Constraining the p-Mode–g-Mode Tidal Instability with GW170817

B. P. Abbott et al.*
(LIGO Scientific Collaboration and Virgo Collaboration)

(Received 13 September 2018; revised manuscript received 30 October 2018; published 13 February 2019)

We analyze the impact of a proposed tidal instability coupling p modes and g modes within neutron stars on GW170817. This nonresonant instability transfers energy from the orbit of the binary to internal modes of the stars, accelerating the gravitational-wave driven inspiral. We model the impact of this instability on the phasing of the gravitational wave signal using three parameters per star: an overall amplitude, a saturation frequency, and a spectral index. Incorporating these additional parameters, we compute the Bayes factor (\(\ln B^p_{pg}\)) comparing our p-g model to a standard one. We find that the observed signal is consistent with waveform models that neglect p-g effects, with \(\ln B^p_{pg} = 0.03^{+0.70}_{-0.58}\) (maximum \textit{a posteriori} and 90% credible region). By injecting simulated signals that do not include p-g effects and recovering them with the p-g model, we show that there is a \(\approx 50\%\) probability of obtaining similar \(B^p_{pg}\) even when p-g effects are absent. We find that the p-g amplitude for 1.4 \(M_\odot\) neutron stars is constrained to less than a few tenths of the theoretical maximum, with maxima \textit{a posteriori} near one-tenth this maximum and p-g saturation frequency \(\sim 70\) Hz. This suggests that there are less than a few hundred excited modes, assuming they all saturate by wave breaking. For comparison, theoretical upper bounds suggest \(\lesssim 10^3\) modes saturate by wave breaking. Thus, the measured constraints only rule out extreme values of the p-g parameters. They also imply that the instability dissipates \(\lesssim 10^{51}\) erg over the entire inspiral, i.e., less than a few percent of the energy radiated as gravitational waves.

DOI: 10.1103/PhysRevLett.122.061104

\textit{Introduction.}—Detailed analysis of the gravitational-wave (GW) signal received from the first binary neutron star (NS) coalescence event (GW170817 [1]) constrains the tidal deformability of NSs and thus the equation of state (EOS) above nuclear saturation density [2–4]. Studies of NS tidal deformation typically focus on the linear, quasi-static tidal bulge induced in each NS by its companion. Such deformations modify the system’s binding energy and GW luminosity and thereby alter its orbital dynamics. The degree of deformation is often expressed in terms of the tidal deformability \(\Lambda \propto (R_i/m_i)^5\) of each component [5], or a particular mass-weighted average thereof (\(\Lambda\)) [2]. The strong dependence on compactness \(R/m\) means that a stiffer EOS, which has larger \(R\) for the same \(m\), imprints larger tidal signals than a softer EOS. Current analyses of GW data from the LIGO [6] and Virgo [7] detectors favor a soft EOS [3,8]. Specifically, Ref. [2] finds \(\Lambda \lesssim 730\) at the 90% credible level for all waveform models considered, allowing for the components to spin rapidly. The pressure at twice nuclear saturation density is also constrained to \(P = 3.5^{+7.7}_{-1.7} \times 10^{34}\) dyn/cm\(^2\) (median and 90% credible region) [3] assuming small component spins. In addition to GW phasing, the EOS dependence of \(\Lambda\) should correlate with postmerger signals [9], possible tidal disruptions, and kilonova observations [10]. Observed light curves for the kilonova suggest a lower bound of \(\Lambda \gtrsim 200\) [11,12].

Although some dynamical tidal effects are incorporated in these analyses (see, e.g., Refs. [2,13]), the impact of several types of dynamical tidal effects are neglected because they are believed to be small or have large theoretical uncertainties. These effects arise because tidal fields, in addition to raising a quasistatic bulge, excite stellar normal modes. Three such excitation mechanisms are (i) resonant linear excitation, (ii) resonant nonlinear excitation, and (iii) nonresonant nonlinear excitation (see, e.g., Ref. [14]). The first occurs when the GW frequency (the oscillation frequency of the tidal field) sweeps through a mode’s natural frequency (see, e.g., Refs. [15–22]). However, since the GW frequency increases rapidly during the late inspiral, the time spent near resonance is too short to excite modes to large amplitudes. As a result, for modes with natural frequencies within the sensitive bands of ground-based GW detectors, the change in orbital phasing is expected to be small (\(\Delta \Psi \lesssim 10^{-2}\) rad) unless the stars are rapidly rotating [17–19]. The impact of resonant nonlinear mode excitation (i.e., the parametric subharmonic instability) is likewise limited by the swiftness of the inspiral [23].

The proposed p-g tidal instability is a nonresonant, nonlinear instability in which the tidal bulge excites a low-frequency buoyancy-supported g mode and a high-frequency pressure-supported p mode [23–26]. It occurs in the inner core of the NS, where the stratification is weak and the shear due to the tidal bulge is especially susceptible to instability. Unlike resonantly excited modes, an unstable...
$p$-$g$ pair continuously drains energy from the orbit once excited, even after the orbital frequency changes significantly. There are many potentially unstable $p$-$g$ pairs, each becoming unstable at a different frequency and growing at a different rate. Although there is considerable uncertainty about the number of unstable pairs, their exact growth rates, and how they saturate, estimates suggest that the impact could be measurable with current detectors [27].

In this Letter, we investigate the possible impact of the $p$-$g$ instability on GW170817 using the phenomenological model developed in Ref. [27]. The model describes the energy dissipated by the instability within each NS, indexed by $i$, in terms of three parameters: (i) an overall amplitude $A_i$, which is related to the number of modes participating in the instability, their growth rates, and their saturation energies, (ii) a frequency $f_i$ corresponding to when the instability saturates, and (iii) a spectral index $n_i$ describing how the saturation energy evolves with frequency. In the section “Phenomenological model,” we describe our models in detail. In the section “Model selection,” we compare the statistical evidence for models that include the $p$-$g$ instability relative to those that do not.

In the section “Parameter inference,” we investigate the constraints on the $p$-$g$ parameters from GW170817, and in the section “Discussion,” we conclude.

Phenomenological model.—Following Ref. [27], we extend a post-Newtonian (PN) waveform by including a parametrized model of the $p$-$g$ instability. For the initial PN model, we use the TaylorF2 frequency-domain approximant (see, e.g., Ref. [28]) terminated at the innermost stable circular orbit, which includes the effects of linear tides ($\tilde{\Lambda}$) and spins aligned with the orbital angular momentum (the impact of misaligned spins on $p$-$g$ effects is not known). Waveform systematics between different existing approximants may be important for small $p$-$g$ effects. However, by comparing the waveform mismatches between several other models (TaylorF2, SEOBNR, PhenomDNRT, and PhenomPNRT, see Ref. [2]), we find these systematics induce waveform mismatches that correspond to $p$-$g$ phenomenological amplitudes roughly an order of magnitude smaller than the upper limits set by our analysis (see section “Parameter inference”). We expect the TaylorF2 approximant to be reasonably accurate and defer a complete analysis of waveform systematics to future work.

Assuming the $p$-$g$ effects are a perturbation to the TaylorF2 approximant, we find that they modify the phase in the frequency domain by

$$\Delta \Psi(f) = -\frac{2C_i}{3B^2(3-n_i)(4-n_i)} \left\{ \Theta_1 \left( \frac{f}{f_{\text{ref}}} \right)^{n_i-3} + (1-\Theta_1) \left( \frac{f}{f_{\text{ref}}} \right)^{n_i-3} \left[ (4-n_i) - (3-n_i) \left( \frac{f}{f_i} \right) \right] \right\} + (1 \leftrightarrow 2),
$$

where $f_i$ is the saturation frequency, $f_{\text{ref}} \equiv 100$ Hz is a reference frequency with no intrinsic significance, $C_i = [2m_i/(m_1 + m_2)]^{2/3}A_i$, $B = (32/5)(GMf_{\text{ref}}/c^3)^{2/3}$, $\mathcal{M} = (m_1 m_2)^{1/3}/(m_1 + m_2)^{1/5}$, and $\Theta_i = \Theta(f-f_i)$ where $\Theta$ is the Heaviside function. This approximant is slightly different than that of Ref. [27] because they incorrectly applied the saddle-point approximation to obtain the frequency-domain waveform from time-domain phasing [29]. This correction renders the $p$-$g$ instability slightly more difficult to measure than predicted in Ref. [27], although the observed behavior is qualitatively similar. Specifically, we find that in order to achieve the same $|\Delta \Psi|$, $A_i$ needs to be larger than Ref. [27] found by a factor of $\sim (4-n_i)$, although the precise factor also depends on the other $p$-$g$ parameters.

The $\Delta \Psi$ expression contains three types of terms: a constant term, a linear term $\sim (1-\Theta_1)f$, and a power-law term $\sim \Theta_1 f^{n_i-3}$. The constant term corresponds to an overall phase offset and is degenerate with the orbital phase at coalescence. The linear term corresponds to a change in the time of coalescence; because the $p$-$g$ instability transfers energy from the orbit to stellar normal modes, the binary inspirals faster than it would if the effect was absent. The power-law term accounts for the competition between the rate of $p$-$g$ energy dissipation and the rate of inspiral, both of which increase as $f$ increases. As argued in Ref. [27], we expect $n_i < 3$, which implies that the phase shift accumulates primarily at frequencies just above the “turn-on” (saturation) frequency $f \gtrsim f_i$.

When $n_i < 3$, $p$-$g$ effects are most important at lower frequencies whereas linear tides ($\tilde{\Lambda}$) and spins ($\chi_i = cS_i/Gm_i^2$, where $S_i$ is the spin-angular momentum of each component) have their largest impact at higher frequencies (see, e.g., Ref. [30]). The priors placed on the latter quantities can, however, affect our inference of $p$-$g$ parameters.

In order to account for a possible dependence on the component masses ($m_i$), we parametrize our model using a Taylor expansion in the $p$-$g$ parameters around $m_i = 1.4 M_\odot$ and sample from the posterior using the first two coefficients. Our model computes $A_i$ as

$$A_i(m_i) = A_0 + \left( \frac{dA}{dm} \right)_{1.4 M_\odot} (m_i - 1.4 M_\odot),
$$

and uses $A_0$ and $dA/dm$ instead of $A_1$ and $A_2$. The model uses similar representations for $f_i$ and $n_i$ in terms of the parameters $f_0$, $df/dm$, $n_0$, and $dn/dm$. We assume a uniform prior on $\log_{10} A_0$ within $10^{-10} \leq A_0 \leq 10^{-5.5}$, a uniform prior in $f_0$ within 10 Hz $\leq f_0 \leq 100$ Hz, and a uniform prior in $n_0$ within $-1 \leq n_0 \leq 3$. The priors on the first-order terms ($dA/dm, df/dm, dn/dm$) are the same as those in Ref. [27]; when $m_i \sim m_2$, they imply $A_1 \sim A_2$, etc.

We investigate GW170817 using data from several different frequency bands and with different spin priors, but unless otherwise noted we focus on results for data
above 30 Hz with $|x| \leq 0.89$. Throughout this Letter, results from GW170817 were obtained using the same data conditioning as Ref. [2], including the removal of a short-duration noise artifact from the Livingston data (Ref. [31] and discussion in Ref. [1]) along with other independently measured noise sources (see, e.g., Refs. [32–35]), calibration [36,37], marginalization over calibration uncertainties, and whitening procedures [38,39]. We use the publicly available LALInference software package throughout [40,41].

Model selection.---Using GW data from GW170817, we perform Bayesian model selection. We compare a model that includes linear tides, spin components aligned with the orbital angular momentum, and PN phasing effects up to 3.5 PN phase terms ($H_{pg}$) to an extension of this model that also includes $p$-$q$ effects ($H_{pq}$). Since we have nested models ($H_{pq}$ is obtained from $H_{pg}$ as $A_i \rightarrow 0$), we use the Savage-Dickey density ratio (see, e.g., Refs. [42–44]) to estimate the Bayes factor ($B_{pq}^{pg} = p(D|H_{pq})/p(D|H_{pg})$, where $D$ refers to the observed data). Because we use a uniform-in-$\log_{10}A_i$ prior, $H_{pq}$ does not formally include $A_i = 0$. Nonetheless, our lower limit on $A_i$ is sufficiently small that $H_{pq}$ is effectively nested in $H_{pg}$. Specifically, we sample from the model’s posterior distribution [40,41] and calculate

$$\lim_{\lambda \rightarrow 0} \left[ \frac{p(A_i|D, H_{pg})}{p(A_i|H_{pg})} \right] = \lim_{\lambda \rightarrow 0} \left[ \frac{1}{p(D|H_{pg})} \int d\theta df dn_1 p(D|\theta, A_i, f, n_1; H_{pg}) p(\theta|H_{pg}) p(f, n_1|A_i, H_{pg}) \right]$$

$$= \frac{1}{p(D|H_{pg})} \int d\theta p(D|\theta; H_{pq}) p(\theta|H_{pq}) \left[ \frac{p(\theta|H_{pq})}{p(\theta|H_{pg})} \right] \int df dn_1 p(f, n_1|A_i, H_{pq})$$

$$= \frac{p(D|H_{pq})}{p(D|H_{pg})} \left( \frac{p(\theta|H_{pq})}{p(\theta|H_{pg})} \right) p(\theta|D; H_{pq})$$

(3)

where $\theta$ refers to all parameters besides the $p$-$q$ phenomenological parameters; we note that $\int df dn p(f, n_1|A_i, H_{pg}) = 1$ $A_i$, and $\langle x \rangle_p$ denotes the average of $x$ with respect to the measure defined by $p$. Assuming that $p(\theta|H_{pq}) = p(\theta|H_{pg})$, we determine $\ln B_{pq}^{pg}$ from the ratio, as $A_i \rightarrow 0$, of the marginal distribution of $A_i$ a priori to the distribution a posteriori:

$$\ln B_{pq}^{pg} = \lim_{A_i \rightarrow 0} \left[ \ln p(A_i|H_{pq}) - \ln p(A_i|D, H_{pg}) \right].$$

(4)

This allows us to directly measure $\ln B_{pq}^{pg}$ by extracting $p(A_i|D, H_{pg})$ from Monte Carlo analyses with a known prior $p(A_i|H_{pq})$. We confirmed that this estimate agrees with estimates from both nested sampling [45] and thermodynamic integration [46].

Figure 1 shows $\ln B_{pq}^{pg}$ as a function of $f_{low}$, the minimum GW frequency considered. At a given $f_{low}$, we show the distribution of $\ln B_{pq}^{pg}$ due to the sampling uncertainty from the finite length of our MCMC chains. The solid and dashed curves correspond to the high-spin ($|x| \leq 0.89$) and low-spin ($|x| \leq 0.05$) priors discussed in Refs. [1–3].

For certain combinations of $f_{low}$ and $|x|$, we find $\ln B_{pq}^{pg} > 0$, suggesting $H_{pq}$ is more likely than $H_{pg}$. In order to assess how likely such values are, we calculate $\ln B_{pq}^{pg}$ for a large number of simulated, high-spin signals with $A_i = 0$ and distinct realizations of detector noise from times near GW170817. We find that simulated signals without $p$-$q$ effects can readily produce $\ln B_{pq}^{pg}$ at least as large as the ones we measured from GW170817. For example, for the 30 Hz high-spin data we obtain $\ln B_{pq}^{pg} = 0.031 \pm 0.58$ (maximum a posteriori and 90% credible region; bottom panel of Fig. 1), whereas approximately half of our simulated signals yield $\ln B_{pq}^{pg}$ at least this large, i.e., a false alarm probability (FAP) $\approx 50\%$. We focus on the 30 Hz, high-spin data because

![FIG. 1](image-url)

FIG. 1. Distributions of $\ln B_{pq}^{pg}$ due to sampling uncertainty when analyzing GW170817 data with different values of $f_{low}$. The solid red curves assume high-spin priors ($|x| \leq 0.89$) and the dashed blue curves assume low-spin priors ($|x| \leq 0.05$).
it corresponds to the largest bandwidth investigated and the largest signal-to-noise ratio. The high-spin prior is the most inclusive prior considered, and therefore allows the most model freedom when fitting p-g effects.

In our model of the instability, the phase shift $\Delta \Psi$ accumulates primarily at frequencies just above the saturation frequency $f \gtrsim f_i$. Therefore, if it is present, its impact should become more apparent as we decrease the minimum GW frequency considered from $f_{\text{low}} \gg f_i$ to $f_{\text{low}} \lesssim f_i$. We do see some indication of this behavior in Fig. 1. However, we note that if our phenomenological model breaks down at weaker in $f_i$, due to poor modeling of the presaturation behavior (e.g., if our step-function turn-on at $f_i$ is not a good approximation to the instability’s induced phase shift), we might expect $\ln B_{\text{pp}}^H$ to decrease as we lower $f_{\text{low}}$ below $f_i$. If the fidelity of our model is sufficiently poor, we could be insensitive to p-g effects even at frequencies above $f_{\text{low}}$.

Parameter inference.—We now investigate the constraints obtained from GW170817. Figure 2 shows the joint posterior distributions for both $\mathcal{H}_{pg}$ and $\mathcal{H}_{ppg}$. We find that $\mathcal{H}_{pg}$ and $\mathcal{H}_{ppg}$ yield similar posterior distributions for all non-p-g parameters, including both extrinsic and intrinsic parameters. The constraints on the chirp mass ($\mathcal{M}$), effective spin $\chi_{\text{eff}} = (m_1 \chi_1 + m_2 \chi_2)/(m_1 + m_2)$, and $\tilde{\Lambda}$ are slightly weaker in $\mathcal{H}_{pg}$ than $\mathcal{H}_{ppg}$. This is because $\mathcal{H}_{pg}$ provides extra freedom to the signal’s duration in the time domain.

![FIG. 2. Posterior distributions for $\mathcal{H}_{ppg}$ (red) and $\mathcal{H}_{pg}$ with Hanford, Livingston, and Virgo data (thick black, gray shading). Hanford data only (dark blue), and Livingston data only (light blue) using GW data above 30 Hz, $|\chi| \leq 0.89$, and a uniform-in-log $A_0$ prior. Left: a subset of parameters shared by $\mathcal{H}_{pg}$ and $\mathcal{H}_{ppg}$. Right: a subset of parameters belonging only to $\mathcal{H}_{ppg}$. We only show one-dimensional posteriors for the single instrument data, although the multidimensional posteriors are similarly consistent with the full $\mathcal{H}_{ppg}$ data. Contours in the two-dimensional distributions represent 10%, 50%, and 90% confidence regions.]

Regarding the $p-g$ parameters, we find a noticeable peak near $A_0 \sim 10^{-7}$ with a flat tail to small $A_0$. We find $A_0 \lesssim 3.3 \times 10^{-7}$ assuming a uniform-in-log $A_0$ prior and $A_0 \lesssim 6.8 \times 10^{-7}$ assuming a uniform-in-$A_0$ prior, both at 90% confidence. The upper limit with a uniform-in-$A_0$ prior is larger only because we weight larger values of $A_0$ more a priori than with a uniform-in-log $A_0$ prior. We also find a peak at $f_0 \sim 70$ Hz. The peaks persist when we analyze the data from each interferometer separately, with reasonably consistent locations and shapes (Fig. 2). However, we find that the simulated signals with $A_i = 0$ can produce similar peaks, suggesting they may be due to noise alone. Similar to Ref. [27], we find that $n_i$ is not strongly constrained and the gradient terms in the Taylor expansions are not measurable.

Theoretical arguments suggest an upper bound of $A_0 \lesssim 10^{-6}$ [27]. Therefore, our $A_0$ constraint only rules out the most extreme values of the $p-g$ parameters.

Discussion.—While GW170817 is consistent with models that neglect p-g effects, it is also consistent with a broad range of p-g parameters. The constraints from GW170817 imply that there are $\lesssim 200$ excited modes at $f = 100$ Hz, assuming all modes grow as rapidly as possible and saturate at their breaking amplitudes ($\lambda = \beta = 1$ in Eq. (7) of Ref. [27]) and that the frequency at which modes become unstable is well approximated by $f_0$. For comparison, theoretical arguments suggest an upper bound of $\sim 10^3$
modes saturating by wave breaking [27]. More modes may be excited if they grow more slowly or saturate below their wave breaking energy.

We can also use the measured constraints to place upper limits on the amount of energy dissipated by the \( p-g \) instability. As Fig. 3 shows, \( p-g \) effects dissipate \( \geq 2.7 \times 10^{41} \) erg throughout the entire inspiral at 90% confidence. In comparison, GWs carry away \( \geq 10^{53} \) erg. This implies time-domain phase shifts \( |\Delta \phi| \lesssim 7.6 \) rad (0.6 orbits) at 100 Hz and \( |\Delta \phi| \lesssim 32 \) rad (2.6 orbits) at 1000 Hz after accounting for the joint uncertainty in component masses, spins, linear tides, and \( p-g \) effects.

A \( g \) mode with natural frequency \( f_g \) is predicted to become unstable at a frequency \( f_{\text{crit}} \approx 45 \) Hz \( (f_{\text{dy}}/10^{-4} \lambda f_{\text{dy}})^{1/2} \), where \( f_{\text{dy}} \) is the dynamical frequency of the NS and \( \lambda \) is a slowly varying function typically between 0.1–1 [25,27]. Since the modes grow quickly, the frequency at which the instability saturates is likely close to the frequency at which the modes become unstable \( (f_0 \approx f_{\text{crit}}) \). If we assume that the observed peak near \( f_0 \approx 70 \) Hz is not due to noise alone, then the maximum \textit{a posteriori} estimate for \( f_0 \) along with approximate values for the masses \( (1.4 M_\odot) \) and radii \( (11 \text{ km}) \) of the components [3] imply \( f_g \approx 0.5 \) Hz.

With several more signals comparable to GW170817, it should be possible to improve the amplitude constraint to \( A_0 \lesssim 10^{-7} \). Obtaining even tighter constraints will likely require many more detections, especially since most events will have smaller SNR. Future measurements will also benefit from a better understanding of how the instability saturates. To date, there have only been detailed theoretical studies of the instability’s threshold and growth rate [23–26], not its saturation. As a result, we cannot be certain of the fidelity of our phenomenological model.

While this Letter was in review, related work was posted [47] with the conclusion that the \( \mathcal{H}_{pg} \) model is strongly favored over the \( \mathcal{H}_{pg} \) model by a factor of at least \( 10^4 \). In Ref. [48], some of the authors of this work investigate the origin of the discrepancy by analyzing publicly available posterior samples from Ref. [47]. Contrary to the claims in Ref. [47], they find that samples from Ref. [47] yield \( B_{pg}^* \sim 1 \) and therefore conclude that their posterior data, like what is presented here, do not disfavor the \( \mathcal{H}_{pg} \) model. Reference [48] suggests that the error stems from using too few temperatures when implementing thermodynamic integration.

The authors gratefully acknowledge the support of the United States National Science Foundation (NSF) for the construction and operation of the LIGO Laboratory and Advanced LIGO as well as the Science and Technology Facilities Council (STFC) of the United Kingdom, the Max-Plank-Society (MPS), and the State of Niedersachsen, Germany for support of the construction of Advanced LIGO and construction and operation of the GEO600 detector. Additional support for Advanced LIGO was provided by the Australian Research Council. The authors gratefully acknowledge the Italian Istituto Nazionale di Fisica Nucleare (INFN), the French Centre National de la Recherche Scientifique (CNRS) and the Foundation for Fundamental Research on Matter supported by the Netherlands Organisation for Scientific Research, for the construction and operation of the Virgo detector and the creation and support of the EGO consortium. The authors also gratefully acknowledge research support from these agencies as well as by the Council of Scientific and Industrial Research of India, the Department of Science and Technology, India, the Science & Engineering Research Board (SERB), India, the Ministry of Human Resource Development, India, the Spanish Agencia Estatal de Investigación, the Vicepresidència i Conselleria d’Innovació, Recerca i Turisme and the Conselleria d’Educació i Universitat del Govern de les Illes Balears, the Conselleria d’Educació, Investigació, Cultura i Esport de la Generalitat Valenciana, the National Science Centre of Poland, the Swiss National Science Foundation (SNSF), the Russian Foundation for Basic Research, the Russian Science Foundation, the European Commission, the European Regional Development Funds (ERDF), the Royal Society, the Scottish Funding Council, the Scottish Universities Physics Alliance, the Hungarian Scientific Research Fund (OTKA), the Lyon Institute of Origins (LIO), the Paris Île-de-France Region, the National Research, Development and Innovation Office Hungary (NKFI), the National Research Foundation of Korea, Industry Canada and the Province of Ontario through the Ministry of Economic Development and Innovation, the Natural Science and Engineering Research Council Canada, the Canadian Institute for Advanced Research, the Brazilian Ministry of Science, Technology, Innovations, and Communications, the International Center.
for Theoretical Physics South American Institute for Fundamental Research (ICTP-SAIFR), the Research Grants Council of Hong Kong, the National Natural Science Foundation of China (NSFC), the Leverhulme Trust, the Research Corporation, the Ministry of Science and Technology (MOST), Taiwan and the Kavli Foundation. The authors gratefully acknowledge the support of the NSF, STFC, MPS, INFN, CNRS and the State of Niedersachsen, Germany for provision of computational resources. N. Weinberg was supported in part by NASA Grant No. NNX14AB40G.

†Deceased.