Studies in History and Philosophy of Science 105 (2024) 99–108

Contents lists available at ScienceDirect

Studies in History and Philosophy of Science

journal homepage: www.elsevier.com/locate/shpsa

Computer simulation in data analysis: A case study from particle physics

Brigitte Falkenburg

TU Dortmund, Faculty 14, Emil-Figge-Str. 50, 44227, Dortmund, Germany

ABSTRACT

The paper presents a case study of the data analysis in the CDHS scattering experiment of particle physics performed in 1983. The case study compares the function of computer simulation in the data analysis with recent philosophical work on the role of simulations in high energy physics (HEP) and the theory-ladenness of the data. In the data analysis of CDHS, computer simulations entered an iterative process of probabilistic data correction. The computer simulation was a crucial ingredient of the data analysis that served to increase the accuracy of the measurement. The way in which simulation was used corresponds in a certain sense to the function of “models as mediators” (Morgan and Morrison), by mediating knowledge about measurement errors and the way of correcting them. I argue that this use of simulation did not give rise to a vicious circle of adjusting data to theory and vice versa but only to a weak, or benign, theory-ladenness of the data compatible with scientific realism. In the publication of the CDHS results, the measurement outcomes are called “observed data”, indicating a realist attitude of the physicists towards the measured quantities which does not exactly fit in with entity realism or theory realism.

1. Introduction

Computer simulation and methods of machine learning are omnipresent in all advanced disciplines, from physics, chemistry, biophysics, evolutionary biology, cognitive neuroscience, to climate science. Typical examples from physics concern models of structure formation in the universe, the generation of “landscapes” of theories beyond the standard model of particle physics, or the simulation of event signatures in the particle detectors of high energy scattering experiments, e.g., at the LHC. In all these cases, computer simulation is a powerful tool of theoretical physics that aims at modelling the dynamics of systems and processes and at exploring theories. In recent philosophy of science, the role of computer simulations in the observations and experiments of astrophysics, condensed matter physics, and particle physics has drawn particular attention.

In this paper, I present a case study from particle physics to shed new light on the function of computer simulations in the scattering experiments of high energy physics (HEP). My starting point are Margaret Morrison’s investigations of how models function as indispensable instruments of experimental inquiry and her claim that computer simulations may have the same epistemic status as experiments (Morrison, 2009, 2015). Morrison’s case studies provide substantial insights into the ways in which simulations mediate between theory and the phenomena, i.e., the measurement outcomes that result from the data analysis of an experiment. Simulations cannot prove the existence of physical phenomena, but they are essential for determining their properties (Morrison, 2009, p. 54):

> “Obviously one cannot determine the existence of physical phenomena using computer simulations, only material experiment is capable of that. But, the ability of those experiments to establish ontological and epistemic claims about those phenomena requires a complex network of modelling assumptions that are an inseparable part of the experimental practice. In that sense the appeal is not to materiality as the source of justification but to the model of that materiality that best accounts for the data.”

Massimi and Bhimji (2015) specify Morrison’s account as follows: Simulations entered the LHC experiments in search of the Higgs boson as a genuine experimental practice that contributes to the data analysis, by providing knowledge about the causal interaction between the apparatus and the target system. They were genuine *parts* of the LHC experiments rather than being experiments (or substitutes of such) on their own. This diagnosis is in accordance with a general analysis of computer simulation and measurement in scientific practice, which was supported by a case study on atmospheric data assimilation (Parker, 2017).

Computer simulations have been entering the data analysis of HEP indeed for about half a century, beginning with the scattering experiments of the 1970s and 1980s performed to test the quark model and the predictions of Quantum Chromodynamics (QCD). Since then and up to the present experiments of HEP, simulations and the models underlying them are not only indispensable for interpreting the data in causal terms, but also for the preceding procedure of data taking and the related steps of data analysis and data correction. The way in which simulations enter the procedure of data taking raises new questions about the theory-ladenness of data and the background theories involved in the data

---

E-mail address: brigitte.falkenburg@tu-dortmund.de.

https://doi.org/10.1016/j.shpsa.2024.03.005
Received 29 August 2022; Received in revised form 8 January 2024; Accepted 25 March 2024

analysis, which only recently began to draw the attention in the philosophy of physics.

Most case studies on the function of computer simulation in the experiments of particle physics focus on the way in which simulations serve to calculate dynamical models of the scattering events in a particle detector. Morrison (2009, 2015) explores the interpretative function of the models behind the simulations and compares them with the mediating role of models in experimental contexts. Massimi and Bhimji (2015) discuss whether computer simulations and material experiments support different kinds of causal stories about what goes on in an experiment and argue that both are interchangeable. In addition, there are case studies investigating how the scattering events in particle detectors are generated by Monte Carlo methods for the purposes of analysing the data in theoretical terms (Merz, 1999), and how the corresponding simulation results are validated (Mättig, 2019, cf. also Morrison, 2014). Recent case studies finally focus on the data models employed in the LHC experiments and the interplay between simulations and data processing and computer simulations (Antoniou, 2021; Karaca, 2018, 2020).

None of these case studies explains *in detail* how computer simulations contribute to the data analysis of high energy scattering experiments. To fill this gap, here I present a detailed case study on an experiment of the CDHS collaboration in which I have participated almost 40 years ago, as a PhD student of particle physics. The respective experiment was run in 1983 as a high-precision measurement of the cross section of deep inelastic neutrino-nucleon scattering (Berge et al., 1991). In the experiment, we recorded and analysed *big data* long before they became a public issue, performing many steps of data analysis "by hand" that are now done automatically by machine learning techniques. After giving an outline of what is a computer simulation and what it is good for in the practice of physics (Sect. 2.), I explain the objective and the setup of the CDHS experiment (Sect. 3), the event reconstruction (Sect. 4), and the role of computer simulation in the data analysis (Sect. 5). The final philosophical discussion will focus on three topics. I will compare the function of computer simulation in data analysis with Morrison's 2015 view of knowledge mediated by models and simulations (Sect. 6), introduce Karaca's 2013 distinctions to defend the claim that the CDHS data are theory-laden in a weak sense (Sect. 7), and draw some conclusions about the realist attitude of the physicists towards the measured quantities which does not fit in with well-established philosophical positions (Sect. 8).

2. What is computer simulation and what is it good for in physics?

Computer simulations are algorithms that run on computers and are programmed to calculate the dynamic evolution of a certain process in a complex system. The calculation can either be based on dynamic equations (which may be analytically solvable or not) or stochastic. In the experiments of HEP and many other areas of science, the simulation is usually stochastic, i.e., the algorithm is based on a random generator that generates aleatoric events. The corresponding methods are called Monte Carlo simulations.

In the philosophical discussion, computer simulations (or the underlying algorithms) have been compared with experiments, thought experiments, logical inferences, and models (cf. The detailed review Saam, 2017). In the vast debate about the nature of computer simulations it is also controversial whether they contribute any new philosophical problems at all to the objectives and methods of the exact sciences. Computer simulations owe their name to the fact that they imitate a process by another processes that runs on a computer (Hartmann, 1996). If we replace "imitate" by "represent" or "model", we come close to Morrison's views. To compare computer simulations either with inferences (Beisbart & Norton, 2012) or with models (Morrison, 2009, 2015) focuses on the way in which they represent the process that is simulated. To consider them inferences means to focus on the underlying algorithms and is close to the good old "syntactic" view of scientific theorizing, whereas to consider them models seems to be closer to the "semantic" view of theories as classes of models. In contrast to the semantic view, however, Morrison considers autonomous models that do not belong to a more comprehensive theory of a certain domain.

The discussion is due to the fact that computer simulations are methodologically hybrid. They combine the methods of computer science with the theories and models of the empirical sciences. The philosophical debate centers on the question of where computer simulations belong in the landscape of scientific methods. The question of what they are is best understood in this sense, I suppose. However, a closer look at the practice of science reveals that it is much more instructive to look at the *functions* of computer simulations in certain scientific contexts. According to Hartmann (1996), simulations are employed in five ways:

(1) as a *technique* of calculating the dynamics of a complex system in detail,
(2) as a *heuristic tool* of developing and exploring hypotheses, models, and theories,
(3) as a *substitute* of material experiments,
(4) as an *instrument* of supporting the performance of experiments, and finally,
(5) as a *didactic tool* of giving insight into opaque processes, e.g., the wave-particle duality of processes of quantum optics such as the double-slit experiment with single photons.

According to my knowledge of several theories and experiments of physics, this list is pretty much complete. I suspect that philosophers who *identify* computer simulations with experiments in the context of physics or claim that experiments and computer simulations are *exchangeable* with experiments, focus on function (3), i.e., on simulations that substitute experiments. But there is much more to say about computer simulation in the practice of physics. In typical experiments of particle physics, functions (1) and (4) are combined, as my case study and its comparison with Morrison's views about measurements, models, mediation, and the function of simulations in experiments (Morrison, 2015) will show.

3. The discovery of point-like structures in the nucleon

In the philosophy of physics, the scattering experiments of the late 1960s supporting the existence of quarks as scattering centers in the proton and neutron have not yet received the attention they deserve. They were no less spectacular than the discovery of the Higgs boson. Prior to them, the quark model proposed by Gell-Mann in 1964 was highly speculative. The unexpected discovery of point-like structures in electron-proton scattering at the Stanford Linear Accelorator Center (SLAC) gave it first evidence. The experiment was performed by Richard E. Taylor, Jerome I. Friedman, and Henry W. Kendall, who in 1990 received the Nobel prize for their discovery. Their experiments initiated the success of the Standard Model of particle physics, to which the detection of the Higgs boson only provided the final piece.

The first hints to the discovery were announced by the director of the SLAC during the high energy physics conference in Vienna (Panofsky, 1968). A detailed report of the results and their theoretical interpretation was later published by Friedman and Kendall (1972). The discovery of point-like structures in the proton resulted from a routine experiment that explored the region of deep inelastic scattering, i.e., subatomic interactions with very high energy transfer from the beam particles to the atomic nuclei within the target. The experimental results were in some ways an unexpected reprise of the results of scattering experiments carried out 60 years earlier in Rutherford's laboratory at much lower scattering energies. The measured cross section did not depend on the energy of the beam and showed in addition unexpected backward scattering, indicating the existence of point-like scattering centers within the atom (in Rutherford's case) or the protons and neutrons (here).

Depending on the energy of the beam and the corresponding spatial resolution of a scattering experiment (which is determined by the de Broglie wavelength of the quantum particles of the beam), the total cross section of a scattering experiment permits inferences to the internal structure of the target. If the measured cross section depends on the energy of the beam, it permits to extract form factors that describe the charge structure of the scattering center. In the quantum mechanics of scattering, the form factor is defined as the Fourier transform of the charge distribution of a scattering center. The definition of the electromagnetic form factors is based on the formula of the cross section of elastic electron-proton scattering given by Rosenbluth (1950). Hofstadter et al. (1958) measured the form factors of the proton and neutron. If the measured cross section does not depend on the beam energy, it is scale invariant, i.e., it does not crucially depend on the energy scale at which the experiment is performed. This phenomenon was interpreted as evidence of point-like scattering centers, in the SLAC experiment (for details of the argument, cf. Falkenburg, 2007, Chap. 4).

The results of 1968 were soon interpreted in terms of generalized form factors, the "structure functions" of the proton, derived in the "parton model" (Bjorken, 1969; Bjorken & Paschos, 1969; Feynman 1969). The "partons" were understood as point-like parts of the proton or neutron. The result of the SLAC experiment and its interpretation in terms of the parton model gave first empirical support to the quark model (cf. Riordan, 1987, 156–188; Riordan, 1992; for a philosophical analysis, cf. Karaca, 2013). Several follow-up scattering experiments measured the structure functions of the proton and their scaling behaviour in much more detail.

The scattering experiments exploring the structure of the proton and neutron from the 1960s to the 1980s were fixed target experiments. A beam generated by a particle accelerator was directed at a target equipped with particle detectors to measure the kinematic and dynamic quantities of the particles produced in a scattering process. The objective of these scattering experiments was to measure the dynamical properties of as many particle reactions of a particular type as possible with the greatest possible precision. Such experiments took big amounts of data, employing Monte Carlo simulations in the data analysis to minimize the measurement errors since the 1970s.

With increasing precision, the experiments could serve to test the predictions of the quantum field theory of strong interaction, Quantum Chromodynamics (QCD), developed in 1973. The quark model of 1964 and the parton model of 1969 were combined to the quark-parton model, which extended the "naïve" parton model of 1969, an idealized model of point-like partons, by including QCD predictions on the dynamics of the quarks and gluons (the field quanta of QCD) (for a philosophical analysis, cf. Massimi, 2004). According to the quark-parton model, quarks and gluons are dynamic parts of the protons and neutrons. The model interprets the structure functions in terms of the momentum distributions of quarks and gluons, which depend on the energy transfer between the scattered "probe" particles of the beam to the protons or neutrons in the target of the experiment.

QCD predicts that at higher scattering energies the increasing energy transfer to nucleons generates increasingly many quark-antiquark pairs and gluons inside the protons or neutrons, giving rise to violations of the scale invariance of the scattering predicted by the naïve parton model. Scaling violations were first detected in muon-nucleon scattering at the Fermi National Accelerator Laboratory (FNAL) (cf. Feltesse, 2012). The experimental results indicated the existence of *non*-point-like structures within the proton or neutron, and they were taken as further confirmation of the quark model. The decisive experiment employed a Monte Carlo simulation to test the scaling behaviour of the scattering more precisely (Chang et al., 1975).

4. The CDHS experiment

The experiment on which my case study focuses made further decisive steps in increasing the measurement precision. The CDHS group performed a high energy scattering experiment at the European Center for Nuclear Research (CERN) with a neutrino beam produced by the Super Proton Synchroton (SPS). It was designed and performed from 1976 to 1983 by physicists from CERN (C), Dortmund (D), Heidelberg (H), Paris-Saclay (S), and in its second phase, also from Warsaw. The group leader was Jack Steinberger, who finally together with Melvin Schwartz and Leon M. Lederman in 1988 received the Nobel Prize for their discovery of the muon neutrino in 1961. The CDHS experiment was designed to validate the standard model, which at that time was still undergoing tests. Compared to the earlier experiments that measured deep inelastic electron-proton scattering at or muon-nucleon scattering at the SLAC and the FNAL, the experimental setup had several advantages. On the one hand, it was designed to perform a high precision measurement of the nucleon structure functions and the scaling violations predicted by QCD. On the other hand, to employ neutrino scattering made it possible to measure also neutral currents and the electroweak parameters of the standard model of particle physics, such as the Weinberg angle of the Salam-Weinberg theory.

The neutrinos resulted from the decay of kaons and pions produced by making the proton beam of the SPS hit at a beryllium target. Depending on the charge selection of the kaons and pions by means of a magnetic horn, the resulting secondary beam consisted either of neutrinos or of anti-neutrinos. About 10^{13} protons from the SPS were extracted in bursts of 2 ms duration every 14.4 s, each resulting in about 70 neutrino events or 30 antineutrino events in the detector (Berge et al., 1991, p. 189).

The CDHS experiment was a fixed target experiment. The detector, which also served as the target of the scattering, was made up of 19 (and later 21) magnetized iron modules at which the neutrinos hit and produced secondary particles in their interactions with the iron atoms. Between the modules, plastic scintillators and wire drift chambers were inserted to measure the energy of hadron showers and to record the tracks of charged particles. In this way, the detector served simultaneously as scattering target, hadron calorimeter, and muon spectrometer. When it was completed in 1976, it was 20 m long and had a weight of 1250 tons. Later, a hydrogen tank was added in front of the detector to measure the scattering at protons only and to perform independent measurements of the structure functions of up and down quarks (Abramowicz et al., 1984). From 1980 on, the detector was upgraded to improve the accuracy of the hadron energy measurement and the spatial resolution of the particle tracks. Several calorimeter modules were exchanged, four more precise drift chambers were added, and the number of modules was increased.

In the following, I focus on the data analysis of the measurement of structure functions with the upgraded detector (Berge et al., 1991), to which I contributed as a PhD student (Falkenburg, 1986). It was based on the measurement of charged current particle interactions:

$$\nu + \mathrm{Fe} \rightarrow \mu + X$$

The neutrino (or antineutrino) – iron interaction results in an outgoing muon μ^- (or μ^+) and the production of a hadron shower X (see Fig. 1). The Feynman diagram of the scattering process in the lowest order of perturbation theory (Born approximation) looks as follows (Fig. 2).

In contrast to the charged current scattering processes measured here, the *neutral current reactions* $\nu + \mathrm{Fe} \rightarrow \nu + X$ measured with the CDHS detector in other experiments did not give rise to a muon track in the drift chambers, only to the deposition of hadronic energy in the calorimeter. On the other hand, *atmospheric* muons *not* due to interactions induced by the neutrino beam from the SPS had the opposite signature: No hadron shower in the calorimeter, only a muon track in the drift chambers. This (omnipresent) natural background of cosmic rays was also measured by taking data *without* beam, between the neutrino beam bursts from the SPS, for purposes of calibration, control of the detector response, and data correction.

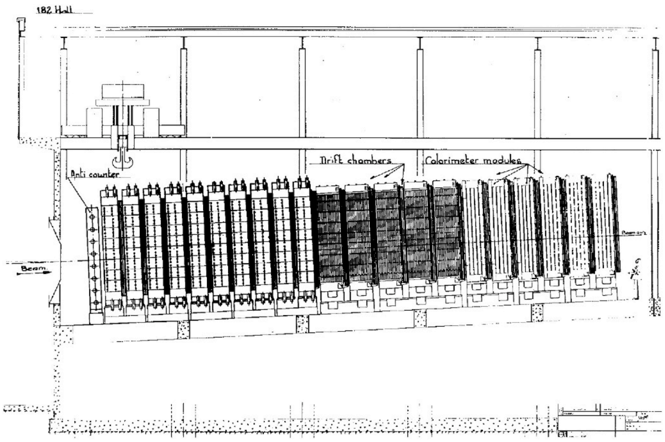

Fig. 1. The CDHS detector (Berge et al., 1991, p. 189).

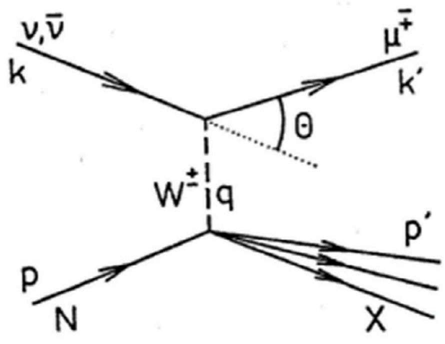

Fig. 2. Deep inelastic scattering in lowest order perturbation theory (Born approximation).

The data taking took place from September to December 1983, with a wide-band neutrino beam (i.e., with a large energy spectrum) resulting from the scattering of a 400 GeV proton beam from the SPS at the beryllium target. All secondary particles except the neutrinos (or antineutrinos) were eliminated by means of a collimator, a magnetic horn, an evacuated decay tunnel of ~300 m, and an iron and earth shield of ~450 m.

The measured quantities of the charged current interactions were the muon momentum p_μ after the scattering, the scattering angle θ of the muon relative to the direction of the neutrino beam, and the hadronic energy transferred from the neutrino to the proton or neutron, E_{had}, in the laboratory frame. From these three magnitudes, the four-momentum transfer q of the neutrino to the nucleon, which is crucial for the measurement of the differential cross sections and the nucleon structure functions, could be calculated for a given Lorentz invariant neutrino four-momentum k_ν, or a corresponding laboratory energy E_ν of the neutrinos. The energy E_ν of the incoming neutrinos in the laboratory system was not precisely known. Their energy spectrum was estimated based on the observed event rates and previous measurements of the total cross sections, which was assumed to increase linearly with the neutrino energy. To determine the neutrino energy spectrum more precisely, an iterative procedure employing Monte Carlo simulations was used (see Sect. 5 below).

The scattering events due to charged-current neutrino or antineutrino interactions had the signature that a curved muon track and the energy attributable to a hadron shower had to originate in the same spatiotemporal region of the detector, i.e., have the same vertex, in correlation with a pulse from the primary proton beam of the SPS. A computer program called RECONU reconstructed the scattering events satisfying these criteria from the raw data taken by the calorimeters and the drift chambers. The program performed pattern recognition of the raw data of all events that met the trigger conditions. It calculated the hadronic energy E_{had} deposited in the calorimeter. The calculation took the event-rate dependent detector response and the measurement errors into account which were obtained from independent measurements. A further correction concerned the muon energy deposited in the calorimeter. In addition, the reconstruction program put a muon track through the measurement points recorded by the drift chambers to calculate the initial momentum of the muon p_μ at the scattering vertex and the scattering angle θ.

The calculation of the muon momentum was carried out iteratively from the curvature of the track, fitting a semi-classical track through the position measurement points recorded by the drift chambers. The program RECONU traced the muon track from its end back to its beginning, accounting for all damping effects including the scattering processes of the muon along the detector. To do so, the classical formula of the Lorentz force for the motion of a charged particle in a magnetic field was combined with semi-classical formula obtained from the quantum electrodynamical expectation values for multiple scattering and energy loss due to processes of ionization, pair creation, or bremsstrahlung along the muon tracks. In addition, the reconstruction accounted for measurement errors due to the finite spatial resolution of the drift chamber measurements and due to multiple scattering of the muons at the iron nuclei. These measurement errors were obtained from theoretical calculations which were adapted to the specific conditions of the CDHS detector. To verify and correct the theoretical measurement errors, the accuracy of the pattern recognition and fitting procedure of the program was checked by reconstructing a large sample of events obtained by the Monte Carlo simulation (see Sect. 5 below). The resulting momentum and angle resolutions were included into the iterative procedure of fitting a muon track through the measurement points, tracing it back to the vertex and calculating the initial muon momentum. The latter was then used to correct the hadronic energy calculated from the calorimeter data, as already mentioned above, because in addition to the hadronic shower the muon also deposited energy in the calorimeter.

A typical event reconstructed in four projections along the detector axis is shown above (Fig. 3). The upper line shows the pulse height recorded in the calorimeter modules along the detector. The cylinder at left symbolizes the detector, the z axis pointing along the detector and the other axes perpendicular to it.

Finally, the reconstructed event distributions were corrected at the probabilistic level for the measurement errors of the hadron energy E_{had}, the initial muon momentum p_μ, and the scattering angle θ. Here, too, an iterative adjustment by Monte Carlo simulations was required, and in addition an unfolding procedure that aimed at recovering an approximately true event distribution from the measured and reconstructed event distribution, which was distorted by the measurement errors (see Sect. 5 below).

Hence, the reconstruction program and the whole procedure of data analysis were highly sophisticated. The basic criteria for "good" charged current events, the corresponding trigger conditions, and the formulae for the propagation of muons in iron entered the reconstruction program. These formulae, in turn, employed the classical Lorentz force law as well as semi-classical corrections of it obtained from quantum electrodynamics (QED), in the usual measurement theory employed in the experiments of particle physics (cf. Falkenburg, 2007, 174–187). Hence, QED served as a well-established background theory of the data analysis.

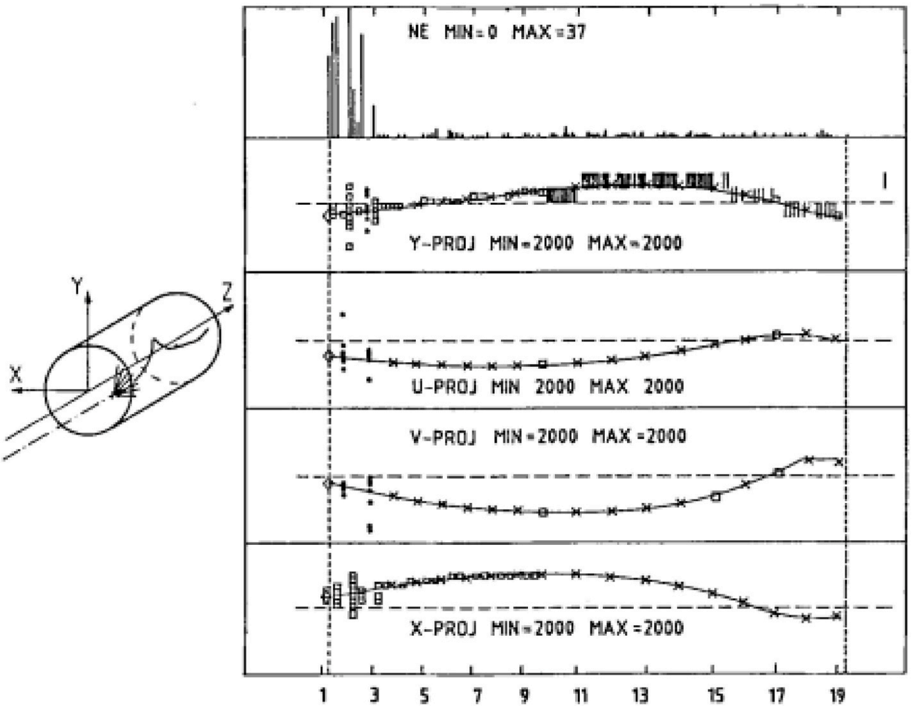

Fig. 3. Typical reconstructed charged scattering event (Berge et al., 1991, p. 191).

The semi-classical formulae derived from it were adapted to the specific behaviour of charged particles in the passage through matter (here: iron) known from independent measurements (Parker et al., 1984; Rossi, 1952). That is, the resulting formulae were *semi-classical* as well as *semi-empirical*, i.e., they were based on a classical model of particle tracks modified by quantum corrections as well as by data. Because particle tracks result from quantum processes that include quantum fluctuations, their semi-classical reconstruction gave rise to systematic measurement errors of the event reconstruction, which added to the usual detector-specific measurement errors.

To estimate the precision of the event reconstruction, a representative sample of several thousand computer-reconstructed events was visually inspected and interactively corrected by members of the CDHS collaboration, comparing the results of the computer reconstruction of the events with the raw data, as recorded by the detector and visualized at a computer screen. The comparison demonstrated an efficiency of the automatic reconstruction of more than 98,5 % (Berge et al., 1991, 191–192). After calibration and further event selection, a sample of 640 000 neutrino events and 550 000 antineutrino events remained for the further data analysis (ibid., 194).

5. Computer simulation in the data analysis

As sketched above, the pattern recognition and event reconstruction by the computer program RECONU already incorporated the results of Monte Carlo simulations in the iterative process of reconstructing the scattering events. Hence, simulations and event reconstruction were intertwined at several stages of the data analysis. Roughly speaking, the Monte Carlo simulation of the neutrino scattering events served to estimate the measurement errors, to correct the reconstructed data for them, and to cross-check the accuracy of the results of the event reconstruction. Let us now have a closer look at the interplay between simulation and data correction, which included iterative procedures of adjusting the simulation results to the data, and vice versa, until the simulation and the reconstructed data converged.

The Monte Carlo simulation and its use for the data correction proceeded in several steps. The objective was to increase the measurement accuracy of the experiment. This gain of precision was iteratively achieved, based on less precise data from previous experiments, theoretical knowledge of what was going on in the experiment, sufficiently large event samples for which the statistical measurement errors lost significance, and detailed studies of the systematic measurement errors. The Monte Carlo program generated charged current scattering events in the detector with a spectrum of the neutrino energy E_ν that was based on an independent previous measurement of the total cross section and a distribution of the crucial kinematic magnitudes p_μ, θ, and E_{had} of the scattering that resulted from previous measurement of the structure functions. The iterative procedures aimed at consistency of the simulation results and the corrected data, in the sense of a coherence criterion for an adequate measurement.

1. *Determination of the energy spectrum of the neutrino beam*: The energy spectrum and radial distribution of the neutrinos or antineutrinos in the wide-band beam were not known, because they resulted from several different secondary decay processes in the decay tunnel behind the Beryllium target at which the SPS protons hit. To determine the energy spectrum and the radial shape of the respective neutrino beam more precisely, previous measurements of the total cross section and the nucleon structure functions were employed. The total cross section, i.e., the total number of charged current interactions $\nu + \mathrm{Fe} \rightarrow \mu + X$ as a function of the neutrino energy E_ν, was obtained from independent measurements with a narrow-band beam of well-defined neutrino energy. It turned out to increase linearly with the neutrino energy E_ν. The Monte Carlo program took this

result and generated scattering events of a linearly increasing number for increasing neutrino energy. The distribution of the kinematic variables of the scattering at a given energy was taken from previous measurements of the structure functions. The program calculated the energy deposited in the hadron calorimeters, compared the resulting event distribution with the input of the simulation, calculated the calorimeter acceptance from the difference, corrected the input, and repeated the procedure until input and output agreed. The resulting distribution was not sensitive to the choice of the measured structure functions taken as input (ibid., 195).

2. *Simulation of the scattering events:* The Monte Carlo program simulated the distribution of the kinematic variables of deep inelastic scattering events, the deposition of hadronic energy in the calorimeters after the scattering, and the propagation of the resulting muons through the magnetized iron and the drift chambers, i.e., the amount of energy deposited by them in the calorimeter and the positions measured by the drift chambers indicating the muon tracks. The simulation of the muon tracks included the quantum electrodynamical effects of the passage of charged particles through matter, namely multiple scattering, and energy loss due to processes of ionization, pair creation, or bremsstrahlung along the track, which were also implemented in the event reconstruction. However, in contrast to the semi-classical reconstruction of the measured muon tracks that included average effects given by quantum theoretical expectation values (cf. Sect. 4), the Monte Carlo simulation generated the relative frequency of these processes for given muon momentum corresponding to the QED quantum probabilities, including quantum fluctuations of energy loss along the tracks. In this way, the Monte Carlo also generated muon tracks for which the semi-classical model underlying the reconstruction program fails in the sense that it is not even approximately true, e.g., for a muon track with a kink that is due to an abrupt energy loss by pair creation or bremsstrahlung. These failures of the semi-classical track reconstruction resulted in deviations of the reconstructed event distributions from the simulated "true" quantum scattering event distributions. Discrepancies between the event distributions obtained from the semi-classical track reconstruction and from the quantum electrodynamical probabilities implemented in the Monte Carlo simulation were especially expected to occur for muons of high energy and large momentum, for which the probability of pair creation and bremsstrahlung increases.
3. *Analysis of the measurement errors:* For a detailed study of the systematic measurement errors due to track reconstruction failures, large samples of events were generated by the Monte Carlo for given values of the kinematic variables p_μ, θ, and E_{had} of the scattering, and then reconstructed by the reconstruction program RECONU (Falkenburg, 1986). The reconstruction of simulated events directly resulted in the distribution of the reconstructed values of muon momentum p_μ and the scattering angle θ around the generated values, i.e., in the distribution of the errors due to failures of the event reconstruction. The resulting error distributions were compared with the theoretical calculation of the measurement errors. It turned out that the above-mentioned discrepancy between the quantum electrodynamical energy loss and the semi-classical track reconstruction did neither significantly affect the theoretical average errors used for the track reconstruction, nor the momentum distributions resulting from the reconstruction. But the comparison between the simulated events and the reconstruction results showed that the error distributions had large non-Gaussian tails that biased the reconstructed event distributions in certain kinematic regions. These distortions had to be taken account in the final data correction for measurement errors, which was performed at the probabilistic level via an unfolding procedure.
4. *Unfolding the event distributions from measurement errors:* To unfold a sample of empirical events (or their reconstruction) from measurement errors is intended to correct for the fact that the relative event frequencies of given measured variables are distorted or smeared by measurement errors. The objective of the unfolding procedure is to recover approximately true event distributions from the convolution of the true event distributions with measurement errors, using statistical methods. Here, the term "unfolding" is to be understood in the technical sense of a well-defined and well-established statistical method, which has been employed in the data analysis of particle physics at least since the early 1980s (for details, cf. Blobel, 1984, 2013).[1] Due to the finite resolution of the measured quantities, the reconstruction program assigns kinematic variables to the scattering events that may deviate from the true values due to measurement errors, with the result that the program puts the reconstructed events into wrong kinematic regions. Mathematically, this problem is described by an integral equation (a Fredholm integral equation of the first kind) that describes the convolution of a distribution (here: the true event distribution) with another distribution (here: the distribution of the measurement errors). The Monte Carlo simulation can easily generate scattering events that correspond to the reconstructed data, by smearing (or convoluting) the simulated events with the measurement error distributions according to this integral equation. The inverse problem, that is, to recover the true event distribution from the smeared events, is more delicate (ibid.). To the experiments of particle physics a discretized model of this continuous mathematical problem applies, which is handled by sorting the events into discrete bins (i.e., intervals) of the kinematic variables and weighting the event numbers of all bins by acceptance factors. The acceptance factors A(z) are defined as the ratio of true to error biased event numbers per bin Δz around the value z of a given variable, and they are calculated as the ratio of unsmeared to smeared Monte Carlo events. As for the CDHS experiment:

$$A(x, y, E_\nu) = MC^0\ (x, y, E_\nu) \,/\, MC^1\ (x, y, E_\nu)$$

Here, MC^0 (x, y, E_ν) and MC^1 (x, y, E_ν) are the numbers of Monte Carlo events without and with smearing, in a bin belonging to certain intervals of the kinematic variables x, y, and E_ν. The bins are expressed in terms of the Lorentz invariant variables x and y, in which the measured differential cross sections and the nucleon structure functions of the quark-parton model are defined (with muon energy $E_\mu \approx p_\mu$):

$$y = E_{had}\ /E_\nu$$

$$x = 4E_\nu\ E_\mu\ \sin^2(\theta\,/\,2)$$

The bin sizes had to be chosen carefully to avoid further distortions. The smeared event distributions were generated with the error distributions obtained from the independent measurements of the resolution, for the hadron energy E_{had}, and the theoretical calculation plus the Monte Carlo analysis explained above, for the muon momentum p_μ and the scattering angle θ. As input for the event generation, again previously measured structure functions (or differential cross sections, respectively) were needed to obtain the relative event frequencies. Hence, the unfolding procedure had to be iterated, taking the unfolded kinematic event distributions calculated with the acceptance factors of one iteration as input for the next iteration. It was checked that the

---

[1] Unfolding in this technical sense is also employed in the LHC experiments (Beauchemin, 2017). It should not be confused with the loose meaning of unfolding or unravelling complex problems, on which the article "Multiplex and Unfolding: Computer Simulation in Particle Physics" (Merz, 1999) focuses, from the point of view of historical epistemology. The article does not mention at all the statistical unfolding method of data analysis in particle physics. It analyses the uses of Monte Carlo event generators, "the program packages that particle physicists construct and use to simulate mechanisms of particle production" (ibid., 293), with the goal "to bring the multiplex and unfolding character of such knowledge objects to the fore: Multiple meanings and functions are embodied in the object and can be drawn out selectively according to the requirements of a work setting" (ibid.).

resulting event distributions depended only weakly on the input structure functions (Berge et al., 1991, p. 195)[2]. When the iteration procedure converged to acceptance factors that no longer changed, the smeared Monte Carlo event distributions reproduced well the data obtained from the reconstruction program (see Fig. 4).

In the publication of the experimental results, these final corrected event distributions are called the "observed data" (ibid.). It is quite remarkable that the physicists call the results of a highly theory-laden procedure of data acquisition and data analysis the "observed data" of the experiment. This sloppy parlance blurs the distinction between the observation of entities and the measurement of physical quantities. This parlance indicates belief in the truth of the measurement results that sheds some light on questions of scientific realism (cf. Sect. 8).

Further data analysis no longer required computer simulations, but several difficult theoretical calculations. The measured structure functions provided substantial information about the quarks and gluons inside the proton and neutron, especially about the so-called "sea quarks", i.e., virtual quark-antiquark pairs generated within the nucleons with increasing scattering. The paper (Berge et al., 1991) was not published until 8 years after the data taking (1983) and 5 years after completing the measurement of the differential cross sections (1986). The main reason for this delay were calculations of electroweak radiative corrections in higher orders of perturbation theory, which were needed to compare the results of the experiment with the predictions of quantum chromodynamics.

6. The functions of computer simulation in experimental practice

Let me now compare the functions of the Monte Carlo simulations of my case study with the recent philosophical work on computer simulation mentioned in the introduction and with Hartmann's classification of the ways in which computer simulations enter scientific practice (cf. Sect. 2). The simulations in the CDHS experiment neither substituted an experiment nor were they like an experiment (function 3). They modelled the dynamic processes occurring in the CDHS detector due to the scattering of neutrinos at the protons and neutrons of the iron nuclei (function 1). Their objective was to support the data analysis of a high precision experiment, i.e., they served to perform an experiment (function 4). Especially, they were needed to determine the energy spectrum of the neutrino beam, to make detailed studies of the measurement errors, and to correct the measured event distributions.

My case study is in good agreement with the recent philosophical studies of data models in HEP mentioned in the introduction. It gives a "fine-grained" description of how Monte Carlo simulations enter the data analysis already at the lowest level of modelling the raw data, in accordance with the more "coarse-grained" studies of the data processing at the LHC given by Karaca (2018, 2020) and Antoniou (2021). The reconstruction procedure of CDHS is indeed very similar to the one described Antoniou (2021, Sect. 5.2) for LHCb, indicating an enormous methodological stability of data analysis in HEP over decades. Another very detailed case study has been presented by Beauchemin (2017). To clarify the theory-ladenness of observation in HEP, he performs an "autopsy" of the measurements at the ATLAS experiment, from the "raw data" to the event reconstruction to the measurement of differential cross sections, including the correction of the data samples by unfolding the event distributions from measurement errors – which is also still the same method as employed in the CDHS experiment. It should be noted that the way in which Beauchemin (2017) and Antoniou (2021) understand the "raw data" of the LHC experiments also agrees with the use of the term at CDHS: The raw data are what is recorded by the electronic devices of the particle detectors and not what is observed in the empiricist sense of immediately perceiving the image displayed on a computer screen.

The detailed structure of the data analysis of the CDHS experiment also fits in with a more general analysis of how computer simulations may be embedded in measurements (Parker, 2017). The CDHS measurement of the differential cross sections, and the extraction of the structure functions out of the measured cross sections, was undoubtedly a (*very*) complex measurement, in terms of Parker's distinction of direct, derived, and complex measurements. Computer simulations entered the measurement at different levels of data analysis and layers of inferences about the data. However, I would not say that the simulations *constituted* the measurement outcomes (cf. Parker, 2017, *passim*). It is only justified to claim that they were constitutive for the *accuracy* of the measurement results. In this regard, the CDHS experiment is like the case of the muon-nucleon scattering at FNAL in the early days of QCD, where the Monte Carlo simulation provided a more precise test for scale violations than the predecessor experiment (Chang et al., 1975, p. 902; cf. end of Sect. 3 above).

In contrast, Morrison (2009, 2015) only claims that computer simulations *may* have the same epistemic status as experiments. This is more differentiated and can cope also with the CDHS and FNAL cases. I completely agree with her general conclusion about the use of simulation in HEP, namely that it

> "casts doubt on the very distinction between experiment and simulation; the latter is simply an integral part of the former. That said, there is still a distinction to be made between signal data … and simulated data. But, as I showed with the LHC, the two are 'processed' together" (Morrison, 2015, p. 361).

But the remaining distinction between simulation and experiment has yet another aspect. In contrast to what is often said about simulations, the Monte Carlo event generation and the event reconstruction of CDHS were *not* based on the same theoretical assumptions about what was going on in the detector. Although the Monte Carlo program and the reconstruction program RECONU were based on the same assumptions of QED, they processed them in completely different ways. The Monte Carlo generated scattering events *ab initio* at a probabilistic level, with aleatory relative frequencies that corresponded to the quantum probabilities. In contrast, the computer reconstruction of the muon momentum from the track shape rested on semi-classical formulae that describe the passage of charged particles through matter by quantum theoretical expectation values plus semi-empirical formulae. The semi-classical track reconstruction was a source of the momentum measurement. The Monte Carlo, in turn, served to correct the semi-classically "misconstrued" data by generating the event frequencies directly in accordance with the well-confirmed quantum probabilities.

The way in which the Monte Carlo simulation was here used to correct the data corresponds to the function of "models as mediators" (Morgan and Morrison 1999), in the sense that the theory underlying the Monte Carlo simulation resulted in a better model of the data. More specifically, the Monte Carlo model was a probabilistic model of the passage of charged particles through the apparatus that mediated knowledge about the measurement errors introduced by the semi-classical reconstruction of the measured particle tracks, and about how to correct for them. In terms of Morrison's 2015 overview, the Monte Carlo model provided a representational model of what was going on in the apparatus, which mediated knowledge in the second sense of mediation, where the starting point is not an explanatory or predictive goal but the theory itself (Morrison, 2015, p. 119), here: the quantum theory of the passage of charged particles through matter. This representational model was a "source of mediated knowledge of physical phenomena … characteristic of the type of knowledge we acquire in

---

[2] The unfolding procedure also accounted for other technical details, which especially concerned kinematic regions with relatively few events, but are not relevant here. For the purposes of the present paper, I also neglect the different cuts applied to the data and the Monte Carlo events, as well as all further corrections applied to the differential cross sections and to the extraction of the structure functions from them.

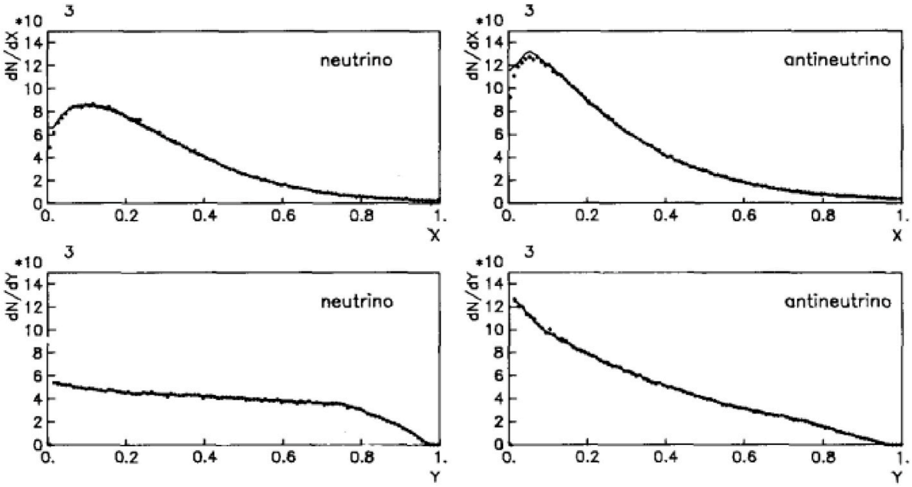

Fig. 4. Comparison of x and y distributions of data (points) and simulated events (solid lines), for neutrinos and antineutrinos (Berge et al., 1991, p. 195).

measurement contexts" (ibid., 205). It helped to improve the accuracy of the measurement. This model belonged to measurement itself in a more sophisticated way than discussed in Morrison's case studies, which serve to distinguish between cases of simulations that can be considered as measuring instruments and cases of simulations that are constitutive for an experimental activity (Morrison, 2015, Chapter 6). In the CDHS case, the Monte Carlo was constitutive for the measurement accuracy. Given this constitutive role, the question arises as to the "materiality" of the results of the data analysis, to put it in Morrison's terms. To put it in other terms: Are the final data theory-laden in a fatal sense or is it justified to assume that they represent what was really going on in the experiment?

7. Theory-ladenness and observation

At a first glance, the way in which computer simulations enter the data analysis of the CDHS experiment may appear a vicious circle, in which the raw data of the experiment were corrected based on computer simulations that vice versa were based on experimental results of the same kind. A closer look at the reconstruction of the scattering events and the computer simulations employed in data correction shows, however, that the corrections of the data result from well-defined iteration procedures rather than a vicious circle, indicating a case of benign circularity in the sense of Ritson and Staley (2021). The input of the respective iteration procedures was based on independent measurements, well-confirmed calculations adapted to the conditions of the CDHS experiment, and careful studies of the measurement errors.

In addition, the accuracy of the reconstruction results was cross-checked by means of a visual inspection of thousands of events. That is, the raw data recorded by the photomultipliers and the drift chambers were displayed on a computer screen together with the reconstructed particle tracks, as shown in Fig. 3, and the reconstruction results (solid lines) were compared with the raw data (discrete measurement points and bars), counting the numbers of events for which both crudely differed and estimating the statistical errors of the reconstruction on this basis. (This was the only stage of the experiment in which observation in an empiricist sense came into play. What was observed in this sense were deviations between the computer representations of the raw data, on the one hand, and the reconstruction results, on the other.) This cross-check of the reconstruction results (plus cross-checks and redundancies at other experimental levels, concerning, e.g., the calibration of the neutrino beam and the particle detectors) accounted for the robustness of the measurement outcomes – in the sense of a procedure robustness, as recently defined by Karaca (2022), that substantially contributes to the robustness of the experimental results.

To clarify the "circularity" vs. "materiality" question of the CDHS results, let me introduce further distinctions concerning the theory-ladenness of measurements, which are beyond the scope of Morrison's works. Karaca (2013) suggests to understand the theory-ladenness of experimental results as an *umbrella concept* that includes three different kinds of theorizing: (i) *background theories* in the sense of well-defined background information (cf. Shapere, 1982), (ii) interaction-specific *model theories* based on the background theories, such as the theory of the scattering matrix, that enter the data analysis of HEP since the beginning of high energy scattering experiments in the early 1960s, and (iii) the *phenomenological models* that guide theory-driven experiments, i.e. experiments that are not exploratory but designed to test a specific theory. Karaca (2013) distinguishes theory-ladenness in a *weak* sense, which is restricted to background theories and model theories, and theory-ladenness in a *strong* sense, i.e., the dependence of theory-driven experiments on phenomenological models.

The SLAC results of 1968 which indicated the "scaling" behaviour of deep inelastic scattering were obtained by exploratory experimentation with weak theory-ladenness, as emphasized by Karaca. In contrast, the follow-up experiments of the 1970s and 1980s, including CDHS, were theory-driven in Karaca's sense, that is, based on the phenomenological quark-parton model. CDHS was designed to obtain a more precise measurement of the nucleon structure functions and the scaling violations predicted by QCD. If assumptions of the quark-parton model or of QCD had been employed in the measurement of the differential cross sections, its results would have been theory-laden in a strong sense. But this was not the case. Only well-established background theories and model theories in Karaca's sense entered the analysis of the raw data: the quantum theory of elastic or inelastic scattering, QED, semi-classical models of particle tracks, the models of the neutrino beam and the particle detectors employed in the experiment, Lorentz invariant kinematics (i.e., special relativity), and mathematical statistics. Therefore, the measured differential cross sections were only theory-dependent in a weak sense. In the data analysis of CDHS, the step from weak to strong theory-ladenness seems to be marked by the extraction of the structure functions from the measured differential cross sections. The measured cross sections only depended in the weak sense on background and model theories, as did the computer simulations entering them. In contrast, the structure functions extracted from them depend on higher orders of electroweak radiative corrections. To calculate them needed several years, delaying the publication of the measurement outcomes of the 1983 experiment until 1991. However, this calculation was based on the Salam-Weinberg theory (SWT) of electroweak interactions, and *not* on the QCD predictions the experiment was to test. The SWT belongs to the background theories of the experiment, too, and not to the theory to be tested, i.e., QCD. Hence, also the radiative corrections needed to extract the structure functions were based on a model theory in Karaca's sense, indicating only weak theory-ladenness. Nevertheless, the SWT obviously enters the phenomenological quark-parton model that made the experiment theory-driven.

My first conclusion is that the differential cross sections measured by CDHS are a clear case of weak theory dependence in Karaca's sense, which supports belief in the truth of the measured data. My second conclusion, however, is that his distinction between weak and strong theory dependence is neither obvious nor sharp. There are clear cases of a weak or benign theory dependence, such as the differential cross sections of CDHS; but it is much harder to figure out which cases of a stronger theory dependence are no longer benign, from a philosophical point of view. Eliminating a vicious stronger theory dependence from experimental results means a lot of hard work for physicists, and if it cannot be eliminated, belief in the truth of the results is no longer justified.

Finally, what does the CDHS experiment tell us about the physicists' account of observation and their attitude(s) towards questions of truth and scientific realism? As mentioned in Sect. 4, the publication of the CDHS experiment calls the final corrected event distributions, i.e., the measured differential cross sections, the "observed data" (Berge et al., 1991, p. 195). In the light of the above considerations, we may understand this sloppy parlance in such a way that the measurement outcomes at this level of the data analysis, i.e., before the extraction of the structure functions, indicate a clear case of a non-vicious, weak theory dependence.

At first glance, this seems to agree with a recent definition, according to which an observation results from the transformation of instrumental signals "to the point at which comparison to theory can be made" by Beauchemin (2017, p. 279). However, the experimental results that could finally be compared to the QCD predictions were not the measured differential cross sections, but the structure functions only extracted from them after years of hard calculation. For good reasons, the CDHS publication only refers to the former as the "observed data", but not the latter, which were only published much later.

Therefore, let me suggest another interpretation of the physicists' parlance. Parker (2017, p. 279) considers measurement an "information-gathering activity, involving physical interaction with the entity measured, which locates the entity in a logical space" of physical quantities. Her approach follows van Fraassen (2008, Chapters 6-7) and Tal (2012). Concerning the question of how the measurement outcomes relate to reality, she emphasizes that "neither [of both philosophers] defends the view that measurement accuracy is a matter of how close an outcome is to a true quantity value" (Parker 2017, 278); she attributes an agnostic attitude to Tal and a model-dependent representationalism to van Fraassen (ibid.). In contrast, by speaking of the "observed data", the physicists implicitly express a firm belief that the measured quantities come close to truth, although (being no philosophers) they do not explicitly defend this belief.

8. Conclusions concerning scientific realism

To call the measurement outcomes "observed data" blurs the distinction between observed entities and measured quantities. This sheds some light on the physicists' attitude towards scientific realism. The observation of an entity supports an existence claim. The measurement of a quantity does not. The "observed data" of CDHS were the differential cross sections of deep inelastic neutrino-nucleon scattering obtained from the values of the hadron energy and the momentum and scattering angle of the outgoing muon measured by the CDHS detector. The experiment was theory-driven in the sense that it presupposed the quark-parton model of what is going on in deep inelastic neutrino-nucleon scattering. However, the measurement of the differential cross sections did not employ any specific assumptions of the quark-parton model (in contrast to the extraction of the structure functions from these "observed data"). The differential cross sections were given in terms of the Lorentz invariant magnitudes calculated from the measured quantities. To consider them approximately true presupposes some belief in the adequacy of the relevant background theories, i.e. the quantum theory of scattering and relativistic kinematics, and the model of deep inelastic neutrino-nucleon scattering based on them. At this point we arrive at the representational character of models emphasized by Morrison (2015, 119–130) and at the questions of scientific realism. Without entering the realism debate, we may leave it open whether the well-established background theories employed in measurements come close to truth, or whether they are just empirically adequate.

Most physicists have no well-elaborated philosophical views and their attitude towards questions of realism may even be incoherent. In the case of astroparticle physics, e.g., they combine instrumentalist views of their models of astrophysical sources, in which they do not yet have sufficient confidence, with certain elements of entity realism, causal realism, and belief in the truth of the laws of nature (cf. Falkenburg, 2012). Such an unclear combination of different elements of scientific realism seems to be also at work in the practice of high energy physics. To perform a high energy scattering experiment at least presupposes a kind of realism about the physical magnitudes of the incoming and outgoing particles of scattering processes, plus belief in the empirical adequacy (if not truth) of relativistic kinematics and the quantum theory of scattering. This means to assume that the measured beam properties and the corrected data distribution of the scattering events come close to the true particle properties and event distributions. To my view, this kind of scientific realism is typical of particle physics from the early cosmic ray studies (which, e.g., resulted in the discovery of the positron as a charged particle with the electron mass, but opposite charge) up to the present day. For the CDHS experiment, it implied the assumption that the neutrinos (or antineutrinos) of the beam have a certain energy which in the scattering is transferred to a proton or neutron, and that the outgoing particles after the scattering detected by the hadron calorimeters and the drift chambers have well-defined values of mass, charge, energy, and momentum.

The studies of the measurement errors as well as the unfolding procedure described above in Sect. 5 presuppose that the detected particles have unknown *true values* of the kinematic variables with a probabilistic distribution that can at least *approximately* be recovered by the unfolding procedure. By referring to the corrected data as "observed data" (Berge et al., 1991, p. 195), the physicists express an implicit scientific realism about an approximately true event distribution recovered by the unfolding procedure.

Epistemically, the iteration procedure is based on a coherence criterion for an adequate measurement, according to which an approximately true distribution of the data is found when the iteration procedure converges, that is, when input and output of the iteration agree. To combine a coherence criterion regarding adequate data correction with a correspondence view regarding the true event distribution does perhaps not conform to standard knowledge about philosophical truth theories. However, the use of coherence *criteria* does not necessarily preclude a correspondence *theory* of truth. This is shown by every legal judgement that is based only on circumstantial evidence, and we should not forget that experiments (following Kant, 1998, B xiii) are just circumstantial trials against nature.

Nor does the physicists' realism about physical magnitudes conform to any of the positions defended or rejected in the debate on scientific realism. Indeed, since the rise of quantum theory the practice of physics has rarely corresponded to philosophical standard knowledge. That is, much of philosophical standard knowledge is at odds with well-established results of contemporary physics. Nevertheless, we should make efforts to connect them. In standard terms of the realism debate, one may say that the physicists' implicit scientific realism includes at least two kinds of belief: belief in the truth of well-confirmed laws, such as the conversation laws that hold for the kinematic and dynamic properties of subatomic particles, and belief in the existence of the causal powers carried by these particles, e.g. in Cartwright's sense of capacities (Cartwright, 1989) or Hacking's entity realism, as expressed in his famous reality criterion "if you can spray them, then they are real" (Hacking, 1983, p. 23).

I should add that this version of scientific realism is complicated by the very fact that due to the quantum nature of particle tracks and scattering events, an entity realism falls short of any adequate interpretation of the measured quantum phenomena of CDHS or any other scattering experiment of HEP. The measured particle tracks and the scattering events which are supposed to have caused them are reconstructed in terms of quasi-classical entities. However, according to quantum field theory as the relevant well-established background theory, the underlying entity is a quantum superposition of scattering amplitudes interpreted in terms of probabilistic expectation values. Therefore, concerning the scattering experiments of HEP the distinction between entity realism and a realism of theories is not only blurred for reasons such as discussed by Massimi (2004), concerning the naïve parton model vs. the dynamic properties attributed to the quarks, but in general. The more modest realism with respect to the physical properties measured in the experiments, which I have suggested above, is also far from being clear. To discuss these problems in more detail is however beyond the scope of the present paper.

CRediT authorship contribution statement

Brigitte Falkenburg: Writing – original draft, Writing – review & editing.

Acknowledgements

I would like to thank three anonymous referees for their very helpful critical comments on earlier versions of the paper and for drawing my attention to additional relevant literature.

References

Abramowicz, H., et al. (1984). Measurement of neutrino and antineutrino structure functions in hydrogen and iron. *Zeitschrift für Physik C: Particles and Fields, 25*, 29–43.

Antoniou, A. (2021). What is a data model? An anatomy of data analysis in high energy physics. *European Journal for Philosophy of Science, 11*, 101. https://doi.org/10.1007/s13194-021-00412-2

Beauchemin, P.-H. (2017). Autopsy of measurements with the ATLAS detector at the LHC. *Synthese, 194*, 275–312.

Beisbart, C., & Norton, J. D. (2012). Why Monte Carlo simulations are inferences and not experiments. *International Studies in the Philosophy of Science, 26*, 403–422.

Berge, P., et al. (1991). A measurement of differential cross-sections and nucleon structure functions in charged-current neutrino interactions on iron. *Zeitschrift für Physik C: Particles and Fields, 49*, 187–223.

Bjorken, J. D. (1969). Asymptotic sum rules at infinite momentum. *Physics Reviews, 179*, 1547–1553.

Bjorken, J. D., & Paschos, E. A. (1969). Inelastic electron and γ –proton scattering and the structure of the nucleon. *Physics Reviews, 185*, 1975–1982.

Blobel, V. (1984). Unfolding methods in high-energy physics experiments. In *1984 CERN school of computing* (Vol. 88, pp. 88–127). https://doi.org/10.5170/CERN-1985-009

Blobel, V. (2013). In O. Behnke, et al. (Eds.), *Data analysis in high energy physics: A practical guide to statistical methodsUnfolding* (pp. 187–225). Weinheim: Wiley-VCH.

Cartwright, N. (1989). *Nature's capacities and their measurement*. Oxford: Clarendon Press.

Chang, C., et al. (1975). Observed deviations from scale invariance in high-energy muon scattering. *Physical Review Letters, 35*, 901–904.

Falkenburg, B. (1986). Bestimmung von Nukleonstrukturfunktionen aus Neutrino-Eisen-Streuung. PhD Thesis, University of Heidelberg. Results published in: Berge et al. (1991).

Falkenburg, B. (2007). *Particle metaphysics. A critical account of subatomic reality*. Heidelberg: Springer.

Falkenburg, B. (2012). Pragmatic unification, observation and realism in astroparticle physics. *Journal for General Philosophy of Science, 43*, 327–345.

Feltesse, J. (2012). Introduction to deep inelastic scattering: Past and present. In *Proceedings of the XX international workshop on deep-inelastic scattering and related subjects*. n: University of Bon. https://indico.cern.ch/event/153252/contributions/1396991/attachments/160442/226595/6-feltesse_joel.pdf. (Accessed 7 June 2023).

Friedman, J. I., & Kendall, H. W. (1972). Deep inelastic electron scattering. *Annual Review of Nuclear and Particle Science, 22*, 203–254.

Hacking, I. (1983). *Representing and intervening*. Cambridge: Cambridge University Press.

Hartmann, S. (1996). In R. Hegselmann, et al. (Eds.), *Simulation and modelling in the social sciences from the philosophy of science point of view. Theory and decision libraryThe world as a process: Simulations in the natural and social sciences* (pp. 77–100). Dordrecht: Kluwer.

Hofstadter, R., Bumiller, F., & Yearian, M. (1958). Electromagnetic structure of the proton and neutron. *Reviews of Modern Physics, 30*, 482–497.

1787 Kant, I. (1998). Riga: Hartknoch. Engl.. In P. Guyer, & A. W. Wood (Eds.) (2nd ed., B)*The critique of pure reasonKritik der reinen Vernunft*. Cambridge: Cambridge University Press.

Karaca, K. (2013). The strong and weak senses of theory-ladenness of experimentation: Theory-driven versus exploratory experiments in the history of high-energy particle physics. *Science in Context, 26*, 93–136.

Karaca, K. (2018). Lessons from the large hadron collider for model-based experimentation: The concept of a model of data acquisition and the scope of the hierarchy of models. *Synthese, 195*, 5431–5452.

Karaca, K. (2020). In S. Leonelli, & N. Tempini (Eds.), *Data journeys in the sciencesWhat data get to travel in high energy physics? The construction of data at the large hadron collider* (pp. 45–57). Cham: Springer Nature.

Karaca, K. (2022). Two senses of experimental robustness: Result robustness and procedure robustness. *The British Journal for the Philosophy of Science, 68*, 279–298.

Massimi, M. (2004). Non-defensible middle ground for experimental realism: Why we are justified to believe in colored quarks. *Philosophy of Science, 71*, 36–60.

Massimi, M., & Bhimji, W. (2015). Computer simulations and experiments: The case of the Higgs boson. *Studies in History and Philosophy of Modern Physics, 51*, 71–81.

Mättig, P. (2019). In C. Beisbart, & N. J. Saam (Eds.), *Computer simulation validation. Fundamental concepts, methodological frameworks, and philosophical perspectivesValidation of particle physics simulation* (pp. 631–660). Cham: Springer Nature.

Merz, M. (1999). Multiplex and unfolding: Computer simulation in particle physics. *Science in Context, 12*, 293–316.

Morgan, M. S., & Morrison, M. (Eds.). (1999). *Models as Mediators: Perspectives on Natural and Social Science*. Cambridge: Cambridge University Press.

Morrison, M. C. (2009). Models, measurement and computer simulation: The changing face of experimentation. *Philosophical Studies, 143*, 33–57.

Morrison, M. C. (2014). Values and uncertainty in simulation modelling. *Erkenntnis, 79*, 939–959.

Morrison, M. C. (2015). Reconstructing reality. In *Mathematics, models and simulations*. Oxford: Oxford University Press.

Panofsky, W. H. K. (1968). In J. Prentki, & J. Steinberger (Eds.), *Proceedings of 14th international conference on high energy physicsLow q2 electrodynamics, elastic and inelastic electron (and muon) scattering* (pp. 23–39). Geneva: CERN.

Parker, W. (2017). Computer simulation, measurement, and data assimilation. *The British Journal for the Philosophy of Science, 68*, 273–304.

Parker, S., et al. (1984). In C. G. Wohl, et al. (Eds.), *Review of particle properties: Particle data group. Rev. Mod. Phys.: Vol. 56. Particle detectors, absorbers, and ranges* (pp. S48–S52).

Riordan, M. (1987). *The hunting of the quark*. New York: Simon & Schuster.

Riordan, M. (1992). The discovery of quarks. *Science, 256*(5061), 1287–1293. and SLAC-PUB-5794 https://www.slac.stanford.edu/pubs/slacpubs/5500/slac-pub-5724.pdf (access: 13.08.1922).

Ritson, S., & Staley, K. (2021). How uncertainty can save measurement from circularity and holism. *Studies in History and Philosophy of Science, 85*, 155–165.

Rosenbluth, M. N. (1950). High energy elastic scattering of electrons on protons. *Physics Reviews, 79*, 615–619.

Rossi, B. (1952). *High energy particles*. New York: Prentice Hall.

Saam, N. J. (2017). What is a computer simulation? A review of a passionate debate. *Journal for General Philosophy of Science, 48*, 293–309.

Shapere, D. (1982). The concept of observation in science and philosophy. *Philosophy of Science, 49*, 485–525.

Tal, E. (2012). *The epistemology of measurement: A model-based account*. Toronto: Dissertation. https://tspace.library.utoronto.ca/bitstream/1807/34936/1/Tal_Eran_201211_PhD_thesis.pdf.

van Fraassen, B. C. (2008). *Scientific representation*. Oxford: Oxford University Press.