CALT-TH-2017-013

Cosmic Equilibration: A Holographic No-Hair Theorem from the Generalized Second Law

Sean M. Carroll and Aidan Chatwin-Davies*
Walter Burke Institute for Theoretical Physics,
California Institute of Technology,
Pasadena, CA 91125

Abstract

In a wide class of cosmological models, a positive cosmological constant drives cosmological evolution toward an asymptotically de Sitter phase. Here we connect this behavior to the increase of entropy over time, based on the idea that de Sitter space is a maximum-entropy state. We prove a cosmic no-hair theorem for Robertson-Walker and Bianchi I spacetimes by assuming that the generalized entropy of a Q-screen ("quantum" holographic screen), in the sense of the cosmological version of the Generalized Second Law conjectured by Bousso and Engelhardt, increases up to a finite maximum value, which we show coincides with the de Sitter horizon entropy. We do not use the Einstein field equations in our proof, nor do we assume the existence of a positive cosmological constant. As such, asymptotic relaxation to a de Sitter phase can, in a precise sense, be thought of as cosmological equilibration.

* seancarroll@gmail.com, achatwin@caltech.edu

CONTENTS

I. Introduction 3

II. The generalized second law for cosmology 5

III. A cosmic no-hair theorem for RW spacetimes 9

IV. A cosmic no-hair theorem for Bianchi I spacetimes 13
- A. 1+2 dimensions 13
  - 1. Showing that $\mathcal{Q}$ squeezes into the coordinate origin $\chi = 0$ as $\eta \to \eta_\infty$. 17
  - 2. Showing that $a(\eta)$ is asymptotically de Sitter. 20
  - 3. Showing that the anisotropy decays. 22
- B. 1+3 dimensions 23

V. Discussion 27

Acknowledgements 28

A. Q-screens – a worked example 28

References 31

I. INTRODUCTION

Like black holes, universes have no hair, at least if they have a positive cosmological constant Λ [1–10]. A cosmic no-hair theorem states that, if a cosmological spacetime obeys Einstein's equation with $\Lambda > 0$, then the spacetime asymptotically tends to an empty de Sitter state in the future. A more precise statement is due to Wald, who proved the following theorem [1]:

Theorem I.1 (Wald) *All Bianchi spacetimes (except for certain type IX spacetimes) that are initially expanding, that have a positive cosmological constant $\Lambda > 0$, and whose matter content besides Λ obeys the strong and dominant energy conditions, tend to a de Sitter state in the future.*

Bianchi spacetimes are cosmologies that are homogeneous but in general anisotropic [11, 12]. For example, the metric of the $1+3$ dimensional Bianchi I spacetime in comoving Cartesian coordinates is given by

$$ds^2 = -dt^2 + a_1^2(t)\, dx^2 + a_2^2(t)\, dy^2 + a_3^2(t)\, dz^2. \tag{1}$$

It is essentially a Robertson-Walker (RW) spacetime in which the scale factor can be different in different directions in space. In this case, when the necessary conditions are satisfied, Wald's theorem implies that each $a_i(t)$ tends to the same de Sitter scale factor, $\exp(\sqrt{\Lambda/3}\, t)$ for a cosmological constant $\Lambda > 0$, as t tends to infinity.

The intuition behind why one would expect a cosmic no-hair theorem to hold is that as space expands, the energy density of ordinary matter decreases while the density of vacuum energy remains constant. As such, the cosmological constant eventually dominates regardless of the initial matter content and geometry, and a universe in which a positive cosmological constant is the only source of stress-energy is de Sitter. For Bianchi I spacetimes, one can make this intuition explicit by writing down a Friedmann equation for the average scale factor, $\bar{a}(t) \equiv [a_1(t)a_2(t)a_3(t)]^{1/3}$, which gives [13, Ch. 8.6]

$$\left(\frac{\dot{\bar{a}}(t)}{\bar{a}(t)}\right)^2 \propto \left(\rho_\Lambda + \rho_{\text{matter}} + \rho_{\text{an}}\right). \tag{2}$$

On the right hand side, ρ_Λ and ρ_{matter} denote the energy densities due to the cosmological constant and matter respectively, while ρ_{an} is an effective energy density due to anisotropy, similarly to how one can think of spatial curvature as an effective source of stress-energy. Crucially, ρ_{an} scales at most like $\bar{a}^{-2}$, and so as the universe expands, only the constant contribution due to ρ_Λ persists. The exception to Wald's theorem is the case of a Bianchi IX spacetime (which has positive spatial curvature) whose initial matter energy density is so high that the spacetime recollapses before the cosmological constant can dominate [1]. Intuitively, we expect not only anisotropies, but also perturbative inhomogeneities to decay away at late times, though this is harder to prove rigorously [2, 9, 14, 15]. Beyond classical general relativity, various generalizations of Wald's theorem attempt to demonstrate analogous no-hair theorems for the quantum states of fields on a curved spacetime background [16–18].

As the universe expands and the cosmological constant increases in prominence with respect to other energy sources, something else is also going on: entropy is increasing. According to the Second Law of Thermodynamics, the entropy of any closed system (such as the universe) will increase or stay constant, at least until it reaches a maximum value. It is interesting to ask whether there is a connection between these two results, the cosmic no-hair theorem and the Second Law. Can the expansion of the universe toward a quiescent de Sitter phase be interpreted as thermodynamic equilibration to a maximum-entropy state? It is well-established that de Sitter has many of the properties of an equilibrium maximum-entropy state, including a locally thermal density matrix with a constant temperature [19, 20], and the relationship between entropy and de Sitter space has been examined from a variety of perspectives [21–29].

In this paper we try to make one aspect of these ideas rigorous, showing that a cosmic no-hair theorem can be derived even without direct reference to Einstein's equation, simply by invoking an appropriate formulation of the Second Law. This strategy of deducing properties of spacetime from the behavior of entropy is reminiscent of the thermodynamic/entropic gravity program [30–34], as well as of the gravity/entanglement connection [35–42]. Though we do not attempt to derive a complete set of gravitational field equations from entropic considerations, it is interesting that a specific spacetime can be singled out purely from the requirement that entropy increases to a maximum finite value.

To derive our theorem, we require a precise formulation of the Second Law that is applicable in curved spacetime, and that includes the entropy of spacetime itself. A step in this direction is Bekenstein's Generalized Second Law (GSL), which notes that the entropy of a black hole with area A is given by $S_{\rm BH} = A/4G$, and adds the entropy of all black holes in a system to the ordinary thermodynamic entropy to define a generalized entropy $S_{\rm gen}$ that is conjectured to increase or remain constant over time [43]. Unfortunately this form of the GSL doesn't immediately help us in spacetimes without any black holes. Recently, Bousso and Englehardt proposed a cosmological version of the GSL [44], building on previous work on holography [45], apparent horizons [46–51], and holographic screens [52, 53]. They define a version of generalized entropy on a hypersurface they call a "Q-screen," a quantum version of a holographic screen, which in turn is a modification of an apparent horizon. Given a Cauchy hypersurface Σ and a codimension-2 spatial surface with no boundary $\sigma \subset \Sigma$ that divides Σ into an interior region and an exterior region, the generalized entropy is the sum of the area entropy of σ, i.e., its area in Planck units, and the entropy of matter in the exterior region:

$$S_{\rm gen}[\sigma, \Sigma] = \frac{A[\sigma]}{4G} + S_{\rm out}[\sigma, \Sigma]. \tag{3}$$

Bousso and Engelhardt's version of the GSL is the statesment that generalized entropy increases strictly monotonically with respect to the flow through a specific preferred foliation of a Q-screen:

$$\frac{dS_{\rm gen}}{dr} > 0\,, \tag{4}$$

where r parameterizes the foliation. Although it is unproven in general, this version of the

GSL is well-motivated and known to hold in specific circumstances (the discussion of which we defer to the next section).

In this work, we use the GSL to establish a cosmic no-hair theorem purely on thermodynamic grounds. In an exact de Sitter geometry, the de Sitter horizon is a holographic screen, and every finite horizon-sized patch is associated with a fixed entropy, proportional to the area of the horizon in Planck units [54]. We therefore conjecture that evolution toward such a state is equivalent to thermodynamic equilibration of a system with a finite number of degrees of freedom, and therefore a finite maximum entropy. Specifically, assuming the GSL, we show that if a Bianchi I spacetime admits a Q-screen along which generalized entropy monotonically increases up to a finite maximum, then the anisotropy necessarily decays and the scale factor approaches de Sitter behavior asymptotically in the future. At no point do we use the Einstein field equations, nor do we assume the presence of a positive cosmological constant. The GSL and that entropy tends to a finite maximum along the Q-screen take the logical place of these two respective ingredients.

The proof essentially consists of first showing that an approach to a finite maximum entropy heavily constrains the possible asymptotic structure of a Q-screen. Second, we show that the spacetime must necessarily be asymptotically de Sitter (and in particular, isotropic as well) in order to admit a Q-screen with the aforementioned asymptotic structure.

The structure of the rest of this paper is as follows. We review Q-screens and the GSL in Sec. II. In Sec. III, we first prove a cosmic no-hair theorem for the simpler case of RW spacetimes using the GSL. Then, in Sec. IV, we move on to the proof for Bianchi I spacetimes, first in $1+2$ dimensions to illustrate our methods, and then in $1+3$ dimensions, which also illustrates how to generalize to arbitrary dimensions. We discuss aspects of the theorems and their proofs as well as some implications in Sec. V.

II. THE GENERALIZED SECOND LAW FOR COSMOLOGY

We begin by briefly reviewing Bousso and Engelhardt's conjectured Generalized Second Law (GSL). The GSL can be thought of as a quasilocal version of Bekenstein's entropy law for black holes [43], but which also applies to cosmological settings. Moreover, the GSL is a natural semiclassical extension of Bousso and Engelhardt's area theorem for holographic screens in the same way that Bekenstein's entropy law extends Hawking's area theorem to evaporating black holes.

An early cornerstone of classical black hole thermodynamics [55, 56] was Hawking's area theorem: in all spacetimes which satisfy the null curvature condition, the total area of all black hole event horizons can only increase, i.e., $dA/dt \geq 0$ [57]. Of course, the area theorem fails for evaporating black holes, the technical evasion being that they do not satisfy the null curvature condition. Bekenstein pointed out, however, that if one instead interprets the area of the black hole event horizon as horizon entropy and includes the entropy of the Hawking radiation outside the black hole, S_{out}, in the total entropy budget, then the generalized entropy, $S_{\text{gen}} = A/4G + S_{\text{out}}$, increases monotonically or stays constant, $dS_{\text{gen}}/dt \geq 0$ [43].

From the perspective of trying to understand the thermodynamics of spacetime, however, both Hawking's and Bekenstein's results suffer from two inconveniences. First, they are fundamentally nonlocal, since identifying event horizons requires that one know the full structure of a Lorentzian spacetime. Second, these results only apply to black holes; it would be desirable to understand thermodynamic aspects of spacetime in other geometries as well. These considerations motivate holographic screens [52, 53], a subset of which obey a classical area theorem, as well as their semiclassical extensions called Q-screens [44], a subset of which are conjectured to obey an entropy theorem. Importantly, both holographic screens and Q-screens are quasilocally defined and are known to be generic features of cosmologies in addition to black hole spacetimes.

Let us first review holographic screens. Following the convention of Bousso and Engelhardt, here and throughout we will refer to a spacelike codimension-2 hypersurface simply as a "surface."

Let σ be a compact connected surface. At every point on σ, there are two distinct future-directed null directions (or equivalently, two distinct past-directed null directions) that are orthogonal to σ: inward- and outward-directed. The surface σ is said to be *marginal* if the expansion of the null congruence corresponding to one of these directions, say k^μ, is zero everywhere on σ. Consequently, σ is a slice of the null sheet generated by k^μ that locally has extremal area. This last point is particularly clear if one observes that the expansion, $\theta = \nabla_\mu k^\mu$, at a point $y \in \sigma$, can be equivalently defined as the rate of change per unit area of the area of the slice, $A[\sigma]$, when a small patch of proper area $\mathcal{A}$ is deformed along the null ray generated by k^μ at y with an affine parameter λ:

$$\theta(y) = \lim_{\mathcal{A}\to 0} \frac{1}{\mathcal{A}} \frac{dA[\sigma]}{d\lambda}\bigg|_y \tag{5}$$

This definition is illustrated in Fig. 1 below.

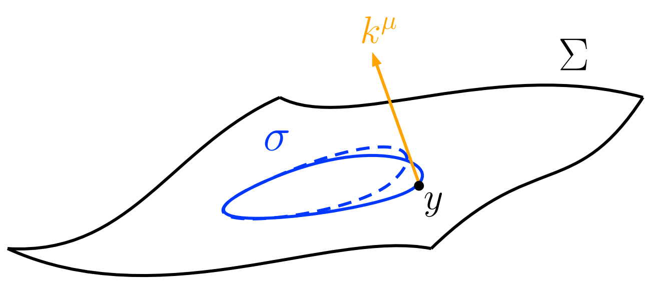

FIG. 1. Given a Cauchy hypersurface Σ, the surface $\sigma \subset \Sigma$ (drawn with a solid line) splits Σ into an interior and exterior. We define deformations of σ (drawn with a dotted line) by dragging σ along the null ray generated by k^μ at any point $y \in \sigma$. More precisely, a deformation is defined by transporting a small area element $\mathcal{A} \subset \sigma$ at y in the k^μ direction.

A *holographic screen* is a smooth codimension-1 hypersurface that can be foliated by marginal surfaces, which are then called its *leaves*. Note that while the leaves σ are spacelike, in general a holographic screen need not have a definite character. A marginal surface σ is said to be *marginally trapped* if the expansion of the congruence in the other null direction is negative everywhere on σ, and a *future holographic screen* is a holographic screen whose leaves are marginally trapped; marginally anti-trapped surfaces and past holographic screens are defined analogously. Then, assuming the null curvature condition as well as a handful of mild generic conditions, Bousso and Engelhardt proved that future and past holographic screens obey the area theorem paraphrased below [52, 53]:

Theorem II.1 (Bousso & Engelhardt) *Let H be a regular holographic screen. The area of its leaves changes strictly monotonically under the flow through the foliation of H.*

Q-screens are related to holographic screens, but with expansion replaced by what is dubbed the "quantum expansion." Let σ again denote a compact connected surface. The quantum expansion at a point $y \in \sigma$ in the orthogonal null direction k^μ is defined as the rate of change per unit proper area of the generalized entropy (3), i.e., the sum of both area and matter entropy, with respect to affine deformations along the null ray generated by k^μ:

$$\Theta_k[\sigma; y] = \lim_{\mathcal{A}\to 0} \frac{4G}{\mathcal{A}} \frac{dS_{\text{gen}}}{d\lambda}\bigg|_y \tag{6}$$

Then similarly to before, a *quantum marginal surface* is a surface σ such that the quantum expansion in one orthogonal null direction vanishes everywhere on σ. Just as a marginal surface locally extremizes area along a lightsheet, a quantum marginal surface locally extermizes the generalized entropy along the lightsheet generated by k^μ.

The adjective "quantum" can be confusing in this context. In this work it denotes a shift from classical general relativity, where one proves theorems about the area of surfaces, to quantum field theory on a semiclassical background, where analogous theorems refer to a generalized entropy that adds the entropy of matter degrees of freedom to such an area. That matter entropy may be be calculated as the quantum (von Neumann) entropy of a density operator, but in the right circumstances (which we will in fact be dealing with below) it is equally appropriate to treat it as a classical thermodynamic quantity. So here "quantum" should always be interpreted as "adding an entropy term to the area of some surface," whether or not quantum mechanics is directly involved.

The remaining constructions have similarly parallel definitions. A *Q-screen* is a smooth codimension-1 hypersurface that can be foliated by quantum marginal surfaces. A quantum marginal surface σ is *marginally quantum trapped* if the quantum expansion in the other null direction is negative everywhere on σ, and a *future Q-screen* is a Q-screen whose leaves are marginally quantum trapped, and analogously for anti-trapped surfaces and past Q-screens. A Q-screen may be timelike, null, spacelike, or some combination thereof in different regions. Future and past Q-screens that also obey certain generic conditions analogous to those for holographic screens are the objects that are conjectured to obey a Generalized Second Law [44]:

Conjecture II.2 (Generalized Second Law) *Let $\mathcal{Q}$ be a regular future (resp. past) Q-screen. The generalized entropy of its leaves increases strictly monotonically under the past and outward (resp. future and inward) flow along $\mathcal{Q}$.*

Note that while the GSL remains unproven in general, it is known to hold in several examples, and it can in fact be shown to hold if one assumes the Quantum Focusing Conjecture [58].

So far we have not said much about the precise definition of generalized entropy, so let us discuss how it is defined in more careful terms. Our context here is quantum field theory in curved spacetime, rather than a full-blown theory of quantum gravity. Given some spacetime, suppose that it comes equipped with a foliation by Cauchy hypersurfaces, and suppose that the spacetime's matter content is described by a density matrix $\rho(\Sigma)$ on each Cauchy hypersurface Σ. Let σ be a compact connected surface that divides a Cauchy hypersurface Σ into two regions: the interior and exterior of σ. The generalized entropy computed with respect to σ and Σ is then the sum of the area of σ in Planck units and S_{out}, the von Neumann entropy of the reduced state of ρ restricted to the exterior of σ, cf. Eq. (3). The reduced state of ρ outside σ, which we denote ρ_{out}, is obtained by tracing out degrees of freedom on Σ in the interior of σ,

$$\rho_{\text{out}} \equiv \text{tr}_{\text{int}\,\sigma}[\rho(\Sigma)]\,, \tag{7}$$

and the Von Neumann entropy of ρ_{out} is

$$S_{\text{out}}[\sigma,\Sigma] = -\text{tr}\left[\rho_{\text{out}} \ln \rho_{\text{out}}\right]. \tag{8}$$

For a general field-theoretic state, the von Neumann entropy $S_{\text{out}}[\sigma,\Sigma]$ is a formally divergent quantity, so there is some subtlety surrounding how it is should be regulated, whether through an explicit ultraviolet cutoff or via subtracting a divergent vacuum contribution [44, 59]. Since we will exclusively be concerned with cosmology, we will work in a regime where the matter content of the spacetime has a conserved "thermodynamic," or coarse-grained entropy s per unit comoving volume. (Entropy per comoving volume is approximately conserved in cosmologies that do not have too much particle production [13, Ch 3.4].) The von Neumann entropy of a quantum mechanical system coincides with the thermodynamic Gibbs entropy in the classical limit where the state ρ_{out} has no coherence, i.e., is diagonal in the energy eigenbasis of Gibbs microstates. We then take the matter contribution to the generalized entropy to be a coarse-grained entropy $S_{\text{CG}}[\sigma,\Sigma]$ in the interior of σ. In other words, we will write the generalized entropy (3) as

$$\begin{aligned} S_{\text{gen}}[\sigma,\Sigma] &= \frac{A[\sigma]}{4G} + S_{\text{CG}}[\sigma,\Sigma] \\ &= \frac{A[\sigma]}{4G} + s \cdot \text{vol}_c[\sigma,\Sigma]\,, \end{aligned} \tag{9}$$

where $\text{vol}_c[\sigma,\Sigma]$ denotes the comoving (coordinate) volume of $\text{int}\,\sigma$ on Σ. (This approach is also taken in the examples of [44].) This expression is appropriate for cosmology, where observers find themselves on the inside of Q-screens and cosmological horizons when present, as opposed to observers who remain outside of a black hole and who are unable to access the interior of the black hole's horizon. Moreover, in the field-theoretic case where $\rho(\Sigma)$ is a pure state, then it follows that $S_{\text{in}} = S_{\text{out}}$, where S_{in} is the Von Neumann entropy of $\rho_{\text{in}} \equiv \text{tr}_{\text{ext}\,\sigma}[\rho(\Sigma)]$.

The fact that each leaf of a Q-screen extremizes the generalized entropy on an orthogonal lightsheet leads to a useful method for constructing Q-screens [45]. Given some spacetime with a foliation by Cauchy surfaces, suppose that one is also supplied with a foliation of the spacetime by null sheets with compact spatial cross-sections. Let each null sheet be labelled by a parameter r, and on each null sheet, let $\sigma(r)$ be the spatial section with extremal generalized entropy, when such exists. (Not every spacetime contains Q-screens, such as Minkowski space. But in Big Bang cosmologies, we expect both the area of, and entropy inside, a light cone to decrease in the very far past, so the generalized entropy will have an extremum somewhere.) It follows that each $\sigma(r)$ is a quantum marginal surface, and so if the quantum expansion has a definite sign in the other orthogonal null direction on each $\sigma(r)$, the union of these surfaces, $\mathcal{Q} = \bigcup_r \sigma(r)$, is by construction a Q-screen.

One way to generate a null foliation of a spacetime is to consider the past light cones of some timelike trajectory. Q-screens constructed from this type of foliation will be particularly useful for our purposes. This construction is illustrated through a worked example in Appendix A.

III. A COSMIC NO-HAIR THEOREM FOR RW SPACETIMES

We can used the notions reviewed above to show that spacetimes that expand and approach a constant maximum entropy along Q-screens will asymptote to de Sitter space. The basic idea of our proof is made clear by the simple example of a metric that is already homogeneous and isotropic, so that all we are showing is that the scale factor approaches e^{Ht} for some fixed constant H. The anisotropic case, considered in the next section, is considerably more complex, but the ideas are the same.

Let $\mathcal{M}$ be a Robertson-Walker (homogeneous and isotropic) spacetime with the line element

$$ds^2 = -dt^2 + a^2(t)\left(d\chi^2 + \chi^2 d\Omega_d^2\right), \tag{10}$$

where $t \in (t_i, \infty)$. Our aim is to show that if $\mathcal{M}$ admits a past Q-screen along which the generalized entropy monotonically increases up to a finite maximum value, then this alone, together with a handful of generic conditions on $\mathcal{M}$, implies that $\mathcal{M}$ is asymptotically de Sitter, or in other words, that

$$\lim_{t\to\infty} a(t) = e^{Ht} \tag{11}$$

for some constant H. In particular, we will neither make use of the Einstein field equations nor assume that there is a positive cosmological constant.

Begin by foliating $\mathcal{M}$ with past-directed light cones whose tips lie at the spatial origin $\chi = 0$, and suppose that $\mathcal{M}$ admits a past Q-screen, $\mathcal{Q}$, constructed with respect to this foliation. In other words, suppose that each light cone has a spatial slice with extremal generalized entropy so that $\mathcal{Q}$ is the union of all of these extremal slices. Past light cones will generically have a maximal entropy slice in cosmologies which, for example, begin with a big bang where $a(t_i) = 0$. An example is portrayed in Figure 2, which shows a holographic screen and a Q-screen in a cosmological spacetime with a past null singularity and a future de Sitter evolution; this example is explained in more detail in Appendix A. The intuition here is that while the past-directed null geodesics that make up a light cone may initially diverge, eventually they must meet again in the past when the scale factor vanishes and space becomes singular. Ultimately, however, we need only assume that the Q-screen exists, and we only remark on its possible origins for illustration.

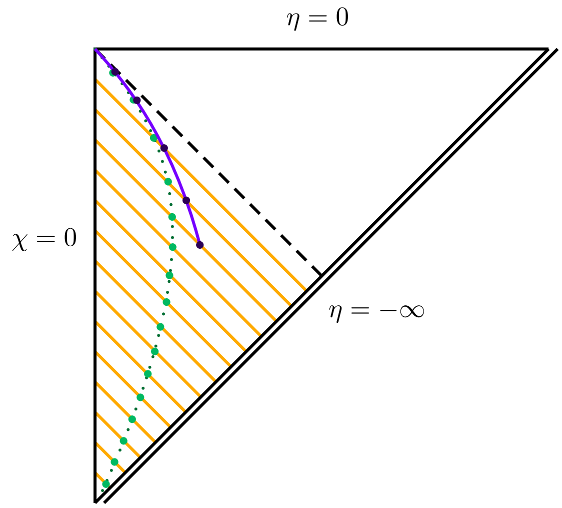

FIG. 2. Holographic screen and Q-screen illustrated on the Penrose diagram for a homogeneous and isotropic spacetime with a positive cosmological constant. Null sheets that make up the foliation by past-directed light cones are shown in yellow, and the cosmological horizon is the dotted black line. The dotted green line and large green dots are the holographic screen and its leaves respectively. The purple line and large purple dots are the Q-screen and its leaves.

Because RW spacetimes are spherically symmetric, the extremal-entropy light cone slices will be spheres, i.e., constant-t slices. If the quantum expansion vanishes in the lightlike direction along the light cone and is positive in the other lightlike direction at a single point on some test sphere, then it maintains these properties at every point on that sphere due to symmetry. This sphere is by construction a marginally quantum anti-trapped surface, or equivalently has extremal generalized entropy on the light cone. We therefore take the Cauchy surfaces Σ with respect to which generalized entropy is defined to be the constant-t surfaces in $\mathcal{M}$, since constant-t slices of light cones are spheres.

We will also make a handful of generic assumptions about $\mathcal{M}$ and $\mathcal{Q}$ without which a cosmic no-hair theorem is not guaranteed. Indeed, Wald's theorem does not hold in completely general cosmologies either; it requires that the spacetime is initially expanding and that it satisfies the strong and dominant energy conditions. Here, we will assume that space continues to expand for all cosmic time.[1] We want to avoid cosmologies that crunch or that otherwise clearly do not admit a no-hair theorem. We will also suppose that $\mathcal{Q}$ satisfies the generic conditions outlined in [44].

With these considerations in mind, the theorem that we wish to prove is the following:

Theorem III.1 *Let $\mathcal{M}$ be a RW spacetime with the line element* (10) *and whose matter content has constant thermodynamic entropy s per comoving volume. Suppose that $\mathcal{M}$ admits*

[1] In principle, the expansion need not be monotonic, but we will show that monotonicity is implied when $\mathcal{M}$ admits a Q-screen such as $\mathcal{Q}$.

a past Q-screen, $\mathcal{Q}$, constructed with respect to a foliation of $\mathcal{M}$ with past-directed light cones that are centered on the origin, $\chi = 0$, and suppose that the Generalized Second Law holds on $\mathcal{Q}$. Suppose that $\mathcal{M}$ and $\mathcal{Q}$ together satisfy the following assumptions:

(*a*) $a(t) \to \infty$ *as* $t \to \infty$,

(*b*) $S_{\text{gen}} \to S_{\text{max}} < \infty$ *along* $\mathcal{Q}$.

Then, $\mathcal{M}$ is asymptotically de Sitter and the scale factor $a(t)$ approaches e^{Ht}, where H is a constant.

Proof: For convenience we work in $d = 3$ spatial dimensions, but the generalization to arbitrary dimensions is straightforward. As discussed above, the leaves of $\mathcal{Q}$ are spheres. Letting the leaves be labelled by some parameter r, the generalized entropy is then given by

$$S_{\text{gen}}[\sigma(r), \Sigma(r)] \equiv S_{\text{gen}}(r) = \frac{\pi}{G}\chi^2(r)a^2(t(r)) + \tfrac{4}{3}\pi\chi^3(r)s\,. \tag{12}$$

The hypersurface $\Sigma(r)$ is the constant-$t(r)$ surface in which the leaf $\sigma(r)$ is embedded, and $\chi(r)$ denotes the radius of the leaf.

First, we need to establish that $\mathcal{Q}$ extends out to future timelike infinity. In principle, $\mathcal{Q}$ could become spacelike and consequently not extend beyond some time t (or in other words, $t(r)$ could have some finite maximum value), but it turns out that this does not happen.

Recall the property of Q-screens that generalized entropy is extremized on each leaf with respect to lightlike deformations. Here we may write

$$k^\mu \partial_\mu S_{\text{gen}} = 0\,, \tag{13}$$

where $k^\mu = (a(t), -1, 0, 0)$ is the lightlike vector that is tangent to the light cone and with respect to which S_{gen} is extremal. (Any point x^μ belongs to a unique sphere on a past-directed light cone and may therefore be associated to a particular value of S_{gen}. This lets us define the partial derivative in Eq. (13) above.) The deformation corresponds to dragging the leaf $\sigma(r)$ up and down the light cone, and by construction $S_{\text{gen}}(r)$ is extremal on the leaf $\sigma(r)$. Note that in principle we should consider deformations with respect to null *geodesics*, since the null generators of the light cone could have different normalizations at different points on $\sigma(r)$, or in other words, the geometry of the leaf $\sigma(r)$ could change as it is dragged by some fixed affine amount along the light cone. Here, however, the spherical symmetry of RW ensures that the null generators on $\sigma(r)$ all have the same normalization, so that k^μ as defined above is proportional to the null generators everywhere on $\sigma(r)$.

Writing out the partial derivatives, (13) becomes

$$\begin{aligned} 0 &= (a\,\partial_t - \partial_\chi)\left(\frac{\pi}{G}\chi^2 a^2 + \tfrac{4}{3}\pi\chi^3 s\right) \\ &= \frac{2\pi}{G}\chi^2 a^2 \dot{a} - \frac{2\pi}{G}\chi a^2 - 4\pi\chi^2 s. \end{aligned} \tag{14}$$

(One must be careful to distinguish between the coordinate t and the value $t(r)$ which labels leaves in the Q-screen.) If $\chi \neq 0$, then it follows that

$$\frac{1}{\chi} = \dot{a}(t) - \frac{2Gs}{a^2(t)}\,. \tag{15}$$

Eq. (15) lays out the criterion for when there is a leaf in a constant-t slice; when the right side is finite and positive, then there must be a leaf in that slice.

Observe that the right side of Eq. (15) does not diverge for any finite $t > t_i$ since $a(t)$ is defined for all $t \in [t_i, \infty)$ and only diverges in the infinite t limit by assumption. Furthermore, if the right side is nonzero for some time t (and consequently there is a leaf $\sigma(r)$ in the $t(r) = t$ slice), then the right side cannot approach zero, since this would cause the radius of subsequent leaves to grow infinitely large, which contradicts the assumption that S_{gen} remains finite. Therefore, if $\mathcal{Q}$ has a leaf at *some* time, then Eq. (15) shows that $\mathcal{Q}$ must have leaves in *all* future slices. $\mathcal{Q}$ is therefore timelike and extends out to future timelike infinity.[2] Furthermore, that the right side of Eq. (15) cannot vanish immediately implies that $\dot{a} > 2Gs/a^2 > 0$, so that the expansion must be monotonic.

Because $\mathcal{Q}$ is timelike, we can label each leaf by the constant-t_1 surface in which it lies, i.e., let the parameter r be a time t_1 (subscripted as such to distinguish it from the coordinate t). Referring to Eq. (12), since $a(t)$ grows without bound by assumption, it must be that $\chi(t_1)$ decreases at least as fast as $a^{-1}(t_1)$ in order for the area term in S_{gen} to remain finite (as it must, since by hypothesis $S_{\text{gen}} \leq S_{\text{max}}$). The matter entropy term therefore becomes irrelevant in the asymptotic future, and so that $S_{\text{gen}} \to S_{\text{max}}$, it must be that

$$\chi(t_1) \to \sqrt{\frac{GS_{\text{max}}}{\pi}}\frac{1}{a(t_1)} \tag{16}$$

as $t_1 \to \infty$.

Next, rearrange Eq. (15) to solve for $\dot{a}$. Using the asymptotic form for $\chi(t_1)$ in Eq. (16), to leading order in a we find that

$$\dot{a} \to \sqrt{\frac{\pi}{GS_{\text{max}}}}a \quad + \quad (\text{subleading})\,. \tag{17}$$

Therefore, it follows that $a(t) \to e^{Ht}$ as $t \to \infty$, where $H = (\pi/GS_{\text{max}})^{1/2}$, demonstrating that the metric approaches the de Sitter form, as desired. The entropy $S_{\text{max}} = \pi/GH^2$ coincides with the usual de Sitter horizon entropy. $\square$

[2] Alternatively, we could instead replace Assumption (*a*) with the assumption that $\mathcal{Q}$ is timelike and extending out to future timelike infinity and argue that $a \to \infty$. The arguments given here show that these two points are logically equivalent.

IV. A COSMIC NO-HAIR THEOREM FOR BIANCHI I SPACETIMES

In a RW spacetime, we demonstrated that the existence of a Q-screen along which entropy monotonically increases to a finite maximum implies that the scale factor tends to the de Sitter scale factor far in the future. Now we will go one step further and show that in the case where the cosmology is allowed to be anisotropic, virtually the same assumptions imply that any initial anisotropies decay at late times as well. Specifically, we will prove a cosmic no-hair theorem for Bianchi I spacetimes. The calculations in the proof for Bianchi I spacetimes are more involved than the RW case, so we will begin with a proof in 1+2 dimensions, where the anisotropy only has one functional degree of freedom. We will then generalize to 1+3 dimensions, which also makes apparent how to generalize to arbitrary dimensions.

A. 1+2 dimensions

Let $\mathcal{M}$ be a Bianchi I spacetime in $1+2$ dimensions with the line element

$$ds^2 = -dt^2 + a_1^2(t)\,dx^2 + a_2^2(t)\,dy^2 \tag{18}$$

where $t \in (t_i, \infty)$. Once again foliate $\mathcal{M}$ with past-directed light cones whose tips lie at $x = y = 0$, and suppose that $\mathcal{M}$ admits a past Q-screen $\mathcal{Q}$, constructed with respect to this foliation, together with an accompanying foliation by Cauchy hypersurfaces. Here we will assume that generalized entropy is locally maximal on each leaf (as opposed to being extremal). Our aim is to show that if generalized entropy tends to a finite maximum along $\mathcal{Q}$, then the GSL implies that $a_1(t), a_2(t) \to e^{Ht}$ as $t \to \infty$ for some constant H.

Here as well we will assume that space expands for all time, with $a_1(t), a_2(t) \to \infty$ as $t \to \infty$. We can similarly argue that $\mathcal{Q}$ must extend out to future timelike infinity provided that there is a leaf on each light cone so that the Q-screen does not simply terminate. A technical complication in the case of an anisotropic spacetime is that the leaves of $\mathcal{Q}$ in general are neither spheres (in proper or comoving coordinates) nor constant-t slices of light cones, and so it is difficult to argue as directly as in the RW case. But, because by definition any leaf $\sigma(r)$ locally maximizes the generalized entropy, any constant-t slice of the lightcone, label it $\varsigma(r)$, such that $\sigma(r) \cap \varsigma(r) \neq \emptyset$ can be used to place a lower bound on S_{gen}:

$$S_{\text{gen}}[\sigma(r), \Sigma(r)] \geq S_{\text{gen}}[\varsigma(r), \Sigma(r)] \tag{19}$$

Schematically, the type of Q-screen that we want to rule out is shown in Fig. 3 below. This is a Q-screen that gets "stuck" and does not progress past some finite time. One one hand, the screen cannot linger at a time greater than t_i, since the area of $\varsigma(r)$ would get larger and larger as its radial size increased throughout the foliation of $\mathcal{Q}$, causing $S_{\text{gen}}[\varsigma(r), \Sigma(r)]$, and hence also $S_{\text{gen}}[\sigma(r), \Sigma(r)]$ via the bound above, to diverge via the area entropy term. On the other hand, the screen cannot wrap back down to $t = t_i$ even if $a_1(t_i) = a_2(t_i) = 0$, since then S_{gen} diverges via the matter entropy term. Therefore, if $\mathcal{M}$ admits a Q-screen with a leaf on every light cone, then $\mathcal{Q}$ is necessary timelike past some t_{time} and extends out to future timelike infinity.

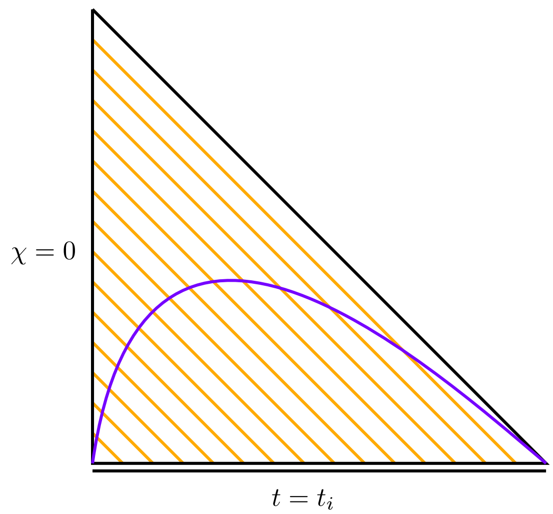

FIG. 3. Sketch of a hypothetical Q-screen (shown in purple) which does not extend past some finite maximum time in the Penrose diagram for a cosmological spacetime with a timelike initial singularity at $t = t_i$. The foliation of past-directed lightcones are the yellow diagonal lines. In such a case, the Q-screen does not reveal any information about the spacetime's asymptotic future.

Next, we introduce *conformal light cone coordinates* [60, 61], a more convenient coordinate system to work in when dealing with anisotropy. First, observe that we may rewrite the line element (18) as

$$ds^2 = -dt^2 + a^2(t)\left[e^{2b(t)}dx^2 + e^{-2b(t)}dy^2\right] \tag{20}$$

with $a_1(t) = a(t)\,e^{b(t)}$ and $a_2(t) = a(t)\,e^{-b(t)}$ [12]. In this parametrization, the "volumetric scale factor" $a(t)$ describes the overall expansion of space while $b(t)$ characterizes the anisotropy. Next, make the coordinate transformation to conformal time defined by $dt = \pm a(\eta)\; d\eta$ so that the line element (20) reads

$$ds^2 = a^2(\eta)\left[-d\eta^2 + e^{2b(\eta)}dx^2 + e^{-2b(\eta)}dy^2\right]. \tag{21}$$

Choose the sign of η so that $\eta(t)$ is a monotonically increasing function of t, and denote the limiting value of $\eta(t)$ as $t \to \infty$ by η_∞. Conformal light cone coordinates are then defined by the following coordinate transformation:

$$x(\eta, \eta_o, \theta) = \cos\theta \int_\eta^{\eta_o} \frac{e^{-2b(\zeta)}}{\sqrt{\cos^2\theta\; e^{-2b(\zeta)} + \sin^2\theta\; e^{2b(\zeta)}}}\; d\zeta \tag{22}$$

$$y(\eta, \eta_o, \theta) = \sin\theta \int_\eta^{\eta_o} \frac{e^{2b(\zeta)}}{\sqrt{\cos^2\theta\; e^{-2b(\zeta)} + \sin^2\theta\; e^{2b(\zeta)}}}\; d\zeta \tag{23}$$

The point with coordinates (η, η_o, θ) is reached by firing a past-directed null geodesic from the spatial origin $x = y = 0$ at an angle $\theta \in [0, 2\pi)$ counterclockwise relative to the x-axis at conformal time η_o and following the light ray in the past down to the conformal time η (Fig. 4). Note that while η is a timelike coordinate, η_o acts as a radial coordinate at each η.

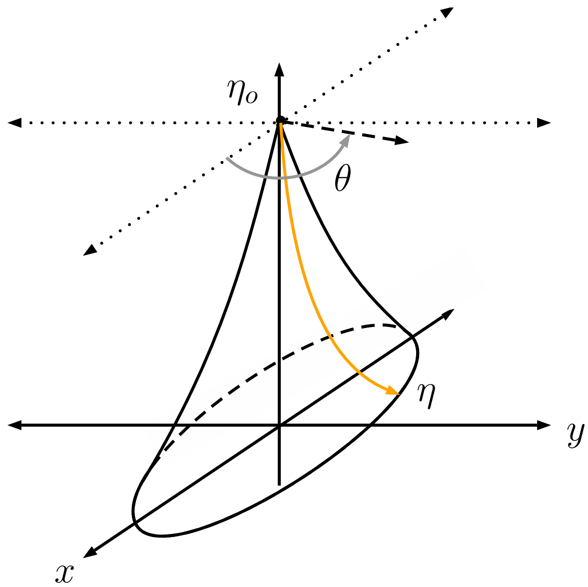

FIG. 4. Conformal light cone coordinates. At the conformal time η_o, fire a past-directed null geodesic (shown in yellow) from the origin at an initial angle θ relative to the positive x-axis and follow it until the conformal time η.

Note that the surfaces of constant η_o are precisely the past-directed light cones with respect to which $\mathcal{Q}$ is constructed. We can therefore label the leaves σ of $\mathcal{Q}$ by the values of η_o corresponding to the light cones on which they lie (Fig. 5):

$$Q = \bigcup_{\eta_o} \sigma(\eta_o) \tag{24}$$

Similarly, label the Cauchy hypersurfaces with respect to which each leaf is defined by $\Sigma(\eta_o)$. In various instances, it will be useful to use another coordinate,

$$\chi = \eta_o - \eta, \tag{25}$$

which may be thought of as a comoving radius in a sense that will be made precise later. We will also sometimes work with (η, χ, θ) and (χ, η_o, θ) in addition to the conformal light cone coordinates (η, η_o, θ).

The no-hair theorem that we will prove is as follows:

Theorem IV.1 *Let $\mathcal{M}$ be a Bianchi I spacetime with the line element (18) and whose matter content has constant thermodynamic entropy s per comoving volume. Suppose that $\mathcal{M}$ admits a past Q-screen, $\mathcal{Q}$, with maximal entropy leaves constructed with respect to a foliation of $\mathcal{M}$ with past-directed light cones that are centered on the origin, $x = y = 0$. Suppose that the Generalized Second Law holds on $\mathcal{Q}$ and that $\mathcal{M}$ and $\mathcal{Q}$ together satisfy the following assumptions:*

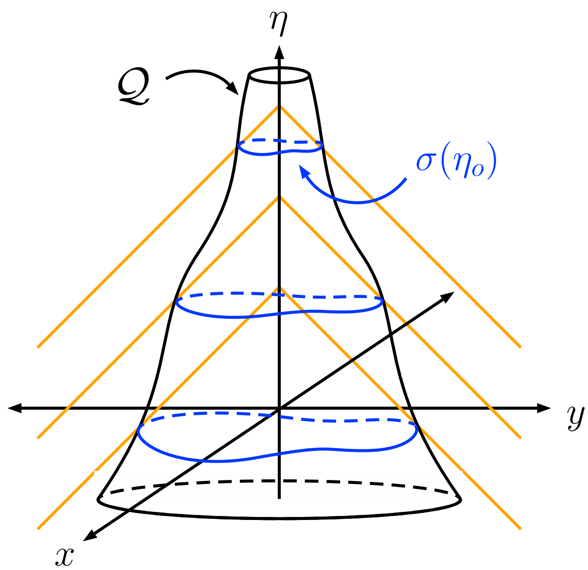

FIG. 5. A Q-screen $\mathcal{Q}$ (the solid black hypersurface) constructed with respected to a foliation by past-directed light cones (sketched in yellow). Each leaf $\sigma(\eta_o)$ (shown in blue) is labelled by the value of η_o where the tip of its parent light cone sits.

(*i*) $a_1(t), a_2(t) \to \infty$ *as* $t \to \infty$,

(*ii*) $\dot{a}_1(t), \dot{a}_2(t) > 0$ *after some* t_{mono},

(*iii*) $S_{\text{gen}} \to S_{\text{max}} < \infty$ *along* $\mathcal{Q}$.

Then, $\mathcal{M}$ *is asymptotically de Sitter and the scale factors* $a_1(t)$ *and* $a_2(t)$ *approach* $C_1 e^{Ht}$ *and* $C_2 e^{Ht}$*, respectively, where* H*,* C_1*, and* C_2 *are constants.*

Notes: To obtain a manifestly isotropic metric, rescale the coordinates x and y by C_1 and C_2, i.e., set $X = C_1 x$ and $Y = C_2 y$. Then, the line element (18) asymptotically reads $ds^2 \to -dt^2 + e^{2Ht}(dX^2 + dY^2)$. Also note that we have introduced an additional assumption compared to the RW case: Assumption (*ii*), that $a_1(t)$ and $a_2(t)$ grow monotonically past some time t_{mono}. Finally, also note that in terms of $a(\eta)$ and $b(\eta)$, Assumption (*i*) becomes:

(i') $a(\eta) \to \infty$ as $\eta \to \eta_\infty$ and $|b(\eta)| < \infty$.

The requirement that $b(\eta)$ remain bounded ensures that no dimension becomes singular. Likewise, the theorem is established by showing that $a(\eta) \to -1/H\eta$ and $b(\eta) \to B$ as $\eta \to 0^-$ (and also that $\eta_\infty = 0$) for some constant B.

Proof: The proof can be broken down into three parts. First, we show that the GSL and Assumption (*iii*) together imply that, asymptotically, $\mathcal{Q}$ squeezes into the comoving coordinate origin. Second, we use this asymptotic squeezing behaviour to demonstrate that the volumetric scale factor $a(\eta)$ tends to the de Sitter scale factor. Finally, we show that the asymptotic behaviour of $a(\eta)$ and Assumption (*ii*) together imply that anisotropy decays.

1. Showing that $\mathcal{Q}$ squeezes into the coordinate origin $\chi = 0$ as $\eta \to \eta_\infty$.

Consider the leaves of $\mathcal{Q}$ and work in $\tilde{x}^\mu = (\eta, \eta_o, \theta)$ coordinates. On the light cone whose tip is at η_o, each leaf $\sigma(\eta_o)$ is a closed path parameterized by:

$$\tilde{x}^\mu(u; \eta_o) = (\eta(u; \eta_o), \eta_o, u) \qquad u \in [0, 2\pi). \tag{26}$$

With u held fixed, $\eta(u; \eta_o)$ must be a monotonically-increasing function of η_o in the regime in which $\mathcal{Q}$ is timelike, i.e., for all conformal times past some $\eta_{\rm time}$. Furthermore, since $\mathcal{Q}$ extends out to timelike infinity, the function $\eta(u; \eta_o)$ with u held fixed takes every value of the conformal time between $\eta_{\rm time}$ and η_∞. In other words, as discussed above, the paths $\tilde{x}^\mu(u; \eta_o)$ never get‘stuck at some value of η. Therefore, in the timelike regime, we are free to relabel the leaves by some value of $\eta(u; \eta_o)$. A useful choice is to label the leaf $\sigma(\eta_o)$ by

$$\eta_{\rm min} = \min_u \left\{\eta(u; \eta_o)\right\}, \tag{27}$$

so that $\sigma(\eta_o) \equiv \sigma(\eta_{\rm min})$. Likewise, relabel the Cauchy hypersurfaces $\Sigma(\eta_o) \equiv \Sigma(\eta_{\rm min})$. Since $\eta_{\rm min}$ over all leaves takes all values between $\eta_{\rm time}$ and η_∞, it follows that $\eta_{\rm min}$ must tend to the same limiting value as η_o when the latter tends to its limiting value as $t \to \infty$. In other words, as $\eta_o \to \eta_\infty$, then $\eta_{\rm min} \to \eta_\infty$ as well.

Let $\tilde{\sigma}_{\eta_o}(\eta)$ denote the constant-η slice of the light cone whose tip is at η_o and let $\tilde{\Sigma}(\eta)$ denote the constant-η hypersurface (Fig. 6). (Recall that leaves σ of the Q-screen are not themselves generally embedded in constant-η surfaces.) Since by definition $\sigma(\eta_{\rm min})$ is a maximal entropy slice of a light cone, it follows that

$$S_{\rm gen}[\tilde{\sigma}_{\eta_o}(\eta_{\rm min}), \tilde{\Sigma}(\eta_{\rm min})] \leq S_{\rm gen}[\sigma(\eta_{\rm min}), \Sigma(\eta_{\rm min})] \leq S_{\rm max}, \tag{28}$$

where $S_{\rm gen}[\sigma(\eta_{\rm min}), \Sigma(\eta_{\rm min})] \to S_{\rm max}$ monotonically as $\eta_{\rm min} \to \eta_\infty$. We will use this bound on $S_{\rm gen}$ to demonstrate the desired squeezing of $\mathcal{Q}$. First we must derive the form of $S_{\rm gen}[\tilde{\sigma}_{\eta_o}(\eta_{\rm min}), \tilde{\Sigma}(\eta_{\rm min})]$.

Lemma IV.2 *The generalized entropy of a constant-η slice is given by*

$$S_{\rm gen}[\tilde{\sigma}_{\eta_o}(\eta), \tilde{\Sigma}(\eta)] = \frac{\tilde{A}(\eta, \chi)}{4G} + c_g(\eta, \chi)\chi^2 s, \tag{29}$$

where $\tilde{A}$ is given by

$$\tilde{A}(\eta, \chi) = a(\eta) \cdot \left[2\pi\chi + O(\chi^3)\right], \tag{30}$$

and $c_g(\eta, \chi)$ is some $O(1)$ geometric factor due to anisotropy that does not depend on $a(\eta)$.

Proof: First we justify the parametrization of the coarse-grained entropy $S_{\rm CG} = c_g(\eta, \chi)\chi^2 s$.

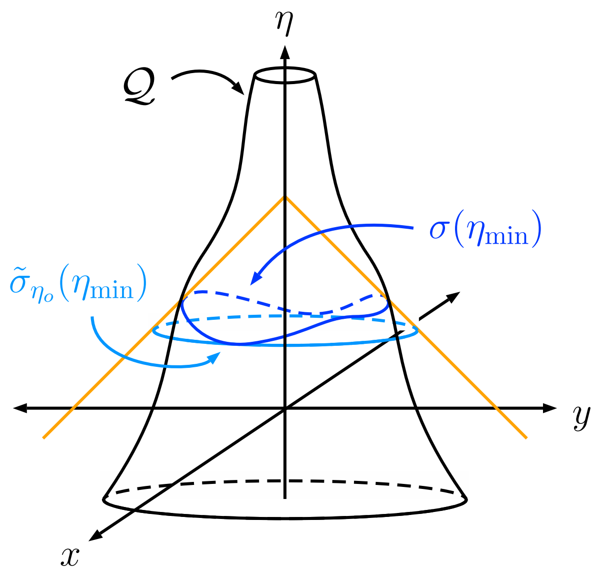

FIG. 6. Given a leaf $\sigma(\eta_{\min})$, the constant-$\eta = \eta_{\min}$ slice of its parent light cone is the surface $\tilde{\sigma}_{\eta_o}(\eta_{\min})$.

In the coordinates of the metric (18), S_{CG} is given by

$$S_{\mathrm{CG}}[\tilde{\sigma}_{\eta_o}(\eta), \tilde{\Sigma}(\eta)] = s \cdot \mathrm{vol}_c[\tilde{\sigma}_{\eta_o}(\eta), \tilde{\Sigma}(\eta)] = s \iint_{\mathrm{int}\,\tilde{\sigma}} dx\, dy\,, \tag{31}$$

where $\mathrm{int}\,\tilde{\sigma}_{\eta_o}(\eta)$ denotes the region on $\Sigma(\eta)$ inside $\tilde{\sigma}_{\eta_o}(\eta)$. In terms of the coordinates (η, χ, θ), S_{CG} is

$$S_{\mathrm{CG}}[\tilde{\sigma}_{\eta_o}(\eta), \tilde{\Sigma}(\eta)] \equiv S_{\mathrm{CG}}(\eta, \chi) = s \int_0^{\chi} \int_0^{2\pi} \left| \frac{\partial(x, y)}{\partial(\chi', \theta)} \right| \, d\theta\, d\chi'\,. \tag{32}$$

Formally, the Jacobian can be calculated from the coordinate transformation (22)-(23) above.[3] Expanding in powers of χ, one finds that

$$S_{\mathrm{CG}}(\eta, \chi) = s \cdot \left(\pi \chi^2 + \frac{\pi}{8} b'(\eta)^2 \chi^4 \right) + O(\chi^5)\,. \tag{33}$$

As such, we can simply define the function $c_g(\eta, \chi)$ to be the function

$$c_g(\eta, \chi) \equiv \frac{S_{\mathrm{CG}}(\eta, \chi)}{\chi^2 s} = \pi + \frac{\pi}{8} b'(\eta)^2 \chi^2 + O(\chi^3)\,. \tag{34}$$

The function $c_g(\eta, \chi)$ is therefore $O(\chi^0)$ by construction, and from the coordinate transformation (22)-(23), in which $a(\eta)$ never appears, we see that c_g cannot depend on $a(\eta)$, as

[3] A Maple worksheet which implements the calculations in this article is available through the online repository [62].

claimed.

Next we compute the proper area $\tilde{A}$. In three dimensions, the induced metric on a surface of constant η and χ has only a single component, given by

$$\gamma = \frac{\partial x^\mu}{\partial \theta}\frac{\partial x^\nu}{\partial \theta} g_{\mu\nu} = a^2(\eta)\left[e^{2b(\eta)}\left(\frac{\partial x}{\partial \theta}\right)^2 + e^{-2b(\eta)}\left(\frac{\partial y}{\partial \theta}\right)^2\right] \equiv a^2(\eta)\,\tilde{\gamma}\,, \tag{35}$$

where the coordinate partial derivatives read:

$$\frac{\partial x}{\partial \theta} = \sin\theta\left[\cos^2\theta \int_\eta^{\eta+\chi} \frac{(e^{-4b(s)}-1)\ ds}{(\cos^2\theta\ e^{-2b(s)} + \sin^2\theta\ e^{2b(s)})^{3/2}} - \int_\eta^{\eta+\chi} \frac{e^{-2b(s)}\ ds}{(\cos^2\theta\ e^{-2b(s)} + \sin^2\theta\ e^{2b(s)})^{1/2}}\right]$$
$$\frac{\partial y}{\partial \theta} = \cos\theta\left[\sin^2\theta \int_\eta^{\eta+\chi} \frac{(1-e^{4b(s)})\ ds}{(\cos^2\theta\ e^{-2b(s)} + \sin^2\theta\ e^{2b(s)})^{3/2}} + \int_\eta^{\eta+\chi} \frac{e^{2b(s)}\ ds}{(\cos^2\theta\ e^{-2b(s)} + \sin^2\theta\ e^{2b(s)})^{1/2}}\right].$$

It follows that

$$\tilde{A}(\eta,\chi) = \int_0^{2\pi} \sqrt{\gamma}\ d\theta = a(\eta)\int_0^{2\pi} \sqrt{\tilde{\gamma}}\ d\theta\,. \tag{36}$$

Expanding $\sqrt{\tilde{\gamma}}$ in powers of χ and integrating yields

$$\tilde{A}(\eta,\chi) = a(\eta)\cdot\left[2\pi\chi + O(\chi^3)\right]\,. \tag{37}$$

Note that Eq. (37) demonstrates the sense in which χ is a comoving radius (at least for small values of χ). ■

Using the result of Lemma IV.2, we can write the inequality (28) as

$$S_{\text{gen}}[\tilde{\sigma}_{\eta_o}(\eta_{\min}), \tilde{\Sigma}(\eta_{\min})] = \frac{\tilde{A}(\eta_{\min},\chi_{\max})}{4G} + c_g(\eta_{\min},\chi_{\max})\chi_{\max}^2 s \le S_{\max}, \tag{38}$$

where we have defined $\chi_{\max} \equiv \eta_o - \eta_{\min}$. At this point, we can re-examine the partial derivatives that enter the definition (35) of $\tilde{\gamma}$. Both $\partial x/\partial\theta$ and $\partial y/\partial\theta$ remain finite *and nonzero* for all η and for all nonzero χ, and correspondingly $\tilde{\gamma}^{1/2}$ is finite and nonzero for all η and θ if $\chi \neq 0$. (Recall that $|b(\eta)| < \infty$ by assumption.) In particular, this means that in order for $\tilde{A}$, and hence S_{gen} as well, to remain finite as $a \to \infty$, we must have that $\chi_{\max} \to 0$. As such, asymptotically, only the leading-order behaviour of $\tilde{A}$ is important, and the S_{CG} term becomes negligible. Therefore, asymptotically it must be true that

$$\chi_{\max} \le \frac{2GS_{\max}}{\pi}\frac{1}{a(\eta_o - \chi_{\max})}\,. \tag{39}$$

Then, as η_o takes its limiting value, $\chi_{\max}$ must tend to zero since it cannot become stuck at a finite value per the logic discussed above. In other words, in the system of coordinates (χ, η_o, θ), the Q-screen squeezes into the origin $\chi = 0$ as η_o takes its limiting value asymptotically in the future.

2. *Showing that $a(\eta)$ is asymptotically de Sitter.*

Now we turn our attention to calculating $S_{\text{gen}}[\sigma(\eta_o), \Sigma(\eta_o)]$ itself, and using its asymptotic properties as $\eta_o \to \eta_\infty$ to demonstrate that $a(\eta) \to -1/H\eta$ for a constant H with $\eta \in (-\infty, 0)$. For this part of the proof, we will work in the coordinates (χ, η_o, θ).

Since the radius of $\mathcal{Q}$ shrinks down to $\chi = 0$, here as well we can demonstrate that the area entropy approaches a constant while $S_{\text{CG}}[\sigma(\eta_o), \Sigma(\eta_o)]$ vanishes asymptotically in the future. The latter point follows by observing that the comoving volume of $\sigma(\eta_o)$ is contained within the comoving volume of $\tilde{\sigma}_{\eta_o}(\eta_{\min})$ (c.f. Fig. 6) which, in the last section, was shown to vanish as $\chi \to 0$.

Next we investigate the asymptotic behaviour of $A[\sigma(\eta_o)]$. Recall that the leaf $\sigma(\eta_o)$ is parametrized by some path $\tilde{x}^\mu(u) = (\chi(u; \eta_o), \eta_o, u)$ with η_o held constant and $0 \leq u < 2\pi$. In the future when S_{CG} becomes negligible, this path is the maximal area (also known as length in 1+2 dimensions) path on the light cone whose tip is at η_o, and so $A[\sigma(\eta_o)]$ satisfies

$$\frac{\delta A[\sigma(\eta_o)]}{\delta \chi(u; \eta_o)} = 0\,. \tag{40}$$

In principle, one can therefore solve the Euler-Lagrange problem above to obtain the path $\chi(u; \eta_o)$ and hence also the maximal area $A[\sigma(\eta_o)]$.

The tangent to the path is $t^\mu = d\tilde{x}^\mu/du = (\dot{\chi}(u; \eta_o), 0, 1)$ (where a dot denotes a derivative with respect to the parameter u). Therefore, the area of $\sigma(\eta_o)$ is given by

$$A[\sigma(\eta_o)] = \int_0^{2\pi} \sqrt{\tilde{g}_{\mu\nu} t^\mu t^\nu}\; du = \int_0^{2\pi} \sqrt{\tilde{g}_{00}\dot{\chi}^2 + 2\tilde{g}_{02}\dot{\chi} + \tilde{g}_{22}}\; du, \tag{41}$$

where $\tilde{g}_{\mu\nu}$ is the metric of Eq. (21) but rewritten in (χ, η_o, θ) coordinates. One finds that $\tilde{g}_{00} = 0$ exactly, but $\tilde{g}_{02}$ and $\tilde{g}_{22}$ do not admit any such simplifications. Because of this, solving the full Euler-Lagrange problem to actually obtain the path $\chi(u; \eta_o)$ is intractable in general.

Nevertheless, we can exploit the fact that $\mathcal{Q}$ squeezes into the coordinate origin and perform a small-χ expansion of $A[\sigma(\eta_o)]$. First, pull out a factor of $a(\eta_o - \chi)$ from the square root:

$$A[\sigma(\eta_o)] = \int_0^{2\pi} a(\eta_o - \chi)\sqrt{2f_{02}\dot{\chi} + f_{22}}\; du \tag{42}$$

In so doing we have defined $\tilde{g}_{\mu\nu} = [a(\eta_o - \chi)]^2 f_{\mu\nu}$. Then, expand the square root in χ. The result is

$$A[\sigma(\eta_o)] = \int_0^{2\pi} a(\eta_o - \chi) \left[\frac{\chi}{R(u;\eta_o)} + \frac{1}{2} b'(\eta_o) \frac{Q(u;\eta_o)}{R^2(u;\eta_o)} \chi^2 + O(\chi^3) \right] \, du, \tag{43}$$

where

$$\begin{aligned} R(u;\eta_o) &= e^{-2b(\eta_o)} \cos^2 u + e^{2b(\eta_o)} \sin^2 u \\ Q(u;\eta_o) &= e^{-2b(\eta_o)} \cos^2 u - e^{2b(\eta_o)} \sin^2 u. \end{aligned} \tag{44}$$

Pulling out the scale factor is necessary to avoid pathologies that arise because both χ and η_o become small in the same limit (see Appendix A for illustration).

Only keeping the first order term, the variation $\delta A/\delta \chi = 0$ gives

$$0 = -a'(\eta_o - \chi) \frac{\chi}{R(u;\eta_o)} + a(\eta_o - \chi) \frac{1}{R(u;\eta_o)} \,, \tag{45}$$

so asymptotically, the maximizing path $\chi(u;\eta_o) = \chi(\eta_o)$ is given implicitly by the solution of

$$\chi = \frac{a(\eta_o - \chi)}{a'(\eta_o - \chi)}. \tag{46}$$

To first order, $A[\sigma(\eta_o)]$ is given by

$$A[\sigma(\eta_o)] = 2\pi \frac{a^2(\eta_o - \chi)}{a'(\eta_o - \chi)}. \tag{47}$$

But the requirement that $S_{\text{gen}} \to S_{\max}$ means that $A[\sigma(\eta_o)]/4G$ must tend to the constant value $S_{\max}$, or in other words,

$$\lim_{\substack{\eta_o \to \eta_\infty \\ \chi \to 0}} \frac{a^2(\eta_o - \chi)}{a'(\eta_o - \chi)} = \frac{2GS_{\max}}{\pi} \equiv \frac{1}{H} \,. \tag{48}$$

Therefore, $a(\eta)$ asymptotically approaches de Sitter, $a(\eta) \to -1/H\eta$ as $\eta \to 0^-$, with $H = \pi/2GS_{\max}$.

Since $\chi(\eta_o)$ is a function of η_o, a technical detail to address is to check that the higher-order coefficients in the expansion (43), which themselves depend on η_o through $b(\eta_o)$ and its derivatives, do not cause the higher-order terms to be larger than the term that is first-order in χ. This we can achieve by bounding the remainder, $r_1(\chi;\eta_o) \equiv \sqrt{2f_{02}\dot{\chi} + f_{22}} - \chi/R$.

Let $F = \sqrt{2f_{02}\dot{\chi} + f_{22}}$. We may write its second derivative with respect to χ as

$$\frac{\partial^2 F}{\partial \chi^2} = b'(\eta_o - \chi) \frac{Q(u;\eta_o - \chi)}{R^2(u;\eta_o - \chi)} + \varepsilon(\chi;\eta_o), \tag{49}$$

where the term $\varepsilon(\chi;\eta_o) \to 0$ as $\chi \to 0$ for any η_o. As such, choose χ and η_o both small enough such that $|\varepsilon(\chi;\eta_o)| < |b'(\eta_o - \chi)|/R(u;\eta_o - \chi)$ for all u.[4] With this choice, and since $|Q/R| \leq 1$, we have that

$$\left|\frac{\partial^2 F}{\partial \chi^2}\right| < \frac{2|b'(\eta_o - \chi)|}{R(u;\eta_o - \chi)}\,. \tag{50}$$

Next we invoke the monotonicity Assumption (ii). Let $\eta_\star = \max\{\eta_{\rm mono}, \eta_{\rm time}\}$. In terms of $a(\eta)$ and $b(\eta)$, Assumption (ii) reads $(a(\eta)e^{+b(\eta)})' > 0$ and $(a(\eta)e^{-b(\eta)})' > 0$, or $|b'(\eta)| < a'(\eta)/a(\eta)$. Therefore, upon additionally requiring $0 > \eta_o - \chi > \eta_\star$ (i.e. possibly making χ and η_o smaller), it follows that

$$\left|\frac{\partial^2 F}{\partial \chi^2}\right| < \frac{2}{R(u;\eta_o - \chi)}\frac{a'(\eta_o - \chi)}{a(\eta_o - \chi)} \;\longrightarrow\; \frac{2}{R(u;\eta_o - \chi)\chi(\eta_o)}\,. \tag{51}$$

So, by Taylor's remainder theorem we have that $|r_1(\chi;\eta_o)| < R(u;\eta_o - \chi_1)^{-1}(\chi^2/\chi(\eta_o))$ on any interval $\chi \in [\chi(\eta_o), \chi_2]$, where $\chi_1 \in [\chi(\eta_o), \chi_2]$ minimizes $R(u;\eta_o - \chi)$. Or, at the edge of the interval,

$$|r_1(\chi(\eta_o);\eta_o)| < \frac{\chi(\eta_o)}{R(u;\eta_o - \chi_1)}\,. \tag{52}$$

Since $\int_0^{2\pi} R(u;\eta_o)^{-1}\; du = 2\pi$ irrespective of the value of η_o, it follows that remainder in the expansion is strictly smaller than the first-order term, so we were safe in restricting our attention to the first-order solution.

3. *Showing that the anisotropy decays.*

The decay of anisotropy is directly implied by Assumption (ii) once we have established that the volumetric scale factor $a(\eta)$ is asymptotically de Sitter. In the far future limit, Assumption (ii) recast as $(a(\eta)e^{\pm b(\eta)})' > 0$ gives

$$|b'(\eta)| < \frac{a'(\eta)}{a(\eta)} \;\xrightarrow{\eta\to 0^-}\; Ha(\eta) = \frac{1}{-\eta}\,. \tag{53}$$

Therefore, to capture the asymptotic scaling of $b'(\eta)$, we can write

$$b'(\eta) = \frac{f(\eta)}{(-\eta)^{1-p}}, \tag{54}$$

where $p > 0$ and where $|f(\eta)| \leq F$ for some bounded constant F when $\eta > \eta_\star$. In other words, $b'(\eta)$ cannot grow faster than $1/\eta$ as $\eta \to 0^-$, so that $(-\eta)^{1-p}b'(\eta)$ is some bounded

[4] The only instance in which this is not possible is if $|b'(\eta_o - \chi)|/R(u;\eta_o - \chi)$ vanishes faster than $|\varepsilon(\chi;\eta_o)|$. But, in this case, the remainder $|r_1(\chi;\eta_o)|$ can be bounded arbitrarily tightly, since $|\partial^2 F/\partial\chi^2|$ can be made arbitrarily small.

function. To establish that the anisotropy decays, and thus complete the proof of the theorem, we need only establish that $b(\eta)$ goes to a fixed limit at late times:

Lemma IV.3 *If $b'(\eta)$ satisfies Eq. (54) on $(\eta_\star, 0)$, then $\lim_{\eta\to 0^-} b(\eta)$ exists.*

Proof: We show that the limit exists by showing that $b(\eta)$ is a Cauchy function. Namely, let $\epsilon > 0$. We must find $\delta > 0$ such that $0 < -\eta_1 < \delta$ and $0 < -\eta_2 < \delta$ implies that $|b(\eta_2) - b(\eta_1)| < \epsilon$. Without loss of generality, suppose that $\eta_\star < \eta_1 < \eta_2$. Then:

$$\begin{aligned} |b(\eta_2) - b(\eta_1)| &= \left| \int_{\eta_1}^{\eta_2} \frac{f(u)}{(-u)^{1-p}}\, du \right| \\ &\leq \int_{\eta_1}^{\eta_2} \frac{|f(u)|}{(-u)^{1-p}}\, du \\ &\leq F \int_{\eta_1}^{\eta_2} \frac{1}{(-u)^{1-p}}\, du \\ &\leq \frac{F}{p}(-\eta_1)^p. \end{aligned} \tag{55}$$

Therefore, let $\delta = (\epsilon F/p)^{1/p}$. Then $|b(\eta_2) - b(\eta_1)| < \epsilon$ as required. ■

□

B. 1+3 dimensions

Now suppose that $\mathcal{M}$ is a Bianchi spacetime in $1+3$ dimensions with the line element

$$ds^2 = -dt^2 + a_1^2(t)\, dx^2 + a_2^2(t)\, dy^2 + a_3^2(t)\, dz^2\,. \tag{56}$$

The case where $\mathcal{M}$ has 3 dimensions of space parallels the $1+2$ dimensional case with only a handful of technical complications. The main difference is that now the anisotropy has two functional degrees of freedom:

$$ds^2 = -dt^2 + a^2(t)\left[e^{2b_1(t)}dx^2 + e^{2b_2(t)}dy^2 + e^{2b_3(t)}dz^2\right] \tag{57}$$

One arrives at the equation above by setting $a_i(t) = a(t)e^{b_i(t)}$ for $i = 1, 2, 3$, where the $b_i(t)$ are subject to the constraint $\sum_{i=1}^3 b_i(t) = 0$. The definition of conformal light cone coordinates $(\eta, \eta_o, \theta, \phi)$ is correspondingly modified:

$$x^j(\eta, \eta_o, \theta, \phi) = D^j(\theta, \phi) \int_\eta^{\eta_o} \frac{e^{-2b_j(\zeta)}}{\sqrt{\sum_{i=1}^3 D^i(\theta,\phi)^2\, e^{-2b_i(\zeta)}}}\, d\zeta, \tag{58}$$

where

$$D^j(\theta, \phi) = (\sin\theta\, \cos\phi, \sin\theta\, \sin\phi, \cos\theta).$$

Nevertheless, the essential construction remains unchanged. We still consider a past Q-screen, $\mathcal{Q}$, constructed with respect to a foliation of $\mathcal{M}$ by past-directed light cones, and the leaves of $\mathcal{Q}$ are still labelled by the conformal time η_o where the tip of their corresponding light cone is located. The no-hair theorem also generalizes in a straightforward way:

Theorem IV.4 *Let $\mathcal{M}$ be a Bianchi I spacetime with the line element (56) and whose matter content has constant thermodynamic entropy s per comoving volume. Suppose that $\mathcal{M}$ admits a past Q-screen, $\mathcal{Q}$, with maximal entropy leaves constructed with respect to a foliation of $\mathcal{M}$ with past-directed light cones that are centered on the origin, $x = y = z = 0$. Suppose that the Generalized Second Law holds on $\mathcal{Q}$ and that $\mathcal{M}$ and $\mathcal{Q}$ together satisfy the following assumptions for $i \in \{1, 2, 3\}$:*

(i) $a_i(t) \to \infty$ *as* $t \to \infty$,

(ii) $\dot{a}_i(t) > 0$ *past some* t_{mono},

(iii) $S_{\text{gen}} \to S_{\max} < \infty$ *along* $\mathcal{Q}$.

Then, $\mathcal{M}$ is asymptotically de Sitter and the axial scale factors $a_i(t)$ approach $C_i e^{Ht}$, where H and C_i are constants.

Note: In terms of $a(\eta)$ and the $b_i(\eta)$, Assumption (i) becomes:

(i') $a(\eta) \to \infty$ as $\eta \to \eta_\infty$ and $|b_i(\eta)| < \infty$.

Proof: The proof for 1+3 dimensions exactly parallels the proof of Theorem IV.1, so we only note the most important modifications. Beginning with Part 1, in $(\eta, \eta_o, \theta, \phi)$ coordinates, the leaves $\sigma(\eta_o)$ are now parametrized surfaces,

$$\tilde{x}^\mu(u, v; \eta_o) = (\eta(u, v; \eta_o), \eta_o, u, v) \qquad u \in [0, \pi] \qquad v \in [0, 2\pi)\,, \tag{59}$$

and so $\eta_{\min}$ is the minimum value of $\eta(u, v; \eta_o)$ over the two parameters u and v:

$$\eta_{\min} = \min_{u,v} \{\eta(u, v; \eta_o)\} \tag{60}$$

The expression for the generalized entropy of a constant-η slice of a light cone is also modified:

Lemma IV.5 *The generalized entropy of a constant-η slice is given by*

$$S_{\text{gen}}[\tilde{\sigma}_{\eta_o}(\eta), \tilde{\Sigma}(\eta)] = \frac{\tilde{A}(\eta, \chi)}{4G} + c_g(\eta, \chi)\chi^3 s. \tag{61}$$

where

$$\tilde{A}(\eta, \chi) = a^2(\eta) \cdot \left[4\pi\chi^2 + O(\chi^4)\right]. \tag{62}$$

and $c_g(\eta, \chi)$ is an $O(1)$ geometric factor due to anisotropy that does not depend on $a(\eta)$.

Proof: Repeating the steps described in Lemma IV.2, one finds that

$$c_g(\eta,\chi) \equiv \frac{S_{\rm CG}(\eta,\chi)}{\chi^3 s} = \frac{4\pi}{3} + \frac{8\pi}{45}\left(b_1'(\eta)^2 + b_1'(\eta)b_2'(\eta) + b_2'(\eta)^2\right)\chi^2 + O(\chi^3). \tag{63}$$

The area of $\tilde{\sigma}_{\eta_o}(\eta)$ is now a surface integral,

$$\tilde{A}(\eta,\chi) = \int_0^{\pi}\int_0^{2\pi}\sqrt{\gamma}\; d\phi\, d\theta\,, \tag{64}$$

where the induced metric on $\tilde{\sigma}_{\eta_o}(\eta)$ is

$$\gamma_{ab} = \frac{\partial x^\mu}{\partial\theta^a}\frac{\partial x^\nu}{\partial\theta^b}g_{\mu\nu} = a^2(\eta)\sum_{j=1}^{3} e^{2b_j(\eta)}\frac{\partial x^j}{\partial\theta^a}\frac{\partial x^j}{\partial\theta^b}\,, \tag{65}$$

and where x^j and $g_{\mu\nu}$ refer to Eq. (57) with $\theta^a \equiv (\theta,\phi)$. Expanding $\sqrt{\gamma}$ in powers of χ and integrating yields Eq. (62) as required. ■

From this point onward, the rest of Part I is the same. The inequality $S_{\rm gen}[\tilde{\sigma}_{\eta_o}(\eta_{\min}),\tilde{\Sigma}(\eta_{\min})] \leq S_{\rm gen}[\sigma(\eta_{\min}),\Sigma(\eta_{\min})] \leq S_{\max}$ still applies, leading to the conclusion that

$$\chi_{\max} \leq \sqrt{\frac{GS_{\max}}{\pi}}\frac{1}{a(\eta_o-\chi_{\max})}\,. \tag{66}$$

In the system of coordinates $(\chi,\eta_o,\theta,\phi)$, the Q-screen squeezes into the origin $\chi=0$ as η_o takes its limiting value asymptotically in the future.

Next we turn to showing that scale factor $a(\eta)$ is asymptotically de Sitter (Part 2). Consider the generalized entropy $S_{\rm gen}[\sigma(\eta_o),\Sigma(\eta_o)]$ once more. First, since the values taken by χ on any leaf tends to zero, we again conclude that $S_{\rm CG}[\sigma(\eta_o),\Sigma(\eta_o)]$ vanishes asymptotically, and so we focus on the area term $A[\sigma(\eta_o)]$.

The leaf $\sigma(\eta_o)$ is parametrized by some surface $\tilde{x}^\mu(u,v) = (\chi(u,v;\eta_o),\eta_o,u,v)$ with η_o held constant and $0\leq u\leq\pi$, $0\leq v<2\pi$. By definition, this surface is the surface on the light cone with tip at η_o with maximal area, and so it is the solution of

$$\frac{\delta A[\sigma(\eta_o)]}{\delta\chi(u,v;\eta_o)} = 0\,. \tag{67}$$

The induced metric on this surface is, as usual, given by

$$h_{ab} = \frac{\partial\tilde{x}^\mu}{\partial u^a}\frac{\partial\tilde{x}^\nu}{\partial u^b}\tilde{g}_{\mu\nu} \tag{68}$$

where $\tilde{g}_{\mu\nu}$ is the metric of Eq. (57) but rewritten in $(\chi,\eta_o,\theta,\phi)$ coordinates. The area of

$\sigma(\eta_o)$ is given by

$$A[\sigma(\eta_o)] = \int_0^{\pi} \int_0^{2\pi} \sqrt{\det h}\; dv\, du\,, \tag{69}$$

and the components of h_{ab} are as follows:

$$\begin{aligned} h_{uu} &= (\partial_u \chi)^2 \tilde{g}_{00} + 2(\partial_u \chi)\tilde{g}_{02} + \tilde{g}_{22} \\ h_{uv} &= (\partial_u \chi)(\partial_v \chi)\tilde{g}_{00} + (\partial_u \chi)\tilde{g}_{03} + (\partial_v \chi)\tilde{g}_{02} + \tilde{g}_{23} \\ h_{vv} &= (\partial_v \chi)^2 \tilde{g}_{00} + 2(\partial_v \chi)\tilde{g}_{03} + \tilde{g}_{33}. \end{aligned} \tag{70}$$

Once more, solving the full Euler-Lagrange problem for $\chi(u,v;\eta_o)$ to obtain the maximal area A is intractable, so we use the same trick where we extract an overall factor of $a^4(\eta_o - \chi)$ from $\det h$ and then expand the square root of the quotient in powers of χ. The result is

$$A[\sigma(\eta_o)] = \int_0^{\pi} \int_0^{2\pi} a^2(\eta_o - \chi) \left[\frac{\sin\theta}{R(u,v;\eta_o)^{3/2}} \chi^2 + \frac{Q(u,v;\eta_o)\sin\theta}{R(u,v;\eta_o)^{5/2}} \chi^3 + O(\chi^4) \right] dv\, du, \tag{71}$$

where

$$\begin{aligned} R(u,v;\eta_o) &= \sum_{i=1}^{3} e^{-2b_i(\eta_o)} D^i(u,v)^2 \\ Q(u,v;\eta_o) &= \sum_{i=1}^{3} b_i'(\eta_o) e^{-2b_i(\eta_o)} D^i(u,v)^2. \end{aligned} \tag{72}$$

Only keeping the lowest order term, the variation $\delta A/\delta\chi = 0$ gives the maximal path $\chi(u,v;\eta_o) = \chi(\eta_o)$ as the solution of

$$\chi = \frac{a(\eta_o - \chi)}{a'(\eta_o - \chi)}. \tag{73}$$

So, to lowest order, $A[\sigma(\eta_o)]$ is given by

$$A[\sigma(\eta_o)] = \chi^2 a^2(\eta_o - \chi) \int_0^{\pi} \int_0^{2\pi} \frac{\sin\theta}{R(u,v;\eta_o)^{3/2}}\, dv\, du = 4\pi \left(\frac{a^2(\eta_o - \chi)}{a'(\eta_o - \chi)} \right)^2 . \tag{74}$$

But the requirement that $S_{\text{gen}} \to S_{\max}$ means that $A[\sigma(\eta_o)]/4G$ must tend to the constant value $S_{\max}$, or in other words,

$$\lim_{\substack{\eta_o \to \eta_\infty \\ \chi \to 0}} \frac{a^2(\eta_o - \chi)}{a'(\eta_o - \chi)} = \sqrt{\frac{G S_{\max}}{\pi}} \equiv \frac{1}{H} \tag{75}$$

Therefore, $a(\eta)$ asymptotically approaches de Sitter, $a(\eta) \to -1/H\eta$ as $\eta \to 0^-$, with $H^2 = \pi/GS_{\max}$. Note that we recover the same Hubble constant as in the Cosmic No-Hair Theorem III.1 for RW spacetimes in $1+3$ dimensions.

Finally, as in the case of $1+2$ dimensions, the condition $(a(\eta)e^{b_i(\eta)})' > 0$ is enough to show that $\lim_{\eta\to 0^-} b_i'(\eta)$ exists for each i. $\square$

V. DISCUSSION

Assuming the Generalized Second Law, we have shown that if a Bianchi I spacetime admits a past Q-screen along which generalized entropy increases up to a finite maximum, then this implies that the spacetime is asymptotically de Sitter. We recover a version of Wald's cosmic no-hair theorem by making thermodynamic arguments about spacetime, without appealing to Einstein's equations.

While the proof of these cosmic no-hair theorems is most tractable (and certainly easiest to visualize) in $1+2$ dimensions, the generalization to $1+3$ dimensions was fairly immediate. In principle, the proof strategy for arbitrary dimensions is the same, albeit more difficult from the perspective of calculation. This is chiefly because calculating area elements of codimension-2 surfaces in arbitrary dimensions is cumbersome. Nevertheless, it is natural to expect that analogous cosmic no-hair theorems hold for Bianchi I spacetimes of arbitrary dimensions.

Within the proof itself, it would be interesting to see if the monotonicity assumptions, $a_i'(\eta) > 0$, could be eliminated. The fact that the Generalized Second Law asserts that S_{gen} increases monotonically along a Q-screen does offer some leverage. In particular, asymptotically this implies that the average scale factor $a(\eta) = (\prod_{i=1}^{d} a_i(\eta))^{1/d}$ increases monotonically; however, we learn nothing about the anisotropies $b_i(\eta)$, since the leading order behaviour of S_{gen} does not depend on the $b_i(\eta)$ in the asymptotic future regime. We also note that the monotonicity assumptions do not trivialize the cosmic no-hair theorems demonstrated in Sec. IV. For example, assuming monotonicity does not rule out exponential expansion with different rates in different directions, nor asymptotically power-law scale factors, nor does it even imply accelerated expansion at all.

An interesting extension would be to try to prove a no-hair theorem for classical cosmological perturbations [63], or for quantum fields in curved spacetime. Given a scalar field on a curved spacetime background, the task is to show that the combined metric and scalar field perturbations settle into the Bunch-Davies state [64] at late times. In principle it would suffice to show that the background spacetime still tends to de Sitter in the future in this case, as one could then simply invoke known no-hair results about scalar fields in curved backgrounds [16–18]. Conceptually, such a calculation would be interesting because one can explicitly write down the quantum state of cosmological perturbations, and so a full treatment of the matter entropy as Von Neumann entropy (modulo ultraviolet divergences) is possible in this case.

To prove our theorem, it wasn't strictly necessary to assume that the gravitational contribution to the entropy was precisely proportional to the surface area. We could imagine choosing some other function of the area, such that

$$S_{\mathrm{gen}}[\sigma,\Sigma] = f(A[\sigma]/G) + S_{\mathrm{CG}}[\sigma,\Sigma]. \tag{76}$$

For example, returning to the RW case, if one sets $f(A/G) = C(A/G)^p$ for some constants C and p, exactly the same analysis as in the proof of Theorem III.1 leads to the conclusion that (cf. Eq. (17))

$$\dot{a}(t) \to \sqrt{\frac{4\pi}{G}\left(\frac{C}{S_{\max}}\right)^{1/p}}\, a(t) \tag{77}$$

in the limit as $t \to \infty$. In other words, one still concludes that the scale factor is asymptotically de Sitter, albeit with a Hubble constant that differs from the usual case of $f(A/G) = A/4G$.

This work can be thought of as part of the more general program of connecting gravitation to entropy, thermodynamics, and entanglement [30–42]. As in attempts to derive Einstein gravity from entropic considerations, we deduce the behavior of the geometry of spacetime from thermodynamics, without explicit field equations. Our result is less general, as we only obtain the asymptotic behavior of the universe, but is perhaps also more robust, as our assumptions are correspondingly minimal. Thinking of spacetime as emerging thermodynamically from a set of underlying degrees of freedom can change our perspective on the knotty problems of quantum gravity; for example, as emphasized by Banks [21], the cosmological constant problem becomes the question of "Why does Hilbert space have a certain number of dimensions?" rather than "Why is this parameter in the low-energy effective Lagrangian so small?" Problems certainly remain (including why the entropy was so low near the Big Bang), but this alternative way of thinking about gravitation may prove useful going forward.

ACKNOWLEDGEMENTS

We would like to thank Cliff Cheung, John Preskill, and Alan Weinstein for helpful discussions. This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of High Energy Physics, under Award Number DE-SC0011632, as well as by the Walter Burke Institute for Theoretical Physics at Caltech and the Gordon and Betty Moore Foundation through Grant 776 to the Caltech Moore Center for Theoretical Cosmology and Physics.

Appendix A: Q-screens – a worked example

In this appendix, we illustrate Q-screens by explicitly constructing one in a RW spacetime that is asymptotically de Sitter. Consider a RW spacetime in $1+3$ dimensions with the line element $ds^2 = -dt^2 + a^2(t)(d\chi^2 + \chi^2 d\Omega_2^2)$ and where the scale factor is $a(t) = \sinh t$, $t \in (0,\infty)$. Conformal time is given by $\eta(t) = -2\,\mathrm{arccoth}(e^t)$, $\eta \in (-\infty, 0)$, and the scale factor in conformal time is

$$a(\eta) = \frac{1}{\sinh(-\eta)}\,. \tag{A1}$$

Foliate the spacetime with past-directed light cones centered at the coordinate origin $\chi = 0$, and let the Cauchy hypersurfaces of the spacetime be the constant-η hypersurfaces. Let us now construct a Q-screen by extremizing the generalized entropy on each light cone.

Consider a past-directed light cone whose tip is at the conformal time η_o. A constant-$\eta < \eta_o$ slice of this light cone is a 2-sphere of coordinate radius $\eta_o - \eta$, and so the generalized entropy computed with respect to this slice is

$$S_{\text{gen}}(\eta;\eta_o) = \frac{\pi}{G}\left(\frac{\eta_o - \eta}{\sinh(-\eta)}\right)^2 + \frac{4}{3}\pi(\eta_o - \eta)^3 s\,. \tag{A2}$$

A plot of $S_{\text{gen}}(\eta;\eta_o)$ as a function of η for several values of η_o is shown in Fig. 7. The area term $A(\eta;\eta_o)/4G$ alone is also overlaid on the plot, which illustrates that it is the dominant contribution to the generalized entropy at late times. Notice that in addition to having a local maximum, $S_{\text{gen}}(\eta;\eta_o)$ also has a local minimum, and beyond a certain critical value η_o^{crit} there is in fact no nonzero value of η which locally extremizes $S_{\text{gen}}(\eta;\eta_o)$. As such, the Q-screen, which is defined as the union of the slices with maximal generalized entropy, is only defined for $\eta_o \geq \eta_o^{\text{crit}}$. This is in contrast to the area $A(\eta;\eta_o)$, which has a locally-maximizing value of η for all η_o. The holographic screen, which is made up of extremal *area* slices, is therefore defined for all times. Both the Q-screen and the holographic screen were schematically illustrated previously in Fig. 2.

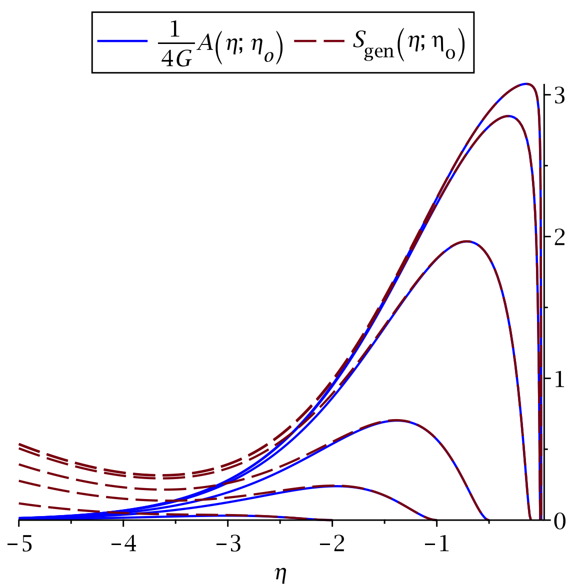

FIG. 7. Plots of area (solid) and generalized entropy (dashed) along light cones. From the lowest peak to the highest peak, the values of η_o are -2, -1, -0.5, -0.1, -0.01, and -0.001. Here we have taken $G = 1$ and we have picked $s = 0.001$.

Generalized entropy is extremal when $\partial S_{\text{gen}}/\partial\eta = 0$. Excluding $\eta = 0$ and $\eta \to -\infty$, the extremizing values of η are the real-valued solutions of

$$\eta_o - \eta = \frac{\sinh(-\eta)}{\cosh(-\eta) - 2Gs\sinh(-\eta)^3} \tag{A3}$$

when they exist. Let $\eta^Q(\eta_o)$ denote the maximizing value, and hence also define the Q-screen leaf radius $\chi^Q(\eta_o) \equiv \eta_o - \eta^Q(\eta_o)$. A plot of $\chi^Q(\eta_o)$ is shown in Fig. 8. As expected, $\chi^Q(\eta_o)$ vanishes as $\eta_o \to 0^-$. For comparison, we also plot the holographic screen radius $\chi^H(\eta_o) \equiv \eta_o - \eta^H(\eta_o)$, where $\eta^H(\eta_o)$ maximizes the area of the light cone slice, i.e., it is the solution of

$$\eta_o - \eta = \tanh(-\eta)\,. \tag{A4}$$

In particular note that $\chi^Q(\eta_o)$ is always slightly larger than $\chi^H(\eta_o)$, but they ultimately coincide in the limit $\eta \to 0^-$ (cf. Fig. 2).

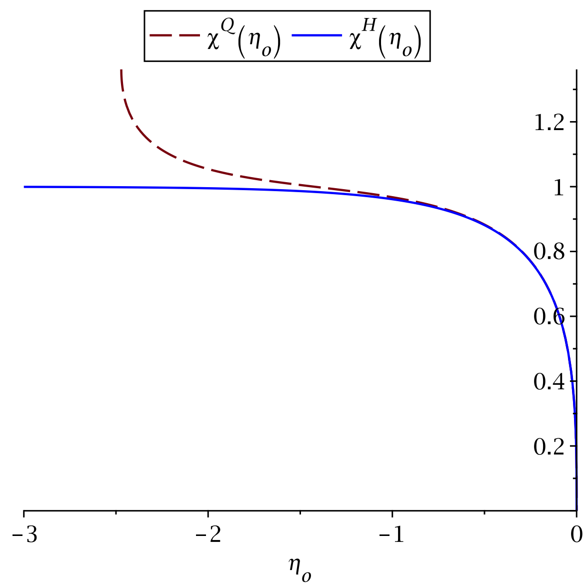

FIG. 8. Asymptotic behaviour of the radius of the Q-screen leaves ($\chi^Q(\eta_o)$, dashed) and holographic screen leaves ($\chi^H(\eta_o)$, solid).

As a final exercise, let us investigate the asymptotic dependence of $\chi^H(\eta_o)$ on η_o (which is also the asymptotic dependence of $\chi^Q(\eta_o)$, since the two coincide as $\eta_o \to 0$) to illustrate some of the subtleties involved in performing asymptotic expansions. Consider Eq. (A4) and let $\eta = \eta_o - \chi$ so that we have $\chi = \tanh(\chi - \eta_o)$. Since, asymptotically, $\chi \to 0$, one may be tempted to expand this last equation for small values of χ:

$$\chi = \tanh(-\eta_o) + (1 - \tanh^2(-\eta_o))\chi + O(\chi^2) \quad \Rightarrow \quad \chi^H(\eta_o) \stackrel{?}{=} \frac{1}{\tanh(-\eta_o)} \tag{A5}$$

Notice, however, that since $0 < \tanh(-\eta_o) < 1$, this expression for $\chi^H(\eta_o)$ cannot be infinitesimally small—the expansion is inconsistent! Rather, χ and η_o are simultaneously infinitesimal. Consider instead the double Taylor series in χ and η_o:

$$\chi = \chi - \eta_o - \tfrac{1}{3}\chi^3 + \eta_o\chi^2 - \eta_o^2\chi + \tfrac{1}{3}\eta_o^3 \quad \Rightarrow \quad \chi^H(\eta_o) = \eta_o + (-3\eta_o)^{1/3} \tag{A6}$$

This last result is the correct asymptotic behaviour of $\chi(\eta_o)$.

Similarly, writing $A = 4\pi\chi^2 a^2(\eta_o - \chi)$, one arrives at the wrong expressions for extremal values if one tries to expand A in small values of χ, η_o, or even both at the same time. The key is to keep $a(\eta_o - \chi)$ intact so that one arrives as Eq. (A4). Doing so leaves just enough nonlinearity to be able to restore the correct asymptotic behaviour of $\chi^H(\eta_o)$. This technique is exploited in Sec. IV A 2.

---

[1] R. M. Wald, "Asymptotic behavior of homogeneous cosmological models in the presence of a positive cosmological constant," *Phys. Rev. D* **28** no. 8, (1983) 2118–2120.

[2] A. A. Starobinskii, "Isotropization of arbitrary cosmological expansion given an effective cosmological constant," *JETP Lett.* **37** no. 1, (1983) 66.

[3] J. D. Barrow, "Cosmic no-hair theorems and inflation," *Phys. Lett. B* **187** no. 1, (1987) .

[4] J. D. Barrow and G. Götz, "The asymptotic approach to de Sitter space-time," *Phys. Lett. B* **231** no. 3, (1989) 228–230.

[5] Y. Kitada and K. Maeda, "Cosmic no-hair theorem in power-law inflation," *Phys. Rev. D* **45** no. 4, (1992) 4–7.

[6] Y. Kitada and K. Maeda, "Cosmic no-hair theorem in homogeneous spacetimes. I. Bianchi models," *Class. Quant. Grav.* **10** (1993) 703.

[7] M. Bruni, S. Matarrese, and O. Pantano, "A local view of the observable universe," *Phys. Rev. Lett.* **74** (1995) 1916–1919, `arXiv:astro-ph/9407054`.

[8] M. Bruni, F. C. Mena, and R. Tavakol, "Cosmic no-hair: non-linear asymptotic stability of de Sitter universe," *Class. Quant. Grav.* **19** no. L23, (2002) 1–8, `arXiv:gr-qc/0107069`.

[9] W. Boucher and G. W. Gibbons, "Cosmic baldness," `arXiv:1109.3535`.

[10] A. Maleknejad and M. M. Sheikh-Jabbari, "Revisiting cosmic no-hair theorem for inflationary settings," *Phys. Rev. D* **85** (2012) 123508, `arXiv:1203.0219`.

[11] R. M. Wald, *General Relativity*. University of Chicago Press, Chicago, IL, 1984.

[12] B. K. Berger, *Singularities in Cosmological Spacetimes*, pp. 437–460. Springer, Berlin, Heidelberg, 2014.

[13] E. Kolb and M. Turner, *The Early Universe*. Westview Press, Boulder, Colorado, 1994.

[14] H. J. Schmidt, "De Sitter space-time as an attractor solution in fourth order gravity," *Class. Quant. Grav.* **5** (1988) 233.

[15] U. Brauer, A. Rendall, and O. Reula, "The cosmic no hair theorem and the nonlinear stability of homogeneous Newtonian cosmological models," *Class. Quant. Grav.* **11** (1994) 2283–2296, arXiv:gr-qc/9403050.

[16] S. Hollands, *Correlators, Feynman diagrams, and quantum no-hair in de Sitter spacetime*, vol. 68. 2010. arXiv:1010.5367.

[17] D. Marolf and I. A. Morrison, "Infrared stability of de Sitter space: Loop corrections to scalar propagators," *Phys. Rev. D* **82** no. 10, (2010) 1–20, arXiv:1006.0035.

[18] D. Marolf and I. A. Morrison, "Infrared stability of de Sitter QFT: Results at all orders," *Phys. Rev. D* **84** no. 4, (2011) 1–15.

[19] G. W. Gibbons and S. W. Hawking, "Cosmological event horizons, thermodynamics, and particle creation," *Phys. Rev. D* **15** (May, 1977) 2738–2751.

[20] M. Spradlin, A. Strominger, and A. Volovich, "Les Houches lectures on de Sitter space," in *Unity from duality: Gravity, gauge theory and strings. Proceedings, NATO Advanced Study Institute, Euro Summer School, 76th session, Les Houches, France, July 30-August 31, 2001*, pp. 423–453. 2001. arXiv:hep-th/0110007.

[21] T. Banks, "Cosmological breaking of supersymmetry?," *Int. J. Mod. Phys. A* **16** (2001) 910–921, arXiv:hep-th/0007146.

[22] E. Witten, "Quantum gravity in de Sitter space," in *Strings 2001: International Conference Mumbai, India, January 5-10, 2001*. 2001. arXiv:hep-th/0106109.

[23] L. Dyson, M. Kleban, and L. Susskind, "Disturbing implications of a cosmological constant," *JHEP* **10** (2002) 011, arXiv:hep-th/0208013.

[24] M. K. Parikh and E. P. Verlinde, "De Sitter holography with a finite number of states," *JHEP* **01** (2005) 054, arXiv:hep-th/0410227.

[25] A. Albrecht and L. Sorbo, "Can the universe afford inflation?," *Phys. Rev. D* **70** (2004) 063528, arXiv:hep-th/0405270.

[26] S. M. Carroll and J. Chen, "Spontaneous inflation and the origin of the arrow of time," arXiv:hep-th/0410270.

[27] A. Albrecht, "De Sitter equilibrium as a fundamental framework for cosmology," *J. Phys.*

Conf. Ser. **174** (2009) 012006, arXiv:0906.1047.

[28] P. B. Krishna and T. K. Mathew, "Holographic equipartition and the maximization of entropy," arXiv:1702.02787.

[29] N. Bao, C. Cao, S. M. Carroll, and L. McAllister, "Quantum circuit cosmology: The expansion of the universe since the first qubit," arXiv:1702.06959.

[30] T. Jacobson, "Thermodynamics of spacetime: The Einstein equation of state," *Phys. Rev. Lett.* **75** (1995) 1260–1263, arXiv:gr-qc/9504004.

[31] T. Padmanabhan, "Thermodynamical aspects of gravity: New insights," *Rept. Prog. Phys.* **73** (2010) 046901, arXiv:0911.5004.

[32] E. P. Verlinde, "On the origin of gravity and the laws of Newton," *JHEP* **1104** (2011) 029, arXiv:1001.0785.

[33] T. Jacobson, "Entanglement equilibrium and the einstein equation," *Phys. Rev. Lett.* **116** (2016) 201101, arXiv:1505.04753.

[34] S. M. Carroll and G. N. Remmen, "What is the entropy in entropic gravity?," *Phys. Rev. D* **93** no. 12, (2016) 124052, arXiv:1601.07558.

[35] B. Swingle, "Entanglement renormalization and holography," *Phys. Rev. D* **86** no. 6, (2012) 065007, arXiv:0905.1317.

[36] M. Van Raamsdonk, "Building up spacetime with quantum entanglement," *Gen. Rel. Grav.* **42** (2010) 2323–2329, arXiv:1005.3035. [Int. J. Mod. Phys.D19,2429(2010)].

[37] N. Lashkari, M. B. McDermott, and M. Van Raamsdonk, "Gravitational dynamics from entanglement 'thermodynamics'," *JHEP* **04** (2014) 195, arXiv:1308.3716.

[38] T. Faulkner, M. Guica, T. Hartman, R. C. Myers, and M. Van Raamsdonk, "Gravitation from entanglement in holographic CFTs," *JHEP* **1403** (2014) 051, arXiv:1312.7856.

[39] B. Swingle, "Constructing holographic spacetimes using entanglement renormalization," arXiv:1209.3304.

[40] B. Swingle and M. Van Raamsdonk, "Universality of gravity from entanglement," arXiv:1405.2933.

[41] J. Maldacena and L. Susskind, "Cool horizons for entangled black holes," *Fortschritte der Physik* **61** (2013) 781–811, arXiv:1306.0533.

[42] C. Cao, S. M. Carroll, and S. Michalakis, "Space from Hilbert space: Recovering geometry from bulk entanglement," *Phys. Rev. D* **95** no. 2, (2017) 024031, arXiv:1606.08444.

[43] J. D. Bekenstein, "Generalized second law of thermodynamics in black-hole physics," *Phys. Rev. D* **9** no. 12, (1974) 3292–3300.

[44] R. Bousso and N. Engelhardt, "Generalized second law for cosmology," *Phys. Rev. D* **93** (2016) 024025, arXiv:1510.02099.

[45] R. Bousso, "Holography in general space-times," *JHEP* **06** no. 028, (1999) 24, arXiv:hep-th/9906022.

[46] S. A. Hayward, "General laws of black-hole dynamics," *Phys. Rev. D* **49** (Jun, 1994) 6467–6474, arXiv:gr-qc/9303006.

[47] A. Ashtekar and B. Krishnan, "Dynamical horizons: energy, angular momentum, fluxes, and balance Laws," *Phys. Rev.Lett.* **89** no. 26, (2002) 261101, arXiv:gr-qc/0207080.

[48] A. Ashtekar and B. Krishnan, "Dynamical horizons and their properties," *Phys. Rev. D* **68** no. 10, (2003) 44, arXiv:gr-qc/0308033.

[49] A. Ashtekar and B. Krishnan, "Isolated and dynamical horizons and their applications," *Living Rev. Relat.* **7** (2004) , arXiv:gr-qc/0407042.

[50] I. Booth, "Black-hole boundaries," *Can. J. Phys.* **83** no. 11, (2005) 1073–1099, arXiv:gr-qc/0508107.

[51] I. Booth, L. Brits, J. A. Gonzalez, and C. V. D. Broeck, "Marginally trapped tubes and dynamical horizons," *Class. Quant. Grav.* **23** no. 2, (2006) 413–439, arXiv:gr-qc/0506119.

[52] R. Bousso and N. Engelhardt, "New area law in general relativity," *Physical Review Letters* **115** no. 8, (2015) , arXiv:1504.07627.

[53] R. Bousso and N. Engelhardt, "Proof of a new area law in general relativity," *Phys. Rev. D* **92** no. 4, (2015) 1–14, arXiv:1504.07660.

[54] F. Sanches and S. J. Weinberg, "Holographic entanglement entropy conjecture for general spacetimes," *Phys. Rev. D* **94** no. 8, (2016) 1–16, arXiv:1603.05250.

[55] J. M. Bardeen, B. Carter, and S. W. Hawking, "The four laws of black hole mechanics," *Commun. Math. Phys.* **31** no. 2, (1973) 161–170.

[56] T. Jacobson, "Introductory lectures on black hole thermodynamics," *Given at Utrecht U. in* (1996) 30–59.

[57] S. W. Hawking, "Black holes in general relativity," *Commun. Math. Phys.* **25** no. 2, (1972) 152–166.

[58] R. Bousso, Z. Fisher, S. Leichenauer, and A. C. Wall, "Quantum focusing conjecture," *Phys.

Rev. D **93** no. 6, (2016) 1–20, arXiv:1506.02669.

[59] R. Bousso, H. Casini, Z. Fisher, and J. Maldacena, "Proof of a quantum Bousso bound," *Phys. Rev. D* **90** no. 4, (2014) 044002, arXiv:1404.5635.

[60] P. T. Saunders, "Observations in some simple cosmological models with shear," *Monthly Notices of the Royal Astronomical Society* **142** (1969) 213–227.

[61] P. Fleury, F. Nugier, and G. Fanizza, "Geodesic-light-cone coordinates and the Bianchi I spacetime," *J. Cosmol. Astropart. Phys.* **2016** no. 06, (2016) 008–008, arXiv:1602.04461.

[62] A. Chatwin-Davies, "Companion calculations for 'cosmic equilibration: A holographic no-hair theorem from the generalized second law'." https://doi.org/10.5281/zenodo.438142, 2017.

[63] V. Mukhanov, H. Feldman, and R. Brandenberger, "Theory of cosmological perturbations," *Phys. Rep.* **215** no. 5, (1992) 203 – 333.

[64] N. Birrell and P. Davies, *Quantum Fields in Curved Space*. Cambridge University Press, Cambridge, 1982.