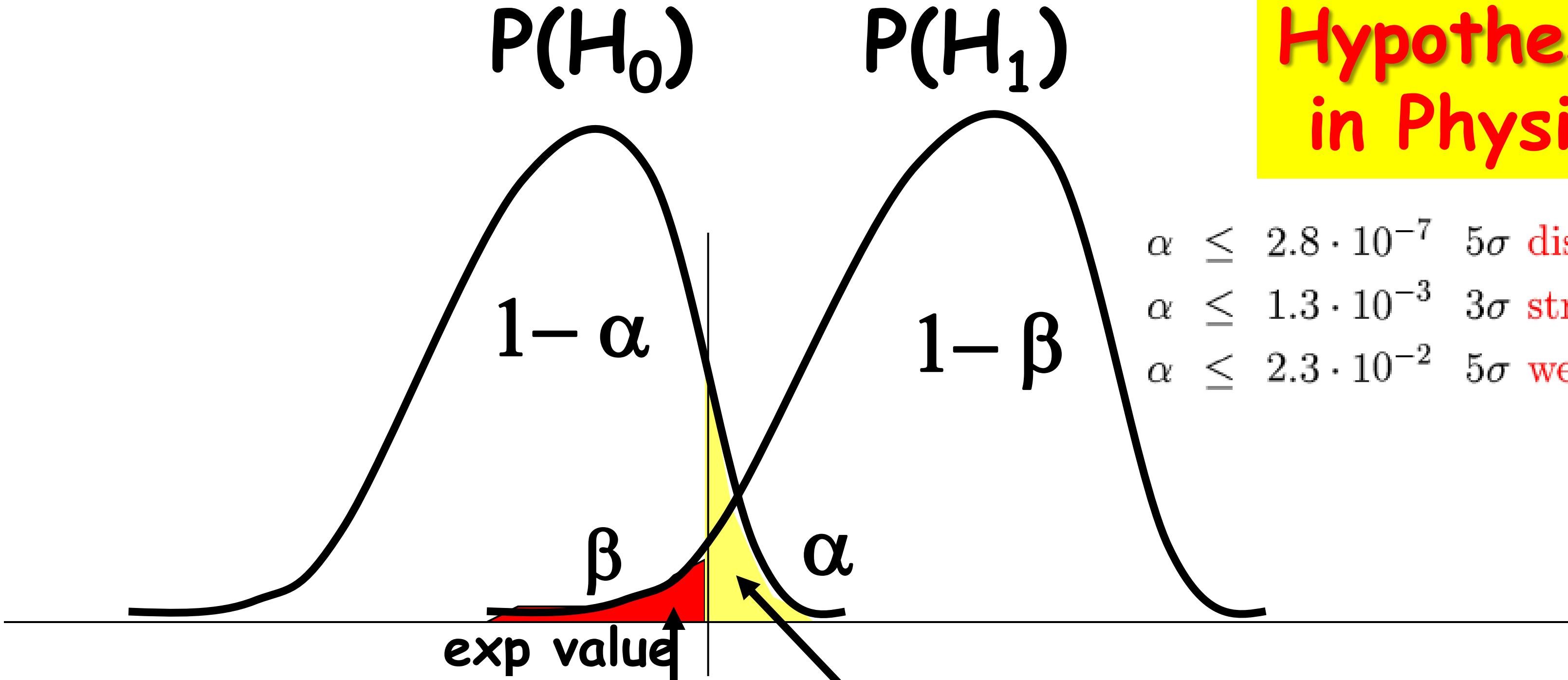
Data Analysis

- statistics
- efficiency
- resolution

- counting
- pile-up effects
- unfolding
- **signal to background ratio**

Alberto Rotondi,
Pavia University and
INFN Sezione di Pavia

Hypothesis test in Physics



$\alpha \leq 2.8 \cdot 10^{-7}$ 5σ discovery

$\alpha \leq 1.3 \cdot 10^{-3}$ 3σ strong evidence

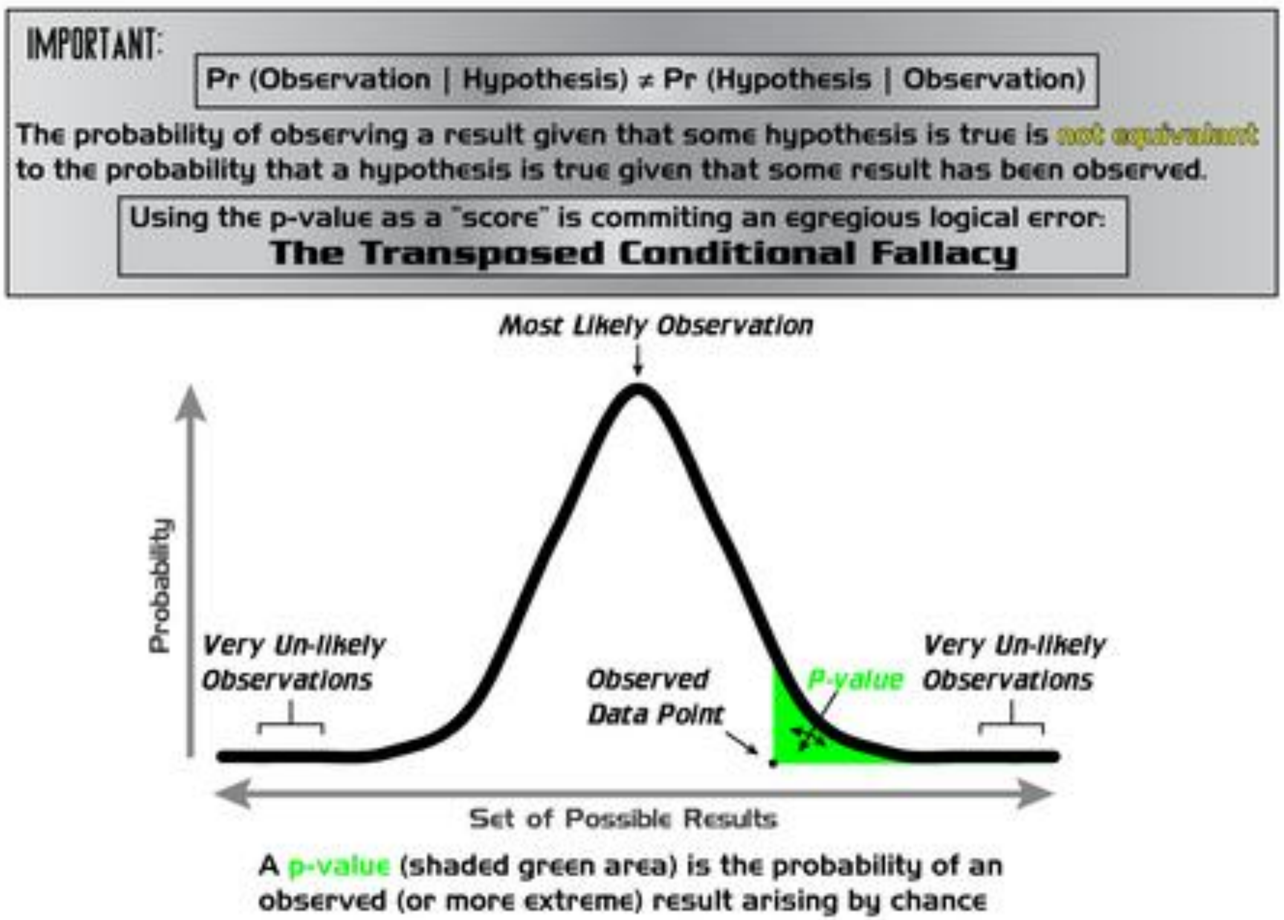
$\alpha \leq 2.3 \cdot 10^{-2}$ 5σ weak evidence

true hypothesis	Decision	
	H_0	H_1
H_0 no effect	correct decision $1-\alpha$ good rejection	type I error α contamination
H_1 effect	type II error β event loss	correct decision $1-\beta$ good acceptance

Power (discovery Potential)

If H_1 is the discovery, the maximum power test maximizes the discovery probability, that is the good acceptance

(...from wikipedia)



Data yielding a p-value of .05 means there is only a 5% chance obtaining the observed (or more extreme) result if no real effect exists. "Common sense" tells us to judge our hypotheses based on how well they fit observed evidence. This is not what a p-value describes. Instead, it describes the likelihood of observing certain data given that the null hypothesis is true.

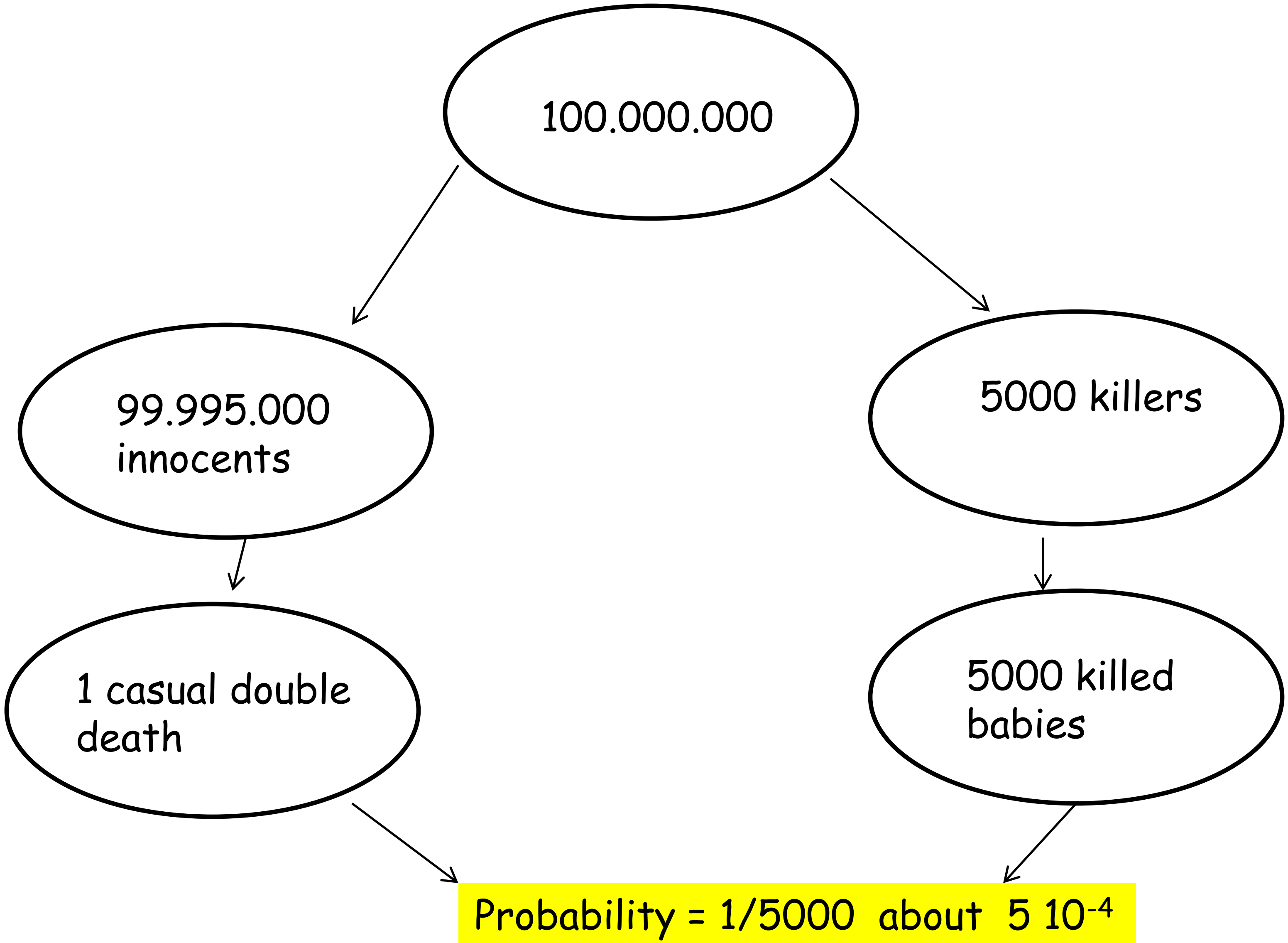
The prosecutor fallacy
(see the Sally Clark case)

- The probability for an innocent mother to have 2 baby sudden deaths is $P(2D|I)=10^{-8}$.
- If $P(2D|I)=P(I|2D)$ $=10^{-8}$, the mother is guilty!!

ERROR !

- The right response is given by the Bayes theorem

$$P(I\,|\,2D)=\frac{P(2D\,|\,I)P(I)}{P(2D\,|\,I)P(I)+P(2D\,|\,G)P(G)}=\begin{cases}1 \text{ if } \ P(G)=0\\ 10^{-3} \text{ if } P(G)=1\text{-}P(I)\\ \qquad\approx 30/640\,000\end{cases}$$


100.000.000
99.995.000 innocents
5000 killers
1 casual double death
5000 killed babies
Probability = 1/5000 about 5 10-4

SALFORD ADVERTISER THURSDAY JUNE 13, 2002

SALFORD UNI MAN SAYS SALLY CLARK CONVICTION MAY BE WRONG

Maths professor challenges double baby murder case

A SALFORD University Maths professor will challenge evidence used to convict a solicitor of murdering her two baby sons at a conference on cot-deaths next week.

Prof Ray Hill, from Eccles, head of the university's Applied and Discrete Mathematics Research Unit said statistical evidence used to convict Sally Clark, from Wilmslow, in October 2000, was not only quoted out of context and unfairly used to imply guilt, but was actually wrong.

Watching the trial on the TV he became furious and told us: "I shouted at the screen 'that figure's wrong!' They took an estimated figure for the likelihood of one cot death and then just squared it to get this one-in-73 million chance. That's not allowed unless you're sure the events are independent. A bookie wouldn't give you those odds."

He has now studied the Confidential Enquiry into Stillbirths and Deaths in Infancy (CESDI) report, which gives detailed figures on the number of deaths from 1993-1996.

He said: "It seems the chances of two cot deaths in the same family are much higher than the prosecution led the jury to believe."

Prof Hill has written to several national newspapers and is working with Sally Clark's defence team on the campaign to free her.

He will present his full criticism of the evidence at a Developmental Physiology Conference on cot deaths organised by Leicester University on June 28.

The Criminal Cases Review Commission has been looking at the case and is expected to report within the next few weeks. With their report imminent, Sally Clark's defence team and family do not feel it is appropriate to comment.

For more information on the Sally Clark campaign visit www.sallyclark.org.uk

Evidence challenge: Prof Ray Hill (2553-S 02)

The prosecutor fallacy

- 10^6 tests are made for AIDS. Probability of a false positive is $P(P|NO)=10^{-4}$.
- If $P(P|NO)=P(NO|P)$ $=10^{-4}$, I'm sick!
- The right response is given by the Bayes theorem

$$P(NO|P)=\frac{P(P|NO)P(NO)}{P(P|NO)P(NO)+P(P|AIDS)P(AIDS)}=\begin{cases}1 \text{ if } P(AIDS)=0\\ 9\% \text{ if there are}\\ \quad 1000 \text{ AIDS cases}\end{cases}$$

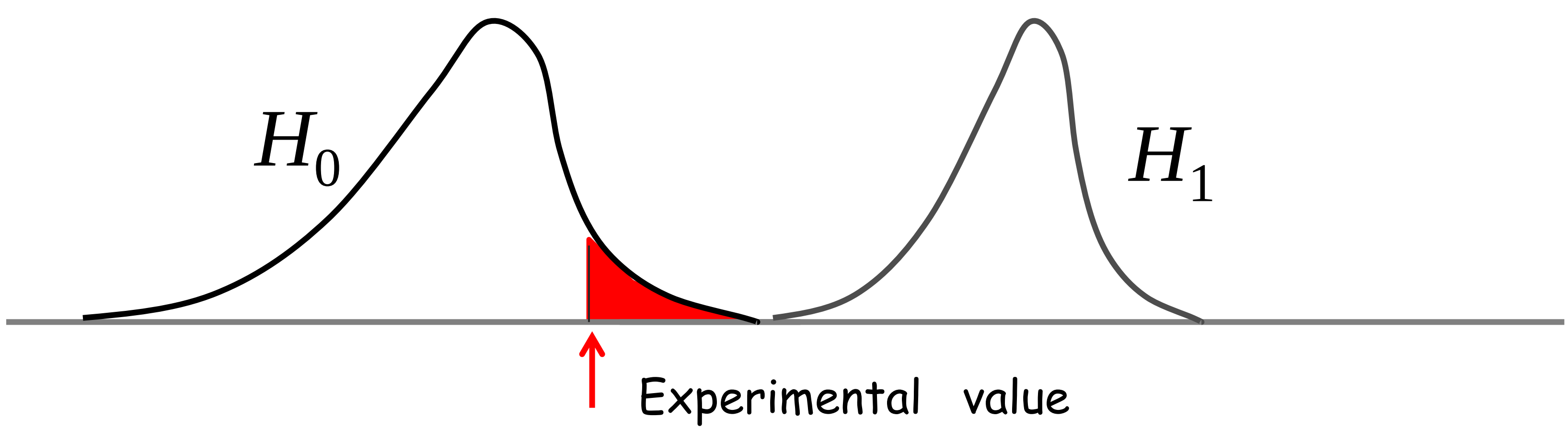
…. the *p*-value….

- is the probability of falsely rejecting the null hypothesis

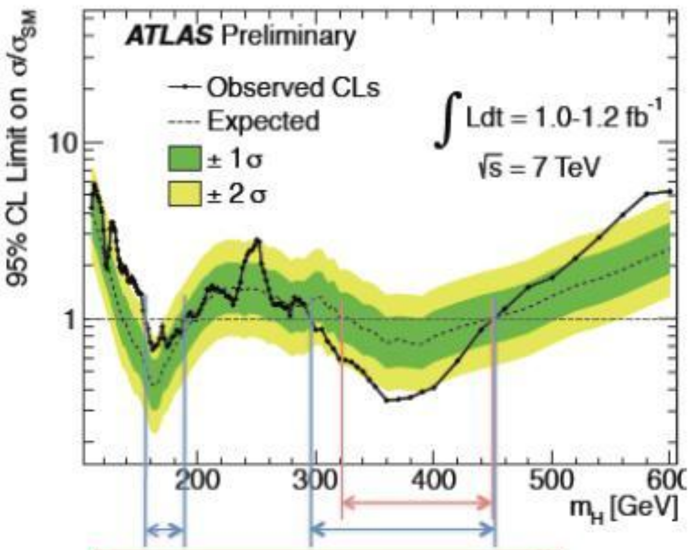
  when it is true $P(p\text{-value}|H_0)$

- is **not** the probability that the null hypothesis is true $P(H_0 \mid p\text{-value})$

$$P(H_0 \mid p-value) = \frac{P(p-value \mid H_0)P(H_0)}{P(p-value \mid H_0)P(H_0) + P(p-value \mid H_1)P(H_1)}$$



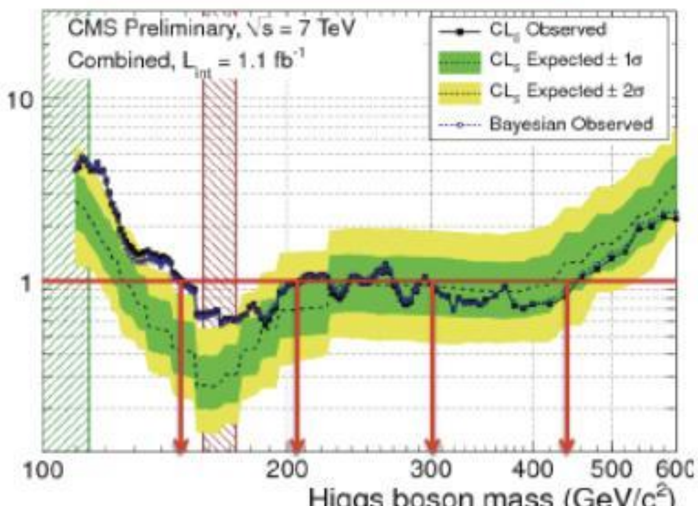
Where we are now !



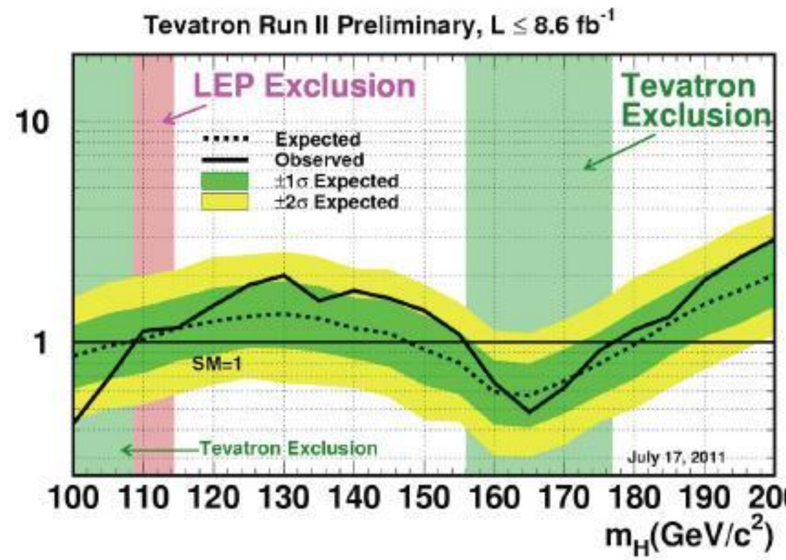
155<MH<190 and 295 <MH< 450 GeV/c²
excluded at @ 95% CL

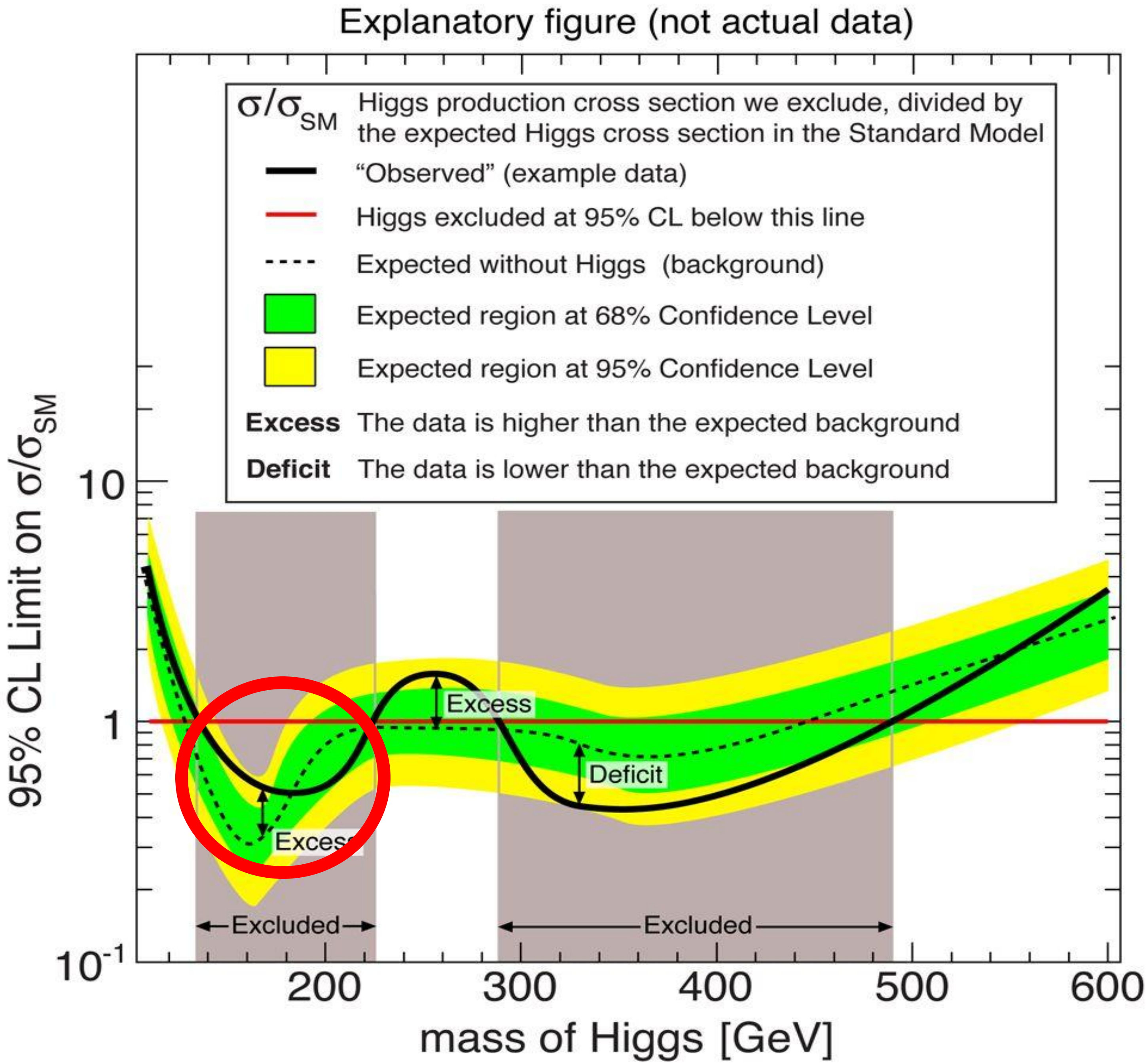
How a theorist feels…..

….like having a banquet after years of diet !



Excluded (GeV)
[149-206] … [300-440]
and 3 short segments
in between





Explanatory figure (not actual data)
σ/σSM Higgs production cross section we exclude, divided by the expected Higgs cross section in the Standard Model
"Observed" (example data)
Higgs excluded at 95% CL below this line
Expected without Higgs (background)
Expected region at 68% Confidence Level
Expected region at 95% Confidence Level
Excess The data is higher than the expected background
Deficit The data is lower than the expected background
95% CL Limit on σ/σSM
10
1
10-1
Excess
Deficit
Excess
Excluded
Excluded
200
300
400
500
600
mass of Higgs [GeV]

Figure A

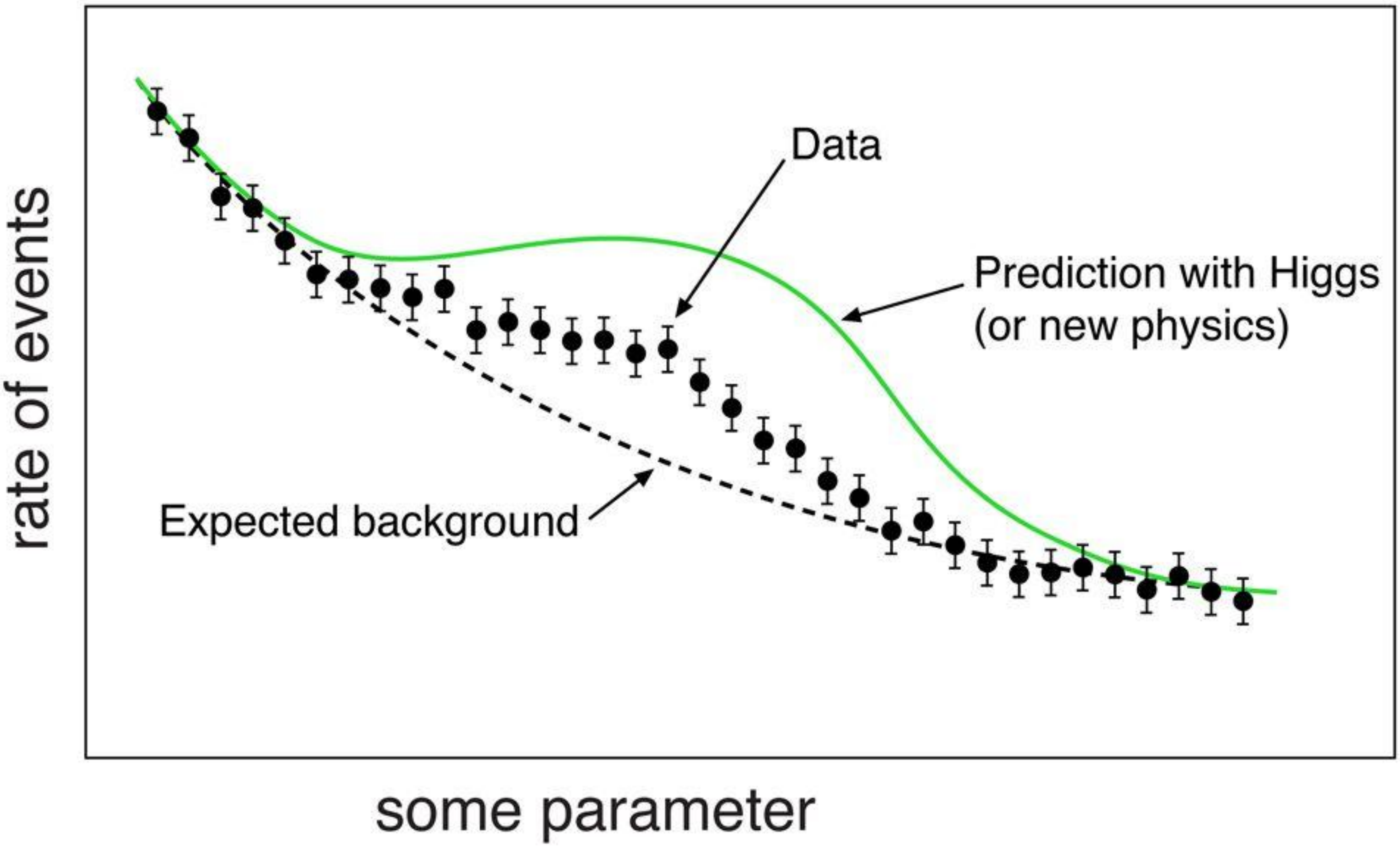
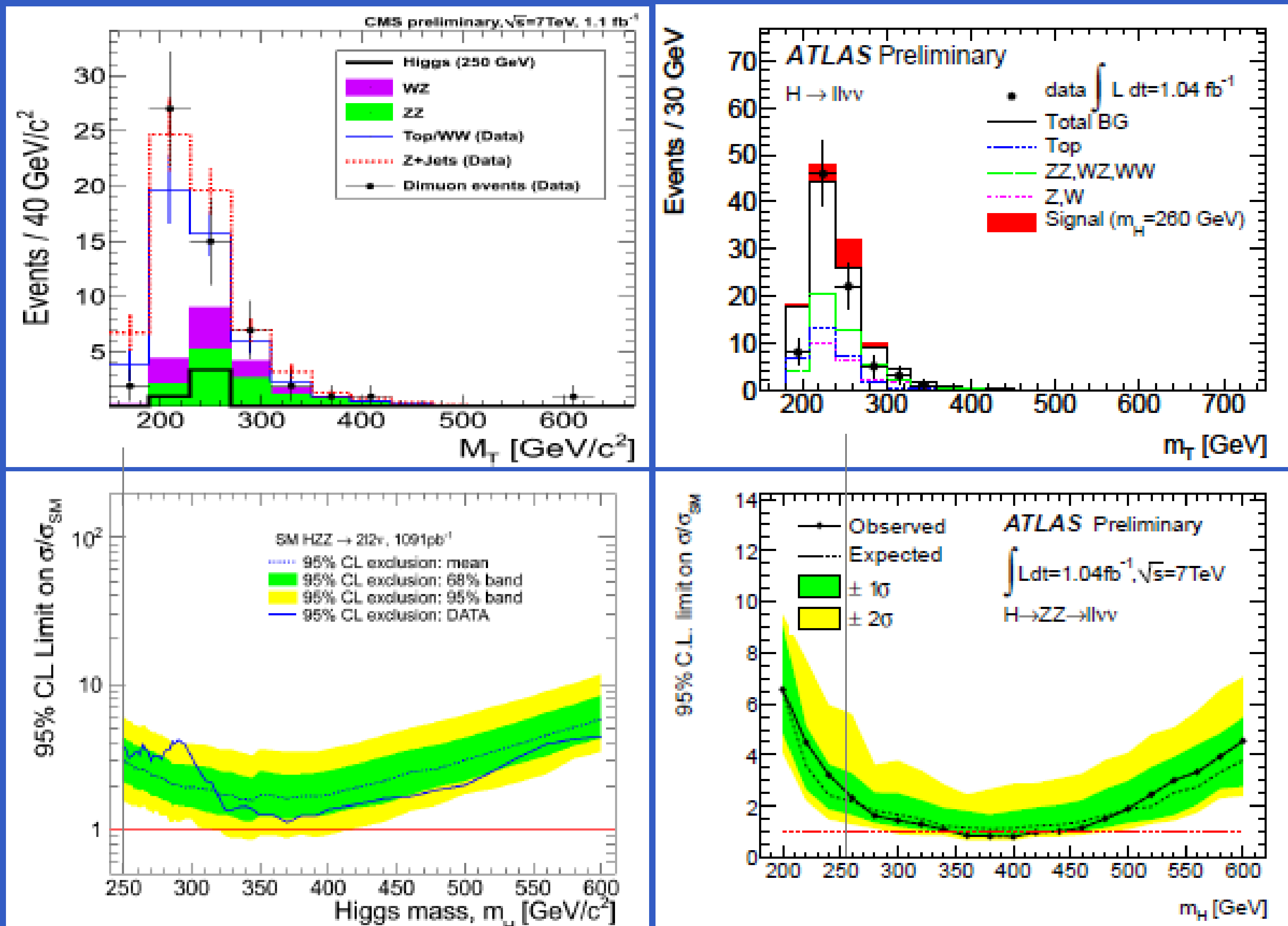
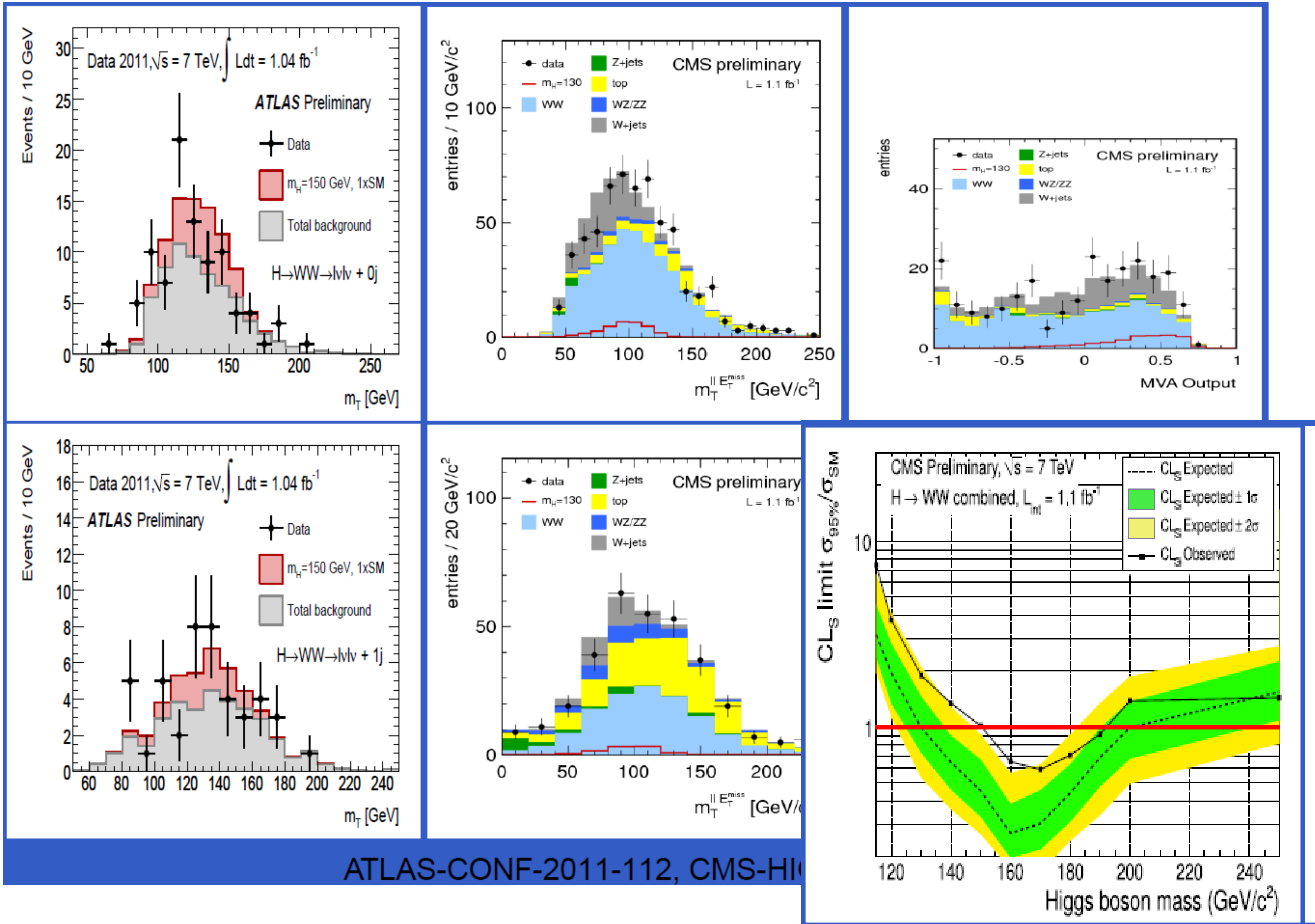

Data
Prediction with Higgs
(or new physics)
Expected background
rate of events
some parameter

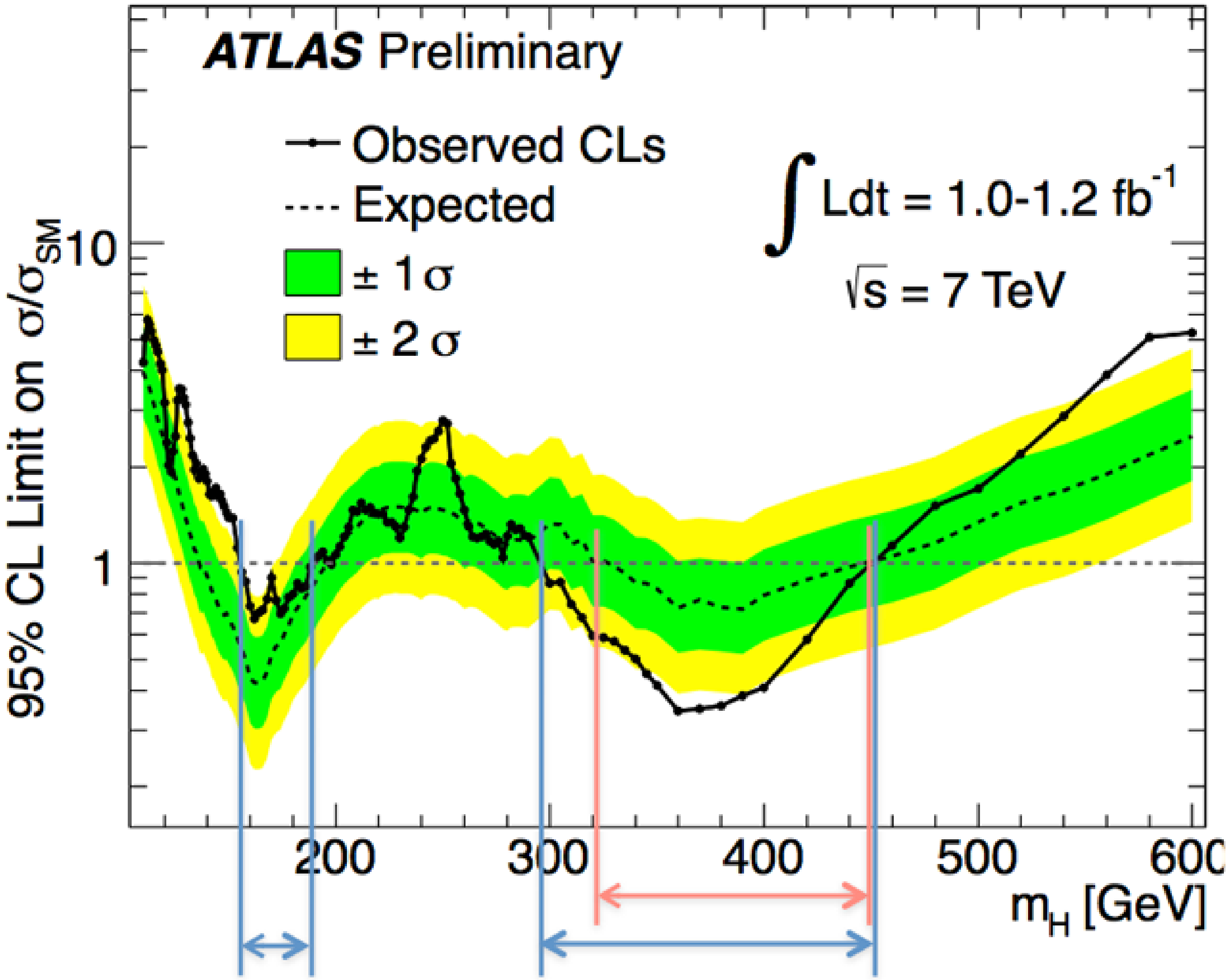
Figure B



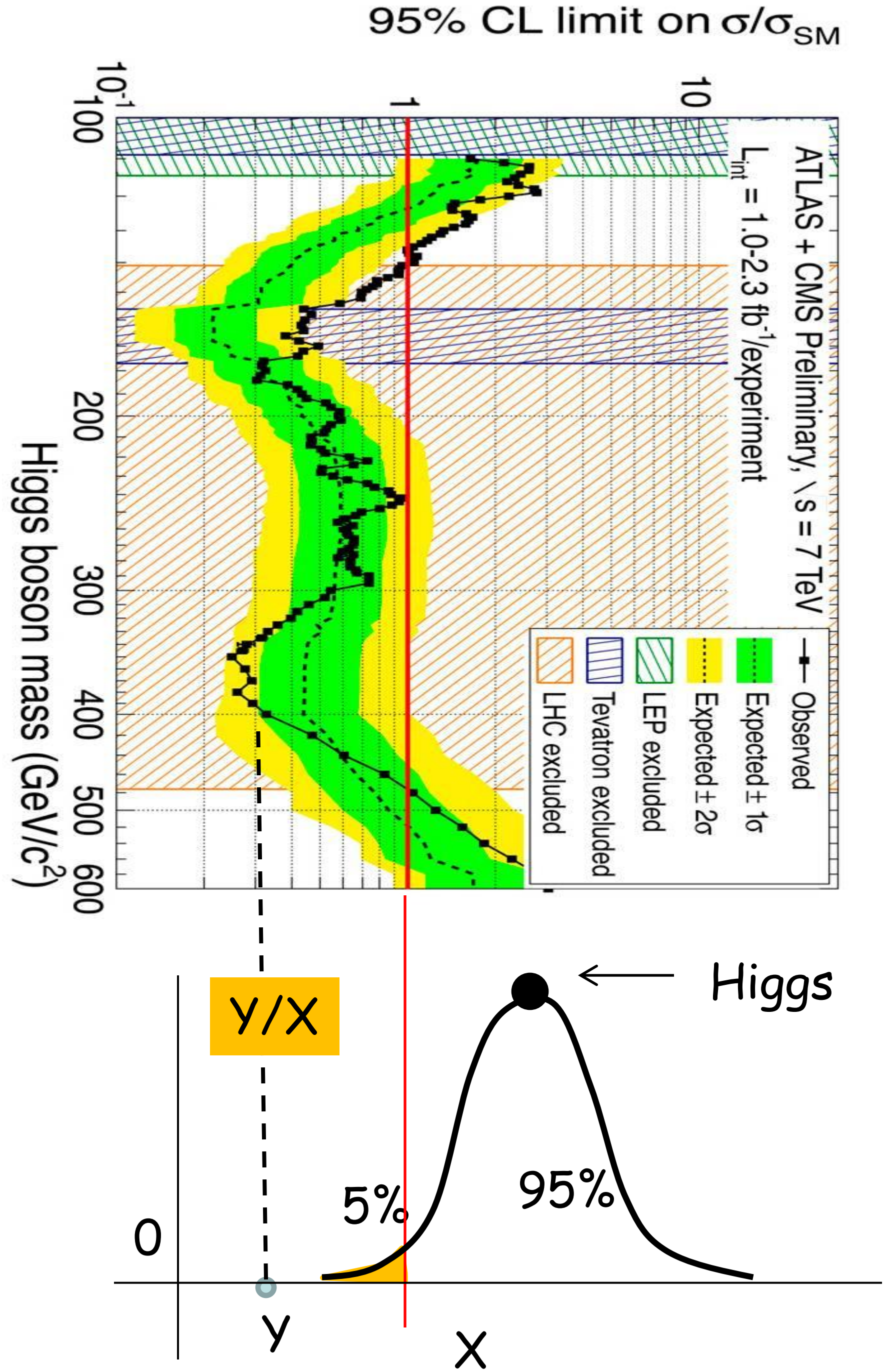
ATLAS-CONF-2011-XYZ, CMS-HIG-11-005



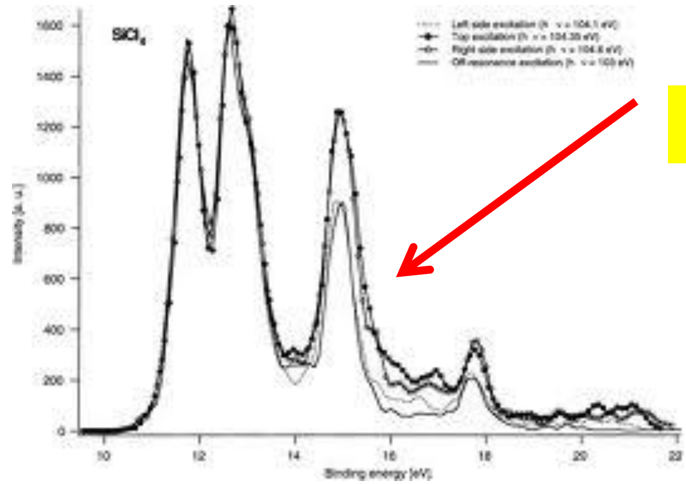
ATLAS-CONF-2011-112, CMS-HI



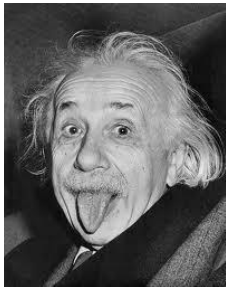
$155 < M_H < 190$ and $295 < M_H < 450$ GeV/c²
excluded at @ 95% CL


95% CL limit on σ/σ_SM
ATLAS + CMS Preliminary, √s = 7 TeV
L_int = 1.0-2.3 fb-1/experiment
Observed
Expected ± 1σ
Expected ± 2σ
LEP excluded
Tevatron excluded
LHC excluded
Higgs boson mass (GeV/c2)
Y/X
Higgs
5%
95%
0
Y
X

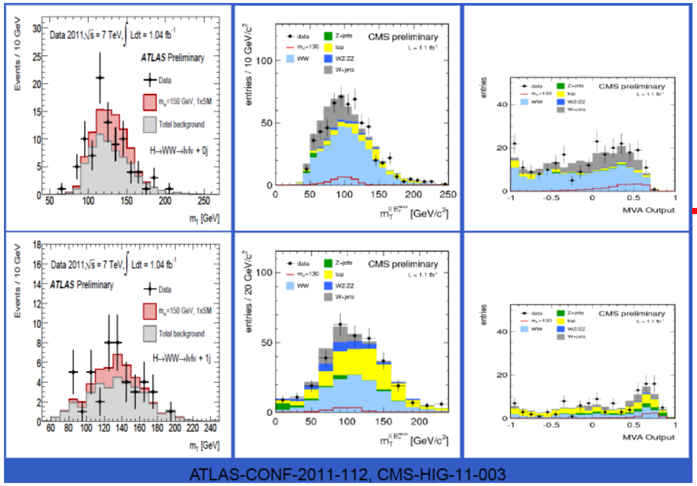
Old physics



Discovery!!

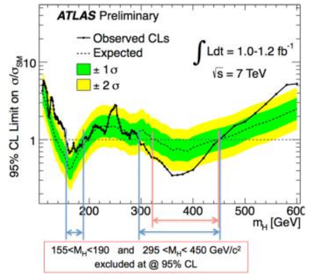


now



Statistical committee

Discovery!!



The Significance of a signal

The right notation

- A random variable before the experiment: X
- A random variable after the experiment: x
- True mean μ
- True sigma: σ
- Measured mean: m
- Measured sigma: s
- True probability: p
- Measured Frequency: f

$$F(X) = \int_0^X p(x)\,\mathrm{d}x \sim U(0,1)$$

$$g(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$g(X) = \frac{1}{\sqrt{2\pi}} e^{-X^2/2}$$

The case of Pentaquark

The pentaquark is a baryon with five valence quarks.
The clearest signature is that of a

$$u\,u\,d\,d\,\bar{s}\ , \qquad S = +1$$

pentaquark, the unique baryon with positive strangeness.

The $\bar{s}$ antiquark cannot annihilate with the u or d quark by the strong interaction.
Some models predict a mass around 1.5 GeV and a very small width ($\simeq 0.015$ GeV)

The recent pentaquark saga began at 2002 PANIC conference when Nakano measured the following reaction on a Carbon nucleus

$$\gamma n \to \Theta^+ K^- \to K^+ K^- n$$

$\gamma p \rightarrow K^+ \Lambda(1250) \rightarrow K^+ K^- p$

$\gamma n \rightarrow K^- \Theta \rightarrow K^- K^+ n$

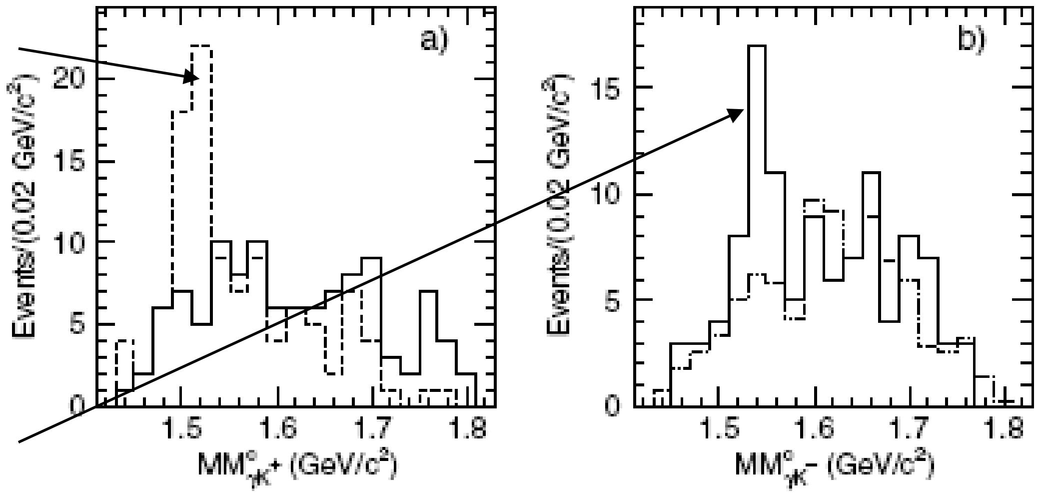


FIG. 3. (a) The $MM^c_{\gamma K^+}$ spectrum [Eq. (2)] for K^+K^- productions for the signal sample (solid histogram) and for events from the SC with a proton hit in the SSD (dashed histogram). (b) The $MM^c_{\gamma K^-}$ spectrum for the signal sample (solid histogram) and for events from the LH_2 (dotted histogram) normalized by a fit in the region above 1.59 GeV/c^2.

The first result

PRL 91(2003)012002

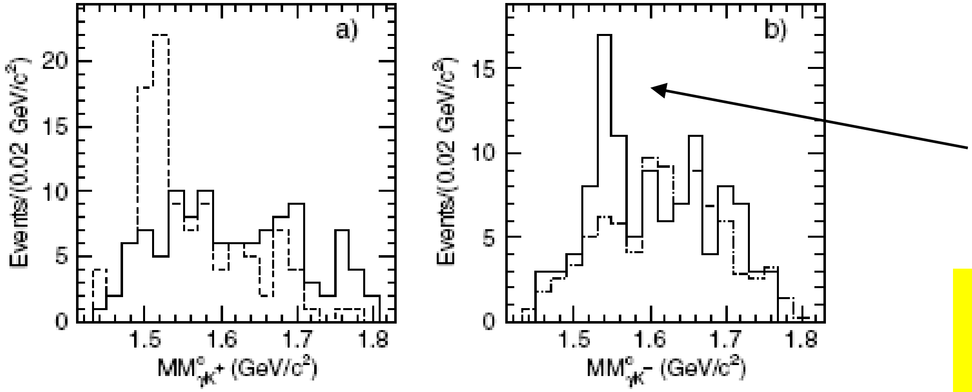
$$\left(\sum_{in} E_{in} - \sum_{fin} E_{fin}\right)^2 - \left(\sum_{in} \vec{p}_{in} - \sum_{fin} \vec{p}_{fin}\right)^2$$

The neutron presence was detected by the $MM_{\gamma K^+K^-}$ missing mass

The $\gamma p \rightarrow K^+K^-p$ reaction was eliminated by direct proton detection.

The neutron was reconstructed from the missing momentum and energy of K^+ and K^-.

The background was measured from a LH_2 target.



4.6 sigma!

Is it convincing???

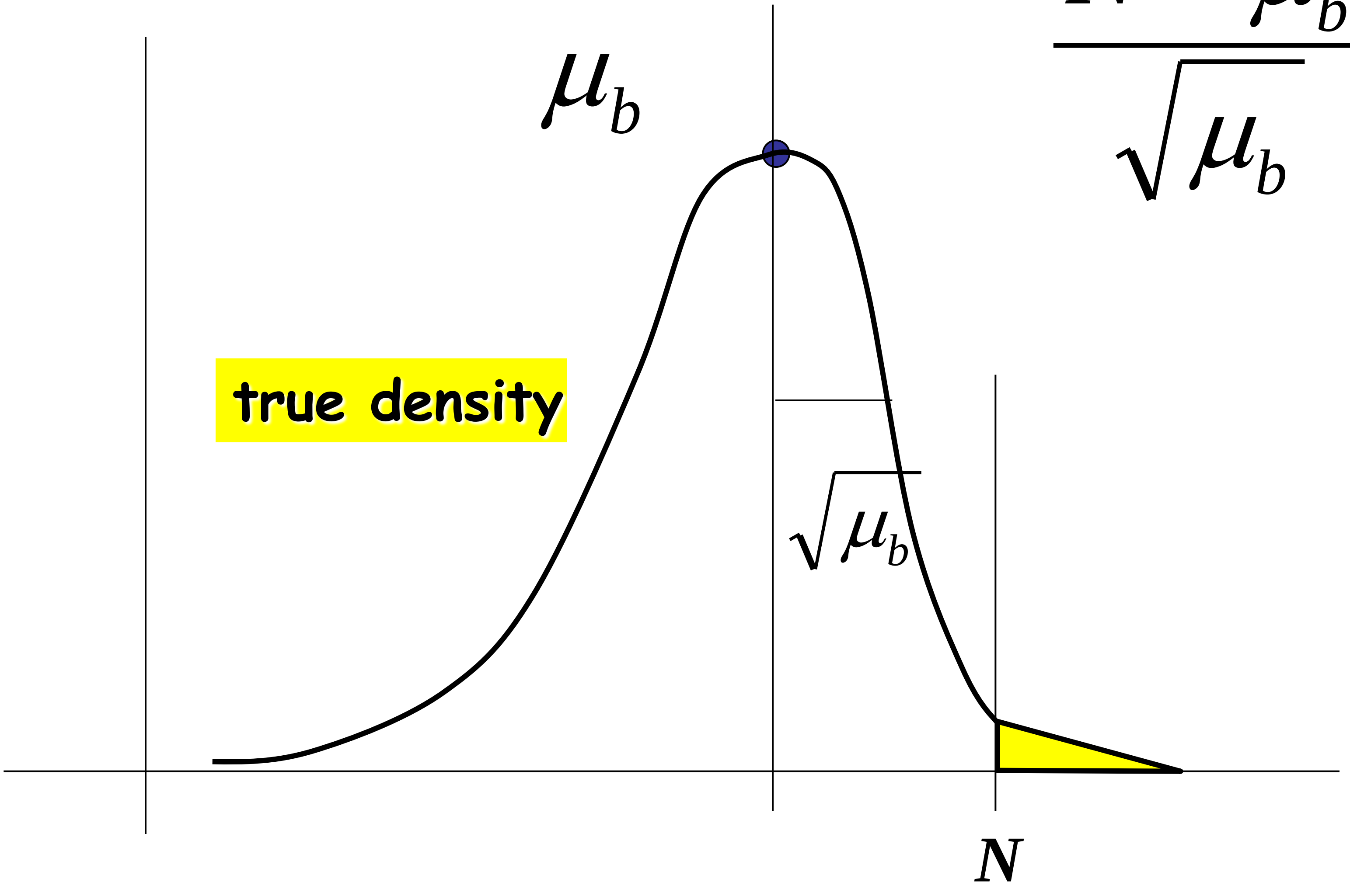
FIG. 3. (a) The $MM^c_{\gamma K^+}$ spectrum [Eq. (2)] for K^+K^- productions for the signal sample (solid histogram) and for events from the SC with a proton hit in the SSD (dashed histogram). (b) The $MM^c_{\gamma K^-}$ spectrum for the signal sample (solid histogram) and for events from the LH_2 (dotted histogram) normalized by a fit in the region above 1.59 GeV/c^2.

012002-3

The background level in the peak region is estimated to be $17.0 \pm 2.2 \pm 1.8$, where the first uncertainty is the error in the fitting in the region above 1.59 GeV/c^2 and the second is a statistical uncertainty in the peak region. The combined uncertainty of the background level is ± 2.8. The estimated number of the events above the background level is 19.0 ± 2.8, which corresponds to a Gaussian significance of $4.6^{+1.2}_{-1.0}\sigma$ $(19.0/\sqrt{17.0} = 4.6)$.

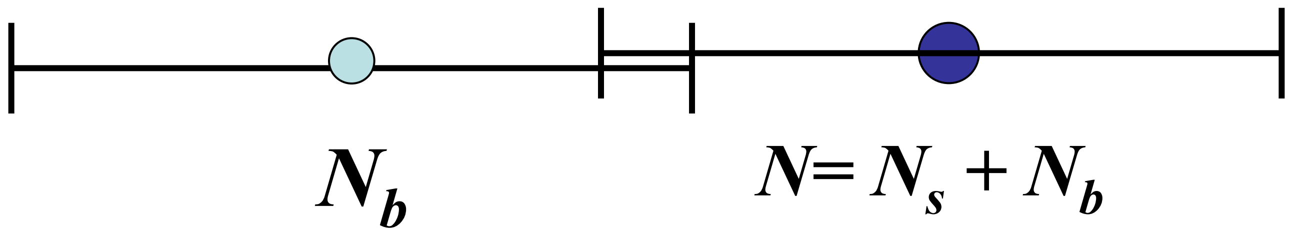
Hypothesis test I

$$\frac{N-\mu_b}{\sqrt{\mu_b}}$$



Parameter estimation

$$N - N_b \pm \sqrt{N \ + N_b} \cong N_s \pm \sqrt{N_s + 2N_b}$$



$$\frac{N_s}{\sqrt{N_s + 2N_b}}$$

“standard” significances

$$\frac{N-N_b}{\sqrt{N_b}} = \frac{N_s}{\sqrt{N_b}}$$

WRONG
The true formula is

$$\frac{N-\mu_b}{\sqrt{\mu_b}}$$

APPROXIMATE
The true formulae are

$$\frac{N-N_b}{\sqrt{N+N_b}} = \frac{N_s}{\sqrt{N_s+2N_b}}$$

$$\frac{N-N_b}{\sqrt{\mu+\mu_b}} = \frac{N_s}{\sqrt{\mu_s+2\mu_b}}$$

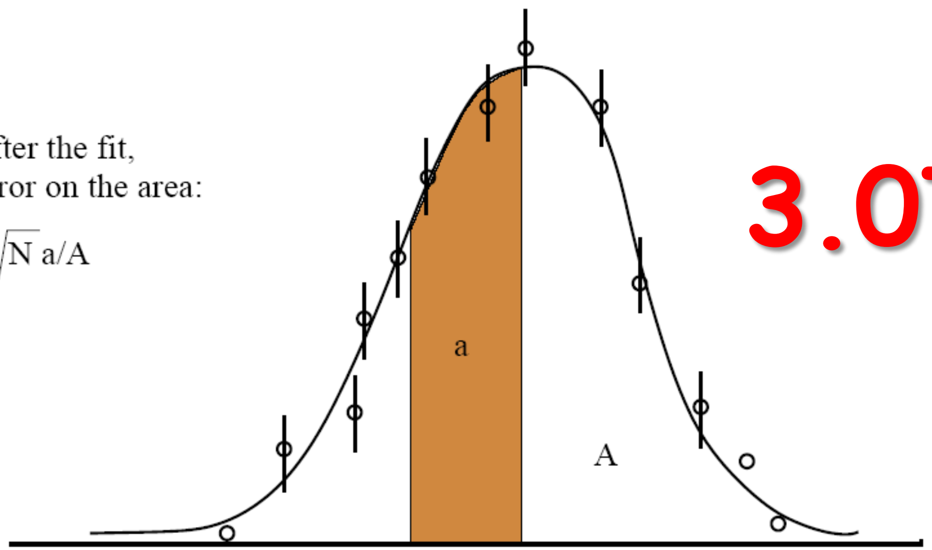
$$\frac{N-N_b}{\sqrt{N+N_b}} \xrightarrow{H_0 : N \approx N_B} = \frac{N_s}{\sqrt{2N_b}}$$

$$\frac{N-N_b}{\sqrt{\mu+\mu_b}} \xrightarrow{H_0 : N \approx N_b} = \frac{N_s}{\sqrt{2\mu_b}}$$

PRL 91(2003)012002

The background level in the peak region is estimated to be $17.0 \pm 2.2 \pm 1.8$, where the first uncertainty is the error in the fitting in the region above 1.59 GeV/c^2 and the second is a statistical uncertainty in the peak region. The combined uncertainty of the background level is ± 2.8. The estimated number of the events above the background level is 19.0 ± 2.8, which corresponds to a Gaussian significance of $4.6^{+1.2}_{-1.0}\sigma$ $(19.0/\sqrt{17.0} = 4.6)$.

after the fit,
error on the area:
$\sqrt{N}\, a/A$



3.07

$$\frac{19}{\sqrt{19+17+17}} = 2.6$$

$$\frac{19}{\sqrt{17+2.8^2}} = 3.8$$

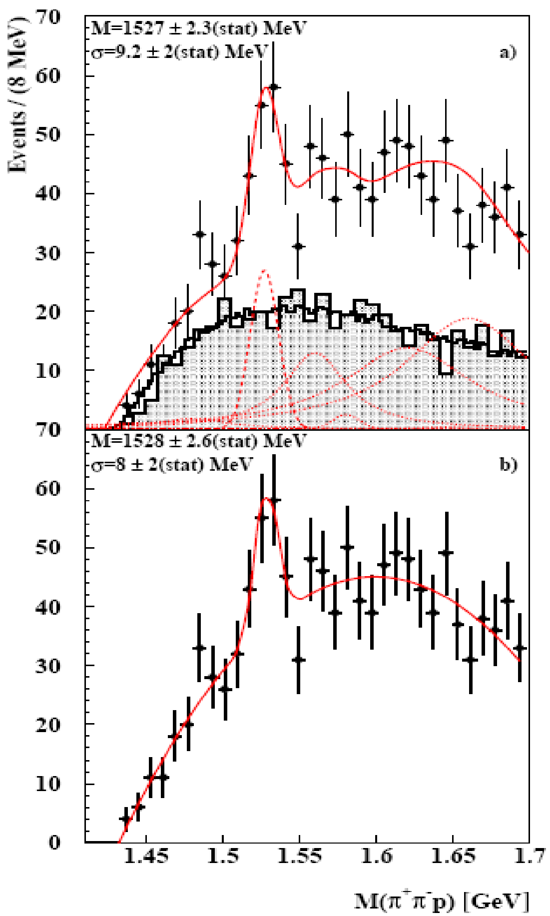
Evidence for a narrow |S| = 1 baryon state at a mass of 1528 MeV in quasi-real photoproduction

A. Airapetian,[31] N. Akopov,[31] Z. Akopov,[31] M. Amarian,[8,31] V.V. Ammosov,[23] A. Andrus,[16] E.C. Aschenauer,[8] W. Augustyniak,[30] R. Avakian,[31] A. Avetissian,[31] E. Avetissian,[12] P. Bailey,[16] D. Balin,[22] V. Baturin,[22] M. Beckmann,[7] S. Belostotski,[22] S. Bernreuther,[10] N. Bianchi,[12] H.P. Blok,[21,29] H. Böttcher,[8] A. Borissov,[18]

V. Vikhrov, M.G. Vincter, C. Vogel, M. Vogt, J. Volmer, C. Weiskopf, J. Wendland, J. Wilbert,
G. Ybeles Smit,[29] Y. Ye,[5] Z. Ye,[5] S. Yen,[27] W. Yu,[4] B. Zihlmann,[21] H. Zohrabian,[31] and P. Zupranski[30]

(The HERMES Collaboration)

[1] Department of Physics, University of Alberta, Edmonton, Alberta T6G 2J1, Canada



Photoproduction on a deuterium target

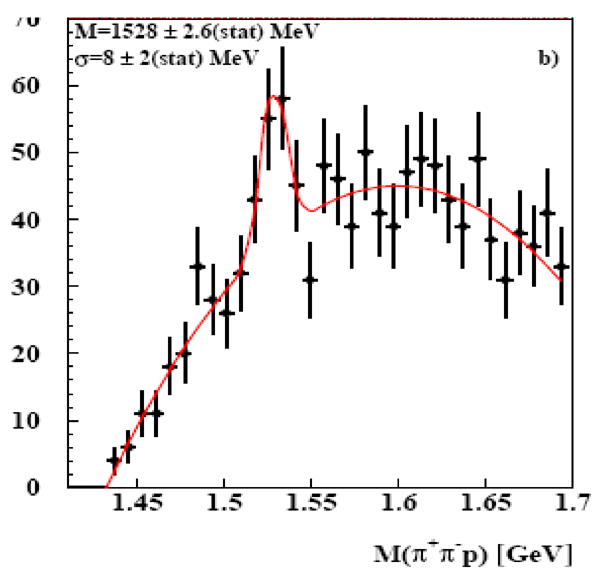
$$\Theta^+ \rightarrow pK^0_S \rightarrow p\pi^+\pi^-$$

FIG. 2: Distribution in invariant mass of the $p\pi^+\pi^-$ system subject to various constraints described in the text. The experimental data are represented by the filled circles with statistical error bars, while the fitted smooth curves result in the indicated position and σ width of the peak of interest. In panel a), the PYTHIA6 Monte Carlo simulation is represented by the gray shaded histogram, the mixed-event model normalised to the PYTHIA6 simulation is represented by the fine-binned histogram, and the fitted curve is described in the text. In panel b), a fit to the data of a Gaussian plus a third-order polynomial is shown.

HERMES : 27.6 GeV positron beam on deuterium

TABLE I: Mass values and experimental widths, with their statistical and systematic uncertainties, for the Θ^+ from the two fits, labelled by a) and b), shown in the corresponding panels of Fig. 2. Rows a') and b') are based on the same background models as rows a) and b) respectively, but a different mass reconstruction expression that is expected to result in better resolution. Also shown are the number of events in the peak N_s and the background N_b, both evaluated from the functions fitted to the mass distribution, and the results for the naïve significance $N_s^{2\sigma}/\sqrt{N_b^{2\sigma}}$ and realistic significance $N_s/\delta N_s$. The systematic uncertainties are common (correlated) between rows of the table.

	Θ^+ mass [MeV]	FWHM [MeV]	$N_s^{2\sigma}$ in $\pm 2\sigma$	$N_b^{2\sigma}$ in $\pm 2\sigma$	naïve signif.	Total $N_s \pm \delta N_s$	signif.
a)	$1527.0 \pm 2.3 \pm 2.1$	$22 \pm 5 \pm 2$	74	145	6.1σ	78 ± 18	4.3σ
a')	$1527.0 \pm 2.5 \pm 2.1$	$24 \pm 5 \pm 2$	79	158	6.3σ	83 ± 20	4.2σ
b)	$1528.0 \pm 2.6 \pm 2.1$	$19 \pm 5 \pm 2$	56	144	4.7σ	59 ± 16	3.7σ
b')	$1527.8 \pm 3.0 \pm 2.1$	$20 \pm 5 \pm 2$	52	155	4.2σ	54 ± 16	3.4σ



$$S_b = \frac{N - \mu_b}{\sqrt{\mu_b}} = \frac{N_b + N_s - \mu_b}{\sqrt{\mu_b}}$$

$$S_0 = \frac{N - N_b}{\sqrt{N + N_b}} = \frac{N_b + N_s - N_b}{\sqrt{N + N_b}} = \frac{N_s}{\sqrt{N + N_b}}$$

Z=3.84

$$74/\sqrt{74+145+74} = 4.3$$

VOLUME 92, NUMBER 3 PHYSICAL REVIEW LETTERS week ending 23 JANUARY 2004

CLAS Collaboration 2004

Observation of an Exotic Baryon with $S = +1$ in Photoproduction from the Proton

V. Kubarovsky,[1,3] L. Guo,[2] D. P. Weygand,[3] P. Stoler,[1] M. Battaglieri,[18] R. DeVita,[18] G. Adams,[1] Ji Li,[1] M. Nozar,[3] C. Salgado,[26] P. Ambrozewicz,[13] E. Anciant,[5] M. Anghinolfi,[18] B. Asavapibhop,[24] G. Audit,[5] T. Auger,[5] H. Avakian,[3] H. Bagdasaryan,[28] J. P. Ball,[4] S. Barrow,[14] K. Beard,[21] M. Bektasoglu,[27] M. Bellis,[1] N. Benmouna,[15] B. L. Berman,[15]

VOLUME 92, NUMBER 3 PHYSICAL REV

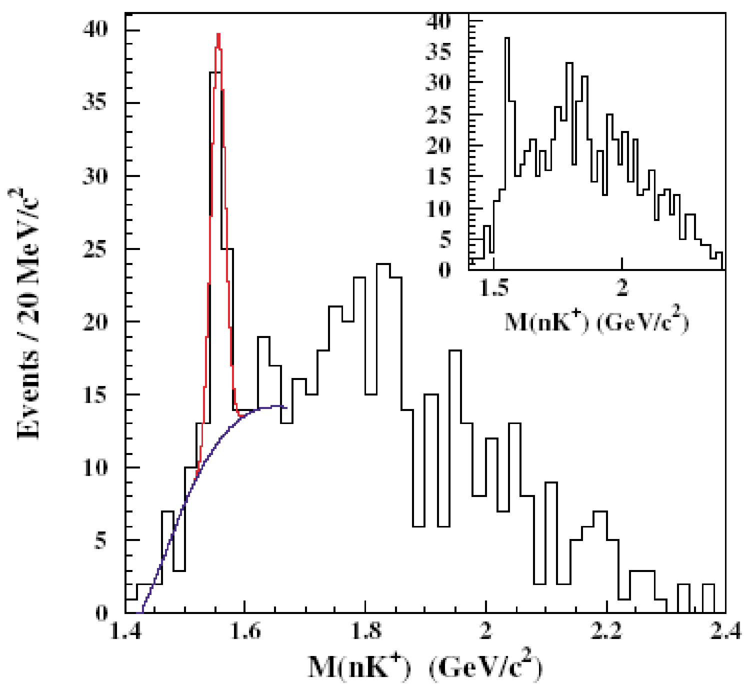
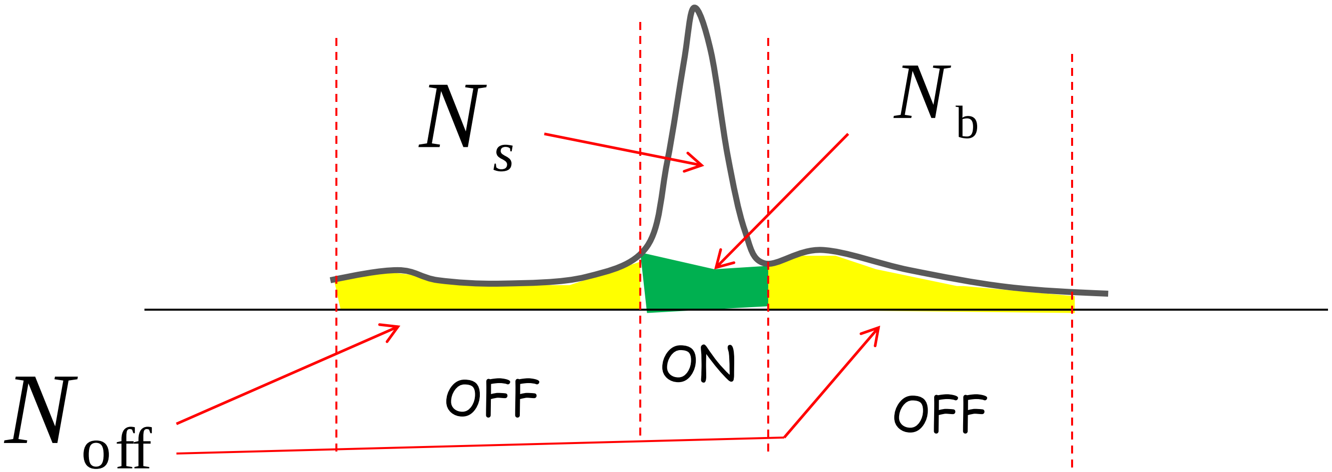


FIG. 4 (color online). The nK^+ invariant mass spectrum in the reaction $\gamma p \to \pi^+ K^- K^+ (n)$ with the cut $\cos\theta^*_{\pi^+} > 0.8$ and $\cos\theta^*_{K^+} < 0.6$. $\theta^*_{\pi^+}$ and $\theta^*_{K^+}$ are the angles between the π^+ and K^+ mesons and photon beam in the center-of-mass system. The background function we used in the fit was obtained from the simulation. The inset shows the nK^+ invariant mass spectrum with only the $\cos\theta^*_{\pi^+} > 0.8$ cut.

$$\frac{41}{\sqrt{27}} = 7.8 \quad ???$$

$$\frac{41}{\sqrt{41+27+27}} = 4.2$$

All these methods estimate true values through measured quantities… but …



$$N_{off} \approx \text{Pois}(\lambda\, \mu_b) \quad \text{with } \lambda \text{ known}$$

Consider

$$N_{on} = N_s + N_b$$

$$N_b \cong \frac{N_{off}}{\lambda} \qquad \lambda = \frac{N_{off}}{N_b} \qquad \sigma_b = \frac{\sqrt{N_{off}}}{\lambda} \qquad \lambda = \frac{N_b}{\sigma_b^2}$$

λ is the ratio between the two surfaces

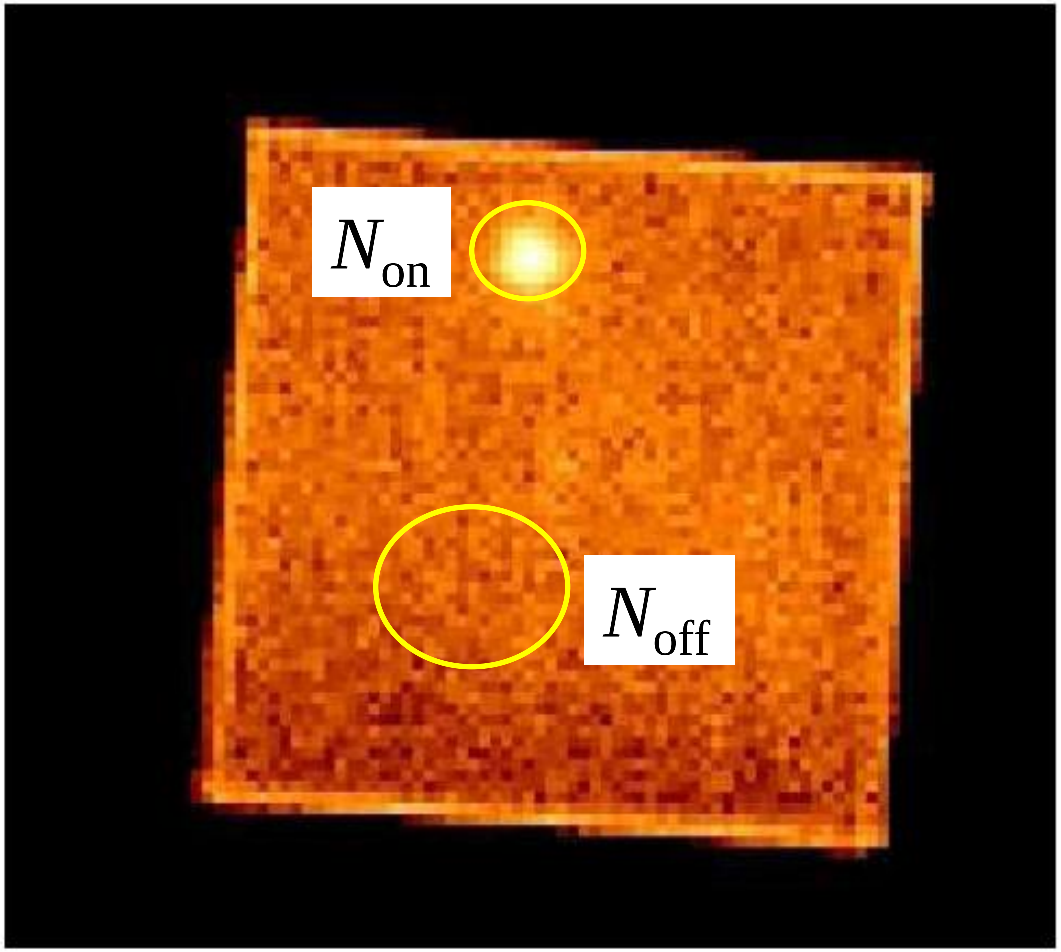


Figura 2.1: Immagine X di PSR J1740+1000, ottenuta dalle due osservazioni di XMM.

A first rigorous solution

R. Cousins et al, NIM A 595(2008)480

The joint probability of observing $n_{\rm on}$ and $n_{\rm off}$ is the product of Poisson probabilities for $n_{\rm on}$ and $n_{\rm off}$, and can be rewritten as the product of a single Poisson probability with mean $\mu_{\rm tot} = \mu_{\rm on} + \mu_{\rm off}$ for the total number of events $n_{\rm tot}$, and the binomial probability that this total is divided as observed if the binomial parameter ρ is $\rho = \mu_{\rm on}/\mu_{\rm tot} = 1/(1+\lambda)$:

$$\lambda = \mu_{\rm off} / \mu_{\rm on} \xrightarrow{H_0} \mu_{\rm off} / \mu_{\rm b}$$

$$P(n_{\rm on}, n_{\rm off}) = \frac{e^{-\mu_{\rm on}} \mu_{\rm on}^{n_{\rm on}}}{n_{\rm on}!} \times \frac{e^{-\mu_{\rm off}} \mu_{\rm off}^{n_{\rm off}}}{n_{\rm off}!}$$

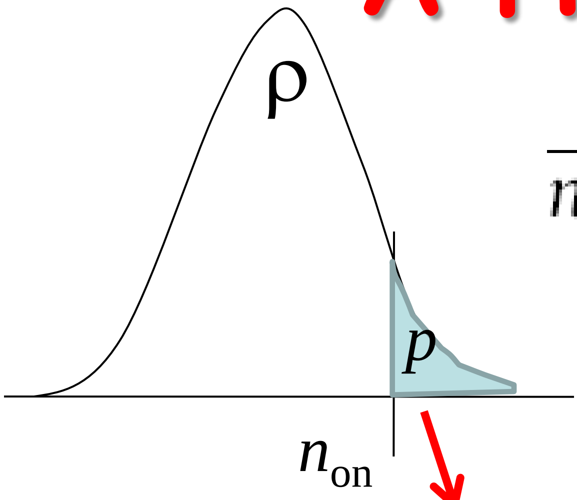
$$= \frac{e^{-(\mu_{\rm on}+\mu_{\rm off})} (\mu_{\rm on}+\mu_{\rm off})^{n_{\rm tot}}}{n_{\rm tot}!} \quad (9)$$

$$\times \frac{n_{\rm tot}!}{n_{\rm on}!(n_{\rm tot}-n_{\rm on})!} \rho^{n_{\rm on}} (1-\rho)^{(n_{\rm tot}-n_{\rm on})}. \quad (10)$$

ρ is known!!!!

λ is the known normalization constant supposing that the **on** measurement does not contain the signal **(H_0 hyp.)**

A rigorous solution



$$\frac{n_{tot}!}{n_{on}!(n_{tot}-n_{on})!}\rho^{n_{on}}(1-\rho)^{(n_{tot}-n_{on})}$$

$$\rho = \mu_{on}/\mu_{tot} = 1/(1+\lambda)$$

$$\lambda = \mu_{off} / \mu_{on}$$

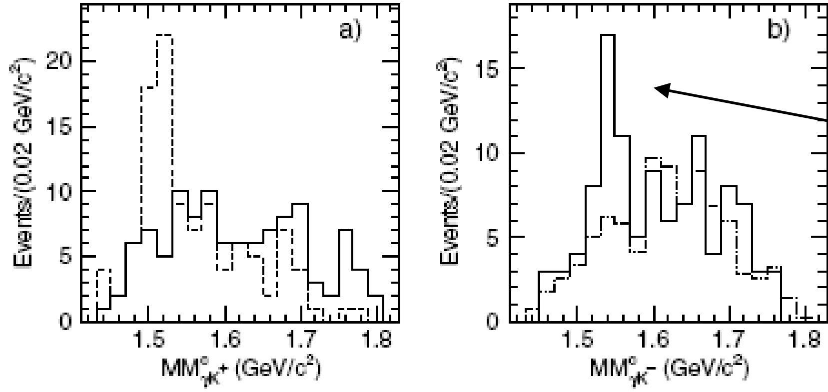
$$p = \sum_{j=n_{on}}^{n_{tot}} B(j; n_{tot}, \rho)$$

$$Z = \Phi^{-1}(1-p) = -\Phi^{-1}(p)$$

where

$$\Phi(Z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{Z} \exp(-t^2/2)\,dt = \frac{1+\mathrm{erf}(Z/\sqrt{2})}{2} = 1-p$$

$$Z = \sqrt{2}\,\mathrm{erf}^{-1}(1-2p)$$



4.6 sigma!

Is it convincing???

FIG. 3. (a) The $MM^c_{\gamma K^+}$ spectrum [Eq. (2)] for K^+K^- productions for the signal sample (solid histogram) and for events from the SC with a proton hit in the SSD (dashed histogram). (b) The $MM^c_{\gamma K^-}$ spectrum for the signal sample (solid histogram) and for events from the LH_2 (dotted histogram) normalized by a fit in the region above 1.59 GeV$/c^2$.

012002-3

The background level in the peak region is estimated to be $17.0 \pm 2.2 \pm 1.8$, where the first uncertainty is the error in the fitting in the region above 1.59 GeV$/c^2$ and the second is a statistical uncertainty in the peak region. The combined uncertainty of the background level is ± 2.8. The estimated number of the events above the background level is 19.0 ± 2.8, which corresponds to a Gaussian significance of $4.6^{+1.2}_{-1.0}\sigma$ $(19.0/\sqrt{17.0} = 4.6)$.

$$p_{\mathrm{Bi}} = \sum_{j=n_{\mathrm{on}}}^{n_{\mathrm{tot}}} P(j|n_{\mathrm{tot}};\rho) \quad = B(\rho, n_{\mathrm{on}}, 1+n_{\mathrm{off}})/B(n_{\mathrm{on}}, 1+n_{\mathrm{off}})$$

$$Z = \sqrt{2}\,\mathrm{erf}^{-1}(1-2p)$$

For the simple on/off problem with $n_{\mathrm{on}} = 140$, $n_{\mathrm{off}} = 100$, and $\tau = 1.2$, the ROOT commands are:

```
double n_on  = 140.
double n_off  = 100.
double tau  = 1.2
double P_Bi  = TMath::BetaIncomplete(1./(1.+tau),n_on,n_off+1)
double Z_Bi  = sqrt(2)*TMath::ErfInverse(1 - 2*P_Bi)
```

Pentaquark: n_on=36, n_off= 89
τ= λ = 89/17 = 5.2, Z=3.49

Likelihood -> Z=3.52

A 2nd "rigorous" solution

$$\lambda = \frac{\mu_{\rm off}}{\mu_{\rm on}} \rightarrow \text{no signal} \rightarrow \tau = \frac{\mu_{\rm off}}{\mu_{\rm b}} \qquad \frac{1}{1+\tau} = \frac{\mu_{\rm b}}{\mu_{\rm b} + \mu_{\rm off}}$$

$$\Lambda = \frac{L(\mu)}{L(\text{best})} = \frac{\mu_b^{N_{\rm on}} \; e^{-\mu_{\rm b}} \; \mu_{\rm off}^{N_{\rm off}} \; e^{-\mu_{\rm off}}}{N_{\rm on}^{N_{\rm on}} \; e^{-N_{\rm on}} \; N_{\rm off}^{N_{\rm off}} \; e^{-N_{\rm off}}}$$

$$\mu_b = \frac{1}{1+\tau}(N_{\rm on} + N_{\rm off}) \qquad \mu_{\rm off} = \frac{\tau}{1+\tau}(N_{\rm on} + N_{\rm off})$$

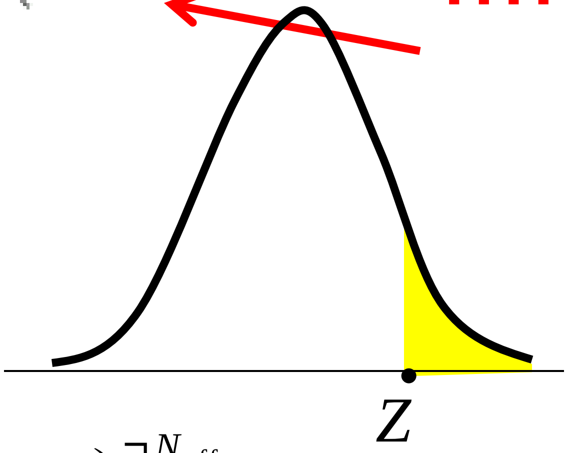
$$\Lambda = \left[\frac{1}{1+\tau}\left(\frac{N_{\rm on} + N_{\rm off}}{N_{\rm on}}\right)\right]^{N_{\rm on}} \left[\frac{\tau}{1+\tau}\left(\frac{N_{\rm on} + N_{\rm off}}{N_{\rm off}}\right)\right]^{N_{\rm off}}$$

A 2nd "rigorous" solution

T.-P. Li, Y.-Q. Ma, Astrophys. J. 272 (1983) 317.

!!!!

R. Cousins et al, NIM A 595(2008)480



Approximated Gaussian significance

$$\Lambda = \left[\frac{1}{1+\tau}\left(\frac{N_{on}+N_{off}}{N_{on}}\right)\right]^{N_{on}} \left[\frac{\tau}{1+\tau}\left(\frac{N_{on}+N_{off}}{N_{off}}\right)\right]^{N_{off}}$$

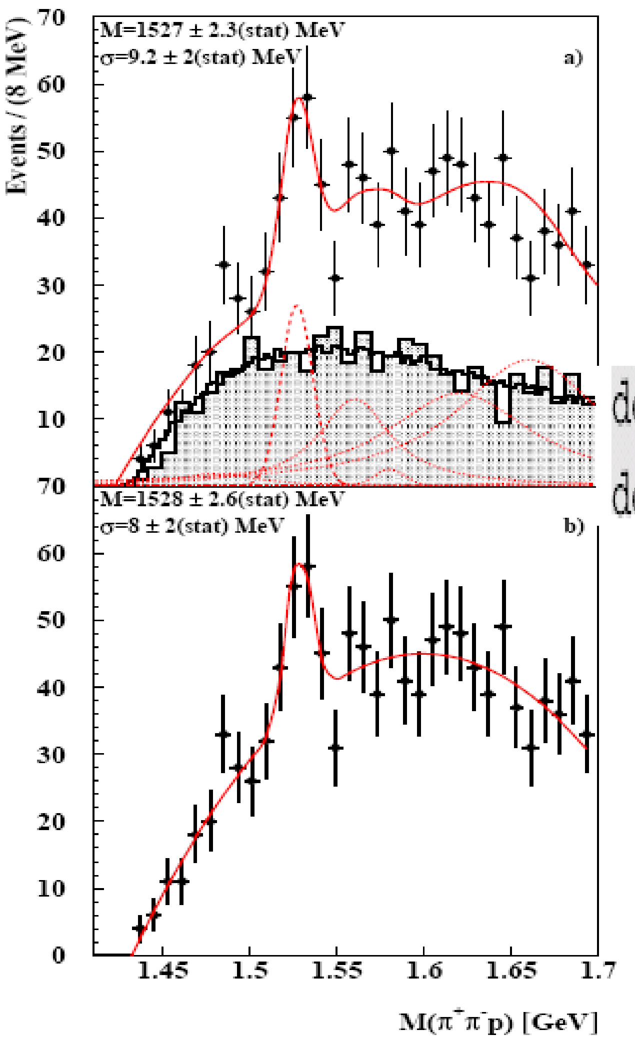
$$Z = \sqrt{\chi^2(1)} = \sqrt{-2\ln\Lambda} =$$

$$\sqrt{2}\left[N_{on}\ln\left(\frac{(1+\tau)N_{on}}{N_{on}+N_{off}}\right) + N_{off}\ln\left(\frac{(1+\tau)N_{off}}{\tau(N_{on}+N_{off})}\right)\right]^{1/2}$$

HERMES : 27.6 GeV positron beam on deuterium

$$S_b = \frac{N - \mu_b}{\sqrt{\mu_b}} = \frac{N_b + N_s - \mu_b}{\sqrt{\mu_b}} \simeq \frac{N_s}{\sqrt{\mu_b}} = 6.1$$

$$74/\sqrt{74+145+74} = 4.3$$



$$S_0 = \frac{N - N_b}{\sqrt{N + N_b}} = \frac{N_b + N_s - N_b}{\sqrt{N + N_b}} = \frac{N_s}{\sqrt{N + N_b}} = 4.3$$

$$N_{off} = 145 \qquad N_{on} = 145 + 74 = 219$$

$$N_{tot} = 145 + 219 = 364 \quad N_b = 145 \quad \lambda = \tau = 1$$

```
double P_Bi = TMath::BetaIncomplete(1./(1.+tau),n_on,n_off+1)
double Z_Bi = sqrt(2)*TMath::ErfInverse(1 - 2*P_Bi)
```

Z=3.84 Binomial

$$Z = \sqrt{2}\left[N_{on}\ln\left(\frac{(1+\tau)N_{on}}{N_{on}+N_{off}}\right) + N_{off}\ln\frac{1+\tau}{\tau}\left(\frac{N_{off}}{N_{on}+N_{off}}\right)\right]^{1/2}$$

Z=3.90 Likelihood

Table 1
Test cases and significance results

Reference	[40]	[41]	[42]	[43]	[44]	[44]	[45]	[46]	[47]	[48]
n_{on}	**4**	**6**	**9**	**17**	**50**	**67**	**200**	**523**	**498 426**	**2 119 449**
n_{off}	**5**	18.78	17.83	40.11	**55**	**15**	**10**	**2327**	**493 434**	23 650 096
τ	**5.0**	14.44	4.69	10.56	**2.0**	**0.5**	**0.1**	**5.99**	**1.0**	**11.21**
$\hat{\mu}_{\mathrm{b}}$	1.0	**1.3**	**3.8**	**3.8**	27.5	30.0	100.0	388.6	493 434	**2 109 732**
$s = n_{\mathrm{on}} - \hat{\mu}_{\mathrm{b}}$	3.0	4.7	5.2	13.2	22.5	37	100	**134**	**4992**	**9717**
σ_{b}	0.447	**0.3**	**0.9**	**0.6**	3.71	7.75	31.6	8.1	702.4	433.8
$f = \sigma_{\mathrm{b}}/\hat{\mu}_{\mathrm{b}}$	0.447	0.231	0.237	0.158	0.135	0.258	0.316	0.0207	0.00142	0.000206
Reported p		**0.003**	**0.027**	**2E−06**						
Reported Z		2.7	1.9	**4.6**				**5.9**	**5.0**	**6.4**
See conclusion										
$Z_{\mathrm{Bi}} = Z_{\Gamma}$ binomial	**1.66**	**2.63**	**1.82**	**4.46**	**2.93**	**2.89**	**2.20**	**5.93**	**5.01**	**6.40**
Z_{N} Bayes Gaussian	1.88	2.71	1.94	4.55	3.08	*3.44*	*2.90*	**5.93**	**5.02**	**6.40**
Z_{PL} profile likelihood	1.95	2.81	1.99	4.57	3.02	3.04	2.38	**5.93**	**5.01**	**6.41**
Z_{ZR} variance stabilization	1.93	2.66	1.98	4.22	3.00	3.07	2.39	5.86	**5.01**	**6.40**
Not recommended										
$Z_{\mathrm{BiN}} = s/\sqrt{n_{\mathrm{tot}}/\tau}$	*2.24*	*3.59*	2.17	*5.67*	3.11	**2.89**	**2.18**	6.16	**5.01**	**6.41**
$Z_{\mathrm{nn}} = s/\sqrt{n_{\mathrm{on}} + n_{\mathrm{off}}/\tau^2}$	1.46	*1.90*	1.66	*3.17*	2.82	3.28	*2.89*	*5.54*	**5.01**	**6.40**
$Z_{\mathrm{ssb}} = s/\sqrt{\hat{\mu}_{\mathrm{b}} + s}$	1.50	*1.92*	1.73	*3.20*	3.18	*4.52*	*7.07*	5.88	*7.07*	6.67
$Z_{\mathrm{bo}} = s/\sqrt{n_{\mathrm{off}}(1+\tau)/\tau^2}$	*2.74*	*3.99*	*2.42*	*6.47*	*3.50*	*3.90*	*3.02*	6.31	**5.03**	**6.41**
Ignore σ_{b}										
Z_{P} Poisson: ignore σ_{b}	2.08	2.84	2.14	4.87	*3.80*	*5.76*	*8.76*	*6.44*	*7.09*	6.69
$Z_{\mathrm{sb}} = s/\sqrt{\hat{\mu}_{\mathrm{b}}}$	*3.00*	*4.12*	*2.67*	*6.77*	*4.29*	*6.76*	*10.00*	*6.82*	*7.11*	6.69
Unsuccessful ad hockery										
Poisson: $\mu_{\mathrm{b}} \to \hat{\mu}_{\mathrm{b}} + \sigma_{\mathrm{b}}$	1.56	2.51	1.64	**4.47**	3.04	*4.24*	*5.51*	6.01	*6.09*	**6.39**
$s/\sqrt{\hat{\mu}_{\mathrm{b}} + \sigma_{\mathrm{b}}}$	*2.49*	*3.72*	*2.40*	*6.29*	*4.03*	*6.02*	*8.72*	*6.75*	*7.10*	6.69

Kolmogorov-Smirnov test

Kolmogorov, A. (1933) "Sulla determinazione empirica di una legge di distribuzione“ *G. Inst. Ital. Attuari*, 4, 83

A rigorous method to compare observed data to a given distribution (or to check if two samples are from the same population):

For a sample of dimension *n*, the empirical cdf is

$$F_n(x^*) = \frac{1}{n}\sum_{i=1}^{n} n\{x_i \le x^*\}$$

This has to be compared with the theoretical cdf of the distribution assumed for the data, $F(x)$, through the evaluation of the maximum difference:

$$D = \max_x \{F_n(x) - F(x)\}$$

Under the hypothesis H_0 that the data are from the assumed distribution, the significance level of a particular value of *D* is given by:

$$p_D = Q(\sqrt{n}D) \text{ , where } Q(x) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2x^2}$$

KS test

- If two samples (of dimensions n_1 and n_2) have to be compared the argument of the Q function becomes

$$D\sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

- The KS test need not know the distribution parameters, but it applies to continuous distributions:

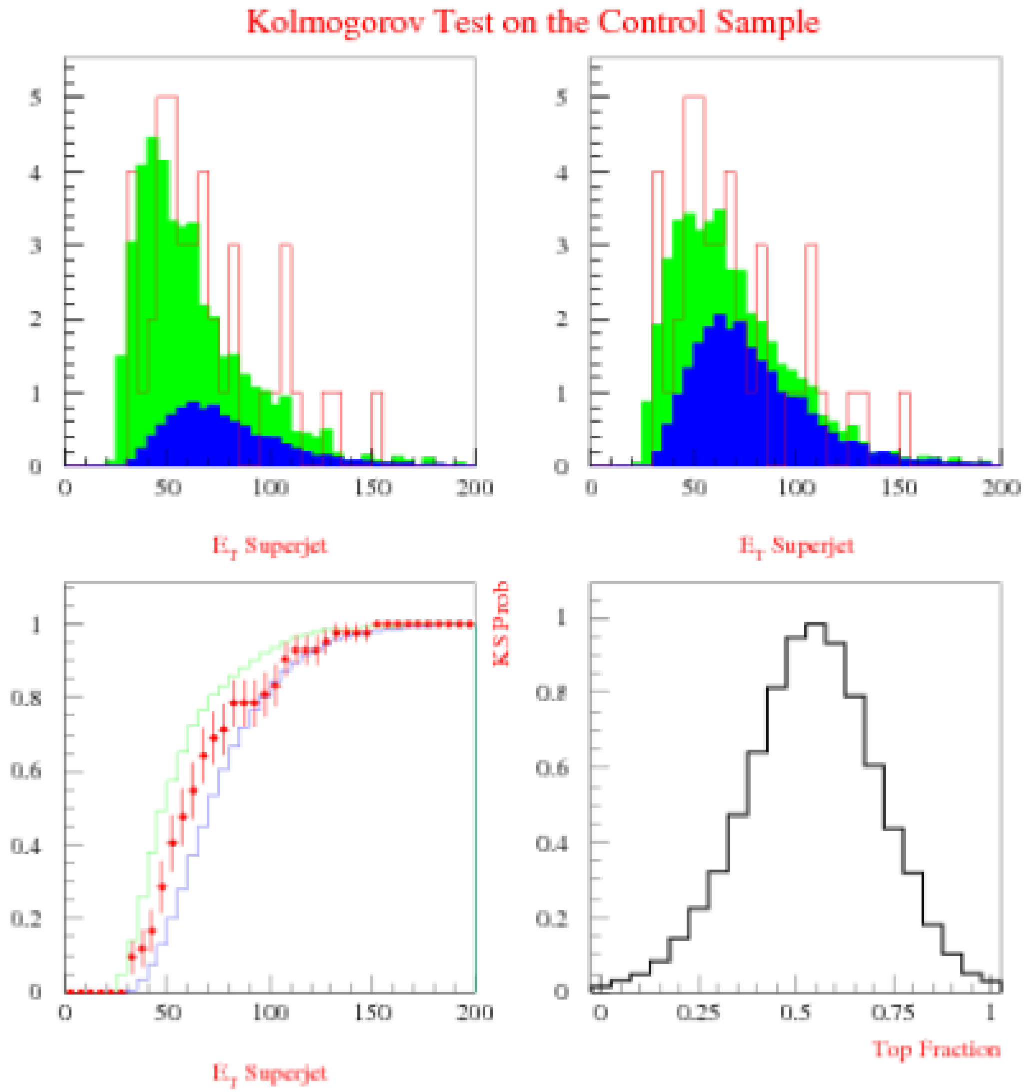
  an attempt to apply KS test to discrete distribution can be found in

  Facchinetti, S., & Chiodini, P.M. (2008)

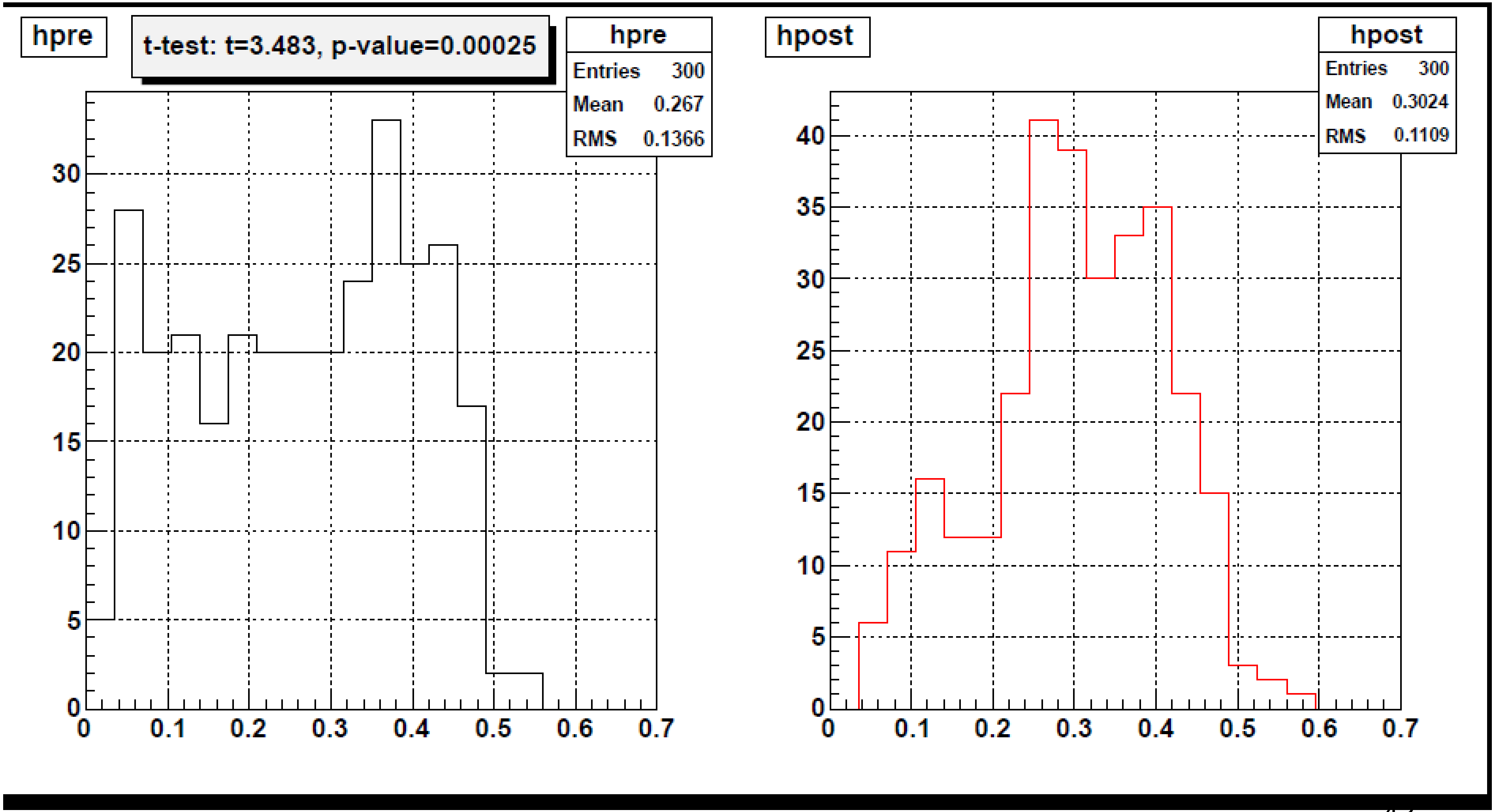
  http://hdl.handle.net/10281/4859

  Note that in this case the procedure is no longer distribution free.

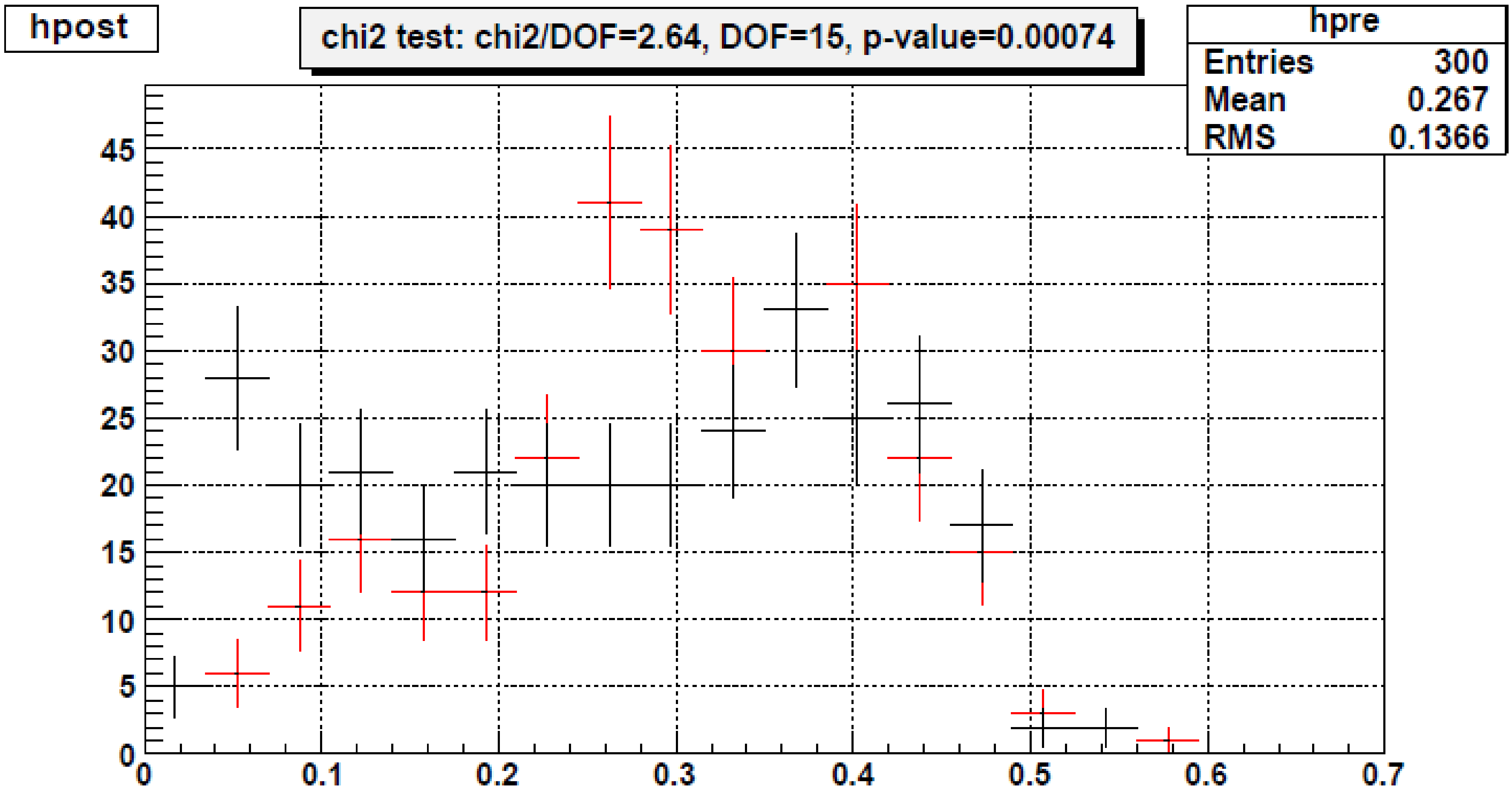
Graphical example of application of the KS test on CDF “superjet events”.
green (QCD) and blue (t-tbar).
ul: expected admixture; ur: best admixture; ll:integral distributions; lr: admixure probability
(see *T. Dorigo, note CDF/ANAL/TOP/CDFR/4861*)



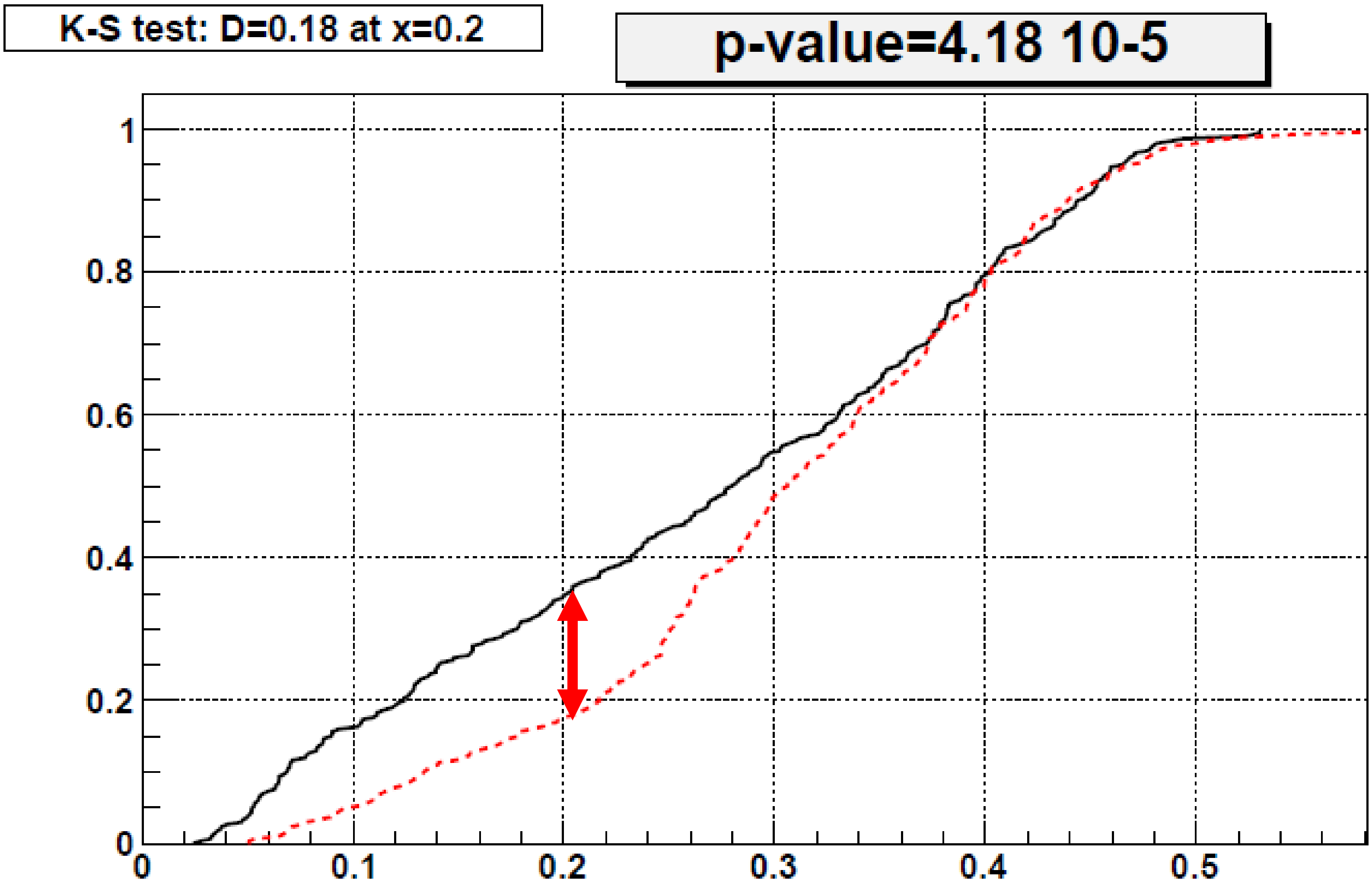
$$t_m = |m_1 - m_2| \left[\frac{s_1^2}{N_1} + \frac{s_2^2}{N_2} \right]^{-1/2}$$



$$\chi^2 = \frac{(n_1 - r_1)^2}{n_1 + r_1} + \frac{(n_2 - r_2)^2}{n_2 + r_2} + \ldots$$



$$D_n = \sup_x | F_n(x) - F(x) |$$



Conclusions

- **Physicists should follow the right statistical notation**

- The usual formulae used by physicists to find signal on a "back of an envelope" should be abandoned

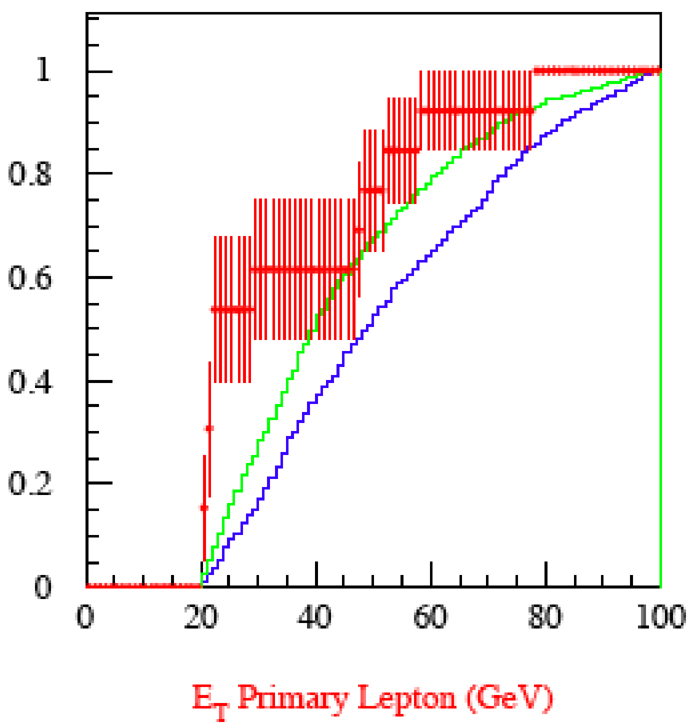
- the RIGHT formulae for the signal significance there exist and should be used (see **R. Cousins et al, NIM A 595(2008)480**)

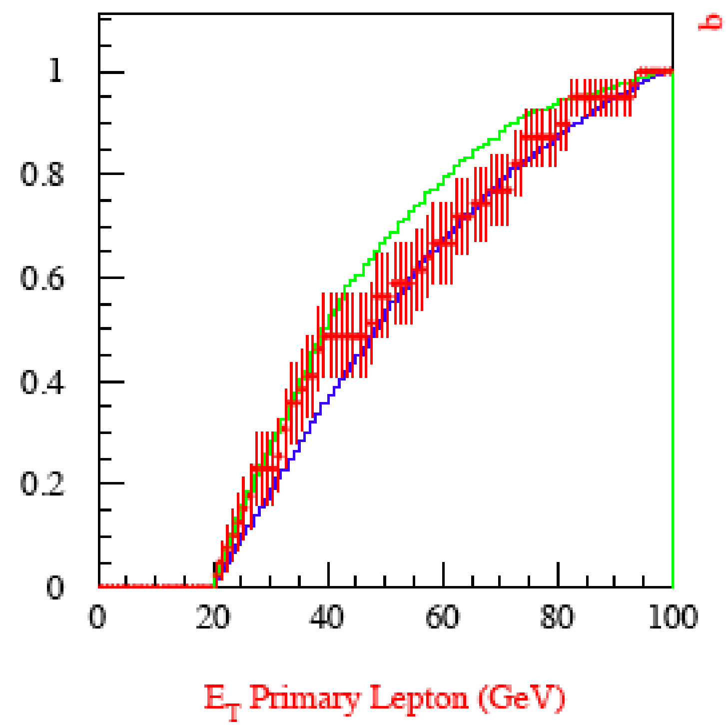
  1 binomial formula
  2 likelihood ratio formula
  3 KS test with unbinned continuous data

END

Graphical example of application of the KS test on CDF “superjet events”.

Left: comparison of superjet sample with expected event distributions; right: comparison of control sample with expected event distributions. (see *T. Dorigo, note CDF/ANAL/TOP/CDFR/4861*)





VOLUME 92, NUMBER 3 PHYSICAL REVIEW LETTERS week ending 23 JANUARY 2004

Observation of an Exotic Baryon with $S = +1$ in Photoproduction from the Proton

V. Kubarovsky,[1,3] L. Guo,[2] D. P. Weygand,[3] P. Stoler,[1] M. Battaglieri,[18] R. DeVita,[18] G. Adams,[1] Ji Li,[1] M. Nozar,[3]
C. Salgado,[26] P. Ambrozewicz,[13] E. Anciant,[5] M. Anghinolfi,[18] B. Asavapibhop,[24] G. Audit,[5] T. Auger,[5] H. Avakian,[3]
H. Bagdasaryan,[28] J. P. Ball,[4] S. Barrow,[14] K. Beard,[21] M. Bektasoglu,[27] M. Bellis,[1] N. Benmouna,[15] B. L. Berman,[15]
C. S. Whisnant,[32] E. Wolin,[3] M. H. Wood,[32] A. Yegneswaran,[3] and J. Yun[28]

VOLUME 92, NUMBER 3 PHYSICAL REV

(CLAS Collaboration)

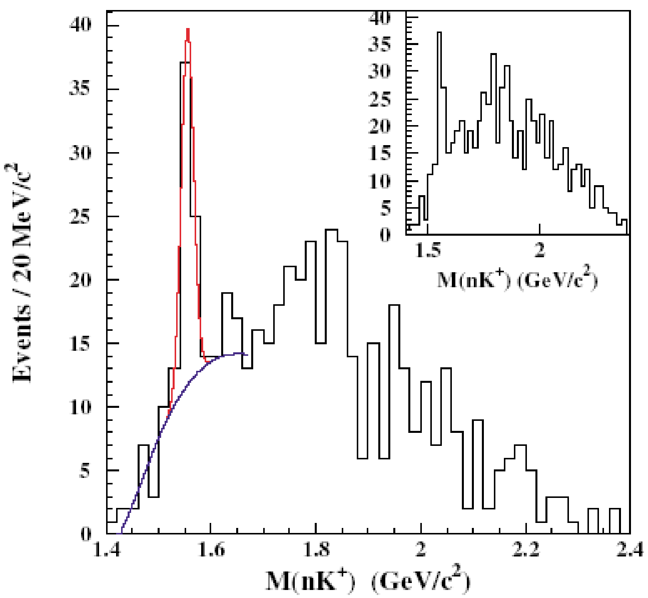


FIG. 4 (color online). The nK^+ invariant mass spectrum in the reaction $\gamma p \rightarrow \pi^+ K^- K^+ (n)$ with the cut $\cos\theta^*_{\pi^+} > 0.8$ and $\cos\theta^*_{K^+} < 0.6$. $\theta^*_{\pi^+}$ and $\theta^*_{K^+}$ are the angles between the π^+ and K^+ mesons and photon beam in the center-of-mass system. The background function we used in the fit was obtained from the simulation. The inset shows the nK^+ invariant mass spectrum with only the $\cos\theta^*_{\pi^+} > 0.8$ cut.

The final nK^+ effective mass distribution (Fig. 4) was fitted by the sum of a Gaussian function and a background function obtained from the simulation. The fit parameters are $N_{\Theta^+} = 41 \pm 10$, $M = 1555 \pm 1\ \mathrm{MeV}/c^2$, and $\Gamma = 26 \pm 7\ \mathrm{MeV}/c^2$ (FWHM), where the errors are statistical. The systematic mass scale uncertainty is estimated to be $\pm 10\ \mathrm{MeV}/c^2$. This uncertainty is larger than our previously reported uncertainty [6] because of the different energy range and running conditions and is mainly due to the momentum calibration of the CLAS detector and the photon beam energy calibration. The statistical significance for the fit in Fig. 4 over a $40\ \mathrm{MeV}/c^2$ mass window is calculated as $N_P/\sqrt{N_B}$, where N_B is the number of counts in the background fit under the peak and N_P is the number of counts in the peak. We estimate the significance to be $7.8 \pm 1.0\sigma$. The uncertainty of 1.0σ is due to

Statistics

$$\xi_1 = \frac{s}{\sqrt{b}}$$

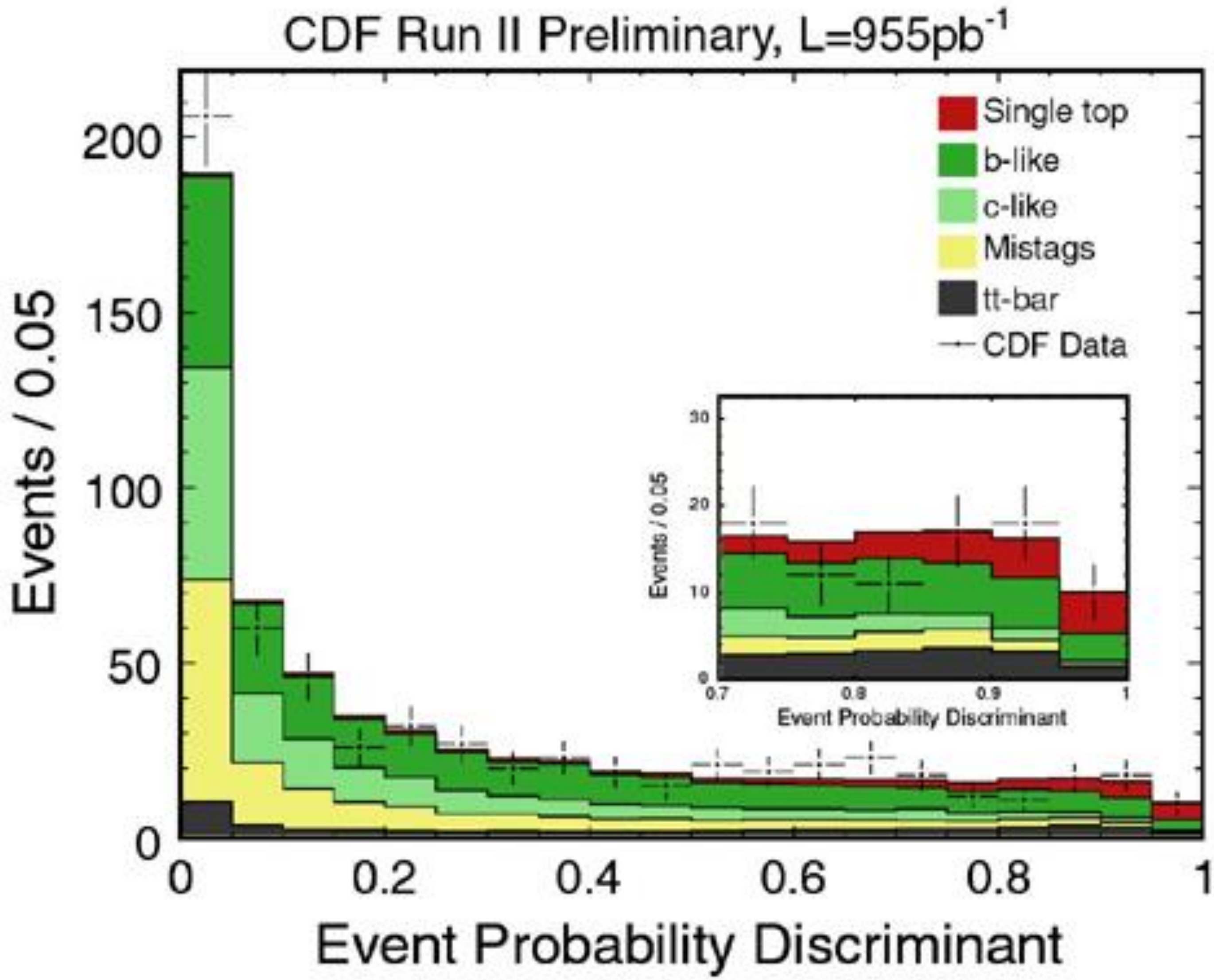
$$\xi_2 = \frac{s}{\sqrt{s+b}}$$

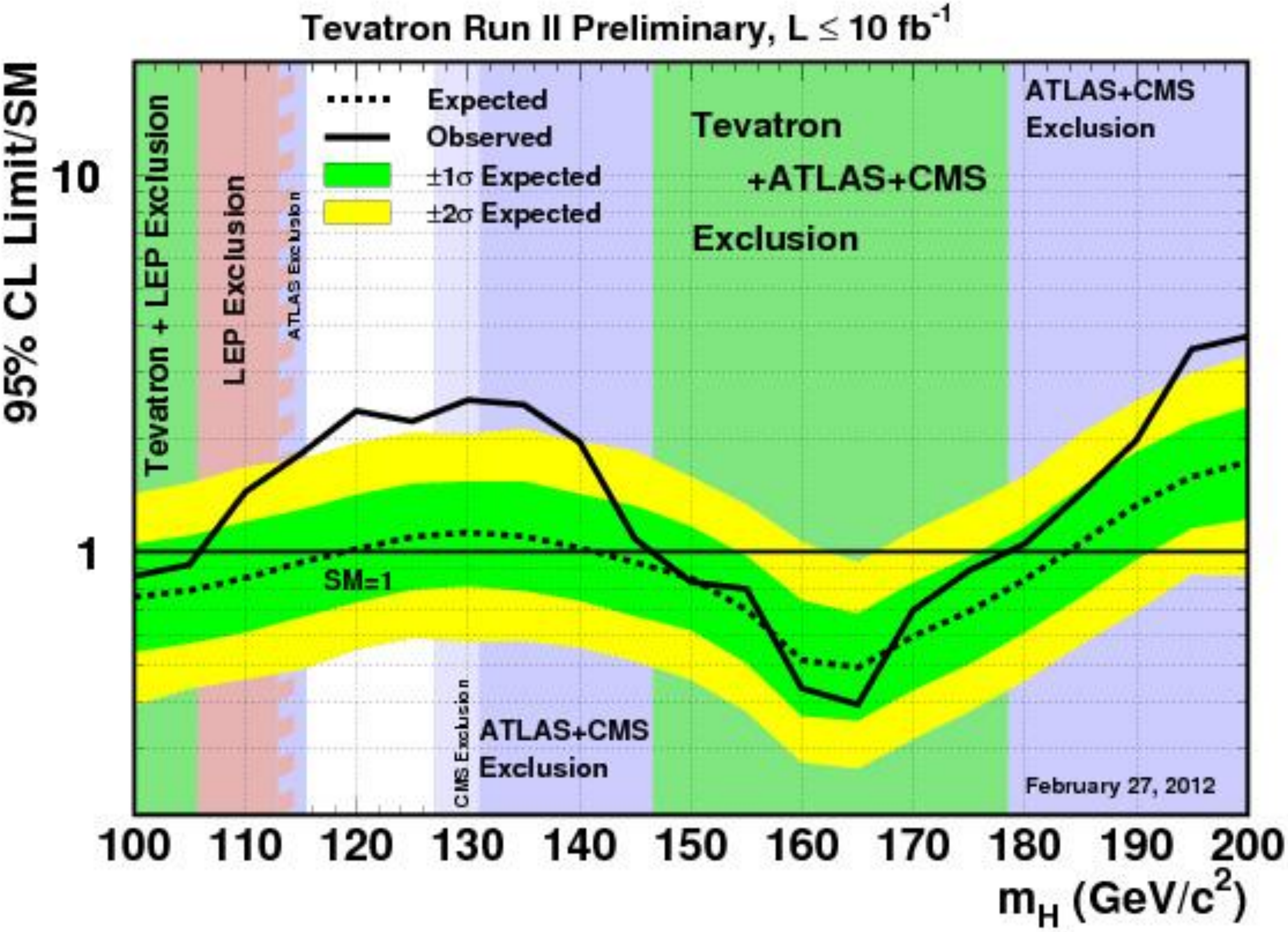
$$\xi_3 = \frac{s}{\sqrt{s+2b}}$$

Experiment	Signal	Background	Publ.	Significance ξ_1	ξ_2	ξ_3
Spring8	19	17	4.6σ	4.6	3.2	2.6
Spring8	56	162		4.4	3.8	2.9
SPAHIR	55	56	4.8σ	7.3	5.2	4.3
CLAS (d)	43	54	5.2σ	5.9	4.4	3.5
CLAS (p)	41	35	7.8σ	6.9	4.7	3.9
DIANA	29	44	4.4σ	4.4	3.4	2.7
ν	18	9	6.7σ	6.0	3.5	3.0
HERMES	51	150	4.3-6.2σ	4.2	3.6	2.7
COSY	57	95	4-6σ	5.9	4.7	3.7
ZEUS	230	1080	4.6σ	7.0	6.4	4.7
SVD	35	93	5.6σ	3.6	3.1	2.4
NOMAD	33	59	4.3σ	4.3	3.4	2.7
NA49	38	43	4.2σ	5.8	4.2	3.4
NA49	69	75	5.8σ	8.0	5.8	4.7
H1	50.6	51.7	5-6σ	7.0	5.0	4.1

No 5σ effect!!

Signal to background




Tevatron Run II Preliminary, L ≤ 10 fb-1
95% CL Limit/SM
Tevatron + LEP Exclusion
LEP Exclusion
ATLAS Exclusion
Expected
Observed
±1σ Expected
±2σ Expected
Tevatron
+ATLAS+CMS
Exclusion
ATLAS+CMS
Exclusion
SM=1
CMS Exclusion
ATLAS+CMS
Exclusion
February 27, 2012
10
1
100
110
120
130
140
150
160
170
180
190
200
mH (GeV/c2)

From coin tossing to physics: the efficiency measurement

$$P(\varepsilon;k,n) = (n+1)\binom{n}{k}\varepsilon^k(1-\varepsilon)^{n-k}$$
$$= \frac{(n+1)!}{k!\,(n-k)!}\,\varepsilon^k(1-\varepsilon)^{n-k}$$

$$\overline{\varepsilon} = \int_0^1 \varepsilon\; P(\varepsilon;k,n)\; d\varepsilon$$
$$= \frac{(n+1)!}{k!\,(n-k)!}\int_0^1 \varepsilon^{k+1}(1-\varepsilon)^{n-k}d\varepsilon$$
$$= \frac{k+1}{n+2}$$

Valid also for *k=0* and *k=n*

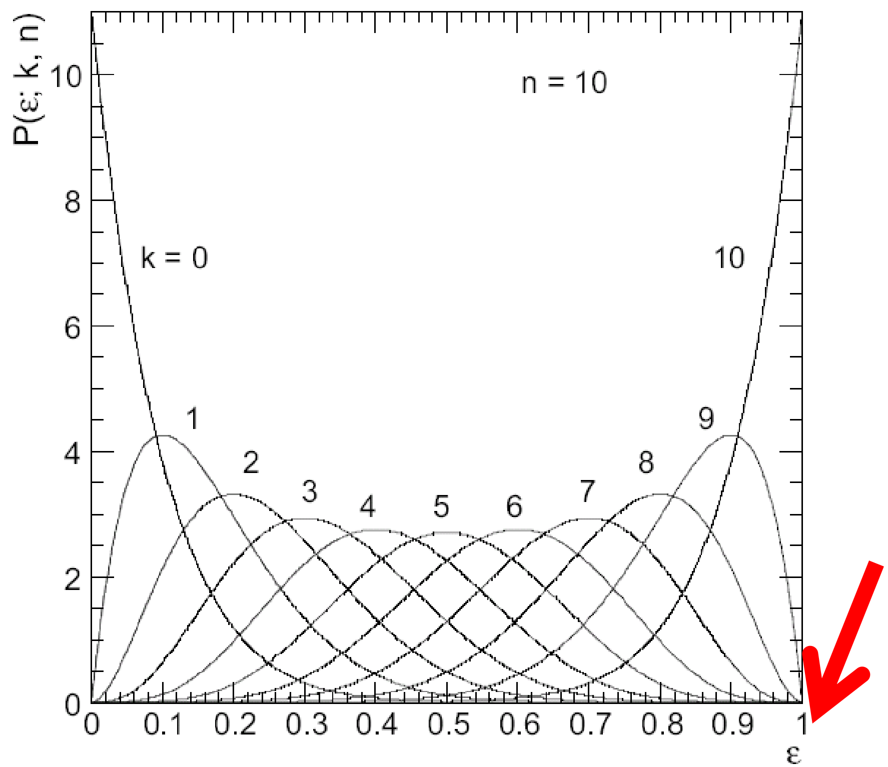
$$V(\varepsilon) = \overline{\varepsilon^2} - \overline{\varepsilon}^2$$
$$= \int_0^1 \varepsilon^2\; P(\varepsilon;k,n)\; d\varepsilon - \overline{\varepsilon}^2$$
$$= \frac{(k+1)\,(k+2)}{(n+2)\,(n+3)} - \frac{(k+1)^2}{(n+2)^2}.$$

ArXiv:physics/0701199v1

Treatment of Errors in Efficiency Calculations

T. Ullrich and Z. Xu
Brookhaven National Laboratory

February 2, 2008



re 1: The probability density function $P(\varepsilon;k,n)$ for $n=10$ and $k=0,1,\ldots,10$.

Frequentist C.I.
right and wrong definitions

RIGHT quotations:

- CL is the probability that the random interval $[T_1, T_2]$ covers the true value θ;
- in an infinite set of repeated identical experiments, a fraction equal to CL will succeed in assigning $\theta \in [\theta_1, \theta_2]$;
- if $\theta \notin [\theta_1, \theta_2]$, one can obtain $\{I = [\theta_1, \theta_2]\}$ in a fraction of experiments $\leq 1 - CL$
- if $H_0 : \theta \notin [\theta_1, \theta_2]$ the probability to reject a true H_0 is $1 - CL$ (falsification).
  see upper and lower limits estimates.

WRONG quotations

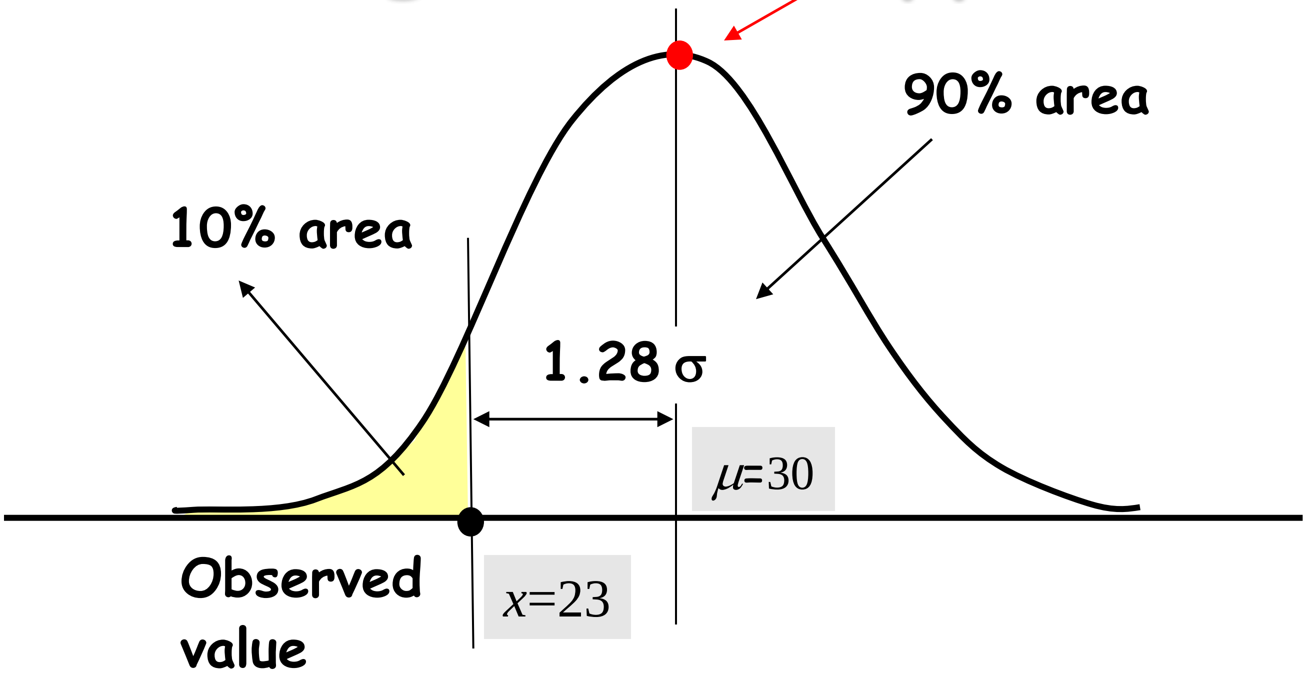
- CL is the degree of belief that the true value is in $[\theta_1, \theta_2]$
- $P\{\theta \in [\theta_1, \theta_2]\} = CL$
  (θ is not a random variable!)

The prosecutor fallacy

- 10^6 tickets are distributed in a lottery. The probability for a Honest (H) to win is $P(W|H)=10^{-6}$.
- If $P(W|H)=P(H|W)=10^{-6}$, the winner is Cheat!!
- The right response is given by the Bayes theorem

$$P(H|W)=\frac{P(W|H)P(H)}{P(W|H)P(H)+P(W|C)P(C)}=\begin{cases}1 \text{ if } P(C)=0\\ 10^{-2} \text{ if there are}\\ \qquad 10\,000 \text{ Cheats}\end{cases}$$

The 90% CL gaussian upper limit..



Meaning: this upper limit should give values less than the observed one in less than 10% of the experiments

$$\frac{\mu - 23}{\sqrt{\mu}} = 1.28 \Rightarrow \mu - 1.28\sqrt{\mu} - 23 = 0 \Rightarrow \mu = 30.0$$

$$(23 + 1.28\sqrt{23} = 29)$$

A 2nd rigorous solution

R. Cousins et al, NIM A 595(2008)480

$$\mathcal{L}_{\mathrm{P}} = \frac{(\mu_s + \mu_b)^{n_{\mathrm{on}}}}{n_{\mathrm{on}}!} e^{-(\mu_s+\mu_b)} \frac{(\tau\mu_b)^{n_{\mathrm{off}}}}{n_{\mathrm{off}}!} e^{-\tau\mu_b} \quad (20)$$

while for the Gaussian-mean background problem with either absolute or relative σ_b, it is

$$\mathcal{L}_{\mathrm{G}} = \frac{(\mu_s + \mu_b)^{n_{\mathrm{on}}}}{n_{\mathrm{on}}!} e^{-(\mu_s+\mu_b)} \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp\left(-\frac{(\hat{\mu}_b - \mu_b)^2}{2\sigma_b^2}\right) \quad (21)$$

where as discussed below we have explored the effect of truncating the Gaussian pdf in $\hat{\mu}_b$ and renormalizing prior to forming $\mathcal{L}_{\mathrm{G}}$.

Using either $\mathcal{L}_{\mathrm{P}}$ or $\mathcal{L}_{\mathrm{G}}$, one obtains the log-likelihood ratio

$$-2\ln\Lambda(\mu_s) < F^{-1}_{\chi^2_1}(1-2\alpha)$$

$$\Lambda(\mu_s) = \frac{\mathcal{L}(\mu_s, \hat{\hat{\mu}}_b)}{\mathcal{L}(\hat{\mu}_s, \hat{\mu}_b)} \quad (22)$$

A 2nd rigourous solution

$$Z_{PL} = \sqrt{-2 \ln \Lambda(\mu_s = 0)} \qquad (24)$$

where the likelihood ratio is computed using $\mathscr{L}_P$ or $\mathscr{L}_G$, as appropriate for the problem.

For the on/off problem and $\mathscr{L}_P$, the explicit result obtained from Eq. (24) was given by Li and Ma (their Eq. 17) [8]:

$$Z_{PL} = \sqrt{2} \left(n_{on} \ln \frac{n_{on}(1+\tau)}{n_{tot}} + n_{off} \ln \frac{n_{off}(1+\tau)}{n_{tot}\tau} \right)^{1/2}. \qquad (25)$$

Pentaquark: n_on=36, n_off= 17*2.17 = 36.7,

$\tau = \lambda = 17/2.8^2 = 2.17$, Z=3.25