Status of 4Top analysis

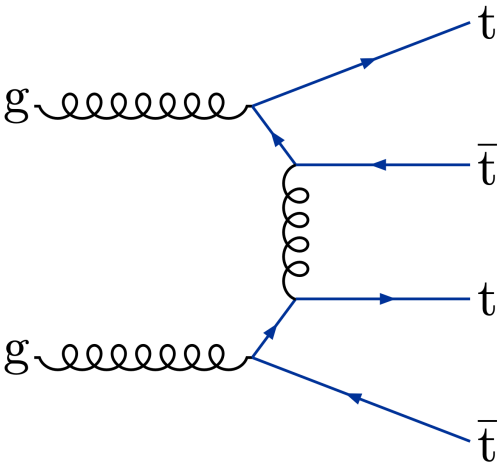
Huiling Hua[1], Fabio Iemmi[1], Hongbo Liao[1], Hideki Okawa[2], Yu Zhang[2]

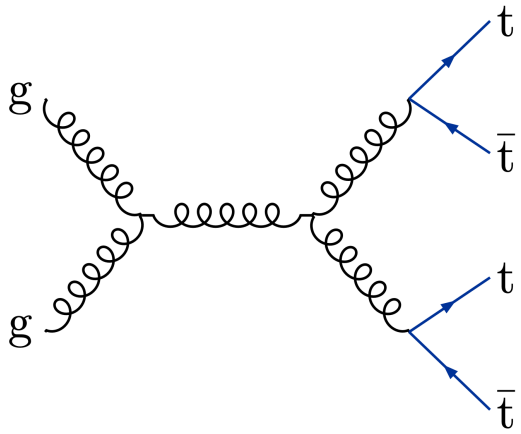
1,IHEP, 2,Fudan University

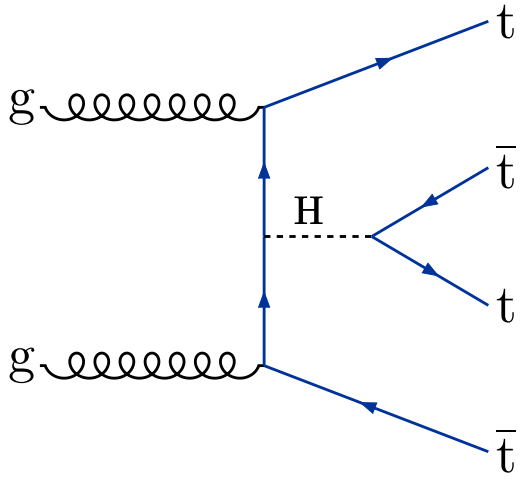
May 26, 2021

On CMS China meeting, at Hangzhou on 28-29th May 2021

Motivation







- It is predicted by SM but not observed.
- The cross section is sensitive to the magnitude and CP properties of top Yukawa coupling Y_t
- BSM such as 2HDM and SUSY can enhance the production
- Any advantage of tau final state?

Review the public result

Expected and observed significance in ATLAS and CMS.

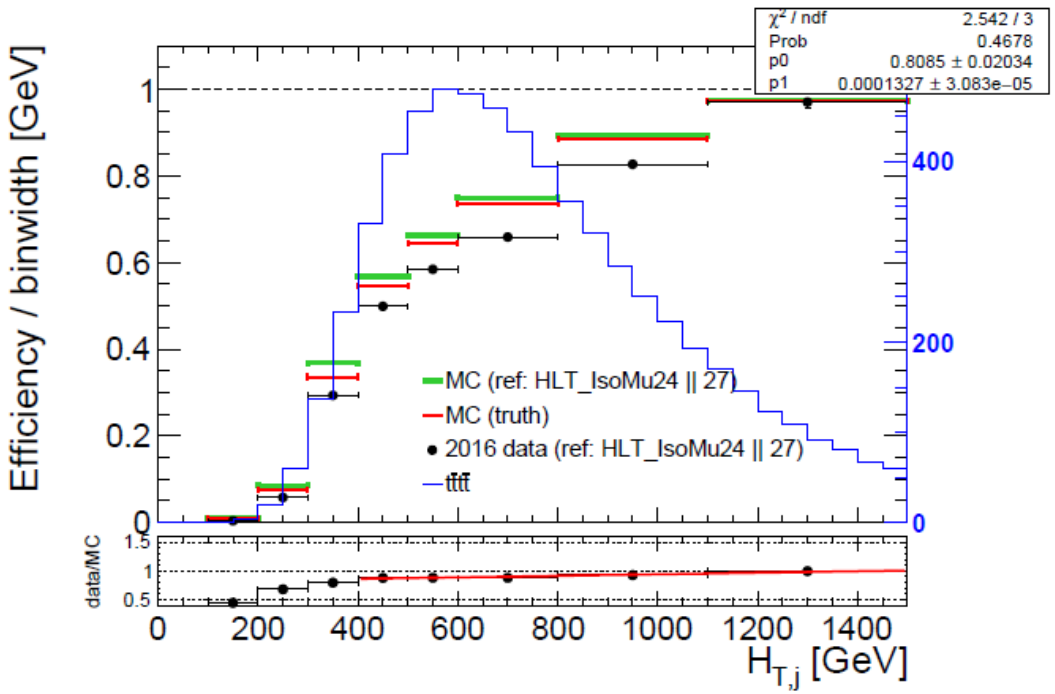
obs(exp)	ATLAS	CMS
full hadronic	-	(0.71)
1L OS2L	1.9(1.0)	(0.87) (0.88)
SS2L/ML	4.3(2.4)	2.7(2.6)
combined	4.7(2.6)	(3.16)

- There is no tau dedicated category.
- We first propose the tau final state in CMS.

Trigger

"OR" of three triggers

- HLT_PFHT450_SixJet40_BTagCSV_p056
- HLT_PFHT400_SixJet30_DoubleBTagCSV_p056
- HLT_PFJet450



- Preselection + 1μ
- **Data and MC are closer**
- **Fitted the ratio with a straight line of the form**

$$y = mx + q$$

$$m = 0.0001327 \pm 3.083\text{e-}05$$
$$q = 0.8085 \pm 0.02034$$

- **Also plotted H_T distribution for signal**

Object definition and pre-selection

Object Selection

	Electron	Muon	Tau	Jet
Kinematic	$pT > 10\ GeV$ $\|\eta\| < 2.5$	$pT > 10\ GeV$ $\|\eta\| < 2.4$	$pT > 20\ GeV$ $\|\eta\| < 2.3$	$pT > 25\ GeV$ $\|\eta\| < 2.4$
ID&ISO	Tight ID Tight ISO	Medium ID Medium ISO	DeepTau VSjet:Medium DeepTau VSmu:VLoose DeepTau VSe:VVVLoose	Loose Jet ID btagging : DeepFlavour Medium

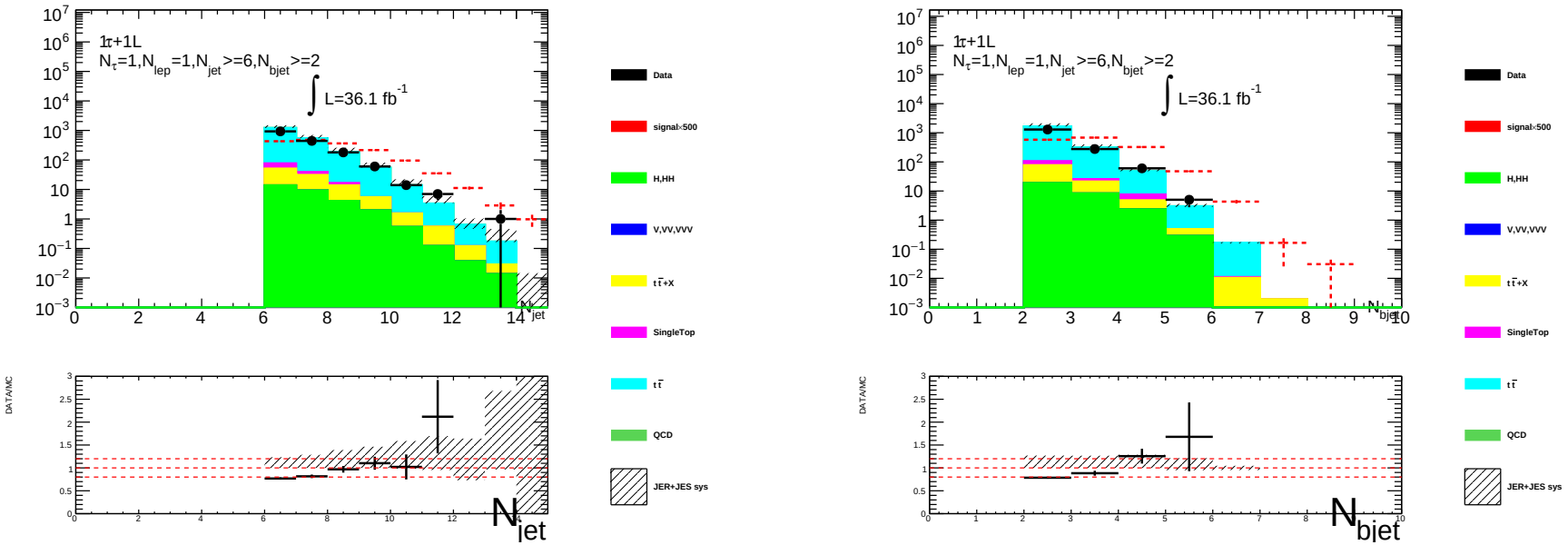
- Pre-selection of 7 tau categories

	Selection
1Tau+0L	$H_T > 400\ GeV, N_{lep} == 0, N_\tau == 1, N_{jet} \geq 8, N_{bjet} \geq 2$
1Tau+1L	$H_T > 400\ GeV, N_{lep} == 1, N_\tau == 1, N_{jet} \geq 6, N_{bjet} \geq 2$
1Tau+2L	$H_T > 400\ GeV, N_{lep} == 2, N_\tau == 1, N_{jet} \geq 4, N_{bjet} \geq 2$
1Tau+3L	$H_T > 400\ GeV, N_{lep} == 3, N_\tau == 1, N_{jet} \geq 2, N_{bjet} \geq 2$
2Tau+0L	$H_T > 400\ GeV, N_{lep} == 0, N_\tau == 2, N_{jet} \geq 6, N_{bjet} \geq 2$
2Tau+1L	$H_T > 400\ GeV, N_{lep} == 1, N_\tau == 2, N_{jet} \geq 4, N_{bjet} \geq 2$
2Tau+2L	$H_T > 400\ GeV, N_{lep} == 2, N_\tau == 2, N_{jet} \geq 2, N_{bjet} \geq 2$

Cut-flow and Data/MC comparison

category	data	signal	$t\bar{t}$	SingleTop	TTX	SM Higgs	HH	QCD	total bkg	(data-bkg)/bkg
1Tau+0L	14017	9.474	6283.881	0.261	105.070	57.212	0.027	7303.656	13750.109	0.019
1Tau+1L	2033	7.694	2336.646	0.649	65.640	35.305	0.003	2.359	2440.602	-0.167
1Tau+2L	78	1.843	63.006	0.568	19.501	6.501	0.002	0.000	89.578	-0.129
1Tau+3L	3	0.146	0.000	0.006	1.950	0.397	0.000	0.000	2.353	0.275
2Tau+0L	280	0.507	247.273	0.391	13.174	6.856	0.000	1.418	269.112	0.040
2Tau+1L	21	0.260	22.571	0.140	6.450	3.553	0.000	0.000	32.715	-0.358
2Tau+2L	1	0.031	0.271	0.004	0.737	0.362	0.000	0.000	1.375	-0.273

The MC is higher than data in some high statistic categories.



- There is a modeling problem in $t\bar{t}$ background.

Background estimation

MC correction : It is known that $t\bar{t}$ MC sample is mismodeling

- $t\bar{t}$ has more jets in MC than that in data.
- The fraction of $t\bar{t} + b\bar{b}$ is higher in data.
- The top pT is harder in MC than that in data.

Data-driven background

- The QCD background in 1Tau/2Tau+0L should be estimated by data-driven.
- We can use the tau ID and N_{bjet} to define the so-called "ABCD" region.
- It is under investigation.

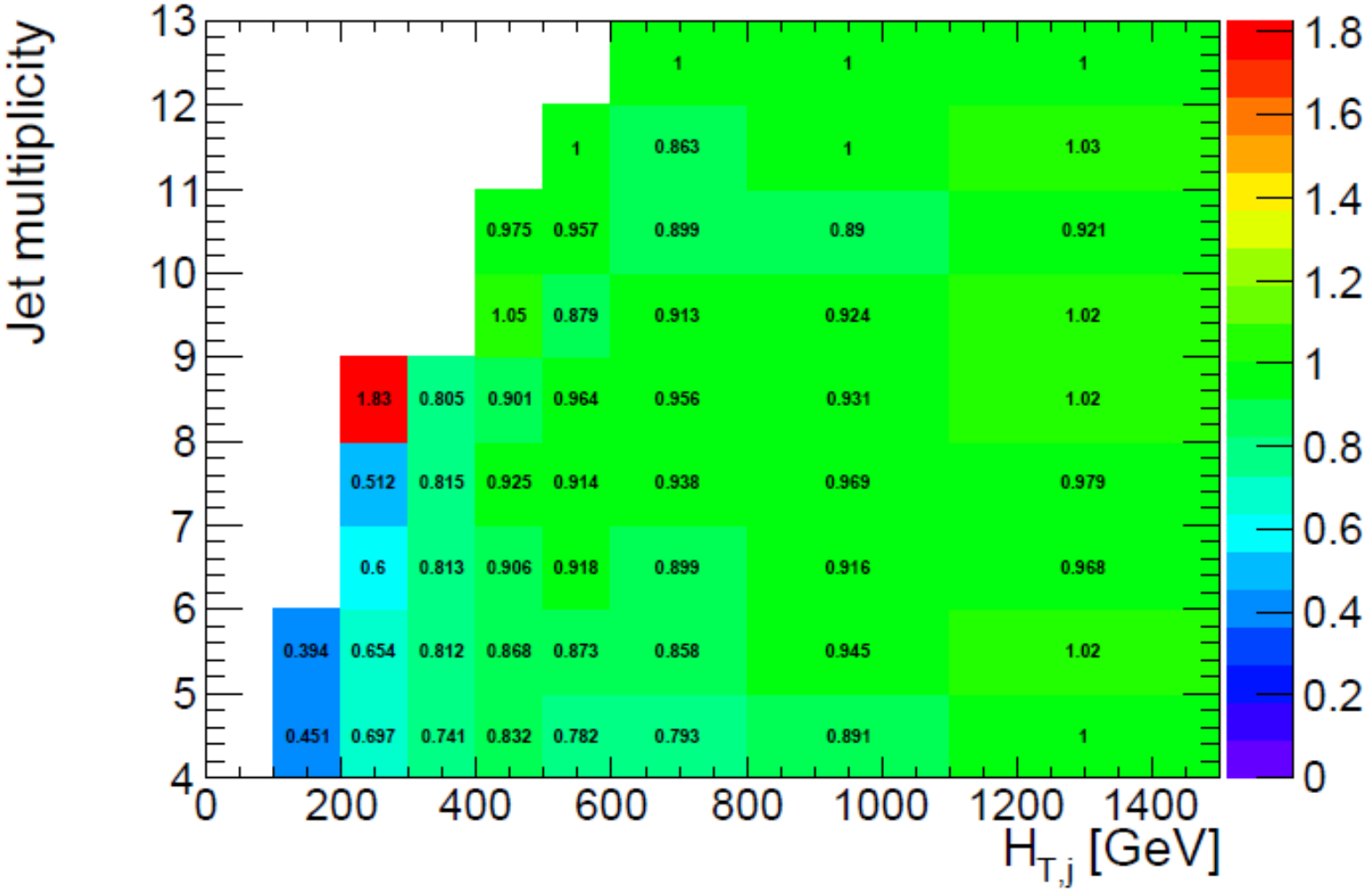
BDT optimization

Summary

- What has been done :
  - Propose the 4Top tau channel in CMS
  - Determine the trigger and pre-selection
  - Finish a very preliminary BDT optimizatation
- Next step :
  - Background estimation
  - BDT optimization
  - Systematics

Backup

Trigger Scale Factor



TauID

	VVT	VT	T	M	L	VL	VVL	VVVL
S	3	5	6	8	10	15	21	19
B	839	1209	1728	2528	4028	7422	13164	12221
Z	0.1178	0.1325	0.1451	0.1561	0.1652	0.1793	0.1867	0.1710
Z_{syst}	0.019	0.018	0.016	0.014	0.012	0.0097	0.0075	0.0072
E(sig)	0.8786	0.8751	0.8613	0.8488	0.8222	0.7675	0.6947	0.6868
E(bkg)	0.9415	0.9302	0.9146	0.8903	0.8537	0.7723	0.6694	0.6688
P(sig)	0.9076	0.8720	0.8278	0.7681	0.6781	0.5339	0.4040	0.3984
P(bkg)	0.7461	0.6648	0.5779	0.4801	0.3611	0.2312	0.1405	0.1403