

University of
Zurich^{UZH}

SEARCH FOR DIJET RESONANCES IN EVENTS WITH THREE JETS

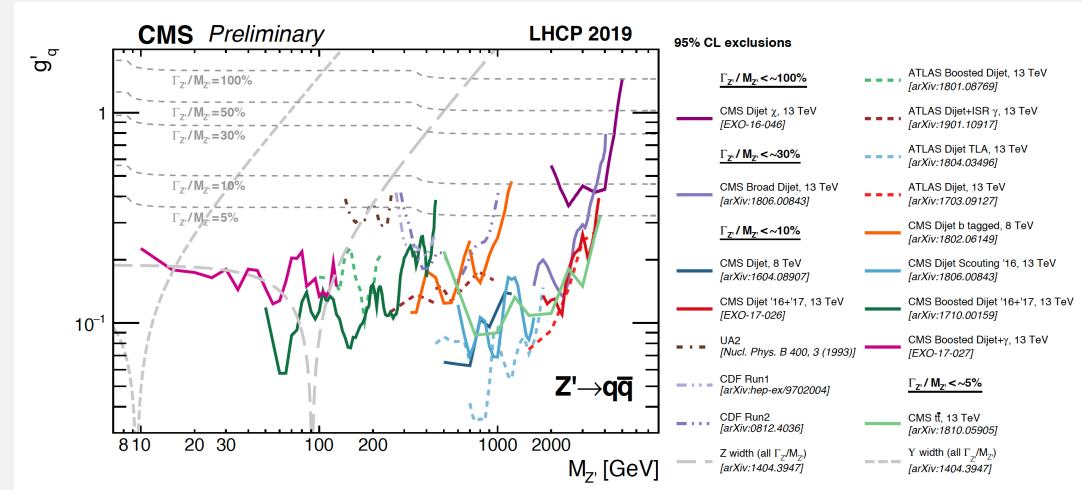
Danyyl Brzhechko

Zurich PhD seminar 2019

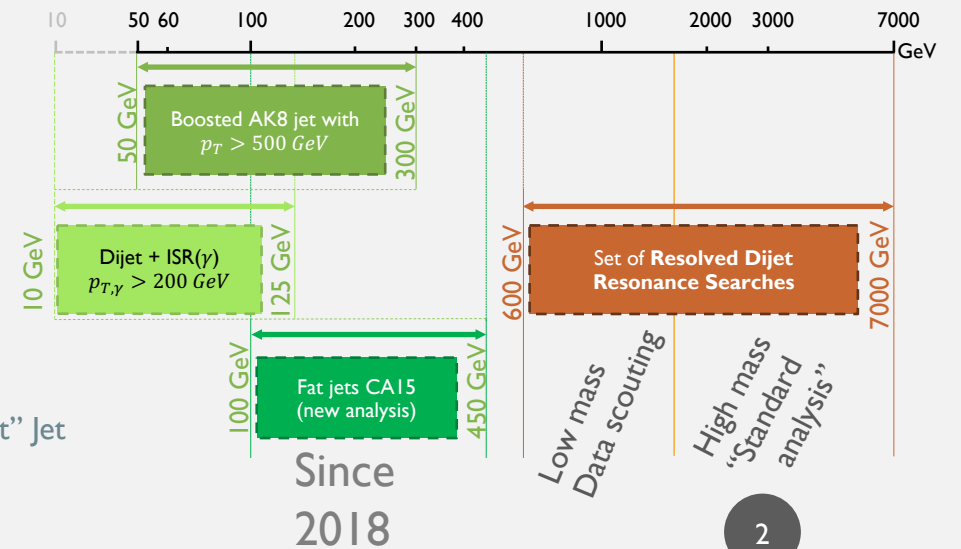
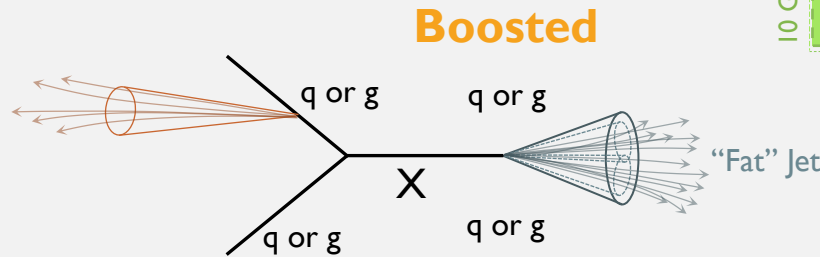
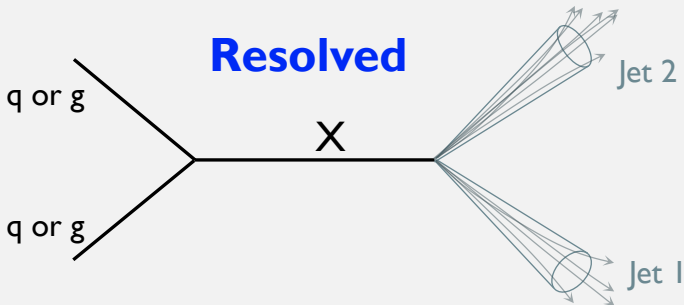
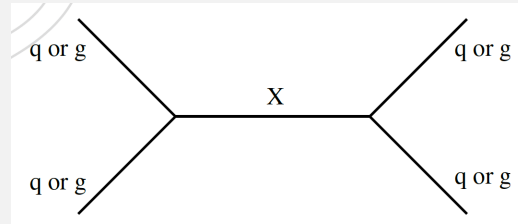
October 10, 2019

DIJET ANALYSIS OVERVIEW

- **Dijet searches** is a big group of analyses, which aims to **discover BSM particles** that can couple to **2 partons**
- Bunch of models considered...
- 2 regimes: boosted (<300 GeV) and resolved (>600 GeV) (as of 2018)
- Dijet mass region of 300-600 GeV was not covered by CMS at $\sqrt{s} = 13 \text{ TeV}$ (up to 2018)

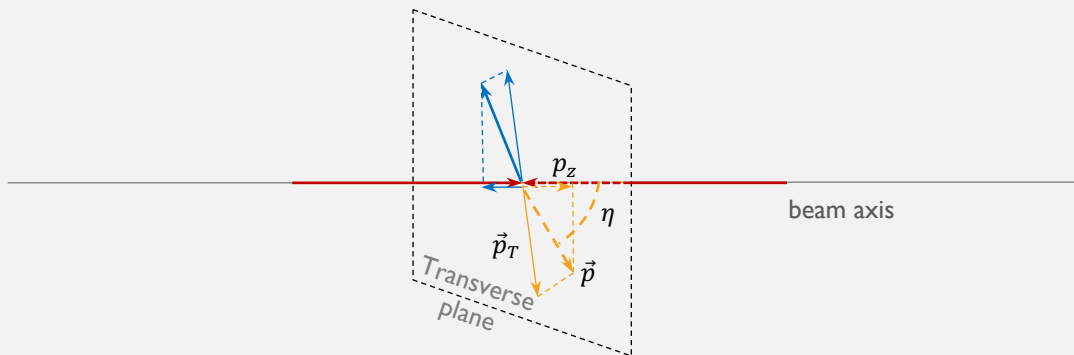
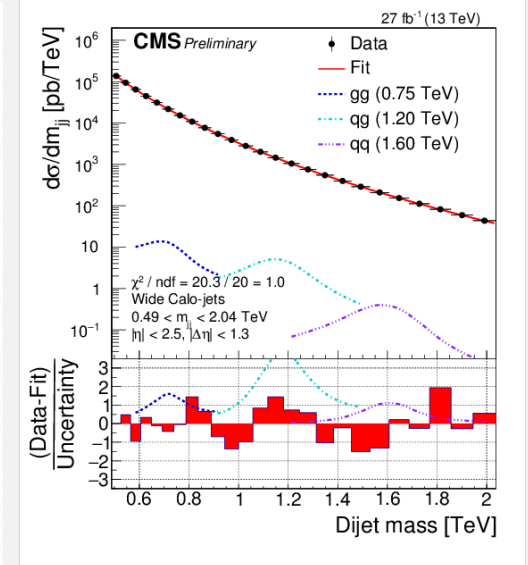
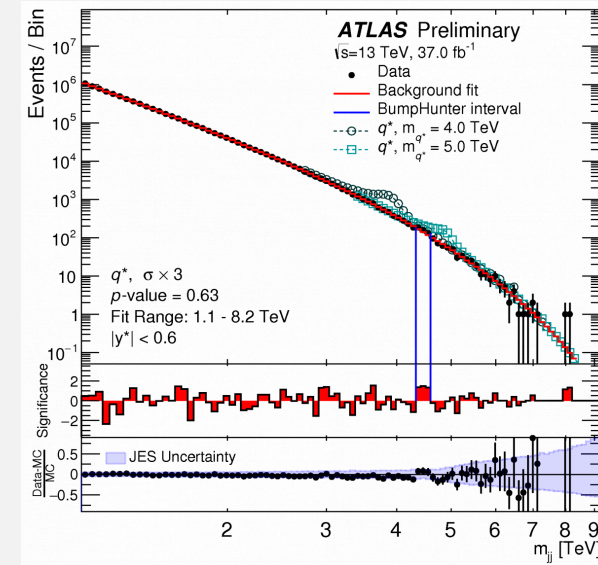


Model Name	X	Color	J^P	$\Gamma / (2M)$	Decay Channel
Axigluon	A	Octet	1^+	0.04	$q\bar{q}$
E_6 Diquark	D	Triplet	0^+	0.004	qq
Coloron	C	Octet	1^-	0.04	$q\bar{q}$
Heavy W	W'	Singlet	1^-	0.01	$q\bar{q}$
Heavy Z	Z'	Singlet	1^-	0.01	$q\bar{q}$
Dark Matter Mediator		Singlet	1^\pm	0.02	$q\bar{q}$



RESOLVED DIJET SEARCH IN THREE-JET FINAL (DIJET+ISR)

- **Resolved dijet search in three-jet final state (Dijet+ISR shortly)** is a resolved dijet analysis requiring **additional jet in the final state**
- Covers the region of 350 to 700 GeV of the mediator mass. 300(450)-600 GeV is not covered by CMS $\sqrt{s} = 13$ TeV analysis
- Data scouting + additional jet in FS \rightarrow 100% trigger efficiency starting from 300 GeV of dijet mass
- ISR jet is the third leading jet in p_T (transverse momentum)
- Dijet is 2 jets with the highest p_T



DATA SCOUTING

- Two-tiered trigger system:
 - **L1 (level 1) trigger:** 40 MHz \rightarrow 100kHz.
 - **HLT (high level trigger):** decreases rate to 1kHz before data storage
 - \rightarrow **Too small rate for some physics analyses !**
- **Data scouting** – higher rate with the suitable performance for physics analysis
 - Reduced amount of data per event
 - Types: **Calo scouting; PF-scouting.**
- **Calo-scouting** technique used in particular by Dijet+ISR jet search. Allows to select events with $H_T = \sum_i p_T > 250 \text{ GeV}$.
 - Ability to extend resolved dijet search coverage from 1.5 TeV to 0.5 TeV.

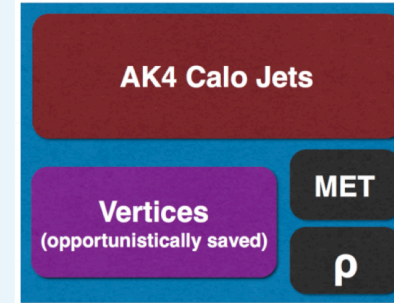
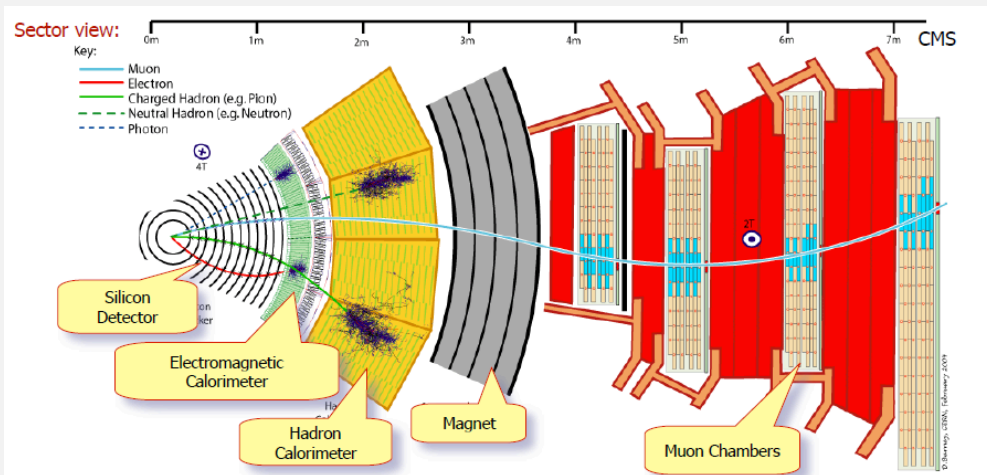
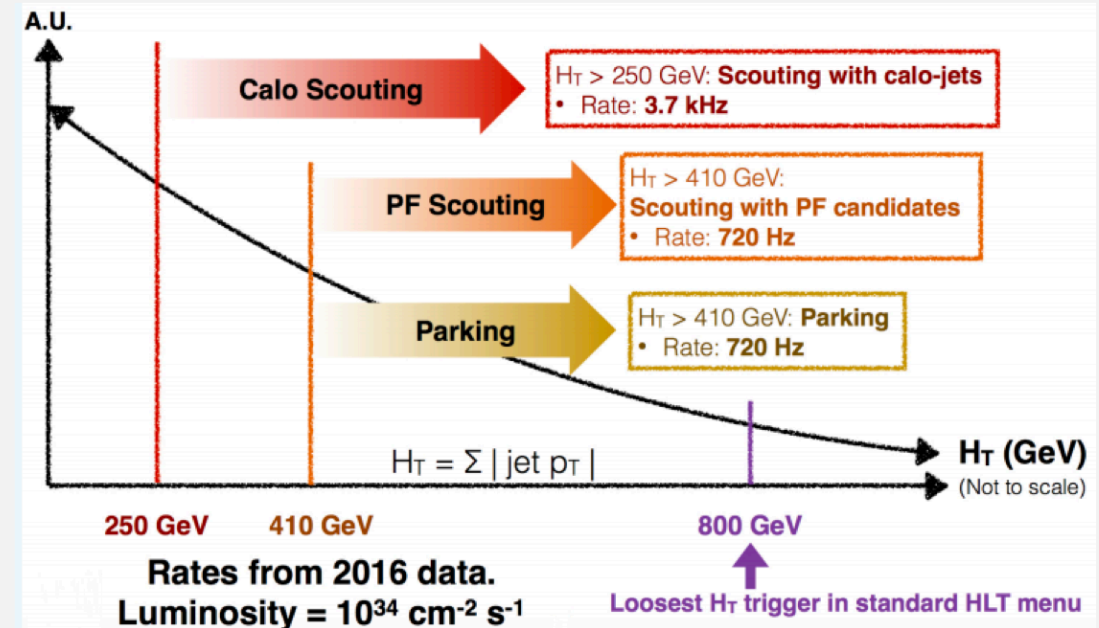


Figure 1: Calo Scouting. Event size \approx 1.5 kB

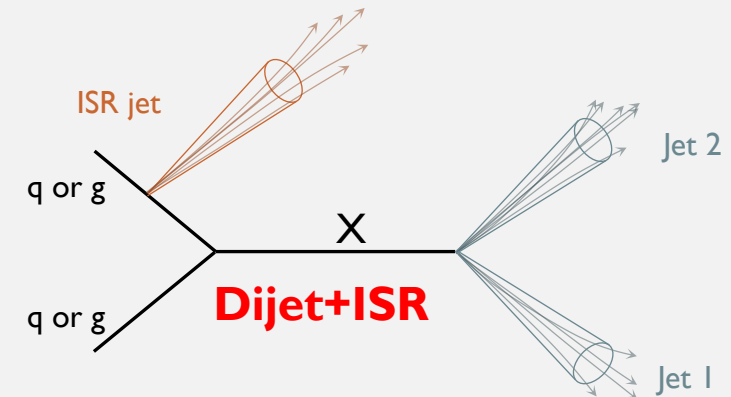
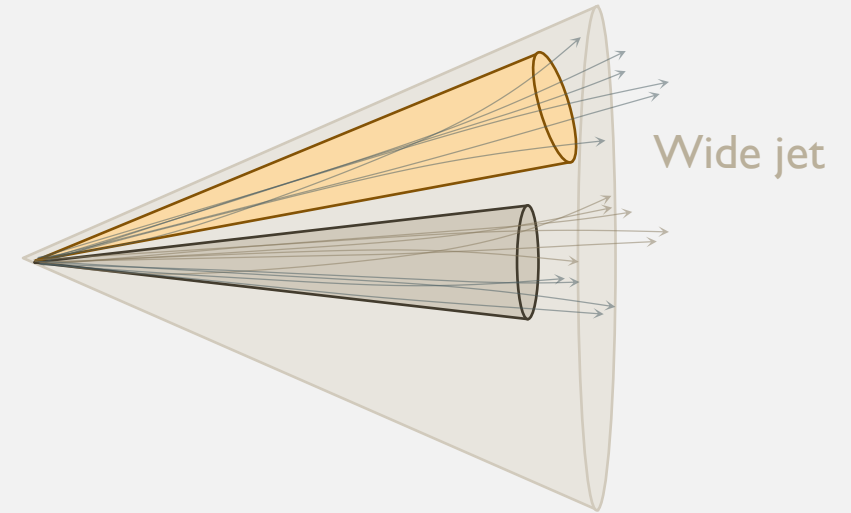


Figure 2: PF Scouting. Event size \approx 10 kB



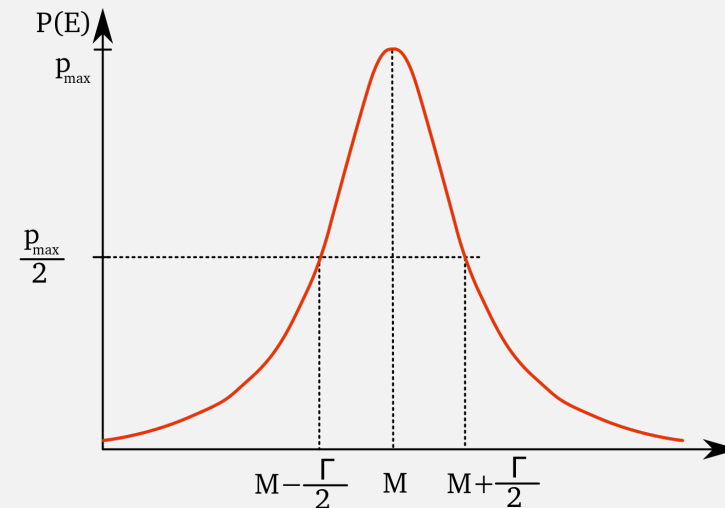
EVENT RECONSTRUCTION AND SELECTION

- Preliminary selections before wide-jet reconstruction:
 - $p_T > 30 \text{ GeV}$ and $|\eta| < 2.5$ for all AK4 Calo
- Wide jet reconstruction:
 - Recovers strong final state radiation (FSR) jets not included into AK4 jet.
 - $\Delta R_{WJ} < 1.1$
- $n_{Jets} = 3$
- The two wide jets with the highest p_T is selected as a dijet. Third jet is initial state radiation (ISR) jet.
- $|\Delta\eta| < 1.1$ – suppresses QCD background and increases trigger efficiency
- Third jet selection: $p_T > 72 \text{ GeV}$ – defined by **automatic procedure**
- Fit range: 290-1000 GeV – lower edge is defined by **automatic procedure**



DIJET MASS SIGNAL SHAPE STUDIES

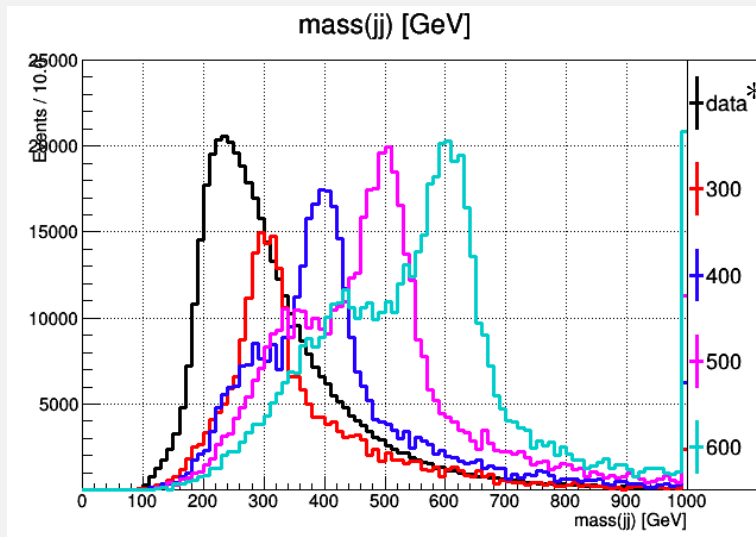
- Signal: narrow Breit-Wigner resonance



DIJET MASS SIGNAL SHAPE STUDIES

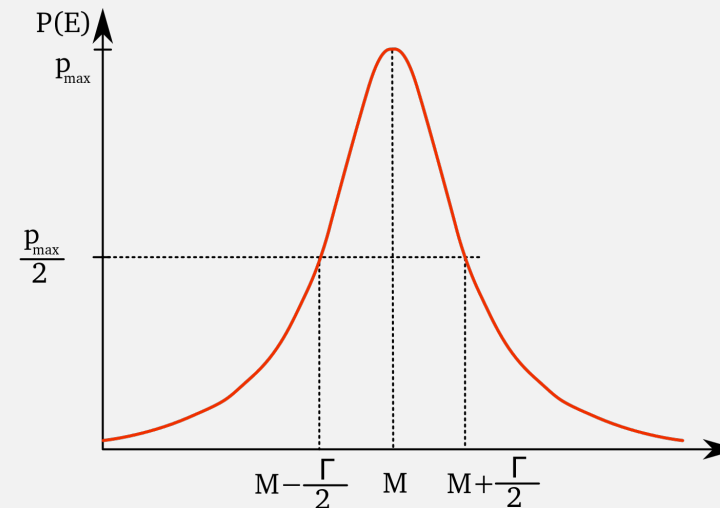
- Signal: narrow Breit-Wigner resonance
- Shape is not gaussian-like due to the following reasons:

Real shape



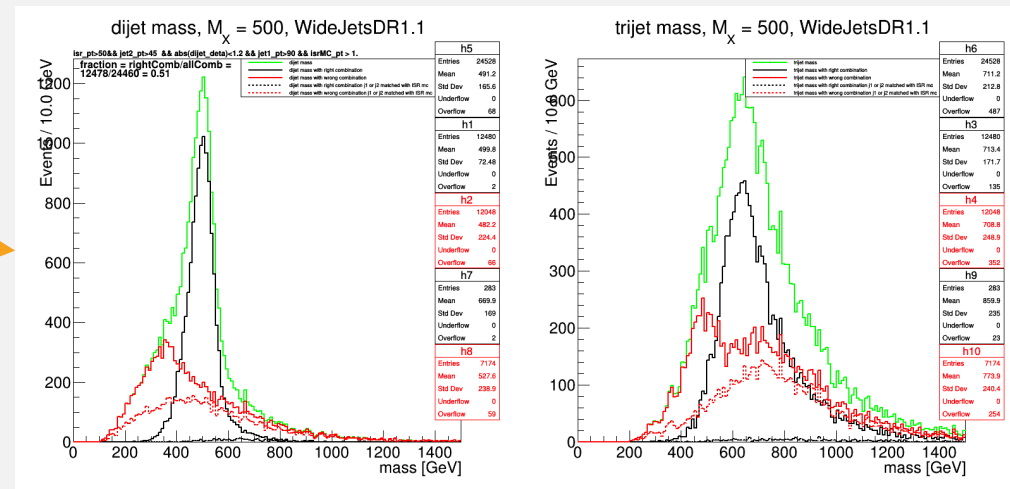
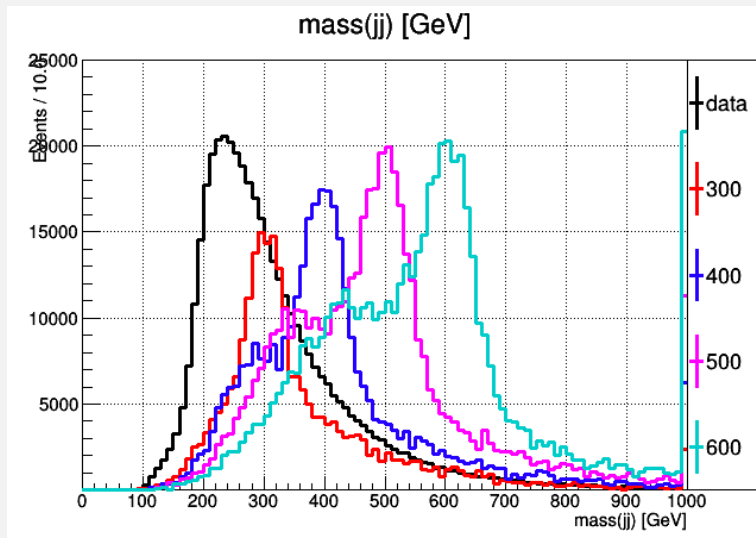
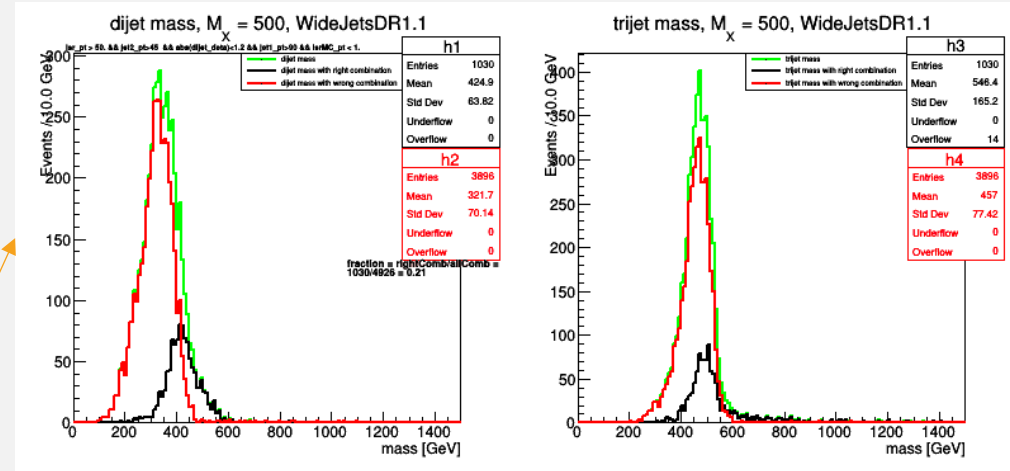
* Shape of the data

Should be similar to ...



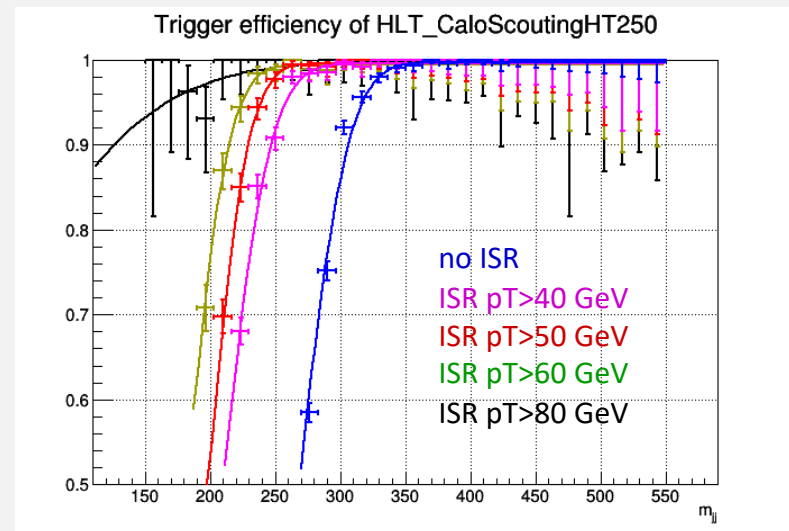
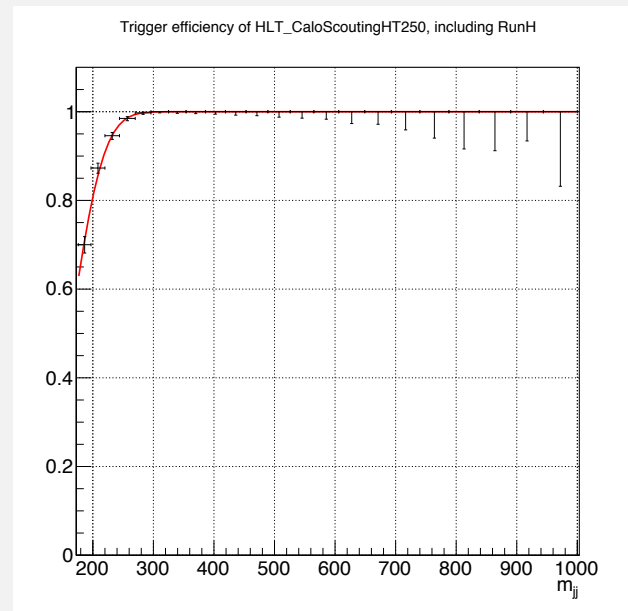
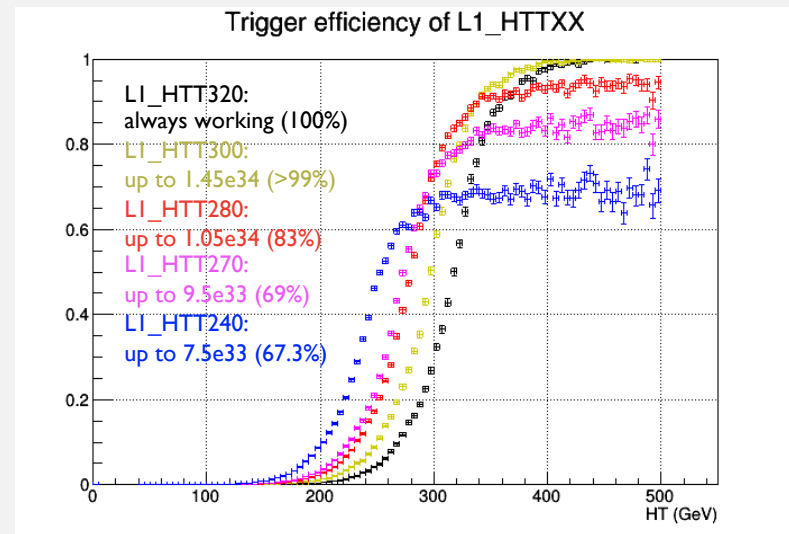
DIJET MASS SIGNAL SHAPE STUDIES

- Signal: narrow Breit-Wigner resonance
- Shape is not gaussian-like due to the following reasons:
 - Not all FSR jets are included inside WJ cone
 - FSR is misinterpreted as ISR
- Change of radius of the WJ cone did not help



TRIGGER

- LI trigger: LI_HTT240 (selects events with $H_T > 240 \text{ GeV}$) in order to go lower in dijet mass
- Was turned on in 67.3% of the run – decreases luminosity to 21.2 fb^{-1} . Issue with LI_HTT in RunH: ended up with 18.3 fb^{-1}
- Events with $H_T > 250 \text{ GeV}$ were selected by HLT trigger.
- Requiring the third jet shifts the trigger “plateau” significantly
- Can we go even lower in dijet mass increasing ISR pT cut?



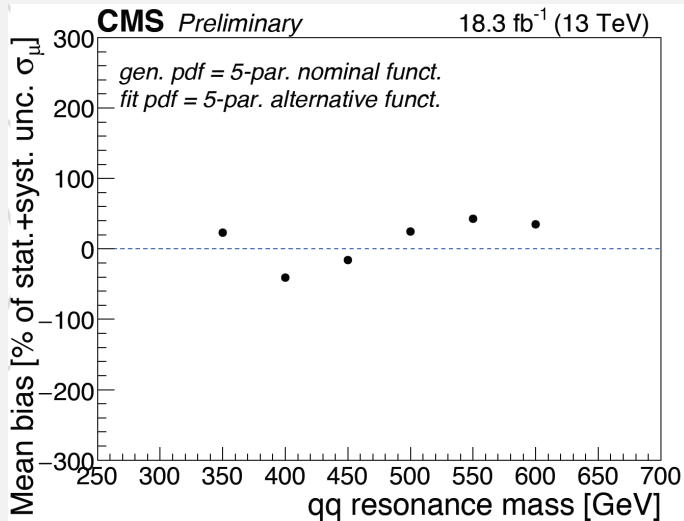
BACKGROUND ESTIMATION

- Background fit is done using a smooth function:

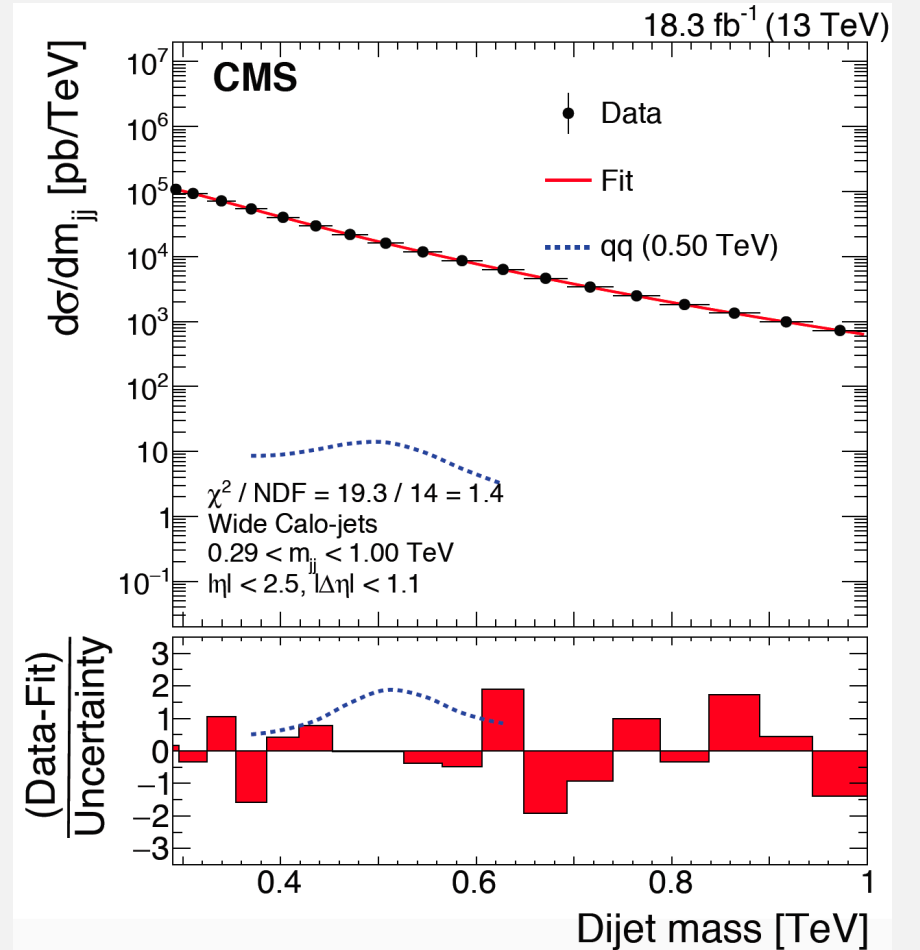
$$\frac{d\sigma}{dm_{jj}} = p_0 \frac{p_2 x - 1}{x^{p_1 + p_3 \log(x) + p_4 \log(x^2)}}$$

starting from 290 GeV of the dijet mass.

- Bias study shows if background function to fit bkg+signal (up to 40%)
- Fisher test is performed to choose the number of parameters (5)
- The data is fitted using Combine Higgs Tool with profile likelihood method
- Fit range, function and ISR pT cut is not trivial at all

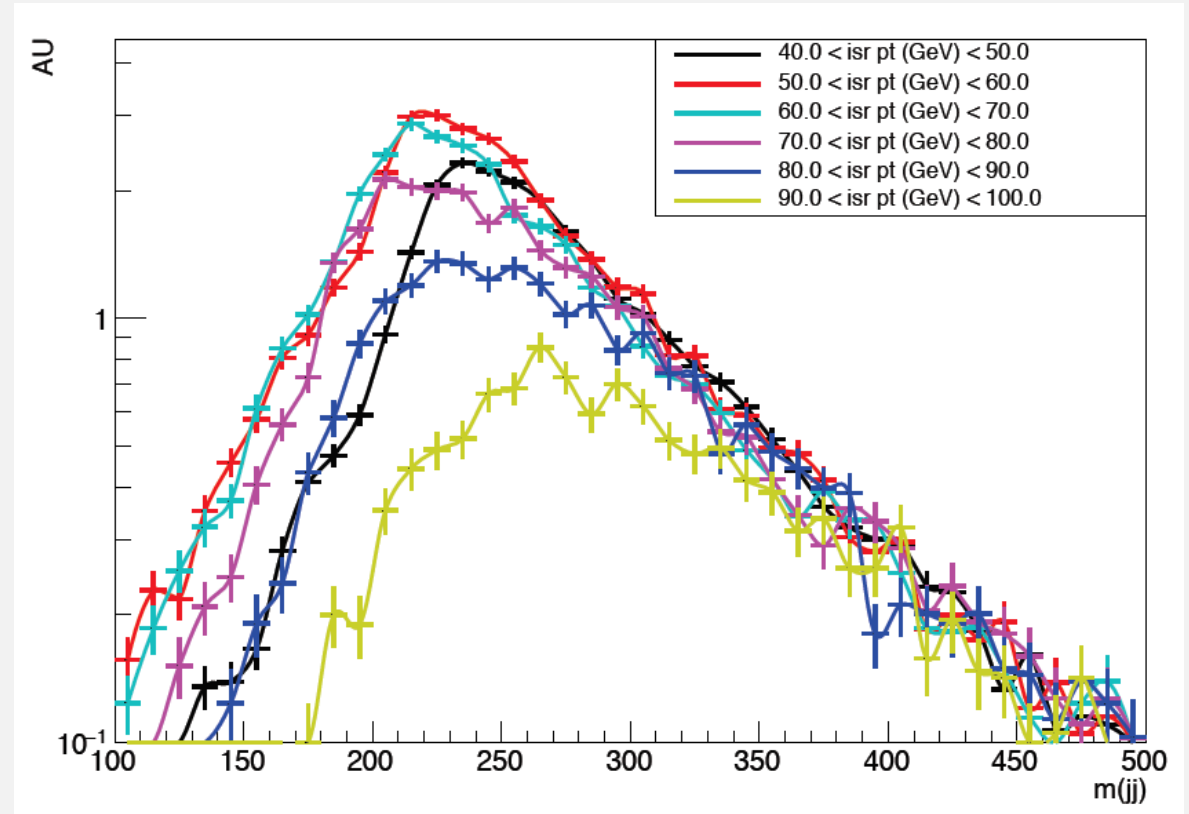


Nominal	
3 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{(p_2 x - 1)}{x^{p_1}}$
4 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{e^{\log(p_2 x - 1) + p_3 x}}{x^{p_1}}$
5 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{e^{\log(p_2 x - 1) + p_3 x + p_4 x^2}}{x^{p_1}}$
6 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{e^{\log(p_2 x - 1) + p_3 x + p_4 x^2 + p_5 x^3}}{x^{p_1}}$
Alternative	
3 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{(p_2 x - 1)}{x^{p_1}}$
4 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{(p_2 x - 1)}{x^{p_1 + p_3 \log x}}$
5 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{(p_2 x - 1)}{x^{p_1 + p_3 \log x + p_4 \log^2 x}}$
6 Par	$\frac{d\sigma}{dm_{jj}} = p_0 \cdot \frac{(p_2 x - 1)}{x^{p_1 + p_3 \log x + p_4 \log^2 x + p_5 \log^3 x}}$



DIJET MASS DISTRIBUTION SHAPE STUDIES

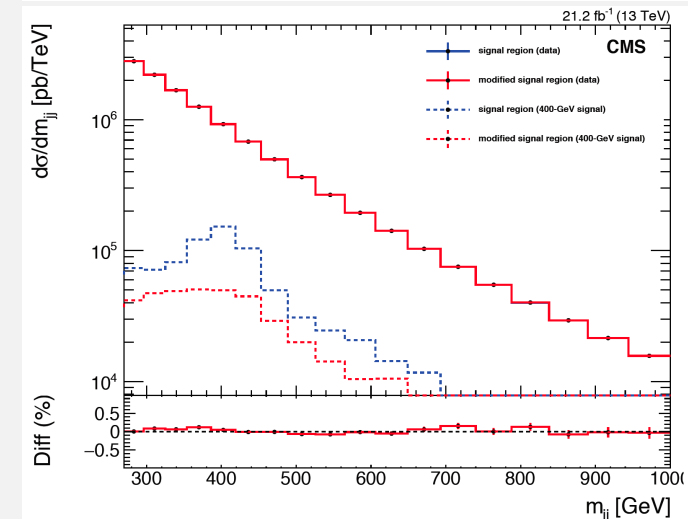
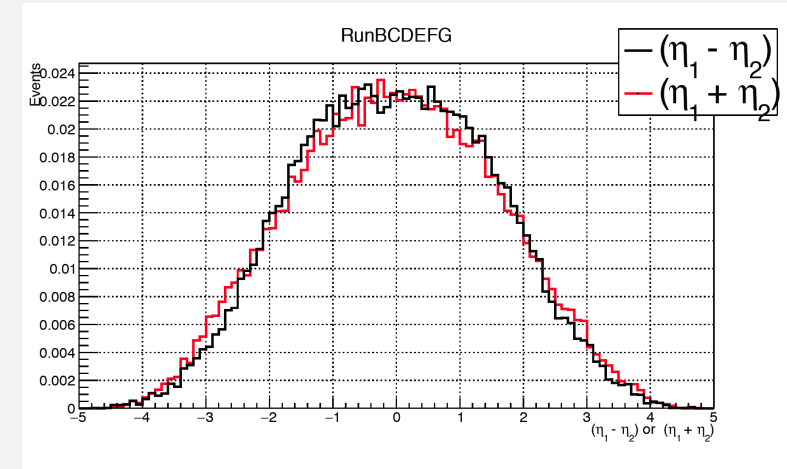
- Dijet spectrum for three jet final state is different than for two-jet FS.
- Dijet mass distribution is shaped because of:
 - Trigger: lower $p_{T,3}$ cut \rightarrow lower H_T \rightarrow less events for lower mass \rightarrow dijet spectrum shifts to the higher values
 - Kinematic selection: $p_{T1} > p_{T2} > p_{T3} >$ ISR p_T cut $\rightarrow M^2 = 2p_{T1}p_{T2}(\cosh(\Delta\eta(jj)) - \cos(\Delta\phi)) \sim 2p_{T1}p_{T2} \cosh(\Delta\eta(jj)) \rightarrow$ higher $p_{T,3}$ cut shifts dijet spectrum to the higher values.
- The selection on ISR p_T and fit range is not trivial at all
- The automatic procedure to find ISR p_T selection and fit range was developed



Dijet mass distribution shapes within different ISR p_T selection.

DEFINING MODIFIED SIGNAL REGION

- Ideal way: make all the tests on background only distribution using MC samples.
- Issues:
 - Huge statistics of calo-scouting data
 - Not able to reproduce our data at the precision level given by statistical uncertainty
- Solution: use of data-driven approach
- By flipping the sign of η_2 , we get similar dijet mass spectrum
- Signal plot $\Delta\eta = \eta_2 - \eta_1$; control plot $\Delta\eta \rightarrow \eta_2 + \eta_1$
- Modified signal region $|\eta_2 + \eta_1| < 1.1$ corrected using bin-by-bin correction that depends on $p_{T,1}p_{T,2}$ using signal region $|\eta_2 - \eta_1| < 1.1$ dijet mass shape:
 - MSR background shape is compatible with SR within $\sim 0.1\%$
 - No clear narrow peak from signal
 - Signal contamination reduced by a factor of 3



AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION

- **Recursive part:**

- **Step-1**

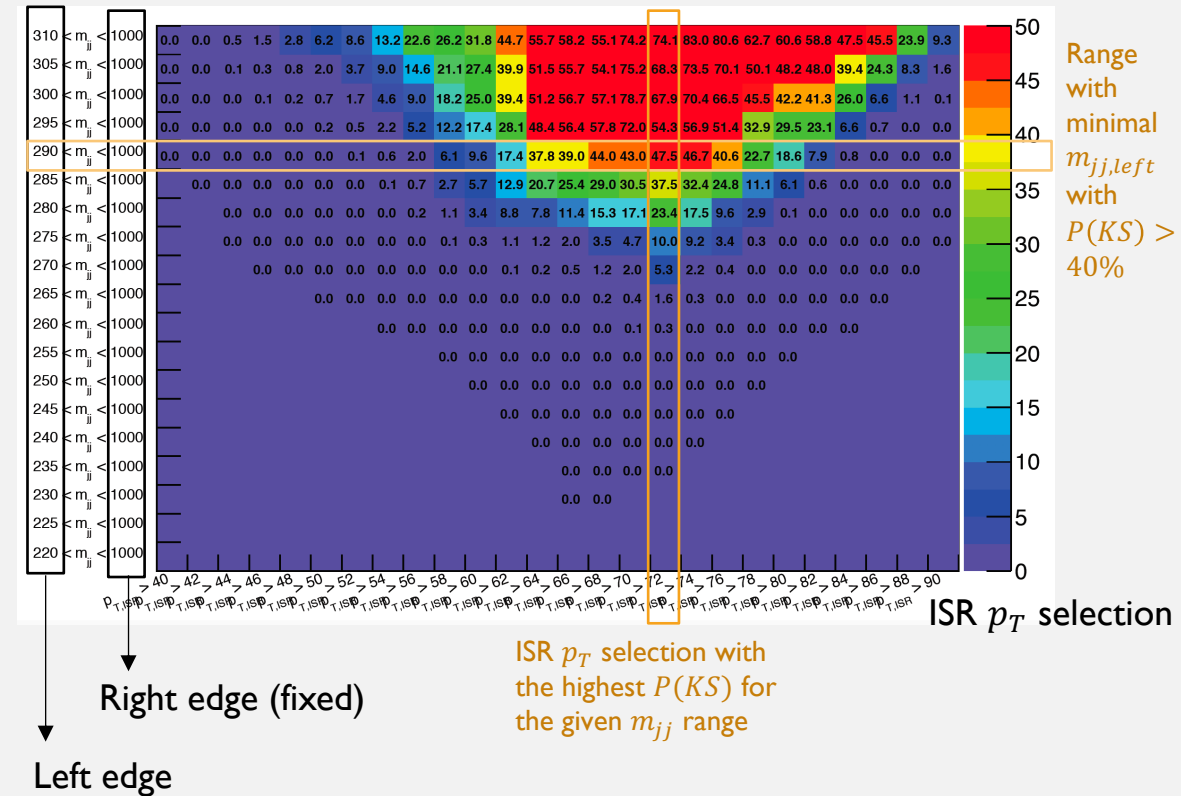
- Select the fit range and ISR p_T cut using MSR by looking at the Kolmogorov-Smirnov test output in (ISR p_T cut)x(fit range) space using N-parameter alternative and nominal function given at Step-2 (N=4 in the beginning)
 - Select the lowest low edge ($m_{jj,left}$) of the fit range with $P(KS) > 40\%$
 - Select the ISR p_T cut with the highest $P(KS)$ (for given $m_{jj,left}$)

- **Step-2**

- Select the fit function using Fisher test + χ^2 criteria.
- Repeat Step-1 and Step-2 until fit range, ISR p_T cut and fit function “converge”.

- **Step-3**

- Use Fisher test to select best fit function in SR



Step-1. Kolmogorov-Smirnov test in m_{jj} -range – ISR p_T selection space

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION

- **Recursive part:**

- **Step-1**

- Select the fit range and ISR p_T cut using MSR by looking at the Kolmogorov-Smirnov test output in (ISR p_T cut)x(fit range) space using N-parameter alternative and nominal function given at Step-2 (N=4 in the beginning)
 - Select the lowest low edge ($m_{jj,left}$) of the fit range with $P(KS) > 40\%$
 - Select the ISR p_T cut with the highest $P(KS)$ (for given $m_{jj,left}$)

- **Step-2**

- Select the fit function using Fisher test + χ^2 criteria.
- Repeat Step-1 and Step-2 until fit range, ISR p_T cut and fit function “converge”.

- **Step-3**

- Use Fisher test to select best fit function in SR

# of par-s	Nominal	Alternative
3	1764	1764
4	11.3	39.6
5	8.88	9.41
6	8.80	9.66
7	8.72	10.03

1) χ^2 for all of the functions for $290 < m_{jj} < 1000$ GeV mediator mass range and ISR $p_T > 72$ GeV.

2) Fisher test results. The best option: Nominal 4-parametric function.

Nominal			Alternative		
1 st -2 nd function	F-value	CL	1 st -2 nd function	F-value	CL
3Par-4Par	2329	0.000	3Par-4Par	653	0.000
4Par-5Par	3.8	0.072	4Par-5Par	44.9	0.000
5Par-6Par	0.11	0.740	5Par-6Par	-0.34	1.000
6Par-7Par	0.10	0.753	6Par-7Par	-0.43	1.000

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION

- **Recursive part:**

- **Step-1**

- Select the fit range and ISR p_T cut using MSR by looking at the Kolmogorov-Smirnov test output in (ISR p_T cut)x(fit range) space using N-parameter alternative and nominal function given at Step-2 (N=4 in the beginning)

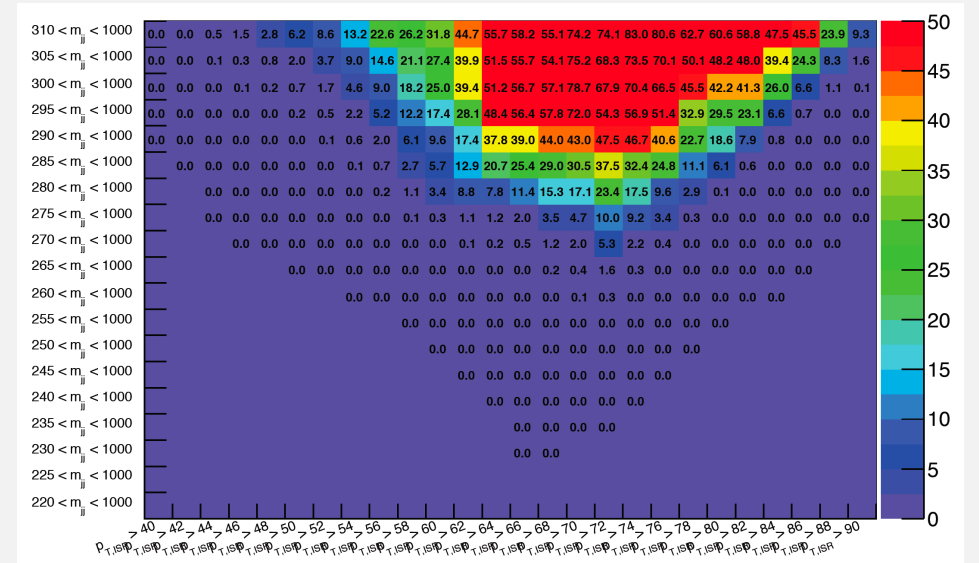
- Select the lowest low edge ($m_{jj,left}$) of the fit range with $P(KS) > 40\%$
- Select the ISR p_T cut with the highest $P(KS)$ (for given $m_{jj,left}$)

- **Step-2**

- Select the fit function using Fisher test + χ^2 criteria.
- Repeat Step-1 and Step-2 until fit range, ISR p_T cut and fit function “converge”.

- **Step-3**

- Use Fisher test to select best fit function in SR



Nominal			Alternative		
1st-2nd function	F-value	CL	1st-2nd function	F-value	CL
3Par-4Par	2329	0.000	3Par-4Par	653	0.000
4Par-5Par	3.8	0.072	4Par-5Par	44.9	0.000
5Par-6Par	0.11	0.740	5Par-6Par	-0.34	1.000
6Par-7Par	0.10	0.753	6Par-7Par	-0.43	1.000

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION

- **Recursive part:**

- **Step-1**

- Select the fit range and ISR p_T cut using MSR by looking at the Kolmogorov-Smirnov test output in (ISR p_T cut)x(fit range) space using N-parameter alternative and nominal function given at Step-2 (N=4 in the beginning)
 - Select the lowest low edge ($m_{jj,left}$) of the fit range with $P(KS) > 40\%$
 - Select the ISR p_T cut with the highest $P(KS)$ (for given $m_{jj,left}$)

- **Step-2**

- Select the fit function using Fisher test + χ^2 criteria.
- Repeat Step-1 and Step-2 until fit range, ISR p_T cut and fit function “converge”.

- **Step-3**

- Use Fisher test to select best fit function in SR

# of par-s	Nominal	Alternative
3	8176	8176
4	13.9	53.1
5	10.06	10.03
6	10.08	10.11
7	10.00	10.34

1) χ^2 for all of the functions for $270 < m_{jj} < 1000$ GeV mediator mass range and ISR $p_T > 70$ GeV.

2) Fisher test results. The best option: Alternative 5-parametric function.

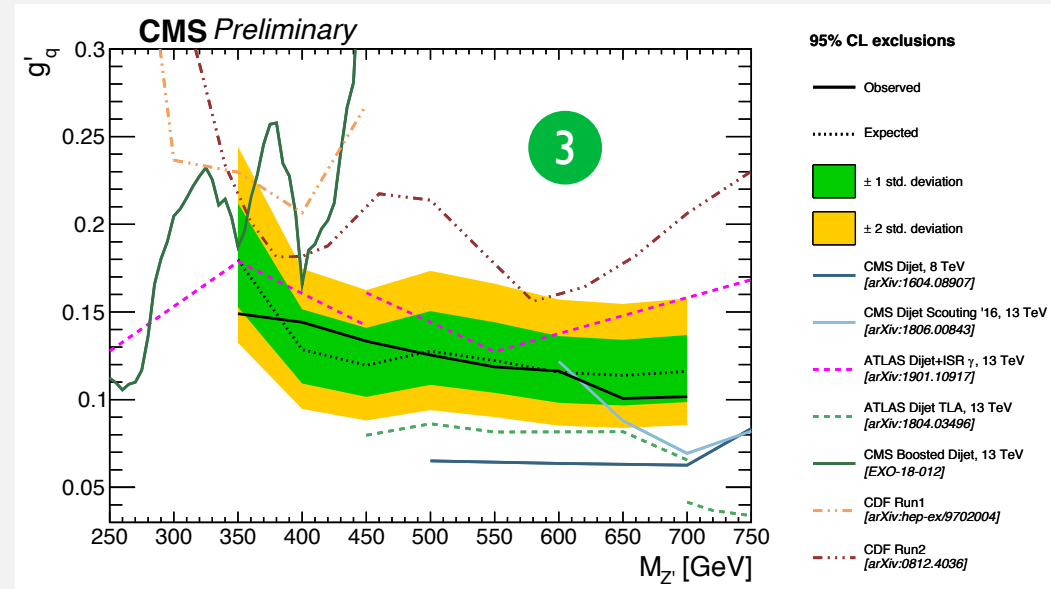
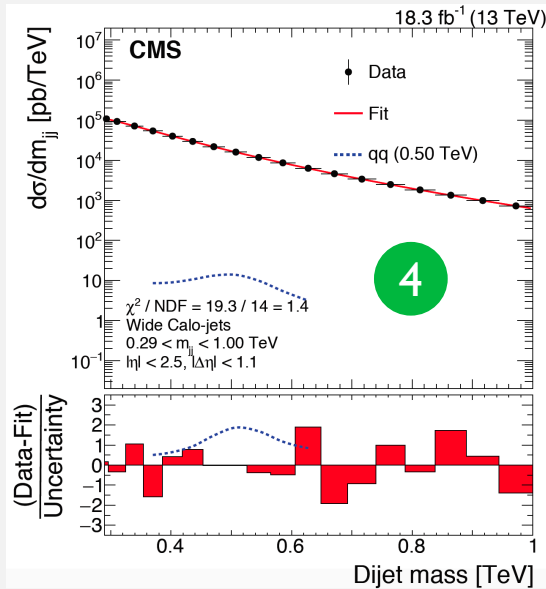
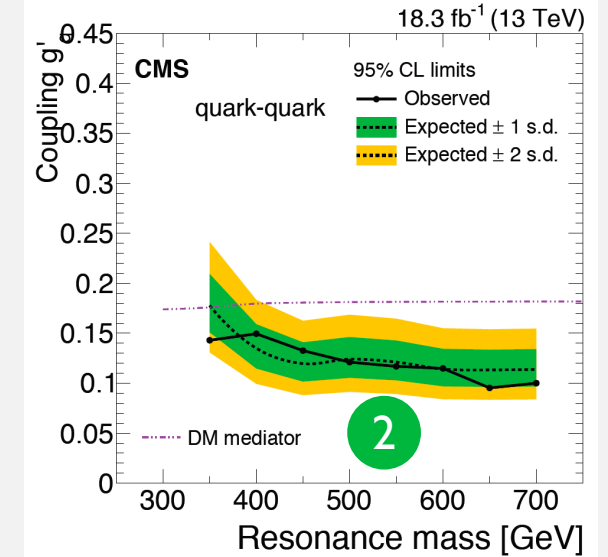
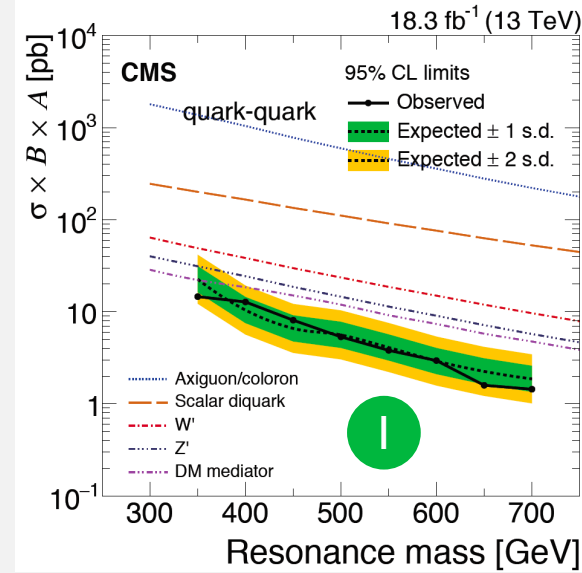
Nominal			Alternative		
1 st -2 nd function	F-value	CL	1 st -2 nd function	F-value	CL
3Par-4Par	8828	0.000	3Par-4Par	2293	0.000
4Par-5Par	5.3	0.037	4Par-5Par	60.2	0.000
5Par-6Par	-0.03	1.000	5Par-6Par	-0.10	1.000
6Par-7Par	0.10	0.761	6Par-7Par	-0.26	1.000

SYSTEMATIC UNCERTAINTIES

Systematic Uncertainty Source	Uncertainty
Jet Energy Resolution	10% of RECO resolution
Jet Energy Scale	$\pm 2\%$ shift of m_{jj}
Luminosity	$\pm 2.5\%$
Parton shower	$\pm 6.0\%$

RESULTS

- There is no significant excess in 2016 dijet in three-jet final state data
- Plot 1 shows the upper limit on the $\sigma \times BR \times A$
- Plot 2 shows the upper limit on the coupling (calculated using formulas on slides 11-13)
- Plot 3 shows upper limit on the coupling in comparison with other analysis
- Plot 4 shows binned fit



g'_q : coupling of Z' to qq of leptophobic Z' that couples only to qq ;
 g_q : coupling of Z' to qq of leptophobic Z' that can couple to qq and DM;
 Assuming equal x section:

$$g'_q = g_q \sqrt{\frac{1}{1 + \frac{1}{3N_f g_q^2}}}, \text{ where}$$

$$N_f = 5 + \left(1 - \frac{4m_{top}^2}{M_{med}^2}\right)^{\frac{1}{2}} \left(1 + \frac{2m_{top}^2}{M_{med}^2}\right).$$

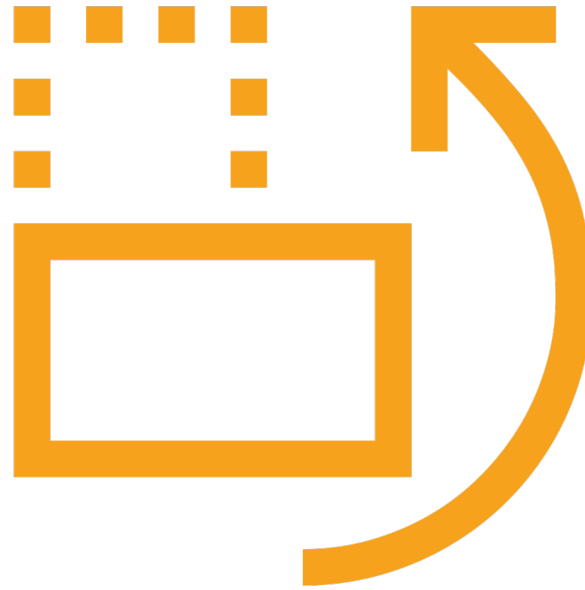


CONCLUSIONS

- Search for a dijet resonance in three-jet final state is finished
- Long road from LI trigger decision to the final result
- A lot of work in between:
 - Dijet background shape studies
 - Automatic procedure development
 - Background fit function searches
 - Wide jets cone radius searches
 - Trigger studies
 - Relation between couplings dependence on mediator mass
 - EE-LI pre-firing issue study
 - Etc.
- The limit on the coupling set to 0.1-0.15 for mass range of 350-700 GeV

FUTURE PLANS

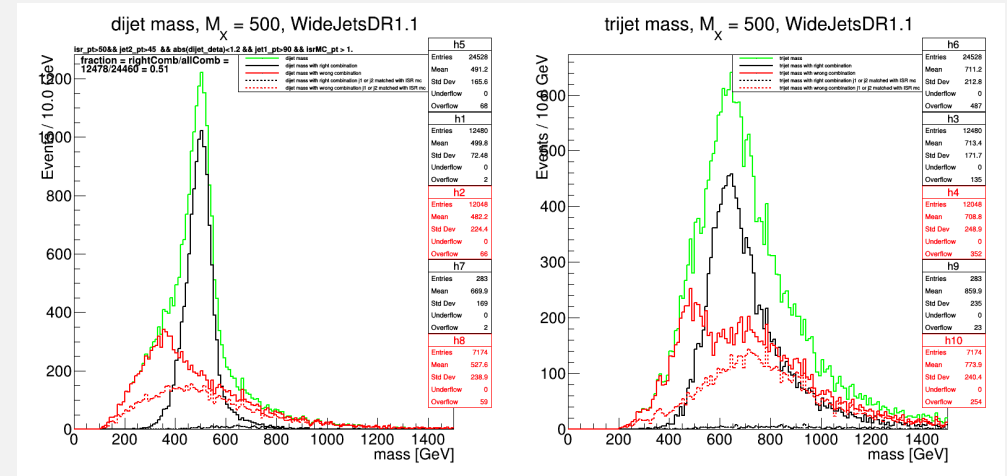
- Cannot perform the analysis for Run 2:
 - LI HTT240 is off after 2016. The only trigger available for whole Run 2 is LI HTT320
 - Cannot increase much ISR p_T cut in order to increase trigger efficiency (kinematic effect)
 - Possible use of LI HTT320 with CR method to determine the background (new analysis!).
 - ISR p_T – the first leading jet, look at the $m_{jj} < 300$ GeV (new analysis!)



BACK UP

SIGNAL

- Samples are produced using the inclusive process $q\bar{q}/qg \rightarrow Z'(q\bar{q}) + q/g$ for 300,400,500,600,800,1000 GeV mediator mass.
- Z' is the Breit-Wigner narrow resonance with $\sim 10\% \Gamma/M$ that decays to quark-antiquark pair, which create dijet. The dijet invariant mass has the shape described by relativistic Breit-Wigner theory and peaks at the mediator mass.
- However, in case of three-jet final state, some mismatching of the 2 leading jets (which we select as a dijet) with the 2 quarks from resonance decay is present.
- There is no trivial way to improve matching or tag somehow FSR (ISR) jet. Studies about possible FSR tagger can be found in chapter 5 of AN (http://cms.cern.ch/iCMS/jsp/db_notes/notelInfo.jsp?cmsnotid=CMS%20AN-2018/269).



Signal shapes for $M = 500$ GeV. Dijet (left) and trijet (right) invariant mass. Black line - good matching ($DR < 0.4$) and red line - bad matching. The green line is the sum of the two categories. The dashed red line represent events with bad matching, with the additional q/g matching of one of the two leading jets.

SIGNAL SHAPE STUDIES

- “Bad” shape: second peak at lower mass than the mass of the resonance

- Reasons:

- ❖ Events without MC ISR parton:

- Wrong combinations: dijet system contains one of the FSR jets

- $dijet = jet(q_1) + jet(FSR)$

- $ISR\ jet = jet(q_2)$

- Right combinations: FSR jet interpreted as ISR jet or no jet

- $dijet = jet(q_1) + jet(q_2)$

- $ISR\ jet = jet(FSR)\ or\ 0$

- ❖ Events with MC ISR:

- Wrong combinations: FSR (dashed red line) or ISR (remaining events) jets are included to the dijet system

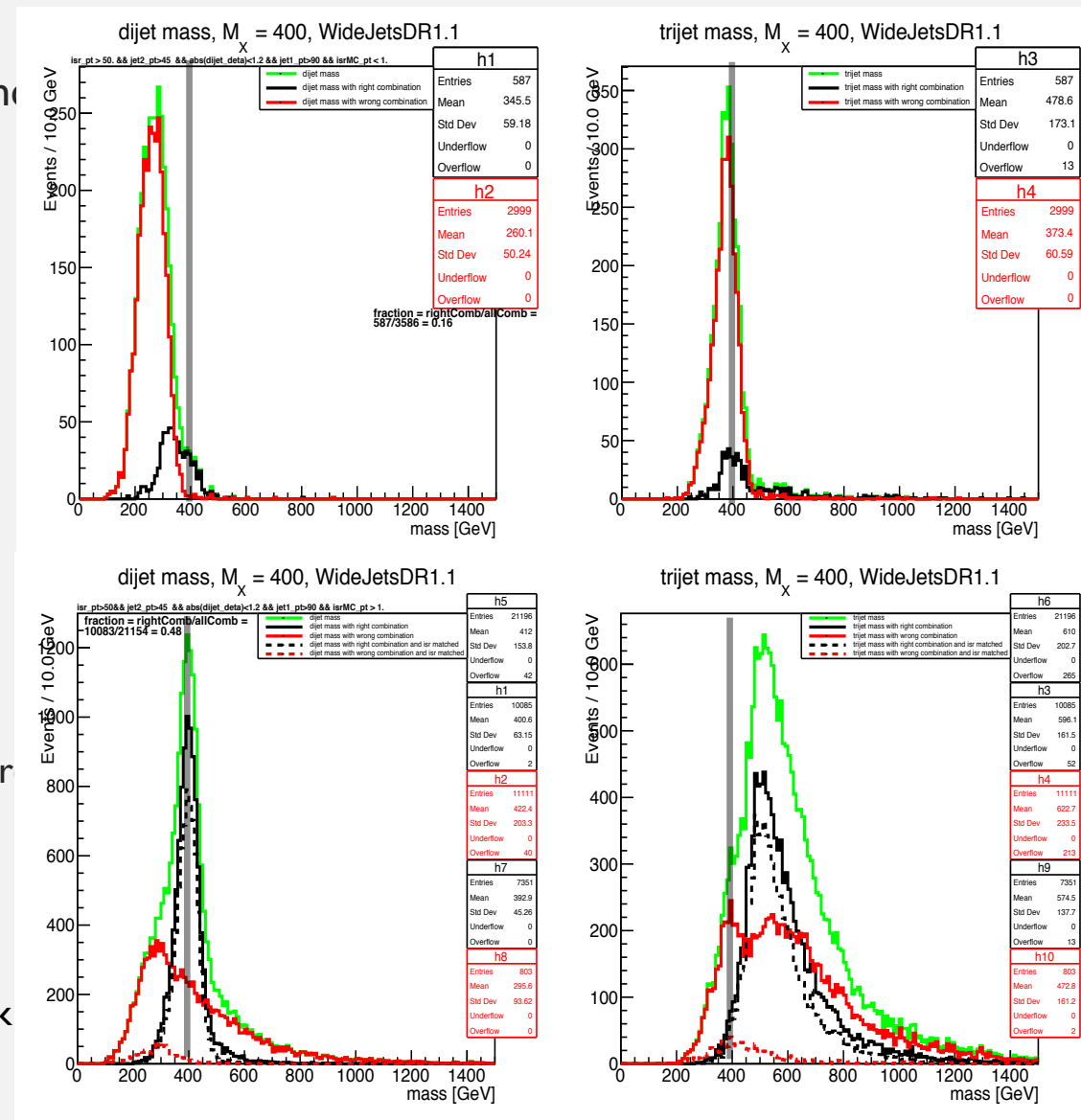
- $dijet = jet(q_1) + jet(FSR)\ or\ jet(q_1) + jet(ISR)$

- $ISR\ jet = jet(q_2)$

- Right combinations: ISR jet reconstructed in a right way in most cases (black dashed line)

- $dijet = jet(q_1) + jet(q_2)$

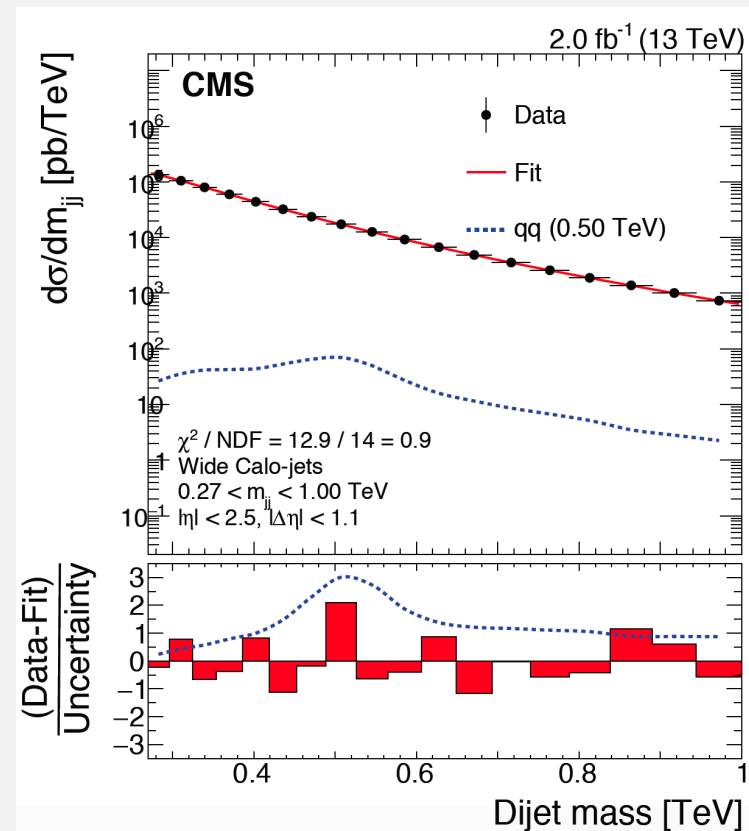
- $ISR\ jet = jet(ISR)\ or\ jet(FSR)$



BACKGROUND ESTIMATION

- To fit the dijet mass spectrum we use analytic smooth function (method used by resolved dijet analysis). However, the functions used by resolved dijet analysis are not capable to perform good fit in our case.
- We defined 2 families of functions that are able to fit dijet mass spectrum. Difference in shape with the standard dijet analysis is caused by kinematic cut (for instance p_T cut) or small trigger inefficiency.
- To define a range of the fit and ISR p_T cut we use procedure that is different from what standard dijet analysis has (see slides 10-15)

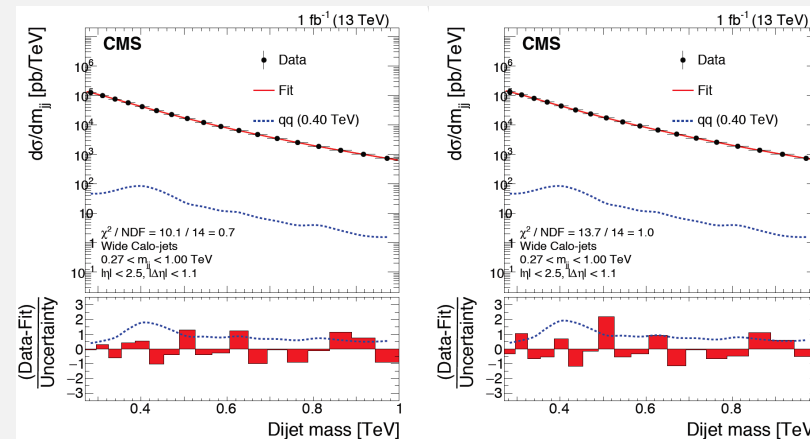
# of par-s	Nominal family	Alternative family
3	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1)}}{x^{p_1}}$	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1)}}{x^{p_1}}$
4	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1) + p_3 x}}{x^{p_1}}$	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1)}}{x^{p_1 + p_3 \log(x)}}$
5	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1) + p_3 x + p_4 x^2}}{x^{p_1}}$	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1)}}{x^{p_1 + p_3 \log(x) + p_4 \log(x)^2}}$
6	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1) + p_3 x + p_4 x^2 + p_5 x^3}}{x^{p_1}}$	$\frac{d\sigma}{dm_{jj}} = \frac{p_0 e^{\log(p_2 x - 1)}}{x^{p_1 + p_3 \log(x) + p_4 \log(x)^2 + p_5 \log(x)^3}}$



Background fit using 10% of the data performed in $270 < m_{jj} < 1000$ GeV mass range.

RUN H CHECK

- Two-stage check of the Run H data:
 - Check the GOF including and excluding Run H data:
 - Should have similar $\chi^2/ndof$
 - We use the ScoutingCaloHT data obtained by DST_HT250_v trigger path (consists of events mainly selected by $H_T > 250 \text{ GeV}$ cut)
 - Check the trigger efficiency with and without Run H data:
 - We use ScoutingCaloCommissioning data samples which allows us to check the trigger efficiency of HLT_CaloScouting_250
- Conclusion:** we can use Run H and the data inefficiency that influence standard dijet analysis does not have impact on our samples. Integrated luminosity increased by $\sim 3 \text{ fb}^{-1}$.



Fit results for the 5-parameter nominal function for 10% of the data obtained excluding (left) and including (right) RunH.

Trigger efficiency.

Right column: data

w/o runH

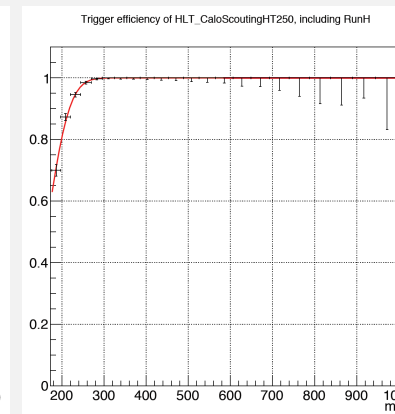
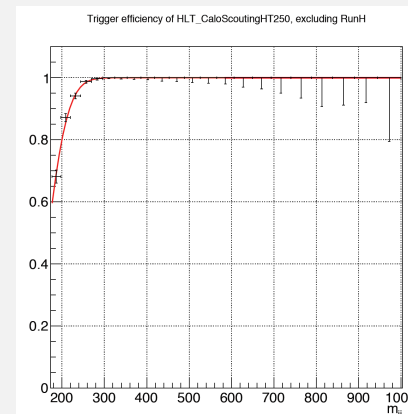
Middle column: only

Run H data

Right column: data

with Run H

Bin	Excluding RunH	Only RunH	Including RunH
197 – 220	87.19 %	87.68 %	87.30 %
220 – 244	94.17 %	96.12 %	94.59 %
244 – 270	98.77 %	97.45 %	98.49 %
270 – 296	99.68 %	100.00 %	99.74 %
296 – 325	100.00 %	100.00 %	100.00 %
325 – 354	100.00 %	100.00 %	100.00 %
354 – 386	100.00 %	100.00 %	100.00 %
386 – 419	100.00 %	100.00 %	100.00 %
419 – 453	100.00 %	100.00 %	100.00 %
453 – 489	100.00 %	100.00 %	100.00 %



Trigger efficiency and fit excluding RunH (left) and including RunH (right) data

EE LI PRE-FIRING ISSUE

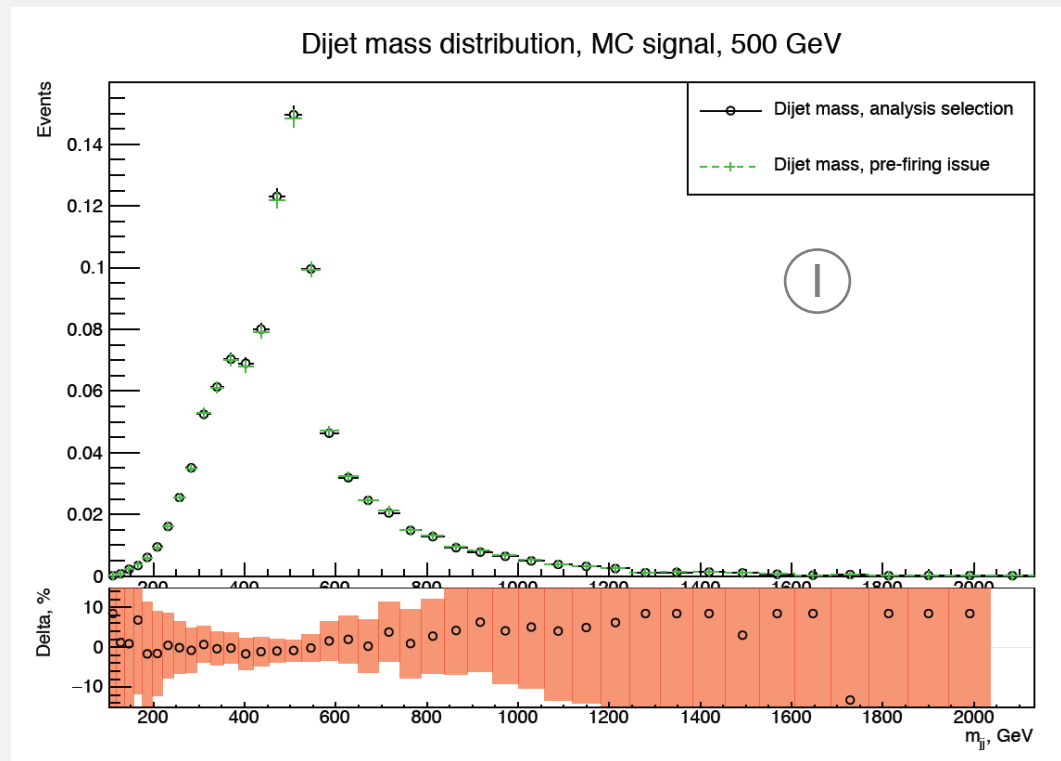
- EE LI pre-firing issue might impact final kinematic distribution even if analysis does not use objects that can fire EE LI triggers.
- Plots I) show the distribution of the signal for 500 GeV of the resonance mass
- From the table below we can conclude that EE LI pre-firing has **no** impact on the final dijet mass distribution and we should apply corrections.
- We did it using the following maps https://lathomas.web.cern.ch/lathomas/TSGStuff/LIPrefiring/PrefiringMaps_2016and2017/ and recipe <https://twiki.cern.ch/twiki/bin/view/CMS/LIECALPrefiringWeightRecipe>.
- The main idea is to implement the weights on the final kinematic distribution, which depends on η and p_T of jets:

$$\omega = 1 - P(\text{prefiring}) = \prod_{i=\text{jets,photons}} (1 - \epsilon_i^{\text{pref}}(\eta, p_T))$$

using efficiencies from the map above.

Mass GeV	Nominal	Pre-firing	Difference
300	11504.0	11376.7	1.11%
400	15235.0	15045.9	1.24%
500	19421.0	19158.4	1.35%
600	21998.0	21704.8	1.33%
800	27725.0	27417.7	1.11%
1000	31260.0	30979.3	0.90%

Number of events for the **Nominal** analysis selection, **Pre-firing** selection (removing events with $p_T > 100$ GeV and $2.25 < \eta < 3.0$) and difference in %. If the last one is $\cong 1\%$ or less, than analysis is not influenced by pre-firing issue.



Distributions of the signal using analysis selection (black line) and signal excluding events using jet selection of $p_T > 100$ GeV and $2.25 < \eta < 3.0$ (green).

RELATION BETWEEN DM COUPLING g' AND g .

INTRODUCTION

- Partial widths for a vector mediator Z' that decays to 2 DM particles and quark-antiquark pair (according to <https://arxiv.org/pdf/1603.04156.pdf>) are:

$$\Gamma_{vector}^{\chi\chi} = \frac{g_{DM}^2 M_{med}}{12\pi} (1 - 4z_{DM})^{1/2} (1 + 2z_{DM}); z_{DM} = \frac{m_{DM}^2}{M_{med}^2}, Z' \rightarrow \chi\bar{\chi}$$

$$\Gamma_{vector}^{qq} = \frac{g_q^2 M_{med}}{4\pi} (1 - 4z_q)^{1/2} (1 + 2z_q); z_q = \frac{m_q^2}{M_{med}^2}, Z' \rightarrow q\bar{q}$$

- In our case, we consider $m_{DM} = 1 \text{ GeV}$ and $g_{DM} = 1$, thus $(1 - 4z_{DM})^{1/2} (1 + 2z_{DM}) \approx 1$ and $\Gamma_{vector}^{\chi\chi} = \frac{1}{12\pi} M_{med}$.
- Furthermore, if $m_q \ll M_{med}$ (as in EXO-16-056), then partial width of Z' going to a quark-antiquark pair is $\Gamma_{vector}^{qq} = \frac{g_q^2 M_{med}}{4\pi}$.
- The total cross section σ for particle production via a Breit-Wigner resonance is approximately proportional to product of partial width in initial and final states divided by full width, and for the dijet process (AN-17-013, Appendix – “Relation of limits on DM Coupling”):

$$\sigma \sim \frac{\Gamma_i \Gamma_f}{\Gamma} = \frac{\Gamma(q\bar{q})^2}{\Gamma}$$

- The partial width of $\Gamma(q\bar{q})$ decaying to any $q\bar{q}$ pair for $m_q \ll M_{med}$ is

$$\Gamma(q\bar{q}) = \sum_{qq} \Gamma_{vector}^{qq} = 6 \cdot \frac{g_q^2}{4\pi} M_{med} = \frac{18g_q^2}{12\pi} M_{med}$$

- Partial width into DM particles (considering only $Z' \rightarrow \chi\bar{\chi}$ decay, $g_{DM} = 1$):

$$\Gamma(\chi\bar{\chi}) = \frac{1}{12\pi} M_{med}$$

- Full width for Z' that couples to $q\bar{q}$ and $\chi\bar{\chi}$:

$$\Gamma = (18g_q^2 + 1) \frac{M_{med}}{12\pi}$$

- Now consider the case when Z' couples only to quarks with the total width Γ' (the analysis is sensitive only to this type of coupling):

$$\Gamma' = \frac{\Gamma_i \Gamma_f}{\Gamma} = \frac{(\Gamma'(q\bar{q}))^2}{\Gamma'(q\bar{q})} = \Gamma'(q\bar{q}) = \frac{18(g'_q)^2}{12\pi} M_{med}$$

- Setting the cross section of the two resonance processes equal to each other, we get

$$\frac{\Gamma(q\bar{q})^2}{\Gamma(q\bar{q}) + \Gamma(\chi\bar{\chi})} = \Gamma'(q\bar{q}) \rightarrow \frac{(18g_q^2)^2}{18g_q^2 + 1} = 18(g'_q)^2 \rightarrow g'_q = g_q \sqrt{\frac{1}{1 + 1/(18g_q^2)}}$$

Eq. 1

RELATION BETWEEN DM COUPLING g' AND g . DEPENDENCE OF THE COUPLING ON THE MEDIATOR MASS

- In a case of dijet analysis in three-jet final state, relation is different, since we are searching for the new particle in the range of 300-600 GeV, than $m_q \ll M_{med}$ is not fulfilled for top quark.

- We start from

$$\Gamma(q\bar{q}) = \sum_{qq} \frac{g_q^2 M_{med}}{4\pi} (1 - 4z_q)^{1/2} (1 + 2z_q) = \frac{g_q^2 M_{med}}{4\pi} \left(5 + (1 - 4z_{top})^{1/2} (1 + 2z_{top}) \right) = \frac{g_q^2 M_{med}}{4\pi} N_f$$

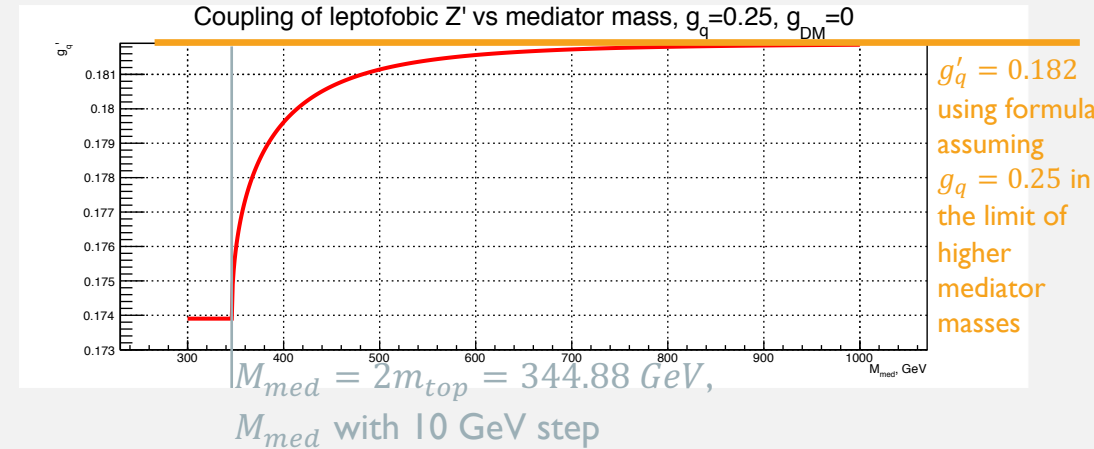
where $z_{top} = \frac{m_{top}^2}{M_{med}^2}$ and $N_f = 5 + (1 - 4z_{top})^{1/2} (1 + 2z_{top})$ is the effective number of quarks. Rewriting Eq. 1 from previous slide we get

$$\frac{\Gamma(q\bar{q})^2}{\Gamma(q\bar{q}) + \Gamma(\chi\bar{\chi})} = \Gamma'(q\bar{q}) \rightarrow \frac{\left(\frac{g_q^2 M_{med}}{4\pi} N_f \right)^2}{\frac{g_q^2 M_{med}}{4\pi} N_f + \frac{1}{12\pi} M_{med}} = \frac{g_q'^2 M_{med}}{4\pi} N_f \rightarrow g'_q = g_q \sqrt{\frac{1}{1 + \frac{1}{3N_f g_q^2}}}$$

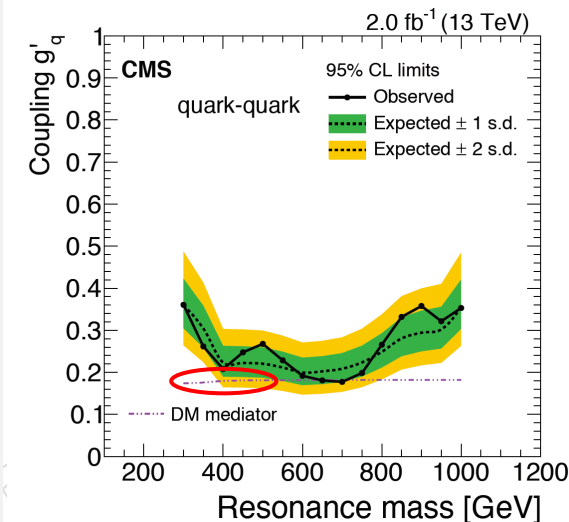
- Also, to find the upper limit on the coupling g'_q , we need to use this formula:

$$(g'_q)_{95\% CL} = (g'_q)_{MC} \sqrt{\frac{(\sigma \times BR \times A)_{95\% CL}}{(\sigma \times BR \times A)_{MC} \frac{A_{reco}}{A_{MC}}}}$$

Where $(g'_q)_{MC}$ is calculated using $g_q = 0.25$ for M_{med} from 300 to 600 GeV.



Dependence of g'_q on mediator mass assuming $g_q = 0.25$ for $m_{DM} = 1$ GeV.



The upper limit on the g'_q coupling. In the red circle you can see small dependence of g'_q coupling obtained theoretically.

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION. STEP-1

- We define (for now) the **control region** within $0.9 < |\Delta\eta| < 1.1$. This region is not orthogonal to SR, but it has:
 - Lower purity
 - Lower sensitivity
 - Distribution is shifted towards larger values
- We use three-step automatic iterative procedure to define ISR p_T cut, fit range and the function we'll use to fit the data:
 - Step-1: select ISR p_T cut and fit range using control region data, looking at the χ^2 probability obtained with N-parameter function (from one of 2 families) in step-2 (at first iteration N=4). Criteria to chose mass range and ISR p_T cut:
 - Select the mass range with the lowest left bound having $Prob(\chi^2, ndof) > 33\%$ among all ISR p_T cuts
 - Select ISR p_T cut for which we have the largest $Prob(\chi^2, ndof)$.
 - The result here illustrated is the following: the lowest left bound is **270-1000 GeV**, obtained with the nominal function (4-parametric). The highest probability we get by selecting the events with $ISR p_T > 70 GeV$.
 - Step-2 ... see the next slide

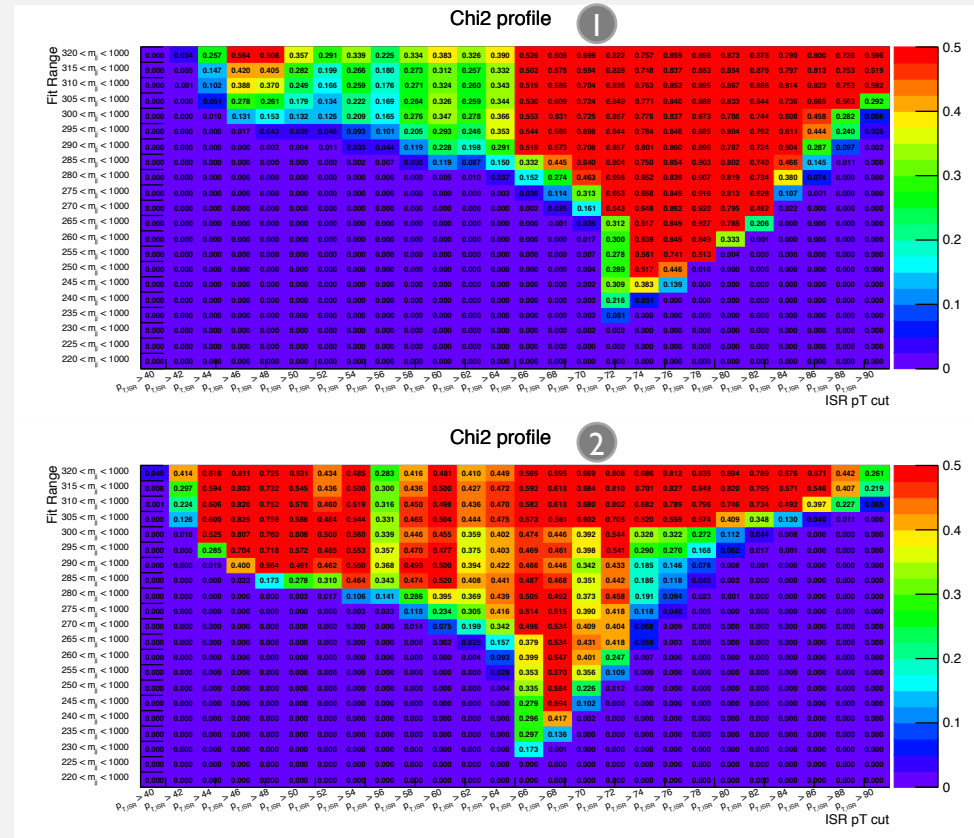


Illustration of step-1 of the iterative automatic procedure.

- Probability of χ^2 using 4-parametric nominal function.
- Probability of χ^2 using 4-parametric alternative function.

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION. STEP-2

- We use three-step automatic iterative procedure to define ISR p_T cut, fit range and the function we'll use to fit the data:
 - Step-2: using the mass range and ISR p_T cut found in step-1, perform a Fisher test to choose the best fitting function (using CR).
 - We choose here 4 parametric nominal function since, since the pair 4par-5par has $CL > 0.05$ and 1st function of this pair has least parameters among all functions.
 - * Step-1 and step-2 are invoked recursively until **ISR p_T cut, fit range and fit function** will be found.
 - Step-3 ... see the next slide

❖ Reminder. Fisher test.

- Fisher test is using to define how much effective parameters we need to perform a fit. First of all we need to define F-value of two given **models** f_1 and f_2 :

$$F_{21} = \frac{\frac{RSS_1 - RSS_2}{n_2 - n_1}}{\frac{RSS_2}{N - n_2}}$$

- where $RSS_i = \sum_{bins} (data_{bin} - fit_{bin})^2$ - residual sum of squares of model i , n_i is the number of parameters of the model i , N is the number of bins.
- Under the null hypotheses that **model 2** does not improve significantly the fit than **model 1**, the F-value is distributed by F-distribution with $(n_2 - n_1, N - n_2)$ degrees of freedom. The observed confidence level defined as:

$$CL_{21} = 1 - \int_{-\infty}^{F_{21}} F(f; n_2 - n_1, N - n_2) df$$

- The null hypotheses is rejected if $CL_{21} < \alpha$, where α assumed to be 0.05.

# of par-s	Nominal	Alternative
3	1764	1764
4	11.3	39.6
5	8.88	9.41
6	8.80	9.66
7	8.72	10.03

1) χ^2 for all of the functions for $270 < m_{jj} < 1000$ GeV mediator mass range and ISR $p_T > 70$ GeV.

2) Fisher test results. The best option: Nominal 4-parametric function.

Nominal			Alternative		
1 st -2 nd function	F-value	CL	1 st -2 nd function	F-value	CL
3Par-4Par	2329	0.000	3Par-4Par	653	0.000
4Par-5Par	3.8	0.072	4Par-5Par	44.9	0.000
5Par-6Par	0.11	0.740	5Par-6Par	-0.34	1.000
6Par-7Par	0.10	0.753	6Par-7Par	-0.43	1.000

AUTOMATIC PROCEDURE TO DEFINE THE EVENT SELECTION AND FUNCTION. STEP-3

- We use three-step automatic iterative procedure to define ISR p_T cut, fit range and the function we'll use to fit the data:
 - Step-3: select the best fit function in the SR using Fisher test
 - We chose alternative 5 parametric function as a fit function
- ! We might add additional fit parameter if CR contains few times less statistics than SR.

# of par-s	Nominal	Alternative
3	8176	8176
4	13.9	53.1
5	10.06	10.03
6	10.08	10.11
7	10.00	10.34

1) χ^2 for all of the functions for $270 < m_{jj} < 1000$ GeV mediator mass range and ISR $p_T > 70$ GeV.

2) Fisher test results. The best option: Alternative 5-parametric function.

❖ Reminder. Fisher test.

- Fisher test is using to define how much effective parameters we need to perform a fit. First of all we need to define F-value of two given models f_1 and f_2 :

$$F_{21} = \frac{\frac{RSS_1 - RSS_2}{n_2 - n_1}}{\frac{RSS_2}{N - n_2}}$$

- where $RSS_i = \sum_{bins} (data_{bin} - fit_{bin})^2$ - residual sum of squares of model i , n_i is the number of parameters of the model i , N is the number of bins.
- Under the null hypotheses that **model 2** does not improve significantly the fit than **model 1**, the F-value is distributed by F-distribution with $(n_2 - n_1, N - n_2)$ degrees of freedom. The observed confidence level defined as:

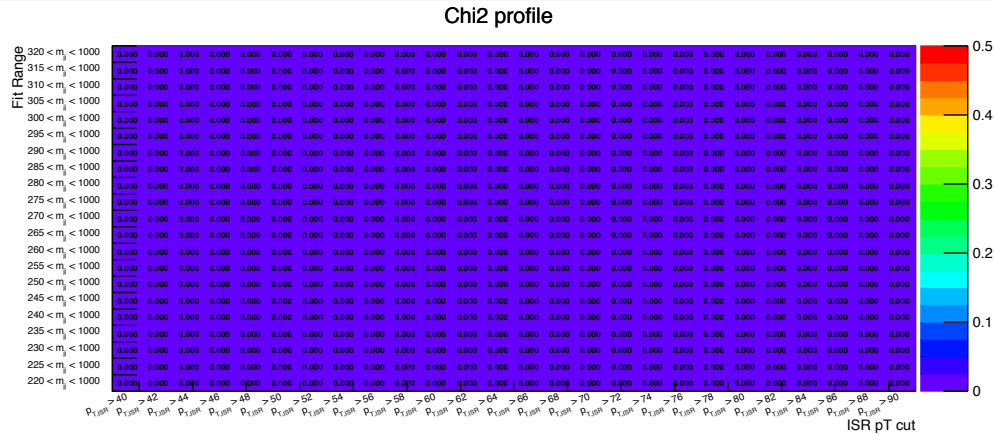
$$CL_{21} = 1 - \int_{-\infty}^{F_{21}} F(f; n_2 - n_1, N - n_2) df$$

- The null hypotheses is rejected if $CL_{21} < \alpha$, where α assumed to be 0.05.

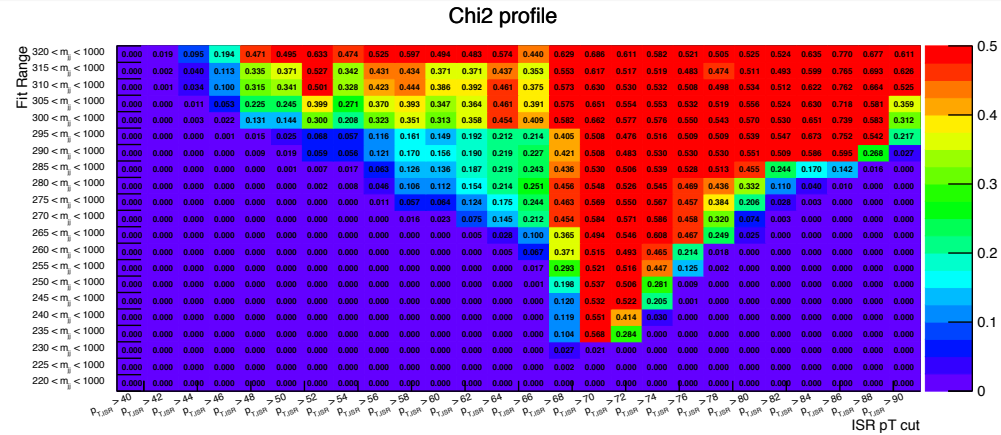
Nominal			Alternative		
1 st -2 nd function	F-value	CL	1 st -2 nd function	F-value	CL
3Par-4Par	8828	0.000	3Par-4Par	2293	0.000
4Par-5Par	5.3	0.037	4Par-5Par	60.2	0.000
5Par-6Par	-0.03	1.000	5Par-6Par	-0.10	1.000
6Par-7Par	0.10	0.761	6Par-7Par	-0.26	1.000

SEARCH FOR A NEW CONTROL REGION. RESULTS OF CHI-SQUARE SCAN FOR THE 1ST CR (FLIPPED η)

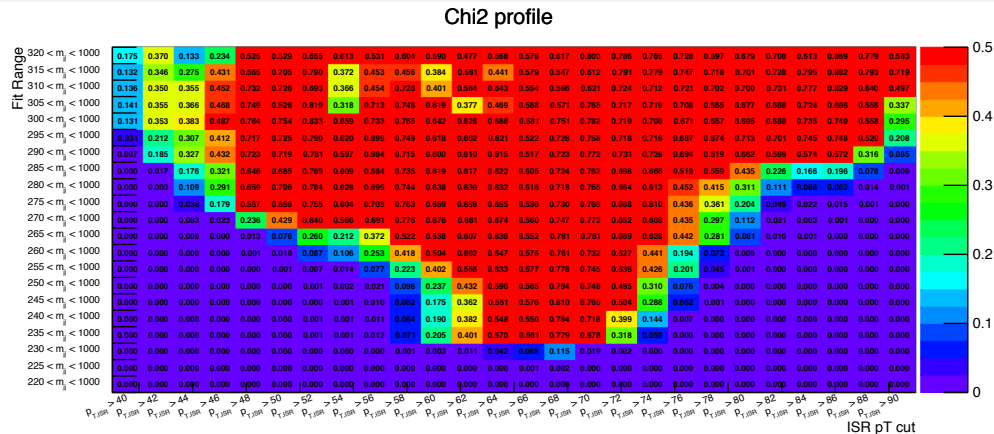
Ait 3



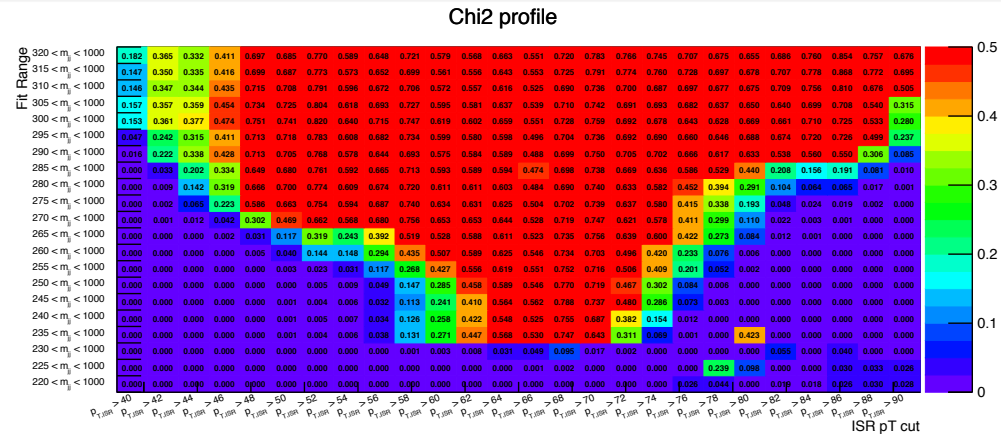
Ait 4



Ait 5



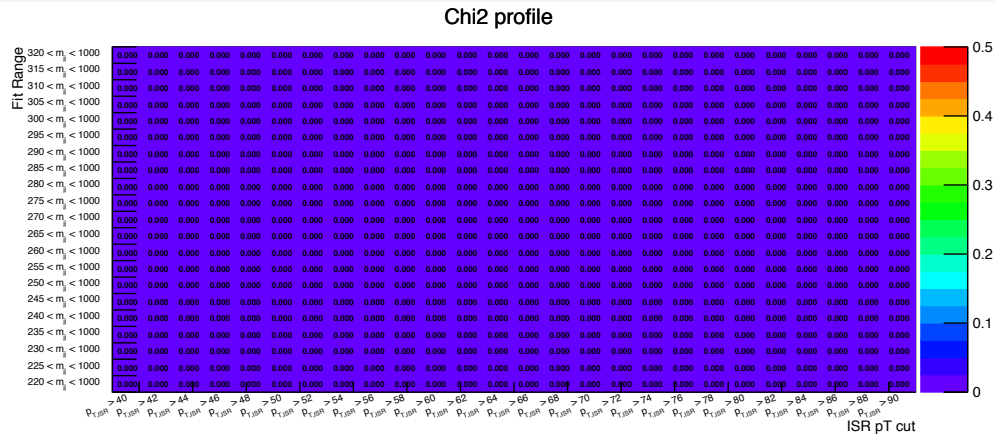
Ait 6



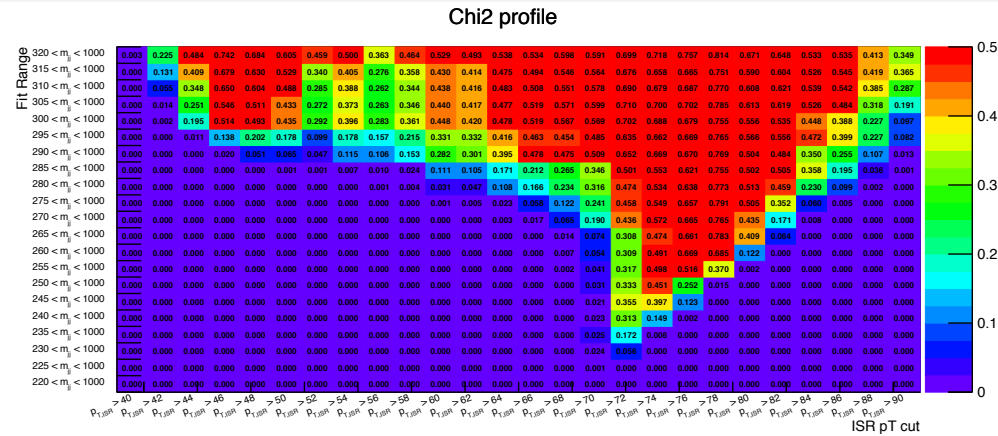
Probability of χ^2 scan for alternative functions for 1st CR (flipped η).

SEARCH FOR A NEW CONTROL REGION. RESULTS OF CHI-SQUARE SCAN FOR THE SR

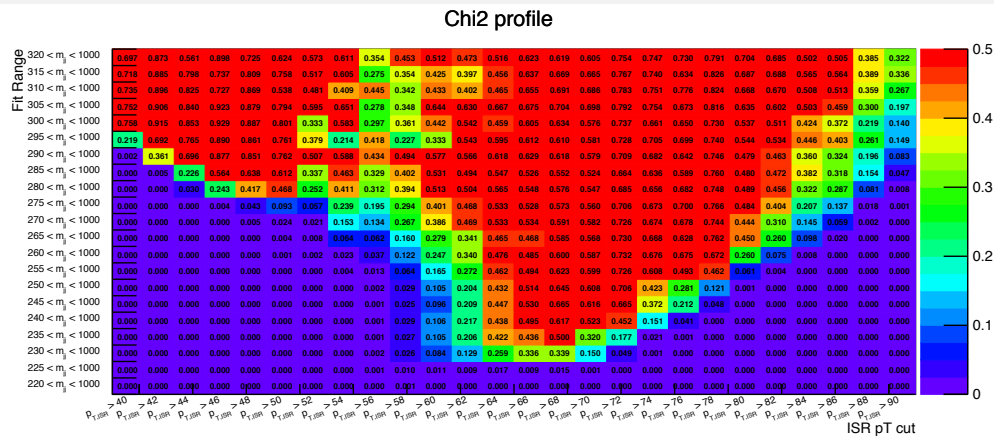
Ait 3



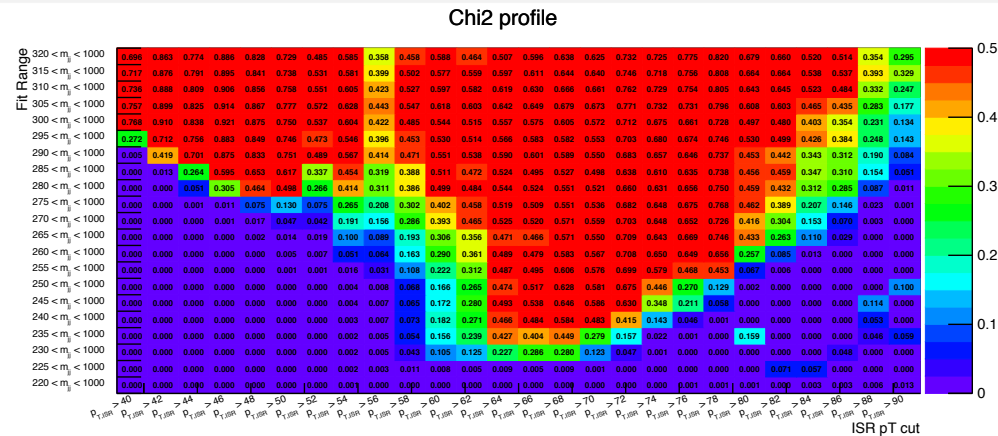
Ait 4



Ait 5



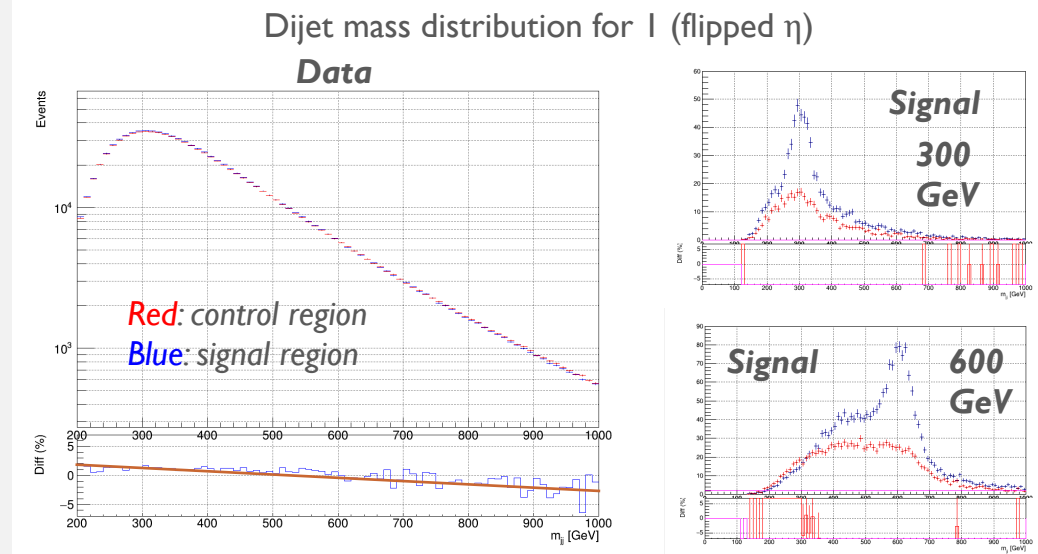
Ait 6



Probability of χ^2 scan for alternative functions for SR. Comparison with the previous slide shows the difference in the regions of good fit in SR and CR. It is caused by difference between shapes of SR and CR, which I mentioned before.

SEARCH FOR A NEW CONTROL REGION

- Until now we have several ideas to implement in a very short time to find the best CR, which we could use in the automatic procedure:
 1. Flip η of the second jet.
 - However, the dijet mass distribution in SR is more steep than for CR.
 2. Define η of the second jet as a function of the η of the first jet and $\Delta\eta$ skimmed using gaussian.
 - With this we try to make dijet spectrum in CR more similar to SR.
- * Problems of the 2 methods above: signal is smaller and peaks at the different mass than the mass of the mediator, but it is not still enough to neglect it's contribution.
- 3. Use jets from CR (some sub-region of $\Delta\eta > 1.1$) and try to reconstruct the background dijet mass spectrum.
 - It might further remove the contribution from the signal.
- 4. *Alternative method:* find the lowest fit range directly using 10% of data in the signal region (if expected limits obtained with 10% of the data are already excluded by other analysis).



Dijet mass distribution for part of the data (left) in SR (blue) and CR (red). Difference is $(SR-CR)/SR$ in %. One can see more steep distribution of SR than CR.

Dijet mass spectrum for the signal sample (right) in SR (blue) and CR (red). The peak at the resonance mass is not presented, but still we have big signal contamination.

❖ For the further details see [1]:
https://indico.cern.ch/event/805513/contributions/3372466/attachments/1818786/2975978/dijet_ISR_1.pdf

DARK MATTER INTERPRETATION

- Dark matter particles: a Dirac fermion χ and a particle mediating interaction Z' (s-channel)
- Four parametric model: $m_{DM}, M_{med}, g_q, g_{DM}$.
- Lagrangian with spin-1 mediator model:

$$\mathcal{L}_{vector} = -g_{DM} Z'_\mu \bar{\chi} \gamma^\mu \chi - g_q \sum_{q=u,d,s,c,b,t} Z'_\mu \bar{q} \gamma^\mu q$$

- The total width is given by the sum of all partial widths:

$$\Gamma_{vector}^{\chi \bar{\chi}} = \frac{g_{DM}^2 M_{med}}{12\pi} (1 - 4z_{DM})^{1/2} (1 + 2z_{DM}); z_{DM} = \frac{m_{DM}^2}{M_{med}^2}, Z' \rightarrow \chi \bar{\chi}$$

$$\Gamma_{vector}^{q \bar{q}} = \frac{g_q^2 M_{med}}{4\pi} (1 - 4z_q)^{1/2} (1 + 2z_q); z_q = \frac{m_q^2}{M_{med}^2}, Z' \rightarrow q \bar{q}$$

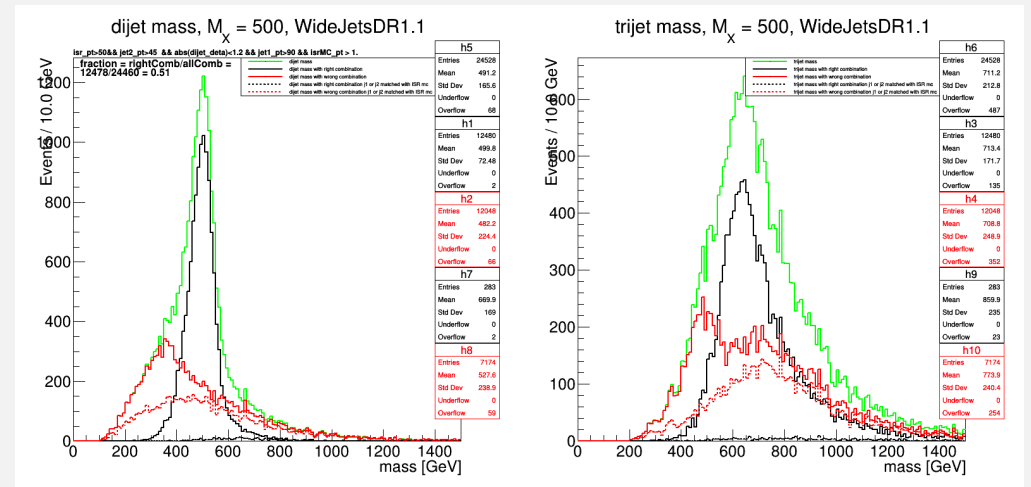
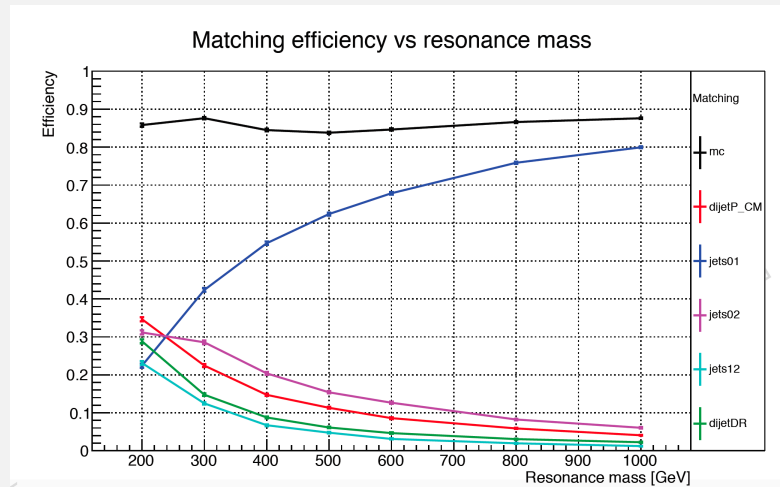
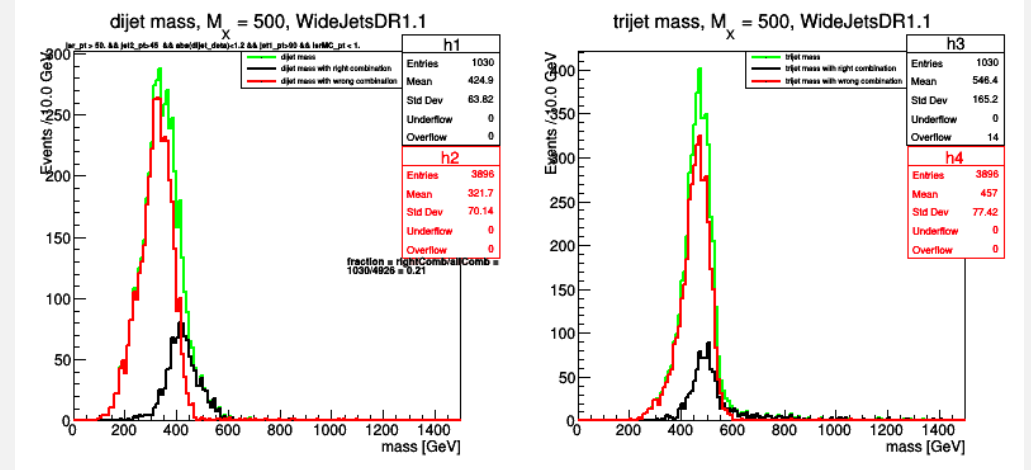
- Simplified DM model (suggested by ATLAS/CMS DM forum): $Z' \rightarrow qq$ or $DM DM$ with unknown $m_{DM} \neq 0$.
- $g_q = 0.25, g_{DM} = 1$: scan the mediator masses from 300 to 600 GeV.
- The cross sections for Z' masses of 300 to 1000 GeV with 50 GeV step was calculated.

DM mediator mass (GeV)	Cross section (pb)
300	230
350	142.7
400	99.3
450	71.1
500	50.4
550	35.6
600	26.4
...	...

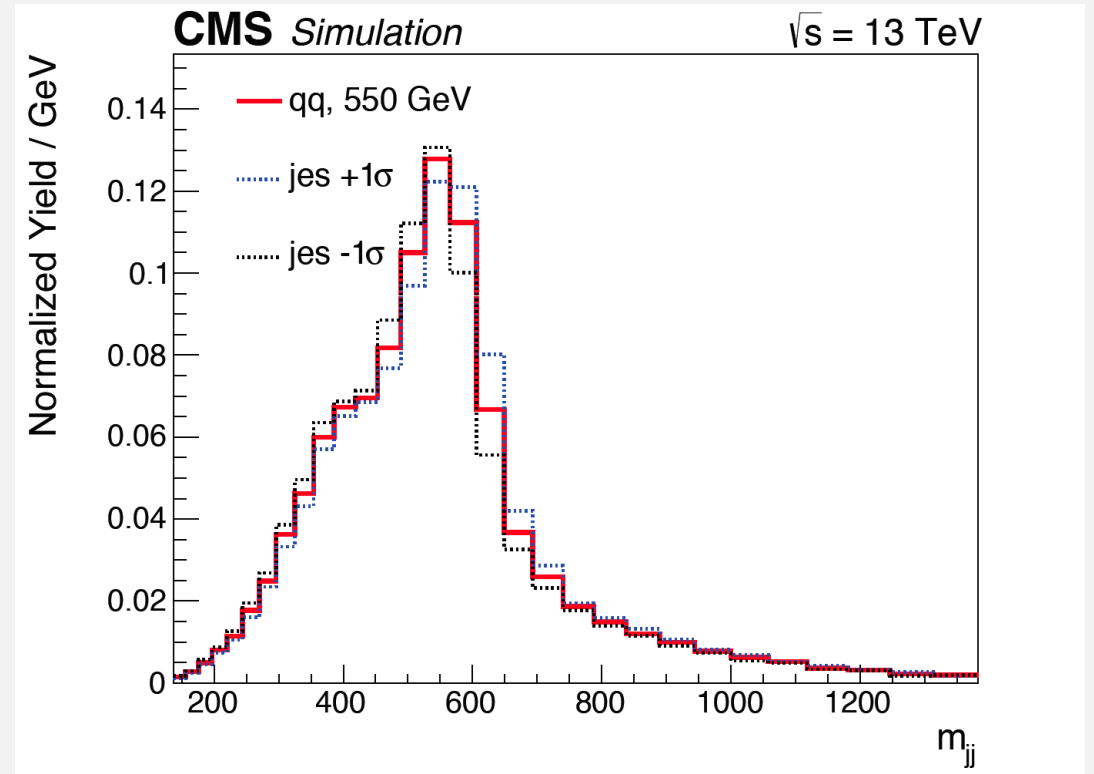
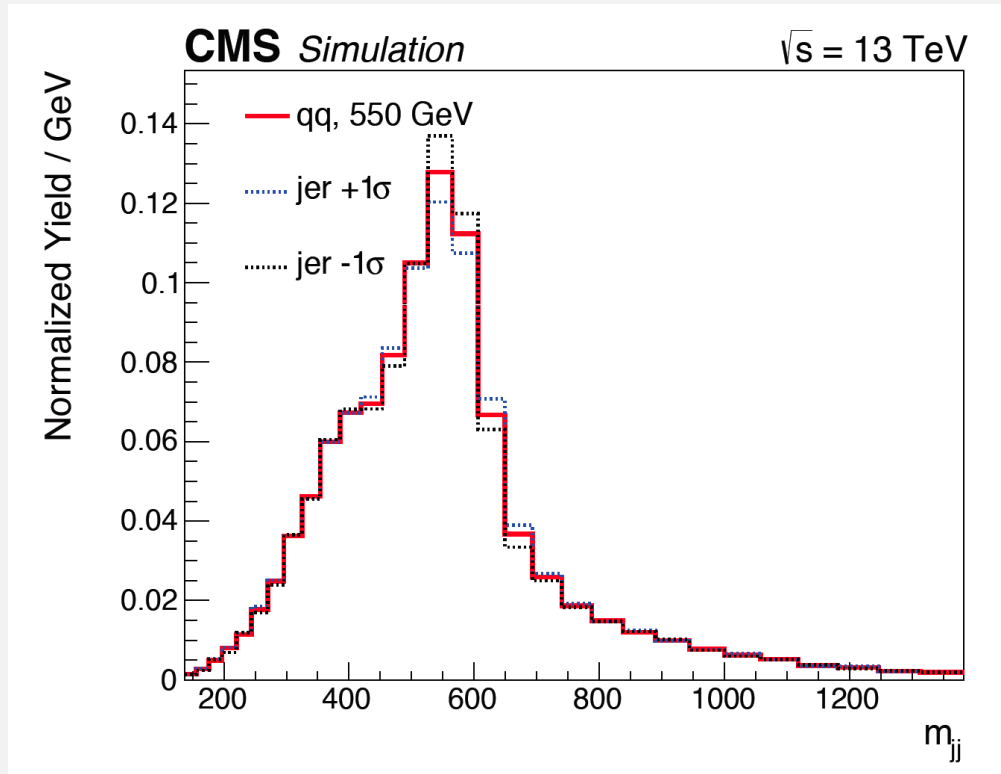
Cross sections using $|\Delta\eta_{jj}| < 1.1$ and $|\eta| < 2.5$

DIJET MASS SIGNAL SHAPE STUDIES

- Signal: narrow Breit-Wigner resonance
- Shape is no gaussian-like due to the following reasons:
 - Not all FSR jets are included inside WJ cone
 - FSR is misinterpreted as ISR



JES AND JER UNCERTAINTIES



Systematic Uncertainty Source	Uncertainty
Jet Energy Resolution	10% of RECO resolution
Jet Energy Scale	$\pm 2\%$ shift of m_{jj}
Luminosity	$\pm 2.5\%$
Parton shower	$\pm 6.0\%$

JET ENERGY CORRECTIONS

- How to compute JES and JER for Calo Scouting analysis?
- Take 2 datasets with RECO and HLT data containing the same events.

JER

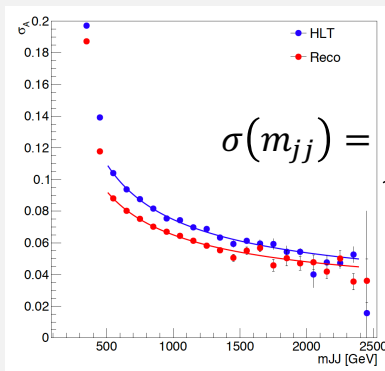
- Dijet pT balancing method. Main variable $A = \frac{p_T^{j1} - p_T^{j2}}{p_T^{j1} + p_T^{j2}}$
- $\sigma(A) \sim \frac{\sigma(m_{jj})}{m_{jj}}$
- $c \rightarrow \text{Gaussian} \left(1, \left(\frac{\sigma(m_{jj})}{m_{jj}} \right)_{RECO} \sqrt{c_{SF}^2 - 1} \right), c_{SF} = \frac{\left(\frac{\sigma(m_{jj})}{m_{jj}} \right)_{HLT}}{\left(\frac{\sigma(m_{jj})}{m_{jj}} \right)_{RECO}}$

JES

- $\frac{\langle p_T(HLT, probe) \rangle}{\langle p_T(RECO, probe) \rangle} \rightarrow f(\langle p_T(HLT, probe) \rangle)$
- $\frac{p_T(HLT, probe)}{p_T(RECO, tag)} \rightarrow f(p_T(RECO, tag))$
- $\frac{p_T(RECO, probe)}{p_T(RECO, tag)} \rightarrow f(p_T(RECO, tag))$
- $\frac{p_T(HLT, probe)}{p_T(RECO, probe)} \rightarrow f(p_T(RECO, tag))$
- $p_T(HLT, probe) \rightarrow f(p_T(RECO, tag))$

Datasets	Objects
/ParkingScoutingMonitor/Run2016B-v2/RAW /ParkingScoutingMonitor/Run2016C-v2/RAW /ParkingScoutingMonitor/Run2016D-v2/RAW	HLT Jets
/ParkingScoutingMonitor/Run2016B-PromptReco-v2/MINIAOD /ParkingScoutingMonitor/Run2016C-PromptReco-v2/MINIAOD /ParkingScoutingMonitor/Run2016D-PromptReco-v2/MINIAOD	RECO Jets

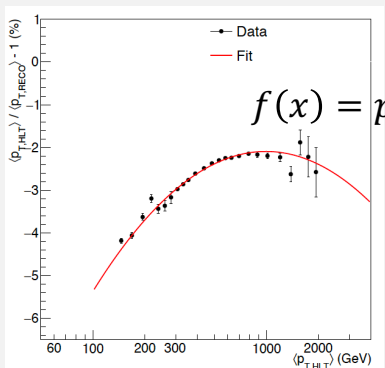
Table 9: Dataset containing both HLT jets and RECO jets.



$$\sigma(m_{jj}) = \sqrt{\frac{N^2}{m_{jj}^2} + \frac{S^2}{m_{jj}} + C^2}$$

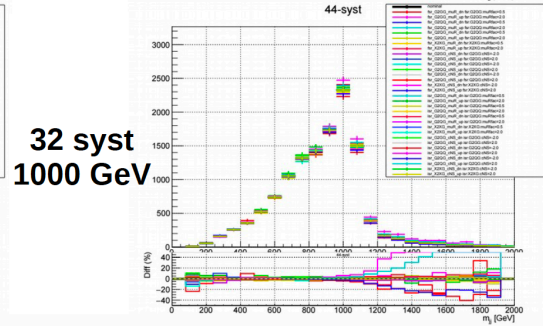
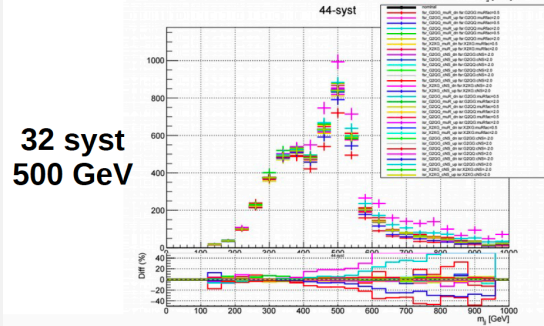
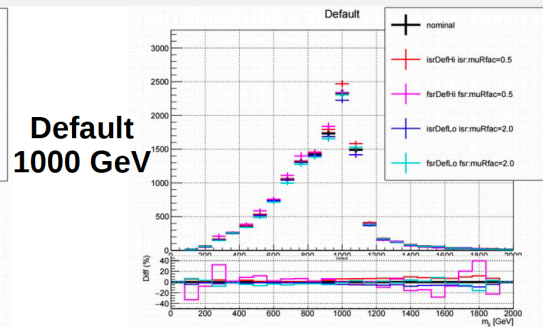
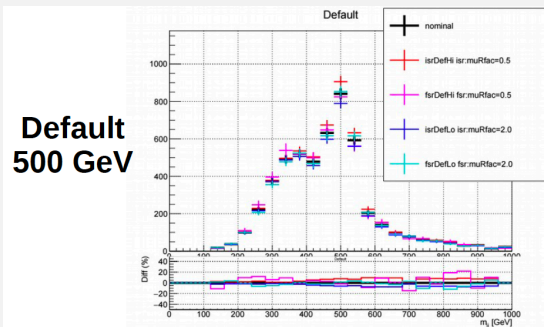
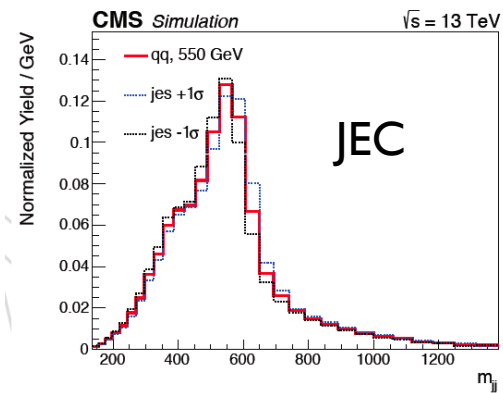
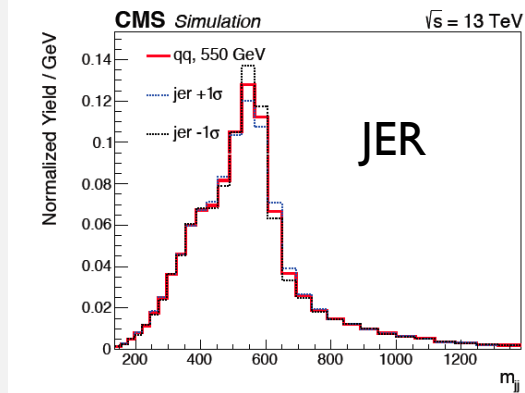
HLT and RECO jets common cuts	
AK4 jets	p_T (leading jet) > 10 GeV p_T (sub-leading jet) > 10 GeV $ \eta < 2.5$
Widejets	p_T (leading widejet) > 10 GeV p_T (sub-leading widejet) > 10 GeV $ \eta < 2.5$ $\Delta\phi_{jj} > 2.8$ $p_T^3 < 0.3 p_{Tave}$
Geometrical Matching	$\Delta R(HLT, RECO) < 0.3$

Table 10: Event selection



$$f(x) = p_0 + p_1 \log(x) + p_2 \log^2(x)$$

SYSTEMATIC UNCERTAINTIES



Parton showering:
 reduced (ISR $mR = 1/\sqrt{2}$, FSR $mR = \sqrt{2}$, FSR $mR = 1/\sqrt{2}$, FSR $mR = \sqrt{2}$);
 default (ISR $mR = 1/2$, FSR $mR = 2$, FSR $mR = 1/2$, FSR $mR = 2$);
 conservative (ISR $mR = 1/4$, FSR $mR = 4$, FSR $mR = 1/4$, FSR $mR = 4$);
 32 decorrelated variations of mR .