MadWeight

Automatic events reweighting with the matrix element method

Mattelaer Olivier

Université Catholique de Louvain

Pierre Artoisenet: UCL-CP3

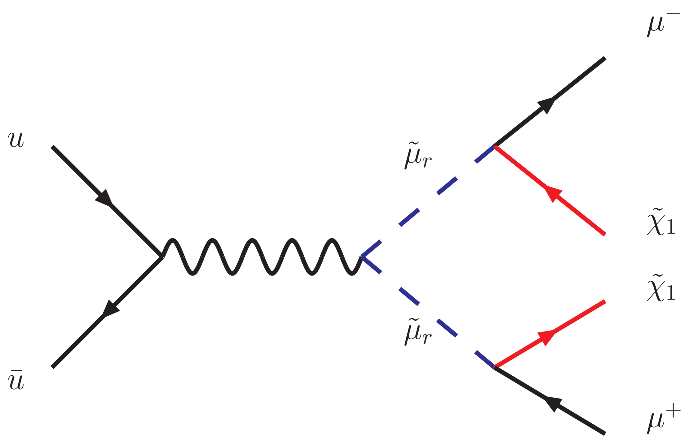
Fabio Maltoni: UCL-CP3

Vincent Lemaître: UCL-CP3

Motivation

- Study of resonances at hadron colliders in decay chains with invisible particles

Example : slepon production



- How to identify this signature ?
- How to measure the properties of (new) particles in the decay chain ?

- Method to maximize the information that you can extract from a sample of events

Motivation

Outline :

- Matrix Element Method : procedure to discriminate between two theoretical assumptions using the maximum amount of information

- MadWeight : automatic procedure to apply matrix element techniques

Matrix Element Method

- given a theoretical assumption α, attach a **weight** $P(\boldsymbol{x}, \alpha)$ to each experimental event $\boldsymbol{x}$ quantifying the validity of the theoretical assumption α for this event.

$$P(\boldsymbol{x}, \alpha) = \qquad |M_\alpha|^2(\boldsymbol{x})$$

where

- $|M_\alpha|^2$ is the squared matrix element

Matrix Element Method

- given a theoretical assumption α, attach a **weight** $P(\boldsymbol{x}, \alpha)$ to each experimental event $\boldsymbol{x}$ quantifying the validity of the theoretical assumption α for this event.

$$P(\boldsymbol{x}, \alpha) = \qquad |M_\alpha|^2(\boldsymbol{y}) W(\boldsymbol{x}, \boldsymbol{y})$$

where

- $|M_\alpha|^2$ is the squared matrix element
- $W(\boldsymbol{x}, \boldsymbol{y})$ is the resolution function

Matrix Element Method

- given a theoretical assumption α, attach a **weight** $P(\boldsymbol{x},\alpha)$ to each experimental event $\boldsymbol{x}$ quantifying the validity of the theoretical assumption α for this event.

$$P(\boldsymbol{x},\alpha) = \frac{1}{\sigma}\int d\phi(\boldsymbol{y})|M_\alpha|^2(\boldsymbol{y})W(\boldsymbol{x},\boldsymbol{y})$$

where

- $|M_\alpha|^2$ is the squared matrix element
- $W(\boldsymbol{x},\boldsymbol{y})$ is the resolution function
- $d\phi(\boldsymbol{y})$ is the parton-level phase-space measure

Matrix Element Method

- given a theoretical assumption α, attach a **weight** $P(\boldsymbol{x},\alpha)$ to each experimental event $\boldsymbol{x}$ quantifying the validity of the theoretical assumption α for this event.

$$P(\boldsymbol{x},\alpha) = \frac{1}{\sigma}\int d\phi(\boldsymbol{y})|M_\alpha|^2(\boldsymbol{y})W(\boldsymbol{x},\boldsymbol{y})$$

where

- $|M_\alpha|^2$ is the squared matrix element
- $W(\boldsymbol{x},\boldsymbol{y})$ is the resolution function
- $d\phi(\boldsymbol{y})$ is the parton-level phase-space measure

The value of the weight $P(\boldsymbol{x},\alpha)$ is the probability to observe the experimental event $\boldsymbol{x}$ in the theoretical frame α.

Weighting experimental events

- combine the weights into a likelihood

$$L(\alpha) = \prod_{i=1}^{N} P(\boldsymbol{x}_i; \alpha)$$

Weighting experimental events

- combine the weights into a likelihood

$$L(\alpha) = e^{-N \int P(\boldsymbol{x},\alpha) d\boldsymbol{x}} \prod_{i=1}^{N} P(\boldsymbol{x}_i;\alpha)$$

Weighting experimental events

- combine the weights into a likelihood

$$L(\alpha) = e^{-N \int P(\boldsymbol{x},\alpha) d\boldsymbol{x}} \prod_{i=1}^{N} P(\boldsymbol{x}_i; \alpha)$$

the best estimation of α is the one that maximizes L

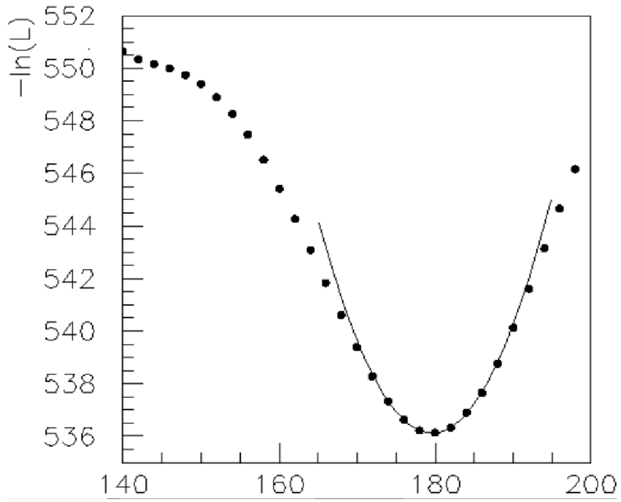
Weighting experimental events

- combine the weights into a likelihood

$$L(\alpha) = e^{-N\int P(\boldsymbol{x},\alpha)d\boldsymbol{x}} \prod_{i=1}^{N} P(\boldsymbol{x}_i;\alpha)$$

the best estimation of α is the one that maximizes L

example : top-quark mass measurement from $t\bar{t} \to l^+X$ sample at DØ

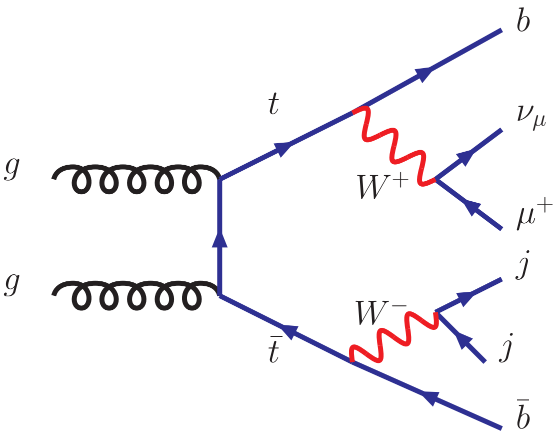


- 72 events
- $M_{top} = 180.1 \pm 3.6_{stat} \pm 4.0_{sys}$ GeV
- J. Estrada : Phd dissertation, University of Rochester (2001)

Examples of Matrix Element analysis

- top-quark mass determination from top-quark pair events

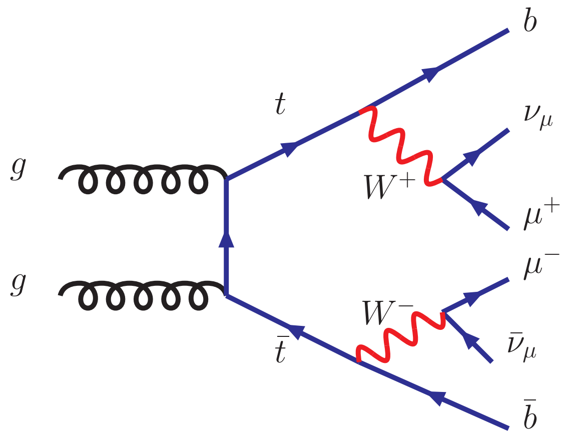
semi-leptonic channel



D0 collaboration : Phys. Rev. D74 092005, 2006

CDF collaboration : Phys. Rev. Lett. 99 182002, 2007

double-leptonic channel



D0 collaboration : Phys. Lett. B655 :7, 2007

CDF collaboration : Phys. Rev. D75 :031105, 2007

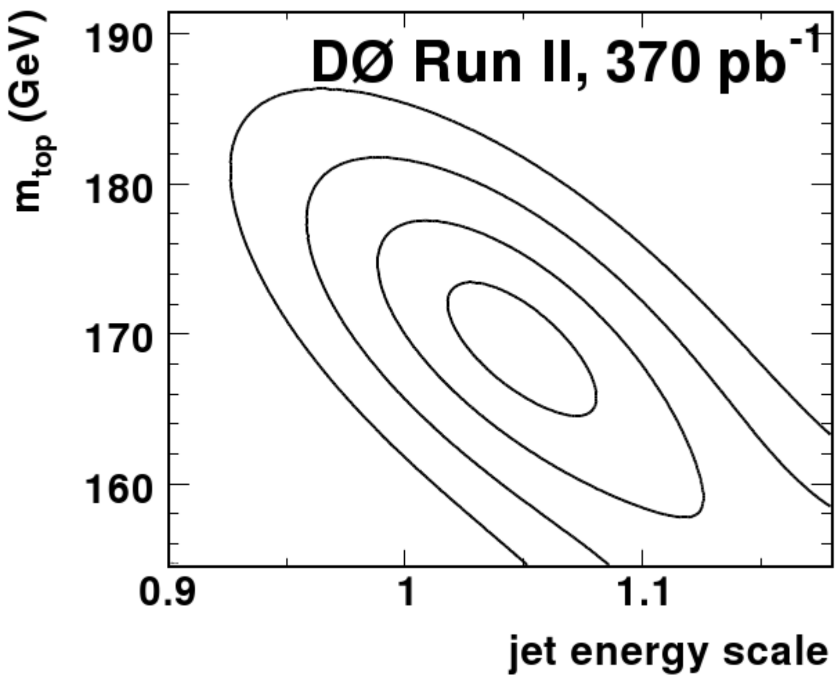
Remarks :

- all jet combinations are considered (info from b-tagging included)
- internal check of the Jet Energy Scale (JES fixed by the maximization of the likelihood)

Examples of Matrix Element analysis

- top-quark mass determination from top-quark pair events

  Results for the semi-leptonic channel (D0 collaboration)



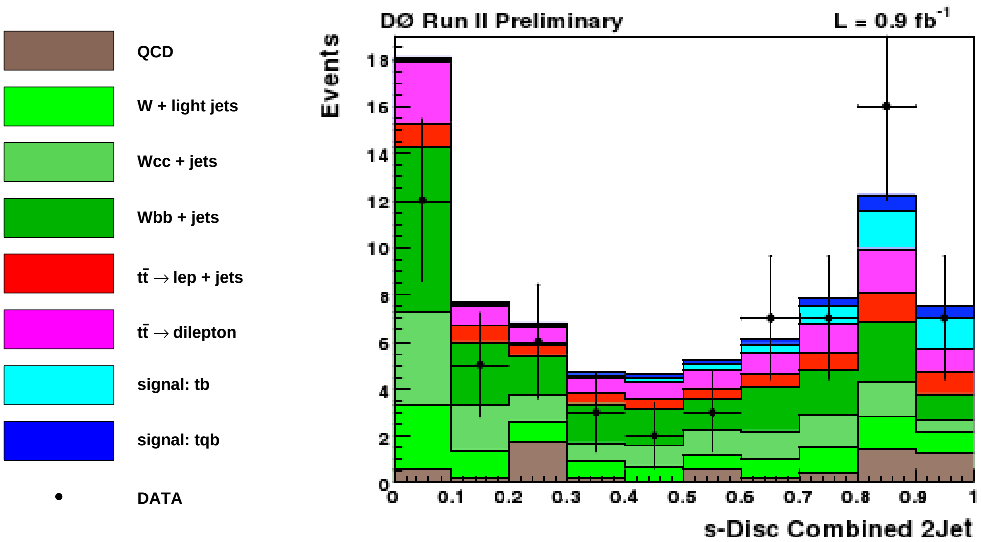
- 0.4 fb^{-1}
- 175 events
- $170.3 \pm 4.5 \pm 1.8$GeV

  Phys.Rev.D74 :092005 (2006)

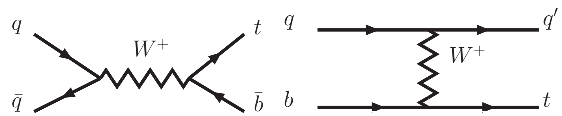
Examples of Matrix Element analysis

- Definition of a event-by-event discriminator

$$D(x) = \frac{P(x|S)}{P(x|S) + P(x|B)}$$



- Single Top cross section
- s- and t-channels



- $\sigma(p\bar{p} \to tb + tbq + X) = 4.8^{+1.6}_{-1.4}$ pb

  J. Mitrevski : D0 note 5392-CONF

- Maximum significance at the LHC :Cranmer & al :hep-ph/0605268

The Matrix Element Method

- advantages :
  - it is conceptually simple
  - the maximum amount of experimental information can be used to discriminate different theoretical hypothesis
  - the events are weighted with the squared matrix element $\longrightarrow$ refined analysis of the decay chain (spin, coupling types, masses, ...)

- drawbacks :
  - difficult to estimate the systematic errors
    in particular : parametrization of the transfer functions ?
  - the evaluation of the weight is time-consuming : one phase-space integration per event and per theoretical assumption.

Practical evaluation of the weights

$$P(\boldsymbol{x}, \alpha) = \frac{1}{\sigma} \int |M_\alpha|^2(\boldsymbol{y}) \;\; W(\boldsymbol{x}, \boldsymbol{y}) \;\; d\phi(\boldsymbol{y})$$

Practical evaluation of the weights

$$P(\boldsymbol{x}, \alpha) = \frac{1}{\sigma} \int |M_\alpha|^2(\boldsymbol{y}) \; W(\boldsymbol{x}, \boldsymbol{y}) \; d\phi(\boldsymbol{y})$$

Amplitude generator (MadGraph)

Practical evaluation of the weights

$$P(\boldsymbol{x}, \alpha) = \frac{1}{\sigma} \int |M_\alpha|^2(\boldsymbol{y}) \, W(\boldsymbol{x}, \boldsymbol{y}) \, d\phi(\boldsymbol{y})$$

Amplitude generator (MadGraph)

fit from MC tuned to the resolution of the detector

Practical evaluation of the weights

$$P(\boldsymbol{x}, \alpha) = \frac{1}{\sigma} \int |M_\alpha|^2(\boldsymbol{y}) \, W(\boldsymbol{x}, \boldsymbol{y}) \, d\phi(\boldsymbol{y})$$

Amplitude generator (MadGraph)

fit from MC tuned to the resolution of the detector

Phase space generator (MadWeight)

Practical evaluation of the weights

Numerical integration : In addition to the propagators in $|M_\alpha(\boldsymbol{y})|^2$, new peaks are introduced with the transfer function

$$W(\boldsymbol{x},\boldsymbol{y}) \approx \prod_i \frac{1}{\sqrt{2\pi}\sigma_{E,i}} e^{-\frac{(E_i^{rec}-E_i^{gen})^2}{2\sigma^2_{E,i}}}$$

resolution

in energy

$$\times \frac{1}{\sqrt{2\pi}\sigma_{\phi,i}} e^{-\frac{(\phi_i^{rec}-\phi_i^{gen})^2}{2\sigma^2_{\phi,i}}}$$

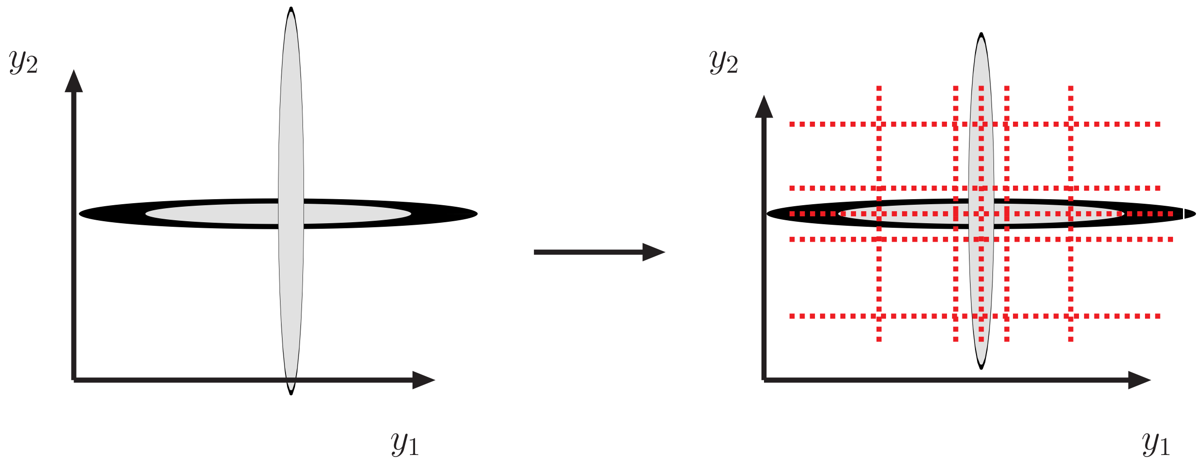
in azimuthal angle

$$\times \frac{1}{\sqrt{2\pi}\sigma_{\eta,i}} e^{-\frac{(\eta_i^{rec}-\eta_i^{gen})^2}{2\sigma^2_{\eta,i}}}$$

in pseudo-rapidity

Monte Carlo integration

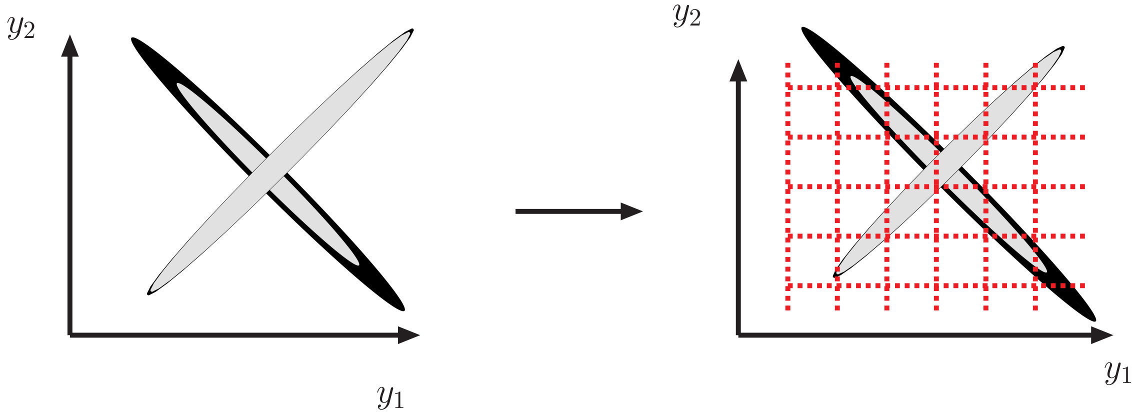
- choice of the phase-space parametrization has a strong impact on the efficiency of the MC integration :
  - any peak is aligned along a single direction of the P-S parametrization



$\longrightarrow$ the adaptive Monte-Carlo P-S integration is very efficient

Monte Carlo integration

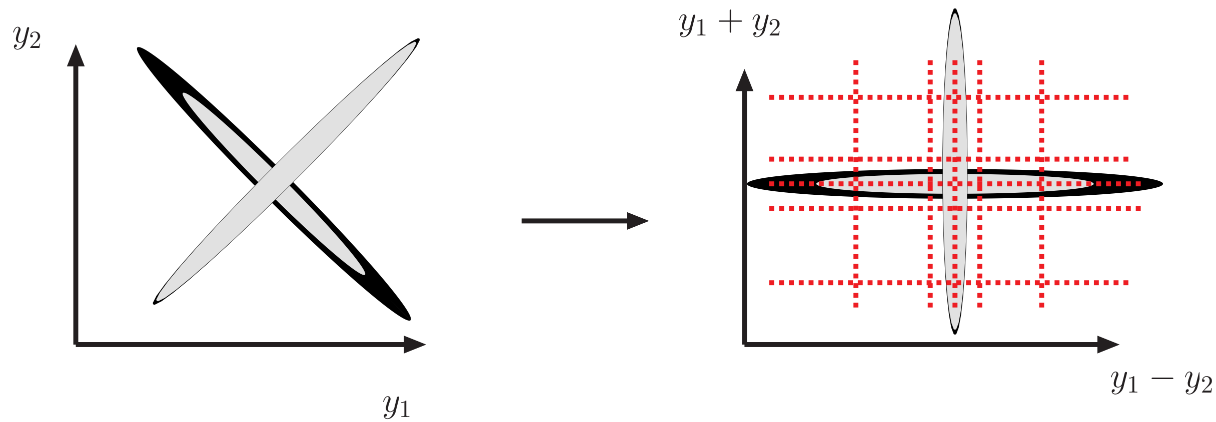
- choice of the phase-space parametrization has a strong impact on the efficiency of the MC integration :
  - some peaks are not aligned along a single direction of the P-S parametrization



$\longrightarrow$ the adaptive Monte-Carlo P-S integration converges slowly

Monte Carlo integration

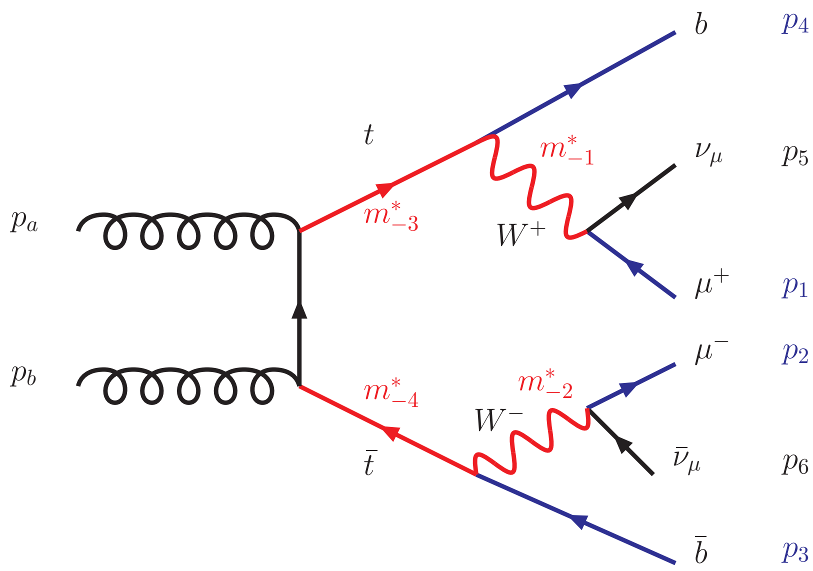
- choice of the phase-space parametrization has a strong impact on the efficiency of the MC integration :
  - solution to the previous case : perform a change of variables in order to align the peaks along a single direction of the P-S parametrization



$\longrightarrow$ the adaptive Monte-Carlo P-S integration is very efficient

Phase-space generation ($t\bar{t}$ production)

- Example : top-quark pair production in the double-leptonic channel



- peaks in $|M_\alpha(\boldsymbol{y})|^2$ controlled by $m^*_{-1}, \ldots, m^*_{-4}$ (4 variables)
- peaks in $W(\boldsymbol{x}, \boldsymbol{y})$ controlled by $\theta_i, \phi_i, |p_i|^2 \quad i \in \{1, 2, 3, 4\}$ (12 variables)
- $dim[d\phi] = 16$, $\rightarrow$ each peak can be aligned along a single variable of integration

Phase-space generation ($t\bar{t}$ production)

optimal parametrization ?

- start from the parametrization

$$d\phi = \prod_{i=1}^{4} \frac{d^3 p_i}{(2\pi)^3 2E_i} \prod_{i=5}^{6} \frac{d^3 p_i}{(2\pi)^3 2E_i} dx_1 dx_2 \delta^4 \left(p_a + p_b - \sum_j p_j \right)$$

  where all the peaks in $W(\boldsymbol{x}, \boldsymbol{y})$ are aligned

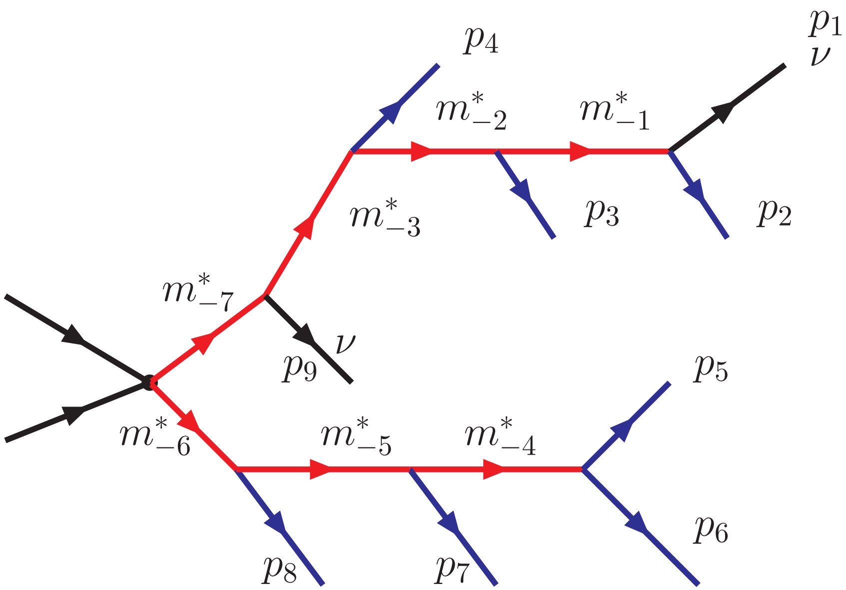
- apply local changes of variables to reach the parametrization

$$d\phi = \prod_{i=1}^{4} d\theta_i d\phi_i d|\boldsymbol{p}_i| \prod_{j=1}^{4} dm^{*2}_{-j} \times J$$

where each Breit-Wigner distribution is aligned

Phase-space generation (generalization)

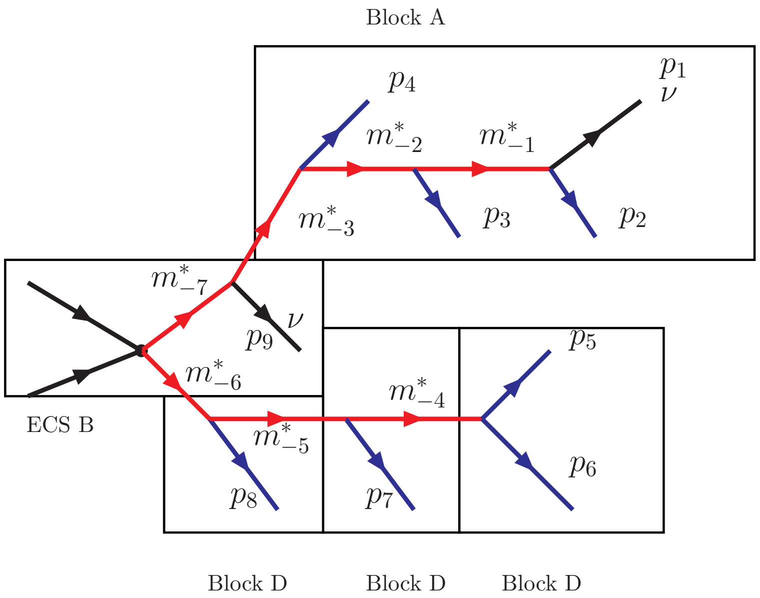
Let us consider a example of generic decay chain :



- peaks in $|M_\alpha(\boldsymbol{y})|^2$ controlled by $m^*_{-1}, \ldots, m^*_{-7}$ (7 variables)
- peaks in $W(\boldsymbol{x}, \boldsymbol{y})$ controlled by $\theta_i, \phi_i, |p_i|^2 \quad i \in \{2, 3, 4, 5, 6, 7, 8\}$ (21 variables)
- $dim[d\phi] = 25$, $\rightarrow$ some peaks must be left misaligned

Phase-space generation

- each local change of variable is performed successively

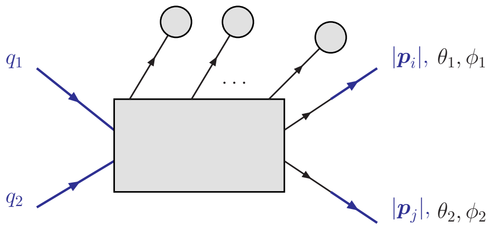


Final parametrization :

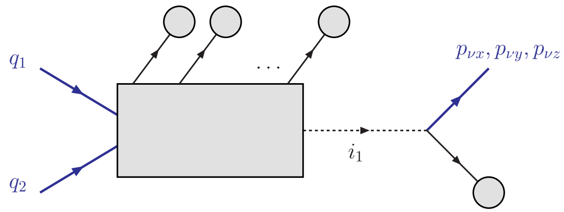
$$d\phi = d|\boldsymbol{p}_2|d|\boldsymbol{p}_3|d|\boldsymbol{p}_4|d|\boldsymbol{p}_6| \prod_{i=2}^{8} d\theta_i d\phi_i \prod_{j=1}^{7} dm^{*2}_{-j} \times J$$

Phase-space generation

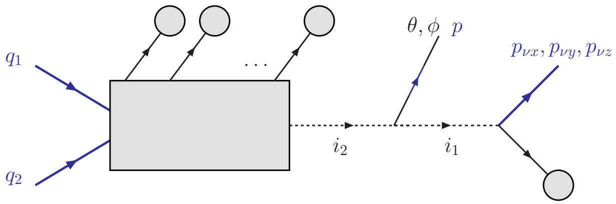
- ECS :



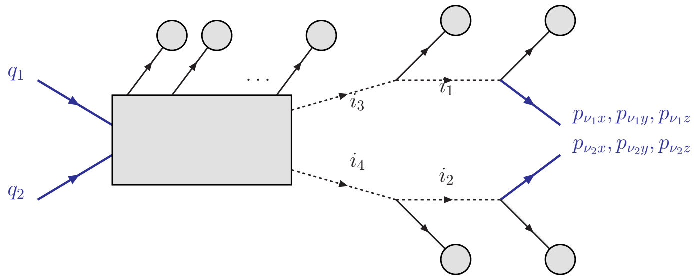
Class A



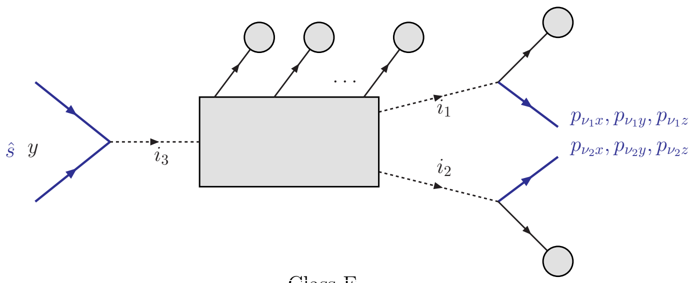
Class B



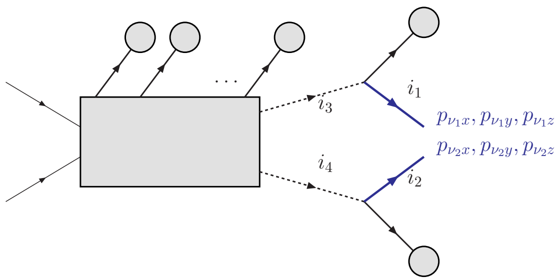
Class C



Class D



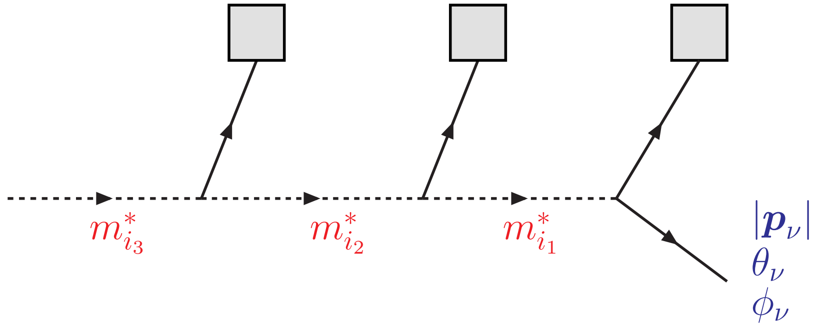
Class E



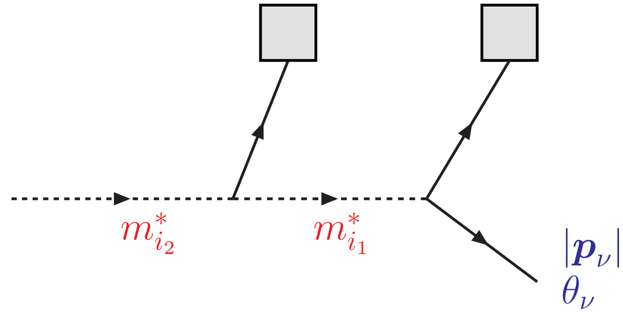
Class F

Phase-space generation

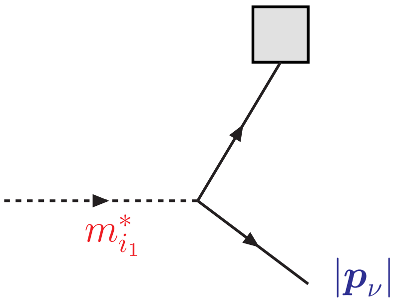
- Blocks :



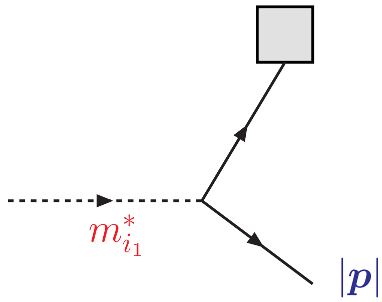
block A



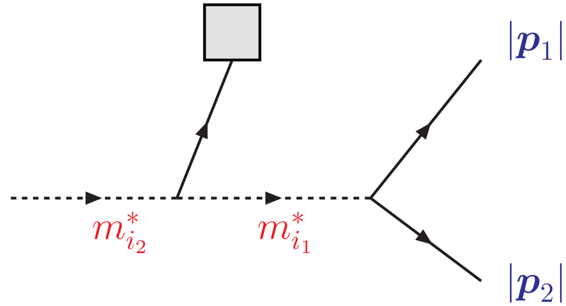
block B



block C



block D



block E

Phase-space generation

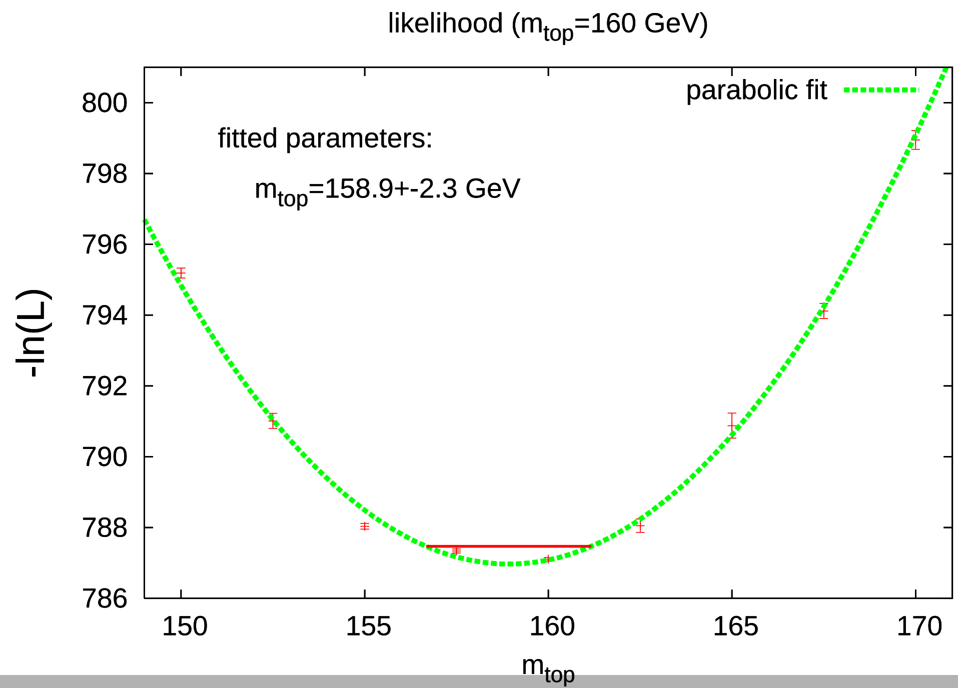
In MadWeight,

- the phase-space is splitted into *blocks*, each of them is associated to a specific local change of variables
- 12 blocks, i.e. 12 analytic changes of variables have been defined in our code
- Given
  - the decay chain under study
  - the resolution function for each visible particle

  MadWeight
  - finds the optimal partition of the PS into blocks
  - computes the weights using the corresponding PS parametrization

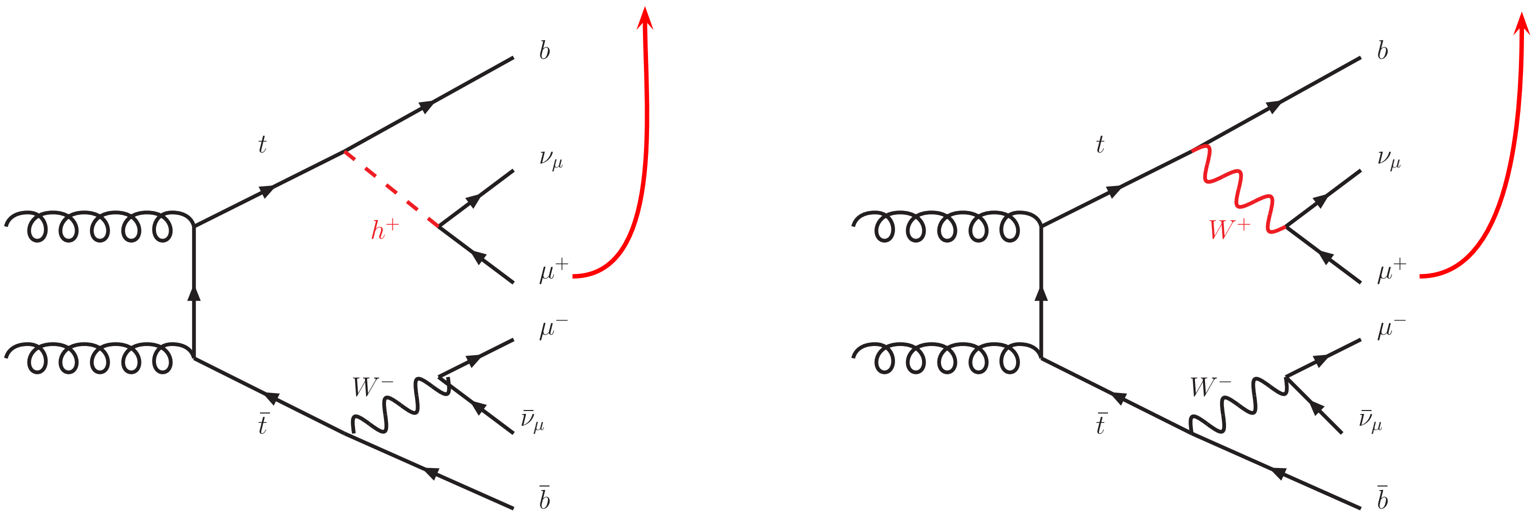
Application

- Example 1 : measurement of the top-quark mass in semi-leptonic channel
  - 20 Monte Carlo events (MadGraph/Pythia/PGS)
  - $L(m_t) = e^{-N \int P(\boldsymbol{x}, m_t) d\boldsymbol{x}} \prod_{i=1}^{N} P(\boldsymbol{x}_i; m_t)$
  - input : $m_{top} = 160$ GeV, output : $m_{top} = 158.9 \pm 2.3$ GeV



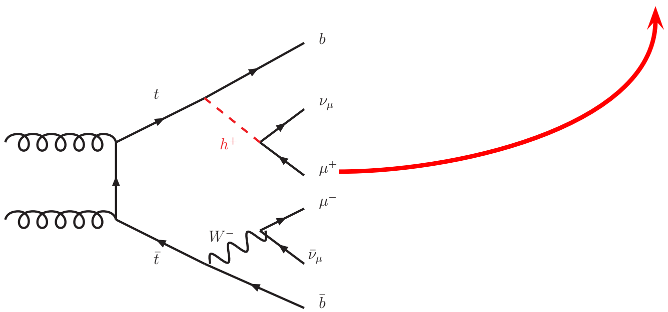
Applications

- Example 2 :
  separate the signal $t \to H^+ b$ from the background $t \to W^+ b$

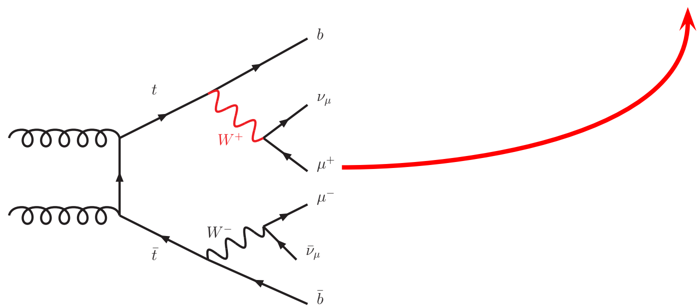


Applications

- Example 2 :
  separate the signal $t \to H^+ b$ from the background $t \to W^+ b$

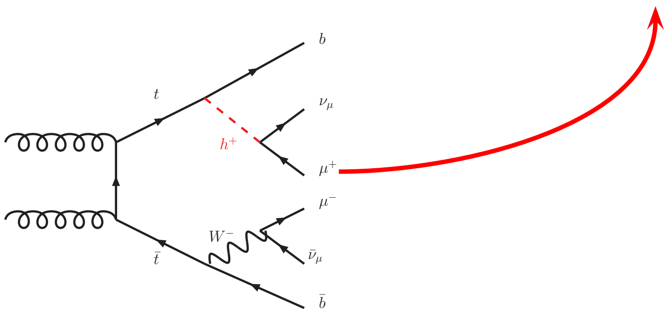


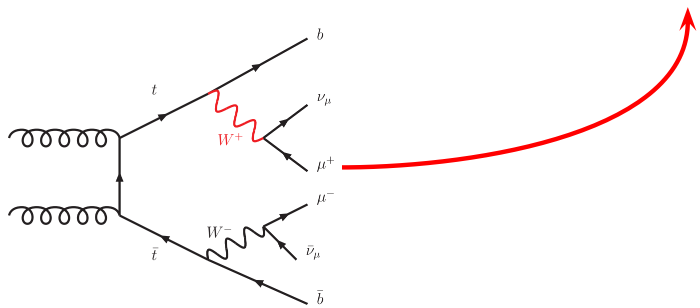
3



Applications

- Example 2 :
  separate the signal $t \to H^{+}b$ from the background $t \to W^{+}b$

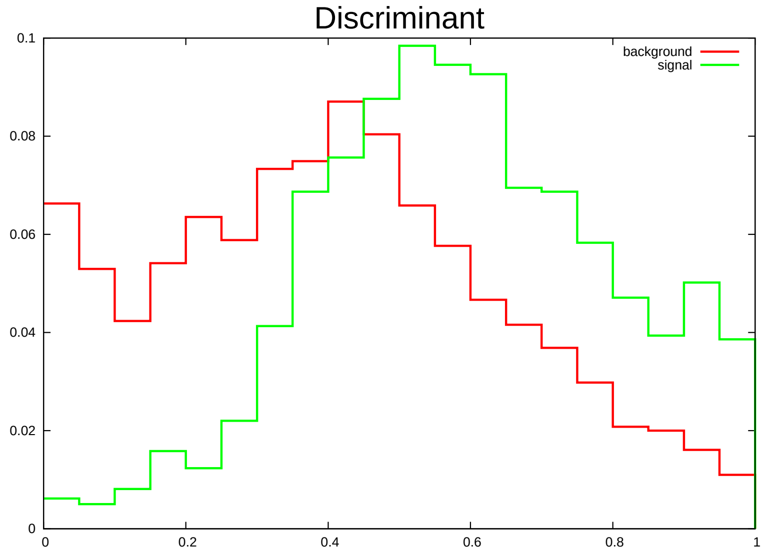




3

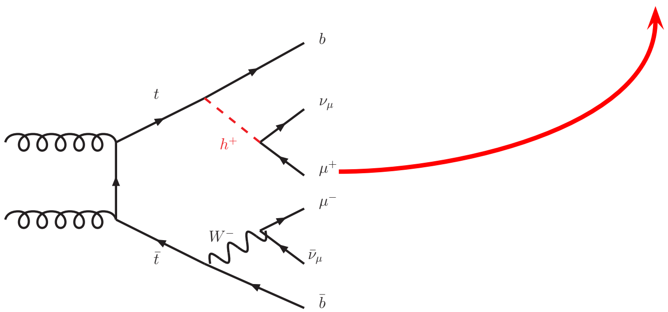
- define the discriminator

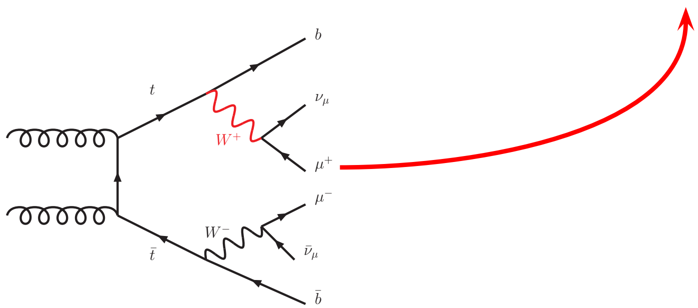
$$d = \frac{P_S}{P_S+P_B}$$



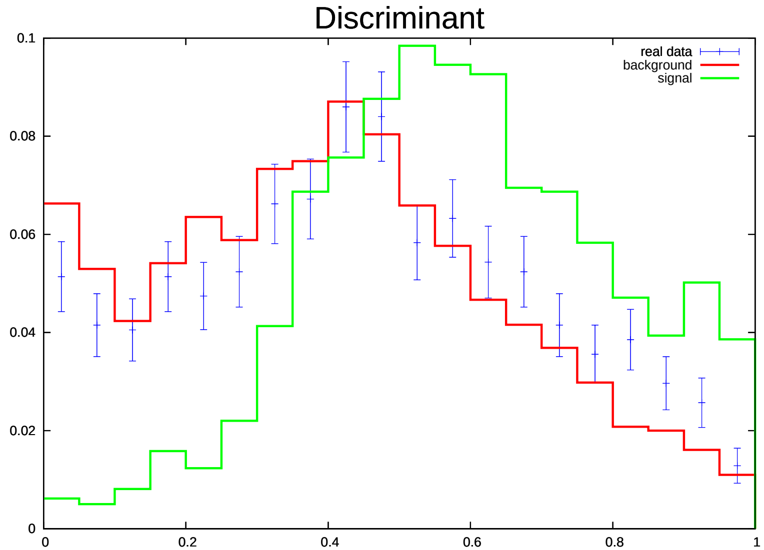
Applications

- Example 2 :
  separate the signal $t \to H^+ b$ from the background $t \to W^+ b$



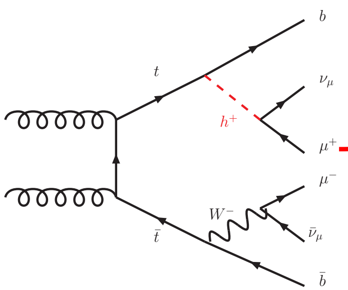


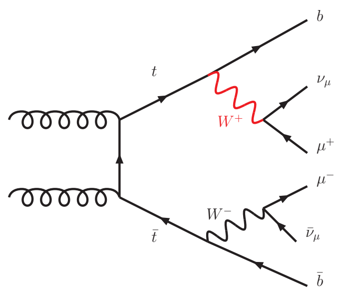
- define the discriminator
  $d = \frac{P_S}{P_S + P_B}$
- 750 background events
- 262 signal events
- $r = 25.9\%$
- $r_{mes} = 21 \pm 4\%$



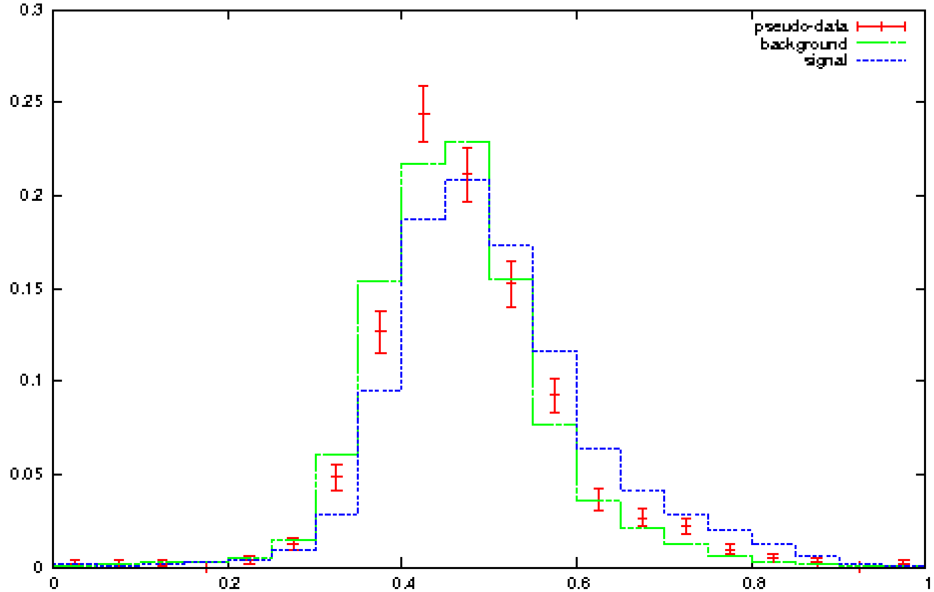
Applications

- Example 2 :
  separate the signal $t \rightarrow H^{+}b$ from the background $t \rightarrow W^{+}b$



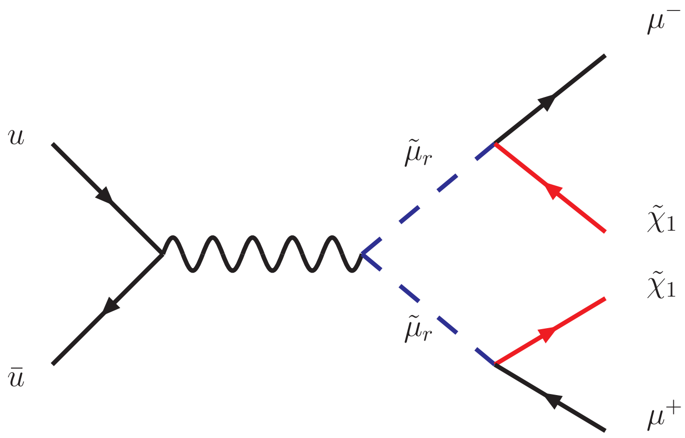


- $M_{H^+} = M_{W^+}$
- $\sigma = 1.632pb$
- $L = 8.5fb^{-1}$
- $\sigma_{mes} = 1.7 \pm 0.4pb$



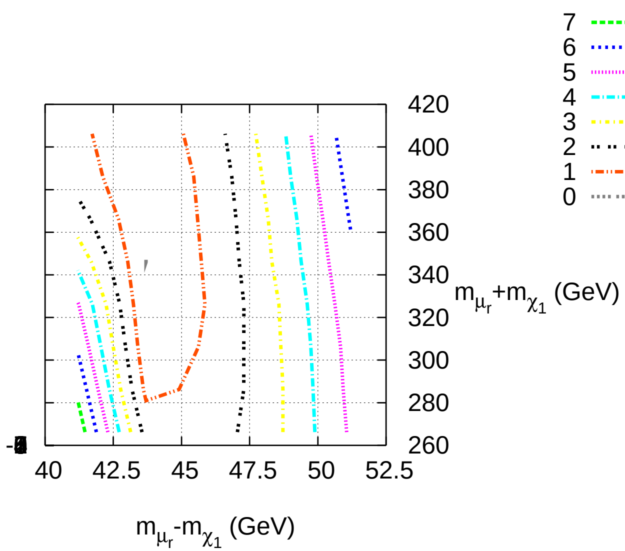
Applications

- Example 3 : slepton pair production



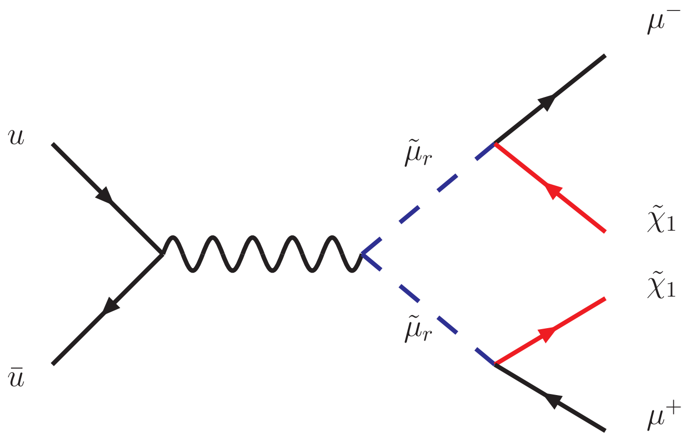
aim : measure of the masses $m_{\tilde{\mu_r}}$, $m_{\tilde{\chi_1}}$ from a pure signal sample

- inputs : $m_{\tilde{\mu_r}} = 144$ GeV, $m_{\tilde{\chi_1}} = 96.7$ GeV, 50 events
- transfer function : delta except on muon energies (5 %)



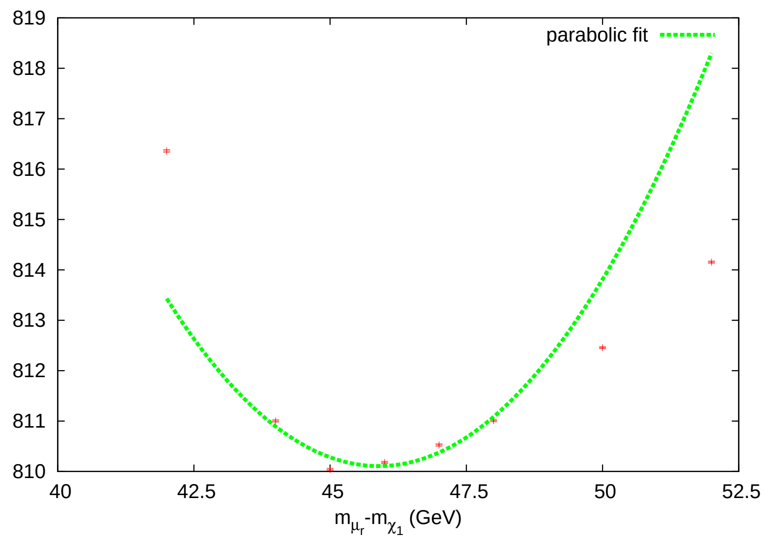
Applications

- Example 3 : slepton pair production



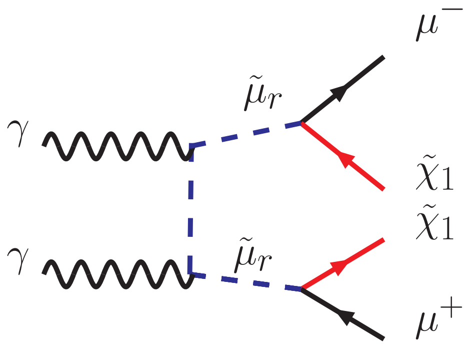
aim : measure of the masses $m_{\tilde{\mu_r}}$, $m_{\tilde{\chi_1}}$ from a pure signal sample

- inputs : $m_{\tilde{\mu_r}} = 144$ GeV, $m_{\tilde{\chi_1}} = 96.7$ GeV, 50 events
- transfer function : delta except on muon energies (5 %)
- $\Delta = 45.9 \pm 1.5 GeV$



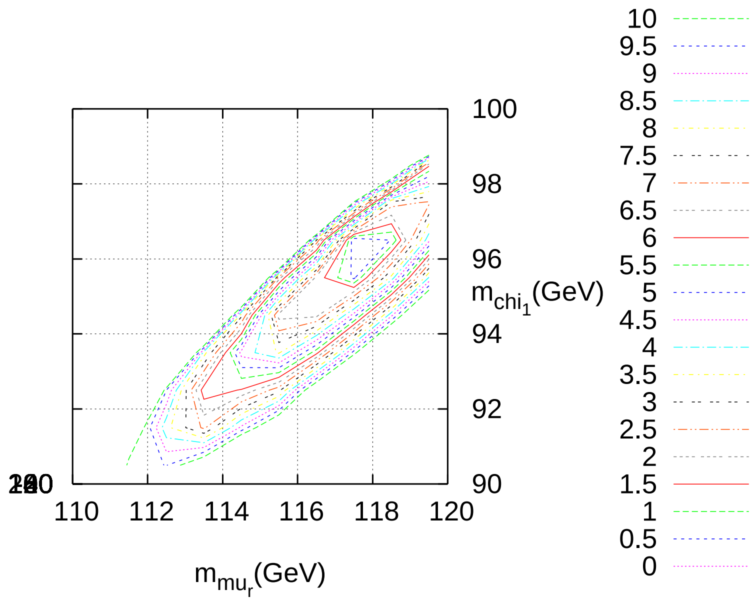
Applications

- Example 4 : slepton photo-production (with Nicolas Schul UCL-CP3)



aim : measure of the masses $m_{\tilde{\mu}_r}$, $m_{\tilde{\chi}_1}$ from a pure signal sample

- inputs : $m_{\tilde{\mu}_r} = 118$ GeV, $m_{\tilde{\chi}_1} = 97$ GeV, 20 events
- transfer function : delta except on muon energies (5 %)



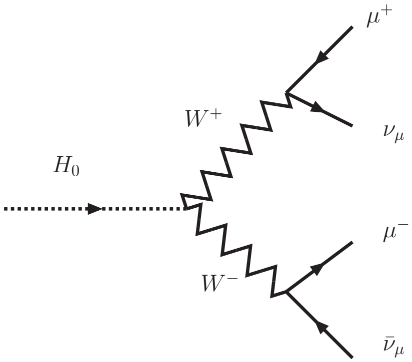
Conclusion

- the Matrix Element method provides the best discriminator on an event-by-event basis
- both theoretical ($|M|^2$) and experimental ($\boldsymbol{x}$, $W(\boldsymbol{x}, \boldsymbol{y})$) information is used
- the computation of the weights requires a specific phase space generator
- Given a arbitrary decay chain and the resolution function, MadWeight
  - determines automatically the best phase-space parametrization(s)
  - computes the weights for each experimental event
- code available on madgraph.phys.ucl.ac.be (on the download page)

Back-up slides

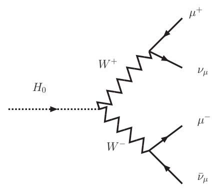
Example of Application (II)

- Higgs mass analysis

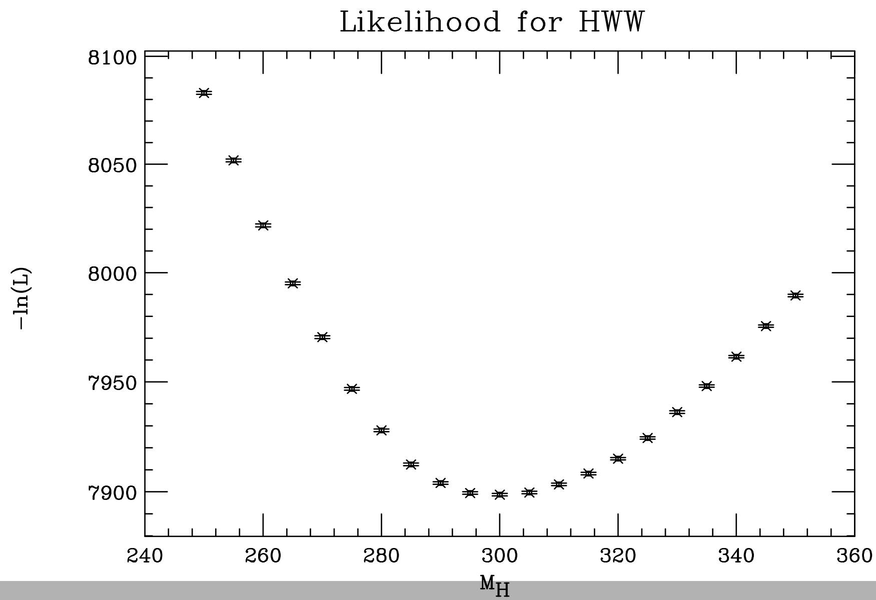


Example of Application (II)

- Higgs mass analysis



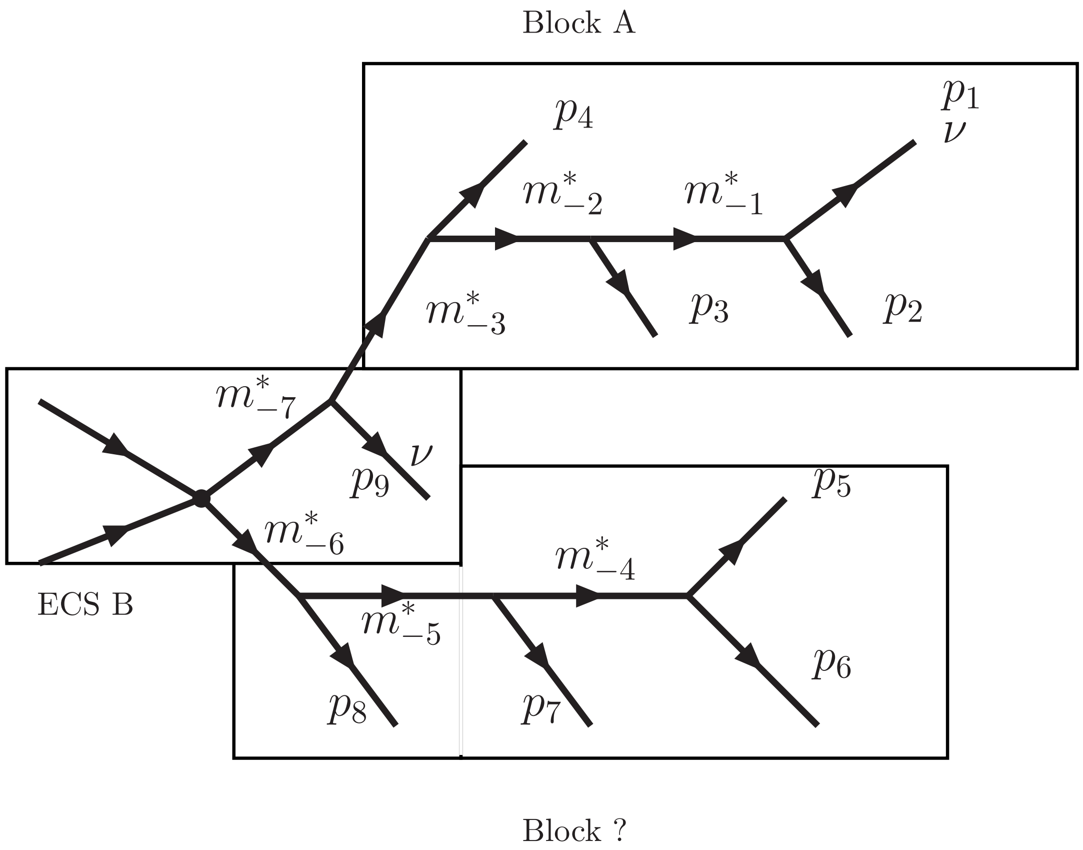
- 500 Monte Carlo events (MadGraph/Pythia/PGS)
- input : $m_{Higgs} = 300$ GeV, output : $m_{Higgs} = 300.9 \pm 3.0$ GeV



Topology identification

- input
  - Transfert function : specify if particle have delta-narrow-large resolution on Energy
  - Parameter : charge the width of all propagator
  - Feynman Diagram : (come from MG)
- choose ECS
  - rule 1a : minimize the content of the Black Box
  - rule 1b : minimize the number of quantities generated randomly on blob
  - systematic comparaison between all type of ECS

Topology identification



Topology identification

- Blob resolution
  - rule 1 : Particle with thin Transfer functions must be generated by TF
  - rule 2 : Absolute priority of neutrino Block A-B-C (in this order)
  - rule 3 : Produce Multi solution for E-D block. Block effenciency must depend trully on cinematics -> tricky

    1. Maximaze alignment of propagator : priority in D-E, use fuse
    2. No additional change of variable
    3. Intermediate solution +local change : priority in D,Nothing, don't une fuse