Look Elsewhere Effect

Look Elsewhere Effect

Eilam Gross

Weizmann Institute of Science

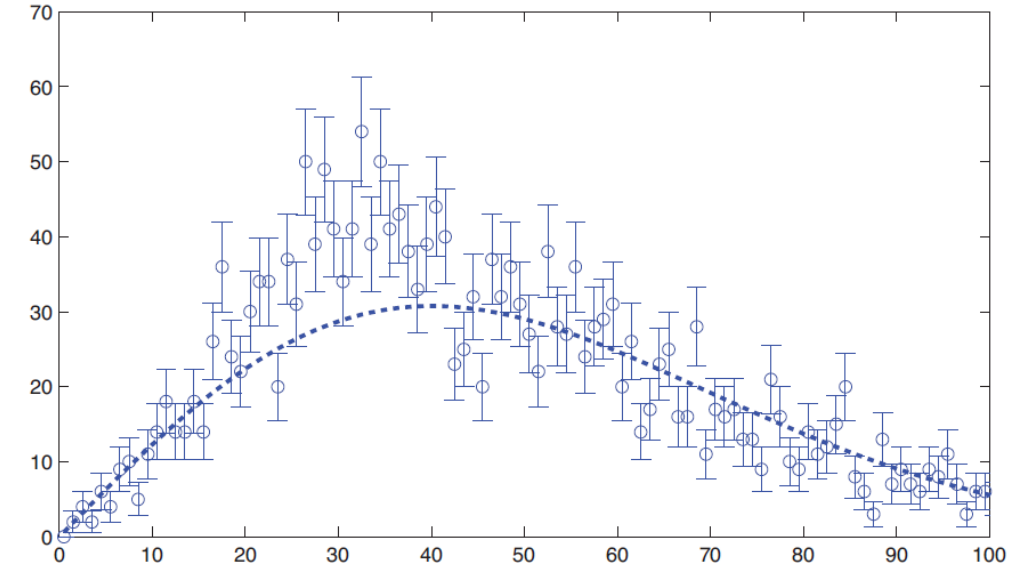
E.G., O. Vitells “Trial factors for the look elsewhere effect in high energy physics”,
Eur. Phys. J. C 70 (2010) 525

O. Vitells and E. G., Estimating the significance of a signal in a multi-dimensional search,
1669 Astropart. Phys. 35 (2011) 230, arXiv:1105.4355

Search for resonances in diphoton events at s√=13 TeV with the ATLAS detector, ATLAS collaboration, arXiv 1606.03833

Look Elsewhere Effect

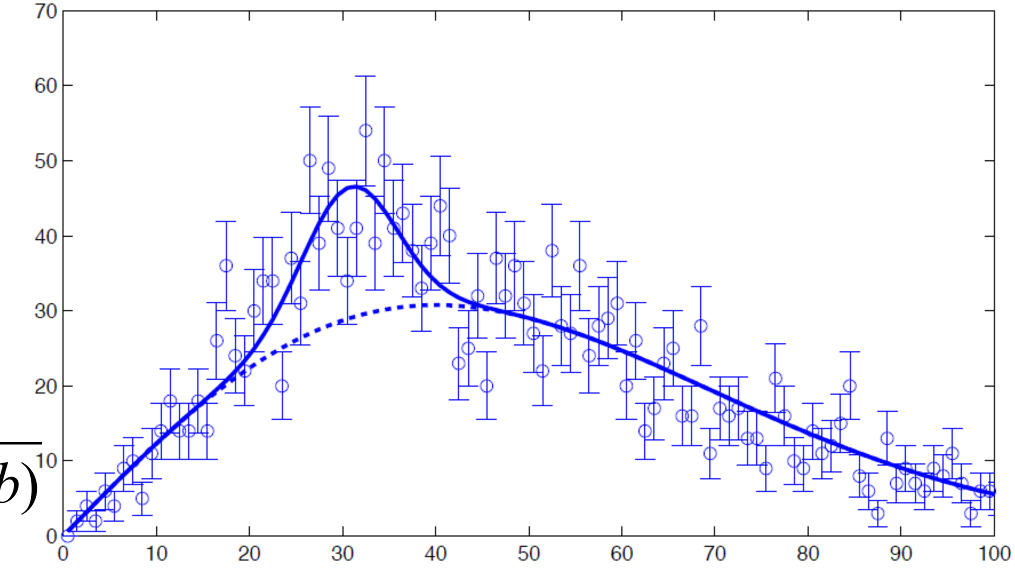
- Is there a signal here?



Look Elsewhere Effect

- Looks like @ m=30
- What is its significance?
- What is your test statistic?

$$q_{fix,obs} = -2\ln\frac{L(b)}{L(\hat{\mu}s(m=30)+b)}$$



Look Elsewhere Effect

- Test statistic

$$q_{fix,obs} = -2\ln\frac{L(b)}{L(\hat{\mu}s(m=30)+b)}$$

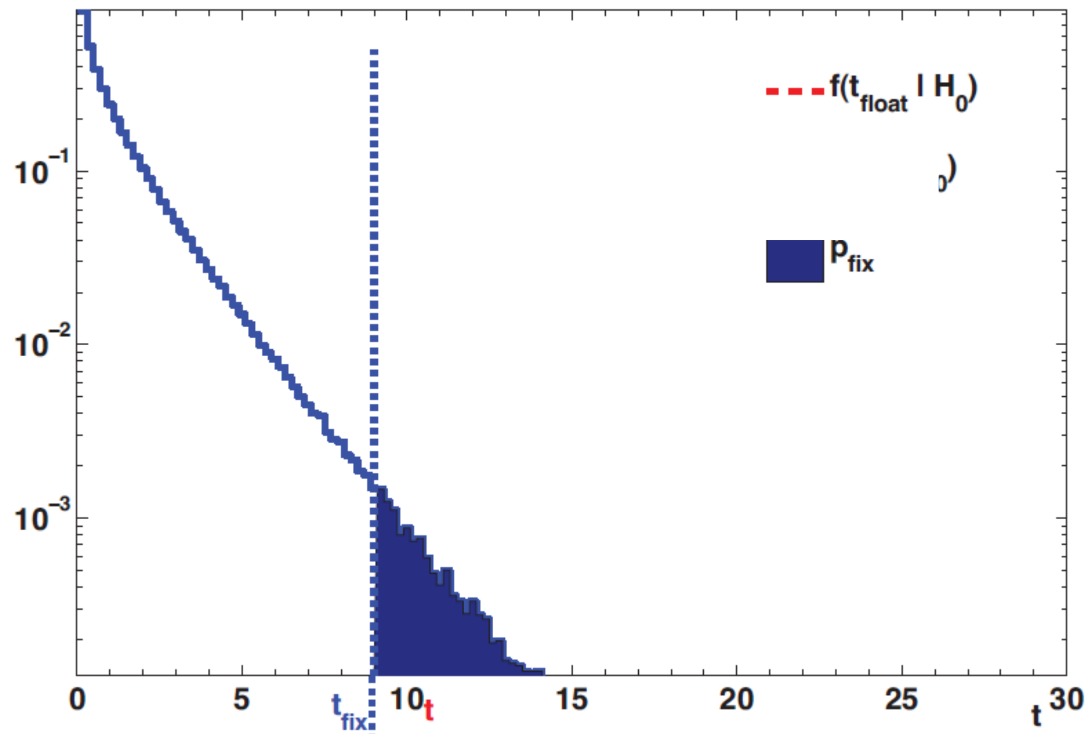
- What is the p-value?
- generate the PDF

$$f(q_{fix} \mid H_0)$$

and find the **p-value**

Wilks theorem:

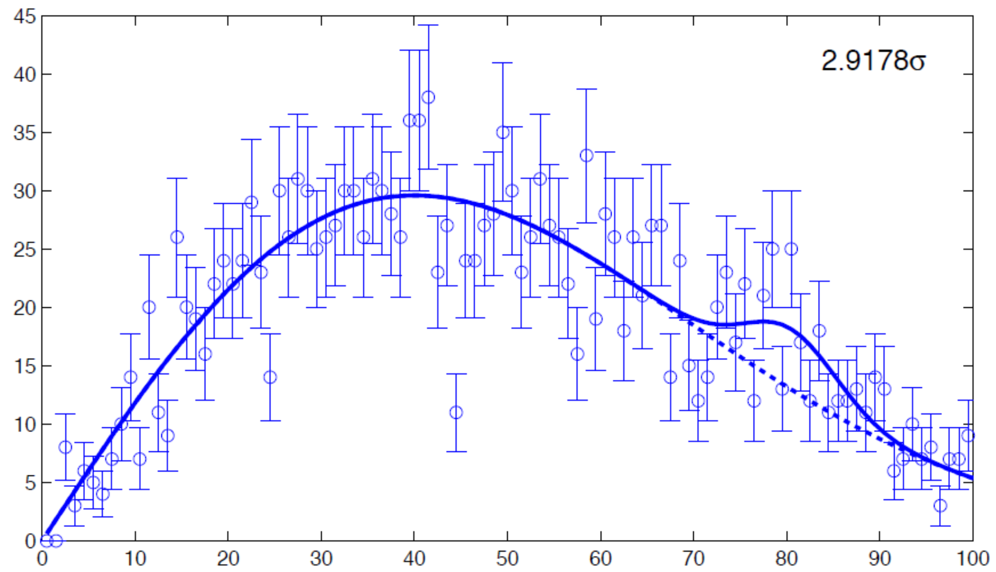
$$f(q_{fix} \mid H_0) \sim \chi_1^2$$



$$p_{fix} = \int_{q_{fix,obs}}^{\infty} f(q_{fix} \mid H_0)\, dq_{fix}$$

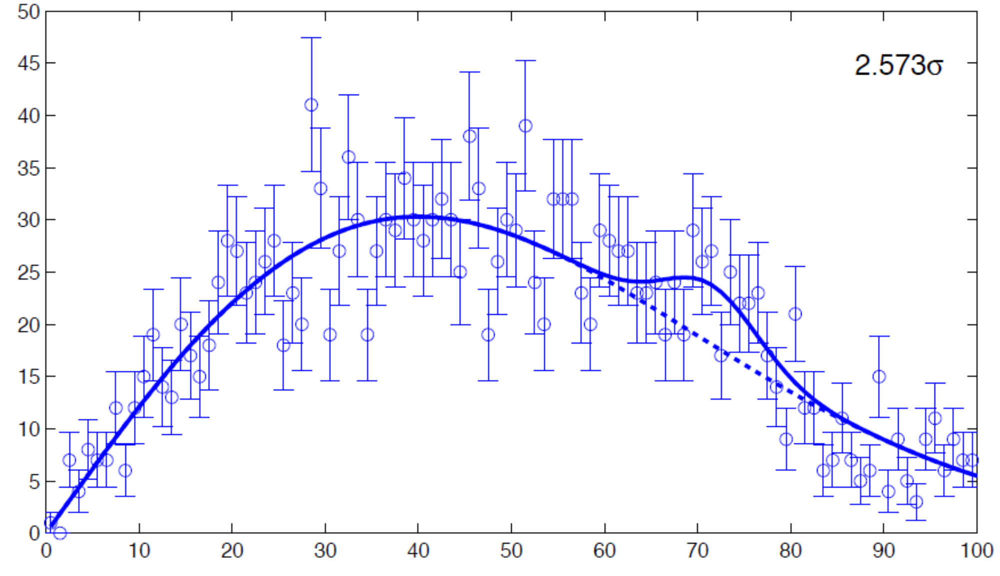
Look Elsewhere Effect

- Would you ignore this signal, had you seen it?



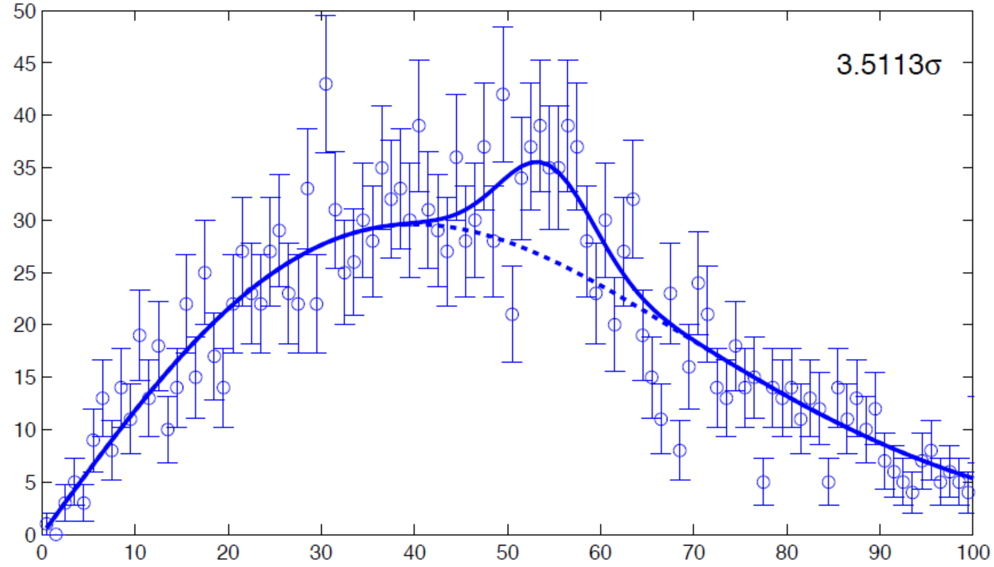
Look Elsewhere Effect

- Or this?



Look Elsewhere Effect

- Or this?

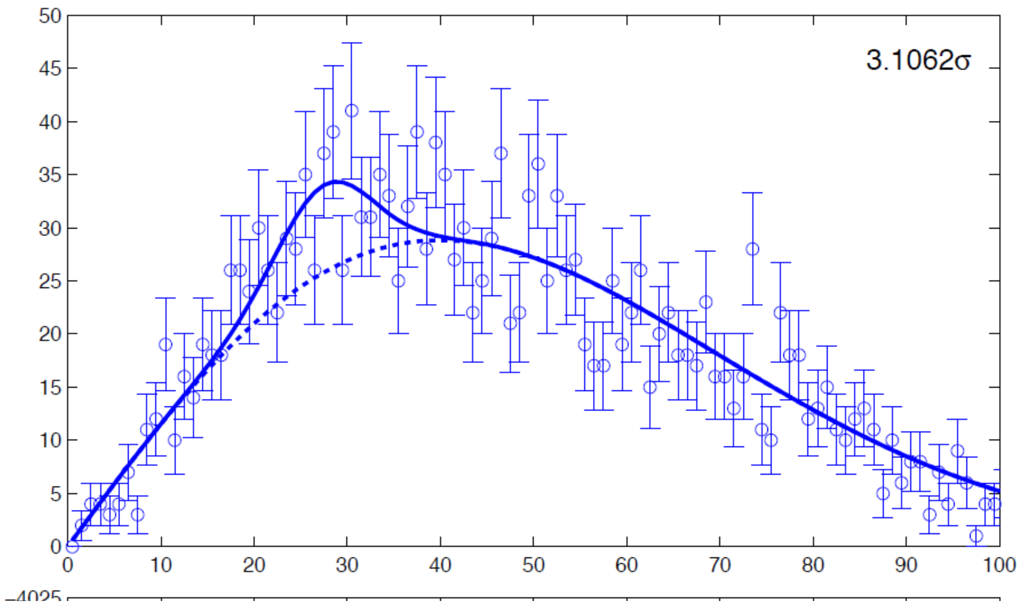


Look Elsewhere Effect

- Or this?

- Obviously NOT!

- ALL THESE “SIGNALS” ARE BG FLUCTUATIONS



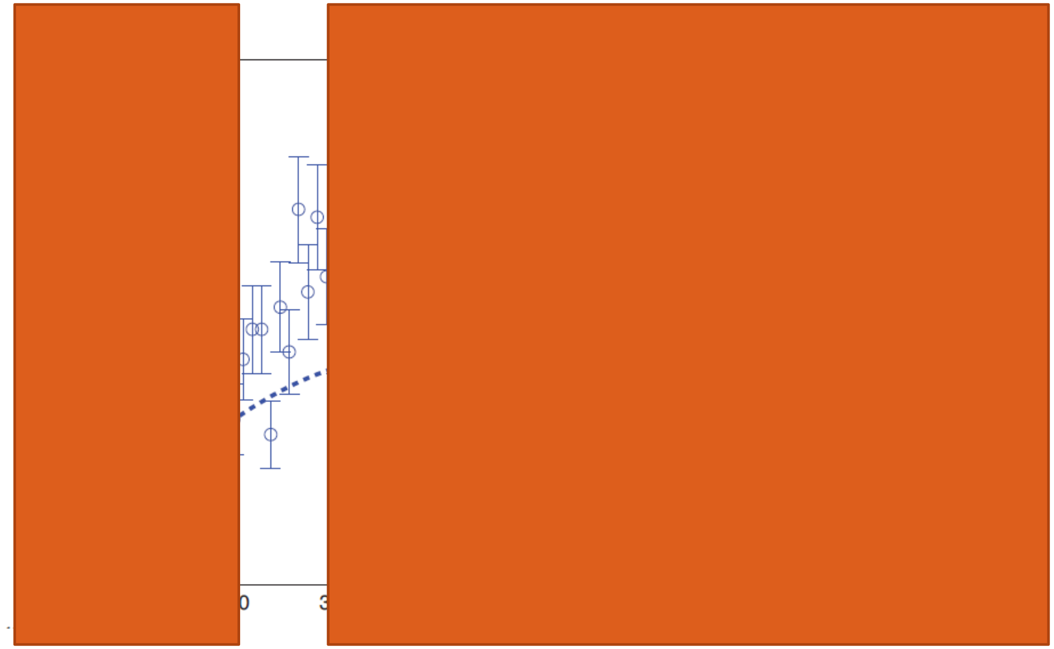
Look Elsewhere Effect

- Having no idea where the signal might be there are two options
- **OPTION I:**

scan the mass range in pre-defined steps and test any disturbing fluctuations

(do not confuse me with the facts)

- Perform a fixed mass analysis at each point



$$q_{fix,obs}(\hat{\mu}) = -2\ln\frac{L(b)}{L(\hat{\mu}s(m)+b)}$$

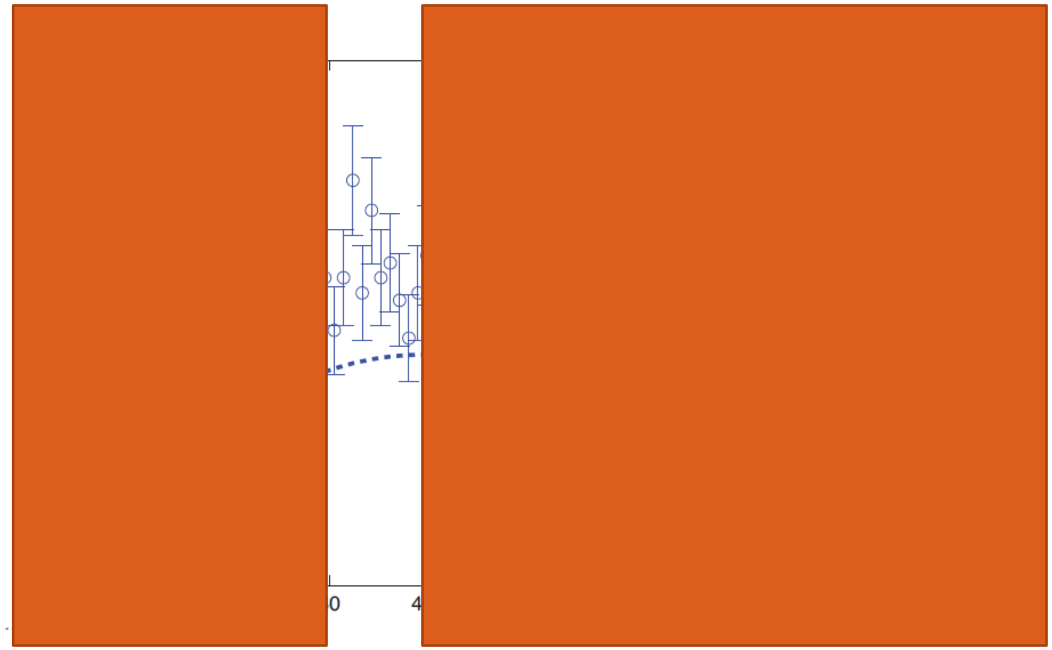
Look Elsewhere Effect

- Having no idea where the signal might be there are two options
- **OPTION I:**

scan the mass range in pre-defined steps and test any disturbing fluctuations

(do not confuse me with the facts)

- Perform a fixed mass analysis at each point

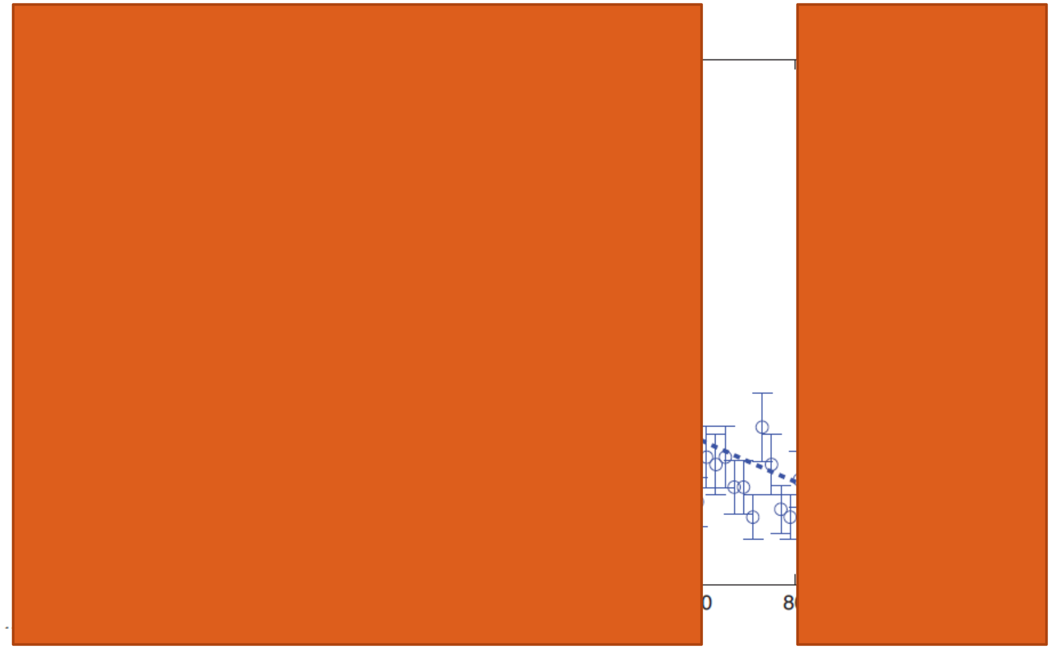


$$q_{fix,obs}(\hat{\mu}) = -2\ln\frac{L(b)}{L(\hat{\mu}s(m)+b)}$$

Look Elsewhere Effect

- The scan resolution must be less than the signal mass resolution

- Assuming the signal can be only at one place, pick the one with the smallest p-value (maximum significance)



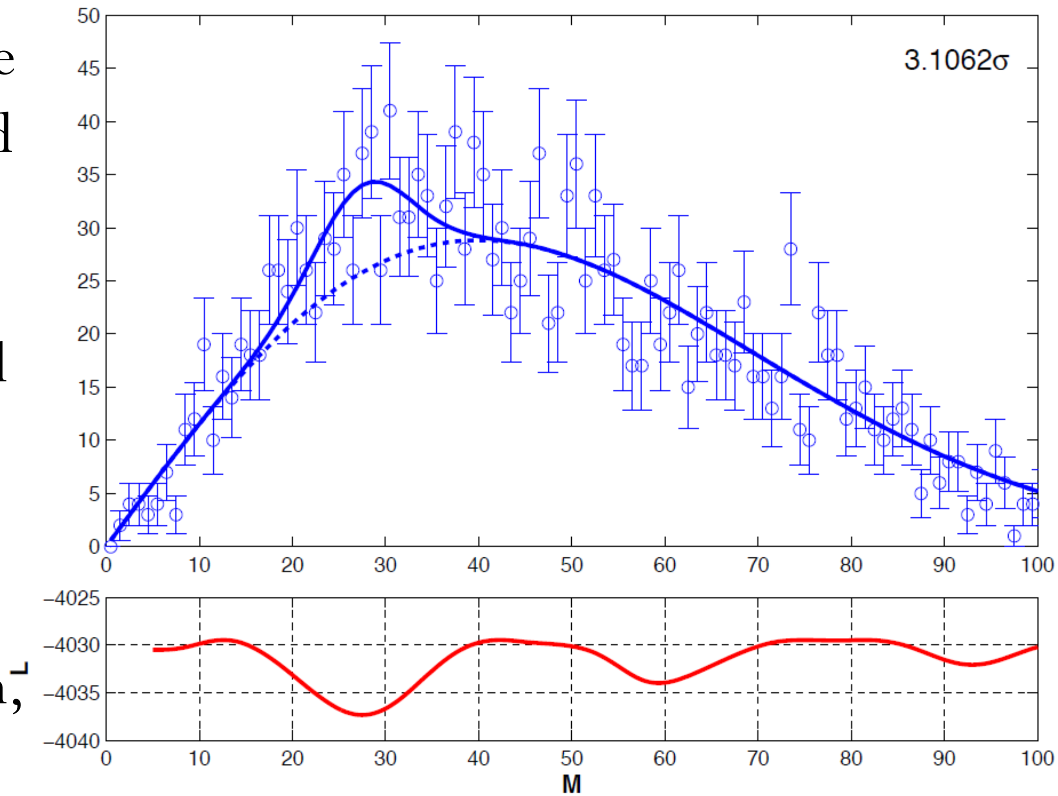
$$q = \max_m \left\{ q_{fix,obs}(\hat{\mu}) \right\} = \max_m \left\{ -2\ln \frac{L(b)}{L(\hat{\mu}s(m)+b)} \right\}$$

$$= \min_m \left\{ p - value \right\}$$

Look Elsewhere Effect: Floating Mass

•Option II:
Leave the mass floating

•Having no idea where the signal might be you would allow the signal to be anywhere in the **search range** and use a modified test statistic



$$q_{float,obs}(\hat{\mu},\hat{m}) = -2\ln\frac{L(b)}{L(\hat{\mu}s(\hat{m})+b)}$$

•For the same observation, the p-value increases because more possibilities are opened

Look Elsewhere Effect

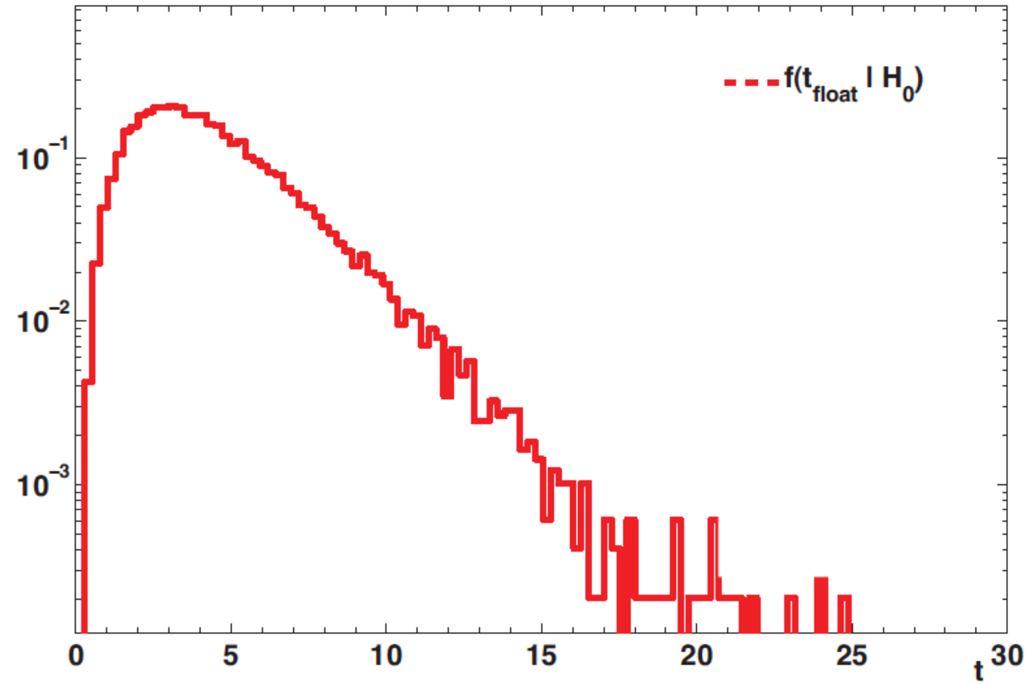
- the test statistis

$$q_{float,obs}(\hat{\mu},\hat{m}) = -2\ln\frac{L(b)}{L(\hat{\mu}s(\hat{m})+b)}$$

- The null hypothesis PDF

$$f(q_{float}\,|\,H_0)$$

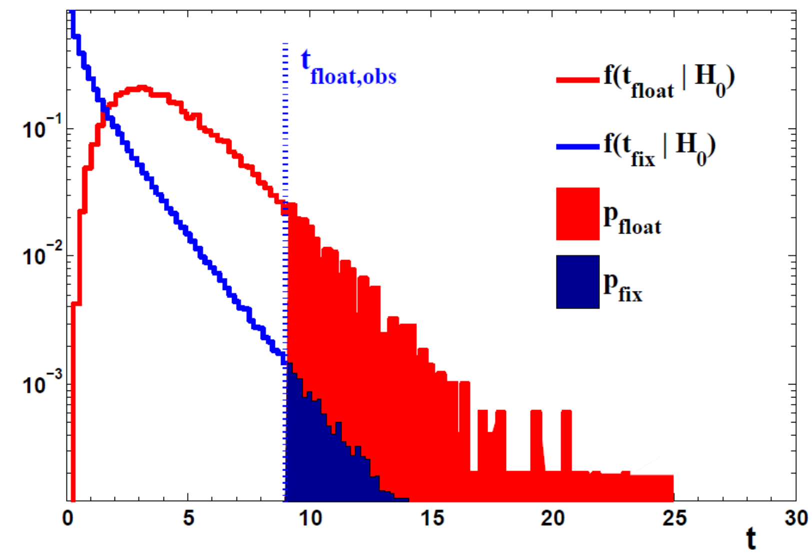
does not follow a chi-squared with 2dof because there are multiple minima depending on the size of the search range and resolution



trial#

- Assume a maximal local fluctuation at mass $\hat{m} = 30$
- We can calculate the following p-value

$$q_{fix,obs} = q_{float,obs} = -2\ln\frac{L(b)}{L(\hat{\mu}s(\hat{m}=m=30)+b)}$$



$$p_{fix} = \int_{q_{obs}} f(q_{fix} \mid H_0)\, dq_{fix} < p_{float} = \int_{q_{obs}} f(q_{float} \mid H_0)\, dt_{float}$$

$$\boxed{trial\# = \frac{\int_{q_{obs}} f(q_{float} \mid H_0)\, dt_{float}}{\int_{q_{obs}} f(q_{fix} \mid H_0)\, dt_{fix}} = \frac{p_{float}}{p_{fix}} > 1}$$

Can we analytically calculate the trial# (or p_float)?

Define the Problem

- Let $n = \mu s(m,\Gamma) + b$
- m,Γ are nuisance parameters undefined under the null hypothesis $\mu = 0$
- What is the pdf of

$$\hat{q}_0 \equiv q_0(\hat{m},\hat{\Gamma}) = -2\log\frac{\mathcal{L}(\mu=0)}{\mathcal{L}(\hat{\mu},\hat{m},\hat{\Gamma})} = \max_{m,\Gamma}[q_0(m,\Gamma)]$$

under the null hypothesis

- To generalize the problem , let θ be the nuisance parameter, undefined under the null hypothesis, and let us try to find out the pdf of

$$\boxed{\hat{q}_0 \equiv q_0(\hat{\theta}) = -2\log\frac{\mathcal{L}(\mu=0)}{\mathcal{L}(\hat{\mu},\hat{\theta})} = \max_{\theta}[q_0(\theta)]}$$

for which we want to calculate $\boxed{\text{p-value}=P(\max_{\theta}[q_0(\theta)] \ge u), u = Z^2}$

The profile-likelihood test statistic

(with a nuisance parameter that is not defined under the Null hypothesis)

- Consider the teststatistic:

$$q_0(\theta) = -2\log\frac{\mathcal{L}(\mu=0)}{\mathcal{L}(\hat{\mu},\theta)} \qquad \begin{aligned} H_0 &: \mu = 0 \\ H_1 &: \mu > 0 \end{aligned}$$

μ="signal strength"

- For some <u>fixed</u> *θ*, *$q_0(\theta)$* has (asymptotically) a chi^2 distribution with one degree of freedom by Wilks' theorem.

- *$q_0(\theta)$* is a <u>*chi^2 random field*</u> over the space of *θ* (a random variable indexed by a continuous parameter(s)). we are interested in

$$\hat{q}_0 \equiv q_0(\hat{\theta}) = -2\log\frac{\mathcal{L}(\mu=0)}{\mathcal{L}(\hat{\mu},\hat{\theta})} = \max_{\theta}[q_0(\theta)]$$

$\hat{\theta}$ is the **global** maximum point

- For which we want to know what is the p-value

$$\text{p-value=P}(\max_{\theta}[q_0(\theta)] \geq u),\ \ u = Z^2$$

The profile-likelihood test statistic

(with a nuisance parameter that is not defined under the Null hypothesis)

- Usually we only look for 'positive' signals

(downward fluctuations of the BG are not considered as evidence against the BG)

$$q_0(\theta) = \begin{cases} -2\log\dfrac{\mathcal{L}(\mu=0)}{\mathcal{L}(\hat{\mu},\theta)} & \hat{\mu} > 0 \\ 0 & \hat{\mu} \le 0 \end{cases}$$

$q_0(\theta)$ is 'half chi^2'

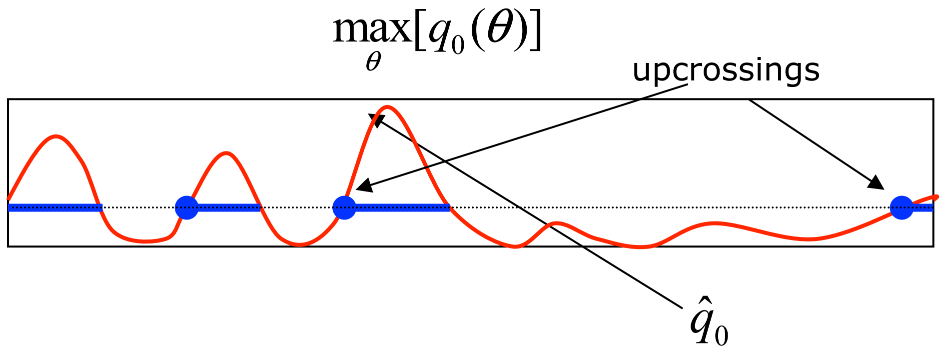
[H. Chernoff, Ann. Math. Stat. 25, 573578 (1954)]

The p-value just get divided by 1/2

- Or equivalently consider $\hat{\mu}$ as a gaussian field

(since $q_0(\theta) = \left(\dfrac{\hat{\mu}(\theta)}{\sigma}\right)^2$)

Random fields (1D)

- In 1 dimension: points where the field values become larger then *u* are called *upcrossings.*
- 

- The probability that the global maximum is above the level u is called *exceedance probability*. (p-value of $\hat{q}_0$)

$$P(\max_{\theta}[q_0(\theta)] \geq u)$$

Random fields

- Fortunately, quite a lot of statistical literature on the properties of random fields

[3] R.J. Adler and A.M. Hasofer, *Level Crossings for Random Fields*, Ann. Probab. **4**, Number 1 (1976), 1-12.

[4] R.J. Adler, *The Geometry of Random Fields*, New York (1981), Wiley, ISBN: 0471278440.

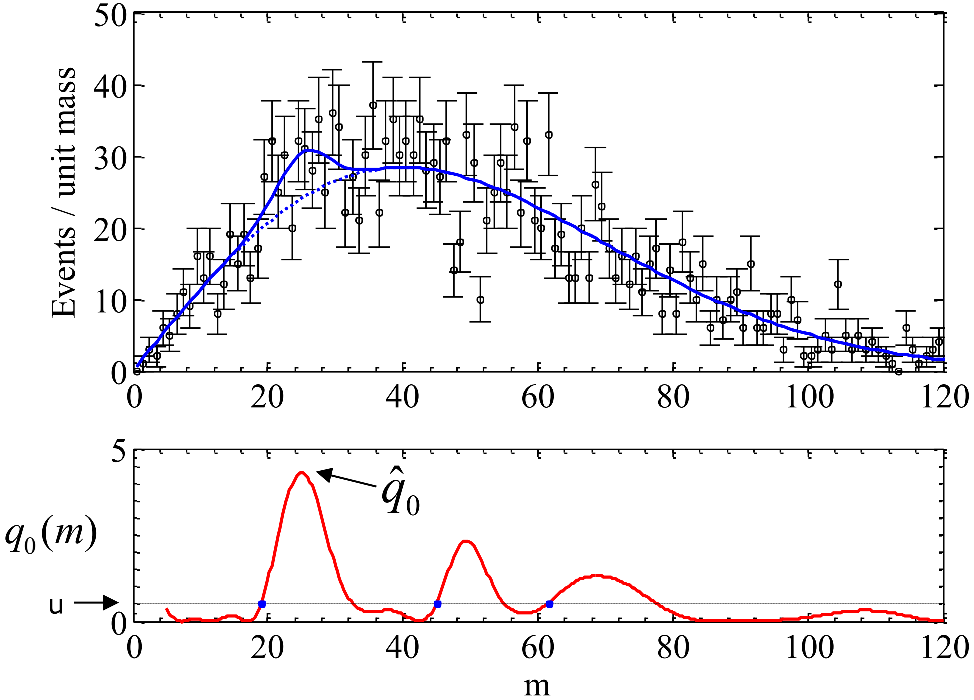
[5] K.J. Worsley, S. Marrett, P. Neelin, A.C. Vandal, K.J. Friston and A.C. Evans, *A Unified Statistical Approach for Determining Significant Signals in Location and Scale Space Images of Cerebral Activation*, Human Brain Mapping **4** (1996) 58-73.

[6] R.J. Adler and J.E. Taylor, *Random Fields and Geometry* , Springer Monographs in Mathematics (2007). ISBN: 978-0-387-48112-8.

[9] J. Taylor, A. Takemura and R.J. Adler, *Validity of the expected Euler characteristic heuristic*, Ann. Probab. 33 (2005) 1362-1396.

Applications in Cosmology, Brain mapping, Oceanography …

The 1-dimensional case



For a chi² random field, the expected number of *upcrossings* of a level *u* is given by: [Davies,1987]

$$E[N_u] = \mathcal{N}_1 e^{-u/2}$$

Note the inequality:

$$E[N_u] \geq \mathrm{P}(N_u > 0)$$

To have the global maximum above a level *u:*

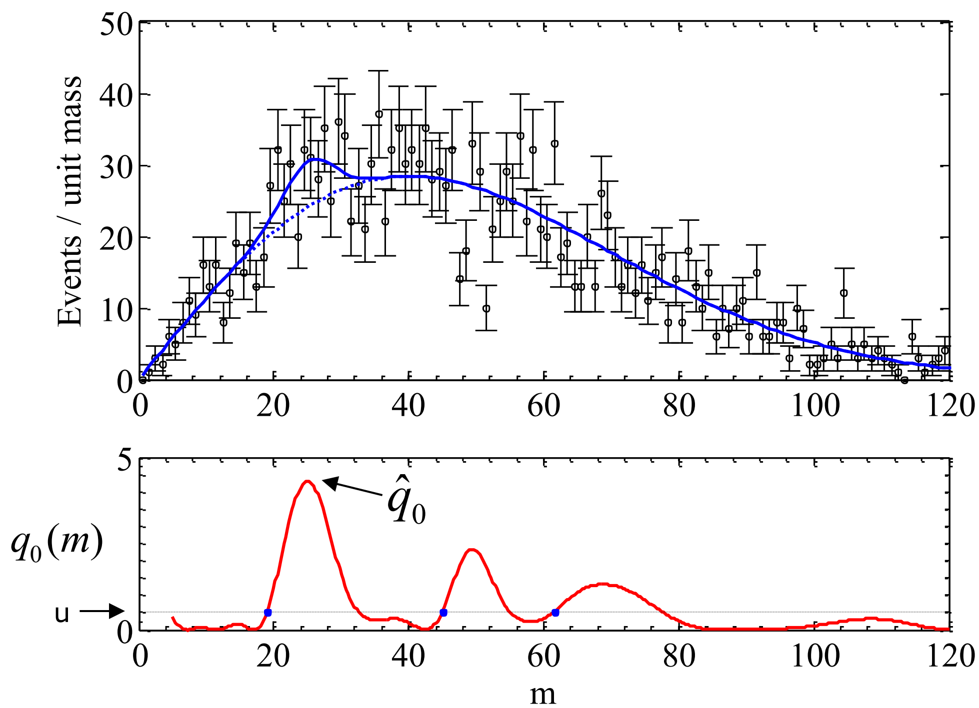
- Either have at least one upcrossing ($N_u>0$) ***or*** have $q_0>u$ at the origin ($q_0(0)>u$) **:**

$$P(\hat{q}_0 > u) \leq P(N_u > 0) + P(q_0(0) > u)$$
$$\leq E[N_u] + P(q_0(0) > u)$$

[R.B. Davies, Hypothesis testing when a nuisance parameter is present only under the alternative. Biometrika **74**, 33–43 (1987)]

Becomes an equality for large *u*

The 1-dimensional case



$$E[N_u] = \mathcal{N}_1 e^{-u/2}$$

The only unknown is $\mathcal{N}_1$, which can be estimated from the average number of upcrossings at some low reference level

$$E[N_u] = \mathcal{N}_1 e^{-u/2}$$

$$E[N_{u_0}] = \mathcal{N}_1 e^{-u_0/2} \Rightarrow$$

$$\mathcal{N}_1 = E[N_{u_0}] e^{u_0/2}$$

The p-value can then be estimated by Davies' formula

$$\mathrm{P}(q_0 > u) \leq \mathcal{N}_1 e^{-u/2} + \frac{1}{2} P(\chi_1^2 > u) \Rightarrow$$

$$\mathrm{P}(q_0 > u) \leq E[N_{u_0}] e^{(u_0 - u)/2} + \frac{1}{2} P(\chi_1^2 > u)$$

$$P_{global}(u) \approx p_{local}(u) + E(n_{u_0}) e^{\frac{u_0 - u}{2}}$$

Spin 0 2015

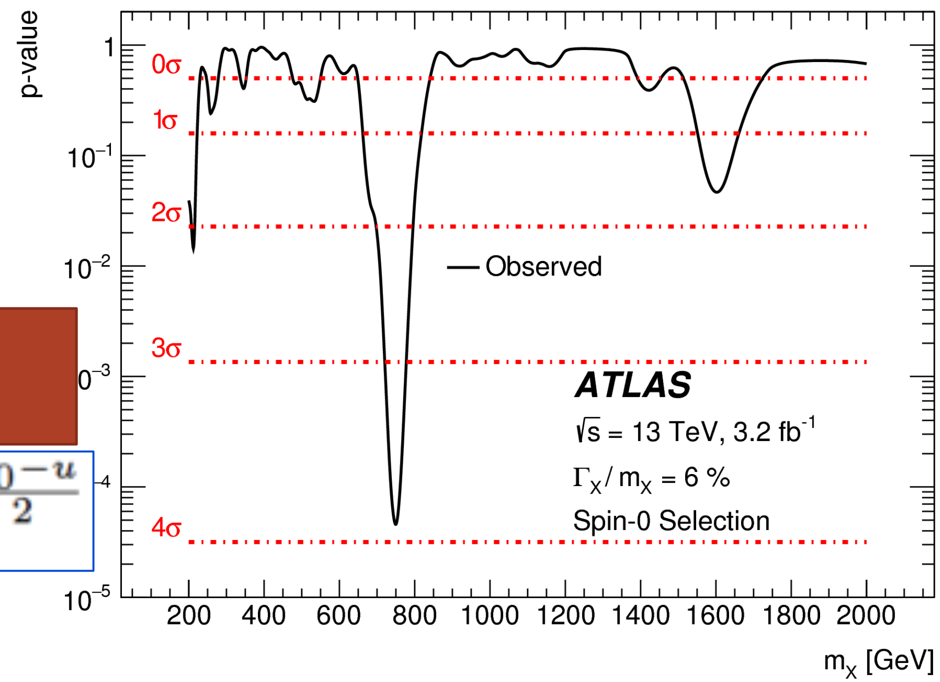
Largest significance

$m_x \sim 750GeV, \Gamma_x \sim 45GeV(6$

Local Z = 3.9σ

Any peak with Z>3.8 σ
with m=500-2000 will draw our attention

$$P_{global}(u) \approx p_{local}(u) + E(n_{u_0})e^{\frac{u_0-u}{2}}$$



$$p_{local} = 5 \cdot 10^{-5}$$

$$u_0 = 0$$

$$n_{u_0} = 7 \pm 2.6$$

$$u = Z^2 = 3.9^2 = 15.2$$

$$p_{global} = 5 \cdot 10^{-5} + (7 \pm 2.6)e^{-15.2/2} = (2.2 - 4.8)10^{-3}$$

$$Z_{global} \sim 2.7 \pm 0.1\sigma$$

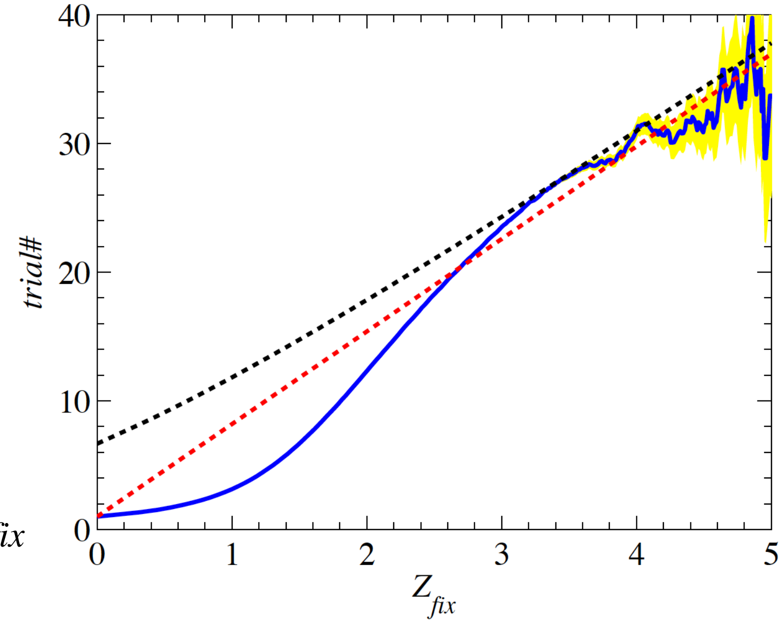
The LEE is even stronger when you consider another dimension (the width range (0-10%m) should also be taken into account)

- The Trial factor is given by

$$u = Z^2$$

$$trial\,\# = \frac{p_{global}}{p_{local}} \le \frac{\mathcal{N}_1 e^{-u/2} + \frac{1}{2}P(\chi_1^2 > u)}{\frac{1}{2}P(\chi_1^2 > u)}$$

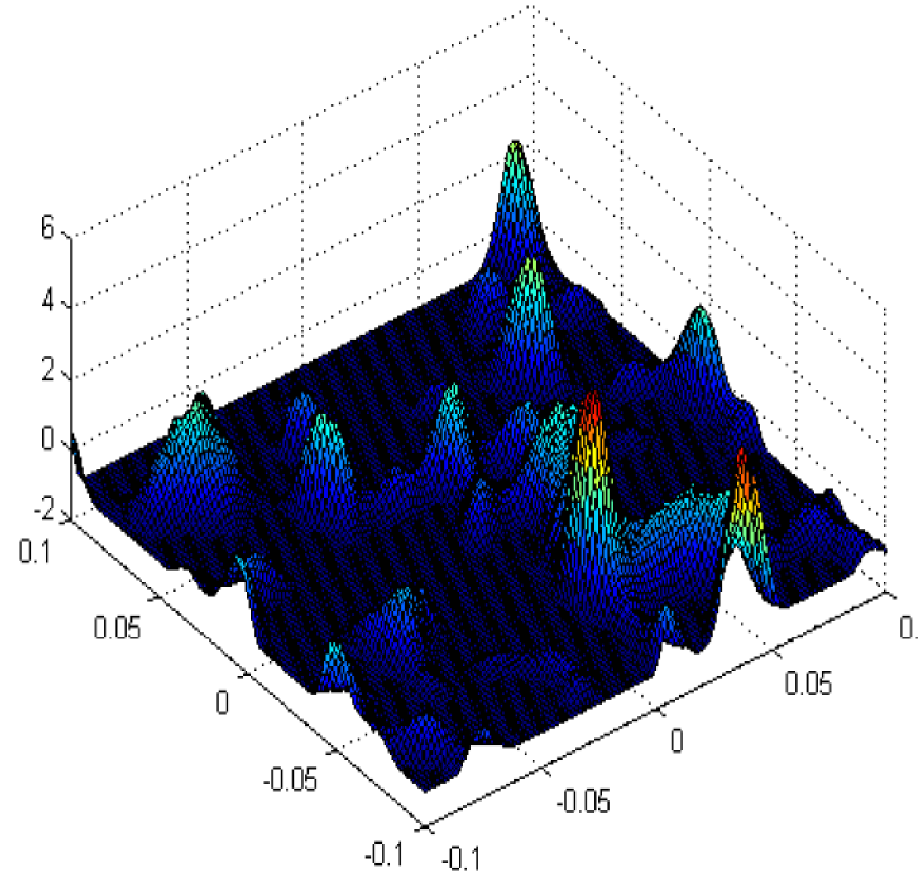
$$trial\,\#(u >> 1) = 1 + \sqrt{\frac{\pi}{2}}\mathcal{N}_1 Z_{fix} \approx \sqrt{\frac{\pi}{2}}\mathcal{N}_1 Z_{fix}$$



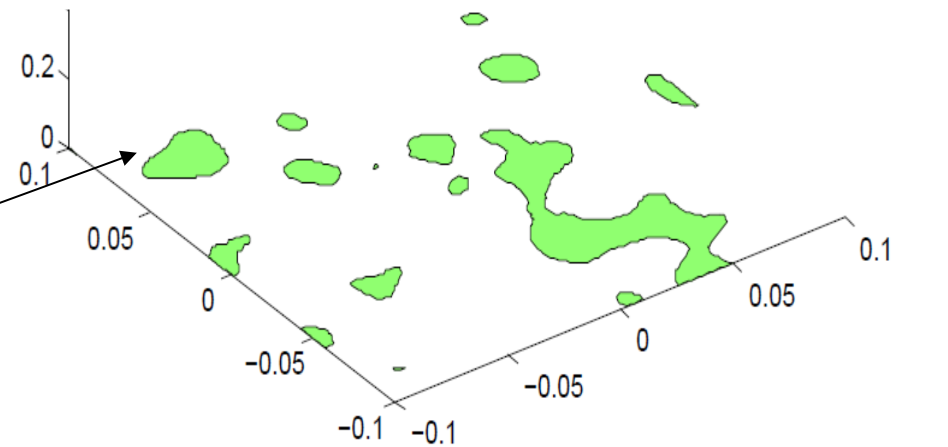
$\mathcal{N}_1$ Is the number of independent search regions

Random fields (>1 D)

- The set of points where the field has values larger then some number u is called the *excursion set* A_u above the level *u.*

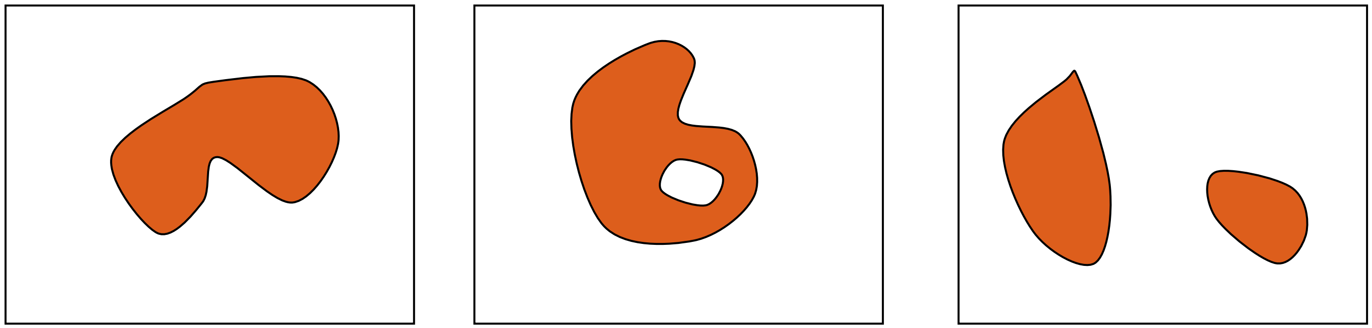


Excursion set



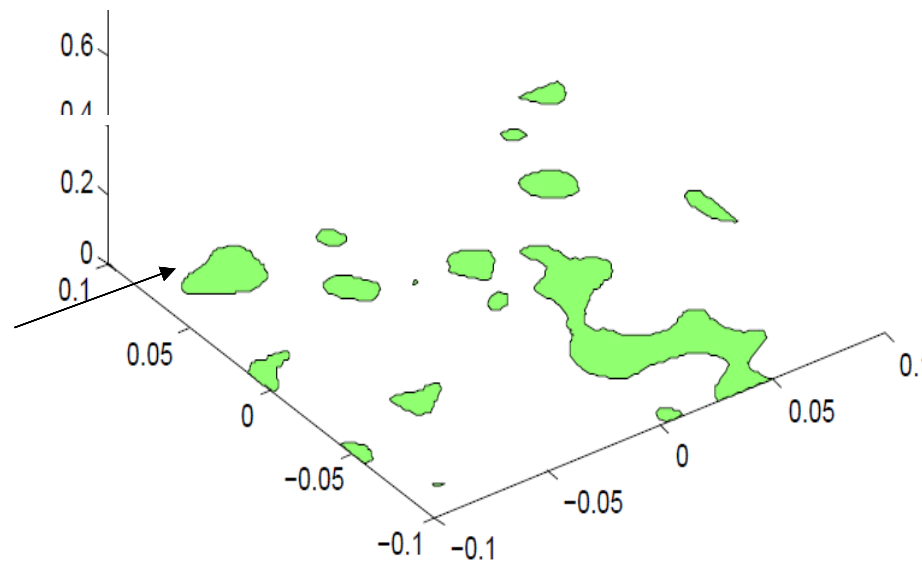
Euler characteristic

- Number of disconnected components minus number of `holes'



φ=1 φ=0 φ=2

Excursion set



The *n*-dimensional case

- The upcrossings formula is a special case of a more general result which gives ***the expectation of the Euler characteristic of the excursion set*** of a random field over a general *n*-dimensional manifold

$$\mathrm{E}[\varphi(A_u)] = \sum_{d=0}^{n} \mathcal{N}_d \rho_d(u)$$

- Here:

  A_u is the excursion set of the field above a level *u* (set of points where $q_0(\theta)>u$)
  $\varphi(A_u)$ is it's Euler characteristic
  ρ_d are 'universal' functions (depend only on the level *u* and the type of distribution)

e.g. for a chi^2 field with *s* degrees of freedom:

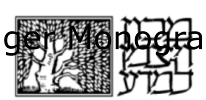
$$\rho_0(u) = \mathrm{P}(\chi_s^2 > u)$$

$$\rho_1(u) = u^{(s-1)/2} e^{-u/2}$$

$$\rho_2(u) = u^{(s-2)/2} e^{-u/2}[u-(s-1)]$$

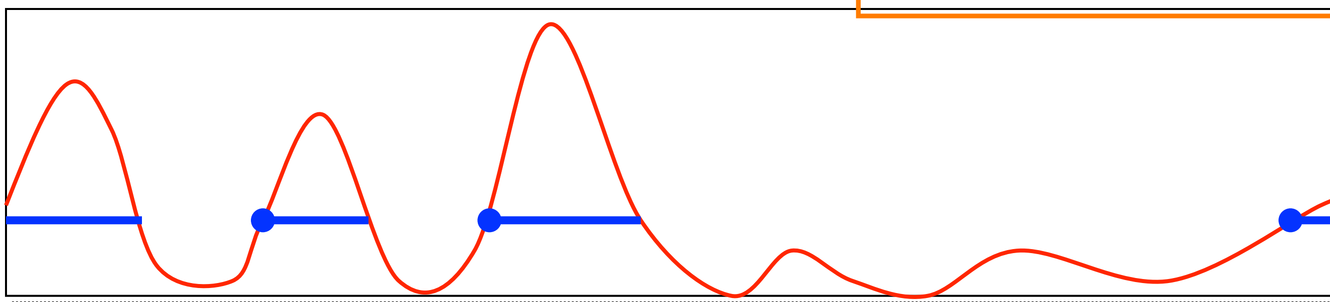
...

[R.J. Adler and J.E. Taylor, *Random Fields and Geometry* (2007 Springer Monographs in Mathematics]



Euler characteristic

$$\mathrm{E}[\varphi(A_u)] = \sum_{d=0}^{n} \mathcal{N}_d \rho_d(u)$$



In 1 dimension:

$$\varphi(A_u) = N_u + \mathbf{1}_{[q_0(0)>u]}$$

$$E[\varphi(A_u)] = E[N_u] + \mathrm{P}(q_0(0) > u)$$

$$= \mathcal{N}_1 e^{-u/2} + \frac{1}{2} P(\chi_1^2 > u)$$

(Davies' Bound)

The general case

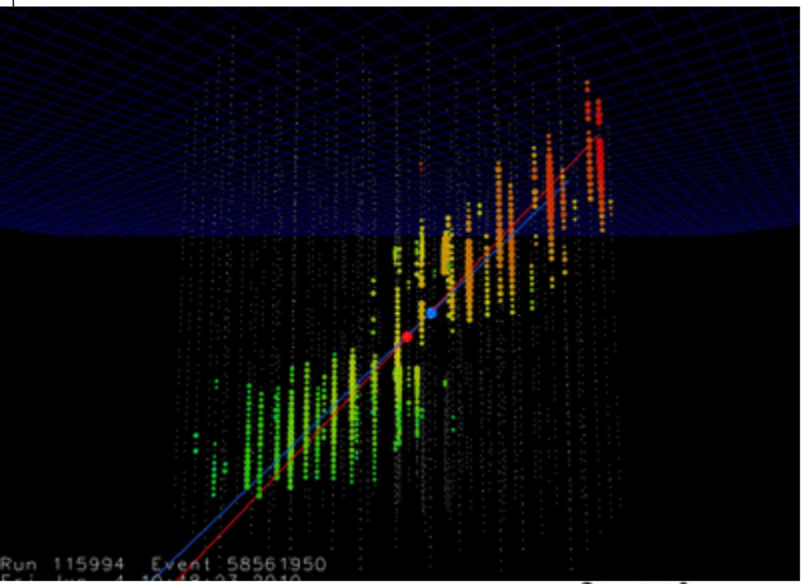
$\mathcal{N}_0 =$ Euler characteristic of the manifold

$$\rho_0(u) = \mathrm{P}(q_0 > u)$$

In general for high-level excursions $E[\varphi(A_u)] \rightarrow \mathrm{P}(\max_{\theta}[q_0(\theta)] \geq u)$

[J. Taylor, A. Takemura, and R. J. Adler, Ann. Probab. 33, 4 (2005)]

2-D example: IceCube search for astrophysical neutrino point sources



IceCube looks for neutrino sources,

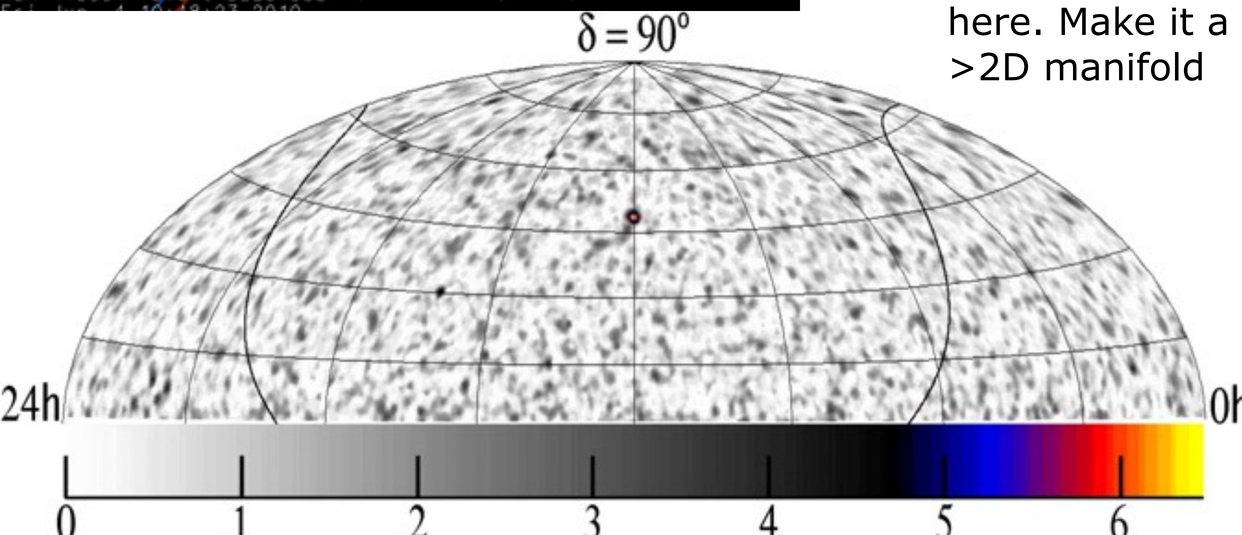
2-D Search over the sky (θ,φ)

Unbinned likelihood:

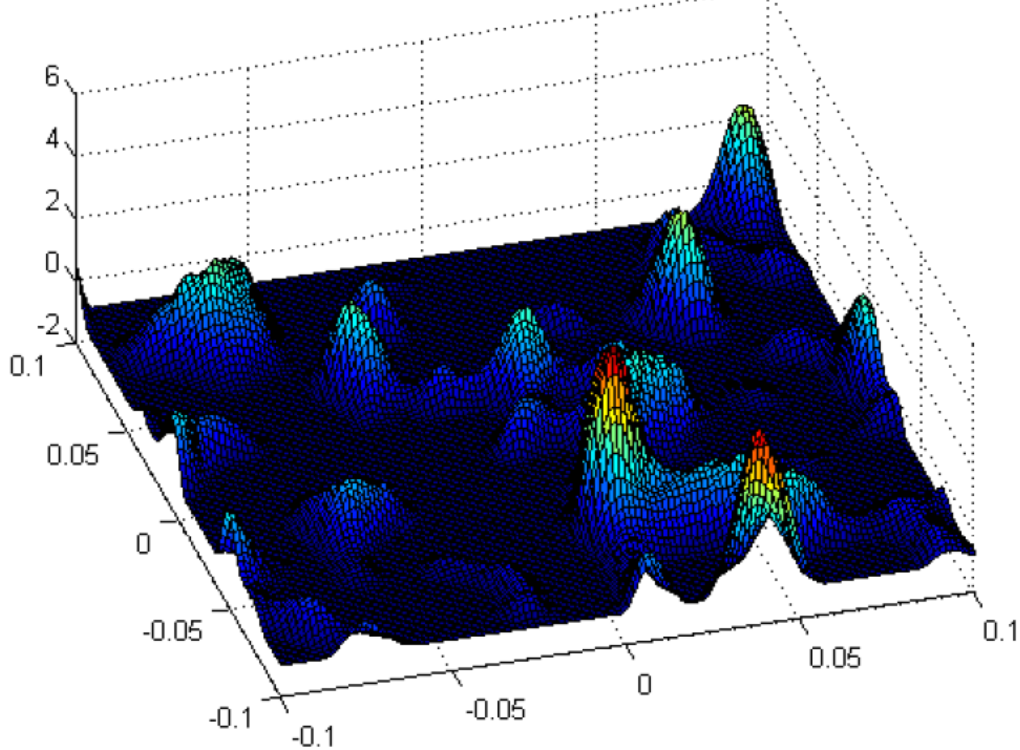
$$\mathcal{L}(\vec{x}_s, n_s) = \prod_i \left(\frac{n_s}{N} f_s(x_i) + (1 - \frac{n_s}{N}) f_b(x_i) \right)$$

$$\vec{x}_s = (\theta, \varphi)$$

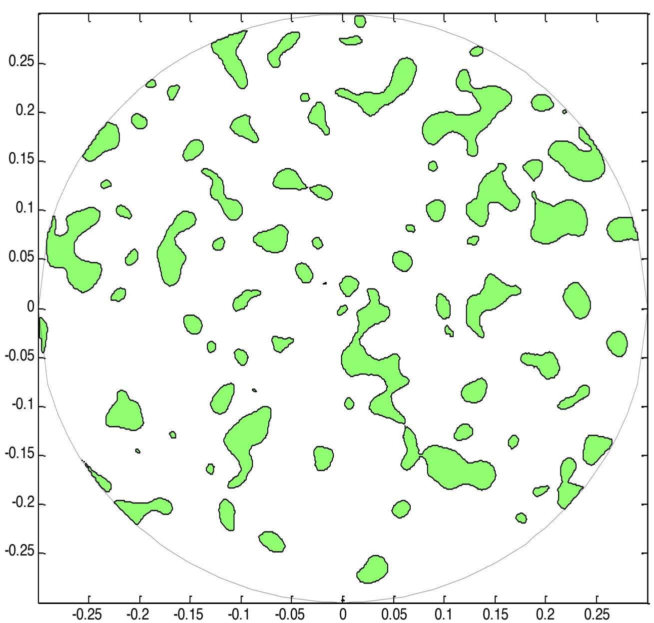
Signal parameters can also include energy and time, not considered here. Make it a multidomensional >2D manifold

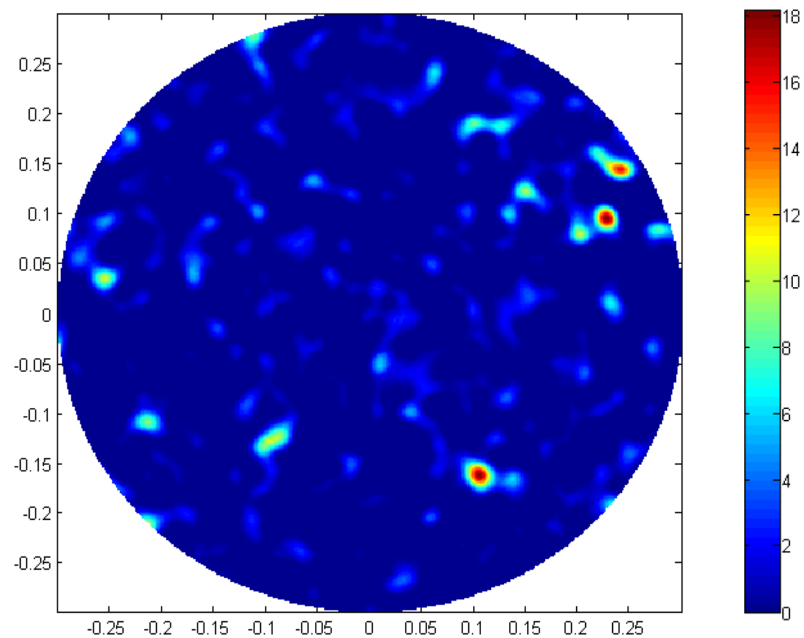


J. Braun, J. Dumm, F. De Palma, C. Finley, A. Karle, and T. Montaruli, Astropart. Phys. 29, 299 (2008); [arXiv:0801.1604]

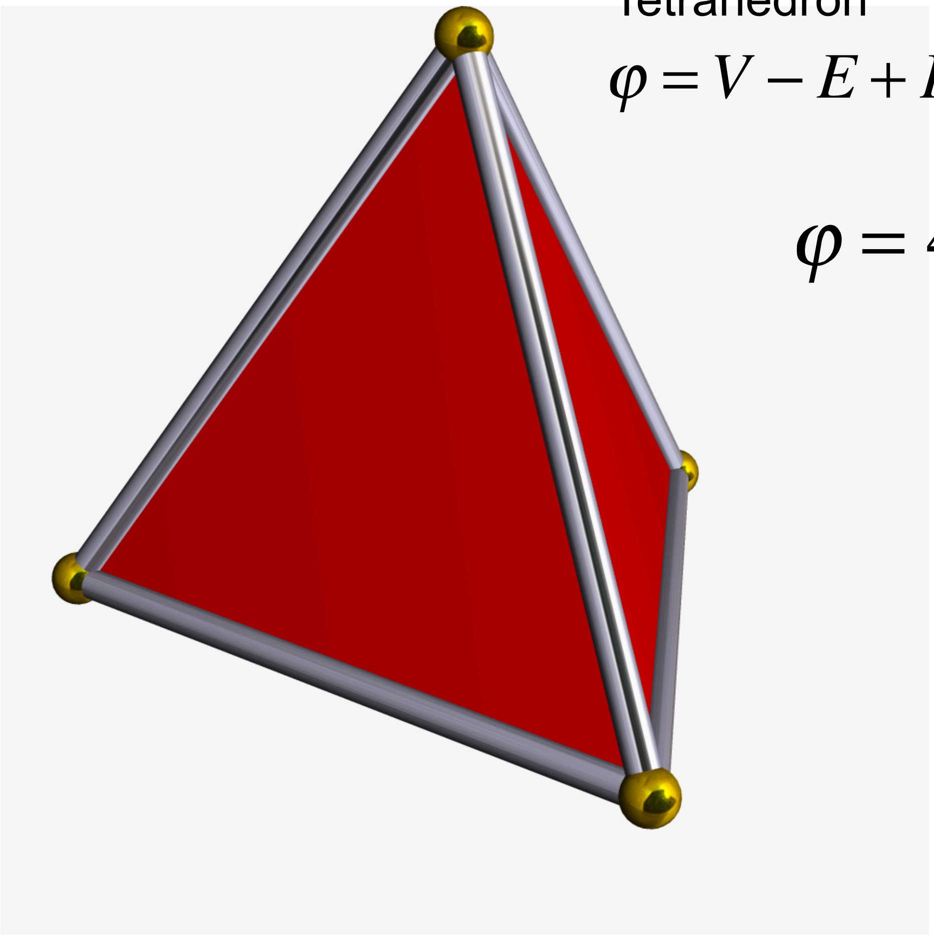


Excursion set
(u=1)





Calculation of the Euler characteristic

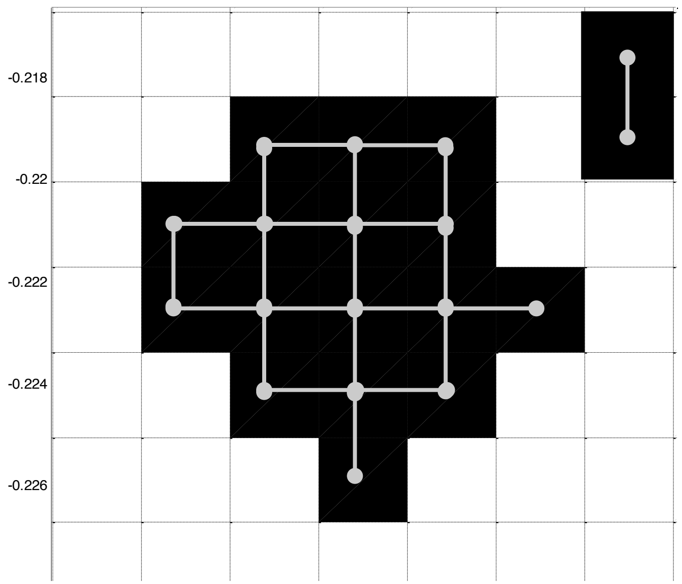


Tetrahedron

$$\varphi = V - E + F = \#vertices - \#edges + \#faces$$

$$\varphi = 4 - 6 + 4 = 2$$

Calculation of the Euler characteristic



- Usually we have $q(\theta)$ calculated on a grid of points
- Calculation of the E.C. is straightforward:
- φ = #verices - #edges + #faces
- Generalizes to higher dimensions

φ = 18(points) – 23(edges) + 7(faces)

= 2

2-d example: search for neutrino sources (IceCube)

For a chi² field in 2 dimensions:

$$E[\varphi(A_u)] = \frac{1}{2}\mathrm{P}(\chi^2 > u) + (\mathcal{N}_1 + \mathcal{N}_2\sqrt{u})e^{-u/2}$$

Estimate E[φ] at two levels, e.g. 0 and 1, and solve for $\mathcal{N}_1$ and $\mathcal{N}_2$

From 20 bkg. Simulations:

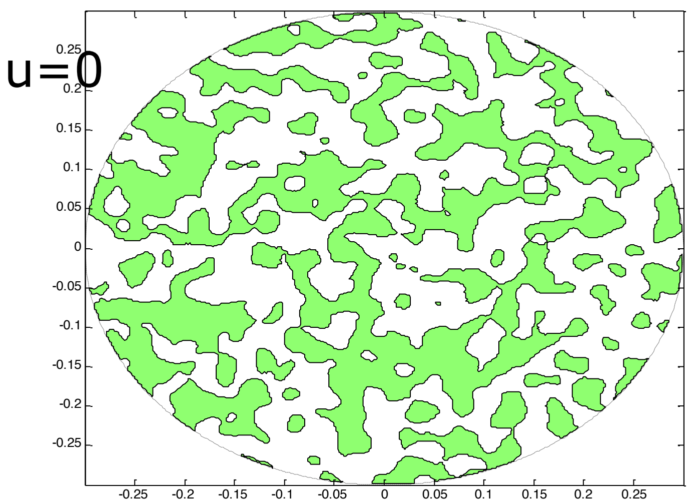
$$\langle \varphi_0 \rangle = 33.5 \pm 2$$

$$\langle \varphi_1 \rangle = 94.6 \pm 1.3$$

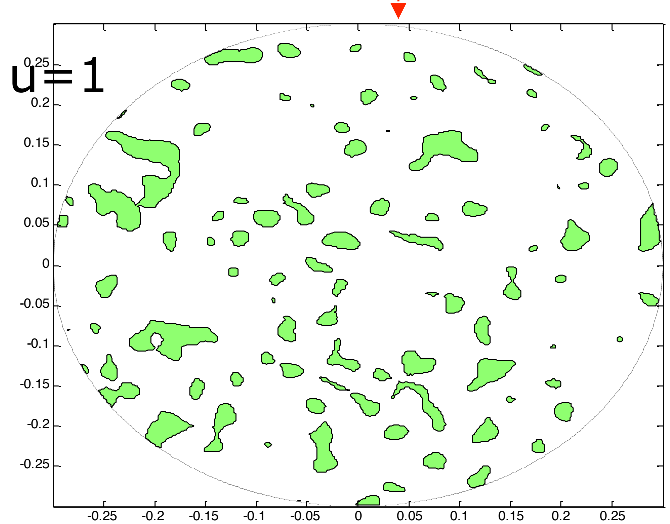
$$\Downarrow$$

$$\mathcal{N}_1 = 33 \pm 2$$

$$\mathcal{N}_2 = 123 \pm 3$$



φ=35

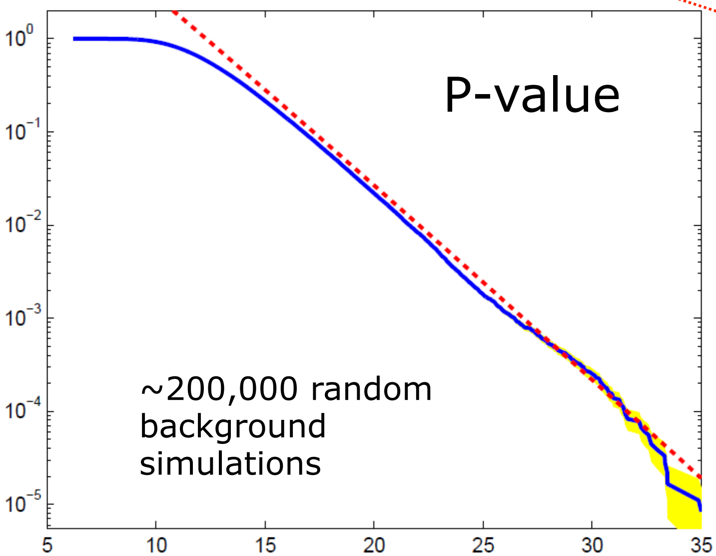


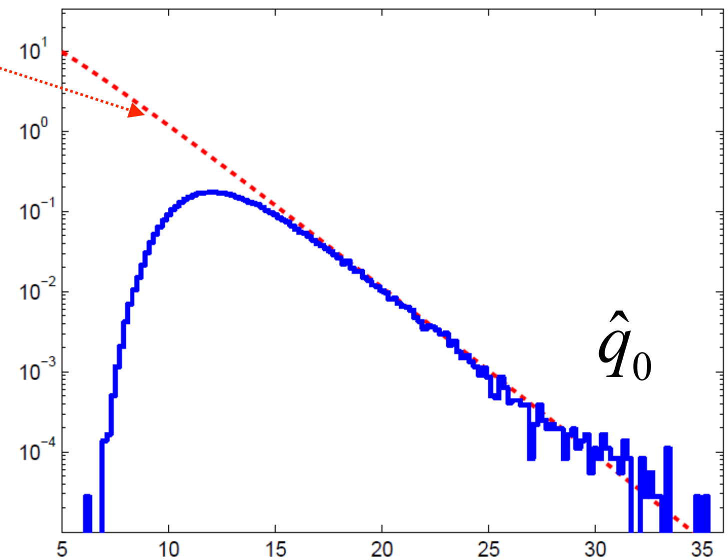
φ=95

2-d example: search for neutrino sources (IceCube)

$$E[\varphi(A_u)] = \frac{1}{2}\mathrm{P}(\chi^2 > u) + (\mathcal{N}_1 + \mathcal{N}_2\sqrt{u})e^{-u/2}$$

$$\mathcal{N}_1 = 33 \pm 2$$
$$\mathcal{N}_2 = 123 \pm 3$$

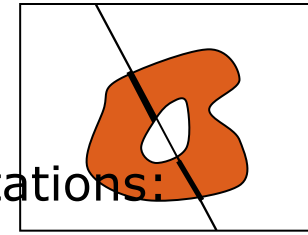




e.g.: P(max q_0>30) = $(2.5 \pm 0.4)\mathrm{x}10^{-4}$ (estimated)
E.C. Formula : $(2.28 \pm 0.06)\mathrm{x}10^{-4}$

Slicing

- Exploit the azimuthal angle symmetry to reduce computations:



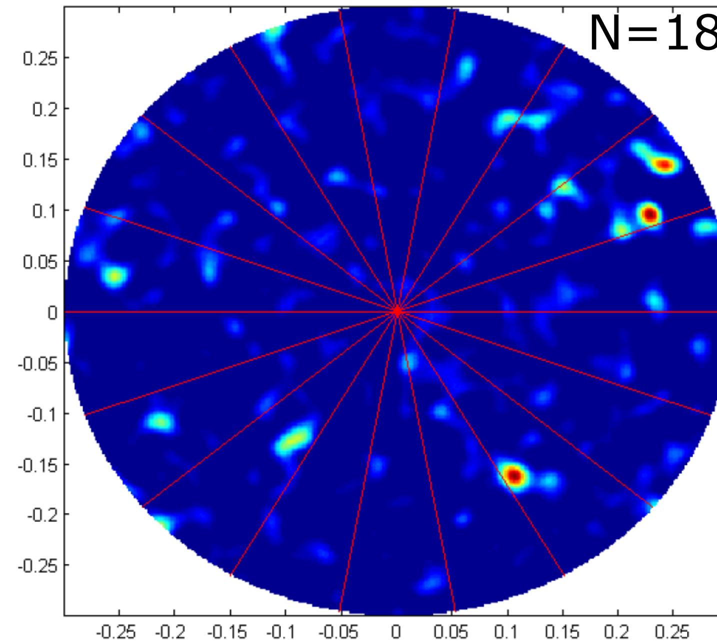
φ=0=1+1-2

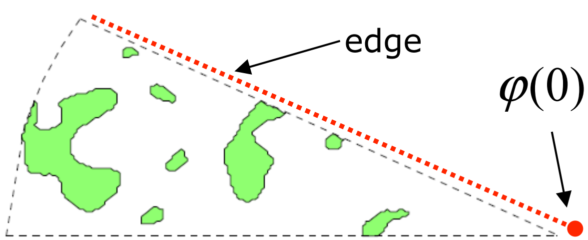
$$\boxed{\varphi(A\cup B)=\varphi(A)+\varphi(B)-\varphi(A\cap B)}$$

Divide to N slices:

$$\varphi=\sum_i[\varphi(\text{slice}_i)-\varphi(\text{edge}_i)]+\varphi(0)$$

$$E[\varphi]=N\times(E[\varphi(\text{slice})]-E[\varphi(\text{edge})])+\varphi(0)$$



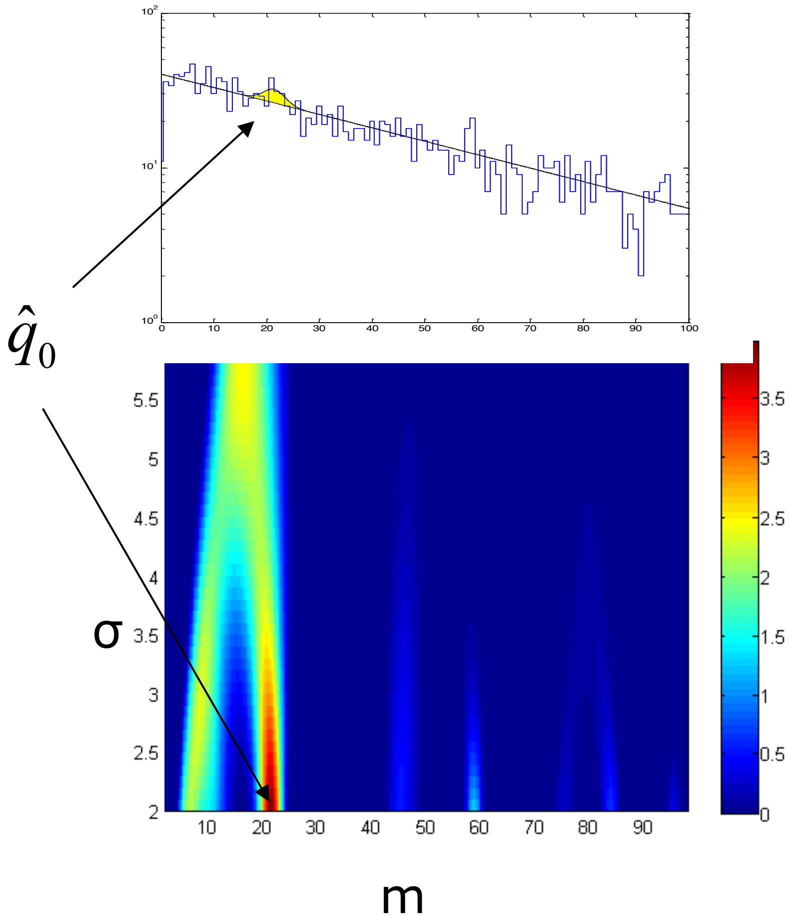


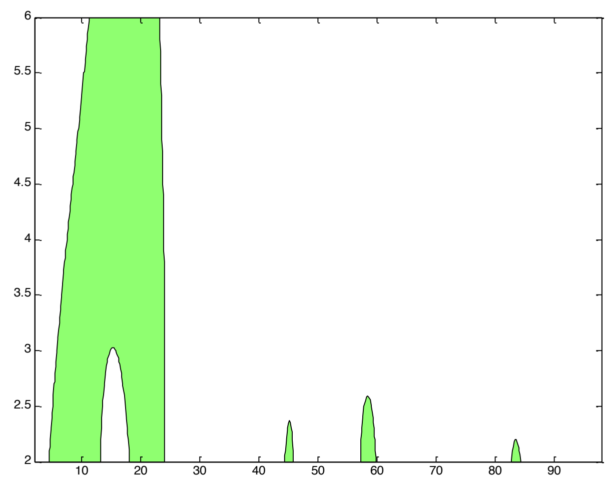
2-D exapmle #2: resonance search with unknown width

- Gaussian signal on exponential backgrou
- Toy model : 0<m<100 ,   2<σ<6
- Unbinned likelihood:

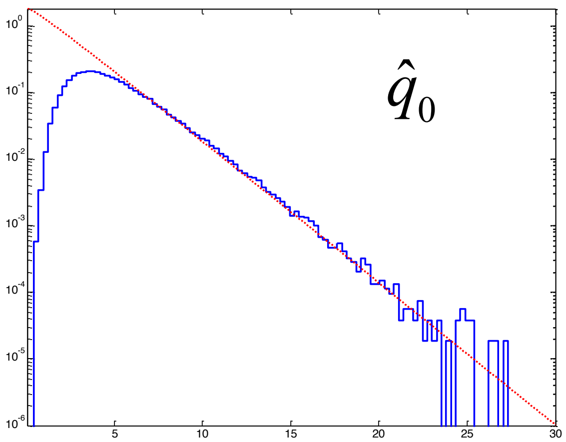
$$\mathcal{L} = \prod_i \frac{N_s f_s(x_i) + N_b f_s(x_i)}{N_s + N_b} \times Poiss(N \mid N_s + N_b)$$

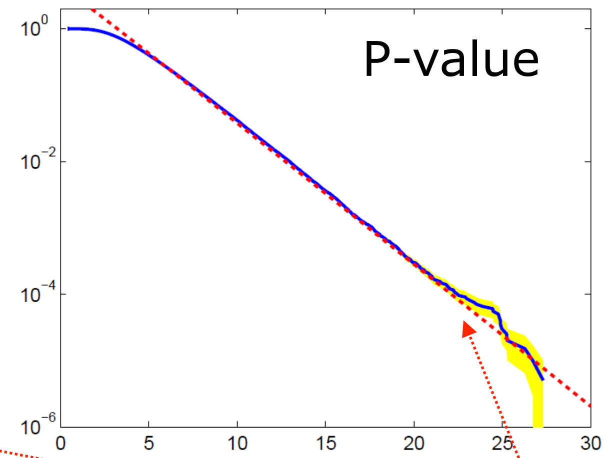
$$f_s(x; m, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-m)^2}{2\sigma^2}} \qquad f_b(x) = ce^{-cx}$$





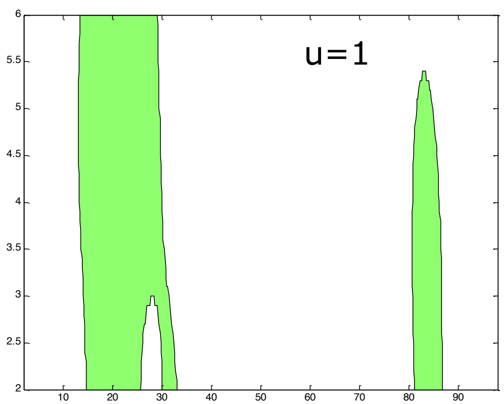
2-D exapmle #2: resonance search with unknown width



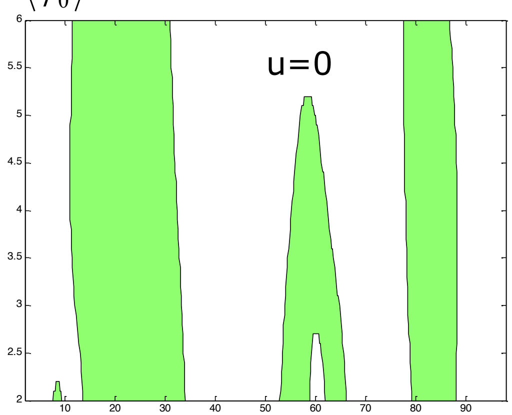


Excellent approximation above the ~2σ level

$\langle \varphi_1 \rangle = 3 \pm 0.16$



$\langle \varphi_0 \rangle = 4.5 \pm 0.2$



$$E[\varphi(A_u)] = \frac{1}{2}\mathrm{P}(\chi^2 > u) + (\mathcal{N}_1 + \mathcal{N}_2\sqrt{u})e^{-u/2}$$

$$\mathcal{N}_1 = 4 \pm 0.2$$

$$\mathcal{N}_2 = 0.7 \pm 0.3$$

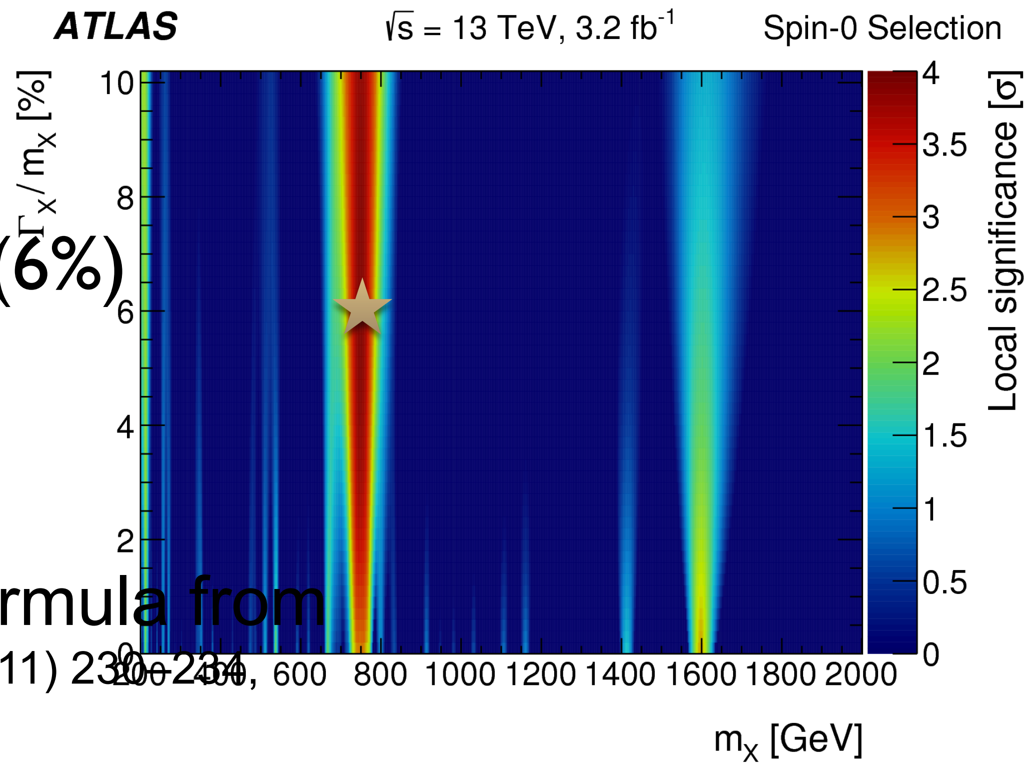
2015

2D Scan
Largest significance

$m_x \sim 750\text{GeV}, \Gamma_x \sim 45\text{GeV}(6\%)$

Local Z = **3.9σ**

m=200-2000 GeV
Γ_X/m_X=0-10%



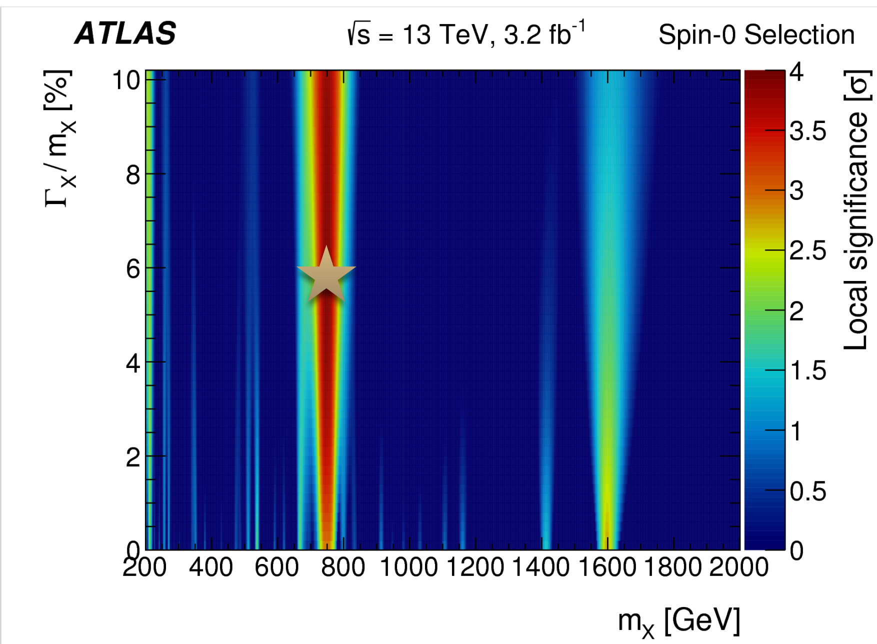
Use toys or asymptotic formula from
O. Vitells et. al. Astropart. Phys. 35 (2011) 230-234,
arXiv:1105.4355

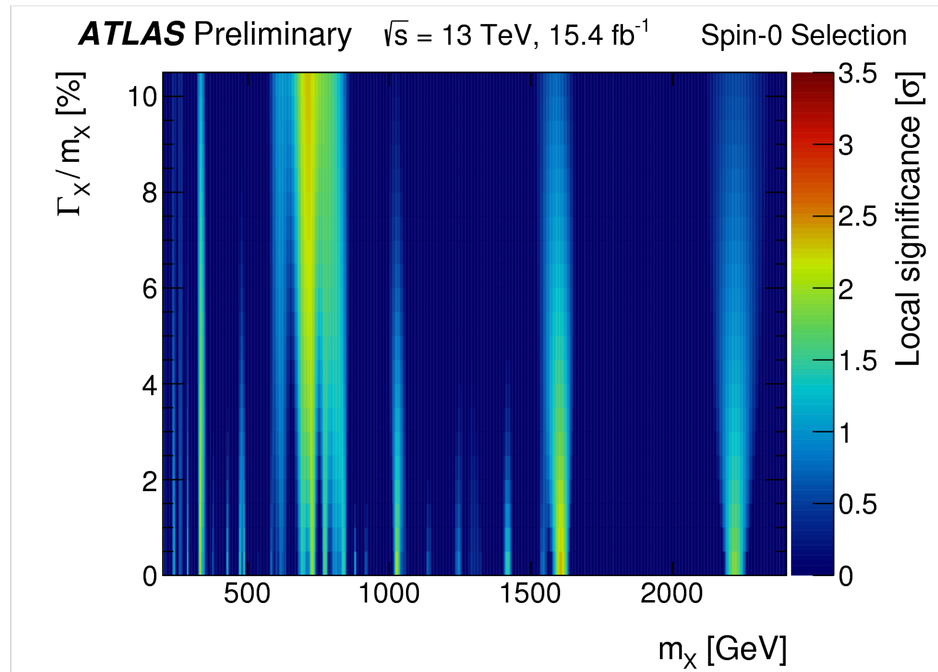
$$Z_{local} = 3.9\sigma$$

$$Z_{global} = 2.1\sigma$$

2.1σ is not something to write home about

And indeed,local 3.9σ turns into a global 2.1σ which is not something to write home about





Summary

p-value ≈

$$E[\varphi(A_u)] = \frac{1}{2}\mathrm{P}(\chi^2 > u) + (\mathcal{N}_1 + \mathcal{N}_2\sqrt{u})e^{-u/2} + \ldots$$

- The Euler characteristic formula provides a practical way of estimating the look-elsewhere effect.
- Applicable in wide range of applications, such as astrophysical searches for neutrino sources or resonance search with unknown width, and in any number of search dimensions.
- The procedure for estimating the p-value is simple and reliable.