Quantum interactions between a laser interferometer and gravitational waves

Belinda Pang and Yanbei Chen

Theoretical Astrophysics and Walter Burke Institute for Theoretical Physics, M/C 350-17,
California Institute of Technology, Pasadena, California 91125, USA

LIGO's detection of gravitational waves marks a first step in measurable effects of general relativity on quantum matter. In their current operation, laser interferometer gravitational-wave detectors are already quantum limited at high frequencies, and planned upgrades aim to decrease the noise floor to the quantum level over a wider bandwidth. This raises the interesting idea of what a gravitational-wave detector—or, more generally, an optomechanical system—may reveal about gravity beyond detecting gravitational waves from highly energetic astrophysical events, such as its quantum versus classical nature. In this paper we develop a quantum treatment of gravitational waves and their interactions with the detector. We show that the treatment recovers known equations of motion in the classical limit for gravity, and we apply our formulation to study the system dynamics, with a particular focus on the implications of gravity quantization. Our framework can also be extended to study alternate theories of gravity and the ways in which their features manifest themselves in a quantum optomechanical system.

I. INTRODUCTION

With LIGO's detection of gravitational waves [1], there has been interest in using gravitational-wave detectors (including, e.g., VIRGO [2] and KaGRA [3]) to study not only astrophysical sources, but the nature of gravity itself, including modified theories [4–6] and quantum gravity [7–9]. There are also important questions related to the quantum nature of the LIGO probe in terms of the implications for its sensitivity as a measurement device [10–13], as well as the possibility of it being a test bed to study the interplay of quantum mechanics and gravity.

However, up to this point, the quantum interaction between gravitational waves and a LIGO-like (optomechanical) system has not been carefully studied from a general-relativistic point of view, despite the interest in using optomechanical systems to study low-energy gravity effects. To date, the literature comprises mainly theoretical studies of its interaction with classical Newtonian gravity [14] and phenomenological models of quantum or semi-classical gravity [15–18]. There are also general-relativistic quantum formulations of weak gravity interactions with bosonic fields [19–21], but these treatments do not easily extend to more complex matter systems. In particular, Oniga *et al.* [22] derived the master equation for a bound scalar field, similar to an optical cavity, but in crucial contrast to LIGO's operation, did not allow for the boundary length to change. Furthermore, because these treatments focus on the decoherence of quantum matter, they take the view of the gravitational field as an equilibrated bath in which the effects of interaction perturbations cannot be observed.

In this paper, we develop a canonical formulation of linear quantum gravity from Einstein's theory of general relativity interacting with a quantum LIGO-like system (probe) in processes involving gravitational waves (GWs), which in principle can be extended to study interactions of quantum LIGO with GWs from other gravitational theories derivable from the action principle, such as scalar-tensor theories [23]. In contrast to the conventional role of GWs as a classical and predetermined signal in LIGO, as well as treatments of quantum gravity as a thermal bath coupled to quantum matter, our formulation treats both the matter probe and the GW field on equal footing with dynamical degrees of freedom (d.o.f.) in an enlarged Hilbert space. Importantly, this treatment allows the matter probe and GW field to act mutually on each other, as compared to the previous scenarios. This paper focuses their dynamics, and in particular examines the testable physical implications of GW quantization. We find that the probe → GW field direction of interaction recovers Einstein's field equations for the generation of GWs, but where the stress-energy tensor $T_{\mu\nu}$ is quantum. Conversely, the GW → probe direction of interaction recovers the same equations for LIGO's output field in the presence of a classical GW signal, as was calculated previously [24–27].

That the formalism in the classical gravity limit recovers well-known equations of GW detection and generation provides a check on its validity, but it additionally predicts effects for which the quantum nature of the GW field

becomes essential. Specifically, we find that the mutual interaction leads to quantum coherent backaction effects on the probe in such a way that requires the presence of quantum GW fluctuations in order to preserve commutation relations. Furthermore, these backaction effects can be shown to be analogous to the radiation reaction damping that comes from classical corrections to the Newtonian potential [28]. This suggests that in order to include the effects of weak GR corrections to Newtonian potential on quantum matter consistently and without violating canonical quantization, those perturbations must themselves be quantum.

As an interpretational tool, our formulation offers an alternative (though physically equivalent) perspective of the detection process—in previous treatments that considered test masses as quantum objects, LIGO's GW detection [24–27] was viewed in the Newtonian gauge, where the GW signal is understood to act as a tidal force on LIGO's cavities' mirrors (test mass), whose motion then modulates the field inside the cavity. Most treatments using the Newtonian gauge were valid in the long-wavelength approximation, where the wavelength of the gravitational wave is much greater than the interferometer's arm length, although it has been shown that such treatments can be extended to full validity, such that they are equivalent to results in the traceless-transverse (TT) gauge [29].

In our formulation in the TT gauge, the GWs interact directly with the cavity field—similar to *earlier* treatments of LIGO, e.g., that given by Ref. [30]. While all measurable quantities are the same in either gauge, the latter facilitates an intuitive and straightforward derivation of the probe's ultimate quantum-limited measurement sensitivity, known as its quantum Cramer-Rao bound (qCRB) [10–13]. This bound has interesting relations with GW radiation and decoherence, which will be discussed in an accompanying paper.

This paper is divided into the following sections: Section II provides a description of the physical system along with definitions and notations used in this paper. Section III expounds our theoretical framework whereby we develop a Hamiltonian formulation of the interacting system, which we apply in Secs. IV and V to separately study the GW field and probe dynamics in the presence of their mutual interaction.

II. DESCRIPTION OF SYSTEM

In this section, we shall describe a second-generation laser interferometer gravitational-wave detector, like those in LIGO, VIRGO, and KaGRA.

A. The optomechanical system

Let us consider a Michelson interferometer containing a Fabry-Perot cavity in each of its two arms shown in Fig. 1a, which additionally has a power-recycling mirror to increase

the power circulating inside the arm cavities, as well as a signal-recycling mirror to increase the bandwidth of detection [1].

It has been shown that the antisymmetric mode of the interferometer which carries the GW signal and the quantum noises, including the power- and signal-recycling and resonant sideband extraction enhancements [31–35], can be mapped to a single detuned Fabry-Perot cavity with an effective input mirror and a perfectly reflective end mirror shown in Fig. 1b [36]. This introduces errors of $O(l_{\rm SRC}/L)$ for signal-recycling cavity length $l_{\rm SRC}$ and cavity arm length L, and so the assumption is valid when $l_{\rm SRC} \ll L$, which is the case in experiment.

Additionally, for simplicity, we attribute the effects of radiation pressure to the end mirror alone, which is possible if we assume that the input mirror is infinitely massive, or by reducing the end mirror to half its actual mass, which introduces errors of $\max\{\Omega L/c, T\}$ for input transmissivity T [36]. In this way, we can assume that the input mirror falls along its geodesic. Then, in the traceless-transverse gauge, if we choose our coordinate frame so that its origin coincides with the position of the input mirror at some point along its worldline and its coordinate velocity is initially zero, then the coordinates of the input mirror are fixed in time.

Finally, we choose the cavity axis to be the x direction, along which we constrain the mirror motion described by its center-of-mass coordinate, thereby allowing us to model the mirror as a massive point particle (valid for $t_m \ll \lambda_{\rm GW}$ for mirror thickness t_m). In reality, the LIGO mirrors are suspended pendulums, but the error in making this assumption is $O(q/l_p)$, where q is mirror displacement due to GW and radiation pressure and l_p is the pendulum length. In summary, with the stated errors, the signal and quantum noise analysis for the LIGO Michelson interferometer can be mapped onto that for a single one-dimensional Fabry-Perot cavity and all the radiation pressure effects attributed to the end mirror. We choose the origin of our coordinate system to be the position of the input mirror, which means that it is also not affected by metric perturbations, and we can therefore hold this coordinate fixed. We now have mapped LIGO to a basic optomechanical system [37].

B. Inclusion of gravity

To describe space-time geometry, we assume weak metric perturbations about flat spacetime, such that the metric is given by $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, where $\eta_{\mu\nu}$ is the Minkowski metric with signature $(-+++)$ and $h_{\mu\nu}$ is a small perturbation which we treat up to linear order. Although LIGO operates in Earth's weak gravitational field, this is a constant longitudinal component whose effect on the test masses is balanced directly by the tension in the pendulums. Since we only consider linear gravitational perturbation, it is safe to ignore Earth's gravity
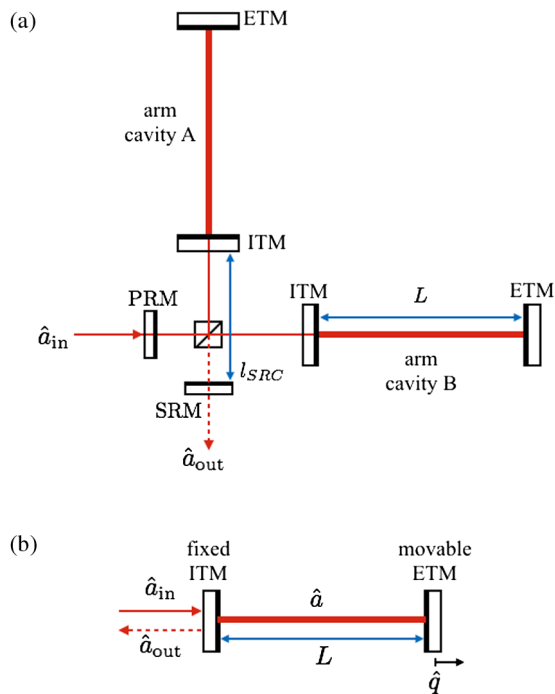
FIG. 1. Schematic of a second-generation laser interferometer gravitational-wave detector. (a) The full Michelson interferometer in its current configuration with power- and signal-recycling mirrors (PRM and SRM) and the two Fabry-Perot arm cavities. Here L denotes the length of the arm cavity, and l_{SRC} denotes the length of the signal-recycling cavity (shown here not to scale). The arm cavities' input mirror (ITM) has transmissivity T, and its end mirror (ETM) is perfectly reflective with $R = 1$. For low frequencies Ω of the GW wave such that $\Omega L/c \ll 1$, and for $T \ll 1$, $l_{SRC} \ll L$, the quantum inputs and outputs of the schematic in (a) can be mapped to those of a single one-dimensional Fabry-Perot cavity as shown in (b).

because it does not couple to gravitational-wave contributions to $h_{\mu\nu}$ at this order.

C. Terminology

Throughout this paper and unless otherwise noted, greek indices μ, ν, ρ, etc. (with the exception of λ) denote spacetime components of vectors and tensors; latin indices i, j, k, etc. denote purely spatial components; λ denotes graviton polarizations; boldface form **v** represents 3-vectors; and $\vec{v}$ represents 4-vectors. We also use Einstein subscript summation notation where contraction is with respect to the background Minkowski metric (e.g., $M_{\mu\nu}M^{\mu\nu}$). Repeated spatial indices denote summation regardless of upper or lower position (e.g., $v_i v_i$).

For terminology, "optical mode" refers to the optical field inside the Fabry-Perot cavity; "test mass" refers to the end mirror of the cavity; "probe" refers to the optomechanical system comprising the optical mode, the test mass, and their interaction; "system" without a modifier refers to the GW field and the probe together; "pump field" refers to

the optical input to the cavity; and "output field" refers the cavity output on which measurement is performed.

III. THEORETICAL FRAMEWORK

To study the interaction of weak gravity with macroscopic matter systems in finite time, we can use the canonical quantization formulation, even though a full quantum theory of gravity is not yet available [38]. Canonical quantization can be quite straightforward in systems with well-defined physical coordinates and velocities that appear in the Lagrangian in quadratic form, but here there are two difficulties.

First is the fact that gravity has coordinate (or gauge) d.o.f., which is mathematically reflected in the singularity of its Lagrangian. This means a physical state can be represented by multiple points that form a trajectory in phase space. Dirac is credited with developing the Hamiltonian formulation of such gauge theories, in which the degeneracies in phase space due to gauge d.o.f. can be eliminated by restricting the representation to a hypersurface which itself is foliated by gauge orbits. Quantization can proceed in the usual way by ignoring the existence of gauge freedom, but physical quantum states must satisfy constraint conditions which ensure that the physical Hilbert space slices across the gauge orbits [39]. Applying this approach to linearized gravity, Gupta derived a Hamiltonian for a pure gravitational field and the constraint conditions that must be satisfied by physical gravitational states. He also demonstrated that a pure gravitational field has only two physical gravitons, although more exist in virtual states in the presence of interaction [40]. However, we can greatly simplify the quantization procedure by noting that, since we are only interested in studying leading-order interactions involving incoming or outgoing gravitons, we can restrict the gravitational field to its two physical polarizations at the level of the action, eliminating its longitudinal and timelike components, which obviates the need for constraint conditions on physical states. While such a method would not give the correct result for interactions mediated by virtual gravitons, i.e., self-gravity of the detector, it is appropriate for studying leading-order interactions between the interferometer and gravitational waves.

The second issue lies in obtaining the Hamiltonian for the optomechanical interaction between the test mass and the optical mode. To our best knowledge, to derive this interaction, there currently only exist procedures which assume that the equations of motion for a cavity are known *a priori*, whereupon a suitable Lagrangian producing those equations is constructed [41]. However, since our purpose is to study the unknown behavior of an interacting system, the equations of motion must follow *from* the action instead of preceding it, and we develop an alternative approach so that the equations for all dynamical quantities of the system follow consistently from a canonical formulation beginning with the action.

A. Gauge fixing for gravitational field

We begin with the linearized Einstein-Hilbert action for the metric in the harmonic gauge with $\partial_\mu h^{\mu\nu} = 0$, and write

$$S_{\rm EH} = -\frac{c^4}{32\pi G} \int d^4x \left[\frac{1}{2} \partial_\mu h_{\alpha\beta} \partial^\mu h_{\alpha\beta} - \frac{1}{4} \partial_\mu h \partial^\mu h \right], \quad (1)$$

where $h = h^\mu{}_\mu$ is the tensor trace. As discussed previously, to leading order the interactions between the interferometer and gravitational waves only involve the physical polarizations of the field, which we have the freedom to express in any gauge. Choosing the traceless-transverse (TT) gauge, we eliminate any timelike component of the field and expand its spatial components in the Fourier domain as

$$h_{ij}^{\rm TT}(t, \mathbf{x}) = \int \frac{d^3\mathbf{k}}{\sqrt{(2\pi)^3}} \tau_{ij}^\lambda(\mathbf{k}) h_\lambda(t, \mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad (2)$$

where the index $\lambda = +, \times$ denotes the polarization. The tensor $\tau_{ij}^\lambda(\mathbf{k})$ is the unit tensor for the $\mathbf{k}$-mode component, and it satisfies orthogonality, transverse, and traceless conditions:

$$\tau_{ij}^\lambda \tau_{jk}^{\lambda'} = \delta_{\lambda,\lambda'} \delta_{ik}, \qquad \mathbf{k} \cdot \boldsymbol{\tau}^\lambda(\mathbf{k}) = 0, \qquad {\rm Tr}[\boldsymbol{\tau}^\lambda] = 0. \quad (3)$$

Finally, the Einstein-Hilbert action can be rewritten as

$$S_{\rm EH} = \frac{c^4}{32\pi G} \int dt \int_{\frac{1}{2}} d^3\mathbf{k} \left[\frac{1}{c^2} |\dot{h}_\lambda(t, \mathbf{k})|^2 - \mathbf{k}^2 |h_\lambda(t, \mathbf{k})|^2 \right]$$

$$\equiv \int dt \int_{\frac{1}{2}} d^3\mathbf{k} \mathcal{L}_{\rm GW}^{(0)}(t, \mathbf{k}). \quad (4)$$

We remark that h is related to h^* by

$$h_\lambda^*(t, \mathbf{k}) \tau_{ij}^\lambda(\hat{\mathbf{k}}) = h_\lambda(t, -\mathbf{k}) \tau_{ij}^\lambda(-\hat{\mathbf{k}}). \quad (5)$$

Therefore, summing $h_\lambda \tau_{ij}^\lambda$ over all of $\mathbf{k}$ space is physically equivalent to summing $h_\lambda \tau_{ij}^\lambda$ and $h_\lambda^* \tau_{ij}^\lambda$ over half of $\mathbf{k}$ space. The latter method allows us to treat h_λ and h_λ^* as independent d.o.f., in a similar approach to that of Ref. [42] for the Hamiltonian formulation of electrodynamics.

B. Optomechanical interaction

In this section, we briefly summarize the first-principles derivation of the optomechanical interaction between the optical mode and test mass from the electromagnetic (EM) field action. The details of the derivations, including relativistic corrections and extension to multimodes, will be published in an accompanying paper. (Note that this derivation does not follow the work of Law [41] because, due to the presence of other interactions, we cannot posit *a priori* equations of motion for the test mass and optical mode as is done in Ref. [41].) For simplicity, we present the

derivation in Minkowski space, although adding the metric perturbation to our analysis is straightforward. As we will show, the optomechanical interaction is hidden in the spatial boundary conditions of the EM field. With the appropriate coordinate transformation, the boundary condition appears as an explicit term in the action instead of being embedded in the integration limits. We consider the ideal case, where the cavity has perfectly reflective boundaries $R = 1$, although we will relax this assumption later on to allow transmission of the pump and output fields.

Let us first write down the EM action in Minkowski space, denoting the EM vector potential by $\mathcal{A}^\mu$. Since there are no charged currents, we can apply the Coulomb gauge and set the time component of the vector potential to zero, or $\mathcal{A}^0 = 0$, and write

$$S_{\rm EM}^\eta = \frac{1}{2\mu_0} \int_{V_{\rm cav}} d^4x \left[\frac{1}{c^2} \dot{\mathcal{A}}_j^2 - (\partial_i \mathcal{A}_j)^2 \right] = \int dt L_{\rm EM}^\eta, \quad (6)$$

with the η superscript denoting Minkowski space. Importantly, $V_{\rm cav}$ specifies the spatial limits of the EM field contained inside the Fabry-Perot cavity. We can approximate this field as having a constant mode profile transverse to the cavity axis, so that only the part of the field which propagates along the cavity axis is dynamical. The vector potential is then separable as $\mathcal{A}_j(t, \mathbf{x}) = u(y, z) A_j(t, x)$, where j can only take values of y, z to satisfy the Coulomb gauge condition. Then, defining the mode volume $\mathcal{U} = \int dy \int dz |u(y, z)|^2$, the Lagrangian $L_{\rm EM}^\eta$ in Eq. (6) can be written as an integral along the cavity axis:

$$L_{\rm EM}^\eta = \frac{\mathcal{U}}{2\mu_0} \int_0^{L'} dx \left[\frac{1}{c^2} \dot{A}_j^2 - \partial_x A_j^2 \right], \quad (7)$$

where L' is the coordinate length of the cavity, or equivalently the position of the test mass, and is a dynamical quantity. The optomechanical interaction derives from this dynamical boundary condition, which comes from the physical constraint that the EM field must vanish at the cavity's perfectly reflecting mirrors, such that $A_j(t, 0) = A_j(t, L') = 0$. We perform the following coordinate transformation so that the test-mass position appears explicitly in the Lagrange density (note that by "coordinate transformation," we are referring to the coordinates which themselves and whose velocities appear in the Lagrangian, and not spacetime coordinates):

$$A_j(t, x) = \frac{c^2}{\omega_0 L'} \Phi_j(t) \sin(\kappa x), \qquad \kappa = \frac{n\pi}{L'}, \qquad n \in \mathbb{Z}, \quad (8)$$

where $\omega_0 = n\pi c / L$ is the resonant frequency of the cavity at its equilibrium length L. The coordinate transformation of Eq. (8) automatically satisfies the boundary conditions and separates the spatial part of the field, which is subject to a time-dependent boundary condition from the naturally

time-varying part. The coordinate $\Phi_j(t)$ is no longer a field defined at each point in spacetime, but now represents excitations of the spatially extended optical mode. The collection of all modes $\Phi_j(t)$ with different values of n contains the same information as $A_j(t, x)$, but in this paper we only consider a single mode which is closest to the pumping frequency.

Then, defining $q = L' - L$ to be the motion of the test mass about equilibrium and substituting Eq. (8) into Eq. (7), to leading order in q/L, we find $L_{\text{EM}}^{\eta} = L_{\text{EM}}^{(0)} + L_{\text{OM}}$, where

$$L_{\text{EM}}^{(0)} = \frac{\mathcal{U}}{2\mu_0} \left[\frac{1}{2\omega_0} \dot{\Phi}_j^2(t) - \frac{\omega_0}{2} \Phi_j^2(t) \right], \tag{9a}$$

$$L_{\text{OM}} = \frac{\mathcal{U}}{2\mu_0} \frac{\omega_0 q}{L} \Phi_j^2(t), \tag{9b}$$

where summation over j (through y and z) is implied. The test-mass position q now appears explicitly in the Lagrangian, and we also find the optomechanical interaction term L_{OM}.

C. Interaction between GWs and probe

The probe consists of the optical mode and the test mass, whose actions in perturbed spacetime can be written as $S_{\text{probe}} = S_{\text{EM}}^{\eta} + S_{\text{EM}}^{h} + S_q$, where S_{EM}^{η} was the EM action in Minkowski space and was defined in Eq. (6), and S_{EM}^{h} is the first-order term in the expansion of $S_{\text{EM}} \propto -\int d^4x \sqrt{-g} g_{\alpha\mu} g_{\beta\nu} F^{\mu\nu} F^{\alpha\beta}$ with respect to $h_{\mu\nu}$, which we write below:

$$S_{\text{EM}}^{h} = -\frac{1}{2\mu_0} \int_{V_{\text{cav}}} d^4x \, h_{ij} \left[\frac{1}{c^2} \dot{\mathcal{A}}_i \dot{\mathcal{A}}_j - (\partial_i \mathcal{A}_k)(\partial_j \mathcal{A}_k) \right.$$
$$\left. - (\partial_k \mathcal{A}_i)(\partial_k \mathcal{A}_j) \right] = \int dt \, L_{\text{EM}}^{h}. \tag{10}$$

Similarly, we expand the action for the test mass $S_q = -mc \int d\tau$ to leading order in $h_{\mu\nu}$ and obtain

$$S_q = \frac{m}{2} \int dt [\dot{\mathbf{x}}_q^2 + h_{ij}(t, \mathbf{x}_q) \dot{x}_q^i \dot{x}_q^j] = \frac{m}{2} \int dt \, \dot{q}^2, \tag{11}$$

where to obtain the second equality we have ignored terms of $O(v^2/c^2)$ as well as the test mass's d.o.f. of motion in the y, z directions, which are noninteracting and trivial. We write the corresponding Lagrangian as $L_q^{(0)} = m\dot{q}^2/2$.

Thus, we find that the interaction between the GW and the probe only concerns the EM field, and, denoting this interaction by $L_{\text{GW}}^{\text{int}}$, we have $L_{\text{GW}}^{\text{int}} = L_{\text{EM}}^{h}$. Substituting h_{ij} in Eq. (10) with its expansion into transverse-traceless Fourier modes given in Eq. (2), and performing the same coordinate transformation to the EM vector potential as given in Eq. (8), we obtain

$$L_{\text{GW}}^{\text{int}} = -\frac{\mathcal{U}}{2\mu_0} \int_{\frac{1}{2}} d^3\mathbf{k} [J_{ij}^{\lambda}(\mathbf{k}) h_{\lambda}(t, \mathbf{k}) + J_{ij}^{\lambda*}(\mathbf{k}) h_{\lambda}^*(t, \mathbf{k})]$$
$$\times \left[\frac{1}{2\omega_0} \dot{\Phi}_i \dot{\Phi}_j - \frac{\omega_0}{2} (\Phi_i \Phi_j + \delta_{ix} \delta_{jx} \Phi_k^2) \right], \tag{12}$$

where $J_{ij}^{\lambda}(\mathbf{k})$ is a GW mode profile function and is given by

$$J_{ij}^{\lambda}(\mathbf{k}) = \frac{-i(e^{ik_xL} - 1)}{k_xL} \frac{\tau_{ij}^{\lambda}(\mathbf{k})}{\sqrt{(2\pi)^3}}. \tag{13}$$

The factor of $-i(e^{ik_xL} - 1)/k_xL$ derives from the variation of the gravitational wave over the spatial extent of the optical mode. For long-wavelength GWs such as those from LIGO's astrophysical sources, $k_xL \ll 1$, and $J_{\lambda}(\mathbf{k})$ simply reduces to the interacting polarization tensor component. However, in order to study backaction due to GW radiation from LIGO itself, one must include this factor to ensure convergence, since the backaction effect requires that radiated GW vary over the length of the cavity.

D. Canonical quantization

The full Lagrangian for the system is then

$$L = L_q^{(0)} + L_{\text{EM}}^{(0)} + \int dt \int_{\frac{1}{2}} d^3\mathbf{k} \mathcal{L}_{\text{GW}}^{(0)}(t, \mathbf{k}) + L_{\text{OM}} + L_{\text{GW}}^{\text{int}}. \tag{14}$$

We then proceed with canonical quantization, first performing a Legendre transform to identify conjugate pairs $\{h_{\lambda}(\mathbf{k}), \Pi_{\lambda}^*(\mathbf{k})\}, \{\Phi_i, \mathbb{P}_i\}, \{q, p\}$. We remark that although the photon polarizations appear coupled in $L_{\text{GW}}^{\text{int}}$ in Eq. (12), the interactions between the different polarizations occur on much shorter timescales than those of interest. To show this, and also to express the Hamiltonian in more familiar variables, we perform the following canonical transformation on the EM conjugate pair by defining $\alpha_1^{(i)}, \alpha_2^{(i)}$ such that

$$\Phi_i = \sqrt{\frac{2\mu_0}{\mathcal{U}}} [\alpha_1^{(i)} \cos \omega_0 t + \alpha_2^{(i)} \sin \omega_0 t], \tag{15a}$$

$$\mathbb{P}_i = -\sqrt{\frac{\mathcal{U}}{2\mu_0}} [\alpha_1^{(i)} \sin \omega_0 t - \alpha_2^{(i)} \cos \omega_0 t], \tag{15b}$$

where the superscript (i) represents photon polarization. Anticipating that the optical mode is driven by a resonant pump field, here we have separated the fast time dependence of the pump which is rotating at ω_0 in the tuned configuration, thereby going into the interaction picture with respect to the optical carrier frequency. In this corotating frame we find terms oscillating at $2\omega_0$. Since these effects occur on much shorter timescales than the

interaction dynamics, for weak interactions we ignore them under the rotating wave approximation (RWA) [43], and we find that the two photon polarizations interact with the test mass and GWs independently and in identical ways. We therefore suppress the photon polarization superscript and obtain the Hamiltonian $H = H_q^{(0)} + H_{GW}^{(0)} + H_{OM} + H_{GW}^{int}$. Here $H_q^{(0)}$ and $H_{GW}^{(0)}$ are the free Hamiltonians for the test mass and metric perturbation, respectively; $H_{OM} = -\omega_0(\hat{\alpha}_1^2 + \hat{\alpha}_2^2)\hat{q}/2L$ is the optomechanical interaction; and H_{GW}^{int} is the interaction between the GW field and the probe, given by

$$H_{GW}^{int} = -\frac{\omega_0}{4}(\hat{\alpha}_1^2 + \hat{\alpha}_2^2) \int d^3\mathbf{k} J_\lambda(\mathbf{k})\hat{h}_\lambda(t, \mathbf{k}), \quad (16)$$

where we have used J_λ to represent J_{xx}^λ. We point out that only the xx component of the GW field interacts with our probe, where x is the direction of propagation of the dynamical photons in the optical mode.

We then quantize canonically by imposing the commutation relations

$$[\hat{q}, \hat{p}] = i\hbar, \qquad [\hat{\alpha}_1, \hat{\alpha}_2] = i\hbar, \quad (17a)$$

$$[\hat{h}_\lambda(\mathbf{k}), \hat{\Pi}_{\lambda'}^\dagger(\mathbf{k}')] = [\hat{h}_\lambda^\dagger(\mathbf{k}), \hat{\Pi}_{\lambda'}(\mathbf{k}')] = i\hbar\delta_{\lambda\lambda'}\delta^2(\mathbf{k} - \mathbf{k}'). \quad (17b)$$

In order to perform measurements on the probe state while maintaining a constant amplitude inside the cavity, we must relax the perfect reflectivity condition to couple the probe to an external pump field whose ingoing photons drive the optical mode and whose outgoing photons are measured by the photodetector. The outgoing photons may also be thought of as ancillae, which ensures that the probe state evolves unitarily during continuous measurement without measurement-based feedback. We enlarge our Hilbert space to include the external pump by adding $H_{ext} = i\sqrt{2\gamma}[\hat{a}^\dagger\hat{c}_{x=0} - \hat{a}\hat{c}_{x=0}^\dagger] - i\int_{-\infty}^{\infty} dx \hat{c}_x^\dagger\partial_x\hat{c}_x$ to account for the interaction between the pump and probe (going directly to the interaction picture with respect to the free evolution of the pump field). Here $\hat{a}$, $\hat{a}^\dagger$ are raising and lowering operators of the optical mode, defined such that $\hat{\alpha}_1 = \sqrt{\hbar/2}(\hat{a} + \hat{a}^\dagger)$ and $\hat{\alpha}_2 = -i\sqrt{\hbar/2}(\hat{a} - \hat{a}^\dagger)$. We shall refer to $\hat{\alpha}_{1,2}$ as the amplitude and phase quadratures.

Relaxing the condition of perfect reflectivity also allows LIGO to operate in the detuned configuration, where the pump field has the off-resonant frequency $\omega_L = \omega_0 + \Delta$. The fast evolution in the canonical transformation in Eq. (15a) is then given by ω_L instead of ω_0, and this results in an additional term in the Hamiltonian given by $H_\Delta = -\Delta(\hat{\alpha}_1^2 + \hat{\alpha}_2^2)/2$.

Assuming large average amplitude inside the cavity, we linearize the Hamiltonian by writing $\hat{\alpha}_1 \to \bar{\alpha} + \delta\hat{\alpha}_1$ and $\hat{\alpha}_2 \to \delta\hat{\alpha}_2$. Then, keeping only terms linear in small quantity δ in the interaction terms (which are already small), we write down the final form of our Hamiltonian:

$$H = H_q^{(0)} + H_{GW}^{(0)} + H_{ext} - \frac{\Delta}{2}(\hat{\alpha}_1^2 + \hat{\alpha}_2^2)$$
$$- \omega_0\bar{\alpha}\hat{\alpha}_1\left[\frac{\hat{q}}{L} + \frac{1}{2}\int d^3\mathbf{k} J_\lambda(\mathbf{k})\hat{h}_\lambda(t, \mathbf{k})\right], \quad (18)$$

where we take $\Delta \to 0$ to recover the tuned configuration.

For notational simplicity, we denote the integral over GW $\mathbf{k}$ modes by $\hat{\mathfrak{h}}(t)$, or

$$\hat{\mathfrak{h}}(t) \equiv \int d^3\mathbf{k} J_\lambda(\mathbf{k})\hat{h}_\lambda(t, \mathbf{k}), \quad (19)$$

and point out that in the long-wavelength approximation where $k_x L \ll 1$ we have $\hat{\mathfrak{h}}(t) = \hat{h}_{xx}^{TT}(t, \mathbf{x} = \mathbf{0})$ according to Eq. (2).

For a strong excitation of the GW field from an astrophysical event such as a BBH merger, we can separate the GW field into a large classical component along with quantum fluctuations and write $\hat{\mathfrak{h}}(t) \to h_s(t) + \hat{\mathfrak{h}}(t)$, and the interaction Hamiltonian becomes

$$H_{int} = -\frac{\omega_0\bar{\alpha}\hat{\alpha}_1}{2}[\mathfrak{h}_s(t) + \hat{\mathfrak{h}}(t)]. \quad (20)$$

We also point out that H_{ext} was not derived from a fundamental action and was added phenomenologically in accordance with the standard formulation for input-output theory in quantum optics [43]. The only concern here would be the interaction between the external pump and the GW field, but that interaction is negligible, since the power in the pump without the amplification effects of a Fabry-Perot cavity is orders of magnitude smaller than that of the optical mode.

E. Equations of motion

We identify the input and output pump field from the external field operators $\hat{c}_x$ [37],

$$\hat{a}_{in} = \hat{c}_{x=0^-}, \quad \hat{a}_{out} = \hat{c}_{x=0^+}, \quad \hat{c}_{x=0} = \frac{\hat{c}_{x=0^-} + \hat{c}_{x=0^+}}{2}, \quad (21)$$

and denote their corresponding amplitude and phase quadratures by $\hat{\alpha}_{in}^{(1,2)}$. From the Hamiltonian in Eq. (18) we derive the following Heisenberg EOM for the probe and the GW field:

$$\dot{\hat{q}} = p/m, \qquad \dot{\hat{p}} = \omega_0\bar{\alpha}\hat{\alpha}_1/L, \quad (22)$$

$$\dot{\hat{\alpha}}_1 = -\Delta\hat{\alpha}_2 - \gamma\hat{\alpha}_1 + \sqrt{2\gamma}\hat{\alpha}_1^{\text{in}}, \tag{23a}$$

$$\dot{\hat{\alpha}}_2 = \Delta\hat{\alpha}_1 - \gamma\hat{\alpha}_2 + \sqrt{2\gamma}\hat{\alpha}_2^{\text{in}}$$
$$- \frac{\omega_0\bar{\alpha}}{2}\left[\mathfrak{h}_s(t) + \int d^3\mathbf{k} J_\lambda(\mathbf{k})\hat{h}_\lambda(\mathbf{k})\right] - \frac{\omega_0\bar{\alpha}}{L}\hat{q}, \tag{23b}$$

$$\dot{\hat{h}}_\lambda(\mathbf{k}) = \frac{1}{M_G}\hat{\Pi}_\lambda(\mathbf{k}), \tag{24a}$$

$$\dot{\hat{\Pi}}_\lambda(\mathbf{k}) = -\omega_k^2 M_G \hat{h}_\lambda(\mathbf{k}) + \frac{\omega_0\bar{\alpha}}{2}J_\lambda^*(\mathbf{k})\hat{\alpha}_1, \tag{24b}$$

where $\omega_k = c|\mathbf{k}|$ and we have defined $M_G = c^2/32\pi G$.

F. Discussion of gauge choice

We point out that, as expected, in the TT gauge the metric perturbations do not affect the coordinate motion of a particle moving along a geodesic [28] to order (v^2/c^2), as is evident in Eq. (11). Additionally, if the particles are initially at rest (such that $v = 0$), and experience no other forces, then linear metric perturbations will not affect coordinate motion for even appreciable values of v/c.
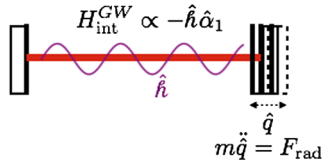
Newtonian Gauge
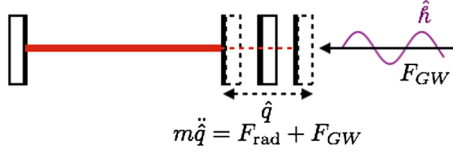
TT gauge

FIG. 2.   Representations of the GW interaction in the Newtonian versus TT gauges. In the Newtonian gauge, the gravitational wave exerts a tidal force F_{GW} so that the test-mass position is driven by both radiation pressure and gravitational-wave forces. In contrast, in the TT gauge, the GW interacts directly with the optical cavity mode, and the test-mass position is driven by the radiation pressure force alone. The two pictures are physically equivalent descriptions of the dynamics of a cavity whose mirrors fluctuate about their geodesic due to radiation pressure. In the presence of an incoming GW, the geodesics of the two mirrors deviate, and the time delay for a photon entering to be reflected back changes. In the Newtonian gauge, the change in time delay is reflected in the test-mass coordinate, while in the TT gauge this effect is directly accounted for by a phase shift in the cavity mode. The TT gauge viewpoint allows for a canonical description of the interaction.

This is because the gauge symmetry for gravitational fields is a diffeomorphism, or a local symmetry, and so the coordinate length between two free-falling particles is gauge dependent. However, all physical quantities must be gauge invariant, and in particular, the proper time elapsed from when a photon enters the cavity to when it is reflected back (i.e., with respect to an observer sitting on the input mirror) is the same in either gauge, despite the difference between the two gauges in terms of the coordinate distance that the photon travels. It should be noted that for our system, the test mass is not actually a free-falling particle, since it interacts optomechanically with the optical mode. This means that its position coordinate may be affected by metric perturbations even in the TT gauge, and in fact is, through the coupling of L_{OM} to h_{ij}. However, this term is of $O(hq/L)$, which is second order in small quantities and therefore ignored in Eq. (12).

As shown in Fig. 2, the consequence of the TT gauge choice is that in this picture, the GWs interact directly with the optical mode, in contrast to the point of view that GWs exert a tidal force onto the test mass, whose motion then causes a phase shift, which has so far been taken in treatments in which the test masses were considered quantum [26]. The TT gauge view allows for straightforward derivation of the qCRB for LIGO, or the fundamental limit to measurement sensitivity. We will discuss the details of this and its implications in an accompanying paper.

IV. GENERATION OF QUANTUM GRAVITATIONAL WAVES

Since our formalism treats the GW field and the probe on equal footing, the interaction between them is bidirectional, meaning that in addition to the Hamiltonian in Eq. (18) describing how the probe evolves under GW interaction, it also governs how the probe affects the dynamics of the GW field. Specifically, we recover the quantum analogue of the classical quadrupole moment formula as derived from Einstein's field equations [28]:

$$\hat{h}_{ij}^{\text{TT}}(t, \mathbf{x}) = \frac{2G}{c^4}\frac{\ddot{\hat{I}}_{ij}^{\text{TT}}(t - |\mathbf{x}|/c)}{|\mathbf{x}|}, \tag{25}$$

where $\hat{\mathbf{I}}^{\text{TT}}$ is the TT projection of the mass quadrupole moment. The result is unsurprising, but it serves as a demonstration of the equivalence in the relations between perturbative quantum gravity interacting with quantum matter and those of the classical scenario, specifically that quantum stress energy radiates GW in the same way as classical matter, and moreover that the radiation is quantum. This is in contrast to semiclassical gravity, which postulates that spacetime is classical while matter fields are quantum, such that the expectation value must be taken over $\ddot{\hat{\mathbf{I}}}$. We emphasize that the semiclassical and fully quantum viewpoints are in nature different and indeed

have testable physical implications, at least in principle. First, Eq. (25) implies that nonclassical states of quantum matter will result in nonclassical states of GWs, and second, the quantum generation of GWs will result in quantum coherent backaction effects on matter, which has not been previously considered and will be discussed further in Sec. V.

To obtain Eq. (25), we solve the set of coupled differential equations in Eqs. (24a) and (24b) in the frequency domain. Defining the Fourier pair using the conventions

$$\hat{O}(\Omega) = \int_{-\infty}^{\infty} dt \hat{O}(t) e^{-i\Omega t}, \qquad \hat{O}(t) = \int_{-\infty}^{\infty} \frac{d\Omega}{2\pi} \hat{O} e^{-i\Omega t},$$

(26)

we have

$$\tau_{ij}(\mathbf{k})\hat{h}_\lambda(\Omega, \mathbf{k}) = \frac{\omega_0 \bar{\alpha}}{2M_G} \frac{\tau_{ij}(\mathbf{k}) J_\lambda(\mathbf{k})}{(\omega_k^2 - \Omega^2)} \hat{\alpha}_1(\Omega). \quad (27)$$

Since the probe is a Newtonian source, we can apply the slow-motion condition and make the simplifying approximation that $|\mathbf{k}|L \ll 1$, so that $J_\lambda(\mathbf{k}) \to \tau_{xx}(\mathbf{k})/\sqrt{(2\pi)^3}$. Remembering that we are already in the TT gauge, the inverse spatial Fourier transform of Eq. (27) will give us the gauge-fixed GW field in configuration space, or $\hat{h}_{ij}^{\mathrm{TT}}(\Omega, \mathbf{x})$. Then, from our equations of motion we obtain

$$\hat{h}_{ij}^{\mathrm{TT}}(\Omega, \mathbf{x}) = \frac{1}{2M_G} \int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\omega_0 \bar{\alpha} \hat{\alpha}_1(\Omega) \tau_{ij}^\lambda(\mathbf{k}) \tau_{xx}^\lambda(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}}{\omega^2 - \Omega^2},$$

(28)

which can also be written as

$$\hat{h}_{ij}^{\mathrm{TT}}(\Omega, \mathbf{x}) = \frac{1}{2M_G} \mathbb{L}^{\mathrm{TT}}(\mathbf{x}) \left[\int \frac{d^3\mathbf{k}}{(2\pi)^3} \frac{\hat{T}_{ij}(\Omega, \mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}}{\omega_k^2 - \Omega^2} \right], \quad (29)$$

where $\hat{T}_{ij}(\Omega, \mathbf{k})$ is the time and spatial Fourier transform of the stress energy tensor of the probe, and, neglecting terms of $O(q/L)$ and $O(v^2/c^2)$, is simply the stress energy tensor of the optical mode $\hat{T}_{ij}^{\mathrm{EM}}$. It is given by

$$\hat{T}_{ij}(\Omega, \mathbf{k}) = \hat{T}_{ij}^{\mathrm{EM}}(\Omega, \mathbf{k}) = \delta_{ix}\delta_{jx} \frac{\omega_0 \bar{\alpha}}{c^2} \hat{\alpha}_1(\Omega). \quad (30)$$

The notation $\mathbb{L}^{\mathrm{TT}}(\mathbf{x})$ is shorthand for the TT projection operator, which projects each Fourier component $\hat{T}_{ij}(\Omega, \mathbf{k})$ to its TT components with respect to propagation vector $\mathbf{k}$, and accounts for the presence of the polarization tensors in Eq. (28). Explicitly, this operation on some general tensor field $f_{ij}(\mathbf{x})$ is given by

$$\mathbb{L}^{\mathrm{TT}}(\mathbf{x})[f_{ij}(\mathbf{x})]$$

$$= \int d^3\mathbf{x}' \int d^3\mathbf{x}'' f_{lm}(\mathbf{x}'') \left[\mathcal{P}_{il}(\mathbf{x}, \mathbf{x}') \mathcal{P}_{jm}(\mathbf{x}', \mathbf{x}'') \right.$$

$$\left. - \frac{1}{2} \mathcal{P}_{ij}(\mathbf{x}, \mathbf{x}') \mathcal{P}_{mn}(\mathbf{x}', \mathbf{x}'') \delta_l^n \right], \quad (31)$$

where $\mathcal{P}_{ij}(\mathbf{x}, \mathbf{x}')$ is the transverse projection operator for vector plane waves, such that $\int d^3\mathbf{x}' \mathcal{P}_{ij}(\mathbf{x}, \mathbf{x}') v_j e^{i\mathbf{k}\cdot\mathbf{x}'} = v_i^\perp e^{i\mathbf{k}\cdot\mathbf{x}}$, and is equal to

$$\mathcal{P}_{ij}(\mathbf{x}, \mathbf{x}') = \delta_{ij}\delta^3(\mathbf{x} - \mathbf{x}') - \partial_i G(\mathbf{x}, \mathbf{x}')\partial_{j'}, \quad (32)$$

where $G(\mathbf{x}, \mathbf{x}') = -1/(4\pi|\mathbf{x} - \mathbf{x}'|)$ is the Green function of the $\mathbf{\nabla}^2$ operator. The equivalence of Eq. (29) to Eq. (28) then follows readily from applying Eqs. (30) and (31).

From Eq. (29), we may obtain the result from Einstein's field equations by identifying

$$\hat{T}_{ij}(\Omega, \mathbf{k}) \approx \int d^3\mathbf{x} \hat{T}_{ij}(\Omega, \mathbf{x}), \quad (33)$$

which holds under the slow-motion approximation and which states that the Fourier $\mathbf{k}$ mode of the stress-energy tensor is approximately equal to its own volume integral. Then, performing the inverse transform on Eq. (29) into the time domain in the far zone, we obtain from the equations of motion in Eq. (24) our final expression for the GW field:

$$\hat{h}_{ij}^{\mathrm{TT}}(t, \mathbf{x}) = \frac{4G}{c^2} \mathbb{L}^{\mathrm{TT}}(\mathbf{x}) \left[\frac{\int d^3\mathbf{x}' \hat{T}_{ij}(t - |\mathbf{x}|/c, \mathbf{x}')}{|\mathbf{x}|} \right]. \quad (34)$$

To see that this is equivalent to the quadrupole formula in Eq. (25), we point out that in the far zone approximation, the argument of the TT projection operator $\mathbb{L}^{\mathrm{TT}}$ depends only on the distance $r = |\mathbf{x}|$, and therefore Eq. (31) reduces to a simple form, where it can be written in terms of the transverse projection operator for radially traveling waves:

$$\mathbb{L}^{\mathrm{TT}}[f_{ij}] = P_{il}f_{lm}P_{mj} - \frac{1}{2}P_{ij}\mathrm{Tr}\{\mathbf{Pf}\}, \quad (35)$$

where $P_{lm} = \delta_{lm} - x_l x_m/r^2$. In this way, it follows that Eqs. (34) and (25) are equivalent (with a final cosmetic step invoking stress-energy conservation [28]), and therefore that our equations of motion are consistent with classical general relativity.

In summary, we have reproduced the quadrupole formula for gravitational-wave generation using a fully quantum formalism. This serves as a check to our theoretical framework, as well as a theoretical basis for discussing quantum GW states generated by quantum matter.

V. PROBE DYNAMICS

Previously, we pointed out that in the presence of a large excitation, the GW field can be decomposed into classical and quantum components, whereby the interaction Hamiltonian takes the form of Eq. (20). Correspondingly, the GW effects that appear in $\hat{\alpha}_2$ consist of a predetermined classical component $\mathfrak{h}_s(t)$ as well as a quantum component involving $\hat{h}_\lambda(\mathbf{k})$. In previous treatments, one considered only $\mathfrak{h}_s(t)$, which acts on the probe as a fixed external force [26]. However, as discussed in Sec. IV and as is evident in Eq. (28), $\hat{h}_\lambda(\mathbf{k})$ has contributions from the probe itself, which is responsible for quantum coherent GW backaction. In this section, we obtain the equations of motion incorporating this quantum contribution and discuss their consequences.

A. Solving equations of motion in the Laplace domain

Let us solve for $\hat{\alpha}_1, \hat{\alpha}_2$'s equations of motion in Eqs. (23a) and (23b) in the Laplace domain, defining $t = 0$ to be the time of the probe's initial state preparation and using the transform pair

$$\hat{O}(s) = \int_0^\infty dt \hat{O}(t) e^{-st}, \quad \hat{O}(t) = \frac{1}{2\pi i} \int_{-i\infty+\epsilon}^{i\infty+\epsilon} ds \hat{O}(s) e^{st}. \tag{36}$$

Then, each $\mathbf{k}$ mode of the GW field is given by

$$\hat{h}_\lambda(s, \mathbf{k}) = \frac{s}{s^2 + \omega_k^2} \hat{h}_\lambda(0, \mathbf{k}) + \frac{1}{M_G} \frac{\hat{\Pi}_\lambda(0, \mathbf{k})}{s^2 + \omega_k^2} + \frac{\omega_0 \bar{\alpha}}{2} \frac{1}{M_G} J_\lambda^*(\mathbf{k}) \frac{\hat{\alpha}_1(s)}{s^2 + \omega_k^2}. \tag{37}$$

The difference between Eq. (37) and our previous result in Eq. (28) is that the Laplace transform solution accounts for the initial values of the field, whereas the Fourier transform solution does not. These initial values, given by the first two terms on the right-hand side of Eq. (37), represent the input quantum GW fluctuations which we collectively denote by $\hat{\mathfrak{h}}_{in}$, while the last term is the part of the GW field generated by the probe that leads to backaction.

Using Eq. (37) and also substituting $\hat{q}(s) = \omega_0 \bar{\alpha} \hat{\alpha}_1 / mLs^2$ into Eq. (23b) (ignoring the initial state of the test mass, which is irrelevant for the interaction dynamics), we solve Eqs. (23a) and (23b) for the optical mode in terms of the input field quadratures $\hat{\alpha}_{1,2}^{in}, \hat{\alpha}_{1,2}^{out}$, which are associated with the mode operators defined in Eq. (21), and obtain

$$\begin{bmatrix} s + \gamma & \Delta \\ -(1 + \xi_{BA})\Delta & s + \gamma \end{bmatrix} \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix}$$
$$= \sqrt{2\gamma} \begin{pmatrix} \hat{\alpha}_1^{in} \\ \hat{\alpha}_2^{in} \end{pmatrix} + \frac{\omega_0 \bar{\alpha}}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix} [\mathfrak{h}_s + \hat{\mathfrak{h}}_{in}], \tag{38}$$

where ξ_{BA} is the modification to the optical mode response due to backaction from the test mass and GW field and is given by

$$\xi_{BA}(s) = \frac{1}{\Delta} \left(\frac{\epsilon_q}{s^2} + \epsilon_{GW} s \right) \tag{39}$$

with

$$\epsilon_q = \frac{(\omega_0 \bar{\alpha})^2}{mL^2}, \tag{40a}$$

$$\epsilon_{GW} = \frac{8G}{15c^5} (\omega_0 \bar{\alpha})^2, \tag{40b}$$

and where we have included the effect of the predetermined classical component of the field $\mathfrak{h}_s$ in our solution. Here ϵ_q arises due to interaction between light and the test mass, while ϵ_{GW} arises due to interaction between light and the gravitational field.

We remark that for GW backaction we have only considered effects due to outgoing gravitational waves. The leading-order Feynman diagrams for backaction processes involve a graviton propagator between in and out matter states, which in principle should include contributions from longitudinal and timelike gravitons if we were to account for all gravitational effects. Restricting our attention to the TT modes ignores time-symmetric self-gravity effects such as the Newtonian self-potential, but those are well separated from the leading-order time-asymmetric term that is the GW backaction.

B. Tuned configuration

At times long enough such that the initial state of the probe is forgotten, for the tuned configuration with $\Delta = 0$ we obtain the cavity mode solutions

$$\hat{\alpha}_1 = \frac{\sqrt{2\gamma}}{s + \gamma} \hat{\alpha}_1^{in}, \tag{41a}$$

$$\hat{\alpha}_2 = \frac{\sqrt{2\gamma}}{s + \gamma} \hat{\alpha}_2^{in} + \xi_{BA} \frac{\sqrt{2\gamma}}{(s + \gamma)^2} \hat{\alpha}_1^{in} + \frac{\omega_0 \bar{\alpha}}{2} (\mathfrak{h}_s + \hat{\mathfrak{h}}_{in}). \tag{41b}$$

Inspecting the above equations, we see that the backaction term does not modify the dynamical response of either quadrature. Instead, it introduces additional fluctuations to the phase quadrature $\hat{\alpha}_2$, which modifies the shape of the output noise ellipse. This is shown explicitly in the noise input-output relations given by

$$\hat{\alpha}_j^{\text{out}} = \hat{\alpha}_j^{\text{in}} - \sqrt{2\gamma}\hat{\alpha}_j \qquad (42)$$

for $j = 1, 2$, and the input-output relations themselves can be derived from the EOM for the external pump field using H_{ext}. Substituting the solutions for $\hat{\alpha}_j$ given in Eqs. (41a) and (41b), we find the fluctuating parts of the outgoing quadratures

$$\hat{\alpha}_1^{\text{out}} = e^{2i\beta}\hat{\alpha}_1^{\text{in}}, \qquad (43a)$$

$$\hat{\alpha}_2^{\text{out}} = e^{2i\beta}[\hat{\alpha}_2^{\text{in}} - (\mathcal{K}_{pd} + i\mathcal{K}_{GW})\hat{\alpha}_1^{\text{in}}] - \sqrt{\frac{\gamma}{2}}\frac{\omega_0\bar{\alpha}}{(s+\gamma)}\hat{\mathfrak{h}}_{\text{in}}, \qquad (43b)$$

where β is an uninteresting overall phase factor, and $\mathcal{K}_{pd}$ and $\mathcal{K}_{GW}$ are the backaction terms due to the test mass and GWs, respectively. Making the identification $s \to -i\Omega$, we have $\beta = \arctan(\Omega/\gamma) + \pi/2$ and

$$\mathcal{K}_{pd}(\Omega) = \frac{1}{m}\left(\frac{\omega_0\bar{\alpha}}{L}\right)^2 \frac{2\gamma}{\Omega^2(\gamma^2+\Omega^2)}, \qquad (44a)$$

$$\mathcal{K}_{GW}(\Omega) = \frac{2\Omega\gamma\epsilon_{GW}}{\gamma^2+\Omega^2}. \qquad (44b)$$

On the right-hand side of Eq. (43b), the term that contains $\mathcal{K}_{pd}$ gives rise to the well-known ponderomotive effect, which causes a rotation and squeezing of the input noise ellipse, with rotation angle $\theta = \arctan(\mathcal{K}_{pd}/2)$, squeeze angle $\phi = \text{arccot}(\mathcal{K}_{pd}/2)/2$, and squeeze factor $r = \text{arcsinh}(\mathcal{K}_{pd}/2)$ [26]. On the other hand, the term containing $\mathcal{K}_{GW}$, which gives rise to GW backaction, is imaginary, and therefore must be associated with additional fluctuations in $\hat{\alpha}_2^{\text{out}}$ in order for $[\hat{\alpha}_2^{\text{out}}(t), \hat{\alpha}_2^{\text{out}}(t')] = 0$. Since $\hat{\alpha}_{1,2}^{\text{out}}(t)$ are outgoing fields at different moments of time, and therefore are independent d.o.f., this commutation relation must be satisfied.

Indeed, we see in Eq. (43b) that there are fluctuations from the GW field which ensure that the output phase quadrature commutes at different times, and additionally enlarges the total area of the output noise ellipse. Unfortunately, these additional fluctuations cannot be removed without access to quantum GW d.o.f. and will therefore introduce additional noise, albeit at $O(\epsilon_{GW}^2)$.

C. Detuned configuration

When the optical drive is detuned from cavity resonance, the amplitude and phase quadratures of the cavity mode rotate into each other. In the presence of backaction, this results in modifications to their dynamical response functions, in contrast to the tuned configuration. For simplicity, we ignore the ponderomotive backaction due to the test mass by taking $m \to \infty$ and focus solely on gravitational effects. Then, solving Eq. (38) for $\Delta \neq 0$, we find

$$\hat{\alpha}_1 = \sqrt{2\gamma}[\chi_1\hat{\alpha}_1^{\text{in}} - \chi_2\hat{\alpha}_2^{\text{in}}] - \frac{\omega_0\bar{\alpha}}{2}\chi_2(\mathfrak{h}_s + \hat{\mathfrak{h}}_{\text{in}}), \qquad (45a)$$

$$\hat{\alpha}_2 = \sqrt{2\gamma}\left[\left(1 + \frac{\epsilon_{GW}s}{\Delta}\right)\chi_2\hat{\alpha}_1^{\text{in}} + \chi_1\hat{\alpha}_2^{\text{in}}\right] + \frac{\omega_0\bar{\alpha}}{2}\chi_1(\mathfrak{h}_s + \hat{\mathfrak{h}}_{\text{in}}) \qquad (45b)$$

for the response functions

$$\chi_1 = \frac{s+\gamma}{(s+\gamma)^2 + \Delta^2 + \epsilon_{GW}\Delta s}, \qquad (46a)$$

$$\chi_s = \frac{\Delta}{(s+\gamma)^2 + \Delta^2 + \epsilon_{GW}\Delta s}. \qquad (46b)$$

The optical mode's response to external drive is modified through ϵ_{GW}, which can be encapsulated by shifts in the effective damping and detuning of the optical mode:

$$\tilde{\gamma} = \gamma + \frac{\epsilon_{GW}\Delta}{2}, \qquad \tilde{\Delta} = \Delta - \frac{\epsilon_{GW}\gamma}{2}. \qquad (47)$$

Since Δ can take positive or negative values, the backaction can either augment or reduce the effective cavity damping rate γ. This is due to the amplitude quadrature being coherently added to itself as the amplitude and phase quadratures rotate into each other through the $\Delta\hat{\alpha}_2$ term in its own EOM. Depending the sign of Δ, it can beat either destructively ($\Delta > 0$ or red-detuned) or constructively ($\Delta < 0$ or blue-detuned). The reddetuned case is shown schematically in Fig 3.

D. Backaction in the Newtonian gauge

To understand the GW backaction more intuitively, we now go to the Newtonian gauge and show that it is in fact the quantum analogue of the radiation reaction potential. The radiation reaction potential, or Φ_{react}, is the leading-order time-asymmetric GR correction to the Newtonian potential and derives from the outgoing GWs radiated by the probe's time-dependent mass quadrupole moment. The correction accounts for the consequent loss of energy which leads to damping of the motion, an effect known as radiative damping [28]. To see this explicitly, we write the EOM of the test mass in the Newtonian gauge under radiation pressure and the radiation reaction force:

$$\ddot{\hat{q}} = -\frac{\partial}{\partial x}\Phi_{\text{react}} + \frac{\omega_0\bar{\alpha}}{mL}\hat{\alpha}_1, \qquad \Phi_{\text{react}}(\mathbf{x}) = \frac{G}{5c^5}\mathcal{I}_{jk}^{(5)}x^jx^k, \qquad (48)$$

where $\mathcal{I}_{jk}$ is the reduced-mass quadrupole moment tensor, and the superscript represents the number of time derivatives. We find that the reaction force on the system evaluates to
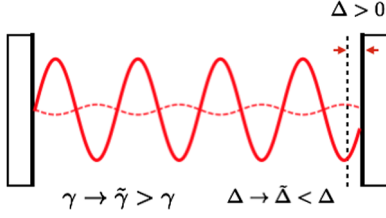
FIG. 3.   Illustration of quantum coherent backaction effects on the cavity mode due to GW interaction in the presence of detuning. The GWs generated by $\hat{\alpha}_1$ acts back on $\hat{\alpha}_2$ in such a way that causes the field to beat coherently with itself. The above shows the case for red-detuning where $\Delta > 0$. The solid red line represents the cavity mode in the absence of backaction, while the dotted red line represents the contribution due to GW backaction. This effect is quantified by changes to the cavity's effective damping and detuning rate so that $\tilde{\gamma} = \gamma + \epsilon_{GW}\Delta/2$ and $\tilde{\Delta} = \Delta - \epsilon_{GW}\gamma/2$.

$$F_{\text{react}} = -\frac{8G}{15c^5}mL^2\hat{q}^{(5)}. \qquad (49)$$

Since the reaction force is dependent on $\hat{q}$ itself, it serves to modify the response of $\hat{q}$ to radiation pressure force. Here we are interested only in the backaction effects of GW interaction and have therefore suppressed the input GW field. Assuming that the probe is operating under a steady state, we may solve the probe EOM in the Fourier domain,

$$\hat{q}(\Omega) = [\chi_q^{(0)} + \delta\chi_q]\frac{\omega_0\bar{\alpha}}{L}\hat{\alpha}_1(\Omega), \qquad (50)$$

for free mass susceptibility $\chi_q^{(0)}$ and its perturbation $\delta\chi_q$:

$$\chi_q^{(0)} = -\frac{1}{m\Omega^2}, \qquad \delta\chi_q = -i\frac{8G}{15c^5}mL^2\Omega^3\chi_q^{(0)}. \qquad (51)$$

In the Newtonian gauge, the optical mode does not interact gravitationally, and the Heisenberg EOM for its quadratures depends only on the input optical field and the test-mass dynamics, which in the tuned configuration is given by $\hat{\alpha}_1(\Omega)(\gamma - i\Omega) = \sqrt{2\gamma}\hat{\alpha}_1^{\text{in}}(\Omega)$ and $\hat{\alpha}_2(\Omega)(\gamma - i\Omega) = (\omega_0\bar{\alpha}/L)\hat{q}(\Omega) + \sqrt{2\gamma}\hat{\alpha}_2^{\text{in}}(\Omega)$. Substituting Eq. (50) into the expression for $\hat{\alpha}_2$, we find

$$\hat{\alpha}_1 = \frac{\sqrt{2\gamma}}{\gamma - i\Omega}\hat{\alpha}_1^{\text{in}}, \qquad (52a)$$

$$\hat{\alpha}_2 = \frac{\omega_0^2\bar{\alpha}^2}{L^2}(\chi_q^{(0)} + \delta\chi_q)\frac{\hat{\alpha}_1}{\gamma - i\Omega} + \frac{\sqrt{2\gamma}}{\gamma - i\Omega}\hat{\alpha}_{\text{in}}^{(2)}, \qquad (52b)$$

where $\chi_q^{(0)} = -(m\Omega^2)^{-1}$ is the free test-mass response and $\delta\chi_q$ is the GW correction. Making the substitution $-i\Omega \to s$, we find that the modification to the test-mass response deriving from Φ_{react} exactly corresponds to the GW backaction term in Eq. (38) for $\Delta \to 0$:

$$\left(\frac{\omega_0\bar{\alpha}}{L}\right)^2\frac{\delta\chi_q}{\gamma - i\Omega} \to \frac{\epsilon_{GW}s}{s + \gamma}. \qquad (53)$$

From here, it follows that identical input-output relations to Eqs. (43a) and (43b) may be obtained. Here we note that the fifth-order time derivative in Eq. (49) makes the radiation reaction force time asymmetric, which results in $\mathcal{K}_{GW}$ being imaginary.

It is straightforward to see that an analysis for the detuned configuration would yield a similar result, since the only difference in the EOMs for $\hat{\alpha}_1$ and $\hat{\alpha}_2$ between the tuned and detuned cases in the Newtonian gauge would be the addition of the terms $-\Delta\hat{\alpha}_2$ to the rhs of Eq. (52a) and $\Delta\hat{\alpha}_1$ to that of Eq. (52b). This yields identical expressions to Eq. (38).

We have confirmed that in both gauges, the same backaction effect appears in the output field which is being measured, as must be the case. However, there are some interesting points that arise from comparing the different interpretations offered by each gauge. First, in the Newtonian gauge, the underlying physical mechanism for the backaction is the modification of test-mass response to radiation pressure due to the radiative damping. Thus, the GW backaction can be interpreted as a correction to the response function of the test mass to external forces (i.e., radiation pressure), given by $\delta\chi_q$ in Eq. (51). Correspondingly, $\mathcal{K}_{GW}$ can be interpreted as a correction to the ponderomotive backaction $\mathcal{K}_{pd}$ in Eq. (44). Viewed in this way, one might intuitively expect that in the limit $m \to \infty$ the infinitely massive mirror would be unperturbed by radiation pressure, and therefore both $\mathcal{K}_{pd}$ and its correction $\mathcal{K}_{GW}$ should go to zero. However, as demonstrated in the TT gauge, the GW backaction appears without consideration of test-mass dynamics, and without regard to its mass. A more careful look at the Newtonian gauge reveals that this result is consistent, due to the fact that F_{react} is mass dependent, and therefore, under weak equivalence, the inertial mass in $\chi_q^{(0)}$ which appears in the denominator of $\delta\chi_q$ cancels the gravitational mass in F_{react} which appears in its numerator. The gravitational radiation reaction therefore ensures that even an infinitely massive object will respond to external forces.

The second point of contrast is that in the Newtonian gauge, the motion of the test mass is indeed damped due to energy loss though GW radiation. However, in the TT gauge, neither the test mass nor the optical mode experiences damping. Again, this seemingly paradoxical observation may be resolved when one recognizes that the energy that drives the test-mass motion is from the optical pump, which also provides the energy radiated away in GWs. This is clear from Eqs. (34) and (38), where the generation of GWs is shown to depend on $T_{ij} \sim \hat{\alpha}_1$, which depends only on the input optical field, as shown in Eq. (52a). That the energy lost to GW radiation is sourced from the external optical pump also addresses the

counterintuitive result that GW backaction can reduce the effective cavity damping rate, since the damping of the cavity mode does not directly correspond to energy exchange with the GW field.

E. Role of quantum GW fluctuations in backaction

Although radiation reaction can be traced back to the classical radiation reaction potential, in order to consistently include its effects on quantum matter, one must necessarily quantize the GW fluctuations. This is clear from Eqs. (45a) and (45b), where including the ϵ_{GW} backaction effect without also including the $\hat{\mathfrak{h}}_{\mathrm{in}}$ fluctuations will result in a modified commutator between the cavity mode quadratures

$$[\hat{\alpha}_1(t), \hat{\alpha}_2(t)] = i\hbar \left[1 - \frac{\epsilon_{\mathrm{GW}}\Delta}{2\gamma}(1 - e^{2\tilde{\gamma}t})\right], \quad (54)$$

and only by also including $\hat{\mathfrak{h}}_{\mathrm{in}}$ will we recover the canonical commutation relation $[\hat{\alpha}_1(t), \hat{\alpha}_2(t)] = i\hbar$. This result furthers our understanding of how gravity interacts with quantum matter and advances the point of view that gravity must ultimately be quantum, since the predictions of semiclassical gravity upon including the radiation damping correction will result in Eq. (54), in violation of the uncertainty principle.

VI. CONCLUSION

To summarize, in this work we developed a framework for the Hamiltonian formulation of the interaction between LIGO or a similar laser interferometer with gravity, such that both the GW field and the matter probe are dynamical d.o.f. of the total system. To do so, we needed to address the issues of (i) gauge fixing for the gravitational field and (ii) formulating the optomechanical interaction using the action principle when the equations of motion are unknown. With regard to the former, we have argued that keeping only leading-order interactions between the detector and gravitational waves allows us to fix the action in the TT gauge. For the latter, we have shown how to derive the optomechanical interaction from the action of an electromagnetic field with a dynamical boundary condition. Our completed formulation allows for the leading-order quantum treatment of the interactions between LIGO and gravitational waves, which, in the limit of classical gravity, recovers known equations of motion for LIGO's output fields as well as the quadrupole moment formula for GW generation. This serves as a verification for the framework, which, allowing for quantum gravity, additionally predicts quantum coherent backaction effects that correspond to the classical radiation reaction potential in general relativity. We have shown that in order for this well-studied potential to be consistently applied to quantum matter such that

commutation relations are preserved, the gravitational field itself must be quantum. This suggests that testing the effects of gravity at post-Newtonian orders may offer some insight into the quantum versus classical nature of gravity.

Even in the classical limit for gravity, our formulation presents an alternative interpretation of LIGO's detections. In the Newtonian gauge used by previous analyses, the interaction occurs between the test mass and the signal which subsequently modulates the cavity mode; in the TT gauge, the interaction occurs directly with the cavity mode. While the signal appears in the same way in the output fields that we measure, the latter point of view provides a rigorous foundation for calculating the fundamental limit to LIGO's detection sensitivity in the form of its qCRB. The qCRB follows readily [10,13] from the Hamiltonian in Eq. (20) and can be expressed in the form of a minimum bound on the noise spectral density for the estimation error on $\mathfrak{h}_s(t)$, or

$$S_{\Delta\mathfrak{h}_s}(\Omega) \geq \frac{\hbar^2}{(\omega_0\bar{\alpha})^2 S_{\alpha_1}(\Omega)}, \quad (55)$$

where $\Delta\mathfrak{h}_s(t)$ is the residual between the signal and our estimate, and the noise spectral density for any stationary process $x(t)$ is given by

$$S_x(\Omega) = \int_{-\infty}^{\infty} d\tau e^{i\Omega\tau} \langle x(t+\tau)x(t)\rangle. \quad (56)$$

We also point out that we have not considered the decohering effects of GW fluctuations, which again follow from the formalism. In fact, it can be shown that the qCRB, decoherence, and radiation are fundamentally and quantifiably related, and we therefore delay a more detailed discussion of these processes and their relations to an accompanying paper. Finally, we remark that our framework offers a theoretical foundation consistent with current observations that can be extended to study alternative theories of gravity in LIGO in the form of modifications to the Einstein-Hilbert action. It provides a basic model for investigating the quantum versus classical nature of GR or modified gravity and how their features manifest themselves in quantum measurement.

ACKNOWLEDGMENTS

[1] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **116,** 061102 (2016)

[2] F. Acernese *et al.* (Virgo .Collaboration), Classical Quantum Gravity **35,** 205004 (2018).

[3] Y. Aso, Y. Michimura, K. Somiya, M. Ando, O. Miyakawa, T. Sekiguchi, D. Tatsumi, and H. Yamamoto (KAGRA Collaboration), Phys. Rev. D **88,** 043007 (2013).

[4] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **119,** 161101 (2017).

[5] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **120,** 091101 (2018).

[6] N. Yunes, K. Yagi, and F. Pretorius, Phys. Rev. D **94,** 084002 (2016).

[7] G. Amelino-Camelia, Nature (London) **398,** 216 (1999).

[8] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), Phys. Rev. Lett. **118,** 221101 (2017).

[9] K. G. Arun and C. M. Will, Classical Quantum Gravity **26,** 155002 (2009).

[10] H. Miao, R. X. Adhikari, Y. Ma, B. Pang, and C. Yanbei, Phys. Rev. Lett. **119,** 050801 (2017).

[11] T. G. Downes, J. R. van Meter, E. Knill, G. J. Milburn, and C. M. Caves, Phys. Rev. D **96,** 105004 (2017).

[12] C. W. Helstrom, J. Stat. Phys. **1,** 231 (1969).

[13] M. Tsang, H. M. Wiseman, and C. M. Caves, Phys. Rev. Lett. **106,** 090401 (2011).

[14] D. Kafri, J. M. Taylor, and G. J. Milburn, New J. Phys. **16,** 065020 (2014).

[15] D. Kleckner, I. Pikovski, E. Jeffrey, L. Ament, E. Eliel, J. van den Brink, and D. Bouwmeester, Phys. Rev. Lett. **116,** 100601 (2016).

[16] H. Yang, H. Miao, D.-S. Lee, B. Helou, and Y. Chen, Phys. Rev. Lett. **110,** 170401 (2013).

[17] I. Pikovski, M. R. Vanner, M. Aspelmeyer, M. S. Kim, and C. Brukner, Nat. Phys. **8,** 393 (2012).

[18] A. Belenchia, D. M. T. Benincasa, S. Liberati, F. Marin, F. Marino, and A. Ortolan, Phys. Rev. Lett. **116,** 161303 (2016).

[19] C. Anastopoulos and B. L. Hu, Classical Quantum Gravity **30,** 165007 (2013).

[20] M. P. Blencowe, Phys. Rev. Lett. **111,** 021302 (2013).

[21] T. Oniga and C. H.-T. Wang, Phys. Rev. D **93,** 044027 (2016).

[22] T. Oniga and C. H.-T. Wang, J. Phys. Conf. Ser. **845,** 012020 (2017).

[23] T. Damour and G. Esposito-Farese, Classical Quantum Gravity **9,** 2093 (1992).

[24] C. M. Caves, Phys. Rev. D **23,** 1693 (1981).

[25] A. Pace, M. Collett, and D. Walls, Phys. Rev. A **47,** 3173 (1993).

[26] H. J. Kimble, Y. Levin, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Phys. Rev. D **65,** 022002 (2001).

[27] V. B. Braginsky, M. L. Gorodetsky, F. Y. Khalili, A. B. Matsko, K. S. Thorne, and S. P. Vyatchanin, Phys. Rev. D **67,** 082001 (2003).

[28] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation* (W. H. Freeman and Company, New York, 1973).

[29] M. Rakhmanov, Phys. Rev. D **71,** 084003 (2005).

[30] R. Weiss, Rev. Mod. Phys. **71,** S187 (1999).

[31] B. J. Meers, Phys. Rev. D **38,** 2317 (1988).

[32] J. Mizuno, K. A. Strain, P. Nelson, J. Chen, R. Schilling, A. Rüdiger, W. Winkler, and K. Danzmann, Phys. Lett. A **175,** 273 (1993).

[33] A. Buonanno and Y. Chen, Phys. Rev. D **64,** 042006 (2001).

[34] A. Buonanno and Y. Chen, Phys. Rev. D **65,** 042001 (2002).

[35] J. Aasi, B. Abbott, R. Abbott, T. Abbott, M. Abernathy, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari *et al.*, Classical Quantum Gravity **32,** 115012 (2015).

[36] A. Buonanno and Y. Chen, Phys. Rev. D **67,** 104025 (2003).

[37] Y. Chen, J. Phys. B **46,** 104001 (2013).

[38] C. Kiefer, *Quantum Gravity* (Oxford University Press, New York, 2012).

[39] P. A. M. Dirac, Can. J. Math. **2,** 129 (1950).

[40] S. N. Gupta, Proc. Phys. Soc. London Sect. A **65,** 161 (1952).

[41] C. K. Law, Phys. Rev. A **51,** 2537 (1995).

[42] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Photons and Atoms: Introduction to Quantum Electrodynamics* (John Wiley and Sons, New York, 1989).

[43] D. F. Walls and G. J. Milburn, *Quantum Optics* (Springer, New York, 2008).