

This is the accepted manuscript made available via CHORUS. The article has been published as:

## Gedanken experiments to destroy a BTZ black hole

Bo Ning, Baoyi Chen, and Feng-Li Lin

Phys. Rev. D **100**, 044043 — Published 23 August 2019

DOI: [10.1103/PhysRevD.100.044043](https://doi.org/10.1103/PhysRevD.100.044043)

# Gedanken Experiments to Destroy a BTZ Black Hole

Baoyi Chen<sup>\*a</sup>, Feng-Li Lin<sup>†b</sup> and Bo Ning<sup>‡c</sup>

<sup>a</sup> *TAPIR, Walter Burke Institute for Theoretical Physics,  
California Institute of Technology, Pasadena, California 91125, USA*

<sup>b</sup> *Department of Physics, National Taiwan Normal University,  
No. 88, Sec. 4, Ting-Chou Road, Taipei 11677, Taiwan*

<sup>c</sup> *College of Physical Science and Technology,  
Sichuan University, Chengdu, Sichuan 610065, China*

## Abstract

We consider gedanken experiments to destroy an extremal or near-extremal BTZ black hole by throwing matter into the horizon. These black holes are vacuum solutions to (2+1)-dimensional gravity theories, and are asymptotically  $\text{AdS}_3$ . Provided the null energy condition for the falling matter, we prove the following—(i) in a Mielke-Baekler model without ghost fields, when torsion is present, an extremal BTZ black hole can be overspun and becomes a naked conical singularity; (ii) in 3-dimensional Einstein gravity and chiral gravity, which both live in the torsionless limits of Mielke-Baekler model, an extremal BTZ black hole cannot be overspun; and (iii) in both Einstein gravity and chiral gravity, a near-extremal BTZ black hole cannot be overspun, leaving the weak cosmic censorship preserved. To obtain these results, we follow the analysis of Sorce and Wald on their gedanken experiments to destroy a Kerr-Newman black hole, and calculate the second order variation of the

---

<sup>\*</sup>baoyi@tapir.caltech.edu

<sup>†</sup>linfengli@phy.ntnu.edu.tw

<sup>‡</sup>ningbo@scu.edu.cn

black hole mechanics. Furthermore, Wald's type of gedanken experiments provide an operational procedure of proving the third law of black hole dynamics. Through the AdS/CFT correspondence, our results on BTZ black holes also indicate that a third law of thermodynamics holds for the holographic conformal field theories dual to 3-dimensional Einstein gravity and chiral gravity.

# Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>BTZ black hole and variational identities</b>	<b>6</b>
2.1	First order variations . . . . .	8
2.2	Second order variations . . . . .	13
<b>3</b>	<b>Gedanken experiment to destroy an extremal BTZ</b>	<b>14</b>
<b>4</b>	<b>Gedanken experiment to destroy a near-extremal BTZ</b>	<b>17</b>
4.1	Chiral gravity . . . . .	18
4.2	Einstein gravity . . . . .	19
<b>5</b>	<b>Conclusions and Discussions</b>	<b>24</b>

## 1 Introduction

Weak cosmic censorship conjecture (WCCC) was formulated by Penrose [1] to postulate that a gravitational singularity should not be naked and should be hidden inside a black hole horizon. A gravitational singularity is usually mathematically ill-defined due to the divergent spacetime curvature. Thus, the WCCC helps to avoid seeing such unphysical part of the universe and retains the predicted power of physical laws. Its philosophical incarnation was summarized by Hawking that “Nature abhors a naked singularity” [3]. In this sense, a special case worthy of consideration is the 3-dimensional Banados-Teitelboim-Zanelli (BTZ) black hole, for which there is no curvature singularity but a conical one. The conical singularity will thus cause no physical divergence as the curvature one. It is then interesting to check if WCCC holds for this case or not, and partly motivates the study of this paper.

The general proof or demonstration of WCCC is notoriously difficult. One way is to find the critical situation in which a black hole almostly turns into a naked singularity by subjecting to small perturbations. This is when a Kerr-Newman black hole is in its near-extremal regime. A super-extremal solution possesses the naked singularity, thus checking WCCC is to see if a sub-extremal black hole in the near-extremal limit can turn into a super-extremal one by throwing some matter. Along this line of thought, a gedanken experiment was firstly proposed by Wald [4]

to demonstrate the impossibility of destroying an extremal Kerr-Newman black hole by throwing the matter obeying the null energy condition. The key ingredient in [4] is the linear variation of black hole mechanics [5, 6], i.e.,

$$\delta M - \Omega_H \delta J - \Phi_H \delta Q \geq 0 \quad (1)$$

where  $M$  is the mass of the black hole,  $J$  the angular momentum,  $Q$  the charge, and  $\Omega_H$  and  $\Phi_H$  are respectively the angular velocity and chemical potential evaluated on the horizon. A similar consideration for the near-extremal Reissner-Nordstrom black hole was examined by Hubeny [7] and found that it can be overcharged to violate WCCC by throwing a charged particle. See [9–12] for the follow-up works.

Recently, it was realized by Sorce and Wald [15] that the analysis in Hubeny's type of gedanken experiments is insufficient at the linear order so that the second order variation must be taken into account to check WCCC for near-extremal black holes. Based on an earlier development of the second order variation of black hole mechanics [16], they went beyond the first order analysis in [7] and derived the following inequality

$$\delta^2 M - \Omega_H \delta^2 J - \Phi_H \delta^2 Q \geq -T_H \delta^2 S \quad (2)$$

with  $T_H$  the Hawking temperature and  $S$  the Wald's black hole entropy [17] [18], which equals to the Bekenstein-Hawking entropy of area law for the case of Einstein's theory of gravity, but receives modifications for non-Einstein theories of gravity [19] (See (40) for the case of Mielke-Barkler gravity). Under the situation that the linear variation is optimally done, i.e., the inequality (1) is saturated, Sorce and Wald adopted (2) to show that the WCCC holds for Kerr-Newman black holes in 4-dimensional Einstein-Maxwell gravity. In [15] it is assumed that the near-extremal black hole is linearly stable, so that at very late time the linear perturbation induced by falling matter becomes the perturbation towards another Kerr-Newman black hole. Thus, the WCCC is to prohibit the possibility of a naked singularity, and can be formally described as the condition for a 1-parameter family of black hole solutions

$$f(\lambda) > 0, \quad \text{for all } \lambda \geq 0 \quad (3)$$

with  $f(\lambda) = 0$  being the condition for extremal black hole. For examples,  $f(\lambda) = M(\lambda)^2 - \frac{J(\lambda)^2}{M(\lambda)^2} - Q(\lambda)^2$  for a Kerr-Newman black hole of mass  $M(\lambda)$ , angular momentum  $J(\lambda)$  and charge  $Q(\lambda)$ , and  $f(\lambda) = M(\lambda)^2 + \Lambda J(\lambda)^2$  for a BTZ black hole in 3-dimensional anti-de Sitter (AdS) space of cosmological constant  $\Lambda < 0$ .

Note that there is no need in this formulation of examining WCCC to consider the self-force effects of the in-falling matter as done in [20–24].

In this paper, we will check WCCC for a BTZ black hole in 3-dimensional torsional Mielke-Barkler gravity (MBG) [28–30] with the general falling matter <sup>1</sup>. In some special limits of MBG we have either Einstein gravity or chiral gravity [31], both of which have the known dual descriptions by a 2-dimensional conformal field theory (CFT) in the context of AdS/CFT correspondence [2]. Especially, the extremal black hole has zero surface gravity, and corresponds to a dual CFT state at zero temperature. The motivation of our study is two folds. First, we would like to see if WCCC holds even for the naked conical singularity such as the one in BTZ, and at the same time extend the formulation of [15] to more general gravity theories. Second, Wald’s type of gedanken experiments provide an operational procedure of proving the third law of black hole dynamics [8, 10]: One cannot turn the non-extremal black hole into an extremal one in the finite time-interval by throwing into the black hole the matter satisfying the null energy condition. We can turn the above third law into the one of black hole thermodynamics if we adopt Bekenstein and Hawking’s point of view. Moreover, through the AdS/CFT correspondence, this third law also corresponds to the third law of thermodynamics for the dual 2-dimensional CFT <sup>2</sup>. Our results indicate that such a third law of thermodynamics holds for the holographic CFTs dual to 3-dimensional Einstein gravity and chiral gravity. Intuitively, the cooling procedure can be holographically understood as throwing the coolant, i.e., matter of spin  $J$  and energy  $E$  with  $J > E$ , into the black hole.

We organize the rest of the paper as follows. In Sec. 2 we derive the linear and second order variational identities for the MB model, with which we can proceed to the consideration of gedanken experiments for three ghost-free limits of MB model, i.e., the Einstein gravity, chiral gravity and torsional chiral gravity. In Sec. 3 we consider the gedanken experiments for an extremal BTZ black hole by using the linear variational identity and the null energy condition. In Sec. 4 we check WCCC for nonextremal BTZ black holes for the chiral gravity and Einstein gravity. Finally in Sec. 5 we summarize our results and conclude with some discussions on the issue of proving the third law and its implication to the holographic dual CFTs.

---

<sup>1</sup>See also recent papers [13, 14] for the related discussion for special falling matter.

<sup>2</sup>See [34] for the earlier discussion for AdS<sub>5</sub> case in the context other than WCCC.

## 2 BTZ black hole and variational identities

BTZ black holes are topologically non-trivial solutions to the 3-dimensional Einstein gravity as well as the topological massive gravity (TMG) [25–27]. In fact, they are solutions to a quite general category of gravity theories with the name Mielke-Baekler (MB) model [28, 29] which also incorporates torsion, with Einstein gravity and TMG arise as special limits in its parameter space. In this section, we derive the variational identities and canonical energy for this model following Wald’s formulation.

In three dimensional spacetime, it is convenient to express the gravity theory in the first order formalism. The Lagrangian of a general chiral gravity with torsion, namely the MB model, is as following:

$$L = L_{\text{EC}} + L_{\Lambda} + L_{\text{CS}} + L_{\text{T}} + L_{\text{M}} , \quad (4)$$

where

$$L_{\text{EC}} = \frac{1}{\pi} e^a \wedge R_a , \quad (5)$$

$$L_{\Lambda} = -\frac{\Lambda}{6\pi} \epsilon_{abc} e^a \wedge e^b \wedge e^c , \quad (6)$$

$$L_{\text{CS}} = -\theta_{\text{L}} \left( \omega^a \wedge d\omega_a + \frac{1}{3} \epsilon_{abc} \omega^a \wedge \omega^b \wedge \omega^c \right) , \quad (7)$$

$$L_{\text{T}} = \frac{\theta_{\text{T}}}{2\pi^2} e^a \wedge T_a , \quad (8)$$

in which  $\Lambda < 0$  is the cosmological constant,  $\theta_{\text{L}}$  and  $\theta_{\text{T}}$  are coupling constants,  $L_{\text{EC}}$  is the Einstein-Cartan term,  $L_{\Lambda}$  is the cosmological constant term,  $L_{\text{CS}}$  is the Chern-Simons (CS) terms for curvature,  $L_{\text{T}}$  is a translational Chern-Simons term, and  $L_{\text{M}}$  is the Lagrangian for the matter.  $T_a$  is the torsion 2-form defined by  $T^a = de^a + \omega^a_b \wedge e^b$  with  $e^a$  the dreibeins. We have also defined the dual spin connection  $\omega^a$  and the dual curvature 2-form  $R^a$  for simplicity:

$$\omega^a = \frac{1}{2} \epsilon^a_{bc} \omega^{bc} , \quad R^a = \frac{1}{2} \epsilon^a_{bc} R^{bc} , \quad (9)$$

Variations of the Lagrangian (4-8) with respect to the dreibeins  $e^a$  and the dual spin connections  $\omega^a$  give rise to the equations of motion  $E_a^{(e)} = 0$  and  $E_a^{(\omega)} = 0$  with

$$E_a^{(e)} = \frac{1}{\pi} \left( R_a + \frac{\theta_{\text{T}}}{\pi} T_a - \frac{\Lambda}{2} \epsilon_{abc} e^b \wedge e^c \right) , \quad (10)$$

$$E_a^{(\omega)} = \frac{1}{\pi} \left( T_a - 2\pi\theta_{\text{L}} R_a + \frac{\theta_{\text{T}}}{2\pi} \epsilon_{abc} e^b \wedge e^c \right) \quad (11)$$

for vanishing matter. For the case  $1 + 2\theta_T\theta_L \neq 0$ , the equations of motion are solved by

$$T^a = \frac{\mathcal{T}}{\pi} \epsilon^a_{bc} e^b \wedge e^c, \quad (12)$$

$$R^a = -\frac{\mathcal{R}}{2\pi^2} \epsilon^a_{bc} e^b \wedge e^c, \quad (13)$$

in which

$$\mathcal{T} \equiv \frac{-\theta_T + 2\pi^2\Lambda\theta_L}{2 + 4\theta_T\theta_L}, \quad \mathcal{R} \equiv -\frac{\theta_T^2 + \pi^2\Lambda}{1 + 2\theta_T\theta_L}. \quad (14)$$

The MB model was originally proposed as a torsional generalization of TMG. It has a Poincaré gauge theory description, and there are propagating massive gravitons just like in TMG. We will be especially interested in three limits:

(i) Einstein gravity (with negative cosmological constant). This could be approached by taking the limit  $\theta_L \rightarrow 0$  and  $\theta_T \rightarrow 0$ .

(ii) Chiral gravity. The torsionless branch of the MB model, which is equivalent to TMG, could be obtained by setting  $\mathcal{T} = 0$  according to (12). It was pointed out in [31] that TMG is only well defined at the critical point in which the dual CFT becomes chiral. In our convention, the critical point is located at  $\theta_L = -1/(2\pi\sqrt{-\Lambda})$ . Hence the chiral gravity is approached by setting  $\mathcal{T} = 0$  first and then taking the limit  $\theta_L \rightarrow -1/(2\pi\sqrt{-\Lambda})$ .

(iii) Torsional chiral gravity. For the branch with non-vanishing torsion, we note from the Lagrangian (4-8) that the torsion field  $T_a$  could not be kinematic since there is no second order derivative of  $\omega^a$ . The torsion field should just contribute to the interaction term in the linearized theory, while the propagators of the gravitons should not be changed compared with TMG. We then expect that the MB model also behaves well with no ghost at the critical point  $\theta_L \rightarrow -1/(2\pi\sqrt{-\Lambda})$ . Note that by taking this limit first, we obtain  $\mathcal{T} \rightarrow \pi\sqrt{-\Lambda}/2$  hence the torsion field could not be vanishing. This is a different limit from the case (ii), and we refer it as the torsional chiral gravity.

An interesting class of solutions to the equations (12)(13) are the BTZ-like solutions with non-vanishing torsion [32]. They are described by the following dreibeins

$$e^0 = Ndt, \quad e^1 = \frac{dr}{N}, \quad e^2 = r(d\phi + N^\phi dt), \quad (15)$$

and the dual spin connections

$$\omega^a = \tilde{\omega}^a + \frac{\mathcal{T}}{\pi} e^a, \quad (16)$$



with the torsion free parts

$$\tilde{\omega}^0 = N d\phi, \quad \tilde{\omega}^1 = -\frac{N^\phi}{N} dr, \quad \tilde{\omega}^2 = -\Lambda_{\text{eff}} r dt + r N^\phi d\phi, \quad (17)$$

in which

$$N^2(r) = -M - \Lambda_{\text{eff}} r^2 + \frac{J^2}{4r^2}, \quad N^\phi(r) = -\frac{J}{2r^2}, \quad (18)$$

where  $M$  and  $J$  are constants corresponding to mass and angular momentum of BTZ black hole, respectively, for the case of Einstein's gravity, and

$$\Lambda_{\text{eff}} \equiv -\frac{\mathcal{T}^2 + \mathcal{R}}{\pi^2}. \quad (19)$$

Taking the torsion free limit  $\mathcal{T} \rightarrow 0$ , the above solutions recover the usual BTZ black holes with  $\Lambda_{\text{eff}} = \Lambda$ . The horizons are located at

$$r_\pm^2 = \frac{1}{2\Lambda_{\text{eff}}} \left( -M \mp \sqrt{M^2 + \Lambda_{\text{eff}} J^2} \right) \quad (20)$$

(note that  $\Lambda_{\text{eff}} < 0$  for asymptotic AdS solutions), and the angular velocity of the outer horizon is

$$\Omega_{\text{H}} = \frac{J}{2r_+^2} = \frac{r_-}{r_+} \sqrt{-\Lambda_{\text{eff}}}. \quad (21)$$

The black hole temperature is fixed by the vanishing of the conical singularity of the corresponding Euclidean metric:

$$T_{\text{H}} = -\frac{\Lambda_{\text{eff}} (r_+^2 - r_-^2)}{2\pi r_+}, \quad (22)$$

and the surface gravity is  $\kappa_{\text{H}} = 2\pi T_{\text{H}}$ .

## 2.1 First order variations

Wald's gedanken experiment to destroy a black hole begins with considering a general off-shell variation of the fields, which in principle incorporates all kinds of possible perturbations of a black hole, including throwing matter into it. From the variational identities one obtains general constraints obeyed by these perturbations.

The first order variation of the Lagrangian (4-8) gives rise to the equations of motion as well as a surface term:

$$\delta L = \delta e^a \wedge E_a^{(e)} + \delta \omega^a \wedge E_a^{(\omega)} + d\Theta(\phi, \delta\phi), \quad (23)$$

in which  $\phi = (e^a, \omega^a)$ ,  $E_a^{(e)}$  and  $E_a^{(\omega)}$  are given by (10) and (11). The surface term  $\Theta(\phi, \delta\phi)$ , called the symplectic potential, is evaluated to be

$$\Theta(\phi, \delta\phi) = \frac{1}{\pi} \delta \omega^a \wedge e_a + \frac{\theta_{\text{T}}}{2\pi^2} \delta e^a \wedge e_a - \theta_{\text{L}} \delta \omega^a \wedge \omega_a. \quad (24)$$

In Wald's approach, the space of field configurations is the phase space of the theory, and the variation  $\delta\phi \equiv (d\phi/d\lambda)|_{\lambda=0}$  is the phase space flow vector associated with a 1-parameter family of field configurations  $\phi(\lambda)$ . For a 2-parameter family of field configurations  $\phi(\lambda_1, \lambda_2)$ , one could define the symplectic current

$$\Omega(\phi, \delta_1\phi, \delta_2\phi) = \delta_1\Theta(\phi, \delta_2\phi) - \delta_2\Theta(\phi, \delta_1\phi), \quad (25)$$

in which  $\delta_1, \delta_2$  denote derivatives with respect to parameters  $\lambda_1, \lambda_2$ :

$$\delta_1 = \left. \frac{\partial}{\partial \lambda_1} \right|_{\lambda=0}, \quad \delta_2 = \left. \frac{\partial}{\partial \lambda_2} \right|_{\lambda=0}. \quad (26)$$

One can show that the symplectic current is conserved when the linearized equations of motion are satisfied:

$$d\Omega = 0. \quad (27)$$

The Noether current 2-form associated with a vector field  $\xi$  is defined by

$$j_\xi = \Theta(\phi, \mathcal{L}_\xi\phi) - i_\xi L, \quad (28)$$

in which  $i_\xi L$  represents the interior derivative which contracts  $\xi^\mu$  into the first index of the 3-form  $L$ . Then,  $j_\xi$  could be written in the form

$$j_\xi = dQ_\xi + C_\xi, \quad (29)$$

in which the Noether charge  $Q_\xi$  and the constraints  $C_\xi$  are given by

$$Q_\xi = \frac{1}{\pi} (i_\xi \omega^a) \wedge e_a + \frac{\theta_T}{2\pi^2} (i_\xi e^a) \wedge e_a - \theta_L (i_\xi \omega^a) \wedge \omega_a, \quad (30)$$

$$C_\xi = -(i_\xi e^a) \wedge E_a^{(e)} - (i_\xi \omega^a) \wedge E_a^{(\omega)}. \quad (31)$$

Suppose the field configuration is a family of asymptotic AdS spacetime. Variation of equations (28)(29) gives rise to the following linear variational identity after integrating over an achronal hypersurface  $\Sigma$ :

$$\int_{\partial\Sigma} \delta Q_\xi - i_\xi \Theta(\phi, \delta\phi) = \int_\Sigma \Omega(\phi, \delta\phi, \mathcal{L}_\xi\phi) - \int_\Sigma \delta C_\xi - \int_\Sigma i_\xi (E\delta\phi). \quad (32)$$

The first term on the right hand side is recognized as the variation of the Hamiltonian  $h_\xi$  associated with the diffeomorphism generated by the vector field  $\xi$

$$\delta h_\xi = \int_\Sigma \Omega(\phi, \delta\phi, \mathcal{L}_\xi\phi). \quad (33)$$

Note that  $\delta h_\xi$  (or the first term on the R.H.S. of (32)) vanishes if  $\xi$  is a Killing field, i.e.,  $\mathcal{L}_\xi \phi = 0$ . If the  $\phi$  is on-shell so that  $E_a = 0$ , then the last two terms on the R.H.S. of (32) also vanish. This then motivates the following definition of the conserved ADM quantity  $H_\xi$  conjugate to the Killing field  $\xi$  for an on-shell  $\phi$  [17] [18]:

$$\delta H_\xi = \int_\infty \delta Q_\xi - i_\xi \Theta(\phi, \delta\phi), \quad (34)$$

where  $\int_\infty$  is the integration over the circle at spatial infinity. For a black hole solution, the boundary of  $\Sigma$  contains also horizon as the “inner boundary” besides the “outer boundary” at spatial infinity, then there will be contribution to the L.H.S. of (32) from the “inner boundary” as well (i.e., the area law term for the Einstein gravity). Combining all the above, (32) finally yields the first law of black hole mechanics/thermodynamics.

For the timelike Killing field  $\partial/\partial t$  and the rotational Killing field  $\partial/\partial\varphi$ , the above integration gives rise to the variation of the total mass  $\mathcal{M}$  and the total angular momentum  $\mathcal{J}$ , respectively. For the BTZ-like black holes (15-17), it could be evaluated that [33]

$$\mathcal{M} = M - 2\theta_L (\mathcal{T}M + \pi\Lambda_{\text{eff}}J), \quad (35)$$

$$\mathcal{J} = J + 2\theta_L (\pi M - \mathcal{T}J). \quad (36)$$

For the case that the equations of motion are satisfied and  $\xi$  is a Killing field, the linear variational identity (32) yields

$$\int_{\partial\Sigma} \delta Q_\xi - i_\xi \Theta(\phi, \delta\phi) = - \int_\Sigma \delta C_\xi. \quad (37)$$

For nonextremal black holes, the boundaries include the infinity as well as the bifurcation surface  $B$ . If  $\xi$  is the horizon Killing field  $\xi^a = \partial/\partial t + \Omega_H \partial/\partial\varphi$ , the boundary integral over infinity is given by

$$\int_\infty \delta Q_\xi - i_\xi \Theta(\phi, \delta\phi) = \delta\mathcal{M} - \Omega_H \delta\mathcal{J}. \quad (38)$$

For Einstein’s theory of gravity, the boundary contribution from the bifurcation surface  $B$  turns out to be proportional to the variation of the Bekenstein-Hawking entropy [17] [18]:

$$\int_B \delta Q_\xi - i_\xi \Theta(\phi, \delta\phi) = T_H \delta S, \quad (39)$$

in which  $S = A_B/4$  where  $A_B$  is the area of the bifurcation surface. We will take the above equation as a rightful definition of the modified black hole entropy in

the MB model so that the first law of the black hole thermodynamics still holds but with  $S$  Wald's generalized black hole entropy. It has been evaluated for the BTZ-like black holes that [33]

$$S = 4\pi r_+ - 8\pi\theta_L \left( \mathcal{T}r_+ - \pi\sqrt{-\Lambda_{\text{eff}}}r_- \right). \quad (40)$$

The equation (37) then takes the form

$$\delta\mathcal{M} - \Omega_H\delta\mathcal{J} - T_H\delta S = - \int_{\Sigma} \delta C_{\xi}. \quad (41)$$

We will consider the special situation that the perturbation vanishes near the internal boundary of the surface  $\Sigma$ , then equation (41) with  $\delta S = 0$  would hold for both extremal and non-extremal black holes. Noting (35)(36) and  $\delta S = 0$ , (41) turns out to be

$$(1 - 2\theta_L\mathcal{T} - 2\pi\theta_L\Omega_H)(\delta M - \Omega_H\delta J) - 2\pi\theta_L\Lambda_{\text{eff}}\left(\frac{r_+^2 - r_-^2}{r_+^2}\right)\delta J = - \int_{\Sigma} \delta C_{\xi}. \quad (42)$$

for BTZ-like black holes in the MB model.

The equations (41)(42) are derived from the Lagrangian without matter. It might be puzzling that the vacuum configuration could be perturbed without matter; however, this is physically possible since there are gravitational waves in the MB model with general couplings. In general,  $\delta M$  and  $\delta J$  should be understood as variations allowed mathematically in the parameter space, rather than consequences of certain physical evolutions. On the other hand, since we didn't enforce the linearized equations of motion to be satisfied, it should be expected that these equations could also be used for considering perturbed configurations due to matter contribution<sup>3</sup>. The right hand side of (42) would be related to the energy-momentum tensor of the matter. To see this explicitly, we first define the “energy-momentum 2-form”  $\Sigma_a$  and “spin current 2-form”  $\tau_a$  as follows:

$$\Sigma_a \equiv \frac{\delta L_M}{\delta e^a}, \quad \tau_a \equiv \frac{\delta L_M}{\delta \omega^a}. \quad (43)$$

The equations of motion with matter would be

$$E_a^{(e)} = -\Sigma_a, \quad E_a^{(\omega)} = -\tau_a. \quad (44)$$

---

<sup>3</sup>For the matter field, we impose Dirichlet condition on asymptotic AdS boundary, as conventionally used in AdS/CFT dictionary for black holes dual to thermal states in CFT. This choice will not affect the argument for WCCC as we only care about the matter that falls in.

Since  $\Sigma_a = \tau_a = 0$  in the background spacetime, from (31) we get

$$\delta C_\xi = (\mathbf{i}_\xi e^a) \wedge \delta \Sigma_a + (\mathbf{i}_\xi \omega^a) \wedge \delta \tau_a. \quad (45)$$

$\Sigma_a$  should be related to the conserved canonical energy-momentum tensor  $\Sigma_a^\mu$  defined by

$$\sqrt{-g} \Sigma_a^\mu \equiv \frac{\partial \mathcal{L}}{\partial e^a_\mu} = e_a^\mu \mathcal{L} - \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} D_a \psi, \quad (46)$$

in which  $\psi$  is the matter field,  $\mathcal{L}$  is the Lagrangian density of the matter related to  $L_M$  by  $L_M = \mathcal{L} d^3x$ , and  $D_a$  is the covariant derivative defined by  $D_a = e_a^\mu (\partial_\mu + \omega_\mu^{bc} f_{cb})$  where  $f_{ab}$  are the representations of the generators of the Lorentz group associated with  $\psi$ . From (46) we obtain

$$\Sigma_a = \frac{1}{2} \epsilon_{\mu\nu\lambda} \Sigma_a^\lambda dx^\mu \wedge dx^\nu. \quad (47)$$

Note that

$$\epsilon_{\mu\nu\lambda} = -3k_{[\mu} \hat{\epsilon}_{\nu\lambda]}, \quad (48)$$

in which  $k^\mu$  is the future-directed normal vector to the horizon, and  $\hat{\epsilon}$  is the volume element on the horizon. The first term on the right hand side of (45) then turns out to be

$$(\mathbf{i}_\xi e^a) \wedge \delta \Sigma_a = -\xi_\mu k_\nu \delta \Sigma^{\mu\nu} \sqrt{-\gamma} d^2x, \quad (49)$$

as  $\xi^\mu \propto k^\mu$ , the contribution of this term to the right hand side of equation (42) is non-negative if and only if the null energy condition of matter energy-momentum tensor  $\delta \Sigma_{\mu\nu}$  is satisfied:

$$k^\mu k^\nu \delta \Sigma_{\mu\nu} \geq 0. \quad (50)$$

For the second term on the right hand side of (45), our “spin current 2-form”  $\tau_a$  is related to the canonical spin angular momentum tensor  $\tau_{ab}^\mu$  defined by

$$\sqrt{-g} \tau_{ab}^\mu \equiv \frac{\partial \mathcal{L}}{\partial \omega_\mu^{ab}} = -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} f_{ab} \psi, \quad (51)$$

Comparing (51) with (43), we obtain

$$\tau_a = -\frac{1}{2} \epsilon_a^{bc} \epsilon_{\mu\nu\lambda} \tau_{bc}^\lambda dx^\mu \wedge dx^\nu, \quad (52)$$

hence the second term on the right hand side of (45) is reduced to

$$(\mathbf{i}_\xi \omega^a) \wedge \delta \tau_a = -(\xi^\sigma \omega^{ab}{}_\sigma) k_\lambda \delta \tau_{ab}^\lambda \sqrt{-\gamma} d^2x. \quad (53)$$

For axially symmetric stationary black holes, in general we have [33]

$$\mathrm{i}_\xi \omega^a|_{\mathcal{H}} = -\frac{1}{2} \kappa_{\mathrm{H}} \epsilon^a_{bc} n^{bc} + \mathrm{i}_\xi K^a|_{\mathcal{H}}, \quad (54)$$

in which  $n^{ab}$  is the binormal to the horizon and  $K^a$  is the dual contorsion 1-form defined by  $T^a = \epsilon^a_{bc} K^b \wedge e^c$ , satisfying the identity  $\omega^a = \tilde{\omega}^a + K^a$ . For BTZ-like black holes, (16) gives

$$K^a = \frac{\mathcal{T}}{\pi} e^a. \quad (55)$$

Using (54)(55), equation (53) turns out to be

$$(\mathrm{i}_\xi \omega^a) \wedge \delta \tau_a = \left( \kappa_{\mathrm{H}} n_{\mu\nu} + \frac{\mathcal{T}}{\pi} \epsilon_{\mu\nu}{}^\sigma \xi_\sigma \right) k_\lambda \delta \tau^{\mu\nu\lambda} \sqrt{-\gamma} d^2x. \quad (56)$$

The first term on the right hand side is vanishing for extremal black holes. We note that the sign of the second term could not be determined for torsional chiral gravity unless the spin angular momentum tensor satisfies  $\epsilon_{\mu\nu}{}^\sigma k_\sigma k_\lambda \delta \tau^{\mu\nu\lambda} \geq 0$ , of which the physical meaning is not clear yet for us.

Combining all the results above, we obtain the following linear variational identity for the BTZ-like black holes in the MB model, with the additional assumption that the perturbation  $\delta\phi$  vanishes near the internal boundary of  $\Sigma$

$$\begin{aligned} & \delta\mathcal{M} - \Omega_{\mathrm{H}} \delta\mathcal{J} \\ &= (1 - 2\theta_{\mathrm{L}} \mathcal{T} - 2\pi\theta_{\mathrm{L}} \Omega_{\mathrm{H}}) (\delta M - \Omega_{\mathrm{H}} \delta J) - 2\pi\theta_{\mathrm{L}} \Lambda_{\mathrm{eff}} \left( \frac{r_+^2 - r_-^2}{r_+^2} \right) \delta J \\ &= \int_{\Sigma} d^2x \sqrt{-\gamma} \left\{ \xi_\mu k_\nu \delta \Sigma^{\mu\nu} - \left( \kappa_{\mathrm{H}} n_{\mu\nu} + \frac{\mathcal{T}}{\pi} \epsilon_{\mu\nu}{}^\sigma \xi_\sigma \right) k_\lambda \delta \tau^{\mu\nu\lambda} \right\}. \end{aligned} \quad (57)$$

For extremal BTZ black holes with  $\kappa_{\mathrm{H}} = 0$  and  $r_+ = r_-$ , the above identity takes the following simpler form:

$$\begin{aligned} \delta\mathcal{M} - \Omega_{\mathrm{H}} \delta\mathcal{J} &= \left( 1 - 2\theta_{\mathrm{L}} \mathcal{T} - 2\pi\theta_{\mathrm{L}} \sqrt{-\Lambda_{\mathrm{eff}}} \right) (\delta M - \sqrt{-\Lambda_{\mathrm{eff}}} \delta J) \\ &= \int_{\Sigma} d^2x \sqrt{-\gamma} \left\{ \xi_\mu k_\nu \delta \Sigma^{\mu\nu} - \frac{\mathcal{T}}{\pi} \epsilon_{\mu\nu}{}^\sigma \xi_\sigma k_\lambda \delta \tau^{\mu\nu\lambda} \right\}. \end{aligned} \quad (58)$$

## 2.2 Second order variations

As pointed out in [15], for near-extremal black holes it is in general not sufficient to consider just the linear order variation due to Hubeny-type violations. We therefore construct further the second order variational identity. A second variation of equation (32) gives rise to

$$\mathcal{E}_\Sigma(\phi; \delta\phi) = \int_{\partial\Sigma} [\delta^2 Q_\xi - \mathrm{i}_\xi \delta \Theta(\phi, \delta\phi)] + \int_{\Sigma} \delta^2 C_\xi + \int_{\Sigma} \mathrm{i}_\xi (\delta E \wedge \delta\phi), \quad (59)$$

in which

$$\mathcal{E}_\Sigma(\phi; \delta\phi) \equiv \int_\Sigma \Omega(\phi, \delta\phi, \mathcal{L}_\xi \delta\phi) \quad (60)$$

is Wald's canonical energy of the off-shell perturbation  $\delta\phi$  on  $\Sigma$ . For the case that the background  $\phi$  is a stationary black hole solution and  $\xi$  is the horizon Killing field, the boundary contribution from infinity is simply

$$\int_\infty \delta^2 Q_\xi - i_\xi \delta\Theta(\phi, \delta\phi) = \delta^2 \mathcal{M} - \Omega_H \delta^2 \mathcal{J} \quad (61)$$

according to (38). The contribution from interior boundary would be vanishing if there's no perturbation in its neighborhood, as supposed before. Then equation (59) turns out to be

$$\delta^2 \mathcal{M} - \Omega_H \delta^2 \mathcal{J} = \mathcal{E}_\Sigma(\phi; \delta\phi) - \int_\Sigma i_\xi (\delta E \wedge \delta\phi) - \int_\Sigma \delta^2 C_\xi. \quad (62)$$

Noting (44)(47)(52), the integrand of the second term on the right hand side is evaluated to be

$$\begin{aligned} i_\xi (\delta E \wedge \delta\phi) &\equiv i_\xi \left( \delta E_a^{(e)} \wedge \delta e^a + \delta E_a^{(\omega)} \wedge \delta \omega^a \right) \\ &= \xi^\tau \Xi_{[\mu\nu\tau]} dx^\mu \wedge dx^\nu, \end{aligned} \quad (63)$$

in which

$$\Xi_{\mu\nu\tau} = -\frac{3}{2} \epsilon_{\mu\nu\lambda} \left( \delta \Sigma_a^\lambda \delta e^a_\tau + \delta \tau_{ab}^\lambda \delta \omega^{ab}_\tau \right). \quad (64)$$

Since  $\xi$  is tangent to the horizon, the pullback of (63) to the horizon vanishes, hence this term gives no contribution. From (31), it turns out that

$$\delta^2 C_\xi = \delta^2 \left\{ -d^2 x \sqrt{-\gamma} \left[ \xi_\mu k_\nu \Sigma^{\mu\nu} - \left( \kappa_H n_{\mu\nu} + \frac{\mathcal{T}}{\pi} \epsilon_{\mu\nu}{}^\sigma \xi_\sigma \right) k_\lambda \tau^{\mu\nu\lambda} \right] \right\}. \quad (65)$$

Substituting the above expression into (62) leads to the following identity for the second order variation:

$$\begin{aligned} \delta^2 \mathcal{M} - \Omega_H \delta^2 \mathcal{J} &= \mathcal{E}_\Sigma(\phi; \delta\phi) \\ &\quad + \delta^2 \int_\Sigma d^2 x \sqrt{-\gamma} \left\{ \xi_\mu k_\nu \Sigma^{\mu\nu} - \left( \kappa_H n_{\mu\nu} + \frac{\mathcal{T}}{\pi} \epsilon_{\mu\nu}{}^\sigma \xi_\sigma \right) k_\lambda \tau^{\mu\nu\lambda} \right\}. \end{aligned} \quad (66)$$

### 3 Gedanken experiment to destroy an extremal BTZ

We now consider our gedanken experiment to destroy a BTZ black hole along the line of Wald's proposals [4, 15]. In this section, we will deal with an extremal BTZ

black hole with mass parameter  $M$  and angular momentum parameter  $J$ . We wish to see if a naked singularity can be made via throwing matter into the extremal black hole. Without losing generality, we take our gravity theory as MB model, and then discuss its three limits, torsional chiral gravity, chiral gravity and three-dimensional Einstein gravity.

Considering a 1-parameter family of solutions  $\phi(\lambda)$ ,  $\phi_0 = \phi(0)$  is an extremal BTZ black hole, which is a vacuum solution in MB model. The existence of event horizon is determined by a function,

$$f(\lambda) = M(\lambda)^2 + \Lambda_{\text{eff}} J(\lambda)^2, \quad (67)$$

If  $f(\lambda) \geq 0$ , the spacetime is a BTZ black hole. If  $f(\lambda) < 0$ , it is a naked conical singularity and WCCC is violated. We now consider perturbations to the extremal black hole  $\phi_0$ . Then, to the first order in  $\lambda$ , we have

$$f(\lambda) = 2\lambda\sqrt{-\Lambda_{\text{eff}}}|J|(\delta M - \sqrt{-\Lambda_{\text{eff}}}\delta J) + \mathcal{O}(\lambda^2), \quad (68)$$

where we have used the extremal condition  $M = \sqrt{-\Lambda_{\text{eff}}}|J|$  to eliminate  $M$ . It is then evident that if  $\delta M < \sqrt{-\Lambda_{\text{eff}}}\delta J$ ,  $f(\lambda)$  can be negative.

We would like to see whether this sort of violation of WCCC is possible if we throw matter into the BTZ black hole in a certain way. Let  $\Sigma_0$  be an asymptotically AdS hypersurface which extends from the future horizon to the spatial infinity. We consider a perturbation  $\delta\phi$  whose initial data for both fields  $\delta e^a$  and  $\delta\omega^a$  on  $\Sigma_0$  vanish in the neighborhood of the intersection between  $\Sigma_0$  and the horizon. We assume that the initial data for the matter sources  $\delta\Sigma_{\mu\nu}$  and  $\delta\tau^{\mu\nu\lambda}$  also vanish in this neighborhood, and, only exist in a compact region of  $\Sigma_0$ . That is, we consider perturbations whose effects at sufficiently early times are negligibly small. To simplify the discussion, we only consider the case where, as we evolve the perturbation, all of the matter will fall through the horizon. Therefore, the whole evolutions of the matter source  $\delta\Sigma_{\mu\nu}$  and  $\delta\tau^{\mu\nu\lambda}$  stay in a shaded region as shown in Fig. 1. As matter falls in, we further define a hypersurface  $\Sigma$  in the following way—it starts on the future horizon in the region where the perturbation vanishes and extends along the future horizon till all matter falls into the horizon; then it becomes spacelike, approaches the spatial infinity and becomes asymptotically AdS. We denote the horizon portion of  $\Sigma$  as  $\mathcal{H}$ , and the spatial portion as  $\Sigma_1$ .

We now use the linear variational identity with vanishing inner boundary contributions (58) for this choice of  $\Sigma$ . As we will show later, this identity constrains



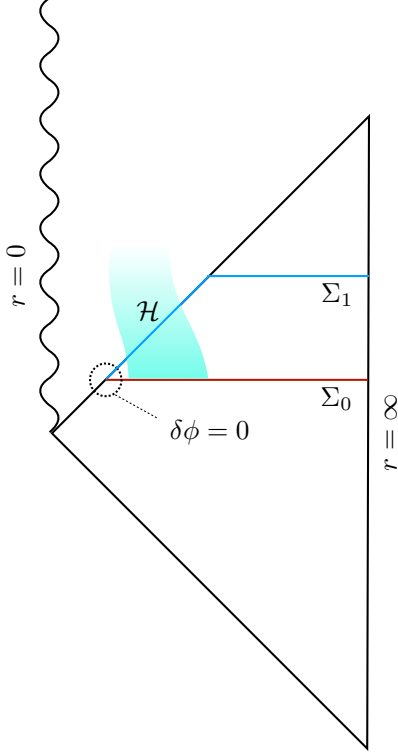


Figure 1: Carter-Penrose diagram of an extremal BTZ black hole. The shaded region consists of the falling matter which all goes into the black hole. The perturbation  $\delta\phi$  vanishes in a neighborhood of  $\Sigma_0 \cap \mathcal{H}$ .

the sign of  $f(\lambda)$ . We notice that in e.q. (58), the integral in the second line is not positive definite due to the spin angular momentum term and its coupling to torsion. That is, in torsional chiral gravity, whether WCCC can hold depends on an additional relation between the spin angular momentum and the torsion. The physical origin of this additional information needed is unclear, and is beyond our scope of this paper. We will leave it to a future work. In the torsionless limit  $\mathcal{T} \rightarrow 0$ , however, this integral would be non-negative as long as the null energy condition is satisfied. From now on, we will focus on this limit, and assume the falling matter satisfies the null energy condition. Then  $f(\lambda)$  is non-negative only if the constant factor on the rhs of the first line of e.q. (58) is non-negative,

$$1 - 2\theta_L \mathcal{T} - 2\pi\theta_L \sqrt{-\Lambda_{\text{eff}}} \geq 0. \quad (69)$$

For chiral gravity, we choose  $\theta_L = -1/(2\pi\sqrt{-\Lambda})$ , and send  $\mathcal{T} \rightarrow 0$ . The inequality (69) is then satisfied. Therefore extremal BTZ black hole in chiral gravity

cannot be destroyed in our experiment, and WCCC is preserved.

For 3-dimensional Einstein gravity with a negative cosmological constant, both torsion and Chern-Simons interaction vanish, thus we set  $\theta_L \rightarrow 0$  and  $\theta_T \rightarrow 0$ . The inequality (69) is then satisfied. Consequently, extremal BTZ black hole in 3-dimensional Einstein gravity cannot be destroyed, leaving WCCC preserved.

## 4 Gedanken experiment to destroy a near-extremal BTZ

For extremal BTZ black holes, we have found that WCCC can be violated in the presence of torsion. With torsion being turned off, we have seen that WCCC is preserved in both chiral gravity and 3-dimensional Einstein gravity, provided that the matter obeys the null energy condition. In 4-dimensional Einstein gravity, Hubeny [7] proposed that violations of WCCC might be possible if one threw matter into a near-extremal black hole in an appropriate manner. In order to examine whether Hubeny-type violations can truly happen, one has to calculate the energy and momentum of the matter beyond the linear order. In this section, we will examine the Hubeny-type violations for a near-extremal BTZ black hole in chiral gravity and 3-dimensional Einstein gravity respectively.

As shown in Fig. 2, we make similar choices of  $\Sigma_0$  and  $\Sigma$  like those for the extremal BTZ case. The only difference is that, the two hypersurfaces now terminate at the bifurcation surface  $B$ . We further assume that the second order perturbation  $\delta^2\phi$  for both fields  $\delta e^a$  and  $\delta\omega^a$  also vanishes in a neighborhood of  $B$ . Again, we simplify our discussions by restricting to the case where all matter falls into the black hole. We will also make the following additional assumption:

*Assumption:* The non-extremal BTZ black hole is linearly stable to perturbations, i.e., any source-free linear perturbation  $\delta\phi$  approaches a perturbation  $\delta\phi^{\text{BTZ}}$  towards another BTZ black hole at sufficiently late times.

Although our perturbations are not source-free in general, we will only apply the above assumption on the late-time spatial surface  $\Sigma_1$  long after all of the matter has fallen in the the black hole. We emphasize that this linear stability assumption does not indicate WCCC which we wish to prove. This is because a finite perturbation is needed to overspin a non-extremal black hole, while a linear perturbation can

always be scaled down to a infinite small one. Hence the linear instability of a non-extremal BTZ black hole should be independent of the instability associated with overspinning, i.e., a linearly stable non-extremal BTZ black hole could possibly be overspun by a finite perturbation, just like the situation for the Kerr-Newman black hole in [15].

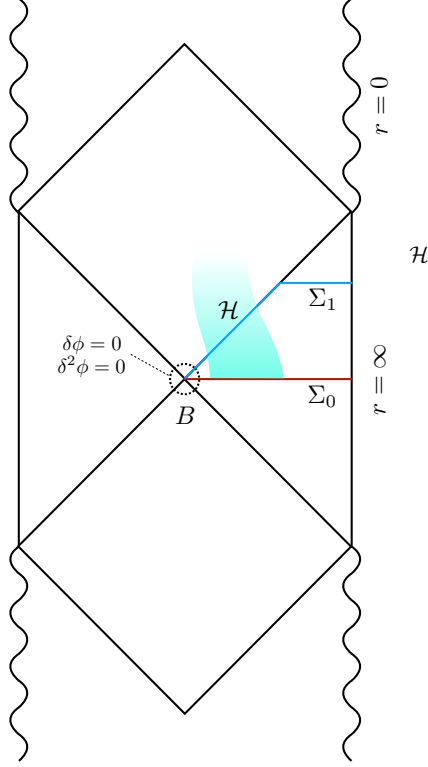


Figure 2: Carter-Penrose diagram of a near-extremal BTZ black hole. The shaded region consists of the falling matter which all goes into the black hole. The perturbation  $\delta\phi$  and  $\delta^2\phi$  vanishes in a neighborhood of  $B$ .

## 4.1 Chiral gravity

We now consider our thought experiment to destroy a near-extremal BTZ black hole  $(M, J)$  in chiral gravity for which  $\mathcal{T} = 0$  and  $\theta_L = -\frac{1}{2\pi\sqrt{-\Lambda}}$ . Thus, using (35-36) it is straightforward to see

$$\delta\mathcal{M} - \Omega_H\delta\mathcal{J} = \left(1 + \frac{\Omega_H}{\sqrt{-\Lambda}}\right) \left(\delta M - \sqrt{-\Lambda}\delta J\right), \quad (70)$$

and the first law of black hole thermodynamics yields

$$T_H \delta S = \delta \mathcal{M} - \Omega_H \delta \mathcal{J}. \quad (71)$$

where the black hole entropy is given by [33]

$$S = 4\pi (r_+ - r_-). \quad (72)$$

Recall (57), the null energy condition for the falling matter yields the first order relation that

$$\delta M \geq \sqrt{-\Lambda} \delta J. \quad (73)$$

Assuming the first order perturbation has been optimally done, i.e.  $\delta S = 0$ , such that

$$\delta M = \sqrt{-\Lambda} \delta J. \quad (74)$$

For some constant entropy  $S$ , we can then plot the line of constant entropy in the parameter space of BTZ black holes, which is shown in Fig. 3.

We are now ready to discuss our experiment to destroy the near-extremal BTZ black hole. Starting from a point  $(M_0, J_0)$  in the parameter space, after a perturbation of the spacetime as induced by falling matter, we will always arrive at another point  $(M_1, J_1)$ . At the linear order, the change from one point to another will correspond to a tangent vector in the parameter space. For any  $S$ , the line of constant entropy is given by

$$M = \left(\sqrt{-\Lambda}\right) J - \frac{\Lambda}{16\pi^2} S^2. \quad (75)$$

The slope of the constant entropy line is then equal to that of the line representing extremal BTZ black holes. Since the tangent to the constant entropy line is a lower bound to all physically-realizable perturbations, a non-extremal BTZ black hole will at most be perturbed to another BTZ black hole with the same entropy. There is no Hubeny-type violation of weak cosmic censorship for the BTZ black hole in three-dimensional chiral gravity, thus WCCC is preserved.

## 4.2 Einstein gravity

The discussions above can be applied to the BTZ black holes in 3-dimensional Einstein gravity as well, for which we turn off both torsion and Chern-Simons interactions in MB model. In this theory, the linear variational identity is given by

$$\delta \mathcal{M} - \Omega_H \delta \mathcal{J} = \delta M - \Omega_H \delta J. \quad (76)$$

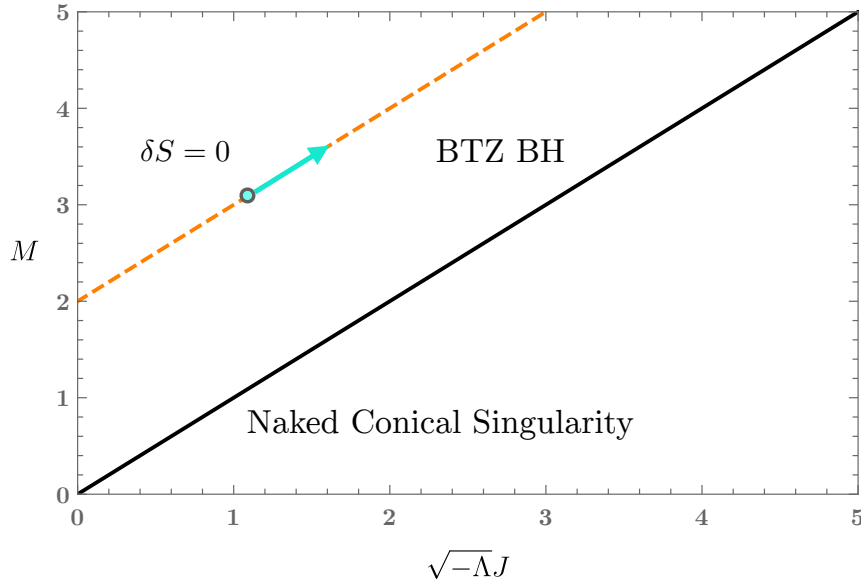


Figure 3: The parameter space of BTZ black holes in chiral gravity. The black solid line corresponds to extremal BTZ black holes. Any point above this line corresponds to a non-extremal BTZ black hole, while any point below the line is a naked conical singularity. The orange dashed line is one of the lines of constant entropy, which is parallel to the line for extremal BTZ black holes. Starting with some point on the constant entropy line, any tangent vector will always be parallel to the extremal BTZ line. That is, there is no Hubeny-type violation that can overspin a near-extremal BTZ black hole in chiral gravity.

Given the material null energy condition, we similarly find that

$$\delta M - \Omega_H \delta J \geq 0. \quad (77)$$

Once a first order perturbation is optimally done by choosing  $\delta M = \Omega_H \delta J$ , according to the first law of black hole thermodynamics, we will also find a lower bound for all perturbations given by  $\delta S = 0$ . In Einstein gravity, the entropy of the BTZ black hole is  $S = 4\pi r_+$ , and the curve of constant entropy is given by

$$M = \frac{4\pi^2}{S^2} J^2 - \frac{\Lambda}{16\pi^2} S^2. \quad (78)$$

We plot one of such curves in Fig. 4.

As shown in Fig. 4, if the initial spacetime is an extremal BTZ black hole, a tangent vector at this point is also tangent to the line representing extremal BTZ

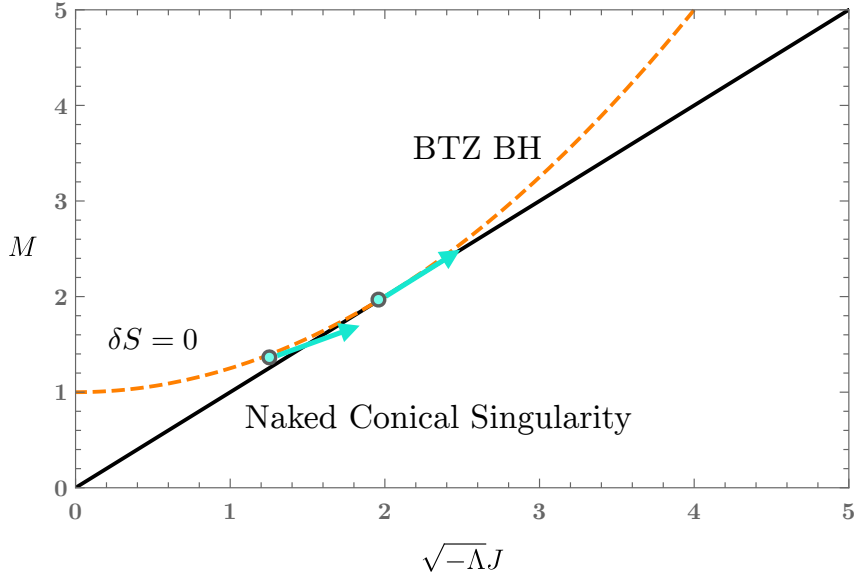


Figure 4: The parameter space of BTZ solutions in the 3-dimensional Einstein gravity. The black solid line corresponds to extremal BTZ black holes. Any point above this line corresponds to a non-extremal BTZ black hole, while any point below the line is a naked conical singularity. The orange dashed curve is one of the curves of constant entropy, which meets the extremal BTZ line tangentially. The tangent vector at the point of an extremal BTZ black hole will always bring it to another extremal BTZ black hole. However, starting from a slightly non-extremal BTZ black hole, to linear order, the tangent vector can perturb the spacetime to become a naked conical singularity.

solutions. Therefore given extremality, the best one can do is to deform the black hole to another extremal BTZ black hole. WCCC is then preserved and no naked singularities will form. However, if one starts at a slightly non-extremal BTZ black hole, the tangent to the curve of constant entropy is possible to move the original point to another point located in the region representing naked conical singularities. This type of violation of WCCC is exactly the Hubeny-type violation, which can be found at the linear order for near-extremal black holes. As we will see in the following discussions, a conclusive answer to whether this type of perturbations truly leads to a violation of WCCC requires calculations to the second order.

Now we consider a 1-parameter family of solutions  $\phi(\lambda)$ ,  $\phi_0 = \phi(0)$  is a nearly extremal BTZ black hole in 3-dimensional Einstein gravity. We then expand  $f(\lambda)$

in e.q. (67) to second order in  $\lambda$ ,

$$f(\lambda) = (M^2 - \alpha^2 J^2) + 2\lambda (M\delta M - \alpha^2 J\delta J) + \lambda^2 [(\delta M)^2 - \alpha^2 (\delta J)^2 + M\delta^2 M - \alpha^2 J\delta^2 J] + \mathcal{O}(\lambda^3), \quad (79)$$

where we have introduced a parameter  $\alpha = \sqrt{-\Lambda}$ . For convenience we also introduce a parameter  $\epsilon$  according to

$$\epsilon = \frac{r_+^2 - r_{+, \text{extremal}}^2}{r_{+, \text{extremal}}^2} = \frac{\sqrt{M^2 - \alpha^2 J^2}}{M}. \quad (80)$$

The background spacetime corresponds to  $\epsilon \ll 1$ , and  $\epsilon \rightarrow 0$  is the extremal limit. The null energy condition for the matter fields yields  $\delta M - \Omega_H \delta J \geq 0$ , which is equivalent to the statement that black hole entropy always increases. If we only consider perturbations to first order in  $\lambda$ , that entropy always increases will constrain  $f(\lambda)$  by

$$f(\lambda) \geq M^2 \epsilon^2 - 2\lambda \epsilon (\alpha^2 J \delta J) + \mathcal{O}(\lambda^2). \quad (81)$$

It is then evident from this inequality that, when  $\delta J \sim \epsilon M / \alpha$ , it is possible to make  $f(\lambda) < 0$  by some careful choice of  $\delta J$ . This is exactly the Hubeny-type violation of WCCC. The problem is that when  $\delta J \sim \epsilon M / \alpha$ , the violation of  $f(\lambda) \geq 0$  is of order  $M^2 \epsilon^2 \sim \alpha^2 (\delta J)^2$ , which is not fully captured to first order in  $\lambda$ . Therefore to determine whether there is a true violation of WCCC, one needs to calculate all quantities in e.q. (81) to the appropriate order.

We now consider the second order variations in order to give a bound for  $f(\lambda)$ . Given the null energy conditions for the falling matter, we can obtain the following relation from the second order variational identity with no inner boundary contributions (66),

$$\delta^2 M - \Omega_H \delta^2 J \geq \mathcal{E}_\Sigma(\phi; \delta\phi), \quad (82)$$

where the canonical energy  $\mathcal{E}_\Sigma$  is given by

$$\begin{aligned} \mathcal{E}_\Sigma(\phi; \delta\phi) &= \mathcal{E}_\mathcal{H}(\phi; \delta\phi) + \mathcal{E}_{\Sigma_1}(\phi; \delta\phi) \\ &= \int_{\mathcal{H}} \Omega(\phi, \delta\phi, \mathcal{L}_\xi \delta\phi) + \int_{\Sigma_1} \Omega(\phi, \delta\phi, \mathcal{L}_\xi \delta\phi). \end{aligned} \quad (83)$$

In (3+1)-dimension, the term  $\mathcal{E}_\mathcal{H}(\phi; \delta\phi)$  is identified as the total flux of gravitational wave energy into the black hole [16]. In (2+1)-dimensional Einstein gravity, however, there is no propagating degree of freedom in the bulk, i.e. there is no gravitational wave solution. Thus  $\mathcal{E}_\mathcal{H}(\phi; \delta\phi) = 0$ . A more rigorous way to see this

can be done by following the calculation of the canonical energy as in [16], and we similarly find that

$$\int_{\mathcal{H}} \Omega(\phi, \delta\phi, \mathcal{L}_\xi \delta\phi) = \frac{1}{4\pi} \int_{\mathcal{H}} (\kappa u) \delta\sigma_{ab} \delta\sigma^{ab} \hat{e} + \frac{1}{16\pi} \int_{\mathcal{H} \cap \Sigma_1} (\kappa u) \delta g^{ab} \delta\sigma_{ab} \hat{e}, \quad (84)$$

where  $\kappa$  is the surface gravity,  $u$  is an affine parameter on the future horizon,  $\delta\sigma_{ab}$  is the perturbed shear of the horizon generators, and  $\hat{e}$  is the volume element. In three dimension, it is found that every null geodesic congruence is shear-free [35], i.e.  $\sigma_{ab} = 0$ , therefore  $\delta\sigma_{ab} = 0$  on  $\mathcal{H}$  and the canonical energy on  $\mathcal{H}$  vanishes.

Then we only need to calculate the canonical energy on  $\Sigma_1$ . According to our assumption, the perturbation  $\delta\phi$ , as induced by the falling matter, approaches a perturbation  $\delta\phi^{\text{BTZ}}$  towards another BTZ black hole on  $\Sigma_1$ . Also since  $\delta\phi^{\text{BTZ}}$  has no gravitational wave energy through  $\mathcal{H}$ , we may replace  $\Sigma_1$  by  $\Sigma$  and obtain that

$$\mathcal{E}_{\Sigma_1}(\phi; \delta\phi) = \mathcal{E}_{\Sigma}(\phi; \delta\phi^{\text{BTZ}}). \quad (85)$$

We use the general second order variational identity (59) on this  $\Sigma$ . As before, we consider a one-parameter family of BTZ black holes,  $\phi^{\text{BTZ}}(\beta)$ . The black hole mass and angular momentum are given by  $M(\beta) = M + \beta\delta M^{\text{BTZ}}$  and  $J(\beta) = J + \beta\delta J^{\text{BTZ}}$ , where  $\delta M^{\text{BTZ}}$  and  $\delta J^{\text{BTZ}}$  are fixed by the first order perturbation for  $\phi(\lambda)$ . Therefore for this family of solutions, we have  $\delta^2 M = \delta^2 J = \delta E = \delta^2 C = 0$ . In Eq. (59), the only nonvanishing contribution in the evaluation of the canonical energy  $\mathcal{E}_{\Sigma}(\phi; \delta\phi^{\text{BTZ}})$  then comes from the integral over the bifurcation surface  $B$ , which yields

$$\mathcal{E}_{\Sigma_1}(\phi; \delta\phi) = -T_{\text{H}} \delta^2 S^{\text{BTZ}}. \quad (86)$$

Here, the minus sign is due to the fact that the bifurcation surface is the inner boundary of  $\Sigma$ .

With the canonical energy being calculated, (82) now reads

$$\delta^2 M - \Omega_{\text{H}} \delta^2 J \geq -T_{\text{H}} \delta^2 S^{\text{BTZ}}. \quad (87)$$

Here the temperature of the BTZ black hole is given by

$$T_{\text{H}} = -\frac{\Lambda(r_+^2 - r_-^2)}{2\pi r_+} = \frac{\alpha M \epsilon}{\pi \sqrt{2M(1+\epsilon)}}. \quad (88)$$

The second order variation of the black hole entropy is calculated as

$$\begin{aligned} \delta^2 S^{\text{BTZ}} = & (\delta J)^2 \left( -\frac{\pi \alpha M [\alpha^2 J^2 (3\epsilon + 2) + 2M^2 \epsilon^2 (\epsilon + 1)]}{\sqrt{2}\epsilon^3 [M^3(\epsilon + 1)]^{3/2}} \right) \\ & + (\delta J \delta M) \left( \frac{\pi \sqrt{2} \alpha J (\epsilon + 2)}{M \epsilon^3 \sqrt{M^3(\epsilon + 1)}} \right) + (\delta M)^2 \left( \frac{\pi (\epsilon - 2)(\epsilon + 1)}{\sqrt{2} \alpha \epsilon^3 \sqrt{M^3(\epsilon + 1)}} \right). \end{aligned} \quad (89)$$



where we have used the relation that for this family of solutions,  $\delta^2 M = \delta^2 J = 0$ . We assume that the first order perturbation is optimally done, i.e.  $\delta M = \Omega_H \delta J$ , and we use the inequality (87) to constrain  $f(\lambda)$  in e.q. (79). We obtain that

$$f(\lambda) \geq M^2 \epsilon^2 - 2\lambda \epsilon (\alpha^2 J \delta J) + \lambda^2 \frac{\alpha^4 J^2 (\delta J)^2}{M^2} + \mathcal{O}(\lambda^3, \epsilon \lambda^2, \epsilon^2 \lambda, \epsilon^3), \quad (90)$$

which can be further written as

$$f(\lambda) \geq \left( M\epsilon - \lambda \frac{\alpha^2 J \delta J}{M} \right)^2 + \mathcal{O}(\lambda^3, \epsilon \lambda^2, \epsilon^2 \lambda, \epsilon^3). \quad (91)$$

Consequently,  $f(\lambda) \geq 0$  when second order variations in  $\lambda$  are also taken into account. Our gedanken experiment cannot destroy a near-extremal BTZ black hole in three-dimensional Einstein gravity, thus WCCC is preserved.

## 5 Conclusions and Discussions

Along the line of Wald's proposals [4, 15] for 4D Einstein gravity, in this paper we have considered the gedanken experiments of destroying a BTZ black hole for three different limits of Mielke-Baekler (MB) model of 3D gravity. They are (i) Einstein gravity, (ii) chiral gravity and (iii) torsional chiral gravity. All three limits are free of perturbative ghosts and show different behaviors in the gedanken experiments. We find that there are Hubeny-type violations for Einstein gravity but none for chiral gravity when trying to destroy a nonextremal BTZ black hole. However, in these two theories, the WCCC holds for both extremal and nonextremal BTZ black holes if the falling matter obeys the null energy condition. It is philosophically interesting to see that WCCC prevails here even though the BTZ singularity is just a conical one.

On the other hand, for the torsional chiral gravity there is an additional contribution to the null energy condition from the spin angular momentum tensor even at the linear order of variations. Thus, that the WCCC will hold or not depends on the imposition of additional null energy-like condition for the spin angular momentum tensor. If WCCC does not hold for the first order variations, one needs to check the second order variation to see if there is Hubeny-type violation. However, the full formalism of deriving the second order variational identity for MB model is out of scope of this paper, and it deserves as a future work.

The third law of black hole dynamics was first proposed by Israel and a sketchy proof was also given [8], which states that one cannot turn a nonextremal black

hole into an extremal one by throwing the matter in a finite time interval. Later, the detailed proof was given by Sorce and Wald [15] as described and adopted in this paper. In the context of AdS/CFT correspondence, the temperature of the boundary CFT is the same as the Hawking temperature of the black hole in the bulk. Thus, our results in this paper can serve as an operational proof of thermodynamic third law by holographically mapping our gedanken experiments around a near-extremal BTZ black hole into the cooling processes of the boundary CFT toward zero temperature. Our generalization to BTZ black holes though seems straightforward, its implication to the third law of thermodynamics for holographic condensed matter systems is nontrivial and deserves further study. Especially, the generalization to the higher dimensional AdS black holes for more general gravity theories will give holographic tests of the third law of thermodynamics for the more realistic systems. We plan to attack this problem in the near future.

Before ending the paper, we comment on one more point about the proof of the third law by noticing that the equality of (91) holds for one particular choice of parameter  $\lambda$ . This implies that one can reach the extremal black hole at the second order for this particular case. To pin down the issue, one needs to check the third order of variation for this particular  $\lambda$  value. This is too involved to carry out just for a measure-zero possibility. However, it is still an interesting issue for the future work.

## Acknowledgements

BC acknowledges the support from the Brinson Foundation, the Simons Foundation (Award Number 568762), and the National Science Foundation, Grants PHY-1708212 and PHY-1708213. FLL is supported by Taiwan Ministry of Science and Technology through Grant No. 103-2112-M-003-001-MY3. BN is supported by the National Natural Science Foundation of China with Grant No. 11505119.

## References

- [1] R. Penrose, “Gravitational collapse: The role of general relativity,” *Riv. Nuovo Cim.* **1**, 252 (1969) [*Gen. Rel. Grav.* **34**, 1141 (2002)].
- [2] J. M. Maldacena, “The Large N limit of superconformal field theories and supergravity,” *Int. J. Theor. Phys.* **38**, 1113 (1999) [*Adv. Theor. Math. Phys.*

- 2**, 231 (1998)] doi:10.1023/A:1026654312961, 10.4310/ATMP.1998.v2.n2.a1 [hep-th/9711200].
- [3] See the story in <https://www.nytimes.com/1997/02/12/us/a-bet-on-a-cosmic-scale-and-a-concession-sort-of.html>
- [4] R. Wald, “Gedanken experiments to destroy a black hole,” [Ann. Phys. **82**, 548 - 556 (1974)].
- [5] G. Z. Toth, “Test of the weak cosmic censorship conjecture with a charged scalar field and dyonic Kerr-Newman black holes,” Gen. Rel. Grav. **44**, 2019 (2012) doi:10.1007/s10714-012-1374-z [arXiv:1112.2382 [gr-qc]].
- [6] J. Natario, L. Queimada and R. Vicente, “Test fields cannot destroy extremal black holes,” Class. Quant. Grav. **33**, no. 17, 175002 (2016) doi:10.1088/0264-9381/33/17/175002 [arXiv:1601.06809 [gr-qc]].
- [7] V. E. Hubeny, “Overcharging a black hole and cosmic censorship,” Phys. Rev. D **59**, 064013 (1999) doi:10.1103/PhysRevD.59.064013 [gr-qc/9808043].
- [8] W. Israel, “Third Law of Black-Hole Dynamics: A Formulation and Proof,” [Phys. Rev. Lett. **57**, no. 4, 397 (1986)]. doi:10.1103/PhysRevLett.57.397
- [9] F. de Felice and Y. Q. Yu, “Turning a black hole into a naked singularity,” Class. Quant. Grav. **18**, 1235 (2001). doi:10.1088/0264-9381/18/7/307
- [10] G. Chirco, S. Liberati and T. P. Sotiriou, “Gedanken experiments on nearly extremal black holes and the Third Law,” Phys. Rev. D **82**, 104015 (2010) doi:10.1103/PhysRevD.82.104015 [arXiv:1006.3655 [gr-qc]].
- [11] A. Saa and R. Santarelli, “Destroying a near-extremal Kerr-Newman black hole,” Phys. Rev. D **84**, 027501 (2011) doi:10.1103/PhysRevD.84.027501 [arXiv:1105.3950 [gr-qc]].
- [12] S. Gao and Y. Zhang, “Destroying extremal Kerr-Newman black holes with test particles,” Phys. Rev. D **87**, no. 4, 044028 (2013) doi:10.1103/PhysRevD.87.044028 [arXiv:1211.2631 [gr-qc]].
- [13] D. Chen, “Weak cosmic censorship conjecture in BTZ black holes with scalar fields,” arXiv:1812.03459 [gr-qc].
- [14] K. Dzta and M. Jamil, “Testing Cosmic Censorship Conjecture for Extremal and Near-extremal (2+1)-dimensional MTZ Black Holes,” arXiv:1808.04711 [gr-qc].

- [15] J. Sorce and R. M. Wald, “Gedanken experiments to destroy a black hole. II. Kerr-Newman black holes cannot be overcharged or overspun,” *Phys. Rev. D* **96**, no. 10, 104014 (2017) [arXiv:1707.05862 [gr-qc]].
- [16] S. Hollands and R. M. Wald, “Stability of Black Holes and Black Branes,” *Commun. Math. Phys.* **321**, 629 (2013) doi:10.1007/s00220-012-1638-1 [arXiv:1201.0463 [gr-qc]].
- [17] R. M. Wald, “Black hole entropy is the Noether charge,” *Phys. Rev. D* **48**, no. 8, R3427 (1993) doi:10.1103/PhysRevD.48.R3427 [gr-qc/9307038].
- [18] V. Iyer and R. M. Wald, “Some properties of Noether charge and a proposal for dynamical black hole entropy,” *Phys. Rev. D* **50**, 846 (1994) doi:10.1103/PhysRevD.50.846 [gr-qc/9403028].
- [19] T. Jacobson, G. Kang and R. C. Myers, “On black hole entropy,” *Phys. Rev. D* **49**, 6587 (1994) [gr-qc/9312023].
- [20] E. Barausse, V. Cardoso and G. Khanna, “Test bodies and naked singularities: Is the self-force the cosmic censor?,” *Phys. Rev. Lett.* **105**, 261102 (2010) doi:10.1103/PhysRevLett.105.261102 [arXiv:1008.5159 [gr-qc]].
- [21] E. Barausse, V. Cardoso and G. Khanna, “Testing the Cosmic Censorship Conjecture with point particles: the effect of radiation reaction and the self-force,” *Phys. Rev. D* **84**, 104006 (2011) doi:10.1103/PhysRevD.84.104006 [arXiv:1106.1692 [gr-qc]].
- [22] P. Zimmerman, I. Vega, E. Poisson and R. Haas, “Self-force as a cosmic censor,” *Phys. Rev. D* **87**, no. 4, 041501 (2013) doi:10.1103/PhysRevD.87.041501 [arXiv:1211.3889 [gr-qc]].
- [23] M. Colleoni and L. Barack, “Overspinning a Kerr black hole: the effect of self-force,” *Phys. Rev. D* **91**, 104024 (2015) doi:10.1103/PhysRevD.91.104024 [arXiv:1501.07330 [gr-qc]].
- [24] M. Colleoni, L. Barack, A. G. Shah and M. van de Meent, “Self-force as a cosmic censor in the Kerr overspinning problem,” *Phys. Rev. D* **92**, no. 8, 084044 (2015) doi:10.1103/PhysRevD.92.084044 [arXiv:1508.04031 [gr-qc]].
- [25] S. Deser, R. Jackiw and S. Templeton, “Topologically Massive Gauge Theories,” *Annals Phys.* **140**, 372 (1982) [*Annals Phys.* **281**, 409 (2000)] Erratum: [*Annals Phys.* **185**, 406 (1988)].

- [26] S. Deser, R. Jackiw and S. Templeton, “Three-Dimensional Massive Gauge Theories,” *Phys. Rev. Lett.* **48**, 975 (1982).
- [27] S. Deser and B. Tekin, “Massive, topologically massive, models,” *Class. Quant. Grav.* **19**, L97 (2002) [hep-th/0203273].
- [28] E. W. Mielke and P. Baekler, *Phys. Lett. A* **156**, 399 (1991).  
doi:10.1016/0375-9601(91)90715-K
- [29] P. Baekler, E. W. Mielke and F. W. Hehl, “Dynamical symmetries in topological 3-D gravity with torsion,” *Nuovo Cim. B* **107**, 91 (1992).
- [30] F. W. Hehl, P. Von Der Heyde, G. D. Kerlick and J. M. Nester, “General Relativity with Spin and Torsion: Foundations and Prospects,” *Rev. Mod. Phys.* **48**, 393 (1976).
- [31] W. Li, W. Song and A. Strominger, “Chiral Gravity in Three Dimensions,” *JHEP* **0804**, 082 (2008) [arXiv:0801.4566 [hep-th]].
- [32] A. A. Garcia, F. W. Hehl, C. Heinicke and A. Macias, “Exact vacuum solution of a (1+2)-dimensional Poincare gauge theory: BTZ solution with torsion,” *Phys. Rev. D* **67**, 124016 (2003) [gr-qc/0302097].
- [33] B. Ning and C. H. Wei, “Quasi-local Energy in 3D Gravity with Torsion,” arXiv:1807.08736 [hep-th].
- [34] E. D’Hoker and P. Kraus, “Charged Magnetic Brane Solutions in AdS (5) and the fate of the third law of thermodynamics,” *JHEP* **1003**, 095 (2010)  
doi:10.1007/JHEP03(2010)095 [arXiv:0911.4518 [hep-th]].
- [35] D. D. K. Chow, C. N. Pope and E. Sezgin, *Class. Quant. Grav.* **27**, 105002 (2010) doi:10.1088/0264-9381/27/10/105002 [arXiv:0912.3438 [hep-th]].