

Duality-invariant higher-derivative corrections to charged stringy black holes

Upamanyu Moitra^{*}

*Institute for Theoretical Physics, Institute of Physics, Universiteit van Amsterdam,
Science Park 904, 1098 XH Amsterdam, The Netherlands*



(Received 9 July 2026; accepted 25 August 2026; published 9 September 2026)

We study duality-invariant higher-derivative corrections to the charged black hole geometry in two-dimensional heterotic string theory. We illustrate how the conventional perturbative approach to determine the corrected geometry breaks down. Using a nonperturbative (in α') parametrization of the solution, we find the corrected charge-to-mass ratio for extremal black holes. We remark on the results in relation to the weak gravity conjecture. We also consider the entropy of the extremal black hole within the attractor mechanism and find that the two-derivative entropy is not renormalized to any order. We make comments on interpretations of the results and their extension to near-extremal black holes.

DOI: [10.1103/s3qf-q9dk](https://doi.org/10.1103/s3qf-q9dk)

Introduction. Black holes (BHs) are essential objects in our efforts to understand various features of quantum gravity (QG). String theory, a leading candidate for a unified theory of all known interactions, has elucidated many aspects of QG. While general questions involving BHs in string theory can be intractably difficult to answer, a consideration of lower spacetime dimensions often simplifies computations, while offering insights about general aspects of gravity. Even in low dimensions, however, it is extremely important to take into account the straitjackets that make string theory compelling as a theory of QG. Among such constraints, target space duality or T duality [1], describing the equivalence of physics on small and large circles, is particularly important. It is known that, at tree level, strings compactified on a d -dimensional torus have $O(d, d; \mathbb{R})$ as an exact duality symmetry group [2].

In the recent years, there has been significant progress in understanding duality-symmetric theories with double field theory-based methods [3]. Inspired by early work [4], there now exists a complete classification of perturbative higher-derivative (HD) corrections for cosmological and $1+1$ -dimensional static backgrounds [5] in the massless sector of closed string theory. In recent work [6], we extended this classification program to include heterotic strings and the associated gauge fields, truncated to the Cartan subgroup of the gauge group. For a single Maxwell field, the duality

group is $O(1, 2; \mathbb{R})$. The extension of the formalism in [6] allows for a detailed study of two-dimensional charged BHs in the heterotic theory, first discovered in [7], and various interesting associated questions. We also showed in [6] how a nonperturbative parametrization [8,9] of the solution in the presence of arbitrary HD terms could be extended to the heterotic case.

The goal of the present work is to study some physically interesting aspects of charged BHs within this framework. Charged BHs with degenerate horizons are known as “extremal.” Extremal BHs have played a crucial role in illuminating various aspects of QG, from microscopic state counting [10] to the ongoing efforts of mapping the landscape of consistent QG theories [11], known as the “Swampland” program. One statement of the weak gravity conjecture (WGC) [12], a central part of this program, is that extremal BHs should be able to decay by emitting particles of sufficiently high charge-to-mass ratio. It was also suggested [13] that BHs themselves could meet the criterion of such particles: consistent HD corrections in the low-energy effective field theory (EFT) were argued to increase the extremal charge-to-mass ratio. This has been explicitly checked in several scenarios [14].

One motivation of this work is addressing such questions in the two-dimensional framework. There are many known subtleties related to the Swampland program in $d = 2$ [15]. For instance, the “no-global-symmetries” conjecture is believed to be generally valid only for $d \geq 4$, when gravity is dynamical. In $d = 2$, both gravity and electromagnetism are nondynamical and the electrostatic potential for a point charge grows linearly with distance. A general statement of WGC cannot be easily extended to $d = 2, 3$; see [16]. Therefore, it is not *a priori* clear what a Swampland-like statement would resemble and whether one can make any

^{*}Contact author: u.moitra@uva.nl

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 SCOAP³.

such universal statement at all. On the other hand, the BH we consider is a genuine string background arising from a world sheet description. It has been argued [17] that continuous global symmetries on a world sheet always translate to gauge symmetries in spacetime. Given this, one does expect an absence of global symmetries in the background and consequently, something akin to WGC. We hope that this work is the first step toward addressing such subtle questions.

In what follows, we explore various physical consequences of duality-invariant HD corrections. First, we calculate the corrected extremal charge-to-mass ratio. While part of the motivation was outlined in the previous paragraph, this calculation, not previously attempted in the literature, is of broader interest in BH physics. We first perform the computation in perturbation theory with only the four-derivative terms and encounter a surprising difficulty. To circumvent this, we use a nonperturbative formalism [6,8,9]. Pleasingly, we are able to write down an explicit formula for the ratio, with a remarkable feature, which we comment upon.

We next turn our attention to the calculation of Wald entropy [18] of the extremal BH with the HD terms within the attractor mechanism [19]. To this end, we apply the entropy function formalism [20] proposed by Sen. We find yet another striking result concerning the nonrenormalization of the two-derivative entropy. We conclude the Letter with some comments on interpretations of the foregoing results, BHs slightly away from extremality, and possible future directions. We have relegated some formulas to the End Matter.

Corrected extremal charge-to-mass ratio. The two-derivative target space action describing the massless sector of heterotic strings involving the metric $g_{\mu\nu}$, string dilaton ϕ , and Maxwell field A_μ is

$$I_{(2)} = \int d^2x \sqrt{-g} e^{-2\phi} \left[\xi^2 + R + 4(\nabla\phi)^2 - \frac{F_{\mu\nu}F^{\mu\nu}}{4} \right], \quad (1)$$

where $\xi^2 \equiv 8/\alpha'$, $\sqrt{\alpha'}$ being the string length. We have absorbed the Newton constant in ϕ . The full action is augmented with HD terms, involving combinations of R , $F_{\mu\nu}$, and $\nabla_\mu\phi$ and their derivatives. While the general HD action is complicated in form, enormous simplifications occur under the assumption of duality in a time-independent background described by

$$\begin{aligned} ds^2 &= -m(x)^2 dt^2 + n(x)^2 dx^2, \\ \phi &= \phi(x), \quad A_\mu = V(x) \delta_\mu^t. \end{aligned} \quad (2)$$

Defining the duality-invariant dilaton

$$\Phi(x) = 2\phi(x) - \log |m(x)| \quad (3)$$

under appropriate field redefinitions [5,6], the full action, dimensionally reduced, can be shown to assume a remarkably simple $O(1, 2; \mathbb{R})$ -invariant form,

$$I_{\text{full}} = \int dx n(x) e^{-\Phi} \left(\xi^2 + \frac{\Phi'(x)^2}{n(x)^2} + F(\rho) \right), \quad (4)$$

where $F(\rho)$ is a function of the single variable ρ , a duality-invariant combination of the curvature and gauge field strength,

$$\begin{aligned} F(\rho) &= -\frac{1}{2}\rho^2 + \sum_{n=2}^{\infty} c_{2n} \rho^{2n}, \\ \rho^2 &\equiv \frac{2m'(x)^2 - V'(x)^2}{n(x)^2 m(x)^2}. \end{aligned} \quad (5)$$

We can generalize the discussion to include r Abelian gauge fields, leading to the symmetry group $O(1, 1 + r; \mathbb{R})$. Remarkably, duality is so constricting that there is only one Wilson coefficient at each derivative order [6]. For definiteness, we consider a single Maxwell field.

The two-derivative action admits the well-known charged BH solution [7]

$$\begin{aligned} \bar{m}(x)^2 &= 1 - 2\mu e^{-\xi x} + Q^2 e^{-2\xi x} = \bar{n}(x)^{-2}, \\ \bar{\phi}(x) &= -\frac{1}{2}\xi x, \quad \bar{V}(x) = -\sqrt{2}Q e^{-\xi x}. \end{aligned} \quad (6)$$

The physical mass and charge are $2\xi\mu$ and $\sqrt{2}\xi Q$, respectively, but we will keep using the parameters μ and Q . The BH is nonextremal for $\mu > |Q|$ and extremal for $\mu = |Q|$. Henceforth, we consider a positive Q . The BH has horizons at $x_{\pm} = \xi^{-1} \log(\mu \pm \sqrt{\mu^2 - Q^2})$.

HD terms are known to modify such a linear mass-charge relationship, characteristic of asymptotically flat extremal BHs in higher dimensions. It was suggested [13] that the modification occurs in such a way that $(Q/\mu)_{\text{ext}} \geq 1$. Finding the nature of HD-corrected solutions in 2D, besides their relevance for a possible formulation of WGC, is a worthwhile goal in itself, a result that seems to be missing from the literature.

Failure of perturbation theory: We employ perturbation theory to calculate the corrected $(Q/\mu)_{\text{ext}}$ due to the duality-invariant HD terms. In the function $F(\rho)$ in (5), we truncate the series at the four-derivative term with $c_4 = \epsilon c$ ($\epsilon \ll 1$) and work to $\mathcal{O}(\epsilon)$. The background fields (6) receive $\mathcal{O}(\epsilon)$ corrections,

$$m = \bar{m} + \epsilon \delta m, \quad n = \bar{n} + \epsilon \delta n, \quad V = \bar{V} + \epsilon \delta V. \quad (7)$$

We choose a gauge where the string dilaton is fixed to $\phi = \bar{\phi}$. The dilaton actually defines the spatial coordinate x : leaving it fixed is analogous to considering the size of the transverse space fixed in higher dimensions. We can solve

for the equations of motion arising from (4) and imposing suitable boundary conditions at asymptotic infinity, $x \rightarrow \infty$, to ensure that the corrections are subleading. The solution is given in Eq. (A2).

Rather surprisingly, perturbation theory breaks down near the horizon. Near the unperturbed event horizon $x = x_+$, the curvature and electric field strength behave as, see Eq. (A3),

$$R \sim \frac{\epsilon c \xi^2}{(x - x_+)^2}, \quad F^2 \sim \frac{\epsilon c \xi^3}{x - x_+}, \quad (8)$$

which indicates a failure of the perturbative approach near the nonextremal horizon. Such a dramatic breakdown of perturbation theory is disconcerting and surprising from an EFT point of view. Thus, any hope of systematically studying the effects of HD terms on BHs would appear to be lost.

Nonperturbative parametrization: There does exist a framework to obtain a solution in presence of a full tower of HD terms. In [6], a nonperturbative parametrization for charged BHs was derived, building on the works of [8,9]. The solution proceeds from the assumption that the derivative $f(\rho) = F'(\rho)$ of the function appearing in (5) has a single-valued inverse $\rho = \rho(f)$. In this f parametrization, the duality-invariant dilaton is simply $\Phi = \log(f/\xi) + \Phi_0$ and in the gauge $n(x) = 1$, the BH exterior is described by [6]

$$\begin{aligned} x &= x_0 + \int_f^\infty df' \frac{1}{f' \sqrt{\xi^2 - G(f')}}, \\ m(f)^2 &= \frac{1}{\mathcal{Q}^2} \text{csch}^2 \left[\int_{f_m}^f df' \frac{\rho(f')}{\sqrt{2f' \sqrt{\xi^2 - G(f')}}} \right], \\ V(f) &= \mathcal{Q} \int_0^f df' \frac{\rho(f') m(f')^2}{f' \sqrt{\xi^2 - G(f')}}, \end{aligned} \quad (9)$$

where x_0 , $\mathcal{Q}$, and $f_m < 0$ are constants and $G(f) = \int_0^f df' \rho(f')$.

For the two-derivative theory, $\rho(f) = -f$. For an HD extension, one considers [9] a perturbation of the form

$$\rho(f) = -f(1 + h(f^2)), \quad (10)$$

where $h(z)$ is a function satisfying $h(0) = 0$ and decaying sufficiently fast at infinity. The function is bounded in a way so that $\xi^2 - G(f) > 0$ between $f = 0$ (spatial infinity) and $f = \infty$ (event horizon). In terms of the function

$$\beta(y) \equiv \int_0^y dy' h(y'), \quad (11)$$

we have

$$G(f) = -\frac{1}{2}(f^2 + \beta(f^2)). \quad (12)$$

Thus, $h(f^2) = \beta'(f^2)$. For a BH solution, $\beta(\infty) \equiv \beta_0$ must be finite [9]. We can now see why the example of the ρ^4 term in the previous subsection failed to work. In this case, for large f , $h(f^2) \sim 1$ and hence $|\beta_0| = \infty$ and there is no BH. A related example was also studied in [9].

We can now readily make use of the general solution (9) to find the extremal charge-to-mass ratio. It is convenient to express the solution in Schwarzschild-Droste-like gauge (6). To be precise, we shall use the various transformation formulas to use the string dilaton $2\phi = \Phi + \log m$ as a coordinate, whereby we obtain a line element of the form

$$ds^2 = -m(\phi)^2 dt^2 + n(\phi)^2 \frac{4d\phi^2}{\xi^2}. \quad (13)$$

We first perform an expansion about asymptotic infinity. The solution is given by (A5); up to the order shown, it is identical in form to the two-derivative result (6). The HD terms change only the terms subleading to (6). We thus identify $\mathcal{Q} = \text{csch} \gamma$ and the charge-to-mass ratio

$$\frac{\mathcal{Q}}{\mu} = \text{sech} \gamma, \quad (14)$$

where

$$\gamma \equiv \int_{f_m}^0 df' \frac{1 + h(f'^2)}{\sqrt{2\xi^2 + f'^2 + \beta(f'^2)}}. \quad (15)$$

The extremality condition, on the other hand, must be deduced from a near-horizon ($f = \infty$) expansion, Eq. (A6). In this case, assuming $\beta(f^2)$ admits an expansion of the form $\sum_{n=0}^\infty \beta_n f^{-2n}$ (neglecting possible exponentially small corrections), we find from the near-horizon metric (A7) the extremality condition to be given by $\Delta = 0$, where

$$\Delta \equiv 4\xi^2 - e^{2(\gamma+\chi)}(\beta_0 + 2\xi^2), \quad (16)$$

with

$$\begin{aligned} \chi &\equiv \int_0^\xi df' \frac{1 + h(f'^2)}{\sqrt{2\xi^2 + f'^2 + \beta(f'^2)}} \\ &+ \int_\xi^\infty df' \left(\frac{1 + h(f'^2)}{\sqrt{2\xi^2 + f'^2 + \beta(f'^2)}} - \frac{1}{f'} \right). \end{aligned} \quad (17)$$

We have to perform a suitable rescaling of the coordinates to transform the metric components (A7) to a more familiar form. Note that, in general, the coordinate system breaks down in the limit $\Delta \rightarrow 0$. This is not unexpected from the discussion in [6]. [However, as clarified in the End Matter, the limit is smooth for exponentially small

corrections to $\beta(\infty)$: when $\beta_n = 0$ for $n > 0$.] Since our goal is to find precisely this limiting value, we need not worry about this possibility. Setting $\Delta = 0$ gives the corrected extremal charge-to-mass ratio,

$$\left(\frac{Q}{\mu}\right)_{\text{ext}} = \frac{4\xi\sqrt{\beta_0 + 2\xi^2}}{4e^{-\chi}\xi^2 + e^\chi(\beta_0 + 2\xi^2)}. \quad (18)$$

This is one of the main results in this Letter, which expresses the extremal charge-to-mass ratio in terms of the parameters of the theory. Notice that this relation is state independent. By the arithmetic mean-geometric mean inequality, we immediately see that, irrespective of the form of the corrections, $(Q/\mu)_{\text{ext}} \leq 1$. This bound is completely general in a duality-invariant theory and cannot be changed by tuning the Wilson coefficients. The two-derivative theory saturates the bound but is not unique in this regard.

We see that the bound is opposite of the usual WGC bound in higher dimensions. This is not too surprising in the present two-dimensional scenario, for reasons mentioned in the Introduction. It would be interesting to further explore related aspects in this setup.

Attractor mechanism and entropy. Extremal BHs have also been crucial in our understanding of BH thermodynamics and statistical mechanics, especially *vis-à-vis* state counting [10]. Extremal BHs often exhibit the attractor mechanism [19], in which the near-horizon properties of the BH are determined only by the charges carried by it and not by asymptotic moduli. In the near-horizon region, the BHs have a two-dimensional anti-de Sitter (AdS_2) geometry and a covariantly constant electric field. The entropy function formalism [20] is an elegant method to calculate the Wald entropy [18] due to the HD terms in the action. We now turn our attention to this problem. The two-dimensional heterotic BH entropy was also considered in [21] in a specific, limited context with the HD corrections involving only the Ricci scalar.

In this formalism, one starts with a two-dimensional Lagrangian density (possibly after integrating over the transverse directions, which does not apply here). Inserting in this Lagrangian, the following *Ansätze* for the metric, gauge field, and scalar,

$$\begin{aligned} ds^2 &= \mathbf{v} \left(-x^2 dt^2 + \frac{dx^2}{x^2} \right), \\ V'(x) &= \mathbf{e}, \quad \phi(x) = \phi_0, \end{aligned} \quad (19)$$

one obtains the function $\mathbf{f}(\phi_0, \mathbf{v}, \mathbf{e})$. One then defines the Legendre transform

$$\mathcal{E}(\phi_0, \mathbf{v}, \mathbf{e}, \mathbf{q}) = 2\pi(\mathbf{e}\mathbf{q} - \mathbf{f}(\phi_0, \mathbf{v}, \mathbf{e})) \quad (20)$$

and sets each of $\partial\mathcal{E}/\partial\mathbf{v}$, $\partial\mathcal{E}/\partial\phi_0$, and $\partial\mathcal{E}/\partial\mathbf{e}$ to zero (known

as the “attractor equations”). Inserting the solution in (20), one obtains the entropy.

It is worth explaining one technical detail at length. The Lagrangian density we use must be supplemented with an additional total-derivative term in comparison with (4),

$$\mathcal{L} = ne^{-\Phi} \left(\xi^2 + \frac{\Phi'^2}{n^2} + F(\rho) \right) - 2 \frac{d}{dx} \left(\frac{e^{-2\Phi} m'}{n} \right). \quad (21)$$

While the inclusion of this term does not affect the equations of motion, it does matter in the entropy computation. At the two-derivative level, the generally covariant Lagrangian (1) reduces precisely to the form (21). In obtaining the duality-invariant form (4), we discard the total derivative [6]. In the derivation of the HD action (4), one performs repeated integration by parts [5,6]. These two procedures are, however, not on equal footing. In the second case, we use the two-derivative equations of motion to perform field redefinitions. These are consistent with covariant field redefinitions; see, for example, [4]. It is known that the entropy evaluated in this formalism is invariant under such field redefinitions [20].

Inserting the *Ansätze* (19) in (21), the suitable combinations of the first two attractor equations yield

$$\begin{aligned} \frac{\mathbf{e}^2 - 2\mathbf{v}}{2\mathbf{v}^2} + \sum_{n=2}^{\infty} c_{2n} \left(\frac{2\mathbf{v} - \mathbf{e}^2}{\mathbf{v}^2} \right)^n - \frac{1 - \mathbf{v}\xi^2}{\mathbf{v}} &= 0, \\ (\mathbf{e}^2 - \mathbf{v}) \left[1 - 2 \sum_{n=2}^{\infty} n c_{2n} \left(\frac{2\mathbf{v} - \mathbf{e}^2}{\mathbf{v}^2} \right)^{n-1} \right] - \mathbf{v} &= 0. \end{aligned} \quad (22)$$

It is not hard to see that the equations admit the exact, physically unique solution $\mathbf{e}^2 = 2/\xi^2 = 2\mathbf{v}$. From the third attractor equation, we also find $e^{-2\phi_0} = |\mathbf{q}|/(\sqrt{2}\xi)$, whereby we find the entropy

$$S = 2\sqrt{2}\pi\xi^{-1}|\mathbf{q}|, \quad (23)$$

which exhibits no dependence on the Wilson coefficients. The HD terms therefore do not renormalize the leading order two-derivative entropy. This nonrenormalization is another key result of this Letter.

It is worth pointing out some interesting features of the entropy. Note that the AdS_2 radius is of order the string length, $L_{\text{AdS}} = \sqrt{\alpha'}/8$. This necessitates that we take into account all HD terms in the effective action. We implicitly assumed in (5) that the coefficients c_{2n} decay sufficiently fast so that the infinite sum converges at least within some finite radius; see [9]. Note that the duality-invariant combination ρ^2 ensures that nonrenormalization persists even if we truncate the sum in (5) to any finite number of terms. The constant dilaton scenario here is very different from the linear dilaton case before in that it is possible to find a consistent BH solution even with a finite number of HD terms. Furthermore, unlike in many typical examples,

the near-horizon electric field e is not a free parameter but fixed by the attractor equations. It would be interesting to examine if the quantization of charge plays any role in this nonrenormalization.

We emphasize that the nonrenormalization is noteworthy for the reason that we worked in an EFT framework with no input apart from duality invariance. One is compelled to wonder whether this EFT “knows” physics beyond this input. A relation between the attractor mechanism and topological strings has been conjectured in [22]. It would be interesting to examine if the entropy is protected because it describes some topological sector of the theory.

Discussion. The different results in the previous sections are intriguing from multiple perspectives. Upon the introduction of a single duality-invariant HD term, perturbation theory breaks down rather dramatically, against EFT expectations, and a singularity seems to develop in the vicinity of the BH horizon, originally present in the two-derivative theory. It would be interesting to examine the formation and dynamical behavior of such a singularity [23].

The situation is ameliorated by the introduction of an infinite number of HD terms, satisfying certain conditions outlined previously. Using the nonperturbative form of the solutions, we calculated an explicit form of the HD correction to the extremal charge-to-mass ratio. Remarkably, we found a specific bound for this ratio, albeit one that goes in the opposite direction compared to higher dimensions [13]. The existence of the bound itself is indicative of universality of some kind: it would be interesting to find similar properties of the two-dimensional string landscape in the spirit of the Swampland program.

In relation to the extremal charge-to-mass ratio, the nonrenormalization of the attractor entropy might seem a little puzzling at first. It was argued in [24] that the difference in entropy due to HD terms is non-negative—a result that must hold, in general—and also proportional to the correction to the charge-to-mass ratio. While we have $\Delta S = 0$, the second part of the argument does not extend to our case. There is no contradiction because the higher-dimensional derivation relied on perturbation theory, which we have found to break down in 2D.

The results can be easily generalized to r Maxwell fields [6]. For instance, the form of (18) would be the same with the replacement $Q \rightarrow |\vec{Q}|$. In the entropy calculation, one finds similarly $\tilde{e}^2 = 2/\xi^2 = 2\mathbf{b}$. This further suggests a

role of charge quantization in the nonrenormalization of the entropy, since only one out of r Maxwell fields seems to be relevant to the story. One would have, in this case, similar to (23), $S = 2\sqrt{2}\pi\xi^{-1}|\vec{q}|$.

The nonrenormalization of the entropy invites a more refined understanding of the underlying mechanism. It is known that [25,26] that the two-dimensional extremal BH can be obtained as an orbifold of an extremal AdS_3 BH. Nonrenormalization of the microscopic entropy in AdS_3 has been argued for in [27]. It would be interesting to see how our low-energy EFT captures such microscopic details irrespective of the values of the Wilson coefficients. It is worth emphasizing that, without insisting on duality invariance, such a nonrenormalization would not hold true: generic HD corrections would modify (23). We have only considered the string tree-level action [28]. A consideration of the quantum entropy [29] would shed more light on the microscopic aspects—the duality group would also be broken down to a discrete subgroup. In this respect, it might be worth considering the torus partition function of strings on orbifolds of AdS_3 [30].

Considering a small departure from extremality at the attractor point would be an interesting direction to explore. While the Jackiw-Teitelboim (JT) model [31], suitable for describing BHs near extremality [32], is typically obtained via dimensional reduction, we can directly obtain the JT model and the corresponding Schwarzschild from the two-derivative action in the near-horizon region without any reduction. One can combine the results of [32,33] and directly obtain the $T^{3/2}$ correction to the BH partition function. The relative simplicity of two-dimensional string theory would perhaps facilitate comparison between the effective and microscopic theories. We leave such promising questions for future investigations.

Acknowledgments. My research is supported by the European Union’s Horizon 2020 Research and Innovation Program (Grant Agreement No. 101115511).

Data availability. There are no publicly available research data or software supporting this manuscript. Requests for further information or data should be sent to the authors.

Appendix: Perturbative and nonperturbative solutions to the equations of motion. For the perturbative analysis, as described in the main text, we consider the HD action of the form (4) with a single four-derivative term,

$$I_{(4)} = \int dx n(x) e^{-\Phi(x)} \left[\xi^2 + \frac{\Phi'(x)^2}{n(x)^2} - \frac{2m'(x)^2 - V'(x)^2}{2n(x)^2 m(x)^2} + \epsilon c \left(\frac{2m'(x)^2 - V'(x)^2}{n(x)^2 m(x)^2} \right)^2 \right], \quad (\text{A1})$$

with $\epsilon \ll 1$ a perturbative parameter and the duality-invariant dilaton Φ related to the string dilaton ϕ by (3). With the $\mathcal{O}(\epsilon^0)$ solution given by (6), we consider perturbative *Ansätze* of the form (7). Solving the resulting equations of motion and suitably choosing constants so as not to change the asymptotically measured mass and charge, we find

$$\begin{aligned}\delta m(x) &= \frac{c\xi^2 e^{-2\xi x} (e^{\xi x} \mu - Q^2)}{\bar{m}(x)} \left[\frac{(2Q^2 - 3\mu^2) e^{-2\xi x} + 2e^{-\xi x} \mu - 1}{\bar{m}(x)^2} + \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \coth^{-1} \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \right], \\ \delta n(x) &= \frac{2c\xi^2 e^{-\xi x} (\mu^2 - Q^2)}{\bar{m}(x)^3} \left[\frac{9e^{-2\xi x} (e^{\xi x} - \mu)}{2\bar{m}(x)^2} - 4e^{-\xi x} - \frac{1}{2\sqrt{\mu^2 - Q^2}} \coth^{-1} \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \right], \\ \delta V(x) &= \sqrt{2} c \xi^2 Q e^{-\xi x} \left[\frac{(2Q^2 - 3\mu^2) e^{-2\xi x} + 2e^{-\xi x} \mu - 1}{\bar{m}(x)^2} + \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \coth^{-1} \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \right].\end{aligned}\quad (\text{A2})$$

The scalar invariants up to $\mathcal{O}(\epsilon)$ are given by

$$\begin{aligned}R &= 2e^{-2\xi x} \xi^2 (e^{\xi x} \mu - 2Q^2) - 2\epsilon c \xi^4 e^{-\xi x} \sqrt{\mu^2 - Q^2} \coth^{-1} \frac{e^{\xi x} - \mu}{\sqrt{\mu^2 - Q^2}} \\ &\quad + 2\epsilon c \xi^4 e^{-2\xi x} (\mu^2 - Q^2) \frac{1 + 21e^{-\xi x} \mu - e^{-2\xi x} (41Q^2 + 36\mu^2) + 87e^{-3\xi x} \mu Q^2 - 32e^{-4\xi x} Q^4}{\bar{m}(x)^4}, \\ F_{\mu\nu} F^{\mu\nu} &= -4\xi^2 Q^2 e^{-2\xi x} - \epsilon c \xi^4 e^{-4\xi x} \frac{64Q^2 (\mu^2 - Q^2)}{\bar{m}(x)^2}.\end{aligned}\quad (\text{A3})$$

Since the redshift factor $\bar{m}(x)^2$, with the unperturbed horizon, is present in the denominator in the scalar invariants, we conclude that perturbation theory breaks down in the vicinity of the horizon.

We next turn to the nonperturbative solution parametrized as (9). In order to determine the mass and charge at asymptotic infinity, we expand various quantities about $f = 0$, which yields

$$\begin{aligned}\frac{dx}{df} &= -\frac{1}{\xi f} + \frac{f}{4\xi^3} - \frac{3 - 4\xi^2 \beta''(0)}{32\xi^5} f^3 + \dots, \\ m(f)^2 &= \frac{\text{csch}^2 \gamma}{Q^2} \left[1 - \frac{\sqrt{2} f}{\xi} \coth \gamma + \frac{2 + \cosh 2\gamma}{2} \text{csch}^2 \gamma \frac{f^2}{\xi^2} + \dots \right], \\ 2\phi &= \log \frac{f}{\xi} + \underbrace{\Phi_0 + \log \frac{\text{csch} \gamma}{Q}}_{\equiv 2\phi_1} - \frac{f \coth \gamma}{\sqrt{2}\xi} + \frac{f^2 \text{csch}^2 \gamma}{4\xi^2} + \dots,\end{aligned}\quad (\text{A4})$$

where γ is defined in Eq. (15).

Inverting the last expansion, we obtain the expansion of the metric and gauge field components in the string dilaton coordinate,

$$\begin{aligned}m(\phi)^2 &= \frac{\text{csch}^2 \gamma}{Q^2} \left[1 - \sqrt{2} \coth \gamma e^{2(\phi - \phi_1)} + \frac{e^{4(\phi - \phi_1)}}{2} \text{csch}^2 \gamma + \dots \right], \\ n(\phi)^2 &= 1 + \sqrt{2} \coth \gamma e^{2(\phi - \phi_1)} + \frac{1 + 2 \cosh 2\gamma}{2} \text{csch}^2 \gamma e^{4(\phi - \phi_1)} + \dots, \\ V(\phi) &= -\frac{\text{csch}^2 \gamma}{Q} e^{2(\phi - \phi_1)} + \dots.\end{aligned}\quad (\text{A5})$$

In a similar vein, the determination of the extremality condition requires a knowledge of the metric close to the BH horizon ($f = \infty$). As mentioned in the main text, we assume the *Ansatz* $\beta(f^2) = \sum_{n=0}^{\infty} \beta_n f^{-2n}$ and expand various quantities about $f = \infty$,

$$\begin{aligned}
\frac{dx}{df} &= -\frac{\sqrt{2}}{f^2} + \frac{2\xi^2}{\sqrt{2}f^4} + \frac{\beta_0}{\sqrt{2}f^4} - \frac{3(2\xi^2 + \beta_0)^2 - 4\beta_1}{4\sqrt{2}f^6} + \dots, \\
m(f)^2 &= \frac{4\xi^2 e^{-2(\gamma+\chi)}}{Q^2 f^2} + \frac{2\xi^2 e^{-4(\gamma+\chi)} \Delta}{Q^2 f^4} + \frac{e^{-6(\gamma+\chi)} \xi^2 [\Delta(5\Delta - 8\xi^2) - 12e^{4(\gamma+\chi)} \beta_1]}{4Q^2 f^6} + \dots, \\
2\phi &= \underbrace{\Phi_0 + \log \frac{2e^{-(\gamma+\chi)}}{Q}}_{\equiv 2\phi_h} + \frac{\Delta e^{-2(\gamma+\chi)}}{4f^2} - \frac{12\beta_1 - e^{-4(\gamma+\chi)} \Delta(3\Delta - 8\xi^2)}{32f^4} + \dots,
\end{aligned} \tag{A6}$$

where Δ and χ are defined in (16) and (17), respectively. Then, inverting the last expression and using the definition of the metric (13), we find the near-horizon expansion of the metric components to be

$$\begin{aligned}
Q^2 m(\phi)^2 &= \frac{16\xi^2}{\Delta} (\phi - \phi_h) + 8\xi^2 \frac{\Delta(\Delta + 8\xi^2) + 12e^{4(\gamma+\chi)} \beta_1}{\Delta^3} (\phi - \phi_h)^2 + \dots, \\
\frac{\xi^2}{n(\phi)^2} &= \frac{\Delta(\phi - \phi_h)}{2e^{2(\gamma+\chi)}} + \frac{\Delta(\Delta + 8\xi^2) - 36e^{4(\gamma+\chi)} \beta_1}{4\Delta e^{2(\gamma+\chi)}} (\phi - \phi_h)^2 + \dots.
\end{aligned} \tag{A7}$$

By making a suitable rescaling of the temporal coordinate, we can bring $m(\phi)^2$ to the more familiar form in which $m(\phi)^2 \sim \Delta(\phi - \phi_h)$. We can thus find the extremality condition to be $\Delta = 0$. Note that, as $\Delta \rightarrow 0$, the expansions become ill defined. However, in case the large- f corrections to $\beta(f^2)$ are exponentially small, the limiting behavior as $\Delta \rightarrow 0$ is smooth.

-
- [1] A. Giveon, M. Porrati, and E. Rabinovici, *Phys. Rep.* **244**, 77 (1994).
[2] K. A. Meissner and G. Veneziano, *Phys. Lett. B* **267**, 33 (1991); A. Sen, *Phys. Lett. B* **271**, 295 (1991); M. Gasperini and G. Veneziano, *Phys. Lett. B* **277**, 256 (1992); J. Maharana and J. H. Schwarz, *Nucl. Phys.* **B390**, 3 (1993).
[3] W. Siegel, *Phys. Rev. D* **48**, 2826 (1993); C. Hull and B. Zwiebach, *J. High Energy Phys.* 09 (2009) 099.
[4] K. A. Meissner, *Phys. Lett. B* **392**, 298 (1997).
[5] O. Hohm and B. Zwiebach, *Phys. Rev. D* **100**, 126011 (2019); T. Codina, O. Hohm, and B. Zwiebach, *Phys. Rev. D* **108**, 026014 (2023).
[6] U. Moitra, *J. High Energy Phys.* 05 (2026) 165.
[7] M. D. McGuigan, C. R. Nappi, and S. A. Yost, *Nucl. Phys.* **B375**, 421 (1992).
[8] M. Gasperini and G. Veneziano, *J. High Energy Phys.* 07 (2023) 144.
[9] T. Codina, O. Hohm, and B. Zwiebach, *Phys. Rev. D* **108**, 126006 (2023).
[10] A. Strominger and C. Vafa, *Phys. Lett. B* **379**, 99 (1996).
[11] C. Vafa, [arXiv:hep-th/0509212](#).
[12] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, *J. High Energy Phys.* 06 (2007) 060.
[13] Y. Kats, L. Motl, and M. Padi, *J. High Energy Phys.* 12 (2007) 068.
[14] P. A. Cano, T. Ortín, and P. F. Ramirez, *J. High Energy Phys.* 02 (2020) 175; P. A. Cano, S. Chimento, R. Linares, T. Ortín, and P. F. Ramirez, *J. High Energy Phys.* 02 (2020) 031; P. A. Cano, T. Ortín, A. Ruipérez, and M. Zatti, *J. High Energy Phys.* 03 (2022) 103.
[15] J. McNamara and C. Vafa, [arXiv:2004.06738](#).
[16] U. Moitra, *Phys. Rev. D* **109**, L041903 (2024).
[17] T. Banks and L. J. Dixon, *Nucl. Phys.* **B307**, 93 (1988).
[18] R. M. Wald, *Phys. Rev. D* **48**, R3427 (1993).
[19] S. Ferrara, R. Kallosh, and A. Strominger, *Phys. Rev. D* **52**, R5412 (1995).
[20] A. Sen, *J. High Energy Phys.* 09 (2005) 038; *Gen. Relativ. Gravit.* **40**, 2249 (2008).
[21] S. Hyun, W. Kim, J. J. Oh, and E. J. Son, *J. High Energy Phys.* 04 (2007) 057; J. Sadeghi, M. R. Setare, and B. Pourhassan, *Acta Phys. Pol. B* **40**, 251 (2009).
[22] H. Ooguri, A. Strominger, and C. Vafa, *Phys. Rev. D* **70**, 106007 (2004).
[23] U. Moitra, *J. High Energy Phys.* 06 (2023) 194.
[24] C. Cheung, J. Liu, and G. N. Remmen, *J. High Energy Phys.* 10 (2018) 004.
[25] A. Achucarro and M. E. Ortiz, *Phys. Rev. D* **48**, 3600 (1993); D. A. Lowe and A. Strominger, *Phys. Rev. Lett.* **73**, 1468 (1994); A. Giveon and A. Sever, *J. High Energy Phys.* 02 (2005) 065.
[26] I thank Y. Zigdon for pointing out a reference in [25].
[27] A. Dabholkar, A. Sen, and S. P. Trivedi, *J. High Energy Phys.* 01 (2007) 096.

- [28] Since the HD action is constructed perturbatively, one can entertain the possibility that the nonrenormalization is nonperturbatively violated.
- [29] A. Sen, *Int. J. Mod. Phys. A* **24**, 4225 (2009).
- [30] A. Dabholkar and U. Moitra, *J. High Energy Phys.* **06** (2024) 053; A. Dabholkar, E. Harris, and U. Moitra, *J. High Energy Phys.* **02** (2026) 136.
- [31] C. Teitelboim, *Phys. Lett.* **126B**, 41 (1983); R. Jackiw, *Nucl. Phys.* **B252**, 343 (1985).
- [32] U. Moitra, S.K. Sake, S.P. Trivedi, and V. Vishal, *J. High Energy Phys.* **11** (2019) 047; U. Moitra, S.P. Trivedi, and V. Vishal, *J. High Energy Phys.* **07** (2019) 055; U. Moitra, Aspects of quantum gravity, holography and entanglement, Ph.D. thesis, Tata Institute of Fundamental Research.
- [33] D. Stanford and E. Witten, *J. High Energy Phys.* **10** (2017) 008.