MADWEIGHT

A tool for Matrix Element Methods

The 2012 FeynRules/MadGraph
School on phenomenology

October 4, 2012

Pierre Artoisenet
NIKHEF

OUTLINE

Morning:

- **What is the Matrix Element Method ?**
- **What is MadWeight ?**

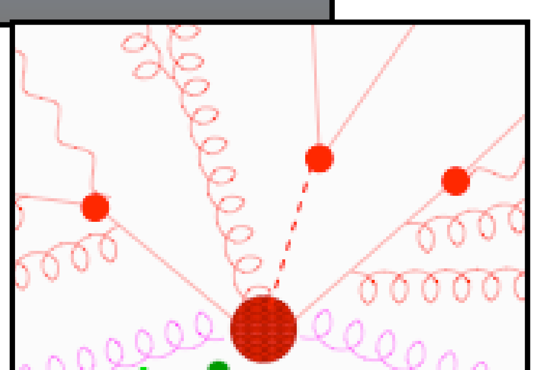
Afternoon:

- **How to use MadWeight ?**

New physics searches at the LHC

Experimental events

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



PROBLEM #1:

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

Two distinct approaches are used at hadron colliders:

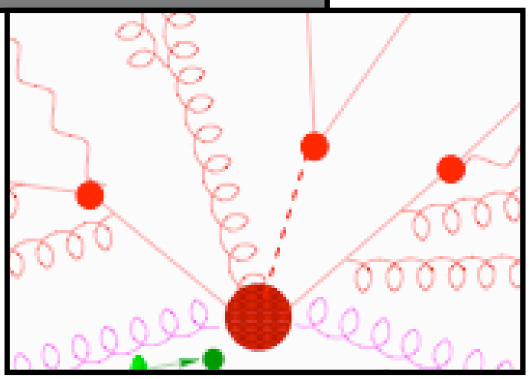
Approach 1: the discriminator is built upon Monte Carlo events only,

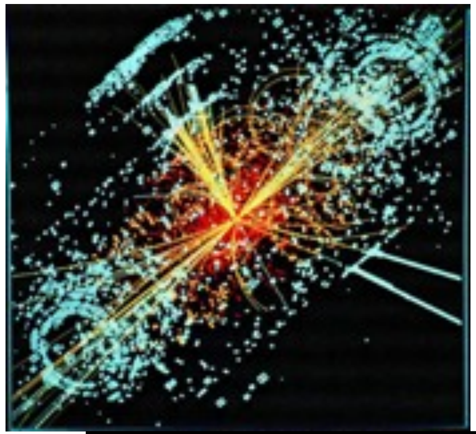
Approach 2: the discriminator is built upon hard-scattering matrix elements and Monte Carlo events

= subject of this lecture

Approach 1

Lagrangian
$L(p_1,p_2,\ldots)$



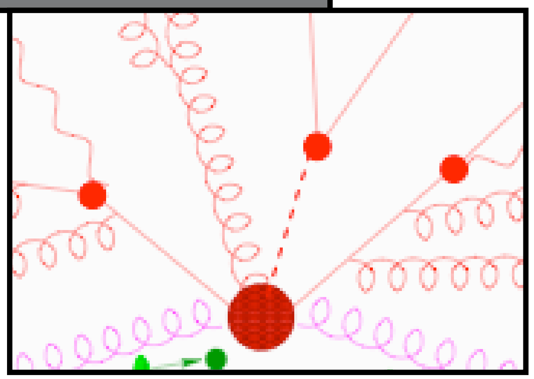


Experimental
events

From the previous lectures, we have learnt that one can simulate Monte Carlo events for any model that can be defined in the form of a Lagrangian

Approach I

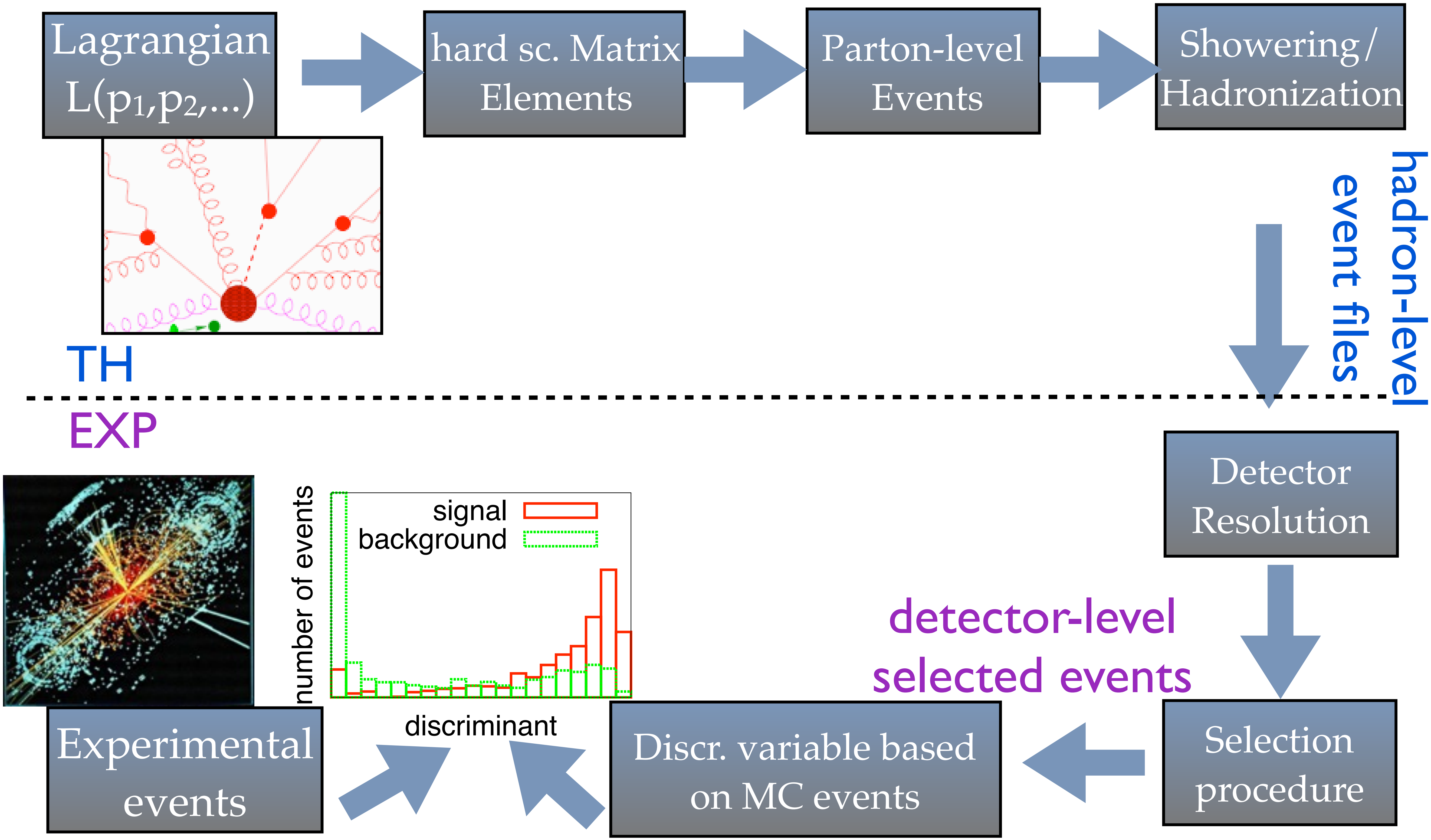
Lagrangian $L(p_1,p_2,\ldots)$ → hard sc. Matrix Elements → Parton-level Events → Showering/ Hadronization



TH output:

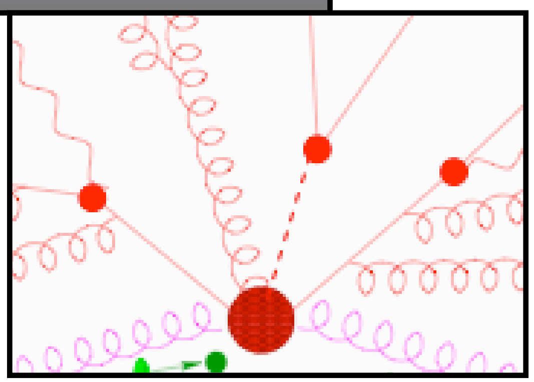
Experimental events

Approach I


Lagrangian
L(p1,p2,...)
hard sc. Matrix
Elements
Parton-level
Events
Showering/
Hadronization
hadron-level
event files
TH
EXP
Detector
Resolution
detector-level
selected events
Selection
procedure
Discr. variable based
on MC events
Experimental
events
signal
background
number of events
discriminant

Approach I

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



hadron-level event files

TH

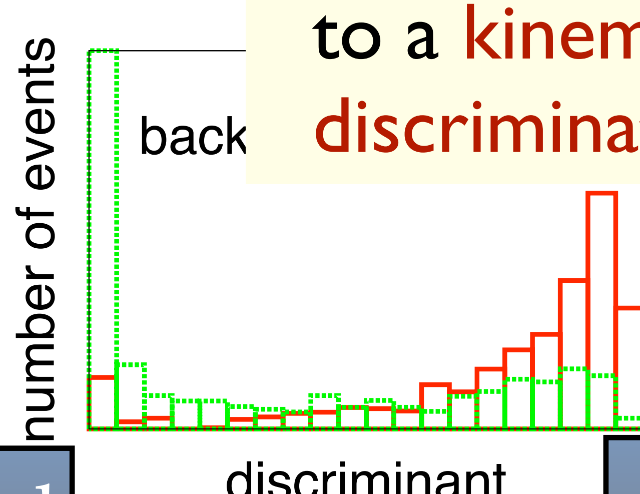
EXP

- 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 → Selection procedure

detector-level selected events

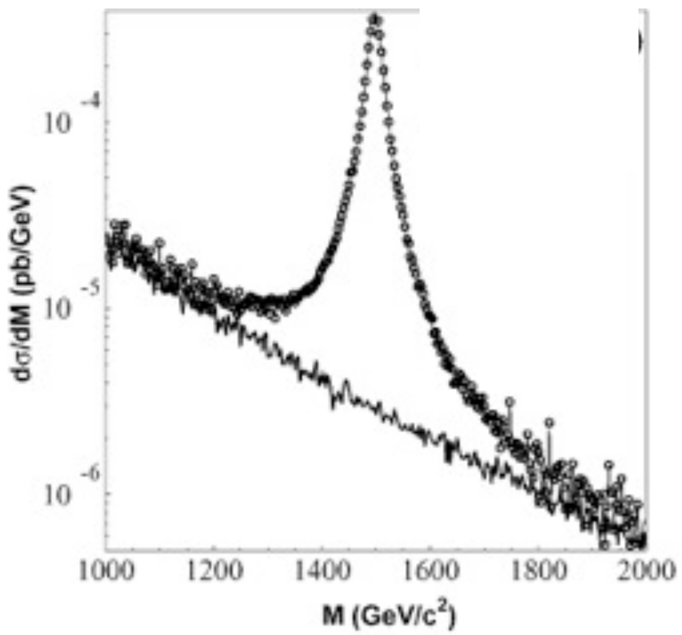
Selection procedure → Discr. variable based on MC events



Experimental events

Approach I

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



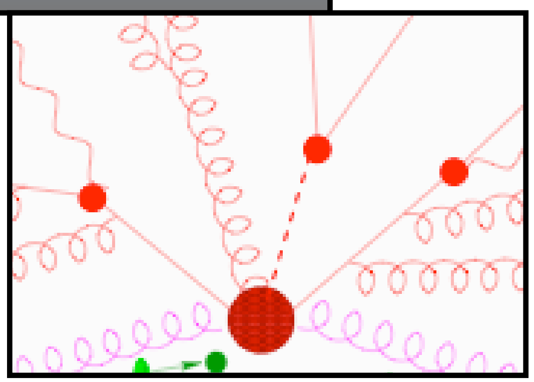
1. Reconstruct the distribution of events with respect to d=m(l^+,l^-) from MC events, under B-only and S+B hypotheses,

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

Approach I

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



hadron-level event files

TH

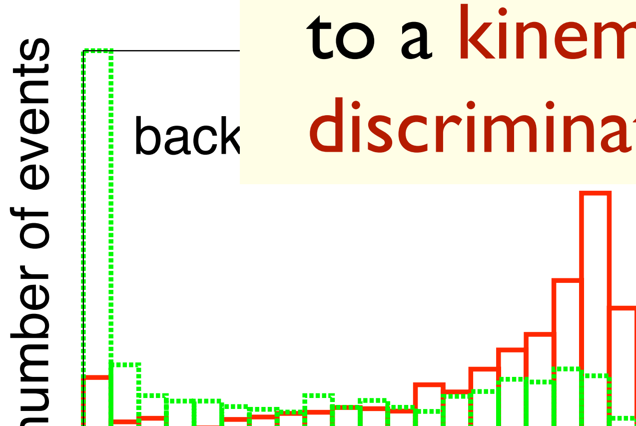
EXP

Detector Resolution

- 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-level selected events

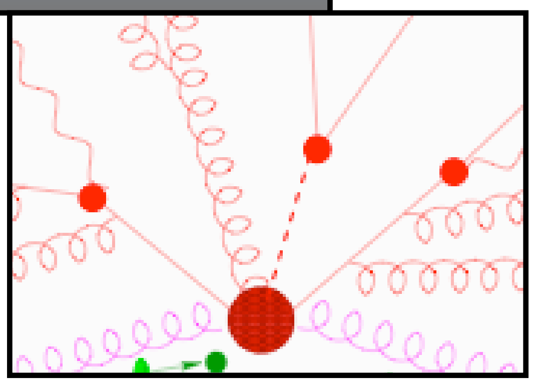
Selection procedure → Discr. variable based on MC events



Experimental events

Approach II

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



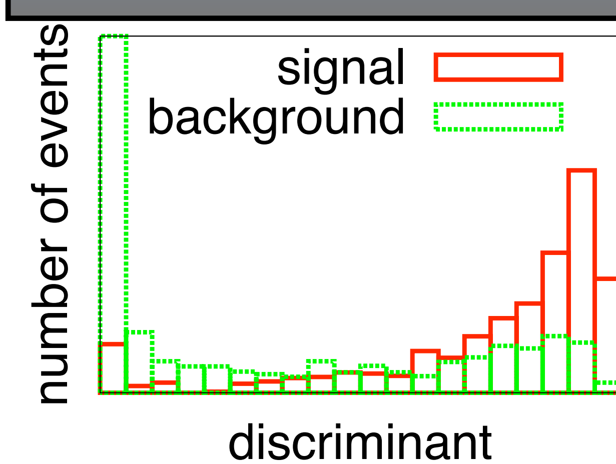
hadron-level event files

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

detector-level selected events

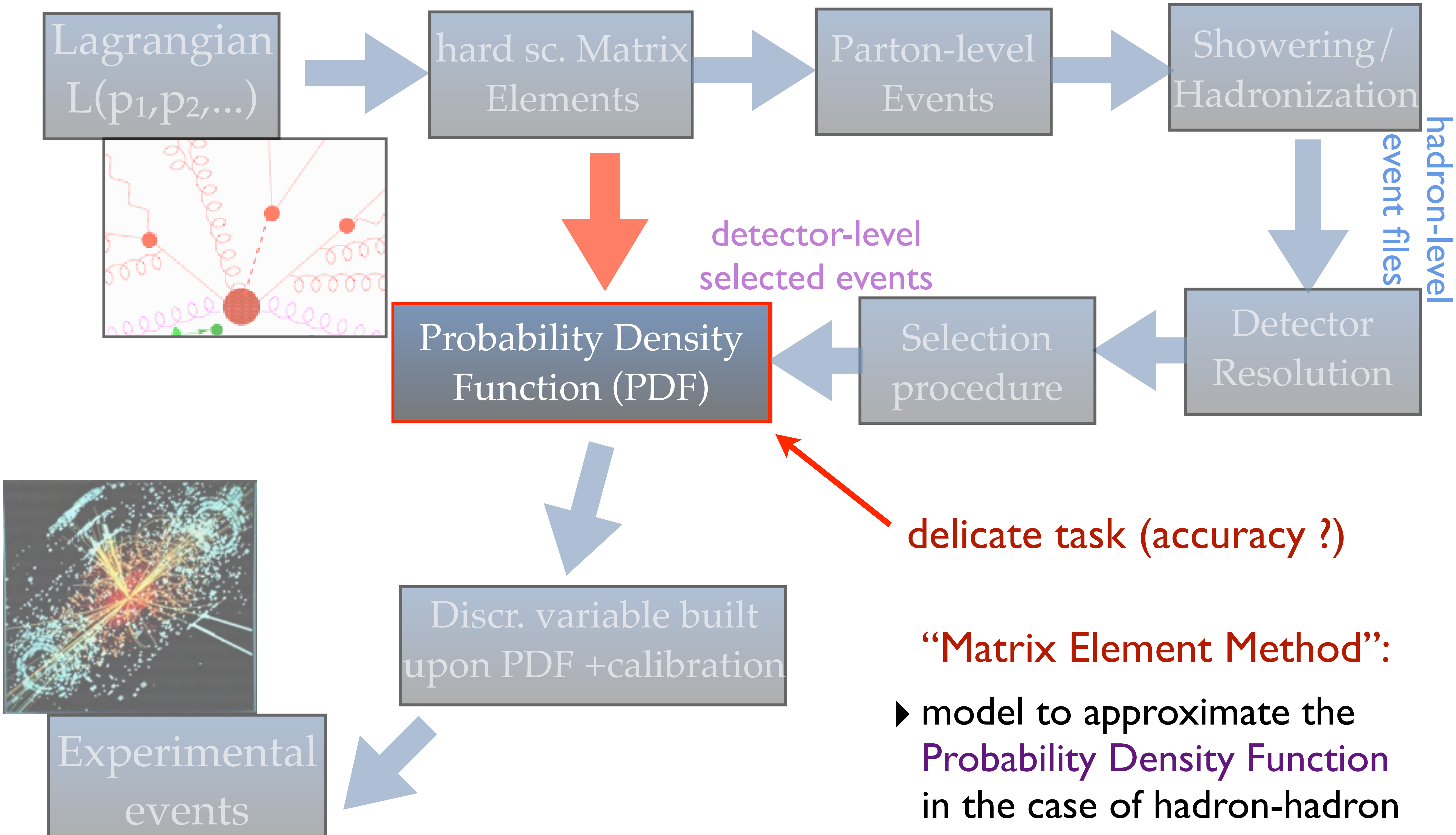
hard sc. Matrix Elements ↓ Probability Density Function (PDF) ↓ Discr. variable built upon PDF +calibration

Experimental events



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

Approach II

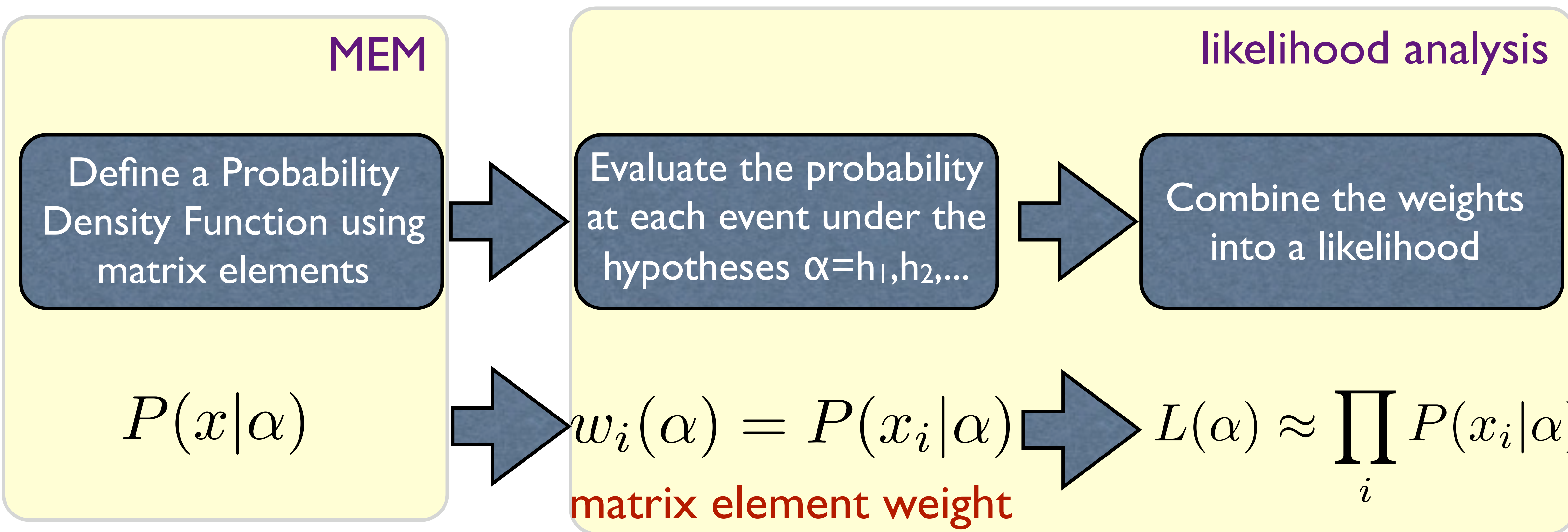


"Matrix Element Method":

- 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



x : kinematics of the reconstructed event

α : theoretical assumption

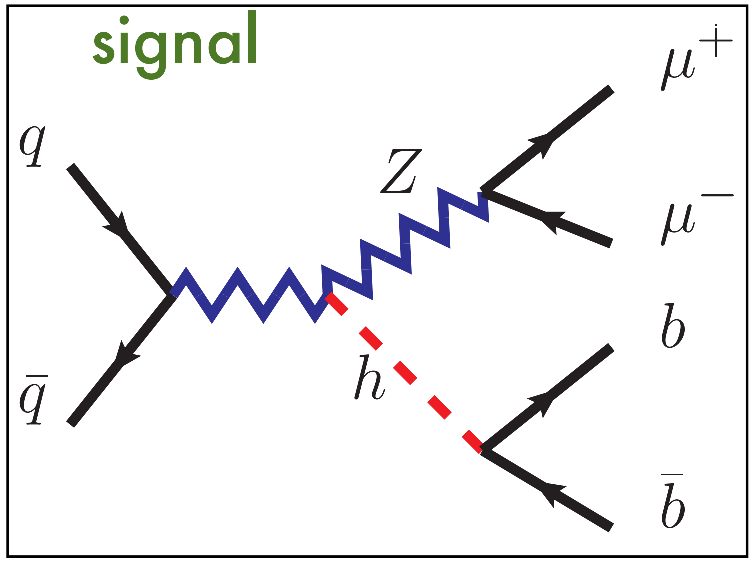
Reweighing events with matrix elements

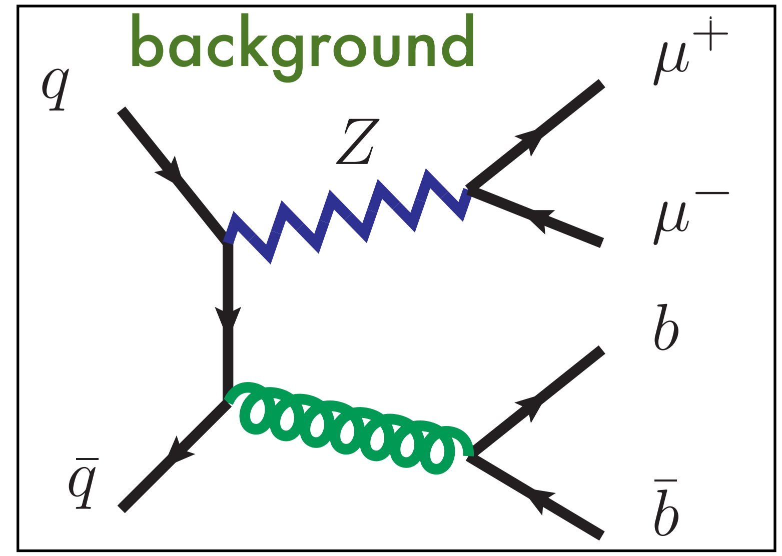
imagine we live in an ideal world, with an
ideal detector that reconstruct

- ✓ all the final state objects
- ✓ at the scale Q= scale of the hard interaction
- ✓ with an infinite resolution

Reweighing events with matrix elements

under these conditions, consider the following Higgs 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

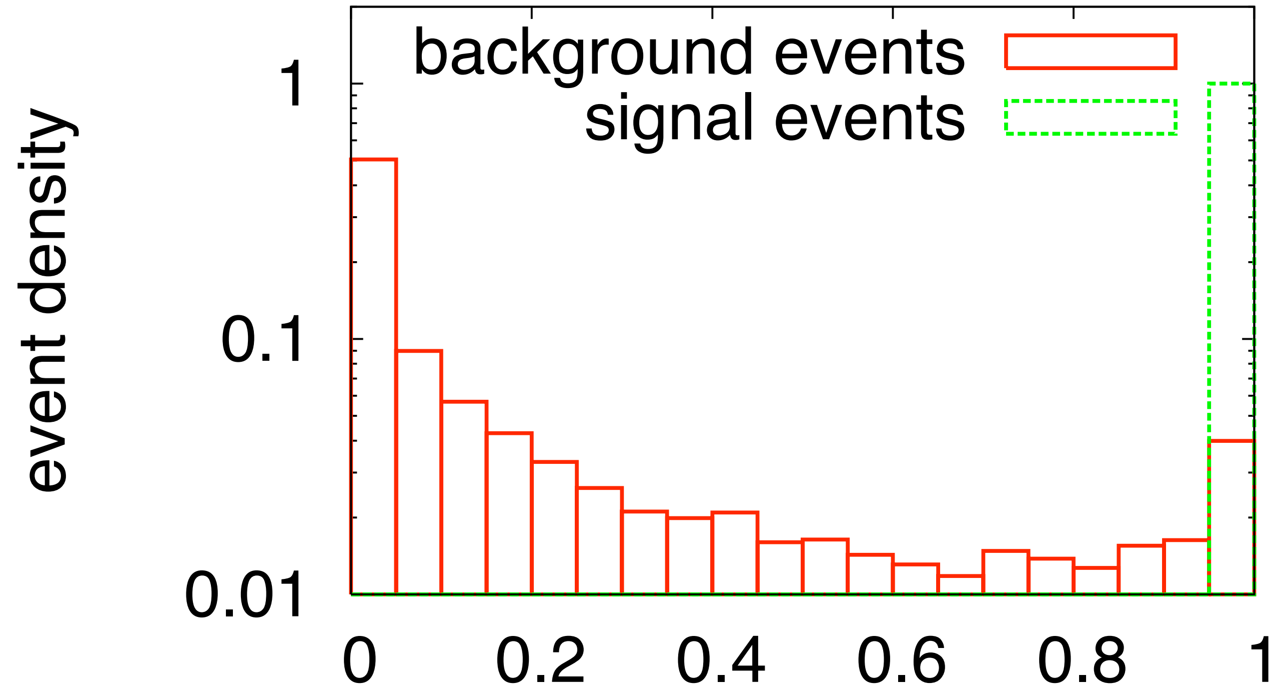
$$P(x|S) = \frac{\phi(x)}{\sigma_S}|M_S(x)|^2 \qquad P(x|B) = \frac{\phi(x)}{\sigma_B}|M_B(x)|^2$$

M_S : matrix element under the signal hypothesis

M_B : matrix element under the background hypothesis

Reweighing events with matrix element

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:

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

see Jorgen's talk

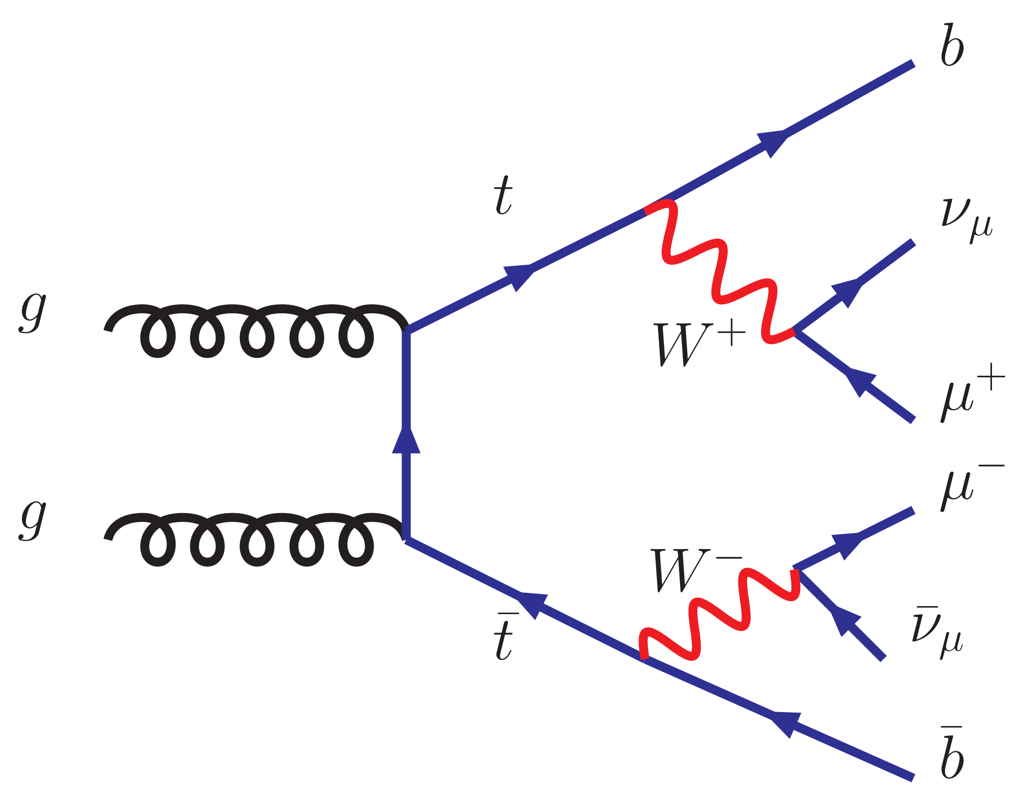
Real experiment

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

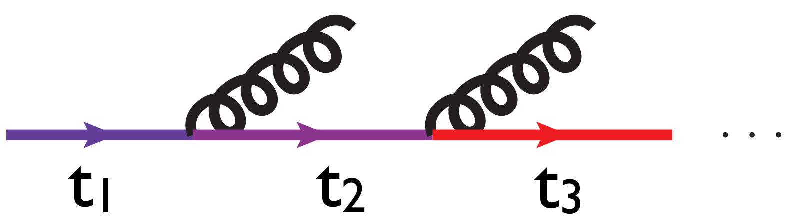
in a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

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 ⟶ ~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

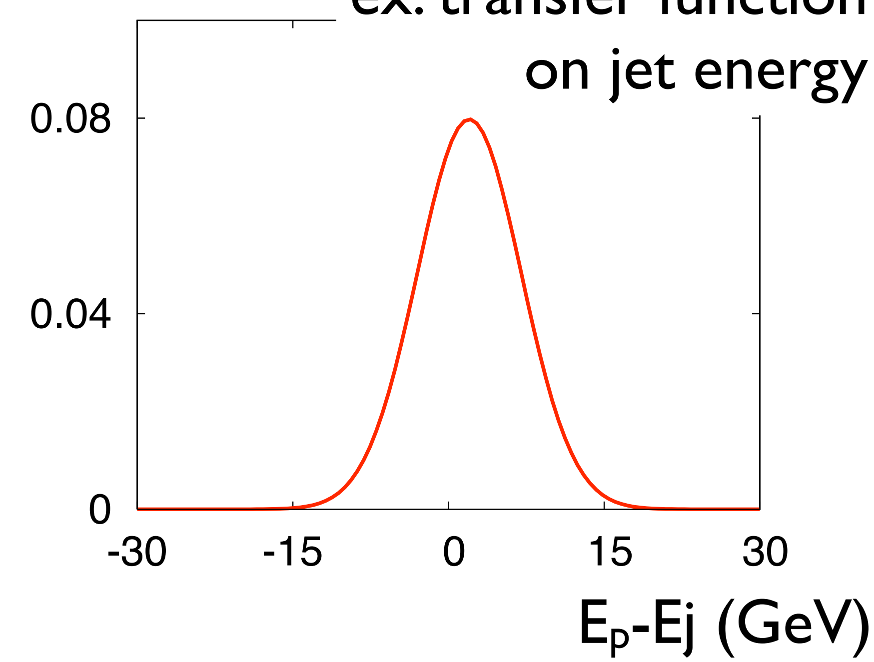
in a real experiment, a reconstructed event cannot be weighted by a unique matrix element:

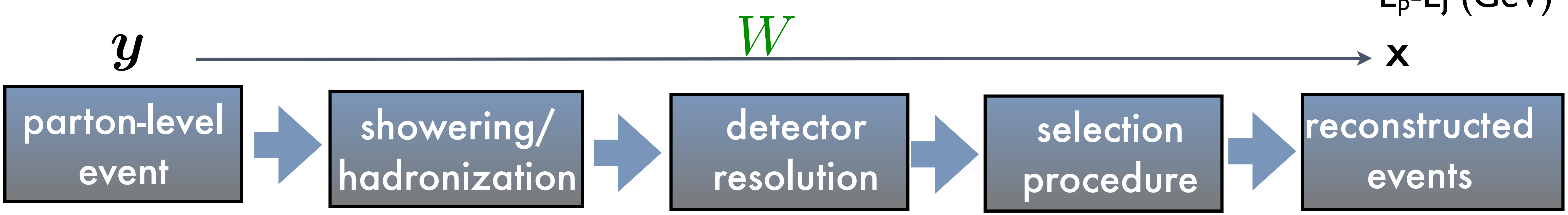
1. missing energy ⟷ $P(x,\alpha)$ must be summed over the unobserved degrees of freedom

2. showering/hadronization effets
3. experimental resolution/reconstruction algorithm

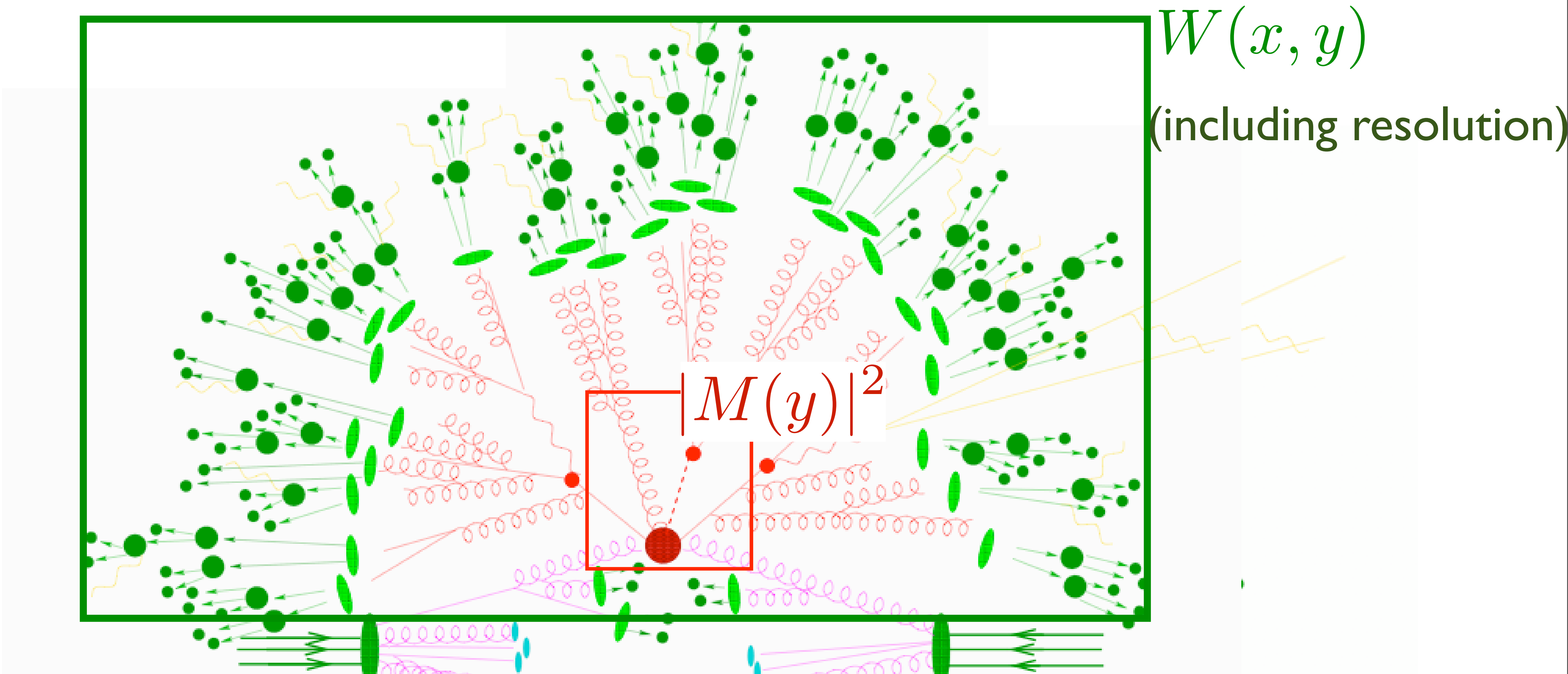
⟷ convolute with a transfer function $W(x,y)$
= probability that x is reconstructed given that y has been produced

ex: transfer function on jet energy





"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: we need to marginalize over unconstrained information and to convolute with the resolution function W for the measured quantities

$$g(x'|\vec{\alpha},\vec{\theta}) = \int R(x,x'|\vec{\alpha}) f_X(x|\vec{\theta}) dx$$

see Jorgen's talk

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

integration on the parton-level phase-space

tree-level matrix element

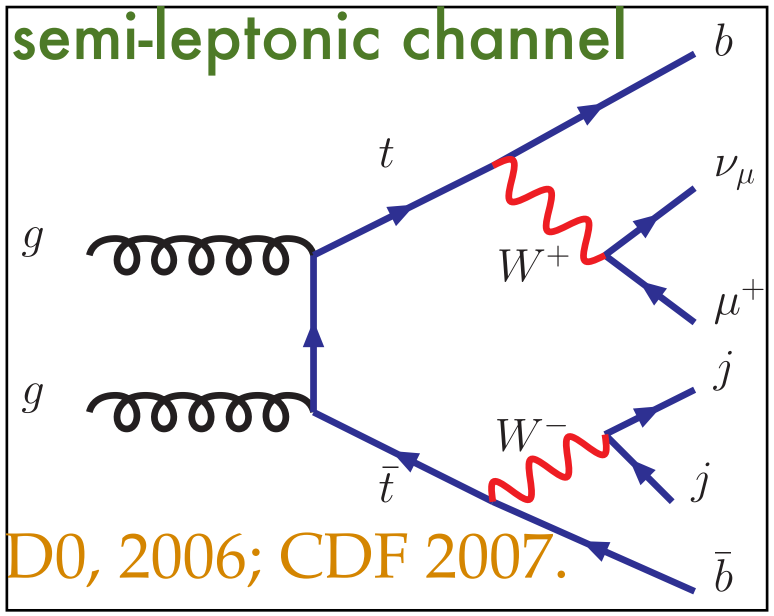
transfer function extracted from MC simulation

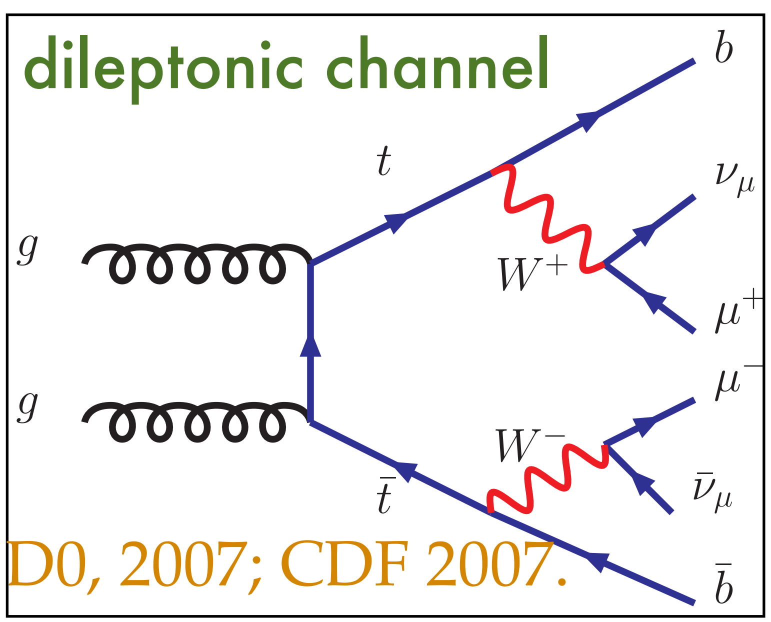
normalization: $\int dx W(x,y) Acc(x) = \epsilon(y)$

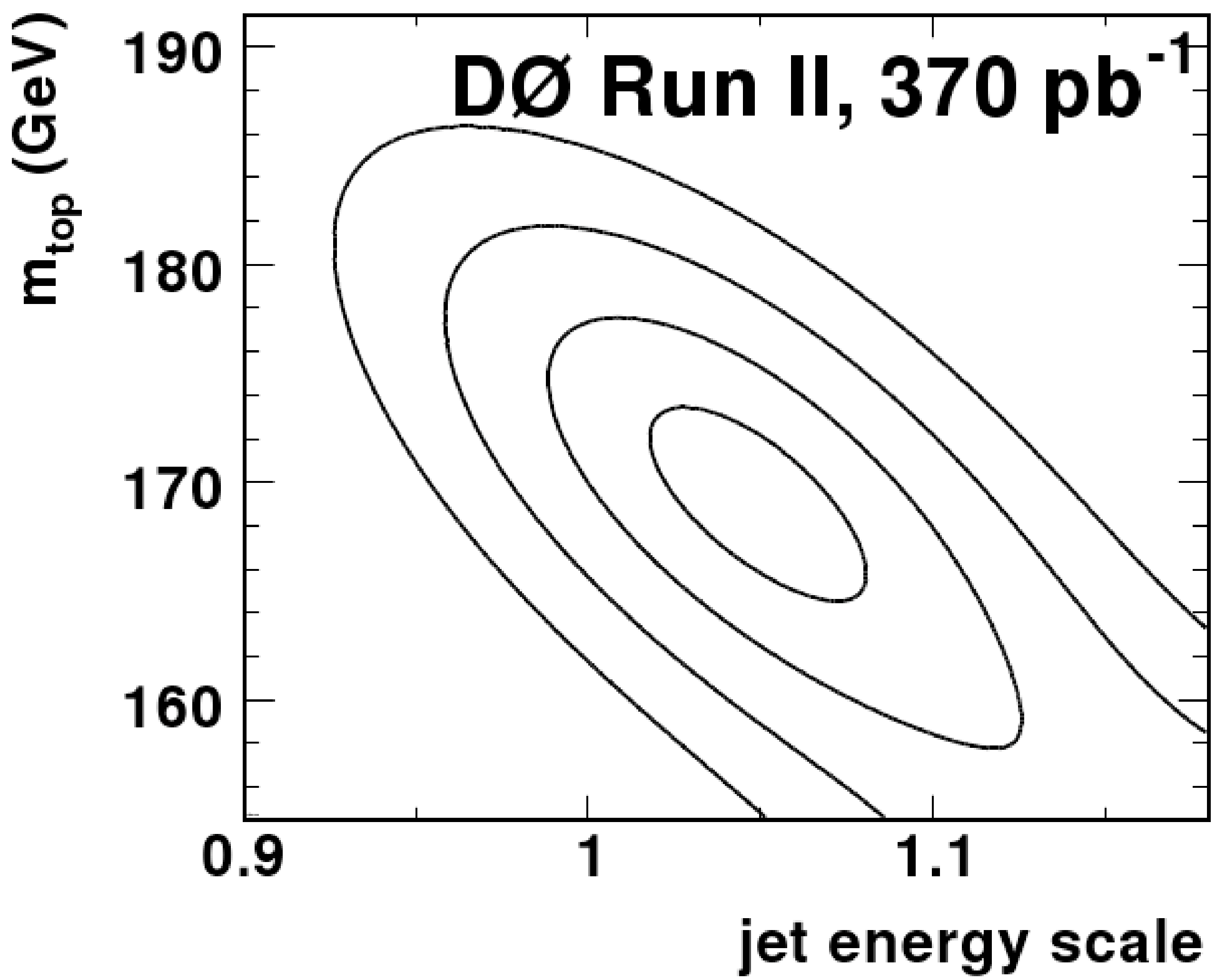
⇨ the probability density P(x| α) is normalized to 1

First MEM analyses at the Tevatron

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



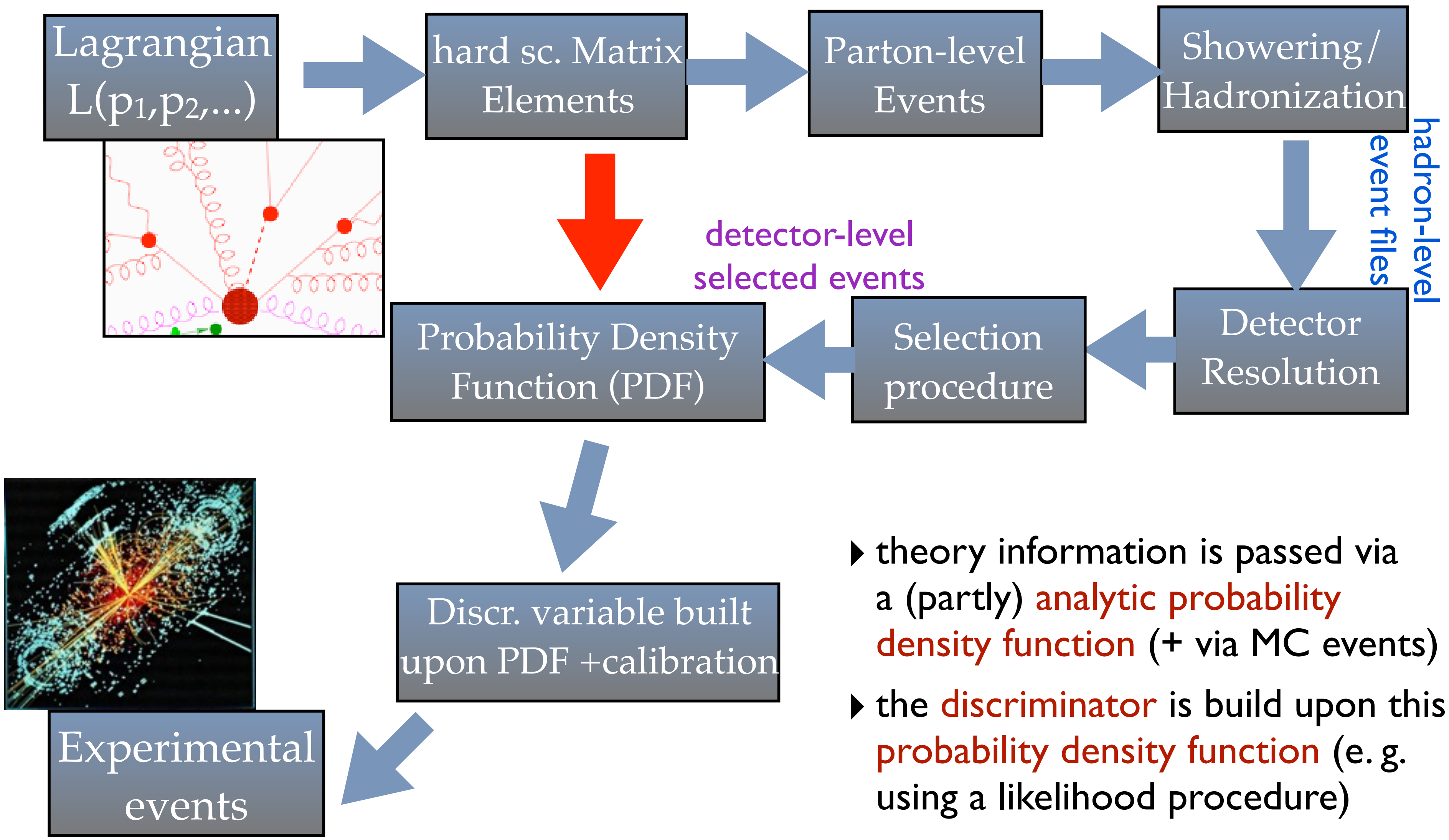




[DO Phys. Rev. D75 092005, 2006]

Significant improvement for the measurement of the top-quark mass

Matrix element method



- 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)

How can we evaluate the probability density function P(x|a) in practice ?

Practical Evaluation of the PDF

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

integration on the parton-level phase-space

tree-level matrix element

transfer function

Practical Evaluation of the PDF

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

integration on the parton-level phase-space

tree-level matrix element

transfer function

✓ available in mg5 for a very large set of processes

Practical Evaluation of the PDF

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

integration on the parton-level phase-space

tree-level matrix element

transfer function

✓ available in mg5 for a very large set of processes

✓ can be extracted from Monte Carlo similations

Practical Evaluation of the PDF

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

integration on the parton-level phase-space

tree-level matrix element

transfer function

- Monte Carlo integration ?

✓ available in mg5 for a very large set of processes

✓ can be extracted from Monte Carlo similations

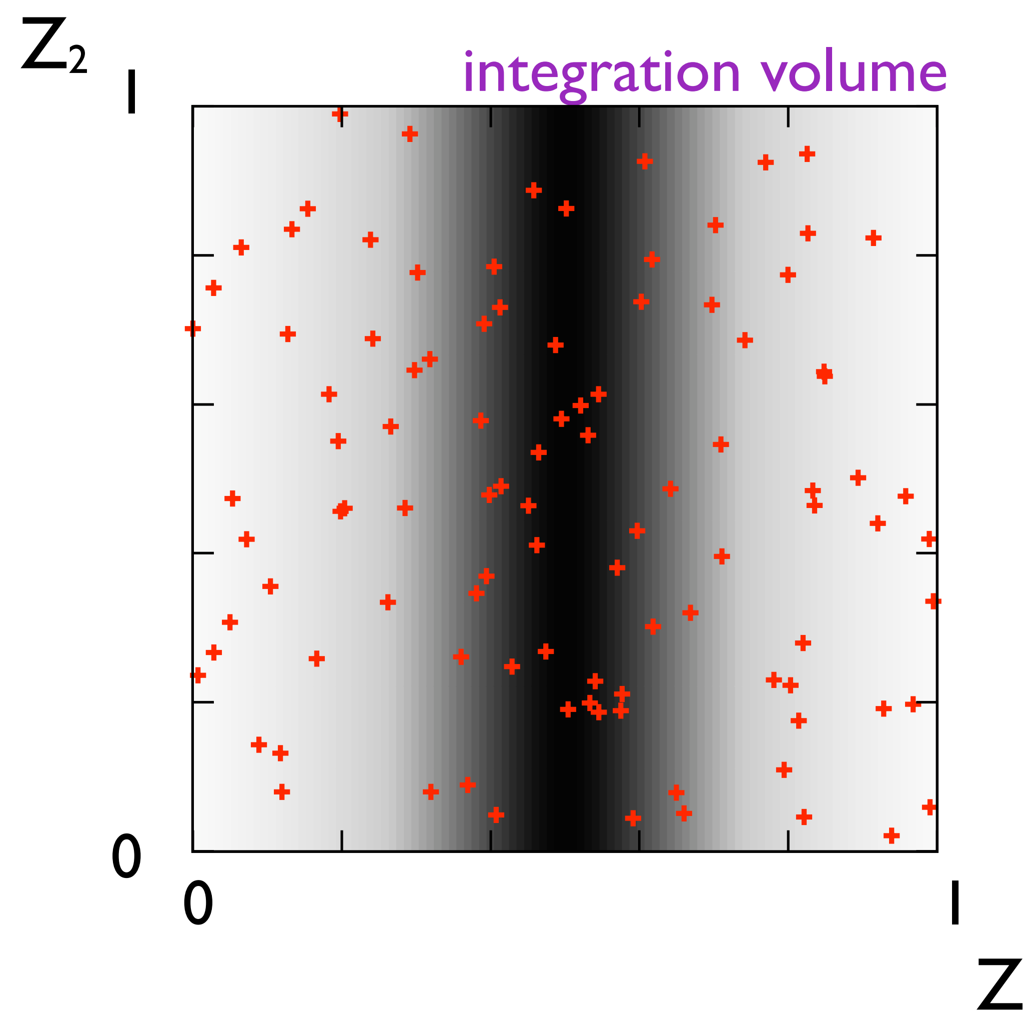
Monte Carlo integration

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: $\sigma_I \approx \frac{S}{\sqrt{N}}$

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

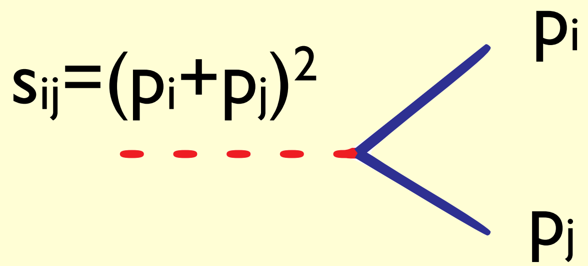
if S large ⇒ poor convergence



Practical Evaluation of the PDF

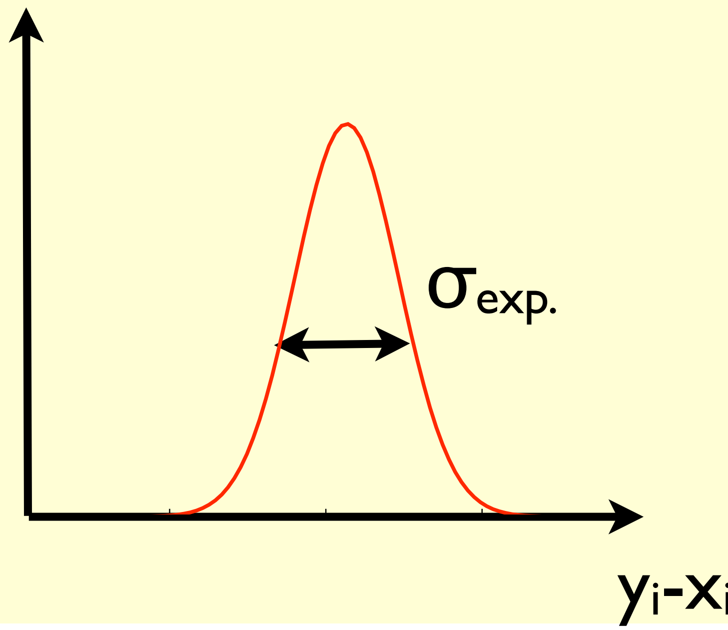
$$P(x,\alpha) \propto \int d\phi_{\boldsymbol{y}} \, |M|^2(\boldsymbol{y}) \, W(\boldsymbol{x},\boldsymbol{y})$$

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 an algorithm that is sufficiently fast (large number of weights must be evaluated)

MADWEIGHT

P. Artoisenet, V. Lemaitre, F. Maltoni, O. Mattelaer

- consider the definition of the PDF in the Matrix Element Method

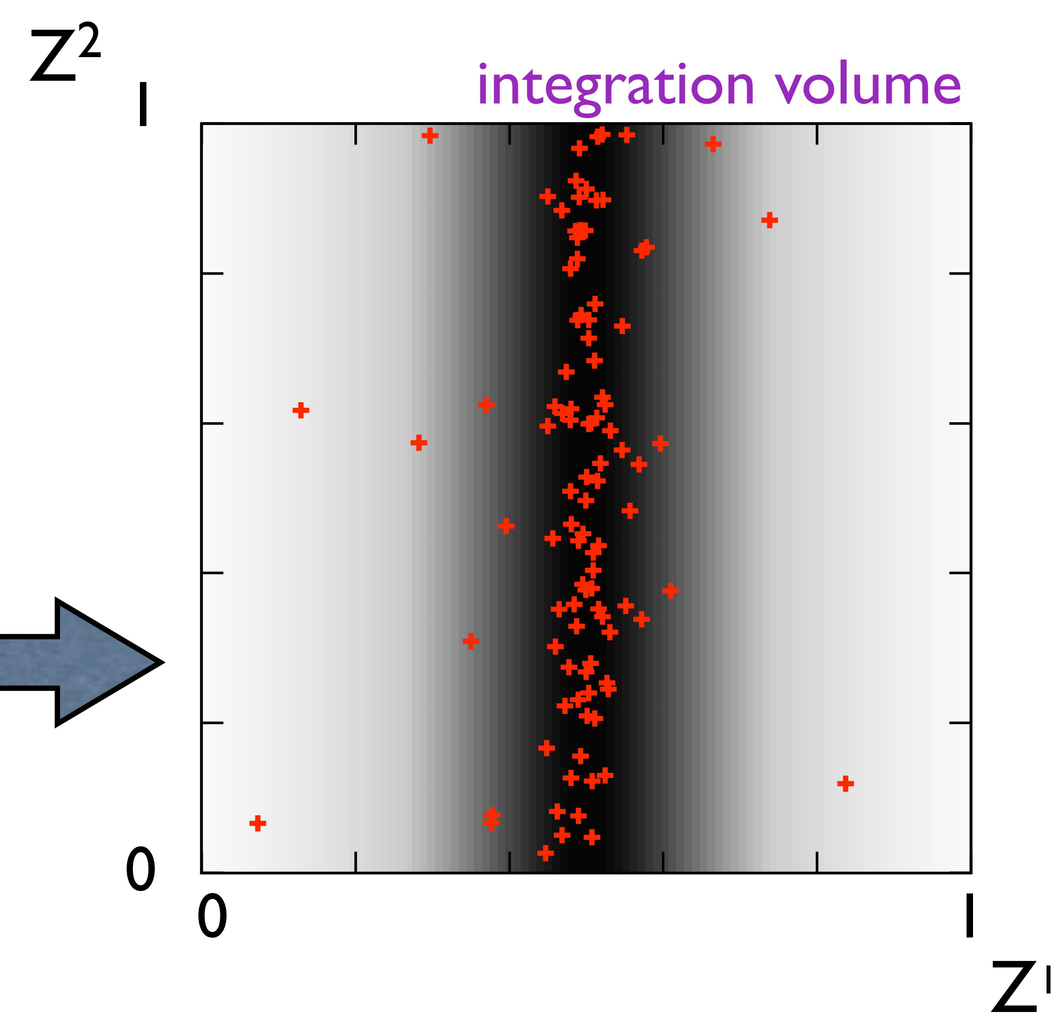
- solve the problem of evaluating the PDF at a specific event in a generic way by using adaptive and multichannel Monte Carlo techniques

Monte Carlo integration

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

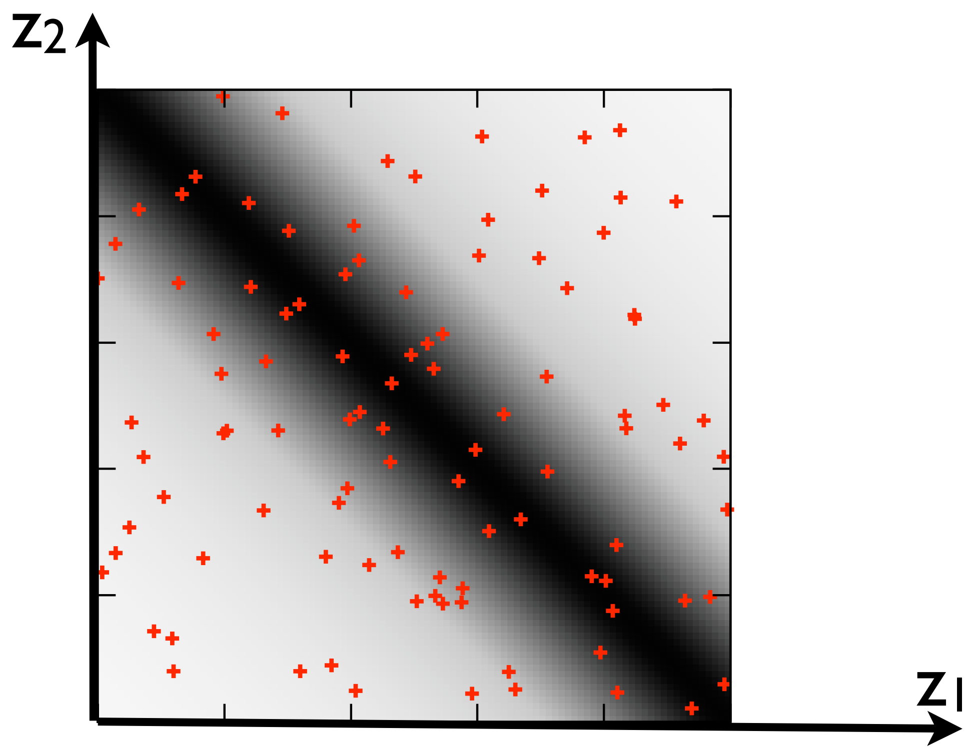
The grid has a factorized dependence in the integration variables

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



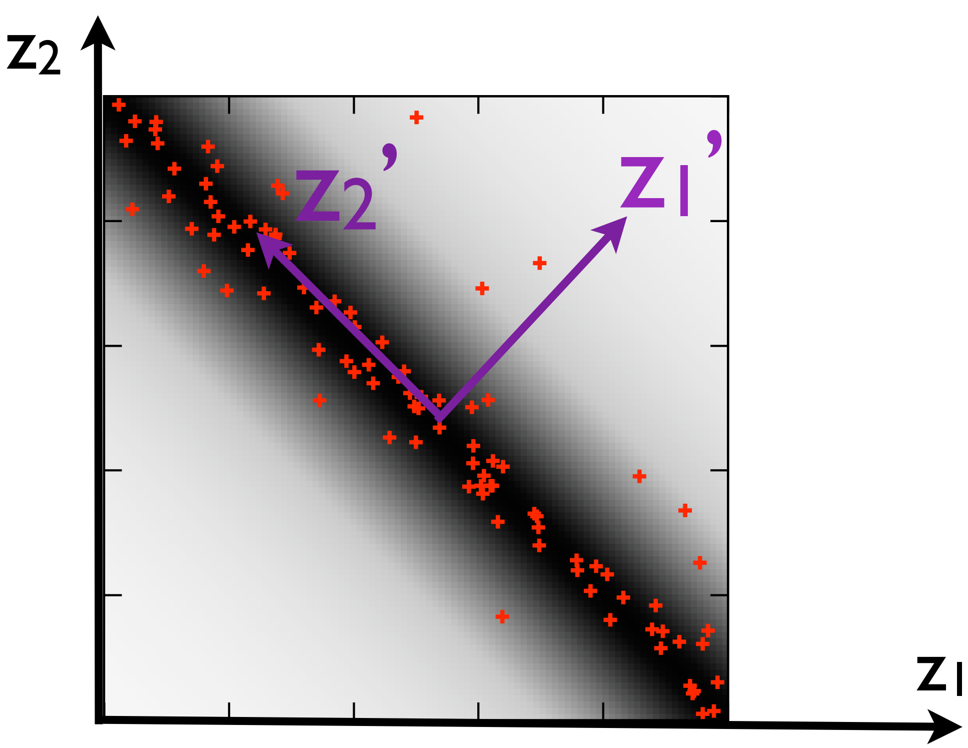
Adaptive Monte Carlo integration

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



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

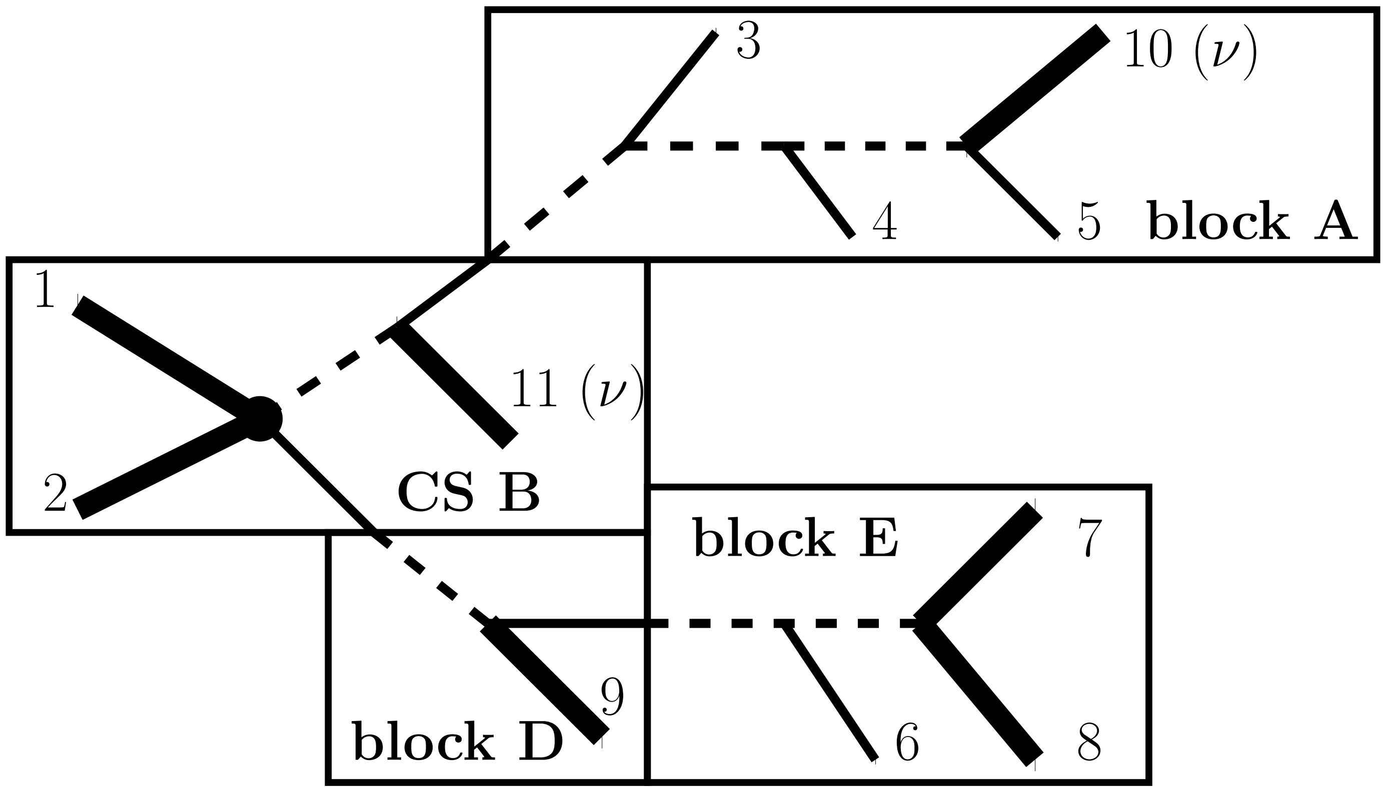


variables z_1', z_2':

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

MadWeight

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



- The phase-space measure is decomposed into "blocks"
- The phase-space measure associated with each block is optimized to map the ME + TF enhancements
- momenta are generated backward (from the end of the decay chain to the interaction point)

- 12 blocks are defined in MadWeight ➪ infinite set of phase-space mappings
- the optimal phase-space mappings are generated automatically and combined in a multichannel approach

MadWeight

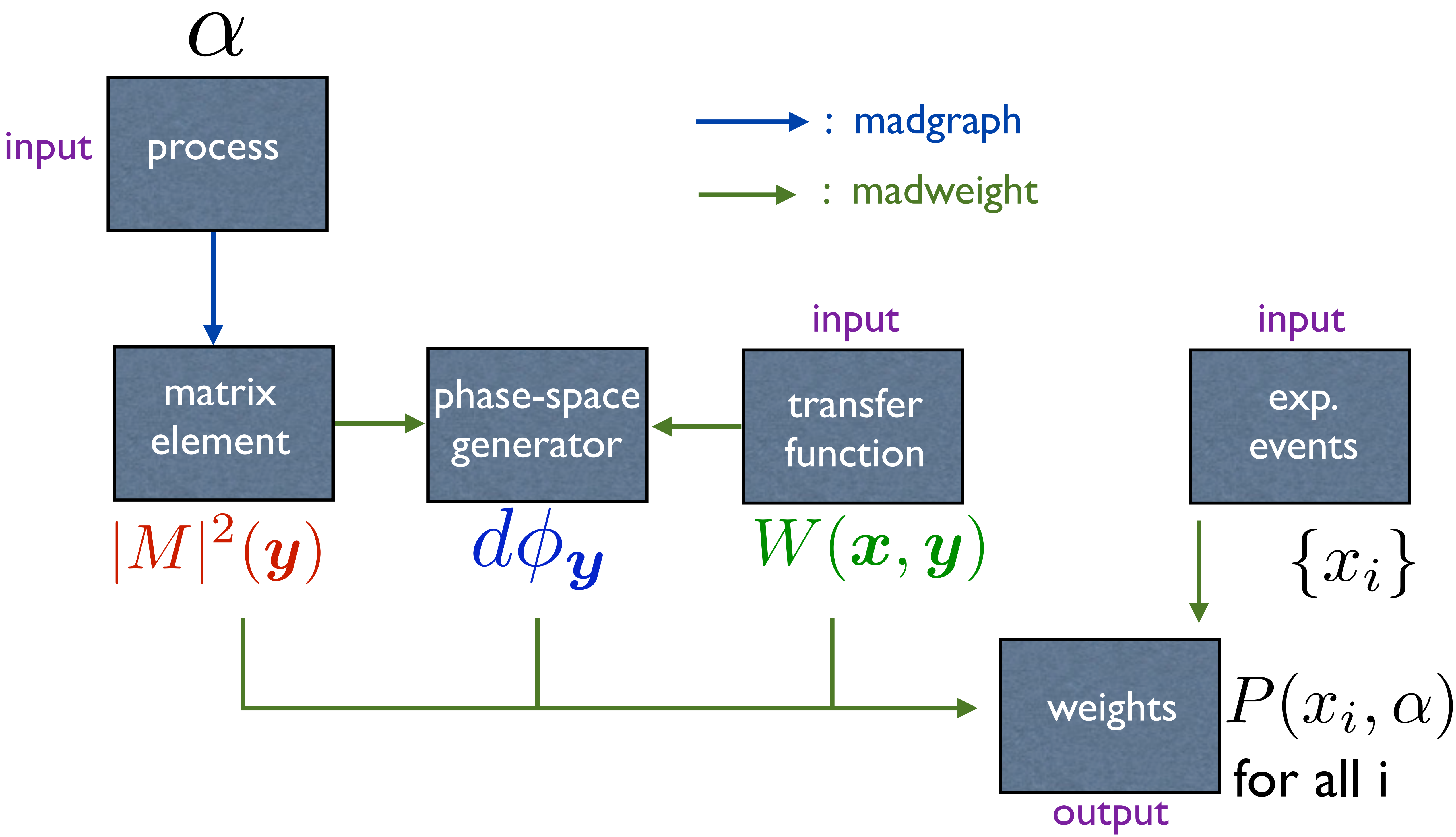
IMPROVEMENTS compared with previous codes:

- generic code for any decay chain and any transfer function (in principle)
- EXACT phase-space measure $d\phi_y$: reproduction of the phase-space volume for a large class of PS parametrizations
- multichannel techniques for overconstrained systems

l	blocks	integrated volume
3	MB A	6.30×10^{-5}
3	MB B	6.30×10^{-5}
3	MB C	6.30×10^{-5}
6	MB D	694 GeV^6
4	MB E	0.0166 GeV^2
4	MB F	0.0166 GeV^2
5	MB B + SB A	3.89 GeV^4
4	MB B + SB B	0.0166 GeV^2
3	MB B + SB C	6.30×10^{-5}
3	MB B + SB D	6.30×10^{-5}
4	MB B + SB E	0.0166 GeV^2

Implementation in madgraph5

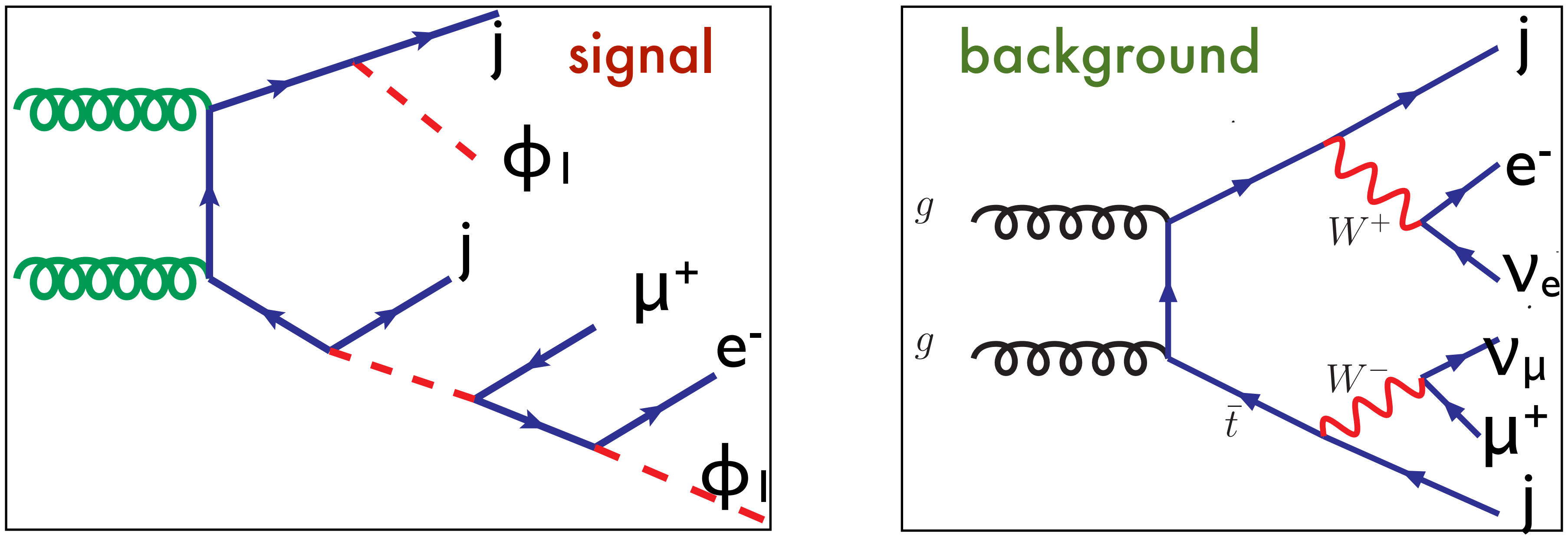
$$P(x,\alpha) \propto \int d\phi_{\boldsymbol{y}}\ |M|^2(\boldsymbol{y})\ W(\boldsymbol{x},\boldsymbol{y})$$



MadWeight: application

let us consider one process presented in the tutorial

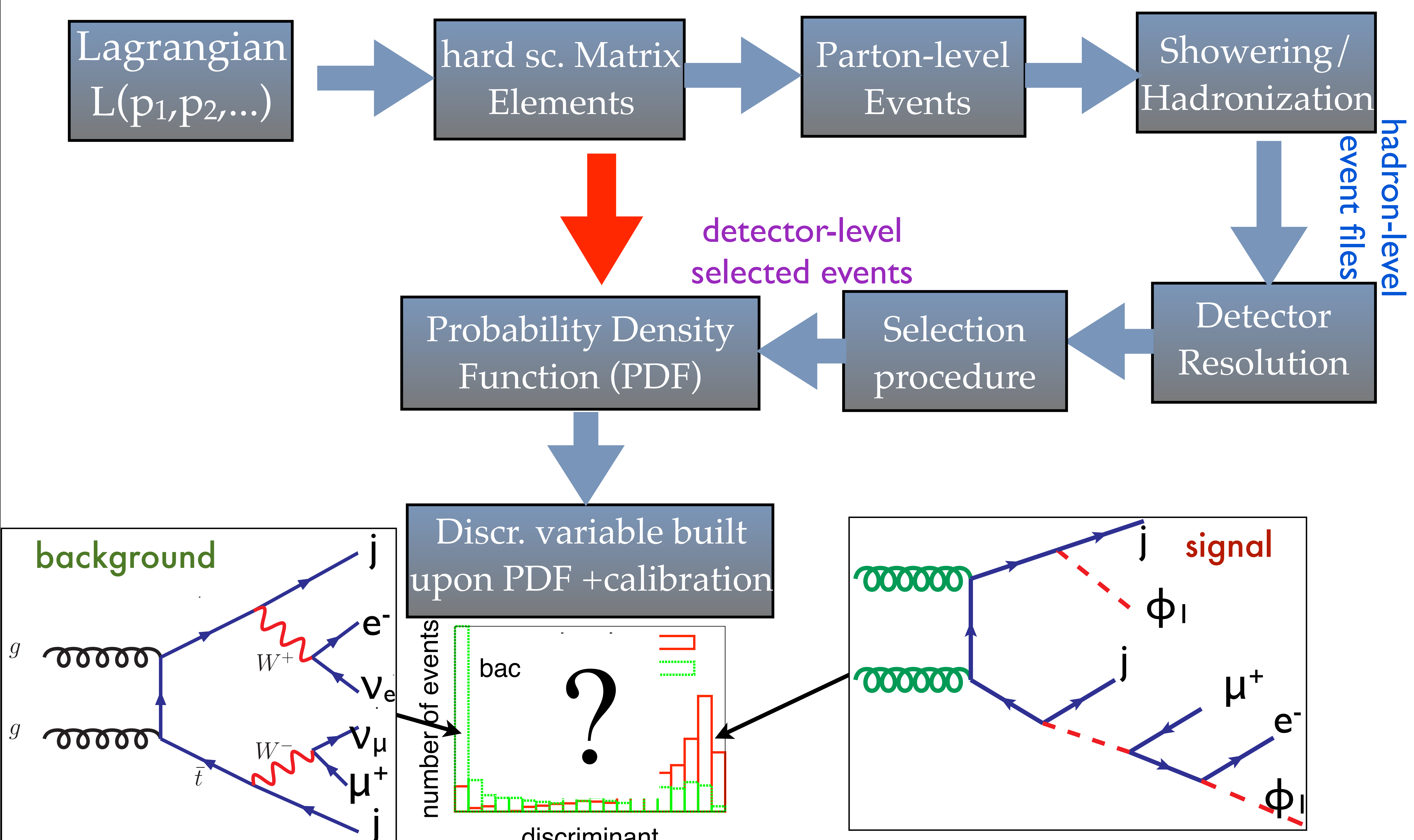
III. $pp \to (U \to j\Phi_1)(\bar{U} \to j\,\ell^+\ell'^-\,\Phi_1)$+h.c , i.e., $pp \to \ell^+\ell^- + 2$ jets + missing E_T.



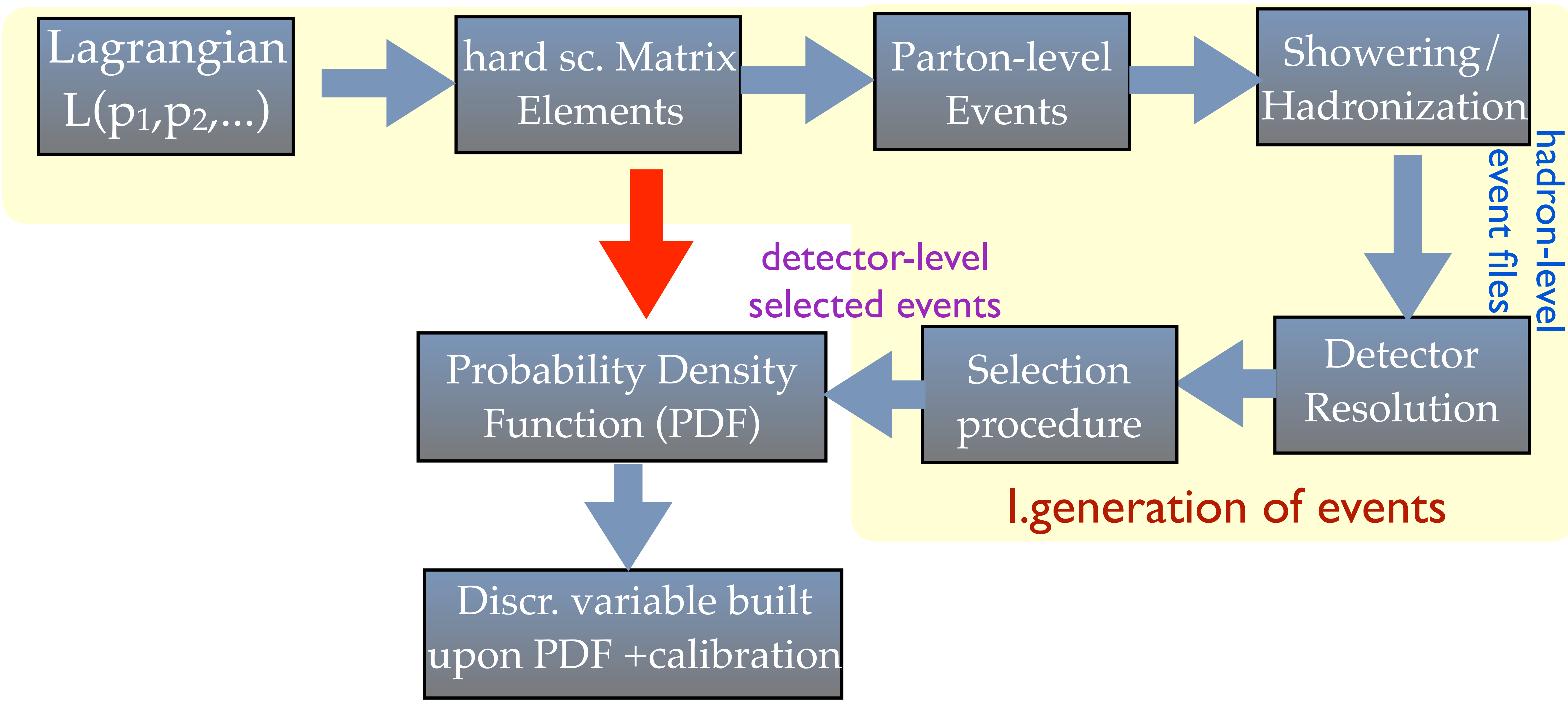
and use the matrix element method under realistic conditions

(showering, hadronization, detector effects, ...)

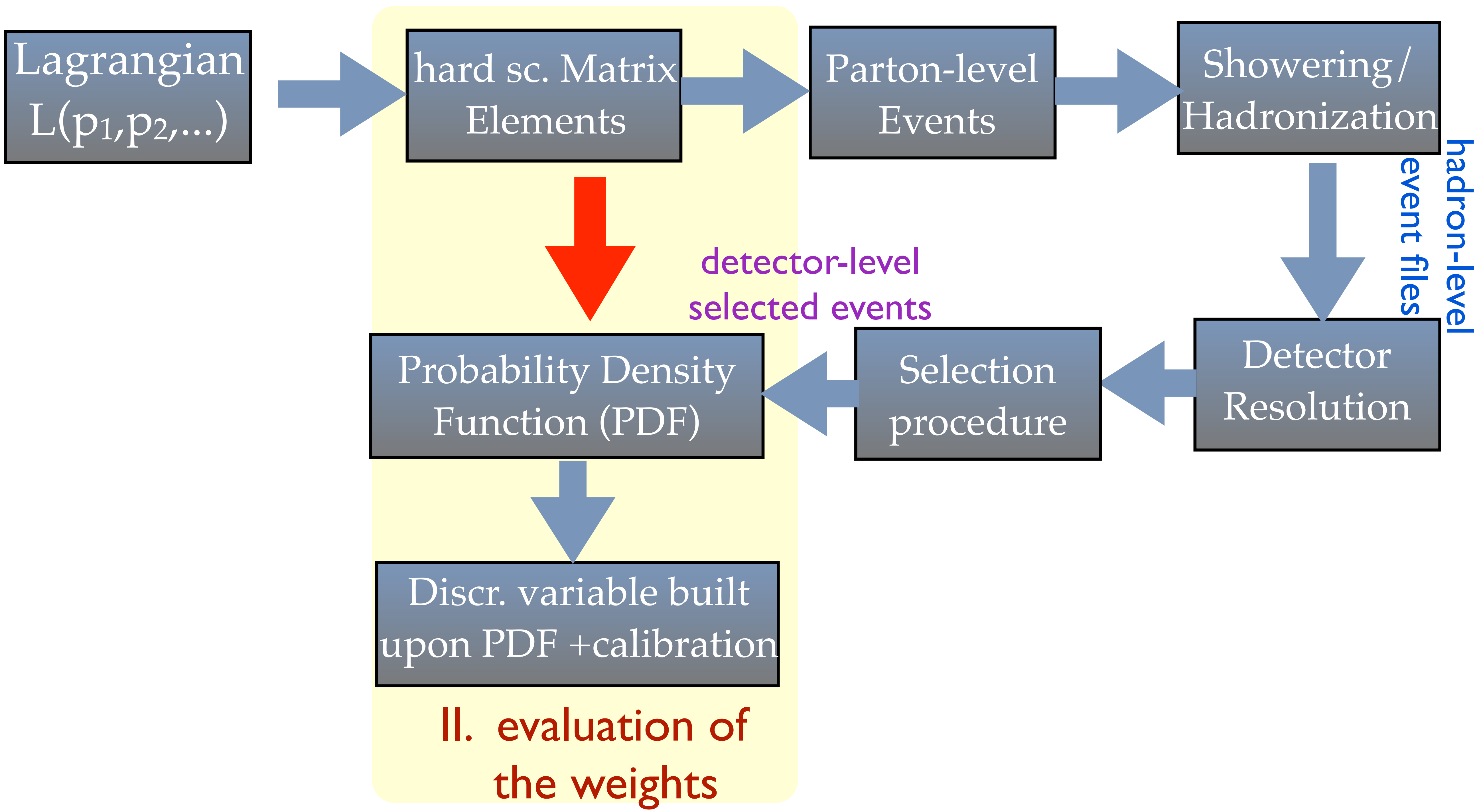
MadWeight: application



MadWeight: application



MadWeight: application

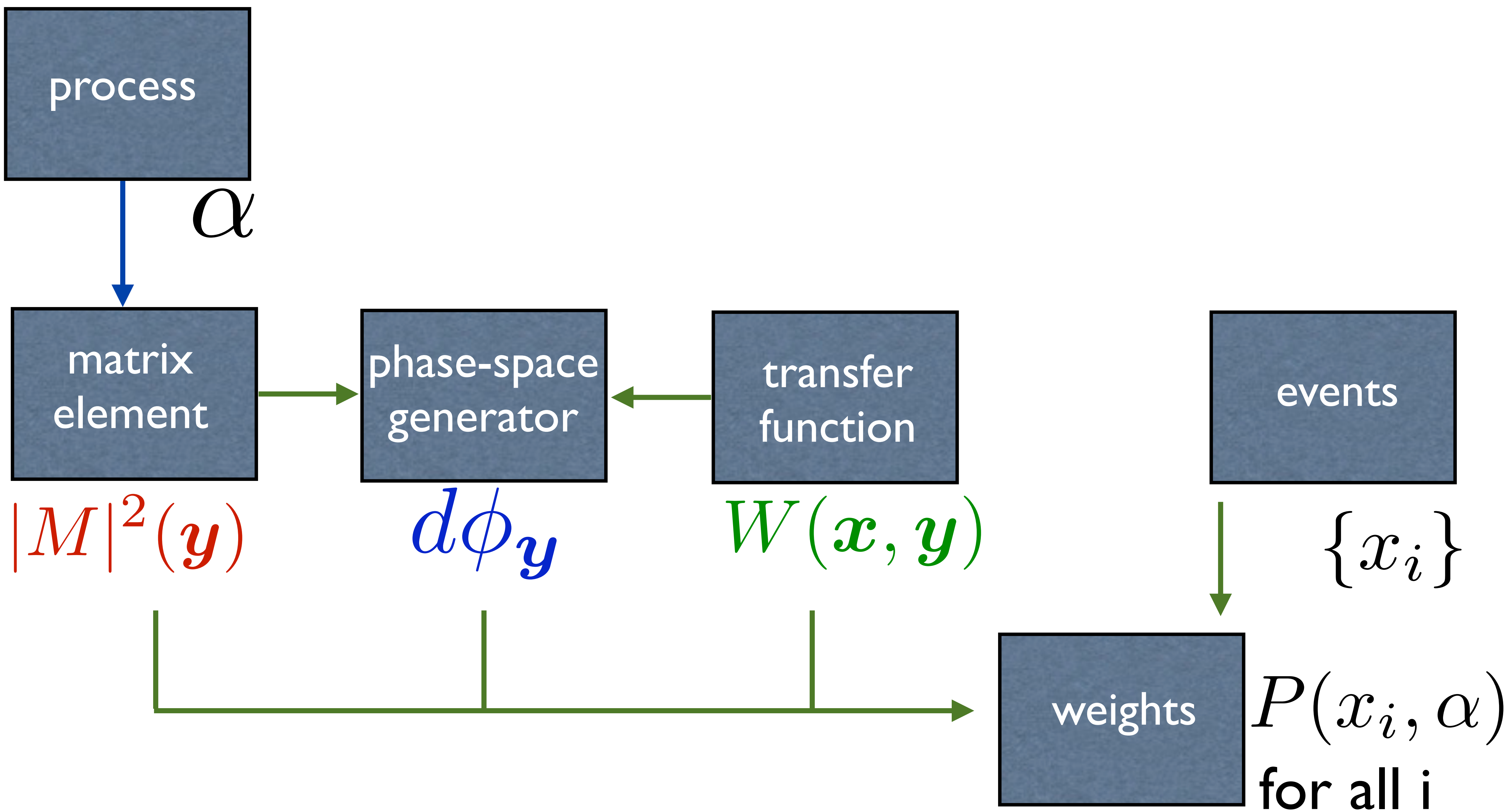


II. evaluation of the weights

using scheme that is fast, reliable, reproducible

➪ load madweight implementation in madgraph 5:

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

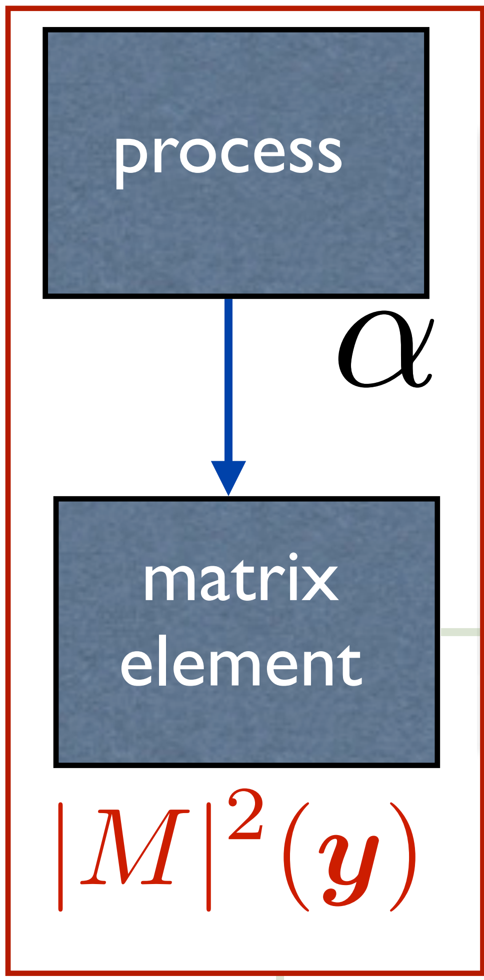


II. evaluation of the weights

using scheme that is fast, reliable, reproducible

➯ load madweight implementation in madgraph 5:

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

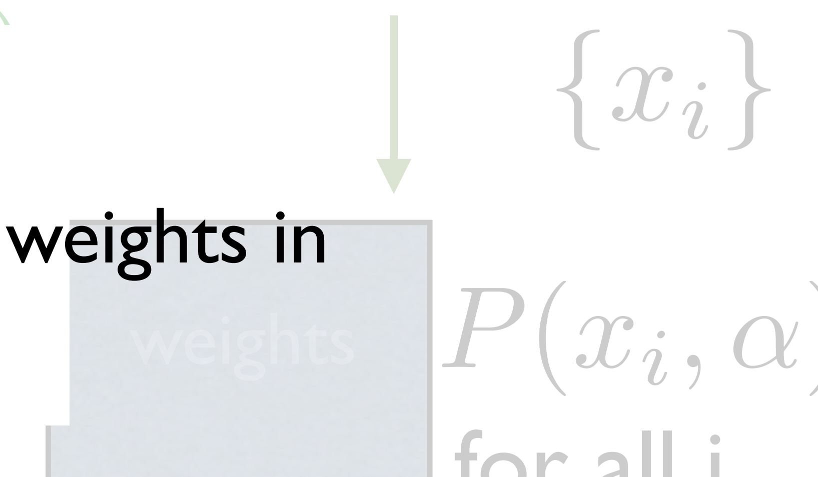


input : proc_card_mg5.dat

```
import model Natal_2012_UFO
generate p p > uv uv~ , uv > p1 u ,
( uv~ > p2 u~ , ( p2 > ev mu+ , ev > mu- p1 ))
output madweight signal_hypothesis
```

output

code for the evaluation of the weights in directory signal_hypothesis



II. evaluation of the weights

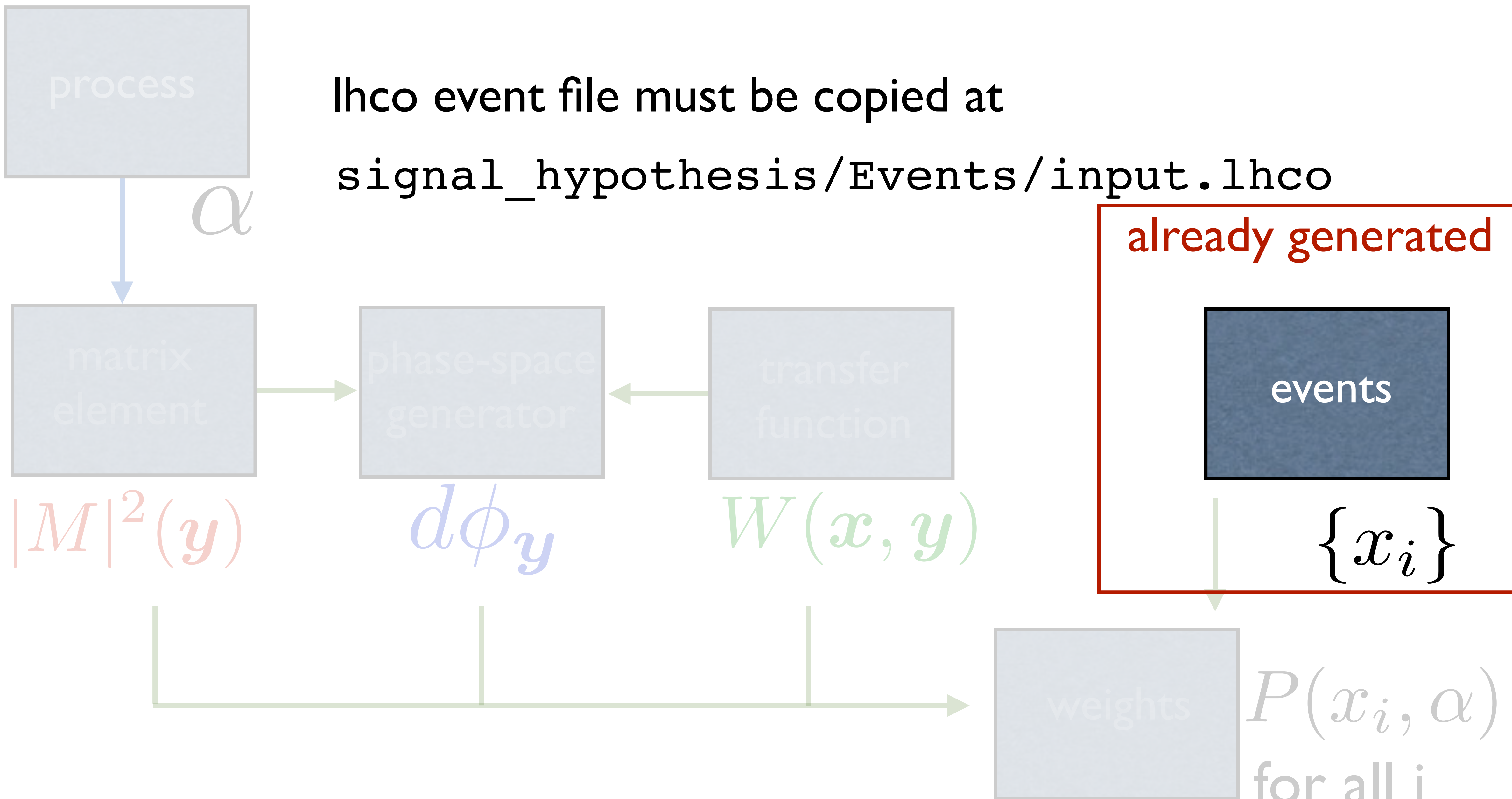
using scheme that is fast, reliable, reproducible

➪ load madweight implementation in madgraph 5:

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

lhco event file must be copied at

```
signal_hypothesis/Events/input.lhco
```



II. evaluation of the weights

using scheme that is fast, reliable, reproducible

➪ load madweight implementation in madgraph 5:

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

input 1: TF parametrization

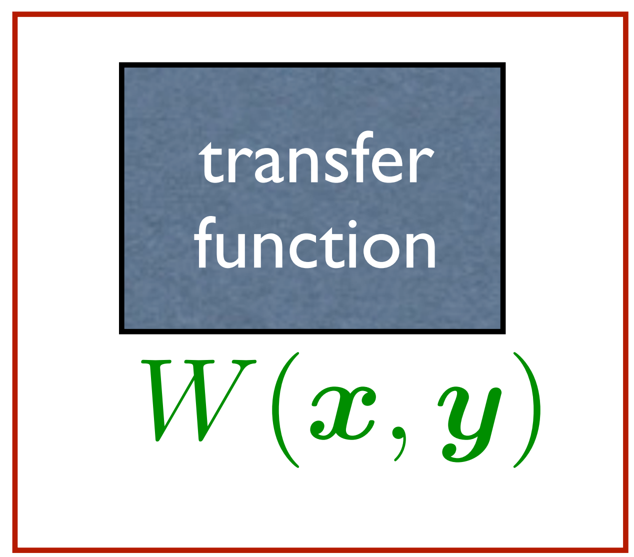
Source/MadWeight/transfer_function/TF_my_tf.dat

load TF:

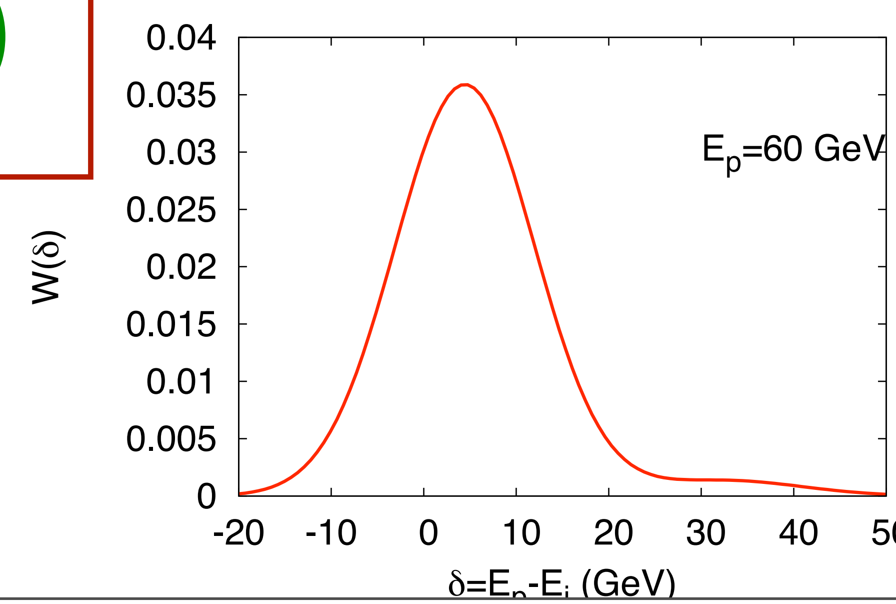
./bin/change_tf.py my_tf

input 2: TF parameters

transfer_card.dat



in this illustration:
double gaussian param.



II. evaluation of the weights

using scheme that is fast, reliable, reproducible

➪ load madweight implementation in madgraph 5:

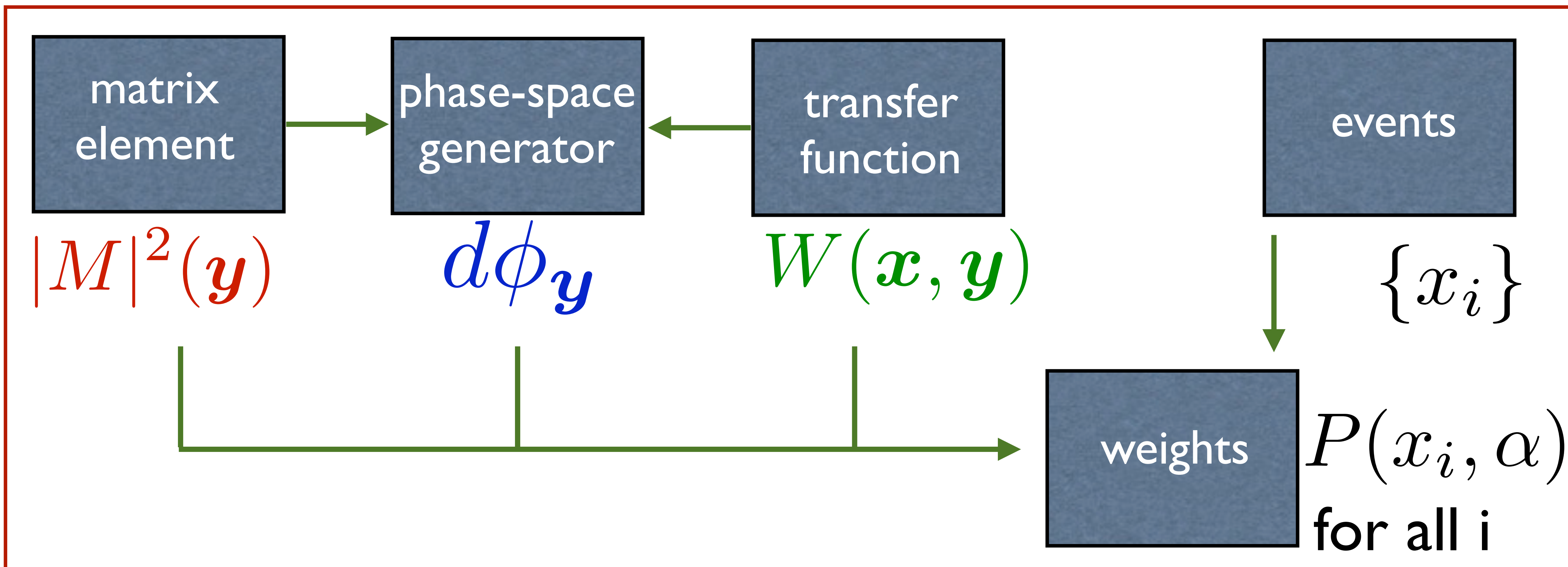
```
bzr branch lp:~maddevelopers/madgraph5/madweight
```

additional inputs:

MadWeight_card.dat, param_card.dat, run_card.dat

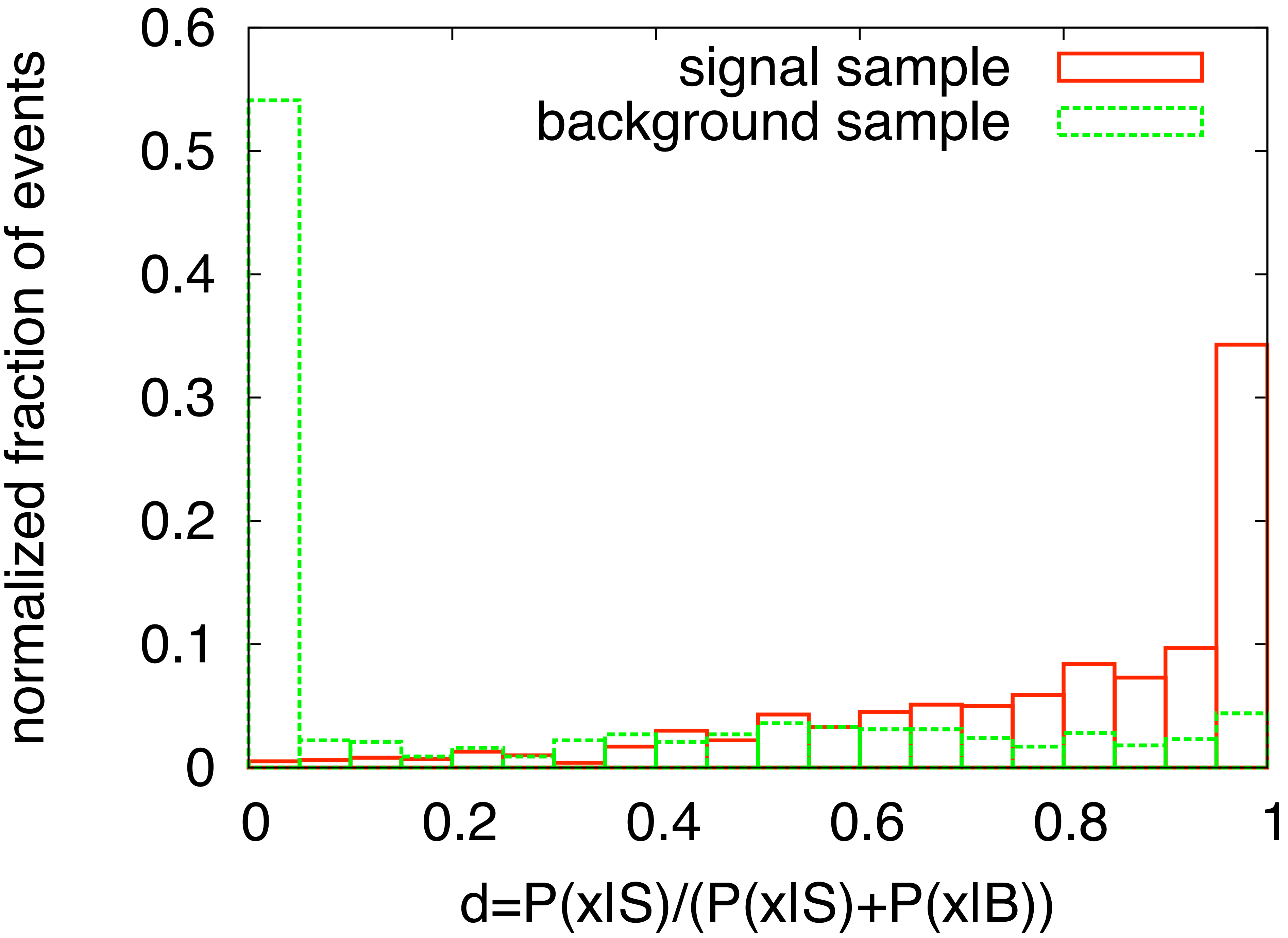
load phase-space generator+ evaluate the weights:

./bin/madweight 6-



Result

once all the weights have been evaluated for each event and each assumption, they can be combined to analyze the discriminating power:



Result

you can reproduce the whole analysis by loading the cards:

event generations

A. proc_card_mg5.dat
B. run_card.dat, param_card.dat
C. pythia_card.dat
D. delphes_card.dat, delphes_trigger.dat
E. selection / MadAnalysis script

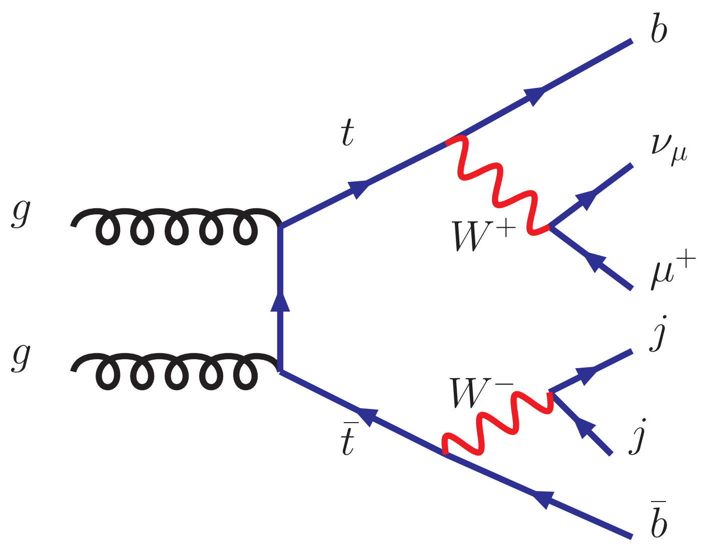
weight evaluation

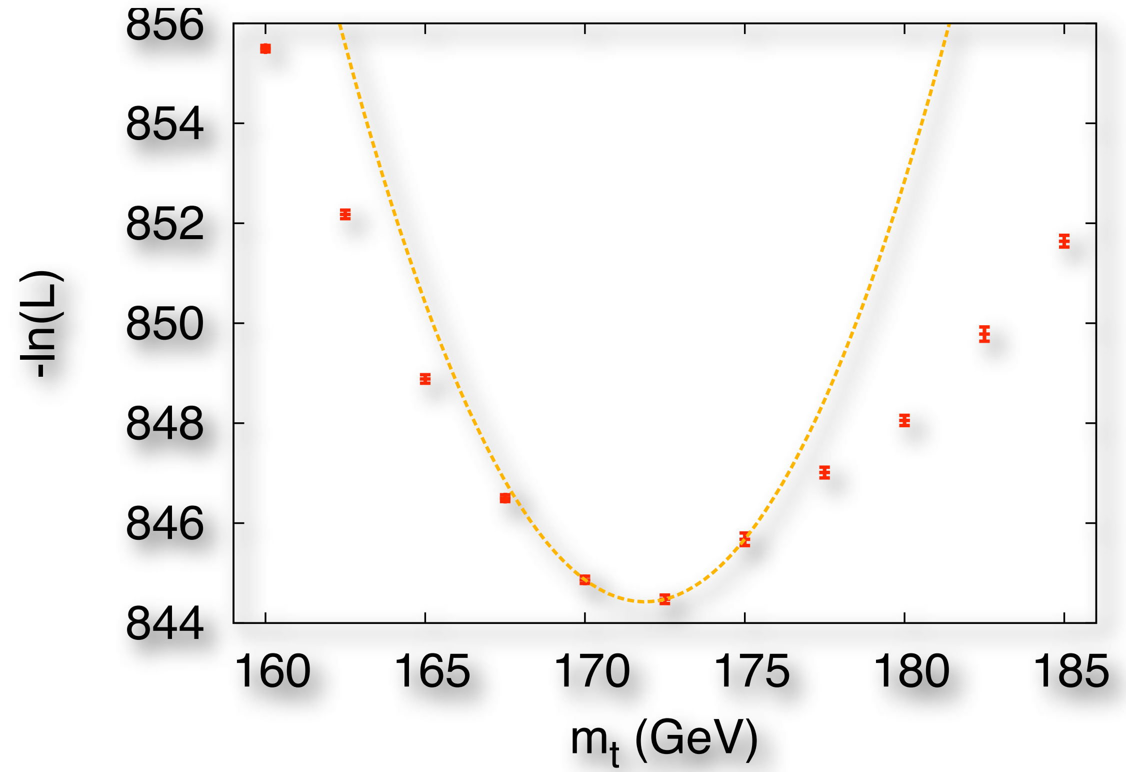
A. proc_card_mg5.dat
B. TF_my_tf.dat,
   transfer_card.dat
C. MadWeight_card.dat
D. run_card.dat, param_card.dat

expected time: ~1min / weight

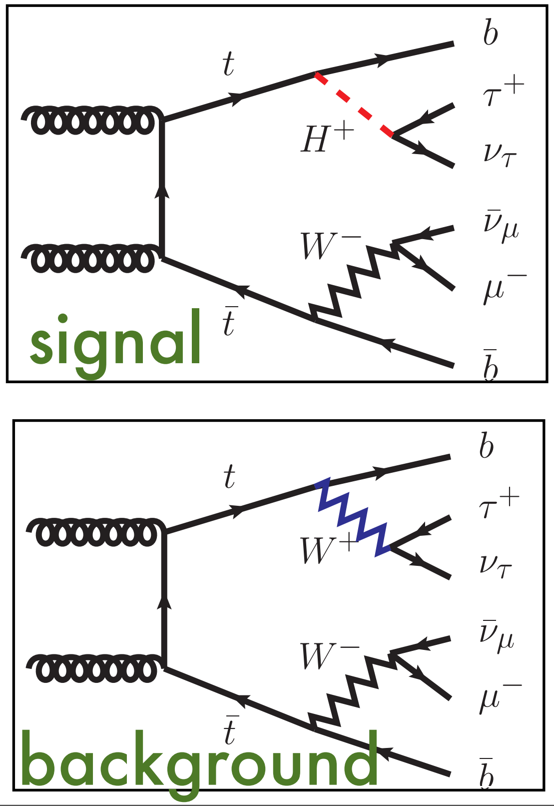
Other examples of application

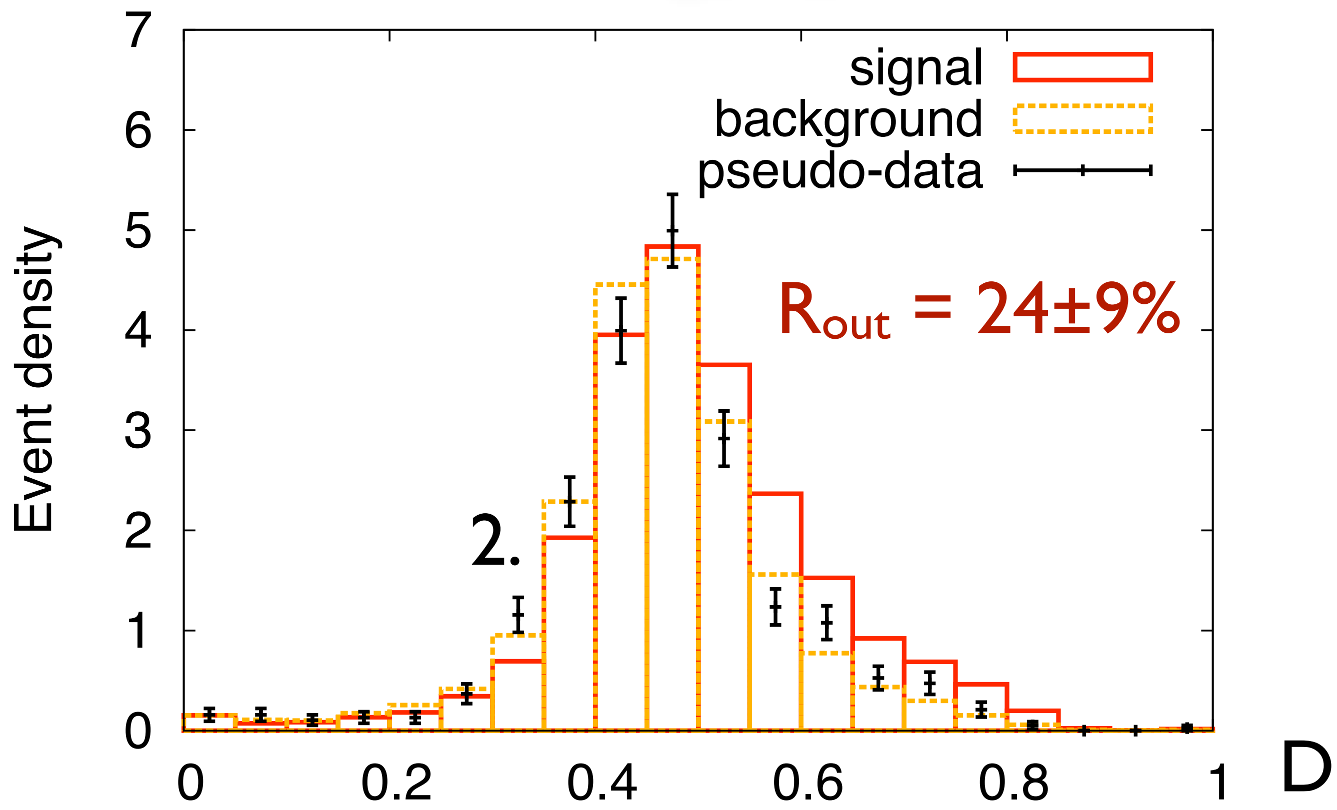
- mass determination:





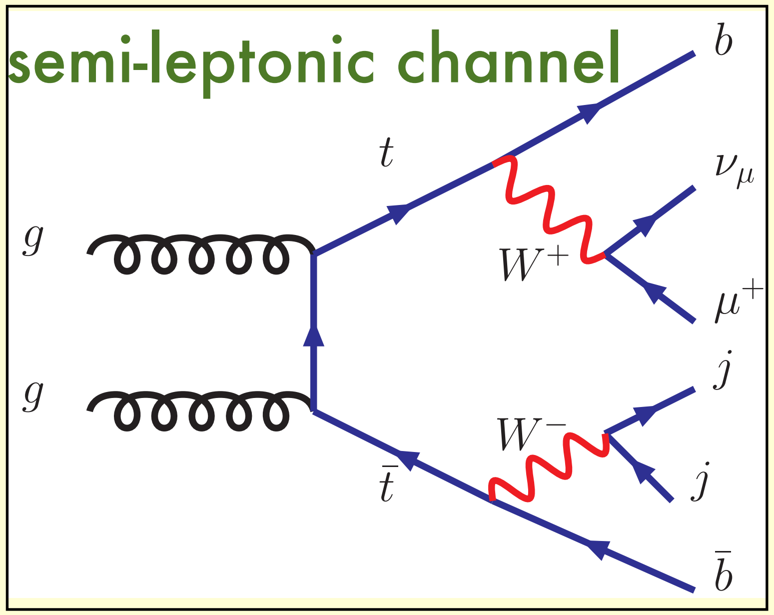
- spin determination:



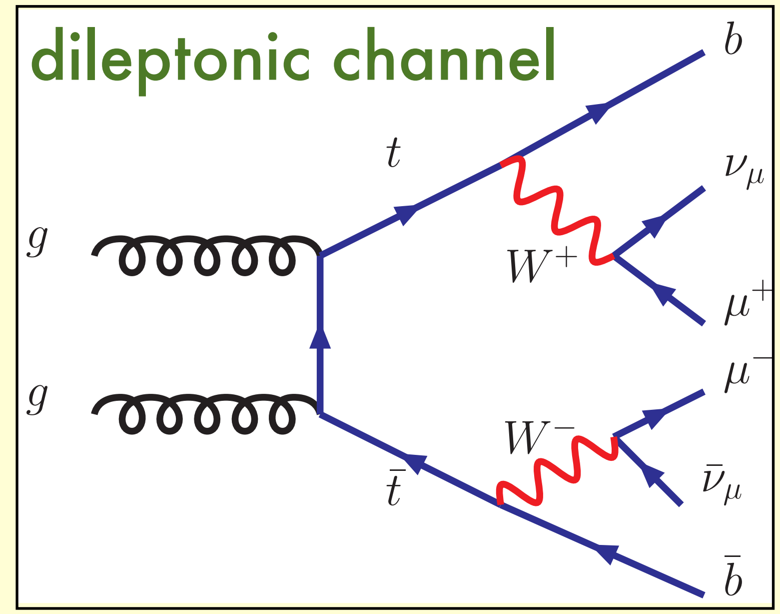


Aspects to investigate when applying MEM

- Evaluation of the weights: convergence of the Monte Carlo integration ?



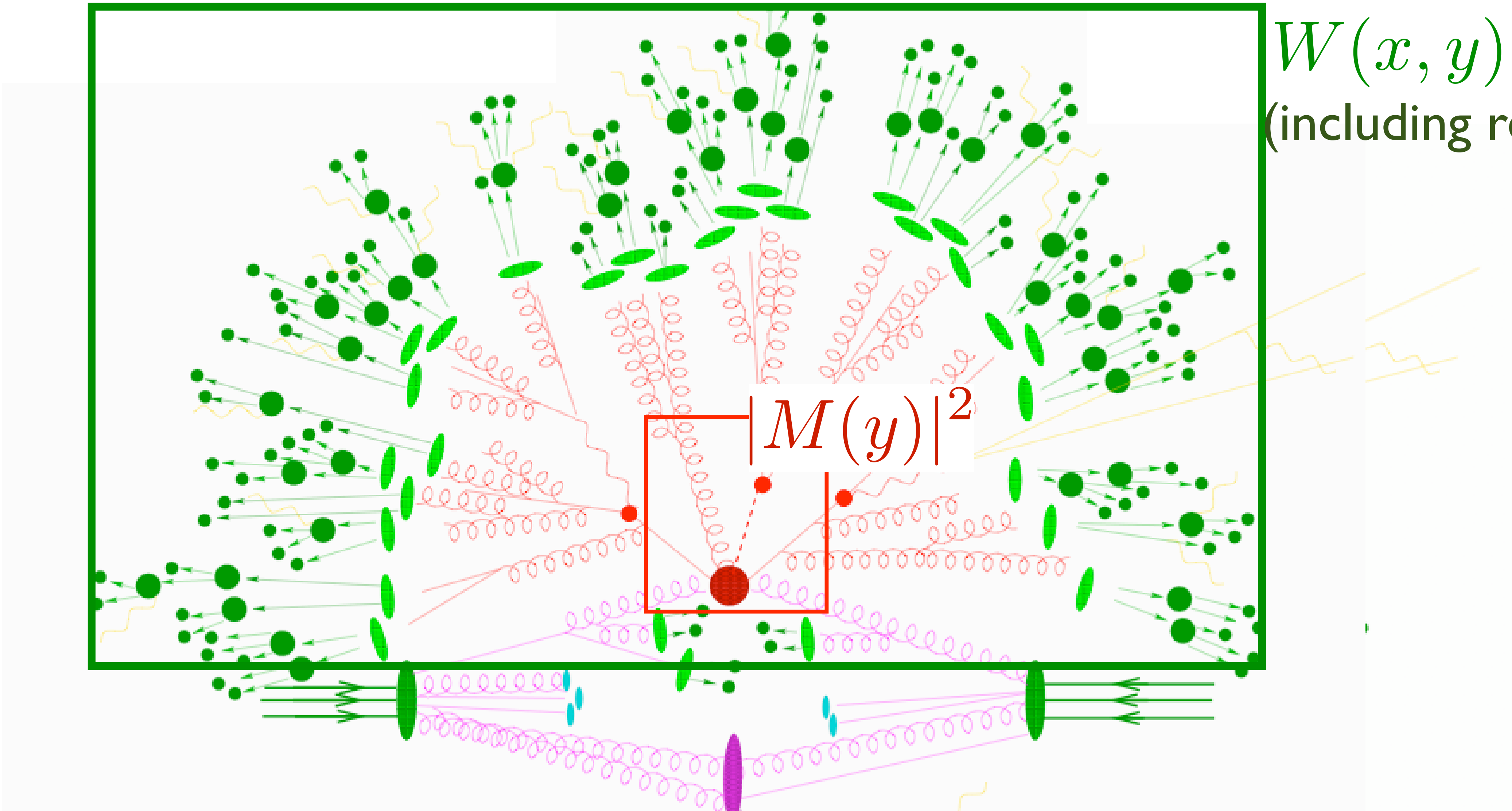
- overconstrained system, need to combine several PS channels
- 12 parton-jet assignements
- time spent on one weight: ~1 hour



- if poor resolution on E(jet), exactly constrained system, need to consider one phase-space channel
- 2 parton-jet assignements
- time spent on one weight: ~1min

Aspects to investigate when applying MEM

- Accuracy of the matrix element weights



- The prescription to extract the transfer function is only valid at LO in α_s ➪ the PDF has not been defined properly beyond LO accuracy.

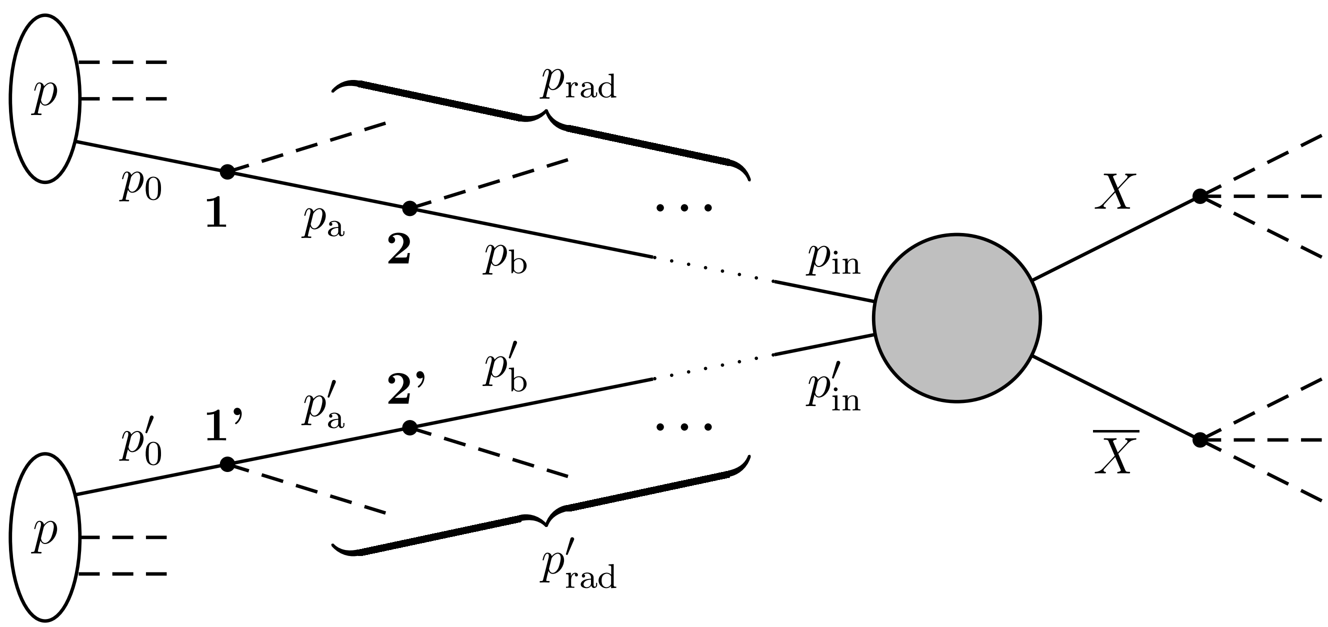
- the omission of higher order corrections can be treated as a systematic effect and is expected to lower the discriminant power

Aspects to investigate when applying MEM

- Accuracy of the matrix element weights

  - some prescription can be adopted to improve the model for the PDF in the Matrix Element Method, e.g. to take into account the dominant effects of ISR

J. Alwall, A. Freitas, O. Mattelaer arXiv:1010.2263



partial implemention in madweight 5
(three different ISR corrections)

Aspects to investigate when applying MEM

- Normalization of the weights: left to the user

  See e.g.

  Matrix Element in HEP: Transfer Functions, Efficiencies and Likelihood Normalization
  I. Volobouev, arXiv:1101.2259

Conclusion

MadWeight is designed to conduct Matrix-Element-based likelihood analyses in an efficient way (fast, reliable, reproducible)

it is an appropriate tool to test a new idea or conduct a pheno/experimental analysis in many instances

for more information on how to use it in practice: see the madgraph wiki

Backup slides

Backup 1: mass reconstruction

Q: assuming that the masses m_1 and m_2 are the only unknown, what is the maximum significance that can be achieved in measuring these masses at a given luminosity ?

Let us consider a specific example:

$$pp \to (\tilde{\mu}_r^+ \to \mu^+ \tilde{\chi}_1)(\tilde{\mu}_r^- \to \mu^- \tilde{\chi}_1)$$

sample of 50 events with $m_{\tilde{\mu}_r} = 150$ GeV

$m_{\tilde{\chi}_1} = 100$ GeV

$$(m^2_{\tilde{\mu}_r} - m^2_{\tilde{\chi}_1})/2m_{\tilde{\mu}_r} = 42 \text{ GeV}$$

possible discriminators:

- keeping only information from $p_T(\mu^+)$, $M(\mu^+, \mu^-)$

$$P(x|\tilde{\mu}_r, \tilde{\chi}_1) = \sigma^{-1}\frac{d\sigma}{dp_{T\mu}}(p_{T\mu}|m_{\tilde{\mu}_r}, m_{\tilde{\chi}_1}) \times \sigma^{-1}\frac{d\sigma}{dM_{\mu\mu}}(M_{\mu\mu}|m_{\tilde{\mu}_r}, m_{\tilde{\chi}_1})$$

- matrix element method (keeps all information): $P(x|\tilde{\mu}_r, \tilde{\chi}_1) =$ matrix element weight

Backup 1: mass reconstruction

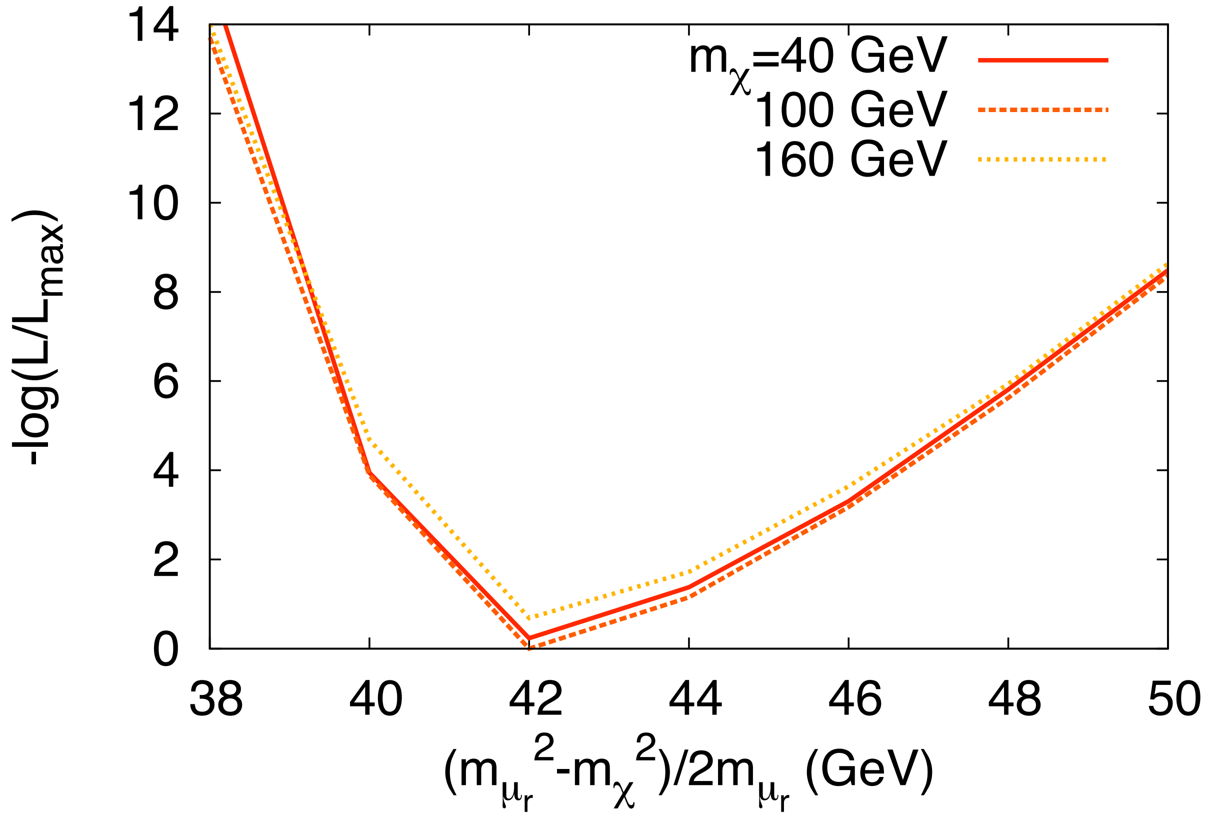
Q: assuming that the masses m_1 and m_2 are the only unknown, what is the maximum significance that can be achieved in measuring these masses at a given luminosity ?

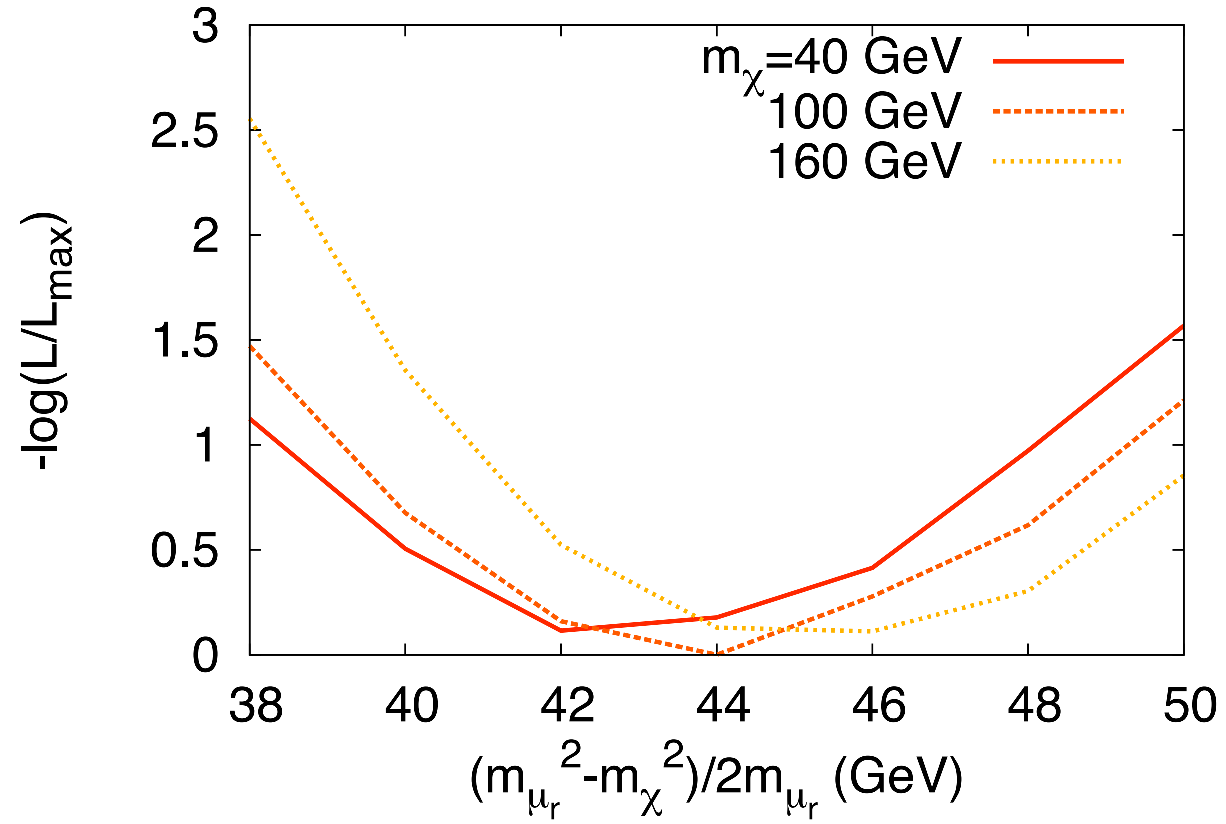
Let us consider a specific example:

$$pp \to (\tilde{\mu}_r^+ \to \mu^+ \tilde{\chi}_1)(\tilde{\mu}_r^- \to \mu^- \tilde{\chi}_1)$$

sample of 50 events with $m_{\tilde{\mu}_r} = 150$ GeV, $m_{\tilde{\chi}_1} = 100$ GeV

- matrix element method (keeps all information)
- keeping only information from $p_T(\mu^+)$, $M(\mu^+, \mu^-)$





Backup 2: Monte Carlo integration

1. 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: $\sigma_I \approx \frac{S}{\sqrt{N}}$

2. importance sampling: $\boldsymbol{z'} = \boldsymbol{P}(\boldsymbol{z}),\ p(\boldsymbol{z}) = Jac[\boldsymbol{P}(\boldsymbol{z})]$

$$\int dz\, f(\boldsymbol{z}) = \int \frac{f[\boldsymbol{P}^{-1}(\boldsymbol{z'})]}{p[\boldsymbol{P}^{-1}(\boldsymbol{z'})]} d\boldsymbol{z'} = \int \boxed{\frac{f(\boldsymbol{z})}{p(\boldsymbol{z})}}\, \boxed{p(\boldsymbol{z}) d\boldsymbol{z}}$$

new integrand

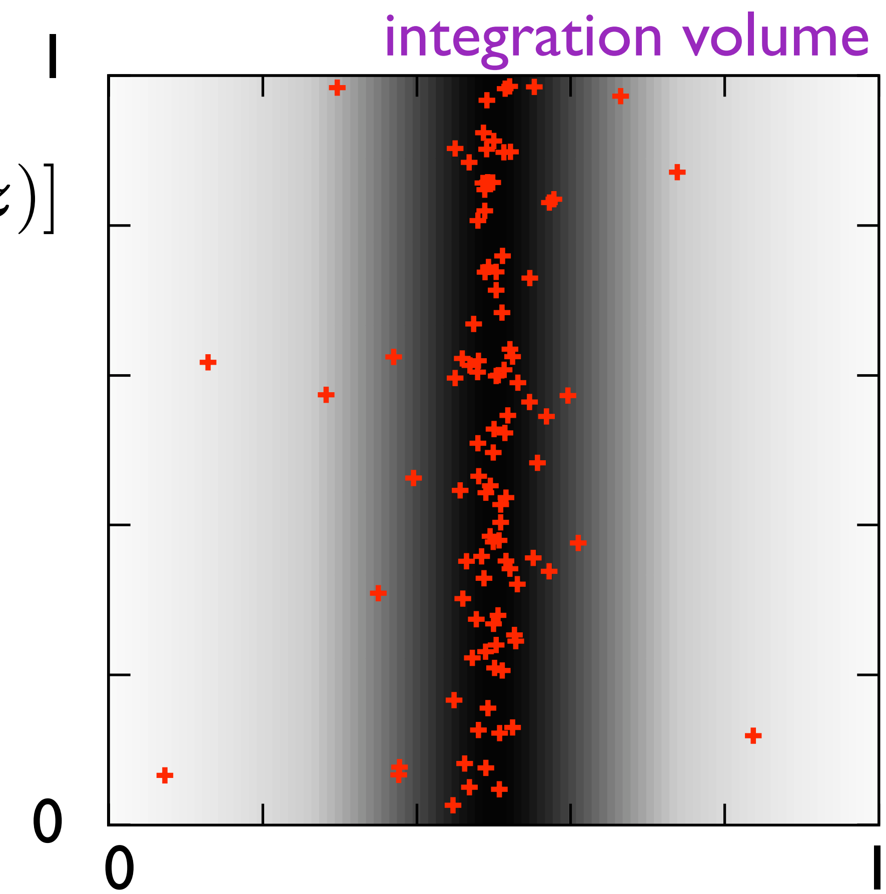
new integr. measure

if $\{\boldsymbol{z}_n\}$ distributed according to $p(\boldsymbol{z})$ then

$$E \to \frac{1}{N}\sum_{n=1}^{N} \frac{f(\boldsymbol{z}_n)}{p(\boldsymbol{z}_n)}$$

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

S is decreased if $p(\boldsymbol{z}) \approx f(\boldsymbol{z})/E$



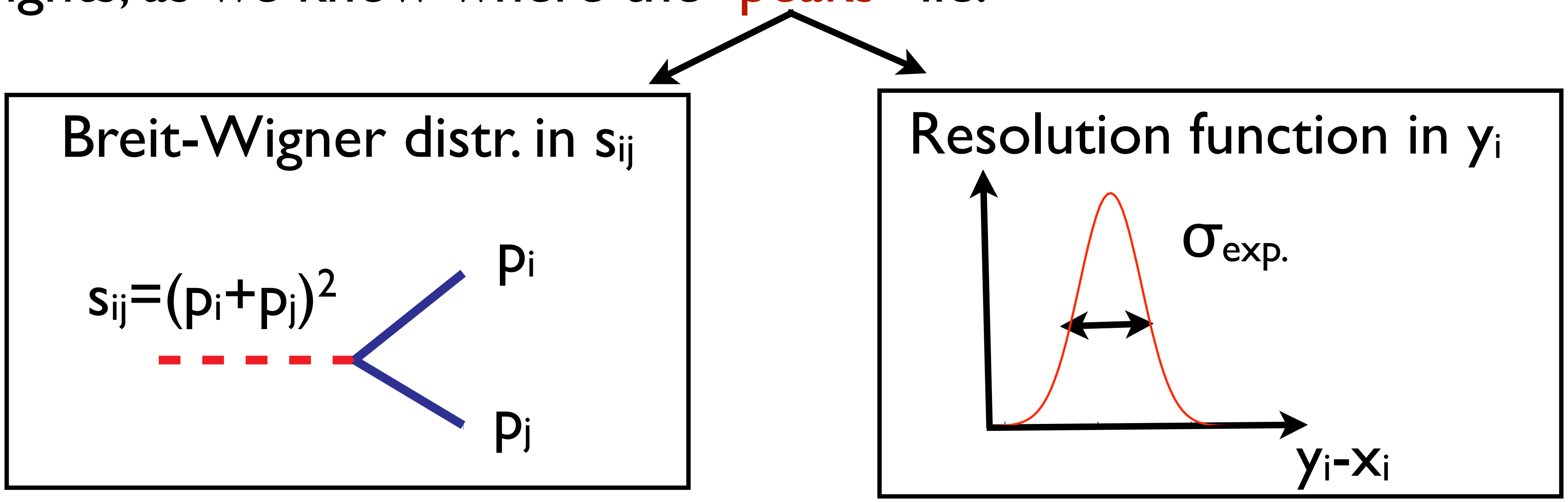
3. adaptive Monte Carlo integration:

$p(\boldsymbol{z}) = p_1(z^1)p_2(z^2)\ldots p_d(z^d)$ (grid)

optimized using an iteration procedure

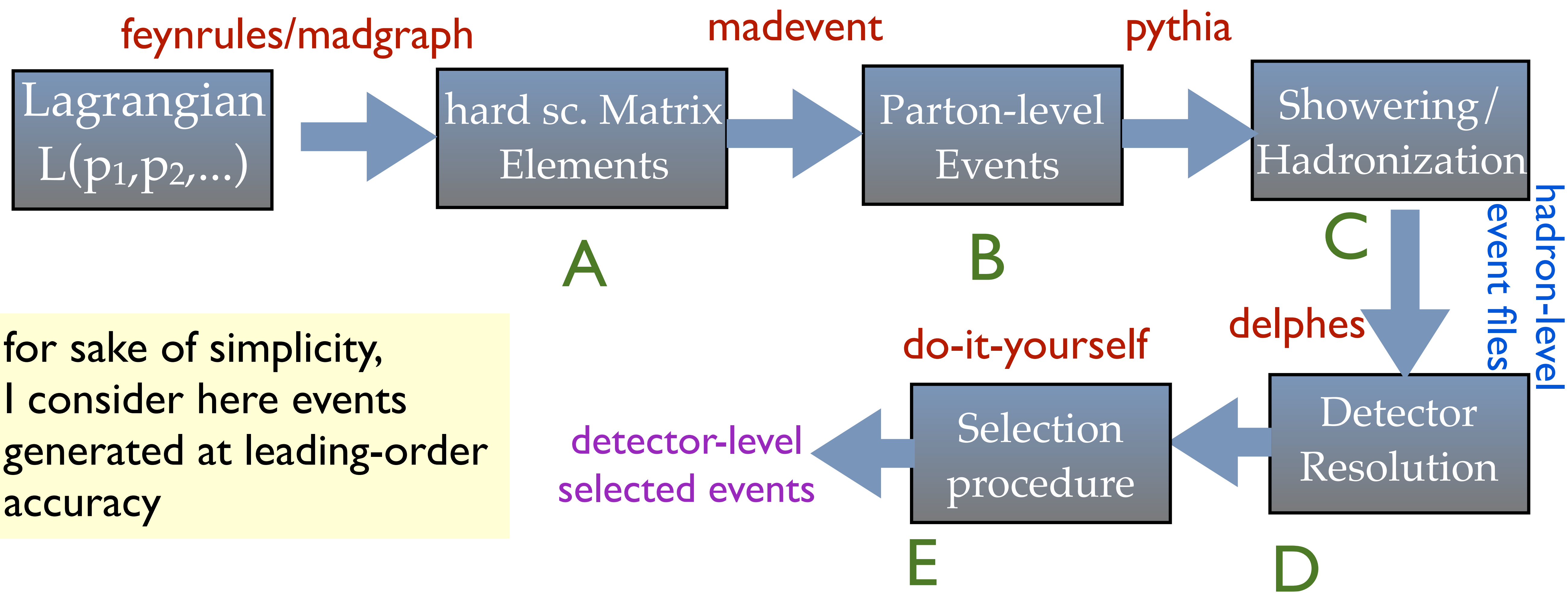
New phase-space mappings

- adaptive MC integration can be used for the computation of the weights, as we know where the "peaks" lie:



- for a given decay chain and a given transfer function, one needs to construct a new parametrization of the phase-space measure
- in the MEM analyses at the Tevatron, this problem was solved on a case-by-case basis

I. generation of the events



for sake of simplicity,
I consider here events
generated at leading-order
accuracy

input parameters:

A. proc_card_mg5.dat

B. run_card.dat, param_card.dat

C. pythia_card.dat

D. delphes_card.dat, delphes_trigger.dat

E. c++ code, existing template

fast, reliable, reproducible