

# Gedanken experiments to destroy a black hole by a test particle: Multiply charged black hole with higher derivative corrections

Keisuke Izumi<sup>1,2</sup>, Toshifumi Noumi,<sup>3</sup> and Daisuke Yoshida<sup>2</sup>

<sup>1</sup>*Kobayashi-Maskawa Institute, Nagoya University, Nagoya 464-8602, Japan*

<sup>2</sup>*Department of Mathematics, Nagoya University, Nagoya 464-8602, Japan*

<sup>3</sup>*Graduate School of Arts and Sciences, The University of Tokyo, Komaba, Meguro-ku, Tokyo 153-8902, Japan*

 (Received 15 April 2024; accepted 2 July 2024; published 2 August 2024)

We investigate a gedanken experiment to destroy an extremally charged black hole by dropping a test particle, provided that there are multiple  $U(1)$  gauge fields coupled with each other through higher derivative interactions. It is known that the Coulomb repulsion prevents a test particle that would break the extremal condition from falling into an extremal black hole and therefore the black hole cannot be destroyed. We extend this observation to include higher derivative corrections. Although the extremal condition is modified by the higher derivative interactions, we find that the repulsive force induced by the higher derivative couplings is responsible for preventing a test particle that would break the modified extremal condition to reach the event horizon. Thus, we confirm that the weak cosmic censorship conjecture holds for extremally charged black holes even in the presence of higher derivative corrections, as long as the test particle approximation is justified.

DOI: [10.1103/PhysRevD.110.044008](https://doi.org/10.1103/PhysRevD.110.044008)

## I. INTRODUCTION

The weak cosmic censorship conjecture [1] states that all of the spacetime singularities formed by the gravitational collapse must be hidden by event horizons. This conjecture is supported by many gedanken experiments. In particular, gedanken experiments trying to destroy a black hole by dropping a test particle have been actively investigated so far. The pioneering work along this direction is Ref. [2], where the motion of the test particle on the extremal Kerr-Newman spacetime is investigated. There, it was found that a test particle that would overcharge and/or overspin the extremal Kerr-Newman black hole is prevented to reach the event horizon due to the Coulomb repulsion, centrifugal force, and spin-spin interactions. Thus, such a test particle cannot destroy the black hole. Following this work, the analysis is extended to other kinds of black holes [3–11]. See also Refs. [12–18], where the violation of the weak cosmic censorship by a test particle dropping to a nearly extremal black hole is discussed. In this paper, we investigate gedanken experiments along the same direction but with including the effective field theory corrections to Einstein gravity with two gauge fields.

If we regard the classical gravitational theory as a low-energy effective field theory, it is natural to expect the presence of higher derivative corrections. For example, starting from the quantum electrodynamics with gravity, massive heavy charged fields generate the higher derivative corrections in the low energy effective action below their mass scale [19]. Implications of such correction terms in the low-energy effective theory are key bypaths to reveal the nature of the high-energy physics. With the presence of higher derivative corrections, the extremal condition of black holes is modified (see, e.g., [20–23]), because the electrovacuum solutions have corrections from the Reissner-Nordström solution. This modification would have the information of the high-energy physics.

In this paper, we investigate whether such an extremal black hole could be overcharged by a test particle. We treat an extremally charged black hole in the theory of Einstein gravity with two  $U(1)$  gauge fields including higher derivative corrections. We use the test particle approximation, which is valid if the backreaction to the test particle by itself is negligible. Hence, this approximation is trustable in the first order in the charge of the test particle. Within this approximation, we confirm that any charged test particle can never destroy an extremal black hole, that is, the weak cosmic censorship conjecture holds.

If the charges of the black hole and the particle are associated with different gauge fields, no Coulomb force acts between them. It implies that the first order in the charge of the test particle vanishes. However, the higher-order

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP<sup>3</sup>.*

contributions are required to be taken into account. The effects beyond the test particle approximation are important even in the case with single  $U(1)$  gauge field, as discussed in Refs. [14,24,25]. In particular, the presence of the repulsive self-force for a charged particle around the Kerr spacetime is demonstrated by Ref. [24]. In addition, the Sorce-Wald formalism [25], which does not rely on the test point particle approximation, enables us to include such effects systematically as the second order increase of the mass of the resultant black hole. The second-order effects for the subextremal Kerr-Newman black hole is investigated in the original paper [25], and it is found that the potential violation of the weak cosmic censorship discussed in Refs. [12–18] is prevented by the second-order effects. Then, the formalism and its higher order extension are studied in Refs. [26–33]. We note also that the gedanken experiments to destroy the extremal black hole with higher derivative corrections with the single gauge field is investigated in Refs. [31,32] based on the Sorce-Wald formalism at the first order in a charge. See also Ref. [33] for the higher-order analysis for the subextremal black hole with higher derivative corrections. The higher-order contributions in the particle charges are not addressed in this paper, but similar results are expected to be obtained.

This paper is organized as follows. In the next section, we show the dynamics of a test charged particle on the static, spherically symmetric spacetime and electric fields. Then, we give the gedanken experiments for the black hole with minimally coupled gauge fields in Sec. III. In Sec. IV, we generalize the analysis to the case with higher derivative corrections. The final section is devoted to the summary and discussion. We describe the detailed equations in Appendix A. In Appendix B, we confirm that the results do not change under the field redefinition. Throughout the paper, we represents two  $U(1)$  gauge fields by  $A_\mu$  and  $B_\mu$ . The field strength, the charge of the particle, and the charge of the black hole with respect to each gauge fields are summarized in Table I. We use the unit  $\hbar = 1$ ,  $c = 1$ , and  $\epsilon_0 = 1$ , where  $\hbar$ ,  $c$ , and  $\epsilon_0$  are the reduced Planck constant, the speed of light, and the permittivity of vacuum.

## II. MOTION OF TEST PARTICLE WITH MULTIPLE CHARGES

In this section we summarize the dynamics of a test charged particle on the static and spherically symmetric spacetime with two  $U(1)$  gauge fields  $A_\mu$  and  $B_\mu$ . The particle that we consider has mass  $m$  and charges  $q$  and  $p$

TABLE I. Summary of the notation.

| Gauge field | Field strength | Electrostatic potential | Particle charge | Black hole charge |
|-------------|----------------|-------------------------|-----------------|-------------------|
| $A_\mu$     | $F_{\mu\nu}$   | $\Phi$                  | $q$             | $Q$               |
| $B_\mu$     | $H_{\mu\nu}$   | $\Psi$                  | $p$             | $P$               |

with respect to the gauge fields  $A_\mu$  and  $B_\mu$ , respectively. Suppose that the particle is minimally coupled with gravity and have no derivative coupling with the background fields. This setup is justified in our analysis where the lowest-order corrections with higher derivative couplings are taken into account, since they are removed by the field redefinition.<sup>1</sup> The action of the test particle is

$$S^{\text{test}}[x^\mu] = \int d\tau \left[ -m \sqrt{-g_{\mu\nu}(x) \dot{x}^\mu \dot{x}^\nu} + q A_\mu(x) \dot{x}^\mu + p B_\mu(x) \dot{x}^\mu \right], \quad (1)$$

where  $x^\mu(\tau)$  is the worldline of the test particle and the dot represents the derivative with respect to  $\tau$ . This action has the symmetry under the reparametrization of  $\tau$ . We fix this degree of freedom by setting  $\tau$  to be the proper time, i.e.,  $g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = -1$ . Then, the equation of motion for the test particle reads,

$$m u^\nu \nabla_\nu u^\mu = q F^\mu{}_\nu u^\nu + p H^\mu{}_\nu u^\nu, \quad (2)$$

$$u_\mu u^\mu = -1, \quad (3)$$

where  $u^\mu$  is the four-velocity of the particle, that is,  $u^\mu := \dot{x}^\mu$ . Here,  $F_{\mu\nu}$  and  $H_{\mu\nu}$  are the field strength of  $A_\mu$  and  $B_\mu$ , defined by  $F_{\mu\nu} := \partial_\mu A_\nu - \partial_\nu A_\mu$  and  $H_{\mu\nu} := \partial_\mu B_\nu - \partial_\nu B_\mu$ .

Let us focus on the static spherically symmetric case, and denote the static Killing vector by  $\xi^\mu$ . Then, we introduce the static time coordinate  $t$  by  $\xi^\mu \partial_\mu = \partial_t$ . With the static time  $t$  and the areal radius  $r$ , the static spherically symmetric metric can be described as

$$g_{\mu\nu} dx^\mu dx^\nu = -f(r) dt^2 + \frac{h(r)}{f(r)} dr^2 + r^2 d\Omega^2, \quad (4)$$

where  $d\Omega^2 := d\theta^2 + \sin^2 \theta d\phi^2$  is the metric on the unit two sphere. In asymptotically flat spacetime, the event horizon  $r_H$  is located at the largest positive root of  $f(r)$ , if it exists. We focus only on the outside of the black hole, that is, the region  $r \in (r_H, \infty)$ , where  $f(r)$  and  $h(r)$  are assumed to be positive. The asymptotic flatness implies  $f(r) \rightarrow 1$  and  $h(r) \rightarrow 1$  in the limit  $r \rightarrow \infty$ .

Suppose that  $U(1)$  electric fields  $A_\mu$  and  $B_\mu$  also enjoy the static, spherically symmetric properties, and then, they take the following form of vector potentials:

$$A_\mu dx^\mu = -\Phi(r) dt, \quad (5)$$

$$B_\mu dx^\mu = -\Psi(r) dt. \quad (6)$$

<sup>1</sup>The direct analysis with keeping the lowest-order corrections is shown in Appendix B 2.

Here,  $\Phi$  and  $\Psi$  are the electrostatic potentials for  $A_\mu$  and  $B_\mu$ , respectively. Note that they are manifestly static, that is,  $(\xi_\xi A)_\mu = 0$  and  $(\xi_\xi B)_\mu = 0$ . We also assume that the electrostatic potentials vanish at asymptotic infinity, i.e.,  $\Phi(r) \rightarrow 0$  and  $\Psi(r) \rightarrow 0$  in the limit  $r \rightarrow \infty$ . The corresponding field strengths are

$$\frac{1}{2}F_{\mu\nu}dx^\mu \wedge dx^\nu = \Phi'(r)dt \wedge dr, \quad (7)$$

$$\frac{1}{2}H_{\mu\nu}dx^\mu \wedge dx^\nu = \Psi'(r)dt \wedge dr, \quad (8)$$

where the prime represents the derivative with respect to  $r$ .

Due to the static symmetry, the energy of the charged particle defined by

$$E := (-mg_{\mu\nu}u^\nu - qA_\mu - pB_\mu)\xi^\mu, \quad (9)$$

is conserved along the worldline, which implies  $u^\mu \nabla_\mu E = 0$ .

Let us investigate a test particle moving along the radial direction, starting from the asymptotic infinity, that is,  $r(\tau) \rightarrow \infty$  at the infinite past  $\tau \rightarrow -\infty$ . We denote the four-velocity  $u^\mu$  by

$$u^\mu \partial_\mu = \dot{t}(\tau)\partial_t + \dot{r}(\tau)\partial_r, \quad (10)$$

with  $\dot{t} > 0$  and  $\dot{r} < 0$ . Then, the proper time condition and the conserved energy can be written as

$$-1 = -f(r)\dot{t}^2 + \frac{h(r)}{f(r)}\dot{r}^2, \quad (11)$$

$$E = mf(r)\dot{t} + q\Phi(r) + p\Psi(r). \quad (12)$$

By eliminating  $\dot{t}$  from the above equations we obtain,

$$E = m\sqrt{f + h\dot{r}^2} + q\Phi + p\Psi. \quad (13)$$

Due to the energy conservation, the value of  $E$  can be fixed in the asymptotic region as

$$\lim_{r \rightarrow \infty} E = m\sqrt{1 + \dot{r}^2}, \quad (14)$$

and hence we focus on the case  $E \geq m$ . By solving Eq. (13) for  $\dot{r}^2$  we obtain,

$$\begin{aligned} \dot{r}^2 &= \frac{1}{m^2 h(r)} ((q\Phi + p\Psi - E)^2 - m^2 f(r)), \\ &= \frac{1}{m^2 h(r)} (E - V_+(r))(E - V_-(r)), \end{aligned} \quad (15)$$

where  $V_\pm$  are defined by

$$V_\pm(r) := q\Phi(r) + p\Psi(r) \pm m\sqrt{f(r)}. \quad (16)$$

From Eqs. (13) and (16), we find that the positivity of  $\dot{r}^2$  requires

$$E \geq V_+(r). \quad (17)$$

This inequality indicates the allowed region of motion for any given energy  $E \geq m$ . Thus,  $V_+(r)$  can be regarded as the effective potential of this system. In particular, we obtain a necessary condition for the particle to reach the event horizon  $r_H$ , which we call the *falling condition*, by

$$E \geq V_+(r_H) = q\Phi(r_H) + p\Psi(r_H), \quad (18)$$

where  $V_+(r_H)$  can be understood as the  $r \rightarrow r_H$  limit of  $V(r)$  and so on. This can be regarded as the condition among  $E$ ,  $p$ , and  $q$  for the test particle falling into the black hole. In the following section, we will compare this condition with the extremal condition of the total system.

### III. BLACK HOLE WITH TWO MINIMALLY COUPLED GAUGE FIELDS

In this section, we investigate the gedanken experiments to destroy a black hole in the theory of Einstein gravity with two minimally coupled  $U(1)$  gauge fields without higher derivative interactions between the gauge fields.

#### A. Static spherically symmetric solution

We consider the Einstein–Hilbert action with two minimally coupled  $U(1)$  gauge fields,

$$S = \int d^4x \sqrt{-g} \left[ \frac{1}{16\pi G} R - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} H_{\mu\nu} H^{\mu\nu} \right]. \quad (19)$$

In general, there exists a mixing of the kinetic terms for  $A_\mu$  and  $B_\mu$ , see e.g. Ref. [34], which can be diagonalized by the redefinition of  $A_\mu$  and  $B_\mu$ . We assume that our gauge fields  $A_\mu$  and  $B_\mu$  are such diagonalized ones after the field redefinition.<sup>2</sup>

The Einstein equation is

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = 8\pi GT_{\mu\nu}, \quad (20)$$

where  $T_{\mu\nu}$  is given by

$$\begin{aligned} T_{\mu\nu} &= F_{\mu\rho}F_{\nu}{}^\rho - \frac{1}{4}g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} \\ &\quad + H_{\mu\rho}H_{\nu}{}^\rho - \frac{1}{4}g_{\mu\nu}H_{\rho\sigma}H^{\rho\sigma}. \end{aligned} \quad (21)$$

<sup>2</sup>It is pointed out in Ref. [34] that in theories with two  $U(1)$  gauge fields, their charges can be irrational numbers after the renormalization. We do not get into this issue, since it is irrelevant to our discussion.

The Maxwell equations for each gauge field are

$$\nabla_\nu F^{\nu\mu} = 0, \quad (22)$$

$$\nabla_\nu H^{\nu\mu} = 0. \quad (23)$$

By solving the above equations with the static, spherically symmetric ansatz (4)–(6), we obtain the Reissner-Nordström solution as usual,

$$f(r) = 1 - \frac{2GM}{r} + \frac{G(Q^2 + P^2)}{4\pi r^2}, \quad (24)$$

$$h(r) = 1, \quad (25)$$

$$\Phi(r) = \frac{Q}{4\pi r}, \quad (26)$$

$$\Psi(r) = \frac{P}{4\pi r}. \quad (27)$$

The roots of  $f(r) = 0$  are

$$r = G \left( M \pm \sqrt{M^2 - \frac{Q^2 + P^2}{4\pi G}} \right). \quad (28)$$

Then, in terms of the extremal mass  $\bar{M}_{\text{ext}}(Q, P)$  defined by<sup>3</sup>

$$\bar{M}_{\text{ext}}(Q, P) := \sqrt{\frac{Q^2 + P^2}{4\pi G}}, \quad (29)$$

the condition for the event horizon to exist, or in other words, the subextremal condition, can be written as

$$M \geq \bar{M}_{\text{ext}}(Q, P). \quad (30)$$

The equality holds for the extremal black hole. The areal radius of the event horizon of the extremal black hole becomes

$$\bar{r}_{\text{H}} = G\bar{M}_{\text{ext}}(Q, P). \quad (31)$$

On the other hand, the condition for the absence of the event horizon, say, the super-extremal condition, can be written as

$$M < \bar{M}_{\text{ext}}(Q, P). \quad (32)$$

<sup>3</sup>The bar of  $\bar{M}_{\text{ext}}(Q, P)$  indicates that the extremal mass for given  $Q$  and  $P$  are evaluated in the minimally coupled theory (19). We will use similar notation for other quantities as well.

## B. Gedanken experiments to destroy an extremal black hole: Single minimally coupled gauge field

Let us first review the Gedanken experiments to destroy an extremally charged black hole with a single gauge field, which is the case with  $P = 0$  and  $p = 0$  in our analysis. This corresponds to the nonrotating version of the analysis on the Kerr-Newman spacetime in Ref. [2].

Let us start with the extremal black hole with  $P = 0$  as a background spacetime

$$(M, Q, P) = (\bar{M}_{\text{ext}}(Q, 0), Q, 0), \quad (33)$$

and introduce a charged test particle with  $p = 0$

$$(E, q, p) = (E, q, 0), \quad (34)$$

where  $\bar{M}_{\text{ext}}(Q, 0)$  is simplified as

$$\bar{M}_{\text{ext}}(Q, 0) = \frac{|Q|}{\sqrt{4\pi G}}. \quad (35)$$

In the test particle approximation,  $E \ll M$  and  $|q| \ll |Q|$  are assumed. The event horizon of the background spacetime is located at  $\bar{r}_{\text{H}} = G\bar{M}_{\text{ext}}(Q, 0)$ .

Let us confirm the subextremal condition for the total system first. The total energy  $\hat{M}$  and charges  $\hat{Q}, \hat{P}$  in this system are

$$(\hat{M}, \hat{Q}, \hat{P}) = (\bar{M}_{\text{ext}}(Q, 0) + E, Q + q, 0). \quad (36)$$

Suppose that this system is settled in a static spacetime and no energy escapes to infinity. The uniqueness theorem states that if the resultant spacetime is a black hole solution, on the one hand, it must be the Reissner-Nordström spacetime with  $(\hat{M}, \hat{Q}, \hat{P})$ , which satisfies the subextremal condition. On the other hand, if the super-extremal condition

$$\hat{M} - \bar{M}_{\text{ext}}(\hat{Q}, \hat{P}) = E - \text{sgn}(Q) \frac{q}{\sqrt{4\pi G}} + \mathcal{O}(q^2) < 0 \quad (37)$$

is satisfied, the resultant object cannot be a black hole, and thus, the black hole will be destroyed.

The falling condition is investigated as follows. Suppose that the particle will be captured by the black hole. Then, the inequality (18) must be satisfied, that is, the falling condition (18) reduces to

$$E \geq q \frac{Q}{4\pi r_{\text{H}}} = \text{sgn}(Q) \frac{q}{\sqrt{4\pi G}}. \quad (38)$$

The super-extremal condition (37) and the falling condition (38) cannot be satisfied at the same time. This leads us to the conclusion that any particle which would destroy the extremal black hole cannot be captured by the black hole.



### C. Gedanken experiments to destroy an extremal black hole: Two minimally coupled gauge fields

Let us move on to the discussion on general cases including  $P \neq 0$  and  $p \neq 0$ . The extremal condition is given by

$$(M, Q, P) = (\bar{M}_{\text{ext}}(Q, P), Q, P). \quad (39)$$

We introduce a test particle with charge  $(q, p)$ , and assume  $E \ll M$  and  $|q|, |p| \ll \sqrt{Q^2 + P^2}$ .

The super-extremal condition for the total system is obtained in the following discussion. The total energy  $\hat{M}$  and charges  $\hat{Q}$  and  $\hat{P}$  of this system are given by

$$(\hat{M}, \hat{Q}, \hat{P}) = (\bar{M}_{\text{ext}}(Q) + E, Q + q, P + p). \quad (40)$$

Assuming  $|q|, |p| \ll \sqrt{Q^2 + P^2}$ , we can evaluate the super-extremal condition as

$$\begin{aligned} \hat{M} - \bar{M}_{\text{ext}}(\hat{Q}, \hat{P}) \\ = E - \frac{1}{\sqrt{4\pi G}} \frac{qQ + pP}{\sqrt{Q^2 + P^2}} + \mathcal{O}(p^2, pq, q^2) < 0. \end{aligned} \quad (41)$$

This inequality is rewritten as

$$E < \frac{1}{\sqrt{4\pi G}} \frac{qQ + pP}{\sqrt{Q^2 + P^2}} + \mathcal{O}(p^2, pq, q^2). \quad (42)$$

Then, let us investigate the falling condition (18). It can be written as

$$E \geq \frac{qQ + pP}{4\pi\bar{r}_H} = \frac{1}{\sqrt{4\pi G}} \frac{qQ + pP}{\sqrt{Q^2 + P^2}}. \quad (43)$$

The super-extremal condition (42) and the falling condition (43) cannot be satisfied at the same time, which is the same result as that in the single gauge field case. Our purpose of this paper is to verify whether the same structure holds even if we include the higher derivative corrections. We investigate it in the next section.

Before closing this section, we would like to mention on the special parameter choice,

$$qQ + pP = 0, \quad (44)$$

which means that the charges of the black hole  $(Q, P)$  are orthogonal to those of the particle  $(q, p)$  and no Coulomb force acts between the black hole and the particle. Without loss of generality, let us focus on the case with  $P = 0$  and  $q = 0$ . Then, the super-extremal condition starts from the second order in  $p^2$ ,

$$\hat{M} - \bar{M}_{\text{ext}}(\hat{Q}, \hat{P}) = E - \frac{1}{2} \frac{|Q|}{\sqrt{4\pi G}} \frac{p^2}{Q^2} + \mathcal{O}(p^4) < 0. \quad (45)$$

Thus, a naive discussion may lead to the result that the condition for the event horizon to be destroyed can be written as

$$E < \frac{1}{2} \frac{|Q|}{\sqrt{4\pi G}} \frac{p^2}{Q^2}. \quad (46)$$

However, no Coulomb force seems to act on the particle, so that the falling condition may reduce to

$$E \geq 0, \quad (47)$$

which is trivially satisfied for  $E \geq m$ . Thus, the conditions (46) and (47), as well as  $E \geq m$ , can be satisfied at the same time for a test particle with the energy

$$m \leq E < \frac{1}{2} \frac{|Q|}{\sqrt{4\pi G}} \frac{p^2}{Q^2}. \quad (48)$$

However, it is too early to conclude that this is a counterexample of the cosmic censorship conjecture: Eq. (45) indicates that the potential violation of the extremal condition is at most the order  $p^2$ , whereas the test particle approximation is valid at the order  $p^1$ . A careful analysis beyond the test particle approximation is therefore required. Indeed, a general formalism treating the effects beyond the test particle approximation was developed in Ref. [25] for the subextremal Kerr-Newman black hole with a single gauge field, showing that those subextremal black holes cannot be overcharged or overspun. Its generalization to the extremal black hole will be necessary to make a conclusion if the cosmic censorship conjecture is satisfied or not in this setup. We will not address this issue further in the present paper. Instead, we will see that the weak cosmic censorship conjecture holds at the level of the test particle approximation, that is, at the order  $p^1$ , with the higher derivative corrections included.

## IV. HIGHER DERIVATIVE CORRECTIONS

In this section, we extend the discussion of the gedanken experiments to the theory with the higher derivative corrections up to four derivatives. We derive the super-extremal conditions (Sec. IV A) and the falling conditions (Sec. IV B) on black holes with two charges with the higher derivative corrections, and then, show that no overlap exists between the two.

### A. Super-extremal conditions

Based on the spirit of the effective field theory approach, we treat all the possible higher derivative corrections up to four derivatives. More explicitly, we consider

$$\begin{aligned}
S = \int d^4x \sqrt{-g} & \left( \frac{1}{16\pi G} R - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{4} H_{\mu\nu} H^{\mu\nu} \right. \\
& + \alpha_1 F_{\mu\nu} F^{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + \alpha_2 F_{\mu}{}^{\nu} F_{\nu}{}^{\rho} F_{\rho}{}^{\sigma} F_{\sigma}{}^{\mu} + \alpha_3 R F_{\mu\nu} F^{\mu\nu} + \alpha_4 R_{\mu\nu} F^{\mu\rho} F^{\nu}{}_{\rho} + \alpha_5 R_{\mu\nu\rho\sigma} F^{\mu\nu} F^{\rho\sigma} \\
& + \beta_1 F_{\mu\nu} F^{\mu\nu} F_{\rho\sigma} H^{\rho\sigma} + \beta_2 F_{\mu}{}^{\nu} F_{\nu}{}^{\rho} F_{\rho}{}^{\sigma} H_{\sigma}{}^{\mu} + \beta_3 R F_{\mu\nu} H^{\mu\nu} + \beta_4 R_{\mu\nu} F^{\mu\rho} H^{\nu}{}_{\rho} + \beta_5 R_{\mu\nu\rho\sigma} F^{\mu\nu} H^{\rho\sigma} \\
& + \gamma_{11} F_{\mu\nu} F^{\mu\nu} H_{\rho\sigma} H^{\rho\sigma} + \gamma_{12} F_{\mu\nu} H^{\mu\nu} F_{\rho\sigma} H^{\rho\sigma} + \gamma_{21} F_{\mu}{}^{\nu} F_{\nu}{}^{\rho} H_{\rho}{}^{\sigma} H_{\sigma}{}^{\mu} + \gamma_{22} F_{\mu}{}^{\nu} H_{\nu}{}^{\rho} F_{\rho}{}^{\sigma} H_{\sigma}{}^{\mu} \\
& + \gamma_3 R H_{\mu\nu} H^{\mu\nu} + \gamma_4 R_{\mu\nu} H^{\mu\rho} H^{\nu}{}_{\rho} + \gamma_5 R_{\mu\nu\rho\sigma} H^{\mu\nu} H^{\rho\sigma} + \kappa_1 F_{\mu\nu} H^{\mu\nu} H_{\rho\sigma} H^{\rho\sigma} + \kappa_2 F_{\mu}{}^{\nu} H_{\nu}{}^{\rho} H_{\rho}{}^{\sigma} H_{\sigma}{}^{\mu} \\
& \left. + \zeta_1 H_{\mu\nu} H^{\mu\nu} H_{\rho\sigma} H^{\rho\sigma} + \zeta_2 H_{\mu}{}^{\nu} H_{\nu}{}^{\rho} H_{\rho}{}^{\sigma} H_{\sigma}{}^{\mu} \right). \tag{49}
\end{aligned}$$

Here, all the coefficients,  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $\kappa_i$ , and  $\zeta_i$ , should be small enough, since the higher derivative terms are corrections to the first line in the right-hand side of Eq. (49). We regard the orders of  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $\kappa_i$ , and  $\zeta_i$  as  $\mathcal{O}(\epsilon)$  in a small constant  $\epsilon$ . Moreover, in the action (49), we omit the terms that are quadratic in the background equations of motion such as

$$\begin{aligned}
& \left( R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R - 8\pi G T_{\mu\nu} \right)^2, R^2, \\
& \nabla_{\mu} F^{\mu\nu} \nabla_{\rho} F^{\rho}{}_{\nu}, \nabla_{\mu} F^{\mu\nu} \nabla_{\rho} H^{\rho}{}_{\nu}, \nabla_{\mu} H^{\mu\nu} \nabla_{\rho} H^{\rho}{}_{\nu}. \tag{50}
\end{aligned}$$

The reason is as follows. First, the effective coupling constants in front of these operators are  $\mathcal{O}(\epsilon)$ . Second, variation of those terms is proportional to the background equations, which gives one more factor of  $\mathcal{O}(\epsilon)$  when evaluated with the full solution. In total their contributions are  $\mathcal{O}(\epsilon^2)$ , which are higher orders ignored in our analysis. Note that the first combination in (50) together with the second one was used to replace  $R_{\mu\nu} R^{\mu\nu}$  term into some terms appearing in (49). Similarly, since the Gauss-Bonnet term,  $R^2 - 4R_{\mu\nu} R^{\mu\nu} + R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}$ , does not contribute to the equations of motion,  $R_{\mu\nu\rho\sigma} R^{\mu\nu\rho\sigma}$  term can be rewritten as well. See Appendix A 1 for more detail. Also we emphasize that the deformation of the action by field redefinitions is not done on the action (49). See Appendix B for more comments on the field redefinition.

Note that the higher derivative interactions generically appear in the low-energy effective action. For instance,  $\alpha_i$  terms, as well as  $\zeta_1$ ,  $\zeta_2$  and  $\gamma_3$ ,  $\gamma_4$ ,  $\gamma_5$ , are nothing but the Drummond-Hathrell effective action [19] which arise from the mediation of heavy massive particles with a single charge. Similarly, if there exist bicharged heavy massive particles,  $\beta_i$  terms and others are generated as the low-energy effective interactions.

The equations of motion can be derived as

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} - 8\pi G T_{\mu\nu} = 8\pi G \delta T_{\mu\nu}, \tag{51}$$

$$\nabla_{\nu} F^{\nu\mu} = \delta J_F^{\mu}, \tag{52}$$

$$\nabla_{\nu} H^{\nu\mu} = \delta J_H^{\mu}, \tag{53}$$

where  $\delta T_{\mu\nu}$ ,  $\delta J_F^{\mu}$ , and  $\delta J_H^{\mu}$  represent the contribution from the higher derivative correction terms and the detailed expressions are given in Appendix A 2.

We solve a static spherically symmetric solution perturbatively by regarding the parameters  $\alpha_i$ ,  $\beta_i$ ,  $\gamma_i$ ,  $\kappa_i$ ,  $\zeta_i$  are the first order in perturbations, say,  $\mathcal{O}(\epsilon)$ . Let us expand the dynamical variables as

$$f(r) = \bar{f}(r) + \delta f(r), \tag{54}$$

$$h(r) = \bar{h}(r) + \delta h(r), \tag{55}$$

$$\Phi(r) = \bar{\Phi}(r) + \delta\Phi(r), \tag{56}$$

$$\Psi(r) = \bar{\Psi}(r) + \delta\Psi(r), \tag{57}$$

where  $\bar{f}$ ,  $\bar{h}$ ,  $\bar{\Phi}$ , and  $\bar{\Psi}$  are  $\mathcal{O}(\epsilon^0)$  and  $\delta f$ ,  $\delta h$ ,  $\delta\Phi$ , and  $\delta\Psi$  are regarded as  $\mathcal{O}(\epsilon)$ . In the background, that is,  $\mathcal{O}(\epsilon^0)$ , the equations of motion reduce to the usual Einstein-Maxwell equations, and hence we obtain Reissner-Nordström solution. Then, in the  $\mathcal{O}(\epsilon)$  equations of motion, we can plug the  $\mathcal{O}(\epsilon^0)$  solutions in the right hand side of Eqs. (51)–(53), because they already involve  $\mathcal{O}(\epsilon)$  coefficients such as  $\alpha_i$ . By solving  $\mathcal{O}(\epsilon)$  equations of motion, we obtain the following perturbative solutions:

$$\begin{aligned}
 f(r) = & 1 - \frac{2GM}{r} + \frac{G(Q^2 + P^2)}{4\pi r^2} - \frac{G}{\pi r^4} \left( Q^2(-2\alpha_3 + \alpha_5) + QP(-2\beta_3 + \beta_5) + P^2(-2\gamma_3 + \gamma_5) \right) \\
 & + \frac{G^2 M}{\pi r^5} \left( Q^2(-6\alpha_3 - \alpha_4 + \alpha_5) + QP(-6\beta_3 - \beta_4 + \beta_5) + P^2(-6\gamma_3 - \gamma_4 + \gamma_5) \right) \\
 & - \frac{G}{80\pi^3 r^6} \left( Q^4 \tilde{\alpha}_1 + Q^3 P \tilde{\beta}_1 + Q^2 P^2 \tilde{\gamma}_1 + Q P^3 \tilde{\kappa}_1 + P^4 \tilde{\zeta}_1 \right) \\
 & + \frac{G^3 \bar{M}_{\text{ext}}(Q, P)^2}{5\pi r^6} \left( Q^2(20\alpha_3 + 4\alpha_4 - \alpha_5) + QP(20\beta_3 + 4\beta_4 - \beta_5) + P^2(20\gamma_3 + 4\gamma_4 - \gamma_5) \right) + \mathcal{O}(\epsilon^2), \quad (58)
 \end{aligned}$$

$$h(r) = 1 + \frac{G}{\pi r^4} \left( Q^2 \tilde{\alpha}_3 + QP \tilde{\beta}_3 + P^2 \tilde{\gamma}_3 \right) + \mathcal{O}(\epsilon^2), \quad (59)$$

$$\begin{aligned}
 \Phi(r) = & \frac{Q}{4\pi r} + \frac{GM(2Q\alpha_5 + P\beta_5)}{2\pi r^4} - \frac{1}{160\pi^3 r^5} \left( 4Q^3 \tilde{\alpha}_1 + 3Q^2 P \tilde{\beta}_1 + 2QP^2 \tilde{\gamma}_1 + P^3 \tilde{\kappa}_1 \right) \\
 & - \frac{1}{10\pi r^5} \left( G^2 \bar{M}_{\text{ext}}(Q, P)^2 \left( 2Q(\alpha_4 + 6\alpha_5) + P(\beta_4 + 6\beta_5) \right) - \frac{GQ}{4\pi} \left( Q^2 \tilde{\alpha}_3 + QP \tilde{\beta}_3 + P^2 \tilde{\gamma}_3 \right) \right) + \mathcal{O}(\epsilon^2), \quad (60)
 \end{aligned}$$

$$\begin{aligned}
 \Psi(r) = & \frac{P}{4\pi r} + \frac{GM(2P\gamma_5 + Q\beta_5)}{2\pi r^4} - \frac{1}{160\pi^3 r^5} \left( Q^3 \tilde{\beta}_1 + 2Q^2 P \tilde{\gamma}_1 + 3QP^2 \tilde{\kappa}_1 + 4P^3 \tilde{\zeta}_1 \right) \\
 & - \frac{1}{10\pi r^5} \left( G^2 \bar{M}_{\text{ext}}(Q, P)^2 \left( Q(\beta_4 + 6\beta_5) + 2P(\gamma_4 + 6\gamma_5) \right) - \frac{GP}{4\pi} \left( Q^2 \tilde{\alpha}_3 + QP \tilde{\beta}_3 + P^2 \tilde{\gamma}_3 \right) \right) + \mathcal{O}(\epsilon^2), \quad (61)
 \end{aligned}$$

where

$$\tilde{\alpha}_1 := 2\alpha_1 + \alpha_2, \quad (62)$$

$$\tilde{\beta}_1 := 2\beta_1 + \beta_2, \quad (63)$$

$$\tilde{\gamma}_1 := 2\gamma_{11} + \gamma_{21} + 2\gamma_{12} + \gamma_{22}, \quad (64)$$

$$\tilde{\kappa}_1 := 2\kappa_1 + \kappa_2, \quad (65)$$

$$\tilde{\zeta}_1 := 2\zeta_1 + \zeta_2, \quad (66)$$

and

$$\tilde{\alpha}_3 := 10\alpha_3 + 3\alpha_4 + 3\alpha_5, \quad (67)$$

$$\tilde{\beta}_3 := 10\beta_3 + 3\beta_4 + 3\beta_5, \quad (68)$$

$$\tilde{\gamma}_3 := 10\gamma_3 + 3\gamma_4 + 3\gamma_5, \quad (69)$$

and  $\bar{M}_{\text{ext}}(Q, P)$  is the extremal mass without the higher derivative corrections defined in Eq. (29). Note that the solutions correspond to the two gauge fields extension of the solution obtained in Ref. [20].

Now let us derive the extremal mass, i.e., the minimum black hole mass, for given charges  $Q, P$ . Recall that in the Einstein-Maxwell theory (19) without higher derivative corrections, the extremal mass for given charges  $Q, P$  was obtained by solving the two conditions,

$$\overline{f(r_{\text{H}})} = 0, \quad \overline{f'(r_{\text{H}})} = 0 \quad (70)$$

for  $r_{\text{H}}$  and  $M$ . Here the first condition means that  $r = r_{\text{H}}$  is a horizon, whereas the second condition is for the two horizons to degenerate. Since higher derivative corrections are supposed to be small, those properties of the metric functions should remain qualitatively the same and in particular the extremal mass should be determined by the conditions (70) as long as the derivative expansion works well. Indeed, one can explicitly confirm this expectation within the perturbative expansion in  $\epsilon$ .

Our task is then to solve Eq. (70) perturbatively. To do so, let us expand the horizon radius  $r_{\text{H}}$  and the extremal mass  $M_{\text{ext}}(Q, P)$  around those of the Einstein-Maxwell theory without higher derivative corrections,

$$r_{\text{H}} = \bar{r}_{\text{H}} + \delta r_{\text{H}}, \quad (71)$$

$$M_{\text{ext}}(Q, P) = \bar{M}_{\text{ext}}(Q, P) + \delta M_{\text{ext}}(Q, P), \quad (72)$$

where as we defined earlier we have

$$\bar{r}_{\text{H}} = G\bar{M}_{\text{ext}}(Q, P), \quad \bar{M}_{\text{ext}}(Q, P) = \sqrt{\frac{Q^2 + P^2}{4\pi G}}. \quad (73)$$

To solve the first condition of Eq. (70), it is convenient to expand  $f(r)$  as

$$\begin{aligned}
 f(\bar{r}_H + \delta r) = & -\frac{2\delta M_{\text{ext}}(Q, P)}{\bar{M}_{\text{ext}}(Q, P)} - \frac{G}{80\pi^3 \bar{r}_H^6} \left( Q^4 \alpha + Q^3 P \beta \right. \\
 & \left. + Q^2 P^2 \gamma + Q P^3 \kappa + P^4 \zeta \right) \\
 & + \mathcal{O}(\epsilon^2, \epsilon \delta r_H, \delta r_H^2), \tag{74}
 \end{aligned}$$

where the coefficients are defined by

$$\alpha := \tilde{\alpha}_1 + 4\pi G(\alpha_4 + \alpha_5), \tag{75}$$

$$\beta := \tilde{\beta}_1 + 4\pi G(\beta_4 + \beta_5), \tag{76}$$

$$\gamma := \tilde{\gamma}_1 + 4\pi G(\alpha_4 + \gamma_4 + \alpha_5 + \gamma_5), \tag{77}$$

$$\kappa := \tilde{\kappa}_1 + 4\pi G(\beta_4 + \beta_5), \tag{78}$$

$$\zeta := \tilde{\zeta}_1 + 4\pi G(\gamma_4 + \gamma_5). \tag{79}$$

Note that  $\mathcal{O}(\delta r_H)$  terms do not appear in the right-hand side because of the extremal condition defining  $\bar{r}_H$  terms. The correction to the extremal mass then reads,

$$\begin{aligned}
 \frac{\delta M_{\text{ext}}(Q, P)}{\bar{M}_{\text{ext}}(Q, P)} = & -\frac{2}{5} \frac{G}{(4\pi)^3 \bar{r}_H^6} \left( Q^4 \alpha + Q^3 P \beta + Q^2 P^2 \gamma \right. \\
 & \left. + Q P^3 \kappa + P^4 \zeta \right) + \mathcal{O}(\epsilon^2). \tag{80}
 \end{aligned}$$

$$\begin{aligned}
 E_C(Q, P; q, p) := & \frac{1}{4\pi G \bar{M}_{\text{ext}}(Q, P)} \left( qQ + pP \right. \\
 & + \frac{2G}{5(4\pi)^3 \bar{r}_H^6} \left( q \left( Q^3(Q^2 - 4P^2)\alpha + Q^2 P(2Q^2 - 3P^2)\beta + Q P^2(3Q^2 - 2P^2)\gamma + P^3(4Q^2 - P^2)\kappa + 5Q P^4 \zeta \right) \right. \\
 & \left. \left. + p \left( 5Q^4 P \alpha + Q^3(-Q^2 + 4P^2)\beta + Q^2 P(-2Q^2 + 3P^2)\gamma + Q P^2(-3Q^2 + 2P^2)\kappa + P^3(-4Q^2 + P^2)\zeta \right) \right) \right). \tag{84}
 \end{aligned}$$

Note that the expansion in Eq. (83) with respect to  $q$  and  $p$  should start from the linear order, since the right-hand side shows the deviation from the extremal condition. Therefore, terms in the order of  $\epsilon^2 q^0$  or  $\epsilon^2 p^0$  never appear in Eq. (83). Moreover, the charges of test particle  $q$  and  $p$  are small such that  $q, p \ll \epsilon$  is satisfied. This justifies the approximation in the right hand side of Eq. (83).

Next, let us derive the falling condition. The general expression has already been shown in Eq. (18). Substituting the electrostatic potentials shown in Eqs. (60) and (61) into (18), we finally have

Substituting this into the second condition  $f'(r_H) = 0$  of Eq. (70) gives the shift  $\delta r_H$  of the horizon position of the extremal black hole as

$$\begin{aligned}
 \frac{\delta r_H}{\bar{r}_H} = & -\frac{2G}{(4\pi)^3 \bar{r}_H^6} \left( Q^4 \tilde{\alpha}_1 + Q^3 P \tilde{\beta}_1 + Q^2 P^2 \tilde{\gamma}_1 \right. \\
 & \left. + Q P^3 \tilde{\kappa}_1 + P^4 \tilde{\zeta}_1 \right) + \frac{G}{\pi \bar{r}_H^4} \left( Q^2 \alpha_3 + Q P \beta_3 + P^2 \gamma_3 \right) \\
 & + \mathcal{O}(\epsilon^2). \tag{81}
 \end{aligned}$$

Here  $\delta r_H$  is  $\mathcal{O}(\epsilon)$  as expected. Note that the areal radius of the event horizon of the extremal black hole is not invariant under the field redefinition. The reason is explained in Appendix B 1.

## B. Falling conditions

Let us compare the super-extremal conditions with the falling conditions. Our system is composed of an extremal black hole  $M = M_{\text{ext}}(Q, P)$  and a test charged particle with the energy  $E$  and charges  $(q, p)$ . The total mass and charges are

$$(\hat{M}, \hat{Q}, \hat{P}) = (M_{\text{ext}}(Q, P) + E, Q + q, P + p). \tag{82}$$

Let us expand the super-extremal condition for  $(\hat{M}, \hat{Q}, \hat{P})$  up to linear order in  $q$  and  $p$ . The super-extremal condition  $\hat{M} < M_{\text{ext}}(\hat{Q}, \hat{P})$  can be expressed as

$$E < E_C(Q, P; q, p) + \mathcal{O}(\epsilon^2 q, \epsilon^2 p, q^2, qp, p^2), \tag{83}$$

with

$$E \geq q\Phi(r_H) + p\Psi(r_H) = E_C(Q, P; q, p), \tag{85}$$

where  $r_H$  is given by Eq. (81). Thus, the super-extremal condition (83) and the falling condition (85) cannot be satisfied at the same time at the leading order in  $q$  and  $p$ . This is the main conclusion of this paper.

To see the implications more explicitly, let us set  $P = 0$  in the above equations. Then, the super-extremal condition  $\hat{M} < M_{\text{ext}}(\hat{Q}, \hat{P})$  can be expressed as



$$E < \frac{\text{sgn}(Q)}{\sqrt{4\pi G}} \left( q + \frac{2(\alpha q - \beta p)}{5G^2 Q^2} \right) + \mathcal{O}(\epsilon^2 q, \epsilon^2 p, p^2, pq, q^2), \quad (86)$$

while the falling condition can be expressed as

$$E \geq \frac{\text{sgn}(Q)}{\sqrt{4\pi G}} \left( q + \frac{2(\alpha q - \beta p)}{5G^2 Q^2} \right) + \mathcal{O}(\epsilon^2 q, \epsilon^2 p). \quad (87)$$

We would like to comment on specific choices of the particle charges. First, if we set  $p = 0$ , our result reduces to the single gauge field case investigated in Refs. [28,31,33].

Next, let us consider the case with  $q = 0$ . Remember the discussion at the end of Sec. III that, in the minimally coupled theory, this parameter choice corresponds to the case where the  $\mathcal{O}(p)$  contribution vanishes and the test particle approximation is violated. Contrary to the minimally coupled case, even though the condition  $q = 0$  indicates no Coulomb force, there is additional repulsive forces that are induced through the higher derivative couplings between two gauge fields. More explicitly, from the expression (61), we can see that the electrostatic potential  $\Psi$  is induced even when  $P = 0$ , through the coupling  $\beta_i$ . This is because even if objects (namely, particles and/or black holes) have different charges interact with each other through the mixing derivative couplings. The condition for the absence of force at the event horizon can be written as

$$q = \frac{2\beta}{5G^2 Q^2} p + \mathcal{O}(\epsilon^2 q, \epsilon^2 p). \quad (88)$$

For the particle with this charges, both the Coulomb force and the force induced by the higher derivative corrections work but exactly cancel at the event horizon. Then,  $\mathcal{O}(q^2, qp, p^2)$  corrections are required to be seriously taken into account, because the effects beyond the point particle approximation might be significant. Thus, the situation is similar to the case of  $q = 0$  in the minimally coupled theory, although the particle feels Coulomb force at the infinity.

## V. SUMMARY AND DISCUSSION

In this paper, we investigate gedanken experiments to destroy the extremal black hole by dropping a test charged particle. In particular, we focus on a theory with two  $U(1)$  gauge fields with higher derivative corrections. The black hole we consider is a spherically symmetric, static extremal one with charges  $(Q, P)$  with respect to two gauge fields. In the case with minimally coupled gauge fields, the black hole solution is described by Reissner-Nordström solution. The super-extremal condition for the total energy and total charges of the system is given by Eq. (42) and the falling condition, which shows that the particle can arrive at the

event horizon, is given by Eq. (43). Since these two conditions does not hold at the same time, extremal black holes cannot be destroyed by a test particle. We extend this analysis in the presence of the higher derivative corrections. Our action is given by Eq. (49). The black hole solution is derived perturbatively in Eqs. (58)–(61). The super-extremal condition and the falling condition are obtained in Eqs. (83) and (85) respectively. As is the case of minimally coupled gauge fields, these conditions are not satisfied at the same time. Hence, we confirm that the extremal black hole with two gauge fields cannot be destroyed by a test particle even when the higher derivative corrections are included.

Due to the validity of the test particle approximation, our result holds only up to the linear order in  $q$  and  $p$ . For a specific choice of the charges, Eq. (44) for the minimally coupled case and Eq. (88) for the case with the higher derivative corrections with  $P = 0$ , linear order contributions disappear both in the super-extremal condition and the falling condition. In such a case, the  $\mathcal{O}(p^2, pq, q^2)$  effects become dominant and the analysis beyond the point particle approximation similarly to the one in Ref. [25] is again required, which is left for future work.

Although the second order analysis in Ref. [25] is applicable to subextremal black holes only and it does not directly apply to extremal black holes, it is instructive to see what kind of effects may be expected from a naive extremal limit of the result there. Here let us focus on the minimally coupled case for simplicity. Assuming that the nonelectromagnetic stress-energy tensor satisfies the null energy condition, Ref. [25] derived a lower bound on the second-order correction to the total energy of the system, which captures the self-force and finite size effects for example. By extending the result there to the minimally coupled two gauge fields case naively, for a charged particle with  $q = 0$ ,  $p \neq 0$  thrown into a near-extremal black hole with  $Q \neq 0$ ,  $P = 0$ , the lower bound reads

$$\hat{M} \geq \bar{M}_{\text{ext}}(Q, 0) + \frac{1}{2} \frac{|Q|}{\sqrt{4\pi G}} \frac{p^2}{Q^2} + \mathcal{O}(\epsilon), \quad (89)$$

where  $\epsilon$  is a small parameter quantifying deviation from the extremality of the original near-extremal black hole. Here we assumed that the first order increase of the total energy is negligible. Note that the second term on the right-hand side is the lower bound on the second-order correction to the total energy. On the other hand, the super-extremal condition on the total system up to  $\mathcal{O}(p^2)$  is

$$\hat{M} - \bar{M}_{\text{ext}}(Q, p) = \hat{M} - \bar{M}_{\text{ext}}(Q, 0) - \frac{1}{2} \frac{|Q|}{\sqrt{4\pi G}} \frac{p^2}{Q^2} < 0. \quad (90)$$

Interestingly, the second-order corrections in Eqs. (89) and (90) precisely match with each other, so that the two bounds cannot be satisfied simultaneously. In other words,

the window (48) for destroying the black hole may be closed by carefully taking into account the second-order correction to the total energy. Again, we emphasize that the bound (89) does not apply directly to the extremal black hole ( $\varepsilon = 0$ ) and so it is not clear yet if the cosmic censorship conjecture is satisfied for the extremal case, but our observation motivates us to generalize the analysis in Ref. [25] to the extremal black hole. We leave this interesting problem to the future work.

### ACKNOWLEDGMENTS

D. Y. would like to thank Hirotaka Yoshino for useful discussion in the 24th workshop on spacetime singularities in Japan. K. I. and D. Y. are supported by Grant-Aid for Scientific Research from Ministry of Education, Science, Sports and Culture of Japan (No. JP21H05189). K. I. and T. N. are supported by JSPS KAKENHI Grant No. JP20H01902. All of the authors are supported by JSPS Bilateral Joint Research Projects (JSPS-Department of Science and Technology Collaboration) (JPJSBP120227705). K. I. is also supported by JSPS(JP21H05182). T. N. is also supported by JSPS KAKENHI Grant No. JP22H01220 and MEXT KAKENHI Grant No. JP21H05184 and No. JP23H04007. D. Y. is supported by JSPS KAKENHI Grant No. JP20K14469.

### APPENDIX A: DETAILED EXPRESSION FOR THE HIGHER DERIVATIVE CORRECTIONS

#### 1. How to rewrite the curvature square terms

We investigate a theory with the higher derivative interactions (49), where the curvature square terms do not show up in the action. This is justified, as we have explained below Eq. (49), by the fact that all curvature square terms are rewritten by the other terms. In this appendix, we explicitly show how to rewrite the curvature square terms with the other higher derivative terms.

The quadratic curvature terms,

$$R^2, \quad R_{\mu\nu}R^{\mu\nu}, \quad R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}, \quad (\text{A1})$$

are expressed by linear combinations of

$$R^2, \quad G_{\mu\nu}G^{\mu\nu}, \quad GB := R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}. \quad (\text{A2})$$

We investigate quadratic curvature terms (A2), instead of (A1). Suppose that an action includes these quadratic curvature terms (A2)

$$\lambda_1 R^2 + \lambda_2 G_{\mu\nu}G^{\mu\nu} + \lambda_3 GB. \quad (\text{A3})$$

This term can be expressed as

$$\begin{aligned} & \lambda_1 R^2 + \lambda_2 G_{\mu\nu}G^{\mu\nu} + \lambda_3 GB \\ &= \lambda_1 R^2 + \lambda_2 (G_{\mu\nu} - 8\pi GT_{\mu\nu})(G^{\mu\nu} - 8\pi GT^{\mu\nu}) \\ & \quad + \lambda_3 GB + 16\pi G\lambda_2 G_{\mu\nu}T^{\mu\nu} - (8\pi G)^2 \lambda_2 T_{\mu\nu}T^{\mu\nu}, \end{aligned} \quad (\text{A4})$$

where  $T_{\mu\nu}$  is the energy momentum tensor without the higher derivative contributions, defined in Eq. (21). Since the Gauss-Bonnet combination  $GB$  is topological invariant, it does not affect the dynamics and thus we can just ignore it. Meanwhile, because  $R$  and  $G_{\mu\nu} - 8\pi GT_{\mu\nu}$  vanish at the order in  $\varepsilon^0$ , i.e., without the higher derivative contributions, they are  $\mathcal{O}(\varepsilon^1)$  quantities. The variations of squares of them gives linear terms of them, that is, they are  $\mathcal{O}(\varepsilon^1)$  quantities. Suppose the coefficients of quadratic curvature terms  $\lambda_i$  is  $\mathcal{O}(\varepsilon)$ , and then the first two terms in the right-hand side of Eq. (A4) gives  $\mathcal{O}(\varepsilon^2)$  contributions in equations of motion, which we ignore in our analysis. As a result, the quadratic curvature (A3) is written in the last two terms in the right-hand side of Eq. (A4). The explicit form of each term is

$$\begin{aligned} & 16\pi G\lambda_2 G_{\mu\nu}T^{\mu\nu} \\ &= 4\pi G\lambda_2 (-RF_{\mu\nu}F^{\mu\nu} + 4R_{\mu\nu}F^{\mu\rho}F^{\nu\rho} \\ & \quad - RH_{\mu\nu}H^{\mu\nu} + 4R_{\mu\nu}H^{\mu\rho}H^{\nu\rho}), \end{aligned} \quad (\text{A5})$$

$$\begin{aligned} & - (8\pi G)^2 \lambda_2 T_{\mu\nu}T^{\mu\nu} \\ &= 16\pi^2 G^2 \lambda_2 (F_{\mu\nu}F^{\mu\nu}F_{\rho\sigma}F^{\rho\sigma} - 4F_{\mu}{}^{\nu}F_{\nu}{}^{\rho}F_{\rho}{}^{\sigma}F_{\sigma}{}^{\mu} \\ & \quad + 2F_{\mu\nu}F^{\mu\nu}H_{\rho\sigma}H^{\rho\sigma} - 8F_{\mu}{}^{\nu}F_{\nu}{}^{\rho}H_{\rho}{}^{\sigma}H_{\sigma}{}^{\mu} \\ & \quad + H_{\mu\nu}H^{\mu\nu}H_{\rho\sigma}H^{\rho\sigma} - 4H_{\mu}{}^{\nu}H_{\nu}{}^{\rho}H_{\rho}{}^{\sigma}H_{\sigma}{}^{\mu}). \end{aligned} \quad (\text{A6})$$

More concretely, if one starts with the action with (A3) terms, such effect can be included by replacing the parameters in Eq. (49) as follows:

$$\alpha_1 \rightarrow \alpha_1 + 16\pi^2 G^2 \lambda_2, \quad (\text{A7})$$

$$\alpha_2 \rightarrow \alpha_2 - 64\pi^2 G^2 \lambda_2, \quad (\text{A8})$$

$$\alpha_3 \rightarrow \alpha_3 - 4\pi G\lambda_2, \quad (\text{A9})$$

$$\alpha_4 \rightarrow \alpha_4 + 16\pi G\lambda_2, \quad (\text{A10})$$

$$\gamma_{11} \rightarrow \gamma_{11} + 32\pi^2 G^2 \lambda_2, \quad (\text{A11})$$

$$\gamma_{21} \rightarrow \gamma_{21} - 128\pi^2 G^2 \lambda_2, \quad (\text{A12})$$

$$\gamma_3 \rightarrow \gamma_3 - 4\pi G\lambda_2, \quad (\text{A13})$$

$$\gamma_4 \rightarrow \gamma_4 + 16\pi G\lambda_2, \quad (\text{A14})$$

$$\zeta_1 \rightarrow \zeta_1 + 16\pi^2 G^2 \lambda_2, \quad (\text{A15})$$

$$\zeta_2 \rightarrow \zeta_2 - 64\pi^2 G^2 \lambda_2. \quad (\text{A16})$$

## 2. Source terms in the equations of motion

In this section, we represent the detailed expressions for  $\delta T_{\mu\nu}$ ,  $\delta J_F^\mu$ , and  $\delta J_H^\mu$  appearing in the right hand side of Eqs. (51)–(53). The expressions are as follows:

$$\begin{aligned} \delta T_{\mu\nu} = & \alpha_1 \mathcal{T}_{1,\mu\nu}[F, F, F, F] + \alpha_2 \mathcal{T}_{2,\mu\nu}[F, F, F, F] + \beta_1 \mathcal{T}_{1,\mu\nu}[F, F, F, H] + \beta_2 \mathcal{T}_{2,\mu\nu}[F, F, F, H] \\ & + \gamma_{11} \mathcal{T}_{1,\mu\nu}[F, F, H, H] + \gamma_{12} \mathcal{T}_{1,\mu\nu}[F, H, F, H] + \gamma_{21} \mathcal{T}_{2,\mu\nu}[F, F, H, H] + \gamma_{22} \mathcal{T}_{2,\mu\nu}[F, H, F, H] \\ & + \kappa_1 \mathcal{T}_{1,\mu\nu}[F, H, H, H] + \kappa_2 \mathcal{T}_{2,\mu\nu}[F, H, H, H] + \zeta_1 \mathcal{T}_{1,\mu\nu}[H, H, H, H] + \zeta_2 \mathcal{T}_{2,\mu\nu}[H, H, H, H] \\ & + \alpha_3 \mathcal{T}_{3,\mu\nu}[F, F] + \alpha_4 \mathcal{T}_{4,\mu\nu}[F, F] + \alpha_5 \mathcal{T}_{5,\mu\nu}[F, F] + \beta_3 \mathcal{T}_{3,\mu\nu}[F, H] + \beta_4 \mathcal{T}_{4,\mu\nu}[F, H] + \beta_5 \mathcal{T}_{5,\mu\nu}[F, H] \\ & + \gamma_3 \mathcal{T}_{3,\mu\nu}[H, H] + \gamma_4 \mathcal{T}_{4,\mu\nu}[H, H] + \gamma_5 \mathcal{T}_{5,\mu\nu}[H, H], \end{aligned} \quad (\text{A17})$$

$$\begin{aligned} \delta J_F^\mu = & \nabla_\nu \left( 8\alpha_1 \mathcal{J}_1^{\mu\nu}[F, F, F] + 8\alpha_2 \mathcal{J}_2^{\mu\nu}[F, F, F] + \beta_1 (2\mathcal{J}_1^{\mu\nu}[F, F, H] + 4\mathcal{J}_1^{\mu\nu}[F, H, F]) \right. \\ & + \beta_2 (4\mathcal{J}_2^{\mu\nu}[F, F, H] + 2\mathcal{J}_2^{\mu\nu}[F, H, F]) + 4\gamma_{11} \mathcal{J}_1^{\mu\nu}[H, H, F] + 4\gamma_{12} \mathcal{J}_1^{\mu\nu}[F, H, H] \\ & + 4\gamma_{21} \mathcal{J}_2^{\mu\nu}[F, H, H] + 4\gamma_{22} \mathcal{J}_2^{\mu\nu}[H, F, H] + 2\kappa_1 \mathcal{J}_1^{\mu\nu}[H, H, H] + 2\kappa_2 \mathcal{J}_2^{\mu\nu}[H, H, H] \\ & \left. + 4\alpha_3 R F^{\nu\mu} + 4\alpha_4 R^{[\nu} F^{\rho\mu]} + 4\alpha_5 R^{\nu\mu\rho\sigma} F_{\rho\sigma} + 2\beta_3 R H^{\nu\mu} + 2\beta_4 R^{[\nu} H^{\rho\nu]} + 2\beta_5 R^{\nu\mu\rho\sigma} H_{\rho\sigma} \right), \end{aligned} \quad (\text{A18})$$

$$\begin{aligned} \delta J_H^\mu = & \nabla_\nu \left( 8\zeta_1 \mathcal{J}_1^{\mu\nu}[H, H, H] + 8\zeta_2 \mathcal{J}_2^{\mu\nu}[H, H, H] + \kappa_1 (2\mathcal{J}_1^{\mu\nu}[H, H, F] + 4\mathcal{J}_1^{\mu\nu}[H, F, H]) \right. \\ & + \kappa_2 (4\mathcal{J}_2^{\mu\nu}[H, H, F] + 2\mathcal{J}_2^{\mu\nu}[H, F, H]) + 4\gamma_{11} \mathcal{J}_1^{\mu\nu}[F, F, H] + 4\gamma_{12} \mathcal{J}_1^{\mu\nu}[H, F, F] \\ & + 4\gamma_{21} \mathcal{J}_2^{\mu\nu}[H, F, F] + 4\gamma_{22} \mathcal{J}_2^{\mu\nu}[F, H, F] + 2\beta_1 \mathcal{J}_1^{\mu\nu}[F, F, F] + 2\beta_2 \mathcal{J}_2^{\mu\nu}[F, F, F] \\ & \left. + 4\gamma_3 R H^{\nu\mu} + 4\gamma_4 R^{[\nu} G^{\rho\mu]} + 4\gamma_5 R^{\nu\mu\rho\sigma} G_{\rho\sigma} + 2\beta_3 R F^{\nu\mu} + 2\beta_4 R^{[\nu} F^{\rho\nu]} + 2\beta_5 R^{\nu\mu\rho\sigma} F_{\rho\sigma} \right), \end{aligned} \quad (\text{A19})$$

where  $\mathcal{T}_{i,\mu\nu}[A, B, C, D]$ ,  $\mathcal{T}_{i,\mu\nu}[A, B]$  and  $\mathcal{J}_i^{\mu\nu}[A, B, C]$  are functionals of antisymmetric tensors  $A_{\mu\nu}$ ,  $B_{\mu\nu}$ ,  $C_{\mu\nu}$ ,  $D_{\mu\nu}$  defined as

$$\begin{aligned} \mathcal{T}_{1,\mu\nu}[A, B, C, D] = & g_{\mu\nu} A_{\alpha\beta} B^{\beta\alpha} C_{\gamma\lambda} D^{\lambda\gamma} - 4A_{(\mu|\alpha} B^{\alpha}_{|\nu)} C_{\beta\gamma} D^{\gamma\beta} \\ & - 4C_{(\mu|\alpha} D^{\alpha}_{|\nu)} A_{\beta\gamma} B^{\gamma\beta}, \end{aligned} \quad (\text{A20})$$

$$\begin{aligned} \mathcal{T}_{2,\mu\nu}[A, B, C, D] = & g_{\mu\nu} A_{\alpha\beta} B^{\beta\gamma} C_{\gamma\lambda} D^{\lambda\alpha} - 2A_{(\mu|\alpha} B^{\alpha\beta} C_{\beta\gamma} D^{\gamma}_{|\nu)} \\ & - 2B_{(\mu|\alpha} C^{\alpha\beta} D_{\beta\gamma} A^{\gamma}_{|\nu)} - 2C_{(\mu|\alpha} D^{\alpha\beta} A_{\beta\gamma} B^{\gamma}_{|\nu)} \\ & - 2D_{(\mu|\alpha} A^{\alpha\beta} B_{\beta\gamma} C^{\gamma}_{|\nu)}, \end{aligned} \quad (\text{A21})$$

$$\begin{aligned} \mathcal{T}_{3,\mu\nu}[A, B] = & -R A_{\alpha\beta} B^{\beta\alpha} g_{\mu\nu} + 2A_{\alpha\beta} B^{\beta\alpha} R_{\mu\nu} + 4R A_{(\mu|\beta} B^{\beta}_{|\nu)} \\ & - 2\nabla_\mu \nabla_\nu (A_{\alpha\beta} B^{\beta\alpha}) + 2g_{\mu\nu} \square (A_{\alpha\beta} B^{\beta\alpha}), \end{aligned} \quad (\text{A22})$$

$$\begin{aligned} \mathcal{T}_{4,\mu\nu}[A, B] = & -R_{\rho\sigma} A^{\rho\alpha} B_{\alpha}{}^\sigma g_{\mu\nu} + 2R_{(\mu}{}^\rho A_{\nu)\alpha} B^{\alpha\rho} \\ & + 2R_{(\mu}{}^\rho B_{\nu)\alpha} A^{\alpha\rho} - 2R_{\rho\sigma} A^{\rho}{}_{(\mu} B^{\sigma}_{\nu)} \\ & + \square (A_{(\mu|\alpha} B^{\alpha}_{|\nu)}) + \nabla_\rho \nabla_\sigma (A^{\rho\alpha} B_{\alpha}{}^\sigma) g_{\mu\nu} \\ & - \nabla_\rho \nabla_{(\mu} (A_{\nu)\alpha} B^{\alpha\rho}) - \nabla_\rho \nabla_{(\mu} (B_{\nu)\alpha} A^{\alpha\rho}), \end{aligned} \quad (\text{A23})$$

$$\begin{aligned} \mathcal{T}_{5,\mu\nu}[A, B] = & R_{\rho\sigma\lambda\tau} A^{\rho\sigma} B^{\lambda\tau} g_{\mu\nu} \\ & - 3R_{(\mu}{}^{\rho\sigma\tau} A_{\nu)\rho} B_{\sigma\tau} - 3R_{(\mu}{}^{\rho\sigma\tau} B_{\nu)\rho} A_{\sigma\tau} \\ & - 2\nabla_\rho \nabla_\sigma (A_{(\mu}{}^\rho B_{\nu)}{}^\sigma + B_{(\mu}{}^\rho A_{\nu)}{}^\sigma), \end{aligned} \quad (\text{A24})$$

$$\mathcal{J}_1^{\mu\nu}[A, B, C] = A_{\alpha\beta} B^{\beta\alpha} C^{\mu\nu}, \quad (\text{A25})$$

$$\mathcal{J}_2^{\mu\nu}[A, B, C] = A^{[\mu}{}_{\alpha} B^{\alpha\beta} C_{\beta}{}^{\nu]}. \quad (\text{A26})$$

## APPENDIX B: FIELD REDEFINITION

Field redefinition changes the description of the theory, for instance, changes the coupling constants, but the physical phenomenon remains the same. Although the theory we consider can be rewritten in a different description by a field redefinition, in this appendix we explicitly confirm that the obtained results are the same. This is obvious, but it becomes a cross-check of the validity of our analysis, and may bring a new perspective on the analysis of theories with the higher derivative terms.

Let us investigate the perturbative redefinition of the metric field  $g_{\mu\nu}$  to  $\hat{g}_{\mu\nu}$  through,

$$g_{\mu\nu} = \hat{g}_{\mu\nu} + \delta g_{\mu\nu}, \quad (\text{B1})$$

with  $\delta g_{\mu\nu}$  defined by

$$\begin{aligned} \delta g_{\mu\nu} = & \epsilon_1 R_{\mu\nu} + \epsilon_2 g_{\mu\nu} R + 8\pi G \epsilon_3 F_{\mu\rho} F_{\nu}{}^{\rho} + 8\pi G \epsilon_4 g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} \\ & + 8\pi G \epsilon_5 F_{\mu\rho} P_{\nu}{}^{\rho} + 8\pi G \epsilon_6 g_{\mu\nu} F_{\rho\sigma} P^{\rho\sigma} + 8\pi G \epsilon_7 P_{\mu\rho} P_{\nu}{}^{\rho} \\ & + 8\pi G \epsilon_8 g_{\mu\nu} P_{\rho\sigma} P^{\rho\sigma}, \end{aligned} \quad (\text{B2})$$

where the coefficients are regarded as  $\mathcal{O}(\epsilon)$ . One can replace the metric  $g_{\mu\nu}$  in Eq. (B2) with  $\hat{g}_{\mu\nu}$  because the difference is higher order in  $\epsilon$ . Let us investigate the

influence of the field redefinition (B1) in the black hole solutions and the falling conditions.

## 1. Black hole solutions

Through this transformation, the action (49), expressing the dependence of the metric and the parameters as  $S[g_{\mu\nu}, \alpha_i]$ , is modified to

$$\begin{aligned} S[g_{\mu\nu}, \alpha_i] &= S[\hat{g}_{\mu\nu}, \alpha_i] - \int d^4x \frac{\sqrt{-g}}{16\pi G} \left( R^{\mu\nu} - \frac{1}{2} g^{\mu\nu} R - 8\pi G T^{\mu\nu} \right) \delta g_{\mu\nu} + \mathcal{O}(\epsilon^2) \\ &= S[\hat{g}_{\mu\nu}, \hat{\alpha}_i] + \int d^4x \frac{\sqrt{-g}}{16\pi G} \left[ -\epsilon_1 \left( R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} - 8\pi G T_{\mu\nu} \right)^2 + \left( \frac{\epsilon_1}{2} + \epsilon_2 \right) R^2 \right] + \mathcal{O}(\epsilon^2). \end{aligned} \quad (\text{B3})$$

Here the parameters  $\hat{\alpha}_i$  are defined as follows:

$$\hat{\alpha}_1 = \alpha_1 - \pi G(\epsilon_1 + \epsilon_3), \quad (\text{B4})$$

$$\hat{\alpha}_2 = \alpha_2 + 4\pi G(\epsilon_1 + \epsilon_3), \quad (\text{B5})$$

$$\hat{\alpha}_3 = \alpha_3 + \frac{1}{8}(\epsilon_1 + 2\epsilon_3 + 4\epsilon_4), \quad (\text{B6})$$

$$\hat{\alpha}_4 = \alpha_4 - \frac{1}{2}(\epsilon_1 + \epsilon_3), \quad (\text{B7})$$

$$\hat{\alpha}_5 = \alpha_5, \quad (\text{B8})$$

$$\hat{\beta}_1 = \beta_1 - \pi G \epsilon_5, \quad (\text{B9})$$

$$\hat{\beta}_2 = \beta_2 + 4\pi G \epsilon_5, \quad (\text{B10})$$

$$\hat{\beta}_3 = \beta_3 + \frac{1}{4}(\epsilon_5 + 2\epsilon_6), \quad (\text{B11})$$

$$\hat{\beta}_4 = \beta_4 - \frac{1}{2}\epsilon_5, \quad (\text{B12})$$

$$\hat{\beta}_5 = \beta_5, \quad (\text{B13})$$

$$\hat{\gamma}_{11} = \gamma_{11} - \pi G(2\epsilon_1 + \epsilon_3 + \epsilon_7), \quad (\text{B14})$$

$$\hat{\gamma}_{12} = \gamma_{12}, \quad (\text{B15})$$

$$\hat{\gamma}_{21} = \gamma_{21} + 4\pi G(2\epsilon_1 + \epsilon_3 + \epsilon_7), \quad (\text{B16})$$

$$\hat{\gamma}_{22} = \gamma_{22}, \quad (\text{B17})$$

$$\hat{\gamma}_3 = \gamma_3 + \frac{1}{8}(\epsilon_1 + 2\epsilon_7 + 4\epsilon_8), \quad (\text{B18})$$

$$\hat{\gamma}_4 = \gamma_4 - \frac{1}{2}(\epsilon_1 + \epsilon_7), \quad (\text{B19})$$

$$\hat{\gamma}_5 = \gamma_5, \quad (\text{B20})$$

$$\hat{\kappa}_1 = \kappa_1 - \pi G \epsilon_5, \quad (\text{B21})$$

$$\hat{\kappa}_2 = \kappa_2 + 4\pi G \epsilon_5, \quad (\text{B22})$$

$$\hat{\zeta}_1 = \zeta_1 - \pi G(\epsilon_1 + \epsilon_7), \quad (\text{B23})$$

$$\hat{\zeta}_2 = \zeta_2 + 4\pi G(\epsilon_1 + \epsilon_7). \quad (\text{B24})$$

The second term in Eq. (B3) is irrelevant for our analysis because it consists of squares of the background equations of motion. Thus, the relevant part of the action for  $\hat{g}_{\mu\nu}$  can be obtained by replacing the parameters in the original action (49) with the hatted ones defined above.

The transformations for the coupling constants are described by the six independent combinations of the parameters,

$$\epsilon_1 + \epsilon_3, \epsilon_3 + 4\epsilon_4, \epsilon_5, \epsilon_6, \epsilon_1 + \epsilon_7, \epsilon_7 + 4\epsilon_8. \quad (\text{B25})$$

One can see that the following combinations are invariant under the field redefinition (B1),

$$\alpha_1 - 2\pi G \alpha_4, \quad (\text{B26})$$

$$\alpha_2 + 8\pi G \alpha_4, \quad (\text{B27})$$

$$\beta_1 - 2\pi G \beta_4, \quad (\text{B28})$$

$$\beta_2 + 8\pi G \beta_4, \quad (\text{B29})$$

$$\gamma_{11} - 2\pi G(\alpha_4 + \gamma_4), \quad (\text{B30})$$

$$\gamma_{21} + 8\pi G(\alpha_4 + \gamma_4), \quad (\text{B31})$$

$$\kappa_1 - 2\pi G \beta_4, \quad (\text{B32})$$

$$\kappa_2 + 8\pi G\beta_4, \quad (\text{B33})$$

$$\zeta_1 - 2\pi G\gamma_4, \quad (\text{B34})$$

$$\zeta_2 + 8\pi G\gamma_4, \quad (\text{B35})$$

as well as  $\alpha_5, \beta_5, \gamma_{12}, \gamma_{22}$ , and  $\gamma_5$ . The parameters  $\alpha, \beta, \gamma, \kappa, \zeta$ , appearing in the extremal condition (80), are invariant thus under the field redefinition.

Next let us investigate the position of the horizon. The horizon is characterized by  $f(r) = 0$ .<sup>4</sup> Through the field redefinition,  $\hat{f} := -\hat{g}_{tt}$  can be calculated as

$$\hat{f} = f - \delta f \quad (\text{B36})$$

with

$$\begin{aligned} \delta f &= -\delta g_{tt} \\ &= -f(r) \frac{G}{4\pi r^4} \left( Q^2(\epsilon_1 + 2\epsilon_3 + 4\epsilon_4) + 2QP(\epsilon_5 + 2\epsilon_6) \right. \\ &\quad \left. + P^2(\epsilon_1 + 2\epsilon_7 + 4\epsilon_8) \right). \end{aligned} \quad (\text{B37})$$

Although  $\hat{f}(r)$  does not coincide with  $f(r)$  itself, a solution of  $f(r) = 0$  is the solution of  $\hat{f}(r) = 0$ , that is, the position of the horizon is invariant.

We note that it does not imply the invariance of the areal radius of the horizon. As we can see from Eq. (81),  $r_H$  is not invariant under the field redefinition. The reason is as follows. After the field redefinition, the function forms of metric change. Due to the spherical symmetry, metric after the field redefinition is written as

$$\hat{g}_{\mu\nu} dx^\mu dx^\nu = -\hat{f}(r) dt^2 + \frac{\hat{h}(r)}{\hat{f}(r)} dr^2 + \hat{r}(r)^2 d\Omega^2. \quad (\text{B38})$$

The invariance of the horizon position means that, in this coordinate with  $r$ , the locus of horizon is the same. However,  $r$  does not coincide with the areal radius  $\hat{r}(r)$  in metric (B38). The areal radius is  $\hat{r}(r)$ , which is calculated as

$$\begin{aligned} \hat{r}^2 &= r^2 - \epsilon_1 \frac{G(Q^2 + P^2)}{4\pi r^2} + \epsilon_4 \frac{GQ^2}{\pi r^2} \\ &\quad + \epsilon_6 \frac{GQP}{\pi r^2} + \epsilon_8 \frac{GP^2}{\pi r^2}. \end{aligned} \quad (\text{B39})$$

<sup>4</sup>In theories with noncanonical kinetic terms, superluminal modes can exist. Thus, an event horizon may not be a null surface. However, in static or stationary case, any horizon is expected to be a Killing horizon [35–37].

This gives us

$$\begin{aligned} \hat{r} &= r - \epsilon_1 \frac{G(Q^2 + P^2)}{8\pi r^3} + \epsilon_4 \frac{GQ^2}{2\pi r^3} + \epsilon_6 \frac{GQP}{2\pi r^3} \\ &\quad + \epsilon_8 \frac{GP^2}{2\pi r^3} + \mathcal{O}(\epsilon^2). \end{aligned} \quad (\text{B40})$$

Thus,  $\hat{r}_H$ , the areal radius of the horizon for  $\hat{g}_{\mu\nu}$ , can be obtained as

$$\begin{aligned} \hat{r}_H &= r_H - \epsilon_1 \frac{1}{2\bar{r}_H} + \epsilon_4 \frac{GQ^2}{2\pi\bar{r}_H^3} + \epsilon_6 \frac{GQP}{2\pi\bar{r}_H^3} \\ &\quad + \epsilon_8 \frac{GP^2}{2\pi\bar{r}_H^3} + \mathcal{O}(\epsilon^2), \end{aligned} \quad (\text{B41})$$

where we have used

$$r_H = \bar{r}_H + \mathcal{O}(\epsilon). \quad (\text{B42})$$

This shift of the areal radius of the horizon is consistent with Eq. (81).

## 2. Falling conditions

The action (1) gives a theory for a charged particle where the particle is minimally coupled with gravity. One may consider a case where the charged particle has the higher-order derivative coupling

$$\begin{aligned} S^{\text{test}}[x^\mu] &= \int d\tau \left[ -m \sqrt{-\left( g_{\mu\nu}(x) + \delta g_{\mu\nu}^{(p)}(x) \right) \dot{x}^\mu \dot{x}^\nu} \right. \\ &\quad \left. + qA_\mu(x) \dot{x}^\mu + pB_\mu(x) \dot{x}^\mu \right], \end{aligned} \quad (\text{B43})$$

where

$$\begin{aligned} \delta g_{\mu\nu}^{(p)} &= \mu_1 R_{\mu\nu} + \mu_2 g_{\mu\nu} R + 8\pi G \mu_3 F_{\mu\rho} F^{\nu\rho} \\ &\quad + 8\pi G \mu_4 g_{\mu\nu} F_{\rho\sigma} F^{\rho\sigma} + 8\pi G \mu_5 F_{\mu\rho} P^{\nu\rho} \\ &\quad + 8\pi G \mu_6 g_{\mu\nu} F_{\rho\sigma} P^{\rho\sigma} + 8\pi G \mu_7 P_{\mu\rho} P^{\nu\rho} \\ &\quad + 8\pi G \mu_8 g_{\mu\nu} P_{\rho\sigma} P^{\rho\sigma}. \end{aligned} \quad (\text{B44})$$

Here,  $\mu_i$ 's are  $\mathcal{O}(\epsilon)$  constants. Although this theory is reduced to that with the action (1) by the field redefinition (B25) with  $\epsilon_i = -\mu_i$ , we can directly show that falling conditions becomes the same as those obtained in Sec. II.

The analysis is very simple. We have a degree of freedom for the parameter  $\tau$  of the worldline  $x(\tau)$ . Instead of  $g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu = -1$ , we set the parameter  $\tau$  by  $(g_{\mu\nu} + \delta g_{\mu\nu}^{(p)}) \dot{x}^\mu \dot{x}^\nu = -1$ . Then, all the equations become those in Sec. II with replacing  $g_{\mu\nu}$  by  $g_{\mu\nu} + \delta g_{\mu\nu}^{(p)}$ . Though we used the areal radius in Sec. II, the same result holds



with any radial coordinate. Thus, we can obtain the result simply by replacing  $f$  and  $h$  with the corresponding components of  $g_{\mu\nu} + \delta g_{\mu\nu}^{(p)}$ . As similar to the discussion in the previous subsection, the event horizon for  $g_{\mu\nu} + \delta g_{\mu\nu}^{(p)}$  is still located at  $r = r_H$  and hence the value of the potential

height at the event horizon, the right hand side of Eq. (18), is unchanged. In addition, since  $\delta g_{\mu\nu}^{(p)}$  disappears at  $r \rightarrow \infty$ , the definition of  $E$  is also unchanged. Therefore, the falling conditions with action (B43) become the same as those with action (1) whose metric is  $g_{\mu\nu} + \delta g_{\mu\nu}^{(p)}$ .

- 
- [1] R. Penrose, Gravitational collapse: The role of general relativity, *Riv. Nuovo Cimento* **1**, 252 (1969).
- [2] R. Wald, Gedanken experiments to destroy a black hole, *Ann. Phys. (N.Y.)* **82**, 548 (1974).
- [3] J. M. Cohen and R. Gautreau, Naked singularities, event horizons, and charged particles, *Phys. Rev. D* **19**, 2273 (1979).
- [4] T. Needham, Cosmic censorship and test particles, *Phys. Rev. D* **22**, 791 (1980).
- [5] I. Semiz, Dyon black holes do not violate cosmic censorship, *Classical Quantum Gravity* **7**, 353 (1990).
- [6] J. D. Bekenstein and C. Rosenzweig, Stability of the black hole horizon and the Landau ghost, *Phys. Rev. D* **50**, 7239 (1994).
- [7] I. Semiz, Dyonic Kerr-Newman black holes, complex scalar field and cosmic censorship, *Gen. Relativ. Gravit.* **43**, 833 (2011).
- [8] B. Gwak and B.-H. Lee, Cosmic censorship of rotating anti-de Sitter black hole, *J. Cosmol. Astropart. Phys.* **02** (2016) 015.
- [9] B. Gwak, Weak cosmic censorship conjecture in Kerr-(anti-)de Sitter black hole with scalar field, *J. High Energy Phys.* **09** (2018) 081.
- [10] R. Ghosh, C. Fairoos, and S. Sarkar, Overcharging higher curvature black holes, *Phys. Rev. D* **100**, 124019 (2019).
- [11] R. Ghosh, A. K. Mishra, and S. Sarkar, Overcharging extremal black holes, *Phys. Rev. D* **104**, 104043 (2021).
- [12] V. E. Hubeny, Overcharging a black hole and cosmic censorship, *Phys. Rev. D* **59**, 064013 (1999).
- [13] F. de Felice and Y.-Q. Yu, Turning a black hole into a naked singularity, *Classical Quantum Gravity* **18**, 1235 (2001).
- [14] S. Hod, Cosmic censorship, area theorem, and selfenergy of particles, *Phys. Rev. D* **66**, 024016 (2002).
- [15] T. Jacobson and T. P. Sotiriou, Over-spinning a black hole with a test body, *Phys. Rev. Lett.* **103**, 141101 (2009); **103**, 209903(E) (2009).
- [16] 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).
- [17] A. Saa and R. Santarelli, Destroying a near-extremal Kerr-Newman black hole, *Phys. Rev. D* **84**, 027501 (2011).
- [18] S. Gao and Y. Zhang, Destroying extremal Kerr-Newman black holes with test particles, *Phys. Rev. D* **87**, 044028 (2013).
- [19] I. T. Drummond and S. J. Hathrell, QED vacuum polarization in a background gravitational field and its effect on the velocity of photons, *Phys. Rev. D* **22**, 343 (1980).
- [20] Y. Kats, L. Motl, and M. Padi, Higher-order corrections to mass-charge relation of extremal black holes, *J. High Energy Phys.* **12** (2007) 068.
- [21] C. Cheung, J. Liu, and G. N. Remmen, Proof of the weak gravity conjecture from black hole entropy, *J. High Energy Phys.* **10** (2018) 004.
- [22] Y. Hamada, T. Noumi, and G. Shiu, Weak gravity conjecture from unitarity and causality, *Phys. Rev. Lett.* **123**, 051601 (2019).
- [23] C. R. T. Jones and B. McPeak, The black hole weak gravity conjecture with multiple charges, *J. High Energy Phys.* **06** (2020) 140.
- [24] B. Leaute and B. Linet, Selfinteraction of a point charge in the Kerr space-time, *J. Phys. A* **15**, 1821 (1982).
- [25] 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**, 104014 (2017).
- [26] J. Jiang, Static charged Gauss-Bonnet black holes cannot be overcharged by the new version of gedanken experiments, *Phys. Lett. B* **804**, 135365 (2020).
- [27] X.-Y. Wang and J. Jiang, Gedanken experiments at high-order approximation: Nearly extremal Reissner-Nordström black holes cannot be overcharged, *J. High Energy Phys.* **05** (2020) 161.
- [28] K. Düztaş, The variational method, backreactions, and the absorption probability in Wald type problems, *Eur. Phys. J. C* **81**, 49 (2021).
- [29] A. Sang and J. Jiang, Gedanken experiments at high-order approximation: Kerr black hole cannot be overspun, *J. High Energy Phys.* **09** (2021) 095.
- [30] X.-Y. Wang and J. Jiang, Gedanken experiments at high-order approximations: Kerr-Newman black hole cannot be overcharged and overspun, *Phys. Rev. D* **106**, 064050 (2022).
- [31] B. Chen, F.-L. Lin, B. Ning, and Y. Chen, Constraints on low-energy effective theories from weak cosmic censorship, *Phys. Rev. Lett.* **126**, 031102 (2021); **126**, 119903(E) (2021).
- [32] J. Jiang, A. Sang, and M. Zhang, Comment on ‘‘Constraints on low-energy effective theories from weak cosmic censorship’’, arXiv:2101.10172.

- [33] F.-L. Lin, B. Ning, and Y. Chen, Weak cosmic censorship and the second law of black hole thermodynamics in higher derivative gravity, *Phys. Rev. D* **108**, 044025 (2023).
- [34] B. Holdom, Two U(1)'s and epsilon charge shifts, *Phys. Lett.* **166B**, 196 (1986).
- [35] K. Izumi, Causal structures in Gauss-Bonnet gravity, *Phys. Rev. D* **90**, 044037 (2014).
- [36] H. Reall, N. Tanahashi, and B. Way, Causality and hyperbolicity of Lovelock theories, *Classical Quantum Gravity* **31**, 205005 (2014).
- [37] H. S. Reall, Causality in gravitational theories with second order equations of motion, *Phys. Rev. D* **103**, 084027 (2021).