

Search for dark matter production in association with top quarks in the dilepton final state at $\sqrt{s} = 13$ TeV

Pablo Martínez Ruíz del Árbol, Jónatan Piedra Gomez, Cédric Prieëls
February 22th 2021

Instituto de Física de Cantabria

Introduction

A search for the production of dark matter particles in association with either one or two top quarks is presented:

- We study the pp collisions produced by the LHC at $\sqrt{s} = 13$ TeV;
- Data collected by the CMS detector;
- Legacy analysis, considering the full Run II dataset (data collected in 2016, 2017 and 2018 and summing around 137 fb^{-1}).

Motivation

- Several (mostly astrophysical) evidences for the existence of dark matter, however **its nature remains unknown** and it has never been detected experimentally;
- If dark matter is made of some kind of particle it might be produced in the high energy collisions.

Main objective

- Consider different dark matter production models to discover or eventually exclude some of them, or **put upper limits on their cross section of production**.

The dark matter case

At the origins of dark matter I

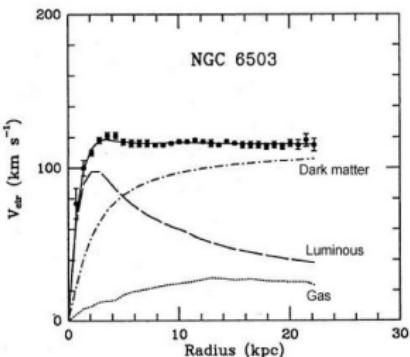
The concept of dark matter can be traced back to the 19th century, and was introduced to **explain several astrophysical evidences**, among which:

Zwicky's calculations

- Measurement of the mass of the Coma Cluster using the virial theorem;
- Concluded that its mass was **400-500 times larger** than the value obtained by Hubble, considering only visible galaxies.

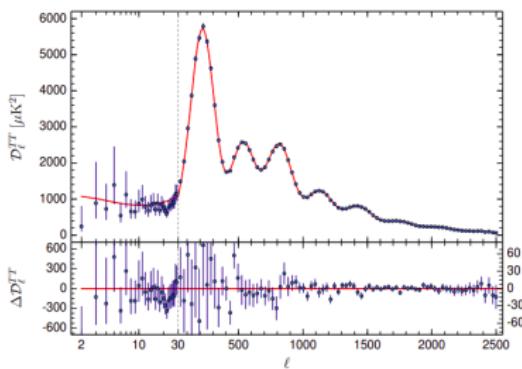
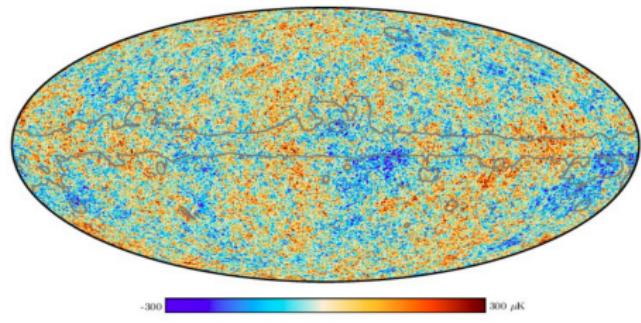
Spiral galaxies rotation curves

- Stars within spiral galaxies should rotate with a velocity depending on the radius to the galactic center, but **this is not what is observed experimentally**.
- Either our understanding of gravity at large scales or our basic understanding of galaxies as a celestial body made of stars has to be revised.



CMB anisotropies

- Background of primary radio waves emitted when the Universe became transparent around 380 000 years after the Big Bang;
- Can be considered as emitting a black body spectrum with a temperature of $(2.72548 \pm 0.00057)\text{K}$, but small anisotropies at the 10^{-5} level are observed.
- Implies that dark matter **accounts for $\sim 27\%$ of the total mass of the Universe.**



Other observations, such as the gravitational lensing effect, **also tend to further support the existence of dark matter** (cf. backup).

Several fundamental properties of dark matter are nowadays known or assumed:

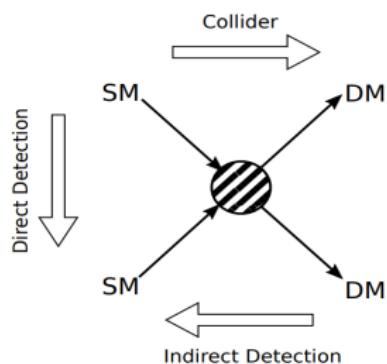
- Dark matter is a particle, given that it is assumed to have a certain mass;
- It should be dark, unable to interact with electromagnetic radiation, otherwise we would have seen it already. It should then also be electrically neutral;
- It is non-baryonic, because the energy density for the baryonic matter estimated from the CMB is too low to account for dark matter;
- We only consider cold dark matter since the widely accepted Λ_{CDM} model is based on this assumption and this helps explaining the presence of large scale structures in the Universe;
- It should have a mass in the electroweak scale, between 10 GeV and 1 TeV, because of the relic density obtained from the thermal freeze-out mechanism.
- Finally, it should be long-lived, since we expect them to have been produced during the Big Bang and they are still present in the Universe.

Weakly Interactive Massive Particles

The WIMPs are the dark matter candidates considered in this work, because of the so-called **WIMP miracle**. Indeed, they:

- Are expected to interact very weakly with ordinary baryonic matter;
- Have a mass in the 100 GeV-1 TeV range for reasonable electroweak production cross-section values;
- Give us a dark matter candidate while being able to solve the **hierarchy problem**.

Main search strategies



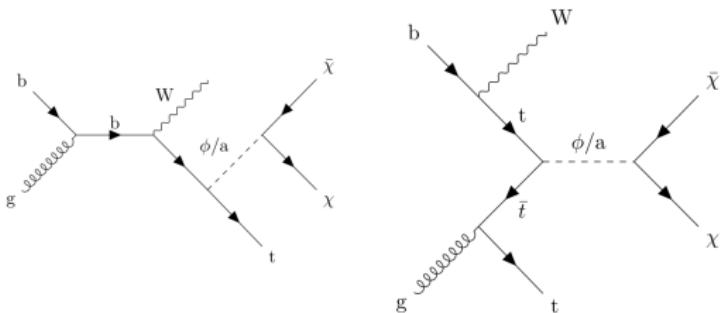
Different strategies are used:

- The **direct and indirect searches**, relying on the production of baryonic matter from the interaction between two DM particles or on the observation of the interaction between the dark and baryonic sectors;
- And the **collider production**, able to probe lower dark matter candidates masses.

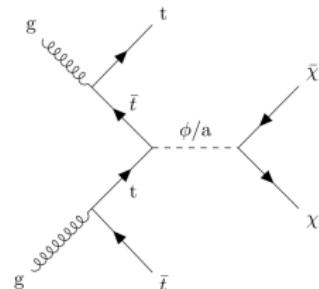
We are searching for **dark matter produced in association with either one or two top quarks**. Several **simplified models** are interesting to consider:

- 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)

$t/\bar{t}+\text{DM}$ tW models



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



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 MET coming from the dark matter and the decay of the Ws;

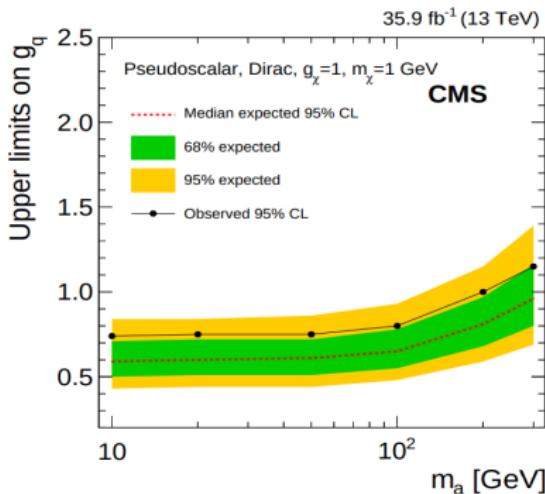
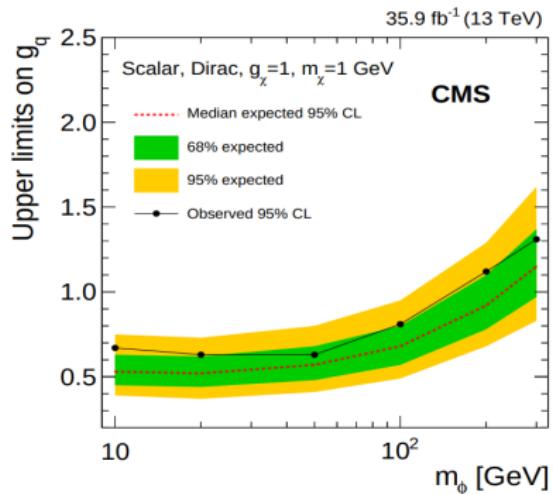
In particular, we are studying the **dilepton final state** in this work:

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

This channel is then **expected to be competitive with the hadronic channel**, especially when considering high mediator masses, which feature a higher global discrimination signal/background.

Previous relevant results I

A similar analysis has already been published by CMS using 2016 data, considering the $t\bar{t}+DM$ signal only and a combination of the three possible final states.



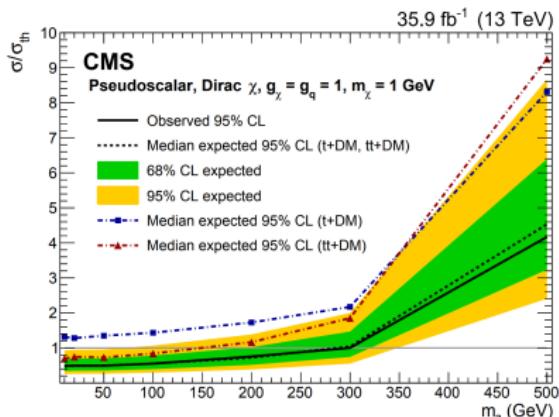
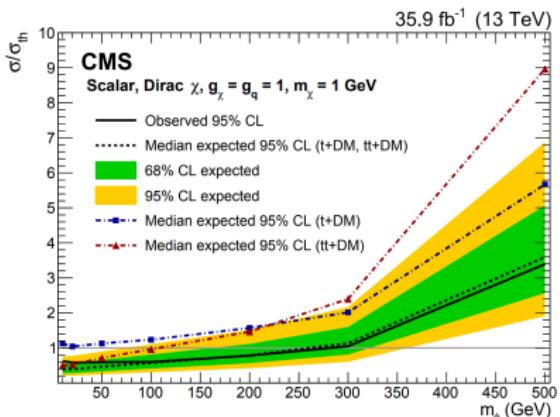
The observed (expected) limits excluded a **pseudoscalar mediator** with mass below 220 (320) GeV, and a **scalar mediator** with mass below 160 (240) GeV.

Previous relevant results II

A combination of both the $t/\bar{t}+DM$ and $t\bar{t}+DM$ processes has also been performed.

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.

Global strategy

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).
→ Paper expected to be approved by LHCP (\sim June).

The effort is **globally common** between the groups 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 $t/\bar{t}+DM/t\bar{t}+DM$ regions

This talk will **be focused on the dilepton final state**, in which we are mostly involved, along with a team of DESY. Deborah Pinna and her team from the University of Wisconsin are focused on the semi-leptonic and hadronic channels.

Object and samples

Objects definition I

Analysis groups have synchronized their frameworks. The following objects are used:

Triggers

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

Leptons

- Analysis relies on the selection of events with two tight leptons;
- Tight Muon POG working point used with a tight muon isolation requirement (< 0.15) for muons, EGamma POG MVA WP90 for electrons;
- Some additional quality cuts (eg., in the d_{xy}/d_z impact parameters) to reduce the impact of the fakes applied on top.

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 $> 98\%$);
- Loose jet PU ID (if $p_T > 50 \text{ GeV}$) applied on top to reject jets coming from PU interactions in 2017 and 2018;
- Corresponding L1 ECAL prefiring corrections applied.

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 $\sim 10\%$);
- B-tagging weight larger than 0.6321, 0.4941 or 0.4184 (2016, 2017 or 2018).

MET

- **PfType1MET** considered by propagating the JECs to the MET;
- All recommended **MET filters applied** to filter anomalous high MET 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.

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 is currently in place, allowing us to only look at 1 fb^{-1} of data per year near the signal regions.

Monte-Carlo

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;
- Others: dibosons, tribosons, non-prompt contamination (data-driven, not MC).

Signal samples

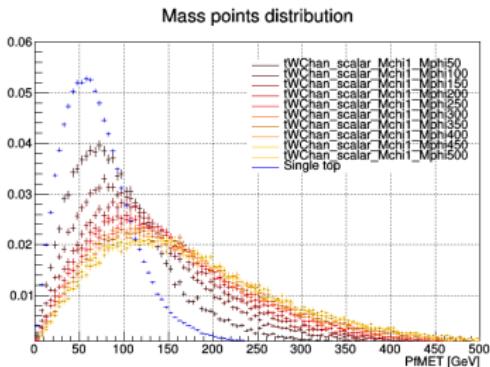
Both signal samples have been generated using MADGRAPH and PYTHIA8 (with the CP5 tune) at LO, while simulated events are then interfaced with a realistic model of the CMS detector using Geant4 [113] and are reconstructed using the official CMS reconstruction algorithms.

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

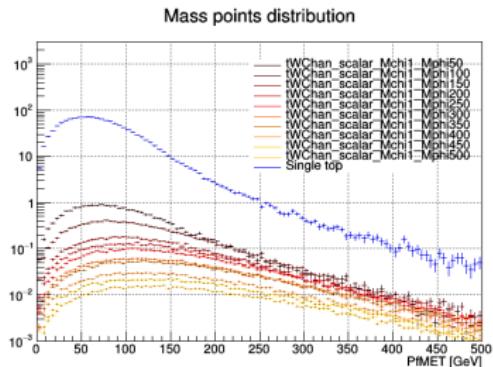
- Both scalar and pseudoscalar mediators are considered;
- 400.000 events were produced for each mediator mass, from 10 to 1000 GeV;
- The dark matter mass was set to 1 GeV, but additional samples ranging from 1 to 55 GeV were also produced;
- All the g_q and g_χ couplings were set to 1.

Scalar mediators

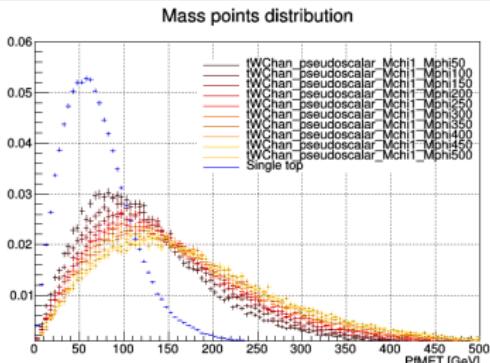
With normalization



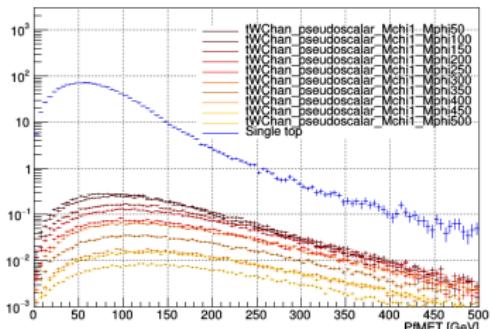
Without normalization



Pseudoscalar mediators

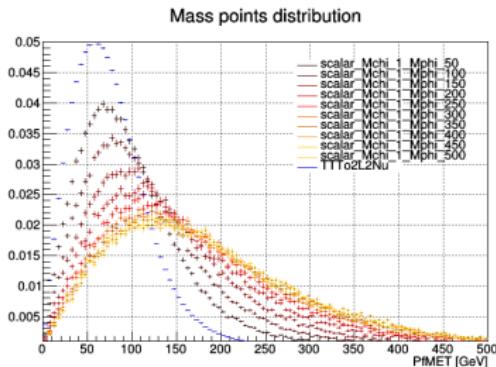


Mass points distribution

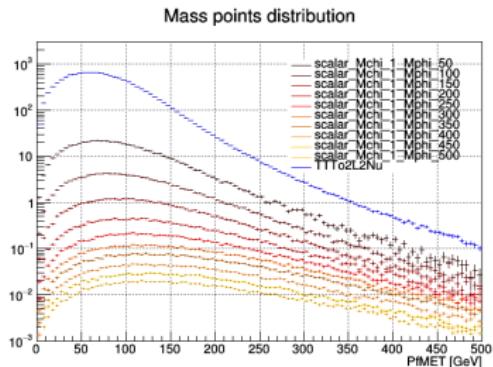


Scalar mediators

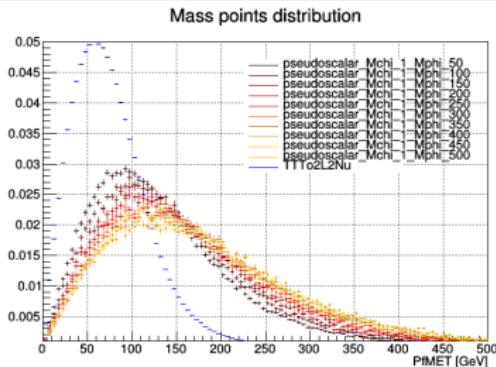
With normalization



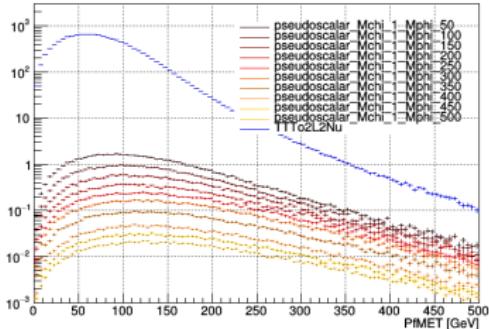
Without normalization



Pseudoscalar mediators



Mass points distribution



Inclusive selection

Inclusive selection

First, distributions in the following inclusive control region mostly enriched in Drell-Yan were studied, in all the different channels available:

- Leading (trailing) lepton $p_T > 25$ (20) GeV
- Third lepton veto ($p_T < 10$ GeV)
- Opposite sign leptons
- $m_{\parallel} > 20$ GeV to avoid low mass resonances
- At least 1 jet
- At least 1 medium deep CSV b-jet

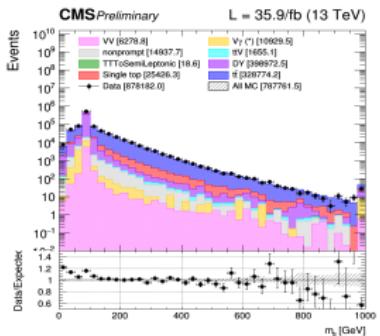
This region is studied to have a look at the most inclusive selection to spot initial issues and problems.

A slight mismodeling of the low mass DY+Jets sample and a general problem of the MET in DY events is observed, but these features are known in CMS and are mitigated with the analysis cuts applied.

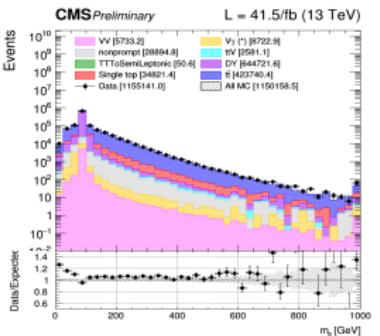
Inclusive control region

// channel

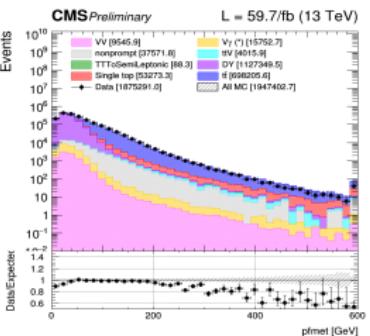
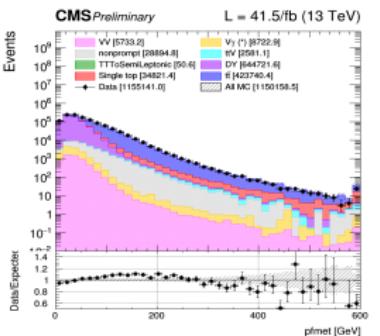
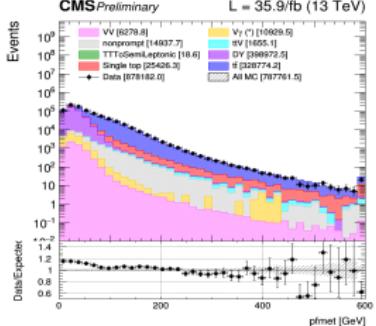
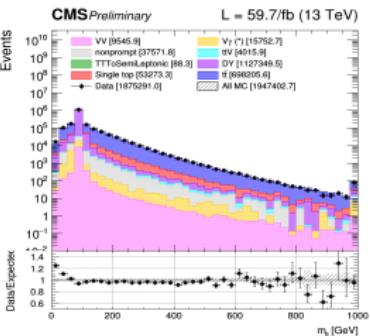
2016



2017

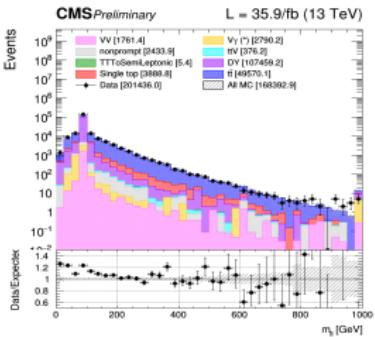


2018

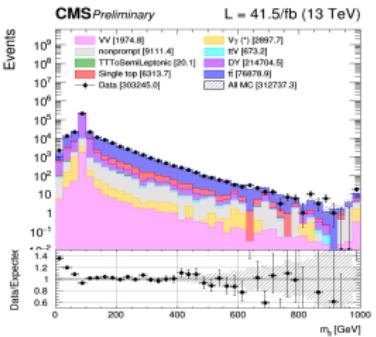


Inclusive control region

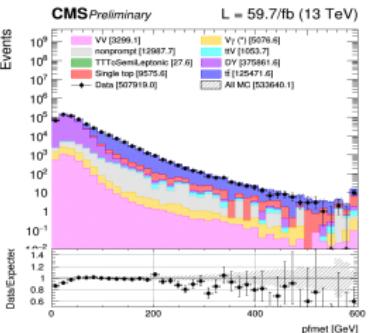
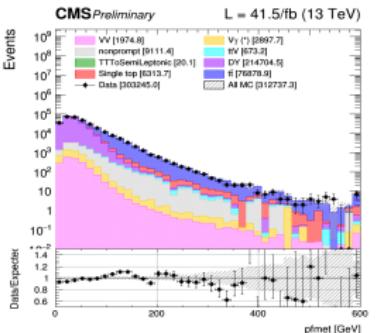
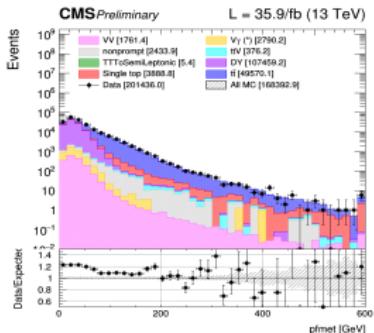
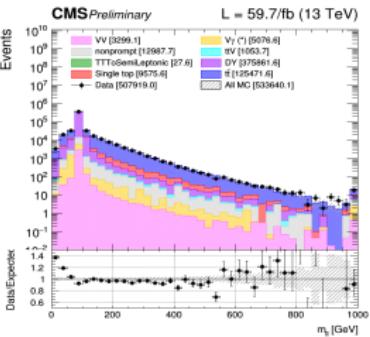
2016



2017



2018



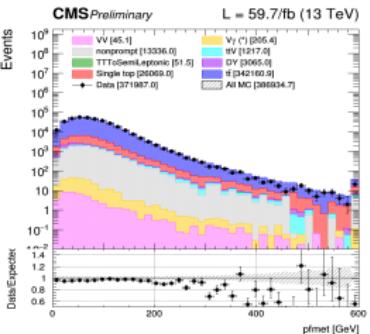
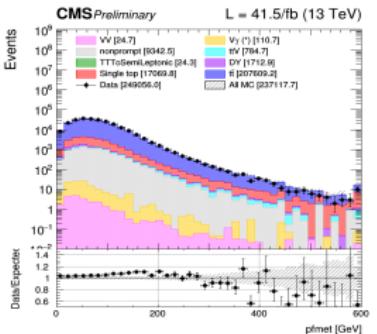
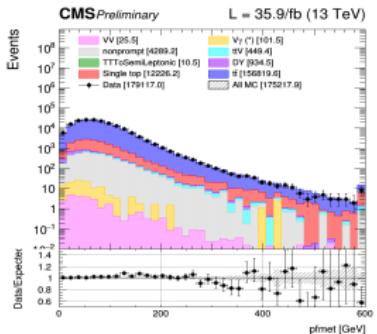
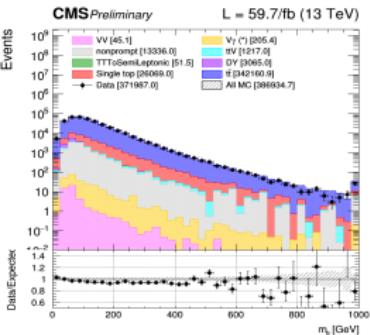
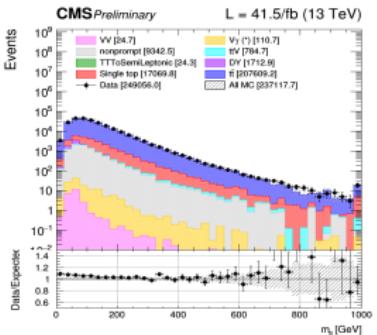
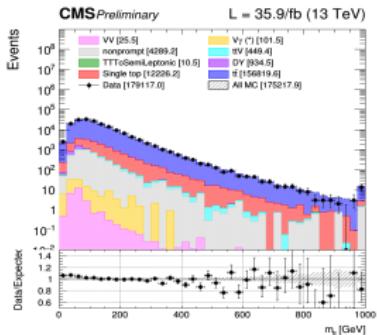
Inclusive control region

df channel

2016

2017

2018



Signal extraction

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 **following pre-selection** are considered for the training:
 - 2 tight leptons: $p_T > 25, 20$ GeV
 - Third lepton: $p_T < 10$ GeV
 - Opposite sign leptons
 - $m_{ll} > 20$ GeV
 - 15 GeV Z-veto in ee and $\mu\mu$ channels
 - At least 1 jet
 - At least 1 b-jet
 - $M_{T2}^H > 80$ GeV, to stay orthogonal to the other channels
- One specific training performed per signal mass point;
- 50% train/test splitting used (~ 40.000 training events in total).

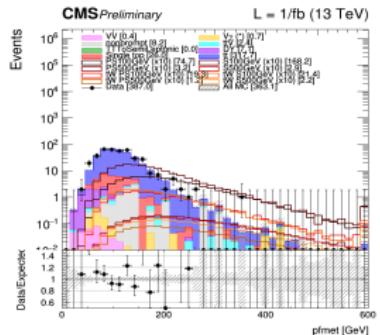
At the end of the day though, the **ANN was chosen for the analysis** over the BDT, given that it gave $\sim 10\%$ better upper limits once optimized. The ANN shape is then used to perform a general **shape analysis**.

The TMVA package was then finally used to study the training performed, as shown in the next few slides for the 2016 scalar 100/500GeV training performed. We plan on combining all the year together.

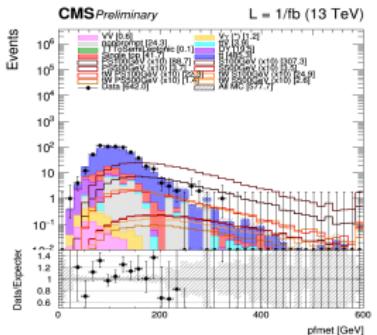
Pre-selection region

// channel

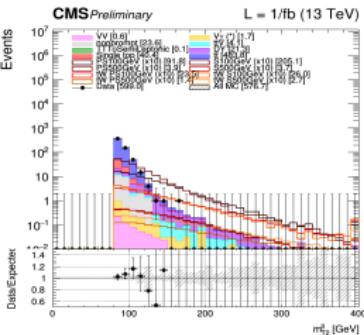
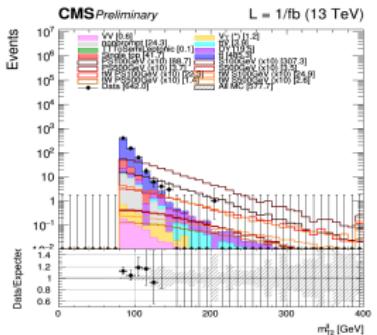
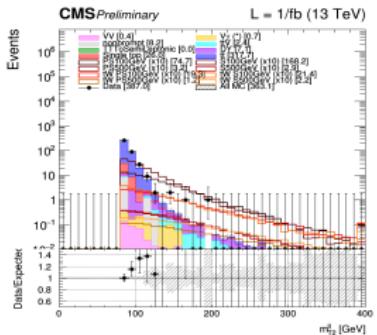
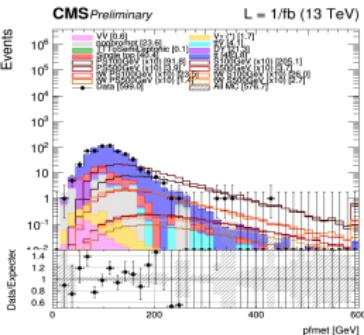
2016



2017



2018



Hyperparameters optimization

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

DNN parameter	Optimized value
Hidden layers neurons	80, 80, 40
Activation functions	Relu (x3), softmax (output)
Error function	Mean square error
Optimizer	Adam
Learning rate	0.005
Training epochs	250
Batch size	250

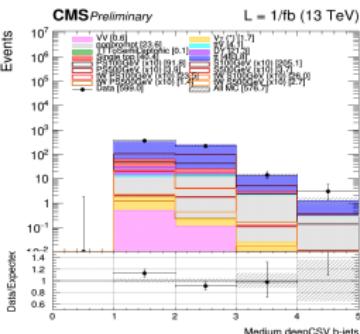
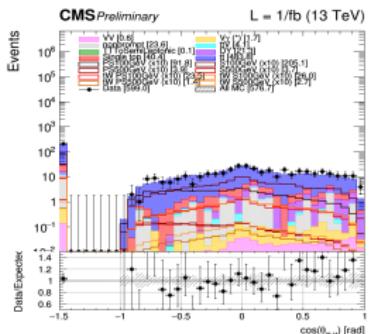
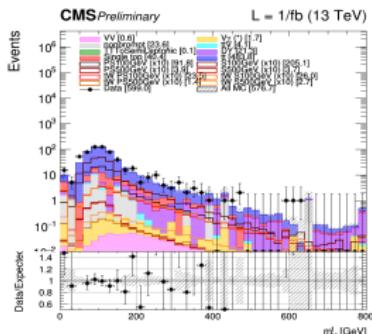
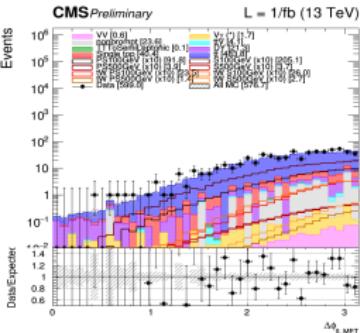
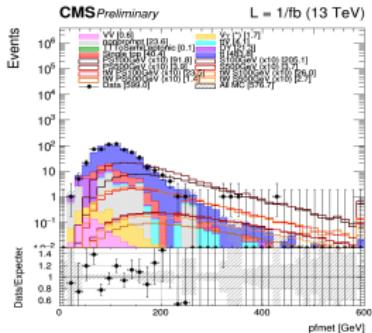
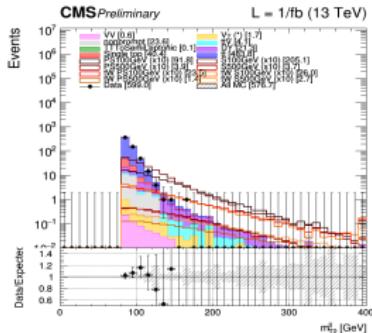
At the end of the day, the ANN using the following input variables was chosen:

Rank	Variable	Importance
1	$M_{T_2}^{II}$	$2.30 \cdot 10^{-1}$
2	pfMET	$1.92 \cdot 10^{-1}$
3	$\Delta\Phi(E_T^{\text{miss}}, II)$	$1.67 \cdot 10^{-1}$
4	m_{bl}^t	$1.38 \cdot 10^{-1}$
5	$\cos(\theta_i) \cos(\theta_j)$	$1.35 \cdot 10^{-1}$
6	nbJet	$6.95 \cdot 10^{-2}$

2018 discriminating variables

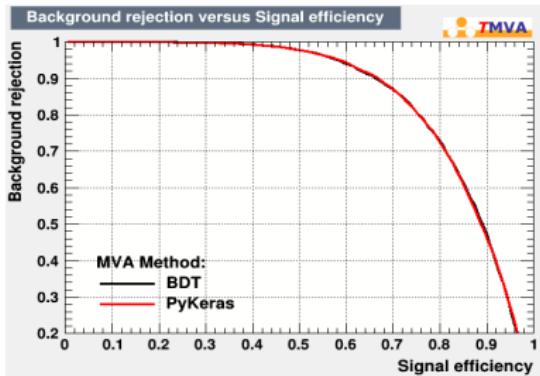
Example distributions shown in the blinded 2018 pre-selection region.

2018, // channel

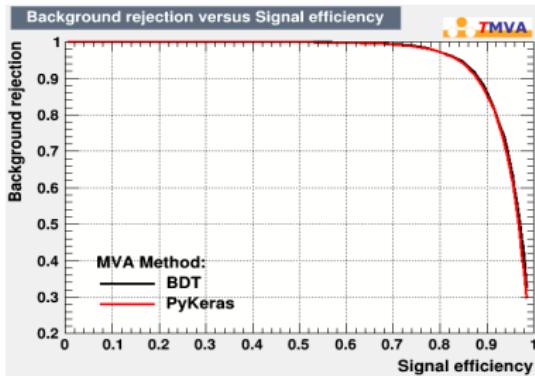


Scalar mediators

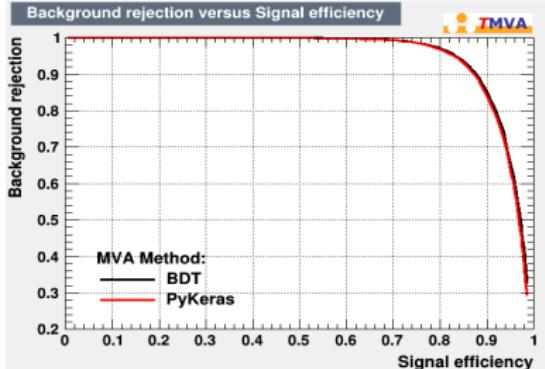
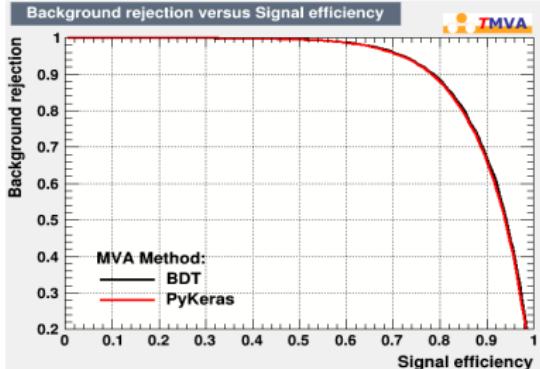
100 GeV



500 GeV



Pseudoscalar mediators

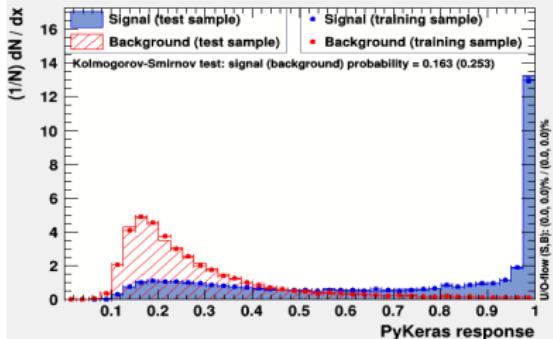


Overtraining plots

Scalar mediators

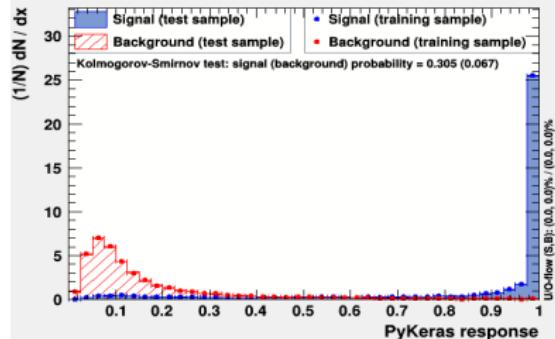
100 GeV

TMVA overtraining check for classifier: PyKeras



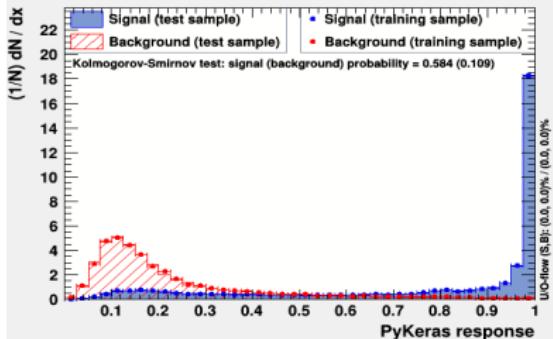
500 GeV

TMVA overtraining check for classifier: PyKeras

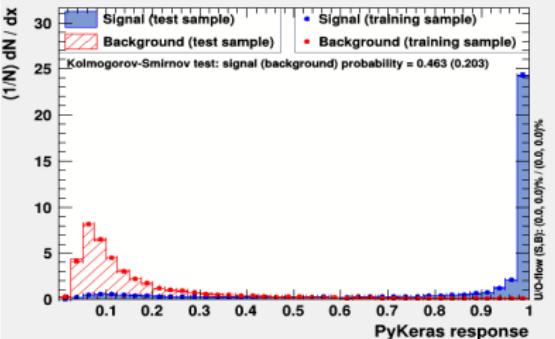


Pseudoscalar mediators

TMVA overtraining check for classifier: PyKeras

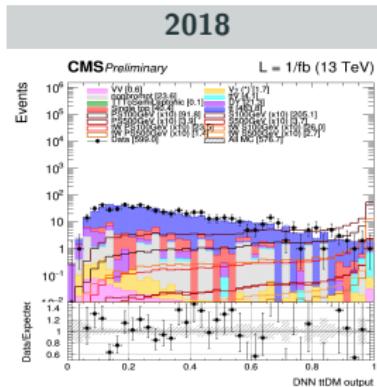
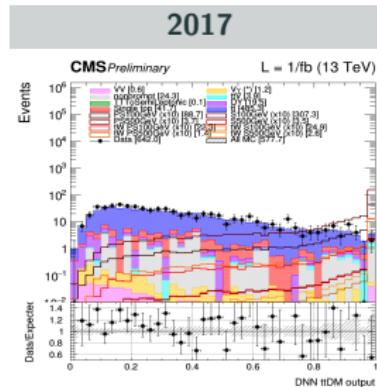
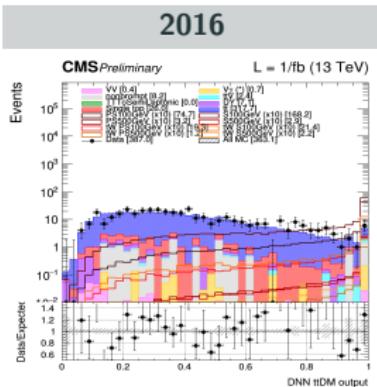


TMVA overtraining check for classifier: PyKeras

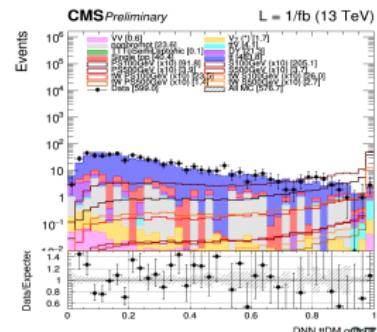
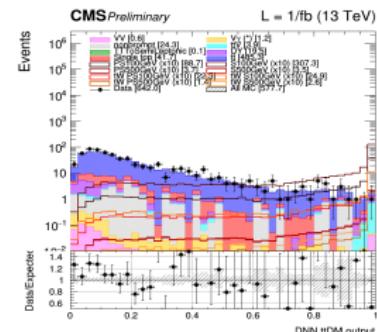
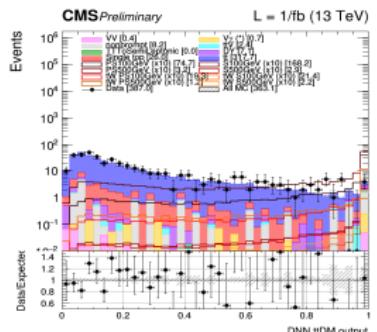


Blinded ANN output shape

Scalar 100 GeV output shape



Scalar 500 GeV output shape



Background prediction methods

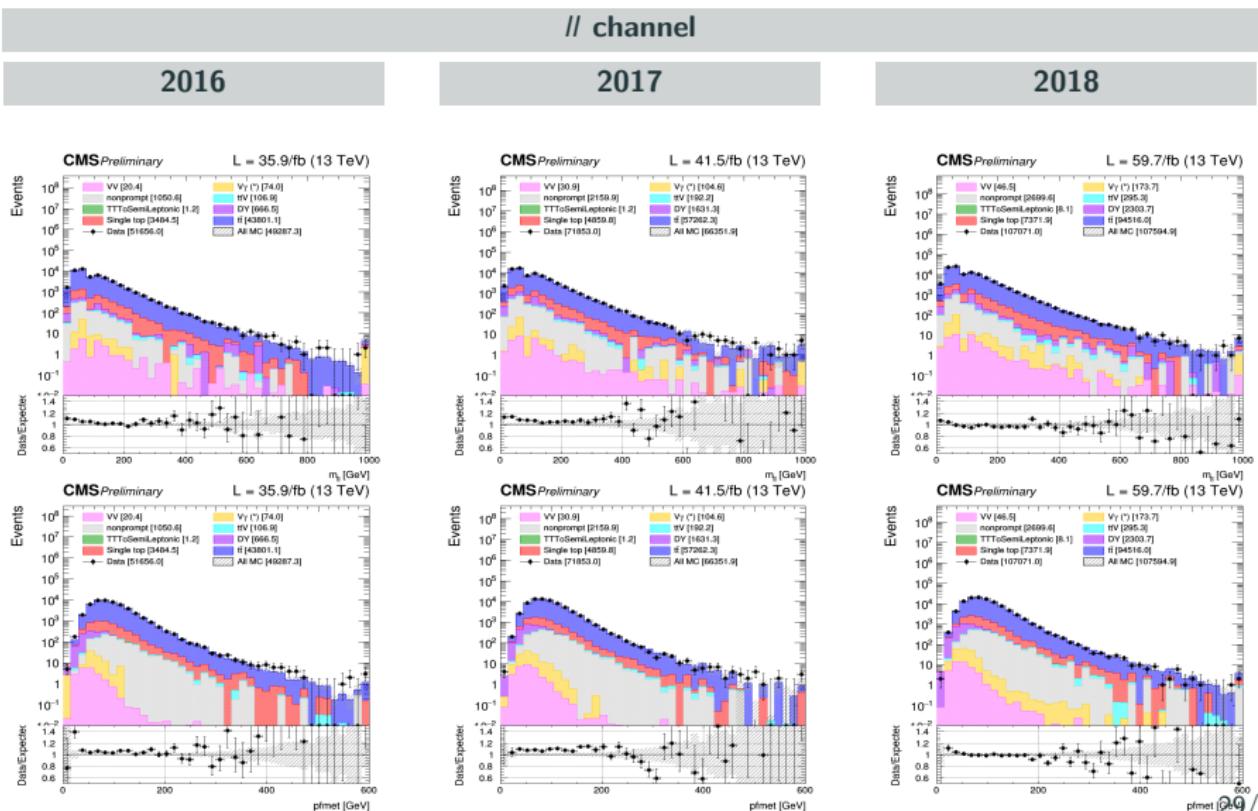
Main background processes

The backgrounds are predicted either directly from Monte-Carlo simulations or from data-driven methods. In order of importance:

- 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).
→ A data validation region (low $m_{T_2}^H$) is explored to ensure the quality of the prediction;
- 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;
- The **non-prompt contamination** is estimated from data control regions and validated in a same sign validation region;
- The irreducible **ttV process** ($ttW + ttZ$) is taken from simulation and checked in a particular validation region;
- **Diboson processes and other minor backgrounds** are taken directly from MC.

Top control region

Same as the signal region but with $60 < m_{T2}^H < 80$ GeV.



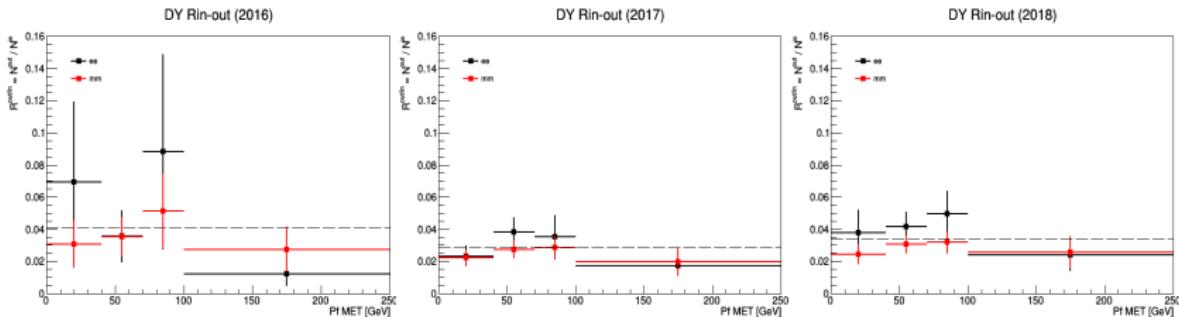
DY Rin-out method

We want to estimate the DY yields outside of the Z-peak from the data:

- 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**:

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

- 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 MET and for each channel (ee , $\mu\mu$) separately**:

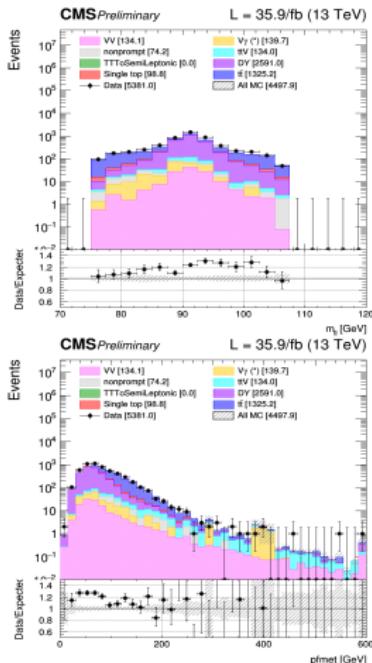


A flat scale factor and a fixed 20% systematic uncertainty is then applied to the DY. This method and the difference in statistics are still being studied.

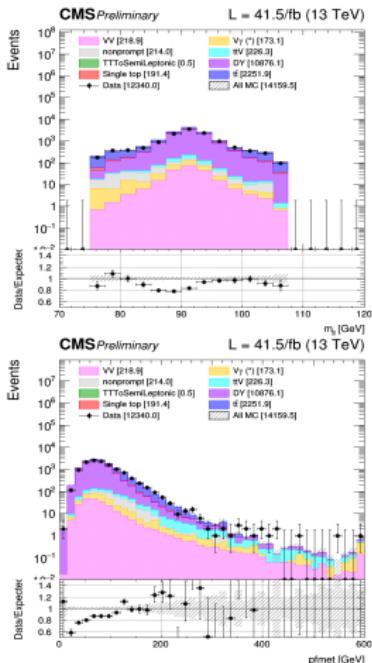
DY control region

// channel

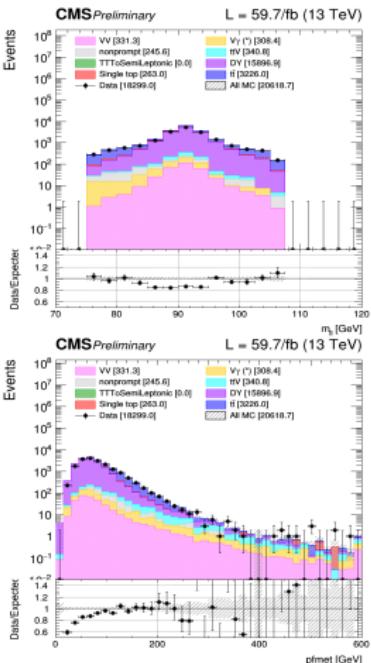
2016



2017



2018



The 0-bjet correction allows us to fix the data/MC discrepancies observed. A large systematic uncertainty is associated to this background, minor in the signal regions.

Non-prompt contamination I

Fake leptons detection (mostly jets misidentified or leptons coming from heavy hadrons decays) in the detector needs to be taken into account properly, through a **data-driven tight-to-loose method** since the Monte-Carlo is not reliable in this case:

Fake rate

- A QCD enriched region is defined with a looser particle selection criteria, where the misidentification should be high;
- Any eventual contamination from electroweak processes in this region is removed;
- The **fake rate** is defined as the ratio between the fakeable object (lepton-like objects passing only the loose isolation requirements) and fully selected objects yields.

Prompt rate

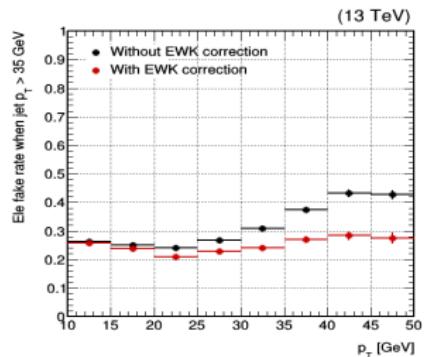
- The **prompt rate**, taking into account the real lepton contamination is calculated in a Z enriched region from a general tag and probe method.

Then, we calculate from data an extrapolation factor to go back to the signal region of the analysis and the results obtained are checked in a **same sign control region**.

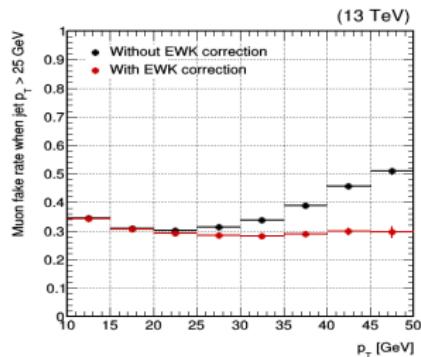
Non-prompt contamination II

2016 fake rate

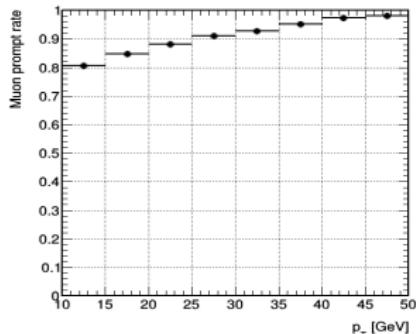
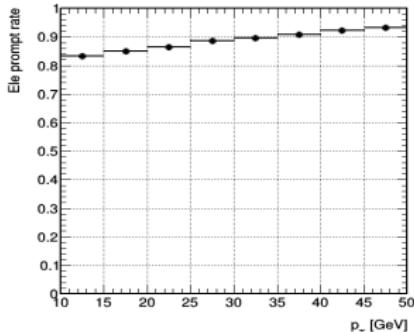
Electron



Muon



2016 prompt rate

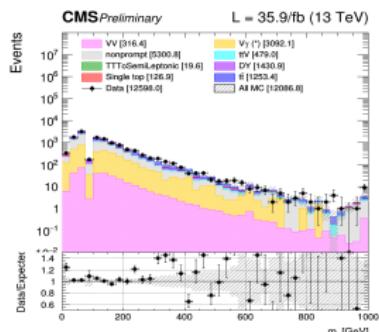


Same sign control region

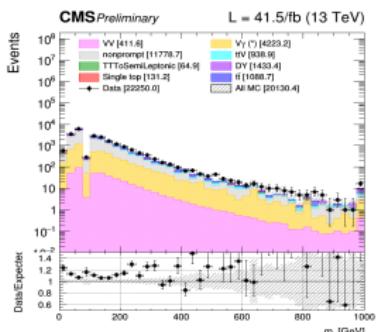
This is checked in a dedicated same sign control region.

// channel

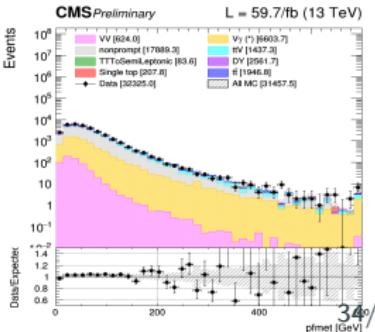
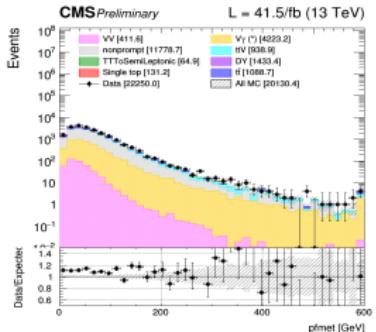
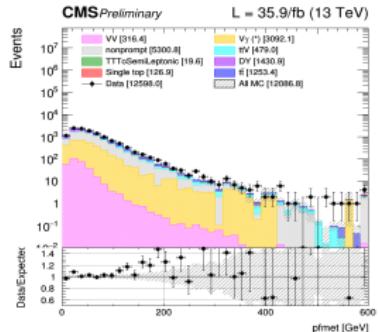
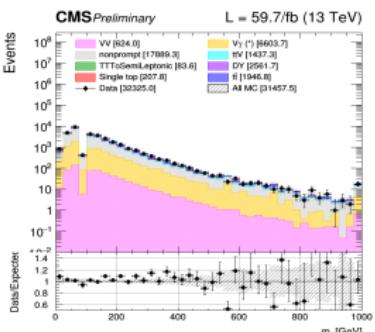
2016



2017



2018

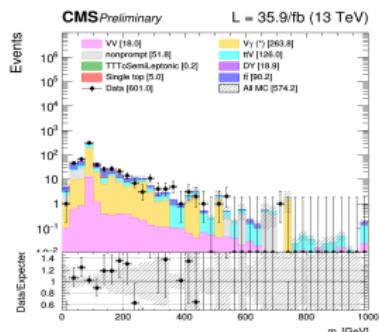


$t\bar{t}V$ prediction

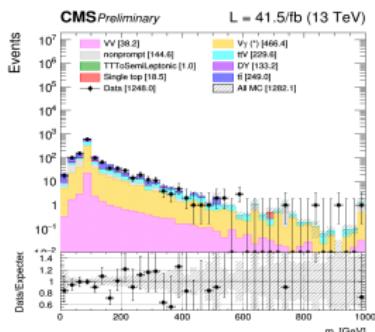
This process is taken directly from MC, and crosschecked in a dedicated control region.

// channel

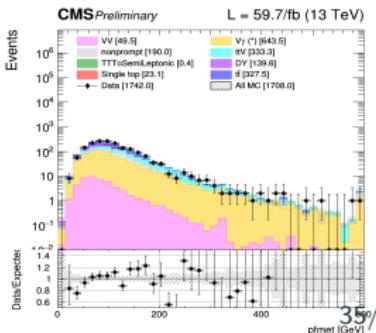
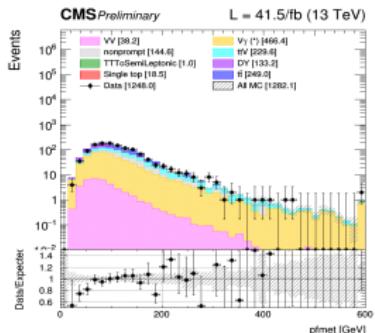
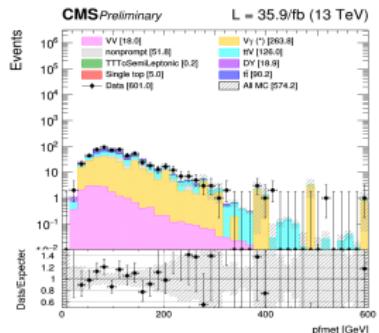
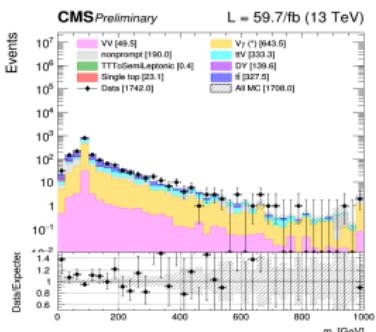
2016



2017



2018



Systematic uncertainties

Systematic uncertainties

Most of the systematics to be considered (on top of the statistical uncertainties) are already in place, such as:

Theoretical uncertainties

- PDF and higher order corrections, underlying event and parton shower, renormalization and factorization scales.

Experimental uncertainties

- Luminosity, pileup modeling, lepton trigger, lepton efficiency and energy scale, jet energy scale, MET mismodelling, b-tagging efficiency, top p_T reweighting.

Background specific uncertainties

- Drell-Yan and non-prompt backgrounds related uncertainties.

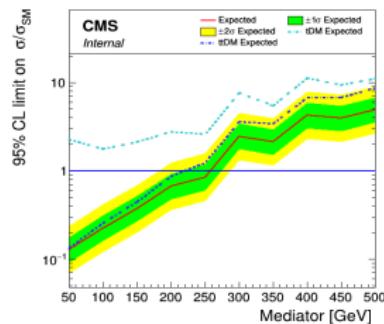
Disclaimer: this part of the analysis still needs to be checked/optimized, so results shown next do not include any systematics for now, even though we are mostly ready to include them when needed.

Results obtained

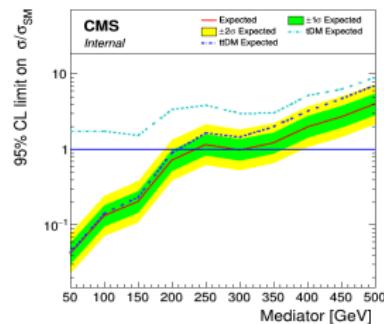
Upper limits without systematics

Scalar upper limits

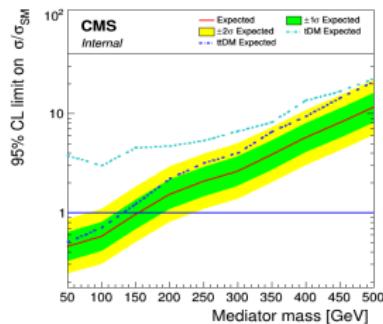
2016



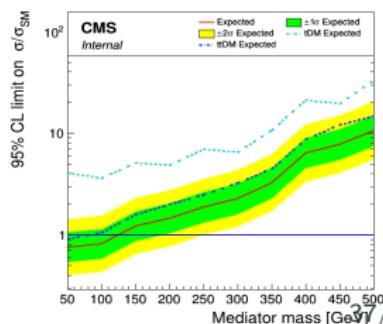
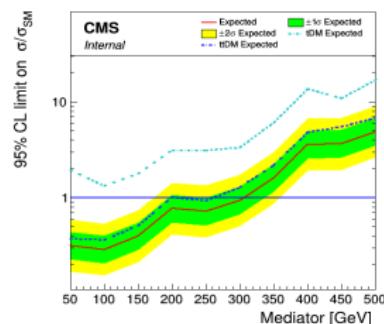
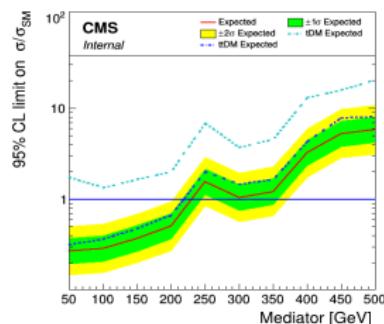
2017



2018

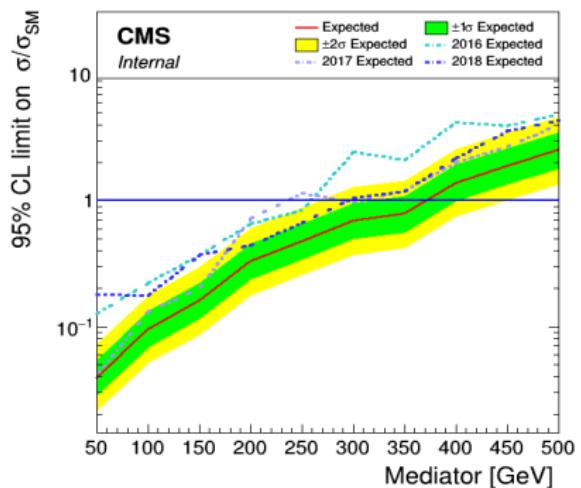


Pseudoscalar upper limits

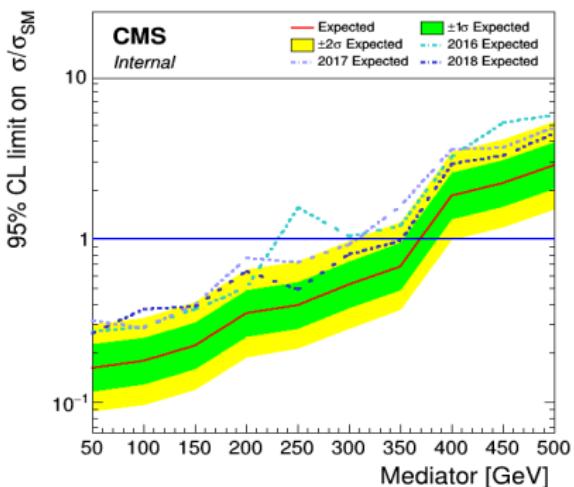


Disclaimer: comparison done without systematics and expected to change a bit.

Scalar mediators



Pseudoscalar mediators



A search for dark matter produced in association with either one or two top quarks is on-going and was presented:

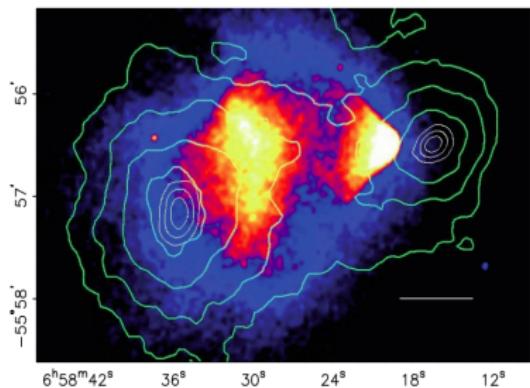
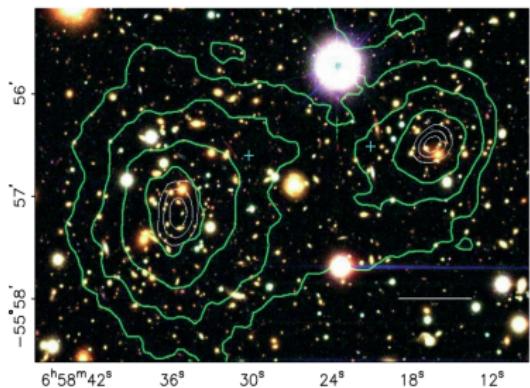
- This search is considering the **Run II legacy dataset** collected by the CMS detector;
- At IFCA, our efforts are entirely focused on the **dilepton final state**;
- This search is performed by defining a ANN, training it to be able to recognize background and signal events, to separate them and **increase the signal efficiency**;
- First time that such a combination will be performed considering this canal, which should increase by a lot the limits published in 2016;
- We expect this analysis **to be approved by June**.

Back up

Gravitational lensing

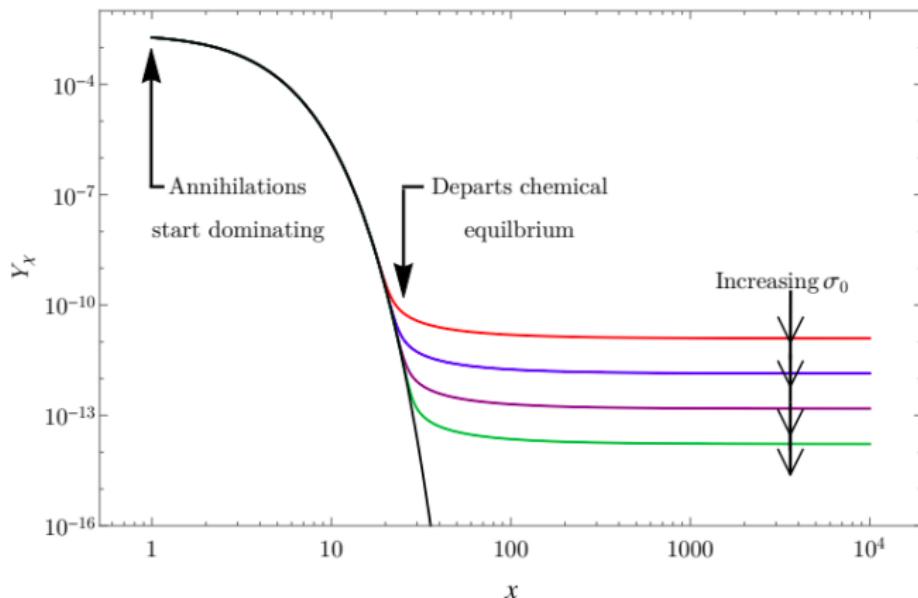
Consequence of the general relativity: massive objects placed between distant sources and the observer should be able to act as lenses and bend the light of the source.

- The deviation of the light is proportional to the mass of the intermediate object, giving us a way to measure its mass;
- The mass distribution obtained has been compared to the luminous distribution of several galaxies, leading to 8σ discrepancies.

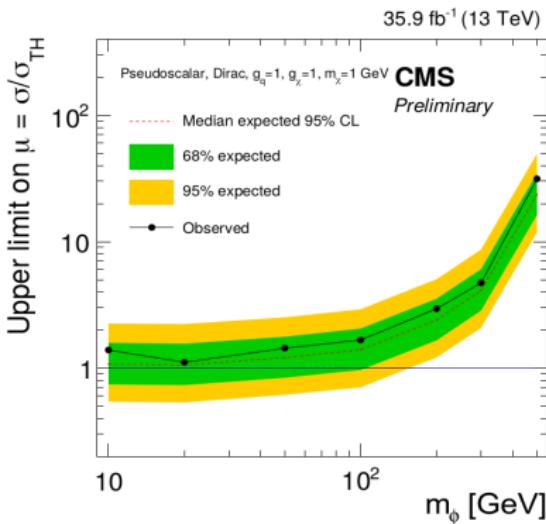
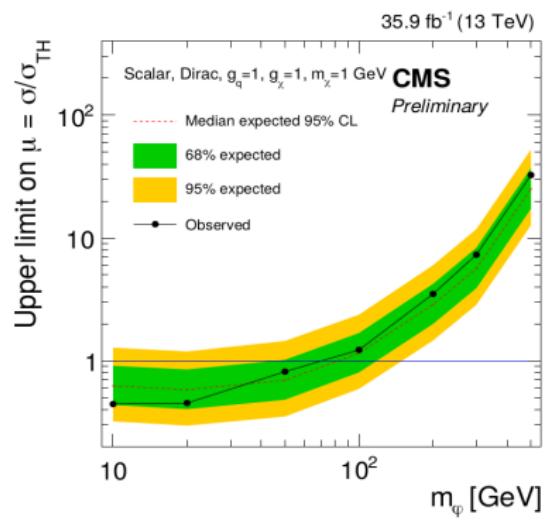


Thermal freeze-out

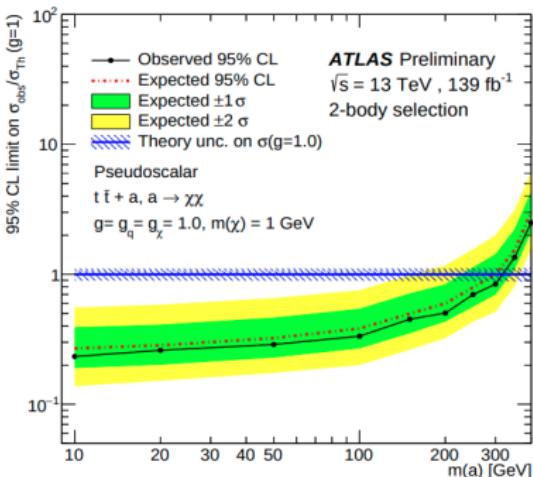
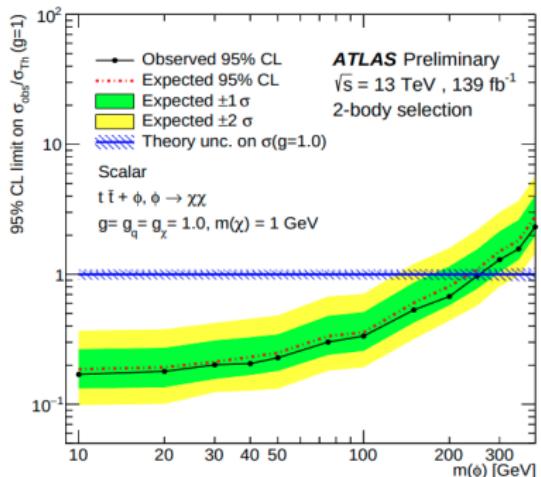
Schematic representation of the freeze-out process, representing the abundance of a 500 GeV dark matter with respect to the time and the impact of increasing cross-section annihilation values on this freeze-out abundance.



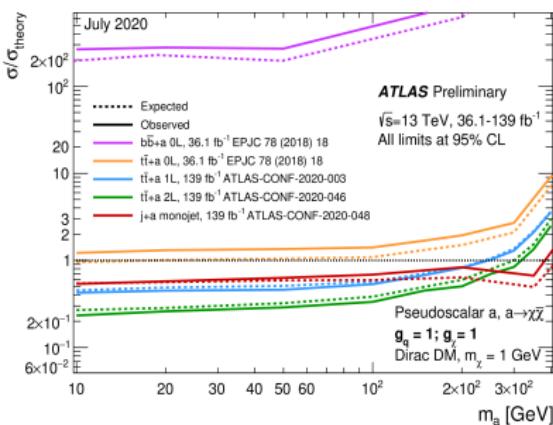
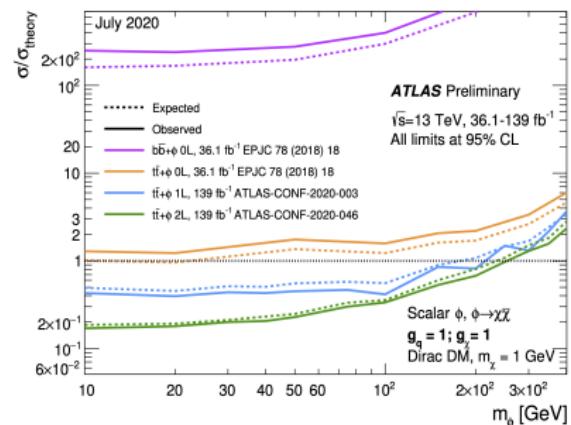
CMS dilepton channel 2016 results published:



ATLAS dilepton channel Run II legacy results from ICHEP 2020:



The ATLAS collaboration does not perform any combination between the different channels though.



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 ($H_T < 70$ GeV) DYJetsToLL_M-5to50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-5to50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-5to50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-5to50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-5to50_HT-600tolInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 ($H_T < 70$ GeV) DYJetsToLL_M-50_HT-70to100_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-100to200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-200to400_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-400to600_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-600to800_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-800to1200_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-1200to2500_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-2500tolInf_TuneCUETP8M1_13TeV-madgraphMLM-pythia8	18610.0 303.8 224.2 37.2 3.581 1.124 6025.20 169.9 147.4 40.99 5.678 1.367 0.6304 0.1514 0.003565
TTTo2L2Nu	TTTo2L2Nu_TuneCUETP8M2_ttHtranche3_13TeV-powheg-pythia8	87.310
Single top	ST_s-channel_4f_leptonDecays_13TeV-amcatnlo-pythia8_TuneCUETP8M1 ST_t-channel_antitop_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 ST_t-channel_top_4f_inclusiveDecays_13TeV-powhegV2-madspin-pythia8_TuneCUETP8M1 ST_tW_antitop_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1 ST_tW_top_5f_inclusiveDecays_13TeV-powheg-pythia8_TuneCUETP8M1	3.360 80.95 136.02 35.60 35.60
ttV	TTWJetsToLNu_TuneCUETP8M1_13TeV-amcatnloFXF-madspin-pythia8 ttZJets_13TeV_madgraphMLM-pythia8 TTZToLLNuNu_M-10_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.2043 0.7826 0.2529
WW	WWTo2L2Nu_13TeV-powheg WWJToLNuNu_EWK_noTop_13TeV-madgraph-pythia8 GluGluWVTo2L2Nu_MCFM_13TeV	12.178 0.34520 0.5905
V γ /V γ^*	WGToLNuG_TuneCUETP8M1_13TeV-madgraphMLM-pythia8 ZGTo2LG_TuneCUETP8M1_13TeV-amcatnloFXF-pythia8 WZTo3LNu_mlmlmin01_13TeV-powheg-pythia8	405.271 131.300 58.59
VZ	ZZTo2L2Nu_13TeV_powheg_pythia8 ZZTo2L2Q_13TeV_powheg_pythia8 ZZTo4L_TuneCP5_13TeV_powheg_pythia8 WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	0.5640 3.22 1.212 5.595
VVV	ZZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 WZZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 WWZ_TuneCUETP8M1_13TeV-amcatnlo-pythia8 WWW_4F_TuneCUETP8M1_13TeV-amcatnlo-pythia8	0.01398 0.05565 0.16510 0.18331
Non-Prompt	Data-driven (tight-to-loose method)	

2017 MC samples

Process	Sample	Cross section [pb]
Drell-Yan	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 ($H_T < 100$ GeV) DYJetsToLL_M-4to50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-4to50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-4to50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-4to50_HT-600tolnf_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-4to50_HT-70 < 70 GeV) DYJetsToLL_M-50_TuneCP5_13TeV-madgraphMLM-pythia8 ($H_T < 70$ GeV) DYJetsToLL_M-50_HT-70to100_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-100to200_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-200to400_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-400to600_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-600to800_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-800to1200_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-1200to2500_TuneCP5_13TeV-madgraphMLM-pythia8 DYJetsToLL_M-50_HT-2500tolnf_TuneCP5_13TeV-madgraphMLM-pythia8	18610 204.0 54.39 5.697 1.85 6189.39 169.9 161.1 48.66 6.968 1.743 0.8052 0.1933 0.003468
TTTo2L2Nu	TTTo2L2Nu_TuneCP5_13TeV-powheg-pythia8	87.310
Single top	ST_s-channel_4f_leptonDecays_mtop1715_TuneCP5_PSweights_13TeV-amcatnlo-pythia8 ST_t-channel_antitop_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8 ST_t-channel_top_4f_inclusiveDecays_TuneCP5_13TeV-powhegV2-madspin-pythia8 ST_tW_antitop_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8 ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	3.360 80.95 136.02 35.60 35.60
ttV	ttWJets_TuneCP5_13TeV-madgraphMLM_pythia8 TTWJetsToLNu_TuneCP5_PSweights_13TeV-amcatnlo_FXFX-madspin-pythia8 ttZJets_TuneCP5_13TeV_madgraphMLM_pythia8 TTZToLLNuNu_M-10_TuneCP5_PSweights_13TeV-amcatnlo-pythia8	0.6105 0.2001 0.7826 0.2529
WW	WWTo2L2Nu_NNPDF31_TuneCP5_PSweights_13TeV-powheg-pythia8 WWJJToLNuLNu_EWK_noTop_TuneCP5_13TeV-madgraph-pythia8 GluGluToWWTo_13TeV_MCFM701_pythia8	12.178 0.34520 0.06387
V γ /V γ^*	WGToLNuG_TuneCP5_13TeV-madgraphMLM-pythia8 ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnlo_FXFX-pythia8 WZTo3LNu_mlmin01_NNPDF31_TuneCP5_13TeV_powheg_pythia8	405.271 58.83 58.59
VZ	ZZTo2L2Nu_13TeV_powheg_pythia8 ZZTo2L2Q_13TeV_amcatnlo_FXFX_madspin_pythia8 ZZTo4L_TuneCP5_13TeV_powheg_pythia8 WZTo2L2Q_13TeV_amcatnlo_FXFX_madspin_pythia8	0.5640 3.22 1.212 5.595
VVV	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8 WZZ_TuneCP5_13TeV-amcatnlo-pythia8 WWZ_4F_TuneCP5_13TeV-amcatnlo-pythia8 WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.01398 0.05565 0.16510 0.18331
Non-Prompt	Data-driven (tight-to-loose method)	

2018 MC samples

Process	Sample	Cross section [pb]
Drell-Yan	DYJetsToLL_M-10to50_TuneCP5_13TeV-madgraphMLM-pythia8 ($H_T < 100$ GeV)	18610.0
	DYJetsToLL_M-4to50_JHT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	204.0
	DYJetsToLL_M-4to50_JHT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	54.39
	DYJetsToLL_M-4to50_JHT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	5.697
	DYJetsToLL_M-4to50_JHT-600to1000_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.85
	DYJetsToLL_M-4to50_JHT-70to100_TuneCP5_13TeV-amcatnloFXFX-pythia8 ($H_T < 70$ GeV)	6189.39
	DYJetsToLL_M-50_JHT-70to100_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	169.9
	DYJetsToLL_M-50_JHT-100to200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	161.1
	DYJetsToLL_M-50_JHT-200to400_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	48.66
	DYJetsToLL_M-50_JHT-400to600_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	6.968
	DYJetsToLL_M-50_JHT-600to800_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	1.743
	DYJetsToLL_M-50_JHT-800to1200_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.8052
	DYJetsToLL_M-50_JHT-1200to2500_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.1933
	DYJetsToLL_M-50_JHT-2500toInf_TuneCP5_PSweights_13TeV-madgraphMLM-pythia8	0.003468
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	35.60
	ST_tW_top_5f_inclusiveDecays_TuneCP5_13TeV-powheg-pythia8	35.60
ttV	ttWJets_TuneCP5_13TeV_madgraphMLM_pythia8	0.6105
	TTWJetsToLNu_TuneCP5_13TeV_amcatnloFXFX-madspin_pythia8	0.2043
	ttZJets_TuneCP5_13TeV_madgraphMLM_pythia8	0.7826
	TTZToLLNuNu_M-10_TuneCP5_13TeV-amcatnlo-pythia8	0.2529
WW	WWTo2L2Nu_NNPDF31_TuneCP5_13TeV-powheg-pythia8	12.178
	WWJJToLNuNu_EWK_TuneCP5_13TeV-madgraph-pythia8	0.4286
	GluGluToWWTo*_TuneCP5_13TeV_MCFM701_pythia8	0.06387
V γ /V γ^*	WGToNuG_TuneCP5_13TeV-madgraphMLM-pythia8	405.271
	ZGToLLG_01J_5f_TuneCP5_13TeV-amcatnloFXFX-pythia8	131.300
	WZTo3LNu_mlmin01_NNPDF31_TuneCP5_13TeV_powheg_pythia8	58.59
VZ	ZZTo2L2Nu_TuneCP5_13TeV_powheg_pythia8	0.5640
	ZZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	3.22
	ZZTo4L_TuneCP5_13TeV_powheg_pythia8	1.212
	WZTo2L2Q_13TeV_amcatnloFXFX_madspin_pythia8	5.595
VVV	ZZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.01398
	WZZ_TuneCP5_13TeV-amcatnlo-pythia8	0.05565
	WWZ_TuneCP5_13TeV-amcatnlo-pythia8	0.16510
	WWW_4F_TuneCP5_13TeV-amcatnlo-pythia8	0.18331
Non-Prompt	Data-driven (tight-to-loose method)	

Mass point	Cross-section [pb]
Scalar mediators	
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi10	0.4632
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi20	0.3021
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi50	0.1236
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi100	0.05262
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi150	0.03173
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi200	0.02204
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi250	0.01627
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi300	0.01240
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi350	0.00951
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi400	0.006274
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi450	0.004236
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi500	0.002995
DMscalar_Dilepton_top_tWChan_Mchi1_Mphi1000	0.0002844
Pseudoscalar mediators	
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi10	0.05745
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi20	0.05482
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi50	0.0462
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi100	0.03417
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi150	0.02572
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi200	0.01959
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi250	0.01509
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi300	0.01170
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi350	0.007333
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi400	0.004083
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi450	0.002891
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi500	0.002168
DMpseudoscalar_Dilepton_top_tWChan_Mchi1_Mphi1000	0.0002607

$t\bar{t} + \text{DM}$ 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	$7.750 \cdot 10^{-4}$

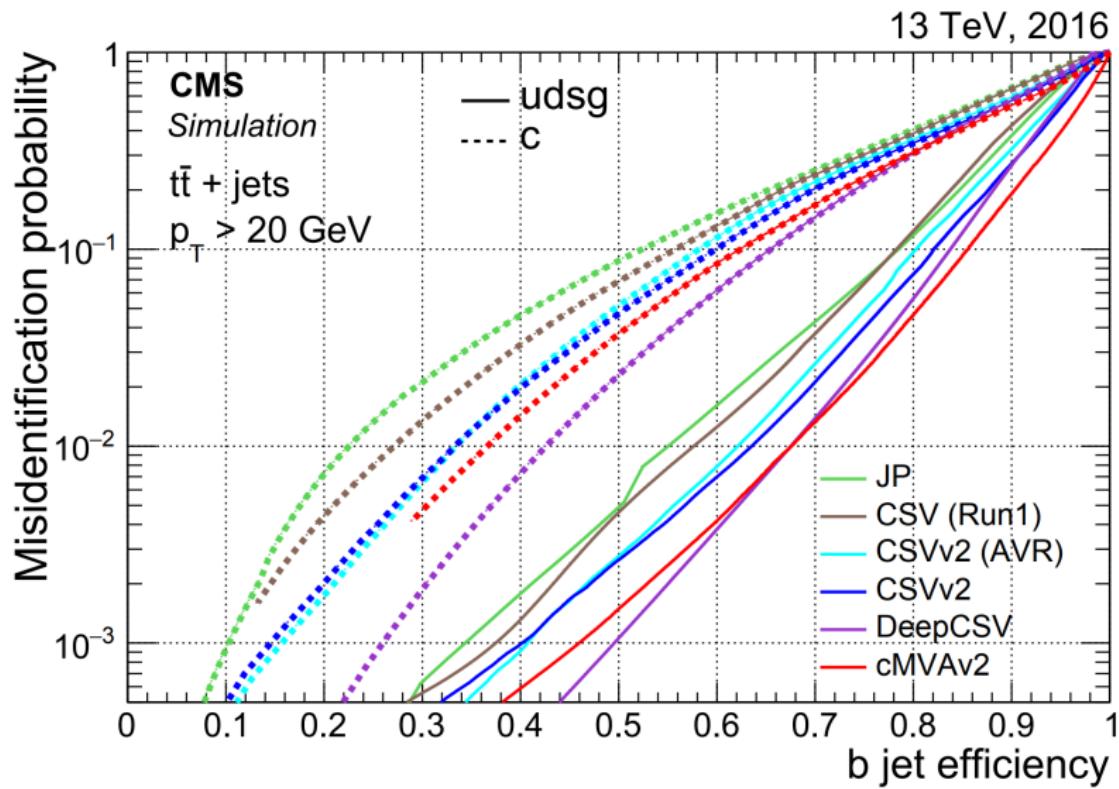
2016 triggers

Dataset	Run range	HLT trigger path
SingleMu	[273158,284044]	HLT_IsoMu24_v* HLT_IsoTkMu24_v*
SingleEle	[273158,284044]	HLT_Ele27_WPTight_Gsf_v* HLT_Ele25_eta2p1_WPTight_Gsf_v*
DoubleEG	[273158,284044]	HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*
DoubleMu	[273158,281612] [281613,284044]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_v* HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_v* HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* HLT_Mu17_TrkIsoVVL_TkMu8_TrkIsoVVL_DZ_v*
MuonEG	[273158,278272] [278273,284044]	HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL HLT_Mu8_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v* HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_DZ_v*

2017 triggers

Dataset	Run range	HLT trigger path
SingleMu	[297046,306462]	HLT_IsoMu27_v*
EGamma	[297046,306462]	HLT_Ele35_WPTight_Gsf_v* HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*
DoubleMu	[297046,299329] [299368,306462]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_v* HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v*
MuonEG	[297046,306462] [297046,299329] [299368,306462]	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	[315252,325172]	HLT_IsoMu24_v*
	[314859,325175]	HLT_Mu5_v*
		HLT_IsoMu27_v*
EGamma	[315252,325172]	HLT_Ele32_WPTight_Gsf_v*
		HLT_Ele35_WPTight_Gsf_v*
		HLT_Ele23_Ele12_CaloIdL_TrackIdL_IsoVL_v*
DoubleMu	[315252,325172]	HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass3p8_v*
		HLT_Mu17_TrkIsoVVL_Mu8_TrkIsoVVL_DZ_Mass8_v*
MuonEG	[315252,325172]	HLT_Mu23_TrkIsoVVL_Ele12_CaloIdL_TrackIdL_IsoVL_v*
		HLT_Mu12_TrkIsoVVL_Ele23_CaloIdL_TrackIdL_IsoVL_DZ_v*



Stransverse mass M_{T2}^{\parallel}

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 MET and the individual contribution of each particle ($\not{p}_{T_1}, \not{p}_{T_2}$) to this missing energy cannot be inferred. M_{T2}^{\parallel} is defined as:

$$\begin{cases} M_{T2}^{\parallel} = \min_{\not{p}_{T_1} + \not{p}_{T_2} = \not{p}_{T_{\text{tot}}}} \left(\max \left(m_T^2(\not{p}_{T_1}, \not{p}_{T_1}), m_T^2(\not{p}_{T_2}, \not{p}_{T_2}) \right) \right) \\ m_T^2(\not{p}_T, \not{p}_T) = 4 |\not{p}_T| |\not{p}_T| \sin^2 \left(\frac{\alpha}{2} \right) \end{cases}$$

Different combinations ($\not{p}_{T_1}, \not{p}_{T_2}$) satisfying the condition $\not{p}_{T_1} + \not{p}_{T_2} = \not{p}_{T_{\text{tot}}}$ 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.

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.

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 actually be inferred from the top quark decay products.

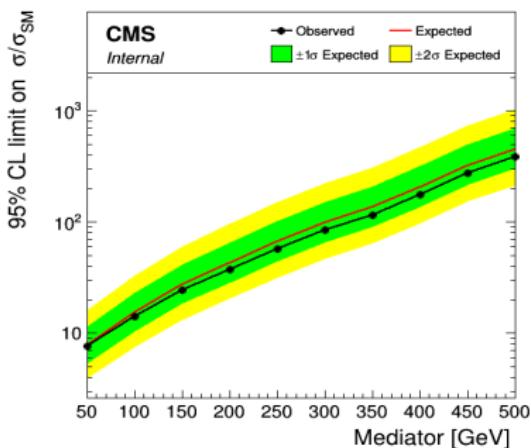
Such variables are interesting because **the spin correlation depend on the production mechanism** and will be influenced by the additional coupling to a scalar or pseudoscalar mediator, making this a perfect candidate to be a good discriminating variable.

In particular, two spin correlated variables are considered:

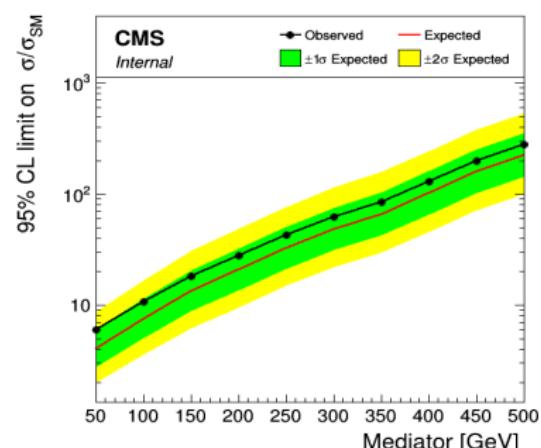
- $\xi = \cos(\theta_i) \cos(\theta_j)$, where i and j are either leptons, b-jets or neutrinos;
- $\cos(\Phi_{i,j})$, the cosine of the full opening angle of such top decay products in their respective parent rest frames.

Disclaimer: comparison done without systematics, considering the 2016 with a single 100 GeV training and for a single fb^{-1} of data given the blinding policy applied.

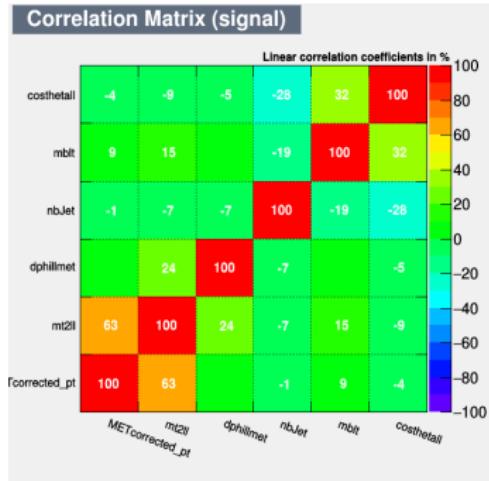
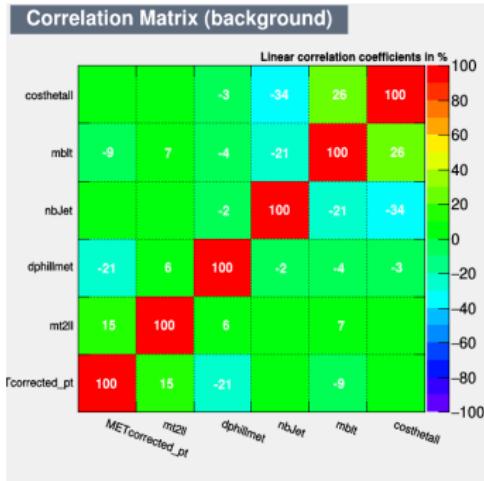
BDT



DNN

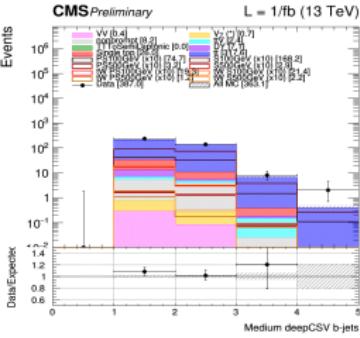
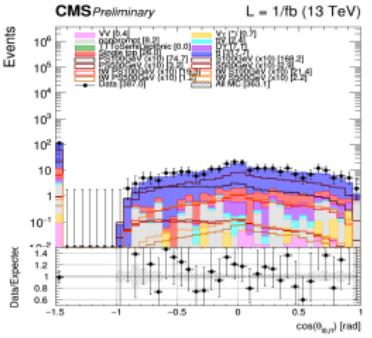
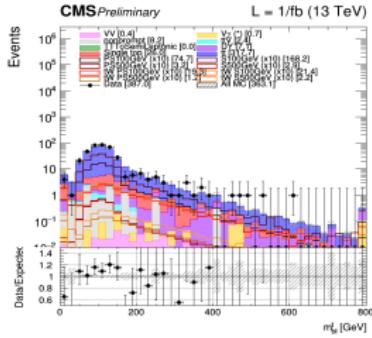
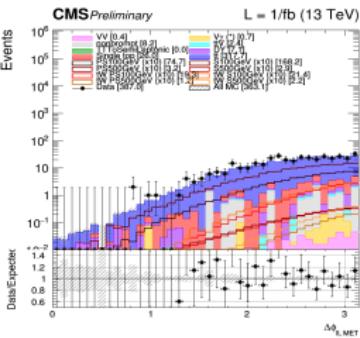
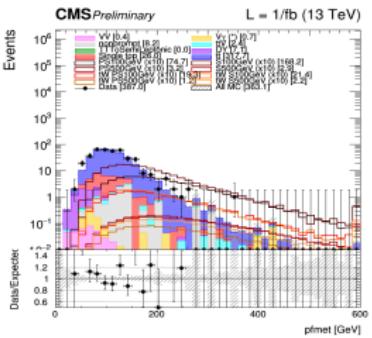
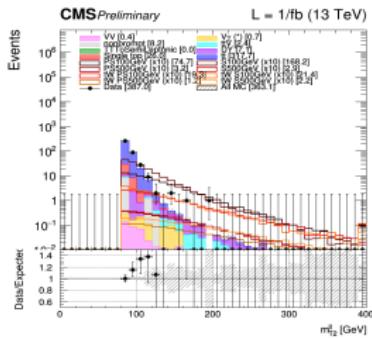


Input variables correlation



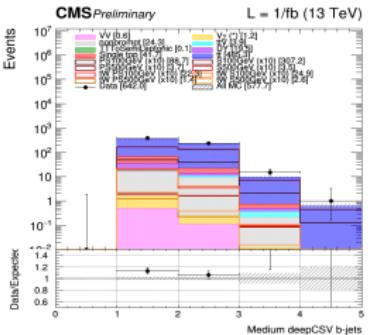
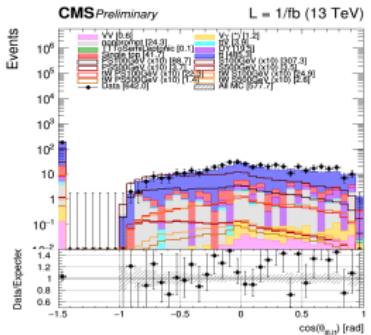
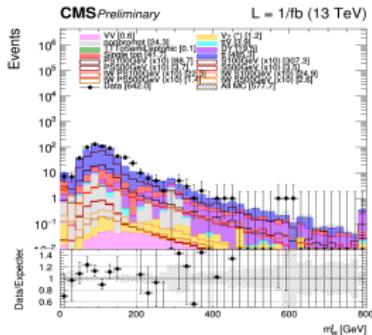
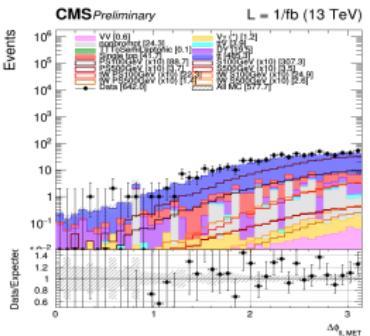
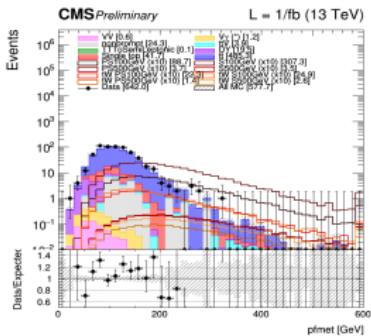
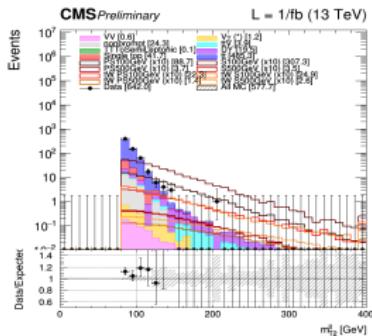
2016 discriminating variables

2016, // channel



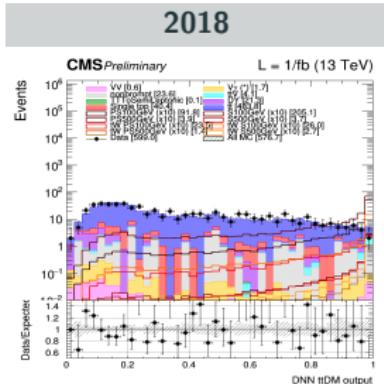
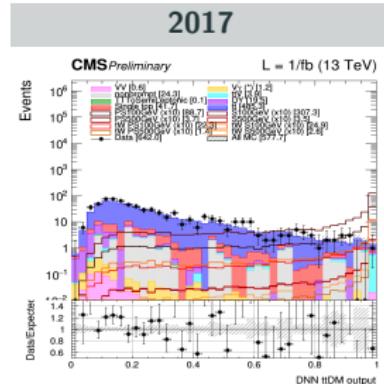
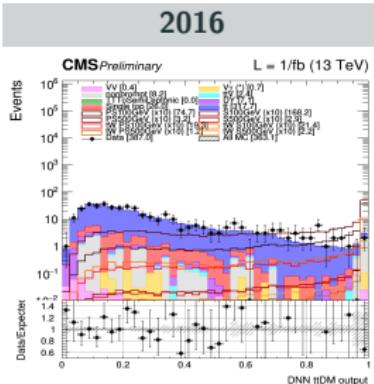
2017 discriminating variables

2017, // channel

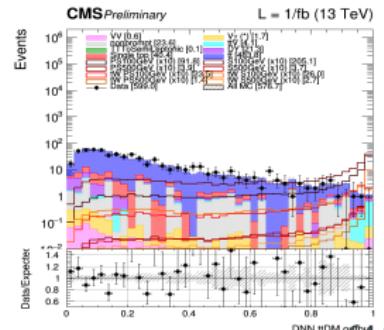
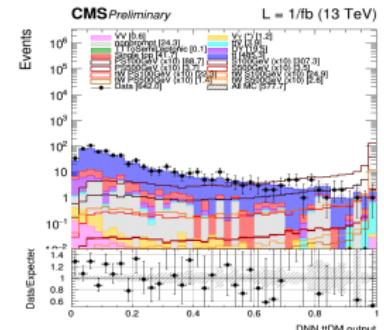
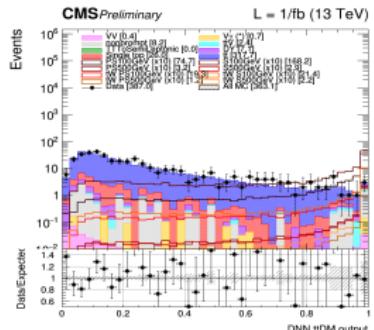


Blinded ANN output shape

Pseudoscalar 100 GeV output shape



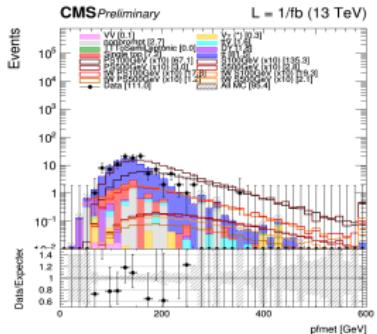
Pseudoscalar 500 GeV output shape



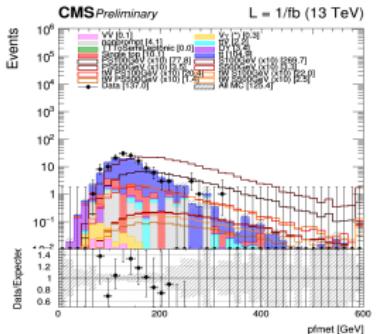
Scalar 100 GeV signal region

// channel

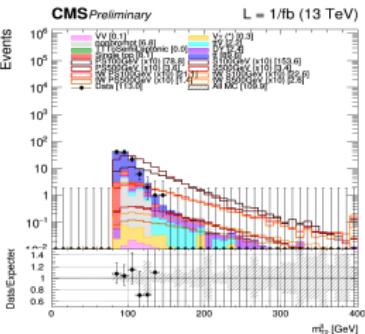
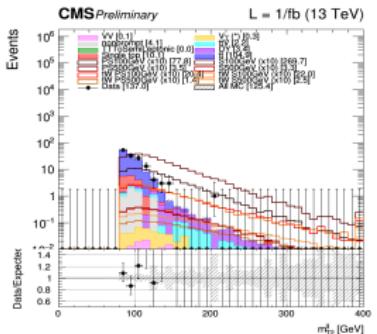
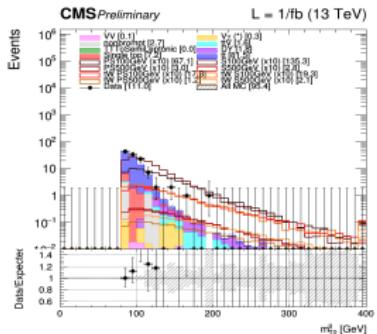
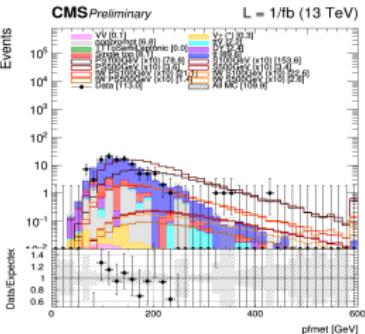
2016



2017



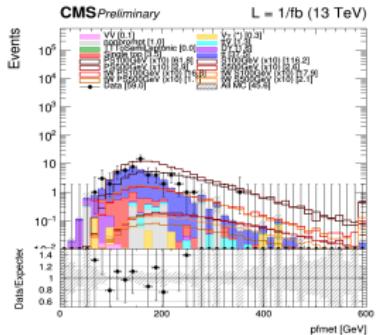
2018



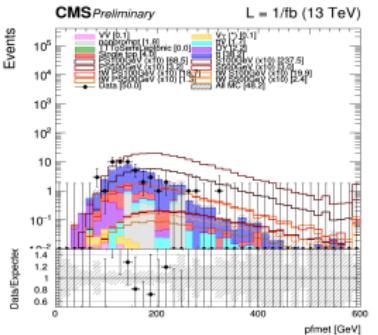
Scalar 500 GeV signal region

// channel

2016



2017



2018

