Nonequilibrium thermodynamics of spacetime: The role of gravitational dissipation

G. Chirco[*] and S. Liberati[†]

SISSA, Via Beirut 2, 34151 Trieste, Italy and INFN sezione di Trieste

In [T. Jacobson, Phys. Rev. Lett. **75**, 1260 (1995).] it was shown that the Einstein equation can be derived as a local constitutive equation for an equilibrium spacetime thermodynamics. More recently, in the attempt to extend the same approach to the case of $f(R)$ theories of gravity, it was found that a nonequilibrium setting is indeed required in order to fully describe both this theory as well as classical general relativity (GR) [C. Eling, R. Guedens, and T. Jacobson, Phys. Rev. Lett. **96**, 121301 (2006).]. Here, elaborating on this point, we show that the dissipative character leading to nonequilibrium spacetime thermodynamics is actually related—both in GR as well as in $f(R)$ gravity—to nonlocal heat fluxes associated with the purely gravitational/internal degrees of freedom of the theory. In particular, in the case of GR we show that the internal entropy production term is identical to the so-called tidal heating term of Hartle-Hawking. Similarly, for the case of $f(R)$ gravity, we show that dissipative effects can be associated with the generalization of this term plus a scalar contribution whose presence is clearly justified within the scalar-tensor representation of the theory. Finally, we show that the allowed gravitational degrees of freedom can be fixed by the kinematics of the local spacetime causal structure, through the specific equivalence principle formulation. In this sense, the thermodynamical description seems to go beyond Einstein's theory as an intrinsic property of gravitation.

I. INTRODUCTION

Over a decade ago, the connection between gravity, thermodynamics, and quantum field theory, developed in 70's for black holes physics [1–3], was strongly tightened up by Jacobson [4], who was able to derive the Einstein equations as equilibrium constitutive equations for spacetime, starting from the thermodynamical properties of local causal horizons and the thermal nature of the Minkowski vacuum.

Since then, this result has given great support to the idea that the well-established black holes' thermodynamics should be in fact extendable to some more general spacetime thermodynamics, where the fundamental ingredients are a mix between causal horizons' stationarity, quantum fields' thermal behavior, and peculiar holographic properties of gravity [5–9].

However, more recently, it was realized that the thermodynamical derivation of the Einstein equations needs a generalization to a nonequilibrium thermodynamical setting [10,11]. This was first noticed in the attempt to reproduce the derivation *à la* Jacobson for a $f(R)$ gravity theory. In that case, making the horizon entropy proportional to a function of the Ricci scalar led to a break down of the local thermodynamical equilibrium.

In order to recover the $f(R)$ gravity field equations from the thermodynamical prescription, it was then necessary to modify the equilibrium entropy balance of the system by

considering some extra entropy production terms. More surprisingly, the same problem was in fact pointed out, in the same works, even for general relativity (GR), the main issue being again substantially related to the definition of entropy for the system. Nevertheless, a clear physical interpretation of the extra entropy production terms which come into play for the nonequilibrium thermodynamics derivation is still missing.

In this work, following [10], we adopt the nonequilibrium description as the general proper setting for the above approach. In such a nonequilibrium setting, the Einstein equations (or their generalization for higher order gravity) can be still derived from a local thermodynamical condition, as a constitutive equation, but only as far as one separates the reversible and irreversible thermodynamical contributions (with the reversible sector being associated with the Einstein field equations). In this context, the general form for the irreversible/viscous contributions and the constitutive equations for their coefficients is identified by exploiting a fluid dynamics description of the local causal horizon kinematics. We then argue that such terms, already interpreted in [10] as dissipative effects, should be identified as heat terms associated with the purely gravitational/internal degrees of freedom of the theory. In particular, for the case of GR we find that the internal entropy production term is exactly the well-known tidal heating term associated with dissipation of black hole horizons' perturbations via gravitational fluxes [12–16]. Furthermore, by applying the same approach to the $f(R)$ theories of gravity, we actually find a generalization of the tidal heating effect of GR together with an extra purely

*chirco@sissa.it
†liberati@sissa.it

scalar dissipative contribution (as expected in this class of theories given their possible representation as scalar-tensor ones [17]).

We start, in Sec. II, by reviewing and analyzing the operative framework for the spacetime thermodynamics' approach. In particular, we further elaborate on the role of the equivalence principle (EP) in selecting gravitational theories in such a framework. We then proceed in Sec. III to the derivation of the Einstein equation of state as in [4,10,11] and, in Sec. IV, we give some remarks on the open issues left in the argument. In Sec. V, we describe the analogy between a general fluid and the horizon congruence kinematics. This allows us to interpret (in a classical irreversible thermodynamics setting) the nonequilibrium features of the congruence as the analogues of the viscous dissipative terms for the fluid. In Secs. VI and VII, we rederive the equations of motion for GR and $f(R)$ gravity, respectively, by starting from a general nonequilibrium entropy balance and looking at its reversible sector. We show that the irreversible contributions can in both cases be associated with the purely gravitational dissipative effects of the theory. In Sec. VIII, we shall finally speculate on the possible lessons one can gather from our results and their hints towards an identification of the microscopic degrees of freedom of gravity.

II. SPACETIME THERMODYNAMICS: THE FRAMEWORK

The idea of spacetime thermodynamics, as first presented in [4], is built on the relation between the thermal character of quantum fields vacuum, as perceived by an accelerated Rindler observer, and the stationarity properties of the respective Killing horizon. By making use of the equivalence principle, the notion of the Rindler frame can be used at a local level as an experimental setting for studying the local spacetime dynamics as the way the variation of the spacetime geometry follows from the energy variation of the matter fields. Here we review and extend the framework of the previous investigations [4,10,11] stressing its crucial points and its implicit assumptions.

A. The local Rindler wedge

The first step in the construction of a spacetime thermal system is the introduction of a local notion of horizon. In analogy with the global definition of a horizon as the boundary of the past of future null infinity, one can generally consider a local horizon at p, in a generic spacetime $(\mathcal{M}, g_{ab})$, as one side of the boundary of the past of a spacelike 2-surface patch B including p. Thereby, near p, the local horizon will be constituted by the congruence of null geodesics orthogonal to B, characterized by the past pointing tangent null vector ℓ^a (see Fig. 1).

With respect to the point p, one can then invoke local Lorentz invariance of spacetime (assumed in both the
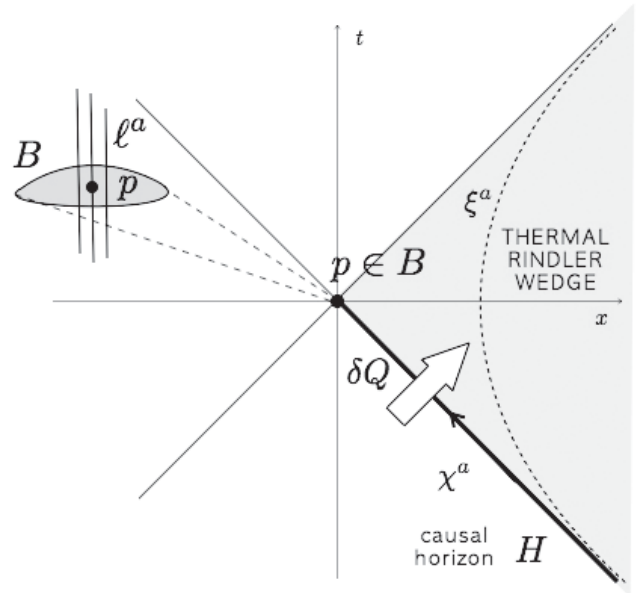

FIG. 1. A schematic view of the thermal Rindler wedge. The local causal horizon H plays the role of a diathermic barrier for the thermal Rindler wedge. The thermal system is constituted by the set of Rindler observers ξ^a moving along their isometry trajectories. For each of them the ensemble is described by $\rho = Z^{-1} \exp(-H_b/T)$ and perturbed by the energy δQ that flows across H.

Einstein and Strong formulations of the equivalence principle [18]) to introduce a local inertial frame. This is always allowed, provided one restricts oneself to a region of size $\ell^2 \ll \mathcal{R}(p)^{-1}$, where $\mathcal{R}(p)$ gives the value of the smaller scale associated with the radius of curvature at p (which will be generically nonzero).

Within this region the metric will be approximately Minkowski, that is

$$g_{ab} = \eta_{ab} + \mathcal{O}(\epsilon^2), \tag{1}$$

the order of approximation being fixed by the local curvature. Finally, the local patch can be described via Riemann normal coordinates $\{x^a\}$, such that p stays at $x^a = 0$. On the introduced local inertial frame, one can construct a local Rindler frame (LRF) by the usual coordinate transformations.[1] It is convenient here to choose $x = \chi \cosh(\eta\kappa)$ and $t = \chi \sinh(\eta\kappa)$, introducing an arbitrary rescaling factor κ for the proper time, in order to have a clear label for the Rindler wedge temperature in the following derivation.

With this choice, in the neighborhood of p, the LRF metric will be

$$ds^2 = -\kappa^2\chi^2 d\eta^2 + d\chi^2 + dy^2 + dz^2, \tag{2}$$

---

[1]Note that the construction of the LRF requires one to fully use the equivalence principle, at least in its Einstein formulation, in order to identify geodesics' motion [19].

corresponding to the action of a Lorentz boost, with acceleration $a = 1/\chi$.

Therefore, within a small neighborhood of p, one can associate the boundary of the past of the patch B to a section of the approximate Killing horizon, centered in p. The future pointing approximate boost Killing vector χ^a, vanishing in p by definition, will be tangent to the null congruence comprising the causal horizon and will leave invariant the tangent plane to B at p.

In terms of the approximate Killing vector χ^a, one can introduce a time[2] label v along the horizon null hypersurface, defined by $\chi^a \nabla_a v = 1$. The Killing parameter v can be expressed in terms of the null congruence affine parameter λ. For a Killing horizon, the relation is generally given by

$$\lambda = -e^{-\kappa v}, \qquad (3)$$

so that the point p is located at an infinite Killing parameter and at $\lambda = 0$.

As a consequence, one gets $\chi^a = (d\lambda/dv)\ell^a$, with $(d\lambda/dv) = -\kappa\lambda$ and, in the same way,

$$\hat{\theta} = \left(\frac{d\lambda}{dv}\right)\theta = -\kappa\lambda\theta \quad \text{and} \quad \hat{\sigma} = \left(\frac{d\lambda}{dv}\right)\sigma = -\kappa\lambda\sigma, \qquad (4)$$

which gives some helpful relations between the Killing expansion $\hat{\theta}$ and shear $\hat{\sigma}$ and the respective affine geodesics' quantities.

B. Rindler wedge thermal character

Together with local flatness, the previous construction actually needs to further assume that the ground state of the fields living in the original spacetime is locally approximated by the Minkowski vacuum. In this case, with respect to the approximate Killing vector flow, associated with the set of observers living in the wedge at $x > |t|$ (Rindler wedge), the vacuum state can be interpreted as an approximate thermal state, with a temperature

$$T \approx T_{\mathrm{un}}\sqrt{-g_{00}} = \frac{\hbar\kappa}{2\pi}, \qquad (5)$$

where $T_{\mathrm{un}} = \hbar a/2\pi$ is the Unruh temperature [20] for the Rindler observer.

The expression in (5) shows that T stays constant throughout the Rindler wedge, because of the gravitational Doppler factor χ associated with the Unruh temperature, and it is well defined on the horizon. Therefore, the thermal character of the Rindler state is effectively extended from the single Rindler observer to the whole wedge.

As a further step, one can introduce an entropy for the system, via an entanglement argument. In the Rindler wedge, an accelerated observer can only access information on spacelike slices bounded by the bifurcation plane. Thereby, since vacuum fluctuation between the inside and the outside of the wedge are correlated, she/he will perceive an entanglement entropy, which scales with the area of the local boundary and diverges with the density of field states in the UV limit. However, with the introduction of an UV cutoff (generically justified via the Planck scale quantum vibrations, the so called horizon *Zitterbewegung*) one can make this entropy become actually proportional to the area, that is

$$S = \alpha A, \qquad (6)$$

where the proportionality factor α can *a priori* depend on the nature of the quantum fields as well as be some complicate function of the position in spacetime [21,22]. Together with the temperature T, this notion of entropy allows one to consider the local Rindler wedge with its Killing horizon as an analogue of a canonical ensemble (Gibbs state) bounded by a diathermic wall.

C. Local equilibrium condition

Since all of the thermal information related to the Rindler wedge is somehow recorded on the horizon boundary [11], one can define a notion of thermal equilibrium for the system in terms of the stationarity properties of the horizon.[3]

In this sense, the equilibrium state for the system is identified with the horizon cross section at p, where the Killing expansion $\hat{\theta}$ and shear $\hat{\sigma}$ vanish and the horizon is instantaneously stationary with respect to λ,

$$dS = \delta(\alpha A) = 0. \qquad (7)$$

Given the relations in (4), it is clear that the equilibrium state can be reached as long as the affine quantities (θ, σ) are not diverging in p. The local curvature of B at p determines the geometrical properties of the horizon null congruence and consequently the value of affine expansion and shear at p. Hence, in order to define an equilibrium surface one just needs to require a suitably smooth curvature for B at p, without fixing *a priori* the values of affine expansion and shear.

This is a very delicate point, as we will see that the properties of B at p, and the corresponding values of the optical scalars of the associated null congruence, actually select the theories of gravity which may arise from the thermodynamical approach by fixing the gravitational de-

---

[2]Notice that there is no relation between the proper time η defined in the wedge and the Killing parameter v along the horizon. Nevertheless we need to keep the same scaling κ for dimensional consistence.

[3]It can be proved formally that the stability of the Rindler metric (Lorentz invariance), hence the Rindler horizon stationarity, actually implies the Kubo-Martin-Schwinger conditions [23], responsible for the thermal character of the vacuum energy fluctuations as measured by a Rindler observer.

grees of freedom of the theory. In this sense, we will show that the choice of B can be directly related to the equivalence principle formulation, which plays a fundamental role in the argument.

III. REVIEW OF THE PREVIOUS RESULTS: THE GR CASE

Moving away from the bifurcation surface at p, the spacetime will become dynamical and the presence of matter will distort the local Rindler causal structure, perturbing the thermal state at the same time.

As long as the departure from equilibrium is small and slow, it can be described in analogy with a quasistatic process where a suitably small amount of energy is thrown through the horizon. Then, one needs some condition to relate the spacetime geometry deformation to the variation of the fields' energy content. In fact, thanks to the thermal properties of the system, this condition can actually be borrowed from classical equilibrium thermodynamics.

Indeed, for a slightly perturbed Gibbs state at temperature T, variations of entropy and internal energy are related by the Clausius relation,

$$dS = \frac{\delta Q}{T}, \qquad (8)$$

where the change in the mean energy is taken to be due to the fluxes into/from an unobservable region of spacetime, which is perfectly thermalized by the horizon system, and hence it is assumed to consist entirely of heat. For a system undergoing a quasistatic process of energy exchange with the surroundings, relation (8) is nothing but a local thermodynamical equilibrium condition.

Now, since the thermal Rindler wedge behaves like a Gibbs state, the first fundamental assumption in [4] consists in using relation (8) to relate the horizon entropy and the boost energy across the Rindler horizon, with T given by Eq. (5).

As a second fundamental postulate, it is assumed a universal entropy density α per unit horizon area δA, such that

$$dS = \alpha \delta A, \qquad (9)$$

thereby implicitly considering a constant UV cutoff for the fields, whereas in general one would have $dS = \delta(\alpha A)$ (i.e., α can be some spacetime function). With this choice, the changes of the entanglement entropy of the fields in the wedge can be effectively described in terms of geometrical variations of the horizon cross section.

Let us highlight here that the $\alpha = $ constant assumption, made in the GR derivation of [4,10], can be indeed recast as an explicit choice of a specific formulation of the equivalence principle. As said, generally the UV cutoff α is fixed at the quantum gravity scale. This can be identified as the scale at which the gravitational action is of the order of the quantum of action $\hbar$. For GR this is the standard Planck

length $l_p = \sqrt{G\hbar/c^3}$ and hence it is directly related to the Newton constant. However, for a general scalar-tensor theory, i.e., a theory compatible just with the Einstein equivalence principle (EEP) [18], G is promoted to a spacetime field. As a consequence of this, one should expect that the cutoff will be generically position dependent. In this sense, assuming $dS = \alpha \delta A$ is equivalent to assume the strong formulation of the equivalence principle (SEP) [18], hence effectively to allow for the only two SEP-compatible gravity theories: GR and Nordström gravity [24].

Given the assumption in (9), a quantitative expression for the system entropy variation is obtained just by applying the definition for the change of the horizon area in terms of the expansion rate of the null geodesics comprising it, that is

$$\delta A = \int_H \tilde{\epsilon}\theta d\lambda, \qquad (10)$$

with $\tilde{\epsilon}$ indicating the 2-surface area element of the horizon cross section. Moving away from the equilibrium surface at $\lambda = 0$, along the null congruence, the infinitesimal evolution of θ is given by a linear expansion around its equilibrium value at p, up to the first order in λ,

$$\theta \approx \theta_p + \lambda \frac{d\theta}{d\lambda}\bigg|_p + \mathcal{O}(\lambda^2). \qquad (11)$$

This first order coefficient will be determined as usual by the Raychaudhuri equation,

$$\frac{d\theta}{d\lambda} = -\frac{1}{2}\theta^2 - \parallel \sigma \parallel^2 - R_{ab}\ell^a\ell^b, \qquad (12)$$

where $\parallel \sigma \parallel^2$ stands for the squared congruence shear $\sigma^{ab}\sigma_{ab}$.[4]

In this way, the entropy variation, up to $\mathcal{O}(\lambda^2)$, is given by

$$dS = \alpha \int_H \tilde{\epsilon}d\lambda\left[\theta - \lambda\left(\frac{1}{2}\theta^2 + \|\sigma\|^2 + R_{ab}\ell^a\ell^b\right)\right]_p. \qquad (13)$$

The mean energy variation of the thermal system, given by the boosted energy current flux of matter, is then described by the heat flux across the horizon as

$$\delta Q = \int_H T_{ab}\chi^a d\Sigma^b, \qquad (14)$$

where T_{ab} is the matter stress-energy tensor, while the volume element is given by $d\Sigma^b = \tilde{\epsilon}d\lambda\ell^b$. The same quantity, with respect to the null congruence parameters, reads

---

[4]Here we consider a vanishing twist, as the null congruence is taken hypersurface orthogonal.

$$\delta Q = \int_H \tilde{\boldsymbol{\epsilon}} d\lambda (-\lambda\kappa) T_{ab} \ell^a \ell^b. \tag{15}$$

At this point, asking for relation (8) to hold for all null vectors ℓ^a, one can equate the $\mathcal{O}(\lambda)$ integrands in (13) and (15). At the zeroth order in λ, the value of heat flux at p is zero, hence one necessarily gets $\theta_p = 0$. Then, to the first order,

$$\frac{2\pi}{\hbar\alpha} T_{ab} \ell^a \ell^b = (\|\sigma\|^2 + R_{ab} \ell^a \ell^b)_p. \tag{16}$$

Moreover, if one further requires that $\sigma_p = 0$, one is left with

$$\frac{2\pi}{\hbar\alpha} T_{ab} = R_{ab} + \Phi g_{ab}, \tag{17}$$

where Φ is an undetermined integration function.

Eventually, by assuming the local energy conservation, that is $\nabla^b T_{ab} = 0$, applying the divergence operator on both sides of (17), and using the contracted Bianchi identity $\nabla^b R_{ab} = \frac{1}{2}\nabla_a R$, one finally gets $\Phi = -\frac{1}{2}R - \Lambda$, hence

$$\frac{2\pi}{\hbar\alpha} T_{ab} = R_{ab} - \frac{1}{2}R g_{ab} - \Lambda g_{ab}, \tag{18}$$

where Λ is some arbitrary integration constant. Once the condition

$$\alpha = \frac{1}{4\hbar G} \tag{19}$$

is imposed, one can easily recognize the familiar Einstein equations. Noticeably, Eq. (19) implies that the entropy density of the local Rindler horizon is the same as the one of a black hole.

This result tells us that, given the entropy and energy conservation for the system, the local thermodynamical equilibrium condition is, in fact, equivalent to the Einstein equation for the local thermal spacetime. Furthermore, the EP implies that the above construction can be done at *any* spacetime point p, and hence that Eq. (18) holds everywhere in spacetime.

On the other hand, by allowing for some shear in p, that is for $\sigma_p \neq 0$, the Einstein equation is no more recovered. As discussed in the previous section, the LRF is not sensitive to the exact value of the affine expansion and shear at p. Therefore, setting $\sigma_p = 0$ is an unjustified arbitrary choice. This was first realized in [10]. In fact, a nonzero affine shear at p may change the way in which the equilibrium is approached by the system in the Killing frame. Given the relations in (4) and the analogue of the expansion (11) for the affine shear, one realizes that the Killing shear falls off to zero at p as $\hat{\sigma} \sim e^{-2\kappa v}$ when σ vanishes, while only as $\hat{\sigma} \sim e^{-\kappa v}$ when σ is nonvanishing. In this sense, for a nonvanishing affine shear, the equilibrium approach can be considered slow enough for the system to be in a nonequilibrium regime.

This argument was used in [10] to recast the thermodynamical derivation in a nonequilibrium setting, where

$$dS > \delta Q/T. \tag{20}$$

In this new context, the Clausius law is replaced by the entropy balance law,

$$dS = \delta Q/T + d_i S, \tag{21}$$

and the extra shear term in (16) is then associated with the internal entropy production $d_i S$, generated by the system out of equilibrium. The internal entropy contribution, $\mathcal{O}(\lambda)$, has the form

$$d_i S = -\frac{4\pi\eta}{\hbar} \int_H \tilde{\boldsymbol{\epsilon}} \lambda d\lambda \|\sigma\|_p^2 \tag{22}$$

and, in analogy to the internal entropy production terms originating from the squared gradients of state variables, a universal property of systems with viscosity in nonequilibrium thermodynamics [25], it is interpreted as an internal entropy production term due to some internal spacetime viscosity, with $\eta = \hbar\alpha/4\pi$. Furthermore, it is also noticed in [10] that by using (19), one gets $\eta = 1/(16\pi G)$ in agreement with the value obtained for the shear viscosity of the stretched horizon of a black hole in the so-called membrane paradigm [26,27]. This result concludes the review of the argument described in [4,10,11].

IV. SOME REMARKS

One can consider the local spacetime thermodynamics developed in [4,10,11] as a useful experimental setting, where the local behavior of gravity is tested in its different formulations. Nevertheless, we believe that the argument still misses a clear understanding.

First of all, it seems that the thermodynamical approach cannot have a detailed control over the degrees of freedom of the resulting gravity theory, in particular, on the way they are effectively involved in the perturbation of the spacetime causal structure. Indeed, by allowing for some shear at p, the local equilibrium condition is formally split in two parts: the ℓ^a part, which leads to the Einstein equation, and the $\partial \ell^a$ part, related to the shear term, which remains unexplained. In this sense, at the $\partial \ell^a$ level, the local equilibrium condition is broken.

Therefore, equilibrium thermodynamics can give a suitable description only under the assumption that the affine congruence orthogonal to B has zero expansion and shear at p. However, we saw that this is equivalent to require that the chosen B (and hence its associated null congruence) is less general than the one allowed by the assumed entropy-area relation (or alternatively by the SEP). Such an ansatz seems too restrictive to consider (18) as a general result.

On the one hand, allowing for nonzero affine shear at p gives rise to further interesting clues. In some way, the shear contribution in (16) brings into the entropy balance process a new degree of freedom, which is not fixed by the

Ricci tensor and so has nothing to do with the local matter energy sources. Actually, the surface shear is generally related to the Weyl tensor and usually associated with the distortion on the geodesics congruence due to a gravitational perturbation.

In fact, this argument opens an issue about the absence so far of any role for gravitational fluxes in the system energy perturbation mechanism. Because of their nonlocal nature, the gravitational energy fluxes cannot be taken into account with a proper stress-energy tensor. However, allowing for nonlocal terms, as the one in $\| \sigma \|^2$ in (16), seems at odds with neglecting the role of these nonlocal energy contributions.

On the other hand, the interpretation of the internal entropy contribution as a by-product of some sort of viscous work on the system, given in [10,11], is very reasonable, because the term (22) is actually related to some mechanical deformation due to the presence of shear in the null congruence generating the horizon.

In this sense, such a spacetime viscosity seems naturally related to the distorsive effect of a gravitational flux, to be intended as a local curvature perturbation which is independent from the Einstein equation. This suggests that gravitational energy fluxes can possibly play a role into the total entropy balance of the system without entering into the Einstein equilibrium relation.

Starting from these remarks, we are led to reconsider the thermodynamical argument of this section in a fully non-equilibrium setting further elaborating on the approach taken in [10]. In particular, for the motivations given above, we argue that the internal entropy production, such as (22), actually indicates the presence of dissipative (irreversible) processes, to be related to the conformal components of the spacetime curvature, which do not take part into the field equations and are associated with purely gravitational degrees of freedom.

V. INTERNAL ENTROPY IN NONEQUILIBRIUM THERMODYNAMICS

In classical nonequilibrium thermodynamics, the rate of change of the entropy is generally written as the sum of two contributions:

$$dS = d_e S + d_i S, \qquad (23)$$

where $d_e S$ is the rate of entropy exchange with the surroundings while $d_i S$ comes from the process occurring inside the system and is a non-negative quantity, accordingly to the second law of thermodynamics. In particular, $d_i S$ is zero for reversible (quasistatic) processes and positive for irreversible processes.

The Clausius relation used for the equilibrium approach in Sec. III is actually equivalent to the Clausius definition of entropy for the equilibrium system, that is

$$d_e S = \delta Q / T \quad \text{and} \quad d_i S = 0, \qquad (24)$$

as, in that case, the horizon perturbation is effectively described as a quasistatic process occurring in continuous equilibrium with the surrounding. However, this definition does not hold true any more as irreversible processes come into play.

Actually, in the nonequilibrium thermodynamical setting, the Clausius definition of entropy is generalized to the expression

$$dS = \frac{\delta Q}{T} + \delta N, \qquad (25)$$

where δQ is classically referred to as *compensated heat*, that is the heat transferred between the system and its surroundings, while δN, the so-called *uncompensated heat*, indicates the amount of entropy associated with the heat which is intrinsic to the system when it undergoes an irreversible process.

Let us stress that the above definition is very general, as it does not require either an *a priori* specification of the nature of the nonequilibrium variable, or the nature of the process involved. It generalizes the notion of local equilibrium by extending the entropy balance to the unbalanced contributions related to the irreversible processes, like dissipation (see, e.g., [28]).

The generalized Clausius relation (25) is helpful in order to clarify the nature of the equilibrium and nonequilibrium contributions defining our system entropy. In fact, by using the definition of the nonequilibrium entropy given in (23), we can write

$$d_e S + d_i S = \frac{\delta Q}{T} + \delta N, \qquad (26)$$

and identify the external and internal entropy in terms of the compensated and uncompensated heat, respectively,

  (i) $d_e S = \delta Q / T$, at the reversible level,

  (ii) $d_i S = \delta N$, at the irreversible level.

With this approach, the argument described in III acquires a straightforward interpretation. Indeed, the extra contribution (22) introduced by the nonvanishing horizon shear is an internal entropy production term allowed by the most general choice of the null congruence associated with B compatible with the area-entropy relation for GR (that we linked to the choice of the EP formulation). Therefore, it has to be seen as a by-product of the presence of internal/ purely gravitational degrees of freedom of the theory which can be responsible for irreversible dissipative processes.

However, in order to physically identify an internal entropy contribution $d_i S$ into the general expression for the horizon entropy given in (13), one needs a clear understanding of the relation between nonequilibrium forces and intrinsic spacetime properties involved.

Since all the thermal information of the Rindler wedge vacuum is recorded on the horizon boundary [11], the internal spacetime variables involved in the nonequilibrium process should be related to the null geodesic con-

gruence kinematics around p. In this sense, a possible way to capture nonequilibrium features of the thermal system is to use the analogy between the congruence bundle comprising the horizon and a classical fluid.

A. Null congruence as a continuous medium

Given the Price and Thorne *membrane* approach [26,27], the local Rindler horizon can be effectively approximated by a timelike hypersurface living just inside the true Rindler wedge, i.e., a timelike *stretched horizon*.

With a $2 + 1$ decomposition, the timelike congruence comprising the stretched horizon is formally equivalent to a $2 + 1$ continuous medium (fluid) living on the spacelike two dimensional cross section of the hypersurface and moving with velocity $\mathbf{v}^i$, defined by the unit timelike vector tangent to the hypersurface [11,26].

As the stretched horizon tends to the Rindler horizon, the unit timelike vector $\mathbf{v}^i$ approximates the Killing vector χ^a. Therefore, from a purely kinematical point of view, we can consider the velocity gradient of the medium, $\nabla_j \mathbf{v}^i$, as the equivalent of the deformation tensor field of our congruence (extrinsic curvature), in the Killing frame,

$$\hat{B}_{ab} = h_a{}^c h_b{}^e \chi_{c;e}, \qquad (27)$$

where the tensor h_{ab} is the projector onto the spacelike cross section transverse to the congruence (transverse metric), given by

$$h_{ab} = g_{ab} + \chi_a k_b + \chi_b k_a \qquad (28)$$

while χ^a is the Killing vector flow, as defined in Sec. II and k is an auxiliary null vector field satisfying $k^a \chi_a = -1$ which spans, together with χ^a the 2-surface horizon cross section. In both cases, the tensor can be decomposed into trace and symmetric trace-free parts,

$$\mathbf{V}_{ij} = \tfrac{1}{3}(\nabla^\rho \mathbf{v}_\rho)\delta_{ij} + \tilde{V}_{ij}, \qquad (29)$$

where $\mathbf{V}_{ij}$ represents the symmetric component of $\nabla_j \mathbf{v}_i$, δ_{ij} is the identity tensor, and

$$\tilde{V}_{ij} = \tfrac{1}{2}(\nabla_i \mathbf{v}_j + \nabla_j \mathbf{v}_i) \qquad (30)$$

is the deviatoric traceless tensor. Similarly

$$\hat{B}_{ab} = \tfrac{1}{2}\hat{\theta} h_{ab} + \hat{\sigma}_{ab}, \qquad (31)$$

where $\hat{\theta}$ and $\hat{\sigma}_{ab}$ are the usual scalar expansion and shear of the null congruence.

Now, for the continuous medium, the presence of $\mathbf{V}_{ik}$ indicates a relative motion between various parts of the fluid and it is responsible for the irreversible viscous transfer of momentum into the medium and for the consequent production of internal entropy. In particular, for the viscous

medium, the mechanical dissipation will contribute to the internal entropy density $d_i s$ by an amount[5]

$$d_i s = -\frac{1}{T}\mathbf{P}^{\text{vis}}_{ij}\mathbf{V}^{ij} = -\frac{1}{T}(\mathbf{p}^{\text{vis}}\nabla^\rho \mathbf{v}_\rho + \tilde{\mathbf{P}}^{\text{vis}}_{ij}\tilde{\mathbf{V}}^{ij}), \quad (32)$$

where $\mathbf{P}^{\text{vis}}_{ij} = \mathbf{p}^{\text{vis}}\delta_{ij} + \tilde{\mathbf{P}}^{\text{vis}}_{ij}$ is the viscous pressure acting on the medium, once decomposed, respectively, in its bulk and traceless components [28].

B. Constitutive equations

The expression for the internal entropy in (32) is a bilinear form, where the external stress $\mathbf{P}^{\text{vis}}$ comes coupled with its conjugate strain $\mathbf{V}$. Typically, $\mathbf{V}$ is a known function of the internal state variable of the medium, while the external forces are unknown. However, in classical irreversible thermodynamics, under a *local equilibrium hypothesis*, that is, as far as it is possible to assume that local and instantaneous relations between the thermal and mechanical properties of the system are the same as for uniform systems at equilibrium, the deformation of the velocity field and the viscous pressure (mechanical stress) can be related by the linear constitutive equations,

$$\mathbf{p}^{\text{vis}} = -\zeta \nabla^\rho \mathbf{v}_\rho, \qquad \text{Stoke's law} \qquad (33)$$

$$\tilde{\mathbf{P}}^{\text{vis}}_{ij} = -2\eta \tilde{\mathbf{V}}_{ij}, \qquad \text{Newton's law} \qquad (34)$$

where ζ and η, respectively, indicate the bulk and shear viscosity of the medium. The dissipation coefficients are exactly related to the time dependence of equilibrium fluctuations for $\mathbf{p}^{\text{vis}}$ and $\tilde{\mathbf{P}}^{\text{vis}}_{ij}$, in the framework of the linear response theory, via the Green-Kubo relations [28].

Thereby, given the linear relation introduced above, the viscous internal entropy density can be defined in terms of velocity gradient components of the medium, that is

$$d_i s = \frac{\zeta}{T}(\nabla^\rho \mathbf{v}_\rho)^2 + \frac{2\eta}{T}\|\tilde{\mathbf{V}}\|^2, \qquad (35)$$

where $\|\tilde{\mathbf{V}}\|^2 = \tilde{\mathbf{V}}_{ij}\tilde{\mathbf{V}}^{ij}$.

Now, given the stretched horizon approximation and by associating the bulk term $(\nabla^\rho \mathbf{v}_\rho)$ to $\hat{\theta}$ and the deviatoric traceless tensor $\tilde{\mathbf{V}}_{ij}$ to $\hat{\sigma}_{ab}$, one can expect a dissipative internal entropy production term for the horizon congruence, of the form

$$(d_i S)^{\text{vis}} = \frac{1}{T}\int_H \tilde{\boldsymbol{\epsilon}}\,dv\,\zeta\hat{\theta}^2 + 2\eta\|\hat{\sigma}\|^2. \qquad (36)$$

The above expression provides two important insights. First, it identifies the congruence kinematical quantities which are responsible for the irreversible viscous transfer of momentum into the system, and for the consequent

---

[5]Among the several dissipative processes which can contribute to the internal entropy production for the medium, here we just consider the viscous stresses.

production of internal entropy. In our setting, the favorite candidates to play the role of conjugate viscous stresses are the spacetime curvature perturbations. Indeed, the evolution of the deformation tensor $\hat{B}_{ab}$ along the approximated Killing flow is driven by the equation

$$\chi^c \nabla_c \hat{B}_{ab} = \kappa \hat{B}_{ab} - \hat{B}_a^c \hat{B}_{cb} - h_a{}^c h_b{}^{cd} R_{cedf} \chi^e \chi^f, \quad (37)$$

whose trace and trace-free part give, respectively, the Raychaudhuri and shear evolution (tidal force) equations along the horizon [29]. In vacuum, the only source term for both the equations is given by the electric part of the Weyl tensor, $C_{ab} = h_a{}^{cc} h_b{}^{ce} C_{cdef} \chi^d \chi^f$ [27,29] which is well known to be associated with gravitational perturbations.

Second, the strict correspondence between Eqs. (35) and (36) seems to suggest the presence of linear constitutive equations, analogues of (33) and (34), between the horizon kinematical quantities $\hat{\theta}$ and $\hat{\sigma}_{ab}$ and some external viscous stress. Therefore, by extending the fluid analogy argument, one should be able to relate the congruence viscosity coefficients to the local curvature fluctuations via some analogue of the Green-Kubo relations. We leave this for future investigations.

VI. NONEQUILIBRIUM SPACETIME THERMODYNAMICS: GR

We now have a clear way to relate the nonequilibrium features, arising in the thermodynamical derivation of the Einstein equations in Sec. III, to those kinematical degrees of freedom of the horizon congruence which are turned on by the local spacetime curvature. The horizon kinematics actually define the intrinsic spacetime properties involved in the irreversible processes.

We can then reproduce the thermodynamical derivation of the Einstein equations with a nonequilibrium irreversible thermodynamics setting, simply by starting, in a quite general way, from a generic spacelike 2-surface patch B at p, with nonvanishing θ_p and σ_p.

As for the previous derivation, the strong equivalence principle allows us to use the entropy-area relation as in (9), that is

$$dS = \alpha \delta A \quad \text{with} \quad \alpha = \text{const.} \quad (38)$$

However, in the new setting, the Clausius relation will be generalized to the expression (25), where δN will now encode all the information about both microscopic properties and irreversible perturbations of the system.

The new argument starts from the same definition of entropy, given in (13),

$$dS = \alpha \int_H \tilde{\epsilon} d\lambda \left[\theta - \lambda \left(\frac{1}{2} \theta^2 + \|\sigma\|^2 + R_{ab} \ell^a \ell^b \right) \right]_p. \quad (39)$$

Since we are now dealing with a nonequilibrium setting, we expect that the entropy can be expressed as a sum of

two different contributions $dS = d_e S + d_i S$. Moreover, for the argument given in VA, we are able to identify the form of the nonequilibrium, unbalanced, entropy terms. Therefore, we can write

$$d_e S = \alpha \int_H \tilde{\epsilon} d\lambda (\theta - \lambda R_{ab} \ell^a \ell^b)_p \quad (40)$$

$$d_i S = -\alpha \int_H \tilde{\epsilon} d\lambda \lambda \left(\frac{1}{2} \theta^2 + \|\sigma\|^2 \right)_p, \quad (41)$$

and separate, as previously argued, the reversible and irreversible levels
  (i) (40) $= \delta Q/T$, at the reversible level,
  (ii) (41) $= \delta N$, at the irreversible level.
  From the first expression above, one has

$$\begin{aligned} d_e S &= \alpha \int_H \tilde{\epsilon} d\lambda (\theta - \lambda R_{ab} \ell^a \ell^b)_p \\ &= -\frac{2\pi}{\hbar} \int_H \tilde{\epsilon} d\lambda \lambda T_{ab} \ell^a \ell^b = \frac{\delta Q}{T}, \quad (42) \end{aligned}$$

where the heat flux is still defined by the expression in (15). Even for the nonequilibrium setting the reversible heat will vanish at $\lambda = 0$. Thereby, at the zero order in λ, one deduces again $\theta_p = 0$, while, at the first order, the relation $R_{ab} + \Phi g_{ab} = (2\pi/\hbar\alpha) T_{ab}$ is recovered for all null vectors ℓ^a. Following the previous discussion this implies, together with the conservation of the matter stress-energy tensor, the Einstein equations if $\alpha = (4\hbar G)^{-1}$.

On the other hand, for the irreversible level, we have, in accordance with (22),

$$\delta N = d_i S = -\alpha \int_H \tilde{\epsilon} d\lambda \lambda \|\sigma\|_p^2. \quad (43)$$

This again identifies the shear contribution as an internal entropy term, associating it to some irreversible dissipative process occurring in the thermal Rindler wedge.

To get a physical interpretation of δN with respect to the thermal properties of the Rindler wedge, it is helpful to express Eq. (43) in terms of the Killing horizon parameters. In the new frame,

$$\delta N = d_i S = \frac{\alpha}{\kappa} \int_H \tilde{\epsilon} dv \|\hat{\sigma}\|_p^2 \geq 0, \quad (44)$$

in accordance with the second law of thermodynamics.

By a comparison with expression (36), one can actually interpret the expression in (44) as the standard entropy production term for a fluid with shear viscosity η, defined by

$$\frac{2\eta}{T} = \frac{\alpha}{\kappa}, \quad (45)$$

that is $\eta = \hbar\alpha/4\pi$, in agreement with the universal relation for the viscosity to entropy density ratio found in the AdS/CFT context [30].

While the previous discussion shows that the spacetime thermodynamics nicely fit into a nonequilibrium setting, we now want to take this argument a step further and ask whether the expression in (44) can be effectively related to some gravitational energy flux.

The expression for the *uncompensated heat* given in (44) quantifies the energy of the system which is effectively dissipated by the viscous process,

$$T \delta N = \frac{\alpha T}{\kappa} \int_H \tilde{\boldsymbol{\epsilon}} dv \|\hat{\sigma}\|_p^2. \qquad (46)$$

Then, by substituting $\alpha = (4\hbar G)^{-1}$ from the reversible sector of the thermodynamical approach, the quantity in (46) reads

$$T \delta N = \frac{1}{8\pi G} \int_H \tilde{\boldsymbol{\epsilon}} dv \|\hat{\sigma}\|_p^2, \qquad (47)$$

which coincides with the Hartle-Hawking formula for the tidal heating of a classical black hole [12–16].[6]

This is a striking result as it defines the internal entropy production as a purely gravitational effect. Indeed, it can be associated with the work done on the horizon by the perturbative tidal field which is described by the electric part of the Weyl curvature tensor.[7] Furthermore, relation (45) suggests that such a work has to be seen as acted upon the internal/microscopic degrees of freedom of the theory rather than on macroscopic quantities [in this sense (47) cannot be interpreted as a standard/reversible work term]. The horizon viscosity implies that such a work will be converted into internal heat. Hence, the presence of the internal entropy term can then be directly related to the process of dissipation via gravitational/internal degrees of freedom. In this sense, the irreversible sector contains the information about the possible activation/propagation of such degrees of freedom of the theory.

VII. NONEQUILIBRIUM SPACETIME THERMODYNAMICS: $f(R)$

A crucial assumption in the previous derivation was the validity of the SEP which allowed us to consider the entropy density α as a constant. One might wonder what are the consequences of relaxing such an assumption in favor of the less restrictive EEP? In this case, one might generically expect that the entropy density is promoted to a spacetime function (basically because the EEP implies a

spacetime dependent Newton constant). However, in the definition for the entanglement entropy of the Rindler wedge, this implies a possibly very complicated spacetime dependence for the UV cutoff. Furthermore, the specific form of such a cutoff is not uniquely fixed by the EEP correspondingly to the fact that the latter generically allows for many generalized theories of gravity.

In order to make the argument as simple as possible, and following the treatment of [10], we consider here the specific case of $f(R)$ gravity, which is known to be equivalent to a single field scalar-tensor theory (more precisely a Brans-Dicke theory with $\omega = 0$ and a specific potential for the scalar field [17]).[8] In this case, the UV cutoff is known to be proportional to some function of the curvature $\mathbf{f}(R) \equiv f'(R)$ (where the prime indicates the derivative with respect to R), playing the role of the inverse of the gravitational coupling. In this case, the area-entropy relation is known to be given by

$$S = \alpha \mathbf{f}(R)\tilde{\boldsymbol{\epsilon}} \qquad (48)$$

where α is still a constant (albeit *a priori* different from the one considered in the previous section).

It is easy to see that in this case the entropy variation along the null congruence will be

$$\frac{dS}{d\lambda} = \alpha \left(\frac{d\mathbf{f}}{d\lambda} \tilde{\boldsymbol{\epsilon}} + \mathbf{f} \frac{d\tilde{\boldsymbol{\epsilon}}}{d\lambda} \right), \qquad (49)$$

where, by definition $\tilde{\boldsymbol{\epsilon}}^{-1} d\tilde{\boldsymbol{\epsilon}}/d\lambda = \theta$.

Consequently, the entropy change along the horizon will read [10]

$$dS = \alpha \int_H \tilde{\boldsymbol{\epsilon}} d\lambda (\dot{\mathbf{f}} + \mathbf{f}\theta), \qquad (50)$$

therefore acquiring, with respect to the previous argument, an extra contribution $\dot{\mathbf{f}}$ coupled to the dynamics of the scalar function $\mathbf{f}$. (Here the dot stands for the differentiation with respect to λ.)

For this reason, in order to set the instantaneous stationarity condition at p, that is $dS = 0$, the affine expansion is no more a good dynamical variable. In this sense, it is helpful to define the quantity $\tilde{\theta} \equiv (\theta \mathbf{f} + \dot{\mathbf{f}})$ as a sort of effective expansion for the congruence. Consequently, the equilibrium surface for the system will be fixed by the condition

$$\tilde{\theta}_p = 0, \qquad (51)$$

that is $\theta_p = -\dot{\mathbf{f}}/\mathbf{f}$, where $\dot{\mathbf{f}} = \mathbf{f}'(R)\ell^a R_{,a}$ is generally nonzero. In particular, this actually provides an example of LRF equilibrium surface, for which θ_p is always nonvan-

---

[6]Note that both in [16] and in [12] the Hartle-Hawking formula expressing the relation between the horizon area variation and the horizon shear is utilized with a definition of the surface gravity ϵ which is half of that used in [14].

[7]Even though the magnetic part of the Weyl tensor is actually necessary in order to define the Weyl curvature and the equations governing its propagation (Bianchi identities), this part does not play any direct role in determining the time derivative (evolution) of the congruence kinematic quantities, as it is just related to their spatial gradients.

[8]In a similar thermodynamical approach, $f(R)$ gravity is alternatively considered in [31–33]. See also [34] for a further extension of the thermodynamical perspective with a Lanczos-Lovelock gravity.

ishing, apart from the trivial case where $\mathbf{f}$ is constant, for which the theory will be equivalent to GR.

From Sec. VA, we could already expect that the presence of the nonvanishing affine expansion would produce a nonequilibrium contribution to the system entropy. In order to get a quantitative expression for the entropy change in the neighborhood of p, one again can consider an infinitesimal deviation of the entropy from its equilibrium value.

Let us then Taylor expand the integrand in (50) around p up to the first order in λ, that is

$$\tilde{\theta} = \tilde{\theta}_p + \lambda\dot{\tilde{\theta}}|_p + \mathcal{O}(\lambda^2), \qquad (52)$$

where

$$\dot{\tilde{\theta}}_p = (\ddot{\mathbf{f}} - \mathbf{f}^{-1}\dot{\mathbf{f}}^2 + \mathbf{f}\dot{\theta})_p. \qquad (53)$$

One can use the Raychaudhuri equation (12) and the geodesic equation $\ell^a\ell^b_{;a} = 0$, to obtain the $\mathcal{O}(\lambda)$ expression for the entropy change

$$dS = \alpha \int_H \tilde{\boldsymbol{\epsilon}}d\lambda\lambda[(\mathbf{f}_{;ab} - \mathbf{f}R_{ab})\ell^a\ell^b - 3/2\mathbf{f}\theta^2 - \mathbf{f}\|\sigma\|^2]_p, \qquad (54)$$

where relation (51) is used to substitute $\mathbf{f}^{-2}\dot{\mathbf{f}}^2 = \theta^2$ at p. Now, keeping the expression in (15) for the heat flux, one can finally reproduce the same approach used in Sec. VI.

At the reversible level, the generalized Clausius relation gives

$$\mathbf{f} R_{ab} - \mathbf{f}_{;ab} + \Psi g_{ab} = (2\pi/\hbar\alpha)T_{ab} \qquad (55)$$

where Ψ is an undetermined function. With the same argument given in Sec. III, one then requires the conservation of the matter stress-energy tensor and use the contracted Bianchi identity to write the commutator of the covariant derivative as $2v^c_{;[ab]} = R_{abd}{}^c v^d$. In this way, one finds

$$(\mathbf{f}R_{ab} - \mathbf{f}_{;ab})^{;a} = (\tfrac{1}{2}f - \Box\mathbf{f})_{,b}, \qquad (56)$$

and thereby

$$\Psi = (\Box\mathbf{f} - \tfrac{1}{2}\mathbf{f}). \qquad (57)$$

Eventually, Eq. (57), together with (55) exactly leads, as expected, to the field equations of $f(R)$ gravity

$$\mathbf{f} R_{ab} - \mathbf{f}_{;ab} + \left(\Box\mathbf{f} - \frac{1}{2}f\right)g_{ab} = \frac{2\pi}{\hbar\alpha}T_{ab}, \qquad (58)$$

with the identification $\alpha = (4\hbar G)^{-1}$. In [10], the same result was obtained starting from the entropy balance relation (21), assuming $\sigma = 0$, and then identifying the extra entropy term in θ in the second line of (55) with a suitable internal entropy term. There, it was also shown that the alternative route of keeping the θ^2 in Eq. (56) is not compatible with the conservation of the matter energy-momentum tensor.

Indeed, following the previous discussion, the above term is expected (together with a shear dependent term) as an unavoidable contribution related to the irreversible sector of the generalized Clausius relation (25)

$$\delta N = -\int_H \tilde{\boldsymbol{\epsilon}}d\lambda\lambda(\alpha\mathbf{f})\left[\frac{3}{2}\theta^2 + \|\sigma\|^2\right]_p, \qquad (59)$$

which, as explained in Sec. V, identifies the internal entropy production terms of the system.

As expected, the internal entropy in (59) now shows a contribution both from scalar and tensorial degrees of freedom. Indeed, by using the same argument as in the GR case, we again have a natural interpretation for the expression in (59) as the dissipative function of the system.

The shear squared contribution is equivalent to the one found for GR, with a shear viscosity coefficient which now takes a factor $\mathbf{f}$,

$$\eta = \frac{\hbar\alpha\mathbf{f}}{4\pi}, \qquad (60)$$

as a consequence of the UV cutoff chosen for the area-entropy relation.

On the other hand, the internal entropy contribution due to the scalar degree of freedom is now given by

$$d_iS_\theta = -\int_H \tilde{\boldsymbol{\epsilon}}d\lambda\lambda(\alpha f)\frac{3}{2}\theta_p^2. \qquad (61)$$

By making use of the kinematical analogy described in Sec. VA, and by expressing the above equation in the Killing frame, one is naturally led to define the bulk viscosity ζ as

$$\frac{\zeta}{T} = \frac{3\alpha\mathbf{f}}{2\kappa}, \qquad (62)$$

that is $\zeta = 3\hbar\alpha\mathbf{f}/4\pi$, as already found in [10].

In order to give a physical interpretation to (61), one can use the equivalence between $f(R)$ and scalar-tensor gravity, thereby interpreting $\mathbf{f}$ as an effective massive dilaton.

The action for $f(R)$ gravity is given by

$$\mathcal{S} = \frac{\hbar\alpha}{4\pi} \int d^4x\sqrt{-g}f(R) + \mathcal{S}_{\text{mat}}. \qquad (63)$$

By introducing an auxiliary field $\varphi \equiv \mathbf{f}(R)$[9] and assuming $f''(R) \neq 0$ for all R, one can take $V(\varphi)$ as the Legendre transform of $f(R)$ so that $R = V'(\varphi)$, thereby rewriting the expression in (63) as

$$\mathcal{S} = \frac{\hbar\alpha}{4\pi} \int d^4x\sqrt{-g}[\varphi R + V(\varphi)] + \mathcal{S}_{\text{mat}}. \qquad (64)$$

The Euler-Lagrange equations, in the Jordan frame, take

---

[9]Note that the fact that in $f(R)$ gravity the associated scalar field is not a generic spacetime function but rather just of R, makes it possible to derive a closed system of equations without having to assume the equations of motion of the scalar field separately.

the form

$$\varphi\left(R_{ab} - \frac{1}{2}g_{ab}R\right) + (g_{ab}\Box - \nabla_a\nabla_b)\varphi + \frac{1}{2}g_{ab}V(\varphi)$$
$$= \frac{2\pi}{\hbar\alpha}T_{ab}, \tag{65}$$

equivalent to field equations given in (58).

In this frame, by using the relation (51), one can express the dissipated energy coupled to the bulk and shear viscosity in (61), in terms of the auxiliary scalar field φ

$$T\delta N = -\int_H \tilde{\boldsymbol{\epsilon}}\,d\lambda\,\lambda(\alpha\varphi)T\left[\frac{3}{2}\varphi^{-2}\dot{\varphi}^2 + \|\sigma\|^2\right]_p. \tag{66}$$

Similarly to the GR case, one expects that relation (66) expresses some purely gravitational energy loss for the system, this time involving both scalar and tensorial fluxes. The interpretation of the term related to the shear is straightforward as it is clearly the generalization to a scalar-tensor theory of the tidal heating already obtained for the GR case. More problematic is the interpretation of the bulk viscosity (purely scalar) contribution as no equivalent derivation as that for the tidal heating term in GR has to be performed (to our knowledge) for scalar-tensor theories of gravity.

In this direction, as a first step, one can look for the effective field source terms which drive the local deformation of the null horizon congruence. Again, we consider the effective expansion $\tilde{\theta}$ as the suitable quantity to describe the scalar perturbations of the horizon given that $\tilde{\theta} = 0$ is the condition for equilibrium.

Let us then use Eq. (53) as the effective expansion rate for the horizon bundle at p. Starting from this equation, where now $\mathbf{f}$ is substituted by φ, then using the Raychaudhuri equation and the Jordan frame field equation given in (66), one gets

$$\dot{\tilde{\theta}}_p = -\frac{\theta_p^2}{2} - \|\sigma\|_p^2 - (2\pi/\hbar\alpha)\varphi^{-1}T_{ab}\ell^a\ell^b$$
$$- \varphi^{-2}\nabla_a\varphi\nabla_b\varphi\ell^a\ell^b. \tag{67}$$

Now, since we are interested only in the scalar contributions to $\dot{\tilde{\theta}}_p$, we can set $T_{ab} = 0$ and $\|\sigma\|_p^2 = 0$. Then, by using the equilibrium relation in the new frame $\theta = -\varphi^{-1}\dot{\varphi}$, we are left with

$$\dot{\tilde{\theta}}_p = -\frac{3}{2}\varphi^{-2}\nabla_a\varphi\nabla_b\varphi\ell^a\ell^b = -\frac{3}{2}\varphi^{-2}\dot{\varphi}^2. \tag{68}$$

We see now that Eq. (68) identifies the quantity $(3/2)\varphi^{-2}\dot{\varphi}^2$ as what one might define as the gravitational energy flux associated with the solely scalar field degrees of freedom.

In conclusion, we can now provide a clean interpretation of the viscous terms in (66) as those representing the thermal system internal energy loss due to both scalar and tensor gravitational energy fluxes through the horizon.

In particular, by moving to the Killing frame, (66) can be rewritten as

$$T\delta N = \frac{1}{8\pi G}\int_H \tilde{\boldsymbol{\epsilon}}\,dv\left[\frac{3}{2}\varphi^{-1}\dot{\varphi}^2 + \varphi\|\sigma\|^2\right]_p, \tag{69}$$

where we have set $\alpha = (4\hbar G)^{-1}$ as required by the equations of motion (58). The striking similarity of the above expression with the energy loss rate due to the gravitational radiation in scalar-tensor gravity [see, e.g., Eq. (10.135) of [35]] further reinforces the above suggested interpretation. Furthermore, we now see that also for the case of $f(R)$ gravity the relations (60) and (62) plus the above interpretation of (69) provide a correlation between the transmission coefficient of the gravitational energy through the horizon and the horizon congruence viscosity.

VIII. DISCUSSION

In nonequilibrium spacetime thermodynamics the viscous dissipative effects appear to be naturally associated with purely gravitational energy fluxes. Their association to the irreversible/dissipative sector of the theory strongly suggests an interpretation of their nature as, nonlocal, internal heat flows associated with the internal spacetime degrees of freedom and clarifies why in GR a local, background independent, description of gravitational waves is precluded.

Noticeably, in order to recover the field equations one always needs to effectively isolate these dissipative contributions by neatly separating the reversible and irreversible sectors of the constitutive equation. The analogy between the horizon null congruence and a continuous medium allows one to recognize the natural terms related to the irreversible sector of the entropy balance. This effective separation of the reversible and irreversible regimes is further supported, at least in GR, by the fact that the energy contributions occurring in the equilibrium constitutive relations have a local nature, being always related to the gravity field sources (Ricci curvature), whereas the nonequilibrium terms are intrinsically nonlocal and related to those curvature components which are independent from the source's distribution (Weyl curvature). This actually shows that the thermodynamical derivation of the gravitational field equations is very general as it is sensitive to the whole spacetime curvature.[10]

However, a different issue is the interpretation of the internal entropy production terms related to dissipation with respect to a particular spacetime solution. While the association between internal entropy and allowed form of gravitational fluxes seems quite clear (e.g., we showed that

---

[10]The situation in $f(R)$ gravity is less clear due to the presence of the scalar field dissipative contribution in (69). It has, however, to be noticed that this is not a part of the stress-energy tensor of the scalar field for the scalar-tensor theory equivalent to $f(R)$.

the energy dissipated in GR *coincides* exactly with the Hartle-Hawking tidal heating term), it might seem however puzzling that the arbitrariness in the choice of B allows for nonzero shear and expansion of the null congruence (and hence for internal entropy production terms) even, for example, if one imagines to have performed the local Rindler wedge construction in a Minkowski spacetime.

In fact, the thermodynamical approach is providing us just with the constitutive equations of the thermal system associated with the local Rindler wedge, not of the spacetime in which the latter is constructed. The arbitrariness of the choice of B (and hence of the thermal system properties) implies that such equations will at most characterize the structure of the gravitational theory selected by the entropy-area relation (the EP formulation). In this sense they will not be associated with physical fluxes or curvatures of the spacetime as a whole. Hence, the possible presence of internal entropy terms, even when the local Rindler wedge is constructed in flat spacetime, does not imply that the latter can be seen as a system in a nonequilibrium state.

Of course, one might take an alternative point of view, and claim that the above discussion actually shows an intrinsic limitation of the standard construction adopted here, as in [4,10]. In addressing this issue, a possibility could consist in a further characterization of the local Rindler wedge construction. In fact, one might choose to construct the 2-surface B in such a way that it will be sensitive to the local curvature at p and reduce to a plane in the flat spacetime case (i.e., B would lead generically to a nonzero θ_p and σ_p but would also reduce to the standard bifurcation surface at p for a Rindler wedge, whose orthogonal null congruence has $\theta_p = \sigma_p = 0$, in the flat spacetime limit). For example, this could be achieved by constructing B as a totally geodesic 2-dimensional spacelike submanifold of the spacetime passing through p [36]. (That is, any geodesic passing through p and there tangent to B would have to be completely contained in B.) Within this alternative approach, while all of the formulas would still pertain to the thermodynamical behavior of the local Rindler wedge at p,[11] they would now be able to specialize to a specific spacetime choice and hence link the dynamical behavior of the wedge to the actual local matter-curvature content of the chosen spacetime.

We do not see at the moment a decisive argument to go in one sense or the other. All in all, the whole point of the thermodynamical approach is not to provide an instrument able to reconstruct the kind of spacetime one is living in. Rather it is aimed to put in evidence the thermodynamical structure and the internal degrees of freedom of gravitational theories. In this sense the traditional construction,

with an arbitrary B, seems sufficient. We plan however to further explore this issue in future work.

Another important aspect of this approach highlighted by this work has to do with the role that the different equivalence principle formulations play in selecting the possible gravitational dynamics.

The *strong* equivalence principle implies, for a generic choice of B, $\theta_p = 0$, thereby leading to the equations of motion of GR (with irreversible level only corresponding to tensorial gravitational fluxes).[12]

The *Einstein* equivalence principle, does not fix either θ nor σ, thereby allowing for a generalized theory of gravity. For the simple case of $f(R)$, we showed how the scalar degree of freedom in fact produces a dissipative contribution which is actually associated with some purely gravitational scalar energy flux through the system boundary (and which is not part of the stress-energy tensor of the scalar field).

The presence of an unavoidable purely scalar gravitational energy flux seems to indicate that GR and scalar tensor are truly separated theories. However, the very same thermodynamical approach might also suggest to look at such gravitational theories as different regimes of some more general effective description of gravity (in the same way as incompressible flows can be seen as a special regime of a general compressible fluid). This seems an intriguing possibility worth further investigation as it might lead to a unified framework which associates different gravitational theories with different hydrodynamical regimes of the analogue flow associated with the horizon null congruence.

For what regards the spacetime viscosity, the quantities η and ζ are found to be always related to the UV cutoff scale of the theory through the entropy density. The UV cutoff scale can be identified as the scale at which the gravitational action is of the order of the quantum of action $\hbar$, therefore where the horizon is subject to some quantum fluctuations, the so-called horizon *Zitterbewegung*. This naturally suggests to interpret the spacetime dissipative effects as a consequence of an underlying fluctuation behavior of spacetime at the UV cutoff scale, as already considered in the seventies by Candelas and Sciama [37] and nowadays explored in different terms in the AdS/CFT correspondence context [11,38]. In this sense, it is interesting to note that while in principle α can *a priori* depend on the nature of the quantum fields and their interactions, consistency of the gravitational equations derived from the thermodynamical approach implies a trivial relation

---

[11]E.g., one can talk about dissipation only with reference to the local Rindler wedge as spacetime as a whole as to be seen as a conservative system.

[12]We suspect that the same formulation, together with the restriction of patches B such that the null congruence shear is zero at any point p (taken now as a defining property of the theory), may imply a restriction to the conformally invariant sector of GR (Einstein equations plus the condition of everywhere vanishing Weyl), which is basically equivalent to Nordström gravity.

to the Newton constant. This might suggest an underlying microscopic interpretation of gravity along the ideas of induced gravity (see, e.g., [39] for a related discussion).

Finally, the separation between reversible and irreversible sectors of gravity seems to indicate that dissipation can occur just at the gravitational level, without any contribution from matter, however phenomena like the Hawking radiation, with its backreaction on the geometry (black hole evaporation), seems to suggest that this separation

might be a by-product of the test field treatment so far adopted.

[1] J. D. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
[2] J. M. Bardeen, B. Carter, and S. W. Hawking, Commun. Math. Phys. **31**, 161 (1973).
[3] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Spacetime* (Cambridge University Press, Cambridge, England, 1973); S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975); **46**, 206 (1976).
[4] T. Jacobson, Phys. Rev. Lett. **75**, 1260 (1995).
[5] T. Padmanabhan, Astrophys. Space Sci. **285**, 407 (2003); Int. J. Mod. Phys. D **17**, 2585 (2008).
[6] T. Padmanabhan, arXiv:0903.1254v1.
[7] T. P. Sotiriou and S. Liberati, Phys. Rev. D **74**, 044016 (2006).
[8] G. 't Hooft, in *Abdus Salam Festschrift: A Collection of Talks* (World Scientific, Singapore, 1993).
[9] L. Susskind, J. Math. Phys. (N.Y.) **36**, 6377 (1995).
[10] C. Eling, R. Guedens, and T. Jacobson, Phys. Rev. Lett. **96**, 121301 (2006).
[11] C. Eling, J. High Energy Phys. 11 (2008) 048.
[12] S. Chandrasekhar, *The Mathematical Theory of Black Holes* (Clarendon Press, Oxford, 1983).
[13] E. Poisson, Phys. Rev. D **70**, 084044 (2004).
[14] S. W. Hawking and J. B. Hartle, Commun. Math. Phys. **27**, 283 (1972).
[15] E. Poisson, Phys. Rev. Lett. **94**, 161103 (2005).
[16] S. A. Teukolsky and W. H. Press, Astrophys. J. **193**, 443 (1974).
[17] T. P. Sotiriou and V. Faraoni, arXiv:0805.1726v3 [Rev. Mod. Phys. (to be published)].
[18] C. M. Will, Living Rev. Relativity **9**, 3 (2006), http://www.livingreviews.org/lrr-2006-3.
[19] T. Sotiriou (private communication).
[20] W. G. Unruh, Phys. Rev. D **14**, 870 (1976); W. G. Unruh and N. Weiss, Phys. Rev. D **29**, 1656 (1984).
[21] G. 't Hooft, Nucl. Phys. **B256**, 727 (1985).
[22] L. Bombelli, R. K. Koul, J. H. Lee, and R. D. Sorkin, Phys. Rev. D **34**, 373 (1986).
[23] R. Haag, D. Kastler, and E. B. Trich-Pohlmeyer, Commun. Math. Phys. **38**, 173 (1974).
[24] J. M. Gerard, Classical Quantum Gravity **24**, 1867 (2007).
[25] S. R. de Groot and P. Mazur, *Non-equilibrium Thermodynamics* (North-Holland Publishing Company, Amsterdam, 1962).
[26] R. H. Price and K. S. Thorne, Phys. Rev. D **33**, 915 (1986).
[27] K. S. Thorne, R. H. Price, and D. A. Macdonald, *Black Holes: The Membrane Paradigm* (Yale University Press, New Haven, Connecticut, 1986).
[28] D. Jou, J. Casas-Vázquez, and G. Lebon, *Extended Irreversible Thermodynamics* (Springer-Verlag, Berlin, 2001), 3rd ed..
[29] E. Poisson, *A Relativist's Toolkit: The Mathematics of Black Hole Mechanics* (Cambridge University Press, Cambridge, England, 2004).
[30] J. M. Maldacena, Adv. Theor. Math. Phys. **2**, 231 (1998).
[31] M. Akbar and R. Cai, Phys. Lett. B **648**, 243 (2007).
[32] K. Bamba, C. Geng, S. Nojiri, and S. D. Odintsov, arXiv:0909.4397v1.
[33] E. Elizalde and P. J. Silva, Phys. Rev. D **78**, 061501(R) (2008).
[34] A. Paranjape, S. Sarkar, and T. Padmanabhan, Phys. Rev. D **74**, 104015 (2006); D. Kothawala and T. Padmanabhan, Phys. Rev. D **79**, 104020 (2009).
[35] C. M. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, Cambridge, England, 1993), Revised edition.
[36] B. O'Neill, *Semi-Riemannian Geometry: With Applications to Relativity* (Academic Press, San Diego, California, 1983).
[37] P. Candelas and D. W. Sciama, Phys. Rev. Lett. **38**, 1372 (1977).
[38] G. Policastro, D. T. Son, and A. O. Starinets, Phys. Rev. Lett. **87**, 081601 (2001).
[39] T. Jacobson, arXiv:gr-qc/9404039v1.