Erkenn (2015) 80:551–574
DOI 10.1007/s10670-013-9541-5

The Best Humean System for Statistical Mechanics

Roman Frigg · Carl Hoefer

Received: 2 October 2013 / Accepted: 2 October 2013 / Published online: 16 October 2013

Abstract Classical statistical mechanics posits probabilities for various events to occur, and these probabilities seem to be objective chances. This does not seem to sit well with the fact that the theory's time evolution is deterministic. We argue that the tension between the two is only apparent. We present a theory of Humean objective chance and show that chances thus understood are compatible with underlying determinism and provide an interpretation of the probabilities we find in Boltzmannian statistical mechanics.

1 Introduction

Classical statistical mechanics (CSM) posits probabilities for various events to occur. Yet the theory's time evolution is deterministic. How can probabilities in a deterministic setting be understood? We begin by outlining our own version of Humean Objective Chance, which we call THOC (Sect. 2). We then briefly review the main tenets of CSM and discuss how it deals with probabilities (Sect. 3). We then argue that these probabilities fit the mould of THOC; and in doing so we qualify various aspects of our approach to chance (Sect. 4). These qualifications are best discussed in a more general setting, and so we widen the scope of our

R. Frigg (✉)
Department of Philosophy, Logic and Scientific Method, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK
e-mail: r.p.frigg@lse.ac.uk

C. Hoefer
ICREA, Universitat Autònoma de Barcelona, Bellaterra, Spain
e-mail: carl.hoefer@uwo.ca

C. Hoefer
Rotman Institute of Philosophy, University of Western Ontario, Stevenson Hall 2150F, London, ON N6A 5B8, Canada

discussion to probabilities in non-fundamental sciences beyond CSM. This brings us to a discussion of various presuppositions of THOC and a comparison with other approaches to chance (Sect. 5). In the last section we draw some conclusions.

2 Humean Objective Chance

Let e be an event, for instance a coin coming up heads when flipped, and let E be the proposition stating that e occurs.[1] The chance of E, $ch(E)$, is a real number in the interval [0, 1] such that:

(1) *ch* satisfies the axioms of probability;[2]
(2) *ch* is the correct plug-in for X in the *Principal Principle*; and
(3) *ch* supervenes on the Humean Mosaic in the right way,

where X is the statement that '$ch(E) = x$'. Chances thus defined are Humean Objective Chances (HOC)—'chances' for short. We call a statement like X a *chance rule*.[3] Chances are linked to setups, which are characterised by setup conditions: relative to a certain physical setup condition, the chance of E is x.[4] In the example of the coin, the setup conditions are that the coin is a (nearly) perfect short cylinder and has a homogenous mass distribution, is spun upwards at moderate speeds (thus avoiding that the coin decomposes), lands on a mushy surface that leaves the coin intact, etc. To keep notation simple we are not introducing a special index for a setup, but it must be remembered that all chance rules are linked to a specific type of setup conditions.

We use 'THOC' to refer to the entire theory of chance presented in this section. We take (1) to be unproblematic;[5] (2) and (3) need unpacking. Chances are guides to action, and the most important requirement on a theory of chance is that it show how chances can do that. This aspect of chances is enshrined in the *Principal Principle* (PP), which establishes a connection between chances and the credences (subjective degree of belief) a rational agent should assign to certain events. Roughly, PP says that a rational agent who knows the chance of e should have credence in e's occurrence equal to the chance of E, as long as the agent has no inadmissible knowledge relating to E's truth. More formally, PP is the rule that for all E, all X and all K

[1] The presentation of HOC in this section is based on Frigg and Hoefer (2010). The classical source is (Lewis 1980). But as we will see below, our version of HOC differs from Lewis' in important respects.

[2] As is well known, there are different axiomatisations of probability. Nothing in what follows depends on which axiomatisation we chose.

[3] We use this term in restrictive way: only statements having exactly that form are chance rules. Existence claims, statements about upper and lower bounds, or specifications of probability intervals are not chance rules in our sense.

[4] This point has been made by Hájek (2007) for propensities. His arguments readily carry over to any notion of chance.

[5] In this we disagree with Lewis, who thought it a major problem to prove that chances satisfy the axioms of probability. THOC *defines* chance, and a function that does not satisfy the axioms of probability cannot be a chance function.

$$cr(E|X \,\&\, K) = x, \qquad (1)$$

where '*cr*' stands for a rational agent's initial credence, and *K* is any *admissible* proposition.[6]

A crucial question is what counts as an 'admissible' proposition. This question has been discussed controversially, and we cannot revisit this discussion here.[7] Following Hoefer (2007), and in line with Lewis's first characterisation in his (1980), we take a proposition *P* to be admissible with respect to an outcome-specifying proposition *E* for chance set-up *S* iff *P* contains only the sort of information whose impact on reasonable credence about *E*, if any, comes entirely by way of impact on credence about the chances of those outcomes. Chance is a guide to action *when information about E's truth or falsity is not available*.

The essence of the requirement of admissibility is to exclude the agent's possession of other knowledge relevant to the truth of *E*, the kind of knowledge the possession of which might make it no longer sensible to set credence equal to objective chance. To use the usual (and silly) example: if you have a crystal ball that you believe reliably shows you future events and if your crystal ball shows you the coin toss landing tails and you trust the ball's revelations, then it would *not* be reasonable to set your credence in tails to 0.5 for that flip—you have inadmissible knowledge. This example helps make the notion of admissibility intuitively clear. It also points toward an important fact in our world: inadmissible evidence concerning future events is not something we typically have—if we did, then chances would be rather useless to have.

Let us now turn to (3). The *Humean Mosaic* (HM) is the collection of everything that actually happens; that is, all occurrent facts at all times. There is a question about what credentials something must have to be considered part of HM, and we will touch upon this question in Sect. 5.2. What matters for now is that irreducible modalities, powers, propensities, necessary connections and so forth are not part of HM.[8] That is the 'Humean' in Humean mosaic.

Supervenience requires that chances be entailed by the overall pattern of events and processes in HM. A simple example of supervenience is actual frequentism (the view that probabilities *are* actual frequencies): the overall pattern of events uniquely determines the relative frequency of an event (if it is well defined), and hence its probability. Note that actual frequentism has no frequency tolerance: there cannot be *any* difference between the chance of *E* and the frequency of *E*. This contrasts with propensity theories, which have maximal frequency tolerance, but fail to satisfy Humean supervenience. THOC strikes a balance between these extremes by

[6] Justifying PP is a thorny issue, and, unsurprisingly, one fraught with controversy. We refer the reader to Frigg and Hoefer (2010, Section 3.4) and references therein for a discussion.

[7] For a discussion see Hoefer (2007, 553–555 and 558–560). Some philosophers maintain that the PP needs no admissibility clause. For counter-arguments in favour of the necessity of an admissibility clause in PP see (Hoefer 2014).

[8] Note that even if you believe that the world does contain necessary connections, powers or propensities, it still *also* has a HM. The HM is just the panoply of actual events understood as purely occurrent, setting aside any modal aspects those events may possess. The Humean about chance then simply maintains that chance facts supervene on this HM.

requiring that chances supervene on HM, but not *simply*: THOC postulates that chances are the numbers assigned to *E* by probability rules that are part of a *Best System* (BS) of such rules, where 'best' means that the system strikes as good a balance as the actual events will allow of *simplicity*, *strength* and *fit*. We then say that a probability rule is Humean *BS*-supervenient on HM ('HBS-supervenes on HM', for short) iff it is part of a Best System. Clause (3) can now be made more precise: the function *ch* HBS-supervenes on HM.

Simplicity and strength are notoriously difficult to explicate. We present a characterisation of these that is precise enough for our purposes in Sect. 4. Throughout Sect. 4 we assume that chance rules are formulated in a natural language, thereby ruling out grue/blean-type predicates. We reflect in Sect. 5.3 on how strong this assumption is and on how damaging residual indeterminacies are to our project. Fortunately, *fit* is a less problematic concept. The *fit* of a system is a matter of the extent to which the actual course of events in HM is likely in light of that system: the more likely a system regards the actual course of history as being, the better its fit. This explication of fit has certain limitations. It is only readily applicable to certain simple sorts of chance system, and Humean mosaics with only finitely many chancy events. For, e.g., systems with continuous chance distributions and/or worlds with infinitely many chancy events, a different approach is required; we set aside this issue because the cases we are interested are covered by the current explication.[9]

3 Thermodynamics and Classical Statistical Mechanics

The behaviour of macroscopic systems like a gas in a box is to a good degree of approximation correctly described by thermodynamics (TD). TD introduces macrostates M_k (where k is an integer which depends on the specific properties of the system, which are characterised by the values of macroscopic parameters such as temperature, pressure and volume).[10] Macrostates pertain to the system as a whole and neither their definition nor their use depends on the microscopic makeup of the system. In a world which is phenomenally equivalent to ours but in which there are no atoms and matter is a Cartesian continuum, the macrostates of a TD system would be the same and all laws of TD would apply.

The most important of these laws is the so-called Second Law of TD, saying that the TD entropy of a closed system cannot decrease. The law needs to be qualified in two ways. First, the strict version of the law is only approximately true; it is an empirical fact that entropy of a system can fluctuate. Second, it is typically the case that entropy increases (as opposed to just not decrease). Taking these qualifications into account, the Second Law says that a system is highly likely to exhibit *thermodynamic-like behaviour* (TD-like behaviour). We have TD-like behaviour *iff* the entropy of a system that is initially prepared in a low-entropy state increases

[9] For a discussion of infinite sequences see (Elga 2004).

[10] The assumption that macrostates can be indexed by an integer k is a common idealisation in this context and we follow this convention here; see (Frigg 2008b) and references therein.

until it comes close to its maximum value and then stays there, exhibiting frequent small and only rare large fluctuations away from equilibrium.[11] Notice the reference to 'highly likely' in the statement of the Second Law. In effect the law introduces a probability $p(TS)$ that a given system S behaves in a TD-like way, and it states that the value of this probability is close to one.

There is an entirely different way of looking at a system such as a gas in a box. As a matter of fact, the gas is a collection of molecules, each governed by the laws of mechanics. In what follows we assume these laws to be the ones of classical mechanics (CM). The microstate of a system consisting of n particles is specified by a point x in its $6n$-dimensional phase space Γ, which is endowed with the Lebesgue measure μ. The dynamics of the system is governed by Hamilton's equations of motion, which define a measure-preserving phase flow ϕ_t on Γ. More precisely, $\phi_t : \Gamma \rightarrow \Gamma$ is a one-to-one mapping for every real number t and $\mu(\phi_t(R)) = \mu(R)$ for every measurable region $R \subseteq \Gamma$. We assume that the relevant physical process begins at a particular instant t_0 and adopt the convention that $\phi_t(x)$ denotes the state of the system at time $t_0 + t$ if it was in state x at t_0, and likewise for $\phi_t(R)$; x is then commonly referred to as the 'initial condition'. The flow ϕ_t is deterministic and hence, together with the initial condition x of a gas completely determines the behaviour of the gas.[12]

Statistical Mechanics (SM) aims to establish a connection between the TD way and the mechanical way of looking at the system and to account for TD behaviour in terms of the dynamical laws governing the microscopic constituents of macroscopic systems and probabilistic assumptions. We now briefly review classical SM and discuss in some detail how probabilities are introduced into a mechanical theory.[13]

It is the basic posit of Boltzmannian SM (BSM) that the M_k *supervene* on the system's microstates. Therefore each macrostate M_k is associated with a macroregion $\Gamma_k \subseteq \Gamma$ so that the system is in macrostate M_k at t iff its microstate x lies within Γ_k at t. The Boltzmann entropy of a macrostate is defined as $S_B(M_k) = k_B \log(\mu(\Gamma_k))$, where k_B is the Boltzmann constant. Since the system is only exactly in one macrostate/macroregion at a given moment in time, we can define the Boltzmann entropy of the *system* simply as the entropy of the current macrostate: $S_B(t) = k_B \log(\mu(\Gamma_t))$, where Γ_t is the macroregion in which the system's microstate is located at time t. It can be shown that the equilibrium state is by far the largest of all states (relative to the measure μ), a fact known as the dominance of the equilibrium state. Since the logarithm is a monotonic function it follows that the entropy is maximal in equilibrium.

---

[11] This definition of TD-likeness is adapted from Lavis (2005). A different way of reformulating the Second Law emerges from (Albert 2000). We prefer an approach based on TD-likeness for the reasons outlined in (Frigg and Werndl 2011) and use it here because it is simpler than Albert's. Noting we say about chance depends on this choice, though, and *mutatis mutandis* our account of chance can also be applied to Albert's transition probabilities.

[12] We base our discussion on the standard possible worlds definition of determinism; see (Earman 1986, Ch. 2).

[13] We restrict attention to Boltzmannian SM. For detailed discussions of that theory, as well as of the Gibbsian approach which we set aside here, see Frigg (2008b) and Uffink (2006).

Deriving the (above statistical version of the) Second Law from the fundamental laws of mechanics and probabilistic assumptions is an important aim of SM. By assumption the systems starts in a low entropy non-equilibrium macro state. That this be the case is the subject matter of the so-called *Past Hypothesis*, and the system's low entropy macro state is called the *past state*.[14] Since the past state and the equilibrium state are of particular importance, we introduce special labels and denote the former by M_p (associated with Γ_p) and the latter by M_{eq} (associated with Γ_{eq}).

It is now time to introduce the concept of ergodic motion.[15] Metaphorically speaking, a trajectory is ergodic if it evenly spreads out on the entire accessible phase space and does not, for instance, get stuck in one corner. More precisely, the motion of a system is ergodic iff for any set $A \subseteq \Gamma$, the proportion of time the trajectory spends in A is equal to the proportion of the measure that A takes up in Γ in the long run (for instance, if A occupies a quarter of Γ, then the system spends a quarter of the time in A). If a trajectory is ergodic, then the system behaves TD-like because the dynamics will carry the system's state into Γ_{eq} and will keep it there most of the time (because Γ_{eq} is much larger than any other macro region). The system will move out of the equilibrium region every now and then and visit non-equilibrium states. Yet since these are small compared to Γ_{eq} it will only spend a small fraction of time there. Accordingly, the entropy is close to its maximum most of the time and fluctuates away from it only occasionally.

As a matter of fact, whether or not a system's motion is ergodic depends on the initial condition: some initial conditions lie on trajectories that are ergodic while others don't. This realisation is the clue to introducing probabilities. Consider an arbitrary measurable subset $C \subseteq \Gamma_p$. We may postulate that the probability that the initial condition x lies within C at time t_0 is

$$p(C) = \frac{\mu(C)}{\mu(\Gamma_p)}, \tag{2}$$

which is well-defined because it is a fact of physics that $\mu(\Gamma_p) > 0$. Let us refer to this principle as the Past Hypthesis Proportionality Postulate (PHPP). We now denote by $\Omega \subseteq \Gamma_p$ the subset of all initial conditions that lie on ergodic trajectories; $p(\Omega)$ is the probability that x lies within Ω at time t_0. We then immediately get:

$$p(TS) = p(\Omega) \tag{3}$$

We now see that $p(TS)$ has two aspects. From TD it inherits a solid connection to observable matters of fact, and via SM it is connected to the fundamental underlying mechanics. This will be important in the next section.

[14] Note that the term 'Past Hypothesis' is usually reserved for approaches in which the system under consideration is the entire universe; it then says that the universe came into being in a low entropy macrostate provided to us by modern Big Bang cosmology. We return to the issue of the nature of systems studied in SM Sect. 5.5.

[15] For an accessible introduction see (Berkovitz et al. 2011). We assume that the relevant systems are ergodic (Frigg and Werndl 2011).

4 Interpreting SM Probabilities as HOC's

It is tempting to interpret the above probabilities as epistemic probabilities. There is nothing chancy about a system's initial condition; there is a matter of fact about the system's microstate at t_0. One just doesn't know what this state is, and $p(TS)$ quantifies our ignorance about the system's true microstate. But this argument is unsatisfactory. The point is simple and has been made by many: these probabilities stem from basic laws of physics and the objective behaviour of natural systems, describe how things are, and are not rooted in what we fail to know (see, for instance, Redhead 1995).

We think that this intuition is correct and now argue that $p(TS)$ can be interpreted as a HOC. This interpretation faces an immediate challenge: how could there be chances in a world that obeys deterministic laws? Lewis (1986, 118–120) famously held that chance and determinism were incompatible, and many share his intuition. We do not. It is instructive, though, to see what line of reasoning leads to a rejection of deterministic chance.

Unfortunately Lewis does not make the reason for his verdict explicit. Hoefer (2007, 558–559) provides the following reconstruction of the incompatibilist's argument. Let H_t be the complete history of the world up to time t; let L be the conjunction of all laws governing our world; and let $ch_t(E) = x$ be the non-trivial chance of E at time t (i.e. $0 < x < 1$). Lewis regards historical facts and laws as admissible (1986, 92). PP then tells us that

$$cr(E|ch_t(E) = x \,\&\, H_t \,\&\, L) = x$$

However, given that L is deterministic, H_t and L jointly imply E (assuming E is true; they imply $\neg E$ if E is false). Hence the axioms of probability dictate that

$$cr(E|ch_t(E) = x \;\&\; H_t \;\&\; L) = 1.$$

So deterministic chances and PP as used here together with the axioms of probability, which credences must satisfy, lead to a contradiction. Lewis' way to avoid this contradiction is to deny that there are chances for events that are governed by deterministic laws: $ch_t(E) = x$ simply doesn't exist.

This response is closely tied to Lewis' metaphysics. The world consists of a manifold of spacetime points, which stand in various relations to one another. These spacetime points are simple particulars, meaning that they are individuals with no proper parts. Each spacetime point instantiates perfectly natural monadic properties, which stand in various spatiotemporal relations to one another. This ontology is based on classical physics and Lewis realised that it was so austere that it may not be able to accommodate modern physics. The point in upholding it nevertheless, he thought, was 'not to support reactionary physics, but rather to resist philosophical arguments that there are more things in heaven and earth than physics has dreamt of' (1994, 474). So while he would accept adjustments to the system that are needed to accommodate, say, quantum theory, these adjustments would have to be such that no properties other than the properties of basic physics would be allowed into it: '"how things are" is fully given by the fundamental, perfectly natural, properties and relations those things instantiate' (*ibid.*). The question of what counts as a perfectly natural property is answered by physics. So all that there is are facts about spacetime points and their spatiotemporal relations along with perfectly natural properties they instantiate—*all* other facts are determined by these basic facts. This is Lewis' physicalism, 'the thesis that physics—something not too different from present-day physics, though presumably somewhat improved—is a comprehensive theory of the world, complete as well as correct. The world is as physics says it is, and there's no more to say' (Lewis 1999, 33–34).

It follows from such a view that the predicates of sciences other than fundamental physics are not perfectly natural and that therefore there are no laws that are expressed in the language of, say, biology, geology, or meteorology. And Lewis, recall, advocates a Best System account of laws plus chances, not merely a Best System account of objective chances. It is then clear why he resolves the problem that deterministic chances do not satisfy PP by denying that there are chances for E in a deterministic world. If E is about fundamental properties, then, on pain of contradiction, there cannot be chance laws for it in a deterministic world (the contradiction being the one sketched just above). If, by contrast, E is about non-fundamental types (such as coins, macrostates, or genes) then a best system will not contain any laws for E-type events (and a fortiori no chance laws): a best system simply does not say anything about E-type events because fundamental physics is complete and comprehensive and everything that can be said about everything that there is can be said in terms of fundamental physics.

We find Lewis' approach overly restrictive. Even if there are deterministic fundamental laws, there is room in a Best System approach for chance rules about events and kinds at non-fundamental levels. In the remainder of this section we show how this is possible.

The starting point of our argument is to deny that everything that can be said about the world can be said in terms of fundamental physics. Probability rules can be formulated in terms pertaining to different levels of discourse such as macro physics, chemistry, genetics, mechanical engineering and meteorology, and probability rules formulated in such terms have equal right to be considered for inclusion in a Best System package of rules, alongside micro-level rules. We call this view *chance-rule pluralism* (CRP). We provide arguments in support of CRP in Sect. 5.2. In this section we argue for the conditional claim that *if* we accept CRP, *then* $p(TS)$, and with it scads of probability rules from non-fundamental physics, biology, and many other sciences, are plausibly part of a Best System.

The first step is to show that the above contradiction in terms of credences can be avoided. The key to a solution is to realise that what drives the contradiction is cross level information: we add information about fundamental physics to rules about coins, macro-states, and genes. But such information must be inadmissible: *chance rules operate at a specific level and evidence pertaining to more fundamental levels is inadmissible*. Far from being an ad hoc stipulation, this requirement is a natural consequence of the definition of admissibility given in Sect. 2. Recall our characterisation of admissibility: a proposition P is admissible with respect to an outcome-specifying proposition E for chance set-up S *iff* P contains only the sort of information whose impact on reasonable credence about E, if any, comes entirely by way of impact on credence about the chances of those outcomes. This means that the complete history H_t together with the deterministic laws L are jointly inadmissible, because they determine what will happen and hence give information maximally relevant for credence about E, and not by way of giving information about the chance of E.[16,17]

So there is no direct contradiction between deterministic laws and chance rules at higher levels; only an apparent contradiction that arises if one applies PP using Lewis' understanding of admissibility. If K contains no evidence pertaining to the micro level, then $cr(E|ch_t(E) = x \,\&\, K) = x$ by PP. If, by contrast, K does contain such evidence, plus the deterministic laws, then $cr(E|ch_t(E) = x \,\&\, K) = 1$ by the axioms of probability. But there is no contradiction because PP itself tells us that our credence ought to be equal to the chance only in the absence of inadmissible evidence, and it remains silent about what our credence ought to be when we have inadmissible knowledge. In this case the answer is obviously 1 because E is implied by K.

This does not render non-fundamental chances useless. In typical situations in which we observe coins or gases, we just don't have inadmissible information, and indeed given our epistemic finitude, couldn't have such information, nor calculate anything based on it if we did have it. So instead, setting our credences via $ch_t(E) = x$ and PP serves us well.

The removal of the apparent contradiction makes non-fundamental chance rules candidates for inclusion into the Best System (it now satisfies requirements 1 and 2 for HOC's). But does their candidacy ever result in membership? We argue that it does. To this end we divide Best System approaches into two classes: those that offer a Best System account of chance alone (leaving non-chance laws of nature, if any, off to the side), and those that offer a unified Best System account of laws plus chances. The former approach is advocated by Hoefer (2007), while the latter has been the default approach of other Best System theorists since Lewis' (1994).

Under the chance-only Best System approach, non-fundamental chance rules obviously make the cut. Since by assumption the fundamental physical laws are deterministic, there are no chancy fundamental laws at all, and hence nothing at the fundamental level that could compete for inclusion in the Best System of chance rules. So all the chance rules that may be considered for inclusion into the Best System will be rules covering events at "higher" levels.

Most philosophers who favour HOC, however, want to give a unified Best System account of laws and chances together, and on such an approach it may seem at first that there is a serious issue as to whether a Best System can contain fundamental deterministic laws and higher-level chance rules as well. The issue is not resolved just by embracing CRP; the question is, with deterministic laws in the System, how could adding any higher-level laws or rules at all make the system better?

16 Note that determining what will happen is not the same as entailing that the objective chance of something happening is equal to one (and *mutatis mutandis* for not happening and chance equal to zero).

17 Note that on our understanding of admissibility, it is not closed under logical conjunction, as Lewis supposed it to be.

To see how this can work we need to say a bit more about how we understand simplicity and strength. As indicated above, these concepts are notoriously difficult to define precisely and we won't try. However, it may help to notice that the very idea of a Humean Best System theory of laws involves a certain element of pragmatism and user-relativity, which makes the impreciseness of simplicity and strength (and their trade-offs) less disquieting.[18] A Best System is an elegant, simple, and powerful way of summarizing and systematizing the patterns of events in the Humean Mosaic. But omniscient beings have no need for a System at all, and Laplace's demon-like super-intelligences have no need of simplicity. A Best System is of use to finite and epistemically limited beings, such as ourselves. It is a "guide to life", just as objective chance is often said to be, for the epistemically handicapped. For a System to be *the laws of nature* is, for a genuine Humean, nothing more than for it to be the *best* such guide to life that our HM allows. And as such, the notions of simplicity and strength that are at issue in the BSA are linked to the powers and perspectives of epistemically handicapped beings such as ourselves. So we think it is legitimate to rely on our own intuitions as we try to say something about simplicity and strength, and also to not worry about whether the imprecision of these notions makes it impossible to be sure that *exactly* one Best System exists out there for our world's HM.[19]

We maintain that the notions of strength and simplicity have several dimensions to them. Strength is a matter of how many things in HM are covered by the laws and rules of a system. "How many things" can refer both to tokens of events and to the types under which the tokens are subsumed. So, *ceteris paribus* (literally: keeping all else equal), adding a rule to a system that applies to 10^9 distinct events in HM adds more strength than adding a rule that applies to only 100 events. And adding a rule that covers 5 different types of chance setups adds more strength, again *ceteris paribus*, than adding a rule that covers just one new kind of chance setup.

Of course adding rules to a system detracts, *ceteris paribus*, from its simplicity. The number of rules is one dimension of simplicity; we call it numerical simplicity. But recalling that Best Systems are meant to be "guides to life" for epistemically limited beings, we maintain that other types of simplicity should be taken into account. So, for example, a system that entails coin flip probabilities, but only via quantum-like micro-chance laws, so that only a Laplace's demon could actually calculate the chance of landing Heads, is in a sense much *less* simple than a system that is otherwise identical, but adds a simple rule stating $Pr(H) = Pr(T) = ½$. We call this derivation-simplicity. For a given natural language, Kolmogorov complexity gives one way to try to make the notion precise,[20] though again we

[18] While Lewis tried to suppress this element of pragmatism and user-relativity as much as he could without actually saying much at all about simplicity and strength, other BSA advocates such as Albert (2011), Callender and Cohen (2009) and Schrenk (2008) have openly embraced it.

[19] That said, we are sympathetic with Lewis' line on this point (1994, 479): we may, not unreasonably in light of actual science, hope that there is one *robustly* Best System for our HM, or a small family of closely resembling cousins, that come out as Best under any reasonable ways of cashing out and weighing up the qualities of simplicity, strength and fit.

[20] Roughly, the Kolmogorov complexity is the length of the shortest computer programme that derives a certain result. With respect to a given language, the Kolmogorov complexity is an objective quantity.

stress that we are not aiming to provide precise definitions. Finally, and relatedly, there is simplicity of formulation: assuming, again, a given natural language, then a law or chance rule that takes many pages to write down is less simple than a rule that can be compactly expressed on a single line.

Let us illustrate some of these points with an example. Consider a fictional world consisting of nothing but 10 coin tosses, 20 die throws and 50 spins of a roulette wheel, and assume that the outcomes of these events are nicely randomly distributed in the usual way so that the usual classical probabilities have good *fit* with the world's history. System 1 contains three separate rules covering coins, dice and roulette wheels, respectively; system 2 contains one rule covering all gambling devices; system 3 contains one rule covering only coins and dice and says nothing about roulette wheels. We then can see that system 2 has a better combination of simplicity and strength than system 1, which in turn beats system 3 because the latter, while having fewer rules, also covers fewer *types* of chancy events and many fewer tokens. There is a premium for covering cases that occur frequently (leaving out roulette wheels is penalised) and there is a reward for having fewer rules while keeping scope constant.

Based on these notions of simplicity and strength, we now argue that non-fundamental rules can be part of a Best System because a system that has these rules in them strikes a better balance between simplicity and strength than ones that don't (throughout we assume that the systems under considerations are on par as regards fit). The conclusion is obviously true for anti-reductionists: non-fundamental rules cover cases that aren't covered by fundamental rules. If instances covered aren't *extremely* rare, then strength increases significantly while simplicity decreases only a little, and hence overall the system with nonfundamental probability is better than the system that is restricted to fundamental rules (we return to the rarity issue below).

The more interesting case is the one in which reductionism in the sense of supervenience is true. To see how the argument plays out we first consider that case where we have inderministic laws at the fundamental level, for instance because the world is at bottom quantum mechanical (and quantum mechanics is understood in the standard way). As an example consider the probability rule discussed in (Lehe et al. 2012), specifying how the probability of a successful mutation increases as a function of distance from the bulk of the population in species that are expanding their territorial range.[21] Introducing such a rule into the system leaves that system's strength unchanged because, *per* supervenience, every case covered by a non-fundamental probability rule is also covered by a fundamental one. At the same time there is a loss in numerical simplicity because the number of rules in the system increases. So the strength versus-numerical simplicity balance would count against the introduction of non-fundamental rules. However, the system's *derivational*

[21] In a phenomenon known as 'gene surfing', genetic drift becomes a much stronger evolutionary force in populations at the edge of a territorial expansion wave, because genes from the individuals at/near the edge of the wave will be disproportionately represented in the gene pool of the newly colonized regions in subsequent generations. Lehe et al. (2012) propose (in our terms) chance rules for the fixation of a favorable mutation as a function of distance of the individual in which the mutation occurs from the edge of the colonization wave.

simplicity is greatly increased. It is hugely costly to start from first principles every time one wants to make a prediction about a non-fundamental object like a genetic mutation in a population, and a system becomes significantly simpler if we write in rules about such objects.[22] As already mentioned, quantifying the gain in derivational simplicity is notoriously difficult. But the fact that no one has yet succeeded in deducing a single biological probability from anything like first principles, let alone from fundamental quantum theory, suggests that the costs of such a derivation are indeed significant, and certainly sufficiently high to compensate for the relatively small loss in numerical simplicity due to the addition of one extra rule.

Let us now turn to the case in which the fundamental laws are deterministic. Not only can we never derive a probability rule from the micro-level laws. We also face the above problem that $ch_t(E) = x$ assigns a non-trivial chance to E while H_t & L imply E. Given this, how can $ch_t(E) = x$ (with $x < 1$) possibly be part of the Best System?

For reasons similar to the ones we considered above in the indeterministic case, such rules can earn their place in a Best System. Even though the fundamental theory and non-fundamental rules make different statements about HM, they cover the same setups (e.g. coins and gasses), which would count against their inclusion. However, while deterministic laws may, with the help of infinitely precise initial conditions, cover macro-level events, the derivational simplicity of a system can be vastly improved, *vis a vis* macro-level events, by the introduction of non-fundamental chance rules, with only a small cost in terms of the number of rules/laws in the system.[23]

The only remaining questions are: Does a proposed chance rule have good *fit* with the events in HM, and does the gain in derivational simplicity outweigh the loss of simplicity caused by adding an additional rule?

For statistical mechanical probabilities such as $p(TS)$, at least, we think the answers to both questions are affirmative, though the first one is mostly empirical and depends on facts about the HM which we can't claim to know with certainty. As we have seen in Sect. 3, $p(TS)$ has two aspects, one macro and one micro. The macro aspect is that it captures an observable pattern of events. The micro aspect is that it is connected to the fundamental underlying theory via its mechanical description and the probability distribution $p(C)$ (which, it is worth emphasising, is not itself part of the fundamental theory). In effect $p(C)$ expresses the probability that a system's microstate is in set $C \subseteq \Gamma_p$.

Now look at this probability from the point of view of THOC. There is a well circumscribed class of objects to which a chance rule like PHPP applies (gases, etc.). Each of these is a classical system and hence has a precise initial condition x at t_0, which, by assumption, lies within Γ_p. Now go through the entire HM and put every single initial condition x into Γ_p. The result of this is a swarm of points in Γ_p.

[22] For a discussion of this point see (Frisch 2011, forthcoming).

[23] This is so even if the output of the rule is merely a chance rather than a yes/no determination. When it comes to the chance $p(TS)$, which is almost always near-1, the information conveyed to the agent is nearly as strong as what is entailed by (but impossible to derive from) the deterministic laws plus IC's.

Then recall from above that THOC is essentially a refinement of finite frequentism and chances should closely track relative frequencies wherever such frequencies are available. Hence the chance of an initial condition being in set C given that it lies in Γ_p is the fraction of points in set Γ_p that lie in C (in the same way in which the chance of heads is, or is close to, the fraction of heads in the set of all coin toss outcomes). But listing all points individually and checking whether they lie in C is extremely cumbersome and won't make for simple system. So we have to reduce the complexity of the system by giving a simple summary of the distribution of points. To this end we approximate the swarm of points with a continuous distribution (which can be done using one of the well-known fitting techniques such as the method of the least mean squares or kernel dressing) and normalise it. The result of this is a probability density function ρ on Γ_p, which can be regarded as an expression of the 'initial condition density' in different subsets C of Γ_p.[24]

The good-fit constraint now is that $\rho(C)$ be equal to (or in very close agreement with) $\mu(C)/\mu(\Gamma_p)$ for all subsets C of Γ_p. This is a non-trivial constraint. For it to be true it has to be the case that the initial conditions are more or less evenly distributed over Γ_p because $\mu(C)/\mu(\Gamma_p)$ is a flat distribution over Γ_p. If it turned out that all initial conditions were crammed into one corner of Γ_p, then Eq. 2 would have poor fit, which would preclude it being part of the Best System, despite its great simplicity. So the requirement $\rho(C) \approx \mu(C)/\mu(\Gamma_p)$ presents a real touchstone for the system. Eq. 3 is then dealt with easily. This requirement has to hold for all C, *a fortiori* it has to hold for E and so Eq. 3 gives the right chances if $\rho(E) \approx \mu(E)/\mu(\Gamma_p)$.

The second question concerns the inclusion of statistical mechanical chances in the Best System is: is the gain in derivational simplicity/real-use strength big enough to compensate the slight loss in rule-count simplicity? We think it is obvious that it is, given how ubiquitous gases and other systems covered by statistical mechanics are, in our world. We would go further and argue that the same goes for chance rules covering things like successful mutations, symmetrical gambling devices, and so on. Here an objection can be raised, however, to the effect that the Earth is such a small corner of the HM that adding rules to cover any type of chance setup (gene mutations, coin flips) found only on Earth can only make a miniscule addition to a system's strength, an addition that therefore intuitively cannot counterbalance the loss of simplicity caused by the addition of an extra rule.

As noted before, it does not affect the question of whether statistical mechanical probabilities ought plausibly to be part of a Best System, since the systems covered are to be found all over the universe. Let us now nevertheless offer a few remarks to

[24] In this section our discussion idealises by pretending that the histories of all sorts of different SM systems could be treated as representable via paths in a single phase space. This is an idealisation because systems with a different particle number N have different phase spaces. We think that this is no threat to our approach. SM systems such as expanding gases and cooling solids are ubiquitous in HM and there will be enough of them for most N to ground a HBS supervenience claim. Those for which this is not the case (probably ones with very large N) can be treated along the lines of rare gambling devices such as dodecahedra: they will be seen as falling into the same class as more common systems and a flat distribution over possible initial conditions will be the best distribution in much the same way in which the $1/n$ rule is the best for all gambling devices.

try to deflate this objection as regards earth-only types. In the first place, as we have seen above, the strength of a system is measured not only in terms of the number of tokens covered; it also depends on the number of types covered, and the increase in this dimension may well be dramatic once we include biology. But even those who are not moved by this consideration may want to avoid the conclusion that the strength of a chance rule is measured by the proportion of space–time in which it is applicable. Indeed, such weighting arguably leads to disaster for the BSA, for a universe like ours. Astrophysicists assure us that we may well live in a universe which will have an infinitely long (in time) future "heat death" state in which, basically, nothing happens. So the simplest and strongest system for a universe like ours—ignoring the tiny corner of spacetime in which interesting things actually happen—would surely be something like "Nothing happens but minor quantum fluctuations in an otherwise cold, dead, slowly expanding space." And finally, of course, we can again invoke the fact that the Best System approach by its very nature should be a "guide to life" for epistemically limited agents—agents for whom, almost by definition, one corner of the HM is going to be more relevant and important than the vast regions with which they have no contact.

In this section we have argued that the standard probability rule of statistical mechanics can plausibly be seen as a rule that would form part of a Humean Best System for a Newtonian world, despite the availability of a fully deterministic underlying dynamics. The rule would merit inclusion for the dramatic simplification (derivational simplicity) it brings to the capture of a pervasive regularity in HM: that most systems behave in a TD-like fashion. In the last section we turn a defense of chance-rule pluralism, and to addressing the worries some philosophers may have about whether determinism and objective chance really are compatible after all.

5 Reverberations

5.1 A Kind of Counterfeit Chance?

Lewis famously quipped that introducing chances in a deterministic setting results in a 'kind of counterfeit chance', which is 'quite unlike genuine chance' (Lewis 1986, 120). And even writers more sympathetic to the cause of the determinist argue that the probabilities of SM and evolutionary theory should be considered a type of *objective probability* but not *objective chance*. Lyon (2011), for instance, argues that they are a third type of probability which coincides neither with chance nor with subjective probability.

There is no ultimate right and wrong in the use of a word, and depending on one's other philosophical commitments one *can* reserve 'chance' to propensities, primitive fundamentally chancy laws of nature, or some kind of fundamental modality (no matter how they are analysed). So our aim is not to prove authors embracing this decision wrong; our aim is to make it plausible that there is no compulsion to follow them in doing so and that there is a different and equally legitimate use of 'chance' that does not come with such restrictions. We see our own analysis as best because it both best satisfies the desideratum of rationalizing PP, and lets us capture all the key uses of chance in science—and lets us do so whether or not the world's fundamental physical laws are deterministic or indeterministic.

THOC is a philosophical analysis of objective chance. 'Objective chance' here is meant to refer to our everyday folk notion of chances as objective facts, "out there in the world" in some sense, which are representable mathematically as probabilities, and which are the kinds of things that gamblers, actuaries and many sorts of scientists are keen to find out about. Like any folk notion (like causation, reference, justice, for example), chance is not a sharply defined concept, and people will have divergent intuitions about which aspects or connotations associated with it are most central. This does not invalidate philosophical analysis, which aims to sharpen vague intuitions and mould them into consistent concepts. But it warns us that other analyses are always possible, and for the notion of chance they are currently proliferating in the literature. The aim in this section is to explicate how THOC differs from other Humean accounts of chance and make its distinctive features explicit. We see in particular THOC's closeness to scientific practice and its ability reconcile chance with determinism as virtues which render it superior to its competitors. But we acknowledge that those who cherish different aspects of chance will disagree on the relative value of the different accounts. The aim in this section is not to adjudicate between different outlooks. The aim is to make explicit wherein the differences lie and make it plausible that THOC is a legitimate and successful analysis of our folk notion of chance.

5.2 Why Accept Chance-Rule Pluralism?

Recall that CRP is the posit that probability rules can be formulated in terms pertaining to different levels of discourse and probability rules formulated in such terms have equal right to be *considered* for inclusion in a Best System package of rules, alongside rules at the fundamental level.[25] What reasons are there to adopt such a principle?

The answer to this question depends on one's stand on reduction. Let us begin with antireductionism. This position comes in different versions, some of which fall under the label of 'emergentism'. Antireductionists of all stripes will hold that the entities and properties at higher levels cannot be reduced (or at any rate not completely) to fundamental entities and properties at a fundamental level. In Anderson's famous words: more is different (1972). In this vein Dupré insists that 'there is a whole hierarchy of increasingly complex things that really exist, and that have causal powers that are not reducible to the mechanical combination of the powers of their constituents' (2006, 15). It is an obvious consequence of such a view that a best system will contain chance rules for various levels.

The more difficult question is why a reductionist would accept CRP. Like antireductionism, reductionism comes in different versions. The weakest version, and also most widely held, is the claim that the existence and properties of entities at

[25] In Sect. 4 we argued that *if allowed to compete*, some higher-level rules or laws may well deserve to make it into the Best System. Here the question is the prior one, which Lewis answered negatively: should such rules even be allowed to compete?

all levels supervenes on the facts about fundamental physical entities and their properties. Supervenience still has bite, because it entails that how non-fundamental entities behave is completely determined by how fundamental entities behave. It is important to notice, though, that the strong form of physicalism envisaged by Lewis—one holding that fundamental physics is a complete theory and that there is no more to say about the world than what physics says—is not a consequence of supervenience: one can consistently hold that all higher level entities/properties supervene on the entities/properties described by fundamental physics and that there theories other than fundamental physics are needed to provide a comprehensive description of what there is. So an argument over and above supervenience is needed to support that kind of physicalism.

At this point it is instructive to notice a parallel with a long-standing debate in the philosophy of mind. Eliminativist materialism submits that folk psychology should not only be reduced to neuroscience, but actually be *replaced* by it: whatever can be said about metal activities can said in neuroscientific terms. The reductionism versus eliminativism debate in the philosophy of mind exactly parallels the controversy around CRP, fundamental physics corresponding neuroscience and non-fundamental science to folk psychology. So Lewis' physicalism really is a form of eliminativism, and it is therefore illustrative to look at the arguments supporting eliminativism in the philosophy of mind.

Arguments in support of eliminativism typically are arguments against folk psychology. Proponents of eliminativism argue that folk psychology is profoundly wrong and predictively unsuccessful because central tenets of folk psychology radically misdescribe cognitive processes, that folk psychology is not a fertile research programme, and that ordinary mental states cannot in any way be reduced to, or be identified with, the brain states of neuroscience (see, for instance, Churchland 1981; Churchland 1986).

Typical arguments against elimintivism therefore take the form of a defence of folk psychology (see, for instance, Horgan and Woodward 1985). Carrying this argument in support of eliminativism over to the case at hand would amount to arguing that non-fundamental theories are profoundly wrong and predictively unsuccessful, and that they form stagnant research programmes. That this would be hopeless enterprise becomes clear as soon as one starts writing a list of theories that would have to be thus discarded: statistical mechanics, thermodynamics, chemistry, molecular genetics, etc. Needless to say, none of these theories is free of internal difficulties, but it is hard to see that these would justify a dismissal of, say, molecular genetics as on par with folk psychology even if one grants that arguments against folk psychology are successful (which is by no means uncontroversial). We regard this position as a non-starter.

And things get worse for the eliminativist. Not only is the main argument in support eliminativism unsuccessful in our context; there are reasons against it. The first is that the supervenience relation is one that holds between two (sets of) properties and this trivially implies that there are non-fundamental properties. That, say, temperature supervenes on average kinetic energy presupposes that a system has a temperature, and we can only express this supervenience relation if our scientific language contains a term referring to temperature. One might try to downplay the importance of non-fundamental properties by endorsing *ontological innocence*, the claim that supervenient properties really are nothing over and above the properties in the supervenience base and have therefore no ontological status.

But this move is of no help. Ontological innocence is controversial (van Inwagen 1994), and even if we assume, for the sake of argument, that the thesis stands, it will not help the eliminativist. To begin with, CRP is a posit about scientific theories, not ontology. It says that theories can be formulated at different levels and refer to properties at different levels; it is not committed to the claim that the properties at non-fundamental levels are in some sense sui generis. And what is more, deflating supervenient properties is successful only as long as there is a type–type identity. But many non-fundamental properties are multiply realisable, meaning that the same non-fundamental property can be realised by a number of distinct fundamental kinds. This is true not only in psychology (the classical example being *pain*); we find multiply realised properties even in close-to-fundamental physics: temperature in gases and temperature in spin systems have completely different micro-realisers. This is a problem for the eliminativist because the only thing these realisers have in common is that they are realisers of a certain non-fundamental property. There is no way to group these together without making reference to non-fundamental kinds, and hence not everything that can be said about these kinds can be said *only* in terms of the fundamental theory.[26]

The second argument against eliminativism is that supervenience claims by themselves do nothing more than assert that a certain pattern of property variation hold (you can't vary the supervening properties without also varying the base properties), but they do not say why this pattern holds and where the dependency comes from. This is widely seen as a problem, and it is natural to look for an explanation (Kim 1998). Whatever such an explanation will look like, it will have to make reference to the higher level properties and hence cannot be eliminativist.

For these reasons elimintivism as not a plausible position. Let us add a further reason for this conclusion, even though it is one that elimativists would dismiss as mere (and misguided) opinion. The point is that eliminativism is little more than a philosopher's dream. Real science has never borne any resemblance to it, and there is no evidence that it ever will. In our view a Best System should not be entirely disconnected from how science as practised by scientists looks like. Science has always operated at different levels; it is widely recognised that disciplines working at a certain level do so with a great deal of autonomy, and that both methods and explanatory practices are closely tied to discipline-specific domains of discourse. A Humean Best System approach should aim at capturing this feature of science and therefore ought to recognise (at least in principle) chances found in *all* the sciences. If this is so, then higher level entities form part of the Humean mosaic and certain statistical generalisations about them will be part of the overall best system. And we are glad to report that even outspoken radical reductionists seem agree. Nobel Prize winning physicist Steven Weinberg first observes that: 'no one doubts that with a

[26] Fodor famously used multiple realisability as an argument against reduction (1974). Following (Dizadji-Bahmani et al. 2010) we think that this is going too far, but multiple realisability does provide an argument against eliminativism.

large enough computer we could in principle explain all the properties of DNA by solving the equations of quantum mechanics for electrons and the nuclei of a few common elements, whose properties are explained in turn by the standard model' (Weinberg 1993, 24–25), but immediately adds that: '[t]here is no reason to suppose that the convergence of explanation [i.e. the solving fundamental equations to explain non-fundamental properties] must lead to a convergence of scientific methods. Thermodynamics and chaos and population biology will each continue to operate with its own language, under its own rules, whatever we learn about elementary particles' (*ibid*., 33).

5.3 Anything Goes?

An evergreen among the criticisms of HBS approaches is that simplicity and strength are psychological concepts and that this opens the flood gates for radical relativism: if we don't like the laws we find, we just change the way we think and get different laws. As noted earlier, Lewis responded to this charge by observing that '[i]f nature is kind, the best system will be *robustly* best—so far ahead of its rivals that it will come out first under any standards of simplicity and strength and balance' (1994, 479). He adds that there is of course no guarantee that nature is kind, but also no evidence to the contrary. And if it turned out that nature wasn't kind then we should blame the demise of a solid notion of lawhood on an unkind nature and not on the HBS analysis (*ibid*.).

We agree with Lewis on this point and set worries about radical relativism aside. In doing so we also assume that laws and chance rules are formulated in a natural language. If we allow artificial constructions, then indeed almost anything goes. We could, for instance, conjoin all chance rules in a system to form a super rule, and thereby bring n-simplicity down to a minimum. We know of no principled way to rule out such manoeuvres, but there is also no evidence that they are a serious problem for science. So we follow Lewis (1983) in assuming that laws and chance rules be formulated in some natural language. Lewis thought that a natural language was the language of 'something not too different from present-day physics, though presumably somewhat improved' (Lewis 1999, 33–34). We agree with that but would add that the same move be made for non-fundamental disciplines: a natural language for micro biology, for instance, would contain terms like 'gene' and be an improved version of something like current molecular biology.

A related criticism is that simplicity and strength are too vague to base a robust notion of law or chance on them. There is no denying that these concepts are bound to have a certain degree of vagueness to them. Does that make THOC (or related Humean accounts by Lewis et al.—for all invoke the same virtues) hopelessly unclear or ill-defined? We think not, obviously, and want to make two points in support of our optimism.

Firstly, assuming that artificial constructs like the above super rule are excluded by the natural language requirement, the vagueness pertaining to these concepts is greatly reduced. In particular, derivational simplicity is non-arbitrary if one commits to a particular language. As we have suggested above, it could, for instance, be measured in terms of Kolmogorov complexity, which is an objective measure if we assume a particular language as given. The details of such proposals need more work, but we believe the situation is not as hopeless as critics want us to believe.

Secondly, the prominence of the role of PP in THOC also provides a sort of guarantee that the imprecision of simplicity and strength can't do too much harm. If someone interprets the Bestness competition in such a way that the winning system is too restrictive, having little strength or covering only fundamental physical interactions (say), then their version can't be what THOC aims at, because THOC aims (clause (2)) to give us the right things to plug into PP. A too-restricted system would leave out potential probability rules that clearly do deserve to be plugged into PP, and do HBS-supervene on the HM, under a better understanding of simplicity, strength etc. Conversely, a too-unrestricted system would include a plethora of chance rules that cover, say, gerrymandered classes of objects and/or many chance rules with only one or a few instantiations in HM. For such a profligate system, the proofs of PP mentioned earlier (see Hoefer 2007; Frigg and Hoefer 2010) would fail, hence they could not be what THOC aims at. So we regard THOC as clearly—and by design—apt for capturing the core of our uses of objective chance, even if the notion of "Best" in "Best System" is not definable with any great precision.

5.4 Platitudes About Chance Reconsidered

Jonathan Schaffer (2007) formulated six platitudes about chance with the aim to convince the reader of the incompatibility of chance and determinism. We now revisit these platitudes from the point of view of THOC. Some of these alleged platitudes in fact aren't platitudes and hence need not worry the compatibilist. Other platitudes can be re-interpreted so that they hold true in THOC. So THOC can capture different but related platitudes to Schaffer's. This lends further support to our claim in Sect. 5.1 that THOC is a legitimate analysis of our folk notion of chance.

The first of Schaffer's platitudes is no *mere* platitude, being precisely PP itself. Needless to say, we agree with the spirit of this platitude but understand PP differently than Schaffer because we have a different notion of admissibility: Schaffer (along with Lewis) regards the total history and laws as admissible while (roughly speaking) we only regard evidence at the chance rule's level as admissible.

Similar remarks pertain to the case of Schaffer's second platitude, the link between chance and (future) possibilities. He starts with the *Basic Chance Principle* (BCP), which asserts that if at t there is a non-trivial objective chance of future event A occurring, then there is a possible world with the same history up to t, in which A has the same objective chance as it actually has, and in which A in fact occurs. Schaffer actually proposes a stronger version of this platitude, his *Realization Principle* (RP) stating that this other possible world in which A occurs should have the same laws as our world.

THOC denies that BCP/RP as formulated by Schaffer is a platitude, but endorses a closely related principle that differs from Schaffer's by restricting 'the same history up to t' to the *admissible* history—in the case of statistical mechanics this is the *macro*-history. Thus understood, there *can be* branching. Of course, the proponents of BCP/RP will reject this reading of 'same history'; they mean the full history in all detail. Fixing the precise initial condition of the world rules out all but one macro history provided the underlying dynamics is deterministic, and hence BCP is violated. In a view that sees chances as operating a certain level and decrees cross level evidence as inadmissible in conjunction with laws, such a choice is not forced upon us.

Schaffer's next principle seems linked to the notion that chance and chanciness have something to do with the "openness" of the future. His formulation of the Futurity Principle (FP) says that if an event has, at time *t*, chance strictly between zero and one, then that event (or its failure to occur) lie to the future of *t*. All past events that occurred have chance 1 (or no chance); all events that might have occurred in the past but didn't now have chance 0 (or no chance).

This platitude is independent of the issues surrounding determinism. THOC could be formulated so that it satisfy FP, and our reasons not do so have nothing to do with our views about determinism and chance. One awkward aspect of FP is that it requires that we make sense of an *objective* past/future division of all events (since we are, after all, discussing *objective* chance). In a relativistic world, this requirement is of course highly problematic. For this reason, among others, Hoefer (2007, 554) proposes that, for events, we make our platitude be "once chancy, always chancy". Your coin flip of yesterday was a chancy event, i.e., an event assigned a non-trivial objective chance by the HOC rules in our world. It still is such an event. Of course, your credence in the coin landing Heads is probably now either zero or 1—because you have inadmissible evidence about the outcome, of course. But the probabilistic nature of a coin flip is the same, whether in your "past" or your "future".

Next in line is the *Lawful Magnitude Principle* (LMP), which codifies the idea that 'chance values should fit with the values projected by the laws of nature. For instance, if the chance that the coin lands heads is 0.5, then the laws should codify that value (via history-to-chance conditionals)' (2007, p. 126). This principle codifies the assumption that chances should in the end be derivable from fundamental, probabilistic physical laws. However, we submit that LM is not (or at any rate should not be) a widely held platitude about chance. There are chances that are not connected to fundamental laws (at least not in an obvious way). The chance of a particular macrostate transitions in SM, the chance of malfunctioning of a turbine, the chance of cancer after being exposed to a certain amount of radiation, and the chance of an increase in global mean temperature of more than two degrees centigrade by 2050 are not connected to fundamental laws and yet, according to THOC, are objective chances.

The *Causal Transition Constraint* (CTC) specifies that chances should be pushing events around directly, as they unfold in time. CTC stipulates that if an objective chance "plays a role in the causal transition from *d* to *e*", then *t* must be temporally between t_d and t_e. This seems intended to capture the propensity theorist's conception of chance propensities as "partial causes" helping to bring about events as time unfolds. As with LMP, we would deny that this constraint has the status of a platitude, except to a person who already thinks that "chance" means "law-governed primitive propensity", and furthermore thinks of propensities along the lines of causal powers. It's status as platitude is also problematic because our best fundamental science makes no mention of this kind of causal transitions at all.[27]

Lastly, there is the *Intrinsicness Requirement* (IR) which stipulates that "[i]f e' is an intrinsic duplicate of e, and the mereological sum of the events at t' is an intrinsic duplicate of the mereological sum of the events at t, then $ch < p_e, w, t > = ch < p_{e'}, w, t'>$"[28] (2007, p. 125). A more intuitive way to read IR is as follows: if you have the same (physically identical) set-up conditions at two different times, then the chances of all their corresponding possible outcomes must be the same. So interpreted, THOC clearly satisfies IR when it comes to BSM chances. The platitude "same set-up, same probabilities" will only be violated by HOC's if HM is such that the Best System must contain some time- or space-variable chance rules, i.e. rules where the rule explicitly states that the chances vary depending *only* on time or place in the universe. There is no reason to suppose that anything like this is true in our world, and certainly the phenomena falling under BSM show no such non-universality. IR does capture a fairly widely held platitude about objective chance, without begging any questions concerning determinism; and THOC has no trouble complying with it.

To sum up: THOC satisfies suitably qualified versions of PP, LMP, IR and BCP/RP; and we deny that FP and CT are platitudes about chance.

5.5 Reinventing the Wheel?

Loewer (2001, 2004) has presented a Humean reconciliation of determinism and chance, and so we want to end by pointing out wherein the differences lie. To this end let us now come back to an assumption we made rather silently but which plays in important role in our discussion, namely that the systems under investigation are 'laboratory systems' like gases in containers, liquids in tanks, and solids on tables. This is what SM is mostly applied to in the hands of working physicists, and so there is nothing objectionable about this restriction. However, following Albert (2000), Loewer considerably extends the domain of application of the theory, and treats the entire universe as one large SM system. Much of what we said about HOC's is independent of what one takes the relevant systems to be. There is one essential difference, though, namely how we understand the Humean supervenience of probability rules. As we have seen above, we interpret the even distributions over the past state (which occur in Eqs. 2 and 3) as an elegant summary of *actual* initial conditions as they occur in the HM of a world like ours. Such an interpretation is not open to those who take BSM to be a theory about the universe as a whole, since there is only exactly one initial condition of the universe.

Our take on the distributions has two advantages over Loewer's approach. First, postulating a probability distribution for an event that not only happens only once in the entire history of the universe, but is in fact the only probabilistic event ever,

[27] Lyon (2011) and Glynn (2010, 25–26) argue convincingly that CTC is unacceptable in any case no matter what view of chance one adopts.

[28] In Schaffer's notation, $ch < p_{e'}, w, t >$ is the chance in world w, assessed at time t, that the proposition p_e (asserting that event e happens) is true.

seems conceptually problematic even if one takes frequency tolerance seriously. There is no demonstration of how the PHPP supervenes on the classical particle motions in HM; if that rule, applied to numerous small subsystems of the world, does happen to make predictions with good fit to the patterns in HM, that is a happy accident, but not something that obviously follows from the global probability rule postulated in Loewer's approach.

Second, one can show that the fit of a system can be improved by choosing a peaked rather than a flat (Lebesgue) distribution over Γ_p (Frigg 2008a, 2010). A peaked distribution, nearly dirac-delta style, over the world's actual initial condition is, *qua* postulate, just as simple as a flat distribution, but it will assign significantly higher probabilities to the actual world's macro-state transitions, and hence give the system to which it belongs a much higher *fit*. So the best system is not one with the flat distribution but one which has a distribution that is peaked over the actual initial condition. This has the undesirable conclusion that the probabilities defined in Eqs. 2 and 3 are not chances. This objection is based on the fact that the distribution over the past state is not dictated by facts about initial conditions (there is only one!) and hence there is a great degree of freedom in choosing the distribution. This is not so if one understands the distribution to be a summary of (many!) actual initial conditions of systems like gases all over the Humean mosaic. If the initial conditions of these systems are roughly evenly distributed, then one cannot choose a peaked distribution. So that problem is effectively solved in an approach focussing on small isolated systems rather than the universe as a whole.

6 Conclusion

We have provided a reformulation of Lewis' HBS approach to chance. Thus reformulated, HOC's are compatible with underlying deterministic laws and provide a viable interpretation of SM probabilities. A key element in the reconciliation of determinism and chance is the realisation that chances are specific to a certain level. This is a consequence of CRP, which we defend by drawing parallels to the situation in the philosophy of mind.

Acknowledgments This paper was presented at the IHPST workshop "Probability in Biology and Physics" in Paris, February 2009. We would like to thank the organisers for the opportunity and the audience for stimulating comments. Furthermore, We would like to thank Nancy Cartwright, José Díez, Jossi Berkovitz, Mathias Frisch, Barry Loewer, Alan Hájek, Aidan Lyon, Kristina Musholt, Huw Price, Josefa Toribio, and Eric Winsberg for helpful discussions. Thanks are also due to two anonymous referees for helpful comments. RF acknowledges financial support from Grant FFI2012-37354 of the Spanish Ministry of Science and Innovation (MICINN). CH acknowledges the generous support of Spanish MICINN grants FFI2008-06418-C03-03 and FFI2011-29834-C03-03, AGAUR grant SGR2009-01528, and MICINN Consolider-Ingenio grant CSD2009-00056.

References

Albert, D. (2000). *Time and chance*. Cambridge/MA and London: Harvard University Press.

Albert, D. (2011). Physics and chance. In Y. Ben-Menahem & M. Hemmo (Eds.), *Probability in physics* (pp. 17–40). Berlin: Springer.

Anderson, P. W. (1972). More is different. *Science, New Series, 177*, 393–396.
Berkovitz, J., Frigg, R., & Kronz, F. (2011). The ergodic hierarchy. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Summer 2011 Edition: http://plato.stanford.edu/archives/sum2011/entries/ergodic-hierarchy/.
Callender, C., & Cohen, J. (2009). A better best system account of lawhood. *Philosophical Studies, 145*, 1–34.
Churchland, P. M. (1981). Eliminative materialism and the propositional attitudes. *Journal of Philosophy, 78*, 67–90.
Churchland, P. (1986). *Neurophilosophy: Toward a unified science of the mind/brain*. Cambridge/MA: MIT Press.
Dizadji-Bahmani, F., Frigg, R., & Hartmann, S. (2010). Who's afraid of Nagelian reduction? *Erkenntnis, 73*, 393–412.
Dupré, J. (2006). *The constituents of life*. http://www.umb.no/statisk/causci/SpinozalecturesDupre.pdf.
Earman, J. (1986). *A primer on determinsim*. Dordrecht: Reidel.
Elga, A. (2004). Infinitesimal chances and the laws of nature. *Australasian Journal of Philosophy, 82*, 67–76.
Fodor, J. (1974). Special sciences and the disunity of science as a working hypothesis. *Synthese, 28*, 77–115.
Frigg, R. (2008a). Chance in Boltzmannian statistical mechanics. *Philosophy of Science, 75*, 670–681.
Frigg, R. (2008b). A field guide to recent work on the foundations of statistical mechanics. In D. Rickles (Ed.), *The Ashgate companion to contemporary philosophy of physics* (pp. 99–196). London: Ashgate.
Frigg, R. (2010). Probability in Boltzmannian statistical mechanics. In G. Ernst & A. Hüttemann (Eds.), *Time, chance and reduction. Philosophical aspects of statistical mechanics*. Cambridge: Cambridge University Press.
Frigg, R., & Hoefer, C. (2010). Determinism and chance from a Humean perspective. In D. Dieks, W. Gonzalez, S. Hartmann, M. Weber, F. Stadler & T. Uebel (Eds.), *The present situation in the philosophy of science* (pp. 351–372). Berlin, New York: Springer.
Frigg, R., & Werndl, C. (2011). *Explaining thermodynamic-like behaviour In terms of epsilon-ergodicity*. Philosophy of Science (Forthcoming).
Frisch, M. (2011). From Boltzmann to Arbuthnot: Higher-level laws and the best system. *Philosophy of Science, 78*, 1001–1011.
Frisch, M. (forthcoming). Physical fundamentalism in a Lewisian best system. In A. Wilson (Ed.), *Asymmetries of chance and time*. Oxford: Oxford University Press.
Glynn, L. (2010). Deterministic chance. *British Journal for the Philosophy of Science, 61*, 51–80.
Hájek, A. (2007). The reference class problem is your problem too. *Synthese, 156*, 563–585.
Hoefer, C. (2007). The third way on objctive probability: A sceptic's guide to objective chance. *Mind, 116*, 549–596.
Hoefer, C. (2014). Consistency and admissibility. In T. Handfield & A. Wilson (Ed.), *Chance and temporal asymmetry*. Oxford: Oxford University Press (Forthcoming).
Horgan, T., & Woodward, J. (1985). Folk psychology is here to stay. *Philosophical Review, 94*, 197–226.
Kim, J. (1998). The mind-body problem after fifty years. In A. O'Hear (Ed.), *Current issues in philosophy of mind* (pp. 3–21). Cambridge: Cambridge University Press.
Lavis, D. (2005). Boltzmann and Gibbs: An attempted reconciliation. *Studies in History and Philosophy of Modern Physics, 36*, 245–273.
Lehe, R., Hallatschek, O., & Peliti, L. (2012). The rate of beneficial mutations surfing on the wave of a range expansion. *PLoS Computatinal Biology, 8*, e1002447.
Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. 2, pp. 83–132). Berkeley: University of California Press (reprinted in Lewis 1986, with postscripts added).
Lewis, D. (1983). New work for a theory of universals. *Australasian Journal of Philosophy, 61*, 343–377.
Lewis, D. (1986). *Philosophical papers*. Oxford: Oxford University Press.
Lewis, D. (1994). Humean supervenience debugged. *Mind, 103*, 473–490.
Lewis, D. (1999). *Papers in metaphysics and epistemology*. Cambridge: Cambridge University Press.
Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Modern Physics, 32*, 609–629.
Loewer, B. (2004). David Lewis' Humean theory of objective chance. *Philosophy of Science, 71*, 1115–1125.

Lyon, A. (2011). Deterministic probability: Neither chance nor credence. *Synthese, 182*, 413–432.
Redhead, M. (1995). *From physics to metaphysics*. Cambridge: Cambridge University Press.
Schaffer, J. (2007). Deterministic chance? *British Journal for the Philosophy of Science, 58*, 113–140.
Schrenk, M. (2008). A theory for special sciences laws. In H. Bohse, K. Dreimann & S. Walter (Eds.), *Selected papers contributed to the sections of GAP.6, 6th international congress of the society for analytical philosophy*. Paderborn: Mentis.
Uffink, J. (2006). Compendium of the foundations of classical statistical physics. In J. Butterfield & J. Earman (Eds.), Philosophy of physics (pp. 923–1047). Amsterdam: North Holland.
van Inwagen, P. (1994). Composition as identity. *Philosophical Perspectives, 8*, 207–220.
Weinberg, S. (1993). *Dreams of a final theory: The search for the fundamental laws of nature*. New York: Vintage.