

Search for dark matter production in association with a single top quark or a top quark pair in the dilepton final state at $\sqrt{s} = 13$ TeV

Pablo Martínez Ruíz del Árbol, Jónatan Piedra Gomez, Cédric Prieëls
February 4th 2022

Thesis Endorsement
Instituto de Física de Cantabria

All the results presented in this work have been documented in the [AN-22-014](#).

Several presentations were regularly made in MET+X meetings:

- 12th of November talk: [link](#)
- 6th of August talk: [link](#)
- 4th of June talk: [link](#)
- 28th of May talk: [link](#)

Available on CMS information server

CMS AN -2022/014

 The Compact Muon Solenoid Experiment
Analysis Note
The content of this note is intended for CMS internal use and distribution only



20 January 2022

Search for dark matter produced in association with top quarks using the full Run II dataset in the dilepton final state

Cedric Prieels, Juanita Piedra Gomez, Pablo Martinez Ruiz del Arbol

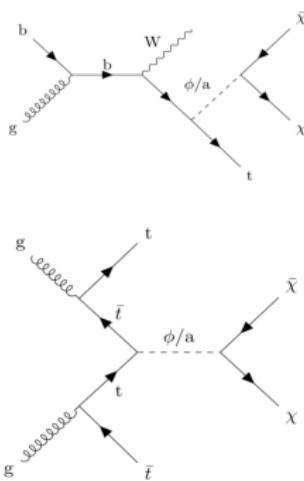
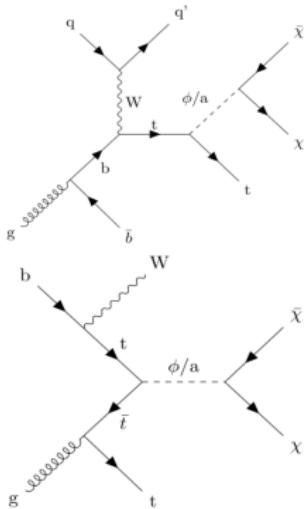
Abstract

A search for dark matter particles produced in association with one or two top quarks in the dilepton final state using Boosted Decision Trees is presented in this work. This analysis has been done by considering the full $(13.7 \pm 2.0) \text{ fb}^{-1}$ of proton-proton collision data collected by the CMS detector during the Run II of operation of the LHC, at a center of mass energy of $\sqrt{s} = 13 \text{ TeV}$. This is the first time that such a search combining the t/\bar{t} and $t/\bar{t} +$ dark matter models is performed in such a final state. No excess in the signal of dark matter has been observed. Upper limits on the signal strength have been obtained by considering different production models. We achieve an experimental (observed) exclusion for scalar and pseudoscalar mediators up to 155 (130) and 150 (105) GeV respectively, at the 95% confidence level. Scalar exclusion limits were therefore improved by a factor of 2 with respect to the previous results obtained in 2016 for the $t/\bar{t} + \text{DM}$ model alone, while a pseudoscalar exclusion has been achieved for the first time when considering this particular final state.

Focus of this thesis I

We are searching for **dark matter produced in association with either one or two top quarks**. Several **simplified models** have been considered:

- Spin 1/2 DM χ ($\in [1, 55]$ GeV, Dirac fermion)
- Spin 0 scalar (S)/pseudoscalar (PS) mediator ϕ/a (Yukawa-like structure of such interactions → **gain from the coupling of the mediator to top quarks**)
- Mediator mass $\in [10, 1000]$ GeV
- Coupling g_χ mediator/DM set to 1 (same for all g_q couplings)



The **typical final state** of such models is made out of:

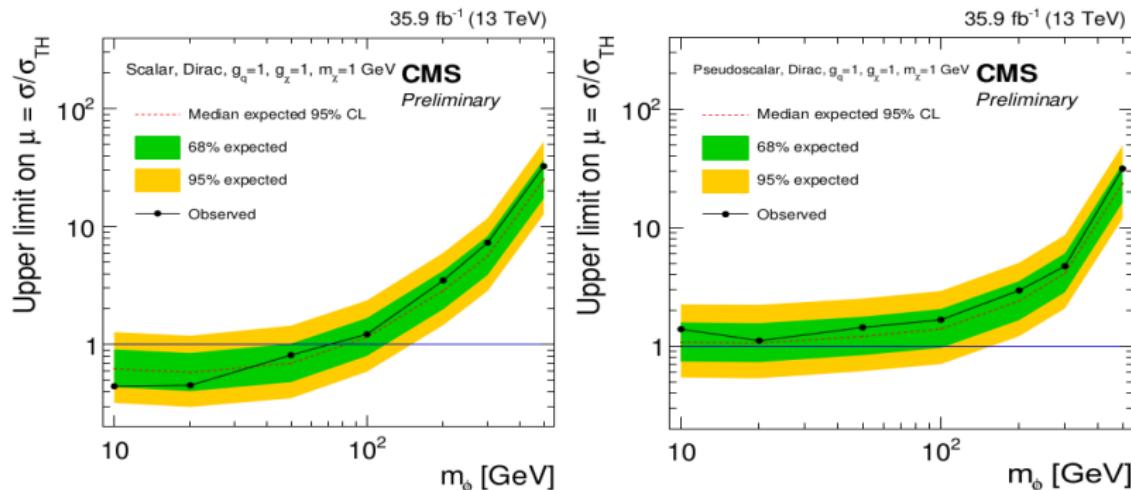
- 1 or 2 b-tagged jets coming from the decay of the top quark(s);
- 2 W bosons, seen as a combination of jets and leptons depending on the channel;
- Some ptmiss coming from the dark matter and the leptonic decay of the Ws;

In particular, we are studying the **dilepton final state** in this work:

- Has the lowest branching ratio: $\text{BR}(W \rightarrow l^+ + \nu_l) = (10.80 \pm 0.09)\%$ for each of the charged leptons (contains only 5% of the signal events);
- But, electrons and muons can usually be reconstructed better than jets, resulting in lower systematic uncertainties;
- And this channel has the lowest number of backgrounds, with cross-sections typically lower, resulting in a better signal isolation.

Previous relevant results I

A similar analysis has already been carried out by CMS using 2016 data, considering only the $t\bar{t}$ +DM signal and the dilepton final state (EXO-17-014).



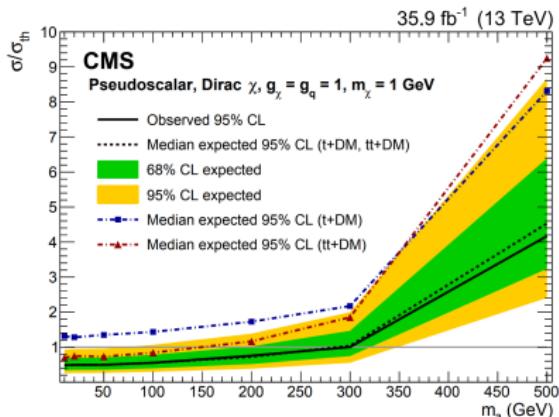
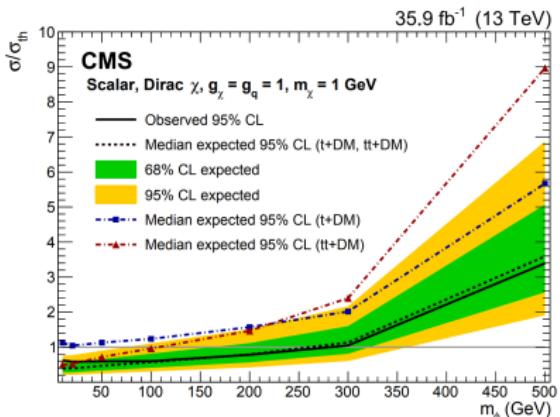
This analysis **excluded scalar mediators** with masses below 80 GeV, while **no exclusion was achieved** when considering pseudoscalar mediators.

A combination of all the different final states was all also performed in EXO-16-049.

Previous relevant results II

A combination of both the t/\bar{t} +DM and $t\bar{t}$ +DM processes has also been performed (EXO-18-010). The inclusion of the single top signal process improved up to a factor 2 the limits obtained by the $t\bar{t}$ analysis on its own. This analysis:

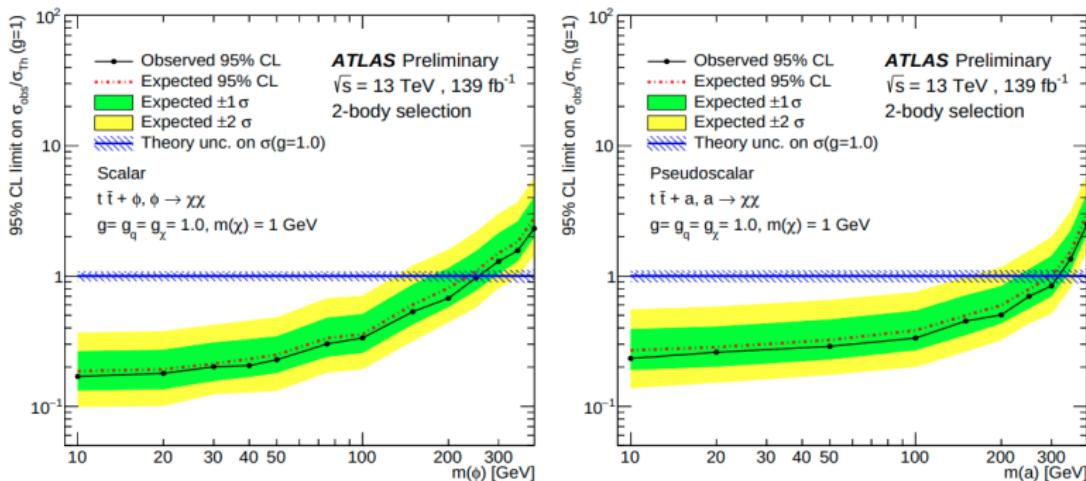
- Only considered the 2016 data-taking period;
- And only considered the semi-leptonic and hadronic final states.



Scalar (pseudoscalar) mediators were with this combination excluded up to 290 (300) GeV at the 95% confidence level.

Previous relevant results III

The ATLAS collaboration also obtained the exclusion limits obtained using the full Run II legacy dataset and considering the dilepton final state only (ATLAS-CONF-2020-046).



They obtained **expected scalar (pseudoscalar) exclusion limits of 250 (300) GeV**, even though they used NLO cross-sections for the signals, around 30% higher than ours.

Analysis context

Analysis strategy

Run II legacy paper being worked on, expected to combine both the $t/\bar{t}+DM$ and $t\bar{t}+DM$ searches, and the 3 possible final states (hadronic, semi-leptonic and dileptonic).
→ Pre-approval process expected to start within a few weeks.

The effort is **globally common** between the groups (Wisconsin, DESY, IFCA) studying the different final states:

- Objects are defined in a common way;
- Control and signal region orthogonal between the channels.
→ Number of leptons and b-jet categorization to improve the sensitivity by defining enriched $t/\bar{t}+DM$ and $t\bar{t}+DM$ regions.

This talk will however **be focused on the dilepton final state only**.

Given that my PhD thesis is now coming to an end, with no further extensions possible, I would like to ask the conveners to endorse the results presented next.

Samples and objects

Data

Single/double leptons datasets built to avoid any eventual double counting, considering the 3 years of the Run II of operation of the LHC:

- $(35.9 \pm 0.9) \text{ fb}^{-1}$
- $(41.5 \pm 1.0) \text{ fb}^{-1}$
- $(59.7 \pm 1.5) \text{ fb}^{-1}$

A blinding policy has been followed at first, allowing us to only look at 1 fb^{-1} of data per year near the signal regions.

Backgrounds

The major backgrounds have been considered from MC and read from NanoAOD. Each year has its corresponding MC samples:

- $t\bar{t}$: decaying to both 1 and 2 leptons;
- Single top: s, t and tW channels considered;
- Drell-Yan: HT-binned samples to increase the statistics, with a correction factor derived from data applied;
- TTZ and TTW: usually grouped together as TTV, and considering both the hadronic and leptonic final states;
- Others, such dibosons and tribosons production, all taken from MC directly.

Signal samples

All the signal samples have been generated using MADGRAPH and PYTHIA8 (with the CP5 tune) at LO, while simulated events are then interfaced with a realistic model of the CMS detector using Geant4 [113] and are reconstructed using the official CMS reconstruction algorithms.

The t/\bar{t} +DM process was **produced privately** (central request has been made but not yet processed), while the $t\bar{t}$ +DM was **generated centrally**. In both cases:

- Both scalar and pseudoscalar mediators are considered;
- 400.000 events were produced for each mediator mass, from 10 to 1000 GeV;
- The dark matter mass was set to 1 GeV, but additional samples ranging from 1 to 55 GeV were also produced;
- All the g_q and g_χ couplings were set to 1.

Recommended correction factors (L1 ECAL prefiring in 2016 and 2017, HEM issue in 2018) are applied to the simulation.

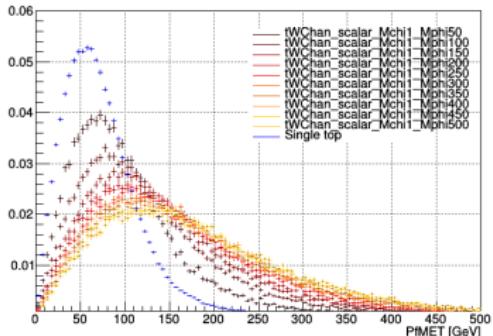
All the samples used and their cross-sections are listed in the backup.

Signal samples

$t/\bar{t}+DM$

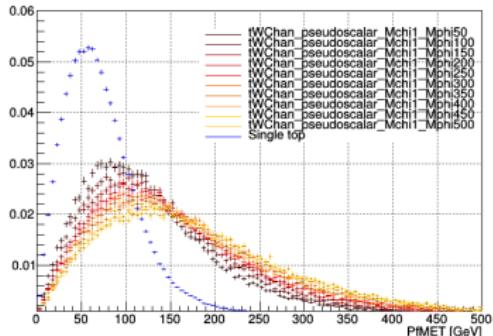
Scalar

Mass points distribution



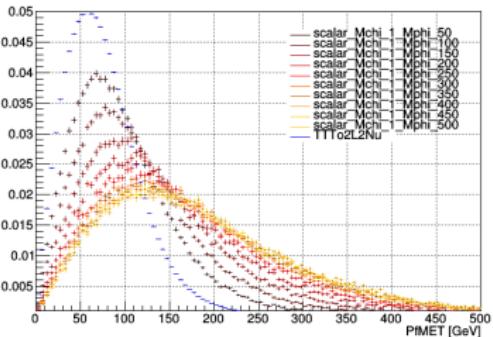
Pseudoscalar

Mass points distribution

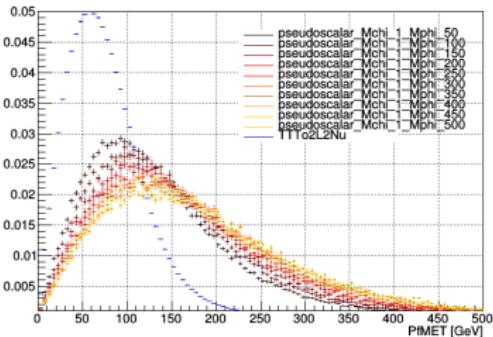


$t\bar{t}+DM$

Mass points distribution



Mass points distribution



Triggers

- Single and double lepton triggers combined to gain statistics, and any possible double counting of events in multiple trigger is taken care of;
- Trigger and lepton p_T carefully chosen to avoid any turn-on effect;
- SingleMuon, SingleEle, DoubleMuon, DoubleEG, MuonEG (2016) and SingleMuon, EGamma, DoubleMuon, MuonEG (2017/2018) data streams considered;
- All the triggers used and their efficiencies (computed using orthogonal MET datasets) are listed in the backup.

Leptons

- Analysis relies on the selection of events with two leptons, with a leading (trailing) $p_T > 25$ (20) GeV and $|\eta| < 2.4$;
- Medium cut based POG WP used for electrons without additional ISO cut;
- Medium cut based POG WP for muons with tight ISO ($\text{pfRelIso04_all} < 0.15$);
- Additional small cuts on the impact parameters to reduce the non-prompt contamination in the ptmiss tail ($|d_0| < 0.05$ cm, $|d_z| < 0.1$ cm, $S_{3D}^d < 4$).

Jets

- Clustered from the PF candidates using the **anti-kT algorithm**;
- Basic selection: $p_T > 30 \text{ GeV}$, $|\eta| < 2.4$;
- **Tight JET/MET POG** working point (efficiency and background rejection $> 98\%$), tight jet PU ID applied to jets with $p_T < 50 \text{ GeV}$ to reject PU jets contamination;
- $\Delta R > 0.4$ away from any lepton passing the criteria established for analysis to prevent signal leptons clustered as jets from entering the jet counting.

B-tag

- B-Tagging and Vertexing POG **deep CSV b-tag medium working point** (high efficiency, misidentification rate for a light jet as a b-jet $\sim 1\%$).

Missing transverse momentum

- **PfType1MET** considered by propagating the JECs to the ptmiss;
- All recommended **filters applied** to filter anomalous high ptmiss events due to several detector issues, such as eventual dead cells in the calorimeters;
- XY-shift (ϕ modulation fix) and EE noise (2017) corrections applied on top.

Event selection

Minimal event selection

We require for the analysis:

- Two opposite sign leptons, with leading (trailing) $p_T > 25$ (20) GeV;
- Third lepton veto ($p_T < 10$ GeV);
- $m_{\parallel} > 20$ GeV to avoid low mass resonances;
- At least 1 jet.

Pre-selection region

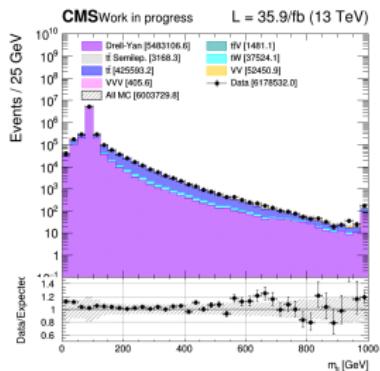
A pre-selection region is then defined by additionally asking for:

- At least 1 medium deep CSV b-jet (misidentification rate of light jets around 1%);
- A 76-106 GeV Z-veto on the ee and $\mu\mu$ channels;
- $\text{ptmiss} > 100$ GeV and $M_{T2}^{\parallel} > 80$ GeV to keep this region orthogonal to the $t\bar{t}$ control regions used by the semi-leptonic channel.

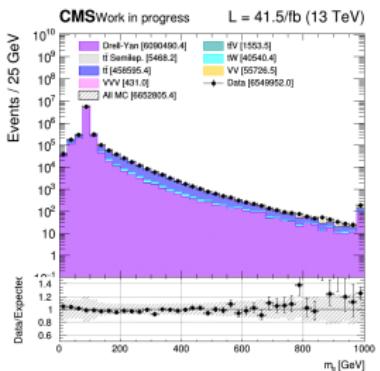
This region is used as the basis for the definition of our signals regions.

Minimal event selection region

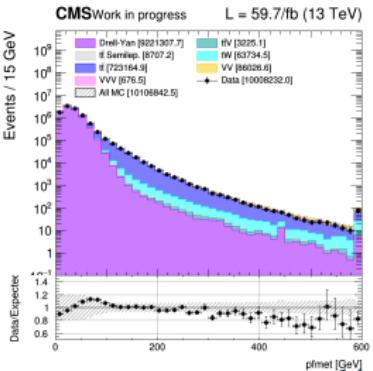
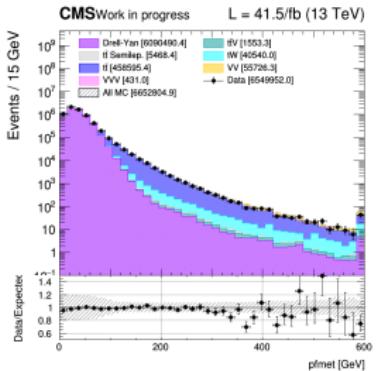
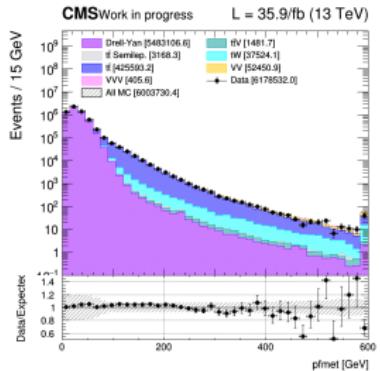
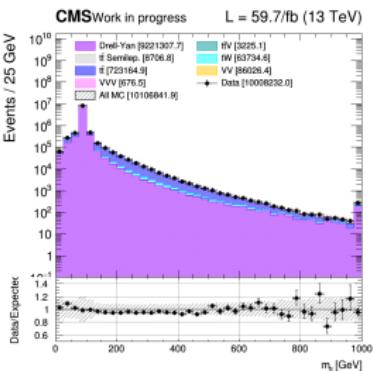
2016



2017

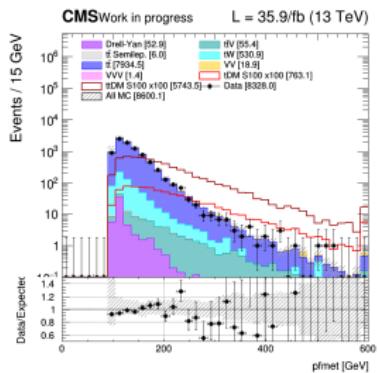


2018

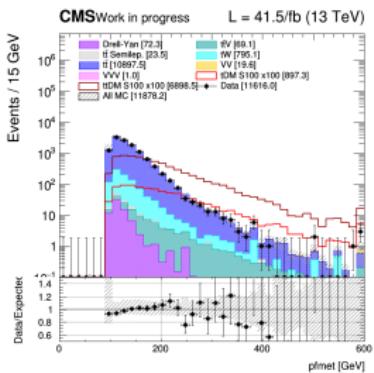


Pre-selection region

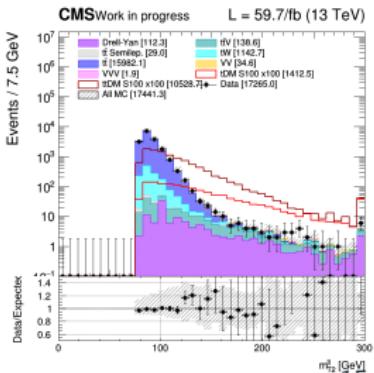
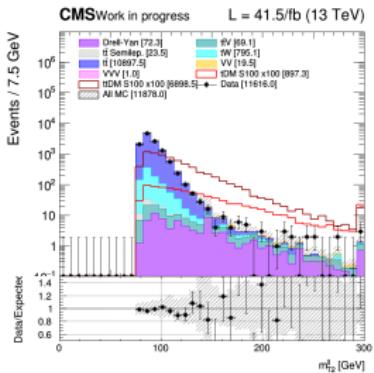
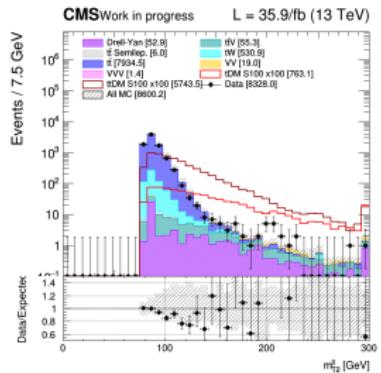
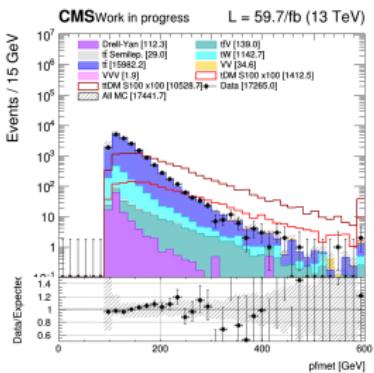
2016



2017



2018



Background prediction methods

Main background processes

The backgrounds are predicted either directly from Monte-Carlo simulations or from semi data-driven methods.

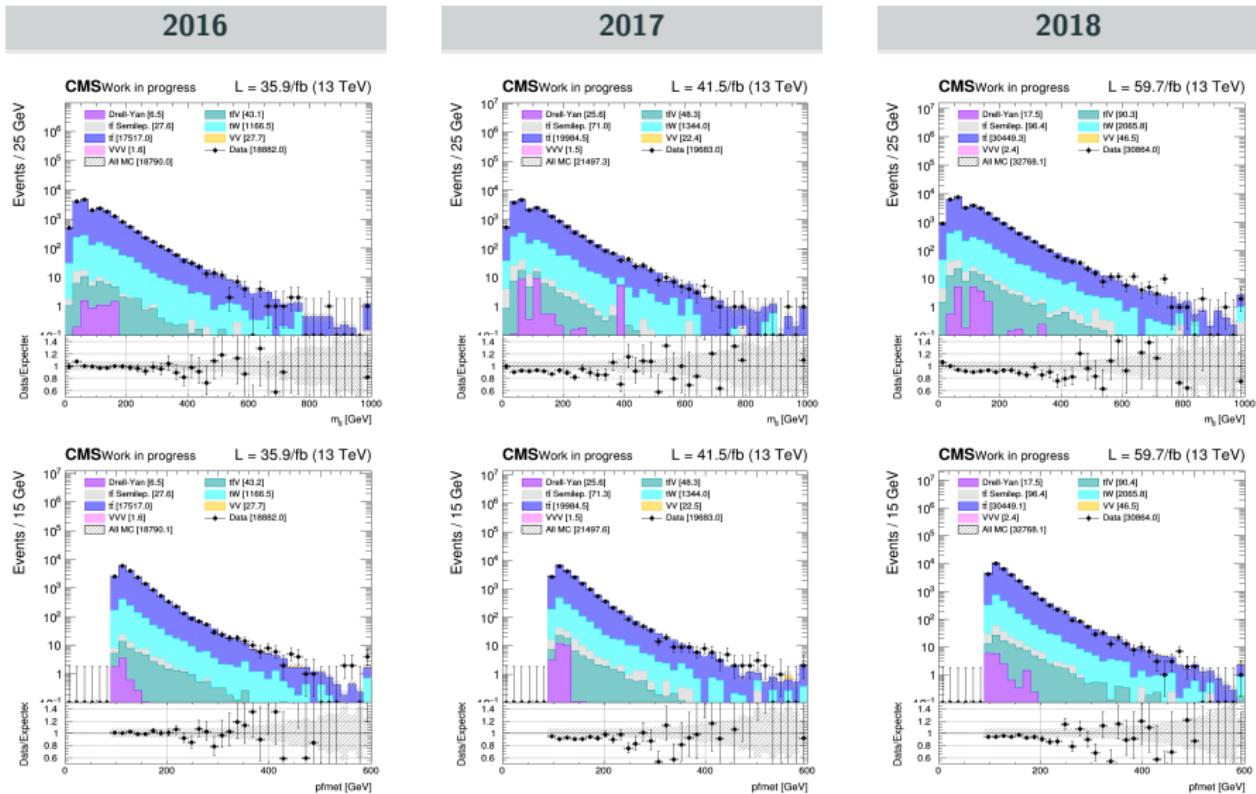
- The **$t\bar{t}$ and the single top** are taken from simulation accounting for all the variations in the generation parameters. Several parameters (QCD scale, PDF variation,...) are varied and included as a systematic (see later);
- The **Drell-Yan** yields are obtained from a semi data-driven method using the excluded same flavor region on the Z peak as control region;
- **ttV , diboson, triboson processes and other minor backgrounds** are taken directly from MC simulations.

Recommended correction factors (L1 ECAL prefiring in 2016 and 2017, HEM issue in 2018) are also applied to the simulation.

Data validation regions enriched in top and Drell-Yan have been explored to ensure the quality of the prediction made.

Top control region

Same as the pre-selection region but with $\text{ptmiss} > 50 \text{ GeV}$ and $60 < M_{T2}^H < 80 \text{ GeV}$.



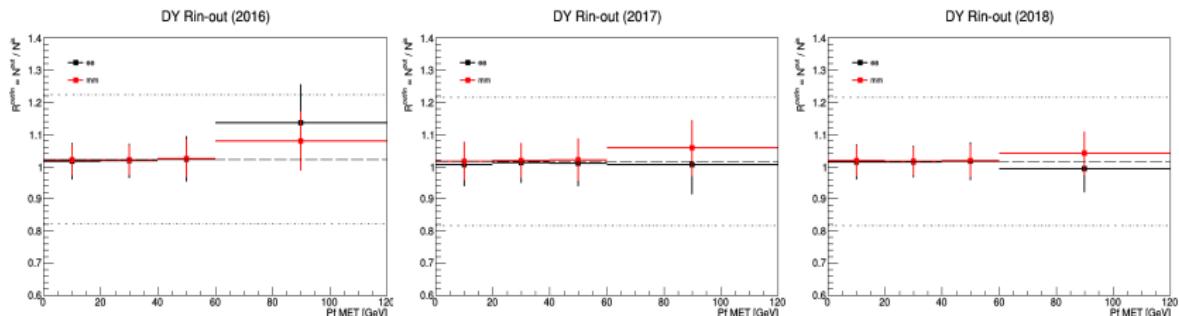
DY Rin-out method

We want to estimate the DY yields outside of the Z-peak from the data in the minimal event selection region:

- Given the presence of large backgrounds (such as $t\bar{t}$) in the analysis region, we go inside of the Z-peak to compute the **Rin-out factor**:

$$N_{DY}^{out} = N_{DY, \text{data}}^{in} \cdot \kappa \cdot \left(\frac{N_{DY, MC}^{out}}{N_{DY, MC}^{in}} \right) \equiv N_{DY, \text{data}}^{in} \cdot \frac{R_{out/in, MC}^{0bj}}{R_{out/in, \text{data}}^{0bj}} \cdot R_{out/in, MC}$$

- To avoid any bias, the contamination of non-peaking backgrounds is removed and we correct this factor by the ratio κ between the data/MC transfer factors in a CR close to the SR (asking for 0 b-jet instead of 1);
- We then get this Rin-out in **bins of ptmiss** and for each channel ($ee, \mu\mu$):

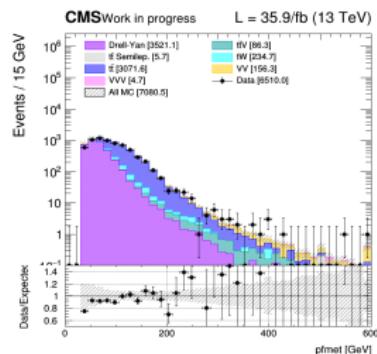
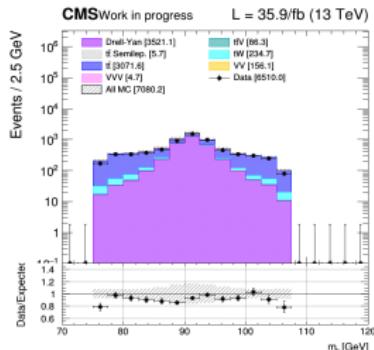


A flat scale factor and a fixed 20% systematic uncertainty is then applied to the DY. 18/36

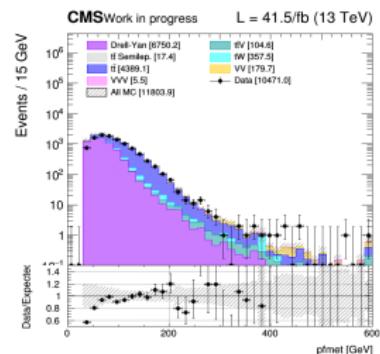
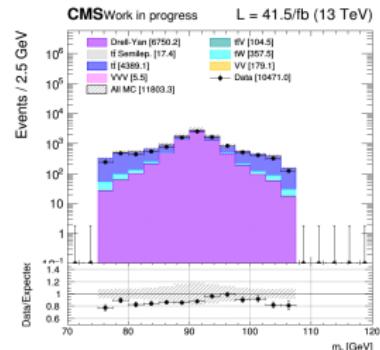
DY control region

Same as the pre-selection region but with $\text{ptmiss} > 30 \text{ GeV}$ and Z-veto reversed.

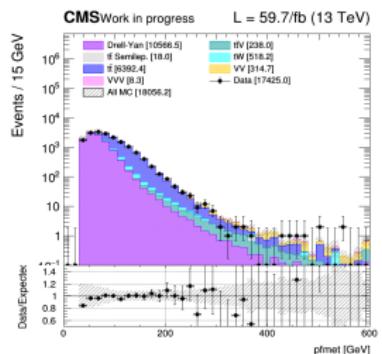
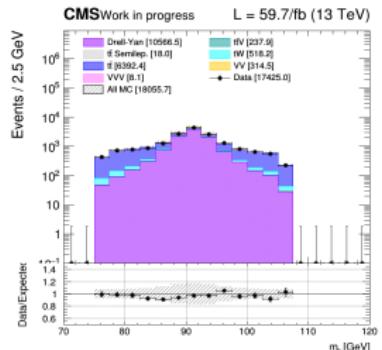
2016



2017



2018



Signal extraction

In this analysis, two different signal regions are being used, targeting each one of our signals of interest, each based on the pre-selection region:

- One targeting the t/\bar{t} +DM signal, by considering events having exactly 1 jet, or exactly 2 jets and 1 b-jet;
- Another one targeting the $t\bar{t}$ +DM signal, by considering events having exactly 2 jets and more than 1 b-jet, or more than 2 jets.

Several different discriminating variables (ptmiss, stranverse mass M_{T2}^H , spin correlated variables, etc.) are being considered.

Many of such variables require knowledge of the top quark and anti-quark 4-momenta, only available after a complete reconstruction of the $t\bar{t}$ system, performed whenever possible (details in the backup).

A BDT combines the discriminating power of all the variables considered, and a ANN is being used as a cross check. A complete optimization was followed in order to select features which maximize the performance.

Discriminating variables I

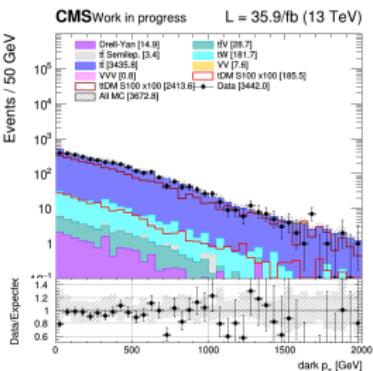
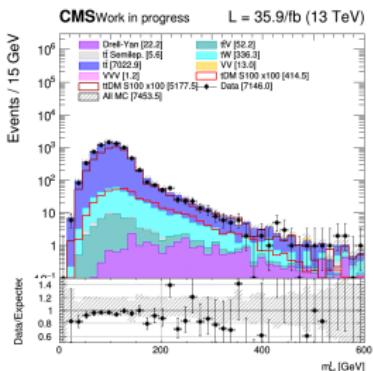
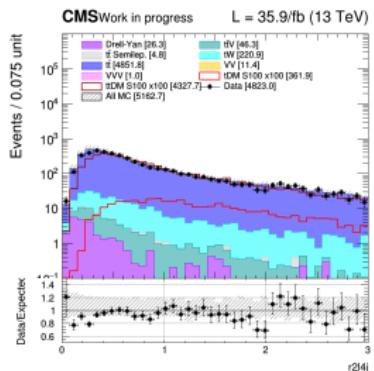
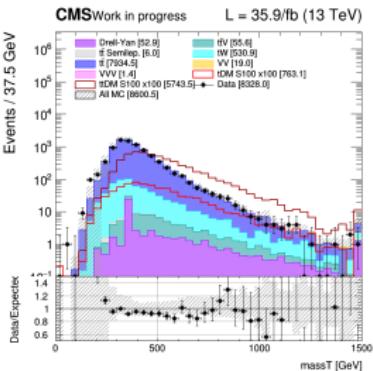
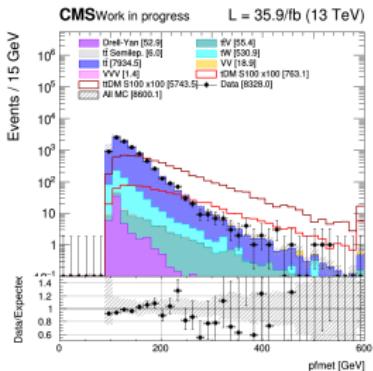
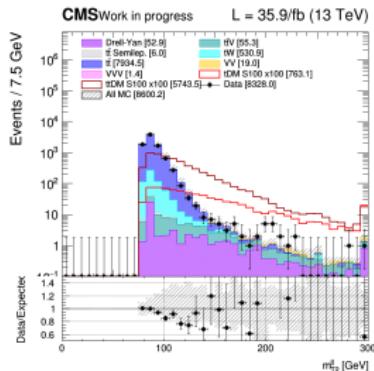
Several discriminating variables (all detailed in the backup) are considered, such as:

- The transverse mass M_{T2}^{\parallel}
- The missing transverse momentum
- The number of b-jets (only in the $t\bar{t}$ +DM signal region) and m_{bl}^t variable, useful to separate our two signals
- Several spin correlated variables
- r_{2l} and r_{2l4j} , defined as the ratio between the ptmiss and the p_T of the leptons (+ the 4 first jets for r_{2l4j})
- The dark p_T and overlapping factor naturally arising from the top reconstruction
- Other variables, such as the angle $\Delta\phi$ between the ptmiss and the two leptons, and total transverse mass massT.

Rank	t/\bar{t} +DM region		$t\bar{t}$ +DM region	
	Variable	Importance	Variable	Importance
1	M_{T2}^{\parallel}	$5.96 \cdot 10^{-1}$	M_{T2}^{\parallel}	$5.87 \cdot 10^{-1}$
2	E_T^{miss}	$5.32 \cdot 10^{-1}$	E_T^{miss}	$5.09 \cdot 10^{-1}$
3	massT	$3.11 \cdot 10^{-1}$	massT	$3.98 \cdot 10^{-1}$
4	r_{2l4j}	$1.70 \cdot 10^{-1}$	r_{2l4j}	$3.45 \cdot 10^{-1}$
5	m_{bl}^t	$7.05 \cdot 10^{-2}$	m_{bl}^t	$1.72 \cdot 10^{-1}$
6	r_{2l}	$2.63 \cdot 10^{-2}$	Dark p_T	$3.92 \cdot 10^{-2}$

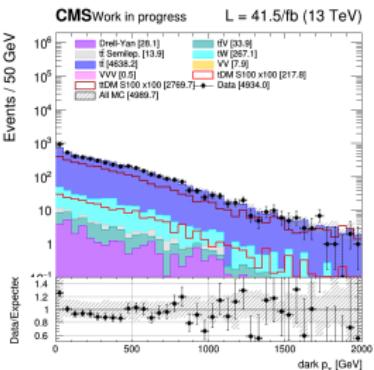
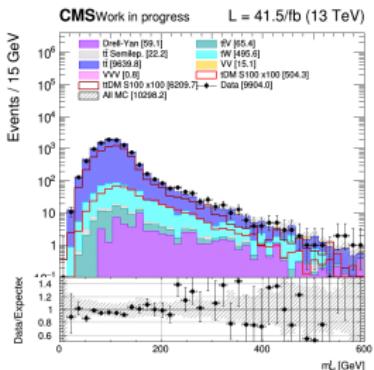
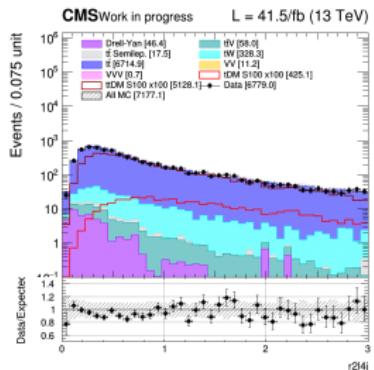
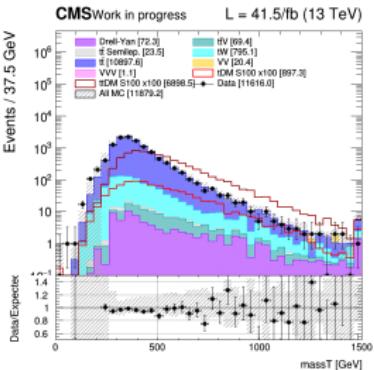
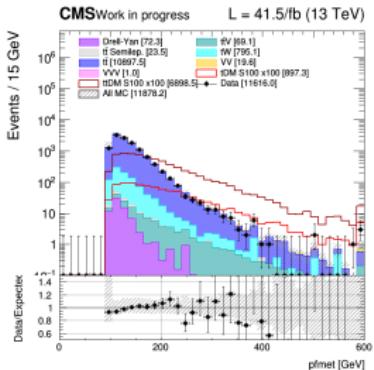
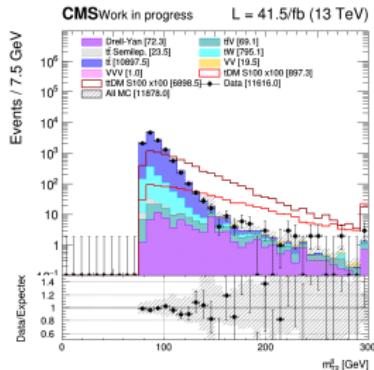
Discriminating variables II

2016



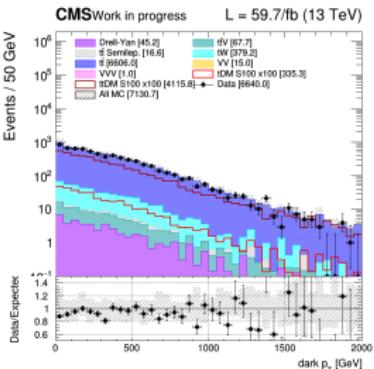
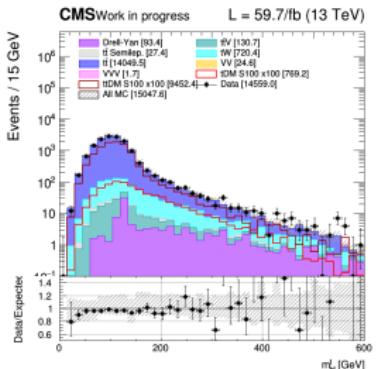
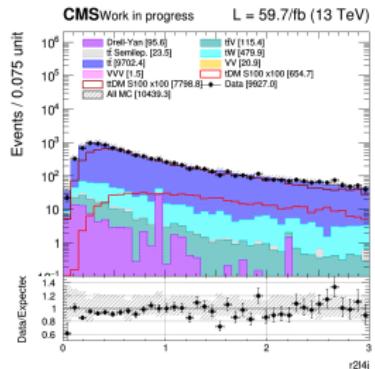
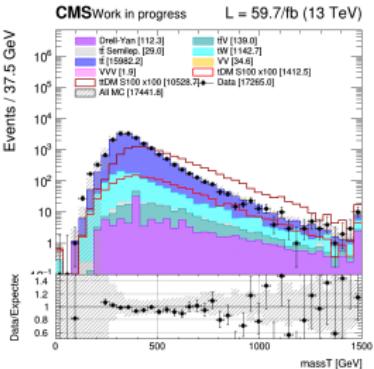
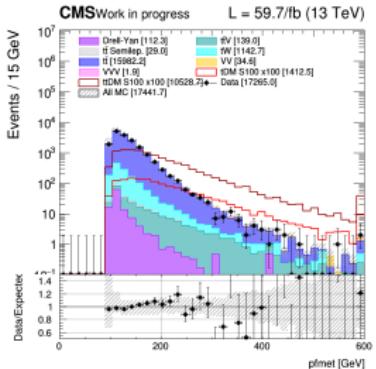
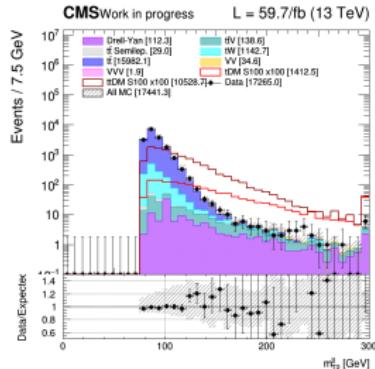
Discriminating variables III

2017



Discriminating variables IV

2018



We trained both a BDT and an ANN, featuring the following common characteristics:

- Mix of standard model $t\bar{t}$ and single top as **backgrounds**, and mix of both $t/\bar{t}+DM$ and $t\bar{t}+DM$ as **signals**;
- Only events passing the **pre-selection cuts** are considered for the training;
- One specific training performed per signal mass point, and per signal region:
 - One targeting the $t/\bar{t}+DM$ signal, by considering events having exactly 1 jet, or exactly 2 jets and 1 b-jet;
 - Another one targeting the $t\bar{t}+DM$ signal, by considering events having exactly 2 jets and more than 1 b-jet, or more than 2 jets.
- 70%/30% train/test splitting used (~ 50.000 training events in total);
- 14 different discriminating variables used as input, all documented in the backup.

At the end of the day though, the **BDT was chosen for the analysis** over the ANN, given that it gave $\sim 10\%$ better upper limits once optimized. The BDT output shape is then used to perform a general **shape analysis**.

Hyperparameters optimization

The hyperparameters of the BDT **have all been fully optimized** one by one, trying each time to minimize the error in the test dataset and the discrimination obtained.

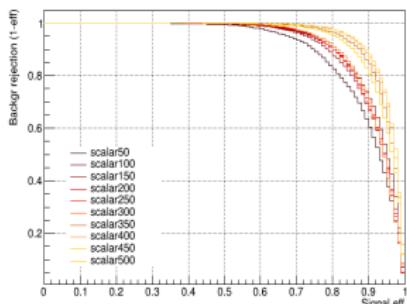
BDT parameter	Optimized value
Maximum depth	4
Minimum samples per leaf	2%
Loss function	Quadratic
Boost algorithm	Gradient descent
Shrinkage	0.3
Grid points n_{cut}	1000
Number of trees	250

ROC curves

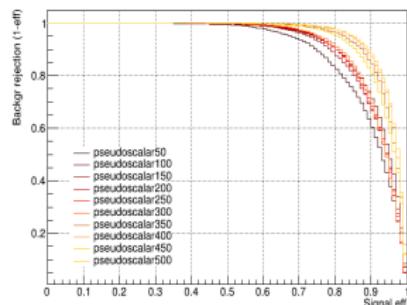
ROC curves have been obtained for all the different mass points available, from 50 to 500 GeV, for both scalar and pseudoscalar mediators, in both signal regions.

$t/\bar{t}+DM$ region

ROC curves for several trainings

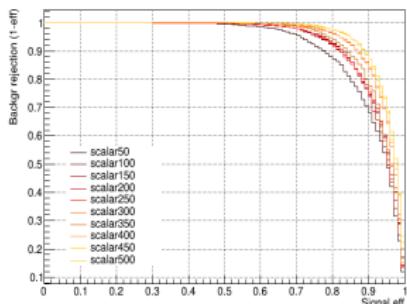


ROC curves for several trainings

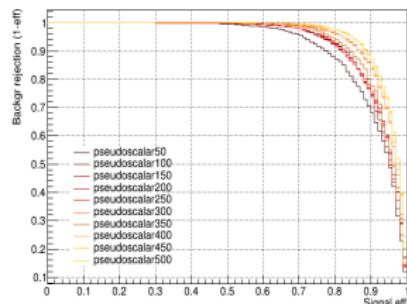


$t\bar{t}+DM$ region

ROC curves for several trainings



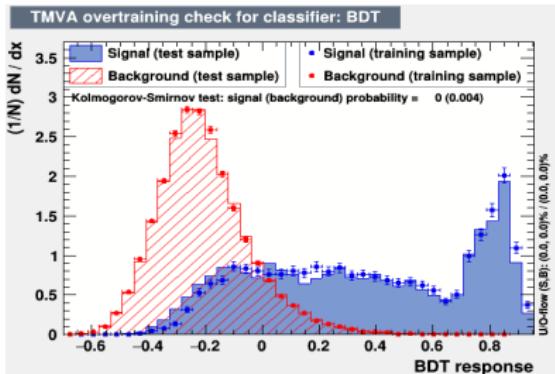
ROC curves for several trainings



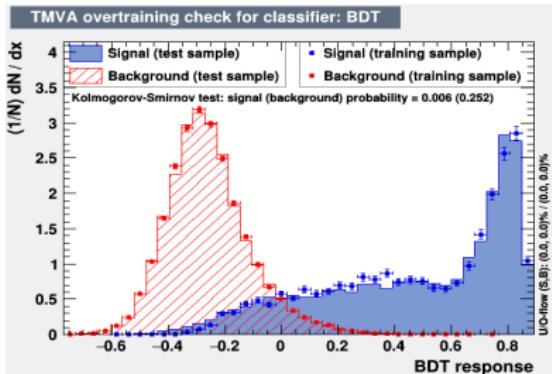
Overtraining check (t/\bar{t} +DM region)

Scalar mediators

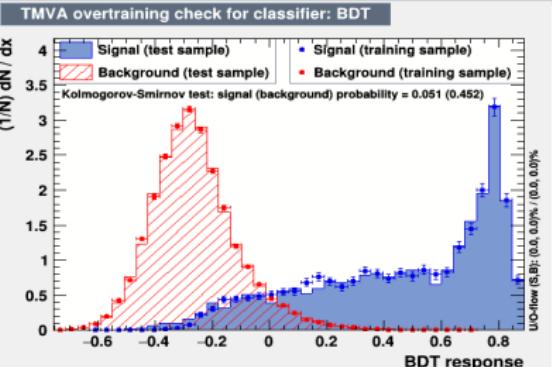
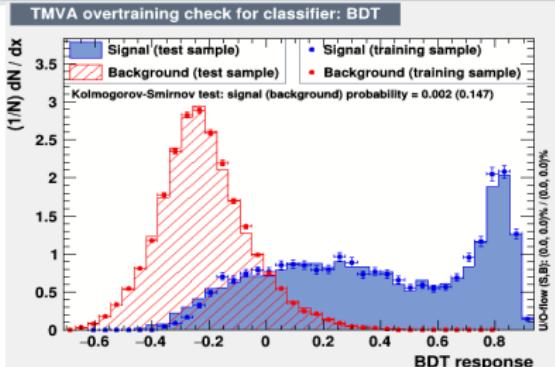
100 GeV



500 GeV



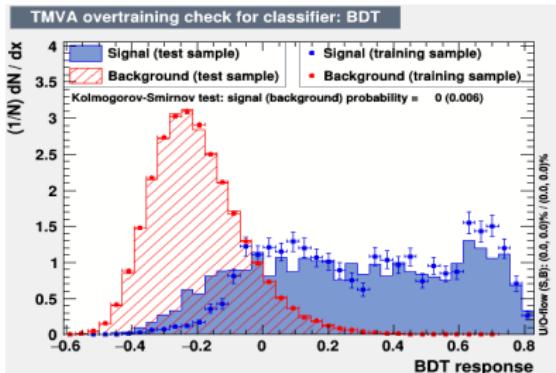
Pseudoscalar mediators



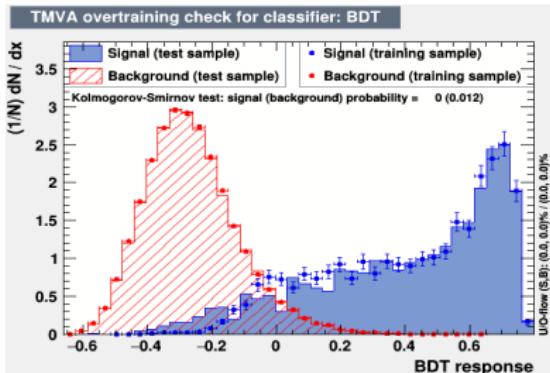
Overtraining check ($t\bar{t}$ +DM region)

Scalar mediators

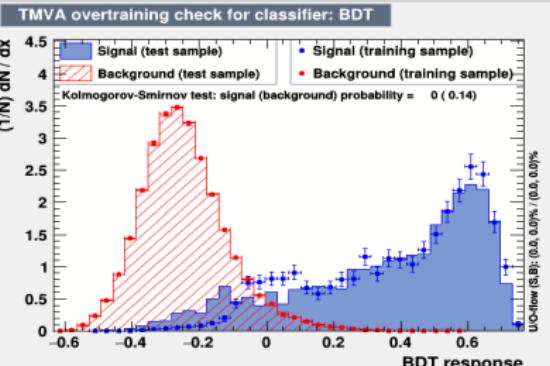
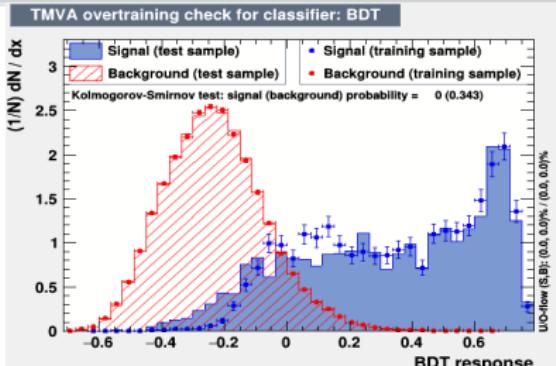
100 GeV



500 GeV



Pseudoscalar mediators

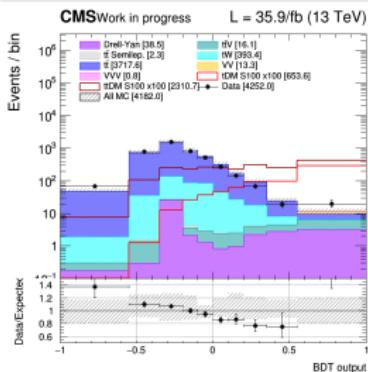


Signal regions

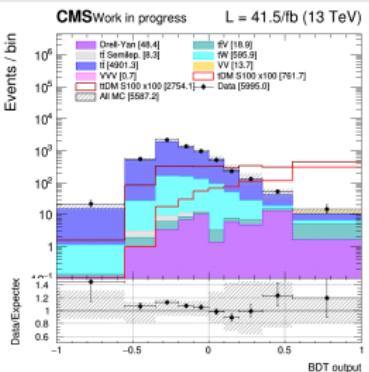
BDT scalar 100 GeV output shape

t/\bar{t} +DM signal region

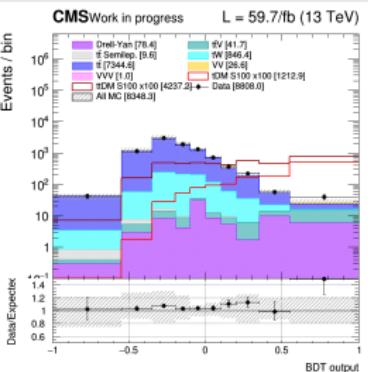
2016



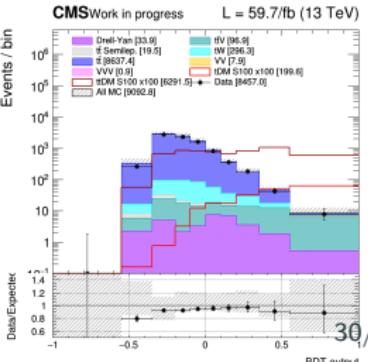
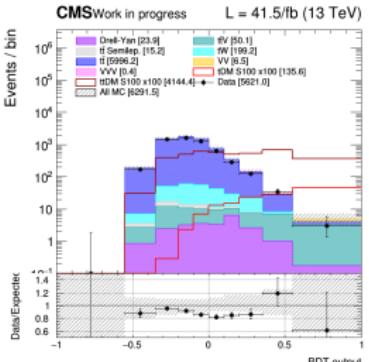
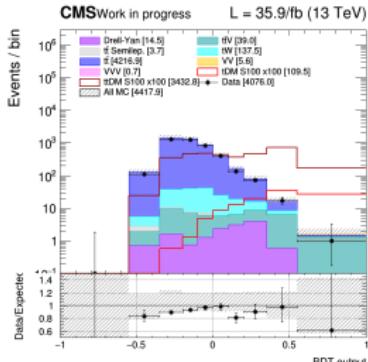
2017



2018



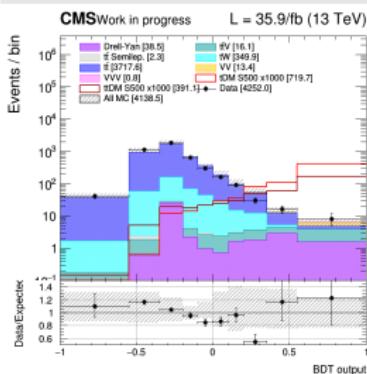
$t\bar{t}$ +DM signal region



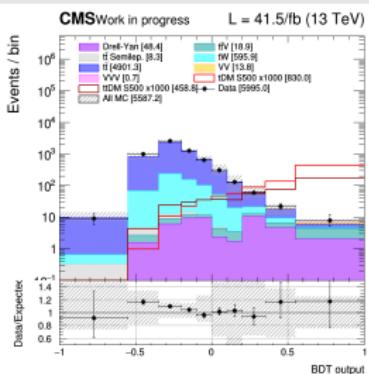
BDT scalar 500 GeV output shape

t/\bar{t} +DM signal region

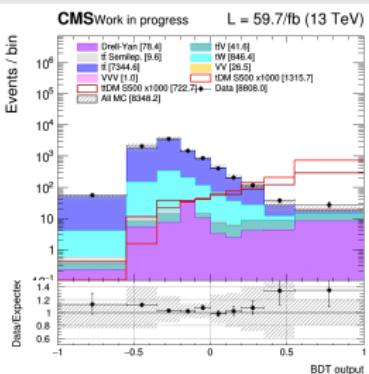
2016



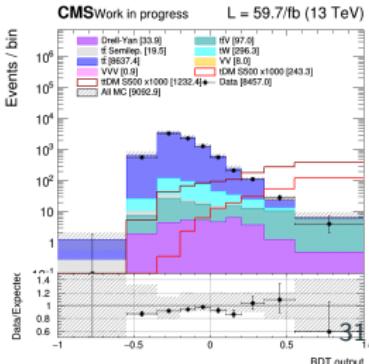
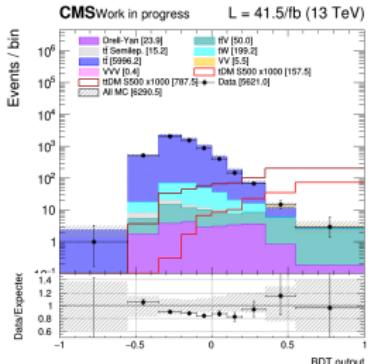
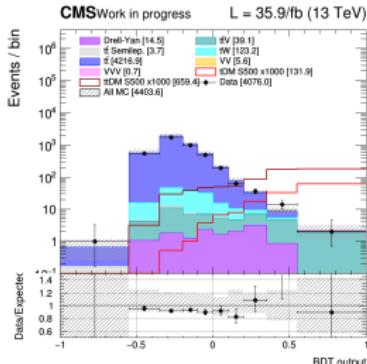
2017



2018



t/\bar{t} +DM signal region



Systematic uncertainties

Systematic uncertainties

On top of statistical uncertainties, many systematics have been considered:

Theoretical uncertainties

- PDF and higher order corrections ($\sim 4\%$), underlying event (1.5%) and parton shower modeling ($\sim 4\%$), renormalization and factorization scales.

Experimental uncertainties

- Luminosity ($\sim 2.5\%$), pileup modeling ($\sim 5\%$), lepton trigger ($\sim 2\%$), lepton efficiency and energy scale ($\sim 2\%$), jet energy scale ($\sim 3\%$), ptmiss mismodelling ($\sim 3\%$), b-tagging efficiency, top p_T reweighting, ECAL prefiring.

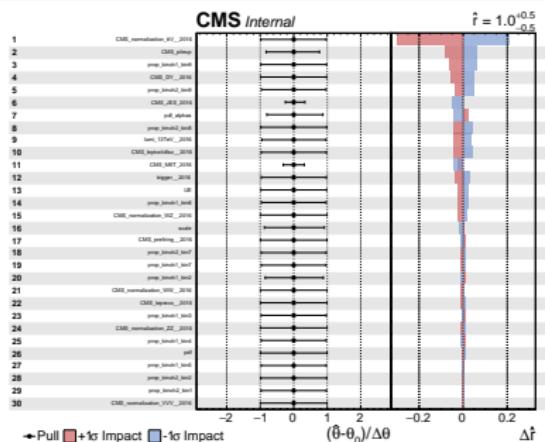
Background specific uncertainties

- MC statistical uncertainties
- 20% systematic uncertainty associated to the DY process in order to cover for the non-flatness of the $R_{\text{in-out}}$ transfer factor;
- 30% uncertainty associated to the normalization of all the minor backgrounds, except for the ttV, for which a 50% systematic uncertainty is associated.

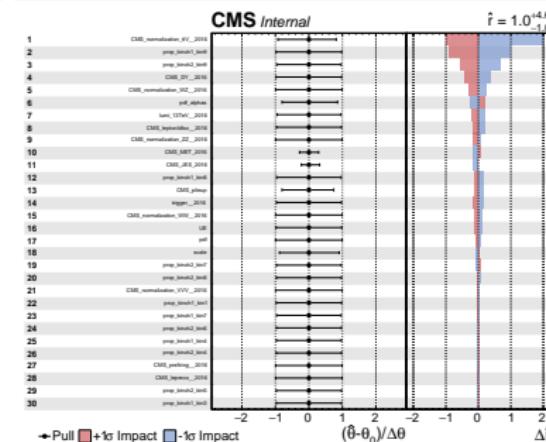
Pulls and impact plots

2016, scalar

100 GeV



500 GeV



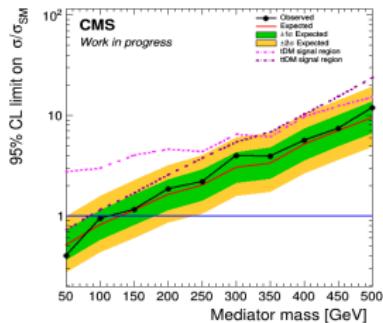
The most important systematics in each case is the normalization of the ttV process.
 Additional impact plots can be found in the backup.

Results obtained

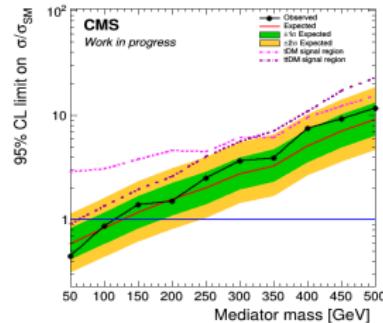
Upper limits

Scalar upper limits

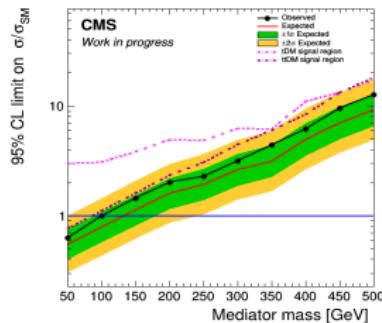
2016



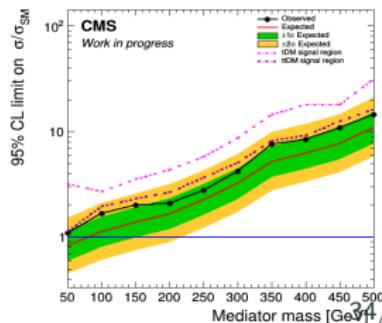
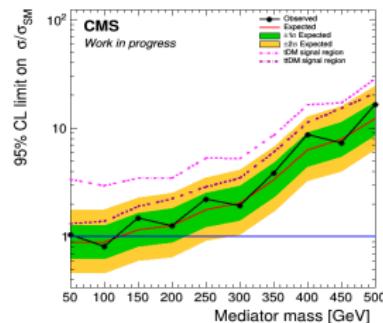
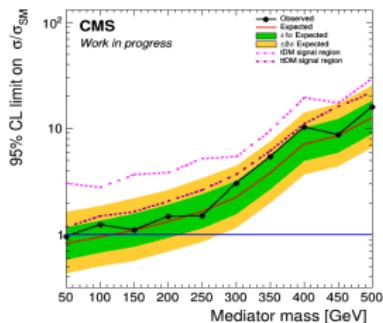
2017



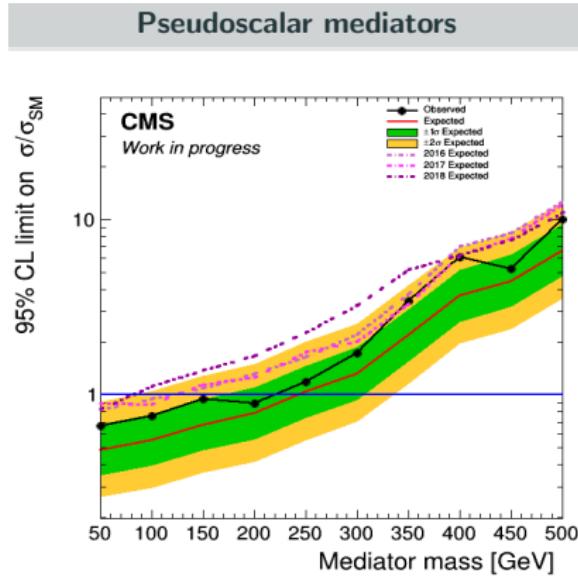
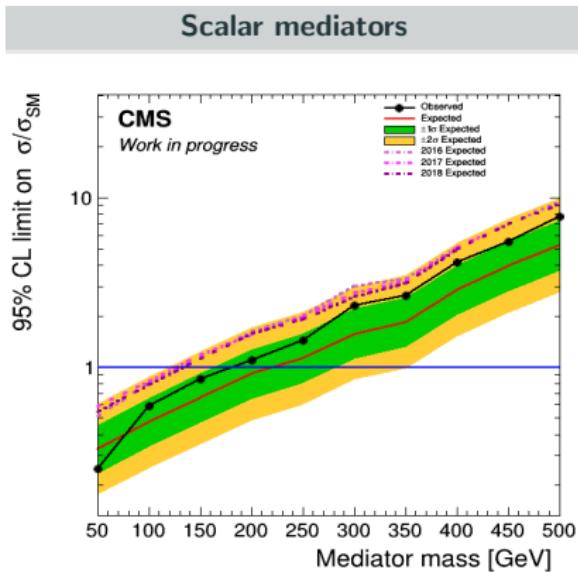
2018



Pseudoscalar upper limits



Run II legacy limits



After combining the different years, the following **expected (observed) exclusion have been achieved**:

- Scalar mediators excluded to 215 (180) GeV;
- Pseudoscalar mediators excluded up to 250 (220) GeV.

Conclusions

A search for **dark matter produced in association with either one or two top quarks** has been performed, considering in particular its dilepton final state, and analyzing the **Run II legacy dataset** collected by the CMS detector at 13 TeV.

This is the **first time that such a combination of two signals of interest** is performed considering this final state.

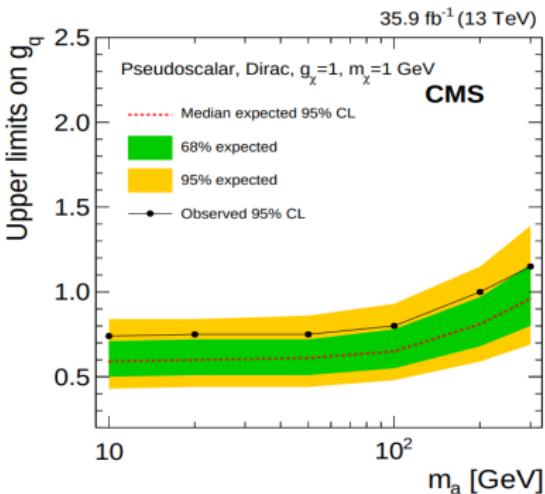
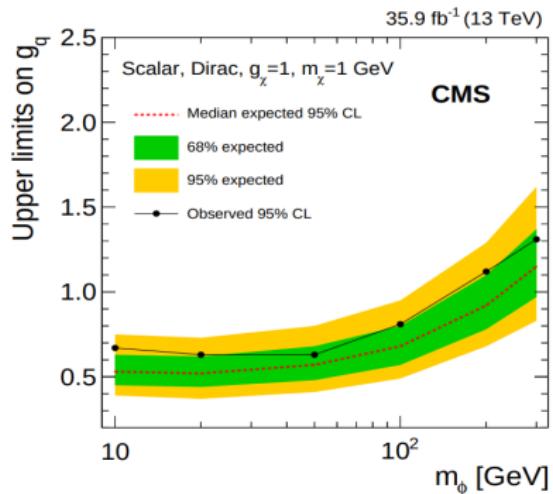
This search **improves by a factor of 2** the scalar exclusion limits obtained in 2016 and **manages to exclude for the first time pseudoscalar mediators** up to 150 GeV.

We would therefore like to ask for the endorsement of the work performed here.

Back up

Additional relevant results

CMS combination of all the different final states published in 2016:



The observed (expected) limits excluded a **pseudoscalar mediator** with mass below 220 (320) GeV, and a **scalar mediator** with mass below 160 (240) GeV.

2016 data samples

Dataset	Events (size)	\mathcal{L} [fb $^{-1}$]
Run 2016B		
/DoubleEG/Run2016B_ver2-Nano02Apr2020_ver2-v1/NANOAOOD	143073268 (99.4Gb)	
/DoubleMuon/Run2016B_ver2-Nano02Apr2020_ver2-v1/NANOAOOD	82535526 (53.2Gb)	
/MuonEG/Run2016B_ver2-Nano02Apr2020_ver2-v1/NANOAOOD	32727796 (26.8Gb)	5.8
/SingleElectron/Run2016B_ver2-Nano02Apr2020_ver2-v1/NANOAOOD	246440440 (167.8Gb)	
/SingleMuon/Run2016B_ver2-Nano02Apr2020_ver2-v1/NANOAOOD	158145722 (96.4Gb)	
Run 2016C		
/DoubleEG/Run2016C-Nano02Apr2020-v1/NANOAOOD	47677856 (35.3Gb)	
/DoubleMuon/Run2016C-Nano02Apr2020-v1/NANOAOOD	27934629 (19.7Gb)	
/MuonEG/Run2016C-Nano02Apr2020-v1/NANOAOOD	15405678 (12.8Gb)	2.6
/SingleElectron/Run2016C-Nano02Apr2020-v1/NANOAOOD	97259854 (69.3Gb)	
/SingleMuon/Run2016C-Nano02Apr2020-v1/NANOAOOD	67441308 (42.4Gb)	
Run 2016D		
/DoubleEG/Run2016D-Nano02Apr2020-v1/NANOAOOD	53324960 (39.6Gb)	
/DoubleMuon/Run2016D-Nano02Apr2020-v1/NANOAOOD	33861745 (24.1Gb)	
/MuonEG/Run2016D-Nano02Apr2020-v1/NANOAOOD	23482352 (19.4Gb)	4.2
/SingleElectron/Run2016D-Nano02Apr2020-v1/NANOAOOD	148167727 (104.4Gb)	
/SingleMuon/Run2016D-Nano02Apr2020-v1/NANOAOOD	98017996 (61.3Gb)	
Run 2016E		
/DoubleEG/Run2016E-Nano02Apr2020-v1/NANOAOOD	49877710 (37.9Gb)	
/DoubleMuon/Run2016E-Nano02Apr2020-v1/NANOAOOD	28246946 (20.8Gb)	
/MuonEG/Run2016E-Nano02Apr2020-v2/NANOAOOD	22519303 (19.0Gb)	4.0
/SingleElectron/Run2016E-Nano02Apr2020-v1/NANOAOOD	117321545 (86.5Gb)	
/SingleMuon/Run2016E-Nano02Apr2020-v1/NANOAOOD	90984718 (58.7Gb)	
Run 2016F		
/DoubleEG/Run2016F-Nano02Apr2020-v1/NANOAOOD	34577629 (26.9Gb)	
/DoubleMuon/Run2016F-Nano02Apr2020-v1/NANOAOOD	20329921 (15.3Gb)	
/MuonEG/Run2016F-Nano02Apr2020-v1/NANOAOOD	16002165 (13.6Gb)	3.1
/SingleElectron/Run2016F-Nano02Apr2020-v1/NANOAOOD	70593532 (51.4Gb)	
/SingleMuon/Run2016F-Nano02Apr2020-v1/NANOAOOD	65489554 (42.4Gb)	
Run 2016G		
/DoubleEG/Run2016G-Nano02Apr2020-v1/NANOAOOD	78797031 (61.6Gb)	
/DoubleMuon/Run2016G-Nano02Apr2020-v1/NANOAOOD	45235604 (34.2Gb)	
/MuonEG/Run2016G-Nano02Apr2020-v1/NANOAOOD	33854612 (29.0Gb)	7.6
/SingleElectron/Run2016G-Nano02Apr2020-v1/NANOAOOD	153363109 (109.2Gb)	
/SingleMuon/Run2016G-Nano02Apr2020-v1/NANOAOOD	149912248 (94.6Gb)	
Run 2016H		
/DoubleEG/Run2016H-Nano02Apr2020-v1/NANOAOOD	85388734 (67.7Gb)	
/DoubleMuon/Run2016H-Nano02Apr2020-v1/NANOAOOD	48912812 (37.3Gb)	
/MuonEG/Run2016H-Nano02Apr2020-v1/NANOAOOD	29236516 (26.0Gb)	8.6
/SingleElectron/Run2016H-Nano02Apr2020-v1/NANOAOOD	128854598 (93.8Gb)	
/SingleMuon/Run2016H-Nano02Apr2020-v1/NANOAOOD	174035164 (110.2Gb)	

2017 data samples

Dataset	Events (size)	\mathcal{L} [fb $^{-1}$]
Run 2017B		
/DoubleEG/Run2017B-Nano02Apr2020-v1/NANOAOD	58088760 (46.6Gb)	
/DoubleMuon/Run2017B-Nano02Apr2020-v1/NANOAOD	14501767 (10.8Gb)	
/SingleElectron/Run2017B-Nano02Apr2020-v1/NANOAOD	60537490 (42.2Gb)	
/SingleMuon/Run2017B-Nano02Apr2020-v1/NANOAOD	136300266 (86.2Gb)	
/MuonEG/Run2017B-Nano02Apr2020-v1/NANOAOD	4453465 (4.1Gb)	
Run 2017C		
/DoubleEG/Run2017C-Nano02Apr2020-v1/NANOAOD	65181125 (53.8Gb)	
/DoubleMuon/Run2017C-Nano02Apr2020-v1/NANOAOD	49636525 (39.5Gb)	
/SingleElectron/Run2017C-Nano02Apr2020-v1/NANOAOD	136637888 (102.5Gb)	
/SingleMuon/Run2017C-Nano02Apr2020-v1/NANOAOD	165652756 (109.5Gb)	
/MuonEG/Run2017C-Nano02Apr2020-v1/NANOAOD	15595214 (15.0Gb)	
Run 2017D		
/DoubleEG/Run2017D-Nano02Apr2020-v1/NANOAOD	25911432 (21.6Gb)	
/DoubleMuon/Run2017D-Nano02Apr2020-v1/NANOAOD	23075733 (18.6Gb)	
/SingleElectron/Run2017D-Nano02Apr2020-v1/NANOAOD	51526710 (38.5Gb)	
/SingleMuon/Run2017D-Nano02Apr2020-v1/NANOAOD	70361660 (47.2Gb)	
/MuonEG/Run2017D-Nano02Apr2020-v1/NANOAOD	9164365 (8.9Gb)	
Run 2017E		
/DoubleEG/Run2017E-Nano02Apr2020-v1/NANOAOD	56233597 (49.8Gb)	
/DoubleMuon/Run2017E-Nano02Apr2020-v1/NANOAOD	51589091 (44.4Gb)	
/SingleElectron/Run2017E-Nano02Apr2020-v1/NANOAOD	102121689 (81.3Gb)	
/SingleMuon/Run2017E-Nano02Apr2020-v1/NANOAOD	154630534 (111.0Gb)	
/MuonEG/Run2017E-Nano02Apr2020-v1/NANOAOD	19043421 (19.2Gb)	
Run 2017F		
/DoubleEG/Run2017F-Nano02Apr2020-v1/NANOAOD	74307066 (67.1Gb)	
/DoubleMuon/Run2017F-Nano02Apr2020-v1/NANOAOD	79756560 (68.0Gb)	
/SingleElectron/Run2017F-Nano02Apr2020-v1/NANOAOD	128467223 (105.2Gb)	
/SingleMuon/Run2017F-Nano02Apr2020-v1/NANOAOD	242135500 (178.3Gb)	
/MuonEG/Run2017F-Nano02Apr2020-v1/NANOAOD	25776363 (26.3Gb)	

2018 data samples

Dataset	Events (size)	\mathcal{L} [fb^{-1}]
Run 2018A		
/DoubleMuon/Run2018A-Nano02Apr2020-v1/NANO AOD	75499908 (62.6Gb)	
/EGamma/Run2018A-Nano02Apr2020-v1/NANO AOD	327843843 (261.8Gb)	
/SingleMuon/Run2018A-Nano02Apr2020-v1/NANO AOD	241608232 (167.7Gb)	13.5
/MuonEG/Run2018A-Nano02Apr2020-v1/NANO AOD	32958503 (32.3Gb)	
Run 2018B		
/DoubleMuon/Run2018B-Nano02Apr2020-v1/NANO AOD	35057758 (28.3Gb)	
/EGamma/Run2018B-Nano02Apr2020-v1/NANO AOD	153822427 (123.1Gb)	
/SingleMuon/Run2018B-Nano02Apr2020-v1/NANO AOD	119918017 (82.3Gb)	6.8
/MuonEG/Run2018B-Nano02Apr2020-v1/NANO AOD	16211567 (15.8Gb)	
Run 2018C		
/DoubleMuon/Run2018C-Nano02Apr2020-v1/NANO AOD	34565869 (27.6Gb)	
/EGamma/Run2018C-Nano02Apr2020-v1/NANO AOD	147827904 (119.2Gb)	
/SingleMuon/Run2018C-Nano02Apr2020-v1/NANO AOD	110032072 (75.7Gb)	6.6
/MuonEG/Run2018C-Nano02Apr2020-v1/NANO AOD	15652198 (15.3Gb)	
Run 2018D		
/DoubleMuon/Run2018D-Nano02Apr2020_ver2-v1/NANO AOD	168605834 (128.6Gb)	
/EGamma/Run2018D-Nano02Apr2020-v1/NANO AOD	751348648 (583.6Gb)	
/SingleMuon/Run2018D-Nano02Apr2020-v1/NANO AOD	513867253 (344.5Gb)	
/MuonEG/Run2018D-Nano02Apr2020_ver2-v1/NANO AOD	71961587 (68.6Gb)	32.0

2016 MC samples

Process	Sample	Cross section [pb]
Drell-Yan	DYJetsToLL_M-10to50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610.0
	DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 ($H_T < 70$ GeV)	6077.22
	DYJetsToLL_M-50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	169.9
	DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	147.4
	DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	40.99
	DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	5.678
	DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	1.367
	DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.6304
	DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.1514
	DYJetsToLL_M-50_HT-2500toInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	0.003565
TTTo2L2Nu	TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	87.310
Single top	ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1	3.360
	ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	80.95
	ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1	136.02
	ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85
	ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	35.85
TTToSemiLeptonic	TTToSemilepton_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	364.35
ttV	TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.2814
	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
VZ	WWTo2L2Nu_13TeV-powheg	12.178
	WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.42965
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595
	ZZTo2L2Nu_13TeV_powheg_pythia8	0.5640
	ZZTo2L2Q_13TeV_powheg_pythia8	3.22
Others	WWWW_WWZ_WZZ_ZZZ_WWG	//

2017 MC samples

Process	Sample	Cross section [pb]
Drell-Yan	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	18610
	DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8 ($H_T < 70$ GeV)	6077.22
	DYJetsToLL_M-50_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8	169.9
	DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	147.4
	DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	40.99
	DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	5.678
	DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.367
	DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.6304
	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1514
	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003565
TTTo2L2Nu	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	87.310
Single top	ST_s-channel_4f_leptonDecays_mtop1715_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	3.360
	ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	80.95
	ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8	136.02
	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
TTToSemiLeptonic	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	364.35
ttV	TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.2814
	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
	TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
VZ	WWTo2L2Nu_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8	12.178
	WZTo3LNu_TuneCUETP8M1_13TeV-powheg-pythia8	4.42965
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595
	ZZTo2L2Nu_13TeV_powheg_pythia8	0.5640
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
Others	WWW, WWZ, WZZ, ZZZ, WWG	//

2018 MC samples

Process	Sample	Cross section [pb]
Drell-Yan	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8	18610
	DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8 ($H_T < 70$ GeV)	6077.22
	DYJetsToLL_M-50_HT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	169.9
	DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8	147.4
	DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8	40.99
	DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8	5.678
	DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8	1.367
	DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8	0.6304
	DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8	0.1514
	DYJetsToLL_M-50_HT-2500toInf_TuneCP5_13TeV-madgraphMLM-pythia8	0.003565
TTTo2L2Nu	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	87.310
Single top	ST_s-channel_4f_leptonDecays_TuneCP5_13TeV-madgraph-pythia8	3.360
	ST_t-channel_antitop_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	80.95
	ST_t-channel_top_4f_InclusiveDecays_TuneCP5_13TeV-powheg-madspin-pythia8	136.02
	ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.85
TTToSemiLeptonic	TTToSemiLeptonic_TuneCP5_13TeV-powheg-pythia8	364.35
ttV	TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.2814
	TTZToQQ_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.5297
	TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnloFXFX-madspin-pythia8	0.2043
	TTWJetsToQQ_TuneCUETP8M1_13TeV-amcatnloFXFX-madspin-pythia8	0.4062
VZ	WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8	12.178
	WZTo3LNu_TuneCP5_13TeV-amcatnloFXFX-pythia8	4.42965
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595
	ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	0.5640
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
Others	WWW, WWZ, WZZ, ZZZ, WWG	//

Mass point	Cross-section [pb]
Scalar mediators	
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi10	$4.959 \cdot 10^{-2}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi20	$3.235 \cdot 10^{-2}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi50	$1.323 \cdot 10^{-2}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi100	$5.633 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi150	$3.397 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi200	$2.359 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi250	$1.720 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi300	$1.328 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi350	$1.018 \cdot 10^{-3}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi400	$6.717 \cdot 10^{-4}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi450	$4.535 \cdot 10^{-4}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi500	$3.206 \cdot 10^{-4}$
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi1000	$3.045 \cdot 10^{-5}$
Pseudoscalar mediators	
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi10	$6.151 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi20	$5.869 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi50	$4.946 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi100	$3.658 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi150	$2.754 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi200	$2.097 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi250	$1.616 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi300	$1.253 \cdot 10^{-3}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi350	$7.851 \cdot 10^{-4}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi400	$4.371 \cdot 10^{-4}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi450	$3.095 \cdot 10^{-4}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi500	$2.321 \cdot 10^{-4}$
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi1000	$2.791 \cdot 10^{-5}$

$t\bar{t} + \text{DM}$ signal samples

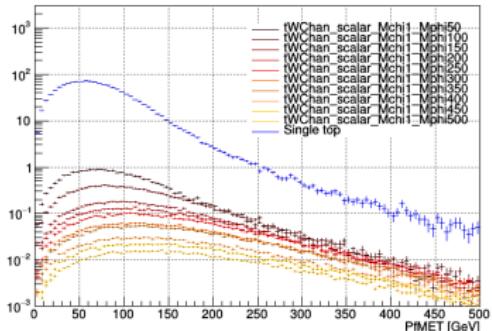
Mass point	Cross-section [pb]
Scalar mediators	
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_50	$3.405 \cdot 10^{-1}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_100	$8.027 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_150	$2.673 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_200	$1.158 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_250	$6.020 \cdot 10^{-3}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_300	$3.579 \cdot 10^{-3}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_350	$2.376 \cdot 10^{-3}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_400	$1.443 \cdot 10^{-3}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_450	$9.025 \cdot 10^{-4}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_500	$6.204 \cdot 10^{-4}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_20.mPhi_100	$7.993 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_30.mPhi_100	$8.052 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_40.mPhi_100	$8.147 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_45.mPhi_100	$8.319 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_49.mPhi_100	$8.304 \cdot 10^{-2}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_51.mPhi_100	$9.735 \cdot 10^{-4}$
TTbarDMJets.Dilepton.scalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_55.mPhi_100	$4.835 \cdot 10^{-4}$
Pseudoscalar mediators	
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_50	$3.440 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_100	$2.164 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_150	$1.414 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_200	$9.773 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_250	$6.753 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_300	$4.808 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_350	$2.742 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_400	$1.409 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_450	$9.302 \cdot 10^{-4}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_1.mPhi_500	$6.618 \cdot 10^{-4}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_20.mPhi_100	$2.166 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_30.mPhi_100	$2.164 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_40.mPhi_100	$2.162 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_45.mPhi_100	$2.180 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_49.mPhi_100	$2.151 \cdot 10^{-2}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_51.mPhi_100	$1.993 \cdot 10^{-3}$
TTbarDMJets.Dilepton.pseudoscalar._LO_TuneCP5_13TeV-madgraph-mcatnlo-pythia8.mChi_55.mPhi_100	$7.750 \cdot 10^{-4}$

Signal samples

$t/\bar{t}+DM$

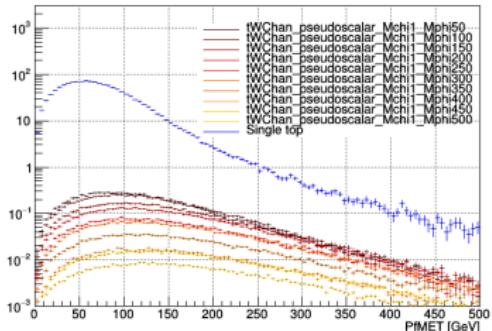
Scalar

Mass points distribution



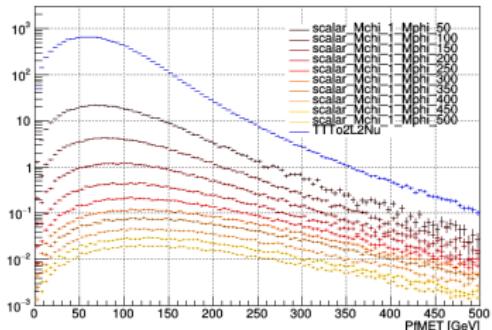
Pseudoscalar

Mass points distribution

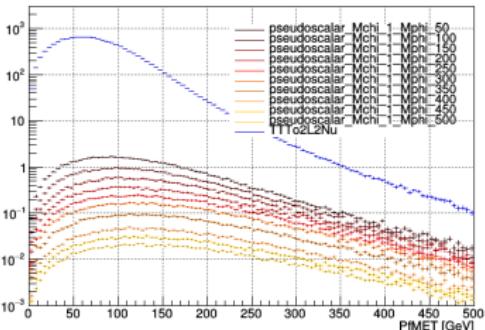


$t\bar{t}+DM$

Mass points distribution



Mass points distribution



2016 triggers

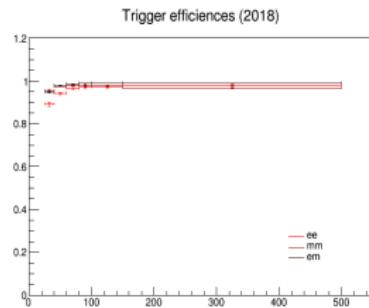
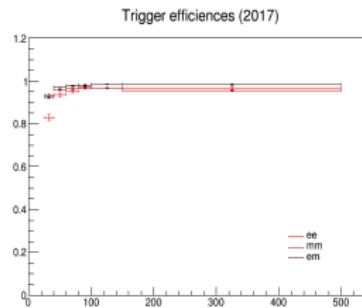
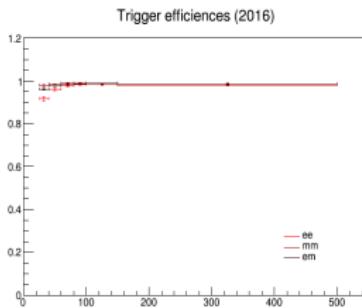
Dataset	Run range	HLT trigger path
SingleMu	[297020,306462]	HLT_IsoMu27_v*
SingleEle	[297020,306462]	HLT_Ele35_WPTight_Gsf_v*
DoubleEG	[297020,306462]	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*
DoubleMu	[297020,299336] [299337,306462]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v*
MuonEG	[297020,306462] [297020,299336] [299337,306462]	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*

2017 triggers

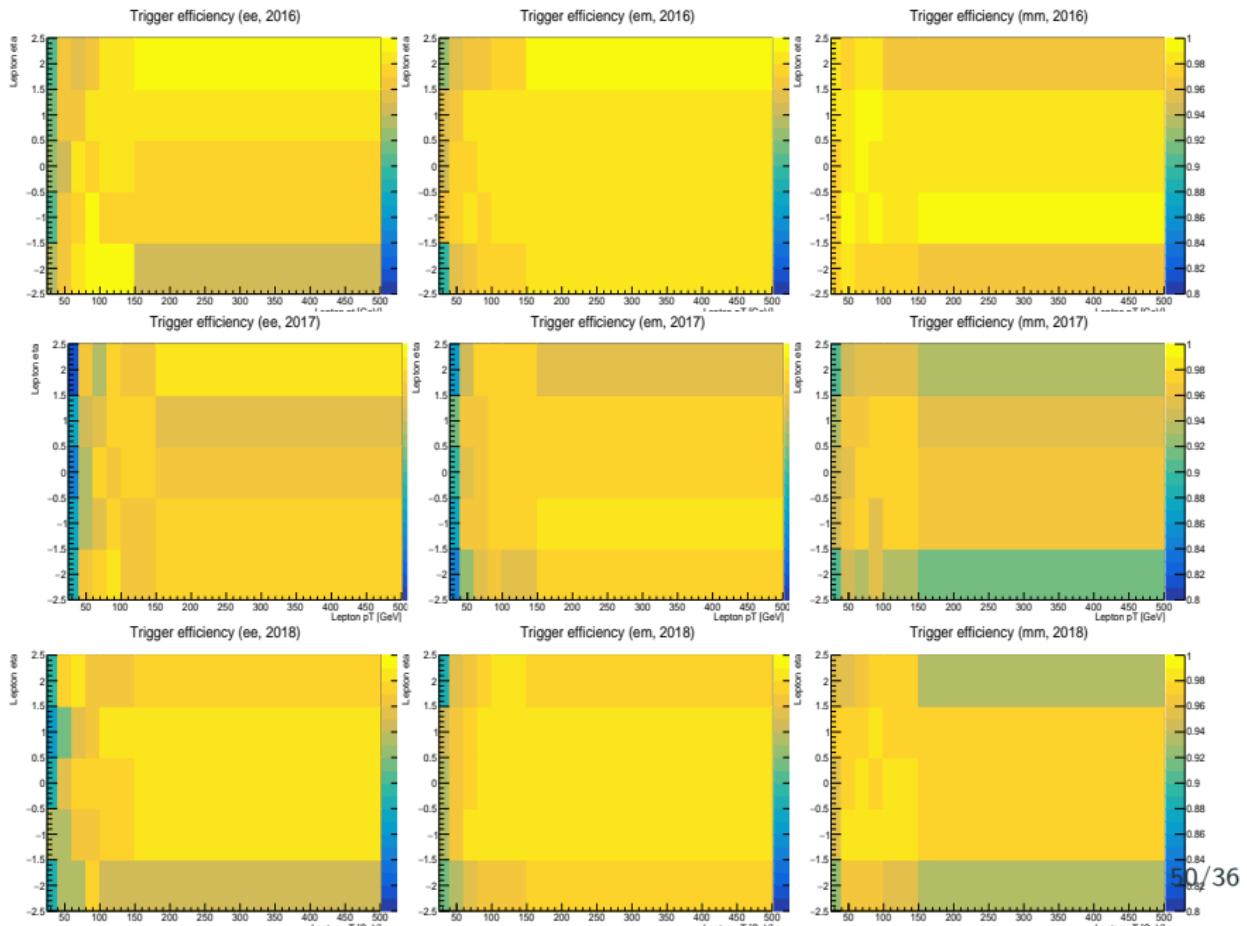
Dataset	Run range	HLT trigger path
SingleMu	[297020,306462]	HLT_IsoMu27_v*
SingleEle	[297020,306462]	HLT_Ele35_WPTight_Gsf_v*
DoubleEG	[297020,306462]	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*
DoubleMu	[297020,299336] [299337,306462]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v*
MuonEG	[297020,306462] [297020,299336] [299337,306462]	HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*

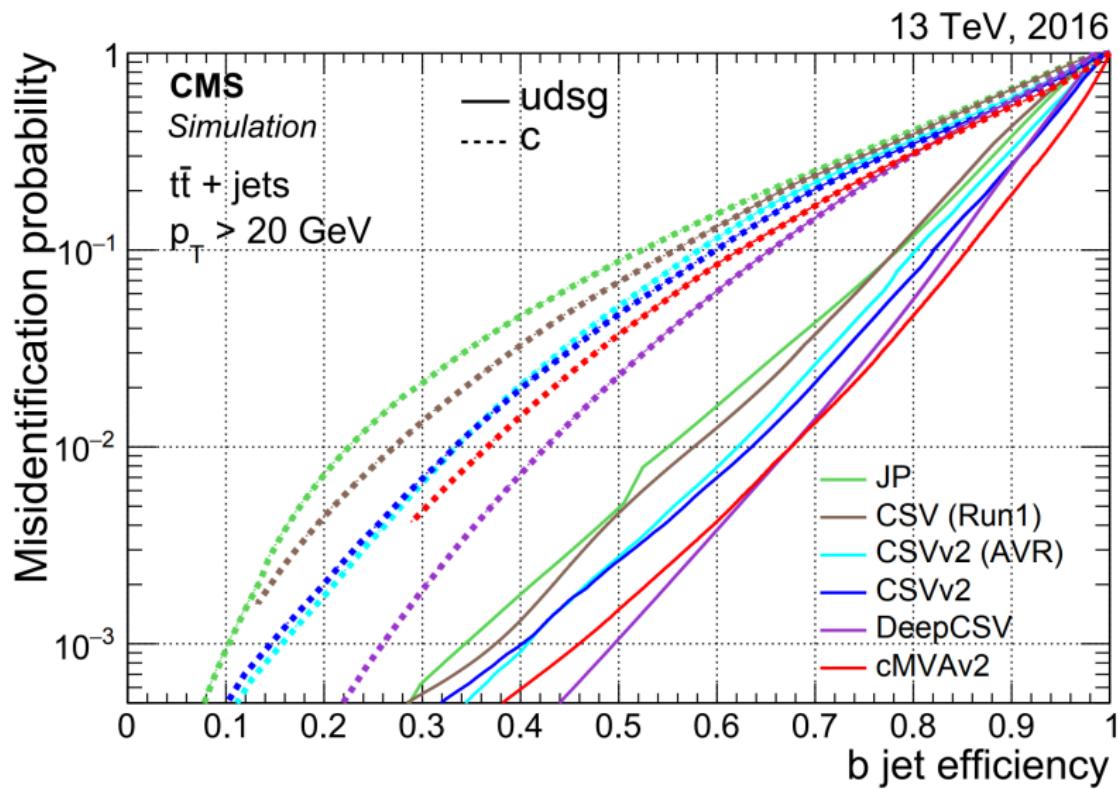
Dataset	Run range	HLT trigger path
SingleMu	[315252,325175]	HLT_IsoMu24_v*
SingleEle	[315252,325175]	HLT_Ele32_WPTight_Gsf_v*
DoubleEG	[315252,325175]	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*
DoubleMu	[315252,325175]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*
MuonEG	[315252,325175]	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
		HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*

Trigger efficiencies computed using orthogonal MET datasets.



2D trigger efficiencies





MET filters

Filter name	Applied to data	Applied to simulation
Flag_goodVertices	✓	✓
Flag_globalSuperTightHalo2016Filter	✓	✓
Flag_HBHENoiseFilter	✓	✓
Flag_HBHENoiselsoFilter	✓	✓
Flag_EcalDeadCellTriggerPrimitiveFilter	✓	✓
Flag_BadPFMuonFilter	✓	✓
Flag_ecalBadCalibFilterV2 [†]	✓	✓

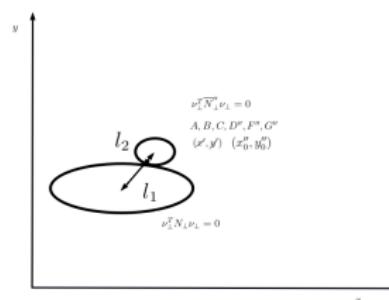
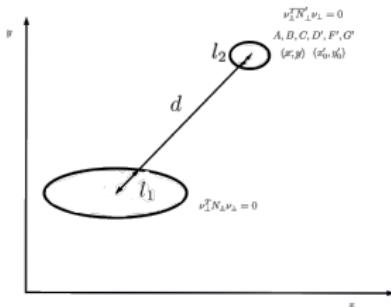
[†] applied only to 2017 and 2018.

Top reconstruction I

Many discriminating variables for our signals require knowledge of the top quark and anti-quark 4-momenta, only available after a **complete reconstruction of the $t\bar{t}$ system**.

The **Batchchart analytical method** was used, relying on a geometrical approach to analytically solve equations constraining the decay of top quarks involving leptons:

- Invariant mass constraints from the top quark and W boson imply that the solution set for each neutrino 4-momentum can be constrained to an ellipse;
- A typical SM $t\bar{t}$ event will give two ellipses close to each other, while ellipses coming from a typical signal event will be further apart and will not intersect.



Two discriminating variables naturally arise from this reconstruction: the **dark p_T** , defined as the distance between the centroids of the ellipses, and the **overlapping factor R** , defined as the ratio between the size of the ellipses and the distance between them.

In practice, **several complications** were taken into account:

- 0, 2 or 4 intersections can be observed while only 1 is physical, therefore considered to be the solution with the lowest possible invariant mass for the $t\bar{t}$ system;
- All the leptons and (b-)jets combinations of the event were taken into account;
- A **smearing process** was followed to evaluate the impact that imperfectly measured kinematic variables can have on this process (100 iterations per event, by updating several parameters such as the energy/direction of the jets and leptons within their respective uncertainties).

These considerations allowed to **increase the mean top reconstruction efficiency from around 70 to 90%**. Non-physical default values are set to all the variables which depend on the value of the neutrino momenta for the events which still fail this reconstruction.

m_{bl}^t variable

If a b-jet is produced in a top-quark decay, its invariant mass is bounded from above by $\sqrt{m_t^2 - m_W^2} = 153$ GeV. Events compatible with two semileptonic top-quark decays can then be selected or rejected by introducing the observable m_{bl}^t :

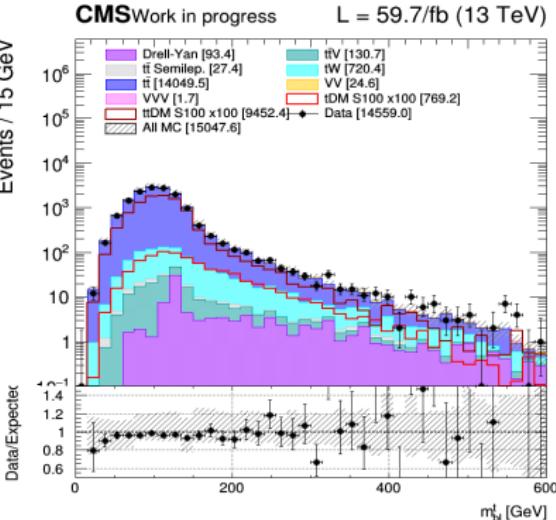
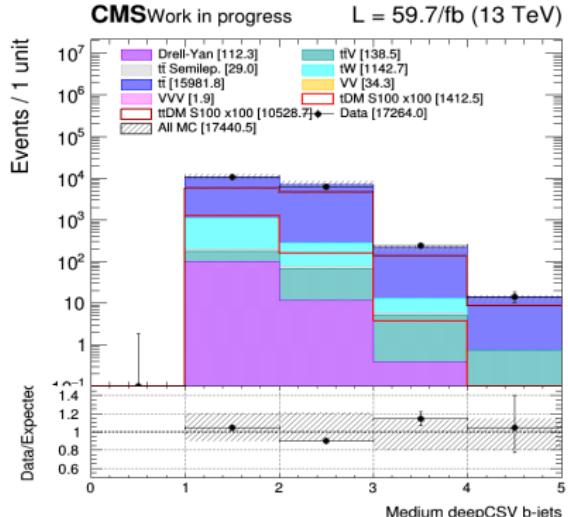
$$m_{bl}^t = \min (\max(m_{l_1 j_a}, m_{l_2 j_b}))$$

In this equation, the minimization is performed either:

- Over all the possible combinations of jets j_a, j_b among the b-jets of the events if three or more j-bets are observed;
- Or over the b-jet(s) observed plus the non b-tagged jet having the highest b-tag weight of the event.

This variable is expected to **give some discrimination** between our two signals of interest.

Number of b-jets and m_{bl}^t variable



Stransverse mass M_{T2}^{ll} and M_{T2}^{bl}

Extension of the transverse mass m_T to cases when pairs of same flavor particles decay into one visible and one invisible particle, such as the double $W \rightarrow l\nu$ decay.

Here, 2 neutrinos contribute to the presence of ptmiss and the individual contribution of each particle ($\cancel{p}_{T1}, \cancel{p}_{T2}$) to this missing energy cannot be inferred. M_{T2}^{ll} is defined as:

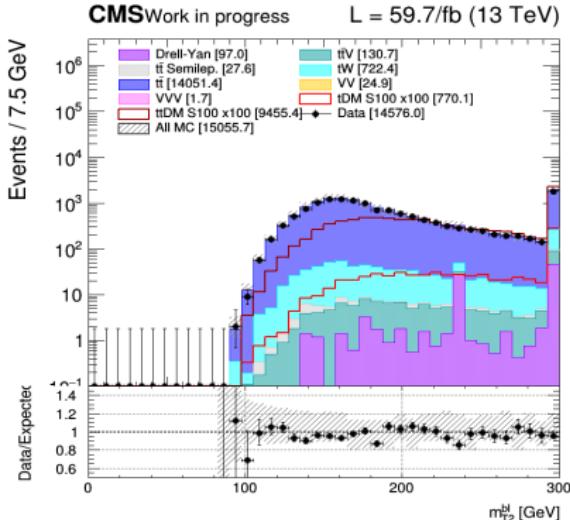
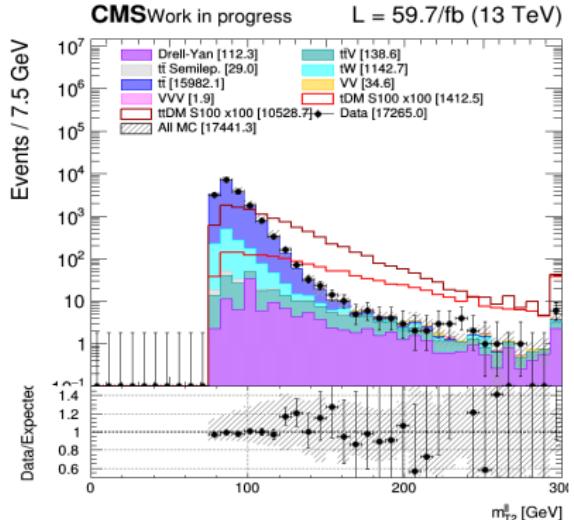
$$\begin{cases} M_{T2}^2 = \min_{\cancel{p}_{T1} + \cancel{p}_{T2} = \cancel{p}_{T\text{tot}}} \left(\max \left(m_T^2(p_{T1}, \cancel{p}_{T1}), m_T^2(p_{T2}, \cancel{p}_{T2}) \right) \right) \\ m_T^2(p_T, \cancel{p}_T) = 4 |p_T| |\cancel{p}_T| \sin^2 \left(\frac{\alpha}{2} \right) \end{cases}$$

Different combinations ($\cancel{p}_{T1}, \cancel{p}_{T2}$) satisfying the condition $\cancel{p}_{T1} + \cancel{p}_{T2} = \cancel{p}_{T\text{tot}}$ then need to be probed, keeping only the combination which results in the lowest possible value.

The $t\bar{t}$ process is expected to have an endpoint exactly at the mass of the W boson, while our eventual signal does not have this limitation because of the pair of dark matter particles produced.

The M_{T2}^{bl} variable is defined in a similar case, except that in this case, the lepton is paired with a b-jet. The jet/lepton permutation giving the smallest value is kept.

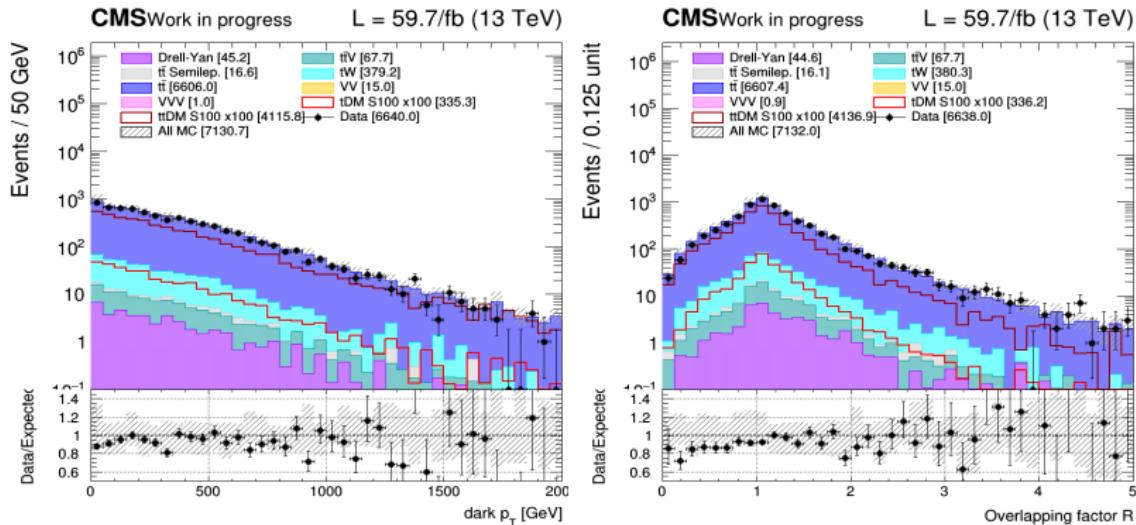
Transverse mass M_{T2}^{ll} and M_{T2}^{bl}



Dark p_T and overlapping factor R

Two discriminating variables naturally arise from the top reconstruction:

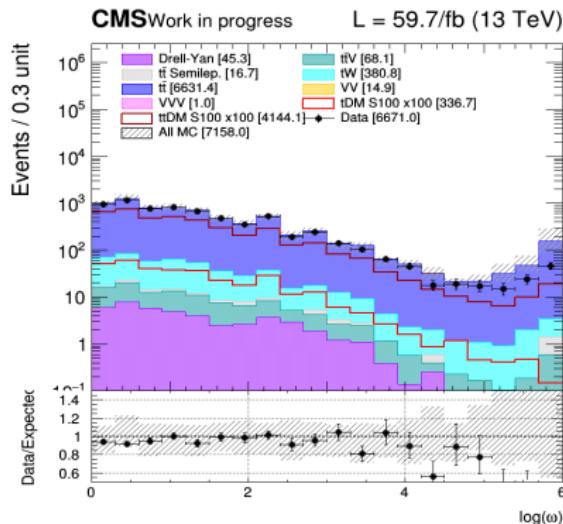
- The **dark p_T** , defined as the distance between the centroids of the ellipses;
- And the **overlapping factor R** , defined as the ratio between the size of the ellipses and the distance between them.



Top reconstruction weight W

The top reconstruction smearing process does introduce a new discriminating variable: the **top reconstruction weight W** .

Indeed, in order to know which lepton/b-jets combination and which smearing iteration performs the best, a weight is assigned to each iteration by comparing the invariant mass obtained m_{lb} with the expected distribution using generation. The combination with the largest weight is then simply considered as the solution of the event.

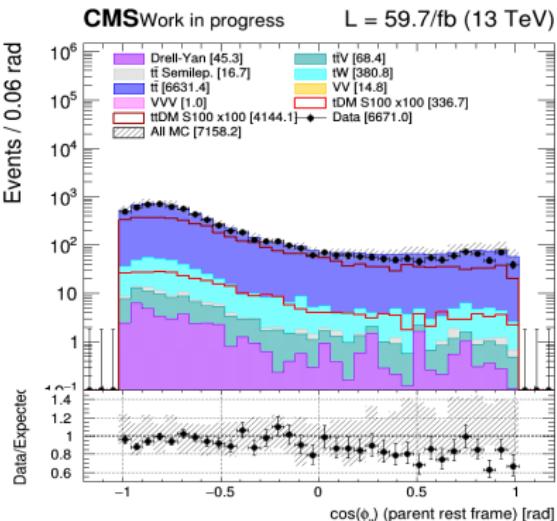
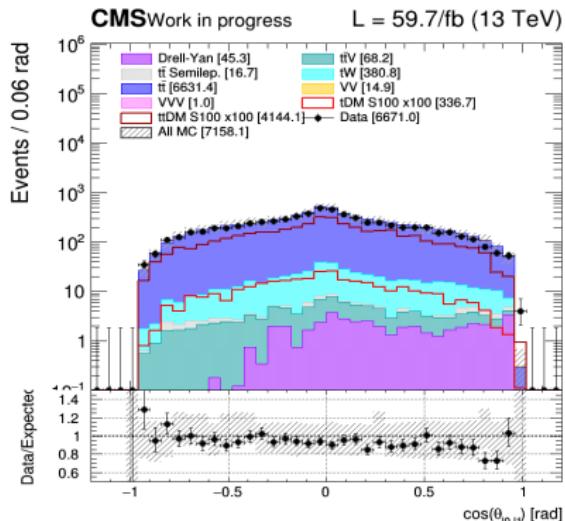


Spin correlated variables

The spin correlation in a $t\bar{t}$ -like event is expected to be conserved, because of the short lifetime of the top quark, and can be inferred from the top quark decay products.

Such variables are interesting because **the spin correlation depend on the production mechanism** and will be influenced by the additional coupling to a scalar or pseudoscalar mediator, making this a perfect candidate to be good discriminating variables:

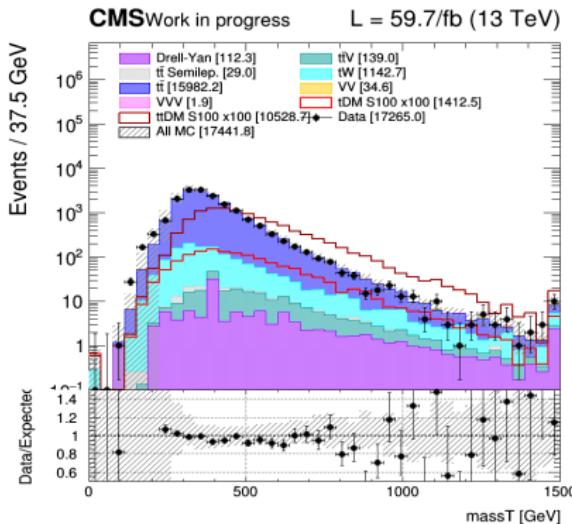
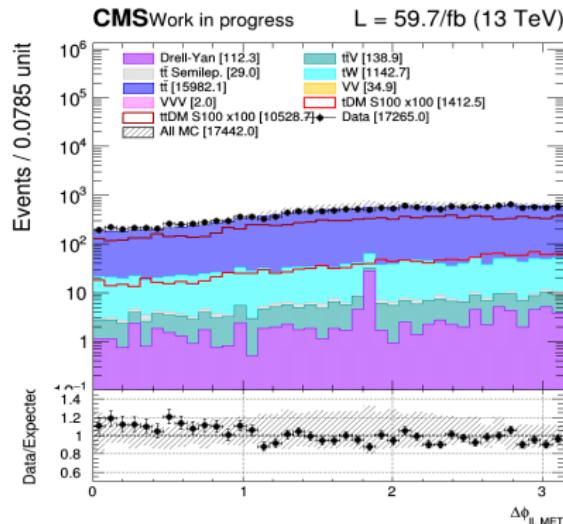
- $\xi = \cos(\theta_i) \cos(\theta_j)$, where i and j are either leptons, b-jets or neutrinos;
- $\cos(\Phi_{i,j})$, the cosine of the full opening angle of such top decay products in their respective parent rest frames.



Other discriminating variables I

Several other variables were considered for their discriminating power:

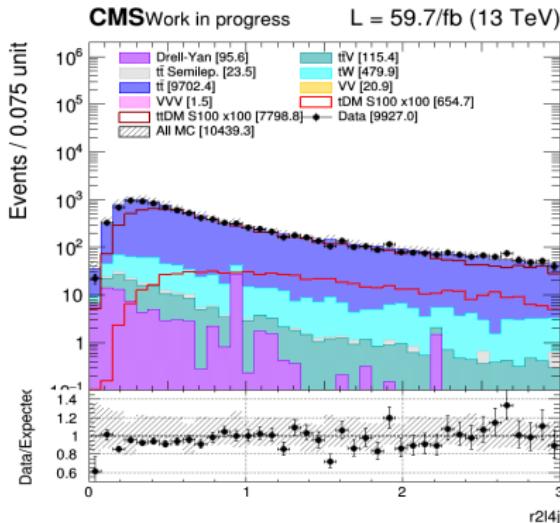
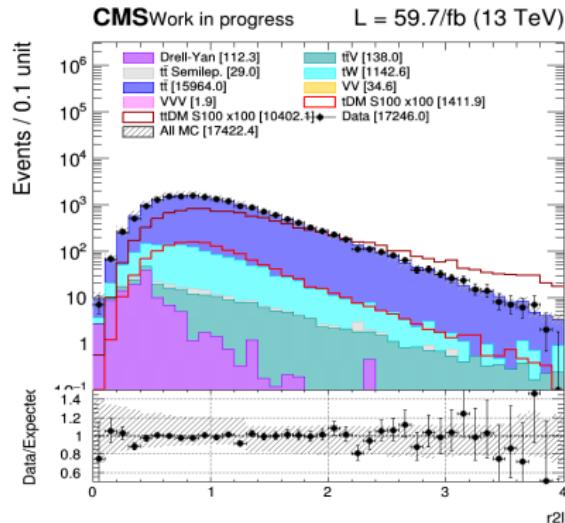
- $\Delta\Phi(E_T^{\text{miss}}, \text{II})$: the distribution in Φ of the two leptons is expected to change depending on the eventual production of DM;
- massT, which corresponds to the scalar sum of the transverse component of the MET, the two leptons and the two b-jets obtained by the top reconstruction process also helps with the discrimination process.



Other discriminating variables II

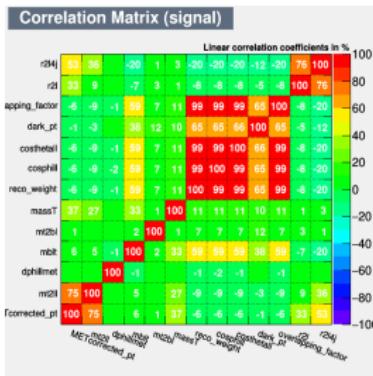
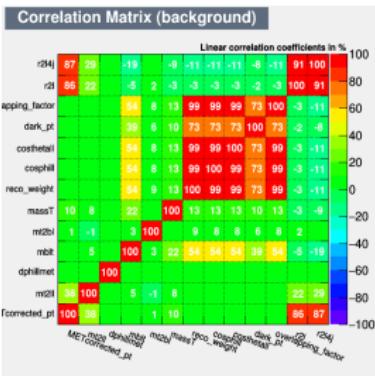
Two other interesting variables used by the ATLAS collaboration for their own analysis:

- $r2l$, defined as the ratio between the ptmiss and the p_T of the two leptons observed;
- And $r2l4j$, defined in a similar way but considering additionally the p_T of the first 4 jets (if they exist) in the sum in the denominator.

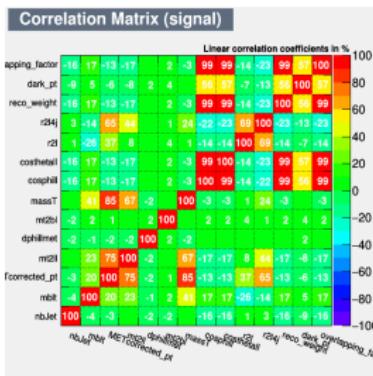
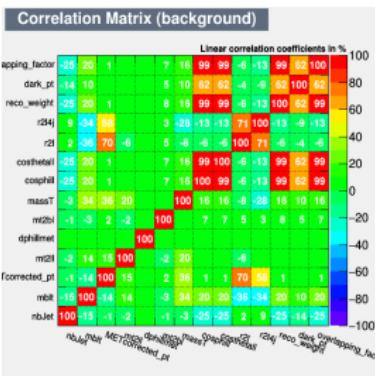


Input variables correlation

t/\bar{t} +DM region



$t\bar{t}$ +DM region

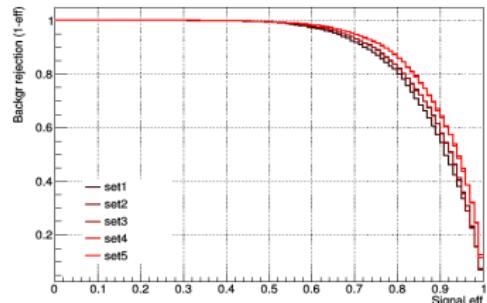


Input variables

BDT ROC curves

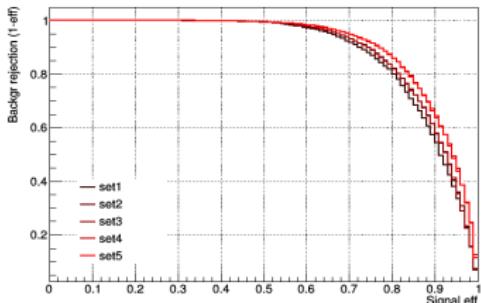
$t/\bar{t}+DM$ region

ROC curves for several trainings



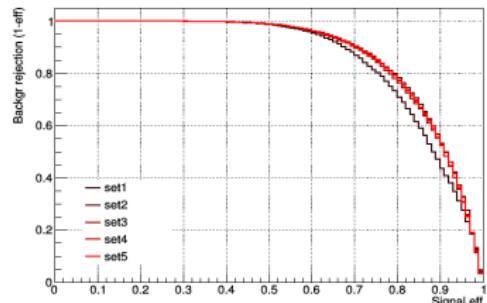
$t\bar{t}+DM$ region

ROC curves for several trainings

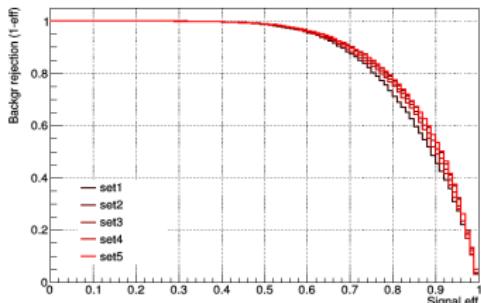


ANN ROC curves

ROC curves for several trainings



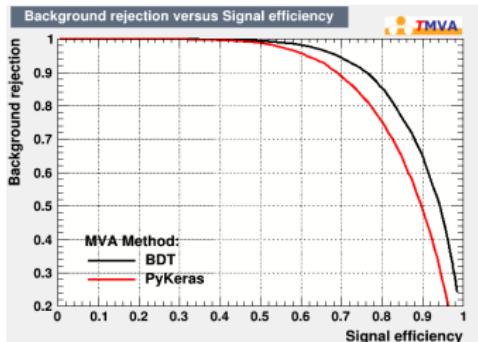
ROC curves for several trainings



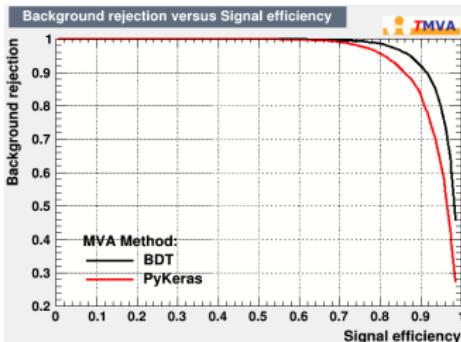
BDT vs ANN (t/\bar{t} +DM region)

Scalar mediators

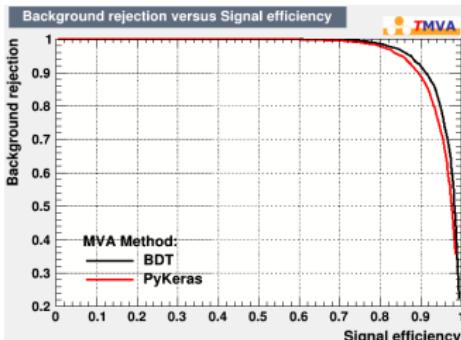
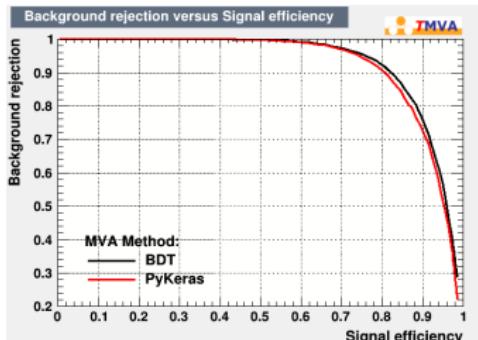
100 GeV



500 GeV



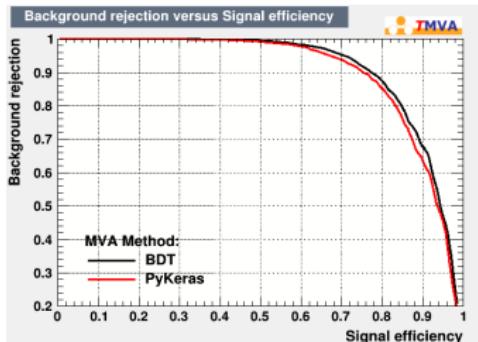
Pseudoscalar mediators



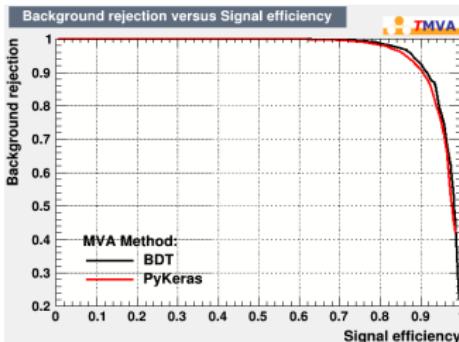
BDT vs ANN ($t\bar{t}$ +DM region)

Scalar mediators

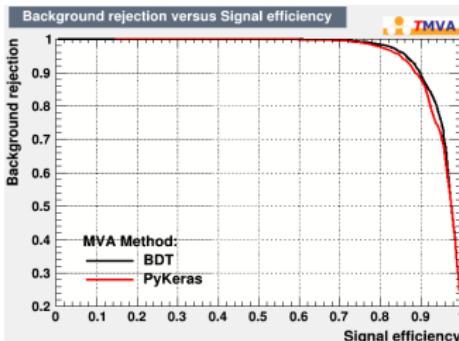
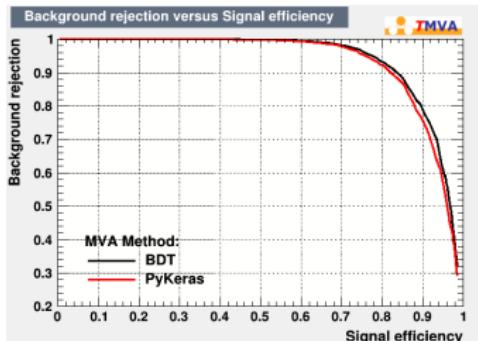
100 GeV



500 GeV



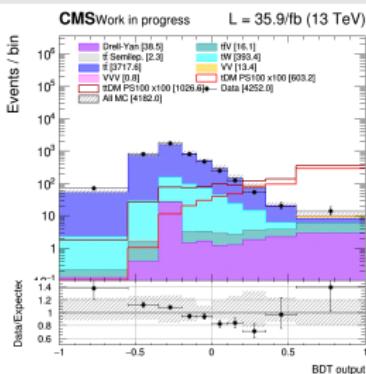
Pseudoscalar mediators



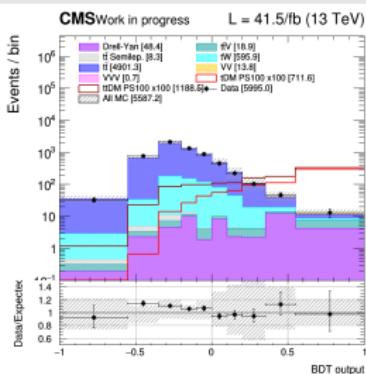
BDT pseudoscalar 100 GeV output shape

t/\bar{t} +DM region

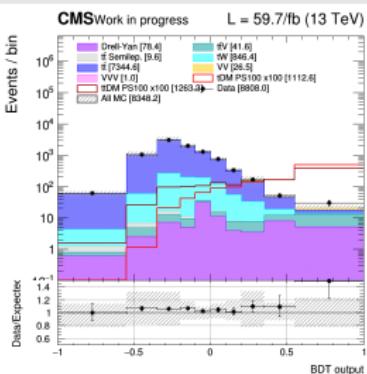
2016



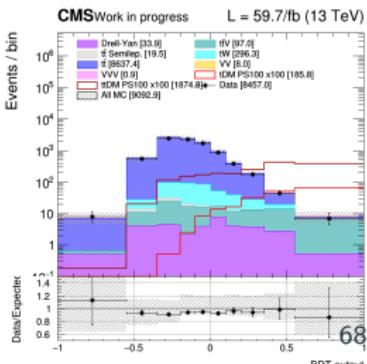
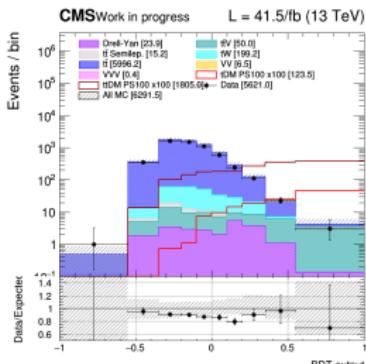
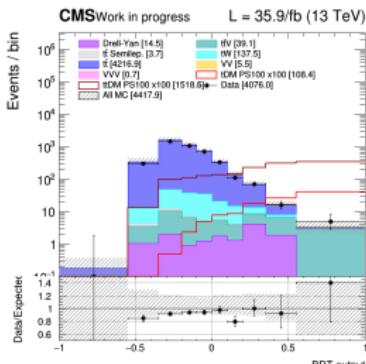
2017



2018



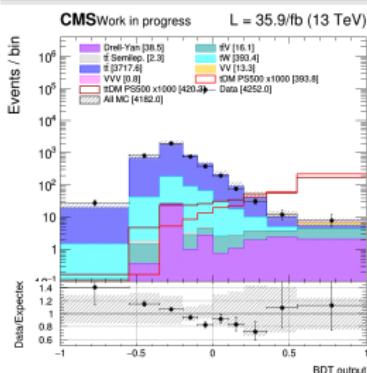
$t\bar{t}$ +DM region



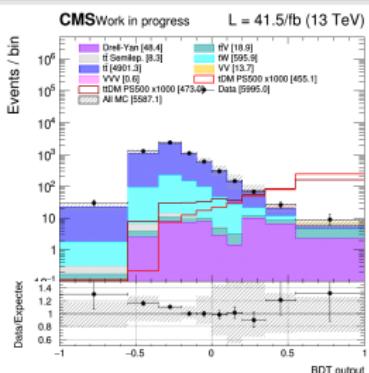
BDT pseudoscalar 500 GeV output shape

$t/\bar{t} + \text{DM}$ region

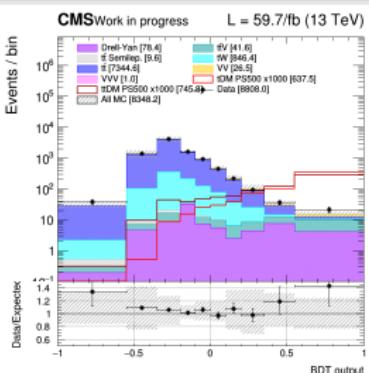
2016



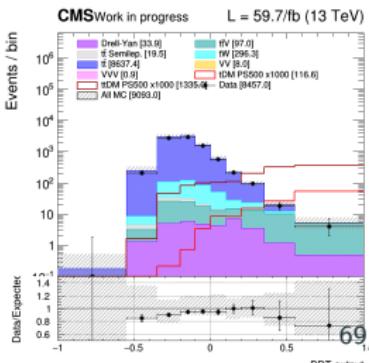
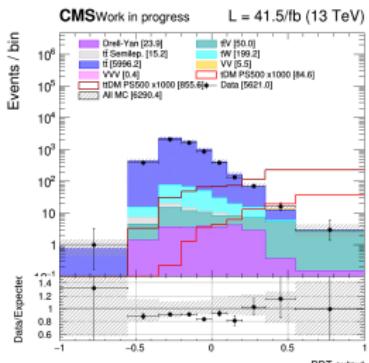
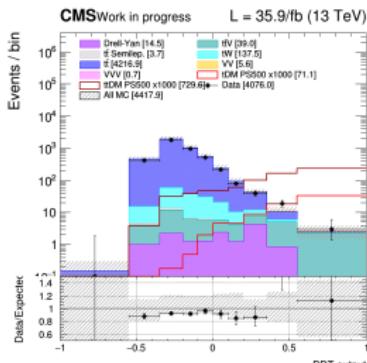
2017



2018



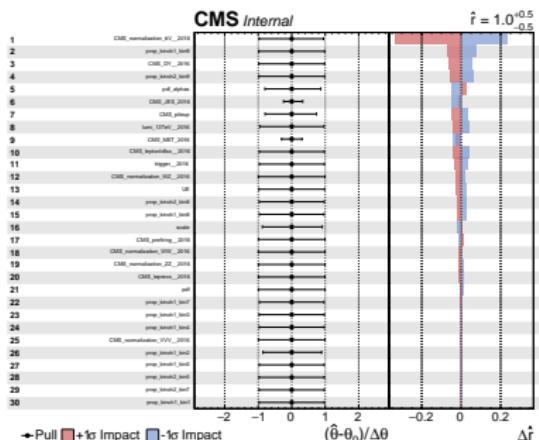
$t\bar{t} + \text{DM}$ region



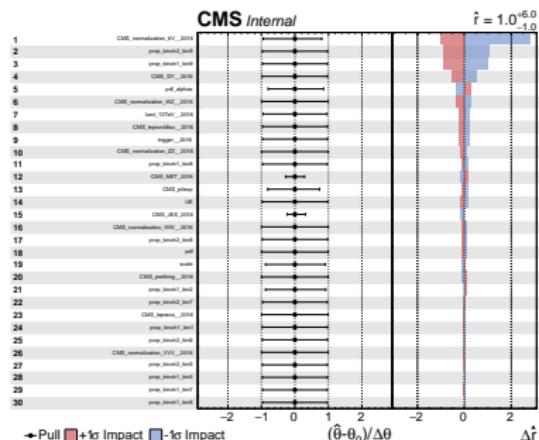
Pulls and impact plots

2016, pseudoscalar

100 GeV



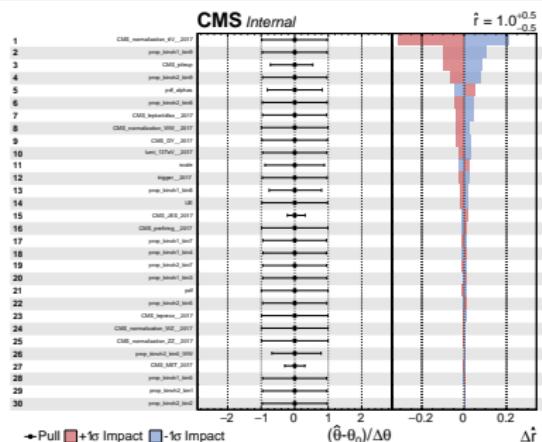
500 GeV



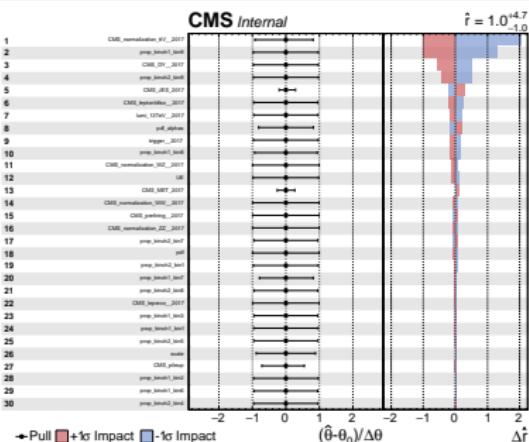
Pulls and impact plots

2017, scalar

100 GeV



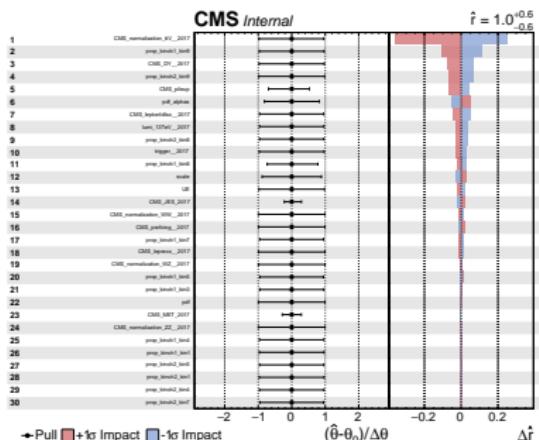
500 GeV



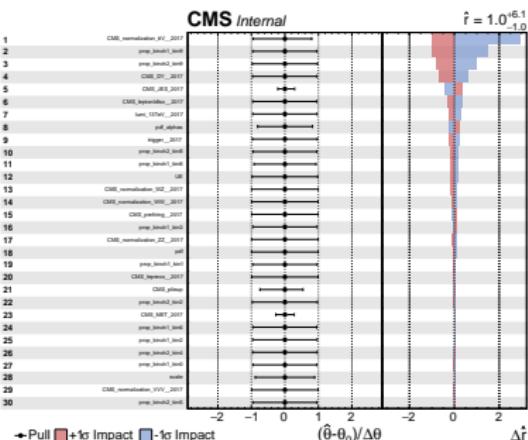
Pulls and impact plots

2017, pseudoscalar

100 GeV



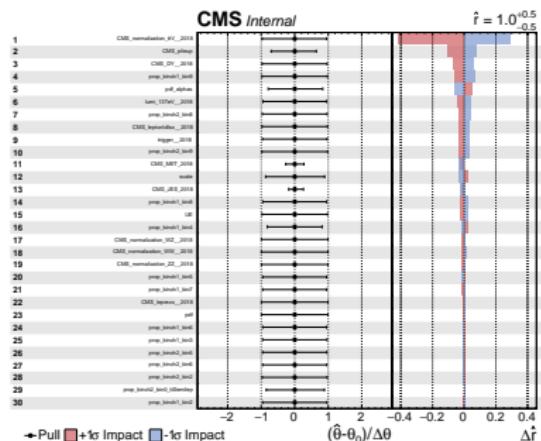
500 GeV



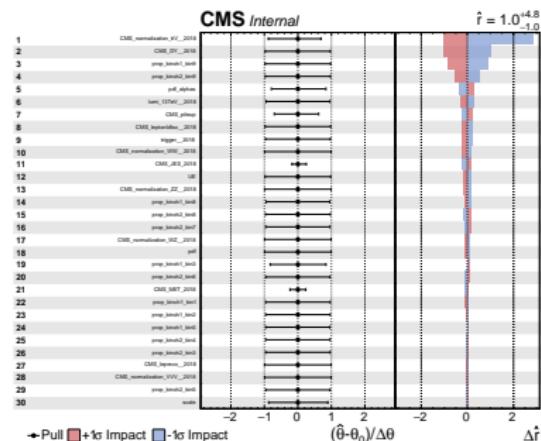
Pulls and impact plots

2018, scalar

100 GeV



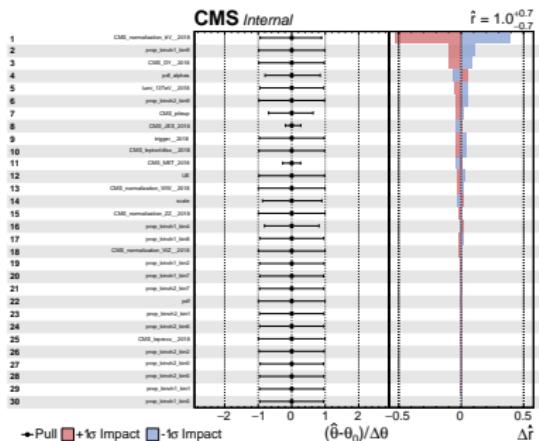
500 GeV



Pulls and impact plots

2018, pseudoscalar

100 GeV



500 GeV

