Order out of Chaos? A Case Study in High Energy Physics

Rafaela Hillerbrand

Department of Values, Technology & Innovation, Delft University of Technology

In recent years, computational sciences such as computational hydrodynamics or computational field theory have supplemented theoretical and experimental investigations in many scientific fields. Often, there is a seemingly fruitful overlap between theory, experiment, and numerics. The computational sciences are highly dynamic and seem a fairly successful endeavor—at least if success is measured in terms of publications or engineering applications. However, for theories, success in application and correctness are two very different things; and just the same may hold for "methodologies" like computer simulations. A lively debate on the epistemic status of computer simulations has thus emerged within the philosophy of science. This paper discusses possible problems when computer simulation and laboratory experiment are intertwined. In present experiments, stochastic methods in the form of Monte Carlo simulations are often involved in generating experimental data. It is questioned as to how far a realistic stance can be maintained when such stochastic elements are involved. Taking experiments in high energy physics as a study case, this paper contends that using these types of entangled material and numerical experiments as a source of new phenomena or for theory testing must presuppose a certain understanding of causality and thus binds us at least to a weak form of realism.

Keywords: philosophy of simulation, causality, experiment

1. Introduction

In recent years, computational sciences are understood as the development, exploration and numerical implementation of computational models supplemented theoretical and experimental investigations in many scientific fields. Particularly in the physical sciences like computational hydrodynamics or computational quantum field theory there is often a seemingly fruit-

Corresponding author's address: Rafaela Hillerbrand, Philosophy Section, Department of Values, Technology & Innovation, TPM-Faculty of Technology, Policy & Management, TU Delft, Jaffalaan 5, NL-2628 BX Delft, Netherlands. Email: r.c.hillerbrand@tudelft.nl.

Online ISSN: 1736–5899
Studia Philosophica Estonica (2012) **5.2**, 61–78
www.spe.ut.ee
Published online: December 2012
http://dx.doi.org/10.12697/spe.2012.5.2.05

ful overlap between theory, experiment, and numerics. The computational sciences are highly dynamic and seem a very successful endeavor—at least when success is measured in terms of publications or engineering applications (cf. Hillerbrand 2012). We know, however, that for theories, success in application and correctness are two very different things; and just the same for "methodologies" like computer simulations. A lively debate on the epistemic status of computer simulations has emerged within the philosophy of science.[1] While some see computer simulations as epistemically on a par with real, i.e. laboratory experiments (e.g. Parker 2009, Morrison 2009), others deny that the empirical flavor that goes along with numerical programming results in an epistemic similarity between material and numerical "experimenting" and thus view computer simulations as means of theoretical inquiry (e.g. Oreskes et al. 1994).

As characteristic for philosophy of science after the practice turn (cf. Rouse 2002, Soler et al. 2012), the debate on the epistemic status of computer simulations is commonly developed by means of case studies. However, one must be aware of hasty inferences from case studies to statements on the epistemic payoff of computer simulations generally: computer simulations are powerful because versatile instruments and any analysis of their epistemic payoff must be sensitive to the various purposes for which they are used in the sciences. I want to distinguish three types of simulations as regards their epistemic content or aim.

Type-I simulations refer to simulations that aim at information about abstract, most often mathematical systems. An example for this type of simulation is the "proof" of the four color theorem, [2] however, most of the time, the systems under investigation are differential equations that cannot or cannot yet be solved analytically. The search for finite-time singularities in the three-dimensional incompressible Euler equations provides an excellent example where the numerical investigation of an abstract system, i.e. the Euler differential equation, is of practical importance for research in the empirical sciences, in this case fluid dynamics (e.g. Grauer et al. 1998). In this first sense, simulations yield a (possibly preliminary) alternative for a lack of theoretic understanding.

Type-II simulations provide information on systems that cannot or cannot yet be accessed experimentally or are simply very hard to access in real laboratory or field experiments. Examples here are very diverse. (a) Physicists may use this type of simulation for analyzing turbulent flows on scales

[1] See, fore example, (Humphreys 1991, 2004, 2009, Hartmann 1996, Hughes 1999, Morgan 2003b,a, Frigg and Reiss 2009, Winsberg 2009, Parker 2009, Morrison 2009, Giere 2009).

[2] (Appel and Haken 1977). Note that this proof is not accepted as such uniformly amongst mathematicians.

too small to access in laboratory experiments, but information on the behavior on these small scales is very important for refining or testing existing theories. (b) In numerical experiments, certain effects can be singled out that cannot be detangled in material experimenting. When analyzing inertial particles in any sort of flow, for example, the numerical simulation has the advantage that one may focus on the particles' inertia only while neglecting effects like gravitational interaction or the particles' finite size. These effects cannot be decoupled in real experiments. Moreover (c) the analysis of some real, i.e. material experimental data may rely on simulations, usually on the form of Monte Carlo simulations. In this sense, simulations may be seen as a (possibly preliminary) replacement of experiments.

Type-III simulations may be seen as a kind of instrument for prognoses or forecasting. Here, simulations are used for predicting the behavior of real, usually complex systems for which (a) no accepted analytic description exists, or for which we are (b) certain that the theoretical description implemented numerically is correct (within the desired precision). Typical examples for the former arise in the engineering sciences, in weather and climate predictions while the latter type is often studied in astrophysics (cp. Morrison 2009) or engineering sciences.

Note that often scientists use the very same simulation for various purposes. Practically in applied sciences or engineering applications type-I and type-II simulations seem to mix fairly commonly. Moreover, one may argue that all simulations give information on mathematical systems as in one way or the other it is an abstract mathematical model that is numerically implemented. Here, a distinction made, for example, by S. Hartmann with the terms 'discrete' and 'continuos' simulation models seems of importance. While for the latter the corresponding dynamic model is conveniently formulated in the language of differential equations (Rohrlich 1991), discrete simulations are based on a discrete space-time structure right from the beginning (Wolfram 1994). A prominent example of a discrete model is the game of life simulation, di Paolo and Bedau call these types of discrete models "simulation models".

Though in scientific practice, all three types of simulations may appear jointly, distinguishing these three epistemic aims is of use when discussing epistemic issues related to computer simulations. In particular it may help to judge the applicability of extreme positions. For example Oreskes *et al.* (1994) claim that computer simulations lack any empirical content—which, if at all, can make sense only for type-I simulations, the same is true for claims that see computer simulations mainly as thought experiments, a claim not applicable, at least not in a straightforward way, to type-II simulations. Moreover it may seem that claims on the epistemic status of computer simulations per se seem unjustified as while some computer simulations may indeed be epistemically on a par with material experiments, other may be very close to theoretical investigation (Hillerbrand 2012). A large part of the philosophical discussion on the epistemic status of computer simulations thereby focuses on the demarcation of material and numerical experiments. So far , little attention has been paid to data-generation where computer simulations and material experimenting are inevitably intertwined (type-II (c) simulations above). Not only are these types of simulations fairly common in today's experimental sciences, examples range from data-processing in fMRI scans (Amaro and Barker 2006) to experiments in high-energy physics, but also the study of whether and in case how, the involvement of computer simulations alters the epistemic status of the gathered data may shed some light on the question as to whether simulations are rather means of theoretical or of experimental inquiry. The question to be addressed in this paper is therefore whether the involvement of simulations in the process of material-experimental data-generation alters the epistemic status of the data or whether it is nothing but a possibly severe case of theory-ladeness of observation. As the involvement of computer simulations occurs already at a very early stage of data generation, this question is of high relevance for a practical realist account of how science works. In this paper, experiments in high-energy physics are used as a study case as here a fairly profound understanding of the involvement of Monte-Carlo simulations and a somewhat elaborate mathematical background theory exists.

In particular, I want to study the so-called HERMES effect (Ackerstaff *et al.* 2000), a prominent and allegedly new experimental observation that turned out to be a numerical artefact some years after its first publication (Airapetian *et al.* 2003). The erratum showed convincingly for everybody in the particle physics community that what became known as the HERMES effect merely resulted from an error in the Monte-Carlo simulation used in processing the data. Before discussing how the involvement of stochastic processes changes the nature of data generation in section 4, let us begin with a brief philosophical account of experiments in section 3. Here I want to draw on Duhem's account of experiments and, in particular, Heidelberger's adaption of Duhem's account. In their words, the involvement of Monte-Carlo methods may spoil the "causal level" of experimenting. A detailed analysis of the HERMES experiment however reveals that though the involvement of computer simulations provides a severe case of theory ladenness, the problems leading to the observation of an allegedly new phenomena had nothing to do with the simulation per se (section 5). In the final section it is asked as to how far a realistic stance can be maintained when such stochastic elements are involved. Taking experiments in high energy physics as a study case, I will contend that using these types of experiments as a source of new phenomena or for theory testing must presuppose a certain understanding of causality and thus binds us at least to a weak form of realism. Before I want to dilate a little the usage of the term computer simulation in section 2 as this topic has recently attracted much attention among philosophers of science.

2. Computer Simulations and Monte Carlo

Computers may be used simply as calculators, but offer far more potential. To distinguish simple number crunching from more sophisticated numerics, S. Hartmann (1996) and others use the term simulation to denote the imitation "of one process by another" (Hartmann 1996, 83; cp. Parker 2009). Here, 'process' refers to some temporal sequence of states of a system, thereby stressing the dynamic aspects of (not only computer) simulations. By contrast, P. Humphreys (1991) adopts a broader notion of computer simulations:

> A computer simulation is any computer-implemented method for exploring the properties of mathematical models where analytic methods are unavailable.

But computer simulations may be of great value even where analytic solutions are known, for example via computer aided visualizations. I thus want to broaden Humphrey's notion and use the term 'numerical experiment' to refer to any computer-implemented method that is non-analytic. It is these numerical experiments that are at the core of this paper. Note that not all numerical investigations aim at simulations and as such at dynamical aspects. Most and the most interesting computer experiments, however, are simulations in the sense that they mimic a dynamic sequence of states. Following the common parlance in the sciences and in philosophy, I will thus sometimes use the terms computer simulation and numerical experiment synonymously though strictly speaking they are not interchangeable.[3] Note particularly that most scientific investigations seem to be concerned with dynamic processes. Even when explaining such stationary phenomena like rock formation, for example, one often falls back on dynamic explanations and thus simulations in Hartman's sense. As the term 'material experiment' is commonly not restrained to simulations in this paper, just the same computer-aided investigations are not reduced to simulations in this paper,

[3] Note in particularly that, as has been noted in the literature, both terms are problematic. While 'numerical experiment' seems to presuppose that numerics is epistemically on a par with material experimenting, the term 'computer simulation' seems too narrow as it raises the connotation of the modeling of a dynamic process, see main body of the text.

though simulations in the narrow sense may indeed be by far the largest and most interesting class amongst numerical experiments.

Note in particular that Monte-Carlo methods as a class of computational algorithms, which rely on repeated random sampling to compute their results, are simulations in the sense of any of the definitions introduced above. These Monte-Carlo simulations are used in data-generation when the experimentally detected data itself is more or less meaningless to the experimenter—scientists use the terms 'uninterpreted' or 'raw data'.[4]. The additional information needed to interpret the data may involve computer simulations, particularly simulations like Monte-Carlo methods that entail genuine stochastic, i.e. random processes.

Such types of experiments are fairly common in today's sciences. Examples range from data detection in particle physics to imaging techniques as they are used in fMRI scans, for example (Ward 2000, Amaro and Barker 2006). Commonly stochastic methods are employed via Monte Carlo simulations. The catch phrase "order out of chaos" in the title of this article refers to the fact that Monte-Carlo simulation and thus a stochastic process is involved in interpreting the data, so with a random process you sort of create order in a seemingly random series. A procedure that seems, at least at first glance, somewhat odd.

Monte-Carlo methods are fairly popular in many parts of the numerical sciences—ranging from mathematics and theoretical physics to economics and experimental sciences. Though the analysis in this papers focusses on the use of Monte-Carlo simulations in data generation from material experiments, I want to first introduce the principle idea behind these type of simulations in a fairly general fashion, the application to experiments is discussed in section 5.[5]

Before the physicists J. von Neumann, S. Ulman and N. Metropolis, who worked at the time at the Manhattan project, introduced Monte Carlo simulations in the 1940s at Los Alamos National Laboratory, statistical sampling in simulations was narrowed down to estimating uncertainties in simulations of deterministic problems. The Monte Carlo simulation inverted this approach and solved a deterministic problem with the help of probabilistic analogues. Though Monte-Carlo methods vary, we may distinguish the fol-

[4] Using Galison's terms, almost all instruments used in the big collider experiments are in the logical or electronic tradition (Galison 1997) and thus the raw data is meaningless to the experimenter. Raw data in less elaborate particle physics experiment, however, such as the electrical signals from a photo multipliers may be perfectly reasonable.

[5] Note that in this paper the notions 'Monte Carlo *method*' and 'Monte Carlo *simulation*' are used in a general fashion and loosely synonymously because the distinction between these terms as drawn by some authors, e.g. Sawilowsky 2003, seems hard to maintain (Kalos and Whitlock 2008).

lowing particular steps that seem together characteristic for all Monte-Carlo simulations (e.g. Kalos and Whitlock 2008):

1. The domain of possible inputs is defined.

2. Input is generated randomly from a probability distribution over the domain.[6]

3. A deterministic computation on the inputs is performed.

4. The results are aggregated.

As an illustration we may consider how this procedure is used to estimate the number π: (1) By drawing a square on the ground and inscribing a circle within it, the domain of input is defined. (2) Scattering objects of uniform size (drops of water, grains of sand, or needles) uniformly over the square, and then (3) counting the number of drops that landed inside the inscribed circle to the total number of drops gives the ratio of the two areas. (4) The ratio of these areas (πr^2 for the inscribed circle, $(2r)^2$ for the square) is calculated analytically as $\pi/4$. This gives a probabilistic estimate of π and a instructive example of how Monte-Carlo methods are used in mathematics. The range of application is fairly diverse and not limited to genuine stochastic modeling as depicted by the estimation of the number π. With the help of the so-called Feynman-Kac formula that links non-stochastic parabolic partial differential equations to stochastic processes, Monte Carlo simulations can be used for the integration to a deterministic differential equation.

Involving Monte-Carlo simulations in material experimenting may bring about a very severe theory-ladeness of our observations. This seems rather uncontroversial. In this paper I want to raise a further question, namely whether involving genuine stochastic methods renders the constructive and productive function of experiments impossible and thus data processed via Monte Carlo would not be able to yield new phenomena, for example. Many high energy physics experiments like the ones performed at the Large Hadron Collider (LHC), which is currently the world largest and highest-energy particle accelerator and is operated by the European Organization for Nuclear Research, known as CERN, distinguish themselves from other

[6] To limit computational time, the input in step (2) is often not generated by random sampling, but by using quasi or pseudo random sequences instead. Often these procedures are referred to as Quasi or Pseudo Monte Carlo simulations. The used sequences are totally deterministic, so the popular name quasi-random may seem misleading. The sequences do, however, exhibit statistical randomness. This article thus does not dwell on this issue any further and acts as if all Monte-Carlo simulations use indeed random sequences. This approach is motivated by the fact that the pseudo random sequences are indeed to mimic statistical randomness.

types of data production by the involvement of stochastic process in generating the data, in particular the involvement of Monte Carlo simulations in calibrating the detectors. And this is indeed no peculiarity of the HERMES experiment, but rather a feature of almost all particle detectors at the large particle physics experiments at the large particle colliders, even at CERN's LHC.

3. Two-level Account of Material Experiments

Before analysing in detail how Monte-Carlo simulations are used in high energy physics, let us first look at how "normal" experiments, i.e. those without computer simulations, work and thereby follow the distinction of two levels of experiment as introduced by Pierre Duhem and systematized by Michael Heidelberger. According to this account, one can distinguish a *causal* and a *symbolic* level of experiments.[7] Duhem distinguishes experiments in, for example, physiology from those in physics. While for the former, experimental results may be understood without a (deep) theoretical understanding and can be captured with the *causal expressions of our ordinary everyday language*, the latter necessarily and unavoidably interpret the observation within a symbolic system, provided by some already existing theory. Following Duhem, only in the latter—i.e. in the mature sciences—theory-ladeness of the observation is a problem.

Physiology is paradigmatic for what Duhem refers to as "…sciences …where the experimenter reasons directly on the facts by a method which is only common sense brought to greater attentiveness but where the mathematical theory has not yet introduced its symbolic representations" (Duhem 1906, 180). For sciences like physiology, experiments serve to improve and

[7] Note that, following Heidelberger (2003), these two levels or "roles" always need to be present irrespective as to whether we consider an experiment in its productive or its constructive or its representative function. Productive experiments generate phenomena that are usually not part of our environment—for example the creation of elementary particles, free nuclei or strange matter in high energy physics experiments, but also the generation of vacua, X-rays, etc. One may distinguish from these productive instruments constructive ones—though there is some overlap and it is a continua between both cases. In its constructive function, an experiment manipulates phenomena in such a way that they become accessible in the laboratory. The constructive instruments coincide with McMullin's (1985) causal idealization in which phenomena are sort of liberated from spurious side-effects as done in imitating experiments that mimic strokes of lightning or the air flow around a car. An example for such an instrument is the Leyden jar or experimental setups that try to mimic or imitate real phenomena. An example for the latter are wind tunnel experiments. The representational function of an instrument becomes important when measurement is involved. For example, when measuring temperature, one sensorial experience, namely heat, is represented by another usually audial one, for example the height of a mercury column.

enhance our understanding of the natural environment in terms of causal relations. It is causal relations that one discovers in these simple experiments. Within sciences like physics where a mathematical theory has already been successfully introduced—Duhem refers to these as "mature sciences"—the aim of experiment is to incorporate new experimental findings into the already existing symbolic notion. So for mature science, experiments cannot reveal new phenomena in the sense of new causal relations. The scientists only inscribes phenomena in terms of an abstract and symbolic structure. The raw data of our every day experience is replaced by abstract and symbolic representations in order to be manageable. "The physicist can no sooner conceive the concrete apparatus without associating with it the idea of the schematic apparatus than a Frenchman can conceive an idea without associating it with the French word expressing it. This radical impossibility, preventing one from dissociating physical theories from the experimental procedures appropriate for testing these theories, complicates this test in a singular way, [...]" (Duhem 1906, 183).

Following Duhem, in the historical evolution of sciences, the experiment first aims to bring information on causal relations and thus the experiment reveals new information about some phenomena—experiments may be constructive or productive. In the more mature sciences, however, the observation is highly theory-laden; the observed only makes sense when interpreted in some symbolic interpretation system. Duhem applies his two-level account of the experiment mainly to the historical evolution of the sciences. Michael Heidelberger has argued more recently that in all experiments, even within the mature sciences, both levels are present (Heidelberger 2003): There is always a causal and always a representational level involved. Moreover, only when the causal level is present, experimental results can have significance without the abstract and symbolic representation provided by some background theory: Only when one understands an experiment as a causal manipulation of instruments, then the experiment may have a productive or constructive function as attributed to it not only by Heidelberger, but famously by Hacking and other new experimentalists (Hacking 1983).

Heidelberger's reconstruction of Ohm's experiments which led to the formulation of Ohm's law relating electrical current I and voltage U via the resistance R of the respective circuit element elucidate how the causal level is present even in experiments in the mature sciences. Ohm's investigation of the interaction of his instrument (which consists in magnetic centerpins) with wires in an electric circuit resulted first in a causal concept, namely that a centerpin is deflected in a certain way when electricity runs through a wire in its close vicinity. This causal connection can be experienced without theoretical background information. Only in a second step did this causal understanding lead to an improvement and modification of the symbolic interpretation. In particular the findings could not be implemented into the existing theoretical (symbolic) apparatus, partly they necessitated a new symbolic representation of resistance. So the causal and thus in Duhem's and Heidelberger's understanding theory-free aspects of an experiment, is necessary to understand an experiment as a causal manipulation of instruments. These instruments then become important in gathering data via their productive and constructive or representative function.

At first glance, involving stochastic processes in generating the data seems to spoil the causal function of experiments. So the question to be addressed in particular in the following when analyzing the data generation that lead to the HERMES effect is, whether involving genuine random processes in the form of Monte-Carlo simulations in the data processing spoils the causal level and hence these types of experiments cannot, for example, yield any information on new phenomena.

4. Monte Carlo Simulations in Data Generation: The HERMES Effect

When computer simulations are involved in data generation it is almost always in the form of so-called Monte-Carlo methods. As detailed in section 2, these are a class of computational algorithms that rely on repeated random sampling to compute their results (Metropolis and Ulam 1949, Manno 1999). As genuine stochastic processes are involved here, it stands to reason whether this stochastic nature of the data-generation spoils the causal level of experimenting that following Heidelberger and Duhem needs to be present for experiments to develop their representative, constructive or productive function. Before addressing this question from a more generic standpoint, this section analyzes a specific misinterpretation of an experimental result, the HERMES effect, that was attributed to the Monte-Carlo simulation involved in data-generation.

The HERMES experiment ran from 1995 till 2007 to investigate the spin structure of the nucleon, i.e. protons and neutrons (e.g. Avakian *et al.* 1998). To study the nucleon structure, electrons or positrons, i.e. point-like, charged elementary particles, were accelerated to high energies in a circular ring of about 6.5 km circumference. At HERMES this beam of high energy electrons was then brought to a collision with gas injected into the beam line. The electrons would collide with the nucleus (the protons and neutrons) of the gas, destroying it and producing new particles. Specifically, the electrons would collide with the quarks and gluons which make up the nucleons. The kind and properties of the the produced particles allow to study the internal structure of the protons and neutrons, that means, the distribution of quarks and gluons inside the nucleon.

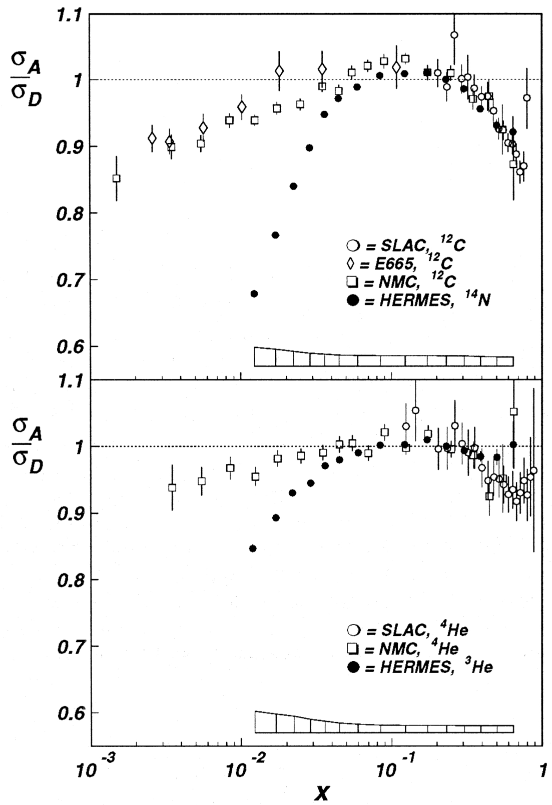

Fig. 1. *The HERMES effect. Cross-sections σ for various nucleons as reported by Ackerstaff et al. 2000 (HERMES Collaboration). Details see text.*

Figure 1 shows the ratio of cross-sections σ for various nucleons is plotted as a function of the variable x-Bjorken. This variable can be understood as the fraction of the nucleon energy carried by the quark on which the electron scattered. Shown is the ratio of the cross section of a heavier target with nuclear mass capital A (σ_A) to the cross section of Deuterium (σ_D). All experiments show a slight decrease of this ratio towards smaller x. Very striking is the fact that in the HERMES data, this ratio drops rather dramatically (for a slightly heavier target (nitrogen) N^{14} instead of (carbon) C^{12}). Obviously, an explanation for this very different behaviour was sought. Theories were developed and published—even in peer-reviewed journals.

5. Monte-Carlo Simulations in Data Processing

When we follow Duhem's and Heidelberger's account of experiment, the causal aspects of manipulating instruments are necessary for an instrument to have some theory-independent output. Now it seems that if the measurement instrument itself invokes genuine random stochastic processes as is the case when Monte-Carlo simulations are involved in data processing, then the causal level of the experiment is endangered, challenging our interpretation of experiments as causal manipulation of instruments. Does the involvement of Monte-Carlo simulations and thus of a genuine stochastic processes spoil the causal level of experimenting and is this why we ended up with the HERMES effect as a numerical artifact? To answer this question, let us look in more detail how Monte-Carlo simulations were used in interpreting data at HERMES. This procedure is fairly general and not peculiar to the HERMES experiment, rather Monte-Carlo are used the same way at all particle accelerators.

Before actually being able to use a detector in high energy physics experiment, one actually runs a simulation of this very detector, whereby two Monte Carlo processes are involved. (i) With a so-called *event generator* the scattering processes as they are expected to occur in the real collision as set up by the collider experiment are simulated. The stochastic nature of the Monte-Carlo method is of importance here as the theory underlying the numerical model, i.e. the Standard Model of particle physics, takes scattering as a stochastic event. The results of this first Monte-Carlo simulation determine "what can be seen" by an ideal detector (particles, energies, momenta, angles, …)—provided that the theoretical model underlying the simulation is indeed correct. (ii) In a second step, a program (originally developed by CERN) is used to build a three-dimensional model of the real detector. Then the input of step 1 is used to determine where the particles that were created in the collision hit the detector material (again assuming that the theoretical model is correct). By using abstract models of the cross-sections, it is determined whether there are interactions between created particles and detector material. It is then simulated how the track of a scattered particle may be modified by the interaction with the detector material. Again, following physical theory, there are stochastic processes involved in the events and they are simulated with the help of Monte-Carlo methods.[8]

[8] Note that the two-step process of modeling and processing the experimental data gives, of course, only a very rough picture of the work of the high-energy physicist. A closer look would acknowledge various different computer simulations, for example, the first step in the main body of the text may be decomposed into a simulation of the actual particle generation in the collision and a simulation of the following parton showers and hadronization. These details are, however, of no relevance for the epistemic discussion.

The HERMES effect seems to indicate a severe theory-ladenness of experiments within high energy physics. Indeed, it illustrates very clearly at least two types of theory-ladenness distinguished by Kuhn. While *semantic theory loading* (Kuhn 1962, 127ff) expresses the idea that theoretical commitments of whatever sort exert a strong influence on observational descriptions, *salience* (Kuhn 1962, 123f) expresses the fear that scientists working with different theoretical frameworks (what Kuhn refers to as paradigms) may not look at the same thing when observing the same experiment. Kuhn illustrates this by Aristotle and Galilei watching pendulum swings. Aristotle would look at the weight of the pendulum, the vertical height to which it rose and the time required for it to achieve rest, while Galilei primarily measured and observed things like radius, angular displacement and time per swing. Galilei and Aristotle would not have collected the same data when looking at the same pendulum experiment.

As the reconstruction of the HERMES experiment above appears to reveal, high energy physics seems to illustrate Kuhn's theory-ladeness of experimental observations very lucidly. The output of the two-step simulation detailed above determines what can actually "be seen" with the help of this very detector. All interactions or possibly generated particles that were not already known by the theories underlying the models in step 1 and 2, are not detected by the observer. Experiments within high energy physics thus seem to be a paradigm case for what Kuhn referred to salience by working within a certain theoretical framework. Just like for salience, high energy physics experiments also provide good examples for semantic theory loading. The interpreted data can only be of the same format as the output of the second step of the simulation. However within Kuhn's or others framework of reasoning that focusses on theory-ladenness the involvement of stochastic processes in determining the interpreted data does not distinguish itself from more familiar types of data generation which do not involve stochastic processes.

But what actually did go wrong at HERMES? The critiques were right, it had to do with the Monte-Carlo simulation, but not actually with the simulation per se. The physics result of the HERMES experiment was a ratio of the cross section of nitrogen divided by the cross section of deuterium. However, events happening in the detector are recorded with a certain efficiency, that means a certain fraction of events is not recorded for various reasons. Also, apart from the scattering events of interest (in this case: scattering on the quarks inside the nucleons), also other events can happen (in this case particularly: elastic scattering of the electron on the whole nucleon, leaving the nucleon intact), which for the experimentalist are indistinguishable from the first kind. These are the so-called background events.

The finite efficiency of the detector as well as the ratio of background to real events are different for deuterium and nitrogen. So the measured cross section ratio between nitrogen and deuterium is not the real ratio: events are missing, others are included although they do not belong there, all that with different probability for numerator and denominator. To correct the measured result, as indicated above, a Monte-Carlo simulation of the scattering and the subsequent measurement in the detector is used. This simulation provides the correction to the measured result.

During the elastic scattering which contributes to the background, the beam electron is deflected (scattered) and in addition a photon with high energy is produced. This photon has a high probability to hit the beam pipe (a solid tube through which the electron beam passes through the detector). In reality, this creates a lot of secondary particles, causing the detector to be blinded and the data of this event being unusable. In reality, background events have rather high probability of not being recorded. In the simulation, however, the beam pipe was ignored and thus photons hitting the beam pipe. So in the simulated detector, no secondary particles were created and thus the detector (in the simulation) had no problem in detecting the scattered electron. The measured result was corrected for the undesired background. Since the faulty simulation showed too high a probability that the background events were indeed detected, too much background was subtracted. And since the likelihood of such background events (compared to the scattering events HERMES looked for) was proportional to the charge of the nucleus, the effect was larger for the heavier nitrogen target. Hence the final result showed a drop of the ratio $\sigma(Nitrogen)/\sigma(Deuterium)$, which in reality was not there. An erratum was published in 2002 and its results are in agreement with the other experiments (Airapetian *et al.* 2003).

6. The Causal Level of Experimenting and the Stochastic Nature of Data Generation

Summarizing the reconstruction of the origin of the HERMES effect in the last section, we may say that what actually went wrong at HERMES seems something very common in experimental practice. In particular, it seems to be no unique feature of the involvement of stochastic processes in the form of a Monte-Carlo simulation. When aiming at the measurement of the brightness of a star, for example, while some object (a fly say) gets in between the star and your measurement instrument without you realizing it, similar problems show up. This however does not answer the original question posed in this paper, namely as to whether the involvement of genuine stochastic processes in the form of Monte Carlo simulations actually results in epistemically different data.

Only if one is willing to buy a certain (though weak) metaphysical assumption, the involvement of Monte Carlo simulation does makes no difference on an epistemic level . Following Duhem and Heidelberger, it was argued that the causal level is indispensable for using the experimental results as hints for new phenomena, and not only for theory testing. The causal interaction between measuring device and target system can be maintained when Monte-Carlo simulations are involved only when we accept that causality is stochastic in nature.

Requiring a causal level in experiments like the one at HERMES requires that we do interpret causality in a certain way, namely: a probabilistic account of causality is required in order to make sense of high energy physics experiments. This may seem easy to buy for the (experimental or theoretical) particle physicist, however, it is more than a side note to philosophy of science debates on the metaphysics of causality. A practical realistic account of how science works must be sensitive to the metaphysical or rather protoscientific assumptions that have to be made to actually interpret measurement results as such. Note that the analysis in this article says nothing about the fundamental nature of causality, whether it is fundamental or may be reduced to other features. Here is where the practical realist account comes in handy: Neither does the world consist in self-identifying objects, nor are humans world makers. Not even the working scientists who uses instruments like particle generators to *generate* phenomena that they subsequently investigate creates the world. Niiniluoto (1999) nicely points out the interconnection between the human world-co-creator and the objective reality she investigates. By referring to the world as cake, Niiniluoto (1999, 222) writes "A cake can be sliced into pieces in a potentially infinite number of ways, and resulting slices are human constructions made out of the parts of the cake". While Niiniluoto and Vihalemm (2012) take the cake to be the world, by using constructive instruments like a particle accelerator in which by collision new particles are created, scientists rather seem to bake their own cake from given substances in the world. The resulting cakes may be different. This holds for every use of a constructive instrument. The investigation in this article however focused on an aspect peculiar for the involvement of Monte-Carlo methods in data generation. If we take the stochastic nature of the underlying process seriously, this implies that when a causal account of experiment is necessary, we have to accept a stochastic interpretation of causality. Using Niiniluoto's comparison, this, however, means that in order to be able to use our knife, i.e. the Monte-Carlo approach to high-energy physics data, we have to assume certain things about the world that we do not test in an experiment. And, moreover, we actually cannot test these in the experiment we are about to undertake.

Summarizing, this paper focusses on a specific form of using computer simulations in present day science, namely involving Monte Carlo simulations in generating experimental data. This implies a very severe case of theory-ladenness of observation. However, the involvement of stochastic processes, in the case of particle physics in the form of Monte Carlo simulations, does not spoil the causal level for which I try to argue that it is necessary to understand an experiment as a productive or constructive instrument in scientific progress. However, in order to be able to use these types of experiments in this way affords a certain metaphysical commitment in that sense that it forces us to take up a probabilistic (i.e. non regularity) view of causation.

Acknowledgments

I want to thank the participants of the Workshops "Practical Realism" (Tartu University, 2011) and "Modeling and LHC" (Wuppertal, 2012), as well as Johannes Lennard and Simon Friedrich for helpful comments on earlier versions of this manuscript.

Bibliography

Ackerstaff, K. *et al.* (2000). Nuclear effects on $R = \sigma_L/\sigma_T$ in deep-inelastic scattering, *Physics Letters* **B475**: 386–394.

Airapetian, A. *et al.* (2003). Erratum to: Nuclear effects on $R = \sigma_L/\sigma_T$ in deep-inelastic scattering, *Physics Letters* **B567**: 339–346.

Amaro, E. J. and Barker, G. (2006). Study design in fMRI: Basic principles, *Brain and Cognition* **60**: 220–232.

Appel, K. and Haken, W. (1977). Every planar map is four colorable. part i: Discharging, *Illinois Journal Mathematics* **21**: 429–490.

Avakian, H. *et al.* (1998). Performance of the electromagnetic calorimeter of the HERMES experiment, *Nuclear Instruments and Methods in Physics Research* **A417**: 69–78.

Duhem, P. (1906). *La Théorie physique — son object, et sa structure*, Chevalier et Rivière, Paris. cited after the English translation *The Aim and Structure of Physical Theory*.

Frigg, R. and Reiss, J. (2009). The philosophy of simulation: Hot new issues or same old stew?, *Synthese* **169**: 593–613.

Galison, P. (1997). *Image and Logic: A Material Culture of Microphysics*, University of Chicago Press, Chicago.

Giere, R. (2009). Is computer simulation changing the face of experimentation?, *Philosphical Studies* **143**: 59–62.

Grauer, R., Marliani, C. and Germaschewski, K. (1998). Adaptive mesh refinement for singular solutions of the incompressible Euler equations, *Physical Review Letters* **84**: 4850–4853.

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

Hartmann, S. (1996). The world as a process, *in* R. Hegselmann, U. Mueller and K. G. Troitzsch (eds), *Modelling and Simulation in the Social Sciences from the Philosophy of Science Point of View*, Springer, Dordrecht, pp. 77–100.

Heidelberger, M. (2003). Theory-ladeness and scientific instruments in experimentation, *in* H. Radder (ed.), *The Philosophy of Scientific Experimentation*, Pittsburgh University Press, Pittsburgh, pp. 138–151.

Hillerbrand, R. (2012). Computer simulation — The future of experiment or theoretical phantasies? Epistemic aspects of numerical experiments.

Hughes, R. (1999). The Sing Model, computer simulations and universal physics, *in* M. Morgan and M. Morrison (eds), *Models as mediators: Perspectives on natural and social science*, Cambridge University Press, Cambridge, pp. 97–145.

Humphreys, P. (1991). Computer simulations, *in* A. Fine, M. Forbes and L. Wessels (eds), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990*, University of Chicago Press, East Lansing, pp. 497–506.

Humphreys, P. (2004). *Extending ourselves. Computer science, empiricisms and scientific method*, Oxford University Press, Oxford.

Humphreys, P. (2009). The philosophical novelty of computer simulation methods, *Synthese* **169**: 615–626.

Kalos, M. and Whitlock, P. (2008). *Monte Carlo Methods*, Wiley-VCH.

Kuhn, T. (1962). *The Structure of Scientific Revolution*, University of Chicago Press.

Manno, I. (1999). *Introduction to the Monte Carlo Method*, Akademiai Kiado, Budapest.

McMullin, E. (1985). Galilean idealization, *Studies in the History and Philosophy of Science* **16**: 247–273.

Metropolis, N. and Ulam, S. (1949). The monte carlo method, *Journal of the*

American Statistical Association **44**: 335–341.

Morgan, M. (2003a). Experiments without material intervention: Model experiments, virtual experiments and virtually experiments, *in* H. Radder (ed.), *The Philosophy of scientific experimentation*, Pittsburgh University Press, Pittsburgh, pp. 216–235.

Morgan, M. (2003b). Model experiments and model in experiments, *in* L. Magnani and N. Nersessian (eds), *Model based-reasoning: Science, technology and values,*, Kluwer, New York, pp. 41–58.

Morrison, M. (2009). Models, measurment and computer simualtion: The changing face of experimentation?, *Philosphical Studies* **143**: 33–57.

Niiniluoto, I. (1999). *Critical Scientific Realism*, Oxford University Press, New York.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994). Verification, validation, and confirmation of numerical models in the earth science, *Science* **263**: 641–646.

Parker, W. (2009). Does matter really matter? Computer simulations, experiments, and materiality, *Synthese* **169**: 483–496.

Rohrlich, F. (1991). Computer simulation in the physical sciences, *in* A. Fine, M. Forbes and L. Wessels (eds), *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1990*, University of Chicago Press, East Lansing, pp. 507–518.

Rouse, J. (2002). *How Scientific Practices Matter: Reclaiming Philosophical Naturalism*, University of Chicago Press, Chicago.

Soler, L., Trizio, E., Nickles, T. and Wimsatt, W. (eds) (2012). *Characterizing the Robustness of Science: After the Practice Turn in Philosophy of Science*, Springer, Dordrecht.

Vihalemm, R. (2012). Practical realism: Against standard scientific realism and anti-realism, *Studia Philosophica Estonica* **5.2**: 123–456.

Ward, B. (2000). Simultaneous inference for FMRI data.
URL: *http://afni.nimh.nih.gov/pub/dist/doc/manual/AlphaSim.pdf*

Winsberg, E. (2009). A tale of two models, *Synthese* **169**: 575–592.

Wolfram, S. (1994). *Cellular Automata and Complexity*, Addison-Wesley, Reading.