Inference by Multiple Identical Observers

M. T. Barlow*

September 3, 2025

Abstract

We consider models for inference made by observers which may have multiple identical copies, such as in the Sleeping Beauty problem. We establish a framework for describing these problems on a probability space satisfying Kolmogorov's axioms. This enables the main competing approaches to be compared precisely.

1   Introduction

This paper is concerned with inference by observers whose existence and multiplicity is affected by the outcome of the random experiment which they are observing. One can argue that implicit in the classical model of a random experiment is the assumption that observations and statistical inferences are made by a Classical Observer (CO), which exists before the start of the experiment, and whose existence is not affected by the outcome of the experiment. Little attention has been paid to this requirement since these conditions are nearly always satisfied in practice.

However non Classical Observers do arise in physics and cosmology. The laws of nature depend on about 20-30 constants, many of which appear "fine tuned' for life – see [1, 26]. Among these is the cosmological constant Λ, which has a value much smaller than that predicted by some theories. One type of explanation is via a multiverse, which contains many universes, each with a different value of Λ. Only a small range of these values will permit the existence of stars, planets, life and observers. The value for our universe Λ_0 has to lie in this range, and we can ask whether this observed value is likely given the multiverse theory. To answer this we need to know how the (random) observers in these universes should make inferences. (For some recent papers in this area see [5, 29].)

Following the terminology introduced by Brandon Carter in [8], I will call the conditional observers arising in examples like this *Anthropic Observers* or AO. (The term is flawed as these observers do not have to be human.) Most examples allow multiple observers, and one needs to know how the multiplicity of these observers affects inferences.

A 'toy model' one can use to investigate these questions is the Sleeping Beauty problem, described by A. Elga in [12] in the philosophy journal *Analysis*. The name is due to Robert Stalnaker, an earlier version of the problem is due to A. Zuboff (see [33]), and an essentially identical problem is given in Example 5 of [25], a paper on the Absentminded

---

*Department of Mathematics, University of British Columbia, Vancouver, BC V6T 1Z2, Canada. barlow@math.ubc.ca

Driver problem in decision theory. Elga introduced the problem in order to examine how observers update probabilities in centered and uncentered worlds.

On Sunday the experiment, as described below, is explained to Sleeping Beauty (SB). On Sunday night she goes to sleep in an isolated cell. A fair coin is tossed. Whatever the outcome she is awakened on Monday morning, and then on Monday evening is given a potion which puts her to sleep and causes her to forget everything that occurred on Monday. If the outcome of the toss was Tails she is then woken on Tuesday, while if the outcome was Heads then she is not woken on Tuesday. We ask two questions:

Question 1. When SB wakes in her cell during the experiment, what probability should she assign to the event that the coin landed Heads?

Question 2. If she is then told that it is Monday, what probability should she now assign to Heads?

The Sleeping Beauty problem has been framed in a fanciful fashion, but it raises questions which are relevant in mainstream research in physics and cosmology. These problems have been actively discussed by physicists and philosophers, but they also deserve attention from the mathematical community.

Philosophers' use the term *credence* to denote a subjective probability. I will use that term as shorthand for "the probability law that SB should use", but without wishing to be tied to any particular philosophical interpretation of probability. There are four positions in the philosophy literature. *Halfers* claim that the answer to Question 1 is $\frac{1}{2}$. *Standard Halfers* claim that the answer to Question 2 is $\frac{2}{3}$, while *Double Halfers* argue that the answer to Question 2 is still $\frac{1}{2}$. *Thirders* argue that the answer to Question 1 is $\frac{1}{3}$, and then deduce that the answer to Question 2 is $\frac{1}{2}$. A final group argue that the problem is ambiguous, inappropriate or undetermined. In his excellent review [32] Winkler suggests that the Thirder view is the majority position among philosophers, but that this is not reflected in the literature due to publication bias. Physics papers such as [5] use a weighting which aligns with the Thirder approach.

A further question is whether this problem can be adequately described using Kolmogorov's axioms for probability. The main goal of this paper is to argue that it can. The framework we construct then allows the competing views of Halfers and Thirders to be evaluated in precise mathematical terms.

Elga [12] proposed a straightforward solution of the SB problem using conditional probabilities. He claimed that

$$\mathbb{P}(\text{Mon}|\text{Tails}) = \tfrac{1}{2}, \tag{1.1}$$

$$\mathbb{P}(\text{Heads}|\text{Mon}) = \mathbb{P}(\text{Tails}|\text{Mon}) = \tfrac{1}{2}. \tag{1.2}$$

Equation (1.1) follows from a "highly restricted" principle of indifference. To justify (1.2), since SB is always woken on Monday morning the coin toss can be deferred until Monday afternoon. If SB is told at midday on Monday that it is Monday, then the (fair) coin toss is still in the future, and so SB should assign probability $\frac{1}{2}$ to Heads. (We can ask SB to toss the coin.) Then

$$\mathbb{P}(\text{Mon \& Tails}) = \mathbb{P}(\text{Mon}|\text{Tails})\mathbb{P}(\text{Tails}) = \mathbb{P}(\text{Tails}|\text{Mon})\mathbb{P}(\text{Mon}), \tag{1.3}$$

so that $\mathbb{P}(\text{Mon}) = \mathbb{P}(\text{Tails}) = 1 - \mathbb{P}(\text{Heads})$. As $\mathbb{P}(\text{Heads}|\text{Tue}) = 0$ we have

$$\mathbb{P}(\text{Heads}) = \mathbb{P}(\text{Heads}|\text{Mon})\mathbb{P}(\text{Mon}) = \tfrac{1}{2}\mathbb{P}(\text{Mon}) = \tfrac{1}{2} - \tfrac{1}{2}\mathbb{P}(\text{Heads}),$$

and so $\mathbb{P}(\text{Heads}) = \tfrac{1}{3}$.

This argument is simple enough that it seems surprising that the conclusion should be disputed. Nevertheless Standard Halfers reject (1.2), while Double Halfers need to redefine what is meant by the conditioning in (1.2).

The starting point for this paper is to ask what probability space the calculations above take place in. In the standard SB model there is only one randomization, the toss of a fair coin, which suggests that it should be possible to define the model on the space $\Omega_O = \{H, T\}$ with $\mathbb{P}(\{H\}) = \mathbb{P}(\{T\}) = \tfrac{1}{2}$. However this cannot be the space used in Elga's calculation: no event in Ω_O has probability $\tfrac{1}{3}$, and the events $\{\text{Mon}\}$ and $\{\text{Tue}\}$ do not lie in this space. These events also appear to be ambiguous – an outside observer of the experiment will experience both Monday and Tuesday.

The set of days is $\Omega_A = \{\text{Mon}, \text{Tue}\}$, and so the state of SB on waking is described by the pair

$$\omega = (\omega_o, \omega_a) \in \Omega_O \times \Omega_A.$$

In the terminology of the philosophy literature elements of Ω_O describe 'possible worlds', while elements of Ω_A give the observer's location in the world. The general tendency has been to regard these two components as being at fundamentally different levels, a point most clearly indicated by the introduction in [15, 22, 6] of a new form of conditioning which acts in different ways on the two components of Ω.

However, when SB is woken, she is uncertain both about ω_o and ω_a, and it seems natural to place her ignorance about these at the same level. In this paper we will define our model on the space $\Omega = \Omega_O \times \Omega_A$. It is viewed by two observers. The classical observer CO only sees the space Ω_O with its objective probability $\mathbb{P}_O$. The AO sees the space Ω, and the problem for the AO is to choose a suitable probability P_A on Ω. This is not an 'extension problem' since we will not in general require that

$$P_A(F \times \Omega_A) = \mathbb{P}_O(F) \quad \text{for } F \subset \Omega_O. \tag{1.4}$$

As we cannot expect to derive P_A from $\mathbb{P}_O$ using the axioms of probability our task divides into two parts.

The first is to formulate reasonable properties (or 'Principles') which restrict the possible class of measures P_A; these Principles have to come from our real world intuitions. In mathematical terms these Principles provide links between the space $(\Omega_O, \mathbb{P}_O)$ seen by the CO and the space (Ω, P_A) seen by the AO. The second problem is to identify the class of measures allowed by these Principles, and work out their properties.

One Principle mentioned in the philosophy literature is Lewis' Principal Principle (see [20]) which roughly states that SB should use a probability P_A which satisfies (1.4). We will see below strong reasons for rejecting a straightforward use of this Principal Principle.

The remainder of this paper is as follows. In Section 2 we introduce a generalised Sleeping Beauty problem (GSB), where the set of occupied cells is a random subset $\mathcal{X}$ of finite

set K. Since we are interested in inferences by the AO we introduce an auxiliary process Z, and we may allow the AO in cell x to observe the value Z_x. We set up the extended space Ω and formulate four Principles which give apparently desirable properties of the AO's probability P_A. The first, denoted (PN), is a mild condition on null sets, and will be assumed throughout this paper. The second, the *Principle of Indifference* (PI), is an extension of (1.1). The third, the *Principle of Equivalent Information* (PEI), generalises (1.2). This principle states roughly that a Classical Observer and an Anthropic Observer who are in communication and have shared all the information they have available should assign the same conditional probabilities to events in the space Ω_O. The final Principle, denoted (PP), is an extension of (1.4). We find that, except when the number of AO is deterministic (conditional on there being any AO at all), there is no probability P_A which satisfies all four Principles – see Corollary 2.7. If we drop (PP) then there is a unique probability, which we denote P_E, which satisfies the remaining three Principles. If we require (PP) then there exists a natural probability P_L which satisfies (PI).

In Example 2.9 we look at the standard Sleeping Beauty problem and show that $P_E(\text{Heads}) = \frac{1}{3}$ and $P_L(\text{Heads}) = \frac{1}{2}$; we will therefore refer to P_E and P_L as the *Thirder* and *Halfer* measures respectively. In Section 2.2 (see Theorem 2.15) we consider sequential experiments such as the original SB experiment, and show that (PEI) is equivalent to the assertion that the AO should assign the objective probability to random events which are still in its future.

Section 3 looks at an AO's inferences based on observation of the process Z. Section 3.2 shows that one needs to be careful when one conditions on an AO's observations, and Section 3.3 looks at the model in [16], a paper where insufficient care was taken. Section 4 looks at three simple examples from astronomy and cosmology, and we compare the inferences made by AO which use P_E and P_L. In the third example we see that in a system with a large number of observers, and many independent components, inferences using P_E and P_L do not differ by much. (These conditions do typically hold in multiverse models.)

The very extensive literature contains many ideas, examples and calculations. In some cases a space such as our extension $\Omega = \Omega_O \times \Omega_A$ is implicit, but the framework given here does not seem to have been set out precisely before, and the mathematical formulation of the key principle (PEI) seems to be new. [14] does consider probability laws on the sample space $\Omega = \Omega_O \times \Omega_A$, but rejects, at least for some probability laws on Ω, the possibility of conditioning on events like {Mon}. Thus (PEI) would be regarded as inadmissible in the setup given there. The calculations in [27] use an assertion similar to (PEI).

The approach of this paper, which is natural from a probabilistic viewpoint, concentrates on information and conditional probabilities, and owes much to Elga's paper [12]. An alternative approach, as seen in for example Bostrom's book [7], is to assume that the observer's location has been randomly selected in some fashion from the set of all possible locations. The type of random selection then determines the outcome: Bostrom's *Self Sampling Assumption* (SSA) leads to Halfer type conclusions, while his *Self Indication Assumption* (SIA) gives Thirder answers.

Two other approaches to determining probabilities are using betting arguments or long run frequencies. Neither of these give a clear answer for the SB problem. In general betting

arguments have to allow for the possibility of multiple bets by SB, and the question then arises how she should regard gains or losses by her other copies. See [2] for the type of difficulties which arise. Given a series of (independent) SB experiments, once a week over many weeks, it is clear that the long run proportion of 'Heads weeks', that is weeks with a Heads toss, will be $\frac{1}{2}$, while the proportion of awakenings in 'Heads weeks' will be $\frac{1}{3}$. The difficulty is how to match these limits with SB's credences in a mathematically precise fashion.

One reason philosophers do not agree on the solution to the Sleeping Beauty problem is that all known approaches lead to paradoxical conclusions. For the Halfers the problems arise quickly: standard Halfers find themselves holding that SB's position gives her information about the future, while Double Halfers have to redefine conditional probabilities. See Example 2.9 for details. The difficulties for Thirders do not arise so immediately. However Bostrom's example of the Presumptuous Philosopher (see [7, p. 124] and Example 2.11) suggests that the methods of inference used by Thirders, together with some plausible extensions, lead to the conclusion that, without needing to make any observations at all, we should assign a credence of 1 to the event that the universe is infinite.

2   The generalized Sleeping Beauty model

2.1   Probability spaces and observers

We generalize the Sleeping Beauty problem as follows. Let K be a finite set with $|K| = M$ where $M \geq 1$. Let $\mathcal{P}(K)$ be the set of all subsets of K, and let $(\Omega_O, \mathcal{F}_O, \mathbb{P}_O)$ be a probability space carrying a random variable $\mathcal{X}^O : \Omega_O \to \mathcal{P}(K)$, and a process $Z^O = (Z_x^O, x \in K)$ taking values in a finite set $\mathcal{Z}$. We assume that $0 \leq \mathbb{P}_O(\mathcal{X} = \emptyset) < 1$. We call $(\Omega_O, \mathcal{F}_O, \mathbb{P}_O)$ the *objective probability space*. Most of our arguments generalize to the case when K is an infinite set and $\mathbb{E}(|\mathcal{X}^O|) < \infty$, but since the main challenges of the problem are already present in the case when K is finite, we will restrict to that case.

Associated with each $x \in K$ is a cell C_x. The cells are identical, and are labelled, but on the outside, so that an occupant does not know its cell label. Identical anthropic observers are placed in each cell C_x with $x \in \mathcal{X}$. Each cell is equipped with a telephone which permits only incoming calls. A Classical Observer (CO) is situated outside the collection of cells, and does not initially know the value of $\mathcal{X}^O$. The CO has a telephone which can be used to call cell C_x. By calling cell C_x the CO is able to verify whether or not $x \in \mathcal{X}^O$. (We assume that the AO always answers the telephone.)

If we wish to implement the model with a single AO, as in the original SB problem, we choose a bijection π between $\{1, \ldots, M\}$ and K. The experiment is run for M days. On each day j, for $1 \leq j \leq M$, the AO is placed in cell $C_{\pi(j)}$ for that day. If $\pi(j) \in \mathcal{X}$ then the AO is woken on day j, and at the end of day j the AO's memory of that day is erased. If $\pi(j) \notin \mathcal{X}$ then the AO is not woken on day j.

To handle the experience of the AO we extend the probability space as follows. Let ∂ be a point not in K, and let

$$\Omega_A = K \cup \{\partial\}, \quad \Omega = \Omega_O \times \Omega_A.$$

Let $S : \Omega \to \Omega_A$ be defined by $S((\omega, x)) = x$ for $\omega \in \Omega_O$, $x \in \Omega_A$. Set

$$\mathcal{F}_W = \{F \times \Omega_A, F \in \mathcal{F}_O\}, \quad \mathcal{F} = \sigma(\mathcal{F}_W, S) = \sigma(F \times \{x\}, F \in \mathcal{F}_O, x \in \Omega_A). \qquad (2.1)$$

It is straightforward to check that $F \in \mathcal{F}$ if and only if there exist $F_x \in \mathcal{F}_W$ such that

$$F = \bigcup_{x \in \Omega_A} \{S = x\} \cap F_x. \qquad (2.2)$$

We call $\mathcal{F}_W$ the *objective σ-field*, and events in $\mathcal{F}_W$ *objective* events. In the terminology in the philosophy literature, events in $\mathcal{F}_W$ tell us about possible Worlds. We extend $\mathbb{P}_O$ to a probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F}_W)$ by setting $\mathbb{P}(F \times \Omega_A) = \mathbb{P}_O(F)$. Throughout this paper we will use a superscript 'O' to denote random variables on the base space Ω_O, and remove it to denote the extension of the random variable to Ω. Thus if Y^O is defined on Ω_O, then $Y((\omega, x)) = Y^O(\omega)$ for $(\omega, x) \in \Omega$.

The space $(\Omega, \mathcal{F}_W, \mathbb{P})$ describes the experience of the classical observer. Over the course of the experiment the CO will observe that a variety of cells are occupied on different days, but as none of these has a distinguished status the random variable S is not observed by, and makes no sense to, the CO.

We now look at the viewpoint of an AO, which wakes in an initially unknown cell. If $G \in \mathcal{F}_O$ and $x \in K$ then the interpretation of the event $G \times \{x\} \in \mathcal{F}$ is that the event G occurs, and the AO is situated in cell C_x. As $S((\omega, x)) = x$ we see that the random variable S gives the location of the AO. (The point ∂ is needed to give the value of S when $\mathcal{X} = \emptyset$, and there are no anthropic observers.)

Many papers in the probability literature look at models where observers have different information or use different probability measures. For example, the 'expansion of filtrations' literature, following [3, 18], looks at stochastic processes viewed by two observers with different information sets. Systems with multiple probability measures are common in the mathematical finance literature.

The specific setup here, with the observers CO and AO, seems to be new. The CO uses the probability space $(\Omega, \mathcal{F}_W, \mathbb{P})$, while the AO observes measurable functions on the space $(\Omega, \mathcal{F})$. The question for the AO is what probability to place on this space. In the theory of Markov chains on a countable state space I one can derive the law $\mathbb{P}^x$ of the process started at any point $x \in I$ by conditioning from a single probability $\mathbb{P}^\mu$ where μ is a probability measure on I with full support. Similarly a law P_A on $(\Omega, \mathcal{F})$ can cover all possible locations of the AO, since the laws $P_A(\cdot | S = x)$ give the credences of the AO if it learns that it is in cell C_x.

Definition 2.1. We define the following Principles.

Principle of null sets (PN).
(a) If $F \in \mathcal{F}_W$ and $\mathbb{P}(F) = 0$ then $P_A(F) = 0$.
(b) $P_A(S \in \mathcal{X}) = 1$. (It follows that $P_A(\mathcal{X} = \emptyset) = 0$.)

Principle of Indifference (PI). For $x \in K$, $B \subset K$, $B \neq \emptyset$, with $P_A(\mathcal{X} = B) > 0$,

$$P_A(S = x | \mathcal{X} = B) = |B|^{-1} 1_B(x). \qquad (2.3)$$

6

Principle of Equivalent Information (PEI). If $F \in \mathcal{F}_W$ and $P_A(S = x) > 0$ then

$$P_A(F|S = x) = \mathbb{P}(F|x \in \mathcal{X}). \tag{2.4}$$

Principal Principle (PP). For $F \in \mathcal{F}_W$ we have

$$P_A(F) = \mathbb{P}(F|\mathcal{X} \neq \emptyset). \tag{2.5}$$

The first of these places mild restrictions on null sets, while other three are extensions of (1.1), (1.2) and (1.4).

The intuition for (PI) is that if an AO is told that $\mathcal{X} = B$ and has no other information, then it should decide it is equally likely to be in each of cells $\{C_y, y \in B\}$. It is difficult to see any reason to choose an alternative distribution.

To see the motivation for (PEI), we fix $x \in K$ before the start of the experiment; this value is known to the AO. The CO phones cell C_x. If the phone is answered than the CO knows that the event $\{x \in \mathcal{X}\}$ occurs. If the AO receives a phone call, it knows that it is in cell x. The AO and CO are in contact, and are allowed to share any information they have on the experiment. (In fact neither has anything to add to what the other already knows.) Since they have the same information they should have identical views on the probability of any $\mathcal{F}_W$-measurable event F, and hence (2.4) should hold. We remark that the point of the phone conversation is just to emphasize the fact that the CO and AO have the same information, and that it may still be reasonable to assume (PEI) in contexts where such a conversation cannot occur. Note that both [15, 24] reject assertions similar to (PEI).

The restriction in (PEI) that $F \in \mathcal{F}_W$ is necessary, since otherwise the right hand side is undefined. While the events $\{S = x\}$ are disjoint, the events $\{x \in \mathcal{X}\}$ are not, and this asymmetry may cause one to ask whether (PEI) gives the right mathematical description of the equivalent information of the CO and AO. However we will see below that for some models there are strong reasons for accepting (PEI).

Remark 2.2. 1. None of these principles involve the bijection π between K and $\{1, \ldots, M\}$. If we had an AI which could make Bayesian inferences at a level comparable with the most able humans we could consider a simultaneous experiment, where identical AO/AI are placed in each cell C_x, for $x \in \mathcal{X}$. Some quibbles about the original SB experiment can be overcome if we use AI. In particular the cells can be virtual, the observations of the AO can be tightly controlled, and memories can be completely erased. 2. It is conceivable that there are two different real world experiments R_1 and R_2 which are described by the same mathematical model (i.e. the same set K and processes $\mathcal{X}$ and Z), but where one would have different levels of confidence in the applicability of the various Principles. For example, it might be very reasonable to assume (PEI) for R_1 but not for R_2.

For $B \subset K$, $x \in K$ define

$$q_B = \mathbb{P}(\mathcal{X} = B), \quad Q_x = \mathbb{P}(x \in \mathcal{X}) = \sum_{A:x \in A} q_A. \tag{2.6}$$

7

We define a graph structure on K by defining $\{x, y\}$ to be an edge if $\mathbb{P}(\{x, y\} \subset \mathcal{X}) > 0$. Let E be the set of edges. We set

$$X = |\mathcal{X}|. \tag{2.7}$$

Theorem 2.3. *(a) Suppose that (K, E) is connected. There is a unique probability P_E on $(\Omega, \mathcal{F})$ which satisfies (PN), (PI) and (PEI). Writing $\lambda = 1/\mathbb{E}(X)$ we have*

$$P_E(F \cap \{S = x\}) = \lambda \mathbb{P}(F \cap \{x \in \mathcal{X}\}) \text{ for } F \in \mathcal{F}_W, x \in K. \tag{2.8}$$

In particular for $x \in K$, $B \subset K$,

$$P_E(S = x) = \lambda Q_x, \quad P_E(\mathcal{X} = B) = \lambda |B| q_B, \tag{2.9}$$

and if $F \in \mathcal{F}_W$ then

$$P_E(F) = \frac{\mathbb{E}(1_F X)}{\mathbb{E}(X)}. \tag{2.10}$$

(b) The probability P_E defined by (2.8) satisfies (PN), (PI) and (PEI).

Proof. (a) We use the same basic strategy as in Elga [12]. Note that the conditions on K and $\mathcal{X}$ imply that $Q_x > 0$ for each $x \in K$. Let P_E be a probability which satisfies the hypotheses, and write

$$t_B = P_A(\mathcal{X} = B), \text{ for } B \subset K, \quad s_x = P_A(S = x) \text{ for } x \in K. \tag{2.11}$$

Let $B \subset K$ with $q_B > 0$ and let $x \in B$. If $t_B > 0$ then by (PI)

$$P_E(S = x, \mathcal{X} = B) = P_E(S = x | \mathcal{X} = B) t_B = \frac{t_B}{|B|} > 0, \tag{2.12}$$

and thus $s_x > 0$. If $s_x > 0$ then by (PEI)

$$P_E(S = x, \mathcal{X} = B) = P_E(\mathcal{X} = B | S = x) s_x = \mathbb{P}(\mathcal{X} = B | x \in \mathcal{X}) s_x = \frac{s_x q_B}{Q_x} > 0, \tag{2.13}$$

so that $t_B > 0$. Hence

$$\frac{t_B}{q_B |B|} = \frac{s_x}{Q_x} \tag{2.14}$$

for any pair (x, B) with $q_B > 0$ and $x \in B$. If $\{x, y\}$ is an edge in the graph (K, E) then it follows that $s_x/Q_x = s_y/Q_y$. As the graph (K, E) is connected, the function s_x/Q_x is equal to a constant λ on K. We then have $t_B = \lambda |B| q_B$ and since $\sum_{B \subset K} t_B = 1$ it follows that $\lambda^{-1} = \mathbb{E}(|\mathcal{X}|)$.

A further application of (PEI) proves (2.8), which gives the uniqueness of P_E. The final assertions (2.9) and (2.10) follow easily from (2.8).

(b) Now suppose that P_A satisfies (2.8). Then $P_A(S = x) = \lambda \mathbb{E}(1_{\mathcal{X}}(x))$, and summing over $x \in K$ it follows that $P_A(S \in K) = 1$, so that $P_A(S = \partial) = 0$. If $F \in \mathcal{F}_W$ with $\mathbb{P}(F) = 0$ then $P_A(F \cap \{S = x\}) = 0$ by (2.8). Hence

$$P_A(F) = P_A(F \cap \{S \in K\}) + P_A(F \cap \{S = \partial\}) = 0,$$

so that part (a) of (PN) holds. Finally, if $x \in K$ then by (2.8)

$$P_A(S = x, x \notin \mathcal{X}) = \lambda \mathbb{P}(\{x \notin \mathcal{X}\} \cap \{x \in \mathcal{X}\}) = 0,$$

and summing over x it follows that $P_A(S \in \mathcal{X}) = 1$. Thus P_A satisfies (PN).

If $B \subset K$ and $x \in B$ then by (2.8)

$$P_A(\{\mathcal{X} = B\} \cap \{S = x\}) = \lambda \mathbb{P}(\mathcal{X} = B).$$

Summing over $x \in B$ we have $P_A(\mathcal{X} = B) = \lambda |B| \mathbb{P}(\mathcal{X} = B)$, and (PI) follows. Finally (PEI) follows easily from (2.8). $\square$

Remark 2.4. 1. (2.9) implies that the P_E-law of $|\mathcal{X}|$ is the size-biased distribution associated with the law of $|\mathcal{X}|$. This reweighting of the law of $|\mathcal{X}|$ is well known – see for example [7, p. 122].

2. If (K, E) is not connected the measure P_E given by (2.8) is no longer the unique measure which satisfies (PN), (PI), and (PEI). See the appendix to [4] (an earlier version of this paper) for more details on this. However this case does not seem to be important in applications: all the examples that the author has seen in the literature, when put into the framework of this paper, have (K, E) connected.

The possibility that there are no observers does not affect the probability P_E. Let

$$\widetilde{\mathbb{P}}(F) = \mathbb{P}(F | \mathcal{X} \neq \emptyset), \quad F \in \mathcal{F}_W, \tag{2.15}$$

and let $\widetilde{P}_E$ be the unique probability which satisfies (PN), (PI), (PEI) with respect to $\widetilde{\mathbb{P}}$.

Corollary 2.5. *Let $\widetilde{P}_E$ be as above. Then $\widetilde{P}_E = P_E$.*

Proof. Let $p = \mathbb{P}(\mathcal{X} \neq \emptyset)$; by hypothesis $p > 0$. Let $\widetilde{\lambda} = \widetilde{\mathbb{E}}(|\mathcal{X}|)$; then $\widetilde{\lambda} = \lambda/p$. By (2.8) we have for $F \in \mathcal{F}_W$

$$\begin{aligned}
\widetilde{P}_E(F \cap \{S = x\}) &= \widetilde{\lambda}\widetilde{\mathbb{P}}(F \cap \{x \in \mathcal{X}\}) \\
&= \lambda p^{-1} \mathbb{P}(F \cap \{x \in \mathcal{X}\} | \mathcal{X} \neq \emptyset) \\
&= \lambda \mathbb{P}(F \cap \{x \in \mathcal{X}\} \cap \{\mathcal{X} \neq \emptyset\}) = \lambda \mathbb{P}(F \cap \{x \in \mathcal{X}\}).
\end{aligned}$$

Hence $\widetilde{P}_E(F \cap \{S = x\}) = P_E(F \cap \{S = x\})$, and it follows that $\widetilde{P}_E = P_E$. $\square$

Let $\mathcal{X}'$ be a random subset of K with the property that $\mathcal{X}' \subset \mathcal{X}$. Suppose that an AO learns that it is located in $\mathcal{X}'$, i.e. that $S \in \mathcal{X}'$, and wishes to estimate the probability of an event F. One natural procedure is to look at $P_E(F | S \in \mathcal{X}')$, but another is to look at the model associated with the random process $\mathcal{X}'$, and then use the probability P_E' obtained from that model. More precisely, we wish to define P_E' on the space $\Omega = \Omega_O \times \Omega_A$ with $\Omega_A = K \cup \{\partial\}$. The graph structure E' associated with $\mathcal{X}'$ is given by taking $\{x, y\} \in E'$ if $\mathbb{P}(\{x, y\} \subset \mathcal{X}') > 0$. Even if the original graph (K, E) is connected, the graph (K, E') may not be connected, so that we cannot use Theorem 2.3 to define P_E'. However, we can still define P_E' by using (2.8):

$$P_E'(F \cap \{S = x\}) = \lambda' \mathbb{P}(F \cap \{x \in \mathcal{X}'\}) \text{ for } F \in \mathcal{F}_W, x \in K. \tag{2.16}$$

Here $\lambda' = 1/\mathbb{E}(X')$, where $X' = |\mathcal{X}'|$.

Proposition 2.6. *Let $\mathcal{X}' \subset \mathcal{X}$ with $\mathbb{P}(\mathcal{X}' = \emptyset) < 1$, let P_E be the probability defined by (2.8), and P'_E be given by (2.16). Then $P_E(S \in \mathcal{X}') > 0$ and for $G \in \mathcal{F}$,*

$$P_E(G|S \in \mathcal{X}') = P'_E(G). \tag{2.17}$$

Proof. It is enough to prove that for $F \in \mathcal{F}_W$, $x \in K$,

$$P_E(F \cap \{S = x\}|S \in \mathcal{X}') = P'_E(F \cap \{S = x\}). \tag{2.18}$$

By (2.16) the right hand side of (2.18) is $\lambda' \mathbb{P}(F \cap \{x \in \mathcal{X}'\})$. For the left side we have

$$\begin{aligned}
P_E(F \cap \{S = x\} \cap \{S \in \mathcal{X}'\}) &= P_E(F \cap \{x \in \mathcal{X}'\}|S = x)P_E(S = x) \\
&= \lambda \mathbb{P}(F \cap \{x \in \mathcal{X}'\}|x \in \mathcal{X})\mathbb{P}(x \in \mathcal{X}) \\
&= \lambda \mathbb{P}(F \cap \{x \in \mathcal{X}'\}).
\end{aligned}$$

Taking $F = \Omega$ and summing over x,

$$P_E(S \in \mathcal{X}') = \lambda \sum_{x \in K} \mathbb{P}(\{x \in \mathcal{X}'\}) = \frac{\mathbb{E}(X')}{\mathbb{E}(X)}.$$

Thus $P_E(S \in \mathcal{X}') > 0$ and

$$P_E(F \cap \{S = x\}|S \in \mathcal{X}') = \lambda' \mathbb{P}(F \cap \{x \in \mathcal{X}'\}),$$

proving (2.18). □

We will call (2.17) the *restriction property* of P_E.

The next result shows that it is only in rather special circumstances that we can find a probability which satisfies all four of the Principles given in Definition 2.1.

Corollary 2.7. *(a) Suppose that there exists a law P_A which satisfies (PN), (PI), (PEI) and (PP). Then there exists k such that $\mathbb{P}(|\mathcal{X}| = k|\mathcal{X} \neq \emptyset) = 1$.*
(b) If there exists k such that $\mathbb{P}(|\mathcal{X}| = k|\mathcal{X} \neq \emptyset) = 1$ then $P_L = P_E$.

Proof. (a) Let $\widetilde{\mathbb{P}}$ be as in (2.15). As P_A satisfies (PN), (PI) and (PEI) it is equal to P_E. Let $F \in \mathcal{F}_W$; using (PP) and (2.10) we have

$$\widetilde{\mathbb{P}}(F) = P_E(F) = \frac{\widetilde{\mathbb{E}}(1_F X)}{\widetilde{\mathbb{E}}(X)} \text{ for all } F \in \mathcal{F}_W. \tag{2.19}$$

Let k be such that $\mathbb{P}(X = k) > 0$. Setting $F = \{X = k\}$ and using (2.19) gives that $\widetilde{\mathbb{E}}(X) = k$, so that there is exactly one such k.
(b) is immediate from (2.10). □

If one wishes to keep (PP) then the most natural probability to use is given by

$$P_L(F \cap \{\mathcal{X} = B\} \cap \{S = x\}) = \begin{cases} |B|^{-1}\widetilde{\mathbb{P}}(F \cap \{\mathcal{X} = B\}) \text{ for } x \in B, B \subset K, \\ 0 \text{ for } x \notin B, B \subset K. \end{cases} \tag{2.20}$$

This satisfies (PN), (PI) and (PP). We can strengthen (PI) by requiring that, for $x \in B$, $B \subset K$, and $F \in \mathcal{F}_W$,

$$P_A(S = x | F \cap \{\mathcal{X} = B\}) = |B|^{-1} \quad \text{if } P_A(F \cap \{\mathcal{X} = B\}) > 0. \qquad (2.21)$$

Proposition 2.8. *The probability P_L is the unique probability on $(\Omega, \mathcal{F})$ which satisfies (PN), (PP) and (2.21).*

Proof. Let P_A satisfy (PN), (PP) and (2.21). Then if x, B and F are as above and $x \in B$,

$$P_A(\{S = x\} \cap F \cap \{\mathcal{X} = B\}) = |B|^{-1} \mathbb{P}(F \cap \{\mathcal{X} = B\} | X \geq 1). \qquad (2.22)$$

If $x \notin B$ then (PN) implies that the left side of (2.22) is zero. Thus P_A satisfies (2.20) and so equals P_L. $\qquad \square$

Easy examples show that P_L does not satisfy the restriction property. (See also Section 4.2.)

Example 2.9. *The original Sleeping Beauty problem.* Let $K = \{1, 2\}$, and $\Omega_O = \{0, 1\}$, with $\mathbb{P}_O(\{0\}) = \mathbb{P}_O(\{1\}) = \frac{1}{2}$, and let $\mathcal{X}^O(0) = \{1\}$, $\mathcal{X}^O(1) = K$. Since $\mathbb{P}_O(\mathcal{X}^O = \emptyset) = 0$ we can dispense with the extra point ∂, and take $\Omega_A = K$, so that $\Omega = \Omega_O \times K$. As usual we write $\mathbb{P}$ and $\mathcal{X}$ for the extensions of $\mathbb{P}_O$ and $\mathcal{X}^O$ to Ω.

Let P_A be a probability on Ω which satisfies (PN), so that $P_A(\{(0, 2)\}) = 0$. Set

$$a = P_A(\{(0, 1)\}), \ b = P_A(\{(1, 1)\}), \ 1 - a - b = P_A(\{(1, 2)\}). \qquad (2.23)$$

We define the events Heads $= \{(0, 1), (0, 2)\}$, Mon $= \{(0, 1), (1, 1)\}$, and let Tails and Tue be the complements of Heads and Mon respectively. Then

$$P_A(\text{Heads} | \text{Mon}) = \frac{a}{a + b}, \ \text{and } P_A(\text{Mon} | \text{Tails}) = \frac{b}{1 - a}. \qquad (2.24)$$

(PEI) states that $P_A(\text{Heads} | \text{Mon}) = \frac{1}{2}$, which implies that $a = b$. (PI) gives $P_A(\text{Mon} | \text{Tails}) = \frac{1}{2}$, so that $a + 2b = 1$. Finally (PP) gives $a = \frac{1}{2}$. So each of these principles imposes a different linear relation between a and b, and thus the three conditions are independent.

Taking (PI) and (PEI) together we obtain $a = b = \frac{1}{3}$, and so we have $P_E(\text{Heads}) = \frac{1}{3}$. If (PI) and (PP) hold then $a = \frac{1}{2}$, $b = \frac{1}{4}$, and $P_A = P_L$. We have $P_L(\text{Heads} | \text{Mon}) = \frac{2}{3}$, so this measure does not satisfy (PEI).

The interpretation of this conditional probability is a difficulty for Halfers. Even though the coin toss can take place on Monday afternoon, on the standard Halfer view SB at midday on Monday should hold that this fair coin toss has probability $\frac{2}{3}$ of coming down Heads. Lewis [21] did adopt this position, but nearly all subsequent authors have found it unsatisfactory. Double Halfers hold that both the probability of Heads, and the probability of Heads given that it is Monday are $\frac{1}{2}$. If we use (2.23) and (2.24) then we obtain $a = \frac{1}{2}$ and $a/(a + b) = \frac{1}{2}$, so that $a = b = \frac{1}{2}$ and $P_A(\text{Tue}) = 0$. Thus Double Halfers would appear to be committed to the view that on waking SB is certain that it is Monday: [17] does adopt that position. But most Double Halfers wish to allow $P_A(\text{Tue}) > 0$, and

some wish to keep (PI) and hold that $P_A(\text{Mon}|\text{Tails}) = \frac{1}{2}$. (Thus they are really 'Triple Halfers'). It is clear from the calculations above that this position cannot be described by the usual axioms of probability. Double Halfers are well aware of this difficulty, and the literature contains a number of strategies for dealing with it. As one might expect, attempts such as those in [15, 22, 6] to modify the standard definition of conditional probability leads to operations with mathematically undesirable properties. See [4, Sections 4, A.2] for more details.

Example 2.10. *Four Beauties.* This elegant example is due to Pittard [24]. The experiment lasts for one night only, with four participants ('Beauties') A, B, C, D. One of the four will be chosen at random to be the 'Waker'. The four participants all go to sleep. Three times in the night a pair of participants are woken at the same time for a period (say 15 minutes) during which they can converse. They then go back to sleep, and as usual are given a memory erasing drug so that they forget the awakening. The Waker is woken all three times, and paired with each other participant exactly once.

We analyse this model by describing the awakenings of a particular participant, A say. We take $\Omega_O = \{A, B, C, D\}$, and let $K = \Omega_A = \{B, C, D\}$ be the set of possible partners of A during an awakening. We write $\omega = (v, x)$ for points in Ω, and with ω of this form we define $W(\omega) = v$, $S(\omega) = x \in K$. So W gives the Waker, and S describes A's partner for the selected awakening. $\mathbb{P}$ is the probability which makes W uniform on Ω_O, and $\mathcal{X}$ is defined by $\mathcal{X} = \{K\}$ if $W = A$, and $\mathcal{X} = \{W\}$ if $W \neq A$.

As $\mathbb{P}(\mathcal{X} = K) > 0$ the model satisfies the hypotheses of Theorem 2.3. We have $Q_x = \frac{1}{2}$ for $x = B, C, D$, and $\mathbb{E}|\mathcal{X}| = \frac{3}{2}$, so that $\lambda = \frac{2}{3}$. Thus $P_E(S = x) = \lambda Q_x = \frac{1}{3}$ for $x = B, C, D$. Suppose that A wakes and sees that the other person woken is B. Then, using (PEI) the probability that A is the Waker is

$$P_E(W = A|S = B) = \mathbb{P}(W = A|B \in \mathcal{X}) = \frac{\mathbb{P}(W = A)}{\mathbb{P}(B \in \mathcal{X})} = \frac{1}{2}.$$

This answer accords with common sense – A and B are in the same situation and by symmetry each regards herself as equally likely to be the Waker.

On the other hand, if we use P_L then $P_L(W = x) = \mathbb{P}(W = x) = \frac{1}{4}$, for each $x \in K$. By (PI) $P_L(S = B|W = A) = \frac{1}{3}$, and so $P_L(S = B, W = A) = \frac{1}{12}$. By symmetry $P_L(S = B) = \frac{1}{3}$, and hence

$$P_L(W = A|S = B) = \frac{P_L(S = B, W = A)}{P_L(S = B)} = \frac{1}{4}.$$

Thus if A and B are woken together then as A has credence $\frac{1}{4}$ that she is the Waker, she must have credence $\frac{3}{4}$ that B is the Waker. B, of course, has the opposite credences. A and B appear to have the same information, and can share all they that know, but still have different credences for an objective event. Although this example would seem to pose a severe challenge for Halfers, [24] does nevertheless adhere to the Halfer position.

Example 2.11. *The Presumptuous Philosopher.* (For an explanation of the terminology see [7].) Here we take $1 \leq N \ll M$, and set $K = \{1, \ldots, M\}$. Let W be a $\text{Ber}(\frac{1}{2})$ r.v and

let $\mathcal{X} = K$ if $W = 1$ and $\mathcal{X} = \{1, \ldots, N\}$ if $W = 0$. Using (2.10) and (2.5) we have

$$P_E(W = 0) = \frac{N}{M + N}, \quad P_L(W = 0) = \tfrac{1}{2}. \tag{2.25}$$

If we take $|\mathcal{X}|$ to be the size of the universe, then an AO which uses P_E strongly favours the larger universe. If say $N = 10^{25}$ and $M = 10^N$ then the P_E probability that $W = 0$ is so small that any observation to the contrary is overwhelmingly likely to be due to experimental error.

Bostrom [7] rejects the Thirder solution on the basis of this example. But, as we have seen, the Halfer position also runs into difficulties, which seem to me to be more immediate and fundamental than the problem here.

2.2 Future information

In the introduction we noted that for the standard SB problem, one argument for (1.2) is that as the coin can be tossed on Monday evening, on Monday morning it is a future event, and so the AO and CO should assign it the same probability.

Our model does not have a natural time structure; to implement it for a SB type experiment we needed to choose a bijection between the set of states K and $\{1, \ldots, M\}$. However, an important subclass of models does have such a structure; these include the original SB problem and the Doomsday problem (see [10]). For these models we have

$$\mathcal{X} = \{k, 1 \le k \le X\}, \tag{2.26}$$

and the interpretation is that on days $1, \ldots, X$ the AO is woken in a cell and then (as usual) has its memories erased at the end of each day. There are no awakenings after day X, at least until the experiment is concluded. We define q_n, Q_n for $0 \le n \le M$ by (2.6), and assume that M is chosen so that $q_M = \mathbb{P}(X = M) > 0$.

We now show that, subject to a mild restriction on the probability space, for models which satisfy (2.26) (PEI) is equivalent to a condition which states that the AO and CO have the same views on the probability of future events. In essence this argument proceeds by pushing every randomisation not needed to determine whether or not $X \ge n$ into the post-n future.

To avoid measure-theoretic technicalities, we make the assumption that the objective probability space $(\Omega_O, \mathcal{F}_O, \mathbb{P}_O)$ has the following structure.

Assumption 2.12. (1) There exist independent Bernoulli random variables $(V_n^O, 0 \le n \le M)$ and U^O, such that U^O is uniform on $[0, 1]$ and

$$\mathbb{P}_O(V_n^O = 1) = 1 - \mathbb{P}_O(V_n^O = 0) = q_n/Q_n, \quad 0 \le n \le M. \tag{2.27}$$

(2) $\mathcal{F}_O = \sigma(U^O, V_n^O, 0 \le n \le M)$.
(3) The r.v X^O satisfies
$$X^O = \min\{n : V_n^O = 1\}. \tag{2.28}$$

13

Remark 2.13. This assumption is much less restrictive than it might appear. A *standard probability space* is a space which is isomorphic to the unit interval with Lebesgue measure. In [19, p. 61] Itô remarks that "all probability spaces appearing in practical applications are standard", which can be restated as saying that any stochastic process appearing in practical applications can be defined as a measurable function of a single r.v. U with uniform distribution on $[0,1]$. In particular, given a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ carrying a random variable X' taking values in $\{0, \ldots, M\}$ and a stochastic process $Z' = (\widetilde{Z}_1, \ldots, \widetilde{Z}_M)$ taking values in a complete separable metric space $\mathcal{K}$, then there exists a probability space $(\Omega_O, \mathcal{F}_O, \mathbb{P}_O)$ satisfying Assumption 2.12 carrying X^O, Z^O such that $(X^O, Z_1^O \ldots, Z_M^O)$ has the same distribution as $(X', Z_1', \ldots, Z_M')$. See [19] for more details.

We write (V_n), U for the extensions of $(V_n^O, n \in \mathbb{Z}_+)$ and U^O to the space $(\Omega, \mathcal{F}, \mathbb{P})$. Assumption 2.12 implies that $\mathcal{F}_W = \sigma(U, V_n, 0 \leq n \leq M)$. From (2.28) we have

$$\{X \geq n\} = \{V_k = 0, 0 \leq k \leq n - 1\}.$$

We regard the randomization V_k as being made at the end of day k, and the randomization U being made at the end of day M.

Definition 2.14. Principle of no future information (PNFI). Let P_A be a probability on $(\Omega, \mathcal{F})$ which satisfies (PN). P_A satisfies (PNFI) if whenever $0 \leq n \leq M$, and $G \in \sigma(V_n, \ldots, V_M, U)$, then

$$P_A(G|S = n) = \mathbb{P}(G). \tag{2.29}$$

We assume that the CO is able to observe the random variables V_k as they occur. The value n is fixed at the start of the experiment. On day n the CO knows the values of $V_0, \ldots, V_{n-1}$, and thus whether or not $X \geq n$. If $X < n$ the CO takes no action. If $V_0 = \cdots = V_{n-1} = 0$, so that $X \geq n$, then as G is independent of $V_0, \ldots, V_{n-1}$ we have $\mathbb{P}(G|X \geq n) = \mathbb{P}(G)$, and the right hand side of (2.29) gives the CO's conditional probability that G will occur. If $X \geq n$ then the CO telephones cell n and talks to the AO there. The AO thus knows that the event $\{S = n\}$ holds, and hence also knows that $\{X \geq n\}$. The event G lies in the future for the AO also, and the left side of (2.29) is the AO's conditional probability that it will occur. (PNFI) then states that the two observers agree on the probability of this future event.

Theorem 2.15. *Suppose that Assumption 2.12 holds. Then (PEI) and (PNFI) are equivalent for probabilities P_A which satisfy (PN).*

Proof. Suppose that (PEI) holds, and let $G \in \sigma(V_n, \ldots, V_M, U)$. Then $G \in \mathcal{F}_W$ and is independent of $\{X \geq n\}$. Hence (2.29) follows immediately from (PEI).

Now suppose that (PNFI) holds. Let $(e_0, \ldots, e_{n-1}) \in \{0, 1\}^n$, $H = \{V_k = e_k, 0 \leq k \leq n - 1\}$, and $H_1 = \{V_k = 0, 0 \leq k \leq n - 1\}$. Let $G \in \sigma(V_n, \ldots, V_M, U)$. If the AO knows that $\{S = n\}$ then it knows that H_1 has occurred, and so $P_A(G \cap H|S = n) = 0$ if $H \neq H_1$, while

$$P_A(G \cap H_1|S = n) = P_A(G|S = n). \tag{2.30}$$

14

Similarly we have $\mathbb{P}(G \cap H | X \geq n) = 0$ if $H \neq H_1$, and $\mathbb{P}(G \cap H_1 | X \geq n) = \mathbb{P}(G)$. We therefore deduce that

$$P_A(G \cap H | S = n) = \mathbb{P}(G \cap H | X \geq n). \tag{2.31}$$

Using the π-λ Lemma (see [11, p. 447]) and property (2) of Assumption 2.12 it follows that (PEI) holds. $\qquad\square$

3 Observations by Anthropic Observers

3.1 Technicolour Beauty

Recall that $Z = (Z_x, x \in K)$ is an $\mathcal{F}_W$ measurable process taking values in a finite set $\mathcal{Z}$. We call Z_x the *colour* of cell C_x and assume that the AO is able to observe this value. This model is often called *Technicolour Beauty* – see [30, 9]. For $y \in \mathcal{Z}$ let

$$L_y = |\{x \in \mathcal{X} : Z_x = y\}| = \sum_{x \in K} 1_{(x \in \mathcal{X})} 1_{(Z_x = y)} \tag{3.1}$$

be the number of AO which see the colour y. Set

$$G_y = \{L_y \geq 1\}; \quad H_y = \{S \in \mathcal{X}, Z_S = y\}. \tag{3.2}$$

Thus G_y is the event that some AO sees y, while H_y is the event that the AO selected by S sees the value y. We have $G_y \in \mathcal{F}_W$, and in general $H_y \notin \mathcal{F}_W$. We also have $H_y \subset G_y$.

Conditioning on the observation H_y does not affect the size biasing obtained by using P_E.

Lemma 3.1. *For $F \in \mathcal{F}_W$ and $y \in \mathcal{Z}$,*

$$P_E(F \cap \{Z_S = y\}) = \lambda \mathbb{E}(1_F L_y). \tag{3.3}$$

Hence

$$P_E(F | Z_S = y) = \frac{\mathbb{E}(1_F L_y)}{\mathbb{E}(L_y)}. \tag{3.4}$$

Proof. We have using (2.9) and (PEI)

$$
\begin{aligned}
P_E(F \cap H_y) &= \sum_{x \in K} P_E(F \cap \{Z_x = y\} | S = x) P_E(S = x) \\
&= \sum_{x \in K} \mathbb{P}(F \cap \{Z_x = y\} | x \in \mathcal{X}) \lambda Q_x \\
&= \lambda \sum_{x \in K} \mathbb{P}(F \cap \{Z_x = y\} \cap \{x \in \mathcal{X}\}) = \lambda \mathbb{E}(L_y 1_F).
\end{aligned}
$$

Setting $F = \Omega$ gives $P_E(H_y) = \lambda \mathbb{E}(L_y)$, and (3.4) follows. $\qquad\square$

We compare this with what happens when we use P_L.

Lemma 3.2. *For $F \in \mathcal{F}_W$ and $y \in \mathcal{Z}$,*

$$P_L(F \cap \{Z_S = y\}) = \mathbb{E}\Big(\frac{1_F L_y}{X}\Big|X \geq 1\Big), \tag{3.5}$$

and

$$P_L(F|Z_S = y) = \frac{\mathbb{E}\big(1_F L_y X^{-1}\big|X \geq 1\big)}{\mathbb{E}\big(L_y X^{-1}\big|X \geq 1\big)}. \tag{3.6}$$

Proof. We have

$$
\begin{aligned}
P_L(F \cap \{Z_S = y\}) &= \sum_{B \neq \emptyset} \sum_{x \in B} P_L(F \cap \{\mathcal{X} = B\} \cap \{Z_x = y\} \cap \{S = x\}) \\
&= \sum_{B \neq \emptyset} \sum_{x \in B} |B|^{-1} \mathbb{P}(F \cap \{\mathcal{X} = B\} \cap \{Z_x = y\}|\mathcal{X} \neq \emptyset) \\
&= \sum_{B \neq \emptyset} |B|^{-1} \mathbb{E}(1_F 1_{\{\mathcal{X}=B\}} L_y\}|\mathcal{X} \neq \emptyset) = \mathbb{E}\big(1_F L_y X^{-1}\big|X > 0\big).
\end{aligned}
$$

Dividing this expression with the one for $P_L(Z_S = y)$ gives (3.6). $\qquad \square$

Remark 3.3. The presence of the term X^{-1} in (3.5) means one seldom obtains simple expressions for P_L probabilities.

Definition 3.4. We say that an auxiliary process Z is *injective* if $\mathbb{P}(L_y \leq 1) = 1$ for each $y \in \mathcal{Z}$.

Lemma 3.5. *Suppose that Z is injective. Then*

$$P_E(F|H_y) = \mathbb{P}(F|G_y) \quad \text{for all } F \in \mathcal{F}_W. \tag{3.7}$$

Proof. As Z is injective $L_y = 1_{G_y}$ and (3.7) follows from (3.4). $\qquad \square$

3.2 Improper conditioning events and symmetry breaking

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and F, G, H be events with $H \subset G$ and $\mathbb{P}(G \backslash H) > 0$. The conditional probability of F given H is

$$\mathbb{P}(F|H) = \frac{\mathbb{P}(F \cap H)}{\mathbb{P}(H)}, \tag{3.8}$$

which is often informally described as being "the probability that F will occur given that the observer knows that H occurs". However if H occurs then G also occurs, so that this informal description does not explain why one could not use $\mathbb{P}(F \cap G)/\mathbb{P}(G)$ instead. This procedure is so obviously incorrect that few if any introductions to conditional probability

16

consider this possibility. (Some works are quite careful in their explanations of conditional probability and avoid the informal description above – see for example [13, p. 115].) In terms of Kolmogorov's definition of conditional expectation, one is interested in

$$\mathbb{E}(1_F|\sigma(1_H)) = 1_H\mathbb{P}(F|H) + 1_{H^c}\mathbb{P}(F|H^c). \tag{3.9}$$

If one replaces H by G in (3.9) then the resulting random variable is not measurable with respect to the agent's observation, which is of 1_H. I will call conditioning events of the kind that G is here *improper conditioning events*. Easy examples show that replacing H by an improper conditioning event G can lead to a wildly incorrect answer.

In the standard SB experiment she has identical awakenings on each day, but one can break this symmetry by introducing an auxiliary process Z as in Section 3.1 and allowing SB to observe the value of Z in her cell. This suggests that one might be able to describe SB's experience by using the objective probability $\mathbb{P}$. One paper which adopts this approach is [9]. Using the notation introduced above [9] gives the definition

$$P_{Z,y}(F) = \mathbb{P}(F|G_y) \text{ for } F \in \mathcal{F}_W. \tag{3.10}$$

The interpretation in [9] is that if the AO observes that the process Z in its cell takes the value y (i.e. the event H_y) then it knows that G_y occurs, and so (3.10) gives the AO's credence that F occurs. While of course the definition (3.10) makes mathematical sense, since G_y is an improper conditioning event it is questionable whether the law $P_{Z,y}$ correctly models the experience of the AO. If Z is injective then $P_{Z,y} = P_E(\cdot|H_y)$ by Lemma 3.5, but as [9] allows for more general Z it is not surprising that by varying the process Z they are able obtain a range of values for $P_{Z,y}(\text{Heads})$. The framework of [9] does not allow one to combine the laws $P_{Z,y}$ into a single probability P_Z which gives SB's credences before she makes an observation. As the authors of [9] remark, they cannot use the law of total probability since the events G_y are not disjoint.

The next section looks at an example of the use of improper conditioning events in the physics literature.

3.3   A model of Hartle and Srednicki

The paper [16] questions arguments, such as those in [23], which assume that we are 'typical' observers. To help clarify the issues they introduce a model of a universe which has N successive cycles, and assume that observers in cycle i have no knowledge of the state of the universe in any other cycle. Each cycle has one of two global properties: red (R) or blue (B). In each cycle the probability that an 'observing system' (such as 'us') exists is $p \in (0,1)$, and is independent of whether the global state is red or blue.

There are two theories of the universe, which are equally likely. The first, AR is that all the cycles are red, and the second, SR, is that exactly M of the cycles are red, and $N - M$ are blue. Here $1 \leq M \leq N - 1$. Hartle and Srednicki do not specify which cycles are red or blue in the SR case; for simplicity we will assume that under SR each possible arrangement of M red and $N - M$ blue is equally likely. For each theory we

write $N_R(T)$ and $N_B(T)$ for the number of red and blue cycles according to the theory T. (Thus $N_R(AR) = N$ and $N_R(SR) = M$.)

The authors of [16] wish to calculate the probability of SR given that 'we' (situated in one of the cycles) observe that our cycle is red. Writing H_{ER} for the event "we exist and observe red", they claim that

$$P(H_{ER}|T) = 1 - (1-p)^{N_R(T)}, \tag{3.11}$$

and then use Bayes' formula to compute $P(T|H_{ER})$. Writing $f(p,n) = 1 - (1-p)^n$, they obtain

$$P(SR|H_{ER}) = \frac{f(p,M)}{f(p,N) + f(p,N)}. \tag{3.12}$$

This conclusion is counterintuitive. If for example $M = 1$, $N = 1000$ and p is very close to 1, then (3.12) gives that $P(SR|H_{ER}) \approx \frac{1}{2}$: our observation of red has given very little information even though, on the hypothesis SR, red universes are very rare.

However (3.11), and hence (3.12), are incorrect. To see that (3.11) cannot be true, let H_E be the event "we exist", and note that by red/blue symmetry we should also have $P(H_{EB}|T) = 1 - (1-p)^{N_B(T)}$. Since $P(H_{ER}|T) + P(H_{EB}|T) = P(H_E|T) \leq 1$, we obtain

$$2 - (1-p)^{N_R(T)} - (1-p)^{N_B(T)} \leq 1, \tag{3.13}$$

and this fails for many values of p, $N_R(T)$ and $N_B(T)$.

The error which led to (3.11) is confusion between the events $H_{ER} = \{$we exist and observe red$\}$ and $G_{ER} = \{$there is a cycle in which an observer exists and observes red$\}$. Since $H_{ER} \subset G_{ER}$ if we observe that H_{ER} occurs then we also observe that G_{ER} occurs. The events G_{ER} and G_{EB} are not disjoint, which is why the sum on the left hand side of (3.13) can be greater than 1. I suspect that the cause of this error is that the authors of [16] were (rightly) uneasy about the event H_{ER}, and wished to find a suitable objective event which they could work with. In the terminology of the previous section, G_{ER} is an improper conditioning event.

We now apply our formalism to this problem. We define

$$\Omega_O = \{AR, SR\} \times (\{R, B\} \times \{0,1\})^N.$$

An element $\omega \in \Omega_O$ is a sequence

$$\omega = (y, a_1, n_1, a_2, n_2, \ldots, a_N, n_N), \tag{3.14}$$

where $y \in \{AR, SR\}$, and for $1 \leq i \leq N$ we have $a_i \in \{R, B\}$ and $n_i \in \{0,1\}$. Thus a_j gives the state (red or blue) of cycle j, and $n_j = 1$ if there are observing systems in cycle j, and $n_j = 0$ otherwise. We define $\Omega_A = \{\partial, 1, \ldots, N\}$. With ω given by (3.14) we define random variables on $\Omega = \Omega_O \times \Omega_A$ by

$$U(\omega, k) = y, \quad S(\omega, k) = k, \quad Z_i(\omega, k) = a_i, \quad \xi_i(\omega, k) = n_i, \text{ for } 1 \leq i \leq N. \tag{3.15}$$

The random set of occupied universes is $\mathcal{X} = \{i : \xi_i = 1\}$. Under the objective probability law $\mathbb{P}$ the random variables U, $(\xi_j, 1 \leq j \leq N)$ are independent with $\mathbb{P}(U = AR) = \mathbb{P}(U =$

18

$SR) = \frac{1}{2}$, and $\mathbb{P}(\xi_j = 1) = 1 - \mathbb{P}(\xi_j = 0) = p$. If $U = AR$ then $Z_i = R$ for all i, while if $U = SR$ then exactly M of the Z_i are equal to R, and the remainder are equal to B, with each ordering of reds and blues being equally likely. We assume that (PN), (PI) and (PEI) hold and write P_E for the probability which satisfies these conditions. Note that $H_{ER} = \{Z_S = R\}$. For $y \in \{R, B\}$ let

$$L_y = |\{j : \xi_j = 1, Z_j = y\}| = |\mathcal{X} \cap \{i : Z_i = y\}|.$$

By Lemma 3.1, for $F \in \mathcal{F}_W$,

$$P_E(F|Z_S = y) = \frac{\mathbb{E}(1_F L_y)}{\mathbb{E}(L_y)}.$$

Taking $F = \{U = SR\}$ we have $\mathbb{E}(1_F L_y) = \frac{1}{2}\mathbb{E}(L_y|U = SR) = \frac{1}{2}pM$, and similarly $\mathbb{E}(L_y) = \frac{1}{2}p(M + N)$. Hence

$$P_E(U = SR|Z_S = R) = \frac{M}{M + N}. \tag{3.16}$$

In the case noted above, when $M = 1$, $N = 1000$, we obtain the natural and unsurprising result that $P_E(U = SR|Z_S = R) = 1/1001$.

By Corollary 2.7 the laws P_E and P_L for this model are not the same, but using Lemma 3.2 we find that one does have $P_L(U = SR|Z_S = R) = M/(M + N)$.

4   Examples – Life, the Universe, and Everything

4.1   Probability of life

Look at M sites (e.g. planets, each associated with a separate star) so that $K = \{1, \ldots, M\}$. Let V be a r.v. on $(0, 1)$ with a continuous density function $f(v)$, and conditional on V let $(\xi_x, x \in K)$ be independent $Ber(V)$ r.v. We set $\lambda = 1/\mathbb{E}(V)$. We define $\mathcal{X} = \{x : \xi_x = 1\}$, and $X = |\mathcal{X}|$. We interpret V as the probability life develops on a planet, and say that the planet x has life if $\xi_x = 1$. For simplicity we assume that life is the same as the existence of observers, and neglect any consideration of how long this might take. (See [31, 28] for more realistic models.)

We have $\mathbb{E}(X|V) = MV$, and the probability that no observers exist is

$$p_M = \mathbb{P}(X = 0) = \int_0^1 (1 - v)^M f(v)dv = \mathbb{E}(1 - V)^M.$$

We now look at the density of V given a variety of possible observations.

(0) An external observer with no observation. In this case the observer just uses the prior density $f(v)$.

(1) An external observer. Suppose that this observer looks at site x and observes $\xi_x = 1$. Standard Bayesian computations give

$$\mathbb{P}(V \leq t|\xi_x = 1) = \frac{\mathbb{P}(V \leq t, \xi_x = 1)}{\mathbb{P}(\xi_x = 1)} = \frac{\int_0^t vf(v)dv}{\int_0^1 vf(v)dv} = \lambda \int_0^t vf(v)dv.$$

So the posterior density of V conditional on $\xi_x = 1$ is

$$f_1(v) = \lambda v f(v). \tag{4.1}$$

(2) An AO using P_E. This AO is on planet S, and observes $\xi_S = 1$. By Theorem 2.3

$$P_E(V \le t) = \frac{\mathbb{E}(X 1_{(V \le t)})}{\mathbb{E}(X)} = \frac{\mathbb{E}(V 1_{(V \le t)})}{\mathbb{E}(V)} = \lambda \int_0^t v f(v) dv.$$

So the AO's 'posterior' density is the same as in (1).

(3) An AO using P_L. By (2.5)

$$P_L(V \le t) = \mathbb{P}(V \le t | \mathcal{X} \neq \emptyset) = (1 - p_M)^{-1} \int_0^t \left(1 - (1-v)^M \right) f(v) dv.$$

In the $M \to \infty$ limit this gives the density $f_{3,\infty}(v) = f(v)$. So in the large M limit for an AO using P_L the observation $\xi_S = 1$ gives no information.

Remark 4.1. 1. The case $M = 1$ can be interpreted as a single universe, where the constants of physics are chosen according to some random distribution and giving rise to observers with probability V. By Corollary 2.7 we have $P_L = P_E$, and the AO which uses either of these will adopt the posterior density f_1 given by (4.1). So we find, somewhat remarkably, that an AO gains information from its existence, even though it is not possible for it to observe that it does not exist.

2. In the standard SB problem the law P_L looks initially plausible (see [27]), but is found to give unreasonable results when one conditions on a suitable event F. We can loosely call these *challenge events* for P_L. In this example there do not seem to be any natural challenge events for P_L, and so, looking at this example on its own, it does not seem easy to decide which of P_E or P_L one should adopt.

4.2 A two zone universe

We modify the example above by considering a universe with two zones. Let $K = \{-M, \ldots, -1\} \cup \{1, \ldots, M\}$, let p_0, p_1, p_2 be strictly positive, $W \sim \text{Ber}(\frac{1}{2})$, let $V_i = p_0$ for $x < 0$ and $V_x = p_1 1_{(W=0)} + p_2 1_{(W=1)}$ for $x > 0$. Conditional on W, let $(\eta_x, x \in K)$ be independent $\text{Ber}(V_x)$ r.v. We set $\mathcal{X} = \{x : \eta_x = 1\}$. We assume that M is large enough so that $\mathbb{P}(X = 0)$ can be neglected. Define $K_0 = K \cap \mathbb{Z}_-$, $K_1 = K \cap \mathbb{Z}_+$, $\mathcal{X}_j = \mathcal{X} \cap K_j$, and $X_j = |\mathcal{X}_j|$, $X = X_0 + X_1$.

We have

$$P_L(V_1 = p_1) = P_L(W = 0) = \mathbb{P}(W = 0 | X \neq 0) \simeq \mathbb{P}(W = 0) = \tfrac{1}{2}. \tag{4.2}$$

Now let $Z_i = 1_{(i \ge 1)}$. By Lemma 3.2, and noting that $L_0(Z) = |\{i \in \mathcal{X} : Z_i = 0\}| = X_0$,

$$P_L(V_1 = p_1 | Z_S = 0) = \frac{\mathbb{E}(X_0 X^{-1} 1_{(W=0)} | X > 0)}{\mathbb{E}(X_0 X^{-1} | X > 0)}.$$

20

For large M we have $X_0/M \simeq p_0$ and $X_1/M \simeq V_1/M$. Hence

$$\mathbb{E}(L_0 X^{-1} 1_{(W=0)}|X > 0) \simeq \tfrac{1}{2} p_0/(p_0 + p_1),$$

and

$$P_L(V_1 = p_1|Z_S = 0) \simeq \frac{p_0/(p_0 + p_1)}{p_0/(p_0 + p_1) + p_0/(p_0 + p_2)} = \frac{p_0 + p_2}{2p_0 + p_1 + p_2}.$$

If $p_0, p_1 \ll p_2$ then this probability is close to 1. Thus using the terminology of the previous example, $\{Z_S = 0\} = \{S \in K_0\}$ is a challenge event for P_L. Initially an observer (in some planet S) has the view that $P_L(V_1 = p_1) \simeq \frac{1}{2}$. If further observation reveals that S is in the negative zone K_0, then the observer does not seem to have gained any information on conditions in the positive zone K_1, but nevertheless now assigns a probability close to 1 to the event $\{V_1 = p_1\}$.

The restriction property given in Proposition 2.6 means that this does not occur for an observer using P_E.

4.3 Anthropic estimates of the cosmological constant

This final example is based on [5, 29], which look at a multiverse, containing a random number of universes, each with a random cosmological constant $\Lambda \in [a, b]$. The proportion of mass in a universe which turns into stars varies according to Λ, and assuming that the number of observers is proportional to the number of stars these papers ask if the value of Λ which we observe is likely.

We simplify the model in [5], in particular by restricting to a finite number of AO. Let $n \geq 1$, $\Theta = \{i/n, 0 \leq i \leq n-1\}$, $\kappa \in (0, n)$ and $p_n = \kappa/n$. Let $m(t)$ be a continuous function on $[0, 1]$, and let $M_n = \sum_{\theta \in \Theta} n^{-1} m(\theta)$. We assume that $M_n > 0$ for all n. We have $M_n \to M$ where

$$M = \int_0^1 m(t) dt.$$

Let $(U_\theta^O, \theta \in \Theta)$ be independent $\text{Ber}(p_n)$ r.v. (For a justification of this 'flat prior' in the cosmological context see [5, p. 3737].) We interpret $U_\theta^O = 1$ as the event that there exists a universe with cosmological constant θ. Let $n_0 \in \mathbb{N}$ and let $(Y_\theta^O, \theta \in \Theta)$ be independent r.v. with values in $\{0, 1, \ldots, n_0\}$ such that

$$\mathbb{E}^O(Y_\theta^O) = m(\theta).$$

If $U_\theta^O = 1$ then Y_θ^O is the number of AO in universe θ. Let $K = \Theta \times \{0, \ldots, n_0\}$; the set of AO is

$$\mathcal{X}^O = \{(\theta, k) : U_\theta^O = 1, 1 \leq k \leq Y_\theta^O\} \subset K.$$

Let $(\Omega_O, \mathcal{F}_O, \mathbb{P}_O)$ be the probability space carrying $(U_\theta^O, Y_\theta^O, \theta \in \Theta)$. As in Section 2 we construct the space $(\Omega, \mathcal{F}, \mathcal{F}_W, \mathbb{P})$, and write $\mathcal{X}, U_\theta, Y_\theta$ for the extensions of $\mathcal{X}^O, U_\theta^O, Y_\theta^O$ to this space. Define $Z : K \to \mathbb{R}$ by setting $Z((\theta, k)) = \theta$; the observation of an AO (sited

at location $S \in \mathcal{X}$) is that it is in a universe with cosmological constant Z_S. We define L_θ by (3.1), so that $L_\theta = U_\theta Y_\theta$. Thus $\mathbb{E}(L_\theta) = p_n m(\theta)$, and

$$\mathbb{E}(|\mathcal{X}|) = \sum_{\theta \in \Theta} p_n m(\theta) = \kappa M_n.$$

Set

$$\pi_n^{(O)}(\theta) = n^{-1}, \quad \pi_n^{(E)}(\theta) = P_E(Z_S = \theta), \quad \pi_n^{(L)}(\theta) = P_L(Z_S = \theta). \tag{4.3}$$

Thus $\pi_n^{(O)}$ gives the probability that a (uniformly randomly selected) universe has constant θ, while $\pi_n^{(E)}$ and $\pi_n^{(L)}$ give the distribution of θ for AO which use P_E or P_L. (One is tempted to call $\pi_n^{(O)}$ the prior, and $\pi_n^{(E)}$ and $\pi_n^{(L)}$ the posterior distributions of θ, but this is not exactly the standard Bayesian setup.)

By Lemma 3.1

$$\pi_n^{(E)}(\theta) = \frac{\mathbb{E}(L_\theta)}{\mathbb{E}(|\mathcal{X}|)} = \frac{m(\theta)}{n M_n}. \tag{4.4}$$

Thus $\pi_n^{(E)}$ is weighted according to $m(\theta)$. Note that this expression does not depend on κ. This agrees with the calculation in [5] – see equation (14).

An interesting point is that if $F \in \sigma(U_t, Y_t, t \neq \theta)$ then as 1_F and L_θ are independent by Lemma 3.1

$$P_E(F \cap \{Z_S = \theta\}) = \mathbb{P}(F)P_E(Z_S = \theta).$$

So the AO has a flat distribution for the constants of every universe other than its own.

If instead the AO uses P_L then by Lemma 3.2,

$$P_L(Z_S = \theta) = \mathbb{E}\left(L_\theta X^{-1} \Big| X > 0\right) = \mathbb{P}(X > 0)^{-1} \mathbb{E}(L_\theta X^{-1} 1_{(X>0)}). \tag{4.5}$$

We now look at (4.5) when κ is either small, or very large. To simplify calculations we assume further that $\mathbb{P}(Y_\theta = 0) = 0$ for all θ, and will take n to be large. Since we obtain different answers, it follows that $P_L(Z_S = \theta)$ does depend on κ. We write $X = |\mathcal{X}|$, $X_0 = X - U_\theta Y_\theta$.

If κ is small then both $\mathbb{P}(X > 0)$ and $\mathbb{P}(X_0 > 0)$ are close to κ. We have, $\mathbb{P}$-a.s., $L_\theta X^{-1} 1_{(X>0)} = U_\theta Y_\theta / (X_0 + Y_\theta)$, so that

$$U_\theta 1_{(X_0=0)} \leq \frac{L_\theta}{X} 1_{(X>0)} \leq U_\theta.$$

It follows that $\pi_n^{(L)}(\theta) \simeq \pi_n^{(O)}(\theta)$, so that for an AO using P_L the cosmological constant in its own universe still has a flat density.

Now let $\alpha \in (0,1)$ and let κ be given by $\kappa_n = n^{1-\alpha}$. Then $q_n = \mathbb{P}(X = 0) = (1-p_n)^n \to 0$ as $n \to \infty$, and

$$\pi_n^{(L)}(\theta) = \frac{1}{1-q_n} \mathbb{E}\left(\frac{U_\theta Y_\theta}{X_0 + Y_\theta}\right) = \frac{1}{n(1-q_n)} \mathbb{E}\left(\frac{\kappa_n Y_\theta}{X_0 + Y_\theta}\right).$$

If $Y \geq 0$ then $\mathbb{E}(1/Y) \geq 1/\mathbb{E}(Y)$ and thus as Y_θ and X_0 are independent,

$$\mathbb{E}\left(\frac{\kappa_n Y_\theta}{X_0 + Y_\theta}\right) \geq \mathbb{E}\left(\frac{\kappa_n Y_\theta}{X_0 + n_0}\right) \geq \frac{\kappa_n m(\theta)}{\mathbb{E}(X_0) + n_0} = \frac{m(\theta)}{M_n + \kappa_n^{-1}(n_0 - m(\theta))},$$

and thus $\liminf_n n\pi_{L,n}(\theta) \geq m(\theta)/M$.

Set $\mu_0 = \mathbb{E}(X_0) = \kappa_n M_n - m(\theta)$. We have

$$\mathbb{E}\Big(\frac{\kappa_n Y_\theta}{X_0 + Y_\theta}\Big) \leq \mathbb{E}\Big(\frac{\kappa_n Y_\theta}{X_0 + 1}\Big) = m(\theta)\mathbb{E}\Big(\frac{\kappa_n}{X_0 + 1}\Big).$$

Straightforward computations give that $\mathbb{E}(X_0 - \mu_0)^4 \leq c\kappa_n^2$. So if $\varepsilon > 0$ and $F_\varepsilon = \{X_0 < (1-\varepsilon)\mu_0\}$, then a fourth moment estimate gives $\mathbb{P}(F_\varepsilon) \leq c'\varepsilon^{-4}\kappa_n^{-2}$. Then

$$\mathbb{E}\Big(\frac{\kappa_n}{1 + X_0}\Big) = \mathbb{E}\Big(\frac{\kappa_n 1_{F_\varepsilon}}{1 + X_0}\Big) + \mathbb{E}\Big(\frac{\kappa_n 1_{F_\varepsilon^c}}{1 + X_0}\Big) \leq \kappa_n \mathbb{P}(F_\varepsilon) + \frac{\kappa_n}{(1 - \varepsilon)\mathbb{E}(X_0)},$$

and it follows that

$$\limsup_n \mathbb{E}\Big(\frac{\kappa_n}{1 + X_0}\Big) \leq \frac{1}{(1 - \varepsilon)M}.$$

Thus we have proved the following Lemma, which shows that in the large universe limit P_E and P_L lead to the same estimates on θ.

Lemma 4.2. *If $\kappa_n = n^{1-\alpha}$ with $\alpha \in (0,1)$ then*

$$\lim_n n\pi_n^{(L)}(\theta) = \lim_n n\pi_n^{(E)}(\theta) = m(\theta)/M. \tag{4.6}$$

Remark 4.3. Many cosmological models involve infinite universes, which may then rise to an infinite number of observers. It is therefore usual in the physics literature to look at observers per unit volume – see for example Section 5 of [5].

References

[1] F. C. Adams. The degree of fine-tuning in our universe – and others. *Physics Reports* **807** (2019), 1–111.

[2] S. Armstrong. Anthropic decision theory for self-locating beliefs, v3. arXiv:1110.6437 (2017).

[3] M. T. Barlow. Study of a filtration expanded to include an honest time. *Z. fur Wahrs.* **44** (1978), 307–323.

[4] M. T. Barlow. Inference for multiple and conditional observers. arXiv:2411.13257v2 (2024).

[5] L. A. Barnes, P. J. Elahi, J. Salcido, R. G. Bower, G. F. Lewis, T. Theuns, M. Schaller, R. A. Crain, J. Schaye. Galaxy formation efficiency and the multiverse explanation of the cosmological constant with EAGLE simulations, *Monthly Notices of the Royal Ast. Soc.* **477** (2018), 3727–3743.

[6] R. Briggs. Putting a value on beauty. In T. S. Gendler and J. Hawthorne (Eds.), Oxford studies in epistemology (Vol. 3, pp. 3-34). Oxford: Oxford University Press, 2010.

[7] N. Bostrom. *Anthropic bias: observation selection effects in science and philosophy.* Routledge, 2002.

[8] B. Carter. Large Number Coincidences and the Anthropic Principle in Cosmology. *Confrontation of Cosmological Theories with Observational Data.* M. S. Longair. Dordrecht, Reidel: 291–298 (1974).

[9] J. Cisewski, J. B. Kadane, M. J. Schervish, T. Seidenfeld, R. Stern. Sleeping Beauty's Credences. *Philosophy of Science*, **83** (2016), No. 3, 324–347.

[10] D. Dieks. Reasoning about the Future: Doom and Beauty. *Synthese* **156** (2007), No. 3, 427–439.

[11] R. Durrett. *Probability: Theory and Examples.* 2nd ed. Duxberry 1996.

[12] A. Elga. Self-locating belief and the Sleeping Beauty problem. *Analysis* **60** (2000), 143–147.

[13] W. Feller. *An Introduction to probability theory and its applications. Vol. 1, 3rd edition.* Wiley, 1968.

[14] U. Grömping. The Sleeping Beauty Problem Demystified. Preprint 2019. Available at: `https://www1.beuth-hochschule.de/FB_II/reports/Report-2019-002.pdf`

[15] J. Y. Halpern. Sleeping Beauty reconsidered: Conditioning and reflection in asynchronous systems. In T. Gendler and J. Hawthorne (Eds.), Oxford studies in epistemology (Vol. 1, pp. 111–142). Oxford: Oxford University Press (2005).

[16] J.B. Hartle, M. Srednicki. Are we typical? *Phys. Rev. D* **75**(2007), 123523.

[17] P. Hawley. Inertia, optimism and Beauty. *Noûs* **1**(2013), no. 1 85–103.

[18] T. Jeulin, and M. Yor. Grossissement d'une filtration et semi-martingales: Formules explicites. Séminaire de Probabilités XII. Lecture Notes in Math. **649** (1978), 78–97. Springer, Berlin.

[19] K. Itô. *Introduction to probability theory.* Cambridge Univ. Press (1984).

[20] D. Lewis. A subjectivist guide to objective chance. In R. C. Jeffrey, ed. *Studies in Inductive logic and Probability*, vol II Berkeley, U. of California Press. (1980).

[21] D. Lewis. Sleeping beauty: reply to Elga. *Analysis* **61** (2001), 171–176.

[22] C. J. G. Meacham. Sleeping beauty and the dynamics of de se beliefs. *Philosophical Studies*, **138** (2008), 245–269.

[23] D. N. Page. Is our Universe likely to decay within 20 billion years? *Phys. Rev. D* **78** (2008), 063535.

[24] J. Pittard. When Beauties Disagree: Why Halfers Should Affirm Robust Perspectivalism. In Tamar Szabó Gendler, and John Hawthorne (eds), Oxford Studies in Epistemology Volume 5, (2015).

[25] M. Piccione, A. Rubenstein. On the interpretation of decision problems with imperfect recall. *Games and Economic Behavior* **20** (1997), 3–24.

[26] M. Rees. *Just six numbers: The deep forces that shape the universe.* Weidenfeld and Nicholson, London 1999.

[27] J. S. Rosenthal. A mathematical analysis of the Sleeping Beauty problem. *Math. Intelligencer.* **31** (2009), 32–37.

[28] A. Sandberg, E. Drexler, T. Ord. Dissolving the Fermi Paradox. arXiv.1806.02404 (2018).

[29] D. Sorini, J. A. Peacock, L. Lombriser. The impact of the cosmological constant on past and future star formation. *Monthly Notices of the Royal Astron. Soc.* **535** (2024), 1449–1474.

[30] M. G. Titelbaum. The relevance of self-locating beliefs. *Philos. Rev.* **177** (2008), 555–605.

[31] V. Verendel, O. Häggström. Fermi's paradox, extraterrestrial life and the future of humanity: a Bayesian analysis. *International Journal of Astrobiology*, **16** (2015), 14–18.

[32] P. Winkler. The sleeping beauty controversy. *The American Mathematical Monthly*, **124** (2017), 579–587.

[33] A. Zuboff. One self: The logic of experience. *Inquiry - An Interdisciplinary Journal of Philosophy*, **33** (1990), 39–68.