Reproducibility In Particle Physics and possible applications to other sciences

Tony Weidberg (Particle Physics, Oxford University)
Talk Outline

- Science & Experiments
- Case Studies to illustrate past problems
  - Wrong claims for discovery of top quark & SUSY.
- Lessons learned
  - How to minimise chances of wrong claims of discovery
- Example of procedures
  - Higgs boson discovery.

- Critical issues and possible applications to other sciences
  - Discussion session not a lecture
  - Aim to provoke discussion.
  - I will act as scientific secretary: document consensus (if any!).
Science & Experiments

Very Naïve Philosophy of Science

• Complex links between experiment & theory but need both!
• “It doesn't matter how beautiful your theory is, it doesn't matter how smart you are. If it doesn't agree with experiment, it's wrong.” Richard P. Feynman
• Science requires reliable experimental results.
  – Discuss criteria for obtaining reliable results in this talk.

2 Way Interactions
Wrong Top Discovery

• Claim for top quark discovery:
  – “The two-jet signal has an over-all invariant mass clustering around the W mass, indicating a novel decay of the W”
  – “They are, however, consistent with the process $W \rightarrow tb$, where t is the sixth "top" quark of the Cabibbo current. If this is indeed the case, then the mass of the top is bounded between 30 and 50 GeV/$c^2$”.
• Now known to be wrong, top quark mass $\sim 175$ GeV
• Fooled by statistical fluctuation.
Wrong SUSY

• **Search for Supersymmetry (SUSY)** G. ARNISON, 1984, Phys. Lett B. 139, p115.
  – Supersymmetry allows for Weakly Interacting Massive Particles
  – WIMPs candidate for dark matter in universe.
  – WIMPs don’t interact in detector ➔ apparent violation of conservation of momentum.

• **Discovery claim:**
  – “We report the observation of five events in which a missing transverse energy larger than 40 GeV is associated with a narrow hadronic jet and of two similar events with a neutral electromagnetic cluster (either one or more closely spaced photons). We cannot find an explanation for such events in terms of backgrounds or within the expectations of the Standard Model.”

• **Result was wrong because backgrounds were not carefully evaluated.**
Right & Wrong Results in Particle Physics

• Several correct discoveries in 1980s-1990s:
  – Overwhelmingly statistically significant.
  – Confirmed by multiple experiments.
• A few wrong results
  – Not confirmed by other experiments
  – Some discovered to be wrong by better background calculations.
• Problem hasn’t gone away completely ...
  – e.g. BICEP 2 claim for B mode polarization of CMB.
  – DAMA claim to observe dark matter.
Rewards & Punishment

• Fame and glory for scientists who are first to make big discoveries
  – Also £££ in some areas

• Wrong discoveries can be published in high impact journals ➔ high citation index

• No sanctions for making wrong discoveries (unless fraudulent).
How to avoid mistakes at LHC

• Discussion between two general purpose experiments ATLAS and CMS before data taking.
  – High statistical power
  – Blind analysis
  – 5 sigma significance for discovery claim
  – Rigorous checking of results before publication
  – Need two independent experiments

• Illustrate this approach with Higgs discovery.
Big Data
~ 10 PetaBytes
Use Grid computing
~ 100,000 CPUs
Case study: Higgs boson discovery

- **Scientific procedures**
  - Statistical
  - Blind analysis
  - 5 sigma threshold for discovery

- **Levels of checking**
  - Low level x-checks
  - Sub-group
  - Working groups
  - Editorial Board
  - (Several further levels)
  - Collaboration
  - Refereed journals
  - Confirmation by another experiment.
Statistical Power Higgs Search

• Assume a value for the mass of the Higgs boson
• Detailed Monte Carlo for signal & background & detector response.
• Expected significance > 5\(\sigma\).
• If experiment works as expected, either confirm or reject Standard Model Higgs boson theory.

Expected mass spectrum for assumed \(m_H\) Bump above smooth background.
Blind Analysis

• Avoid finding spurious signals in very large data sets ➔ use blind analysis:
  – Monte Carlo simulations for signals and backgrounds
  – Optimise analysis (separation of signal from background) using Monte Carlo samples
  – Review analysis and then “open box” and look at data without changing analysis
  – Warning: this is a very simplified description!
Higgs Boson

- In the Standard Model of particle physics, Higgs boson gives mass to other elementary particles.
- Use high energy proton-proton collisions to try to produce Higgs bosons.
- Reconstruct decay products and use $E=mc^2$.
- Look for peak in mass spectra at mass of Higgs boson ($m_H$).
Higgs boson discovery

• One channel: \( H \rightarrow \gamma\gamma \) is the "bump" a statistical fluctuation or significant evidence for a Higgs boson?
Statistical Procedures

- Frequentist approach:
  - Define probability $p_0$ that if experiment were repeated infinite number of times that we would see a larger discrepancy with the no-signal model than in the actual data set.

- Combine all channels

- Plot $p_0$ vs Higgs mass ($m_H$)

- Look Elsewhere Effect $\rightarrow$ Global significance.

- $5 \sigma$ rule
Checking Results

• Internal
  – Low level x-checks
  – Sub-group
  – Working groups
  – Editorial Board
  – (Several further levels)
  – Collaboration

• Refereed journals

• Confirmation by another experiment.

Works well because of scientific culture in which everybody is encouraged to give critical feedback.
Applicability to other sciences (1)

• Don’t do low power experiments?
Applicability to other sciences (2)

• **Blind analysis**
Applicability to other sciences (3)

• High level of statistical significance for claiming discovery
  – $5\sigma$ may be very high but is 95% c.l. appropriate?
Applicability to other sciences (4)

• Internal checking by collaboration before publication
  – Requires healthy scientific culture in which junior PhD student can criticize results.
Applicability to other sciences (5)

• Confirmation of claim by at least one other independent experiment