ETH Institute for Particle Physics

Unlocking the Higgs with the matrix element method

Joosep Pata

HiggsTools Journal Club

01.12.2015

ETH Zürich

higgstools

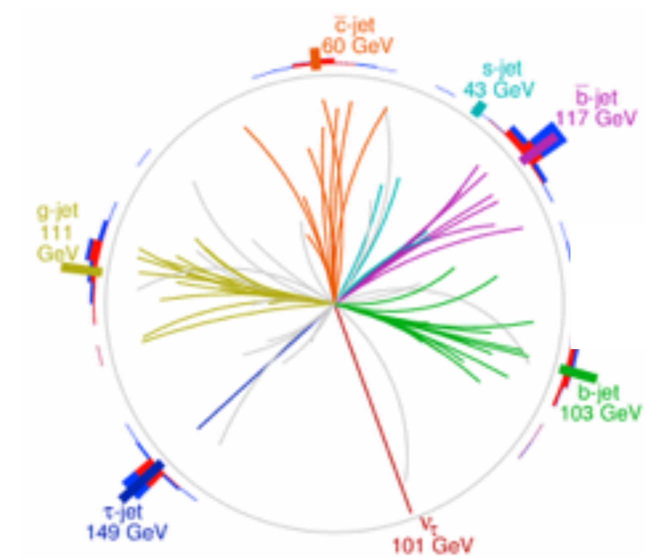
Overview

- Statistical context
- How we use the Matrix Element Method for hypothesis testing
- Why it is (specially) useful for Higgs studies
- Future developments of MEM

Not a lecture! Let's ask questions and discuss!

The statistical problem

- Given an observed event (data) **Y**, extract interesting quantities (e.g. theory parameters **Θ**)

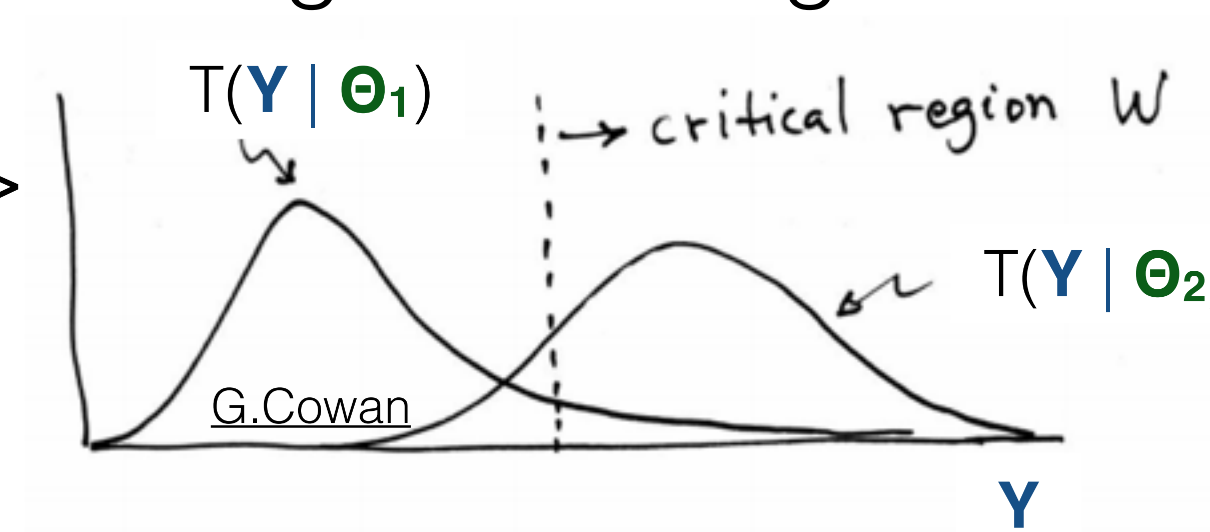


=**Y**

- sufficient statistic T(**Y**): no loss of information about data

“bkg” “signal”

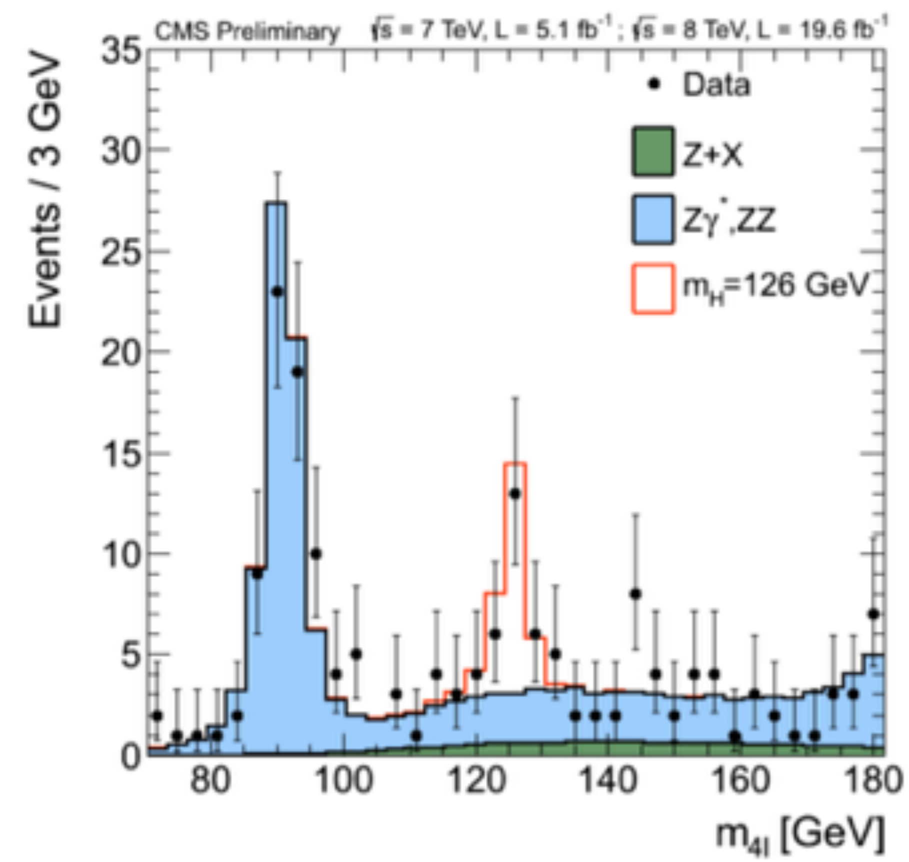
- maximize T(**Y** | **Θ**) wrt. **Θ** -> maximum likelihood is asymptotically sufficient



How to find a good test statistic T(**Y** | **Θ**)?

The usual approach

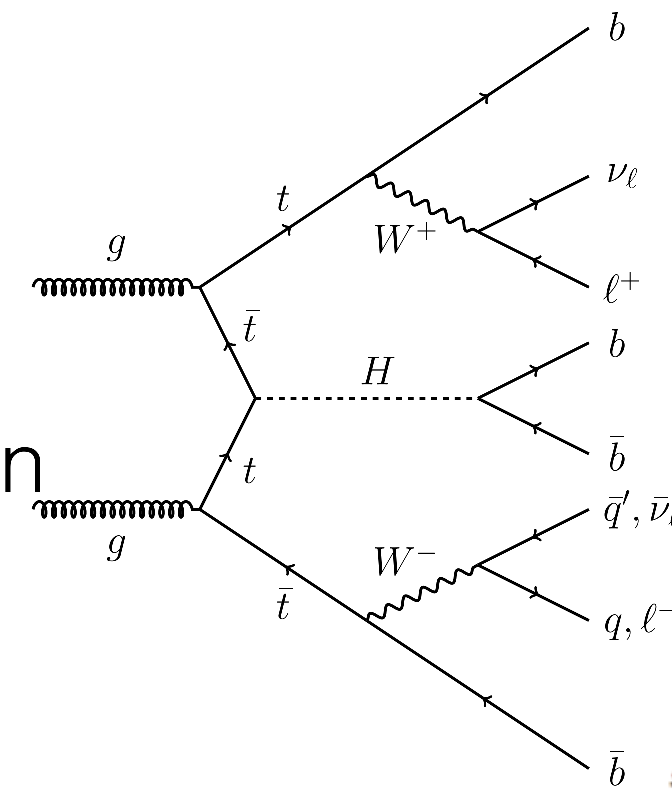
- Construct a statistic T that is sensitive to **Θ**, e.g Higgs candidate mass **M** if **Θ** is the theory mass
- Estimate the distribution by generating MC events
    - detector simulation -> cuts -> differential distribution T(**M | Θ**)
    - maximize wrt. **Θ** (template method), but need new templates (simulation) for each **Θ** and nuisance

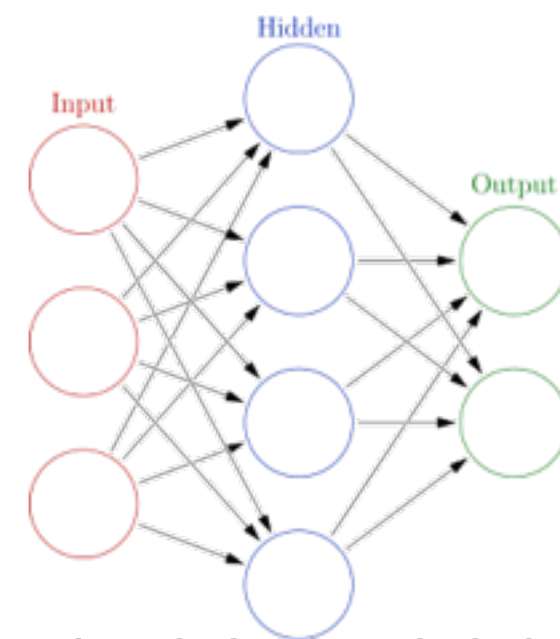


The difficulty

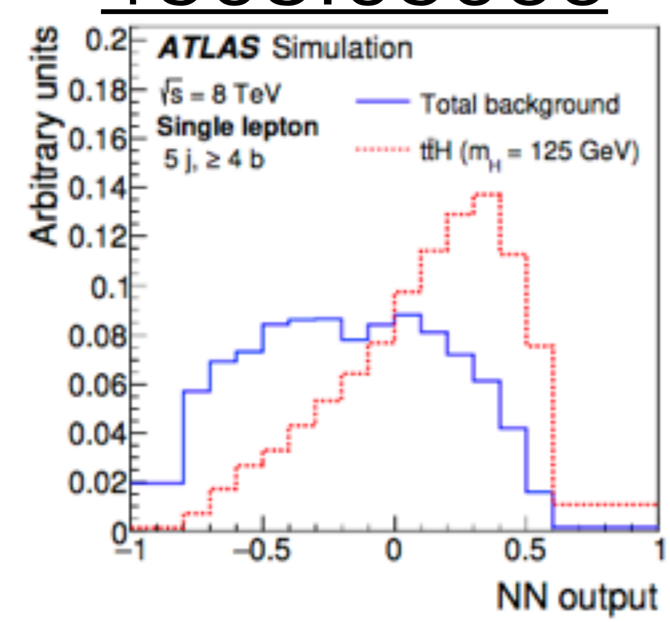
- complicated final states, high dimension, simple N-dimensional likelihood fails
- Combine measured quantities **Y**->T(**Y**) to minimize information loss: MVA
- High MC simulation statistics required to estimate T(**Y**)!
- **Can we calculate an event weight directly?**

ttH(bb): 2 -> 8 scattering





1503.05066



building the MEM

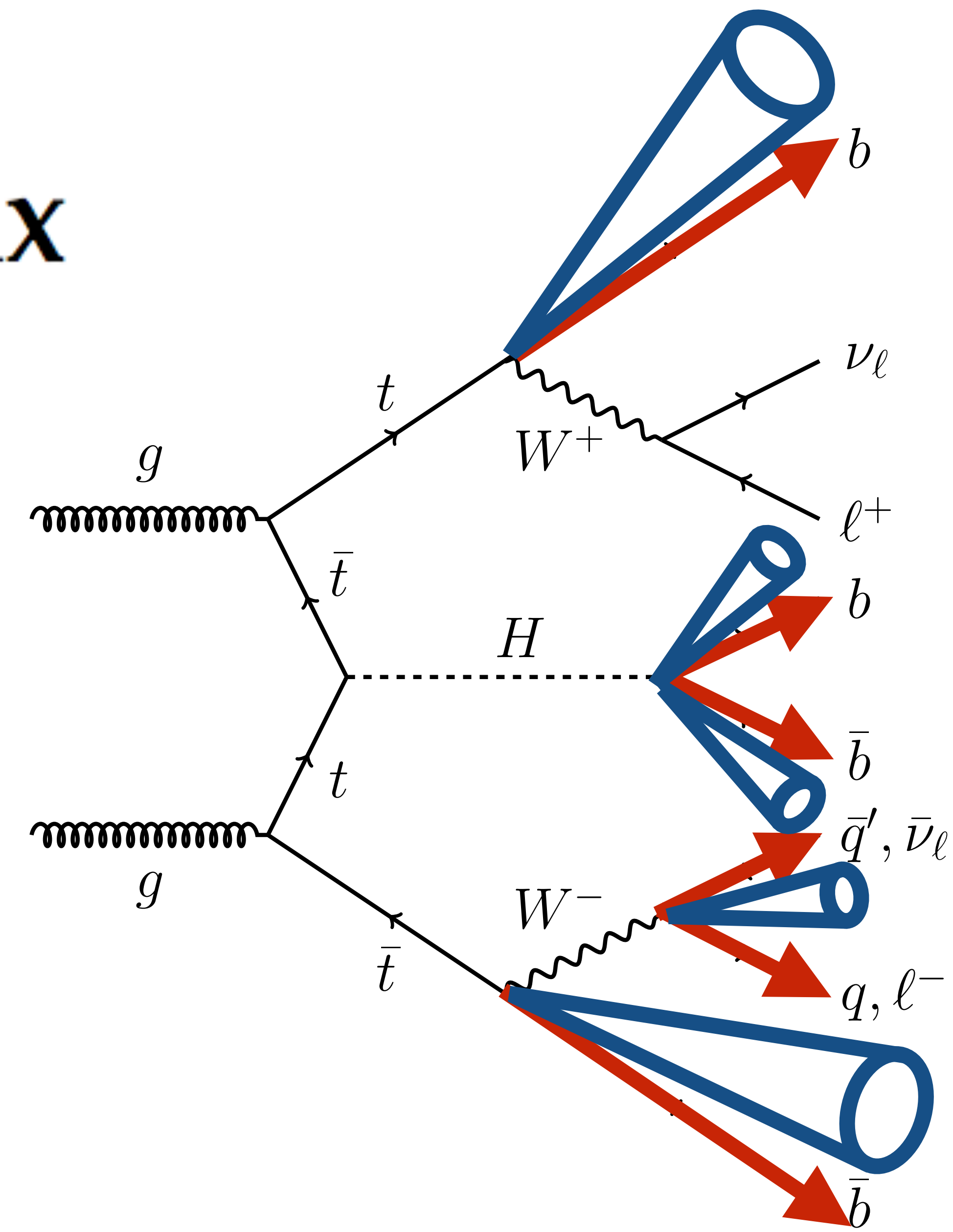
- Parton-level **X**: 4-momenta of the scattering

$$dP_{\mathbf{X}}(H_\theta) \propto f_a(x_1) f_b(x_2) |\mathcal{M}(\mathbf{X}|H_\theta)|^2 d\mathbf{X}$$

- detector transfer function $\mathcal{W}$(**Y**,**X**) from simulation

$$dP_{\mathbf{Y}|\mathbf{X}} = \mathcal{W}(\mathbf{Y},\mathbf{X}) d\mathbf{Y}$$

- Measured quantities **Y**: jet, lepton momenta, MET



$$dP_{\mathbf{Y}}(H_\theta) = \int \Big[dP_{\mathbf{X}}(H_\theta) \times \mathcal{W}(\mathbf{Y},\mathbf{X}) \Big] d\mathbf{Y} = p_{H_\theta}(\mathbf{Y}) d\mathbf{Y}$$

Interpreting as weight

- Given differential probability

$$\mathrm{d}P_{\mathbf{Y}}(H_\theta) = \int \left[\mathrm{d}P_{\mathbf{X}}(H_\theta) \times \mathcal{W}(\mathbf{Y},\mathbf{X}) \right] \mathrm{d}\mathbf{Y} = p_{H_\theta}(\mathbf{Y})\mathrm{d}\mathbf{Y}$$

- Interpret as differential cross-section

$$p_{H_\theta}(\mathbf{Y}) = \frac{1}{\sigma_{H_\theta}} \frac{\mathrm{d}\sigma_{H_\theta}}{\mathrm{d}\mathbf{Y}}$$

- Which is just an event weight, given hypothesis H_Θ

$$\mathcal{P}(\mathbf{Y}|H_\theta) = \frac{1}{\sigma_{H_\theta}} \int \mathrm{d}\mathbf{X} \int \mathrm{d}x_a \mathrm{d}x_b \, f(x_a, x_b) \, |\mathcal{M}(\mathbf{X}|H_\theta)|^2 \, \mathcal{W}(\mathbf{Y},\mathbf{X}).$$

Detector information

- Measure detector response to tree-level process
- For jets: parton shower and hadronization: E_q, Ω_q -> E_j, Ω_j, energy-dependent double-gaussian
- Factorize objects, energy/direction, assume continuity
- Capture simplified detector & simulation properties in a closed (albeit parametric) form
- **Question: didn't we just move the simulation problem? Which hadronization stage to correct to?**

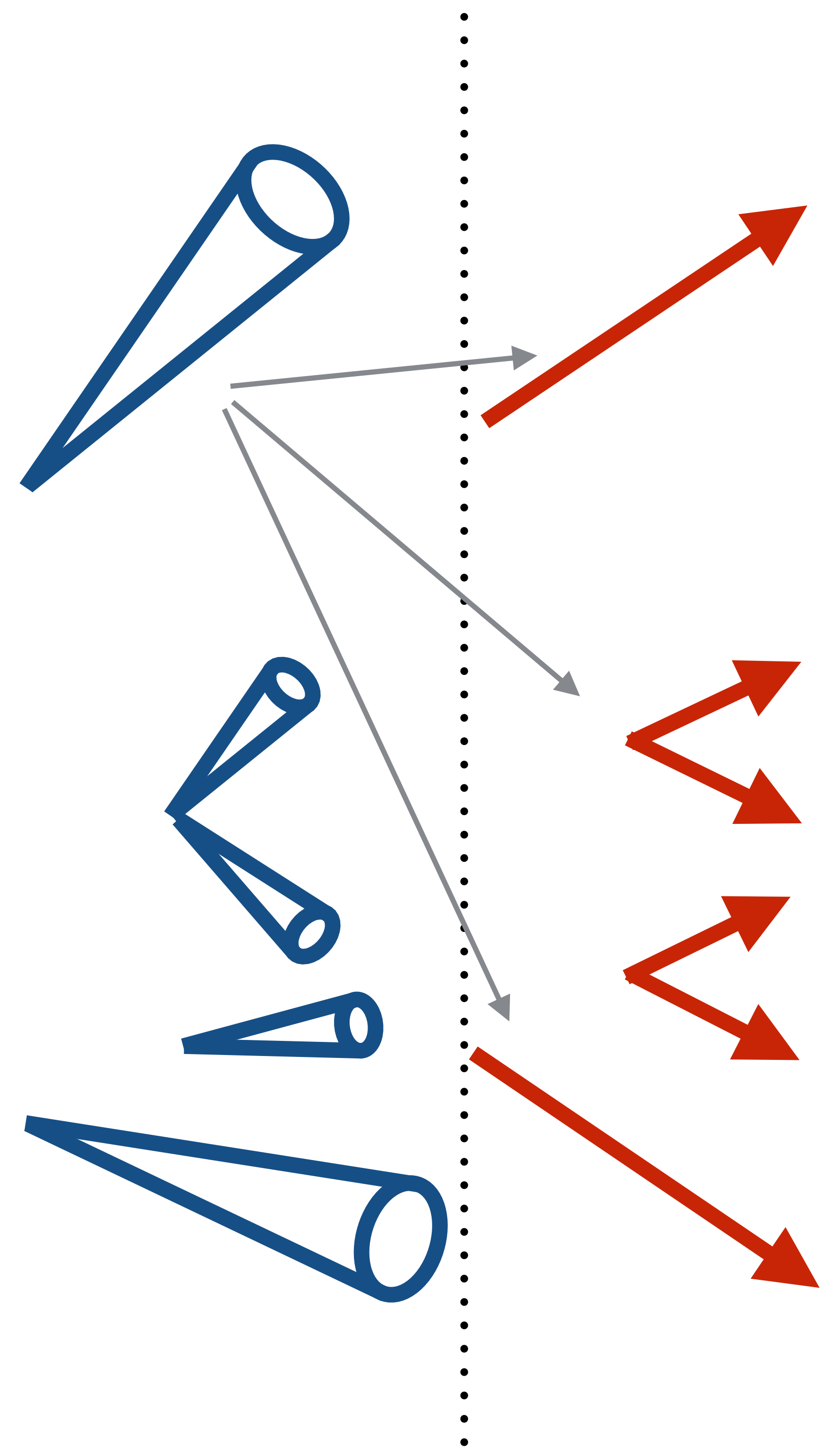
Association

- Associating parton-level (**N_q quarks**) to detector-level (**N_j jets**): permutations (up to hundreds)

- Can (and should!) use additional information to restrict

  - e.g. is jet from b-quark or light quark?

- **Question: what is the optimal way of associating?**



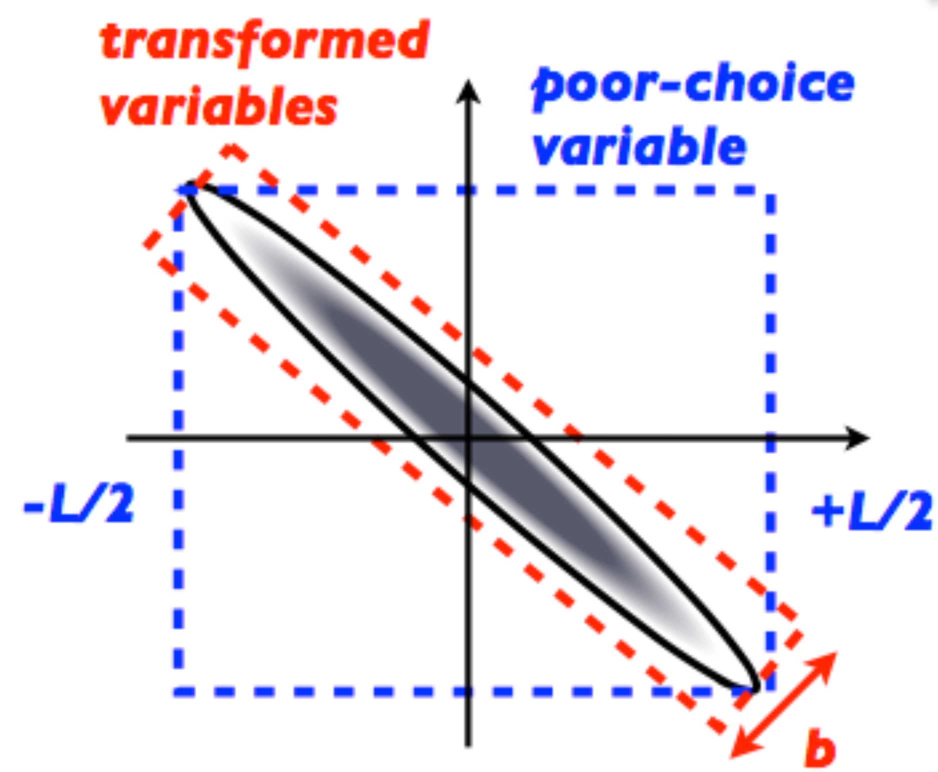
Integration

- Integrand an oscillating function: transfer functions, resonant propagators, 4-momentum conservation

$$\frac{1}{(q_k^2 - m_k^2)^2 + m_k^2\Gamma_k^2},$$

- Derive Jacobian factor for each topology: currently artigianal

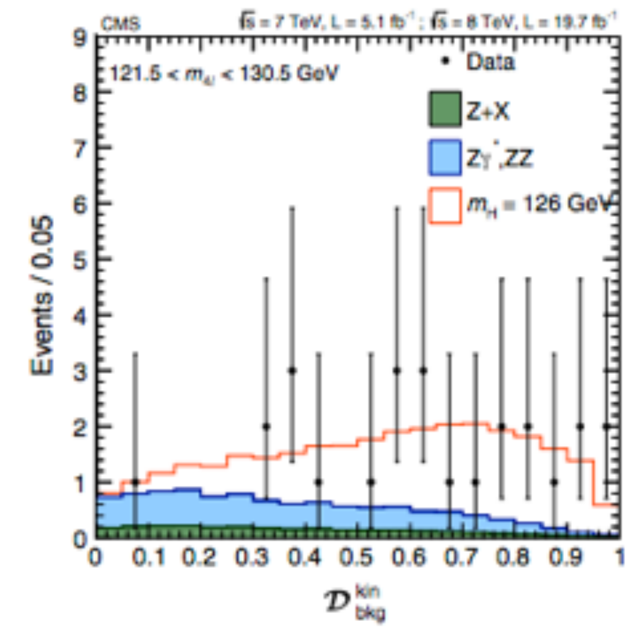
- **Question: how to benefit from existing technology?**



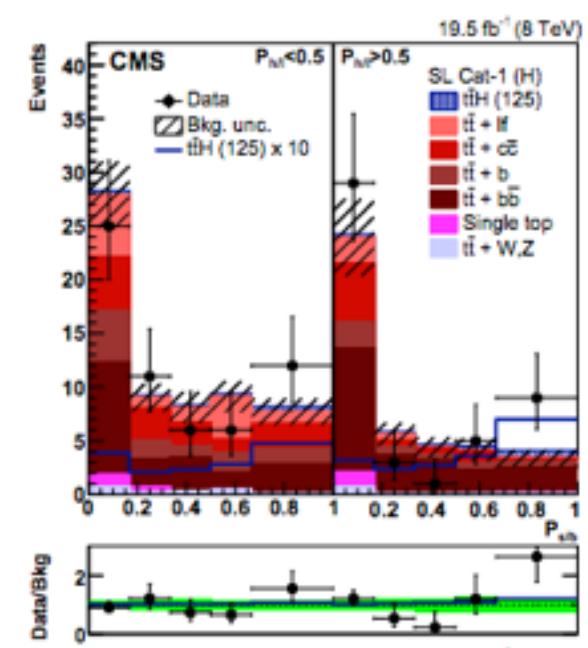
Applications

- **Point estimation**: mass, spin, parity
- **Searches**: tt+H, single-top, WH
- perfect resolution -> simplify transfer functions, PDFs: kinematic discriminant (MELA, MEKD)
- Framework for extending theorey concepts to experiment

PRD 89 092007



EPJ C 75 (2015) 251



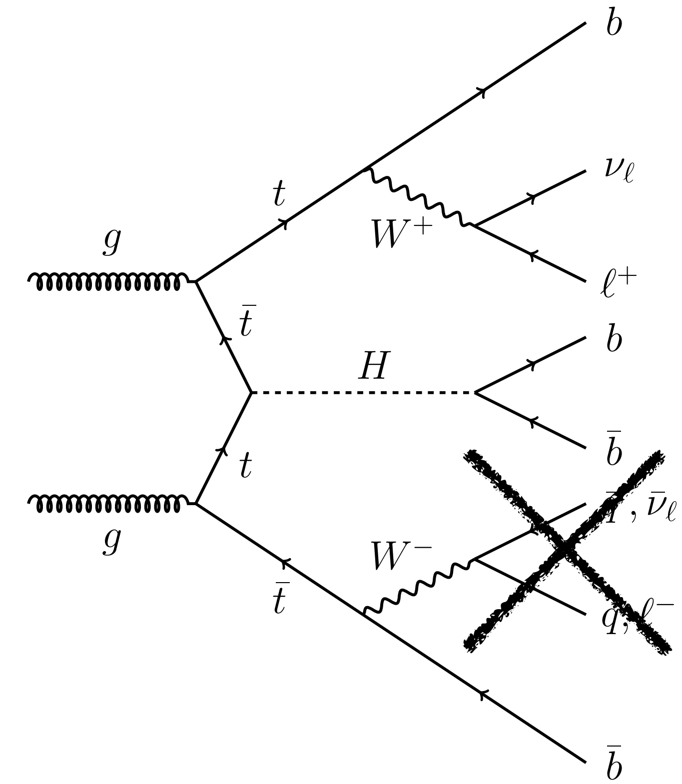
H -> ZZ -> 4l

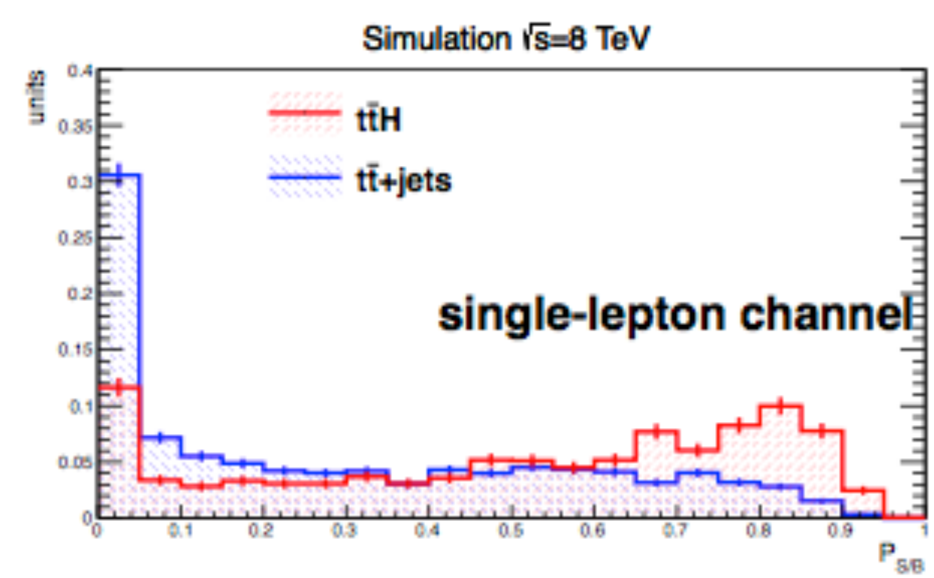
MEKD: 1210.0896

MELA: 1208.4018

Ongoing in run II

- Extending to unreconstructed final states (resolve P~0)
- Multiple event hypotheses: integrate over unknowns
- Leveraging machine-learning: MEM/ML based on final state topology
- ttH(bb) fully-hadronic, substructure
- NLO extensions
- distributed computing

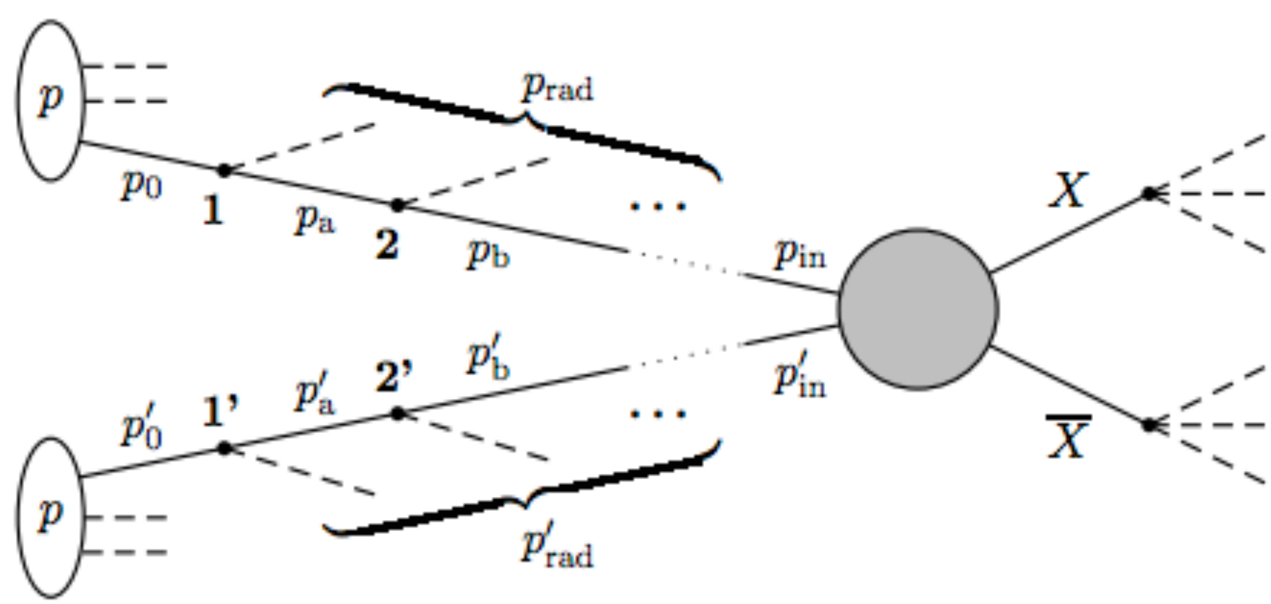




Moving to NLO

- **ISR**: kinematics changed, possible bias in e.g. measured mass
  - boost with ISR momentum to "MEM frame"
  - Sudakov factors @ parton level: $\mathcal{P}_j(p^2_{\mathrm{T},E}, z)$ $\quad z \approx \frac{p_{\mathrm{in,z}}}{p_{\mathrm{in,z}} + p_{\mathrm{rad,z}}}$,
  - transfer function for ISR at hadron level $\quad W_{\mathrm{ISR}}(p_{\mathrm{T}}, p^{\mathrm{vis}}_{\mathrm{T}})$

1010.2263

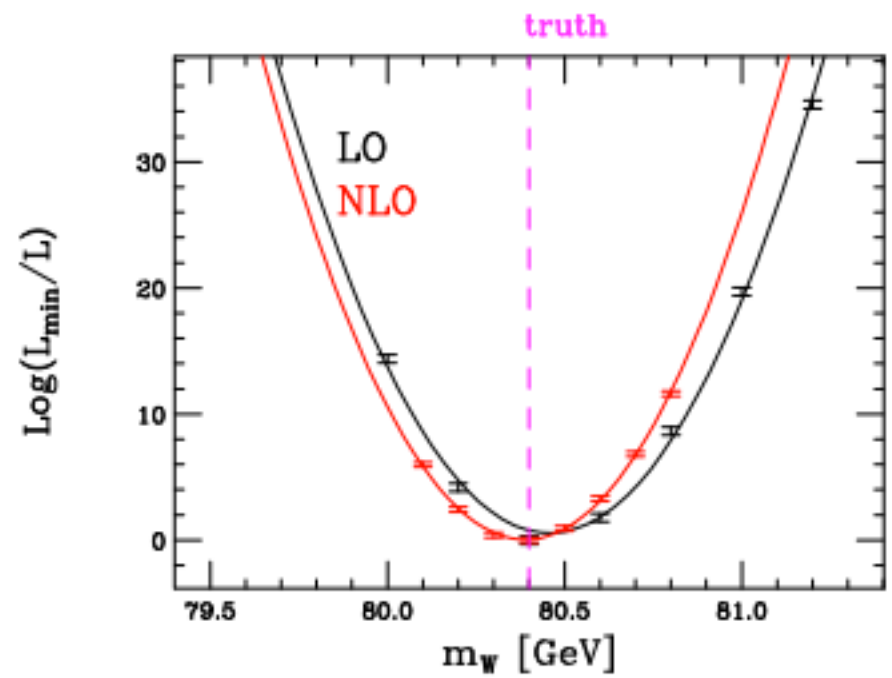


Question: how to benefit from automated NLO?

Real & virtual

1204.4424
1301.7086

- If NLO known, can "just" integrate real & virtual
- Real emissions boost events back to acceptance region
- NLO covers larger phase space, sensitivity improved by ~O(10%)
- But phase space is non-trivial



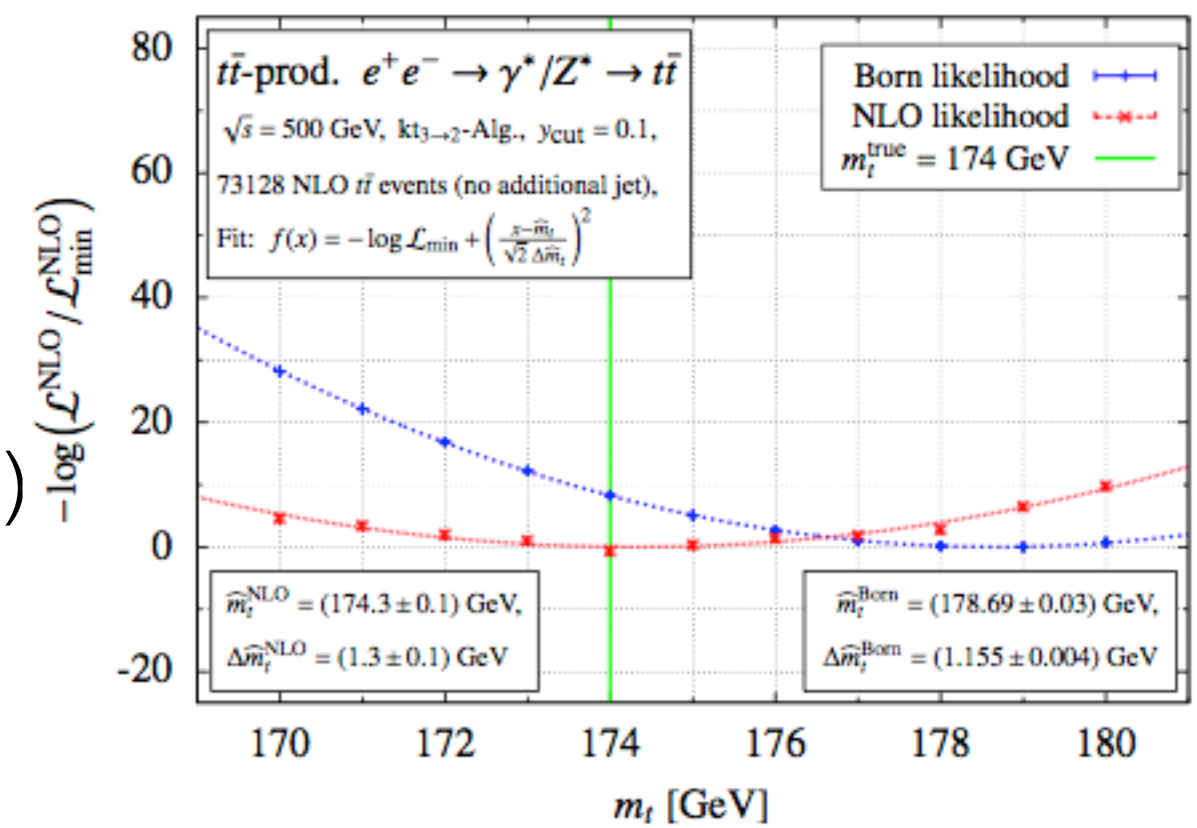
$$\tilde{P}_{LO}(\tilde{\phi}) = \frac{1}{\sigma_{LO}} \int dx_1\, dx_2\, d\phi\, \delta(x_1 x_2 s - Q^2) \times f^j(x_1) f^i(x_2) \mathcal{B}_{ij}(x_1, x_2, \phi) W(\phi, \tilde{\phi}).$$

$$\tilde{P}_{NLO}(\tilde{\phi}) = \frac{1}{\sigma_{NLO}} \int d\phi (V(\phi) + R(\phi)) W(\phi, \tilde{\phi}).$$

Jet algorithms

- Real: *n* jets in final state, *n+1* parton phase space
- recombine jets: massive jets from massless partons!
- need a modified clustering algorithm (3->2 instead of 2->1)
- factorize PS as *n*-jet + unresolved

1506.08798



$$\frac{d\sigma^{\rm NLO}_{n\text{-jet}}(\Omega)}{dJ_1\dots dJ_n} = \frac{d\sigma^{\rm B+V}(\Omega)}{dJ_1\dots dJ_n} + \int dR_{\rm unres}(\Phi)\,\frac{d\sigma^{\rm R}(\Omega)}{dp_1\dots dp_{n+1}}.$$

Machine learning

- **Standard ML**: throw simulation to a BDT/ANN
- **Standard MEM**: build best possible (N)LO discriminant, simulation in TF-s
- Combined approach: ML to choose MEM hypothesis based on reconstructed properties, combine with best possible simulation of processes

or

- Use ML to identify and calibrate objects, feed to best analytical discriminant

Variable	Definition	NN rank			
		$\geq$ 6j, $\geq$ 4b	$\geq$ 6j, 3b	5j, $\geq$ 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	-
$p_{\mathrm{T}}^{\mathrm{jet5}}$	p_{T} of the fifth leading jet	3	7	-	-
$H1$	Second Fox–Wolfram moment computed using all jets and the lepton	4	3	2	-

Summary

- MEM can replace “black-box” multivariate thinking
- Can be more sensitive, less simulation needed to test different hypotheses
- But there are subtleties to be discussed and understood with theorists
- Finding an optimal discriminant is hard
- No single automatic state-of-the-art package, many home-made implementations