Naturalness and the Weak Gravity Conjecture

Clifford Cheung[*] and Grant N. Remmen[†]
Walter Burke Institute for Theoretical Physics, California Institute of Technology, Pasadena, California 91125, USA
(Received 19 February 2014; revised manuscript received 23 April 2014; published 29 July 2014)

The weak gravity conjecture (WGC) is an ultraviolet consistency condition asserting that an Abelian force requires a state of charge q and mass m with $q > m/m_{Pl}$. We generalize the WGC to product gauge groups and study its tension with the naturalness principle for a charged scalar coupled to gravity. Reconciling naturalness with the WGC either requires a Higgs phase or a low cutoff at $\Lambda \sim q m_{Pl}$. If neither applies, one can construct simple models that forbid a natural electroweak scale and whose observation would rule out the naturalness principle.

Introduction.—The naturalness principle asserts that operators not protected by symmetry are unstable to quantum corrections induced at the cutoff. As a tenet of effective field theory, naturalness has provided a key motivation for new physics at the electroweak scale. However, the discovery of the Higgs boson [1,2] together with null results from direct searches has led many to revisit naturalness as a fundamental principle. Rather than amend naturalness to fit the data, we instead explore its interplay with established concepts in quantum field theory.

Our focus will be the weak gravity conjecture (WGC) [3], which states that a consistent theory of gravity coupled to an Abelian gauge theory must contain a state of charge q and mass m satisfying

$$q > m/m_{Pl}; \qquad (1)$$

i.e., gravity is the weakest force. [We define the Planck mass, m_{Pl}, such that Eq. (1) is saturated for an extremal black hole.] While Eq. (1) is certainly true of electromagnetism, Ref. [3] convincingly argued that it is a universal consistency condition of all healthy quantum field theories.

However, in theories with fundamental scalars, Eq. (1) runs afoul of naturalness because it bounds a quadratically divergent mass by a logarithmically divergent charge. For small charge, Eq. (1) forbids a natural spectrum in which scalars have masses near the cutoff. We illustrate this contradiction with scalar quantum electrodynamics (QED) coupled to general relativity, but this tension is a ubiquitous feature of any model with a hierarchy problem and a small charge. We also generalize Eq. (1) to the case of multiple forces and particles.

As we will show, reconciling naturalness with Eq. (1) requires a revision of the original theory: either the gauge symmetry is spontaneously broken or new degrees of freedom enter prematurely at the cutoff

$$\Lambda \sim q m_{Pl}. \qquad (2)$$

Reference [3] conjectured Eq. (2) with the stronger interpretation that Λ signals the complete breakdown of four-dimensional quantum field theory. Supporting this claim with compelling string theoretic examples, Ref. [3] fell short of a general argument. However, if one asserts the primacy of naturalness, then our logic provides a reason from quantum field theory for new states at Λ.

To illustrate these ideas we present simple, concrete extensions of the standard model (SM) in which a natural value of the electroweak scale—at the Planck scale—is incompatible with Eq. (1) due to a new millicharged force. These models offer the unique opportunity to test naturalness experimentally. Indeed, either naturalness reigns, in which case Eq. (2) demands a low cutoff, or it fails. Absent additional ultralight states, a discovery of this millicharged force would then invalidate naturalness and mandate an unnatural electroweak scale. In particular, Eq. (1) would disallow a natural electroweak scale and the hierarchy problem would arise from as-yet-unknown ultraviolet dynamics. More generally, a fifth force discovery of any kind would invalidate the interpretation of Λ advocated in Ref. [3] as the cutoff of four-dimensional quantum field theory. If, as conjectured in Ref. [3], this breakdown is a universal feature of all string compactifications, such an observation would also falsify string theory.

Evidence for the WGC.—Let us summarize the justification for the WGC [3]. Consider a $U(1)$ gauge theory with charged species labeled by i, each representing a particle (anti-particle) of charge q_i ($-q_i$) and mass m_i. We define dimensionless charge-to-mass ratios,

$$z_i = q_i m_{Pl}/m_i, \qquad (3)$$

so Eq. (1) implies that there exists some particle i with $z_i > 1$. The authors of Ref. [3] offered theoretical evidence in support of Eq. (1). They presented many examples from field theory and string theory, all satisfying Eq. (1). Further, they argued that Eq. (1) reconciles the inherent inconsistency of exact global symmetries with the naïvely

innocuous $q \to 0$ limit of a gauge theory. This limit yields an exact global symmetry; however, such charges are not conserved by quantum gravity [4,5] because, in accordance with no-hair theorems [6], a stationary black hole is fully characterized by its mass, spin, and charge.

Of course, examples and consistency with no-hair theorems only provide circumstantial evidence for Eq. (1). Importantly, Ref. [3] also argues for Eq. (1) via *reductio ad absurdum*, drawing only on general relativity, conservation of charge and energy, and minimal assumptions about the ultimate theory of quantum gravity. Consider a black hole of charge Q and mass M decaying solely to particles of species i, which can occur via Hawking radiation or Schwinger pair production [7,8]. By charge conservation, Q/q_i particles are produced. Conservation of energy dictates that the total rest mass of the final state, $m_i Q/q_i$, be less than M. In terms of the black hole charge-to-mass ratio $Z = Q m_{\mathrm{Pl}}/M$ this implies $z_i > Z$. An extremal black hole corresponds to $Z = 1$ and is stable unless some state i exists for which $z_i > 1$. If Eq. (1) fails, the spectrum contains a large number of stable black hole remnants, in tension with holographic bounds [9,10] and afflicted with various quantum gravitational and thermodynamic pathologies [11,12].

The Limits of Naturalness.—The WGC is straightforward at tree level, but radiative corrections introduce subtleties. In fermionic QED, q and m run with renormalization scale, as does their ratio, naïvely making q/m ambiguous; however, as Ref. [3] notes, the appropriate scale to evaluate q/m is the physical mass of the particle. This is the mass scale that is relevant to the kinematics of extremal black hole decay, which provides the justification for the WGC.

However, the radiative stability question becomes more interesting in scalar QED:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}^2 + |D_\mu \phi|^2 - m^2|\phi|^2 - \frac{\lambda}{4}|\phi|^4, \qquad (4)$$

where $D_\mu = \partial_\mu + iqA_\mu$ is the gauge covariant derivative. As for any effective field theory, we assume an ultraviolet cutoff Λ, above which new physics enters. Since ϕ is a fundamental scalar, its mass is radiatively unstable and corrected by $m^2 \to m^2 + \delta m^2$, where

$$\delta m^2 = \frac{\Lambda^2}{16\pi^2}(aq^2 + b\lambda). \qquad (5)$$

Here, a and b are dimensionless numerical coefficients. We assume that δm^2 is positive so that the theory remains in the Coulomb phase. In a natural theory, the physical mass of ϕ cannot be parametrically smaller than its radiative corrections. Equivalently, the counterterm for the scalar mass should not introduce a delicate cancellation. This is formally equivalent to requiring that the coefficients a and b take on $\mathcal{O}(1)$ values.

Let us set the physical mass squared for ϕ to its natural value, δm^2, which the WGC forbids from exceeding its charge in Planck units. The charge-to-mass ratio of ϕ is

$$z = \frac{4\pi m_{\mathrm{Pl}}}{\Lambda}\frac{1}{\sqrt{a + b\lambda/q^2}}, \qquad (6)$$

where the WGC implies that $z > 1$. If $q^2 \gg \lambda$, then

$$\Lambda < \frac{4\pi m_{\mathrm{Pl}}}{\sqrt{a}}, \qquad (7)$$

which is the reasonable requirement that the cutoff not exceed the Planck scale.

Turning to the opposite hierarchy, $q^2 \ll \lambda$, which is also radiatively stable, we find that the WGC implies

$$\Lambda < 4\pi m_{\mathrm{Pl}}\sqrt{\frac{q^2}{b\lambda}}. \qquad (8)$$

As $q^2/\lambda \to 0$, a sensible cutoff requires $b \to 0$, indicating mandatory fine-tuning in order to satisfy the WGC. We are left with a remarkable conclusion: scalar QED with $q^2 \ll \lambda$ and natural masses fails Eq. (1) and is thus inconsistent with a quantum theory of gravity.

We have not traded a mass scale hierarchy problem for an equivalent hierarchy problem of couplings. Small charges are radiatively stable and thus technically natural. In principle, $q^2 \ll \lambda$ is no worse than the small electron Yukawa coupling.

To reconcile naturalness with Eq. (1), one alternative is to argue that the original theory—scalar QED with $q^2 \ll \lambda$—is impossible. For example, this would be true if nature does not permit fundamental scalars or if a hierarchy among couplings is somehow strictly forbidden. However, there are far less drastic options, elaborated below, if one modifies the original scalar QED theory.

(i) *Radiative corrections induce the Higgs phase.*—It is possible that quantum effects generate a tachyon for ϕ, Higgsing the theory. Charge becomes ill-defined; the charge and mass eigenbases need not commute, leaving q/m ambiguous. Further, the WGC is not justified in the Higgs phase. The original argument for the WGC [3] relied on stable extremal black holes. However, no-hair theorems imply that there are no stationary black hole solutions supporting classical hair from a massive photon [13], independent of the size of the black hole relative to the Higgs scale. If a black hole accretes a massive-$U(1)$-charged particle, it briefly supports an associated electric field, but after a time of order the photon Compton wavelength, it balds [14] when the gauge field is radiated away to infinity or through the horizon.

(ii) *New physics enters below the Planck scale.*—The simplest way to reconcile the WGC with naturalness is for the effective field theory to break down at a cutoff defined

by Eq. (8). There could be new light states regulating quadratic divergences of ϕ, effectively lowering Λ. This option resolves the contradiction tautologically by eliminating the hierarchy problem altogether. However, a more interesting alternative occurs when the new states do not couple to ϕ. The quadratic divergence of ϕ is robust and m is large. If one of these new states satisfies Eq. (1), then ϕ is irrelevant: the WGC and naturalness are reconciled. Thus, asserting naturalness offers Eq. (8) as a more precise version of the low cutoff conjecture of Ref. [3] stated in Eq. (2).

More Forces, More Particles.—Extending our results to various charged species of different spins, the WGC implies that at least one state in the spectrum must satisfy Eq. (1) after taking into account radiative corrections. Naturalness is violated in parameter regions with a hierarchy between charges and couplings that generate quadratic divergences (quartic couplings, Yukawa couplings).

The story becomes more interesting for product gauge symmetries. Consider a gauge group $\prod_{a=1}^{N} U(1)_a$ and particles i with charges q_{ia} and masses m_i. We represent the charges, $\vec{q}_i = q_{ia}$, and charge-to-mass ratios, $\vec{z}_i = q_{ia} m_{\text{Pl}}/m_i$, as vectors of $SO(N)$, the symmetry transforming the N photons among each other. If present, photon kinetic mixing can be removed by a general linear transformation on the photons, which is equivalent to redefining charge vectors of states in the theory.

To generalize the WGC for multicharged particles, Eq. (1) is inadequate and requires upgrading to a constraint on $\vec{q}_i$ and m_i. Reference [3] briefly alluded to this scenario, but detailed analysis will reveal quantitative differences between the WGC as applied to a single $U(1)$ versus many. By symmetry, the proper generalized WGC must be $SO(N)$ invariant. Naïvely, the WGC could require at least one species i with $|\vec{z}_i| > 1$. However, this is insufficient—it guarantees the existence of one particle of large total charge, but preserves stability for orthogonally charged extremal black holes. A stricter alternative is that for each $U(1)$ there exists a species i charged under that $U(1)$ with $|\vec{z}_i| > 1$. Curiously, this is still actually weaker than the true generalized WGC.

To determine the proper generalized WGC, we revisit black hole decay kinematics. Consider a black hole of charge $\vec{Q}$, mass M, and charge-to-mass ratio $\vec{Z} = \vec{Q} m_{\text{Pl}}/M$ decaying to a final state comprised of n_i particles of species i. Charge and energy conservation imply $\vec{Q} = \sum_i n_i \vec{q}_i$ and $M > \sum_i n_i m_i$. If $\sigma_i = n_i m_i/M$ is the species i fraction of the total final state mass, then $\vec{Z} = \sum_i \sigma_i \vec{z}_i$ and $1 > \sum_i \sigma_i$; decay requires that $\vec{Z}$ be a subunitary weighted average of $\vec{z}_i$. This criterion has a geometric interpretation in charge space. Draw the vectors $\pm \vec{z}_i$ corresponding to the charge-to-mass ratio of each fundamental particle in the spectrum. A weighted average of $\vec{z}_i$ defines the convex hull spanned by the vectors, delineating the space of $\vec{Z}$ that is unstable to

decay. Any state outside the convex hull is stable. Since extremal black holes correspond to $|\vec{Z}| = 1$, the generalized WGC requires that the convex hull spanned by $\pm \vec{z}_i$ contain the unit ball.

Consider a model of two Abelian factors and two charged states. The left and right panels of Fig. 1 represent two possible choices for the charge-to-mass ratios of the particles. Black holes of all possible charges are represented by the unit disc. The left panel of Fig. 1 depicts a theory that is consistent with the WGC: the unit disc is contained in the convex hull. Extremal black holes, the boundary of this disc, can decay. However, the right panel of Fig. 1 depicts a theory that violates the WGC: there are regions of the unit disc not within the convex hull, corresponding to stable black hole remnants. Remarkably, this theory fails the WGC despite the fact that $|\vec{z}_1| > 1$ and $|\vec{z}_2| > 1$. Simple geometry shows that the WGC imposes the more stringent constraint:

$$(\vec{z}_1^2 - 1)(\vec{z}_2^2 - 1) > (1 + |\vec{z}_1 \cdot \vec{z}_2|)^2. \tag{9}$$

For example, given orthogonal charges of equal magnitude, $|\vec{z}_1| = |\vec{z}_2| = z$ and $\vec{z}_1 \perp \vec{z}_2$, Eq. (9) implies $z > \sqrt{2}$, manifestly stronger than the $z > 1$ condition required for theories with a single $U(1)$. Note that the WGC places constraints on $\vec{z}_1$ and $\vec{z}_2$ that are not mathematically independent. Were a particular value of $\vec{z}_1$ experimentally observed, this would fix a bound $\vec{z}_2^2 > (1 - 1/\vec{z}_1^2)^{-1}$.

A similar analysis can be applied for N Abelian factors and N charged states. Suppose each particle is charged under a single $U(1)$, with equal magnitude charge-to-mass ratios, so $z_{ia} = \delta_{ia} z$ for some z. The convex hull defined by $\pm \vec{z}_i$ is an N-dimensional cross-polytope of circumradius z. The largest ball contained in the cross-polytope has radius $z/\sqrt{N}$. Requiring that the radius of this ball be greater than unity then implies $z > \sqrt{N}$, parametrically stronger than the condition required for a single Abelian factor.
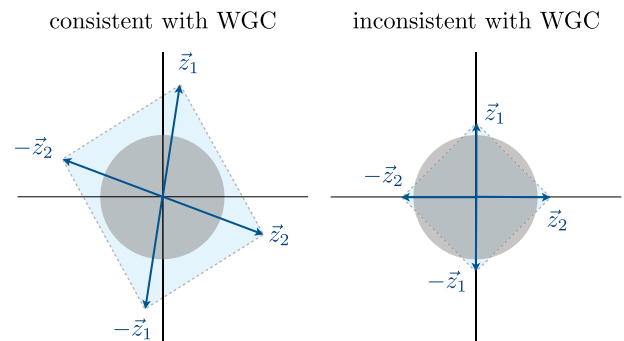

FIG. 1 (color online). Vectors representing charge-to-mass ratios for two species charged under two Abelian gauge symmetries. When the convex hull defined by these vectors contains the unit ball, then extremal black holes can decay to particles and the condition of the WGC is satisfied.

The WGC constraint grows at large N for fixed physical Planck scale m_{Pl}. However, the presence of N additional species generally renormalizes the strength of gravity [15–17] as $\delta m_{Pl}^2 \sim N\Lambda^2/16\pi^2$. If corrections enhance m_{Pl} by a factor of $\sqrt{N}$, all factors of N encountered in our earlier analyses cancel. That is, in a theory with fixed Lagrangian parameters and cutoff, the limit from the WGC is N independent at large N. A similar phenomenon was discussed in Ref. [3] for N Abelian factors Higgsed to a $U(1)$ subgroup. The large-N limit introduces a Z_2^N symmetry, which is subject to the large-N species bounds considered in Ref. [15].

The multicharge generalized WGC has implications for naturalness. Consider a $U(1)^N$ gauge theory with scalars ϕ_i of charges $\vec{q}_i$ and masses m_i,

$$\mathcal{L} = -\frac{1}{4}\sum_a F_{\mu\nu a}^2 + \sum_i |D_\mu\phi_i|^2 - m_i^2|\phi_i|^2 - \frac{\lambda_i}{4}|\phi_i|^4, \quad (10)$$

where $D_\mu\phi_i = (\partial_\mu + i\sum_a q_{ia}A_{\mu a})\phi_i$. Radiative corrections send $m_i^2 \to m_i^2 + \delta m_i^2$, where

$$\delta m_i^2 = \frac{\Lambda^2}{16\pi^2}(a_i\vec{q}_i^2 + b_i\lambda_i) \quad (11)$$

and a_i and b_i are $\mathcal{O}(1)$ ultraviolet-sensitive coefficients. The charge-to-mass ratio vector for ϕ_i is

$$\vec{z}_i = \frac{4\pi m_{Pl}}{\Lambda}\frac{\vec{q}_i}{|\vec{q}_i|}\frac{1}{\sqrt{a_i + b_i\lambda_i/\vec{q}_i^2}}. \quad (12)$$

A necessary albeit insufficient condition for the WGC is that, for each $U(1)$, there is a state i charged under that Abelian factor such that $|\vec{z}_i| > 1$. This implies

$$\Lambda < 4\pi m_{Pl} \times \begin{cases} \frac{1}{\sqrt{a_i}}, & \vec{q}_i^2 \gg \lambda_i \\ \sqrt{\frac{\vec{q}_i^2}{b_i\lambda_i}}, & \vec{q}_i^2 \ll \lambda_i \end{cases}. \quad (13)$$

As for the single Abelian case, $\vec{q}_i^2 \gg \lambda_i$ corresponds to the reasonable requirement of a sub-Planckian cutoff, while $\vec{q}_i^2 \ll \lambda_i$ implies tension with naturalness. However, the most stringent requirement of the WGC—that the convex hull spanned by $\pm\vec{z}_i$ contain the unit ball—places a stronger limit than Eq. (13) by a factor of order $\sqrt{N}$ for fixed m_{Pl}.

The Hierarchy Problem.—We have presented explicit models in which naturalness contradicts Eq. (1). We now construct theories in which natural values of the electroweak scale—at the cutoff—are similarly incompatible. In these models, strict adherence to naturalness implies either a Higgs phase or a parametrically low cutoff given by Eq. (2).

The obvious path is to relate the electroweak scale to the mass m of a particle that carries a tiny charge q. The SM gauge couplings are $\mathcal{O}(1)$, so we require an additional $U(1)$ gauge symmetry beyond the SM. It is tempting to charge the Higgs boson, but this will spontaneously break the $U(1)$, invalidating the applicability of the WGC.

However, we can charge the SM fermions under a very weakly gauged unbroken $U(1)_{B-L}$ symmetry. Current limits on $U(1)_{B-L}$ require $q \lesssim 10^{-24}$ [18,19] and will likely be improved by several orders of magnitude by astrophysical [20], lunar ranging [21], and satellite-based [22–24] tests of apparent equivalence principle violation. To cancel anomalies we introduce a right-handed neutrino ν_R that combines with the left-handed neutrino ν_L to form a $U(1)_{B-L}$ preserving Dirac mass term of the form $m_\nu\bar{\nu}_L\nu_R + \text{H.c.}$, where $m_\nu \sim y_\nu v$ is controlled by the electroweak symmetry breaking scale. The particle with the largest charge-to-mass ratio is the lightest neutrino. Assuming its mass is of order the neutrino mass scale, $m_\nu \lesssim 0.1$ eV [25,26], we fix the charge to a technically natural albeit tiny value: $q \sim m_\nu/m_{Pl} \sim 10^{-29}$. For this value of q, Eq. (1) is just marginally satisfied by the lightest neutrino. While such a charge is permitted in quantum field theory, it may be difficult to engineer in string theory if q arises from a string coupling constant requiring dilaton stabilization at large field values. Similar issues arise in theories of large extra dimensions and it is a detailed question of string moduli stabilization whether this is possible. In any case, at fixed Yukawa coupling y_ν, were the electroweak scale any higher than its measured value, Eq. (1) would fail. In this model, regions of parameter space favored by naturalness — and an electroweak scale at the cutoff—are inconsistent with Eq. (1). Strictly speaking, this logic hinges on the absence of additional $U(1)_{B-L}$ charged states lighter than the neutrino. Depending on the cosmological history, however, such particles may be constrained experimentally by primordial nucleosynthesis.

Our model offers a direct experimental test of naturalness by virtue of a very specific prediction: a new gauge boson very weakly coupled to the SM. As discussed earlier, the assumption of naturalness mandates either a Higgs phase or a low cutoff. The discovery of a fifth force would rule out the former, while current sensitivities would for the latter imply $\Lambda \lesssim$ keV from Eq. (2). In the absence of such ultralight states, the observation of a $U(1)_{B-L}$ gauge boson at $q \sim 10^{-29}$ would then simultaneously falsify the naturalness principle and suggest an ultraviolet-dependent reason for the why the weak scale takes an unnatural value. Moreover, given present sensitivities, a fifth force discovery of any kind would falsify string theory to the extent to which it predicts the strong interpretation of Λ as the scale at which four-dimensional quantum field theory breaks down [3].

This mechanism can generally be incorporated into any theory where the electroweak scale sources the mass of a $U(1)$ millicharged state, e.g., dark matter [27] charged under an unbroken $U(1)$ dark force. For weak scale dark matter, a charge of $q \sim 10^{-16}$ is sufficient to saturate Eq. (1).

---

[*]clifford.cheung@caltech.edu

[†]gremmen@theory.caltech.edu

[1] G. Aad *et al.* (ATLAS Collaboration), Phys. Lett. B **716**, 1 (2012).

[2] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Lett. B **716**, 30 (2012).

[3] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, J. High Energy Phys. 06 (2007) 060.

[4] J. D. Bekenstein, Phys. Rev. D **5**, 1239 (1972).

[5] J. D. Bekenstein, Phys. Rev. Lett. **28**, 452 (1972).

[6] S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of Space-Time* (Cambridge University Press, Cambridge, England, 1973).

[7] S. R. Coleman, J. Preskill, and F. Wilczek, Nucl. Phys. **B378**, 175 (1992).

[8] G. W. Gibbons, Commun. Math. Phys. **44**, 245 (1975).

[9] G. 't Hooft, arXiv:gr-qc/9310026.

[10] R. Bousso, Rev. Mod. Phys. **74**, 825 (2002).

[11] L. Susskind, arXiv:hep-th/9501106.

[12] S. B. Giddings, Phys. Rev. D **46**, 1347 (1992).

[13] S. L. Adler and R. B. Pearson, Phys. Rev. D **18**, 2798 (1978).

[14] J. Preskill, Phys. Scr. **T36**, 258 (1991).

[15] G. Dvali, Fortschr. Phys. **58**, 528 (2010).

[16] N. Arkani-Hamed, S. Dimopoulos, and S. Kachru, arXiv:hep-th/0501082.

[17] S. Dimopoulos, S. Kachru, J. McGreevy, and J. G. Wacker, J. Cosmol. Astropart. Phys. 08 (2008) 003.

[18] T. A. Wagner, S. Schlamminger, J. H. Gundlach, and E. G. Adelberger, Classical Quantum Gravity **29**, 184002 (2012).

[19] S. Schlamminger, K.-Y. Choi, T. A. Wagner, J. H. Gundlach, and E. G. Adelberger, Phys. Rev. Lett. **100**, 041101 (2008).

[20] S. M. Ransom *et al.*, Nature (London) **505**, 520 (2014).

[21] T. W. Murphy, Jr., E. G. Adelberger, J. B. R. Battat, C. D. Hoyle, N. H. Johnson, R. J. McMillan, C. W. Stubbs, and H. E. Swanson, Classical Quantum Gravity **29**, 184005 (2012).

[22] J. Overduin, F. Everitt, P. Worden, and J. Mester, Classical Quantum Gravity **29**, 184012 (2012).

[23] A. M. Nobili *et al.*, Classical Quantum Gravity **29**, 184011 (2012).

[24] P. Touboul, G. Métris, V. Lebat, and A. Robert, Classical Quantum Gravity **29**, 184010 (2012).

[25] E. Giusarma, R. de Putter, S. Ho, and O. Mena, Phys. Rev. D **88**, 063515 (2013).

[26] R. de Putter *et al.*, Astrophys. J. **761**, 12 (2012).

[27] C. Cheung and Y. Nomura, Phys. Rev. D **86**, 015004 (2012).