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

Cédric Prieels

Director - Jónatan Piedra Gómez
Co-director - Alicia Calderón Tazón

IFCA - Instituto de Física de Cantabria

UC - UNIVERSIDAD DE CANTABRIA

CSIC - Consejo Superior de Investigaciones Científicas

Instituto de Física de Cantabria
Universidad de Cantabria

- **June 26th, 2017** -

Outline

- Theoretical introduction

- Experimental device

- Objects, datasets, triggers and samples

- Event reconstruction and selection

- Multi-variate analysis and neural network

- Systematic uncertainties

- Limits of the analysis

- Conclusions

Introduction

In this work is presented a search for dark matter production in proton-proton collisions, using the data collected by the CMS detector at CERN during 2016. The data analyzed has been taken at a center of mass energy $\sqrt{s} = 13$ TeV.

Motivations for the analysis

- There are several **astrophysical evidences** for the existence of dark matter, but **no direct nor indirect detection** so far.

- We hope **to be able to produce WIMPs** (Weakly Interactive Massive Particles) in high energy collisions of ordinary particles within the LHC.

Main objective

- Consider different dark matter production models to **search for dark matter**, and to eventually exclude some of these models or at least put upper limits on their cross section production.

Standard Model

- The most accepted and used model to describe the elementary particles and the fundamental interactions between them is the **Standard Model**.

- This model is quite elegant and is working really well, and has successfully made different predictions, such as the existence of the Higgs boson, finally discovered in 2012.
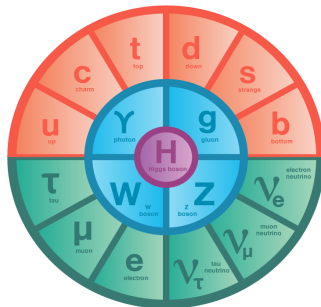
$\rightarrow$ However, this model is known to have **several shortcomings** which require further investigation. For example, eventual exotic particles which do not fit within this model (such as dark matter) are extensively searched for nowadays.

Dark matter

The dark matter hypothesis was introduced as a way to explain the strange behavior of the **rotation curves** of the galaxies at high radius, and the apparent **missing mass** in the Universe.
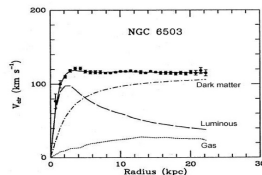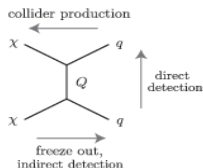
K.G. Begeman, A.H. Broels, R.H. Sanders. 1991. Mon.Not.RAS 249, 523.

Ordinary baryonic matter constitutes only 5% of the total mass of the Universe, while the dark matter accounts for 27%.

We also assume that dark matter is **made out of cold particles** and interact only weakly with ordinary matter or itself, making it **extremely difficult to detect**.

$\rightarrow$ Since we do not know any baryonic candidate fulfilling these conditions, we can postulate the **existence of new elementary particles**, the WIMPs. This is typically the kind of particles we **hope to find at the LHC**.

Our channel of interest

This work will be focused on **dark matter production in association with a top and an anti-top quarks**, in the dilepton final state. The dark matter χ is produced through the interaction of two gluons and through the apparition of a **mediator** Φ for which we will consider **different masses and couplings** (scalar and pseudoscalar).
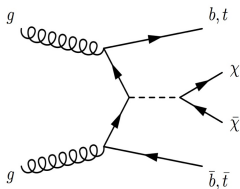

We are looking at this channel mainly because...

- The **signature left by the top quarks** can be isolated well with respect to the other backgrounds' leftovers.
- If we assume that the dark matter mediator couples in the same way than the Higgs does, it should have **stronger couplings with heavier particles** (and the top is the most massive fermion known).

The dilepton final state is interesting because, even though this is the channel with the **smallest branching ratio** (fraction of times for which a particle decays by an individual decay mode with respect to the total number of decays), this is also the channel having the **least number of backgrounds**.

The Large Hardon Collider
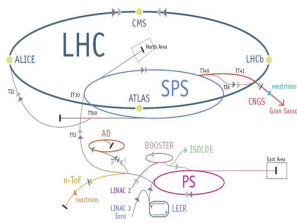
The data has been taken by the Large Hardron Collider, a **circular underground proton-proton collider**, situated at CERN. With its 27 kilometers of circumference, it is currently the **most powerful accelerator** in the world.
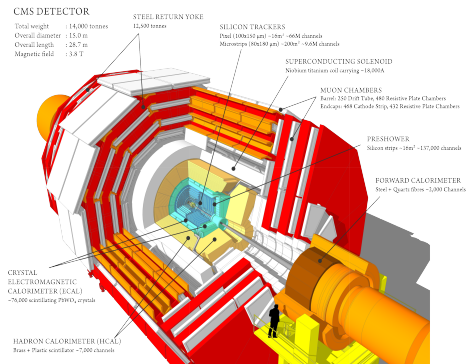
It is the result of a collaboration of 22 countries, and has been built in order to study and reproduce the Universe at its origin.

The LHC in a nutshell

- Beams of protons circulate in opposite directions at velocities close to the speed of light and collide in 4 different points, where the detectors are.
- Accelerates 2808 bunches of 10^{11} protons up to a center of mass energy of 13 TeV so far (but has been designed to reach 14 TeV).
- Can reach an instantaneous luminosity of 10^{34} cm^{-2}s^{-1} and a collision rate of 40 MHz.

The Compact Muon Solenoid

CMS is one of the two **polyvalent detectors** of the LHC, designed to make measurements in most of the major different fields of particle physics, from precision measurements to the search of new exotic processes.

CMS DETECTOR
Total weight : 14,000 tonnes
Overall diameter : 15.0 m
Overall length : 28.7 m
Magnetic field : 3.8 T

STEEL RETURN YOKE
12,500 tonnes

SILICON TRACKERS
Pixel (100x150 μm) ~16m² ~66M channels
Microstrips (80x180 μm) ~200m² ~9.6M channels

SUPERCONDUCTING SOLENOID
Niobium titanium coil carrying ~18,000A

MUON CHAMBERS
Barrel: 250 Drift Tube, 480 Resistive Plate Chambers
Endcaps: 468 Cathode Strip, 432 Resistive Plate Chambers

PRESHOWER
Silicon strips ~16m² ~137,000 channels

FORWARD CALORIMETER
Steel + Quartz fibres ~2,000 Channels

CRYSTAL
ELECTROMAGNETIC
CALORIMETER (ECAL)
~76,000 scintillating PbWO₄ crystals

HADRON CALORIMETER (HCAL)
Brass + Plastic scintillator ~7,000 channels

CMS in a nutshell

- Is a **compact** detector (its 14.000 tons are compacted in a relatively small volume).

- Has a **powerful tracker and muon detection system** allowing us to measure some of the properties of the leptons throughout a large range of energies.

- Has a **huge solenoid** as central piece able to produce a 3.8T magnetic field, to curve the particles and study their properties.

- Is made out of **different layers** (such as the tracker, the calorimeters and the muon chambers), each having its own purpose.

What are we looking for?

For an event to be considered interesting, we basically require the presence of exactly **two energetic leptons**, at least **two b-jets** and some **missing transverse energy** (corresponding to the unbalanced transverse momentum of the collision).

Two leptons and two b-jets

We are looking for dark matter produced with a top and an anti-top, which are not stable and decay immediately to a $W^{\pm}$ and a bottom quark. We are then searching for the results of the **hadronisation of the bottom** quarks, and for **two leptons coming from the $W^{\pm}$**.
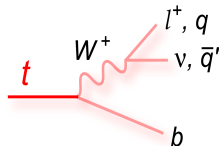
Missing transverse energy

Moreover, since dark matter particles do not interact with ordinary matter, we do not expect this kind of particles to **interact at all with the detector**. We are then also looking for a large amount of **missing transverse energy**.

Datasets and triggers

The datasets...

- Have been recorded during 7 different periods and have been selected to match the current **blinding policy** (we look at the already published 2.4 fb^{-1} for the signal region, and at the full 35.9 fb^{-1} for the control regions).

- This blinding policy consists in optimizing the analysis considering a **limited dataset** only, to avoid any **conscious or unconscious bias** when one tries to optimize his analysis based on what has already been seen.

Era	Luminosity (fb^{-1})
Run2016B	5.748
Run2016C	2.573
Run2016D	4.248
Run2016E	4.009
Run2016F	3.102
Run2016G	7.540
Run2016H	8.606
Total	35.827

The triggers...

- Are used to **store only interesting events** out of the 40 MHz of collisions produced.

- Have been chosen and combined to only select events with at least 1 or 2 leptons while trying to maximize the efficiency of selection.

Background processes

Different backgrounds have similar final state, distributions and/or signatures than the ones expected for our signal. Here are the most important ones.

- $t\bar{t} \rightarrow$ is our main background and has kinematics close to the expected ones for our signal. The biggest challenge of the analysis consists in finding ways to reduce this background while leaving the signal as it is.
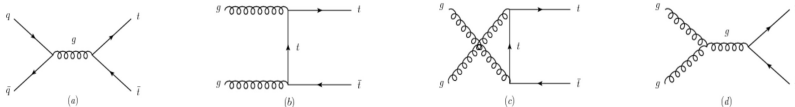
- $tW^{\pm} \rightarrow$ having a lower cross section and different kinematics, this background has a lower impact on the final results, even though it features a two b-jets and missing transverse energy in the final state as well.

- ttV (ttW/ttZ) $\rightarrow$ having two top quarks in the final state, some of this process is also expected to survive the selection of the analysis and to have an impact on the final results.

- Other backgrounds $\rightarrow$ such as the Drell-Yan ($Z/\gamma*$) or the non-prompt background also have different kinematics and only a limited impact on the final results.

Monte Carlo samples

- The different backgrounds have either been simulated with **Monte Carlo simulations** and checked in **control regions** (regions enriched in a single background) or measured using **data-driven methods** (estimated directly from the data).

- The dark matter samples have also been generated using Monte Carlo simulations. Different samples have been produced, for different **dark matter mediator masses** (10, 20, 50, 100, 200, 300 and 500 GeV), considering both the **scalar and pseudoscalar mediator couplings**.

Background	σ [pb]
$t\bar{t} \to ll \, \nu\nu$	87.31
Single top	35.6
Single antitop	35.6
$t\bar{t}W \to l \, \nu$	0.2043
$t\bar{t}W \to q\bar{q}$	0.4062
$t\bar{t}Z \to ll \, \nu\nu$	0.2529
$t\bar{t}Z \to q\bar{q}$	0.5297
$Z \to ll$ (low m_{ll})	18610
$Z \to ll$ (high m_{ll})	6025.2
$W\gamma \to l \, \nu \, \gamma$	586
$Z\gamma \to ll \, \gamma$	131.3

Scalar mediators	
m_S [GeV]	σ [pb]
10	19.950
20	10.480
50	2.941
100	0.672
200	0.093
300	0.029
500	0.005

Pseudoscalar mediators	
m_P [GeV]	σ [pb]
10	0.441
20	0.399
50	0.303
100	0.191
200	0.084
300	0.040
500	0.005

MC samples used

Event reconstruction

Reconstruction of the objects

A particular algorithm, the Particle Flow, is usually used in order to **gather the information** coming from the different parts of the detector and to **reconstruct the tracks** of the different particles.
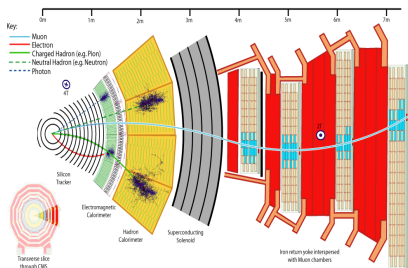
The geometry of the detector is quite simple. Usually, the z-axis is defined as the axis followed by the beams, meaning that the O_{xy} plane (also called Φ plane) corresponds to the transverse plane of the detector.

The pseudorapidity η is a Lorentz invariant quantity defined as $-ln\left(tan(\frac{\theta}{2})\right)$ where θ is the angle between the z- and x-axes.

In this analysis, only **isolated tight leptons** (complying with a lot of different requirements, resulting in better leptons) having a p_T larger than 10 GeV and $|\eta|$ smaller than 2.5 (2.4) for electrons (muons) are considered. Jets are required have a p_T larger than 30 GeV and $|\eta|$ smaller than 4.

Top reconstruction

General idea

The top reconstruction is a general method developed to calculate the p_T of the tops in any $t\bar{t}$-like event. We extended this method so that it is also **able to measure the p_T of the mediator** produced in the dark matter case.

Having a way to determine the p_T of the mediator is interesting, because this variable is expected to **give us a good discrimination** between the $t\bar{t}$ and our signal.
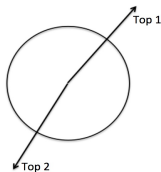
In the usual $t\bar{t}$, both the tops are expected to leave the primary vertex back-to-back.
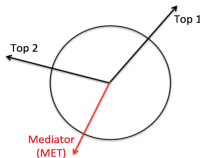
In the $t\bar{t}$ + DM case, the mediator leaves the primary vertex with a p_T equal to the difference in p_T of the tops.

An event passing the usual top reconstruction gets a value of dark p_T exactly equal to 0, while an event passing moreover our improved reconstruction (and therefore more likely to be a signal event) gets a value higher than 0.

Event preselection

We want to be able to reduce as much as we possible can the background, while keeping as much signal as possible by applying several cuts on different variables. This is first done by cutting in several general variables.

Number	Cut level	Cut	Comment
0	Preselection	$p_T^{lep_1} > 25$ GeV $\\ p_T^{lep_2} > 20$ GeV $\\ p_T^{lep_3} < 10$ GeV $\\ q_l^{lep_1} \cdot q_l^{lep_2} < 0$	On the leading lepton On the trailing lepton Third lepton veto Opposite charge requirement
1	First level	$m_{ll} > 20$ GeV $\\ |m_{ll} - m_Z| > 15$ GeV $\\ n_{jet} \geq 2 \\ n_{b_{jet}} \geq 1$	Only applied to ee and $\mu\mu$ channels

All these cuts allow us to reduce strongly different backgrounds (such as the Drell-Yan). However, we still need to find other variables able to remove some of the $t\bar{t}$, the most problematic background.

Discriminating variables

In the following plots have been represented the distributions of the number of event for four different variables. These plots have been obtained by applying the cuts corresponding to the preselection of the analysis, and both the 10 (red line) and 500 GeV (brown line) scalar signals have been represented (and **rescaled** by a factor 10 and 1000, respectively).

Dark p_T

m_{T2}^{ll}
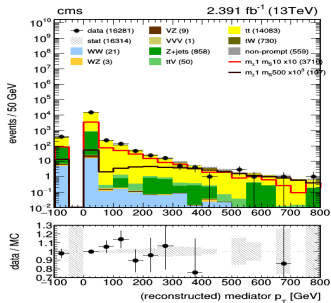
Reconstructed mediator p_T, obtained by applying our top reconstruction method
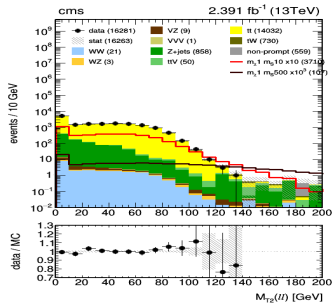
Transverse mass of the pair of leptons produced (cf. backup)

Discriminating variables II

E_T^{miss}
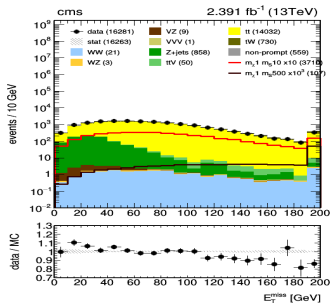
Missing transverse energy, equivalent
to the imbalance of vector momentum
in the O_{xy} plane

$\Delta\Phi_{ll, E_T^{miss}}$
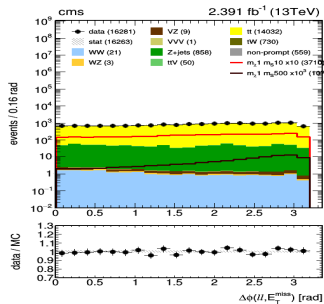
Angle between the two leptons coming from
the top quarks and
the E_T^{miss} in the O_{xy} plane

Our complete analysis actually lays on the discriminating power of four different variables that will
be used as input for our different neural networks, as we will see later.

Final event selection

We can then use the previous discriminating variables to introduce a third level of selection, to gain even more in purity.

Number	Cut level	Cut	Comment
0	Preselection	$p_T^{lep_1} > 25$ GeV $p_T^{lep_2} > 20$ GeV $p_T^{lep_3} < 10$ GeV $q_l^{lep_1} \cdot q_l^{lep_2} < 0$	On the leading lepton On the trailing lepton Third lepton veto Opposite charge requirement
1	First level	$m_{ll} > 20$ GeV $|m_{ll} - m_Z| > 15$ GeV $n_{jet} \geq 2$ $n_{b_{jet}} \geq 1$	Only applied to ee and $\mu\mu$ channels
2	Second level	$E_T^{\mathrm{miss}} > 80$ GeV $m_{T2}^{ll} > 80$ GeV Dark $p_T > 0$ GeV	Our top reconstruction works

Multi-Variate Analysis

Performing a general **cut and count analysis** (applying several cuts on different variables to check for the yields obtained) is not optimal when looking for dark matter, since the **significance** of the signal is expected to be really low.

We then need to come up with solutions to **extract as much information as we can** from the data. This is why we decided to perform a **Multi-Variate Analysis**, by using general machine learning techniques and neural networks.

The main idea of the MVA consists in using a **neural network** to combine the information coming from the four previous variables into a single output.

This should give a way to classify any single event as either background or signal, after training correctly the networks created.
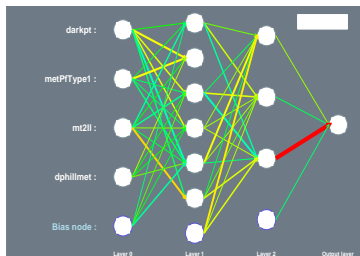
Artificial Neural Network

Our MVA is performed by defining several different **neural networks**.

- Each network created contains **two hidden layers** made out of respectively 6 and 3 neurons. We considered two different kind of neurons, having different **activation functions**.
- We train a **different network for each mediator mass** point considered, and for both the scalar and pseudoscalar mediators. The neural networks are trained using 200 events and are trained for the moment only against the $t\bar{t}$.

10 GeV scalar mediator
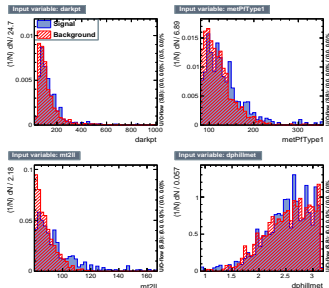
500 GeV scalar mediator
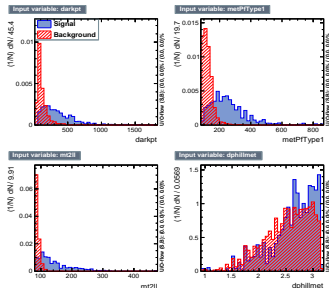
Input distributions

Artificial Neural Network II

It is also important to check the presence of eventual **correlations** between the input variables, to make sure not to complicate the networks without valid reason.

$m_\Phi = 10$ GeV

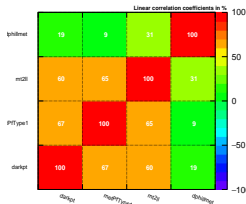Correlation Matrix (background)
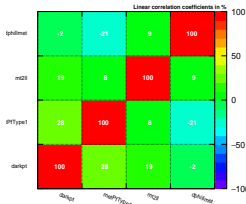
Correlation Matrix (signal)
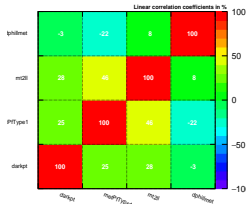
The strongest correlation is obtained for the signal, between the m_{T2}^{ll} and the E_T^{miss} variables (46 %).

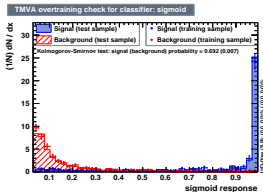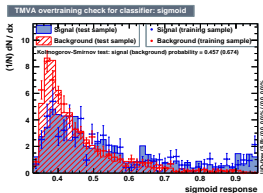$m_\Phi = 500$ GeV

Correlation Matrix (background)
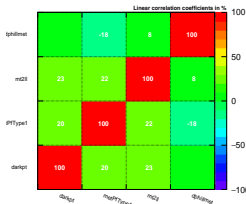
Correlation Matrix (signal)

The strongest correlations are obtained for the signal, between the m_{T2}^{ll}, the E_T^{miss} and the dark p_T variables.
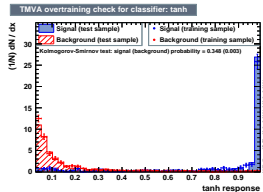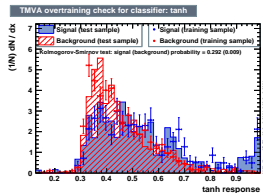
Artificial Neural Network III

Finally we can study the **actual output** of the different networks considered so far which is, at least for the moment, used in a simple count and count analysis.

Activation function 1

Activation function 2

$m_\Phi = 10$ GeV

$m_\Phi = 500$ GeV

We clearly observe a **better discrimination** of the output in the 500 GeV case, and the Kolmogorov-Smirnov test indicates that the first activation function output presents less overtraining.

Systematic uncertainties

Several sources of **systematic uncertainties** have been considered, and can be classified into two different categories. The estimation of these systematics is a crucial point of the analysis since we need to get a precise idea about the error we are committing if we want to be able to make some conclusions about the eventual existence of dark matter.

- The **theoretical uncertainties**, such as the Parton Density Functions normalization ($< 5\%$) and shape ($< 3\%$), and the high-order corrections of the MC simulations uncertainties.

- The **experimental uncertainties** are usually coming from the detector itself. In this category are grouped several different uncertainties, such as:

  ▶ Determination of the luminosity
  ▶ Lepton RECO and ID normalization ($< 2\%$) and shape ($< 1\%$)
  ▶ B-tagging efficiency normalization ($< 7\%$) and shape ($< 7\%$)

  ▶ Jet energy scale normalization ($< 4\%$) and shape ($< 10\%$)
  ▶ E_T^{miss} modeling and resolution
  ▶ Non-prompt (30%), $t\bar{t}$ (15%) and Drell-Yan normalizations (7%)

The other backgrounds estimated directly from Monte Carlo get an usual 50% systematic uncertainty associated. All of the previously mentioned systematics have of course been taken into account when plotting the limits of the analysis (and for most of them, we did considered both their normalization and shape).

Yields table

We can now represent the yields obtained for the different backgrounds with the associated systematics, for the different background processes and different dark matter mediator mass points (and therefore, different neural networks).

10 GeV scalar (ANN output > 0.75)			
Process	Yields	Statistical error	Systematic error
$t\bar{t}$	11.08	0.18	1.72
Non-prompt	1.30	0.92	0.39
ttV	0.75	0.03	0.19
Single top	0.48	0.07	0.16
Drell-Yan	0.46	0.35	0.34
WW	0.02	0.02	0.02
Total background	14.09	1.00	1.81
Signal	3.21	0.18	2.28
Total with signal	17.30	1.02	2.91
Data	10	-	-

500 GeV scalar (ANN output > 0.98)			
Process	Yields	Statistical error	Systematic error
$t\bar{t}$	1.82	0.14	0.28
Non-prompt	0.00	0.00	0.00
ttV	0.57	0.03	0.15
Single top	0.12	0.06	0.04
Drell-Yan	0.00	0.00	0.00
WW	0.01	0.01	0.01
Total background	2.52	0.16	0.32
Signal	0.03	0.01	0.01
Total with signal	2.55	0.16	0.32
Data	2	-	-

$\rightarrow$ The yields tables for the pseudoscalar mediator can be found in the backup.

Scalar limits

We can now represent the **limits** of the analysis, for both the scalar and pseudoscalar mediators.

The limit plots consist in plotting the value of the upper limit on the cross-section for which we expect to get some sensitivity, divided by the theoretical cross-section. This way, obtaining a value smaller than 1 for a given mass point then means that with the luminosity considered, we would **expect to get some sensitivity** to the signal.

Scalar mediator
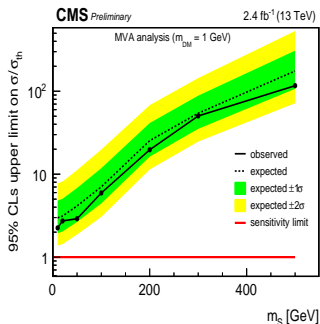
m_S [GeV]	Best AAN cut	Expected limit	Observed limit
10	0.75	2.99	2.22
20	0.80	3.12	2.73
50	0.90	4.14	2.90
100	0.90	7.03	5.94
200	0.90	25.37	19.69
300	0.80	54.25	50.34
500	0.98	175.50	116.65

The green and yellow lines on the plot correspond to the 1 and 2σ confidence intervals for the value of the expected limits.

Pseudoscalar limits

Pseudoscalar mediator
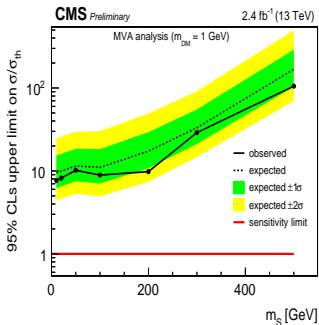
m_P [GeV]	Best ANN cut	Expected limit	Observed limit
10	0.80	9.47	7.66
20	0.85	9.94	8.18
50	0.85	11.47	10.18
100	0.90	11.03	8.91
200	0.95	17.31	9.81
300	0.80	33.12	28.98
500	0.98	167.94	105.27

With our analysis and considering for the moment the 2.4 fb^{-1} blinded dataset, we **cannot exclude** any dark matter production model.

However, we expect to be **able to exclude the 10 and 20 GeV scalar mediators** once considering the complete 2016 dataset, since we know that our sensitivity scales with the change in luminosity.

Conclusions

A search for dark matter production in association with top quark pairs in the dilepton final state has been performed. The analysis has been made using a 2.4 fb^{-1} blinded dataset for the signal region, and the complete 35.9 fb^{-1} dataset for the control regions, both taken by the CMS detector at CERN, at a center of mass energy of $\sqrt{s} = 13$ TeV.

The results obtained **do not allow us to exclude any dark matter production model** so far, with the limited luminosity considered, but we do **expect to be able to exclude the 10 and 20 GeV scalar mediators** when unblinding the analysis.

We did think about several ways we have to improve the analysis. First of all, a **better understanding** of the different **backgrounds** (mainly the $t\bar{t}$) and the different **systematics** is a crucial point since studying them in more detail and reducing them would allow us to improve our results. Moving to a complete shape study of the output of the MVA instead of performing a simple cut and count analysis is also expected to improve the final results, and will be done as soon as possible.

Thank you
for your attention!

Any questions?

Top reconstruction addendum

We can get 6 different equations from the kinematics expected for our signal.

$$\begin{cases} M(b_1 + W_1) = M_t \\ M(b_2 + W_2) = M_t \end{cases}$$

$$\begin{cases} M(\nu_1 + l_1) = M_W \\ M(\nu_2 + l_2) = M_W \end{cases}$$

$$\begin{cases} \nu_{1_x} + \nu_{2_x} = (E_T^{miss})_x \\ \nu_{1_y} + \nu_{2_y} = (E_T^{miss})_y \end{cases}$$

...because both the tops produced decay to a bottom and a W boson.

...because we only consider the leptonic decay of the W in this analysis.

...at least if we assume that the E_T^{miss} is coming from the neutrinos only.

$\rightarrow$ We have 6 equations and 6 unknowns (the three components of the momenta of the two neutrinos), meaning that this is a problem which can be solved. However, we have to find a way to correct the efficiency drop appearing when considering the signal, since the last hypothesis is not correct in this case (the E_T^{miss} is not coming only from the neutrinos).

This variable is expected to give some discrimination between the usual $t\bar{t}$ and our signal because in the dark matter case, we expect the p_T of the tops to be higher, since they will be accommodating for the extra missing transverse energy which appears (at least, if we manage to correct the efficiency drop issue since in this latest case, the previous equations are actually not exactly true).

E_T^{miss} variable definition

The E_T^{miss}, the missing transverse energy, corresponds to the imbalance of vector momentum in the plane perpendicular to the beam direction. It is defined in the following way:

$$E_T^{\text{miss}} = - \sum_i \overrightarrow{p}_T(i)$$

We have to consider the missing <u>transverse</u> energy and not directly the missing energy simply because the LHC is a proton-proton collider, and because each proton has three quarks (and the totat energy is distributed between these quarks in unknown proportions). This means that, for any given collision, we cannot actually know the exact initial longitudinal energy of collision. However, we do know that the initial transverse energy of the collision is equal to 0, so we can only calculate differences in transverse energy.

We expect this variable to give some discrimination between the $t\bar{t}$ and our signal because we know that the eventual dark matter particles produced should escape the detector while being undetected, since they almost do not interact with ordinary matter. The only way to detect the production of this kind of particle is therefore to look at the missing transverse energy.

$\Delta\Phi_{ll, E_T^{\mathrm{miss}}}$ variable definition

This variable corresponds to the angle in the Φ plane, between the two leptons coming from the tops and the missing transverse energy.

| $t\bar{t}$ background | $t\bar{t}$ + DM (low m_Φ) | $t\bar{t}$ + DM (high m_Φ) |
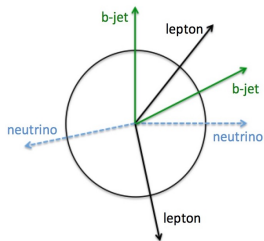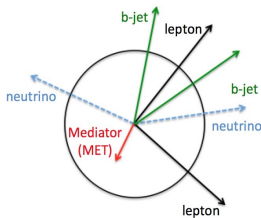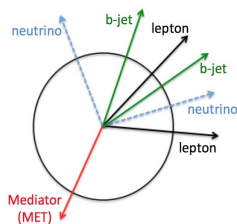
We expect that the tops will be much closer to each other when the mass of the mediator produces raises, so this variable is expected to give some separation between the $t\bar{t}$ and the signal we are looking for, at least for high mediator masses.

m_{T2}^{ll} variable definition

The variable m_{T2}^{ll} has been introduced to measure the mass of a pair of particle produced in the particular case where both these particles decay into a final state including an undetected particle (neutrinos in this case).

The problem is that in this case we can only know the total amount of E_T^{miss} of the event considered and there is no way to know the exact contribution to this value given by each neutrino separately. We then need to calculate the transverse mass for the two pairs of particles, for different repartitions of E_T^{miss}. The repartition giving rise to the smallest possible mass is kept as the value of m_{T2}^{ll}.

$$\begin{cases} \left(M_T^{(i)} \right)^2 = \left(m_{vis}^{(i)} \right)^2 + m_\chi^2 + 2 \cdot \left(E_T^{vis(i)} E_T^{\chi(i)} - \overrightarrow{p}_T^{\,vis(i)} \cdot \overrightarrow{p}_T^{\,\chi(i)} \right) \\ m_{T2} \left(m_\chi \right) = \min_{\sum_i \overrightarrow{p}_T^{\,\chi(i)}} \left[\max \left(M_T^{(1)}, M_T^{(2)} \right) \right] \end{cases}$$

We expect this variable to give some discrimination between $t\bar{t}$ and our signal because we know that the transverse mass obtained this way for the W has to take values less or equal than the actual mass of the W. However, the presence of additional E_T^{miss} coming for the dark matter can break this condition.

$t\bar{t}$ control region

A **control region** is usually defined by several cuts in order to check for the validity of a process simulated by Monte Carlo. The control region then has to be defined in such a way that this region is **enriched in the process we are interested in**.

This is why we created a control region for the $t\bar{t}$, our most important background. This region has been defined by removing the E_T^{miss} and by reversing the cut in m_{T2}^{ll}. As always, the 10 GeV signal is represented with the red line and has been rescaled by a factor 10.

The main objective of this control region is to calculate a **scale factor**, a factor by which we need to scale the considered process to get a perfect agreement between data and Monte Carlo.

In this case, we obtain a scale factor for the $t\bar{t}$ of 0.97 ± 0.01, the error being statistical only.
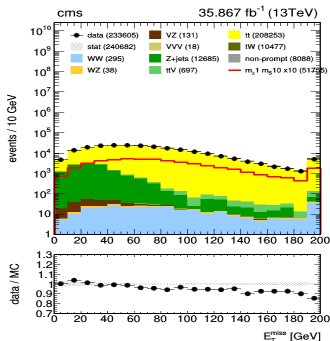
Rin-out method

To estimate the Drell-Yan (DY), we calculated its importance using a data-driven method. This method consists basically in estimating the number of DY outside of the peak of the Z directly from the number of data events inside of the peak.
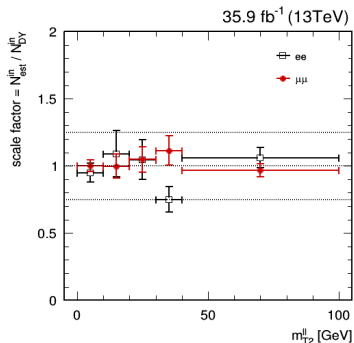
The scale factor obtained (1.07 ± 0.07) with this method has been applied to the usual MC simulations for the Drell-Yan process.
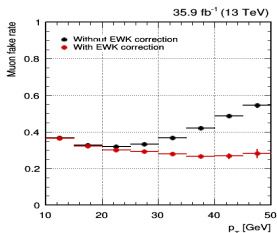
Fake rate

The fake rate corresponds to the probability to see a fake lepton (usually misidentified from a jet) passing the cuts corresponding to the final selection level of the analysis. This rate is estimated for both electrons and muons, within a QCD-enriched control region, by removing the contamination coming from the Z+jets and W+jets processes.
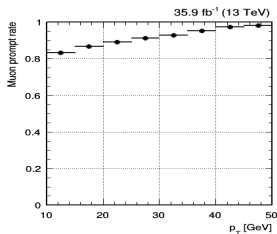
Electron FR
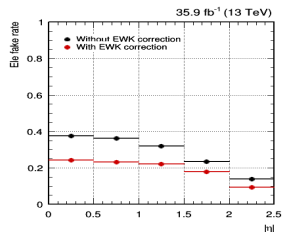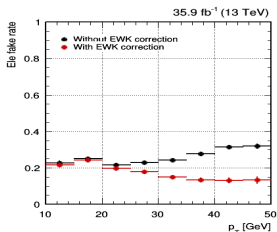
Muon FR
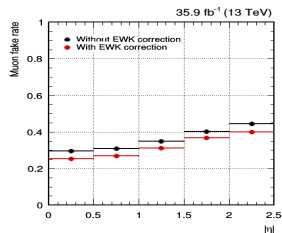
Prompt rate

The prompt rate is calculated thanks to a general tag and probe method and is useful to take into account the real lepton contamination in the control region we defined to calculate the fake rate.
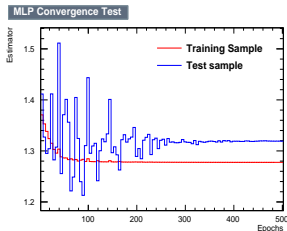
Electron PR
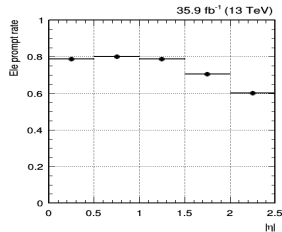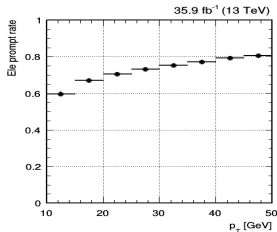
Muon PR
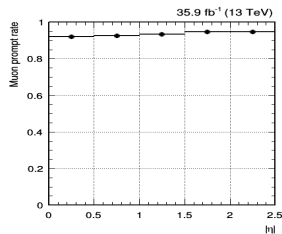
Overtraining issue of the ANN

We can check for eventual signs of overtraining for the different neural networks built. This can be done thanks to the convergence plots.
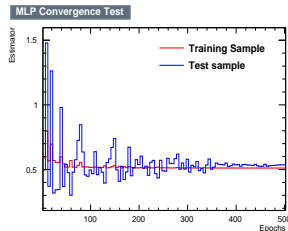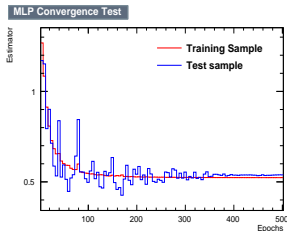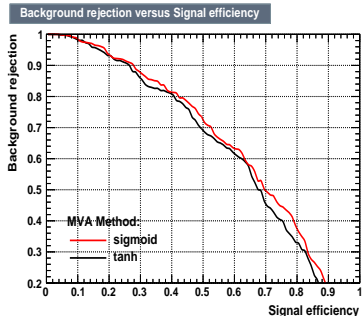
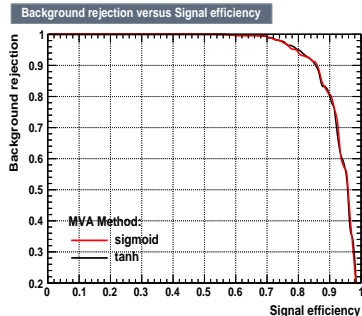$m_\Phi = 10$ GeV
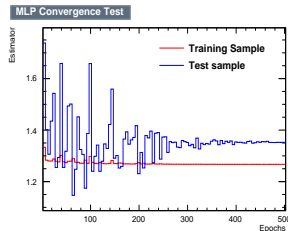
$m_\Phi = 500$ GeV

ROC curves of the ANN

The ROC curves represent the signal efficiency possible to achieve with the neural network considered, given any background rejection we are interested in. Here are compared the ROC curves obtained for both the sigmoid and tanh activation functions, and for both the 10 and 500 GeV networks. As we can see, we can achieve a much better background rejection for any given signal efficiency in the 500 GeV case.
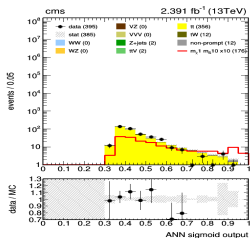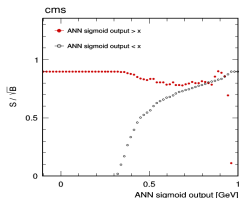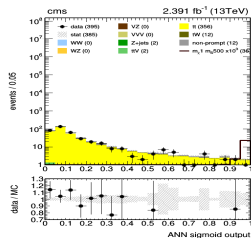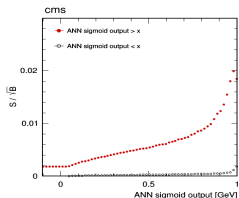
| 10 GeV network | 500 GeV network |

Output of the ANN

We can also study the output of the different neural networks, and its significance curve which are helpful to find the optimal cut to apply to the analysis, to get the best significance possible.

10 GeV network

500 GeV network

Yields table (pseudoscalar mediator)

We can now represent the yields obtained for the different backgrounds with the associated systematics, for the different background processes and different dark matter mediator mass points (and therefore, different neural networks).

10 GeV pseudo (ANN output > 0.80)

Process	Yields	Statistical error	Systematic error
$t\bar{t}$	8.28	0.15	0.42
Non-prompt	0.30	0.65	0.10
ttV	0.74	0.03	0.18
Single top	0.35	0.06	0.01
Drell-Yan	0.39	0.28	0.12
WW	0.02	0.02	0.01
Total background	10.08	0.72	0.48
Signal	0.89	0.02	0.45
Total with signal	10.97	0.73	0.66
Data	8	-	-

500 GeV pseudo (ANN output > 0.98)

Process	Yields	Statistical error	Systematic error
$t\bar{t}$	0.81	0.05	0.04
Non-prompt	0.10	0.31	0.03
ttV	0.46	0.02	0.12
Single top	0.04	0.02	0.01
Drell-Yan	0.00	0.00	0.00
WW	0.01	0.01	0.01
Total background	1.42	0.32	0.13
Signal	0.03	0.01	0.01
Total with signal	1.45	0.32	0.13
Data	1	-	-

Complete 2016 dataset expected limits

From the 2.4 fb^{-1} limits we calculated, it is actually easy to estimate the expected limits when considering the complete 35.9 fb^{-1} dataset, since we know that the limits are expected to decrease as the square root of the increase in luminosity. An increase of a factor 15 in the luminosity should then result in limits 4 times smaller.

Scalar mediator		Pseudoscalar mediator	
m_S [GeV]	Expected limit	m_P [GeV]	Expected limit
10	0.77	10	2.44
20	0.81	20	2.57
50	1.07	50	2.96
100	1.82	100	2.84
200	6.55	200	4.47
300	14.00	300	8.55
500	45.31	500	43.36

By considering the complete 2016 dataset, we therefore expect to be able to exclude with our analysis the 10 and 20 GeV scalar mediators.