

**Relative information entropy and Weyl curvature of the inhomogeneous Universe**Nan Li<sup>\*</sup>*Department of Physics, College of Sciences, Northeastern University, Shenyang 110819, China*Thomas Buchert<sup>†</sup>*Université de Lyon, Observatoire de Lyon, Centre de Recherche Astrophysique de Lyon,  
CNRS UMR 5574: Université Lyon 1 and École Normale Supérieure de Lyon,  
9 Avenue Charles André, Saint-Genis-Laval F-69230, France*Akio Hosoya<sup>‡</sup>*Department of Physics, Tokyo Institute of Technology, Oh-Okayama, Meguro, Tokyo 152-8551, Japan*Masaaki Morita<sup>§</sup>*Department of Integrated Arts and Science, Okinawa National College of Technology, 905 Henoko, Nago, Okinawa 905-2192, Japan*Dominik J. Schwarz<sup>||</sup>*Fakultät für Physik, Universität Bielefeld, Universitätsstraße 25, Bielefeld D-33615, Germany  
(Received 16 August 2012; published 22 October 2012)*

Penrose conjectured a connection between entropy and Weyl curvature of the Universe. This is plausible, as the almost homogeneous and isotropic Universe at the onset of structure formation has negligible Weyl curvature, which then grows (relative to the Ricci curvature) due to the formation of large-scale structure, and thus reminds us of the second law of thermodynamics. We study two scalar measures to quantify the deviations from a homogeneous and isotropic space-time: the relative information entropy and a Weyl tensor invariant, and show their relation to the averaging problem. We calculate these two quantities up to second order in standard cosmological perturbation theory and find that they are correlated and can be linked via the kinematical backreaction of a spatially averaged universe model.

DOI: [10.1103/PhysRevD.86.083539](https://doi.org/10.1103/PhysRevD.86.083539)

PACS numbers: 98.80.Jk, 02.40.Ky, 89.70.Cf

**I. INTRODUCTION**

The starting point of cosmology is commonly based on the simple Friedmann-Lemaître-Robertson-Walker (FLRW) model, which describes a homogeneous and isotropic space-time, as a mathematical realization of the strong cosmological principle. However, due to gravitational instability, our Universe obviously deviates from this simple model locally, and hosts various large-scale structures, e.g., the anisotropies in the cosmic microwave background. Spatial homogeneity and isotropy may be valid only on scales larger than hundreds of Mpc, and the effects due to inhomogeneities within these scales are in the focus of many investigations. While the conventional approach assumes that the FLRW model also provides a valid description *on average* of the inhomogeneous Universe, in general we expect structure formation to affect the average evolution: the FLRW model is in general not only locally, but also globally gravitationally unstable [1]. As a first step, we shall here confront

average properties of inhomogeneous models with a perturbative modeling of inhomogeneities using standard perturbation theory.

Penrose conjectured that some scalar invariant of the Weyl curvature tensor is a monotonically growing function of time and could be identified with the gravitational entropy of the Universe [2,3]. This conjecture, also known as the “Weyl curvature hypothesis,” to our knowledge was neither formulated in a rigorous way, nor has the precise notion of “entropy” been specified. However, it is very plausible that some entropy measure exists, as a FLRW space-time has vanishing Weyl curvature and nonvanishing Ricci curvature, whereas black hole solutions of space-time have nonzero Weyl and (for the simplest models) zero Ricci curvatures. As it is assumed that the Universe evolves from an almost homogeneous and isotropic space-time towards an ensemble of randomly distributed black holes in the far future, the Weyl curvature seems to grow relative to the Ricci curvature of the Universe. A monotonically increasing function reminds us of the second law of thermodynamics, so Penrose wondered whether the Weyl tensor somehow describes or is related to some notion of “gravitational entropy” of the Universe. An attempt to define the gravitational entropy has been made for a linearly perturbed FLRW solution, taking tensor perturbations as purely gravitational ones [4].

<sup>\*</sup>linan@mail.neu.edu.cn<sup>†</sup>buchert@obs.univ-lyon1.fr<sup>‡</sup>ahosoya@th.phys.titech.ac.jp<sup>§</sup>morita@okinawa-ct.ac.jp<sup>||</sup>dschwarz@physik.uni-bielefeld.de

We propose in this paper to identify the notion of “entropy” with an information theoretical measure. On a finite patch of the Universe such a measure can be defined, once one has specified an averaging procedure. In the context of cosmology, such an averaging procedure was proposed by one of the present authors [5,6], and it has been applied to the Kullback-Leibler information entropy by some of the present authors in Ref. [7]. Here we combine those findings with cosmological perturbation theory, motivated by the Weyl curvature hypothesis.

In Sec. II, we first introduce two scalar measures to quantify inhomogeneities in the Universe: the relative information entropy and a Weyl tensor invariant, and next show that they both can be related to the averaging problem in a perturbed universe model in Sec. III. In Sec. IV, we calculate these two quantities by means of cosmological perturbation theory up to second order and prove that they can be correlated via the kinematical back-reaction of a spatially averaged universe model in Sec. V. We conclude in Sec. VI.

## II. MEASURES OF INHOMOGENEITY IN THE PERTURBED UNIVERSE

Since our Universe is regionally inhomogeneous and anisotropic, the value of a physical observable  $O(t, \mathbf{x})$  at a specific time and place is of little significance to us, and only its average  $\langle O(t, \mathbf{x}) \rangle_{\mathcal{D}}$  over a domain  $\mathcal{D}$  is meaningful and practical in cosmology. The importance of the averaging problem in the perturbed Universe lies in the fact that many cosmological observables are averaged quantities. The most obvious example is the Hubble constant  $H_0$ . We pick  $N$  galaxies in a local volume  $V$  and measure their luminosity distances  $d_i$  and recession velocities  $v_i$ , and  $H_0 = \frac{1}{N} \sum_{i=1}^N \frac{v_i}{d_i}$ . Therefore, in the limit of a big sample, it naturally turns into a volume average  $H_0 = \frac{1}{V} \int_{\mathcal{D}} d^3x$ . (We should state that cosmological information is generally encoded on the past light cone, so the averaging problem should be done over a light-cone volume [8,9]. However, for objects with redshifts  $\ll 1$ , spatial averaging on a hypersurface of constant time is already a good approximation because the Hubble rate does not change significantly on the temporal scale involved.)

There are many ways to quantify the deviation of  $O(t, \mathbf{x})$  from  $\langle O(t, \mathbf{x}) \rangle_{\mathcal{D}}$ . For instance, the variance  $\text{var}(O) := \langle (\delta O)^2 \rangle_{\mathcal{D}} = \langle (O - \langle O \rangle_{\mathcal{D}})^2 \rangle_{\mathcal{D}} = \langle O^2 \rangle_{\mathcal{D}} - \langle O \rangle_{\mathcal{D}}^2$  is a good choice. In this section, we will present two different methods to describe this deviation: one is the Kullback-Leibler relative entropy in information theory, and the other is the Weyl tensor in differential geometry. They both can characterize the degree of inhomogeneity of the perturbed Universe.

### A. Kullback-Leibler relative information entropy

The relative information entropy was introduced by Kullback and Leibler [10] first in information theory as

$$S\{p \parallel q\} := \sum_i p_i \ln \frac{p_i}{q_i},$$

where  $\{p_i\}$  and  $\{q_i\}$  denote the actual and presumed probability distributions, respectively. It is easy to see that this entropy vanishes if the two distributions agree, and will be positive definite if the actual distribution departs from the presumed one. Therefore, the relative information entropy may quantify the divergence from the actual distribution to the theoretical one, which is usually used in an idealized model such as an assumed background.

This idea was generalized to the study of the evolution of an inhomogeneous Universe by Hosoya, Buchert, and Morita. In Ref. [7], a straightforward and reasonable extension of the Kullback-Leibler relative information entropy from a discrete system to a continuum in cosmology was derived from the basic principle of noncommutativity of the two operations “spatial averaging” and “temporal evolution”:

$$\frac{\mathcal{S}_{\mathcal{D}}\{\rho \parallel \langle \rho \rangle_{\mathcal{D}}\}}{V_{\mathcal{D}}} := \left\langle \rho \ln \frac{\rho}{\langle \rho \rangle_{\mathcal{D}}} \right\rangle_{\mathcal{D}}, \quad \text{with}$$

$$\langle \rho \rangle_{\mathcal{D}} - \langle \dot{\rho} \rangle_{\mathcal{D}} = -\frac{\dot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}}. \quad (1)$$

Here,  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$  is the relative information entropy density in a domain  $\mathcal{D}$  with the volume  $V_{\mathcal{D}}$ , and  $\rho$  and  $\langle \rho \rangle_{\mathcal{D}}$  are the actual and averaged energy densities in the Universe, as the analogues of  $\{p_i\}$  and  $\{q_i\}$ . (For a general discussion of the Kullback-Leibler relative information entropy and its applications in different models in cosmology, we refer to Refs. [11,12].)

### B. Weyl tensor

The Weyl tensor is a measure of the curvature of a pseudo-Riemannian manifold. In four-dimensional space-time, it is defined as

$$C_{\mu\nu\lambda\rho} := R_{\mu\nu\lambda\rho} + \frac{1}{2}(g_{\mu\rho}R_{\nu\lambda} + g_{\nu\lambda}R_{\mu\rho} - g_{\mu\lambda}R_{\nu\rho} - g_{\nu\rho}R_{\mu\lambda}) + \frac{1}{6}(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda})R, \quad (2)$$

where  $R_{\mu\nu\lambda\rho}$  and  $R_{\mu\nu}$  are the coefficient functions of the Riemann and Ricci tensors, and  $R_{\mu\nu} := R^{\lambda}_{\mu\lambda\nu}$ . Our sign convention is based on the metric signature  $(-, +, +, +)$ . The Weyl tensor has the same symmetries as the Riemann tensor:  $C_{\mu\nu\lambda\rho} = -C_{\nu\mu\lambda\rho} = -C_{\mu\nu\rho\lambda}$ ,  $C_{\mu\nu\lambda\rho} = C_{\lambda\rho\mu\nu}$ , and  $C_{\mu\nu\lambda\rho} + C_{\mu\lambda\rho\nu} + C_{\mu\rho\nu\lambda} = 0$ . But the Weyl tensor is traceless, i.e., the contraction on any pair of indices yields zero, e.g.,  $C^{\lambda}_{\mu\lambda\nu} = 0$ . So the contraction of the Weyl tensor  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  is the principal scalar invariant we can construct. (Strictly speaking, there is another principal scalar invariant  $\star C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$ , with  $\star C_{\mu\nu\lambda\rho}$  being the left dual of the Weyl tensor, but in this paper we restrict our exploration to  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  for simplicity.)

The Weyl tensor possesses an important property: if the Weyl tensor vanishes, the metric of space-time is locally conformally flat. This means that we may transform the metric tensor to a constant tensor in a local coordinate system by a conformal transformation. Two metric tensors are said to be conformally equivalent if  $\tilde{g}_{\mu\nu}(\tilde{x}) = \Omega^2(x)g_{\mu\nu}(x)$ . In this circumstance, they have the same Weyl tensor, and their null geodesics coincide. (For an introduction to conformal transformation and conformal equivalence, see Ref. [13].)

We assume that at early times when the Universe is still almost homogeneous and perturbations have not grown significantly, the metric of space-time can be approximately represented by the zero-curvature FLRW one,  $ds^2 = -dt^2 + a^2(t)\delta_{ij}dx^i dx^j$ . By a conformal transformation,  $dt = a(t)d\eta$ , we obtain  $ds^2 = a^2(\eta)(-d\eta^2 + \delta_{ij}dx^i dx^j)$ . Therefore, the metric is conformally equivalent to a flat Minkowski one, and its Weyl tensor vanishes automatically. Whereas, in the late Universe, space-time becomes highly inhomogeneous, and its metric departs from the FLRW one and is thus not conformally flat anymore. Hence, the Weyl tensor appears in the perturbed Universe.

Penrose conjectured an analogue [2,3] of the emergence of the Weyl tensor to the entropy increasing in thermodynamics of black holes. For a black hole with mass  $M$ , the Schwarzschild metric is  $ds^2 = -(1 - \frac{2GM}{r})dt^2 + (1 - \frac{2GM}{r})^{-1}dr^2 + r^2 d\Omega^2$ , and  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} = 48(GM)^2/r^6$ . Meanwhile, the entropy  $S$  of the Schwarzschild black hole is  $S = \frac{k_B}{4G\hbar}A = \frac{k_B}{4G\hbar} \times 4\pi(2GM)^2$ , with  $A = 4\pi(2GM)^2$  being its area. We observe clearly that  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  is proportional to  $S$ , so there may be some latent relation of the Weyl tensor to the entropy of black holes and gravitational fields.

In cosmology, during the process of structure formation, the cosmic structures decouple from the global expansion of the Universe little by little and become gravitationally bound systems. Generally speaking, these systems will eventually end their evolution as black holes. Consequently, black holes and their corresponding Weyl tensors will together arise here and there in the Universe, and the average of  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  can thus be considered as another measure of structure formation, or another kind of entropy in some sense. But we should point out that Penrose's conjecture is not a well formulated mathematical statement, and the aim of this paper is to explore its relation to the relative information entropy and other relevant physical quantities.

In this section, we have discussed two different methods to measure the degree of inhomogeneity of the Universe from the unperturbed background: one is the relative entropy in information theory, and the other is the Weyl tensor in differential geometry. They root from different branches of science and are seemingly uncorrelated at first look. However, besides these apparent distinctions, they have essential points in common. First, they are

both related to the entropy increasing in the evolution of the Universe; second, and more importantly, they are both related on the level of spatially averaged quantities in the inhomogeneous Universe, as we shall demonstrate. Therefore, in the next sections, we shall explain the averaging problem in the perturbed Universe, calculate  $\mathcal{S}_{\mathcal{D}}$  and  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  in cosmological perturbation theory, and finally show their relation in Sec. V.

### III. AVERAGING PROBLEM AND COSMOLOGICAL PERTURBATION THEORY

In this section, we first recall the averaging procedure in an inhomogeneous Universe together with the effective Friedmann equations for an irrotational dust universe model. The matter model ‘‘irrotational dust’’ is assumed throughout the paper and implied when we talk about ‘‘cosmic continuum.’’ Next, we turn to cosmological perturbation theory and provide all the necessary mathematical preparations for Secs. IV and V, the main part of this paper.

#### A. Averaging problem in cosmology

The essence of the averaging problem in cosmology lies in the noncommutation of spatial averaging and temporal evolution. As a consequence, inhomogeneities and anisotropies influence the evolution of the background (averaged) universe model. Let us start from a compact domain  $\mathcal{D}$  with metric perturbations at initial time  $t_i$ . If we first smooth out the fluctuations of the metric at  $t_i$ , what remains is merely a simple FLRW model. Afterwards, the evolution of this averaged model to time  $t$  is nothing but a pure expansion. In contrast, if we exchange the order of these two operations: first follow the evolution of the perturbed space-time from  $t_i$  to  $t$ , and then take the average of physical observables in the resulting domain at  $t$ , we clearly arrive at a different result. This mechanism is usually named as cosmological backreaction [14,15]. (See Ref. [16] for a comprehensive review of the backreaction formalism, and Ref. [17] for a recent overview.)

To understand the averaging problem mathematically, we start from the kinematics of the perturbed Universe. The covariant derivative of the four-velocity of the cosmic continuum can be decomposed as  $u_{\mu;\nu} = \frac{1}{3}h_{\mu\nu}\theta + \sigma_{\mu\nu} + \omega_{\mu\nu}$ , where  $h_{\mu\nu}$  is the projection operator,  $\sigma^{\mu}_{\nu} := h^{\mu}_{\alpha}h^{\beta}_{\nu}[\frac{1}{2}(u^{\alpha}_{;\beta} + u^{\beta}_{;\alpha}) - \frac{1}{3}h^{\alpha}_{\beta}u^{\lambda}_{;\lambda}]$  is the shear tensor,  $\sigma^2 := \frac{1}{2}\sigma^{\mu}_{\nu}\sigma^{\nu}_{\mu}$  is the shear scalar, and  $\theta := u^{\lambda}_{;\lambda}$  is the expansion scalar. In the following, the vorticity tensor  $\omega_{\mu\nu}$  has to vanish by construction in a flow-orthogonal foliation of space-time; we assume that this is a good approximation on large scales, since vorticity decays in perturbation theory due to the expansion of the Universe. Moreover, we neglect radiation at the late times, and use the dust approximation. Last, in the following, we perform the concrete calculation for the simplest Einstein-de Sitter dust

model in which the cosmological constant is supposed to vanish. (It is reasonable because the Einstein-de Sitter model with small perturbations is believed to be an excellent approximation in most of the matter-dominated epoch, during which our calculation applies. A generalization with the cosmological constant is straightforward, but that will not change our following statements in a qualitative way.)

Below, we follow the averaging formalism proposed in Refs. [5,6]. The metric of the perturbed Universe can be written in the synchronous gauge as  $ds^2 = -dt^2 + g_{ij}(t, \mathbf{x})dx^i dx^j$ , and the spatial Riemannian volume average of an observable  $O(t, \mathbf{x})$  in a comoving domain  $\mathcal{D}$  at time  $t$  is defined as

$$\langle O \rangle_{\mathcal{D}} := \frac{1}{V_{\mathcal{D}}(t)} \int_{\mathcal{D}} O(t, \mathbf{x}) \sqrt{\det g_{ij}} d^3x,$$

with  $V_{\mathcal{D}}(t) := \int_{\mathcal{D}} \sqrt{\det g_{ij}} d^3x$  being the Riemannian volume of  $\mathcal{D}$ . An effective volume scale factor,  $a_{\mathcal{D}}(t)/a_{\mathcal{D}}(t_0) := (V_{\mathcal{D}}(t)/V_{\mathcal{D}}(t_0))^{1/3}$ , can thus be introduced. For later evaluation within perturbation theory, we define the Euclidian average  $\langle O \rangle := \int O(t, \mathbf{x}) d^3x / \int d^3x$ , i.e., the integral without  $\sqrt{\det g_{ij}}$  on the background cosmology.

By averaging the energy constraint and the Raychaudhuri equation, we arrive at the effective Friedmann equations for the irrotational dust universe model (the Buchert equations) [5,6],

$$\begin{aligned} \left(\frac{\dot{a}_{\mathcal{D}}}{a_{\mathcal{D}}}\right)^2 + \frac{k_{\mathcal{D}}}{a_{\mathcal{D}}^2} &= \frac{8\pi G}{3} \langle \rho \rangle_{\mathcal{D}} - \frac{\mathcal{Q}_{\mathcal{D}} + \mathcal{W}_{\mathcal{D}}}{6}, \\ \frac{\ddot{a}_{\mathcal{D}}}{a_{\mathcal{D}}} &= -\frac{4\pi G}{3} \langle \rho \rangle_{\mathcal{D}} + \frac{\mathcal{Q}_{\mathcal{D}}}{3}, \end{aligned}$$

$$\langle \rho \rangle_{\mathcal{D}} + 3 \frac{\dot{a}_{\mathcal{D}}}{a_{\mathcal{D}}} \langle \rho \rangle_{\mathcal{D}} = 0.$$

In these equations, we observe that besides the ordinary entries in the Friedmann equations for the FLRW model, there appear two additional terms: the deviation of the averaged spatial curvature from a constant-curvature model,  $\mathcal{W}_{\mathcal{D}} := \langle \mathcal{R} \rangle_{\mathcal{D}} - 6k_{\mathcal{D}}/a_{\mathcal{D}}^2$ , and the kinematical backreaction  $\mathcal{Q}_{\mathcal{D}}$ ,

$$\mathcal{Q}_{\mathcal{D}} := \frac{2}{3} (\langle \theta^2 \rangle_{\mathcal{D}} - \langle \theta \rangle_{\mathcal{D}}^2) - 2 \langle \sigma^2 \rangle_{\mathcal{D}}. \quad (3)$$

These two terms are not independent and can be linked via an integrability condition  $(a_{\mathcal{D}}^6 \mathcal{Q}_{\mathcal{D}}) + a_{\mathcal{D}}^4 (a_{\mathcal{D}}^2 \mathcal{W}_{\mathcal{D}}) = 0$ . In the following, we shall be using a background model with  $k_{\mathcal{D}} = 0$  so that the variable  $\mathcal{W}_{\mathcal{D}}$  is equal to  $\langle \mathcal{R} \rangle_{\mathcal{D}}$ .

## B. Cosmological perturbation theory

For the perturbative calculations of the relative information entropy and the Weyl tensor invariant in Sec. IV, we briefly list the basics of cosmological perturbation theory.

Here we concentrate on the scalar modes at the first order for a flat irrotational dust model.

In the comoving synchronous gauge, the perturbed metric of space-time reads

$$ds^2 = -dt^2 + a^2(t) [(1 - 2\Psi)\delta_{ij} + D_{ij}\chi] dx^i dx^j. \quad (4)$$

Here the scale factor  $a(t) = (t/t_0)^{2/3}$  ( $a(t_0) := 1$ ) is slightly different from the effective scale factor  $a_{\mathcal{D}}(t)$ , and their relation was provided in Ref. [18].  $\Psi(t, \mathbf{x})$  and  $\chi(t, \mathbf{x})$  are the scalar metric perturbations,  $D_{ij} := \partial_i \partial_j - \frac{1}{3} \delta_{ij} \Delta$ , and  $\Delta$  is the three-dimensional Laplace operator. Substituting this perturbed metric into the Einstein equations, we obtain the solutions for  $\Psi$  and  $\chi$  (only the growing modes are taken into account) [18,19],

$$\begin{aligned} \Psi(t, \mathbf{x}) &= \frac{1}{2} \Delta \varphi(\mathbf{x}) t_0^{4/3} t^{2/3} + \frac{5}{3} \varphi(\mathbf{x}), \\ \chi(t, \mathbf{x}) &= -3 \varphi(\mathbf{x}) t_0^{4/3} t^{2/3}, \end{aligned} \quad (5)$$

where  $\varphi(\mathbf{x})$  is the time-independent peculiar-gravitational potential defined from the Poisson equation,  $\Delta \varphi(\mathbf{x}) := 4\pi G a^2(t) \rho^{(1)}(t, \mathbf{x})$ . Similarly, we may directly find the energy density of the background universe model  $\rho^{(0)}(t)$  and the perturbation  $\rho^{(1)}(t, \mathbf{x})$  over it,

$$\rho^{(0)}(t) = \frac{1}{6\pi G t^2}, \quad \rho^{(1)}(t, \mathbf{x}) = \frac{1}{4\pi G} \left(\frac{t_0}{t}\right)^{4/3} \Delta \varphi(\mathbf{x}). \quad (6)$$

We should emphasize that in the following sections we shall perform all the perturbative calculations up to second order. However, here we only mention the first order results, but this is already enough, as we shall see immediately that both the relative information entropy and contraction of the Weyl tensor are second order quantities and do not involve zeroth order quantities, and can thus be constructed by using only the first order perturbations. All this will be carefully explained in Secs. IV A and IV B [20].

The above discussion completes all the mathematical preparations for the forthcoming perturbative calculations.

## IV. PERTURBATIVE CALCULATIONS OF THE RELATIVE INFORMATION ENTROPY AND THE WEYL TENSOR INVARIANT

We now move on to the detailed perturbative calculations of the behavior of the relative information entropy  $\mathcal{S}_{\mathcal{D}}$  and the spatial average of the contraction of the Weyl tensor  $\langle C_{\mu\nu\lambda\rho} C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  in the evolution of the perturbed Universe.

### A. Perturbative calculation of the relative information entropy

At the onset of structure formation, when energy density perturbations are small, we expand  $\rho$  to second order, and the relative information entropy density is

$$\begin{aligned}
\frac{\mathcal{S}_{\mathcal{D}}}{V_{\mathcal{D}}} &= \left\langle \rho \ln \frac{\rho}{\langle \rho \rangle_{\mathcal{D}}} \right\rangle_{\mathcal{D}} \\
&= \left\langle (\rho^{(0)} + \rho^{(1)} + \rho^{(2)}) \ln \left( \frac{\rho^{(0)} + \rho^{(1)} + \rho^{(2)}}{\rho^{(0)} + \langle \rho^{(1)} \rangle_{\mathcal{D}} + \langle \rho^{(2)} \rangle_{\mathcal{D}}} \right) \right\rangle_{\mathcal{D}} \\
&= \frac{1}{2} \frac{\langle (\rho^{(1)})^2 \rangle - \langle \rho^{(1)} \rangle^2}{\rho^{(0)}} + \text{higher order terms.} \quad (7)
\end{aligned}$$

We see from Eq. (7) that although the second order perturbation  $\rho^{(2)}$  enters the expression of  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$ , it does not explicitly enter into the final result at second order. This is understandable, as the leading term of the entropy density is the variance of the energy density, which is already a second order term. Moreover, at second order, we are entitled to use the Euclidian average  $\langle \cdot \cdot \rangle$  to approximate  $\langle \cdot \cdot \rangle_{\mathcal{D}}$  since their differences are terms of even higher order [20].

Substituting the solution of  $\rho^{(1)}$  from Eq. (6) into Eq. (7), we obtain

$$\begin{aligned}
\frac{\mathcal{S}_{\mathcal{D}}}{V_{\mathcal{D}}} &= \frac{3}{16\pi G} \frac{t_0^{8/3}}{t^{2/3}} [ \langle (\Delta\varphi)^2 \rangle - \langle \Delta\varphi \rangle^2 ] \\
&= \frac{3}{16\pi G} \frac{t_0^{8/3}}{t^{2/3}} \text{var}(\Delta\varphi) \propto \frac{1}{t^{2/3}}. \quad (8)
\end{aligned}$$

We find clearly from Eq. (8) that  $\mathcal{S}_{\mathcal{D}}$  is positive, which is consistent with its definition. Since  $V_{\mathcal{D}} \propto a_{\mathcal{D}}^3 \propto a^3 \propto t^2$ , we observe from Eq. (8) that

$$\mathcal{S}_{\mathcal{D}} \propto V_{\mathcal{D}}^{2/3} \propto a_{\mathcal{D}}^2 \propto a^2 \propto t^{4/3}. \quad (9)$$

Equation (9) shows that  $\mathcal{S}_{\mathcal{D}}$  increases in the evolution of the perturbed Universe, and thus it indeed is a measure that characterizes the degree of structure formation, as we expect. Furthermore, the results for  $\dot{\mathcal{S}}_{\mathcal{D}}/V_{\mathcal{D}}$  and  $\ddot{\mathcal{S}}_{\mathcal{D}}/V_{\mathcal{D}}$  can directly be obtained,

$$\frac{\dot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{4}{3t} \frac{\mathcal{S}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{1}{4\pi G} \frac{t_0^{8/3}}{t^{5/3}} \text{var}(\Delta\varphi), \quad (10)$$

$$\frac{\ddot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{4}{9t^2} \frac{\mathcal{S}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{1}{12\pi G} \frac{t_0^{8/3}}{t^{8/3}} \text{var}(\Delta\varphi). \quad (11)$$

We find that both of them are positive, meaning that  $\mathcal{S}_{\mathcal{D}}$  not only increases monotonically, but in an accelerated manner.

The general result for the evolution of the relative information entropy has a profound relation to the non-commutation of spatial averaging and temporal evolution in the averaging problem. It was proved in Ref. [7] that

$$\begin{aligned}
\frac{\dot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} &= \langle \dot{\rho} \rangle_{\mathcal{D}} - \langle \rho \rangle_{\mathcal{D}} \\
&= \langle \rho \rangle_{\mathcal{D}} \langle \dot{\theta} \rangle_{\mathcal{D}} - \langle \rho \dot{\theta} \rangle_{\mathcal{D}} \\
&= -\langle \delta\rho \delta\theta \rangle_{\mathcal{D}}. \quad (12)
\end{aligned}$$

Therefore, in the process of structure formation, whether for an overdense or underdense region, we have both

- (1) Overdense region:  $\delta\rho > 0$  and  $\delta\theta < 0$ ,  $-\langle \delta\rho \delta\theta \rangle_{\mathcal{D}} > 0$ , so  $\dot{\mathcal{S}}_{\mathcal{D}} > 0$ .
- (2) Underdense region:  $\delta\rho < 0$  and  $\delta\theta > 0$ ,  $-\langle \delta\rho \delta\theta \rangle_{\mathcal{D}} > 0$ , so  $\dot{\mathcal{S}}_{\mathcal{D}} > 0$ .

Thus, generally speaking,  $\mathcal{S}_{\mathcal{D}}$  increases monotonically, and this is in agreement with our linear perturbative results.

Last, we can prove that the time convexity of relative information entropy  $\frac{\ddot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}}$  is

$$\begin{aligned}
\frac{\ddot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} &= 4\pi G \text{var}(\rho) + \frac{1}{3} \langle \rho (\delta\theta)^2 \rangle_{\mathcal{D}} + 2 \langle \rho \sigma^2 \rangle_{\mathcal{D}} \\
&\quad + \langle \rho \rangle_{\mathcal{D}} \mathcal{Q}_{\mathcal{D}} - \frac{2}{3} \frac{\dot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} \langle \theta \rangle_{\mathcal{D}}.
\end{aligned}$$

Unfortunately, it is a highly nontrivial task to figure out the sign of this exact result. But we see from Eq. (11) that in a perturbative approach, up to second order,  $\ddot{\mathcal{S}}_{\mathcal{D}}/V_{\mathcal{D}}$  is still positive.

At this level of calculation, we can safely assume that the relative information entropy increases monotonically during the evolution of the Universe, and it thus serves as a measure of structure formation.

## B. Perturbative calculation of the Weyl tensor invariant

Starting from the perturbed metric in Eq. (4),

$$ds^2 = -dt^2 + a^2(t)[(1 - 2\Psi)\delta_{ij} + D_{ij}\chi]dx^i dx^j,$$

the calculation for the contraction of the Weyl tensor  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  is straightforward, although a little bit tedious. Here we show some intermediate steps before giving the final result. Due to the symmetry of the Weyl tensor, 112 of its 256 components vanish automatically, and all the remaining ones are the derivatives of a function  $C$ , a combination of the metric perturbations  $\Psi$  and  $\chi$ ,

$$C := \Psi + \frac{1}{6}\Delta\chi - \frac{1}{2}(aa\dot{\chi} + a^2\ddot{\chi}).$$

For example, some typical components are

$$\begin{aligned}
C_{0101} &= \frac{1}{2}D_{11}C, & C_{0102} &= \frac{1}{2}D_{12}C, \\
C_{1212} &= \frac{a^2}{2}(D_{11} + D_{22})C, & C_{1213} &= \frac{a^2}{2}D_{23}C.
\end{aligned}$$

Therefore, direct calculation shows

$$\begin{aligned}
C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} &= \frac{1}{a^4} \{ 2[D_{11}CD_{22}C + D_{22}CD_{33}C + D_{33}CD_{11}C] \\
&\quad + 3[(D_{11}C)^2 + (D_{22}C)^2 + (D_{33}C)^2] \\
&\quad + 4[(D_{12}C)^2 + (D_{23}C)^2 + (D_{31}C)^2] \}.
\end{aligned}$$

In a dust universe model, from the solutions in Eq. (5), we have  $C = 2\varphi$ , so

$$C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} = 8\left(\frac{t_0}{t}\right)^{8/3} \left[ \partial^i(\partial_j\varphi\partial^j\partial_i\varphi) - \partial^i(\partial_i\varphi\Delta\varphi) + \frac{2}{3}(\Delta\varphi)^2 \right],$$

and

$$\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}} = 8\left(\frac{t_0}{t}\right)^{8/3} \left[ \langle \partial^i(\partial_j\varphi\partial^j\partial_i\varphi) \rangle - \langle \partial^i(\partial_i\varphi\Delta\varphi) \rangle + \frac{2}{3}\langle (\Delta\varphi)^2 \rangle \right]. \quad (13)$$

Equation (13) implies similarly to  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$  that  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  is also a second order quantity. Since all the components of  $C_{\mu\nu\lambda\rho}$  are zero at the FLRW background, i.e., there are no product terms of second order perturbations with (here vanishing) zeroth order terms, we only need the linear perturbation theory to calculate  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  up to second order.

Equations (8), (10), (11), and (13), display all the perturbative results for the relative information entropy and the Weyl tensor invariant. However, their relation seems still unclear, but we shall prove immediately that they actually can be related via the kinematical backreaction term  $\mathcal{Q}_{\mathcal{D}}$ .

## V. RELATION OF THE KINEMATICAL BACKREACTION

In this section, we shall finally show the relation of the relative information entropy and the Weyl tensor invariant with the kinematical backreaction term  $\mathcal{Q}_{\mathcal{D}}$ ,

$$\mathcal{Q}_{\mathcal{D}} := \frac{2}{3}(\langle \theta^2 \rangle_{\mathcal{D}} - \langle \theta \rangle_{\mathcal{D}}^2) - 2\langle \sigma^2 \rangle_{\mathcal{D}}.$$

$\mathcal{Q}_{\mathcal{D}}$  consists of two parts: the variance of the expansion scalar  $\theta$  and the variance of the shear scalar  $\sigma$  (where the average of  $\sigma$  itself could be added to the backreaction term and accordingly added to the kinematical part of the equations, which then would feature Bianchi-type kinematics). Since both of them are second order terms, we can again calculate them with only the first order perturbed metric in Eq. (4). Detailed perturbative calculations can be found in Refs. [18,19], and here we only list the necessary results,

$$\begin{aligned} \theta &= 3\frac{\dot{a}}{a} - 3\dot{\Psi}, & \sigma^i_j &= \theta^i_j - \frac{1}{3}\theta\delta^i_j = \frac{1}{2}D^i_j\dot{\chi}, \\ \sigma^2 &= \frac{1}{2}\sigma^i_j\sigma^j_i = \frac{1}{8}D^i_j\dot{\chi}D^j_i\dot{\chi} \\ &= \frac{t_0^{8/3}}{2t^{2/3}} \left[ \partial^i\partial_j\varphi\partial^j\partial_i\varphi - \frac{1}{3}(\Delta\varphi)^2 \right]. \end{aligned}$$

(These results may be understood along the following lines: the Weyl curvature can be irreducibly split into the

electric part  $E_{\mu\nu}$  and the magnetic part  $H_{\mu\nu}$ . If only linear scalar perturbations are treated, the magnetic part vanishes, so the Weyl curvature squared  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  should be proportional to the electric part squared  $E_{\mu\nu}E^{\mu\nu}$ . Also in the linear approximation, the shear tensor is proportional to the electric part, and consequently  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  is proportional to the shear squared  $\sigma^2$ . This fact is easily seen by comparing  $C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho}$  in Eq. (13) and  $\sigma^2$  above.) Therefore, up to second order, we have

$$\mathcal{Q}_{\mathcal{D}} = \frac{t_0^{8/3}}{t^{2/3}} \left[ \langle \partial^i(\partial_i\varphi\Delta\varphi) \rangle - \langle \partial^i(\partial_j\varphi\partial^j\partial_i\varphi) \rangle - \frac{2}{3}\langle (\Delta\varphi)^2 \rangle \right]. \quad (14)$$

We also remark that due to the vanishing of the magnetic part, and due to the fact that in this perturbative approach the perturbations are propagating on the background, this result is formally equivalent to the corresponding Newtonian result [21].

Taking Eq. (14), we eventually find that the difference between the relative information entropy  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$  in Eq. (8) and the contraction of the Weyl tensor  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  in Eq. (13) is nothing but the kinematical backreaction term  $\mathcal{Q}_{\mathcal{D}}$ . These three quantities are correlated in the following relation (up to second order):

$$\frac{\mathcal{S}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{9}{32\pi G} \left( \frac{t^2}{8} \langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}} + \mathcal{Q}_{\mathcal{D}} \right). \quad (15)$$

We may get further insight from the result in Eq. (15), especially for the limit in which the volume of the averaging domain  $V_{\mathcal{D}}$  goes to zero. From Eqs. (13) and (14), we observe that  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  and  $\mathcal{Q}_{\mathcal{D}}$  are almost the same, except for the difference,  $\langle (\Delta\varphi)^2 \rangle - \langle \Delta\varphi \rangle^2$ , i.e., the variance of the energy density perturbation  $\rho^{(1)}$  [see Eq. (6)]. So in the limit  $V_{\mathcal{D}} \rightarrow 0$ , the averaging effects from inhomogeneities turn out to be negligible, and we should recover in a nice way the pointwise defined parameters for the unperturbed Universe. Therefore,  $\langle (\Delta\varphi)^2 \rangle$  and  $\langle \Delta\varphi \rangle^2$  cancel each other, and the relative information entropy  $\mathcal{S}_{\mathcal{D}}$  thus reduces to zero, as it should be for the coincidence of the actual distribution with the assumed one. The above argument provides a self-consistency check of the averaging framework [22].

The content of our relation, Eq. (15), may also be restated by using the relative entropy production density  $\dot{\mathcal{S}}_{\mathcal{D}}/V_{\mathcal{D}}$  of Eq. (10),

$$\frac{\dot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{3}{8\pi G} \left( \frac{t}{8} \langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}} + \frac{\mathcal{Q}_{\mathcal{D}}}{t} \right), \quad (16)$$

or by using  $\ddot{\mathcal{S}}_{\mathcal{D}}/V_{\mathcal{D}}$  of Eq. (11),

$$\frac{\ddot{\mathcal{S}}_{\mathcal{D}}}{V_{\mathcal{D}}} = \frac{1}{8\pi G} \left( \frac{1}{8} \langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}} + \frac{\mathcal{Q}_{\mathcal{D}}}{t^2} \right). \quad (17)$$

We should point out that the sign of  $\mathcal{Q}_{\mathcal{D}}$  is not explicit in perturbation theories, and is even more difficult to

determine in exact solutions. At last we should state that even though the results in Eqs. (15)–(17) are obtained from the perturbative calculations in the synchronous gauge, they are gauge invariant because all the three quantities  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$ ,  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$ , and  $\mathcal{Q}_{\mathcal{D}}$  are second order ones and vanish at both the zeroth (background) and first orders [23].

## VI. CONCLUSIONS

The study of inhomogeneous models in relativistic cosmology and the relevant averaging problems have attracted more and more attention in recent years [16,24–29], among which one of the central issues is to seek a simple and reasonable measure of the inhomogeneous distribution of the cosmic continuum. In this paper, we first discuss two such measures as an assessment of the degree of structure formation in the perturbed Universe: the relative information entropy  $\mathcal{S}_{\mathcal{D}}$  and the contraction of the Weyl tensor  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$ ; we calculate these two seemingly uncorrelated quantities in the standard perturbative approach up to second order. We find that both of them are related to the general averaging problem in the inhomogeneous Universe and show their relation via the kinematical backreaction term  $\mathcal{Q}_{\mathcal{D}}$ .

Equations (15)–(17) are the main results of our paper. From these relations, we prove that  $\mathcal{S}_{\mathcal{D}}/V_{\mathcal{D}}$  is proportional to  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$ , iff  $\mathcal{Q}_{\mathcal{D}}$  is negligible, i.e., iff the averaged Universe is described by an idealized FLRW cosmology. However, since kinematical backreaction measures

the deviation from a FLRW cosmology, an improved background including backreaction could eventually result in a direct proportionality if perturbations on the general average are considered, rather than on a FLRW background. The answer to this question is beyond the scope of this paper, but the formalism to address this more general issue has been outlined in a recent work [30].

Regarding the Weyl curvature hypothesis, we can state that it is not the Weyl curvature alone that is monotonically growing, but a specific combination with the kinematical backreaction, a result that has been proved here to second order in perturbation theory. In other words, the information entropy  $\mathcal{S}_{\mathcal{D}}$  consists of the gravitational entropy  $\langle C_{\mu\nu\lambda\rho}C^{\mu\nu\lambda\rho} \rangle_{\mathcal{D}}$  proposed by Penrose, and the kinematical backreaction  $\mathcal{Q}_{\mathcal{D}}$ , which contributes to the observed cosmic acceleration. It will be interesting to find out if this could be extended to the fully nonlinear theory.

## ACKNOWLEDGMENTS

We are very grateful to Yun-He Li, Bo-Qiang Ma, and Xin Zhang for fruitful discussions, and to Xavier Roy and Alexander Wiegand for valuable comments on the manuscript. The work of N.L. is supported by the National Natural Science Foundation of China (Grant No. 11105026). N.L. and D.J.S. thank the Deutsche Forschungsgemeinschaft (DFG) for financial support via Grant No. IRTG 881. T.B.'s work was conducted within the Lyon Institute of Origins under Grant No. ANR-10-LABX-66.

- 
- [1] X. Roy, T. Buchert, S. Carloni, and N. Obadia, *Classical Quantum Gravity* **28**, 165004 (2011).
  - [2] R. Penrose, in *General Relativity, an Einstein Centenary Survey*, edited by S.W. Hawking and W. Israel (Cambridge University Press, Cambridge, 1979), p. 581.
  - [3] R. Penrose, *Proceedings of EPAC 2006* (2006), p. 2759, <http://accelconf.web.cern.ch/accelconf/e06/>.
  - [4] F.C. Mena and P. Tod, *Classical Quantum Gravity* **24**, 1733 (2007).
  - [5] T. Buchert, *Gen. Relativ. Gravit.* **32**, 105 (2000).
  - [6] T. Buchert, *Gen. Relativ. Gravit.* **33**, 1381 (2001).
  - [7] A. Hosoya, T. Buchert, and M. Morita, *Phys. Rev. Lett.* **92**, 141 302 (2004).
  - [8] M. Gasperini, G. Marozzi, F. Nugier, and G. Veneziano, *J. Cosmol. Astropart. Phys.* **07** (2011) 008.
  - [9] I. Ben-Dayan, M. Gasperini, G. Marozzi, F. Nugier, and G. Veneziano, *J. Cosmol. Astropart. Phys.* **04** (2012) 036.
  - [10] S. Kullback and R.A. Leibler, *Ann. Math. Stat.* **22**, 79 (1951).
  - [11] M. Morita, T. Buchert, A. Hosoya, and N. Li, *AIP Conf. Proc.* **1241**, 1074 (2010).
  - [12] N. Akerblom and G. Cornelissen, *J. Math. Phys. (N.Y.)* **53**, 012502 (2012).
  - [13] M.P. Dąbrowski, J. Garecki, and D.B. Blaschke, *Ann. Phys. (Leipzig)* **18**, 13 (2009).
  - [14] G.F.R. Ellis, in *General Relativity and Gravitation*, edited by B. Bertotti, F. de Felice, and A. Pascolini (D. Reidel Publishing Company, Dordrecht, 1984), p. 215.
  - [15] G.F.R. Ellis and T. Buchert, *Phys. Lett. A* **347**, 38 (2005).
  - [16] T. Buchert, *Gen. Relativ. Gravit.* **40**, 467 (2008).
  - [17] T. Buchert and S. Räsänen, *Annu. Rev. Nucl. Part. Sci.* **62**, 57 (2012).
  - [18] N. Li and D.J. Schwarz, *Phys. Rev. D* **76**, 083011 (2007).
  - [19] N. Li and D.J. Schwarz, *Phys. Rev. D* **78**, 083531 (2008).
  - [20] C. Clarkson, K. Ananda, and J. Larena, *Phys. Rev. D* **80**, 083525 (2009).
  - [21] J. Ehlers and T. Buchert, *Gen. Relativ. Gravit.* **41**, 2153 (2009).
  - [22] A. Wiegand and D.J. Schwarz, *Astron. Astrophys.* **538**, A147 (2012).
  - [23] J.M. Stewart and M. Walker, *Proc. R. Soc. A* **341**, 49 (1974).

- [24] G. F. R. Ellis, [Classical Quantum Gravity](#) **28**, 164001 (2011).
- [25] C. Clarkson, G. F. R. Ellis, J. Larena, and O. Umeh, [Rep. Prog. Phys.](#) **74**, 112901 (2011).
- [26] E. W. Kolb, [Classical Quantum Gravity](#) **28**, 164009 (2011).
- [27] S. Räsänen, [Classical Quantum Gravity](#) **28**, 164008 (2011).
- [28] D. L. Wiltshire, [Classical Quantum Gravity](#) **28**, 164006 (2011).
- [29] T. Buchert, [Classical Quantum Gravity](#) **28**, 164007 (2011).
- [30] X. Roy and T. Buchert, [Classical Quantum Gravity](#) **29**, 115004 (2012).