Are Entropy Bounds Epistemic?

Emily Adlam

March 19, 2023

Contents

1 Introduction

Discussions about entropy bounds and holography often focus on the strong entropy bound, which conjectures that the entropy associated with some spacelike volume is upper bounded by a quantity proportional to the surface area of the region. For example, holography is often described in terms of a map from space to the spatial boundaries of the world [1]. But there are known counterexamples to the strong entropy bound, including some which are believed to occur in our actual world; to fix these problems we must turn to the covariant entropy bound (also known as the Bousso bound), which tells us that the entropy associated with a null surface known as a 'light-sheet' is upper bounded by a quantity proportional to the surface area of the spacelike region bounding the light-sheet. In Bousso's words, the covariant bound '*evades these counterexamples because of the special properties of light-sheets. Thus, the notion of light-sheets (rather than spatial volumes, or light-cones lacking a nonexpansion condition), appears to be crucial.*'

So it seems that if we are to understand the significance of entropy bounds and holography we should be focusing on the light-sheet construction used in the covariant bound. In this article we set out to understand the meaning of this construction. We observe that there is a possible interpretation of the covariant entropy bound which would suggest that it encodes an epistemic limitation rather than an objective count of the true number of degrees of freedom on a light-sheet; thus we distinguish between ontological and epistemic interpretations of the covariant bound. We discuss how these interpretations relate to Jacobson's derivation of the Einstein equations, and we consider the evidence for each interpretation. Our aim here is not to advocate for either the ontological or epistemic approach in particular, but rather to articulate both possibilities clearly and explore some arguments for and against them.

Note that this paper will largely be focused on entropy bounds in a semiclassical context - in a companion paper we will address the significance of the entropy bounds in the context of quantum mechanics and quantum gravity.

2 Background: Entropy Bounds

Entropy bounds have their origins in an argument due to Bekenstein [2] and Hawking [3] suggesting that if the second law of thermodynamics is to be obeyed in the vicinity of black holes, we must generalize it to include an entropy associated with black holes equal to $\frac{A}{4G\hbar}$, where A is the surface area of the black hole; this value is known as the Bekenstein-Hawking entropy. Subsequently Bekenstein argued that if the generalized second law is to remain true when we lower systems into a black hole, it must be the case that the maximum entropy contained in a spherical system of radius R and total gravitating energy E is $S \leq \frac{2\pi ER}{\hbar}$ [4]; this is known as the Bekenstein bound. Furthermore, for a weakly gravitating system E will always be significantly less than $\frac{2\pi R}{4G}$, so for weakly gravitating systems this bound reduces to the form $S \leq \frac{A}{4\hbar G}$, where A is the area of a spherical surface fully containing the system.

Although this result was derived originally for the weak gravity regime, it inspired a more general conjecture known as the 'strong entropy bound' which says that the maximum entropy which can be contained in any region of space is upper bounded by $\frac{A}{4G\hbar}$, where A is the region of any surface bounding the region. For after all, if we could find a region with entropy greater than $\frac{A}{4G\hbar}$, we could add energy to it adiabatically without changing its entropy, thus transforming it into a black hole of area A - but then this black hole would have entropy greater than the Bekenstein-Hawking entropy. So if the Bekenstein-Hawking entropy formula is correct, the entropy of a black hole of surface area A must be the maximum possible entropy for any region of space with surface area A.

The strong entropy bound initially seems quite surprising. First of all, we may be surprised that there is any finite bound on the entropy inside a region of spacetime, since quantum field theory tells us that the number of degrees of freedom in any given region of space is infinite. There are ways we might make sense of this - for example, in quantum gravity, states above a certain energy can be expected to collapse to a black hole, and this mechanism might be expected to lead to a finite space of physically possible states [5]. Alternatively, several approaches to quantum gravity postulate that spacetime is discrete [6–9], which would also explain why there should be a finite bound on the entropy of a spacetime region. However, the natural way of discretizing spacetime would seem to suggest that the degrees of freedom inside a region should scale with its volume, not the surface area of its surface, so why does the entropy content of spacetime seem to be so much lower than we would naturally expect?

One way of understanding the scaling behaviour is to argue that although the number of states does in principle scale with the volume, in fact many of those potential states have energies so high that they will cause the region to become a black hole, so there are significantly fewer states than we might naively expect (though it remains to be explained why the number should scale with the surface area in particular). In any case, the strong bound seems to be correct in many physically relevant situations [4], and it has led to a flourishing research field aiming to understand quantum gravity in terms of a 'holographic' theory where the gravity theory in a 'bulk' is dual to a field theory defined on a 'boundary' [10–12].

However, the strong bound is not universally true: the heuristic arguments used to motivate it presume a static background and a static entropy-containing region subject to only a small perturbation, but in situations which are not static the strong entropy bound can be violated [13, 14], and several of these scenarios (such as cosmic nflation and gravitational collapse) are believed to occur in our actual universe. Of course, the scenarios in which the bound fails are all quite unusual and so the strong bound is still of great practical relevance. But nonetheless, these special scenarios are surely quite important if one hopes to use the bound to learn something profound about fundamental physics, as many physicists apparently do - for a bound that is not universally valid cannot translate straightforwardly to a universal constraint on fundamental physics. Furthermore, from a relativistic point of view it is immediately clear that there is something wrong with the strong entropy bound, since the specification of the volume enclosed by a given surface is not relativistically covariant and thus strictly speaking the strong entropy bound is not even well-defined in a relativistic context. These considerations suggest that we should really be employing some other entropy bound which is universally valid and relativistically covariant, and which implies the strong entropy bound as a special case in appropriate circumstances, and then we should use *that* bound to make inferences about fundamental physics.

And in fact, such a bound has been found. The 'covariant entropy bound' was originally proposed by Bousso [14] (and hence is also known as the 'Bousso bound'): instead of bounding the information inside a volume enclosed by the relevant surface, it bounds the information on a 'light-sheet' associated with the relevant surface. A light-sheet is defined as a set of null geodesics leaving the surface orthogonally such that the expansion of the set in the direction going away from the surface is zero or negative, i.e. the geodesics are remaining parallel or coming closer together as they get further from the surface. The light-sheet continues up until the geodesics intersect at a 'caustic' (i.e. crossing-point) or encounter a singularity of spacetime. Bousso's bound then simply says that the entropy on the light-sheet associated with a surface of area A is upper bounded by $\frac{A}{4G\hbar}$; and it seems to be the case that this bound is satisfied everywhere in our actual universe, at least at the semiclassical level[1].

Let us illustrate the meaning of a 'light-sheet' by applying it to the case of a closed sphere. Consider the sphere at a fixed moment of time t, according to a frame of reference which is co-moving with the sphere. We can think of the null geodesics leaving the surface as the possible paths for rays of light leaving this surface; they must be emitted at the time t and must be orthogonal to the surface of the sphere. In this case there are four possible directions in which the rays could be emitted - inside the sphere into the past, outside of the sphere into the past, inside the sphere into the future, or outside of the sphere into the future[2]. Clearly the rays emitted outside of the sphere are not converging, so these sets of rays have positive expansion and thus do not define light-sheets; but the rays emitted into the interior of the sphere are indeed converging, so the sphere will have one interior light-sheet oriented into the past and another interior light-sheet oriented into the future. In the co-moving reference frame, the light emitted inwards will propagate towards the centre of the sphere and all of the rays of light will meet at the centre of the sphere at the same time. This meeting point is a 'caustic' and thus the light-sheet terminates at this point.

This example makes it easy to see why the covariant entropy bound implies the strong entropy bound whenever the state inside the relevant surface is static. For in the co-moving reference frame, every spatial point inside the sphere is traversed exactly once by a ray belonging to the light-sheet, though different points will in general be traversed at different times. Thus if the state inside the sphere is not changing, the entropy traversed by these light rays is simply the same as the entropy of the state inside the sphere on any spacelike hypersurface, and thus the covariant bound implies the strong entropy bound in this case.

Now, one may have questions about the line of reasoning by which this bound was arrived at - for example, Dougherty and Callender [16] express various doubts about the validity of the thermodynamic analogy for black holes which originally inspired the entropy bounds. On the other hand, Wallace [17] makes a compelling case that the

[1]It has been shown [15] that in the $k = 0$, radiation dominated Friedmann-Robertson-Walker (FRW) space-time the covariant entropy bound will be violated near the big bang - but this kind of case belongs to the strong gravity regime, so it does not threaten the correctness of the bound at the semiclassical level.

[2]We don't really have to postulate any light travelling backwards in time - light rays emitted 'into the past' can be thought of as light rays coming from the past and terminating on the surface, but it makes the construction more intuitive to think of the light-sheets as being emitted from the surface.

analogy is a solid one. And besides, although the bound may originally have been formulated on the basis of black hole thermodynamics, we now have independent reasons to think that it might be correct at least at the semiclassical level. For unlike the Bekenstein-Hawking entropy, the covariant entropy bound is supposed to apply also to ordinary systems which we can inspect directly and where we think we have a reasonably good grip on the physics, so it's possible to make an independent assessment about whether the bound is correct for these systems. Of course, even for very well-understood systems we don't have enough control over all of the degrees of freedom to literally count them, and theoretical calculations have their limitations since there may be degrees of freedom (particularly quantum gravitational ones) that we're missing; but nonetheless, there are indications that the bound does hold quite generally at least at the semiclassical level. For example, ref [18] proves the bound under some broad conditions - we need only assume that entropy can be modelled as a local fluid obeying certain conditions, and it is known that this fluid approximation is a good one in many regimes. Not only that, we can identify quite specific mechanisms which are responsible for preserving the correctness of the bound in certain cases - in particular, the focussing of geodesics by matter and gravitational collapse both play an important part in ensuring that light-sheets come to an end before they could contain more entropy than the bound allows. So our best current understanding of physics does seem to provide quite good evidence for the covariant bound, at least in the classical limit, and moreover it seems to be linked to gravitational phenomena in intriguing ways.

What about the quantum regime? It has been argued that the bound will not hold if matter fields fail to satisfy the null energy condition, and expectation values of the matter stress-energy tensor in quantum field theory may fail to satisfy this condition. But ref [19] argues that the requirement that light-sheets are non-expanding at all points protects the bound even in cases where the null energy condition is violated, and thus provides a proof of the bound for free fields in the limit where gravitational backreaction is small [19]; an extension to interacting fields appears in ref [20]. So it seems that the bound can still hold in the quantum regime with weak gravity. However, it seems likely that difficulties will arise once we get to regimes which involve both quantum effects and strong gravity. Indeed Smolin [13] argues that the covariant bound cannot possibly continue to be valid in that regime, because there will not in general be any unique well-defined light-sheet at all once we have superpositions of spacetime structures. We will return to this point in part II of this paper, but for now, it is enough to note that even if it turns out to be the case that the covariant entropy bound is only valid in the semi-classical limit, that is still a striking fact about reality and one which seems to demand an explanation.

3 Does the Covariant Bound Count Degrees of Freedom?

The entropy bounds discussed in section 2 have led to speculations about the total amount of information contained in regions of spacetime, leading to the formulation of various 'holographic principles' holding that the physics inside a region can be completely characterised by physics defined on a boundary. However, as Smolin notes, '*most forms of the holographic principle which have been discussed assume the strong form of the entropy bound'* [13][3] - for example, holographic principles are often described as telling us that the '*three dimensional world (can be described as) an image of data that can be stored on a two dimensional projection much like a holographic image*' [1]. In this section, we seek to understand what kinds of claims are justified if we employ the covariant bound rather than the strong entropy bound.

3.1 Holographic Proposals

We will mostly assume in this article that the covariant bound is in some sense correct, at least in the regime of well-understood semiclassical physics. But even if the correctness of the bound is granted, some care must be taken in attempting to infer from it statements about the number of degrees of freedom in some region of space, because entropy does not straightforwardly correspond to an 'amount of information' or 'number of degrees of freedom.' For a start, the idea that entropy is a measure of a number of degrees of freedom seems to draw on an information-theoretic notion of entropy (e.g. the Shannon entropy, which is given by $\sum_x p(x)\log(p(x))$ where the sum is taken over all

[3]Smolin made this comment in 2000 but as far as we can tell it continues to be an accurate assessment of the situation - in particular, AdS/CFT remains the most well-known approach to holography and it is most naturally understood in terms of the strong entropy bound, rather than the covariant one.

possible values of a variable x and $p(x)$ gives the probability for some particular value). Indeed, Bekenstein himself often talks about the entropy of a black hole in informational terms: '*the entropy of a thermodynamic system which is not in equilibrium increases because information about the internal configuration of the system is being lost ... and black hole entropy is understood "as the measure of the inaccessibility of information (to an exterior observer) as to which internal configuration of the black hole is actually realized*' [2]. But Bekenstein's argument for black hole entropy is based on *thermodynamic* entropy, and thermodynamic entropy is distinct from information-theoretic entropy - for example, Dougherty and Callender [16] give the example of an isolated gas, whose thermodynamical entropy is presumably unchanged if we are simply told its microstate, even though this will cause us to assign a zero Shannon entropy to it. It is true that thermodynamic entropy does often coincide with information-theoretic entropy, but one might worry that we have little evidence that this continues to be case in black holes, given that we don't yet have a firm understanding of the state space of black holes.

Moreover, it has often been noted that thermodynamic entropy is not generally an observer-independent quantity. For example, in ref [21] Myrvold argues for a view of thermodynamics as a science which involves designating a certain set of 'manipulable' variables and then studying the responses of physical systems to manipulations of those variables. According to this understanding of thermodynamics, entropy '*must be relativized to the means we have available for gathering information about and manipulating physical systems,*' [22] and hence this entropy is observer-relative. And if we look more closely at the arguments for the strong entropy bound, it is striking that these arguments do indeed proceed by considering the processes that can be performed by agents who have access to only a certain set of 'manipulable' variables (specifically, variables which can be accessed entirely outside of the black hole or relevant spatial region) - for example, Bekenstein's original argument for the Bekenstein-Hawking entropy involved an assumption that entropy is 'lost' once it enters the black hole and becomes inaccessible to external observers, and the argument for the strong entropy bound invoked a process in which an external observer lowers a system into a black hole [4]. So it seems quite plausible that the entropies relevant to these bounds are indeed observer-relative in the same way as ordinary thermodynamic entropies, and if this is so it's unclear that they can be regarded as literally counting degrees of freedom[4].

However, anticipating this criticism, ref [18] makes the following argument: '*there is an apparent tension between the fact that these statements are supposed to have the status of fundamental laws and the fact that entropy is a quantity whose definition is coarse-graining dependent. However, this tension is resolved by noting that the number of degrees of freedom should be an upper bound for the entropy S, irrespective of choice of coarse-graining.*' The idea here is that even though thermodynamical entropy is observer-relative, nonetheless if we discover a universal bound on the observer-relative thermodynamical entropy which applies to all possible observers, we have good reason to infer that this bound reflects the true number of degrees of freedom of the system, since no observer can assign to a system an entropy greater than its total number of degrees of freedom[5].

Thus we will for now entertain the possibility that the covariant bound can indeed be regarded as providing a count of the total number of degrees of freedom in some region of spacetime. But *which* region of spacetime? There are two opposing views on this point. Bousso himself [26] infers from the covariant entropy bound that the total number of degrees of freedom *on a light-sheet* is proportional to the area of the surface with which the light-sheet is associated; we will refer to this as **the light-sheet postulate**. On the other hand, others [13, 27–29] advocate a weaker principle which says only that the total number of degrees of freedom *associated with a surface* is proportional to the area of that surface; we will refer to this as **the surface postulate**. For example, Smolin gives a quantum formulation of the surface postulate that he calls the 'weak entropy bound' which stipulates that for a system Σ defined by the identification of a fixed boundary $\delta\Sigma = S$, and a Hilbert space H_S defined as the the smallest faithful representation of the algebra of observables measurable on the boundary only, then if the area $A(S)$ of the boundary is fixed, we must have $\log(dim(H_S)) \leq \frac{A(S)}{4\hbar G}$ [13].

[4]Even more complications arise once we move to the quantum regime, because it is often assumed that these bounds also apply to the von Neumann entropy, and yet there are ongoing controversies about whether or not the von Neumann entropy is equivalent to the thermodynamic entropy and/or Shannon entropy [23–25]. We will discuss the von Neumann entropy further in Part II of this paper.

[5]This implicit assumption in the argument of Flanagan et al is actually not quite true, since it is possible for observers to be mistaken about the number of degrees of freedom of a system, and if an observer believes that a system has more degrees of freedom than it really does, then she may assign to it an entropy higher than the number of its degrees of freedom. But nonetheless, it seems fair to say that except under very unusual circumstances, the true number of degrees of freedom of a system is an upper bound on the entropy that will be assigned to that system by physically well-informed, rational agents, so the idea that the covariant bound reflects the true number of degrees of freedom doesn't seem unreasonable.

Smolin's argument for the surface postulate [13] is based on the fact that, as we have already noted, the entropies featuring in the derivation of the Bekenstein-Hawking entropy are thermodynamic ones; but rather than concluding on this basis that the covariant bound does not count degrees of freedom at all, he argues that since these thermodynamic arguments pertain specifically to exchanges of matter and radiation between the black hole and external observers, the resulting bound must pertain to the degrees of freedom on the surface rather than the interior. Similarly, Sorkin notes that '*the internal dynamics of a black hole ought to be irrelevant to its exhibited entropy because - almost by definition - the exterior is an autonomous system for whose behavior one should be able to account without ever referring to internal black hole degrees of freedom.*' On this basis Smolin [13], and also Jacobson [27], Rovelli [28] and Sorkin [29] all contend that the Bekenstein-Hawking entropy describes only the amount of information an observer external to the black hole can gain about its interior from measurements made outside the horizon, not the actual amount of information inside the black hole. Smolin then concludes [13] that since the arguments for the strong entropy bound are derived from the Bekenstein-Hawking entropy bound, it follows that this bound also applies only to observables measurable on a boundary, not observables measurable in the interior. And since the covariant bound is intended as a generalization of the strong entropy bound, a similar argument suggests that it too applies only to observables measurable on a boundary.

But if the surface postulate is in fact the only inference about degrees of freedom justified by the covariant entropy bound, this would be a little disappointing, for it doesn't seem to have much to do with 'holography' in the usual sense. After all, it seems quite natural that the number of degrees of freedom on a surface should be proportional to the area of that surface - for example, the surface postulate would seem to follow immediately from the hypothesis that spacetime is discrete. This is exactly Smolin's conclusion: '*It is then difficult to escape the conclusion that the holographic principle, in its weak form, is telling us that nature is fundamentally discrete. The finiteness of the information available per unit area of a surface is to be taken simply as an indication that fundamentally, geometry must turn out to reduce to counting.*' [13] So if entropy bounds only tell us about the degrees of freedom on a surface, they could perhaps be regarded as evidence for the idea that spacetime is fundamentally discrete, but that is not a particularly novel idea - as we have noted, the entropy bounds attracted so much interest precisely because they seemed to be pointing to something *more* than mere finiteness or discreteness.

However, there are reasons to think that the surface postulate does not exhaust the content of the entropy bounds. For the surface postulate does not acknowledge the difference between the original strong entropy bound and the covariant entropy bound - if these bounds really pertained only to the information on a bounding surface, it should make no difference whether we identify the region bounded by the surface with a spacelike slice or a null surface. Yet we seem to have reasonably good evidence that the covariant bound holds everywhere in our actual universe, at least in the semiclassical regime, and meanwhile there are counterexamples indicating that the strong bound does not hold everywhere in our actual universe, even in the semiclassical regime. This suggests that there is in fact a meaningful difference between these two bounds, which provides some justification for thinking that we should be able to use the covariant bound to make some inferences about degrees of freedom pertaining specifically to light-sheets rather than merely bounding surfaces.

So perhaps *both* the surface postulate and the light-sheet postulate are correct, at least at the semiclassical level? Yet if they are both correct, we have some explaining to do, because it surely cannot be a coincidence that the total number of degrees of freedom on a light-sheet are always related in this specific way to the total number of degrees of freedom on its bounding surface - it seems natural to think that these bounds must be connected in some way. We will now examine one interesting way in which they could be connected.

3.2 Thought Experiment: Learning About Physics Inside a Region

We note that the surface postulate has a natural corollary. Our best current physics seems to indicate that information can be transferred between spacetime points only by local processes. That is to say, although we do observe non-local effects in quantum physics, those non-local effects cannot be used to transfer information (e.g. to send superluminal signals) - to the best of our knowledge all processes which do transfer information in this sense are mediated by physical signals travelling along continuous paths in spacetime. Moreover, presumably information passing through a surface via a local process can be understood as being located *on* that surface, at least instantaneously, during its passage. It follows that the rate at which information passes through a surface is limited by the amount of information that can be stored on that surface, and hence the surface postulate implies that there is an upper bound on the rate at which information can pass through a surface which is proportional to the area of the surface. Therefore we will henceforth take it that the surface postulate constrains not only the information on a surface but also the information flux through that surface.

Inspired by this, consider the following thought experiment. Suppose you are presented with a large opaque sphere covered with n LEDs. You are told that the interior of this sphere is not governed by the usual laws of physics - some completely new physics is going on inside it. The LEDs on the outside are linked to detectors on the interior, and the LEDs are reset at one second intervals, lighting up if their detector has received a 'signal' from the inside during the most recent interval. You have no information about what kind of signal is involved, but you have been assured that this is the only way in which information can emerge from the sphere. Your task is to use the information displayed by the LEDs to learn something about the physics on the inside.

How can you go about this task? Presumably, you will do something like the following. First, you will observe the sphere for a time, creating a record of the state of the LEDs every second. Then you will look at the series of states you have written down, attempting to discern some pattern. Hopefully you will eventually be able to write down some kind of 'time evolution' law which relates past states to future states - perhaps a Markovian law which simply relates one state to the next, or perhaps a non-Markovian law which uses some larger number of past states to predict the next state. Having done this, what more can you say about the physics inside the sphere?

The simplest option would be to postulate a one-to-one correspondence between states of the LEDs and states on the interior. Then the evolution law you arrived at for the LED states can be regarded as the evolution law for the states on the interior too. And clearly the states that you assign to the interior will be describable in exactly n bits, which is to say they will have a Shannon entropy of no greater than $\log(n)$.

You could perhaps do something more complex - for example, you could come up with a hypothesis about the nature of the signals which cause the LEDs to switch on or off, and thereby work backwards from the surface states to interior states. In this case there may no longer be a straightforward one-to-one mapping from the surface states to the interior states, because you have to worry about the possibility that the signals interact with one another or with other things in the interior while on their way to the surface. However, it seems likely that whatever model you come up with, the states you arrive at will still be describable in n bits or less, since you can get no more than n bits of information out of the sphere at any given time.

Of course, it could be the case that you find a way to greatly simplify the evolution laws you've come up with by adding some additional bits with unknown values. But it should be noted that postulating additional bits with unknown values is, at least at a mathematical level, equivalent to simply postulating indeterministic evolution laws (the unknown bits would play the role of 'hidden varibables,' making the theory indeterministic at the level of observables) and arguably such indeterministic laws would always be just as simple as the alternative with additional bits. Of course, much depends here on the specifics of the laws you discover, so it is not our intention to suggest that there could never be good scientific reasons for assigning to the system a set of interior states which require more than n bits to describe. Rather our point is that it would not be in any way *surprising* if you were to arrive at a description which assigns the interior only n degrees of freedom, regardless of how many degrees of freedom there really are inside - that would be a natural outcome of the epistemic limitations you are subject to in this situation.

We might make the thought experiment more complex by allowing you to 'intervene,' as in a normal laboratory situation. This would involve a version of the scenario where the LEDs on the outside can both receive signals and also emit signals into the inside of the sphere, so you can prepare a 'state' of the interior by setting the LEDs to some desired configuration, then let the system evolve and observe its response. But of course you will only be able to prepare interior states up to a resolution of n bits, since you only have n degrees of freedom on which you can intervene, and then just as before you will only be able to learn n bits about the succession of interior states that follows, so even with the ability to intervene it would not be surprising if you were to arrive at a description which assigns the interior only n degrees of freedom.

3.3 Light-sheets and accessible information

The upshot of all this is that your epistemic situation with regards to the sphere leads to a natural 'entropy bound' - there will always be compelling reasons to choose a description in which the entropy assigned to the interior is upper bounded by the number of LEDs on the surface. Moreover, if we take it that the information flux through a surface is upper bounded by its surface area, and we assume that information can be transferred only by physical signals which travel along continuous paths in spacetime, then it would seem that the sphere example is a fairly close analogy for the situation described by the covariant entropy bound.

In light of this, consider again the specific role played by the light-sheet construction in the covariant entropy bound. First note that a 'null surface' is really the appropriate relativistically covariant way of describing the region to which we assign a state based on information received at a single time. For example, suppose I observe a football field in order to assign a state to it. Evidently information from different parts of the field will take different times to reach me; so although I may have the impression that I am observing the whole region simultaneously, that is not really true - since photons from one side of the field will have to pass through the other side on their way to my eyes, there is no frame of reference in which all the photons which reach me at a single time also were emitted at the same time. So each 'snapshot' that I see of the field corresponds not to a spacelike hyperplane but to a null surface (assuming that all the information available reaches me at the speed of light).

What about light-sheets? Well, they are simply a special kind of null surface. Specifically, a past-directed light-sheet is the region to which we end up assigning a state if instead of collecting information at a single point, we collect it from a bounding surface. For example, if the surface is a sphere and we collect all the information that arrives at the surface at a single instant of time in the co-moving reference frame, then assuming that the information all arrived at the surface at the speed of light, we get a 'snapshot' of the interior of the sphere which corresponds to the associated light-sheet. So in a sense, the role of the light-sheet construction is simply to identify the appropriate null surface to which we should assign a state based on information collected at a spacelike surface; and as we have just seen in our thought-experiment, it is natural under these circumstances to assign states from a state space which has no more than n degrees of freedom, where n is the number of bits which can pass through the surface in one instant of time. And of course if the surface postulate is correct, it follows that n will be proportional to the area of the surface.

Thus there is a sense in which the surface postulate implies the light-sheet postulate - but only if we interpret the 'degrees of freedom on the light-sheet' as referring to the states that external observers will assign to the light-sheet, rather than the true number of degrees of freedom. So although we argued earlier that the hypothesis of discretized spacetime doesn't explain why the true number of degrees of freedom inside a region should scale with the surface area, arguably it *does* explain why the number of bits of information *that external observers can obtain* about the physics inside a region should have an upper bound proportional to the area of the bounding surface. From this point of view, it looks as though the covariant entropy bound is not telling us anything about the true number of degrees of freedom inside a region, but rather describing the state space which will naturally arise if the states are assigned by agents observing from the outside who can learn about the system only via information which passes through its bounding surface.

Now, one might object that this argument only applies to a closed surface, whereas the covariant entropy bound applies also to open surfaces. But we can formulate a similar argument for an open surface too. Suppose I am standing on one side of a screen of area A and trying to gain information about a system on the other side of the screen. Again, if the surface postulate is correct, the total amount of information about the system which can pass through the screen to me has an upper bound proportional to the area of the surface, so again we'd expect the number of bits in my description of the system on the other side of the screen to have an upper bound proportional to the region of a bounding surface. Of course, in this case it's possible for information to go around the screen in order to reach me, but if I am to learn anything from that information I'd have to know about the state of the region above around the screen so I can account for any further interactions the signals might undergo on the way. Thus we need to consider the process by which I learn the state of the surrounding regions as well, which requires us to extend the screen to cover information coming from those regions too, and so on, and therefore it's hard to see how I could do any better than learning a number of bits proportional to the area of the surface, unless I'm allowed to assume that I already know the state of some region around the screen. So the epistemic interpretation of the bound seems to work for open surfaces too: the light-sheet of such a surface still represents the locus of all the information which could pass through the screen in a given instant, which can be expected to yield an upper bound proportional to the area of the screen.

The special role of surfaces as interfaces through which observers access systems has been pointed out by various authors. Oeckl [30] has emphasized this feature of experimental design: '*Consider for example a scattering experiment in high energy physics. A typical detector has roughly the form of a sphere with the scattering happening inside (e.g.*

a collision of incoming beams). The entries for particles and the individual detection devices are arranged on the surface.' We will return to Oeckl's general boundary formalism in Part II of this paper. Similarly Crane [31] observed that '*if we divide the universe into "system" and "observer" the observer no longer measures the state of the system, but only that part of it which impinges instantaneously on the observer,*' and thus he proposed a system of observable algebras and Hilbert spaces, one associated with every possible splitting of the universe into system and observer; he therefore came up with a categorical framework to describe the association of Hilbert spaces with boundaries. And May [32] has argued that holography can be expressed as the requirement that the asymptotic tasks which are possible using the bulk dynamics should coincide with tasks that are possible using the boundary; here 'asymptotic tasks' are those which can be stated in terms of inputs and outputs located on the spacetime boundary. Put this way, it does not seem surprising that our physics should have this feature: for after all we never really access a bulk directly, we only ever manipulate systems through a boundary, i.e. the spatial boundary separating the the system from us, so it is to be expected that the physical theories we come up with describing possible tasks should be limited to the tasks which are possible using the boundary.

4 Epistemic or Ontological

The thought experiment of section 3.2 leaves us with a dilemma. For it suggests that, regardless of the number of degrees of freedom there really are on light-sheets, observers will naturally end up describing the physics on a light-sheet with a number of bits no greater than the amount of information which can pass through the associated surface - which according to the surface postulate will be proportional to the area of that surface. So how can we ever know whether or not the covariant bound describes the true number of degrees of freedom on a light-sheet? Won't we inevitably end up with descriptions of light-sheets having a number of degrees of freedom proportional to the area of the associated surface, no matter whether or not the true degrees of freedom are really bounded in this way? It seems that anyone who accepts the surface postulate has good reason to think of the covariant bound not as describing the true numbers of degrees of freedom on a light-sheet, but simply as an epistemic restriction on the number of degrees of freedom accessible to an external observer.

It should be reinforced that the 'epistemic' account we have explored here is largely motivated by relatively simple cases such as a system enclosed in a sphere or behind a screen. It is possible that in more complicated cases the epistemic interpretation would break down in ways that are not evident from the simple cases, and if so this would presumably indicate that the ontological view is correct. But for now, we will suppose that the epistemic interpretation continues to work in more complex cases. Indeed, there are indications within the literature on entropy bounds that this is the right away to think about the covariant entropy bound - in particular, some of the counterexamples to the original strong entropy bound make it clear that what matters in the formulation of an entropy bound for surface S is not just its area, but also how much information can reach it. For example, suppose we have two spherical shells, one contained inside the other. The large sphere has bigger area and thus the strong entropy bound tells us that it can contain more information, which is what we would expect given that it contains all the information that is inside the smaller sphere plus whatever information lies in the region between the two shells. But now we can imagine constructing a 'wiggly surface' obtained by placing a third shell in between these two shells and then folding or distorting it in order to make it wiggly. Evidently we can arrange for the wiggly surface to have area significantly greater than the outer sphere, and yet it is inside the outer sphere and thus presumably contains less information. On the basis of this example, Smolin argues that '*the area that is relevant for the measure of information is not the actual area of the surface Ω. Rather it must correspond to the information reaching Ω from its interior.*' [13] [6] That is to say, some of the cases motivating the covariant bound do indeed seem to suggest that the bound should be regarded as describing the amount of information which is able to escape the region through its bounding surface, rather than necessarily the total amount of information inside the region.

In light of these considerations, we now set out several different attitudes that one might take to the covariant bound, assuming that one accepts the surface postulate (which, recall, is an independently plausible postulate, particularly if we already have reason to think that the number of degrees of freedom on a surface should be finite).

[6] In the original Smolin used S for the surface, but we have switched the symbol to Ω to avoid confusion with entropy.

1. **Ontological:** views which hold that the covariant entropy bound is a constraint on the number of real number of ontological degrees of freedom on a light-sheet:

   (a) **Pragmatic-ontological:** views which invoke some form of pragmatism or empiricism to say that since we can never have any information about degrees of freedom beyond the covariant entropy bound, we should assume there aren't any further degrees of freedom, but the bound doesn't have any deep meaning beyond that.

   (b) **Positive-ontological:** views which postulate that the covariant entropy bound is a fundamental constraint or follows from aspects of the fundamental ontology, so it tells us something profound about the nature of our reality.

2. **Epistemic:** views which hold that the covariant entropy bound is merely epistemic, and the real number of degrees of freedom on a light-sheet may be greater than the amount specified by the bound:

   (a) **Pragmatic-epistemic:** views which postulate that there may exist degrees of freedom beyond those allowed by the covariant entropy bound, but we will probably never know anything about them and they will have no impact on anything we can observe, so we can get on with physics as if they don't exist at all.

   (b) **Positive-epistemic:** views which postulate that there may exist degrees of freedom beyond those allowed by the covariant entropy bound, and it may be important to take them into consideration, even though we can't directly observe them.

Now, there may appear to be a natural connection between interpretations of the covariant bound and the interpretation of the entropy appearing in the covariant bound. If this entropy is understood to be a thermodynamic entropy which is observer-relative in the sense described by Myrvold, then it seems natural to think of the covariant bound as a claim about the information which can be obtained by an observer with access to certain manipulable variables, so it is epistemic. Whereas if one takes this entropy to be an information-theoretic entropy, then it seems natural to think of the covariant bound as counting the number of degrees of freedom in the region, so it is ontological. Moreover, this identification may appear to settle the question in favour of epistemic views, since as we have noted, the thermodynamic arguments for the Bekenstein bound do seem to pertain to observer-relative thermodynamic entropies rather than information-theoretic entropies, and the covariant bound is a generalization of the Bekenstein bound.

However, we would caution against making the epistemic/thermodynamical and ontological/information-theoretic identification in an overly simple way. For proponents of an ontological view of the bound need not deny that the entropies employed in arguments for the Bekenstein-Hawking entropy and the strong entropy bound are observer-relative thermodynamical entropies - *in any particular case*, the entropy that we assign to a region of spacetime is a measure of the information available to some specific (possibly hypothetical) observer. But the Bekenstein-Hawking entropy and the strong and covariant bounds do not describe the entropy assigned in any particular case, rather they represent a universal upper bound on the entropies assigned by all possible observer in all possible cases. And the universality of this bound means that it cannot be understood as a measure of the information available to any one specific observer, so we still have the possibility of interpreting it ontologically - for example, we have already mentioned the argument of Flanagan et al [18] that the universality of the bound indicates that it is a count of the true number of degrees of freedom. On the other hand, we have noted that a universal entropy bound could also be explained by the fact that all realistic agents are all subject to some kind of epistemic restriction which prevents them from gaining more than a certain number of bits of information about it. So really, nothing can be inferred one way or another from the observer-relative nature of the entropies used in the derivation of the bounds, because the nature of those entropies does not tell us the origin on the *universal bound* on those entropies. Thus we would urge that, contrary to the arguments made by Smolin and others [13], the thermodynamical nature of Bekenstein's reasoning does not settle the question of whether the covariant bound (or the strong entropy bound, or even the Bekenstein-Hawking entropy) is epistemic or ontological.

We can classify some of the views we have previously examined within this schema. Bousso clearly takes a positive-ontological view, writing of the covariant entropy bound: '*The bound's simplicity, in addition to its generality, makes the case for its fundamental significance compelling*' [26]. Other proponents of holography in the conventional sense, such as Susskind [1] and t'Hooft [33], also seem to take this kind of view. On the other hand Smolin [13], Jacobson [27] and Sorkin [29] all appear to take an epistemic view. For example, these kinds of epistemic considerations are implicit in Smolin's criticisms of the original Bekenstein bound [13] - he emphasizes that the thermodynamic bounds on the entropy inside a black hole pertain specifically to the *accessible* entropy, and we shouldn't assume that the information accessible from outside a surface constitutes a bound on the information inside that surface, so it isn't clear that the arguments for Bekenstein's bound necessarily constrain the 'real' degrees of freedom.

Meanwhile, it seems likely that the idea that entropy is 'lost' when information falls into a black hole is partly grounded on an implicit acceptance of a pragmatic view, since this would justify the practice of simply ignoring the possibility of degrees of freedom beyond the accessible ones. For example, Dougherty and Callender offer the following account of the reasons why physicists believe that entropy is lost when it crosses a black hole horizon: '*it is held that the entropy vanishes when it passes behind the event horizon because we can't gain access to it. The system itself doesn't vanish; indeed, it had better not because its mass is needed to drive area increase. But for the ordinary entropy, when it crosses the event horizon it's "out of sight, out of world," or at least, out of physics.*' [16] So it seems possible that some physicists are really agnostic about whether there is additional information in the black hole but are willing to disregard any such information for practical purposes, and presumably such physicists would also be inclined to take a pragmatic-epistemic or pragmatic-ontological approach entropy bounds more generally.

However, there is potential for the pragmatic approach to lead to confusion: if we simply ignore the existence of some degrees of freedom, we may get strange effects where information appears to mysteriously vanish. Indeed, it has been argued that this is exactly what is going on in the case of the black hole information paradox. As noted by Rovelli [28], the Page argument for the information loss paradox '*is based on the fact that if the number of black hole states is determined by the area, then there are no more available states to be entangled with the Hawking radiation when the black hole shrinks*' and thus '*if there are more states available in a black hole than $e^{A/4}$, then the Page argument for the information loss paradox fails.*' Therefore this apparent paradox may simply be the result of ignoring the existence of degrees of freedom beyond the number allowed by the entropy bounds. Similarly, it seems very possible that simply ignoring some degrees of freedom may be an obstacle to progress on quantum gravity. So although the pragmatic views have merits in certain contexts, we will henceforth focus on the positive views; let us look more closely at the consequences of the positive-epistemic view and the positive-ontological view.

4.1 Positive-Epistemic Views

It is common to understand the term 'epistemic' as connoting 'not objective.' For example, sometimes a distinction is made between 'objective' and 'epistemic' conceptions of entropy, as in ref [16]. But we emphasize that entropy bounds can be 'epistemic' while still being objective, for the facts about the amount of information that can be obtained by physical systems like us are just as objective as any other physical facts. So really 'epistemic' here should be understood as meaning something like 'relational' - if the covariant bound is epistemic, that simply means it doesn't describe the degrees of freedom of an individual physical system, but rather says something about a relation between physical systems and observers, which themselves are a kind of physical system. These sorts of objective, relational epistemic facts are familiar from thermodynamics - for example, Myrvold emphasizes that although thermodynamic concepts like entropy are observer-relative, at the same time '*it would be misleading to call them subjective, as we are considering limitations on the physical means that are at the agents' disposal.*' [22]

Similarly, we emphasize that a 'positive-epistemic' understanding of the entropy bounds need not lead us to operationalism, positivism, a perspectival account of reality [34], participatory realism [35], or any other such view - the recognition that our epistemic limits may place constraints on the physics that we arrive at is still compatible with a standard realist view postulating a mind-independent external reality. Indeed, the idea that there can be nontrivial epistemic limits actually *requires* the postulation of a mind-independent reality, for if reality consists entirely of agents' perspectives or experiences, then nothing can exist beyond their epistemic limits and hence there cannot be any nontrivial epistemic limits. Of course, if we come to the conclusion that our epistemic limits are so severe that significant portions of reality are forever unknowable to us, this might motivate us to adopt only a very qualified realist stance towards the theoretical entities appearing in physics, understanding them as bearing some systematic relation to the external world but accepting that the external world may not literally contain anything that corresponds to them, but this does not prevent us from believing that there exists a mind-independent external reality, even if we only able to dimly grasp its nature.

To understand what follows from an positive-epistemic view, let us disregard established physics for a while and simply ask what effects we might expect to observe if we did in fact live in a world where every region of space contains some degrees of freedom which are inaccessible to any external observers. For a start, presumably we would observe phenomena which would look indeterministic, even if the underlying processes are really deterministic: the inaccessible degrees of freedom would essentially act as 'hidden variables,' since they would determine the outcomes of various interactions but we would never be able to observe them directly. Thus it may be tempting to elide the hidden degrees of freedom with the 'hidden variables' invoked in some interpretations of quantum mechanics, thus potentially providing a novel explanation for quantum indeterminism.

Second, perhaps we might get behaviour superficially similar to the uncertainty relations in quantum mechanics [36]. For suppose a particle is localised with position uncertainty δx in each direction. Thus the particle is definitely to be found within a region of surface area $4\pi\delta x^2$, and the surface postulate implies that the total amount of information we can get out of this region per unit time is proportional to $4\pi\delta x^2$. In particular, in order to find out the momentum of the particle we must get the information about its momentum at a given time out of the region, and thus the smaller the region, the less information we can obtain about its momentum, or equivalently the greater the momentum uncertainty. So, roughly speaking, we'd expect that as the particle becomes better localised in spacetime our uncertainty about its momentum would increase, and vice versa[7]

That said, one might also worry that an epistemic interpretation of the bound would predict effects that are inconsistent with known science. For example, it seems possible that the existence of inaccessible degrees of freedom would lead to violations of the conservation of energy, because energy could disappear into or emerge from invisible degrees of freedom. Of course, in general relativity energy is not strictly conserved in any case, except along a Killing vector; however, the relation $\nabla^v T_{uv} = 0$ expresses a local form of conservation of energy-momentum, so it seems possible that if energy were indeed disappearing into or emerging from invisible degrees of freedom this relation would sometimes be violated. Since $\nabla^v T_{uv} = 0$ in fact seems to be true universally, this looks like an argument against the epistemic interpretation of the bound. However, we note the existence of a proposal that the cosmological constant could be understood as a consequence of energy 'leaking' into unobservable degrees of freedom associated with the discretization of spacetime [38], so the violations of energy conservation would not be observable in local laboratory searches in approximately flat spacetime, but would manifest themselves on a larger scale in the form of the curvature of spacetime. Of course this idea is quite speculative and needs further development, but at least it may reassure us that the empirical evidence is not necessarily incompatible with the possibility of energy disappearing into or emerging from hidden degrees of freedom.

These examples show that an epistemic interpretation of the entropy bound may offer interesting new possibilities for understanding quantum mechanics. Indeed, the bound on the amount of information which we can get out of a spacetime region looks a lot like the kind of 'epistemic limitation' postulated by Spekkens, and as he shows in ref [39], the existence of epistemic limitations can explain many of the puzzling features of quantum mechanics. Similarly, Rovelli [40] has argued that we can get quite a long way towards deriving quantum mechanics using two simple postulates: '*1) There is a maximum amount of relevant information that can be extracted from a system, and 2) It is always possible to acquire new information about a system.*' And in fact these postulates do seem to describe the kind of behaviour we might expect to follow from an epistemic entropy bound: there is a maximum amount of information that can be extracted from a system because only a finite amount of information can emerge through the surface bounding the system at a given time, but it's always possible to acquire new information about the system because the information extracted at a single time doesn't exhaust the information inside the bounding surface. So the epistemic interpretation of the covariant bound could be regarded as one possible realisation of the ideas of Spekkens and Rovelli: in this picture the limitation on how much information we can obtain from a system need not be treated as a fundamental, unanalysable physical principle, but rather can be explained as a simple consequence of physical limitations on how much information observers like us can obtain. So there may indeed be value in thinking explicitly about the possibility of additional degrees of freedom beyond the covariant bound, rather than merely disregarding them on pragmatic grounds - even if we can't learn anything very concrete about these degrees of freedom, simply acknowledging their existence can provide novel explanations of various features of physics.

[7]In ref [37] Bousso also suggests a derivation of the uncertainty principle from the generalized covariant bound. However, this derivation mostly follows the standard quantum-mechanical argument to obtain $\delta x \delta p \geq \frac{c}{2}$ and then invokes the Bekenstein bound only to argue that $c = \hbar$; whereas the epistemic approach we have suggested here aims to do more than just fix the value of the constant - it is a putative explanation of why there should be an uncertainty relation at all, independent of the formal structure of quantum mechanics.

More generally, the possibility of an epistemic interpretation of the covariant bound provides new motivation for scepticism about those interpretations of quantum mechanics which insist that quantum mechanics is 'complete,' i.e. it describes the full content of reality. For quantum mechanics has been formulated on the basis of the information about a system that can arrive at us through its bounding surfaces, which need not exhaust all the facts about the system if we accept that the covariant bound is epistemic. Furthermore, proponents of the idea that quantum mechanics is complete often argue that this view is supported by the great empirical success of quantum mechanics - but if quantum mechanics is a complete description of all *accessible information*, great empirical success is exactly what we would expect. No experiment will ever lead to observations which directly contradict a theory that is a correct description of all information accessible to observers like ourselves, and yet such a theory may not be a complete description of reality, so we may well find that despite the empirical success there are conceptual puzzles that seem unresolvable - as noted earlier, the black hole information paradox may be an example of this.

4.2 Positive-Ontological Views

We begin our discussion of positive-ontological approaches to the covariant bound by noting that the very possibility of an epistemic interpretation of the bound poses quite a serious problem for attempts to understand it in an ontological way, because proponents of such views will be vulnerable to the criticism that they are simply mistaking an epistemic constraint for an ontological one. For example, the fact that four different methods of calculating the black hole entropy all give the Bekenstein-Hawking entropy formula [17] might initially be taken as evidence that the formula reflects the true number of degrees of freedom in a black hole, but once we start thinking about possible epistemic restrictions, we might worry that the result is the same in all four cases not because it reflects the true, ontological number of degrees of freedom of the black hole but merely because all four approaches are based on physics developed by agents subject to a certain kind of universal epistemic restriction.

This looks like a problem for approaches which attempt to account for the entropy bounds by postulating a gravity theory and then showing by explicit counting of degrees of freedom that the theory agrees with the entropy bound - this has been done in several ways for the specific case of black hole entropy [41, 42]. Such a strategy is perfectly reasonable if one takes a pragmatic view which specifies that the theory is only intended as a description of accessible degrees of freedom, but it may be too limited if one's aim is to go beyond a purely operational description. For even if the entropy bounds seem to arise quite naturally from the structure of a gravity theory, proponents of an epistemic view may always retort that the theory looks this way precisely because it has been arrived at by formalising the kinds of data which are available to agents accessing systems through an interfacing surface, and therefore it fails to account for degrees of freedom which are inaccessible to external agents. In particular one might worry that such a theory would not be suitable for situations where there is no external observer, so it could not be employed to do cosmology or to describe the universe as a whole,

How can proponents of an ontological view resist this kind of undermining? Well, rather than simply postulating a theory in which the number of degrees of freedom just happen to match the entropy bound, they could instead take a principled approach by calling into question some of the ontological and metaphysical presuppositions which enter implicitly into the epistemic view. In particular, the argument that there could be more degrees of freedom inside a given region than observers are able to access is founded on the assumption that we inhabit a three-dimensional space (defined by a foliation of a four-dimensional spacetime) with independent autonomous degrees of freedom at each point of space. But the thought experiment of section 3.2 makes it clear that this assumption far outstrips our evidence - we don't actually receive data in three-dimensional chunks but rather on two-dimensional surfaces, so there is a significant mismatch between our representation of the data in 3D space and the amount of information we actually obtain about the contents of space. Thus rather than concluding that there exists a whole realm of invisible, inaccessible physics, perhaps we should adopt a different ontology which is a better match to the data.

For example, one might imagine an ontology where the number of sites for information storage available in a region scales with the volume, as would naturally expect, but the information at the sites is not in general independent, so the total number of effective degrees of freedom scales instead with the surface area. As t'Hooft puts it: '*This suggests that physical degrees of freedom in three-space are not independent but, if considered at Planckian scale, they must be infinitely correlated*' [33]. Alternatively, one could imagine moving to an ontology composed entirely of two-dimensional surfaces, or something isomorphic to them. After all, since we always interact with bounding surfaces rather than the light-sheets associated with those surfaces, it seems that only the bounding surfaces are strictly necessary to account for our experiences, so we could potentially take it that a bounding surface and its associated light-sheet are really just equivalent representations of the same element of ontology. Indeed, there are some indications from within General Relativity itself that this may be the right way to think about its ontology - for diffeomorphism invariance together with the equivalence principle prevent us from defining quantities like energy, momentum and angular momentum at individual spacetime points, and thus these quantities must instead be defined in terms of integrals over two dimensional surfaces[8] [43]; so it seems quite natural to associate the theory not with an ontology of locally defined quantities at individual spacetime points, but rather with an ontology in which all meaningful quantities are associated with surfaces.

However, some care is required here, because if we choose an ontology specifically to match the nature of the data available to observers like us, we run the risk of arriving at on ontology which works for the kinds of systems we observe but which depends on a split between systems and observers and hence does not allow us to accommodate observers or to explain the relation between systems and observers in physical terms. For example, it is easy enough to say that whenever an observer observes a region through a bounding surface, only the surface is really ontological and the region inside is just a representation of that surface created by the observer. But what if there is another observer inside the region, observing a smaller region through another bounding surface? Should we then say that the inner observer is not real? What about the case where two observers observe one another through different bounding surfaces - whose surface trumps whose? Moreover we can't simply put all possible bounding surfaces into the ontology, because then the ontology would contain the whole of spacetime and we would be back to the 3D picture. So evidently if we don't wish to resort to operationalism we will need some principled approach which allows us to identify an ontology of surfaces without appealing to external observers, in order that we can integrate observed systems and observers into a unified physical description.

There are two obvious ways to achieve this. The first is to accept that the choice of surface is always relativized to an observer, and adopt a relational approach similar to relational quantum mechanics (RQM) [44, 45] in which we say that *all* valid physical descriptions must be relative to an observer. RQM avoids collapsing to operationalism because it tells us that *every* physical system counts as an observer and thus each physical system defines its own observer-system split. So we could potentially add to RQM the specification that for every physical system, the description of the world relative to that system at a given time is defined on a two-dimensional surface, and that surface contains everything that belongs to the ontology defined relative to that observer[9]. Moreover, provided that we include a postulate ensuring the existence of links between the perspectives of different observers, as described in ref [46], RQM still gives rise to a shared macroscopic reality common to large classes of observers, and thus we can potentially arrive at something close to an integrated description of 'the universe as a whole,' even though this approach does not allow such a thing as a completely observer-independent gods-eye view of the universe.

Alternatively, we could define a unique surface, or small set of surfaces, to which all of spacetime can be reduced - that surface thus acting as something like a 'hologram' (in Part II we will take a more detailed look at holography). For example, a picture like this arises within Kent's solution to the Lorentzian quantum reality problem [47, 48] which is intended as a solution to the measurement problem: in this model the wavefunction undergoes its usual unitary evolution until the end of time, and then we imagine something like a measurement being performed on the final state, with the actual course of history being determined by the result of the measurement (the 'measurement' is to be understood as a mathematical device for extracting probabilities, rather than a literal physical operation)[10]. This approach looks like it could be a good fit with the covariant bound because it has the consequence that the information in a region (e.g. on a light-sheet) depends on the state of its bounding surfaces in its lightlike or timelike future: a beable can exist at some spacetime point only if there is a record of its existence somewhere in the final state, and thus the total amount of information in a given region cannot be greater than the amount of information which can escape

[8] In spacetimes with boundaries we can take the surfaces to the boundary and thus obtain asymptotic definitions of energy, momentum and angular momentum; in spacetimes without boundaries we can instead define quasi-local observables, which means that the energy, momentum and angular momentum of an arbitrary region are defined in terms of integrals over its boundary.

[9] It is not entirely clear how one arrives at a specification of which observers exist in the first place within this kind of approach, given that everything is supposed to be relativized to an observer so there can't be observer-independent facts about the set of observers. But this is a generic problem for all relational and perspectival approaches to quantum mechanics - adding the specification that data is defined on surfaces doesn't make the problem worse, although it doesn't offer a solution either.

[10] Kent also makes allowance for the possibility that there is no end of time - in this case we simply take a limit as $t \to \infty$, making some assumptions which ensure that the limit is well-defined.

that region to arrive at timelike infinity, so something like the covariant entropy bound will naturally arise[11]. And in this picture we have no need to appeal to external observers to decide which surfaces should feature in the ontology, because in a sense the ontology includes only the surface on which the result of the final measurement is defined; the contents of the rest of spacetime, including observers themselves, are merely a projection from that surface.

5 Gravity

Having classified possible views on the covariant bound, we may now ask whether the evidence we have for the correctness of the covariant bound favours either the epistemic or ontological approach. To start with, note that if we had arrived at the covariant entropy bound purely by counting the number of degrees of freedom in the observations that we make about various physical systems, the thought-experiment in section 3.2 would be a strong argument against taking the bound too seriously - after all, if you know in advance that regardless of how many degrees of freedom there are in a system you will never be able to extract more than n bits of information about it, then when you subsequently observe that you extract n bits of information, you have grounds to conclude only that the system has at least n degrees of freedom, not that it has exactly n degrees of freedom. However, we did not actually arrive at the covariant entropy bound by directly counting degrees of freedom, but rather by a variety of theoretical arguments. This seems quite puzzling - why should these theoretical arguments have led us to the exact same bound that our epistemic restrictions would have imposed if we had literally counted degrees of freedom?

In the case of the thermodynamic arguments for the bound (i.e. the original motivation for it as a generalization of the Bekenstein bound, which was proposed in order to ensure the correctness of the second law of black hole thermodynamics) this puzzle is quite easily resolved. For we have already argued that thermodynamic entropy is often relativized to a choice of manipulable variables and that the thermodynamic arguments for the Bekenstein-Hawking entropy and strong entropy bound are naturally understood as pertaining to this observer-relative notion of the entropy. So in fact, it would not be at all surprising if this sort of reasoning led us to a general bound not on the total degrees of freedom in regions of spacetime, but on the *accessible* degrees of freedom in regions of spacetime, as encoded in the epistemic view of the bound.

That said, we have also seen that there exists evidence for the covariant bound which is not based on black hole thermodynamics, or indeed any other kind of thermodynamics. As noted by Bousso [14], there is no single mechanism which is responsible for the correctness of the covariant entropy bound in all cases (that is exactly why it is considered so mysterious) but in many examples it is the general-relativistic focussing theorem which does most of the work. In Bousso's own words: '*Gravitational backreaction plays a crucial role in preventing violations of this bound. In realistic systems, an increase in entropy is accompanied by an increase in energy. Energy focusses light rays by an amount proportional to G. Thus it hastens the termination of a light-sheet at caustic points, preventing it from "seeing" too much entropy.*' [37] This leaves us with something of a dilemma: the covariant bound looks like an epistemic restriction, and yet some of the evidence for it comes from gravitational effects, which are surely not epistemic. Does this disprove the epistemic view, or can we argue that some epistemic considerations are somehow being smuggled into the gravitational evidence?

5.1 Focussing

To answer this question, let us take a closer look a the focussing theorem. The theorem arises from Raychauduri's equation [50], which describes the evolution of a 'null congruence' (i.e. a family of geodesics, or freely propagating light-rays) with affine parameter λ in a D-dimensional manifold[12]:

$$\frac{d\theta}{d\lambda} = -\frac{1}{D-2}\theta^2 - \sigma_{ab}\sigma^{ab} - R_{ab}k^a k^b$$

[11] The version of this approach presented by Kent in ref [49] is slightly different because it has the consequence that the information at a point is determined only by parts of the final state which are *outside* the future light-cone of that point, so this formulation of Kent's view does not postulate the same tight connection between the information in a region and its bounding surfaces in the lightlike or timelike future. In future work we hope to consider in more detail the relationship between entropy bounds and these two formulations of Kent's approach.

[12] There is a more general form of the Raychauduri equation which also works for non-geodesics - it includes a quantity known as the 'twist' which is always zero along geodesics - but we will only be concerned with the geodesic version here.

Here, θ is the expansion of the bundle, representing the change in the area spanned by infinitesimally neighbouring geodesics: formally, $\theta = \lim_{A \to 0} \frac{1}{A}\frac{dA}{d\lambda}$. Meanwhile σ_{ab} is the shear, representing the change in the shape of the area spanned by infinitesimally neighbouring geodesics. Using Raychauduri's equation, we can arrive at the focussing theorem [51], which says that in a spacetime satisfying the null curvature condition ($R^{ab}k^a k^b \geq 0$ for all null vectors k^a) the expansion is non-increasing at all points where θ is finite (i.e. points which are not singularities or caustics) of a surface-orthogonal null congruence. That is to say, if the null curvature condition is satisfied geodesics can focus, i.e. come together to a point, but they can never anti-focus, i.e. diverge away from one another.

It is important to emphasize that the focussing theorem is a theorem of differential geometry, so it does not depend on the Einstein equations or on any features of general relativity specifically. However, Einstein's equations imply that the null curvature condition is equivalent to the null energy condition, i.e. $T_{ab}k^a k^b \geq 0$. So in a general relativistic spacetime, we can further conclude that if the null *energy* condition is satisfied then geodesics can focus but not anti-focus. And since the null energy condition is true at least on average for all realistic classical matter[13], we can conclude that in classical general relativity geodesics will always focus.

It is evident that the covariant bound would not work without gravitational focussing. In Bousso's own words, '*Entropy costs energy, energy focusses light, focussing leads to the formation of caustics, and caustics prevent light-sheets from going on forever.*' [26] This leads to a natural speculation - could we turn this reasoning around and explain gravitational phenomena as a consequence of a holographic principle? For if we are prepared to take for granted that the information on a light-sheet (or at least, the *accessible* information on a light-sheet) is upper bounded by the information that can pass through its bounding surface, then Bousso's reasoning can be employed to argue that geodesics must be focussed by the presence of mass-energy, so the existence of some gravity-like behaviour is an inevitable consequence of the covariant entropy bound.

One may wonder exactly how far we can get with this approach: does the covariant bound single out the Einstein equations specifically, or does it merely lead us to a class of theories all exhibiting some sort of focussing of geodesics? In fact there have been suggestions that entropy relations can be used to derive the Einstein equations - most famously Jacobson's 1995 'thermodynamical' derivation of the Einstein equations [52], which employs an entropy bound as one of its ingredients. Jacobson's argument was published four years before Bousso published his covariant bound, so his derivation associates an entropy with a surface rather than a light-sheet; however, in appendix A we suggest a way of rewriting Jacobson's derivation using the light-sheet formulation. The derivation is not fully general - in appendix A.1 we discuss some of its limitations - but it provides at least some support for the idea that Einstein's equations could be regarded as a consequence of the covariant entropy bound.

In what follows we will not assume that the exact form of the Einstein equations can be derived from the covariant entropy bound, but we will take it that, as indicated by Bousso's comments above, the covariant bound does at least entail the existence of some kinds of gravitational phenomena. So we have a two-way relation: the covariant bound implies the existence of certain gravitational phenomena, and meanwhile certain gravitational phenomena imply the covariant bound, or at least play a major role in protecting the bound in many cases. It seems natural, therefore, to suppose that one of these things can be regarded as explaining the other - but how do we decide in which direction the explanation should go? Ref [53] addresses this question, criticizing Jacobson's proof on the grounds that we must invoke the Einstein equations to arrive at the universal relation between entropy and area on which the proof is based, so the derivation is circular. To make this point, ref [53] assumes that the entropy in question is an entanglement entropy and then notes that the entanglement is usually thought to depend on the types of fields present, via the relation $S = \frac{nA}{L^2}$, where n is the number of types of fields and L is the cutoff length at which we reach the Planck regime. But matter fields interact gravitationally and thus if there are more fields present the energy density will be higher, so we will reach the Planck scale at a higher cutoff length L; it can be shown that the two effects exactly cancel out, so we end up with a universal relation between area and entropy that does not depend on the number of fields. Evidently if this is the right way to think about the relation between entropy and area, then this relation is a *consequence* of the Einstein equations, and thus it would indeed be circular to derive the Einstein equations from it.

In part II of this paper we will examine the relationship between entropy bounds and entanglement entropy in greater detail, but for now it is enough to note that this criticism is begging the question: someone who believes that the covariant bound should be understood as explaining the existence of gravitational phenomena would presumably

[13] The null energy condition can be violated classically in the Casimir effect, but the averaged null energy condition $\langle T_{ab}\rangle k^a k^b \geq 0$, which seems to be correct for all realistic classical matter, is adequate for the conclusion here.

argue that the entropy bounds should *not* be understood as a consequence of the Einstein equations, but should be accepted as true for some other reason - for example, because they are regarded as a fundamental fact about ontology. From this point of view the reasoning of ref [53] is the wrong way round: it is precisely *because* there is a universal relation between entropy and area that the gravitational interaction between matter fields must act in such a way as to cancel the dependence on the number of fields, and indeed this can be regarded as simply an alternative version of the proof that a universal entropy-area relation entails the existence of gravitational phenomena.

Indeed, explaining gravitational phenomena by appeal to an entropy-area relation rather than vice versa arguably leads to a better explanation. For if we start from the Einstein equations and give an explanation similar to the one in ref [53], we have to conclude that it just happens to turn out that gravitational effects manage to cancel the dependence on the number of species, in a way that one might regard as suspiciously coincidental. The coincidence disappears if we run the explanation in the other direction, with gravity required to work in such a way as to respect the universality of the entropy bound. Moreover, if we start from the Einstein equations and give an explanation similar to the one in ref [53], we have to accept that a large range of physical mechanisms are responsible for protecting the entropy bound in different circumstances: '*they differ according to the physical situation studied, and they can involve combinations of different effects more reminiscent of a conspiracy than of an elegant mechanism.*' [14] This piecewise explanation seems quite unsatisfactory, whereas we can give a more unifying explanation if we see the entropy bound as a fundamental feature of the ontology and derive all of these effects from it.

Therefore we do not think it can be taken for granted that the correct direction of explanation is from gravitational phenomena to the covariant bounds rather than vice versa. Let us now consider whether the epistemic and ontological views of the covariant bound can shed any light on the matter.

5.2 Epistemic View

If we take an epistemic view of the covariant bound, we know that the gravitational phenomena do not explain the covariant bound, because the bound is already fully explained by the facts about how much information observers can obtain out of various regions of spacetime. So, if we accept that there is an explanatory relation between the covariant bound and gravitational phenomena, we can straightforwardly conclude that gravitational phenomena must be explained by the bound. And furthermore, since it is plausible that scientific explanations are usually transitive (as argued by Lange in ref [54]), it would seem to follow that gravitational phenomena must be explained by the facts about how much information observers can obtain out of various regions of spacetime. This provides an answer to the question of how we could possibly have obtained a bound which is really just an epistemic restriction on the amount of information we can obtain out of various regions of spacetime, even though we arrived at this bound by consideration of gravitational effects rather than explicitly counting the number of states in various regions. For if we understand gravitational phenomena as a consequence of the same epistemic restrictions which give rise to the covariant bound itself, then reasoning based on gravitational effects (including black holes, focussing and gravitational collapse) can be expected to lead to the same results as we would have obtained had we explicitly counted states.

Now, it is likely that explaining gravitational phenomena as a consequence of an epistemic restriction in this way would invite criticisms similar to Albert's emphatic reaction to the idea that the thermodynamic entropy could be epistemic: '*Can anybody seriously think that it is somehow* ***necessary*** *. . . that the particles that make up the material world must arrange themselves in accord with* ***what we know****, with* ***what we happen to have looked into?*** *Can anybody seriously think that our merely being ignorant of the exact microconditions of thermodynamic systems plays some part in* ***bringing it about****, in* ***making it the case****, that (say)* ***milk dissolves in coffee****? How could that* ***be****?*' [55]

However, in response to this kind of criticism it is important to recall that, as noted in section 4, 'epistemic' is not the same as 'subjective.' If indeed it is the case that only a certain amount of information about a system can emerge through its bounding surface, that is an objective fact about reality and not just an illusion: so if gravitational phenomena are to be understood as a consequence of limitations on the amount of information that can pass through a surface, then they are indeed an objective fact about a certain regime of reality. From this point of view gravitational focusing is in a sense an illusion, but the fact that all realistic observers are subject to that illusion is very much objective.

One might also criticize the claim that gravitational phenomena are explained by epistemic limitations on the grounds that the argument we have given leans too heavily on some questionable assumptions about explanation: in addition to assuming the transitivity of scientific explanation, it also implicitly assumes that once we have a complete explanation of some physical phenomenon, we shouldn't posit another one. But many philosophers believe that explanation must be relativized to an audience or context of explanation [56, 57], so really this principle should say that we shouldn't posit another explanation *relative to the same audience or context of explanation*, and thus one might hope to avoid the conclusion by arguing that in one context of explanation the covariant bound is explained by epistemic limitations, and in another context of explanation the covariant bound is explained by gravitational phenomena, so there is no single context of explanation within which we can apply the transitivity of explanation to infer that gravitational phenomena are explained by epistemic limitations.

However, this response may be blocked by bypassing the transitivity argument and showing directly that epistemic limitations on the information obtainable out of various regions of spacetime can give rise to 'illusory' gravitational effects. To do so, we will give an argument inspired by some ideas put forward by Susskind early in the development of the holographic principle [1].Take a screen with area A, and an object O behind the screen with surface area A'; assume the surface postulate is correct, so the information on the surface of O has an upper bound proportional to A', and the information passing through the screen has an upper bound proportional to A. Following Susskind, we will also assume the constant of proportionality is the same for both the surface and screen and that information on the surface of O saturates the bound. Now let light rays carry the information about the state of the surface of the object towards the screen. Suppose all the information in some area element dA' is carried by a bundle of light rays B, which intersect the screen orthogonally and which have cross-sectional area dA by the time they reach the screen. Suppose that gravitational focusing does not occur for some particular area element dA'_x on the object, i.e. $\frac{d\theta}{d\lambda}$ is positive for the corresponding bundle B_x. Since θ is zero at the screen, this implies that θ is negative for B_x between the object and the screen, and hence the corresponding area element dA_x on the screen is smaller than dA'_x, so some of the information on the area element dA'_x will not be able to pass through the screen. Thus we may perhaps make the case that the *perceived* area of dA'_x will be less than its real area, since the external observer can only judge the area of dA'_x on the basis of the information that emerges about dA'_x through the screen; so an external observer will always perceive the area of any area element dA' on the surface of the object as being less than or equal to the corresponding area dA on the screen, and thus it will always look as though gravitational focusing has occurred. This argument suggests a way in which gravitational focusing could indeed somewhat illusory: it can be understood as a function of the fact that we must always perceive objects as having less surface area than the screens through which we observe them.

Of course, this argument needs further elaboration before it can be a real contender for an account of gravitational phenomena. First, more needs to be said about the idea that we assign areas to objects as a function of the amount of information that we receive about them. Arguably this could be a natural consequence of the surface principle, and indeed, an idea of this kind is espoused by Smolin in ref [13]; similarly, approaches such as ref [58] which aim to derive spacetime structure from entanglement postulate a close connection between the areas we assign and the information we have access to. But the details are yet to be fully worked out - in particular, since the scale of the discretization is presumably at the Planck scale, we would need to understand how this Planck-scale effect could be responsible for our observation of gravitational focusing occurring at scales much larger than the Planck scale. Second, there is a need to understand the consequences of this way of thinking about gravitational focusing for other gravitational phenomena and also for various non-gravitational phenomena. After all, gravity is not epiphenomenal, and thus we can't simply declare gravity to be a consequence of our epistemic limitations without also declaring many other things to be a consequence of our epistemic limitations: for example, when an egg is brought into contact with the ground as a result of a gravitational force acting on it, the result is a significant alteration to all of the egg's properties, not merely its area, so if we believe that gravitational forces are only a manifestation of our epistemic limitations, then presumably all of the egg's other properties would also have to be regarded as manifestations of our epistemic limitations - the egg doesn't 'really' smash, it just appears to smash because we are unable to extract any further information about what the constituents of the egg are really doing. And the same goes for any other entities which are subject to a gravitational interaction (which in fact includes *all* known entities, since the gravitational force is universal). So accepting that gravity is really an 'illusion,' as the epistemic account seems to suggest, would potentially force us to accept that there are similar 'illusory' elements in many non-gravitational phenomena as well.

Nonetheless, this line of argument does suggest intriguing possibilities for explaining gravitational phenomena as consequences of epistemic restrictions. Thus it both explains how gravitational phenomena could lead us to a purely epistemic bound, and also further reinforces our conclusion that there is value in thinking explicitly about the possibility of additional degrees of freedom beyond the covariant bound, rather than merely disregarding them on pragmatic grounds.

5.3 Ontological View

The question of whether gravitational phenomena are explained by the covariant bound or vice versa in an ontological picture will ultimately hinge on the nature of the ontological picture one adopts. But as a concrete example, suppose we adopt an ontological view where the ontology is composed entirely of bounding surfaces. In that case, the covariant bound is explained by the fact that light-sheets are nothing more than an alternative encoding of the information on bounding surfaces, and then gravitational phenomena are explained by the ontology of light-sheets: geodesics must be focussed by matter because otherwise the light-sheets would contain more information than their bounding surfaces, which is impossible if the light-sheets are just alternative representations of bounding surfaces.

Smolin proposes something of this kind when he describes the significance of his own formulation of the entropy bound: '*Its role is to constrain the quantum causal structure of a quantum spacetime in a way that connects the geometry of the surfaces on which measurements may be made with a measure of the information that those measurements may produce ... the notion of area is reduced fundamentally to a measure of the flow of quantum information*' [13]. It seems that Smolin is imagining something like the following: spacetime should be understood as a construction - presumably on the part of human observers, or subconscious information-processing going on in their brains - whose geometry and dimensionality is determined by the nature of the information that the observers receive from the outside world, which need not itself be spatiotemporal in the ordinary sense. So in this picture, it is indeed the case that light-sheets are essentially just representations of the information on surfaces. Markopoulou and Smolin [11] have constructed a model of this kind, based on information flowing in a network of 'holographic screens' which take the form of a causal set, with screens composed by a quadruple of events. This approach implements the basic idea of the ontological view of the bound, i.e. that '*all observables in a quantum theory of cosmology are associated with two-surfaces, and represent information reaching a surface from its causal past.*' Markopoulou and Smolin intend their model only as an implementation of the surface postulate (the area of each screen is by definition proportional to its Hilbert space) but one might hope that a similar approach could ultimately be used to implement the light-sheet postulate as well.[14]

It should be noted that the derivation of the Einstein equations from the covariant bound relies on various approximations about entropy (particularly that it can be regarded as a fluid) and it is likely that these approximations do not hold at the most fundamental level. Moreover, the classical focussing theorem does not carry over precisely to quantum mechanics, because it relies on the null energy condition, which can be violated by physically reasonable states in a quantum field theory[15]. Thus if gravitational phenomena are really to be explained as a consequence of the covariant bound, it is likely that gravity and spacetime structure should be thought of as not fundamental but emergent - the fundamental description of reality pertains to the underlying degrees of freedom, and the curvature of spacetime emerges in the limit as the entropy-fluid approximation and the classical focussing theorem become valid, as a kind of higher-level description of the way in which these degrees of freedom depend on each other. The ontological picture thus falls in line with a long tradition of proposals about 'emergent gravity' and 'emergent spacetime' within both

[14] Markopolou and Smolin actually express scepticism about the possibility of implementing something like the light-sheet postulate in a background independent formulation. However, it's possible that the difficulties they raise are due to the fact that their screens do not play quite the same role as the screens employed by the covariant bound - in Markopoulou and Smolin's construction both sides of a surface count as a 'screen,' whereas the covariant bound tells us that usually only one side of a surface counts as a screen and indeed the whole point of the light-sheet construction is to decide which side of a surface is the 'inside' and which is the 'outside' so we know which region of space the entropy bound refers to (for example if the surface is a closed sphere, the 'inside' turns out to be the interior of the sphere, as one might expect). Therefore we would expect that in a holographic theory which reflects the role of surfaces as encoded in the covariant bound, only one side of a surface would count as a 'holographic screen.' Smolin argues against this approach [13] on the grounds that in the context of quantum cosmology there cannot be any well-defined way of picking out one side of the screen rather than another. However, this argument holds only if one assumes that the area, convergence and volume are all defined by quantum operators, whereas the special epistemic role of screens requires them to be defined only in the semiclassical limit in which observers interact with them, so we may not need to formulate them in terms of quantum operators at all.

[15] Bousso et al [51] have formulated a generalized 'quantum focussing conjecture' which then allows them to arrive at a quantum version of the covariant entropy bound. However, it is less straightforward to give this conjecture a transparent physical interpretation, since what is 'focussed' here is a combination of the area and the von Neumann entropy, which does not obviously correspond to the structure of spacetime. In any case, if one is willing to accept the picture of spacetime and gravity as emergent, we don't need such an interpretation - what is important is that the quantum focussing bound gives rise to the classical one in the appropriate limit, so that classical spacetime will indeed emerge as needed.

physics [52, 59, 60] and philosophy [61][16].

We note that Jacobson himself described his result as a demonstration that gravity is 'thermodynamical.' However, Chirco et al [53] argue that this interpretation is incorrect, because the entropy used in Jacobson's proof pertains only to the microscopic degrees of freedom of the quantum gravitational field - it is not the entropy of some hitherto unrecognised degrees of freedom of spacetime itself. Our version of the proof is in accordance with this conclusion: the entropy featuring in the derivation is simply the entropy of whatever is described by the ordinary stress-energy tensor of general relativity, not the entropy of spacetime itself. Moreover the entropy is only being used as a convenient way of estimating the true number of degrees of freedom in the relevant region, so the thermodynamic nature of the entropy is not a particularly crucial point. Thus although the possibility of deriving gravitationalphenomena from the covariant bound does suggest that gravity and spacetime may be emergent, the approach offers no reason to think they are particularly thermodynamical, except insofar as thermodynamics is an illustrative example of emergence.

6 Philosophical issues

Having examined the empirical evidence for the different possible views of the covariant entropy bound, we will now consider if there are philosophical arguments which might support one view or another. Philosophical arguments are relevant here because the epistemic and ontological views are naturally associated with certain metaphysical pictures. In particular, an epistemic view of the bound follows naturally if we are committed to a view that might be described as 'orthodox reductionism,' (i.e. the idea that spacetime is completely filled with autonomous degrees of freedom from which macroscopic phenomena emerge in some appropriate limit) because the assumption of autonomy at each spacetime point leads naturally to the idea that there may be information inside any given spacetime region which is unable to get out of the region. Conversely an ontological view of the bound requires us to accept that the content of a spacetime region is somehow dependent on its bounding surfaces in the past and future, which is hard to reconcile with the orthodox reductionist picture - it seems to entail that the degrees of freedom at different spacetime points are not generically autonomous, since they depend in some sense on higher-level phenomena pertaining to bounding surfaces. So it seems that an orthodox reductionist would have little option but to take an epistemic view.

Now, one might feel that there is something worrying about this dialectic. A metaphysical conviction (i.e. orthodox reductionism) combined with certain empirical results (i.e. the evidence that spacetime is discretized and/or that information flux through surfaces is bounded) apparently compels us to postulate a realm of unobservable physics which we can never observe or learn anything definite about. It may seem worrying that we end up committed to a view so ontologically excessive - from a naturalistic point of view, one might argue that rather than simply accepting the existence of invisible, undetectable physics, we should consider abandoning the metaphysical convictions that led to this conclusion. A historical comparison may help make the point: in the case of the ether hypothesis, it was the metaphysical conviction that light must be propagated in some substance which forced scientists of the day to postulate an invisible omnipresent substance which was undetectable in all experiments, and in hindsight it is easy to see that the right thing to do was to abandon the metaphysical conviction.

That said, some care must be taken here, for the reason to get rid of the ether was not simply the fact that it was not directly detectable - we can have good theoretical reasons to believe in things that are not directly detectable. Rather, the ether was ultimately discarded because it served no theoretical function: there was no reason to believe in it other than the metaphysical conviction that light must be propagated in a substance, so this metaphysical conviction was forcing scientists of the day to make their theories more complex for no increase in explanatory or predictive power. This suggests that in order to decide whether or not we should accept the unobservable degrees of freedom postulated by an epistemic account of the bound, it is important to assess what theoretical function they might serve. Evidently they do not *currently* serve any theoretical function, since our best current theories obey the covariant bound and thus they don't include any unobservable degrees of freedom, but that leaves open the question of whether we could increase the explanatory or predictive power of our theories by taking these degrees of freedom into account.

[16] It should be emphasized that there are two distinct kinds of emergence proposals: some approaches, like loop quantum gravity [6] and causal set theory [9], postulate that spacetime emerges from some underlying substratum as a part of the process of quantizing gravity, whereas others, like Jacobson and Verlinde, suggest that because gravity is emergent it should not be quantised at all. The approach suggested here seems more naturally associated with the second kind of approach, but we defer further discussion of the differences between these approaches to Part II of this paper.

On the one hand, we saw in section 4.1 that the existence of these degrees of freedom could potentially explain some features of our current physics. On the other hand, no *specific* details about those degrees of freedom can possibly be explanatory, since ex hypothesi the world outside the relevant surfaces cannot depend on those details in any way. And one might question whether an entity really has theoretical power if only its existence and no specific features of its state or configuration are relevant to the predictions of the theory. So arguably the *literal* understanding of these degrees of freedom required by orthodox reductionism is indeed not serving any theoretical function - we would perhaps be better off thinking of these degrees of freedom as something like gauge, in that their existence is a necessary feature of our representation of the theory but their specific configuration is unimportant. That said, ultimately an answer to this question must await a better understanding of what a theory which explicitly includes unobservable degrees of freedom might look like - it's possible that the existence of specific states and configurations of these degrees of freedom might play some crucial theoretical function that is not immediately evident in advance of writing down the theory, and if that were the case then we would clearly be justified in being committed to their existence despite the fact that they are not directly detectable.

On the other hand, abandoning reductionism comes with its own conceptual problems. In particular, when A can be derived from B but also B can be derived from A, how do we know the correct direction of explanation? Orthodox reductionism provides a simple answer - the smaller always explains the larger and not vice versa - so if we are moving away from reductionism we will need alternative ways to settle these questions. We saw something of this dialectic in the discussion of Jacobson's derivation in section 5.3: the argument of ref [53] implicitly invokes the reductionist intuition that higher-level effects like the universality of the entropy-area relation must be given constructive explanations based on smaller-scale phenomena, such as the gravitational behaviour of fundamental fields, and yet a supporter of the view that entropy bounds explain gravitational phenomena would presumably want to insist that the gravitational behaviour of the fundamental fields is a consequence of the universal entropy bound, which would mean that a higher-level feature of reality actively constrains the behaviour of the smaller-scale entities. If this kind of reasoning is to become accepted in science, we will need better criteria for judging the fitness of non-reductionist explanations and for deciding the correct direction of explanation in disputed cases.

7 Conclusion

So, assuming that the covariant bound is in some sense correct at least at the semiclassical level, is it epistemic or ontological? For those who are happy to take a purely operational view where we aim only to describe the degrees of freedom which are accessible to observers, this question may not be very important, but for those who hope to do cosmology, describe the universe as a whole or simply understand what is really going on, the issue must be resolved.

We have seen that on the one hand, the thermodynamic arguments for the bound can easily be understood in an epistemic way, but on the other hand the arguments based on gravitational focussing don't seem particularly epistemic, and thus to maintain the epistemic interpretation it would seem that we would have to adopt quite a radically revisionary view which sees gravity itself as being in some sense epistemic. Thus in a sense, entropy bounds may be regarded as indications that physics is starting to come up against the limits of reductionism. This provides motivation for an ontological view, but on the other hand we have argued that an ontological view will likely also be quite revisionary in its account of the nature of spacetime - in order to avoid being undermined by the possibility of an epistemic view, ontological approaches must show what is wrong with the assumptions underpinning the epistemic account, which will most likely involve proposing an alternative ontology which is a better fit to the nature of the data available to us, such as an ontology based on surfaces rather than autonomous degrees of freedom in a three-dimensional space.

The issues discussed in this article also have interesting consequences for philosophy of science more generally. For we have emphasized the way in which the physics we arrive at is limited by features of our epistemic situation, and the 'epistemic' interpretation of the covariant bound presents us with a vision of a world in which these limits are actually very severe, so significant portions of reality may be forever beyond our grasp. This poses some new challenges for scientific realism. First, how do we distinguish between parts of our theories which reflect features of reality and parts of our theories which are just consequences of our epistemic situation? And having done so, what kind of attitude should we take to parts of physics that we suspect are in fact consequences of our epistemic situation? For example, if we conclude that the entropy bound is just epistemic, it is certainly a part of reality and in that sense is compatible with scientific realism, and yet it would seem that we should not adopt a naively realist attitude towards its proposed count of degrees of freedom, since that count is known to be objectively wrong. We hope to explore some of these issues in future work.

8 Acknowledgements

Thanks to Jacob Barandes, Niels Linneman, and the UWO philosophy of physics reading group for very helpful comments on this article. This publication was made possible through the support of the ID 61466 grant from the John Templeton Foundation, as part of the "The Quantum Information Structure of Spacetime (QISS)" Project (qiss.fr). The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

A Deriving Einstein's Equations from an Entropy Bound

Here we make a small modification to Jacobson's 1995 derivation of the Einstein equations from a horizon entropy bound, in order to argue that the covariant entropy bound can be used to derive Einstein's equations. (We note that Jacobson himself advocates a version of the surface postulate rather than the light-sheet postulate, so he presumably would not endorse this modification of his proof).

Proof. At a given point in spacetime, consider a 2-surface Ω of area A whose past directed null normal congruence $\{k^a\}$ to one side has vanishing expansion and shear at the surface (as noted by Jacobson, it is always possible to choose such a surface provided it is sufficiently small). The past-directed null normal congruence has negative or zero expansion, and thus is a light-sheet L; we parametrize the rays $\{k^a\}$ on L with an affine parameter λ which vanishes at A. Because Ω is small, we will take it that the Ricci tensor R_{ab} and the stress-energy tensor T_{ab} are approximately constant in the neighbourhood of Ω.

Now let Ω' be the infinitesimally smaller surface obtained from Ω by moving a distance $d\lambda$ along each light-ray. Ω' has area A' and bounds the light-sheet L' which is infinitesimally smaller than L. We will make a locality assumption to the effect that the curvature of spacetime in the region sandwiched by Ω' and Ω depends only on the contents of this region, and in particular, it is independent of the total entropy on L'. Thus the curvature of spacetime between Ω' and Ω must be large enough that the entropy bound is satisfied even if the entropy on L' takes its maximal value, $\frac{A'}{4\pi G}$. This entails that the total entropy between Ω' and Ω must be less than or equal to $\frac{A-A'}{4G\hbar}$.

Note that this expression can be recognised as the generalized version of Bousso's bound specified in ref [18], where we cut off the light-sheet associated with a surface of area A before it terminates. This bound is also known to be true in many circumstances, although not universally - its validity depends on the assumption that entropy can be regarded as something like a locally-defined fluid, so it will not necessarily hold in scenarios where the nonlocal nature of the entropy becomes important, where of course our locality assumption cannot be expected to hold.

Let $\theta(x)$ denote the expansion of the light-rays in L at value x of the affine parameter λ; thus $A - A' = -\int_{\Omega'} \theta(d\lambda) dA d\lambda$.

Now we may employ the Raychauduri equation. Having chosen Ω such that the shear on it is zero, we have that $\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - R_{ab}k^a k^b$, where R_{ab} is the Ricci tensor and k is a unit vector field which at every point in L is tangent to the light-ray in L passing through that point, oriented from Ω towards Ω'. Since θ is zero on the surface Ω, we can work in the limit of small θ, and then since we have assumed that the Ricci tensor is approximately constant over the region, we infer that everywhere on Ω' we have $\theta = -R_{ab}k^a k^b d\lambda$.

Hence $A - A' = -\int_{\Omega'} -R_{ab}k^a k^b d^2\lambda dA = A' R_{ab}k^a k^b d^2\lambda$.

Now we must calculate the entropy between Ω' and Ω. It is evident that the entropy should be connected to the stress-energy tensor in some way; roughly speaking, higher energy systems have more possible states and hence higher entropy. But writing down an explicit expression will require some approximations. Here we will follow Flanagan, Marolf and Wald [18] in assuming that the entropy can be expressed as an entropy flux $s^a(x)$, so that the entropy between Ω' and Ω is given by $S = \int_{x\in\Omega'} s^a(x) k^a dx d\lambda$. As argued by ref [18], this is a good approximation in many regimes, though not all. Bousso writes: '*Generally speaking, the notion of an entropy flux assumes that entropy can* *be treated as a kind of local fluid. This is often a good approximation, but it ignores the non-local character of entropy and does not hold at a fundamental level*' [26].

Moreover, from thermodynamics we have that $dS = \frac{dQ}{T}$, so we will take it that $s^a(x)$ can be written in the form of a heat flux divided by a temperature. We are interpreting S as an upper bound on the total amount of information in the region, and therefore we can do this calculation from the point of view of any possible observer in the region, because S must be at least as great as the entropy seen by any observer. Thus we may select an observer of acceleration κ whose Rindler horizon coincides approximately with L in this region. This observer sees a temperature given by the Unruh temperature $\frac{\hbar\kappa}{2\pi}$, as well as an energy flux given by the boost-energy current of matter, $T_{ab}\chi^b$, where χ^b is the (approximate) Killing vector field generating Lorentz boosts in the Rindler frame, which will coincide with a suitably defined null vector on the light-sheet. As noted by ref [62], in any spacetime, around any event, there exists a class of (possible) local Rindler observers who will perceive this boost-energy current and Unruh temperature, so we can be sure that such a possible observer will indeed be well-defined in the relevant region.

As explained in ref [63], χ^b and k^b can be related as follows. Since we are performing an infinitesimal calculation to first order only, we can work on a stationary background. Then we have $\lambda k^a = \lambda(\frac{dv}{d\lambda})\chi^a$, where v is the Killing parameter. In general the relation between affine and Killing parameters on a Killing horizon is $\lambda = ae^{\kappa v}$, where a and b are arbitrary constants, and b can be chosen to be zero since we are free to shift the affine parameter. So we find that $\lambda\frac{dv}{d\lambda} = \frac{1}{\kappa}$ and hence we conclude that $\chi^a = \lambda\kappa k^a$; thus at Ω' we have that $\chi^a = d\lambda\kappa k^a$.

Thus we infer that $S = 2\pi\int_{\Omega'}\frac{T_{ab}k^ak^b}{\hbar}dAd^2\lambda$. Since we have assumed that the stress-energy tensor is approximately constant over the region, we conclude that $S = 2\pi A'\frac{T_{ab}k^ak^b}{\hbar}d^2\lambda$. Thus in order to have $S \leq \frac{A-A'}{4G\hbar}$, we require:

$$2\pi A'\frac{T_{ab}k^ak^b}{\hbar}d^2\lambda \leq A'\frac{R_{ab}k^ak^b}{4G\hbar}d^2\lambda$$

and thus $8\pi GT_{ab}k^ak^b \leq R_{ab}k^ak^b$.

Let us now assume that the the bound is in fact saturated - i.e. space is always focussed exactly enough to ensure that the entropy bound is always satisfied, and no more. Thus $8\pi GT_{ab}k^ak^b = R_{ab}k^ak^b$. As noted by Jacobson, this implies that $8\pi GT_{ab} = R_{ab} + fg_{ab}$ for some f. Local conservation of energy and momentum implies that T_{ab} is divergence free and therefore, using the contracted Bianchi identity, that $f = -\frac{R}{2} + \Lambda$ for some constant Λ. So we get $8\pi GT_{ab} = R_{ab} - \frac{Rg_{ab}}{2} + \Lambda g_{ab}$ which is the Einstein equation.

□

A.1 Limitations of the derivation

The derivation given above has a number of limitations. In particular, although it is clear that entropy must be related to the stress-energy tensor in some way, there is no universally valid expression giving the relation between them, so the 'entropy flux' expression used in this calculation will not be valid in all situations - for example, it may fail to hold when the non-local character of entropy is relevant, or when matter is not in thermodynamic equilibrium. One might also question whether the Unruh temperature used in the calculation of the entropy is the right temperature to use here - what if the matter in the region between S and S' is just ordinary classical matter with a normal thermodynamical temperature not associated with any acceleration? Perhaps the Unruh temperature should be thought as encoding something to do with the maximum amount of entropy which could possibly be present for a given value of T_{uv} (since after all it includes Planck's constant) but the details are hazy. But nonetheless, it seems reasonable to think that a relation of a similar form must hold in other scenarios, since it is clear that entropy should be linked to the stress-energy tensor, so it seems plausible to think the argument should generalize. (Note that Jacobson did subsequently generalize his argument to the case of non-equilibrium thermodynamics [64] so one might think that similar strategies will work here).

Also, we assumed at the final stage that the bound $8\pi GT_{ab}k^ak^b \leq R_{ab}k^ak^b$ is in fact saturated. This assumption is somewhat questionable because we justified our entropy calculation on the basis that the entropy observed by the Rindler observer must be a lower bound on the true number of degrees of freedom in the region - we have not ruled out the possibility that some other observer could see more entropy in the region, and if that were the case the true number of degrees of freedom would have to be larger. However, note that the specific choice of acceleration κ cancels out during the entropy calculation, so the same calculation will work for any accelerated observer in the same location; so if one assumes that non-accelerated observers necessarily see less entropy than accelerated observers, one may argue that the entropy observed by the Rindler observer also provides a suitable *upper bound* in this situation.

In addition, we have assumed that the entropy inside the region is given only by the entropy of matter; but in principle there should also be entropy associated with gravitational degrees of freedom. Flanagan et al make this point in ref [18], but they then leave out gravitational contributions because it's unclear how to calculate them. However, one might worry that once gravitational degrees of freedom are included, a derivation of this kind will potentially be subject to some kind of infinite regression, and thus leaving out gravitational degrees of freedom is a non-trivial approximation. In Part II of this paper we will look in more detail at the question of whether or not an entropy bound can still hold once gravitational degrees of freedom become nontrivial.

References

[1] Leonard Susskind. The world as a hologram. *Journal of Mathematical Physics*, 36(11):6377–6396, 1995.

[2] Jacob D. Bekenstein. Black holes and entropy. *Phys. Rev. D*, 7:2333–2346, Apr 1973.

[3] S. W. Hawking. Particle creation by black holes. *Communications in Mathematical Physics*, 43(3):199 – 220, 1975.

[4] Jacob D. Bekenstein. Universal upper bound on the entropy-to-energy ratio for bounded systems. *Phys. Rev. D*, 23:287–298, Jan 1981.

[5] Ning Bao, Sean M. Carroll, and Ashmeet Singh. The hilbert space of quantum gravity is locally finite-dimensional. *International Journal of Modern Physics D*, 26(12):1743013, oct 2017.

[6] C. Rovelli and F. Vidotto. *Covariant Loop Quantum Gravity: An Elementary Introduction to Quantum Gravity and Spinfoam Theory*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2015.

[7] Carlo Rovelli. *Quantum Gravity*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2004.

[8] Rodolfo Gambini and Jorge Pullin. Discrete space-time, 2005.

[9] R. Sorkin. Geometry from order: Causal sets. *Einstein Online*, 2006.

[10] Juan Maldacena. *International Journal of Theoretical Physics*, 38(4):1113–1133, 1999.

[11] Fotini Markopoulou and Lee Smolin. Holography in a quantum spacetime, 1999.

[12] Sanved Kolekar and T. Padmanabhan. Holography in action. *Physical Review D*, 82(2), jul 2010.

[13] Lee Smolin. The strong and weak holographic principles. *Nuclear Physics*, 601:209–247, 2001.

[14] Raphael Bousso. A covariant entropy conjecture. *Journal of High Energy Physics*, 1999(07):004–004, jul 1999.

[15] Abhay Ashtekar and Edward Wilson-Ewing. Covariant entropy bound and loop quantum cosmology. *Physical Review D*, 78(6), sep 2008.

[16] John Dougherty and Craig Callender. Black Hole Thermodynamics: More Than an Analogy?, October 2016.

[17] David Wallace. The case for black hole thermodynamics, part i: phenomenological thermodynamics, 2017.

[18] É anna É. Flanagan, Donald Marolf, and Robert M. Wald. Proof of classical versions of the bousso entropy bound and of the generalized second law. *Physical Review D*, 62(8), sep 2000.

[19] Raphael Bousso, Horacio Casini, Zachary Fisher, and Juan Maldacena. Proof of a quantum bousso bound. *Phys. Rev. D*, 90:044002, Aug 2014.

[20] Raphael Bousso, Horacio Casini, Zachary Fisher, and Juan Maldacena. Entropy on a null surface for interacting quantum field theories and the bousso bound. *Physical Review D*, 91(8), apr 2015.

[21] Wayne C. Myrvold. The science of $\Theta\Delta^{cs}$. *Foundations of Physics*, 50(10):1219–1251, Aug 2020.

[22] Wayne C. Myrvold. Statistical mechanics and thermodynamics: A maxwellian view, May 2011.

[23] Meir Hemmo and Orly Shenker. Von neumann?s entropy does not correspond to thermodynamic entropy. *Philosophy of Science*, 73(2):153–174, 2006.

[24] Carina E. A. Prunkl. On the equivalence of von neumann and thermodynamic entropy. *Philosophy of Science*, 87(2):262–280, 2020.

[25] Eugene Y. S. Chua. Does von neumann's entropy correspond to thermodynamic entropy?, December 2019. Forthcoming at Philosophy of Science. Copyright The Philosophy of Science Association 2020. Preprint (not copyedited or formatted). Please use DOI when citing or quoting. DOI: https://doi.org/10.1086/710072.

[26] Raphael Bousso. The holographic principle. *Reviews of Modern Physics*, 74(3):825–874, aug 2002.

[27] Ted Jacobson. On the nature of black hole entropy. In *Eighth Canadian conference on general relativity and relativistic astrophysics*. ASCE, 1999.

[28] Carlo Rovelli. Black holes have more states than those giving the bekenstein-hawking entropy: a simple argument, 2017.

[29] Rafael D. Sorkin. The statistical mechanics of black hole thermodynamics. 1997.

[30] Robert Oeckl. A "general boundary" formulation for quantum mechanics and quantum gravity. *Physics Letters B*, 575(3):318–324, 2003.

[31] Louis Crane. Clock and category: Is quantum gravity algebraic? *Journal of Mathematical Physics*, 36(11):6180–6193, 1995.

[32] Alex May. Quantum tasks in holography. *Journal of High Energy Physics*, 2019(10):233, October 2019.

[33] G. 't Hooft. Dimensional reduction in quantum gravity, 1993.

[34] Dennis Dieks. Quantum mechanics and perspectivalism, 2018.

[35] Christopher A. Fuchs. On participatory realism, 2016.

[36] W. Heisenberg. Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik. *Zeitschrift fur Physik*, 43(3-4):172–198, March 1927.

[37] Raphael Bousso. Flat space physics from holography. *Journal of High Energy Physics*, 2004(05):050–050, may 2004.

[38] Alejandro Perez, Daniel Sudarsky, and James D. Bjorken. A microscopic model for an emergent cosmological constant. *International Journal of Modern Physics D*, 27(14):1846002, oct 2018.

[39] R. W. Spekkens. Evidence for the epistemic view of quantum states: A toy theory. *Phys Rev A*, 75(3):032110, March 2007.

[40] Carlo Rovelli. Relational quantum mechanics. *International Journal of Theoretical Physics*, 35(8):1637–1678, aug 1996.

[41] Gary T. Horowitz. The Origin of black hole entropy in string theory. *Astrophys. Space Sci. Libr.*, 211:46, 1997.

[42] Carlo Rovelli. Black hole entropy from loop quantum gravity. *Physical Review Letters*, 77(16):3288–3291, oct 1996.

[43] Tullio Regge and Claudio Teitelboim. Role of surface integrals in the hamiltonian formulation of general relativity. *Annals of Physics*, 88(1):286–318, 1974.

[44] Federico Laudisa and Carlo Rovelli. Relational Quantum Mechanics. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Winter 2021 edition, 2021.

[45] C. Rovelli. Relational quantum mechanics. *International Journal of Theoretical Physics*, 35:1637–1678, August 1996.

[46] Emily Adlam and Carlo Rovelli. Information is physical: Cross-perspective links in relational quantum mechanics, 2022.

[47] A. Kent. Solution to the Lorentzian quantum reality problem. *Phys Rev A*, 90(1):012107, July 2014.

[48] Adrian Kent. Lorentzian quantum reality: postulates and toy models. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 373(2047):20140241, Aug 2015.

[49] Adrian Kent. Quantum reality via late-time photodetection. *Physical Review A*, 96(6), Dec 2017.

[50] George F. R. Ellis. On the Raychaudhuri equation. *Pramana*, 69(1):15, July 2007.

[51] Raphael Bousso, Zachary Fisher, Stefan Leichenauer, and Aron C. Wall. Quantum focusing conjecture. *Phys. Rev. D*, 93:064044, Mar 2016.

[52] Ted Jacobson. Thermodynamics of spacetime: The einstein equation of state. *Physical Review Letters*, 75(7):1260–1263, aug 1995.

[53] Goffredo Chirco, Hal M. Haggard, Aldo Riello, and Carlo Rovelli. Spacetime thermodynamics without hidden degrees of freedom. *Physical Review D*, 90(4), aug 2014.

[54] Marc Lange. Transitivity, self-explanation, and the explanatory circularity argument against humean accounts of natural law. *Synthese*, 195(3):1337–1353, 2018.

[55] David Z. Albert. *Time and Chance*. Harvard University Press, 2000.

[56] B.C. van Fraassen, Oxford University Press, and P.P.B.C. Van Fraassen. *The Scientific Image*. Clarendon Library of Logic and Philosophy. Clarendon Press, 1980.

[57] Henk W. de Regt. 87A Contextual Theory of Scientific Understanding. In *Understanding Scientific Understanding*. Oxford University Press, 11 2017.

[58] ChunJun Cao, Sean M. Carroll, and Spyridon Michalakis. Space from hilbert space: Recovering geometry from bulk entanglement. *Physical Review D*, 95(2), jan 2017.

[59] Erik Verlinde. Emergent gravity and the dark universe. *SciPost Physics*, 2(3), may 2017.

[60] Erik Verlinde. On the origin of gravity and the laws of newton. *Journal of High Energy Physics*, 2011(4), apr 2011.

[61] Eleanor Knox. Effective spacetime geometry. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(3):346–356, 2013.

[62] T. Padmanabhan. Exploring the nature of gravity, 2016.

[63] Ted Jacobson and Renaud Parentani. Horizon entropy. *Foundations of Physics*, 33(2):323–348, 2003.

[64] Christopher Eling, Raf Guedens, and Ted Jacobson. Nonequilibrium thermodynamics of spacetime. *Physical Review Letters*, 96(12), mar 2006.