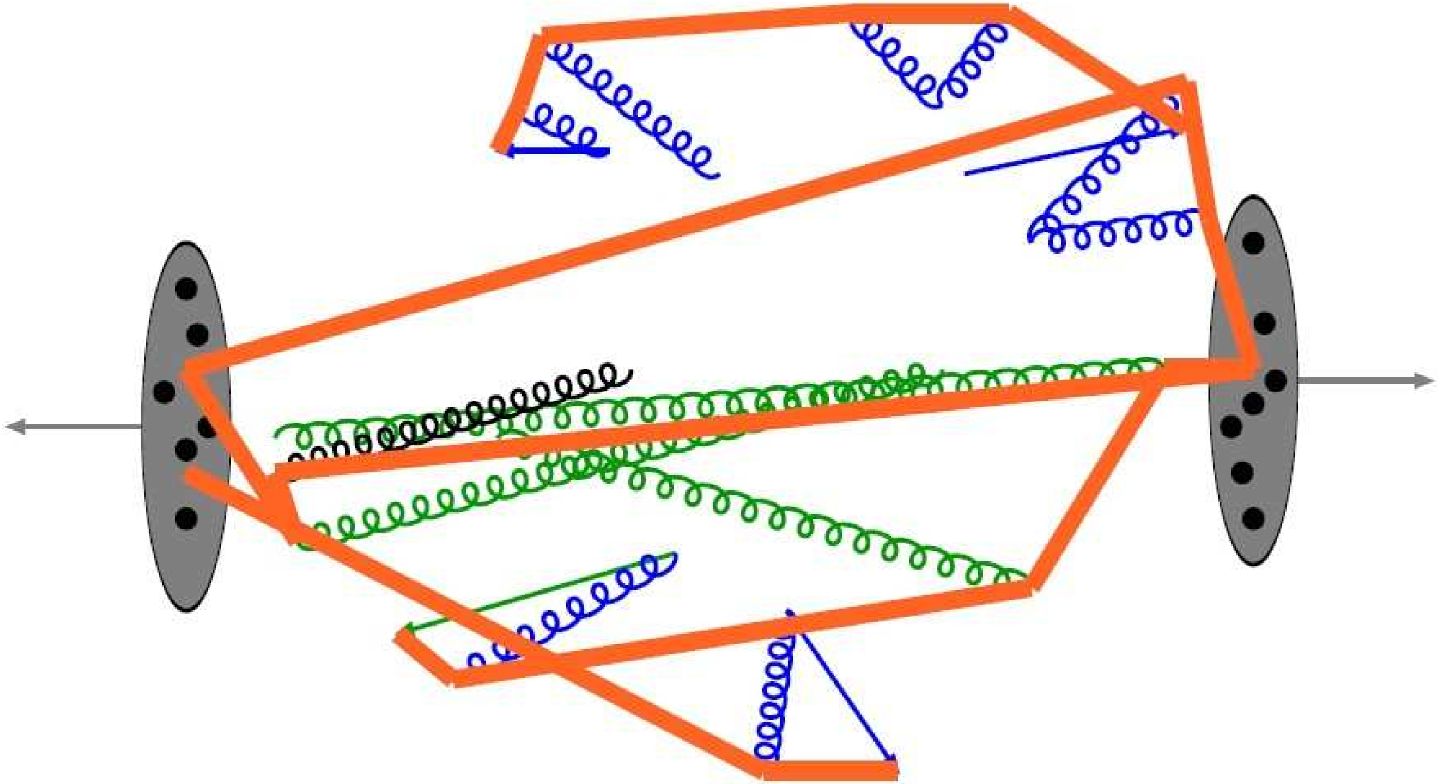


Beam remnants: coloured remains of the proton not taking part in the hard process, but they are colour connected to the hard process.



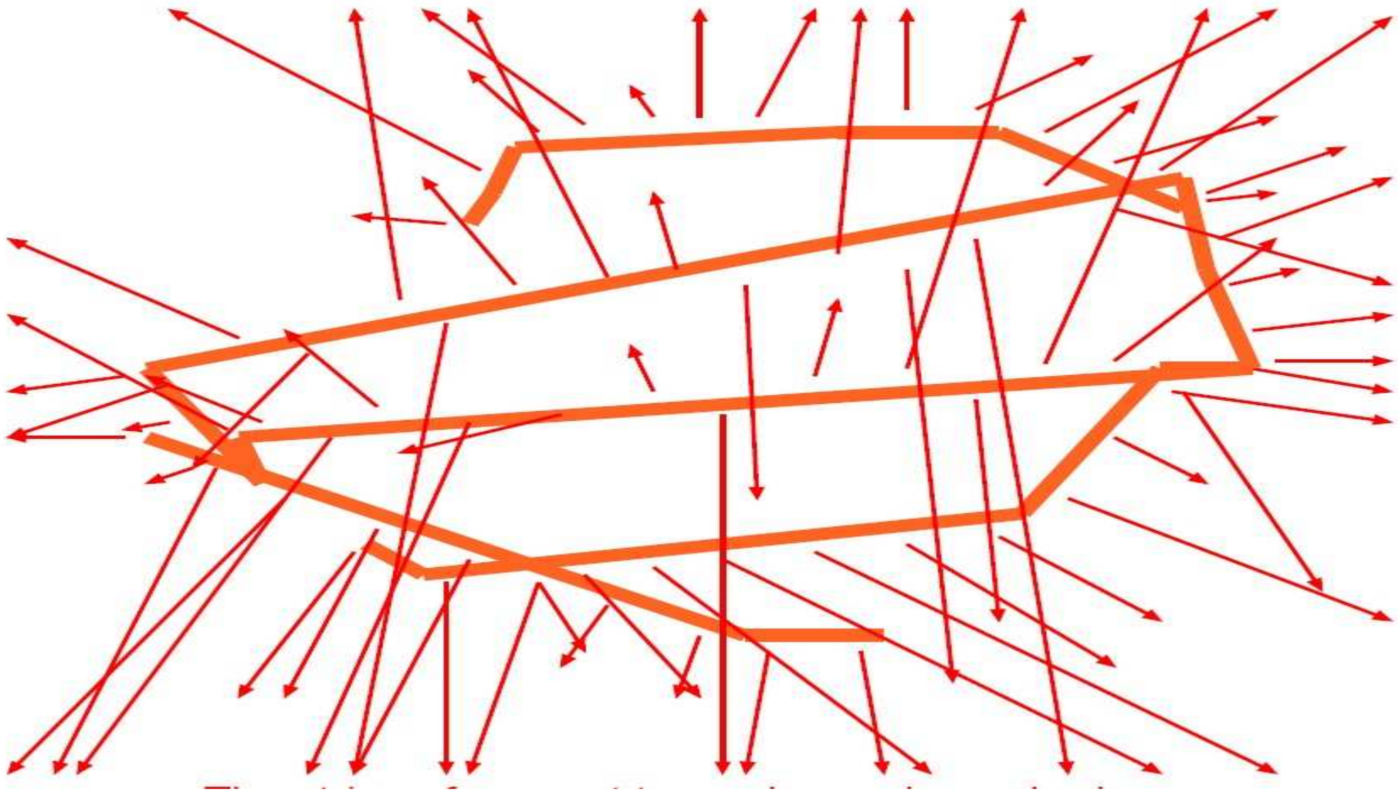
Beam remnants and other outgoing partons !

The muck: UE is not understood. It is modelled through some data that do not give strong constraints and hence extrapolation can be doubted.

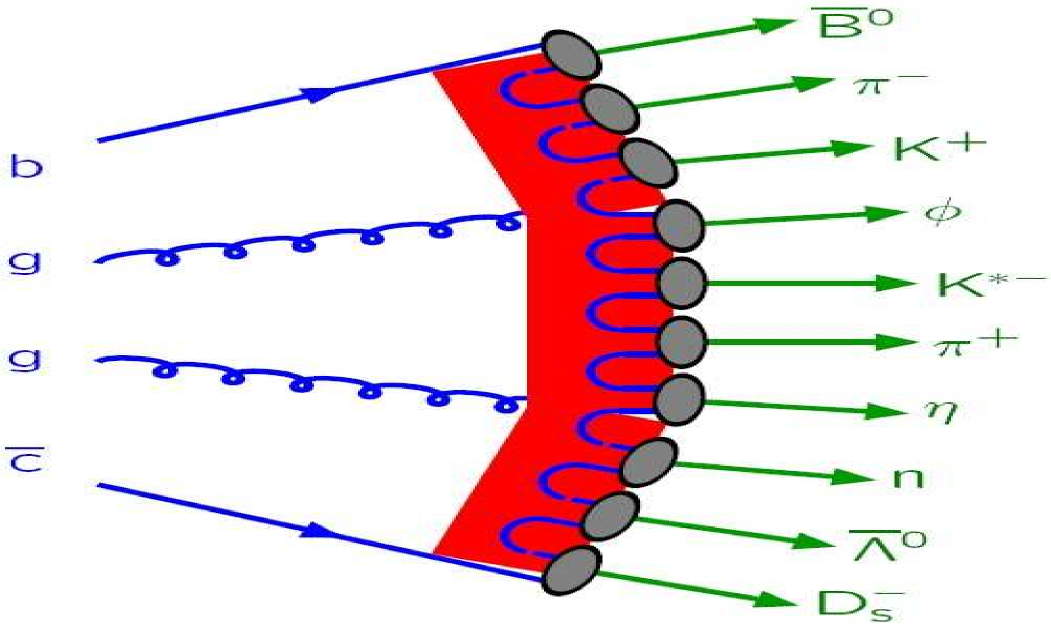


Everything is connected by colour confinement (here strings)

Movie: The structure of an event

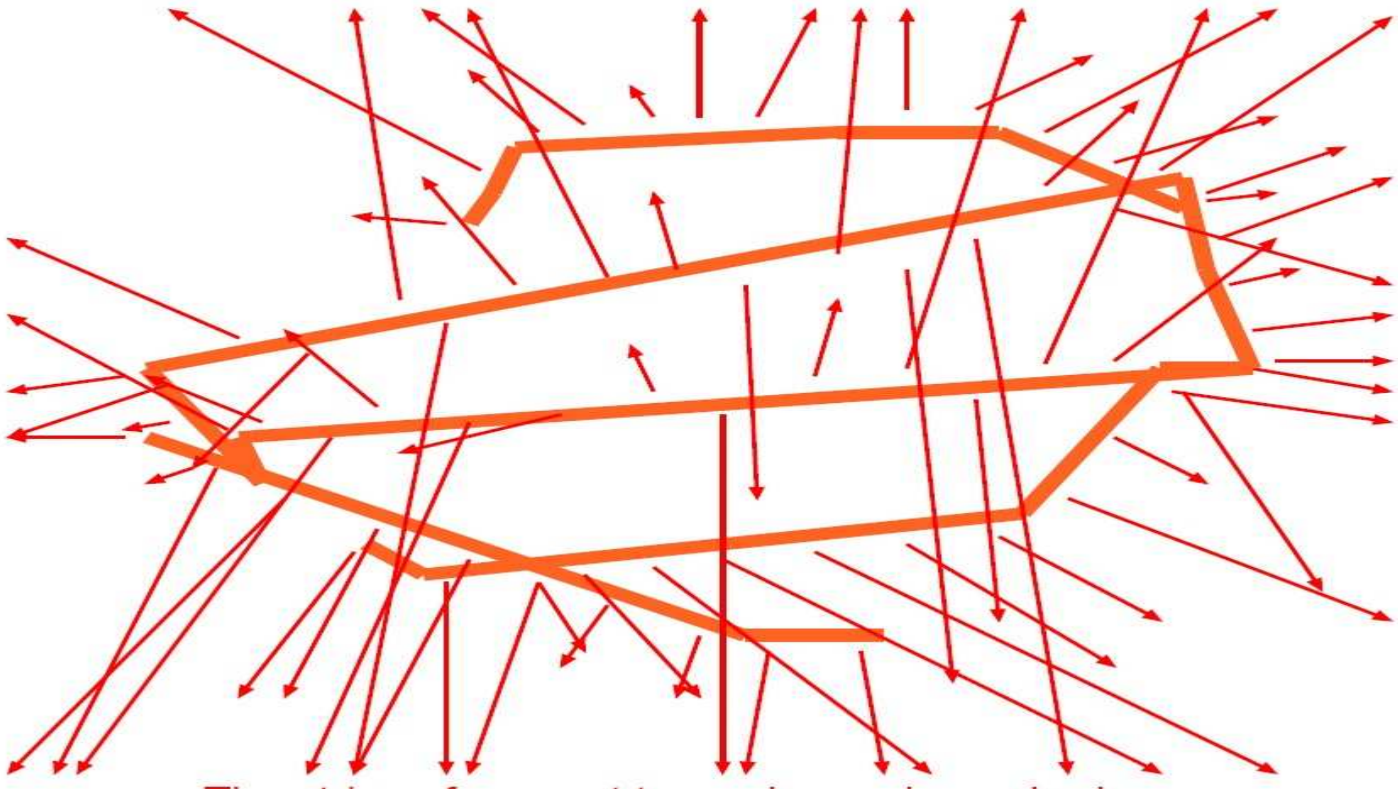


The strings fragments to produce hadrons



Hadronisation: Clusters to produce hadrons (Cluster Model)

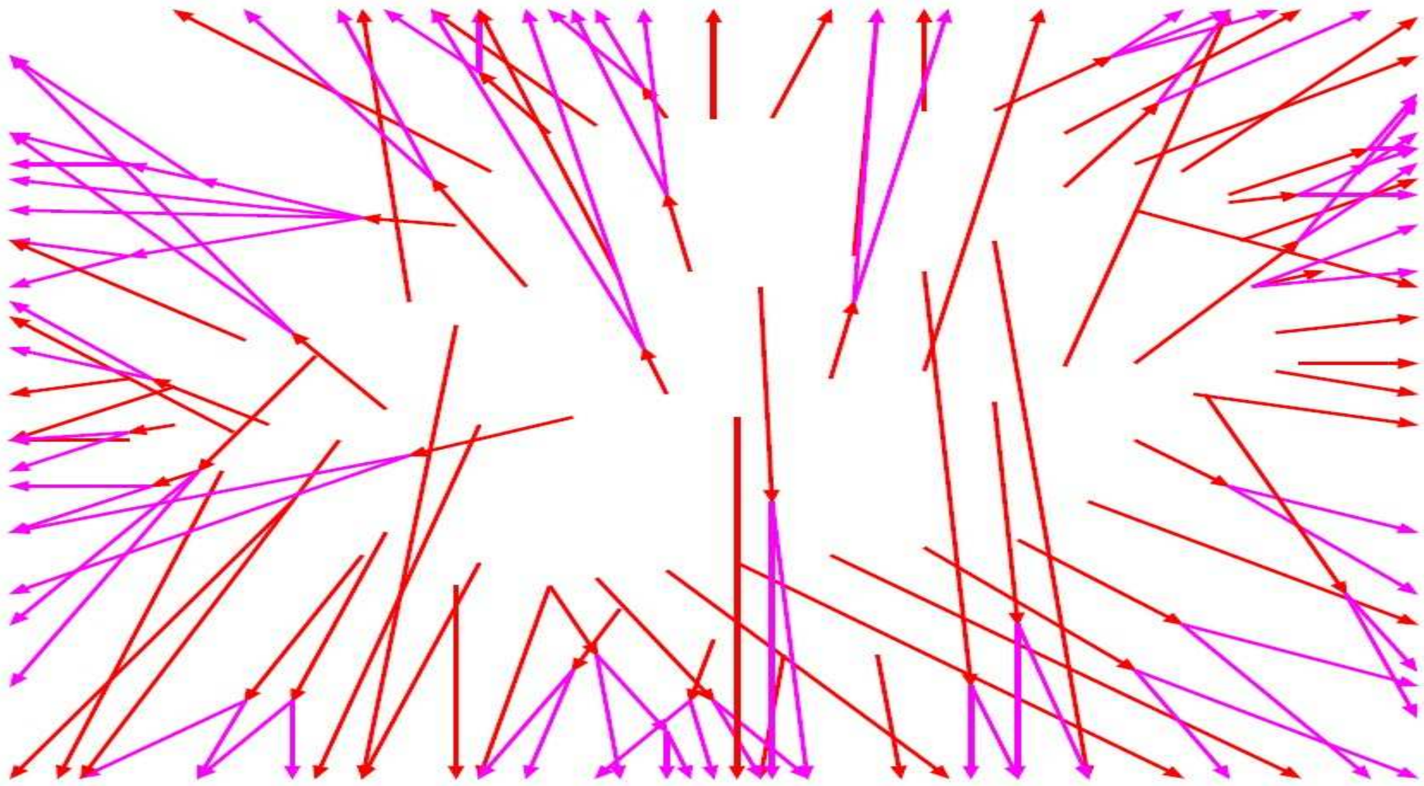
Movie: The structure of an event



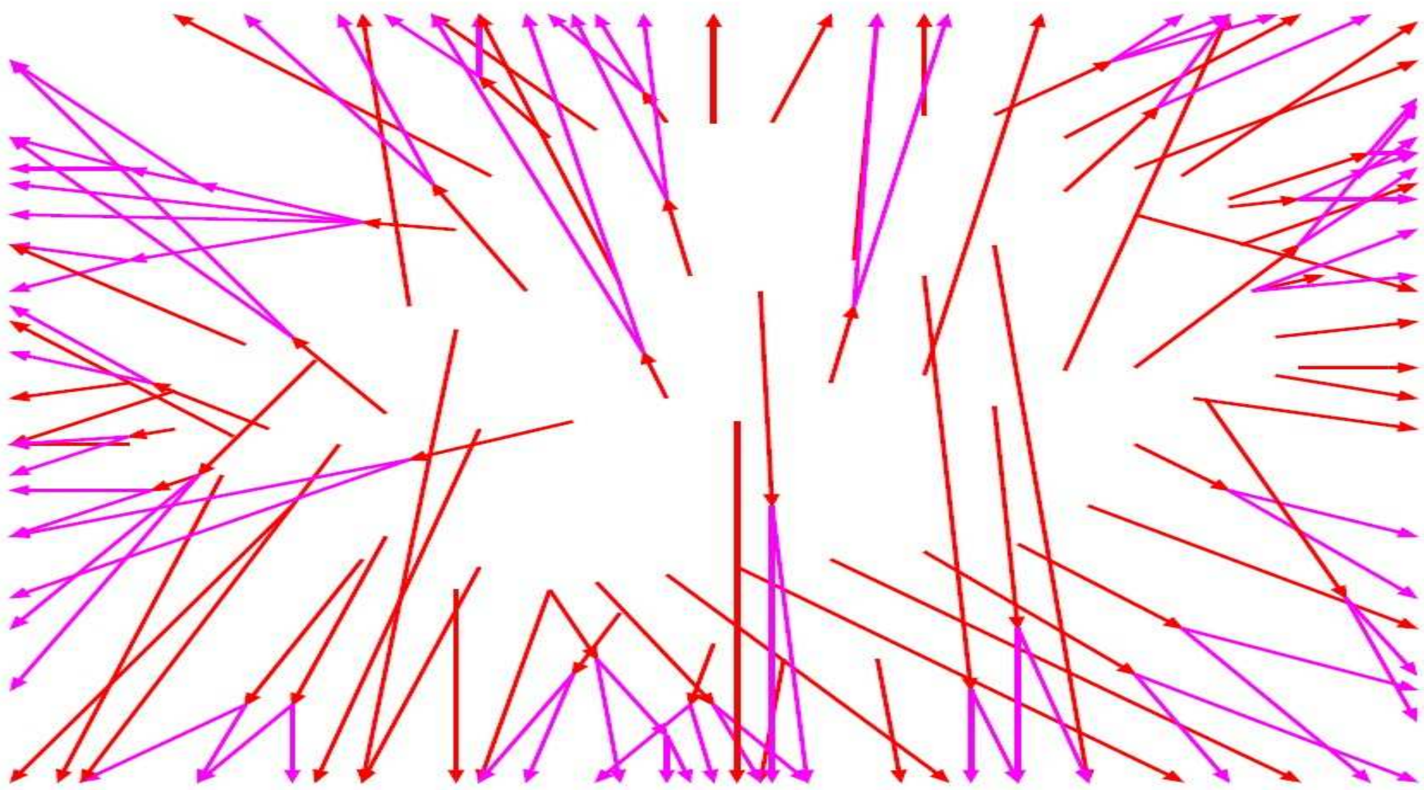
The strings fragments to produce hadrons (strings model)

Hadronisation is not really calculated from first principles, however it is modelled through various data and hence it is considered reliable

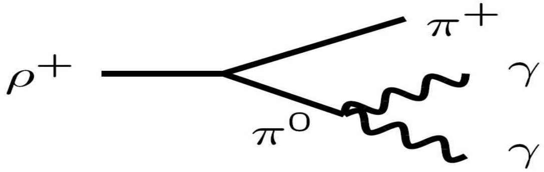
Movie: The structure of an event



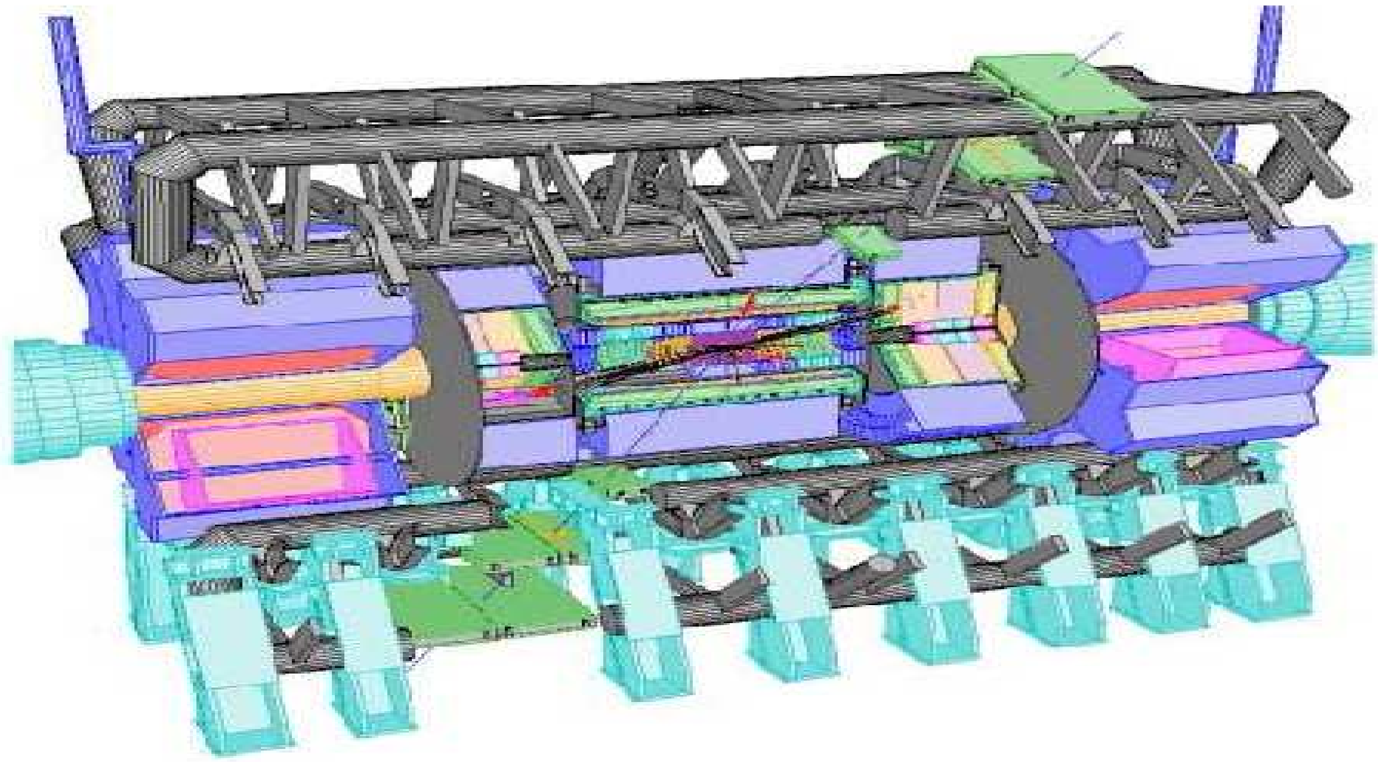
Hadrons decay



Hadrons decay



Movie: The structure of an event



These are the particles that hit the detector

- Parton Shower is well understood , perturbation theory with a few approximations
- Hadronisation is not really calculated from first principles, however it is modelled through various data and hence it is considered reliable
- UE is not understood. It is modelled through some data that do not give strong constraints and hence extrapolation can be doubted.(minimum bias??)
- Important to have a “clear” picture of the physical situation

- generate events with as much details as possible:
  W will decay.
  To τ?, τ will decay,
  there is no quark,
  only hadrons,...
  production comes with non negligible radiation
- $\sigma_{\text{final state}} = \sigma_{\text{hard process}} \, \mathcal{P}_{\text{tot}}$
- $\mathcal{P}_{\text{tot}} = \mathcal{P}_{\text{decay}} \mathcal{P}_{\text{ISR}} \mathcal{P}_{\text{FSR}} \mathcal{P}_{\text{remnants}} \mathcal{P}_{\text{hadronise}} \mathcal{P}_{\text{ord. dec.}}$
- **Divide and Conquer : each $\mathcal{P}_{\text{i}}$ handled in turn**
- an event with n particles involve about 10n random choices (flavour, mass, momentum,spin,...). At the LHC expect about 100 charged and 200 neutral particles, thus totalling a few thousand choices

MC is probabilistic, divide and conquer

- generate events with as much details as possible:

  W will decay.

  To τ?, τ will decay,

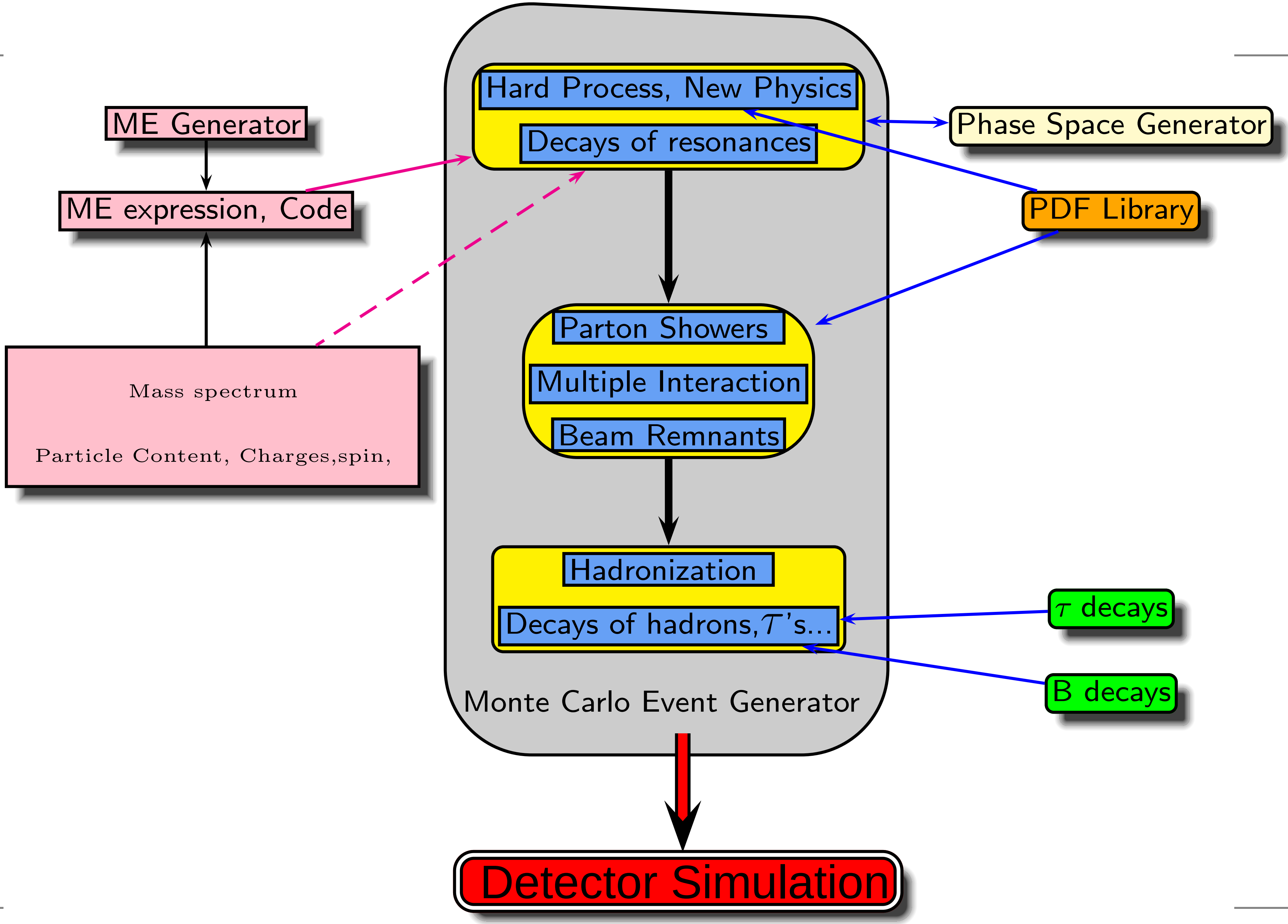
  there is no quark,

  only hadrons,...

  production comes with non negligible radiation
- $\sigma_{\rm final\ state} = \sigma_{\rm hard\ process}\ \mathcal{P}_{\rm tot}$
- $\mathcal{P}_{\rm tot} = \mathcal{P}_{\rm decay}\mathcal{P}_{\rm ISR}\mathcal{P}_{\rm FSR}\mathcal{P}_{\rm remnants}\mathcal{P}_{\rm hadronise}\mathcal{P}_{\rm ord.\ dec.}$
- **Divide and Conquer : each $\mathcal{P}_{\rm i}$ handled in turn**
- an event with n particles involve about 10n random choices (flavour, mass, momentum,spin,...). At the LHC expect about 100 charged and 200 neutral particles, thus totalling a few thousand choices

Divide and Conquer : each $\mathcal{P}_{\rm i}$ handled in turn $\rightarrow$ Modular Structure

Putting all together



MEG vs General purpose MC EG

ME Generators

Hard Process and Decay

Many

....

(see later)

General Purpose MC EG

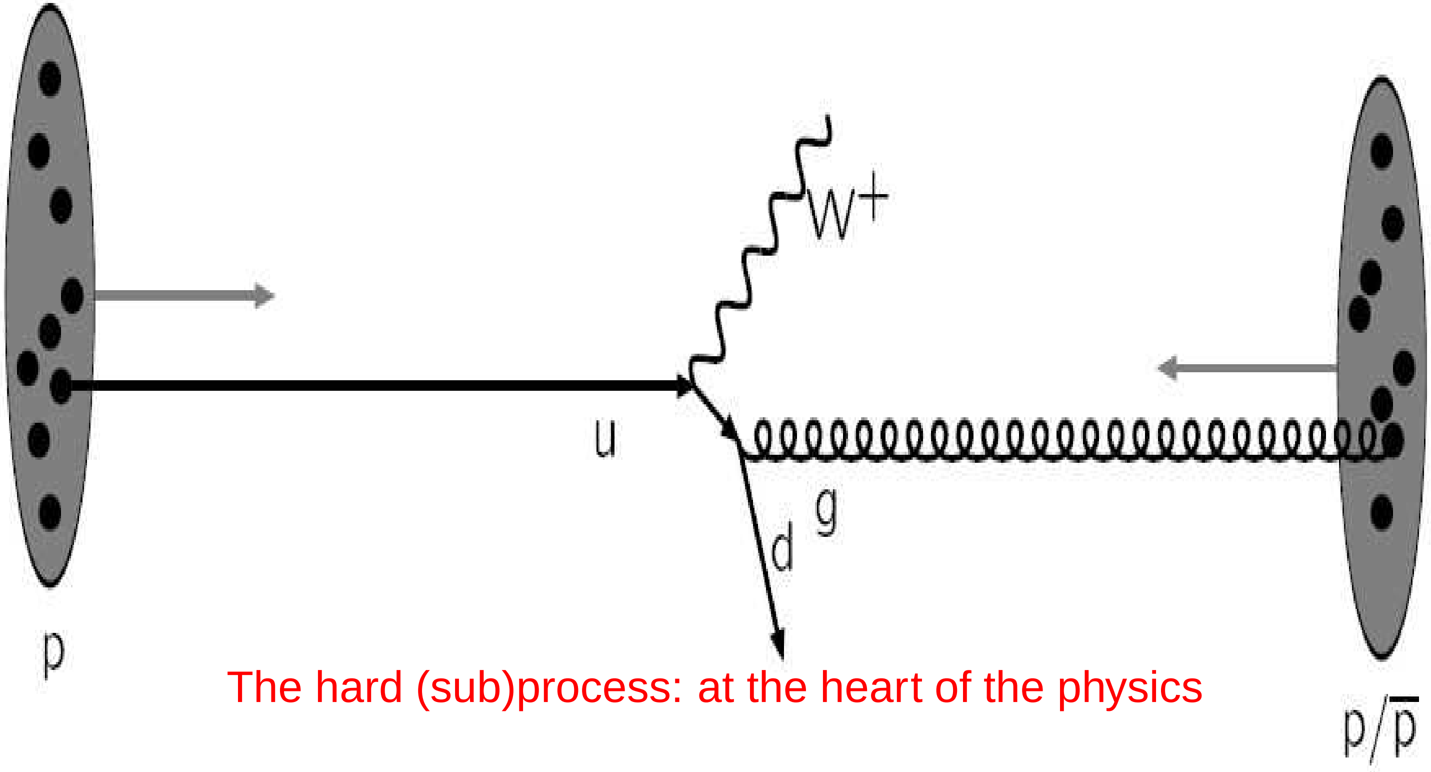
HERWIG++

PYTHIA

SHERPA

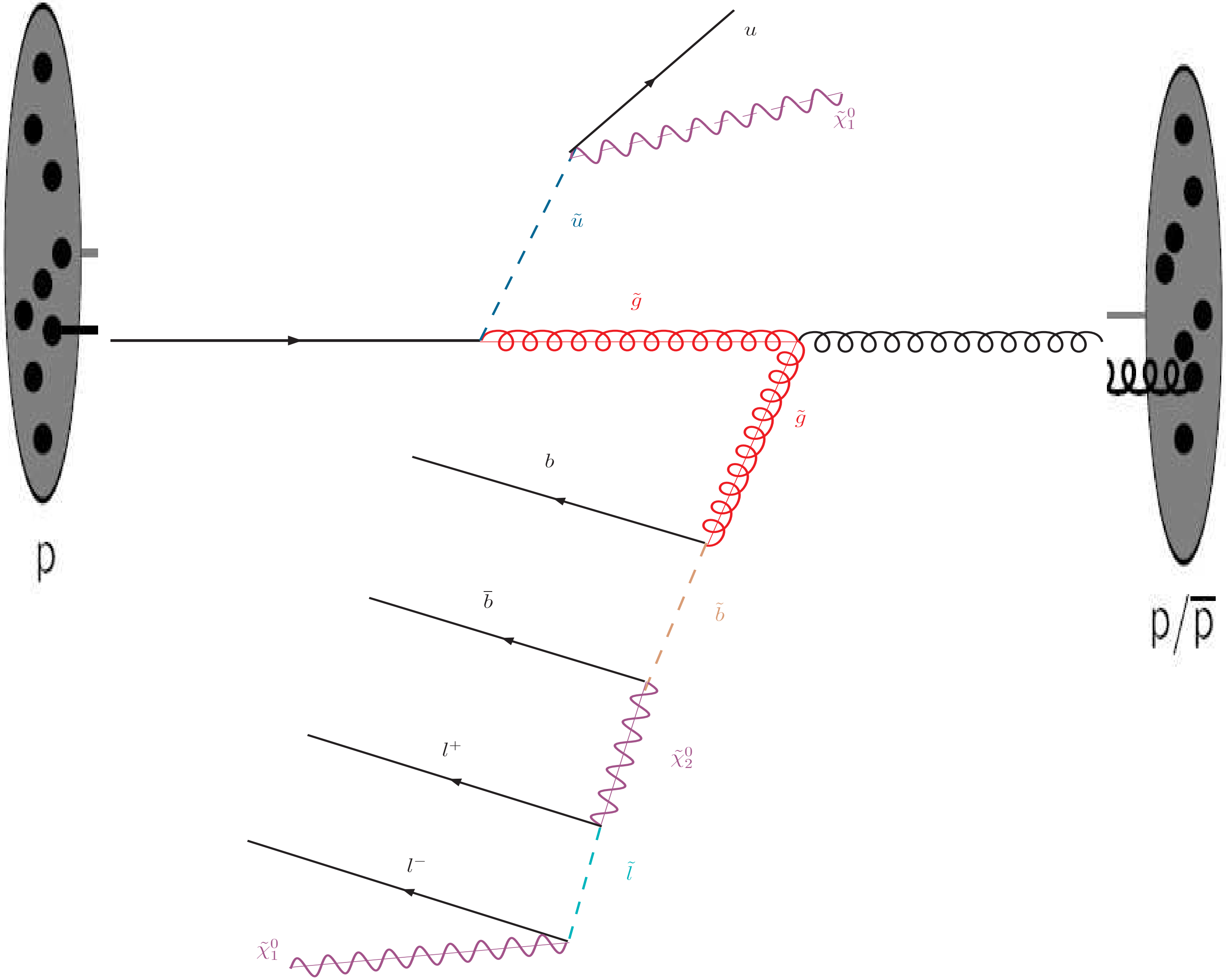
(ISAJET)

Integration: PDF and Cross sections

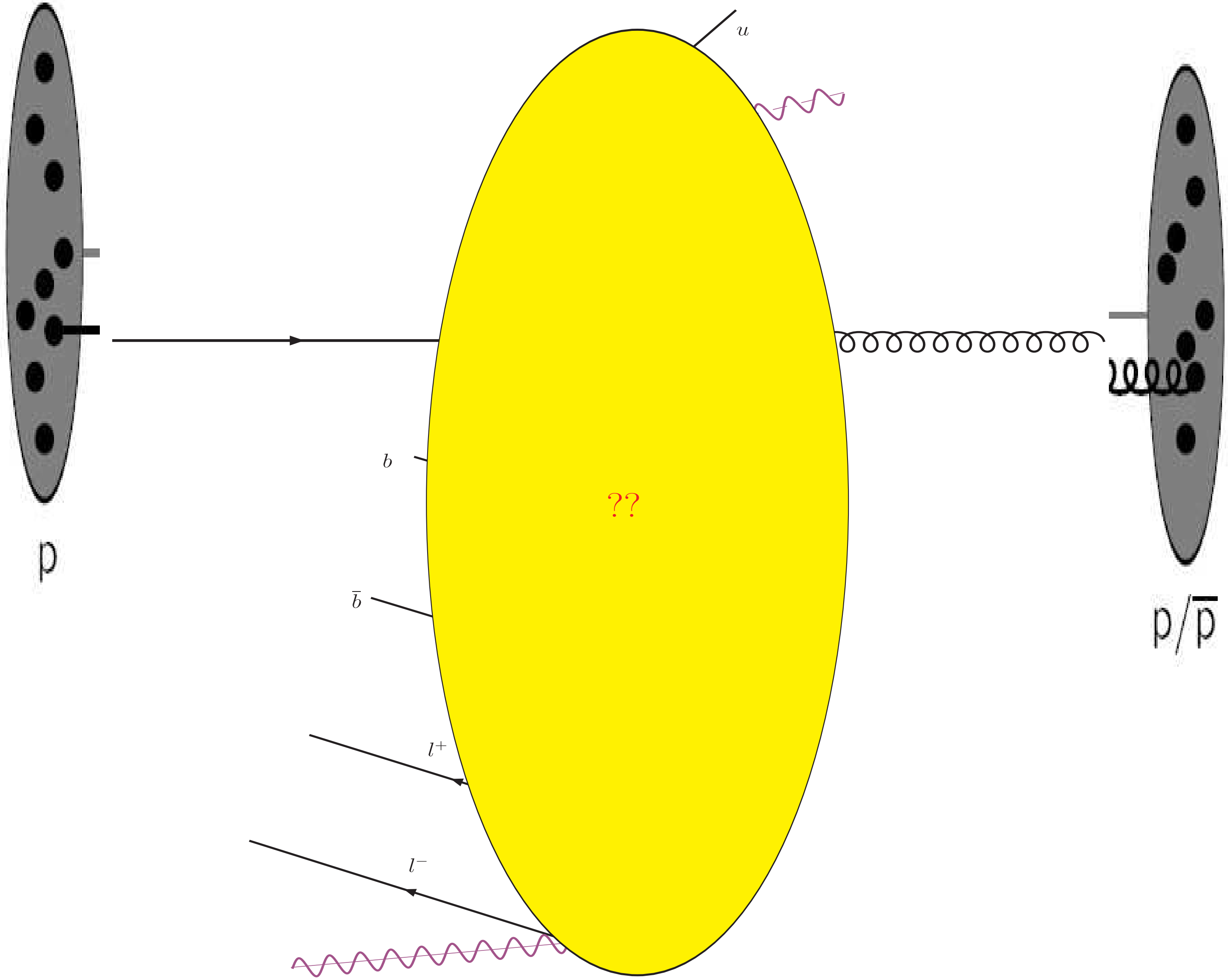


The hard (sub)process: at the heart of the physics

Integration: PDF and Cross sections



Integration: PDF and Cross sections



Factorisation and Parton Distribution Functions

$$\sigma_{pp\to X} = \sum_{a,b=q,\bar{q},g} \int \mathrm{d}x_1 \mathrm{d}x_2 f_a(x_1,\mu^2) f_b(x_2,\mu^2) \hat{\sigma}_{ab\to X}(\hat{s},\mu^2)$$

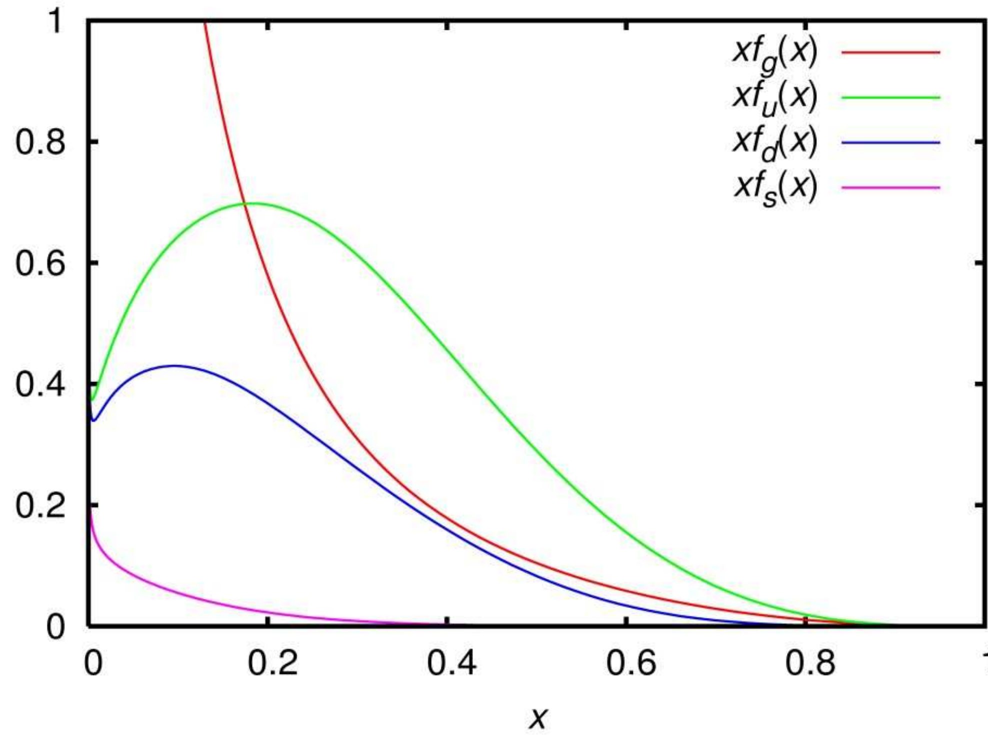
$f_i(x,\mu^2)$ is the **P**arton **D**istributions **F**unction

μ^2 is the factorisation scale !

Many libraries exist (CTEQ, MRSx)

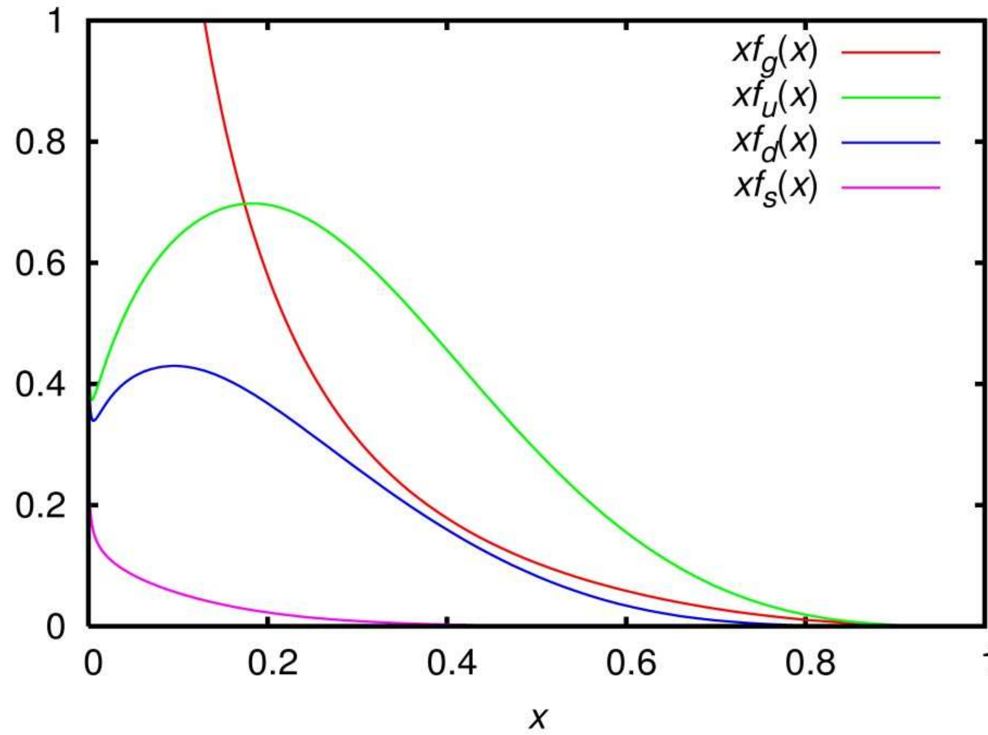
reliable in the range

$10^{-3} < x < 0.8 \;\; (2\mathrm{GeV})^2 < \mu^2 < (1TeV)^2$



$$\sigma_{pp\to X} = \sum_{a,b=q,\bar{q},g} \int \mathrm{d}x_1 \mathrm{d}x_2 f_a(x_1,\mu^2) f_b(x_2,\mu^2) \hat{\sigma}_{ab\to X}(\hat{s},\mu^2)$$

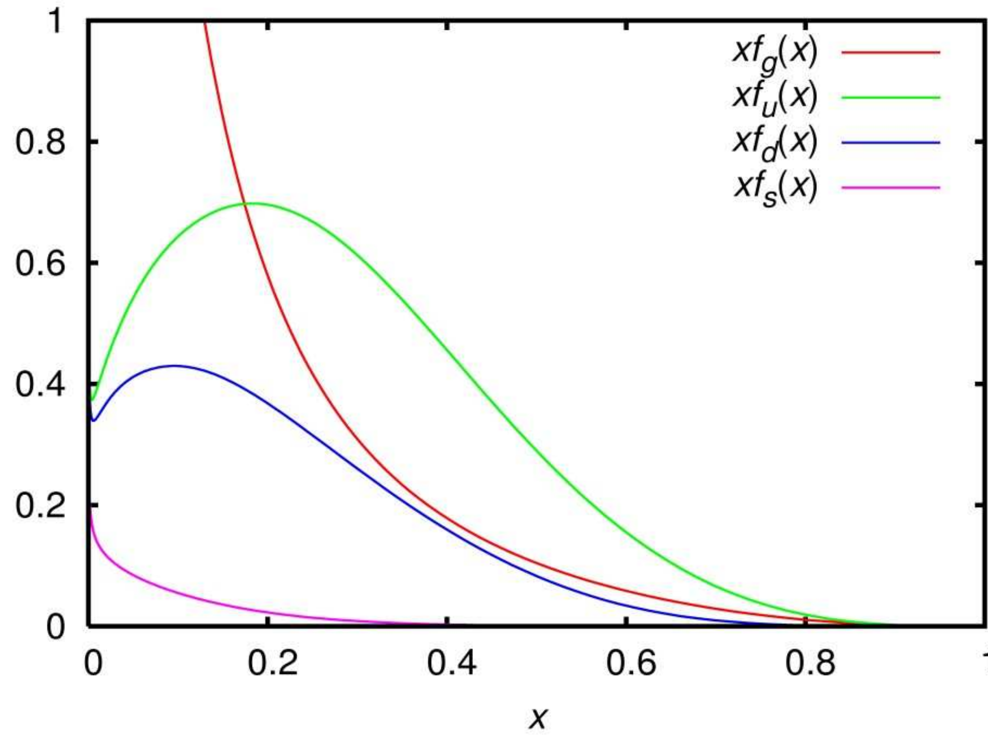
$f_i(x,\mu^2)$ is the **P**arton **D**istributions **F**unction
μ^2 is the factorisation scale !
Many libraries exist (CTEQ, MRSx)
reliable in the range
$10^{-3} < x < 0.8 \;\; (2\mathrm{GeV})^2 < \mu^2 < (1TeV)^2$



Phase Space

$$\hat{\sigma}_{ab\to X} = \frac{1}{2\hat{s}} \sum_{spin,..} \int_{\Phi_N} |\mathcal{M}|^2 \mathrm{d}\Phi_N$$

$$\sigma_{pp\to X} = \sum_{a,b=q,\bar{q},g} \int \mathrm{d}x_1 \mathrm{d}x_2 f_a(x_1,\mu^2) f_b(x_2,\mu^2) \hat{\sigma}_{ab\to X}(\hat{s},\mu^2)$$



$f_i(x,\mu^2)$ is the **P**arton **D**istributions **F**unction
μ^2 is the factorisation scale !
Many libraries exist (CTEQ, MRSx)
reliable in the range
$10^{-3} < x < 0.8 \quad (2\mathrm{GeV})^2 < \mu^2 < (1TeV)^2$

Phase Space

$$\hat{\sigma}_{ab\to X} = \frac{1}{2\hat{s}} \sum_{spin,..} \int_{\Phi_N} |\mathcal{M}|^2 \mathrm{d}\Phi_N$$

Integrals $\longrightarrow \int$

At the heart of the ME is the hard process, that is where the physics lies and that is what gives the probability of a particular event

For the hard process

- amplitude $\mathcal{M} \longrightarrow |\mathcal{M}|^2$
- $N_{\rm evt,cuts} \propto \int d\sigma = \int |\mathcal{M}|^2 {\rm d}\Phi(n)$
- Integration over a phase space with of large number n of dimensions, each particle $\rightarrow 3$ variables (momenta)
- $Dim[d\Phi(n)] \sim 3n$

$$
{\rm d}\Phi(n) = \left(\prod_i n \frac{{\rm d}^2 p_i}{(2\pi)^3 (2E_i)} \right) (2\pi)^4 \delta \left(P_{in} - \sum_i^n p_i \right)
$$

Monte-Carlo Definition

- MC is a numerical method for calculating/estimating an integral based on a random evaluation of the integrand
- Particularly useful because one deals with a large number of (integration) variables (momenta of particles)
- Limits of integration (cuts) are often complicated
- Integrand is a convolution of different functions

$$I = \int_{x_1}^{x_2} f(x)dx = (x_2 - x_1) < f(x) > \quad \text{(usually } \; x_1 = 0, x_2 = 1)$$

The average can be calculated by selecting N values *randomly* $x_i, i = 1, \cdots N$ from *uniform distribution*, calculate $f(x_i)$

$$I = I_N = \frac{1}{N}(x_2 - x_1) \sum_{i=1}^{i=N} f(x_i) = \frac{1}{N} \sum_{i=1}^{i=N} W(x_i) \quad W(x_i) = \text{weight}$$

- Sum is invariant under reordering (randomize)
- Obviously approximation better if number of points N is larger
- Error given by the Central Limit Theorem

One dimension, example. The variance

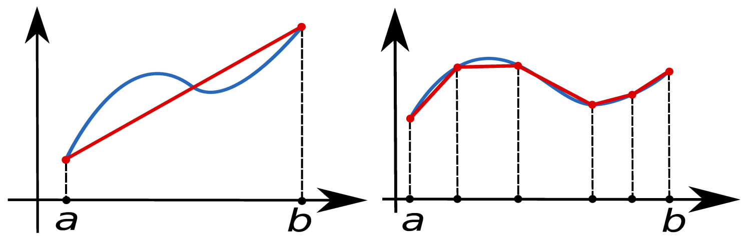
$$I \simeq I_N \pm \sqrt{\frac{V_N}{N}} = I_N \pm \sqrt{\frac{< f^2 > - < f >^2}{N}}$$

- MC converges as $1/\sqrt{N}$

One dimension, example. The variance

$$I \simeq I_N \pm \sqrt{\frac{V_N}{N}} = I_N \pm \sqrt{\frac{< f^2 > - < f >^2}{N}}$$

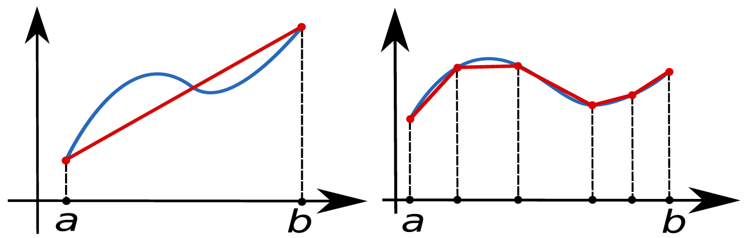
- MC converges as $1/\sqrt{N}$
- compare to trapezium rule convergence $\propto 1/N^2$ (if derivative exists)



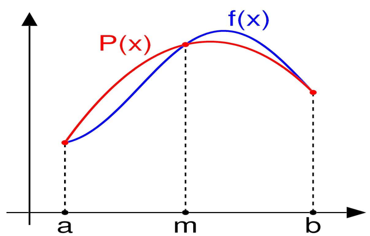
One dimension, example. The variance

$$I \simeq I_N \pm \sqrt{\frac{V_N}{N}} = I_N \pm \sqrt{\frac{< f^2 > - < f >^2}{N}}$$

- MC converges as $1/\sqrt{N}$
- compare to trapezium rule convergence $\propto 1/N^2$ (if derivative exists)



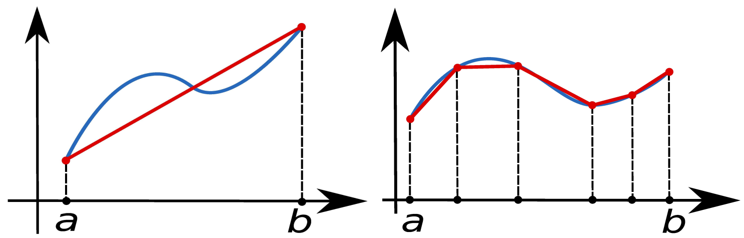
- Simpson (quadratic interpolation) $\propto 1/N^4$ (if derivative exists)



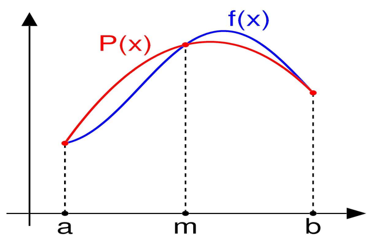
One dimension, example. The variance

$$I \simeq I_N \pm \sqrt{\frac{V_N}{N}} = I_N \pm \sqrt{\frac{< f^2 > - < f >^2}{N}}$$

- MC converges as $1/\sqrt{N}$
- compare to trapezium rule convergence $\propto 1/N^2$ (if derivative exists)



- Simpson (quadratic interpolation) $\propto 1/N^4$ (if derivative exists)



- but this is only in one dimension!

- Convergence may seem slow $\sqrt{1/N}$, but it can be estimated easily
- MC error does not depend on $\#$ of dimensions, d, $\propto 1/\sqrt{N}$

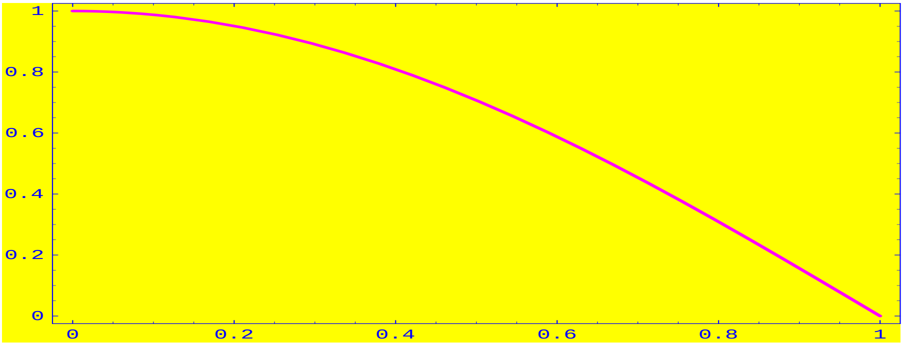
  Trapeze $\propto \longrightarrow 1/N^{2/d}$

  Simpson $\propto \longrightarrow 1/N^{4/d}$
- in MC one can improve convergence by minimising V_N while keeping the same number of points N
- **Importance Sampling**: non uniform sampling more efficient
- Convergence improved by putting more samples in regions where function is largest (where variance is largest)
- Hint: observe that if $f(x) = cste$ then $V_N = 0 \longrightarrow$ **make f as a close to a constant as possible!**

$$I \simeq I_N \pm \sqrt{\frac{V_N}{N}} = I_N \pm \sqrt{\frac{< f^2 > - < f >^2}{N}}$$

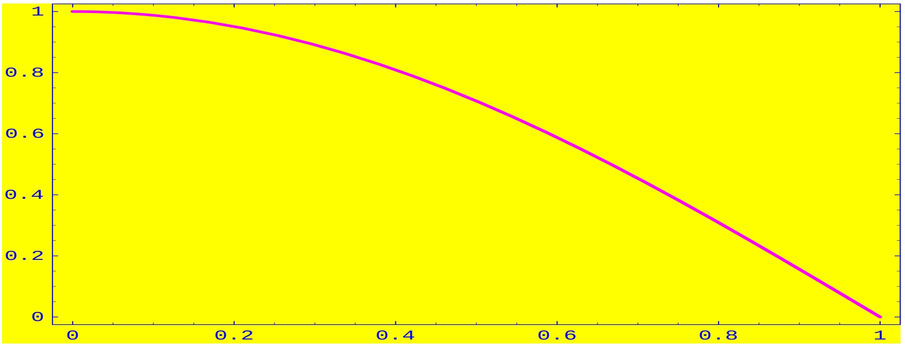
Example: Importance Sampling

Take $f(x) = cos \pi x/2$ then
$I = 2/\pi = 0.637$
MC, $I_N = 0.637 \pm 0.308/\sqrt{N}$
$(0.308 = \sqrt{V_N} = \sqrt{1/2 - (2/\pi)^2})$

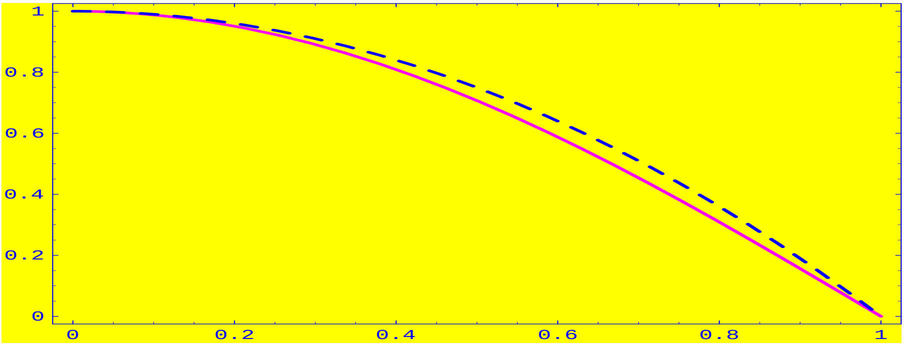


Example: Importance Sampling

Take $f(x) = cos\pi x/2$ then
$I = 2/\pi = 0.637$
MC, $I_N = 0.637 \pm 0.308/\sqrt{N}$
$(0.308 = \sqrt{V_N} = \sqrt{1/2 - (2/\pi)^2})$

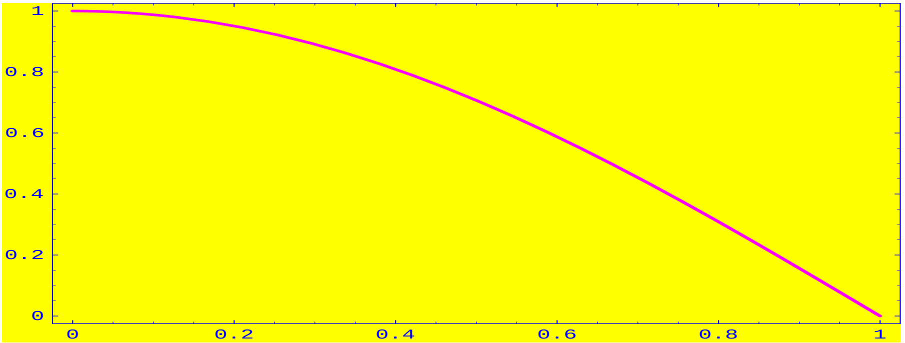


$$I = \int_0^1 dx(1-x^2)\frac{cos\pi x/2}{1-x^2}$$
$$= \int_{y_1}^{y_2} dy \frac{cos\pi x[y]/2}{1-x[y]^2}$$

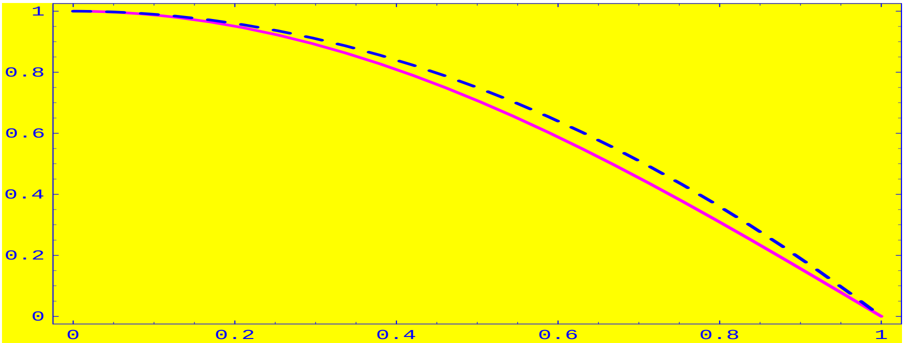


MC, $I_N = 0.637 \pm 0.031/\sqrt{N}$

Take $f(x) = cos\pi x/2$ then
$I = 2/\pi = 0.637$
MC, $I_N = 0.637 \pm 0.308/\sqrt{N}$
$(0.308 = \sqrt{V_N} = \sqrt{1/2 - (2/\pi)^2})$



$$I = \int_0^1 dx(1-x^2)\frac{cos\pi x/2}{1-x^2} = \int_{y_1}^{y_2} dy \frac{cos\pi x[y]/2}{1-x[y]^2}$$



MC, $I_N = 0.637 \pm 0.031/\sqrt{N}$

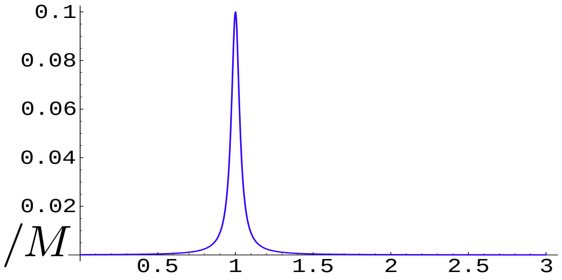
- For the same accuracy $N \to N/100$ events
- We have in fact made a change of variables
- Note however that change of variables may be not so trivial and requires that one knows the function, here is relatively ok
  $y = x - x^3/3$!

in HEP many sharp peaks from resonances, apart from peaks due to forward scattering,..

$$
\begin{aligned}
I &= \int_{m^2_{min}}^{m^2_{max}} \mathrm{d}m^2 \frac{1}{(m^2 - M^2)^2 + M^2\Gamma^2}, \quad \Gamma/M \ll 1 \\
&= \frac{1}{M^2} \int_{x_{min}}^{x_{max}} \mathrm{d}x \frac{1}{(x-1)^2 + \epsilon^2}, \quad x = m^2/M^2, \quad \epsilon = \Gamma/M
\end{aligned}
$$

in HEP many sharp peaks from resonances, apart from peaks due to forward scattering,..

$$
\begin{aligned}
I &= \int_{m^2_{min}}^{m^2_{max}} \mathrm{d}m^2 \frac{1}{(m^2 - M^2)^2 + M^2\Gamma^2}, \quad \Gamma/M \ll 1 \\
&= \frac{1}{M^2} \int_{x_{min}}^{x_{max}} \mathrm{d}x \frac{1}{(x-1)^2 + \epsilon^2}, \quad x = m^2/M^2, \quad \epsilon = \Gamma/M
\end{aligned}
$$



in HEP many sharp peaks from resonances, apart from peaks due to forward scattering,..

$$\begin{aligned} I &= \int_{m^2_{min}}^{m^2_{max}} \mathrm{d}m^2 \frac{1}{(m^2-M^2)^2+M^2\Gamma^2}, \quad \Gamma/M \ll 1 \\ &= \frac{1}{M^2}\int_{x_{min}}^{x_{max}} \mathrm{d}x \frac{1}{(x-1)^2+\epsilon^2}, \quad x = m^2/M^2, \quad \epsilon = \Gamma/M \end{aligned}$$

change of variable $x = \varepsilon \tan\theta + 1, dx = \varepsilon(1+\tan^2\theta)dt$

$$I = \frac{1}{M\Gamma}\int_{\theta_{min}}^{\theta_{max}} d\theta$$

The peak has been smoothed out completely.

Flat distribution, the error has been reduced to 0

in HEP many sharp peaks from resonances, apart from peaks due to forward scattering,..

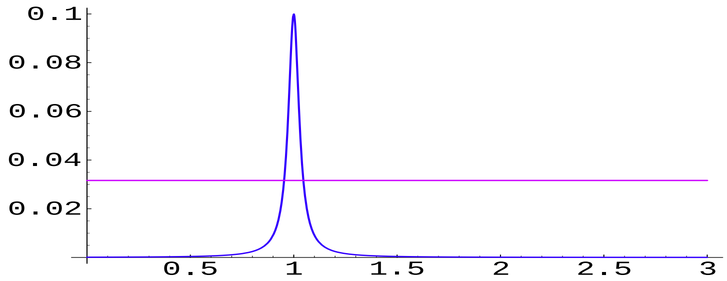
$$I = \int_{m^2_{min}}^{m^2_{max}} \mathrm{d}m^2 \frac{1}{(m^2 - M^2)^2 + M^2\Gamma^2}, \quad \Gamma/M \ll 1$$
$$= \frac{1}{M^2} \int_{x_{min}}^{x_{max}} \mathrm{d}x \frac{1}{(x-1)^2 + \epsilon^2}, \quad x = m^2/M^2, \quad \epsilon = \Gamma/M$$

change of variable $x = \varepsilon \tan\theta + 1, dx = \varepsilon(1 + \tan^2\theta)dt$

$$I = \frac{1}{M\Gamma} \int_{\theta_{min}}^{\theta_{max}} d\theta$$

The peak has been smoothed out completely.

Flat distribution, the error has been reduced to 0



Unfortunately we can not always do the Jacobian trick efficiently, we do not always know $f(x)$

However, as we have seen, finding a simple function, $p(x)$, that approximate $f(x)$ reduces the error drastically

(up to normalisation) take

$$p(x), \int_{x_1}^{x_2} p(x) = 1, \qquad \rightarrow I = \int_{x_1}^{x_2} dx p(x) \frac{f(x)}{p(x)}$$

$$I = \langle \frac{f}{p} \rangle \pm \frac{1}{\sqrt{N}} \sqrt{\langle \frac{f^2}{p^2} \rangle - \langle \frac{f}{p} \rangle^2}$$

Sample according to $p(x)$ and make f/p as small as possible.

Unfortunately we usually do not know much about $f(x)$

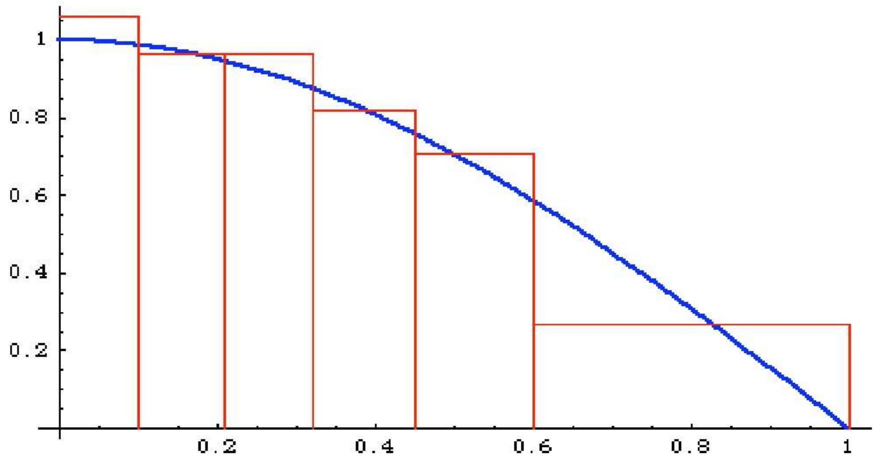
But as we sample we can know more, reconstruct $p(x)$ piecemeal, with step function

$$p(x) = \frac{1}{N_b}\Delta x_i \quad \text{for} \quad x_i - \Delta x_i \leq x \leq x_i$$

Unfortunately we usually do not know much about $f(x)$

But as we sample we can know more, reconstruct $p(x)$ piecemeal, with step function

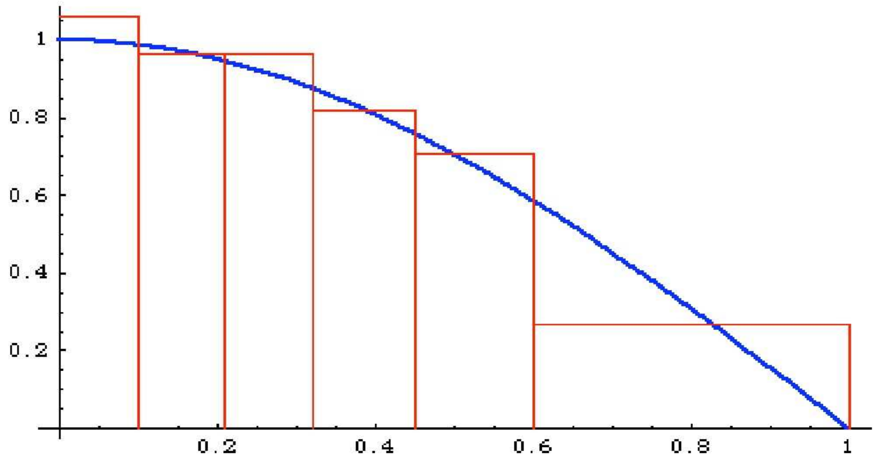
$$p(x) = \frac{1}{N_b}\Delta x_i \quad \text{for} \quad x_i - \Delta x_i \leq x \leq x_i$$



Unfortunately we usually do not know much about $f(x)$

But as we sample we can know more, reconstruct $p(x)$ piecemeal, with step function

$$p(x) = \frac{1}{N_b}\Delta x_i \quad \text{for} \quad x_i - \Delta x_i \leq x \leq x_i$$



- Improve the fit by generating more points where $f(x)$ is large, *i.e* where the variance is large
- Adjust the bin size so that each bin has the same area

Iterative algorithm: VEGAS

- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$

Many variables, VEGAS bis

- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$
- This assumes that we have the correct grid: the peaks are localised and are aligned along the axes!

- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$
- This assumes that we have the correct grid: the peaks are localised and are aligned along the axes!
- Not so obvious to pick up the correct choice of integration variables

Many variables, VEGAS bis

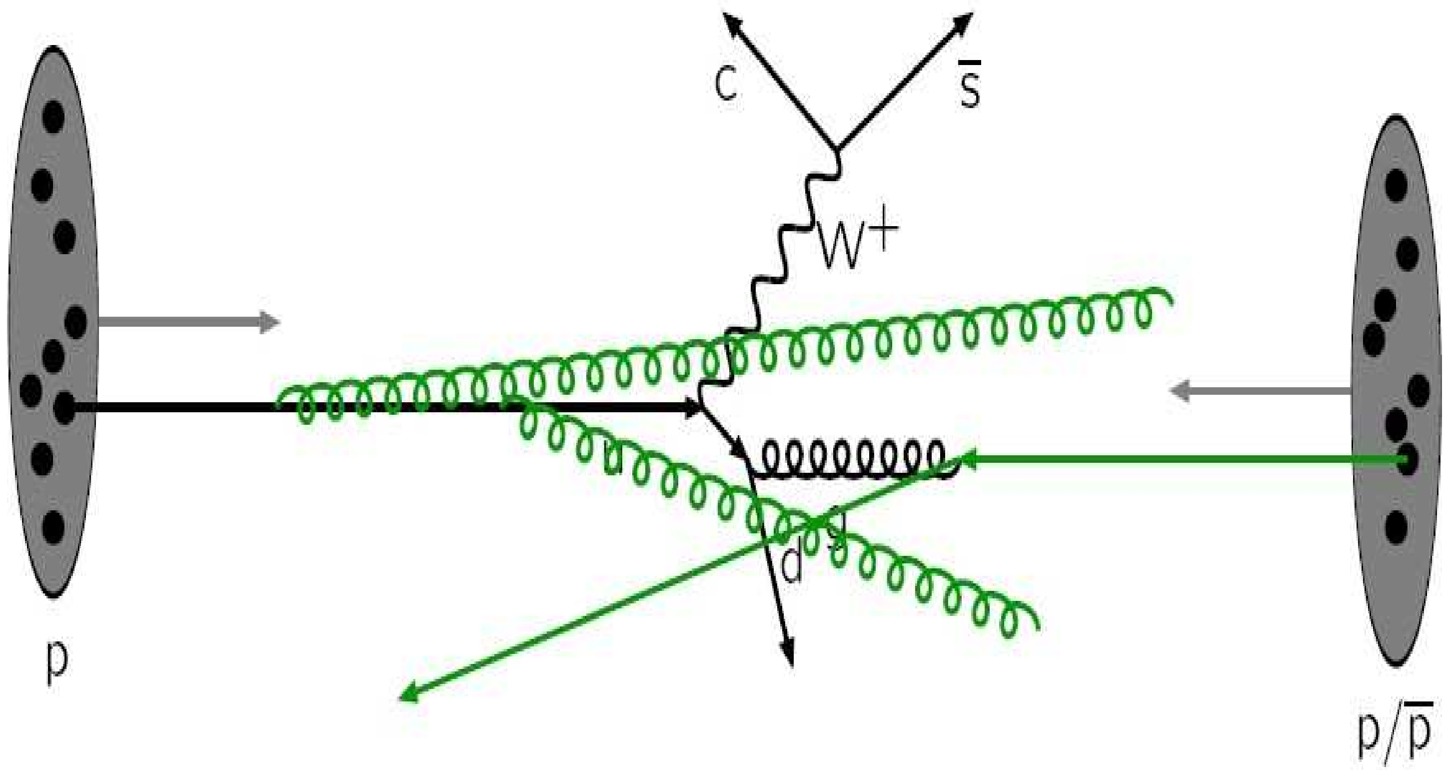
- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$
- This assumes that we have the correct grid: the peaks are localised and are aligned along the axes!
- Not so obvious to pick up the correct choice of integration variables
- dimension of phase space is ~ 3

Many variables, VEGAS bis

- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$
- This assumes that we have the correct grid: the peaks are localised and are aligned along the axes!
- Not so obvious to pick up the correct choice of integration variables
- dimension of phase space is ~ 3
- this means 2^n possible kinematical invariants

Many variables, VEGAS bis

- The approach can be directly generalised to d dimensions if one can write the factorised from $p(\vec{x}) = p(x) \times p(y) \times \cdots$
- This assumes that we have the correct grid: the peaks are localised and are aligned along the axes!
- Not so obvious to pick up the correct choice of integration variables
- dimension of phase space is ~ 3
- this means 2^n possible kinematical invariants
- A scattering amplitude may have many peaks each aligned on a different invariant

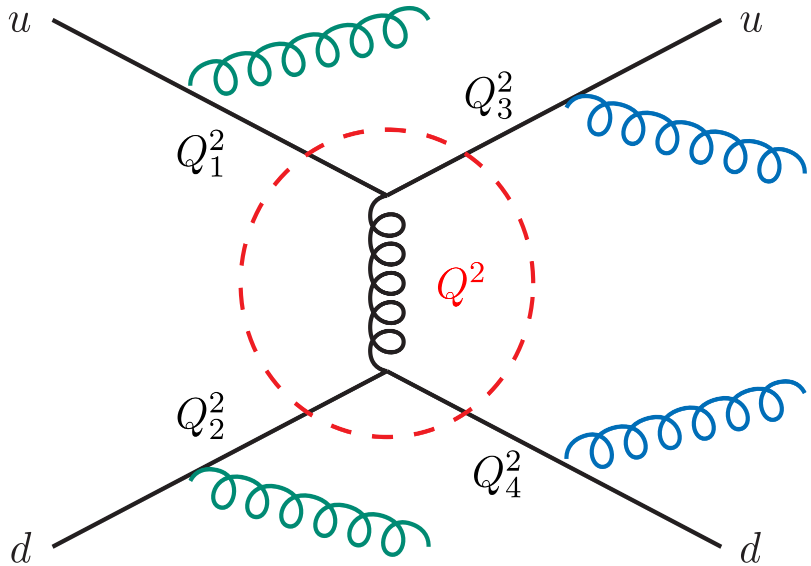


ISR: Initial State Radiation

Parton Shower Approach

$\mathcal{P}_{\mathrm{ISR/FSR}}$ Accelerated charged particles radiate

$2 \to n = (2 \to 2)_{\mathrm{On\ Shell}}$ + ISR + FSR

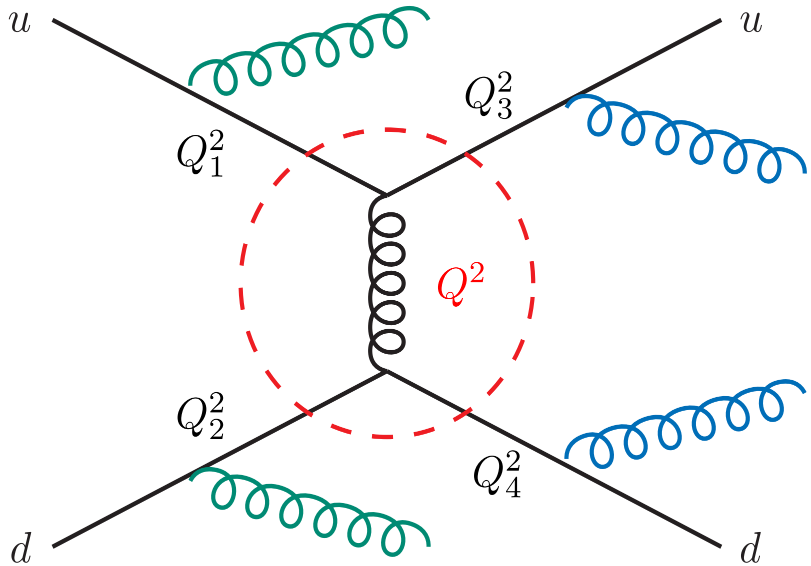


- FSR is time-like shower $Q_i^2 > 0$ decreasing, relatively simple
- ISR is space-like shower $Q_i^2 < 0$ increasing, physics complicated

Parton Shower Approach

$\mathcal{P}_{\mathrm{ISR/FSR}}$ Accelerated charged particles radiate

$2 \to n = (2 \to 2)_{\mathrm{On\ Shell}}$ + ISR + FSR



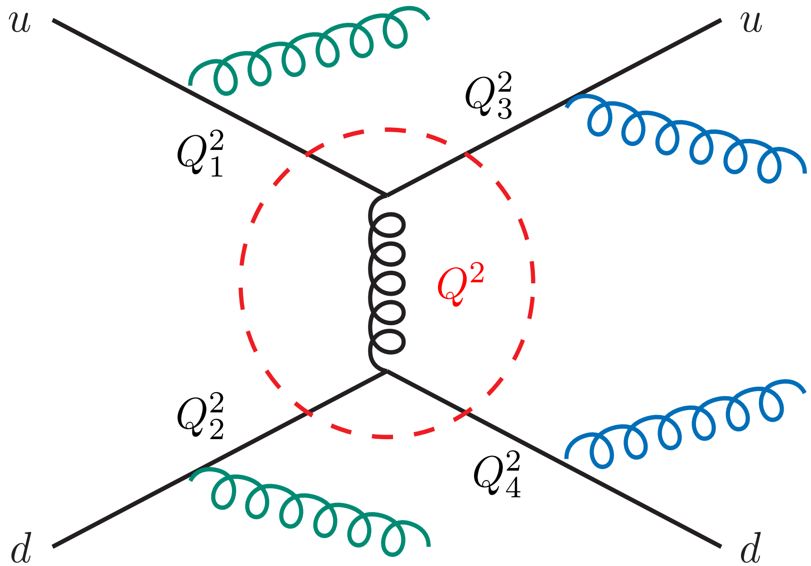
- FSR is time-like shower $Q_i^2 > 0$ decreasing, relatively simple
- ISR is space-like shower $Q_i^2 < 0$ increasing, physics complicated

- Shower is viewed as a probabilistic process which occurs with unit probability
- The (total) cross section is not affected but indirectly it is since the event shape is changed
- Obviously it is an approximation to the full process

Parton Shower Approach

$\mathcal{P}_{\mathrm{ISR/FSR}}$ Accelerated charged particles radiate

$2 \longrightarrow n = (2 \longrightarrow 2)_{\mathrm{On\ Shell}}$ + ISR + FSR

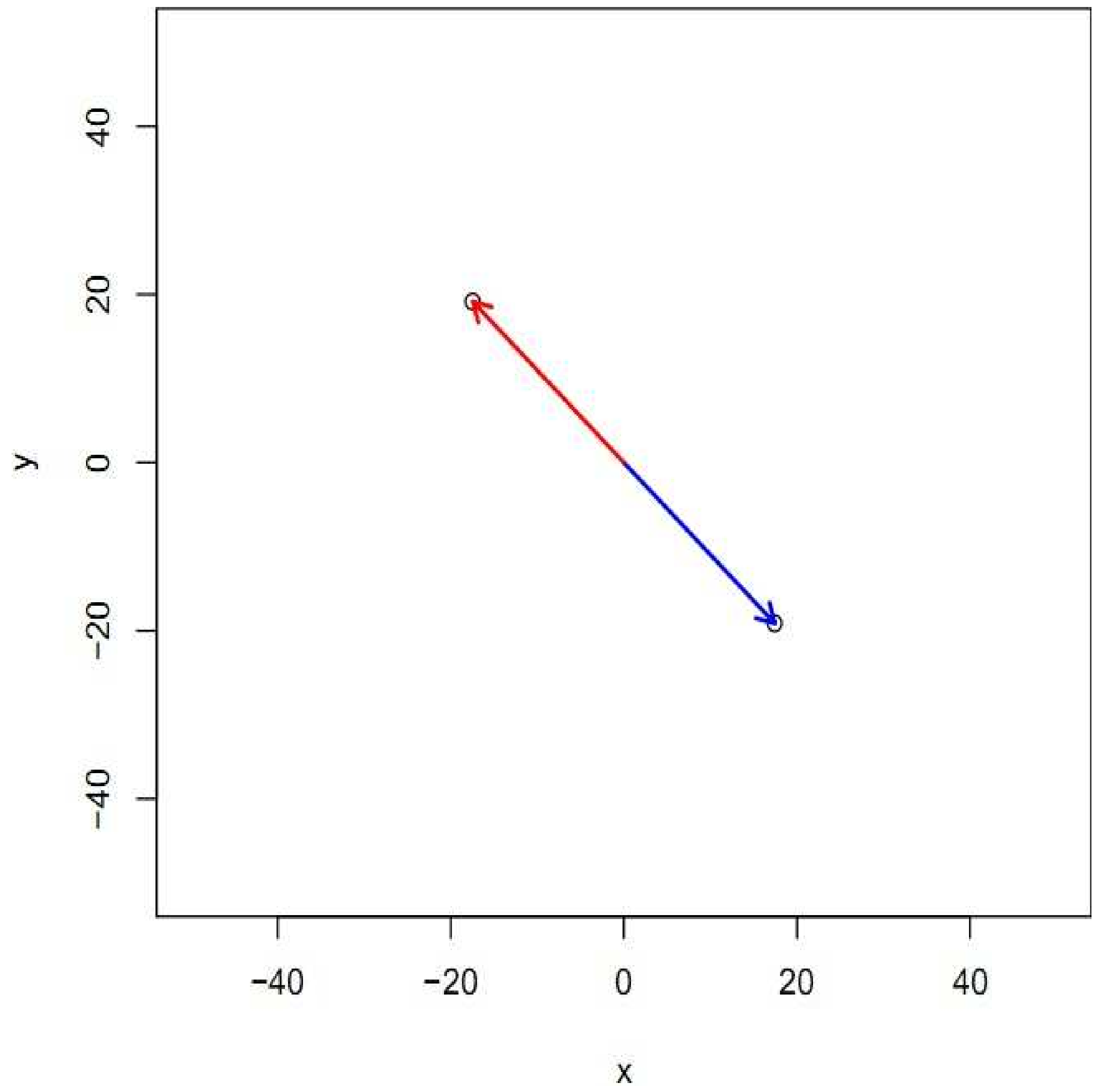


- FSR is time-like shower $Q_i^2 > 0$ decreasing, relatively simple
- ISR is space-like shower $Q_i^2 < 0$ increasing, physics complicated

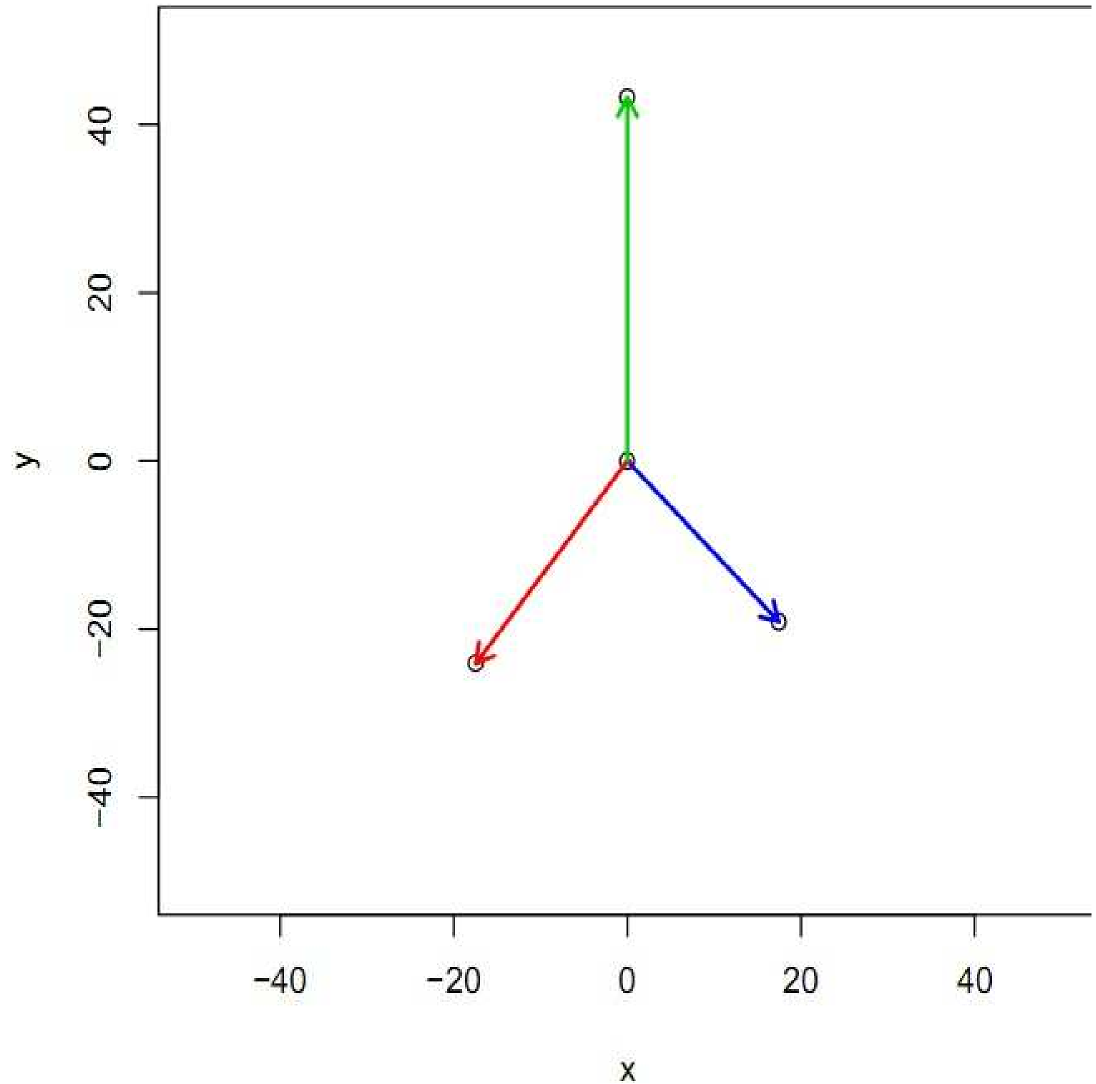
- Shower is viewed as a probabilistic process which occurs with unit probability
- The (total) cross section is not affected but indirectly it is since the event shape is changed
- Obviously it is an approximation to the full process

watch

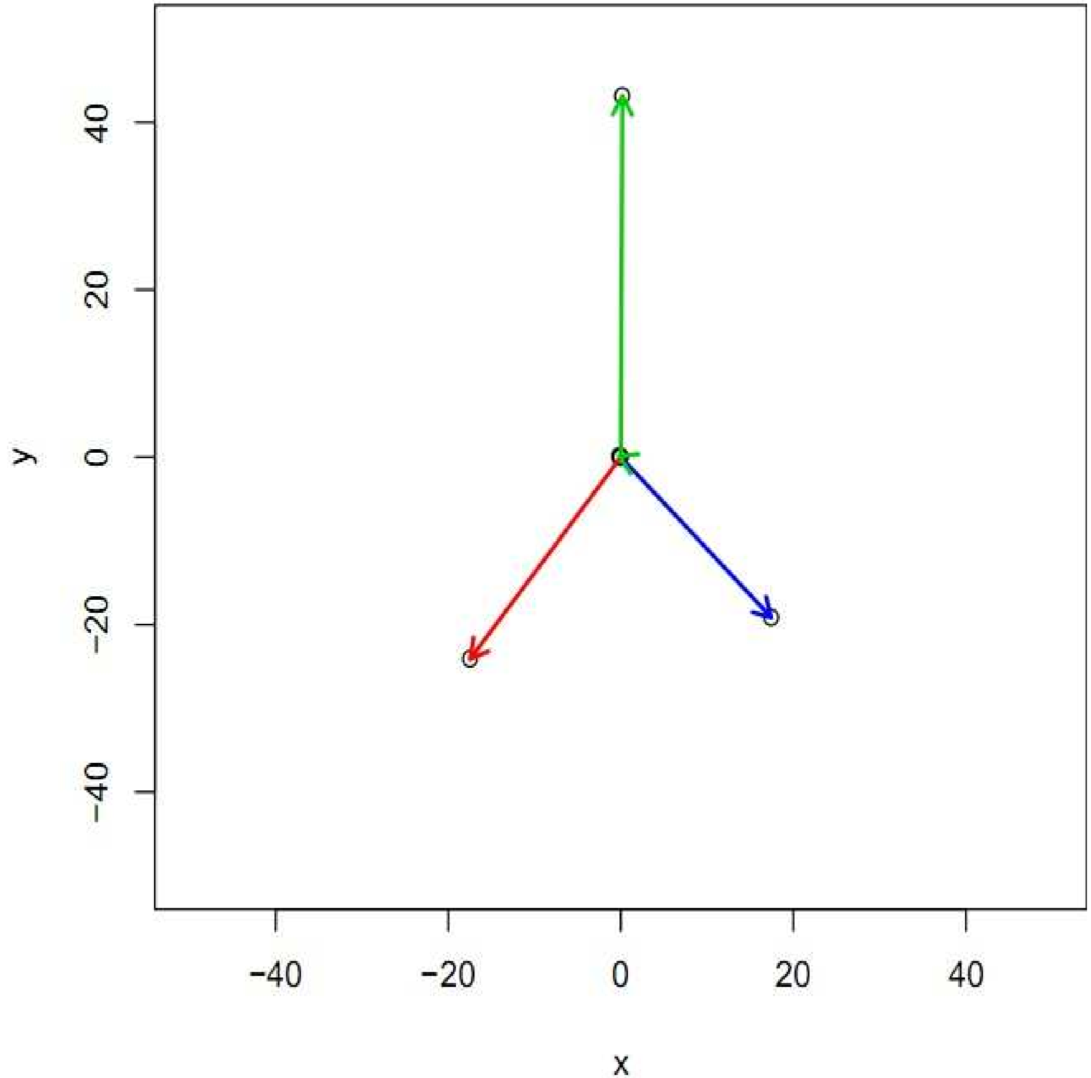
Parton Shower movie



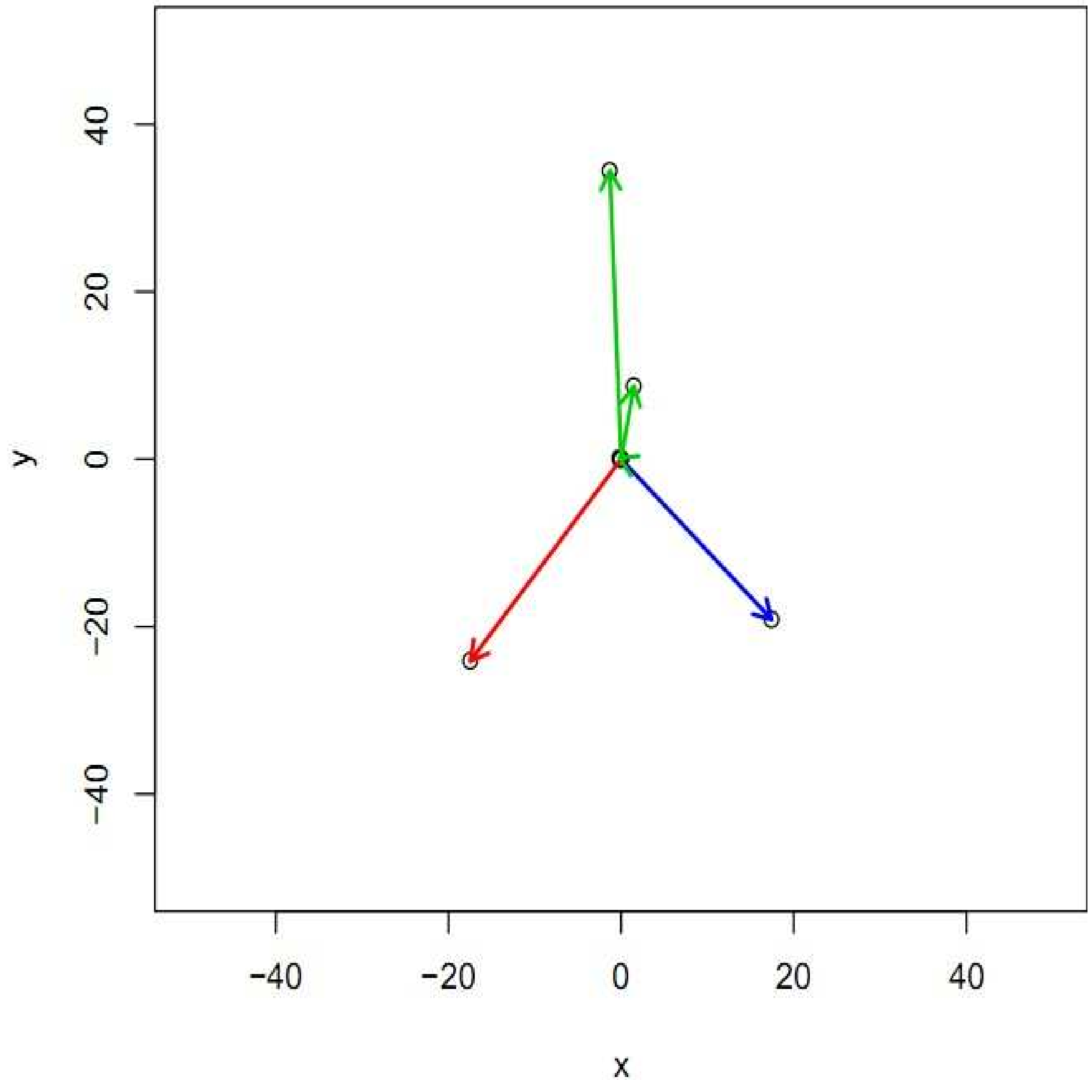
Parton Shower movie



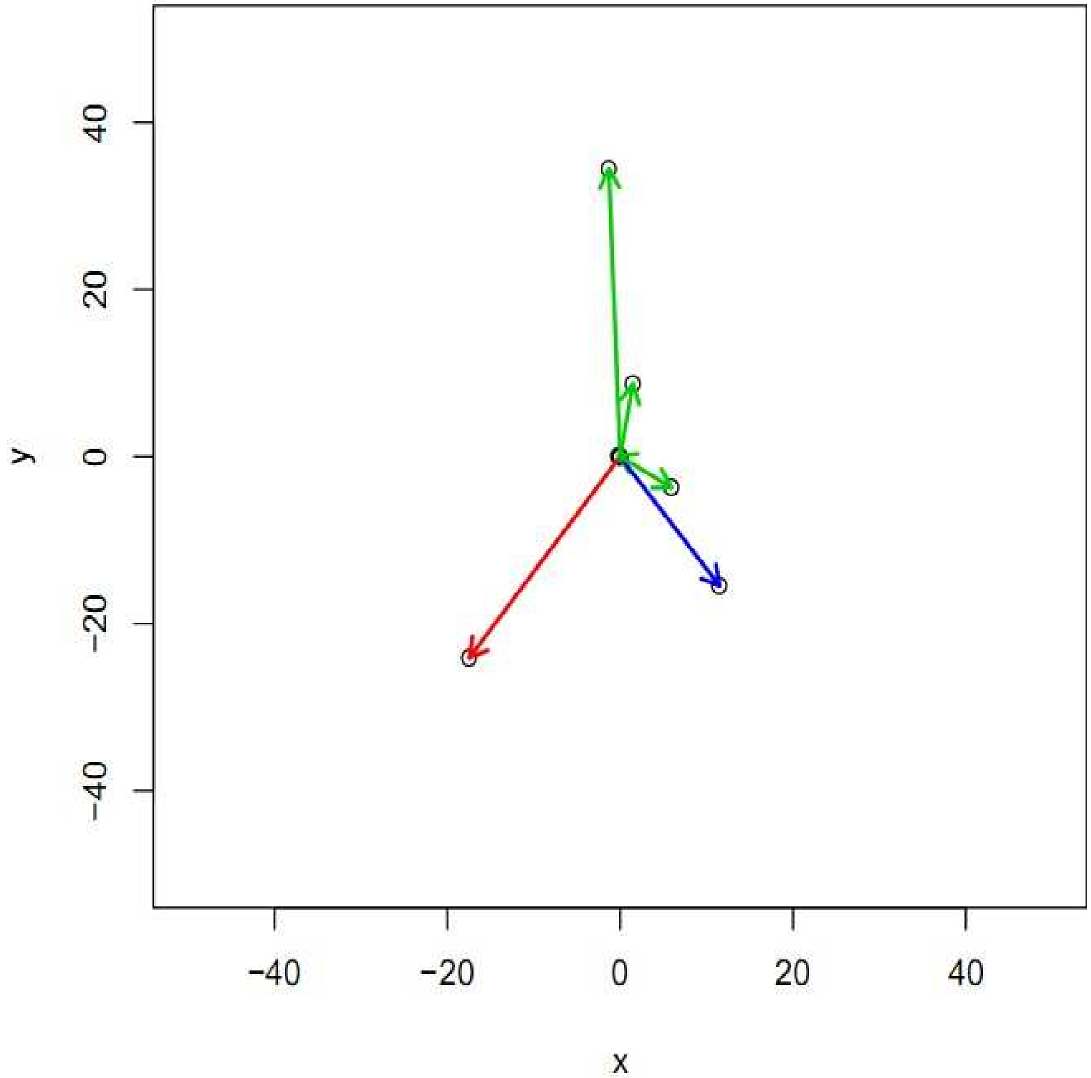
Parton Shower movie



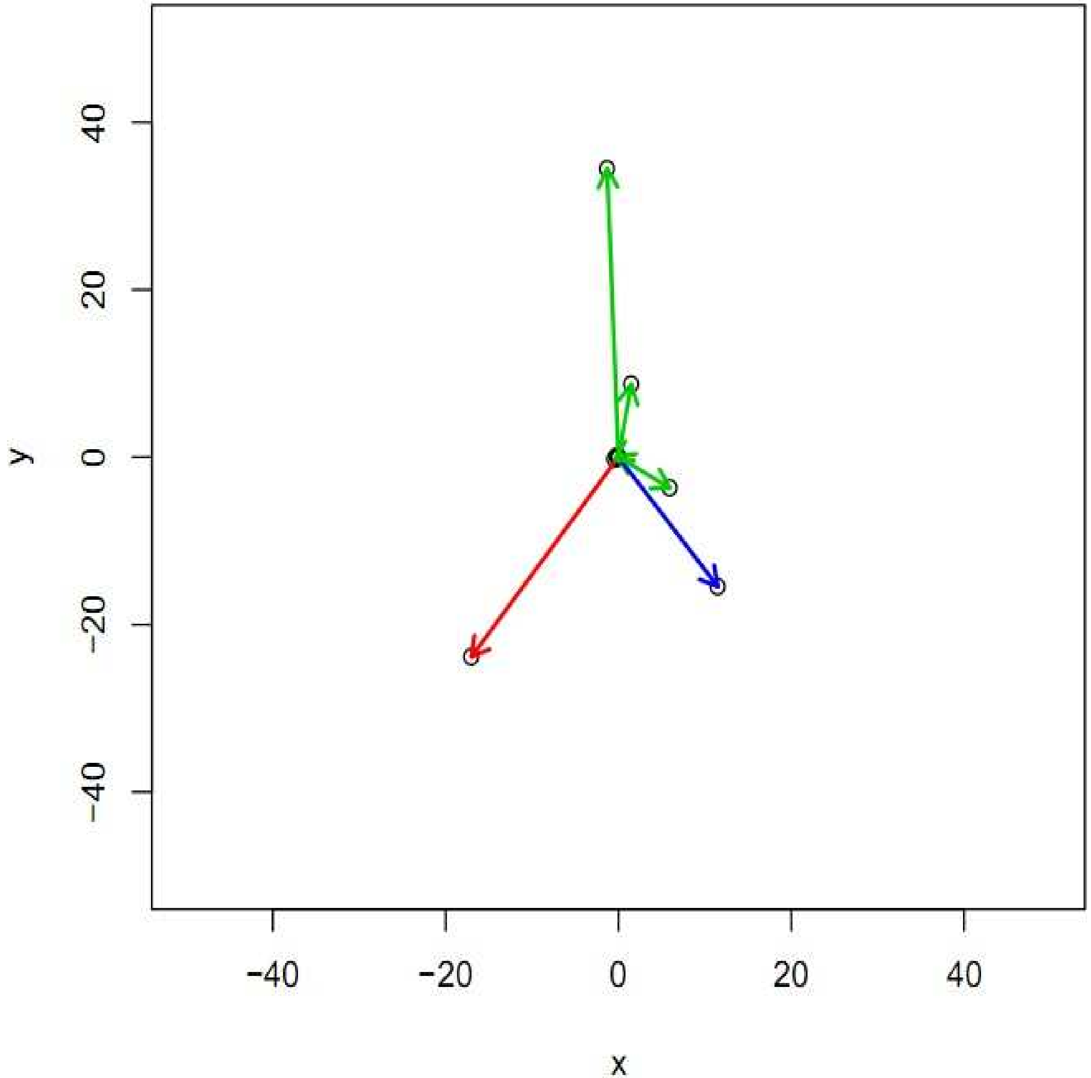
Parton Shower movie



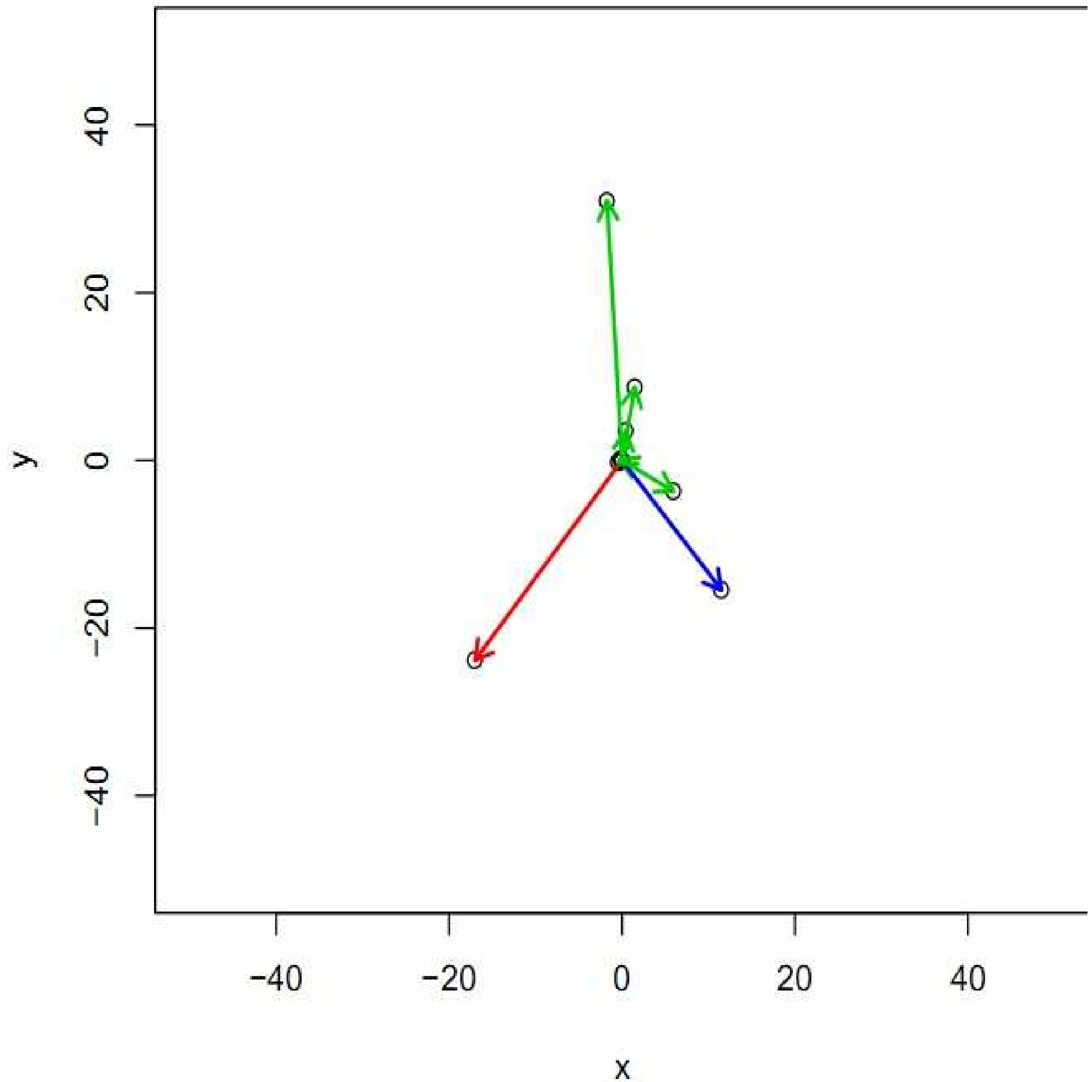
Parton Shower movie



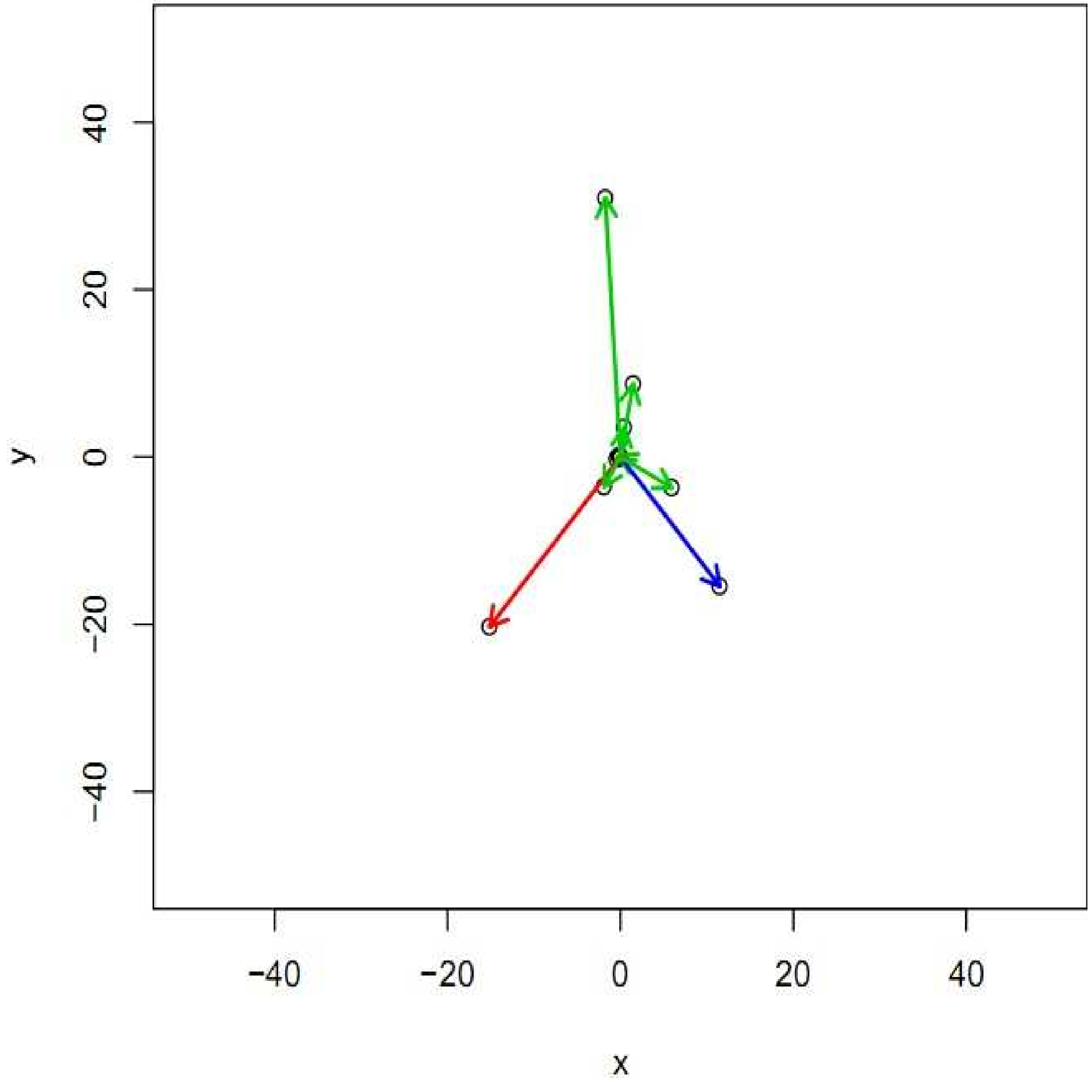
Parton Shower movie



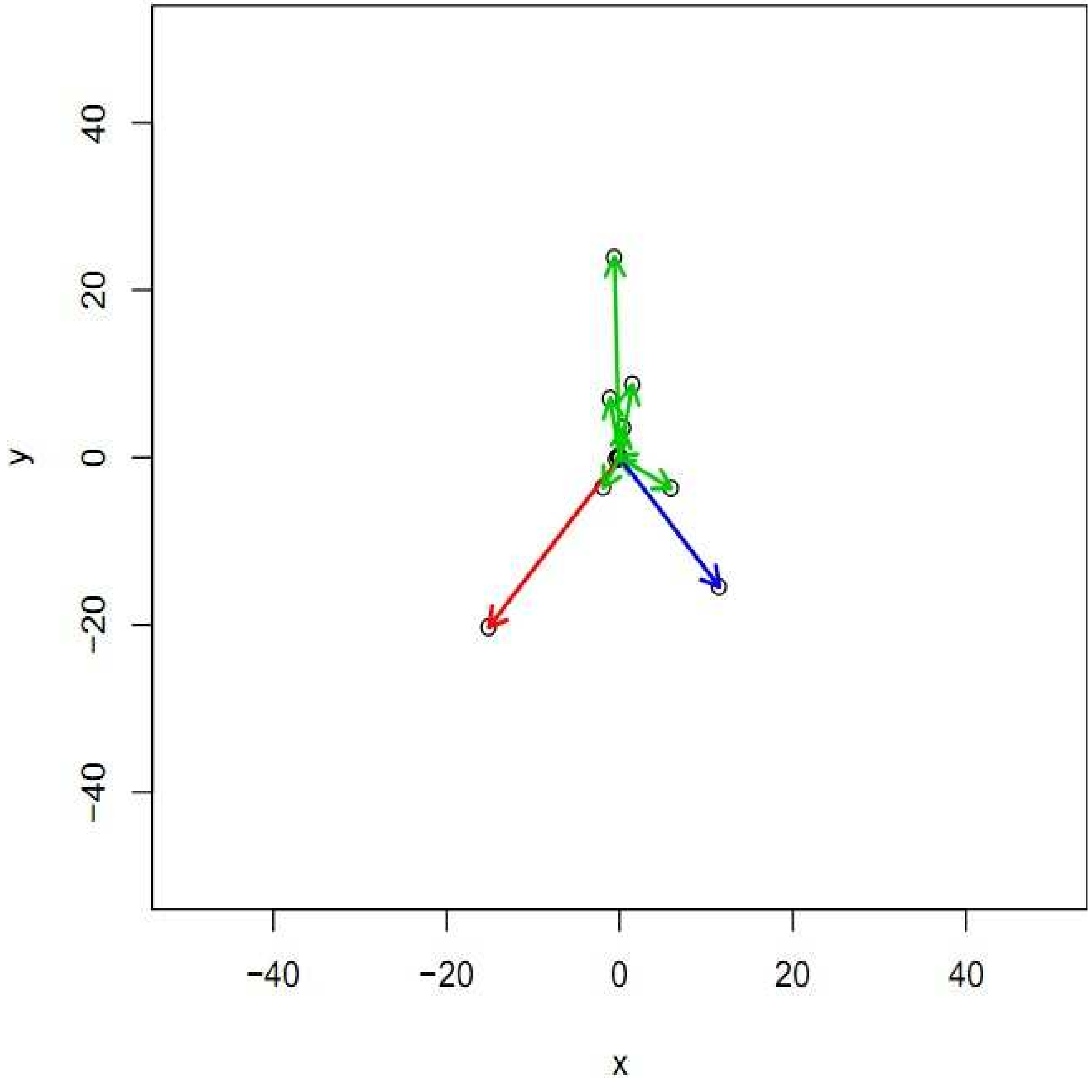
Parton Shower movie



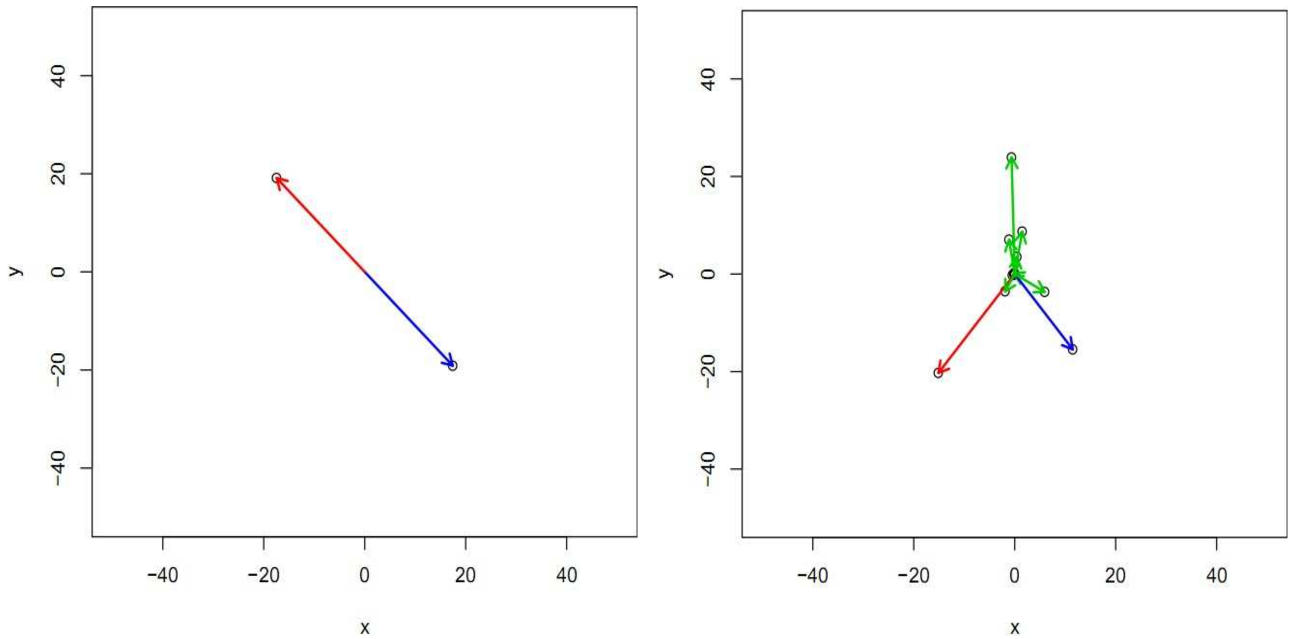
Parton Shower movie



Parton Shower movie

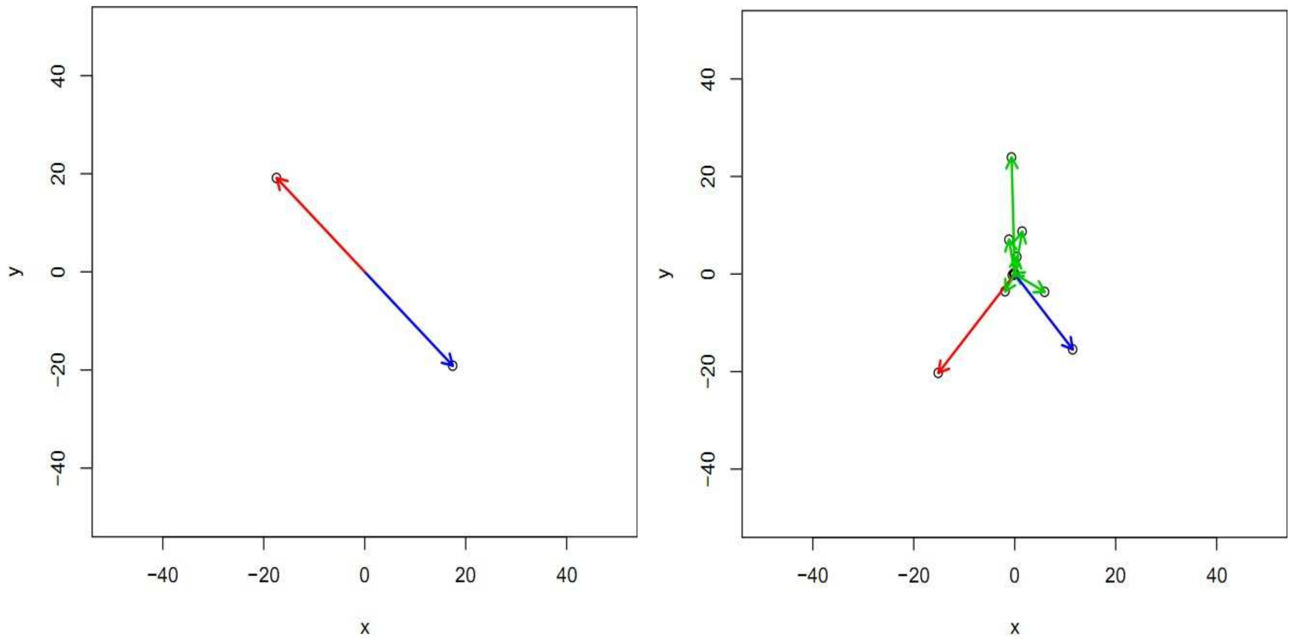


Parton Shower movie



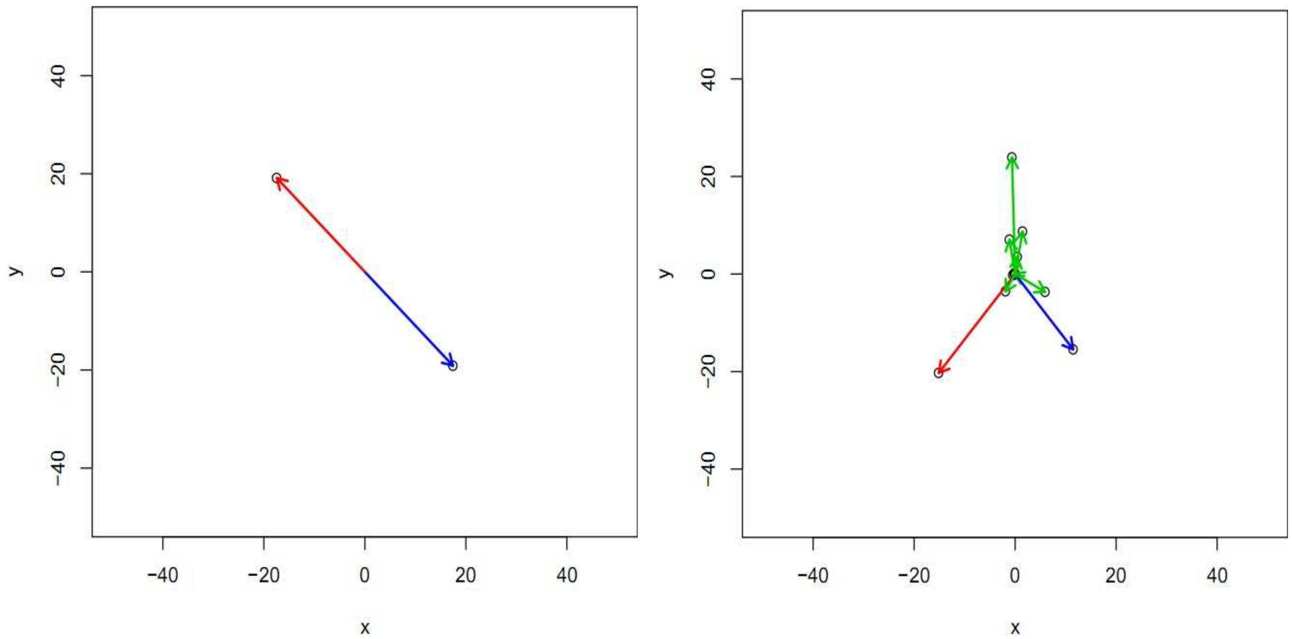
- The topology generated by the PS can be quite complicated

Parton Shower movie

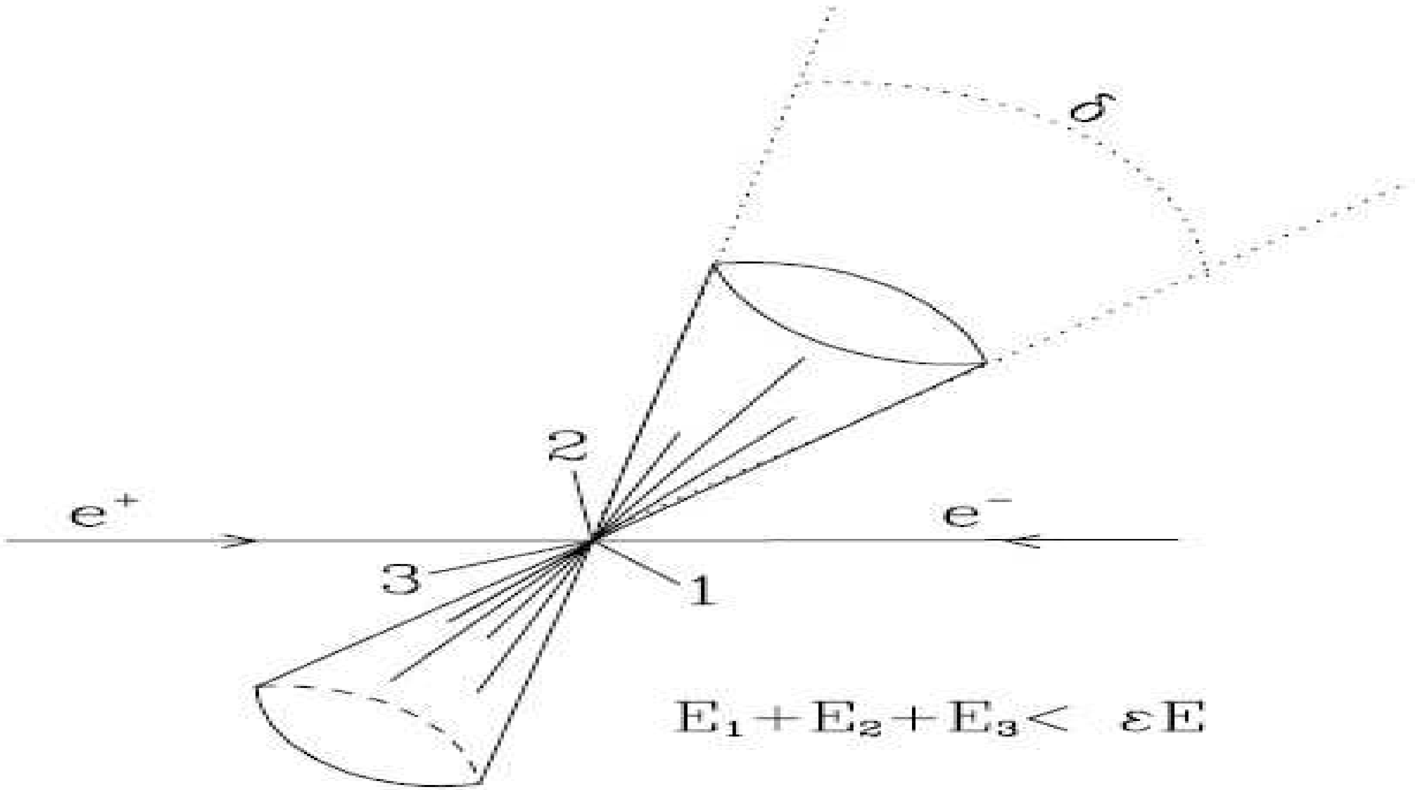


- The topology generated by the PS can be quite complicated
- These are events shape that can not be described by fixed order pert. calculations

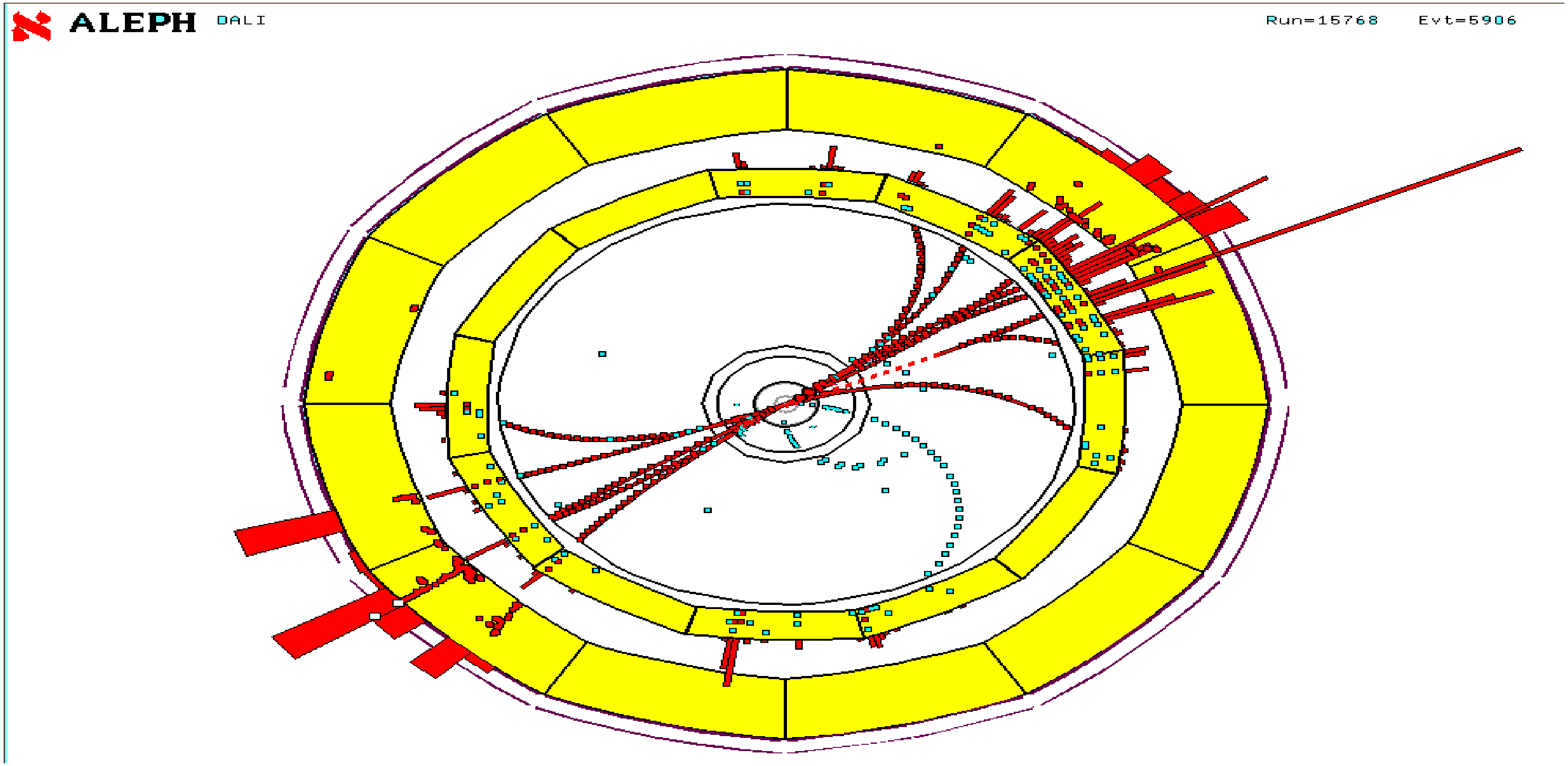
Parton Shower movie



- The topology generated by the PS can be quite complicated
- These are events shape that can not be described by fixed order pert. calculations
- Total cross section still given by hard scattering (usually LO), experiments usually normalise to data


δ
2
e^+
e^-
3
1
$E_1 + E_2 + E_3 < \varepsilon E$

Jets 1.



Jets 2.