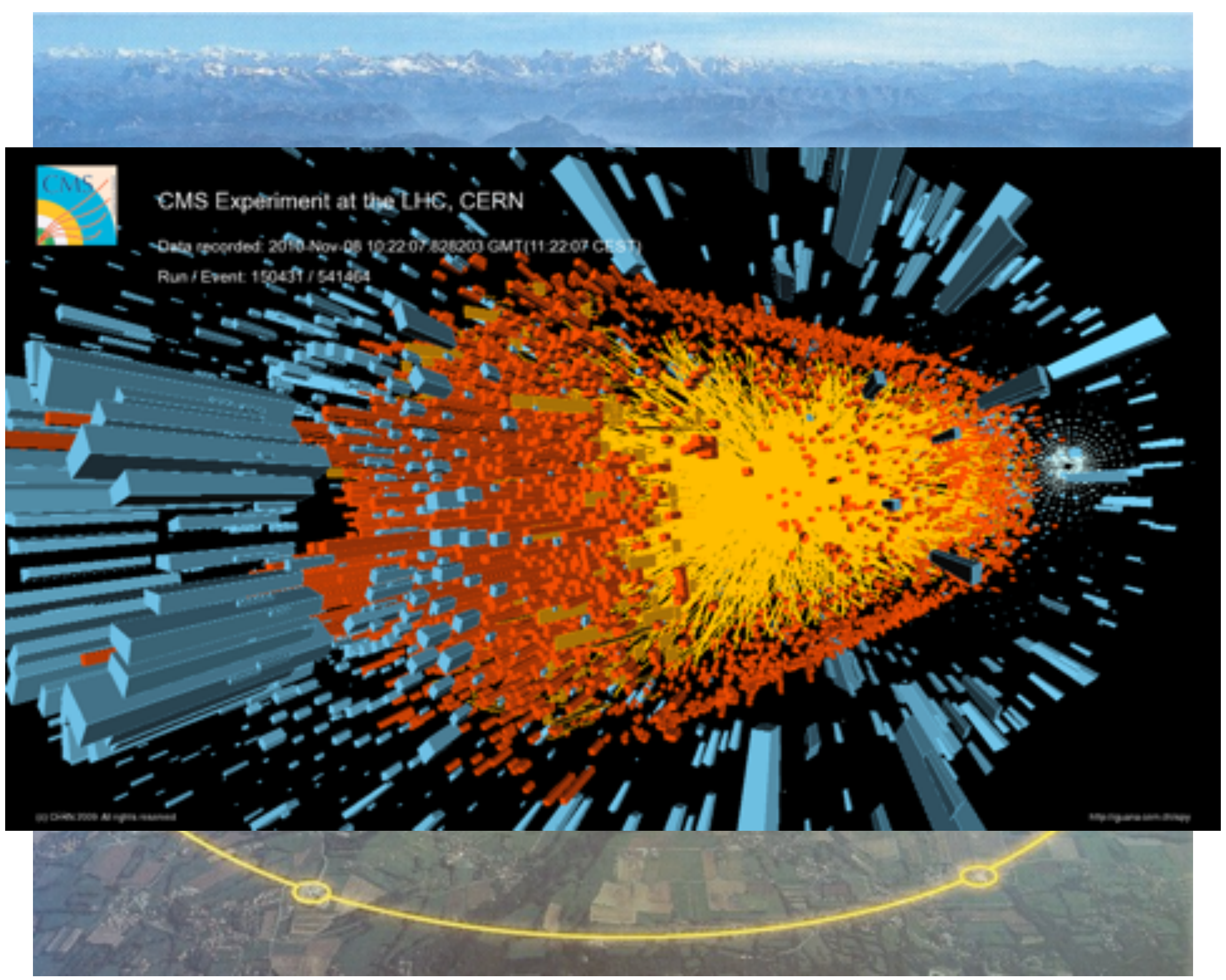
Monte-Carlo Generation

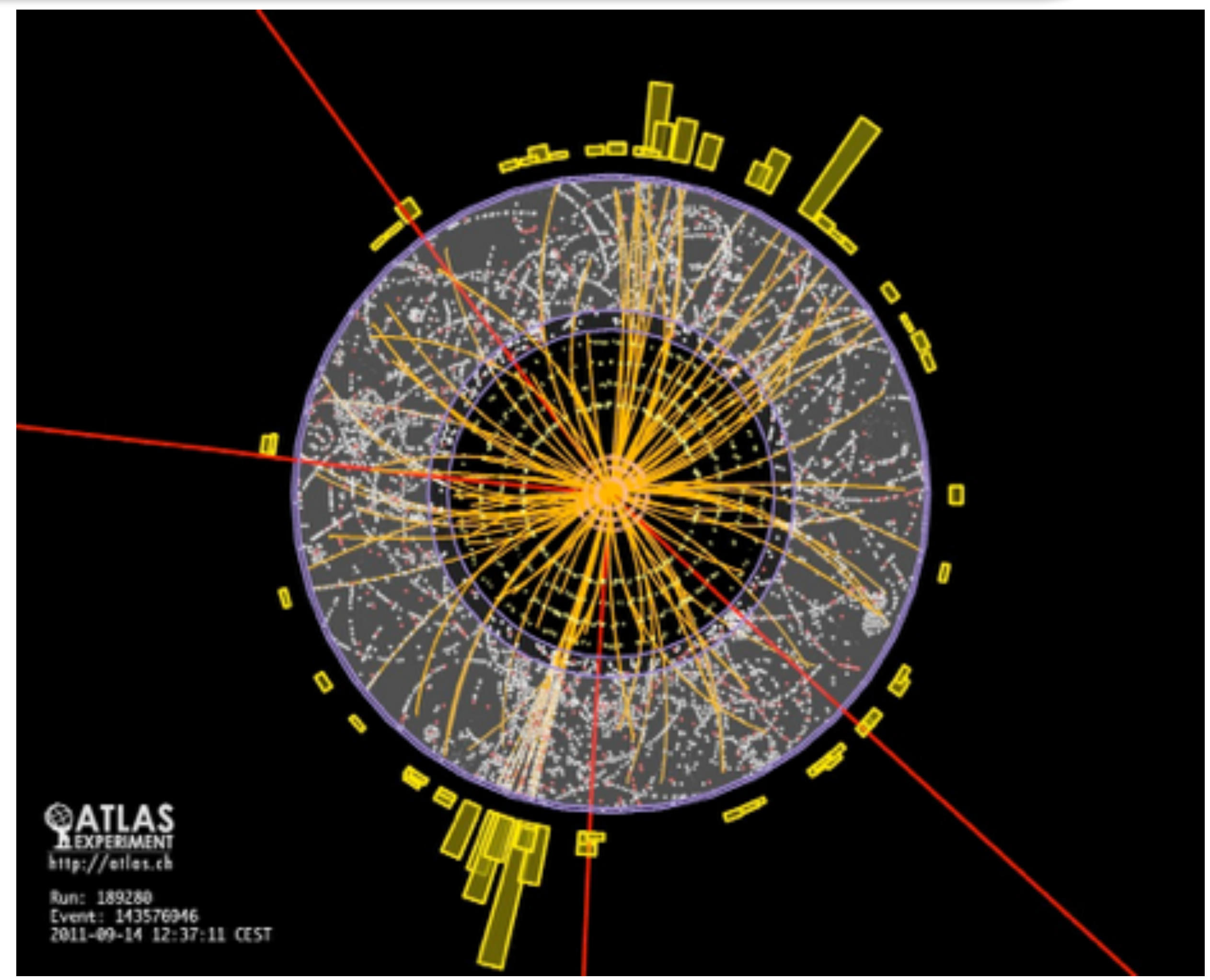
Olivier Mattelaer
IPPP/Durham

Introduction

Topic

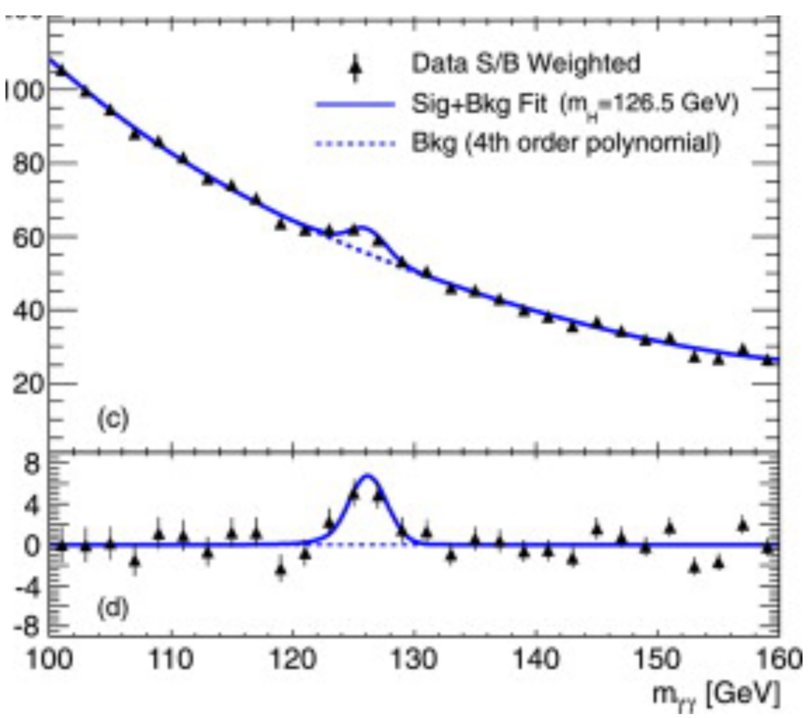
- Collider Physics
  - accelerating particle -> High Energy collision
- What do we need to predict/understand such collision?





Kind of measurement

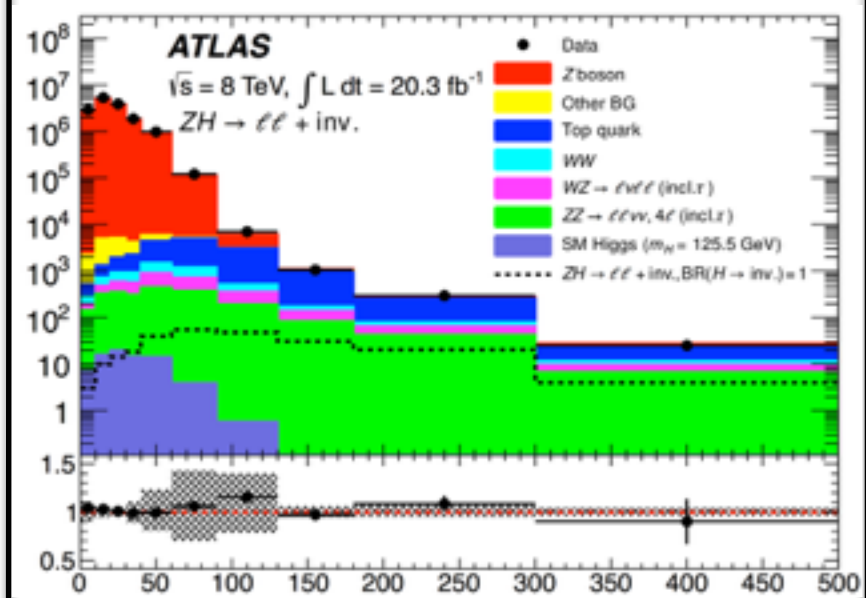
Peak



“EASY”

Background directly measured from **data**.
Theory needed only for parameter extraction

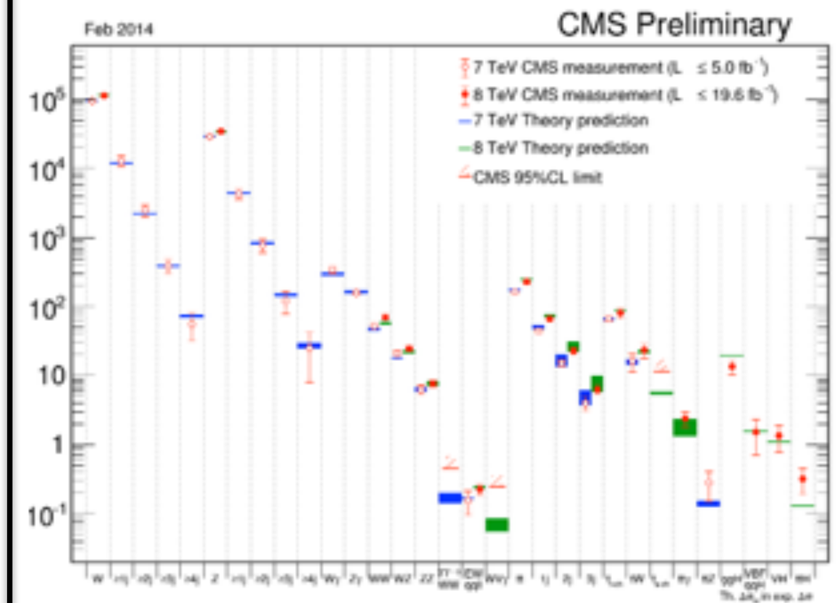
Shape



“HARD”

Background **SHAPE** needed.
Flexible MC for both signal and background validated and tuned to data

Rate

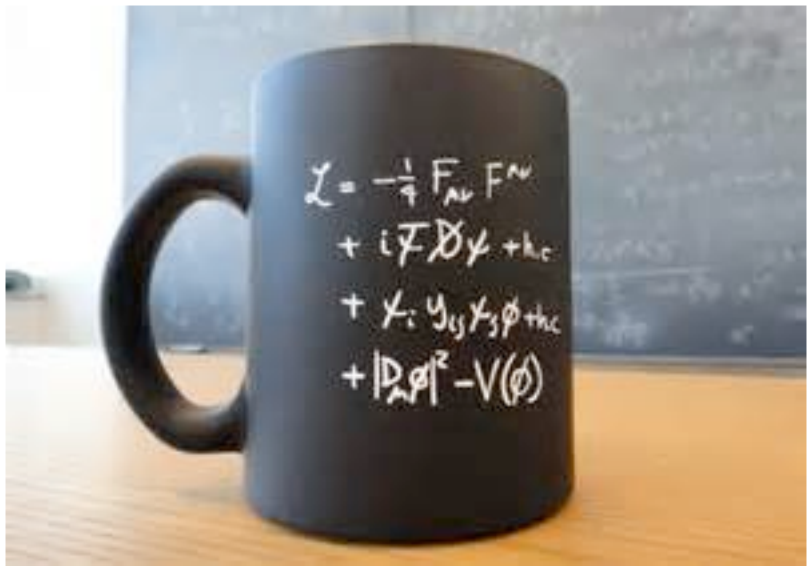


“VERY HARD”

Relies on prediction for both **shape** and **normalization**.
Complicated interplay of best simulations and data

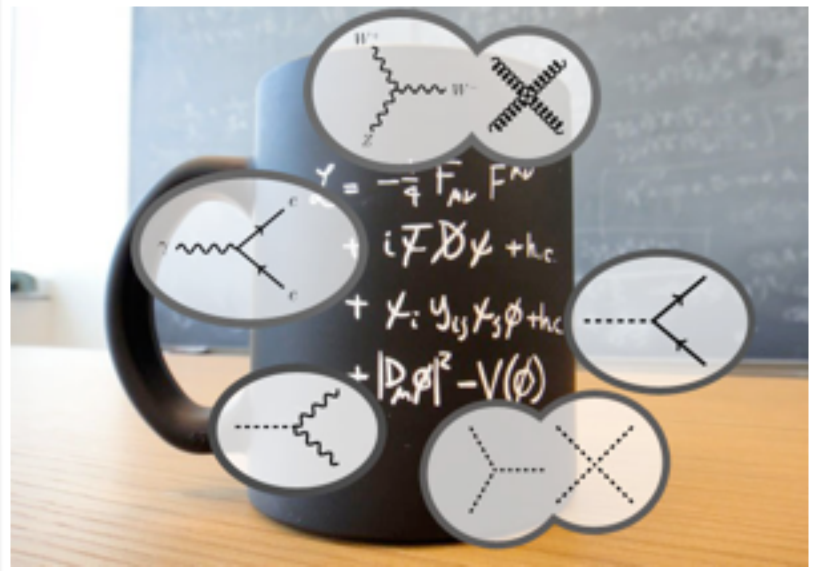
Theory side

Lagrangian



- This is Where the new idea are expressed

Feynman Rule

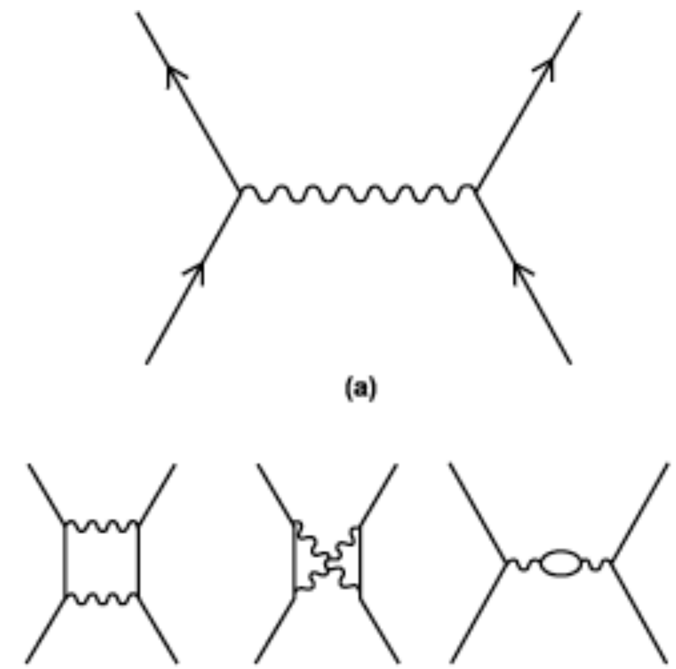


- Same information as the Lagrangian

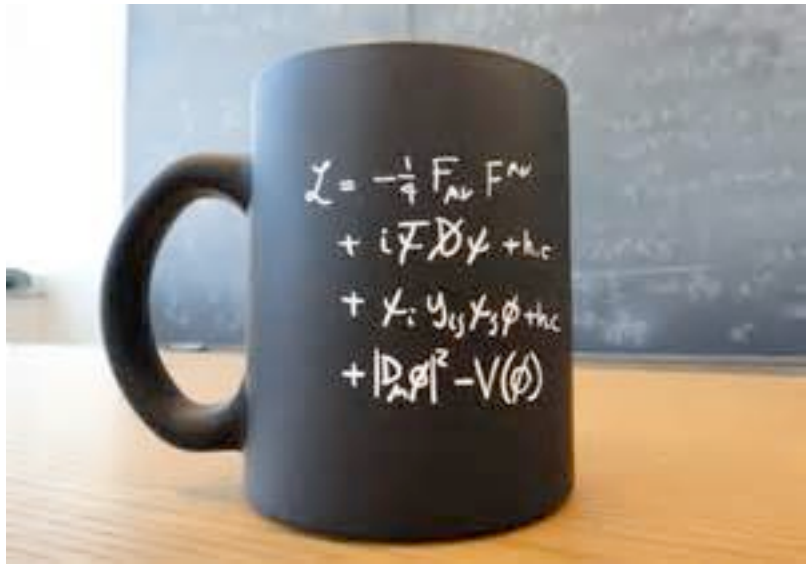
FeynRules

Cross-section

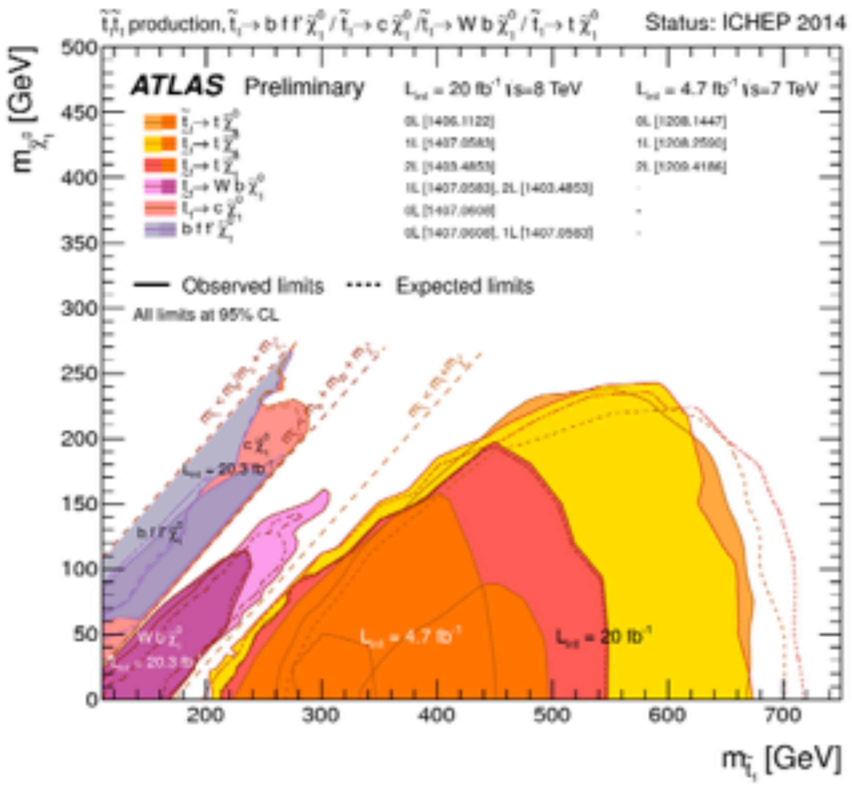
$$\frac{d\sigma}{d\ \cos\theta} = \left(\frac{d\sigma}{d\ \cos\theta}\right)_R \frac{(1+\cos\theta)/2}{\left[1+\frac{(1-\cos\theta)KE}{Mc^2}\right]}$$



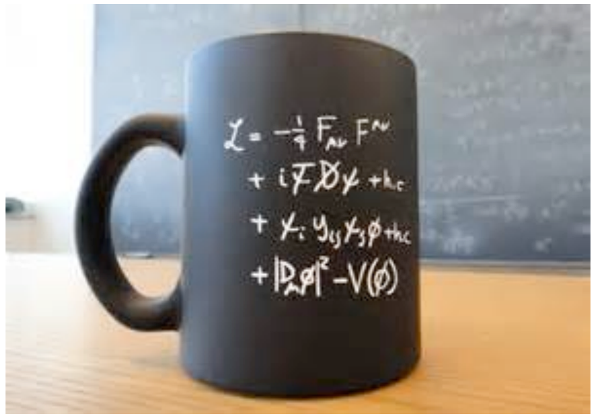
- What is the precision?

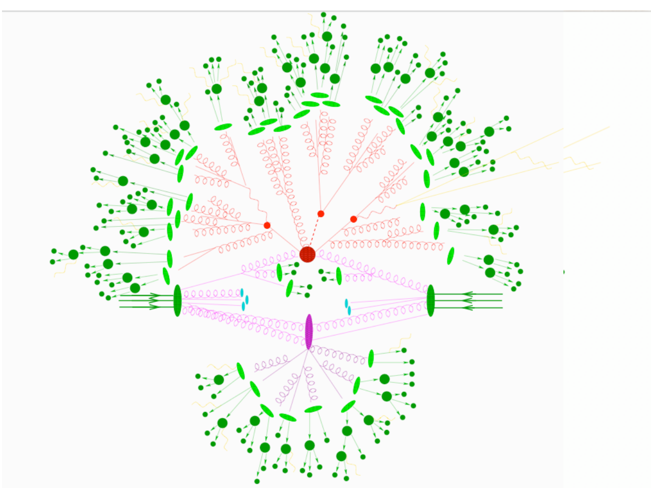


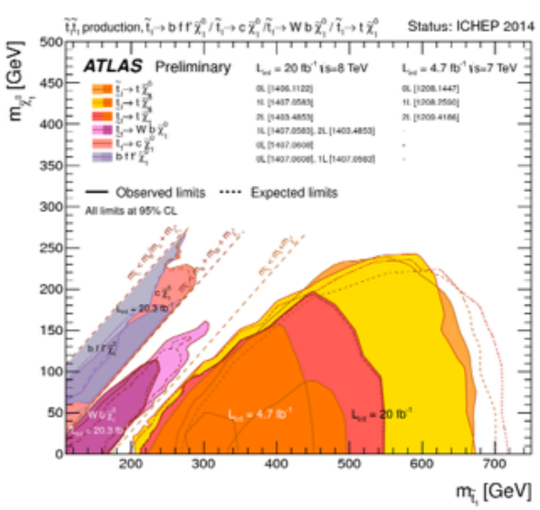
Monte-Carlo Physics



Filling the gap





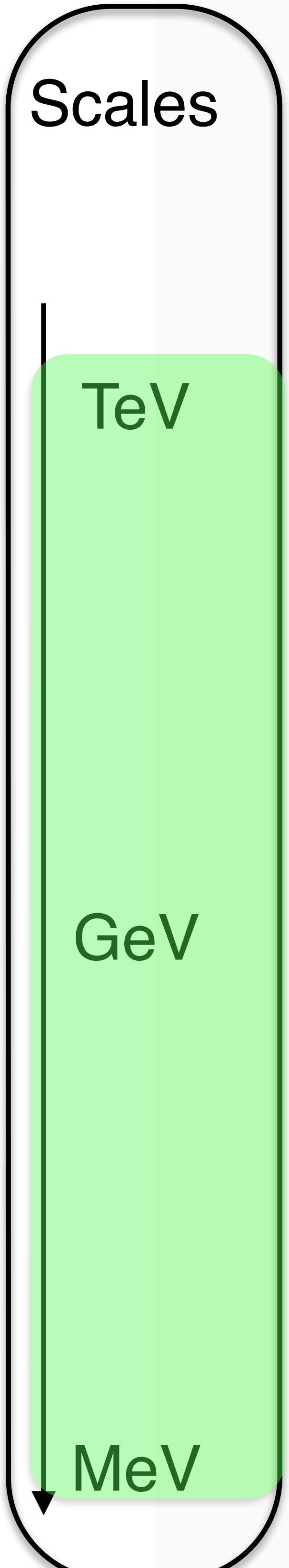


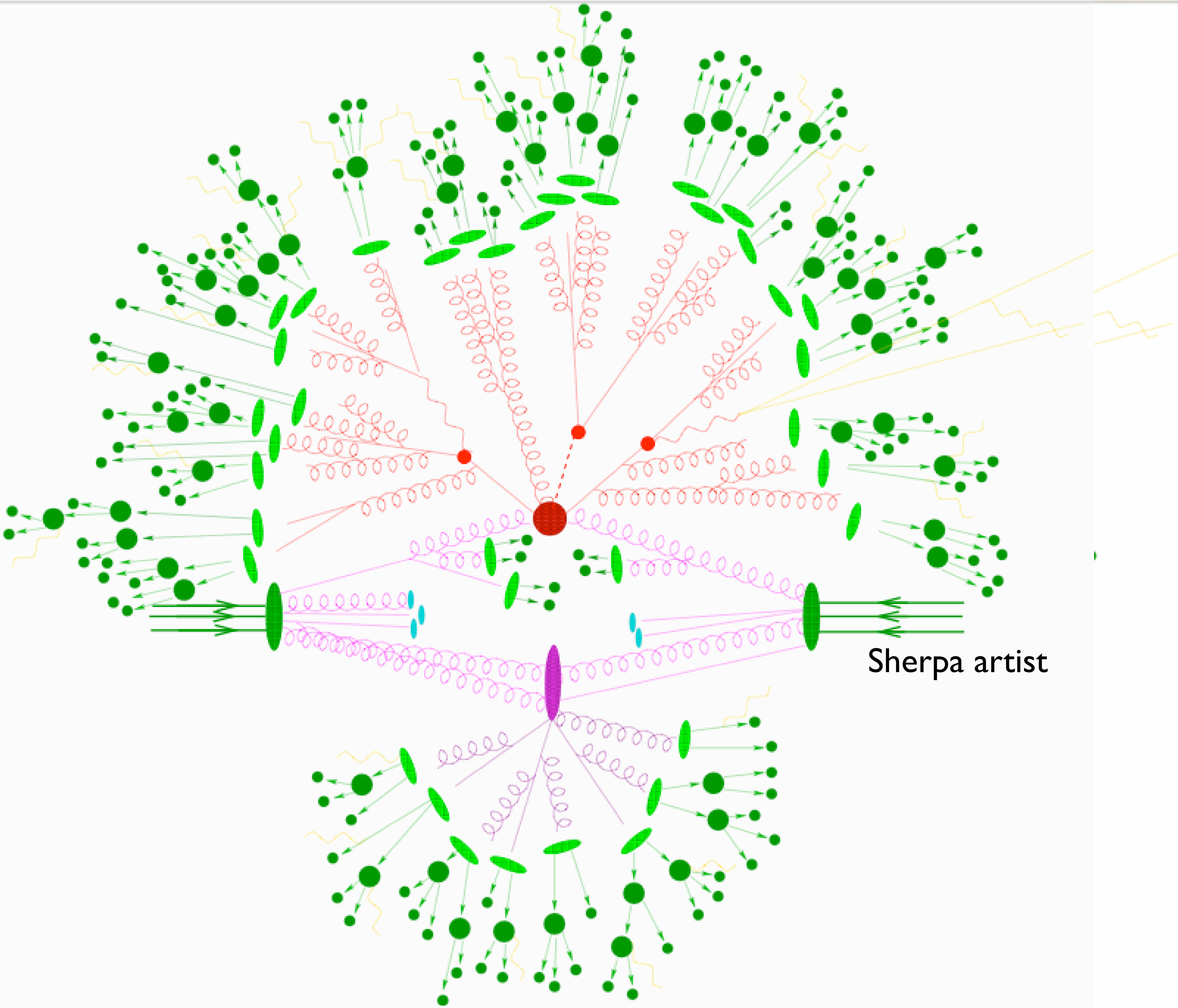
- Lagrangian
- Feynman Rules
- matrix element
- parton events
- shower events
- hadronized events
- detector simulation
- exclusion plot

- FeynRules / Sarah
- MadGraph / Comix
- MadEvent/Sherpa
- Pythia/Herwig
- Pythia/Herwig
- Delphes/Full Sim
- CheckMate/ MadAnalysis5

Simulation of collider events

What are the MC for?





What are the MC for?

Scales

TeV

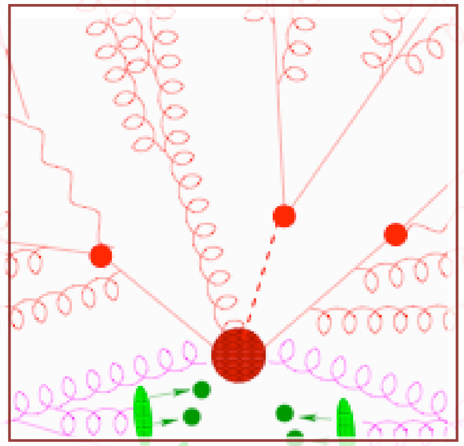
GeV

MeV

1. High-Q^2 Scattering

2. Parton Shower

☞ where BSM physics lies



☞ process dependent

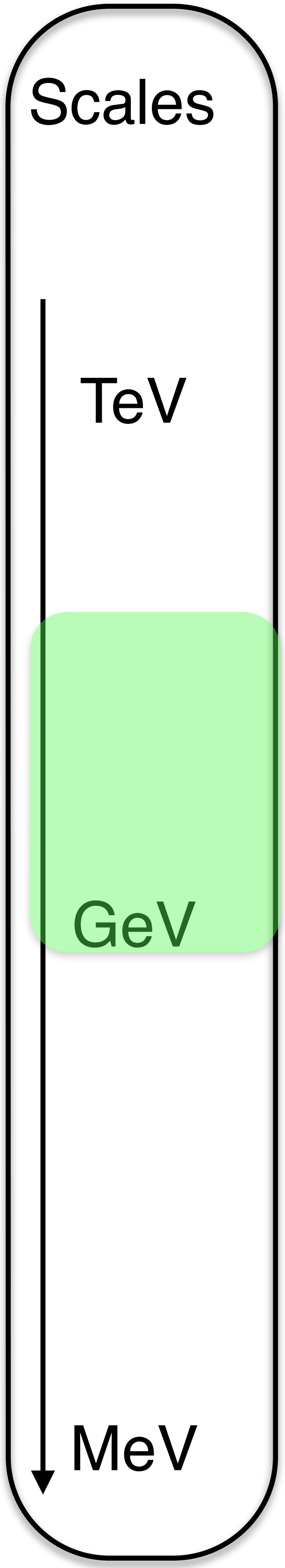
☞ first principles description

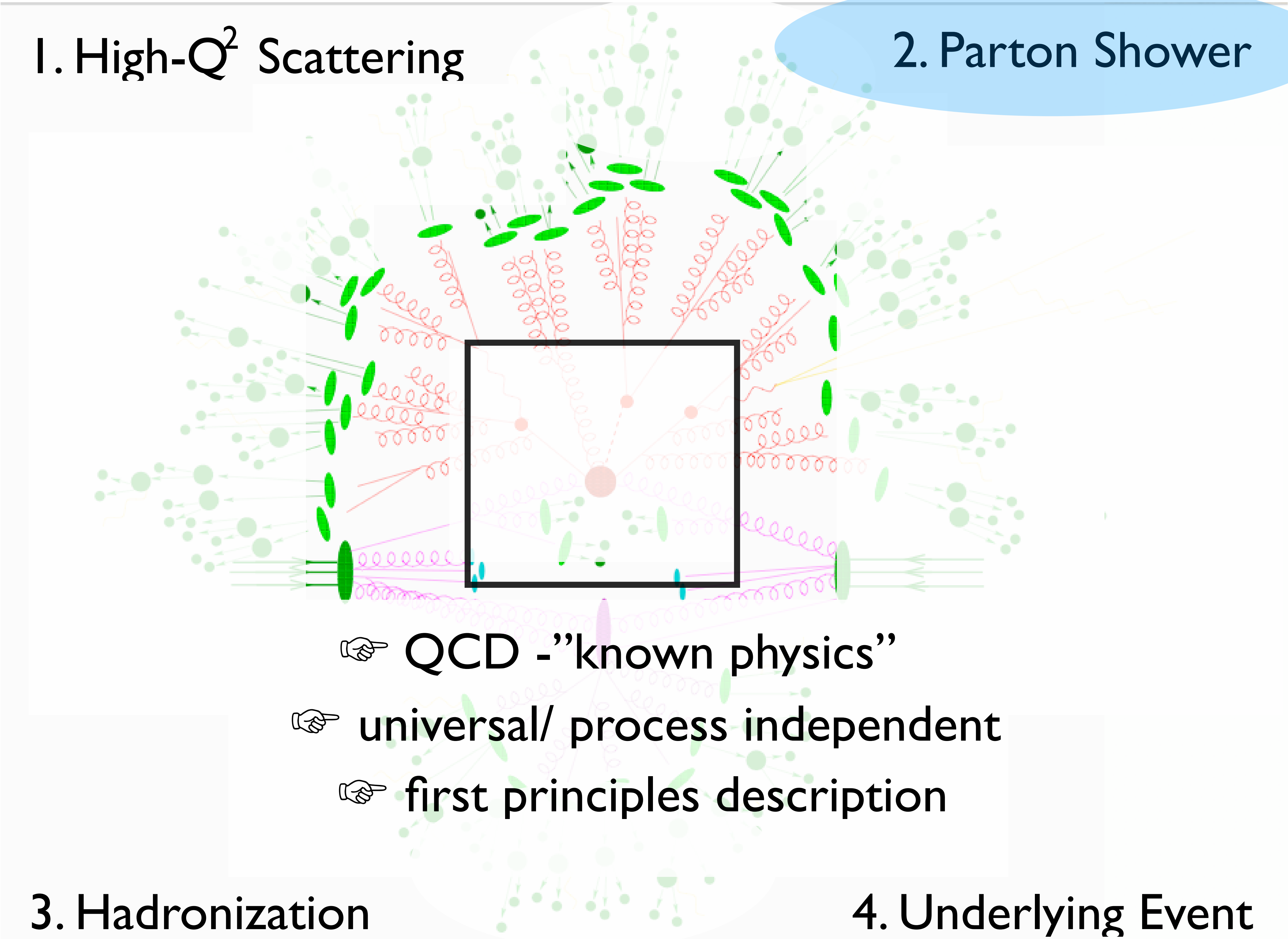
☞ it can be systematically improved

3. Hadronization

4. Underlying Event

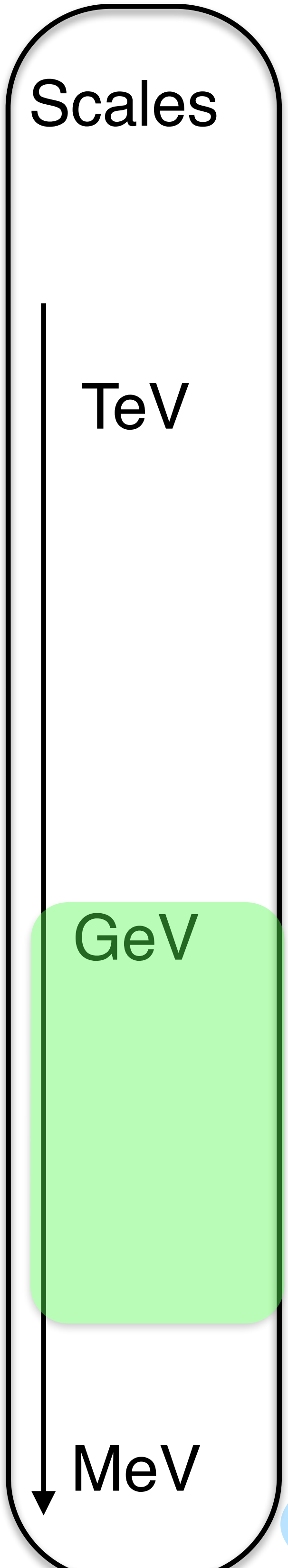
What are the MC for?

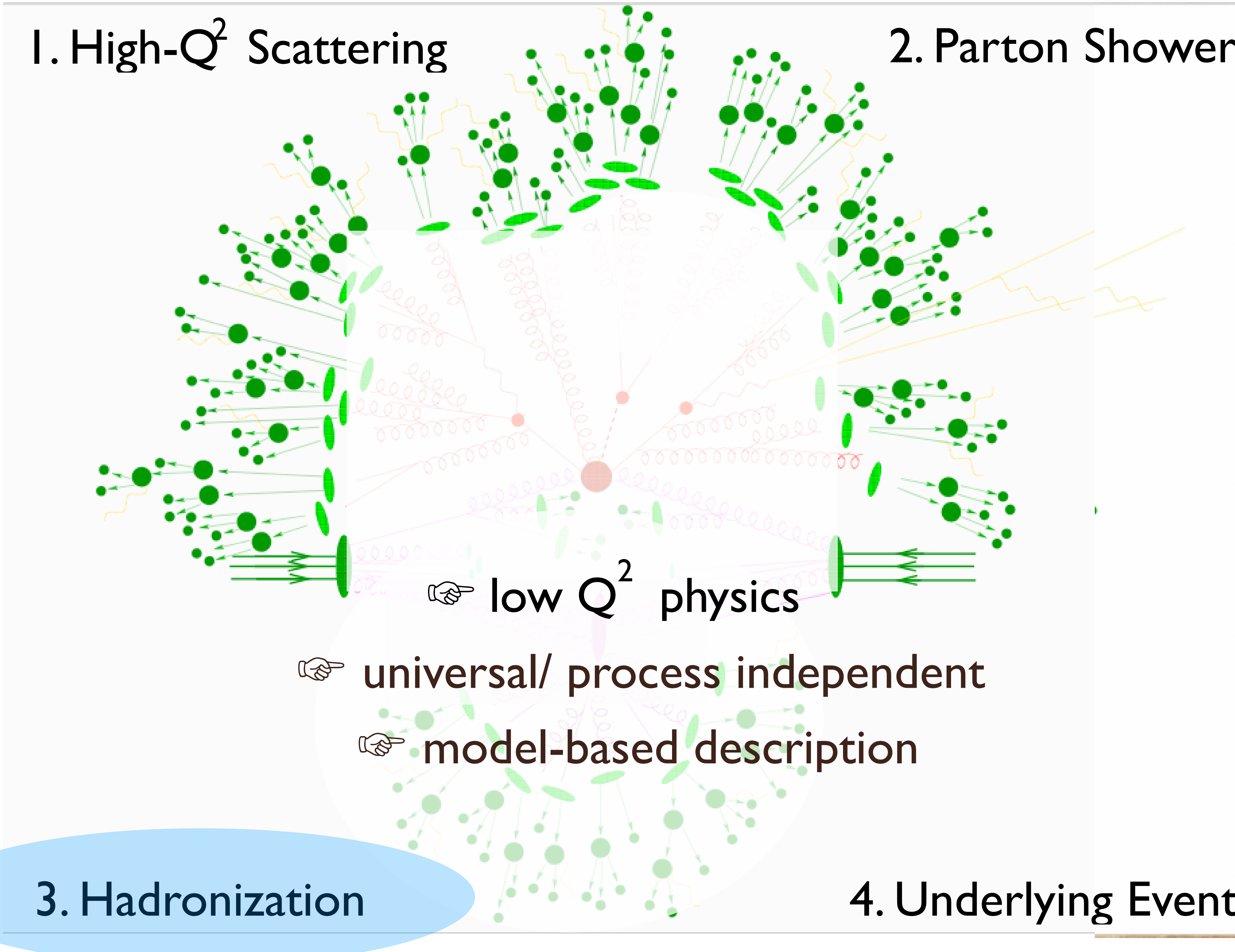




- ☞ QCD -"known physics"
- ☞ universal/ process independent
- ☞ first principles description

What are the MC for?





- ☞ low Q^2 physics
- ☞ universal/ process independent
- ☞ model-based description

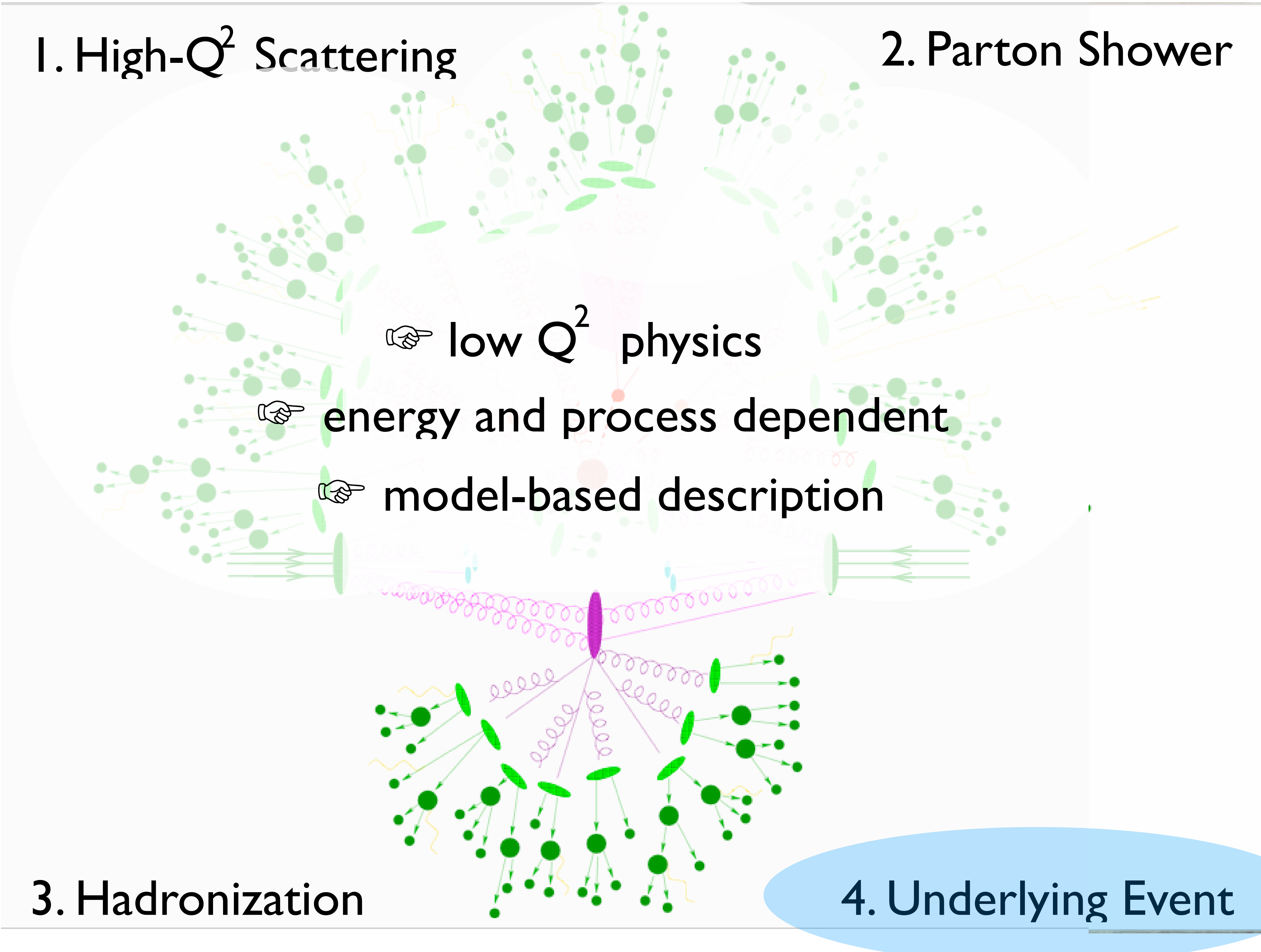
What are the MC for?

Scales

TeV

GeV

MeV



- ☞ low Q^2 physics
- ☞ energy and process dependent
- ☞ model-based description

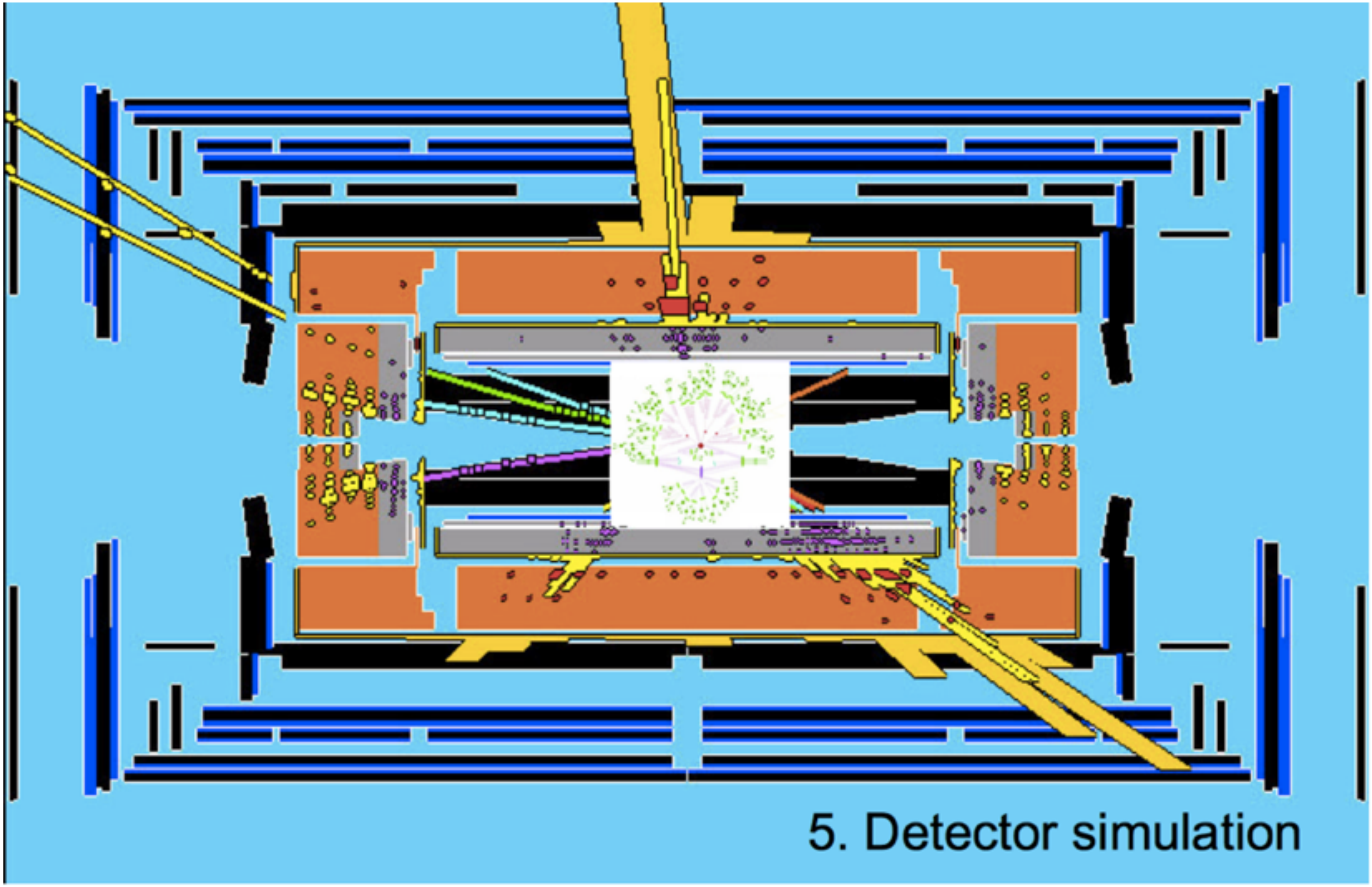
What are the MC for?

Scales

TeV

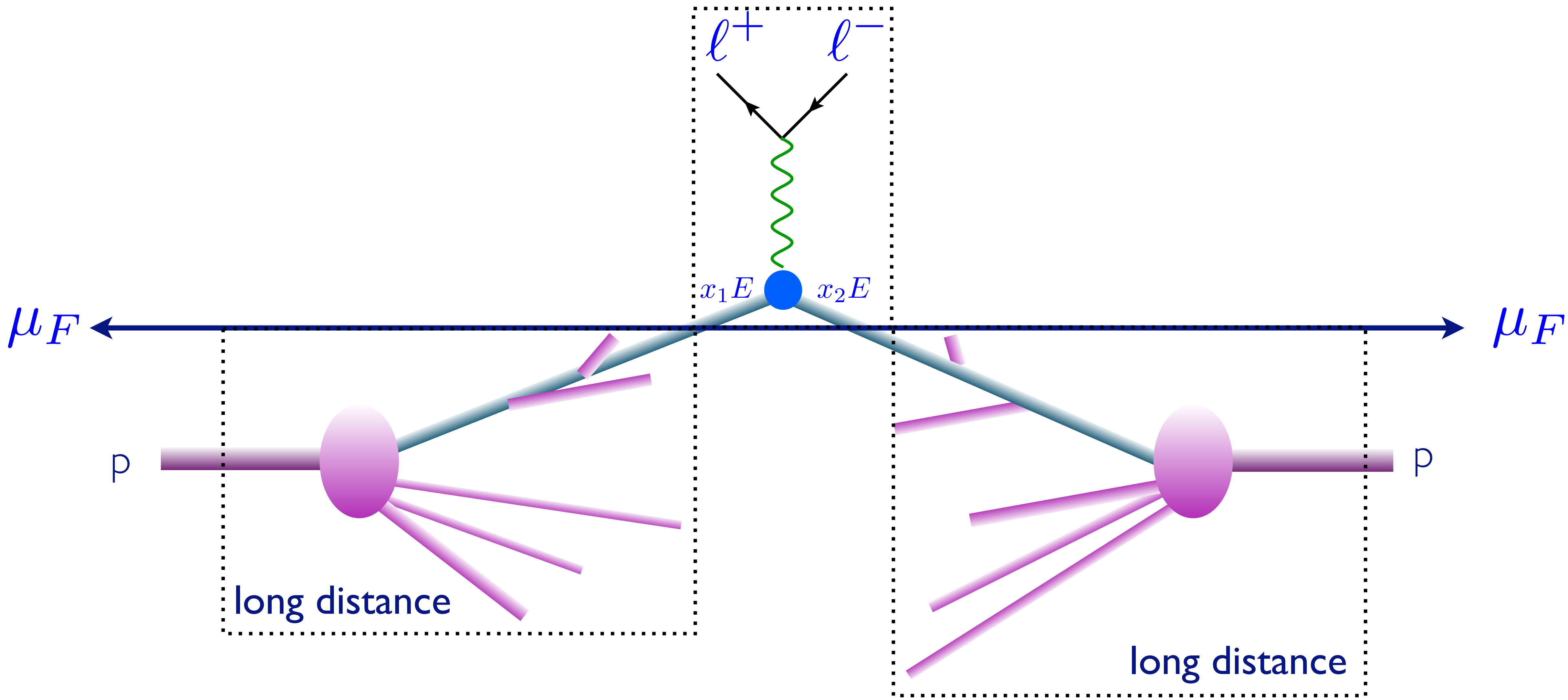
GeV

MeV



To Remember

- Multi-scale problem
  - ➡ New physics visible only at High scale
  - ➡ Problem split in different scale



$$\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

Parton-level cross section

Perturbative expansion

$d\hat{\sigma}_{ab\to X}(\hat{s},\mu_F,\mu_R)$ **Parton-level cross section**

- The parton-level cross section can be computed as a series in perturbation theory, using the coupling constant as an expansion parameter, schematically:

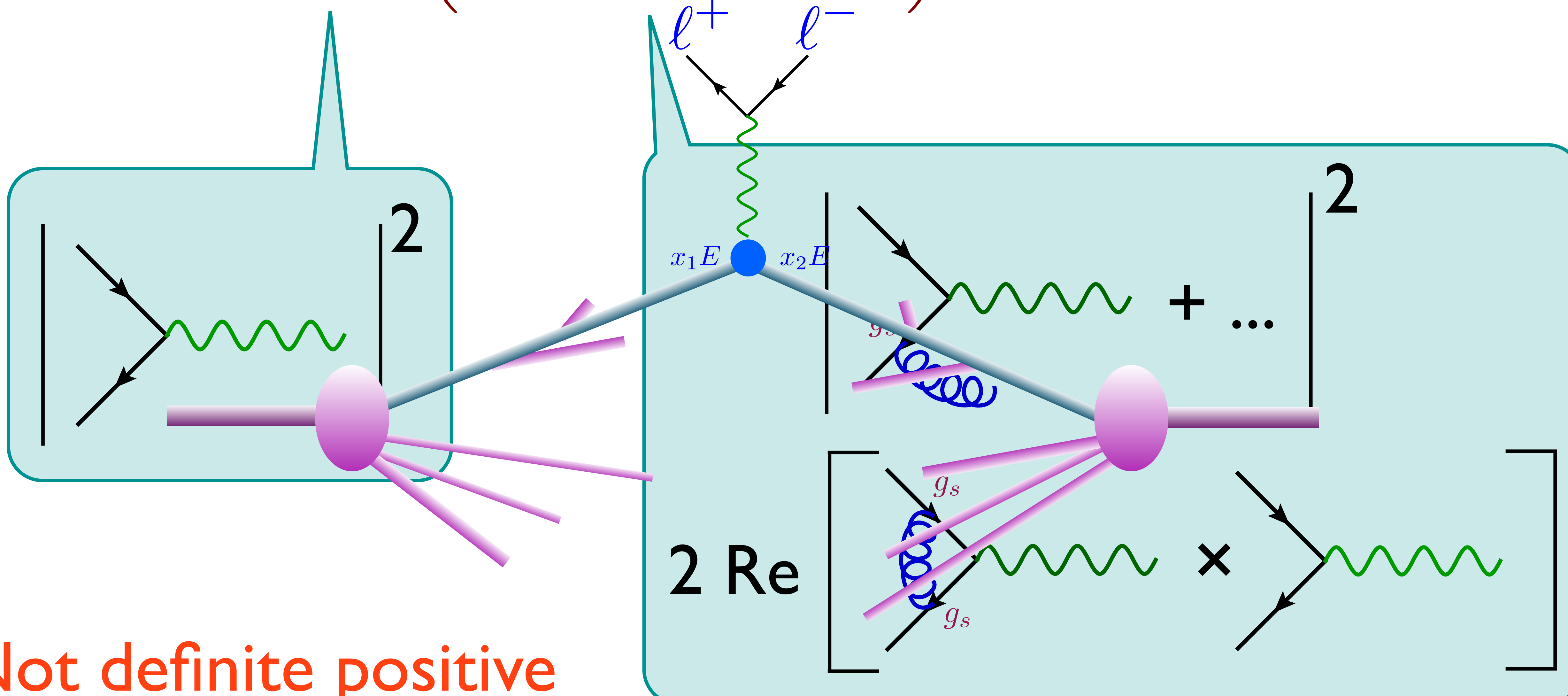
$$\hat{\sigma} = \underbrace{\sigma^{\text{Born}}}_{\text{LO predictions}}\left(1+\underbrace{\frac{\alpha_s}{2\pi}\sigma^{(1)}}_{\text{NLO corrections}}+\underbrace{\left(\frac{\alpha_s}{2\pi}\right)^2\sigma^{(2)}}_{\text{NNLO corrections}}+\underbrace{\left(\frac{\alpha_s}{2\pi}\right)^3\sigma^{(3)}}_{\text{N3LO or NNNLO corrections}}+\ldots\right)$$

- Including higher corrections improves predictions and reduces theoretical uncertainties

NLO predictions

- As an example, consider Drell-Yan Z/γ* production

$$\hat{\sigma} = \sigma^{\text{Born}}\left(1 + \frac{\alpha_s}{2\pi}\sigma^{(1)} + \ldots\right)$$



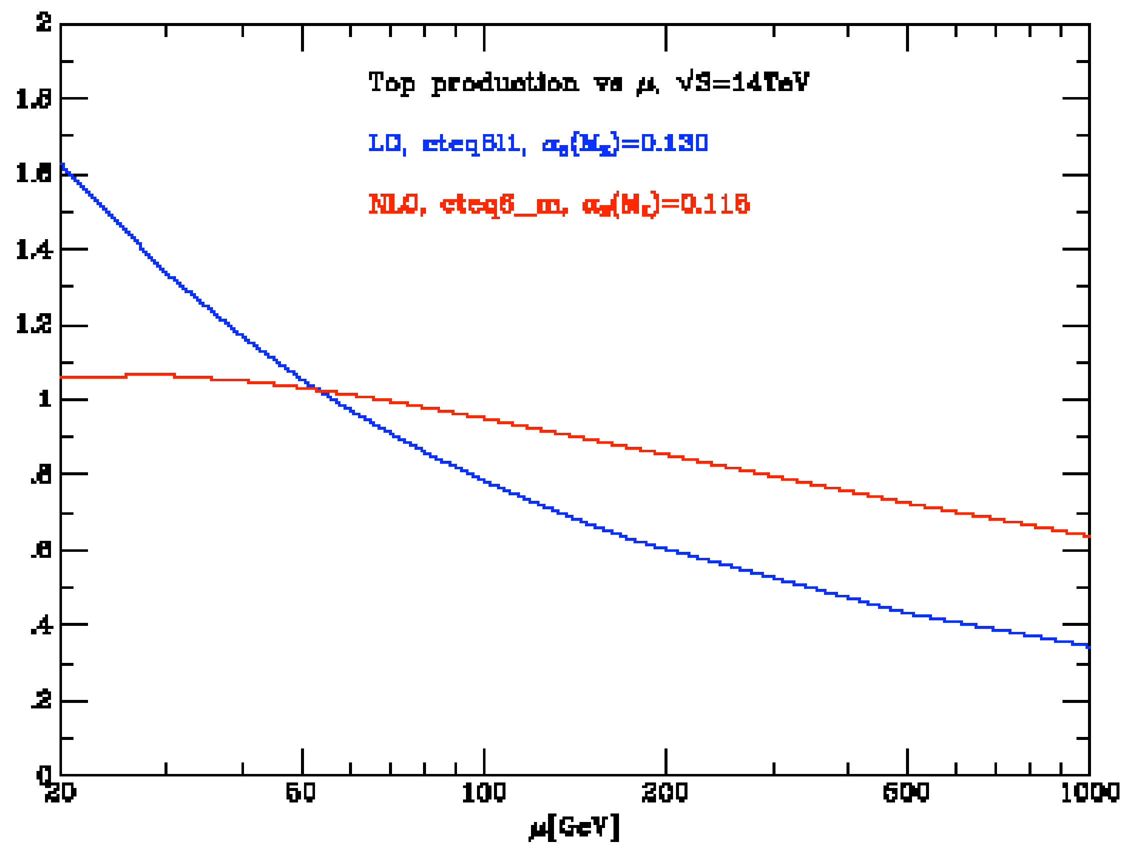
Not definite positive

Improved predictions

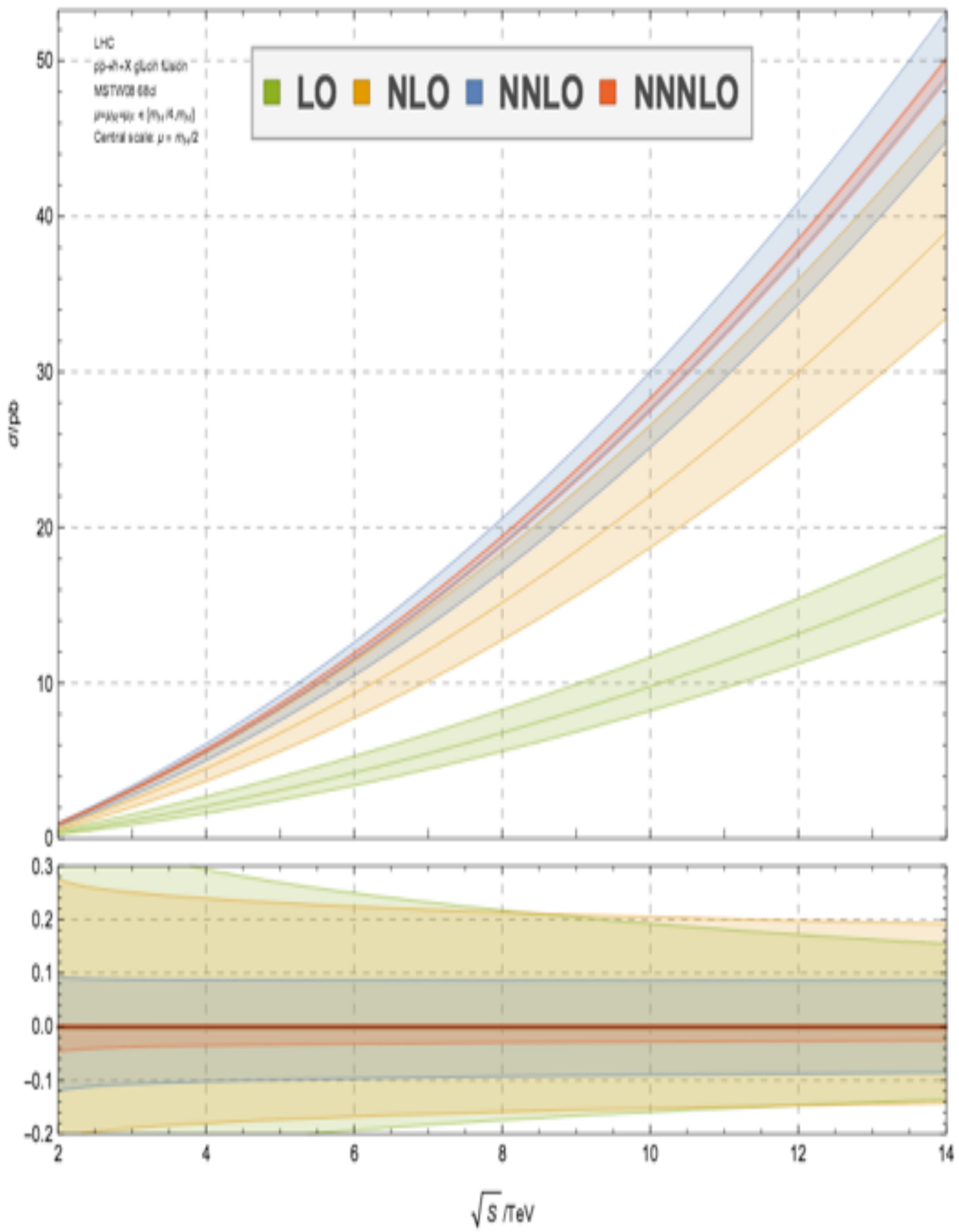
$$d\sigma = \sum_{a,b} \int dx_1 dx_2 \;\; f_a(x_1, \mu_F) f_b(x_2, \mu_F) \, d\hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

$$\hat{\sigma} = \sigma^{\text{Born}} \left(1 + \frac{\alpha_s}{2\pi} \sigma^{(1)} + \left(\frac{\alpha_s}{2\pi}\right)^2 \sigma^{(2)} + \left(\frac{\alpha_s}{2\pi}\right)^3 \sigma^{(3)} + \dots \right)$$

- Leading Order predictions can depend strongly on the renormalization and factorization scales
- Including higher order corrections reduces the dependence on these scales



Higgs at N3LO



- LO calculation is not reliable,
- but the perturbative series stabilises at NNLO/N3LO
- NLO estimation of the uncertainties (by scale variation) works reasonably well

Let's focus on NLO

Infrared safe observables

- For an observable to be calculable in fixed-order perturbation theory, the observable should be infrared safe, i.e., it should be insensitive to the emission of soft or collinear partons.

- In particular, if p_i is a momentum occurring in the definition of an observable, it most be invariant under the branching

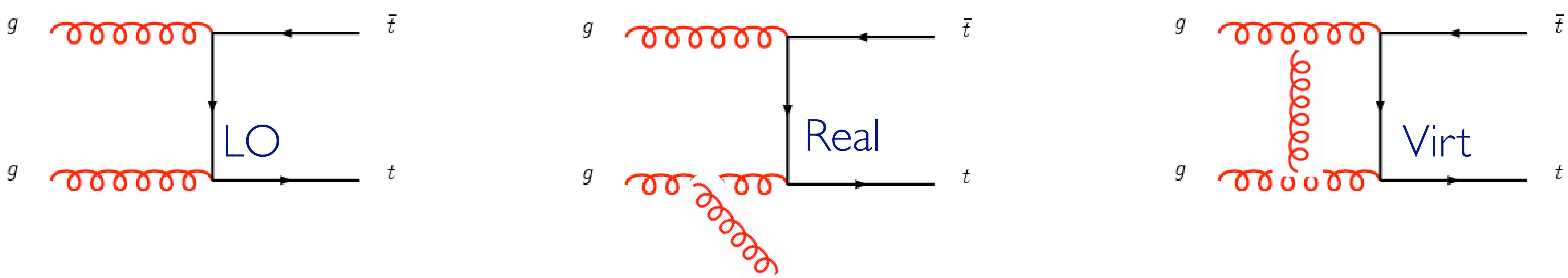
$$p_i \longrightarrow p_j + p_k,$$

whenever p_j and p_k are collinear or one of them is soft.

- Examples
  - "The number of gluons" produced in a collision is not an infrared safe observable
  - "The number of hard jets defined using the k_T algorithm with a transverse momentum above 40 GeV," produced in a collision is an infrared safe observable

NLO...?

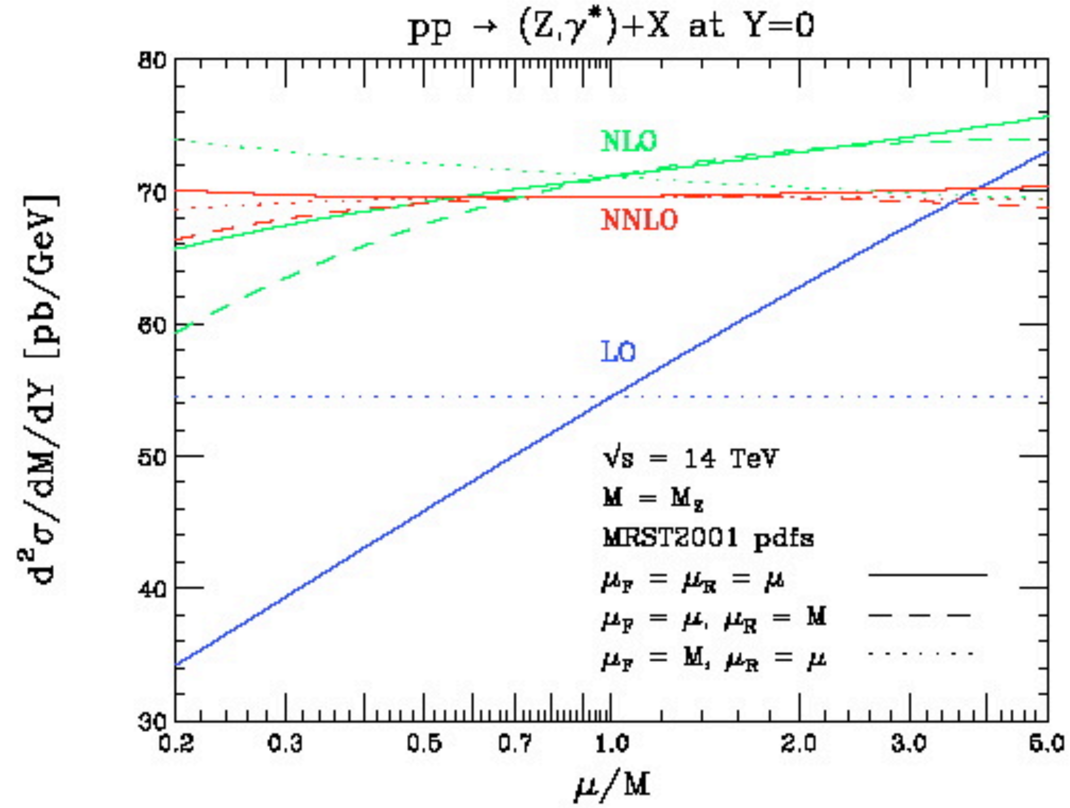
- Are all (IR-safe) observables that we can compute using a NLO code correctly described at NLO? Suppose we have a NLO code for pp ⟶ ttbar



	NLO?
• Total cross section	✔
• Transverse momentum of the top quark	✔
• Transverse momentum of the top-antitop pair	✘
• Transverse momentum of the jet	✘
• Top-antitop invariant mass	✔
• Azimuthal distance between the top and anti-top	✘

Going NNLO...?

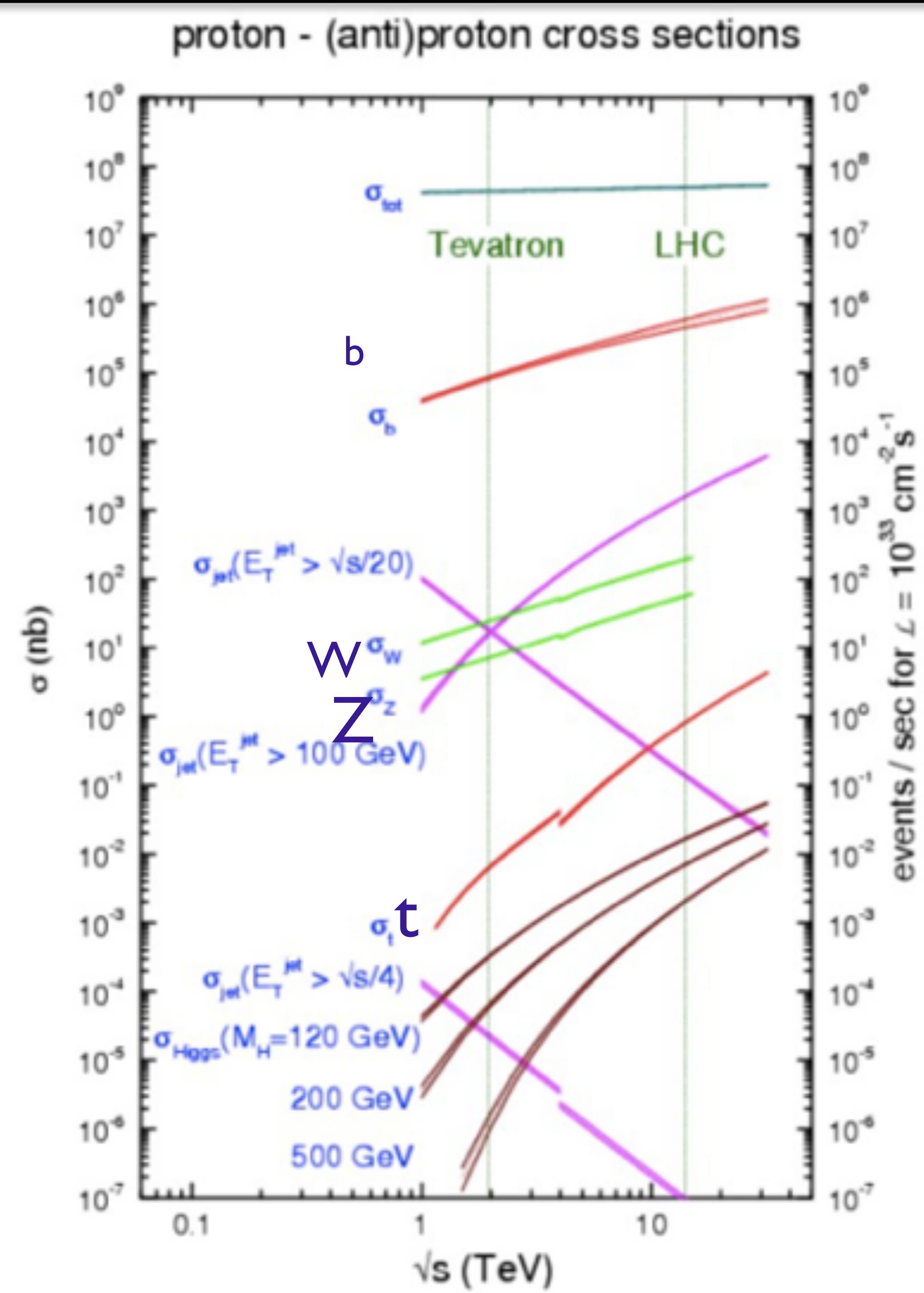
- **NNLO is the current state-of-the-art. There are only a few results available: Higgs (N3LO available), Drell-Yan, ttbar**
- **Why do we need it?**
  - control of the uncertainties in a calculation
  - It is "mandatory" if NLO corrections are very large to check the behavior of the perturbative series
  - It is needed for Standard Candles and very precise tests of perturbation theory, exploiting all the available information, e.g. for determining NNLO PDF sets



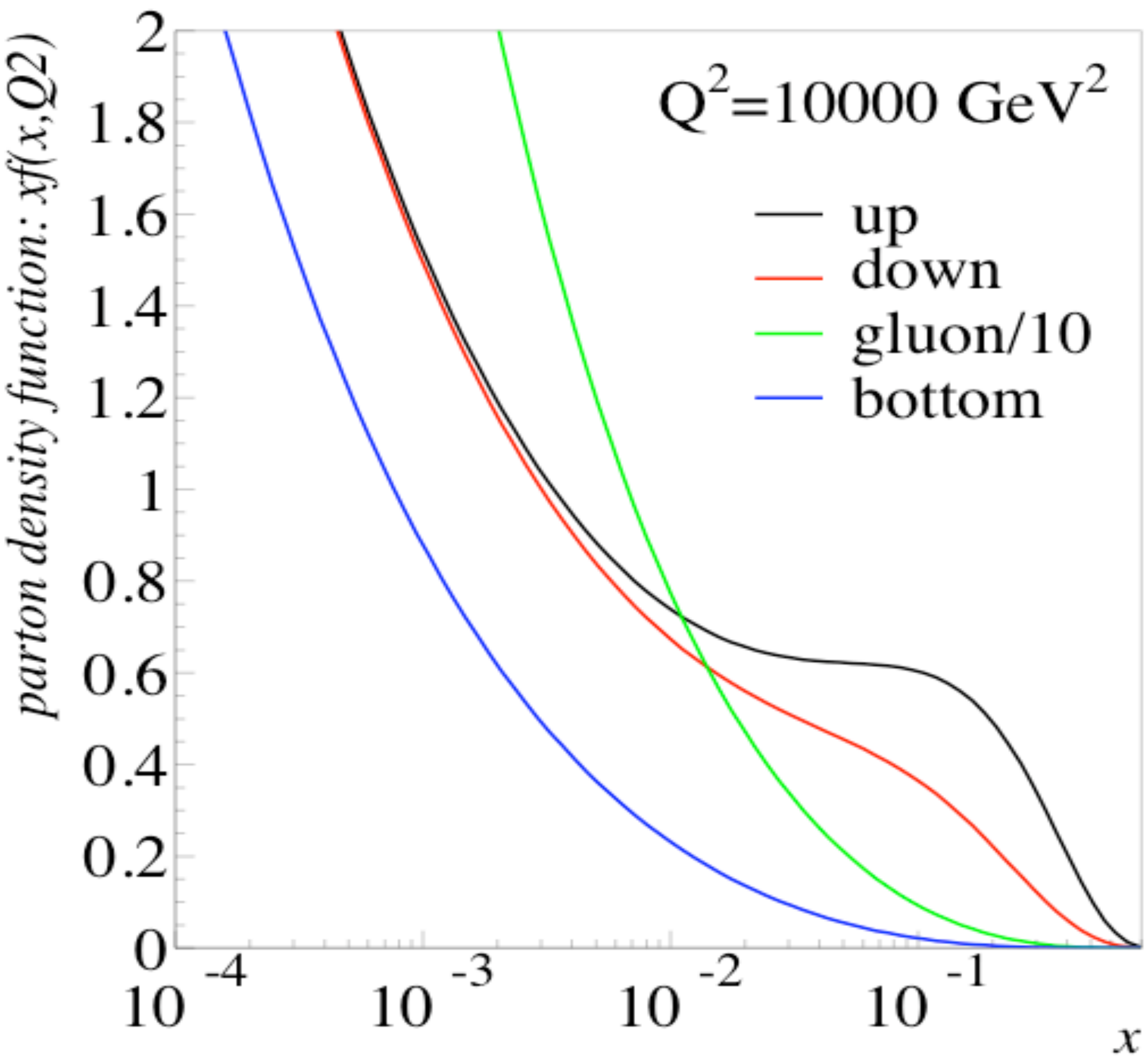
Let's focus on LO

Hadron Colliders

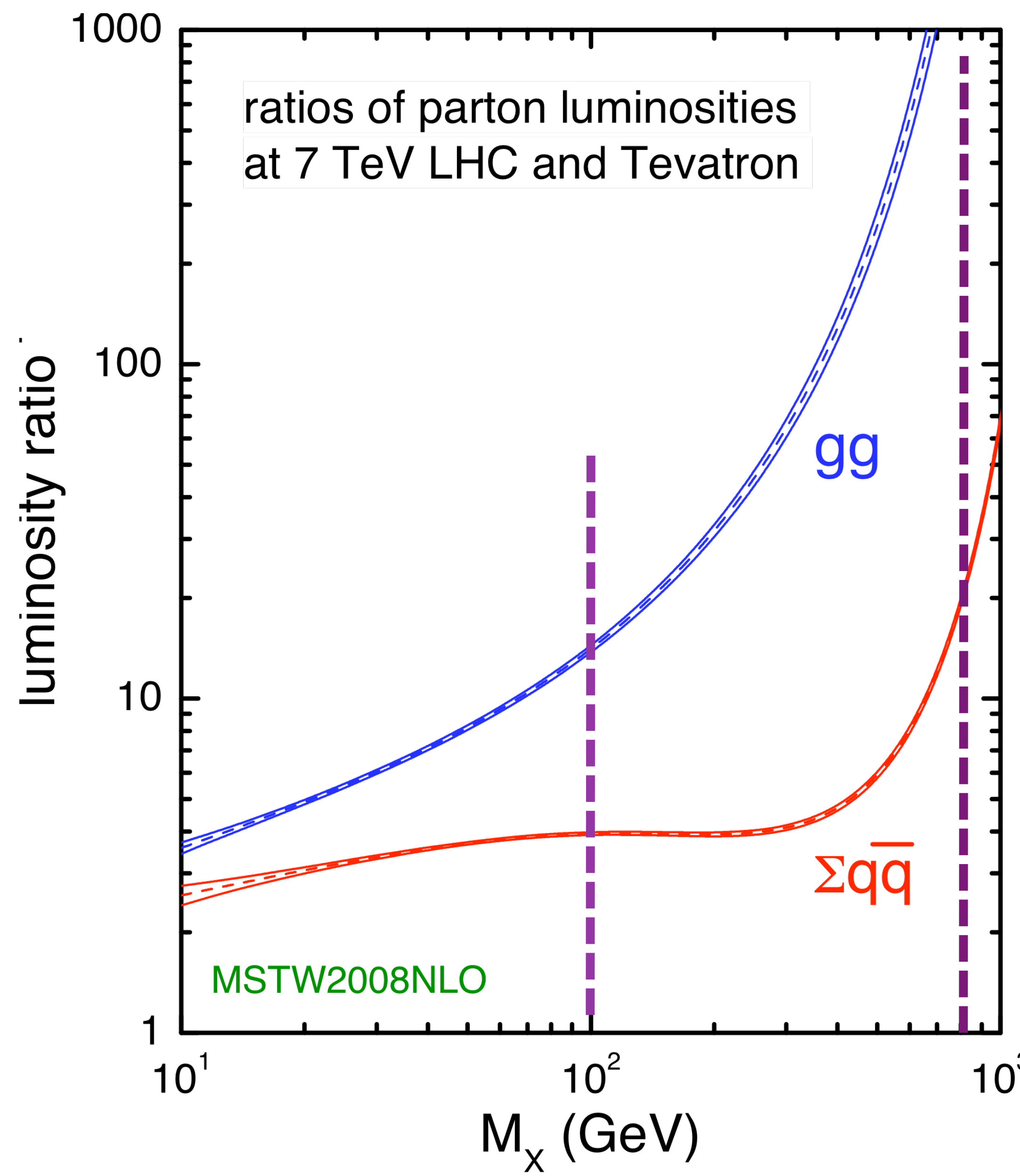
$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1,\mu_F) f_b(x_2,\mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s},\mu_F,\mu_R)$$



Parton densities

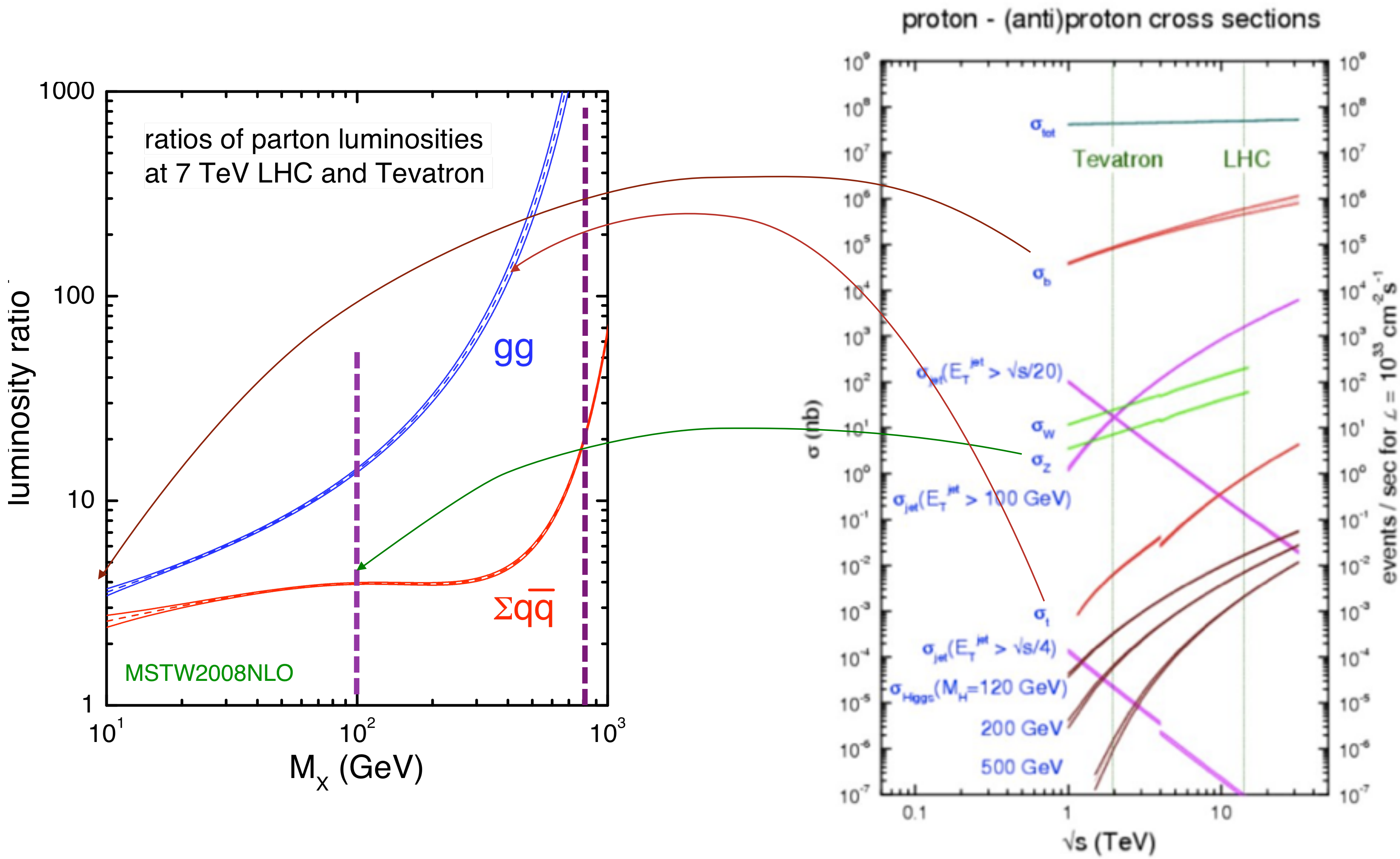


At small x (small ŝ), gluon domination.
At large x valence quarks



LHC formidable at large mass –
For low mass, Tevatron backgrounds smaller

Back to the processes



To Remember

$$\sum_{a,b} \int dx_1 dx_2 d\Phi_{\mathrm{FS}}\, f_a(x_1, \mu_F) f_b(x_2, \mu_F)\, \hat{\sigma}_{ab\to X}(\hat{s}, \mu_F, \mu_R)$$

Phase-space integral

Parton density functions

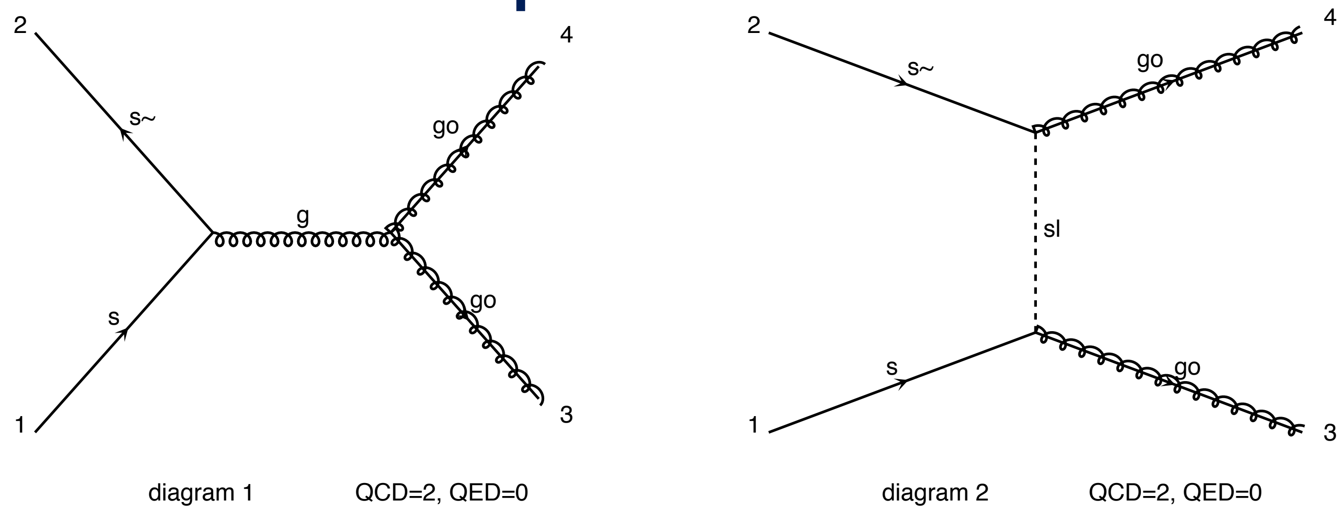
Parton-level cross section

- PDF: content of the proton
  - ➡ Define the physics/processes that will dominate on your accelerator
- NLO/NNLO: Reduce scale uncertainty linked to your division of your multi-scale problem

Matrix-Element

Calculate a given process (e.g. gluino pair)

- Determine the production mechanism



- Evaluate the matrix-element

$|\mathcal{M}|^2$ ➡Need Feynman Rules!

- Phase-Space Integration

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

Easy enough

Hard

Thuesday

Very Hard (in general)

Now

Monte Carlo Integration and Generation

Monte Carlo Integration

Calculations of cross section or decay widths involve integrations over high-dimension phase space of very peaked functions:

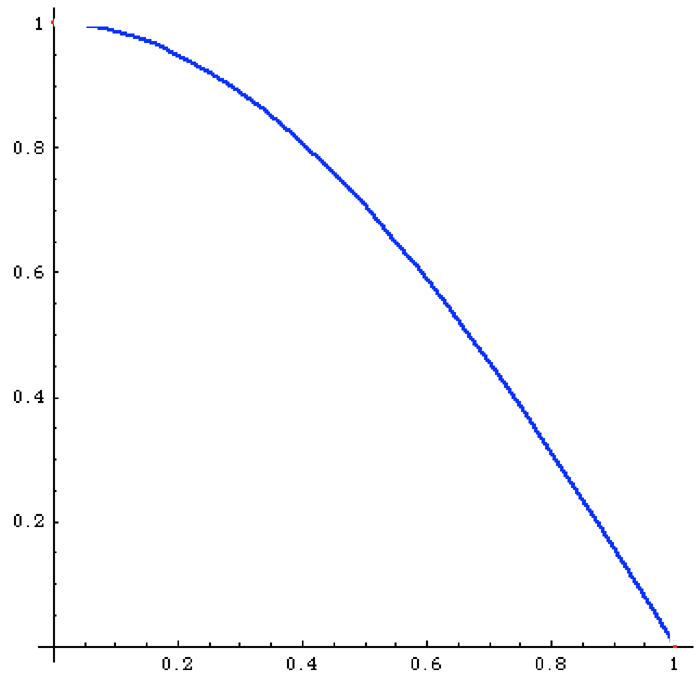
$$Dim[\Phi(n)] \sim 3n$$

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

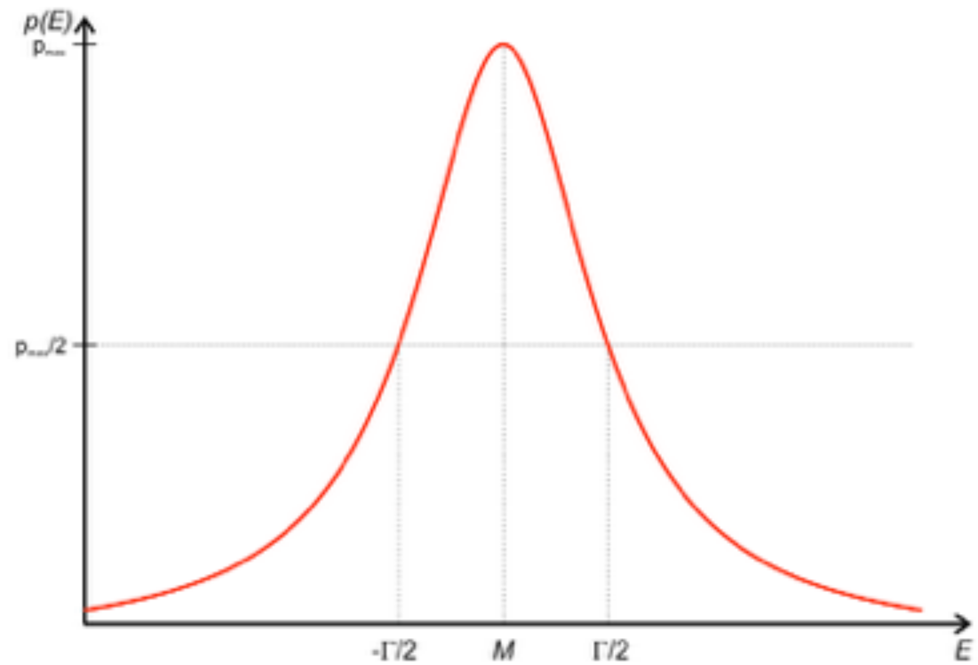
General and flexible method is needed

Integration

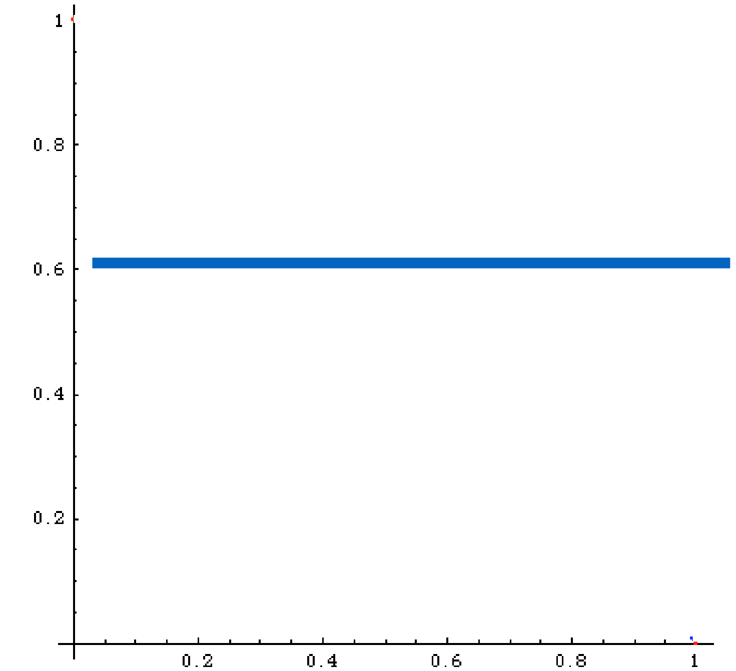
$$I = \int_0^1 dx \cos \frac{\pi}{2} x$$



$$\int \frac{dq^2}{(q^2 - M^2 + iM\Gamma)^2}$$



$$\int dx\, C$$



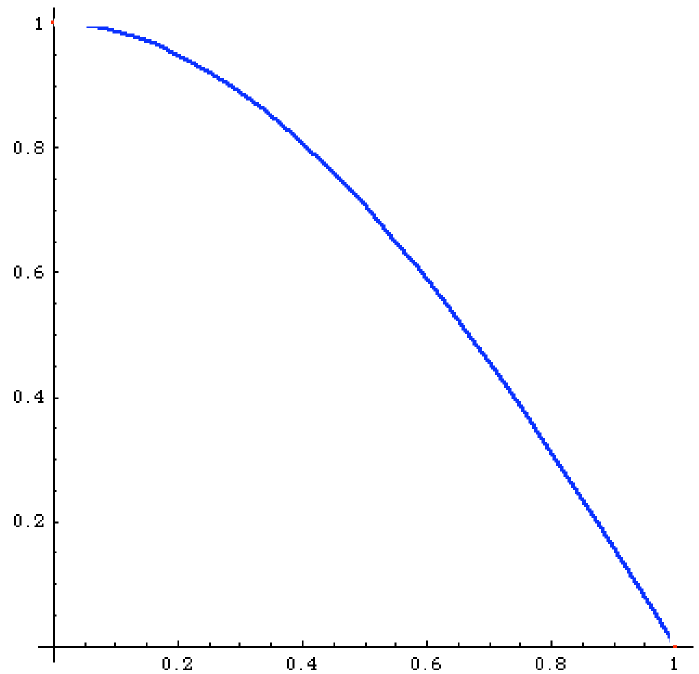
	simpson	MC
3	0.638	0.3
5	0.6367	0.8
20	0.63662	0.6
100	0.636619	0.65
1000	0.636619	0.636

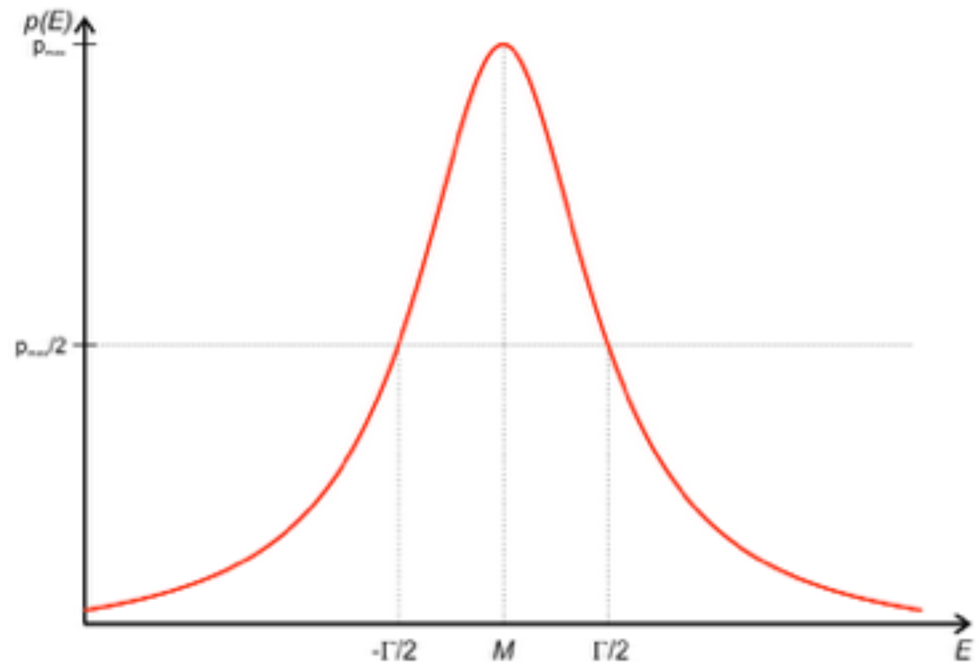
Method of evaluation

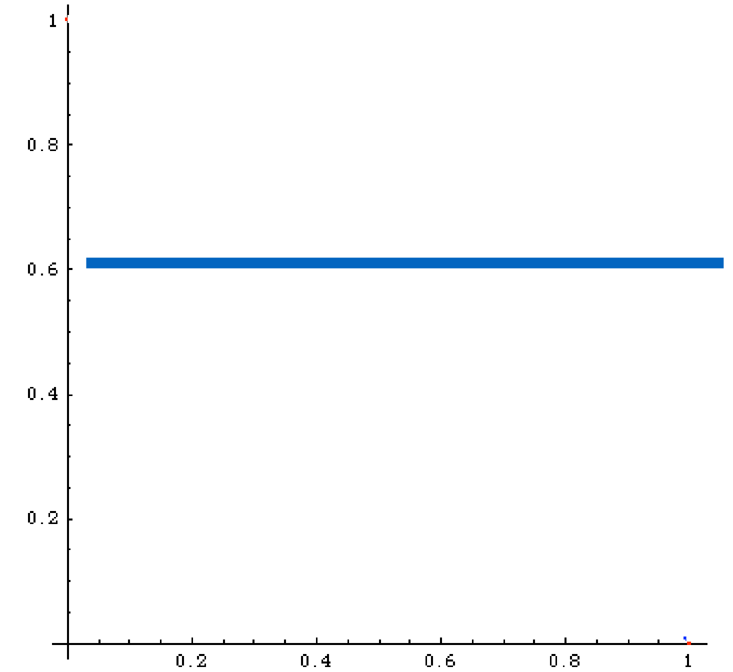
- MonteCarlo $1/\sqrt{N}$
- Trapezium $1/N^2$
- Simpson $1/N^4$

Integration

$$I = \int_0^1 dx \cos\frac{\pi}{2}x \qquad \int \frac{dq^2}{(q^2 - M^2 + iM\Gamma)^2} \qquad \int dx\, C$$







Method of evaluation

- MonteCarlo $1/\sqrt{N}$
- Trapezium $1/N^2$
- Simpson $1/N^4$

More Dimension

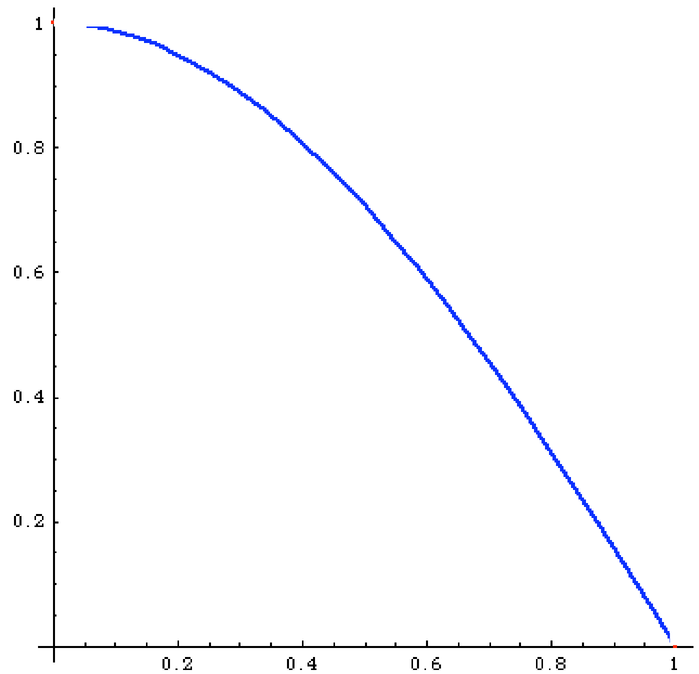
$1/\sqrt{N}$

$1/N^{2/d}$

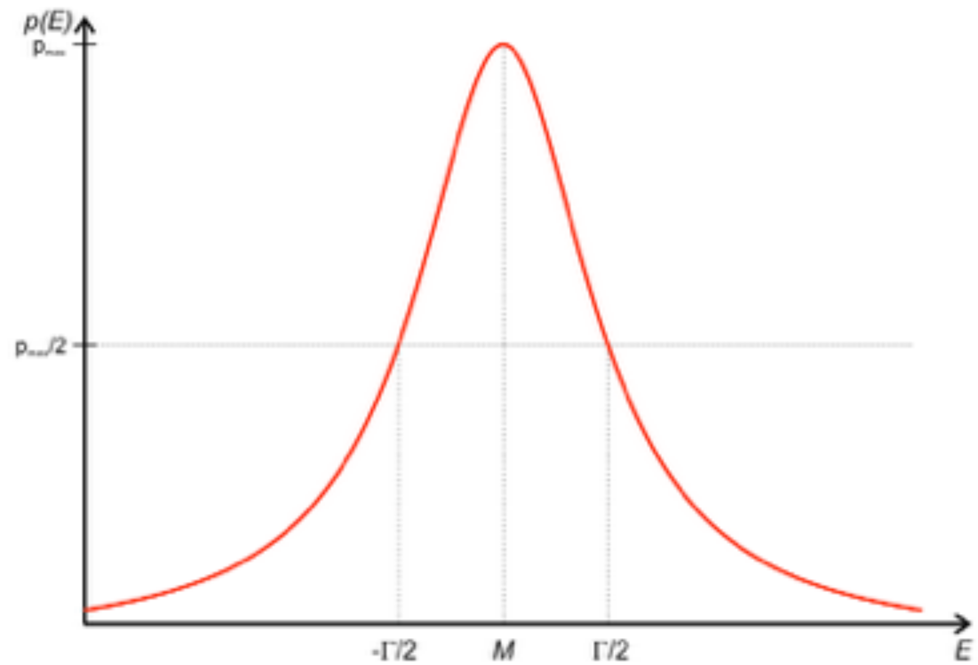
$1/N^{4/d}$

Integration

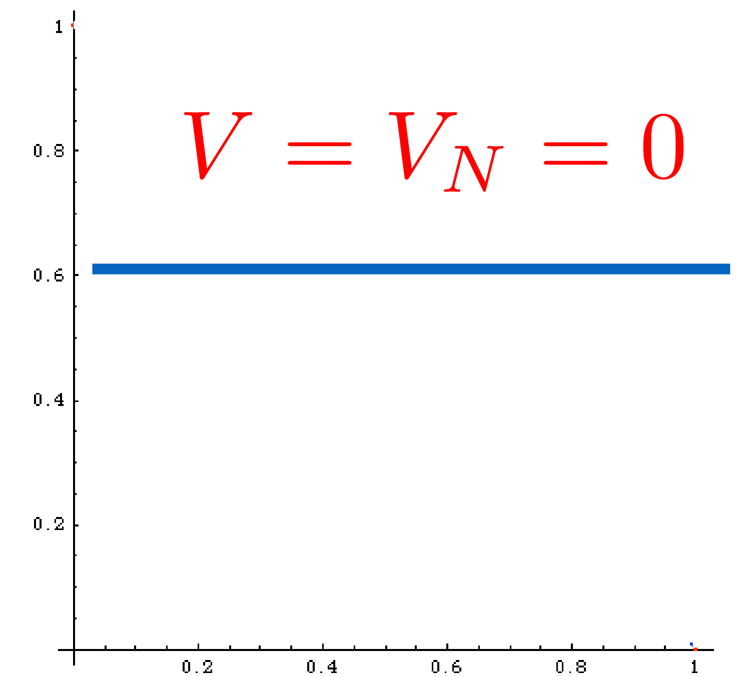
$$I = \int_0^1 dx \cos\frac{\pi}{2}x$$



$$\int \frac{dq^2}{(q^2 - M^2 + iM\Gamma)^2}$$



$$\int dx\, C$$



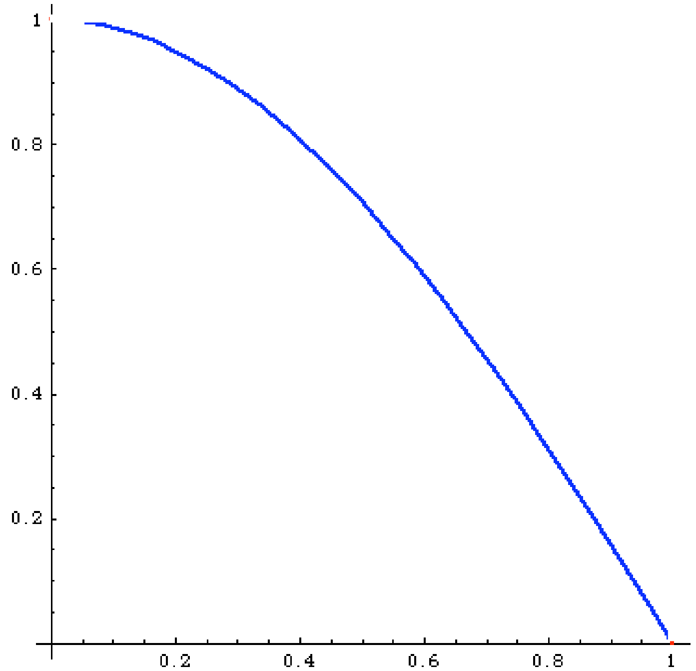
$$I = \int_{x_1}^{x_2} f(x)dx \quad \Rightarrow \quad I_N = (x_2 - x_1)\frac{1}{N}\sum_{i=1}^{N} f(x)$$

$$V = (x_2 - x_1)\int_{x_1}^{x_2} [f(x)]^2 dx - I^2 \quad \Rightarrow \quad V_N = (x_2 - x_1)^2 \frac{1}{N}\sum_{i=1}^{N}[f(x)]^2 - I_N^2$$

$$I = I_N \pm \sqrt{V_N/N}$$

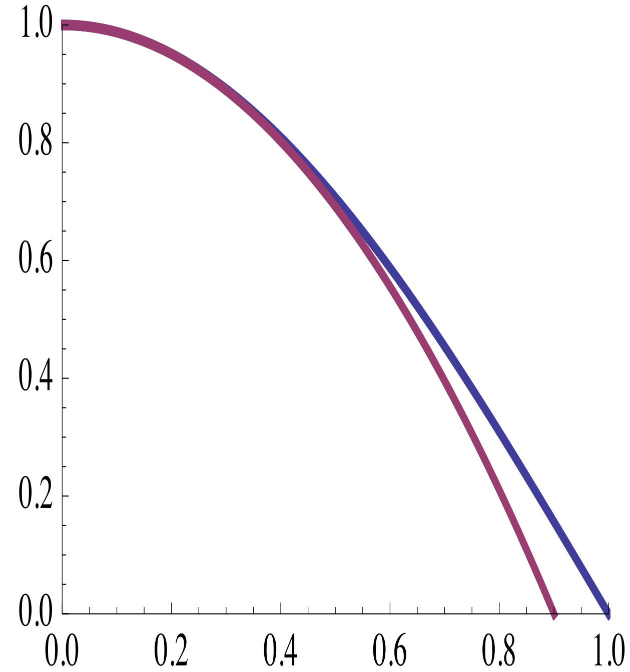
Can be minimized!

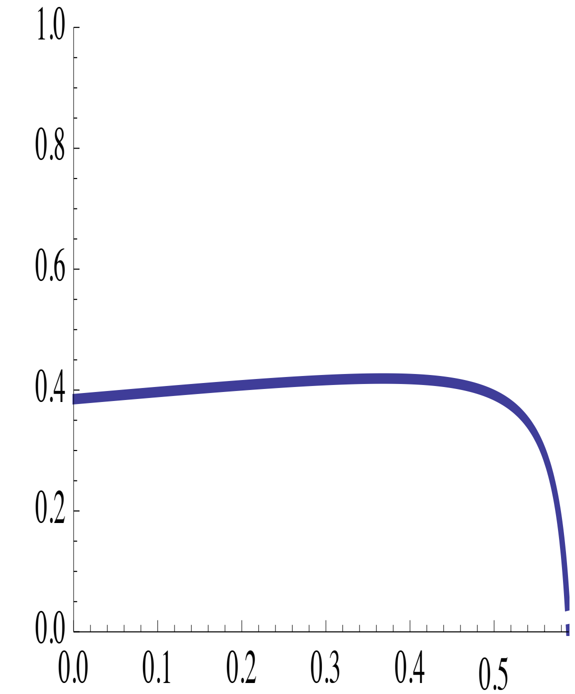
Importance Sampling



$$I = \int_0^1 dx \cos \frac{\pi}{2} x$$

$$I_N = 0.637 \pm 0.307/\sqrt{N}$$





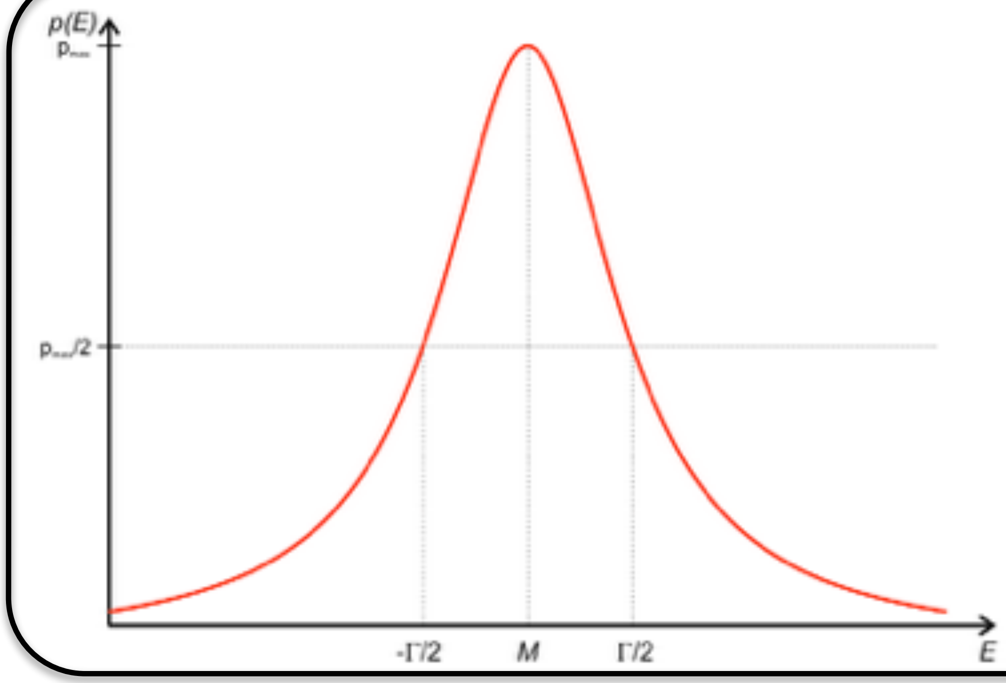
$$I = \int_0^1 dx(1 - cx^2) \frac{\cos\left(\frac{\pi}{2} x\right)}{(1 - cx^2)} = \int_{\xi_1}^{\xi_2} d\xi \frac{\cos \frac{\pi}{2} x[\xi]}{1 - x[\xi]^2 c}$$

$$\simeq 1$$

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

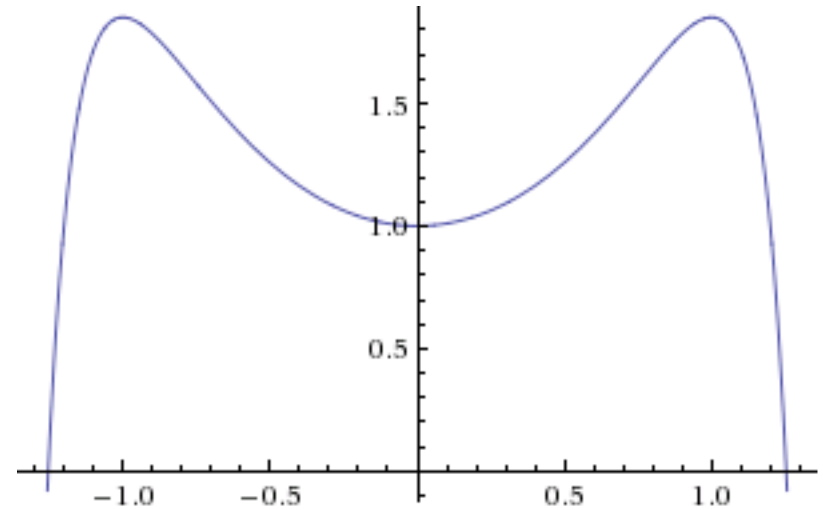
The Phase-Space parametrization is important to have an efficient computation!

Importance Sampling

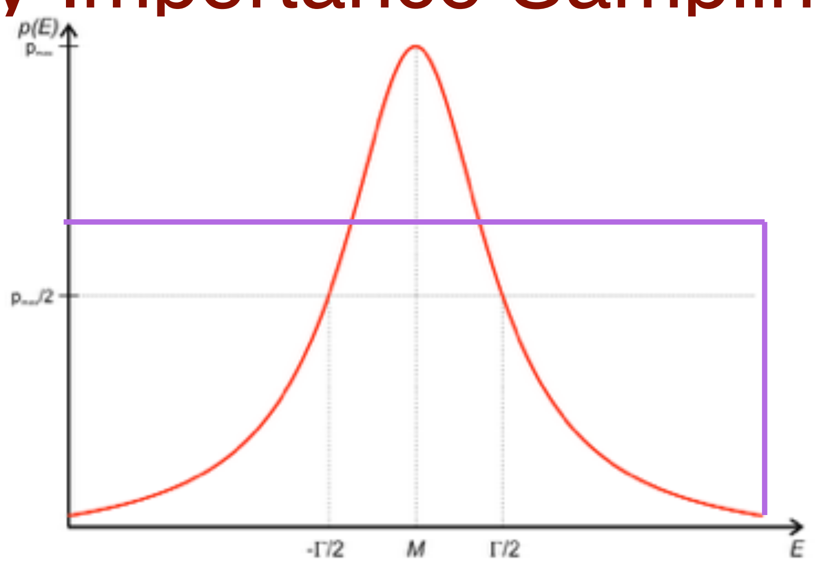


$$\int \frac{dq^2}{(q^2 - M^2 + iM\Gamma)^2}$$

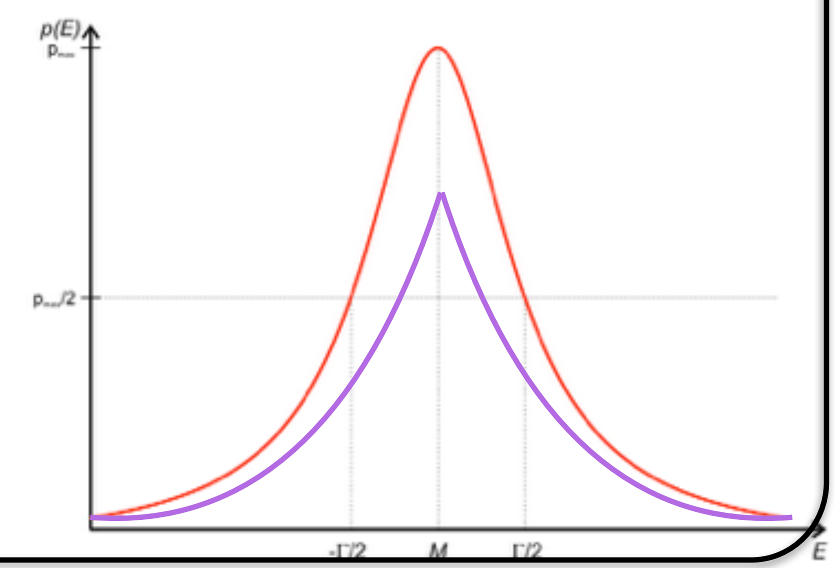
$$\xi = arctan\left(\frac{q^2 - M^2}{\Gamma\, M}\right)$$



Why Importance Sampling?



Probability of using
that point p(x)



The change of variable ensure that the evaluation of the function is done where the function is the largest!

Importance Sampling

Key Point

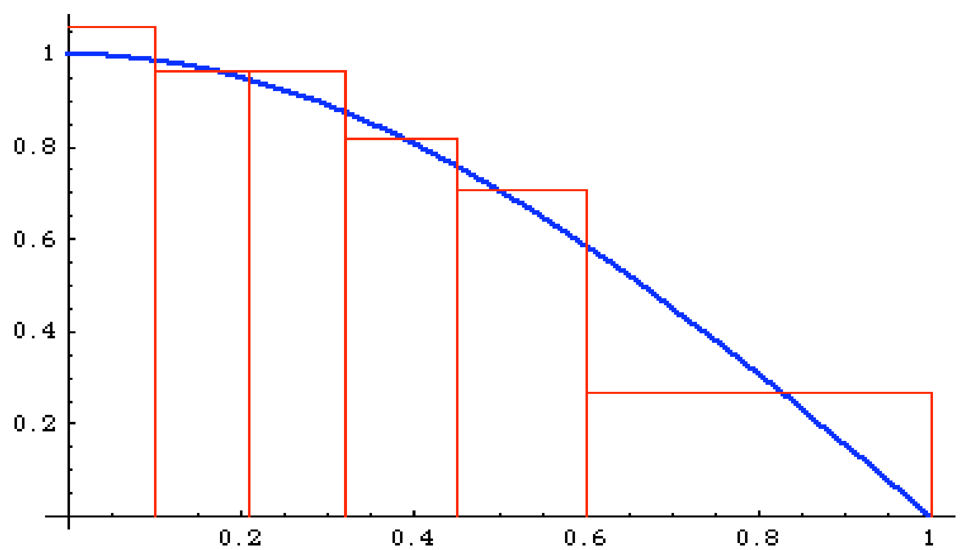
- Generate the random point in a distribution which is close to the function to integrate.
- This is a change of variable, such that the function is flatter in this new variable.
- Needs to know an approximate function.

Adaptative Monte-Carlo

- Create an approximation of the function on the flight!

Adaptative Monte-Carlo

- Create an approximation of the function on the flight!



Algorithm

1. Creates bin such that each of them have the same contribution.

   ➡Many bins where the function is large

2. Use the approximate for the importance sampling method.

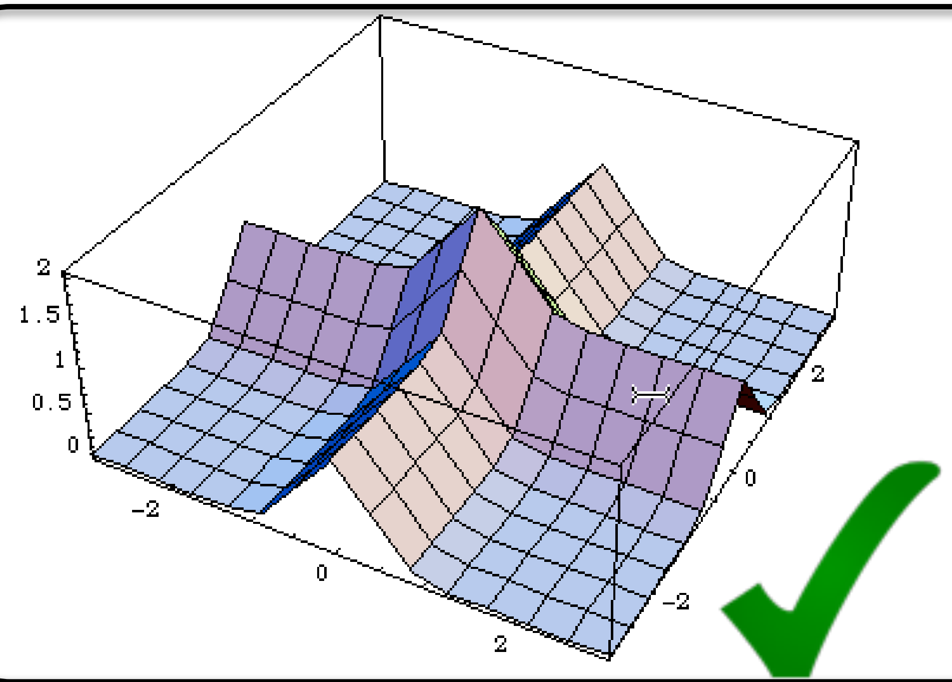
More than one Dimension

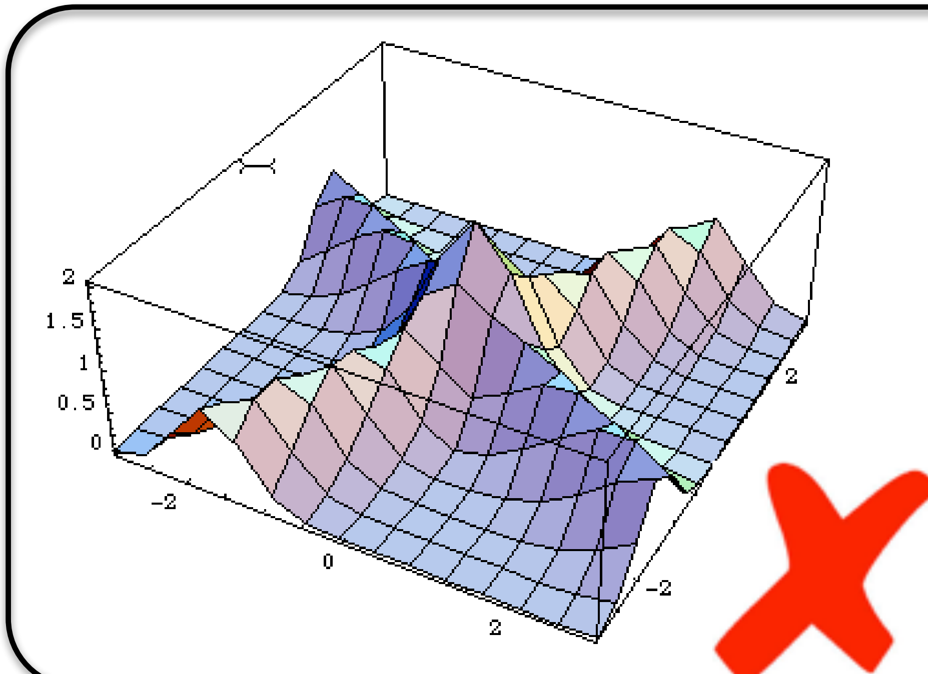
- VEGAS works only with 1(few) dimension
  - ➡memory problem

Solution

- Use projection on the axis

$$\vec{p}(x) = p(x) \cdot p(y) \cdot p(z) \ldots$$



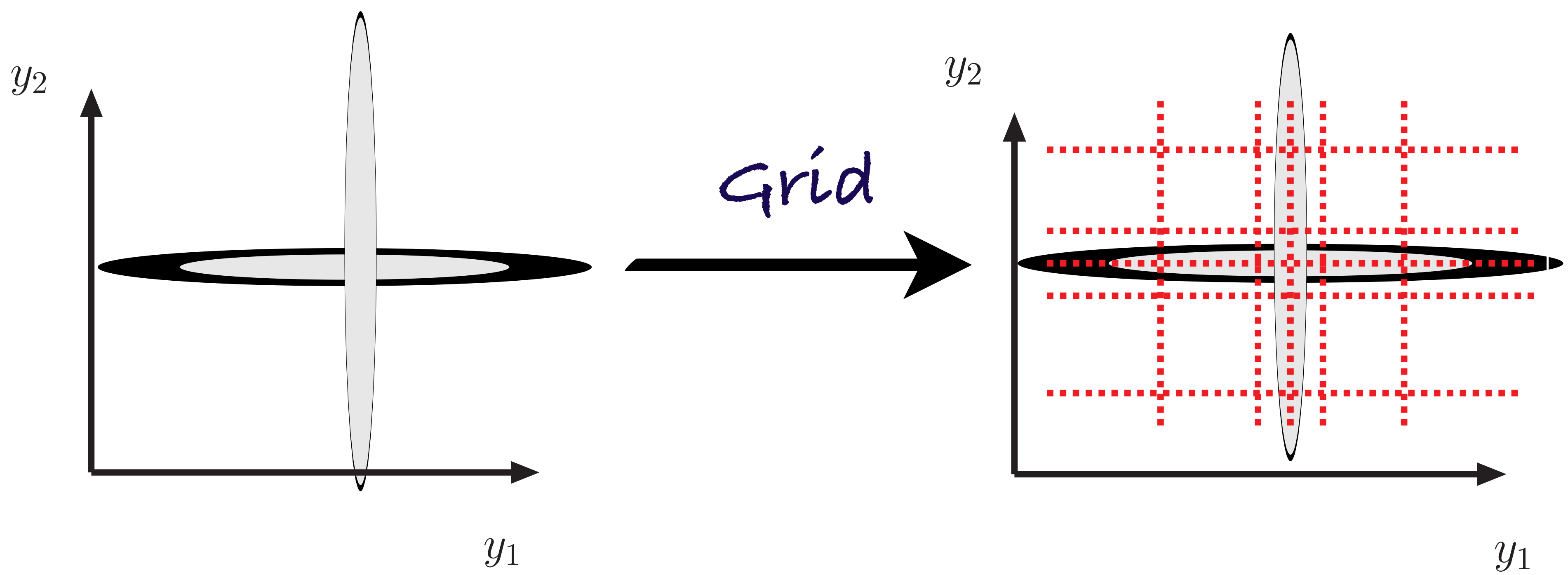


- We need to ensure the factorization !

➡Additional change of variable

Monte-Carlo Integration

- The choice of the parameterisation has a strong impact on the efficiency

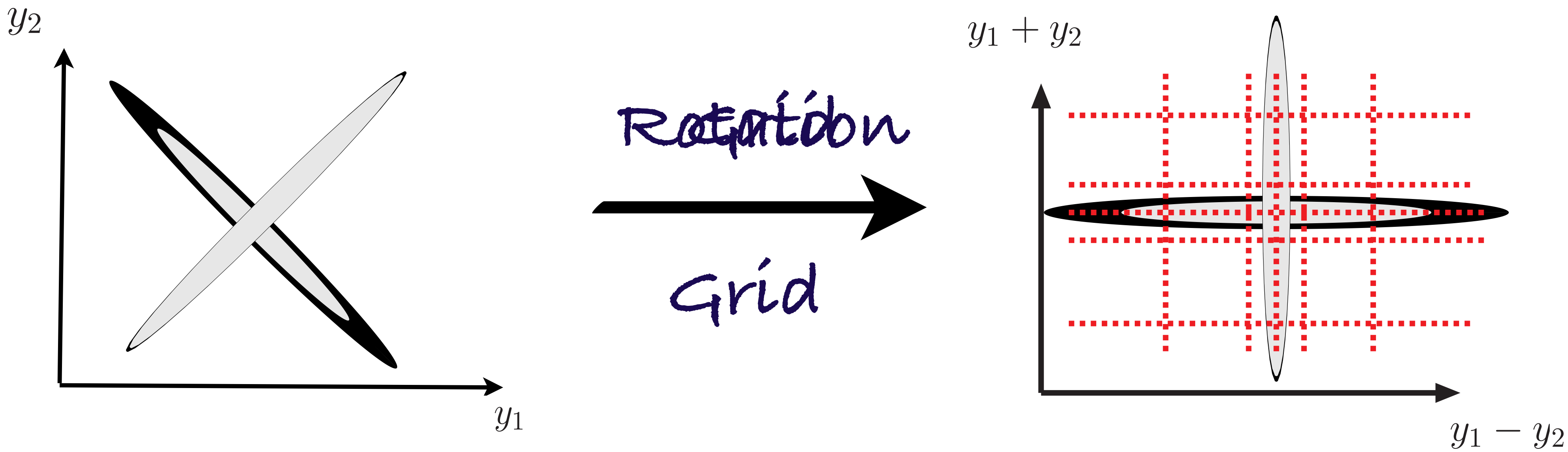


- ☐ The adaptive Monte-Carlo Technique picks point in interesting areas

→ The technique is efficient

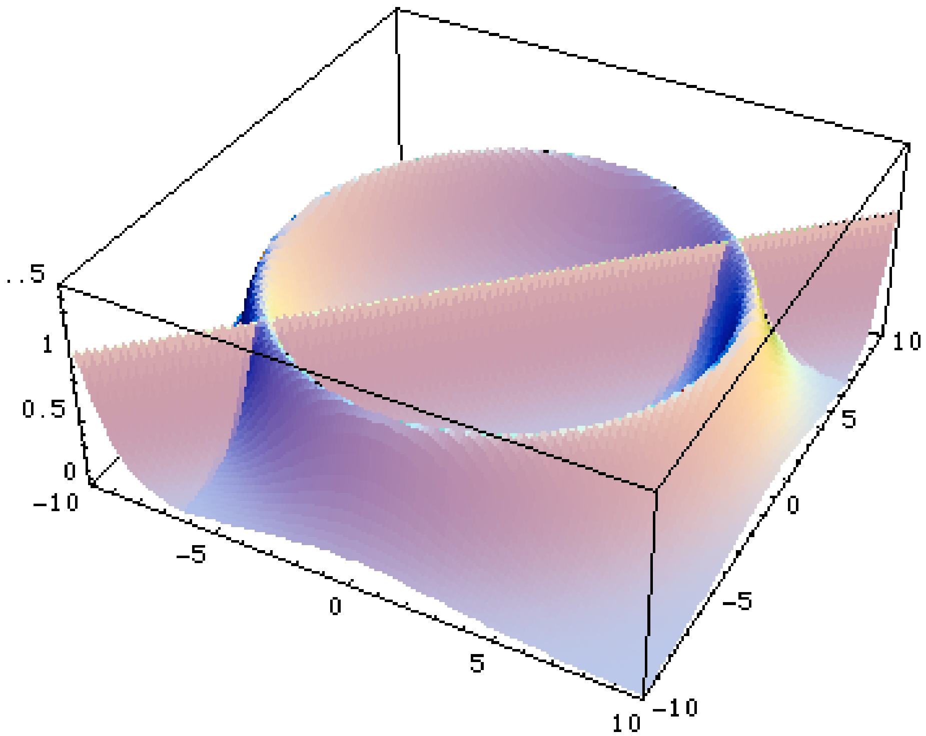
Monte-Carlo Integration

- The choice of the parametrization has a strong impact on the efficiency



- The adaptive Monte-Carlo Techniques picks point
  points everywhere interesting areas
  → The technique is efficient converges slowly

Multi-channel



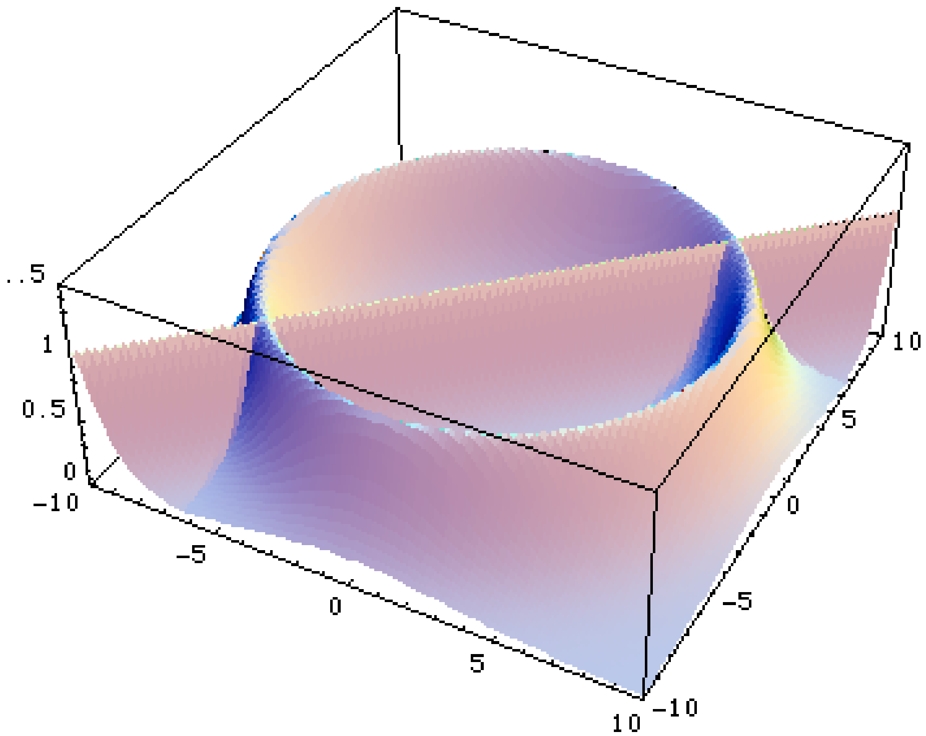
What do we do if there is no transformation that aligns all integrand peaks to the chosen axes? Vegas is bound to fail!

Solution: use different transformations = channels

$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x) \qquad \text{with} \qquad \sum_{i=1}^{n} \alpha_i = 1$$

with each $p_i(x)$ taking care of one "peak" at the time

Multi-channel

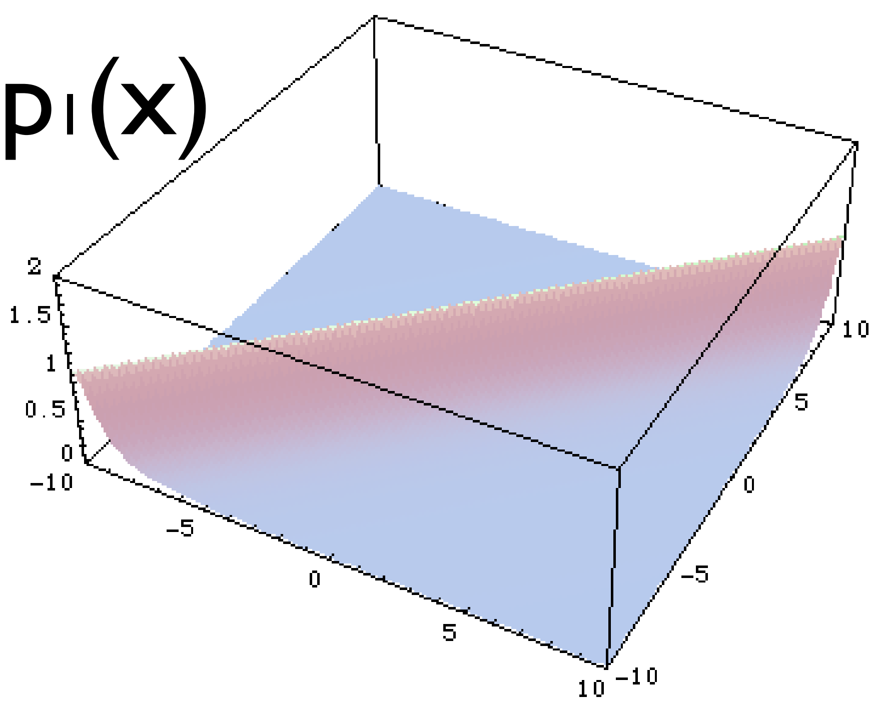


$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x)$$

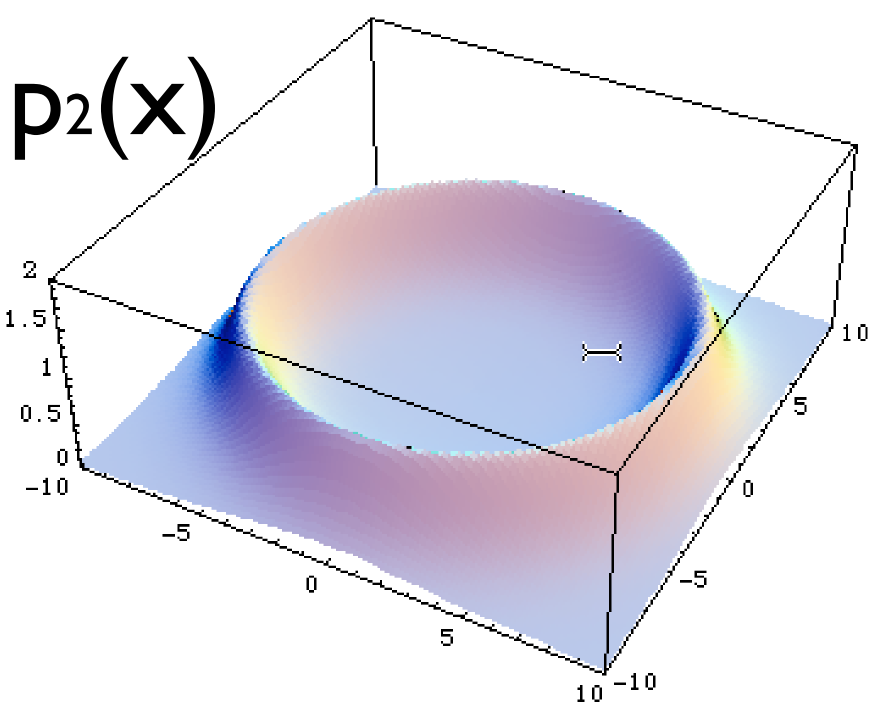
with

$$\sum_{i=1}^{n} \alpha_i = 1$$

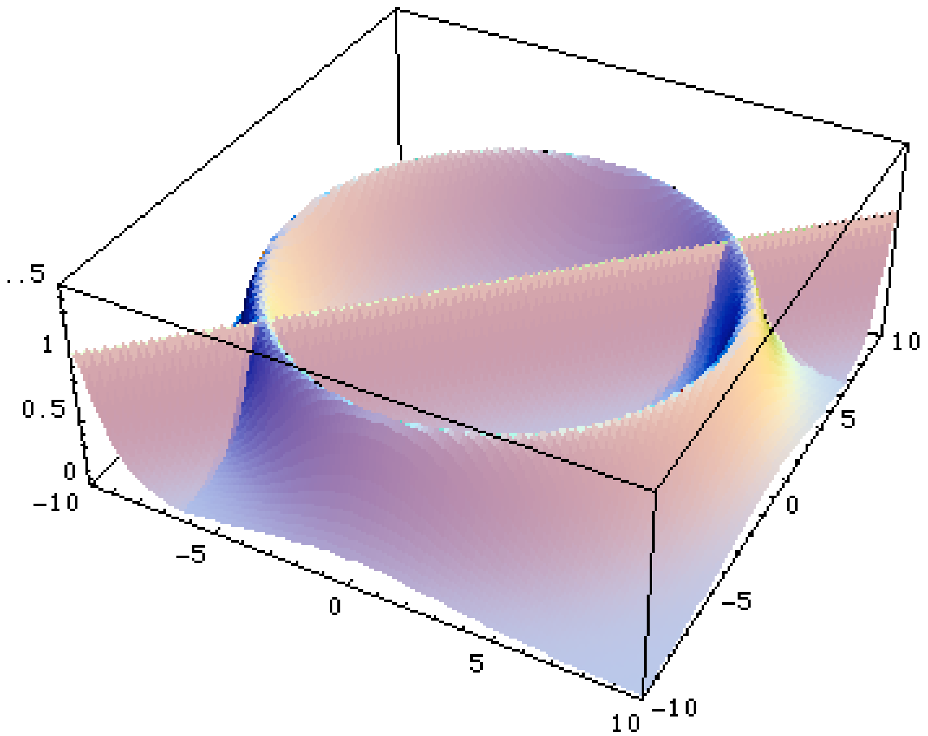
$p_1(x)$



$p_2(x)$



Multi-channel



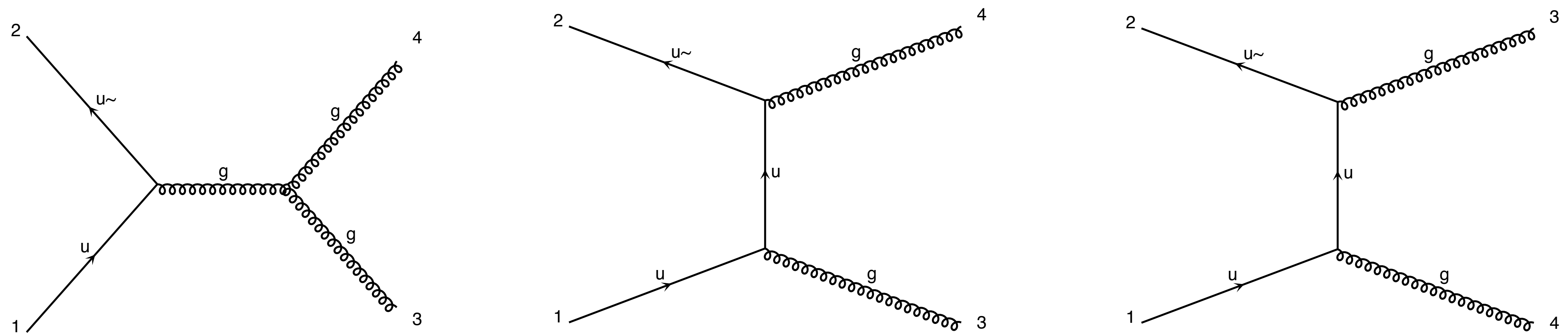
$$p(x) = \sum_{i=1}^{n} \alpha_i p_i(x)$$

with $$\sum_{i=1}^{n} \alpha_i = 1$$

Then,

$$I = \int f(x)dx = \sum_{i=1}^{n} \alpha_i \int \underbrace{\frac{f(x)}{p(x)}}_{\approx 1} p_i(x)dx$$

- Phase-Space integration
- Parton Shower
- Hadronization



$$\propto \frac{1}{\hat{s}} = \frac{1}{(p_1+p_2)^2} \qquad \propto \frac{1}{\hat{t}} = \frac{1}{(p_1-p_3)^2} \qquad \propto \frac{1}{\hat{u}} = \frac{1}{(p_1-p_4)^2}$$

Three very different pole structures contributing to the same matrix element.

Multi-channel based on single diagrams*

*Method used in MadGraph

Does a basis exist?

$$\int |M_{tot}|^2 = \int \frac{\sum_i |M_i|^2}{\sum_j |M_j|^2} |M_{tot}|^2 = \sum_i \int \frac{|M_i|^2}{\boxed{\sum_j |M_j|^2} } \boxed{|M_{tot}|^2}$$

≈ 1

Key Idea

- Any single diagram is "easy" to integrate (pole structures/ suitable integration variables known from the propagators)
- Divide integration into pieces, based on diagrams
- All other peaks taken care of by denominator sum

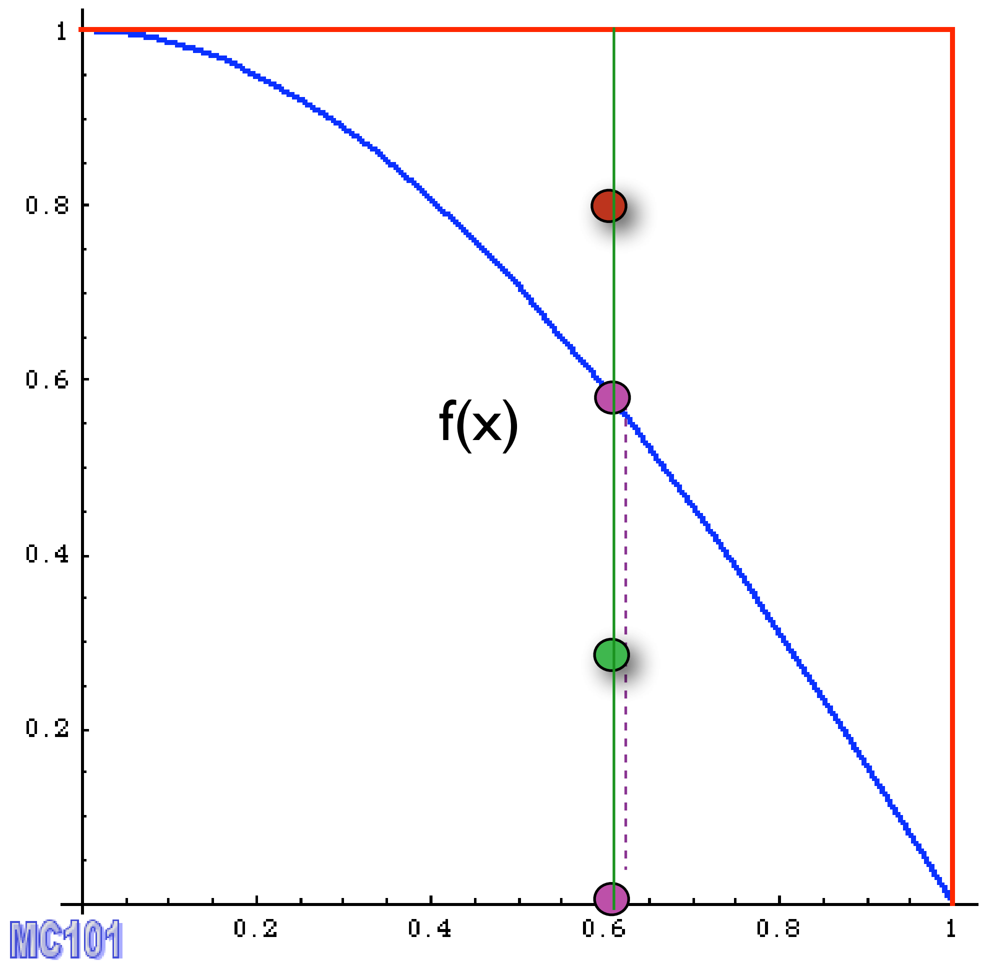
N Integral

- Errors add in quadrature so no extra cost
- "Weight" functions already calculated during $|M|^2$ calculation
- Parallel in nature

To Remember

- Phase-Space integration are difficult
- We need to know the function
  - ➡ Be careful with cut (they change the function)
- Split the function in a sum (one for each structure) and integrate each of those separately
  - ➡ This splitting should not be physical

Event generation

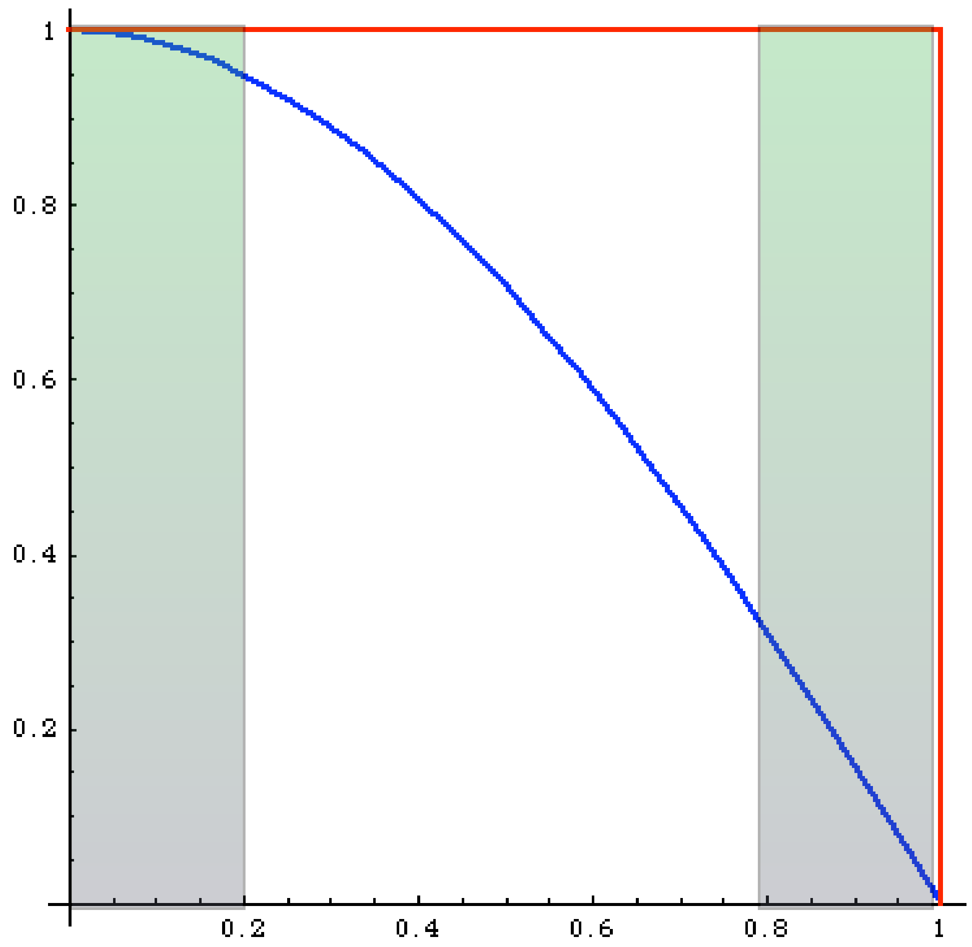


1. pick x
2. calculate f(x)
3. pick 0<y<fmax
4. Compare:
   if f(x)>y accept event,

   else reject it.

$$I = \frac{\text{accepted}}{\text{total tries}} = \text{efficiency}$$

Event generation

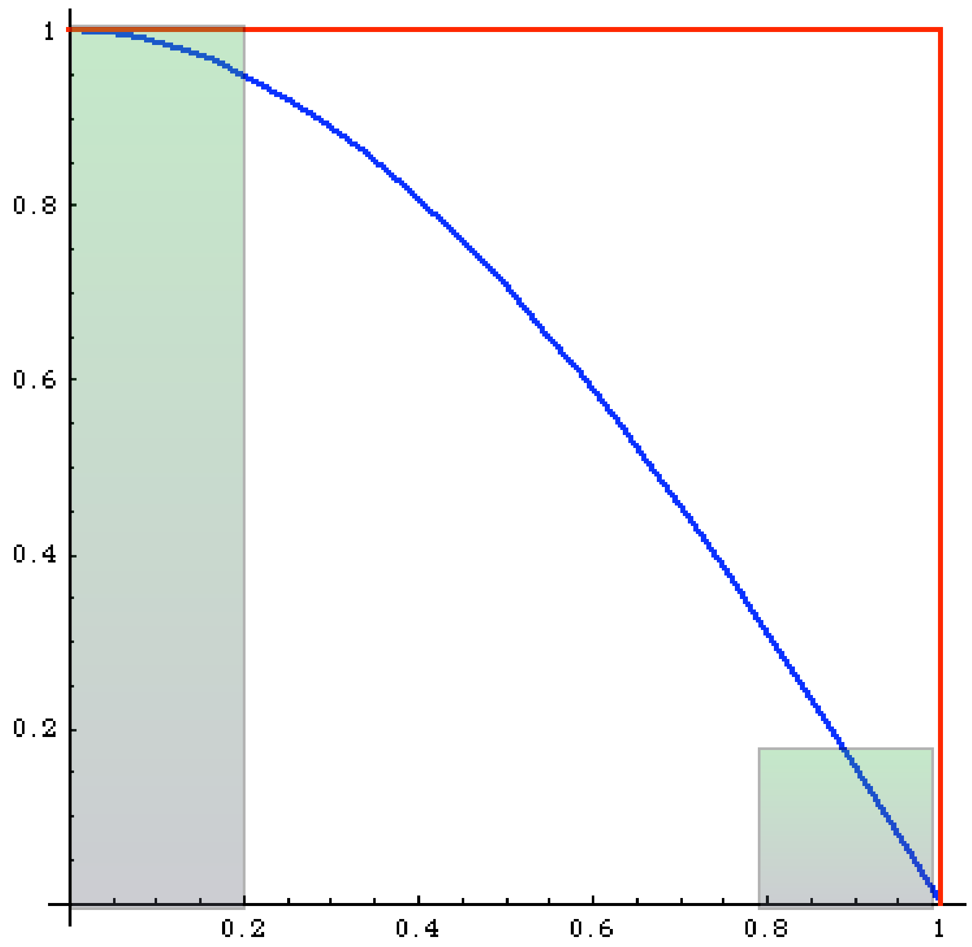


What's the difference between weighted and unweighted?

Weighted:

Same # of events in areas of phase space with very different probabilities: events must have different weights

Event generation



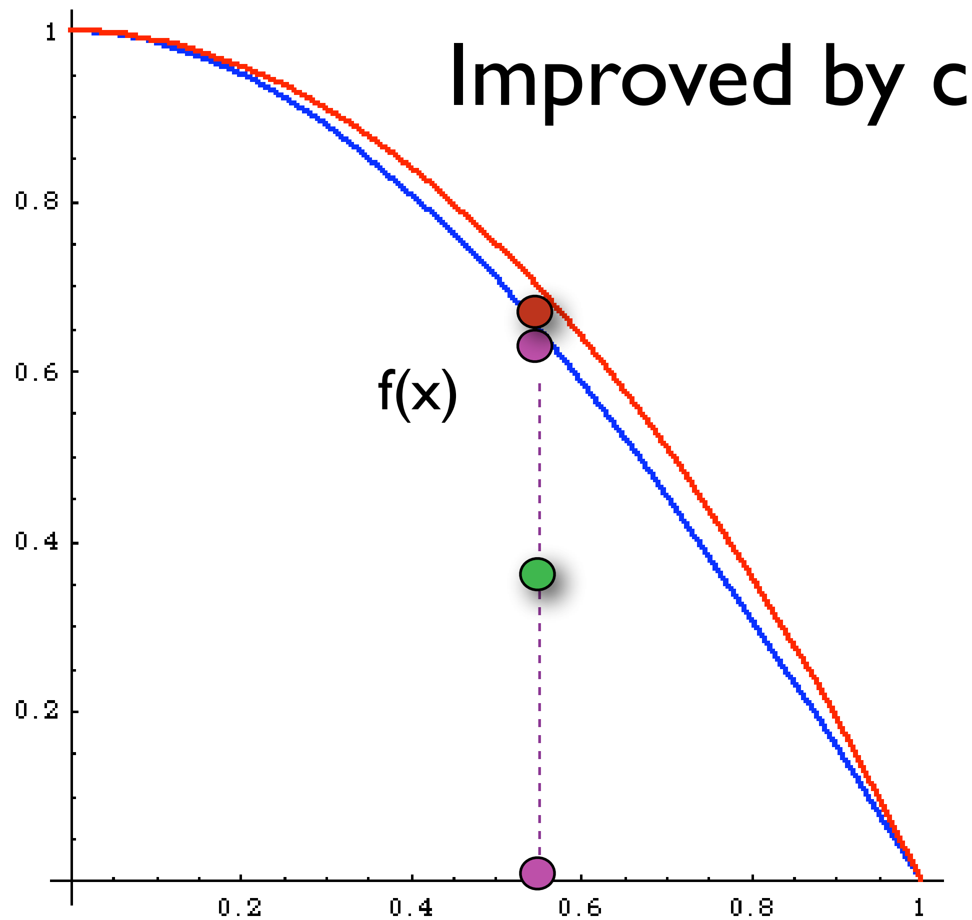
What's the difference between weighted and unweighted?

Unweighted:

events is proportional to the probability of areas of phase space:
events have all the same weight ("unweighted")

Events distributed as in nature

Event generation

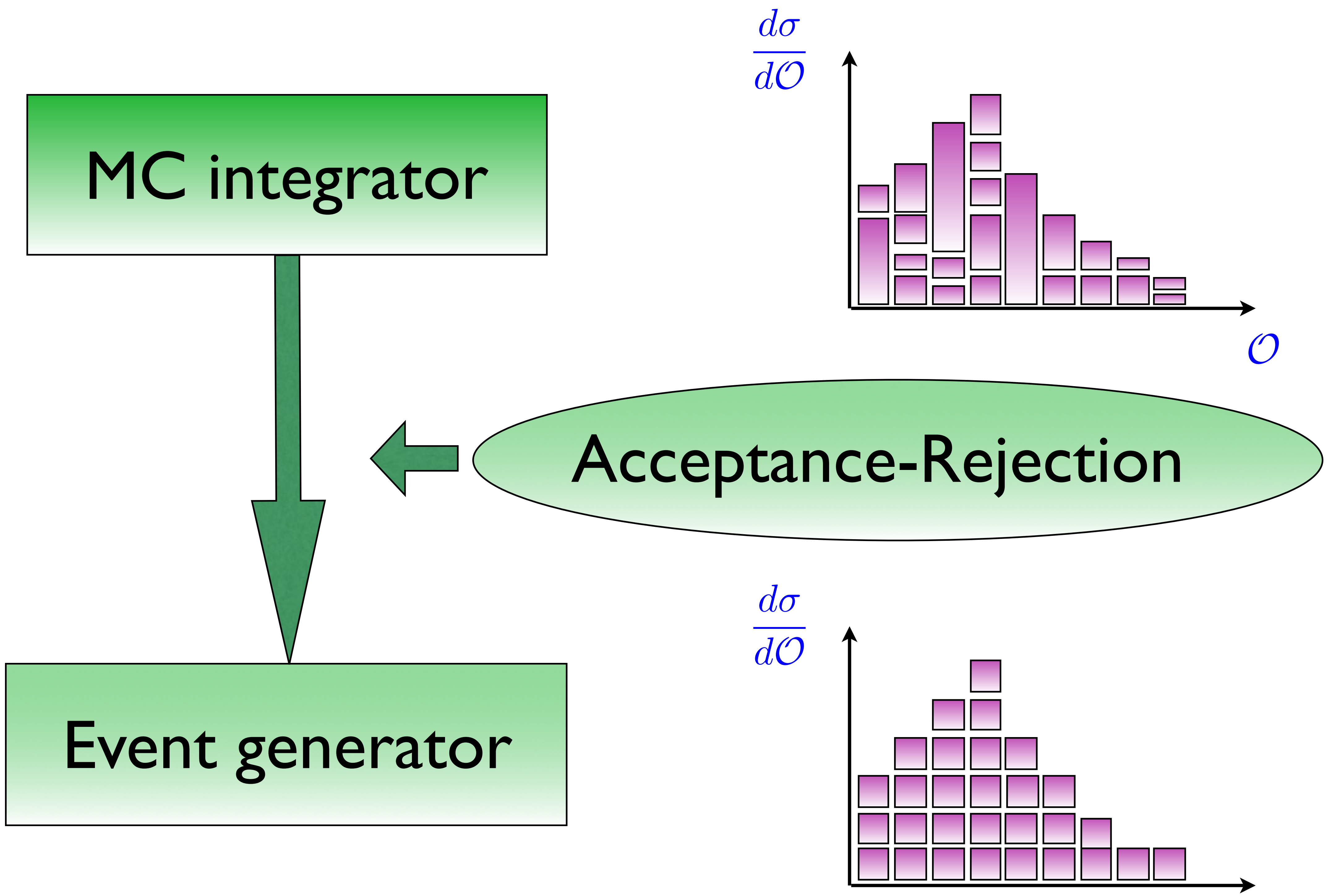


Improved by combining with importance sampling:

1. pick x distributed as p(x)
2. calculate f(x) and p(x)
3. pick 0<y<f/p(max)
4. Compare:
   if f(x)>y p(x) accept event,
   else reject it.

much better efficiency!!!

Event generation



☞ **This is possible only if f(x)<∞ AND has definite sign!**

To Remember

- Sample of unweighted events
  - ➡ Events distributed like nature
  - ➡ Need the function to be
    - ❑ Borned
    - ❑ Always positive
  - ➡ More efficient if the integration is more efficient
    - ❑ Same dependencies in the cut

Monte-Carlo Summary

Bad Point

- Slow Convergence (especially in low number of Dimension
- Need to know the function
  - Impact on cut

Good Point

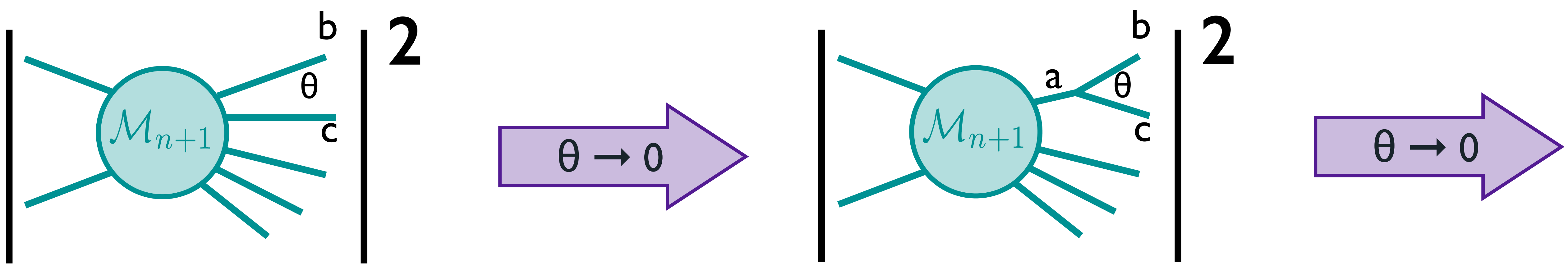
- Complex area of Integration
- Easy Error estimate
- quick estimation of the integral
- Possibility to have unweighted events

Type of MC Simulation

Parton shower

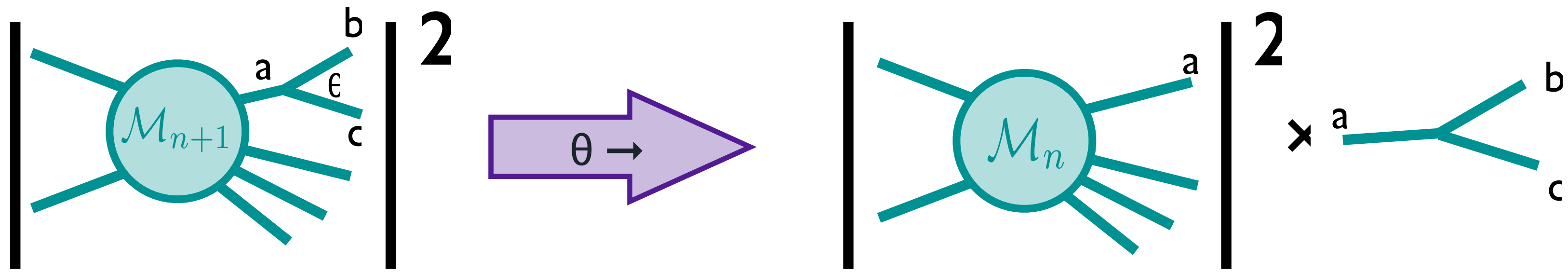
- We need to be able to describe an arbitrarily number of parton branchings, i.e. we need to 'dress' partons with radiation
- This effect should be **unitary:** the inclusive cross section shouldn't change when extra radiation is added
- Remember that parton-level cross sections for a hard process are inclusive in anything else.
  E.g. for LO Drell-Yan production **all** radiation is included via PDFs (apart from non-perturbative power corrections)
- And finally we want to turn partons into hadrons (hadronization)....

Collinear factorization



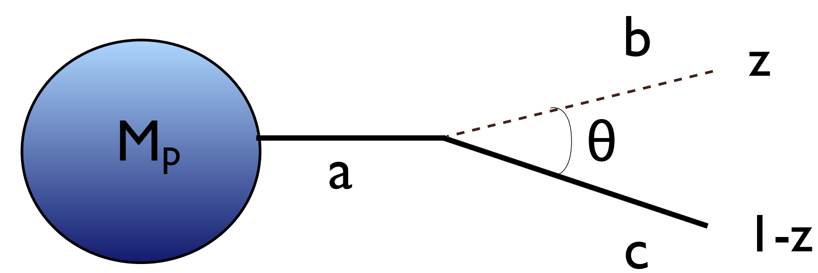
- Consider a process for which two particles are separated by a small angle θ.
- In the limit of θ → 0 the contribution is coming from a single parent particle going on shell: therefore its branching is related to time scales which are very long with respect to the hard subprocess.
- The inclusion of such a branching cannot change the picture set up by the hard process: the whole emission process must be writable in this limit as the simpler one times a branching probability.

Collinear factorization



- The process factorizes in the collinear limit. This procedure it universal!

$$\frac{1}{(p_b+p_c)^2} \simeq \frac{1}{2E_bE_c(1-\cos\theta)} = \frac{1}{t}$$



$z = E_b/E_a$

soft and collinear

divergencies

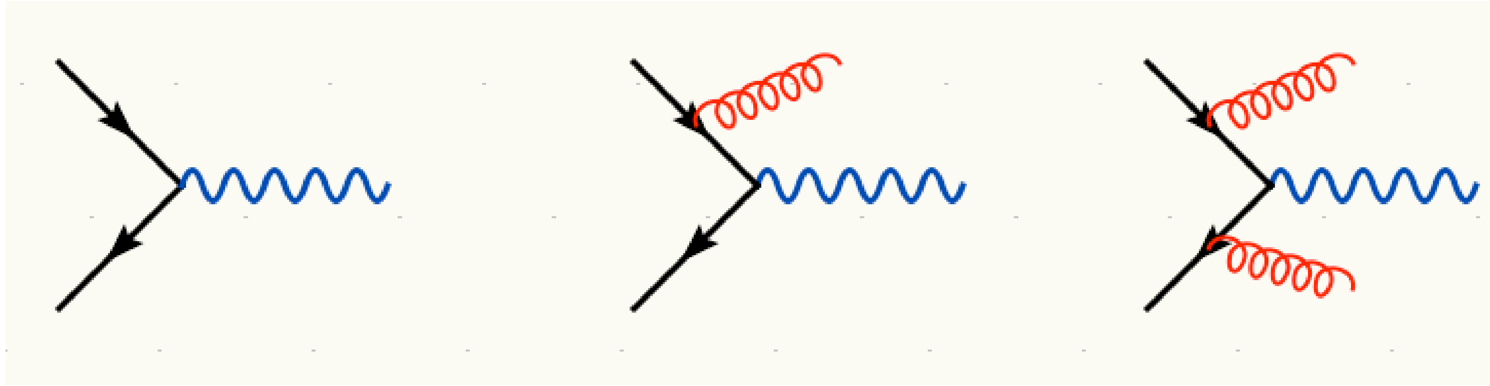
Collinear factorization:

$$|\mathcal{M}_{n+1}|^2 d\Phi_{n+1} \;\simeq\; |\mathcal{M}_n|^2 d\Phi_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_S}{2\pi} P_{a\to bc}(z)$$

when θ is small.

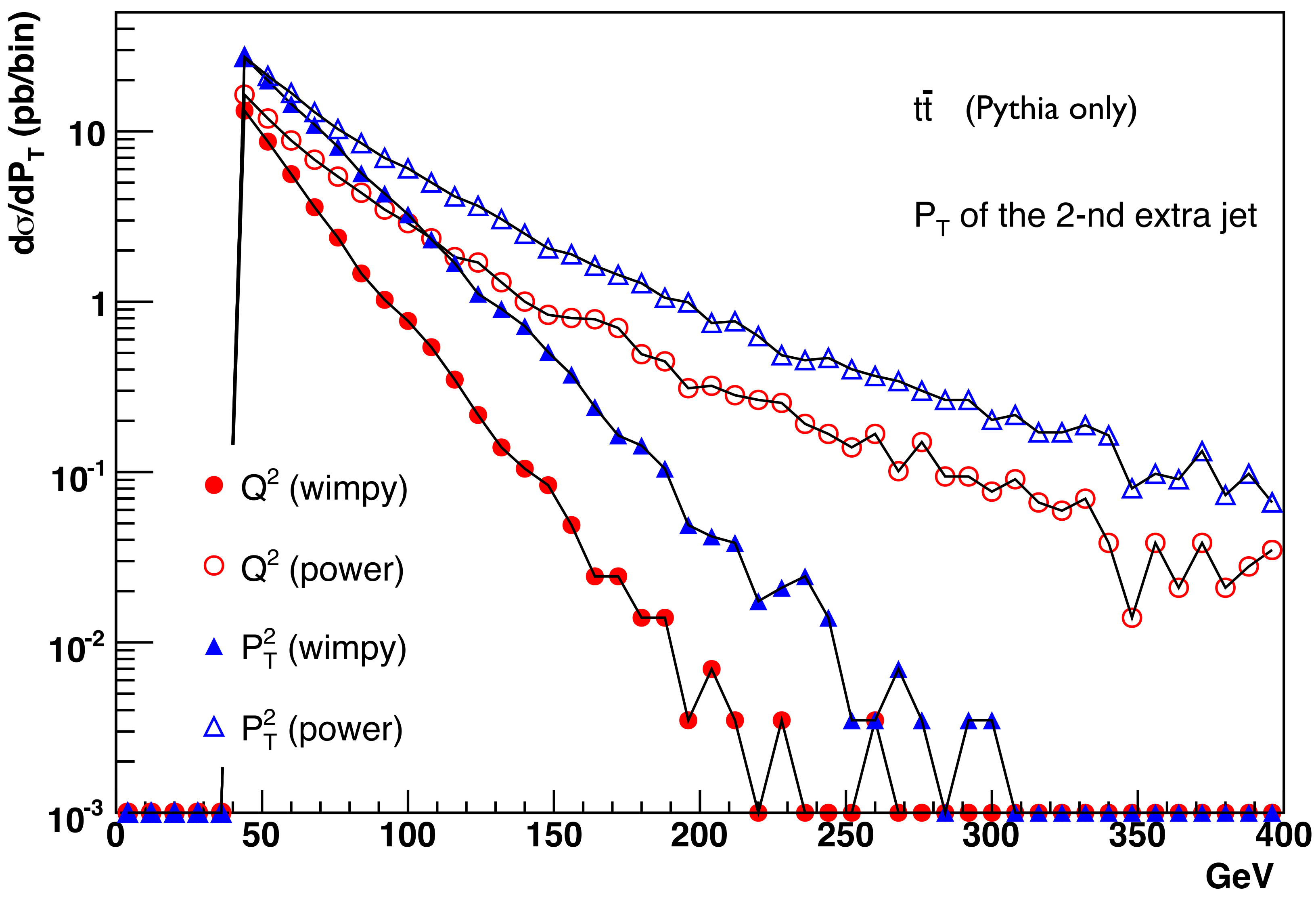
Merging ME with PS

PS →



PS alone vs matched samples

In the soft-collinear approximation of Parton Shower MCs, parameters are used to tune the result ⇒ Large variation in results (small prediction power)



Matrix Elements vs. Parton Showers

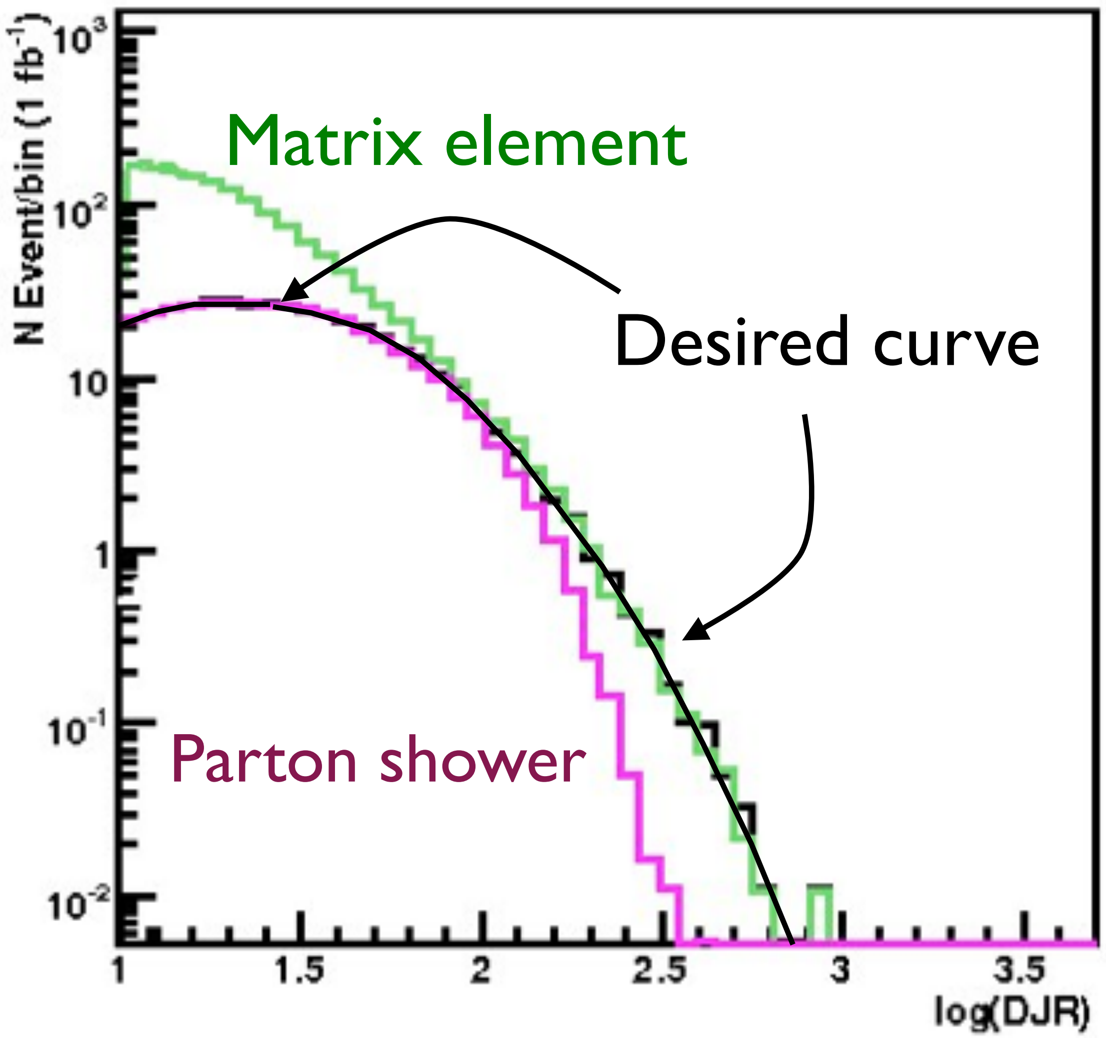
ME	Shower MC
1. Fixed order calculation 2. Computationally expensive 3. Limited number of particles 4. Valid when partons are hard and well separated 5. Quantum interference correct 6. Needed for multi-jet description	1. Resums logs to all orders 2. Computationally cheap 3. No limit on particle multiplicity 4. Valid when partons are collinear and/or soft 5. Partial interference through angular ordering 6. Needed for hadronization

Approaches are complementary: merge them!

Difficulty: avoid double counting, ensure smooth distributions

Goal for ME-PS merging/matching

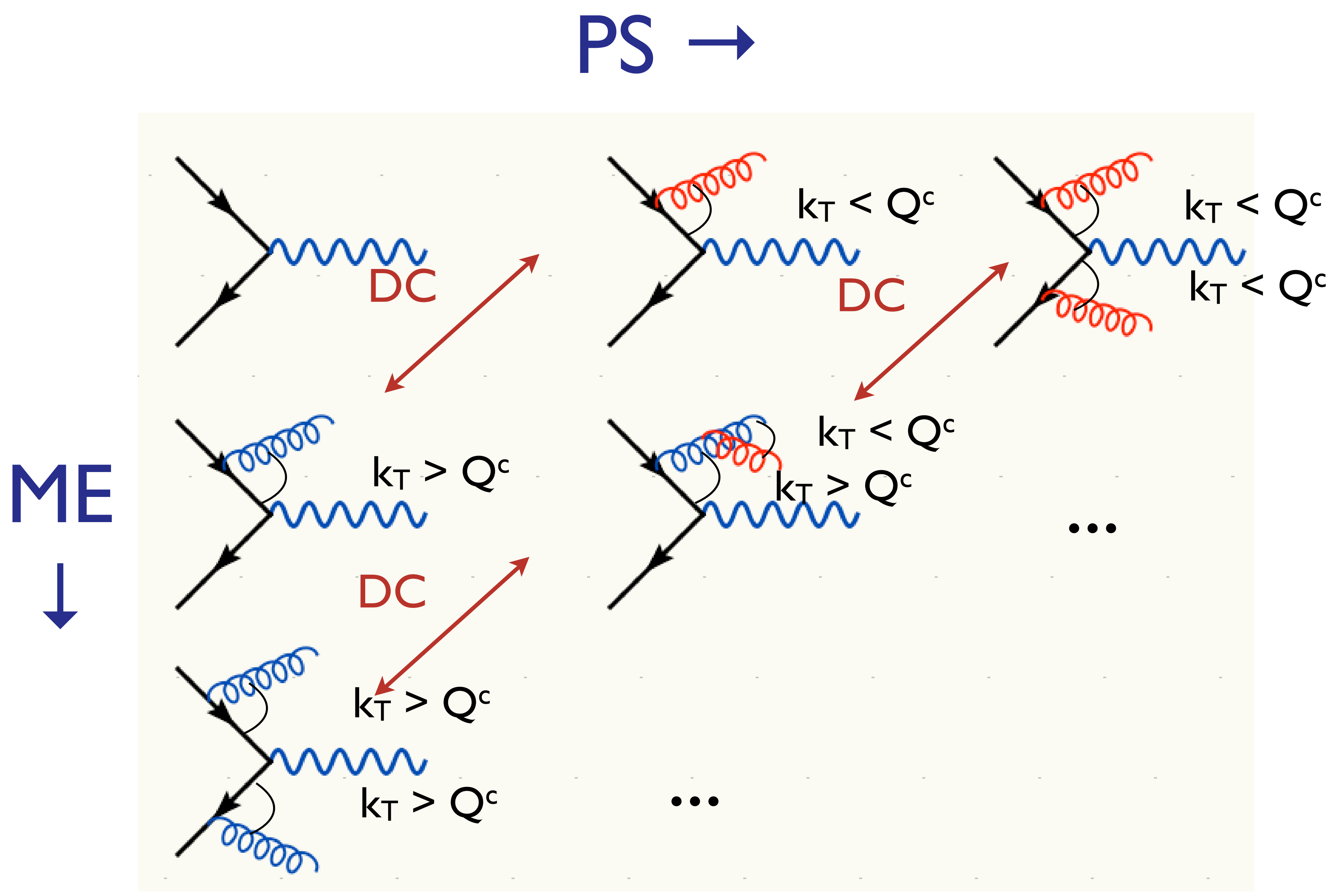
- Regularization of matrix element divergence
- Correction of the parton shower for large momenta
- Smooth jet distributions



2nd QCD radiation jet in top pair production at the LHC, using MadGraph + Pythia

Merging ME with PS

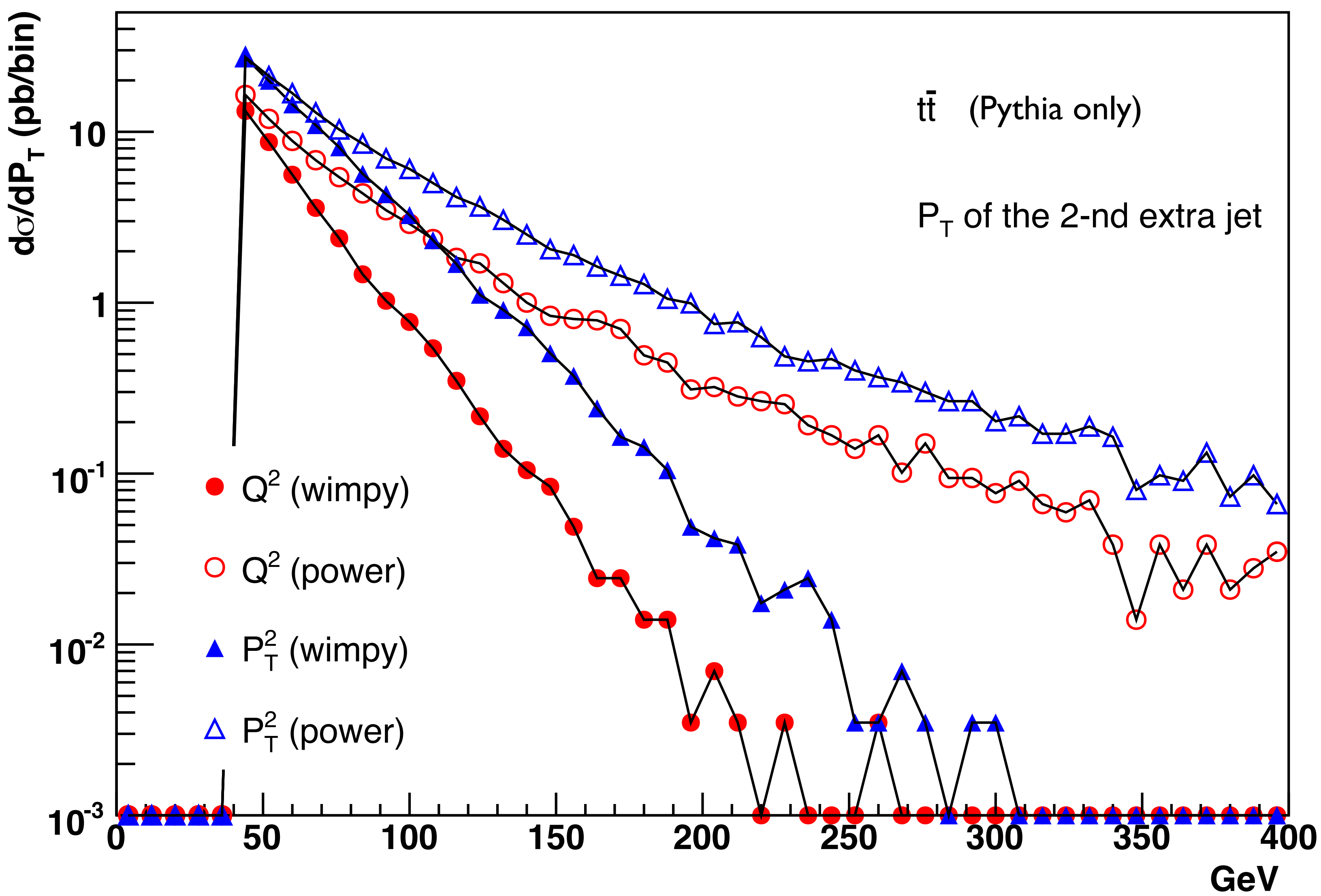
[Mangano]
[Catani, Krauss, Kuhn, Webber]
[Lönnblad]



Double counting between ME and PS easily avoided using phase space cut between the two: PS below cutoff, ME above cutoff.

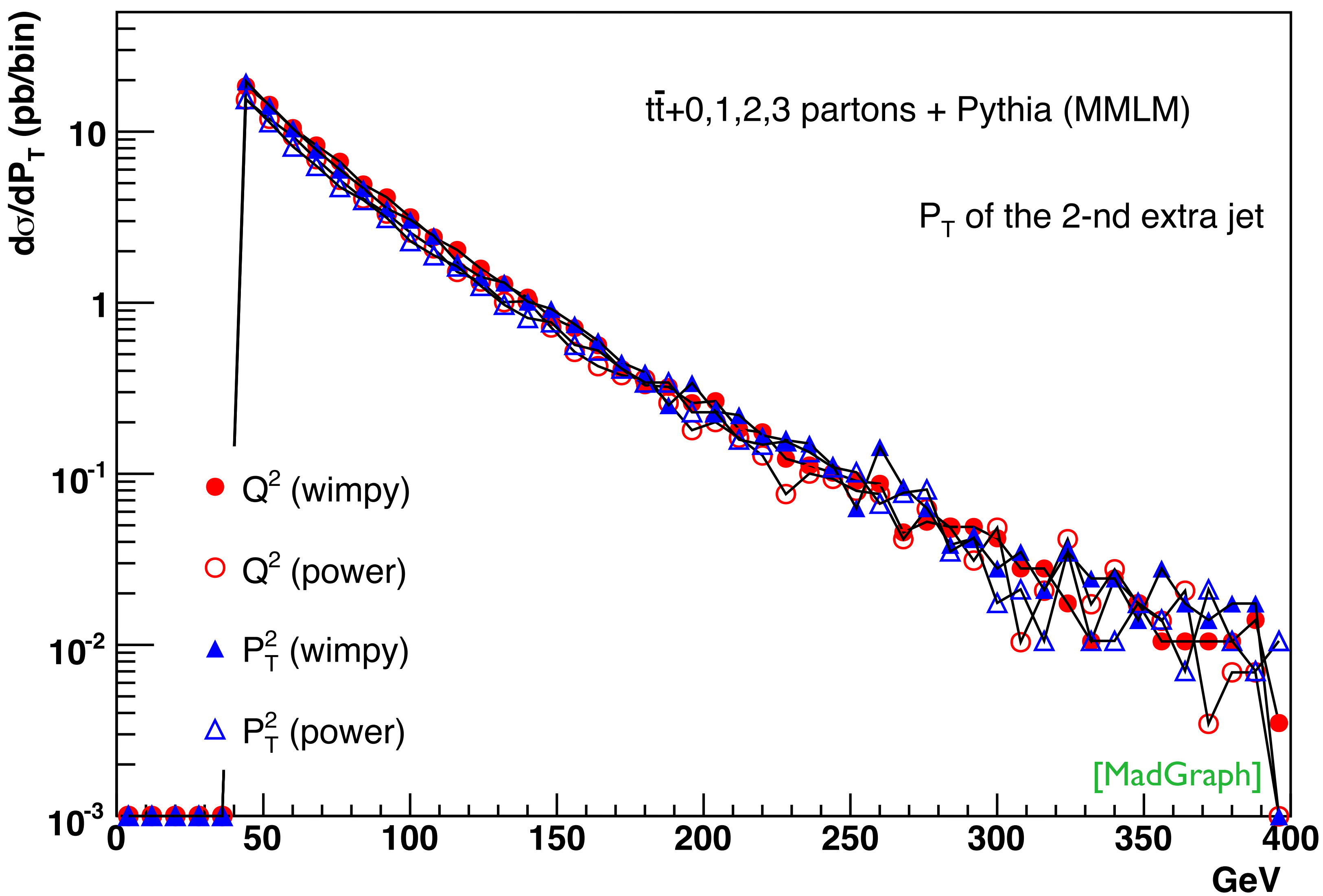
PS alone vs matched samples

In the soft-collinear approximation of Parton Shower MCs, parameters are used to tune the result ⇒ Large variation in results (small prediction power)



PS alone vs ME matching

In a matched sample these differences are irrelevant since the behavior at high pt is dominated by the matrix element.



Tools for MC Simulation

Which kind of MC?

- LO
  - ➡ fix order (plus parton-shower)
  - ➡ matched-merged
- NLO
  - ➡ POWHEG / MC@NLO
  - ➡ merged sample
- NNLO / re-summation / N3LO
- Default:
  - ➡ Do the most advanced possible generation.
  - ➡ Speed issue? check faster possibilities

NLO: merged sample

MadGraph5_aMC@NLO | Sherpa | ... + Matchbox

- NLO in QCD
- Only SM support
- free list of process

MadGraph5_aMC@NLO and Sherpa:

- MC@NLO method

... + Matchbox:

- MC@NLO and POWHEG

MadGraph5_aMC@NLO:

- Need to link to a Shower program (Pythia8)
- FxFx / UnLOPS Merging

Sherpa:

- Need to link the one loop tool
- UNLOPS merging

... + Matchbox:

- Need to provide events
- Need to link to matrix element (both tree and loop)

NLO (one multiplicity)

POWHEG

VBF@NLO

- Fixed list of processes
- Some BSM

- POWHEG matching
  - Less negative events
  - Not pure NLO

- QED@NLO
- very dedicated

MadGraph5_aMC@NLO

- BSM possible

Sherpa

- SM

NLO (cross-section)

MCFM

- Fixed list of processes
- Some BSM
- No events generation

MadGraph5_aMC@NLO

Sherpa

Powheg

VBF@NLO

HPAIR

NJETS

LO (matched/merged)

Sherpa

- Fully built in
- Starts BSM supports
- CKKW-L
- ATLAS Default

… + Pythia

- MLM / UMEPS / CKKW CKKWL
- Full BSM supports
- CMS default (with MG5_aMC)

… + Herwig

- MLM /CKKW CKKWL
- Full BSM supports

LO

CODE	Main advantage	highest multiplicity
MG5_aMC	BSM	normal: 6 decay: 14
Sherpa	fast for QCD muli-leg	normal: 7 decay: 7
CalcHep	very fast for 2 > 2	normal: 3/4 decay: 6
Whizard	ILC physics	normal: 6 decay: 10
pythia	low multiplicity	normal: 3 decay: 100
herwig	low multiplicity	normal: 3 decay: 100