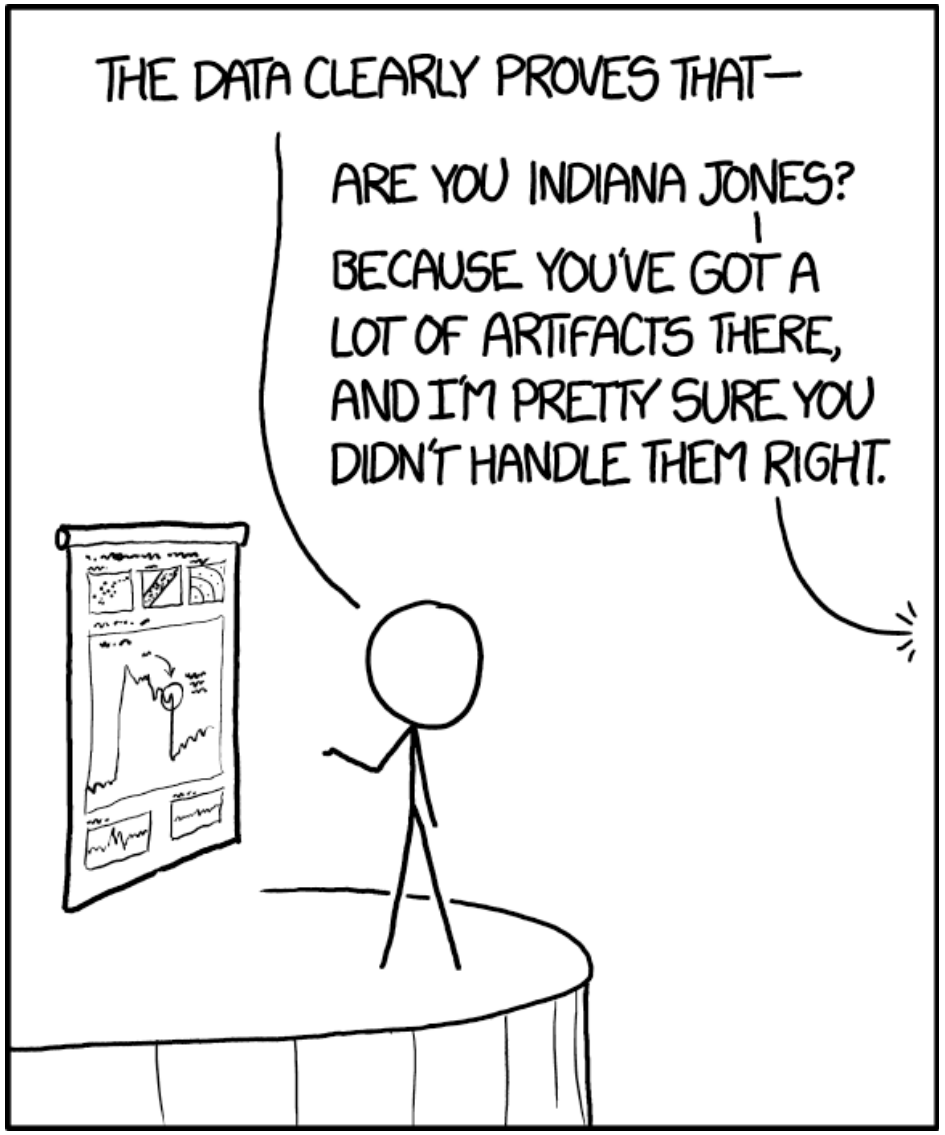
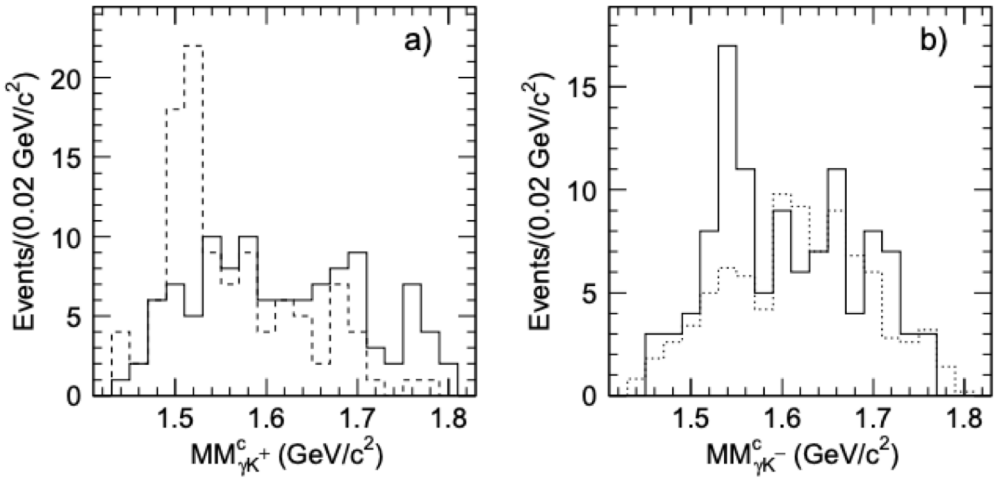
Blind analysis: People can convince themselves anything!

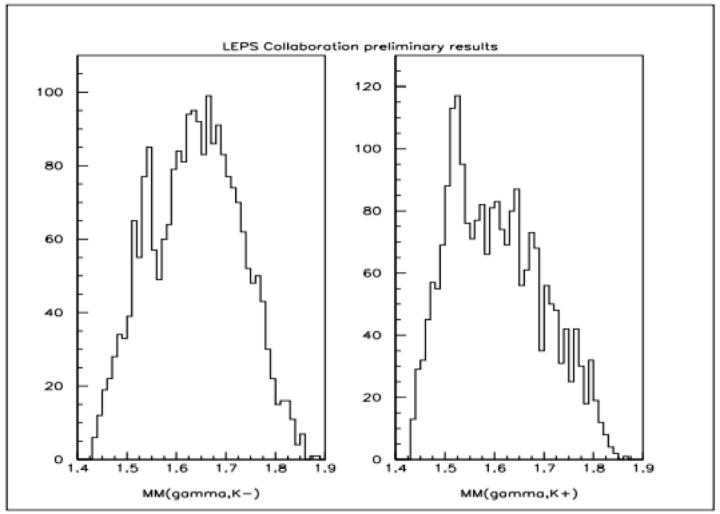


Blinding an analysis: real world implications in discovery

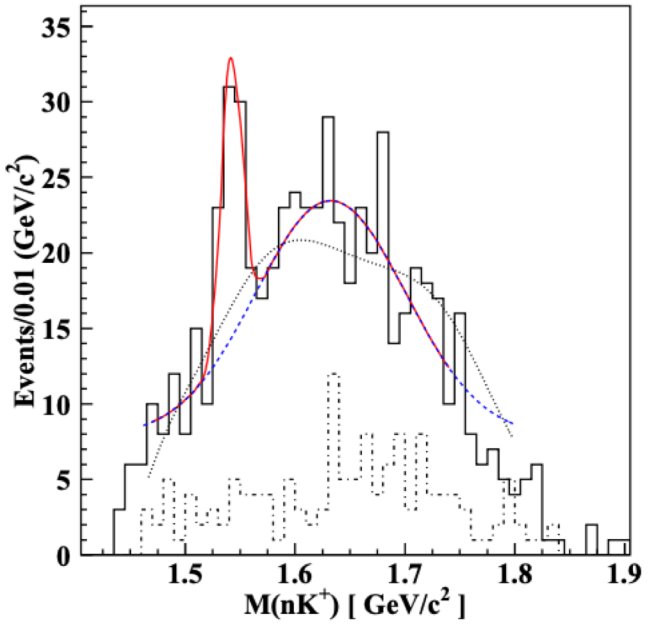
Pentaquark “discovery”

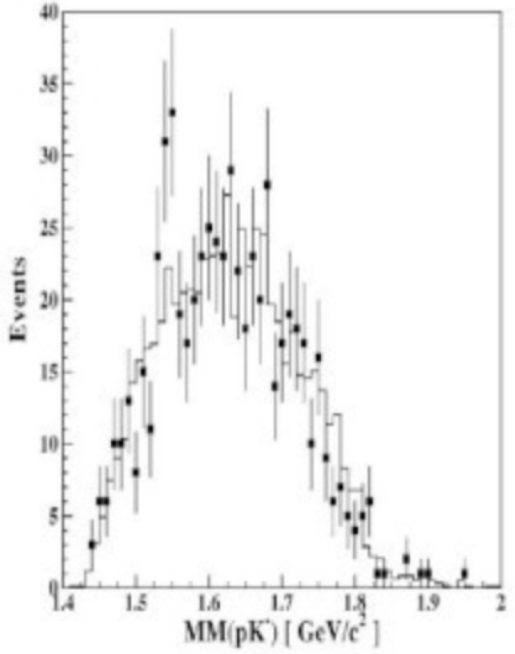
LEPS 1st results

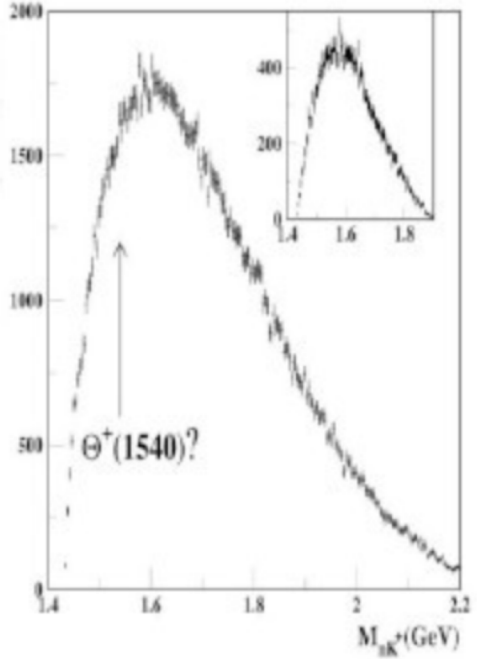




CLAS results

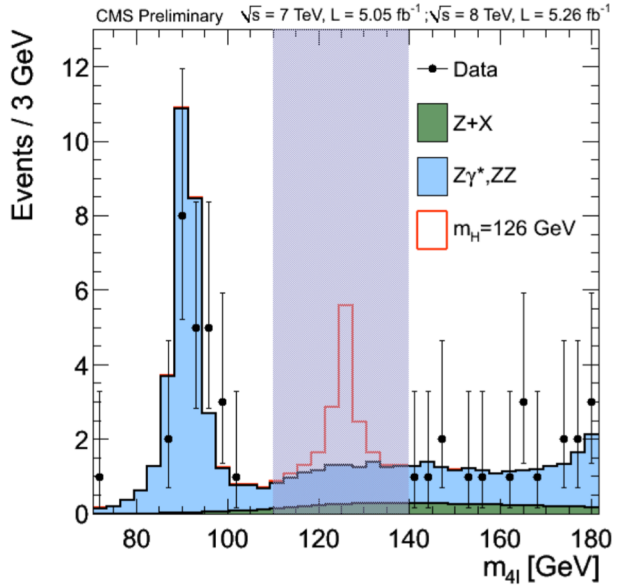






Blinding goals (for our purpose):

1. Fine-tune simulations, helps determine the criteria for selecting signal
2. Develop methodology for rejecting or quantifying background events using data from the region where there is no excess



Relevant techniques (list and combinations somewhat endless):

1. Black/hidden box method
   - Numerous examples: CMS/Higgs
   - must be able to predict background in the box

2. Add or remove events
   - spoil event count in unknown way (neutrino flux by SNO)

3. Data pre-scaling
   - removes statistical bias of tuning of cuts to enhance fluctuations
   - Pre-scaling done in unbiased way, assumes any data sample is the same as any other (reasonably sample)
   - Statistics must be big enough to id backgrounds and small enough not to bias the result of full set

Data pre-scaling, my experience:

- Tune MC to simulate resolutions and backgrounds in 10% of the data
- Tune cuts on simulation and verify against 10% of the data
  (how different do they look? And do we know why?)
- Generate large set of pseudo-data, insert fake signal, validate bump hunting procedure
- Document procedure in formal note, present to collaboration for review and approval to unblind

After unblinding:

- Quantify limits (Feldman-Cousins, Optimum Interval, …)
- Use pseudo-data to quantify Look Elsewhere Effect
- Systematics
- Publish!