# UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

## Colegio de Ciencias e Ingenierías

## Top Quark radius measurement with the CMS experiment
### Proyecto de investigación

## Jorge Stalin Martínez Armas

## Licenciatura en Física

Trabajo de titulación presentado como requisito
para la obtención del título de Licenciado en Física

Quito, 02 de diciembre de 2016

# UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

# COLEGIO DE CIENCIAS E INGENIERÍAS

## HOJA DE CALIFICACIÓN

## DE TRABAJO DE TITULACIÓN

### Top Quark radius measurement with the CMS experiment

# Jorge Stalin Martínez Armas

Calificación:

Nombre del profesor:     Edgar Carrera Ph.D.

Firma del profesor:                  ---------------------------------------

Quito, 02 de diciembre de 2016

# DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en el Art. 144 de la Ley Orgánica de Educación Superior.

Firma del estudiante: _____

Nombres y apellidos:    Jorge Stalin Martínez Armas

Código:            00112098

Cédula de identidad:    172204440-9

Lugar y fecha:        Quito, 02 de diciembre de 2016

# Resumen

En el Modelo Estándar, todas las partículas fundamentales son consideradas puntuales ¿Que pasaría si el quark top tuviera un radio $(R_t)$ y momento dipolar magnético $(\kappa_V)$ anómalo? En este análisis, simulaciones al nivel de generador son usadas para definir observables experimentales sensibles a la presencia de un radio $(R_t)$ y momento magnético anómalo $(\kappa_V)$. Luego, un análisis estadístico es desarrollado para poner límites a dichos parámetros. Usando aproximadamente 4 fb$^{-1}$ de datos de colisiones protón-protón adquiridos usando el experimento CMS en 2016, se excluye de manera preliminar $R_t > 0.001$ TeV$^{-1}$ a 95% CL.

*Palabras clave:* Modelo Estándar, quark top, radio anómalo, momento magnético dipolar anómalo, límite esperado.

# Abstract

In the Standard Model, all particles are considered to be point-like. What would happen if the top quark had an anomalous radius ($R_t$) and magnetic dipole moment ($\kappa_V$)? In this analysis, generator-level simulations are used to define experimental observables sensitive to these two parameters, and later, a statistical analysis is performed to set limits on their contributions. Using up to 4 fb$^{-1}$ of proton-proton collision data acquired by the CMS experiment in 2016, we preliminary exclude $R_t > 0.001$ TeV$^{-1}$ at 95% CL.

*Key words:* Standard Model, top quark, Anomalous radius, Anomalous magnetic dipole moment, Expected limit.

# AGRADECIMIENTOS

*De manera especial agradezco a mi familia. Agradezco todo lo que me han enseñado y los valores que me han inculcado. Este trabajo y yo, son un ejemplo de todo lo que he vivido con ustedes y se los agradezco infinitamente.*

*Agradezco a la Universidad San Francisco de Quito, en donde estudié la carrera que más amo, física. En esa universidad tuve la oportunidad de conocer a algunos de los mejores profesores que he tenido, y conocer a algunos de los amigos y amigas que más valoro.*

*Finalmente, agradezco a mis supervisores Martijn Mulders y Pedro Silva, por su enseñanza y su paciencia. Y de manera muy especial a mi supervisor Édgar Carrera, que me acompañó durante toda la tesis. Desde los inicios en donde no estaban las ideas muy claras, a lo largo de mi estadía en el CERN, y en los difíciles pero entretenidos meses finales de este trabajo.*

# Index

# Figure Index

# 1 Theoretical introduction and motivation

## 1.1 Introduction to the Standard Model of particle physics

The Standard Model is a theory that explains the interactions between the fundamental blocks of nature, called quarks and leptons, and three fundamental forces, which are the electromagnetic, weak, and the strong interactions. Each force is associated with a particle that mediates the interactions with quarks, leptons, or other mediators. The Standard Model divides particles into two big groups: fermions and bosons. The basic difference between these two kinds of particles is a quantum property called spin, which emerges whenever quantum mechanics and special relativity are combined. Fermions have half integer spins, while bosons have integer-valued spins.

From all the quarks and leptons in the Standard Model, only the electron and the up and down quarks are stable at our current temperature scale. The main difference between quarks and leptons is that quarks interact with the strong force, and the other ones do not. Quarks are divided into 3 generations, with 2 quarks in each one of them. These are up, and down for the first generation; charm, and strange for the second one; and top, and bottom for the third generation. Quarks have color, electric and flavor charge. Because they interact by the strong force, they are confined into colorless buckets called baryons and mesons. On the other hand, leptons are also divided into 3 generations. These are electrons, muons, taus, and their respective neutrino per generation. Leptons do not interact via the strong interaction and only electrons, muons, and taus have electric charge. Neutrinos only interact via the weak interaction, which it is why events involving neutrinos are difficult to observe.

Quarks and leptons make up almost all matter we are in touch with. Nevertheless, with only these particles, the universe would be unbelievably boring. The ingredients that make our universe so fascinating are the 4 fundamental forces: gravity, the electromagnetic force, and the strong and weak force. Gravity is not incorporated yet into the Standard Model; it is the aim of most theorists and phenomenologists to find a quantum theory of gravity. Electromagnetism and the weak force are two fragments of a more fundamental force, the electroweak interaction. The particles that mediate this interaction are the photon, the $W^{\pm}$, and $Z^0$. All of them are bosons of spin 1. Photons are massless while the $W^{\pm}$ and the $Z^0$ are very heavy. Photons are remarkably abundant and were the first kind of boson discovered. On the other hand, $W^{\pm}$, and $Z^0$ are the mediators of nuclear reactions such as the beta decay. The last force, the strong interaction, is mediated by 8 gluons; they can have either one of 8 different charges called colors. Quarks can have one of 3 colors and anti-colors. The gluons are massless and have spin 0. They act as springs inside the protons and neutrons, bounding them into stable particles. Gluons are the glue that hold various protons and neutrons together to make the stable elements. The

strong interaction is so intense that can overcome the electric repulsion of protons; however, the scope of this interaction is limited to the small distances of the atomic nucleus All particles in the Standard Model can be seen in Figure 1. It is important to comment that there is another form of matter called "antimatter". It shares the same properties of matter with the exception of charge, which is opposite to the one of normal matter. So, the Standard Model is made of the combination of normal matter as in Figure 1, and a similar table of antimatter.

Finally, the last ingredient of the Standard Model is the Higgs field, which gives mass to all fundamental particles of the Standard Model via the Brout-Englert-Higgs mechanism. In the 70's, physicists were concerned that the Standard Model was only successfully in the scenario of massless particles. Nevertheless, we, the Earth, the Sun, stars, and galaxies have mass. In classical mechanics, mass is defined in the context of Newton's laws; gravity, which was the first force humans could describe, involves mass or energy in each equation. So, it was very embarrassing that the Standard Model was not successful unless particles had mass, otherwise the theory collapsed. Peter Higgs and many other physicists suggested the existence of another fundamental field of nature, and the interaction of this field with all particles of the Standard Model (excluding photons and gluons) gives them mass.



Figure 1: Table showing all the fundamental particles of the Standard Model of particle physics [1].

## 1.2 Beyond the Standard Model

We know that the Standard Model is one of the most precise and exact theories nowadays, but there are some inconsistencies in the theory that are not fully explained. We mention the most important for the purpose of our work.

As we mention in section 1.1, gravity is not incorporated into the Standard Model. General Relativity, which is currently the theory that best explains gravity, it is not a quantum field theory. The graviton, which is the particle that would mediate the interaction, has not been observed in the laboratory. Indeed, the fact that the gravitational interaction is so weak is also another problem. The discrepancy between all other forces and gravity is still a mystery. One would suppose that in the early stages of the universe all interactions were unified, so all of them had the same intensity. However, the differences between each interactions are not explained by the Standard Model. These differences between forces come with another problem. Why are the masses of fundamental particles so different between each other? If we compute the predicted value for the masses of particles in the Standard Model, in most of the cases one gets values close to the Plank scale ($10^{19}$ GeV). It is possible to overcome the problem through a recipe, which is called renormalization, that in several cases gives a value near the measured one.

In 2012, the CMS and ATLAS experiments independently published the discovery of a new particle, compatible with the Higgs boson. The announcement showed a higgs mass in the order of the GeV, which was well below the Planck scale. This was a problem because if one computes the Higgs mass, its value turns out to lay on the order of the Planck scale; however, that mass value would be inconsistent with the measured masses of quarks and leptons. The cancellation of high order terms would either imply extreme fine tuning, or the existence of supersymmetric particles (or other fields), not discovered so far.

## 1.3 Top quark radius and anomalous magnetic moment

Now that the particle discovered at the LHC in 2012 has been proved to share similar properties with the Higgs boson, highlights of new physics have vanished. It is still unclear why the energy scales of the electroweak interaction are extremely low compared with the Planck scale. The top quark, which is the most massive fundamental particle yet discovered, shares the same problem. People do not know the mechanism that makes its mass so big compared to other particles (even the Higgs), but so low compared to the Planck scale. Due to all these questions about the connection of the top mass and the electroweak scale, different theories are trying to look at the top not as a point-like particle, as the rest of particles of the Standard Model, but as a linear mixing of states in the fermionic sector [2]. If proved, this would be a way to lower down

the electroweak scale to the one showed experimentally. So, new research on compositeness models of the top quark would derive in new physics.

In the context of effective field theories, the addition of an intrinsic anomalous radius $R_t$ and anomalous magnetic moment $\kappa_V$ means the introduction of high order operators in the Lagrangian. The following Lagrangians, denoted as $\mathcal{L}_R$ and $\mathcal{L}_\kappa$, represent the Lagrangians that characterize the presence of an anomalous radius and magnetic dipole moment for the top. In equation 1, the radius of the top is introduced for first time into the Standard Model while in equation 2, the magnetic dipole moment is introduced. In both Lagrangians the gluon field $G_\mu$ couples with the top field [2]:

$$\mathcal{L}_R = -g_s \frac{R_t^2}{6} \bar{t} \gamma^\mu G_{\mu\nu} D^\nu t + h.c. \tag{1}$$

$$\mathcal{L}_\kappa = g_s \frac{1}{4m_t} \bar{t} \sigma^{\mu\nu} \left( \kappa_V + i\kappa_A \gamma^5 \right) G_{\mu\nu} t \tag{2}$$

Here, $G_{\mu\nu} = D_\nu G_\mu - D_\mu G_\nu$ is the field strength and $D^\mu = \partial^\mu + ig_s G^\mu$ the usual covariant derivative. The term concerning the anomalous electric moment $\kappa_A$ is omitted because it is a higher order operator. For simplicity, only leading order operators are kept. Due to the presence of these terms in the Standard Model, an increase in the top cross section and its kinematics (extra radiation jets) is expected. In addition, the relation between these operators with the new physic scale is given by $\Lambda$ [2], which is going to be discussed in section 1.3.2:

$$R_t = \frac{\sqrt{6}}{\Lambda} \qquad \kappa_V = \rho_V \frac{m_t^2}{\Lambda^2} \tag{3}$$

### 1.3.1 Natural units

It is important to state that for the above Lagrangians and for the rest of this document, we are using natural units. In natural units, $h = c = 1$, so energy, momentum and mass are dimensionally equivalent. Length and time are also dimensionally equivalent, and are inversely proportional to energy. All of this is explained if we look the following equations.

$$E = mc^2 = pc \qquad E = h\nu = \frac{h}{t} \qquad x = ct$$

If $h = c = 1$, the above equations give the following results.

$$E = m = p \qquad E = \frac{1}{t} \qquad x = t$$

So, $[E] = [M] = [P]$, and $[L] = [T] = [E^{-1}]$. In particle physics, the preferred unit is the eV (1 eV $= 1.602 \times 10^{-19}$ J). For that reason, all physical quantities are in units of eV or eV$^{-1}$.

In equation 3, $\Lambda$ and $\rho_V$ have units of energy, so in this way, the dimensions of $R_t$ and $\kappa_V$ are inverse energy. It is important to keep in mind that $1\,\mathrm{TeV}^{-1} \approx 0.001$ fm.

### 1.3.2  The scale parameter $\Lambda$

In quantum field theory, the dimension of renormalizable operators is $4^1$. Any other corrections implying that high order operators (5, 6, and so on) can be incorporated to the theory together with a power of $\Lambda$. For example, if we have a Lagrangian of the form

$$\mathcal{L} = \Phi\partial_\mu\Phi\partial^\mu\Phi$$

Each $\Phi$ has dimensions of energy. Each $\partial_\mu$ counts as energy as well. Derivatives act as divisions, and every time we divide something by a spatial coordinate ($t$, or $\vec{x}$), we add another dimension of energy to the Lagrangian. In order to incorporate the Lagrangian above to the Standard Model, we have to multiply it by $\Lambda^\alpha$. Here, $\Lambda$ has dimensions of energy, and $\alpha$ is an integer. In order to make the appropriate correction, $\alpha$ must be -1. So, we get the following expression.

$$\mathcal{L} = \frac{1}{\Lambda}\Phi\partial_\mu\Phi\partial^\mu\Phi$$

The important fact to remember is that operators of dimensions higher than 4 are non-renormalizable. The introduction of this kind of operators create divergences in the theory, and divergences in a physical theory are non sensible. The origin of these divergences are linked with the strength of coupling constants[2] at different energy levels. For example, in the case of quantum electrodynamics QED, the coupling constant increases with energy, so higher non-renormalizable operators will introduce infinities when the energy is high enough. On the other hand, in quantum chromodynamics QCD, coupling constants decrease with energy. In order to introduce these "effective" Lagrangians, we use $\Lambda$ to create a cut-off which makes the term vanish at some energy level. Scale factors protect the theory from the running[3] of the coupling constants.

In the case of the present analysis, the effective Lagrangians introduced in equations 1 and 2 are of dimension 6. Therefore, the scale parameter $\Lambda$ is of power of -2. The importance of $\Lambda$ in effective field theories, such as the one in this analysis, is that the presence of these parameters in data is going to give us more information about the scale of energy if physics beyond the Standard Model emerges.

---

[1] Dimension up to 4 means that the exponent in the energy is 4 $\left[\mathrm{E}^4\right]$

[2] A coupling constant $g$ is a dimensionless number which describes the strength of an interaction. In the case of electromagnetism, $g = \frac{e^2}{4\pi}$. The replacement of dimensionless constants with dimensionful constants, such as $\Lambda$, is called dimensional transmutation [3].

[3] Running of coupling constants means the dependence of a coupling constants in the energy scale.

# 2 Experimental overview

In this thesis, we use data taken by the Compact Muon Solenoid (CMS). This experiment is one of the main detectors of the Large Hadron Collider (LHC). For that reason, a brief description of the LHC, the CMS experiment, and its subdetectors will be given. A transverse cut of the CMS experiment can be seen in Figure 2.

## 2.1 The Large Hadron Collider

The LHC is currently the biggest and most energetic accelerator in the world. It is made of 27 kilometers of superconducting magnets; and along with them, special structures that carry proton bunches, accelerating them at almost the speed of light. It sits underground between the border of Switzerland and France near the city of Geneve, and it is managed by CERN. The accelerator sends proton beams in different beam pipes in different directions (clockwise and anticlockwise), making them collide across 4 points along the ring. In each of those places sophisticated particle detectors are built: ATLAS, CMS, ALICE, and LHCb.

## 2.2 The CMS experiment

The Compact Muon Solenoid (CMS) is one of the biggest experiment in the world. It is made of several layers of different subdetectors that measure different physical quantities such as the momentum and energy of incoming particles. The reason why it is called compact is because it is relatively small compared to other experiments such as ATLAS. Nevertheless, it weighs almost 14,000 tons. We will give a briefly description of the main subdetectors. For a further explanation of the detector and each subdetector, reference [4].

### 2.2.1 The Tracker

The tracker is composed of thousands of silicon pixels and silicon strips that track the path of charged particles released in each collision. This information is used to reconstruct the momentum of those particles. Uncharged particles, such as neutrons and photons, will pass by the tracker unchanging their path. For charged particles, such muons and charged hadrons, their path is going to be bended by the powerful magnetic field generated by a solenoid. This curved path is precisely recorded by the tracker down to the order of a micron. The curvature of the path will provide the information of the momentum of the charged particle, which is fundamental in order to reconstruct the particle.

### 2.2.2 Electromagnetic calorimeter

The electromagnetic calorimeter or ECAL is made of lead-tungsten crystals that are arranged next to the tracker. They shine when an electron or a photon pass through them. All electrons and photons are supposed to deposit most of their energy there. The function of the ECAL is to measure the energy of high energy photons and electrons, as well as their position.

### 2.2.3 Hadronic calorimeter

The Hadronic Calorimenter HCAL is placed in the inner part of the CMS detector, just after the ECAL system, and it is made of brass and plastic scintillators. Hadrons, such as pions, kaons or any other hadron, which escape from the ECAL, are supposed to deposit their energy in the HCAL.

### 2.2.4 Superconducting magnet

The reason the CMS is called Compact Muon Solenoid is precisely because the huge super-conducting solenoid which is placed roughly in the middle of the detector. The function of the largest superconducting solenoid on Earth is to produce a magnetic field of 3.8 T inside the detector to try to bend, as much as possible, the path of high energy charged particles. The more energetic the particle, the less its path will be bent. In order to produce such magnetic field, the temperature of the solenoid is lowered to almost -217 °C, which makes the coil a superconductor. The superconducting magnet provides the magnetic field of the inner subdetectors of the CMS experiment: the tracker, the electromagnetic calorimeter, and the hadronic calorimeter.

### 2.2.5 Muon spectrometer

The muon spectrometer is placed at the outermost part of the detector. It weighs almost 12,000 tons. It is made of iron, and it tries to keep all particles inside the detector. Only muons and neutrinos are able to go beyond the muon spectrometer. Neutrinos, in fact, interact so little with matter that they can not be detected by any of the detectors in the CMS. They are merely accounted as missing energy, that is the energy that is not measured when you balance the incoming energy and the outgoing energy in each collision. The concept of missing energy will be defined in section 3.3.

Because the design of the CMS detector, muons are the easiest particles to detect.

### 2.2.6 Data acquisition and trigger systems

In the LHC almost 40,000 collisions happen each second. The amount of computational power to process and store that quantity of information is by far too large. However, only a small fraction of all the events are indeed potentially interesting. For those and other reasons, the

CMS detector incorporates a trigger decision system (Level 1 L1), which briefly takes the most interesting 100,000 events of the billions available. The rest are thrown away. The next trigger (High Level Trigger HLT) takes a bit of extra time in order to process complex physics algorithms on data and select around 1000 of the most interesting events from the previous 100,000. Even with triggers, each second thousands of storage drives are filled up.
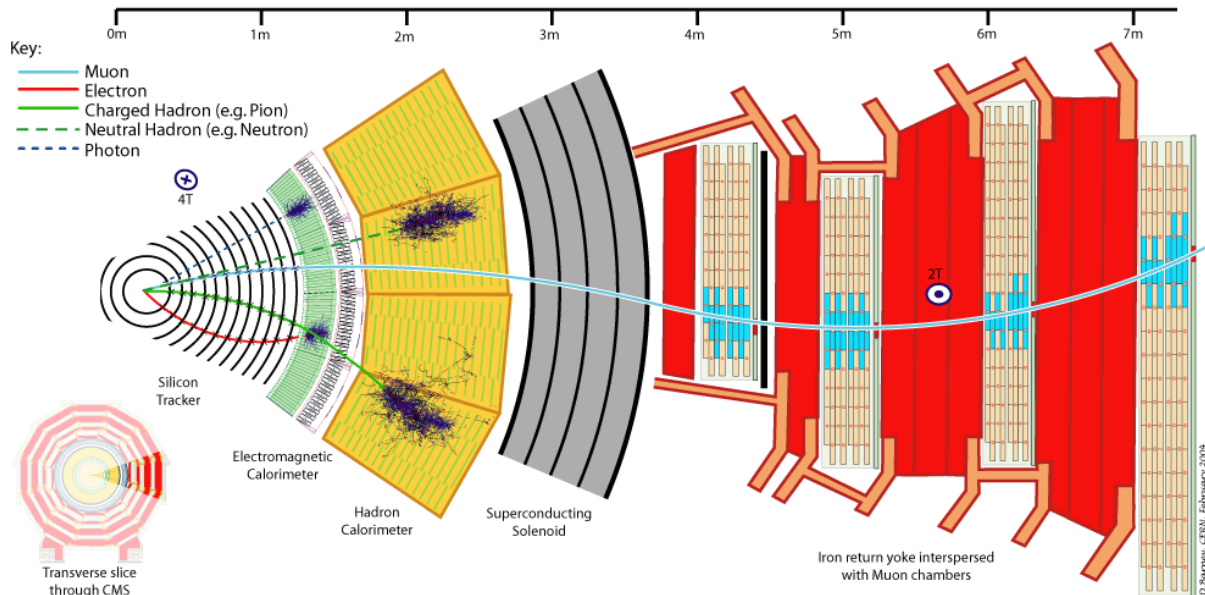


Figure 2: CMS experiment slice. Most of the subdetectors and their role on detecting particles is represented [5].

# 3   Event simulation and reconstruction

## 3.1   Selection criteria

In order to measure the radius of the top quark, we are interested in events involving the production of tops and anti-tops ($t\bar{t}$). A top can be the product of many kind of processes, and also, a lot of processes could look like a top. For the reasons explained above, we choose the dilepton
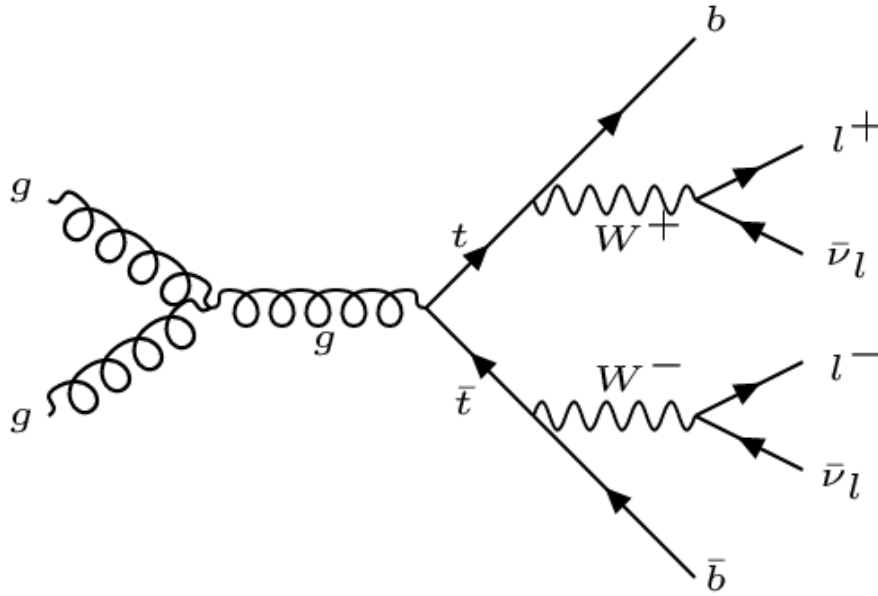


Figure 3: Top quark creation and decay in the dilepton channel.

channel (Figure 3) to reconstruct the top quark because there is not too much background [4] for it. This channel is obtained by the selection of 2 b-jets [5] and 2 leptons. The jets are b-tagged [6] and they must have a $p_T > 30$ [7] GeV and $|\eta| < 2.5$. On the other hand, leptons must have final state $p_T > 20$ GeV and $|\eta| < 2.5$.

Here, $\eta$ is called the pseudorapidity and is a coordinate variable that is defined as

---

[4]Any process different from signal that could be wrongly identified as signal.

[5]A jet is an avalanche of particles coming from a quark or gluon that was created after a collision. Because of confinement, quarks quickly decay into various kind of particles. The cluster of particles is called a jet. A b-jet is a jet which was originated by a b-quark.

[6]b-tagging states if a jet was identified as a b-quark. There are several algorithms that decide whether a jet comes from a b-quark.

[7]The transverse momentum is defined as follows:

$$\vec{p}_T = p_x\hat{x} + p_y\hat{y}$$

$$\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$$

In the formula above, $\theta$ is the angle between the momentum of the particle and the positive direction of the beam (polar angle in spherical coordinates) [6]. The beam travels at the center of the beam pipes. If we place the experiment horizontal, we define the z-axis along the beam pipe from left to right (Figure 4). The y-axis is defined upward, toward the top of the detector, and the x-axis is defined to the right as in the left picture of Figure 4.
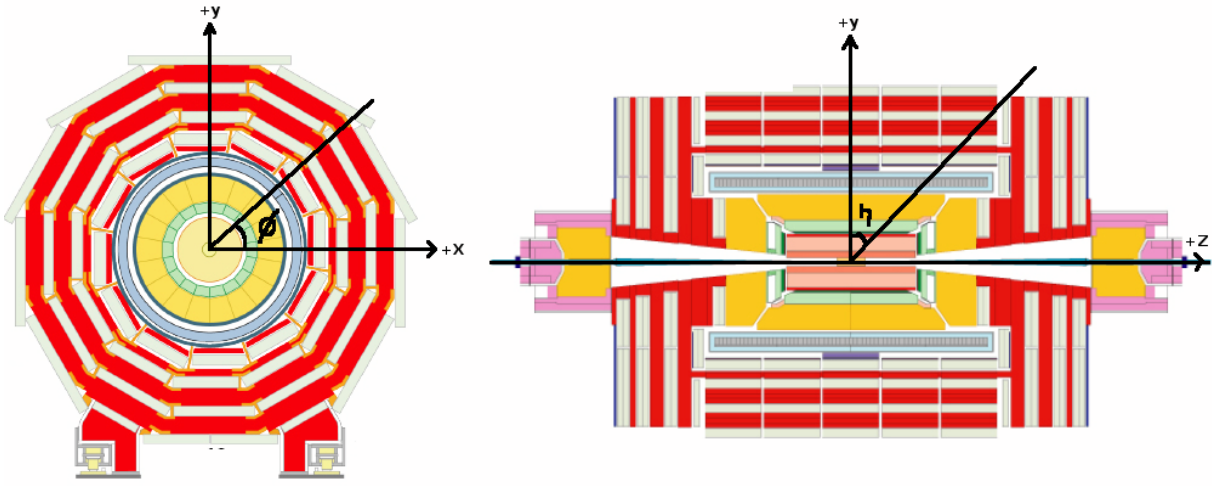


Figure 4: The coordinate system is defined with the help of these pictures. In the left picture, the CMS experiment is shown from a cut perpendicular to the direction of the beam (transversal plane). The right picture is a view of the CMS experiment from a horizontal cut through the experiment, along the direction of the beam pipe, which are the conduit where particles are traveling [7].

The pseudorapidity $\eta$ lays in the yz-plane. As it is seen in Figure 5, the smallest the pseudorapidity, the most perpendicular the particle will go out across the detector. $|\eta| < 2.5$ means that the particle should be inside a cone close to an axis perpendicular to the beam, which implies it has a high transversal momentum, so it is interesting.

Finally, for the purpose of this analysis, we need to define the angular variable $\phi$ with the role of the azimuthal angle. As seen in Figure 4, it is the angle between the x-axis to the transversal projection of the vector of the outgoing particle.

## 3.2   Simulation of anomalous radius and magnetic dipole moment

In order to create simulations of $t\bar{t}$ events implementing equations 1 and 2, we use the FeynRules software to get the Feynman rules associated with the new structure. We employ different
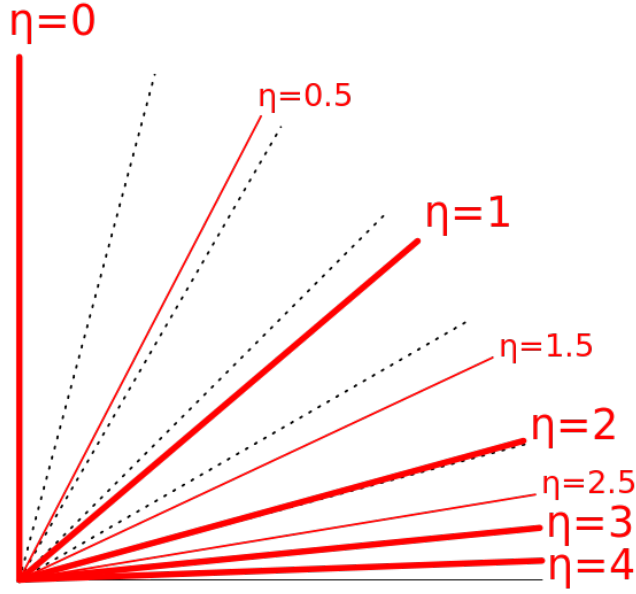
Figure 5: Pseudorapidity seen from the yz-plane. Dotted lines divide the image in same size pieces, and red lines count for different pseudorapidities [8].

type of matrix element calculators, but mostly the software MADGRAPH5 [9] in order to generate the events at parton level. We use PYTHIA8 [10] to shower [8] the events generated with MADGRAPH. In this step, radiation of initial and final states are taken into account; similarly, underlying events during the hadronization are simulated. To simulate events matching the selection criteria, final dilepton flavor states in the simulations were selected to be: electron-electron (EE), muon-muon (MM), and electron-muon (EM), which are the expected combinations of lepton flavors.

Implementing the steps above, we generated different simulations varying the value of the anomalous radius and magnetic moment. Nearly 62 simulations were created. The range of the radius ran from -0.005 to 0.005 $\text{TeV}^{-1}$. In the case of the anomalous magnetic moment, the range varied from -0.5 to 0.5 $\text{TeV}^{-1}$. The simulations files are stored in a specific area in the so-called EOS system at CERN.

$$\texttt{/eos/store/cmst3/user/psilva/CompositeTop}$$

## 3.3 Reconstruction

Reconstruction is a process where data collected from all the different subdetectors are used to identify the particle that left that signal. Using different algorithms, the linear momentum,

---

[8]Shower means to simulate the hadronization of quarks after the collision. Hadronization is a series of processes where quarks decay into more stable particles such as electrons, lighter hadrons, photons, etc.

the energy, the primary vertex, and the id of the particle (in case of simulations), among other variables, are computed.

Each kind of particle: muons, electrons, quarks, photons have their own signature in each subdetector. In the case of photons, their energy is totally absorbed in the ECAL without presenting any signal in the tracking system. In the case of electrons, their path is shifted by the magnetic field, so they deliver a signal in the tracking system. Then, in the ECAL, they are totally absorbed. So, it is only needed to have a track associated with an energy cluster in the ECAL in order to identify an electron. For quarks, because of confinement, they rapidly hadronize into more stable particles such as pions, photons, electrons, etc. Hadronized particles are absorbed gradually in the ECAL and the HCAL. If the hadron is charged, we should find a track, and energy deposits in the ECAL and the HCAL. Uncharged hadrons will be reconstructed matching clustered energy in the ECAL and the HCAL, one in front of the other. Muons are the easiest particles to detect in the CMS experiment. They deliver a track in the inner tracker and energy and an outer track in the muon chambers. Neutrinos do not interact with any subdetector, so their presence only can be inferred by missing energy in each event [9] [11].

As particles travel through the subdetectors, their energy is partially lost until they hit the ECAL or HCAL, therefore energy corrections are applied. These energy corrections are inferred via simulations of known events.

---

[9]Before two bunches of particles collide, the total transverse momentum (in the transverse plane of the detector) was initially zero, so after the collision the total transverse momentum, must be zero. When we reconstruct events, often the total transverse momentum is not zero because of particles that we can not detect as neutrinos. The transverse momentum is very important in particle physics because particles with high transversal momentum probably hold interesting events.

The total missing transverse momentum is defined as

$$\vec{P}_T^{miss} = -\sum_i \vec{p}_T(i)$$

Where $P_T^{miss}$ is the total transverse momentum and $\vec{p}_T(i)$ is the transverse momentum of particle $i$. The negative sign on the expression is written in order to have a total transverse momentum equal to 0 after the collision.

Sometimes, it is easier to just know the total missing transverse energy MET of a collision because it is an scalar. First, the transverse energy is defined as

$$E_T = \sqrt{m^2 + |\vec{p}_T|^2}$$

Just taking the transverse momentum. So, the MET is defined as

$$E_T^{miss} = -\sum_i E_T(i)$$

Interesting events involving neutrinos or any other unidentified particle will have a big MET.

In the case of the present analysis, we used reconstructed-level Monte Carlo simulations of $t\bar{t}$ events and their main backgrounds, such as Single top (tW), diboson (WW,WZ,ZZ), Drell-Yan ($Z \to l^+l^-$ Figure 6), $t\bar{t}V$ (V=W,Z), and W+jets. All Monte Carlo and data used are stored in the EOS system at CERN:

```
eos/store/cmst3/group/top/summer2016/TopWidth_era2016
```
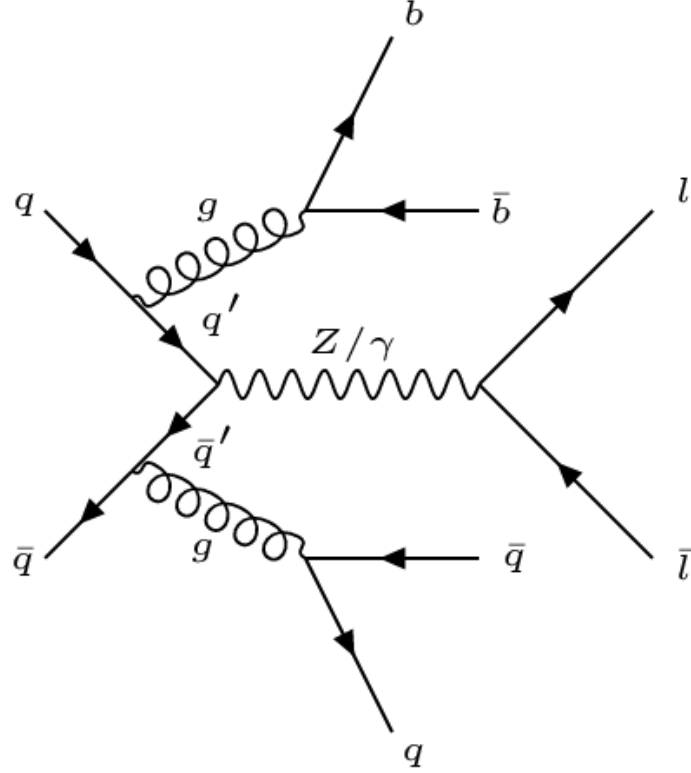


Figure 6: Example of a Drell-Yan event. Due to radiation, it exhibits similar final states required for our analysis. There are 2 leptons, 2 b-tagged jets, and because of chance, the two other jets might not be measured, so there is also missing energy. Drell-Yan background is the largest of all selected backgrounds.

The corresponding luminosity [10] generated for the above Monte Carlo is 1185 fb$^{-1}$, which is nearly 34 times the actual integrated luminosity of the experiment. We make simulations with a bigger number of events, in order to reduce statistical fluctuations that could wrongly lead to a false result [11] . Afterwards, for the purpose of the analysis, we can scale the integrated luminosity to the one present in the current experiment.[12] .

---

[10]Integrated luminosity is defined as the number of events over cross section ($\sigma$), $L = \frac{N}{\sigma}$.

[11]Statistical fluctuations occur when we do not have enough data. In 2015, the CMS and ATLAS experiments announced an excess of less than 3.4 and 3.9 standard deviation at 750 GeV in the diphoton channel. Many speculations were made toward the apparent discovery of a new and unexpected particle. Sadly, the excess vanished with the dataset of 2016, and the abnormality of 2015 was accounted as statistical fluctuations of the dataset. This is an example of why we have to be careful, as they indeed were in that analysis, with statistical fluctuations.

[12]The cross section is a measure of the probability that a certain process could happen, for example, AB into CD.

# 4 Data analysis

We use simulations of $t\bar{t}$ production, which incorporate the radius and the anomalous magnetic moment of the top quark, to try to find variables sensitive to the effects due to the presence of these elements. The purpose of these simulations is to try to model how the production would look like if a top with a radius were present. In order to characterize these simulations with respect to the Standard Model, we created several histograms of $t\bar{t}$ events for different variables. We tried about 36 different variables in order to find interesting, representative sample. We found many promising variables, but for a preliminary analysis, we used the most promising and simple ones in order to reconstruct simulations. The reason of the reconstruction is to find how we would measure the production of these events at the subdetectors of the CMS experiment. Finally, comparing reconstruction simulations with an accurate Monte Carlo simulation of the Standard Model, we test if actually one of those simulations could be detected at the experiment with the current number of events at the LHC. For this, we use sophisticates statistical methods to set limits on the radius of the top quark, if it exists.

## 4.1 Simulation and identification of sensitive variables

First, we use the simulations to make several histograms of different variables in order to find the most sensitive ones. This means, we want variables whose histograms have significant differences from the Standard Model. After plotting several of them, we found that angular variables suffer appreciable changes, specially $\Delta\phi\left(l,l'\right)$ [13] and $\Delta\phi\left(b,b'\right)$. We also explore more complex variables in different frames, but as a preliminary analysis we just keep the simplest ones.

In Figure 7, $\Delta\phi\left(l,l'\right)$ and $\Delta\phi\left(b,b'\right)$ are plotted for the case $R_t = -0.005$ TeV$^{-1}$ and $\kappa_V = 0.00$ TeV$^{-1}$. Note that at generator-level, both histograms show an appreciable change, specially at angles close to $\pi$. More analyzed variables can be seen in the Annex A.

## 4.2 Reconstruction of signal using sensitive variables

Once we select the variables, we try to re-weigh Standard Model $t\bar{t}$ Monte Carlo, so we could emulate the reconstructed-level MC for the signal. In order to perform the re-weighing process, we made a serialization (Figure 8), which consists in measuring $\Delta\phi\left(l,l'\right)$ in categories of $\Delta\phi\left(b,b'\right)$. The serialization process can be seen in Figure 8. Serialization helps us to plot a 2D

---

Where AB are the particles in the initial state, and CD are the particles in the final state. The dimension of cross section is area (the bigger the "effective area" for a process to occur, the bigger the probability of that process). Because 'areas' in particle physics are very small, it is often used the unit of a "barn" $b = 10^{-28}$ m$^2$. A femtobarn is fb $= 10^{-15}$b.

[13]$\phi$ is the azimuthal angle measured from the +x-axis in the xy-plane. $\Delta\phi$ is azimuthal angle between two different vectors (in this case particles).
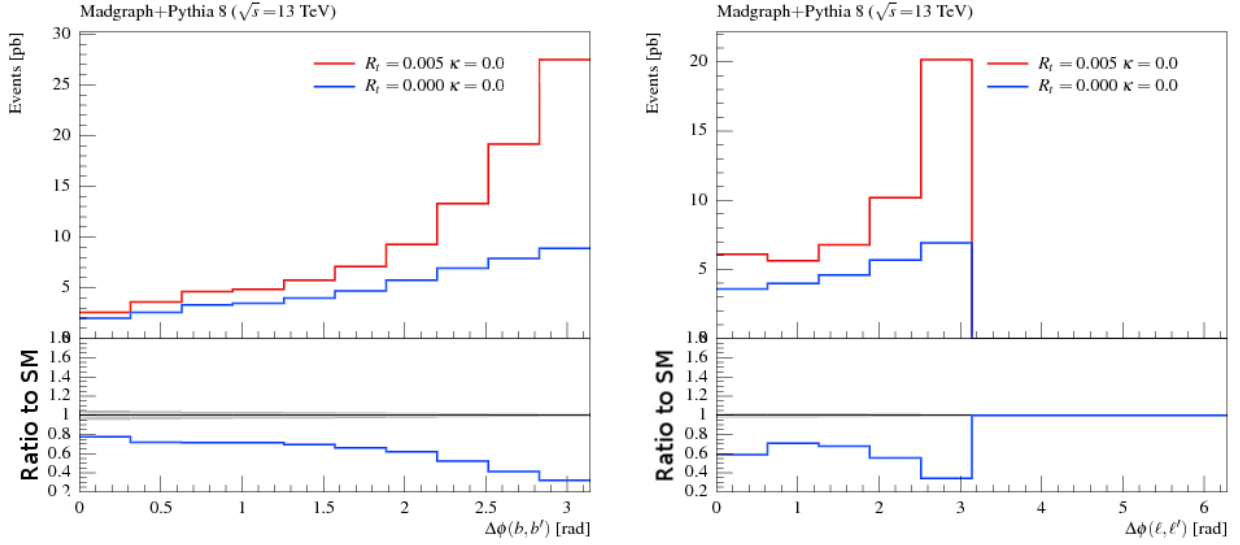
Figure 7: Histograms of selected variables comparing signal and Standard Model.

histogram into a 1D. Using serialized histograms, we compute weights for each bin $w_i$ in the histogram using equation 4.

$$w_i = \frac{[n_i]_s}{[n_i]_b}, \tag{4}$$

where $[n_i]_s$, and $[n_i]_b$ are the number of events in the bin $i$ for signal and for background, respectively. These weights will help us get Monte Carlo simulation of data, scaling a Monte Carlo simulation of the Standard Model.

Histograms in Figure 9 are examples of two different reconstructions of our set of simulations. There, two re-weighed models (there are 4 plots; they will be explained below) are compared against the Standard Model Monte Carlo of $t\bar{t}$ showed in gray, and experimental data points, which are the black dots shown in the plots. The red line is the re-weighed signal, which represents a model with some value of $R_t$ and $\kappa_V$. For Figure 9, the graph in the top right represents a model where $R_t = 0.005$ TeV$^{-1}$ and $\kappa_V = 0.00$ TeV$^{-1}$. Because the black dots are data, the signal simulated at reconstruction level is clearly far beyond the data measured. This implies that at current luminosity (4 fb$^{-1}$ in this analysis) the signal does not represent what we observe, so it can be excluded. On the other hand, the top left graph represents a model where $R_t = 0.001$ TeV$^{-1}$ and $\kappa_V = -0.25$ TeV$^{-1}$. This model, simulated at reconstruction level, matches data almost perfectly. We have to recall that in this model, part of the events were attributed to background (SM) and part to signal (top substructure). In order to evaluate the model, we need to calculate the signal strength $r$ to get a measure of how likely we can exclude the model. The signal strength will be explained in section 4.3.

For the bottom plots in Figure 9, the difference between them and the first plots is that for the first plots, we implemented a cut to select electron-muon (EM) as lepton final states. For the bottom ones, the selected lepton final states are either electron-electron (EE) or muon-
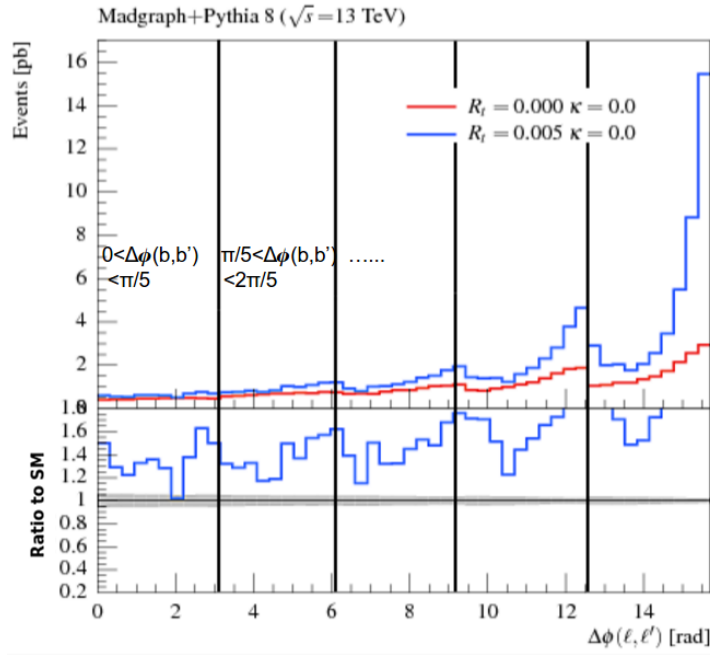
Figure 8: Serialization of signal and the SM. Each rectangular piece has a length of $\pi$. In the first piece, $\Delta\phi_{(l,l')}$ is plotted having the additional constraint that $0 \leq \Delta\phi_{(b,b')} \leq \frac{\pi}{5}$. In the second piece, it is the same, but with the constraint that $\frac{\pi}{5} \leq \Delta\phi_{(b,b')} \leq \frac{2\pi}{5}$, and so on. So, this 1D histogram represents a 2D histogram between the two variables.

muon (MM). Because of the nature of this selection, Drell-Yan events are also present and they represent the biggest background of our selection.

## 4.3 Limit setting methods

This subsection is dedicated to introduce the statistical methods used in the present analysis in order to set limits on the radius of the top quark that are going to be used in the next subsection. Also, the important definition of signal strength is going to be introduced, and we will explain how the computation of the signal strength can help us to exclude or not a new hypothesis.

In particle physics, because of the large amount of data that is used, complex statistical tools have been developed in order to obtain parameters that can tell us whether some data give insights of a new discovery or a new fundamental property of a particle or field. In the present analysis, the information of the signal strength is important in order to infer some information about signal and background.

First, we need to describe what is the null hypothesis, $H_0$, and the alternative hypothesis, $H_a$. Most of the time, the null hypothesis is the hypothesis we want to test and we assume to be true. If, for example, we say that the average grade in a physics exam in a certain class is 75%, that is our null hypothesis. The alternative hypothesis could be the complement, that, in this
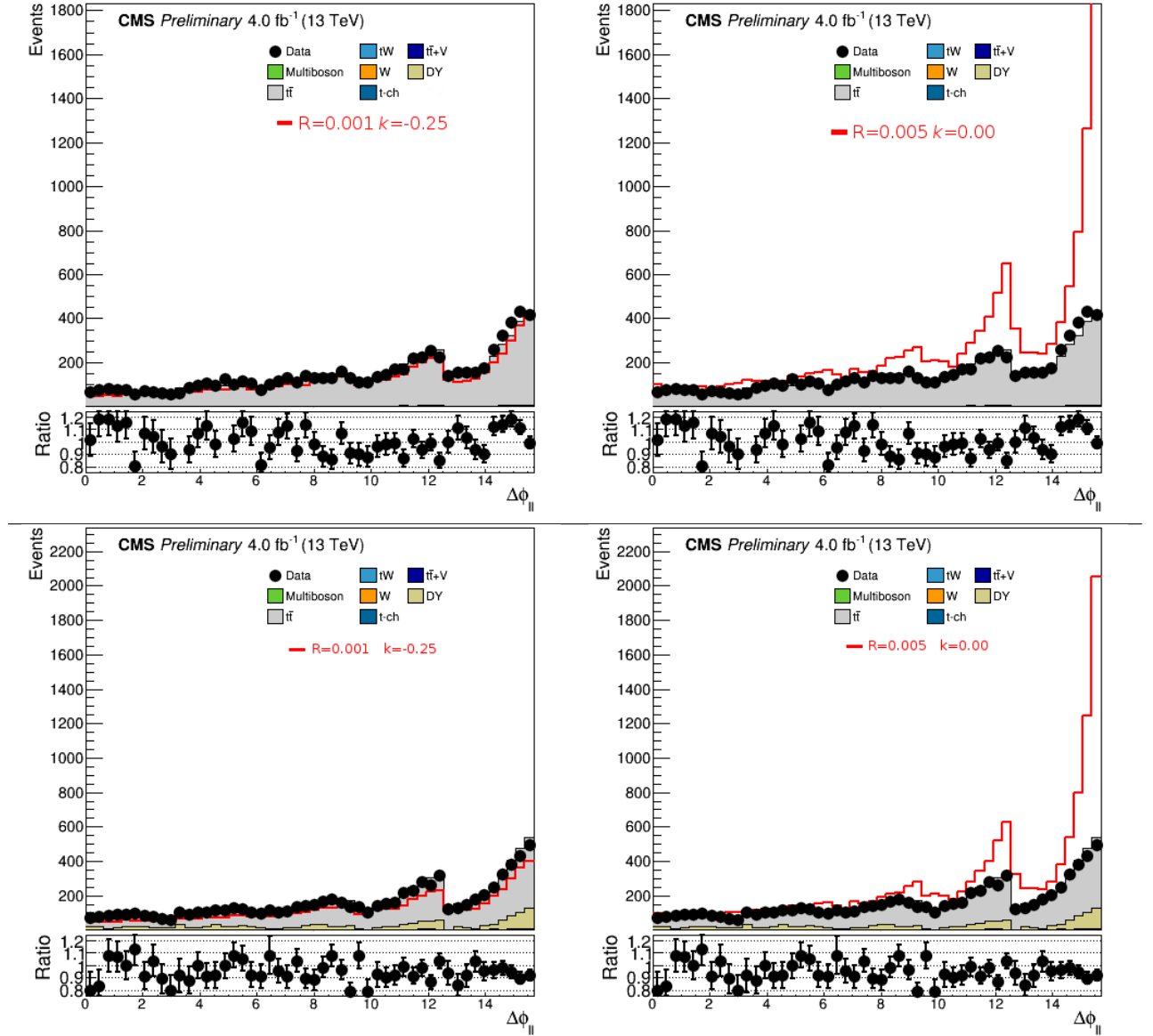
Figure 9: SM $t\bar{t}$ Monte Carlo simulation re-weighted at generator level on the BSM/SM ratio generated with MADGRAPH5 and hadronized with PYTHIA8. The two plots on top were generated selecting electron-muon (EM) as final lepton flavor states, while the bottom ones were constructed selecting electron-electron (EE) or muon-muon (MM) as final lepton flavor states (LL). Backgrounds are plotted as well but they are small relative to the rest of the histogram. Because of the nature of the process, Drell-Yan background is more notorious in the bottom histograms.

case, the average grade is not 75%. In particle physics, depending on what test is performed, the null hypothesis could be the question of whether some measured parameter comes from the background; the alternative hypothesis would be the question of whether that parameter comes from a mixture of a kind of signal (usually involving new physics) and background. The other way around is also frequently used and it depends on the type of analysis.

If we want to measure a kinematic variable $x$ from a sample, we can generate a histogram of $N$ bins[14]. In the bin $i$, there are $n_i$ events. The expected value of $n_i$, $E[n_i]$, would be described by the total amount of events that come from background $b_i$ and the total number of events that come from signal $s_i$ in that specific bin:

$$E[n_i] = rs_i + b_i, \tag{5}$$

where $r$ is known as the signal strength and it has values from $0 \leq r \leq 1$. If $r$ is 0, it means that the size of the bin comes only from background; on the other hand, if $r$ is 1, then it is said that the size of the bin truly comes from signal and background.

$$s_i = s_{tot} \int_{bin\ i} f_s(x; \boldsymbol{\theta}_s)\, dx \tag{6}$$

$$b_i = b_{tot} \int_{bin\ i} f_b(x; \boldsymbol{\theta}_b)\, dx \tag{7}$$

The above expressions are used to compute $s_i$ and $b_i$ [12]. Those expressions are integrated over the required bin; $f_s(x; \boldsymbol{\theta}_s)$ and $f_b(x; \boldsymbol{\theta}_b)$ are the probability density functions (pdf) for signal and background. The pdfs characterize the variable $x$ as belonging to one of the hypothesis by taking into account some constrains that are given by $\boldsymbol{\theta}_s$ and $\boldsymbol{\theta}_b$. Those parameters characterize the shape of the pdfs. Finally, $s_{tot}$ and $b_{tot}$ represent the total average number of signal and background events.

The process to calculate $r$ is rather complicated, but the most important remark is the understanding that not all values of $r$ are equally trustworthy. The actual value of $r$ is only meaningful for some confidence level. One of the most frequently used thresholds of confidence level in particle physics is 95%. It means that there is 5% chance that the hypothetical $r$ would be wrong. The confidence level (CL) is an agreement from all members in the Collaboration.

In this project, we are testing signal from a hypothetical theory against the Standard Model. It represents the probability of wrongly rejecting the hypothetical anomalous radius and magnetic moment of the top quark. If we define $r$ as the expected limit or signal strength, then an expected limit $r \ll 1$ at 95% CL will mean that we are able to exclude the signal upon that condition. It turns that in an asymptotic approximation [12], $r$ is the number of signal events ($n_{obv}$) over the number of expected signal events if the null hypothesis were true, $n_{th}$.

$$r = \frac{n_{obv}}{n_{th}} \tag{8}$$

So, if the number of signal events of a theory introducing an anomalous radius and magnetic

---

[14]Number of divisions of the parameter tested in the histogram.

moment of the top quark are greater than the number of observed events, then we can reject the hypothetical new theory.

It is important to understand that in accordance to the meaning of confidence level, it is only statistical significant if $r < 1$ because in that case we can exclude the null theory; but if $r = 1$, then we can not say that the null theory is right. Such result could be explained by other reasons.

## 4.4 Expected limits of $R_t$ and $\kappa_V$

After we got all reconstructed-level MC for every signal, we evaluated expected statistical power to exclude different hypotheses. It is important to mention that systematic uncertainties are not yet included in the analysis. For some signals, whose distribution resembles pretty much the SM, we got an observed limit on the signal strength $r \gg 1$. We suspect that the picture will change when taking into account the systematic uncertainties of the MC simulation samples.

The following figure, Figure 10, represents the signal strength of simulations with a color scale. Hotter colors represent signal strength closer to 1, while colder colors are values closer to 0. The closer the value to 0, the most likely to exclude the signal. The x-axis represents the anomalous radius $R_t$, while the y-axis represents the anomalous magnetic moment $\kappa_V$. In equation 1, the anomalous radius is squared, so we simulate negative top radius in order to test the symmetry of the radius of the top quark. For that reason, a symmetry in negative and positive values of $R_t$ is expected . As it can be appreciated in Figure 10, not all simulations are represented. The reason for those holes is that some simulations behaved anomalously when we computed the signal strength. Further development is needed in order to find the reason of the behavior of those simulations, but we suspect that the problem raised when we implemented the new model into the Matrix Element generator.

Specific sections of the plot above are represented in Figure 11. In those plots, black squares represent the computed signal strength. The observed and expected signal strength are plotted as a dashed and continuous black line, respectively. The observed signal strength is computed when comparing simulations with real data, and the expected limit is set when comparing simulations with the Standard Model Monte Carlo. They may differ in some cases, specially when there are statistical fluctuations or in the case there is some new physics in data. Yellow and green areas correspond to the $2\sigma$ and $1\sigma$ values of the signal strength for 68% and 95% CL. In the top plot of Figure 11, the signal strength of simulations fixing $\kappa_V = 0.05$ are plotted. It can be noticed easily that for those values of $R_t$, $r$ is well below 1, so they are excluded. In the second plot of Figure 11, it is represented the signal strength fixing $R_t = 0.005$. In the same way, in all those simulations, $r$ is below to 1, so excluded as well.
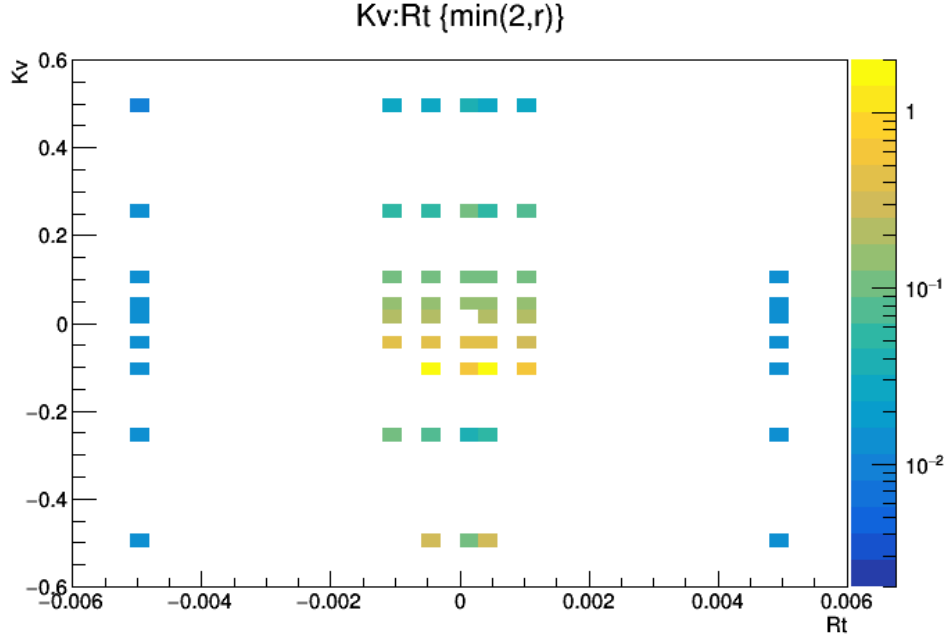
Figure 10: Expected 95% CL limit on the signal strength $r$, indicated by the color scale for different signals. For $r < 1$, we expect to exclude that model.

Finally, values of $R_t$ and $\kappa_V$ close to zero should have the highest signal strength[15]; nevertheless, in Figure 10 the highest values are in the negative y-axis. The reason for this behavior is still unclear, but it could be related to the nature of $\kappa_V$. Negative and positive values could produce non-symmetric physical scenarios. However, we do not know why signals near the center have a higher signal strength than the ones with negative $\kappa_V$.

For further analyses, systematic uncertainties such as pile up, jet energy correction, b-tagging selection, jet energy scale, and lepton energy scale must be taken into account. Also, uncertainties from the theory such as ME/PS[16] matching threshold, QCD factorization and normalization scales, PDFs, etc., need to be considered.

It is important to point out that even though angular variables, as the ones used in this analysis, are not expected to change with energy corrections, the cross-section can change. For a future work, it would be necessary to explore other variables as well.

---

[15]Following equation 1, if $R_t = \kappa_V = 0$, the effective Lagrangians introduced go to zero. So, we recover the SM, which agrees with data.

[16]ME stands for matrix element generator, PS means parton shower. The systematic ME/PS is the paring between matrix element generator and parton shower. When events are simulated, there is an energy threshold for extra radiation process created in PYTHIA, rather than the ones created using the matrix element generator, in this case MADGRAPH. So, extra samples are created with variation in the energy threshold with a factor of 2 [13].
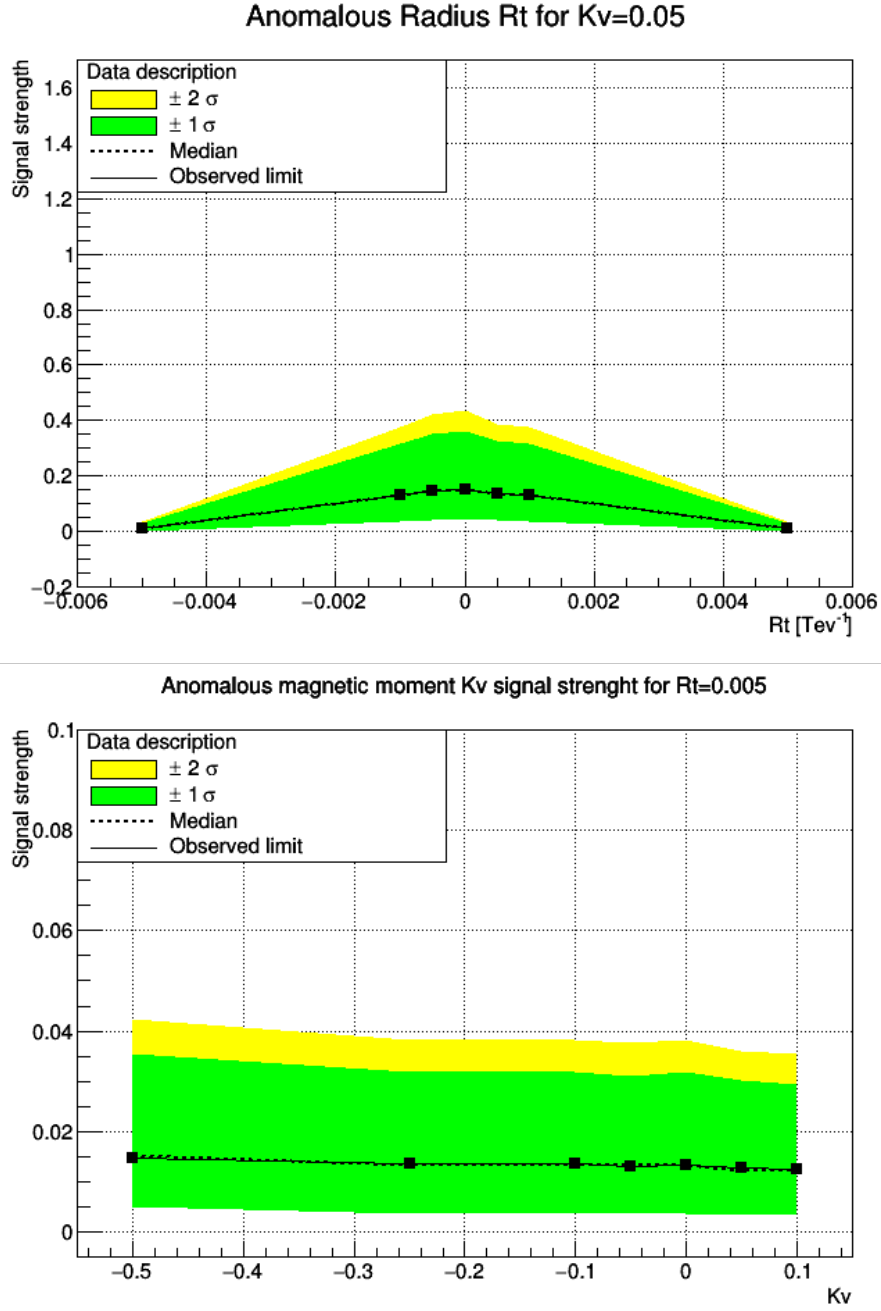
Figure 11: In the first plot, the signal strength $r$ for constant $\kappa_V = 0.05$ is plotted while in the second plot signal strength is plotted for constant $R_t = 0.005$. These are a blind plots, so data has not been used to compute the signal strength, it is only computed by comparison between simulations.

# 5 Conclusions

Using generator-level simulations of $t\bar{t}$ events where anomalous radius and magnetic moment of the top quark were implemented, we analyze different variables sensitive to the presence of those parameters. From the nearly 36 different variables analyzed, we select angular variables $\Delta\phi_{(l,l')}$ and $\Delta\phi_{(b,b')}$ to continue with the analysis. We used these variables because they

are simple and exhibit an interesting behavior against the Standard Model. Using those variables we reconstructed a Standard Model Monte Carlo to simulate signal at reconstruction-level. Implementing statistical methods we were able to set limits on the radius of the top, $R_t < 0.001 \text{ TeV}^{-1} \approx 10^{-6} \text{ fm}$, without taking account systematic uncertainties. Even if the limit changes after implementing systematics, as a preliminary analysis, a better limit than the theoretical one (using cross sections; $R_t \leq 0.25 \text{ TeV}^{-1}$) [2] is expected. As far a we know, this is the first experimental attempt to try to measure directly the radius of the top quark, which potentially could lead to an indication of physics beyond the Standard Model.

This work is in process. We are trying to implement systematic uncertainties in order to see how limits change. In addition to the systematics, the Drell-Yan background from Monte Carlo should be corrected to avoid statistical fluctuations. For this purpose, we are working on getting scale factors taking out the $Z$ pole mass, and implementing them. After checking how these modifications affect the shape of histograms, we will compute again the signal strength for all signal in order to set new limits on the radius of the top. Now, with a new limit, we will explore other kind of variables, in order to get a stronger limit. Our objective is to publish, with the CMS collaboration, this work in 2017.
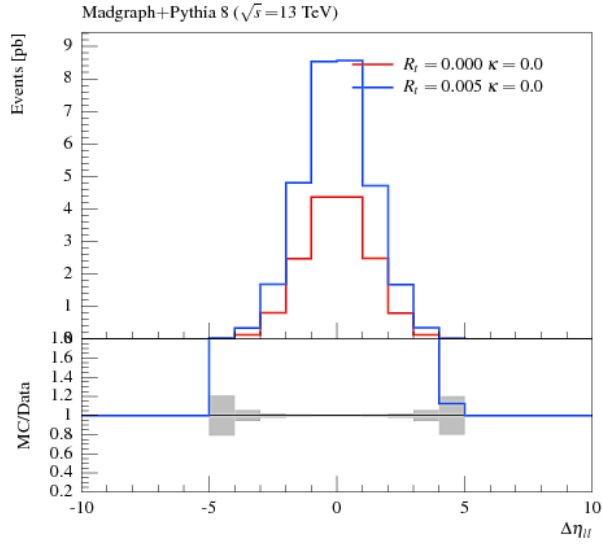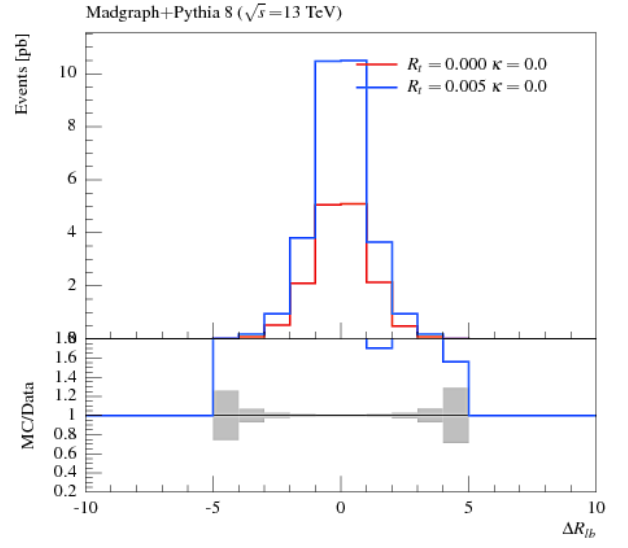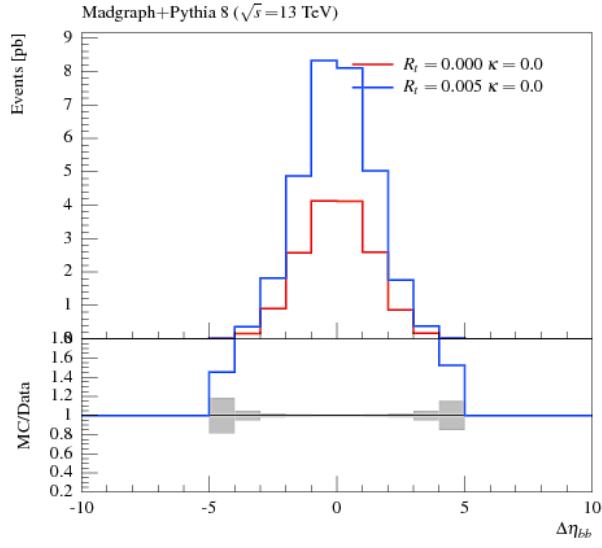
# References

[1] MissMJ,. (2006). *Standard Model of Elementary Particles*. Retrieved from `https://upload.wikimedia.org/wikipedia/commons/0/00/Standard_Model_of_Elementary_Particles.svg`

[2] Englert, C (University of Glasgow)., Gonçalves, D (Durham University)., & Spannowsky, M (Durham University). (2014). *Non-standard top substructure* (1st ed.). Retrieved from `DOI:10.1103/PhysRevD.89.074038, arXiv:1401.1502.`

[3] Camblong, H., Epele, L., Fanchiotti, H., & Canal, C. (2000). *Dimensional Transmutation and Dimensional Regularization in Quantum Mechanics. I. General Theory* (1st ed.). Retrieved from `https://arxiv.org/pdf/hep-th/0003255v2.pdf`

[4] S. Chatrchyan *et al.* [CMS Collaboration], JINST **3**, S08004 (2008). `doi:10.1088/1748-0221/3/08/S08004`

[5] International Particle Physics Outreach Group,. (2010). CMS Slice. Retrieved from `http://ippog.web.cern.ch/resources/2011/cms-slice-july-2010-version`

[6] Wong, C. (1994). *Introduction to high-energy heavy-ion collisions* (1st ed.). Singapore: World Scientific.

[7] S. Chatrchyan *et al.* [CMS Collaboration], JINST **5**, T03003 (2010) `doi:10.1088/1748-0221/5/03/T03003` [`arXiv:0911.4893` [physics.ins-det]].

[8] Mets501,. (2012). *Different values of pseudorapidity shown against a polar grid*. Retrieved from `https://en.wikipedia.org/wiki/File:Pseudorapidity_plot.svg`

[9] J. Alwall, M. Herquet, F. Maltoni, O. Mattelaer and T. Stelzer, JHEP **1106**, 128 (2011) doi:10.1007/JHEP06(2011)128 [arXiv:1106.0522 [hep-ph]].

[10] T. Sjostrand, S. Mrenna and P. Z. Skands, Comput. Phys. Commun. **178**, 852 (2008) doi:10.1016/j.cpc.2008.01.036 [arXiv:0710.3820 [hep-ph]].

[11] Alison, J. (2012). *The Road to Discovery: Detector Alignment, Electron Identification, Particle Misidentification, WW Physics, and the Discovery of the Higgs Boson* (Doctor of Philosophy). University of Chicago.

[12] Cowan, G., Cranmer, K., Gross, E., & Vitells, O. (2013). *Asymptotic formulae for likelihood-based tests of new physics* (2nd ed.). Retrieved from `DOI:10.1140/epjc/s10052-011-1554-0, arXiv:1007.1727.`
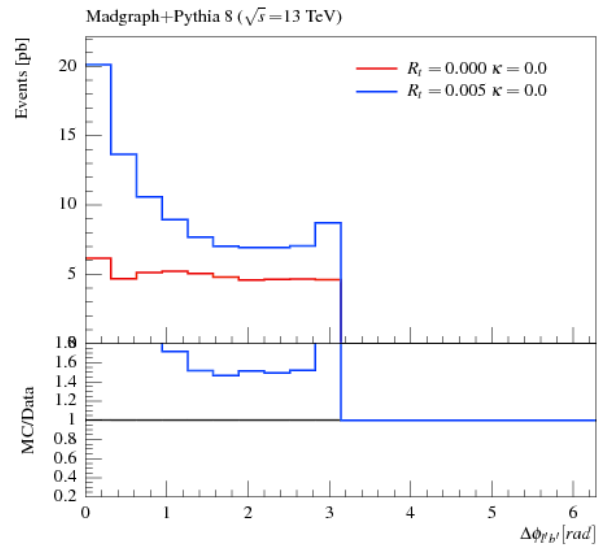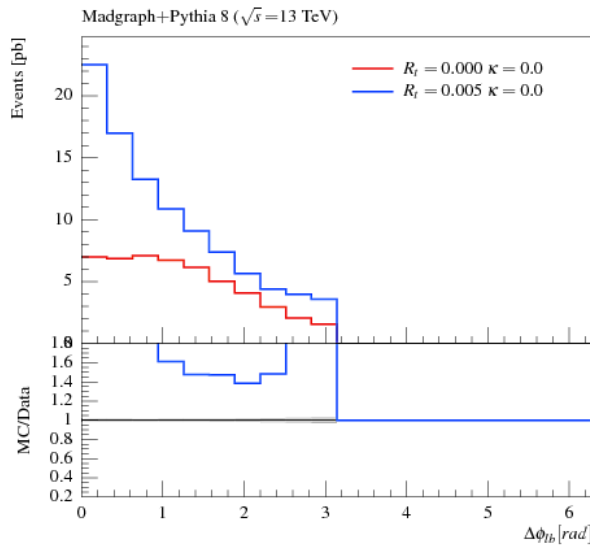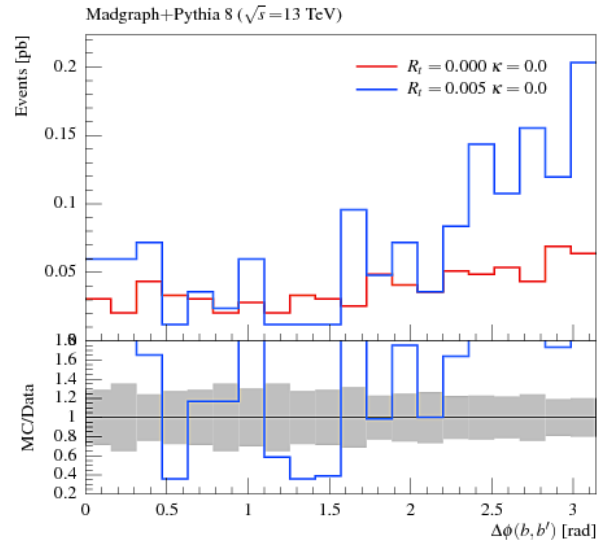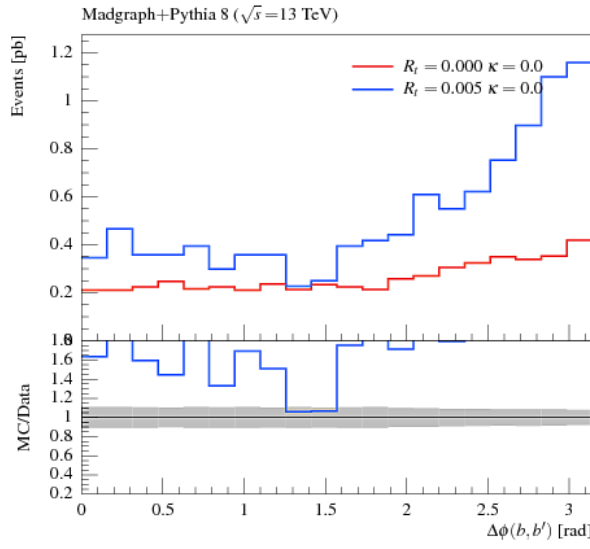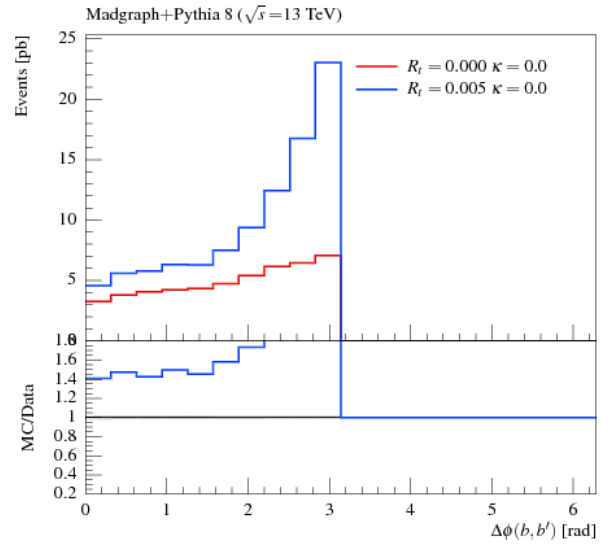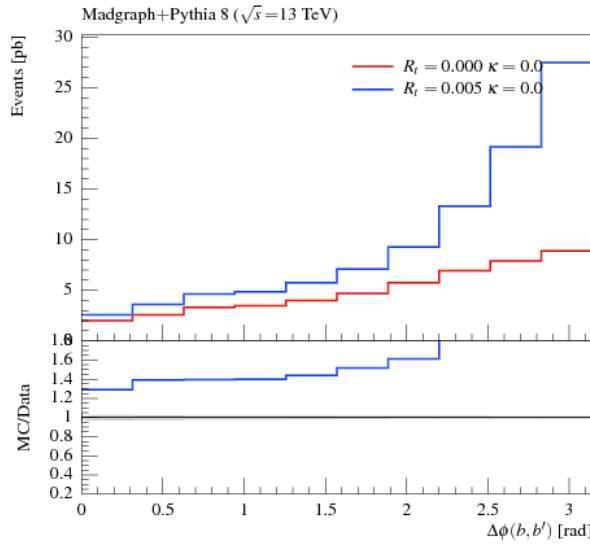
[13] Guerrero, D. (2016). *Determinación de la masa del quark top mediante el análsis del espectro de energía de b-jets con el detector solenoide compacto de muones (CMS)* (B. Sc. Physics). Escuela Politécnica Nacional.
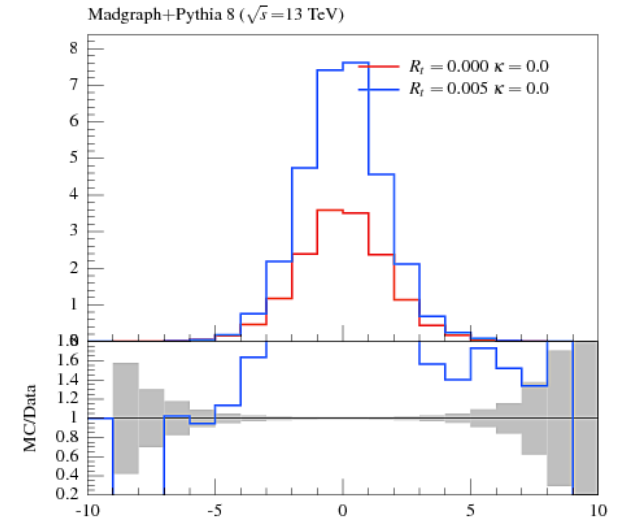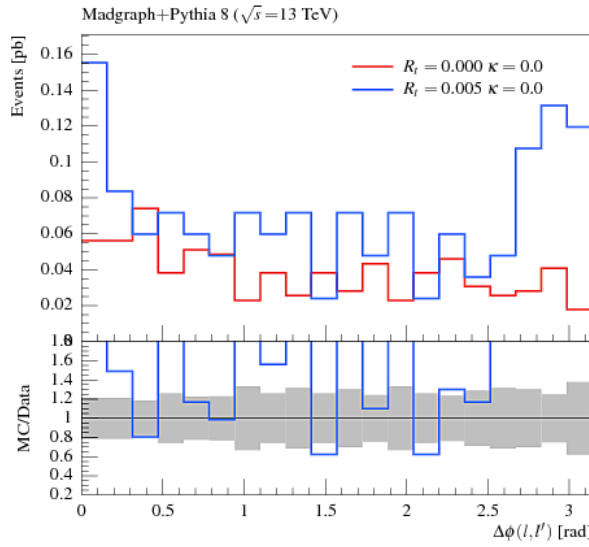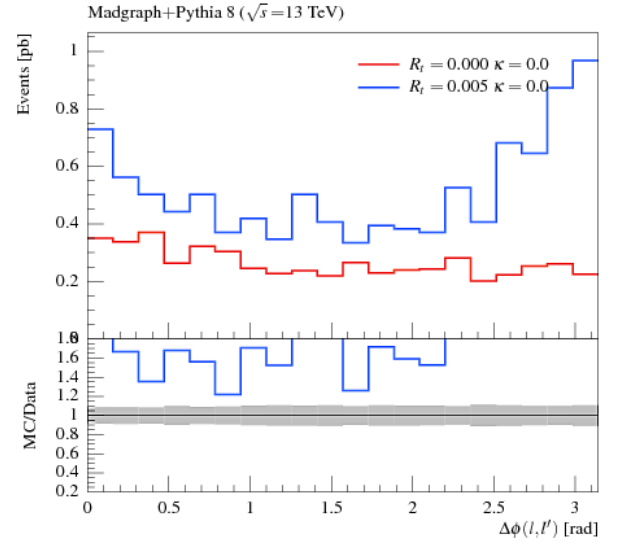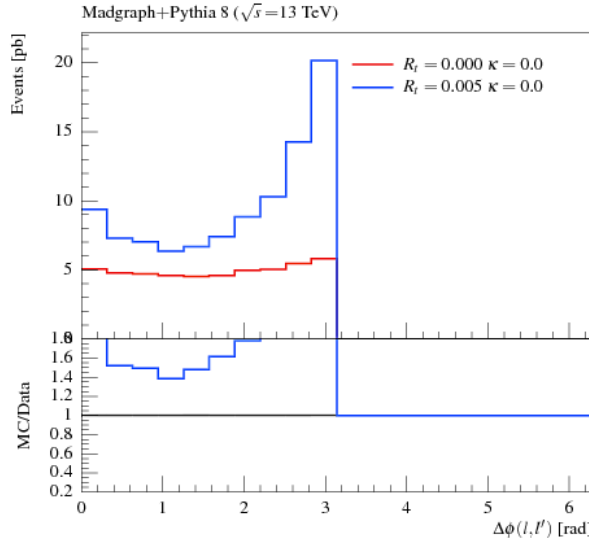
# 6 Annexe

## 6.1 Annex A: Other variables identified at generator-level

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi(b, b')$ [rad]

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi(b, b')$ [rad]

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi(b, b')$ [rad]

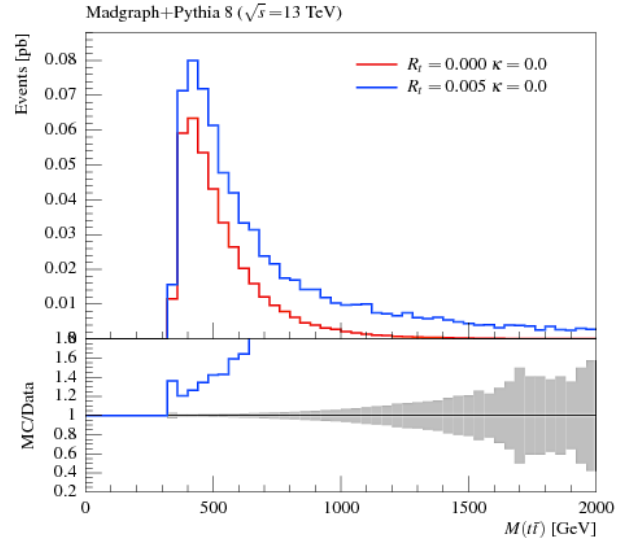Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi(b, b')$ [rad]

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi_{lb}$ [rad]

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$R_t = 0.000$ $\kappa = 0.0$
$R_t = 0.005$ $\kappa = 0.0$

Events [pb]

MC/Data

$\Delta\phi_{l'b'}$ [rad]

Madgraph+Pythia 8 ($\sqrt{s}=13$ TeV)

$R_t = 0.000 \ \kappa = 0.0$
$R_t = 0.005 \ \kappa = 0.0$

Madgraph+Pythia 8 ($\sqrt{s}$ =13 TeV)

$H_\mathrm{T} = p_\mathrm{T}(b_1) + p_\mathrm{T}(\ell_1) + p_\mathrm{T}(b_2) + p_\mathrm{T}(\ell_2) + E_\mathrm{T}^{\mathrm{miss}}$ [GeV]

$m1_{bl} \frac{GeV}{c^2}$

$m2_{bl} \frac{GeV}{c^2}$

$M(t\bar{l})$ [GeV]

$\Delta\phi_b^*$

$\Delta\phi_{l'}^*$
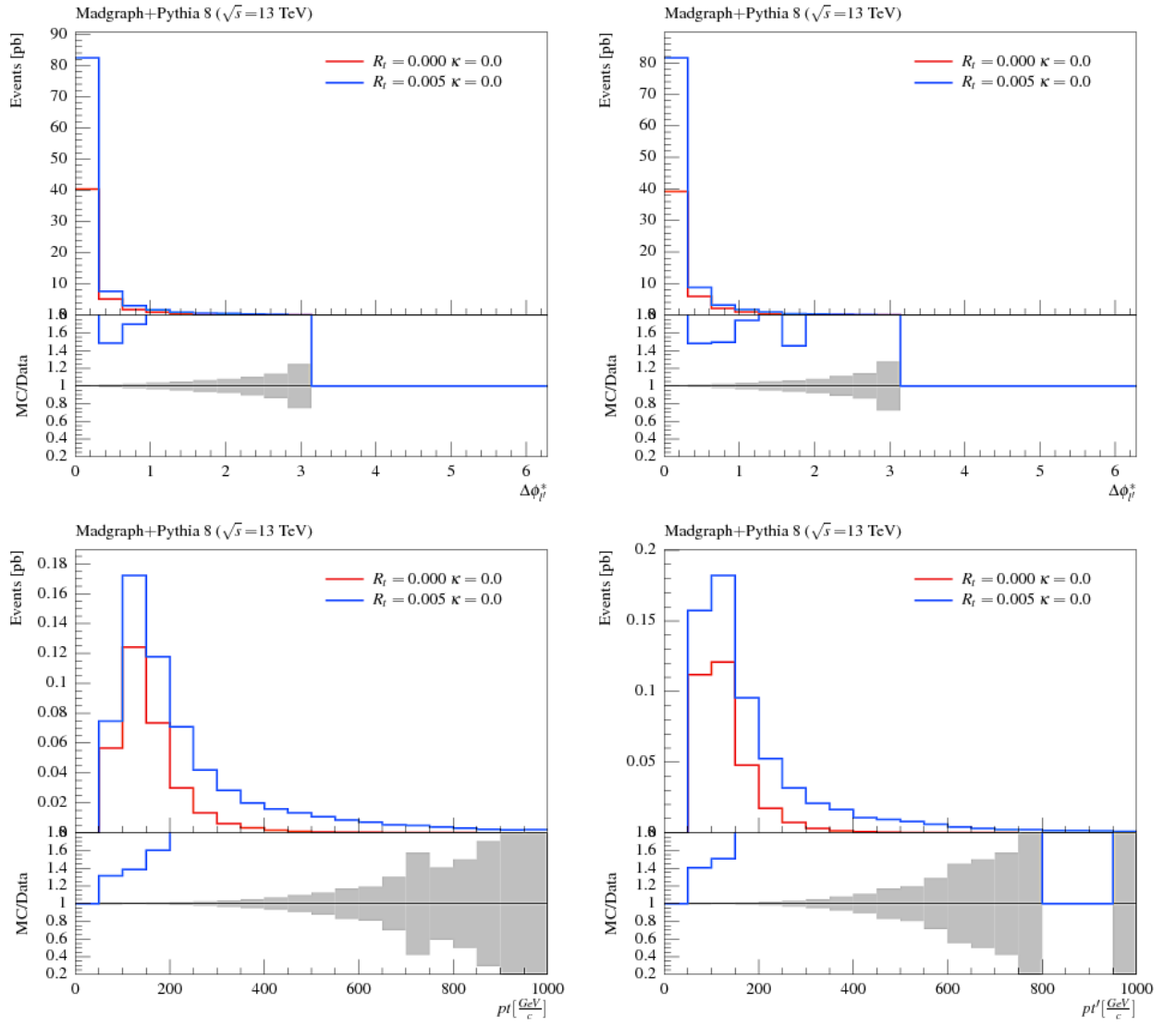
$R_t = 0.000\ \kappa = 0.0$

$R_t = 0.005\ \kappa = 0.0$

Events [pb]

MC/Data

## 6.2 Annex B: Reconstructed SM MC for varius signal