

# Memory, the Fork Asymmetry, and the Initial State

## **Abstract**

Why do we have records of the past and not the future? Entropic explanations for this ‘record asymmetry’ have been popular ever since Boltzmann. Foremost amongst these is Albert and Loewer’s account, which explains the record asymmetry using a low-entropy initial macrostate (the ‘Past Hypothesis’) plus an initial probability distribution. However, the details of how this initial state underpins the record asymmetry are not fully specified. In this paper I attempt to plug this explanatory gap in two steps. First, I suggest the record asymmetry is more immediately explained by the ‘fork asymmetry’, which their picture omits. Second, by relating the fork asymmetry to an initial state that’s metaphysically similar to theirs, I clarify how this ultimately underpins the record asymmetry.

# 1 Introduction

From the movements of Napoleon’s armies to yesterday’s weather, we have access to a wealth of facts about the past. Our forecasts of the future offer no real match. One natural explanation for this is that we have *records* of the past but not the future. If this is right, then we can expect this ‘record asymmetry’ to be grounded in a yet more fundamental time-asymmetry (as the adage goes, ‘no asymmetry in, no asymmetry out’).

Nowadays, entropy-based explanations predominate, and foremost amongst these is Albert and Loewer’s theory. This has faced a number of objections.<sup>1</sup> However, there exists an alternative and somewhat forgotten line of explanation that puts the ‘fork asymmetry’ at centre stage. Although early sketches appeared with Lewis (1979) and Horwich (1988), this tradition has nowadays fallen out of favour due largely to objections from Arntzenius (1990). I shall develop and defend an account in this vein, and relate it to the universe’s initial state. My initial state is metaphysically similar to Albert and Loewer’s, but has a tighter link to the record asymmetry via the fork asymmetry, which their picture omits. Hence, my proposal plugs a gap in theirs.

This paper is structured as follows. In section 2 I outline Albert and Loewer’s account in the context of Loschmidt’s reversibility objection. In section 3 I introduce the basic idea of how the fork asymmetry explains the record asymmetry, and refine it. In section 4 I explain how this can capture the informativeness of *individual* records. In section 5 I flag up an observation made by Arntzenius about deterministic systems, from which two objections to this fork asymmetry explanation follow, and address the first of these. In section 6 I present the second objection, which relates to Horwich’s account in particular - the most salient of its sort. In section 7 I offer a revised account that avoids this objection. In section 8 I derive the fork asymmetry from a

---

<sup>1</sup>See Winsberg (2004), Earman (2006), Parker (2005), and Frisch (2005a, 2010).

‘smooth’ initial distribution, and relate my theory to Albert and Loewer’s. In section 9 I answer some possible objections. I conclude in section 10.

## 2 The Entropic Paradigm

In this paper I shall develop an explanation for the record asymmetry based on the fork asymmetry, and defend it from various objections. I shall understand the record asymmetry simply in terms of our having records of the past but not the future, and not in terms of any associated subjective phenomena such as feelings of passage or directionality.<sup>2</sup> At its core, my fork-asymmetry-based proposal follows an alternative tradition to the entropic one, which has long dominated the literature. Nevertheless, I believe it fills an explanatory gap in today’s most salient entropic account, namely Albert and Loewer’s theory.<sup>3</sup> A quick dip into the history of statistical mechanics will help us understand their account, and how my own proposal might supplement it.

A major motive for statistical mechanics was to provide a microphysical foundation for the laws of thermodynamics.<sup>4</sup> Amongst these is the Second Law, which in its statistical mechanical form states that the entropy of a closed system is overwhelmingly likely to increase toward the future. This is usually explained as follows: since macrostates with higher entropy are more probable, a system is likely to keep evolving into higher-entropy macrostates. This culminates in the equilibrium macrostate, whose entropy has a maximum value, where it is likely to remain for vast timescales.

However, it was soon apparent that there was a tension between the time-symmetric underlying dynamics on the one hand, and the time-*asymmetric*

---

<sup>2</sup>Hemmo and Shenker (2020) tackle such issues.

<sup>3</sup>Other noteworthy entropic accounts include Reichenbach (1956), Grünbaum (1963), Smart (1967), Hawking (1985, 1994), Wolpert (1992), Hartle (2004), Schulman (2005), Mlodinow and Brun (2014), and Smith (2014).

<sup>4</sup>See Uffink (2007) and Brown *et al.* (2009) for the history of this relationship.

Second Law on the other. This was made stark by Loschmidt's objection, which runs as follows. Just as a system is likely to evolve into a series of higher-entropy macrostates toward the *future*, it is also - by the same reasoning - likely to evolve into such a series toward the *past*. When we apply this logic to the universe as a whole, this suggests all macroscopic objects probably just fluctuated into existence. Crucially, this extends to all our *records*, which are therefore likely to be completely spurious. In Albert's words (2000, 116), we face a "full-blown skeptical catastrophe".

This objection is usually blocked with the 'Past Hypothesis', the posit that the universe began in a very low-entropy macrostate.<sup>5</sup> This rules out a Loschmidt-style fluctuation scenario by fiat, taking the 'skeptical catastrophe' off the table. But it's one thing to say that the Past Hypothesis blocks a particular scenario in which all our records are wildly misleading, and quite another to say that it explains the record asymmetry. Albert (2000, Ch. 6) and Loewer (2011, 2012) believe that the Past Hypothesis, supplemented with a uniform initial distribution (the 'Statistical Postulate'), *does* in fact manage this. I outline their reasoning below in skeleton form.

In order for a record at  $t_3$  to reliably register an event at  $t_2$ , it must have started out in its proper 'ready state' at  $t_1$ , or else it would be spurious. For example, if I observe a rolling 8-ball on a frictionless billiard table at  $t_3$ , this reliably records a collision at  $t_2$  only if it was stationary at  $t_1$ , or else it would have been rolling all along. So, why do I assume the proper ready state obtained at  $t_1$ ? The answer is by falling back on yet another record, such as my memory. But this itself requires another, *even earlier* ready state: I must assume my brain started out healthy, untampered with by an evil scientist, and so on. So, why do I assume my brain's ready state obtained? An infinite regress looms: each new record I employ as evidence that some ready state

---

<sup>5</sup>This idea is long-running; see Feynman (1965, 116) and Penrose (1990, Ch. 7).

obtained will *itself* demand its own, even earlier ready state.

This regress eventually leads us to the early universe, and which point it is terminated by what Albert and Loewer call the ‘mother of all ready states’. Since this ‘mother state’ is *the* earliest state, we can’t infer it using records, for it lacks prior ready states to enable this. So instead, we use abductive inference. If the mother state did *not* obtain, then all our records would lose their validating bedrock, undermining our empirical grasp of the world. But if it *did* obtain, then records are veridical in the way we normally imagine. Hence, we believe in the mother state as an inference to the best explanation for our experience: it occupies a central role in our understanding of things. Crucially, Albert and Loewer identify the mother state as the low-entropy macrostate associated with the Past Hypothesis, i.e. the ‘Past State’.<sup>6</sup>

For all its ingenuity, this argument is hand-waving in a key respect: it doesn’t tell us *why* the Past State occupies its role of the mother state. In other words, even if we grant that the Past State successfully terminates their regress, it’s unclear whether it does so because of its low entropy, “certain further symmetry conditions” (Loewer, 2007), or whatever else. This highlights an explanatory gap in their picture: the details of *how* their initial state (Past Hypothesis plus Statistical Postulate) ultimately underpins the record asymmetry are not fully specified.

In this paper I endorse a metaphysically similar initial state (Past Hypothesis plus ‘Smoothness Postulate’), and plug this explanatory gap using the fork asymmetry. This lets us pinpoint which features of the initial state do which bits of explanatory work with regards to the record asymmetry. Since the Past Hypothesis blocks Loschmidt’s objection and facilitates the existence of macroscopic objects, it provides the foundation, but doesn’t fully

---

<sup>6</sup>I borrow this term from Callender (2010).

explain the record asymmetry. However, the character of the initial distribution (its ‘smoothness’) finishes the job, since this is what underpins the fork asymmetry.

Now that we know where we’re headed, I wish to begin by showing why the fork asymmetry is fertile ground for explaining the record asymmetry.

### 3 The (many-pronged) Fork Asymmetry

In this section I begin by introducing the fork asymmetry, and explaining why it’s sometimes thought to imply a record asymmetry. I then refine this model to improve its explanatory capacity.

The fork asymmetry refers to the fact that whenever two observable events  $A$  and  $B$  are correlated, there is often some earlier observable event  $C$  whose occurrence (and non-occurrence) renders  $A$  and  $B$  statistically independent, whereas there is never a later observable  $C$  that plays the same role.<sup>7</sup> This generation of statistical independence between  $A$  and  $B$  on the part of  $C$  (and not- $C$ ) is called ‘screening off’.

More formally, the fork asymmetry amounts to this. Whenever we find:

$$\Pr(A \wedge B) > \Pr(A) \times \Pr(B) \tag{1}$$

$$\Pr(A | C) > \Pr(A | \neg C) \tag{2}$$

$$\Pr(B | C) > \Pr(B | \neg C) \tag{3}$$

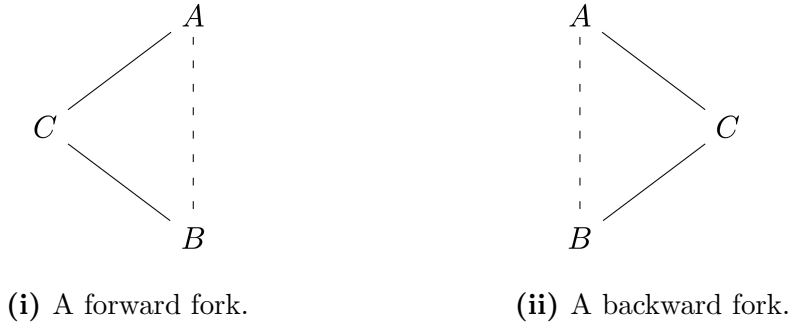
$$\Pr(A \wedge B | C) = \Pr(A | C) \times \Pr(B | C) \tag{4}$$

$$\Pr(A \wedge B | \neg C) = \Pr(A | \neg C) \times \Pr(B | \neg C) \tag{5}$$

---

<sup>7</sup>This was first described by Reichenbach (1956) as the ‘Common Cause Principle’.

then  $C$  is always in the past of  $A$  and  $B$ , and never in their future. If we think of such cases as ‘forks’ with  $C$  as the fork-point and  $A$  and  $B$  as the fork-tips, then the fork asymmetry refers to the fact that the world contains many ‘forward forks’ but not ‘backward forks’ (see Figure 1).



**Figure 1:** Solid lines represent unscreened-off correlations, dashed lines represent screened-off correlations, and the time axis runs from left to right. Whereas (i) is common, we never seem to find (ii).

For example, suppose  $A$  refers to ‘cancer’,  $B$  to ‘yellow fingers’, and  $C$  to ‘smoking’, where  $C$  of course lies in the past of  $A$  and  $B$ . This would satisfy equations (1)-(5), for even though smoking raises the probability of having cancer, and also of having yellow fingers, having cancer and having yellow fingers are nonetheless statistically independent *given* that someone smokes (and also *given* that they do not). Hence, (1)-(5) are satisfied, and we have a forward fork.

But now suppose  $A$  refers to ‘asbestos exposure’,  $B$  to ‘smoking’, and  $C$  to ‘cancer’, so that  $C$  now lies in the *future* of  $A$  and  $B$ . Equations (2) and (3) are satisfied, since someone’s having cancer raises the chance of asbestos exposure, and also of smoking. But the probabilities of asbestos exposure and smoking are *not* statistically independent amongst cancer patients, for the occurrence of one would *reduce* the probability of the other. This violation

of equation (4) is enough to show that we don't have a backward fork here.

Starting with Reichenbach (1956), many have thought that an account of causation can be recovered from the fork asymmetry, or more recently from the general structure of probabilistic asymmetries via key assumptions (e.g. the Causal Markov Condition) that underpin causal modelling using Bayesian networks.<sup>8</sup> The possibility of such a reduction is the topic of much debate. But so as to remain neutral on this issue, I shall frame my account purely in terms of correlations, leaving causal talk aside.

Many have recognised that a forward fork can be interpreted as a single event leaving behind records,<sup>9</sup> since (2) and (3) imply the following:

$$\Pr(C | A) > \Pr(C | \neg A) \tag{6}$$

$$\Pr(C | B) > \Pr(C | \neg B) \tag{7}$$

Since  $A$  and  $B$  each raise  $C$ 's probability, they resemble records of that event. This is apparent in our original example: when  $A$  refers to 'cancer',  $B$  to 'yellow fingers', and  $C$  to 'smoking', it seems plausible that  $A$  and  $B$  serve as records of  $C$ . This sums up the very basic idea of why the fork asymmetry is sometimes thought to imply a record asymmetry.

As it stands, this model has a serious shortcoming. To see this, let's start by asking the following question: to what extent do  $A$  and  $B$  raise  $C$ 's probability? Lewis' 'overdetermination arrow' (1979, 49-50) led him to suppose that  $A$  and  $B$  are each sufficient, given the laws of nature, to guarantee  $C$ . But this was surely far too strong, for we want to allow that records are sometimes spurious. Indeed, it often happens that any *given* record of  $C$  raises its probability only modestly:

---

<sup>8</sup>See Hitchcock (2009) for an overview of these developments.

<sup>9</sup>See for instance Papineau (1985) and Horwich (1988).

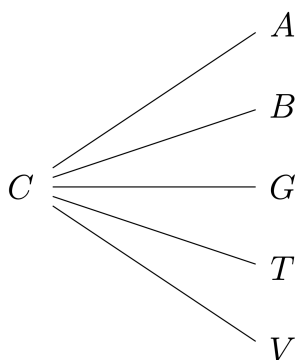


$$\Pr(C \mid A) \ll 1 \tag{8}$$

$$\Pr(C \mid B) \ll 1 \tag{9}$$

As a matter of fact, (8) and (9) apply in our own example. Since cancer can result from asbestos, poor diet, or genetics, and since yellow fingers can result from Raynaud’s disease, carotenemia, or jaundice, neither cancer nor yellow fingers is a very reliable record of smoking. And yet, we *can* sometimes tell with near-certainty that someone was a smoker even when we lack a highly reliable record like a photo. So here is the puzzle: how can we make sense of the fact that  $C$  can be highly probable even when any *given* record is only a mild probability-raiser of that event?

We can answer this by recognising that events often leave behind many more records than just two. In our example, smoking may well lead to not only ‘cancer’ ( $A$ ) and ‘yellow fingers’ ( $B$ ), but also ‘gum disease’ ( $G$ ), ‘stained teeth’ ( $T$ ), and ‘varicose veins’ ( $V$ ). Since *any* of these records can stand in for  $A$  and  $B$  in (1)-(5), it turns out that forward forks can have many more fork-tips than just two (see Figure 2).



**Figure 2:** A many-pronged forward fork (dashed lines ignored).

This ‘many-pronged’ forward fork is a more realistic model of how multiple records form following an event. But more to the point, it allows us to resolve our earlier issue. Because  $C$ ’s records are *correlated*, we often encounter a whole *cluster* of them, so we’re often in a position to conditionalise on *many*. The probability-raising of  $C$  by any given record might only be modest, but when we combine all their piecemeal contributions,  $C$  may become extremely likely. For instance, when we conditionalise not only on cancer ( $A$ ) and yellow fingers ( $B$ ), but also on gum disease ( $G$ ), stained teeth ( $T$ ), and varicose veins ( $V$ ), the evidence of smoking becomes overwhelming:

$$\Pr(C \mid A \wedge B \wedge G \wedge T \wedge V) \approx 1 \tag{10}$$

It’s worth briefly confirming that the inferences which drop out of this picture are indeed time-asymmetric. Just as each of  $C$ ’s records raises the probability of  $C$  itself, likewise  $C$  raises the probability of each of these records. This is captured in equations (2) and (3). However, this does not make  $C$  hugely informative about the future, for (in general) no particular record is immensely likely given  $C$ . For example, given that someone smokes, neither cancer nor yellow fingers (nor gum disease, etc.) is anywhere near guaranteed; each is just more probable. The record asymmetry stems from the fact that certain events are reliably triangulated by *multiple* probability-raisers all bearing on a *single* event, and whereas  $C$  is a lone probability-raiser of its later records, *all* these records are probability-raisers of  $C$  itself.

Once we envision forward forks as many-pronged, we can get a lot more mileage out of the fork asymmetry than is usually supposed. Specifically, the *numerousness* of  $C$ ’s records allows us to explain how records provide us with information about the past that’s relatively a) reliable, b) detailed, c) far-reaching, and d) easily accessible as compared with our predictions about the future.

First, let's consider the issue of *reliability*. If the number of putative records of  $C$  is greater, then the likelihood of  $C$  having occurred will also be greater, for each putative record raises its probability - even if only slightly. This is because the more putative records of  $C$  that we have, the less likely it is that *all* of them are spurious. If for instance we observe that someone has cancer, yellow fingers, gum disease, stained teeth, and varicose veins, and then infer that they were a smoker on the basis of *all* this evidence, our inference is far more reliable than if we were to base it on cancer alone.

Second, let's think about why many records can convey a *detailed* picture. Following  $C$ , the matter of which records end up forming hinges on which 'background events' obtain in their vicinity. For example, if a smoker had a certain mutation on chromosome 5, then this will affect the probability of cancer forming. Likewise, if they regularly handled bleach, then this will affect the probability of them developing yellow fingers. Since these background events are not mutually exclusive, each of  $C$ 's records reveals a unique aspect of what occurred in the vicinity. For example, if a smoker has cancer, this implies they lacked a protective mutation on chromosome 5. Likewise, if I see that they have yellow fingers, this implies they didn't regularly handle bleach. By aggregating these snippets of information, we can assemble a more detailed picture of the past.

Third, why do numerous records imply *far-reaching* information? Given that background events are not mutually exclusive, the formation of many records following  $C$  easily translates to the formation of many *types* of records; their multiplicity facilitates their diversity. But if  $C$  leaves behind many types of records, then this helps safeguard against their wholesale extinction by destructive events in the future. This is because different sorts of calamities tend to wipe out different sorts of records. For instance, a computer virus might destroy records of smoking in the form of medical records but not yellow fingers, whilst a chainsaw accident will do the reverse. If there is more

opportunity for *some* records of a given event to persist for long periods of time, then there is more opportunity for the present to contain records of things that happened *longer ago*, i.e. more opportunity for records to be far-reaching.

Fourth, we can account for the matter of *accessibility* as follows. Other things behind equal, the more records an event leaves behind, the more likely we are to discover at least some of them. As a result, it is more likely that we will do so spontaneously, and not have to search high and low. This is again clear in our example: if records of smoking exist in the form of physical ailments, medical bills, and smoky furniture, it's going to be all the more likely that we'll run into at least some of them without much effort. Hence, the records will be more easily accessible.

## 4 A Closer Look at Records

In this section I shall present and address two apparent shortfalls in the above explanation for the record asymmetry: the first concerns our epistemic access to the fork asymmetry, whilst the second concerns the effectiveness of individual records.<sup>10</sup>

The first objection is this. In order for the fork asymmetry to entail a record asymmetry, it isn't enough that the former merely exists. On top of this, we need some way of actually *discovering* that the fork asymmetry's relata (*A*, *B*, *C*, etc.) have the probabilistic relationships captured in equations (1)-(5). Without this caveat, all that would follow is that fork-tips occur more frequently than their fork-points; it wouldn't follow that we can recog-

---

<sup>10</sup>These doubts appear in Horwich (1988, 81-82). As we shall see, Horwich himself employs the fork asymmetry in his own account of the record asymmetry. But his two misgivings lead him to propose additional mechanisms underlying records (Ibid., 84-90), which I avoid the need for.

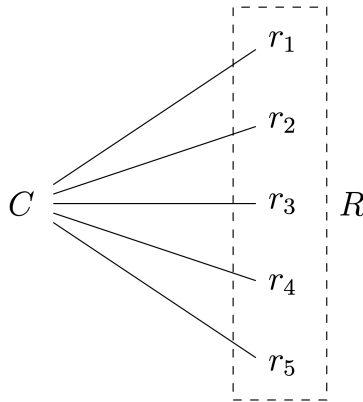
nise these events as constituting forward forks. In a nutshell, what allows the probabilities that characterise the fork asymmetry to become apparent?

In order to establish these probabilities, the events in question must certainly be observable in the first place. But since I've characterised the fork asymmetry in terms of observable events right from the start, we can be assured of this as a built-in property, where their observability amounts to the fact that they are *macrostates* (rather than microstates). Equally importantly, the fork asymmetry's relata are generally *not* one-off events, for we have observed *many* instances of (say) people smoking, having cancer, and having yellow fingers. Their probabilistic relationships may then be established empirically by observing their frequencies. For example, we can establish that  $A$  and  $B$  are correlated, i.e. equation (1), by observing that more people have yellow fingers as a proportion of cancer patients than as a proportion of all people, and also that more people have cancer as a proportion of yellow fingered people than as a proportion of all people. We can infer equations (2)-(5) through similar means.

The second objection runs as follows. So far, we've been characterising fork-tips as *discrete records*. Whilst this might explain the *collective* effectiveness of records, this analysis by its nature cannot explain their *individual* effectiveness. But this is one of the most striking aspects of records: a single photo, for instance, can provide us with a) reliable, b) detailed, c) far-reaching, and d) accessible information about (say) Tank Man's protest on Tiananmen Square in 1989. Unless we can account for this, our picture will be incomplete to say the least.

Fortunately, we can answer this by interpreting the fork asymmetry in a novel way: as operating *within* individual records. A highly reliable record  $R$  generally consists of many sub-components  $r_i$ : a photo consists of many ink blobs, a footprint consists of many sand grains, and a fingerprint consists

of many sweat particles. If we envisage the recorded event  $C$  as a fork-point and the various  $r_i$  as fork-tips, then their statistical relationships imply a miniature forward fork (Figure 3). This is because the  $r_i$  are correlated,  $C$  raises their probabilities, and  $C$ 's occurrence (and non-occurrence) screens them off from each other, satisfying (1)-(5). Each  $r_i$  therefore raises  $C$ 's probability. The probability-raising by any single  $r_i$  might only be modest, but since they are numerous, they can still render  $C$  highly probable when operating in concert.



**Figure 3:** A many-pronged fork whose fork-tips  $r_i$  are the sub-components of  $R$ , a highly effective record of  $C$ .

I shall now argue that this interpretation of the fork asymmetry allows us to explain the effectiveness of individual records. Specifically, it allows us to explain why a single item can afford us a) reliable, b) detailed, c) far-reaching, and d) accessible information about the past. My arguments are analogous to those in the previous section, but applied at a smaller scale. Just like before, the explanatory work in each case is done by the *numerousness* of fork-tips for a given  $C$ .

First, let's consider how individual records can be so *reliable*. Suppose we have a record  $R$  consisting of many sub-components  $r_i$ . Although any given  $r_i$  might be spurious, it is unlikely that all of them are. For example, consider our photo of Tank Man. Any given ink blob could have a spurious origin: a dust particle in the camera, a minor printing malfunction, or some debris embedded in the paper. But because there are thousands of ink blobs in play, it's highly unlikely for all of them to be spurious, so the photo is very reliable.

Second, the *detailed* character of certain records can be explained straightforwardly as follows. If a record consists of many  $r_i$ , then it can capture  $C$  in high resolution. This is clear in our example: a vast number of ink blobs allows me to infer not just that Tank Man stood in front of a column of tanks, but also that he was holding two bags, that the tanks had red stars painted on them, and so on. Relatedly, because these ink blobs are crammed onto a relatively small object, i.e. a small piece of paper, the record is *information-dense*.

Third, we can explain why some records are very *far-reaching*. If a record consists of many  $r_i$ , then it may continue to operate properly even when many of these are destroyed, for enough will still remain to tell the tale. Again, our example illustrates this. Over the years, a photo may deteriorate due to sunlight, stains, creasing, and so on. But so long as this doesn't involve anything too drastic like being caught in a house fire, the photo won't be appreciably worse at telling us what it always did: we can still infer Tank Man himself, his bags, the tanks with their red stars, and so on. Since  $R$  started out with countless  $r_i$ , plenty will remain even if many are lost, making it stable over time and hence far-reaching.

Fourth, this picture lets us explain why the information encoded in certain records is easily *accessible*. If a record consists of many  $r_i$ , then the fact that

records are localised means it's scarcely possible for someone to observe only one or two sub-components. Instead, one tends to observe a vast number *as a bundle*. In our example, it's hard to imagine a realistic scenario in which someone observes just one or two ink blobs in the photo without seeing the rest. It is far more likely for them to observe thousands all at once, and hence see the image for what it is.

In summary, we can think about the many-pronged fork asymmetry in two different ways. On a more orthodox interpretation, its fork-tips are distinct records, allowing us to explain the collective power of records: multiple items afford us a) reliable, b) detailed, c) far-reaching, and d) accessible information about the past. But by interpreting the fork-tips as sub-components of a single record, we can explain their individual power: single items can account for a)-d) all on their own.

## 5 Determinants: The First Objection

So far, I have argued that the fork asymmetry is a better explanation for the record asymmetry than is usually thought. I shall now present an observation made by Arntzenius (1990) about deterministic systems, from which two objections to this fork asymmetry explanation follow. In this section I shall present and respond to the first of these objections, outlining the second in the following section.

Arntzenius' central observation runs as follows. In a deterministic system, every event has a 'determinant' at every other time - a set of physical circumstances whose occurrence necessitates the event, and whose non-occurrence forbids the event. If  $C$  is our event of interest and  $D$  is an associated determinant, the following relations therefore obtain:



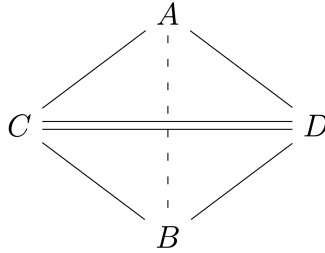
$$\Pr(D | C) = 1 \tag{11}$$

$$\Pr(D | \neg C) = 0 \tag{12}$$

$$\Pr(C | D) = 1 \tag{13}$$

$$\Pr(C | \neg D) = 0 \tag{14}$$

From this observation follows a simple first corollary. Consider some  $D$  that lies in the future of not only  $C$ , but also of its associated  $A$  and  $B$ . Equations (11) and (12) imply  $A$  and  $B$  have exactly the same probabilistic relationships to  $D$  as they do to  $C$ . But since  $C$  is a screener-off of  $A$  and  $B$ , so must  $D$  be. This implies that every forward fork shares its fork-tips with a backward fork, annulling the fork asymmetry altogether (see Figure 4).



**Figure 4:** The ‘fork symmetry’. The double line represents the determinant relationship, which is symmetric; equations (13) and (14) mirror (11) and (12). Just as  $A$  and  $B$  form a forward fork with  $C$ , they form a backward fork with  $D$ .

Papineau (1992) and Frisch (2005b) have responded by arguing that  $D$  is unobservable, and since we only care about the fork asymmetry insofar as it characterises observable events, it poses no threat. They argue as follows. Since  $C$  is observable, it is a macrostate. Following Arntzenius (1990, 82), its future determinant is represented in phase space by time-evolving  $C$ ’s microstates forward under the dynamics. Liouville’s Theorem, a consequence of

determinism, states that because trajectories can neither merge nor branch, the Lebesgue measure of any given set of phase points remains constant, so it behaves like an incompressible fluid. However, its shape can change dramatically, and it is generally assumed that it will *fibrillate* across phase space.<sup>11</sup> Since it would be nothing short of miraculous for  $D$  to perfectly coincide with the borders of one or more macrostates,  $D$  is likely not a macrostate, and hence unobservable.<sup>12</sup>

Before moving on, it's worth analysing macrostates one step further, as this shapes how we ought to envisage the fork asymmetry. So far, I've been taking the familiar macrostates that we all observe for granted. But if a macrostate is a set of microstates that take on the same value for some macrovariable (this being a *partial description* of any given microstate),<sup>13</sup> then the fact that one can concoct infinite possible macrovariables implies that there are infinite possible ways of partitioning phase space, each yielding a *different* set of macrostates. So, is there anything special about the macrostates *we* happen to observe? As Hemmo and Shenker (2016) point out, their salience is not a theorem of mechanics, but a by-product of our own physical makeup, as this dictates which macrovariables our perceptual states (e.g. brain states) correlate with. Whether or not a fork asymmetry presents itself is therefore observer-dependent. It exists for humans and other entities that also seem to experience a record asymmetry (animals, computers, etc.), but there could in principle exist other entities that are sensitive to *different* macrostates, and

---

<sup>11</sup>To be clear, fibrillation is an unproven dynamical assumption that doesn't drop out of fundamental physics. See Berkovitz *et al.* (2006) for discussion.

<sup>12</sup>A commentator has claimed that this fibrillation doesn't preclude the later event from being observable, as we can sometimes steer  $C$  into a later bona fide macrostate (or disjunction thereof) in a reliable way. For example, consider bringing the macrostate 'cream poured into coffee' into the equilibrium macrostate 'creamy coffee'. Nevertheless, this later macrostate is *not* synonymous with  $D$ . Rather,  $D$  is a highly fibrillated region with *largely overlaps* with the equilibrium macrostate in phase space;  $D$  is itself is unobservable. We will address the topic of highly reliable records in section 9, but the key point for now is that the fork asymmetry (qua observable phenomenon) still stands.

<sup>13</sup>See Hemmo and Shenker (2015a).

therefore *don't* experience a fork asymmetry - or perhaps even experience a reversed one.

## 6 Determinants: The Second Objection

Having established the fork asymmetry's existence (at least for humans), I shall now discuss Arntzenius' second objection to its role in an explanation for the record asymmetry. In order to do so, we will need to back up a bit, as this objection relates to Horwich's (1988) account of *why* there is a fork asymmetry at all.

As with other time-asymmetric phenomena, we can expect the fork asymmetry to be explained by a more fundamental asymmetry. Like Albert, Loewer, and many others, Horwich looks to the universe's initial state for an explanation. However, one can accept this logic without appealing to the character of the initial *macrostate* (e.g. its low entropy), for perhaps some *other* aspect of the early universe explains the record asymmetry. Horwich's account, outlined below, proceeds on this basis.

Horwich's (1988, 73-74) explanation for the fork asymmetry is essentially an explanation for why a correlation between  $A$  and  $B$  requires a  $C$  in its past and not in its future, where  $C$  is an event satisfying (1)-(5). To this end, he employs two key ideas: an explicit postulate, and a hidden assumption. Let me present these ideas in turn, and then explain how they fit together.

The explicit postulate is very straightforward, and amounts to the following boundary condition:

**Initial Micro-Chaos:** The universe's initial state contained no correlations.

Meanwhile, the hidden assumption amounts to a certain conjecture about what's required in order for correlations to vanish. The idea is that whenever there exists a correlation between  $A$  and  $B$  at some moment in time (say,  $t_3$ ), a correlation remains in existence under forward or backward time-evolution *until and unless* an event playing the role of  $C$  occurs. In this respect,  $C$  is necessary and sufficient to *annihilate* correlations. I shall express this assumption as follows:

**Annihilation Assumption:** If  $A$  and  $B$  are correlated at  $t_3$ , a correlation remains under forward or backward time-evolution until and unless  $C$  occurs, whereupon the correlation is *annihilated*.

Note that  $C$ 's role as an 'annihilator' in this sense doesn't simply follow from (1)-(5), for those equations are silent about whether or not a correlation remains when we time-evolve from  $t_3$  to *beyond* the time of  $C$ 's occurrence.

To be clear, Horwich doesn't explicitly flag up the Annihilation Assumption. But by seeing how it combines with Initial Micro-Chaos to explain the fork asymmetry, the need for such an assumption will hopefully become apparent. To understand his explanation, it is useful to take  $t_3$  as our focal point, and consider the implications of time-evolving backward and forward in turn.

At  $t_3$ , there exists a correlation between  $A$  and  $B$ . When we time-evolve backward, a correlation remains for some period. But there must eventually come a time  $t_2$  when the correlation vanishes, for Initial Micro-Chaos tells us the universe's initial state (at  $t_1$ ) contained *no* correlations. This correlation-annihilation event is therefore necessary in order to satisfy the boundary condition. But since the Annihilation Assumption says  $C$  is necessary and sufficient to annihilate correlations,  $A$  and  $B$  must have a  $C$  in their past. Hence, a correlation between  $A$  and  $B$  implies a forward fork.

But what happens when we time-evolve forward from  $t_3$ ? Again, since  $A$  and  $B$  are correlated at  $t_3$ , a correlation remains for some period. But because there is no analogous ‘Future Micro-Chaos’ boundary condition, it’s not being stipulated that the universe’s final state ( $t_5$ ) is correlation-free, and so there’s no particular need for the correlation to vanish in this temporal direction (say, at  $t_4$ ). A correlation-annihilation event *could* occur at  $t_4$ , but it isn’t *necessary* like in the previous scenario. This means  $A$  and  $B$  needn’t have a  $C$  in their future. Hence, a correlation between  $A$  and  $B$  doesn’t imply a backward fork.

In summary, given a correlation between  $A$  and  $B$  at  $t_3$ , the Annihilation Assumption implies we require a  $C$  at  $t_2$  to satisfy Initial Micro-Chaos at  $t_1$ , whereas we don’t require a  $C$  at  $t_4$  since there’s no analogous Future Micro-Chaos at  $t_5$ . This means correlations implicate forward forks but not backward forks, giving rise to a fork asymmetry.

So far, so good. This account, however, is undermined by a second corollary of Arntzenius’ determinants, which runs as follows. If every event has a determinant at every other time, then correlated events  $A$  and  $B$  must each have a determinant in the initial state of the universe. Let us call these determinants  $A^*$  and  $B^*$  respectively. Furthermore, since  $A$  and  $A^*$  are related by equations (11)-(14), and since the same is true of  $B$  and  $B^*$ , the fact that  $A$  and  $B$  are correlated means that  $A^*$  and  $B^*$  must also be correlated. But this result flatly contradicts Initial Micro-Chaos. Why? Because in harbouring the correlated events  $A^*$  and  $B^*$ , the initial state can’t have been correlation-free after all.

Moreover, there can be no such mechanism as the Annihilation Assumption, for the fact that correlated events have correlated determinants at every other time means  $C$  cannot destroy them - indeed, *nothing* can. Hence, Arntzenius (1990, 82) writes: “Correlations are not born and do not die,

they merely change variables.’ In view of this it is simply false to claim that all initial properties are uncorrelated.” Horwich’s account is therefore false on two counts: it postulates an incorrect initial state (Initial Micro-Chaos), and an impossible role for  $C$  (given by the Annihilation Assumption).

## 7 Correlation Scrambling

In this section I shall offer a revised account. My picture is crudely analogous to Horwich’s, but it is weak enough to avoid Arntzenius’ second corollary yet strong enough to bear out a fork asymmetry.

Section 5’s take-home message was that since we are only interested in the fork asymmetry insofar as it might explain the record asymmetry, we are right to restrict our attention to observable events. With this in mind, Horwich’s impossible proposal that  $C$  destroys correlations is far stronger than what we need. All  $C$  really has to do is transform correlations between observable events into correlations between *unobservable* events. Upon this transformation, the correlation’s relata would transform from macrostates into *non*-macrostates, i.e. *fibrillated regions* in phase space. In other words, we might say that the correlation has been ‘scrambled’ at this point: it has transformed from an observable correlation, i.e. a correlation between macrostates, into an unobservable correlation, i.e. a correlation between fibrillated regions in phase space.

To reflect this weaker demand on  $C$ , I propose we replace the Annihilation Assumption with following weaker assumption:

**Scrambler Assumption:** If  $A$  and  $B$  are correlated at  $t_3$ , an *observable* correlation remains under forward or backward time-evolution until and unless  $C$  occurs, whereupon the correlation is *scrambled*.

Again, note that  $C$ 's role as a 'scrambler' in this sense doesn't follow from (1)-(5), for those equations say nothing about whether or not an *observable* correlation remains when we time-evolve from  $t_3$  to beyond the time of  $C$ 's occurrence.

Having suggested a weaker role for  $C$ , let's try and figure out what we must say about the early universe in order to explain the fork asymmetry. Arntzenius showed that we're stuck with the existence of initial correlations. However, all we really need to banish is a certain *sort* of initial correlation: the sort that produces observable correlated events later down the line which materialise not only at around the same time, but also in around the same location, and *without* a common observable precursor. This would involve (say) cancer and yellow fingers frequently popping up in roughly the same spatiotemporal locations, i.e. in the same humans, and without smoking (or something else observable) to precede them. If the initial state contained many of these 'latent observable correlations', then observable correlations would often appear out of the blue without the need for a prior  $C$ , undermining the fork asymmetry. So, here is what we need to say: the initial state contained very *few* of these latent observable correlations.

To accommodate this requirement, I suggest we replace Initial Micro-Chaos with the following weaker conjecture:

**Initial Low Latency:** The universe's initial state contained very few latent observable correlations.

To assure ourselves that the Scrambler Assumption and Initial Low Latency yield a fork asymmetry, let's again take  $t_3$  (a time when  $A$  and  $B$  exist) as our focal point and consider the implications of time-evolving backward and forward in turn.

When we time-evolve backward from  $t_3$ , an *observable* correlation remains for some period. Let me be more precise: this correlation could involve the original  $A$  and  $B$ , or it could involve their physical precursors. For example, if  $A$  is ‘cancer’ and  $B$  is ‘yellow fingers’, then the respective precursors might be ‘DNA damage’ and ‘nicotine absorption via skin’. But whatever this correlation involves, it is observable. When we keep time-evolving backward however, there comes a time  $t_2$  when the correlation becomes *unobservable* - a process for which  $C$  is necessary and sufficient due to the Scrambling Assumption. Why does this correlation-scrambling occur as we time-evolve backward? Because if it doesn’t occur, then we’d be led to a latent observable correlation in the initial state ( $t_1$ ), which Initial Low Latency virtually forbids. As Arntzenius demonstrated, a correlation of *some* sort must exist prior to  $C$ . But the relata won’t be macrostates, and hence won’t be observable: they will amount to past determinants of  $A$  and  $B$ , i.e. fibrillated regions in phase space. So here’s the upshot: since a correlated  $A$  and  $B$  generally have a  $C$  in their past, this implies a forward fork.

Let’s now consider the implications of time-evolving forward from  $t_3$ . Once again, since  $A$  and  $B$  are correlated, an observable correlation remains for some period. But since there is no analogous ‘Future Low Latency’ conjecture, the universe’s final state ( $t_5$ ) may well harbour latent observable correlations, i.e. the germs of observable correlations that would materialise from thin air if we were to time-evolve *backward* from  $t_5$ . This means there’s no need for the observable correlation at  $t_3$  to get scrambled as we time-evolve forward (say, at  $t_4$ ) and thus become unobservable. In principle, they may remain observable indefinitely. Since a correlated  $A$  and  $B$  don’t necessarily have a  $C$  in their future, this doesn’t imply a backward fork.

In summary, given an observable correlation between  $A$  and  $B$  at  $t_3$ , the Scrambling Assumption implies we require a  $C$  at  $t_2$  to satisfy Initial Low Latency at  $t_1$ , whereas we don’t require a  $C$  at  $t_4$  since there’s no analogous



Future Low Latency at  $t_5$ . This means observable correlations implicate forward forks but not backward forks, yielding a fork asymmetry.

A brief comparison with Horwich’s account is in order. We are only in a position to verify the existence or non-existence of observable correlations. The Annihilation Assumption transcends our empirical evidence by making claims about *all* correlations - specifically, by claiming they are annihilated. By contrast, the Scrambler Assumption stays within our empirical remit by only making claims about observable correlations. This prevents conflict with Arntzenius’ second corollary, for I allow that correlations of some sort exist in the initial state: all I’ve avoided is a certain *sort* of initial correlation, i.e. the latent observable sort. But since we’re only interested in the fork asymmetry qua observable phenomenon, limiting our claims to observable events does no harm.

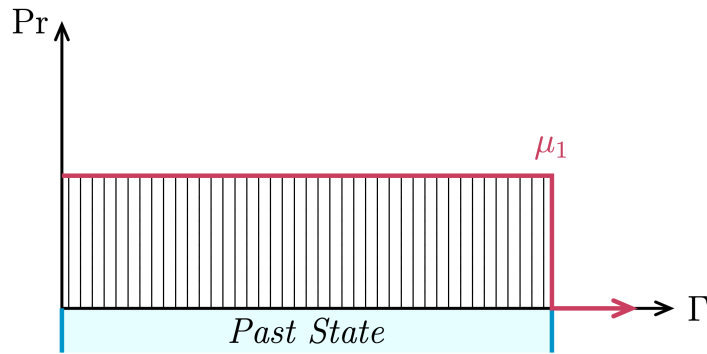
## 8 Characterising the Initial State

In this section I characterise the initial state in more conventional terms, i.e. as the Past Hypothesis plus an initial distribution, and delineate the explanatory role of each component. I proceed in four steps: first, I argue that we need a ‘smooth’ initial distribution; second, I respond to an objection by Hemmo and Shenker; third, I relate my picture to Albert and Loewer’s; and fourth, I compare my account to others in the literature.

What sort of initial distribution do I require? My account rests on Initial Low Latency, which says latent observable correlations were scarce. However, we have good reason to doubt their feasibility, for they produce observable events later down the line (say, ‘cancer’ and ‘yellow fingers’) that materialise simultaneously (of all possible temporal separations), *and* in close proximity (of all possible spatial separations), all in the absence of a common fork-

point. It seems reasonable to suppose that initial microstates rife with latent observable correlations are highly atypical. Therefore, I require a distribution that ascribes these atypical microstates a low probability.

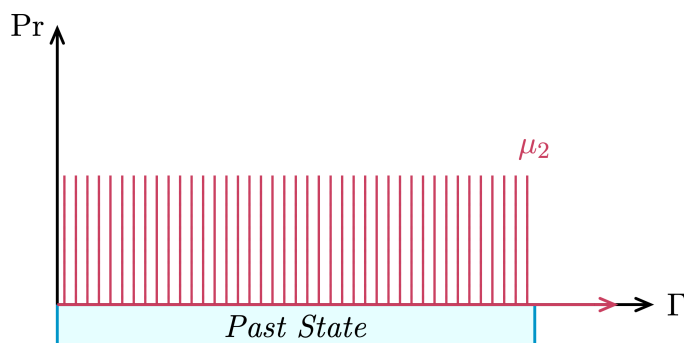
The standard way of giving atypical microstates this treatment is via a uniform initial distribution  $\mu_1$  (with respect to the Lebesgue measure), as proposed by Albert and Loewer’s Statistical Postulate.<sup>14</sup> In thermodynamic contexts the atypical microstates are Second-Law-violators, whereas in this context they’re fork-asymmetry-violators, but the principle is the same: atypical microstates have a small Lebesgue measure, so  $\mu_1$  ascribes them a low probability (see Figure 5).



**Figure 5:** A representation of  $\mu_1$  over phase space  $\Gamma$ . Thin black stripes correspond to atypical microstates along the x-axis, whilst thick white stripes correspond to typical microstates. Clearly, the latter have the lion’s share of likelihood.

<sup>14</sup>An alternative tradition, going back to Lebowitz (1993), interprets the Lebesgue measure simply as a *typicality* measure. The motive here is to avoid various pitfalls surrounding probability. Whether or not this improves on the more standard, probability-based approach is a controversial question; see Frigg (2009), Goldstein (2012), and Hemmo and Shenker (2015b) for discussion. However, I’m sticking with probabilities for two reasons. First, it demystifies the origins of the macroscopic probabilities associated with the fork asymmetry: their source is the initial distribution. And second, it relates more clearly to Albert and Loewer’s own probability-based account.

But in fact, we can get away with a weaker posit. Poincaré and others have argued that in certain deterministic scenarios like coin tosses or spinning roulette wheels, only very atypical initial microstates would yield macro-probabilities that deviate from what's familiar. Therefore, in order to explain the frequencies we observe, all we need is an initial distribution that *isn't* sharply peaked over these atypical microstates.<sup>15</sup> But this describes our own situation, for what *we* want to explain is the fact that we *don't* observe the fruits of latent observable correlations (which would undermine the fork asymmetry). Since these are associated with atypical initial microstates, we just need an initial distribution that *didn't* look like this:

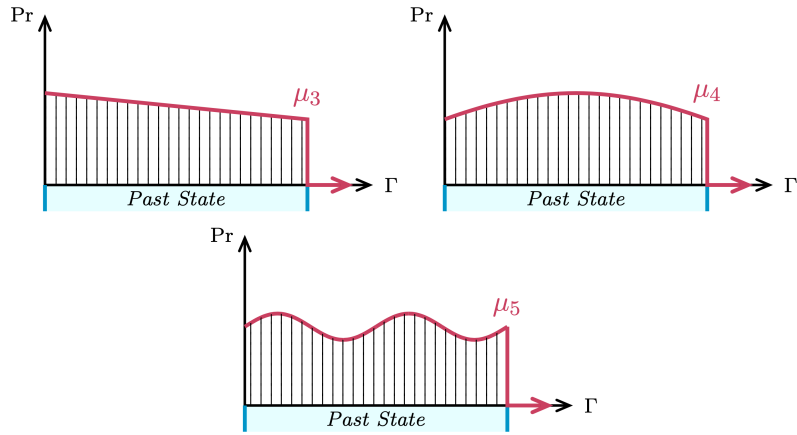


**Figure 6:** The finely-tuned distribution  $\mu_2$  delicately tracks the vertical black lines of Figure 5.

For this purpose,  $\mu_1$  is just one member of a whole family of suitable distributions, others of which are shown below.

---

<sup>15</sup>This approach is known as the ‘method of arbitrary functions’. See von Plato (1983) for some background.



**Figure 7:** Some suitable distributions besides  $\mu_1$ . Each of them ascribes my atypical microstates a very low probability.

How can we gloss these distributions? Following Strevens (2016), I shall call them ‘smooth’ on the grounds that their values don’t vary tremendously over small intervals in phase space (unlike  $\mu_2$ ).<sup>16</sup> I therefore propose that the fork asymmetry is underpinned by the ‘Smoothness Postulate’, with states that the initial distribution over the Past State was smooth. The Smoothness Postulate is weaker than the Statistical Postulate, but just as adequate for my purposes. I therefore prefer it over the Statistical Postulate - not because smoothness has distinctive explanatory features over and above uniformity, but simply because it’s a less committal conjecture.<sup>17</sup>

So far, I’ve outlined a broad similarity between my account and Albert and Loewer’s: we’re in general agreement that an initial distribution is needed to explain the record asymmetry. However, Hemmo and Shenker (2012, Ch. 8; 2016) have made the following objection against any explanation that

<sup>16</sup>This might tally with Wallace’s (2012, 349) notion of a ‘simple’ initial state, though he works in the Gibbsian framework.

<sup>17</sup>Myrvold (2011) uses a similar approach to explain the Second Law.

proceeds in this manner. As is well known, neither mechanics nor a priori arguments can deliver a unique probability distribution for the initial state. Instead, our decision to adopt a uniform (or smooth) one is guided by the assumption that the observed frequencies of events have in fact been probable. But whilst this distribution may be empirically adequate, it allegedly can't *explain* the observed frequencies in a non-circular way, for their assumed likelihood is what steered us toward that distribution in the first place. Since past frequencies are disclosed through records, this criticism undermines the idea that an initial distribution can help explain the record asymmetry, which is central to my account (and Albert and Loewer's).

Whilst I acknowledge a circularity here, I think it is benign. If my goal was to convince a sceptical reader that the observed frequencies *are* probable, and hence that records *are* reliable, then my account would be doomed. However, I'm making this assumption in order to tell a story about how all this comes about, i.e. how an initial distribution gives rise to a fork asymmetry and hence a record asymmetry. If my account is successful in joining these dots, then it can still be contentful (and to that extent explanatory) despite containing a circularity of sorts. This situation seems analogous to a biologist taking her senses to be reliable in the course of constructing a physiological theory that explains *how* they're reliable. The explanation begins in medias res, but so long as it's expository rather than suasive, its circularity isn't fatal.

At this point, it's natural to wonder what the explanatory role of the Past Hypothesis is in my proposal. As is quite standard in Boltzmannian accounts, I require the Past Hypothesis to block Loschmidt's reversibility objection - a situation in which all our records would be fake. Additionally, the Past State's low entropy might help explain why there are macroscopic objects *at all* rather than ongoing heat death since the beginning - a precondition for the existence of records. For these reasons, the Past Hypothesis is the

explanatory foundation for the record asymmetry. However, it alone doesn't fully explain this phenomenon: the job is finished by the Smoothness Postulate, for the smooth character of the initial distribution underpins the fork asymmetry.

It's now possible to see how my proposal plugs the explanatory gap in Albert and Loewer's picture. Broadly speaking, they believe the Past Hypothesis plus an initial distribution ultimately explain the record asymmetry. But as we saw in section 2, they don't fill in the details of *how* the initial state achieves this. My account fleshes this out as follows: the Past Hypothesis provides the explanatory bedrock for the record asymmetry, whilst the initial distribution completes the explanation by underpinning the fork asymmetry. I prefer to call the initial distribution 'smooth' rather than 'uniform', but this is really a side-issue. Interestingly, Albert (2016, 58) acknowledges that although he thinks the Past Hypothesis (presumably, plus the Statistical Postulate) underpins the fork asymmetry, "the business of making that clear will require some further work". Perhaps I've found this missing piece.

Before wrapping up this section, one final point of comparison is worth making. Arntzenius (1992, 234) himself and others have attempted to derive the fork asymmetry from deterministic laws plus independence assumptions in the initial state.<sup>18</sup> However, our emphases are different. Their focus is on explaining the statistical independence of  $A$  and  $B$  (given  $C$ ), whereas I'm less concerned with this screening-off aspect, and more interested in explaining why  $A$  and  $B$  (qua correlated, observable events) are temporally and spatially proximate. There's no reason to think our projects are inconsistent, and they may even be complementary. However, both face an overhanging question: why do  $C$  events materialise in the first place? In other words, why should unobservable correlations ever become observable? Answering this would

---

<sup>18</sup>See also Pearl (2009) and Frisch (2014, 224-228).

fully explain how the fork asymmetry emerges from the initial state. For now, I can only gesture at the ‘radiative arrow’ as a promising lead. But it’s time we returned to the more concrete aspects of my proposal.

## 9 Two Possible Objections

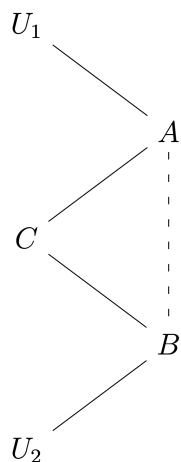
In this section I shall confront two possible objections to my account. The first threatens the Scrambler Assumption, whilst the second threatens to undermine the fork asymmetry all over again.

The first objection runs as follows. Given enough time, it looks as though observable correlations become unobservable (i.e. get scrambled) in a completely spontaneous manner - that is, *without* having a  $C$  in their future. To take one example, the correlated records ‘cancer’ and ‘yellow fingers’ have finite lifespans, and their disappearance doesn’t seem to require any sort of collision into a future  $C$  event. Does this not refute the Scrambler Assumption, which claims that the only way for observable correlations to become unobservable is via some  $C$ ?

My response is that observable correlations *do* in fact persist; it is just that the associated macrostates get harder to actually *use* as records of  $C$ . This is for two reasons: the become more motley, and also less reliable in their own right. I unpack these claims below.

Let us start with the issue of ‘motleyness’. Following some event, which particular records form hinges on which background events  $U_i$  obtain, such that certain records are correlated with certain  $U_i$ . For example, following ‘smoking’ ( $C$ ), ‘cancer’ ( $A$ ) will correlate with ‘ordinary chromosome 5’ ( $U_1$ ), whilst ‘yellow fingers’ ( $B$ ) will correlate with ‘low bleach exposure’ ( $U_2$ ). Moreover, there is no reason to expect these  $U_i$  to correlate with  $C$  (see

Figure 8). Because these  $U_i$  vary wildly, so too will the selection of records that might be realised, with the result that no *particular* of record (say,  $A$ ) is guaranteed to form.



**Figure 8:** The background events of records.

Applying some numbers to things, we might find:

$$\Pr(A \mid C) = 0.5 \tag{15}$$

$$\Pr(B \mid C) = 0.5 \tag{16}$$

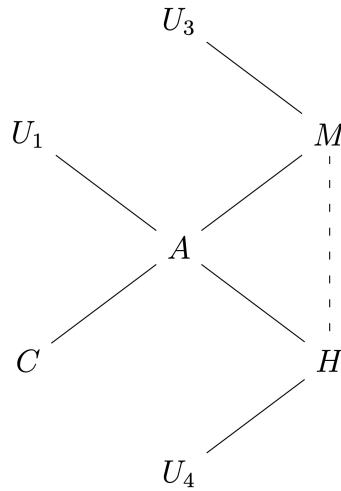
This whole story can be reiterated by taking any of the records ( $A$  or  $B$ ) to themselves act like  $C$ . As an illustration (see Figure 9), let us take this to be cancer ( $A$ ). This event may yield various possible ‘second-generation’ records, depending on which  $U_i$  obtain. For example, mourners ( $M$ ) will correlate with ‘having loved ones’ ( $U_3$ ), whilst ‘happy dogs at Battersea Dogs Home’ ( $H$ ) will correlate with ‘charitable will’ ( $U_4$ ).<sup>19</sup> Clearly, this can be

---

<sup>19</sup>Needless to say, there are many more possible  $U_i$  (and associated resultant records) than those described.



reiterated yet again for any of the new records ( $M$  or  $H$ ).



**Figure 9:** The background events of ‘records of a record’.

Keeping the numbers simple, we might find:

$$\Pr(M \mid A) = 0.5 \tag{17}$$

$$\Pr(H \mid A) = 0.5 \tag{18}$$

But (15), (17), and (18) then imply that given smoking, no particular second-generation record is very likely:

$$\Pr(M \mid C) = 0.25 \tag{19}$$

$$\Pr(H \mid C) = 0.25 \tag{20}$$

Even later records will of course be *even less* likely to materialise given  $C$ , so that the longer we wait around after  $C$ , the less likely it is that any given record will form.

So, with time, older records of  $C$  are replaced by a selection from a growing roster of possible newer records, amongst which observable correlations persist. As we saw, ‘cancer’ may yield ‘mourners’ or ‘happy dogs’, whilst ‘yellow fingers’ may yield (say) ‘hydrogen peroxide treatment’ or ‘exfoliation’. But even though these second-generation records could *in principle* be used to triangulate  $C$ , they are less useful for this purpose, for their appearances depend on a longer list of the right  $U_i$  obtaining. As a comparison, whereas (following  $C$ ) the first-generation record  $A$  hinges only on  $U_1$ , the second-generation record  $M$  hinges on both  $U_1$  and  $U_3$ . So whilst observable correlations proliferate following  $C$ , the associated macrostates become ever more motley, and hence ever less symptomatic of the original event  $C$ . Eventually, we just don’t know what records to look out for any more.

Before moving on, I wish to make a clarificatory remark. According to the Scrambler Assumption,  $C$  is necessary for the scrambling of observable correlations - a process that generally only happens toward the past. For this to be consistent with the idea that motleyness increases toward the future, then there needs to be a sharp distinction between a collection of motley events, which *can* arise spontaneously from observable events like  $A$  and  $B$ , and an unobservable event, which *cannot* arise in this fashion (and hence requires  $C$ ’s intervention). Otherwise, the motleyness of later-generation records would translate to their unobservability, making  $C$  obsolete for bringing about scrambling.

But I think this distinction has a solid basis. A collection of motley events is a union of macrostates whose macrovariables we enter into correlations with. By contrast, an unobservable event is a gerrymandered region in phase space which *isn’t* characterised by macrovariables that we enter into correlations with. So, motleyness and unobservability present different barriers to learning about other times: motley events that exist long after  $C$  are perfectly observable, but just hard to recognise as records, whereas unobservable

events that exist prior to  $C$  aren't the sorts of things we can even register in the first place.

Let us now turn to the second issue, which concerns unreliability. As mentioned earlier, putative records of  $C$  do not guarantee that event, but raise its probability (to a greater or lesser degree). For instance, rather than resulting from smoking, cancer could indicate a poor diet, whilst yellow fingers could indicate Raynaud's disease. Using familiar labelling and keeping the numbers simple, our imperfect records might tell us:

$$\Pr(C \mid A) = 0.5 \tag{21}$$

$$\Pr(C \mid B) = 0.5 \tag{22}$$

The same reasoning applies to our second-generation records. Mourners ( $M$ ) might indicate a car crash, whilst the dogs at Battersea might be happy ( $H$ ) because they were taken for a walk; neither guarantees a past cancer tragedy. So, we might find:

$$\Pr(A \mid M) = 0.5 \tag{23}$$

$$\Pr(A \mid H) = 0.5 \tag{24}$$

From equations (21), (23), and (24), we can see that inferring  $C$  from our second-generation records is unreliable:

$$\Pr(C \mid M) = 0.25 \tag{25}$$

$$\Pr(C \mid H) = 0.25 \tag{26}$$

The probabilities will be smaller still if we conditionalise on *even later* putative records - say, not on mourners, but on black clothing that might have served a non-funerary purpose. This is because they have longer genealogies

reaching back to  $C$ , and hence have more opportunities for a spurious origin. So, the longer we wait around after  $C$ , the less reliable records become. This issue of unreliability is distinct from the issue of motleyhood: it implies that even if knew which very-late-generation records of  $C$  to look out for, they wouldn't be of much use anyway, since they'd be so unreliable.

Let's now turn to the second objection, which runs as follows. It can sometimes happen that  $C$  very reliably produces a certain record  $R$ , and likewise  $R$  very reliably indicates  $C$ . We can visualise this by interpreting  $C$  as 'elephant wandering Soho', and  $R$  as 'photo of an elephant in Soho'.  $R$  is not a true determinant of  $C$ , for  $R$  could be formed by photo-editing software, and likewise the Soho scene could have occurred without anyone producing  $R$ . But since these circumstances are very unlikely,  $R$  stands in for  $D$  in watered-down versions of equations (11)-(14):

$$\Pr(R | C) \approx 1 \tag{27}$$

$$\Pr(R | \neg C) \approx 0 \tag{28}$$

$$\Pr(C | R) \approx 1 \tag{29}$$

$$\Pr(C | \neg R) \approx 0 \tag{30}$$

We might therefore say that  $R$  is a 'pseudo-determinant' of  $C$ . So here is the worry: granted that  $R$  is not a full-blown determinant of  $C$ , might its status as a pseudo-determinant still constitute a problem for my account? More pointedly, if  $C$  leads to correlated events  $A$  and  $B$  (respectively, 'astonished bystander' and 'trampled pub blackboard') that serve as records of  $C$ , and if  $R$  lies in the future of  $A$  and  $B$ , then do we not have something *very much like* a backward fork, with  $A$  and  $B$  as the fork-tips and  $R$  as the fork-point? Arntzenius' first objection threatens to resurface in weakened form.

One response is that even if there are occasional instances where we can use  $A$  and  $B$  to make highly reliable inferences about the future, this would not threaten my account so long as such cases are rare. Since the record asymmetry is very palpable, it seems clear that such cases *are* in fact rare. Therefore, the existence of pseudo-determinants doesn't pose a threat.

But we should also remind ourselves that even in such scenarios, the ability of  $A$  and  $B$  to inform us about some future  $R$  piggybacks on their ability to inform us about some past  $C$  *even more* reliably. Since 'pseudo-records' of the future (in this case,  $A$  and  $B$ ) are unusual, and in any case double up as *full-blown* records of the past, they do not undermine the record asymmetry.

## 10 Conclusion

The fork asymmetry is a *many-pronged* phenomenon which we can think about in two different ways. By envisioning its fork-tips as discrete records, we can explain their collective effectiveness: multiple records working in tandem afford us a) reliable, b) detailed, c) far-reaching, and d) accessible information about the past. But by envisioning its fork-tips as sub-components of a single record, we can explain their individual effectiveness, i.e. the fact that a single record can account for a)-d). Hence, the fork asymmetry can do real justice to the record asymmetry.

My fork-asymmetry-based proposal revives a tradition that's fallen by the wayside due to Arntzenius objections, which I've avoided. Although this takes a different tack to entropy-based explanations, it nevertheless seems to plug an explanatory gap in the most salient amongst these, namely Albert and Loewer's theory. The gap amounts to the fact that it's not totally clear how and why the initial state underpins the record asymmetry. I've remedied this with the following idea: the Past Hypothesis provides the

foundation by blocking Loschmidt's objection and facilitating the existence of macroscopic objects, whilst the initial distribution's character - in my book, its smoothness rather than its uniformity - completes the explanation by underpinning the fork asymmetry.

Despite their obvious interconnectedness, time's various arrows stubbornly resist integration into a single explanatory sequence. But by tightening up a few fragments of this tapestry, I hope to have brought their estrangement a step closer to an end.

## References

Albert, D. Z. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.

Albert, D. Z. (2016). *After Physics*. Cambridge, MA: Harvard University Press.

Arntzenius, F. (1990). Physics and Common Causes. *Synthese* 82(1), 77–96.

Arntzenius, F. (1992). The Common Cause Principle. In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1992, pp. 227–237. Philosophy of Science Association.

Berkovitz, J., R. Frigg, and F. Kronz (2006). The Ergodic Hierarchy, Randomness and Hamiltonian Chaos. *Studies in History and Philosophy of Modern Physics* 37(4), 661–691.

Brown, H. R., W. C. Myrvold, and J. Uffink (2009). Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics. *Studies In History and Philosophy of Modern Physics* 40(2), 174–191.

- Callender, C. (2010). The past hypothesis meets gravity. In G. Ernst and A. Hüttemann (Eds.), *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*, pp. 34–58. Cambridge: Cambridge University Press.
- Earman, J. (2006). The “Past Hypothesis”: Not even false. *Studies In History and Philosophy of Modern Physics* 37(3), 399–430.
- Feynman, R. P. (1965). *The Character of Physical Law*. Cambridge, MA: MIT Press.
- Frigg, R. (2009). Typicality and the approach to equilibrium in Boltzmannian statistical mechanics. *Philosophy of Science* 76(5), 997–1008.
- Frisch, M. (2005a). Counterfactuals and the Past Hypothesis. *Philosophy of Science* 72(5), 739–750.
- Frisch, M. (2005b). *Inconsistency, Asymmetry, and Non-Locality: A Philosophical Investigation of Classical Electrodynamics*. Oxford: Oxford University Press.
- Frisch, M. (2010). Does a Low-Entropy Constraint Prevent Us from Influencing the Past? In G. Ernst and A. Hüttemann (Eds.), *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*, pp. 13–33. Cambridge: Cambridge University Press.
- Frisch, M. (2014). *Causal Reasoning In Physics*. Cambridge: Cambridge University Press.
- Goldstein, S. (2012). Typicality and Notions of Probability in Physics. In Y. Ben-Menahem and M. Hemmo (Eds.), *Probability in Physics*, The Frontiers Collection, pp. 59–71. Berlin Heidelberg: Springer-Verlag.
- Grünbaum, A. (1963). *Philosophical Problems of Space and Time*. New York: Alfred A Knopf.

- Hartle, J. (2004). The Physics of ‘Now’. *American Journal of Physics* 73(2), 101–109.
- Hawking, S. (1985). The arrow of time in cosmology. *Physical Review D* 32(10), 2489–2495.
- Hawking, S. (1994). The No Boundary Condition and the Arrow of Time. In J. J. Halliwell, J. Pérez-Mercader, and W. Zurek (Eds.), *Physical Origins of Time Asymmetry*, pp. 346–357. Cambridge: Cambridge University Press.
- Hemmo, M. and O. R. Shenker (2012). *The Road to Maxwell’s Demon: Conceptual Foundations of Statistical Mechanics*. Cambridge: Cambridge University Press.
- Hemmo, M. and O. R. Shenker (2015a). The emergence of macroscopic regularity. *Mind & Society Society* 14(2), 221–244.
- Hemmo, M. and O. R. Shenker (2015b). Probability and Typicality in Deterministic Physics. *Erkenntnis* 80(3), 575–586.
- Hemmo, M. and O. R. Shenker (2016). *Maxwell’s Demon*. Oxford Handbooks Online. DOI: 10.1093/oxfordhb/9780199935314.013.63: Oxford University Press.
- Hemmo, M. and O. R. Shenker (2020). Can the Past Hypothesis Explain the Psychological Arrow of Time? In V. Allori (Ed.), *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature*, pp. 255–287. World Scientific.
- Hitchcock, C. (2009). Causal Modelling. In H. Beebe, C. Hitchcock, and P. Menzies (Eds.), *The Oxford Handbook of Causation*, pp. 299–314. New York: Oxford University Press.
- Horwich, P. (1988). *Asymmetries in Time: Problems in the Philosophy of Science* (2 ed.). Cambridge, MA: MIT Press.



- Lebowitz, J. L. (1993). Boltzmann's Entropy and Time's Arrow. *Physics Today* 46, 32–38.
- Lewis, D. K. (1979). Counterfactual Dependence and Time's Arrow. *Noûs* 13, 455–467.
- Loewer, B. (2007). Counterfactuals and the Second Law. In H. Price and R. Corry (Eds.), *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited*, pp. 293–326. New York: Oxford University Press.
- Loewer, B. (2011). The emergence of time's arrows and special science laws from physics. *Interface Focus* 2(1), 13–19.
- Loewer, B. (2012). Two Accounts of Laws and Time. *Philosophical Studies* 160(1), 115–137.
- Maroney, O. J. E. (2010). The (absence of a) relationship between thermodynamic and logical irreversibility. *Studies In History and Philosophy of Modern Physics* 36(2), 355–374.
- Mlodinow, L. and T. A. Brun (2014). On the Relation between the Epistemic and the Thermodynamic Arrow of Time. *Physical Review E* 89(5), 052102.
- Myrvold, W. C. (2011). Probabilities in Statistical Mechanics: Subjective, Objective, or a Bit of Both? Preprint available at <http://philsci-archive-dev.library.pitt.edu/8642/>.
- Papineau, D. (1985). Causal Asymmetry. *British Journal for the Philosophy of Science* 36(3), 273–289.
- Papineau, D. (1992). Can We Reduce Causal Direction to Probabilities? In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, Volume 1992, pp. 238–252. Philosophy of Science Association.
- Parker, D. (2005). Thermodynamic Irreversibility: Does the Big Bang Explain What It Purports to Explain? *Philosophy of Science* 72(5), 751–776.

- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2 ed.). Cambridge: Cambridge University Press.
- Penrose, R. (1990). *The Emperor's New Mind: Concerning Computers, Minds, and the Laws of Physics*. London: Vintage.
- Reichenbach, H. (1956). *The Direction of Time*. Los Angeles, CA: University of California Press.
- Schulman, L. S. (2005). A Computer's Arrow of Time. *Entropy* 7(4), 221–233.
- Smart, J. J. C. (1967). Time. In P. Edwards (Ed.), *Encyclopedia of Philosophy*, Volume 8, pp. 126–132. New York: Macmillan.
- Smith, R. (2014). Do Brains Have an Arrow of Time? *Philosophy of Science* 81(2), 265–275.
- Strevens, M. (2016). The Reference Class Problem in Evolutionary Biology: Distinguishing Selection from Drift. In C. H. Pence and C. Ramsey (Eds.), *Chance in Evolution*, pp. 145–175. Chicago: University of Chicago Press.
- Uffink, J. (2007). Compendium of the Foundations of Classical Statistical Physics. In J. Butterfield and J. Earman (Eds.), *Handbook for Philosophy of Physics*, pp. 923–1047. Amsterdam: Elsevier.
- von Plato, J. (1983). The method of arbitrary functions. *British Journal for the Philosophy of Science* 34(1), 37–47.
- Wallace, D. (2012). *The Emergent Multiverse: Quantum Theory according to the Everett Interpretation*. New York: Oxford University Press.
- Winsberg, E. (2004). Can Conditionalising on the “Past Hypothesis” Mitigate Against the Reversibility Objections? *Philosophy of Science* 71(4), 489–504.

Wolpert, D. H. (1992). Memory Systems, Computation, and the Second Law of Thermodynamics. *International Journal of Theoretical Physics* 31(4), 743–785.