

« FDR » Point Sources search

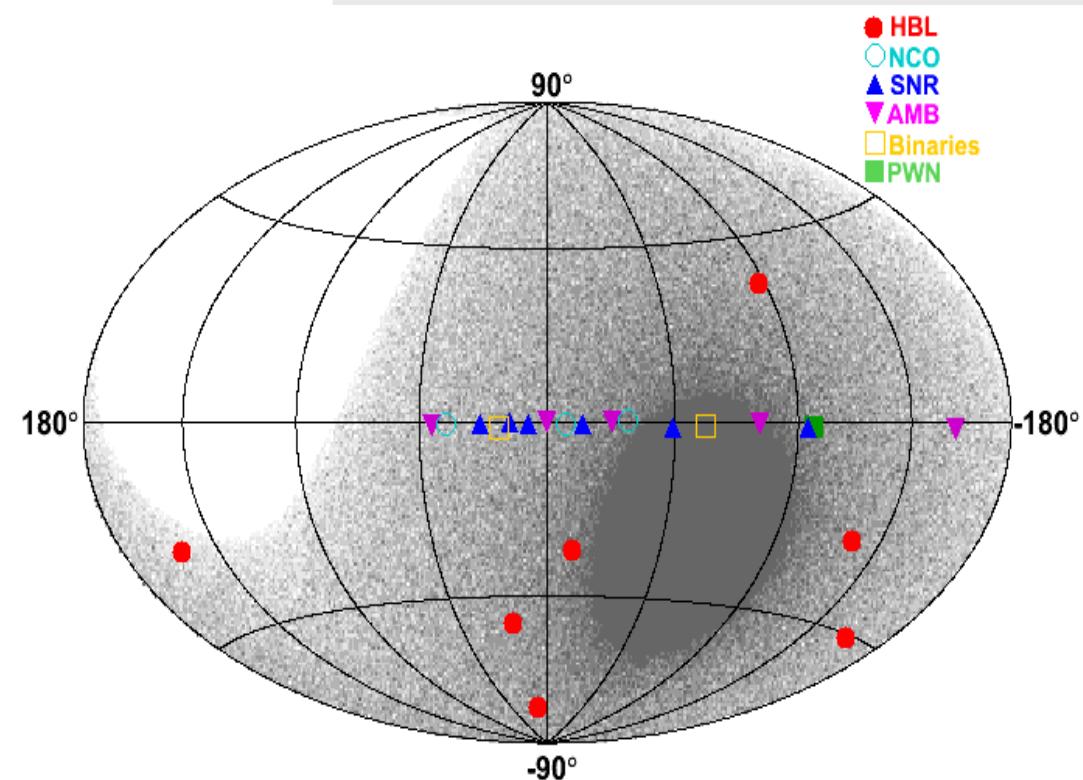
Reminder and update

Bruny Baret -APC



Sources ponctuelles: les suspects

Source	DECL	AR
PSR B1259-63	-63.8339	195.703
RCW 86	-62.4833	220.679
ESO 139-G12	-59.9414	264.414
HESS J1023-575	-57.7639	155.825
Cir X-1	-57.1667	230.171
HESS J1614-518	-51.82	243.579
PKS 2005-489	-48.8219	302.372
GX 339	-48.7897	255.704
RX J0852.0-4622	-46.3667	133
Centaurus A	-43.0191	201.364
RX J1713.7-3946	-39.75	258.25
PKS 0548-322	-32.2712	87.6692
H 2356-309	-30.6275	359.784
PKS 2155-304	-30.2217	329.721
Galactic Center	-29.0061	266.421
1ES 1101-232	-23.4919	165.909
W28	-23.335	270.425
LS 5039	-14.825	276.562
1ES 0347-121	-11.9908	57.3459
HESS J1837-069	-6.95	279.408
3C 279	-5.78917	194.046
RGB J0152+017	1.78861	28.1667
SS 433	4.98278	287.958
HESS J0632+057	5.80556	98.2416
IceCube HotSpot	11	153



2.Les neutrinos: Sources Ponctuelles – Principe de la recherche

« blind analysis »

Analyse optimisée sur bruit de fond déterminé
à partir des données «randomisées» en asc. droite.

Pixelisée:

- ▶ Optimiser un cône autour de la source pour la meilleure limite sup.
- ▶ compter n_{obs} avec n_b attendu

Significativité «locale»:
 $s = 1 - P_{\text{Poisson}}(n_{\text{obs}} | n_b)$

Non-pixelisée

- ▶ Algorithme de «clusterisation» autour de la source
- ▶ Maximisation du rapport de vraisemblance de Signal/Bruit L

Vraie significativité:

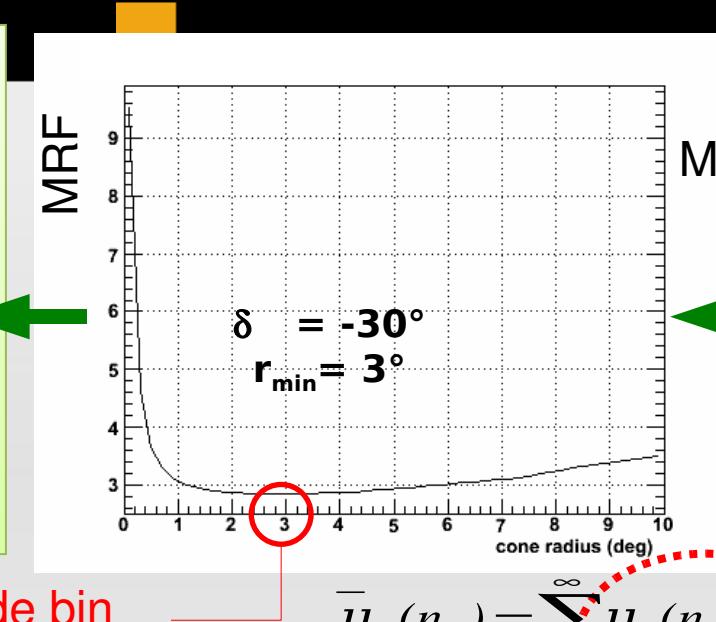
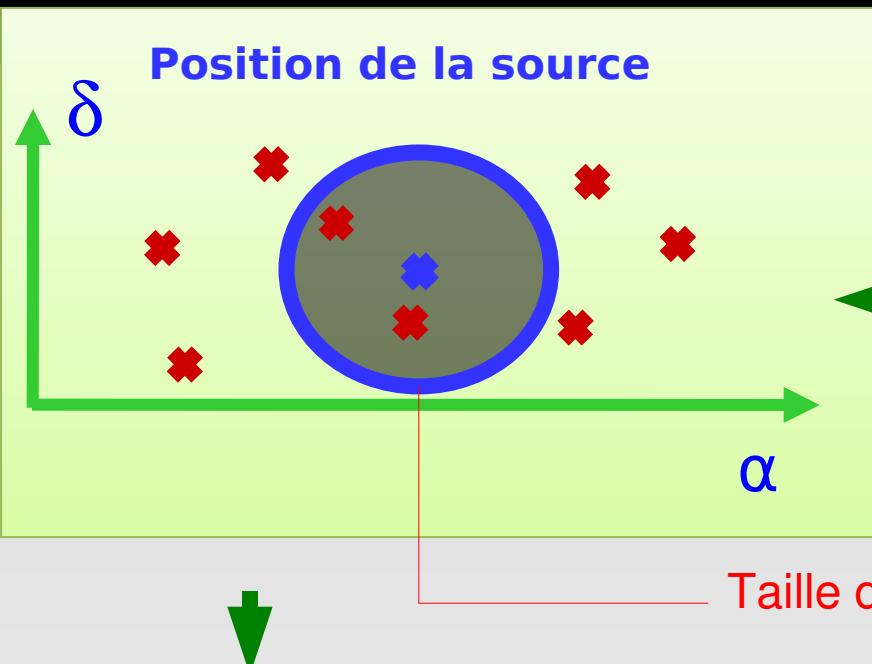
vérifiée sur $0(10^4)$ réalisation du ciel sans source
Probabilité d'avoir au moins s ou L

«unblinding»

Détection à n_{d}

Limites supérieures

Sources Ponctuelles – Principe de la recherche pixelisée



Model Rejection Factor

$$MRF = \frac{\bar{\mu}_{90}}{n_b}$$

$$\bar{\mu}_{90}(n_b) = \sum_{n_{obs}}^{\infty} \mu_{90}(n_{obs}, n_b) \frac{(n_b)^{n_{obs}}}{(n_{obs})!} \exp(-n_b)$$

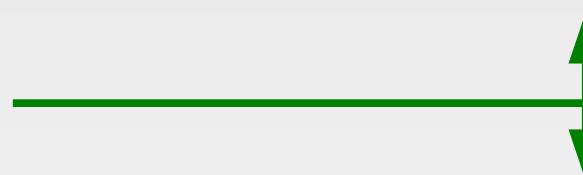
limite sup. à 90% Poids poissonnier

Significativité «locale»:
 $s = 1 - P_{\text{Poisson}}(n_{\text{obs}} | n_b)$

« blind analysis »

Détection à n_0

Vraie significativité:
 vérifiée sur $0(10^4)$ réalisation du ciel sans source
 Probabilité d'avoir au moins s



Limite supérieure

Non pixelisée

The EM method is a pattern recognition algorithm that **analytically** maximizes the likelihood in finite mixture problems, which are described by different density components (pdf).

$$p(\mathbf{x}) = \pi_{BG} P_{BG}(\delta) + \pi_s P_s(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma})$$

mixing proportions

position of event: $\mathbf{x} = (\alpha_{RA}, \delta)$

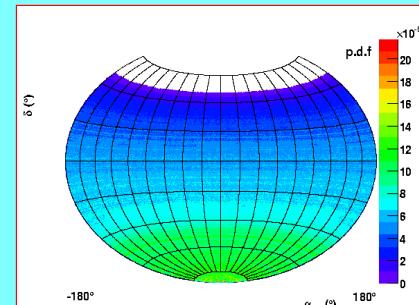
bg: only δ

signal: α_{RA}, δ

Signal pdf model is selected to be 2D-

Gaussians

$$P_s(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{2\pi \sigma_\alpha \sigma_\delta} \exp\left(-\frac{(\alpha - \mu_\alpha)^2 \cos^2 \delta}{2\sigma_\alpha^2}\right) \exp\left(-\frac{(\delta - \mu_\delta)^2}{2\sigma_\delta^2}\right) \cos \delta$$



The background pdf is extracted from MC or

Flowchart of the EM-based method

Initial values for the

signal pdf

- $\boldsymbol{\mu}$ source coordinates
(α, δ)

- σ det. angular resolution

- π_s cluster elements

EM algorithm

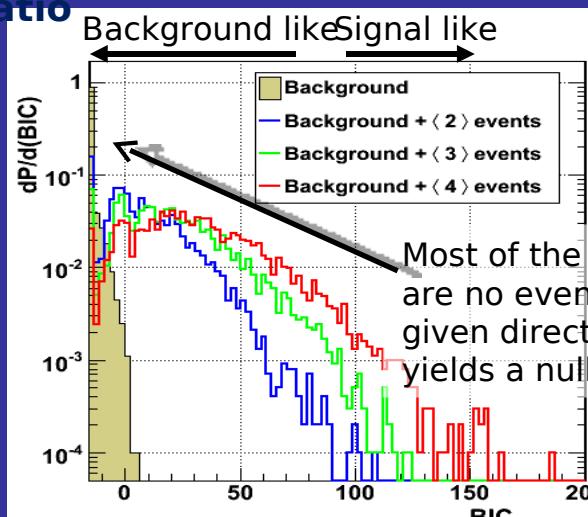
Final pdf parameters that maximize the likelihood



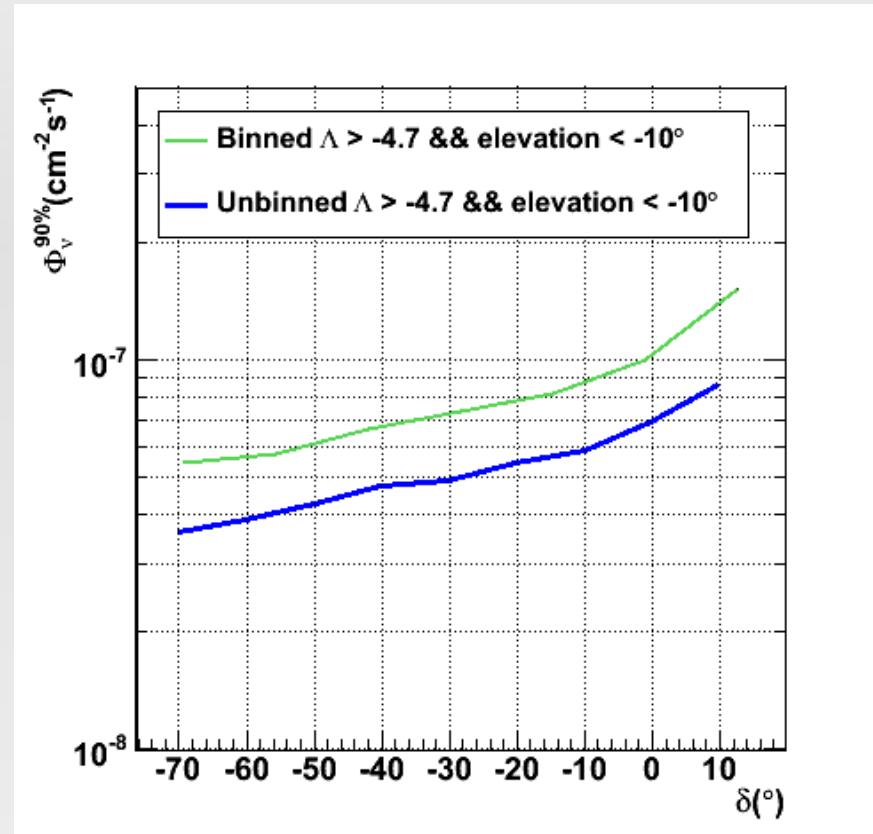
$$\text{BIC} = 2 \log p(D | \Psi_1^{(m)}, M_1) - 2 \log p(D | M_0) + v_1 \log(n)$$

penalty

Likelihood ratio



Comparaison



2.Les neutrinos: Sources ponctuelles - Limites

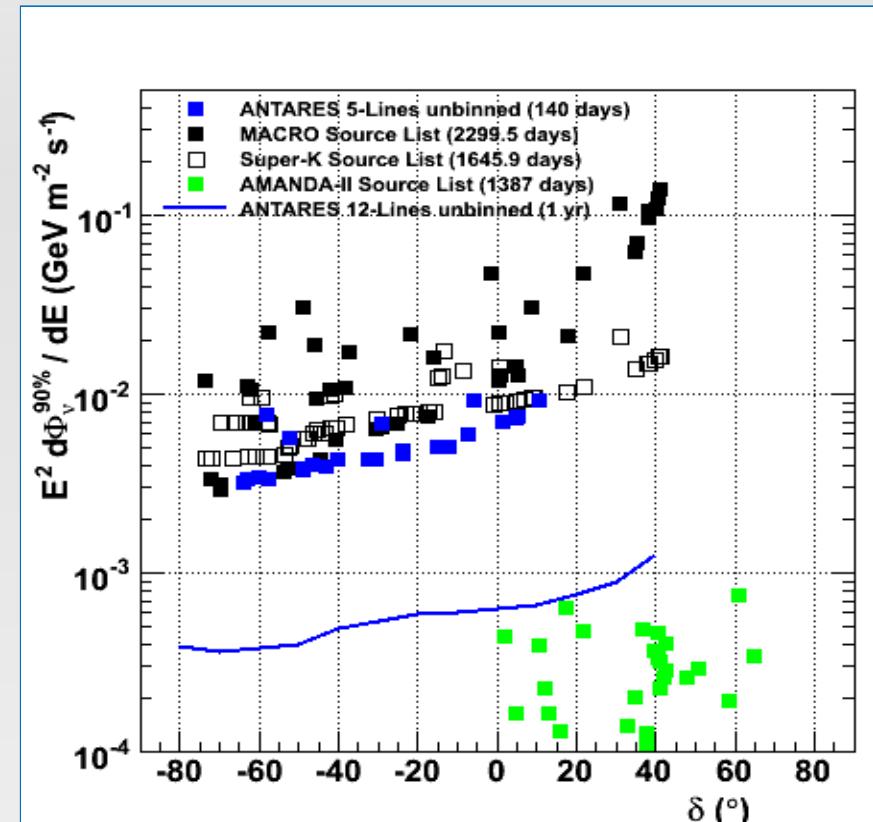
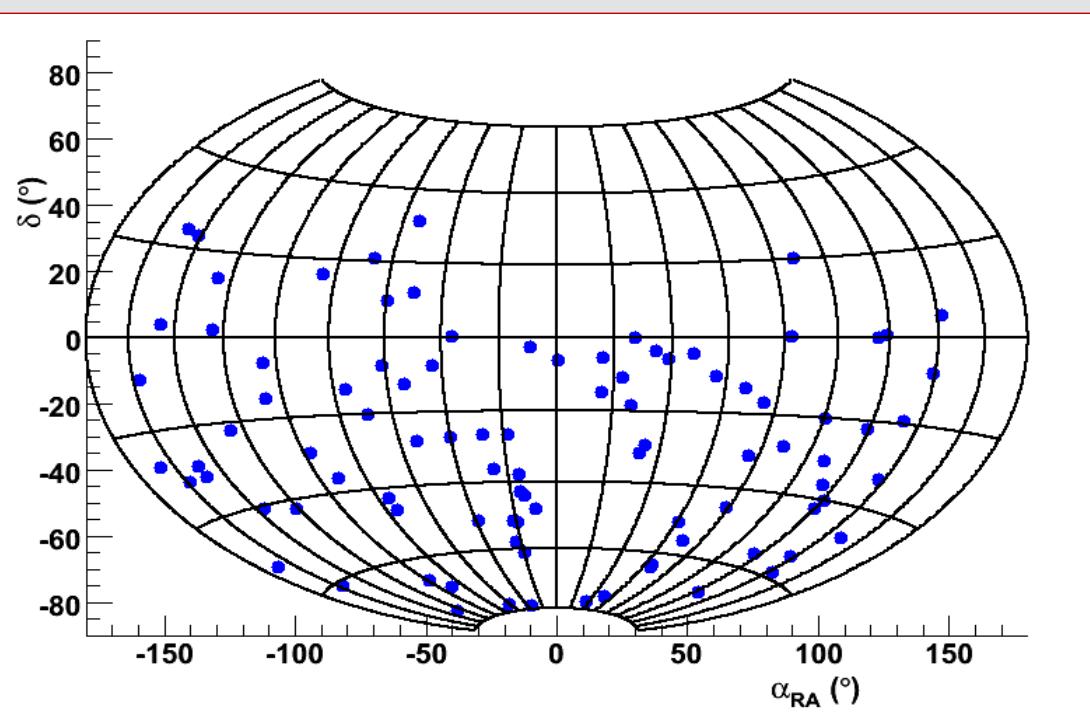


esé

- ▶ 94 évts sélectionnés à haute énergie sur données 5 lignes
- ▶ meilleure limite dans le ciel nord (140jrs!)

5-Line sky map

Carte du ciel



Recherche en aveugle sur tout le ciel en cours

The Idea

What it does:

Blind sky survey

Maximise Detection Potential
while controlling C.L. of
detections

Use multiple hypothesis testing
procedure in a model independent
way

What it does NOT:

Optimise for best upper limit
Use a source flux model
(Usual method:
likelihood ratio maximisation)
Use MC to check significance
a posteriori

Developped for Amanda/Ice3 with M.Labare

Hypothesis testing basics

- Random variable X , Confidence level κ , « null » hypothesis H_0 with pdf $f(x)$

- Test H_0 on each realisation x_i of X :

► compute the p-value(i) = $\int_{x_i}^{\infty} f(x) dx$

► If $p\text{-value}(i) < 1 - \kappa \Rightarrow H_0$ rejected

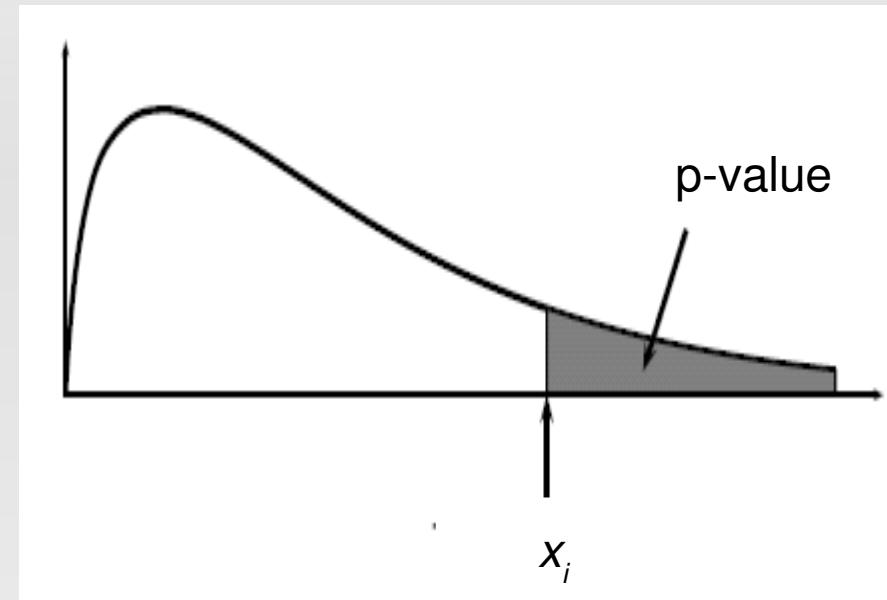
Problem:

If $N > 1$ tests, the confidence level decreases

« Trial factor effect »

Conservative «Bonferroni» Solution:

$$1 - \kappa \rightarrow (1 - \kappa)/N$$



New problem:

Detection power decreases

«FDR» controlling procedure

FDR = False Discovery Rate = $\frac{\# \text{ false discoveries}}{\# \text{ discoveries}}$

Principle:

- Choose a desired FDR α (C.L.=1-FDR)
- Adaptative rejection threshold guarantees $\text{FDR} \leq \alpha$

Advantages:

- meant for « sparse heterogenous mixtures »
- Automated and control of the final confidence level
- Combines power and conservatism

Applications in astro./cosmo.:

- Baryonic spectrum fluctuations
- Galaxy clusters (SDSS)
- SZ clusters with WMAP

a lot of x_i under H_0
very few sources

Mathematical proof

4. Proof of theorem. For ease of exposition let us denote the set of constants in (1), which define the procedure, by

$$(4) \quad q_i = \frac{i}{m}q, \quad i = 1, 2, \dots, m.$$

Let $A_{v,s}$ denote the event that the Benjamini Hochberg procedure rejects exactly v true and s false null hypotheses. The FDR is then

$$(5) \quad E(\mathbf{Q}) = \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{v}{v+s} \Pr(A_{v,s}).$$

In the following lemma, $\Pr(A_{v,s})$ is expressed as an average.

LEMMA 4.1.

$$(6) \quad \Pr(A_{v,s}) = \frac{1}{v} \sum_{i=1}^{m_0} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}).$$

PROOF. For a fixed v and s , let ω denote a subset of $\{1 \dots m_0\}$ of size v , and $A_{v,s}^\omega$ the event in $A_{v,s}$ that the v true null hypotheses rejected are ω . Note that $\Pr\{P_i \leq q_{v+s} \cap A_{v,s}^\omega\}$ equals $\Pr\{A_{v,s}^\omega\}$ if $i \in \omega$, and is otherwise 0.

$$(7) \quad \begin{aligned} \sum_{i=1}^{m_0} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}) &= \sum_{i=1}^{m_0} \sum_{\omega} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}^\omega) \\ &= \sum_{\omega} \sum_{i=1}^{m_0} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}^\omega) \\ &= \sum_{\omega} \sum_{i=1}^{m_0} I(i \in \omega) \Pr\{A_{v,s}^\omega\} \\ &= \sum_{\omega} v \cdot \Pr\{A_{v,s}^\omega\} = v \cdot \Pr\{A_{v,s}\}. \end{aligned}$$

Combining equation (5) with Lemma 4.1, the FDR is

$$(8) \quad \begin{aligned} E(\mathbf{Q}) &= \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{v}{v+s} \left\{ \sum_{i=0}^{m_0} \frac{1}{v} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}) \right\} \\ &= \sum_{i=0}^{m_0} \left\{ \sum_{s=0}^{m_1} \sum_{v=1}^{m_0} \frac{1}{v+s} \Pr(\{P_i \leq q_{v+s}\} \cap A_{v,s}) \right\} \end{aligned}$$

Now that the dependency of the expectation on v is only through $A_{v,s}$, we reconstruct $A_{v,s}$ from events that depend on i and $k = v+s$ only, so the FDR may be expressed similarly.

For $i = 1 \dots m_0$, let $\mathbf{P}^{(i)}$ be the remaining $m - 1$ p -values after dropping P_i . Let $C_{v,s}^{(i)}$ denote the event in which if P_i is rejected then $v - 1$ true null hypotheses and s false null hypotheses are rejected alongside with it. That is, $C_{v,s}^{(i)}$ is the projection of $\{P_i \leq q_{v+s}\} \cap A_{v,s}$ onto the range of $\mathbf{P}^{(i)}$, and expanded again by cross multiplying with the range of P_i . Thus we have

$$(9) \quad \{P_i \leq q_{v+s}\} \cap A_{v,s} = \{P_i \leq q_{v+s}\} \cap C_{v,s}^{(i)}.$$

Denote by $C_k^{(i)} = \bigcup\{C_{v,s}^{(i)}: v + s = k\}$. For each i the $C_k^{(i)}$ are disjoint, so the FDR can be expressed as

$$(10) \quad E(\mathbf{Q}) = \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr(P_i \leq q_k \cap C_k^{(i)}),$$

where the expression no longer depends on v and s , as desired.

In the last part of the proof we construct an expanding series of increasing sets, on which we use the PRDS property to bound the inner sum in (8) by q/m . For this purpose, define $D_k^{(i)} = \bigcup\{C_j^{(i)}: j \leq k\}$ for $k = 1 \dots m$. $D_k^{(i)}$ can also be described using the ordered set of the p -values in the range of $\mathbf{P}^{(i)}$, $\{p_{(1)}^{(i)} \leq \dots \leq p_{(m-1)}^{(i)}\}$, in the following way:

$$(11) \quad D_k = \{p: q_{k+1} < p_{(k)}^{(i)}, q_{k+2} < p_{(k+1)}^{(i)} \dots q_m < p_{(m-1)}^{(i)}\}$$

for $k = 1 \dots m - 1$, and $D_m^{(i)}$ is simply the entire space. Expressing $D_k^{(i)}$ as above, it becomes clear that for each k , $D_k^{(i)}$ is a nondecreasing set.

We now shall make use of the PRDS property, which states that for $p \leq p'$,

$$(12) \quad \Pr(D | P_i = p) \leq \Pr(D | P_i = p').$$

Following Lehmann (1996), it is easy to see that for $j \leq l$ since $q_j \leq q_l$,

$$(13) \quad \Pr(D | P_i \leq q_j) \leq \Pr(D | P_i \leq q_l),$$

for any nondecreasing set D , or equivalently,

$$(14) \quad \frac{\Pr(\{P_i \leq q_k\} \cap D_k^{(i)})}{\Pr(P_i \leq q_k)} \leq \frac{\Pr(\{P_i \leq q_{k+1}\} \cap D_k^{(i)})}{\Pr(P_i \leq q_{k+1})}.$$

Invoking (14) together with the fact that $D_{j+1}^{(i)} = D_j^{(i)} \cup C_{j+1}^{(i)}$ yields for all $k \leq m - 1$,

$$(15) \quad \begin{aligned} &\frac{\Pr(\{P_i \leq q_k\} \cap D_k^{(i)})}{\Pr(P_i \leq q_k)} + \frac{\Pr(\{P_i \leq q_{k+1}\} \cap C_{k+1}^{(i)})}{\Pr(P_i \leq q_{k+1})} \\ &\leq \frac{\Pr(\{P_i \leq q_{k+1}\} \cap D_k^{(i)})}{\Pr(P_i \leq q_{k+1})} + \frac{\Pr(\{P_i \leq q_{k+1}\} \cap C_{k+1}^{(i)})}{\Pr(P_i \leq q_{k+1})} \\ &= \frac{\Pr(\{P_i \leq q_{k+1}\} \cap D_{k+1}^{(i)})}{\Pr(P_i \leq q_{k+1})}. \end{aligned}$$

Now, start by noting that $C_1 = D_1$, and repeatedly use the above inequality for $i = 1, \dots, m - 1$, to fold the sum on the left into a single expression,

$$(16) \quad \sum_{k=1}^m \frac{\Pr(\{P_i \leq q_k\} \cap C_k^{(i)})}{\Pr(P_i \leq q_k)} \leq \frac{\Pr(\{P_i \leq q_m\} \cap D_m^{(i)})}{\Pr(P_i \leq q_m)} = 1,$$

where the last equality follows because $D_m^{(i)}$ is the entire space.

Going back to expression (10) for the FDR,

$$(17) \quad \begin{aligned} E(\mathbf{Q}) &= \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{1}{k} \Pr(\{P_i \leq q_k\} \cap C_k^{(i)}) \\ &\leq \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{q}{m} \cdot \frac{\Pr(\{P_i \leq q_k\} \cap C_k^{(i)})}{\Pr(P_i \leq q_k)}, \end{aligned}$$

OK, bad joke

because $\Pr(P_i \leq q_k) \leq q_k = \frac{k}{m}q$ under the null hypothesis (with equality for continuous test statistics where each P_i is uniform), so finally, invoking (16),

$$(18) \quad \frac{q}{m} \sum_{i=1}^{m_0} \sum_{k=1}^m \frac{\Pr(\{P_i \leq q_k\} \cap C_k^{(i)})}{\Pr(P_i \leq q_k)} \leq \frac{m_0}{m} q.$$

Relax and have a look at:

Y. Benjamini & Y. Hochberg, *J.R. Stat. Soc., B* (1995) 57

Y. Benjamini & Yekutieli The Annals of Statistics 2001, Vol. 29, No. 4, 1165–1188

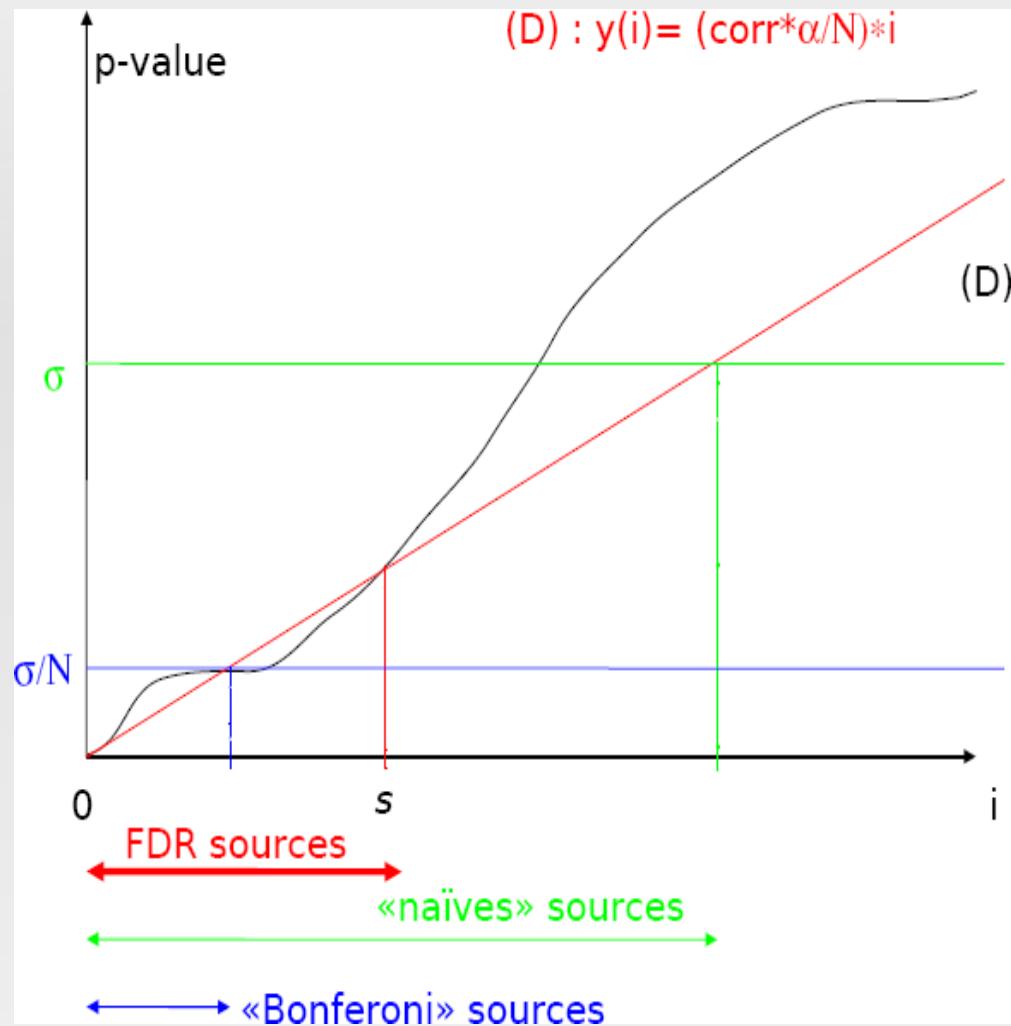
Controlling the False Discovery Rate in Astrophysical Data Analysis

C. J. Miller et al. astro-ph/0107034v1

↗ Begin with this

“False discovery rate controlling procedure for astrophysical high energy neutrino point source search” (or something like that) to be submitted soon (I promise!)

FALSE DISCOVERY RATE : BASICS



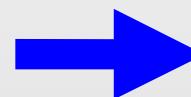
- Set of N pvalues (probability to be bg v)
→ sort increasing order
- Find the last crossing (s) with the line (D)
- Reject H_0 for $i < s$ with a Confidence Level $1-\alpha$
- If H_0 is rejected
→ we have a signal

corr : coefficient accounting for data correlations

Tests on MC simulation

Generate 10^3 final level skies with:

- ~3000 atmospherical ν
- 1 source with 1 to 20 ν and E^{-2} spectrum



Detection power
&
control of C.L.

MC files from K. Fratini PS analysis:

- Bari/Valencia production reprocessed for 5 lines detector in Bologna
`/in2p3/mc/neutrino/mu/prod0401/nu_low/l05_c09_s00/tab_nim/`
- $7 \times 5 \times 10^{10} \nu$ events
- Energy spectrum ($E\nu$) -1.4
 $10 \text{ GeV} < E\nu < 100 \text{ TeV}$ low energy neutrino production
 $0 < \cos\theta\nu < 1$ upgoing neutrinos
- Noise from run 25920 (golden list run)
- Nim table for the angular acceptance
- Trigger Trig3D_5L1
- Aart Strategy

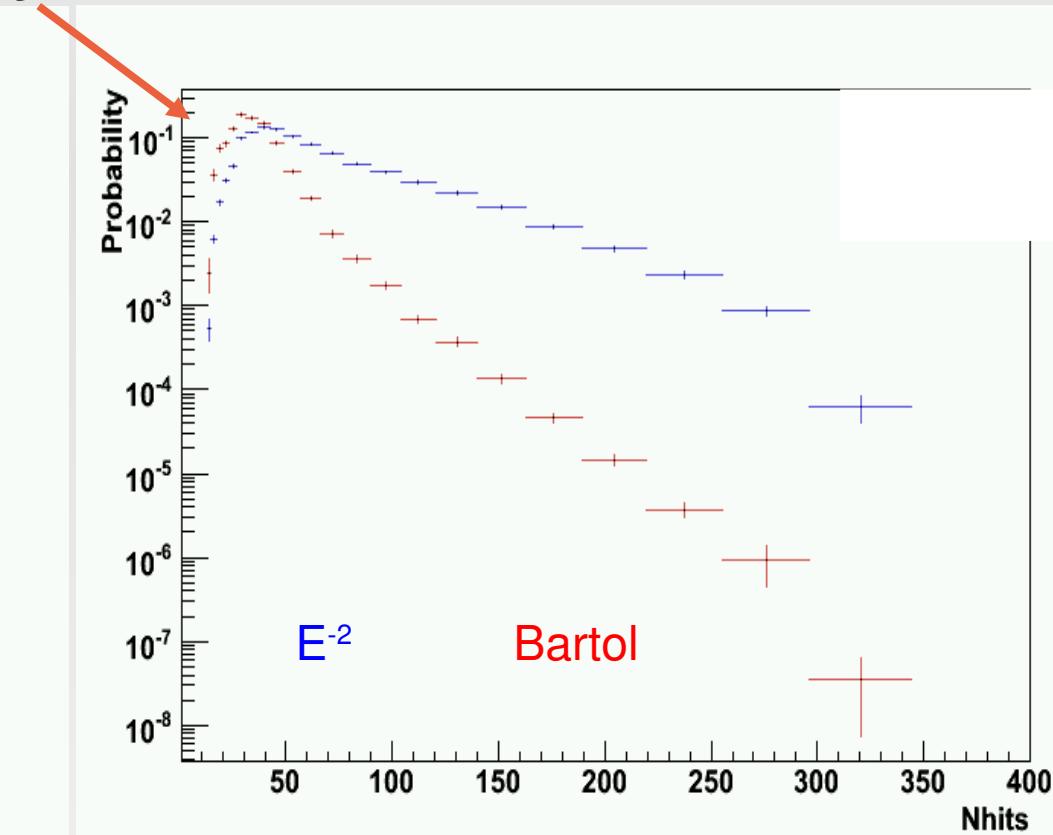
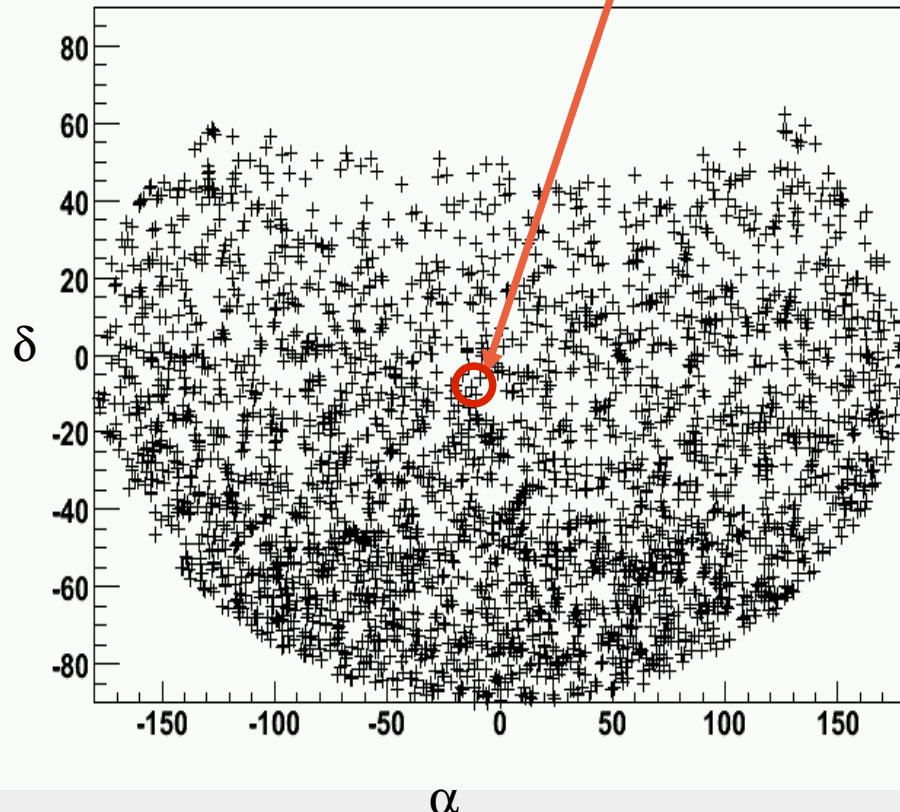
$\Lambda > -5$

FDR PROCEDURE FOR POINT SOURCE SEARCH

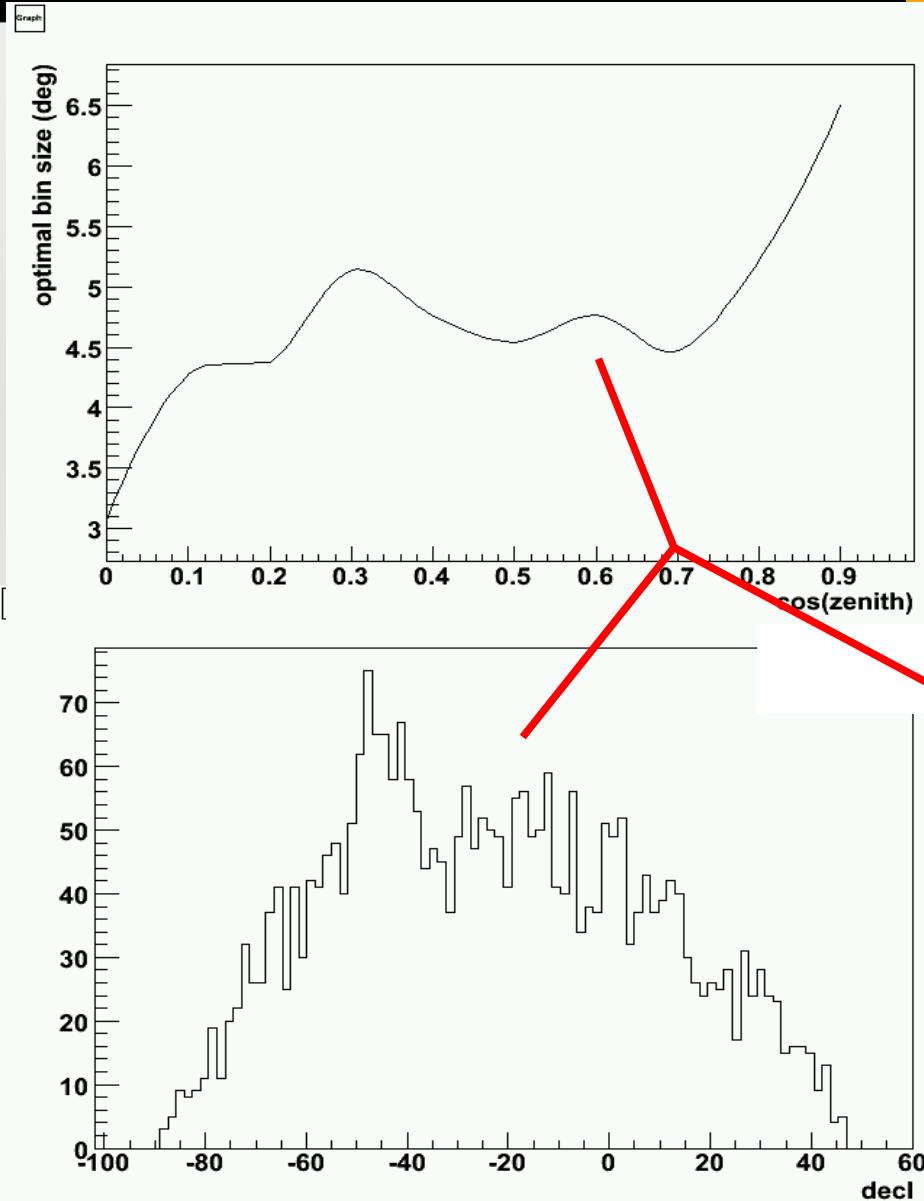
Final set of heatmaps : 0000 v → 0000 p values → 0000 tests

depending on :

- v density (position) in neighborhood
- energy (Number of hits)



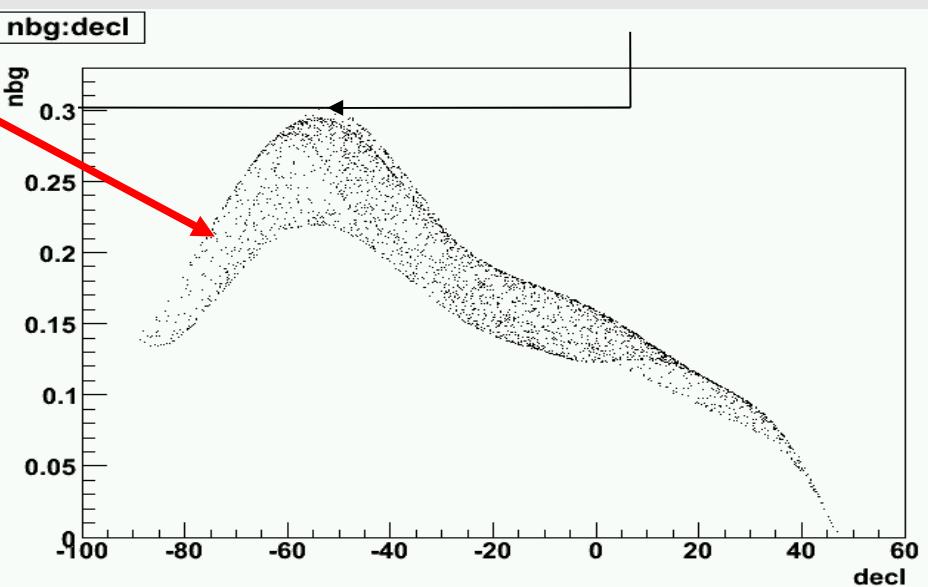
BIN SIZE DETERMINATION



Request

- bin size $\geq R_{\text{opt}} = 1.6 \times \text{resolution}$
(optimal S/B)
- same expected bgd. everywhere N_{BG}

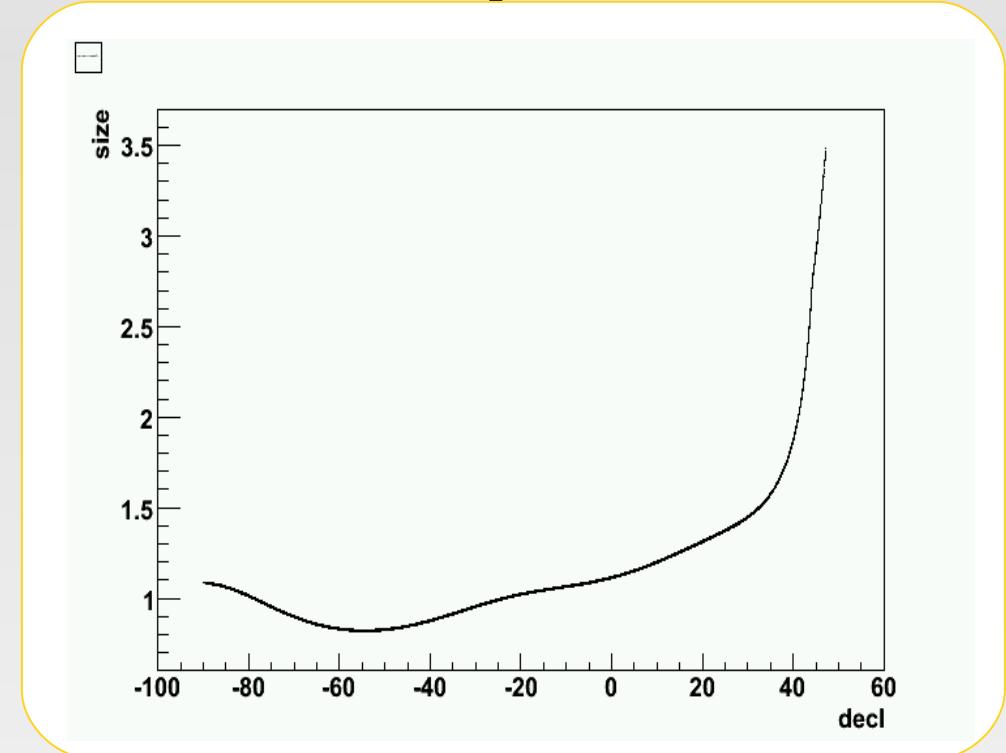
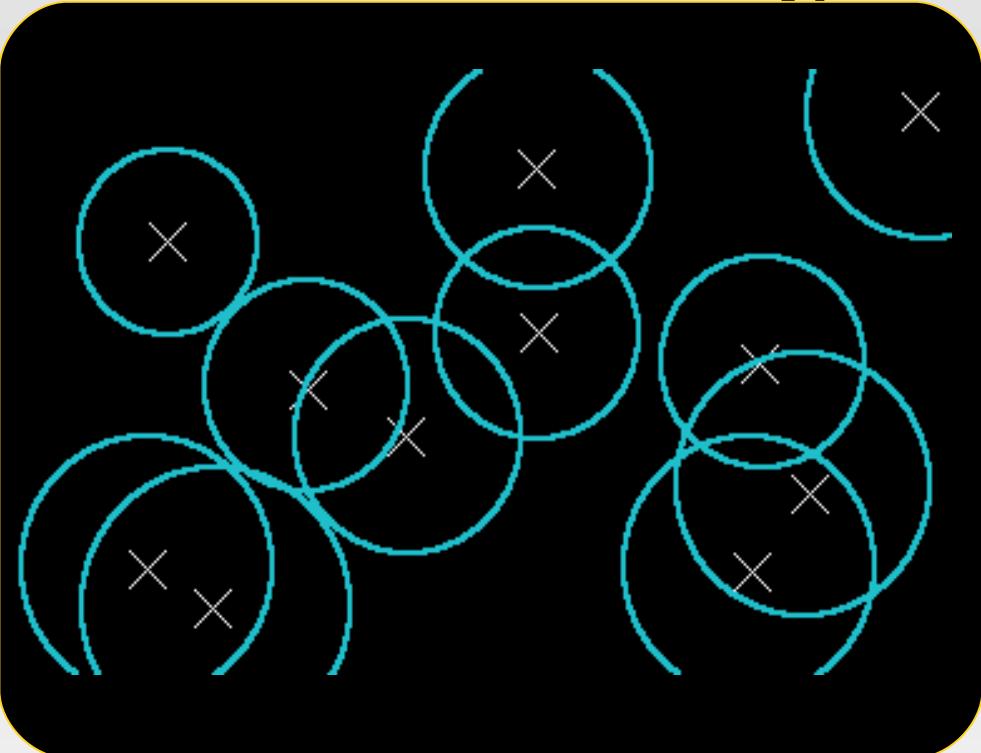
$$N_{\text{BG}} = \text{Max}(\text{expected bgd in } R_{\text{opt}})$$



FDR PROCEDURE FOR POINT SOURCE SEARCH

P-value computation

1. Definition of search bin around the neutrino, which the size (>resolution) depends on declination.
→ Same Null Hypothesis all over the sky

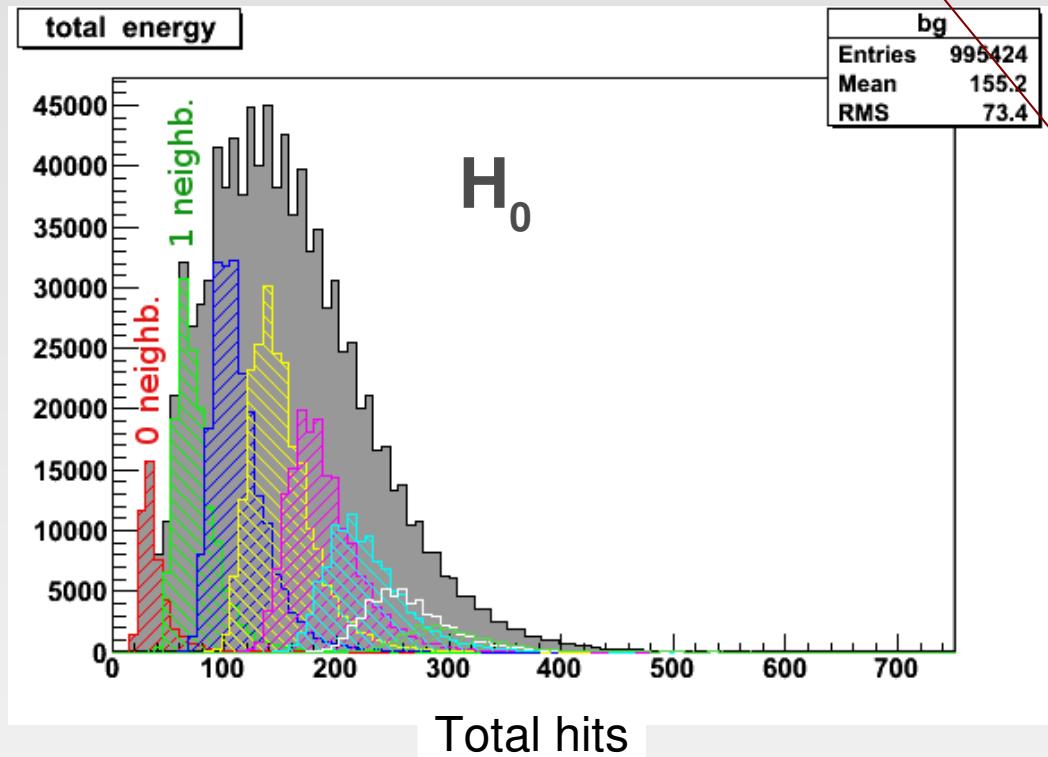


FDR PROCEDURE FOR POINT SOURCE SEARCH

P-value Computation

2. Definition of variable that takes into account both
density and **energy** criteria → **TOTAL ENERGY (Nhits**

$$P(i, n_{bg}) = \frac{e^{-n_{bg}} n_{bg}^i}{i!}$$



$$E_{TOT}(x) = \sum_{i=0}^{\infty} P(i, n_{bg}) E_i(x)$$

with :

$E_0(x)$ fit from data

$$E_1(x) = \sum_0^a E_0(x) E_0(x-a)$$

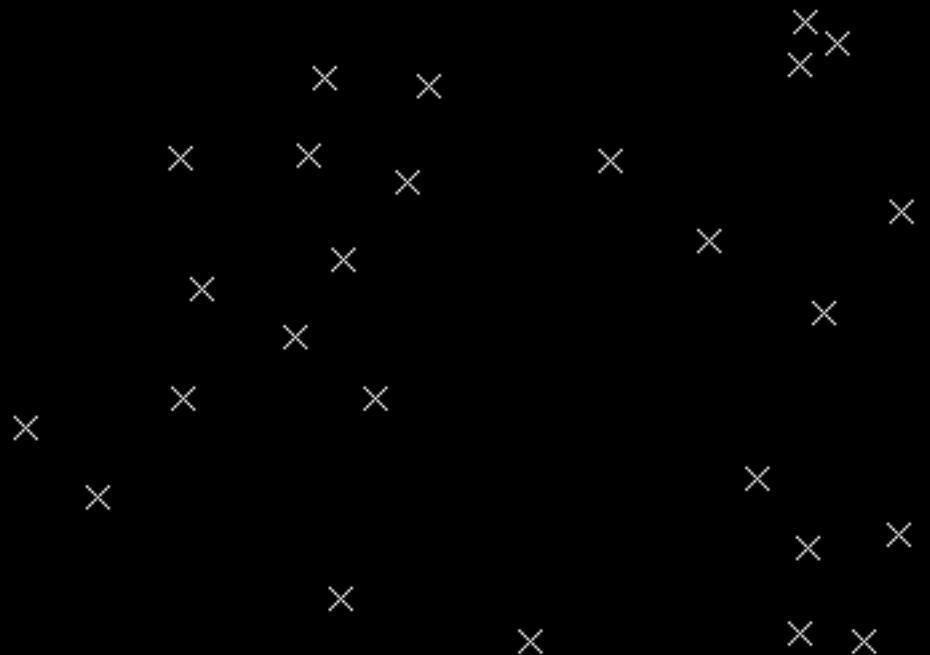
⋮

$$E_n(x) = \sum_0^a E_0(x) E_{n-1}(x-a)$$

FDR PROCEDURE FOR POINT SOURCE SEARCH

Hypothesis tested for each v

Set of p-values



FDR PROCEDURE FOR POINT SOURCE SEARCH

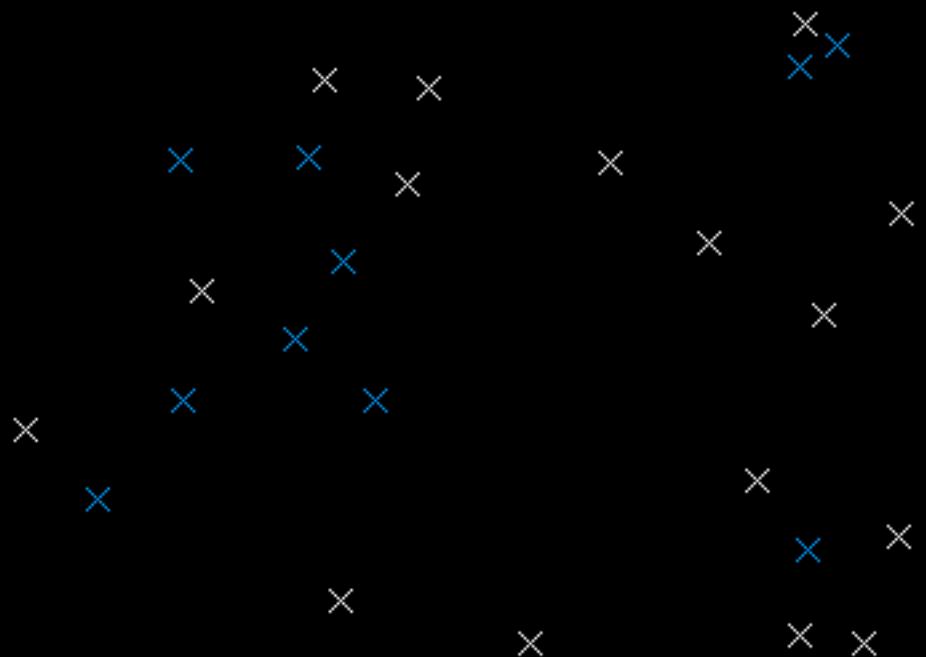
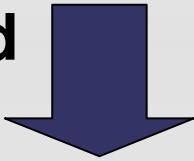
Hypothesis tested for each v

Set of p-values



FDR Procedure

Set of rejected
p-values



FDR PROCEDURE FOR POINT SOURCE SEARCH

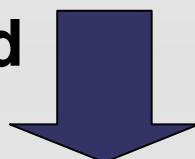
Hypothesis tested for each v

Set of p-values



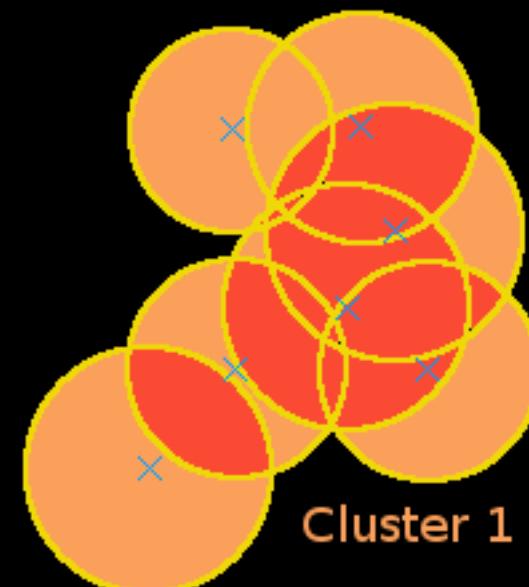
FDR Procedure

Set of rejected
p-values



Hierarchical Clustering

Set of clusters



FDR PROCEDURE FOR POINT SOURCE SEARCH

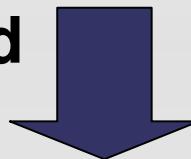
Hypothesis tested for each v

Set of p-values



FDR Procedure

Set of rejected
p-values

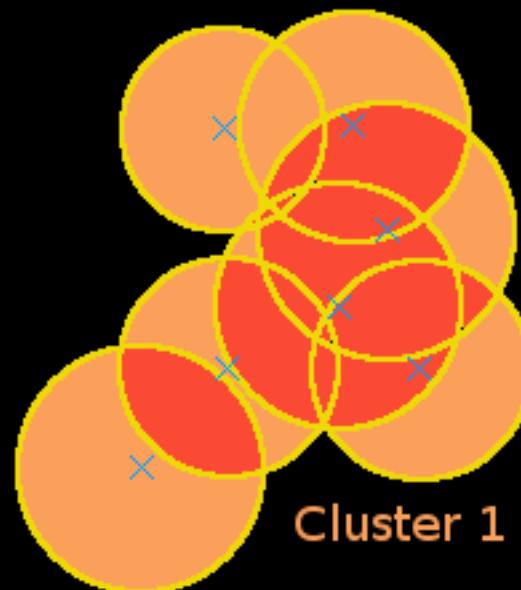


Hierarchical Clustering

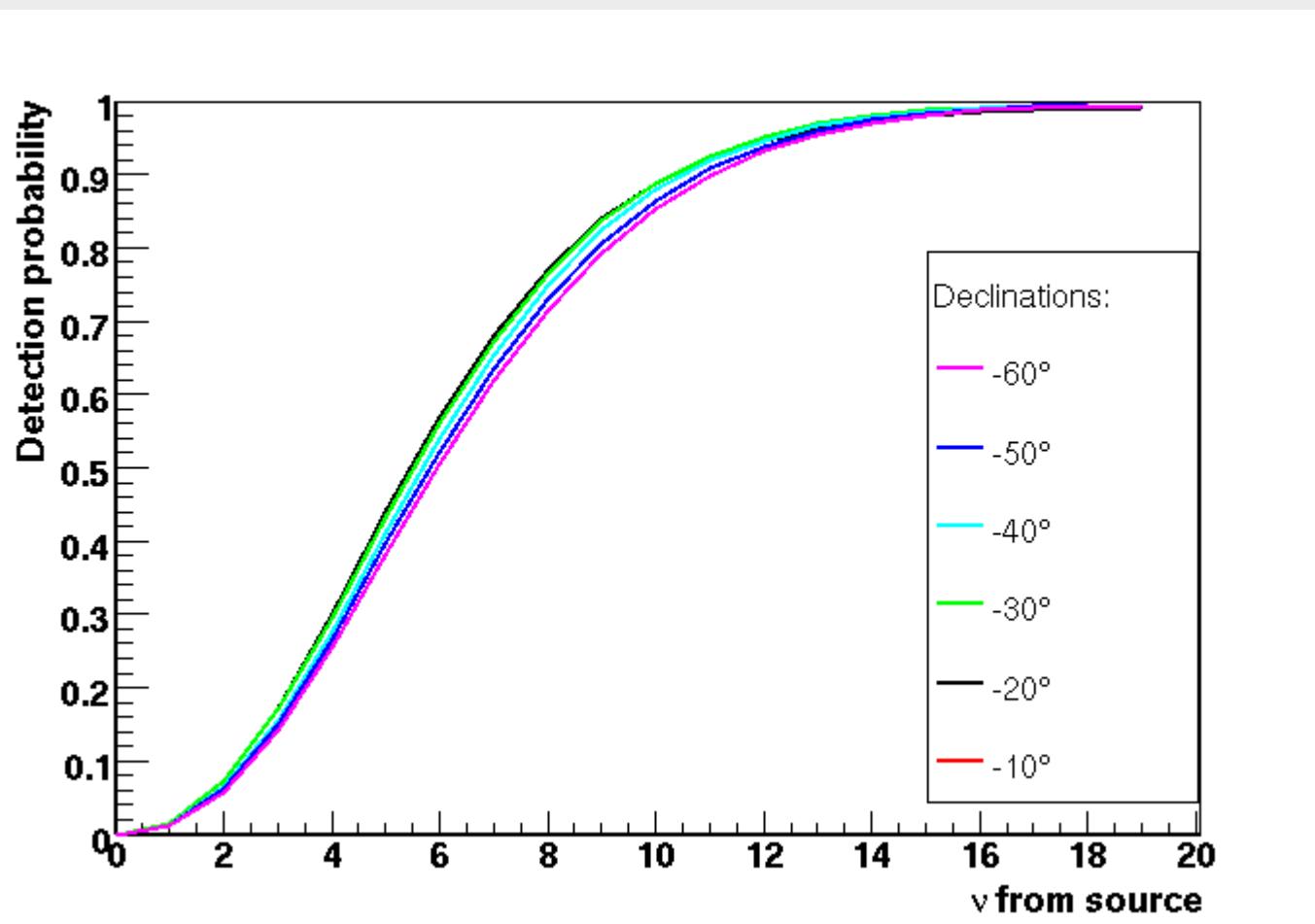
Set of clusters



If more than 1 cluster
& at least 1 cluster with more than 1 v
 \rightarrow cut 1 v clusters

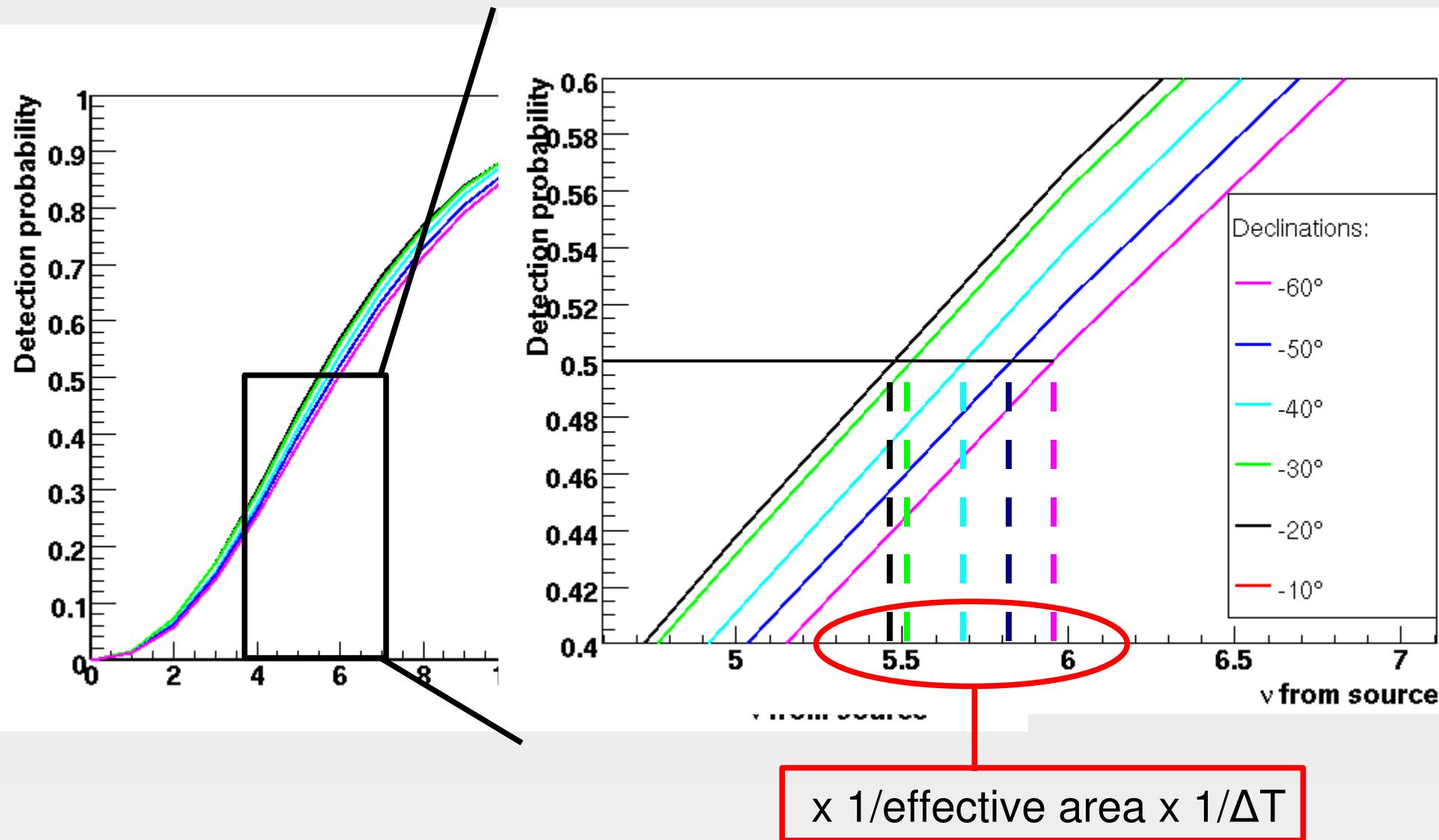


Discovery potential

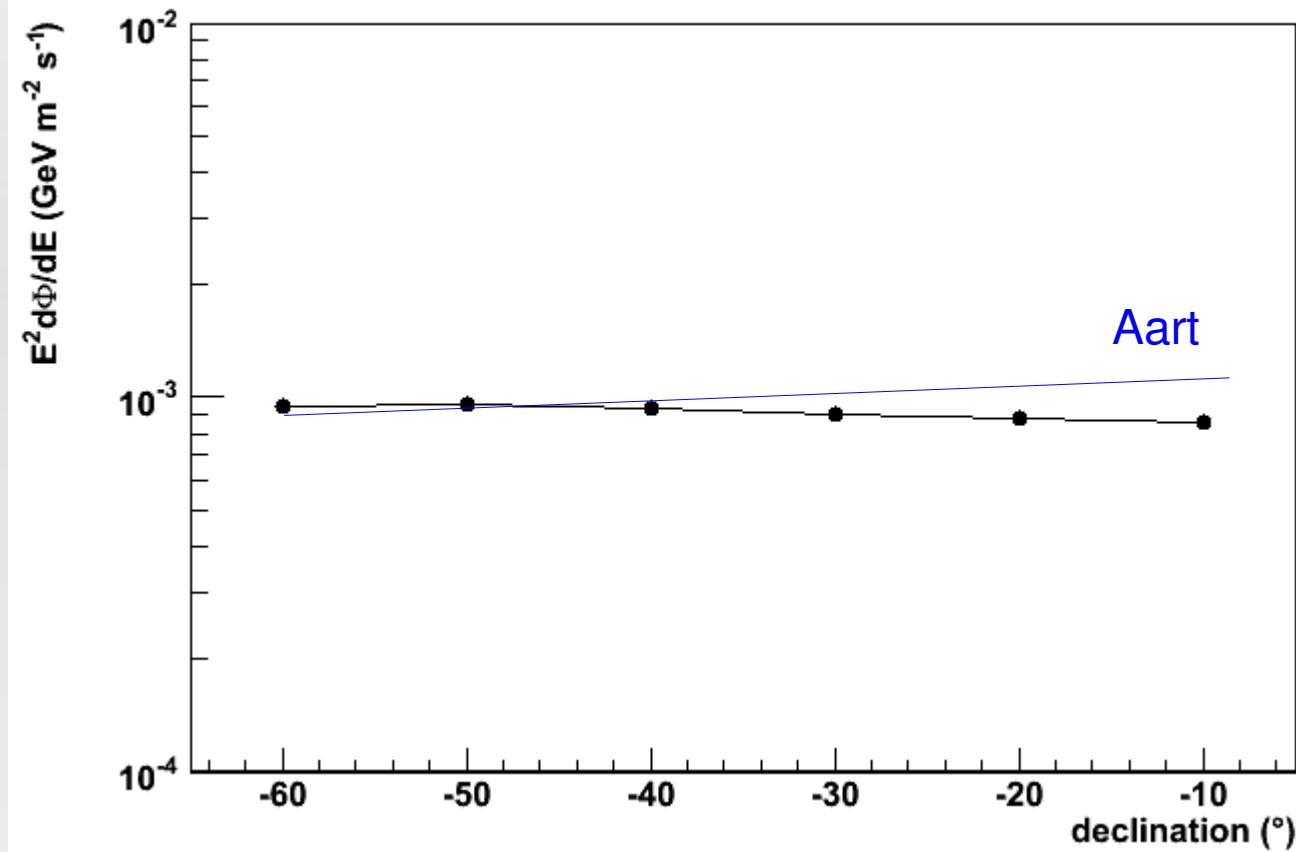


FDR controled at 0.5% \leftrightarrow C.L. At 99.5% ($\sim 3\sigma$)

Flux at 50% chance discovery



Flux at 50% chance discovery



Building limits/sensitivity

Feldman Cousin style

Not Poissonian nor Gaussian PDF ► not tabulated



Use simulations w. source flux (e.g. E^{-2})
&
FC ordering principle algorithm

→ FC confidence belts at 90% C.L.

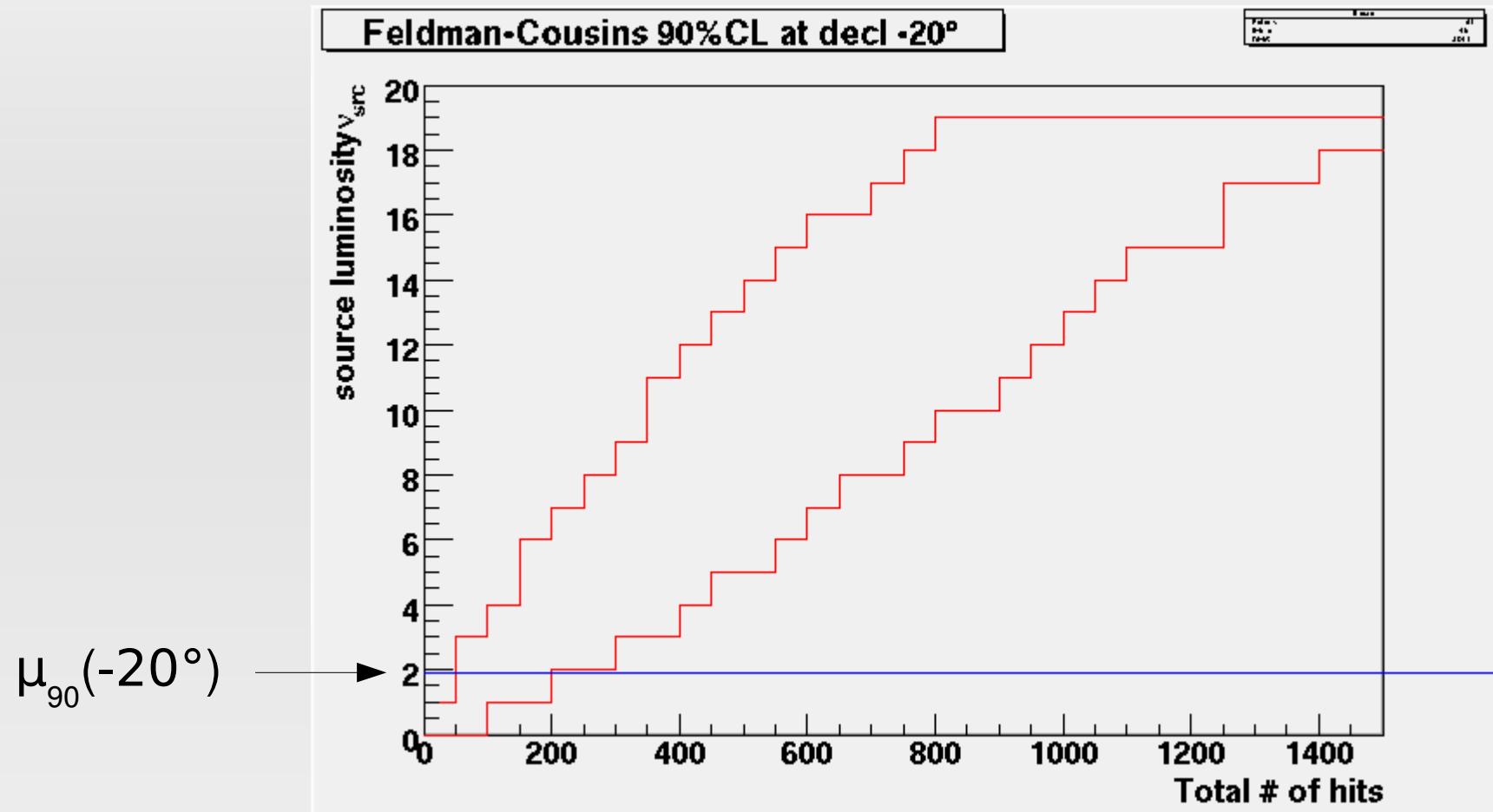


limit at 90% C.L.

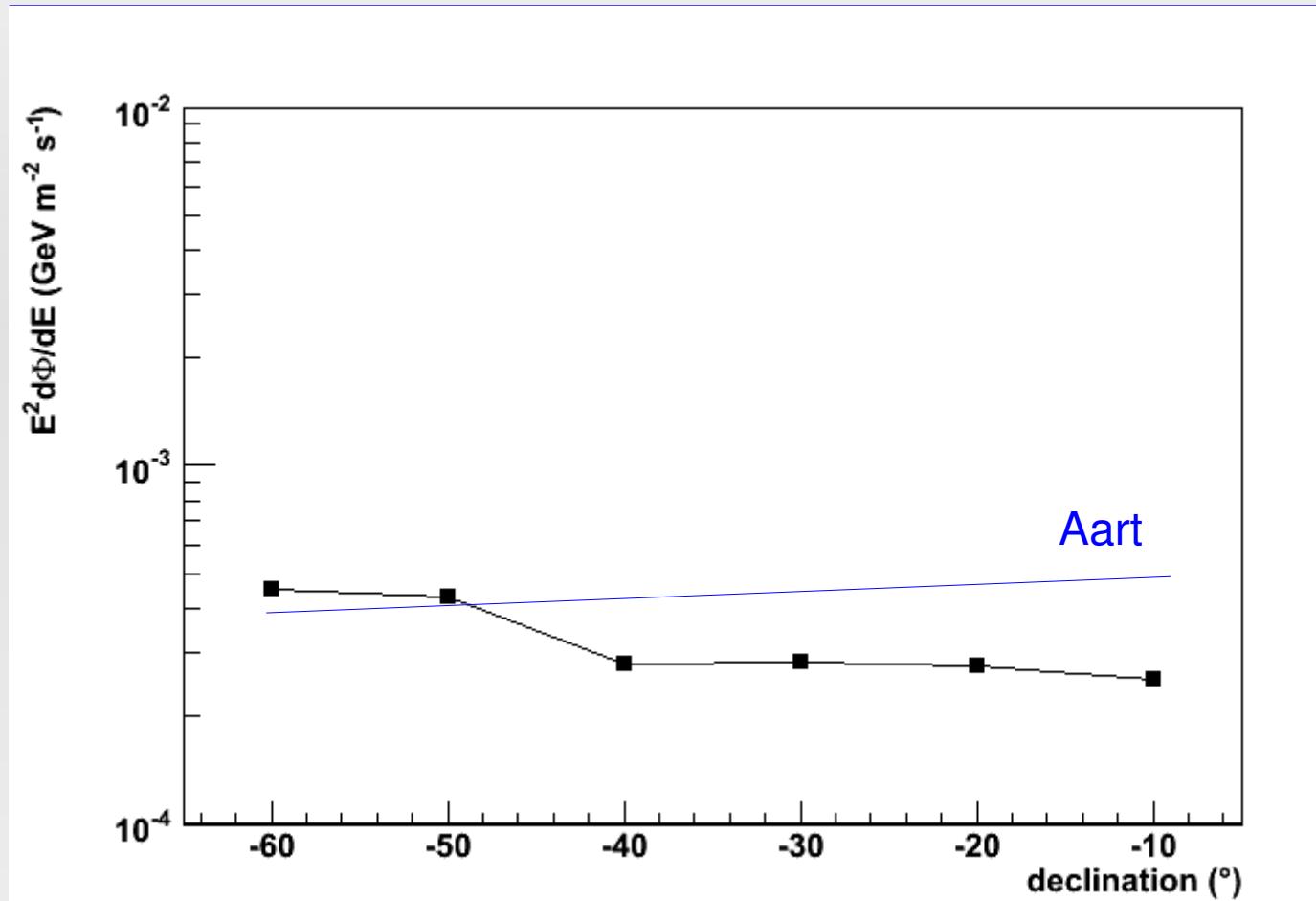


$\mu_{90}(\delta)$

Example of FC confidence belt



Limit for a E^{-2} flux at 90% C.L.



Summary and outlook



Main characteristics:

- ▶ Multiple Hypothesis tests approach
- ▶ « Built in » Confidence Level control
- ▶ Model independent

First results:

- ▶ 1 yr sensitivity and discovery potential comparable to Aart
- ▶ Alternative analysis => Cross check

Outlook:

- ▶ 5 lines analysis (on going)
- ▶ Inclusion of time information (well... much later....)