Búsqueda de fuentes puntuales con el algoritmo EM en el telescopio de Neutrinos ANTARES

XXX Reunión Bienal Real Sociedad Española de Física

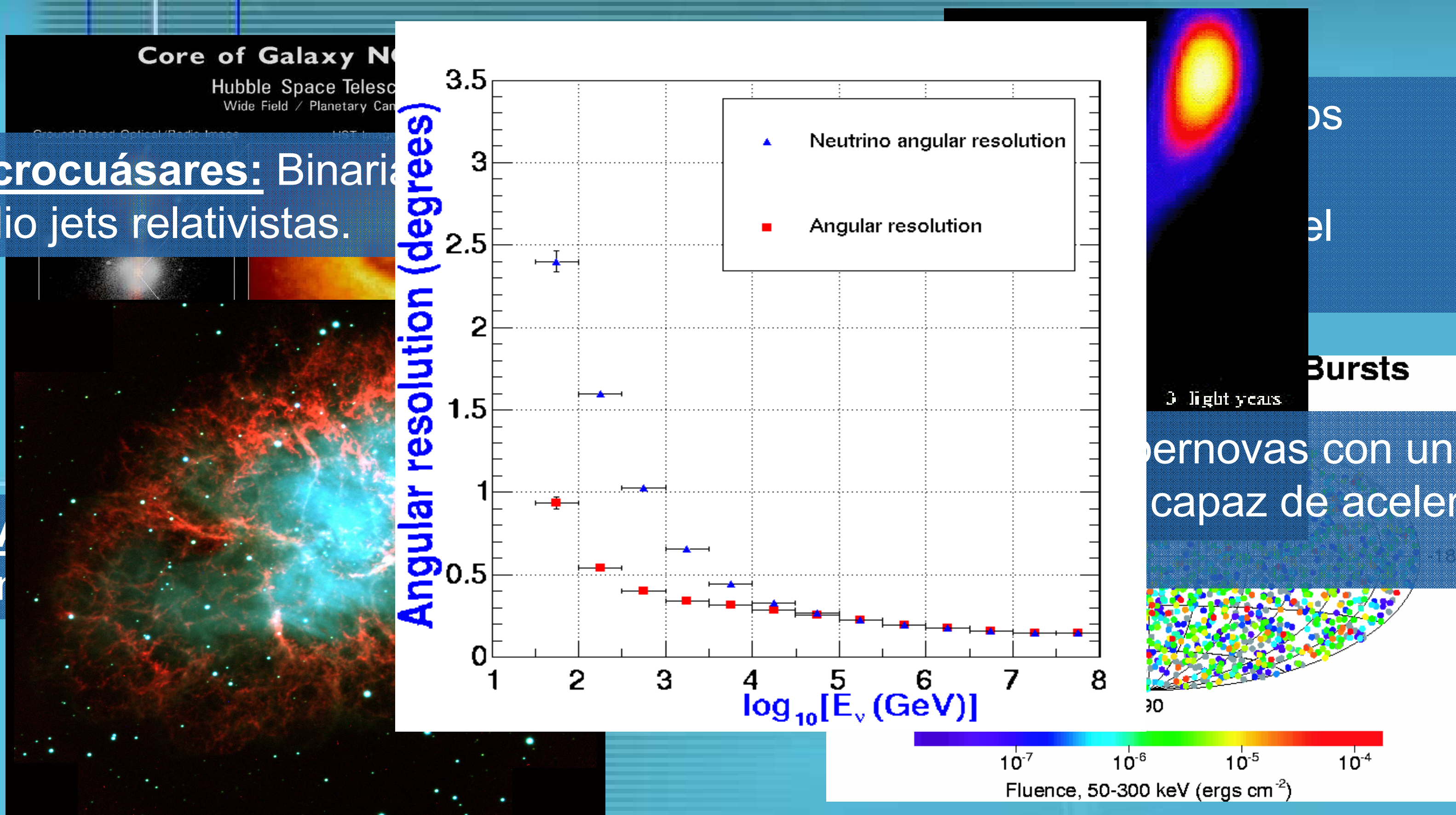
Física de Altas Energías

Juan Antonio Aguilar Sánchez, IFIC

Motivación

- La detección de fuentes emisoras de neutrinos sería un evidencia de los modelos de aceleración hadrónica en los procesos de aceleración de rayos cósmicos (mecanismo Fermi).
- ANTARES tiene una gran resolución angular por lo que tiene un gran potencial para la búsqueda de fuentes de neutrinos.

Microcuásares: Binari
radio jets relativistas.

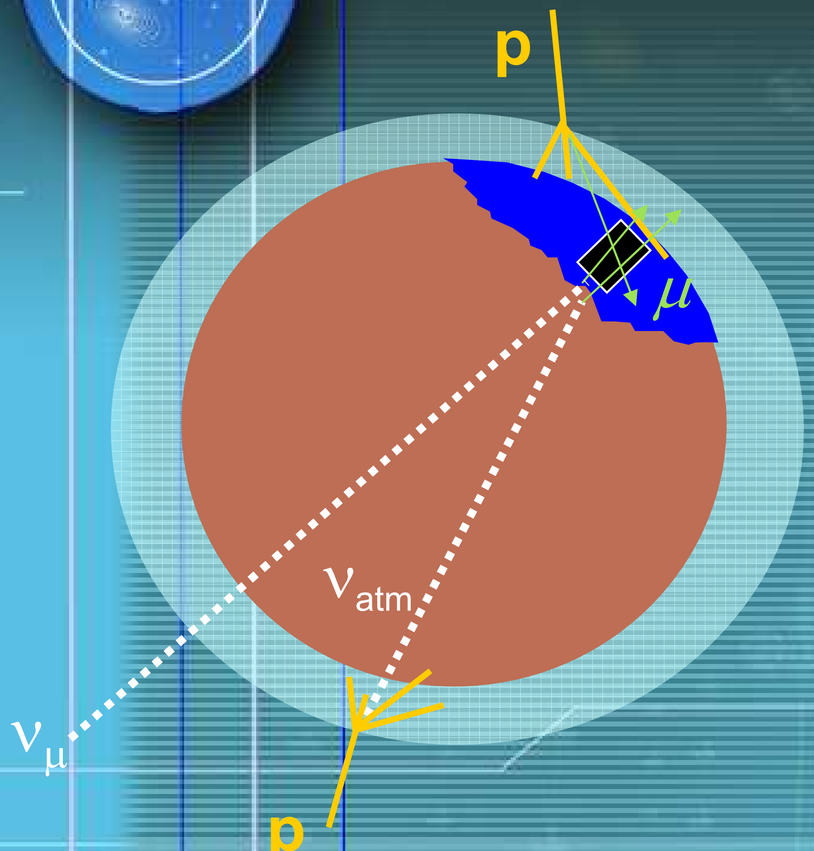


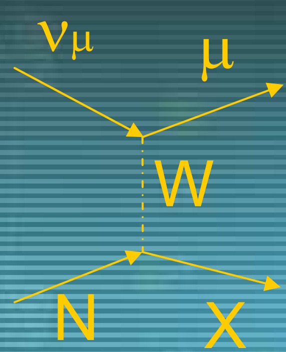
os

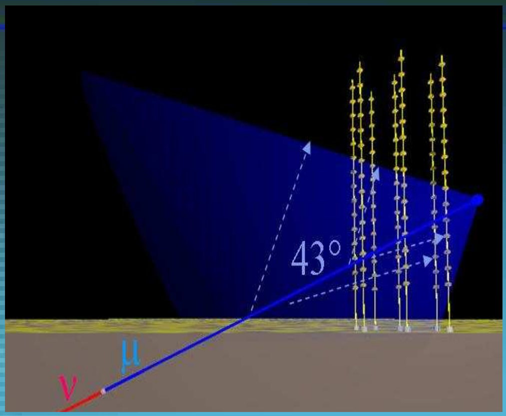
el

ernovas con una
capaz de acelerar

Principio de detección





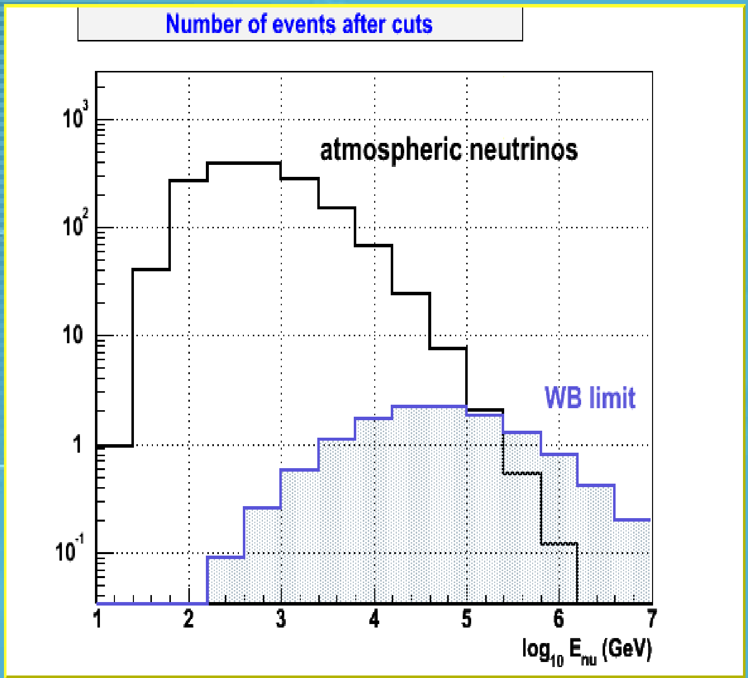


$$p \rightarrow \pi^{+}(+K^{+}...) \rightarrow \mu^{+} + \nu_{\mu}$$

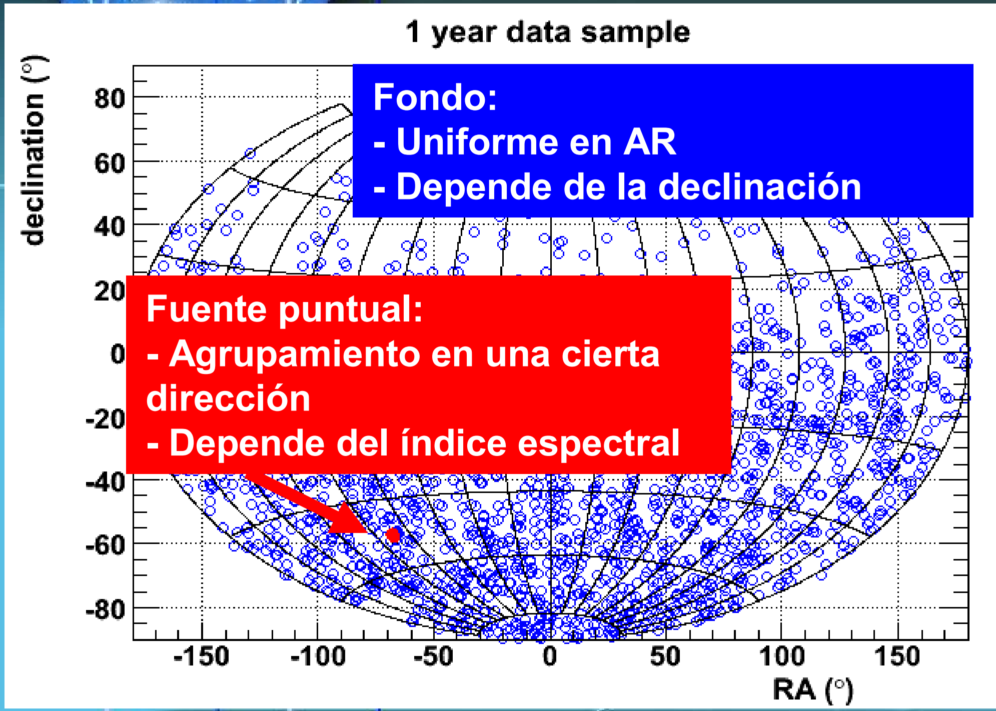
$$\alpha \quad e^{+} + \bar{\nu}_{\mu} + \nu_{e}$$

$$n \rightarrow \pi^{-}(+K^{-}...) \rightarrow \mu^{-} + \bar{\nu}_{\mu}$$

$$\alpha \quad e^{-} + \nu_{\mu} + \bar{\nu}_{e}$$



Búsqueda de fuentes puntuales



- Neutrinos atmosféricos constituyen el mayor fondo en un telescopio de neutrinos.

  ~1832 atmospheric ν / year
  + 138 single μ / year
  + 98 multi-μ / year
  = ~ 2068 bg. events

- El flujo de neutrinos cósmicos es muy pequeño, pero se concentran en ciertas direcciones.

- Búsqueda de fuentes puntuales: identificación de agrupamientos de sucesos sobre un fondo de neutrinos atmosféricos.

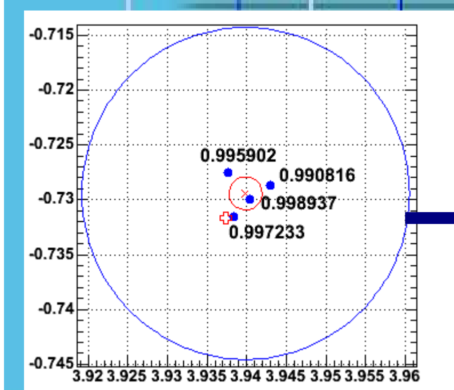
Algoritmo EM
– Búsqueda de clusters.
– Fuentes con distribuciones Gausianas sobre una distribución de fondo.
– Significancia basada en probabilidad del fondo para producir acumulación de sucesos.

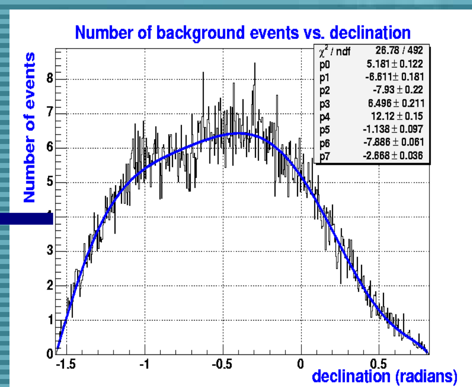
Algoritmo EM: Modelo de Mezcla

Construimos la pdf basándonos en modelo de mezclas:

$$p(\mathbf{x}) = \sum_{j=1}^{g} \pi_j \, p(\mathbf{x};\boldsymbol{\theta}_j)$$

- g es el número de componentes en el modelo
- $\pi_j \geq 0$, proporciones de mezcla ($\Sigma\, \pi_j = 1$)
- $p(\mathbf{x};\theta_j)$, j=1,…,g *funciones densidad* de las componentes





Nuestro caso: 1) Pdf = fondo + señal

2) $x = (\alpha_{RA}, \delta)$

$$\sum_{j=1}^{g} \pi_j p(\mathbf{x};\boldsymbol{\theta}_j) = \underbrace{\pi_1 \frac{1}{2\pi} \mathrm{P}_7(\delta)}_{\text{Fondo}} + \underbrace{\sum_{j=2}^{g} \pi_j p_{Gauss}(\alpha_{\mathrm{RA}}, \delta \,|\, \boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)}_{\text{Fuentes}}$$

No se usa la energía de los sucesos!

Asumimos que las distribuciones de densidad de las fuentes son Gausianas:

$$p_{Gauss}(\mathbf{x};\boldsymbol{\mu}_j,\boldsymbol{\Sigma}_j) = \frac{\exp\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_j)^T \boldsymbol{\Sigma}_j^{-1}(\mathbf{x}-\boldsymbol{\mu}_j)\}}{\sqrt{\det(2\pi\boldsymbol{\Sigma}_j)}}$$

Posición de la fuente

$\boldsymbol{\theta}_j = (\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j)$ donde $\boldsymbol{\mu} = (\mu_x, \mu_y)$ $\boldsymbol{\Sigma} = (\sigma_x, \sigma_y, \sigma_{xy})$

Algoritmo EM: Método General

Dado un conjunto de n observaciones la verosimilutud es:

$$L(\boldsymbol{\Psi}) = p(\{\mathbf{x}\},\boldsymbol{\Psi}) = \prod_{i=1}^{n}\sum_{j=1}^{g}\pi_j p(\mathbf{x}_i;\boldsymbol{\theta}_j)$$

– Donde $\boldsymbol{\Psi}$ es un vector que denota el conjunto de parámetros $\{\pi_1,\ldots,\pi_g,\boldsymbol{\theta}_1,\ldots,\boldsymbol{\theta}_g\}$

– **Esta verosimilitud en principio no es maximizable analíticamente!**

Conjunto INCOMPLETO		Conjunto COMPLETO
$\{\mathbf{x}\} = (\alpha_{RA},\delta)$	$\rightarrow$	$\{\mathbf{y}\}$ $\mathbf{y}_i=(\mathbf{x}_i,\mathbf{z}_i)$

$$z_{ik}\begin{cases} 1 \text{ si } \mathbf{x}_i \text{ pertenece al grupo k} \\ 0 \text{ otros} \end{cases}$$

Las nuevas funciones de densidad: $g(\mathbf{y};\boldsymbol{\Psi}) = g(\mathbf{x},\mathbf{z};\boldsymbol{\Psi}) = f(\mathbf{x}\,|\,\mathbf{z},\boldsymbol{\Psi})\,p(\mathbf{x};\boldsymbol{\theta})$

Verosimilitud del conjunto completo: $L'(\boldsymbol{\Psi}) = g(\{\mathbf{y}\},\boldsymbol{\Psi})$

La verosimilitud "incompleta" se obtiene integrando sobre los valores posibles de {**y**} donde {**x**} está embebido:

$$L(\boldsymbol{\Psi}) = p(\{\mathbf{x}\},\boldsymbol{\Psi}) = \int\prod_{i=1}^{n} g(\mathbf{x}_i,\mathbf{z};\boldsymbol{\Psi})\,dz$$

E-Step and M-Step

Dos pasos: Expectation-step y Maximization-step:

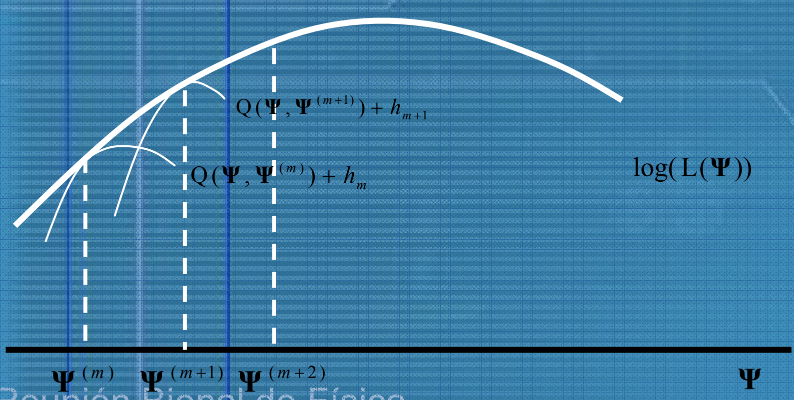
E-Step:

- Empezamos con un conjunto inicial de parámetros $\boldsymbol{\Psi}^{(m)}$
- Calculamos el valor esperado de la log-likelihood del conjunto completo, condicionado en los datos observados $\{\mathbf{x}\}$

$$Q(\boldsymbol{\Psi},\boldsymbol{\Psi}^{(m)}) = E[\log(g(\{\mathbf{y}\};\boldsymbol{\Psi})) \mid p(\{\mathbf{x}\};\boldsymbol{\Psi}^{(m)})]$$

M-Step:

- Encontrar $\boldsymbol{\Psi}=\{\boldsymbol{\Psi}^{(m+1)}\}$ que maximiza $Q(\boldsymbol{\Psi},\boldsymbol{\Psi}^{(m)})$



Maximizaciones sucesivas de la función $Q(\boldsymbol{\Psi},\boldsymbol{\Psi}^{(m)})$ llevan a un incremento en la log-likelihood

Selección de modelo (BIC)

1. Conjunto de datos D
2. Varios modelos $M_1, \ldots, M_k$

$\Rightarrow$

1. M_0 = fondo
2. M_1 = fondo + señal

Necesitamos un parámetro λ(D) que nos diga si nuestros datos se ajustan mejor al modelo M_0 o el modelo M_1

Integración sobre el espacio de parámetros desconocidos θ_k

Probabilidad de darse M_k dado D

$$p(D|M_k) = \int p(D|\theta_k, M_k)\, p(\theta_k|M_k)\, d\theta_k$$

Bayesian Information Criterion o BIC*:

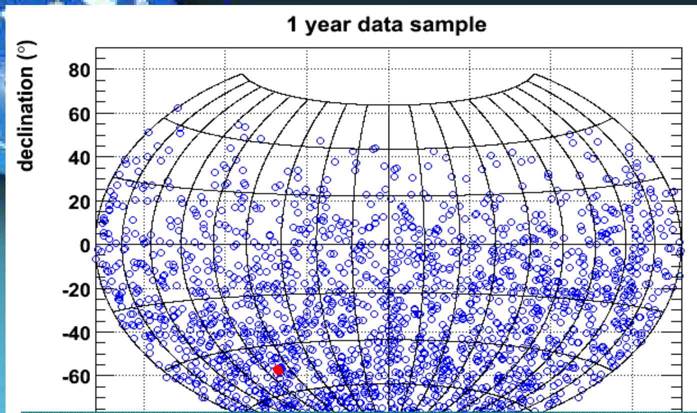
g.d.l (parámetros a ajustar)

$$2\log[p(D|M_k)] \approx 2\log p(D|\hat{\theta}_k, M_k) - v_k \log(n) = \mathrm{BIC}_k$$

$$\mathrm{BIC}_k = \underbrace{2\sum_{i=n}^{n} \log\left[p_{Gauss}(\hat{\alpha}_{RAi}, \hat{\delta}_i) + p_{BG}(\hat{\delta}_i)\right]}_{\text{Fondo + señal}} - \underbrace{2\sum_{i=n}^{n} \log\left[p_{BG}(\hat{\delta}_i)\right]}_{\text{Fondo}} - \underbrace{6\log(n)}_{\text{6 parámetros}}$$

*G. Schwarz. *Estimating the dimension of a model.* Ann. of Statist. 6, 461-464 (1978)

Procedimiento



Realizamos 3000 experimentos cada uno equivalente a 1 año de toma de datos

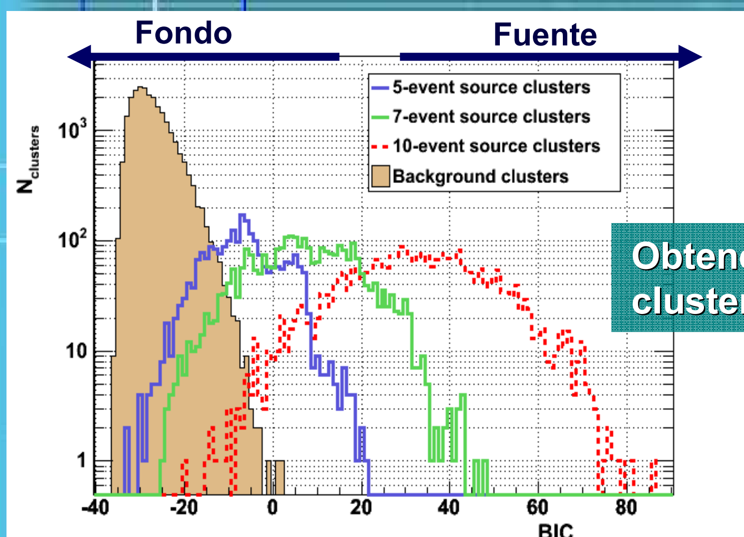
Algoritmo sencillo que identifica cada *cluster* de la muestra

Cada cluster se ajusta por el algoritmo EM

Algoritmo EM:

Expectation

Maximization



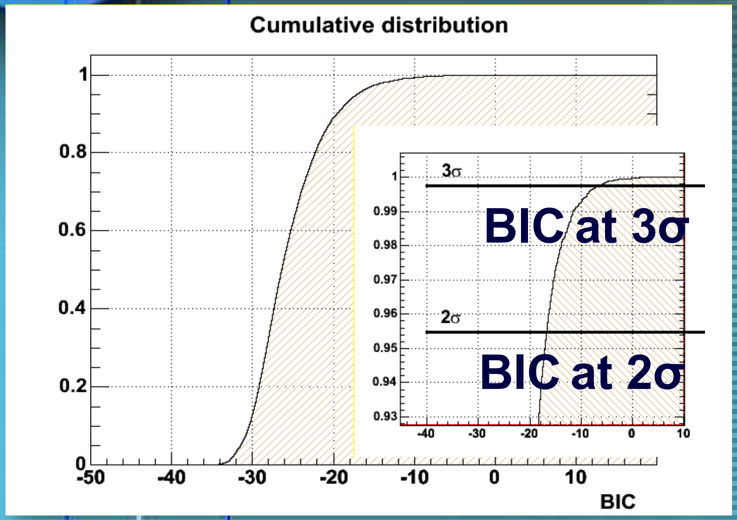
Obtenemos un valor del BIC por cada cluster candidato de la muestra

¿Converge?

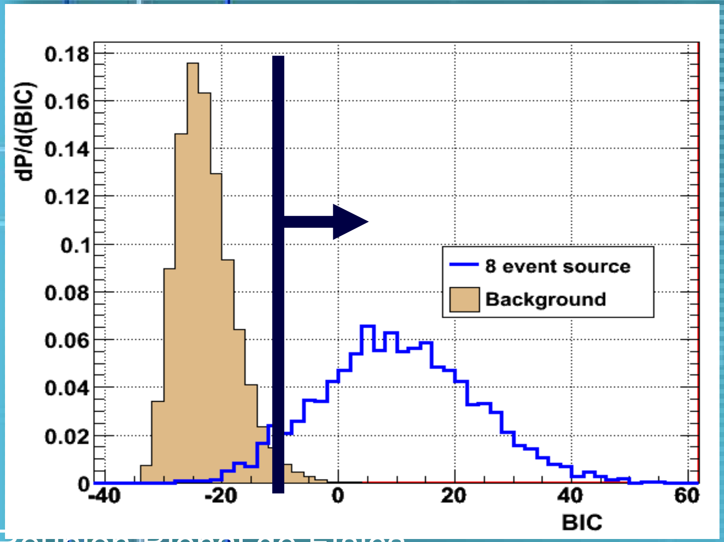
SI

NO

Probabilidad de descubrimiento



- Se calcula la función acumulativa de la distribución de BIC en el caso de sólo fondo
- Valores del BIC para distintos niveles de confianza.

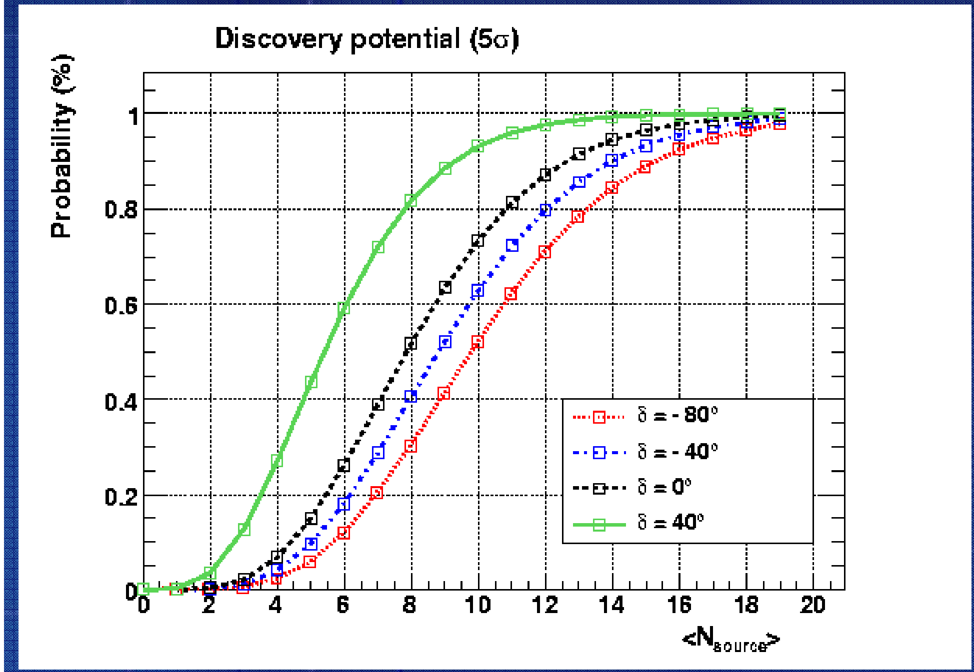


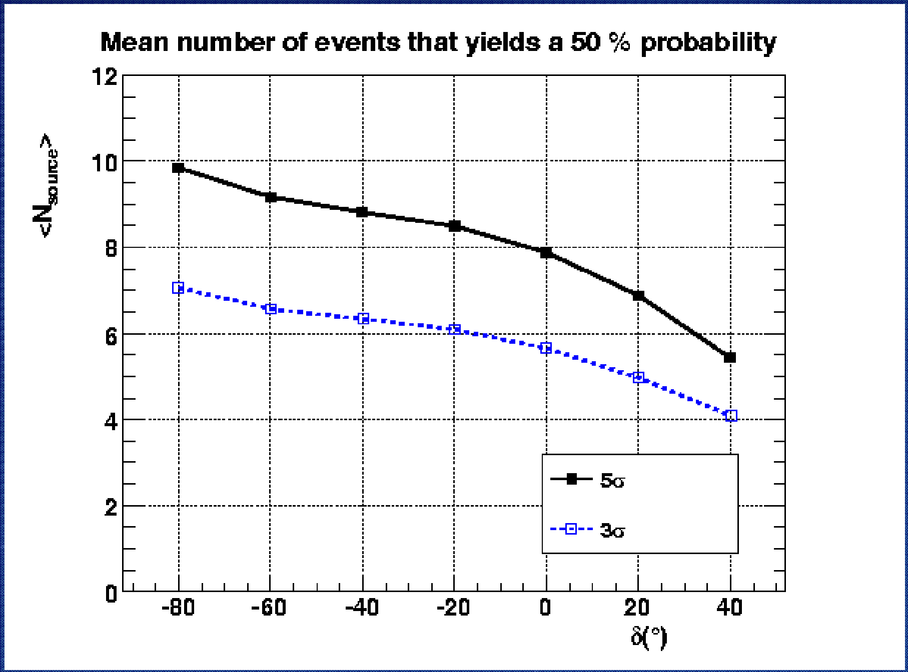
Definición de la probabilidad de descubrimiento es:

$$P(\mathrm{N}\sigma) = \frac{\int_{\mathrm{BIC}^{\mathrm{N}\sigma}}^{\infty} f_{src}(\mathrm{BIC})}{N_{\exp}}$$

Donde Nσ= 2σ, 3σ, 5σ

Probabilidad de descubrimiento: Resultados





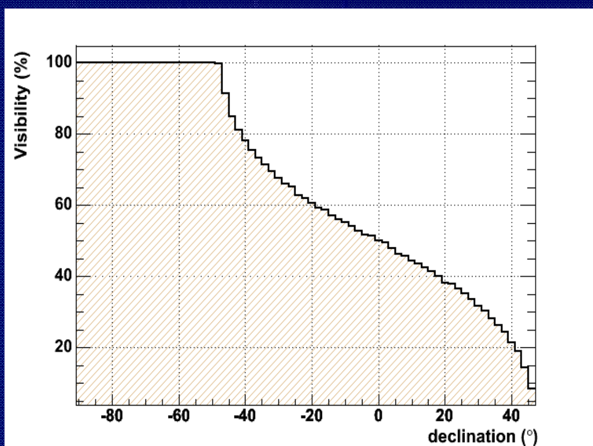
Probabilidad en función del número de sucesos para distintas declinaciones (5σ)

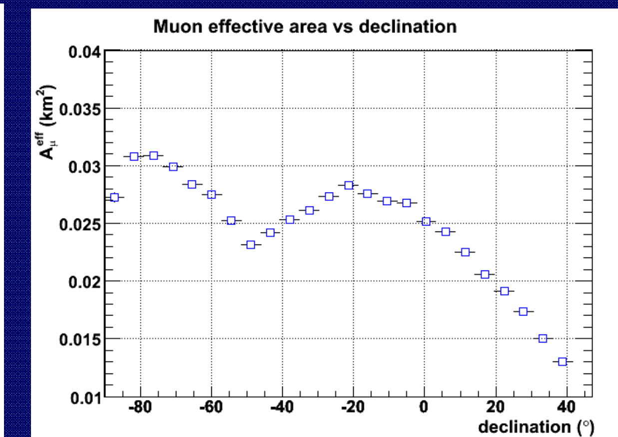
Número de sucesos emitidos por la fuente que dan lugar a una probabilidad del 50% (3σ y 5σ)

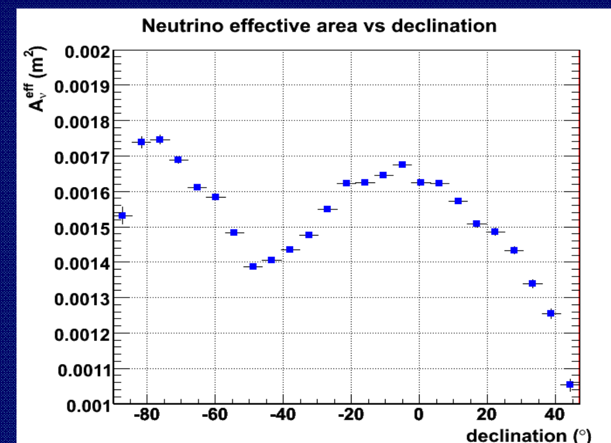
Flujo de descubrimiento al 50%

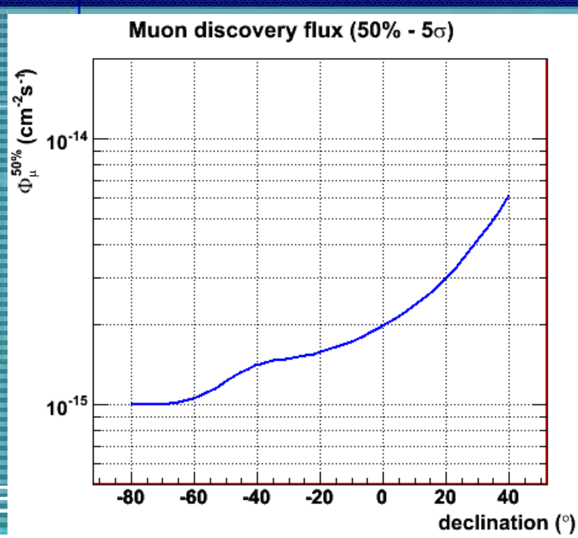
$$\Phi_i(\gamma,\delta) = \frac{N_{events}(\delta)}{A_i^{eff}(\gamma,\delta)V(\delta)T_{live}}$$

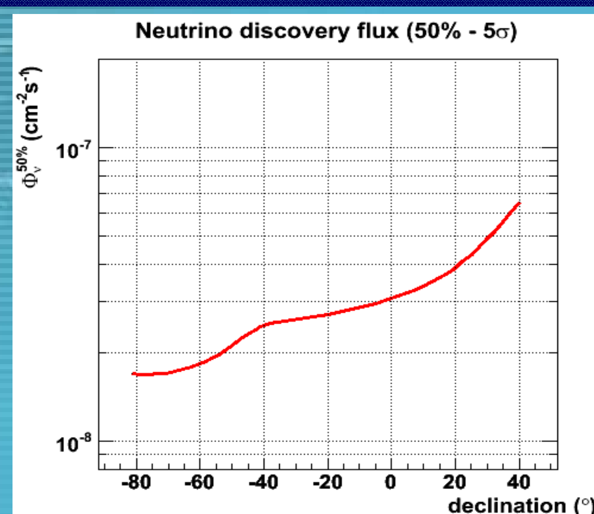
- $N_{events}(\delta)$ Número de sucesos al 50%
- $A^{eff}(\gamma,\delta)$ Area efectiva (i = μ, ν) para γ = 2
- $V(\delta)$ Visibilidad
- T_{live} Tiempo de adquisición del detector





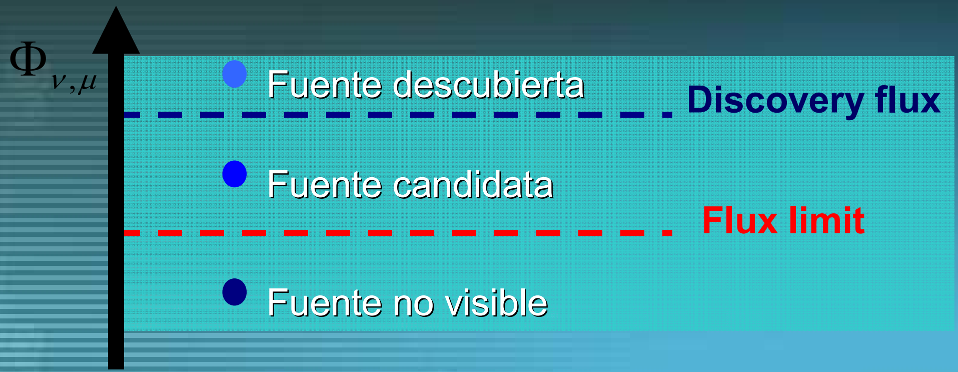


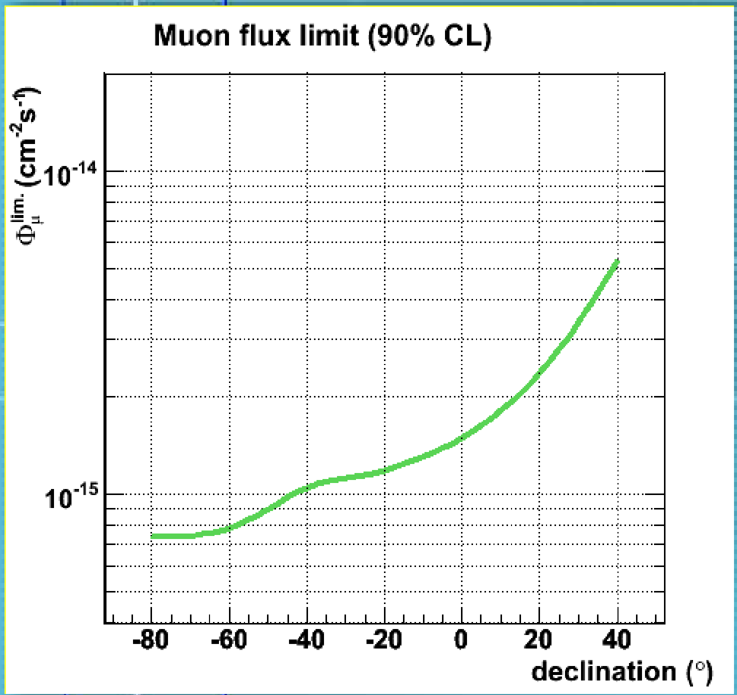


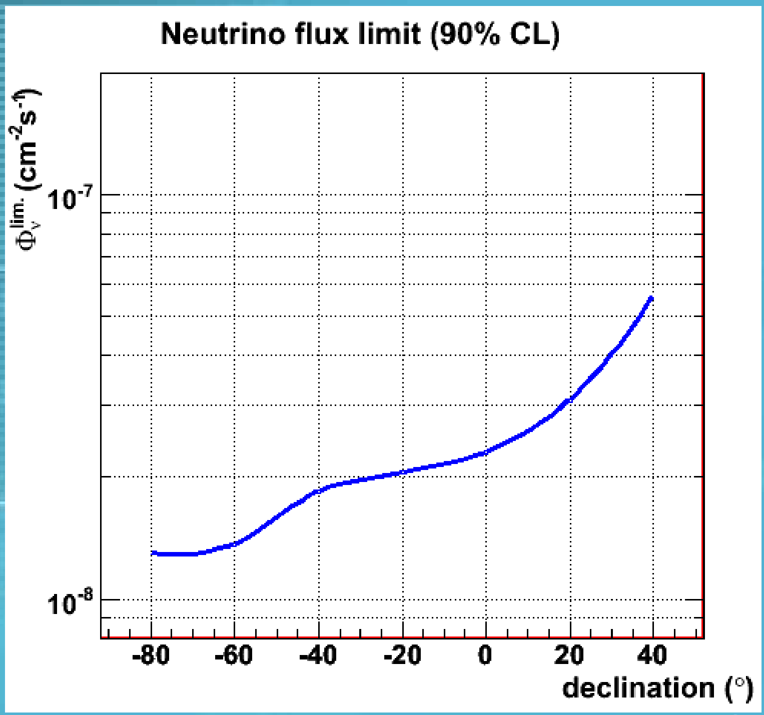


Sensibilidad (flux limit)

Si no hemos visto ninguna fuente, podemos ser capaces de dar un flujo por debajo del cual no podemos decir si existe o no una fuente







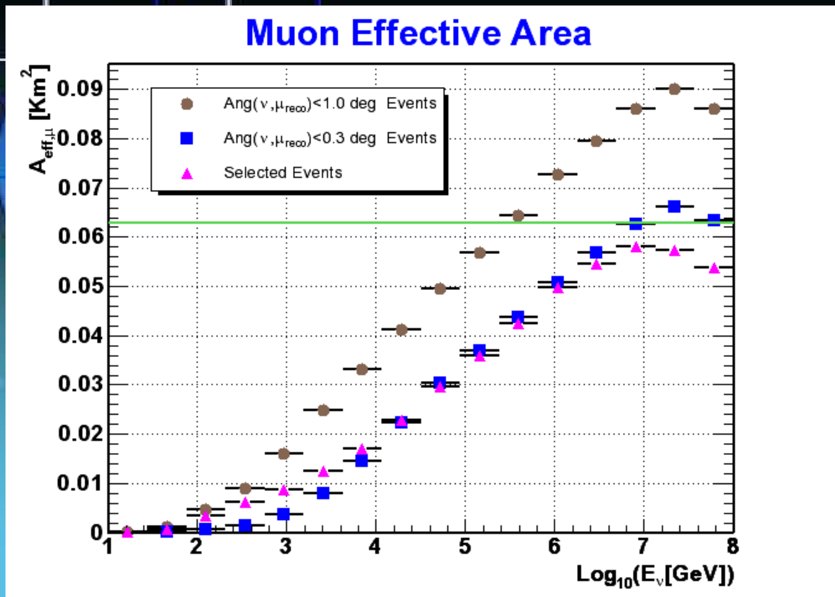
Conclusiones

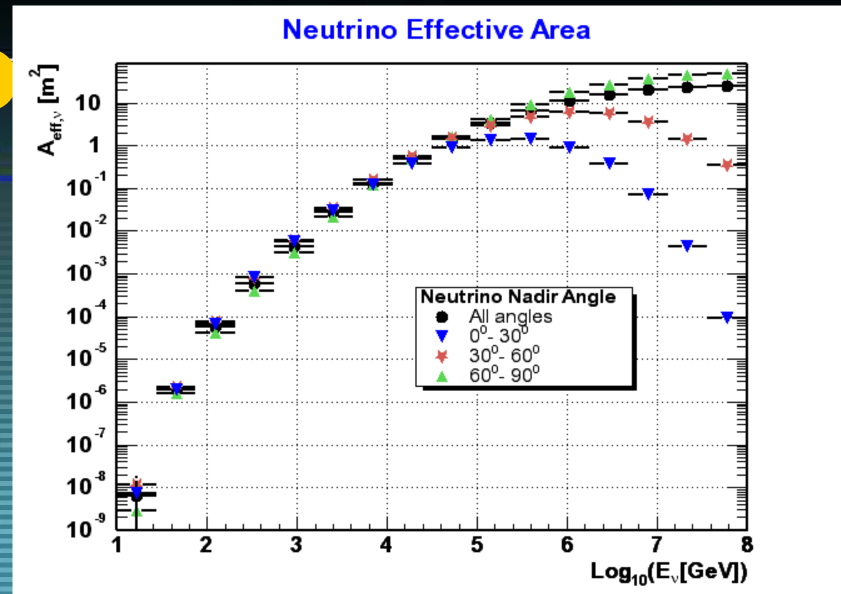
- Se ha estudiado el potencial de **ANTARES** para la búsqueda de fuentes puntuales usando para ello un algoritmo basado en la búsqueda de *clusters*.
- El método basado en el **algoritmo EM** presenta un mayor potencial de descubrimiento que los métodos basado en bines sin la necesidad de usar información estimada del Monte Carlo como otros métodos de búsqueda.
- El **flujo de descubrimiento** al 50% (5σ):

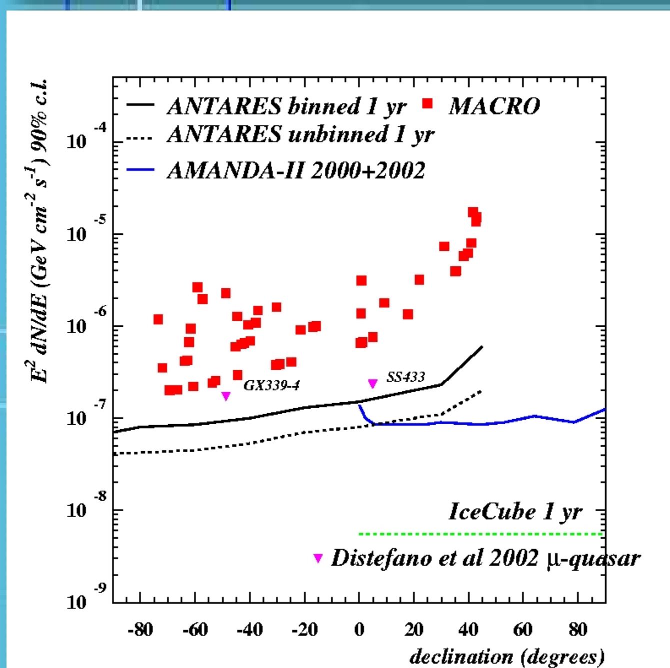
$$\Phi_{\mu}(\delta = 0) \approx 2 \cdot 10^{-15} \;\; cm^{-2} s^{-1} \quad \text{(1 año)}$$

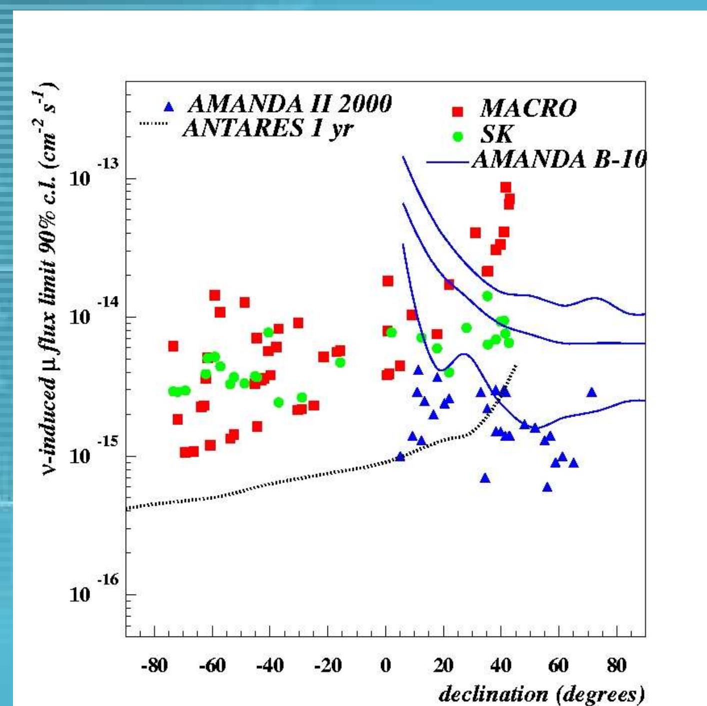
- La **sensibilidad** también se ha calculado en términos del flujo (90% CL):

$$\Phi_{\mu}(\delta = 0) \approx 1.5 \cdot 10^{-15} \;\; cm^{-2} s^{-1} \quad \text{(1 año)}$$


Muon Effective Area
A_eff,μ [Km²]
Ang(ν,μ_reco)<1.0 deg Events
Ang(ν,μ_reco)<0.3 deg Events
Selected Events
Log10(Eν[GeV])


Neutrino Effective Area
A_eff,ν [m²]
Neutrino Nadir Angle
All angles
0°- 30°
30°- 60°
60°- 90°
Log10(Eν[GeV])


ANTARES binned 1 yr
MACRO
ANTARES unbinned 1 yr
AMANDA-II 2000+2002
GX339-4
SS433
IceCube 1 yr
Distefano et al 2002 μ-quasar
E² dN/dE (GeV cm⁻² s⁻¹) 90% c.l.
declination (degrees)


AMANDA II 2000
ANTARES 1 yr
MACRO
SK
AMANDA B-10
ν-induced μ flux limit 90% c.l. (cm⁻² s⁻¹)
declination (degrees)

Selección del modelo (BIC)

Definición general:

1. Conjunto de datos D
2. Varios modelos $M_1, \ldots, M_k$

Teorema de Bayes: $$p(M_k \mid D) \propto p(D \mid M_k)\, p(M_k)$$

Probabilidad darse M_k dado D =

Probabilidad de M_k de reproducir D
X
Probabilidad a priori de M_k

Si $p(M_1) = \ldots = p(M_k)$ basta con $p(D|M_k)$ (factor de Bayes)

Integración sobre el espacio de parámetros desconocidos θ_k

$$p(D \mid M_k) = \int p(D \mid \theta_k, M_k)\, p(\theta_k \mid M_k)\, d\theta_k$$

Bayesian Information Criterion o BIC*:

g.d.l (parámetros a ajustar)

$$2\log[p(D \mid M_k)] \approx 2\log p(D \mid \hat{\theta}_k, M_k) - v_k \log(n) = \mathrm{BIC}_k$$

$$\mathrm{BIC}_k = 2\sum_{i=n}^{n} \log\left[p_{Gauss}(\hat{\alpha}_{RAi}, \hat{\delta}_i) + p_{BG}(\hat{\delta}_i)\right] - 2\sum_{i=n}^{n} \log\left[p_{BG}(\hat{\delta}_i)\right] - 6\log(n)$$

* G. Schwarz. *Estimating the dimension of a model.* Ann. of Statist. 6, 461-464 (1978)

Métodos de búsquedas

- Método de bines:

  Los métodos de bines buscan una acumulación de sucesos en una pequeña región del cielo (bin). La significancia se estima en comparación con la distribución de fondo.

- Método sin bines:

  1) Máxima verosimilitud. Maximiza la verosimilitud de una cierta función densidad de acuerdo a dos hipótesis. Una llamada nula (solo fondo) y la hipotesis de una fuente. Se escoge un parámetro (estadística) que es sensible al hecho de que los datos se distribuyan según una hipótesis u otra.