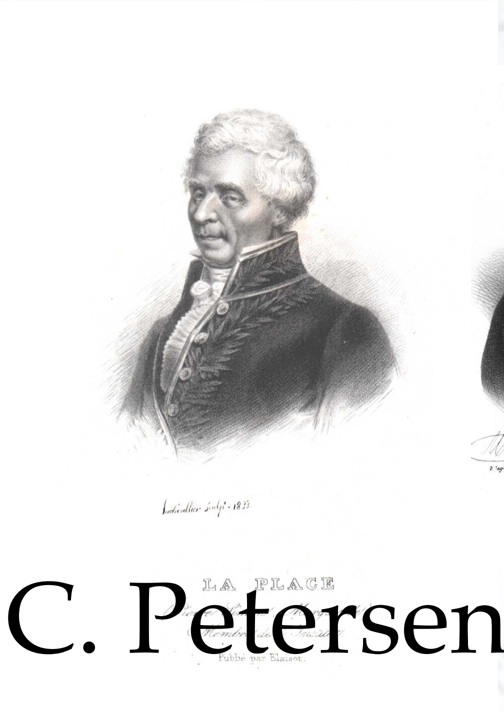
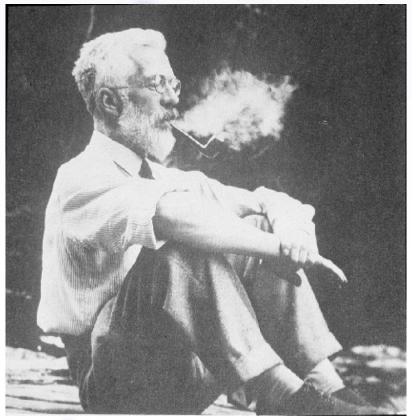
Applied Statistics

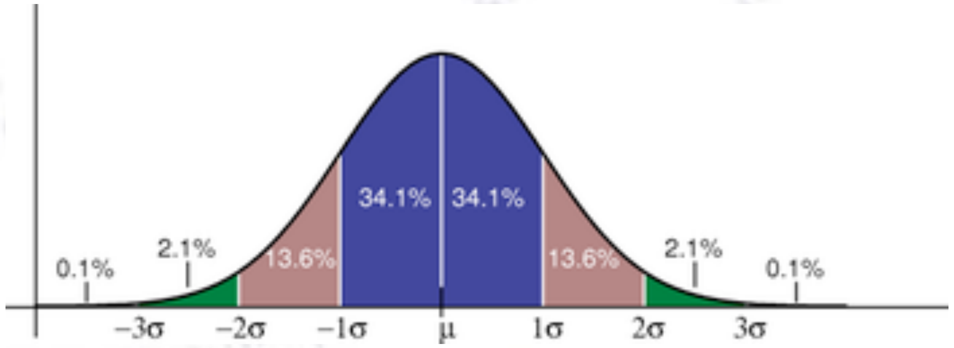
Systematic Uncertainties








Troels C. Petersen & Jason Koskinen



"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.

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.

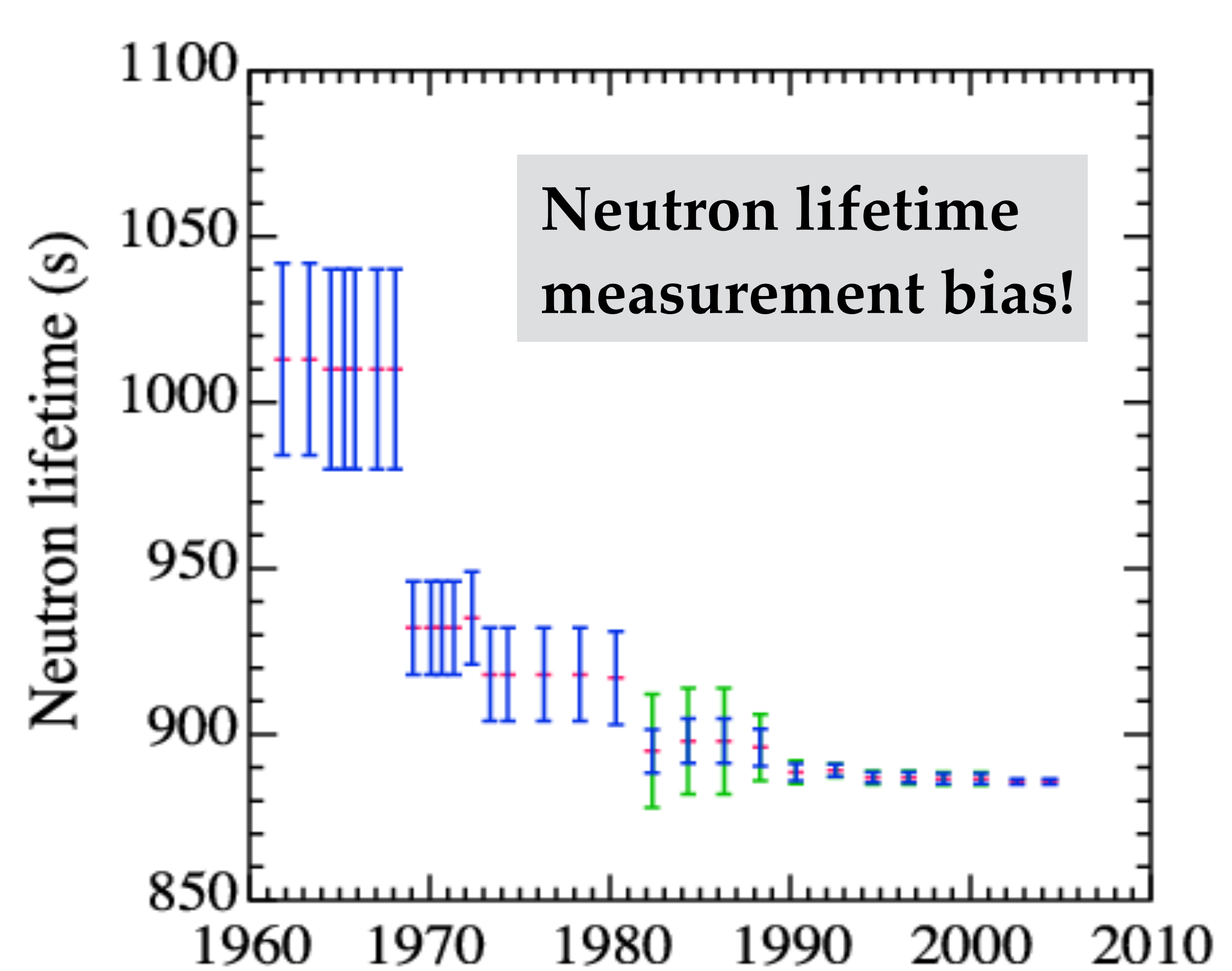
Everything is vague to a degree you do not realise till you have tried to make it precise. [Bertrand Russell]

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]

Blinding of results

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

One method is for a computer to add 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)?
```

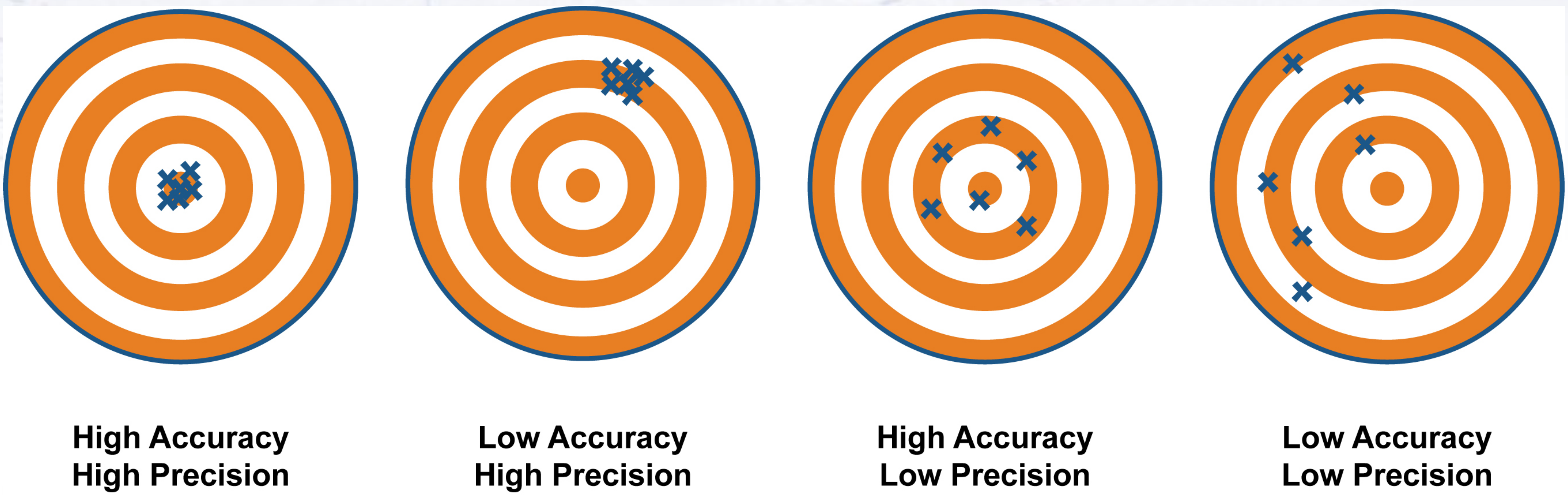
This was used in the epsilon-prime measurements, and has since become standard procedure in many particle physics experiments.

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.

How to find systematic errors?

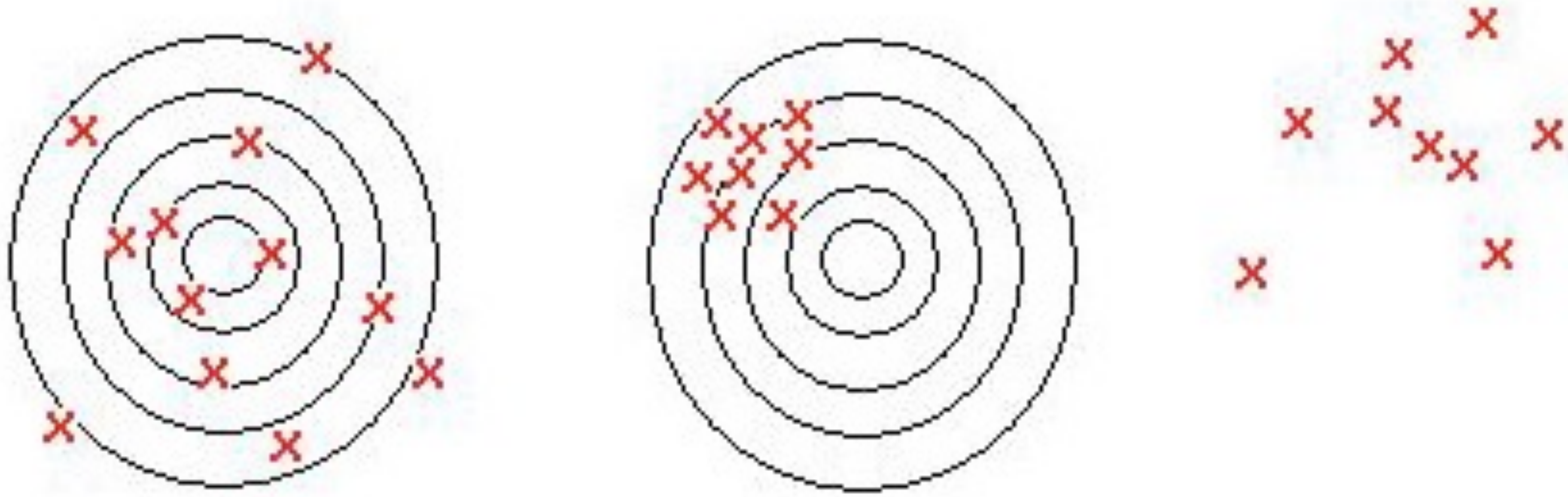
Look for *ANY* effect that can have an influence on your results.

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

Large statistical error
Small systematic error

Small statistical error
Large systematic error

Medium stat. error
??? syst. error



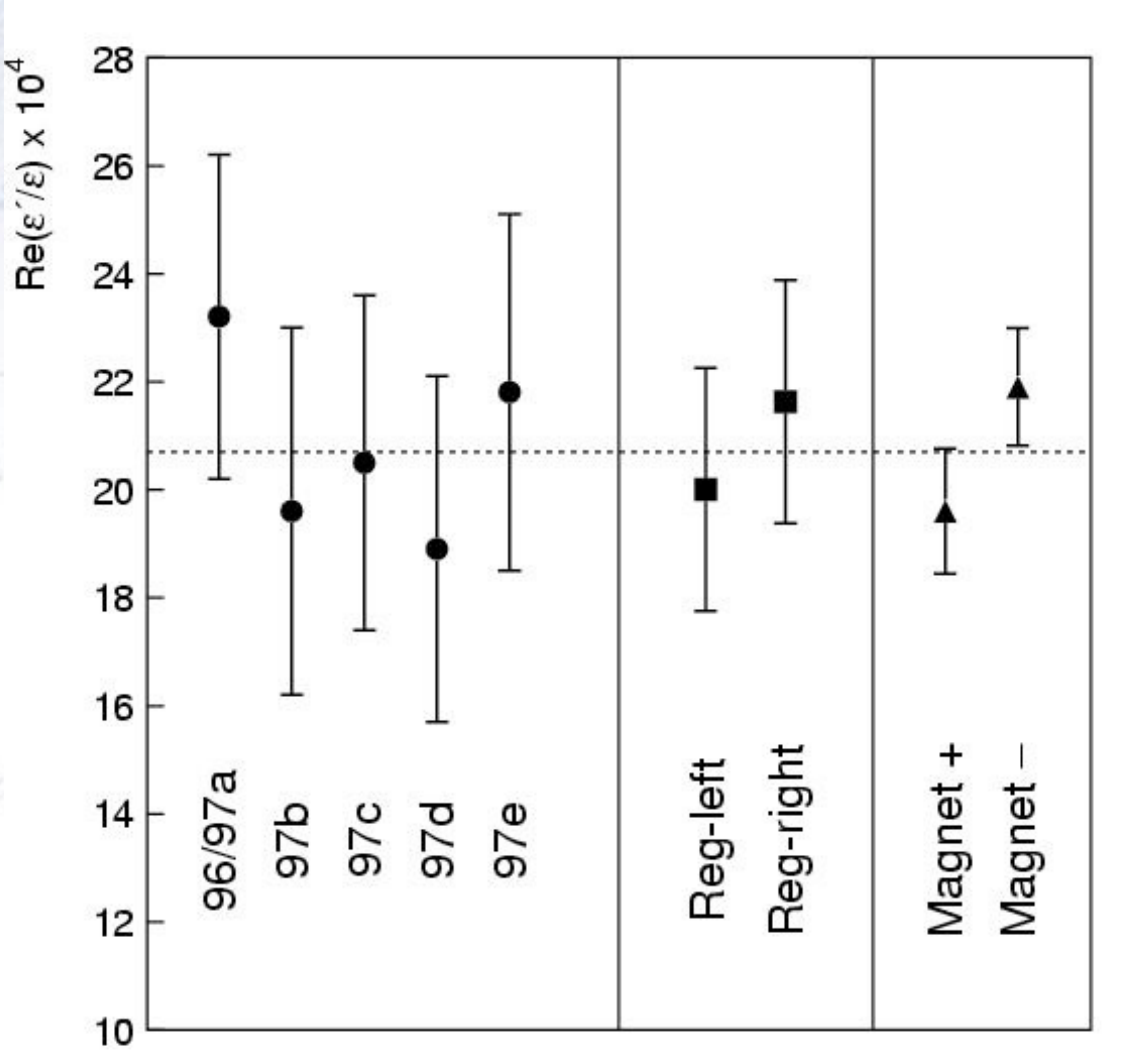
Not **precise**,
but **accurate**

Not **accurate**,
but **precise**

Medium **precise**,
Accurate???

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 measurement is done at 22 C.

This is a systematic **bias** and not only a systematic **uncertainty**! To neglect this effect is a systematic **mistake**.

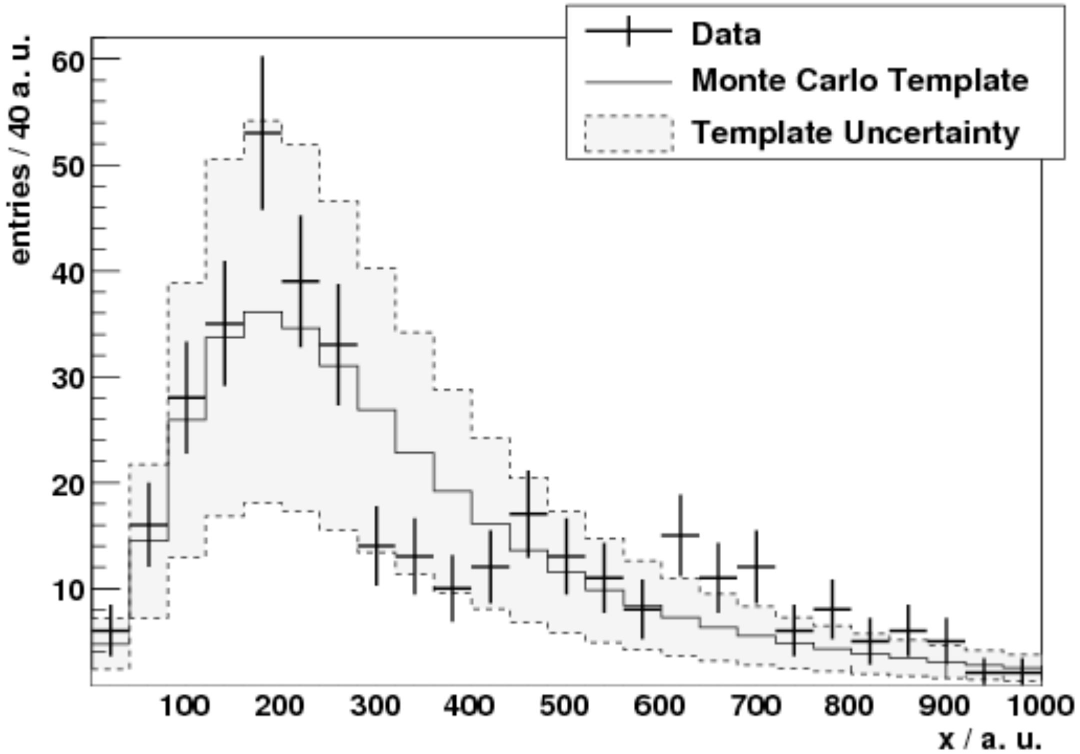
Effects can be corrected for! If the temperature coefficient and lab temperature is known (exactly), then there is no systematic uncertainty.

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

In practice (unfortunately): Often not corrected for such effects, but then just "included in sys. uncertainties".

Often, 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.



Evaluating systematic errors

<u>Known sources:</u>

- 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 RMS.

<u>Unsuspected sources:</u>

Repeating the analysis in a 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.

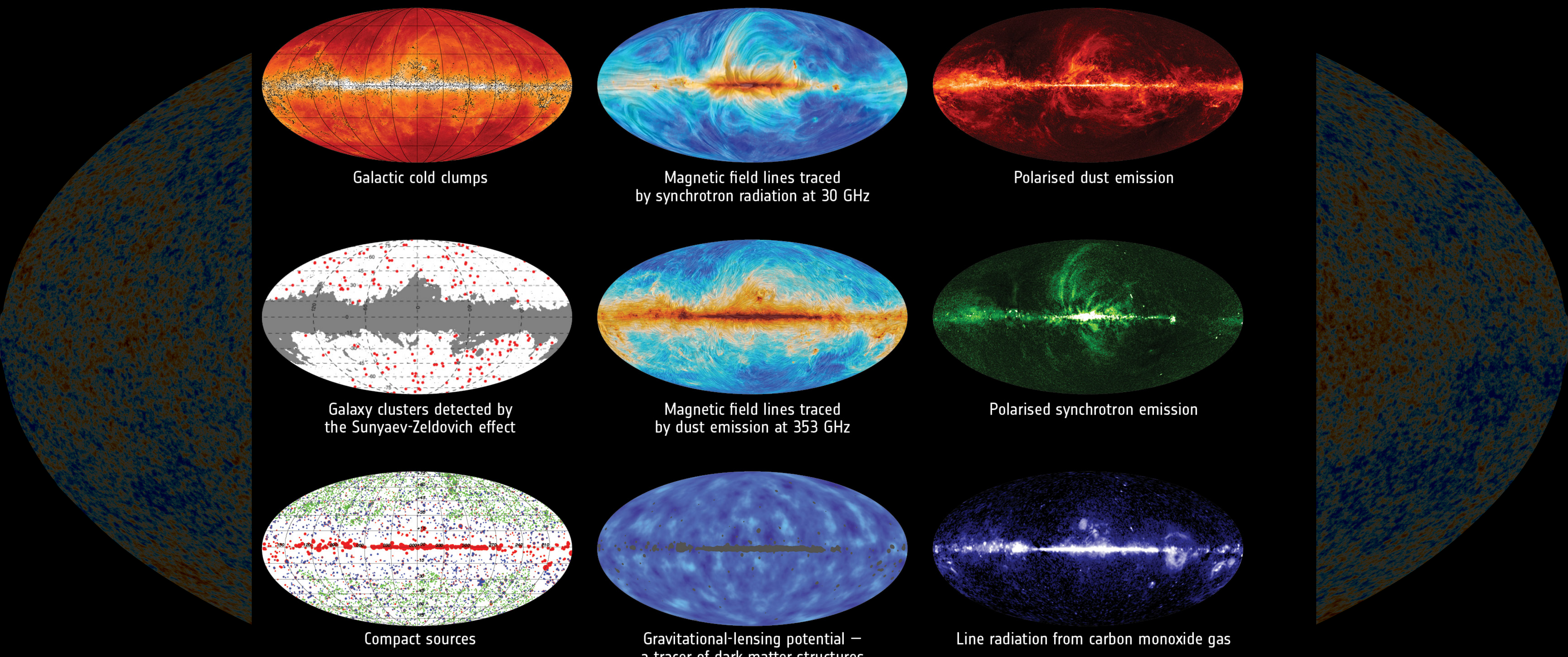
If there is a deviation, try to understand where the mistake is and fix it!

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

Unchecked biases

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

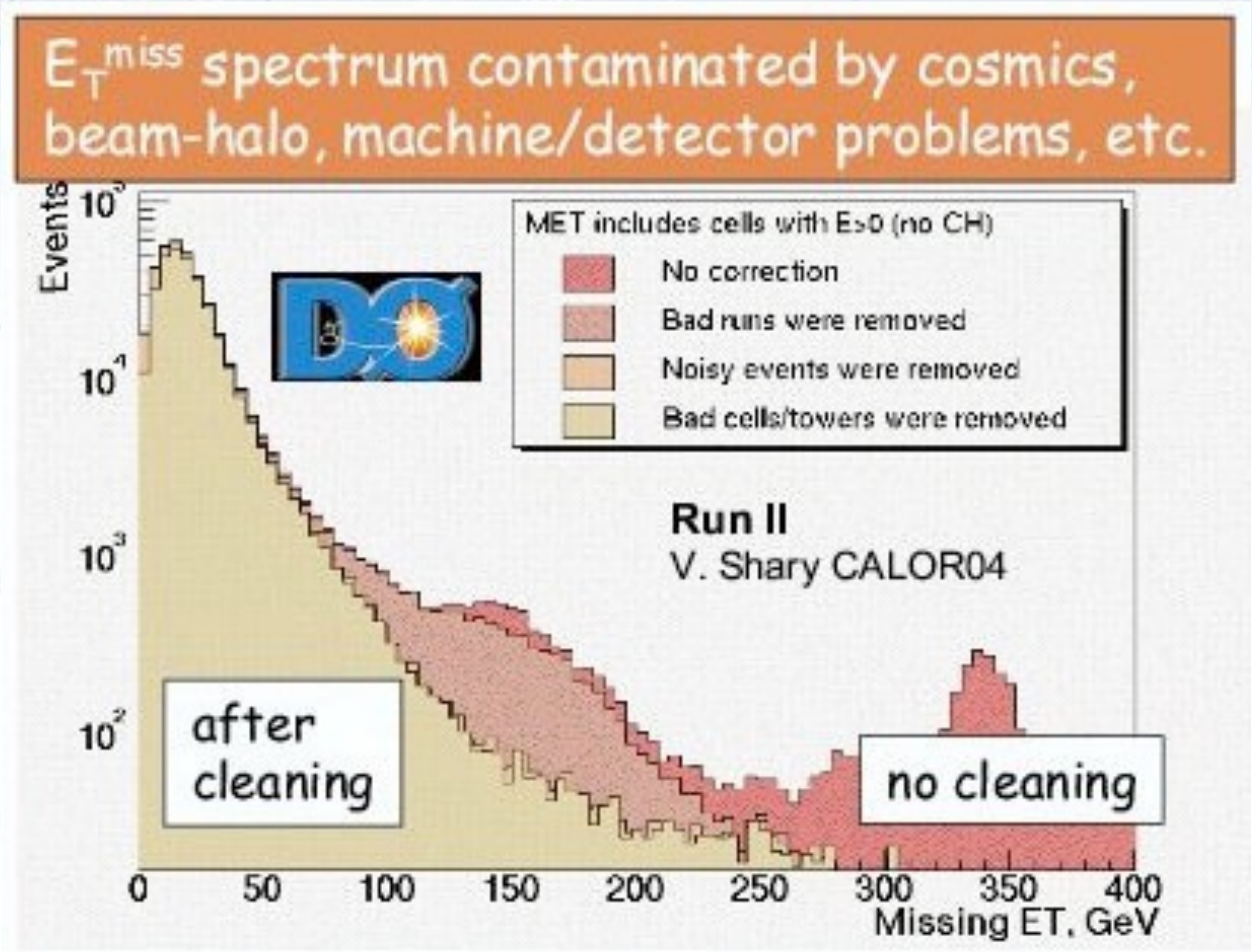
Measuring the cosmic microwave background requires many subtractions of unwanted foregrounds. Missing a single systematic contribution ruins results.



Credit:ESA & Planck

Cleaning data

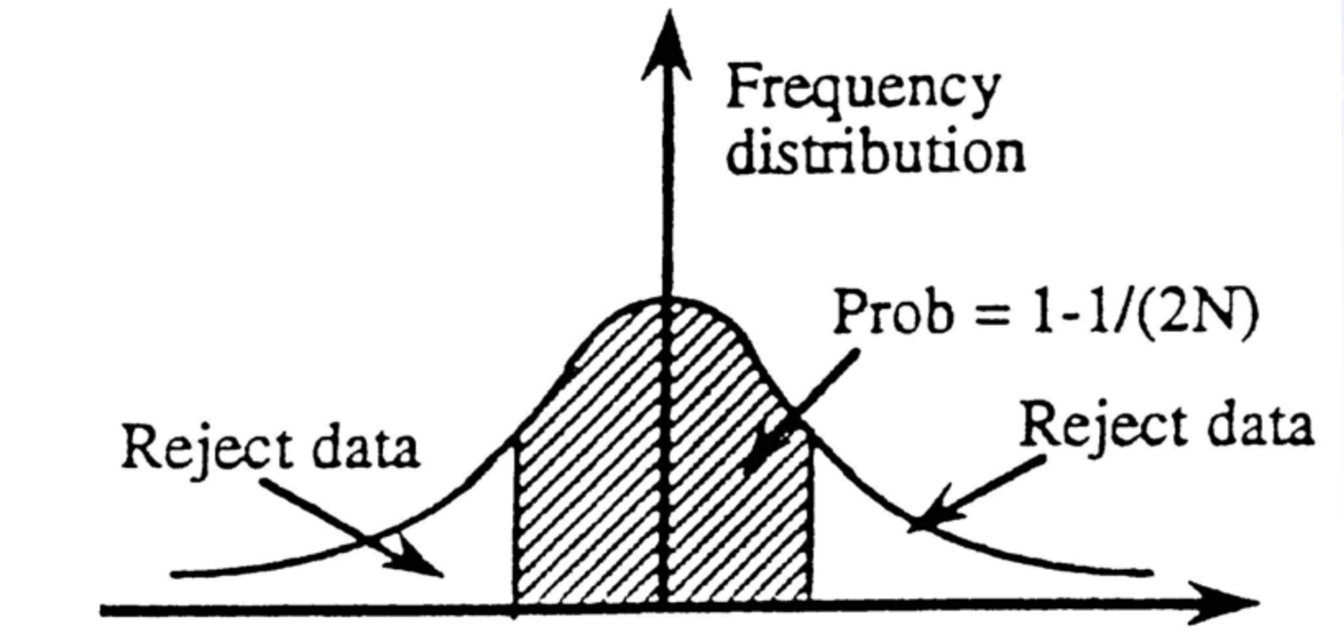
Example of experimental error, which would be a disaster if not corrected for.



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.



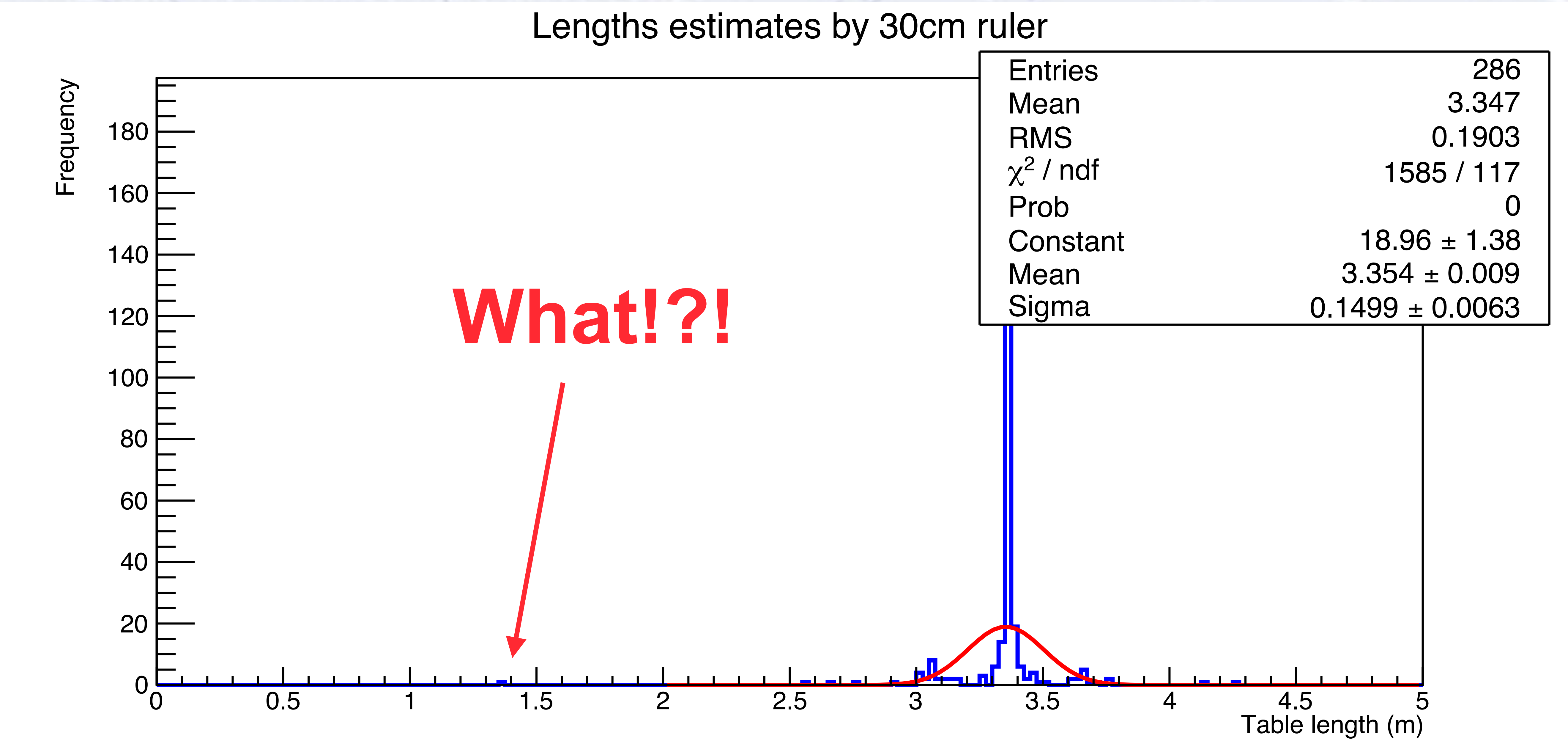
Removing improbable data points is formalised in **Chauvenet's Criterion,** though many other methods exists (Pierce, Grubbs, etc.)

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 today's data…



The good experimenter

The good experimenter

<u>The good experimenter will always:</u>

- inspect data visually.
- test assumptions.
- keep an accurate record.
- perform cross checks.
- do a ChiSquare test (also).
- plan the experiments carefully.
- try to "blind" results until final.

<u>The good experimenter will never:</u>

- rely on untested assumptions.
- "just let someones program do it".
- make changes in data.
- look for only some effects.
- not look at the raw data.