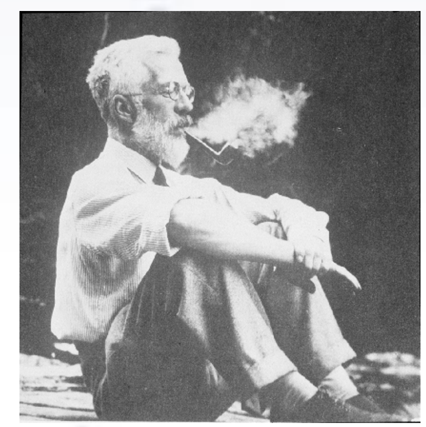
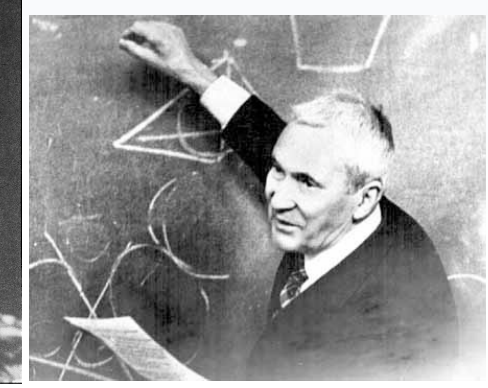
Applied Statistics

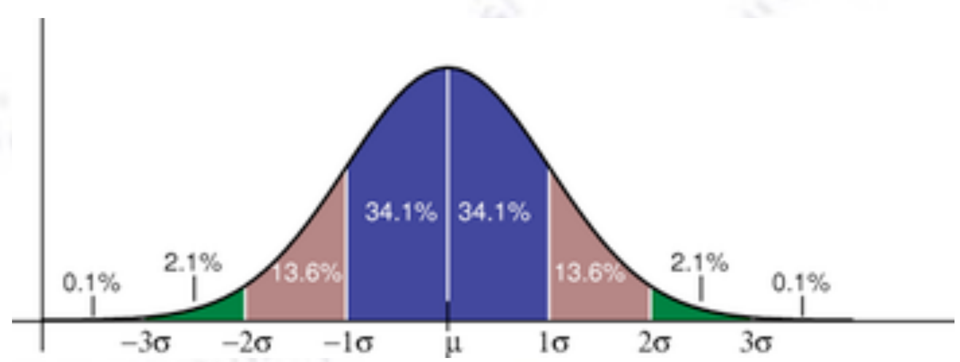
Systematic Uncertainties





Troels C. Petersen (NBI)



"Statistics is merely a quantisation of common sense"

Systematic uncertainties

"Everything is vague to a degree you do not realise till you have tried to make it precise."

[Bertrand Russell, 1872-1970]

Systematic Errors

Even with *infinite* statistics, the error on a result will never be zero!

Such errors are called "systematic uncertainties", and typical origins are:

- Imperfect modeling / simulation
- Lacking understanding of experiment
- Uncertainty in parameters involved
- Uncertainty associated with corrections
- Theoretical uncertainties / limitations

While the ***statistical uncertainty*** is Gaussian and scales like $1/\sqrt{N}$, the ***systematic uncertainties*** do not necessarily follow this rule.

Statistical errors are random, Systematic errors are not.

When **statistical** uncertainty is largest, more **data** will improve precision.
When **systematic** uncertainty is largest, more **understanding** will improve precision.

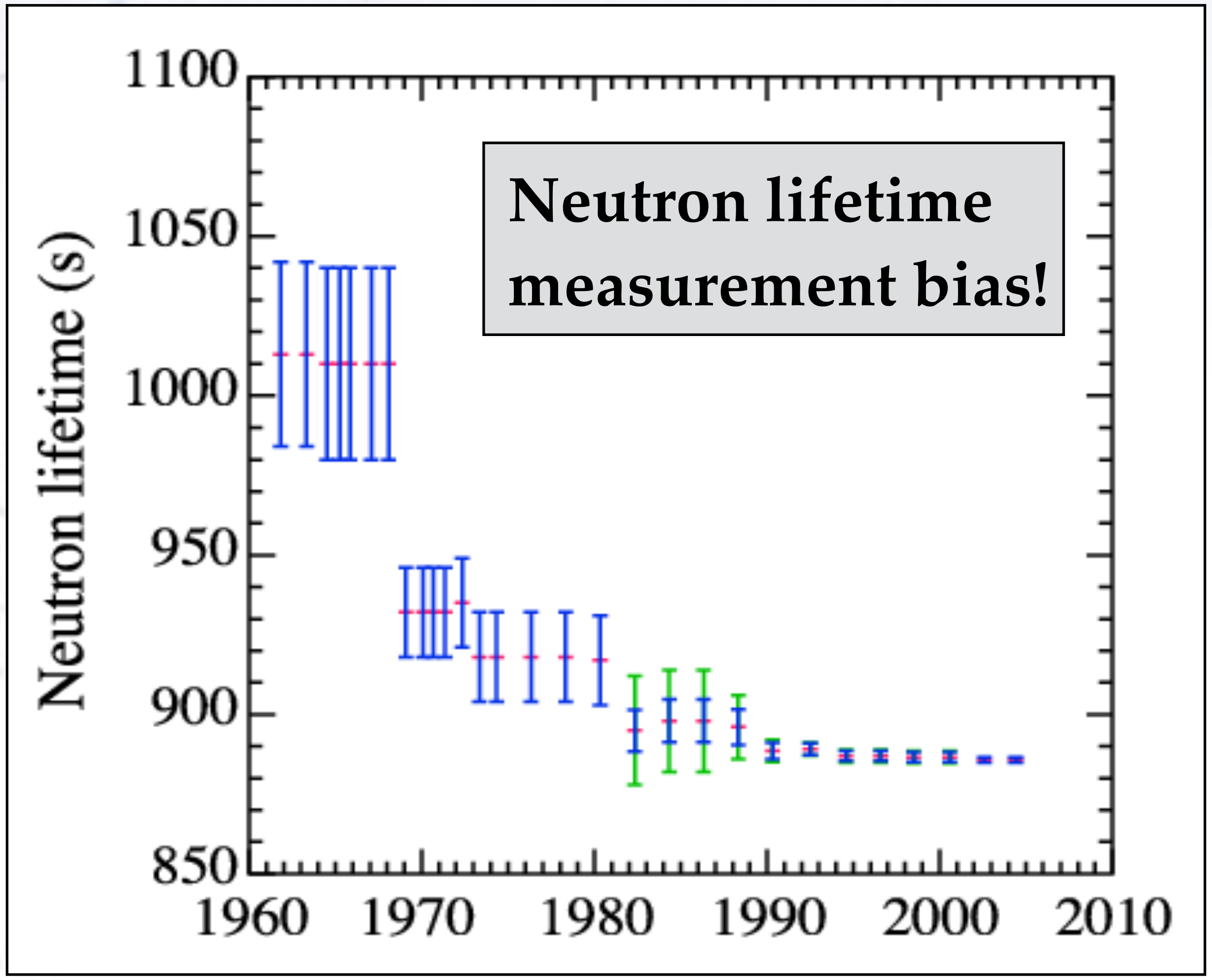
The finding / calculation of systematic errors is hard work.

Biased measurements

Why does my experiment find a lower value than others?

It is questions like these, that makes you start looking for effects that could yield a higher value, leading to…

Biases!



When measuring a parameter for which there are already expectations/predictions, the result can be biased. Examples:

- Millikan's oil-drop experiment.
- Epsilon prime (CERN vs. FNAL).
- Most politically influenced decisions!

Those who forget good and evil and seek only the facts are more likely to achieve good, than those who view the world through the distorting medium of their own desires. [Bertrand Russell]

The charge of an electron

***We have learned a lot from experience about how to handle some of the ways we fool ourselves**. One example: Millikan measured the charge on an electron by an experiment with falling oil drops, and got an answer which we now know not to be quite right. It's a little bit off because he had the incorrect value for the viscosity of air. It's interesting to look at the history of measurements of the charge of an electron, after Millikan. If you plot them as a function of time, you find that one is a little bit bigger than Millikan's, and the next one's a little bit bigger than that, and the next one's a little bit bigger than that, until finally they settle down to a number which is higher.*

*Why didn't they discover the new number was higher right away? It's a thing that scientists are ashamed of—this history—because it's apparent that people did things like this: **When they got a number that was too high above Millikan's, they thought something must be wrong—and they would look for and find a reason why something might be wrong.** When they got a number close to Millikan's value they didn't look so hard. And so they eliminated the numbers that were too far off, and did other things like that …*

[Richard Feynmann]

Blinding of results

To avoid experimenters biases, **blinding** has been introduced.

This means that the computer adds a random number to the result, which is not removed before the analysis has been thoroughly checked.

Example:

```
> ./FitSin2beta
Result is: sin(2beta) = x.xx +- 0.37
Do you wish to unblind (y/n)?
```

Emblem used by the BaBar experiment to label blinded analysis

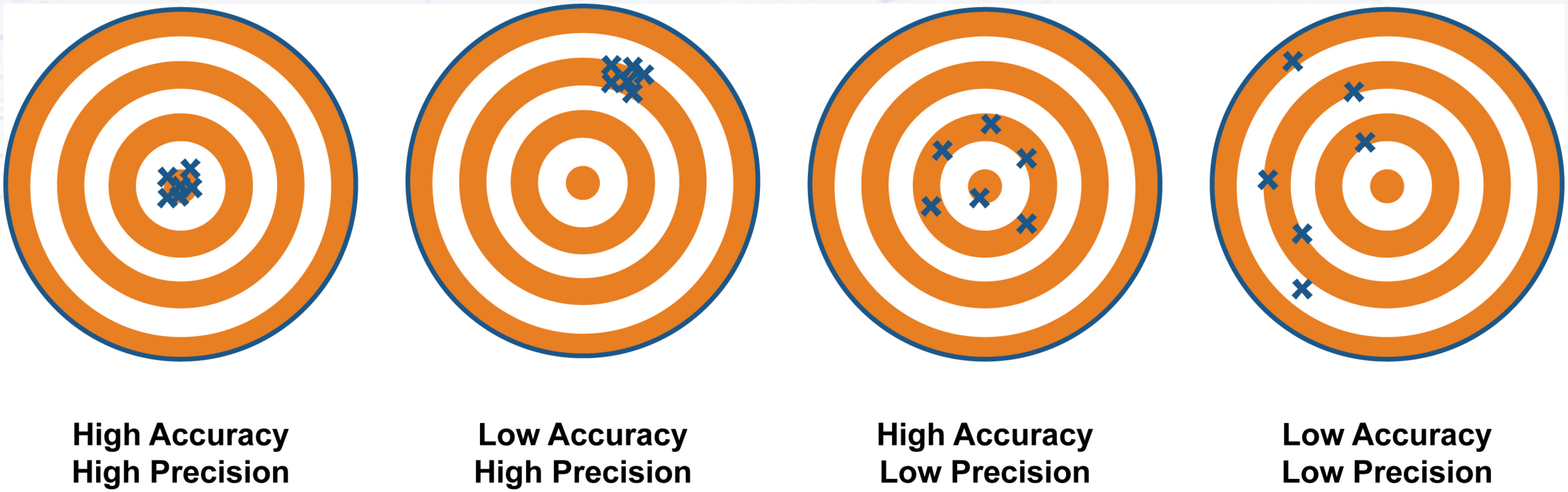
This was first introduced by the French Academy of Science (1784), and has since become standard procedure in most science and medical experiments.

In this way experimenters bias is removed, and the results become truly independent and unaffected by wishful thinking and "common belief".

How to find systematic errors?

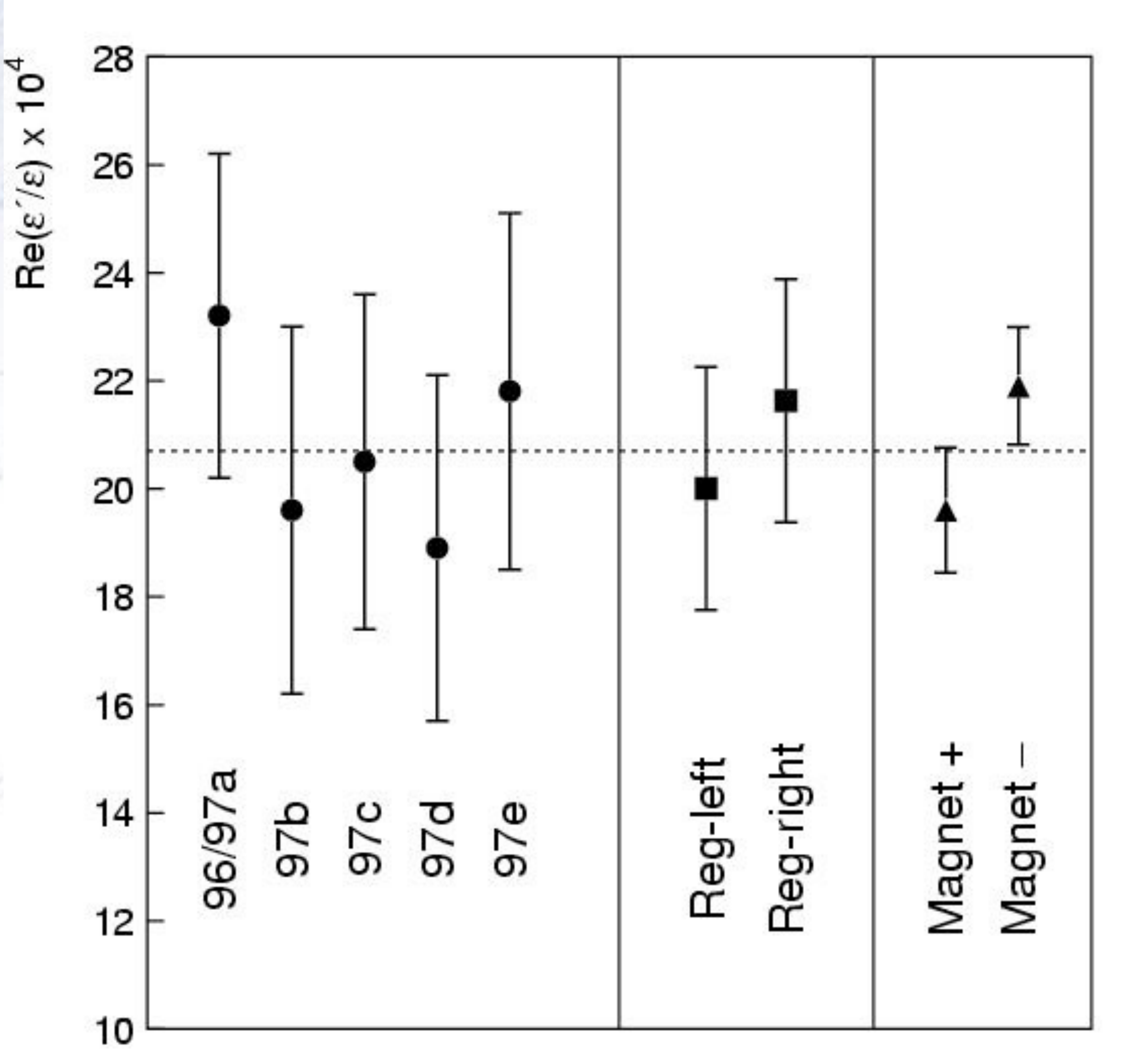
Look for ***ANY*** effect that can have an influence on your results.

Divide your data in any way you can (space, period, condition, analysis, etc.).



Often, systematic errors are also studied using simulation. However, this requires that the simulation is accurate! To check this, one studies known phenomena.

Cross check of data



Classic check of systematic errors, by dividing the data according to:

- Period of data taking
- Direction of regulator
- Direction of B-field

If any of these showed an inconsistency between the subsamples, one would know that this had an impact on the result.

This type of cross checks is at the heart of data analysis.

Example of systematic error

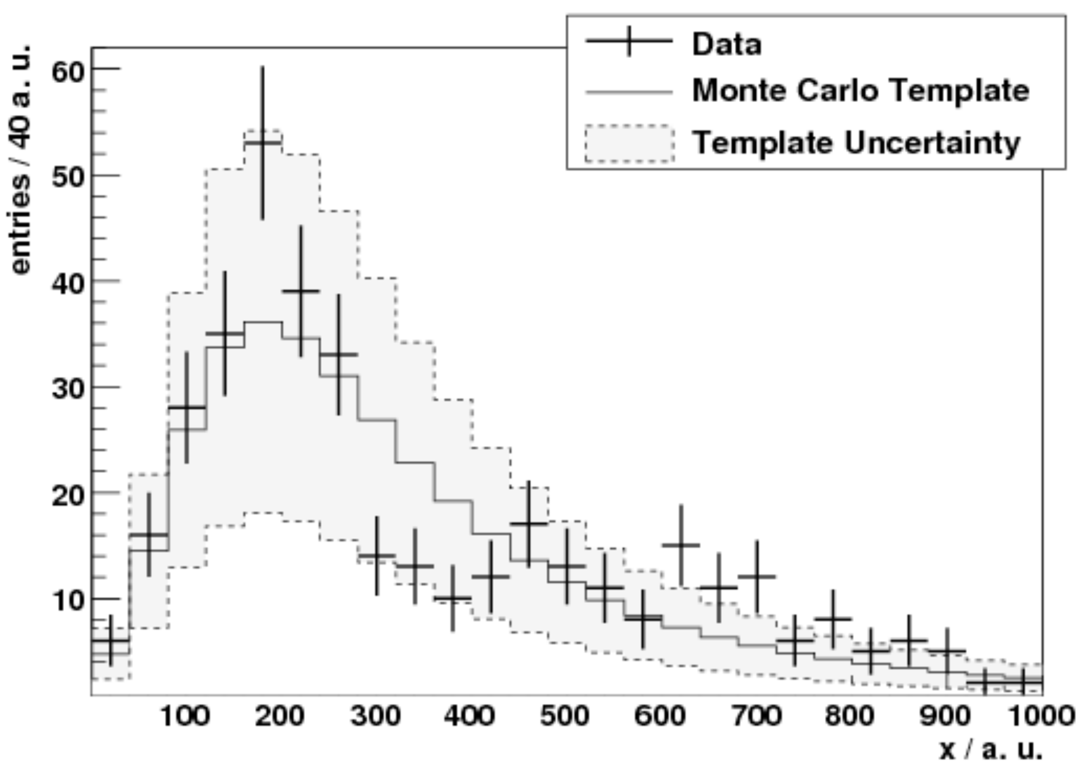
Measurements are taken with a steel ruler, the ruler was calibrated at 15°C, the measurements done at 22°C. This is a systematic **bias** and not only a systematic **uncertainty**! To neglect such an effect is a systematic **mistake**.

Effects can be corrected for! If the temperature coefficient and lab temperature is known (exactly), then there is no systematic uncertainty.

If we correct for effect, but corrections are not known exactly, then we have to introduce a systematic uncertainty (error propagation!).

A sign of a systematic error (or bug), is that one can see in data, that "something" strange is going on.

One should of course work hard to understand the effect, but occasionally one must give up, and suffer a large systematic uncertainty.

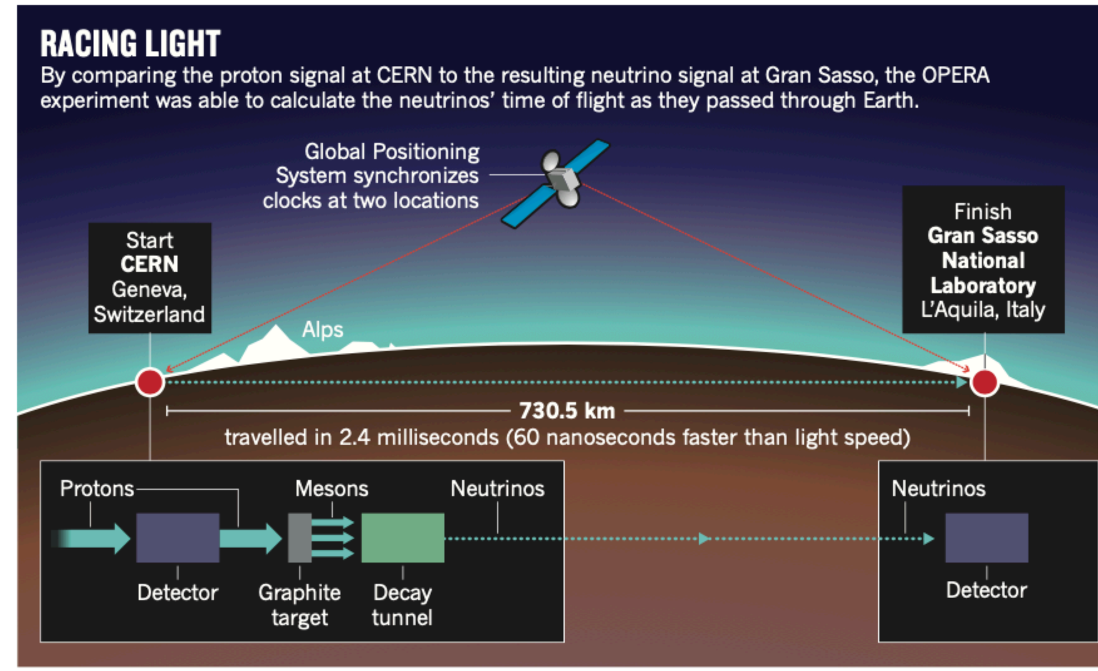


(Another) Example of systematic error

One of the best "recent" examples is the case of physicists measuring neutrinos to travel faster than speed of light.

This would (if true) put the foundations of physics in ruins…
After 6 months of intense studies, the researchers found two possible systematic errors:

- A link from a GPS receiver to the OPERA master clock was loose, which increased the delay through the fiber.
- A clock on an electronic board ticked faster than its expected 10 MHz frequency, lengthening the reported flight-time of neutrinos, thereby somewhat reducing the seeming faster-than-light effect.



PARTICLE PHYSICS

Speedy neutrinos challenge physicists

Experiment under scrutiny as teams prepare to test claim that particles can beat light speed.

BY EUGENIE SAMUEL REICH

The joke begins with the barman saying: "I'm sorry, we don't serve neutrinos." Then the punch line: a neutrino walks into a bar.

Such causality-bending humour has been rife on the Internet in the past week, following the news that an experiment at the Gran Sasso

worth of physics upended, starting with Albert Einstein's special theory of relativity. This sets the velocity of light as the inviolable and unattainable limit for matter in motion, and links it to deeper aspects of reality, such as causality.

Physicists, for the most part, suspect that an unknown systematic error lies behind OPERA's startling result. But nothing obvious has emerged, and many see the experiment as

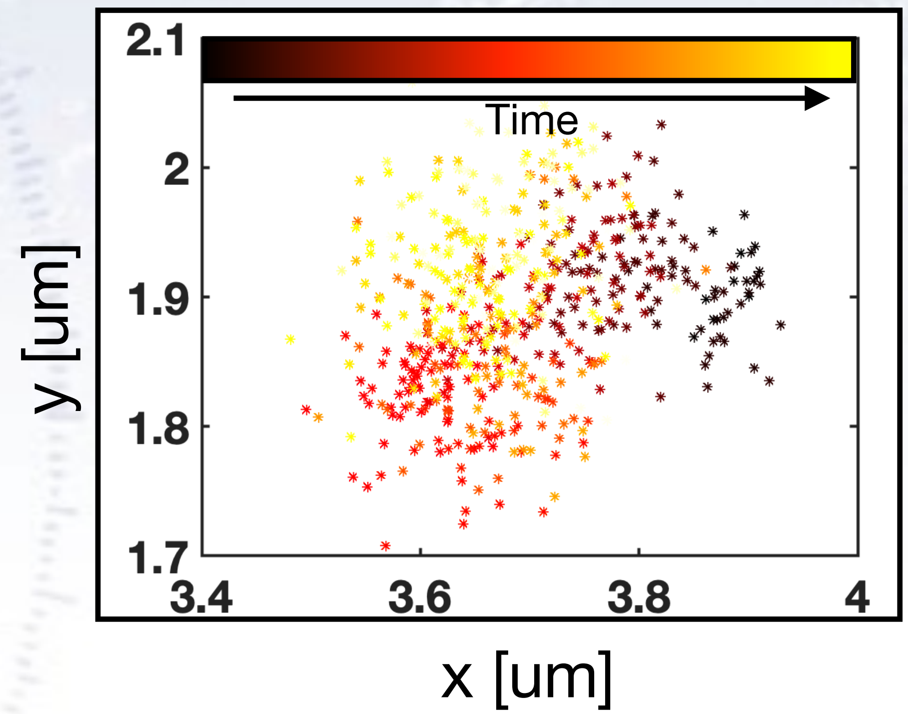
(Another) Example of systematic error

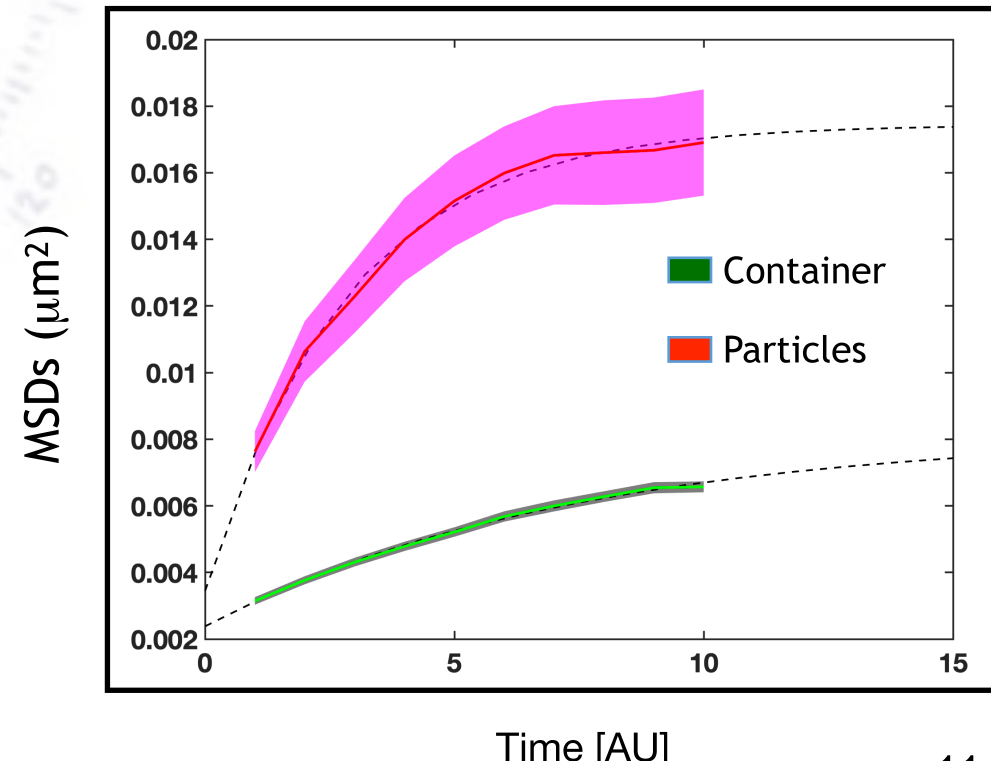
Imagine you have a set of measurements (trapped particles) , and you want to measure the size of the container.

You look at them and think you can just measure their circumference and use that as an estimator.

Next you realise that the container is not constant in time! This leads to a serious overestimation of that container.

In order to resolve this, you need to come up with new ways on analysing the data with methods that do not assume constant position of that container.





Evaluating systematic errors

Known sources:

- Error on factors in the analysis, energy calibration, efficiencies, corrections, ...
- Error on external input: theory error, error on temperature, masses, ...

Evaluate from varying conditions, and compute result for each. Error is RMSE.

Unsuspected sources:

Repeating the analysis in different form helps to find such systematic effects.

- Use subset of data, or change selection of data used in analysis.
- Change histogram binning, change parameterisations, change fit techniques.
- Look for impossibilities.

If you do not a priori expect a systematic effect and if the deviation is not significant, then do not add this in the systematic error.

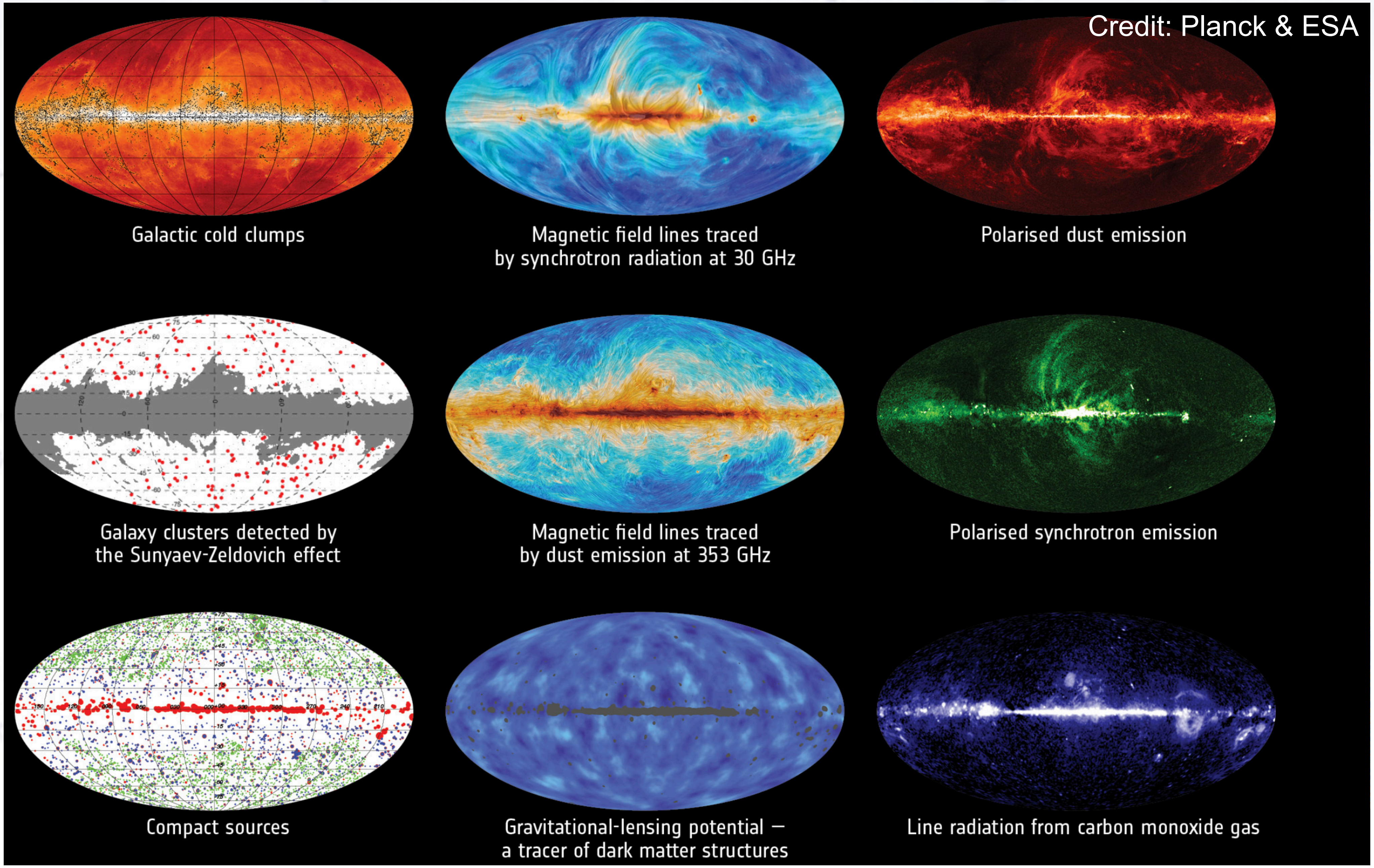
If there is a deviation, try to understand, where the mistake is and fix it!

Only as a last resort include non-understood discrepancy as systematic error.

Unchecked biases

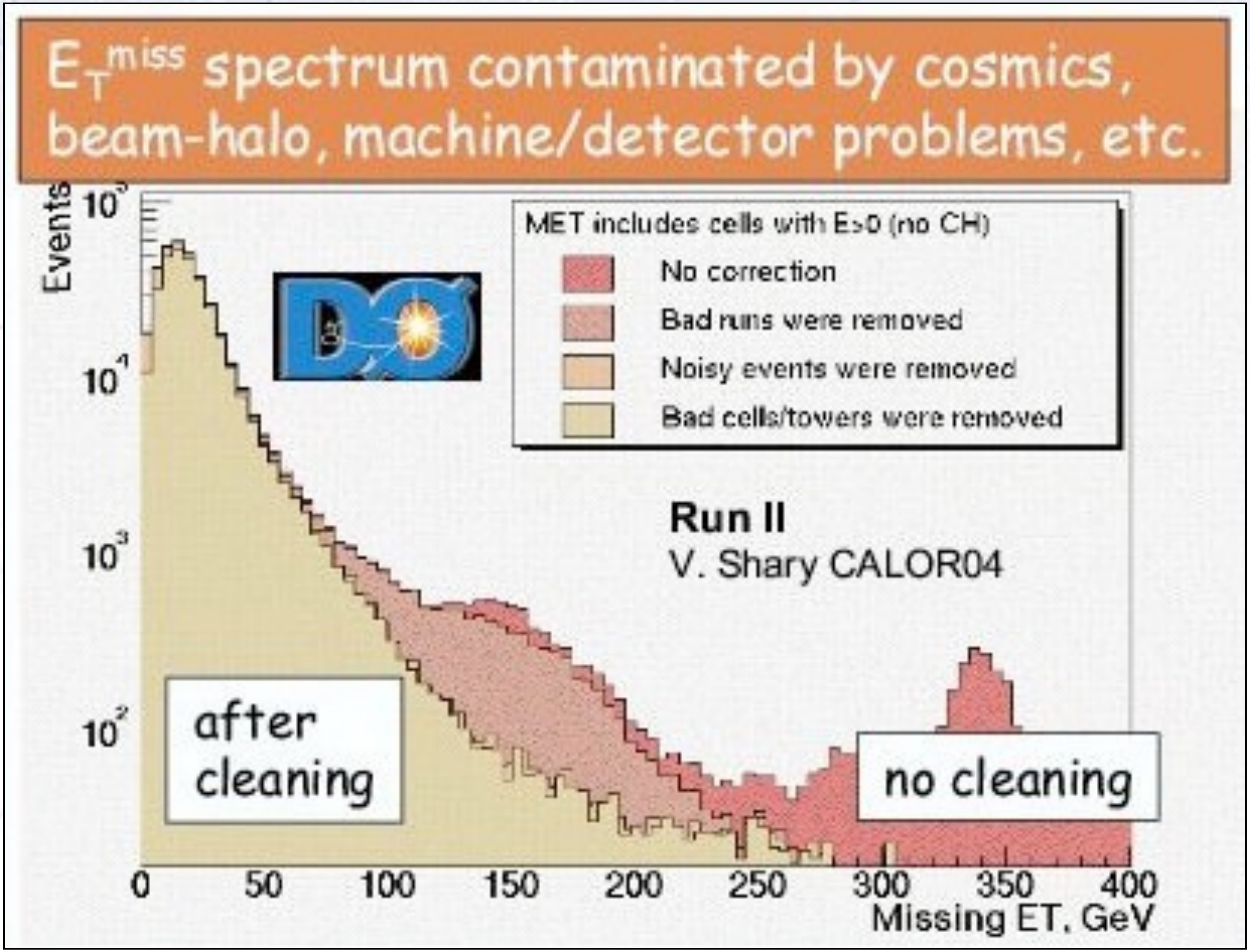
No method of checking for biases or systematics errors is fool proof. Overconfidence that all dominant systematic errors are included can result in wrong results.

Measuring the cosmic microwave background requires many subtractions of unwanted foregrounds. Missing a single systematic contribution ruins results.



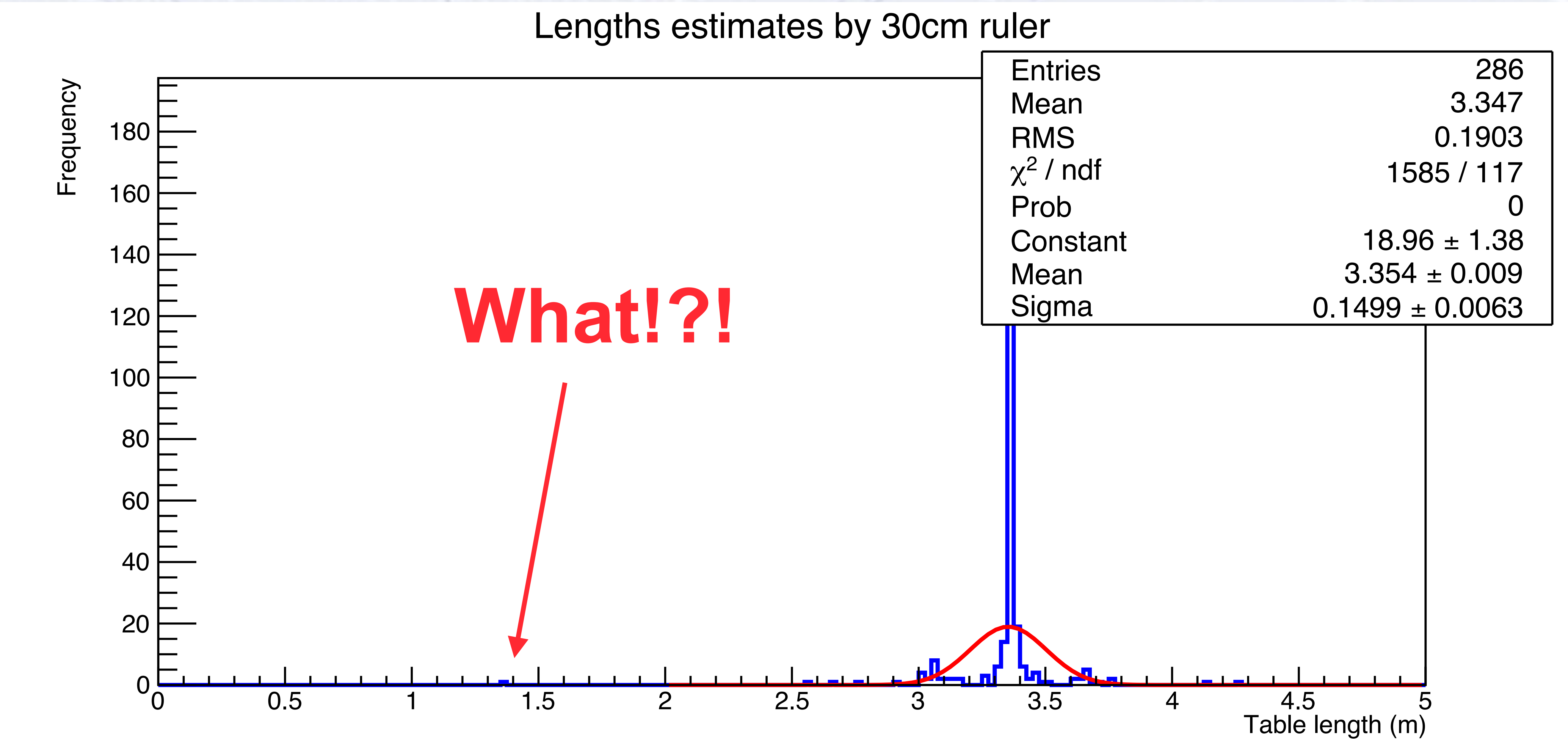
Cleaning data

Example of experimental error, which would be a disaster if not corrected for.



Removing data points

An example could be in some of the Table Measurement exercise data…



Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

4.3 Combination of errors 55

The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history.

If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

4.3 Combination of errors 55

The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history.

If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

Yes!

Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history.

If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

Yes!

Hmm…

Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

4.3 Combination of errors 55

The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history.

Yes!

If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

Hmm…

Yes!

Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that you can catch these rogues at an early stage, before their origins get lost in the mists of history.

If you cannot find an obvious mistake, then you probably have no choice but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

Yes!

Hmm…

Yes!

YES!

Caution discarding data!

The following passage (p. 55 of Barlow) is an intersting read:

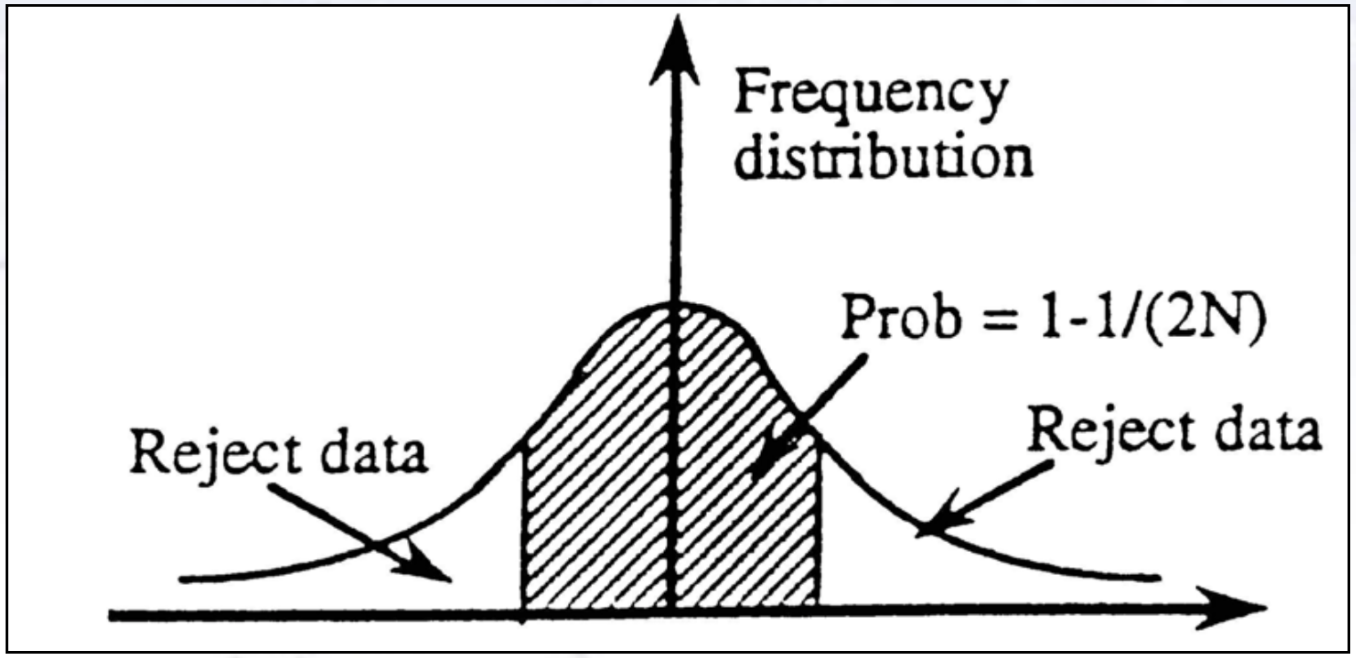
> The first thing to do is go back as far as you can and check the readings. You are very likely to find a misplaced decimal point, or a pair of numbers transposed in the notebook. If you can easily retake the measurement then this should be done—and the moral is to plot your points as you go, so that **Yes!** you can catch these rogues at an early stage, before their origins get lost in the mists of history.
>
> If you cannot find an obvious mistake, then you probably have no choice **Hmm…** but to throw the point away. However you should always do so with reluctance. If you have several such points, and/or if there are more points than you would expect with large ($>2\sigma$) deviations, then you should be extremely suspicious, as there is probably some effect at work that you do **Yes!** not understand, and you should understand. It is usually a trivial matter, but it could be something new and fundamental. Distrust all algorithms that **YES!** advise the automatic rejection of points outside certain limits as they can rapidly get out of hand; points should only be condemned after giving them a fair hearing.

You have to be the good judge!

Removing data points

One should always be careful about removing data points, yet at the same to be willing to do so, if very good arguments can be found:

- It is an error measurement.
- Measurement is improbable.



Removing improbable data points is formalised in **Chauvenet's Criterion,** though many other methods exists (Pierce, Grubbs, etc.)

The idea is to assume that the distribution is Gaussian, and ask what the probability of the farthest point is. If it is below some value (which is preferably to be determined ahead of applying the criterion), then the point is removed, and the criterion is reapplied until no more points should be removed.

However, **ALWAYS keep a record of your original data**, as it may contain more effects than you originally thought.

Removing data points

An example could be in some of the Table Measurement exercise data…