

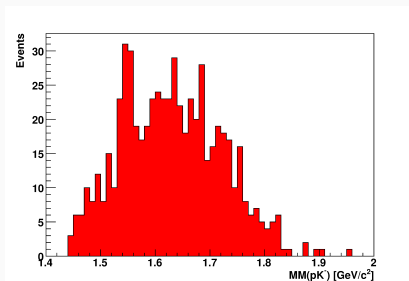
Bayesian Analysis Methods for Hadronic Spectroscopy

D.G. Ireland

Exotic Hadron Spectroscopy Workshop,
September 26-27, 2016

Statistical Evidence from a Spectrum

Given a spectrum with a **feature**, what is the evidence that there is a signal?



- Do you know where the peak should be?
- Can you model the background?
- Are you prepared to make a bet?

Two years later...



New Experiment Casts Doubt on Elusive Pentaquark

By Kate Tiedt

A dedicated hunt for the pentaquark has found nothing. Further, selling into question previous reports of pentaquark sightings. An experiment at the JLab says findings go to counter 2003 report of the "theta collaboration."

The CLAS experiment has produced the first higher than expected results. Because only a few pentaquarks, but they do provide a clue, and so on. The CLAS experiment has produced the first higher than expected results. Because only a few pentaquarks, but they do provide a clue, and so on.

news@nature.com
The best of science journalism

Published online: 18 April 2005 | doi:10.1038/nature03411

Doubt is cast on pentaquarks

Mark Peacock

Exotic subatomic particles may be an experimental phantom.

Physicists have come home empty-handed from a thorough hunting expedition for pentaquarks. The lack of evidence has led some to doubt that these odd subatomic particles, first sighted in 2003, actually exist.

The pentaquark was discovered at the SPb-Synchrotron in Helsinki, Japan. The particles, thought to be made up of five quarks, is so unstable that physicists inferred its existence from the orders of collisions between gamma rays and carbon atoms.

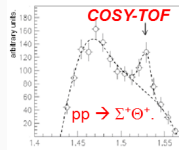
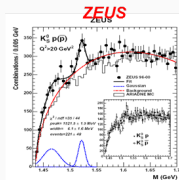
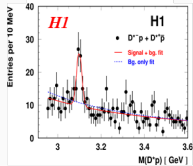
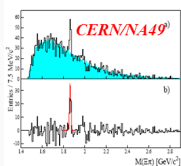
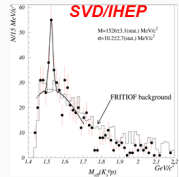
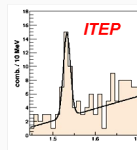
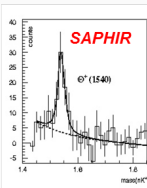
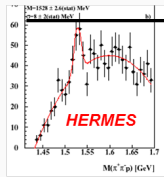
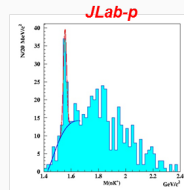
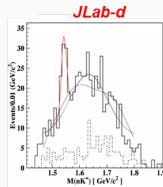
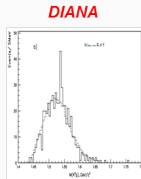
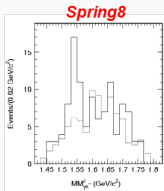
Trails of quarks make up protons and neutrons that are the basic building blocks of atomic nuclei. Particles made up of five quarks would be extremely unusual, and were listed as a new form of matter.

But experiments at the Thomas Jefferson National Accelerator Facility in Newport News, Virginia, now suggest that the discovery was a mistake. "We just didn't see it in three experiments," says Raffaella De Vita, of the National Institute of Nuclear Physics in Giardin, Italy, who presented the results at the American Physical Society conference in Tampa, Florida, on 16 April.

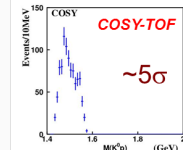
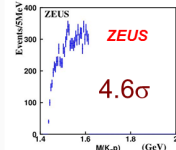
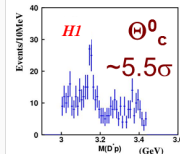
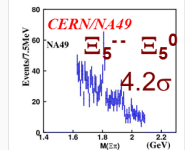
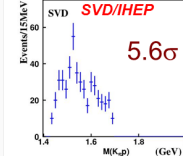
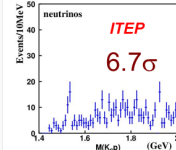
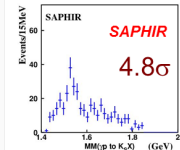
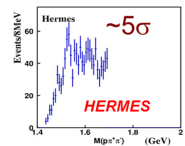
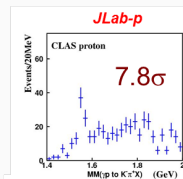
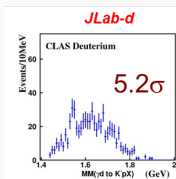
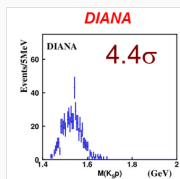
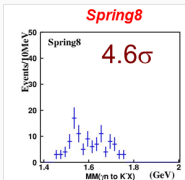


Trail of quarks make up protons and neutrons, but the pentaquark would be extremely unusual.

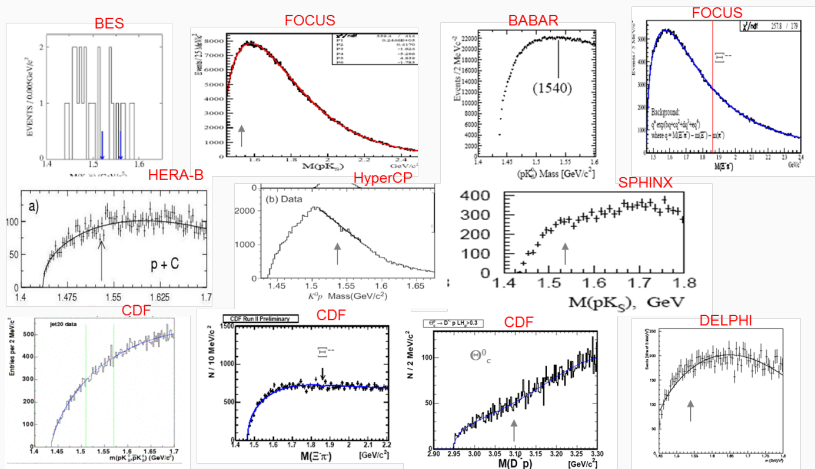
Positive Evidence



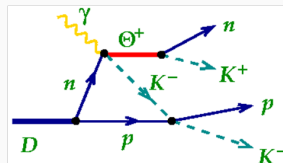
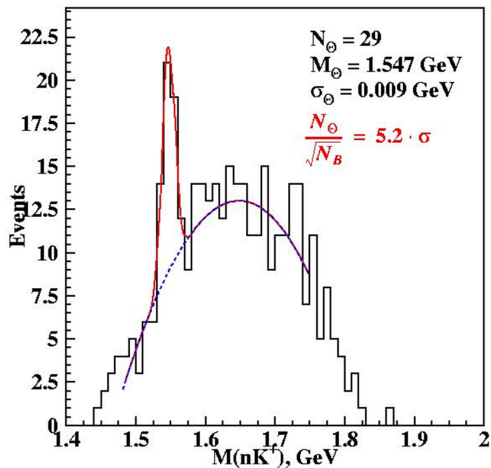
Data Undressed



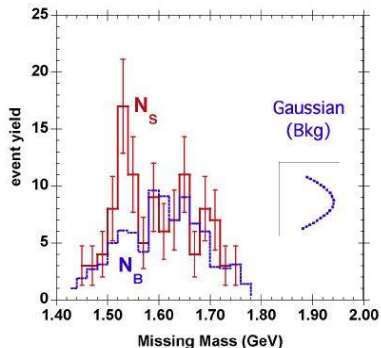
Non-evidence



Evidence from CLAS



Statistical Significance



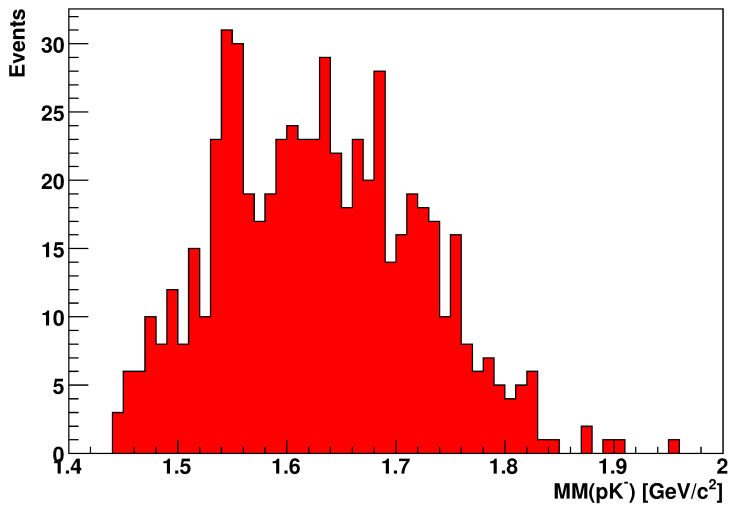
A simple statistic is

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

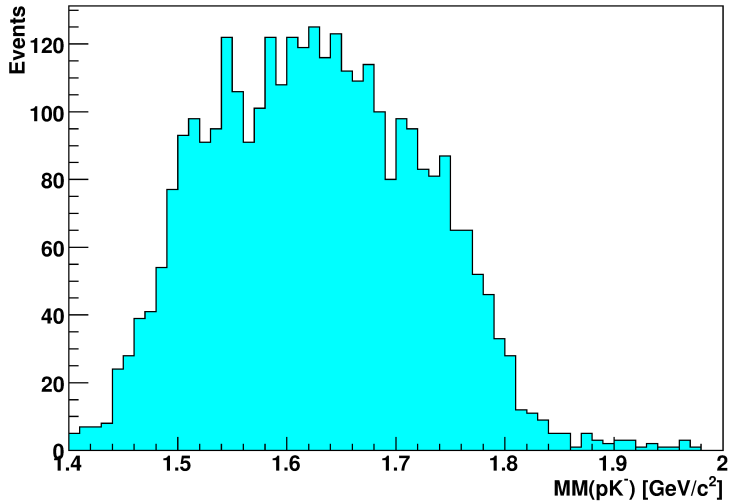
If this is > 5 then the probability of feature being a fluctuation is 10^{-6}

Do you want to bet £1M against £1?

First data

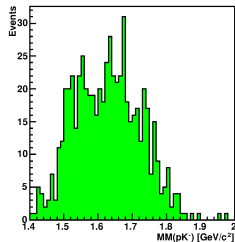


Higher Statistics Run

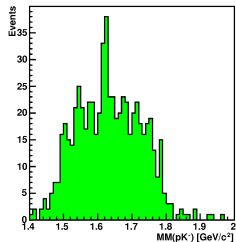


Samples from Higher Statistics Run

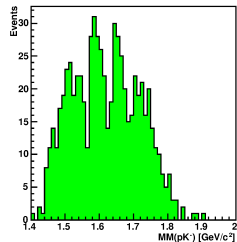
g10 sample 0



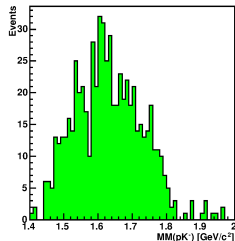
g10 sample 1



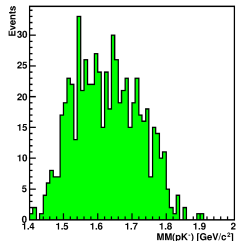
g10 sample 2



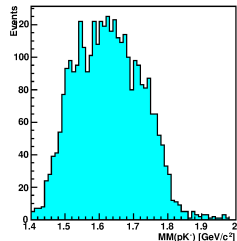
g10 sample 3



g10 sample 4



g10 full



Probability Distribution: Peak Height

Set up a data model M_0 for no signal:

$$S_0(x_i) = \beta f_b(x_i)$$

and a model M_P for a peak and background:

$$S_P(x_i) = \beta (f_b(x_i) + \rho f_s(x_i)),$$

where $f_b(x_i)$ and $f_s(x_i)$ are functions describing the shape of background and signal distributions. x_i are the histogram bin centres.

f_b depends on parameters c_j , and if f_s is modelled by a gaussian, it has parameters μ and σ .

The spectrum is a histogram with counts n_i in each bin i .

The **likelihood** of obtaining a spectrum $\{n_i\}$, given a set of parameters can be written:

$$P(\{n_i\}|\rho, \beta, \mu, \sigma, c_j, M_P) = \prod_i \frac{S_i^{n_i} \exp(-S_i)}{n_i!}$$

Probability Distribution: Peak Height

Bayes Theorem allows the definition of a joint (posterior) probability:

$$P(\rho, \beta, \mu, \sigma, c_j | \{n_i\}, M_P) \propto P(\{n_i\} | \rho, \beta, \mu, \sigma, c_j, M_P) P(\rho, \beta, \mu, \sigma, c_j | M_P),$$

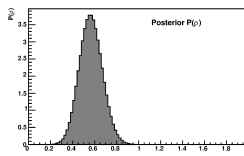
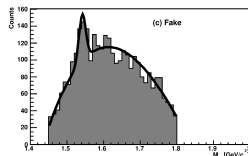
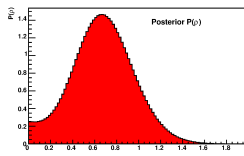
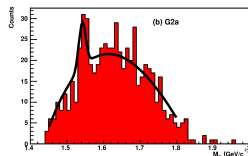
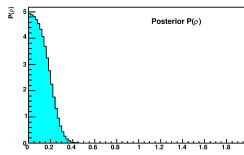
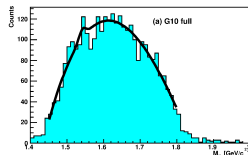
where the last factor is the **prior**.

Marginalisation allows the integration over parameters that are not of interest:

$$\begin{aligned} P(\rho | \{n_i\}, M_P) &= \int P(\rho, \beta, \mu, \sigma, c_j | \{n_i\}, M_P) d\beta d\mu d\sigma dc_j \\ &\propto \int P(\{n_i\} | \rho, \beta, \mu, \sigma, c_j, M_P) P(\rho, \beta, \mu, \sigma, c_j | M_P) d\beta d\mu d\sigma dc_j, \end{aligned}$$

Hence, evidence for a peak will be apparent in the distribution $P(\rho | \{n_i\}, M_P)$ showing a maximum at a non-zero value of ρ .

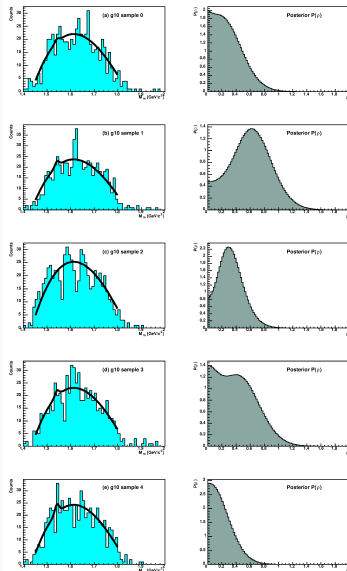
Probability Distribution: Peak Height



$P(\rho|\{n_i\}, M_P)$
distributions for

1. High stats
(original $\times 5$) run
2. Original run
3. Fake spectrum
(original $\times 5$
stats; keep
original signal to
background)

Peak Height

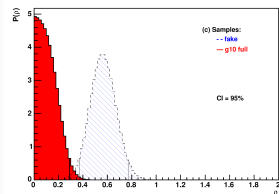
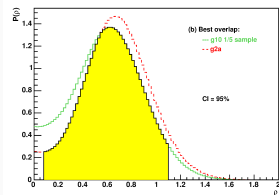
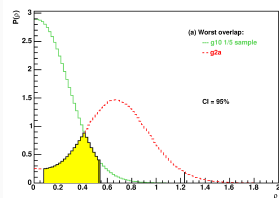


- Evaluate $P(\rho|\{n_i\}, M_P)$ distributions for the subsamples
- At least two show suggestion of non-zero peak heights.

Peak Height

Plots show overlaps of $P(\rho|\{n_i\}, M_P)$ distributions for combinations:

1. Original run and subsample with no peak
2. Original run and subsample with peak
3. High stats run and fake



Final thoughts...

- Inference based on information obtained from joint posterior probability, given a data model.
- Modern computing allows the use of Markov Chain Monte Carlo methods.
- Marginalisation over peak position $\equiv$ look elsewhere effect.
- Can use Bayes factors

$$\frac{P(M_P|\{n_i\})}{P(M_0|\{n_i\})}$$

- Alternatively, mixture models

$$S_M = \alpha S_0 + (1 - \alpha) S_P$$

can show which hypothesis is more likely

- Both Bayesian and frequentist methods should be deployed.