Matrix Element Analyses

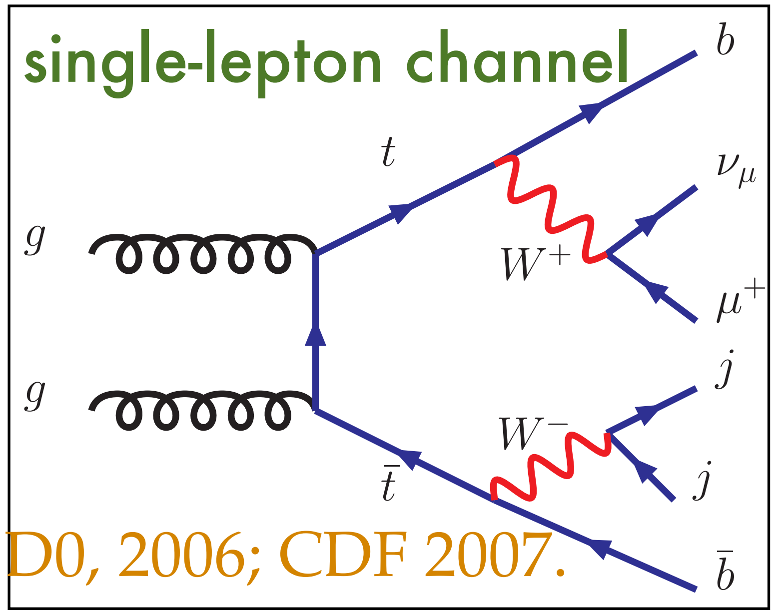
BND2013 School

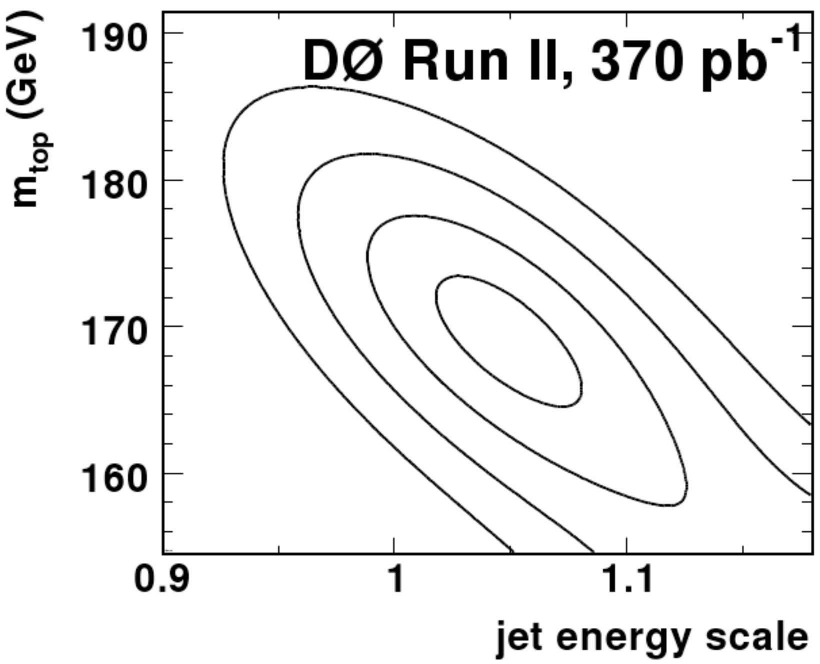
Pierre Artoisenet
NIKHEF

What is the Matrix Element Method (MEM) ?

- Event-by-event discriminator built upon matrix elements

- First introduced at LEP; in hadron-hadron collisions, first application at the Tevatron to measure the mass of the top quark





- Applied subsequently to many other analyses: all tt channels, single top, WH, H to WW at the Tevatron, H to ZZ, H to WW at the LHC

- Still subject to developments to improve its formulation / extend the range of application

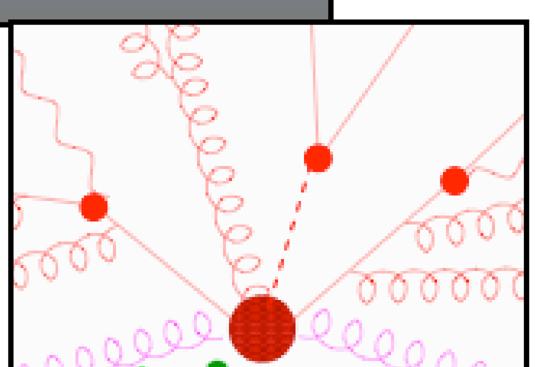
OUTLINE

- Part I: Introduction to the Matrix Element Method
  - General idea
  - Definition
  - Current developments
- Part II: Examples of application
  - m_{top} reconstruction
  - Characterization of a scalar resonance
  - Search for ttH

New physics searches at the LHC

Experimental events

Lagrangian $L(m_1, g_1, \ldots)$



PROBLEM STATEMENT:

from a sample of experimental events, how can we learn more about the structure and the parameters of the Lagrangian ?

Due to

- the complexity of the signatures,
- small S/B expected ratios,

this may be very complicated !

⇨ Need for a sophisticated procedure to discriminate between different theoretical assumptions (e.g. for m_1, g_1, ...) from a sample of experimental events

Why is the MEM so special ?

Two distinct approaches are used at hadron colliders:

I. Monte-Carlo-based approach

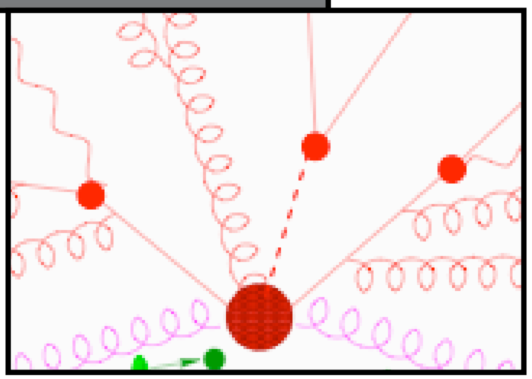
the discriminator is built upon Monte Carlo events ONLY

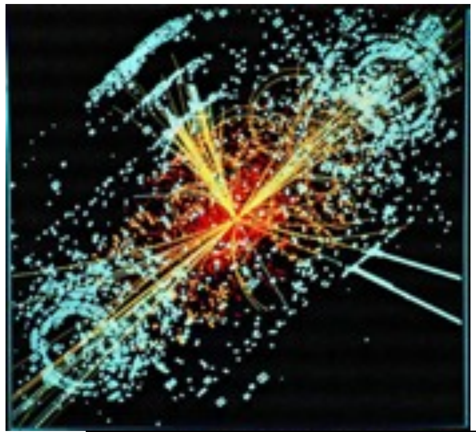
II. Matrix-Element-based approach = subject of this lecture

the discriminator is built upon hard-scattering matrix elements (and also Monte Carlo events)

Monte-Carlo-based approach

Lagrangian
$L(p_1,p_2,...)$



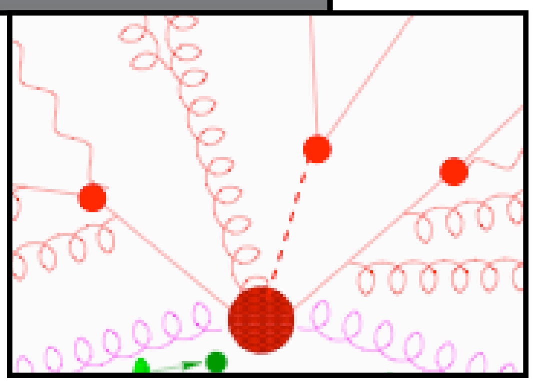


Experimental
events

In the last decade, sophisticated tools has been developed to simulate hard scattering events based on Monte Carlo techniques for any model that can be defined in the form of a Lagrangian at leading/next-to-leading order.

Monte-Carlo-based approach

Lagrangian $L(p_1,p_2,...)$ → hard sc. Matrix Elements → Parton-level Events → Showering / Hadronization



TH output:

Experimental events

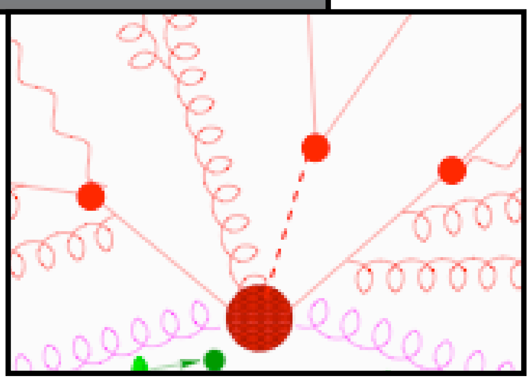
Monte-Carlo-based approach

Lagrangian $L(p_1,p_2,...)$

hard sc. Matrix Elements

Parton-level Events

Showering/ Hadronization

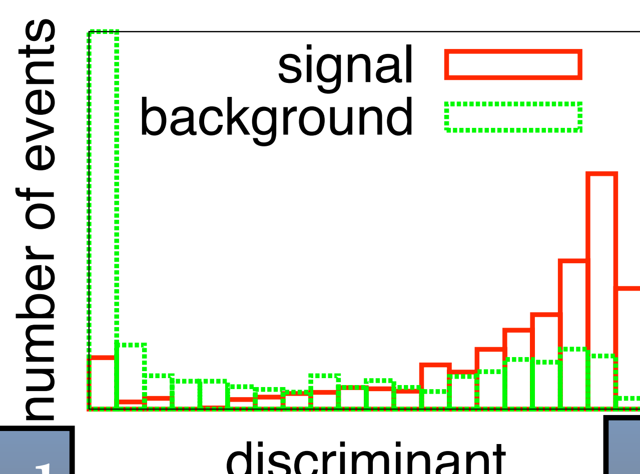


TH

hadron-level event files

EXP

Detector Resolution



detector-level selected events

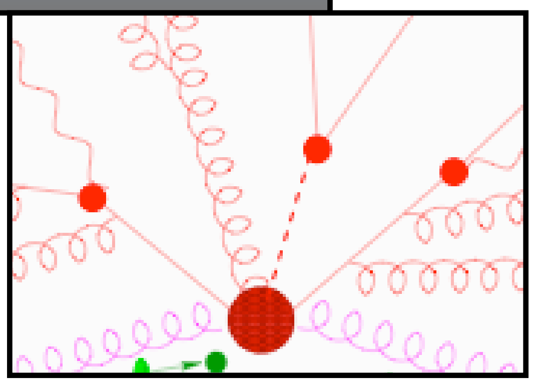
Selection procedure

Discr. variable based on MC events

Experimental events

Monte-Carlo-based approach

Lagrangian L(p_1,p_2,...) → hard sc. Matrix Elements → Parton-level Events → Showering/ Hadronization



hadron-level event files

TH

EXP

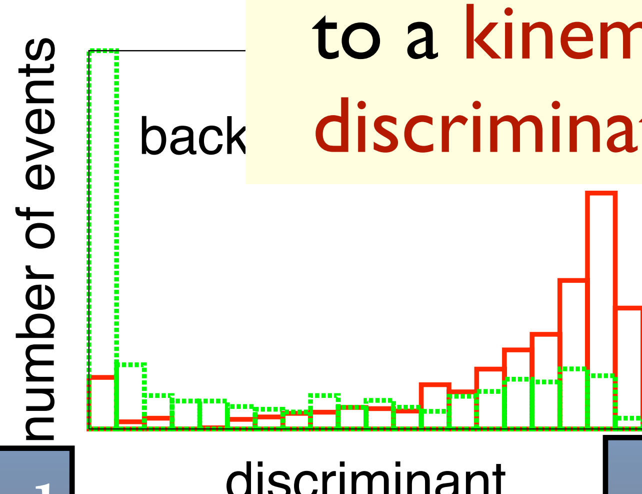
- Theory information is passed through Monte Carlo events **only**
- Samples of events serve as an input to a kinematic method to build the discriminator

Detector Resolution

detector-level selected events

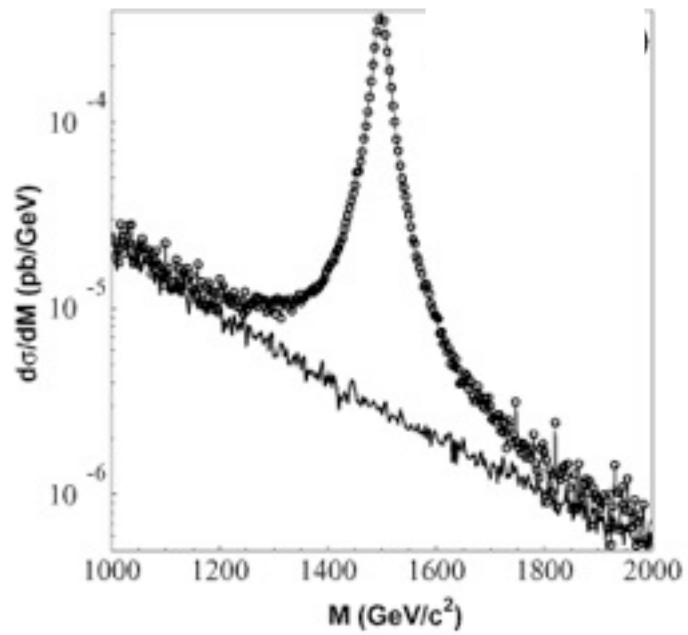
Selection procedure → Discr. variable based on MC events

Experimental events



Monte-Carlo-based approach

- Simple case: discriminator built on one reconstructed observable, e.g. the invariant mass of two leptons



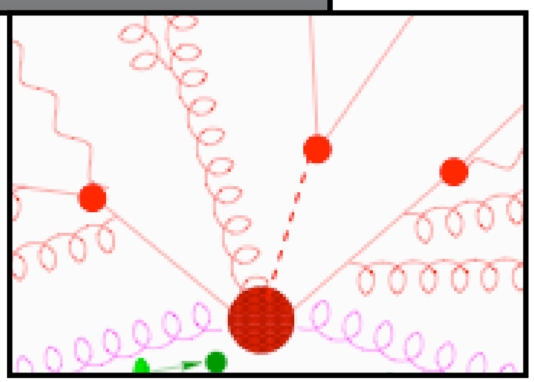
1. Reconstruct the distribution of events with respect to d=m(l^+,l^-) from MC events, under B-only and S+B hypotheses,

2. Compare with the distribution of exp. events with respect to d

- The discriminant power can be enhanced by using a sophisticated algorithm (NN, BDT) which analyses the distribution of MC events with respect to a large number of observables

Monte-Carlo-based approach

Lagrangian L(p1,p2,...) → hard sc. Matrix Elements → Parton-level Events → Showering/ Hadronization



hadron-level event files

TH

EXP

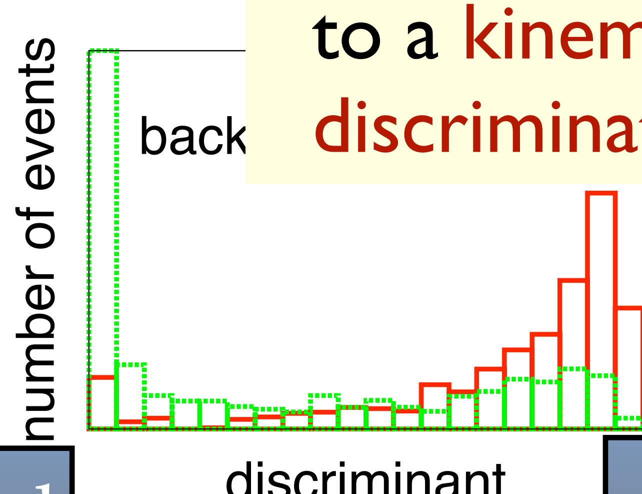
- Theory information is passed through Monte Carlo events **only**
- Samples of events serve as an input to a kinematic method to build the discriminator

Detector Resolution

Selection procedure

detector-level selected events

Discr. variable based on MC events



Experimental events

Matrix-Element-based approach

Lagrangian L(p_1,p_2,...)

hard sc. Matrix Elements

Parton-level Events

Showering / Hadronization

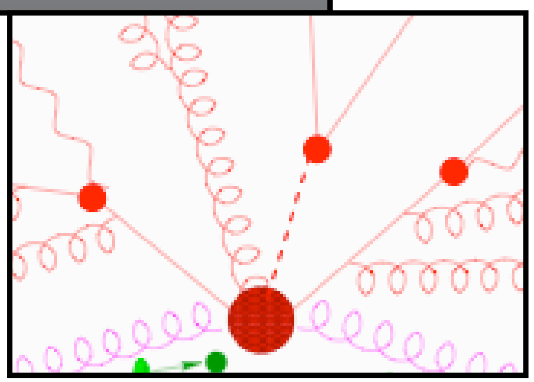
hadron-level event files

Detector Resolution

Selection procedure

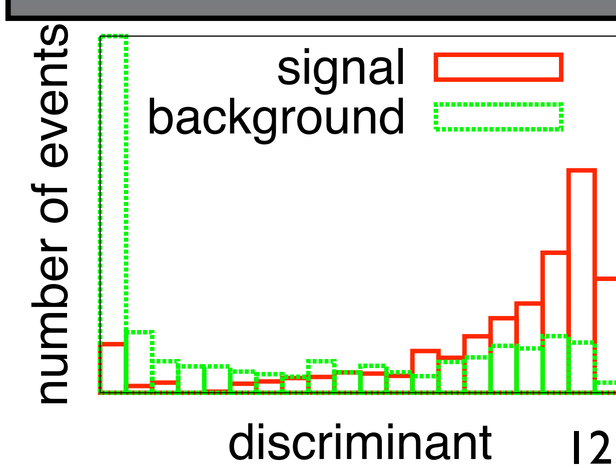
detector-level selected events

Probability Density Function (PDF)



Experimental events

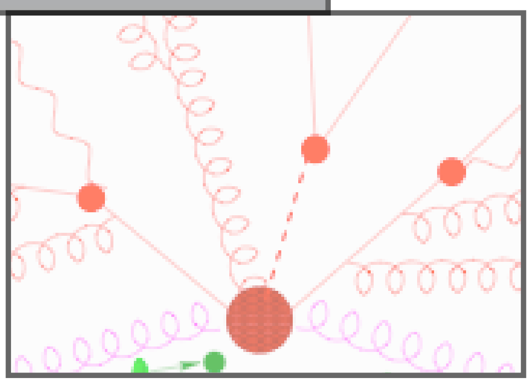
Discr. variable built upon PDF +calibration



- Theory information is passed via a (partly) analytic probability density function (+ via MC events)
- The discriminator is build upon this probability density function (e.g. using a likelihood procedure)

Matrix-Element-based approach

Lagrangian $L(p_1,p_2,\dots)$ → hard sc. Matrix Elements → Parton-level Events → Showering/ Hadronization



hadron-level event files

Detector Resolution → Selection procedure → Probability Density Function (PDF)

detector-level selected events

delicate task (accuracy ?)

Probability Density Function (PDF) → Discr. variable built upon PDF +calibration → Experimental events

"Matrix Element Method":

- Model to approximate the Probability Density Function in the case of hadron-hadron collisions

Matrix Element Method

- Construction of the PDF based on hard scattering matrix elements
- Definition of the discriminating variable: likelihood built upon this PDF

MEM

Define a Probability Density Function using matrix elements

$$P(x|\alpha)$$

likelihood analysis

Evaluate the probability at each event under the hypotheses $\alpha = h_1, h_2, \ldots$

$$w_i(\alpha) = P(x_i|\alpha)$$

matrix element weight

Combine the weights into a likelihood

$$L(\alpha) \approx \prod_i P(x_i|\alpha)$$

x : kinematics of the reconstructed event

α : theoretical assumption

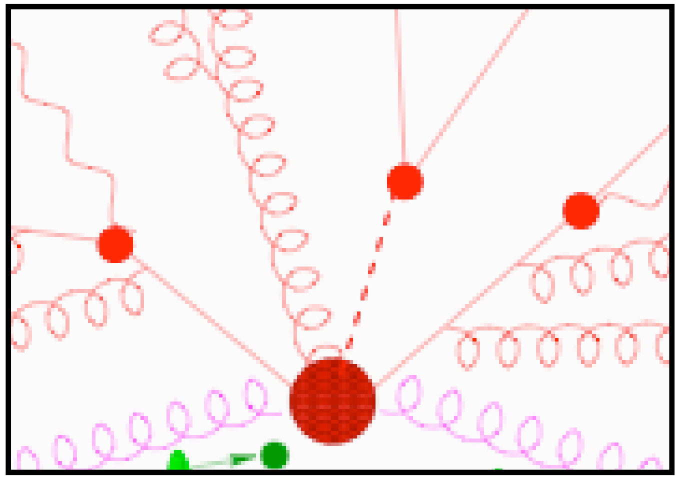
Reweighing events with matrix elements

- Imagine we live in an ideal world, with an ideal detector able to reconstruct

✓ all the final state objects

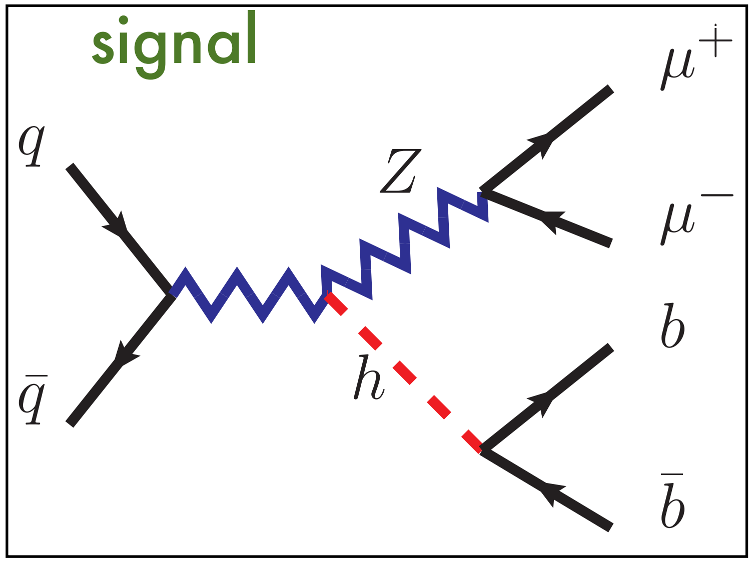
✓ at the scale Q= scale of the hard interaction

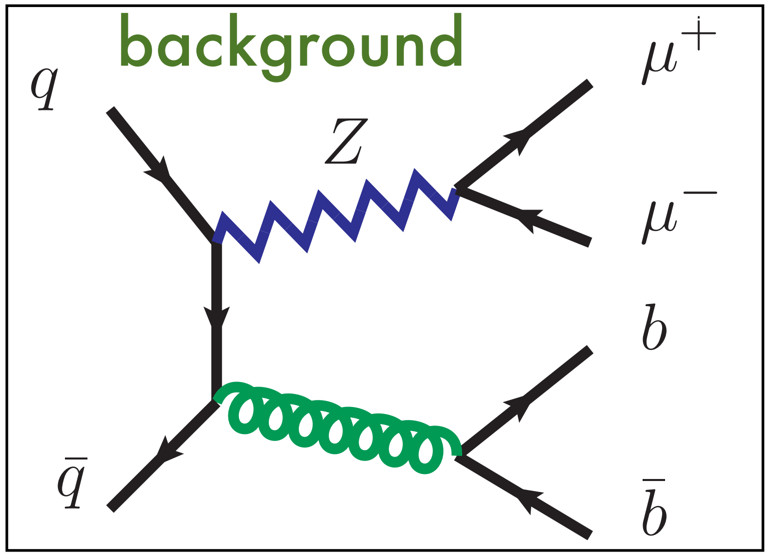
✓ with an infinite resolution



Reweighing events with matrix elements

- Assuming this ideal world/detector, consider the following search:





In this analysis, an event x corresponds to $p_{\mu^+}, p_{\mu^-}, p_b, p_{\bar{b}}$

Define a probability density function using matrix elements

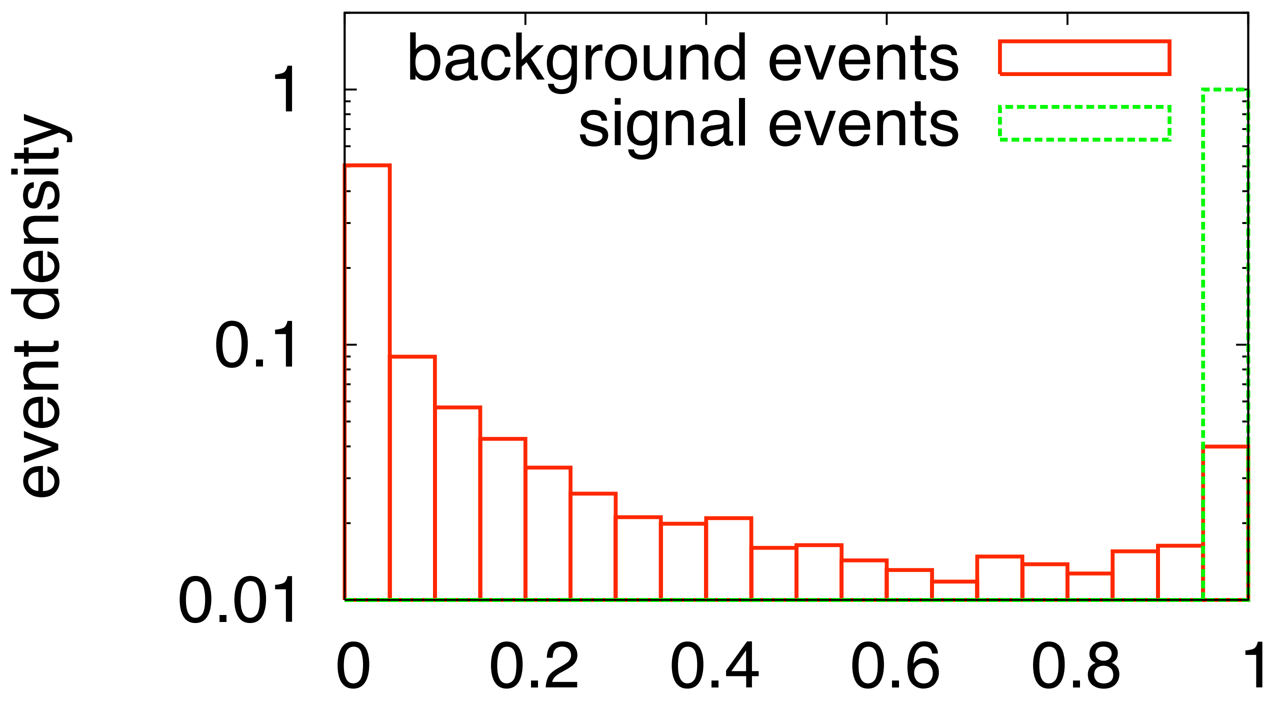
$$P(x|S) \propto |M_S(x)|^2 \qquad P(x|B) \propto |M_B(x)|^2$$

M_S : matrix element under the signal hypothesis

M_B : matrix element under the background hypothesis

Reweighing events with matrix elements

Evaluate the probability for each event under the hypotheses α=S or B



$$D(x) = \left[1 + \frac{P(x|B)}{P(x|S)}\right]^{-1}$$

D is a discriminator based on the phase-space distribution of the events

Defining the likelihood

Combine the weights into one likelihood

Given N experimental events, you can test the S+B hypothesis versus the B-only hypothesis

If s,b =expected numbers of signal and background events is known, you can also use this information to improve the discriminating power

Likelihood for the B-only hypothesis: $$\mathrm{Pois}(N|b)\prod_{i=1}^{N} P(x_i|B)$$

Likelihood for S+B hypothesis:

$$\mathrm{Pois}(N|s+b)\prod_{i=1}^{N}[sP(x_i|S)+bP(x_i|B)]/(s+b)$$

K. Cranmer, T. Plehn, Eur. Phys. J. C 51, 415-420

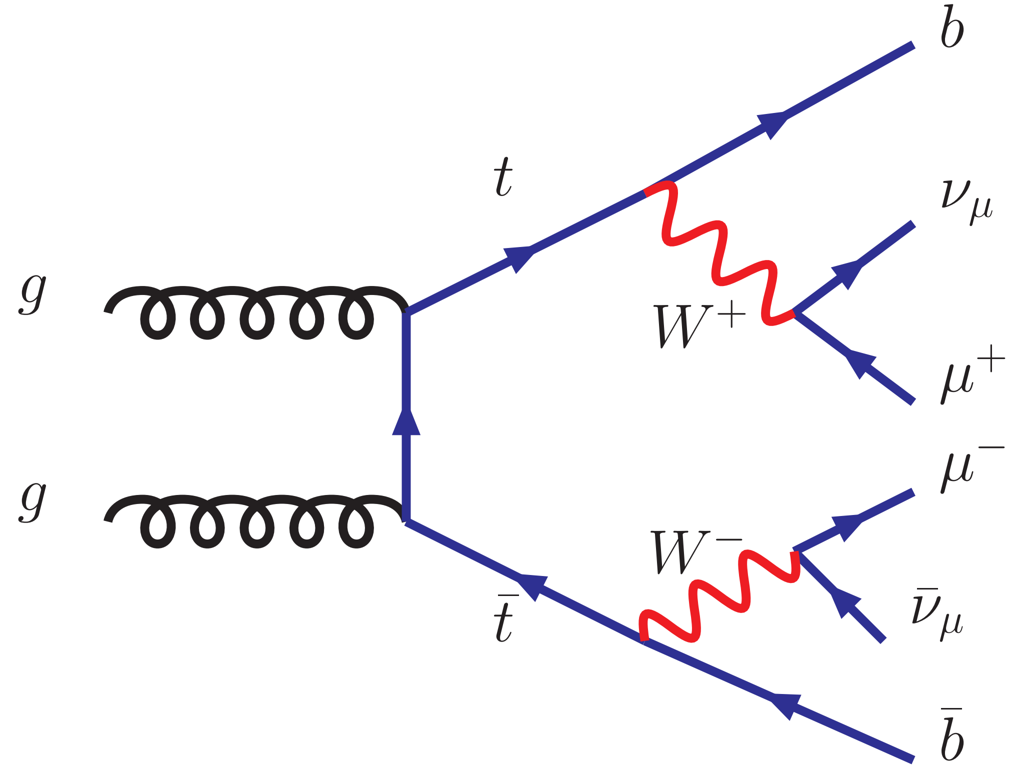
Real experiment

In a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

1. Missing energy

Some particles escape from the detector without any interaction (neutrino, wimp, ...)

example: top-quark pair production, di-leptonic channel



Real experiment

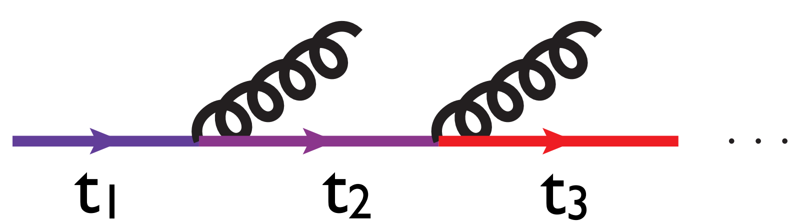
In a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

2. Showering/hadronization effects

A high energy collision is a multi-scale process, but a fixed-order matrix element provides a relevant description only for the hard scale Q

$Q \longrightarrow \sim 1$ GeV

physics	hard scattering	showering	hadronization
description tool	matrix element at fixed order in α_s	Sudakov form factors	simulation model tuned to the data



$$\Delta(t_1, t_2) = \exp\left\{-\int_{t_2}^{t_1} \frac{dt'}{t'} \int dz \frac{\alpha_s(t')}{2\pi} P(z)\right\}$$

non-branching probability between scales t_1 and t_2

Real experiment

In a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

3. Experimental resolution/reconstruction algorithm

the final state objects (hadrons, leptons) are reconstructed with a finite resolution

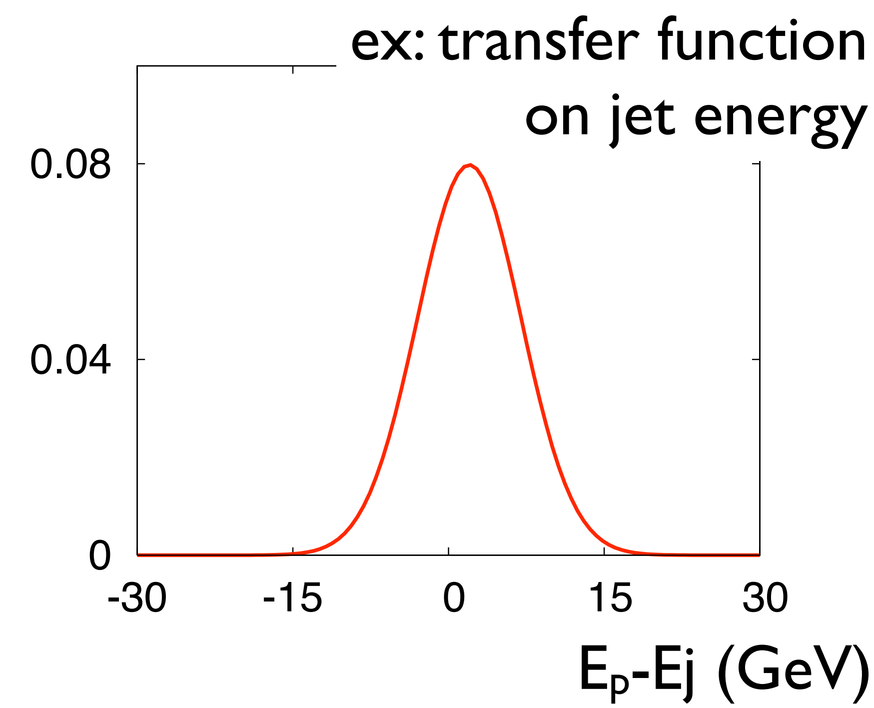
MEM prescription for the PDF

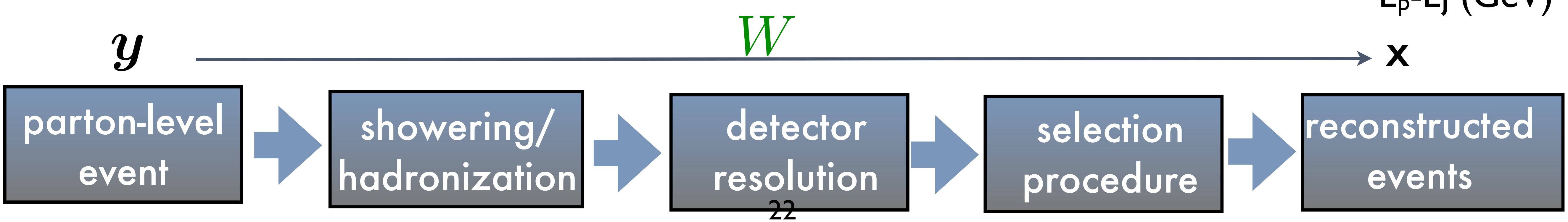
In a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

1. Missing energy ⟷ $P(x,\alpha)$ must be summed over the unobserved degrees of freedom

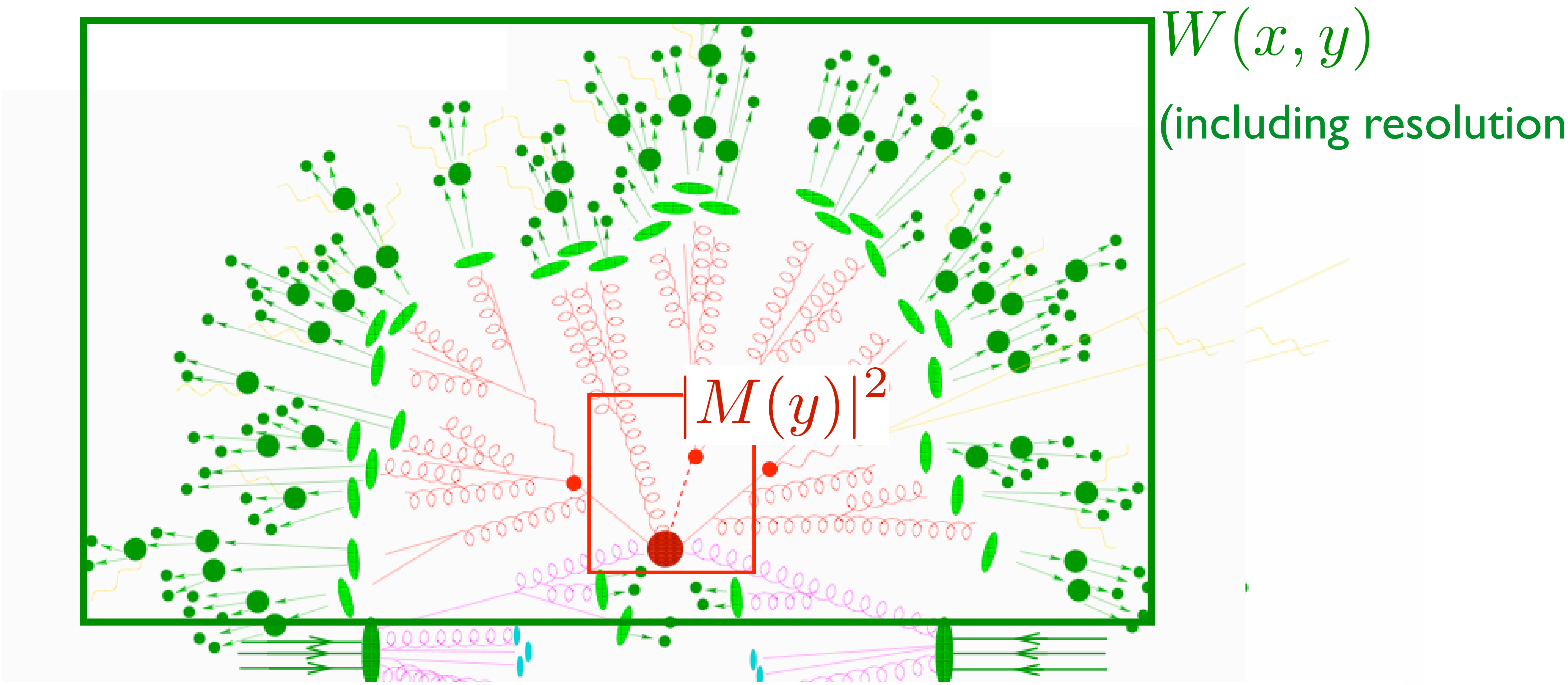
2. Showering/hadronization effets
3. Experimental resolution/reconstruction algorithm

⟷ convolute with a transfer function $W(x,y)$
= probability that x is reconstructed given that y has been produced





“Assumed” factorization in MEM:



The prescription to extract the transfer function relies on a one-to-one assignment between reconstructed jets and partons

- This prescription is ambiguous beyond LO
- Current definition of the pdf in the MEM has LO accuracy only

Definition of the PDF in the MEM

- Real detector: need marginalize over unconstrained information and to convolute with the resolution function W for the measured quantities

$$P(\boldsymbol{x}_i, \alpha) = \frac{1}{\sigma^{obs}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y})$$

integration on the parton-level phase-space

tree-level matrix element

transfer function extracted from MC simulation

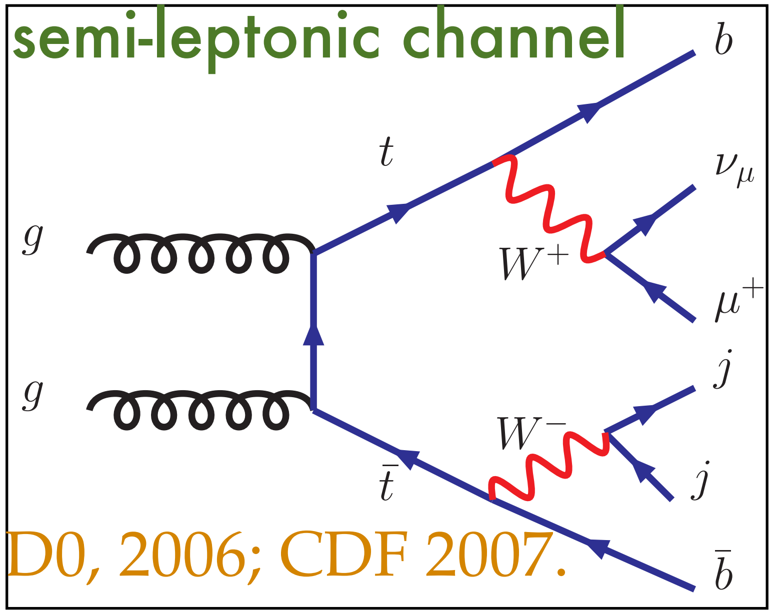
normalization: $\int dx W(x,y) = 1$

⇨ the probability density P(x| α) is normalized to 1

First MEM analyses at the Tevatron

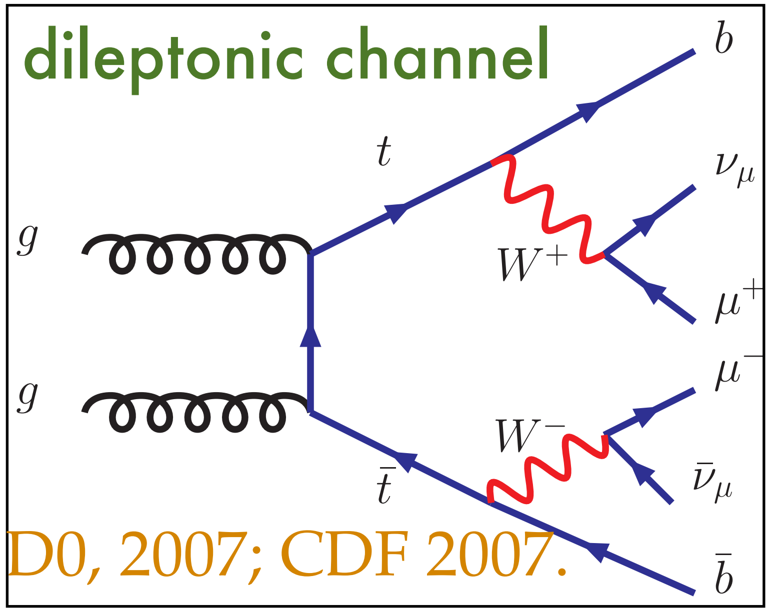
Top-quark mass measurement from $t\bar{t}$ production in hadron collisions

semi-leptonic channel

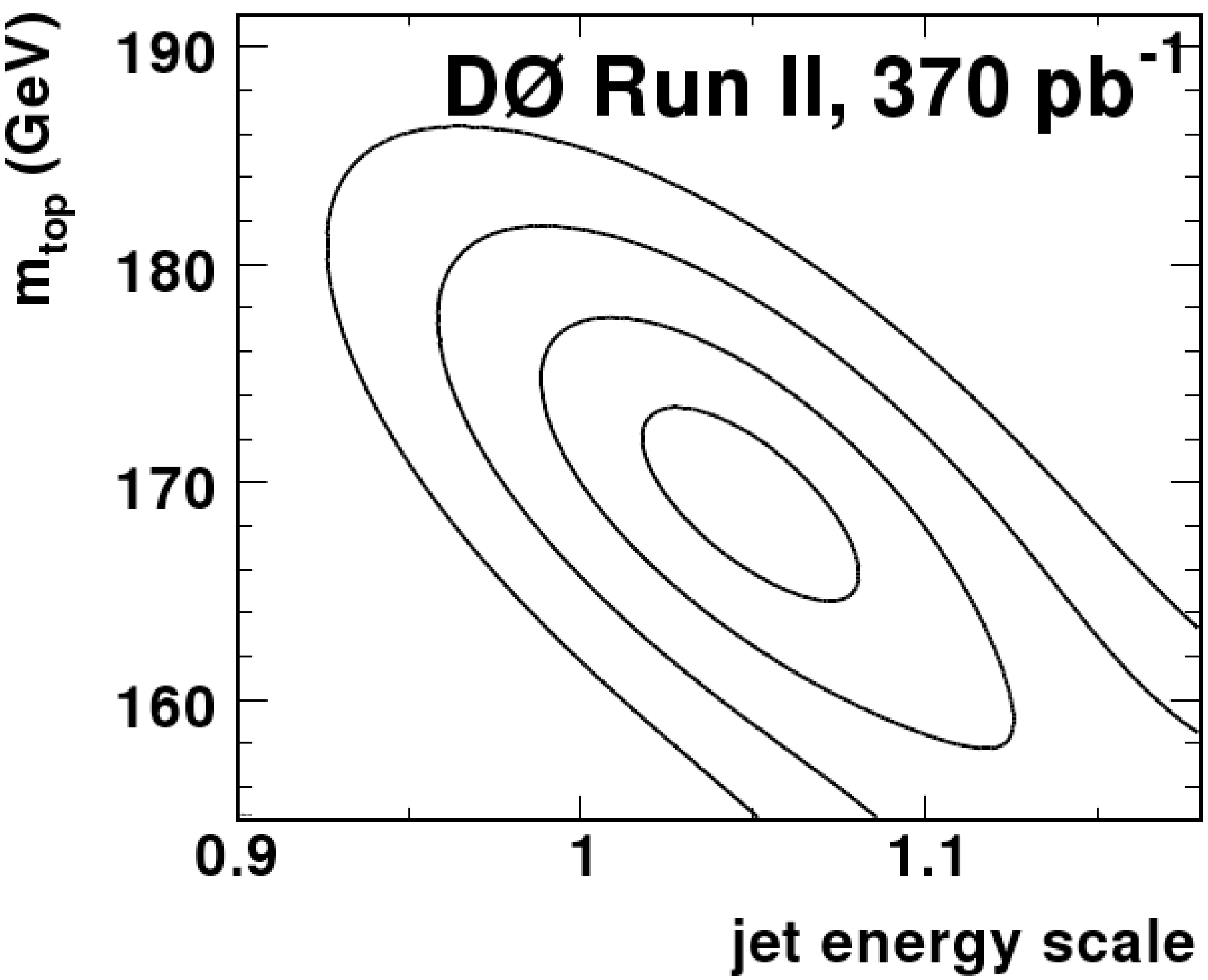


D0, 2006; CDF 2007.

dileptonic channel



D0, 2007; CDF 2007.



[DO Phys. Rev. D75 092005, 2006]

Significant improvement for the measurement of the top-quark mass

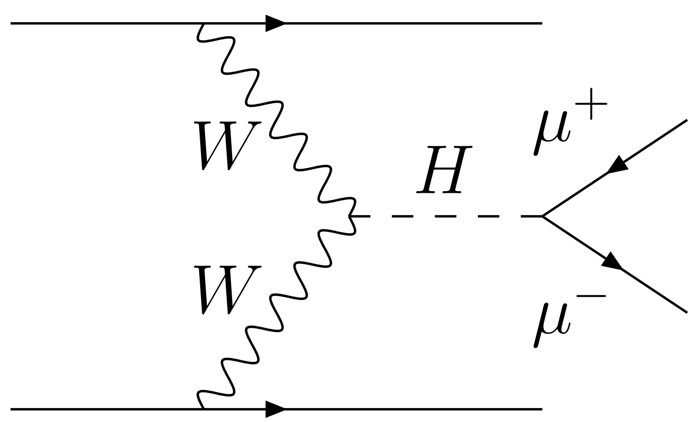
Recent developments in the MEM

Research activities in several directions:

1. Maximum statistical significance / treatment of the systematics
2. MEM beyond Leading order
3. Practical calculation of the MEM weights
4. Exploring new applications of the MEM

I. Maximum statistical significance / systematics

- Formal expected maximum significance based on the Neyman-Pearson Lemma T. Plehn, K. Cranmer, 2006



For a simple hypothesis test (H_0, H_1), the variable that maximizes power is the likelihood ratio

$$Q(\mathbf{x}) = \frac{L(\mathbf{x}|H_1)}{L(\mathbf{x}|H_0)}$$

Fourier transformation to obtain the expected Likelihood profile of a sample of events (including Poisson fluctuations) from the single-event Likelihood profile ⇒ no need to generate pseudo-experiments

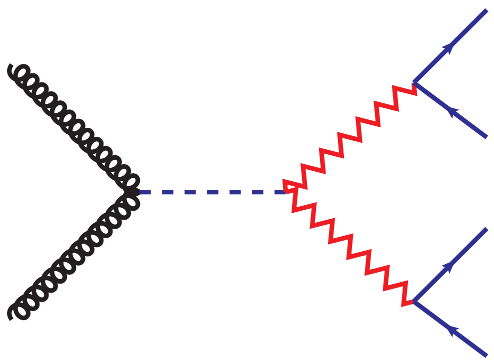
- Generalization: profile likelihood with systematic uncertainties included as nuisance parameters in the likelihood G. Cowan, K. Cranmer, E. Gross, O. Vitells, 2010

II. MEM at NLO

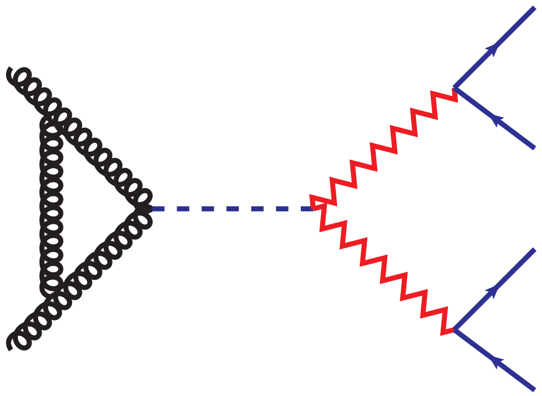
J. M. Campbell, W. T. Giele, C. Williams, 2012

- Define an NLO weight for EW production processes in the unresolved region (veto on events with extra jets)

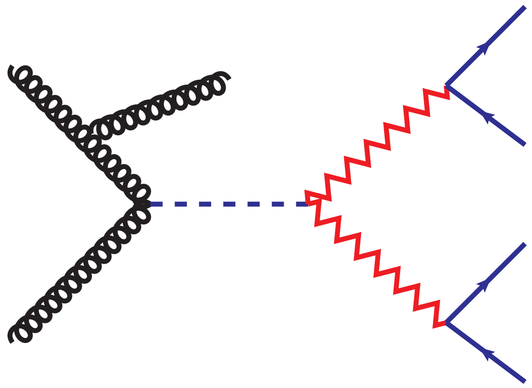
- Example of application: Higgs boson production in the 4-lepton channel



Born



1-loop

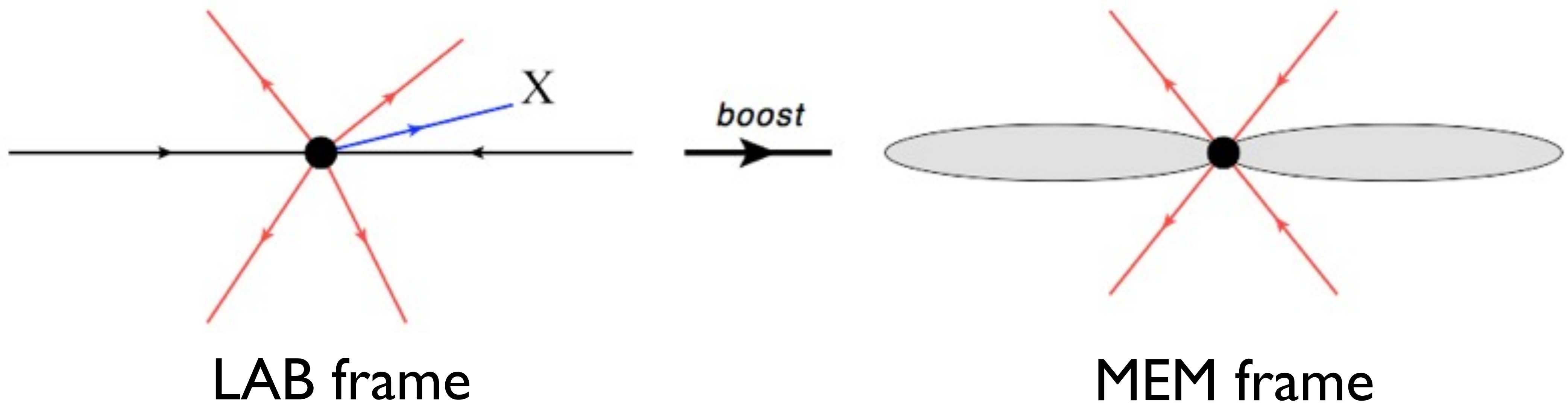


real-emission

II. MEM at NLO

J. M. Campbell, W.T. Giele, C. Williams, 2012

- NLO correction may affect the kinematics of the Higgs boson in the lab frame (nonzero pT)



Step 1: Boost the exp. events in a frame where pT(h) = 0 (MEM frame)

⇒ one can readily associate a LO weight to the event

$$\mathcal{P}_{LO}(\Phi_B) = \frac{f(x_1)f(x_2)}{2x_1x_2s}|\mathcal{M}^{(0)}(\Phi_B)|^2$$

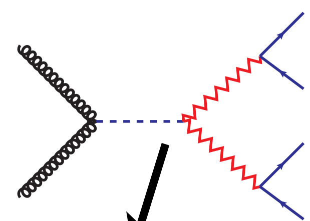
II. MEM at NLO

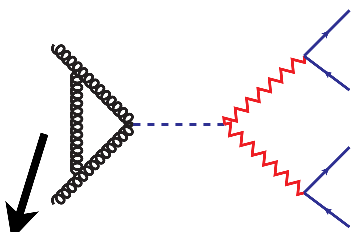
J. M. Campbell, W. T. Giele, C. Williams, 2012

Step2: Define an IR-safe weight at NLO accuracy in the unresolved region [pT(radiation)<s_{min}]

Crucial idea: use the Forward-Branching Phase-Space factorization to integrate the real-emission amplitude over the unresolved degrees of freedom

$$d\Phi(p_a + p_b \to Q + p_r) = d\,\Phi(\widehat{p}_a + \widehat{p}_b \to Q) \times d\,\Phi_{\mathrm{FBPS}}(p_a, p_b, p_r) \times \theta_{\mathrm{veto}}$$



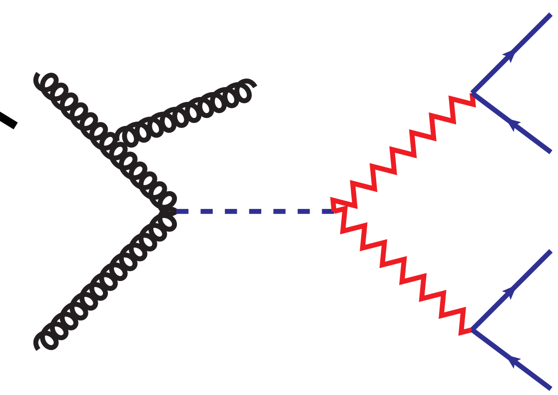


divergent

divergent

$$\mathcal{P}_{NLO} = \frac{f(x_1)f(x_2)}{2x_1x_2s}\left(\left(1+\boxed{\mathcal{R}_v(s_{min})}\right)|\mathcal{M}^{(0)}(\Phi_B)|^2 + \boxed{2\mathrm{Re}\left\{\mathcal{M}^{(0)}\mathcal{M}^{(1)^\dagger}(\Phi_\mathrm{B})\right\}}\right)$$
$$\boxed{+\int_{s_{min}} d\,\Phi^{IS}_{\mathrm{FBPS}}(\Phi_B)J_x\frac{f(x_a)f(x_b)}{2x_ax_bs}|M_R^{(0)}(\Phi_R(\Phi_B))|^2 + \mathcal{O}(s_{min})}$$

divergent (IR)

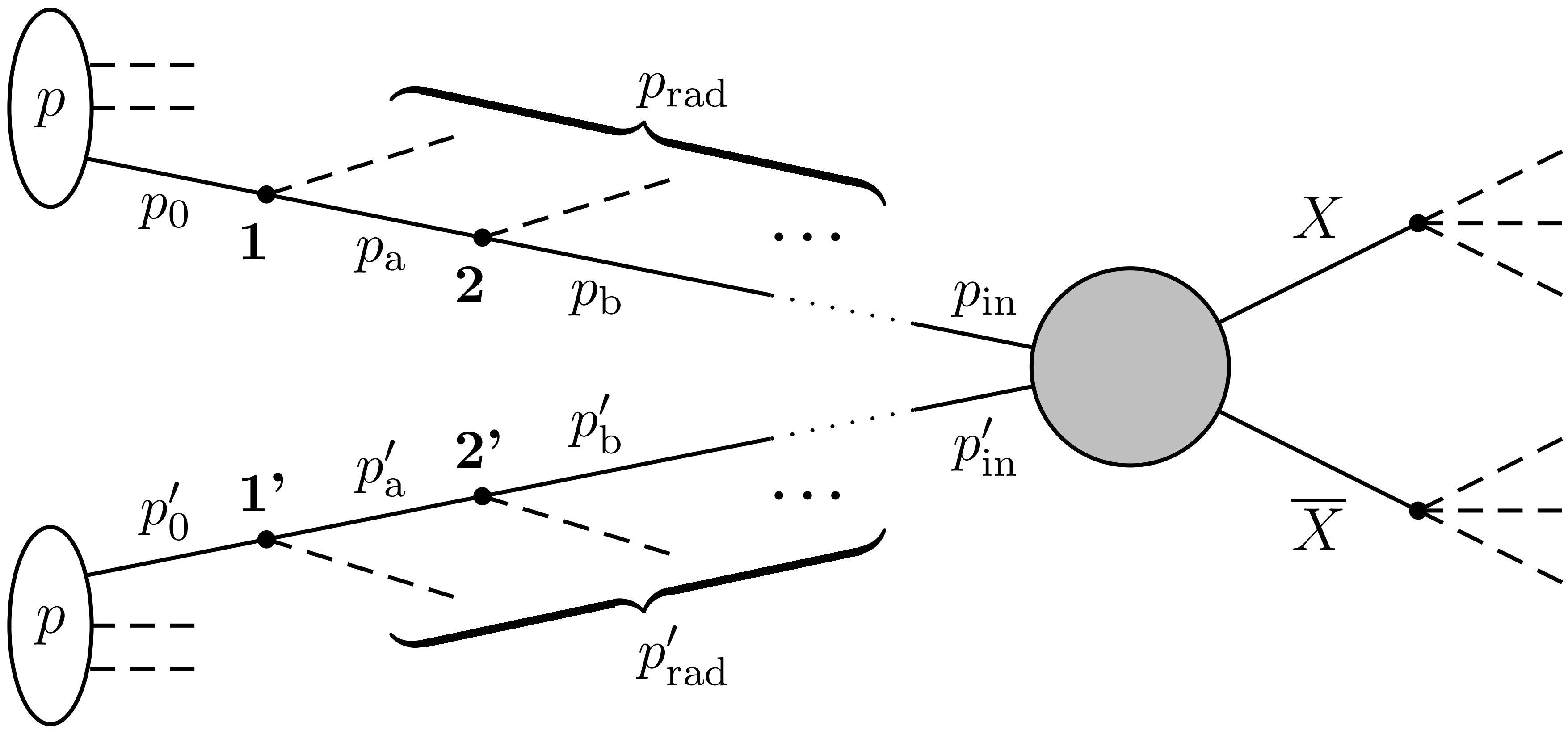


Remarks:

1. fixed-order MEM weight (no resummation), in particular the scale S_{min} cannot be chosen to small
2. marginalization of the information of $p_T(H)$

III. Effective treatment of extra QCD radiations

J. Alwall, A. Freitas, O. Mattelaer, 2011



Step 1: boost correction (correct for the fact that ISR affects Born-level kinematics)

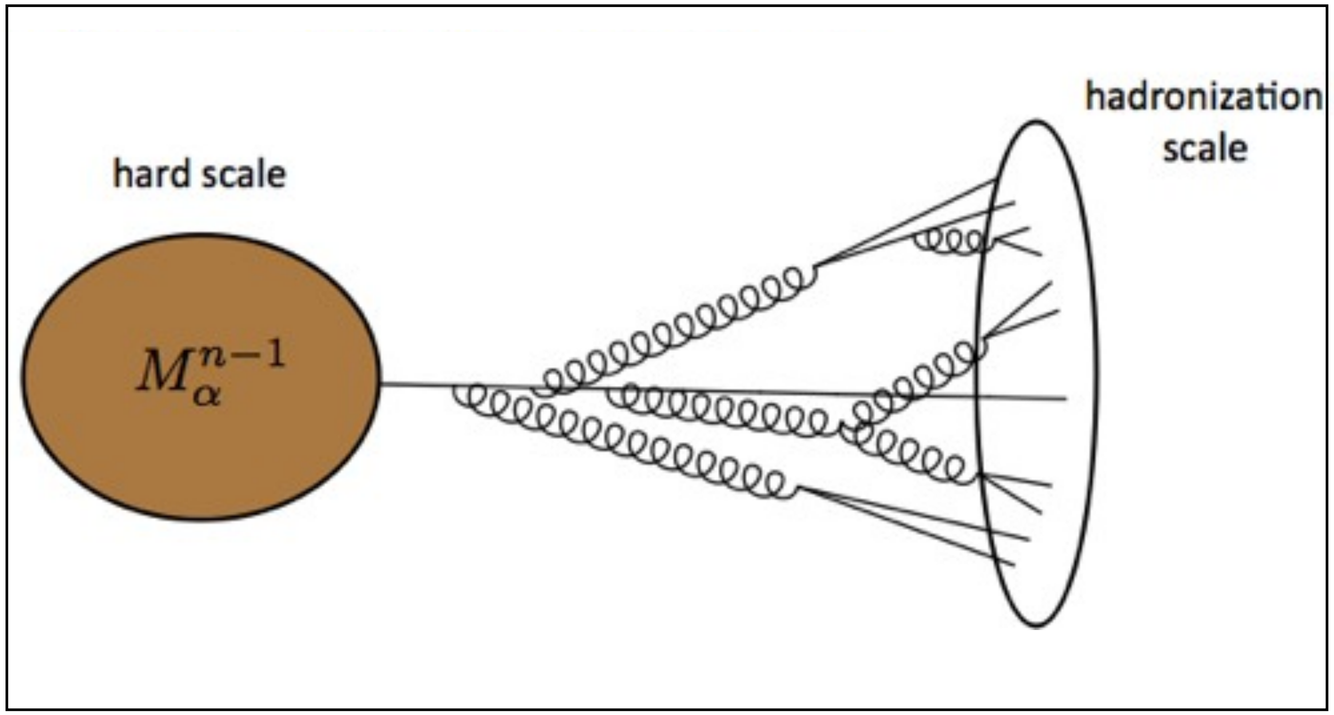
Step 2: sudakov reweighting for the ISR

Remarks: 1. information on p_T (H) is also folded in the definition of the weight (resummation of the logs in the Sudakov weight)

2. no information from the one-loop amplitude

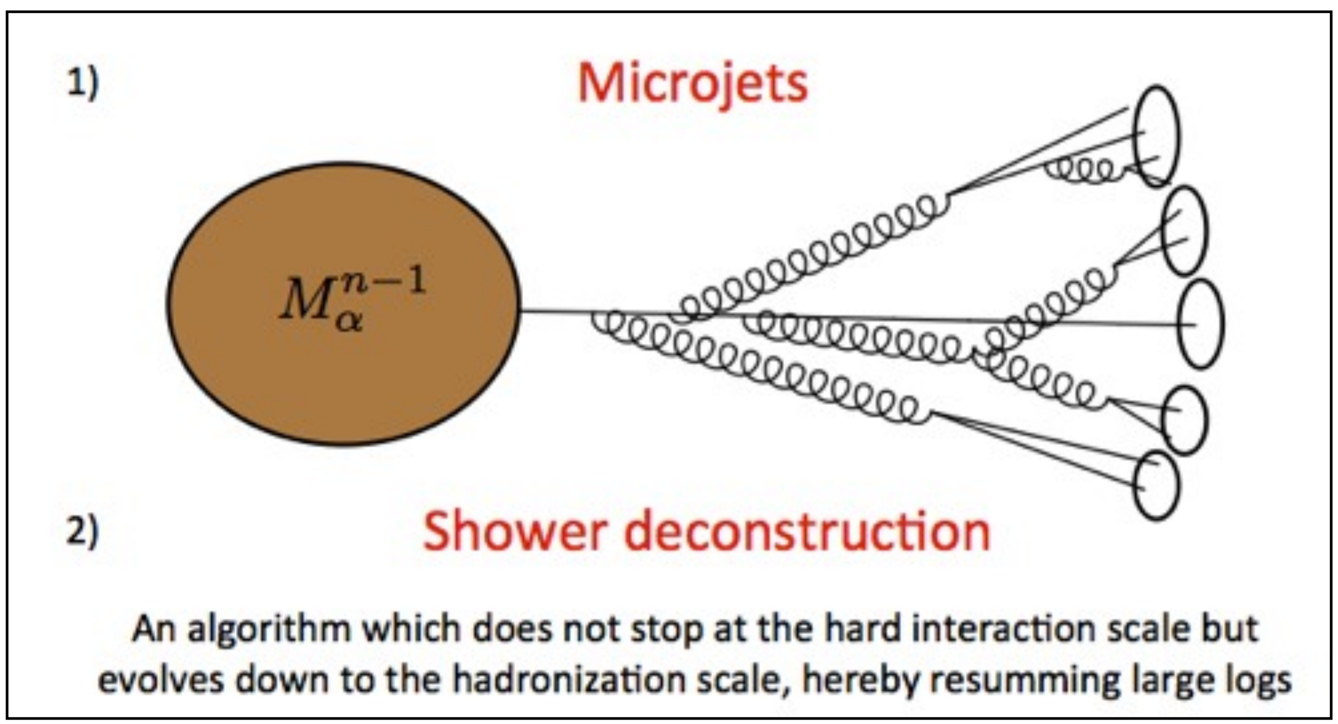
IV. Shower deconstruction

D. E. Soper, M. Spannowsky 2012



Standard MEM:

transfer function W(y,x)
= map between the parton-level kinematics and the reconstructed jets



Shower deconstruction:

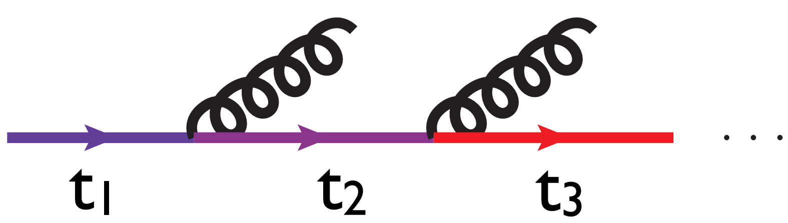
reconstruct the microjet config.
(kt, R=0.15, p_T > 5 GeV,)
and consider the probability density function of these objects

IV. Shower deconstruction

How do you get the weight ? (large logs prevent the use of fixed-order matrix element)

step 1: reconstruct all possible branching histories leading to the observed micro-jet configuration

step 2: relative weight for each branching history is a product of Sudakov form factors



$$\Delta(t_1, t_2) = \exp\left\{ - \int_{t_2}^{t_1} \frac{dt'}{t'} \int dz \frac{\alpha_s(t')}{2\pi} P(z) \right\}$$

Practical Evaluation of the PDF

Let's go back to the definition of the weights at leading order:

$$P(\boldsymbol{x}_i, \alpha) = \frac{1}{\sigma^{obs}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y}) Acc(x)$$

integration on the parton-level phase-space

tree-level matrix element

transfer function

Practical Evaluation of the PDF

Let's go back to the definition of the weights at leading order:

$$P(\boldsymbol{x}_i, \alpha) = \frac{1}{\sigma^{obs}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y}) Acc(x)$$

integration on the parton-level phase-space

tree-level matrix element

transfer function

✓ available with the use of a matrix element generator (madgraph, sherpa, ..)

Practical Evaluation of the PDF

Let's go back to the definition of the weights at leading order:

$$P(\boldsymbol{x}_i, \alpha) = \frac{1}{\sigma^{obs}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y}) Acc(x)$$

integration on the parton-level phase-space

tree-level matrix element

transfer function

✓ available with the use of a matrix element generator (madgraph, sherpa, ..)

✓ can be extracted from Monte Carlo similations

Practical Evaluation of the PDF

Let's go back to the definition of the weights at leading order:

$$P(\boldsymbol{x}_i, \alpha) = \frac{1}{\sigma^{obs}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y}) Acc(x)$$

integration on the parton-level phase-space

tree-level matrix element

transfer function

- Monte Carlo integration ?

✓ available with the use of a matrix element generator (madgraph, sherpa, ..)

✓ can be extracted from Monte Carlo similations

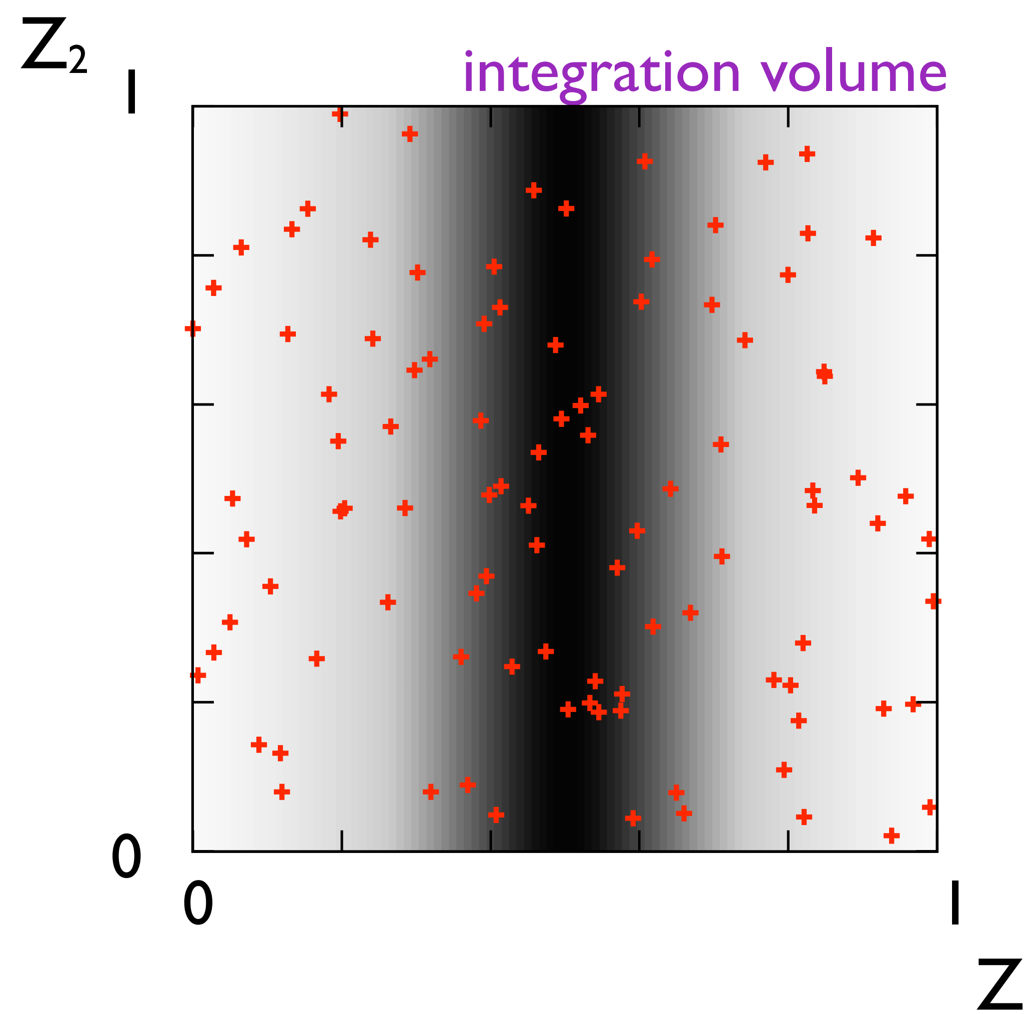
Monte Carlo integration

Basic idea: $I = \int_V dz\, f(\boldsymbol{z})$ is estimated by sampling the volume V=[0,1]^d with N uniformly distributed random points: $E = \frac{1}{N}\sum_{n=1}^{N} f(\boldsymbol{z}_n)$

Std deviation: $\sigma_I \approx \frac{\boxed{S}}{\sqrt{N}}$

$$S^2 = \mathrm{var}(f) = \frac{1}{N-1}\sum_{n=1}^{N}\left[f(\boldsymbol{z}_n) - E\right]^2$$

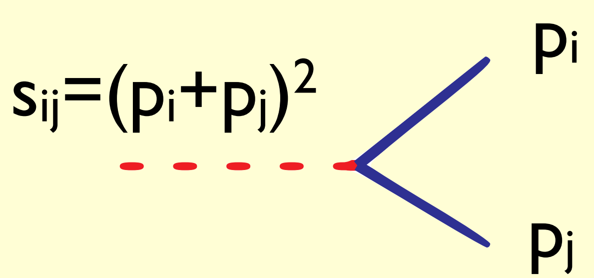
if S large ➡ poor convergence



Monte Carlo integration

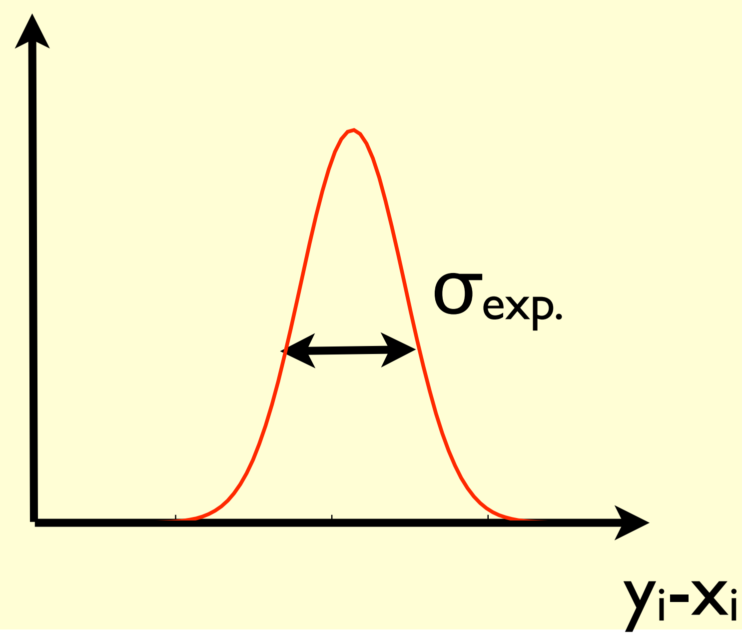
$$P(x,\alpha) \propto \int d\phi_{\boldsymbol{y}} \boxed{|M|^2(\boldsymbol{y})} \boxed{W(\boldsymbol{x},\boldsymbol{y})}$$

Highly non-uniform, especially in the presence of resonances



Breit-Wigner distr. in s_{ij}

Highly non-uniform, especially when the resolution associated with a reconstructed quantity x_i is high:



When the dimension of the phase-space is large, this structure in "peaks" complicates the numerical evaluation of the weights

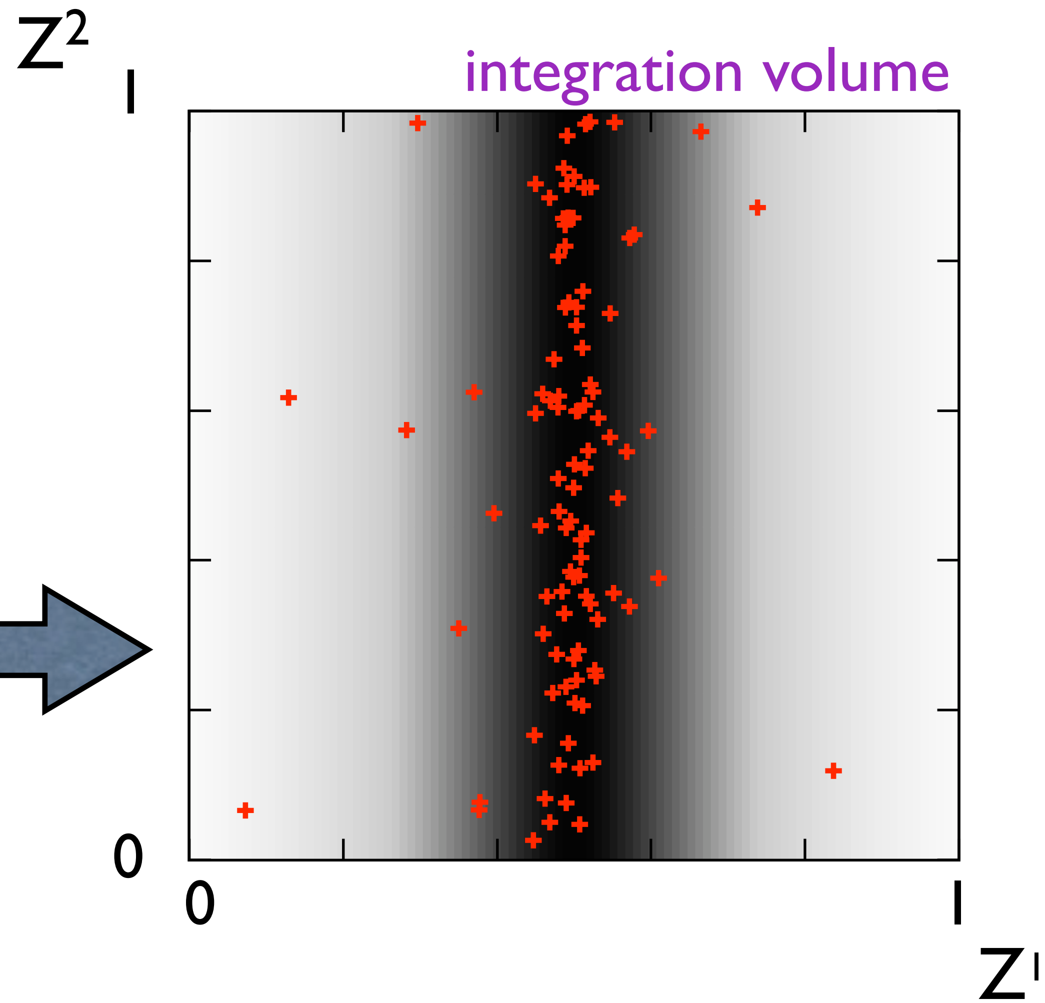
Need for a procedure to speed up the convergence (large number of weights must be evaluated)

Monte Carlo integration

Adaptive MC integration: probe the phase-space volume according to a probability density function $p(\boldsymbol{z}) = p_1(z^1)p_2(z^2)\dots p_d(z^d)$ (grid) that is adapted iteration after iteration

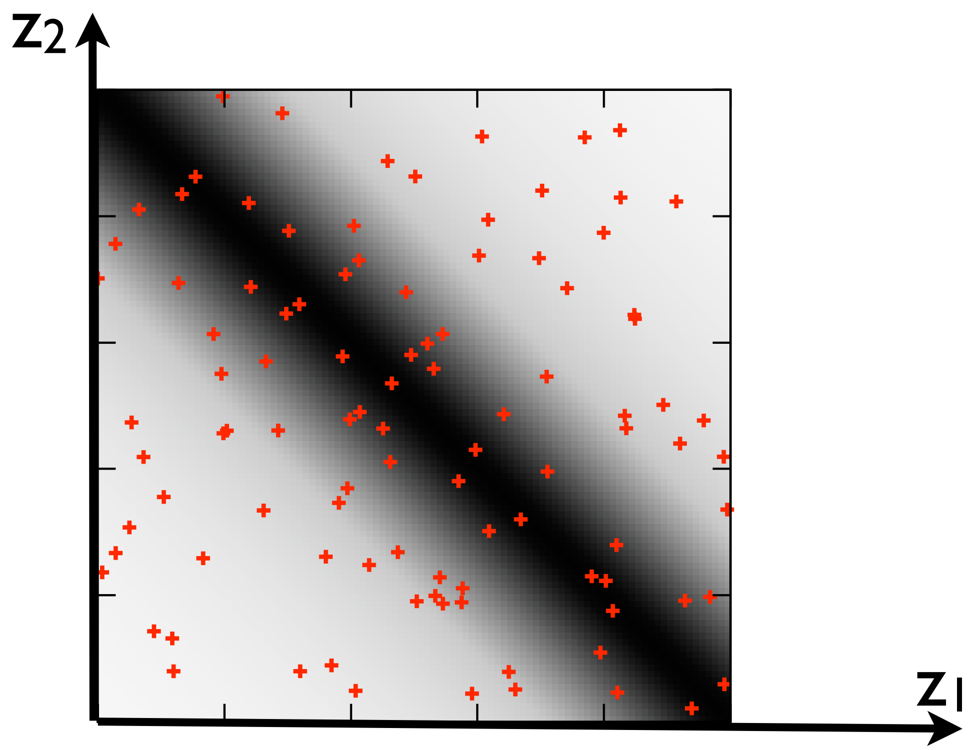
The grid has a factorized dependence in the integration variables

Here: adapt the expected density of points along the direction Z^1 to resolve the "peak"



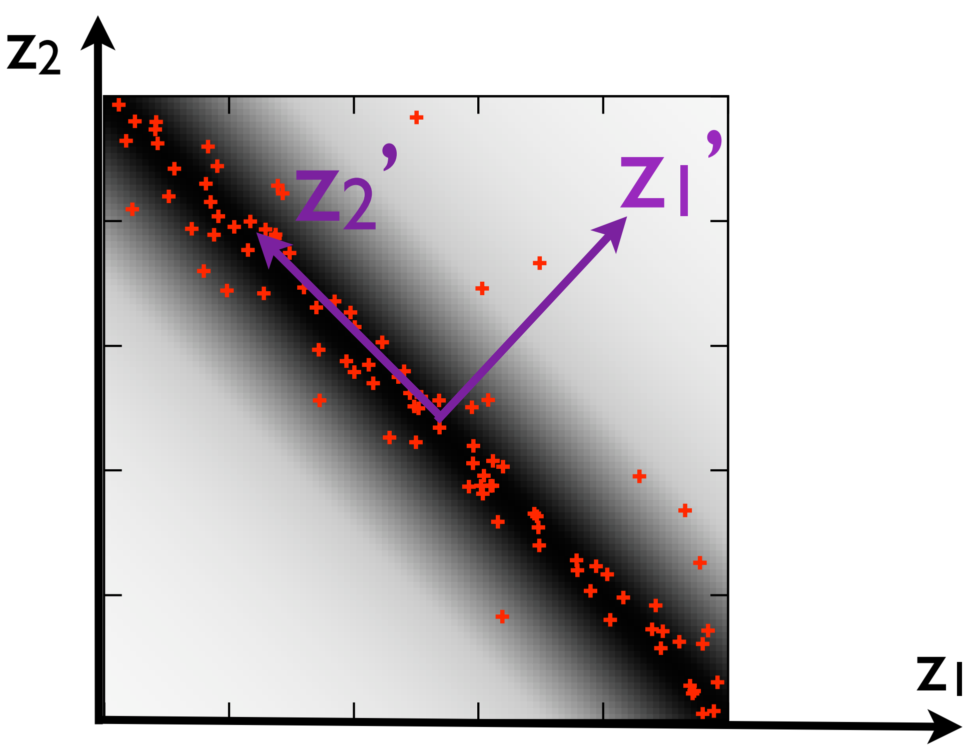
Adaptive Monte Carlo integration

The efficiency of the adaptive MC integration depends on the choice of variables of integration



Variables z_1, z_2:

The grid cannot be adjusted efficiently to the shape of the integrand because the strength of the "peak" in the integrand is not controlled by a single variable of integration



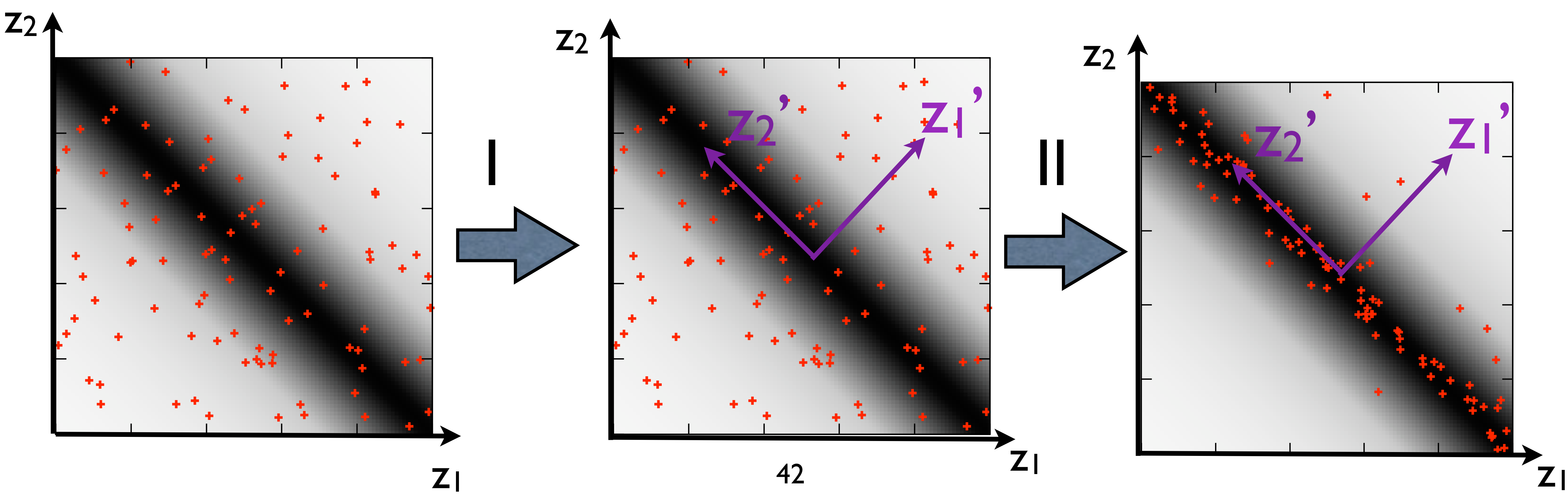
Variables z_1', z_2':

The probability density along z_1' (= variable that controls the strength of the "peak") can be adapted to probe the integration region where the integrand is the largest

III. Practical Calculation of the weights

I. Phase space mapping: parametrize the phase-space in such a way that the strength of each peak is mapped onto a single variable of integration

II. Adaptive MC integration: probe the phase-space volume according to a probability density function $p(\boldsymbol{z}) = p_1(z^1)p_2(z^2)\ldots p_d(z^d)$ (grid) that is adapted iteration after iteration



MadWeight

Full automation of the calculation of the MEM weights (LO only) in the madgraph framework P.A., F. Maltoni, V. Lemaitre, O. Mattelaer, 2011

MadWeight = generator of optimized phase-space mappings $d\phi_y$ for the evaluation of the weights in the Matrix Element Method

- Multichannel integrator
- Narrow Width Approximation optional
- Effective treatment for ISR radiation
- Grouping of subprocesses
- Pre-training of the grid
- Monte Carlo over parton-jet assignements

The code is available on the launchpad:

```
bzr branch lp:~maddevelopers/madgraph5/madweight_mc_perm
```

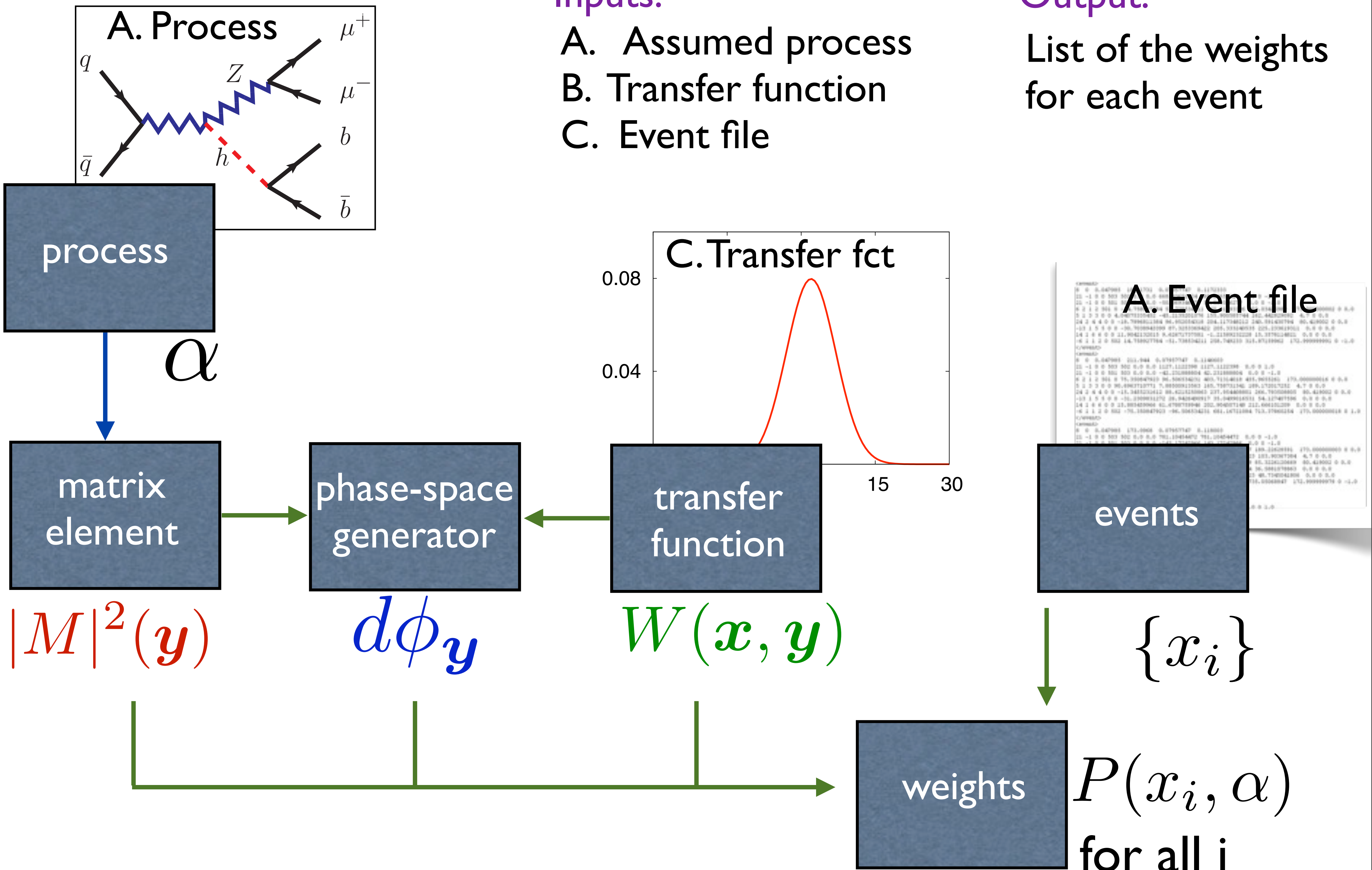
MadWeight: how does it works ?

Inputs:

A. Assumed process
B. Transfer function
C. Event file

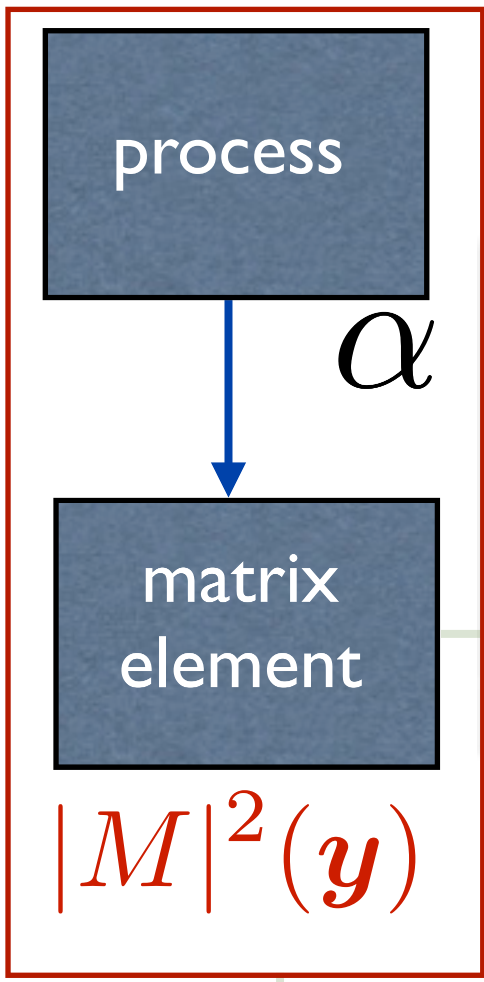
Output:

List of the weights for each event



MadWeight: how does it works ?

Step 1: generate the matrix elements
type ./bin/mg5 to open the prompt



input : proc_card_mg5.dat

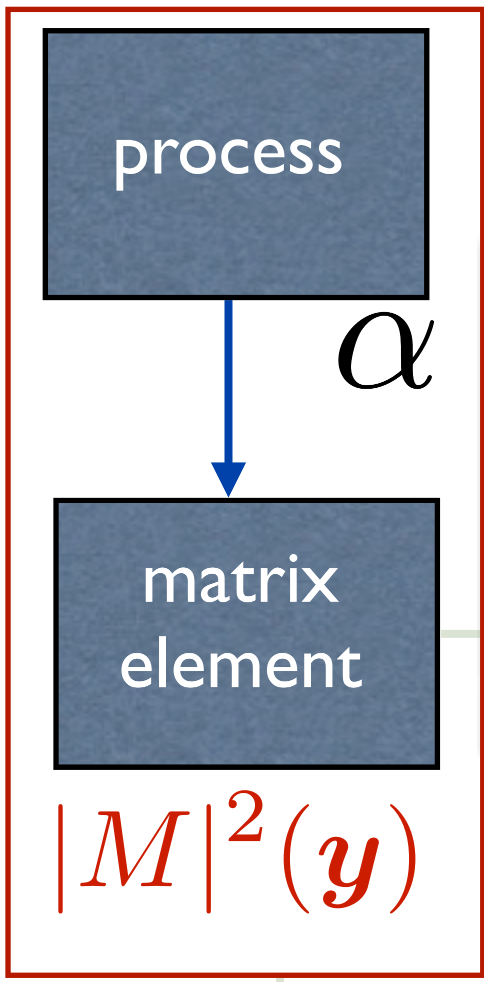
```
import model sm
generate p p > Z H , Z > mu+ mu- , H > b b~
output madweight HZ_hypo
```

output :

code for the evaluation of the weights in directory HZ_hypo

MadWeight: how does it works ?

Step 1: generate the matrix elements
type ./bin/mg5 to open the prompt



input : proc_card_mg5.dat

```
import model sm
generate p p > Z H , Z > mu+ mu- , H > b b~
output madweight HZ_hypo
```

output :
code for the evaluation of the weights in directory HZ_hypo

type ‘launch’ for the evaluation of the weights

and follow the instructions

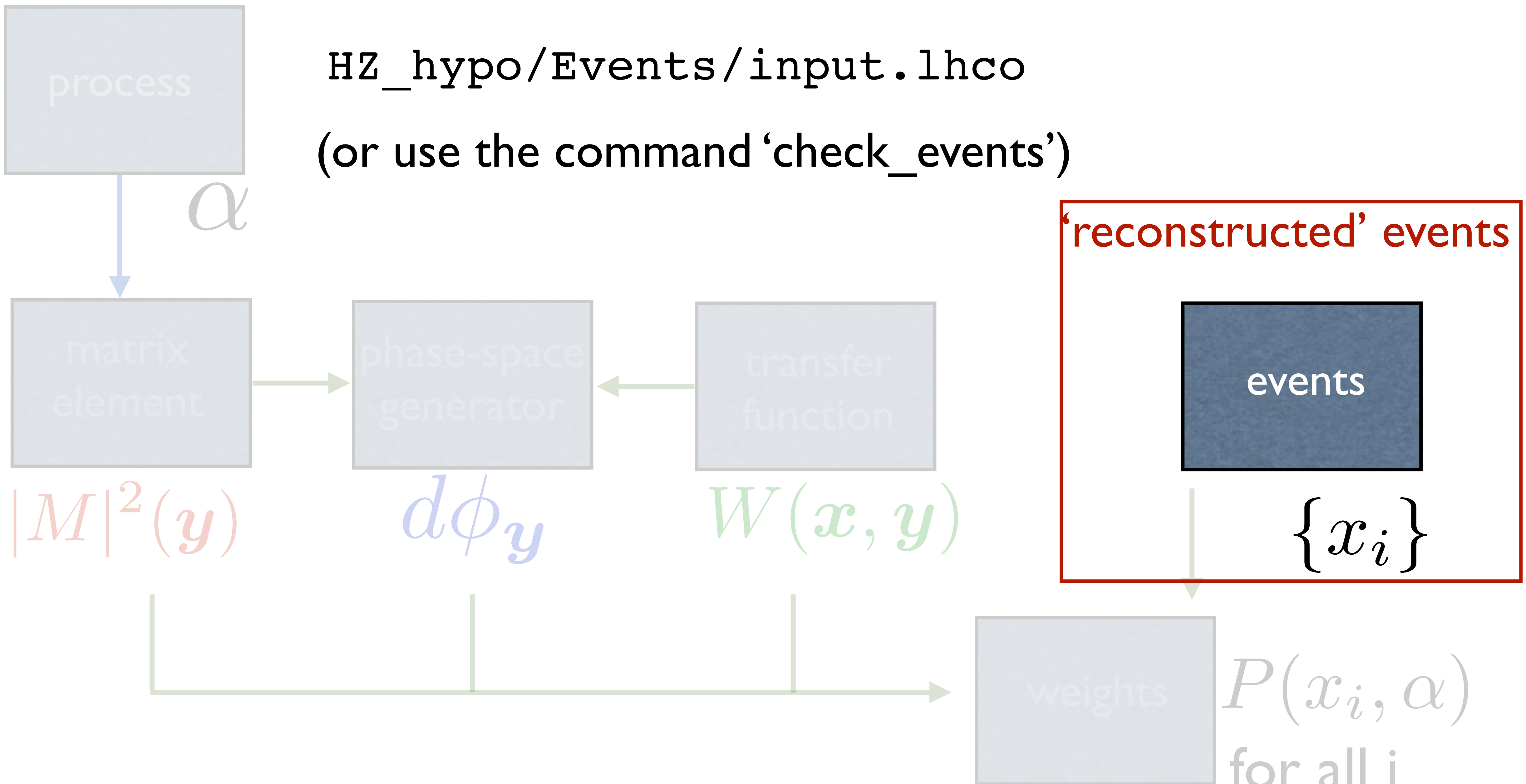
MadWeight: how does it works ?

framework for MEM that is reliable, user-friendly, reproducible, fast

Step 2: event file (LHCO format) must be copied at

`HZ_hypo/Events/input.lhco`

(or use the command 'check_events')



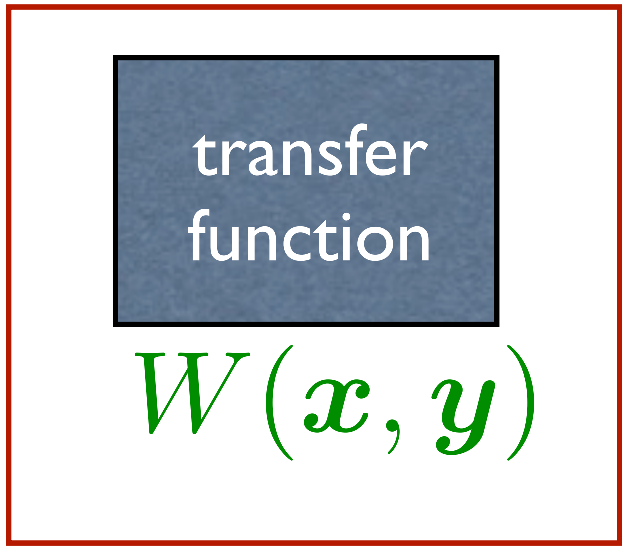
MadWeight: how does it works ?

Step 3: define your own TF parametrization

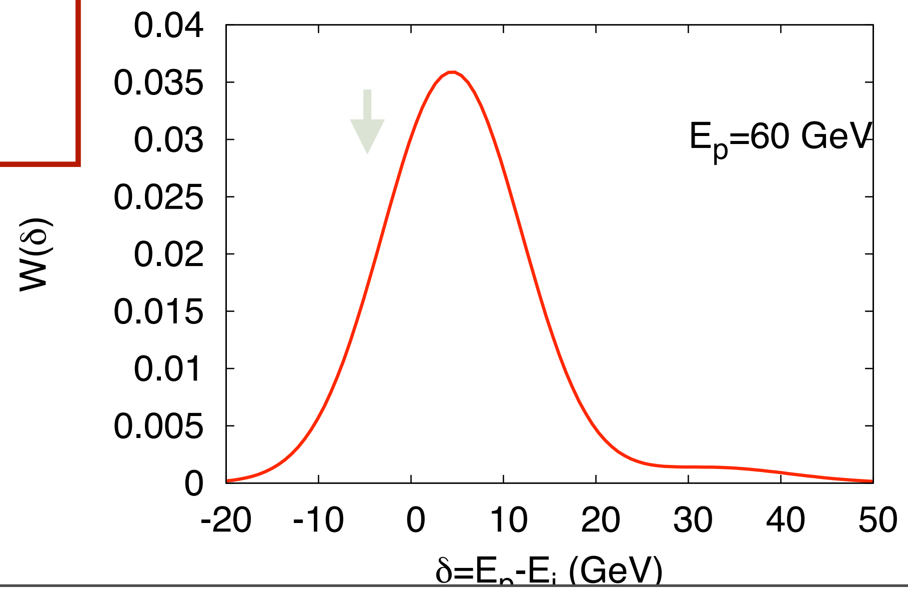
Source/MadWeight/transfer_function/data/TF_my_tf.dat

load TF: type ‘change_tf.py ’

```
Please choose your transfer_function
 0 / all_delta
 1 / dbl_gauss_pt_jet
 2 / gauss_on_leptons
 3 / single_gaussian
 4 / uniform
 5 / user
```



several predefined parametrizations

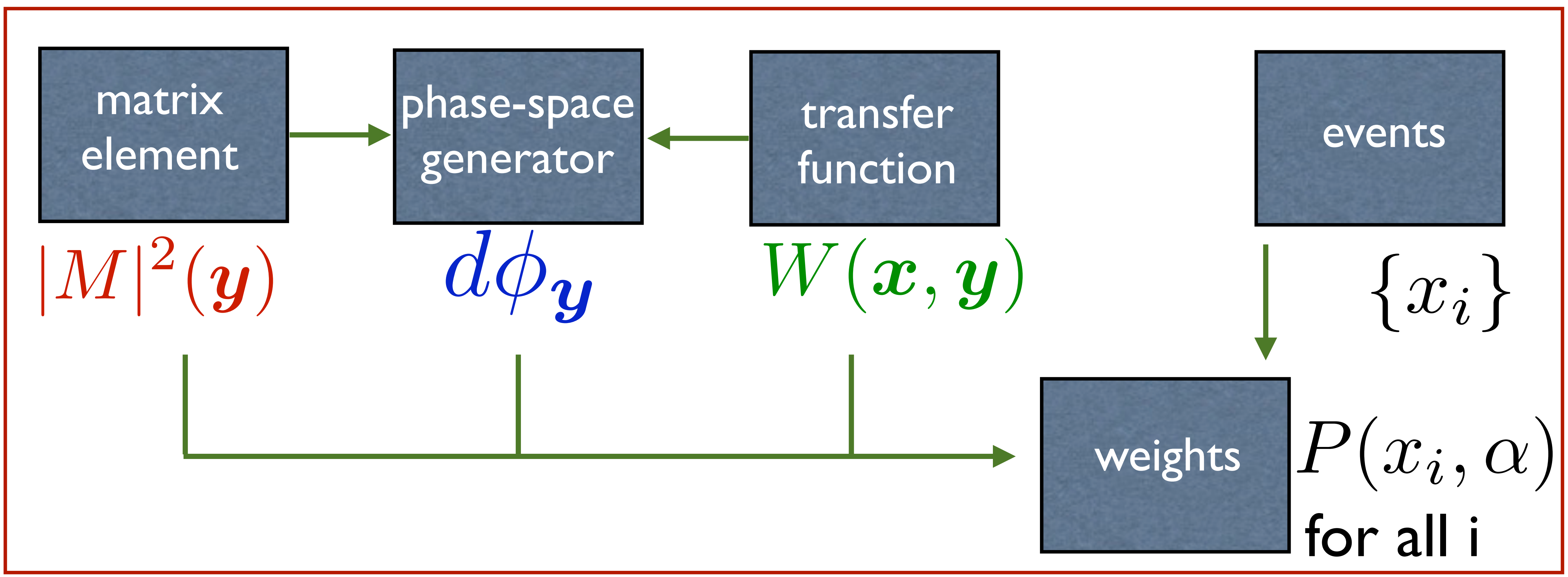


MadWeight: how does it works ?

Step 4: edit the cards with the input parameters associated with (1) the model, (2) the collider, (3) the submission of the jobs and (4) the transfer function

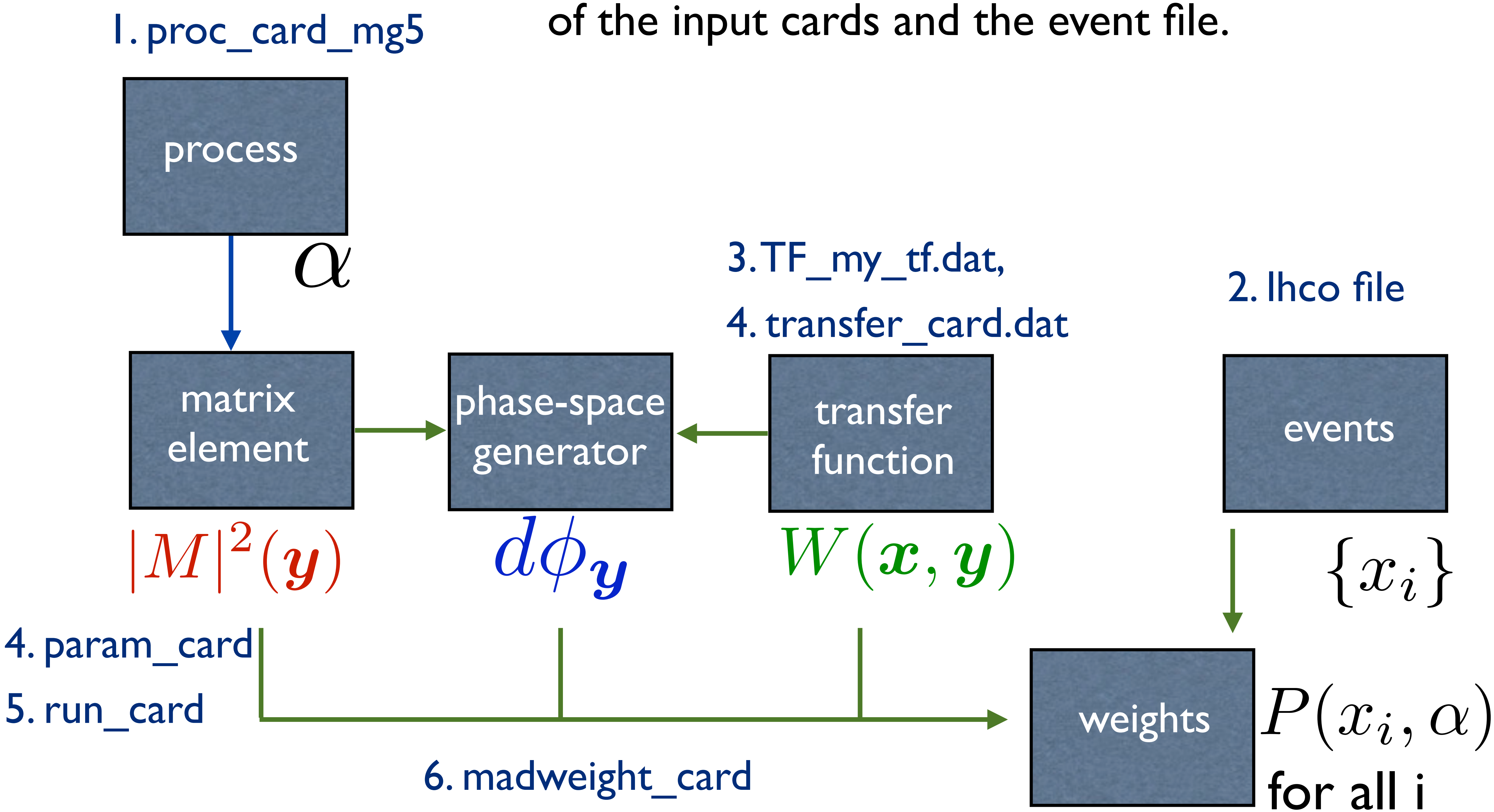
```
Do you want to edit one card (press enter to bypass editing)?
  1 / param      : param_card.dat
  2 / run        : run_card.dat
  3 / madweight  : madweight_card.dat
  4 / transfer   : transfer_card.dat
```

the code will load phase-space generator, evaluate the weights, collect the results



MadWeight: how does it works ?

The output is reproducible if you keep track of the input cards and the event file.



Normalization of the weights

MadWeight returns the values of the weight without the $1/\sigma^{obs}$ normalization

$$P(\boldsymbol{x}_i, \alpha) = \boxed{\frac{1}{\sigma^{obs}}} \frac{1}{N} \sum_{\text{jet perm.}} \int d\phi_{\boldsymbol{y}} |M|^2(\boldsymbol{y}) W(\boldsymbol{x}_i, \boldsymbol{y}) Acc(x)$$

To estimate σ^{obs} :

- Generate parton-level distributed according to $|M|^2(\boldsymbol{y})$ (parton-level cross section = σ)
- Smear the energies/angles of the reconstructed particles according to the transfer function $W(\boldsymbol{x}, \boldsymbol{y})$
- Count the number of events (after smearing) falling inside the acceptance region and record the associated efficiency ϵ

$$\sigma^{obs} = \sigma \times \epsilon$$

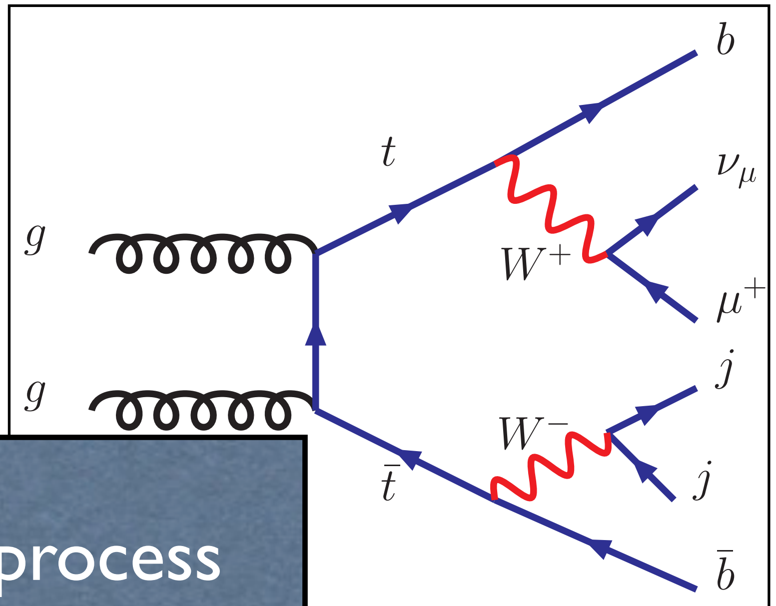
Part II

- Reconstruction of m_t (single-lepton tt samples)
  JHEP 1012 (2010) 068

- Characterization of a scalar resonance arXiv1306.6464

- Search for ttH Phys. Rev. Lett. 111, 091802 (2013)

Reconstruction of m_{top}

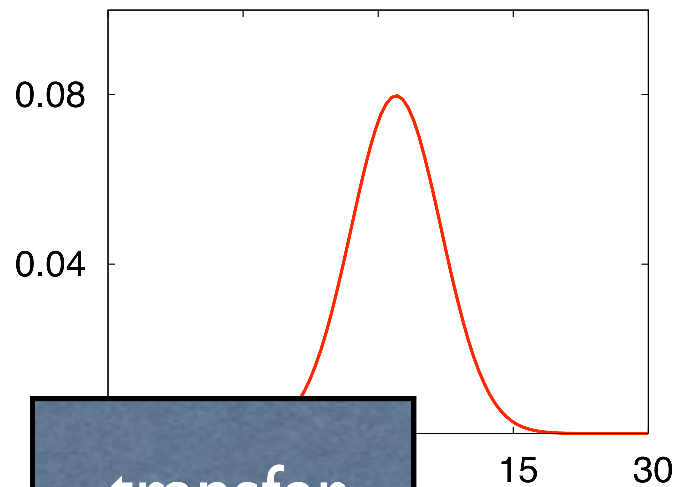


A. PROC: tt, single-lepton channel (LHC, 14 TeV)

process

α

B. TF: double Gaussian distribution on jet energies



C. MC sample: 20 events, generated with m_{top}=170 GeV

matrix element

phase-space generator

transfer function

events

$|M|^2(\boldsymbol{y})$

$d\phi_{\boldsymbol{y}}$

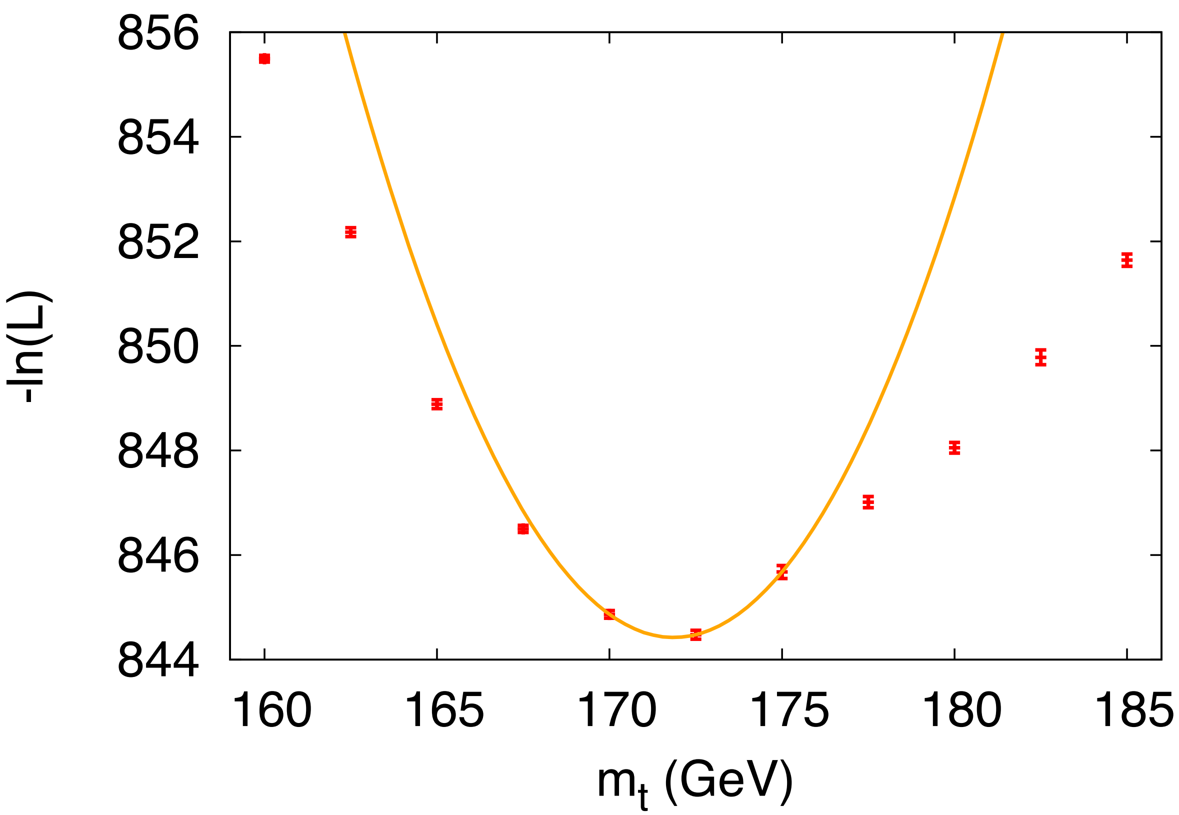
$W(\boldsymbol{x},\boldsymbol{y})$

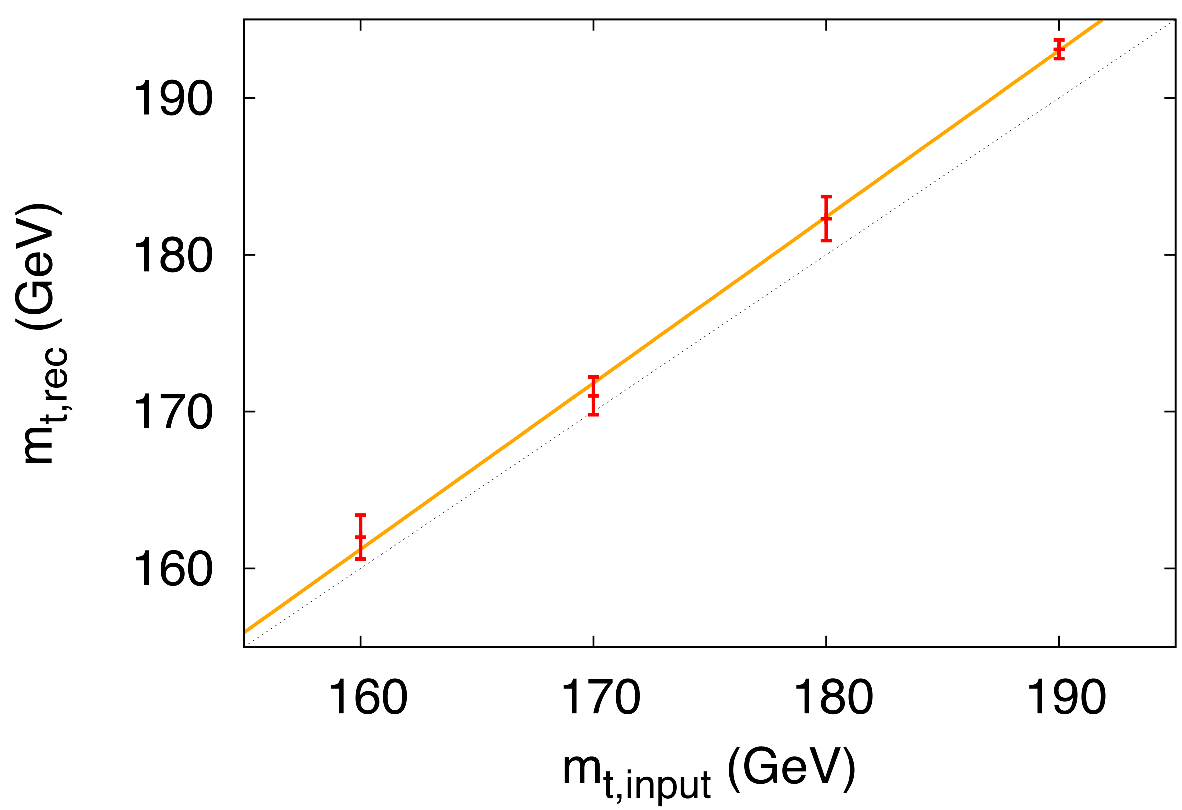
$\{x_i\}$

weights

$P(x_i, m_t)$
for each event,
m_t=160, 162.5, ...

Reconstruction of m_{top}





- Top quark mass can be reconstructed by minimizing

$$-\log[L(m_{\rm top})] = -\sum_{i=1}^{20} P(x_i|m_{\rm top})$$

  with respect to m_{top}

- The statistic error can be estimated by the half width of the distribution at $\log(L/L_{\max}) = 0.5$ ➡ $m_t = 171.9 \pm 2.0_{stat}$ GeV

- Bias: the probability density function assumed in the MEM does not exactly describe the phase-space distributions of the Monte Carlo events ➡ calibration procedure

Higgs Characterization in the 4l channel

- MELA approach: study of $h \to ZZ^* \to 4l$ with the MEM for discovery and characterization of the new 125 GeV resonance

Y. Gao et al, 2010
S. Bolognesi et al, 2012

- Here we consider the characterization of a scalar boson using the framework of an Effective Field Theory valid up to a scale Λ

  - Only one new state X(J^P) at the electroweak scale v (all other new states are heavier than Λ)

  - Includes all effects coming from the complete set of dimension 6 operators the for spin-0 case (above the EW symmetry breaking scale)

Higgs Characterization with the use of EFT

- Interaction of a spin 0 state with Z bosons

$$\mathcal{L}_{\rm eff} = \boxed{\frac{1}{2} c_\alpha \kappa_{SM} g_{HZZ} Z_\mu Z^\mu X_0}$$

SM case 0^+ state,

$$\boxed{-\frac{1}{4}\frac{1}{\Lambda}\left[c_\alpha \kappa_{HZZ} Z_{\mu\nu} Z^{\mu\nu} + s_\alpha \kappa_{AZZ} \tilde{Z}_{\mu\nu} Z^{\mu\nu}\right] X_0}$$

dimension-6 operators (above the EW scale)

Adjustable parameters:

κ_i in front of each parameter

angle α parametrizing the CP-mixing between 0^+ and 0^- states

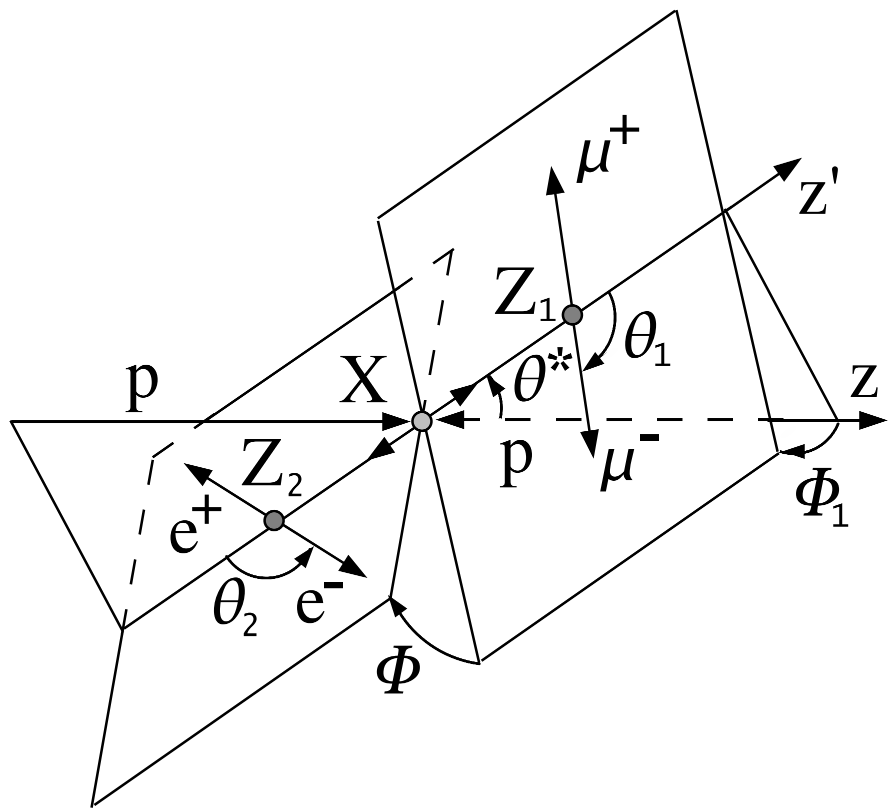
parameter	default value	description
Λ [GeV]	10^3	cutoff scale
$c_\alpha (\equiv \cos\alpha)$	1	mixing between 0^+ and 0^-
κ_i	1 or 0	dimensionless coupling parameter

```
######################################
## INFORMATION FOR FRBLOCK
######################################
Block FRBlock
    1 1.000000e+03 # Lambda
    2 1.000000e+00 # ca
    3 1.000000e+00 # kSM
    4 1.000000e+00 # kHff
    5 1.000000e+00 # kAff
    6 1.000000e+00 # kHaa
    7 1.000000e+00 # kAaa
      1.000000e+00 # kHza
      1.000000e+00 # kAza
      1.000000e+00 # kHgg
      1.000000e+00 # kAgg
      0.000000e+00 # kHzz
      0.000000e+00 # kAzz
      0.000000e+00 # kHww
      0.000000e+00 # kAww
      0.000000e+00 # kHda
      0.000000e+00 # kHdz
      0.000000e+00 # kHdw
```

Specific channel: X_0 into 4 charged leptons

- Kinematics: 5 variables

$$(\theta_1, \theta_2, \theta^*, m_1^*, m_2^*)$$



- Confidence level to reject hypothesis HD if hypothesis SM is realized
  a) significance with 1-dimension distribution ?
  b) significance with the matrix element weights ?

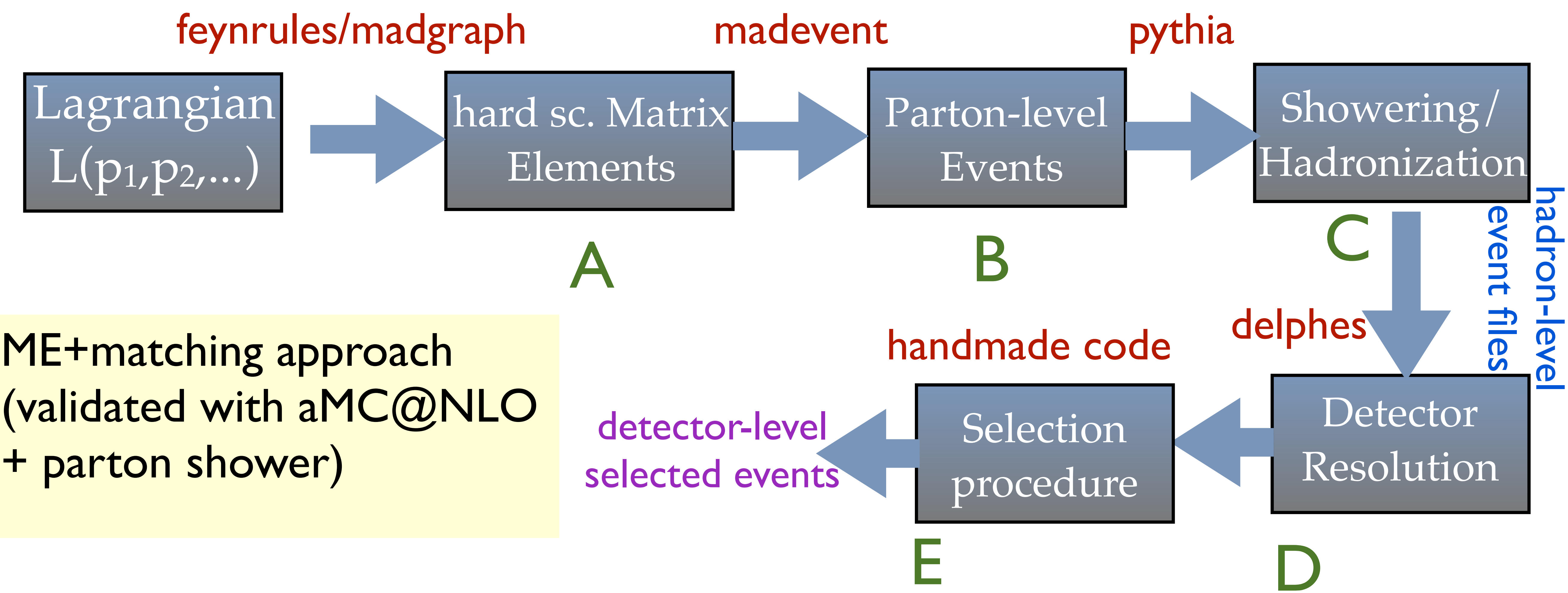
$$L_{\mathcal{O}} = \prod_i^N \frac{\sigma^{-1}_{\mathrm{HD}(c_\alpha)} \frac{d\sigma_{\mathrm{HD}(c_\alpha)}}{d\mathcal{O}}(\mathcal{O}_i)}{\sigma^{-1}_{\mathrm{SM}} \frac{d\sigma_{\mathrm{SM}}}{d\mathcal{O}}(\mathcal{O}_i)}\,.$$

$$L_{\mathrm{MEM}} = \prod_i^N \frac{|M_{HD(c_\alpha)}(i)|^2}{|M_{\mathrm{SM}}(i)|^2}$$

a) likelihood ratio based on 1-dim. distribution

b) likelihood ratio based on matrix elements

Step 1: generation of events



ME+matching approach (validated with aMC@NLO + parton shower)

input parameters:

A. model+proc_card_mg5.dat

B. run_card.dat, param_card.dat

C. pythia_card.dat

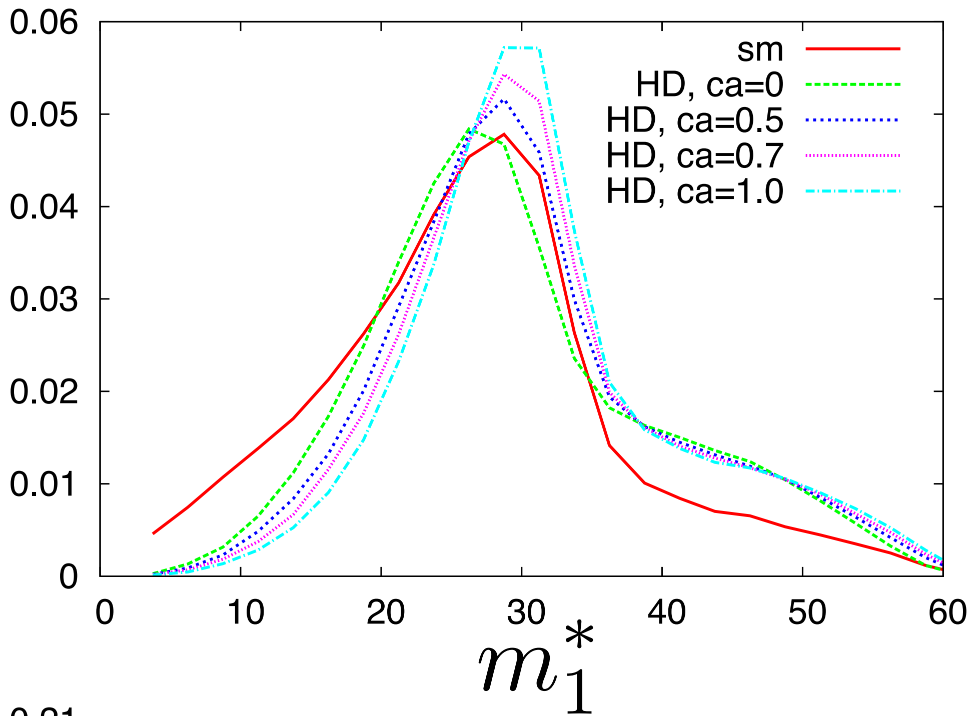
D. delphes_card.dat, delphes_trigger.dat

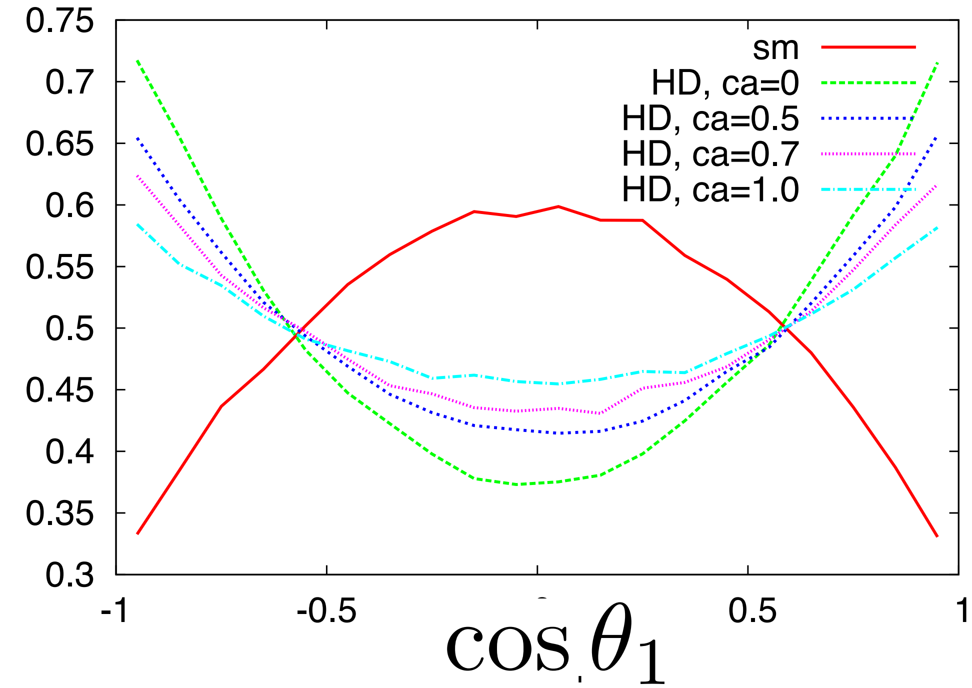
E. p_T > 7 GeV, |y|<7 GeV (python code)

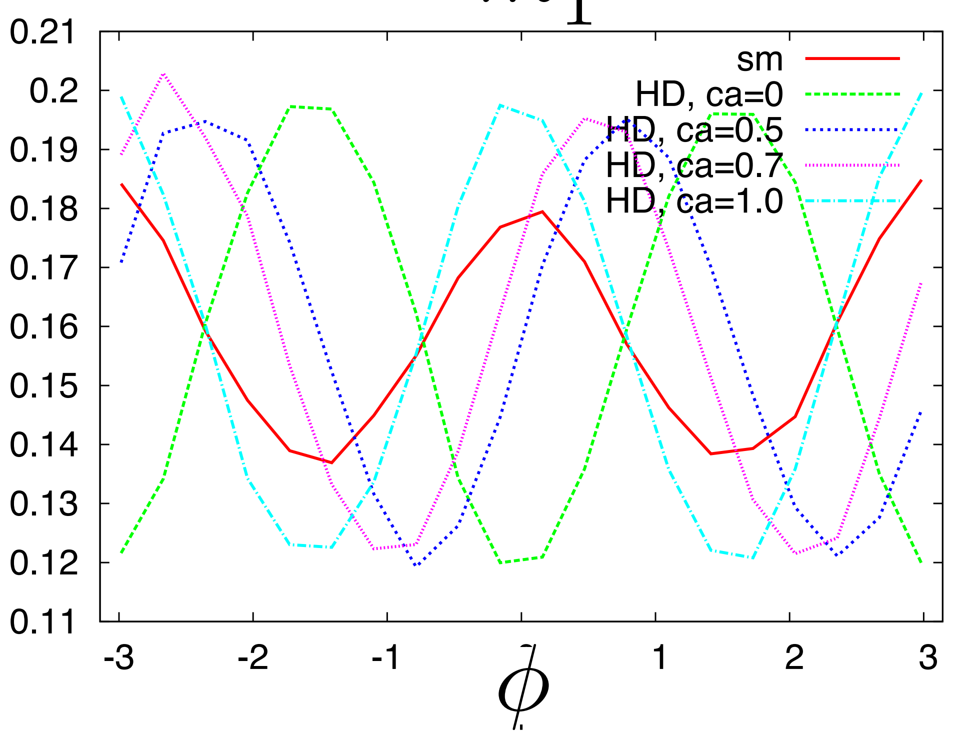
Step 2: distributions of events w.r.t. discriminant

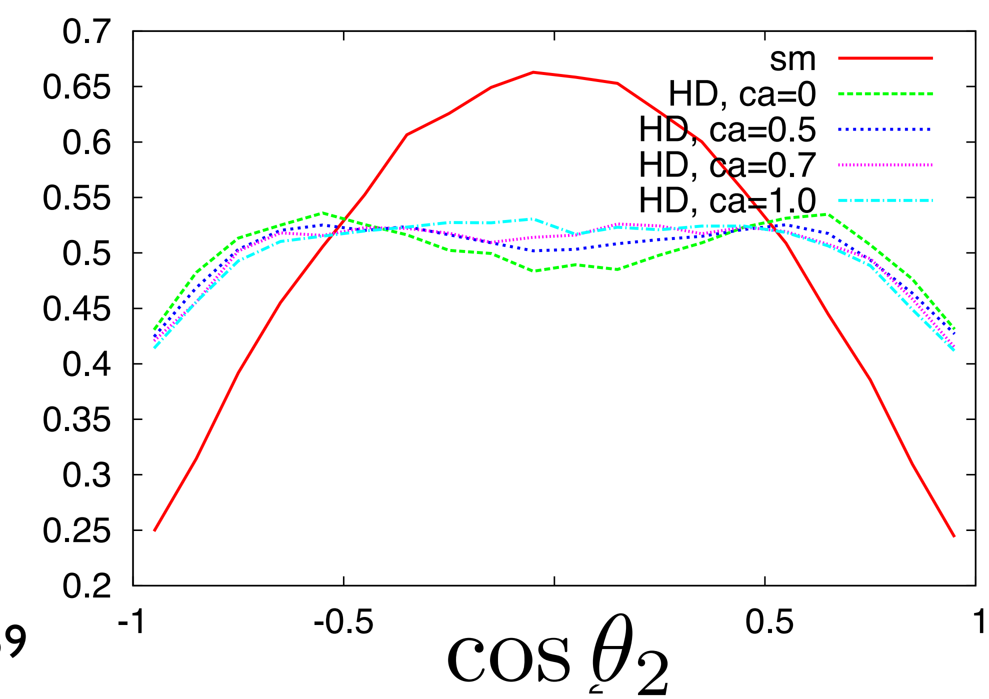
a) 1-dimension distributions:

MadAnalysis: reads the event files and build histograms with respect to each variable $(\theta_1, \theta_2, \theta^*, m_1^*, m_2^*)$









Step 2: distributions of events w.r.t. discriminant

b) MEM-based distributions:

MEM weights: set $W(x,y) = \delta(x-y)$
(infinite resolution on lepton kinematics)

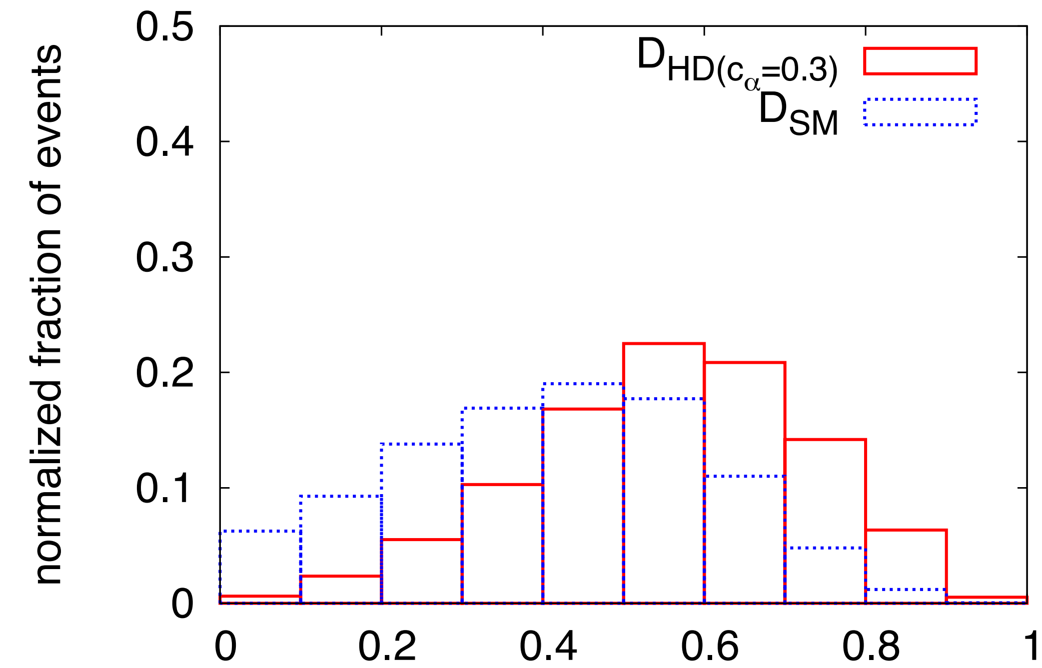
$$P(x|\mathrm{SM}) \propto |M_{\mathrm{SM}}(x)|^2$$

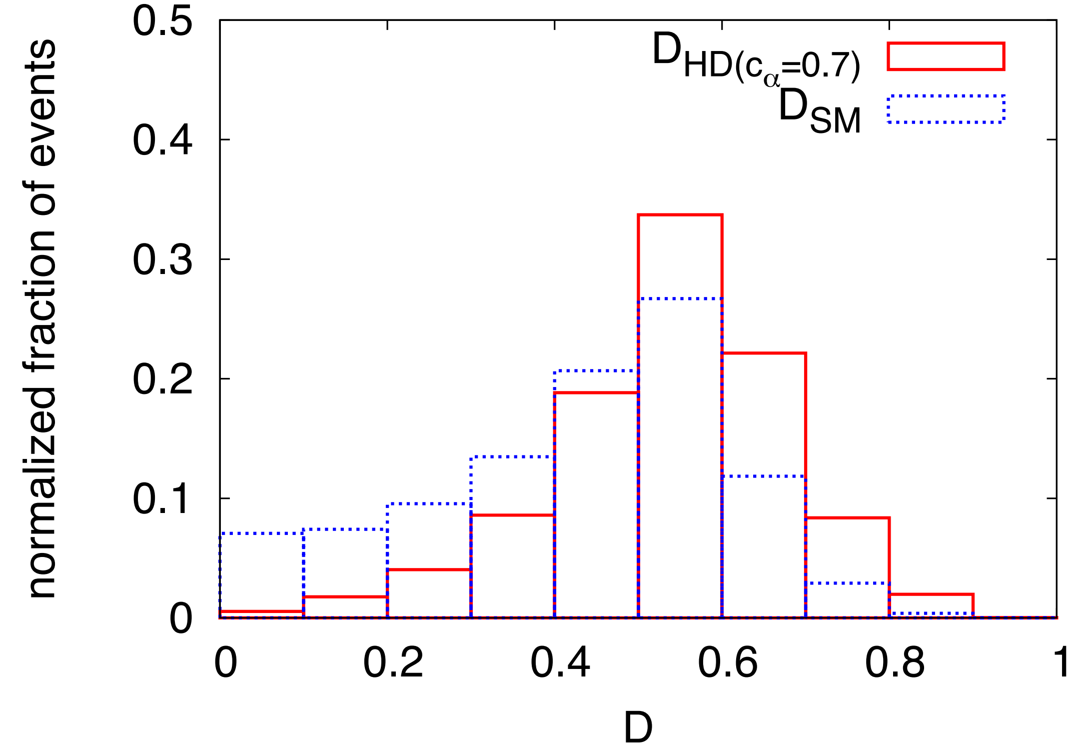
$$P(x|\mathrm{HD}) \propto |M_{\mathrm{HD}}(x)|^2$$

the calculation of the MEM weights is straightforward (madgraph standalone)

distribution of events w.r.t the ratio of the weights

$$D(x) = \left[1 + \frac{|M_{\mathrm{HD}}|^2(x)}{|M_{\mathrm{SM}}|^2(x)}\right]^{-1}$$





Step 3: generate pseudo-experiments

- 1M pseudo-experiments with N=10 events under each assumptions (SM of HD)

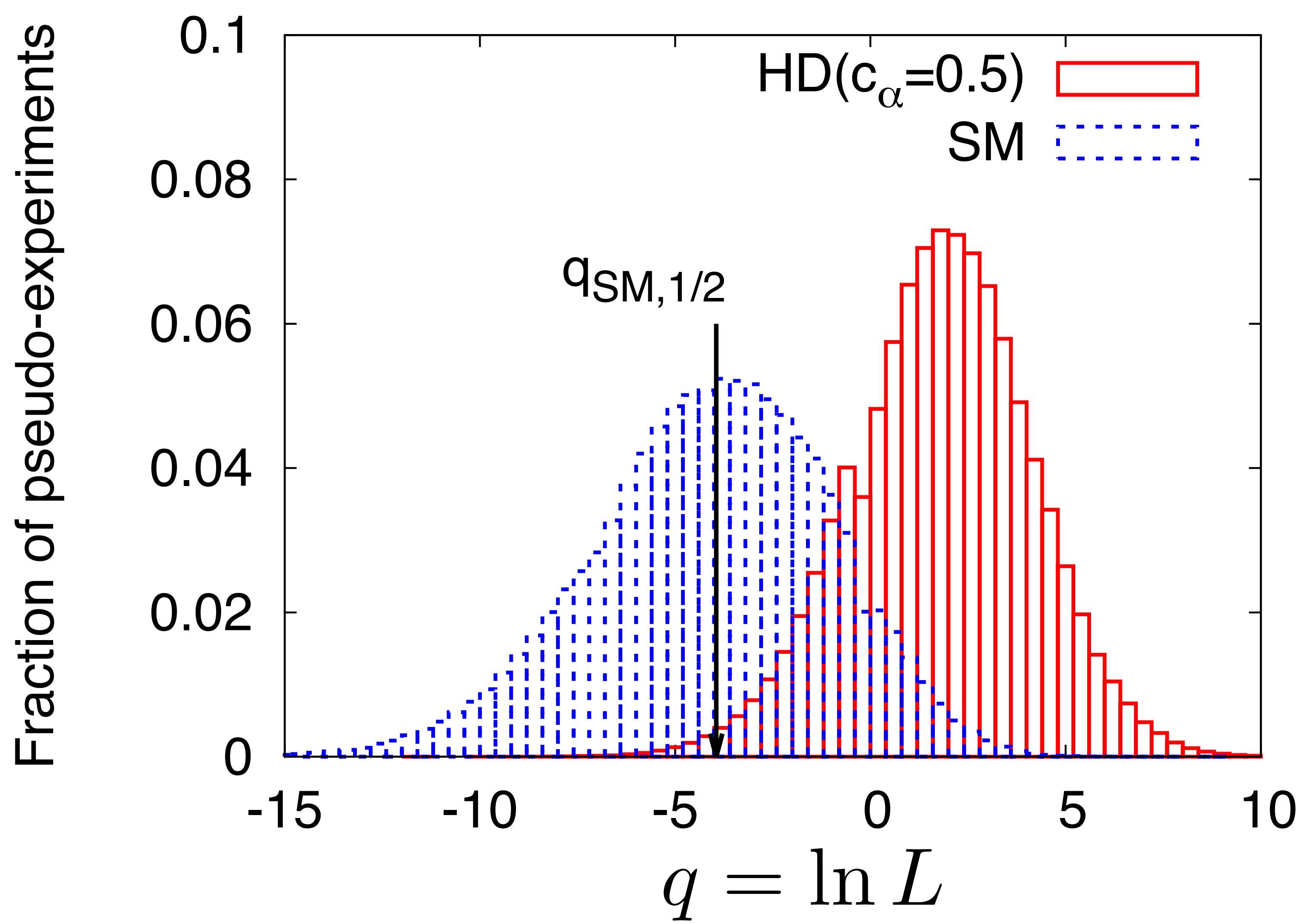
Two options: (1) generate a 10^7 MC events for each assumptions SLOW

(2) generate the kinematic variables of interest according to the previous distributions FAST

For each pseudo-experiment, the likelihood functions can be evaluated:

$$L_{\mathcal{O}} = \prod_{i}^{N} \frac{\sigma^{-1}_{\mathrm{HD}(c_\alpha)} \frac{d\sigma_{\mathrm{HD}(c_\alpha)}}{d\mathcal{O}}(\mathcal{O}_i)}{\sigma^{-1}_{\mathrm{SM}} \frac{d\sigma_{\mathrm{SM}}}{d\mathcal{O}}(\mathcal{O}_i)} . \qquad L_{\mathrm{MEM}} = \prod_{i}^{N} \frac{|M_{HD(c_\alpha)}(i)|^2}{|M_{\mathrm{SM}}(i)|^2}$$

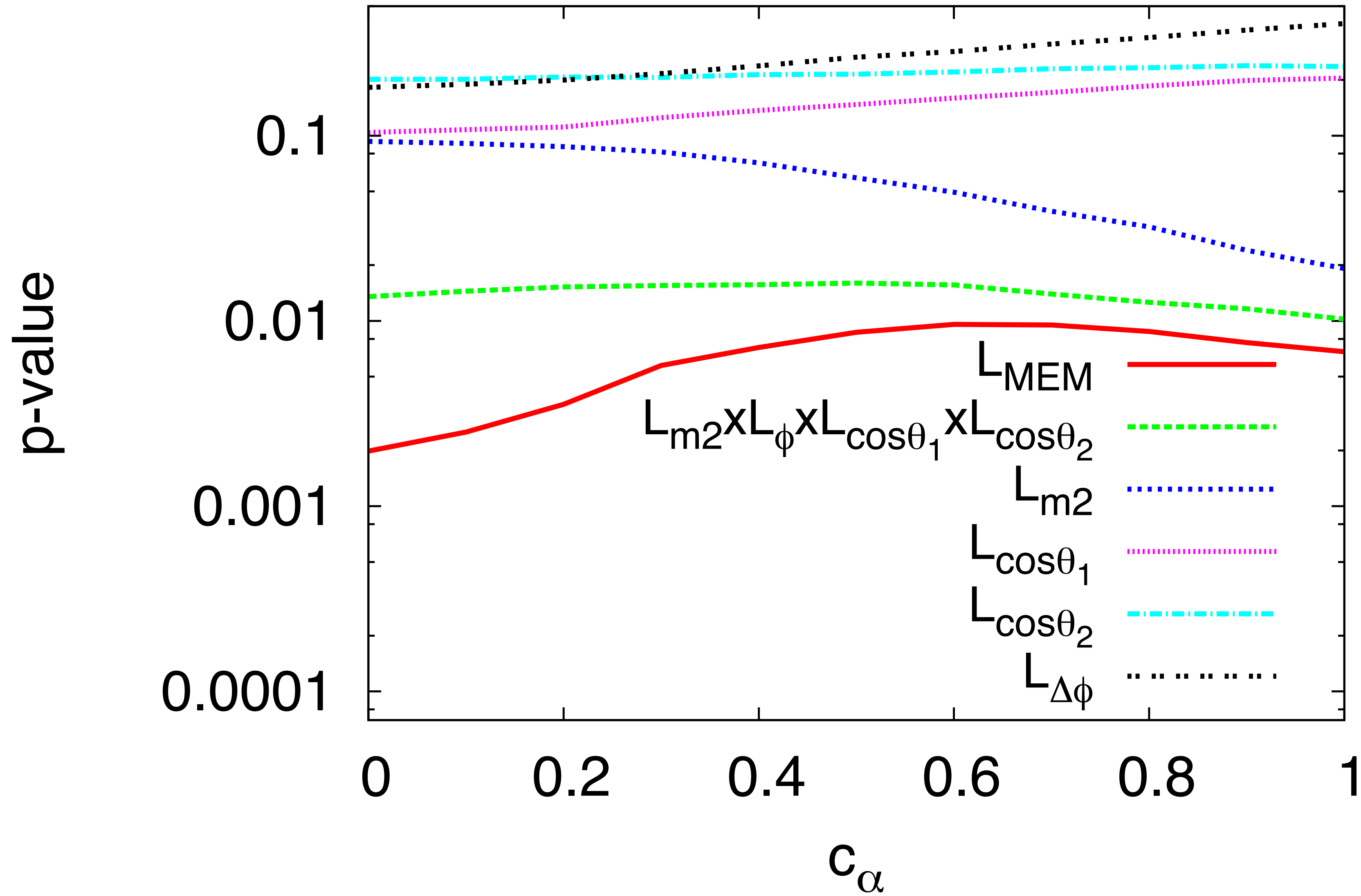
Step 4: estimate the significance



Significance estimated by calculating the median $q_{\mathrm{SM},1/2}$ of the SM distribution and by counting the fraction of pseudo-experiments in the HD distribution with $q < q_{\mathrm{SM},1/2}$

This fraction = expected p-value associated with the test of rejecting hypothesis HD if the SM hypothesis is realized.

Step 4: estimate the significance

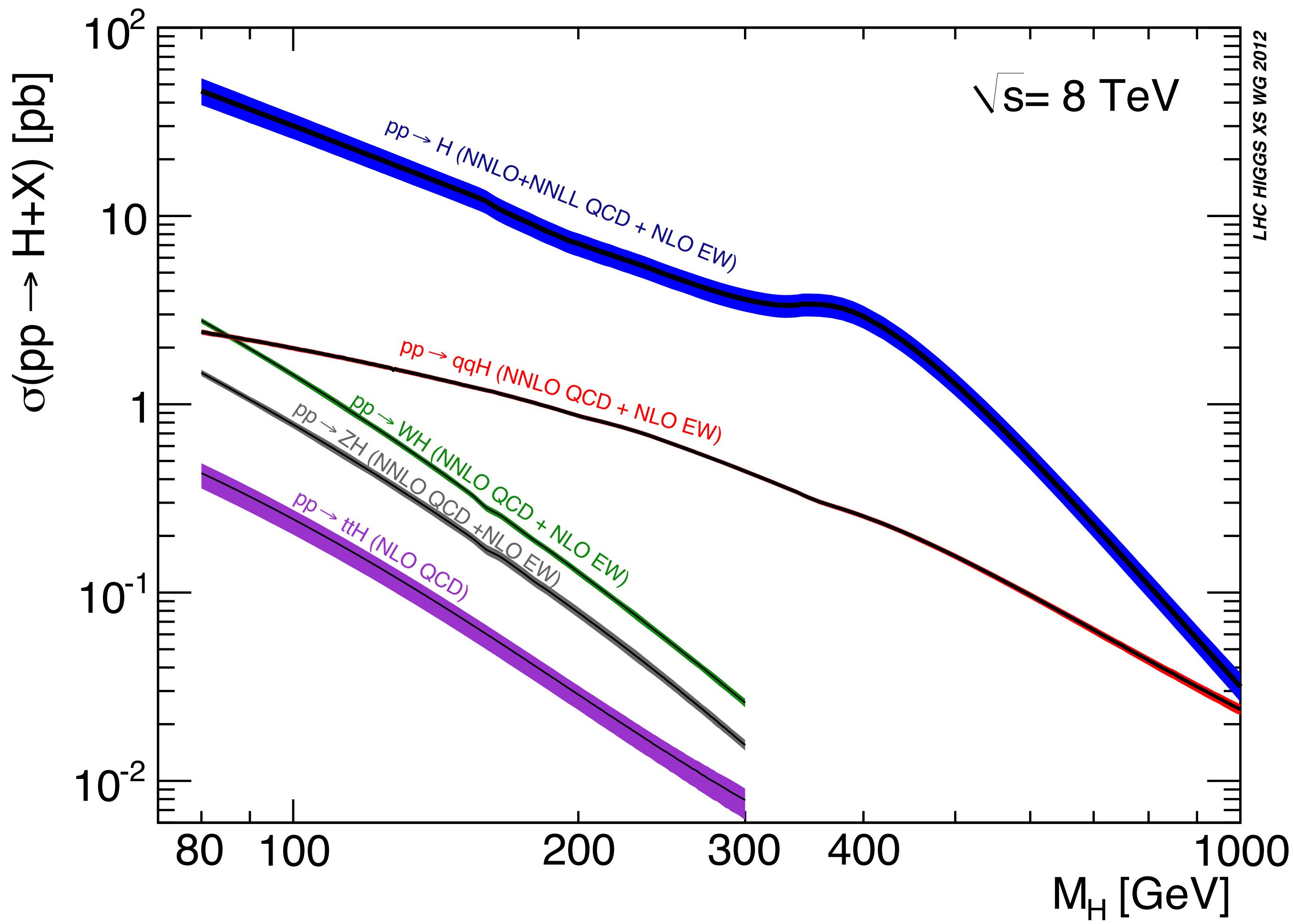


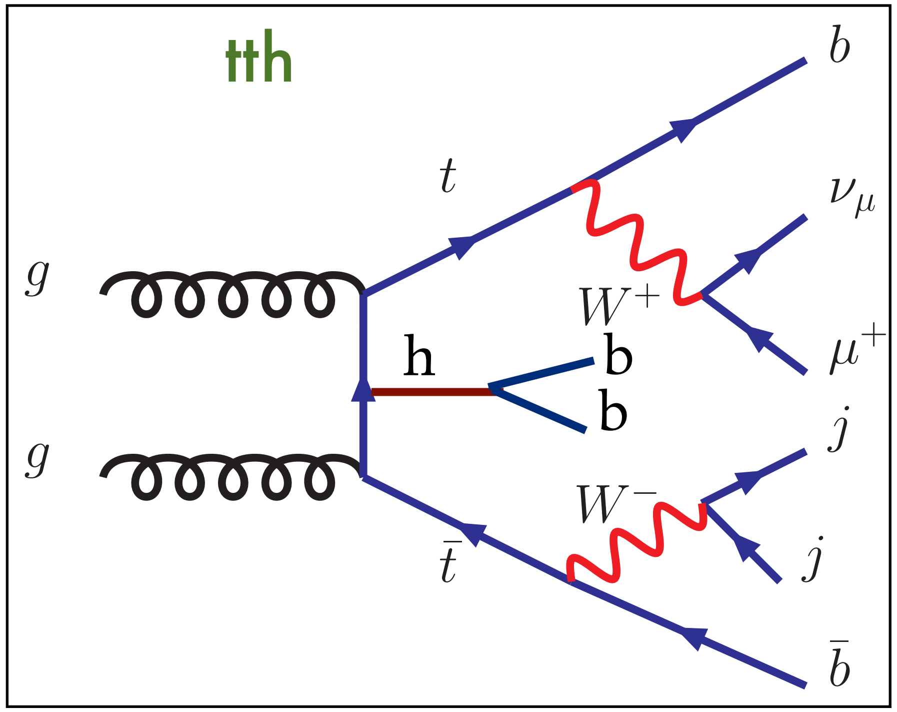
PA et al, 2013

The optimal significance is reached with the MEM-based likelihood approach

tt + scalar boson at the LHC with the MEM

P.A. P. de Aquino, F. Maltoni, O. Mattelaer





W⁺ W⁻ b $\bar{b}$ b $\bar{b}$ final state

1) The production rate is small

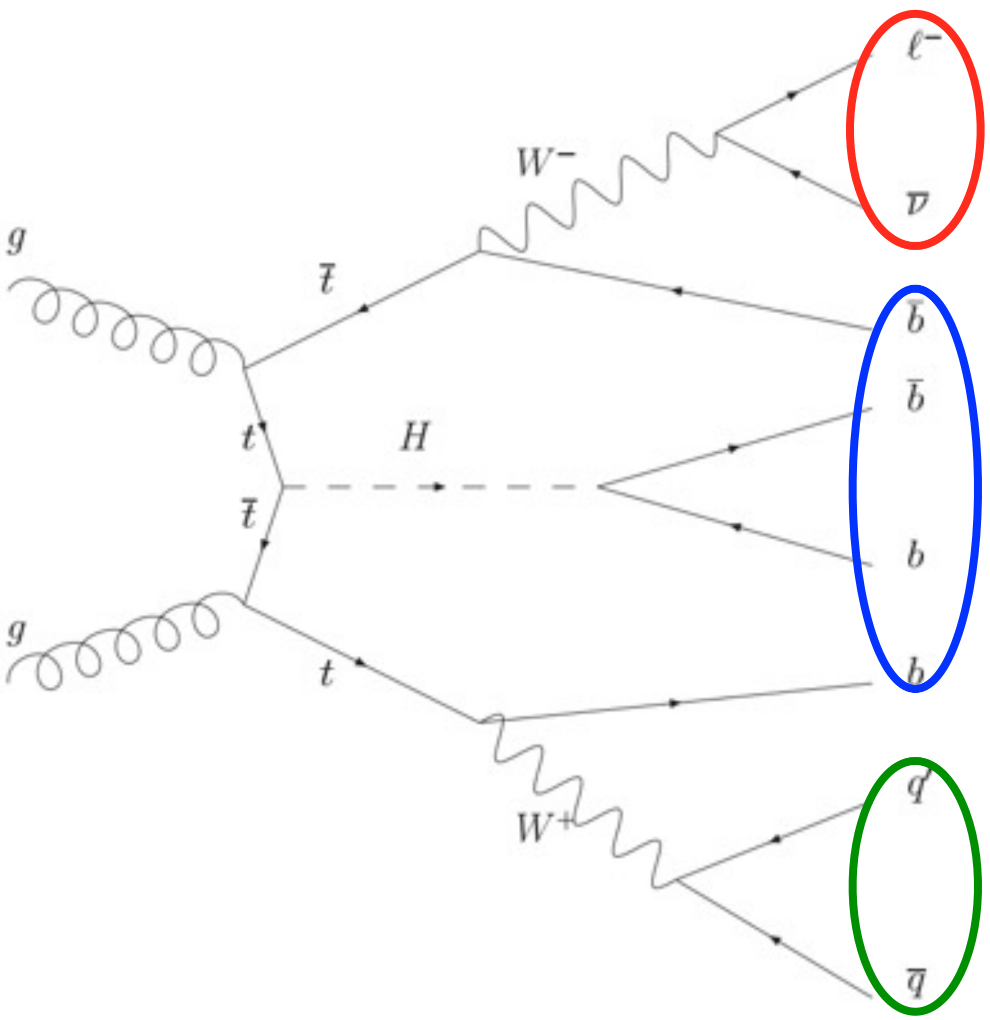
NLO prediction @ LHC:

8 TeV	14 TeV
0.137 pb	0.632 pb

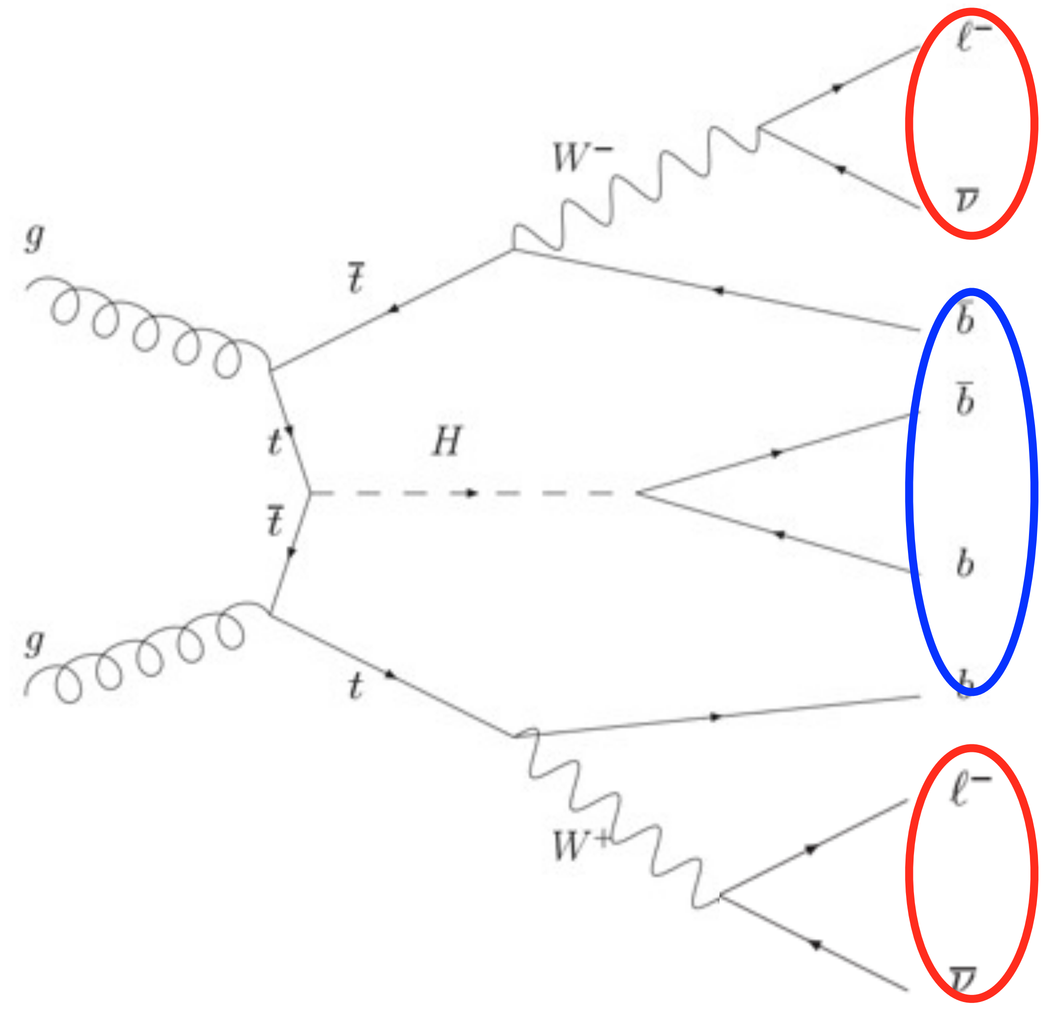
2) Challenging backgrounds:

- t $\bar{t}$ + jets
- combinatorial background (identification of the b-jets coming from the scalar boson)

Decay channels



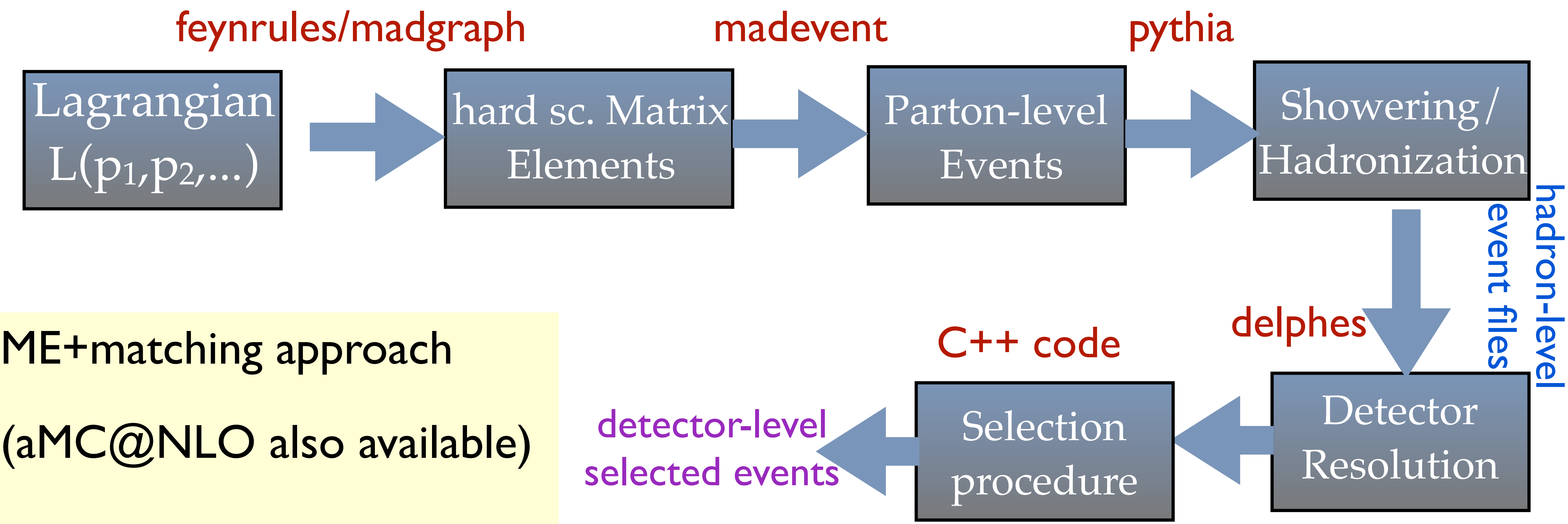
single-lepton final state



di-lepton final state

Q: is the discriminating power in the di-lepton channel higher or less than the one in the semi-lepton channel ?

Step 1: event generation



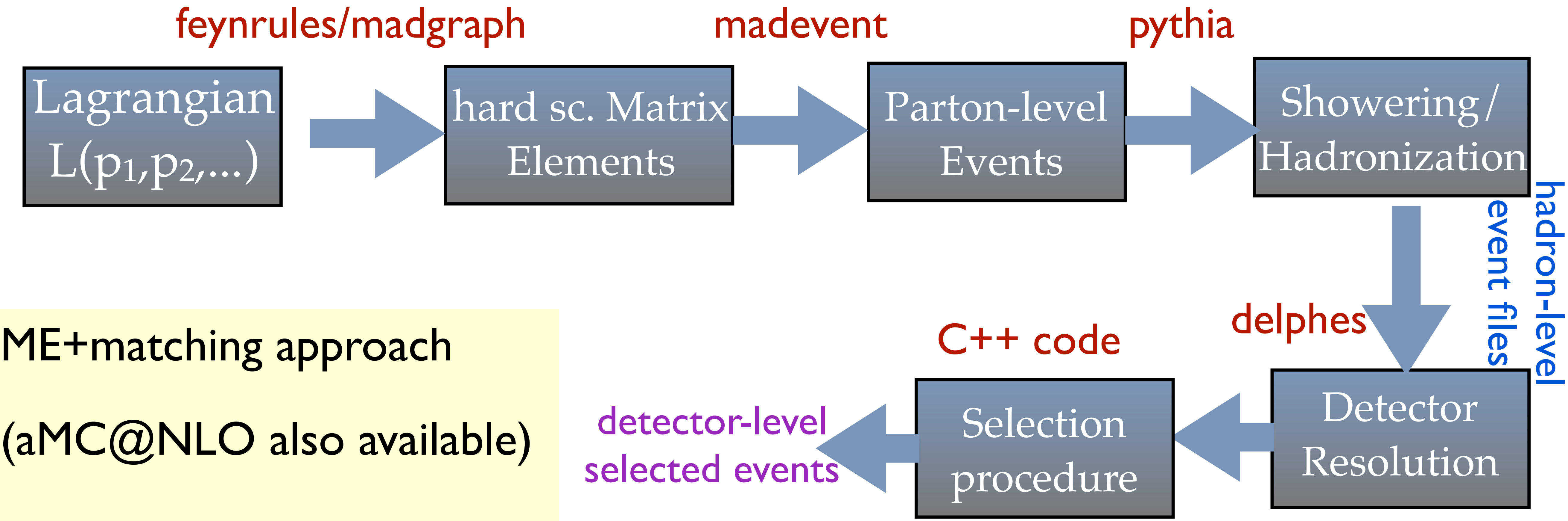
ME+matching approach

(aMC@NLO also available)

- Leptons: P_T > 20 GeV and $|\eta|$ < 2.4
- Jets: anti-k_T with R=0.5 , P_T > 30 GeV and $|\eta|$ < 2.5
- At least 4 b-jets required

process	incl. σ	efficiency	σ^{rec}
$t\bar{t}h$, single-lepton	111 fb	0.0485	5.37 fb
$t\bar{t}h$, di-lepton	17.7 fb	0.0359	0.634 fb
$t\bar{t}$+jets, single-lepton	256 pb	0.463×10^{-3}	119 fb
$t\bar{t}$+jets, di-lepton	40.9 pb	0.168×10^{-3}	6.89 fb

Step 1: event generation



ME+matching approach

(aMC@NLO also available)

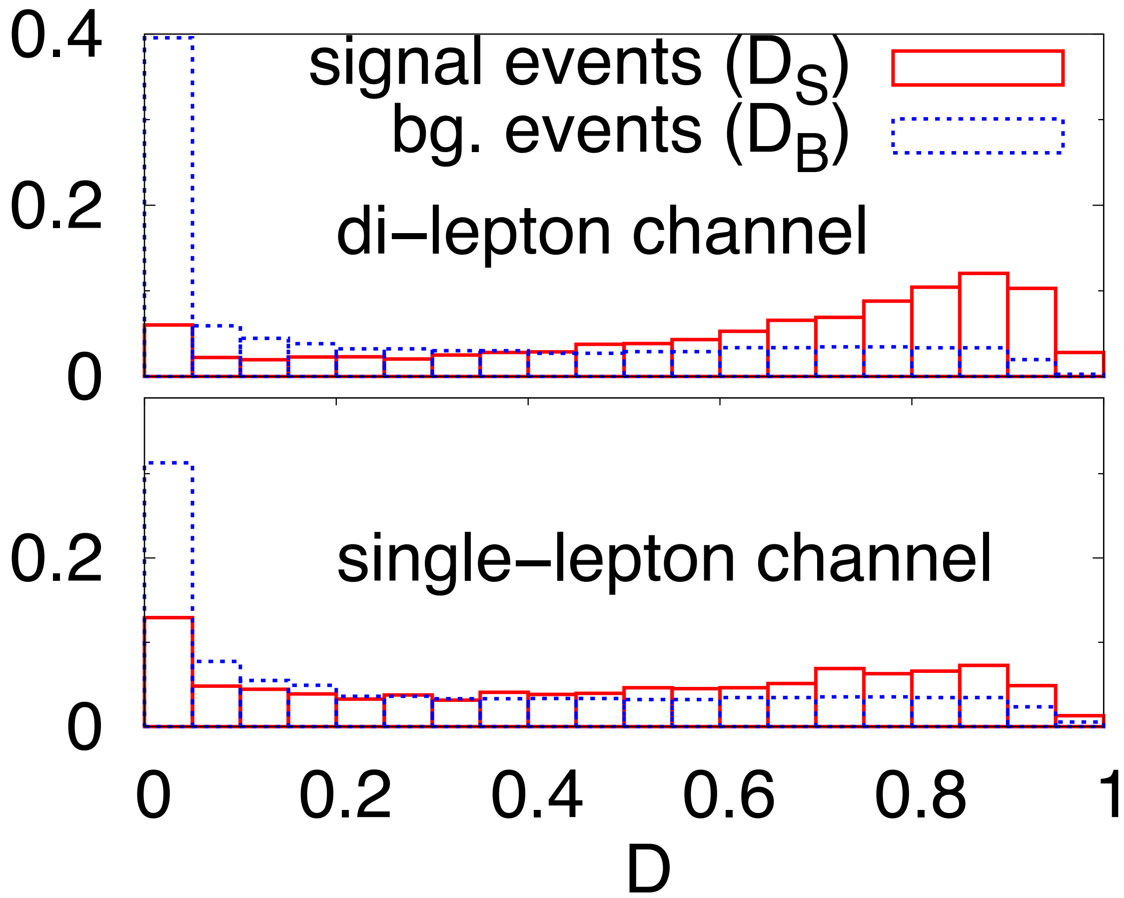
single-lepton: S/B ~ 1/22

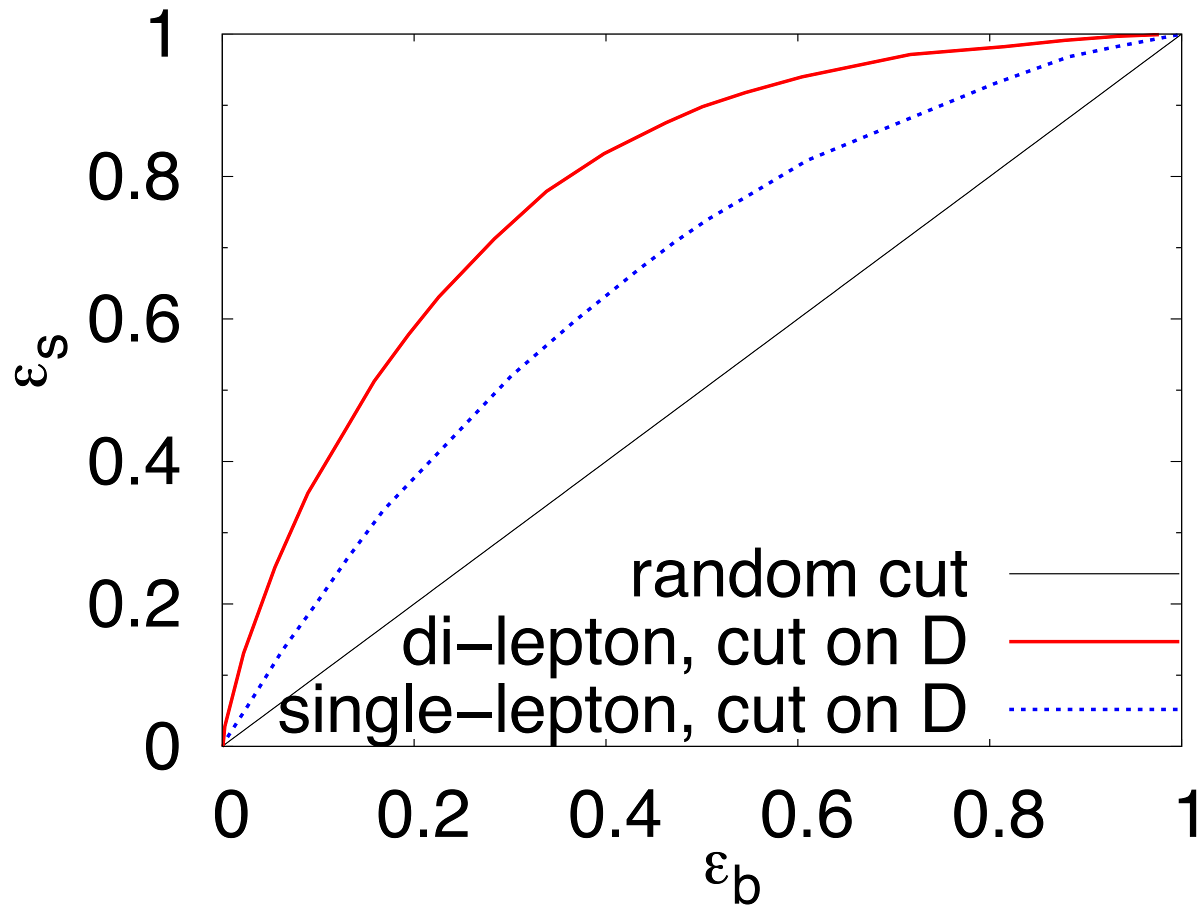
di-lepton: S/B ~ 1/11

- Leptons: $P_T > 20$ GeV and $|\eta| < 2.4$
- Jets: anti-k_T with R=0.5 , $P_T > 30$ GeV and $|\eta| < 2.5$
- At least 4 b-jets required

process	incl. σ	efficiency	σ^{rec}
$t\bar{t}h$, single-lepton	111 fb	0.0485	5.37 fb
$t\bar{t}h$, di-lepton	17.7 fb	0.0359	0.634 fb
$t\bar{t}$+jets, single-lepton	256 pb	0.463×10^{-3}	119 fb
$t\bar{t}$+jets, di-lepton	40.9 pb	0.168×10^{-3}	6.89 fb

Step II: event distributions w.r.t. MEM discriminant



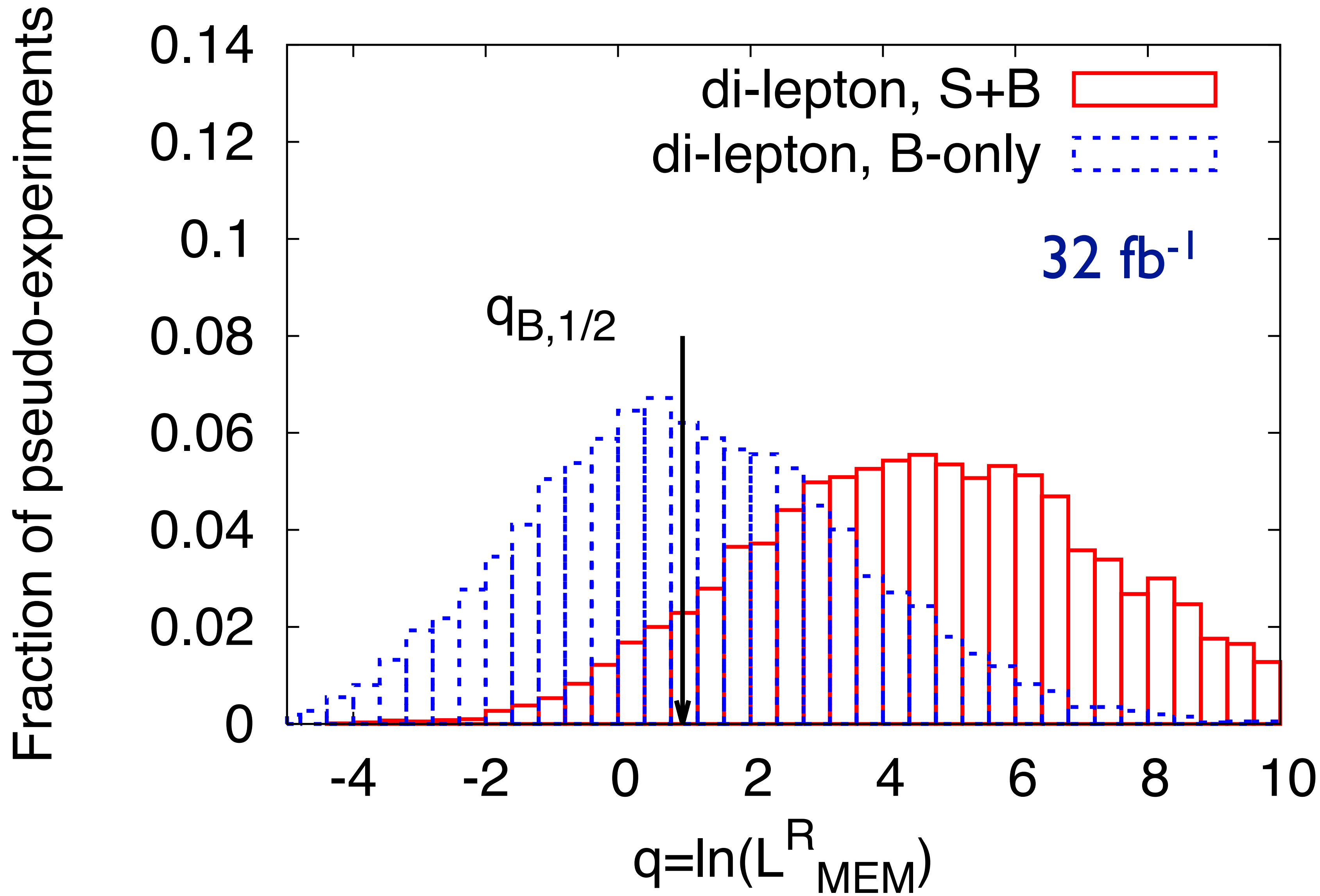


Left-hand plot: distributions of events with respect to $D(x) = \left[1 + \frac{P(x|B)}{P(x|S)}\right]^{-1}$

Right-hand plot: signal vs. background efficiencies resulting from a cut on $D > D_{min}$

➡ dilepton-channel is cleaner, more manageable combinatorial background

Step IV: significance

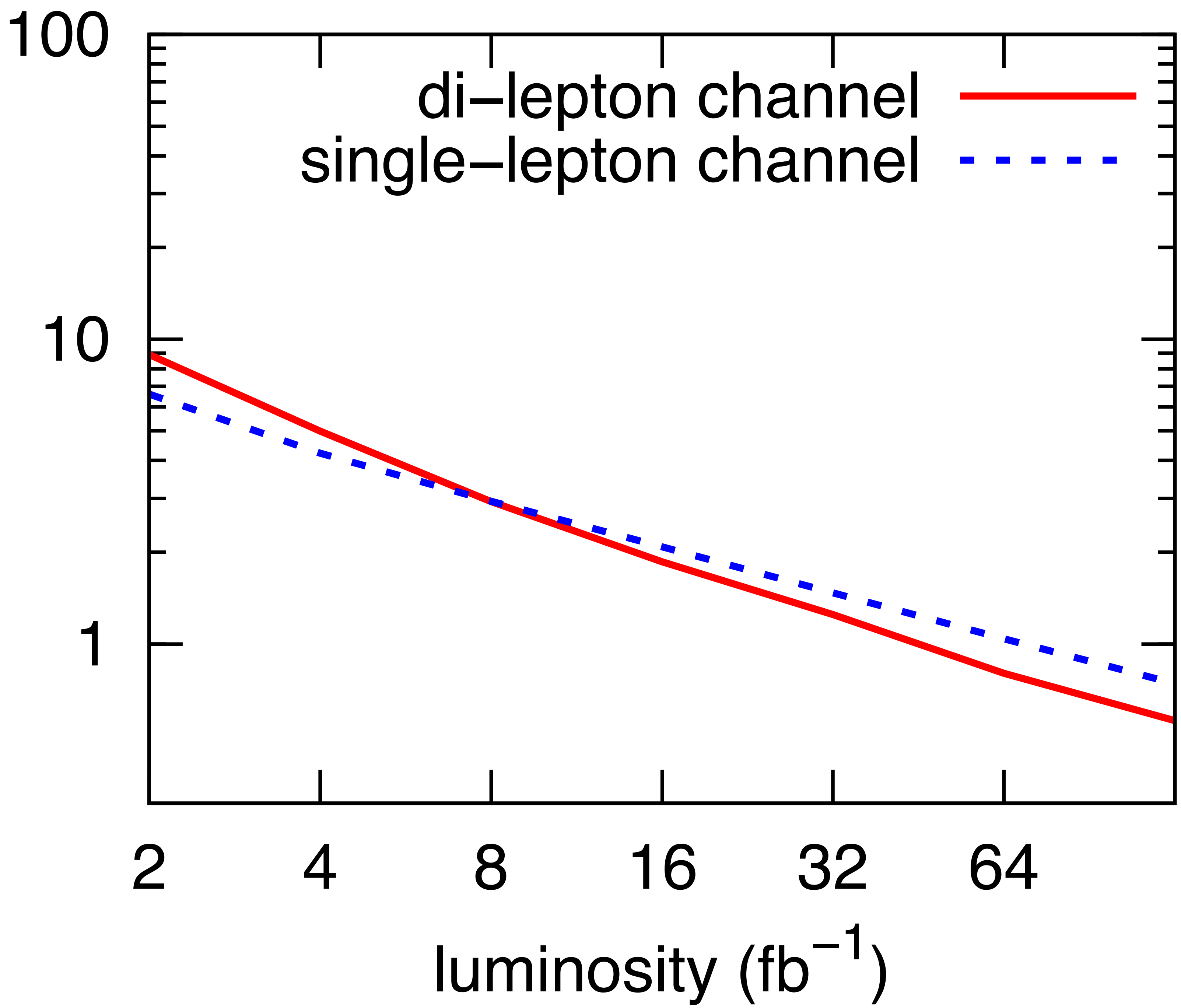


Test: confidence level in rejecting S+B hypothesis if B-only hypothesis is realized

- compute $q_{B,1/2}$ = median of the B-only distribution
- estimate the p-value as the faction of events in the S+B distribution satisfying $q < q_{B,1/2}$
- C.L. = 1-p

Step IV: significance

- Rescale the signal cross section by a factor μ such that S+B is excluded at 95% C.L



MEM versus counting

- Redo the analysis with a likelihood that is based on the number of events

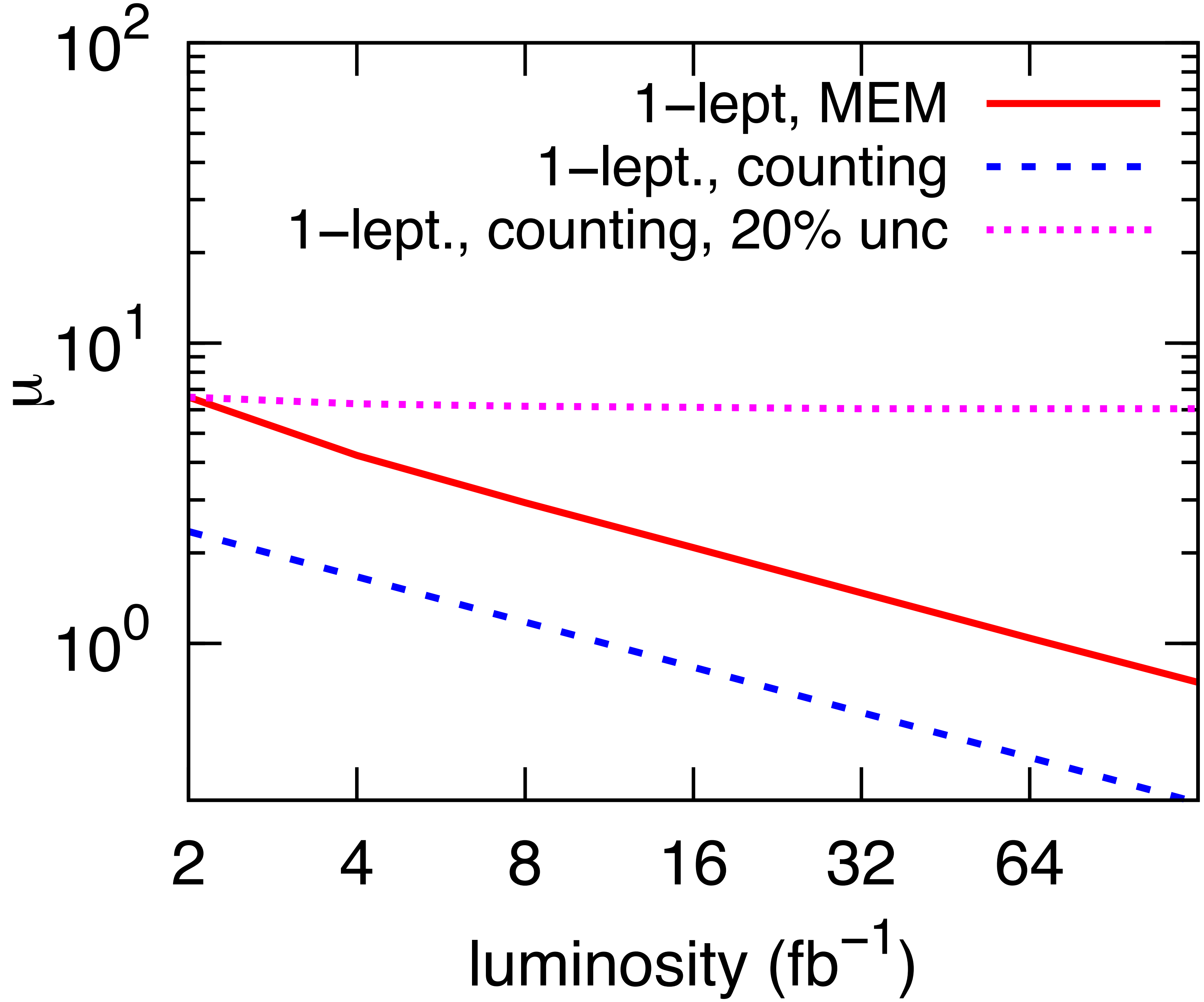
$$L^R_{\text{counting}} = \frac{\text{pois}(s+b|s_0+b_0)}{\text{pois}(b|b_0)}$$

with s_0, b_0 the expected number of signal and background events

- Also consider the effect of 20% systematic uncertainties on b_0 (=the expected number of background events):

This can be done by smearing the value of b_0 according to a log-normal distribution (mean=b_0, std=$0.2b_0$) before drawing the number of background events according to a Poisson distribution in each pseudo-experiment

MEM versus counting



Already a 20% uncertainty on b_0 hampers the counting analysis

Other applications of the MEM

- VBF scalar boson production J.R. Anderson, C. Englert, M. Spannowsky ‘12
- Characterization of a scalar boson S. Bolognesi et al ’12
- b-charge identification Gedalia, Isidori, Maltoni, Perez, Selvaggi, Soreq ‘12
- Stops searches P. Van Mulders et al.
- Z’ searches S. Basegmez, G. Bruno
- Differential weight A. Pin, O. Mattelaer
- ...

Conclusion

- I discussed several aspects of the Matrix Element method:
  - The use of MEM to establish a formal maximum significance,
  - The inclusion of beyond leading-order corrections in the definition of the weights,
  - The practical evaluation of the MEM weights.
- I presented three examples of application:
  - m_{top} reconstruction
  - Higgs characterization
  - Search for ttH