ETH zürich

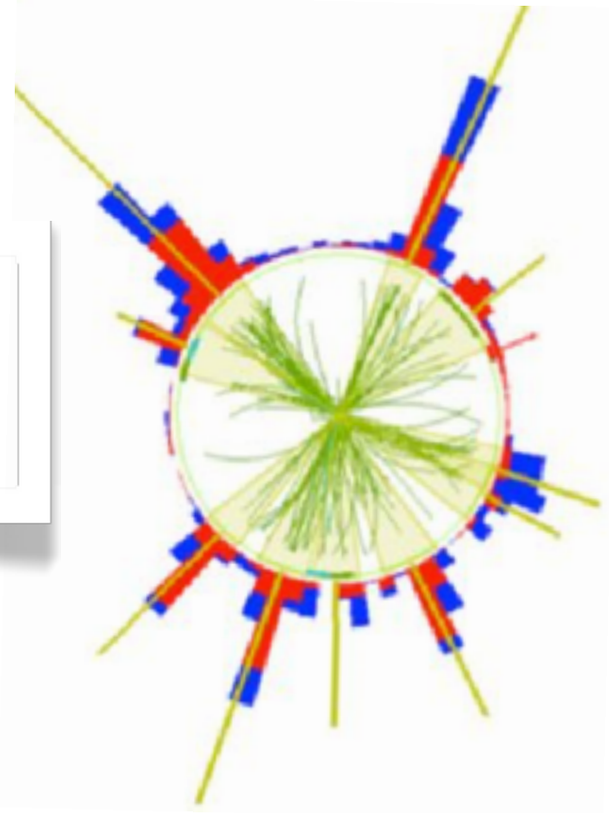
ETH Institute for Particle Physics

The MATRIX element method in HEP

Lorenzo Bianchini
ETH Zürich

A general problem

How can we maximally extract information
(on physical parameters θ) from an HEP event **Y***?*



- In HEP, dim(**Y**) » 1*
  - need data reduction with no loss of information
    - find a statistic T = T(**Y**) that is *sufficient* for θ
  - "The Maximum Likelihood estimator is asymptotically sufficient for θ"

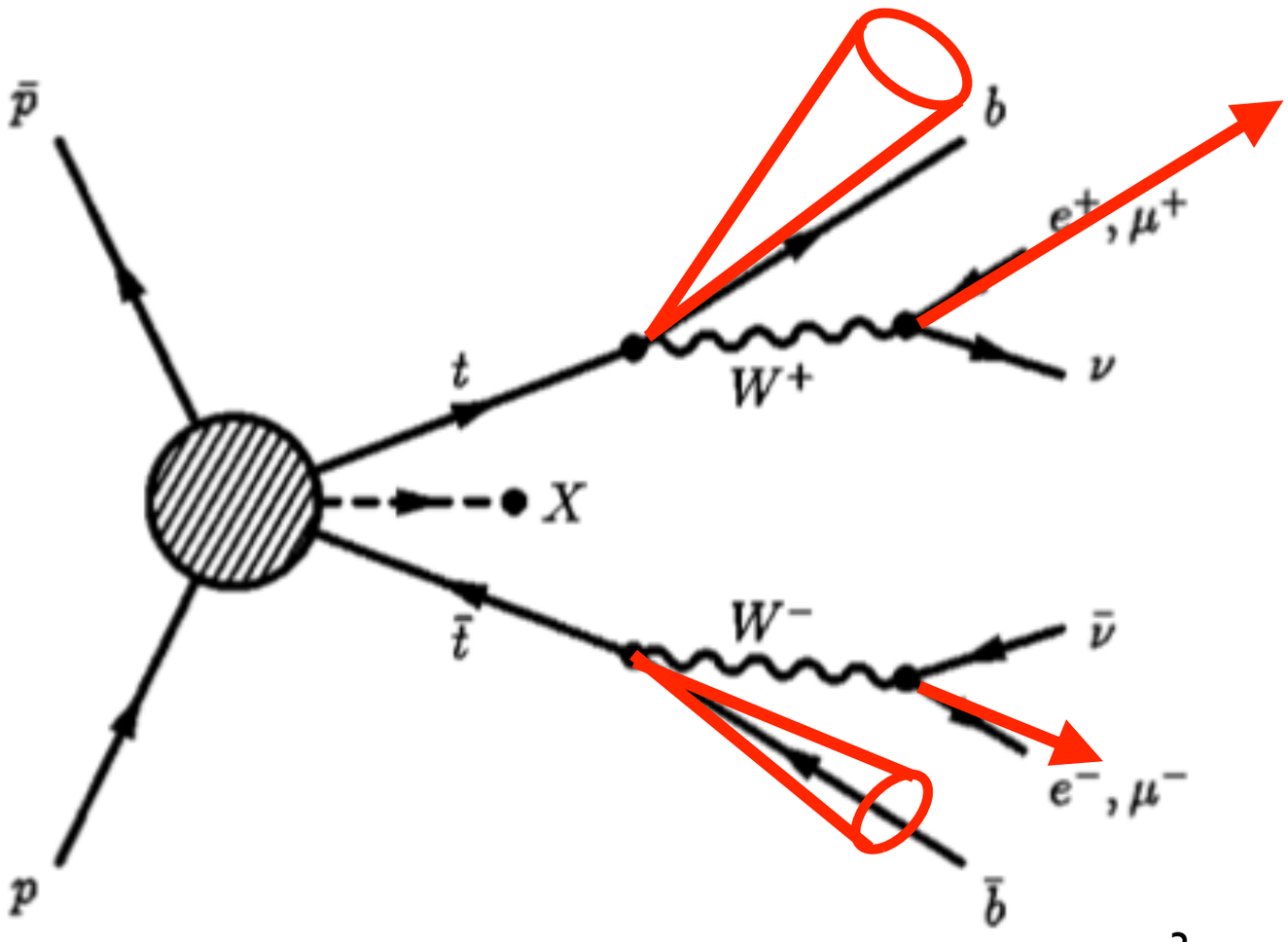
- How to derive the likelihood of an event ?
  - if dim(**Y**) small: templates using MC simulation
  - otherwise:
    - machine learning **
    - analytical approach

(*) Vincenzo's talk
(**) Mauro's and Tobias' talks

Canonical example: top mass

Y = 4-vectors of jets and leptons

M_t = top quark mass

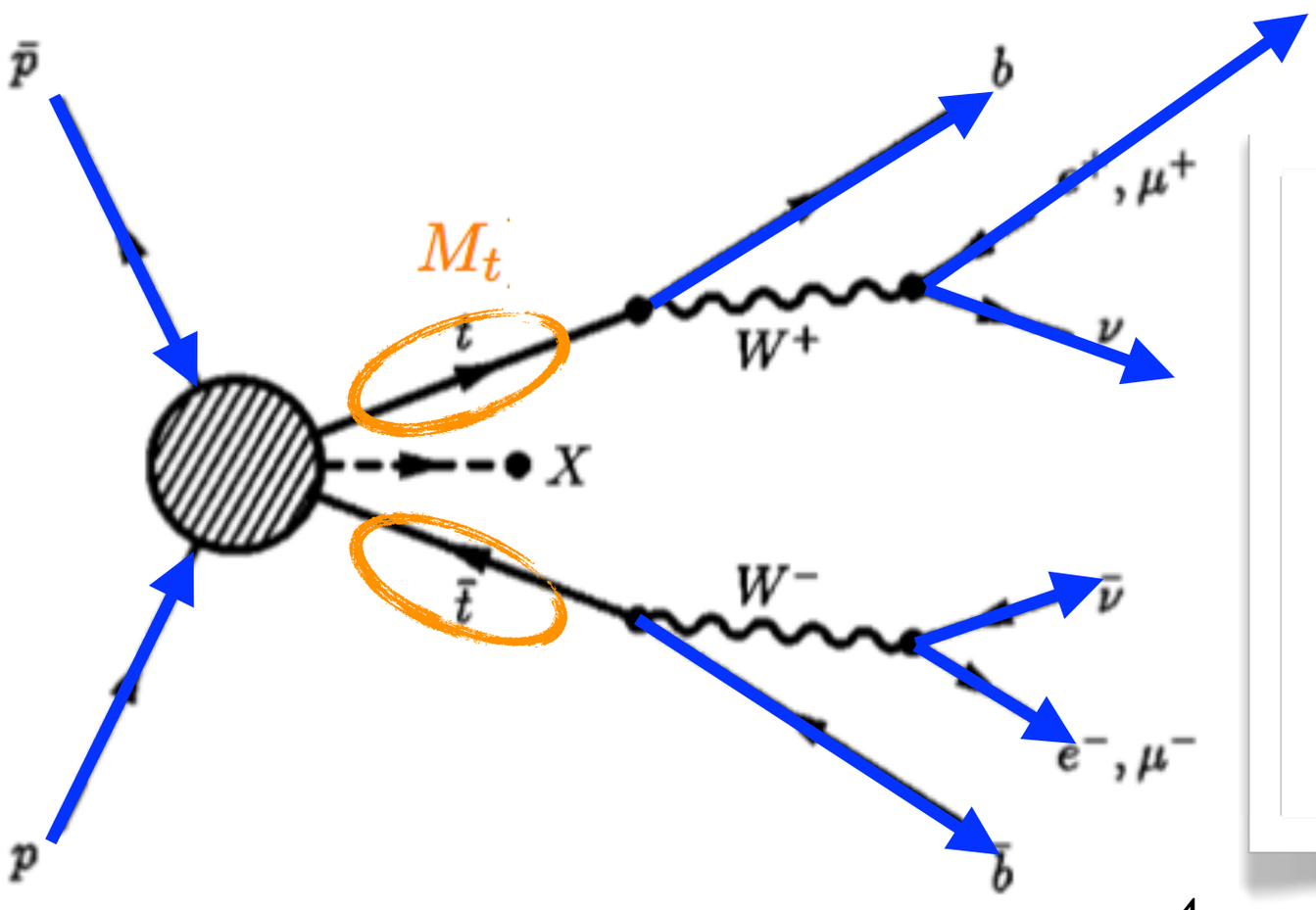


Canonical example: top mass

Y = 4-vectors of jets and leptons

M_t = top quark mass

X = 4-vectors of the 2 → 6 scattering (ancillary variables)



$$dP_X(M_t) \propto f_p(x_1) f_{\bar{p}}(x_2) \, |\mathcal{M}(X; M_t)|^2 \, dX$$

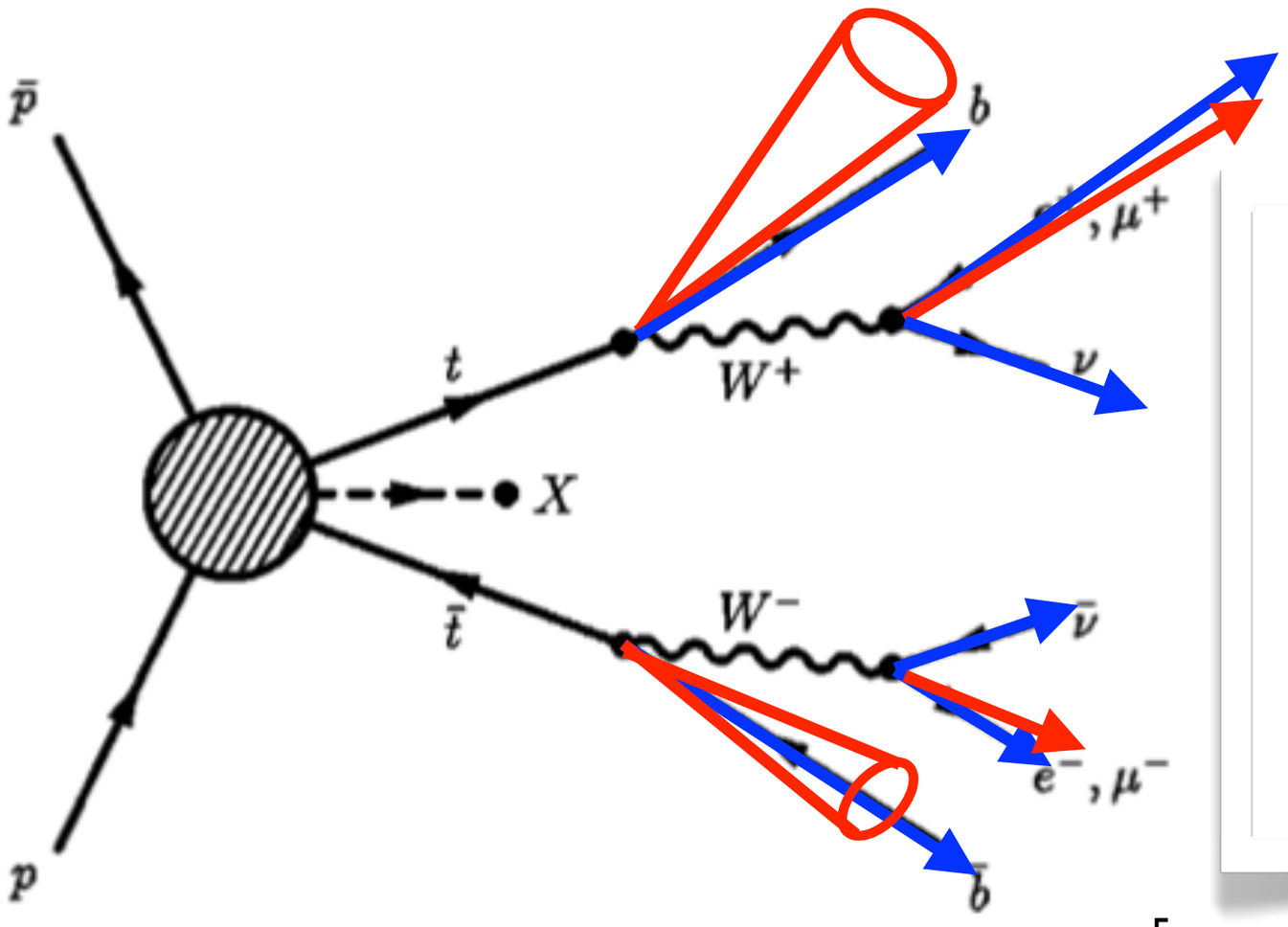
Canonical example: top mass

Y = 4-vectors of jets and leptons

M_t = top quark mass

X = 4-vectors of the 2 → 6 scattering (ancillary variables)

W(**Y**|**X**) = detector response

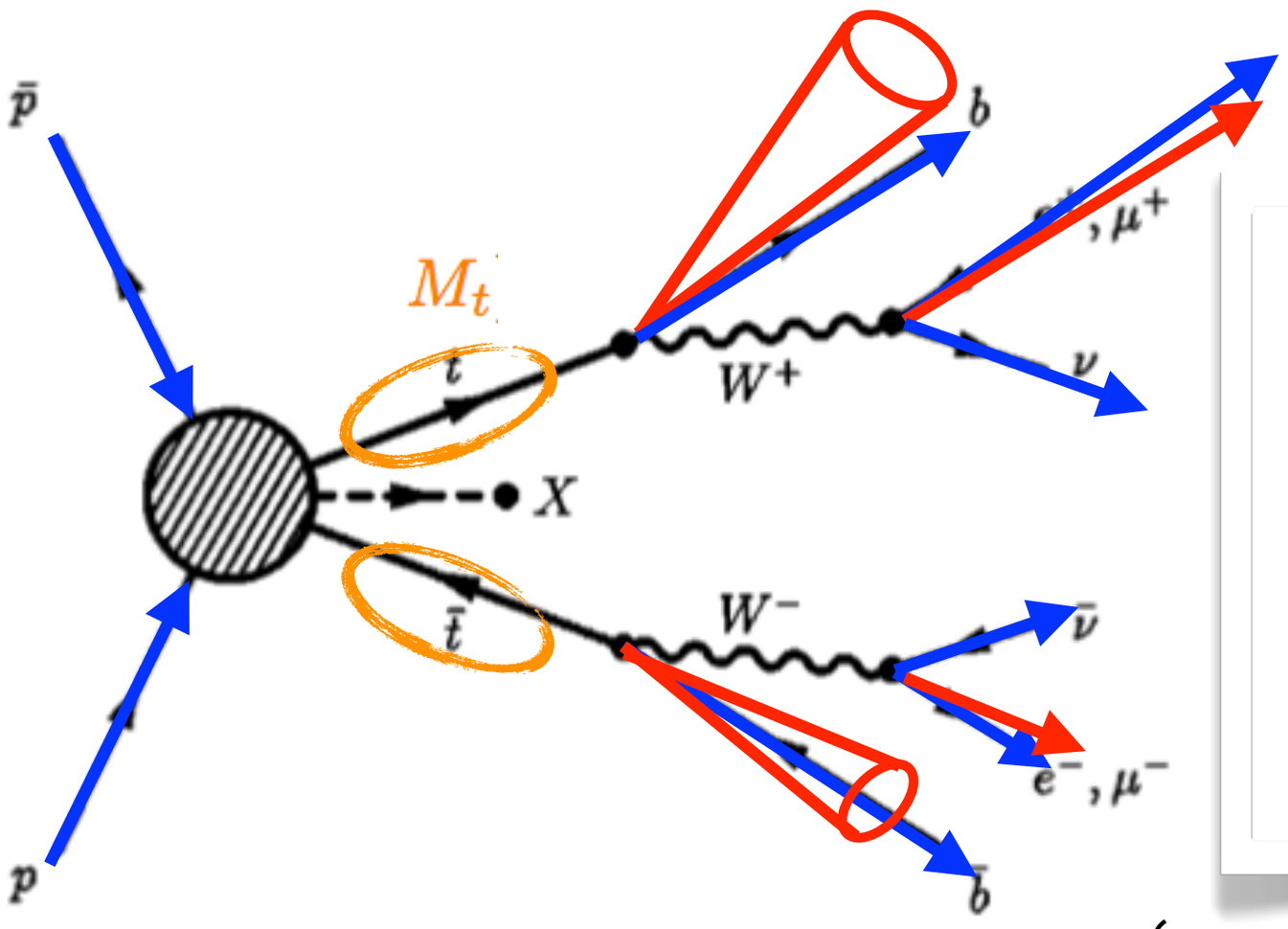


$$dP_X(M_t) \propto f_p(x_1) f_{\bar{p}}(x_2) \, |\mathcal{M}(X; M_t)|^2 \, dX$$

$$dP_{Y|X} = W(Y|X) dY$$

Canonical example: top mass

Y = 4-vectors of jets and leptons
M_t = top quark mass
X = 4-vectors of the 2 → 6 scattering (ancillary variables)
W(**Y**|**X**) = detector response



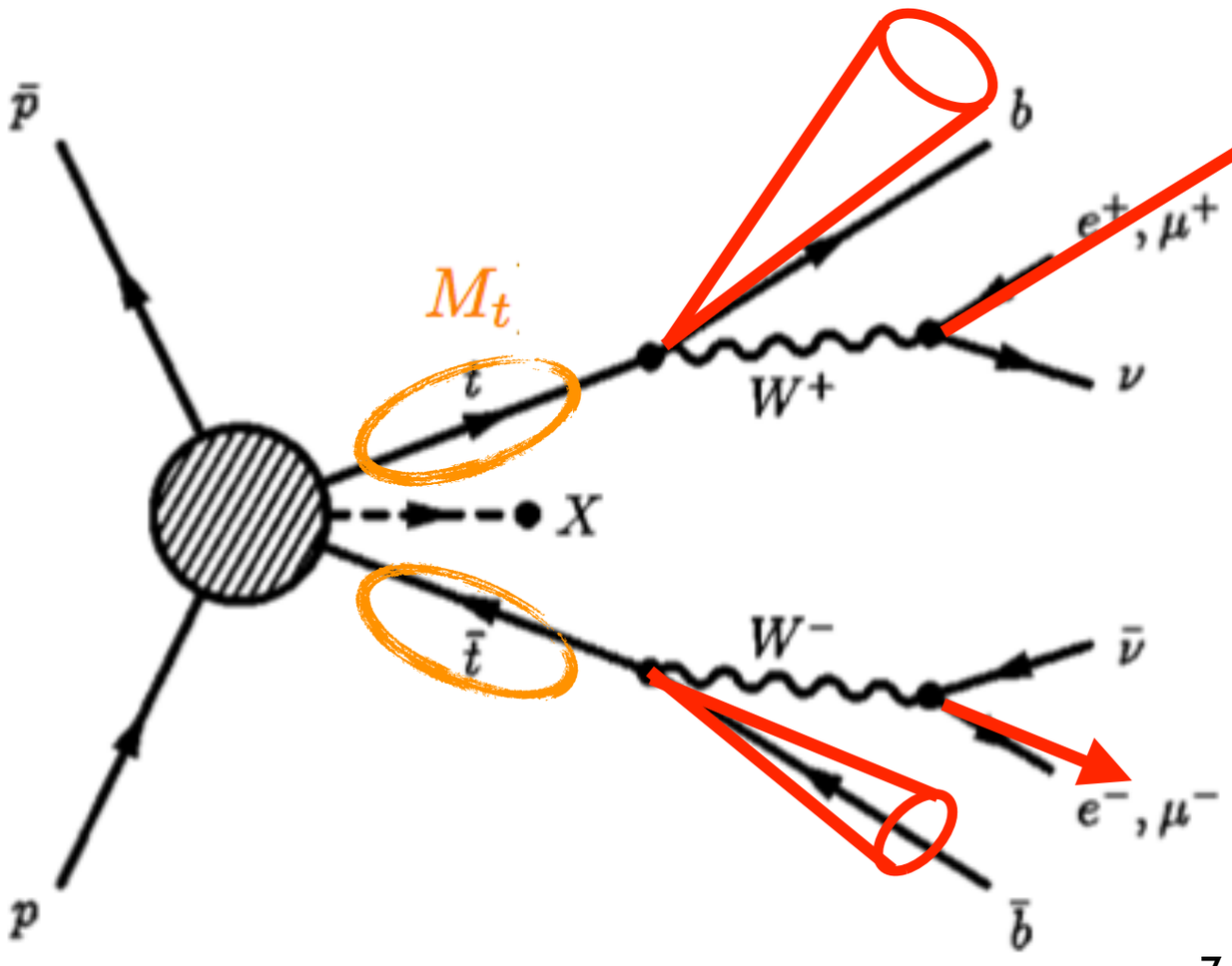
$$dP_X(M_t) \propto f_p(x_1) f_{\bar{p}}(x_2) \, |\mathcal{M}(X; M_t)|^2 \, dX$$

$$dP_{Y|X} = W(Y|X) dY$$

$$dP_{X,Y} \propto dP_{Y|X} \times dP_X(M_t)$$

Canonical example: top mass

Y = 4-vectors of jets and leptons
M_t = top quark mass
X = 4-vectors of the 2 → 6 scattering (ancillary variables)
W(**Y**|**X**) = detector response



$$dP_X(M_t) \propto f_p(x_1) f_{\bar{p}}(x_2) \, |\mathcal{M}(X; M_t)|^2 \, dX$$

$$dP_{Y|X} = W(Y|X) dY$$

$$dP_{X,Y} \propto dP_{Y|X} \times dP_X(M_t)$$

$$dP_Y(M_t) = \left[\int dP_X(M_t) \times W(Y|X) \right] dY$$

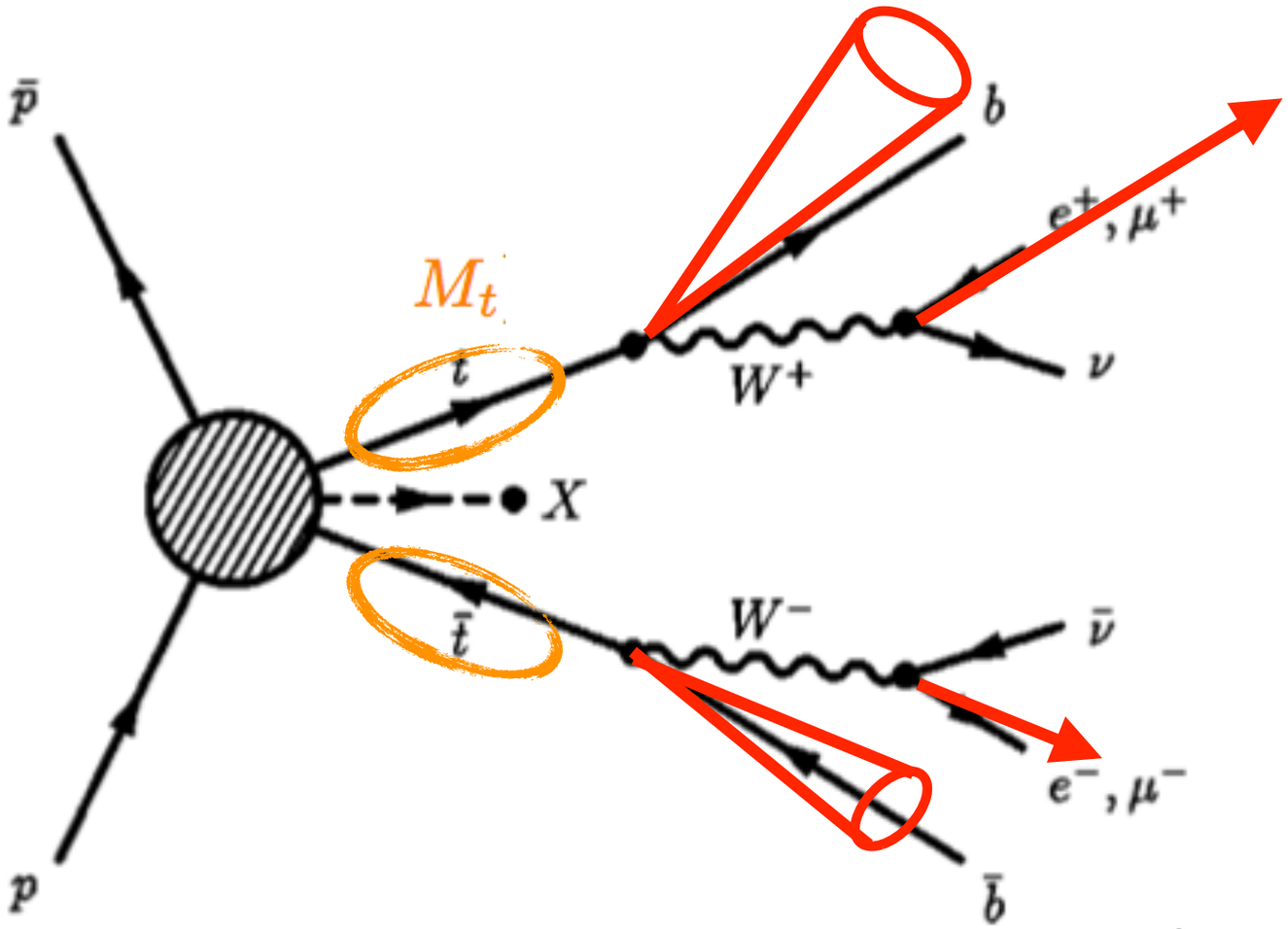
Canonical example: top mass

Y = 4-vectors of jets and leptons

M_t = top quark mass

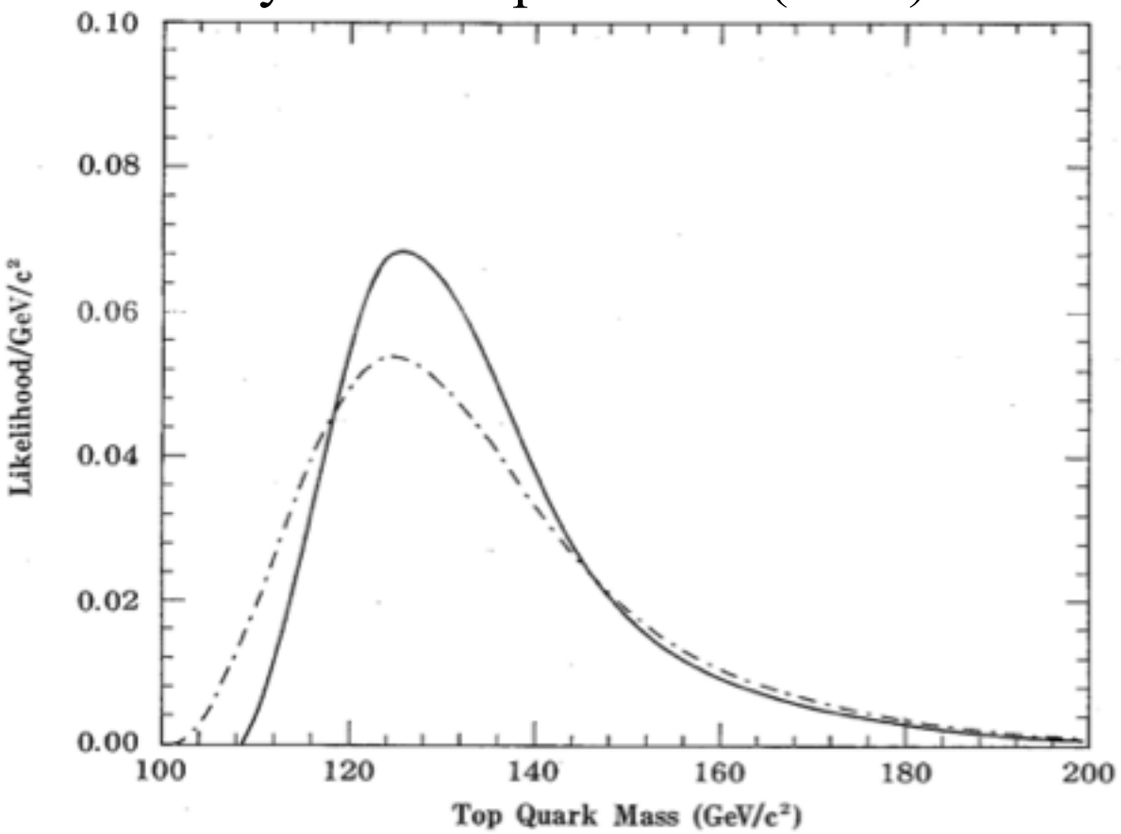
X = 4-vectors of the 2 → 6 scattering (ancillary variables)

W(**Y**|**X**) = detector response



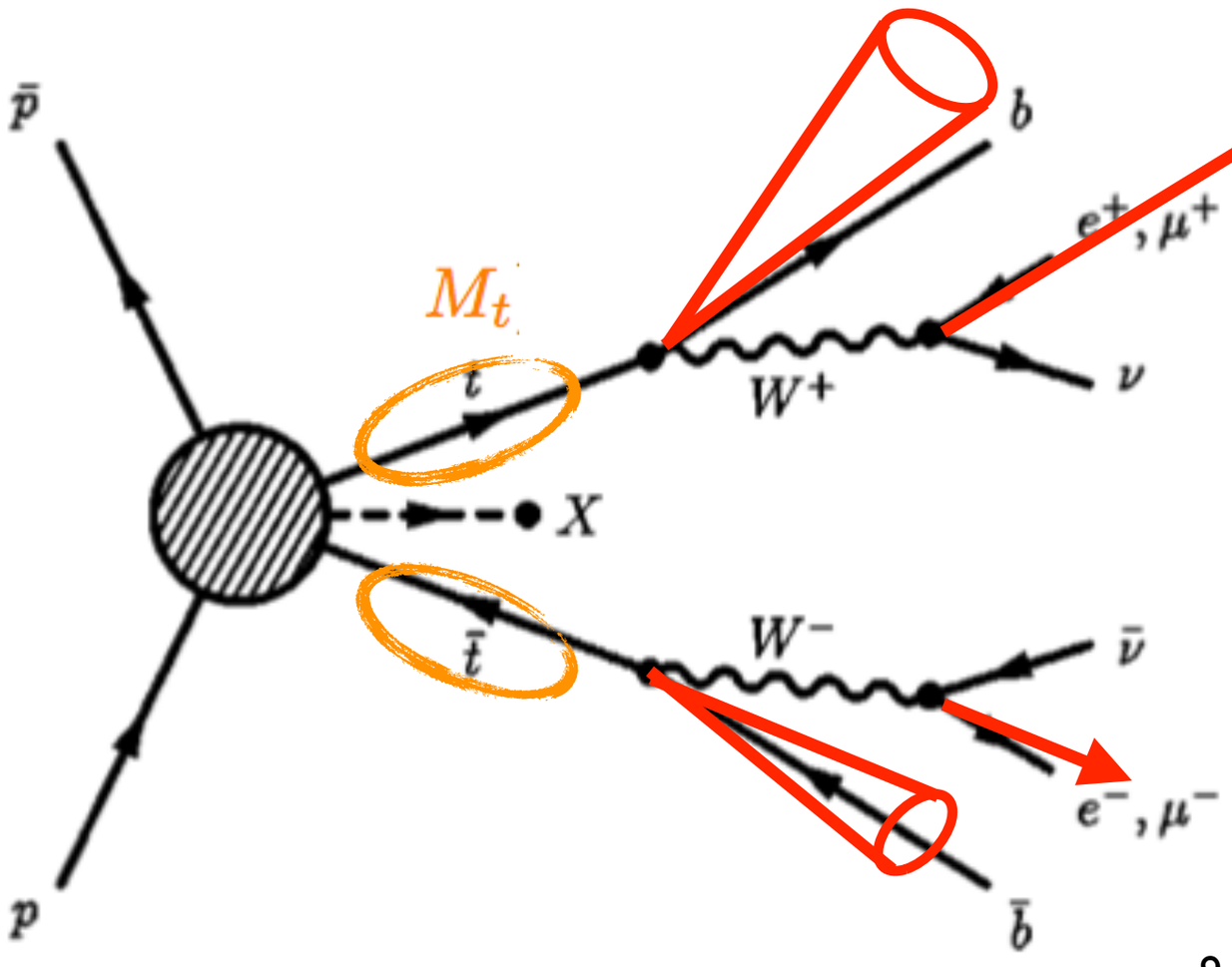
Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. III. Analysis of a CDF High P_T $e\mu$ Event as $t\bar{t}$ Production

K. Kondo et al.,
J. Phys. Soc. Jap 62 1177 (1993)



Canonical example: top mass

Y = 4-vectors of jets and leptons
M_t = top quark mass
X = 4-vectors of the 2 → 6 scattering (ancillary variables)
W(**Y**|**X**) = detector response



Matrix Element Method
(MEM)

$$dP_X(M_t) \propto f_p(x_1) f_{\bar{p}}(x_2) |\mathcal{M}(X; M_t)|^2 dX$$

$$dP_{Y|X} = W(Y|X) dY$$

$$dP_{X,Y} \propto dP_{Y|X} \times dP_X(M_t)$$

$$dP_Y(M_t) = \left[\int dP_X(M_t) \times W(Y|X) \right] dY$$

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y}$$

$$\int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y} \qquad \int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

- (normalised) differential cross section:

$$p_S(\vec{y}\,; \theta) = \frac{1}{\sigma_S(\theta)} \frac{d\sigma_S}{d\vec{y}}(\vec{y}\,; \theta)$$

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y} \qquad \int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

- (normalised) differential cross section:

$$p_S(\vec{y}\,; \theta) = \frac{1}{\sigma_S(\theta)} \frac{d\sigma_S}{d\vec{y}}(\vec{y}\,; \theta)$$

$$d\sigma_S(\vec{y}\,; \theta) = \left[\int d\Phi(\vec{x}) dx_\mathrm{a} dx_\mathrm{b} \sum_{i,j} \frac{f_i(x_\mathrm{a}) f_j(x_\mathrm{b})}{(1+\delta_{ij}) x_\mathrm{a} x_\mathrm{b} s} |\mathcal{M}_S(\vec{x}, \theta)|^2 W(\vec{y}, \vec{x}; \theta) \right] d\vec{y}$$

$$\prod_i \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i}$$

Numerical integration

Parton density functions

Scattering amplitude

Detector transfer function

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y} \qquad \int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

- (normalised) differential cross section:

$$p_S(\vec{y}\,; \theta) = \frac{1}{\sigma_S(\theta)} \frac{d\sigma_S}{d\vec{y}}(\vec{y}\,; \theta)$$

$$d\sigma_S(\vec{y}\,; \theta) = \left[\int d\Phi(\vec{x}) dx_\mathrm{a} dx_\mathrm{b} \sum_{i,j} \frac{f_i(x_\mathrm{a}) f_j(x_\mathrm{b})}{(1+\delta_{ij}) x_\mathrm{a} x_\mathrm{b} s} |\mathcal{M}_S(\vec{x}, \theta)|^2 W(\vec{y}, \vec{x}; \theta) \right] d\vec{y}$$

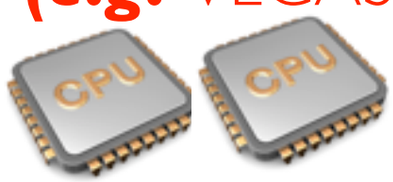
$$\prod_i \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i}$$

Numerical integration (e.g. VEGAS)

Parton density functions (e.g. LHAPDF)

Scattering amplitude (e.g. MCFM, MG5)

Detector transfer function (MC simulation)



[...]

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y} \qquad \int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

- (normalised) differential cross section:

$$p_S(\vec{y}\,; \theta) = \frac{1}{\sigma_S(\theta)} \frac{d\sigma_S}{d\vec{y}}(\vec{y}\,; \theta)$$

$$d\sigma_S(\vec{y}\,; \theta) = \left[\int d\Phi(\vec{x}) dx_\mathrm{a} dx_\mathrm{b} \sum_{i,j} \frac{f_i(x_\mathrm{a}) f_j(x_\mathrm{b})}{(1+\delta_{ij}) x_\mathrm{a} x_\mathrm{b} s} |\mathcal{M}_S(\vec{x}, \theta)|^2 W(\vec{y}, \vec{x}; \theta) \right] d\vec{y}$$

- multi-channel processes:

$$p(\vec{y}\,; \theta) = \lambda_S(\theta) p_S(\vec{y}\,; \theta) + \sum_j \lambda_{Bj} p_{Bj}(\vec{y})$$

General formulation of the MEM

$$\text{Prob}\{\vec{y} \in [\vec{y}_0, \vec{y}_0 + d\vec{y}] \mid S, \theta\} = p_S(\vec{y}_0; \theta) d\vec{y} \qquad \int_{\mathcal{A}} p_S(\vec{y}\,; \theta) d\vec{y} = 1$$

- (normalised) differential cross section:

$$p_S(\vec{y}\,; \theta) = \frac{1}{\sigma_S(\theta)} \frac{d\sigma_S}{d\vec{y}}(\vec{y}\,; \theta)$$

$$d\sigma_S(\vec{y}\,; \theta) = \left[\int d\Phi(\vec{x}) dx_\mathrm{a} dx_\mathrm{b} \sum_{i,j} \frac{f_i(x_\mathrm{a}) f_j(x_\mathrm{b})}{(1+\delta_{ij}) x_\mathrm{a} x_\mathrm{b} s} |\mathcal{M}_S(\vec{x}, \theta)|^2 W(\vec{y}, \vec{x}; \theta) \right] d\vec{y}$$

- multi-channel processes:

$$p(\vec{y}\,; \theta) = \lambda_S(\theta) p_S(\vec{y}\,; \theta) + \sum_j \lambda_{Bj} p_{Bj}(\vec{y})$$

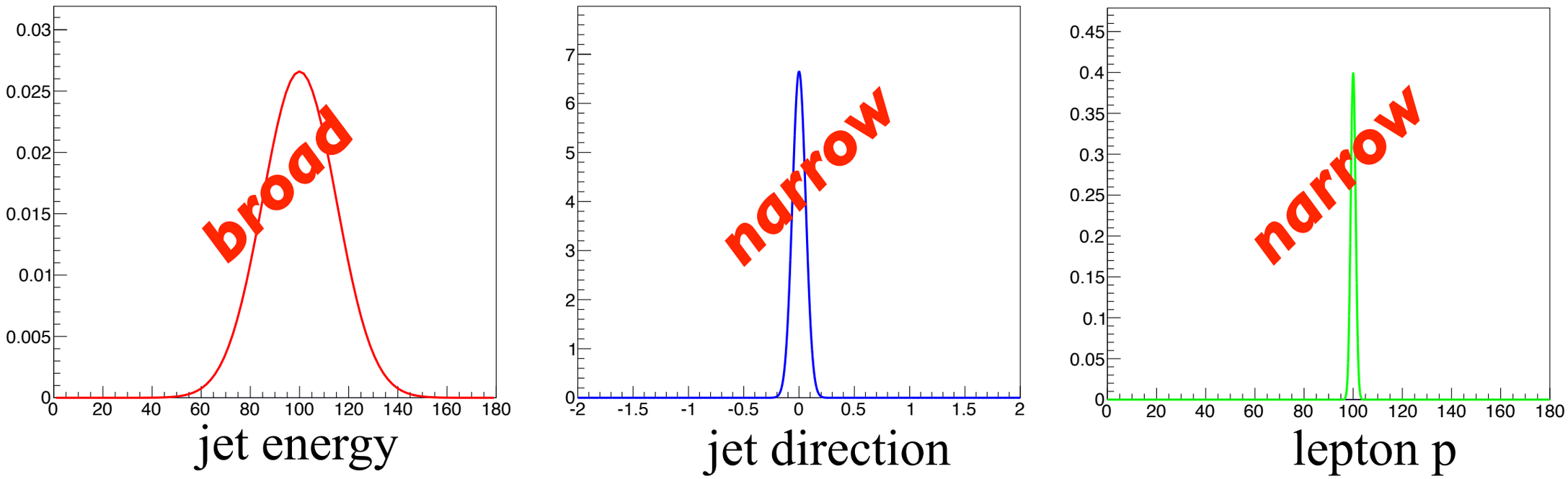
- sample likelihood:

$$\mathcal{L}(\vec{y}\,; \theta) = \left[\prod_{i=1}^{N} p(\vec{y}_i\,; \theta) \right] \frac{e^{-\mathcal{N}(\theta)} \mathcal{N}(\theta)^N}{N!}$$

Transfer function

- Factorized
  - normalisation condition
  - broad vs narrow:

$$W(\vec{y},\vec{x}) \approx \prod_i W(y_i, x_i)$$



- Particle ⟺ detector association not always univocal !!!
  - need to permute all possible associations and sum them up

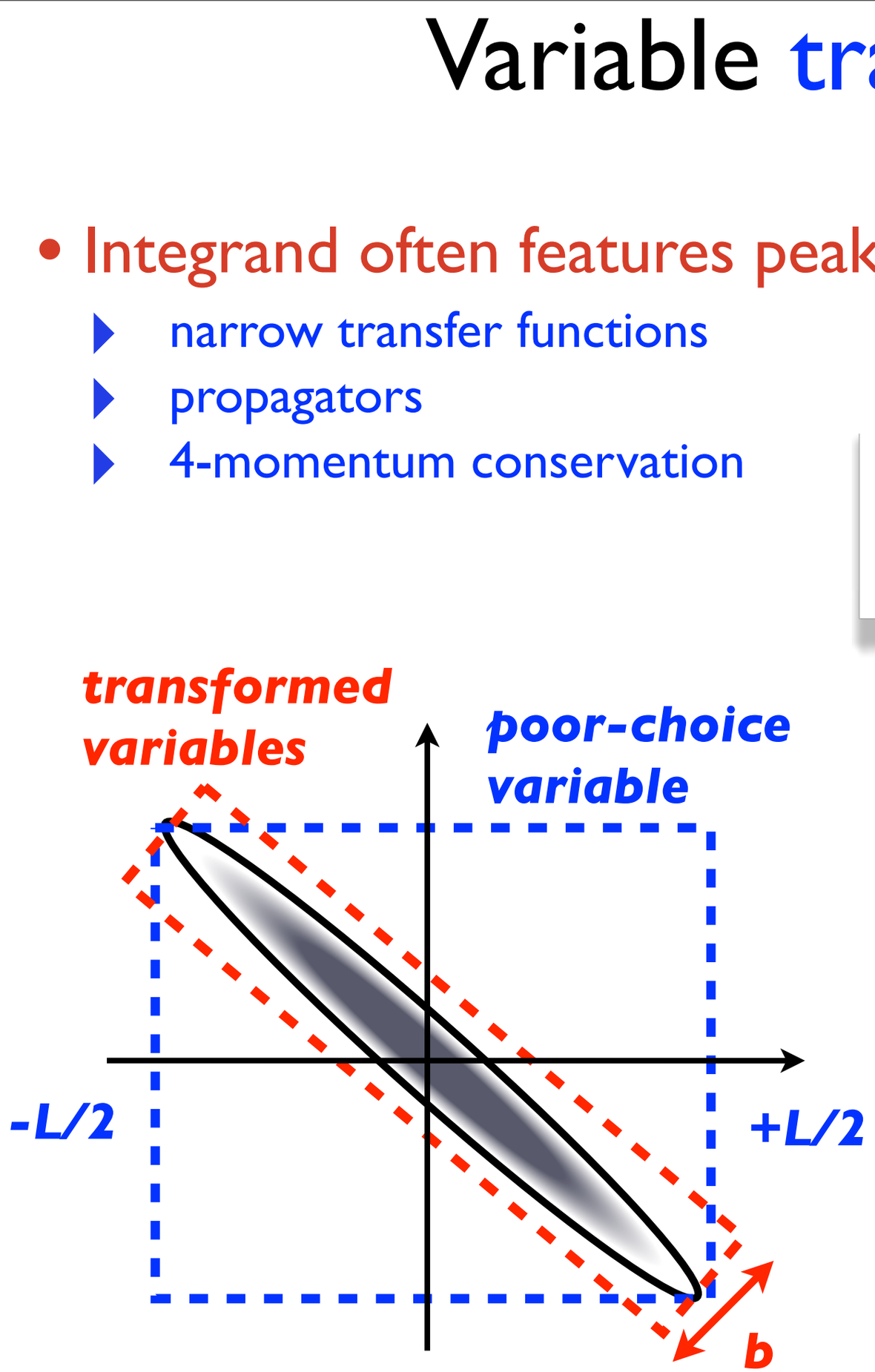
F. Fiedler et al., NIM A 624 (2010) 203
I. Voloubouev, arXiv:1101.2259

Variable transformation

- Integrand often features peaks
  - narrow transfer functions
  - propagators
  - 4-momentum conservation

$$\frac{1}{(s^2-M^2)^2+\Gamma^2M^2} \to \frac{\pi}{M\Gamma}\delta(s^2-M^2)$$

$$d\Phi(\vec{x}) = (2\pi)^4\delta(x_a P_a + x_b P_b - \sum_i p_i)\left[\prod_i \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i}\right]$$



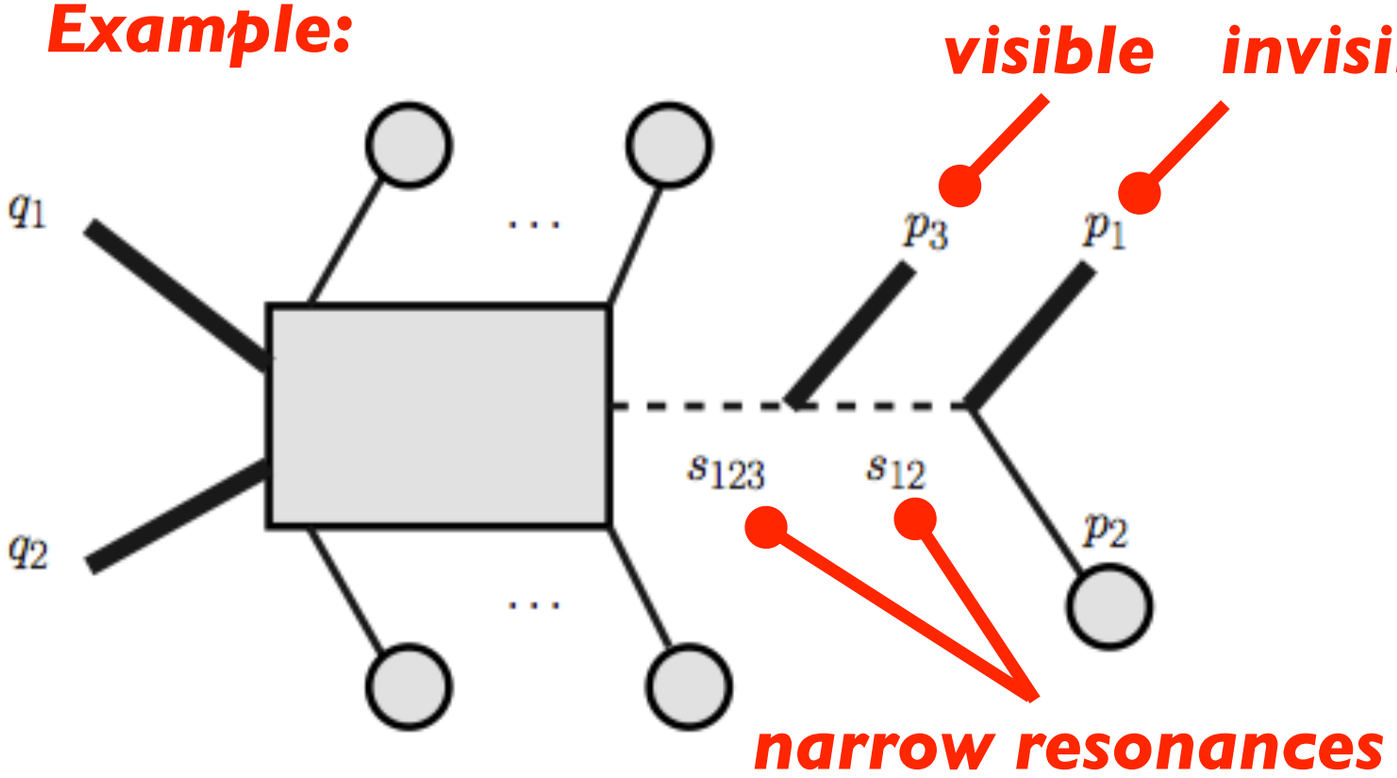
- Align integration variables along peaking directions

Automation of the MEM*

Variable alignment can be automatised (MadWeight, BEM)

- variable transformation for each topology
- solve system of equations
- determine Jacobian factor

Example:



F. Maltoni et al., JHEP 10 (2012) 068

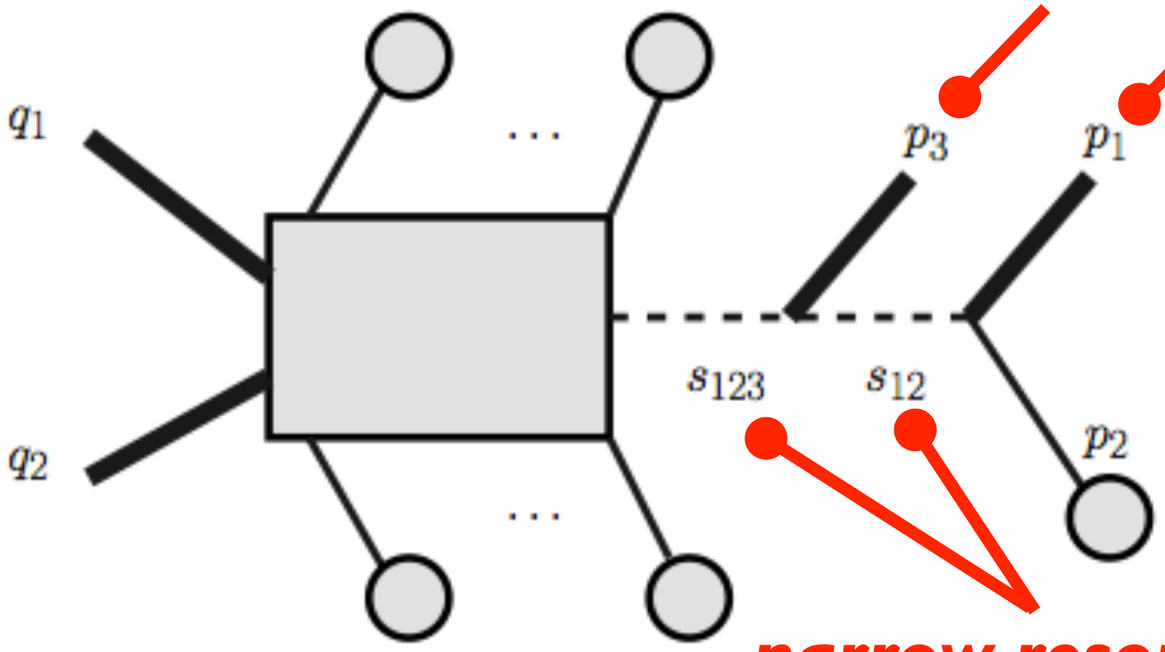
(*) this afternoon tutorial

Automation of the MEM*

Variable alignment can be automatised (MadWeight, BEM)

- variable transformation for each topology
- solve system of equations
- determine Jacobian factor

Example:



$$J = \sin\theta_3 \frac{E_3^2 E_1}{s} \Big| \chi E_2 p_{1z} - \chi E_1 p_{2z} -$$
$$2\cos(\phi_3)\cos(\theta_3) E_2 p_{1x} E_3 \sin(\theta_3) + 2\cos(\phi_3)\cos(\theta_3) E_1 p_{2x} E_3 \sin(\theta_3) -$$
$$2\cos(\phi_3) p_{1z} p_{2x} E_3 \sin(\theta_3) + 2\cos(\phi_3) p_{1x} p_{2z} E_3 \sin(\theta_3) -$$
$$2\cos(\theta_3) E_2 p_{1y} E_3 \sin(\phi_3)\sin(\theta_3) + 2\cos(\theta_3) E_1 p_{2y} E_3 \sin(\phi_3)\sin(\theta_3) -$$
$$2 p_{1z} p_{2y} E_3 \sin(\phi_3)\sin(\theta_3) + 2 p_{1y} p_{2z} E_3 \sin(\phi_3)\sin(\theta_3) +$$
$$2\cos(\phi_3)^2 E_2 p_{1z} E_3 \sin(\theta_3)^2 - 2\cos(\phi_3)^2 E_1 p_{2z} E_3 \sin(\theta_3)^2 +$$
$$2 E_2 p_{1z} E_3 \sin(\phi_3)^2 \sin(\theta_3)^2 - 2 E_1 p_{2z} E_3 \sin(\phi_3)^2 \sin(\theta_3)^2 \Big|^{-1},$$

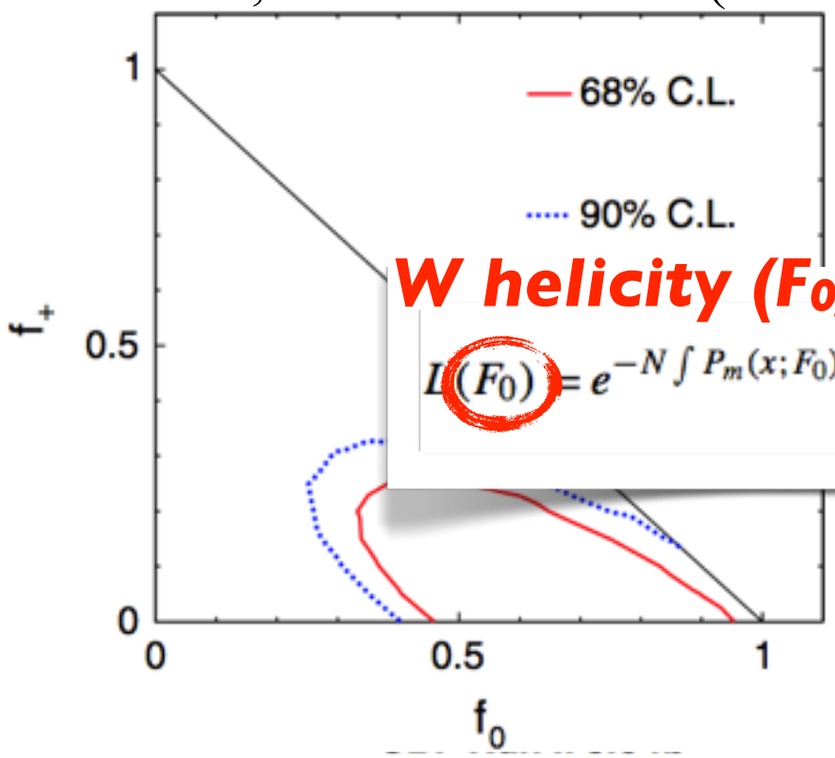
$$dq_1 dq_2 \frac{d^3p_1}{(2\pi)^3 2E_1} \frac{d^3p_3}{(2\pi)^3 2E_3} (2\pi)^4 \delta^4 (P_{\text{in}} - P_{\text{fin}}) \Rightarrow \frac{1}{16\pi^2 E_1 E_3} d\phi_3 d\theta_3 ds_{12} ds_{123} \times J$$

F. Maltoni et al., JHEP 10 (2012) 068

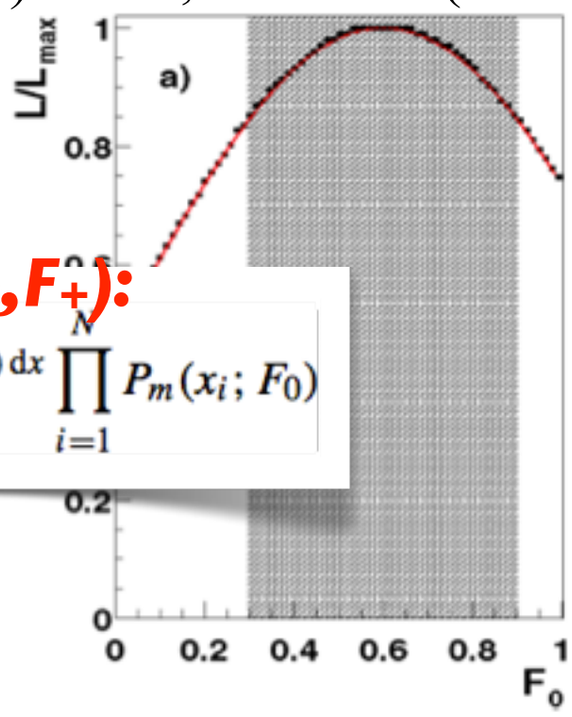
(*) this afternoon tutorial

MEM for point estimation

CDF, PRL 105 042002 (2010)



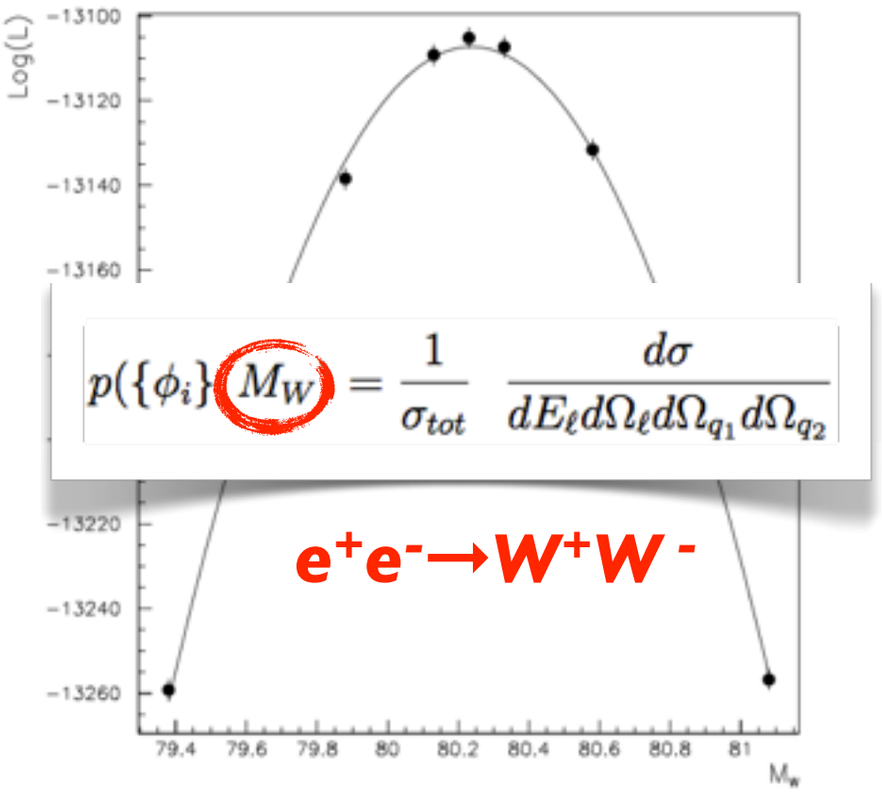
D0, PLB 617 (2005) 1



W helicity (F_0,F_+):

$$L(F_0) = e^{-N\int P_m(x;F_0)\,dx}\prod_{i=1}^{N} P_m(x_i;F_0)$$

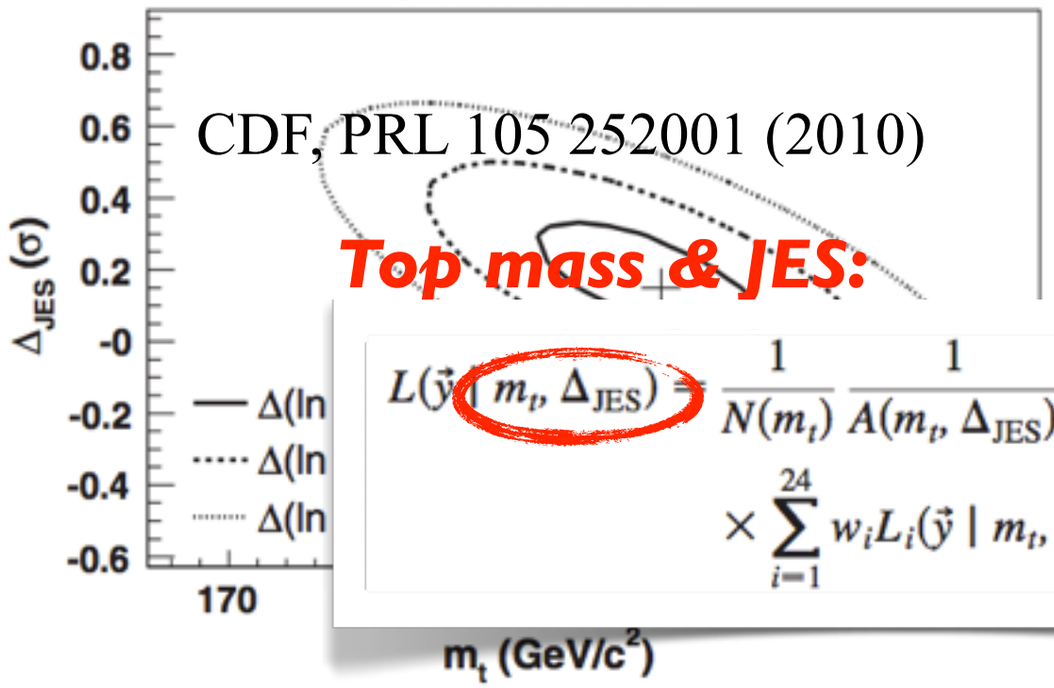
Berends et al. PRL B417 (1998) 385



$$p(\{\phi_i\}|M_W) = \frac{1}{\sigma_{tot}}\frac{d\sigma}{dE_\ell d\Omega_\ell d\Omega_{q_1} d\Omega_{q_2}}$$

$e^+e^- \rightarrow W^+W^-$

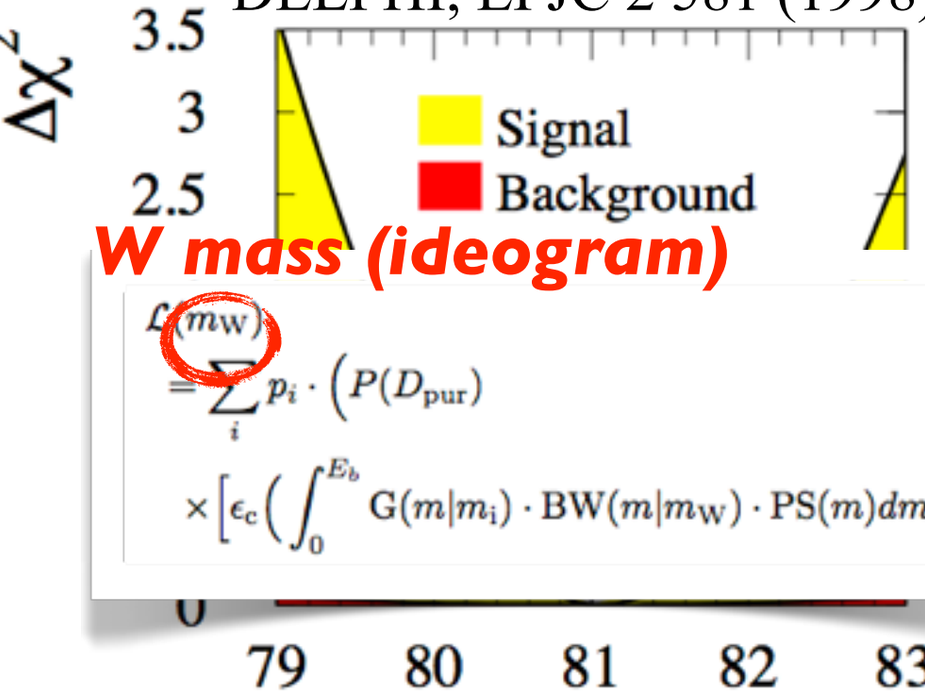
CDF, PRL 105 252001 (2010)



Top mass & JES:

$$L(\vec{y}\,|\,m_t,\Delta_{JES}) = \frac{1}{N(m_t)}\frac{1}{A(m_t,\Delta_{JES})} \times \sum_{i=1}^{24} w_i L_i(\vec{y}\,|\,m_t,\Delta_{JES}) L_i(\vec{y}\,|\,m_t,\Delta_{JES})$$

DELPHI, EPJC 2 581 (1998)



W mass (ideogram)

$$\mathcal{L}(m_W) = \sum_i p_i \cdot \Big(P(D_{pur}) \times \Big[\epsilon_c\Big(\int_0^{E_b} G(m|m_i)\cdot BW(m|m_W)\cdot PS(m)dm\Big)$$

A. Juste, PhD thesis (1998)
J. Estrada, FERMILAB-THESIS-2001-07 (2001)
F. Canelli, FERMILAB-THESIS-2003-22 (2003)

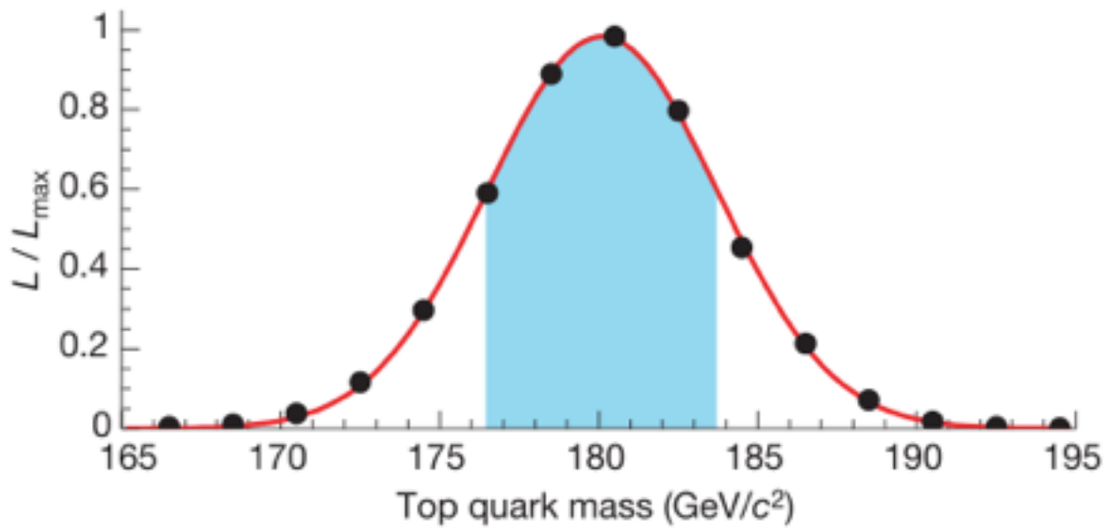
An efficient estimator

D0, PRD 58 052001 (1998)

> We determine the top quark mass m_t using tt pairs produced in the DØ detector by $\sqrt{s}=1.8$ TeV pp collisions in a 125 pb^{-1} exposure at the Fermilab Tevatron. We make a two constraint fit to m_t in $tt\to bW^+bW^-$ final states with one W boson decaying to qq and the other to $e\nu$ or $\mu\nu$. Likelihood fits to the data yield $m_t(l+\text{jets})=173.3\pm5.6\,(\text{stat})\pm5.5\,(\text{syst})$ GeV/c^2. When this result is combined with an analysis

D0, Nature 479 (2004)

> smaller, and discussed in detail elsewhere[21,22]. It should be noted that the new mass measurement method provides a significant (about 40%, from ±5.5 to ±3.9 GeV/c^2) reduction in systematic uncertainty, which is ultimately dominated by the measurement of jet energies. For details on the new analysis, see the Methods.
>
> The final result is $M_t = 180.1 \pm 3.6\,(\text{stat})\pm3.9\,(\text{syst})$ GeV/c^2. The improvement in statistical uncertainty over our previous measurement is equivalent to collecting a factor of 2.4 as much data. Combining the statistical and systematic uncertainties in



Main virtues of the MEM evident here:

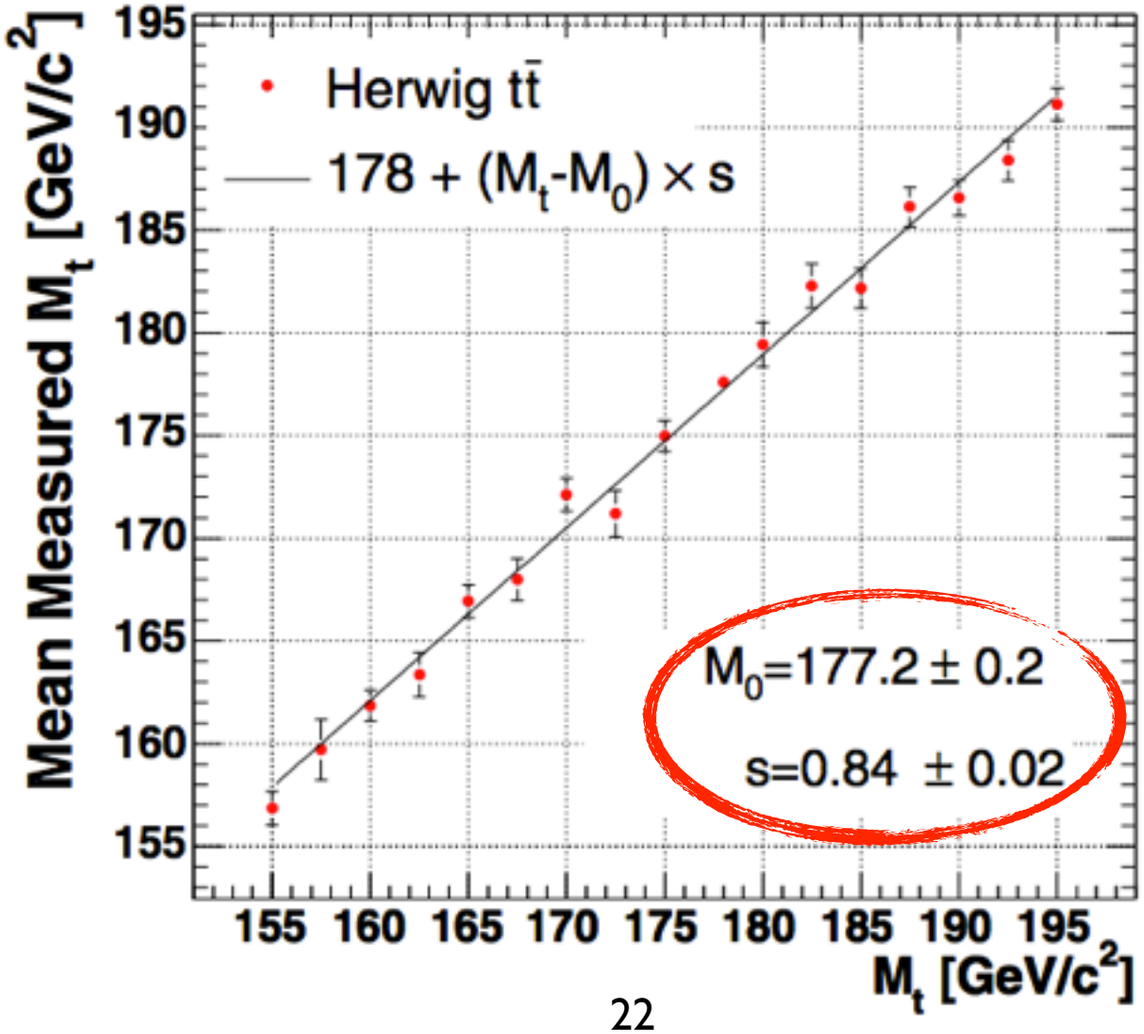
- optimise statistical power of available data
- systematic uncertainties (nuisance parameters) calculated event-by-event
  - each event contributes with optimal weight!

Linearity

Linearity test using MC simulation

- check for bias due to model approximations
- determine calibration function
  - approximations in MEM don't hamper its applicability, just make it less optimal!

CDF, PRD 74, 032009 (2006)

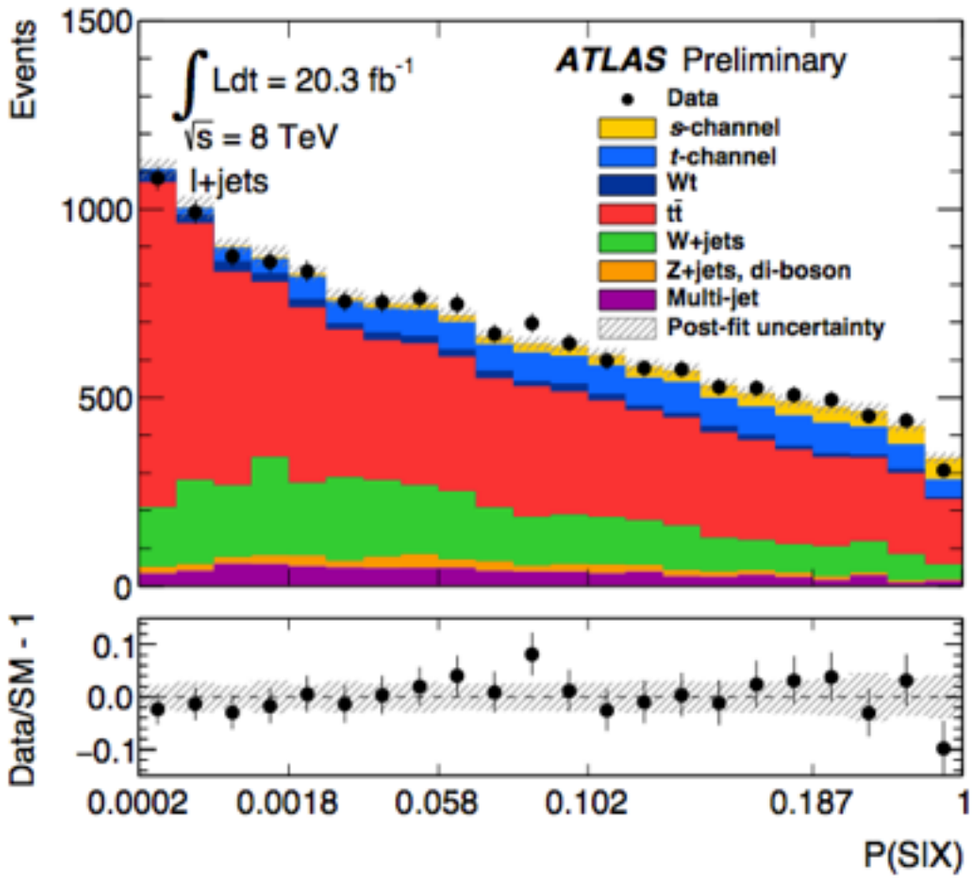


MEM for hypothesis testing

- Ratio of MEM probability densities is a powerful test statistic
  - optimal properties (largest power at fixed error) L. Neyman, Phil. Trans. Royal Soc. London 236 (1937) 333

$$\mathrm{LR} = \frac{\lambda_S p_S}{\sum_j \lambda_{B_j} p_{B_j}} \quad \text{or} \quad \frac{1}{1+\mathrm{LR}} \in [0,1]$$

ATLAS-CONF-2015-047

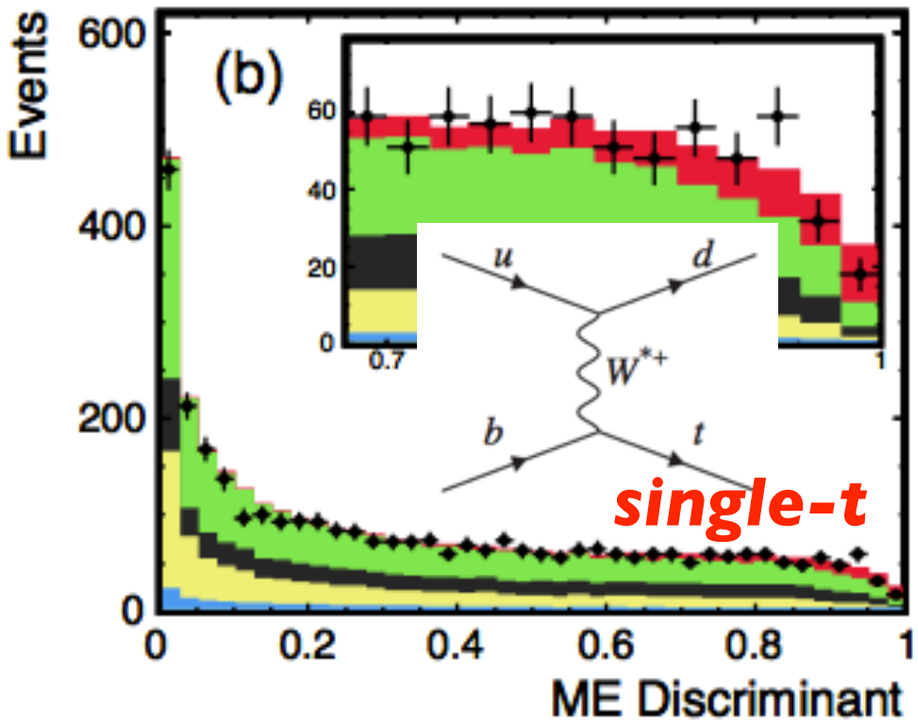


- No need of training sample
  - relevant for searches in rare topologies

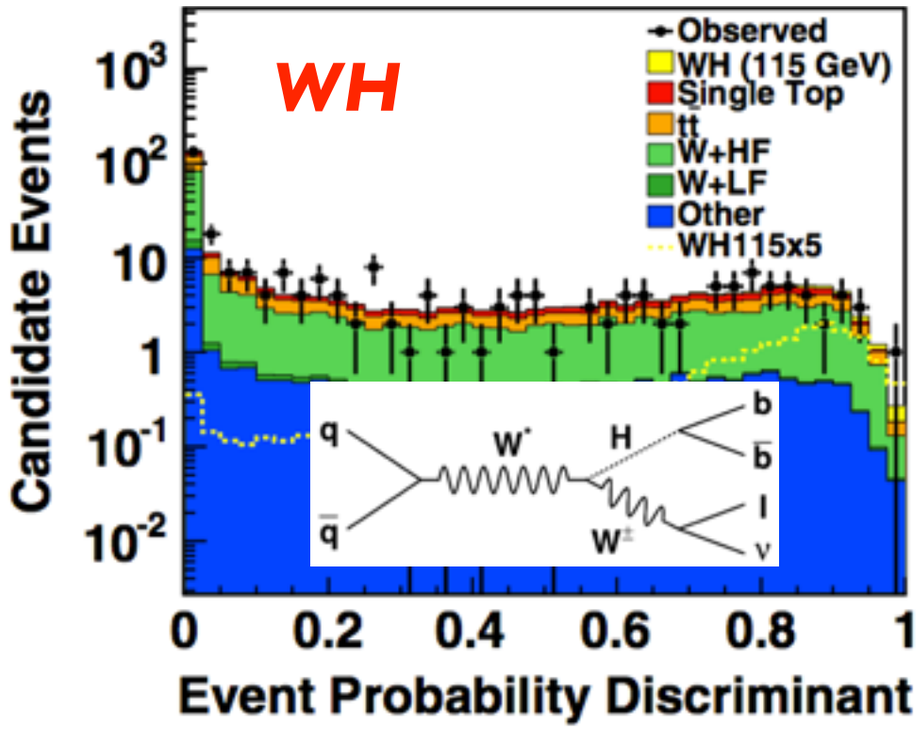
“In contrast to the aforementioned BDT-based analysis [11], the signal extraction in this analysis is based on the Matrix Element (ME) Method [18, 19]. The used data set is still the same, while some of the simulation samples are replaced by enhanced ones which reduces the uncertainty due to the simulation statistics and gives a better description of the data. Furthermore, updated calibrations for the 2012 data are used, resulting in a reduction of systematic uncertainties. The event selection is improved by adding a veto on di-leptonic events which leads to a significant suppression of the background for top-quark pair ($t\bar{t}$) production (c. f. Sec. 5). The combination of all these measures results in a significant improvement in the sensitivity to the *s*-channel process. Approximately half of this improvement can be attributed to the change in method from BDT to ME. In particular, the BDT technique applied to this analysis is limited by the sample sizes available for the training, while the ME approach is not sensitive to this limitation.”

Examples: MEM for searches

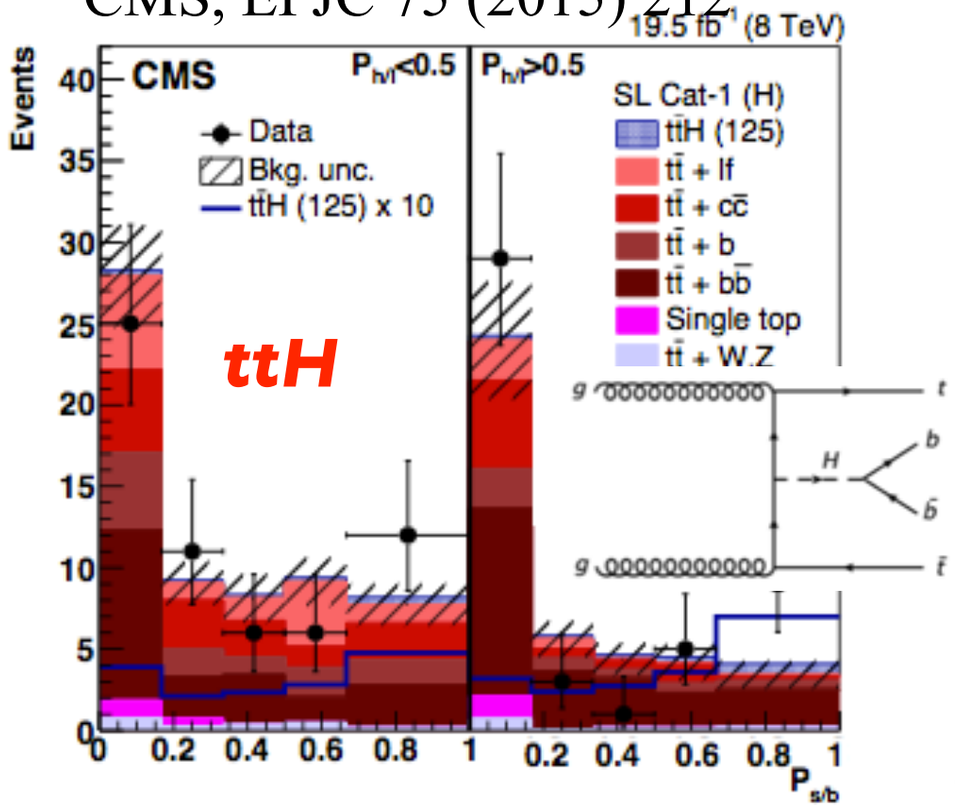
CDF, PRL 103 (2009) 092002



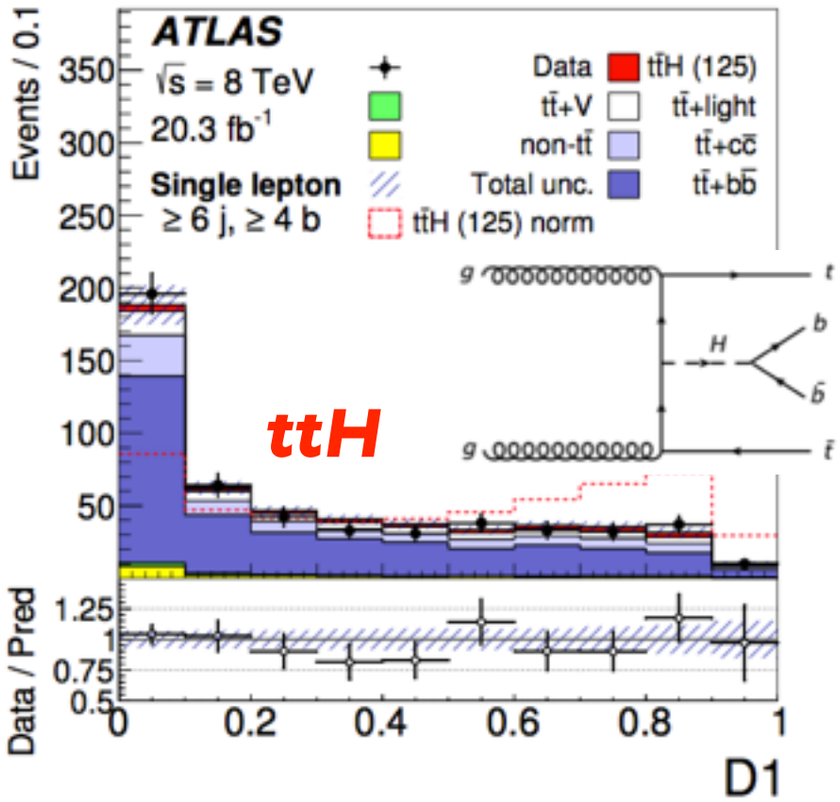
CDF, PRD 85 072001 (2012)



CMS, EPJC 75 (2015) 212

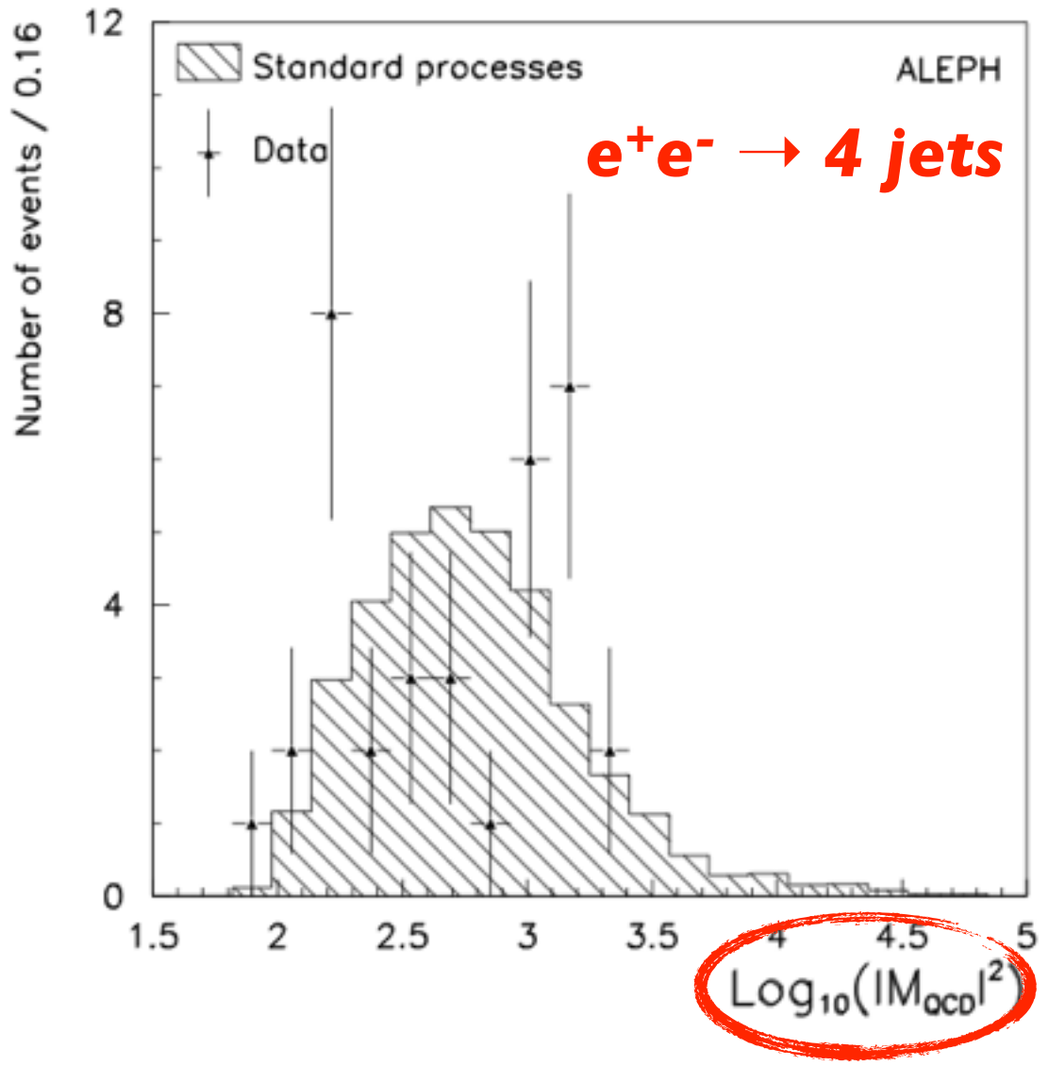


ATLAS, EPJC 75 (2015) 349

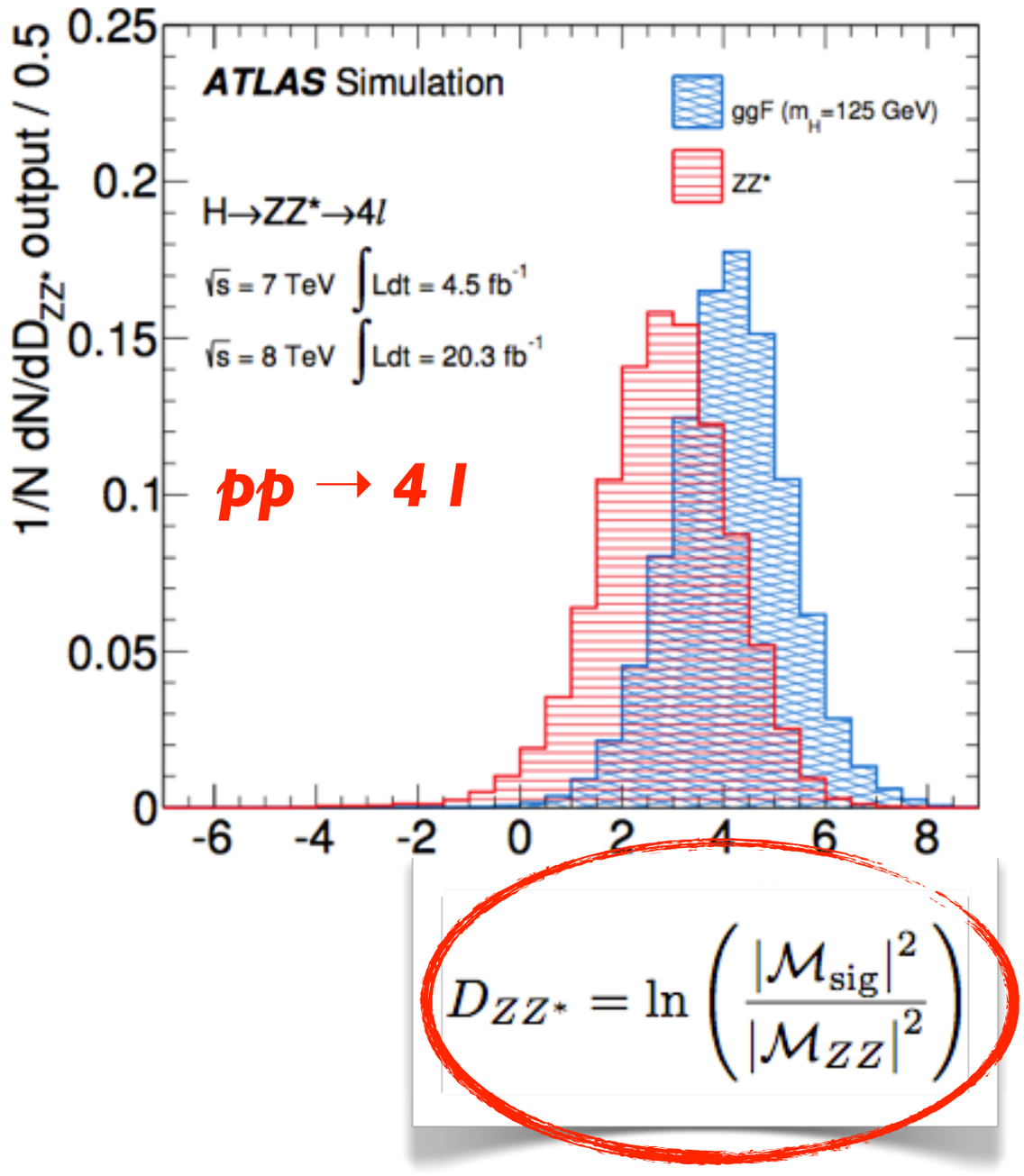


The ME Likelihood Approach

ALEPH, Z. Phys.C 71 (1997) 179



ATLAS, PRD 91 (2015) 012006



$$D_{ZZ^*} = \ln\left(\frac{|\mathcal{M}_{\mathrm{sig}}|^2}{|\mathcal{M}_{ZZ}|^2}\right)$$

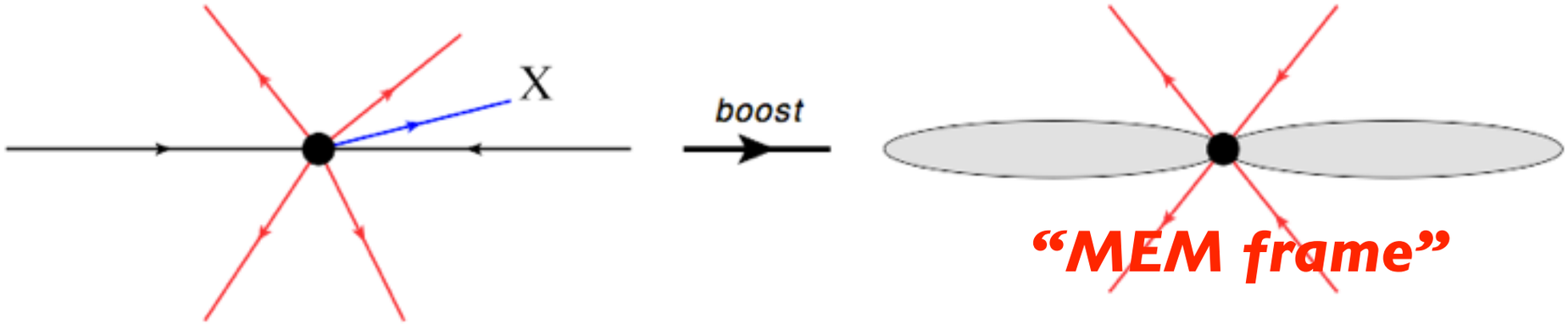
ATLAS, EPJC 75 (2015) 476
CMS, PRD 89 (2014) 092007
CMS, PRD 92 (2015) 012004

Going beyond the leading order

$\mathscr{M}$ typically calculated at leading order accuracy (LO)

- parton shower accounted for by jet transfer function
- ISR not accounted for
  - boost to Born-like frame:

J. Alwall et al., PRD 83 (2011) 074010
CDF, PRD 79 (2009) 072101

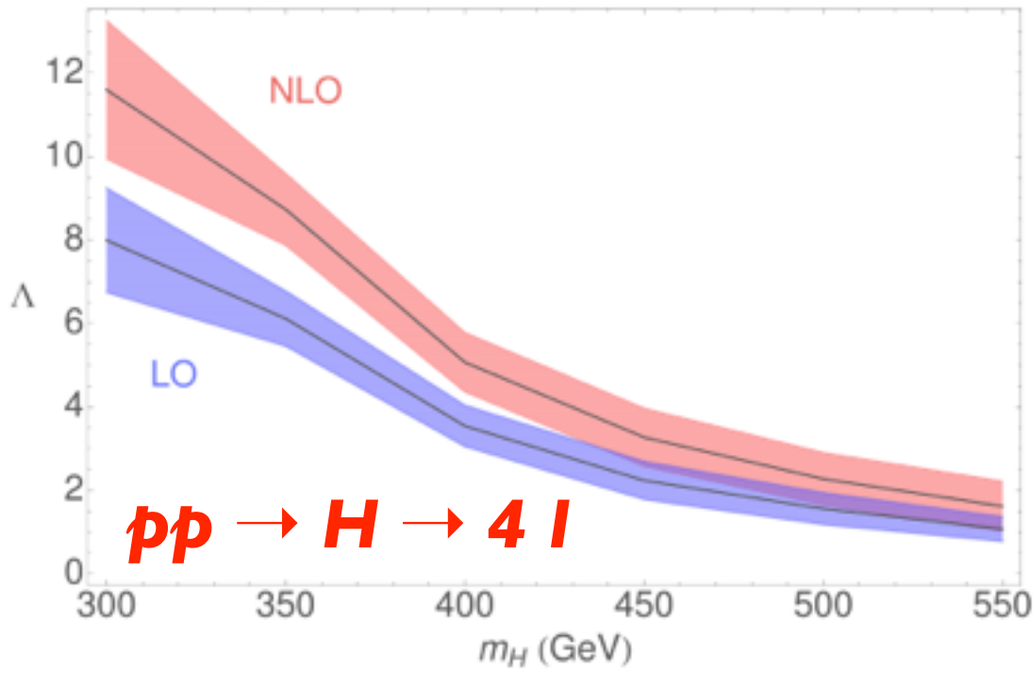


$$P_{x,y} = 0$$

  - Sudakov reweighting, p_T priors
  - MEM @ NLO:

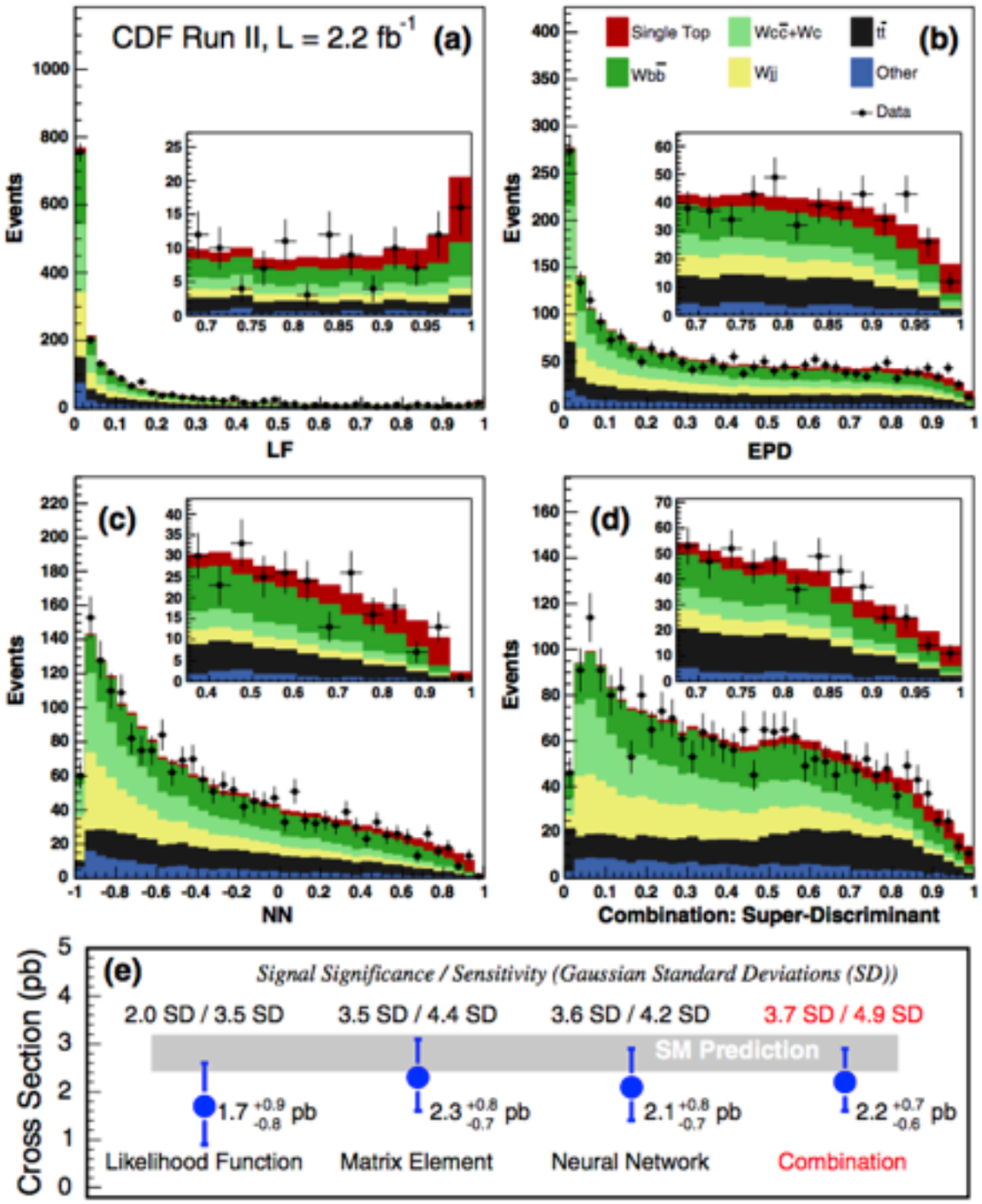
$$\mathcal{P}(\mathbf{x}|\Omega) = \frac{1}{\sigma_{\Omega}^{NLO}}\Big(V_{\Omega}(\mathbf{x}) + R_{\Omega}(\mathbf{x})\Big)$$

J. Campbell et al., JHEP 11 (2012) 043



MEM meets ML*

CDF, PRL 101 (2008) 252001



σ(single-t): 10% improvement compared to single-best discr.

ATLAS, EPJC 75 (2015) 349

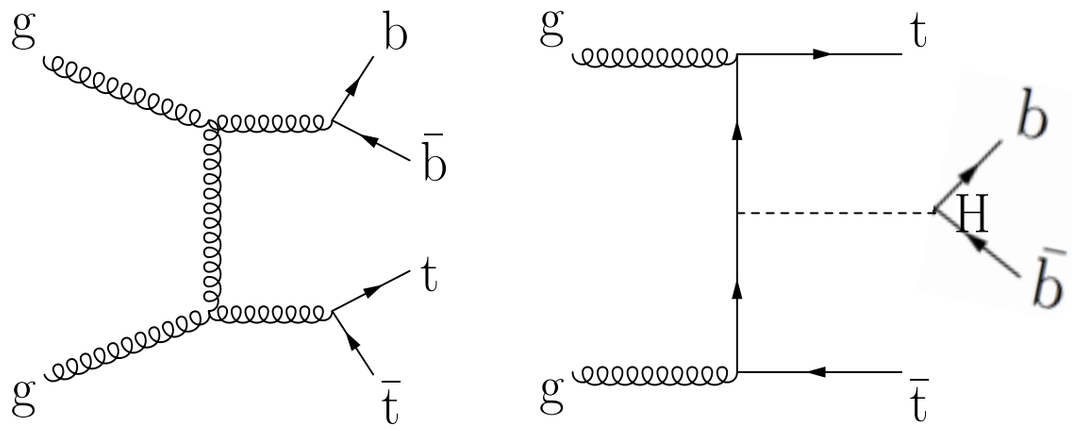
Variable	Definition	NN rank			
		≥ 5j, ≥ 4b	≥ 6j, 3b	5j, 4b	5j, 3b
$D1$	Neyman–Pearson MEM discriminant (Eq. (4))	1	10	-	-
Centrality	Scalar sum of the p_T divided by sum of the E for all jets and the lepton	2	2	1	-
p_T^{jet5}	p_T of the fifth leading jet	3	7	-	-
$H1$	Second Fox–Wolfram moment computed using all jets and the lepton	4	3	2	-
ΔR_{bb}^{avg}	Average ΔR for all b-tagged jet pairs	5	6	5	-
SSLL	Logarithm of the summed signal likelihoods (Eq. (2))	6	4	-	-
$m_{bb}^{min\ \Delta R}$	Mass of the combination of the two b-tagged jets with the smallest ΔR	7	12	4	4
$m_{bj}^{max\ p_T}$	Mass of the combination of a b-tagged jet and any jet with the largest vector sum p_T	8	8	-	-
$\Delta R_{bb}^{max\ p_T}$	ΔR between the two b-tagged jets with the largest vector sum p_T	9	-	-	-
$\Delta R_{lep-bb}^{min\ \Delta R}$	ΔR between the lepton and the combination of the two b-tagged jets with the smallest ΔR	10	11	10	-
$m_{uu}^{min\ \Delta R}$	Mass of the combination of the two untagged jets with the smallest ΔR	11	9	-	2
$Aplan_{b-jet}$	$1.5\lambda_2$, where λ_2 is the second eigenvalue of the momentum tensor[92] built with only b-tagged jets	12	-	8	-
N_{40}^{jet}	Number of jets with $p_T \geq 40GeV$	-	1	3	-
$m_{bj}^{min\ \Delta R}$	Mass of the combination of a b-tagged jet and any jet with the smallest ΔR	-	5	-	-
$m_{jj}^{max\ p_T}$	Mass of the combination of any two jets with the largest vector sum p_T	-	-	6	-
H_T^{had}	Scalar sum of jet p_T	-	-	7	-
$m_{jj}^{min\ \Delta R}$	Mass of the combination of any two jets with the smallest ΔR	-	-	9	-
$m_{bb}^{max\ p_T}$	Mass of the combination of the two b-tagged jets with the largest vector sum p_T	-	-	-	1
$p_{T,uu}^{min\ \Delta R}$	Scalar sum of the p_T of the pair of untagged jets with the smallest ΔR	-	-	-	3
$m_{bb}^{max\ m}$	Mass of the combination of the two b-tagged jets with the largest invariant mass	-	-	-	5
$\Delta R_{uu}^{min\ \Delta R}$	Minimum ΔR between the two untagged jets	-	-	-	6
m_{jjj}	Mass of the jet triplet with the largest vector sum p_T	-	-	-	7

***ttH: combination of* MEM *discriminant + other variables using an artificial* NN**

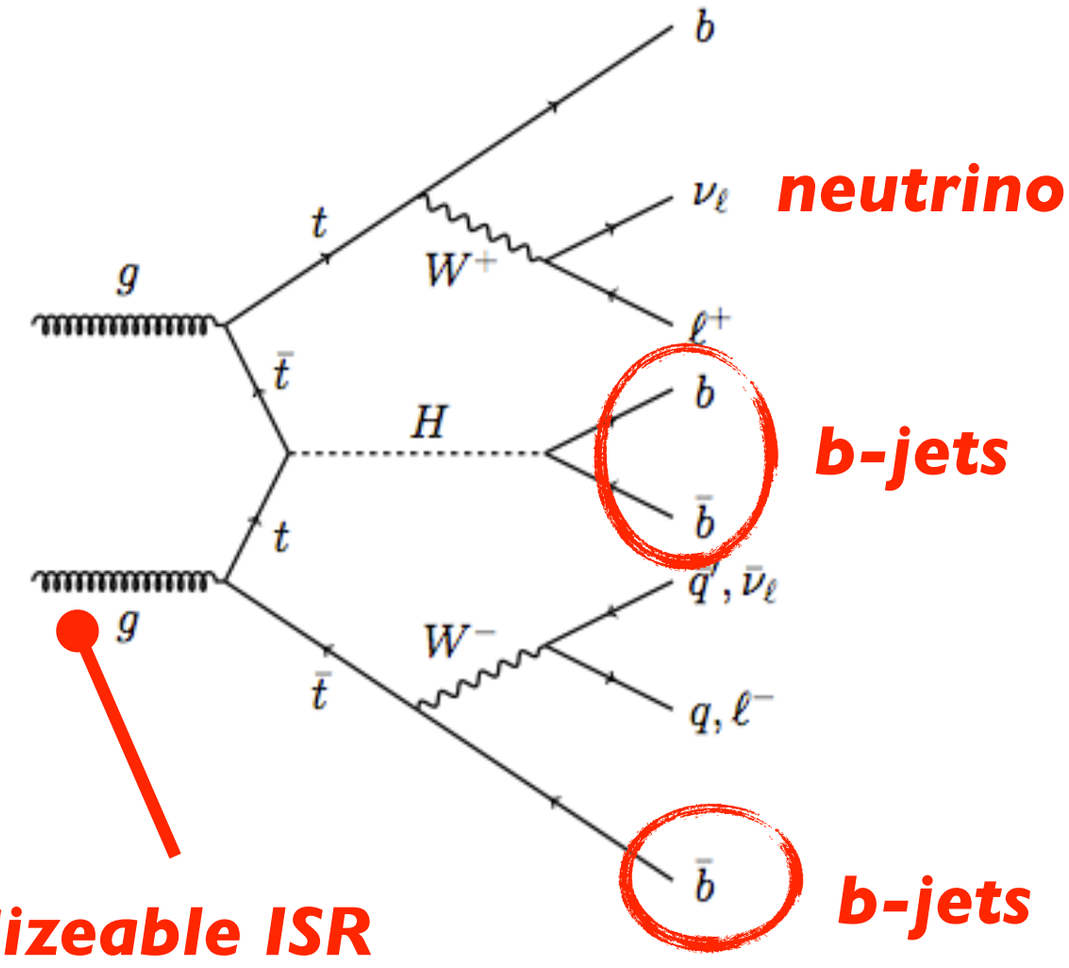
(*) this afternoon session

A complete example: ttH→bb

- Large irreducible background
  - same final state, similar diagrams
    - MEM to maximally exploit kinematical properties

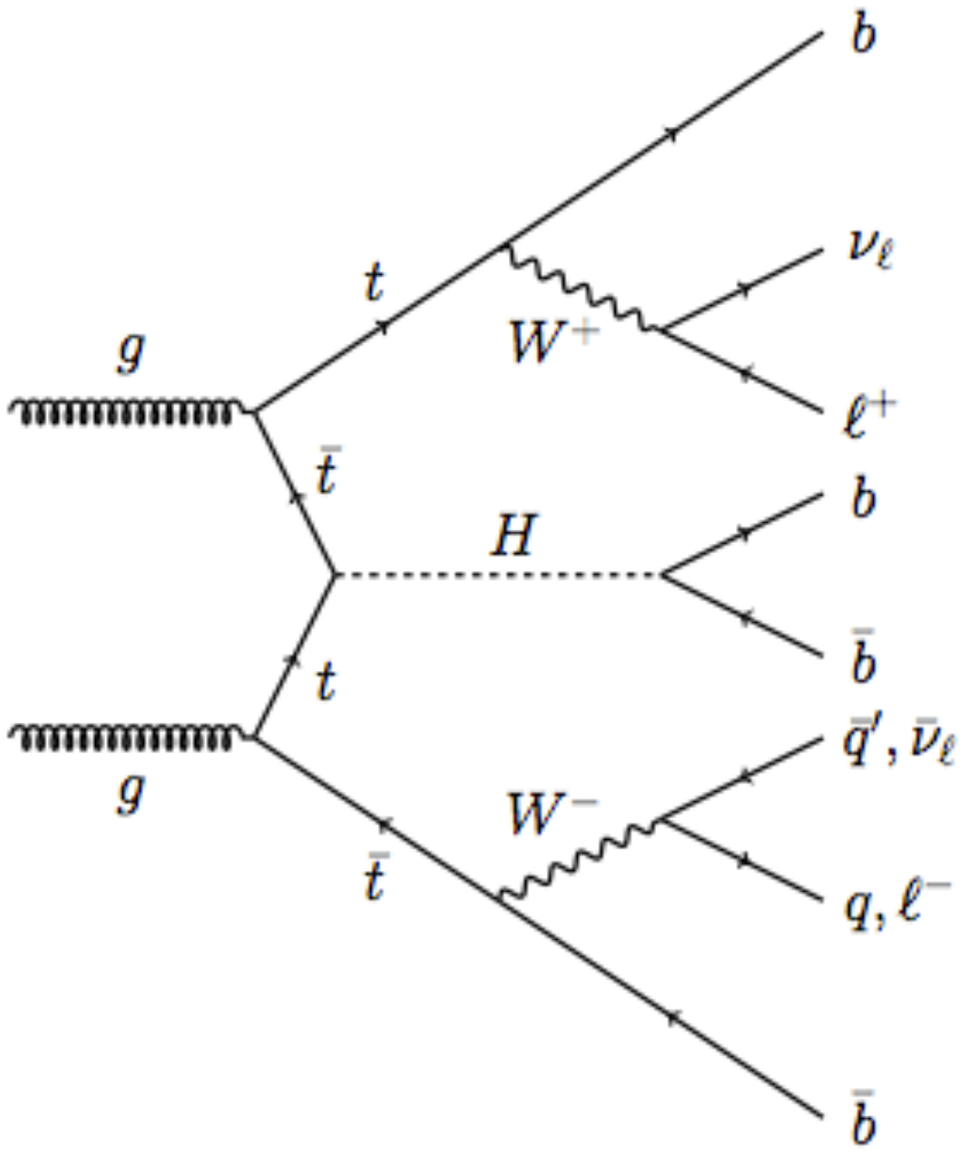


- Challenges:
  - 2 → 8 process
  - combinatorics
  - jets not always arise from top/ Higgs decay
    - only a fraction of events correspond to this picture:
    - allow for ISR jets and partial event reconstruction



P. Artoisenet et al., PRL 111 (2013) 091802
CMS, EPJC 75 (2015) 212
ATLAS, EPJC 75 (2015) 349

Phase-space integration



CASE 1:

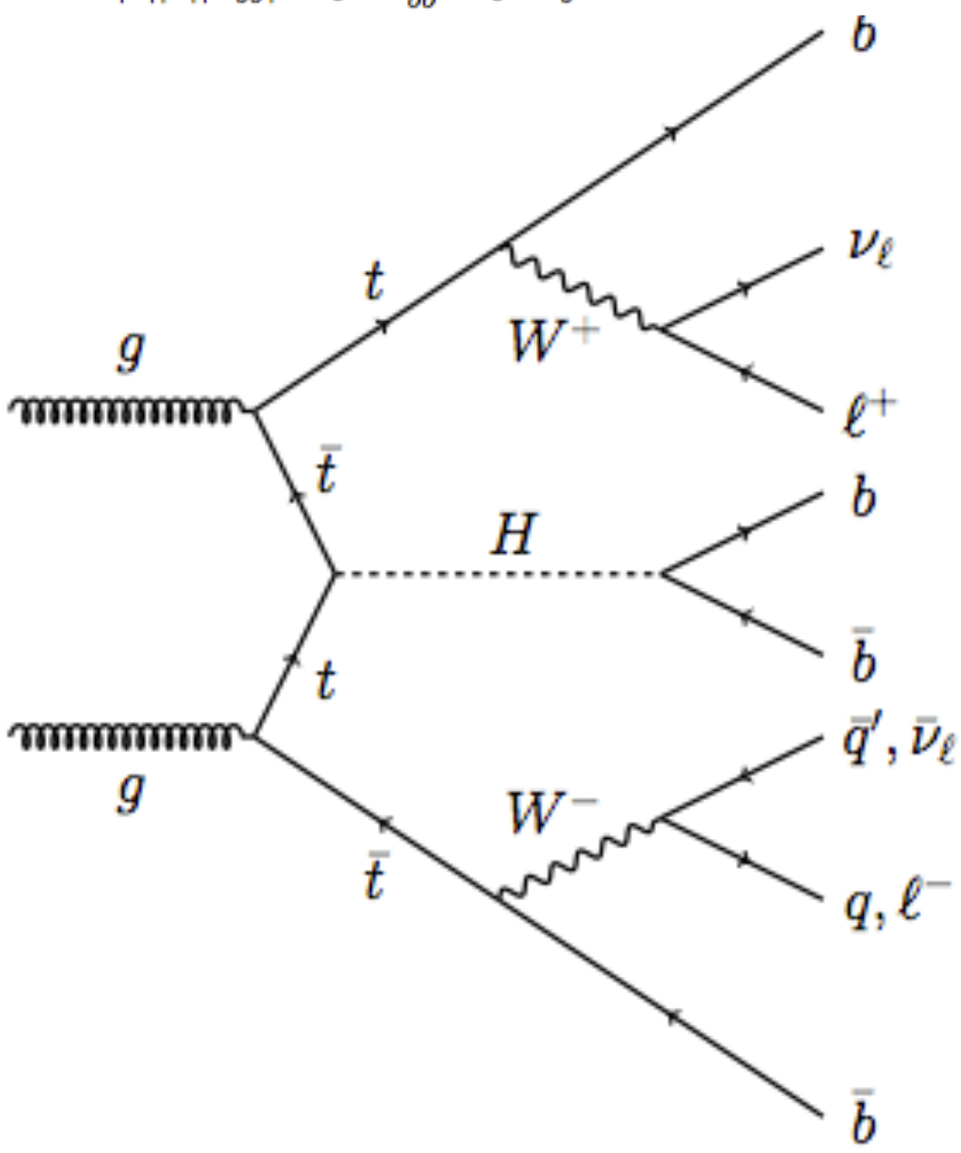
Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dm^2_{q\bar{q}\,i}dm^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{bi}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dm^2_{b\bar{b}}d\Omega_b d\Omega_{\bar{b}}$$

alignment along peaks



 # variables = 24

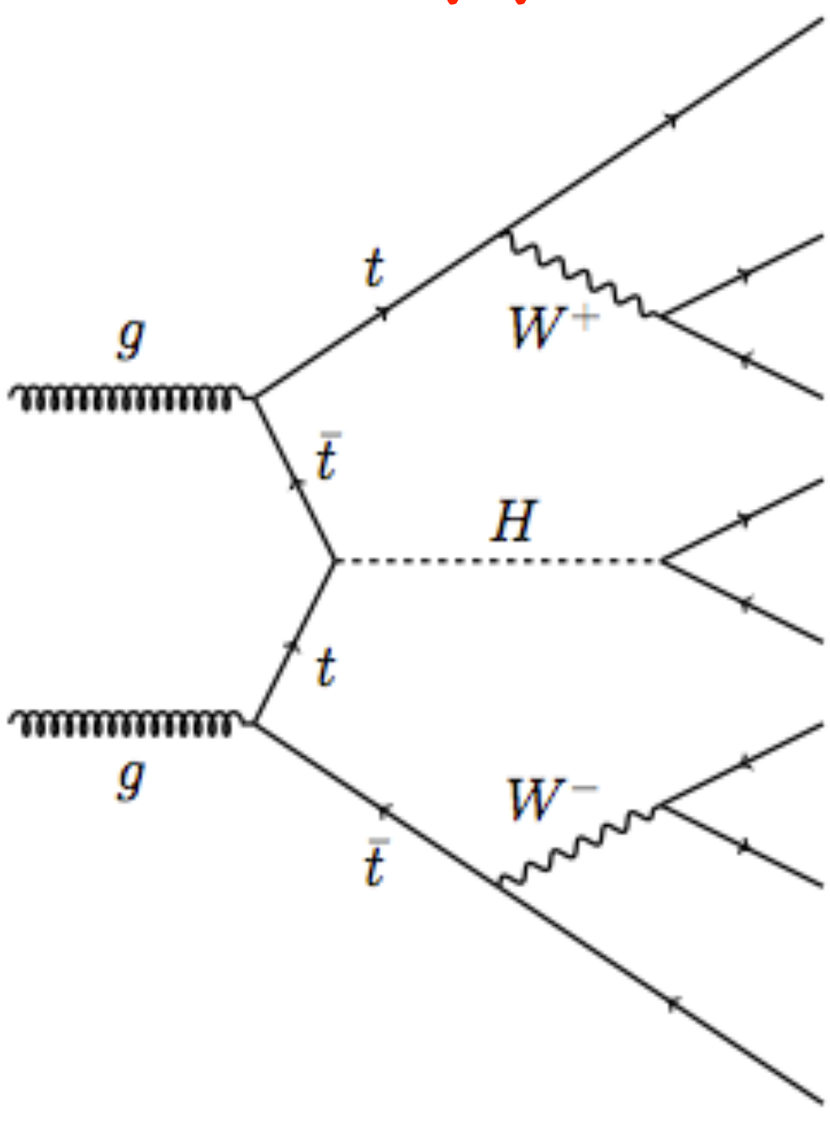
CASE 1:

Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dm^2_{q\bar{q}\,i}dm^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{b_i}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dm^2_{b\bar{b}} d\Omega_b d\Omega_{\bar{b}}$$



$$b \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

$$\nu_\ell \dashrightarrow \delta^3(\vec{p}_{rec} - \vec{p}_{gen})$$

$$\ell^+ \dashrightarrow \delta^3(\vec{p}_{rec} - \vec{p}_{gen})$$

$$b \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

$$\bar{b} \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

$$\bar{q}', \bar{\nu}_\ell \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

$$q, \ell^- \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

$$\bar{b} \dashrightarrow \delta^2(\Omega_{rec} - \Omega_{gen})$$

 # variables = 24 - 15

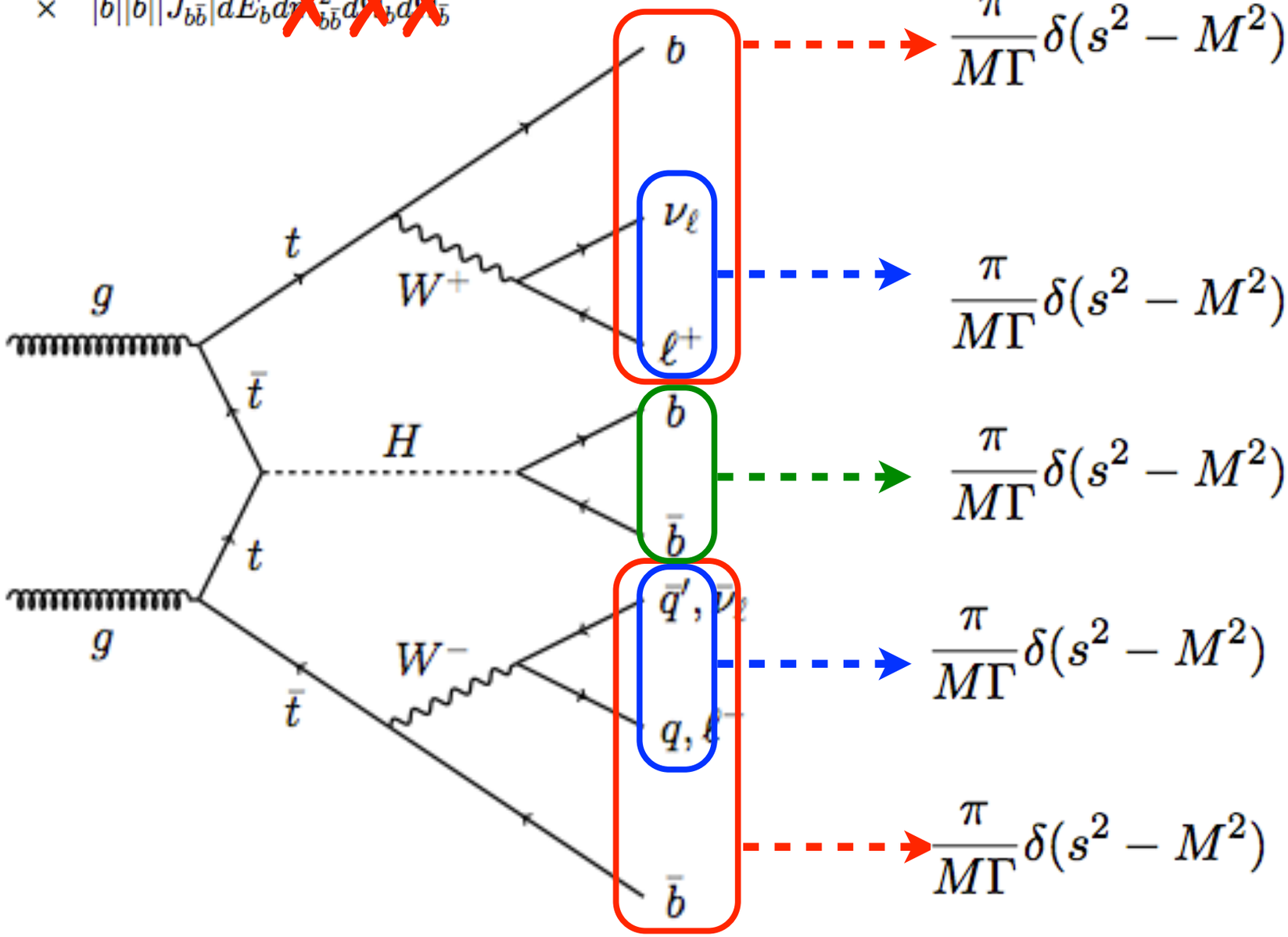
CASE 1:

Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dm^2_{q\bar{q}\,i}dm^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{b_i}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dm^2_{b\bar{b}} d\Omega_b d\Omega_{\bar{b}}$$



$$\frac{\pi}{M\Gamma}\delta(s^2 - M^2)$$

$$\frac{\pi}{M\Gamma}\delta(s^2 - M^2)$$

$$\frac{\pi}{M\Gamma}\delta(s^2 - M^2)$$

$$\frac{\pi}{M\Gamma}\delta(s^2 - M^2)$$

$$\frac{\pi}{M\Gamma}\delta(s^2 - M^2)$$

 # variables = 24 - 15 - 5

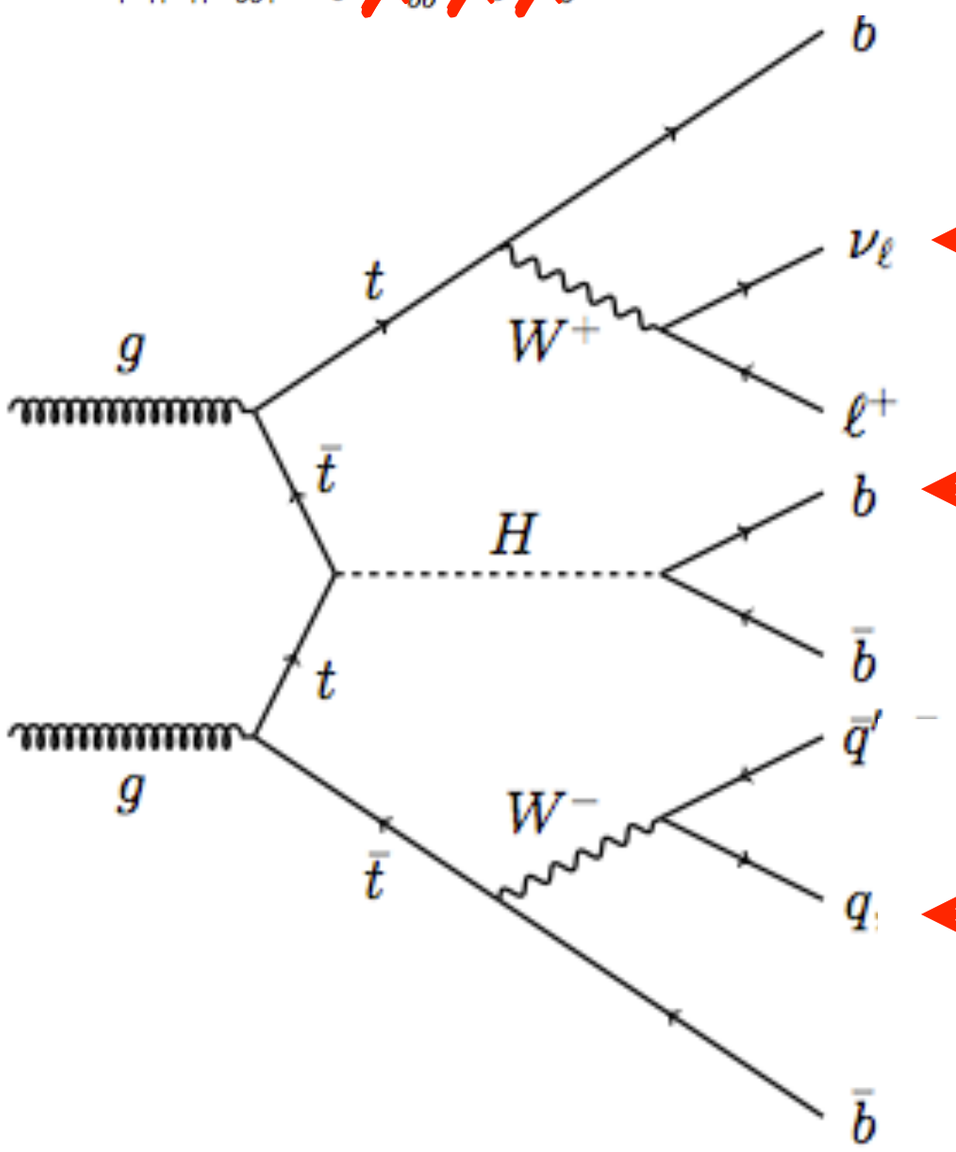
CASE 1:

Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dm^2_{q\bar{q}\,i}dm^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{b_i}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dm^2_{b\bar{b}} d\Omega_b d\Omega_{\bar{b}}$$



neutrino direction

b energy

u energy

 # variables = 24 - 15 - 5 = 4

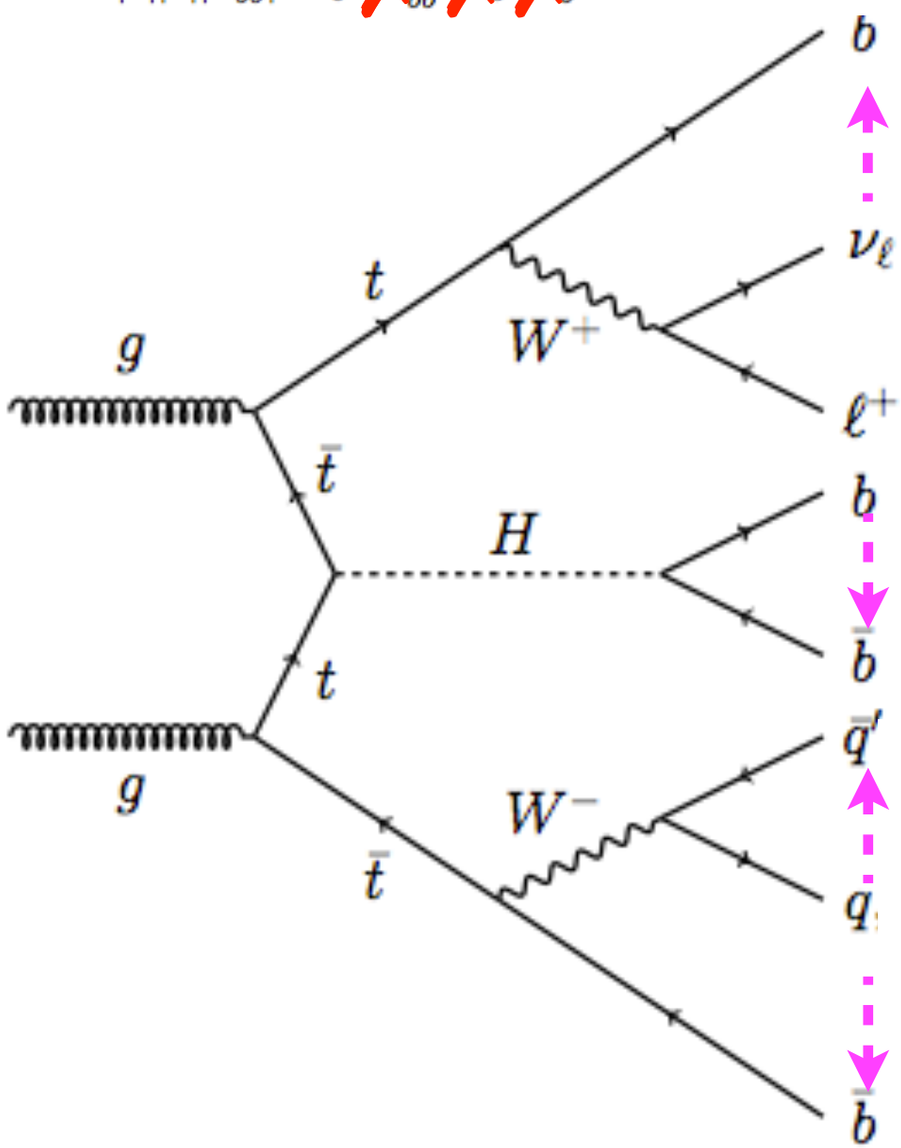
CASE 1:

Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dr^2_{q\bar{q}\,i}dr^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{b_i}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dr^2_{b\bar{b}} d\Omega_b d\Omega_{\bar{b}}$$



b quark energy

neutrino direction

b energy

anti-b energy

anti-d energy

u energy

anti-b energy

 # variables = 24 - 15 - 5 = 4

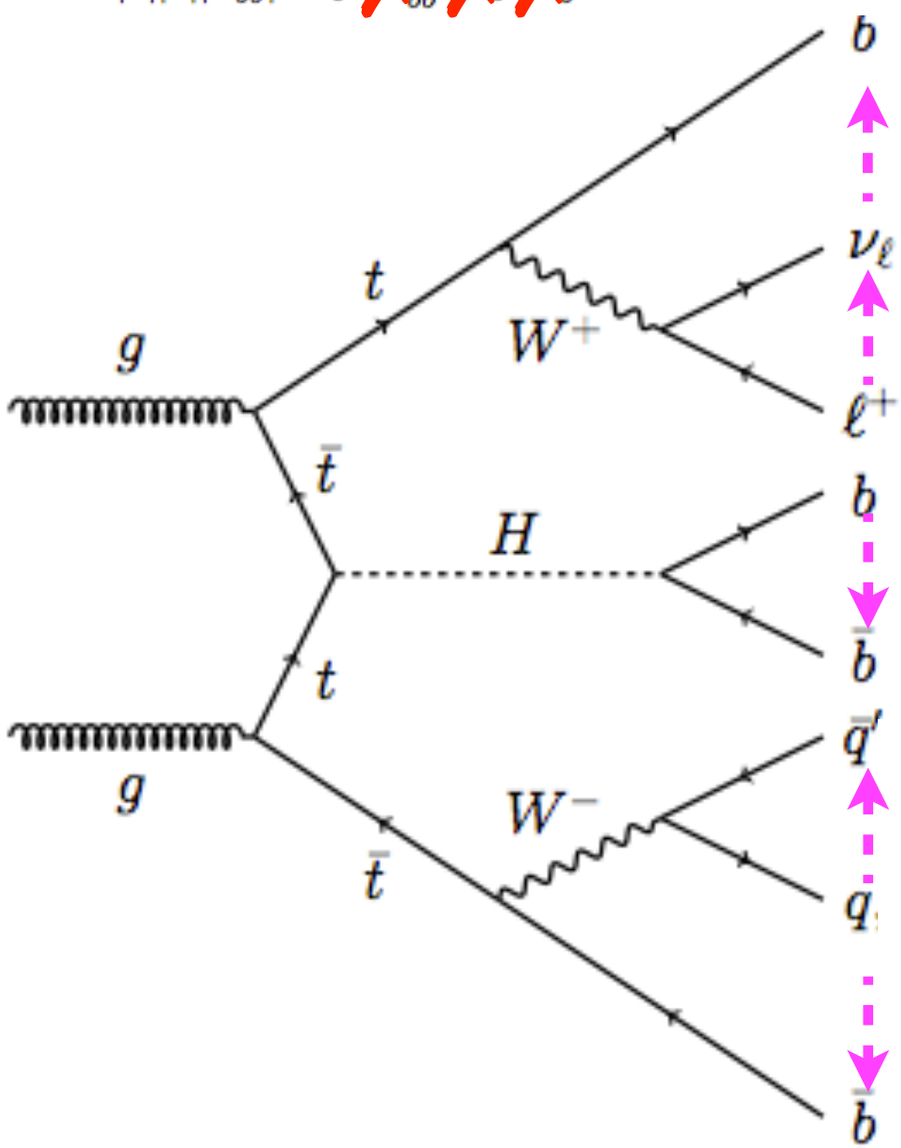
CASE 1:

Assume full reconstruction

Phase-space integration

$$d\Phi_S = \left(\frac{(2\pi)^{-3}}{2}\right)^8 \prod_i \left[|\vec{q}_i||\vec{\bar{q}}_i||\vec{b}_i||J_{t_i}|dE_{q_i}dm^2_{q\bar{q}\,i}dm^2_{q\bar{q}b\,i}d\Omega_{q_i}d\Omega_{\bar{q}_i}d\Omega_{b_i}\right] \times$$

$$\times \quad |\vec{b}||\vec{\bar{b}}||J_{b\bar{b}}|dE_b dm^2_{b\bar{b}} d\Omega_b d\Omega_{\bar{b}}$$



b ----> ***b quark energy***

ν_ℓ <---- ***neutrino direction***

ℓ^+

b <---- ***b energy***

$\bar{b}$ ----> ***anti-b energy***

$\bar{q}'$ ----> ***anti-d energy***

<---- ***anti-d direction***

$\bar{q}$ <---- ***u energy***

$\bar{b}$ ----> ***anti-b energy***

 # variables = 24 - 13 - 5 = 4+2

CASE 2:

Marginalise qbar direction

Sketching the code

Example of MEM implementation for a ttH search (~3k lines)

- requirements: C++ compiler, ROOT, LHAPDF
- features:
  - factorisation of MEM can be easily implemented in object-oriented languages
  - channel-specific optimisations

```
int main(){

  // create a new instance of the class
  MEMCalculator* my_mem_example = new MEMCalculator();

  // add observables as members of the class
  RecoObject j1;
  j1.SetPtEtaPhiM( /* ... */ );
  j1.AddObservable( /* ... */);
  j1.AddTransferFunction( tf );
  my_mem_example->push_back_object( &j1 );
  /* ... */

  p_s = my_mem_example->run( Hypothesis::TTH,  {Particle::qbar1} );
  p_b = my_mem_example->run( Hypothesis::TTBB, {Particle::qbar1} );

  double mem_discriminant = p_s / (p_s + p_b)
}
```

A C++ class; observables are class members

S or B, nonreconstructed particles, ...

```
double MEMCalculator::run( /* ... */ ){

  // build integrand from Eval() method of this class
  Math::Functor toIntegrate(this, &MEMCalculator::Eval, npar);

  // GSL integrator, attach toIntegrate() to it
  GSLMCIntegrator* vegas = new GSLMCIntegrator(
          IntegrationMultiDim::kVEGAS,
          abs, rel, n_max_calls);

    // attach the integrand to the integrator
  vegas->SetFunction(toIntegrate);

  // integration ranges: npar = # of integration variables
  // need a routine to determine the integral domain
  double xL[npar], xU[npar];
  set_integration_range(xL, xU);

  // that's it, just call Integral()!!!
  return vegas->Integral(xL,xU);

  /* ... */
}
```

Eval() must return the integrand

GSL implementation of VEGAS

This function does what it says!

```
double MEMCalculator::Eval(const double* x) const {

  // define a new phase-space point
  PhaseSpacePoint ps = create_PS(x);

  // calculate the integrand
  double p{1.0};

  p *= theory(ps);
  p *= transfer(ps);

  return p;
}
```

double* x = integration variables

Each VEGAS ***point "x" maps to a phase-space point "ps"***
This is process-specific

`theory(ps)` `transfer(ps)`

$$\int d\Phi(\vec{x}) dx_a dx_b \sum_{i,j} \frac{f_i(x_a) f_j(x_b)}{(1+\delta_{ij}) x_a x_b s} |\mathcal{M}_S(\vec{x}, \boldsymbol{\theta})|^2 W(\vec{y}, \vec{x}; \boldsymbol{\theta})$$

```
double MEMCalculator::Eval(const double* x) const {
  /* ... */
  PhaseSpacePoint ps = create_PS(x);
  /* ... */
}

PhaseSpacePoint MEMCalculator::create_PS( const double* x ) const {

  // deal with Particle::q1
  if( isMatched(Particle::q1) ){
    dir     = jets[ Particle::q1 ].p4().Vect().Unit();
  }
  else{
    dir.SetTheta( x[Variable::theta_q1]);
    dir.SetPhi  ( x[Variable::phi_q1]);
  }
  // E_q1 is one of the integration variables
  E_q1 = x[Variable::E_q1];

  /* ... */

}
```

if a quark is matched to a jet: use jet direction

else: integrate over quark direction

```
double MEMCalculator::Eval(const double* x) const {
  /* ... */
  p *= theory(ps);
  p *= transfer(ps);
  /* ... */
}

double MEMCalculator::theory(const PhaseSpacePoint& ps) const {

  // read particle 4-vector from the phase-space point
  LorentzVector q1 = ps.lv(Particle::q1);
  /* ... */

  m *= pdf( x1, x2 );
  m *= amplitude2( t, tx, b, bx );

  /* ... */

}

double MEMCalculator::transfer(const PhaseSpacePoint& ps) const {

  // loop over all quarks/leptons/neutrino
  for( PSMap::const_iterator p = ps.begin() ; p != ps.end() ; ++p ){
     if( isJet( p ) ){
       w *= TMath::Gaus( jets[p]->E(), p->E(), 0.15*p->E() );
      /* ... */
      }
  }
  /* ... */
}
```

Loop over all particles and evaluate W(x,y)

Assume simple Gaussian resolution

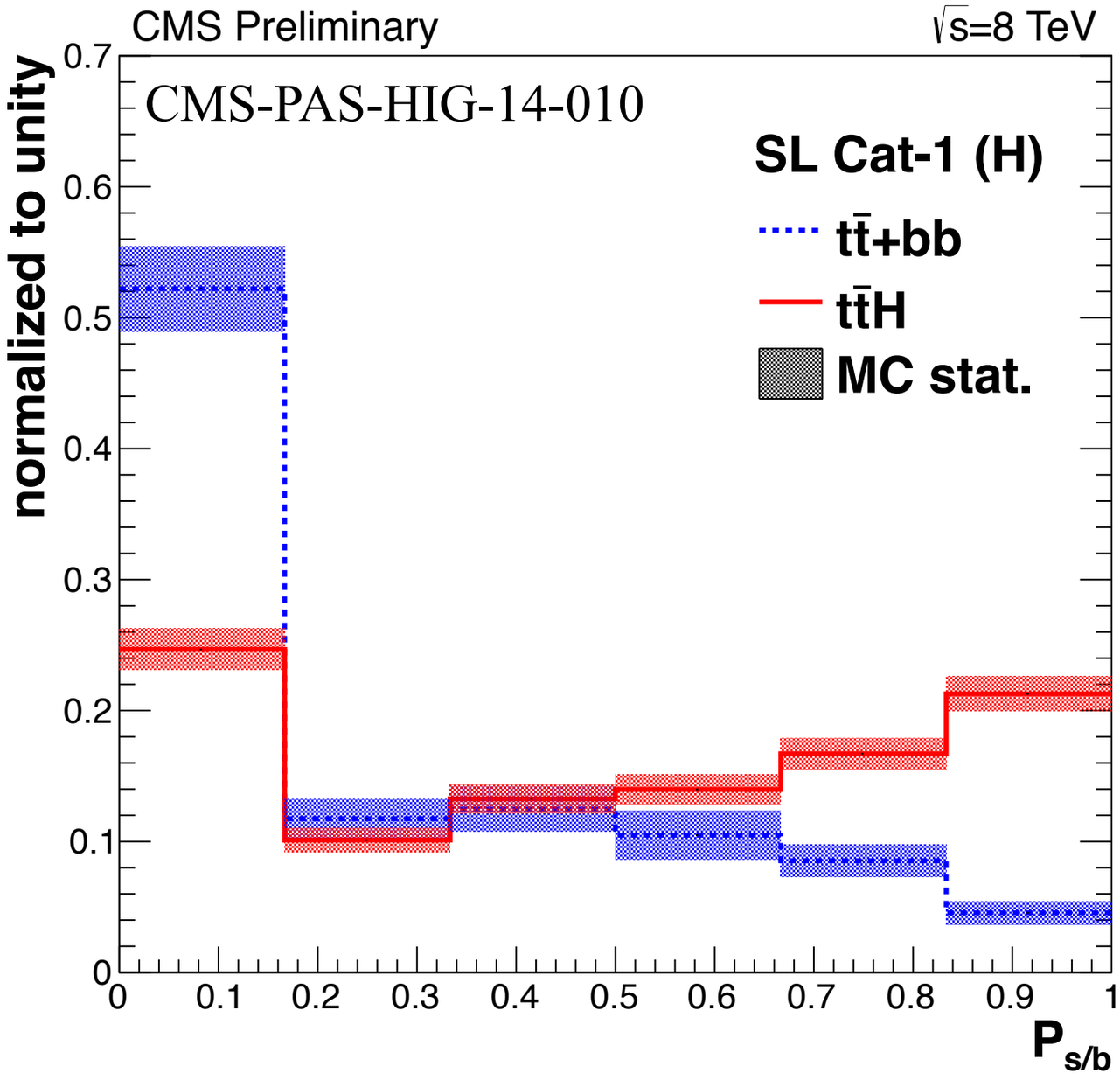
```
double MEMCalculator::theory(const PhaseSpace& ps) const {
  /* ... */
  m *= pdf( x1, x2 );
  m *= amplitude( t, tx, b, bx );
  /* ... */
}

double MEMCalculator::amplitude2( /* ... */ ) const {

  // Boost to "MEM frame"
  TVector3 boostPt( vSum.Px()/vSum.E(), vSum.Py()/vSum.E(), 0.0 );
  t.Boost( -boostPt ); tx.Boost( -boostPt ); h.Boost( -boostPt );

  /* ... */

  // interface to Fortran routine (OpenLoops)
  if( hypo==Hypothesis::TTH )
     pphttxcallme2born_(g1, g1, t, tx, b, bx, MTOP, MH);
  else:
     ppttxbbxcallme2born_(g1, g1, t, tx, b, bx, MTOP);
  /* ... */
}

double MEMCalculator::pdf(const double& x1, const double& x2, const double& Q)
   const {
  /* ... */
  double f1 =  LHAPDF::xfx(1, x1, Q, 0)/x1;
  double f2 =  LHAPDF::xfx(1, x2, Q, 0)/x2;
  /* ... */
}
```

Lorentz algebra (e.g. boost to MEM-frame)

Routines that compute the matrix element

Call to LHAPDF ***functions***

Results


CMS Preliminary
√s=8 TeV
CMS-PAS-HIG-14-010
SL Cat-1 (H)
tt̄+bb
tt̄H
MC stat.
normalized to unity
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
P s/b

Summary & outlook

- The MEM in High Energy Physics
  - compelling arguments behind the MEM approach for
    - parameter estimation
    - hypothesis testing
  - pioneered at Tevatron
    - top mass, W helicity, single-top measurements, Higgs searches
  - ramping up at LHC
    - single-top, ttH

- A field where technologies for data analysis matter
  - MEM remains a CPU-intensive business!
    - accelerating plateforms (e.g. GPU)
    - distributed systems (e.g. MPI)

$$\mathcal{P}(\{o\}|m_t) \propto \int f(x)f(x)|\mathcal{M}|^2 p(\{o\}|\{v\})\,\delta^4 \mathrm{d}^{18}\{v\}\mathrm{d}x\mathrm{d}x, \quad (6)$$

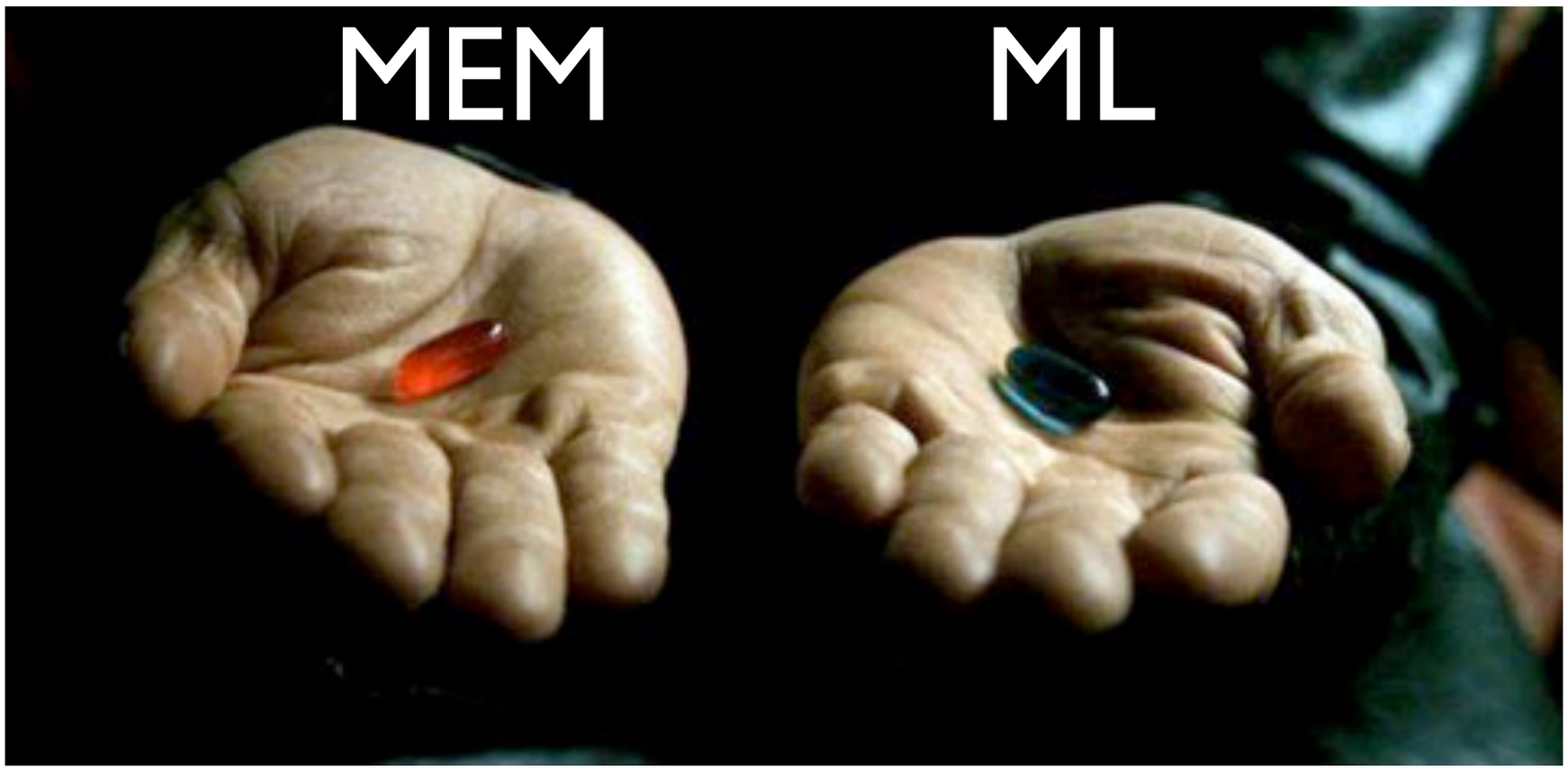
“Unfortunately this expression involves a multidimensional integral that has to be evaluated numerically and is complicated by the need to include initial and final state gluon radiation. Such higher order effects complicate the reconstruction of the top quark mass substantially and cannot be neglected. We therefore do not attempt to compute the exact probability density given in Eq. (6). Rather, we construct simpler weights that retain sensitivity to the value of the top quark mass but can be evaluated with the available computing resources. We calibrate the effect of the simplifi-”

D0, PRD 60, 052001 (**1999**)

Summary & outlook

A field where ML can have some complementarity

- higher-order predictions difficult to integrate into the MEM
  - LO vs NLO, parton shower, transfer function



Summary & outlook

A field where ML can have some complementarity

- higher-order predictions difficult to integrate into the MEM
  - LO vs NLO, parton shower, transfer function
- ML can help where MEM falls short
  - several examples already exist
- squeezing every bit of information out of LHC data is our mandate!

Thanks for your attention!

Many discussions and ideas shared with P. Artoisenet, G. Dissertori, B. Kilminster, O. Nackenhorst, J. Pata, A. Quadt, E. Shabalina, C. Veelken

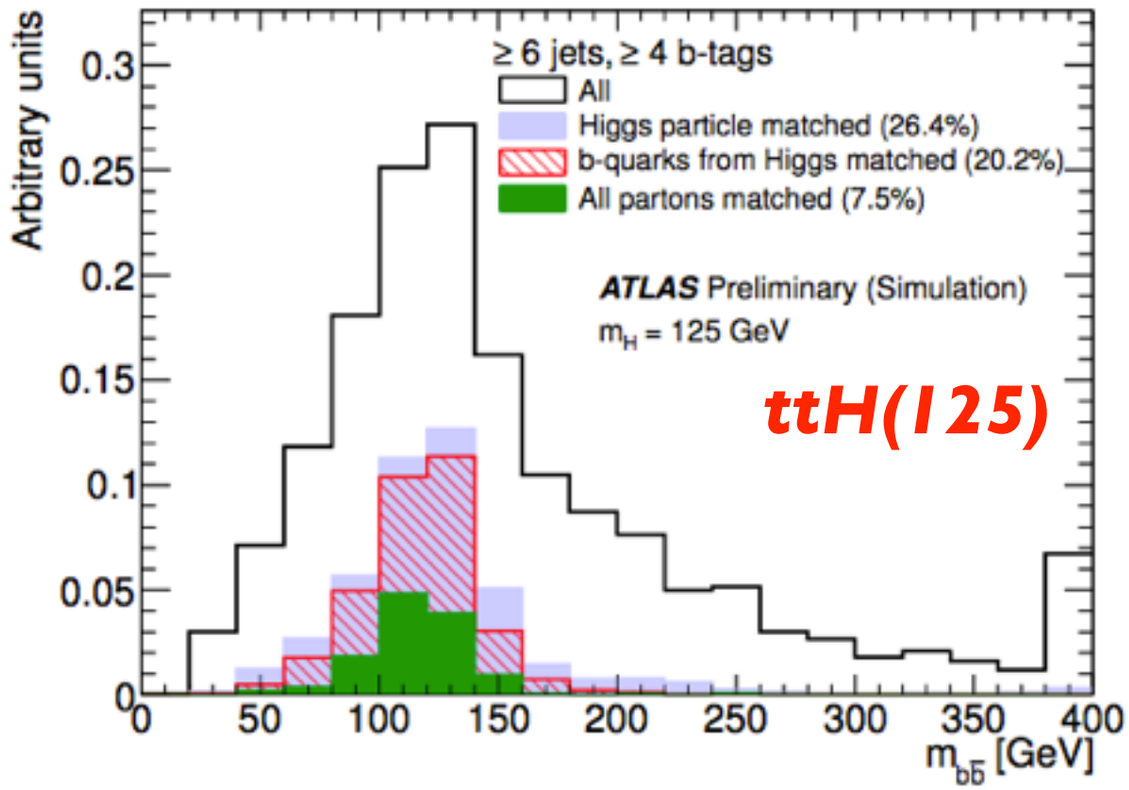
Special thanks to F. Canelli for illuminating discussions on this field.

Back up

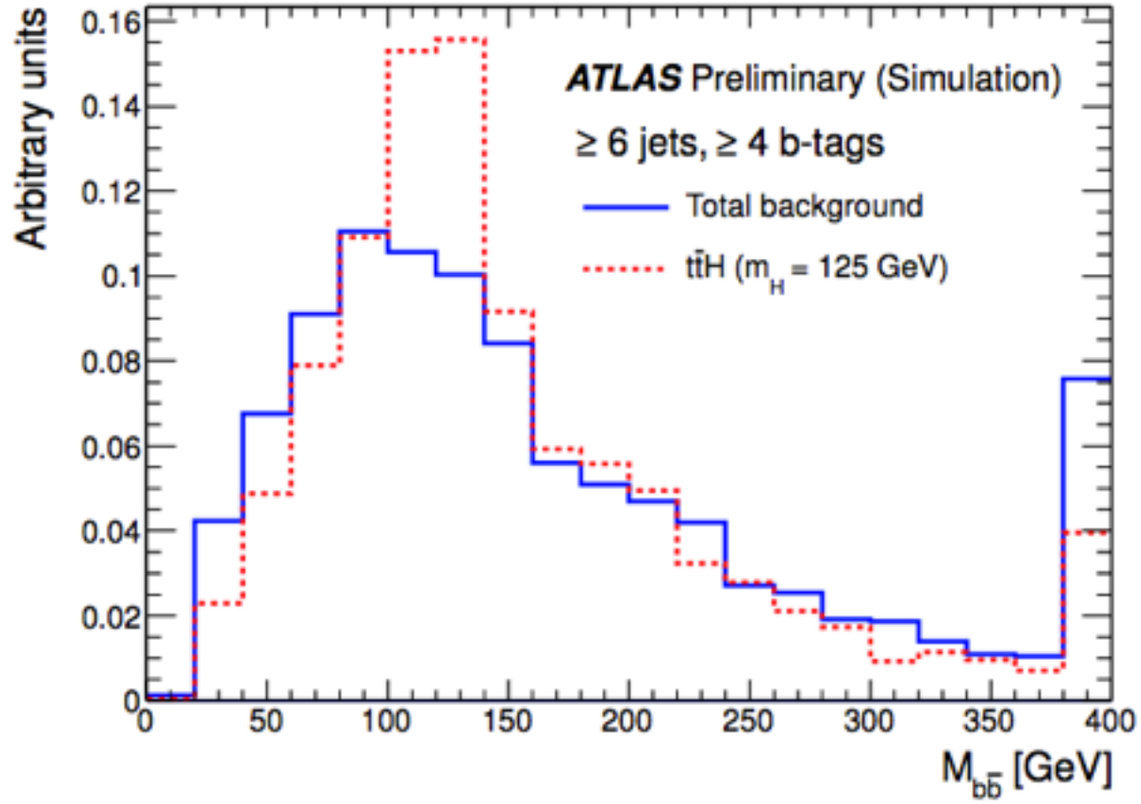
A kinematic fit?

Consider e.g. a kinematic fit

- jet ⟷ quark assignment based on minimum χ^2
- works for fully-reconstructed events
- matching efficiency is low!



⇒ *"combinatorial background" sculpted around m_H*

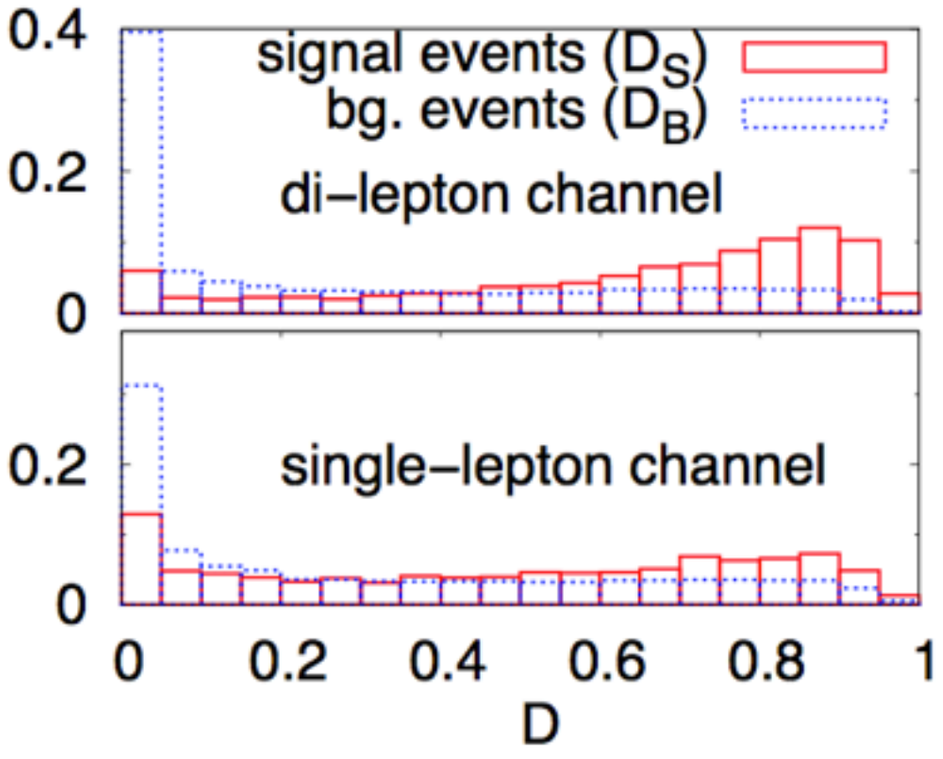


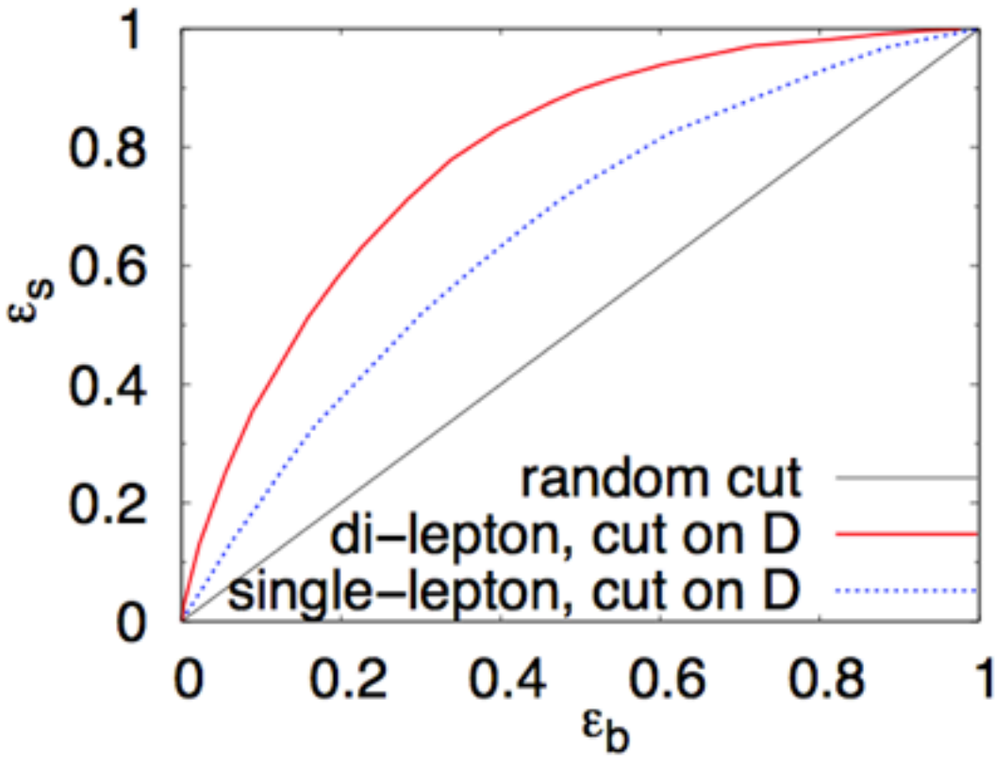
⇒ *background shape ~degenerate (poor separation)*

ATLAS-CONF-2012-135

A possible analysis strategy

- Need a multi-dimensional analysis of particle kinematics
- MEM ideally suited for the task
  - maximise separation against ttbb
  - address combinatorial self-background
- Proof of principle worked out on 14 TeV simulation
  - likelihood analysis to $D_i = \frac{P(\boldsymbol{x}_i|S)}{P(\boldsymbol{x}_i|S) + P(\boldsymbol{x}_i|B)}$ discriminant



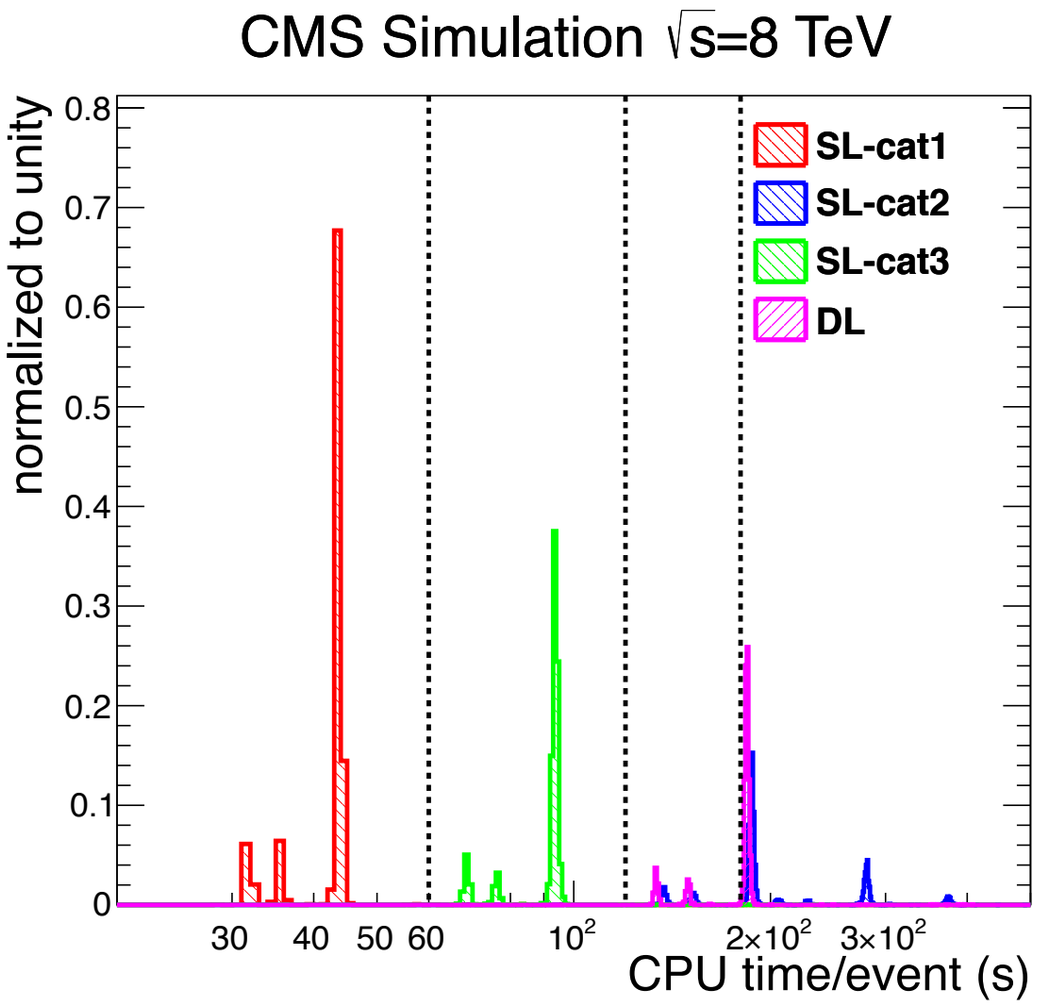


P. Artoisenet et al.,
PRL 111 (2013) 091802

FAQ: is CPU time an issue ?

Not really; compromise between **performance** and **timing**:

- run only on the good "events"
- filter-out permutations using b tagging (6! → 4!)
- test one background hypothesis only
  - optimize separation against tt+bb
- parallelize (by event) as much as possible
- neglect spin/correlations
- JEC/JER systematics: bookkeep VEGAS grid result from "nominal" for faster evaluation
- ...

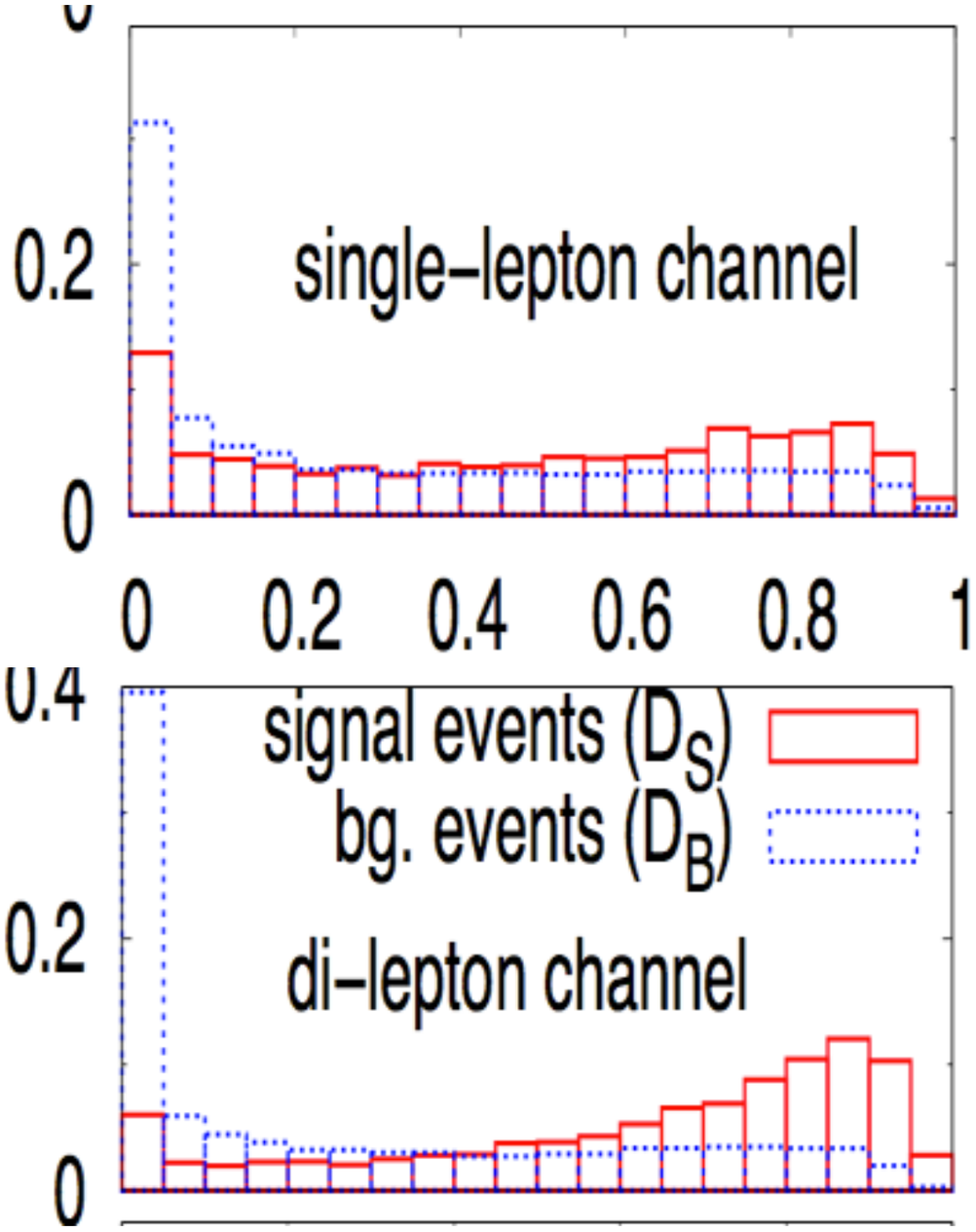


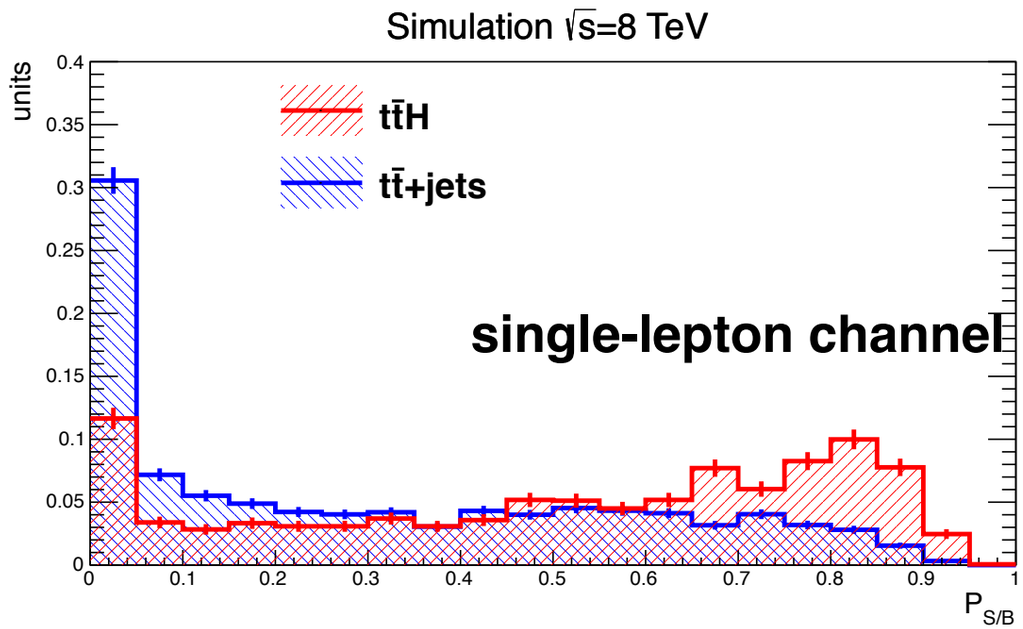
Number of variables	4 (+1)	6 (+1)	5 (+1)
Number of iterations	5	5	5
Function calls	2000	4000	10000
Numerical precision (mode of σ_w/w)	0.8%	1.2%	0.8%
CPU-time per integral (mean)	0.5 (1.5) s	1.1 (3.2) s	2.3 (6.2) s
Time budget for $\|\mathcal{M}\|^2_{\mathrm{ME}}$	30% (80%)	30% (80%)	30% (80%)

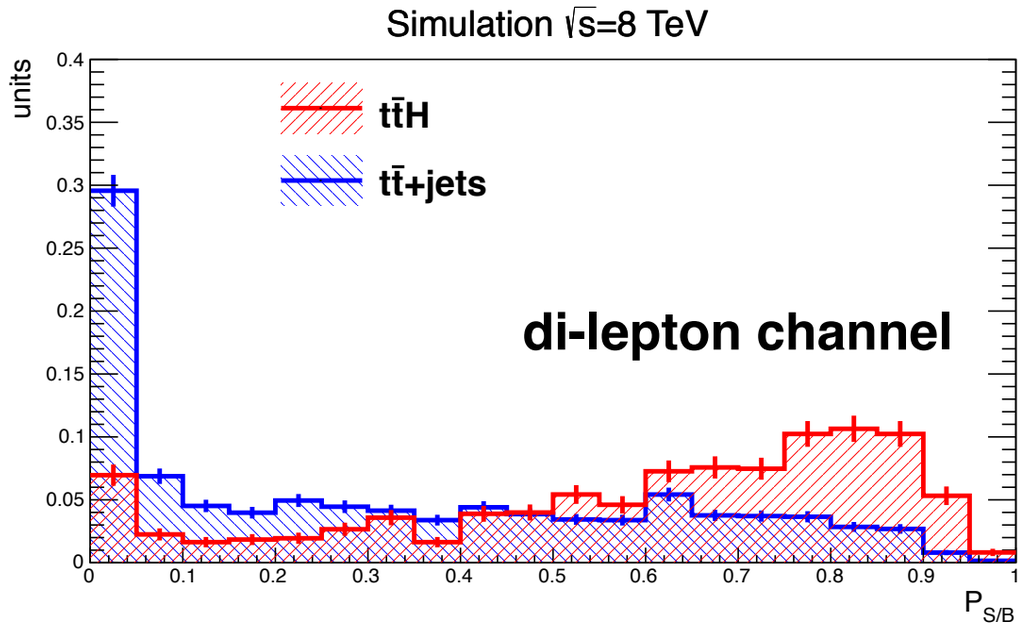
S (B) integration

Comparing with MadWeight

arXiv:1304.6414

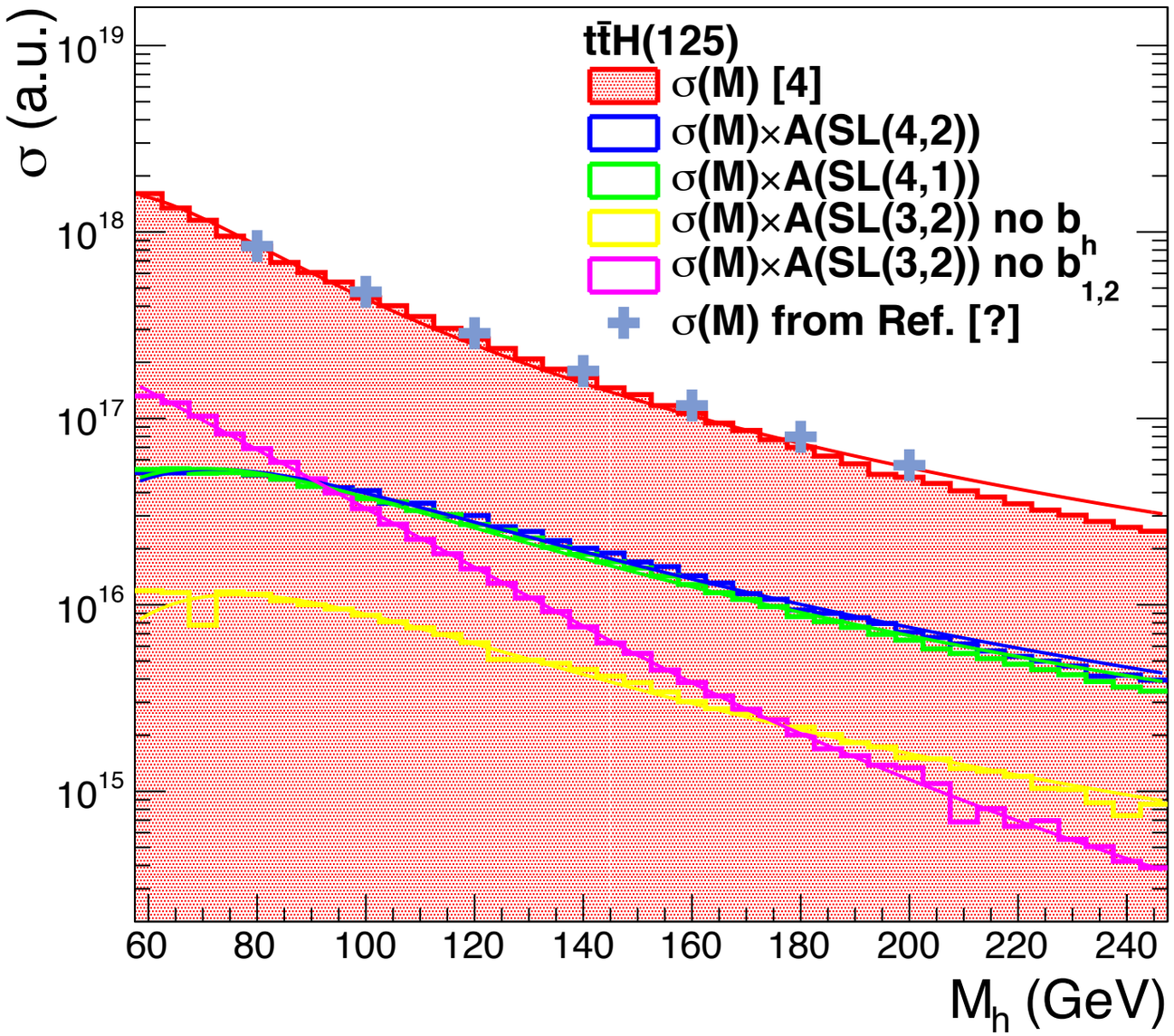






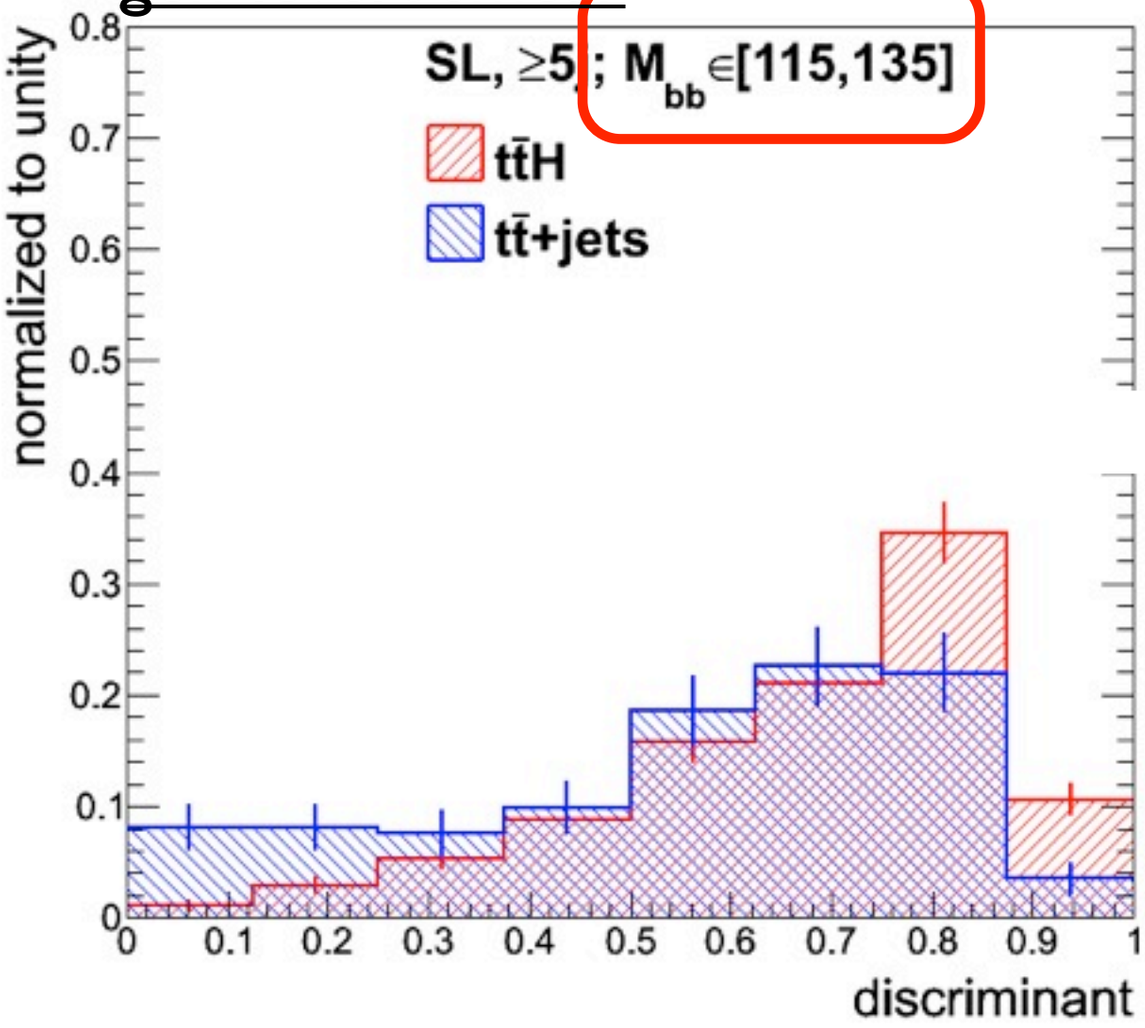
* looser *b*-tag cuts to enhance MC stat.

Cross-section

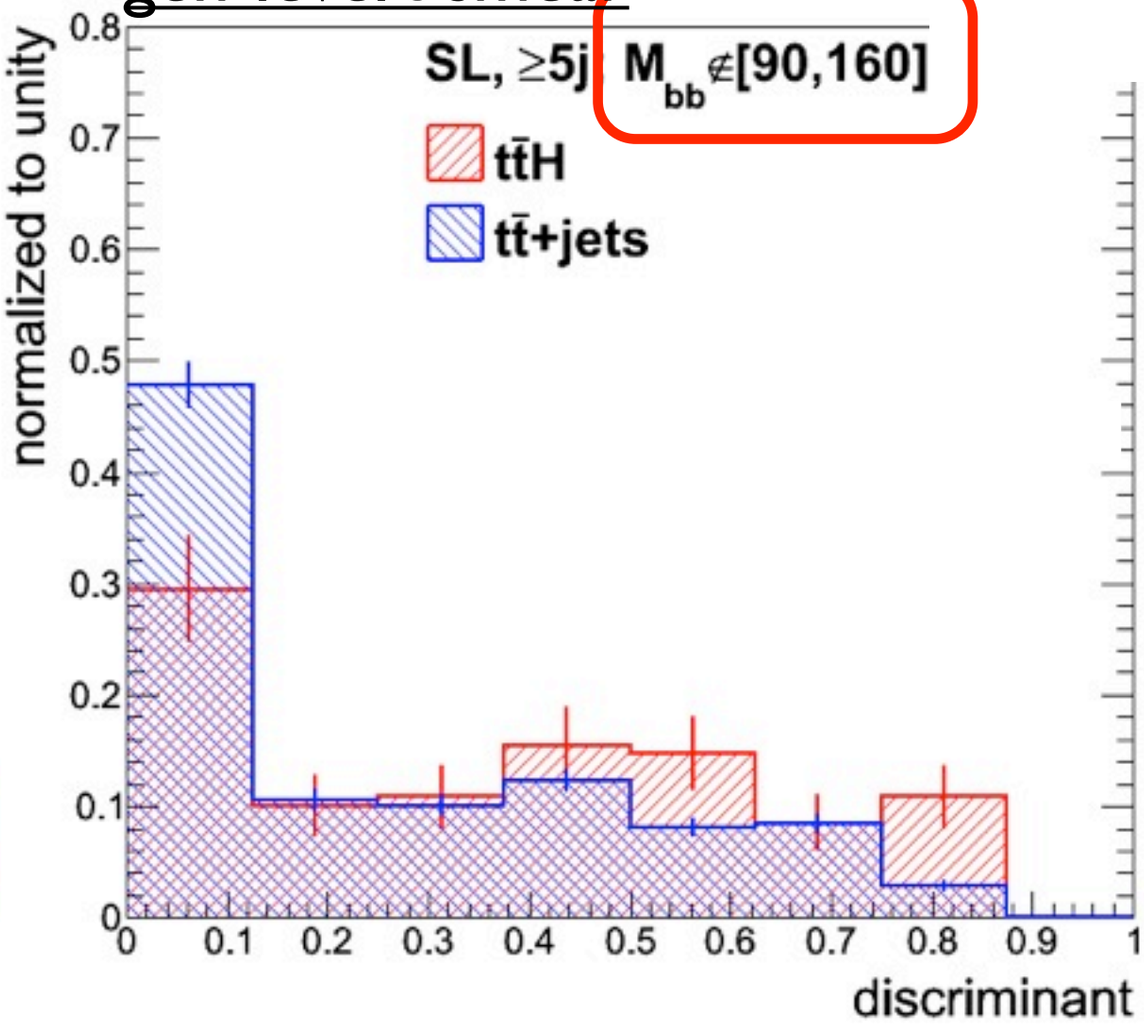


Digression I: the Higgs mass

gen-level⊗smear



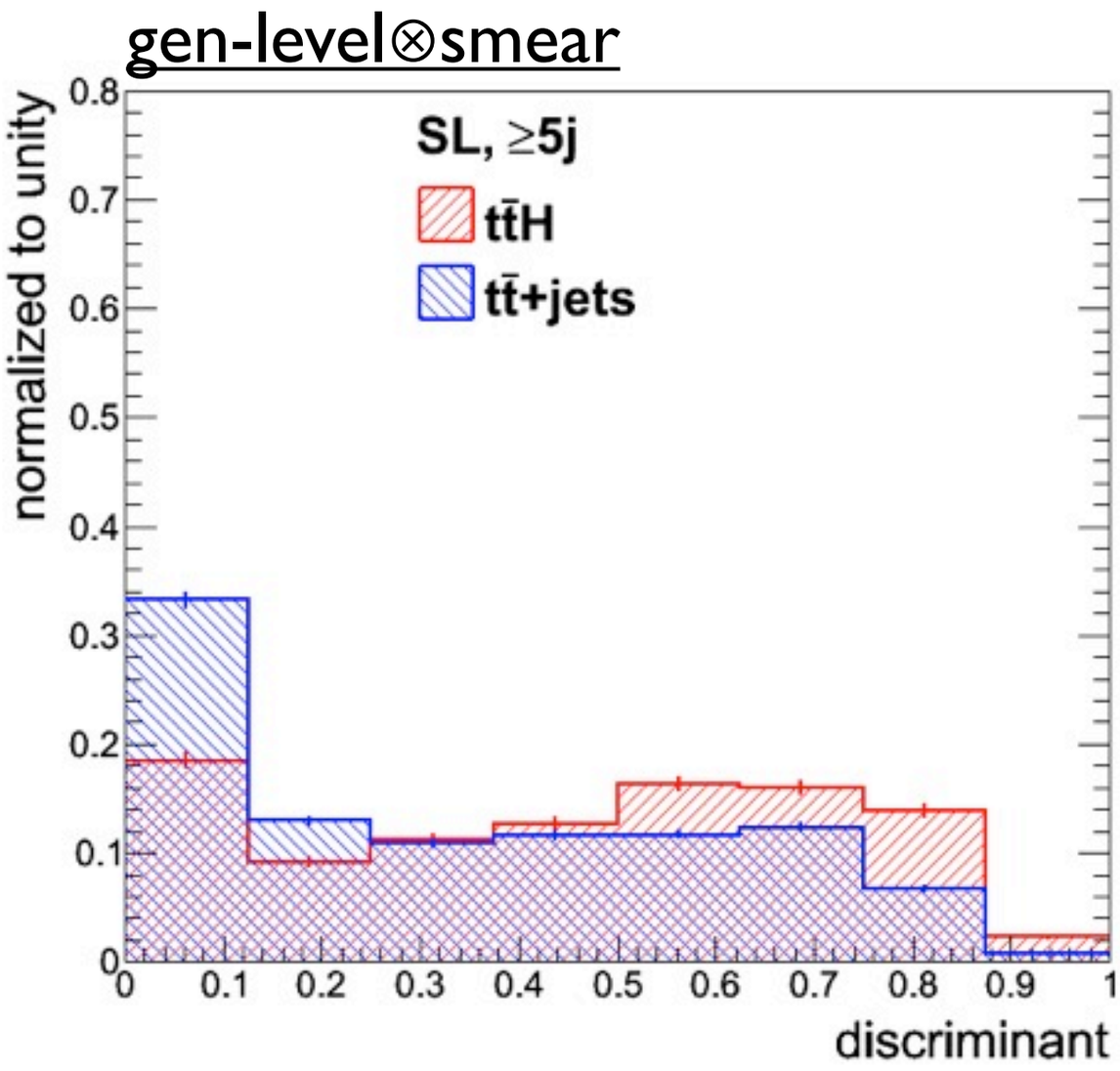
gen-level⊗smear



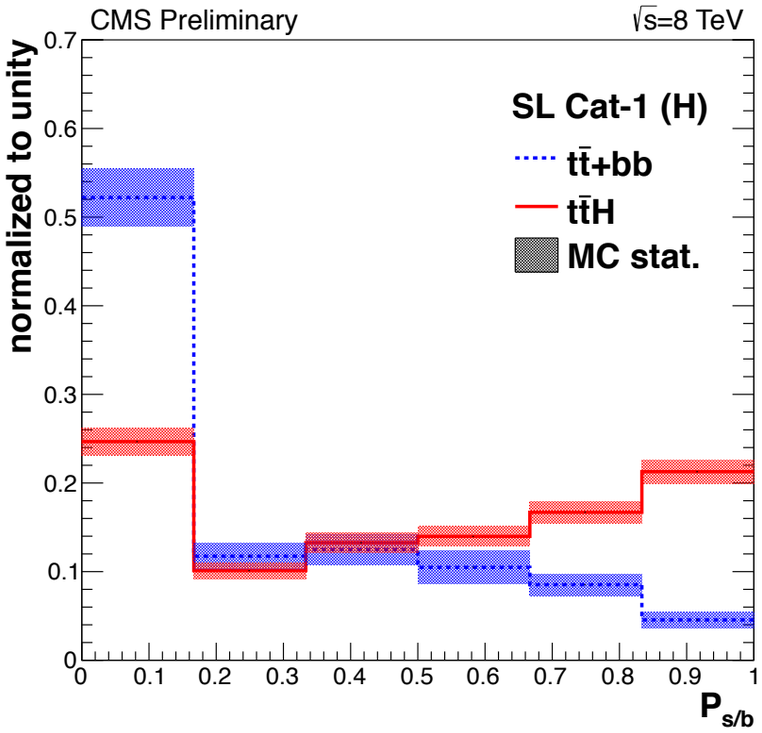
- *ttbb* events w/ M(*bb*) ≈ 125 indeed look like *ttH*!
  - but not identical ⇔ the ME is sensitive also to the other variables
- *ttH* events w/ M(*bb*) ≠ 125 (e.g. poor resolution) undistinguishable from *ttbb*

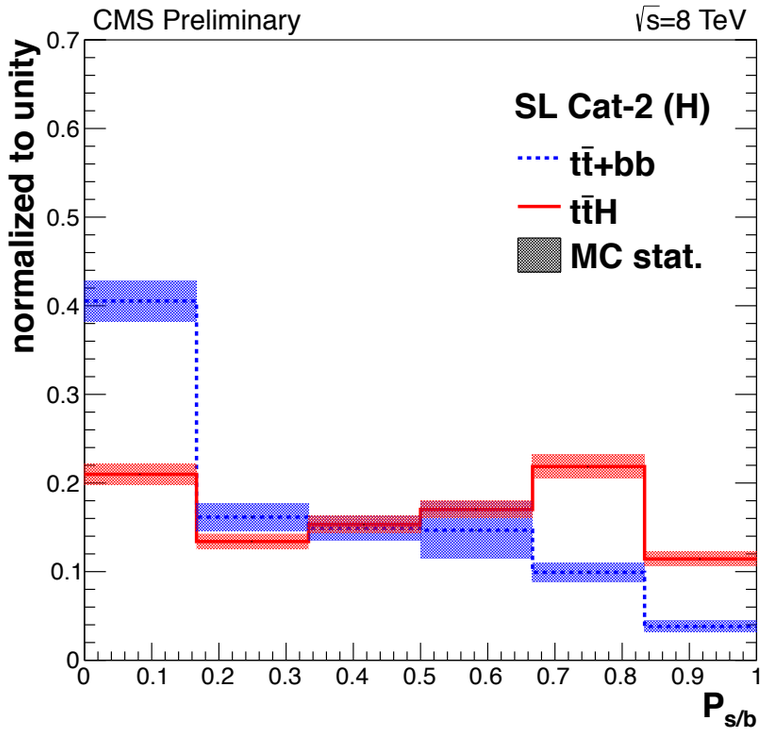
Digression II: wrong hypothesis

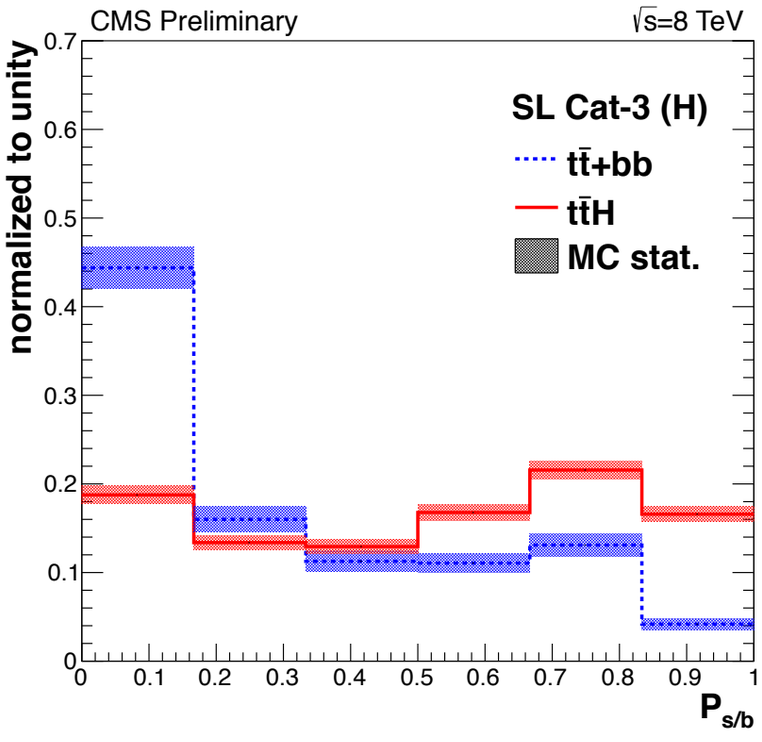
- If the event does not fulfill the *tested* ME hypo, the weight is broadly distributed
  - yet, *ttH* remains slightly more "signal-like"

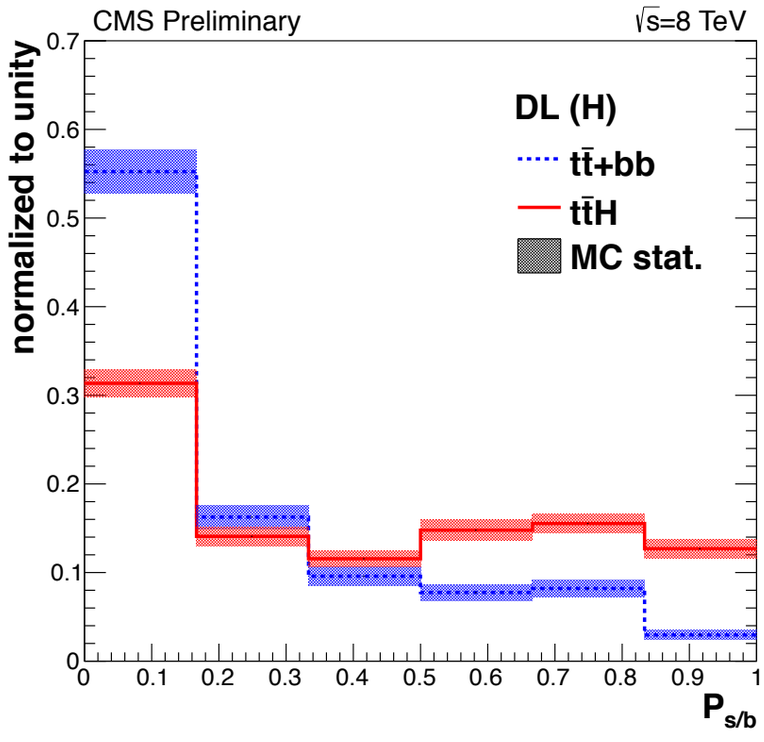


Shape comparison: tt+bb vs ttH









DLM continued

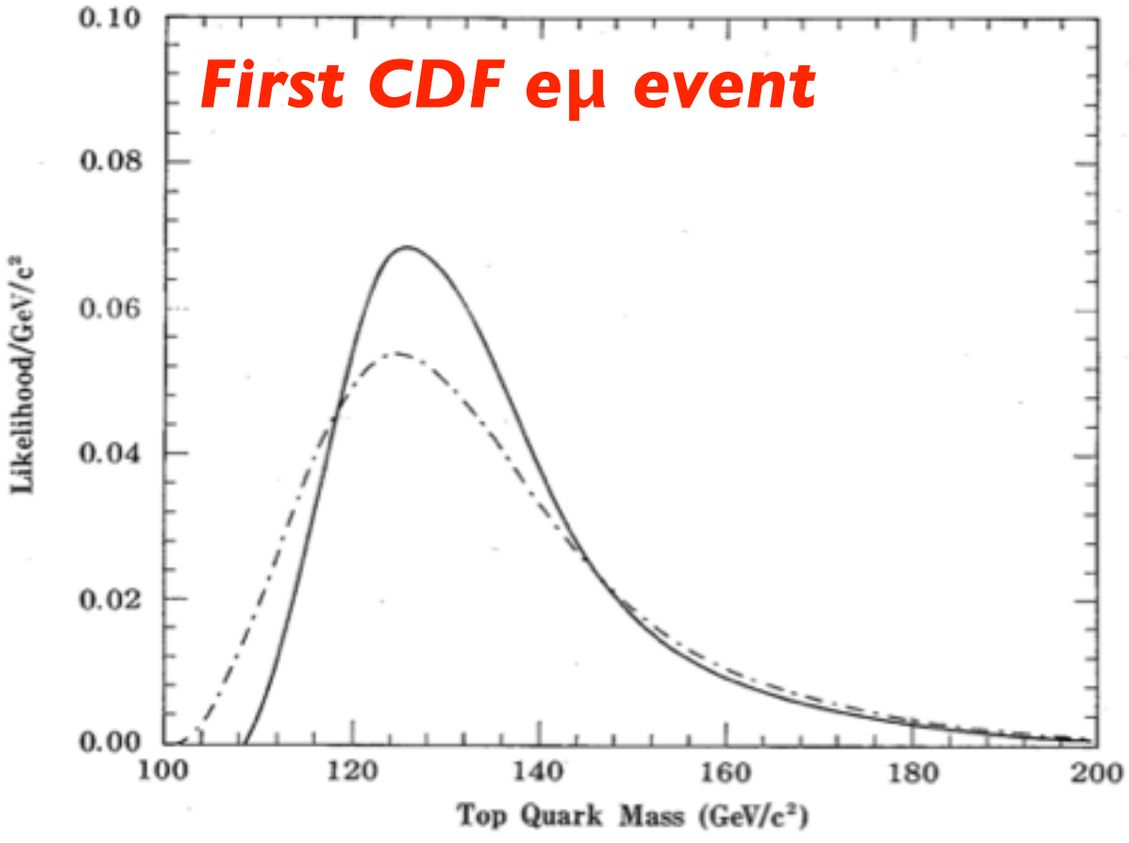
First complete formulation of a dynamical mass likelihood

- A +B → C + D → invisible

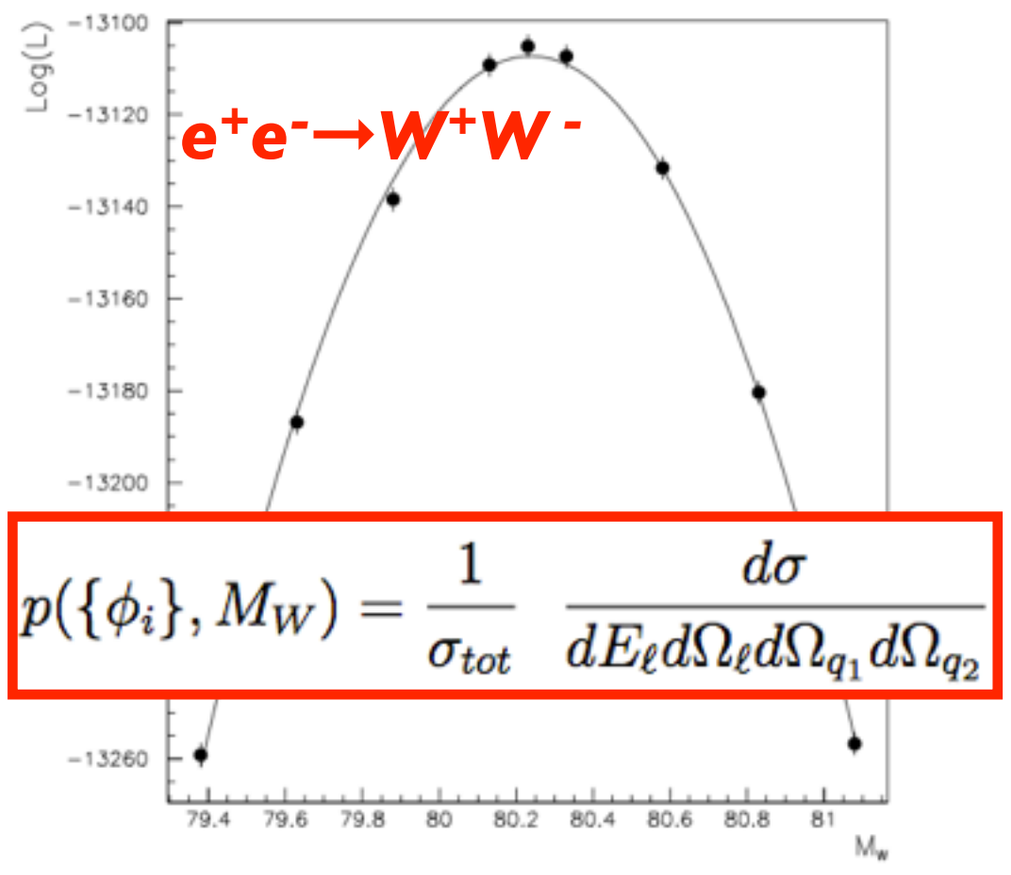
K. Kondo,
J. Phys. Soc. Jap 60 836 (1991)

$$\mathrm{d}P = N' \frac{1}{|(A\cdot B)|}\delta^4(A+B-C-D)|M|^2\cdot\left(\prod_{i=1}^{l}\frac{1}{c_{i0}}\prod_{j=1}^{m}\frac{1}{d_j}\right)\mathrm{d}\Xi\,\mathrm{d}\zeta_f^*$$
$$\times F_{1A}(A_x, A_y, z_a, Q^2)F_{2B}(B_xB_y, z_b, Q^2)\,\mathrm{d}A_x\,\mathrm{d}A_y\,\mathrm{d}z_a\,\mathrm{d}B_x\,\mathrm{d}B_y\,\mathrm{d}z_b\cdot\mathrm{d}m_c\,\mathrm{d}m_d,$$

K. Kondo, J. Phys. Soc. Jap 62 1177 (1993)



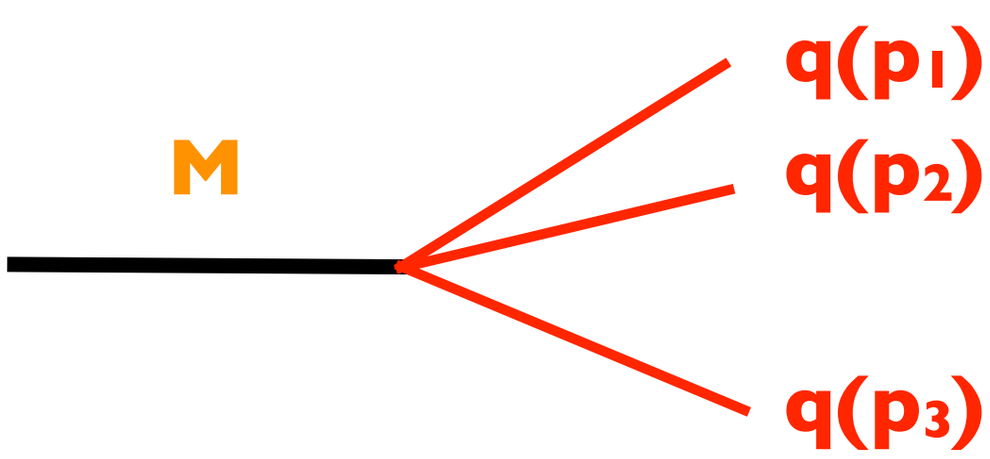
F. Berends et al., PLB 417 (1998) 385

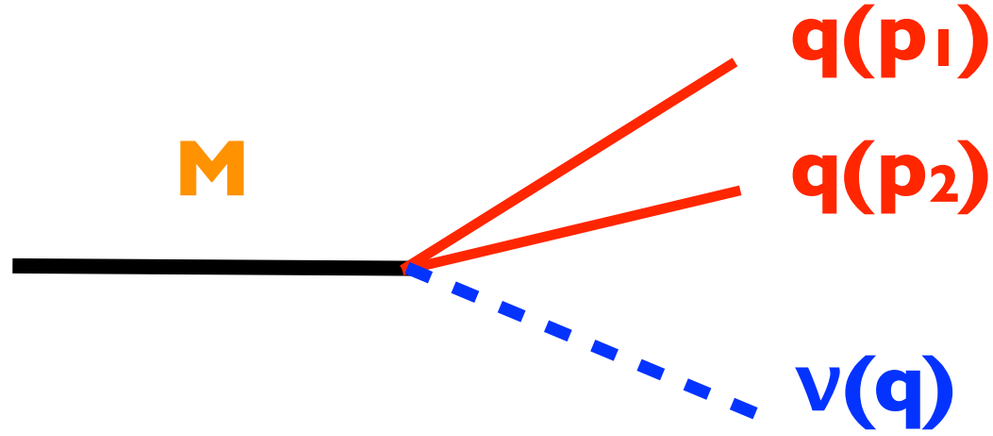


$$p(\{\phi_i\}, M_W) = \frac{1}{\sigma_{tot}}\frac{d\sigma}{dE_\ell d\Omega_\ell d\Omega_{q_1} d\Omega_{q_2}}$$

A toy example

K. Kondo, J. Phys. Soc. Jap 57 4126 (1988)
K. Kondo, J. Phys. Soc. Jap 60 836 (1991)





$$M = \sqrt{(p_1 + p_2 + p_3)^2}$$

$$M(\vec{q}) = \sqrt{(p_1 + p_2 + q)^2}$$

If decay governed by some amplitude, not all configurations equally likely:

$$dP(\vec{q}) = |\mathcal{M}(p_1, p_2, q)|^2 d\vec{q}$$

Recover full kinematics by applying the concept of probability:

$$\frac{dP}{dM}(M) = \int d\vec{q} |\mathcal{M}(p_1, p_2, q)|^2 \times \delta(M - M(p_1, p_2, q))$$

The event likelihood

Complexity of the problem scales with dimensions

- if both dim(**Y**) and dim(θ) small ⇒ MC methods
  - e.g. fill histograms using MC simulation at different θ values
- else MC methods impractical
  - analytical approach

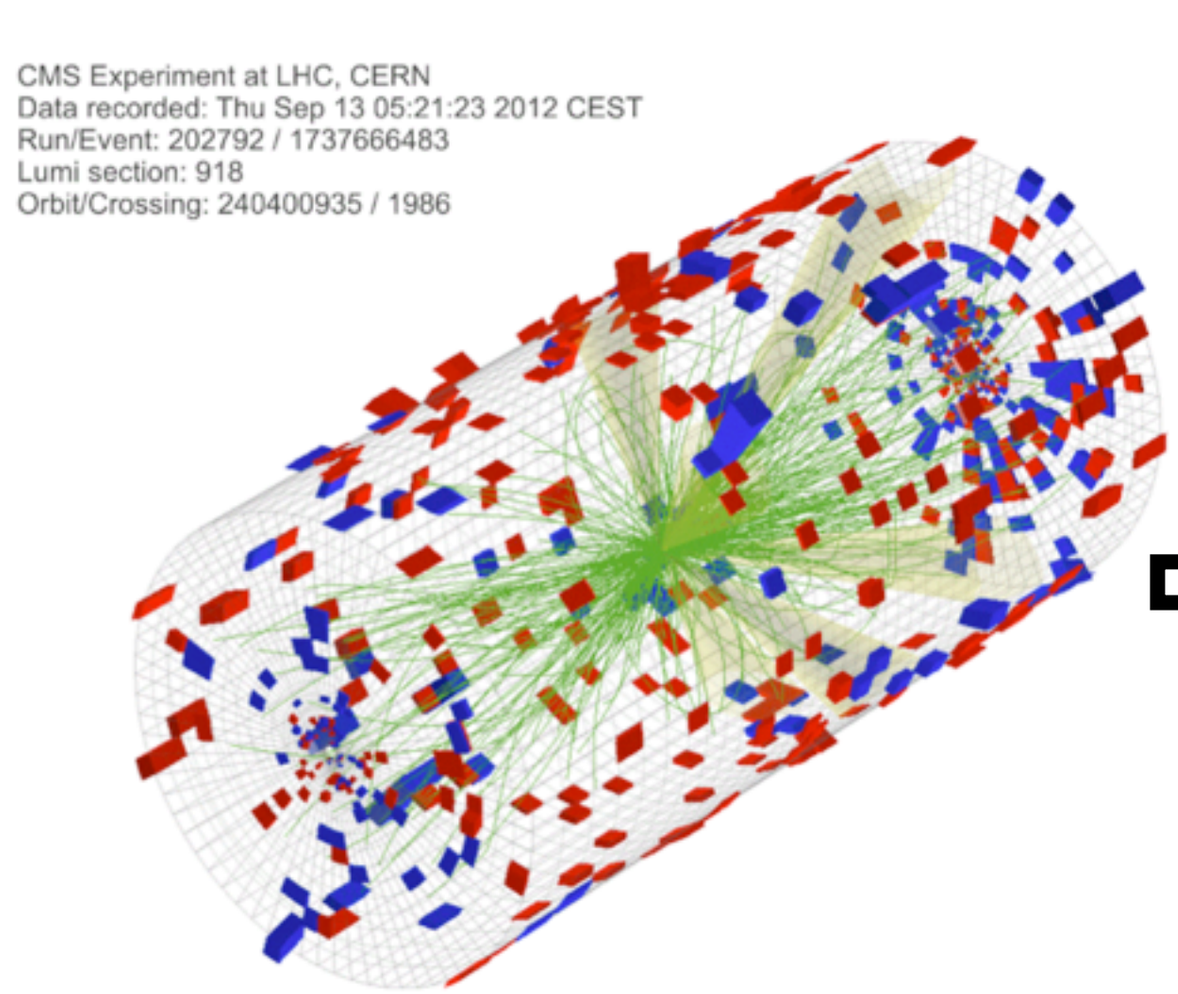
- A set of ancillary variables **X**
- A prior π(**X**;θ)
- A conditional pdf W(**Y**|**X**)

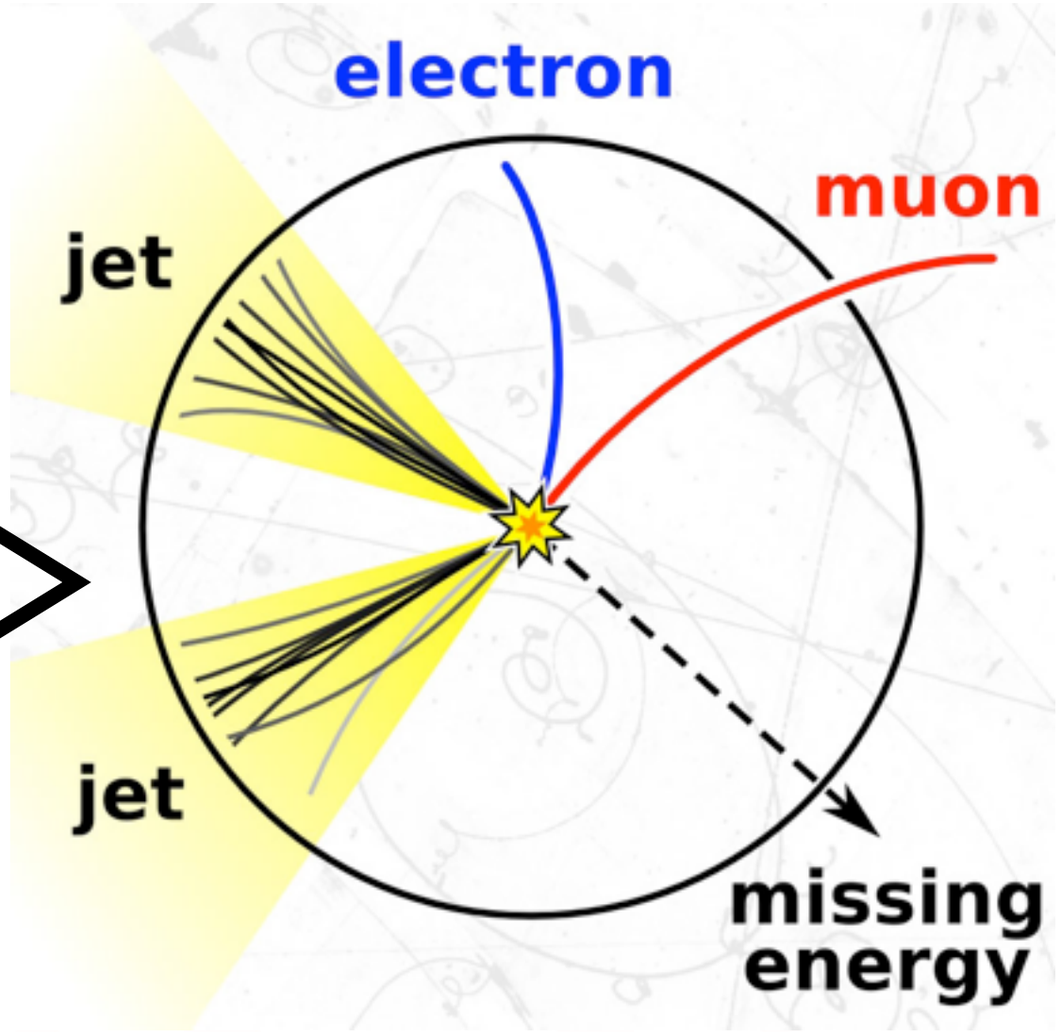
$$\mathcal{L}(Y;\theta) = \int \pi(X;\theta) \cdot W(Y|X) dX$$

Encorporates all of the prior knowledge

any new measurement adds a bit of info

The event likelihood





Clearly not feasible to determine a full "detector likelihood"

- but detector can be reduced down to the level of "physics objects"
  - proxies of ideal asymptotic states (quarks, leptons, ...)
  - still, every particle is at least a 4-dimensional variate

Dynamical Likelihood Models

Problem: retain maximal information on dynamical parameters even in presence of non-reconstructed particles

- especially relevant at hadron colliders

$$\frac{dP}{dX_m} = \int dX_n \pi(X_m, X_n)$$

Probability density on the best measured variables

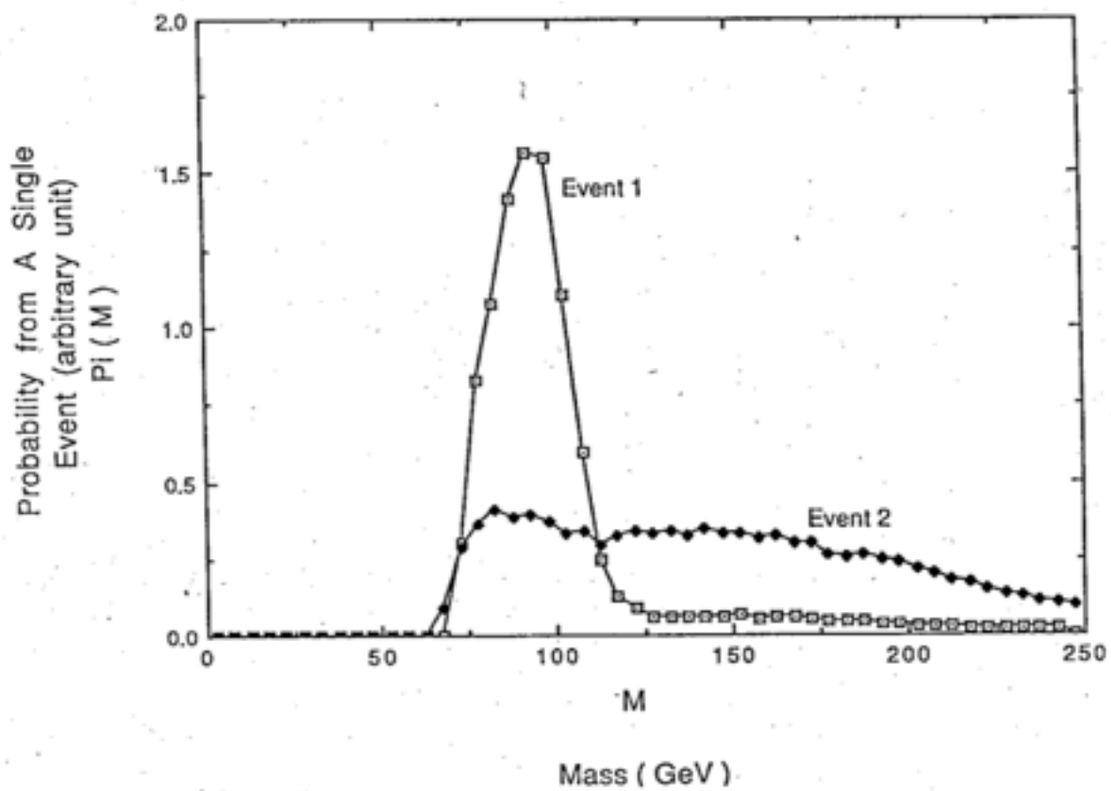
Integrate out poorly measured ones using their theory prior

Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. I. Method and Toy Models

Kunitaka KONDO

University of Tsukuba, Ibaraki 305

(Received July 22, 1988)



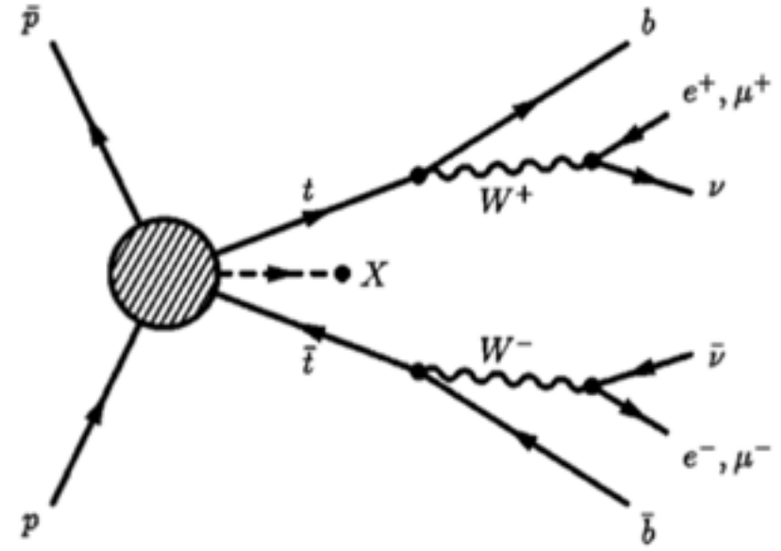
K. Kondo, J. Phys. Soc. Jap 57 4126 (1988)
K. Kondo, J. Phys. Soc. Jap 60 836 (1991)

DLM continued

First six di-lepton events in D0

- possibility of using full mass likelihood mentioned, but not yet pursued

D0, PRD 60, 052001 (1999)



$$\mathcal{P}(\{o\}|m_t) \propto \int f(x) f(x) |\mathcal{M}|^2 p(\{o\}|\{v\}) \, \delta^4 \mathrm{d}^{18}\{v\} \mathrm{d}x \mathrm{d}x, \tag{6}$$

> Unfortunately this expression involves a multidimensional integral that has to be evaluated numerically and is complicated by the need to include initial and final state gluon radiation. Such higher order effects complicate the reconstruction of the top quark mass substantially and cannot be neglected. We therefore do not attempt to compute the exact probability density given in Eq. (6). Rather, we construct simpler weights that retain sensitivity to the value of the top quark mass but can be evaluated with the available computing resources. We calibrate the effect of the simplifi-

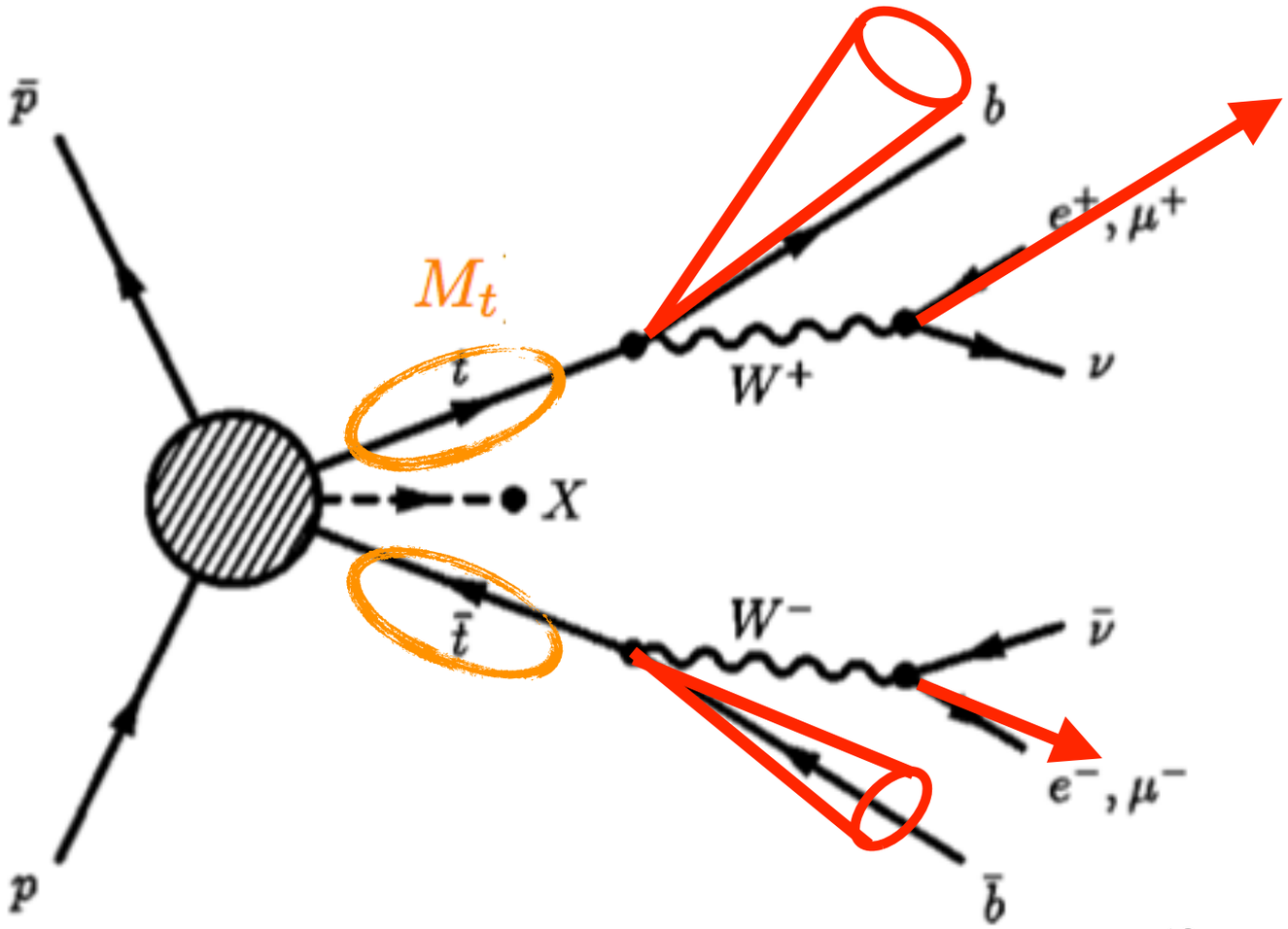
An example: top mass

Y = 4-vectors of jets and leptons

M_t = top quark mass

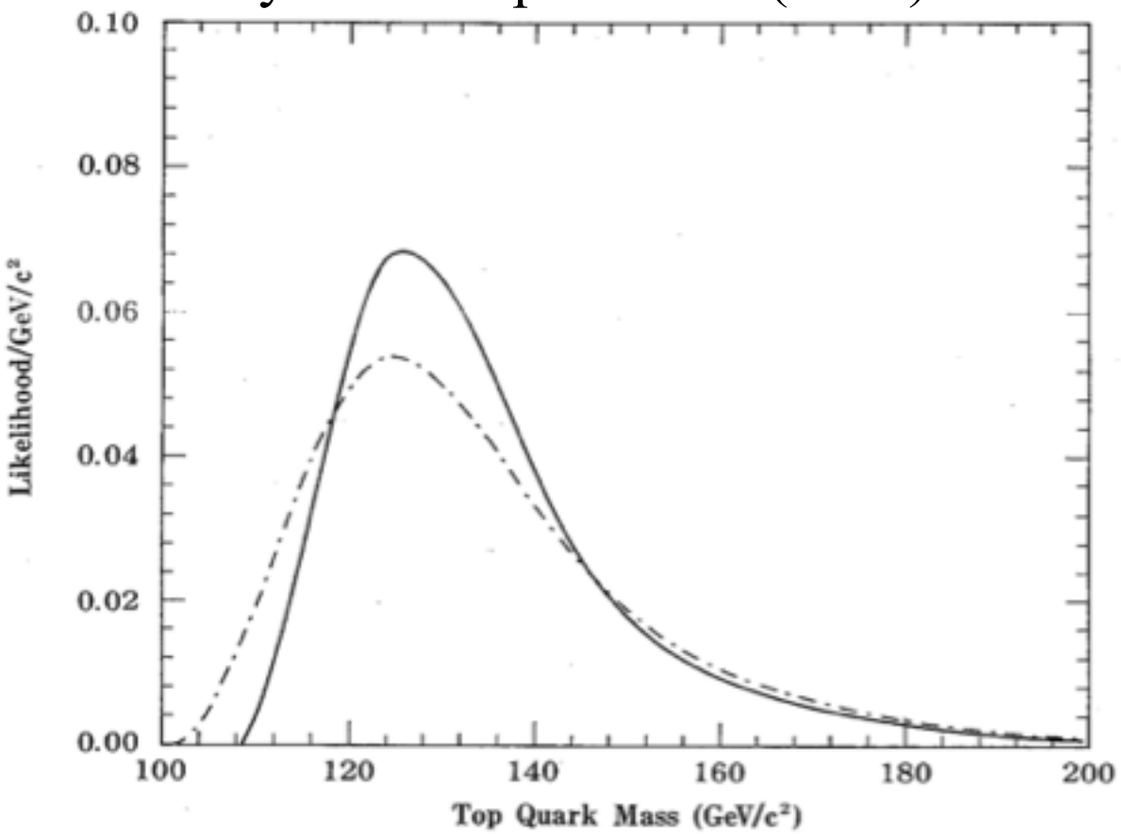
X = 4-vectors of the 2 → 6 scattering (ancillary variables)

W(**Y**|**X**) = detector response



Dynamical Likelihood Method for Reconstruction of Events with Missing Momentum. III. Analysis of a CDF High P_T $e\mu$ Event as $t\bar{t}$ Production

K. Kondo et al.,
J. Phys. Soc. Jap 62 1177 (1993)



A typical problem

Typical problem:

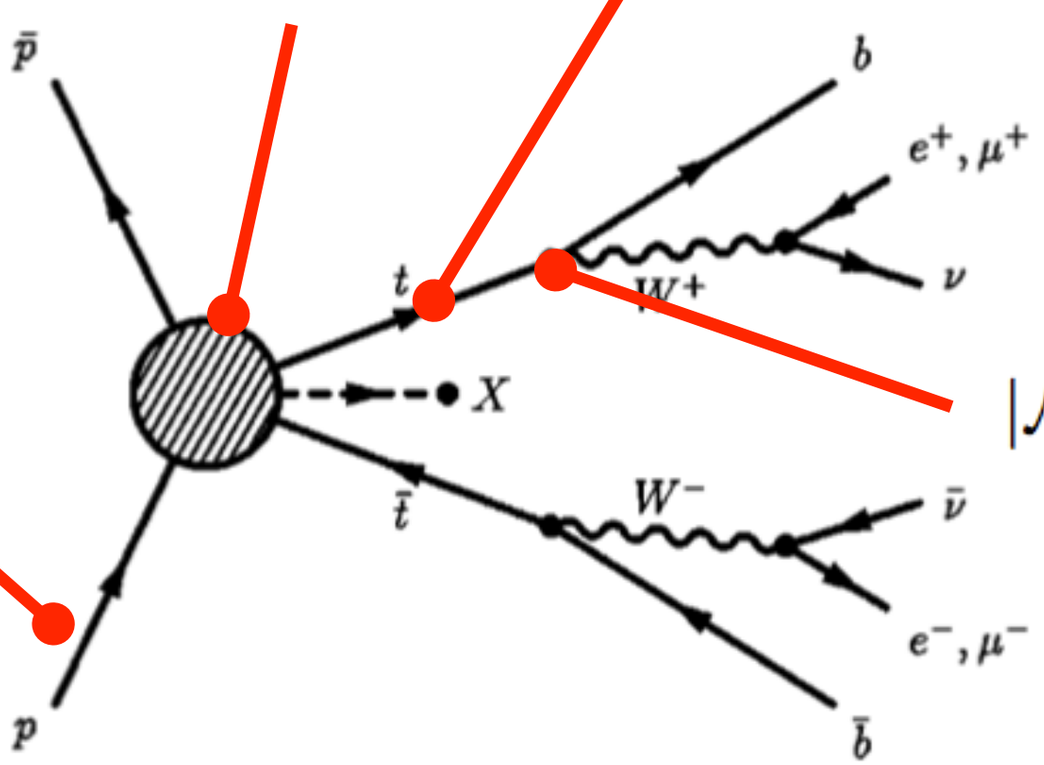
"Mass of resonances decaying to invisible or poorly reconstructed particles"

X = 4-vector of scattering particles

W(**Y**|**X**) = detector response

$$\frac{1}{(t^2 - M_t{}^2)^2 + \Gamma_t^2 M_t{}^2}$$

$$|\mathcal{M}(gg \to t\bar{t}; M_t)|^2$$



$$f_p(x_1(M_t); Q) f_{\bar{p}}(x_2(M_t); Q)$$

$$|\mathcal{M}(t \to b\ell\nu; M_t)|^2$$

$$dP(\mathbf{X}; M_t) \propto f_p(x_1; Q) f_{\bar{p}}(x_2; Q) \, |\mathcal{M}(gg \to \ell\nu b\ell\nu b; M_t)|^2 \, d\mathbf{X}$$

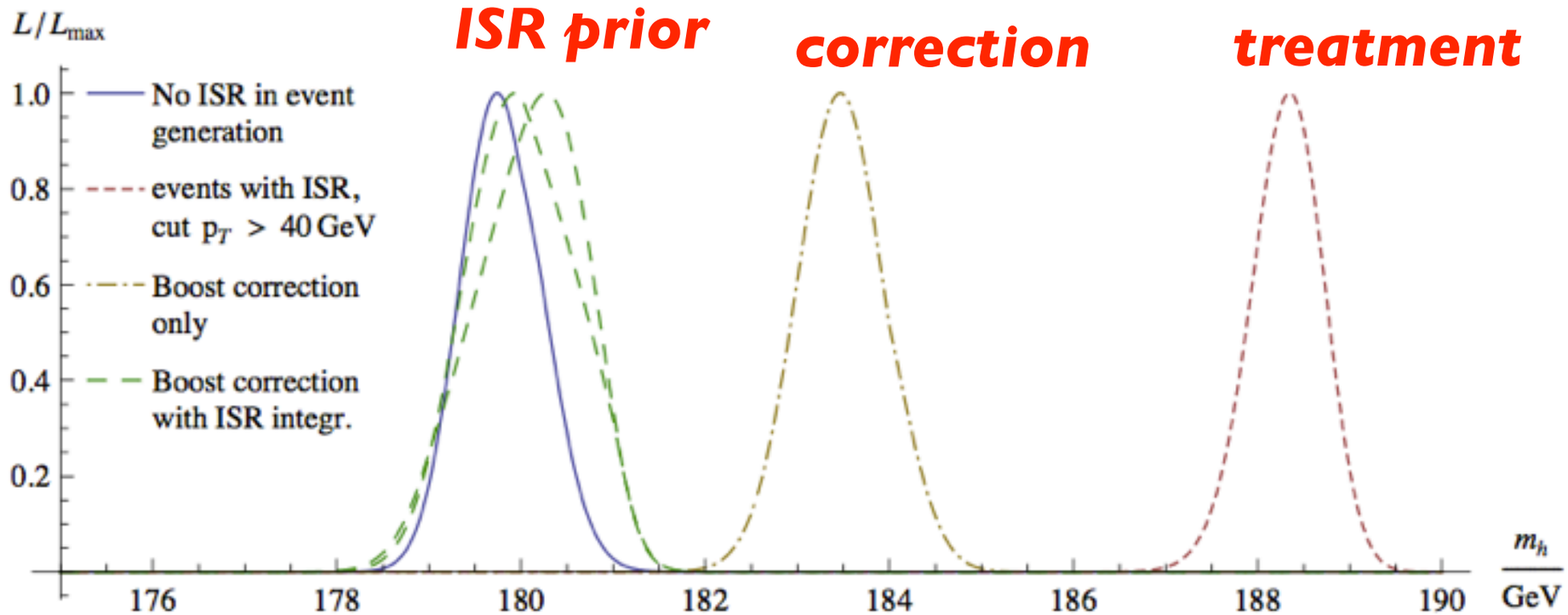
Summary

Strengths	Weaknesses
Analyticity densities better handling of systematics training-free	**Complexity** CPU time variable transformation
Information maximal usage of theory model factorisation theory ⊗ experiment	**Scalability** multiple-channel processes combinatorics
Properties optimal performances	**Fixed order accuracy** bias loss of information

QCD radiation

Dealing with leading order ME and ISR

- assuming P_T balance in real events can be inefficient and/or biasing
- heuristic approaches based on
  - simple boost, Sudakov reweighting, p_T priors



J. Alwall et al., PRD 83 (2011) 074010
CDF, PRD 79 (2009) 072101

The integral

Calculation of $d\sigma$ requires multi-dimensional integral

- numerical integration through MC methods
  - integrand often rapidly changing ⟺ adaptive MC's (e.g. VEGAS)

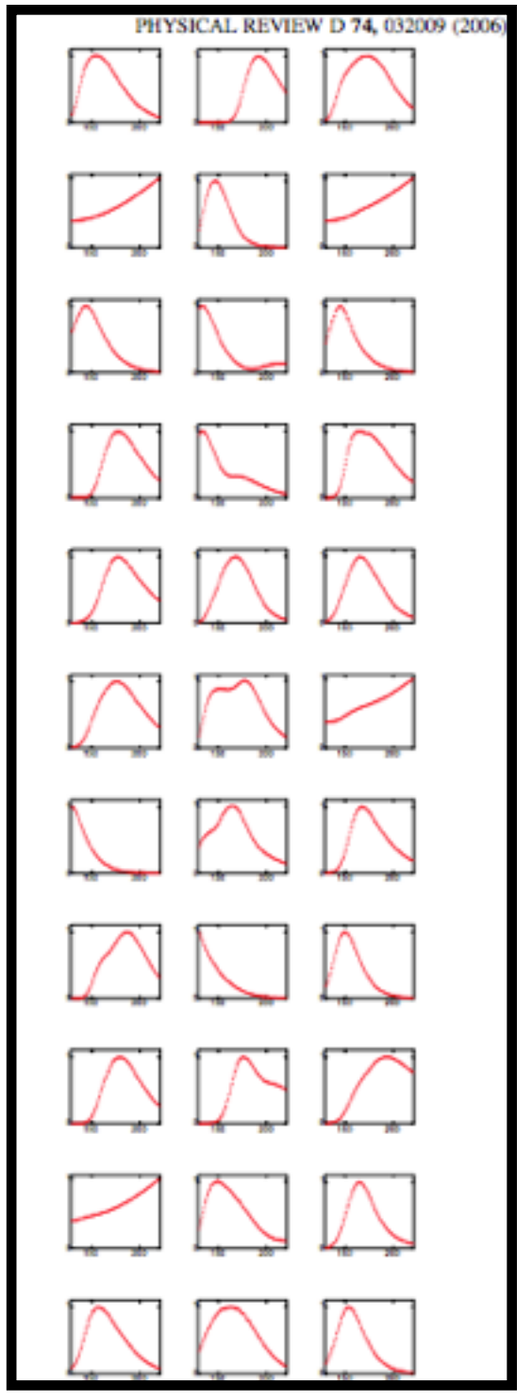
$$\int d\Phi(\vec{x}) dx_a dx_b \sum_{i,j} \frac{f_i(x_a) f_j(x_b)}{(1+\delta_{ij}) x_a x_b s} |\mathcal{M}_S(\vec{x},\theta)|^2$$

$$\prod_i \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i}$$

PDF	S-matrix
• Parametrisations - fast evaluation	• Analytical expressions - whenever manageable
• Discretised data files - slower - larger samples of PDF (e.g. LHAPDF)	• Numerical routines - e.g. MG5, MCFM - CPU-intensive for 2 → N processes with N>>1

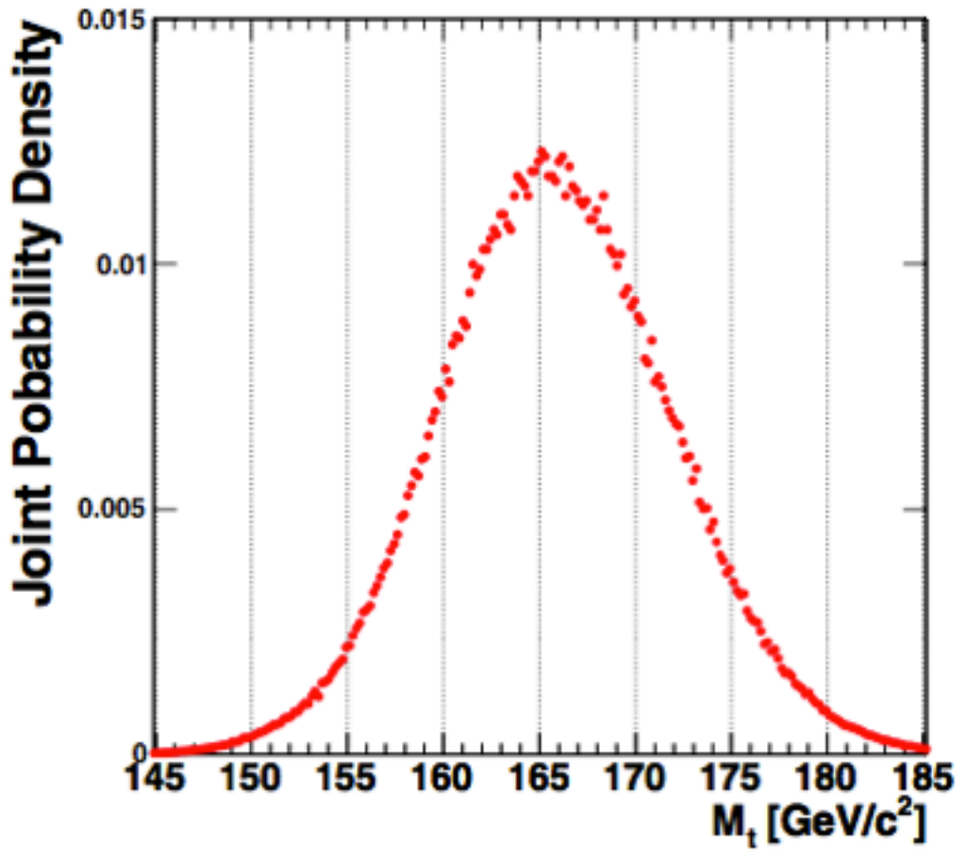
Sample likelihood

Event likelihoods:



CDF, PRD 74, 032009 (2006)

Sample likelihood:



Dimensional reduction

Calculation of *p(y)* requires multi-dimensional numerical integral

- reduce dimensionality using approximations
  - often a trade-off between CPU-cost and performances

1. without ISR, initial partons have longitudinal momenta

$$d\Phi(\vec{x}) = (2\pi)^4 \delta(x_a P_a + x_b P_b - \sum_i p_i) \left[\prod_i \frac{d^3\vec{p}_i}{(2\pi)^3 2E_i} \right]$$

2 inital partons (x_a, x_b) + **N final-state particles**

$$P_{x,y} = 0$$

$$x_{a,b} = \frac{E \pm P_z}{\sqrt{s}}$$

3N-2 variables

2. narrow-width approximation

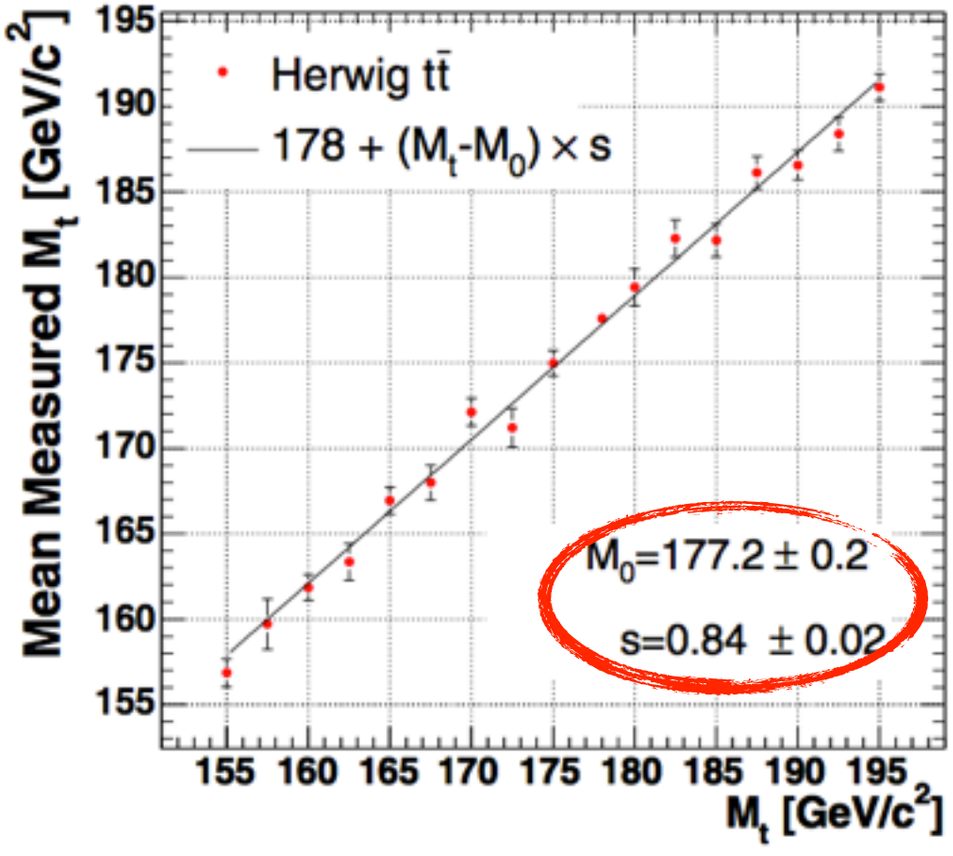
$$\frac{1}{(s^2 - M^2)^2 + \Gamma^2 M^2} \to \frac{\pi}{M\Gamma} \delta(s^2 - M^2)$$

Systematics and bias

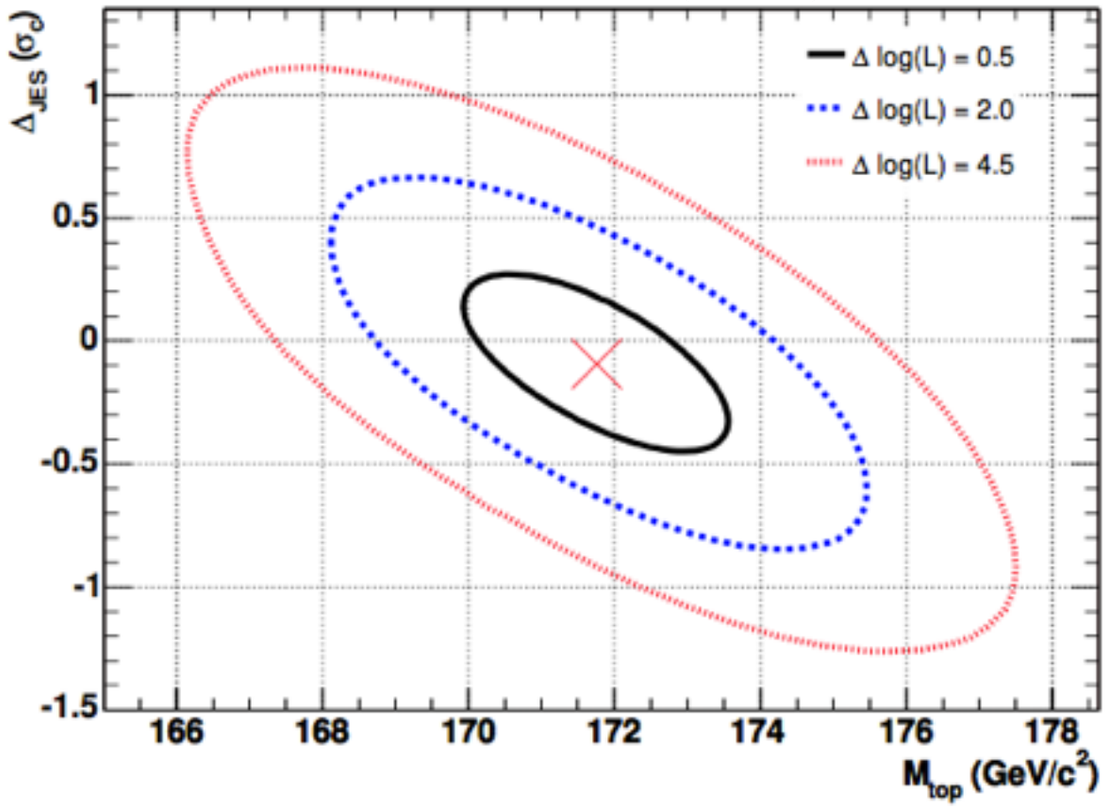
Systematic uncertainties:

- included as nuisance parameters
- calculated event-by-event
  - each event contributes with optimal weight

CDF, PRD 74, 032009 (2006)



CDF, PRD 79, 092005 (2009)



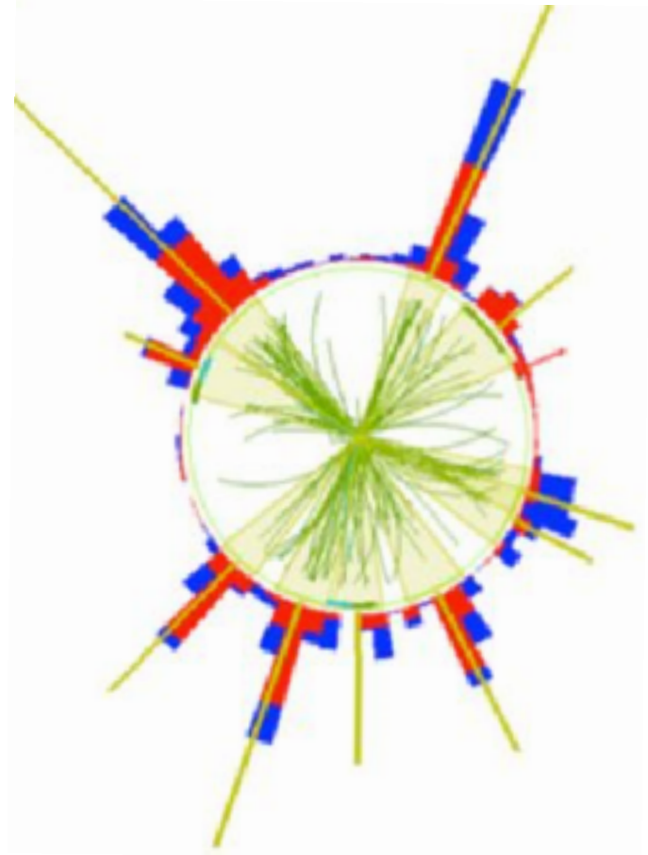
Bias:

- approximations (underlying model, detector simulation)
- MC calibration requires
  - approximations in MEM don't hamper its applicability, just make it less optimal

Jet/quark association

Jet ⇔ quark ambiguity ⇒ multiple associations

- growing as the factorial of the jet multiplicity
- ambiguity can be reduced by
  - jet ⇔ flavour assignment (e.g. b-tagging)
  - quark-exchange symmetry
- ambiguity increased by ISR/FSR and acceptance



DL ***SL*** ***# reco's light-quarks***

# of jets (4 tags)

	$N_q = 0$	$N_q = 1$	$N_q = 2$	$N_q = 3$	$N_q = 4$
$J = 4$	12	–	–	–	–
$J = 5$	12	12	–	–	–
$J = 6$	12	24	12	–	–
$J = 7$	12	36	36	36	–
$J = 8$	12	48	72	144	72

Ideogram

$$\mathcal{L}_{\rm evt}\left(x; m_t, JES, f_{\rm top}\right) = f_{\rm top} \cdot P_{\rm sgn}\left(x;\, m_t, JES\right) + \left(1 - f_{\rm top}\right) \cdot P_{\rm bkg}\left(x;\, JES\right) . \quad (4)$$

$$w_i = \exp(-\frac{1}{2}\chi_i^2) \cdot w_{{\rm btag},i}.$$

$$P_{\rm sgn}\left(x_{\rm fit};\, m_t, JES\right) \equiv \sum_{i=1}^{24} w_i \left[f_{\rm correct}^{\rm ntag} \cdot \int_{m_{\rm max}}^{m_{\rm min}} \mathbf{G}(m_i, m', \sigma_i) \cdot \mathbf{BW}(m', m_t) dm' + \left(1 - f_{\rm correct}^{\rm ntag}\right) \cdot S_{wrong}^{\rm ntag}(m_i, m_t) \right]. \quad (9)$$

Scattering amplitude

OpenLoops (LO)

$$|\mathcal{M}_{\mathrm{S}}(g_1, g_2, q_1, \bar{q}_1, b_1, q_2, \bar{q}_2, b_2, b, \bar{b})|^2 = |\mathcal{M}(g_1, g_2, t, \bar{t}, h)|^2 \times$$

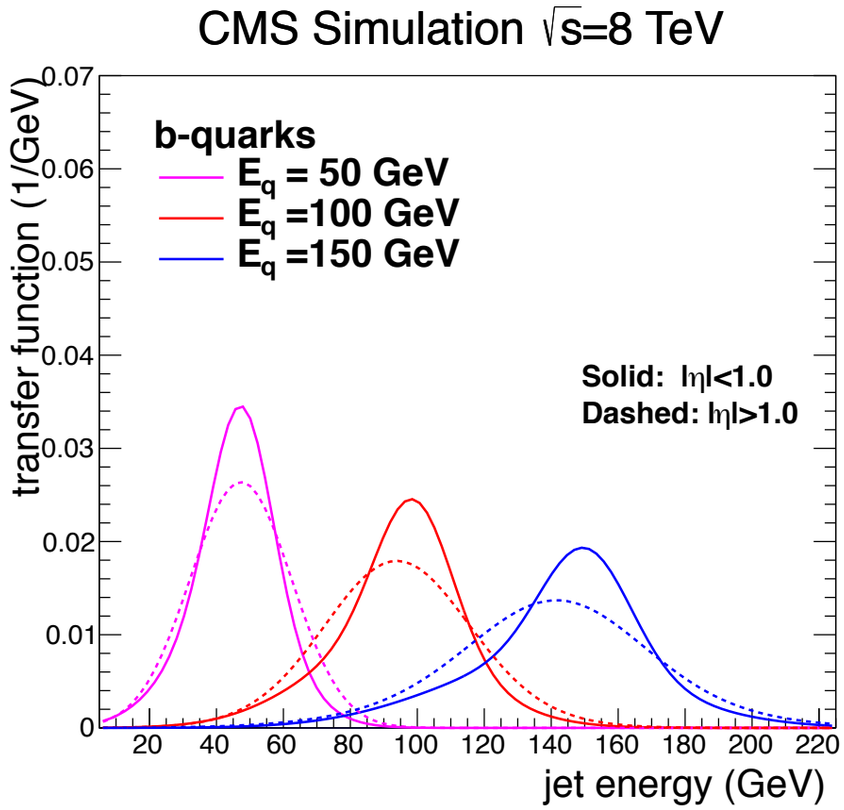
$$\times \prod_{i=1,2} \left[\frac{\pi}{M_t \Gamma_t} \delta(t_i^2 - M_t^2) |\mathcal{M}(q_i, \bar{q}_i, b_i)|^2 \right] \times \frac{\pi}{M_H \Gamma_H} \delta(h^2 - M_H^2) |\mathcal{M}(b, \bar{b})|^2$$

Narrow-width approximation

analytical (no spin-correlations)

Transfer function

$$
\begin{aligned}
\text{jet directions} &\equiv \Omega_j = (\cos\theta_j, \phi_j) \\
\text{jet energies} &\equiv E_j \\
\text{lepton directions} &\equiv \Omega_\ell = (\cos\theta_\ell, \phi_\ell) \\
\text{lepton energies} &\equiv E_\ell \\
\text{missing } p_T &\equiv \vec{\not{E}}_T \\
\text{transverse recoil} &\equiv \vec{\rho}_T = \\
&= -\vec{\not{E}}_T - \sum_{\ell \in \text{decay}} \vec{\ell}_T - \sum_{j \in \text{decay}} \vec{j}_T
\end{aligned}
$$



$$(2\pi)^4 \delta(x_a P_a + x_b P_b - \sum_i p_i)$$

$$(2\pi)^4 \, \delta^{(E,z)} \left(p_a + p_b - \sum_{k=1}^{8} p_k \right) \mathcal{R}^{(x,y)} \left(\vec{\rho}_T, \sum_{k=1}^{8} p_k \right)$$

x_a and x_b in MEM frame

transverse momentum constrained via TF ⇒ account for detector mismeauserement

An example

Example of ME implementation for a ttH search

- requirements: C++11 compiler, ROOT, LHAPDF
- ~3k lines of code
  - show how factorisation of MEM can be easily implemented in object-oriented languages
  - highlight features of ttH channel
- Run I analysis (19.5 fb^{-1}) required ~70k events (data+MC)
  - full analysis still manageable in ~1 night on a 3k nodes cluster

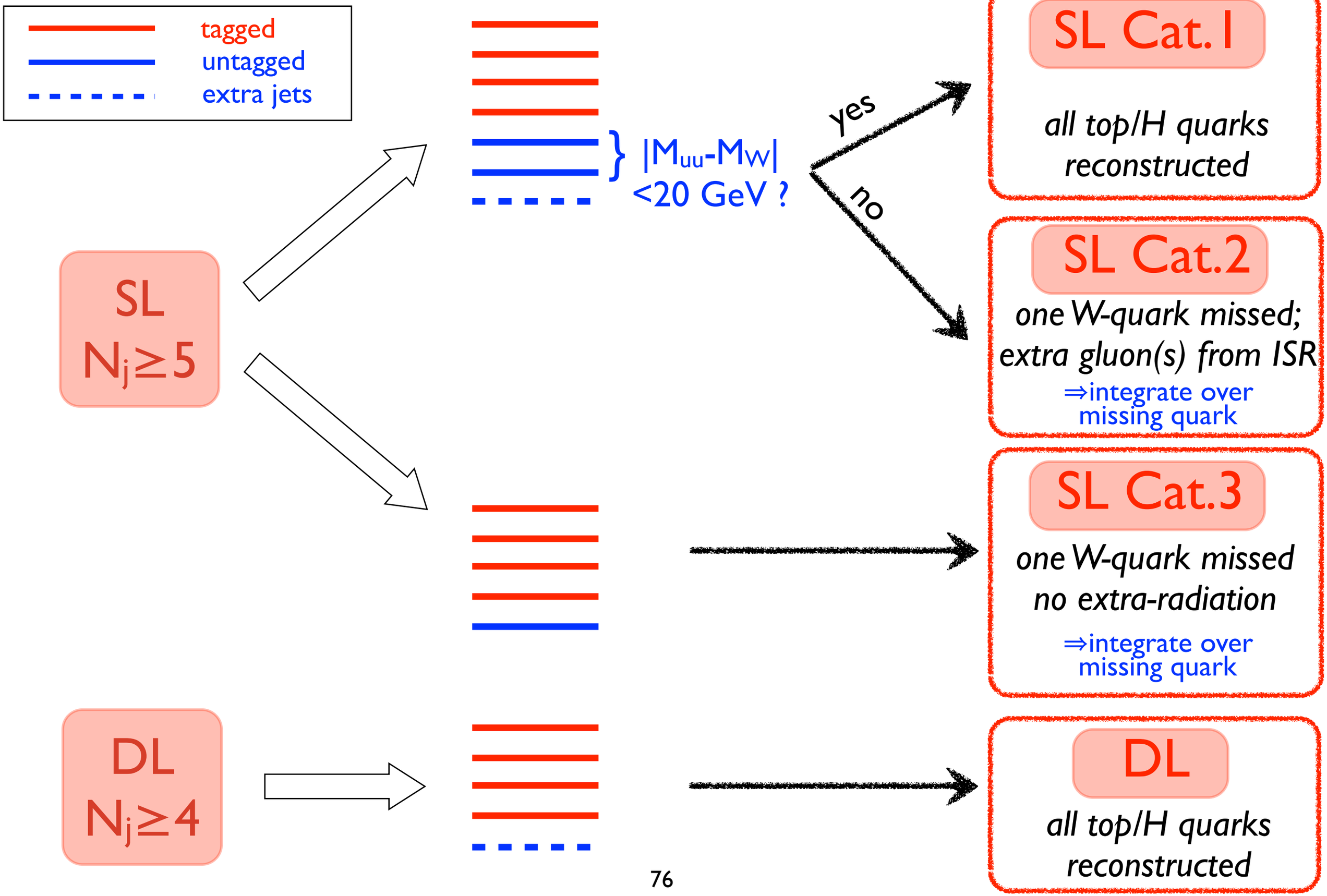
Number of variables	4 (+1)	6 (+1)	5 (+1)
Number of iterations	5	5	5
Function calls	2000	4000	10000
Numerical precision (mode of σ_w / w)	0.8%	1.2%	0.8%
CPU-time per integral (mean)	0.5 (1.5) s	1.1 (3.2) s	2.3 (6.2) s
Time budget for $\|\mathcal{M}\|^2_{\mathrm{ME}}$	30% (80%)	30% (80%)	30% (80%)

× # permutations (~ 12)

× # systematic shifts (~4)

CPU-time dominated $|\mathcal{M}_b|^2$

Event interpretation

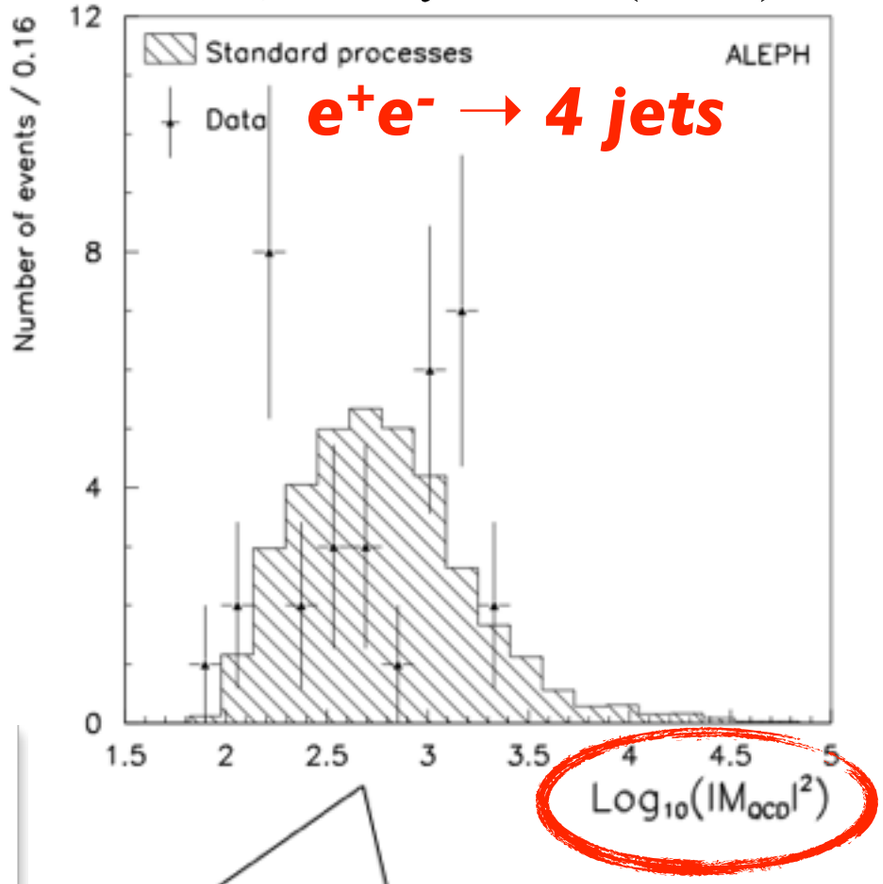


The ME-likelihood approach

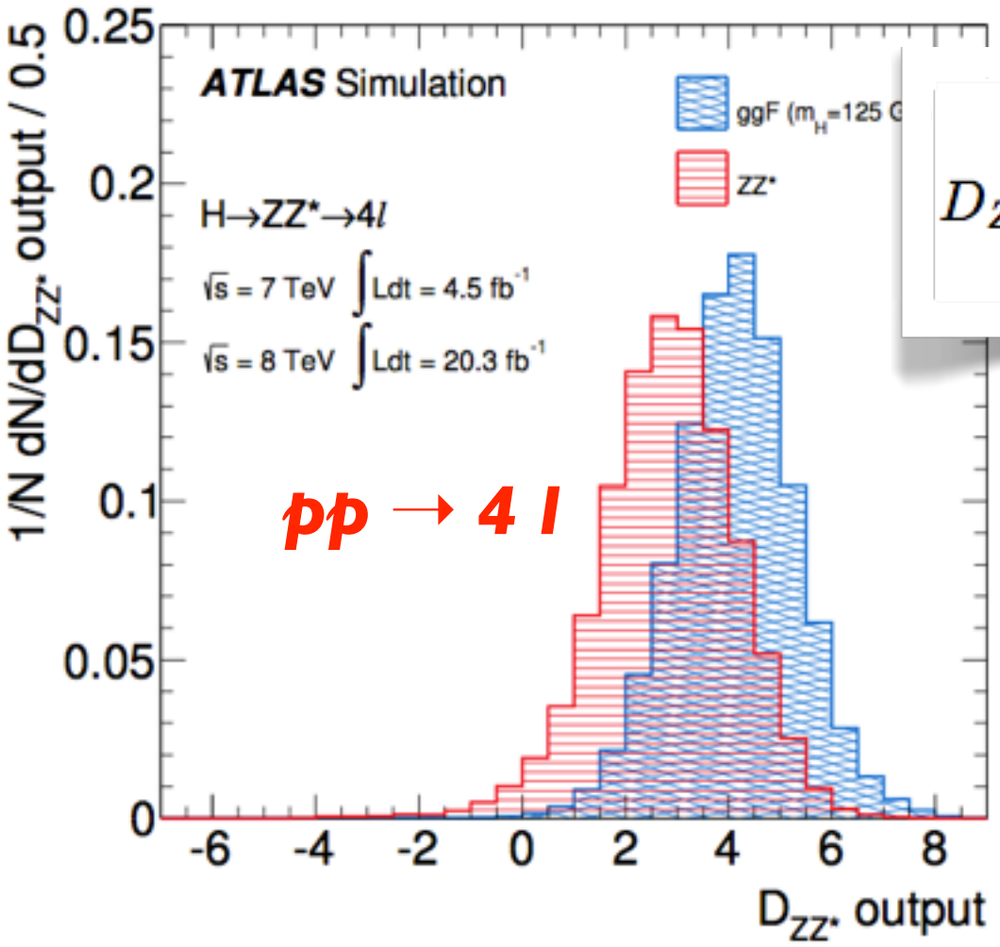
Using $|\mathscr{M}|^2$ as discriminating variable

- not a complete event likelihood though
  - no TF, no phase-space, no PDF

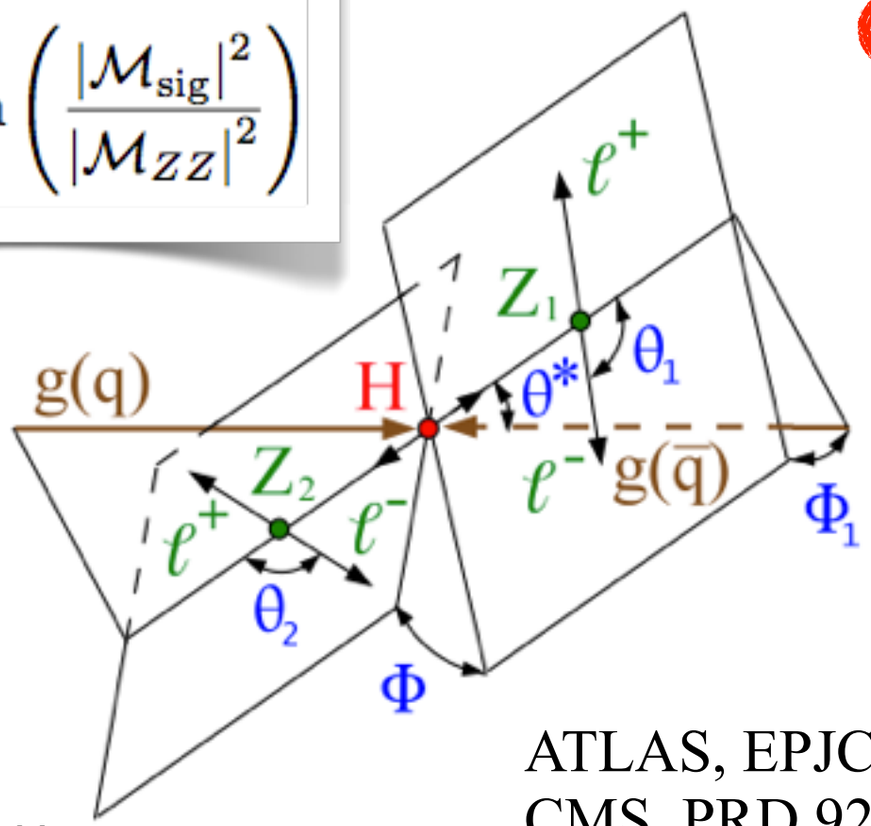
ALEPH, Z. Phys.C 71 (1997) 179



CMS, PRD 89 (2014) 092007
ATLAS, PRD 91 (2015) 012006



$$D_{ZZ^*} = \ln\left(\frac{|\mathcal{M}_{\text{sig}}|^2}{|\mathcal{M}_{ZZ}|^2}\right)$$



ATLAS, EPJC 75 (2015) 476
CMS, PRD 92 (2015) 012004

Future developments

Future developments on the MEM

- NLO / Parton showers
  - extending the MEM to NLO

J. Campbell et al., JHEP 11 (2012) 043
J. Alwall et al., PRD 83 (2011) 074010

- Jet substructure
  - Shower Deconstruction

D. Soper et al., PRD 84 (2011) 074002

- Underlying event, Hadronisation
  - Effective propagators

CDF, PRD 71 (2009) 072001

- Advanced computing
  - GPU
  - distributed systems

D. Schouten et al., Comp. Phys. Comm. 00 (2014) 1
L.B. et al., proceedings of CHEP2015

J. S. Gainer et al., arXiv:1307.3546