PHYSICAL REVIEW D **94,** 044014 (2016)

Canonical energy and hairy AdS black holes

Seungjoon Hyun,* Sang-A Park,† and Sang-Heon Yi‡

Department of Physics, College of Science, Yonsei University, Seoul 120-749, Korea

(Received 7 April 2016; published 9 August 2016)

We propose the modified version of the canonical energy which was introduced originally by Hollands and Wald. Our construction depends only on the Euler-Lagrange expression of the system and thus is independent of the ambiguity in the Lagrangian. After some comments on our construction, we briefly mention on the relevance of our construction to the boundary information metric in the context of the AdS/CFT correspondence. We also study the stability of three-dimensional hairy extremal black holes by using our construction.

DOI: 10.1103/PhysRevD.94.044014

I. INTRODUCTION

Black hole stability has been one of the important persisting issues in black hole physics, whose study has realistic implications, for instance, that the observed black holes would be stable ones if they are found experimentally. This also has interesting applications in the context of the AdS/CFT correspondence since the stability or unitarity of the finite temperature field theory system would be dual to the stability of black holes or branes according to the AdS/CFT dictionary. It has been well known that there are at least two kinds of stability concepts in black holes, one of which is known as the dynamical stability and the other is the thermodynamic one. At the linearized level, the dynamical stability is determined by mode analysis for perturbing black hole solutions, while the thermodynamics stability is concerned about the stability of black holes relative to other thermodynamic states in an appropriate ensemble. Interestingly, it has been conceived that two kinds of stability have different natures and so do not coincide in general. However, there was a conjecture by Gubser and Mitra that the thermodynamics instability implies the dynamical one at least for black branes [1,2]. Since this conjecture relates two different kinds of analysis on black holes, it may indicate a possibility of another approach to the dynamical stability different from the standard mode analysis.

Recently, another method for the linear dynamical stability of black holes was developed by Hollands and Wald (HW) [3], which uses the machinery in the covariant phase space through the second variation of the covariant symplectic form, named the *canonical energy.* By using the canonical energy method, HW have proved the Gubser-Mitra conjecture and showed the consistency of their method with another important criterion for the black hole stability known as the local Penrose inequality [4]. This canonical energy method has been applied successfully to the extremal black holes and the asymptotic AdS space [5–7]. More recently, the HW canonical energy in AdS space is conjectured to be dual to the Fisher information metric on the dual quantum system [8]. This conjecture is based on several interesting properties of the HW canonical energy and is checked explicitly in concrete examples.

Though the HW canonical energy method has a great advantage in some aspects over the standard mode analysis, it raises the following question. Basically, the HW canonical energy method is not based on the equations of motion (EOM) but on the Lagrangian, while other methods for the stability criterion utilize (linearized) EOM or the solutions themselves. Therefore, it seems to be better if we could construct the canonical energy or the modified canonical energy by using EOM or the Euler-Lagrange expression, not the Lagrangian. This construction may be relevant especially when only the EOM are known, for example, as in the type IIB supergravity case. In this paper, we attempt to construct the modified version of the HW canonical energy by using the Euler-Lagrange expression only. Through this construction, it is realized that the canonical energy may have some freedom in its definition, which may be relevant in its interpretation as the dual to the Fisher information metric.

The paper is organized as follows. In Sec. II, we review on the quasilocal formalism for charges developed in [9–13], which may be regarded as the EOM alternative to the covariant phase space approach. Then we introduce a modified canonical energy based on the quasilocal Abbott-Deser-Tekin (ADT) formalism and see its relation to the canonical energy introduced by HW [3] in Sec. III. In pure Einstein gravity, we show that our modified canonical energy leads to the bulk expression only, in contrast to the HW canonical energy. In fact, it turns out that one can add the boundary terms without destroying the properties for the canonical energy and so this seems to indicate that there is freedom in the definition of the canonical energy. In Sec. IV, we study the stability of extremally rotating hairy black holes in three dimensions by using the canonical energy method.

*sjhyun@yonsei.ac.kr
†sangapark@yonsei.ac.kr
‡shyi@yonsei.ac.kr

 © 2016 American Physical Society

II. REVIEW: QUASILOCAL FORMALISM FOR CHARGES

In this section, we review the quasilocal ADT formalism based on the identically conserved current and summarize some properties of the generalized ADT current [9–14], which relies on EOM or Euler-Lagrange expression. To summarize our conventions and present results succinctly, let us denote the metric and matter fields collectively as $\Psi = (g, \psi)$. The Euler-Lagrange expression can also be written collectively as $\mathcal{E}_\Psi = (\mathcal{E}_{\mu\nu}, \mathcal{E}_\psi)$. The Bianchi or Noether identity for a diffeomorphism parameter ζ^μ may be written as

$$\nabla_\mu (2\mathbf{E}^{\mu\nu}\zeta_\nu) = -\mathcal{E}_\Psi \pounds_\zeta \Psi, \tag{1}$$

where $\pounds_\zeta$ denotes the Lie derivative along the vector field ζ^μ and $\mathbf{E}^{\mu\nu} \equiv \mathcal{E}^{\mu\nu} - \frac{1}{2}\mathcal{Z}^{\mu\nu}$. Here, the $\mathcal{Z}$-tensor is given by a linear combination of the product of the matter field ψ and the matter Euler-Lagrange expression $\mathcal{E}_\psi$. Concretely, for an ℓth rank tensor field $\psi_{\mu_1\ldots\mu_\ell}$, one can see that the $\mathcal{Z}$-tensor is given by

$$\mathcal{Z}^{\mu\nu}(\mathcal{E}_\psi, \psi) = \mathcal{E}_\psi^{\mu\alpha_2\alpha_3\ldots\alpha_\ell}\psi^\nu_{\alpha_2\alpha_3\ldots\alpha_\ell} + \mathcal{E}_\psi^{\alpha_1\mu\alpha_3\ldots\alpha_\ell}\psi^\nu_{\alpha_1\alpha_3\ldots\alpha_\ell} + \cdots, \tag{2}$$

which will be represented schematically as $\mathcal{Z}^{\mu\nu} = \mathcal{E}^\mu_\psi \circ \psi^\nu$. Note that the $\mathcal{Z}$-tensor vanishes for minimally coupled scalar fields.

In the quasilocal ADT formalism, the on-shell ADT current [15–18] is generalized to the off-shell level. The off-shell conserved current $\mathbf{J}^\mu$ is composed of two pieces, the generalized off-shell ADT current $\mathcal{J}^\mu_{\rm ADT}$ and the additional term $\mathcal{J}^\mu_\Delta$, as

$$\mathbf{J}^\mu(\zeta;\Psi,\delta\Psi) = \mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi) + \mathcal{J}^\mu_\Delta(\Psi|\pounds_\zeta\Psi,\delta\Psi), \tag{3}$$

where the ADT current is defined by (see [11,12] for some details)

$$\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi) = \delta(\sqrt{-g}\mathbf{E}^{\mu\nu}\zeta_\nu) - \sqrt{-g}\mathbf{E}^\mu_\nu\delta\zeta^\nu + \frac{1}{2}\sqrt{-g}\zeta^\mu\mathcal{E}_\Psi\delta\Psi. \tag{4}$$

The additional term $\mathcal{J}^\mu_\Delta$, which vanishes when ζ is a Killing vector, is introduced to preserve the off-shell conservation property of the current even for an asymptotic Killing vector ζ^μ. This additional term is given by the iterative integration by parts on a specific combination of the Euler-Lagrange expression [11] and the current $\mathbf{J}^\mu$ can be shown to be conserved at the off-shell level (see Appendix A for its derivation). The additional current $\mathcal{J}^\mu_\Delta$ is symplectic and vanishes for a Killing vector ζ^μ, which can be related to the symplectic current ω^μ in the covariant phase space [19] as given in Eq. (A5). Despite its relation to the symplectic current, we would like to emphasize that the additional current $\mathcal{J}^\mu_\Delta$ is constructed solely from the Euler-Lagrange expression.

Since the off-shell current $\mathbf{J}^\mu$ is conserved identically, it can be written in terms of the (antisymmetric) off-shell potential $\mathbf{Q}^{\mu\nu}$ as

$$\sqrt{-g}\mathbf{J}^\mu = \partial_\nu(\sqrt{-g}\mathbf{Q}^{\mu\nu}), \tag{5}$$

and the linearized quasilocal ADT charges for an asymptotic Killing vector ζ^μ are defined by

$$\delta Q(\zeta) = \frac{1}{8\pi G}\int dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(\zeta;\Psi,\delta\Psi), \tag{6}$$

where this expression should be evaluated on shell at the final stage of computation. Here, we would like to emphasize that the off-shell current $\mathbf{J}^\mu$ and potential $\mathbf{Q}^{\mu\nu}$ are constructed from the Euler-Lagrange expression and thus they are free from the ambiguity or the noncovariance in the Lagrangian. As is derived in Appendix A, the off-shell potential $\mathbf{Q}^{\mu\nu}$ can be related to the Noether potential $K^{\mu\nu}$ and the surface term Θ^μ of the Lagrangian, up to a total derivative term, as

$$2\sqrt{-g}\mathbf{Q}^{\mu\nu}(\zeta;\Psi,\delta\Psi) = \delta K^{\mu\nu}(\zeta) - K^{\mu\nu}(\delta\zeta) - 2\zeta^{[\mu}\Theta^{\nu]}(\delta\Psi) + \sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi). \tag{7}$$

For a Killing vector, the charge expression is completely consistent with the conventional ADT expression at the asymptotic infinity and is also consistent with the covariant phase space formalism, for instance, in the computation of the black hole entropy. Furthermore, for asymptotic Killing vectors, this charge expression can be used to obtain asymptotic symmetry generators [11].

Before going ahead, we would like to remark the special properties of the ADT current $\mathcal{J}^\mu_{\rm ADT}$ and the additional current $\mathcal{J}^\mu_\Delta$. Firstly, one may note that the ADT current, $\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi)$, is off-shell conserved when ζ is a Killing vector, and that it depends linearly on ζ. This ADT current can be written in terms of a differential operator $\mathcal{D}_\Psi$ acting on the field variation $\delta\Psi$ as

$$\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi) = \zeta_\nu(\mathcal{D}_\Psi\delta\Psi)^{\mu\nu}. \tag{8}$$

When the background configuration Ψ satisfies EOM, $\mathcal{E}_\Psi = 0$, the current reduces to

$$\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi)|_{\mathcal{E}_\Psi=0} = \delta\mathbf{E}^{\mu\nu}\zeta_\nu = \left(\delta\mathcal{E}^{\mu\nu} - \frac{1}{2}\delta\mathcal{E}^\mu_\psi\circ\psi^\nu\right)\zeta_\nu, \tag{9}$$

which vanishes for an arbitrary ζ^μ when we impose that $\delta\Psi$ also satisfies linearized EOM, $\delta\mathcal{E}_\Psi = 0$. This property will be called the "on-shell" vanishing property of the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ in the following. In summary, the generalized off-shell ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ has the off-shell conservation property for a Killing vector and the on-shell vanishing property for an arbitrary ζ as

$$\nabla_{\mu}\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta\Psi)=0 \quad \text{off shell} \quad \text{for a Killing vector}\,\zeta, \tag{10}$$

$$\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta\Psi)=0 \quad \text{on shell} \quad \text{for arbitrary}\,\zeta. \tag{11}$$

Since $\mathcal{J}^{\mu}_{\Delta}=0$ for a Killing vector K, the above off-shell conservation property for a Killing vector can be rephrased as

$$\mathcal{J}^{\mu}_{\rm ADT}(K)=\mathbf{J}^{\mu}(K)=\nabla_{\nu}\mathbf{Q}^{\mu\nu}(K).$$

Now, let us consider the second variation of the ADT current for an arbitrary vector ζ^{μ}, which would be relevant for the construction of the bilinear form on the first order variation space. Note that the variation of the ADT current $\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_1\Psi)$ can be written in terms of three pieces as

$$\begin{aligned}\delta_2(\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_1\Psi)) &= \mathcal{J}^{\mu}_{\rm ADT}(\delta_2\zeta;\Psi,\delta_1\Psi)\\ &\quad+\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\delta_2\Psi,\delta_1\Psi)\\ &\quad+\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_2\delta_1\Psi),\end{aligned} \tag{12}$$

where the first and the last term in the right-hand side are given through the representation of the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ in Eq. (8), explicitly by

$$\begin{aligned}\mathcal{J}^{\mu}_{\rm ADT}(\delta_2\zeta;\Psi,\delta_1\Psi)&=\delta_2\zeta^{\nu}(\mathcal{D}_{\Psi}\delta_1\Psi)^{\mu}_{\nu},\\ \mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_2\delta_1\Psi)&=\zeta^{\nu}(\mathcal{D}_{\Psi}\delta_2\delta_1\Psi)^{\mu}_{\nu}.\end{aligned} \tag{13}$$

Therefore the second term in the right-hand side may be thought of as defined by the variations of the ADT current as

$$\begin{aligned}\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\delta_2\Psi,\delta_1\Psi)&\equiv\delta_2(\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_1\Psi))\\ &\quad-\mathcal{J}^{\mu}_{\rm ADT}(\delta_2\zeta;\Psi,\delta_1\Psi)\\ &\quad-\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_2\delta_1\Psi),\end{aligned} \tag{14}$$

which would be a good candidate for a symmetric bilinear form on the space of on-shell first order variations or linear perturbations. Note also that this current expression $\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\delta_2\Psi,\delta_1\Psi)$ depends linearly on the vector ζ^{μ} while it is independent of the derivatives of ζ^{μ}, as can be inferred from the representation of the ADT current given in Eq. (8).

III. MODIFIED CANONICAL ENERGY

By using the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$, we would like to introduce a modified canonical energy $\mathcal{E}(\delta_1\Psi,\delta_2\Psi)$, whose original form was defined in [3]. The essential properties of the canonical energy proposed in [3] may be summarized as follows. It is a symmetric bilinear form on the first field variation $\delta\Psi$, gauge-invariant, monotonic along the "time evolution" and conserved in the sense that it does not depend on the choice of the Cauchy surface for given boundaries. As will be shown in the following, one may use the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ instead of the symplectic current ω^{μ} to construct a canonical energy with the alluded properties.

Let us denote the Killing vector for the background Ψ as K. The modified canonical energy for an exact Killing vector K can be introduced through the second variation of the ADT current by[1]

$$\mathcal{E}(K;\delta_1\Psi,\delta_2\Psi)\equiv-\frac{1}{8\pi G}\int_{\Sigma}dx_{\mu}\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi), \tag{15}$$

which is different, at least apparently, from the definition given in Ref. [3]. Here, Σ denotes a Cauchy surface extending from the bifurcation surface B to the spacelike infinity. Nevertheless, the essential features for the canonical energy will be shown to be satisfied. To this purpose, let us look into the properties of the current expression $\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)$: symmetric form, conservation, gauge invariance, and the monotonicity. In the following, we take K as the horizon Killing vector whenever the choice is convenient to present.

A. Symmetric bilinear form

By using EOM, $\mathcal{E}_{\Psi}=0$ and linearized EOM, $\delta\mathcal{E}_{\Psi}=0$ with the interchangeability of two generic variations δ_1 and δ_2, such as $\delta_2\delta_1\Psi=\delta_1\delta_2\Psi$ and $\delta_2\delta_1\mathcal{E}_{\Psi}=\delta_1\delta_2\mathcal{E}_{\Psi}$, one can see that the second variations of the ADT current satisfy the relation as[2]

$$\begin{aligned}\delta_2(\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_1\Psi))|_{\text{on-shell}}&=(\delta_2\delta_1\mathbf{E}^{\mu}_{\nu})\zeta^{\nu}|_{\text{on-shell}}\\ &=\delta_1(\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_2\Psi))|_{\text{on-shell}},\end{aligned} \tag{16}$$

where we have used the definition of the (off-shell) ADT current given in Eq. (4). Because of the interchangeability of two generic variations, $\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta_2\delta_1\Psi)$ is also symmetric over two variations, δ_1 and δ_2. Thus, by using

[1]In the context of the stability of black holes or branes [3,5], we need to impose the axisymmetric condition for linear perturbations and the Killing vector K can be chosen as the stationary one $K_T=\frac{\partial}{\partial t}$ in the asymptotically flat case. In the asymptotically AdS case, we can take the Killing vector K as the horizon Killing vector K_H without imposing the axisymmetric conditions.

[2]Here, "on-shell" means that EOM $\mathcal{E}_{\Psi}=0$ and linearized EOM $\delta\mathcal{E}_{\Psi}=0$ are satisfied without requiring the second order variation of EOM to be satisfied.

the defining relation for $\mathcal{J}^\mu_{\rm ADT}(\zeta;\delta_2\Psi,\delta_1\Psi)$ in Eq. (14) and by using the "on-shell" vanishing property of the ADT current $\mathcal{J}^\mu_{\rm ADT}(\delta_2\zeta;\Psi,\delta_1\Psi)$ in Eq. (11), one can see that the current expression $\mathcal{J}^\mu_{\rm ADT}(\zeta;\delta_2\Psi,\delta_1\Psi)$ is a symmetric bilinear form on the space of on-shell first order variations for an arbitrary vector ζ^μ, while it is independent of the second order field variations.

B. Conservation

Recall that, for the exact Killing vector K for the arbitrary background Ψ with the arbitrary $\delta\Psi$, the additional current $\mathcal{J}^\mu_\Delta(\pounds_K\Psi,\delta\Psi)$ vanishes and so the off-shell current $\mathbf{J}^\mu(K;\Psi,\delta\Psi)$ reduces to the ADT current $\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)$. Thus, just by replacing the arbitrary $\delta\Psi$ with $\delta_2\delta_1\Psi$ in Eq. (5), one obtains

$$\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta_2\delta_1\Psi)=\partial_\nu(\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta_2\delta_1\Psi)),\tag{17}$$

which shows us that the last term in the right-hand side of Eq. (14) is identically conserved in the Killing vector case. The second term in the right-hand side of Eq. (14) vanishes identically when the on-shell condition is imposed. By using the Killing property of the background $\pounds_K\Psi=0$ and using Eq. (A4), one can see that the variation of the ADT current for the Killing vector K is also on-shell conserved:

$$\begin{aligned}&\partial_\mu[\delta_2(\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta_1\Psi))]_{\text{on-shell}}\\&\quad=\sqrt{-g}(\delta_2\delta_1\mathcal{E}_\Psi)\pounds_K\Psi|_{\text{on-shell}}=0,\end{aligned}\tag{18}$$

where the on-shell vanishing property of the current $\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)$ is also used. Note that the variation of the ADT current $\delta(\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi))$ is conserved without the requirement that the second order variation of field Ψ satisfy EOM.[3] Collecting the above conservation and/or vanishing properties of each term in the right-hand side in Eq. (14), one can see that the current expression $\mathcal{J}^\mu_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)$ is conserved at the on-shell level.

C. Gauge invariance

The previously established two properties of the current expression, $\mathcal{J}^\mu_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)$, motivate the introduction of the modified canonical energy $\mathcal{E}(K;\delta_1\Psi,\delta_2\Psi)$ for a Killing vector K in Eq. (15). To see the diffeomorphism transformation property of the current expression $\mathcal{J}^\mu_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)$ and the gauge invariance of the modified canonical energy, note that the Lie derivative of the ADT current can be written as

$$\begin{aligned}\pounds_\epsilon\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)=&\,\mathcal{J}^\mu_{\rm ADT}(\pounds_\epsilon K;\Psi,\delta\Psi)\\&+\mathcal{J}^\mu_{\rm ADT}(K;\pounds_\epsilon\Psi,\delta\Psi)\\&+\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\pounds_\epsilon\delta\Psi),\end{aligned}\tag{19}$$

where the first term in the right-hand side vanishes when the on-shell conditions are imposed because of the on-shell vanishing property of $\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi)$ for an arbitrary ζ^μ. One may note that the Lie derivative of the ADT current can also be written as

$$\begin{aligned}\pounds_\epsilon\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)=&\,\nabla_\nu(2\epsilon^{[\nu}\mathcal{J}^{\mu]}_{\rm ADT}(K;\Psi,\delta\Psi))\\&-\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)\nabla_\nu\epsilon^\nu\\&+\epsilon^\mu\nabla_\nu\mathcal{J}^\nu_{\rm ADT}(K;\Psi,\delta\Psi),\end{aligned}\tag{20}$$

where the second term in the right-hand side vanishes when the on-shell conditions are imposed and the last term vanishes because of the conservation property of the ADT current $\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)$ for the Killing vector K. Combining two expressions in Eq. (19) and (20), one can show that

$$\begin{aligned}-\int dx_\mu\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(K;\pounds_\epsilon\Psi,\delta\Psi)|_{\text{on-shell}}&=\int dx_\mu\sqrt{-g}[\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\pounds_\epsilon\delta\Psi)-\pounds_\epsilon\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\delta\Psi)]_{\text{on-shell}}\\&=\int dx_{\mu\nu}\sqrt{-g}[\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)+2\epsilon^{[\mu}\mathcal{J}^{\nu]}_{\rm ADT}(K;\Psi,\delta\Psi)]_{\text{on-shell}},\end{aligned}$$

where we have used $\mathcal{J}^\mu_{\rm ADT}(K;\Psi,\pounds_\epsilon\delta\Psi)=\nabla_\nu\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)$ for the Killing vector K. Note also that the last term in the second equality also vanishes by the on-shell vanishing condition of the ADT current $\mathcal{J}^\mu_{\rm ADT}$. In the end, one obtains

$$\mathcal{E}(K;\pounds_\epsilon\Psi,\delta\Psi)=\frac{1}{8\pi G}\int_{\partial\Sigma}dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)|_{\text{on-shell}},\tag{21}$$

which shows us that a generic diffeomorphism transformation leads to a total boundary term and the modified canonical energy becomes invariant for the local, compactly supported diffeomorphism parameter ϵ^μ on Σ.

More generically, ϵ^μ may not be local, compactly supported, which could be present in our setup. In pure

[3] When the second order variation of the field Ψ also satisfies EOM, the variation of the ADT current itself vanishes, i.e., $\delta(\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(\delta\Psi))=0$.

Einstein gravity, as was done in Ref. [3], one may choose gauges of the metric near the horizon such that the linear perturbation does not change the expansion of the bifurcation surface B as $\delta\vartheta|_B = 0$ and $\delta A = 0$ for the area A of the surface B. Concretely, on the near horizon, one can take the metric in the Gaussian null coordinates [3,20,21] generically as

$$ds^2_{\rm NH} = 2du(dr - r^2\alpha du - r\beta_a dx^a) + \mu_{ab}dx^a dx^b, \qquad (22)$$

where u and r correspond to affine parameters along the null directions and $\mu_{\alpha\beta}$ denotes the metric on the sphere part in Gaussian null coordinates. Here, two null coordinate vectors $n \equiv n^\mu\partial_\mu = \frac{\partial}{\partial u}$ and $\ell \equiv \ell^\mu\partial_\mu = \frac{\partial}{\partial r}$ commute and n^μ is normal to the horizon with the relation $n^\mu\ell_\mu = 1$. In the chosen gauge, the horizon Killing vector field for the background near the horizon is given by

$$K_H = \kappa\left(u\frac{\partial}{\partial u} - r\frac{\partial}{\partial r}\right), \qquad (23)$$

where κ denotes the surface gravity and the perturbed metric becomes

$$h_{\mu\nu}dx^\mu dx^\nu = -2r^2\delta\alpha du^2 - 2r\delta\beta_a dx^a du + \delta\mu_{ab}dx^a dx^b, \qquad (24)$$

where $h_{\mu\nu} \equiv \delta g_{\mu\nu}$. At infinity in the asymptotic flat case, the "Bondi gauge" is chosen with a further choice as was done by Geroch-Xanthopoulos [21,22], where the "unphysical metric" at infinity may be taken as

$$d\tilde{s}^2 = \tilde{\Omega}^2 ds^2 = 2d\tilde{\Omega}d\tilde{u} + \tilde{\mu}_{ab}d\tilde{x}^a d\tilde{x}^b + \mathcal{O}(\tilde{\Omega}), \qquad (25)$$

and the linearized metric satisfies $\tilde{g}^{\mu\nu}\tilde{h}_{\mu\nu} = \mathcal{O}(\tilde{\Omega})$. In the case of the asymptotically AdS boundary conditions, the unphysical metric is taken as [6,23]

$$d\tilde{s}^2 = d\tilde{\Omega}^2 + \tilde{\gamma}_{\mu\nu}dx^\mu dx^\nu + \mathcal{O}(\tilde{\Omega}^2), \qquad (26)$$

where $\tilde{\gamma}_{\mu\nu}$ denotes the metric on the boundary, i.e., the metric of the Einstein static universe.

And then, further conditions are imposed such that the charges $Q(\zeta)$ for the asymptotic symmetry generator ζ are not changed under linear perturbations. This imposition is related to the perturbation *toward stationary black holes*, which would generate unwanted contributions [3]. There are still remnant gauge transformations preserving the chosen gauge conditions, of which gauge parameter ϵ^μ becomes tangent to the horizon and corresponds to the asymptotic symmetry generators near infinity. For such a noncompact gauge parameter ϵ^μ, one can show the gauge invariance of the modified canonical energy by using the results given in Appendix B. To phrase it simply, under the same assumptions as in Ref. [3] on the horizon and asymptotic behavior of the gauge parameter ϵ, one can argue that

$$\begin{aligned}\mathcal{E}(K;\pounds_\epsilon\Psi,\delta\Psi) &= \frac{1}{8\pi G}\left(\int_\infty - \int_B\right)dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)|_{\text{on-shell}} \\ &= 0. \end{aligned} \qquad (27)$$

In summary, the modified canonical energy is also gauge invariant when the appropriate conditions are taken.

D. Monotonic property

To consider the monotonic property along the time evolution of the canonical energy, Hollands *et al.* evaluated the canonical energy on the four sectors[4] $\mathcal{I}(t_1) \cup \mathcal{I}(t_2) \cup \mathcal{H}_{12} \cup \mathcal{J}_{12}$ with $t_1 < t_2$, where $\mathcal{H}_{12}$ and $\mathcal{J}_{12}$ denote the regions in the future horizon and future null/spacelike infinity, respectively. In order to obtain rigorous statements about the behavior of the canonical energy, some machinery is utilized for taking care of the null infinity and the horizon. Briefly speaking, one needs to choose appropriate gauges and falloff conditions [3,5,6]. In the asymptotic flat spacetime, the null infinity should be managed carefully and the statement is proven only for the even-dimensional case. As was alluded before, K is taken as the stationary Killing vector K_T in this case. To the contrary, in the asymptotic AdS spacetime, the asymptotic infinity is timelike and flux cannot leak away. In this case, the contribution from the asymptotic infinity is trivial and there is no restriction on the dimensionality.

Instead of performing this analysis directly in our construction, we would like to relate our modified canonical energy to the original expression of the canonical energy given in [3] and borrow the monotonicity property of the HW canonical energy to show the monotonicity of the modified canonical energy in our construction. Recall that the symplectic form W_Σ in the covariant phase space is defined by

$$W_\Sigma(\Psi|\delta_1\Psi,\delta_2\Psi) \equiv \frac{1}{16\pi G}\int_\Sigma dx_\mu\omega^\mu(\Psi|\delta_1\Psi,\delta_2\Psi)|_{\text{on-shell}}, \qquad (28)$$

and that the HW canonical energy for the Killing vector K is defined by

$$\mathcal{E}_{\rm HW}(K;\delta_1\Psi,\delta_2\Psi) \equiv W_\Sigma(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi), \qquad (29)$$

where $\delta K^\mu = 0$ is assumed as before for simplicity. Therefore, the difference between our modified canonical energy and its expression in the HW canonical energy turns

[4]See the second figure in Ref. [3] and Fig. 2 in Ref. [5].

out to be just surface terms, as is shown in Appendix C. Explicitly, the difference is given by

$$\begin{aligned}\mathcal{E}_{\rm HW}(K;\delta\Psi,\delta\Psi)-\mathcal{E}(K;\delta\Psi,\delta\Psi)\\ =\frac{1}{16\pi G}\int_{\partial\Sigma}dx_{\mu\nu}[2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\delta\Psi,\delta\Psi)\\ -\sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_K\delta\Psi,\delta\Psi)]_{\text{on-shell}},\end{aligned} \tag{30}$$

where the boundary $\partial\Sigma$ is composed of two parts, $\int_{\partial\Sigma}=\int_\infty-\int_B$. Basically, these boundaries are sphere parts of geometry on the horizon and the infinity. The contribution from the infinity vanishes as can be inferred from the linearized charge expression in Eq. (6), which is finite and taken to vanish for linear perturbations. In pure Einstein gravity, one can infer from Eq. (C3) that the above boundary term at $B(t)$ is given by[5]

$$\begin{aligned}\int_{B(t)}&\equiv\frac{1}{8\pi G}\int_{B(t)}dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K_H;\delta g,\delta g)|_{\text{on-shell}}\\ &=\frac{1}{8\pi G}\int_{B(t)}dx\sqrt{\mu}\left[\kappa\delta\mu^{\alpha\beta}\delta\mu_{\alpha\beta}-\frac{3}{2}\delta\mu^{\alpha\beta}\pounds_{K_H}\delta\mu_{\alpha\beta}\right]_{\text{on-shell}}.\end{aligned} \tag{31}$$

Note that the second term in the right-hand side of the last equality vanishes on the bifurcation surface $B(t=0)$ because of $K_H\to 0$, while it would vanish at $B(t)$ when the perturbed shear vanishes as was argued in Ref. [3].

The absence of the contribution from the infinity implies that there would no difference between $\mathcal{E}$ and $\mathcal{E}_{\rm HW}$ on the sector $\mathcal{J}_{12}$, while they may be different on other sectors $\mathcal{I}(t_1),\mathcal{I}(t_2)$, and $\mathcal{H}_{12}$ up to boundary terms given by the integral over $B(t_1)$ and $B(t_2)$. Schematically, the difference on these sectors can be written as

$$\begin{aligned}\mathcal{E}_{\rm HW}|_{\mathcal{I}(t_{1,2})}&=\mathcal{E}|_{\mathcal{I}(t_{1,2})}-\int_{B(t_{1,2})},\\ \mathcal{E}_{\rm HW}|_{\mathcal{H}_{12}}&=\mathcal{E}|_{\mathcal{H}_{12}}+\int_{B(t_2)}-\int_{B(t_1)},\end{aligned} \tag{32}$$

where we have abused the notation to denote the symplectic form and its counterpart on the sector $\mathcal{H}_{12}$ as $\mathcal{E}_{\rm HW}$ and $\mathcal{E}$, respectively. This is consistent with the individual conservation of $\mathcal{E}_{\rm HW}$ and $\mathcal{E}$.

To show the monotonicity of the canonical energy in Ref. [3], the appropriate boundary terms are subtracted from the canonical energy to modify $\mathcal{E}_{\rm HW}$ as $\bar{\mathcal{E}}_{\rm HW}$ and then it was shown that this barred canonical energy satisfies for $t_1\leq t_2$

$$\bar{\mathcal{E}}_{\rm HW}(t_2)-\bar{\mathcal{E}}_{\rm HW}(t_1)\leq 0. \tag{33}$$

Simply by defining our barred quantity as $\bar{\mathcal{E}}(t)=\bar{\mathcal{E}}_{\rm HW}(t)$, one can establish immediately the monotonicity property of our barred canonical energy. The actual application of the canonical energy comes from the fact that $\bar{\mathcal{E}}_{\rm HW}(t)\to\mathcal{E}_{\rm HW}(t=0)$ as $t\to 0$, which can be used to argue that $\mathcal{E}_{\rm HW}(t=0)<0$ implies the instability.

In order to see how to use our canonical energy for the stability, it is useful to introduce the boundary term composed solely of the first term in the last equality in Eq. (31) by denoting it as $\int'_{B(t)}$, which is negative semidefinite. Note that our barred canonical energy is different from the unbarred one even at $t=0$:

$$\bar{\mathcal{E}}=\bar{\mathcal{E}}_{\rm HW}=\mathcal{E}_{\rm HW}=\mathcal{E}-\int'_B, \tag{34}$$

where we have used the relation between $\mathcal{E}_{\rm HW}$ and $\mathcal{E}$ and the fact that $\bar{\mathcal{E}}_{\rm HW}=\mathcal{E}_{\rm HW}$ at $t=0$. Because of this relation, $\mathcal{E}(t=0)<0$ may not imply the instability for a generic noncompact perturbation. However, we may prepare the initial data at $t=t_1$, which are compactly supported such that $\mathcal{E}(t_1)=\bar{\mathcal{E}}(t_1)=\bar{\mathcal{E}}_{\rm HW}(t_1)$. Then, at later time we can conclude that

$$\bar{\mathcal{E}}(t_2)=\bar{\mathcal{E}}_{\rm HW}(t_2)\leq\bar{\mathcal{E}}_{\rm HW}(t_1)=\bar{\mathcal{E}}(t_1)=\mathcal{E}(t_1), \tag{35}$$

Thus, for the compactly supported initial data, the instability argument works in our modified canonical energy: $\mathcal{E}(t=0)<0$ implies the instability just as $\mathcal{E}_{\rm HW}(t=0)<0$ does.

Now, we would like to give various comments on our construction and its meaning. In our construction the symmetric bilinear property of the modified canonical energy on the first order variations is manifest even on $\Sigma(t)$, in contrast to the expression of the canonical energy in the HW construction. In the HW construction, the appropriate gauges at the asymptotic infinity and the horizon are used to show this property on $\Sigma(t=0)$, only. As can be inferred from Eq. (14) and the property of the ADT current, our modified canonical energy can also be written, when the second order variation of field Ψ satisfies second order EOM, as

$$\begin{aligned}&\mathcal{E}(K;\delta\Psi,\delta\Psi)\\ &\quad=\frac{1}{8\pi G}\int_{\partial\Sigma}dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta^2\Psi)|_{\mathcal{E}_\Psi=\delta\mathcal{E}_\Psi=\delta^2\mathcal{E}_\Psi=0},\end{aligned} \tag{36}$$

which means that the bulk expression of the modified canonical energy expression becomes the total surface term. More concretely, the second order variation of conserved charges in our construction can be read from Eq. (6) as

[5]The similar expression at $B(t=0)$ was obtained in Eq. (85) in Ref. [3].

$$\delta^2 Q(K) = \frac{1}{8\pi G}\int_{\partial\Sigma} dx_{\mu\nu}\delta[\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta\Psi)]$$
$$= \frac{1}{8\pi G}\int_{\partial\Sigma} dx_{\mu\nu}\sqrt{-g}[\mathbf{Q}^{\mu\nu}(K;\delta\Psi,\delta\Psi) + \mathbf{Q}^{\mu\nu}(K;\Psi,\delta^2\Psi)], \quad (37)$$

where we used $\delta K = 0$ and $\delta\sqrt{-g} = 0$ at $\partial\Sigma$. By using the second order perturbation satisfying EOM and falloff conditions of $\delta\Psi$, one can see that

$$\mathcal{E}(K_H;\delta\Psi,\delta\Psi) = \delta^2 M_\infty - \Omega_H\delta^2 J_\infty - \frac{\kappa}{2\pi}\delta^2\mathcal{S}_{\mathrm{BH}} - \frac{1}{8\pi G}\int_B dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K_H;\delta\Psi,\delta\Psi), \quad (38)$$

where the last term in pure Einstein gravity is given by Eq. (C3).

On the sector $\mathcal{J}_{12}$ in Einstein gravity without matter fields, one may obtain directly the expression of our modified canonical energy, which is nothing but the second order Einstein tensor in this case, by choosing the Geroch-Xanthopoulos gauge with additional falloff conditions as in [24]. The final expression given in Eq. (4.13) in Ref. [24] shows us that our modified energy also gives the same expression as the original canonical energy on this sector. Therefore, there is no difference between $\mathcal{E}_{\mathrm{HW}}$ and $\mathcal{E}$ on the sector $\mathcal{J}_{12}$ in Einstein gravity without matter fields, indeed. This direct computation reinforces our argument for the absence of the difference at the infinity between our modified canonical energy and the HW canonical energy. One may perform the similar direct computation on $\mathcal{H}_{12}$, since the expressions on $\mathcal{H}_{12}$ and $\mathcal{J}_{12}$ may be parallel. In the above, the difference is indirectly shown to reside only on the surface $B(t_{1,2})$ in Appendix C.

Though our expression of linearized conserved charges is completely consistent with the one in the covariant phase space approach and the one from the conventional ADT formalism, the second order variation of conserved charges in our construction may be different from the one in the covariant phase space approach since

$$\delta_2\delta_1 Q(K) = \delta_2\delta_1 Q_{\mathrm{cov}}(K) + \frac{1}{16\pi G}\int_{\partial\Sigma} dx_{\mu\nu}\sqrt{-g}\mathbf{A}^{\mu\nu}(\pounds_K\delta_2\Psi,\delta_1\Psi), \quad (39)$$

where Q_{cov} denotes the charge in the covariant phase space approach and we have used the relation given in Eq. (A8) with the condition $\delta K^\mu = 0$. Indeed, in some higher derivative gravity it was noticed that the additional contribution to the covariant phase space charge expression is important in the context of the Kerr/CFT correspondence [25]. The combination in the second order variations in the canonical energy may also be affected by this difference. This is reflected in the following representation of the HW canonical energy,

$$\mathcal{E}_{\mathrm{HW}}(K;\delta\Psi,\delta\Psi) = \delta^2 M^{\mathrm{cov}}_\infty - \Omega_H\delta^2 J^{\mathrm{cov}}_\infty - \frac{\kappa}{2\pi}\delta^2\mathcal{S}^{\mathrm{cov}}_{\mathrm{BH}},$$

which is consistent with the difference given in Eq. (30) and Eq. (38).

Practically, the modified canonical energy can be obtained simply by keeping the first order variation terms in the expression of $\delta_2\delta_1\mathbf{E}^{\mu\nu}$. By using EOM $\mathcal{E}_\Psi = 0$ and linearized EOM $\delta\mathcal{E}_\Psi = 0$, respectively, one can see that

$$\delta_2(\sqrt{-g}\mathcal{J}^\mu_{\mathrm{ADT}}(K;\Psi,\delta_1\Psi)) = \sqrt{-g}(\delta_2\delta_1\mathbf{E}^{\mu\nu})K_\nu|_{\text{on-shell}}. \quad (40)$$

As a result, the modified canonical energy is given by

$$\mathcal{E}(K;\delta_1\Psi,\delta_2\Psi) \equiv -\frac{1}{8\pi G}\int_\Sigma dx_\mu\sqrt{-g}(\delta_2\delta_1\mathbf{E}^{\mu\nu})K_\nu|^{\delta^2\Psi=0}_{\text{on-shell}}, \quad (41)$$

where $\delta^2\Psi = 0$ denotes that we should keep the first order variation terms. This expression clearly shows us that our modified canonical energy is related to the direct generalization of the second order Einstein tensor. Our construction may be regarded as providing the generalization of the construction by the second order Einstein tensor in Ref. [24] beyond Einstein gravity. Furthermore, our construction shows clearly the dependence of the modified canonical energy only on EOM and it makes the covariance of the expression manifest. It would be useful to deal with the odd-dimensional case with the gravitational Chern-Simons term. Besides, our expression of the canonical energy clarifies the relation between the traditional ADT expression and the canonical energy by Hollands and Wald.

One may worry that our form of the canonical energy might have some drawbacks compared to the HW construction since it differs from the HW canonical energy in the boundary term and the HW canonical energy is shown to be consistent with other criteria of the black hole stability. Nevertheless, as was shown in the above, all the essential properties of the canonical energy also hold in our modified version and there are various cases in which such a boundary term does not contribute. In fact, our construction reveals the interesting aspect that the additional boundary term at the horizon may be allowed in defining the canonical energy. Namely, when we define the canonical energy, we may relax its relation to the Hessian in thermodynamic stability, up to the boundary term. More concretely, we may have defined the modified canonical energy by adding the boundary term coming from $\mathbf{J}^\mu(K;\delta\Psi,\delta\Psi)$. For instance, we may have introduced the modified canonical energy by using the current expression $-\mathcal{J}^\mu_{\mathrm{ADT}} + \mathbf{J}^\mu = \mathcal{J}^\mu_\Delta$ as

$$\tilde{\mathcal{E}}(K;\delta\Psi,\delta\Psi) \equiv \frac{1}{8\pi G}\int_\Sigma dx_\mu \sqrt{-g}\mathcal{J}^\mu_\Delta(\Psi|,\pounds_K\delta\Psi,\delta\Psi),$$

which also satisfies all the properties discussed in the above. This canonical energy can be shown to be different from the HW canonical energy as

$$\begin{aligned}&\mathcal{E}_{\rm HW}(K;\delta\Psi,\delta\Psi) - \tilde{\mathcal{E}}(K;\delta\Psi,\delta\Psi)\\ &\quad = \int_B dx_{\mu\nu}\sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_K\delta\Psi,\delta\Psi)|_{\text{on-shell}},\end{aligned}$$

and satisfies the relation, when the second order EOM are imposed, as

$$\tilde{\mathcal{E}}(K_H;\delta\Psi,\delta\Psi) = \delta^2 M_\infty - \Omega_H\delta^2 J_\infty - \frac{\kappa}{2\pi}\delta^2\mathcal{S}_{\rm BH}, \quad (42)$$

which is consistent even with the Hessian in thermodynamic stability consideration.

Recently, there was a suggestion that the canonical energy is dual to the so-called Fisher quantum information metric in the context of the AdS/CFT correspondence [8]. As is clear from our construction, our modified canonical energy $\mathcal{E}$ or $\tilde{\mathcal{E}}$ is also a good candidate like those dual to the information metric, since our modified canonical energy does not give any difference from the HW canonical energy on the pure AdS background. The difference between them comes from the boundary contribution at the bifurcation surface B or more correctly at $B(t)$, which is related to the deep infrared physics in the boundary theory. The freedom of adding the boundary term $\int_B$ to the canonical energy with arbitrary coefficients may be useful in this duality.

In the following section, we consider hairy black holes in the asymptotic AdS space. In the asymptotic AdS space, the roles of boundary terms at infinity are irrelevant because of the AdS nature and one can take K as the horizon Killing vector. We apply our modified canonical energy to study the stability issue on hairy extremally rotating black holes.

IV. HAIRY ADS BLACK HOLES

In this section we consider three-dimensional extremally rotating hairy AdS black holes admitted in Einstein gravity with a cosmological constant and a scalar field, whose analytic solutions are given in [26–28]. Interestingly, there are two arguments for the stability of the above extremally rotating hairy black holes that could give us opposite conclusions. The argument for their stability may be given as follows. Since there are no propagating degrees of freedom in three-dimensional Einstein gravity and the scalar field involved in the above solutions satisfies the Breitenlohner-Freedman bound [29], the extremal hairy black holes should be stable, at least, perturbatively. Moreover, there seems to be no mechanism for the instability in this extremal configuration in the AdS/CFT context since it is dual to the renormalization group flow interpolating two CFTs, which does not seem to allow the other end points. The opposite argument comes from the no-hair conjecture for AdS black holes [30,31], which was made only for the four-dimensional case but seems to hold even in the three-dimensional case. Though there is a numerical attempt to construct rotating hairy black holes deformed from BTZ black holes [32], those hairy black holes require special conditions on the asymptotic behavior of the scalar field [33] that are not satisfied by the hairy extremal black holes under consideration. Furthermore, the extremally rotating black holes in higher than four dimensions are shown to be unstable [5,34]. Of course, all the opposite arguments rely on the higher-dimensional analogues and so may not be so persuasive. In the following we adopt the canonical energy method and show the stability of three-dimensional extremally rotating hairy black holes.

Before presenting the specific models under consideration, let us present some general setup for the Einstein gravity with $U(1)$ gauge and scalar fields φ^I and summarize some results. The Lagrangian for this system consists of three parts, the Einsten-Hilbert one $\mathcal{L}_{\rm EH}$, the scalar one, and the $U(1)$ gauge part, respectively, as

$$\begin{aligned}&\mathcal{L}_{\rm EH} = R - 2\Lambda, \qquad \mathcal{L}_\varphi = -\frac{1}{2}G_{IJ}\partial_\mu\varphi^I\partial^\mu\varphi^J - V(\varphi),\\ &\mathcal{L}_A = -\frac{1}{4}\mathcal{N}(\varphi)F_{\mu\nu}F^{\mu\nu}. \end{aligned} \quad (43)$$

The Euler-Lagrange expressions for metric, gauge and scalar fields are given by

$$\begin{aligned}&\mathcal{E}_{\mu\nu} = \mathcal{G}^\Lambda_{\mu\nu} - T_{\mu\nu}, \qquad \mathcal{E}^\mu_A = \nabla_\mu(\mathcal{N}F^{\mu\nu}),\\ &\mathcal{E}_\varphi = G_{IJ}(\varphi)(\Box\varphi^J + \Gamma^J_{KL}\partial_\mu\varphi^K\partial^\mu\varphi^L) - \partial_{\varphi^I}V(\varphi)\\ &\qquad - \frac{1}{4}\partial_{\varphi^I}\mathcal{N}F_{\mu\nu}F^{\mu\nu},\end{aligned}$$

where $\mathcal{G}^\Lambda_{\mu\nu} = R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu}$ and the energy-momentum tensor $T_{\mu\nu}$ is composed of $T^\varphi_{\mu\nu}$ and $T^A_{\mu\nu}$ as

$$\begin{aligned}T^\varphi_{\mu\nu} &= \frac{1}{2}[G_{IJ}\partial_\mu\varphi^I\partial_\nu\varphi^J + g_{\mu\nu}\mathcal{L}_\varphi],\\ T^A_{\mu\nu} &= \frac{1}{2}[\mathcal{N}F_{\mu\alpha}F_\nu{}^\alpha + g_{\mu\nu}\mathcal{L}_A].\end{aligned}$$

Note that there is the off-shell identity for a generic diffeomorphism parameter ζ as

$$\begin{aligned}0 &= -2T^{\mu\nu}_A\zeta_\nu + \zeta^\mu\mathcal{L}_A - \mathbf{\Theta}^\mu_A(\pounds'_\zeta A),\\ 0 &= -2T^{\mu\nu}_\varphi\zeta_\nu + \zeta^\mu\mathcal{L}_\varphi - \mathbf{\Theta}^\mu(\pounds_\zeta\varphi),\end{aligned}$$

where $\pounds'_\zeta A = -F_{\mu\nu}\zeta^\nu$ denotes the Lie derivative augmented by a gauge transformation and the surface terms for the generic variations are given by

$$\Theta^\mu_g(\delta g) = 2g^{\alpha[\mu}\nabla^{\beta]}\delta g_{\alpha\beta}, \qquad \Theta^\mu(\delta\varphi) = -G_{IJ}(\varphi)\delta\varphi^I\partial^\mu\varphi^J,$$
$$\Theta^\mu_A = -\mathcal{N}F^{\mu\nu}\delta A_\nu.$$

By using this identity under the assumption $\delta K^\mu = 0$, one can see that the scalar field and the Abelian gauge field parts for the modified canonical energy can be extracted from the surface term Θ^μ as

$$\begin{aligned}&\mathcal{J}^\mu_{\rm ADT}(K;\delta\varphi,\delta\varphi)|_{\text{on-shell}}\\ &\quad= \frac{1}{2}[-K^\mu\nabla_\nu\delta_\varphi\Theta^\nu_\varphi(\delta\varphi) + \delta^2_\varphi\Theta^\mu(\pounds_K\varphi)]|^{\delta^2\varphi=0}_{\text{on-shell}},\\ &\mathcal{J}^\mu_{\rm ADT}(K;\delta A,\delta A)|_{\text{on-shell}}\\ &\quad= \frac{1}{2}[-K^\mu\nabla_\nu\delta_A\Theta^\nu_A(\delta A) + \delta^2_A\Theta^\mu(\pounds_K A)]|^{\delta^2 A=0}_{\text{on-shell}}.\end{aligned}$$

As was emphasized before, one can obtain the same expression solely from the expression of $\mathbf{E}^{\mu\nu}$, i.e. a combination of EOM, but we have provided the shortcut to the results by using the relation between EOM and the surface term Θ^μ.

Now, let us stick to the three-dimensional Einstein gravity with a minimally coupled scalar field, whose Lagrangian can be written as

$$\mathcal{L} = R - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi - V(\varphi). \tag{44}$$

Our interest is in the hairy deformed three-dimensional extremal black holes [26,27]. One can obtain the solutions by assuming that the scalar potential V is taken in the form of

$$V(\varphi) = \frac{1}{2L^2}(\partial_\varphi\mathcal{W})^2 - \frac{1}{2L^2}\mathcal{W}^2, \qquad \mathcal{W} = \mathcal{W}(\varphi). \tag{45}$$

By taking the generic ansatz for the metric and scalar as

$$\begin{aligned}&ds^2 = -e^{2A(r)}dt^2 + e^{2B(r)}dr^2 + r^2(d\theta + e^{C(r)}dt)^2,\\ &\varphi = \varphi(r),\end{aligned} \tag{46}$$

where the radius of the asymptotic AdS_3 space is taken to be unity, one can show that the metric functions and the scalar field satisfying the following first order ordinary differential equations solve the full EOM:

$$\begin{aligned}&\varphi' = -e^B\partial_\varphi\mathcal{W}, \qquad A' = \frac{1}{r} + e^B\mathcal{W},\\ &(e^C)' = \left(\frac{1}{r}e^A\right)', \qquad A' + B' = \frac{r}{2}\varphi'^2.\end{aligned} \tag{47}$$

For instance, the simplest case among analytic solutions is

$$\varphi(r) = \frac{\varphi_0}{r^2}, \qquad \mathcal{W} = \alpha[4 + \varphi^2(r)] + \beta e^{\frac{\varphi^2}{4}}, \tag{48}$$

where the coefficients α and β are given, in terms of the constant φ_0 and the position of the horizon r_H, by

$$\alpha = \frac{1}{2}\frac{1}{1 - e^{-\varphi_0^2/4r_H^2}}, \qquad \beta = -\frac{2e^{-\varphi_0^2/4r_H^2}}{1 - e^{-\varphi_0^2/4r_H^2}}. \tag{49}$$

In this case, the metric functions can be obtained as

$$\begin{aligned}&e^A = r\left[2\alpha e^{-\varphi_0^2/4r^2} + \frac{\beta}{2}\right], \qquad e^B = e^{-\varphi_0^2/4r^2}e^{-A},\\ &e^C = \frac{1}{r}e^A.\end{aligned} \tag{50}$$

The near horizon geometry of all these configurations satisfying the first order EOM is given generically by

$$ds^2_{\rm NH} = L^2_{\rm NH}\left[-\rho^2dt^2 + \frac{1}{\rho^2}d\rho^2\right] + r_H^2\left(d\theta - \frac{L_{\rm NH}}{r_H}\rho dt\right)^2, \tag{51}$$

which is known as the self-dual orbifold of AdS_3 space [35]. Here, $L_{\rm NH}$ denotes the radius of the orbifold of AdS_3 space defined by

$$L_{\rm NH} = \frac{1}{\mathcal{W}(\varphi_H)}.$$

The scalar potential near the horizon can be expanded as

$$V = -\frac{1}{2}\mathcal{W}(\varphi_H)^2 + \mathcal{W}(\varphi_H)^2(\varphi - \varphi_H)^2 + \cdots, \tag{52}$$

which comes from the generic expansion of the superpotential $\mathcal{W}$ as

$$\mathcal{W} = \mathcal{W}(\varphi_H) - \frac{1}{2}\mathcal{W}(\varphi_H)(\varphi - \varphi_H)^2 + \cdots. \tag{53}$$

Note that the first term of the superpotential $\mathcal{W}$ plays the role of the cosmological constant on the near horizon geometry. The horizon Killing vector is taken by $K = \frac{\partial}{\partial t}$ in these coordinates. The effective Lagrangian on the near horizon geometry is given by

$$\mathcal{L} = R - 2\Lambda_{\rm NH} - \frac{1}{2}(\partial_\mu\tilde{\varphi})^2 - \frac{1}{2}m^2_{\rm NH}(\tilde{\varphi})^2 + \cdots, \tag{54}$$

where $\Lambda_{\rm NH}$ and $m^2_{\rm NH}$ denote the near horizon effective cosmological constant and effective mass square of the scalar field $\tilde{\varphi} \equiv \varphi - \varphi_H$ as

$$\Lambda_{\rm NH} \equiv -\frac{1}{4L^2_{\rm NH}} = -\frac{\mathcal{W}(\varphi_H)^2}{4}, \qquad m^2_{\rm NH} = 2\mathcal{W}(\varphi_H)^2 = \frac{2}{L^2_{\rm NH}}.$$

The effective mass of the scalar field is greater than the Breitenlohner-Freedman bound in the near horizon geometry and so the scalar field could be thought of as stable in the near horizon geometry. We would like to confirm this argument explicitly by using the canonical energy method.

Let us consider linear perturbations of the metric and scalar fields and compute the modified canonical energy on the near horizon geometry in order to see the stability of the above hairy deformed extremal black holes. Combined with the stability argument at infinity, one may say that the whole configuration is stable for linear perturbations. As is obvious from the three-dimensional nature of our configurations, the metric variation should be just pure gauge and so its role is trivial. From now on, let us take all the functions to depend on the radial coordinate ρ instead of r on the near horizon geometry. Indeed, by taking the metric perturbation as

$$\delta g_{\mu\nu} = \nabla_{(\mu}\zeta_{\nu)}, \qquad \zeta^\mu = \zeta^\mu(t,\rho,\theta), \tag{55}$$

we will show that the metric perturbation does not contribute to the canonical energy.

Instead of solving the linearized EOM on the given background, we would like to analyze the form of modified canonical energy itself, which is composed of four parts as

$$\mathcal{E}(K;\delta_1\Psi,\delta_2\Psi) = \mathcal{E}(K;\delta_1 g,\delta_2 g) + \mathcal{E}(K;\delta_1 g,\delta_2\varphi) + \mathcal{E}(K;\delta_1\varphi,\delta_2 g) + \mathcal{E}(K;\delta_1\varphi,\delta_2\varphi). \tag{56}$$

By using our result given in Eq. (41), one can compute each term directly without difficulty. Firstly, the contribution from the metric perturbation to the canonical energy is given by[6]

$$\mathcal{E}(K;\delta g,\delta g) = -\frac{1}{8\pi G}\int_\Sigma dx_\mu\sqrt{-g}\bigg[\bigg[-h^{\rho\sigma}\nabla^\mu\nabla^\nu h_{\rho\sigma} + 2h^{\rho\sigma}\nabla_\rho\nabla^{(\mu}h^{\nu)}{}_\sigma - \frac{1}{2}\nabla^\mu h^{\rho\sigma}\nabla^\nu h_{\rho\sigma} - 2\nabla^\rho h^{\sigma\mu}\nabla_{[\rho}h_{\sigma]}{}^\nu - \nabla_\rho(h^{\rho\sigma}\nabla_\sigma h^{\mu\nu}) + \frac{1}{2}\nabla^\rho h\nabla_\rho h^{\mu\nu} + 2\left(\nabla_\rho h^{\rho\sigma} - \frac{1}{2}\nabla^\sigma h\right)\nabla^{(\mu}h^{\nu)}{}_\sigma\bigg] - \frac{1}{2}g^{\mu\nu}[\mathrm{trace}]\bigg]K_\nu\bigg|^{\delta^2\Psi=0}_{\text{on-shell}}, \tag{57}$$

where $h_{\mu\nu} \equiv \delta g_{\mu\nu}$, $h \equiv g^{\mu\nu}h_{\mu\nu}$, and $[trace]$ denotes the trace of the expression in front of it. This part is consistent with Eq. (85) in [3]. Since the metric perturbation is given by a pure gauge transformation, i.e., $\delta g_{\mu\nu} = \pounds_\zeta g_{\mu\nu}$, this part has nothing to do with canonical energy and can be checked to vanish by a direct computation, as was shown generically in Eq. (27). The cross terms can be shown to vanish as follows:

$$\begin{aligned}&\mathcal{E}(K;\delta g,\delta\varphi) + \mathcal{E}(K;\delta\varphi,\delta g)\\ &\quad= \frac{1}{8\pi G}\int_\Sigma dx_\mu\sqrt{-g}[m_H^2(2h_{\mu\nu} - g_{\mu\nu}h)\varphi\delta\varphi]K_\nu|^{\delta^2\Psi=0}_{\text{on-shell}}\\ &\quad= 0.\end{aligned} \tag{58}$$

Hence, the contribution to canonical energy comes only from the scalar perturbation part as

$$\begin{aligned}\mathcal{E}(K;\delta\varphi,\delta\varphi) &= -\frac{1}{8\pi G}\int_\Sigma dx_\mu\sqrt{-g}\left[-\nabla^\mu\delta\varphi\nabla^\nu\delta\varphi + \frac{1}{2}g^{\mu\nu}(\nabla_\lambda\delta\varphi\nabla^\lambda\delta\varphi + m_H^2\delta\varphi^2)\right]K_\nu\bigg|^{\delta^2\Psi=0}_{\text{on-shell}}\\ &= \frac{1}{8\pi G}\int_\Sigma d\rho d\theta\frac{r_{\mathrm{NH}}}{2\rho^2}\left[2\rho^2\delta\varphi^2 + \rho^4\left(\frac{\partial\delta\varphi}{\partial\rho}\right)^2 + \left(\frac{\partial\delta\varphi}{\partial t}\right)^2\right].\end{aligned} \tag{59}$$

Since $r_{\mathrm{NH}} > 0$, $\mathcal{E}(K;\delta\varphi,\delta\varphi)$ could not be negative at any time. This confirms the linear stability of the extremally rotating hairy black holes under consideration.

V. CONCLUSION

We have constructed the modified version of the canonical energy that was introduced originally by HW in [3]. Our construction is based on the off-shell adaptation of the ADT current and so connects the various conceptually different constructions. Briefly speaking, it can be regarded as the generalization of the second order Einstein tensor method in pure Einstein gravity or the effective energy-momentum tensor method in the original ADT approach. By showing explicitly the relation between our construction and the original HW one, we have showed that one may construct a quantity that differs from the HW canonical energy in the boundary term over the spatial section of the future horizon. Through this relation, we have also explained clearly why the second order Einstein tensor method in the literature could give the same information as the HW canonical energy at the asymptotic infinity. In other words, our results imply

[6]It would be meaningful to check the gauge invariance explicitly since the gauge choice may be different in the extremal case and in our form of the near horizon geometry.

that the second order contribution to the Bondi energy can be computed by using the ADT current expression.

In fact, the modified canonical energy can be constructed while sharing all the properties of the HW canonical energy as given in Eq. (42) and may be distinguished from the HW canonical energy only in the higher derivative theory of gravity. The essential point of our construction is that one may have freedom in constructing the *canonical energy* equipped with the relevant properties. This possibility would give us a better chance to match the canonical energy to the Fisher information metric in the context of the AdS/CFT correspondence. Our results show that one may be able to use the freedom in the construction of the canonical energy with the required properties under consideration.

We have also considered the three-dimensional extremally rotating hairy AdS black holes that were not yet proven to be stable or not. Since there are conflicting arguments about their stability, it would be a good exercise to use the (modified) canonical energy method in this example, as is done in the main text. We have verified that the canonical energy is positive definite on the near horizon geometry and concluded that the extremally rotating hairy black holes in three dimensions are stable at least under the linear perturbations. It would be very interesting to explore whether or not one can distinguish the various possible forms of the canonical energy from the physical consideration in the context of the AdS/CFT correspondence, especially as a dual to the Fisher information metric. In the context of the black hole stability, it would also be interesting to consider a higher derivative theory of gravity and the stability of its black hole solutions by using the (modified) canonical energy method.

ACKNOWLEDGMENTS

We would like to thank Nakwoo Kim for some discussion. S. H. was supported by the National Research Foundation of Korea (NRF) Grant No. NRF-2013R1A1A2011548. S. Y. was supported by the National Research Foundation of Korea (NRF) Grant No. NRF-2015R1D1A1A09057057. We are grateful to CERN for hospitality, and to CERN-Korea theory collaboration committee for supporting the workshop “CERN-CKC TH Institute on Duality Symmetries in String and M-Theories”.

APPENDIX A: RELATION TO SYMPLECTIC CURRENT

Generic variation of the action can be expressed as

$$\delta I[\Psi] = \frac{1}{16\pi G}\int d^D x \delta(\sqrt{-g}\mathcal{L}) = \frac{1}{16\pi G}\int d^D x[\sqrt{-g}\mathcal{E}_\Psi \delta\Psi + \partial_\mu \Theta^\mu(\Psi,\delta\Psi)]. \quad (A1)$$

Identifying the diffeomorphism transformation of the parameter ζ with the above generic variation leads to the following relation:

$$\partial_\mu(\zeta^\mu\sqrt{-g}\mathcal{L}) = \sqrt{-g}\mathcal{E}_\Psi \pounds_\zeta \Psi + \partial_\mu \Theta^\mu(\Psi|\pounds_\zeta\Psi). \quad (A2)$$

The symplectic current with the variation of the diffeomorphism parameter ζ^μ under a generic variation δ can be defined by

$$\begin{aligned}\omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi) &\equiv \pounds_\zeta\Theta^\mu(\Psi,\delta\Psi) \\ &\quad - [\delta\{\Theta^\mu(\Psi,\pounds_\zeta\Psi)\} - \Theta^\mu(\Psi,\pounds_{\delta\zeta}\Psi)] \\ &= \Theta^\mu(\pounds_\zeta\Psi,\delta\Psi) - \Theta^\mu(\delta\Psi,\pounds_\zeta\Psi),\end{aligned}$$

which reduces to the conventional one when $\delta\zeta^\mu = 0$. By combining the definition of the symplectic current $\omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi)$, the double variation of the action as $(\delta\pounds_\zeta - \pounds_\zeta\delta)I[\Psi] = \pounds_{\delta\zeta}I[\Psi]$, and the relation in (A2) for the diffeomorphism parameter $\delta\zeta^\mu$, one obtains

$$\begin{aligned}\partial_\mu\omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi) &= [\delta(\sqrt{-g}\mathcal{E}_\Psi\pounds_\zeta\Psi) - (\sqrt{-g}\mathcal{E}_\Psi\pounds_{\delta\zeta}\Psi)] \\ &\quad - \pounds_\zeta(\sqrt{-g}\mathcal{E}_\Psi\delta\Psi). \quad (A3)\end{aligned}$$

By using the identity in Eq. (1), the relation of $\delta\partial_\mu = \partial_\mu\delta$ for a generic variation δ, and the property of the Lie derivative on the scalar density $\pounds_\zeta(\sqrt{-g}\mathcal{E}_\Psi\delta\Psi) = \partial_\mu(\zeta^\mu\sqrt{-g}\mathcal{E}_\Psi\delta\Psi)$, one can see that

$$\begin{aligned}\partial_\mu(\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}) &= -\frac{1}{2}[\delta(\sqrt{-g}\mathcal{E}_\Psi\pounds_\zeta\Psi) - \sqrt{-g}\mathcal{E}_\Psi\pounds_{\delta\zeta}\Psi] \\ &\quad + \frac{1}{2}\pounds_\zeta(\sqrt{-g}\mathcal{E}_\Psi\delta\Psi). \quad (A4)\end{aligned}$$

The additional current $\mathcal{J}^\mu_\Delta$ can be related to the symplectic current ω^μ in the covariant phase space [19] as

$$\begin{aligned}2\sqrt{-g}\mathcal{J}^\mu_\Delta(\Psi|\pounds_\zeta\Psi,\delta\Psi) &= \omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi) \\ &\quad + \partial_\nu(\sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi)), \quad (A5)\end{aligned}$$

where $\mathbf{A}^{\mu\nu}$ is an antisymmetric tensor defined by

$$\begin{aligned}\delta\Theta^\mu(\pounds_\zeta\Psi) &= \pounds_\zeta\Theta^\mu(\delta\Psi) + \sqrt{-g}\nabla_\nu(\mathbf{A}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi) \\ &\quad - 2\mathbf{S}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi)) + \delta\Psi[\cdots],\end{aligned}$$

where $\mathbf{S}^{\mu\nu} \equiv \mathbf{S}^{(\mu\nu)}$ and $[\cdots]$ denotes the irrelevant expressions in our presentation. As a result, the additional current $\mathcal{J}^\mu_\Delta$ is symplectic just as ω^μ and vanishes for a Killing vector. The relation in Eq. (A5) between the additional current term $\mathcal{J}^\mu_\Delta$ and the symplectic current ω^μ implies that

$$\partial_\mu(\sqrt{-g}\mathcal{J}^\mu_\Delta) = \frac{1}{2}\partial_\mu\omega^\mu. \quad (A6)$$

Now, the identical conservation of the current $\mathbf{J}^\mu$ follows from the identity given in Eq. (A3),

$$\begin{aligned}\partial_\mu(\sqrt{-g}\mathbf{J}^\mu) &= \partial_\mu(\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}) + \partial_\mu(\sqrt{-g}\mathcal{J}^\mu_\Delta)\\ &= -\frac{1}{2}[\delta(\sqrt{-g}\mathcal{E}_\Psi \pounds_\zeta\Psi) - \sqrt{-g}\mathcal{E}_\Psi\pounds_{\delta\zeta}\Psi]\\ &\quad + \frac{1}{2}\pounds_\zeta(\sqrt{-g}\mathcal{E}_\Psi\delta\Psi) + \frac{1}{2}\partial_\mu\omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi)\\ &= 0.\end{aligned}$$

For a covariant Lagrangian $L(\Psi)$, the off-shell Noether current and potential may be introduced as

$$\begin{aligned}J^\mu &= \zeta^\mu\sqrt{-g}L(\Psi) + 2\sqrt{-g}\mathbf{E}^{\mu\nu}\zeta_\nu - \Theta^\mu(\pounds_\zeta\Psi)\\ &= \partial_\nu K^{\mu\nu}(\zeta).\end{aligned} \tag{A7}$$

After some manipulation by using the relation in Eq. (A5), one can obtain the off-shell relation

$$\begin{aligned}2\sqrt{-g}\mathbf{Q}^{\mu\nu}(\zeta;\Psi,\delta\Psi) &= \delta K^{\mu\nu}(\zeta) - K^{\mu\nu}(\delta\zeta) - 2\zeta^{[\mu}\Theta^{\nu]}(\delta\Psi)\\ &\quad + \sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi).\end{aligned} \tag{A8}$$

Note that there may be additional terms in the right-hand side in the above relation when the Lagrangian contains noncovariant terms [10]. By recalling the relations Eqs. (3), (A5), and (5), one may note that

$$\begin{aligned}&\omega^\mu(\Psi|\pounds_\zeta\Psi,\delta\Psi) + 2\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(\zeta;\Psi,\delta\Psi)\\ &\quad = \partial_\nu[2\sqrt{-g}\mathbf{Q}^{\mu\nu}(\zeta;\Psi,\delta\Psi) - \sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_\zeta\Psi,\delta\Psi)]\\ &\quad = \partial_\nu[\delta K^{\mu\nu}(\zeta) - K^{\mu\nu}(\delta\zeta) - 2\zeta^{[\mu}\Theta^{\nu]}(\delta\Psi)],\end{aligned} \tag{A9}$$

where we used the off-shell identity (A8) in the second equality.

It is straightforward to repeat the same procedure in [36] to derive the first law of black hole thermodynamics. Let us introduce the integral $\mathcal{V}$ of the off-shell current $\mathbf{J}^\mu$ as[7]

$$\begin{aligned}\mathcal{V}_\Sigma(\zeta;\Psi,\delta\Psi) &\equiv \frac{1}{8\pi G}\int_\Sigma dx_\mu\sqrt{-g}\mathbf{J}^\mu(\zeta;\Psi,\delta\Psi)\\ &= \frac{1}{8\pi G}\int_\infty dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}\\ &\quad - \frac{1}{8\pi G}\int_B dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}.\end{aligned} \tag{A10}$$

To see the implication of this integral, take the on-shell condition and ζ as a Killing vector K. Then, the off-shell current $\mathbf{J}^\mu_{\rm ADT}$ reduces to the ADT current $\mathcal{J}^\mu_{\rm ADT}$. The on-shell condition implies $\mathcal{J}^\mu_{\rm ADT} = 0$ and the Killing condition leads to $\mathcal{J}^\mu_\Delta = 0$ and $\mathbf{A}^{\mu\nu} = 0$, and so the integral $\mathcal{V}_\Sigma$ vanishes in this case. For the horizon Killing vector K_H,

$$K_H \equiv \frac{\partial}{\partial t} + \Omega_H\frac{\partial}{\partial\theta},$$

the conserved charges are given by

$$\begin{aligned}\frac{1}{8\pi G}\int_\infty dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K_H;\Psi,\delta\Psi) &= \delta M_\infty - \Omega_H\delta J_\infty\\ \frac{1}{8\pi G}\int_B dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K_H;\Psi,\delta\Psi) &= \frac{\kappa}{2\pi}\delta\mathcal{S}_{\rm BH},\end{aligned}$$

and so one can see that the integral gives us the first law of black hole thermodynamics as

$$0 = \delta M_\infty - \Omega_H\delta J_\infty - \frac{\kappa}{2\pi}\delta\mathcal{S}_{\rm BH}. \tag{A11}$$

APPENDIX B: GAUGE INVARIANCE

Just like two expressions of the Lie derivative of the ADT current in Eq. (19) and (20), the Lie derivative of the potential $\mathbf{Q}^{\mu\nu}$ can be written in two ways. Firstly, it can be written as

$$\begin{aligned}\pounds_\epsilon\mathbf{Q}^{\mu\nu}(K;\Psi,\delta\Psi) &= \mathbf{Q}^{\mu\nu}(\pounds_\epsilon K;\Psi,\delta\Psi) + \mathbf{Q}^{\mu\nu}(K;\pounds_\epsilon\Psi,\delta\Psi)\\ &\quad + \mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi).\end{aligned} \tag{B1}$$

Secondly, it can also be written as

$$\begin{aligned}\pounds_\epsilon\mathbf{Q}^{\mu\nu}(K;\Psi,\delta\Psi) &= \nabla_\alpha(3\epsilon^{[\alpha}\mathbf{Q}^{\mu\nu]}) - \mathbf{Q}^{\mu\nu}\nabla_\alpha\epsilon^\alpha\\ &\quad - \epsilon^\mu\nabla_\alpha\mathbf{Q}^{\nu\alpha} - \epsilon^\nu\nabla_\alpha\mathbf{Q}^{\alpha\mu},\\ &= \nabla_\alpha(3\epsilon^{[\alpha}\mathbf{Q}^{\mu\nu]}) - \mathbf{Q}^{\mu\nu}\nabla_\alpha\epsilon^\alpha\\ &\quad - 2\epsilon^{[\mu}\mathcal{J}^{\nu]}_{\rm ADT},\end{aligned} \tag{B2}$$

where we have used that $\mathcal{J}^\mu_{\rm ADT}(K) = \mathbf{J}^\mu(K) = \nabla_\nu\mathbf{Q}^{\mu\nu}(K)$ for a Killing vector K. Combining the above two expressions for the Lie derivative of the potential $\mathbf{Q}^{\mu\nu}$, one obtains

$$\begin{aligned}\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)|_{\text{on-shell}} &= [\nabla_\alpha(3\epsilon^{[\alpha}\mathbf{Q}^{\mu\nu]}(K;\Psi,\pounds_\epsilon\delta\Psi))\\ &\quad - \mathbf{Q}^{\mu\nu}(K;\Psi,\delta\Psi)\nabla_\alpha\epsilon^\alpha\\ &\quad - \mathbf{Q}^{\mu\nu}(K;\pounds_\epsilon\Psi,\delta\Psi)\\ &\quad - \mathbf{Q}^{\mu\nu}(\pounds_\epsilon K;\Psi,\delta\Psi)]_{\text{on-shell}},\end{aligned} \tag{B3}$$

where the on-shell vanishing condition of the ADT current $\mathcal{J}^\mu_{\rm ADT}$ is used. By inserting this equality in Eq. (21), one can see that there are contributions from two boundaries: the spacelike infinity and the bifurcation surface B. The first term in the right-hand side of the above equality does not

[7]Since we use a Killing vector in the derivation of the first law, one can employ $\mathcal{J}^\mu_{\rm ADT}$ instead of $\mathbf{J}^\mu$.

contribute to the modified canonical energy since we integrate over the closed space at both boundaries. Now, let us consider the leftover terms in the right-hand side.

Since ϵ corresponds to the asymptotic symmetry generators at infinity, one can see that $\nabla_\alpha \epsilon^\alpha \to 0$ and $\pounds_\epsilon \Psi \to 0$ sufficiently fast near infinity compared to the field Ψ itself, which would come from the definition of the asymptotic symmetry generators. As can be inferred from the definition of the charge in Eq. (6), it turns out that $\delta Q(K) \simeq \int dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta\Psi)$ is finite (in fact, taken as zero for linear perturbations) at the spacelike infinity. Therefore, the second and third terms in the right-hand side vanish at the spacelike infinity. Furthermore, $\pounds_\epsilon K = [\epsilon, K] = \epsilon'$ corresponds to another asymptotic Killing vector and $\delta Q(\epsilon') \simeq \int dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(\epsilon';\Psi,\delta\Psi) = 0$ under the chosen condition that the charge is invariant for the linear perturbations. As a result, the last term does not contributes at the spacelike infinity.

Since we are taking the same gauge conditions in Ref. [3], our gauge parameter ϵ satisfies the same property as there. Thus, at the bifurcation surface B, the gauge transformation satisfies $\nabla_\alpha \epsilon^\alpha|_B = \mu^{\alpha\beta}\nabla_\alpha \epsilon_\beta|_B = 0$ (see Remark below Lemma 1 in Ref. [3]). Therefore, the second term in the right-hand side in Eq. (B3) does not contribute. At the bifurcation surface B, therefore, the relevant expression in Eq. (B3) can be written as

$$2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)|_{\text{on-shell}} = -2\sqrt{-g}[\mathbf{Q}^{\mu\nu}(K;\pounds_\epsilon\Psi,\delta\Psi) + \mathbf{Q}^{\mu\nu}(\pounds_\epsilon K;\Psi,\delta\Psi)]_{\text{on-shell}}. \tag{B4}$$

For simplicity, let us focus on pure Einstein gravity, in which the potential $\mathbf{Q}^{\mu\nu}$ is given by [11]

$$\mathbf{Q}^{\mu\nu}(\zeta;g,h) = \frac{1}{2}h\nabla^{[\mu}\zeta^{\nu]} - \zeta^{[\mu}\nabla_\alpha h^{\nu]\alpha} + \zeta_\alpha\nabla^{[\mu}h^{\nu]\alpha} + \zeta^{[\mu}\nabla^{\nu]}h - \frac{1}{2}h^{\alpha[\mu}\nabla_\alpha\zeta^{\nu]} + \frac{1}{2}h^{\alpha[\mu}\nabla^{\nu]}\zeta_\alpha. \tag{B5}$$

By using the metric perturbation near the horizon given in Eq. (24), one can see that the perturbation metric $h_{\mu\nu}$ satisfies $n_{[\mu}\ell_{\nu]}h^{\mu\alpha} \to 0$ as $r \to 0$, that is to say, $n_{[\mu}\ell_{\nu]}h^{\mu\alpha}$ vanishes on the future horizon. Therefore, the last two terms in the above potential $\mathbf{Q}^{\mu\nu}(\zeta;g,\delta g)$ do not contribute after the integration over the spatial section $B(t)$. This is the case even for the second variation in the form of $\mathbf{Q}^{\mu\nu}(\zeta;\delta_2 g,\delta_1 g)$.

Let us consider the contribution from the first term $\mathbf{Q}^{\mu\nu}(K;\pounds_\epsilon g,\delta g)$ in the right-hand side of Eq. (B4). By incorporating $K_H \to 0$ at the bifurcation surface B, this term reduces to

$$\mathbf{Q}^{\mu\nu}(K_H;\pounds_\epsilon g,h)|_B = -\frac{1}{2}h^{\alpha\beta}\pounds_\epsilon g_{\alpha\beta}\nabla^{[\mu}K_H^{\nu]}|_B. \tag{B6}$$

Note that the gauge parameter ϵ^μ is tangent to the future horizon and, in fact, its admissible form is given by [5,6]

$$\epsilon^\mu = f_\epsilon n^\mu + rY^\mu_\epsilon, \qquad n^\mu\nabla_\mu f_\epsilon = 0. \tag{B7}$$

Then, the direct computation in the chosen coordinates in Eq. (24) shows us that $h^{\mu\nu}\nabla_{[\mu}\epsilon_{\nu]} \to 0$ as $r \to 0$. Hence, the first term gives zero contribution.

In the chosen gauge near the horizon, $\xi \equiv \pounds_\epsilon K = -[K,\epsilon]$ becomes normal to the future horizon at the surface B, since the gauge parameter ϵ^μ is tangent to the horizon. By using the property of ξ, one can set [3]

$$\xi^\mu = fn^\mu + uX^\mu + rY^\mu. \tag{B8}$$

Noting that $n^\mu\nabla_\alpha h^\alpha_\mu \to 0$ as $r \to 0$, with the expression of ξ near the horizon, one can show that the second term $\mathbf{Q}^{\mu\nu}(\pounds_\epsilon K;\Psi,\delta\Psi)$ in the right-hand side of Eq. (B4) reduces at the surface B to

$$\mathbf{Q}^{\mu\nu}(\xi;g,\delta g) = \frac{1}{2}\delta\mu^\alpha_\alpha\nabla^{[\mu}\xi^{\nu]} + \xi_\alpha\nabla^{[\mu}h^{\nu]\alpha} + \xi^{[\mu}\nabla^{\nu]}h. \tag{B9}$$

Since $h^{\mu\nu}\xi_\nu = h^{\mu\nu}n_\nu = 0$ at the surface B, one can see that $n_{[\mu}\ell_{\nu]}\xi_\alpha\nabla^{[\mu}h^{\nu]\alpha} = 0$ for the metric perturbation $h_{\mu\nu}$ at B and that $n^\mu\nabla_\mu h \propto \mu^{ab}\partial_u\delta\mu_{ab} \propto \delta\vartheta = 0$. Thus, we obtain the following result:

$$2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;g,\pounds_\epsilon\delta g)|_{\text{on-shell}} = -\sqrt{-g}\delta\mu^\alpha_\alpha\nabla^{[\mu}\xi^{\nu]}. \tag{B10}$$

Since the gauge is chosen as $\mu^{\alpha\beta}\delta\mu_{\alpha\beta} = \delta\mu^\alpha_\alpha = 0$, we immediately see that

$$\mathcal{E}(K;\pounds_\epsilon g,\delta g) = \frac{1}{16\pi G}\int_B dx_{\mu\nu}\sqrt{-g}\delta\mu^\alpha_\alpha\nabla^{[\mu}\xi^{\nu]} = 0. \tag{B11}$$

Now, we would like to give comments on the relation to the derivation in Ref. [3]. In short, our derivation is completely parallel and consistent to the one in Appendix A of Ref. [3]. In fact, one can show that

$$2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\pounds_\epsilon\delta\Psi)|_{\text{on-shell}} = -[2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\pounds_\epsilon\Psi,\delta\Psi) - \delta K^{\mu\nu}(\xi) + 2\xi^{[\mu}\Theta^{\nu]}(\delta\Psi) - \sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_\xi\Psi,\delta\Psi)]_{\text{on-shell}}.$$

In pure Einstein gravity, the first term is already shown to give no contribution. Note that the $\mathbf{A}^{\mu\nu}$-tensor is given by

$$\mathbf{A}^{\mu\nu}(\pounds_\xi g,\delta g) = -(g^{\mu(\alpha}g^{\beta)(\rho}g^{\sigma)\nu} - g^{\nu(\alpha}g^{\beta)(\rho}g^{\sigma)\mu}) \times (\pounds_\xi g_{\alpha\beta}h_{\rho\sigma} - h_{\alpha\beta}\pounds_\xi g_{\rho\sigma}), \tag{B12}$$

from which one can see that the above $\mathbf{A}^{\mu\nu}$-tensor term does not contribute to the canonical energy through Eqs. (B3), (B4), and (21). The absence of the contribution from

$\delta K^{\mu\nu}(\xi) - 2\xi^{[\mu}\Theta^{\nu]}(\delta\Psi)$ is the main result in Appendix A in [3].

APPENDIX C: RELATION TO HW CONSTRUCTION

In the case of the Killing vector K with $\delta K = 0$, one can show that the current expression $\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)$ is related to the symplectic current as follows. The variation of the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ can also be written under the condition $\delta K = 0$ as

$$\begin{aligned}\delta_2(\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\Psi,\delta_1\Psi)) &= \sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)\\ &\quad+\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\Psi,\delta_2\delta_1\Psi)\\ &= \sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)\\ &\quad+\partial_\nu[\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta_2\delta_1\Psi)],\end{aligned}$$

where we used in the first equality the on-shell vanishing condition of $\mathcal{J}^{\mu}_{\rm ADT}(\zeta;\Psi,\delta\Psi)$ and used in the second equality $\mathcal{J}^{\mu}_{\rm ADT}(K) = \nabla_\nu\mathbf{Q}^{\mu\nu}(K)$. Under the condition $\delta K = 0$, the generic variation leads to $\delta\pounds_K\Psi = \pounds_K\delta\Psi$. And thus, the variation of the symplectic current becomes $\delta_2\omega^\mu(\Psi|\pounds_K\Psi,\delta_1\Psi) = \omega^\mu(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)$, because of $\pounds_K\Psi = 0$. By taking into account the second variation of the relation in Eq. (A9) with the above second variation of the ADT current $\mathcal{J}^{\mu}_{\rm ADT}$ for the Killing vector K, one can see that

$$\begin{aligned}&[\omega^\mu(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)+2\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}}\\ &\quad=\partial_\nu[2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\delta_2\Psi,\delta_1\Psi)+2(\delta_2\sqrt{-g})\mathbf{Q}^{\mu\nu}(K;\Psi,\delta_1\Psi)\\ &\qquad-\sqrt{-g}\mathbf{A}^{\mu\nu}(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}},\end{aligned}\tag{C1}$$

where we have used $\mathbf{A}^{\mu\nu}(\pounds_K\Psi,\delta\Psi) = 0$. Under the chosen gauges near the horizon and the asymptotic infinity, it turns out that $\delta\sqrt{-g}|_{B(t)} = 0$ at the future horizon and all the terms vanish at infinity because $\delta\Psi$ decays sufficiently fast at infinity. Thus, one concludes that the difference is written eventually as

$$\begin{aligned}&\int_\Sigma dx_\mu[\omega^\mu(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)+2\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}}\\ &\quad=-\int_B dx_{\mu\nu}\sqrt{-g}[2\mathbf{Q}^{\mu\nu}(K;\delta_2\Psi,\delta_1\Psi)-\mathbf{A}^{\mu\nu}(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}},\end{aligned}\tag{C2}$$

which holds for any Cauchy surface $\Sigma(t)$, not just at $\Sigma(t=0)$. As a result, one obtains the relation given in Eq. (30). It is interesting to observe that the above relation reproduces the same expression given by Eq. (B3) and Eq. (B4) by taking $\delta_2 = \pounds_\epsilon$, which might be just a coincidence not warranted from the construction. In pure Einstein gravity, one can show, by the explicit computation as done in Appendix B, that the $\mathbf{A}^{\mu\nu}$-tensor term does not contribute at the surface $B(t)$ in the chosen coordinates near the horizon as (22) and (24). By using the form of the perturbed metric in Eq. (24) and the fact that $n^{[\mu}\ell^{\nu]}h_{\nu\alpha}\to 0$ at the future horizon, i.e., at $r=0$, one can see that the relevant potential term for the horizon Killing vector K_H is given by

$$\begin{aligned}\mathbf{Q}^{\mu\nu}(K_H;\delta_2 g,\delta_1 g)|_{B(t)} &= \frac{1}{2}\delta_2 g^{\alpha\beta}\delta_1 g_{\alpha\beta}\nabla^{[\mu}K_H^{\nu]}\\ &\quad-\frac{1}{2}\delta_1 g^{\alpha\beta}K_H^{[\mu}\nabla^{\nu]}\delta_2 g_{\alpha\beta}\\ &\quad+K_H^{[\mu}\nabla^{\nu]}(\delta_2 g^{\alpha\beta}\delta g_{\alpha\beta})|_{B(t)},\end{aligned}$$

where we have used that K_H is normal to the future horizon and $n_{[\mu}\ell_{\nu]}h^{\mu\alpha}\to 0$ as $r\to 0$. Thus, we obtain

$$\begin{aligned}&2n_{[\mu}\ell_{\nu]}\mathbf{Q}^{\mu\nu}(K_H;\delta g,\delta g)|_{B(t)}\\ &\quad=\left[\kappa\delta\mu^{\alpha\beta}\delta\mu_{\alpha\beta}-\frac{3}{2}\delta\mu^{\alpha\beta}\pounds_{K_H}\delta\mu_{\alpha\beta}\right]_{B(t)}.\end{aligned}\tag{C3}$$

The above relation may also be written as

$$\begin{aligned}&\int_{\partial\Sigma}dx_{\mu\nu}[\omega^\mu(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi)+2\sqrt{-g}\mathcal{J}^{\mu}_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}}\\ &\quad=\partial_\nu[\delta_2\delta_1K^{\mu\nu}(K)-2K^{[\mu}\delta_2\Theta^{\nu]}(\delta_1\Psi)-2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta_2\delta_1\Psi)]_{\text{on-shell}}.\end{aligned}\tag{C4}$$

From the relation in Eq. (A8), one can see that the last term in the right-hand side in the equality cancels the second order perturbation terms in the proceeding terms. Explicitly, it can be written as

$$2\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\Psi,\delta_2\delta_1\Psi) = \delta_2\delta_1 K^{\mu\nu}(K)|_{\delta_2\delta_1\Psi} - 2K^{[\mu}\Theta^{\nu]}(\Psi,\delta_2\delta_1\Psi), \qquad \text{(C5)}$$

where the subscript $\delta_2\delta_1\Psi$ means that we should keep the second order variations. Schematically, one can write the above relation of the current expression as

$$[\omega^\mu(\Psi|\pounds_K\delta_2\Psi,\delta_1\Psi) + 2\sqrt{-g}\mathcal{J}^\mu_{\rm ADT}(K;\delta_2\Psi,\delta_1\Psi)]_{\text{on-shell}} = \partial_\nu[\delta_2\delta_1 K^{\mu\nu}(K) - 2K^{[\mu}\delta_2\Theta^{\nu]}(\delta_1\Psi)]^{\delta_2\delta_1\Psi=0}_{\text{on-shell}}, \qquad \text{(C6)}$$

where $\delta_2\delta_1\Psi = 0$ in the superscript denotes the absence of second order variations in the expressions.

On the bifurcation surface B, the contribution comes from the first term $\delta^2 K^{\mu\nu}$, only. In fact, the essentially same relation has already been obtained in Ref. [3] [see Eq. (81) there], though its derivation and interpretation seem to be different. In the end, the difference between $\mathcal{E}_{\rm HW}$ and $\mathcal{E}$ is given by

$$\mathcal{E}_{\rm HW}(K;\delta\Psi,\delta\Psi) - \mathcal{E}(K;\delta\Psi,\delta\Psi) = -\frac{1}{16\pi G}\int_B dx_{\mu\nu}[\delta^2 K^{\mu\nu}(K)]^{\delta^2\Psi=0}_{\text{on-shell}}, \qquad \text{(C7)}$$

which can also be written, through Eq. (C3), at least in Einstein gravity as

$$\mathcal{E}_{\rm HW}(K;\delta\Psi,\delta\Psi) - \mathcal{E}(K;\delta\Psi,\delta\Psi) = -\frac{1}{8\pi G}\int_B dx_{\mu\nu}\sqrt{-g}\mathbf{Q}^{\mu\nu}(K;\delta\Psi,\delta\Psi)|_{\text{on-shell}}. \qquad \text{(C8)}$$

By noting that the modified canonical energy differs from the HW canonical energy only on the bifurcation surface B, whenever $\delta_1\Psi$ or $\delta_2\Psi$ are taken in such a way that $\delta_{1,2}Q(K) = 0$ at infinity, one can see that our modified canonical energy satisfies the same properties with the HW canonical energy for the perturbation *toward stationary black holes*.

[1] S. S. Gubser and I. Mitra, Instability of charged black holes in anti-de Sitter space, arXiv:hep-th/0009126.

[2] S. S. Gubser and I. Mitra, The evolution of unstable black holes in anti-de Sitter space, J. High Energy Phys. 08 (2001) 018.

[3] S. Hollands and R. M. Wald, Stability of black holes and black branes, Commun. Math. Phys. **321**, 629 (2013).

[4] P. Figueras, K. Murata, and H. S. Reall, Black hole instabilities and local Penrose inequalities, Classical Quantum Gravity **28**, 225030 (2011).

[5] S. Hollands and A. Ishibashi, Instabilities of extremal rotating black holes in higher dimensions, Commun. Math. Phys. **339**, 949 (2015).

[6] S. R. Green, S. Hollands, A. Ishibashi, and R. M. Wald, Superradiant instabilities of asymptotically anti-de Sitter black holes, Classical Quantum Gravity **33**, 125022 (2016).

[7] J. Keir, Stability, instability, canonical energy and charged black holes, Classical Quantum Gravity **31**, 035014 (2014).

[8] N. Lashkari and M. Van Raamsdonk, Canonical energy is quantum Fisher information, J. High Energy Phys. 04 (2016) 153.

[9] W. Kim, S. Kulkarni, and S.-H. Yi, Quasilocal Conserved Charges in a Covariant Theory of Gravity, Phys. Rev. Lett. **111**, 081101 (2013); **112**, 079902(E) (2014).

[10] W. Kim, S. Kulkarni, and S.-H. Yi, Quasilocal conserved charges in the presence of a gravitational Chern-Simons term, Phys. Rev. D **88**, 124004 (2013).

[11] S. Hyun, S.-A. Park, and S.-H. Yi, Quasi-local charges and asymptotic symmetry generators, J. High Energy Phys. 06 (2014) 151.

[12] S. Hyun, J. Jeong, S.-A. Park, and S.-H. Yi, Quasilocal conserved charges and holography, Phys. Rev. D **90**, 104016 (2014).

[13] S. Hyun, J. Jeong, S.-A. Park, and S.-H. Yi, Frame-independent holographic conserved charges, Phys. Rev. D **91**, 064052 (2015).

[14] Y. S. Myung and T. Moon, Thermodynamic and classical instability of AdS black holes in fourth-order gravity, J. High Energy Phys. 04 (2014) 058; Y. S. Myung, Phase transitions of the BTZ black hole in new massive gravity, Adv. High Energy Phys. **2015**, 1 (2015); J. J. Peng, Conserved charges of black holes in Weyl and Einstein-Gauss-Bonnet gravities, Eur. Phys. J. C **74**, 3156 (2014); Off-shell Noether current and conserved charge in Horndeski theory, Phys. Lett. B **752**, 191 (2016); E. Ayn-Beato, M. Bravo-Gaete, F. Correa, M. Hassane, M. M. Jurez-Aubry, and J. Oliva, First law and anisotropic Cardy formula for three-dimensional Lifshitz black holes, Phys. Rev. D **91**, 064006 (2015); M. Bravo-Gaete and M. Hassaine, Thermodynamics of charged Lifshitz black holes with quadratic corrections, Phys. Rev. D **91**, 064038 (2015); M. Bravo-Gaete, S. Gomez, and M. Hassaine, Towards the Cardy formula for hyperscaling violation black holes, Phys. Rev. D **91**, 124038 (2015); Cardy formula for charged black holes with anisotropic scaling, Phys. Rev. D **92**, 124002 (2015); S. Q. Wu and S. Li, Thermodynamics of static dyonic AdS black holes in the ω-deformed Kaluza-Klein gauged supergravity theory, Phys. Lett. B **746**, 276 (2015); M. R. Setare and H. Adami, Black hole conserved charges in generalized minimal massive gravity, Phys. Lett. B **744**, 280 (2015); Quasi-local conserved charges in Lorentz-diffeomorphism covariant theory of gravity, arXiv:1511.00527; Quasi-local conserved charges of spin-3 topologically massive gravity, arXiv:1601.00171; J. H. Park, S. J. Rey, W. Rim, and Y. Sakatani, O(D, D) covariant Noether currents and global charges in double field theory, J. High Energy Phys. 11 (2015)

131; K. A. Moussa, G. Clément, and H. Guennoune, Chern-Simons dilaton black holes in $2+1$ dimensions, Classical Quantum Gravity **33**, 065008 (2016).
[15] L. F. Abbott and S. Deser, Stability of gravity with a cosmological constant, Nucl. Phys. **B195**, 76 (1982).
[16] L. F. Abbott and S. Deser, Charge definition in nonabelian gauge theories, Phys. Lett. **116B**, 259 (1982).
[17] S. Deser and B. Tekin, Gravitational Energy in Quadratic Curvature Gravities, Phys. Rev. Lett. **89**, 101101 (2002).
[18] S. Deser and B. Tekin, Energy in generic higher curvature gravity theories, Phys. Rev. D **67**, 084009 (2003).
[19] J. Lee and R. M. Wald, Local symmetries and constraints, J. Math. Phys. (N.Y.) **31**, 725 (1990).
[20] V. Moncrief and J. Isenberg, Symmetries of cosmological Cauchy horizons, Commun. Math. Phys. **89**, 387 (1983).
[21] S. Hollands, A. Ishibashi, and R. M. Wald, A higher dimensional stationary rotating black hole must be axisymmetric, Commun. Math. Phys. **271**, 699 (2007).
[22] R. P. Geroch and B. C. Xanthopoulos, Asymptotic simplicity is stable, J. Math. Phys. (N.Y.) **19**, 714 (1978).
[23] S. Hollands, A. Ishibashi, and D. Marolf, Comparison between various notions of conserved charges in asymptotically AdS-spacetimes, Classical Quantum Gravity **22**, 2881 (2005).
[24] C. X. Habisohn, Calculation of radiated gravitational energy using the second-order Einstein tensor, J. Math. Phys. (N.Y.) **27**, 2759 (1986).
[25] T. Azeyanagi, G. Compere, N. Ogawa, Y. Tachikawa, and S. Terashima, Higher-derivative corrections to the asymptotic Virasoro symmetry of 4d extremal black holes, Prog. Theor. Phys. **122**, 355 (2009).
[26] Y. Kwon, S. Nam, J. D. Park, and S.-H. Yi, Extremal black holes and holographic C-theorem, Nucl. Phys. **B869**, 189 (2013).
[27] S. Hyun, J. Jeong, and S.-H. Yi, Fake supersymmetry and extremal black holes, J. High Energy Phys. 03 (2013) 042.
[28] K. Hotta, Y. Hyakutake, T. Kubota, T. Nishinaka, and H. Tanida, The CFT-interpolating black hole in three dimensions, J. High Energy Phys. 01 (2009) 010.
[29] P. Breitenlohner and D. Z. Freedman, Stability in gauged extended supergravity, Ann. Phys. (N.Y.) **144**, 249 (1982).
[30] T. Hertog, Towards a novel no-hair theorem for black holes, Phys. Rev. D **74**, 084008 (2006).
[31] T. Hertog and K. Maeda, Stability and thermodynamics of AdS black holes with scalar hair, Phys. Rev. D **71**, 024001 (2005).
[32] M. Banados, C. Teitelboim, and J. Zanelli, The Black Hole in Three-Dimensional Space-Time, Phys. Rev. Lett. **69**, 1849 (1992).
[33] N. Iizuka, A. Ishibashi, and K. Maeda, A rotating hairy AdS_3 black hole with the metric having only one Killing vector field, J. High Energy Phys. 08 (2015) 112.
[34] M. Durkee and H. S. Reall, Perturbations of near-horizon geometries and instabilities of Myers-Perry black holes, Phys. Rev. D **83**, 104044 (2011).
[35] M. Banados, M. Henneaux, C. Teitelboim, and J. Zanelli, Geometry of the $(2+1)$ black hole, Phys. Rev. D **48**, 1506 (1993); **88**, 069902(E) (2013).
[36] R. M. Wald, Black hole entropy is the Noether charge, Phys. Rev. D **48**, R3427 (1993).