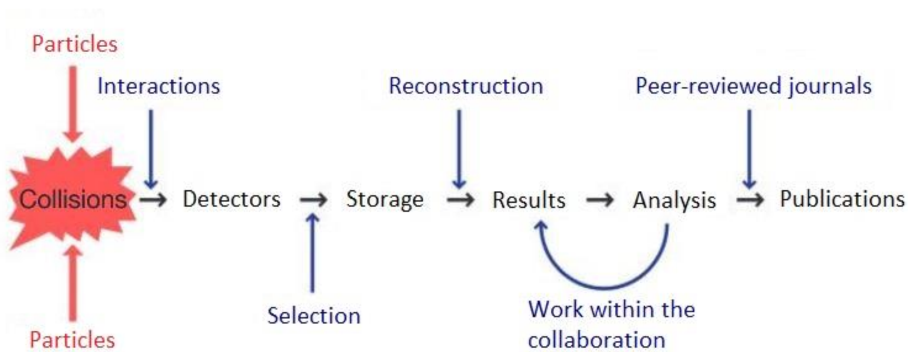
International Particle Physics Outreach Group

PASSPORT TO THE INFINITE

From Detection to Measurement

Analysing data from a particle physics detector requires a thorough understanding of the measuring instruments, the writing of complex computer programmes, and a command of cutting-edge mathematical tools.



Credit: N. Arnaud

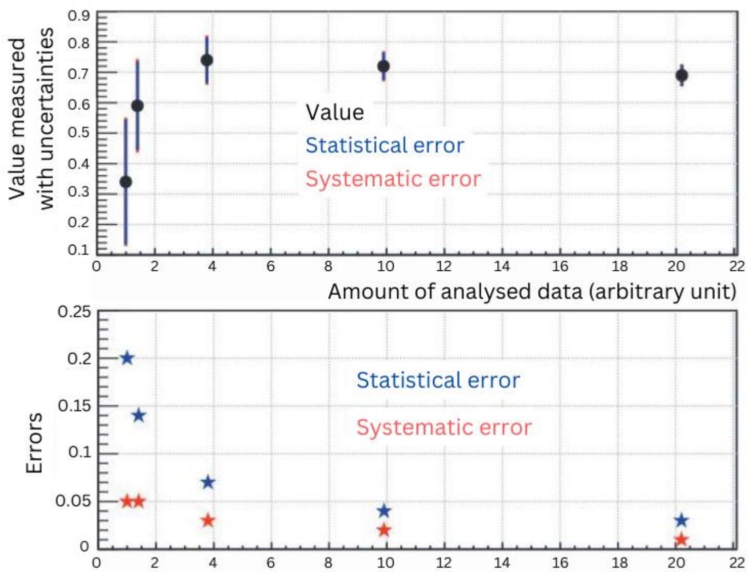
The different stages in a data analysis

Many stages separate particle collisions in a detector from the publication of physics results. Once selected, the relevant events are stored on a hard drive, interpreted (reconstruction stage), then analysed extensively. This latter stage can last a long time: the worth of the analysis has to be proven within the collaboration before the results are made public.

In subatomic physics we try to measure the properties of particles, such as their masses or the way they decay, as accurately as possible. A successful measurement cooks up like a good dish. You need quality ingredients – the particles, supplied by the accelerator-producer. The utensils – detectors – are extremely performant and known down to the finest detail.

As for the recipes, they constantly evolve according to the scientists' requests and the instruments' performance. Some improve results that have been achieved elsewhere while others venture into unchartered territory, guided by theoretical calculations. As soon as a new study is made public, it is reviewed by researchers from all over the world, eager to understand it and assess its quality. The same requirements apply internally: as long as a collaboration – i.e. all the people working on a same experiment, whose number can reach several thousands – is not convinced by a result, the product does not leave its 'kitchen'!

Despite their complexity and variety, physics analyses all follow the same stages. Upstream, an accelerator prepares large quantities of particles that will collide at the centre of detectors,



producing new particles that will eventually be detected. Measuring instruments are structured as Russian dolls, with a series of complementary devices from the centre (collision point) outward. Their function is to measure the properties (energy, speed, mass) of the particles travelling through them, enabling us to identify particles and trace back the physical processes that led to their existence.

Despite the gigantism of today's detectors, they are still designed for a single purpose: to understand the phenomena occurring at their centre during collisions. Their data is processed by powerful computers and stored in large hard drives that can be accessed from all around the world via the Internet. IT is everywhere: each measurement requires great computing power, which is shared between hundreds of networked machines. The few searched-for events – the signal – have to be separated from all the others – background noise. A measurement (e.g. the probability for a particle to decay into two lighter particles) is always associated with an error, or uncertainty. Contrary to common parlance, by 'error', scientists do not mean that they may have made a mistake. Conversely, they try to prove that they are aware of the limits of their findings and that they are able to estimate the distance between the true – unknown – value and the value they have calculated. While interactions between particles are governed by precise laws of probability, during each collision, Nature randomly 'picks' among all the possible combinations, in the rather same way as a lottery draw. This randomness is found in the statistical uncertainty, which results from the limited number of events on which the analysis is based. The systematic error considers the detector's precision and the characteristics of the methodology used to obtain the result.

Improving measurement over time

This shows how a real experimental measurement improved over time thanks to the accumulation of data and advances in analysis techniques. The abscissa axis (arbitrary scale) shows the amount of data used to obtain the results. Between the first and last points, which are 8 years apart, the number of analysed events was multiplied by over 20! The top graph displays the measured values (black) with the associated statistical errors (blue) and systematic errors (red). The variation of uncertainties over time is detailed below. The statistical error decreases as the number of events increases, from 20% to 3%. The systematic error remains almost constant. While initially negligible, in the last measurement its contribution is of the same order as the statistical error.

Translated from 'Du détecteur à la mesure',* **Passeport pour les deux infinis,** *IN2P3/Dunod, https://pass2i.ijclab.in2p3.fr/ (French); updated with ICHEP2024 conference (https://ichep2024.org) for IPPOG