Matrix Element Method at ATLAS and CMS (France)

Top LHC France - 24/05/2017, LPNHE

Nicolas Chanon - IPNL, CNRS/IN2P3
for the ATLAS and CMS Collaboration

Introduction

Matrix Element Method (MEM)

- Historically the **MEM was used first at Tevatron**, in top quark mass measurement, and single top quark discovery
- Today, an **increasing number of analyses performed at the LHC** are using the MEM, in **Higgs boson** and **top quark** measurements

- This talk **focuses on contribution from ATLAS and CMS France**, and prefers **top quark** examples

What is MEM ?

- **Basic idea:** relate reconstructed quantities to parton quantities under the hypothesis for a given process, included as exact **matrix element**

- Can be used to **discriminate between hypotheses,** or to **measure an observable** of interest

Matrix Element Method

MEM weight → $w_{i,\alpha}(\Phi')$

Integration → $\int$

Matrix Element at LO → $|\mathcal{M}_\alpha(p_k^\mu)|^2$

$$w_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha} \int d\Phi_\alpha \cdot \delta^4\Big(p_1^\mu + p_2^\mu - \sum_{k\geq 2} p_k^\mu\Big) \cdot \frac{f(x_1,\mu_F) f(x_2,\mu_F)}{x_1 x_2 s} \cdot \left|\mathcal{M}_\alpha(p_k^\mu)\right|^2 \cdot W(\Phi'|\Phi_\alpha)$$

Phase-space enforcing 4-momentum conservation

Parton distribution function

Transfer functions relating pardon-level to reconstructed quantities

Interpretation: The **MEM weight** is the cross section, for a given hypothesis, evaluated at the phase space point of the event, convolved with the transfer functions

MEM likelihood ratio

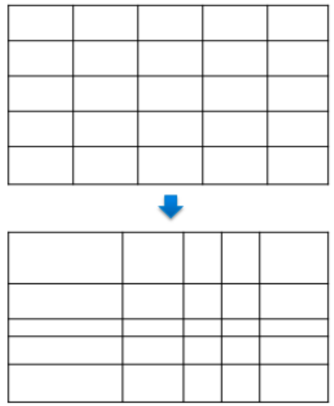
- **Neyman-Pearson Lemma:** Maximum **discrimination between two hypotheses** with a likelihood ratio

$$D_i = \frac{P(\boldsymbol{x}_i|S)}{P(\boldsymbol{x}_i|S) + P(\boldsymbol{x}_i|B)}$$

Phase space and integration

General expression for phase-space:

$$d\Phi = \left(\prod_{i=3}^{n} \frac{|\boldsymbol{p}_i|^2 d|\boldsymbol{p}_i| \sin\theta_i d\theta_i d\phi_i}{2E_i(2\pi)^3}\right) dq_1 dq_2 (2\pi)^4 \delta^4 \left(p_1+p_2-\sum_{j=3}^{n} p_j\right)$$



Integration

- As in monte carlo simulation, integration is performed with **importance sampling (VEGAS algorithm)**
- Preference to the regions with highest values of the integrand

Phase-space needs to be optimised

- **Aligning integration variables with the peaks of the cross section** improves the integration variance

Exemple:

$$d\Phi_{top,lep} \propto dE_b d\theta_b d\phi_b \cdot dE_l d\theta_l d\phi_l \cdot d\phi_\nu dm_W$$

Angles of jets, leptons and lepton energy are assumed to be known

Integrate over **W mass** with a change of variable to make it flat

Integrate over **jet and b-jet energy**, constrained by a **transfer function**

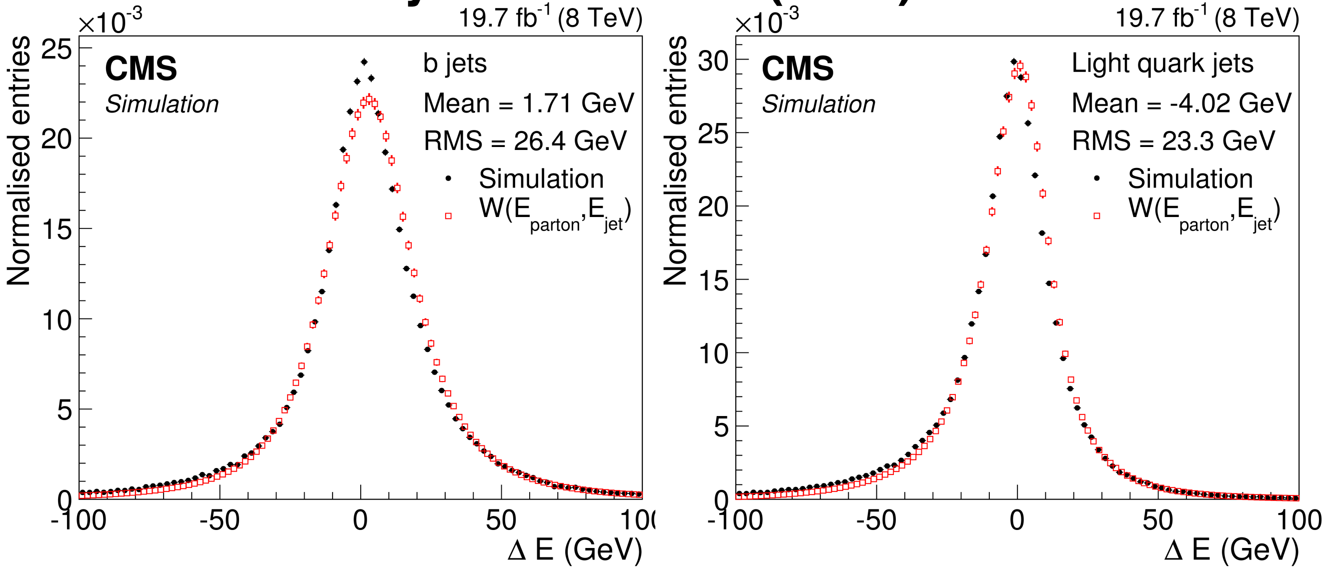
Integrate over **neutrino variables**, **unknown (constrained by mET TF)**

Transfer functions (TF)

Encodes the smearing of parton energies to jets energies

- Transfer functions must be normalised to 1
- Usually independent from another
- **Usual assumptions**:
  - Lepton energy resolution is small relative to jet energy resolution => assign a Dirac TF to lepton energy (speeds up the integration !)
  - A Dirac is also used for the angles

Phys. Lett. B 758 (2016) 321

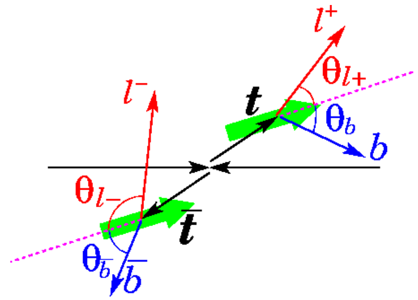


Exemple of parameterization:

$$W(E_{\text{parton}}, E_{\text{jet}}) = \frac{1}{\sqrt{2\pi}\left(\sigma_0 + \sigma_1 E_{\text{parton}} + \sigma_2\sqrt{E_{\text{parton}}}\right)} \times \exp\left[-\frac{1}{2}\left(\frac{\Delta E + m_0 + m_1 E_{\text{parton}}}{\sigma_0 + \sigma_1 E_{\text{parton}} + \sigma_2\sqrt{E_{\text{parton}}}}\right)^2\right]$$

Top quark spin correlation with MEM

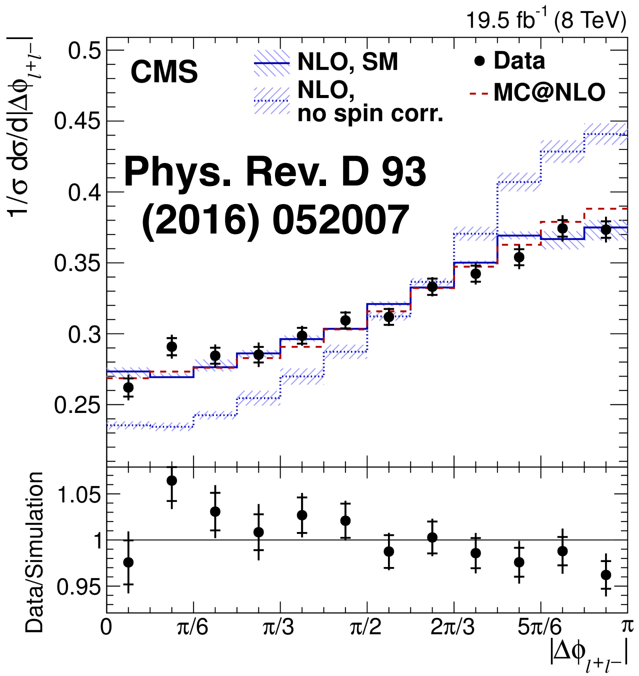
CMS, Phys. Lett. B 758 (2016) 321

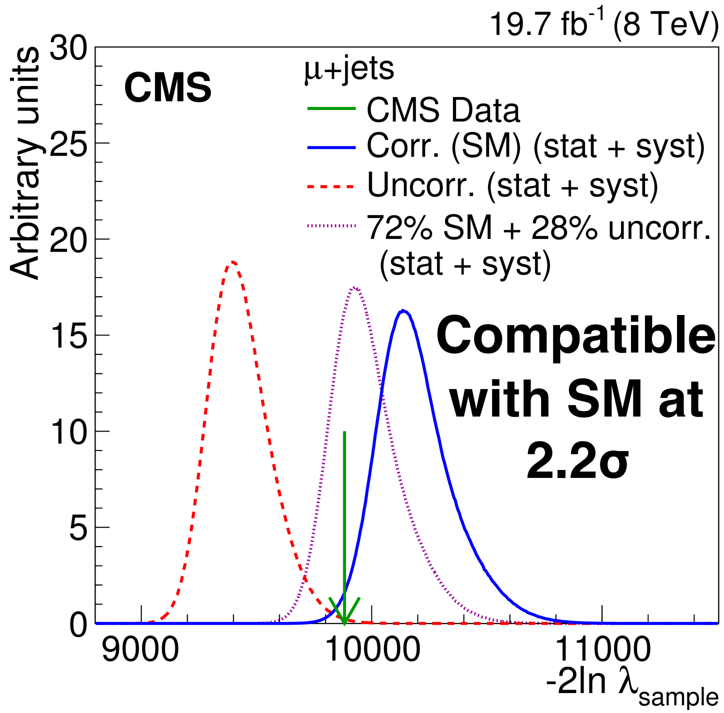


Discriminate ttbar model with/without spin correlation

- Framework used: **Madweight** *(JHEP 1012:068,2010)*
- **MEM used to measure a parameter:** the **amount of spin correlation** (Breit-Wigner for top mass) vs No spin correlation (narrow width for top in the ME)

$$\lambda_{\text{event}} = \frac{P(H_{\text{uncor}})}{P(H_{\text{cor}})}$$





$$A = \frac{(N_{\uparrow\uparrow} + N_{\downarrow\downarrow}) - (N_{\uparrow\downarrow} + N_{\downarrow\uparrow})}{(N_{\uparrow\uparrow} + N_{\downarrow\downarrow}) + (N_{\uparrow\downarrow} + N_{\downarrow\uparrow})}$$

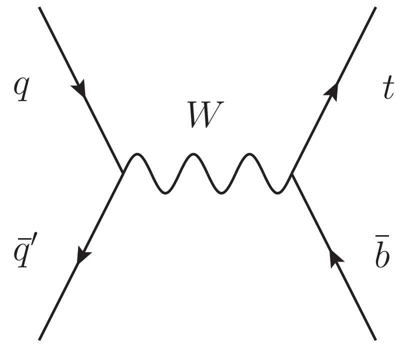
$$A_{\text{hel}}^{\text{measured}} = 0.23 \pm 0.03\,(\text{stat})^{+0.05}_{-0.04}\,(\text{syst})$$

$$A_{\text{hel}}^{\text{SM}} = 0.319$$

Other work to measure parameters with MEM: **top mass at ATLAS, LPNHE (2 PhD thesis)**

Top quark s-channel with MEM

ATLAS, Phys. Lett. B756 (2016) 228



MEM used in s-channel observation at ATLAS

- **MEMTk framework** (Humboldt University, Berlin)
- **Include many hypotheses** in the MEM likelihood: s-channel signal, t-channel (4FS), ttbar (semi-leptonic and dileptonic), W+jets, W+c, W+bb
- **Observation: 3.2σ** (expected 3.9σ)

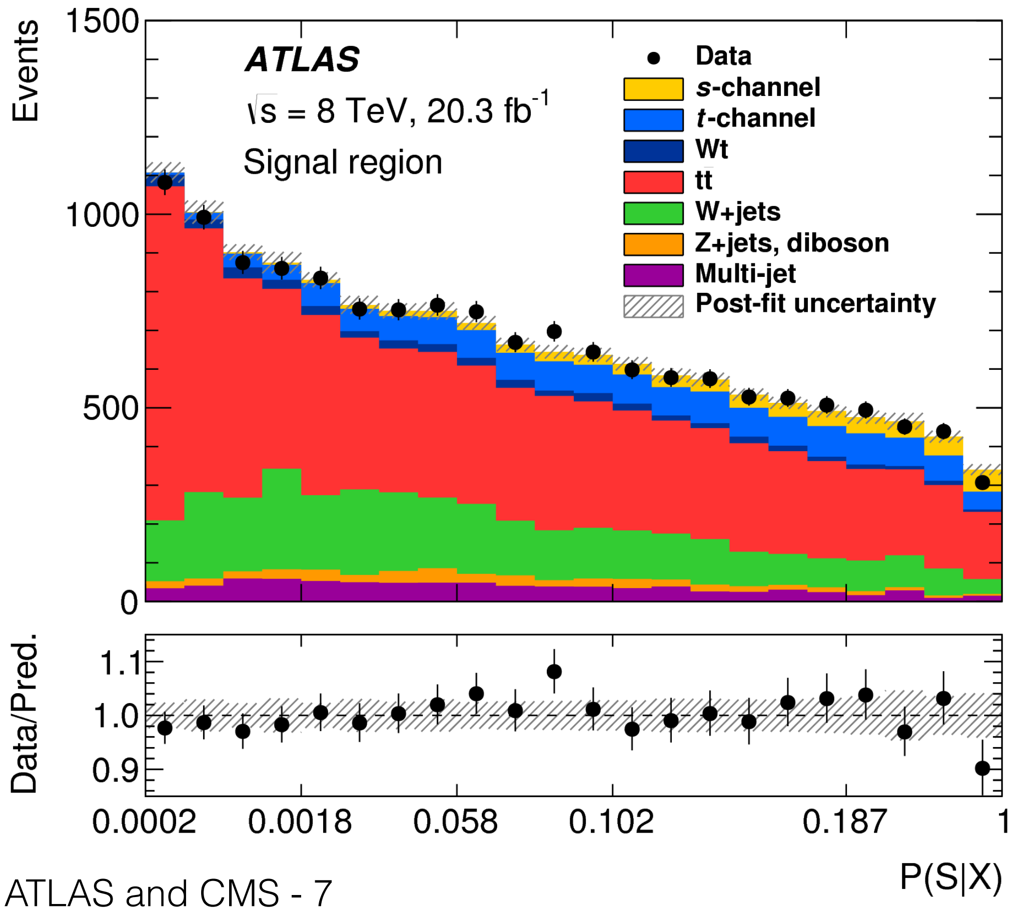
$$\sigma_s = 4.8 \pm 0.8(\text{stat.})^{+1.6}_{-1.3}(\text{syst.})\,\text{pb}$$

$$\sigma_s^{\text{th}} = 5.61 \pm 0.22\,\text{pb}$$

Previous analysis (same dataset):

$$\sigma_s = 5.0 \pm 1.7\,(\text{stat.}) \pm 4.0\,(\text{syst.})\,\text{pb}$$

- **MEM is responsible for half of the improvement** relative to previous analysis using BDT (Phys.Lett. B740 (2015) 118)



MEM, higher order and machine learning

Higher orders are accounted for in MEM with an effective way

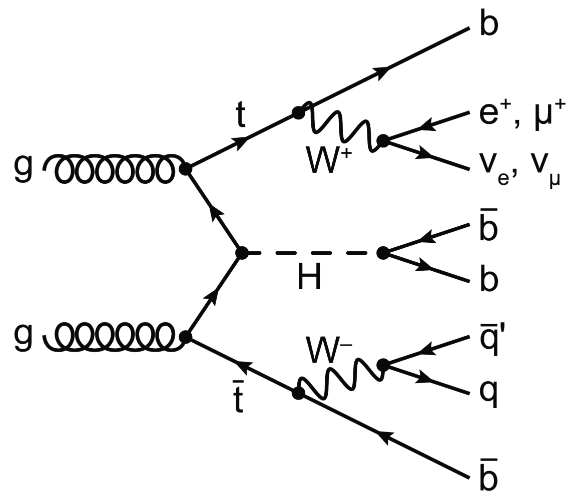
- **Option 1:** Can consider radiation of 1 jet in the ME (adds computing time)
- **Option 2:** Correct momenta of each particles, with the inverse boost of the total momentum projected on Z-axis => correct for higher order.

MEM and Machine learning are complementary:

MEM	Machine learning (NN, BDT…)
Exact computation (limited by integration accuracy)	Learn features from data sample
Exact LO kinematics with effective higher order correction	Fixed order NLO kinematics + parton shower (~leading log)
Integration over many variables	Training needs large samples
Evaluation needs integration at each event	Evaluation is a simple function of the input variables

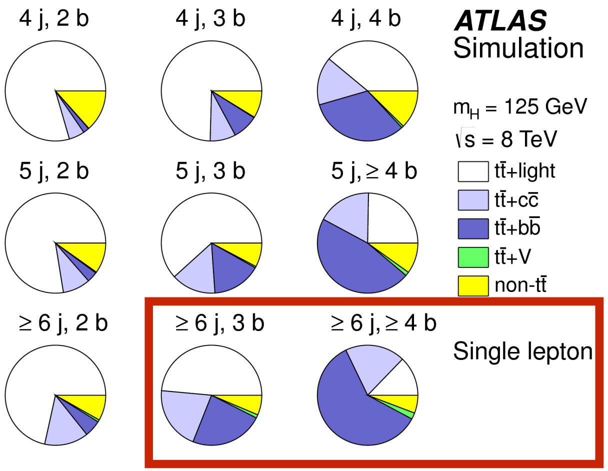
ATLAS $t\bar{t}H, H \rightarrow b\bar{b}$

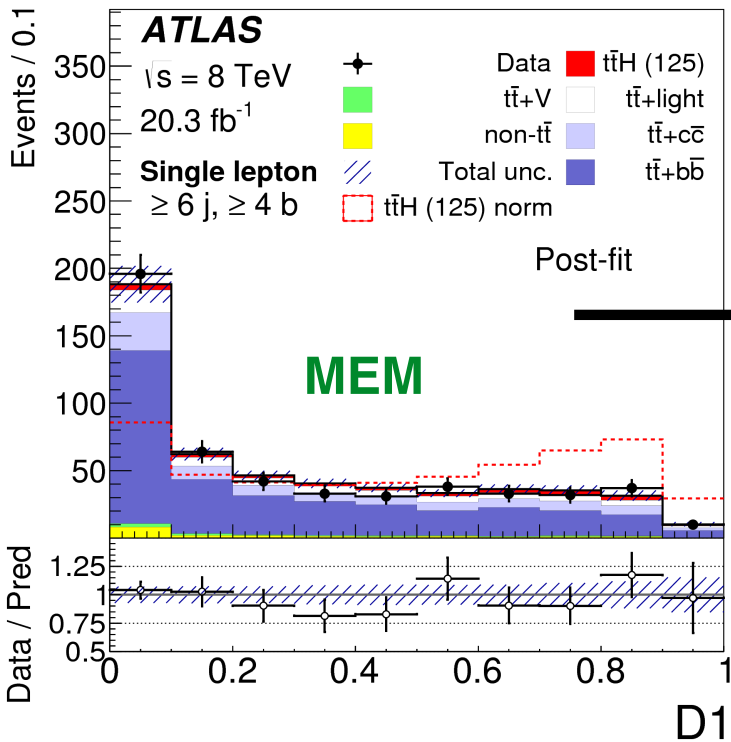
Eur. Phys. J. C (2015) 75:349

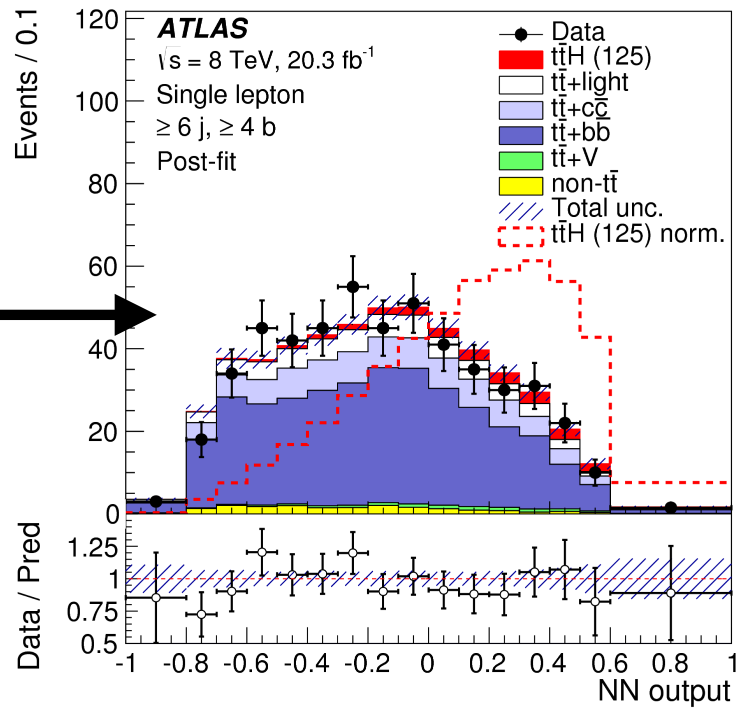


8 TeV analysis targeting lepton+jets and dileptons

- **MEM** is included in **≥6j 3b** and **≥6j ≥4b single lepton** categories.
- Background hypotheses: **tt+bb** (main background)
- MEM likelihood is included as input into a neural network



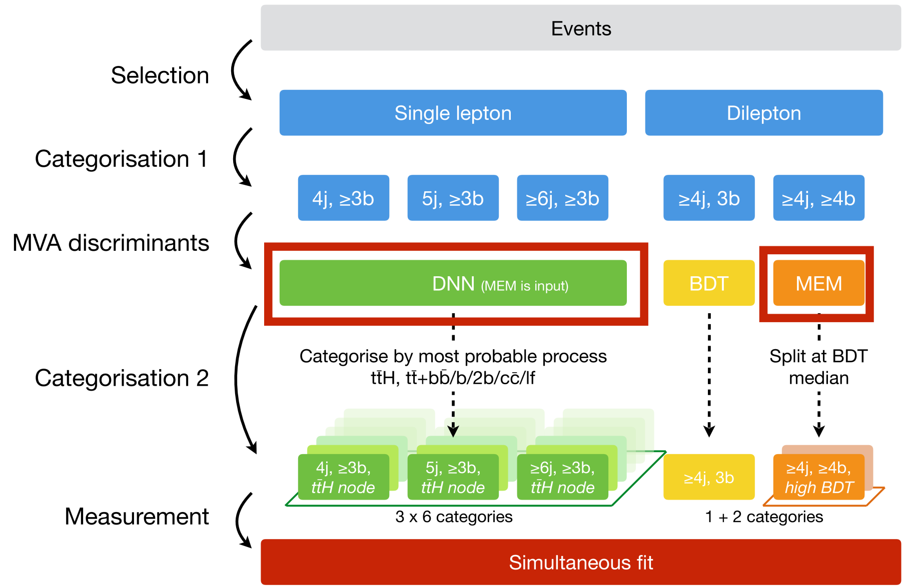




CMS $t\bar{t}H, H \to b\bar{b}$

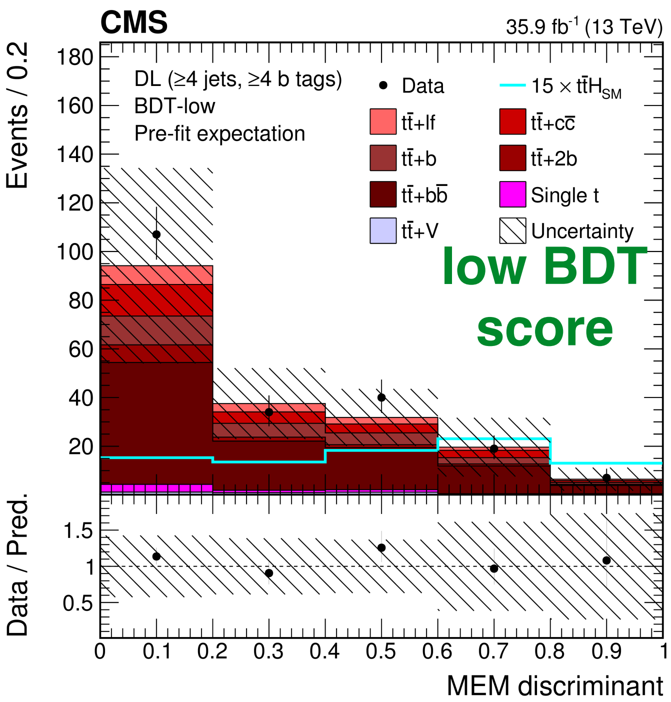
NEW

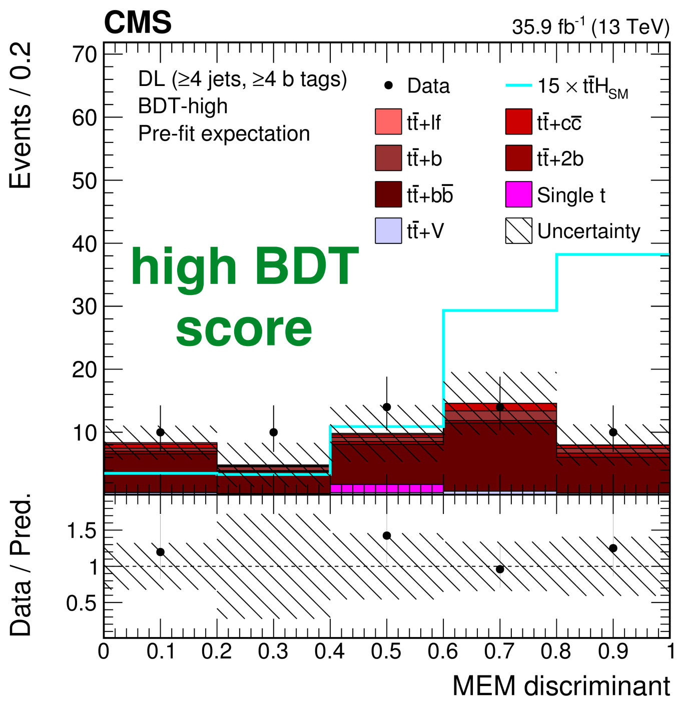
arxiv:1804.03682, submitted to JHEP

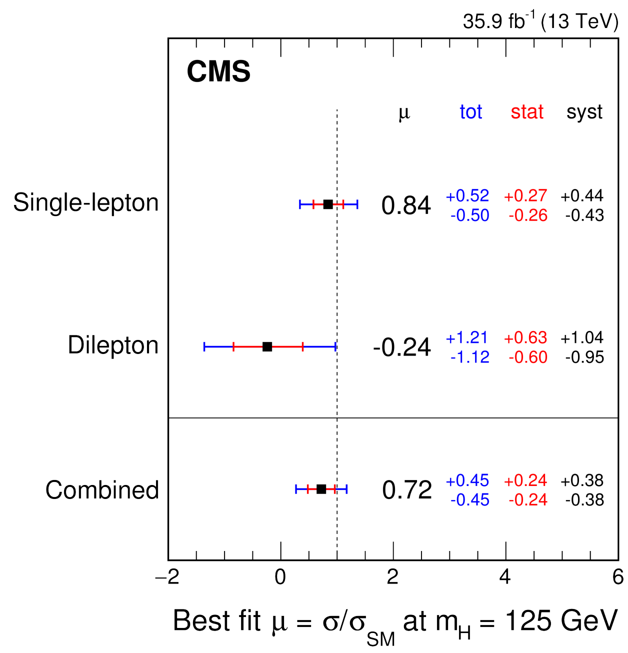


Analysis targeting lepton+jets and dileptons

- **Single lepton : MEM** is an input to a **Deep Neural Network**
- **Dilepton:** Use **MEM** as final discriminant in **low/high BDT** score categories



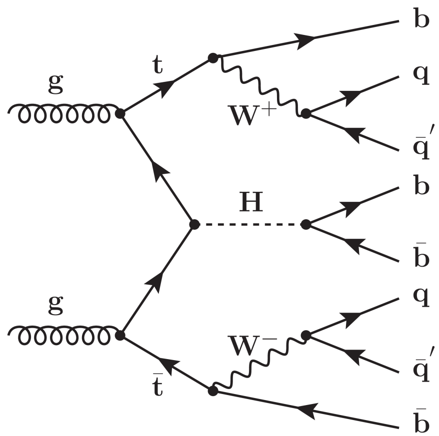




CMS $t\bar{t}H, H \to b\bar{b}$ fully hadronic

NEW

arxiv:1803.06986, submitted to JHEP



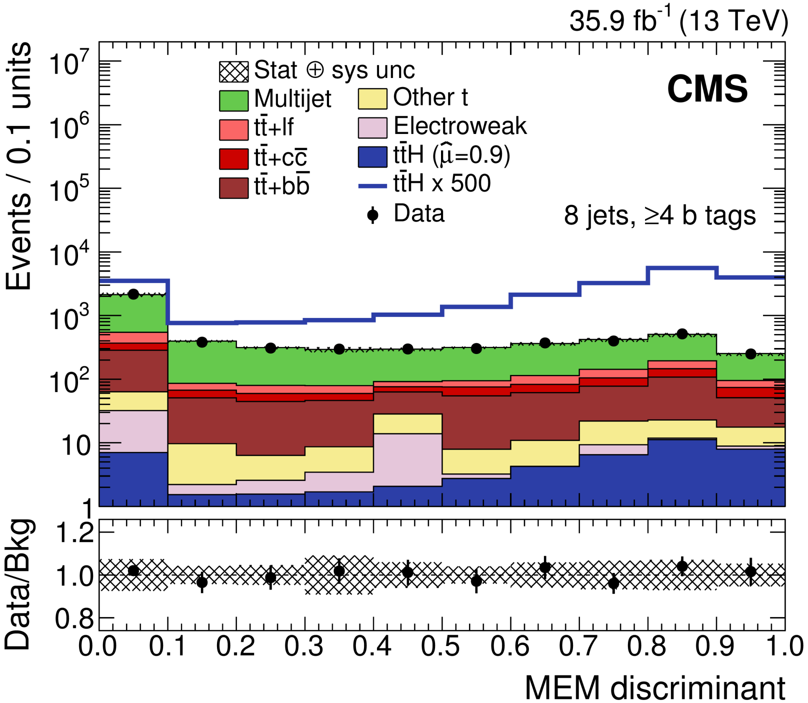
MEM used as discriminant in a busy environment

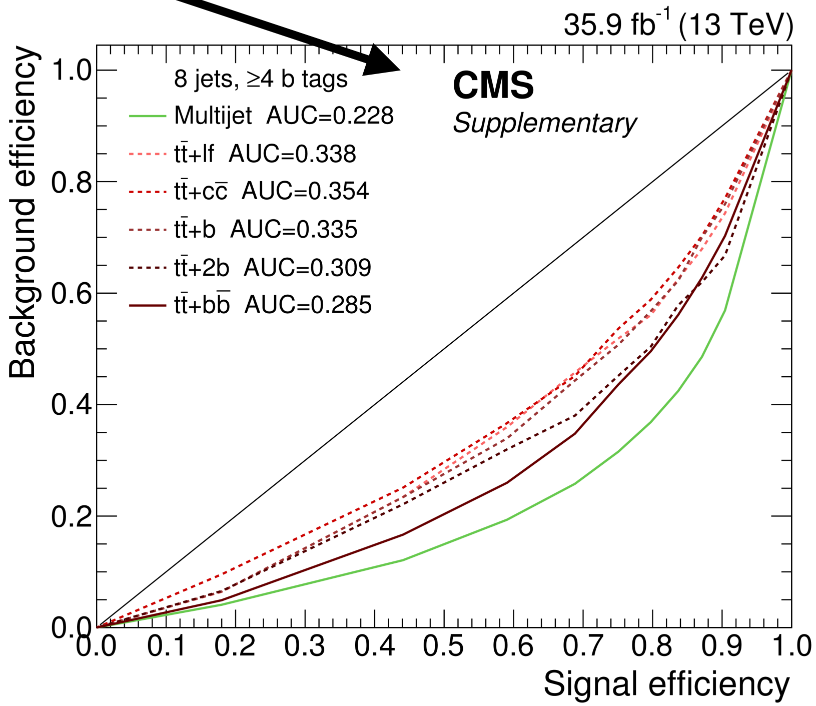
- **MEM** is included in all **of the analysis categories**
- **Many jet permutations !**
- Background hypotheses: **tt+bb** (main background)

Combination result:

$$\hat{\mu} = 0.9 \pm 0.7\,(\text{stat}) \pm 1.3\,(\text{syst})$$

Good discrimination against tt+cc, …

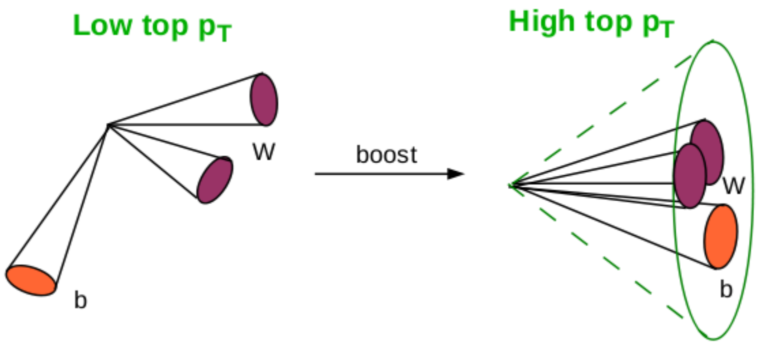


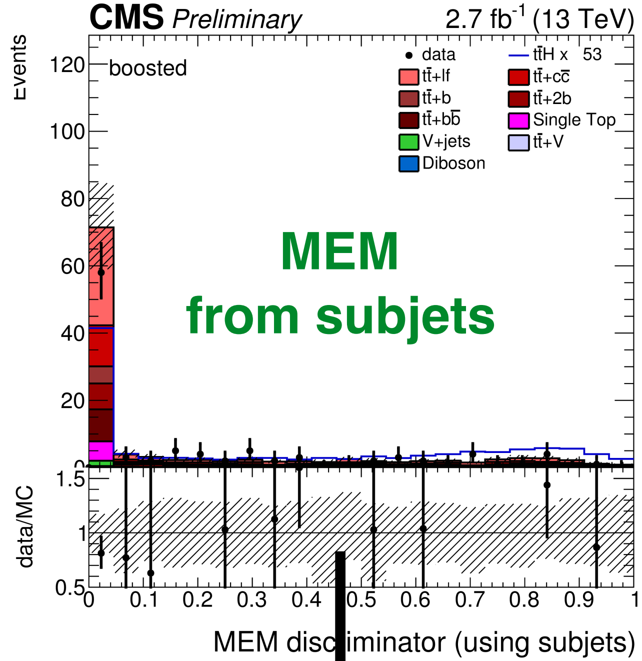


CMS $t\bar{t}H, H\to b\bar{b}$: boosted category

2.7 fb-1

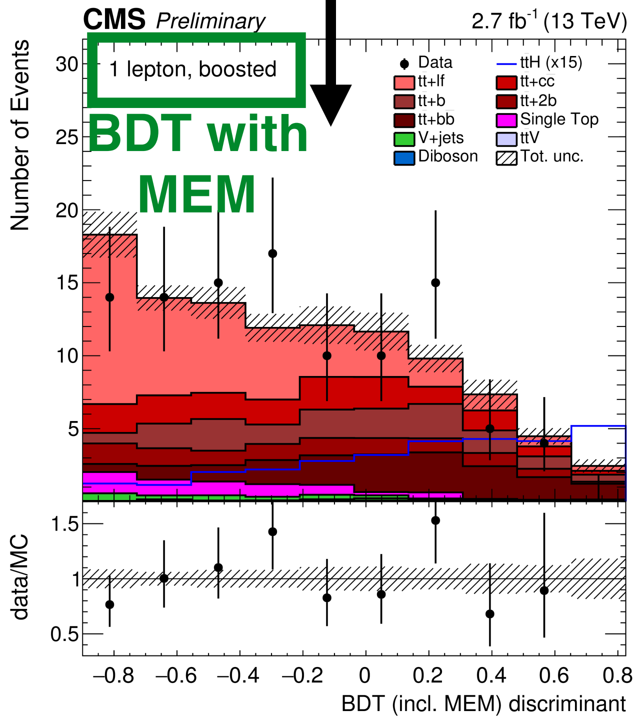
CMS HIG-16-004





Analysis with 2.7fb-1
(Moriond2016, 2015 data)

- **Fat jet substructure** (C/A ΔR=1.5)
- **Subjets** are reconstructed using filtering and mass drop requirement
- Resolved subjets are matched to parton level in MEM



Category	Observed	Expected
4 jets, 3 b-tags	14.5	$18.6^{+8.2}_{-5.5}$
4 jets, ≥ 4 b-tags high BDT output	35.7	$25.6^{+13.4}_{-8.1}$
4 jets, ≥ 4 b-tags low BDT output	86.6	$84.2^{+41.3}_{-25.8}$
5 jets, 3 b-tags	16.0	$12.3^{+5.5}_{-3.6}$
5 jets, ≥ 4 b-tags high BDT output	7.5	$10.3^{+5.6}_{-3.4}$
5 jets, ≥ 4 b-tags low BDT output	35.2	$31.9^{+16.1}_{-9.9}$
≥ 6 jets, 2 b-tags	25.4	$41.1^{+21.1}_{-13.1}$
≥ 6 jets, 3 b-tags	9.6	$7.6^{+3.3}_{-2.2}$
≥ 6 jets, ≥ 4 b-tags high BDT output	9.2	$8.3^{+4.4}_{-2.7}$
≥ 6 jets, ≥ 4 b-tags low BDT output	15.4	$18.3^{+9.6}_{-5.8}$
≥ 4 jets, ≥ 2 b-tags, boosted	7.5	$10.7^{+5.9}_{-3.5}$
lepton+jets combined	4.0	$4.1^{+1.8}_{-1.2}$

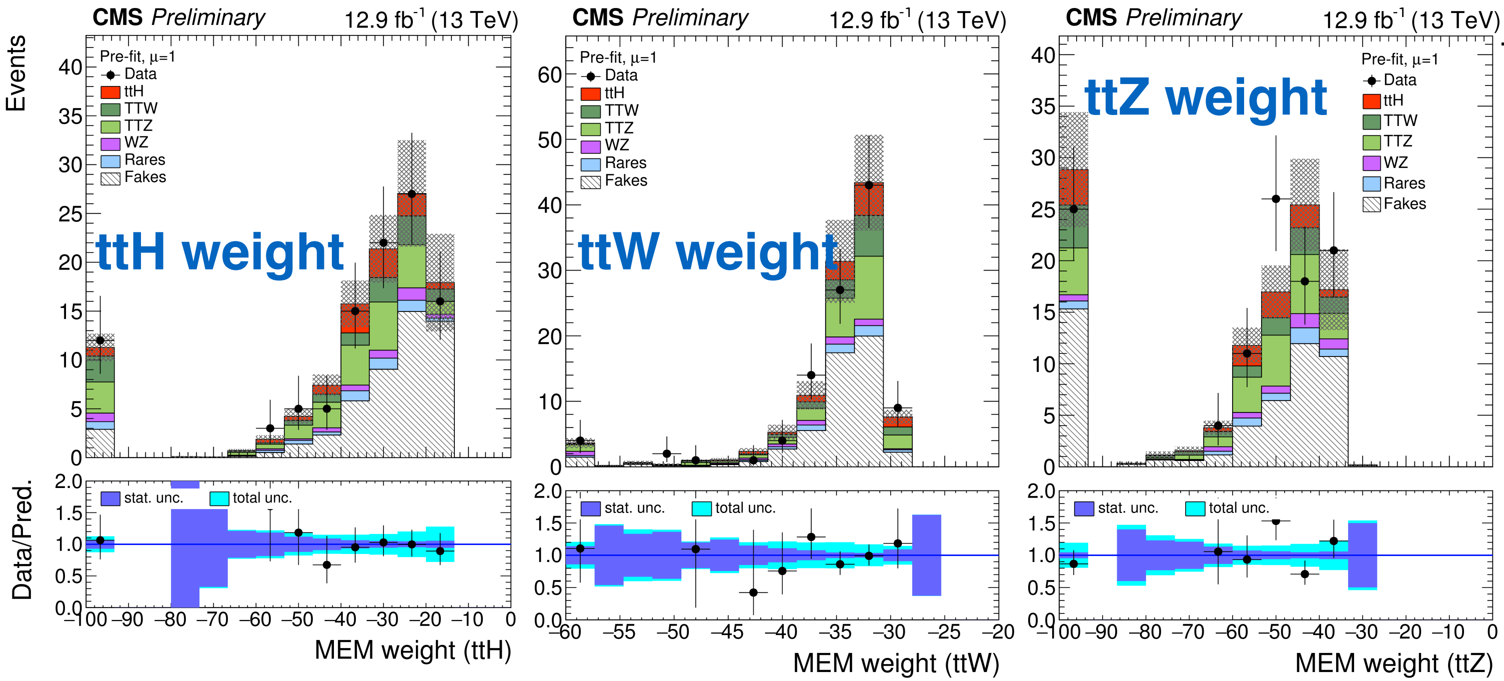
- Was one of the categories with the best performance
- Hopefully, will be included back soon

tt̄H multilepton : MEM in 3l category

CMS HIG-16-022, HIG-17-004

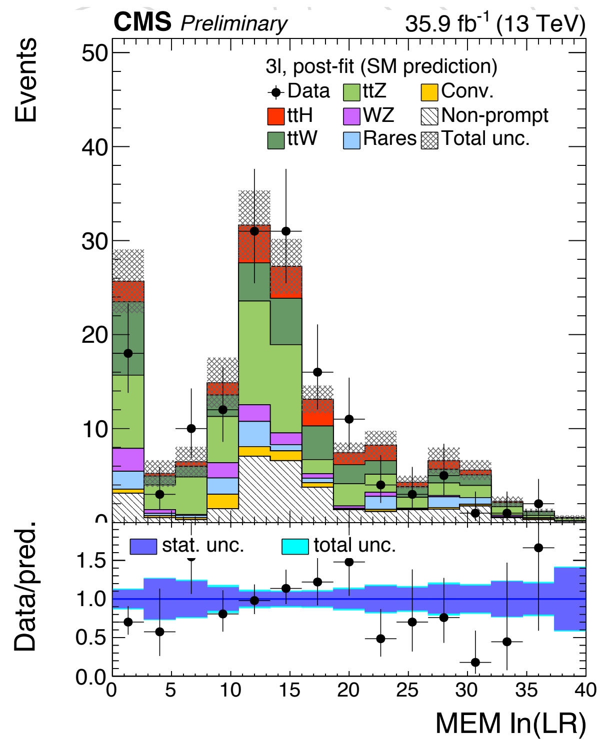
IPHC, IPNL

MEM weights under ttH, ttW, ttZ/γ* hypotheses



HIG-16-022 (ICHEP 2016):

- **improved discrimination by 10% in 3ℓ category**
- Include log(weights) as input to a kinematic BDT trained against ttV



HIG-17-004 (Moriond 2017): include the likelihood of ttH vs ttV weights inside the ttW/Z BDT

Likelihood ratio of ttV vs ttH+ttV

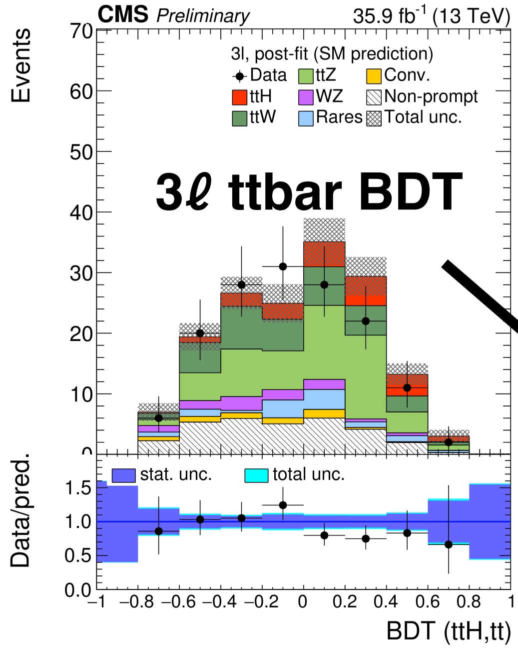
$$-log\left(\frac{\sigma_{t\bar{t}Z}w_{t\bar{t}Z} + k\cdot\sigma_{t\bar{t}W}w_{t\bar{t}W}}{\sigma_{t\bar{t}H}w_{t\bar{t}H} + \sigma_{t\bar{t}Z}w_{t\bar{t}Z} + k\cdot\sigma_{t\bar{t}W}w_{t\bar{t}W}}\right)$$

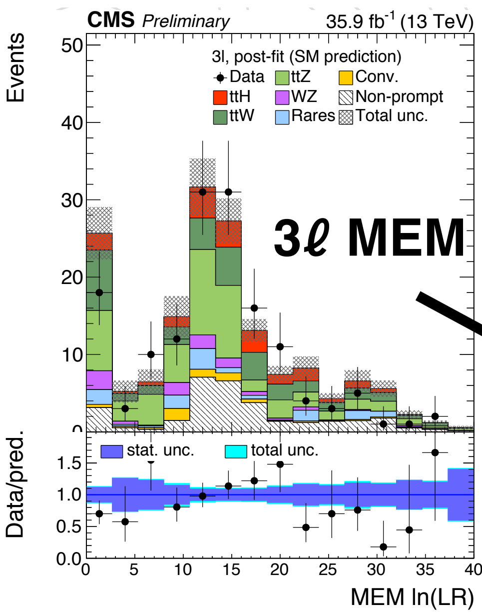
$t\bar{t}H$ multilepton discriminants

CMS HIG-17-004

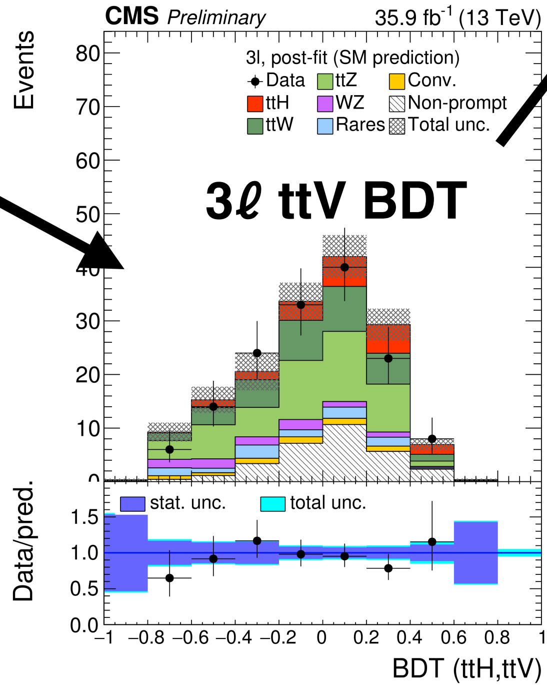
IPHC, IPNL

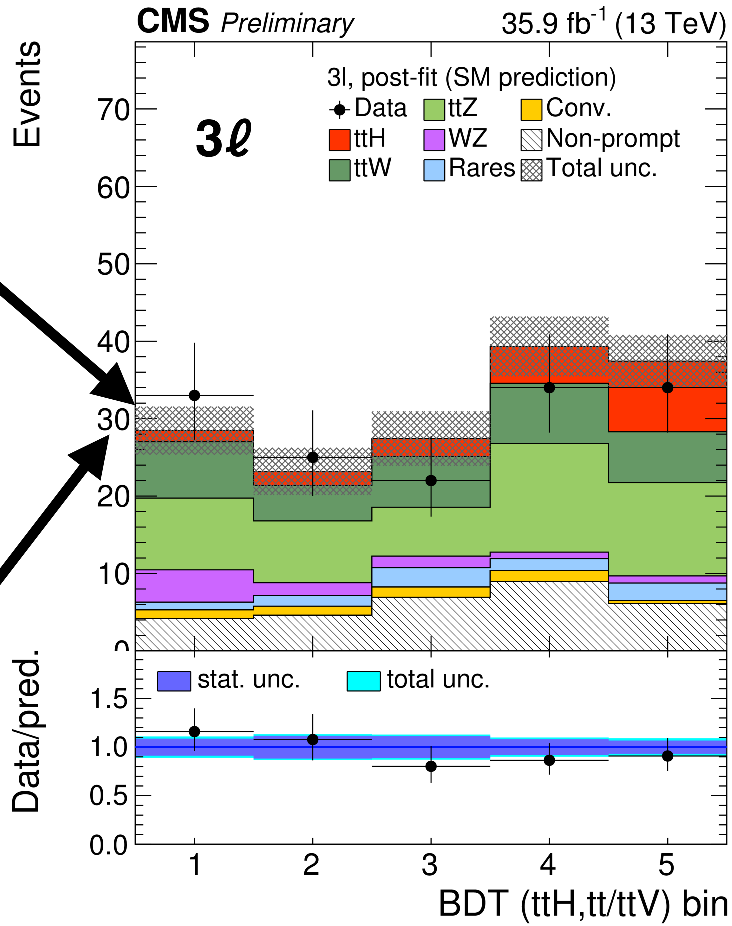
- **3ℓ category uses MEM.**





- 3ℓ vs ttW/Z: Includes **Matrix Element Method** likelihood ratio of ttH vs ttW+ttZ





LLR

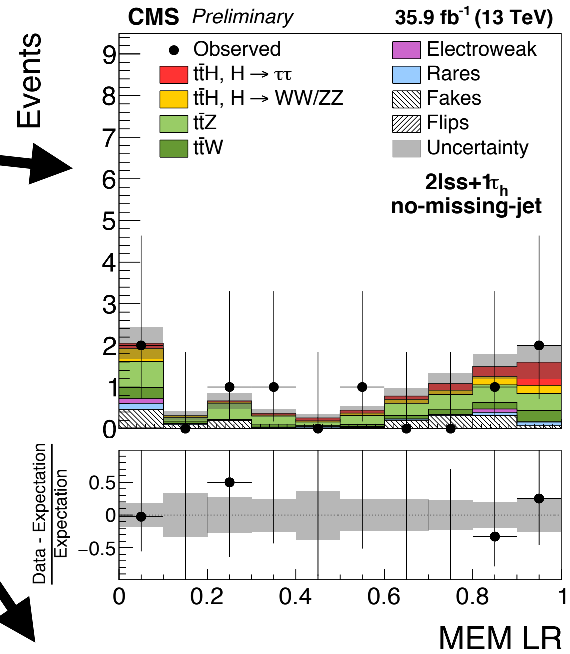
$t\bar{t}H, H\rightarrow\tau\tau$

CMS HIG-17-003

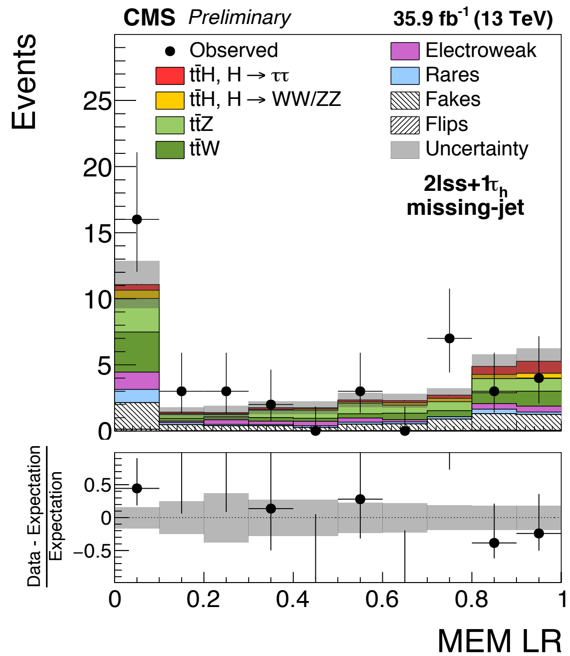
MEM is the final discriminant in the category $2\ell ss+1\tau_h$

- Similar MEM framework as CMS ttH(bb) and ttH multi lepton
- **MEM likelihood ratio** with ttH vs ttZ ($Z\rightarrow\tau_h\tau_h$ or $Z\rightarrow\ell\ell$ with one ℓ misidentified as τ_h) and ttbar hypotheses ($1\ell 1\tau_h$ with an additional lepton from b-decay)

no missing jet



with missing jet



Running MEM on GPU

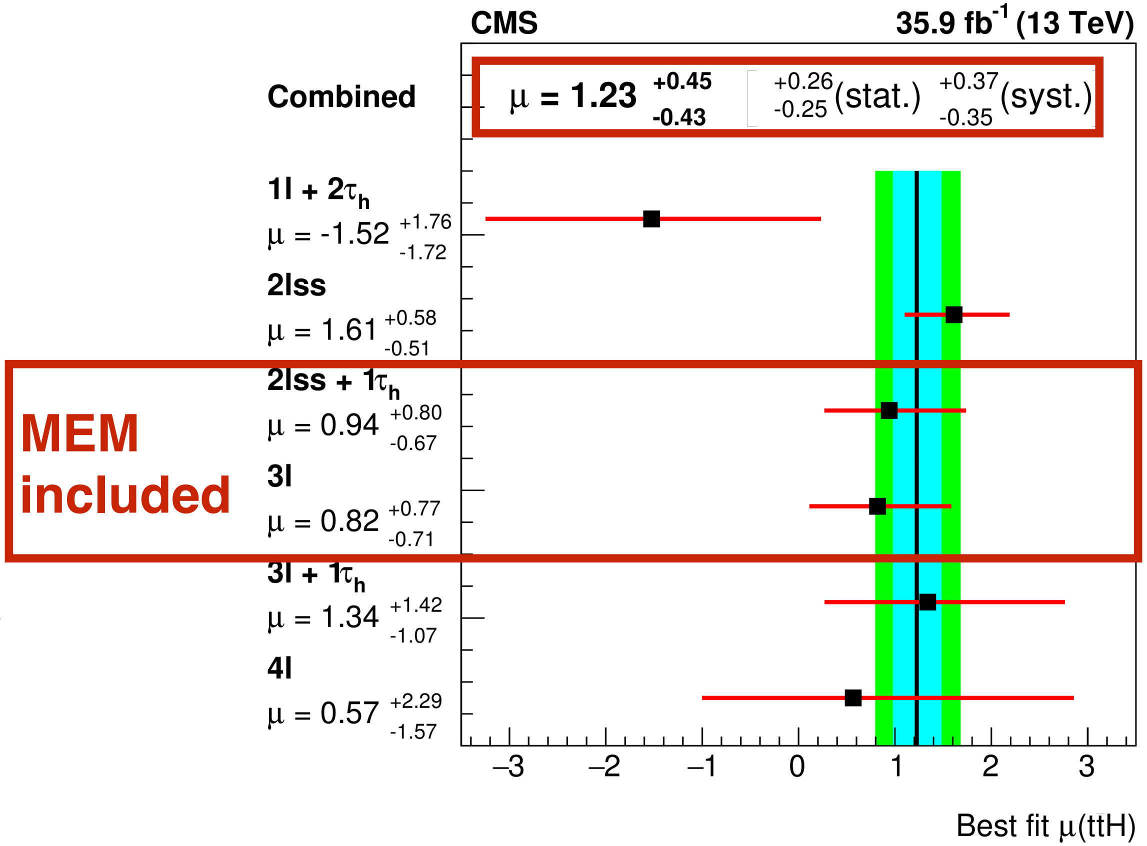
- **LLR** group implemented MEM for this process on GPU
- Each call of VEGAS is composed of many independent iterations of function evaluation: **highly parallelizable**
- Code written in **OpenCL**, tests on NVIDIA K80 GPU
- **Significant speed-up: 1 GPU is equivalent to ~4 CPU nodes of 20 cores** *[G. Grasseau, to be presented to CHEP'18]*

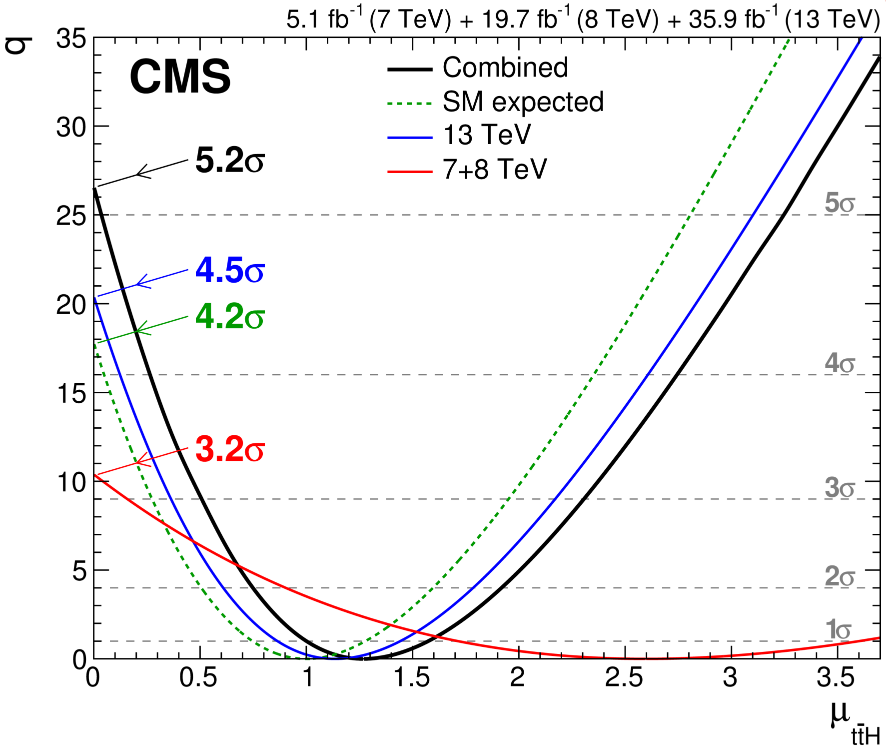
$t\bar{t}H$ combinations

NEW

arxiv:1803.05485, submitted to JHEP

Significance, combining ttH multilepton and tau analyses:
Observation: 3.2σ (2.8σ expected)





$t\bar{t}H$ observation

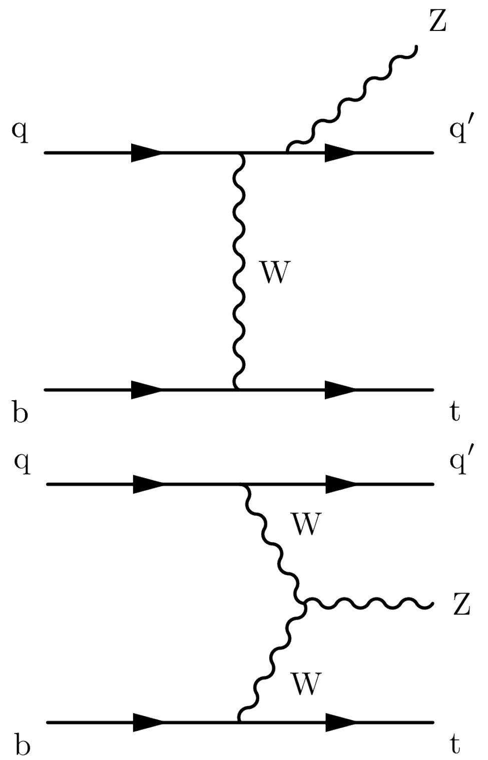
arxiv:1804.02610, accepted by Phys. Rev. Lett.

Full combination (Run 1, Run 2 ttH multilepton, τ_h+X, γγ, ZZ and bb):
5.2σ observed (4.2σ expected)

MEM in single top + Z

CMS, Phys. Lett. B 779 (2018) 358

IPHC, IPNL. Talk from Nicolas Tonon



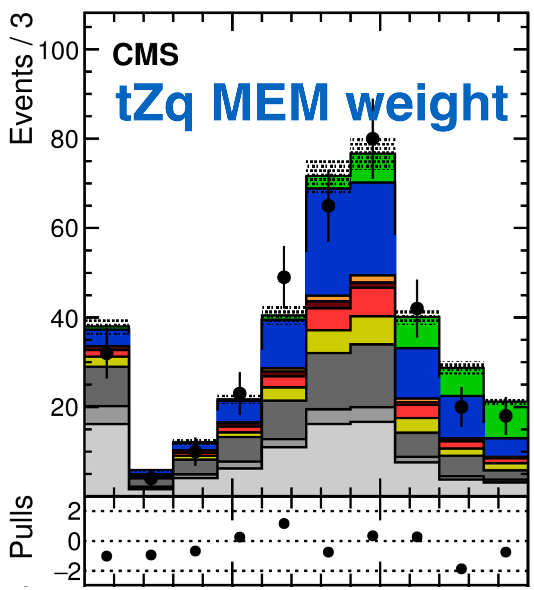
- Same MEM framework as developed for ttH multi lepton
- MEM makes use of the forward jet in tZq to discriminate
- Include **MEM weights and MEM as a kinematic fit**

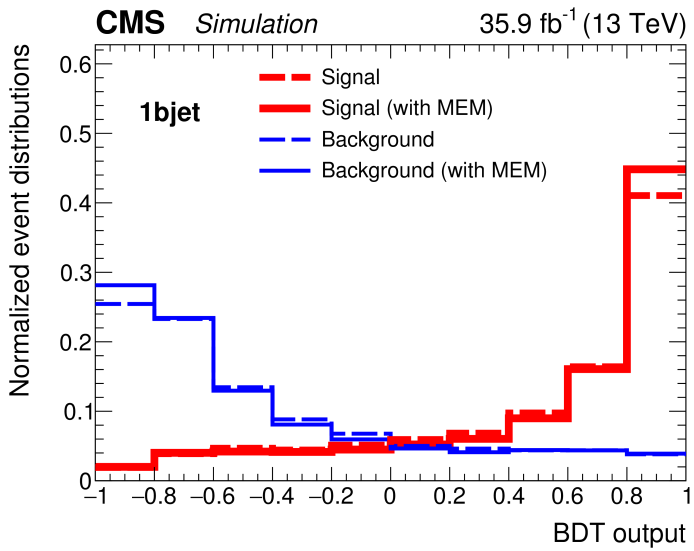
$$w_{i,\alpha}(\Phi') = \frac{1}{\sigma_\alpha}\int d\Phi_\alpha \cdot \delta^4\left(p_1^\mu + p_2^\mu - \sum_{k\geq 2} p_k^\mu\right) \cdot \frac{f(x_1,\mu_F)f(x_2,\mu_F)}{x_1 x_2 s} \cdot \left|\mathcal{M}_\alpha(p_k^\mu)\right|^2 \cdot W(\Phi'|\Phi_\alpha)$$

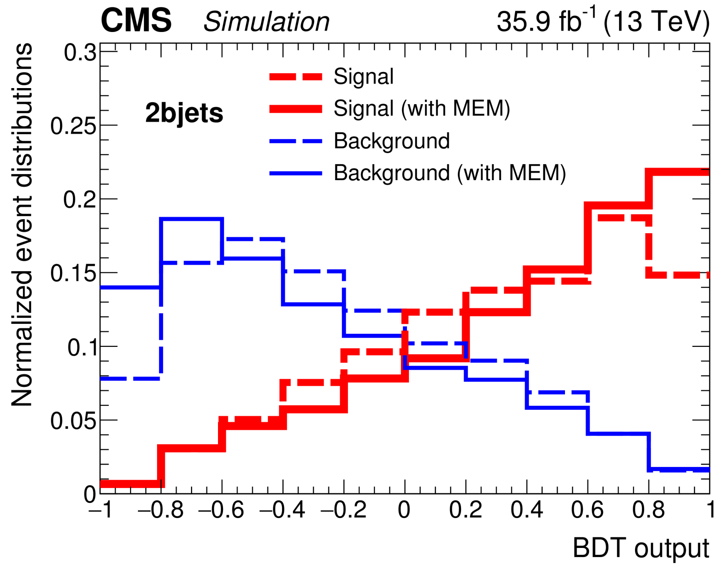
Observation 3.7σ
(3.2σ expected)

Maximize instead of integrating

MEM improves the analysis significance by 20%







Conclusions and perspectives

The **Matrix Element Method** is **used extensively at the LHC** in top quark sector

- Can be used to **measure observables** (top mass, spin correlation), or to **discriminate against background**

- **MEM is complementary to machine learning**: many analyses combine them to get maximum information from the detector

- **Many new developments:** MEM and Deep Neural Network, MEM using subjets, MEM as kinematic fit, MEM running on GPU…

Perspectives

- How to **speed up MEM** evaluation ? GPU, use of regressions….

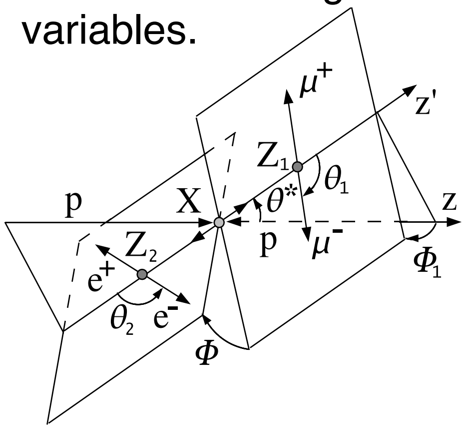
- Theory for **MEM at NLO** is now available *[e.g. JHEP1211(2012)043, JHEP09(2015)083]*, and remains to be tested in experiments

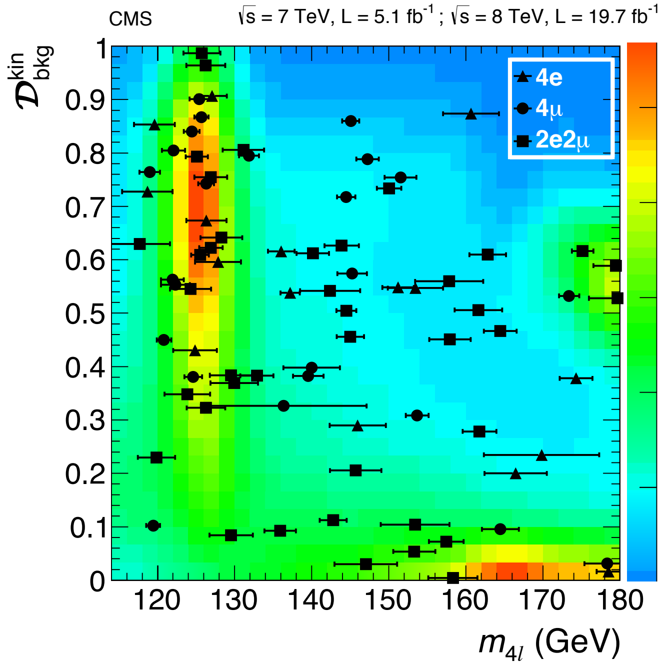
Back-up slides

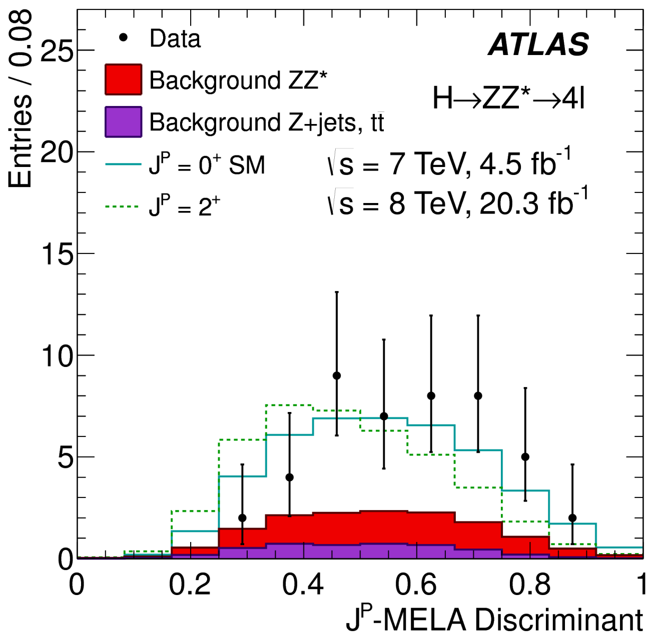
Kinematic discriminant with MELA in H→ZZ

ATLAS and CMS

4l decay kinematics: can be fully reconstructed. Most of the information in invariant mass of Z1 and Z2 and 5 angular variables.



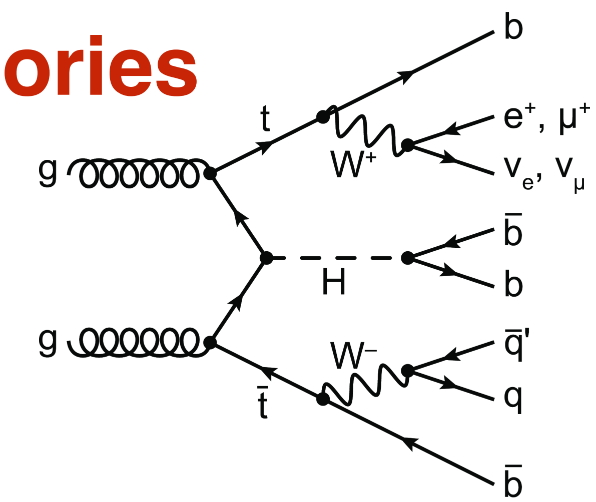




CMS $t\bar{t}H, H\rightarrow bb$: categories

12.9 fb-1

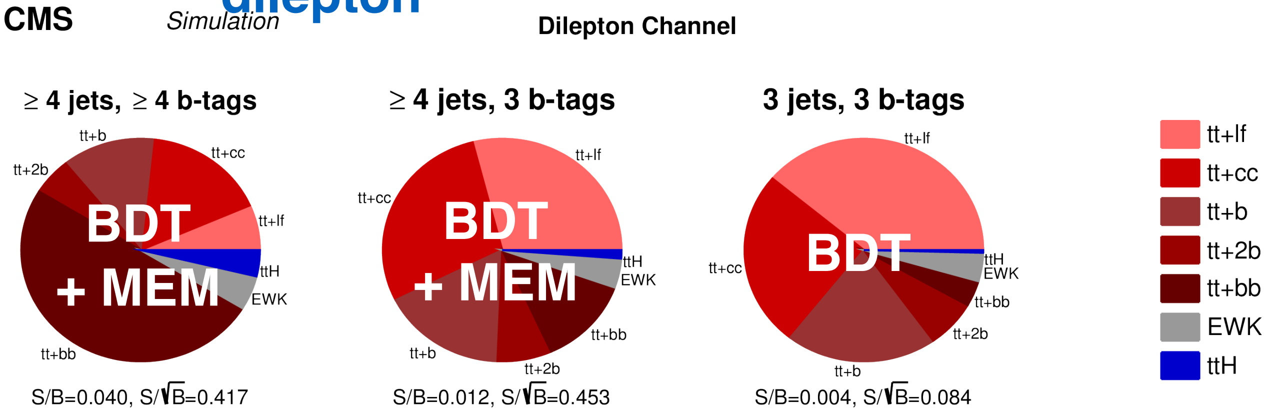
CMS HIG-16-038



Analysis targeting lepton+jets and dileptons

- **l+jets:** =1 lepton, ≥4 jets, ≥3 b-tag
- **2l:** 2 opposite sign lepton, ≥3 jets, ≥3 b-tag

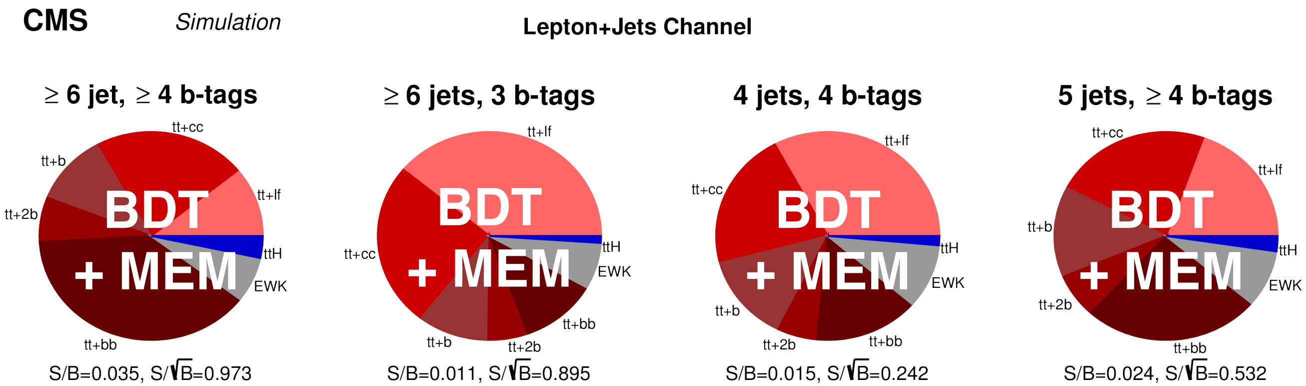
dilepton



Change relative to Moriond 2016 (2.3 fb-1):

- Re-optimize, remove low significance categories
- Use **Matrix Element Method** (MEM) as final discriminant in **low/high BDT** score categories

lepton+jets



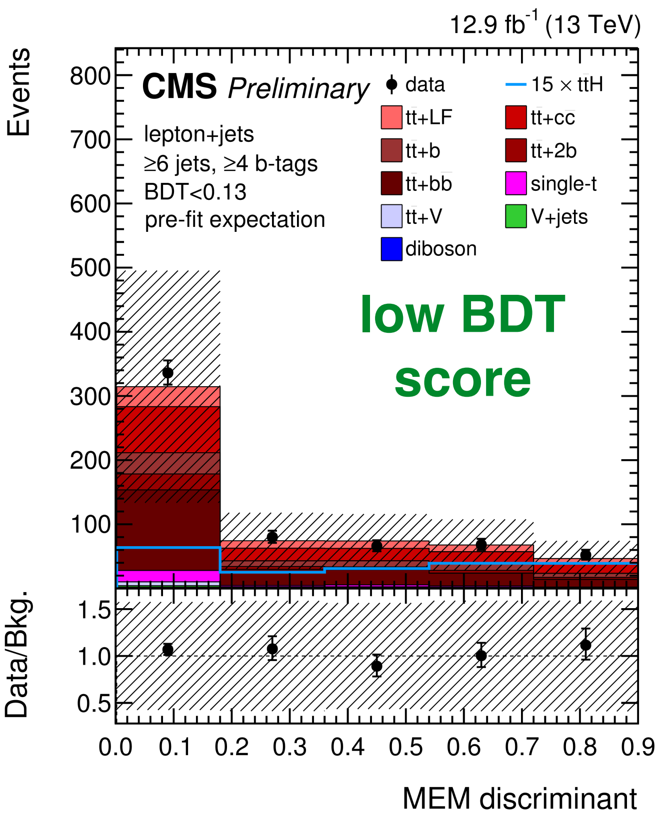
CMS $t\bar{t}H, H \rightarrow b\bar{b}$

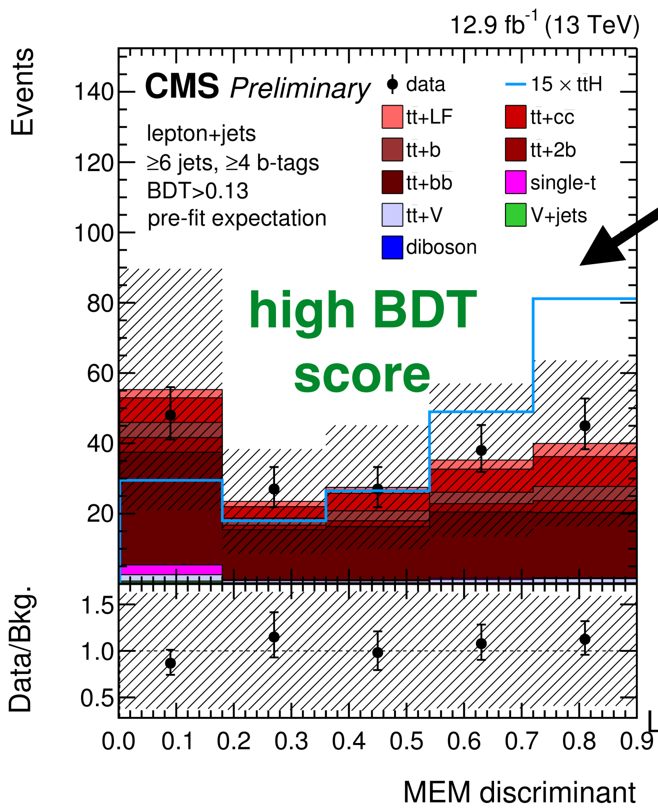
CMS HIG-16-038

12.9 fb-1

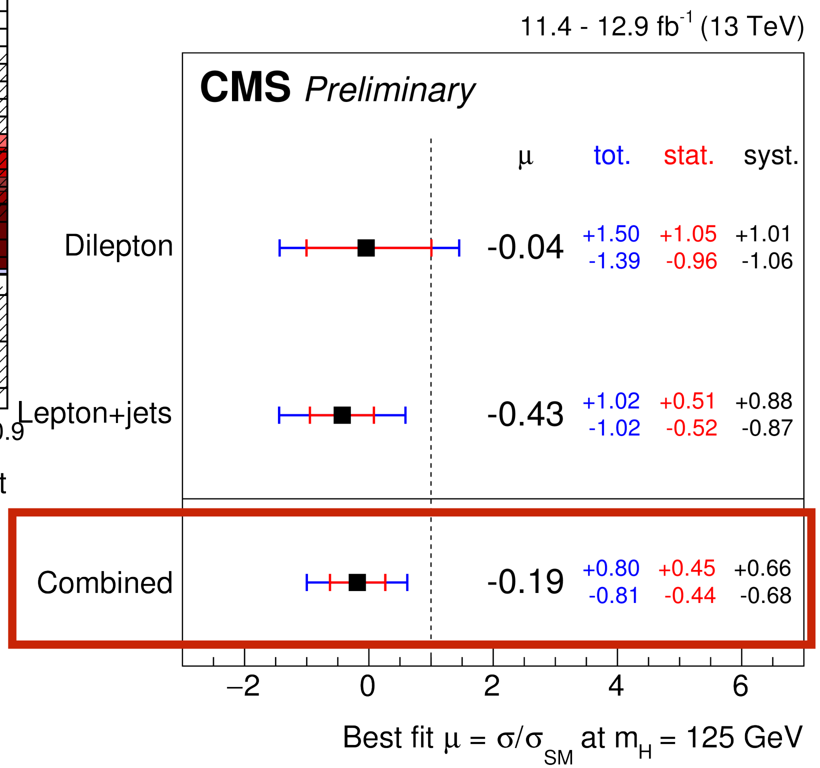
Analysis strategy:

- Split signal regions in low/high BDT parts
- Use Matrix Element Method as discriminant





Example of discriminants for 6j, 4b category (most discriminant among l+jets)



Systematics dominated:
50% uncertainty on ttbar+heavy flavour

CMS $t\bar{t}H, H \rightarrow b\bar{b}$

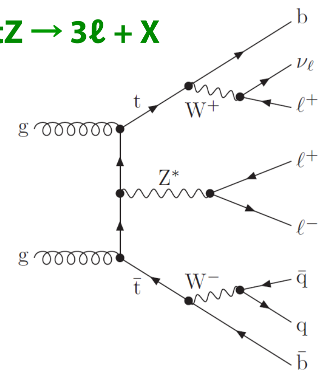
arxiv:1804.03682, submitted to JHEP

Channel	Method	Best-fit μ $\pm$tot ($\pm$stat $\pm$ syst)
Single-lepton	BDT+MEM	$1.0^{+0.69}_{-0.66}$ $\left(^{+0.31}_{-0.30}\ ^{+0.62}_{-0.59}\right)$
Single-lepton	DNN	$1.0^{+0.58}_{-0.55}$ $\left(^{+0.30}_{-0.29}\ ^{+0.50}_{-0.47}\right)$
Dilepton	BDT+MEM	$1.0^{+1.22}_{-1.12}$ $\left(^{+0.65}_{-0.62}\ ^{+1.04}_{-0.93}\right)$
Dilepton	DNN	$1.0^{+1.38}_{-1.36}$ $\left(^{+0.71}_{-0.69}\ ^{+1.18}_{-1.18}\right)$
Combined	BDT+MEM	$1.0^{+0.60}_{-0.57}$ $\left(^{+0.28}_{-0.27}\ ^{+0.53}_{-0.51}\right)$
Combined	DNN	$1.0^{+0.55}_{-0.51}$ $\left(^{+0.27}_{-0.27}\ ^{+0.47}_{-0.44}\right)$

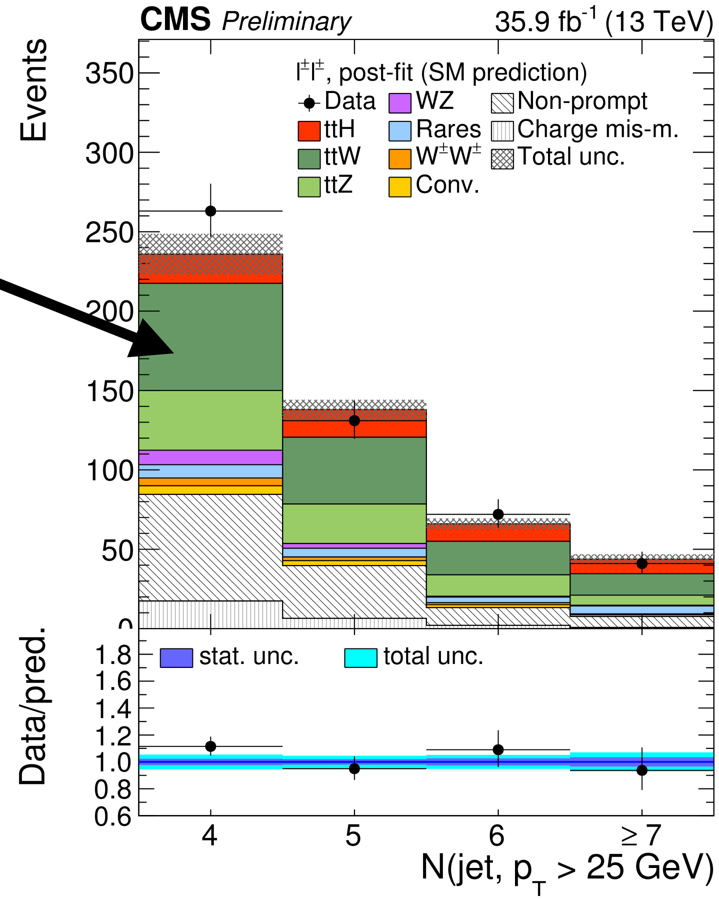
$t\bar{t}$H multilepton : Matrix Element Method

IPHC, IPNL

ttZ → 3ℓ + X



Irreducible: tt+W/Z/γ*
- from Monte Carlo,
- O(10%) uncertainty



MEM for ttH, ttW, ttZ/γ* hypotheses:

- **Custom framework in C++**
- Assume narrow-width for Top quark and Higgs boson
- Treat final-state b from top as massive
- Keep full W and Z propagators in the top ME: follows a Breit-Wigner
- Z and γ* contributions included

- **If jets are needed at ME level and are not reconstructed ("mising jets"):** included, as supplementary phase space to integrate
- MEM weight is the average weight of all possible **lepton, jets, b-jets permutations**