Weak gravity conjecture

Daniel Harlow*

Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Ben Heidenreich†

Department of Physics, University of Massachusetts Amherst, Amherst, Massachusetts 01003, USA

Matthew Reece‡

Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

Tom Rudelius§

Department of Physics, University of California, Berkeley, California 94720, USA

(published 6 September 2023)

The weak gravity conjecture holds that in a theory of quantum gravity any gauge force must mediate interactions stronger than gravity for some particles. This statement has surprisingly deep and extensive connections to many different areas of physics and mathematics. Several variations on the basic conjecture have been proposed, including statements that are much stronger but are nonetheless satisfied by all known consistent quantum gravity theories. These related conjectures and the evidence for their validity in the string theory landscape are reviewed. Also reviewed are a variety of arguments for these conjectures, which tend to fall into two categories: qualitative arguments that claim the conjecture is plausible based on general principles and quantitative arguments for various special cases or analogs of the conjecture. The implications of these conjectures for particle physics, cosmology, general relativity, and mathematics are also outlined. Finally, important directions for future research are highlighted.

DOI: 10.1103/RevModPhys.95.035003

CONTENTS

*harlow@mit.edu
†bheidenreich@umass.edu
‡mreece@g.harvard.edu
§rudelius@berkeley.edu

I. INTRODUCTION

The weak gravity conjecture is a simple statement about theories of quantum gravity. In essence, it says that any gauge force must be stronger than gravity. More precisely, in its mildest form the weak gravity conjecture holds that any U(1) gauge theory must have at least one object satisfying

$$\frac{|q|}{m} \geq \frac{|Q|}{M}\bigg|_{\mathrm{ext}}, \tag{1}$$

where $|Q|/M|_{\mathrm{ext}}$ is the charge-to-mass ratio of a large extremal black hole. This simple statement has profound consequences that touch virtually every aspect of modern fundamental physics, including string theory, cosmology, particle physics, algebraic geometry, black holes, quantum information, holography, and scattering amplitudes.

The original paper on the weak gravity conjecture (WGC) from Arkani-Hamed, Motl, Nicolis, and Vafa (AMNV) (Arkani-Hamed, Motl *et al.*, 2007) is now more than 15 years old. It sparked a flurry of research shortly after it was released, which slowly tapered off over the course of the next several years. The middle of the 2010s, however, saw a resurgence of interest in the conjecture, which has continued to the present day.

This resurgence of interest was driven in part by the hope that quantum gravity may have something to say about testable low-energy physics, despite the fact that quantum gravitational effects are naively suppressed by powers of energy divided by the Planck mass. Originally it was hoped that this problem could be circumvented using string theory to predict low-energy parameters such as Yukawa couplings or the scale of supersymmetry breaking, but the gradual acceptance that string theory has a vast landscape of four-dimensional vacua has posed a major challenge to this idea: the more possibilities one has, the harder it is to make a unique prediction.

Nonetheless, there may be some simple rules that conclusively exclude particular low-energy actions. The WGC is one such rule, and as we see in this review it potentially constrains certain models of particle physics and cosmology and thus offers hope that quantum gravity may yet make decisive predictions for IR physics in the near future.[1]

The WGC has many interesting theoretical implications. In the context of AdS/CFT, it implies nontrivial statements for conformal field theories. In the context of string compactifications, it implies nontrivial statements about Calabi-Yau geometry. In the context of black hole physics, it is intimately related to the preservation of cosmic censorship. These connections, and others that we later review, suggest that the WGC is pointing us toward deep fundamental principles of quantum gravity.

However, despite recent progress we are still far from a concrete understanding of such principles, and some of the most basic questions about the WGC remain unanswered. First and foremost, we emphasize that the WGC is not really a single universally-agreed-upon conjecture, but rather a family of distinct but related "weak gravity conjectures," each of which attempts to formalize the idea that "any gauge force must be stronger than gravity" in a different way. The various conjectures have different consequences for particle physics, cosmology, etc. Some versions of the WGC have been discarded as counterexamples have been identified, while other versions have seen a growing body of evidence in their favor. Some of the most promising versions of the conjecture are known as the tower weak gravity conjecture and the sublattice weak gravity conjecture, and we elaborate on them here.

Moreover, thus far no nontrivial version of the conjecture has actually been proven in the sense of being derived from some accepted general principle. A number of promising routes toward a proof of some version of the WGC have been proposed in recent years, but these routes all suffer from at least one of the following two drawbacks: either they establish some statement that is qualitatively like the WGC, but without the correct O(1) factors included (i.e., "no gauge force can be much weaker than gravity"), or they argue for a precise version of the WGC but rely on additional unproven assumptions. In particular, in the original paper AMNV advocated for the WGC using black hole physics: the requirement that any nonsupersymmetric black hole should be able to decay necessitates some version of the WGC. It is not clear, however, why any nonsupersymmetric black hole must be able to decay, and it is also not clear that black hole decay is the fundamental principle underlying the WGC, as opposed to an accidental consequence of it. In particular, there is strong evidence for some versions of the conjecture (such as the sublattice WGC), with sharp consequences going beyond the minimal requirements of black hole instability. A proof of some form of the WGC (even a mild one) would represent a significant development in our understanding of the conjecture.

[1]The set of low-energy actions that cannot be realized in quantum gravity has been called the swampland (Vafa, 2005), and many more rules for ruling out such actions have been proposed. Some of these proposals are closely related to the WGC, while others are not. In this review we focus on the WGC specifically, so our discussion of other parts of the swampland program is subjective and incomplete. See Brennan, Carta, and Vafa (2018), Palti (2019), Graña and Herráez (2021), and van Beest *et al.* (2021) for broader discussions.

Without a proof of the conjecture or a deeper understanding of why the conjecture must be true, it is difficult to be sure which version or versions of the conjecture are correct, so it is difficult to determine how strong the constraints imposed by the WGC on particle physics, cosmology, geometry, etc., are. This means that despite the immense progress in our understanding of the WGC in recent years, the most important discoveries may lie ahead.

The remainder of this review is structured as follows. In Sec. II, we review arguments for the absence of global symmetries in quantum gravity, which may be viewed as a sort of precursor to the WGC. In Sec. III, we introduce the weak gravity conjecture in its mild and stronger variants. In Sec. IV, we outline evidence for different versions of the WGC, focusing on concrete examples in string theory and Kaluza-Klein theory. In Sec. V, we present qualitative arguments for approximate versions of the WGC, i.e., without precise O(1) factors included. In Sec. VI, we review the attempted derivations of the WGC, explaining why (in our opinion) each of them falls short of a "proof" of the WGC. In Sec. VII, we discuss broader implications of the WGC for phenomenology, mathematics, and other areas of theoretical physics. In Sec. VIII, we finish with conclusions and an outlook. In the Appendix we describe a general procedure for determing the black hole extremality bound (needed to correctly normalize the WGC bound) in theories with moduli.

II. NO GLOBAL SYMMETRIES

The WGC has its origins in an older conjecture that says that theories of quantum gravity admit no global symmetries of any kind. One motivation for this conjecture is the following. An evaporating black hole emits all particles in a theory without regard to their global charges (Hawking, 1975). This differs from gauge charge, where (at least for a continuous gauge group) the electric field outside of a charged black hole provides a chemical potential that favors discharge during evaporation. This insensitivity of black hole evaporation to global charges suggests that black holes can violate global symmetries and destroy global charge (Zeldovich, 1976, 1977).

A more precise argument (Banks and Seiberg, 2011) is that a continuous global symmetry would violate the Bekenstein-Hawking formula for black hole entropy. For example, suppose that we had a quantum gravity theory with a U(1) global symmetry. By colliding objects that are charged under this symmetry, one could produce large black holes of arbitrarily large global charge Q. The semiclassical calculation of Hawking evaporation implies that these black holes will decay, at least until they reach a radius $r_{\rm sc} \gg \ell_{\rm Pl}$ below which the effective field theory description is invalid. A black hole of initial charge Q will have a final charge $Q' \sim Q$: the Hawking evaporation process may emit charged particles, but it does not preferentially discharge the black hole. Thus, we can prepare black holes of size $r_{\rm sc}$ but an arbitrarily large charge. The information stored in this charge is arbitrarily large and, in particular, exceeds the Bekenstein-Hawking entropy $\pi r_{\rm sc}^2/G$. This argument, which is illustrated in Fig. 1, extends directly to any continuous global symmetry and implies a bound on

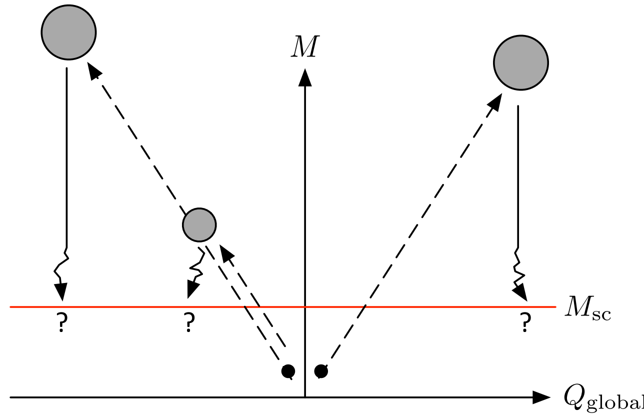

FIG. 1. Gravitational collapse of global-charged objects creates black holes of arbitrarily large global charge. If subsequently left alone, effective field theory dictates that the resulting black holes decay to objects of a size $r \sim r_{\rm sc}$ and a corresponding mass $M \sim M_{\rm sc}$ via Hawking radiation without appreciably changing the expected value of their global charge. This implies an infinite number of microstates for black holes of any fixed mass $M \gg M_{\rm sc}$, in violation of the Bekenstein-Hawking entropy formula. (Whether this process eventually results in stable remnants is immaterial.)

the size of a finite global symmetry group, albeit one that is exponentially weak in $r_{\rm sc}/\ell_{\rm Pl}$, which can be a large number in a weakly coupled theory (Banks and Seiberg, 2011).

Another, somewhat more vague, argument for global symmetry violation in quantum gravity is that if certain "Euclidean wormholes" are included in the gravitational path integral, then an apparent global symmetry violation is a consequence (Giddings and Strominger, 1988; Abbott and Wise, 1989; Coleman and Lee, 1990; Kallosh *et al.*, 1995); see also Hawking (1996) for an alternative view. The basic idea is that if there is a finite amplitude for adding a closed connected spatial component to the Universe, usually called a baby universe, then global charge can end up in such a baby universe, and therefore charge conservation can appear to be violated in the part of the Universe that we can actually access; see Fig. 2. This statement does not apply to gauge charge, as the gauge charge of a closed universe must be zero.[2]

Such general arguments about black hole physics or Euclidean gravity have been supplemented by observations about concrete theories of quantum gravity. In perturbative string theory, given a putative continuous global symmetry, one can create a vertex operator on the world sheet that creates

[2]This argument for the violation of global symmetries is similar to the semiclassical argument that black holes destroy quantum information, so it may be surprising that the modern consensus is that global symmetries are indeed violated but information is not lost. The difference is that the global charge of Hawking radiation is a "simple" observable, which is the kind the low-energy effective field theory needs to get right, while any extraction of information about the initial state of a black hole requires "complex" observables with the capability to invalidate the semiclassical picture. See Harlow and Shaghoulian (2021) for details on why global symmetries are not allowed in theories where black hole evaporation is unitary.

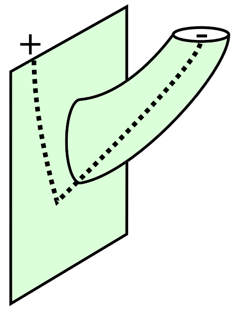

FIG. 2. Global symmetry violation by a Euclidean wormhole: a pair of charged particles is created from the vacuum, with the positive charge staying in the asymptotically flat region but the negative charge ending up in a baby universe. For someone living in the asymptotic region, this apparently violates the symmetry. Such a process cannot happen for a gauge symmetry, since the baby universe is closed and compact, so its gauge charge must be zero.

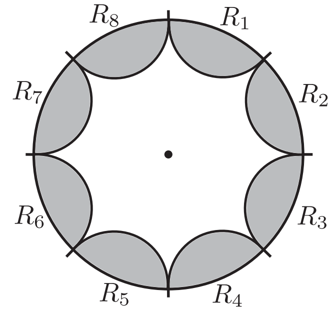

FIG. 3. An AdS/CFT contradiction between global symmetry and entanglement wedge reconstruction. The symmetry operators are products of operators supported in the regions $R_1, R_2, \ldots$, but no such operator can implement the symmetry on a charged operator in the center of the space.

a gauge field in spacetime coupling to the symmetry current, demonstrating that the would-be global symmetry is, in fact, gauged (Banks and Dixon, 1988). Similarly, in AdS/CFT a conserved current for a continuous global symmetry of the conformal field theory (CFT) implies the existence of a corresponding gauge symmetry in the bulk quantum gravity theory (Witten, 1998).

In the context of AdS/CFT, a holographic argument against global symmetries, both continuous and discrete, was presented by Harlow and Ooguri (2019, 2021). Here a symmetry generator U_g associated with a group element G acting on the boundary R is split into the following product:

$$U_g(R) = \prod_i U_g(R_i) U_{\mathrm{edge}}, \tag{2}$$

where $R = \cup_i R_i$, each $U_g(R_i)$ acts only in the region R_i, and U_{edge} acts at the boundaries of the R_i. A charged operator localized in the center of the bulk should transform under $U_g(R)$, but since the entanglement wedge of each R_i will not contain the center of the bulk for R_i sufficiently small, the charged operator cannot transform under the right-hand side of Eq. (2), which is a contradiction; see Fig. 3. We conclude that such a global symmetry cannot exist under the assumption that entanglement wedge reconstruction holds valid. This argument also applies to the higher-form global symmetries of Gaiotto *et al.* (2015), under which the charged objects are strings or branes instead of particles.

The use of AdS/CFT by Harlow and Ooguri (2019, 2021) is restrictive, but more recently it was observed by Harlow and Shaghoulian (2021) that essentially the same argument can be used to exclude global symmetries in any theory of quantum gravity where entanglement wedge reconstruction can be applied to an auxiliary reservoir coupled to an evaporating black hole. This assumption is the essential feature of recent calculations of the "Page curve" for an evaporating black hole, and thus is closely related to the unitarity of black hole evaporation (Almheiri *et al.*, 2019; Penington, 2020). Moreover, it was observed, following Lewkowycz and Maldacena (2013), that semiclassically this calculation can be interpreted as arising from the appearance of certain Euclidean wormholes in the gravitational path integral (Almheiri *et al.*, 2020; Penington *et al.*, 2022). Finally, Chen and Lin (2021) and Hsin, Iliesiu, and Yang (2021) showed that these Euclidean wormholes can indeed lead to concrete violations of global symmetry, thereby quantifying global symmetry violation in evaporating black hole backgrounds.

Finally, we remark that the absence of global symmetries in quantum gravity is closely related to another swampland conjecture, the *completeness hypothesis* (Polchinski, 2004). This hypothesis holds that in any gauge theory coupled to gravity there must be charged matter in every representation of the gauge group. The existence of such states is supported by black hole arguments (Banks and Seiberg, 2011) and holographic arguments in the context of AdS/CFT (Harlow, 2016; Harlow and Ooguri, 2019). In G gauge theory, if G is compact and connected or finite and Abelian, then the presence of charged matter in every representation is equivalent to the absence of a 1-form symmetry under which Wilson lines are charged. If G is compact but disconnected or finite and non-Abelian, then the presence of charged matter in every representation is equivalent to the absence of "noninvertible" global symmetries, which are associated with certain codimension-2 topological operators in the gauge theory (Rudelius and Shao, 2020; Heidenreich *et al.*, 2021b). This close connection between the absence of global symmetries and the completeness hypothesis means that arguments for one conjecture serve as indirect evidence for the other. An interesting quantitative approach to completeness based on algebraic ideas was developed by Casini *et al.* (2020, 2021) and Casini and Magan (2021), who used relative entropy and conditional expectations to diagnose to what extent field theories obey the completeness hypothesis.

The strongest arguments against the existence of global symmetries in quantum gravity are arguments against exact global symmetries. For applications, it is important to refine these arguments to ask to what extent approximate global symmetries are allowed. Recent general arguments along these lines were formulated by Daus *et al.* (2020), Fichet and Saraswat (2020), and Nomura (2020). As we later see, the weak gravity conjecture is one attempt to address this question: the weak coupling limit of a gauge theory has a global symmetry and should be forbidden in quantum gravity. As we discuss in Sec. V.B, the weak gravity conjecture is also related to the breaking of approximate 1-form global symmetries associated with the absence of charged particles.

III. WEAK GRAVITY CONJECTURES

We now consider quantum gravity theories coupled to a U(1) gauge field in $D > 3$ spacetime dimensions, with low-energy actions of the form

$$S = \int d^D x \sqrt{-g}\left(\frac{R}{2\kappa^2} - \frac{1}{4e^2(\phi)} F_{\mu\nu}F^{\mu\nu} + \cdots\right). \tag{3}$$

In Eq. (3) $e^2(\phi)$ is some function of the scalar fields ϕ^i in the theory and the omitted terms include kinetic terms for these scalars, as well as other possible terms involving additional matter fields and/or higher-derivative terms for the gauge field and the metric.[3] The compactness of the gauge group requires the charge to be quantized, and we normalize the gauge field so that the covariant derivative on a field of unit charge is $\partial_\mu - iA_\mu$. We then define the electric charge as

$$Q = \int_{S_\infty^{D-2}} \frac{1}{e^2(\phi)} \star F, \tag{4}$$

where S_∞^{D-2} is a sphere at spatial infinity, in which case the charge is quantized in integer units (i.e., the canonically normalized electrostatic potential is proportional to eQ).

The mildest version of the weak gravity conjecture then says the following.

Conjecture 1.—Mild weak gravity conjecture. Given any U(1) gauge field coupled to gravity as in Eq. (3), there must be an object of charge q and mass m satisfying

$$\frac{|q|}{m} \geq \frac{|Q|}{M}\bigg|_{\text{ext}}. \tag{5}$$

In Eq. (5) $|Q|/M|_{\text{ext}}$ indicates the charge-to-mass ratio of an extremal black hole of arbitrarily large size [in general there are finite-size corrections to this ratio that are not included in Eq. (5)]. We refer to any object obeying Eq. (5) as *super-extremal* in the review. It is convenient to parametrize the extremal charge-to-mass ratio as

$$\frac{e|Q|}{M}\bigg|_{\text{ext}} \equiv \gamma^{1/2}\kappa, \tag{6}$$

where $\kappa > 0$ is the gravitational coupling constant appearing in Eq. (3), related to the Planck mass M_{Pl} and the Newton constant G through

$$\kappa^2 = 8\pi G = \frac{1}{M_{\text{Pl}}^{D-2}}, \tag{7}$$

and $e^2 = e^2(\langle\phi\rangle)$ denotes the gauge coupling in the vacuum when written without an argument. The dimensionless parameter γ depends in general on the function $e^2(\phi)$ and on the metric on moduli space (see the Appendix). If $e^2(\phi)$ is independent of the moduli, then we simply have

$$\gamma = \frac{D-3}{D-2}. \tag{8}$$

Throughout this review, we have set $\hbar = c = 1$, but we emphasize that even if we restore them there are no factors of $\hbar$ in Eq. (6) since the extremality bound is a classical notion.

The original motivation for the conjecture is that it provides a kinematic condition that would allow an extremal black hole to shed its charge, which can happen even at zero Hawking temperature via Schwinger pair production (Gibbons, 1975; Johnson, 2020). However, there is no obvious pathology in a theory that admits infinitely many stable extremal black holes; due to the extremality bound, this would not lead to infinite entropy at finite mass as in the global-charge case in Fig. 1. Hence, this motivation falls far short of a proof or even a strong argument.

Although the mild weak gravity conjecture has an appealing simplicity, in practice it is too weak to imply anything interesting. The object that obeys Eq. (5) could be heavy, in which case it would have no substantive consequences for particle physics or cosmology. Moreover, it would not even be sufficient to allow "medium-sized" near-extremal black holes to decay, and thus would not address the original motivation for the conjecture. The mild weak gravity conjecture is nonetheless useful to consider, as it is a consequence of all of the various stronger versions of the WGC that have been proposed, which do have other more interesting implications, and thus an argument that shows that the mild WGC holds would hopefully also lead to an argument for one or more of the stronger versions. We now turn to discussing these possible generalizations.

A. WGC for *P*-form gauge fields

The mild WGC can be generalized in an obvious way from particles charged under an ordinary 1-form gauge field to $(P-1)$-branes charged under a P-form gauge field, with the restrictions $1 \leq P \leq D-3$. Instead of bounding the charge-to-mass ratio $|q|/m$ of such a particle, the WGC instead bounds the charge-to-tension ratio of the $(P-1)$-brane.

Conjecture 2.—Mild WGC for P-form gauge fields. Given a P-form gauge field coupled to gravity, there must exist a $(P-1)$-brane of charge Q and tension T_P satisfying

[3] For $D = 4$ if massless charged particles exist, then several aspects of this discussion need to be modified due to the logarithmic running that eventually drives the renormalized gauge coupling e to vanish in the deep infrared. The mild WGC still holds in such theories since, after all, there are massless charged particles, but to simplify our exposition we assume that for $D = 4$ all charged particles are massive.

$$\frac{|Q|}{T_P} \geq \frac{|Q|}{T_P}\bigg|_{\rm ext}. \tag{9}$$

In Eq. (9) $|Q|/T_P|_{\rm ext}$ is the charge to tension of an extremal black brane. It is useful to consider a concrete low-energy theory with an action

$$S = \int d^D x \sqrt{-g}\left(\frac{R}{2\kappa^2} - \frac{1}{4\kappa^2}(\nabla\phi)^2 - \frac{1}{2e_P^2}e^{-\alpha_P\phi}F_{P+1}^2\right). \tag{10}$$

In Eq. (10) $F_{P+1} = dA_P$ is the field strength for a P-form gauge field $A_{\mu_1,\ldots,\mu_P}$, with

$$F_q^2 := \frac{1}{q!}F_{\mu_1,\ldots,\mu_q}F^{\mu_1,\ldots,\mu_q}, \tag{11}$$

and by convention we shift ϕ to set $\langle\phi\rangle = 0$, so e_P is indeed the gauge coupling in the vacuum. In this theory we can write the extremal charge-to-tension ratio as

$$\frac{e_P|Q|}{T_P}\bigg|_{\rm ext} = \gamma_P^{1/2}\kappa, \tag{12}$$

with

$$\gamma_P = \frac{\alpha_P^2}{2} + \frac{P(D-P-2)}{D-2}. \tag{13}$$

If we replace $e_P^2 e^{\alpha_P\phi}$ with some more general function $e_P^2(\phi)$, then γ_P is modified as appropriate; see the Appendix. For future reference we write in one place the following super-extremality bound:

$$e_P^2 Q^2 \geq \gamma_P \kappa^2 T_P^2. \tag{14}$$

B. Magnetic WGC

The magnetic version of the mild WGC is simply the ordinary mild WGC applied to the electromagnetic dual gauge field. For the case of a P-form gauge field, this implies the existence of a superextremal magnetically charged $(D-P-3)$-brane with magnetic charge $|\tilde{Q}|$ and tension T_{D-P-2} satisfying

$$\frac{|\tilde{Q}|}{T_{D-P-2}} \geq \frac{|\tilde{Q}|}{T_{D-P-2}}\bigg|_{\rm ext}. \tag{15}$$

In four dimensions, for $p = 1$ this becomes a statement about the charge-to-mass ratio of a magnetic monopole. The monopole mass can be estimated in terms of the energy stored in its magnetic field. This energy is UV divergent, but if we cut it off at the semiclassical radius $r_{\rm sc} \sim 1/\Lambda_{\rm NP}$ associated with the "new physics" scale $\Lambda_{\rm NP}$ at which the low-energy effective field theory (EFT) breaks down, we obtain

$$m_{\rm mon} \gtrsim \frac{\Lambda_{\rm NP}}{e^2} \tag{16}$$

in the absence of a finely tuned cancellation between the field energy and the bare mass, where e is the electric gauge coupling.[4] By Dirac quantization, the magnetic gauge coupling is given by $\tilde{e} = 2\pi/e$, so the magnetic WGC bound (15) becomes

$$\Lambda_{\rm NP} \lesssim e M_{\rm Pl}. \tag{17}$$

In other words, the magnetic WGC places a cutoff on the new physics scale of the Abelian gauge theory, which vanishes (in Planck units) in the limit $e \to 0$. The magnetic WGC thus quantifies the extent to which effective field theory breaks down in the limit of weak gauge coupling. Without imposing the WGC itself, the conclusion (17) can also be obtained by requiring that the magnetic monopole is not a black hole, i.e., that its Schwarzschild radius is smaller than $r_{\rm sc}$ (Arkani-Hamed, Dubovsky et al., 2007; de la Fuente, Saraswat, and Sundrum, 2015).

We emphasize that the new physics scale $\Lambda_{\rm NP} \sim 1/r_{\rm sc}$ is not a cutoff on effective field theory altogether. The Abelian gauge theory may be embedded in another effective field theory with a higher cutoff, such as a Kaluza-Klein theory, a non-Abelian gauge theory, etc. In Sec. III.D we introduce several strong forms of the WGC, and in Sec. V we see that some of these strong forms provide a bound not only on $\Lambda_{\rm NP} \sim 1/r_{\rm sc}$ but also on the energy scale $\Lambda_{\rm QG}$ at which gravity becomes strongly coupled. The latter energy scale represents a cutoff on low-energy effective field theory in any form above which quantum gravity effects cannot be neglected.

Finally, note that a similar argument can be applied to $(D-P-3)$-branes magnetically charged under a P-form gauge field in D dimensions (Hebecker, Henkenjohann, and Witkowski, 2017b). The tension of such an object can be approximated as

$$T_{D-P-3} \sim \frac{\Lambda^P}{e_P^2}, \tag{18}$$

where $r_{\rm sc} = \Lambda^{-1}$ is again the semiclassical radius of the brane and e_P is the electric coupling constant. On the other hand, the tension of a black brane is given by

$$T_{\rm BB} \sim M_{\rm Pl}^{D-2} R_S^P, \tag{19}$$

where R_S is the Schwarzschild radius of the black brane. If we then demand that the magnetic brane is not itself a black hole, such that $\Lambda^{-1} = r_{\rm sc} \gtrsim R_S$, we then have

$$\Lambda \lesssim (e_P^2 M_{\rm Pl}^{D-2})^{1/2P}. \tag{20}$$

[4]This logic is not valid for electrically charged particles, because the self-energy should be cut off at the Compton radius, which is much larger than $\Lambda_{\rm NP}^{-1}$. In other words, the classical radius of an electric charge is less than its Compton wavelength, whereas the reverse is usually true for a magnetic charge, unless it is exceptionally light due to a finely tuned cancellation between the bare mass and the field energy.

Equation (20) reduces to Eq. (17) in the familiar case of $D = 4$ and $P = 1$.

C. The convex hull condition

Thus far we have focused on theories with a single gauge field. In general, however, a quantum gravity theory will have more than one gauge field, so the statement of the WGC must be generalized to this case. For simplicity, we focus on the case of particles charged under 1-form gauge fields, though analogous statements hold for branes charged under higher-form gauge fields.

In a theory of N Abelian gauge fields, the charge of a given particle may be represented by an N-vector $\vec{Q}$, where Q_i is the charge under the ith gauge field. The set of all possible charges $\vec{Q}$ consistent with charge quantization forms a lattice $\Gamma \simeq \mathbb{Z}^N \subset \mathbb{R}^N$. We define a *charge direction* $\hat{Q}$ as a unit vector in $\mathbb{R}^N$, and we say that such a charge direction is *rational* if $\lambda\hat{Q} \in \Gamma$ for some $\lambda \in \mathbb{R}$.

Finally, we define a *multiparticle state* as consisting of one or more actual particles in the theory with a mass m and charge $\vec{q}$ equal to the sums of the masses and charges of the constituent particles. This corresponds to a limit where the particles in question are taken infinitely far from each other, so they do not interact. A multiparticle state is superextremal if $\vec{z} \coloneqq \vec{q}/m$ has a length that is greater than or equal to the charge-to-mass ratio of an extremal black hole in the $\hat{Q}$ charge direction. The length of this vector is measured using the inverse of the kinetic matrix of the $U(1)$ gauge fields; i.e., given a Lagrangian $-(1/4)K_{ij}F^i_{\mu\nu}F^{j\mu\nu}$, the length of $\vec{z}$ is $(K^{ij}z_i z_j)^{1/2}$, where $K^{ij}K_{jk} = \delta^i{}_k$.

With this, we may define a mild WGC in such a theory as follows.

Conjecture 3.—Mild WGC for multiple gauge fields. For every rational direction $\hat{Q}$ in charge space, there is a superextremal multiparticle state with $\vec{z} \propto \hat{Q}$.

When there are a finite number of stable particles in the theory, this statement admits an equivalent, geometric formulation known as the convex hull condition (CHC) (Cheung and Remmen, 2014b). The CHC considers the set of all charge-to-mass vectors $\vec{z}_i \coloneqq \vec{q}_i/m_i$ for the particles in the theory, and it holds that the convex hull of this set should contain the region in $\vec{Z}$ space where black holes live. This condition is depicted graphically in Fig. 4. Note that in the absence of massless scalar fields the black hole region is simply the interior of an ellipsoid $K^{ij}z_i z_j \le \gamma\kappa^2$. If massless scalar fields are added to the theory, the black hole region will generically grow in size, and it may change its shape as well. Thus, the CHC gives stronger bounds in theories with massless scalar fields than those without.

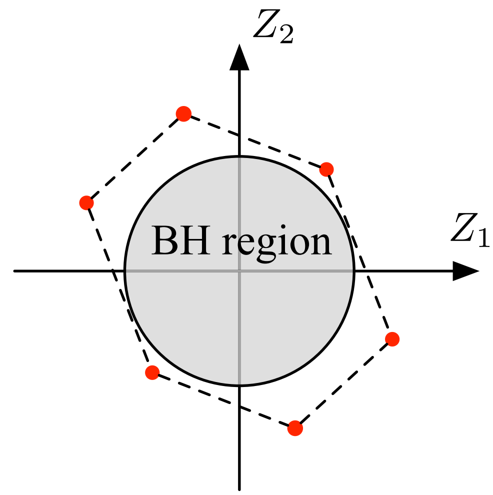

FIG. 4. Convex hull condition. In theories with multiple U(1)'s, the WGC is equivalent to the statement that the convex hull of the charge-to-mass vectors of the various particle species must contain the black hole region.

D. Strong forms of the WGC

Thus far all versions of the WGC that we have discussed are still mild in the sense of not having particularly interesting implications. From the first paper on the WGC, however, there has been interest in stronger versions of the WGC. This interest is not just wishful thinking: as we see in Sec. IV, all known examples in string theory seem to satisfy stronger statements than the mild WGC. Moreover, the heuristic arguments that we review in Sec. V also give support to the idea that something stronger than the mild WGC is true.

A first strong form to mention that is at times implicit in AMNV is the statement that the WGC should be satisfied by superextremal particles which are not themselves black holes. Higher-dimension operators in the action can modify the extremality bound of finite-sized black holes, as we discuss further in Sec. VI. If the charge-to-mass ratio of these finite-sized extremal black holes decreases as their mass is taken to infinity, the mild form of the WGC can be satisfied by stable, finite-sized black hole states. This scenario satisfies the letter of the WGC law but not the spirit of it, which holds that all black holes should be able to decay by emitting charged particles. This points to a first strong form of the WGC: the particles satisfying the WGC bound should not be black holes.

AMNV suggested two additional possible strong forms of the WGC: The first held that the lightest charged particle should be superextremal. The second held that the particle of smallest charge should be superextremal. Neither of these statements holds in general, however: they are violated, for instance, in certain T^n orbifold compactifications of type II and heterotic string theory (Heidenreich, Reece, and Rudelius, 2017).

However, a growing body of evidence points to another pair of strong forms (Montero, Shiu, and Soler, 2016; Heidenreich, Reece, and Rudelius, 2017, 2019; Andriolo *et al.*, 2018).

Conjecture 4.—Tower weak gravity conjecture. For every site in the charge lattice $\vec{q} \in \Gamma$ there is a positive integer n such that there is a superextremal particle of charge $n\vec{q}$.

Conjecture 5.—Sublattice weak gravity conjecture. There is a positive integer n such that for any site in the charge lattice $\vec{q} \in \Gamma$ there is a superextremal particle of charge $n\vec{q}$.

A few remarks about these conjectures are in order. First, note that the tower WGC implies that in any charge direction $\hat{q}$ there must be an infinite tower of superextremal particles.

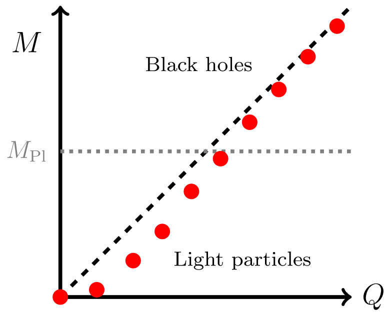

FIG. 5. Schematic illustration of WGC-satisfying particles (red dots) if the tower or sublattice WGCs hold. The black hole extremality bound is the dashed diagonal line. At small Q, the WGC is satisfied by light particles described by EFT. At large Q, black holes with small corrections obey the WGC; these asymptotically approach the extremality bound at large Q.

Indeed, the tower WGC is often defined by the latter statement. In Sec. IV, however, we see that consistency under dimensional reduction requires the formal definition that we have given here.

Second, note that the sublattice WGC is strictly stronger than the tower WGC: the sublattice WGC implies that the integer n appearing in the definition of the tower WGC can be chosen independently of $\vec{q}$. The sublattice WGC is equivalent to the statement that there is a full-dimensional sublattice of the charge lattice such that there is a superextremal particle at each site in the sublattice. The integer n is sometimes referred to as the *coarseness* of the sublattice. If $n = 1$, we say that the theory satisfies the lattice WGC. However, the lattice WGC is false in general; we exhibit a counterexample in Sec. IV.C.3.

Third, note that the tower WGC and the sublattice WGC require an infinite set of superextremal particles in each rational charge direction, whereas the ordinary WGC may be satisfied in a given charge direction by multiparticle states. We see in Sec. IV that the existence of superextremal particles, rather than merely multiparticle states, is required for consistency under dimensional reduction. For small charges, the necessary particles are ordinary quantum-mechanical particles represented by fields in the effective field theory. Far out on the charge lattice, the "particles" are actually black holes. The tower and sublattice WGCs thus interpolate between the effective quantum field theory regime and the gravitational regime of the quantum gravity theory in question. This is schematically illustrated in Fig. 5.

Fourth, note that it is possible (and, in fact, common in string theory examples) for the particles satisfying the tower or sublattice WGCs to be unstable resonances rather than stable states of the theory. Unstable resonances are not as easy to define as stable single-particle states, since they correspond not to states in the Hilbert space of the theory but rather to localized peaks in the S matrix of some scattering process. If the theory is weakly coupled, such a peak will be localized at a particular energy scale (the mass of the unstable particle) and the lifetime of this particle will be long. If the theory is strongly coupled, however, such a peak will be spread out across a range of energy scales, and it is not easy to define the mass of the resonance. Correspondingly, the tower WGC and sublattice WGC are not easy to define in this case.

Fifth, and finally, note that the tower or sublattice WGCs are modified in the presence of a few light charged particles in four dimensions due to the logarithmic running of the gauge coupling. Such charged particles appear near special loci in the moduli space where they become massless (for instance, where the Coulomb and Higgs branches of an $\mathcal{N} = 2$ theory intersect). For $D \geq 5$, this has a mild effect (generating finite threshold corrections), but in four dimensions the log running reduces the infrared gauge coupling gradually to zero as the massless locus is approached. A naive reading of the tower or sublattice WGCs would then suggest that an infinite tower of charged particles becomes light near the massless locus, but this does not always occur, particularly when the massless locus lies at finite distance in the moduli space.[5] While this seems to be a counterexample to the 4D tower or sublattice WGCs, as originally stated replacing the infrared gauge coupling in the WGC bound with its renormalized value resolves the problem (Heidenreich, Reece, and Rudelius, 2018b), suggesting that the conjectures are subtly modified rather than being invalidated in four dimensions. By contrast, this problem is absent in $D \geq 5$ and no modification seems to be needed there; see Alim, Heidenreich, and Rudelius (2021).[6]

In closing, we mention one other proposed "strong form" of the WGC: a superextremal state can saturate the WGC bound (i.e., be extremal) only if the theory is supersymmetric and the state in question is a Bogomol'nyi-Prasad-Sommerfield (BPS) state (Ooguri and Vafa, 2017). This conjecture is a mild extension of the ordinary WGC since there is no good reason why the mass of a superextremal particle should be tuned precisely to extremality unless the state is a BPS state in a supersymmetric theory. Nonetheless, this extension is interesting, as it suggests that extremal black holes can be marginally stable only if they are in a BPS state. When applied to the WGC for p-form gauge fields, the analogous statement further implies that any nonsupersymmetric anti–de Sitter (AdS) vacuum supported by fluxes must be unstable.

E. WGC for non-Abelian gauge fields

Thus far our definition of the WGC has dealt exclusively with particles charged under continuous Abelian gauge groups. We now discuss its extension to continuous non-Abelian gauge groups. For $D = 4$ this discussion is complicated by the fact that non-Abelian gauge fields are often confined, in which case the notion of a charged particle is not well defined, so this topic is of most interest for $D > 4$.

The mild form of the WGC extends in a rather trivial way: one simply decomposes the irreducible representations of the gauge group G into charges under the $\mathrm{U}(1)^{\mathrm{rk}(G)}$ Cartan and demands that the ordinary WGC should be satisfied with

[5]The absence of an infinite tower of light charged particles in such cases agrees with the emergence proposal (Grimm, Palti, and Valenzuela, 2018; Heidenreich, Reece, and Rudelius, 2018a, 2018b).

[6]The difference between the 4D and higher-dimensional cases can also be explained by noting that the tower or sublattice WGCs are related to the mild WGC in one lower dimension (see Sec. IV.A), whereas the mild WGC requires modification in three dimensions, if it continues to exist at all, due to the absence of asymptotically flat black holes.

respect to this Cartan subgroup. This requirement is automatically satisfied by the massless gluon fields of the theory. The sublattice WGC, on the other hand, is somewhat more subtle to define in the non-Abelian context. We use the following definition in this review (Heidenreich, Reece, and Rudelius, 2018b).

Conjecture 6.—Sublattice WGC for non-Abelian gauge fields. Given G gauge theory (with G a connected Lie group) coupled to quantum gravity, there is a finite-index Weyl-invariant sublattice Γ_0 of the weight lattice Γ_G such that for every dominant weight $\vec{Q}_R \in \Gamma_0$ there is a superextremal resonance transforming in the G irreducible representation R with highest weight $\vec{Q}_R$.

Conjecture 6 is stronger than simply requiring that the Abelian sublattice WGC should be satisfied with respect to the Cartan of G, as the latter can be satisfied by particles transforming under a sparse set of representations provided that they are sufficiently light. One argument for this stronger statement is that it is satisfied in perturbative string theory; this follows from the modular invariance argument discussed in Sec. IV.D. Conjecture 6 has also been shown to hold in certain 6D F-theory compactifications (Cota, Klemm, and Schimannek, 2021).

A natural question, now that we have defined the sublattice WGC for continuous Abelian and non-Abelian gauge groups, is whether there are further extensions for finite groups (or disconnected groups more generally). Thought experiments involving the evaporation of black holes carrying charge under finite gauge groups suggest bounds on UV cutoffs that are similar in spirit to WGC bounds (Craig, Garcia, and Koren, 2019a; Dvali and Redi, 2008; Dvali *et al.*, 2008; Dvali, 2010). WGC bounds can also be applied separately to the A and B fields associated with a massive gauge field in BF theory (Reece, 2019), which can lead to conclusions consistent with black hole thought experiments in $\mathbb{Z}_N$ gauge theory (Craig, Garcia, and Koren, 2019a). These considerations may hint at the existence of a formulation of the WGC encompassing all gauge groups.

F. WGC in asymptotically AdS spacetimes

Thus far we have focused on the WGC in flat Minkowski spacetimes. It is also worthwhile to define the conjecture in spacetimes with nontrivial curvature. Here, with an eye toward AdS/CFT, we restrict ourselves to possible definitions of the WGC in AdS spacetimes.

The flat-space definition (5) depends on the mass m of the particle, but in AdS_D with AdS radius R a more natural quantity is its rest energy Δ/R (in AdS/CFT Δ is the scaling dimension of the CFT operator that is dual to the field creating the particle). The relation between m and Δ depends on the dimensionality of spacetime and the spin of the particle; for a scalar field in AdS_D the relationship is

$$\Delta = \frac{D-1}{2} + \sqrt{\frac{(D-1)^2}{4} + R^2 m^2}. \tag{21}$$

A minimal requirement of any WGC bound in AdS_{d+1} is that it reduces to the flat-space bound in the limit where $R \to \infty$. One obvious proposal that does this was noted by Nakayama and Nomura (2015),

$$e^2 q^2 \geq \gamma \kappa^2 \frac{\Delta^2}{R^2}. \tag{22}$$

As in Eq. (5), $\gamma = (D-3)/(D-2)$ in the absence of massless scalar fields. Using the AdS/CFT correspondence, this bound can be recast in terms of data of the CFT_{D-1} as a bound on the charge q and dimension Δ of the operator $\mathcal{O}$ dual to the charged field. For $D = 5$, the CFT bound is (Nakayama and Nomura, 2015)

$$\frac{q^2}{b} \geq \frac{\Delta^2}{12c}, \tag{23}$$

where $c \sim \langle TT \rangle$ is the central charge of the CFT and $b \sim \langle JJ \rangle$ is the beta function coefficient of the conserved current associated with the gauge field in the bulk. On the other hand, there is no particular reason why Eq. (22) is more likely than some other expression with the same flat-space limit, so the proper formulation of the WGC in AdS remains an open problem.

The weak gravity conjecture in AdS/CFT is closely related to the recently formulated *Abelian convex charge conjecture* (Aharony and Palti, 2021). Given a CFT with a U(1) global symmetry, if we define $\Delta(n)$ to be the dimension of the lowest dimension operator of charge n, then this conjecture holds that

$$\Delta(n_1 q_0 + n_2 q_0) \geq \Delta(n_1 q_0) + \Delta(n_2 q_0) \tag{24}$$

for $q_0 \geq 1$ as an order-1 integer. A similar statement is conjectured to hold for non-Abelian gauge groups. Semiclassical tests of this statement were carried out by Antipin *et al.* (2021). If true, this conjecture implies that there must be a particle in the AdS bulk theory with non-negative self-binding energy, which is similar to the later-discussed repulsive force conjecture. Strong forms in which q_0 is 1 or is the charge of the lowest dimension charged operator were considered by Aharony and Palti (2021), but such statements (as currently formulated) are in tension with a flat-space example, as we discuss in Sec. IV.C.3.

When comparing the convex charge conjecture and various strong forms of the WGC, it is important to remember that not every CFT operator corresponds to a single-particle state in AdS. A convex spectrum of charged *single-trace* operators would have important implications for moduli stabilization. Consider a theory in which the gauge coupling $e(\phi)$ is a function of a stabilized modulus ϕ with mass m_ϕ and which has a separation of length scales $L \gg m_\phi^{-1} \gg r_{\mathrm{sc}}$, where L is the curvature radius of an AdS (or dS) vacuum and r_{sc} is the size of the smallest black hole we can treat as semiclassical. In this case, there are black hole solutions that can be approximated as flat-space black holes with a massless modulus ϕ when the black hole radius r obeys $m_\phi^{-1} \gg r \gg r_{\mathrm{sc}}$, and as flat-space black holes with no modulus when $L \gg r \gg m_\phi^{-1}$. Consequently, the black hole spectrum includes a range of extremal black holes that effectively have a modulus-dependent constant γ_ϕ in the extremality bound (6), and another range with the modulus-independent value γ_0 [Eq. (8)]. The modulus-dependent constant γ_ϕ is larger, as in Eq. (13), so the WGC becomes weaker in the infrared than in the UV. As a result, the minimum mass as a function of charge for any black hole spectrum that interpolates between these limits must fail to be convex, as illustrated in Fig. 6. On the other hand, at large $|Q|$ one could consider states consisting of multiple small black holes instead of a single large black hole, which could then have a lower mass following the "unstabilized" line. From the CFT viewpoint, these would correspond to multitrace, rather than single-trace, operators. A better understanding of the convex charge conjecture in CFTs and its relationship to large-N expansions,therefore, could potentially have important implications for the existence of vacua with stabilized moduli and scale separation.

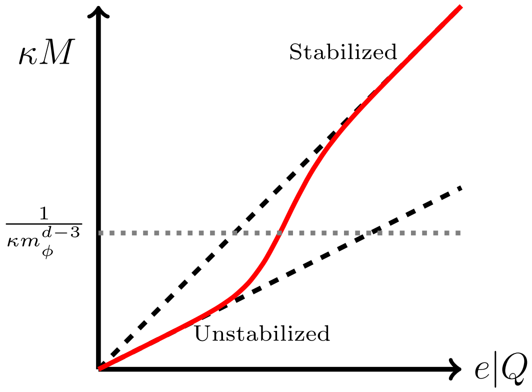

FIG. 6. Modulus stabilization and a nonconvex spectrum of charged black holes. In a theory where a modulus ϕ is stabilized with mass m_ϕ, the extremal black hole spectrum (red curve) should interpolate between small black holes that follow the unstabilized extremality bound (lower dashed black line) with slope $\gamma_\phi^{-1/2}$ and large black holes that follow the "stabilized" extremality bound (upper dashed black line) with larger slope $\gamma_0^{-1/2}$. The red curve indicates the smallest possible mass for a given charge. The detailed shape depends on the potential and couplings of ϕ, but any spectrum that interpolates between the two linear regimes must fail to be convex for some intermediate values of $|Q|$.

G. WGC for axions and axion strings

In Sec. III.A, we extended the WGC to the case of a P-form gauge field. An interesting case to consider is $P = 0$, in which the gauge field A_0 is a periodic scalar field ($A_0 \sim A_0 + 2\pi$), which is also known as an axion.

This case is somewhat degenerate, however, since the objects charged under this gauge field must be (-1)-branes, also known as instantons, with tension given by the instanton[7] action $T_0 \equiv S_{\rm inst}$. The instanton charge, also called the instanton number, is given in Euclidean signature by

$$n = i \int_{S^{d-1}} f^2 \star dA_0, \tag{25}$$

where $f \equiv 1/e_0$ is sometimes called the axion decay constant and S^{d-1} is a small sphere surrounding the instanton. In attempting to formulate an axion version of the WGC, however, we run into the problem that there is no immediately obvious notion of extremality. Indeed, naively plugging $P = 0$ into Eq. (13) (assuming the absence of massless scalar moduli), we see that γ_0 is zero, so the naive WGC bound (9) is trivial. Most likely this does not indicate the absence of any sort of axion WGC bound but rather that the O(1) coefficient γ_0 must be fixed by some other means. Absent a clear notion of extremality, the axion WGC bound is typically written simply as follows.

Conjecture 7.—Axion WGC. Given an axion (i.e., a periodic scalar) with an axion decay constant f coupled to quantum gravity, there must be an instanton of instanton number n satisfying

$$\frac{n}{f} \gtrsim S_{\rm inst}\kappa. \tag{26}$$

Note, in particular, that the sharp bound in the P-form WGC (9) has been replaced by $\gtrsim$ to account for the unknown O(1) coefficient γ_0.

There have, however, been proposals for what this O(1) coefficient should be. In the case of a 1-form, the WGC bound is the opposite of the black hole extremality bound, which sets the maximal charge-to-mass ratio of a macroscopic object in the low-energy theory (namely, a black hole). When it comes to instantons charged under an axion gauge field, there is once again a family of macroscopic solutions in the low-energy theory, known as gravitational instantons, that can ostensibly be used to fix γ_0 and define the extremality bound.

How exactly this should be done is not clear, however, and there are at least two proposals on the table. The confusion deals with the question of which class of gravitational instanton should be used to define the extremality bound, as there are three such classes:

(1) Solutions with a singular core, also known as cored solutions.
(2) Solutions with a flat metric (which we refer to as extremal solutions).
(3) Wormhole solutions, which have two different asymptotic regions connected by a smooth throat.

The metric for these solutions takes the form

$$ds^2 = \left(1 + \frac{C}{r^{2D-4}}\right)^{-1} dr^2 + r^2 d\Omega_{D-1}^2, \tag{27}$$

[7]A potential source of confusion here is that in general these instantons have nothing to do with the topologically nontrivial gauge field configurations introduced by Belavin *et al.* (1975), but they coincide for the particular case of the QCD axion in four dimensions. However, there can be axions without gauge fields and gauge fields without axions, and for $D \neq 4$ these two meanings of instanton do not even correspond to objects with the same dimensionality. The instantons discussed here are always zero-dimensional dynamical objects in the Euclidean path integral, with the property that their instanton number as defined by Eq. (25) is nonzero.

where $d\Omega_{D-1}^2$ is the metric on the unit $(D-1)$-sphere and C is positive, vanishing, and negative for cored, extremal, and wormhole solutions, respectively.

These solutions can all be obtained when we consider theories with a massless dilatonic modulus. Starting with the action (10) for $P = 0$, the action of the extremal instanton solution is given by

$$S_{\rm ext} = \frac{\sqrt{2}|n|}{\alpha f \kappa}, \tag{28}$$

where n is the instanton number. Meanwhile, the lower bound on the action of a cored solution is given by (Bergshoeff et al., 2004, 2005)

$$S_{\rm min} = \frac{\sqrt{2}|n|}{f\kappa} \times \begin{cases} \frac{1}{\alpha}, & \alpha \geq \tilde{\alpha}, \\ \frac{1}{\tilde{\alpha}}\sqrt{\frac{2\tilde{\alpha}}{\alpha} - 1}, & \alpha < \tilde{\alpha}, \end{cases} \tag{29}$$

where

$$\tilde{\alpha} \coloneqq \sqrt{\frac{2(D-2)}{D-1}}. \tag{30}$$

Finally, the instanton action for half of a wormhole solution is given by (Gutperle and Sabra, 2002)

$$S_{(1/2)\rm wh} = \frac{\sqrt{2}|n|}{\alpha f \kappa} \sin\left(\frac{\pi}{2}\frac{\alpha}{\tilde{\alpha}}\right), \tag{31}$$

where $\tilde{\alpha}$ is as previously defined.

After this review, we are now in a position to ask this question: What is the $O(1)$ coefficient for the axion WGC bound in this theory? It is natural to suppose that the extremal instanton should set the axion WGC bound, just as the extremal black hole sets the ordinary WGC bound. From the instanton action (28), this gives the following bound:

$$\frac{|n|}{fS} \geq \frac{|n|}{fS_{\rm ext}} = \frac{\alpha\kappa}{\sqrt{2}}, \qquad \alpha > \tilde{\alpha}. \tag{32}$$

Equation (32) is a plausible candidate for the axion WGC when $\alpha \geq \tilde{\alpha}$. Equation (28) suggests that cored instantons have a larger action than the extremal instanton of the same instanton number, just as subextremal black holes have a larger mass than an extremal black hole of the same charge.

For $\alpha < \tilde{\alpha}$, however, things become more complicated. Cored instantons now have a smaller action than the extremal solution. Thus, the axion WGC bound should perhaps be given by the cored instanton of smallest action, which means that

$$\frac{|n|}{fS} \geq \frac{|n|}{fS_{\rm min}} = \frac{\tilde{\alpha}\kappa}{\sqrt{2}}\frac{1}{\sqrt{2\tilde{\alpha}/\alpha - 1}}, \qquad \alpha < \tilde{\alpha}. \tag{33}$$

However, the half-wormhole solution has an even smaller action than the cored and extremal instanton solutions. If the WGC bound is to be set by the macroscopic object of smallest action, then perhaps the axion WGC bound should be set by the half-wormhole solution so that

$$\frac{|n|}{fS} \geq \frac{|n|}{fS_{\rm min}} = \frac{\alpha\kappa}{\sqrt{2}}\frac{1}{\sin\left[(\pi/2)\alpha/\tilde{\alpha}\right]}. \tag{34}$$

Note that the right-hand side of Eq. (34) remains finite in the $\alpha \to 0$ limit.

It is not clear which of these bounds should be viewed as the "correct" version of the axion WGC. Heidenreich, Reece, and Rudelius (2016) proposed the bounds (32) and (33), whereas Hebecker et al. (2017) and Hebecker, Mikhail, and Soler (2018) suggested the bound (34). One difference in viewpoint is that Heidenreich, Reece, and Rudelius (2016) assumed that only true instanton solutions, not wormholes, can contribute to an axion potential because a wormhole is effectively an instanton–anti-instanton pair with no net charge. Hebecker et al. (2017) and Hebecker, Mikhail, and Soler (2018) argued that, because the instanton and anti-instanton ends of the wormhole can be distant from each other in Euclidean time, they do in fact generate an axion potential. The latter perspective has a close affinity with the heuristic argument that wormholes violate the global symmetries discussed in Sec. II.

Just as the precise statement of the axion WGC is somewhat difficult to define, so too is its magnetic version. Naively we want to say that there must be a $(D-3)$-brane whose charge-to-tension ratio is greater than or equal to that of a large, extremal black $(D-3)$-brane (such as a string in $D = 4$). Such objects do not exist in asymptotically flat spacetime, as we further discuss shortly. Hence, rather than assuming an inequality with an exact coefficient determined by an extremality bound, it is natural to suppose that the WGC should imply the existence of some charged $(D-3)$-brane (i.e., a vortex) of charge $\tilde{Q}$ and tension T_{D-2}, satisfying

$$\frac{e_{D-2}|\tilde{Q}|}{T_{D-2}} \gtrsim \kappa, \tag{35}$$

where $e_{D-2} = 2\pi f$ is the magnetic coupling. The inequality again has a $\gtrsim$ sign, and there is an $O(1)$ coefficient that remains to be fixed. Specializing to $D = 4$ for convenience, an argument for this has been given in terms of axionic black holes, i.e., those with a nonvanishing integral $\int_\Sigma B$ of the axion's dual B field over the horizon (Bowick et al., 1988). It has been argued that axionic strings obeying Eq. (35) are needed to allow this axionic charge to change and avoid a remnant problem in black hole evaporation (Hebecker and Soler, 2017; Montero, Uranga, and Valenzuela, 2017).

Because the magnetically charged object in this case has codimension 2 (e.g., a string in $D = 4$ or a 7-brane in $D = 10$), the classical tension stored in the winding axion field is logarithmically divergent in both the IR and the UV, whereas our discussion of the magnetic WGC in Sec. III.B incorporated only a UV divergence. Consequently Eq. (20) is not valid in the case of $P = 0$. Revisiting the logic by estimating the classical self-energy with UV and IR cutoffs and requiring it to satisfy the magnetic axion WGC bound (35), we have

$$T \sim f^2 \log\frac{\Lambda_{\rm UV}}{\Lambda_{\rm IR}} \lesssim \frac{f}{\kappa} \tag{36}$$

or, in other words,

$$\frac{\Lambda_{\rm UV}}{\Lambda_{\rm IR}} \lesssim \exp\frac{O(1)}{\kappa f}. \tag{37}$$

This is compatible with the idea that instantons will generate an IR scale $\Lambda_{\rm IR} \sim e^{-S}\Lambda_{\rm UV}$ together with the electric axion WGC (26), which implies $S \lesssim 1/\kappa f$. Indeed, once an axion potential is generated through instantons, the axion vortex becomes the boundary of a domain wall, such that the winding of the axion field is localized inside the wall and there is no significant energy density outside the wall. When an axion vortex is attached to a semi-infinite domain wall, we would view the energy outside the axion vortex core as reflecting the finite domain wall tension rather than an infinite correction to the axion vortex tension. In this way, domain walls naturally provide an IR cutoff to the estimate of the axion vortex tension, and there is a relationship between the magnetic and electric WGC that has the same spirit, although more complicated details, as in the cases $1 \le P \le D-3$.

The previous estimate neglects gravitational backreaction, which is significant for objects of low codimension. In particular, static vortices in gravitational theories produce a deficit angle. Implications of gravitational backreaction on axion strings (in $D=4$) for the magnetic axion WGC were considered by Dolan *et al.* (2017) and Hebecker, Henkenjohann, and Witkowski (2017b). The static axion string solution in general relativity (in the case with zero axion potential so that the strings are not confined by domain walls) was first found by Cohen and Kaplan (1988). The IR and UV divergences of the string without gravity are reflected in singularities of this solution. When $f < \sqrt{2}M_{\rm Pl}$, the IR singularity lies exponentially far away in Planck units from the core of the string and the deficit angle is positive. Hence, one could consider, for example, large loops of closed string, which would be well behaved in the IR and potentially completed by UV physics in the string core. When $f > \sqrt{2}M_{\rm Pl}$, the deficit angle becomes negative, the singularity is inside the core of the string, and there is no longer a sensible interpretation of stringlike objects in approximately asymptotically flat spacetime with sensible UV completions in the string core. This suggests $f < \sqrt{2}M_{\rm Pl}$ as a possible consistency condition on 2-form gauge theory in four dimensions.

Although the physics of static axion strings is relatively straightforward, one could consider whether the magnetic axion WGC could be satisfied by time-dependent rather than static objects (Dolan *et al.*, 2017; Hebecker, Henkenjohann, and Witkowski, 2017b). Nonsingular, time-dependent string solutions were written for a complex scalar Φ with a U(1) global symmetry by Gregory (1996) that features a Lagrangian of the form

$$\mathcal{L} = -\frac{1}{2}|\partial\Phi|^2 - \frac{\lambda}{4}(|\Phi|^2 - f^2)^2, \tag{38}$$

such that the phase of Φ is an axion with decay constant f in the low-energy theory. For f smaller than some critical $f_{\rm crit}$, there are nonsingular axion string spacetimes that inflate along the string direction but that have a static field configuration along slices orthogonal to the string. For $f > f_{\rm crit}$, the field Φ itself becomes time dependent and the theory undergoes "topological inflation." This occurs when the core region of a topological defect of size $R_{\rm core}$ has a potential energy density $V_{\rm core}$ sufficiently large to sustain a Hubble expansion rate $H \sim \sqrt{V_{\rm core}}/M_{\rm Pl}$ with $HR_{\rm core} \gtrsim 1$ (Linde, 1994; Vilenkin, 1994). The numerical analysis of Cho (1998) found $f_{\rm crit} = 1.63 M_{\rm Pl}$ as the critical value for the onset of topological inflation. Dolan *et al.* (2017) pointed out that the computation with $f \gtrsim M_{\rm Pl}$ is not necessarily under control, but a scenario with axion strings of winding number $n \gg 1$ such that $f \ll M_{\rm Pl} \ll nf$ provides a controlled setting with similar conclusions. Numerical studies by Dolan *et al.* (2017) confirmed exponential expansion in this scenario. They also demonstrated power-law expansion in a different model in which the axion was the holonomy of a higher-dimensional gauge field. In this case, the radial mode associated with the axion is the radion modulus R of an extra dimension. The string core sees a decompactification limit $R \to \infty$ that lies at infinite distance in field space where $V(R) \to 0$ (hence no exponential expansion). In both the $|\Phi|^4$ case and the radion case, there is no obvious pathology associated with the time-dependent infinite, straight string configurations. However, Dolan *et al.* (2017) argued that the topological inflation of a closed loop of an axion string with $nf \gtrsim M_{\rm Pl}$ would violate the "topological censorship theorem" of Friedman, Schleich, and Witt (1993), suggesting that it would always collapse into a black hole. It would then be impossible for an observer to traverse a loop linking with a closed axion string to measure the field excursion. The impossibility of such a scenario is one candidate for a magnetic axion WGC.

H. Repulsive force conjecture

The WGC was inspired by the idea that gravity should be weaker than any gauge force. The previously descibed definition, however, deals not with the relative strength of gravity against other forces but rather with the notion of superextremal particles. These two notions agree if the only forces are gravity and electromagnetism: a particle is superextremal if and only if the long-range electromagnetic repulsion between a pair of such particles is stronger than their gravitational attraction. In theories with massless scalar fields, however, this correspondence breaks down, and the question of whether a particle is superextremal is distinct from the question of whether or not a pair of such particles will repel each other at long distances. With this in mind, we thus define a particle to be self-repulsive if a pair of such particles repel one another at long distances, and we define the repulsive force conjecture (RFC) as follows.

Conjecture 8.—RFC. In any theory of a single Abelian gauge field coupled to gravity, there is a self-repulsive charged particle (Palti, 2017).

After being emphasized by Palti (2017), Conjecture 8 was further studied by Lee, Lerche, and Weigand (2019a) and Lüst and Palti (2018). This statement can be easily generalized from particles charged under 1-form gauge fields to $(P-1)$-branes charged under P-form gauge fields. The generalization to theories with more than one gauge field is somewhat subtle;

see Heidenreich, Reece, and Rudelius (2019) for a further explanation.

While the RFC and the WGC are distinct conjectures in the presence of massless scalar fields, close connections remain; for instance, at the two-derivative level extremal black holes have a vanishing long-range self-force (Heidenreich, 2020), and the same towers of charged particles typically satisfy the tower or sublattice versions of both conjectures (Heidenreich, Reece, and Rudelius, 2019; Heidenreich and Lotito, 2022).

The idea of gravity as the weakest force has also led to several variations on a scalar weak gravity conjecture, proposing that light scalars should always mediate forces stronger than gravity for some particles (Li *et al.*, 2007; Palti, 2017; Lüst and Palti, 2018; Gonzalo and Ibáñez, 2019). Such conjectures can lead to interesting consequences, including for phenomenology and cosmology. However, because they do not involve gauge fields and have no connection to black hole extremality, we do not discuss them further here. Similarly, we do not discuss weak gravity statements related to higher-spin particles, for which there are sharp bounds from causality (Kaplan and Kundu, 2021).

Here we focus primarily on the WGC and the notion of superextremality, but much of our analysis applies equally well to the RFC and the notion of self-repulsiveness. We stress once again that these conjectures are equivalent (and the notions of superextremality and self-repulsiveness are equivalent) in the absence of scalar fields.

IV. EVIDENCE FOR THE WGC

The WGC was originally inspired by the idea that nonsupersymmetric extremal black holes should be able to decay. As we have discussed, this motivation is not compelling, since there is no obvious reason why stable extremal black holes present a problem for a theory. Nonetheless, this motivation seems to have gotten people to start digging in the right place since there have been a number of lines of evidence to date that support the WGC and its variants. In this section, we focus on four such lines: an argument from dimensional reduction, examples in string theory, a general argument from modular invariance in perturbative string theory, and the relation between the WGC and the swampland distance conjecture (Ooguri and Vafa, 2007).

A. Dimensional reduction

One approach to assessing the validity of the WGC is to examine its internal consistency under dimensional reduction (Heidenreich, Reece, and Rudelius, 2016); similar checks under T duality were carried out by Brown *et al.* (2015). Our starting point is the Einstein-Maxwell-dilaton action (10) for a P-form gauge field $A_{\mu_1,\ldots,\mu_P}$ in $D = d+1$ dimensions. We could in principle include additional terms in the low-energy action, such as Chern-Simons terms, but for our purposes the previous action will suffice.

1. Preservation of the p-form WGC bound

We consider a dimensional reduction ansatz of the form

$$ds^2 = e^{\lambda(x)/(d-2)} d\hat{s}^2(x) + e^{-\lambda(x)} dy^2, \tag{39}$$

where $y \sim y + 2\pi R$. For now, we do not include a Kaluza-Klein photon in our dimensional reduction ansatz, but we do so later in this section. The coefficients of $\lambda(x)$ in the exponentials have been carefully chosen so that the dimensionally reduced action is in the Einstein frame; i.e., there is no kinetic mixing between λ and the d-dimensional metric,

$$\frac{1}{2\kappa_D^2}\int d^D x\sqrt{-g}\mathcal{R}_D \to \frac{1}{2\kappa_d^2}\int d^d x\sqrt{-\hat{g}}\mathcal{R}_d - \frac{1}{2}\int d^d x\sqrt{-\hat{g}}G_{\lambda\lambda}(\nabla\lambda)^2, \tag{40}$$

where

$$\frac{1}{\kappa_d^2} = M_d^{d-2} = (2\pi R) M_D^{D-2}, \tag{41}$$

$$G_{\lambda\lambda}^{(d)} = \frac{d-1}{4\kappa_d^2(d-2)} = M_d^{d-2}\frac{d-1}{4(d-2)}. \tag{42}$$

Upon dimensional reduction, the P-form gauge field in D dimensions gives rise to both a P-form gauge field and a $p = (P-1)$-form gauge field in d dimensions, obtained, respectively, by taking all of the legs of the gauge field to lie in noncompact directions, or by taking one leg to wrap the compact S^1 direction. The gauge couplings of the two gauge fields are given, respectively, by

$$e_{P;d}^2 = \frac{1}{2\pi R} e_{P;D}^2, \qquad e_{p;d}^2 = (2\pi R) e_{P;D}^2. \tag{43}$$

Similarly, a charged $(P-1)$-brane in D dimensions reduces to both a $(P-1)$-brane and a $(p-1)$-brane, obtained, respectively, by taking the brane to lie exclusively in noncompact dimensions or by taking the brane to wrap the compact direction. The tensions of these branes are given, respectively, by

$$T_P^{(d)} = T_P^{(D)}, \qquad T_p^{(d)} = (2\pi R) T_P^{(D)}. \tag{44}$$

Recall from Eqs. (14) and (13) that the WGC bound is modified by the exponential coupling of the radion to the Maxwell term. The Maxwell term of the P-form gauge field couples to a linear combination of both ϕ and the radion λ, and it is useful to rewrite these scalar fields in terms of two canonically normalized fields σ and ρ, the former of which decouples from the Maxwell term and the latter of which couples to it as $e^{-\alpha_{P;d}\rho} F_{P+1}^2$. The coefficient $\alpha_{P;d}$ is then given by

$$\alpha_{P;d}^2 = \alpha_{P;D}^2 + \frac{2P^2}{(d-1)(d-2)}. \tag{45}$$

Equation (45) can be rewritten as

$$\frac{\alpha_{P;d}^2}{2} + \frac{P(d-P-2)}{d-2} = \frac{\alpha_{P;D}^2}{2} + \frac{P(D-P-2)}{D-2}, \quad (46)$$

which by Eq. (13) implies

$$\gamma_{P;d}(\alpha_{P;d}) = \gamma_{P;D}(\alpha_{P;D}), \quad (47)$$

from which we conclude that the $(P-1)$-brane satisfies the P-form WGC bound (14) in D dimensions if and only if it satisfies the P-form WGC bound in d dimensions: in other words, the WGC is exactly preserved under dimensional reduction.

A similar story applies to the case of the wrapped brane: a particular linear combination of ϕ and λ couples to F_{p+1}^2, which ultimately leads to the coefficient

$$\alpha_{p;d}^2 = \alpha_{P;D}^2 + \frac{2(d-p-2)^2}{(d-1)(d-2)}. \quad (48)$$

Equation (48) can be rewritten as

$$\frac{\alpha_{p;d}^2}{2} + \frac{p(d-p-2)}{d-2} = \frac{\alpha_{P;D}^2}{2} + \frac{P(D-P-2)}{D-2}, \quad (49)$$

which by Eq. (13) implies

$$\gamma_{p;d}(\alpha_{p;d}) = \gamma_{P;D}(\alpha_{P;D}), \quad (50)$$

so again the WGC bound is exactly preserved. The $(P-1)$-brane satisfies the P-form WGC bound in D dimensions if and only if it satisfies the p-form WGC bound in d dimensions after wrapping on S^1.

2. Kaluza-Klein modes and a violation of the CHC

We now add a Kaluza-Klein photon to our dimensional reduction ansatz,

$$ds^2 = e^{\lambda(x)/(d-2)} d\hat{s}^2(x) + e^{-\lambda(x)}(dy + RB_1)^2, \quad (51)$$

where $y \cong y + 2\pi R$ and B_1 is normalized so that the Kaluza-Klein (KK) modes carry integral charges. The dimensionally reduced action is then given by

$$S = \int d^d x \frac{\sqrt{-\hat{g}}}{2\kappa_d^2} \left[\hat{\mathcal{R}}_d - \frac{d-1}{4(d-2)} (\nabla\lambda)^2 - \frac{R^2}{2} e^{[-(d-2)/(d-2)]\lambda} H_2^2 \right], \quad (52)$$

where $H_2 = dB_1$. From Eq. (52), we can read off the KK photon gauge coupling and the radion-KK photon coupling such that

$$\frac{1}{e_{\mathrm{KK}}^2} = \frac{1}{2} R^2 M_d^{d-2}, \qquad \alpha_{\mathrm{KK}} = \sqrt{\frac{2(d-1)}{d-2}}. \quad (53)$$

In Eq. (53) α_{KK} is defined by the coupling to the normalized radion $\hat{\lambda} = \sqrt{[(d-1)/2(d-2)]}\lambda$.

The WGC bound for a particle with n units of KK charge is then given by Eq. (14) as

$$\left[\frac{\alpha_{\mathrm{KK}}^2}{2} + \frac{d-3}{d-2} \right] m^2 \le e_{\mathrm{KK}}^2 n^2 M_d^{d-2}. \quad (54)$$

This means that $\gamma_{\mathrm{KK}} = 2$, and the WGC bound is simply

$$m^2 \le \frac{n^2}{R^2}. \quad (55)$$

This can be compared to the spectrum of KK modes for a particle of mass m_0 in the parent theory,

$$m^2 = m_0^2 + \frac{n^2}{R^2}, \qquad n \in \mathbb{Z}, \quad (56)$$

where the KK charge n specifies the momentum n/R of the particle along the compact circle. We see, therefore, that KK modes of massless particles saturate the WGC bound, whereas KK modes of uncharged massive particles violate the WGC bound. The D-dimensional parent theory necessarily has at least one massless particle (namely, the graviton), so the dimensionally reduced theory necessarily has superextremal particles charged solely under the KK photon. Indeed, each of the KK modes of the graviton is superextremal, so there is actually an infinite tower of superextremal KK modes, as required by the tower or sublattice WGC.

What happens, however, if we include a $U(1)$ in the parent theory in D dimensions? The resulting d-dimensional theory will now have two $U(1)$ gauge fields, and the WGC is equivalent to the CHC introduced in Sec. III.C.

In the parent theory, a particle of charge q and mass m is superextremal when the dimensionless charge-to-mass ratio $Z_D \coloneqq (q/m) e_D \gamma_D^{-1/2} M_D^{(D-2)/2}$ has a magnitude $|Z_D| \ge 1$. Likewise, in the dimensionally reduced theory a particle of charge (q, q_{KK}) and mass m is superextremal when the dimensionless charge-to-mass ratio vector

$$\vec{Z}_d \coloneqq \left(\frac{q}{m} e_d \gamma_d^{-1/2} M_d^{(d-2)/2}, \frac{q_{\mathrm{KK}} - q\theta/2\pi}{mR} \right) \quad (57)$$

has a length $|\vec{Z}_d| \ge 1$, where $\theta = \oint A$ is the vacuum expectation value (VEV) of the axion descending from the gauge field and $\gamma_d = \gamma_D$, thereby accounting for the radion coupling as previously discussed.

The nth KK mode of a particle of charge q and mass m_0 in the parent theory has a mass $m^2 = m_0^2 + (1/R^2)(n - q\theta/2\pi)^2$, and hence the charge-to-mass vector

$$\vec{Z}_d^{(n)} = \frac{(\mu Z_D, x_n)}{\sqrt{\mu^2 + x_n^2}}, \qquad \mu = m_0 R, \qquad x_n = n - \frac{q\theta}{2\pi}. \quad (58)$$

The charge-to-mass vectors of the KK modes, along with the convex hull that they generate, are plotted in Fig. 7 (top panel). The vectors lie on the ellipsoid $Z_{d1}^2/Z_D^2 + Z_{d2}^2 = 1$, which lies outside the unit disk provided that $|Z_D| \ge 1$, so each KK mode of a particle that was superextremal in the parent theory is superextremal.

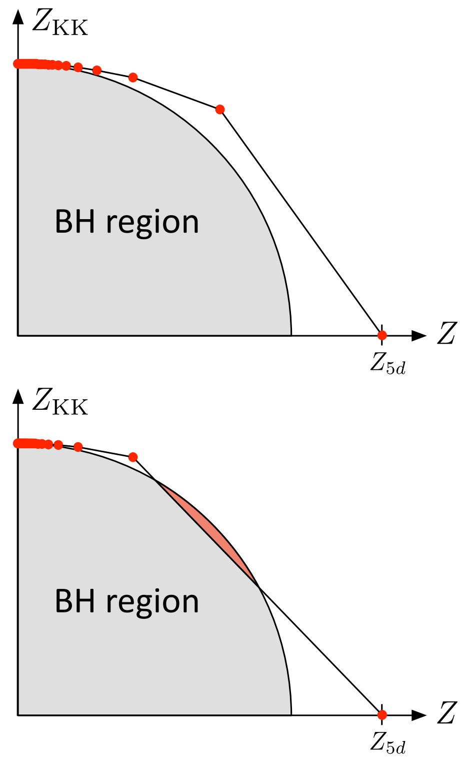

FIG. 7. Top panel: the Kaluza-Klein modes of a superextremal particle with charge q_F in $d+1$ dimensions are superextremal after reduction on S^1, as their charge-to-mass vectors $\vec{Z}$ lie outside the elliptical black hole region. Bottom panel: if the S^1 is sufficiently small, the convex hull condition is violated and the Kaluza-Klein modes of the superextremal particle in question do not satisfy the WGC.

However, the fact that each individual KK mode is superextremal does not ensure that the convex hull condition is satisfied. As shown in Fig. 7 (bottom panel), as we take the limit $R \to 0$ the KK modes of the particle are pushed closer and closer to the point $(0,1)$. Below some critical value of R, the convex hull condition is violated. In fact, if $Z_D = 1$, thus saturating the WGC bound, then the convex hull condition will be violated for any value of R. Starting with a theory that satisfies the WGC in D dimensions, we have arrived at a theory that violates the WGC in d dimensions.

It is important to realize that this does not represent a counterexample to the WGC, because there is no good reason to think that the D-dimensional theory that we started with is in the landscape as opposed to the swampland. Rather, we showed that the WGC in D dimensions alone is not sufficient to ensure that the WGC holds in d dimensions. If we want the WGC to hold in d dimensions, we need to impose a stronger constraint than the WGC in D dimensions.

To identify such a constraint, note that a violation of the convex hull condition for sufficiently small R will arise whenever the number of superextremal particles in D dimensions is finite. To satisfy the WGC in d dimensions for all R, therefore, requires an infinite number of superextremal particles in D dimensions. Indeed, it is not difficult to see that the tower WGC, as defined in Sec. III.D, is a sufficient condition for ensuring that the WGC is satisfied in the dimensionally reduced theory. Indeed, this observation is what originally inspired the tower or sublattice WGC. There are at present no known counterexamples to either of these conjectures in string theory.

One can further verify that the tower WGC is satisfied in d dimensions provided that it is satisfied in D dimensions (and likewise for the sublattice WGC), so the tower WGC and sublattice WGC are automatically preserved under dimensional reduction, unlike the mild WGC. The general idea that a proposed consistency criterion in quantum gravity should apply not only to a single vacuum but to all of its compactifications, whose application to the WGC discussed here originated in the work of Heidenreich, Reece, and Rudelius (2016), was later named the total landscaping principle by Aalsma et al. (2021) and has been successfully applied in several contexts (Montero, Uranga, and Valenzuela, 2017; Heidenreich, Reece, and Rudelius, 2019; Aalsma et al., 2021; Cremonini et al., 2021; Rudelius, 2021).

B. Higgsing

We have just seen that the mild WGC is not automatically preserved under compactification: starting with a theory that satisfies the WGC, we can produce a theory that violates the WGC by Kaluza-Klein reduction on a circle. This points toward a stronger version of the WGC, such as the tower or sublattice WGCs, which are automatically preserved.

In this section, we see that a similar issue arises due to the process of Higgsing. Starting with a theory that satisfies certain forms of the WGC, we can produce a theory that violates these forms of the conjecture by Higgsing. However, other versions of the WGC will be preserved. In particular, we show, following Saraswat (2017), that the mild WGC, tower WGC, and sublattice WGC are preserved (barring a special fine-tuning that we do not expect to occur). In contrast, the statements that the lightest charged particle should be superextremal or that a particle of smallest charge should be superextremal are not preserved under Higgsing.

Consider a theory with two U(1) gauge fields A and B. For simplicity, we assume that their gauge couplings are identical ($g_A = g_B = g$) and that we are working in four dimensions. Suppose that there are two superextremal particles with masses with $m_1 < m_2$ and charges $q_1 = (1,0)$ and $q_2 = (0,1)$, respectively. Let these be the lightest charged particles in the theory.

Next suppose that there is a scalar field of charge $(n,1)$ that acquires a VEV v. Under this process, the gauge boson $H = nA + B$ acquires a mass $m_H = gv\sqrt{n^2+1}$ and the gauge boson $L = (A - nB)/(n^2+1)$ remains massless, with gauge coupling $g_{\rm eff} = g/\sqrt{n^2+1}$. After Higgsing, the particle of charge $q_1 = (1,0)$ has quantized charge 1 under the massless gauge field L. It still has mass m_1, so for n sufficiently large we find that this particle is no longer superextremal after Higgsing, since $g_{\rm eff} \simeq g/n \ll m_1/M_{\rm Pl}$.

On the other hand, the particle of charge $q_2 = (0,1)$ has quantized charge $-n$, so it remains superextremal after Higgsing. Thus, the mild version of the WGC remains satisfied in this theory. However, since we assumed that $m_1 < m_2$, the lightest charged particle is no longer superextremal and there is no longer a superextremal particle of charge 1. We see that the strong forms of the WGC that demand that either the lightest charged particle or the particle of smallest charge should be superextremal are not automatically preserved under Higgsing: they are violated here in the Higgsed theory even though they were satisfied in the un-Higgsed theory.

As with the previous reasoning that led to the tower or sublattice WGCs, one might be tempted to search for stronger conjectures that ensure that the lightest charge particle and/or particle of smallest charge automatically remain superextremal, even after Higgsing. However, in Sec. IV.C, we see an explicit example in string theory in which the latter versions of the WGC are violated, so these conjectures should simply be discarded rather than fixed up with an even stronger consistency condition.

In the Higgsing example considered here, the mild form of the WGC is preserved by the Higgsing process. Similarly, if we assume that the tower or sublattice WGCs are satisfied before Higgsing, we find that they are still satisfied after Higgsing by the tower of particles with charge proportional to $(n, 1)$. However, this is no longer automatically true when we generalize our theory. If we assume that there is mixing in the charge lattice between the two U(1)'s, such that the canonically normalized charge vectors take the forms $(g_A, 0)$ and (g_B^1, g_B^2), with g_B^1/g_A irrational. If we assume that the sublattice WGC is exactly saturated before Higgsing so that there are no particles in the theory charged under B strictly below the WGC bound, then by giving a VEV to a scalar field with charge 0 under B we find that there are no superextremal particles in the Higgsed theory. In this special case, the tower, sublattice, and mild forms of the WGC are all violated after Higgsing.

However, this special scenario is not likely in practice. It is true that the WGC bound may be exactly saturated in much or all of the charge lattice: this happens, for instance, in theories with extended supersymmetry, where BPS bounds may forbid strictly superextremal particles in certain directions in the charge lattice. However, the same BPS bound ensures that any Higgs field with the required charges is massive; hence, the previously discussed problematic Higgsing scenario does not arise.

We conclude that the tower or sublattice WGC and the mild form of the WGC are unlikely to be violated by Higgsing in any UV-complete theory of quantum gravity. However, even in the example considered previously in this section, the sublattice of superextremal particles after Higgsing may be much sparser than the sublattice of superextremal particles before Higgsing, with indices that differ by a factor of n. Relatedly, the superextremal particle of charge $q_2 = (0, 1)$ may have a mass m_2 that is well above the magnetic WGC scale of the IR theory: $\Lambda \sim g_{\text{eff}} M_{\text{Pl}} \sim g M_{\text{Pl}}/n$ (Saraswat, 2017); see also Furuuchi (2018). To ensure that the lightest superextremal particles do not have parametrically large charge, one must argue for an O(1) upper bound on the charge n of the Higgs field in this theory. Little work has gone into arguing for such an upper bound [outside of specific string theory contexts (Ibáñez and Montero, 2018)], although it would be a worthwhile direction for future research.

C. String theory examples

1. Heterotic string theory

As a first example, we consider SO(32) heterotic string theory in ten dimensions. The low-energy effective action in the Einstein frame is given by (Polchinski, 2007)

$$\frac{1}{2g_s^2\kappa_{10}^2}\int d^{10}x\sqrt{-g}\left(R - \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi\right) - \frac{1}{2g_s^2 g_{10}^2}\int d^{10}x\sqrt{-g}e^{-\phi/2}\text{Tr}_V(|F_2|^2), \tag{59}$$

where Tr_V is the trace in the fundamental representation, normalized such that $\text{Tr}_V(T^aT^b) = 2\delta^{ab}$ for the basis of generators T^a.

We may then define

$$8\pi G_N = M_{10}^{-8} = g_s^2\kappa_{10}^2 = \tfrac{1}{2}g_s^2(2\pi)^7\alpha'^4, \tag{60}$$

$$e^2 = \tfrac{1}{2}g_s^2 g_{10}^2 = g_s^2(2\pi)^7\alpha'^3, \tag{61}$$

where e^2 is the coupling constant associated with any single U(1) in the maximal torus. Notice that our dilaton coupling parameter is $\alpha = 1/2$, which by Eq. (13) gives $\gamma = 1$.

The charge lattice of the SO(32) heterotic string consists of all charge vectors of the following form:

$$\vec{q} = (q_1, q_2, \ldots, q_{16}) \quad \text{or} \quad \vec{q} = (q_1 + \tfrac{1}{2}, \ldots, q_{16} + \tfrac{1}{2}), \quad \text{with} \quad q_i \in \mathbb{Z}, \quad \sum_i q_i \in 2\mathbb{Z}. \tag{62}$$

This lattice is even, i.e., $|q|^2 \in 2\mathbb{Z}$ for any $\vec{q}$ in the lattice. States must satisfy the level-matching condition

$$\frac{\alpha'}{4}m^2 = N_L + \frac{1}{2}|\vec{q}|^2 - 1 = N_R - \frac{1}{2}, \tag{63}$$

where $N_{L,R}$ are the occupation numbers of the left- and right-moving oscillators, with N_L a non-negative integer and N_R a positive half integer. Given any choice of $N_L \geq 0$ and $\vec{q} \neq 0$, we can always choose N_R to satisfy the level-matching condition. Thus, the lightest state with a given $\vec{q} \neq 0$ has

$$m^2 = \frac{2}{\alpha'}(|\vec{q}|^2 - 2). \tag{64}$$

The charge-to-mass vector of this state then obeys

$$|\vec{Z}|^2 = \frac{2}{\alpha'}\left|\frac{\vec{q}}{m}\right|^2 = \frac{|\vec{q}|^2}{|\vec{q}|^2 - 2} > 1, \tag{65}$$

which shows that the state is superextremal. This means that there is a superextremal particle in every representation of the SO(32) gauge group, so the theory satisfies the non-Abelian sublattice WGC (in fact, it satisfies even the lattice WGC). Compactifying this theory on T^n and turning on Wilson lines, the gauge group is generically broken to its Cartan subgroup, and the resulting theory will satisfy the lattice WGC for

Abelian gauge groups. The same is true for $E_8 \times E_8$ heterotic string theory.

2. *F* theory

Consider F theory compactified to six dimensions on an elliptically fibered Calabi-Yau threefold Y_3 with base B_2. A gauge symmetry G arises from a stack of 7-branes wrapping a holomorphic curve C in the base, and the gauge coupling and 6D Planck scale are related to the volume of B_2 and C via

$$M_6^4 \sim \mathrm{vol}(B_2), \qquad \frac{1}{g_{\mathrm{YM}}^2} \sim \mathrm{vol}(C). \tag{66}$$

Lee, Lerche, and Weigand (2018) showed that the limit $g_{\mathrm{YM}} \to 0$ with M_6 finite can be achieved only if the base B_2 contains a rational curve C_0 whose volume goes to zero as $\mathrm{vol}(C) \to \infty$. A $D3$-brane wrapping C_0 gives rise to a string with string states charged under the gauge group, and in the tensionless limit $\mathrm{vol}(C_0) \to 0$ this string is identified with a heterotic string in a dual description.

In some cases, such as when the base B_2 is a Hirzebruch surface, the dual heterotic string is weakly coupled. In this case, the sublattice WGC follows from modular invariance, as we show in Sec. IV.D. If the heterotic string is strongly coupled, however, it is not so simple to compute the spectrum of string states, and the best that one can do is to compute an index of charged BPS string states using the elliptic genus. Using properties of the elliptic genus, Lee, Lerche, and Weigand (2018) argued that the sublattice WGC is necessarily satisfied with respect to the gauge group G. This was strengthened by Cota, Klemm, and Schimannek (2021) to an argument that 6D F-theory compactifications on Calabi-Yau threefolds in fact satisfy the non-Abelian sublattice WGC of Sec. III.E.

Subsequent work (Lee, Lerche, and Weigand, 2019b; Klaewer *et al.*, 2021) analyzed the elliptic genera of tensionless strings coming from wrapped $D3$-branes in 4D theories coming from F theory compactified on elliptically fibered Calabi-Yau fourfolds. For generic fluxes, properties of the elliptic genus suffice to prove the sublattice WGC. For nongeneric fluxes, there are still superextremal string states, but (unlike in six dimensions) these superextremal particles do not necessarily furnish a sublattice. Indeed, Lee, Lerche, and Weigand (2019b) identified an example of an F-theory compactification for which the elliptic genus detects no superextremal string states of charge $4\vec{q}$ for any $\vec{q}$ in the charge lattice. This does not necessarily imply a counterexample to the sublattice WGC, however, as there are other sectors of charged states not visible to the elliptic genus, and it is conceivable that these sectors may contain the requisite superextremal particles to satisfy the sublattice WGC.

3. Counterexample to the lattice WGC

We have seen a number of examples that satisfy the WGC, as well as its stronger variants. We now present an example that violates a number of proposed strong forms of the WGC. Nonetheless, it still satisfies the WGC, tower WGC, and sublattice WGC.

The example in question comes from compactifying type II string theory on the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orbifold with orbifold action defined by the two generators:

$$\begin{aligned} &\theta\colon\ \theta_4 \mapsto \theta_4 + \pi, \qquad \theta_5 \mapsto \theta_5 + \pi, \\ &\omega\colon\ \theta_6 \mapsto \theta_6 + \pi, \quad \theta_i \mapsto -\theta_i, \quad i = 1, \ldots, 4. \end{aligned} \tag{67}$$

In Eq. (67) the T^6 in question is parametrized by the angles $\theta^i \cong \theta^i + 2\pi$, where $i = 1, \ldots, 6$. For simplicity we take the metric to be diagonal in the θ_i basis. Note that the ω generator acts as a "rototranslation": a rotation combined with a translation in a different direction. This rototranslation acts freely, and thus the orbifold geometry is smooth. As a result, the compactification can be understood within supergravity, as well as on the string world sheet.

For our purposes, it suffices to concentrate on the θ_4, θ_5, and θ_6 dimensions of T^6. Each of these dimensions has a gauge field associated with Kaluza-Klein momentum around the S^1; we denote them, respectively, as A_μ^4, A_μ^5, and A_μ^6. The action of ω projects the first of these fields out of the spectrum, leaving A_μ^5 and A_μ^6 as the only Kaluza-Klein gauge bosons in the theory.

Next consider a field ϕ on T^3 parametrized by θ_4, θ_5, and θ_6. Its field decomposition is given by

$$\phi(x^\mu, \theta_4, \theta_5, \theta_6) = \sum \phi_{n_4,n_5,n_6}(x^\mu) e^{in_4\theta_4 + in_5\theta_5 + in_6\theta_6}. \tag{68}$$

The orbifold action imposes the identifications

$$\begin{aligned} &\theta\colon\ \phi_{n_4,n_5,n_6}(x) = (-1)^{n_4+n_5}\phi_{n_4,n_5,n_6}(x), \\ &\omega\colon \phi_{n_4,n_5,n_6}(x) = (-1)^{n_6}\sigma(\phi)\phi_{-n_4,n_5,n_6}(x). \end{aligned} \tag{69}$$

In Eq. (69) $\sigma(\phi)$ denotes an additional sign that may arise, depending on the nature of the field ϕ: for instance, the graviton, A_μ^5, and A_μ^6 have $\sigma = 1$, whereas W_μ has $\sigma = -1$ due to the action of ω (and is therefore projected out of the spectrum).

Now we look at the sublattice of the charge lattice consisting of the charges (n_5, n_6) under the surviving Kaluza-Klein fields A_μ^5 and A_μ^6. For n_5 and n_6 both even, Kaluza-Klein modes of the graviton with $n_4 = 0$ are projected in and saturate the extremality bound, so there are indeed superextremal particles of these charges. For n_6 odd and n_5 even, a mode will be projected out unless it has $\sigma = -1$, but KK modes of the gauge field A_μ^4 satisfy this condition and similarly saturate the extremality bound. For odd n_5, however, the action of θ imposes the constraint that n_4 must be odd, which leads to an additional contribution of $(n_4/R)^2$ to the mass squared of such a mode:

$$m^2 = \left(\frac{n_5}{R_5}\right)^2 + \left(\frac{n_4}{R_4}\right)^2, \qquad n_5 \text{ odd}, \ n_4 \text{ odd}. \tag{70}$$

This additional contribution renders such modes subextremal: there are no superextremal particles of charge (n_5, n_6) for n_5 odd. This result is summarized in Table I.

TABLE I. Superextremal particles in type II compactified on the $T^6/(\mathbb{Z}_2 \times \mathbb{Z}_2')$ orbifold exist for n_5 even but not for n_5 odd. As a result, the lattice WGC is violated, whereas the sublattice WGC is satisfied with coarseness 2.

n_6 or n_5	Even	Odd
Odd	✗	✗
Even	✓	✓

This theory represents a counterexample to the lattice WGC: there are charges in the charge lattice without superextremal particles, namely, any charge (n_5, n_6) with n_5 odd. By moving in the moduli space of the theory, the sizes of the R_i of the cycles of the torus can be adjusted freely, and for certain values of the moduli additional proposed strong forms of the WGC may also be violated. For example, taking $R_4 \gg R_5 > R_6$ and $R_I \gg \sqrt{\alpha'}$, the winding modes become heavy and the lightest charged particle in the spectrum is subextremal with $(n_5, n_6) = (1, 0)$. This particle is also the state of smallest charge in its direction in the lattice. Thus, this theory represents a counterexample to both of the strong forms of the WGC considered by AMNV (Arkani-Hamed, Motl *et al.*, 2007): neither the lightest charged particle nor the particle of smallest charge in the n_5 direction in the lattice are superextremal. Furthermore, the masses of the particles of odd A^5 charge n violate the convexity condition:

$$2m_n \geq m_{n+1} + m_{n-1}, \tag{71}$$

where m_n is the mass of the lightest particle of charge n. If there is an AdS analog of this example, it would violate the strong forms of the Abelian convex charge conjecture of Aharony and Palti (2021) introduced in Sec. III.F.

On the other hand, the tower WGC and the sublattice WGC are satisfied in this example: given any charge $\vec{q}$, there is a superextremal particle of charge $2\vec{q}$. The sublattice of superextremal particles therefore has coarseness 2.

Finally, we remark on a puzzling feature of this example: at tree level in string perturbation theory, the lightest particles with odd n_5 are in fact stable (Heidenreich, Reece, and Rudelius, 2017). This suggests that black holes of odd n_5 charge cannot decay, in violation of the original motivation of the WGC. It is possible that loop corrections could modify the spectrum so that this conclusion could be avoided; more work is needed to see whether this possibility is actually realized.

A number of other counterexamples to the lattice WGC were identified by Heidenreich, Reece, and Rudelius (2017). These counterexamples all involve orbifold compactifications of string theory, and all of them satisfy the sublattice WGC with a superextremal sublattice of coarseness no larger than 3. Furthermore, in all such examples the majority of sites in the charge lattice have superextremal particles (even sites outside the superextremal sublattice). This means that when it comes to the existence of superextremal particles quantum gravity seems to impose even stronger constraints than the tower or sublattice WGC; such constraints are seldom discussed, simply because it is not easy to formulate them as precise mathematical statements.

4. Axions in string theory

Recall that the WGC for axions (26) implies an upper bound on the axion decay constant f in terms of the instanton action S,

$$f \lesssim \frac{M_{\rm Pl}}{S}. \tag{72}$$

Within string theory, the condition $S \gtrsim 1$ is typically required for perturbative control. For instance, the instanton action may represent the size of some compactification cycle in string units, so the α' expansion breaks down when this cycle is smaller than the string scale. The WGC for axions thus amounts to the condition that $f \lesssim M_{\rm Pl}$ within the perturbative regime of string theory.

In fact, this condition was famously pointed out by Banks *et al.* (2003) even before the original AMNV paper on the WGC. In particular, Banks *et al.* (2003) considered axions in heterotic, type I, type IIA, type IIB, and M theory compactified to four dimensions. In all cases, these axions arise either as the periods of a p-form C_p over a p-cycle Σ_p of the compactification manifold $\oint_{\Sigma_p} C_p$ or as the dual of a 2-form gauge field $B_{\mu\nu}$ in four dimensions. In all cases, their decay constants were found to be bounded from above[8] as $f \lesssim M_{\rm Pl}$.

More recently a number of works have taken advantage of an improved understanding of Calabi-Yau compactifications to investigate the prospects for super-Planckian decay constants in type IIA or IIB string theory with greater precision (Montero, Uranga, and Valenzuela, 2015; Palti, 2015; Rudelius, 2015a; Bachlechner, Long, and McAllister, 2016; Brown *et al.*, 2016; Conlon and Krippendorf, 2016; Junghans, 2016; Hebecker, Henkenjohann, and Witkowski, 2017a; Long, McAllister, and Stout, 2017; Zyla *et al.*, 2020). In all cases, the axion decay constants appear to be bounded from above by $f \lesssim M_{\rm Pl}$. This remains true even in theories with multiple axions; see Sec. VII.A.5.

The size of decay constants allowed in string theory is important because models of so-called natural inflation require $f > M_{\rm Pl}$. From the perspective of inflation, however, we are interested not only in the kinematic question of how large an axion decay constant can be but also in the dynamical question of how an axion rolls in its potential. Obtaining the potential of an axion in a type IIB compactification is not simple, as it requires information about the sheaf cohomology of curves or divisors of the Calabi-Yau compactification manifold, which is not known in general. Some progress in understanding the relevant sheaf cohomology was made by Braun *et al.* (2020). Finally, note that bounds on axion decay constants do not directly constrain axion monodromy models (Silverstein and Westphal, 2008; McAllister, Silverstein, and Westphal, 2010).

The axion WGC can also be studied outside of the context of specific string constructions. In general, the breakdown of the instanton expansion that arises when $S \lesssim 1$ is always due to the presence of new light states. This has been argued to

[8] Banks *et al.* (2003) incorrectly claimed the model-independent heterotic axion, i.e., the 4D dual of $B_{\mu\nu}$, has $f = M_{\rm Pl}$. In fact, it has $f \sim M_s^2/M_{\rm Pl}$; see Svrcek and Witten (2006).

follow from the general Lee-Yang theory of phase transitions (Stout, 2022). If the potential $V(\theta)$ is a smooth function of the axion value θ, then its harmonics asymptotically decay as $\exp(-nS)$, where S is determined by the location of the nearest singularity ζ_* to the unit circle for the complexified coordinate $\zeta = \exp(i\theta)$. This asymptotic notion of S has been suggested to define the correct formulation of the axion WGC in the limit when instantons are not well-defined semiclassical objects (Stout, 2022).

D. Modular invariance

Thus far we have seen that the WGC and its tower versions hold true in a large class of examples in string or M theory. In this section, we present a general argument for the sublattice WGC in 2D CFTs on the basis of modular invariance. This result can be viewed as either (a) a proof of the sublattice WGC in perturbative string theory viewing the 2D CFT as the world sheet theory (Heidenreich, Reece, and Rudelius, 2017) or (b) a proof of the sublattice WGC in AdS_3 viewing the 2D CFT as the boundary dual of an AdS_3 theory (Montero, Shiu, and Soler, 2016).

A 2D CFT has a partition function of the form[9]

$$Z(\mu, \tilde{\mu}; \tau, \bar{\tau}) \equiv \mathrm{Tr}(q^{\Delta} \bar{q}^{\tilde{\Delta}} y^{Q} \tilde{y}^{\tilde{Q}}), \tag{73}$$

where $\Delta = L_0 - c/24$, $\tilde{\Delta} = \tilde{L}_0 - \tilde{c}/24$, Q and $\tilde{Q}$ are the charges carried by left and right movers under a conserved current, $q = e^{2\pi i\tau}$, $y = e^{2\pi i\mu}$, and $\tilde{y} = e^{2\pi i\tilde{\mu}}$. The partition function satisfies

$$Z(\mu + \rho) = Z(\mu) \qquad \forall\ \rho \in \Gamma_Q^*, \tag{74}$$

where $\Gamma_Q^* = \{(\rho, \tilde{\rho}) | \rho Q - \tilde{\rho}\, \tilde{Q} \in \mathbb{Z}\}$ is the dual lattice to the charge lattice.

Modular transformations form a group $\mathrm{SL}(2, \mathbb{Z})$ generated by the T transformation $\tau \to \tau + 1$ and the S transformation $\tau \to -1/\tau$. Following Benjamin *et al.* (2016), we find that modular transformations act on the partition function as

$$\begin{aligned} Z(\mu; \tau + 1) &= Z(\mu; \tau), \\ Z(\mu/\tau; -1/\tau) &= e^{\pi i k(\mu^2/\tau) - \pi i \tilde{k}(\tilde{\mu}^2/\bar{\tau})} Z(\mu; \tau). \end{aligned} \tag{75}$$

In Eq. (75) k and $\tilde{k}$ are related to the leading term in the current-current operator product expansion,

$$J_L(z) J_L(0) \sim \frac{k}{z^2} + \cdots, \qquad J_R(\bar{z}) J_R(0) \sim \frac{\tilde{k}}{\bar{z}^2} + \cdots. \tag{76}$$

Unitarity implies that k and $\tilde{k}$ are non-negative, and positive for nontrivial currents. For the case of multiple currents, this becomes

$$J_L^a(z) J_L^b(0) \sim \frac{k^{ab}}{z^2} + \cdots, \qquad J_R^{\tilde{a}}(\bar{z}) J_R^{\tilde{b}}(0) \sim \frac{\tilde{k}^{\tilde{a}\tilde{b}}}{\bar{z}^2} + \cdots. \tag{77}$$

k^{ab} and $\tilde{k}^{ab}$ can be thought of as metrics, which raise and lower indices and define inner products. Thus, we may write $\mu \cdot \rho \equiv \mu_a k^{ab} \rho_b$, $\mu \cdot Q \equiv \mu_a Q^a$, and $\tilde{Q}^2 \equiv \tilde{Q}^{\tilde{a}} \tilde{k}^{-1}_{\tilde{a}\tilde{b}} \tilde{Q}^{\tilde{b}}$.

Next we combine the periodicity condition (74) with the S-duality transformation (75), which implies

$$\begin{aligned} Z(\mu + \tau\rho; \tau) = \exp[&-2\pi i(\mu \cdot \rho) - \pi i \tau \rho^2 + 2\pi i(\tilde{\mu} \cdot \tilde{\rho}) \\ &+ \pi i \tau \tilde{\rho}^2] Z(\mu; \tau). \end{aligned} \tag{78}$$

The partition function encodes the spectrum of the theory, which means that the quasiperiod $\mu \to \mu + \tau\rho$ must map the spectrum to itself. This occurs thanks to a rearrangement of simultaneous changes in charge and conformal weight, a phenomenon known as spectral flow (Schwimmer and Seiberg, 1987). To describe this, we define

$$T \equiv \Delta - \tfrac{1}{2} Q^2, \qquad \tilde{T} \equiv \tilde{\Delta} - \tfrac{1}{2} \tilde{Q}^2, \tag{79}$$

which allows us to rewrite Eq. (78) as

$$\begin{aligned} Z &= \mathrm{Tr}(q^{T + (1/2)Q^2} \bar{q}^{\tilde{T} + (1/2)\tilde{Q}^2} y^{Q} \tilde{y}^{\tilde{Q}}) \\ &= \mathrm{Tr}(q^{T + (1/2)(Q+\rho)^2} \bar{q}^{\tilde{T} + (1/2)(\tilde{Q}+\tilde{\rho})^2} y^{Q+\rho} \tilde{y}^{\tilde{Q}+\tilde{\rho}}), \end{aligned} \tag{80}$$

where we have introduced the shorthand notation

$$y^Q \equiv \exp[2\pi i \mu_a Q^a], \qquad \tilde{y}^{\tilde{Q}} \equiv \exp[-2\pi i \tilde{\mu}_a \tilde{Q}^a]. \tag{81}$$

By expanding the traces in Eq. (80) in powers of Q and matching the first and second lines of the equation, we find that the spectrum must be invariant under

$$Q \to Q + \rho, \qquad \tilde{Q} \to \tilde{Q} + \tilde{\rho}, \tag{82}$$

with T and $\tilde{T}$ held fixed. This implies

$$\Gamma_Q^* \subseteq \Gamma_Q, \tag{83}$$

i.e., the dual lattice Γ_Q^* is a sublattice of the charge lattice Γ_Q. From here, beginning with the graviton, which has $\Delta = \tilde{\Delta} = 0$ and $Q = \tilde{Q} = 0$, we use invariance of the spectrum under the transformation (82) to deduce the existence of a state with

$$\Delta = \tilde{\Delta} = \frac{\alpha'}{4} m^2 \le \max\left(\frac{1}{2} Q^2, \frac{1}{2} \tilde{Q}^2\right) \tag{84}$$

for all $Q \in \Gamma_Q^*$. One can show that these states are superextremal (Heidenreich and Lotito, 2022), which means that the sublattice Γ_Q^* is entirely populated by superextremal particles, and the sublattice WGC is satisfied.

E. Relation to the swampland distance conjecture

Ooguri and Vafa (2007) introduced a number of swampland conjectures regarding the moduli space $\mathcal{M}$ of a consistent

[9]Here and to follow we use $\tilde{\mu}$ rather than $\bar{\mu}$ as a reminder that μ and $\tilde{\mu}$ are independent variables. Indeed, in many cases, such as the world sheet CFT of heterotic string theory, the number of left-moving and right-moving currents is different, so there is no way to identify the chemical potentials in complex conjugate pairs.

theory of quantum gravity. First, they conjectured that such a moduli space is parametrized by vacuum expectation values of massless scalar fields. Second, they conjectured that such a moduli space has an infinite diameter: there are points at arbitrarily large geodesic distance. Third, they introduced what is now known as the swampland distance conjecture (SDC); see also Klaewer and Palti (2017).

Conjecture 9.—Swampland distance conjecture. Compared to the theory at some point $p_0 \in \mathcal{M}$, the theory at a point $p \in \mathcal{M}$ has an infinite tower of particles, each with a mass scaling as

$$m \sim \exp[-\lambda d(p, p_0)], \tag{85}$$

where $d(p, p_0)$ is the geodesic distance in $\mathcal{M}$ between p and p_0 and λ is some order-1 number in Planck units. Consequently, in the infinite distance limit $d(p, p_0) \to \infty$ an infinite tower of states becomes light.

The appearance of an infinite tower of light particles is reminiscent of the tower WGC, as well as the tower RFC, both of which require that an infinite tower of particles become light in the limit of vanishing gauge coupling. Indeed, in many contexts these conjectures are satisfied simultaneously by the same tower of particles. In 4D theories with $\mathcal{N} = 2$ supersymmetry, for instance, Gendler and Valenzuela (2021) proved that in any infinite distance limit with a vanishing gauge coupling there is an infinite tower of charged particles satisfying

$$\left(\frac{q}{m}\right)^2 \geq \left(\frac{q}{m}\right)^2\Big|_{\text{ext}} = \frac{1}{2}\frac{1}{M_4^2} + g^{ij}\mu_i\mu_j, \tag{86}$$

where μ_i is the scalar charge of the particle with respect to the ith massless scalar field and g^{ij} is the inverse metric on moduli space. The equality on the right-hand side of Eq. (86) implies an equivalence between the WGC bound and the RFC bound for these particles: they are superextremal precisely when they are self-repulsive. This in turn implies $g^{ij}\mu_i\mu_j = \alpha^2 m^2$ for some $O(1)$ constant α. If the moduli space is one dimensional, this implies that these states must also satisfy the SDC, with $\lambda = \alpha$. If the moduli space has a dimension greater than 1, the relationship between λ and α is more complicated, as it depends on the path taken. However, in a number of examples considered by Gendler and Valenzuela (2021) the tower satisfying the tower WGC also satisfies the SDC for some $O(1)$ constant λ.

It is natural to expect that the towers of particles stipulated by the tower WGC and the SDC should agree whenever a point in moduli space of vanishing gauge coupling is also at infinite distance. It has been shown that all infinite distance limits in Kähler moduli space in 4D and 5D supergravity theories descending from type II or M theory on a Calabi-Yau threefold are associated with vanishing gauge coupling (Corvilain, Grimm, and Valenzuela, 2019; Heidenreich and Rudelius, 2021). More generally, it has been conjectured that every infinite distance limit in moduli space should correspond to the vanishing of some p-form gauge coupling (Gendler and Valenzuela, 2021) and, more concretely, it has been conjectured that every infinite distance limit should correspond either to a decompactification limit (at which point a tower of KK modes becomes light) or to a tensionless string limit (Lee, Lerche, and Weigand, 2022b). The latter conjecture goes by the name of the emergent string proposal and, if true, suggests that either a 1-form gauge field or a 2-form gauge field (or both) must become weakly coupled at any infinite distance limit in moduli space. Further evidence for this conjecture in the context of 4D $\mathcal{N} = 1$ string compactifications was provided by Lanza *et al.* (2021a, 2021b). They postulated the closely related distant axionic string conjecture, which states that any infinite distance limit of a 4D $\mathcal{N} = 1$ quantum gravity theory corresponds to the tensionless limit of a string that is charged under some 2-form gauge field.

Conversely, in compactifications of M theory to five dimensions, it has been shown that every point of vanishing gauge coupling is at infinite distance in moduli space (Heidenreich and Rudelius, 2021). An analogous statement holds in the context of AdS/CFT: in any superconformal field theory in $d > 2$ dimensions, every point on the conformal manifold at which some sector of the theory becomes free is at an infinite distance in the Zamolodchikov metric (Perlmutter *et al.*, 2021). This translates to the statement that the vanishing of a gauge coupling must occur at infinite distance in the moduli space of the bulk AdS dual theory. Additional progress toward classifying infinite distance limits in string compactifications was discussed by Grimm, Palti, and Valenzuela (2018), Grimm, Li, and Palti (2019), Marchesano and Wiesner (2019), Grimm and Heisteeg (2020), Grimm, Li, and Valenzuela (2020), Baume and Infante (2021), Álvarez-García, Kläwer, and Weigand (2022), and Lee, Lerche, and Weigand (2022a, 2022c).

In the case of infinite distance, weak coupling limits, therefore, the WGC and SDC can be essentially unified. More general, qualitative arguments can be given in support of some sort of unification between these two conjectures: we elaborate on these arguments in Sec. V.

V. QUALITATIVE ARGUMENTS FOR THE WGC

Having introduced various versions of the WGC and given some empirical evidence for them, we now turn to the question of why any of them might be true. In our view the most compelling arguments of this type are those that are more qualitative. We do not attempt to reproduce some version of the WGC in detail but instead share some intuition for why a statement of this type should hold. In this section we review several such arguments, and in Sec. VI we review attempts at a more precise derivation.

A. Emergence

It has long been suspected that spacetime itself must be emergent in any theory of quantum gravity that is nontrivial enough to have some kind of black hole thermodynamics. One simple argument in this vein is that the Bekenstein-Hawking formula

$$S = \frac{\text{area}}{4G}$$

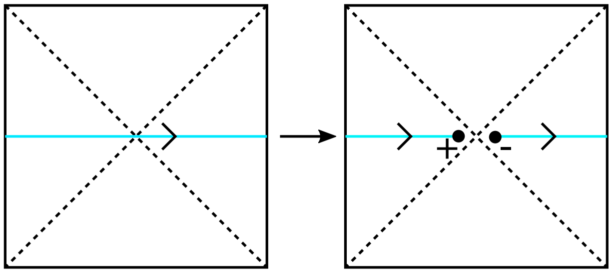

FIG. 8. Wilson line threading an AdS wormhole. For such an operator to respect factorization, we need to be able to split it into a product of "left" and "right" pieces, each consisting of a partial Wilson line ending on a charged operator.

tells us that the maximal entropy in a region of spacetime scales only like the surface area of the region, which is different from the volume scaling we have in quantum field theory ('t Hooft, 1993; Susskind, 1995; Bousso, 2002). This idea is concretely realized in the AdS/CFT correspondence, which formulates quantum gravity in asymptotically AdS spacetime as the quantum mechanics of a dual CFT living on the asymptotic boundary (Maldacena, 1998). Harlow (2016) observed that this emergence can be used to motivate a qualitative version of the WGC. The basic idea is that if spacetime itself is emergent, then surely any bulk gauge fields must also be emergent. It is impossible, however, to have an emergent gauge field without the presence of charged particles, and moreover these charged particles cannot be too heavy.

More concretely, we can consider the maximally extended AdS-Schwarzschild geometry, which in AdS/CFT is dual to the thermofield double state

$$|\mathrm{TFD}\rangle \equiv \frac{1}{\sqrt{Z(\beta)}}\sum_i e^{-\beta E_i/2}|i^*\rangle_L|i\rangle_R \tag{87}$$

that lives in the tensor product Hilbert space

$$\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R \tag{88}$$

of two copies of the CFT on a spatial sphere (Maldacena, 2003a). In the gravity picture this geometry describes a spatial wormhole connecting two AdS boundaries. If there is a U(1) gauge field in the gravity description, we therefore can have Wilson line operators

$$W_{LR} = e^{i\int_{c_{LR}} A}$$

that are integrated on a curve c_{LR} that connects the two boundaries through the wormhole; see Fig. 8. These Wilson lines, however, are somewhat puzzling from the point of view of the tensor product Hilbert space (88). In a tensor product Hilbert space every operator can be written as a sum of product operators, but W_{LR} does not seem to have such a decomposition. Indeed, if we try to view it as a product of two "half-Wilson lines," these parts are not gauge invariant and thus should not act on the physical Hilbert space (88). The way out of this puzzle is that if charged objects exist, we can indeed split the Wilson line into two gauge-invariant operators, each of which consists of a Wilson line connecting an asymptotic boundary to a charged operator; see Fig. 8.

Thus far this has only been an argument for some version of the completeness hypothesis; we have not yet really used the emergence of the gauge field. The idea of Harlow (2016) is as follows: since the gauge field is emergent, there must be some scale $\Lambda_{\mathrm{U}(1)} \le \Lambda_{\mathrm{QG}}$ at which the coefficient of the Maxwell term in the Wilson action flows to zero. Here Λ_{QG} is the scale at which gravity becomes strongly coupled, which in general can be much less than M_{Pl} if there are many degrees of freedom. If the infrared value of the Maxwell coupling e is small, then it must undergo a substantial renormalization group (RG) flow between the scale $\Lambda_{\mathrm{U}(1)}$, where it is large, and the mass scale m of the lightest charge particles since below the scale m the gauge coupling can no longer flow. Therefore, there can be a small infrared gauge coupling if and only if there are light charged particles, which is the essence of the stronger versions of the WGC. Quantifying this argument in general is difficult, but it can be illustrated in concrete models of an emergent U(1) gauge field. For example, the $\mathbb{CP}^{N-1}$ σ model is a theory of N complex scalar fields $z_a(x)$ obeying the constraint $\sum_a z_a^* z_a = 1$ and interacting with the Lagrangian

$$\mathcal{L} = -\frac{N}{g^2}(D^\mu z)^\dagger D_\mu z, \tag{89}$$

where the gauge field A_μ appearing in the covariant derivative $D_\mu = \partial_\mu - iA_\mu$ is given by

$$A_\mu \equiv \frac{1}{2i}(z^\dagger\partial_\mu z - \partial_\mu z^\dagger z). \tag{90}$$

This theory is renormalizable for $D = 2$, while for general D it can be understood as a lattice model with cutoff energy $\Lambda_{\mathrm{U}(1)}$. Either way, there is a critical value of g near which (for large N) its infrared description is as N charged scalars of mass m interacting via Maxwell interactions of strength (Harlow, 2016)

$$\frac{1}{e^2} = \begin{cases} N/6\pi m^2, & D = 2, \\ N/12\pi m, & D = 3, \\ (N/12\pi^2)\log(\Lambda_{\mathrm{U}(1)}/m), & D = 4, \\ N\Lambda_{\mathrm{U}(1)}^{D-4}, & D > 4. \end{cases} \tag{91}$$

If we now couple the model to gravity, for it to work we need $\Lambda_{\mathrm{U}(1)} \le \Lambda_{\mathrm{QG}}$, where Λ_{QG} is the scale where gravity becomes strongly coupled. In a large-N theory this is related to the infrared Newton constant G by (Veneziano, 2002; Arkani-Hamed, Dimopoulos, and Kachru, 2005; Distler and Varadarajan, 2005; Dimopoulos et al., 2008; Dvali and Redi, 2008; Dvali, 2010; Kaplan and Kundu, 2019)

$$\frac{1}{G} \sim \frac{1}{G_{\mathrm{bare}}} + N\Lambda_{\mathrm{QG}}^{D-2}. \tag{92}$$

For the massive scalars to be qualitatively superextremal in the sense of Eq. (5), we need to have

$$\frac{1}{e^2} \lesssim \frac{1}{m^2 G}, \tag{93}$$

which indeed follows from Eqs. (91) and (92) together with $\Lambda_{U(1)} \leq \Lambda_{QG}$ and $G_{bare} > 0$.

There are a few things to note about this argument. First, it is not sufficient to have only one heavy superextremal particle. To break the Wilson line in the unit charge representation as in Fig. 8, we need to have objects of unit charge. This thus gives some qualitative support to the stronger versions of the WGC such as the tower and sublattice WGCs. Second, one might worry that this verification of the WGC is accidental since only a few parameters were involved. In fact, it is robust. In particular, one can consider M copies of the $\mathbb{CP}^{N-1}$ model, each with different values of N and m, and a similar argument shows that the convex hull version of the multiple-U(1) WGC that we discussed in Sec. III.C holds throughout a large parameter space of theories (Harlow and Ooguri, 2021). Moreover, one can also show that the "Grassmannian" generalization of the model, which flows to an SU(N) gauge theory in the infrared, has objects in the fundamental representation that obey the non-Abelian WGC discussed in Sec. III.E.

We can develop this idea further to make it less dependent on the details of the $\mathbb{CP}^{N-1}$ model. Specializing for convenience to $D = 4$, at one-loop order the gauge coupling e_{UV} at an energy scale Λ_{UV} is related to the low-energy gauge coupling e according to

$$\frac{1}{e_{UV}^2} = \frac{1}{e^2} - \sum_i \frac{b_i}{8\pi^2} q_i^2 \log\frac{\Lambda_{UV}}{m_i}. \tag{94}$$

In Eq. (94) m_i and q_i are the mass and charge of the particles in the tower and b_i is a beta function coefficient. For Λ_{UV} sufficiently large, the right-hand side of Eq. (94) vanishes and, correspondingly, e_{UV} diverges: this is the well-known Landau pole of U(1) gauge theory coupled to charged matter. The energy scale $\Lambda_{U(1)}$ of the Landau pole thus represents a UV cutoff on the U(1) gauge theory. Gravity has a similar UV cutoff Λ_{QG} that can be thought of as a result of divergent one-loop corrections to the Einstein-Hilbert term. This is described by Eq. (92), which for $D = 4$ implies that

$$M_{Pl}^2 \gtrsim N_{d.o.f.} \Lambda_{QG}^2. \tag{95}$$

The energy scale Λ_{QG} is the energy scale at which quantum effects significantly modify gravity. It is also sometimes referred to as the species bound scale, as it scales inversely with the number of light particle species $N_{d.o.f.}$ in the theory.

Now suppose our theory has a superextremal particle of each integer charge, so it satisfies the sublattice WGC (and, in fact, the lattice WGC). The number of particles below a mass scale Λ is then given by $N(\Lambda) \geq \Lambda / e M_{Pl}$, so the species bound satisfies

$$M_{Pl}^2 \gtrsim N(\Lambda_{QG}) \Lambda_{QG}^2 \geq \frac{\Lambda_{QG}}{e M_{Pl}} \Lambda_{QG}^2 \tag{96}$$

or, equivalently,

$$\Lambda_{QG} \lesssim e^{1/3} M_{Pl}. \tag{97}$$

We see that in the weak coupling limit $e \to 0$ the species bound scale tends to zero and effective field theory breaks down due to the tower of superextremal particles.

This tower of charged particles also affects the Landau pole of the gauge theory $\Lambda_{U(1)}$. Treating the logarithms and numerical prefactors as parametrically order 1, the gauge coupling e_{UV} in Eq. (94) diverges when

$$\frac{1}{e^2} \sim \sum_{q=1}^{Q} q^2 \sim Q^3, \tag{98}$$

where $Q \sim \Lambda_{U(1)} / e M_{Pl}$ is the largest charge in the tower. This again leads to the conclusion

$$\Lambda_{U(1)} \sim e^{1/3} M_{Pl}. \tag{99}$$

Thus, we see that the tower of superextremal particles leads to UV cutoffs on both gauge theory and gravity. Moreover, for the simple spectrum of charged particles we have considered here, the UV cutoffs for gauge theory and gravity are at parametrically the same energy scale,[10] namely, $\Lambda_{QG} \sim e^{1/3} M_{Pl}$. In a sense, gauge theory and gravity are "unified" at this energy scale, as both of them emerge in the infrared from a strongly coupled theory at the energy scale Λ_{QG} by integrating out a tower of charged states.

Conversely, we now assume that the gauge theory becomes strongly coupled at or below the energy scale Λ_{QG},

$$\frac{1}{e^2} \sim \sum_{i | m_i < \Lambda_{gauge}} q_i^2 \qquad \text{for } \Lambda_{gauge} \lesssim \Lambda_{QG}, \tag{100}$$

where again we are ignoring O(1) factors. We can rewrite this in terms of the average charge squared $\langle q^2 \rangle_\Lambda$ of the particles with mass below Λ as

$$\frac{1}{e^2} \sim N(\Lambda_{gauge}) \langle q^2 \rangle_{\Lambda_{gauge}} \lesssim \frac{1}{\Lambda_{gauge}^2} M_{Pl}^2 \langle q^2 \rangle_{\Lambda_{gauge}}, \tag{101}$$

where we have used the definition of the species bound (95). Finally, we may rearrange this result in a form reminiscent of the WGC bound,

$$\Lambda_{gauge}^2 \lesssim e^2 \langle q^2 \rangle_{\Lambda_{gauge}} M_{Pl}^2. \tag{102}$$

Since all of the particles contributing to $\langle q^2 \rangle_{\Lambda_{gauge}}$ have a mass below Λ_{gauge}, we see that, in a sense, the "average" charged particle in the theory is superextremal. This is not the same as the condition that the theory satisfies the tower WGC, but it points in that direction.

[10]The same parametric cutoff has appeared in a significantly different EFT context involving photons with a Stueckelberg mass (Craig, Garcia, and Kribs, 2020). It would be interesting to determine whether this is a coincidence or something deeper.

We have considered here only one simple case: U(1) gauge theory in four dimensions with a single superextremal particle of each integer charge. However, as shown by Heidenreich, Reece, and Rudelius (2018b), this phenomenon of gauge-gravity unification generalizes to theories in $d \geq 4$ spacetime dimensions, multiple U(1)'s, non-Abelian gauge groups satisfying the sublattice WGC for non-Abelian WGC (see Sec. III.E), theories that satisfy the tower and sublattice WGCs, but not the lattice WGC (provided that the tower of superextremal states is not too sparse), and theories with a more general density of states (provided that the density of states is sufficiently well behaved). In this wide array of theories, gauge theory and gravity become strongly coupled at parametrically the same energy scale Λ_{QG}. Conversely, demanding that gauge theory and gravity become strongly coupled at parametrically the same energy scale implies that, in the same sense as Eq. (102), the average particle should satisfy the WGC bound.

Finally, we remark that a similar emergence picture applies to scalar field theories that satisfy the SDC: just as loop effects from a tower of superextremal particles lead to a strongly coupled gauge theory at the scale Λ_{QG}, loop effects from a tower of particles satisfying the SDC lead to a strongly coupled scalar field theory at the scale Λ_{QG} (Grimm, Palti, and Valenzuela, 2018; Heidenreich, Reece, and Rudelius, 2018a). This concept of emergence thereby unifies the SDC and the sublattice WGC. Indeed, in many cases the tower of particles that satisfies the sublattice WGC also satisfies the SDC, and integrating out this tower of particles produces both a weakly coupled gauge theory and a weakly coupled scalar field theory in the IR.

B. No approximate global symmetries

As is familiar from a first-year course on electromagnetism, Gauss's law holds that the total electric flux through a closed two-dimensional surface S is equal to the charge enclosed. In particular, the size and shape of the surface is irrelevant: the surface may be continuously deformed, and the total electric flux will not change provided that the charge enclosed remains constant.

The modern notion of a higher-form global symmetry offers another perspective on this scenario. The fact that such deformations of the surface S do not affect the total electric flux through it signals the existence of a family of topological surface operators in the theory, which are labeled by an angle $\alpha \in [0, 2\pi)$ and given by the exponentiated electric flux integral,

$$U_\alpha(S) = \exp\left(i\frac{\alpha}{e^2}\oint_S \star F\right), \qquad \alpha \in [0, 2\pi). \tag{103}$$

This surface operator signals the existence of a 1-form global symmetry, which is associated with a conserved charge: namely, the electric flux through S counts the charge of any probe particles contained in such a surface. The conserved Noether current associated with this symmetry is given by the electric flux density $J = 1/e^2 \star F$.

This symmetry is broken in the presence of dynamical charged particles, which screen the charge of a probe particle. At long distances, however, the charge is approximately conserved: the divergence of the Noether current $\partial_\mu F^{\mu\nu}$ is small, and the flux through a closed surface enclosing a probe particle has only weak dependence on the size of the surface. This is encoded by the effective electromagnetic potential in QED at distances that are large compared to the mass of the electron, also known as the Uehling potential (Uehling, 1935),

$$V(r) = \frac{-e^2}{4\pi r}\left(1 + \frac{e^2}{16\pi^{3/2}}\frac{e^{-2mr}}{(mr)^{3/2}} + \cdots\right), \qquad rm \gg 1. \tag{104}$$

In Eq. (104) e is the renormalized coupling constant in the IR. We see that corrections to the leading-order Coulomb potential are exponentially suppressed at long distances, and the charge $\oint_{S^2(r)} \star F \propto \oint_{S^2(r)} V'(r)$ is approximately conserved.

At distances $r \sim 1/m$, on the other hand, one starts to penetrate the polarization cloud and see the bare charge. The gauge coupling runs logarithmically, and the corrections to the effective Coulomb potential from the electron are $O(e^2)$.

More generally, corrections to the Coulomb potential at a distance scale $r = 1/\Lambda$ from a tower of charged particles are given roughly by

$$\Pi(\Lambda^2) = \sum_{i|m_i<\Lambda} e^2 q_i^2. \tag{105}$$

Equation (105) is the same expression we saw in our previous discussion of emergence, so the corrections become O(1) precisely when the gauge theory becomes strongly coupled. We showed there that a U(1) gauge theory satisfying the tower or sublattice WGCs will become strongly coupled at the scale Λ_{QG} at which gravity becomes strongly coupled, so by the same token the approximate 1-form symmetry of such a gauge theory will be badly broken at Λ_{QG}. This gives us a new intuitive understanding of the tower or sublattice WGCs. These conjectures are intimately tied to the absence of global symmetries in quantum gravity, including higher-form and approximate global symmetries. Other versions of the WGC, including the magnetic version and the 0-form version, can be similarly related to the absence of approximate global symmetries in quantum gravity (Córdova, Ohmori, and Rudelius, 2022).

C. Axion strings

As a final qualitative check, we review an argument for the WGC in the presence of Chern-Simons terms (Heidenreich, Reece, and Rudelius, 2021). This argument is somewhat circular from the point of view of establishing the WGC, as it assumes the WGC for axions and charged strings in order to prove the ordinary WGC for charged particles. Nonetheless, it demonstrates an important phenomenon: in the presence of Chern-Simons terms involving multiple gauge fields, the WGC bounds for these different gauge fields are "mixed up" with one another. This offers a bottom-up criterion for determining when the tower of superextremal particles demanded by the tower or sublattice WGCs are modes of some fundamental string, which aligns with recent work (reviewed previously in Sec. IV.E) examining emergent strings in infinite distance limits (Lanza *et al.*, 2021b; Lee, Lerche, and Weigand, 2022b).

The argument relies on five simple assumptions. First, we assume a 4D theory of axion electrodynamics, in which an axion couples to the gauge field via a $\theta F \wedge F$ Chern-Simons coupling,

$$S = \int \left[-\frac{1}{2g^2} F \wedge \star F - \frac{1}{2} f_\theta^2 d\theta \wedge \star d\theta + \frac{1}{8\pi^2} \theta F \wedge F \right]. \tag{106}$$

Second, we assume the axion WGC,

$$f_\theta S \lesssim M_{\rm Pl}, \tag{107}$$

where S is the instanton action. Third, we assume the WGC for a string of tension T charged magnetically under the axion, also known as an axion string,

$$T \lesssim 2\pi f_\theta M_{\rm Pl}. \tag{108}$$

Fourth, we assume that the instanton action takes the form

$$S = \frac{8\pi^2}{g^2}. \tag{109}$$

This form of the instanton action is most familiar from Yang-Mills theory, but Abelian gauge theories also feature instantons with actions of this type, in the form of monopole loops with dyonic winding (Fan *et al.*, 2021), as a consequence of the Witten effect (Witten, 1979). Finally, we assume that the axion θ is a fundamental axion, meaning that the core of the axion string probes physics in the deep ultraviolet. [For more details on the distinction of fundamental versus nonfundamental strings, see Dolan *et al.* (2017) and Reece (2019).]

From here we may combine Eqs. (107)–(109) to get a bound on the string scale of the axion string,

$$M_{\rm str} \coloneqq \sqrt{2\pi T} \lesssim g M_{\rm Pl}, \tag{110}$$

which is precisely the WGC scale associated with the gauge field A. Next our assumption of the Chern-Simons coupling $\theta F \wedge F$ ensures that the higher-spin string excitations of the axion string carry charge under the gauge field A, which follows from anomaly inflow on the string world sheet (Callan and Harvey, 1985). From Eq. (110), we learn that the excitations of the axion string satisfy the WGC [up to O(1) factors].

Finally, invoking our assumption that the axion is a fundamental axion, we further conclude that there is an entire tower of string excitations. This establishes [up to O(1) factors] not only the WGC but also the tower WGC for the gauge field A. Additionally, local quantum field theory breaks down at the axion string scale $M_{\rm str} \sim g M_{\rm Pl}$, which for g small is parametrically below the emergence energy scale $g^{1/3} M_{\rm Pl}$ discussed in Sec. V.A. This can have important consequences for phenomenology in that it leads to tension between effective field theories that require a high-energy scale and those that require a small gauge coupling.

As noted, this argument represents more of a consistency check on the WGC than an argument for it, as it assumes the WGC for axions and axion strings. Furthermore, note that it relies heavily on the presence of the Chern-Simons term: without this term, there is no reason for the form of the instanton action in Eq. (109) to hold, and there is no guarantee that the excitations of the axion string will carry electric charge. Indeed, the circle compactification of a 5D gravity theory yields a Kaluza-Klein photon that does not couple to the axion via a $\theta F \wedge F$ coupling. Consequently the Kaluza-Klein modes are not excitations of the axion string, and effective field theory breaks down at the larger scale $e_{\rm KK}^{1/3} M_4 = M_5$. For more details, see Heidenreich, Reece, and Rudelius (2021).

The bottom-up argument of this section coheres nicely with studies of infinite distance limits in string theory discussed in Sec. IV.E. In particular, the emergent string conjecture (Lee, Lerche, and Weigand, 2022b) implies that every weak coupling limit should correspond to either an emergent string limit or a decompactification limit. We see here that these two cases are distinguished at low energies by the presence or absence of a Chern-Simons coupling. Similarly, Lanza *et al.* (2021b) found that in a large class of 4D $\mathcal{N} = 1$ string compactifications any infinite distance limit yields a fundamental axion string whose tension scales with the mass of a tower of light particles as $T^w \sim m^2$, where $w = 1$, 2, or 3. Here we see that the case $w = 1$ corresponds to the case where the light particles are charged and a $\theta F \wedge F$ coupling is present. The large-radius limit of a Kaluza-Klein compactification of minimal 5D supergravity, where there is no such Chern-Simons coupling involving the KK photon, but there is one involving the 4D descendant of the 5D graviphoton and corresponding to the $w = 3$ case.

Thus far the argument we have sketched in this section is unique to four dimensions, as Eq. (109) does not have a well-known higher-dimensional parallel. However, supergravity constraints in higher dimensions impose similar relations, so the argument of this section does admit higher-dimensional analogs within the supergravity context. For further details, see Heidenreich, Reece, and Rudelius (2021) for the 5D case and Kaya and Rudelius (2022) for even higher-dimensional cases.

VI. ATTEMPTED DERIVATIONS OF THE WGC

A. The WGC from holography

The weak gravity conjecture is a proposed restriction on nonperturbative quantum gravity, and thus it is natural to ask whether we can show that it holds in the theories of nonperturbative quantum gravity we currently possess. In particular, we can ask whether the WGC holds within AdS/CFT, which is our best-understood set of quantum gravity theories. Thus far this has not been established, but a holographic argument for something closely related to the WGC was given by Montero (2019). In this section we sketch this argument and make a few related observations.

One of the main motivations for the WGC is the idea that near-extremal black holes in flat space should be unstable. In AdS/CFT that would be a statement about "small" black holes, whose size is small compared to the AdS radius, but small black holes are not well understood in AdS/CFT. Montero instead argued that large near-extremal black holes in AdS must be unstable, as otherwise the thermofield double state of the dual CFT at large chemical potential and small temperature would have rather surprising (and likely impossible) entropic properties. This argument does not amount to a proof of the WGC, as we later see that there are other ways these black holes could decay besides emitting charged particles obeying Eq. (5), but the argument is still suggestive, and we are optimistic that more could be learned from it.

Charged AdS black holes that are large compared to the AdS scale asymptotically become charged black branes, which in D bulk Euclidean dimensions have a gauge field

$$A_\tau = \frac{i\rho}{D-3}\left(\frac{1}{r^{D-3}} - \frac{1}{r_+^{D-3}}\right) \tag{111}$$

and a metric

$$ds^2 = f(r)d\tau^2 + \frac{dr^2}{f(r)} + r^2 d\vec{x}^2_{D-2}, \tag{112}$$

with

$$f(r) \equiv r^2 - \frac{2\kappa^2\epsilon}{(D-2)r^{D-3}} + \frac{\kappa^2\rho^2}{(D-2)(D-3)r^{2D-6}}. \tag{113}$$

In Eqs. (111)–(113) ϵ and ρ are the boundary energy and charge densities, respectively, r_+ is the largest positive zero of $f(r)$, and we have set the AdS radius to 1. If we work at fixed inverse temperature β and chemical potential μ, then we have

$$\begin{aligned} r_+ &= \frac{2\pi}{(D-1)\beta}\left(1+\sqrt{1+\frac{(D-1)(D-3)^2\kappa^2\beta^2\mu^2}{4\pi^2(D-2)}}\right),\\ \rho &= (D-3)r_+^{D-3}\mu,\\ \epsilon &= \frac{(D-2)r_+^{D-3}}{2\kappa^2}\left(r_+^2 + \frac{(D-3)\kappa^2\mu^2}{D-2}\right). \end{aligned} \tag{114}$$

This black brane approaches extremality when $\kappa\beta\mu \gg 1$, with the extremal radius given by

$$r_+|_{\beta=\infty} \equiv r_e = \frac{(D-3)\kappa|\mu|}{\sqrt{(D-1)(D-2)}}. \tag{115}$$

At extremality the function $f(r)$ has a double zero at $r = r_e$, so the radial geodesic distance from r_e to any large but finite radius r_c is logarithmically divergent.

The main point of Montero (2019) is that if we study the Hartle-Hawking state of two such extremal black branes obtained by slicing the Euclidean path integral, there is tension between two facts that apparently follow from the bulk picture together with the holographic correspondence.[11]

- *Exponential correlators*.—the fixed-time correlators of boundary operators decay exponentially with distance.
- *Volume-law entanglement*.—the union of a boundary subregion A_R in the right CFT and the same subregion A_L in the left CFT has a von Neumann entropy that grows like the volume of the subregion.

There are various ways to understand why the bulk picture implies these results. One way is based on the idea that black hole horizons are extremal surface barriers (Engelhardt and Wall, 2014). What this means is that an extremal surface of any codimension greater than or equal to 2 cannot be smoothly deformed from a surface that does not cross a horizon to a surface that does. The reason is simple: If such a deformation were possible, then at some point the surface would have to be tangent to the horizon; however, then the extremality equations would imply that the surface would be entirely contained in the horizon. In particular, for Euclidean states such as the Hartle-Hawking state, we can approximate the two-point function of as massive field as

$$\langle \mathcal{O}(x)\mathcal{O}(y)\rangle \sim e^{-m|x-y|}, \tag{116}$$

where $|x-y|$ is the geodesic distance between x and y. As shown in Fig. 9, the geodesic that is relevant for computing the correlator of two boundary fields necessarily has a length that grows like the boundary distance between the fields. The horizon at $r = r_e$ is an extremal surface barrier, so the geodesic has no choice but to involve a large extensive component that lies just outside the horizon. Moreover, any geodesics that cross the horizon must have infinite length, and therefore there is no correlation between boundary operators on opposite sides.

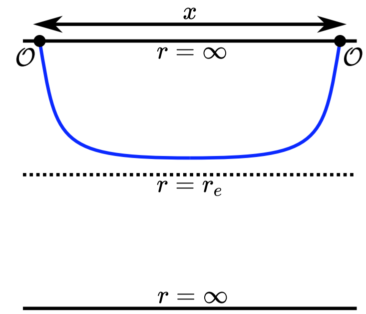

FIG. 9. Exponential decay of spatial correlators for the extremal black hole back brane.

[11]It has gradually been understood that the semiclassical picture of the bulk needs to be used with some care at low temperatures, as quantum effects eventually become important (Preskill et al., 1991; Maldacena, Michelson, and Strominger, 1999; Page, 2000Almheiri and Kang, 2016; Iliesiu and Turiaci, 2021; Heydeman et al., 2022). It would be worthwhile to revisit this argument from the point of view of the modern understanding of these quantum effects via a dimensional reduction to Jackiw-Teitelboim gravity, as they could potentially change the conclusions, but we will not attempt to do so here.

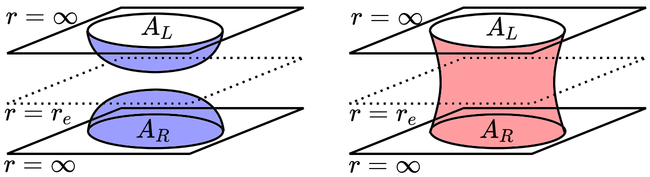

FIG. 10. Competing minimal surfaces to compute the von Neumann entropy of $A_L \cup A_R$ in the thermofield double state. At nonzero temperature the red area-law surface on the right always wins for large enough volume, but at zero temperature its area is infinite so the blue "volume-law" surface on the left prevails for any region size.

The volume-law entanglement can be understood along similar lines. We can compute the von Neumann entropy of the dual CFT on $A_L \cup A_R$ using the Ryu-Takayanagi formula, which tells us that it is given by the area of the minimal-area surface that is homologous to $A_L \cup A_R$. There are two possible candidates for the Ryu-Takayanagi surface (see Fig. 10), only one of which gives a volume law, and at finite temperature the "area-law surface" eventually wins for large enough regions. At zero temperature, however, the area-law surface always has infinite area due to the infinite distance to the horizon, so the "volume-law surface" always wins.

The reason that exponential decay of correlators and volume-law entanglement are in tension is that the former suggests only short-range entanglement is present, while the latter requires long-range entanglement. If the entropy of $A_R \cup A_L$ is growing like the volume, then since the total state is pure all this entanglement must be purified by something in the complementary region. Such a purification is unlikely given the exponential decay of correlation with distance. This intuition has been formalized in $1+1$ dimensions into a precise theorem (Hastings, 2007), and it plausibly holds in general.

If the tension just described indeed constitutes a contradiction, then the only way out is for the thermodynamic description of the extremal black brane to break down. There are two ways that this can happen. The first way is that there could be an exactly BPS particle, which turns out to lead to power-law correlators and thus removes the tension. This is the situation that is realized for BPS branes in supersymmetric theories. The second possibility, which thus far has been manifested for all nonsupersymmetric extremal branes, is that there is some kind of matter present that causes the brane to be unstable. Had we been discussing small black holes, such an instability would have immediately required the existence of superextremal charged particles and thus given a derivation of some version of the WGC. For large black holes, however, there are more possibilities for the instability. Indeed, this topic has a long history in the literature on applications of AdS/CFT to condensed matter physics, where the various possibilities go under the names holographic superconductor (Gubser, 2008; Hartnoll, Herzog, and Horowitz, 2008a, 2008b) or holographic Fermi surface (Hartnoll *et al.*, 2010; Faulkner *et al.*, 2011; Liu, McGreevy, and Vegh, 2011) depending on whether the particle causing the instability is a boson or a fermion. The rough idea for the bosonic case is as follows: in the near-horizon region the gauge kinetic term

$$-(\nabla^\mu \phi - inA^\mu \phi)^\dagger (\nabla_\mu \phi - inA_\mu \phi) - m^2 \phi^\dagger \phi \qquad (117)$$

for a boson of charge n leads to an effective mass

$$m_{\text{eff}}^2 = m^2 - \frac{n^2 e^2}{\kappa^2} \qquad (118)$$

in the near-horizon AdS_2 region. This leads to an instability if m_{eff}^2 violates the AdS_2 Breitenlohner-Freedman bound $m^2 > -(D-1)(D-2)/4$, so, in other words, there is an instability for masses in the range

$$-\frac{(D-1)^2}{4} < m^2 \le \frac{n^2 e^2}{\kappa^2} - \frac{(D-1)(D-2)}{4}. \qquad (119)$$

The first inequality in Eq. (119) is the D-dimensional Breitenlohner-Freedman bound, which is necessary for the vacuum to be stable. When $n \neq 0$ the first term on the right-hand side of Eq. (119) gives something like the WGC inequality, as first noticed by Denef and Hartnoll (2009), but it is missing the factor of γ. And, moreover, owing to the second term it is possible to have an instability even if $n = 0$, so the brane can be unstable even if there are no charged particles at all.[12] Thus, any argument that requires an instability only of large extremal black holes in AdS is not sufficient to imply the validity of the WGC, although it is certainly suggestive. Various other types of instability for this system have been discussed in the AdS/CMT literature, and the connection to the WGC was also discussed by Henriksson, Hoyos, and Jokela (2020).

B. The WGC from thermodynamics

1. WGC and quasinormal mode frequencies

An interesting argument linking the WGC to a bound on the imaginary part of the frequencies of black hole quasinormal modes was given by Hod (2017). The argument relies on the "universal relaxation bound" that was previously proposed by Hod (2007b). To derive this bound, Hod began with Bekenstein's bound on information transfer (Bekenstein, 1981a) [which itself was derived from Bekenstein's entropy bound (Bekenstein, 1981b)] and placed the following upper bound on the rate at which an observer can receive information:

$$\dot{I} \le \frac{\pi E}{\log 2}. \qquad (120)$$

In Eq. (120) I is the information and E is the energy of the package containing the information. Using

[12] In the absence of charged particles, it is not clear what the end point of this instability might be. It would need to be inhomogeneous since the homogeneous brane solution is unstable. The final end point would likely be theory dependent.

$$S = I \log 2, \quad \dot{S} = \frac{\Delta S}{\Delta \tau}, \quad T = \frac{\Delta E}{\Delta S} \tag{121}$$

one can rewrite the bound as

$$\Delta\tau \geq \frac{1}{\pi T}, \tag{122}$$

which Hod interpreted as a bound on the time $\Delta\tau$ over which a system can relax to equilibrium, deemed the universal relaxation bound. Finally, he applied this bound to the quasinormal modes of a black hole by setting T as the temperature of the black hole and $\Delta\tau$ as the inverse of the smallest imaginary part of a quasinormal mode frequency $\Delta\tau = \{\min[\mathrm{Im}(\omega)]\}^{-1}$, eventually arriving at the bound

$$\min[\mathrm{Im}(\omega)] \leq \pi T. \tag{123}$$

In response, note that while the derivation of Hod's bound (122) follows straightforwardly from Bekenstein's bound (120), its interpretation as a universal bound on relaxation times is more suspect. Bekenstein derived his bound by imagining a scenario in which one observer sends a package full of information to another. It is not clear how this scenario can be translated into the case of interest at hand, in which a black hole is perturbed and relaxes to equilibrium. A sharper derivation of the proposed universal relaxation bound, particularly in the context of quasinormal mode frequencies, is clearly desirable. However, note that Hod and others have given both numerical and analytical evidence in favor of the bound (123) (Gruzinov, 2007; Hod, 2007a, 2007b) and bounds similar to (122), but without the precise O(1) factors have been argued for in other contexts; see Lucas (2019) and references therein.

Assuming Eq. (123), Hod argued for the WGC as follows. None of the quasinormal modes arising from gravitational and electromagnetic perturbations of a nearly extremal Reissner-Nordström black hole obey the bound. However, if we assume that the bound merely requires that some mode in the black hole background obeys Eq. (123), then the bound could be satisfied by a quasinormal mode of a matter field. In particular, Hod showed analytically that, to leading order in T in the extremal limit, a charged scalar field has a quasinormal mode that obeys the bound (123) precisely when the scalar satisfies the WGC (Hod, 2017). A similar conclusion was also obtained analytically in the asymptotically $\mathrm{AdS}_2 \times S^2$ near-extremal, near-horizon limit (Urbano, 2018). These are interesting results that call for further study. Black holes that are far from extremality have modes that comfortably satisfy the bound. Near extremality, black hole quasinormal modes split into two families: damped modes, which have $\mathrm{Im}(\omega)$ of the order of the inverse black hole radius, and zero-damped modes (ZDMs), which have $\mathrm{Im}(\omega) \to 0$ as $T \to 0$ (Yang et al., 2013). The bound (123) can be obeyed only by a ZDM. Much of the literature on the numerical computation of quasinormal modes focuses on damped modes, whereas ZDMs are less well studied (Detweiler, 1980; Konoplya and Zhidenko, 2013; Yang et al., 2013; Richartz and Giugno, 2014). ZDMs for a Reissner-Nordström black hole have been found in the pure Einstein-Maxwell theory (Zimmerman and Mark, 2016), so the precise numerical coefficient in Eq. (123) is important for the link to the WGC. Future work could determine whether Eq. (123) is equivalent to the WGC away from the $T \to 0$ limit. It would also be of interest to explore this correspondence for more general black holes (for instance, dilatonic black holes). Strong numerical evidence for a WGC–quasinormal mode connection would provide the motivation for further study of the quasinormal mode relaxation bound itself.

The inequalities (122) and (123) bear a superficial resemblance to the well-studied chaos bound (Maldacena, Shenker, and Stanford, 2016), which requires a Lyapunov exponent $\lambda \leq 2\pi T$. However, the relaxation rate and the rate of growth of chaos are generally not the same. For example, the Sachdev-Ye-Kitaev model saturates the chaos bound, but its thermal two-point function falls exponentially with a timescale $\tau = q/2\pi T$ (Maldacena and Stanford, 2016), where q is a positive even integer. This is consistent with Eq. (122) but does not saturate the bound,[13] except when $q = 2$.

2. WGC and entropy

As discussed in Sec. II, one argument against continuous global symmetries is based on the existence of finite-mass black hole states of arbitrarily large global charge, leading to infinite entropy in a finite-size region, in violation of entropy bounds in quantum gravity. This has inspired studies relating the WGC to entropy bounds. For small but nonzero gauge coupling, a WGC-violating theory can have a large but finite number of stable extremal black holes in a finite-mass range. For example, Banks, Johnson, and Shomer (2006) suggested that the uncertainty in a measurement of the charge of a black hole is of the order of $1/e$, leading to an entropy scaling as $\log(1/e)$ and eventually violating entropy bounds for sufficiently small e. However, it is unclear why one would not be able to measure charge more precisely than $1/e$ (such as by measuring the motion of charged particles in the long-range electric field outside the black hole) or why the Bekenstein-Hawking entropy should be the relevant bound for an ensemble with such a large range of possible charges. Furthermore, examples in which exactly stable BPS charged black holes exist (with moduli spaces such that e can be made arbitrarily small) illustrate that the existence of many marginally stable species is not in itself in contradiction with quantum gravity.

Subsequent studies have examined logarithmic corrections to black hole entropy in the presence of WGC-violating matter (Fisher and Mogni, 2017; Shiu, Cottrell, and Soler, 2017; Shiu, Soler, and Cottrell, 2019). These corrections have interesting properties, but their computation has not led to an undisputed proof of the WGC. In particular, Andriolo et al. (2018) claimed that the argument of Fisher and Mogni (2017) relies on applying a formula outside its regime of validity.

C. WGC from corrections to large black holes

The extremality bound for black holes is derived from the two-derivative effective action. From the beginning, it was

[13] B. H. thanks Zachary Fisher and Ziqi Yan for discussions on this point.

understood that the WGC could potentially be satisfied by large black holes when higher-derivative corrections to the effective action are taken into account (Arkani-Hamed, Motl et al., 2007). Schematically, these modify the extremality bound to take the form $|Q|/M \geq (|Q|/M)|_{\mathrm{ext}}(1 + c/Q^2)$, where c is a linear combination of Wilson coefficients of four-derivative operators and $(|Q|/M)|_{\mathrm{ext}}$ is the charge-to-mass ratio of asymptotically large extremal black holes, which is computed with the two-derivative action. We continue to define a superextremal state as one for which $|Q|/M \geq (|Q|/M)|_{\mathrm{ext}}$, such that finite-size black holes are superextremal when $c \geq 0$. The corrections to the extremality bound from general four-derivative operators added to Einstein-Maxwell theory were calculated (Kats, Motl, and Padi, 2007), and it was found that certain black holes in heterotic string compactifications do in fact become superextremal.[14]

Before discussing the technical details, it is useful to describe two different interpretations that one might attach to the observation that small corrections to large black holes can allow them to become superextremal. The first is that this trivializes the weak gravity conjecture. The WGC in its most mild form requires that some state in the theory be superextremal. If this state is a large black hole, the WGC is simply a statement about the signs of some higher-dimension operators in the effective action, and does not imply the existence of any light charged particles below the Planck scale. If a general positivity proof can be constructed for the linear combination of operator coefficients appearing in the corrected extremality bound, the WGC will follow, and as such will be reduced to a statement about gravitational effective field theory. The second viewpoint is that the evidence we have for the WGC, as already discussed, favors the much stronger tower or sublattice WGCs, involving an infinite tower of charged particles of increasing charge and mass, all of which are superextremal. For large values of $|Q|$, the charged "single-particle states" are simply black holes, so the tower or sublattice WGCs require that they be superextremal (as depicted in Fig. 5). From this perspective, an EFT argument could explain superextremality far out in the charge lattice, but the tower or sublattice WGCs will also imply the existence of superextremal states at smaller Q, where the states are no longer well described as black holes in EFT. Arguments in favor of superextremality from higher-derivative corrections cannot decisively favor the former perspective (that EFT is everything) over the latter (that swampland constraints go beyond EFT). However, if we find consistent theories of quantum gravity (not just EFTs) in which large black holes are subextremal, this would immediately falsify the tower or sublattice WGCs.

[14]However, note that the results of Giddings, Polchinski, and Strominger (1993) and Natsuume (1994), upon which the work of Kats, Motl, and Padi (2007) relied, are obtained at string tree level. Because the string coupling diverges at the horizon of the black holes in question, the string loop expansion may not be under control; see also Cvetic and Tseytlin (1996).

1. The corrected extremality bound

There are several possible four-derivative operators that may be added to the Lagrangian of Einstein-Maxwell theory, built out of $R_{\mu\nu\rho\sigma}$ and $F_{\mu\nu}$. For this discussion, we work with the normalization $(1/2\kappa^2)R - (1/4)F_{\mu\nu}F^{\mu\nu}$ for the two-derivative Lagrangian. If we limit our attention to CP-conserving terms, there are four independent physical four-derivative terms. Their contribution to the effective action can be parametrized as

$$S_{4\partial} = \int d^D x\sqrt{-g}[c_{\mathrm{GB}} O_{\mathrm{GB}} + c_{RF} R_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} + c_T T_{\mu\nu}T^{\mu\nu} + c_F (F_{\mu\nu}F^{\mu\nu})^2], \tag{124}$$

where $T_{\mu\nu}$ is the two-derivative Maxwell stress tensor $F_{\mu\rho}F_{\nu}{}^{\rho} - (1/4)g_{\mu\nu}F_{\rho\sigma}F^{\rho\sigma}$, and $O_{\mathrm{GB}} = R^2 - 4R_{\mu\nu}R^{\mu\nu} + R^{\mu\nu\rho\sigma}R_{\mu\nu\rho\sigma}$ is the Gauss-Bonnet term (for $D = 4$, this is a topological term that does not affect the extremality bound). This is the basis favored by the discussion given by Arkani-Hamed et al. (2022). All other four-derivative operators can be related to these four terms (up to terms that are of higher order in the derivative expansion) via equations of motion (or, equivalently, field redefinitions). For example, terms involving the Maxwell stress tensor $T_{\mu\nu}$ can be traded for terms involving $R_{\mu\nu}$ using the Einstein equations, while terms involving $\nabla_\mu F_{\rho\sigma}$ can be transposed via integration by parts into terms that vanish in pure Einstein-Maxwell theory as well as terms involving a commutator of two covariant derivatives, which can be eliminated in favor of the Riemann tensor.

The condition for extremal Reissner-Nordström black holes to become strictly superextremal due to four-derivative terms is (Kats, Motl, and Padi, 2007)

$$(D-3)[(D-2)(Dc_T + 16c_F) + 8(D-3)c_{RF}\kappa^2] - 4(D-4)(3D-7)c_{\mathrm{GB}}\kappa^4 > 0. \tag{125}$$

For $D = 4$ this simplifies[15] to $c_T + 4c_F + c_{RF}\kappa^2 > 0$.

As discussed by Charles (2019), another convenient basis related to familiar anomalies is

$$\tilde{S}_{4\partial} = \int d^D x\sqrt{-g}[\tilde{c}_W W_{\mu\nu\rho\sigma}W^{\mu\nu\rho\sigma} + \tilde{c}_{RF} R_{\mu\nu\rho\sigma}F^{\mu\nu}F^{\rho\sigma} + \tilde{c}_{\mathrm{GB}} O_{\mathrm{GB}} + \tilde{c}_F (F_{\mu\nu}F^{\mu\nu})^2]. \tag{126}$$

In Eq. (126) $W_{\mu\nu\rho\sigma}$ is the Weyl tensor. The relationship between the bases (126) and (124) is

$$\tilde{c}_F = c_F + c_T\frac{(D-4)^2}{16(D-1)}, \qquad \tilde{c}_W = \frac{c_T}{\kappa^4}\frac{D-2}{4(D-3)},$$
$$\tilde{c}_{\mathrm{GB}} = c_{\mathrm{GB}} - \frac{c_T}{\kappa^4}\frac{D-2}{4(D-3)}, \qquad \tilde{c}_{RF} = c_{RF}. \tag{127}$$

[15]In the 4D case, stronger constraints can be obtained by considering dyonic black holes; see Etheredge and Heidenreich (2022).

A more detailed discussion of the field redefinitions that can convert between operator bases was given in Appendix B of the work of Cheung, Liu, and Remmen (2018).

Early work on this subject derived the condition (125) by directly solving the modified equations of motion in the presence of higher-derivative operators and extracting the corrected extremality bound from the perturbed black hole solution. Recently calculations have been streamlined by the discovery of formulas relating the change in the extremality bound to integrals evaluated on the uncorrected black hole solution. Specifically, the shift in the charge-to-mass ratio $\zeta = |Q|/\sqrt{\gamma}\kappa M$ of an extremal black hole away from 1 is given by

$$\Delta\zeta = \frac{1}{M}\lim_{\zeta\to 1}\int d^{D-1}x N\sqrt{h}\Delta\mathcal{L}\Big|_{\text{two-deriv}}, \qquad (128)$$

where $\Delta\mathcal{L}$ consists of the higher-derivative corrections to the leading-order Lagrangian, N and h are the lapse function and spatial metric associated with a fixed-time slice (extending from the horizon to infinity), and the subscript two-deriv signals that the expression is to be evaluated on the two-derivative solution.

Expressions of this form have been derived in multiple ways. One approach (used mainly for 4D Reissner-Nordström black holes) begins with the Wald entropy of the black hole (Wald, 1993), which is related through standard thermodynamic arguments to the Euclidean action evaluated on the solution. This receives corrections only from the action evaluated on the uncorrected solution since the evaluation of the uncorrected action on corrections to the solution vanishes at first order due to the extremality of the uncorrected action at an uncorrected solution (Cheung, Liu, and Remmen, 2018; Reall and Santos, 2019).

The correction to the extremal charge-to-mass ratio is then shown to be related to the change in the black hole entropy (Cheung, Liu, and Remmen, 2018). This has been generalized to rotating and dyonic black holes (Cheung, Liu, and Remmen, 2019), dilatonic black holes (Loges, Noumi, and Shiu, 2020b), AdS black holes (Cremonini et al., 2020), and dyonic Kaluza-Klein black holes (Cremonini et al., 2021); see Arkani-Hamed et al. (2022) for further discussion. In fact, the extremality-entropy relationship was proven by Goon and Penco (2020) using general thermodynamic considerations, which imply that when there is a minimal mass for a given charge $M > M_{\text{ext}}(\vec{Q})$ sensitive to a parameter ϵ (like the coefficient of a four-derivative operator),

$$\frac{\partial M_{\text{ext}}(\vec{Q},\epsilon)}{\partial\epsilon} = \lim_{M\to M_{\text{ext}}}\left[-T\left(\frac{\partial S(M,\vec{Q},\epsilon)}{\partial\epsilon}\right)\Big|_{M,\vec{Q}}\right], \qquad (129)$$

even outside the black hole context. Recently similar results have been derived using the Iyer-Wald covariant phase space formalism (Aalsma, 2022); see also Aalsma et al. (2021).

Note that it is crucial that the partial derivative on the right-hand side of Eq. (129) is evaluated at fixed mass, rather than at fixed temperature (Etheredge and Heidenreich, 2022; McPeak, 2022). Thus, the mass correction at fixed (zero) temperature, i.e., at extremality, is related to the entropy correction at fixed mass, which takes the black hole away from extremality (since the extremal mass is corrected). A more natural quantity is the extremal entropy correction, evaluated at fixed (zero) temperature. However, this is not related to the extremal mass correction, as has been noted for various explicit stringy black holes (both asymptotically flat and asymptotically AdS) (Charles and Larsen, 2016; Cano, Ortín, and Ramirez, 2020; Cano et al., 2020; Bobev et al., 2021). For example, in four dimensions the Gauss-Bonnet term is topological, and it contributes to the black hole entropy but does not affect the extremality bound. This is consistent with Eq. (129) since nonzero contributions to $\partial S/\partial\epsilon$ that are independent of temperature in the extremal limit make no contribution to the right-hand side of Eq. (129) due to the explicit T prefactor.

Recently Eq. (128) was obtained without reference to the Wald entropy via a direct attack on the equations of motion combined with some general reasoning about the Lorentz invariance of the Lagrangian. In this context, the formula was shown to hold for extremal black holes coupled to arbitrary moduli in any dimension (Etheredge and Heidenreich, 2022).

Similar techniques have been adapted to study not only extremality but also long-range forces to assess whether the repulsive force conjecture is satisfied by corrected black holes (Cremonini et al., 2022); see also Etheredge and Heidenreich (2022). The results suggest that the RFC may not be automatically satisfied by four-derivative corrections. However, they are obtained in EFT examples rather than explicit string theory compactifications, so further work should investigate whether these examples can be realized in a full quantum gravity setting (and hence provide a counterexample to the RFC for corrected black holes). A study of the effects of higher-derivative corrections on the force between dyonic strings was given by Ma, Pang, and Lü (2022).

2. Overview of arguments

The previously sketched thermodynamic arguments have provided an efficient tool for computing the correction to the extremal charge-to-mass ratio in a given EFT as a function of the Wilson coefficients of higher-dimension operators. Such calculations lead to superextremality conditions that take the form of positivity bounds like Eq. (125). A variety of attempts have been made to prove such bounds from general principles.

In many cases, the general structure of a theory implies that the c_T and c_F terms in Eq. (124), which involve only photons and not gravitons, give the dominant corrections to the extremality bound. When these coefficients are explicitly calculable, they are often positive. Indeed, in quantum field theory (without gravity), one can prove rigorous positivity bounds on the coefficients of four-derivative operators involving $F_{\mu\nu}$ (Adams et al., 2006; Cheung and Remmen, 2014a). In a regime where the dominant contributions to c_T and c_F arise from low-energy QFT effects that would persist in the $M_{\text{Pl}} \to \infty$ limit, this is sufficient to prove Eq. (125), as noted by Arkani-Hamed, Motl et al. (2007).

We are interested in gravitational theories, where completely general, rigorous positivity arguments are more elusive. In four dimensions, loop effects involving gravitons can provide dominant contributions to four-derivative operators in

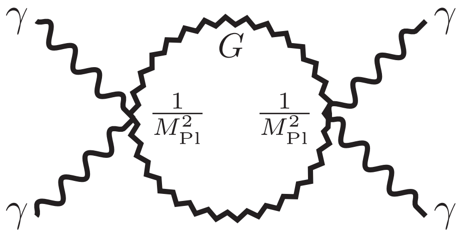

FIG. 11. Example of a loop diagram leading to logarithmic running of a four-derivative operator in four dimensions. Photons scatter via a loop of gravitons. Owing to two couplings each scaling as $1/M_{\rm Pl}^2$, this gives rise to a contribution that schematically behaves as $(1/M_{\rm Pl}^4)F_{\mu\nu}^4\log(E)$.

the IR, so taking the $M_{\rm Pl}\to\infty$ limit obscures important physics. In general, when the WGC is not satisfied by light charged particles but is parametrically saturated (or even violated) by all light particles, EFT proofs of Eq. (125) are more difficult to obtain. Next we summarize three broad categories of arguments for positivity: those that explicitly compute the coefficients in Eq. (124) within a given EFT; those that rely on analyticity, unitarity, and/or causality; and those based on entropy.

3. Explicit computations within low-energy EFTs

In four dimensions, the Wilson coefficients in Eq. (124) exhibit logarithmic renormalization group evolution. This follows from dimensional analysis; for example, $[c_T]=[c_F]=-D$ and $[\kappa^4]=4-2D$, which agree precisely when $D=4$. An explicit example of a loop diagram contributing to this running is shown in Fig. 11. For exponentially large black holes, we expect the Wilson coefficients (evaluated at a renormalization scale corresponding to the black hole's size) to be dominated by RG running. As a result, the sign of the correction should be determined by such RG effects, regardless of the details of the UV completion and the operator coefficients at the cutoff scale. The consequences were explored first by Charles (2019) and more recently by Arkani-Hamed *et al.* (2022). The case of multiple U(1)'s has also been considered (Jones and McPeak, 2020).

In the basis (126) there are logarithmic corrections to $\tilde{c}_W$ and $\tilde{c}_{\rm GB}$, which are determined by the well-known Weyl anomaly coefficients c and a, respectively (Charles, 2019). The coefficients $\tilde{c}_{RF}$ and $\tilde{c}_F$ do not run. In the basis (126), c_T and $c_{\rm GB}$ run (Arkani-Hamed *et al.*, 2022). In both bases, the running of $O_{\rm GB}$ is not relevant to the extremality bound. In Einstein-Maxwell theory plus any minimally coupled matter of spin $<3/2$, $c>0$, thus ensuring the validity of Eq. (125). Spin-$3/2$ fields contribute negatively to c, but a single spin-$3/2$ field is insufficient to drive the running negative. However, nonminimal couplings, such as dipole moments for fermions, also contribute negatively to c for a small range of Planck-suppressed couplings. In cases with $\mathcal{N}\geq 2$ supersymmetry where extremal black holes are BPS, these negative contributions precisely cancel positive ones such that the black hole extremality bound remains uncorrected. However, there are low-energy nonsupersymmetric effective Lagrangians with no obvious pathologies in which multiple fields with finely tuned nonminimal couplings could lead to a negative running for c, and hence to large black holes that cannot satisfy the WGC.

If the tower or sublattice WGC is true, then the corrections to large black holes must allow them to become superextremal. Thus, there must be a bound on the number of fermionic fields with dipole couplings in the limited range where the running of c is negative; such theories would lie in the swampland. In pure quantum field theory (QFT), negative coefficients of the F^4 operators would violate causality. However, precisely because the negative contributions arise from gravitational-strength interactions, there is no violation of causality in the gravitational context (Arkani-Hamed *et al.*, 2022).

Moving beyond log running in the deep IR, we can also consider the threshold corrections induced by integrating out specific massive particles. Neutral bosons coupling to $F_{\mu\nu}F^{\mu\nu}$ or $F_{\mu\nu}\tilde{F}^{\mu\nu}$, exchanged at tree level, generate positive four-derivative operator coefficients consistent with Eq. (125) (Hamada, Noumi, and Shiu, 2019). Loops of charged particles with a sufficiently large charge-to-mass ratio (obeying the WGC themselves by a safe enough margin) also satisfy Eq. (125) (Cheung and Remmen, 2014a). The challenging case, then, is when there are no light neutral bosons and all of the charged particles have $m\gtrsim eM_{\rm Pl}$. In that case, gravitational-strength ultraviolet contributions can be competitive and the sign is not obviously determined.

4. Arguments from analyticity, unitarity, and/or causality

In EFTs embedded within UV-complete quantum field theories, positivity bounds on certain combinations of operator coefficients (or, more invariantly, on derivatives of low-energy scattering amplitudes) may be proven using analyticity, unitarity, and causality (Pham and Truong, 1985; Adams *et al.*, 2006; de Rham *et al.*, 2017, 2018; Zhang and Zhou, 2020; Arkani-Hamed, Huang, and Huang, 2021). A prototypical example is the positivity of the $(\partial\phi)^4$ operator coefficient in the theory of a massive scalar field derived from a forward dispersion relation. This term contributes an $s^2+t^2+u^2$ term in the low-energy amplitude $\mathcal{A}(s,t)$ for $\phi\phi\to\phi\phi$ scattering. The coefficient of this term can be read off from a second derivative and in turn related to a contour integral in the complexified s plane using Cauchy's theorem,

$$\frac{1}{2}\mathcal{A}''(s_0,t=0)=\frac{1}{2\pi i}\oint_\gamma\frac{\mathcal{A}(s,0)}{(s-s_0)^3} =\frac{1}{\pi}\int_{\rm cuts}ds\,\frac{s\sigma_{\rm tot}(s)}{(s-s_0)^3}>0. \quad (130)$$

In the last step, the contour γ around s_0 has been deformed to enclose the s- and u-channel branch cuts and two large arcs at large s, as illustrated in Fig. 12. The integrals along the branch cuts in the $t\to 0$ limit are related to positive total cross sections using the optical theorem. The contour at infinity does not contribute, because the Froissart bound (in conjunction with a Phragmén-Lindelöf theorem) constrains the large-s amplitude to obey $\mathcal{A}(s,t=0)<s^2\log s$. This argument, which was given by Adams *et al.* (2006), can be extended to positive t (below the branch cut) (de Rham *et al.*, 2017, 2018). A version

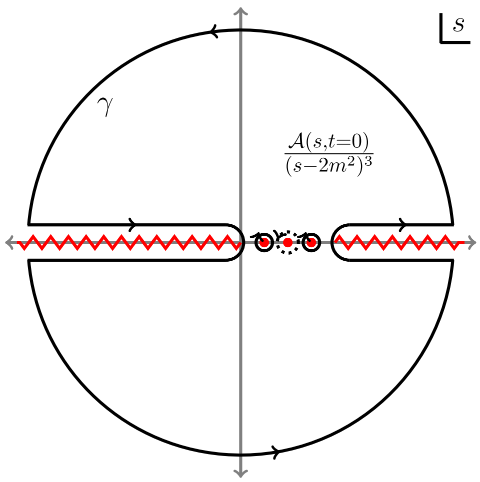

FIG. 12. Contour integral for a dispersive proof of positivity of four-derivative operators. Illustrated is the case of $2 \to 2$ scattering of a particle of mass m. The amplitude $\mathcal{A}(s, t = 0)$ has poles at $s = m^2, 3m^2$ and branch cuts at $s \le 0$ and $s \ge 4m^2$. The dashed contour around the singularity inserted at $s_0 = 2m^2$ can be deformed to the solid contour γ surrounding the subtractable pole contributions, the positive branch cut contributions, and a negligible contour at infinity.

of the argument can also be derived in AdS using CFT crossing relations (Hartman, Jain, and Kundu, 2016). Causality constraints arising from superluminal propagation in nontrivial field backgrounds lead to similar conclusions (Adams *et al.*, 2006). Notice that the positivity bound on the $(\partial\phi)^4$ coefficient is a strict inequality provided that ϕ is not free.

Corrected black holes satisfy the WGC if the inequality (125) holds. This inequality involves four-derivative operators that contribute to scattering amplitudes of photons and gravitons, so it is natural to seek a general argument similar to that for $(\partial\phi)^4$ that implies positivity independent of the details of the UV completion. For example, in a theory of only photons, the bounds derived from unitarity of forward scattering of linearly polarized photons (of all possible polarizations) imply

$$c_T \ge 0, \qquad Dc_T + 16c_F \ge 0 \tag{131}$$

in the notation of Eq. (124). However, the arguments immediately become more difficult in a gravitational context. The relationship between superluminality and causality is more subtle because light cones are not rigid; see the discussions given by Cheung and Remmen (2014a), Goon and Hinterbichler (2017), de Rham and Tolley (2020), and Bellazzini *et al.* (2022). Unitarity arguments based on forward dispersion relations face the difficulty that graviton exchange contributes a term $\propto -G_N s^2/t$ to scattering amplitudes, rendering the $t \to 0$ limit ill defined. Furthermore, high-energy scattering in gravitational theories can produce large black holes, so QFT bounds on the asymptotic UV behavior of amplitudes do not necessarily hold. The fact that the $1/t$ graviton-exchange pole scales as s^2 poses a particular problem for bounding four-derivative operators. For example, as shown by Bellazzini, Cheung, and Remmen (2016), if one carries out a contour integral to read off $O(s^4)$ coefficients and then sends $t \to 0$, one obtains candidate positivity constraints on operators involving four Riemann tensors that are compatible with known string theory examples. On the other hand, one cannot isolate the $O(s^2)$ contributions from local operators from those of graviton exchange in this way. Furthermore, if one simply discards the $-G_N s^2/t$ term and follows the logic of the unitarity bound, one would conclude that (in $D > 4$, where it affects $2 \to 2$ graviton scattering) the coefficient of the Gauss-Bonnet term must be both ≥ 0 and ≤ 0 (Bellazzini, Cheung, and Remmen, 2016). Theories are known in which this coefficient is nonzero, so it is clear that discarding the t-channel pole is not a strictly correct procedure. A plausible interpretation of this result is that the coefficient of the Gauss-Bonnet term cannot be too large with either sign, as further argued by Camanho *et al.* (2016) on causality grounds.

In QFT, we can deform a theory by adding relevant operators without changing the UV behavior. This provides a method for addressing problematic IR divergences. In quantum gravity, we do not have this luxury. Quantum gravity theories are rigid: we cannot simply add terms to the Lagrangian without modifying the entire theory. On the other hand, we can study a consistent theory on different backgrounds. This motivated a novel argument aiming to eliminate the problematic s^2/t pole by compactifying it to three spacetime dimensions, where there is no propagating graviton mode (Bellazzini, Lewandowski, and Serra, 2019). A subtlety is that, in resolving the IR problem of gravity, a new UV problem arises: 3D flat-space gravity does not admit localized states of an arbitrarily high mass, because a massive particle has a deficit angle that eventually eats up the entire space. In other words, the physics of the 3D theory resembles that of the 4D theory over a range of high energies but strongly deviates at truly asymptotic energies. Thus, the meaning of $\mathcal{A}(s, t)$ becomes obscure in high-energy regions, where it seems not even to be well defined, much less analytic. This was suggested by Alberte *et al.* (2020, 2021) as a possible culprit behind their observation that the t-channel subtracted positivity bounds derived from compactification appear to be overly strong. They require new physics to appear at prematurely small energies, in contradiction to known consistent theories. These works, reinforcing previous similar arguments made by Hamada, Noumi, and Shiu (2019) [see also Tokuda, Aoki, and Hirano (2020)], suggest that positivity arguments can forbid terms of the form $-c^2 s^2/M^4$, with $c \sim \mathrm{O}(1)$ and M held fixed in the limit $M_{\mathrm{Pl}} \to \infty$, but not terms of the form $-c^2 s^2/M^2 M_{\mathrm{Pl}}^2$, which tend to zero when gravity is decoupled.

Given the subtleties associated with making completely general and rigorous arguments in gravitational theories, much of the work on this subject has focused attention on identifying a sufficient set of conditions to prove Eq. (125). As discussed, explicit computations show that tree-level exchange of light bosons interacting with photons and loops of light charged particles both produce corrections to c_T and c_F that satisfy Eq. (125). In these cases, effects from c_{RF} and c_{GB} are subdominant. This is often the case, as large contributions to c_{RF} or c_{GB} induce causality violation in the absence of a tower of high-spin states (Camanho *et al.*, 2016; see also Li, Meltzer, and Poland (2017) and Afkhami-Jeddi, Kundu, and Tajdini (2018) for holographic, CFT-based arguments. The most difficult case to assess is when all contributions to the four-derivative operators arise from an ultraviolet scale like the string scale. In this case, additional assumptions have been invoked. If Regge states associated with the photon have effects dominating over those associated with the graviton, Eq. (125) can again be derived (Hamada, Noumi, and Shiu, 2019). Similar arguments were explored for dilatonic black holes by Loges, Noumi, and Shiu (2020b). Constraints from duality have also been shown to imply positivity conditions (Andriolo et al., 2020; Loges, Noumi, and Shiu, 2020a).

Recently new positivity bounds have been derived (Caron-Huot et al., 2021) that, following Camanho et al. (2016), avoid the t-channel pole problem by studying scattering at a fixed impact parameter rather than a fixed t. It remains to be seen whether such an approach can offer a new perspective on the WGC. A crucial test of any completely general future proof of a positivity bound is that it must be compatible with exactly zero correction in the case of BPS black holes.

5. Arguments from entropy

As discussed in Sec. VI.C.1, recent work has shown that the higher-derivative correction to the black hole extremality bound is related in a general way to the shift in the Wald entropy of the black hole due to higher-derivative terms. This raises the interesting prospect of proving Eq. (125) by proving that such corrections to the entropy must be positive (Cheung, Liu, and Remmen, 2018). In particular, an argument based on the Euclidean path integral for black holes with positive specific heat (including Reissner-Nordström black holes of sufficiently large charge) establishes that the correction to the Wald entropy $\Delta S_{4\partial}$ from four-derivative operators is positive whenever the correction $\Delta F_{4\partial}$ to the free energy of the black hole at fixed temperature is negative. The restriction to positive specific heat allows one to conclude that the classical solution minimizes (not just extremizes) the Euclidean action. Under these conditions, one can show that any four-derivative operators generated at tree level lead to $\Delta S_{4\partial} > 0$, which in turn implies that corrected black holes satisfy the WGC. Extending these considerations to rotating dyonic black holes leads to a range of inequalities generalizing Eq. (125) (Cheung, Liu, and Remmen, 2019).

The assumptions in this argument are not universally valid, even for tree-level exchange (Hamada, Noumi, and Shiu, 2019). For example, a massive spin-2 field $h_{\mu\nu}$ with a coupling $h^\lambda_\lambda F_{\mu\nu}F^{\mu\nu}$ generates a negative shift in the entropy. It evades the assumptions because the Euclidean action is not a local minimum with respect to h^λ_λ. Although this example evades the entropy argument, it violates unitarity and thus cannot be embedded in a consistent quantum gravity theory to provide a counterexample to Eq. (125). This connection between unitarity and positive contributions to the Wald entropy may hold more generally and hints at an argument that could extend beyond tree level (Cheung, Liu, and Remmen, 2018). Modular invariance is another supplementary assumption that has been invoked to extend the range of validity of entropy arguments for positivity (Aalsma, Cole, and Shiu, 2019).

VII. IMPLICATIONS AND CONNECTIONS

A. Implications for phenomenology and cosmology

1. Direct application of the WGC

Neither the WGC nor the toweror sublattice WGCs have immediate novel implications for the standard model of particle physics. The electromagnetic coupling constant at low energies is $e = \sqrt{4\pi\alpha} \approx 0.30$, so the electron satisifies the WGC by more than 20 orders of magnitude. Furthermore, because e is an order-1 number, the charged particles in the tower predicted by the tower or sublattice WGCs could all have masses near or above the Planck scale. If one applies the WGC to the non-Abelian gauge groups of the standard model (above the QCD scale or the electroweak scale, such that the gauge bosons appear massive), then the gauge bosons themselves obey the WGC, and again a WGC tower could consistently lie around the Planck scale because the coupling constants are of order 1. Perhaps a more interesting statement is that the WGC implies that a magnetic monopole should exist with a mass near the Planck scale or below (assuming that the bound is not obeyed only by monopoles of large magnetic charge), but this is not a statement that is readily falsifiable by any conceivable experiment at this time.

Interesting direct applications of the WGC, then, should be sought in new gauge interactions beyond the standard model. These could be previously undetected forces through which known particles interact or hidden sector interactions among particles that are currently unknown (or perhaps detected only indirectly through their gravitational effects in the form of dark matter).

Given the minimal standard model matter content (without right-handed neutrinos), the theory can be extended with a single additional $U(1)$ gauge interaction, coupling to one of the differences of lepton numbers for different generations: $L_e - L_\mu$, $L_\mu - L_\tau$, or $L_e - L_\tau$. At most one of these symmetries can be consistently gauged, due to mixed 't Hooft anomalies (Foot, 1991; He et al., 1991). The case of $L_\mu - L_\tau$ is of particular interest, as it can explain the nearly maximal mixing of muon and tau neutrinos (Ma, Roy, and Roy, 2002). Any of these gauge symmetries must be spontaneously broken. The regime of greatest phenomenological interest involves relatively large gauge couplings, where the WGC has little power, even assuming that it applies in the Higgs phase.

A more compelling example is the standard model with Dirac neutrino masses, which admits a different $U(1)$ extension, gauging the difference $B - L$ of baryon and lepton number (Chanowitz, Ellis, and Gaillard, 1977; Deshpande and Iskandar, 1980). In this case, the associated gauge field could be exactly massless without contradicting the experimental results provided that it is extraordinarily weakly coupled: $e_{B-L} \lesssim 10^{-24}$ (Wagner et al., 2012; Heeck, 2014). The combination of the Planck constraint on the sum of neutrino masses $\sum m_\nu < 0.12$ eV (Aghanim et al., 2020) with the values of the neutrino mass-squared differences inferred from neutrino oscillations $\Delta m^2_{21} \ll |\Delta m^2_{31}| \approx 2.4 \times 10^{-3}$ eV^2 (Zyla et al., 2020) implies that the lightest neutrino has a mass $\lesssim$0.03 eV. Thus, the lightest neutrino will obey the WGC for $B - L$ provided that $e_{B-L} \gtrsim 9 \times 10^{-30}$. This provides about 5 orders of magnitude in allowed $B-L$ coupling in which the mild form of the WGC would be satisfied. The tower or sublattice WGCs, however, provide a significant constraint: an infinite tower of $(B-L)$-charged particles should exist, beginning at masses of order $e_{B-L}M_{\rm Pl} \lesssim 1$ keV and extending up indefinitely. This implies that if $B-L$ is an unbroken gauge symmetry in our Universe, then billions of undetected particles that interact (albeit weakly) with ordinary matter exist below the TeV scale. Although this would be surprising, it is not obviously ruled out by data; it would have phenomenology akin to the large extra dimension scenario (Arkani-Hamed, Dimopoulos, and Dvali, 1998). A minimal WGC tower of $(B-L)$-charged particles would suggest a breakdown of local quantum field theory at energies $\lesssim e_{B-L}^{1/3}M_{\rm Pl} \lesssim 10^{10}$ GeV, a scaling analogous to that of Kaluza-Klein theory. (However, because standard model fermions carry $B-L$ charge and are not accompanied by low-mass excitations of higher $B-L$ charge, we would not expect the $B-L$ gauge group to literally arise as a Kaluza-Klein gauge field from a circle compactification.) The tower or sublattice WGCs put the existence of a massless $B-L$ gauge field in tension with conventional models of grand unified theories or of high-scale inflation, which postulate local new physics at energy scales above 10^{10} GeV but are not ruled out by experimental data.

The WGC might also be applied to possible gauge forces in hidden sectors, possibly related to the dark matter in our Universe. One might expect that forces weak enough to have significant WGC constraints would also be too weak to have observable consequences. Somewhat surprisingly, it turns out that weak forces between dark matter particles can sometimes have observable consequences in astrophysics or cosmology. Dark matter charged under a massless Abelian gauge field (or "dark photon") has been considered as a simple QFT with rich phenomenology (Feng, Tu, and Yu, 2008; Ackerman *et al.*, 2009; Feng *et al.*, 2009). Constraints on the strength of such a coupling arise from evidence that dark matter is approximately collisionless. However, even for small couplings there can be collective dark plasma effects (Ackerman *et al.*, 2009; Heikinheimo *et al.*, 2015). These lead to density fluctuations in the plasma on a timescale on the order of the inverse plasma frequency $\omega_p^{-1} \sim (m_d/e_d)\rho^{-1/2}$, where m_d is the mass of an individual dark matter particle, e_d is the dark photon coupling, and ρ is the mass density of dark matter (which is directly inferred from observations). If we suppose that the dark matter particles themselves obey the WGC for the dark U(1), this can lead to interesting consequences, as discussed by Craig, Garcia, and Koren (2019b). In this case, the dark WGC implies that $m_d/e_d \lesssim M_{\rm Pl}$. Quantitative estimates show that dark plasma fluctuations can then lead to shock waves developing on the timescale of a merger of colliding galaxy clusters. Thus, it is conceivable that observations of cluster mergers could reveal dynamical evidence of weak gauge forces between dark matter particles that approximately saturate the WGC. If the dark matter particles are sufficiently light, then the tower or sublattice WGCs could in turn imply important constraints on the UV cutoff of physics in our Universe. Dedicated work, including numerical simulations, would be necessary to make more precise statements about observable dark plasma effects.

The tower or sublattice WGCs can also have interesting implications for non-Abelian gauge groups in the dark sector. For example, dark matter charged under such a gauge group can have distinctive cosmological signatures even for weak couplings because the dark gluons constitute a form of interacting dark radiation (Buen-Abad, Marques-Tavares, and Schmaltz, 2015). The toweror sublattice WGC cutoff on such theories is at most $g^{1/2}M_{\rm Pl}$ (Heidenreich, Reece, and Rudelius, 2018b). Thus, there can potentially be a tension between cosmological observables associated with interacting dark radiation and theories of high-scale inflation.

Another topic of substantial recent phenomenological interest has been kinetic mixing between a dark U(1) and ordinary electromagnetism (Holdom, 1986). Such a mixing can be generated by loops of particles that carry both kinds of U(1) charge. The tower or sublattice WGCs imply the existence of such particles and hence suggest a minimum kinetic mixing, at least in the absence of gauged charge conjugation symmetries that enforce an exact cancellation. The size of kinetic mixing required for such an argument was recently explored and compared to concrete string theory examples by Benakli, Branchina, and Lafforgue-Marmet (2020) and Obied and Parikh (2021).

Finally, a direct application of the magnetic WGC that is only indirectly relevant for phenomenology was made by Cribiori, Dall'Agata, and Farakos (2021) and Dall'Agata *et al.* (2021), who argued that de Sitter critical points in certain gauged supergravity models are incompatible with the magnetic WGC, since by Eq. (17) their associated Hubble scale is larger than the scale of new physics: $\Lambda_{\rm NP} \lesssim eM_{\rm Pl} \lesssim H$.

2. Bounding the electroweak hierarchy

A long-standing challenge in particle physics is the electroweak hierarchy problem: Why is the electroweak energy scale ($v \approx 246$ GeV) so many orders of magnitude below the Planck scale ($M_{\rm Pl} \approx 2.4 \times 10^{18}$ GeV)? An enormous hierarchy between the Planck scale and the masses of the electron, proton, and neutron is necessary in order to have stable, large objects like stars and planets. However, in the standard model the electroweak hierarchy is large in only a small subset of the UV parameter space, unlike the hierarchy between the Planck and QCD scales, which is naturally exponentially large due to asymptotic freedom. This has led to a number of suggested extensions of the standard model in which the electroweak hierarchy can naturally become large, ranging from scenarios where electroweak breaking is triggered by dynamical supersymmetry breaking to those where the Higgs boson is a composite particle of a strongly interacting sector. Traditionally these models all share the feature that they relate the electroweak hierarchy to a scale generated by dimensional transmutation, and they predict new particles with masses near the electroweak scale. LHC measurements have informed us that the Higgs boson appears to be approximately elementary (i.e., it has standard-model-like interactions with other fields), and additional electroweak-scale particles have not yet been discovered. This has motivated theorists to pursue novel explanations of the electroweak hierarchy problem.

Because the WGC gives rise to an upper bound on particle masses, it is tempting to wonder if it could produce an upper bound on the Higgs scale v, thereby explaining why v is so small compared to $M_{\rm Pl}$. This idea was first discussed by Cheung and Remmen (2014b) and was further explored by Lüst and Palti (2018) and Craig, Garcia, and Koren (2019b). The simplest original version of the idea is to suppose that $B-L$ is gauged and that neutrinos, which acquire mass only from electroweak symmetry breaking, are the particles responsible for satisfying the WGC. The Dirac neutrino mass must then obey $m_\nu = y_\nu v/\sqrt{2} < \sqrt{2} e_{B-L} M_{\rm Pl}$. If we fix y_ν to its standard model value ($\sim 10^{-12}$) and we postulate a $B-L$ gauge coupling $e_{B-L} \sim 10^{-28}$ (consistent with the experimental limits), this inequality tells us that $v \lesssim 10^{-16} M_{\rm Pl}$ and thus requires an electroweak hierarchy of the order that we observe in nature.

While this offers an interesting perspective on how quantum gravity might affect low-energy particle physics in surprising ways, several elements of this argument are unsatisfactory. One is that it seeks to explain the origin of a mysterious factor of the order of 10^{-16} in terms of another small number of the order of 10^{-28}, which is unexplained. This is viewed as progress because the electroweak hierarchy is not robust against quantum corrections [the Higgs mass acquires additive corrections of the order of $(h^2/16\pi^2)M^2$ when coupled to heavy fields of mass M via interactions of size h, which must be "tuned away" through cancellations against other contributions], whereas the smallness of the gauge coupling e_{B-L} is "technically natural" (its corrections are all proportional to e_{B-L} itself). Nonetheless, if our goal is to understand the origin of small numbers in our theory of nature, this at best shifts the problem to explaining the origin of the small number e_{B-L}. One might hope that such a problem has a solution, for instance, in terms of a dynamical mechanism of moduli stabilization. Nonetheless, this shift of the hierarchy problem toward a problem of explaining an exponentially small e_{B-L} is in some tension with the spirit of the WGC itself. A small value of e_{B-L} restores a global symmetry of the theory, so quantum gravity should resist attempts to generate exponentially small gauge couplings. This suggests that perspectives rooted in a literal interpretation of technical naturalness may encounter obstacles in a quantum gravity setting. A sharper version of this concern is that the magnetic WGC tells us that $e_{B-L} M_{\rm Pl}$ serves as an ultraviolet cutoff on our EFT. This is particularly problematic from the viewpoint of the tower or sublattice WGCs, which postulate a tower of $(B-L)$-charged particles appearing at this mass scale. If infinitely many particles in such a tower obey the WGC, then it was unnecessary to require that the neutrinos obey the WGC, thus destroying the link between a small e_{B-L} and the Higgs scale v.

One refinement of the argument (Lüst and Palti, 2018; Craig, Garcia, and Koren, 2019b) drew on the repulsive force conjecture in the presence of a scalar field (Palti, 2017), arguing that the bound assumes the schematic form $m \le \sqrt{g^2 - \mu^2} M_{\rm Pl}$, where g is a gauge coupling and μ is a coupling to scalars. In cases where $g^2 \approx \mu^2$, this can be a much stronger bound than simply $m \le g M_{\rm Pl}$. This offers the opportunity to push the magnetic WGC scale $g M_{\rm Pl}$ up to higher energies, where it has less effect on the argument. On the other hand, it introduces yet another small number, the ratio $\sqrt{g^2 - \mu^2}/g$, which requires explanation. One must postulate a specific form of the scalar couplings of the light, WGC-obeying matter fields in order to make this argument. It should be different from the scalar couplings of black holes; otherwise, the tower or sublattice WGC tower would begin at the same scale $\sqrt{g^2 - \mu^2} M_{\rm Pl}$ rather than at $g M_{\rm Pl}$. This application of the RFC requires the scalar coupling providing the additional force to remain light, which itself requires explanation and can lead to additional naturalness constraints on the EFT (Craig, Garcia, and Koren, 2019b).

The most recent variations on the argument (Craig, Garcia, and Koren, 2019b) have explored new forces under which no standard model particle is charged. One could consider a scalar field Φ charged under a new $\mathrm{U}(1)_X$ gauge interaction, which should satisfy the WGC for $\mathrm{U}(1)_X$. If we further posit that Φ couples to the Higgs boson through a quartic coupling $\kappa |\Phi|^2 |h|^2$, then the additive shift of the Φ mass squared by $(1/2)\kappa v^2$ could cause Φ to fail to obey the WGC if v is too large. Similar models can be constructed with fermionic fields. These models make distinctive phenomenological predictions relative to the original $B-L$ model and could have implications for dark matter dynamics.

These attempts to bound the electroweak scale v using weak gravity arguments all invoke a similar set of assumptions. We must assume the existence of small (but technically natural) couplings. We also assume that specific particles in the theory, which happen to interact with the Higgs boson, are the ones that satisfy the WGC. If the WGC were satisfied by an independent set of particles, not interacting with the Higgs, then the link to the electroweak hierarchy would be severed. Finally, we suppose that this restricted set of theories is relevant for the world that we live in. If the landscape of quantum gravity contains many universes resembling our own that do not contain the postulated $\mathrm{U}(1)_{B-L}$ or $\mathrm{U}(1)_X$ force and the specific connections assumed between these forces and the electroweak scale, then there is no reason why we would expect our Universe to obey the assumptions. The argument that the WGC constrains the electroweak scale would be plausible only if such vacua were overwhelmingly more common than others, or overwhelmingly more likely to be populated by cosmology. In recent years, there has been a proliferation of models that link cosmology to particle physics by postulating the existence of a landscape that takes a specific form, where, for instance, certain couplings are assumed to exist and take on fixed values in all vacua. Only a limited set of parameters "scan" from one vacuum to another. These have been referred to as artificial landscapes (Strassler, 2014, 2016) and, in the absence of evidence that they resemble the true landscape of quantum gravity, it is unclear what lessons one can draw from them.

Finally, we emphasize that the examples in which we have checks of the WGC are cases where we compute the mass at leading order in a perturbative expansion, or where the mass is protected by supersymmetry. As a result, we have no explicit examples in which the WGC is satisfied by a state whose mass is fine-tuned to be light due to a cancellation. Indeed, such examples would be extremely difficult to generate. If one could find such examples in the string theory landscape, they would at least serve as an interesting proof of principle that the WGC could require a fine-tuning that would appear accidental from the viewpoint of low-energy effective field theory.

3. Other applications to the hierarchy problem

Graham, Kaplan, and Rajendran (2015) proposed a dynamical mechanism known as cosmological relaxation as a solution to the hierarchy problem. In this scenario, the Higgs field h is coupled to a real scalar field ϕ through a potential of the form

$$V = (-M^2 + g\phi)|h|^2 + (gM^2\phi + g^2\phi^2 + \cdots) + \Lambda^4 \cos(\phi/f), \tag{132}$$

where M is the cutoff of the effective field theory and Λ depends on the VEV of h. Initially, the dynamics of ϕ are dominated by the polynomial terms and the cosine term is negligible. When $\phi \sim M^2/g$, however, the Higgs field acquires a VEV and the scale Λ for the cosine terms grows, creating a barrier that stabilizes the axion and leaves the Higgs with a mass well below the EFT cutoff M. For this mechanism to work, however, the cosine terms must eventually be able to compete with the $gM^2\phi$ term. In typical relaxation scenarios, this requires g to be roughly of the order of 10^{-34}. Furthermore, inflation must last long enough for ϕ to scan the entire range of the Higgs mass. This places an additional bound on the cutoff given by $M \lesssim (\Lambda M_{\rm Pl})^{1/2}$, which yields $M \lesssim 10^9$ GeV for $\Lambda = \Lambda_{\rm QCD}$.

The small coupling $g \sim 10^{-34}$ is "technically natural," but this does not necessarily mean that the model can be UV complete. In particular, as we take $g \to 0$ the theory (132) has an exact global symmetry $\phi' = \phi + 2\pi f$. If the arguments against exact global symmetries have any robustness, they should also rule out sufficiently small values of g. One possible approach to avoiding this problem is to view $\phi' = \phi + 2\pi f$ as a gauge symmetry or, in other words, to turn ϕ into an axion. This, however, forbids most of the terms in Eq. (132) (a small explicit violation of a gauge symmetry is just as bad as a large one), and thus kills the feasibility of the model.

To date the most promising proposal for obtaining a large scalar field excursion that is consistent with all versions of the WGC is the "axion monodromy" proposal of Silverstein and Westphal (2008) and McAllister, Silverstein, and Westphal (2010).[16] The most basic version of this proposal (Kaloper and Sorbo, 2009) uses an axion coupled to a 3-form gauge field A_3 via the Lagrangian (which we here write as a 4-form),

$$L = -\frac{1}{2} d\phi \wedge \star d\phi - \frac{1}{2e_3^2} F_4 \wedge \star F_4 + \frac{g}{e_3} \phi F_4, \tag{133}$$

[16]This was originally proposed as a model of inflation, as we discuss later in this section, but it can also be used as a mechanism to implement cosmological relaxation, as we discuss here.

with $F_4 = dA_3$. Naively one might think that the coupling g should be zero to respect the axion periodicity $\phi \sim \phi + 2\pi f$, but as is usual for Chern-Simons type interactions the fact that the integral of F_4 obeys the quantization

$$\int F_4 = 2\pi m, \qquad m \in \mathbb{Z}, \tag{134}$$

means that it is enough that we have

$$g = \frac{k e_3}{2\pi f}, \qquad k \in \mathbb{Z}. \tag{135}$$

We have normalized g here so that it matches the g in Eq. (132), so we now have two ways to get a small g: either we can take f large in Planck units or we can take e_3 small. To avoid trouble with the axion WGC (26), we do not want to take f large in Planck units (we discuss this further in Sec. VII.A.5), so our task is to understand how constrained we are by the WGC for the 3-form gauge field A_3. Before discussing this, however, we explain in more detail how the theory (133) allows for a super-Planckian field excursion. The equations of motion following from Eq. (133) are

$$\star d \star d\phi + \frac{g}{e_3} \star F_4 = 0, \qquad d(g e_3 \phi - \star F_4) = 0, \tag{136}$$

so the quantity

$$\tilde{F}_0 \equiv \star F_4 - g e_3 \phi \tag{137}$$

is constant. In fact, it is quantized: the integral of A_3 over space is a periodic variable, and $\tilde{F}_0$ is proportional to its canonical conjugate. Working this out gives the quantization

$$\frac{1}{e_3^2} \tilde{F}_0 = n, \tag{138}$$

with $n \in \mathbb{Z}$. Substituting this back into the first equation of motion we find that

$$\star d \star d\phi + g(e_3 n + g\phi) = 0, \tag{139}$$

which is the equation of motion for a scalar field with a potential that is a second-order polynomial just as in Eq. (132). Equation (139) may appear to not be gauge invariant, but it actually is since n is a dynamical variable and the relevant gauge transformation is

$$\phi' = \phi + 2\pi f, \qquad n' = n - k. \tag{140}$$

In Eq. (140) k is the integer defined by Eq. (135). One way to think about the apparent nonperiodicity of the potential is to observe that although ϕ is periodic, $\star F_4$, which is not periodic, is also rolling in order to ensure that $\tilde{F}_0$ is constant. Indeed, one can say that $\star F_4$ is really the gauge-invariant scalar which is rolling in axion monodromy.

There is an interesting subtlety in this model that is worth mentioning explicitly: although the gauge invariance of the action prevents us from adding arbitrary powers of ϕ to the action, $\star F_4$ is perfectly gauge invariant, and thus there is nothing that prevents us from introducing a potential $V(\star F_4)$. Such a potential presumably is generated by quantum gravity effects, so why does it not ruin the model? To the extent that axion monodromy can be realized in string theory (which seems unlikely for the relaxion scenario but plausible for inflation), such a potential does exist but is typically of the form

$$V(\star F_4) = \frac{1}{e_3^2 \ell_s^8} v(\ell_s^4 \star F_4). \tag{141}$$

In Eq. (141) ℓ_s is the string scale and $v(\cdot)$ is a dimensionless function of a dimensionless variable that is expected to be $O(1)$. [In models, the form of this function is known from, for instance, the Dirac-Born-Infeld action (McAllister, Silverstein, and Westphal, 2010).] The correction to the equations of motion (136) arising from this potential does not become important until $\star F_4 \sim \ell_s^{-4}$. An axion excursion $\Delta\phi$ gives rise to a change in $\star F_4$ that is of the order of

$$\Delta \star F_4 \sim g e_3 \Delta\phi \sim k e_3^2 \frac{\Delta\phi}{f}, \tag{142}$$

so we can have an axion excursion that is large compared to f without feeling the potential V provided that

$$k e_3^2 \ell_s^4 \ll 1. \tag{143}$$

In string theory, the dimensionless number $e_3^2 \ell_s^4$ is often small: it can be proportional to positive powers of g_s, inverse powers of volumes, or warp factors; whatever the reason, as long as it is small we can achieve $\Delta\phi \gg f$ without being sensitive to $V(\star F_4)$. The robustness of axion monodromy thus relies on high-energy information from string theory: it cannot be established purely using low-energy power counting and symmetries.

We now turn to applying the WGC for 3-form gauge fields to axion monodromy (Ibáñez *et al.*, 2016). The objects to which it applies are domain walls of tension T_3, across which $(1/e_3^2)\tilde{F}_0$ changes by an integer Q, and the WGC says there should be such domain walls with

$$T_3 \leq \frac{4\pi f g M_{\rm Pl} Q}{k}. \tag{144}$$

The danger here is that an upper bound on the domain wall tension also likely gives some sort of lower bound on the rate for bubbles bounded by the domain wall to nucleate, and if this happens too often it destroys the relaxation mechanism. The domain walls separate regions whose potential energy differs by

$$\Delta V \sim g e_3 \phi \sim f g^2 \phi / k. \tag{145}$$

The bounce action computed by Ibáñez *et al.* (2016) is not accurately described by the thin-wall approximation but involves important gravitational backreaction (Coleman and Luccia, 1980). The result is a bubble nucleation probability

$$P \sim \exp(-B), \qquad B \approx w(b) \frac{2\pi^2 T_3}{H^3}, \tag{146}$$

where H is the Hubble scale during inflation. The parameter b is defined as

$$b = \frac{\Delta V}{H T_3}, \tag{147}$$

and it turns out that in the parameter range of interest $w(b) \sim O(1)$ and $b \lesssim 1$. Using b, we can rewrite the bounce action estimate as

$$B \sim 2\pi^2 \frac{T_3^4 b^3}{(\Delta V)^3}. \tag{148}$$

The WGC provides a constraint, following Ibáñez *et al.* (2016), because we require $B \gg 1$ for an exponentially suppressed tunneling probability, but the WGC implies that $B < B_{\rm max}$, where $B_{\rm max}$ is obtained when T_3 saturates Eq. (144). These can be consistent only when $B_{\rm max} \gg 1$. Together with the estimates $\phi \sim M^2/g$ and $gM^2 \sim \Lambda^4/f$ required for consistent relaxion phenomenology, this inequality translates into the following bound on the EFT scale:

$$M \lesssim (4\pi^2 b^3)^{1/8} \sqrt{\Lambda M_{\rm Pl}}. \tag{149}$$

For $\Lambda = \Lambda_{\rm QCD}$, this bound becomes

$$M \lesssim b^{3/8} \times 2.5 \times 10^9 \ {\rm GeV}, \tag{150}$$

which for $b \sim 1$ rivals the bound for consistency of the previously discussed relaxion model.

Thus, the 3-form WGC provides an interesting constraint on cosmological relaxation implemented via axion monodromy. On the other hand, a similar constraint may also be derived independently of the WGC, and in the original paper it was shown that this constraint could be satisfied without spoiling the model. Therefore, neither the axion WGC nor the 3-form WGC seem to pose a fatal challenge to the axion monodromy version of the cosmological relaxation model. Note, however, that embedding the model in string theory nonetheless seems to be challenging, if not impossible. In particular, McAllister *et al.* (2018) argued that within a string compactification the large winding number of the relaxion corresponds to a large charge carried by branes or fluxes [this is already apparent in the model (133) since $\star F_4$ is rolling]. This charge backreacts on the compactification geometry and eventually spoils the relaxation mechanism. The relaxion scenario may lie in the swampland, but if so the most stringent top-down constraints do not come from the WGC. Even if axion monodromy does not give a viable realization of the relaxion model in string theory, however, it is a plausible candidate for realizing inflation. We return to this in Sec. VII.A.5.

4. Mass of the photon or dark photons

Conventionally, we assume that the photon is a massless gauge field. However, theories of massive, Abelian spin-1 particles are perfectly consistent, either with a simple mass term and no gauge invariance at all (Proca, 1936) or with a real scalar field added to provide the longitudinal mode, together with a gauge invariance to eliminate the redundant degree of freedom (Stueckelberg, 1938). If the photon has a small mass, the longitudinal mode is extremely weakly coupled, so it is difficult to experimentally distinguish from a massless photon despite the change from two to three independent propagating polarization states (Bass and Schrödinger, 1955). The photon in our Universe must be extremely light. There is a large literature on experimental constraints on the mass, to which a few interesting entry points are Adelberger, Dvali, and Gruzinov (2007), Goldhaber and Nieto (2010), and Wu et al. (2016).

In four dimensions a massive photon in the Stueckelberg regime can be described by BF theory: we have a 1-form gauge field A with field strength $F = dA$ and a 2-form gauge field B with field strength $H = dB$ interacting via the Lagrangian

$$S = \int \left(-\frac{1}{2f^2} H \wedge \star H - \frac{1}{2e^2} F \wedge \star F + \frac{k}{2\pi} B \wedge F \right), \qquad (151)$$

where $k \in \mathbb{Z}$ just as in the axion monodromy discussion of Sec. VII.A.3. The gauge coupling f of the 2-form field has a mass of dimension 1. This theory describes a massive gauge field with a mass

$$m = \frac{k}{2\pi} e f. \qquad (152)$$

For $k \neq 0$, taking the gauge field mass to zero requires either $e \to 0$ or $f \to 0$. In either case, we are taking a gauge coupling to zero, so the WGC imposes some constraint. In particular, if we send $e \to 0$, the magnetic WGC tells us that there is a UV cutoff on the theory at the scale eM_{Pl}, and the tower or sublattice WGCs suggest that effective field theory breaks down irrevocably there at some scale, possibly a higher one like $e^{1/3} M_{\mathrm{Pl}}$ (as in Kaluza-Klein theory). If we send $f \to 0$, the mild WGC for the 2-form gauge field B implies that strings charged under B should exist with a tension $T \lesssim f M_{\mathrm{Pl}}$.

In the case where the photon mass arises from the Higgs mechanism, there is no fundamental obstruction to sending $f \to 0$. This corresponds to turning off the Higgs VEV, which can be accomplished just by giving the Higgs a positive mass term around the origin. In this case, the B field may be thought of as an emergent gauge field in the IR below the scale of the Higgs VEV, and the charged strings predicted by the 2-form WGC are simply Abrikosov-Nielsen-Olesen (ANO) strings (Abrikosov, 1957; Nielsen and Olesen, 1973). In the core of an ANO string, the Higgs VEV is zero; in the limit that the Higgs VEV is taken to zero, an ANO string simply becomes more and more diffuse and fades away. The WGC, then, is compatible with small masses arising from the Higgs mechanism.

By contrast, there are massive gauge theories that are fundamentally of the Stueckelberg type. In this case, the strings charged under the B field are fundamental (for instance, the F strings or D strings of string theory). The core of the string is not well described by effective field theory, and there is no finite-distance point in field space at which the gauge boson mass can be sent to zero. In this case, the limit $f \to 0$ corresponds to a theory of fundamental, tensionless strings, signaling a complete breakdown of local effective field theory. In such a case, the fundamental quantum gravity cutoff energy is bounded, $\Lambda_{\mathrm{QG}} \lesssim \sqrt{2\pi f M_{\mathrm{Pl}}}$.

The WGC, then, imposes an ultraviolet cutoff on theories of a massive gauge boson with mass arising from a fundamental Stueckelberg term (Reece, 2019). This can also be understood as a consequence of the swampland distance conjecture: for mass terms of a fundamental Stueckelberg type, the $m \to 0$ limit is an infinite distance limit, so an infinite tower of light states appears when one approaches this limit.

This constraint on massive, Abelian gauge bosons has potentially important implications for the standard model photon and for potential dark photons (Reece, 2019). Consider first the standard model photon. A conservative bound, obtained from the arrival time of different frequencies from fast radio bursts, is that $m_\gamma \lesssim 10^{-14}$ eV (Wu et al., 2016). Stronger bounds exist but involve more assumptions, so we work with this simple kinematic bound; our conclusions can be readily adapted to other constraints. If we assume that electromagnetic charge is quantized in the usual way, the only way to obtain a small standard model photon mass of a fundamental Stueckelberg type is by taking f to be small: $ef/(2\pi) \lesssim 10^{-14}$ eV requires $f \lesssim 10^{-22}$ GeV. But then the weak gravity conjecture would require fundamental strings with a tension $T \lesssim f M_{\mathrm{Pl}} \lesssim (20\ \mathrm{MeV})^2$. However, we know that gravity does not become strongly coupled near the MeV scale, so we cannot have a fundamental string with its associated tower of high-spin modes at such a low scale. This strongly suggests that the only way for the standard model photon to be massive is if it is Higgsed.

Could the standard model photon be Higgsed? The short answer is "probably not," and it becomes "no" provided that we assume that the ratios among charges of light particles in the theory are at most $\mathrm{O}(1)$ (such an assumption is common in discussions of phenomenological implications of the WGC). With this assumption, if the standard model photon were Higgsed, then we would already have discovered the corresponding Higgs boson. The only way to avoid this is for the standard model photon to obtain a mass from a Higgs field with a charge that is a small fraction of the electron's charge, in which case the associated Higgs boson could remain hidden from experiments. For example, suppose that the fundamental unit of electric charge is not e but rather some $e_0 = e/N$ where N is a large integer. Our calculation then becomes significantly different: we have $m_\gamma = e_0 f/2\pi$ and, rather than a small f, we are free to take a small e_0. Furthermore, if f is a Higgs VEV rather than a Stueckelberg scale, then there are no associated fundamental strings providing a UV cutoff. As an example, the choices of $f \sim \mathrm{eV}$ and $e_0 \sim 10^{-14}$ could be consistent with experimental bounds on millicharged particles. This implies a WGC tower of states with small electric charge beginning at the scale $e_0 M_{\mathrm{Pl}} \sim 10$ TeV, which is allowed by data. Apart from a new hierarchy puzzle associated with the small mass of the new Higgs field, the cost of considering such a theory is the introduction of the enormous integer $N \sim 10^{14}$. The standard model fermions would carry electric charge on this order in units of the fundamental charge. There are no known consistent theories of quantum gravity that can produce such large ratios of charges among light particles. On the other hand, there are phenomenological models in which such a large integer could be obtained as a product of smaller integers, as in the clockwork scenario (Choi, Kim, and Yun, 2014; Choi and Im, 2016; Kaplan and Rattazzi, 2016), which was adapted to this context by Craig and Garcia (2018). It remains to be seen if such scenarios can be found in the landscape.

The WGC can constrain not only the possibility that the standard model photon is massive but also the possibility that a dark photon A' has a fundamental Stueckelberg mass. One application is to dark photon dark matter, which is now the target of many dedicated experiments. Dark photon dark matter can arise from the primordial fluctuations of a massive vector field A' during inflation, which can account for the observed dark matter relic abundance if $m_{A'} \gtrsim 10^{-5}$ eV (Graham, Mardon, and Rajendran, 2016). However, for such a light dark photon, if the mass is of a fundamental Stueckelberg type then the WGC implies a cutoff lower than the inflationary Hubble scale assumed in the calculation of the dark matter relic abundance. This constraint excludes a substantial part of the parameter space of such models (Reece, 2019).

5. Axion inflation

The spectra of temperature and polarization fluctuations in the cosmic microwave background (CMB) radiation strongly suggest that the Universe experienced an early period of accelerated expansion known as inflation (Peiris *et al.*, 2003; Akrami *et al.*, 2020). This idea (Starobinsky, 1980; Guth, 1981) is most easily realized by a scalar field rolling slowly down a potential (Linde, 1982). The action describing this scenario is

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R - \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - V(\phi)\right], \quad (153)$$

where the scalar ϕ is called the inflaton. At leading order the metric $g_{\mu\nu}$ is taken to be the Friedmann-Robertson-Walker metric,

$$ds^2 = -dt^2 + a^2(t)d\vec{x}^2, \quad (154)$$

and the scalar field ϕ is taken to be homogeneous in space, $\phi(t,\vec{x}) = \phi(t)$. The equations of motion are

$$\ddot{\phi} + 3H\dot{\phi} + V'(\phi) = 0, \quad \frac{1}{3M_{\rm Pl}^2}\left(\frac{1}{2}\dot{\phi}^2 + V(\phi)\right) = \left(\frac{\dot{a}}{a}\right)^2, \quad (155)$$

and inflation happens when $H \equiv \dot{a}/a$ is approximately constant. Requiring $|\dot{H}| \ll H^2$ implies that $\dot{\phi}^2 \ll V$, and it is usually also assumed that the acceleration of ϕ is small ($|\ddot{\phi}| \ll \dot{\phi}H$); see Weinberg (2008). Together these requirements are equivalent to the "slow roll conditions"

$$\frac{|V'|}{V}M_{\rm Pl} \ll 1, \qquad \frac{|V''|}{V}M_{\rm Pl}^2 \ll 1. \quad (156)$$

Inflation ends when these conditions are violated, after which the field is usually expected to oscillate about its current minimum and in some manner (called reheating) decay into the dense gas of hot particles that we usually call the big bang.

CMB observables give us data about the inflaton potential V. Especially noteworthy for our purposes are the following primordial scalar and tensor power spectra:

$$k^3P_s(k) = \frac{H^4}{2\dot{\phi}^2} \approx \frac{V^3}{6M_{\rm Pl}^6V'^2}, \quad k^3P_t(k) = \frac{4H^2}{M_{\rm Pl}^2} \approx \frac{4V}{3M_{\rm Pl}^4}, \quad (157)$$

where the quantities on the right are evaluated at the value of ϕ corresponding to the time when modes of wave number k were exiting the inflationary horizon; see Maldacena (2003b). The scalar power spectrum has been well measured by the temperature anisotropy of the CMB, so it is the tensor spectrum that causes anisotropy in the B-mode polarization of the CMB, which is of most importance in learning more about the physics of inflation. In particular, from Eq. (157) we see that a measurement of tensor modes would give us direct information about the overall scale of the inflationary potential. The tensor amplitude is usually expressed via the tensor-to-scalar ratio

$$r \equiv \frac{P_t(k_*)}{P_s(k_*)} = \frac{8\dot{\phi}^2}{H^2M_{\rm Pl}^2} \approx \frac{8M_{\rm Pl}^2V'^2}{V^2}, \quad (158)$$

where k_* is a typical wave number of large scale structure, say, 0.05 Mpc^{-1}. For high-scale inflation that might lead to observable tensor modes, this corresponds to the value of the inflaton about 60 e-foldings before the end of inflation (i.e., at the time t_* when $\log[a(t_{\rm end})/a(t_*)] \approx 60$).

From Eq. (158), we see that the tensor-to-scalar ratio measures how quickly the inflaton is rolling. We also know how long inflation lasts (60 e-folds for high-scale inflation). Putting these together and integrating $\dot{\phi}$ over time, we obtain the following rough estimate on the distance traveled by the field ϕ during the course of inflation, which is known as the Lyth bound (Lyth, 1997):

$$\Delta\phi \approx \mathcal{O}(1)\left(\frac{r}{0.01}\right)^{1/2}M_{\rm Pl}. \quad (159)$$

The Lyth bound tells us that if tensor modes can be detected in the CMB in the near future (which would require $r \gtrsim 0.001$–0.01), then the inflaton must traverse a distance of order 1 in Planck units.

This "super-Planckian" field range is important because such large-field excursions can run into trouble with effective field theory. Quantum gravity sets an EFT cutoff no larger than the Planck scale, so a perturbative expansion of $V(\phi)$ in powers of ϕ suppressed by $M_{\rm Pl}$ will cause trouble when $\Delta\phi$ is $O(M_{\rm Pl})$ or larger. If tensor modes are observed in the CMB, therefore, we will need a model of large-field inflation that is not destroyed by large corrections arising from quantum gravity.

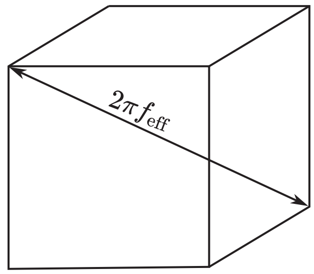

FIG. 13. In N-flation, a large number N of axions with individually sub-Planckian decay constants give rise to an effective decay constant $f_{\rm eff}$ that can be arbitrarily large and are realized by traveling along the space diagonal of the N-dimensional hypercube.

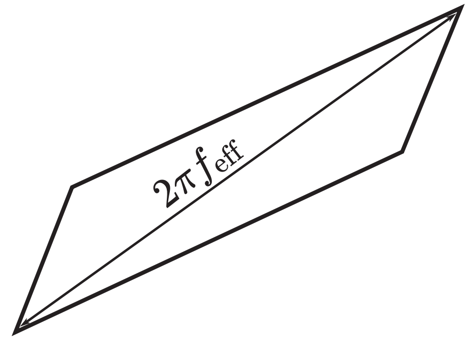

FIG. 14. In decay constant alignment, two axions with individually sub-Planckian decay constants are aligned such that their diagonal can be arbitrarily large.

Historically, axions have been considered the most promising route for circumventing this issue. An axion has a discrete gauge symmetry $\phi \to \phi + 2\pi$. This shift symmetry protects the axion potential from Planck-suppressed operators $\phi^n/M_{\rm Pl}^{n-4}$. The dynamics are instead controlled by a periodic potential, which is often assumed to be dominated by instanton effects, leading to an action of the form

$$S = \int d^4x[-\tfrac{1}{2}f^2(\partial_\mu\phi)^2 - V(\phi)], \tag{160}$$

where

$$V(\phi) = \Lambda_{\rm UV}^4 e^{-S_{\rm inst}}[1-\cos(\phi)] + O(e^{-2S_{\rm inst}}). \tag{161}$$

In Eq. (160) f is the axion decay constant, $S_{\rm inst}$ is the instanton action, and higher harmonics of the potential are suppressed by additional powers of $e^{-S_{\rm inst}}$. The resulting model of inflation is called natural inflation (Freese, Frieman, and Olinto, 1990), and it yields phenomenologically viable models of large-field inflation with detectable tensor modes ($r > 0.01$) for $f \gtrsim 10 M_{\rm Pl}$.

Although natural inflation is an appealing model, it has proven difficult to implement in string theory when $f \gtrsim M_{\rm Pl}$ and $S_{\rm inst} \gg 1$, with the basic problem being that such large excursions in scalar field space often lead to additional light degrees of freedom appearing that spoil inflation (Banks et al., 2003). This difficulty is part of what led AMNV to propose the WGC in 2007: the axion WGC (26) gives an upper bound on f, which in $D=4$ and assuming the instantons satisfying the WGC have a small instanton number tells us that

$$f S_{\rm inst} \lesssim M_{\rm Pl}. \tag{162}$$

Thus, any natural inflation model with observable tensor modes and a computable potential (meaning that $f \gtrsim 10 M_{\rm Pl}$ and $S_{\rm inst} \gg 1$) is in strong tension with the axion WGC.

By the time the WGC was introduced, various works had already considered possible ways to get around the previously mentioned difficulties and realize a model of large-field natural inflation consistent with quantum gravity. One such proposal is N-flation (Liddle, Mazumdar, and Schunck, 1998; Dimopoulos et al., 2008). As its name suggests, N-flation invokes not just 1 but N axion fields. If each field has a decay constant f, then by traveling along the diagonal in field space one sees an effective decay constant of $f_{\rm eff} = \sqrt{N} f$ (using the simple fact that an N-dimensional hypercube of side length f has a diagonal of length $\sqrt{N} f$; see Fig. 13). A related idea is decay constant alignment (Kim, Nilles, and Peloso, 2005): here only two axions are needed, but their decay constants are "aligned" so that the fundamental axion domain is not a square, but rather an elongated parallelogram, as shown in Fig. 14. Even though each individual axion may have a sub-Planckian decay constant, the diagonal direction in field space may be much larger than $M_{\rm Pl}$, thereby generating a model of natural inflation with a super-Planckian effective decay constant $f_{\rm eff} \gg M_{\rm Pl}$.

However, in their simplest incarnations, neither N-flation nor decay constant alignment evade the axion WGC. The individual instantons involved may be superextremal, but together they do not satisfy the convex hull condition, as shown in Fig. 15. In other words, there are no superextremal instantons associated with the diagonal directions of field space.

We can make a more general argument. Suppose that our theory features n instantons with an action

$$S = \int d^4x[-\tfrac{1}{2}\partial_\mu\vec{\phi}\cdot\mathbf{K}\cdot\partial^\mu\vec{\phi} - V(\vec{\phi})], \tag{163}$$

where $\mathbf{K}$ is the kinetic matrix for the axions. We further suppose that instantons generate a leading-order potential of the form

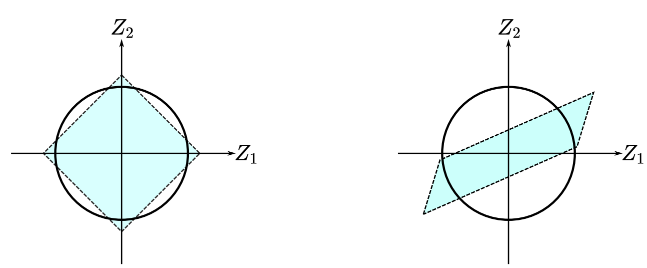

FIG. 15. In their simplest incarnations, both N-flation and decay constant alignment violate the axion WGC: the charge-to-action vectors $\vec{z}_k = \vec{Q}_k/S_k$ of the instantons are superextremal, but their convex hull does not contain the unit ball.

$$V(\vec{\phi}) = \sum_k \Lambda_{\rm UV}^4 e^{-S_k}[1-\cos(\vec{Q}_k\cdot\vec{\phi})], \qquad (164)$$

Next suppose that we want to inflate in the $\vec{e}$ direction of field space such that the inflaton starts at the point $\vec{\phi}=\phi_0\hat{e}$ and rolls to the minimum at the origin in approximately a straight line in field space. We assume that the largest value of ϕ_0 allowed satisfies

$$\phi_0(\vec{Q}_k\cdot\hat{e}) \le \pi \quad \text{for all } k \qquad (165)$$

since otherwise the inflaton sits in a cosine well of the kth potential term that does not contain the origin, which will presumably lead the inflaton to roll into a neighboring vacuum rather than the vacuum at $\vec{\phi}=0$.

Next we assume the axion WGC, which implies that for the given direction $\hat{e}$ there is a superextremal instanton satisfying

$$\frac{\vec{Q}_k\cdot\hat{e}}{fS_k} \gtrsim \frac{1}{M_{\rm Pl}}, \qquad (166)$$

where $f := \sqrt{\hat{e}\cdot\mathbf{K}\cdot\hat{e}}$ is the axion decay constant for the direction $\hat{e}$. Finally, we assume that $S_k > 1$ for perturbative control of the instanton expansion. Together with Eq. (165), this gives a bound on the physical displacement of the field,

$$\|\vec{\phi}\| := \sqrt{\vec{\phi}\cdot\mathbf{K}\cdot\vec{\phi}} = f\phi_0 \lesssim \pi M_{\rm Pl}. \qquad (167)$$

Thus, the axion WGC constrains the maximum axion field range to be $O(\pi M_{\rm Pl})$, which is too small to generate a successful model of natural inflation with observable tensor modes (Montero, Uranga, and Valenzuela, 2015; Rudelius, 2015a, 2015b; Brown et al., 2015, 2016). The generality of this observation reinvigorated provides hope that the consistency of quantum gravity might lead to testable predictions for cosmology and generate renewed interest in the WGC and the swampland program more generally.

There are several caveats to the previously mentioned argument that need to be discussed. One is that it assumes the axion WGC, which as we have seen is on somewhat shakier footing than the higher-form versions of the WGC. In particular, it is not immediately related to black hole evaporation. Relatedly the O(1) coefficient in the axion WGC has thus far not been decisively fixed; hence, our argument produced only a "squiggly" $\lesssim$ statement rather a sharper $\le$ statement. Various possibilities for the precise O(1) coefficient in the axion WGC bound were suggested by Heidenreich, Reece, and Rudelius (2016) and Andriolo et al. (2020). Moreover, the bound (162) relies on assuming that the instantons obeying the axion WGC have an instanton number that is O(1). This is natural from the point of view of the tower or sublattice WGCs, and also from the point of view of the idea that there should be objects obeying the WGC that are not black holes (or in the axion case that do not have large gravitational backreaction), but it follows from tower or sublattice WGCs only if we assume that the relevant tower or sublattice is not too sparse.

Another related issue is that the bound (165) assumes that every instanton whose charge-to-action vector $\vec{Q}_k/S_k$ contributes to the convex hull also contributes significantly to the axion potential. However, it is conceivable that the dominant contributions to the axion potential could come from instantons that violate the axion WGC, whereas the instantons that satisfy the WGC give only subleading, unimportant contributions to the potential. In this case, the inflationary dynamics are unconstrained by the axion WGC. This "extra instanton loophole" has driven a lot of interest in strong forms of the WGC (Hebecker et al., 2015; Bachlechner, Long, and McAllister, 2016). However, even the lattice WGC is not sufficient to close this loophole (Heidenreich et al., 2020). On the other hand, threading the extra instanton loophole seems to require a fair bit of tuning (Heidenreich et al., 2020), and to date super-Planckian axion decay constants have yet to be realized in string theory (Long, McAllister, and Stout, 2017).

It is also possible to try to relax the assumption that $S_k \gg 1$, which we suggested is required for perturbative control of the instanton expansion. In string compactifications, S_k is typically the size of some cycle of the Calabi-Yau in string units, so the α' expansion of string theory is not valid unless $S_k \gg 1$. However, in "extranatural inflation" (Arkani-Hamed et al., 2003) instantons in four dimensions come from the particles in 5D wrapping the compactification S^1. The instanton action for a particle of mass m wrapping a circle of radius R is given by $S_{\rm inst} = 2\pi mR$, and the contribution to the axion potential is given by

$$V(\phi) = \frac{3(-)^S}{4\pi^2}\frac{1}{(2\pi R)^4}\sum_{n\in\mathbb{Z}} c_n e^{-2\pi nRm_5} e^{in\phi},$$
$$c_n = \frac{(2\pi Rm_5)^2}{3n^3} + \frac{2\pi Rm_5}{n^4} + \frac{1}{n^5}. \qquad (168)$$

In this context, there is no problem with taking $S_{\rm inst} \ll 1$. This corresponds simply to a light particle with $m \ll 1/R$. Likewise, there is no problem with perturbative control of the potential: the $1/n^5$ term in c_n suppresses higher harmonics even for $S_{\rm inst} \ll 1$ (de la Fuente, Saraswat, and Sundrum, 2015). This significantly weakens the WGC bound (162) on the axion decay constant. By imposing the convex hull condition on both the U(1) of the parent 5D theory and the Kaluza-Klein U(1), one can strengthen the bound to $f \lesssim M_{\rm Pl}/S_{\rm inst}^{1/2}$, but this is still insufficient to close this "small action loophole" in the context of extranatural inflation (Heidenreich, Reece, and Rudelius, 2015).

Finally, and perhaps most importantly, these arguments apply solely to theories where the only fields relevant during inflation are axions and the metric. Including nonperiodic scalars would only reintroduce the UV sensitivity that we avoided with axions, but we saw in our previous discussion of the axion monodromy model (133) that the inclusion of a 3-form gauge field A_3 coupled to an axion by a Chern-Simons term offers a simple mechanism whereby an axion with $f \lesssim M_{\rm Pl}$ can nonetheless lead to a scalar $\star F_4$ that rolls down a potential for many Planck distances. Moreover, we saw that the 3-form WGC applied to A_3 and stringy corrections of the form $V(\star F_4)$ lead to only weak constraints on the range of this rolling. We presented this model in the context of cosmological relaxation, but its most compelling application is actually inflation, which indeed is what it was originally proposed for (Silverstein and Westphal, 2008; McAllister, Silverstein, and Westphal, 2010). Moreover, proposals have been given for embedding this model in a consistent string compactification, although there are still certainly details remaining to be worked out and to date no detailed model has been given where a large-field range is realized (Baumann and McAllister, 2015; McAllister et al., 2018; Kim and McAllister, 2020).

It is sometimes suggested that string theory does not allow for models with observable tensor modes, and the axion WGC was proposed in part to give an explanation for this claim, but axion monodromy casts serious doubt on this. Natural inflation with observable tensor modes may well be in the swampland, but the right lesson from this may just be that we should see what kind of predictions follow from the more general axion models that do seem to work. In particular, given the ever-improving observational upper bounds on r, it is natural to ask whether axion monodromy models exist which are not excluded but nonetheless predict observable tensor modes. The simplest potential, a quadratic one [see Eq. (139)], as well as simple extensions with other powers (McAllister et al., 2014), is now already excluded by the Planck satellite and ground-based experiments including BICEP and Keck (Ade et al., 2021; Kallosh and Linde, 2021), but variations on the model are possible (D'Amico, Kaloper, and Westphal, 2021; D'Amico, Kaloper, and Westphal, 2022). Perhaps the detailed issues remaining to be resolved in realizing axion monodromy in a genuine string compactification may yet lead to distinctive predictions. If so, the various forms of the WGC will likely be important tools in guiding us toward models that work. Either way it is noteworthy that ongoing observations are teaching us concrete things about physics near the Planck scale.

B. Implications for mathematics

The WGC is a statement about the charges and masses of particles in effective field theory. In string or M theory, supersymmetric effective field theories arise from compactifying on Calabi-Yau manifolds. Charged particles arise from p-branes wrapping p-cycles of the Calabi-Yau manifold. The charge of such a particle is determined by the homology class Σ wrapped by the brane, and the mass of the particle is determined by the volume of the wrapped cycle. Thus, the WGC translates into geometric statements about the volumes of representatives of various cycles in a Calabi-Yau manifold.

For concreteness, we now consider the case of M theory on a Calabi-Yau 3-fold X. This produces a 5D supergravity theory, and charged particles arise from $M2$-branes wrapping 2-cycles of X. The charge lattice of the theory is identified with the homology lattice $H_2(X,\mathbb{Z})$.

The resulting supergravity theory has a BPS bound: the mass of a particle of charge q_I is constrained to satisfy

$$m \geq \left(\frac{(2\pi)^2}{2\kappa_5}\right)^{1/3} |\zeta_q|, \tag{169}$$

where ζ_q is the "central charge," a quantity that depends linearly on q_I. It is sometimes said that the BPS bound is a sort of converse to the WGC bound, and there is a precise sense in which this is true: if there are BPS black holes in a given direction $\hat{q}$ in the charge lattice, then the BPS bounds and extremality bounds coincide in this direction in the lattice. The only way a particle of charge $q_I \propto \hat{q}$ can satisfy both the WGC bound and the BPS bound is if it saturates both bounds. A particle that saturates the BPS bound is called a BPS particle. Therefore, the tower or sublattice WGCs require an infinite tower of BPS particles of increasing mass or charge in every direction in the charge lattice for which the BPS bound coincides with the extremality bound.

Geometrically BPS particles arise from $M2$-branes wrapping "holomorphic" curves of X. Here a curve Σ is holomorphic if its volume is given by integrating the Kähler form J over it, $V_\Sigma = \int_\Sigma J$. Equivalently, we say that the curve is calibrated by the Kähler form.

The upshot of this is that the tower or sublattice WGCs imply the existence of an infinite tower of holomorphic curves in any direction $\hat{q}$ of the homology lattice $H_2(X,\mathbb{Z})$ for which the BPS bound coincides with the extremality bound. In fact, the condition that the BPS bound and extremality bounds coincide in the direction $\hat{q}$ can also be given a geometric interpretation: these bounds necessarily coincide for any $\hat{q}$ that resides in the so-called cone of moving curves $\mathcal{K}^\vee \subset H_2(X,\mathbb{R})$, which is equal to the cone dual of the "cone of effective divisors" (Alim, Heidenreich, and Rudelius, 2021).[17] Thus, the tower or sublattice WGCs translate to the nontrivial geometric statement that there must be an infinite tower of holomorphic curves for any rational direction $\hat{q}$ within the cone of moving curves $\mathcal{K}^\vee$.

This statement is powerful in that (a) it is a purely geometric statement with no reference to physics and (b) it can actually be verified in examples. Gopakumar-Vafa (GV) invariants (Gopakumar and Vafa, 1998a, 1998b, 1999) count the number of BPS particles (i.e., holomorphic curves) of a given charge q_I in the lattice,[18] and these can be computed for many Calabi-Yau hypersurfaces (Hosono et al., 1995a) and complete intersection Calabi-Yau manifolds (Hosono et al., 1995b). At the same time, the cone of moving curves $\mathcal{K}^\vee$ can often be computed (with some input from the 5D supergravity theory) using the methods of Alim, Heidenreich, and Rudelius (2021). Together with the automated computation of GV invariants introduced by Demirtas, Rios-Tascon, and McAllister (2022), this has enabled thorough checks of this geometric version of the tower or sublattice WGCs in more than 1400 Calabi-Yau manifolds (Gendler et al., 2022).

[17] If the theory in question allows certain U(1) gauge fields to be enhanced to a larger non-Abelian group, the effective cone of divisors may change under a phase transformation representing a Weyl reflection of the non-Abelian gauge group. In this case, the BPS and extremality bounds coincide in an even larger cone in the charge lattice, which is geometrically the cone dual to the intersection of the effective cone of divisors over all phases of the theory (Gendler et al., 2022).

[18] More accurately, GV invariants compute an index of BPS particles of a given charge, meaning that a nonzero GV invariant implies a nonzero number of BPS particles, whereas a vanishing GV invariant could result from an equal number of BPS hypermultiplets and BPS vector multiplets of a given charge.

Thus far we have focused our attention on BPS particles, which are required by the tower or sublattice WGCs in certain directions in the charge lattice where the BPS bound and extremality bound coincide. In a theory with eight supercharges, however, there will generically be some directions for which the BPS bound and the extremality bound do not coincide. The tower or sublattice WGCs still require the existence of superextremal particles, which means that they still impose constraints on the volumes of cycles of the Calabi-Yau manifold. To be more precise, the charge-to-mass vector of a particle associated with a p-cycle Σ of a Calabi-Yau manifold is given by (Hebecker, Rompineve, and Westphal, 2016)

$$\vec{z} = \frac{V_X^{1/2} \vec{q}_\Sigma}{V_\Sigma}, \tag{170}$$

where V_X is the volume of the Calabi-Yau X, V_Σ is the volume of Σ, and $\vec{q}_\Sigma$ labels the charge vector associated with the homology class. The norm $\|\vec{q}_\Sigma\|$ is the norm of the harmonic form related to Σ using the metric on X.

The particle is superextremal when $\|\vec{z}\| \geq \gamma_4^{1/2}$. In general, γ_4 depends on the direction $\hat{q}$ as well as the massless scalar fields in the theory, but it necessarily satisfies $\gamma_4 \geq 1/2$, which means that for a superextremal particle

$$\frac{V_X^{1/2} \|\vec{q}_\Sigma\|}{V_\Sigma} \geq \frac{1}{\sqrt{2}}. \tag{171}$$

The tower WGC therefore implies that for a given cycle $[\Sigma]$, there is an integer n and a representative of $n[\Sigma]$ satisfying Eq. (171). The sublattice WGC further implies that there is a universal n that is independent of $[\Sigma]$.

In some cases, this bound leads to surprising, nontrivial mathematical results. In particular, any 4-cycle $[\Sigma]$ in a Calabi-Yau 3-fold X with $h^{2,0}(X) = 0$ can be represented as a union of holomorphic and antiholomorphic representatives. Upon wrapping D4-branes on these representatives, these correspond to BPS and anti-BPS particles, respectively. The minimal volume representative of $[\Sigma]$ therefore has a volume no larger than the sum of the volumes of these holomorphic and antiholomorphic representatives, which we denote as $V(\Sigma_\cup)$. But satisfying Eq. (171) may require this union of holomorphic and antiholomorphic representatives to recombine into a new representative $\Sigma_{\min}$ whose volume is significantly smaller than $V(\Sigma_\cup)$. More precisely, the "recombination fraction"

$$\tau_\Sigma \coloneqq \frac{V(\Sigma_\cup) - V(\Sigma_{\min})}{V(\Sigma_{\min})} \tag{172}$$

may be much larger than 1 (Demirtas *et al.*, 2020). Physically this recombination corresponds to D4-branes wrapping these representatives recombining and fusing, and particles in 4D binding to form bound states of significantly smaller energy. Mathematically the existence of representatives $\Sigma_{\min}$ with a large recombination fraction $\tau_\Sigma \gg 1$ is a nontrivial consequence of the WGC that has been verified in some examples (Long *et al.*, 2021).

Finally, we remark that within the context of 5D M-theory compactifications there are interesting connections among the tower WGC, the WGC for strings, and the swampland distance conjecture and various mathematical conjectures about Calabi-Yau manifolds known as cone conjectures (Morrison, 1993, 1994). See Heidenreich and Rudelius (2022) for more discussion of these connections.

C. Implications for general relativity

The weak cosmic censorship hypothesis holds that for generic initial data the maximal Cauchy development possesses a complete future null infinity (Penrose, 1969). Colloquially there can be no naked singularities visible at future null infinity: any such singularity must be hidden behind a horizon.

There is strong numerical evidence that cosmic censorship is violated in more than four spacetime dimensions (Lehner and Pretorius, 2012), as black strings may pinch off and develop singularities due to the Gregory-Laflamme instability (Gregory and Laflamme, 1993). In four dimensions, however, such instabilities do not exist, and violation of cosmic censorship is much less certain.[19]

A promising class of counterexamples to cosmic censorship in four dimensions was proposed by Horowitz, Santos, and Way (2016), and strong numerical evidence for these counterexamples was subsequently provided by Crisford and Santos (2017). These examples involve a U(1) gauge field coupled to gravity in asymptotically AdS space, with an action (as usual we set the AdS radius to 1)

$$S = \frac{1}{2\kappa^2} \int d^4x \sqrt{-g} (R + 6 - F^{\mu\nu} F_{\mu\nu}), \tag{173}$$

with $F = dA$. The boundary metric is chosen to be flat,

$$ds_\partial^2 = -dt^2 + dr^2 + r^2 d\varphi^2, \tag{174}$$

and the only nonzero component of the potential at the boundary is the time component,

$$A_\partial = + \frac{a(t) dt}{(1 + r^2)^{n/2}}, \tag{175}$$

where n is an integer controlling the falloff of the field at large r.

Apparent violations of cosmic censorship arise when $a(t)$ is chosen to vanish at $t = 0$ but increases to a constant value larger than some critical value $a_{\max}$. In this case, there is no smooth static end point of the evolution, so one expects the curvature F^2 to grow indefinitely at late times. Numerical simulations confirm this expectation for $n = 1$ (Crisford and

[19]Even in higher dimensions, known violations of cosmic censorship have zero mass and occur in Planck-sized regions where quantum gravitational effects become important. It has been argued that such quantum effects restore some notion of cosmic censorship, so these counterexamples to cosmic censorship are relatively benign (Emparan, 2020).

Santos, 2017). Note that the curvature does not diverge in finite time, so this example does not violate the letter of cosmic censorship, although it does violate the spirit of it.

This class of counterexamples disappears, however, in the presence of a superextremal scalar field (Horowitz, Santos, and Way, 2016; Crisford, Horowitz, and Santos, 2018). In particular, suppose that we add a charged scalar Φ to the action,

$$S_\Phi = -\frac{1}{4\pi G}\int d^4x\sqrt{-g}[(D_\mu\Phi)(D^\mu\Phi)^\dagger + m^2\Phi\Phi^\dagger], \tag{176}$$

with $D_\mu = \nabla_\mu - i\tilde{q}A_\mu$ [here $\tilde{q} = (\sqrt{2}e/\kappa)q$, where q is the integral charge used throughout].[20] The proposed WGC bound in AdS (22) then becomes

$$\tilde{q} \geq \Delta = \tfrac{3}{2} + \sqrt{\tfrac{9}{4} + m^2}. \tag{177}$$

When Eq. (177) is satisfied, for all choices of n perturbations of Φ become unstable before a grows to the critical value $a_{\max}$, and cosmic censorship is restored. When this bound is violated, the solution with the scalar field is still singular, and it is once again likely that cosmic censorship is violated (Horowitz, Santos, and Way, 2016). Thus, there is evidently a one-to-one correspondence between satisfying the WGC and obeying cosmic censorship in this setup.

This result is already suggestive. But the connection between the WGC and cosmic censorship becomes even more impressive when dilatonic couplings and multiple scalar fields are included, as was done by Horowitz and Santos (2019). In the dilatonic case, one begins with the action

$$S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}(R + 6 - e^{-2\alpha\phi}F^{\mu\nu}F_{\mu\nu} - 2\nabla_\mu\phi\nabla^\mu\phi) \tag{178}$$

in place of Eq. (173). In Eq. (178) ϕ is a massless, uncharged scalar field that we refer to as the dilaton, not to be confused with the massive, charged scalar field Φ. In the presence of this dilatonic coupling, the WGC bound for Φ is modified to

$$\tilde{q} \geq \tilde{q}^W \equiv \Delta(1+\alpha^2)^{1/2}, \tag{179}$$

where Δ is given by Eq. (177). Notably the minimal charge-to-mass ratio $\tilde{q}/\Delta$ varies continuously with the parameter α.

Horowitz and Santos (2019) constructed numerical solutions to the equations of motion in the presence of the dilaton, with a boundary vector potential given by

$$A_\partial = \frac{a dt}{(1+r^2)^n} \tag{180}$$

focusing, in particular, on the case of $n = 4$. They found that for $\alpha < 1$ cosmic censorship is again preserved precisely when the WGC bound (179) is satisfied, as shown in Fig. 16. This is noteworthy in that it establishes a WGC–cosmic censorship connection over a one-parameter family of theories indexed by α. For $\alpha > 1$, numerical solutions suggest that it may be possible to preserve cosmic censorship even when $\tilde{q}/\tilde{q}^W$ (or, equivalently, q/q^W) is slightly smaller than 1, as shown in Fig. 17. It is possible that this conclusion could be modified at large values of a, and the one-to-one correspondence between the WGC and cosmic censorship could be restored.

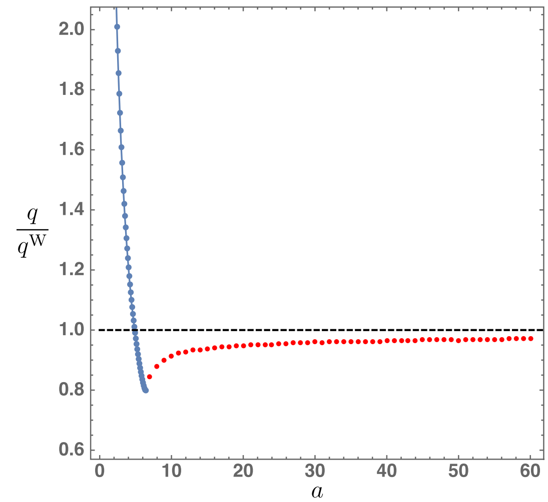

FIG. 16. For fixed $n = 4$, $\Delta = 2$, and dilatonic coupling $\alpha = 0.9$, the condition to preserve cosmic censorship is precisely the WGC bound (179). The blue dots indicate the onset of solutions with $\Phi \neq 0$, and the red dots indicate the approximate location of singular solutions. From Horowitz and Santos, 2019.

Finally, Horowitz and Santos (2019) also considered the relationship between cosmic censorship and the WGC in theories with two gauge fields. In their analysis, the asymptotic profile of the gauge fields is taken to be

$$A_{I\partial} = \frac{a_I}{(1+r^2)^n}dt, \tag{181}$$

focusing again on the case of $n = 4$, with $a_1 = \lambda a_2$. There are now two massive scalar fields Φ_1 and Φ_2. The former has charge $\tilde{q}_1$ under the first gauge field and is uncharged under the second gauge field, whereas the latter carries charge $\tilde{q}_2$ under the second gauge field and is uncharged under the first gauge field. Since there are multiple gauge fields, the WGC bound is equivalent to the convex hull condition (see Sec. III.C), which is given by

$$\frac{1}{z_1^2} + \frac{1}{z_2^2} \leq 1, \tag{182}$$

with $z_I = \tilde{q}_I/\Delta_I$. By constructing numerical solutions to the equations of motion for various choices of λ, $\tilde{q}_I$, and Δ_I,

[20]We observed after Eq. (8) that the WGC bound (6) does not involve any powers of $\hbar$ and thus potentially has classical consequences, but those remarks applied to the case of a classical particle. A similar statement applies for a charged classical field such as Φ, but we need to be more careful since, if we restore c and $\hbar$, then (in Heaviside-Lorentz units) the mass m and charge $\tilde{q}$ appearing in Eq. (176) both have units of inverse length. The true charge and mass of a particle appearing after this field is quantized are related to these by powers of $\hbar$ and c, but in writing the WGC inequality the powers of $\hbar$ and c drop out since m and $\tilde{q}$ have the same units and κ has already been absorbed into $\tilde{q}$.

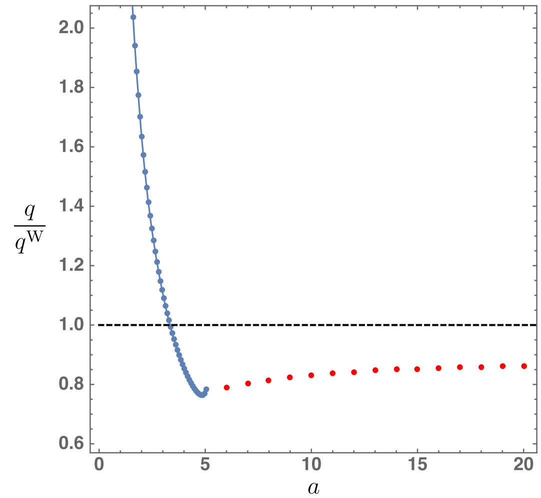

FIG. 17. For fixed $n = 4$, $\Delta = 2$, and dilatonic coupling $\alpha = \sqrt{3}$, the condition to preserve cosmic censorship does not seem to match with the WGC bound (179), although it is possible that modifications at large a could restore the correspondence. The blue dots indicate the onset of solutions with $\Phi \neq 0$, and the red dots indicate the approximate location of singular solutions. From Horowitz and Santos, 2019.

Horowitz and Santos (2019) provided strong evidence that cosmic censorship is preserved precisely when the convex hull condition is satisfied. Figure 18 depicts this correspondence for one choice of λ, $\tilde{q}_2$, and Δ_1 Δ_2 and a varying $\tilde{q}_1$.

Horowitz, Santos, and Way (2016), Crisford and Santos (2017), and Horowitz and Santos (2019) thus revealed a noteworthy and surprising correspondence between the WGC and the weak cosmic censorship conjecture. A couple of aspects of this correspondence are worth further attention. First, note that the mildest form of the WGC, which simply requires the existence of a superextremal state, is not sufficient to preserve cosmic censorship. The mild WGC could be satisfied in principle by a finite-size black hole state due to subleading corrections that slightly increase Q/M relative to Q/M for an infinitely large, extremal black hole. In the present scenario, however, preserving cosmic censorship requires a superextremal field Φ (and thus quantum mechanically a superextremal particle) and not merely a superextremal black hole.

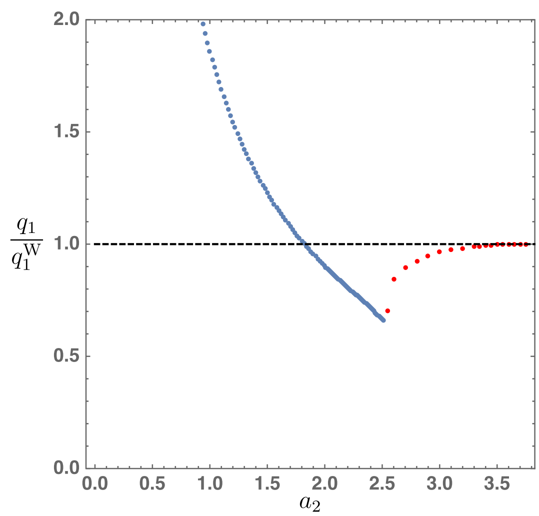

FIG. 18. For fixed $n = 4$, $a_1 = 3a_2$, $q_2 = 4$, $\Delta_1 = \Delta_2 = 2$, the condition to preserve cosmic censorship is precisely the convex hull condition (182). The blue dots indicate the onset of solutions with $\Phi_1 \neq 0$, and the red dots indicate the approximate location of singular solutions. From Horowitz and Santos, 2019.

Second, we observe that in all examples thus far the superextremal field that saves cosmic censorship is a scalar. The usual formulations of the WGC do not put any restrictions on the spins of the superextremal particles that are required, but it is interesting to consider whether there might be some such restriction. One natural question for future study is whether or not superextremal bosonic fields of nonzero spin similarly prevent violations of cosmic censorship. More generally one can also consider the question of whether or not fermions can do the same once quantum effects are included. All string compactifications that we are currently aware of in fact do have a superextremal scalar, but this may be an accident of supersymmetry, so there is no strong evidence yet for an interplay between the WGC and spin.

Finally, we mention an interesting connection between the WGC and a different kind of gravitational censorship, namely, of super-Planckian spatial field variations. A simple dimensional analysis suggests that such a field configuration could collapse into a black hole. Indeed, this is often true (Nicolis, 2008). However, one case in which a classically stable field configuration with an arbitrarily large scalar field variation, not screened by a horizon, can be constructed is a charged Kaluza-Klein bubble stabilized by flux (Horowitz, 2005). The scalar field is the radion, which traverses an infinite distance in field space to the bubble wall where $R \to 0$. Thus, arbitrarily large-field values are not classically censored. In quantum theory, the solution becomes unstable if charged matter satisfying the WGC exists due to Schwinger pair production, which dynamically censors the large-field excursion (Draper and Farkas, 2019).

The general picture painted by these examples is that the WGC can play an important role in ensuring the validity of effective field theory. It prevents a low-energy observer from accessing arbitrarily high-energy scales (in the case of cosmic censorship) or field values (in the case of super-Planckian censorship).

VIII. OUTLOOK

In this review, we saw that the weak gravity conjecture potentially offers a deep organizing principle for unlocking the puzzle of quantum gravity. In particular, the landscape of string vacua is large, yet to our knowledge the WGC is obeyed in all of them. Suitably strong forms of the WGC place meaningful constraints on particle physics and cosmology and have further consequences for black holes, pure mathematics, conformal field theories, etc. Thus, whereas most swampland conjectures fall into either the “rigorous but uninteresting” category or the “interesting but not rigorous” category, the WGC has claims to both rigor and importance.

Nonetheless, despite all that we have learned about it, the WGC remains shrouded in mystery. We saw in Sec. V that several arguments point qualitatively to the validity of the WGC: it is plausible that the WGC is satisfied up to $O(1)$ coefficients. However, examples in string theory suggest something stronger: in all such examples, the WGC is satisfied with the precise $O(1)$ coefficient determined by the black hole extremality bound. This suggests that the WGC may be required for consistency in black hole physics, but it is not yet clear what goes wrong if the WGC is violated.

The evidence for the WGC coming from string theory is strong, but there is still a chance that the WGC suffers from the lamppost effect: the known string examples necessarily involve either (a) weak gauge coupling or (b) BPS particles, the only particles that we currently know how to track in a strongly coupled regime. An example in which the WGC is satisfied by nonsupersymmetric states in a regime of strong coupling is highly desirable, although perhaps unfeasible. Without this, a more compelling black hole argument is likely required to rule out the possibility of the lamppost effect.

Another interesting direction for future research involves the classification of weak coupling limits, as initiated in by Grimm, Palti, and Valenzuela (2018), Corvilain, Grimm, and Valenzuela (2019), Klaewer *et al.* (2021), Lanza *et al.* (2021a, 2021b), Perlmutter *et al.* (2021), and Lee, Lerche, and Weigand (2022b). The emergent string conjecture of Lee, Lerche, and Weigand (2022b), in particular, suggests that any infinite distance–weak coupling limit must be either a decompactification limit or an emergent string limit in which a fundamental string becomes tensionless. In the former case, a tower of light, superextremal Kaluza-Klein modes emerge.[21] In the latter case, the modular invariance argument of Sec. IV.D ensures that the particles are superextremal. Thus, at weak coupling the tower or sublattice WGCs follow from the emergent string conjecture (which could have stronger phenomenological implications).

In Sec. IV, we encountered examples in which the sublattice WGC is satisfied by a sublattice of superextremal particles of coarseness $n > 1$ (and index greater than 1). This raises the question of how large the coarseness may become or, equivalently, how sparse the sublattice is allowed to be. This question is important because the consequences of the sublattice WGC for low-energy physics can be arbitrarily weak if the coarseness is allowed to be arbitrarily large. The maximum coarseness we encountered was $n = 3$: the sublattice of superextremal particles is not sparse, and it is plausible that the sublattice WGC will always be satisfied with a coarseness of $O(1)$.

Before we can attempt to place any sort of universal upper bound on the coarseness or index of the sublattice, however, we must first understand how this sublattice shows up in the low-energy data of the theory. All of the examples with a coarseness $n > 1$ constructed thus far are orbifold models, and n divides the order of the orbifold group. However, we do not yet have a clear understanding of how this UV relationship manifests in the IR. One possibility is that the coarseness is related to the global structure of the low-energy gauge group, but more work is needed to clarify this picture.

As a final direction for future research, we remark that the statement of the WGC in the presence of Chern-Simons terms is not yet well understood. Heidenreich, Reece, and Rudelius (2021) presented a recent example of WGC mixing in which the WGC for different p-form gauge fields were mixed up in the presence of Chern-Simons terms; see also Montero, Uranga, and Valenzuela (2017), Heidenreich *et al.* (2021a), Brennan and Cordova (2022), and Kaya and Rudelius (2022). These Chern-Simons terms imply that the gauge symmetry acquires a higher-group structure (Sharpe, 2015; Córdova, Dumitrescu, and Intriligator, 2019), and it seems likely that the full statement of the WGC is modified in the presence of such higher-group symmetries, which is reminiscent of how the WGC is modified to the convex hull condition in theories with multiple $U(1)$'s. Understanding this better might have interesting implications for axion monodromy.

The weak gravity conjecture has produced no shortage of surprises over the course of its 15-year existence, providing us with new insights into quantum gravity and unexpected connections between disparate areas of theoretical physics. Yet, many of the most important questions remain open: Is the weak gravity conjecture true? If so, why? And which versions of the conjecture are the right ones? The answers to these questions may well lead us to even greater surprises than the ones we have already encountered.

ACKNOWLEDGMENTS

We are grateful to many physicists for conversations that have shaped our understanding of the weak gravity conjecture over the years. We particularly want to thank Murad Alim, Nima Arkani-Hamed, Tom Banks, Clifford Cheung, Clay Córdova, Patrick Draper, Thomas Dumitrescu, Netta Engelhardt, Muldrow Etheredge, Gary Horowitz, Isabel Garcia Garcia, Manki Kim, Cody Long, Matteo Lotito, Javier Magan, Juan Maldacena, Liam McAllister, Jacob McNamara, Miguel Montero, Kantaro Ohmori, Hirosi Ooguri, Grant Remmen, Steve Shenker, Gary Shiu, John Stout, Leonard Susskind, Cumrun Vafa, and Irene Valenzuela for collaboration or discussions, but this is by no means an exhaustive list of those to whom we owe thanks. D. H. is supported by the Simons Foundation as a member of the "It from Qubit" collaboration, the Sloan Foundation as a Sloan Fellow, the Packard Foundation as a Packard Fellow, the Air Force Office of Scientific Research under Grant No. FA9550-19-1-0360, and the U.S. Department of Energy (DOE) under Grants No. DE-SC0012567 and No. DE-SC0020360. The work of B. H. is supported by NSF Grant No. PHY-1914934. The work of M. R. is supported in part by NASA Grant No. 80NSSC20K0506 and DOE Grant No. DE-SC0013607. The work of T. R. is supported by NSF Grant No. PHY-1820912, the Simons Foundation, and the Berkeley Center for Theoretical Physics.

APPENDIX: THE BLACK HOLE EXTREMALITY BOUND

In a gravitational theory with no naked singularities, a black hole of nonzero charge Q cannot be arbitrarily light,

[21] In the supersymmetric case, these KK modes are BPS and saturate the WGC bound. In the nonsupersymmetric case, these KK modes satisfy the WGC bound with room to spare after the radion is stabilized; see Sec. IV.A for more details.

$$M_{\rm BH} \geq M_{\rm ext}(Q) > 0, \tag{A1}$$

where the extremal mass $M_{\rm ext}(Q)$ is defined as the infinimum of the set of possible masses for black holes of charge Q. This bound arises because the energy stored in the electromagnetic field is a positive-energy source for the gravitational field, and in a theory with matter obeying reasonable energy conditions this gravitational flux cannot be canceled without introducing a naked singularity as a negative-energy source. A black hole saturating this bound is extremal, whereas all others are subextremal.

To determine the extremal mass $M_{\rm ext}(Q)$, and thereby the extremality bound (A1), it would suffice to find an extremal black hole solution of charge Q and read off its mass. However, an extremal black hole solution of a given charge does not always exist because taking the extremal limit sometimes generates a singularity at the event horizon. Furthermore, identifying whether a given solution is extremal is not straightforward. Experience with Reissner-Nordström black holes suggests that vanishing surface gravity is closely connected to extremality, but as we see in Appendix 4, not every black hole solution with vanishing surface gravity is extremal.

To solve the first problem, we expand our field of interest to include charged solutions that are merely limits of black holes and not necessarily black holes themselves. We call such limiting cases singular black holes. Familiar examples of singular black holes include the background generated by $N \gg 1$ $D0$-branes in type IIA string theory. Understanding what happens near a singular black hole usually requires UV information that goes beyond effective field theory. For example, near a stack of $D0$-branes the string coupling becomes large and there is a dual M-theory description.

To solve the second problem, note that Hawking radiation must shut off in the extremal limit to satisfy cosmic censorship, so either the surface gravity g_h (the Hawking temperature) or the horizon area A_h (the Bekenstein-Hawking entropy) must go to zero in this limit. We refer to black holes with either of these two properties as quasiextremal, whereas those with $A_h \to 0$ are singular. Cosmic censorship requires extremal black holes to be quasiextremal, but as previously noted quasiextremal black holes are not always extremal; see Appendix 4 for examples. We call black holes that are neither extremal nor quasiextremal nonextremal.

In this Appendix, we review general techniques for determining the extremality bound in the large-Q limit (where derivative corrections can be ignored) for theories with multiple U(1) gauge fields A^A, massless scalars ϕ^i, and a vanishing cosmological constant. For simplicity, we assume that the lightest black holes of a given charge Q are spherically symmetric.[22] Temporarily ignoring the possibility of magnetic charge, the relevant terms in the low-energy Einstein-frame effective action at two-derivative order are

[22]Without this assumption, the problem is unsolved in general, except in special cases where a BPS-like bound can be derived using spinor methods, an approach taken by Gibbons et al. (1983).

$$S_0 = \int d^d x \sqrt{-g} \left[\frac{1}{2\kappa^2} R - \frac{1}{2} t_{AB}(\phi) F_2^A \cdot F_2^B - \frac{1}{2} G_{ij}(\phi) \nabla\phi^i \cdot \nabla\phi^j \right], \tag{A2}$$

where $F_2^A = dA_1^A$. The types of two-derivative terms omitted from Eq. (A2) do not affect spherically symmetric black holes with a purely electric charge (Heidenreich, 2020).

1. Black hole solutions

In a convenient gauge, a spherically symmetric metric ansatz takes the form

$$ds^2 = -e^{2\psi(r)} f(r) dt^2 + e^{-2\psi(r)/(d-3)} \left[\frac{dr^2}{f(r)} + r^2 d\Omega_{d-2}^2 \right] \tag{A3}$$

for functions $\psi(r)$ and $f(r)$ to be determined, where $d\Omega_{d-2}^2$ is the round metric of unit radius on S^{d-2}. The electric charge of the solution is

$$Q_A = \oint_{S^{d-2}} t_{AB}(\phi) \star F_2^B, \tag{A4}$$

where the integral is taken over a sphere enclosing the horizon. Spherical symmetry fixes the electric field at

$$F_2^A = -\frac{t^{AB}(\phi) Q_B}{V_{d-2}} \frac{e^{2\psi} dt \wedge dr}{r^{d-2}}, \tag{A5}$$

where $t^{AB}(\phi)$ is the inverse of $t_{AB}(\phi)$ and $V_{d-2} = 2\pi^{(d-1)/2}/\Gamma[(d-1)/2]$ is the volume of S^{d-2}.

One component of Einstein's equations is now

$$f''(r) + \frac{3d-8}{r} f'(r) + 2\frac{(d-3)^2}{r^2}[f(r) - 1] = 0, \tag{A6}$$

with the solution $f(r) = 1 + A/r^{d-3} + B/r^{2(d-3)}$. To interpret A and B, we switch to ingoing Eddington-Finkelstein coordinates,

$$ds^2 = -\frac{F(\rho) dv^2}{R^{2(d-3)}(\rho)} + \frac{2 dv d\rho}{(d-3) R^{d-4}(\rho)} + R^2(\rho) d\Omega_{d-2}^2, \tag{A7}$$

where $\rho = r^{d-3}$, $R(\rho) = r e^{-\psi/(d-3)}$, and $F(\rho) = r^{2(d-3)} f(r)$. A smooth event horizon occurs when $F \to 0$ with R finite. If one exists, $F(\rho) = \rho^2 + A\rho + B$ can be factored as

$$F(\rho) = (\rho - \rho_+)(\rho - \rho_-), \qquad \rho_+ \geq \rho_-, \tag{A8}$$

leading to an outer (inner) horizon at $\rho = \rho_+$ ($\rho = \rho_-$). We set $\rho_- = 0$ using the residual gauge symmetry $\rho \to \rho + \text{const}$ with $F(\rho)$ and $R(\rho)$ held fixed such that

$$F(\rho) = \rho(\rho - \rho_h) \Leftrightarrow f(r) = 1 - \frac{r_h^{d-3}}{r^{d-3}}, \tag{A9}$$

with $\rho_h = r_h^{d-3}$.

In terms[23] of $z := 1/(d-3)V_{d-2}r^{d-3}$, $f(z) = 1 - z/z_h$ and the remaining equations of motion are

$$\frac{d}{dz}[f\dot{\phi}^j] + f\Gamma^i_{jk}\dot{\phi}^j\dot{\phi}^k = \frac{1}{2}G^{ij}Q^2_{,j}(\phi)d^{2\psi}, \tag{A10}$$

$$k_N^{-1}\frac{d}{dz}[f\dot{\psi}] = e^{2\psi}Q^2(\phi), \tag{A11}$$

$$k_N^{-1}\dot{\psi}(f\dot{\psi} + \dot{f}) + fG_{ij}(\phi)\dot{\phi}^i\dot{\phi}^j = e^{2\psi}Q^2(\phi), \tag{A12}$$

where $\dot{} = d/dz$, $G^{ij}(\phi)$ is the inverse of $G_{ij}(\phi)$, $\Gamma^i{}_{jk} = (1/2)G^{il}(G_{lj,k} + G_{lk,j} - G_{jk,l})$ are the associated Christoffel symbols,

$$Q^2(\phi) := t^{AB}(\phi)Q_AQ_B, \tag{A13}$$

and

$$k_N := \frac{d-3}{d-2}\kappa^2 \tag{A14}$$

is the rationalized Newton force constant [such that $F_{\rm grav} = (-k_N/V_{d-2})mm'/r^{d-2}$].

Note that Eqs. (A10)–(A12) are z-translation invariant,[24] so for any solution passing through a point $\phi^i_0 = \phi^i(z_0)$ there is a corresponding solution $\phi^{i\prime}(z) = \phi^i(z + z_0)$, with $\phi^{i}_\infty{}' = \phi^i_0$. Moreover, Eq. (A12) is a consistent constraint in that the derivative of f times Eq. (A12) is a linear combination of Eqs. (A10) and (A11).

When $r_h > 0$ (z_h is finite), a smooth horizon requires $\dot{\psi}(z_h) = -k_N z_h e^{2\psi_h}Q^2(\phi_h) \le 0$ by evaluating Eq. (A11) at $z = z_h$, where $\psi_h = \psi(z_h)$. Likewise, when $r_h = 0$ ($z_h = \infty$), a smooth horizon requires $ze^{\psi} \propto R^{-(d-3)}$ to approach a nonzero constant as $z \to \infty$. Hence, $\psi \to -\log z + \text{const}$, implying that $\dot{\psi} \to -1/z < 0$. Combining Eqs. (A11) and (A12), we obtain

$$\ddot{\psi} = \dot{\psi}^2 + k_N G_{ij}(\phi)\dot{\phi}^i\dot{\phi}^j. \tag{A15}$$

Hence, $\ddot{\psi} \ge 0$, and we conclude that

$$\dot{\psi} \le 0 \quad \text{for all } z \le z_h \tag{A16}$$

is required for a smooth horizon. Conversely, Eq. (A16) together with $\ddot{\psi} \ge 0$ and $\psi_\infty = \psi(z=0) = 0$ gives $z\dot{\psi}_\infty \le \psi(z) \le 0$ for $0 \le z \le z_h$, so when $r_h > 0$ (z_h is finite), ψ_h is finite and the horizon is smooth.

Because a condition of Eq. (A16) is preserved under limits, singular black holes must also satisfy Eq. (A16). Likewise, because any $r_h > 0$ solution satisfying Eq. (A16) is smooth, $r_h = 0$ solutions satisfying Eq. (A16) are limits of smooth solutions, so solutions to Eqs. (A10)–(A12) are (possibly singular) black holes if and only if Eq. (A16) holds.

[23]Compared to (Heidenreich, 2020) $z^{(\rm here)} = z^{(\rm there)}/V_{d-2}$.

[24]To preserve the boundary condition $\psi_\infty = \psi(z=0) = 0$, we then shift ψ and rescale z to compensate.

Such black hole solutions have an Arnowitt-Deser-Miser mass

$$M = k_N^{-1}\left[-\dot{\psi}_\infty + \frac{1}{2z_h}\right] \tag{A17}$$

[positive by Eq. (A16)] and a surface gravity and horizon area

$$g_h = \frac{d-3}{2r_h}e^{[(d-2)/(d-3)]\psi_h}, \qquad A_h = V_{d-2}r_h^{d-2}e^{-[(d-2)/(d-3)]\psi_h}$$
$$\Rightarrow g_hA_h = (d-3)V_{d-2}r_h^{d-3} = \frac{1}{2z_h}. \tag{A18}$$

Therefore, $r_h = 0$ ($z_h = \infty$) is the quasiextremal case, with coincident (possibly singular) inner and outer horizons, and $r_h > 0$ (z_h finite) is the nonextremal (invariably smooth) case.

a. Magnetic charge

In four dimensions, spherical symmetry allows black holes to carry both electric and magnetic charges. The theta term, which had no effect on purely electrically charged black holes, then becomes important,

$$S = S_0 - \frac{1}{8\pi^2}\int \theta_{AB}(\phi)F_2^A \wedge F_2^B. \tag{A19}$$

The electric and magnetic charges are defined as

$$Q_A = \oint\left[t_{AB}\star F^B + \frac{\theta_{AB}}{4\pi^2}F^B\right], \qquad \tilde{Q}^A = \frac{1}{2\pi}\oint F^A. \tag{A20}$$

The black hole equations (A10)–(A12) take the same form as before [see Heidenreich (2020)] but with

$$Q^2(\phi) = t^{AB}(\phi)\left[Q_A - \frac{\theta_{AC}(\phi)}{2\pi}\tilde{Q}^C\right]\left[Q_B - \frac{\theta_{BD}(\phi)}{2\pi}\tilde{Q}^D\right]$$
$$+ 4\pi^2 t_{AB}(\phi)\tilde{Q}^A\tilde{Q}^B. \tag{A21}$$

b. Black branes

Generalizing to homogeneous, isotropic, and spherically symmetric black $(p-1)$-branes, the relevant effective action is

$$S = \int d^dx\sqrt{-g}\left[\frac{1}{2\kappa^2}R - \frac{1}{2}t_{AB}(\phi)F^A_{p+1}\cdot F^B_{p+1}\right.$$
$$\left. - \frac{1}{2}G_{ij}(\phi)\nabla\phi^i\cdot\nabla\phi^j\right]. \tag{A22}$$

With the appropriate ansatz [see Heidenreich (2020) with $z^{(\rm here)} = z^{(\rm there)}/V_{d-p-1}$] the black hole equations again take the form of Eqs. (A10)–(A12), with k_N replaced by the rationalized gravitational force constant for $(p-1)$-branes $k_{(p)} = [p(d-p-2)/(d-2)]\kappa^2$. Magnetic charge can be added consistent with spherical symmetry when $d = 2p+2$; see Heidenreich (2020) for details.

2. Quasiextremal black holes

The quasiextremal case, in which $f(r) = 1$, has several interesting properties that play an important role in determining the extremality bound.

a. Vanishing self-force

Evaluating Eq. (A12) at $z = 0$ ($r = \infty$), we obtain

$$k_N M^2 + G^{ij}_\infty \mu_i \mu_j = t^{AB}_\infty Q_A Q_B, \tag{A23}$$

where μ_i is the scalar charge appearing in $\phi^i(z) = \phi^i_\infty - G^{ij}_\infty \mu_j z + O(z^2)$. Thus, the long-range self-force between an identical pair of quasiextremal black holes vanishes (Heidenreich, 2020). As a corollary, $M = \sqrt{(1/k_N)(t^{AB}_\infty Q_A Q_B - G^{ij}_\infty \mu_i \mu_j)} \le \sqrt{(1/k_N) t^{AB}_\infty Q_A Q_B}$; hence, quasiextremal black holes coupled to moduli are no heavier than an extremal Reissner-Nordström black hole of the same charge that would result if the moduli were artificially frozen in place at their asymptotic values.

At this point note that even the short-range forces between identical quasiextremal black holes vanish if we restrict our attention to the classical, two-derivative effective action. This can be shown by explicitly constructing the static, multicenter solutions corresponding to several such black holes at rest near each other. While numerous examples of such solutions date back many years (Majumdar, 1947; Papaetrou, 1947; Breitenlohner, Maison, and Gibbons, 1988), a summary was recently given by Van Riet (2021). Generic quantum and derivative corrections not only alter the short-range forces but also change the extremal black hole solutions, and thereby the long-range forces; see Sec. VI.C.

b. Attractor mechanism

A smooth horizon requires $R(\rho)$ to remain finite as $\rho \to 0$. When we define $\chi \coloneqq \psi + \log z$, $e^{-\chi} = (d-3) V_{d-2} R^{d-3}$, so $\chi(z)$ must remain finite as $z \to \infty$. Written in terms of χ and $\tau = -\log z$, the equations of motion become

$$\frac{d^2\phi^i}{d\tau^2} + \Gamma^i_{jk} \frac{d\phi^j}{d\tau} \frac{d\phi^k}{d\tau} = -\frac{d\phi^i}{d\tau} + \frac{1}{2} G^{ij} Q^2_{,j} e^{2\chi}, \tag{A24}$$

$$\frac{d^2\chi}{d\tau^2} = -\frac{d\chi}{d\tau} + k_N e^{2\chi} Q^2 - 1, \tag{A25}$$

$$\left[\frac{d\chi}{d\tau}\right]^2 + k_N G_{ij} \frac{d\phi^i}{d\tau} \frac{d\phi^i}{d\tau} = -2 \frac{d\chi}{d\tau} + k_N e^{2\chi} Q^2 - 1. \tag{A26}$$

Equations (A24) and (A25) describe generalized Newtonian motion in the potential $V(\chi, \phi^i) = k_N^{-1} \chi - (1/2) e^{2\chi} Q^2(\phi^i)$ with metric $G_{IJ} = \mathrm{diag}(k_N^{-1}, G_{ij})$ and a linear drag force. A smooth horizon requires $\chi(\tau)$ and $\phi^i(\tau)$ to approach finite values χ_h and ϕ^i_h, respectively, as $\tau \to -\infty$, which can occur only at a critical point of $V(\chi, \phi^i)$, i.e.,

$$k_N e^{2\chi_h} Q^2(\phi_h) = 1, \qquad Q^2_{,i}(\phi_h) = 0, \tag{A27}$$

from which the constraint (A26) automatically follows. Thus, ϕ_h is a critical point of $Q^2(\phi)$ and $Q^2(\phi_h)$ determines the horizon area,

$$A_h = V_{d-2} R_h^{d-2} = V_{d-2} \left[\frac{\sqrt{k_N} Q(\phi_h)}{(d-3) V_{d-2}}\right]^{(d-2)/(d-3)}. \tag{A28}$$

This is the attractor mechanism (Ferrara, Kallosh, and Strominger, 1995; Cvetic and Tseytlin, 1996; Ferrara and Kallosh, 1996a, 1996b; Strominger, 1996). The trivial solution $\phi^i(z) = \phi^i_h$ is a Reissner-Nordström one with mass $M_0 = k_N^{-1/2} Q(\phi_h)$. All other solutions are strictly heavier since $\mathcal{W}(z) = k_N^{-1} (d/dz)(e^{-\psi})$ evaluates to $M_{\mathrm{BH}} = k_N^{-1}(-\dot\psi_\infty)$ and $M_0 = k_N^{-1} e^{-\chi_h}$ at $z = 0$ and $z = \infty$, respectively, and $\dot{\mathcal{W}} = -e^{-\psi} G_{ij}(\phi) \dot\phi^i \dot\phi^j \le 0$ per[25] Eq. (A15).

If ϕ^i_h is a local minimum of $Q^2(\phi)$ (the attractor point is "stable"), then (χ_h, ϕ^i_h) is a local maximum of $V(\chi, \phi^i)$ and we can roll off the hill in any direction. Hence, there are attractor solutions for any nearby choice of ϕ^i_∞. This is not necessarily the case farther from the attractor point, where the family of solutions to Eqs. (A24) and (A25) beginning at ϕ_h may encounter turning points and/or caustics.

Unstable critical points of $Q^2(\phi)$ also admit attractor solutions, but by the same reasoning these do not exist for generic values of ϕ^i_∞ and thus do not really play a role in determining the extremality bound.

c. Fake superpotentials and a Bogomol'nyi bound

Combining the preceeding observations, we see that there are families of quasiextremal solutions corresponding to each stable attractor point ϕ^i_h, with mass M determined by the choice of vacuum ϕ^i_∞. The resulting mass function $M = W(\phi_\infty)$ is also known as the fake superpotential (Andrianopoli *et al.*, 2007; Ceresole and Dall'Agata, 2007; Andrianopoli, D'Auria, Ferrara, and Trigiante, 2010; Andrianopoli, D'Auria, Orazi, and Trigiante, 2010; Trigiante, Riet, and Vercnocke, 2012) associated with the attractor point in question. Because each member of the family has the same charge and horizon area (entropy), we can identify $\partial W/\partial \phi^i$ as the scalar charge μ_i via the first law $\delta M = \mu_i \delta\phi^i_\infty + \Phi^A_h \delta Q_A + (1/\kappa^2) g_h \delta A_h$ (Gibbons, Kallosh, and Kol, 1996). Therefore, owing to the no-force condition (A23), the fake superpotential satisfies

$$k_N W^2(\phi) + G^{ij}(\phi) \partial_i W(\phi) \partial_j W(\phi) = Q^2(\phi), \tag{A29}$$

where $W(\phi)$ has a global minimum at ϕ^i_h. Solutions to the nonlinear first-order differential equation (A29) are in general highly nonunique. However, the condition that $W(\phi)$ has a minimum at $\phi = \phi_h$ is enough to fix this ambiguity, at least locally.

This can be shown using a Bogomol'nyi bound as follows. Consider any black hole solution $\psi(r), \phi^i(r)$, and $f(r)$ (not necessarily quasiextremal). The functional

[25] Thus, $k_N^{-1/2} Q(\phi_h) \le M \le k_N^{-1/2} Q(\phi_\infty)$.

$$I[\psi,\phi,f] := \int_0^{z_h}\left[\frac{1}{2k_N}(f\dot\psi+\dot f)^2+\frac{1}{2}f^2G_{ij}(\phi)\dot\phi^i\dot\phi^j + \frac{1}{2}e^{2\psi}Q^2(\phi)\right]dz \tag{A30}$$

evaluates to the black hole mass upon imposing the equations of motion,

$$I[\psi,\phi,f] = k_N^{-1}\left[\int_0^{z_h}\frac{d}{dz}\left(\frac{1+f}{2}f\dot\psi\right)dz+\frac{1}{2z_h}\right] = k_N^{-1}\left[-\dot\psi_\infty+\frac{1}{2z_h}\right] = M_{\rm BH}, \tag{A31}$$

where the horizon boundary term vanishes in the quasiextremal ($f=1$) case because $\ddot\psi \geq \dot\psi^2$ and $\dot\psi \leq 0$ from Eqs. (A15) and (A16) imply $\dot\psi \to 0$ as $z\to\infty$.

Given any function $W(\phi)$ satisfying $k_N W(\phi)^2 + G^{ij}W_{,i}W_{,j} \leq Q^2(\phi)$ along the entire trajectory $\phi^i(r)$, I can be factored as follows:

$$I[\psi,\phi,f] = \int_0^{z_h}\left(\frac{k_N}{2}\left[\frac{f\dot\psi+\dot f}{k_N}+e^\psi W(\phi)\right]^2 + \frac{1}{2}G_{ij}[f\dot\phi^i+e^\psi G^{ik}W_{,k}][f\dot\phi^j+e^\psi G^{jl}W_{,l}]\right)dz + \frac{1}{2}\int_0^{z_h}e^{2\psi}[Q^2(\phi)-k_N W(\phi)^2 - G^{ij}W_{,i}W_{,j}]dz + W(\phi_\infty), \tag{A32}$$

where we use $\int_0^{z_h}(d/dz)[fe^\psi W(\phi)]dz = -W(\phi_\infty)$ with the horizon boundary term vanishing in the quasiextremal ($f=1$) case because $e^\psi W(\phi) \leq k_N^{-1/2}e^\psi|Q(\phi)| = k_N^{-1}\sqrt{\ddot\psi}$, $\ddot\psi \geq \dot\psi^2$, and $\dot\psi\leq 0$ likewise imply $\ddot\psi\to 0$ as $z\to\infty$. Every term in Eq. (A32) but the last is positive definite, so we conclude that

$$M_{\rm BH} \geq W(\phi_\infty). \tag{A33}$$

Saturating the bound requires Eq. (A29) along the trajectory together with $f=1$ (quasiextremality)[26] and the Bogomol'nyi equations

$$\dot\psi = -e^\psi k_N W(\phi), \qquad \dot\phi^i = -e^\psi G^{ij}W_{,j}. \tag{A34}$$

These equations, which imply Eqs. (A10)–(A12), have a unique solution for each choice of ϕ^i_∞ provided that the $W(\phi)$ gradient flow remains entirely within the region R_0 where Eq. (A29) is satisfied, and the solution has a smooth horizon provided that the gradient flow ends at a critical point of $W(\phi)$ with $W(\phi_{\rm crit})>0$ (ensuring that $e^\psi \propto 1/z$ as $z\to\infty$). Thus, given $W(\phi)$ and ϕ^i_∞, the black hole solution saturating Eq. (A33) is unique if it exists.

If both $W_1(\phi)$ and $W_2(\phi)$ satisfy Eq. (A29) in a region R_0 encompassing coincident local minima[27] at $\phi=\phi_h$, W_1 gradient flows ending at ϕ_h produce quasiextremal solutions of mass $M_{\rm BH}=W_1(\phi_\infty)$, which must satisfy $M_{\rm BH} = W_1(\phi_\infty)\geq W_2(\phi_\infty)$ per Eq. (A33). By the same token $W_2(\phi_\infty)\geq W_1(\phi_\infty)$, hence $W_1(\phi_\infty)=W_2(\phi_\infty)$ for all ϕ_∞ in R_0 flowing to ϕ_h.

Thus, the fake superpotential $W(\phi)$ associated with a given stable attractor point ϕ^i_h is uniquely fixed near the attractor point by Eq. (A29) and the condition that $W(\phi)$ has a minimum at ϕ^i_h, and the corresponding attractor solutions can be obtained from $W(\phi)$ by solving the gradient flow equations (A34). [The family of attractor solutions beginning at ϕ^i_h may have turning points and/or caustics farther from the attractor point, so $W(\phi)$ does not necessarily extend uniquely throughout moduli space.]

d. Asymptotic attractors

A special case of the attractor mechanism occurs when ϕ^i_h lies at infinite distance in the moduli space, usually in a direction[28] where $Q^2(\phi)\to 0$. Such "asymptotic attractors" technically do not lead to smooth black hole solutions [for instance, $Q^2(\phi_h)=0$ implies vanishing horizon area] but can be understood as the limit of a family of smooth nonextremal solutions, and therefore play a role in determining the extremality bound.

Asymptotic attractors are typically also characterized by a fake superpotential. Heuristically, because modifying $Q^2(\phi)$ far out in the moduli space can turn the asymptotic attractor into a standard attractor at finite distance with an associated unique fake superpotential, taking a limit where the new attractor point is sent off to infinity, while restoring $Q^2(\phi)$ to its original form should yield a fake superpotential for the original asymptotic attractor. This argument could fail in several ways when the asymptotic behavior of $Q^2(\phi)$ is sufficiently strange, but within the realm of actual quantum gravities we know of no issue with it.

3. The extremality bound

Cosmic censorship requires extremal black holes to be quasiextremal, and therefore for a given choice of Q_A and ϕ^i_∞ the lightest quasiextremal black hole should be extremal. Combined with the discussion of fake superpotentials in Appendix 2, this suggests that the extremality bound is given by

$$M_{\rm BH} \geq W(\phi) := \min_{\{a\}} W_a(\phi), \tag{A35}$$

where $W_a(\phi)$ are the fake superpotentials associated with the various stable or asymptotic attractors and the minimum is taken among all the fake superpotentials defined at the point in question.

[26]The Bogomol'nyi equations $f\dot\psi+\dot f = -e^\psi k_N W$, $f\dot\phi^i = -e^\psi G^{ij}W_{,j}$, and Eq. (A29) imply $(d/dz)[fe^\psi W] = -e^{2\psi}Q^2(\phi)$, whereas $(d/dz)[e^\psi W] = -k_N^{-1}(d/dz)[f\dot\psi+\dot f] = -e^{2\psi}Q^2$ using Eq. (A11). Thus, $(d/dz)[(1-f)e^\psi W]=0$, implying $(1-f)e^\psi W = 0$ by integrating from $z=0$. Since $e^\psi W>0$, $f=1$ follows.

[27]This implies that $Q^2(\phi)$ also has a local minimum at ϕ_h.

[28]In principle, ϕ^i_h could lie at an infinite distance point where $Q^2(\phi_h)>0$, but this does not occur anywhere in the landscape to our knowledge.

To verify Eq. (A35), we assume that the "global fake superpotential" $W(\phi)$ specified in Eq. (A35) is defined and continuous everywhere in moduli space. Thus, the moduli space is partitioned into different "attractor basins" associated with the various stable or asymptotic attractors, with $W(\phi)$ equal to the corresponding fake superpotential $W_a(\phi)$ within each basin. Since each constituent fake superpotential $W_a(\phi)$ satisfies Eq. (A29), $W(\phi)$ also satisfies Eq. (A29), except possibly at the boundaries between attractor basins. However, as these boundaries are sets of measure zero and the continuity of $W(\phi)$ precludes delta-function contributions to $\nabla_i W$, the argument leading to the Bogomol'nyi bound (A33) is unaffected; hence, $M_{\rm BH} \geq W(\phi)$. As this bound can be saturated by construction, it is indeed the extremality bound.

Naively, to construct $W(\phi)$ we must first classify all quasiextremal solutions. However, this is not the case: given any possibly incomplete collection of local fake superpotentials $W_a(\phi)$ such that $\hat{W}(\phi) \coloneqq \min_{\{a\}} W_a(\phi)$ is everywhere defined and continuous, the same reasoning as previously employed implies that $M_{\rm BH} \geq \hat{W}(\phi)$ is the extremality bound; hence, $\hat{W}(\phi) = W(\phi)$.

4. Examples

A simple example that has played an outsized role in the development of the WGC is Einstein-Maxwell dilaton theory, with the effective action

$$S = \int d^d x \sqrt{-g}\left[\frac{1}{2\kappa^2}\left(R - \frac{1}{2}(\nabla\phi)^2\right) - \frac{1}{2\hat{e}^2} e^{-\alpha\phi}|F_2|^2\right]. \tag{A36}$$

Equation (A29) then becomes

$$\xi\kappa^2[W(\phi)]^2 + 2\kappa^2[W'(\phi)]^2 = e^{\alpha\phi}(\hat{e}Q)^2, \qquad \xi \coloneqq \frac{d-3}{d-2}. \tag{A37}$$

There is an asymptotic attractor at $\phi_h = -\infty$. Guessing a solution of the form $W(\phi) = \hat{M} e^{\alpha\phi/2}$, we obtain

$$\left[\xi + \frac{\alpha^2}{2}\right](\kappa\hat{M})^2 = (\hat{e}Q)^2,$$
$$\Rightarrow \kappa\hat{M} = \gamma^{-(1/2)}\hat{e}|Q|, \qquad \gamma = \xi + \frac{\alpha^2}{2}. \tag{A38}$$

Since $W(\phi)$ is globally defined, is positive, and satisfies Eq. (A29) everywhere, it defines a global fake superpotential and the extremality bound is

$$\kappa M_{\rm BH} \geq \kappa W(\phi_\infty) = \gamma^{-1/2} e|Q|, \tag{A39}$$

where $e = \hat{e} e^{\alpha\phi_\infty/2}$ is the vacuum gauge coupling. The same result applies to the corresponding $(p-1)$-brane theory with $\xi \to \xi_{(p)} = p(d-p-2)/(d-2)$, as in Eqs. (12) and (13).

As a somewhat less trivial example, consider the following two gauge fields with different dilaton couplings:

$$S = \int d^d x \sqrt{-g}\left[\frac{1}{2\kappa^2}\left(R - \frac{1}{2}(\nabla\phi)^2\right) - \frac{1}{2\hat{e}_1^2} e^{-\alpha_1\phi}|F_2|^2 - \frac{1}{2\hat{e}_2^2} e^{\alpha_2\phi}|H_2|^2\right]. \tag{A40}$$

Equation (A29) then becomes

$$\xi\kappa^2[W(\phi)]^2 + 2\kappa^2[W'(\phi)]^2 = e^{\alpha_1\phi}(\hat{e}_1 Q_1)^2 + e^{-\alpha_2\phi}(\hat{e}_2 Q_2)^2. \tag{A41}$$

Provided that $\alpha_1\alpha_2 > 0$ and $Q_{1,2} \neq 0$, there is a stable attractor point at $\phi_h = [1/(\alpha_1+\alpha_2)]\log[\alpha_2(\hat{e}_2 Q_2)^2/\alpha_1(\hat{e}_1 Q_1)^2]$.

Guessing a solution of the form $W(\phi) = \hat{M}_1 e^{\alpha_1\phi/2} + \hat{M}_2 e^{-\alpha_2\phi/2}$, the left-hand side of Eq. (A41) has cross terms proportional to $e^{(\alpha_1-\alpha_2)\phi}$, whose cancellation requires $\alpha_1\alpha_2 = 2\xi$. With this condition, we obtain

$$\kappa\hat{M}_{1,2} = \gamma_{1,2}^{-1/2}\hat{e}_{1,2}|Q_{1,2}|, \qquad \gamma_{1,2} = \xi + \frac{\alpha_{1,2}^2}{2}, \tag{A42}$$

so the extremality bound in this case is

$$\kappa M_{\rm BH} \geq \gamma_1^{-1/2} e_1|Q_1| + \gamma_2^{-1/2} e_2|Q_2|, \tag{A43}$$

where $e_1 = \hat{e}_1 e^{\alpha_1\phi_\infty/2}$ and $e_2 = \hat{e}_2 e^{-\alpha_2\phi_\infty/2}$ are the gauge couplings in the vacuum in question. When $\alpha_1\alpha_2 \neq 2\xi$, the fake superpotential solving Eq. (A41) is not known in closed form (apart from some special cases) but is easily found using numerical integration.

Thus far we have considered examples with a single attractor basin. A simple (if contrived) example that exhibits multiple attractor basins is

$$Q^2(\phi) = k_N M_0^2\{1 + [(\phi/\phi_0)^2 - \lambda]^2\}, \qquad G_{\phi\phi} = k_N^{-1},$$
$$\Rightarrow W(\phi)^2 + W'(\phi)^2 = M_0^2\{1 + [(\phi/\phi_0)^2 - \lambda]^2\}, \tag{A44}$$

with attractor points at $\phi = \pm\sqrt{\lambda}\phi_0$. The associated fake superpotentials can be found by numerical integration; see Figs. 19 and 20. Note that for some values of ϕ_0 and λ there are quasiextremal solutions that are not extremal, due to finite overlap between the domains of the local fake superpotentials $W_\pm(\phi)$ associated with the attractor points $\phi = \pm\sqrt{\lambda}\phi_0$. Some examples where numerical integration was used to determine the extremality bound in an actual theory of quantum gravity were discussed by Alim, Heidenreich, and Rudelius (2021).

5. Closing comments

To derive the extremality bound (A35), we assumed that $W(\phi) \coloneqq \min_{\{a\}} W_a(\phi)$ was everywhere defined and continuous. This is easily proved for a one-dimensional moduli space since (1) the domain $\mathcal{D}_a$ of the fake superpotential $W_a(\phi)$ associated with a minimum $\phi_{(a)}$ of $Q^2(\phi)$ is at least as large as the interval between the two adjacent maxima of $Q^2(\phi)$ and (2) $W_a(\phi) = |Q(\phi)|/\sqrt{k_N}$ at the boundary of this domain (see Fig. 19), saturating the upper bound $W_b(\phi) \leq |Q(\phi)|/\sqrt{k_N}$ on

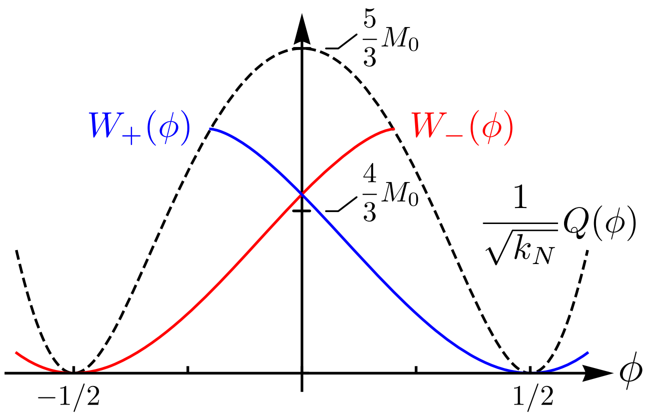

FIG. 19. Fake superpotentials for the two stable attractor points of Eq. (A44) with $\lambda = 4/3$ and $\phi_0 = \sqrt{3}/4$. Because $W_\pm(\phi)$ cross over at $\phi = 0$, $W_-(\phi)$ gradient flow solutions with $\phi_\infty > 0$ are quasiextremal but not extremal, as are $W_+(\phi)$ gradient flow solutions with $\phi_\infty < 0$.

all fake superpotentials, where these two properties ensure (1) the existence and (2) the continuity of $W(\phi)$ at each point. While we do not know a general proof for higher-dimensional moduli spaces, the existence and continuity of $W(\phi)$ can be verified on a case-by-case basis.

Phase transitions in the moduli space can create additional subtleties. While $t_{AB}(\phi)$ and $G_{ij}(\phi)$ need not be analytic at a phase transition, in itself this has little effect on the foregoing analysis. More importantly different branches of moduli space can meet at a phase transition, opening up the possibility of black hole solutions that cross from one branch to another. This is a rather complicated question that has not been worked out in the literature to our knowledge, but it seems probable that some version of fake superpotentials will still be applicable. Yet more drastically, the moduli space could have finite-distance boundaries where a strongly coupled CFT appears, as in some examples given by Alim, Heidenreich, and Rudelius (2021). Black hole solutions that reach this CFT boundary outside their event horizon lie outside the regime of validity of the weakly coupled EFT that our analysis is based on and will require a separate analysis.

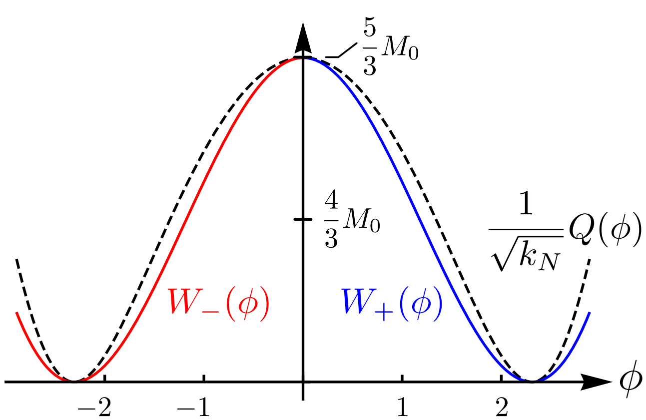

FIG. 20. Fake superpotentials for the two stable attractor points of Eq. (A44) with $\lambda = 4/3$ and $\phi_0 = 2$. In this case, all quasiextremal solutions are extremal.

REFERENCES

Aalsma, Lars, 2022, "Corrections to extremal black holes from Iyer-Wald formalism," Phys. Rev. D **105**, 066022.

Aalsma, Lars, Alex Cole, Gregory J. Loges, and Gary Shiu, 2021, "A new spin on the weak gravity conjecture," J. High Energy Phys. 03, 085.

Aalsma, Lars, Alex Cole, and Gary Shiu, 2019, "Weak gravity conjecture, black hole entropy, and modular invariance," J. High Energy Phys. 08, 022.

Abbott, L. F., and Mark B. Wise, 1989, "Wormholes and global symmetries," Nucl. Phys. **B325**, 687–704.

Abrikosov, A. A., 1957, "On the magnetic properties of superconductors of the second group," Sov. Phys. JETP **5**, 1174–1182.

Ackerman, Lotty, Matthew R. Buckley, Sean M. Carroll, and Marc Kamionkowski, 2009, "Dark matter and dark radiation," Phys. Rev. D **79**, 023519.

Adams, Allan, Nima Arkani-Hamed, Sergei Dubovsky, Alberto Nicolis, and Riccardo Rattazzi, 2006, "Causality, analyticity and an IR obstruction to UV completion," J. High Energy Phys. 10, 014.

Ade, P. A. R., *et al.* (BICEP and Keck Collaborations), 2021, "Improved Constraints on Primordial Gravitational Waves Using Planck, WMAP, and BICEP/Keck Observations through the 2018 Observing Season," Phys. Rev. Lett. **127**, 151301.

Adelberger, Eric, Gia Dvali, and Andrei Gruzinov, 2007, "Photon Mass Bound Destroyed by Vortices," Phys. Rev. Lett. **98**, 010402.

Afkhami-Jeddi, Nima, Sandipan Kundu, and Amirhossein Tajdini, 2018, "A conformal collider for holographic CFTs," J. High Energy Phys. 10, 156.

Aghanim, N., *et al.* (Planck Collaboration), 2020, "Planck 2018 results. VI. Cosmological parameters," Astron. Astrophys. **641**, A6.

Aharony, Ofer, and Eran Palti, 2021, "Convexity of charged operators in CFTs and the weak gravity conjecture," Phys. Rev. D **104**, 126005.

Akrami, Y., *et al.* (Planck Collaboration), 2020, "Planck 2018 results. X. Constraints on inflation," Astron. Astrophys. **641**, A10.

Alberte, Lasma, Claudia de Rham, Sumer Jaitly, and Andrew J. Tolley, 2020, "Positivity bounds and the massless spin-2 pole,"-Phys. Rev. D **102**, 125023.

Alberte, Lasma, Claudia de Rham, Sumer Jaitly, and Andrew J. Tolley, 2021, "QED positivity bounds," Phys. Rev. D **103**, 125020.

Alim, Murad, Ben Heidenreich, and Tom Rudelius, 2021, "The weak gravity conjecture and BPS particles," Fortschr. Phys. **69**, 2100125.

Almheiri, Ahmed, Netta Engelhardt, Donald Marolf, and Henry Maxfield, 2019, "The entropy of bulk quantum fields and the entanglement wedge of an evaporating black hole," J. High Energy Phys. 12, 063.

Almheiri, Ahmed, Thomas Hartman, Juan Maldacena, Edgar Shaghoulian, and Amirhossein Tajdini, 2020, "Replica wormholes and the entropy of Hawking radiation," J. High Energy Phys. 05, 013.

Almheiri, Ahmed, and Byungwoo Kang, 2016, "Conformal symmetry breaking and thermodynamics of near-extremal black holes," J. High Energy Phys. 10, 052.

Álvarez-García, Rafael, Daniel Kläwer, and Timo Weigand, 2022, "Membrane limits in quantum gravity," Phys. Rev. D **105**, 066024.

Andrianopoli, L., R. D'Auria, S. Ferrara, and M. Trigiante, 2010, "Fake superpotential for large and small extremal black holes," J. High Energy Phys. 08, 126.

Andrianopoli, L., R. D'Auria, E. Orazi, and M. Trigiante, 2010, "First order description of $D = 4$ static black holes and the Hamilton-Jacobi equation," Nucl. Phys. **B833**, 1–16.

Andrianopoli, Laura, Riccardo D'Auria, Emanuele Orazi, and Mario Trigiante, 2007, "First order description of black holes in moduli space," J. High Energy Phys. 11, 032.

Andriolo, Stefano, Tzu-Chen Huang, Toshifumi Noumi, Hirosi Ooguri, and Gary Shiu, 2020, "Duality and axionic weak gravity," Phys. Rev. D **102**, 046008.

Andriolo, Stefano, Daniel Junghans, Toshifumi Noumi, and Gary Shiu, 2018, "A tower weak gravity conjecture from infrared consistency," Fortschr. Phys. **66**, 1800020.

Antipin, Oleg, Jahmall Bersini, Francesco Sannino, Zhi-Wei Wang, and Chen Zhang, 2021, "More on the weak gravity conjecture via convexity of charged operators," J. High Energy Phys. 12, 204.

Arkani-Hamed, Nima, Hsin-Chia Cheng, Paolo Creminelli, and Lisa Randall, 2003, "Extra Natural Inflation," Phys. Rev. Lett. **90**, 221302.

Arkani-Hamed, Nima, Savas Dimopoulos, and G. R. Dvali, 1998, "The hierarchy problem and new dimensions at a millimeter," Phys. Lett. B **429**, 263–272.

Arkani-Hamed, Nima, Savas Dimopoulos, and Shamit Kachru, 2005, "Predictive landscapes and new physics at a TeV," arXiv:hep-th/0501082.

Arkani-Hamed, Nima, Sergei Dubovsky, Alberto Nicolis, and Giovanni Villadoro, 2007, "Quantum horizons of the standard model landscape," J. High Energy Phys. 06, 078.

Arkani-Hamed, Nima, Tzu-Chen Huang, and Yu-Tin Huang, 2021, "The EFT-hedron," J. High Energy Phys. 05, 259.

Arkani-Hamed, Nima, Yu-tin Huang, Jin-Yu Liu, and Grant N. Remmen, 2022, "Causality, unitarity, and the weak gravity conjecture," J. High Energy Phys. 03, 083.

Arkani-Hamed, Nima, Lubos Motl, Alberto Nicolis, and Cumrun Vafa, 2007, "The string landscape, black holes and gravity as the weakest force," J. High Energy Phys. 06, 060.

Bachlechner, Thomas C., Cody Long, and Liam McAllister, 2016, "Planckian axions and the weak gravity conjecture," J. High Energy Phys. 01, 091.

Banks, Tom, Michael Dine, Patrick J. Fox, and Elie Gorbatov, 2003, "On the possibility of large axion decay constants," J. Cosmol. Astropart. Phys. 06, 001.

Banks, Tom, and Lance J. Dixon, 1988, "Constraints on string vacua with space-time supersymmetry," Nucl. Phys. **B307**, 93–108.

Banks, Tom, Matt Johnson, and Assaf Shomer, 2006, "A note on gauge theories coupled to gravity," J. High Energy Phys. 09, 049.

Banks, Tom, and Nathan Seiberg, 2011, "Symmetries and strings in field theory and gravity," Phys. Rev. D **83**, 084019.

Bass, Ludvik, and Erwin Schrödinger, 1955, "Must the photon mass be zero?," Proc. R. Soc. A **232**, 1–6. https://www.jstor.org/stable/99679.

Baumann, Daniel, and Liam McAllister, 2015, *Inflation and String Theory*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England).

Baume, Florent, and José Calderón Infante, 2021, "Tackling the SDC in AdS with CFTs," J. High Energy Phys. 08, 057.

Bekenstein, Jacob D., 1981a, "Energy Cost of Information Transfer," Phys. Rev. Lett. **46**, 623–626.

Bekenstein, Jacob D., 1981b, "Universal upper bound on the entropy-to-energy ratio for bounded systems," Phys. Rev. D **23**, 287–298.

Belavin, A. A., Alexander M. Polyakov, A. S. Schwartz, and Yu. S. Tyupkin, 1975, "Pseudoparticle solutions of the Yang-Mills equations," Phys. Lett. **59B**, 85–87.

Bellazzini, Brando, Clifford Cheung, and Grant N. Remmen, 2016, "Quantum gravity constraints from unitarity and analyticity," Phys. Rev. D **93**, 064076.

Bellazzini, Brando, Giulia Isabella, Matthew Lewandowski, and Francesco Sgarlata, 2022, "Gravitational causality and the self-stress of photons," J. High Energy Phys. 05, 154.

Bellazzini, Brando, Matthew Lewandowski, and Javi Serra, 2019, "Positivity of Amplitudes, Weak Gravity Conjecture, and Modified Gravity," Phys. Rev. Lett. **123**, 251103.

Benakli, Karim, Carlo Branchina, and Gaëtan Lafforgue-Marmet, 2020, "$U(1)$ mixing and the weak gravity conjecture," Eur. Phys. J. C **80**, 1118.

Benjamin, Nathan, Ethan Dyer, A. Liam Fitzpatrick, and Shamit Kachru, 2016, "Universal bounds on charged states in 2D CFT and 3D gravity," J. High Energy Phys. 08, 041.

Bergshoeff, E., Andres Collinucci, U. Gran, D. Roest, and S. Vandoren, 2005, "Non-extremal instantons and wormholes in string theory," Fortschr. Phys. **53**, 990–996.

Bergshoeff, E., Andres Collinucci, U. Gran, D. Roest, and S. Vandoren, 2004, "Non-extremal D-instantons," J. High Energy Phys. 10, 031.

Bobev, Nikolay, Anthony M. Charles, Kiril Hristov, and Valentin Reys, 2021, "Higher-derivative supergravity, AdS_4 holography, and black holes," J. High Energy Phys. 08, 173.

Bousso, Raphael, 2002, "The holographic principle," Rev. Mod. Phys. **74**, 825–874.

Bowick, Mark J., Steven B. Giddings, Jeffrey A. Harvey, Gary T. Horowitz, and Andrew Strominger, 1988, "Axionic Black Holes and a Bohm-Aharonov Effect for Strings," Phys. Rev. Lett. **61**, 2823.

Braun, Andreas P., Cody Long, Liam McAllister, Michael Stillman, and Benjamin Sung, 2020, "The Hodge numbers of divisors of Calabi-Yau threefold hypersurfaces," Fortschr. Phys. **68**, 2000087.

Breitenlohner, Peter, Dieter Maison, and Gary W. Gibbons, 1988, "Four-dimensional black holes from Kaluza-Klein theories," Commun. Math. Phys. **120**, 295.

Brennan, T. Daniel, Federico Carta, and Cumrun Vafa, 2018, "The string landscape, the swampland, and the missing corner," Proc. Sci. TASI2017, 015 [arXiv:1711.00864].

Brennan, T. Daniel, and Clay Cordova, 2022, "Axions, higher-groups, and emergent symmetry," J. High Energy Phys. 02 145.

Brown, Jon, William Cottrell, Gary Shiu, and Pablo Soler, 2015, "Fencing in the swampland: Quantum gravity constraints on large field inflation," J. High Energy Phys. 10, 023.

Brown, Jon, William Cottrell, Gary Shiu, and Pablo Soler, 2016, "On axionic field ranges, loopholes and the weak gravity conjecture," J. High Energy Phys. 04, 017.

Buen-Abad, Manuel A., Gustavo Marques-Tavares, and Martin Schmaltz, 2015, "Non-Abelian dark matter and dark radiation," Phys. Rev. D **92**, 023531.

Callan, Jr., Curtis G., and Jeffrey A. Harvey, 1985, "Anomalies and fermion zero modes on strings and domain walls," Nucl. Phys. **B250**, 427–436.

Camanho, Xian O., Jose D. Edelstein, Juan Maldacena, and Alexander Zhiboedov, 2016, "Causality constraints on corrections to the graviton three-point coupling," J. High Energy Phys. 02, 020.

Cano, Pablo A., Samuele Chimento, Román Linares, Tomás Ortín, and Pedro F. Ramírez, 2020, "α' corrections of Reissner-Nordström black holes," J. High Energy Phys. 02 031.

Cano, Pablo A., Tomás Ortín, and Pedro F. Ramirez, 2020, "On the extremality bound of stringy black holes," J. High Energy Phys. 02, 175.

Caron-Huot, Simon, Dalimil Mazac, Leonardo Rastelli, and David Simmons-Duffin, 2021, "Sharp boundaries for the swampland," J. High Energy Phys. 07, 110.

Casini, Horacio, Marina Huerta, Javier M. Magan, and Diego Pontello, 2021, "Entropic order parameters for the phases of QFT," J. High Energy Phys. 04, 277.
Casini, Horacio, Marina Huerta, Javier M. Magán, and Diego Pontello, 2020, "Entanglement entropy and superselection sectors. Part I. Global symmetries," J. High Energy Phys. 02, 014.
Casini, Horacio, and Javier M. Magan, 2021, "On completeness and generalized symmetries in quantum field theory," Mod. Phys. Lett. A **36**, 2130025.
Ceresole, Anna, and Gianguido Dall'Agata, 2007, "Flow equations for non-BPS extremal black holes," J. High Energy Phys. 03, 110.
Chanowitz, Michael S., John R. Ellis, and Mary K. Gaillard, 1977, "The price of natural flavor conservation in neutral weak interactions," Nucl. Phys. **B128**, 506–536.
Charles, Anthony M., 2019, "The weak gravity conjecture, RG flows, and supersymmetry," arXiv:1906.07734.
Charles, Anthony M., and Finn Larsen, 2016, "Kerr-Newman black holes with string corrections," J. High Energy Phys. 10, 142.
Chen, Yiming, and Henry W. Lin, 2021, "Signatures of global symmetry violation in relative entropies and replica wormholes," J. High Energy Phys. 03, 040.
Cheung, Clifford, Junyu Liu, and Grant N. Remmen, 2018, "Proof of the weak gravity conjecture from black hole entropy," J. High Energy Phys. 10, 004.
Cheung, Clifford, Junyu Liu, and Grant N. Remmen, 2019, "Entropy bounds on effective field theory from rotating dyonic black holes," Phys. Rev. D **100**, 046003.
Cheung, Clifford, and Grant N. Remmen, 2014a, "Infrared consistency and the weak gravity conjecture," J. High Energy Phys. 12, 087.
Cheung, Clifford, and Grant N. Remmen, 2014b, "Naturalness and the Weak Gravity Conjecture," Phys. Rev. Lett. **113**, 051601.
Cho, Inyong, 1998, "Inflation and nonsingular space-times of cosmic strings," Phys. Rev. D **58**, 103509.
Choi, Kiwoon, and Sang Hui Im, 2016, "Realizing the relaxion from multiple axions and its UV completion with high scale supersymmetry," J. High Energy Phys. 01, 149.
Choi, Kiwoon, Hyungjin Kim, and Seokhoon Yun, 2014, "Natural inflation with multiple sub-Planckian axions," Phys. Rev. D **90**, 023545.
Cohen, Andrew G., and David B. Kaplan, 1988, "The exact metric about global cosmic strings," Phys. Lett. B **215**, 67–72.
Coleman, Sidney R., and Ki-Myeong Lee, 1990, "Wormholes made without massless matter fields," Nucl. Phys. **B329**, 387–409.
Coleman, Sidney R., and Frank De Luccia, 1980, "Gravitational effects on and of vacuum decay," Phys. Rev. D **21**, 3305.
Conlon, Joseph P., and Sven Krippendorf, 2016, "Axion decay constants away from the lamppost," J. High Energy Phys. 04, 085.
Córdova, Clay, Thomas T. Dumitrescu, and Kenneth Intriligator, 2019, "Exploring 2-group global symmetries," J. High Energy Phys. 02, 184.
Córdova, Clay, Kantaro Ohmori, and Tom Rudelius, 2022, "Generalized symmetry breaking scales and weak gravity conjectures," arXiv:2202.05866.
Corvilain, Pierre, Thomas W. Grimm, and Irene Valenzuela, 2019, "The swampland distance conjecture for Kähler moduli," J. High Energy Phys. 08, 075.
Cota, Cesar Fierro, Albrecht Klemm, and Thorsten Schimannek, 2021, "State counting on fibered CY 3-folds and the non-Abelian weak gravity conjecture," J. High Energy Phys. 05, 030.
Craig, Nathaniel, and Isabel Garcia Garcia, 2018, "Rescuing massive photons from the swampland," J. High Energy Phys. 11, 067.
Craig, Nathaniel, Isabel Garcia Garcia, and Seth Koren, 2019a, "Discrete gauge symmetries and the weak gravity conjecture," J. High Energy Phys. 05, 140.
Craig, Nathaniel, Isabel Garcia Garcia, and Seth Koren, 2019b, "The weak scale from weak gravity," J. High Energy Phys. 09, 081.
Craig, Nathaniel, Isabel Garcia Garcia, and Graham D. Kribs, 2020, "The UV fate of anomalous U(1)s and the swampland," J. High Energy Phys. 11, 063.
Cremonini, Sera, Callum R. T. Jones, James T. Liu, and Brian McPeak, 2020, "Higher-derivative corrections to entropy and the weak gravity conjecture in anti–de Sitter space," J. High Energy Phys. 09, 003.
Cremonini, Sera, Callum R. T. Jones, James T. Liu, Brian McPeak, and Yuezhang Tang, 2021, "NUT charge weak gravity conjecture from dimensional reduction," Phys. Rev. D **103**, 106011.
Cremonini, Sera, Callum R. T. Jones, James T. Liu, Brian McPeak, and Yuezhang Tang, 2022, "Repulsive black holes and higher-derivatives," J. High Energy Phys. 03, 013.
Cribiori, N., G. Dall'Agata, and F. Farakos, 2021, "Weak gravity versus de Sitter," J. High Energy Phys. 04, 046.
Crisford, Toby, Gary T. Horowitz, and Jorge E. Santos, 2018, "Testing the weak gravity–cosmic censorship connection," Phys. Rev. D **97**, 066005.
Crisford, Toby, and Jorge E. Santos, 2017, "Violating the Weak Cosmic Censorship Conjecture in Four-Dimensional Anti–de Sitter Space," Phys. Rev. Lett. **118**, 181101.
Cvetic, Mirjam, and Arkady A. Tseytlin, 1996, "Solitonic strings and BPS saturated dyonic black holes," Phys. Rev. D **53**, 5619–5633.
Dall'Agata, Gianguido, Maxim Emelin, Fotis Farakos, and Matteo Morittu, 2021, "The unbearable lightness of charged gravitini," J. High Energy Phys. 10, 076.
D'Amico, Guido, Nemanja Kaloper, and Alexander Westphal, 2021, "Double monodromy inflation: A gravity waves factory for CMB-S4, LiteBIRD and LISA," Phys. Rev. D **104**, L081302.
D'Amico, Guido, Nemanja Kaloper, and Alexander Westphal, 2022, "General double monodromy inflation," Phys. Rev. D **105**, 103527.
Daus, Tristan, Arthur Hebecker, Sascha Leonhardt, and John March-Russell, 2020, "Towards a swampland global symmetry conjecture using weak gravity," Nucl. Phys. **B960**, 115167.
de la Fuente, Anton, Prashant Saraswat, and Raman Sundrum, 2015, "Natural Inflation and Quantum Gravity," Phys. Rev. Lett. **114**, 151303.
Demirtas, Mehmet, Cody Long, Liam McAllister, and Mike Stillman, 2020, "Minimal surfaces and weak gravity," J. High Energy Phys. 03, 021.
Demirtas, Mehmet, Andres Rios-Tascon, and Liam McAllister, 2022, "`CYTools`: A software package for analyzing Calabi-Yau manifolds," arXiv:2211.03823.
Denef, Frederik, and Sean A. Hartnoll, 2009, "Landscape of superconducting membranes," Phys. Rev. D **79**, 126008.
de Rham, Claudia, Scott Melville, Andrew J. Tolley, and Shuang-Yong Zhou, 2017, "Positivity bounds for scalar field theories," Phys. Rev. D **96**, 081702.
de Rham, Claudia, Scott Melville, Andrew J. Tolley, and Shuang-Yong Zhou, 2018, "UV complete me: Positivity bounds for particles with spin," J. High Energy Phys. 03, 011.
de Rham, Claudia, and Andrew J. Tolley, 2020, "Causality in curved spacetimes: The speed of light and gravity," Phys. Rev. D **102**, 084048.
Deshpande, Nilendra G., and David Iskandar, 1980, "Study of the gauge group $\mathrm{SU}(2)_L \times T(3)_R \times \mathrm{U}(1)_V$," Nucl. Phys. **B167**, 223–236.

Detweiler, Steven L., 1980, "Black holes and gravitational waves III. The resonant frequencies of rotating holes," Astrophys. J. **239**, 292–295.

Dimopoulos, S., S. Kachru, J. McGreevy, and Jay G. Wacker, 2008, "*N*-flation," J. Cosmol. Astropart. Phys. 08, 003.

Distler, Jacques, and Uday Varadarajan, 2005, "Random polynomials and the friendly landscape," arXiv:hep-th/0507090.

Dolan, Matthew J., Patrick Draper, Jonathan Kozaczuk, and Hiren Patel, 2017, "Transplanckian censorship and global cosmic strings," J. High Energy Phys. 04, 133.

Draper, Patrick, and Szilard Farkas, 2019, "Gravitational instabilities and censorship of large scalar field excursions," J. High Energy Phys. 05, 158.

Dvali, Gia, 2010, "Black holes and large *N* species solution to the hierarchy problem," Fortschr. Phys. **58**, 528–536.

Dvali, Gia, and Michele Redi, 2008, "Black hole bound on the number of species and quantum gravity at LHC," Phys. Rev. D **77**, 045027.

Dvali, Gia, Michele Redi, Sergey Sibiryakov, and Arkady Vainshtein, 2008, "Gravity Cutoff in Theories with Large Discrete Symmetries," Phys. Rev. Lett. **101**, 151603.

Emparan, Roberto, 2020, "Predictivity lost, predictivity regained: A Miltonian cosmic censorship conjecture," Int. J. Mod. Phys. D **29**, 2043021.

Engelhardt, Netta, and Aron C. Wall, 2014, "Extremal surface barriers," J. High Energy Phys. 03, 068.

Etheredge, Muldrow, and Ben Heidenreich, 2022, "Derivative corrections to extremal black holes with moduli," arXiv:2211.09823.

Fan, JiJi, Katherine Fraser, Matthew Reece, and John Stout, 2021, "Axion Mass from Magnetic Monopole Loops," Phys. Rev. Lett. **127**, 131602.

Faulkner, Thomas, Hong Liu, John McGreevy, and David Vegh, 2011, "Emergent quantum criticality, Fermi surfaces, and AdS(2)," Phys. Rev. D **83**, 125002.

Feng, Jonathan L., Manoj Kaplinghat, Huitzu Tu, and Hai-Bo Yu, 2009, "Hidden charged dark matter," J. Cosmol. Astropart. Phys. 07, 004.

Feng, Jonathan L., Huitzu Tu, and Hai-Bo Yu, 2008, "Thermal relics in hidden sectors," J. Cosmol. Astropart. Phys. 10, 043.

Ferrara, Sergio, and Renata Kallosh, 1996a, "Supersymmetry and attractors," Phys. Rev. D **54**, 1514–1524.

Ferrara, Sergio, and Renata Kallosh, 1996b, "Universality of supersymmetric attractors," Phys. Rev. D **54**, 1525–1534.

Ferrara, Sergio, Renata Kallosh, and Andrew Strominger, 1995, "$N = 2$ extremal black holes," Phys. Rev. D **52**, R5412–R5416.

Fichet, Sylvain, and Prashant Saraswat, 2020, "Approximate symmetries and gravity," J. High Energy Phys. 01, 088.

Fisher, Zachary, and Christopher J. Mogni, 2017, "A semiclassical, entropic proof of a weak gravity conjecture," arXiv:1706.08257.

Foot, Robert, 1991, "New physics from electric charge quantization?," Mod. Phys. Lett. A **06**, 527–530.

Freese, Katherine, Joshua A. Frieman, and Angela V. Olinto, 1990, "Natural Inflation with Pseudo–Nambu-Goldstone Bosons," Phys. Rev. Lett. **65**, 3233–3236.

Friedman, John L., Kristin Schleich, and Donald M. Witt, 1993, "Topological Censorship," Phys. Rev. Lett. **71**, 1486–1489.

Furuuchi, Kazuyuki, 2018, "Weak gravity conjecture from low energy observers' perspective," Fortschr. Phys. **66**, 1800016.

Gaiotto, Davide, Anton Kapustin, Nathan Seiberg, and Brian Willett, 2015, "Generalized global symmetries," J. High Energy Phys. 02, 172.

Gendler, Naomi, Ben Heidenreich, Liam McAllister, Jakob Moritz, and Tom Rudelius, 2022, "Moduli space reconstruction and weak gravity," arXiv:2212.10573.

Gendler, Naomi, and Irene Valenzuela, 2021, "Merging the weak gravity and distance conjectures using BPS extremal black holes," J. High Energy Phys. 01, 176.

Gibbons, G. W., 1975, "Vacuum polarization and the spontaneous loss of charge by black holes," Commun. Math. Phys. **44**, 245–264.

Gibbons, G. W., S. W. Hawking, Gary T. Horowitz, and Malcolm J. Perry, 1983, "Positive mass theorems for black holes," Commun. Math. Phys. **88**, 295.

Gibbons, Gary W., Renata Kallosh, and Barak Kol, 1996, "Moduli, Scalar Charges, and the First Law of Black Hole Thermodynamics," Phys. Rev. Lett. **77**, 4992–4995.

Giddings, Steven B., Joseph Polchinski, and Andrew Strominger, 1993, "Four-dimensional black holes in string theory," Phys. Rev. D **48**, 5784–5797.

Giddings, Steven B., and Andrew Strominger, 1988, "Loss of incoherence and determination of coupling constants in quantum gravity," Nucl. Phys. **B307**, 854–866.

Goldhaber, Alfred Scharff, and Michael Martin Nieto, 2010, "Photon and graviton mass limits," Rev. Mod. Phys. **82**, 939–979.

Gonzalo, Eduardo, and Luis E. Ibáñez, 2019, "A strong scalar weak gravity conjecture and some implications," J. High Energy Phys. 08, 118.

Goon, Garrett, and Kurt Hinterbichler, 2017, "Superluminality, black holes and EFT," J. High Energy Phys. 02, 134.

Goon, Garrett, and Riccardo Penco, 2020, "Universal Relation between Corrections to Entropy and Extremality," Phys. Rev. Lett. **124**, 101103.

Gopakumar, Rajesh, and Cumrun Vafa, 1998a, "*M*-theory and topological strings—I," arXiv:hep-th/9809187.

Gopakumar, Rajesh, and Cumrun Vafa, 1998b, "*M*-theory and topological strings—II," arXiv:hep-th/9812127.

Gopakumar, Rajesh, and Cumrun Vafa, 1999, "On the gauge theory/geometry correspondence," Adv. Theor. Math. Phys. **3**, 1415–1443.

Graham, Peter W., David E. Kaplan, and Surjeet Rajendran, 2015, "Cosmological Relaxation of the Electroweak Scale," Phys. Rev. Lett. **115**, 221801.

Graham, Peter W., Jeremy Mardon, and Surjeet Rajendran, 2016, "Vector dark matter from inflationary fluctuations," Phys. Rev. D **93**, 103520.

Graña, Mariana, and Alvaro Herráez, 2021, "The swampland conjectures: A bridge from quantum gravity to particle physics," Universe **7**, 273.

Gregory, R., and R. Laflamme, 1993, "Black Strings and *p*-Branes are Unstable," Phys. Rev. Lett. **70**, 2837–2840.

Gregory, Ruth, 1996, "Nonsingular global strings," Phys. Rev. D **54**, 4955–4962.

Grimm, Thomas W., and Damian Van De Heisteeg, 2020, "Infinite distances and the axion weak gravity conjecture," J. High Energy Phys. 03, 020.

Grimm, Thomas W., Chongchuo Li, and Eran Palti, 2019, "Infinite distance networks in field space and charge orbits," J. High Energy Phys. 03, 016.

Grimm, Thomas W., Chongchuo Li, and Irene Valenzuela, 2020, "Asymptotic flux compactifications and the swampland," J. High Energy Phys. 06, 009.

Grimm, Thomas W., Eran Palti, and Irene Valenzuela, 2018, "Infinite distances in field space and massless towers of states," J. High Energy Phys. 08, 143.

Gruzinov, A., 2007, "Kerr quasinormal modes and Hod's time-temperature bound," arXiv:0705.1725.

Gubser, Steven S., 2008, "Breaking an Abelian gauge symmetry near a black hole horizon," Phys. Rev. D **78**, 065034.

Guth, Alan H., 1981, "Inflationary universe: A possible solution to the horizon and flatness problems," Phys. Rev. D **23**, 347–356.

Gutperle, Michael, and Wafic Sabra, 2002, "Instantons and wormholes in Minkowski and (A)dS spaces," Nucl. Phys. **B647**, 344–356.

Hamada, Yuta, Toshifumi Noumi, and Gary Shiu, 2019, "Weak Gravity Conjecture from Unitarity and Causality," Phys. Rev. Lett. **123**, 051601.

Harlow, Daniel, 2016, "Wormholes, emergent gauge fields, and the weak gravity conjecture," J. High Energy Phys. 01, 122.

Harlow, Daniel, and Hirosi Ooguri, 2019, "Constraints on Symmetries from Holography," Phys. Rev. Lett. **122**, 191601.

Harlow, Daniel, and Hirosi Ooguri, 2021, "Symmetries in quantum field theory and quantum gravity," Commun. Math. Phys. **383**, 1669–1804.

Harlow, Daniel, and Edgar Shaghoulian, 2021, "Global symmetry, Euclidean gravity, and the black hole information problem," J. High Energy Phys. 04, 175.

Hartman, Thomas, Sachin Jain, and Sandipan Kundu, 2016, "Causality constraints in conformal field theory," J. High Energy Phys. 05, 099.

Hartnoll, Sean A., Christopher P. Herzog, and Gary T. Horowitz, 2008a, "Building a Holographic Superconductor," Phys. Rev. Lett. **101**, 031601.

Hartnoll, Sean A., Christopher P. Herzog, and Gary T. Horowitz, 2008b, "Holographic superconductors," J. High Energy Phys. 12, 015.

Hartnoll, Sean A., Joseph Polchinski, Eva Silverstein, and David Tong, 2010, "Towards strange metallic holography," J. High Energy Phys. 04, 120.

Hastings, M. B., 2007, "An area law for one-dimensional quantum systems," J. Stat. Mech. P08024.

Hawking, S. W., 1975, "Particle creation by black holes," Commun. Math. Phys. **43**, 199–220; **46**, 206(E) (1976).

Hawking, S. W., 1996, "Virtual black holes," Phys. Rev. D **53**, 3099–3107.

He, X. G., Girish C. Joshi, H. Lew, and R. R. Volkas, 1991, "New-Z' phenomenology," Phys. Rev. D **43**, R22(R).

Hebecker, Arthur, Philipp Henkenjohann, and Lukas T. Witkowski, 2017a, "Flat monodromies and a moduli space size conjecture," J. High Energy Phys. 12, 033.

Hebecker, Arthur, Philipp Henkenjohann, and Lukas T. Witkowski, 2017b, "What is the magnetic weak gravity conjecture for axions?," Fortschr. Phys. **65**, 1700011.

Hebecker, Arthur, Patrick Mangat, Fabrizio Rompineve, and Lukas T. Witkowski, 2015, "Winding out of the swamp: Evading the weak gravity conjecture with F-term winding inflation?," Phys. Lett. B **748**, 455–462.

Hebecker, Arthur, Patrick Mangat, Stefan Theisen, and Lukas T. Witkowski, 2017, "Can gravitational instantons really constrain axion inflation?," J. High Energy Phys. 02, 097.

Hebecker, Arthur, Thomas Mikhail, and Pablo Soler, 2018, "Euclidean wormholes, baby universes, and their impact on particle physics and cosmology," Front. Astron. Space Sci. **5**, 35.

Hebecker, Arthur, Fabrizio Rompineve, and Alexander Westphal, 2016, "Axion monodromy and the weak gravity conjecture," J. High Energy Phys. 04, 157.

Hebecker, Arthur, and Pablo Soler, 2017, "The weak gravity conjecture and the axionic black hole paradox," J. High Energy Phys. 09, 036.

Heeck, Julian, 2014, "Unbroken B-L symmetry," Phys. Lett. B **739**, 256–262.

Heidenreich, Ben, 2020, "Black holes, moduli, and long-range forces," J. High Energy Phys. 11, 029.

Heidenreich, Ben, Cody Long, Liam McAllister, Tom Rudelius, and John Stout, 2020, "Instanton resummation and the weak gravity conjecture," J. High Energy Phys. 11, 166.

Heidenreich, Ben, and Matteo Lotito, 2022, "Proving the weak gravity conjecture in perturbative string theory" (to be published).

Heidenreich, Ben, Jacob McNamara, Miguel Montero, Matthew Reece, Tom Rudelius, and Irene Valenzuela, 2021a, "Chern-Weil global symmetries and how quantum gravity avoids them," J. High Energy Phys. 11, 053.

Heidenreich, Ben, Jacob McNamara, Miguel Montero, Matthew Reece, Tom Rudelius, and Irene Valenzuela, 2021b, "Non-invertible global symmetries and completeness of the spectrum," J. High Energy Phys. 09, 203.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2018a, "Emergence of Weak Coupling at Large Distance in Quantum Gravity," Phys. Rev. Lett. **121**, 051601.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2018b, "The weak gravity conjecture and emergence from an ultraviolet cutoff," Eur. Phys. J. C **78**, 337.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2015, "Weak gravity strongly constrains large-field axion inflation," J. High Energy Phys. 12, 108.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2016, "Sharpening the weak gravity conjecture with dimensional reduction," J. High Energy Phys. 02, 140.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2017, "Evidence for a sublattice weak gravity conjecture," J. High Energy Phys. 08, 025.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2019, "Repulsive forces and the weak gravity conjecture," J. High Energy Phys. 10, 055.

Heidenreich, Ben, Matthew Reece, and Tom Rudelius, 2021, "The weak gravity conjecture and axion strings," J. High Energy Phys. 11, 004.

Heidenreich, Ben, and Tom Rudelius, 2021, "Infinite distance and zero gauge coupling in 5D supergravity," Phys. Rev. D **104**, 106016.

Heidenreich, Ben, and Tom Rudelius, 2022, "The weak gravity conjecture and BPS strings" (to be published).

Heikinheimo, Matti, Martti Raidal, Christian Spethmann, and Hardi Veermäe, 2015, "Dark matter self-interactions via collisionless shocks in cluster mergers," Phys. Lett. B **749**, 236–241.

Henriksson, Oscar, Carlos Hoyos, and Niko Jokela, 2020, "Brane nucleation instabilities in non-AdS/non-CFT," J. High Energy Phys. 02, 007.

Heydeman, Matthew, Luca V. Iliesiu, Gustavo J. Turiaci, and Wenli Zhao, 2022, "The statistical mechanics of near-BPS black holes," J. Phys. A **55**, 014004.

Hod, Shahar, 2007a, "Near-extreme black holes and the universal relaxation bound," Classical Quantum Gravity **24**, 4235–4238.

Hod, Shahar, 2007b, "Universal bound on dynamical relaxation times and black-hole quasinormal ringing," Phys. Rev. D **75**, 064013.

Hod, Shahar, 2017, "A proof of the weak gravity conjecture," Int. J. Mod. Phys. D **26**, 1742004.

Holdom, Bob, 1986, “Two U(1)’s and epsilon charge shifts,” Phys. Lett. **166B**, 196–198.
’t Hooft, Gerard, 1993, “Dimensional reduction in quantum gravity,” Conf. Proc. C **930308**, 284–296, https://inspirehep.net/literature/36137.
Horowitz, Gary T., 2005, “Tachyon condensation and black strings,” J. High Energy Phys. 08, 091.
Horowitz, Gary T., and Jorge E. Santos, 2019, “Further evidence for the weak gravity–cosmic censorship connection,” J. High Energy Phys. 06, 122.
Horowitz, Gary T., Jorge E. Santos, and Benson Way, 2016, “Evidence for an electrifying violation of cosmic censorship,” Classical Quantum Gravity **33**, 195007.
Hosono, S., A. Klemm, S. Theisen, and Shing-Tung Yau, 1995a, “Mirror symmetry, mirror map and applications to Calabi-Yau hypersurfaces,” Commun. Math. Phys. **167**, 301–350.
Hosono, S., A. Klemm, S. Theisen, and Shing-Tung Yau, 1995b, “Mirror symmetry, mirror map and applications to complete intersection Calabi-Yau spaces,” Nucl. Phys. **B433**, 501–554.
Hsin, Po-Shen, Luca V. Iliesiu, and Zhenbin Yang, 2021, “A violation of global symmetries from replica wormholes and the fate of black hole remnants,” Classical Quantum Gravity **38**, 194004.
Ibáñez, Luis E., and Miguel Montero, 2018, “A note on the WGC, effective field theory and clockwork within string theory,” J. High Energy Phys. 02, 057.
Ibáñez, Luis E., Miguel Montero, Angel Uranga, and Irene Valenzuela, 2016, “Relaxion monodromy and the weak gravity conjecture,” J. High Energy Phys. 04, 020.
Iliesiu, Luca V., and Gustavo J. Turiaci, 2021, “The statistical mechanics of near-extremal black holes,” J. High Energy Phys. 05, 145.
Johnson, George, 2020, “Tunnelling of charged particles from black holes,” J. High Energy Phys. 03, 038.
Jones, Callum R. T., and Brian McPeak, 2020, “The black hole weak gravity conjecture with multiple charges,” J. High Energy Phys. 06, 140.
Junghans, Daniel, 2016, “Large-field inflation with multiple axions and the weak gravity conjecture,” J. High Energy Phys. 02, 128.
Kallosh, Renata, and Andrei Linde, 2021, “BICEP/Keck and cosmological attractors,” J. Cosmol. Astropart. Phys. 12, 008.
Kallosh, Renata, Andrei D. Linde, Dmitri A. Linde, and Leonard Susskind, 1995, “Gravity and global symmetries,” Phys. Rev. D **52**, 912–935.
Kaloper, Nemanja, and Lorenzo Sorbo, 2009, “A Natural Framework for Chaotic Inflation,” Phys. Rev. Lett. **102**, 121301.
Kaplan, David E., and Riccardo Rattazzi, 2016, “Large field excursions and approximate discrete symmetries from a clockwork axion,” Phys. Rev. D **93**, 085007.
Kaplan, Jared, and Sandipan Kundu, 2019, “A species or weak-gravity bound for large N gauge theories coupled to gravity,” J. High Energy Phys. 11, 142.
Kaplan, Jared, and Sandipan Kundu, 2021, “Closed strings and weak gravity from higher-spin causality,” J. High Energy Phys. 02, 145.
Kats, Yevgeny, Lubos Motl, and Megha Padi, 2007, “Higher-order corrections to mass-charge relation of extremal black holes,” J. High Energy Phys. 12, 068.
Kaya, Sami, and Tom Rudelius, 2022, “Higher-group symmetries and weak gravity conjecture mixing,” J. High Energy Phys. 07, 040.
Kim, Jihn E., Hans Peter Nilles, and Marco Peloso, 2005, “Completing natural inflation,” J. Cosmol. Astropart. Phys. 01, 005.
Kim, Manki, and Liam McAllister, 2020, “Monodromy charge in D7-brane inflation,” J. High Energy Phys. 10, 060.
Klaewer, Daniel, Seung-Joo Lee, Timo Weigand, and Max Wiesner, 2021, “Quantum corrections in 4D $N = 1$ infinite distance limits and the weak gravity conjecture,” J. High Energy Phys. 03, 252.
Klaewer, Daniel, and Eran Palti, 2017, “Super-Planckian spatial field variations and quantum gravity,” J. High Energy Phys. 01, 088.
Konoplya, R. A., and A. Zhidenko, 2013, “Massive charged scalar field in the Kerr-Newman background: Quasinormal modes, late-time tails and stability,” Phys. Rev. D **88**, 024054.
Lanza, Stefano, Fernando Marchesano, Luca Martucci, and Irene Valenzuela, 2021a, “Swampland conjectures for strings and membranes,” J. High Energy Phys. 02, 006.
Lanza, Stefano, Fernando Marchesano, Luca Martucci, and Irene Valenzuela, 2021b, “The EFT stringy viewpoint on large distances,” J. High Energy Phys. 09, 197.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2018, “Tensionless strings and the weak gravity conjecture,” J. High Energy Phys. 10, 164.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2019a, “A stringy test of the scalar weak gravity conjecture,” Nucl. Phys. **B938**, 321–350.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2019b, “Modular fluxes, elliptic genera, and weak gravity conjectures in four dimensions,” J. High Energy Phys. 08, 104.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2022a, “Emergent strings, duality and weak coupling limits for two-form fields,” J. High Energy Phys. 02, 096.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2022b, “Emergent strings from infinite distance limits,” J. High Energy Phys. 02, 190.
Lee, Seung-Joo, Wolfgang Lerche, and Timo Weigand, 2022c, “Physics of infinite complex structure limits in eight dimensions,” J. High Energy Phys. 06, 042.
Lehner, Luis, and Frans Pretorius, 2012, “Final state of Gregory-Laflamme instability,” in *Black Holes in Higher Dimensions*, edited by Gary T. Horowitz (Cambridge University Press, Cambridge, England).
Lewkowycz, Aitor, and Juan Maldacena, 2013, “Generalized gravitational entropy,” J. High Energy Phys. 08, 090.
Li, Daliang, David Meltzer, and David Poland, 2017, “Conformal bootstrap in the Regge limit,” J. High Energy Phys. 12, 013.
Li, Miao, Wei Song, Yushu Song, and Tower Wang, 2007, “A weak gravity conjecture for scalar field theories,” J. High Energy Phys. 05, 026.
Liddle, Andrew R., Anupam Mazumdar, and Franz E. Schunck, 1998, “Assisted inflation,” Phys. Rev. D **58**, 061301.
Linde, Andrei D., 1982, “A new inflationary universe scenario: A possible solution of the horizon, flatness, homogeneity, isotropy and primordial monopole problems,” Phys. Lett. **108B**, 389–393.
Linde, Andrei D., 1994, “Monopoles as big as a universe,” Phys. Lett. B **327**, 208–213.
Liu, Hong, John McGreevy, and David Vegh, 2011, “Non-Fermi liquids from holography,” Phys. Rev. D **83**, 065029.
Loges, Gregory J., Toshifumi Noumi, and Gary Shiu, 2020a, “Duality and supersymmetry constraints on the weak gravity conjecture,” J. High Energy Phys. 11, 008.
Loges, Gregory J., Toshifumi Noumi, and Gary Shiu, 2020b, “Thermodynamics of 4D dilatonic black holes and the weak gravity conjecture,” Phys. Rev. D **102**, 046010.
Long, Cody, Liam McAllister, and John Stout, 2017, “Systematics of axion inflation in Calabi-Yau hypersurfaces,” J. High Energy Phys. 02, 014.

Long, Cody, Artan Sheshmani, Cumrun Vafa, and Shing-Tung Yau, 2021, "Non-holomorphic cycles and non-BPS black branes," arXiv:2104.06420.

Lucas, Andrew, 2019, "Operator Size at Finite Temperature and Planckian Bounds on Quantum Dynamics," Phys. Rev. Lett. **122**, 216601.

Lüst, Dieter, and Eran Palti, 2018, "Scalar fields, hierarchical UV/IR mixing and the weak gravity conjecture," J. High Energy Phys. 02, 040.

Lyth, David H., 1997, "What Would We Learn by Detecting a Gravitational Wave Signal in the Cosmic Microwave Background Anisotropy?," Phys. Rev. Lett. **78**, 1861–1863.

Ma, Ernest, D. P. Roy, and Sourov Roy, 2002, "Gauged $L_\mu - L_\tau$ with large muon anomalous magnetic moment and the bimaximal mixing of neutrinos," Phys. Lett. B **525**, 101–106.

Ma, Liang, Yi Pang, and H. Lü, 2022, "α'-corrections to near extremal dyonic strings and weak gravity conjecture," J. High Energy Phys. 01, 157.

Majumdar, Sudhansu Datta, 1947, "A class of exact solutions of Einstein's field equations," Phys. Rev. **72**, 390.

Maldacena, Juan, Stephen H. Shenker, and Douglas Stanford, 2016, "A bound on chaos," J. High Energy Phys. 08, 106.

Maldacena, Juan, and Douglas Stanford, 2016, "Remarks on the Sachdev-Ye-Kitaev model," Phys. Rev. D **94**, 106002.

Maldacena, Juan Martin, 1998, "The large *N* limit of superconformal field theories and supergravity," Adv. Theor. Math. Phys. **2**, 231–252.

Maldacena, Juan Martin, 2003a, "Eternal black holes in anti–de Sitter," J. High Energy Phys. 04, 021.

Maldacena, Juan Martin, 2003b, "Non-Gaussian features of primordial fluctuations in single field inflationary models," J. High Energy Phys. 05, 013.

Maldacena, Juan Martin, Jeremy Michelson, and Andrew Strominger, 1999, "Anti–de Sitter fragmentation," J. High Energy Phys. 02, 011.

Marchesano, Fernando, and Max Wiesner, 2019, "Instantons and infinite distances," J. High Energy Phys. 08, 088.

McAllister, Liam, Pedro Schwaller, Geraldine Servant, John Stout, and Alexander Westphal, 2018, "Runaway relaxion monodromy," J. High Energy Phys. 02, 124.

McAllister, Liam, Eva Silverstein, and Alexander Westphal, 2010, "Gravity waves and linear inflation from axion monodromy," Phys. Rev. D **82**, 046003.

McAllister, Liam, Eva Silverstein, Alexander Westphal, and Timm Wrase, 2014, "The powers of monodromy," J. High Energy Phys. 09, 123.

McPeak, Brian, 2022, "Higher-derivative corrections to black hole entropy at zero temperature," Phys. Rev. D **105**, L081901.

Montero, Miguel, 2019, "A holographic derivation of the weak gravity conjecture," J. High Energy Phys. 03, 157.

Montero, Miguel, Gary Shiu, and Pablo Soler, 2016, "The weak gravity conjecture in three dimensions," J. High Energy Phys. 10, 159.

Montero, Miguel, Angel M. Uranga, and Irene Valenzuela, 2015, "Transplanckian axions!?," J. High Energy Phys. 08, 032.

Montero, Miguel, Angel M. Uranga, and Irene Valenzuela, 2017, "A Chern-Simons pandemic," J. High Energy Phys. 07, 123.

Morrison, David R., 1993, "Compactifications of moduli spaces inspired by mirror symmetry," J. Geom. Algébrique Orsay **218**, 243–271, http://www.numdam.org/item/AST_1993__218__243_0/.

Morrison, David R., 1994, "Beyond the Kähler cone," in *Proceedings of the Hirzebruch 65 Conference on Algebraic Geometry*, edited by Mina Teicher (Bar-Ilan University, Ramat Gan, Israel), pp. 361–376.

Nakayama, Yu, and Yasunori Nomura, 2015, "Weak gravity conjecture in the AdS/CFT correspondence," Phys. Rev. D **92**, 126006.

Natsuume, Makoto, 1994, "Higher order correction to the GHS string black hole," Phys. Rev. D **50**, 3949–3953.

Nicolis, Alberto, 2008, "On super-Planckian fields at sub-Planckian energies," J. High Energy Phys. 07, 023.

Nielsen, Holger Bech, and P. Olesen, 1973, "Vortex line models for dual strings," Nucl. Phys. **B61**, 45–61.

Nomura, Yasunori, 2020, "Spacetime and universal soft modes—Black holes and beyond," Phys. Rev. D **101**, 066024.

Obied, Georges, and Aditya Parikh, 2021, "A tale of two U(1)'s: Kinetic mixing from lattice WGC states," arXiv:2109.07913.

Ooguri, Hirosi, and Cumrun Vafa, 2007, "On the geometry of the string landscape and the swampland," Nucl. Phys. **B766**, 21–33.

Ooguri, Hirosi, and Cumrun Vafa, 2017, "Non-supersymmetric AdS and the swampland," Adv. Theor. Math. Phys. **21**, 1787–1801.

Page, Don N., 2000, "Thermodynamics of near extreme black holes," arXiv:hep-th/0012020.

Palti, Eran, 2015, "On natural inflation and moduli stabilisation in string theory," J. High Energy Phys. 10, 188.

Palti, Eran, 2017, "The weak gravity conjecture and scalar fields," J. High Energy Phys. 08, 034.

Palti, Eran, 2019, "The swampland: Introduction and review," Fortschr. Phys. **67**, 1900037.

Papaetrou, Achille, 1947, "A Static solution of the equations of the gravitational field for an arbitrary charge distribution," Proc. R. Ir. Acad., Sect. A **51**, 191–204, https://www.jstor.org/stable/20488481.

Peiris, H. V., *et al.* (WMAP Collaboration), 2003, "First year Wilkinson Microwave Anisotropy Probe (WMAP) observations: Implications for inflation," Astrophys. J. Suppl. Ser. **148**, 213–231.

Penington, Geoff, Stephen H. Shenker, Douglas Stanford, and Zhenbin Yang, 2022, "Replica wormholes and the black hole interior," J. High Energy Phys. 03, 205.

Penington, Geoffrey, 2020, "Entanglement wedge reconstruction and the information paradox," J. High Energy Phys. 09, 002.

Penrose, R., 1969, "Gravitational collapse: The role of general relativity," Riv. Nuovo Cimento **1**, 252–276, https://inspirehep.net/literature/54979.

Perlmutter, Eric, Leonardo Rastelli, Cumrun Vafa, and Irene Valenzuela, 2021, "A CFT distance conjecture," J. High Energy Phys. 10, 070.

Pham, T. N., and Tran N. Truong, 1985, "Evaluation of the derivative quartic terms of the meson chiral Lagrangian from forward dispersion relation," Phys. Rev. D **31**, 3027(R).

Polchinski, J., 2007, *String Theory, Vol. 2: Superstring Theory and Beyond*, Cambridge Monographs on Mathematical Physics (Cambridge University Press, Cambridge, England).

Polchinski, Joseph, 2004, "Monopoles, duality, and string theory," Int. J. Mod. Phys. A **19**, 145–156.

Preskill, John, Patricia Schwarz, Alfred D. Shapere, Sandip Trivedi, and Frank Wilczek, 1991, "Limitations on the statistical description of black holes," Mod. Phys. Lett. A **06**, 2353–2362.

Proca, Alexandru, 1936, "Sur la theorie ondulatoire des electrons positifs et negatifs [On the wave theory of positive and negative electrons]," J. Phys. Radium **7**, 347–353.

Reall, Harvey S., and Jorge E. Santos, 2019, "Higher derivative corrections to Kerr black hole thermodynamics," J. High Energy Phys. 04, 021.

Reece, Matthew, 2019, "Photon masses in the landscape and the swampland," J. High Energy Phys. 07, 181.

Richartz, Maurício, and Davi Giugno, 2014, "Quasinormal modes of charged fields around a Reissner-Nordström black hole," Phys. Rev. D **90**, 124011.

Rudelius, Tom, 2015a, "Constraints on axion inflation from the weak gravity conjecture," J. Cosmol. Astropart. Phys. 09, 020.

Rudelius, Tom, 2015b, "On the possibility of large axion moduli spaces," J. Cosmol. Astropart. Phys. 04, 049.

Rudelius, Tom, 2021, "Dimensional reduction and (anti) de Sitter bounds," J. High Energy Phys. 08, 041.

Rudelius, Tom, and Shu-Heng Shao, 2020, "Topological operators and completeness of spectrum in discrete gauge theories," J. High Energy Phys. 12, 172.

Saraswat, Prashant, 2017, "Weak gravity conjecture and effective field theory," Phys. Rev. D **95**, 025013.

Schwimmer, A., and N. Seiberg, 1987, "Comments on the $N = 2, 3, 4$ superconformal algebras in two dimensions," Phys. Lett. B **184**, 191–196.

Sharpe, Eric, 2015, "Notes on generalized global symmetries in QFT," Fortschr. Phys. **63**, 659–682.

Shiu, Gary, William Cottrell, and Pablo Soler, 2017, "Weak gravity conjecture and black holes in $N = 2$ supergravity," Proc. Sci. CORFU2016, 130.

Shiu, Gary, Pablo Soler, and William Cottrell, 2019, "Weak gravity conjecture and extremal black holes," Sci. China Phys. Mech. Astron. **62**, 110412.

Silverstein, Eva, and Alexander Westphal, 2008, "Monodromy in the CMB: Gravity waves and string inflation," Phys. Rev. D **78**, 106003.

Starobinsky, Alexei A., 1980, "A new type of isotropic cosmological models without singularity," Phys. Lett. **91B**, 99–102.

Stout, John, 2022, "Instanton expansions and phase transitions," J. High Energy Phys. 05, 168.

Strassler, Matthew J., 2014, "On strings from things and things from strings," lecture, KITP Conference: From the Renormalization Group to Quantum Gravity, Celebrating the Science of Joe Polchinski, Santa Barbara, 2014, https://online.kitp.ucsb.edu/online/joefest_c14/.

Strassler, Matthew J., 2016, "Digging deep at the LHC," lecture, ICTP Conference: A First Glance beyond the Energy Frontier, Trieste, Italy, 2016, https://indico.ictp.it/event/7627/.

Strominger, Andrew, 1996, "Macroscopic entropy of $N = 2$ extremal black holes," Phys. Lett. B **383**, 39–43.

Stueckelberg, E. C. G., 1938, "Interaction energy in electrodynamics and in the field theory of nuclear forces," Helv. Phys. Acta **11**, 225–244, https://www.e-periodica.ch/digbib/view?pid=hpa-001:1938:11::636#227.

Susskind, Leonard, 1995, "The world as a hologram," J. Math. Phys. (N.Y.) **36**, 6377–6396.

Svrcek, Peter, and Edward Witten, 2006, "Axions in string theory," J. High Energy Phys. 06, 051.

Tokuda, Junsei, Katsuki Aoki, and Shin'ichi Hirano, 2020, "Gravitational positivity bounds," J. High Energy Phys. 11, 054.

Trigiante, Mario, Thomas Van Riet, and Bert Vercnocke, 2012, "Fake supersymmetry versus Hamilton-Jacobi," J. High Energy Phys. 05, 078.

Uehling, E. A., 1935, "Polarization effects in the positron theory," Phys. Rev. **48**, 55–63.

Urbano, Alfredo, 2018, "Towards a proof of the weak gravity conjecture," arXiv:1810.05621.

Vafa, Cumrun, 2005, "The String landscape and the swampland," arXiv:hep-th/0509212.

van Beest, Marieke, José Calderón-Infante, Delaram Mirfendereski, and Irene Valenzuela, 2021, "Lectures on the swampland program in string compactifications," arXiv:2102.01111.

Van Riet, Thomas, 2021, "A comment on no-force conditions for black holes and branes," Classical Quantum Gravity **38**, 077001.

Veneziano, G., 2002, "Large N bounds on, and compositeness limit of, gauge and gravitational interactions," J. High Energy Phys. 06, 051.

Vilenkin, Alexander, 1994, "Topological Inflation," Phys. Rev. Lett. **72**, 3137–3140.

Wagner, T. A., S. Schlamminger, J. H. Gundlach, and E. G. Adelberger, 2012, "Torsion-balance tests of the weak equivalence principle," Classical Quantum Gravity **29**, 184002.

Wald, Robert M., 1993, "Black hole entropy is the Noether charge," Phys. Rev. D **48**, R3427–R3431.

Weinberg, Steven, 2008, *Cosmology* (Oxford University Press, New York).

Witten, Edward, 1979, "Dyons of charge $e\theta/2\pi$," Phys. Lett. **86B**, 283–287.

Witten, Edward, 1998, "Anti–de Sitter space and holography," Adv. Theor. Math. Phys. **2**, 253–291.

Wu, Xue-Feng, Song-Bo Zhang, He Gao, Jun-Jie Wei, Yuan-Chuan Zou, Wei-Hua Lei, Bing Zhang, Zi-Gao Dai, and Peter Mészáros, 2016, "Constraints on the photon mass with fast radio bursts," Astrophys. J. Lett. **822**, L15.

Yang, Huan, Fan Zhang, Aaron Zimmerman, David A. Nichols, Emanuele Berti, and Yanbei Chen, 2013, "Branching of quasinormal modes for nearly extremal Kerr black holes," Phys. Rev. D **87**, 041502.

Zeldovich, Ya. B., 1976, "A new type of radioactive decay: Gravitational annihilation of baryons," Phys. Lett. **59A**, 254.

Zeldovich, Ya. B., 1977, "A novel type of radioactive decay: Gravitational baryon annihilation," Zh. Eksp. Teor. Fiz. **72**, 18–21.

Zhang, Cen, and Shuang-Yong Zhou, 2020, "Convex Geometry Perspective on the (Standard Model) Effective Field Theory Space," Phys. Rev. Lett. **125**, 201601.

Zimmerman, Aaron, and Zachary Mark, 2016, "Damped and zero-damped quasinormal modes of charged, nearly extremal black holes," Phys. Rev. D **93**, 044033; **93**, 089905(E) (2016).

Zyla, P. A., *et al.* (Particle Data Group), 2020, "Review of particle physics," Prog. Theor. Exp. Phys. 083C01.