

7

ESTIMACIÓN DEL FONDO

Aun luego de aplicar los criterios de identificación *tight* descriptos en el Cap. 5, existe cierta contaminación de fotones del fondo en los eventos seleccionados y su estimación es crucial para la medida de la sección eficaz final. El fondo remanente, debido principalmente al decaimiento de mesones livianos en un par de fotones, es estimado estadísticamente a partir de la distribución de la energía de aislamiento calorimétrico (E_T^{iso}). En este capítulo, se discute en primer lugar la extracción del perfil de E_T^{iso} para los fotones del fondo directamente de los datos. En la Sec. 7.2 se describe el método de conteo bidimensional utilizado para la estimación de la pureza de la muestra seleccionada, i.e. la fracción de eventos de señal, seguido del tratamiento de las incertezas sistemáticas asociadas al mismo. Luego de discutir otras posibles fuentes de contaminación en la Sec. 7.4, se presentan los resultados finales considerados en el capítulo siguiente para la medida de la sección eficaz de producción de fotones directos aislados.

7.1 Caracterización del fondo

El fondo principal de la producción de fotones aislados proviene de los decaimientos de mesones livianos, particularmente $\pi^0 \rightarrow \gamma\gamma$. La contaminación en la muestra de candidatos a señal, aun luego de aplicar los criterios de identificación, es estimada haciendo uso del aislamiento calorimétrico (E_T^{iso}). Es importante entonces entender el comportamiento del fondo en términos de esta variable, extrayendo su distribución directamente de los datos bajo estudio. Las variables de identificación *tight* construidas en la primera capa del ECAL descritas en la Sec. 5.2.1, han sido diseñadas para discriminar este tipo de decaimientos con dos fotones en el estado final, buscando subestructuras en el perfil lateral de la deposición de energía en la dirección de η (Fig. 5.3). Aún más, como el núcleo central ($\eta \times \phi = 5 \times 7$) de las celdas en la segunda capa del ECAL es removido del cono de aislamiento, estas variables de identificación —extraídas en general en una región más chica de la primera capa— están prácticamente no correlacionadas a E_T^{iso} , como se ha mostrado

en la Fig. 6.11. Así, para obtener una distribución no sesgada del aislamiento para fotones del fondo, se construye una muestra «enriquecida» en este tipo de candidatos revirtiendo algunos de los criterios de la identificación *tight*. Es decir, se seleccionan aquellos candidatos que pasen todos los criterios *tight* excepto (al menos) uno de los cuatro siguientes: $w_{s,3}$, F_{side} , E_{ratio} y ΔE . Esta identificación, conocida como «*non-tight*», ha sido elegida para asegurar suficiente estadística en la muestra de fondo seleccionada, manteniendo una correlación débil con la variable de aislamiento que se busca modelar. La distribución de aislamiento obtenida mediante esta inversión se puede observar en la Fig. 7.1 para datos y simulaciones MC de señal+fondo, en comparación con la distribución esperada en simulaciones exclusivas de jets QCD. En general, ambas distribuciones en el MC tienen un comportamiento similar, indicando un buen modelaje del fondo mediante inversión de los criterios de identificación. Sin embargo, se pueden observar algunas diferencias entre la distribución de aislamiento extraída de los datos y aquellas simuladas en el MC, lo que indica la necesidad de estimar la contribución del fondo directamente de los primeros. El método aplicado en el análisis de esta tesis se describe a continuación.

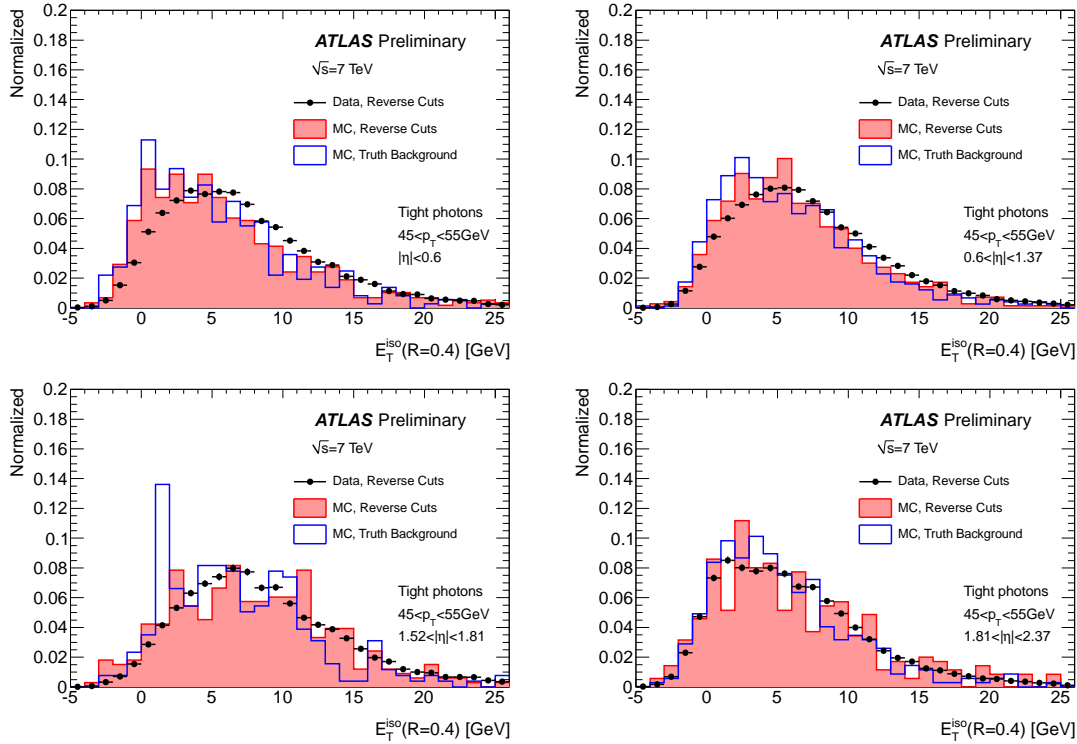


Figura 7.1. Distribución de la variable de aislamiento calorimétrico (E_T^{iso}) extraída de las muestras de datos y MC enriquecidas en fondo mediante la inversión de los cortes de identificación, en comparación a la observada en muestras MC de fondo (jets QCD), para candidatos con $45 \leq E_T < 55$ GeV como función de η .

7.2 El método de conteo bidimensional (ABCD)

Para estimar la cantidad de fondo en la muestra de candidatos seleccionada para la medida final de la sección eficaz de producción de fotones directos aislados, se ha utilizado el llamado método ABCD o de conteo bidimensional. Esta técnica está basada en la definición —en un dado espacio de variables— de regiones de control (o *sidebands*¹), pobladas esencialmente por eventos de fondo. Luego, a partir de la cantidad de eventos observados en estas regiones, se puede inferir la cantidad de fondo remanente en la muestra de interés (i.e. en la región con población de señal dominante), cuya composición es desconocida. En este análisis, las cuatro regiones han sido delimitadas en el espacio bidimensional definido por el aislamiento calorimétrico (E_T^{iso}) y por los criterios de identificación:

- REGIÓN A : candidatos aceptados por la identificación *tight*, con $E_T^{\text{iso}} < 3$ GeV
- REGIÓN B : candidatos aceptados por la identificación *tight*, con $E_T^{\text{iso}} > 5$ GeV
- REGIÓN C : candidatos aceptados por la identificación *non-tight*, con $E_T^{\text{iso}} < 3$ GeV
- REGIÓN D : candidatos aceptados por la identificación *non-tight*, con $E_T^{\text{iso}} > 5$ GeV

En la Fig. 7.2, se puede observar la distribución de eventos en este espacio (Id vs E_T^{iso}) para los datos correspondientes a este estudio, en donde se señalan la región de señal (A) y las distintas regiones de control (B,C,D).

Las ventajas de este método es que no se requiere un conocimiento preciso de la señal y que el fondo es modelado directamente de los datos, según la prescripción de la sección anterior. La estimación de la pureza de señal en la región A, se basa en dos suposiciones principales:

Sea N_λ^s (N_λ^b) el número de eventos de señal (fondo) observados en la región λ ($\lambda = A, B, C, D$), y fijando $N_\lambda = N_\lambda^s + N_\lambda^b$, entonces:

1. la contaminación de señal en las regiones de control es despreciable

$$\rightarrow N_{(B,C,D)} = N_{(B,C,D)}^b \quad (7.1)$$

2. el fondo no está correlacionado entre las distintas regiones

$$\rightarrow N_B^b/N_A^b = N_D^b/N_C^b \quad (7.2)$$

Bajo estas hipótesis, el número de eventos de señal esperado en la región A se puede estimar como:

$$N_A^s = N_A - N_A^b = N_A - \frac{N_B^b \cdot N_C^b}{N_D^b} = N_A - \frac{N_B \cdot N_C}{N_D} \quad (7.3)$$

¹Esta técnica es también referida como «*2D-sideband method*» dentro de ATLAS.

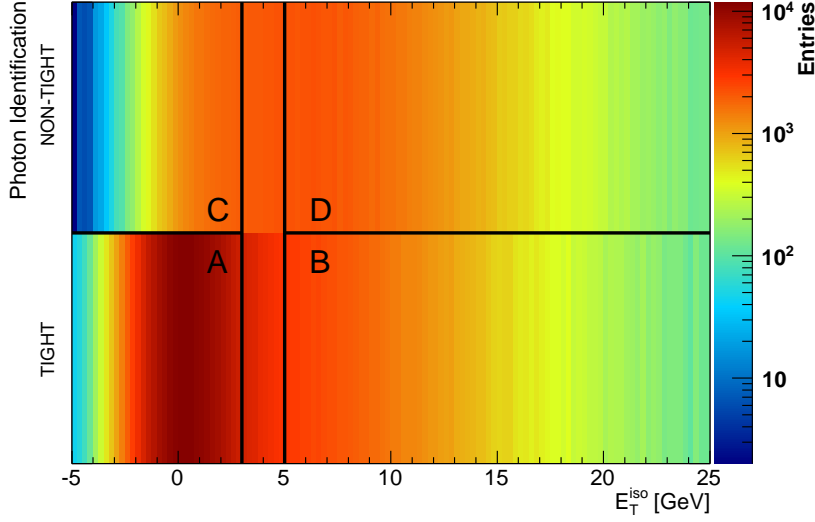


Figura 7.2. Método de conteo bidimensional (ABCD) : el número de eventos de fondo en la región A —donde yace la mayoría de los eventos de señal— es determinado por extrapolación de los eventos observados en las regiones de control dominadas por el fondo (B,C,D). La distribución de eventos en el espacio (Id vs E_T^{iso}) incluye todos los datos utilizados en este análisis (períodos G-I), aceptados por la preselección detallada en la Sec. 4.3.

simplemente a partir del número total de eventos observado en cada cuadrante. La pureza de señal de la muestra seleccionada queda definida entonces como:

$$P = \frac{N_A^s}{N_A} = 1 - \frac{N_B \cdot N_C}{N_A \cdot N_D} \quad (7.4)$$

En la Fig. 7.3 se muestra la aplicación del método, en distintos intervalos de E_T y η . La distribución de E_T^{iso} de los candidatos *non-tight* es normalizada a la de los candidatos *tight* en la zona no aislada ($E_T^{\text{iso}} > 5$ GeV), es decir, multiplicada por el factor de escala N_B/N_D . El acuerdo observado entre ambas distribuciones muestra cualitativamente la validez de las hipótesis originales. El exceso de la distribución para los candidatos *tight* en la región aislada $E_T^{\text{iso}} < 3$ GeV indica claramente la presencia de señal. Así, el número de eventos de señal en esta región es estimado como la diferencia entre la distribución de E_T^{iso} para los candidatos *tight* y *non-tight*, luego de la apropiada normalización de esta última.

En la sección próxima se describen los estudios de validez de las hipótesis del método, y las modificaciones introducidas en la aplicación del mismo para tener en cuenta la aproximación efectiva en cada caso.

7.2.1 Corrección por fuga de señal y correlación entre las regiones de control

La estimación de la pureza depende claramente de la separación entre la señal y el fondo, y de las regiones de control y normalización escogidas. Una contribución de la

7.2 EL MÉTODO DE CONTEO BIDIMENSIONAL (ABCD)

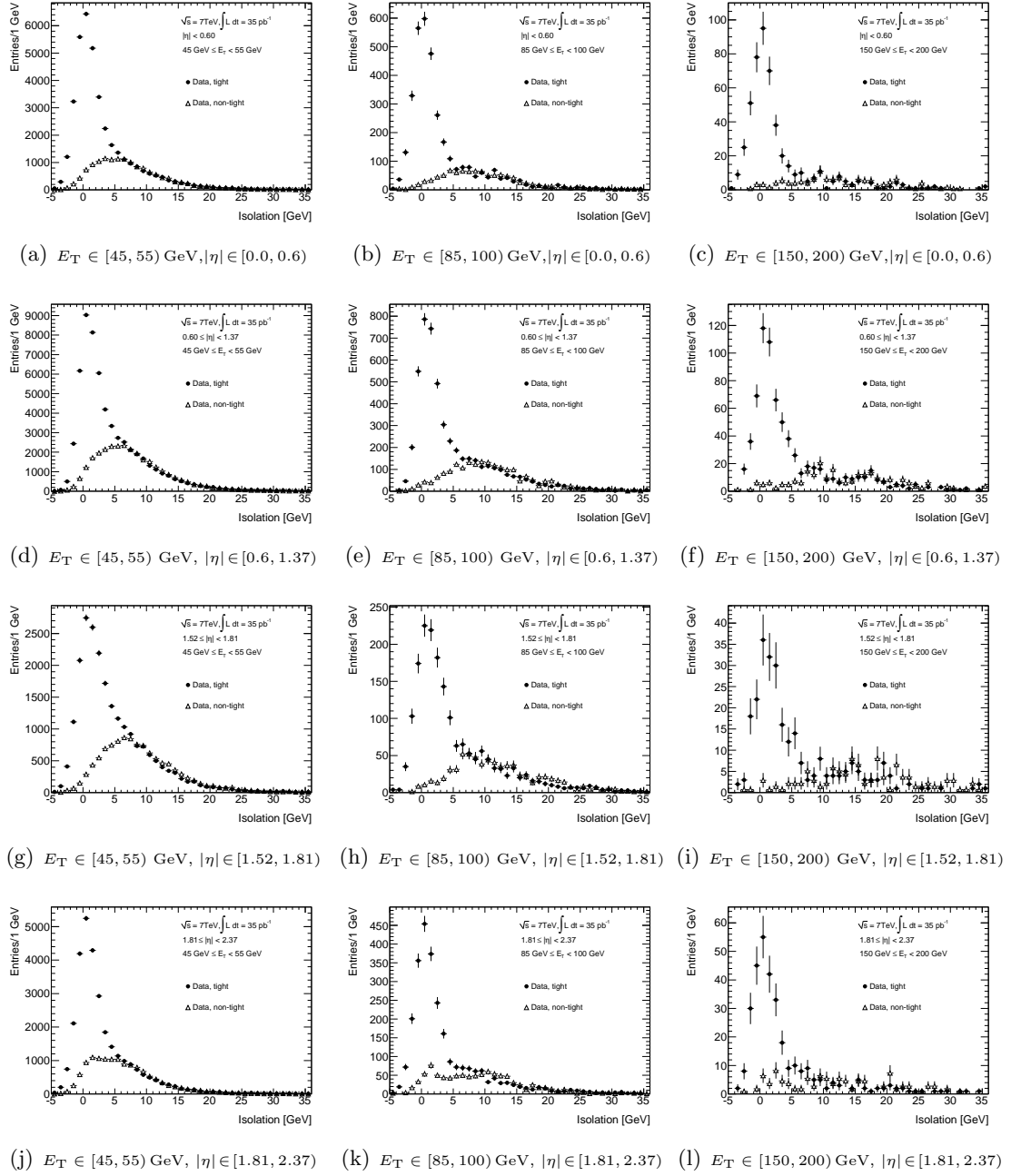


Figura 7.3. Distribución de la energía de aislamiento, E_T^{iso} , como función de E_T (columnas) y η (filas), observada en datos para fotones *tight* (puntos) y *non-tight* (triángulos). Esta última es normalizada a la primera en la región no aislada, $E_T^{\text{iso}} > 5 \text{ GeV}$, dominada por el fondo de QCD.

señal en las regiones de control *non-tight* (i.e. C,D) afectará la forma de la distribución del fondo a extraer. Asimismo, la contaminación de señal en la región de normalización (i.e. B,D) causará una sobrestimación de los eventos de fondo. La fracción de eventos de señal en cada región de control, $c_\lambda = N_\lambda^s/N_\lambda$ ($\lambda = \text{B,C,D}$), ha sido evaluada en muestras MC de fotones directos generadas con PYTHIA y figura en la Tabla 7.1. El efecto más importante ocurre en el cuadrante B (i.e. *tight*–no aislado), en donde la contaminación de

Tabla 7.1. Fracción de señal observada en cada una de las regiones de control, según las predicciones de PYTHIA, en los distintos intervalos de E_T y η . Las incertezas son en todos los casos < 0.01 .

E_T [GeV]	$ \eta < 0.6$			$0.6 \leq \eta < 1.37$			$1.52 \leq \eta < 1.81$			$1.81 \leq \eta < 2.37$		
	c_B	c_C	c_D	c_B	c_C	c_D	c_B	c_C	c_D	c_B	c_C	c_D
[45; 55)	0.07	0.03	0.00	0.08	0.03	0.00	0.12	0.04	0.02	0.08	0.05	0.01
[55; 70)	0.08	0.03	0.01	0.09	0.03	0.00	0.13	0.03	0.02	0.08	0.05	0.01
[70; 85)	0.08	0.02	0.00	0.10	0.03	0.00	0.14	0.03	0.02	0.09	0.05	0.02
[85; 100)	0.09	0.02	0.01	0.11	0.02	0.01	0.14	0.03	0.02	0.10	0.05	0.02
[100; 125)	0.09	0.02	0.01	0.10	0.02	0.01	0.15	0.03	0.02	0.10	0.05	0.02
[125; 150)	0.09	0.02	0.01	0.11	0.02	0.01	0.15	0.03	0.03	0.09	0.05	0.02
[150; 200)	0.09	0.02	0.01	0.11	0.02	0.01	0.17	0.03	0.03	0.11	0.06	0.02
[200; 400)	0.09	0.02	0.01	0.12	0.02	0.01	0.17	0.03	0.04	0.10	0.06	0.03

c_B : fuga de señal hacia la región *tight*, no aislada.

c_C : fuga de señal hacia la región *non-tight*, aislada.

c_D : fuga de señal hacia la región *non-tight*, no aislada.

señal observada alcanza un $\sim 17\%$ a alta E_T en el intervalo $1.52 \leq |\eta| < 1.81$, y es $\sim 10\%$ en el resto del detector. Para tener esto en cuenta, los factores c_λ son utilizados en la Ec. 7.3 para corregir el número de eventos observado en cada región de control:

$$N_\lambda \rightarrow N_\lambda - c_\lambda \cdot N_A, \quad \lambda = (B, C, D) \quad (7.5)$$

Por otro lado, la correlación entre las dos variables de discriminación causa una distorsión en la forma esperada de la distribución del fondo y, consecuentemente, se puede sub- o sobrestimar significativamente la contaminación de éste en la región de señal (A). La presencia de correlación implica una diferencia en los perfiles de aislamiento de los candidatos aceptados por la identificación *tight* o *non-tight*. Es decir, la cantidad

$$R^{bkg} = \frac{N_B^b \cdot N_C^b}{N_A^b \cdot N_D^b} \quad (7.6)$$

resulta $\neq 1$. Para tener en cuenta ambas desviaciones respecto de las suposiciones originales del método, la Ec. 7.3 es modificada, introduciendo las correcciones de las Ec. 7.5 y 7.6:

$$N_A^s = N_A - N_A^b = N_A - R^{bkg} \frac{N_B^b \cdot N_C^b}{N_D^b} = N_A - R^{bkg} \frac{(N_B - c_B N_A^s)(N_C - c_C N_A^s)}{N_D - c_D N_A^s} \quad (7.7)$$

que es ahora una ecuación cuadrática en N_A^s .

El efecto de correlación entre las variables ha sido evaluado en muestras de dijets QCD generadas con PYTHIA (círculos azules), como se observa en la Fig. 7.4, para los distintos intervalos de E_T y η . Aunque los resultados sugieren cierta correlación residual entre los distintos cuadrantes, las fluctuaciones estadísticas como función de E_T dificultan una apreciación concluyente.

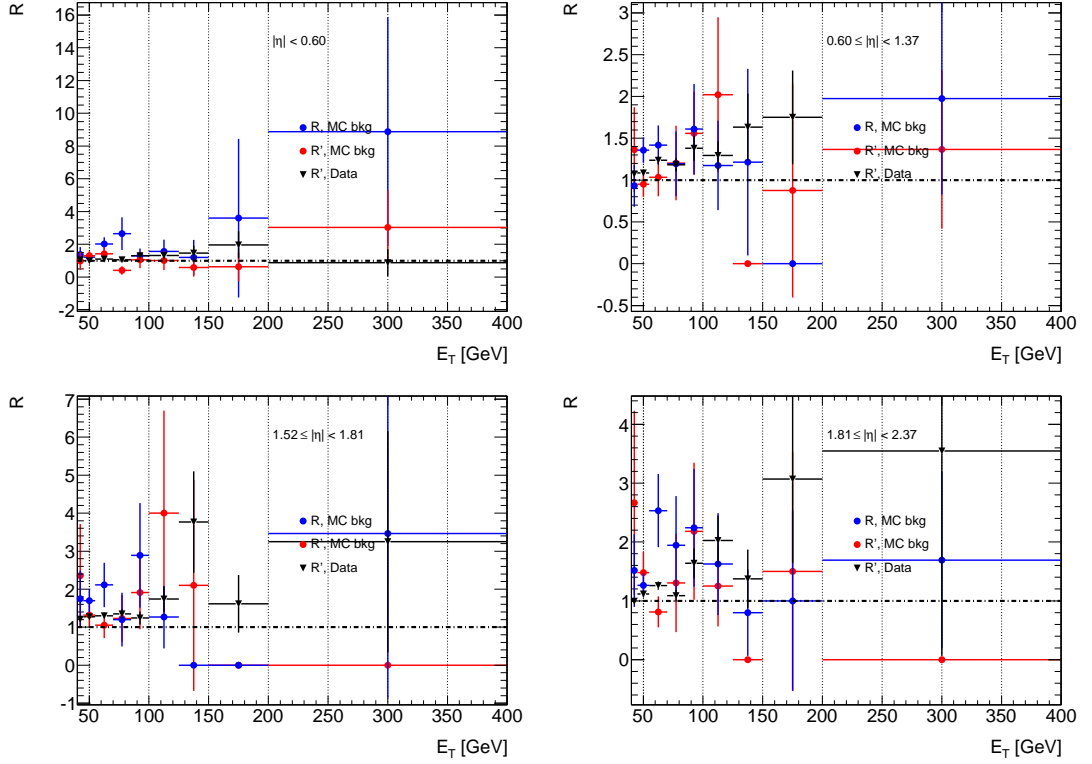


Figura 7.4. Factores de correlación (R_{bkg} , R') como función de E_T y η . Los valores de R_{bkg} (azul) y R' (rojo) son calculados en eventos de dijets generados con PYTHIA (Véase Tabla 4.3). En negro se muestran los valores de R' extraídos directamente de los datos. Las incertezas verticales son sólo estadísticas, las líneas horizontales indican el tamaño del intervalo de E_T considerado.

A fin de comparar estas predicciones, la presencia de correlación ha sido analizada también en datos. Sin embargo, el factor R^{bkg} no puede ser evaluado en este caso, debido a la componente de señal esperada (aún desconocida) en la región A. Para ello, se ha definido una cantidad similar

$$R' = \frac{N_{B'} \cdot N_{C'}}{N_{A'} \cdot N_{D'}} \quad (7.8)$$

en la región de alta energía de aislamiento ($E_T^{\text{iso}} > 6$ GeV), en donde se espera que la contaminación de señal sea despreciable (Fig. 7.5). Los valores de R' determinados tanto en datos como en MC se muestran en la Fig. 7.4. Nuevamente, las incertezas estadísticas no permiten una interpretación significativa de su comportamiento, especialmente a alta E_T . Tales fluctuaciones, de ser consideradas para la corrección de la Ec. 7.7, resultarían en un similar (e indeseable) comportamiento de la pureza estimada. Por esta razón, para el cálculo de los

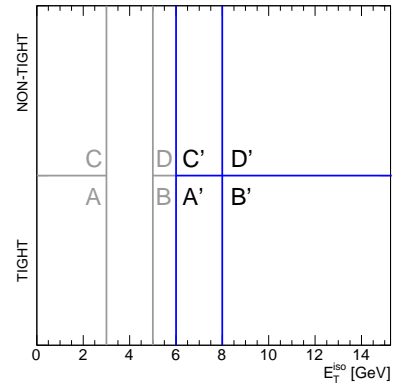


Figura 7.5. Esquema de las regiones de control redefinidas (A',B',C',D'), para estudiar la correlación del fondo directamente en datos.

valores centrales de la pureza se ha fijado $R^{bkg} = 1$. Las desviaciones observadas son consideradas en la estimación de las incertezas sistemáticas detallada en la sección siguiente.

7.3 Incertezas sistemáticas

Las principales fuentes de incerteza sobre la estimación del número de eventos de señal y de la pureza correspondiente en la muestra final son discutidas a continuación.

Definición de las regiones de control . $\sigma_{\text{bkg}}^{\text{nontight}}$, $\sigma_{\text{bkg}}^{\text{noniso}}$

La composición del fondo en las regiones de control, depende de los cortes de identificación invertidos y de la energía de aislamiento mínima requerida para la definición de las regiones de control *non-tight* y no aislada, respectivamente. En el primer caso, se requiere que los candidatos no satisfagan al menos uno de los cortes sobre $w_{s,3}$, F_{side} , E_{ratio} y ΔE . El efecto de tal elección sobre la estimación de la señal y de la pureza de la muestra final ha sido evaluado comparando los valores nominales con aquellos obtenidos por la inversión de sólo dos variables ($w_{s,3}$, F_{side})² o de todas las variables construidas en la primera capa del ECAL (i.e. agregando $w_{s,\text{tot}}$ a las cuatro variables nominales). Las diferencias observadas ($\sigma_{\text{bkg}}^{\text{nontight}}$) son $< 5\%$, decreciendo con E_{T} . La definición de la zona no aislada, $E_{\text{T}}^{\text{iso}} > 5 \text{ GeV}$, es tal que minimiza la contaminación de señal en las regiones de control, manteniendo una estadística suficiente en cada una de ellas para una estimación significativa de la pureza (Fig. 7.3). Una variación de $\pm 1 \text{ GeV}$ en el corte de aislamiento se traduce en una diferencia $\sigma_{\text{bkg}}^{\text{nontight}} < 1\%$ en la pureza estimada.

Correlación entre las regiones de control . $\sigma_{\text{bkg}}^{\text{corr}}$

La inclusión del coeficiente de correlación (R^{bkg}) resulta en un incremento (decremento) de la pureza estimada, si éste es < 1 (> 1). Dado que, como se ha mostrado en la Sec. 7.2.1, los valores de R^{bkg} varían significativamente alrededor de 1 con grandes incertezas estadísticas, se ha considerado sólo su valor global por región de η (integrando sobre todo el rango de E_{T}). La incerteza sistemática asociada ($\sigma_{\text{bkg}}^{\text{corr}}$) es determinada como la diferencia entre la pureza calculada conforme a la Ec. 7.7 usando este valor, y la pureza nominal (i.e. $R^{bkg} = 1$). El efecto es $< 5\%$ ($< 10\%$) en la zona del *barrel* (*endcap*), decreciendo con E_{T} según lo esperado, ya que la contribución del fondo disminuye como función de E_{T} ($\sim 1/P$).

Identificación y composición de la señal . $\sigma_{\text{sig}}^{\text{id}}$, $\sigma_{\text{sig}}^{\text{d/f}}$

La cantidad de señal presente en las regiones de control es estimada en muestras MC de fotones directos, incluyendo tanto fotones de la interacción dura (γ_d) como de

²aquellas que usan el menor número de celdas, dentro de la ventana espacial excluida en el cono de aislamiento.

fragmentación (γ_f). Sin embargo, como se ha mencionado en el Cap. 6, el MC no modela correctamente las (generalmente más anchas) distribuciones características de las lluvias electromagnéticas de fotones observadas en los datos. Por esta razón, la eficiencia nominal de la identificación *tight* ha sido modificada en un -3% , de acuerdo a la incerteza total sobre la eficiencia de identificación discutida en la Sec. 6.4. Esta variación incrementa la fracción de señal en la región *non-tight*, afectando principalmente al cálculo de c_C . El efecto estimado sobre la pureza ($\sigma_{\text{sig}}^{\text{id}}$) es $\sim 2\%$ en el *barrel* interno ($|\eta| < 0.6$), y varía entre 1.5% y 5% en el resto del detector, decreciendo con la energía transversa.

Además, el efecto de la composición de la señal (γ_d, γ_f) ha sido estudiado en muestras MC, en donde la fracción de γ_f varía conservativamente entre 0% y 100% . La diferencia en la pureza para las distintas mezclas ($\sigma_{\text{sig}}^{\text{d/f}}$), con respecto al valor nominal, es en todos los casos $< 8\%$, decreciendo con E_T .

Escala de energía . σ^{ES}

Como se ha mencionado en la Sec. 4.4, para corregir la escala de energía electromagnética medida por el ECAL en los datos se aplican ciertos factores de calibración. Las incertezas totales sobre estos factores ($\lesssim 0.5\%$ en el *barrel*, $\lesssim 1\%$ en el *endcap*) son utilizadas para variar el valor nominal y recalculando tanto la energía transversa del fotón como las variables de identificación sensibles a la escala de energía (i.e. $R_{\text{had}(1)}$ y ΔE). El efecto resultante sobre el número de eventos de señal varía entre 2 y 8% , mientras que el impacto sobre la pureza es $< 1\%$.

Concentración de material frente al ECAL . σ^{material}

Al igual que en el tratamiento de la Sec. 6.4.3, la cantidad de material previo al ECAL afecta la reconstrucción e identificación de fotones, y por ende a la población de eventos en las regiones de señal y de control. La sensibilidad del método de estimación del fondo a este efecto ha sido estudiada recalculando los factores de corrección c_λ , en muestras de señal MC con una geometría del detector «distorsionada» (Tabla 4.3). La variación de la pureza obtenida, respecto de la estimada con los factores derivados de la simulación nominal, es $< 1\%$ en todo el rango de E_T y η .

Corrección del aislamiento calorimétrico . $\sigma_{\text{iso}}^{\text{leak}}, \sigma_{\text{iso}}^{\text{ED}}$

La energía de aislamiento (E_T^{iso}) es calculada según la prescripción de la Sec. 5.3.1. Para tener en cuenta la energía de la lluvia electromagnética del propio fotón en el cono de aislamiento se aplican factores de corrección derivados de muestras MC, que remueven la dependencia de E_T^{iso} con E_T para fotones directos simulados. En datos, sin embargo, la dependencia es removida sólo parcialmente, producto del mayor desarrollo lateral de las lluvias observado con respecto al MC (Sec. 6.2.1). Correcciones adicionales han sido derivadas, en consecuencia, para compensar este comportamiento. La distribución de E_T^{iso}

para la señal, por cada intervalo de E_T y η , se define en los datos como la diferencia de aquella correspondiente a los fotones *tight* en la muestra y el perfil de aislamiento del fondo construido mediante la inversión de cortes detallada en la Sec. 7.1, normalizados al mismo número de eventos en la zona $E_T^{\text{iso}} > 8$ GeV (i.e. donde se espera sólo eventos de fondo, Fig. 7.3). El perfil de aislamiento resultante es ajustado a una función Gaussiana en la zona de alta pureza de señal, -2 GeV $\leq E_T^{\text{iso}} \leq 3$ GeV, para extraer la media (μ). En la Fig. 7.6 se pueden observar los valores medios estimados como función de E_T , en las distintas regiones de η , en donde la dependencia residual de la energía de aislamiento con E_T es evidente (aun luego de las correcciones aplicadas calculadas en el MC). La pendiente de la función lineal ajustada en cada caso ($c_{\text{leak}}^{\text{data}}$) ha sido utilizada para corregir E_T^{iso} como función de E_T , de acuerdo a la Ec. 5.9, donde ahora:

$$\mathcal{I}_{\text{leak}} = (c_{\text{leak}}^{\text{MC}} + c_{\text{leak}}^{\text{data}}) \times E_T \quad (7.9)$$

Las incertezas sobre los coeficientes $c_{\text{leak}}^{\text{data}}$ ($\lesssim 20\%$) son utilizadas para variar el valor central de los mismos, en cada intervalo de E_T y η . Una diferencia entre 1% y 4% es observada en el número de eventos de señal estimado, con respecto al valor nominal. El efecto en la pureza es $< 1\%$.

Los efectos remanentes del *in-time pile-up* sobre la energía de aislamiento corregida fueron también evaluados en datos. En la Fig. 7.7 se puede observar el valor medio de la distribución de E_T^{iso} —obtenido del ajuste gaussiano antes descrito— para electrones producto del decaimiento $W \rightarrow e\nu$ en datos y MC, como función del número de vértices primarios reconstruidos en el evento (N_{PV}). La dependencia observada en algunas regiones de η demuestra cierta insuficiencia de las correcciones por la actividad total subyacente aplicadas para remover los efectos no perturbativos que afectan la medida de E_T^{iso} (Sec. 5.3.1). Estas desviaciones son tratadas como incertezas sistemáticas en la definición del aislamiento, variando las correcciones nominales en un factor 1.5 tanto en los datos como en el MC, suficiente para reducir el comportamiento lineal con N_{PV} . El impacto sobre el número de eventos de señal oscila entre 2 y 8% en la región $1.52 \leq |\eta| < 1.81$, y entre 4 y 5% en el resto del detector. El efecto sobre la pureza es $< 1\%$, decreciendo con E_T .

Generador MC . σ^{gen}

La fracción de señal en las distintas regiones de control (c_λ) es estimada usando eventos MC generados con PYTHIA, incluyendo fotones de la interacción dura y de fragmentación. Para evaluar la dependencia del método a la elección particular de un generador, los factores c_λ fueron recalculados en muestras generadas con HERWIG (con la misma composición de señal anterior). El número de eventos de señal y la pureza correspondiente fueron estimados con estos nuevos factores, y comparados con los valores nominales. Sólo una pequeña diferencia $\sigma^{\text{gen}} < 1\%$ es observada, para todo E_T y η .

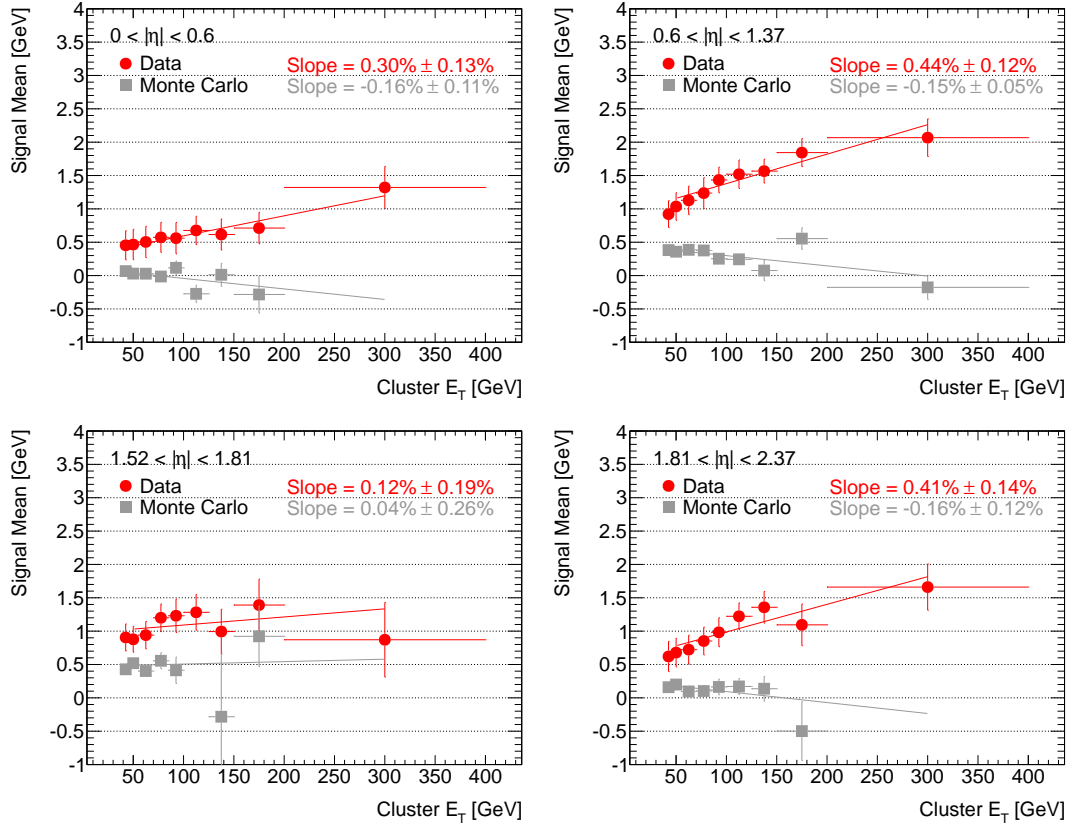


Figura 7.6. Valor medio (μ) de la energía de aislamiento estimado en los datos para fotones *tight*, como función de E_T en las distintas regiones de pseudorapidez. Las distribuciones de E_T^{iso} fueron ajustadas a una función Gaussiana en la región de alta pureza de señal ($-2 \text{ GeV} \leq E_T^{\text{iso}} \leq 3 \text{ GeV}$), de la cual se extrae el valor de μ . La dependencia de E_T^{iso} con E_T es clara evidencia de una fuga lateral de energía dentro del cono de aislamiento, más allá de aquella predicha por el MC.

En la Tabla 7.2 se resumen todas las incertezas sistemáticas sobre la estimación del número de eventos de señal estudiadas. La incerteza total es calculada como la suma cuadrática de todas las fuentes de incertezas, suponiendo que no existe correlación entre ellas.

7.4 Otras fuentes de fondo

Si bien el fondo de la producción de fotones directos aislados está ampliamente dominado por fotones del decaimiento de mesones livianos en jets de QCD, se han estudiado otras posibles fuentes de contaminación como la incorrecta identificación de electrones y la reconstrucción de falsos candidatos en eventos en donde no se ha producido una colisión pp .³

³ aunque el requerimiento sobre el vértice primario del evento elimina efectivamente esta última contribución.

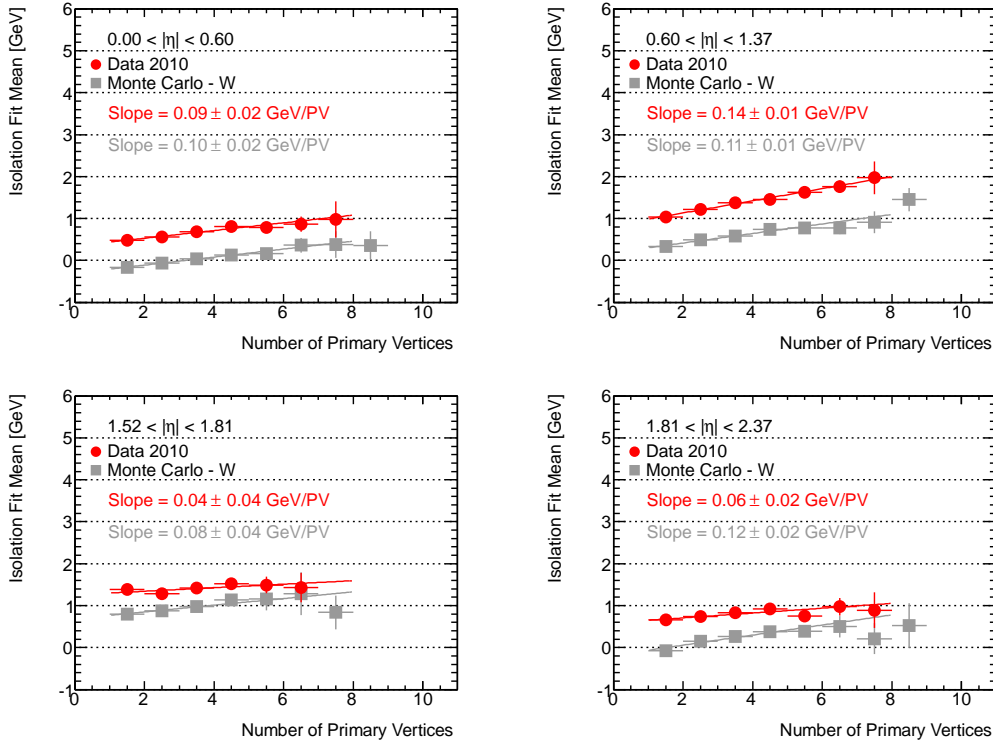


Figura 7.7. Valor medio de la energía de aislamiento estimado para electrones en eventos $W \rightarrow e\nu$ como función del número de vértices primarios reconstruidos, en las distintas regiones de pseudo-rapidez. Las distribuciones de E_T^{iso} fueron ajustadas a una función Gaussiana en la región de alta pureza de señal ($-2 \text{ GeV} \leq E_T^{\text{iso}} \leq 3 \text{ GeV}$), de la cual se extrae el valor de la media (μ). El impacto residual del *in-time pile-up* aun luego de las correcciones derivadas en MC es evidente, tanto en datos como en la simulación.

7.4.1 Electrones

Debido a la similitud de las lluvias electromagnéticas iniciadas por fotones y electrones, estos últimos pueden ser incorrectamente clasificados. En particular, esto ocurre cuando la traza del electrón no tiene un impacto asociado en la capa del detector interno más cercana al eje del haz (*B-layer*) y es entonces clasificado como un fotón convertido (Sec. 5.1).⁴ La principal contaminación de electrones luego de aplicar los criterios de identificación *tight* proviene del decaimiento de los quarks pesados *c* y *b*. Sin embargo, su energía de aislamiento es similar a la de los π^0 y otros mesones livianos, por lo que son tenidos en cuenta automáticamente en la estimación del fondo mediante el método ABCD. El decaimiento electrodébil de los bosones *W* y *Z* producen, en cambio, electrones aislados que pueden satisfacer todos los criterios de identificación y aislamiento aplicados para la selección de fotones directos, y requieren por lo tanto un tratamiento especial. Para estimar su contribución, se calcula primeramente la probabilidad de que un electrón sea reconstruido como un fotón ($\rho_{e\gamma}$), a partir del número de pares $e\gamma$ reconstruidos cuya

⁴ referido como «*single track conversion*» dentro de ATLAS.

Tabla 7.2. Incertezas sistemáticas (σ_i) sobre el número de fotones de señal estimado mediante el método de conteo bidimensional, en las distintas regiones de E_T y η . (Véase el texto para más detalles)

E_T [GeV]	$\sigma_{bkg}^{nontight}$	σ_{bkg}^{noniso}	σ_{sig}^{id}	$\sigma_{sig}^{d/f}$	σ_{bkg}^{corr}	σ^{ES}	$\sigma^{material}$	σ_{iso}^{ED}	σ_{iso}^{leak}	σ^{gen}
$ \eta < 0.6$										
[45, 55)	+379 -378	+58 -74	± 432	+1340 -448	± 1200	+218 -318	± 160	± 805	+124 -141	± 264
[55, 70)	+149 -183	+21 -30	± 287	+697 -190	± 484	+115 -244	± 6	± 566	+123 -58	± 125
[70, 85)	+32 -68	+1 -6	± 110	+243 -49	± 114	+50 -70	± 10	± 180	+40 -29	± 37
[85, 100)	+0 -22	+1 -1	± 50	+89 -19	± 33	+0 -28	± 1	± 84	+21 -18	± 15
[100, 125)	+2 -13	+1 -1	± 32	+64 -11	± 20	+36 -25	± 4	± 53	+17 -17	± 6
[125, 150)	+1 -1	+0 -0	± 11	+19 -3	± 4	+3 -15	± 3	± 15	+5 -5	± 3
[150, 200)	+0 -7	+0 -0	± 7	+13 -1	± 2	+5 -8	± 0	± 12	+8 -3	± 1
[200, 400)	+2 -2	+0 -0	± 3	+0 -0	± 0	+2 -5	± 0	± 8	+6 -5	± 0
$0.6 \leq \eta < 1.37$										
[45, 55)	+1120 -831	+62 -145	± 1086	+1429 -509	± 1517	+521 -1109	± 13	± 1454	+228 -218	± 216
[55, 70)	+523 -204	+34 -46	± 684	+611 -175	± 605	+270 -589	± 27	± 900	+165 -166	± 99
[70, 85)	+111 -125	+9 -12	± 291	+219 -56	± 136	+118 -262	± 7	± 303	+74 -84	± 50
[85, 100)	+1 -12	+2 -4	± 119	+55 -14	± 25	+52 -87	± 2	± 138	+27 -51	± 14
[100, 125)	+10 -0	+2 -4	± 71	+49 -9	± 20	+22 -100	± 12	± 88	+28 -43	± 7
[125, 150)	+0 -6	+0 -0	± 27	+9 -1	± 3	+21 -14	± 0	± 36	+16 -17	± 2
[150, 200)	+4 -0	+1 -1	± 13	+13 -1	± 3	+9 -18	± 3	± 24	+11 -7	± 2
[200, 400)	+1 -1	+0 -0	± 6	+6 -0	± 1	+4 -15	± 0	± 11	+5 -13	± 0
$1.52 \leq \eta < 1.81$										
[45, 55)	+60 -195	+42 -38	± 273	+309 -113	± 1016	+127 -390	± 47	± 664	+162 -153	± 77
[55, 70)	+27 -322	+12 -16	± 146	+160 -52	± 392	+141 -264	± 39	± 352	+96 -111	± 29
[70, 85)	+0 -44	+4 -5	± 52	+58 -15	± 87	+26 -110	± 4	± 109	+37 -56	± 10
[85, 100)	+6 -0	+1 -2	± 18	+15 -4	± 23	+20 -55	± 1	± 73	+30 -22	± 1
[100, 125)	+0 -3	+0 -1	± 11	+8 -2	± 8	+11 -26	± 2	± 33	+17 -10	± 1
[125, 150)	+0 -2	+0 -0	± 4	+4 -1	± 4	+13 -14	± 2	± 11	+5 -11	± 1
[150, 200)	+0 -0	+0 -0	± 3	+3 -0	± 3	+2 -6	± 1	± 11	+7 -4	± 0
[200, 400)	+0 -0	+0 -0	± 1	+0 -0	± 0	+1 -2	± 0	± 1	+4 -7	± 1
$1.81 \leq \eta < 2.37$										
[45, 55)	+594 -821	+21 -83	± 495	+1351 -483	± 1792	+215 -527	± 165	± 666	+94 -109	± 218
[55, 70)	+307 -494	+58 -43	± 333	+616 -189	± 785	+95 -344	± 46	± 401	+87 -91	± 89
[70, 85)	+96 -199	+11 -22	± 131	+236 -61	± 219	+52 -137	± 16	± 147	+41 -46	± 29
[85, 100)	+35 -56	+2 -6	± 55	+113 -22	± 76	+24 -79	± 7	± 78	+22 -22	± 12
[100, 125)	+25 -50	+2 -1	± 31	+69 -13	± 38	+19 -37	± 7	± 42	+17 -18	± 6
[125, 150)	+3 -5	+0 -0	± 8	+14 -3	± 7	+6 -11	± 2	± 13	+10 -6	± 1
[150, 200)	+3 -9	+1 -1	± 5	+10 -2	± 5	+0 -15	± 1	± 9	+5 -3	± 4
[200, 400)	+0 -1	+0 -0	± 1	+0 -0	± 0	+1 -0	± 0	± 2	+2 -2	± 0

masa invariante es compatible con la del bosón Z (i.e. en decaimientos $Z \rightarrow ee$) [138]. Los valores de $\rho_{e\gamma}$ obtenidos son ~ 0.08 en el *barrel* y ~ 0.13 en el *endcap*. El aumento de la probabilidad en el *endcap* ($|\eta| > 1.52$) responde a una mayor concentración de material inactivo en esta región (i.e. mayor probabilidad de conversión de los fotones).

La fracción de electrones reales esperada en la muestra de candidatos reconstruidos (I), luego de aplicar los criterios de identificación y aislamiento (i.e. en la región A), es

entonces:

$$I = \frac{\rho_{e\gamma} \cdot L \cdot \sigma_e}{N_A} \quad (7.10)$$

donde L es la luminosidad integrada correspondiente a los datos analizados, N_A es el número de fotones observados y $\sigma_e = \sigma_{eW}(\sigma_{eZ})$ es la sección eficaz MC de producción de electrones vía el decaimiento $W \rightarrow e\nu$ ($Z \rightarrow ee$). La contaminación de electrones en la muestra de fotones seleccionada varía entre 1.6% (1.7%) en el *barrel* (*endcap*) a bajo p_T y $\sim 1\%$ a alto p_T en todo el rango de η , como se observa en la Fig. 7.8. Estos valores han sido utilizados para corregir el número de eventos de señal estimado (Tabla 7.3), al momento de medir la sección eficaz de fotones directos aislados detallada en el capítulo siguiente.

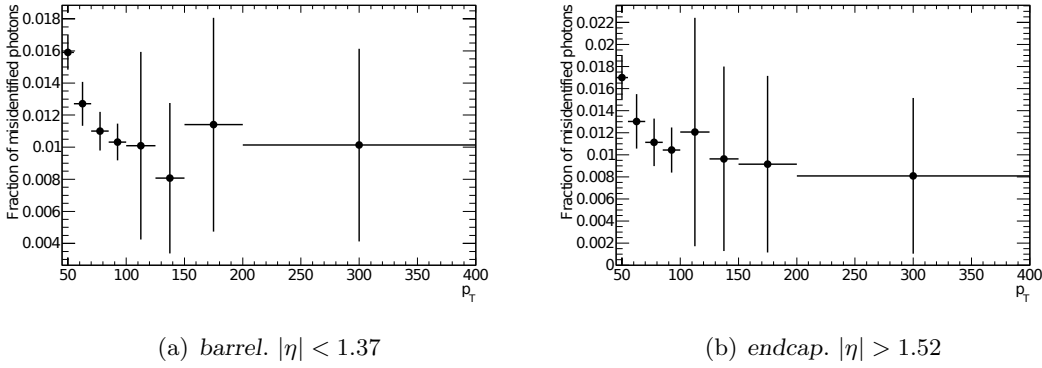


Figura 7.8. Fracción de candidatos a fotón correspondientes a electrones mal identificados provenientes del decaimiento de los bosones W y Z , como función de p_T , separadamente para la región del a) *barrel* y del b) *endcap*. La contribución más importante proviene del decaimiento $W \rightarrow e\nu$.

7.4.2 Rayos cósmicos

Las deposiciones de energía en el ECAL causadas por rayos cósmicos que atraviesan el detector —en su mayoría muones— pueden dar lugar a la reconstrucción de un fotón aislado. La contribución de estos falsos candidatos a la muestra de fotones final ha sido evaluada en eventos seleccionados por un *trigger* especial,⁵ el cual asegura que ninguno de los haces de protones atraviesa el punto de interacción y requiere además un cluster reconstruido en el ECAL con $E_T > 3$ GeV. En la Fig. 7.9(a) se puede observar la baja tasa de contaminación estimada, aun cuando sólo los criterios de identificación *loose* han sido aplicados. Luego de imponer los criterios de aislamiento e identificación *tight*, y requerir la presencia de un vértice primario en el evento, la contribución de este fondo es completamente despreciable.

⁵L1_EM3_EMPTY en la nomenclatura del trigger de ATLAS.

7.4.3 *Beam-gas y beam halo*

Producto de las imperfecciones del alto vacío en el tubo en que circulan los haces de protones (*beam pipe*), el haz puede interactuar con moléculas de gas remanentes, principalmente H_2 , CH_4 , CO y CO_2 [139]. Asimismo, la lluvia de partículas secundarias generadas por la interacción de los protones del haz más externos (*beam halo*) con el material de los colimadores dispuestos a lo largo del LHC a ambos lados del detector puede generar deposiciones de energía significativas en el ECAL, mayormente en el *endcap*. La reconstrucción de falsos candidatos debido a estos procesos es evaluada en los datos de eventos seleccionados por un *trigger* dedicado,⁶ donde se requiere la presencia de sólo un paquete de protones atravesando el punto de interacción (aislado de cualquier otro cruce de haces). La tasa de contaminación total calculada se observa en la Fig. 7.9(b), para todos los candidatos reconstruidos como función de p_T , en las distintas regiones de η . Luego de aplicar los criterios de identificación y aislamiento, aun sin requerir un vértice primario en el evento, la fracción es $< 0.1\%$. Esta contribución del fondo es, por tanto, ignorada en la estimación final de la pureza.

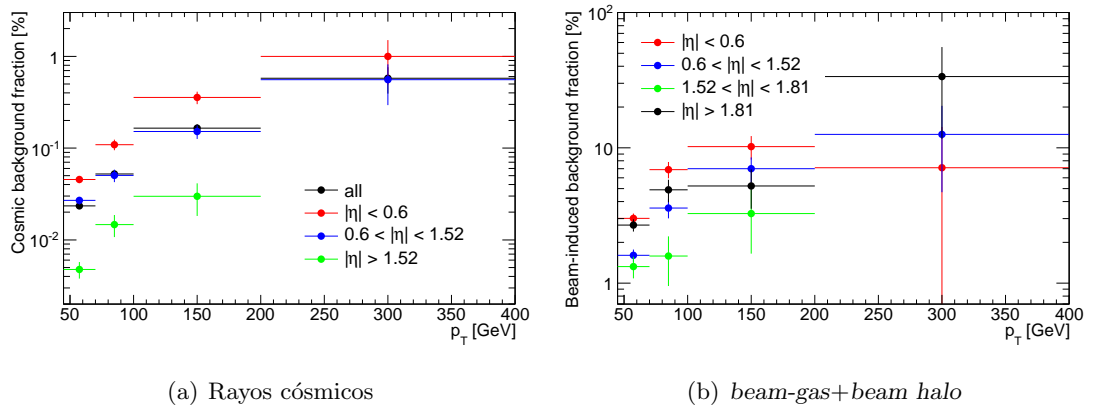


Figura 7.9. Fracción de los candidatos a fotón originados a) por rayos cósmicos o b) en eventos sin colisión pp (i.e. *beam-gas+beam-halo*), estimada en datos como función de p_T para las distintas regiones de η . En a) se han aplicado sólo los criterios de identificación *loose*, en b) se incluyen todos los candidatos reconstruidos.

7.5 Resultados

El número de eventos de señal y la pureza correspondiente estimados en la muestra final de candidatos seleccionada para la medida de la sección eficaz de producción de fotones directos aislados se listan en las Tabla 7.3, y se muestran en las Fig. 7.10 y 7.11, respectivamente. En ambos casos, las incertezas estadísticas son calculadas mediante la propagación de aquellas observadas en el número de eventos y los coeficientes c_λ calculados en cada

⁶L1_EM2_UNPAIRED_ISO en la nomenclatura del trigger de ATLAS.

cuadrante, según la Ec. 7.7. Las incertezas sistemáticas totales se han obtenido como la suma cuadrática de todas aquellas detalladas en la sección anterior y listadas en la Tabla 7.2.

Tabla 7.3. Número de eventos de señal y la correspondiente pureza en la muestra de datos final, estimada mediante el método de conteo bidimensional (ABCD), para cada intervalo de E_T y η . La primera incerteza es estadística, la segunda es la sistemática total.

Eventos señal				
E_T [GeV]	$ \eta < 0.6$	$0.6 \leq \eta < 1.37$	$1.52 \leq \eta < 1.81$	$1.81 \leq \eta < 2.37$
[45; 55)	$22991 \pm 177^{+1832}_{-1336}$	$28166 \pm 215^{+3077}_{-2842}$	$10226 \pm 118^{+892}_{-938}$	$17110 \pm 173^{+1886}_{-1579}$
[55; 70)	$13663 \pm 131^{+1066}_{-849}$	$16071 \pm 155^{+1760}_{-1667}$	$5795 \pm 85^{+505}_{-615}$	$9750 \pm 127^{+1024}_{-969}$
[70; 85)	$5061 \pm 77^{+364}_{-295}$	$6269 \pm 92^{+578}_{-585}$	$2212 \pm 51^{+166}_{-199}$	$3608 \pm 75^{+386}_{-371}$
[85; 100)	$2333 \pm 51^{+134}_{-122}$	$2748 \pm 57^{+183}_{-195}$	$923 \pm 32^{+81}_{-94}$	$1608 \pm 48^{+177}_{-159}$
[100; 125)	$1532 \pm 41^{+91}_{-78}$	$1899 \pm 47^{+122}_{-151}$	$596 \pm 25^{+43}_{-47}$	$1006 \pm 37^{+96}_{-98}$
[125; 150)	$571 \pm 25^{+24}_{-27}$	$713 \pm 28^{+42}_{-47}$	$225 \pm 16^{+27}_{-27}$	$331 \pm 20^{+24}_{-23}$
[150; 200)	$362 \pm 20^{+19}_{-19}$	$403 \pm 21^{+30}_{-32}$	$140 \pm 12^{+17}_{-19}$	$208 \pm 16^{+18}_{-25}$
[200; 400)	$158 \pm 13^{+13}_{-13}$	$189 \pm 15^{+17}_{-26}$	$55 \pm 7^{+12}_{-14}$	$56 \pm 7^{+8}_{-8}$

Pureza				
E_T [GeV]	$ \eta < 0.6$	$0.6 \leq \eta < 1.37$	$1.52 \leq \eta < 1.81$	$1.81 \leq \eta < 2.37$
[45; 55)	$0.91 \pm 0.00^{+0.07}_{-0.05}$	$0.87 \pm 0.00^{+0.09}_{-0.07}$	$0.91 \pm 0.00^{+0.06}_{-0.06}$	$0.87 \pm 0.00^{+0.10}_{-0.08}$
[55; 70)	$0.93 \pm 0.00^{+0.07}_{-0.05}$	$0.90 \pm 0.00^{+0.09}_{-0.07}$	$0.94 \pm 0.00^{+0.06}_{-0.07}$	$0.89 \pm 0.00^{+0.09}_{-0.08}$
[70; 85)	$0.96 \pm 0.00^{+0.04}_{-0.05}$	$0.94 \pm 0.00^{+0.06}_{-0.07}$	$0.96 \pm 0.00^{+0.04}_{-0.05}$	$0.92 \pm 0.00^{+0.08}_{-0.09}$
[85; 100)	$0.97 \pm 0.00^{+0.03}_{-0.04}$	$0.97 \pm 0.00^{+0.03}_{-0.04}$	$0.98 \pm 0.01^{+0.02}_{-0.04}$	$0.93 \pm 0.01^{+0.07}_{-0.08}$
[100; 125)	$0.97 \pm 0.00^{+0.03}_{-0.04}$	$0.97 \pm 0.00^{+0.03}_{-0.04}$	$0.99 \pm 0.00^{+0.01}_{-0.03}$	$0.95 \pm 0.01^{+0.05}_{-0.08}$
[125; 150)	$0.98 \pm 0.01^{+0.02}_{-0.03}$	$0.99 \pm 0.00^{+0.01}_{-0.03}$	$0.98 \pm 0.01^{+0.02}_{-0.04}$	$0.97 \pm 0.01^{+0.03}_{-0.05}$
[150; 200)	$0.99 \pm 0.01^{+0.01}_{-0.03}$	$0.98 \pm 0.01^{+0.02}_{-0.03}$	$0.98 \pm 0.01^{+0.02}_{-0.04}$	$0.96 \pm 0.01^{+0.04}_{-0.08}$
[200; 400)	$1.00 \pm 0.00^{+0.00}_{-0.03}$	$0.98 \pm 0.01^{+0.02}_{-0.03}$	$1.00 \pm 0.00^{+0.00}_{-0.03}$	$1.00 \pm 0.00^{+0.00}_{-0.04}$

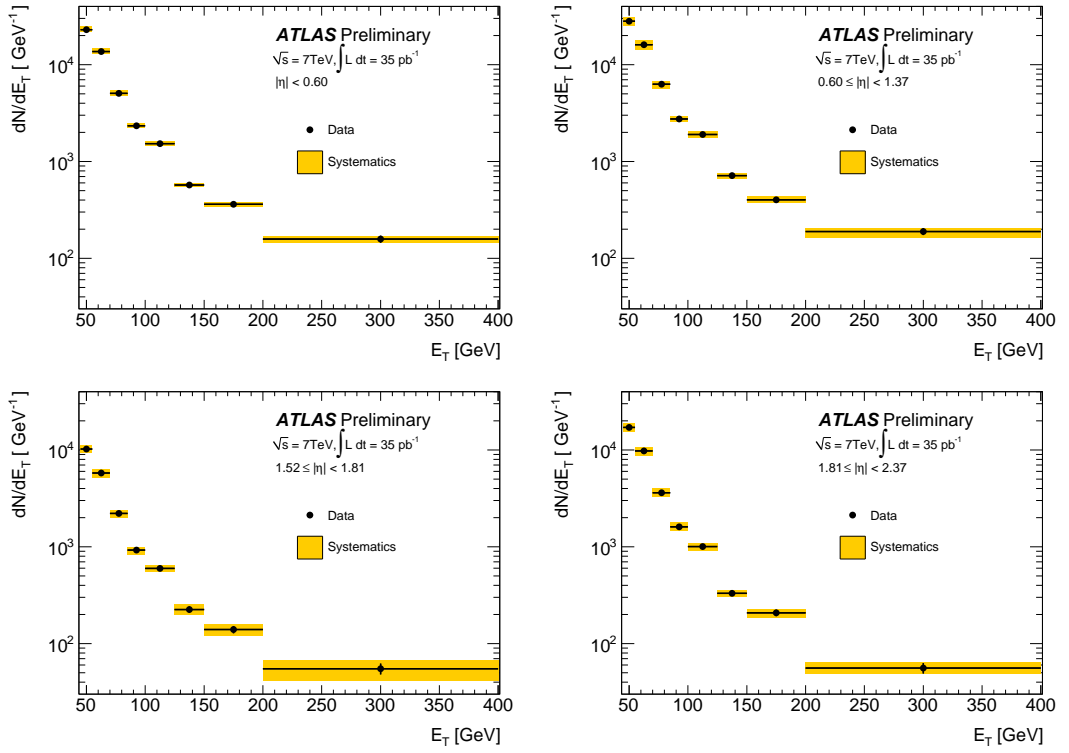


Figura 7.10. Número de eventos de señal estimados con el método de conteo bidimensional (ABCD), como función de E_T y η . Las bandas representan las incertezas sistemáticas totales.

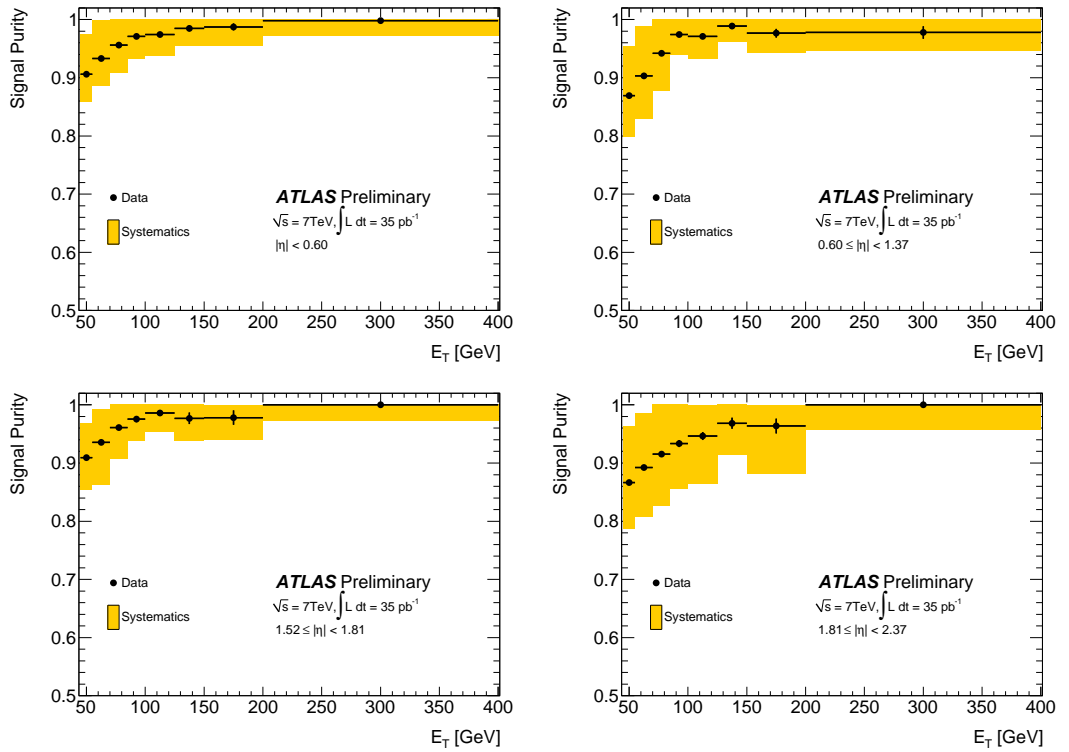


Figura 7.11. Pureza de señal en la muestra final estimada con el método de conteo bidimensional (ABCD), como función de E_T y η . Las bandas representan las incertezas sistemáticas totales.

