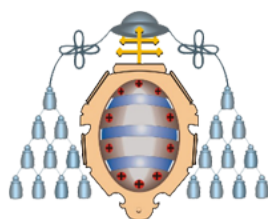


Práctica: Medida de la sección eficaz de producción de pares *top-antitop* en CMS

Física de Altas Energías y Aceleradores - Grado en Física
Universidad de Oviedo



1. Introducción

En esta práctica se pretende realizar una medida de la sección eficaz del proceso de producción de pares *top-antitop* en colisiones protón-protón con el detector CMS, a una energía en centro de masas de $\sqrt{s} = 7$ TeV, en el estado final con un muon. A primer orden, los diagramas que contribuyen a este proceso son la fusión de gluones, con cerca de un 80 % de la contribución total, y la aniquilación *quark-antiquark*, con el 20 %. Estos diagramas se observan en la figura 1.

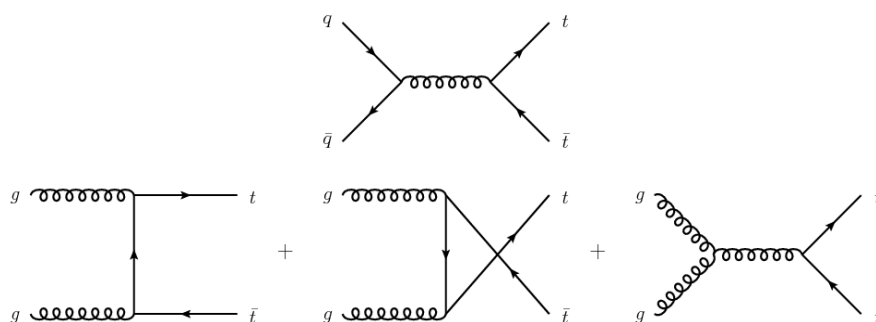


Figura 1: Diagramas de Feynman a orden más bajo para la producción de sucesos $t\bar{t}$ en colisiones de protones.

El *quark top* es la partícula más masiva del modelo estándar (o SM, del inglés *Standard Model*). Además, es el único *quark* que se desintegra antes de poder hadronizarse, lo cual lo hace esencial para probar con precisión las predicciones de la cromodinámica cuántica (QCD, *idem Quantum Chromodynamics*). La medida de la sección eficaz de producción de $t\bar{t}$ es una prueba perfecta para comprobar las predicciones del SM a NNLO (*next-to-next-to leading order*, en teoría de perturbaciones), además de que es sensible a la masa del *quark top* (m_{top}), la constante de acoplo fuerte (α_S) y las funciones de distribución de los partones en el protón (PDF).

Además, debido a la alta sección eficaz de producción de este proceso en colisiones pp a estas energías en el centro de masas, se dice habitualmente que el LHC es una «*top factory*». La mayoría de las búsquedas de nuevas partículas o medidas de precisión se ven afectadas por la presencia de partículas provenientes de desintegraciones de *quarks top*. El estudio de este proceso es fundamental para entender el fondo en estas búsquedas.

En CMS se ha medido esta sección eficaz a diferentes energías en centro de masa en colisiones de protones (7 TeV, 8 TeV, 13 TeV y 5.02 TeV). La medida más precisa a 7 TeV se ha realizado en el canal dileptónico (dos leptones en el estado final) y puede consultarse en la referencia [1]. También puedes consultar la medida a 7 TeV utilizando el canal $\ell+jets$ [2].

2. Datos y material para la práctica

Para realizar la medida se utilizará una luminosidad de 50 pb^{-1} de datos tomados por CMS en 2011. Estos datos son públicos, y pueden utilizarse libremente con propósito académico.

Se usarán muestras simuladas con algoritmos de Monte Carlo (MC) para predecir el fondo debido a otros procesos del SM y para estudiar la aceptación del detector y la eficiencia de reconstrucción e identificación de los diferentes objetos (partículas, *jets*, así como distintos observables físicos) reconstruidos. Estas simulaciones han sido diseñadas con el propósito de realizar esta práctica, por lo que contienen información limitada pero suficiente para poder llevar a cabo todas las partes de la misma.

La práctica está basada en este ejercicio académico propuesto por el *International Particle Physics Outreach Group*. En el siguiente enlace se puede consultar la información acerca de cómo han sido generadas las muestras, número de sucesos, normalización, etcétera.

<http://ippog.web.cern.ch/resources/2012/cms-hep-tutorial>

3. Medida de la sección eficaz

Para medir la sección eficaz utilizaremos la expresión

$$\sigma_{t\bar{t}} = \frac{N - N_{bkg}}{BR \cdot \int \mathcal{L} dt \cdot \mathcal{A} \cdot \varepsilon}, \quad (1)$$

donde:

- N : número de sucesos observados.

- N_{bkg} : número esperado de sucesos de fondo, estimados a partir del MC.
- BR : fracción de ramificación (*branching ratio*) de $t\bar{t}$ al estado final considerado ($\mu+jets$).
- $\int \mathcal{L} dt$: luminosidad integrada total. Son 50 pb^{-1} .
- \mathcal{A} : aceptancia. Incluye la aceptancia geométrica (restricción en $|\eta|$, etcétera) y de selección (cortes cinemáticos, p_T del muon...).
- ε : producto de todas las eficiencias de reconstrucción e identificación para todas las partículas detectadas. Incluye la eficiencia de *trigger*, identificación de *quarks* b (si procede), eficiencia de reconstrucción de muones, etcétera.

Típicamente, en la medida de la sección eficaz se citan tres incertidumbres diferentes: estadística (depende del número de sucesos observados), la debida a la luminosidad (viene dada por el acelerador y cómo CMS la mida) y sistemática (el resto de fuentes). La incertidumbre total será la suma cuadrática de las tres subcomponentes.

3.1. Selección de sucesos

La primera parte del diseño del análisis consiste en entender la topología de la señal ($t\bar{t}$ en nuestro caso) y el fondo (resto de procesos físicos) y diseñar una selección de sucesos que optimice el número de aquellos que sean de señal respecto a los que provengan de un proceso de fondo (siempre intentando mantener una aceptancia lo más alta posible). Para ello, primero puedes dibujar distribuciones cinemáticas para todos los procesos simulados. Algunas de ellas: momento transverso del muón y los *jets*, número de *jets*, número de b -tags, energía faltante en el plano transverso (MET), ángulo entre el muon y el *jet* más energético (o cada uno de los b -tags), etcétera. Ten en cuenta el estado final del proceso, reflejado en la figura 2.

Tras examinar las distribuciones, decide qué variable es más discriminante (separa mejor entre señal y fondo) y aplica un criterio de selección en esa variable (un «corte»). Recomendación: observa los productos de desintegración y haz una lista de los observables que esperarías. Realiza **cortes secuenciales**, es decir, tras decidir y aplicar un criterio de selección vuelve a pintar todos los histogramas de las variables significativas e, iterativamente, elige otra variable que discrimine entre señal y fondo para aplicar otro corte.

Ten en cuenta las siguientes incertidumbres de normalización para los fondos:

- $W+jets$: 50 %.
- QCD: 100 %.
- WW , WZ , ZZ : 50 %.
- Drell-Yan: 15 %.
- *Single-top*: 30 %.

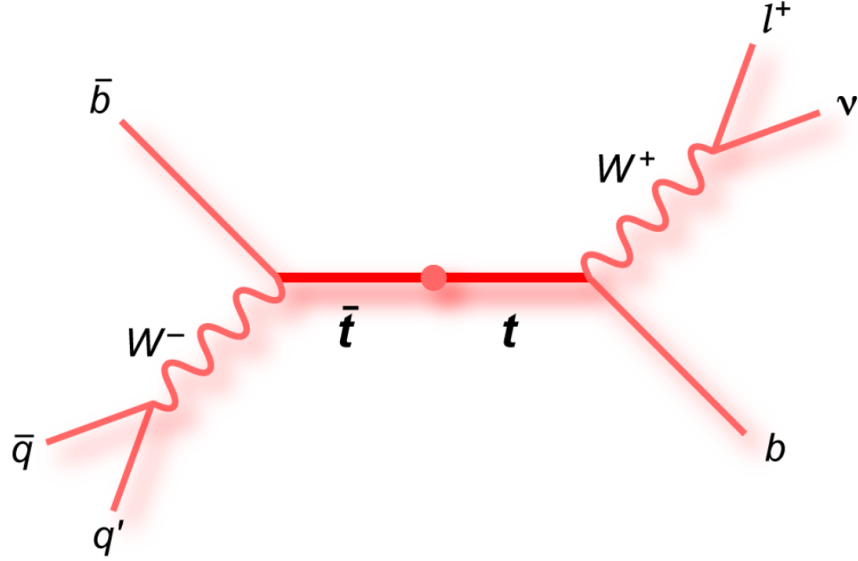


Figura 2: Desintegración de un par $t\bar{t}$ en el estado final con un único muon.

Ten en cuenta también, en caso de que sean grandes, las incertidumbres estadísticas debida al número de sucesos generados: recuerda que en las simulaciones de Monte Carlo se genera un número de sucesos que puede ser mayor o menor que el número de sucesos esperados en las colisiones, por lo que se escala a la luminosidad dada y utilizando la mejor estimación de la sección eficaz teórica utilizando un peso ω aplicado a cada suceso:

$$\omega = \frac{\int \mathcal{L} dt \cdot \sigma}{\text{Número de sucesos generados}}.$$

(Esto no es realmente así, ya que por cómo se producen algunos de los procesos, no todos los sucesos tienen un peso idéntico, pero para lo que necesitamos saber, da igual). Pero la incertidumbre estadística en el número de sucesos esperados (tras los cortes) es proporcional a la raíz cuadrada del número de sucesos *sin peso*, tras la selección final.

3.1.1. Identificación del *b-jets*

Los *jets* procedentes de la hadronización de *quarks* b , *b-jets*, presentan un registro experimental ligeramente diferente de los *jets* que proceden de gluones o *quarks* ligeros, debido fundamentalmente a que se desplazan ligeramente del vértice de interacción principal, dando lugar a vértices secundarios. Gracias a la magnífica resolución del detector de píxeles de CMS, se pueden extrapolar las trazas detectadas de las partículas cargadas de los *jets* y determinar si provienen del vértice primario o no.

Utilizando un algoritmo complejo (llamado CSV), que incluye no solo la distancia del vértice reconstruido respecto el vértice primario, sino también información de cada una de las partículas del *jet* y otras propiedades, se puede asignar una probabilidad a cada *jet*, que refleja cuán seguros

podemos estar de que un *jet* provenga de un *quark* b . Este proceso se conoce como etiquetado de b -jets (*b-tagging*) y los *jets* etiquetados se denominan *b-tags*.

En esta práctica, la variable toma el valor -1 para los *jets* cuya probabilidad de proceder de un *quark* b es muy baja y toma valores mayores que 1 para posibles *b-tags*, siendo este valor más alto cuanto más probable sea que el *b-tag* sea realmente un *b-jet*. La eficiencia de *b-tagging* dependerá del valor que escojamos en la variable para considerar a un *jet* etiquetado, pero es del orden del 80 % para los *jets* medidos. Además, debido a que este proceso es difícil de simular, la eficiencia de *b-tagging* en la simulación de MC no es similar a la eficiencia real (en datos), sino que es mayor que esta. Para corregir este efecto, se aplica un factor de 0.9 a la normalización del suceso simulado por cada *jet* identificado que se requiera. Esto se puede hacer, de manera sencilla, multiplicando el valor del peso del suceso (*EventWeight*) por este factor (las veces requeridas). Si el número de sucesos requerido es 2 (o más), pesa cada sucesos simulado por un factor 0.86 (en lugar de un factor $0.9 \times 0.9 = 0.81$).

La incertidumbre (conservadora) de la eficiencia de *b-tagging* es de 10 %. Puedes tratar de medir la eficiencia de *b-tagging* en simulación y tomar como eficiencia de *b-tagging* (en datos): $\varepsilon_b^{\text{datos}} = 0.9 \cdot \varepsilon_b^{\text{MC}}$. Esto es aconsejable, sobre todo, si escojes un valor alto en la variable de *b-tagging* para identificar los *b-jets* o si seleccionas más de un *b-tag*. Si haces esto, puedes estimar la incertidumbre de *b-tagging* como la diferencia entre la eficiencia de *b-tagging* en datos y en MC.

3.2. Eficiencia de *trigger*

El *trigger* es el mecanismo que tiene CMS para decidir si guardar o no un suceso. El ancho de banda es limitado (aproximadamente de 1kHz, es decir, 1000 sucesos por segundo), sin embargo se producen millones de cruces entre protones cada segundo. La mayoría de las colisiones no resultan en la producción de partículas de interés y, por tanto, no son almacenadas. Hay varias estrategias de *trigger* diseñadas con diferentes propósitos en CMS: basados en energía faltante en el plano transversal, en muy alta energía de *jets*, en fotones, en muones, etcétera. En particular, en este análisis utilizaremos un *trigger* basado en la presencia de un muon aislado con un p_T mayor de 24 GeV.

Todos los sucesos observados han activado el *trigger* (si no, no los podríamos haber guardado). La reconstrucción que se hace a nivel de *trigger* ha de ser rápida y, por tanto, poco precisa. Por ello, puede haber muones con momento menor de 24 GeV que disparan el *trigger*, y muchos otros con momento mayor que no lo hacen. Este efecto se estudia con la eficiencia de *trigger* en función del p_T del muon. A la hora de comparar con la teoría, hay que simular este efecto. Esto se simula en las muestras de MC utilizando una variable que toma un valor booleano (0 o 1) para cada suceso dependiendo de si el *trigger* se ha activado para este suceso o no. La eficiencia de *trigger*, aunque también depende de variables angulares, depende fuertemente del p_T de los muones, siendo muy baja para muones con p_T menor que el p_T de diseño (24 GeV) y tendiendo a un valor fijo a altos valores de p_T .

Para calcular la eficiencia, haremos uso de la muestra de $t\bar{t}$ MC. Dibuja, en un mismo *pad*, la distribución del p_T del muon en un rango amplio de p_T para sucesos de la muestra de $t\bar{t}$ con y sin el requisito de *trigger*. Haz el cociente entre ambas distribuciones para obtener la eficiencia en bins (celdas) de p_T . Integra la distribución a partir del p_T que utilices en el análisis para

calcular un valor integrado (promedio) de la eficiencia de *trigger* (será el valor que utilices en el análisis).

La incertidumbre de la eficiencia de *trigger* será $(1 - \varepsilon_{tr})/2$.

3.3. Aceptancia

Para calcular la aceptancia, utiliza las variables a nivel de generación para desligar el efecto de aceptancia de selección del efecto del detector (la reconstrucción). Utiliza los cortes cinemáticos de tu análisis en estas variables. Ten en cuenta que a este nivel puedes acceder a la información de los *quarks* y neutrinos. Calcula la aceptancia como:

$$\mathcal{A} = \frac{\text{Numero de sucesos en la region sensible}}{\text{Numero total de sucesos} \cdot BR}. \quad (2)$$

El número total de sucesos simulados es de 36941. La incertidumbre en la aceptancia se puede calcular a partir de la incertidumbre estadística debida al limitado número de sucesos simulados (tanto en el numerador como en el denominador).

3.4. Otros valores

Tomaremos la eficiencia total:

$$\varepsilon = \varepsilon_{tr} \cdot \varepsilon_b \cdot \varepsilon_\mu \quad (3)$$

con $\varepsilon_\mu = 0.99 \pm 0.01$.

La luminosidad integrada total tiene una incertidumbre del 10 %.

El BR puede tomarse a partir de los datos experimentales de desintegración de los bosones W tabulados por el PDG [3] (hay que tener en cuenta el caso de $W \rightarrow \tau \rightarrow \mu$) y la incertidumbre es despreciable.

4. Medida de la masa del *quark top*

Se puede determinar la masa del *quark top* a partir de sus productos de desintegración, reconstruyendo primero los bosones W para, posteriormente, reconstruir los *quarks top*. Para ello, primero empezaremos por realizar una selección que nos permita identificar todos los objetos (al menos 4 *jets*, eligiendo dos de ellos como b-*jets*). Tendremos que reconstruir primero los bosones W: uno hadrónico, a partir de dos *jets*, y otro leptónico, a partir del muon y la MET. Posteriormente, utilizaremos dos *b-tags* para reconstruir cada uno de los *quarks top* del suceso y crearemos un histograma con su masa.

Empieza primero reconstruyendo la masa del *quark top* a nivel de generación, utilizando toda la información de las partículas generadas. Después, reconstruye la masa para todos los sucesos (datos y simulación). Puedes seguir los siguientes pasos:

- Ordena los *jets* del suceso (al menos 4) según el valor del discriminante de *b-tagging*. Toma como b-tags los *jets* con valor más alto de este discriminante.

- Utiliza el resto de *jets* para reconstruir la masa del bosón W que sufrió desintegración hadrónica y utiliza la pareja que reconstruya una masa más próxima a 81 GeV. Si quieres simplificar, puedes seleccionar los dos más energéticos.
- Utiliza la energía faltante en el plano transversal y el muón para reconstruir la masa del bosón W que sufrió desintegración leptónica. Puedes forzar a que la masa reconstruida sea próxima a la del bosón W para restringir la componente z de la energía faltante.
- Combina cada uno de los bosones W reconstruidos con los dos b-tags. Elige la combinación que den valores más próximos al teórico.
- Almacena estos valores en un histograma.
- Realiza histogramas para, al menos: la masa del W hadrónico, la masa del W leptónico, la masa del *top* hadrónico, la masa del *top* leptónico y la masa del mejor *top* (la que más se acerque, en cada suceso, a la masa teórica del *quark top*).

Presenta el estudio a nivel de generación (usando la muestra de $t\bar{t}$) y, posteriormente, a nivel de reconstrucción, con datos y simulación. Ajusta las masas utilizando TH1::Fit. El resultado de la medida de la masa será el valor dado por el ajuste de los datos.

5. Objetivos y elaboración del informe

En el informe se debe realizar una descripción completa de la práctica, incluyendo todas las fórmulas y números utilizados, el método, la bibliografía utilizada, etcétera. Debe incluir una descripción de los objetos utilizados (rangos de p_T y pseudorapidez de los diferentes objetos, *trigger* usado, etcétera), selección de sucesos (justificada con histogramas, etcétera), distribuciones significativas, estimación de fondos (¿por qué cada uno de los procesos estimados puede contribuir al fondo?) y medida de cada uno de los parámetros necesarios para calcular la sección eficaz (eficiencias, aceptación). Se deben propagar todas las incertidumbres a la medida.

La media de la sección eficaz habrá que compararla con el valor teórico y con las medidas realizadas por ATLAS y CMS.

La medida de la masa se obtiene tras ajustar la distribución de la masa reconstruida en datos. Compara siempre con la simulación y presenta los histogramas de las distintas variables que intervienen en el proceso (masas invariantes de diferentes objetos, momentos transversos o ángulos entre diferentes objetos, etcétera). Compara también esta medida con lo obtenido experimentalmente por las colaboraciones del CERN.

Referencias

- [1] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in the $e\mu$ channel in proton-proton collisions at $\sqrt{s} = 7$ and 8 TeV*, **10.1007/JHEP08(2016)029**.

- [2] CMS Collaboration, *Measurement of the $t\bar{t}$ production cross section in pp collisions at $\sqrt{s} = 7$ TeV with $\ell + \text{jets}$ final state*, **10.1016/j.physletb.2013.02.021**.
- [3] Web del *Particle Data Group*: <http://pdg.lbl.gov/>