Systematic calibration error requirements for gravitational-wave detectors via the Cramér–Rao bound

Evan D. Hall,[1, *] Craig Cahillane,[2] Kiwamu Izumi,[3] Rory J. E. Smith,[4] and Rana X Adhikari[2]

[1]*LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[2]*LIGO Laboratory, California Institute of Technology, Pasadena, CA 91125, USA*
[3]*LIGO Hanford Observatory, Richland, WA 99352, USA*
[4]*OzGrav, School of Physics & Astronomy, Monash University, Clayton 3800, Victoria, Australia*

Gravitational-wave (GW) laser interferometers such as Advanced LIGO [1] transduce spacetime strain into optical power fluctuation. Converting this optical power fluctuations back into an estimated spacetime strain requires a calibration process that accounts for both the interferometer's optomechanical response and the feedback control loop used to control the interferometer test masses. Systematic errors in the calibration parameters lead to systematic errors in the GW strain estimate, and hence to systematic errors in the astrophysical parameter estimates in a particular GW signal. In this work we examine this effect for a GW signal similar to GW150914, both for a low-power detector operation similar to the first and second Advanced LIGO observing runs and for a higher-power operation with detuned signal extraction. We set requirements on the accuracy of the calibration such that the astrophysical parameter estimation is limited by errors introduced by random detector noise, rather than calibration systematics. We also examine the impact of systematic calibration errors on the possible detection of a massive graviton.

I. INTRODUCTION

Making astrophysical inferences from gravitational-wave detections like GW150914 [2], GW151226 [3], and GW170104 [4] requires detector data that is calibrated into spacetime strain with sufficient precision and accuracy [5]. Statistical error in the strain calibration has the potential to weaken astrophysical inferences—for example, by increasing the statistical error on the estimated masses of a particular binary system, or weakening constraints on the graviton mass. Systematic error in the strain calibration, on the other hand, distorts the detector data and therefore has the potential to produce incorrect astrophysical inferences—for example, the graviton mass could be estimated spuriously to be inconsistent with zero. Constraining systematic calibration errors will become only more pressing with time, as new and improved gravitational-wave detectors see events with ever higher signal-to-noise ratio (SNR), and these events are in turn used to achieve more stringent parameter estimation,.

In certain situations, the effect of systematic calibration errors is straightforward. In the simple case that one wishes to determine the luminosity distance D of a source using an interferometer whose calibration comprises a simple proportionality constant k, then a calibration error $\Delta k/k$ corresponds directly to the error in the estimate of $\Delta D/D$. However, in the more realistic case that the interferometer calibration function and the GW signal each involve multiple parameters, the relationship between systematic calibration errors and systematic calibration errors is less straightforward.

The effect of calibration errors on GW detection and parameter estimation have focused on placing frequency-dependent constraints on calibration accuracy, without assumptions about the underlying calibration parameters. Lindblom [6] derived requirements on the magnitude and phase errors of the inter-ferometer calibration so as to avoid missed signal detections. Vitale et al. [7] examined the effect systematic calibration errors on GW parameter estimation by modeling calibration errors as smooth, random frequency-dependent fluctuations in the interferometer's calibration.

In this work, we first lay out the basic ingredients to Advanced LIGO calibration, including a quasi-zero-pole-gain representation of the optomechanical plant that is valid even in a detuned resonant-sideband-extraction configuration. We use a semianalytic approach to explicitly relate systematic errors in calibration parameters (e.g., gains, poles and zeros) to systematic errors in GW signal parameters (e.g., masses and distances). This approach is similar to the approach of Cutler and Vallisneri [8], who examined how astrophysical parameter estimates are affected by systematic errors in waveform models.

To set requirements on the systematic calibration errors, we compare the systematic calibration-induced errors on the GW signal parameters to the errors induced by the detector's random noise. Here we use a Fisher-matrix method to estimate the errors due to detector noise via the Cramér–Rao bound. Although this method is known to have limitations [9], its results generally coincide (to within a factor of 2) with Monte Carlo methods for the kind of GW signals expected in Advanced LIGO, so long as the signal-to-noise ratio (SNR) is above ~20 [10, 11].

II. INTERFEROMETER MODEL

A basic diagram of an Advanced LIGO interferometer and its differential arm length feedback system is shown in Fig. 1.
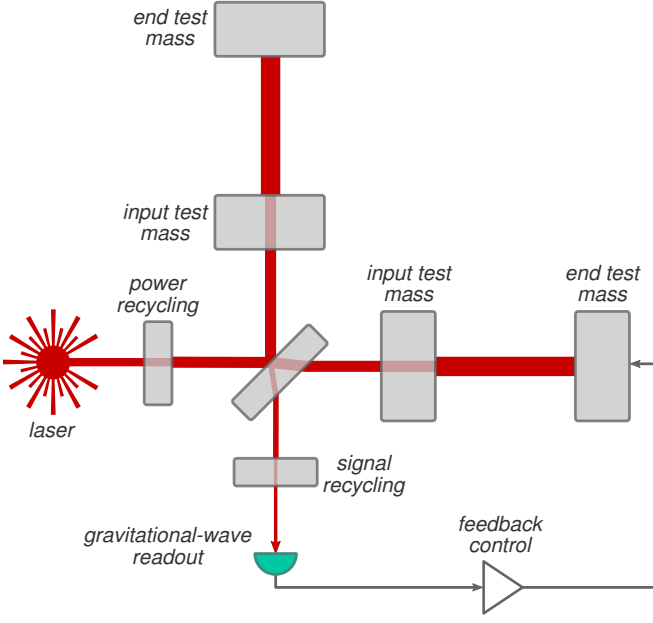
* evanhall@mit.edu

FIG. 1: Simplified diagram of Advanced LIGO, showing the optical layout, the GW readout at the antisymmetric port, and the differential arm-length feedback control system.

A. Optomechanical response

In this section we present the parameters which characterize the optical response of Advanced LIGO (Fig. 2). Each Advanced LIGO detector is a Michelson interferometer whose arms are Fabry–Pérot cavities; since the end test-masses are negligibly transmissive ($T_e \sim 5$ ppm), the bandwidth of each cavity is set by the input test mass transmissivity $T_i = 1.4\%$, yielding $f_a = cT_i/8\pi L = 42$ Hz. The power circulating in arms is enhanced by the inclusion of a power-recycling mirror (PRM) between the laser and the beamsplitter, and the PRM transmissivity T_p is chosen to maximize the circulating arm power.

With these six mirrors alone—four test masses, a beamsplitter, and a PRM—the interferometer's optical response in the GW band (between 10 Hz and 7 kHz) would be (to good approximation) a single-pole low-pass filter, with a gain g set by the arm power and the optical losses, and a single pole p equal to the arm bandwidth f_a. However, each Advanced LIGO detector additionally employs a scheme called resonant sideband extraction (RSE), in which a signal recycling mirror (SRM) is placed at the detector's antisymmetric port to alter the detector's optical response [12] and its quantum-limited noise performance [13]. The exact nature of the alteration depends on the SRM's power transmissivity T_s and the microscopic signal-recycling-cavity (SRC) length ℓ_s. Additionally, the interferometer optical response is affected by the homodyne angle ζ, which describes the audio-band demodulation quadrature of the GW readout.

In the end, the interferometer optical response is determined by the following five physical parameters:

1. the power $P = \varepsilon \sqrt{P_{bs} P_{lo}}$, which depends on the circulat-

ing power P_{bs} impinging on the beamsplitter, the local oscillator power P_{lo} used to detect the GW signal, and any optical losses ε;

2. the arm bandwidth f_a;

3. the power transmissivity T_s of the SRM;

4. the microscopic one-way SRC phase $\phi = (2\pi\ell_s/\lambda_0)$ mod 2π; and

5. the homodyne angle ζ.

However, the optical response is characterized more directly via a quasi-zero-pole-gain representation comprising the following parameters.

1. The optical gain g (with units of watts per meter), which depends on the beamsplitter power P_{bs}, the local oscillator power P_{lo}, and any optical losses in the system.

2. The homodyne zero

$$z = f_a \times \frac{\cos(\phi + \zeta) - r_s \cos(\phi - \zeta)}{\cos(\phi + \zeta) + r_s \cos(\phi - \zeta)}. \tag{1}$$

3–4. The pole

$$p = f_a \times \frac{1 - r_s e^{2i\phi}}{1 + r_s e^{2i\phi}}, \tag{2}$$

which (being complex) comprises two independent parameters. We will find it most convenient to work with the magnitude

$$|p| = f_a \times \left(\frac{1 - 2r_s \cos 2\phi + r_s^2}{1 + 2r_s \cos 2\phi + r_s^2} \right)^{1/2} \tag{3}$$

and the Q factor

$$Q_p = \frac{|p|}{2 \operatorname{Re} p} = \frac{1}{2} \times \frac{\left(1 - 2r_s^2 \cos 4\phi + r_s^4 \right)^{1/2}}{1 - r_s^2}, \tag{4}$$

which attains a minimum value of $1/2$ when p is real.

5. The squared spring frequency

$$\xi^2 = f_a^2 \times \frac{2r_s \sin 2\phi}{1 - 2r_s \cos 2\phi + r_s^2} \times \frac{cP_{bs}/\lambda_0}{2\pi^3 f_a^4 ML^2}, \tag{5}$$

where $\lambda_0 = 1064$ nm is the laser wavelength, and $M = 40$ kg is the mass of the interferometer test masses. ξ^2 is positive in the presence of optical spring and negative in the presence of optical antispring.

With these five parameters—g, z, $|p|$, Q_p, and ξ^2—the interferometer's optomechanical response $C(f)$ is

$$C(f) = \frac{g\, e^{-2\pi i f L/c} \times (1 + i f/z)}{1 + i f/|p|Q_p - f^2/|p|^2 - \xi^2/f^2}, \tag{6}$$

where the factor $e^{-2\pi i f L/c}$ accounts for the time delay of signals propagating down the arms.
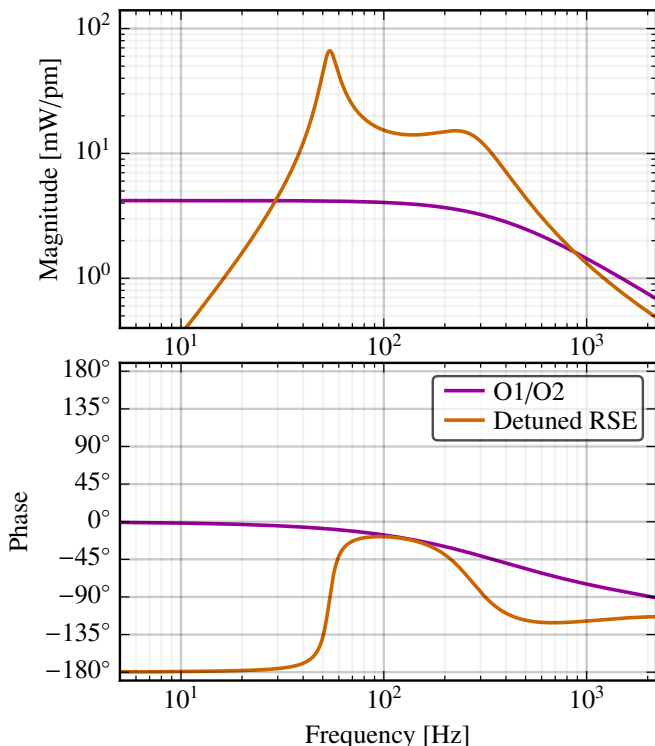
FIG. 2: Optomechanical response C of the differential arm length read-out for Advanced LIGO, shown for the two configurations considered in this work.

TABLE I: Detector calibration parameters and astrophysical signal parameters for the simulations performed in this paper. Note that for the zero-detuned O1/O2 configuration, the quantities z, Q_p and ξ^2 are not included as separate parameters in the optical response model: z is assumed to be equal to $|p|$, and Q_p and ξ^2 are assumed to be known exactly.

Quantity	O1/O2	Detuned RSE	Unit
g	4.2	9.8	mW/pm
z	365	749	Hz
$|p|$	365	276	Hz
Q_p	0.50	1.33	—
ξ^2	0.0	$+53.4^2$	Hz^2
a	0.11	0.11	μN/V
$\mathcal{M}$		30.7	$M_\odot$
η		0.247	—
t_c		2.50	ms
ϕ_c		0.600	rad
D		756	Mpc
m_G		0	kg
ρ	23	63	—

In this work we consider two different realizations of the optomechanical response for Advanced LIGO. These are shown in Fig. 2, and the corresponding parameters are given in Tab. I.

The first configuration corresponds to the Advanced LIGO detectors as they were operated during the first and second observing runs (O1 and O2). Here the detectors are operated with extremal RSE ($\phi = 90°$) [14], with 25 W of input power, an SRM with transmissivity $T_s = 37\%$, and a homodyne readout angle $\zeta = 90°$. In this configuration, the optical response $C(f)$ reduces to

$$C^{(\text{O1/O2})}(f) = \frac{g\,e^{-2\pi i f L/c}}{1 + i f/|p|}, \tag{7}$$

and hence requires characterization of only g and $|p|$, since the remaining parameters z, Q_p, and ξ^2 are then known.

The second configuration corresponds to a possible future observing run in which the detectors may be operated with detuned RSE to optimize the signal-to-noise ratio for compact binary coalescence signals. Here the detectors are operated with a one-way SRC phase $\phi = 82°$, with 125 W of input power, $T_s = 20\%$, and a homodyne angle $\zeta = 100°$. The optomechanical plant in this configuration is given by the full expression in Eq. 6, and hence requires characterization of g, z, $|p|$, Q_p, and ξ^2.

B. Feedback control loop

The optomechanical response of the interferometer is not the only transfer function that must be accounted for to produce an estimated strain signal. Because the interferometer's differential arm length is actively servoed by feeding the GW readout back to the test masses, the effect of this servo loop must be accounted for. A diagram of the servo loop is shown in Fig. 5.

We consider the loop in three parts. The first is the optomechanical response $C(f)$ (already described in Sec. II A) which converts differential arm length displacement $L_-(f)$ into power fluctuation $P(f)$ at the interferometer's dark port. The second is a set of electronic transfer functions $\{D, D_1(f), D_2(f), D_3(f)\}$, which take the power fluctuation $P(f)$ and produce a set $\{k_1(f), k_2(f), k_3(f)\}$ of control voltages intended to be fed back to three of the test mass suspension actuators. The third is a set $\{A_1(f), A_2(f), A_3(f)\}$ of actuator transfer functions which describe how each control voltage produces displacement of the test mass. The total open-loop transfer function $G(f)$ of the interferometer's differential arm length servo is

$$G(f) = CD \times (D_2 A_2 + D_3 A_3), \tag{8}$$

where we have ignored the first-stage suspension actuation term $D_1 A_1$, since its main effect is to suppress length fluctuations at frequencies below ~1 Hz, which is below the GW band.

In this paper we assume that all of the electronic, digital, and mechanical parameters that characterize D, D_2, D_3, A_2, and A_3 are fixed and known to negligible uncertainty *except* the actuation strength a (in newtons per volt) that determines the overall magnitude of the bottom-stage test mass transfer function A_3. The bottom-stage actuator is an electrostatic drive, and is affected by the accumulation of free charges on or near the test mass [15]. The distribution of charge near the test mass is known to vary with time. Therefore, the actuation strength a is included along with g, z, $|p|$, Q_p, and ξ^2 as a
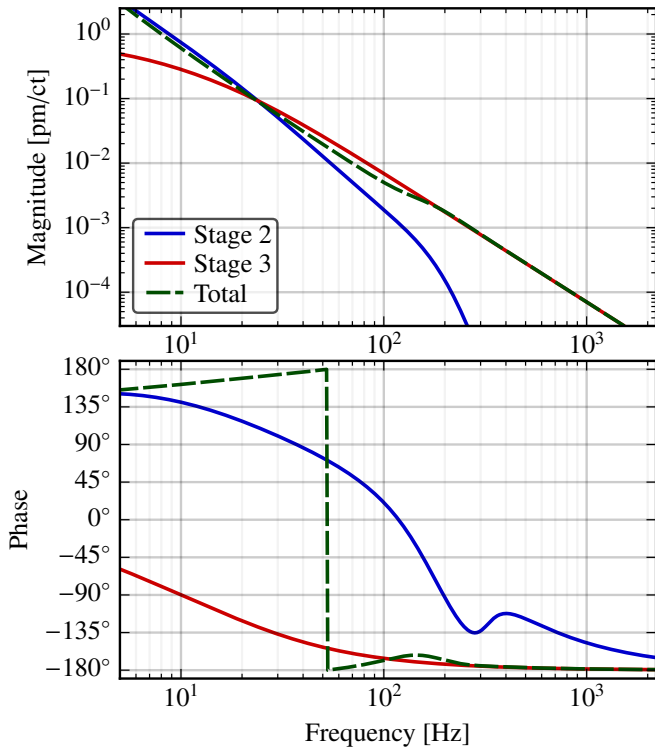
FIG. 3: Actuation transfer function $D_2A_2 + D_3A_3$ for the differential arm length loop (the low-frequency first stage D_1A_1 has been omitted). The overall strength of stage 3 is known to drift because of the charge distribution in the vicinity of the test mass; therefore, this strength a (in newtons per volt) must be carefully measured as part of the calibration process.

FIG. 4: Open-loop transfer function G of the differential arm length loop for the two configurations considered here (Eq. 8).

detector calibration parameter. In this work, its magnitude is fixed at $a = 0.11\,\mu\text{N/V}$ for both the O1/O2 and detuned-RSE configurations.

III. STRAIN ESTIMATION

The estimated freerunning strain $d(f)$ is produced from the measured power fluctuation $P(f)$ via

$$d(f) = \frac{1}{L} \times \frac{1 - G(f)}{C(f)} \times P(f), \qquad (9)$$

where $C(f)$ and $G(f)$ have been described in Sec. II, and $L = 3995$ m is the average arm length. In the GW literature, the quantity $[1 - G(f)]/C(f)$ is frequently called the "response function" and is denoted $R(f)$ [16].

We rewrite Eq. 9 as

$$d(f) = \frac{R(f)}{L} \times [P_h(f) + P_n(f)], \qquad (10)$$

where we have split $P(f)$ into two components: $P_h(f)$, which accounts for power fluctuation induced by astrophysical strain, and $P_n(f)$, which accounts for power fluctuation induced by detector noise. In this equation, it is apparent that the estimated strain $d(f)$ is not equal to the *true* strain $h(f)$ incident on the
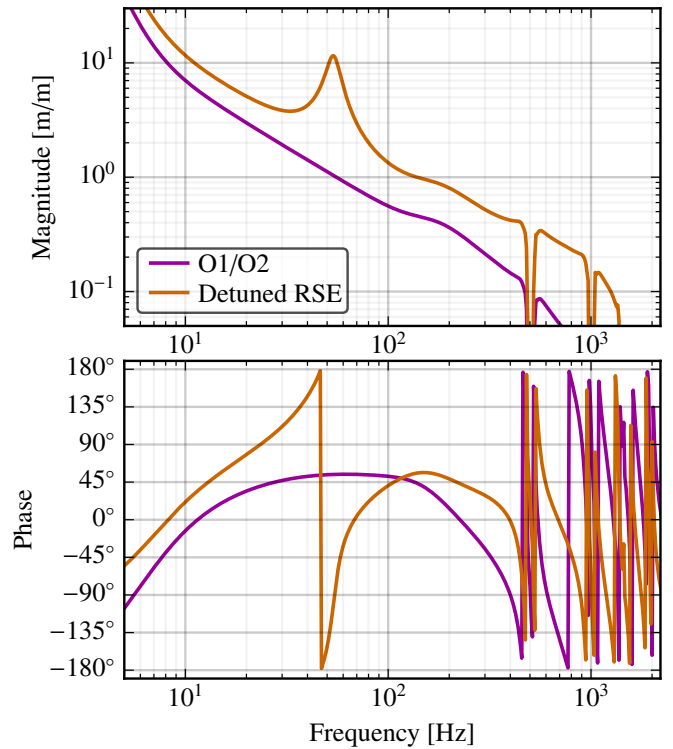
detector, for two reasons. First, the presence of P_n introduces random noise in the data on top of the true astrophysical fluctuation P_h. Second, the estimated response function $R(f)$ may differ from the true response function $R_t(f)$ because of systematic calibration errors. Errors in the estimate of R will therefore cause errors in the estimated freerunning strain d.

Similarly, the estimated strain noise power spectral density (PSD) $S_{nn}(f)$ is related to the detector's power noise PSD $S_{P_nP_n}(f)$ by

$$S_{nn}(f) = \frac{|R(f)|^2}{L^2} \times S_{P_nP_n}(f), \qquad (11)$$

and hence is similarly susceptible to systematic errors in $R(f)$.

IV. PARAMETER ESTIMATION

We now examine how systematic errors in the interferometer calibration affect the estimation of astrophysical parameters from a compact binary coalescence.

A. General formalism

We suppose we have some frequency-domain data $d(f)$, collected from one detector only, that is known to contain a coalescence signal. On the other hand, we have a model that produces a frequency-domain waveform $h(f; \boldsymbol{\theta})$; here $\boldsymbol{\theta}$ is a vector of parameters (component masses, coalescence time,
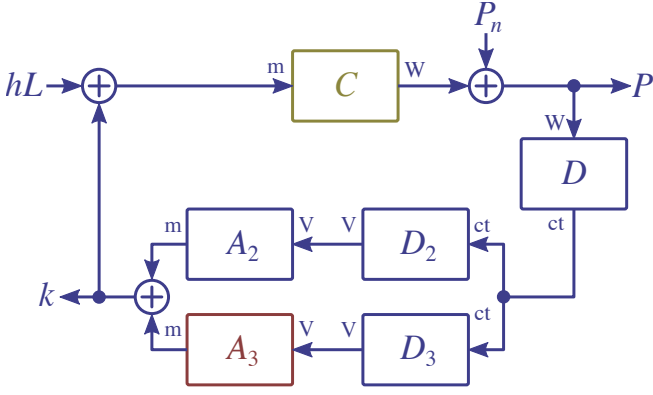
FIG. 5: Loop diagram of the differential arm length feedback control. The interferometer's optomechanical response C converts astrophysically-induced displacement hL into power fluctuation P_h, which is summed with the detector's noise P_n to produce the total observed power fluctuation P. This fluctuation is digitized and filtered by the functions D, D_2, and D_3 to produce a set of control voltages for the multi-stage test mass suspensions. These voltages are applied to the electromechanical length actuators for each suspension stage (A_2 and A_3), producing a displacement control signal k which acts to suppress differential arm length fluctuations in the interferometer.
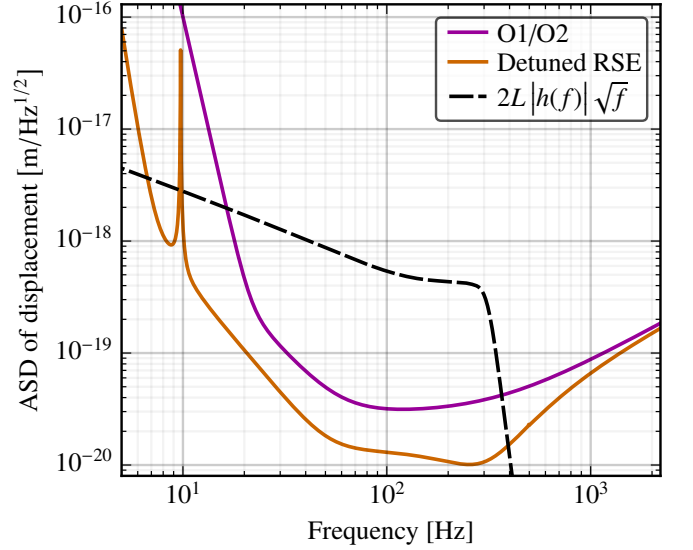

FIG. 6: Displacement-referred amplitude spectral density (ASD) $\sqrt{S_{nn}(f)} \times L$ for differential arm length noise, both for an O1/O2 configuration and for a possible future detuned RSE configuration. To facilitate comparison with the strength of the GW signal $h(f)$ described in Sec. V, we also plot the quantity $2L\left|h(f)\right|\sqrt{f}$.

etc.) describing the astrophysical event. The goal of parameter estimation is to find the parameters $\hat{\boldsymbol{\theta}}$ that best match the detector data d. In this work we will focus on choosing $\hat{\boldsymbol{\theta}}$ via maximum-likelihood (ML) estimation, although the procedure described below could be extended to Bayesian estimation [17]. ML estimation requires the construction of a likelihood function $\mathcal{L}$, which is proportional to $p(d|\boldsymbol{\theta})$, the probability of observing the detector data d given that the astrophysical event has parameters $\boldsymbol{\theta}$.

If the detector's noise is stationary and Gaussian, the logarithm of $\mathcal{L}$ is given by

$$\ell \equiv \ln \mathcal{L} = -\frac{1}{2} \int_0^\infty \mathrm{d}f \, \frac{\left|d(f) - h(f;\boldsymbol{\theta})\right|^2}{S_{nn}(f)}, \tag{12}$$

where $d(f)$ is the frequency-domain detector data, $h(f;\boldsymbol{\theta})$ is the frequency-domain waveform model (which is a function of the parameters $\boldsymbol{\theta}$), and $S_{nn}(f)$ is the power spectral density (PSD) of the detector's noise. ML estimation proceeds by maximizing the log-likelihood ℓ with respect to $\boldsymbol{\theta}$, yielding the ML estimate $\hat{\boldsymbol{\theta}}$.

Several quantities we consider later will depend on defining the usual noise-weighted, frequency-domain inner product between two sets $s_1(f)$ and $s_2(f)$ of data: [18]

$$\langle s_1 | s_2 \rangle = 4 \int_0^\infty \mathrm{d}f \, \frac{\mathrm{Re}\left[s_1^*(f) \, s_2(f)\right]}{S_{nn}(f)}. \tag{13}$$

In particular, the best matched-filter SNR that could be obtained for a waveform $h(f)$ is

$$\rho = \sqrt{\langle h | h \rangle}. \tag{14}$$

Additionally, the Fisher matrix $\boldsymbol{\Gamma}$ for a waveform $h(f)$ which depends on parameters $\boldsymbol{\theta}$ has elements [18]

$$\Gamma_{ij} = \left\langle \frac{\partial h}{\partial \theta_i} \middle| \frac{\partial h}{\partial \theta_j} \right\rangle. \tag{15}$$

The Fisher matrix describes the extent to which detector noise introduces random errors in the estimate of $\boldsymbol{\theta}$. The Cramér–Rao bound implies that the covariance matrix $\boldsymbol{\Sigma}$ of these errors is bounded elementwise by the inverse of $\boldsymbol{\Gamma}$: [18]

$$\Sigma_{ij} \geq (\boldsymbol{\Gamma}^{-1})_{ij}. \tag{16}$$

B. Effect of calibration errors

Both $d(f)$ and $S_{nn}(f)$ are also implicitly functions of the calibration parameters, which we will denote $\boldsymbol{\lambda}$. Because of systematic errors, the calibration parameters $\boldsymbol{\lambda}$ which are used to produce the strain data from power fluctuations at the GW readout port will *not* necessarily correspond to the detector's true calibration parameters $\boldsymbol{\lambda}_t$; rather, they will differ by some amount $\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda} - \boldsymbol{\lambda}_t$. Using Eqs. 10 and 11, we can write the estimated strain data and estimated strain PSD as

$$d(f;\boldsymbol{\lambda}) = \frac{R(f;\boldsymbol{\lambda})}{R(f;\boldsymbol{\lambda}_t)} \times d(f;\boldsymbol{\lambda}_t) \tag{17}$$

and

$$S_{nn}(f;\boldsymbol{\lambda}) = \left|\frac{R(f;\boldsymbol{\lambda})}{R(f;\boldsymbol{\lambda}_t)}\right|^2 \times S_{nn}(f;\boldsymbol{\lambda}_t), \tag{18}$$

where $d(f;\boldsymbol{\lambda}_t)$ and $S_{nn}(f;\boldsymbol{\lambda}_t)$ are the estimated strain data and estimated strain PSD that would have been produced in the absence of systematic calibration errors.

With these equations we can rewrite the log-likelihood explicitly as a function of both $\boldsymbol{\theta}$ and $\boldsymbol{\lambda}$:

$$\ell = -\frac{1}{2} \int\limits_0^\infty \frac{\mathrm{d}f}{S_{nn}(f;\boldsymbol{\lambda}_t)} \times \left| d(f;\boldsymbol{\lambda}_t) - \frac{R(f;\boldsymbol{\lambda}_t)}{R(f;\boldsymbol{\lambda})} h(f;\boldsymbol{\theta}) \right|^2. \quad (19)$$

Viewed in this way, we see that if $\boldsymbol{\lambda}$ is shifted from its true value $\boldsymbol{\lambda}_t$ by an amount $\Delta\boldsymbol{\lambda}$, then the log-likelihood ℓ will shift, and hence the ML estimate $\hat{\boldsymbol{\theta}}$ will shift from its true value $\hat{\boldsymbol{\theta}}_t$

by an amount $\Delta\hat{\boldsymbol{\theta}}$.

We now compute how $\Delta\hat{\boldsymbol{\theta}}$ depends on $\Delta\boldsymbol{\lambda}$. For this calculation, we assume that the signal-to-noise ratio of the GW signal is strong, so that $|P_h(f)| \gg |P_n(f)|$, and that the waveform model $h(f;\boldsymbol{\theta})$ used to match the detector data $d(f;\boldsymbol{\lambda})$ is free of any modeling errors or unmodeled parameters. This ensures that in the absence of calibration errors, we have $d(f;\boldsymbol{\lambda}) = d(f;\boldsymbol{\lambda}_t) = h(f;\boldsymbol{\theta}_t)$, where $\boldsymbol{\theta}_t$ are the system's true parameters. The ML estimate will then be $\hat{\boldsymbol{\theta}} = \boldsymbol{\theta}_t$, and hence ℓ attains its maximum value of 0.

We now examine how the ML estimate $\Delta\hat{\boldsymbol{\theta}}$ shifts in the presence of nonzero systematic calibration errors $\Delta\boldsymbol{\lambda}$. For definiteness we say that $\boldsymbol{\theta}$ consists of M parameters, and $\boldsymbol{\lambda}$ consists of N parameters. First, we allow both $\boldsymbol{\lambda}$ and $\boldsymbol{\theta}$ to vary freely, so that to second order the total change in the log-likelihood is

$$\Delta\ell = \sum_i \frac{\partial\ell}{\partial\mu_i}\Delta\mu_i + \frac{1}{2}\sum_i\sum_j \frac{\partial^2\ell}{\partial\mu_i\partial\mu_j}\Delta\mu_i\,\Delta\mu_j, \quad (20)$$

where $\mu_i, \mu_j \in \{\theta_1, \theta_2, \ldots, \theta_M, \lambda_1, \lambda_2, \ldots, \lambda_N\}$. More explicitly,

$$\Delta\ell = \sum_i \frac{\partial\ell}{\partial\lambda_i}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\lambda_i + \sum_i \frac{\partial\ell}{\partial\theta_i}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\theta_i + \frac{1}{2}\sum_i\sum_j \frac{\partial^2\ell}{\partial\lambda_i\partial\lambda_j}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\lambda_i\,\Delta\lambda_j + \frac{1}{2}\sum_i\sum_j \frac{\partial^2\ell}{\partial\theta_i\partial\theta_j}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\theta_i\,\Delta\theta_j + \sum_i\sum_j \frac{\partial^2\ell}{\partial\lambda_i\partial\theta_j}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\lambda_i\,\Delta\theta_j. \quad (21)$$

We now fix the systematic calibration errors $\Delta\lambda_1, \Delta\lambda_2, \ldots, \Delta\lambda_N$. Since we noted earlier that ℓ attains its maximum value of 0 when $\Delta\boldsymbol{\lambda} = 0$, it must be the case that $\Delta\ell$ is nonpositive, and hence the shift $\Delta\hat{\boldsymbol{\theta}}$ will maximize $\Delta\ell$. The particular values $\Delta\hat{\boldsymbol{\theta}}$ will satisfy the M equations

$$0 = \frac{\partial\Delta\ell}{\partial\Delta\theta_k} = \frac{\partial\ell}{\partial\theta_k}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t} + \sum_j \frac{\partial^2\ell}{\partial\theta_k\partial\theta_j}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\theta_j + \sum_i \frac{\partial^2\ell}{\partial\theta_k\partial\lambda_i}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t}\Delta\lambda_i, \qquad k \in \{1, 2, \ldots, M\}. \quad (22)$$

We assumed earlier that ℓ attains its maximum for $\boldsymbol{\theta} = \boldsymbol{\theta}_t$ and $\boldsymbol{\lambda} = \boldsymbol{\lambda}_t$. Therefore, we must have $\partial\ell/\partial\theta_k\big|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t} = 0$. Therefore, the relation between $\Delta\hat{\boldsymbol{\theta}}$ and $\Delta\boldsymbol{\lambda}$ is

$$-\underbrace{\begin{bmatrix} \ell_{\theta_1\theta_1} & \ell_{\theta_1\theta_2} & \cdots & \ell_{\theta_1\theta_M} \\ \ell_{\theta_2\theta_1} & \ell_{\theta_2\theta_2} & \cdots & \ell_{\theta_2\theta_M} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{\theta_M\theta_1} & \ell_{\theta_M\theta_2} & \cdots & \ell_{\theta_M\theta_M} \end{bmatrix}}_{\equiv\mathcal{H}} \begin{bmatrix} \Delta\hat{\theta}_1 \\ \Delta\hat{\theta}_2 \\ \vdots \\ \Delta\hat{\theta}_M \end{bmatrix} = \underbrace{\begin{bmatrix} \ell_{\theta_1\lambda_1} & \ell_{\theta_1\lambda_2} & \cdots & \ell_{\theta_1\lambda_N} \\ \ell_{\theta_2\lambda_1} & \ell_{\theta_2\lambda_2} & \cdots & \ell_{\theta_2\lambda_N} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{\theta_M\lambda_1} & \ell_{\theta_M\lambda_2} & \cdots & \ell_{\theta_M\lambda_N} \end{bmatrix}}_{\equiv\mathcal{M}} \begin{bmatrix} \Delta\lambda_1 \\ \Delta\lambda_2 \\ \vdots \\ \Delta\lambda_N \end{bmatrix}, \quad (23)$$

with

$$\ell_{\mu_i\mu_j} = \frac{\partial^2\ell}{\partial\mu_i\partial\mu_i}\bigg|_{\boldsymbol{\theta}_t,\boldsymbol{\lambda}_t} \quad (24)$$

and $\mu_i, \mu_j \in \{\theta_1, \theta_2, \ldots, \theta_M, \lambda_1, \lambda_2, \ldots, \lambda_N\}$.

In Eq. 23 we note that the matrix $\mathcal{H}$ on the left-hand side is the Hessian of ℓ with respect to $\boldsymbol{\theta}$. Additionally, to the matrix on the right-hand side we have assigned the letter $\mathcal{M}$. This allows us to write the relationship between calibration errors $\Delta\boldsymbol{\lambda}$ and parameter estimation errors $\Delta\hat{\boldsymbol{\theta}}$ as

$$\Delta\hat{\boldsymbol{\theta}} = \mathbf{J}\,\Delta\boldsymbol{\lambda} \quad (25)$$

with

$$\mathbf{J} = -\mathcal{H}^{-1}\mathcal{M}. \quad (26)$$

As desired, the matrix $\mathbf{J}$ in Eq. 26 quantifies how systematic calibration errors $\Delta\boldsymbol{\lambda}$ produce a shift $\Delta\hat{\boldsymbol{\theta}}$ in the ML estimate of the astrophysical signal parameters. The rest of this work will be dedicated to computing $\mathbf{J}$ for a GW150914-like coalescence signal.

C. Choice of waveform

We use a family of phenomenological waveforms [19, 20] that incorporates the dynamics of the inspiral, merger, and ringdown phases of the coalescence. This family also includes effects from the spins of the components, but in this work we constrain the spins to be zero. The five parameters we consider are

1. the chirp mass $\mathcal{M} = (m_1 m_2)^{3/5}/(m_1 + m_2)^{1/5}$, where m_1 and m_2 are the component masses;

2. the symmetric mass ratio $\eta = m_1 m_2/(m_1 + m_2)^2$;

3. the coalescence time t_c;

4. the coalescence phase ϕ_c; and

5. the effective distance D, which depends on both the actual system distance and the orientation of the source relative to the detector.

We additionally assume that the source is oriented so that the detector senses the strain from only the plus-polarization of the GW.

In parts of the next section, we will additionally consider the problem of GW parameter estimation under the assumption of a massive graviton. The effect of a massive graviton on a post-Newtonian waveform was considered by Will [21]. The energy E, momentum p, and mass m_G of the graviton are related by the dispersion relation $E^2 = p^2 c^2 + m_G^2 c^4$, and from this it can be shown that the GW waveform $h(f)$ acquires an extra phase term

$$\psi_G(f) = -\frac{\mathcal{D}m_G^2 c^3}{4\pi\hbar^2(1+Z)f} = -\frac{\pi\mathcal{D}c}{\Lambda^2(1+Z)f} \equiv -\frac{B}{f}, \quad (27)$$

where $\Lambda = 2\pi\hbar/m_G c$ is the Compton wavelength of the graviton, $\mathcal{D}$ is a cosmological distance quantity (not equal in general to the luminosity distance D), Z is the source redshift, and we have defined $B = \pi\mathcal{D}c/\Lambda^2(1+Z)$.

V. RESULTS

We consider a coalescence signal with parameters $\mathcal{M} = 30.7\,M_\odot$, $\eta = 0.247$, $t_c = 2.50\,\text{ms}$, $\phi_c = 0.600\,\text{rad}$, $D = 756\,\text{Mpc}$. The values of $\mathcal{M}$ and η are similar to the binary-black-hole signal GW150914 [2, 22]; the value of D is chosen to give an overall strain amplitude similar to that of GW150914; and the values of t_c and ϕ_c are chosen arbitrarily. Additionally, we assume the system is at a redshift $Z = 0.09$ (also similar to GW150914) and that the graviton is massless ($m_G = B = 0$; $\Lambda = \infty$).

A. O1/O2 configuration

Here we consider a coalescence detection when the detector is operating in its extremal RSE configuration, for which the optomechanical plant is given by Eq. 7. In this case, the vector of calibration parameters is $\boldsymbol{\lambda} = \begin{pmatrix} g & |p| & a \end{pmatrix}^\top$, where g and $|p|$ are the gain and pole of the optomechanical plant, and a is the actuation strength of the test mass actuator. The true calibration parameters are assumed to be $g = 4.2\,\text{mW/pm}$, $|p| = 365\,\text{Hz}$, and $a = 0.11\,\mu\text{N/V}$ (Table I). In this configuration, the signal has a matched-filter SNR $\rho = 23$.

In the rest of this section we calculate the effect of systematic calibration errors on the parameter estimation of this signal in two cases: first, in the case that our parameter estimation assumes a massless graviton; and second, in the case that our parameter estimation includes the graviton mass as a parameter to be estimated from the signal.

1. Massless graviton

Here we consider parameter estimation assuming a massless graviton ($\Lambda = \infty$ and $B = 0$), so that the vector of astrophysical parameters is $\boldsymbol{\theta} = \begin{pmatrix} \mathcal{M} & \eta & t_c & \phi_c & D \end{pmatrix}^\top$.

Applying Eq. 26 to this scenario yields the following relationship between $\Delta\hat{\boldsymbol{\theta}}$ and $\Delta\boldsymbol{\lambda}$:

$$\begin{bmatrix} \Delta\mathcal{M}/\mathcal{M} \\ \Delta\eta/\eta \\ \Delta t_c/(1\,\text{ms}) \\ \Delta\phi_c/(1\,\text{rad}) \\ \Delta D/D \end{bmatrix} = \frac{1}{10^3} \times \begin{bmatrix} -12 & -21 & 1 \\ -93 & -80 & -181 \\ -1745 & -121 & -2827 \\ -5058 & -10497 & 7770 \\ 736 & -90 & -537 \end{bmatrix} \begin{bmatrix} \Delta g/g \\ \Delta|p|/|p| \\ \Delta a/a \end{bmatrix}, \quad (28)$$

where we have expressed the relationships fractionally where appropriate.

We want to use the result in Eq. 28 to set reasonable goals on the systematic errors in g, $|p|$, and a. As noted by Lindblom [6], once the calibration-induced systematic errors $\Delta\hat{\boldsymbol{\theta}}$ are made sufficiently small, the parameter estimation will be dominated by systematics induced by detector noise, and more stringent calibration will not help. In this work, we will therefore set the calibration requirements so that the $\Delta\hat{\boldsymbol{\theta}}$ is no more than one third of these noise-induced systematic errors.

To estimate the typical size of noise-induced systematic errors, we compute the Cramér–Rao bound on the covariance matrix of $\hat{\boldsymbol{\theta}}$ via Eqs. 15 and 16. In this case, the bound implies that the diagonal elements of the covariance matrix can be no smaller than the following values:

$$\Sigma_{\mathcal{M}\mathcal{M}}^{1/2}/\mathcal{M} = 0.015 \quad (29\text{a})$$

$$\Sigma_{\eta\eta}^{1/2}/\eta = 0.051 \quad (29\text{b})$$

$$\Sigma_{t_c t_c}^{1/2} = 0.45\,\text{ms} \quad (29\text{c})$$

$$\Sigma_{\phi_c\phi_c}^{1/2} = 8.3\,\text{rad} \quad (29\text{d})$$

$$\Sigma_{DD}^{1/2}/D = 0.054. \quad (29\text{e})$$

We then use these values to set a requirement on the systematic calibration errors $\Delta\boldsymbol{\lambda}$ by requiring that the calibration errors introduce a systematic error $\Delta\hat{\boldsymbol{\theta}}$ that is less than one third that the Cramér–Rao limit. The results of this requirement are given in Tab. II.

TABLE II: Requirements on systematic calibration errors for a GW150914-like signal. These requirements are set such that the systematic calibration errors induce an error in the astrophysical parameter estimation that is less than one third of the error introduced by detector noise (as determined from the Cramér–Rao bound as described in the text).

Quantity	O1/O2, no MG	O1/O2, MG	Detuned RSE, no MG	Detuned RSE, MG
$\Delta g/g$	2 %	3 %	4 %	4 %
$\Delta z/z$	—	—	7 %	7 %
$\Delta |p|/|p|$	20 %	18 %	1.1 %	1.0 %
$\Delta Q_p/Q_p$	—	—	4 %	3 %
$\Delta \xi^2/\xi^2$	—	—	3 %	3 %
$\Delta a/a$	3 %	3 %	0.6 %	0.6 %

2. Massive graviton

Here we consider parameter estimation in which the mass of the graviton (via the parameter B defined earlier) is included in the parameter estimation, so that the vector of compact binary coalescenece signal parameters is $\boldsymbol{\theta} = \begin{pmatrix} \mathcal{M} & \eta & t_c & \phi_c & D & B \end{pmatrix}^\top$.

$$\begin{bmatrix} \Delta\mathcal{M}/\mathcal{M} \\ \Delta\eta/\eta \\ \Delta t_c/(1\,\text{ms}) \\ \Delta\phi_c/(1\,\text{rad}) \\ \Delta D/D \\ \Delta B/(1\,\text{kHz}) \end{bmatrix} = \frac{1}{10^3} \times \begin{bmatrix} -29 & -34 & -84 \\ -100 & -85 & -215 \\ -1988 & -307 & -4066 \\ -16317 & -19135 & -49625 \\ 720 & -103 & -621 \\ -26 & -20 & -131 \end{bmatrix} \begin{bmatrix} \Delta g/g \\ \Delta|p|/|p| \\ \Delta a/a \end{bmatrix}. \tag{30}$$

The Cramér–Rao-limited standard deviations on the astrophysical parameters are

$$\Sigma_{\mathcal{M}\mathcal{M}}^{1/2}/\mathcal{M} = 0.022 \tag{31a}$$

$$\Sigma_{\eta\eta}^{1/2}/\eta = 0.051 \tag{31b}$$

$$\Sigma_{t_c t_c}^{1/2} = 0.51\,\text{ms} \tag{31c}$$

$$\Sigma_{\phi_c\phi_c}^{1/2} = 13.0\,\text{rad} \tag{31d}$$

$$\Sigma_{DD}^{1/2}/D = 0.056 \tag{31e}$$

$$\Sigma_{BB}^{1/2} = 0.024\,\text{kHz} \tag{31f}$$

Since our parameter estimation should return $B = 0$ in the absence of error, the error $\Sigma_{BB}^{1/2}$ induced by the noise can be converted into an error $\Sigma_{m_G m_G}^{1/2}$ on the graviton mass or an error $\Sigma_{\Lambda\Lambda}^{1/2}$ on the graviton Compton wavelength via Eq. 27:

$$\Sigma_{m_G m_G}^{1/2} = 2\hbar \sqrt{\frac{1+Z}{1-Z} \times \frac{\pi\Sigma_{BB}^{1/2}}{Dc^3}} \tag{32}$$

$$\Sigma_{\Lambda\Lambda}^{1/2} = \sqrt{\frac{1-Z}{1+Z} \times \frac{\pi Dc}{\Sigma_{BB}^{1/2}}}, \tag{33}$$

where we have used the result from Will [21] that $\mathcal{D} \simeq (1 - Z)D$ for $Z \ll 1$. Therefore, the noise-induced error $\Sigma_{BB}^{1/2} = 0.024\,\text{kHz}$ translates to an error on the graviton mass of $\Sigma_{m_G m_G}^{1/2} = 8.0 \times 10^{-59}\,\text{kg} = 4.5 \times 10^{-23}\,\text{eV}/c^2$, and an error on the graviton Compton wavelength of $\Sigma_{\Lambda\Lambda}^{1/2} = 2.8 \times 10^{13}\,\text{km}$ [23]. These Cramér–Rao limits are combined with the matrix in Eq. 30 to produce the calibration limits given in Tab. II; these requirements are almost identical to the requirements for the massless graviton analysis, indicating that an analysis with a massive graviton does not require a more stringent calibration effort.

Eqs. 32 and 33 apply equally well as relations between the errors ΔB, Δm_G, and $\Delta\Lambda$ introduced by calibration systematics; thus, if the calibration systematics are kept sufficiently small that $\Delta B/\Sigma_{BB}^{1/2} < 1/3$, the corresponding error on the graviton mass is $\Delta m_G/\Sigma_{m_G m_G}^{1/2} = \sqrt{\Delta B/\Sigma_{BB}^{1/2}} > \sqrt{1/3} \simeq 0.6$, and on the graviton wavelength is $\Delta\Lambda/\Sigma_{\Lambda\Lambda}^{1/2} > \sqrt{3} \simeq 1.7$.

B. Detuned signal extraction configuration, massless graviton

This configuration employs both detuned sideband extraction and a non-90° homodyne angle, with parameters given in Table I. Here the vector of calibration parameters is $\boldsymbol{\lambda} = \begin{pmatrix} g & z & |p| & Q_p & \xi^2 & a \end{pmatrix}^\top$, and the vector of astrophysical parameters is $\boldsymbol{\theta} = \begin{pmatrix} \mathcal{M} & \eta & t_c & \phi_c & D \end{pmatrix}^\top$.

The resulting matrix relation between $\Delta\hat{\boldsymbol{\theta}}$ and $\Delta\boldsymbol{\lambda}$ is

$$\begin{bmatrix} \Delta\mathcal{M}/\mathcal{M} \\ \Delta\eta/\eta \\ \Delta t_c/(1\,\text{ms}) \\ \Delta\phi_c/(1\,\text{rad}) \\ \Delta D/D \end{bmatrix} = \frac{1}{10^3} \times \begin{bmatrix} -16 & -2 & 15 & -15 & -11 & -21 \\ -83 & -16 & 132 & -66 & -82 & -133 \\ 29 & -514 & 3376 & 419 & -937 & 93 \\ -7003 & -629 & 5319 & -7329 & -3909 & -8253 \\ 134 & 73 & -512 & -114 & -194 & -1031 \end{bmatrix} \begin{bmatrix} \Delta g/g \\ \Delta z/z \\ \Delta|p|/|p| \\ \Delta Q_p/Q_p \\ \Delta\xi^2/\xi^2 \\ \Delta a/a \end{bmatrix}. \tag{34}$$

The corresponding Cramér–Rao limited estimates of the astrophysical parameters are

$$\Sigma_{\mathcal{M}\mathcal{M}}^{1/2}/\mathcal{M} = 0.0020 \tag{35a}$$

$$\Sigma_{\eta\eta}^{1/2}/\eta = 0.0099 \tag{35b}$$

$$\Sigma_{t_c t_c}^{1/2} = 0.11\,\mathrm{ms} \tag{35c}$$

$$\Sigma_{\phi_c \phi_c}^{1/2} = 0.98\,\mathrm{rad} \tag{35d}$$

$$\Sigma_{DD}^{1/2}/D = 0.017, \tag{35e}$$

and the resulting limits on the systematic errors $\Delta\boldsymbol{\lambda}$ are again given in Tab. II. Additionally, the requirements are given for the additional case of parameter estimation with a massive graviton, although there is little difference compared to the massless graviton requirements.

VI. DISCUSSION AND CONCLUSION

Tab. II shows that in order to achieve noise-limited systematics for a GW150914-like signal detected by a single Advanced LIGO instrument running in an O1/O2 configuration, calibration accuracy of a few percent is required for the optical gain g and actuation strength a, and about 20 % accuracy is required on the optical pole $|p|$. If the detector is instead running in a detuned configuration with a higher sensitivity, there are a greater number of calibration parameters that must be characterized, and the required accuracy ranges from 7 % at the least stringent (for the homodyne zero z) to 0.6 % at the most stringent (for the actuation strength). The inclusion of a model with nonzero graviton mass does not significantly alter these requirements.

There are several obvious extensions of this work. The results presented here rely on a single detector, with a waveform model that does not include the spins, location, and orientation of the system. The effects of these additional parameters should be examined, using a multiple-detector configuration. The analysis should be repeated on a wide variety of coalescence systems (different component masses, spins, distances, etc.) in order to set calibration requirements that are sufficient for the majority of the coalescences that are expected to be detected with Advanced LIGO.

This semianalytical approach could be complemented by a fully numerical analysis in which one examines the effect of systematic errors in the calibration parameters on the full Advanced LIGO parameter estimation pipeline [24]. This would allow requirements on the calibration parameters to be set even for signals with SNR $\rho \lesssim 20$, where the Cramér–Rao analysis is no longer valid.

Finally, this analysis could be useful in determining calibration requirements for future generations of gravitational-wave detectors [25, 26]. These detectors are expected to be more sensitive by a factor of 10 or more in amplitude compared to Advanced LIGO, requiring similar improvements in their calibration accuracy.

ACKNOWLEDGMENTS

[1] The LIGO Scientific Collaboration, Class. Quantum Grav. 32, 074001 (2015).

[2] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. 116 (2016), 10.1103/physrevlett.116.061102.

[3] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, et al., Phys. Rev. Lett. 116, 241103 (2016).

[4] B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, et al. (LIGO Scientific and Virgo Collaboration), Phys. Rev. Lett. 118, 221101 (2017).

[5] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, K. Ackley, C. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, et al. (LIGO Scientific Collaboration), Phys. Rev. D 95, 062003 (2017).

[6] L. Lindblom, Phys. Rev. D 80, 042005 (2009).

[7] S. Vitale, W. Del Pozzo, T. G. F. Li, C. Van Den Broeck, I. Mandel, B. Aylott, and J. Veitch, Phys. Rev. D 85, 064034 (2012).

[8] C. Cutler and M. Vallisneri, Phys. Rev. D 76, 104018 (2007).

[9] M. Vallisneri, Phys. Rev. D 77, 042001 (2008).

[10] T. Cokelaer, Class. Quantum Grav. 25, 184007 (2008).

[11] S. Vitale and M. Zanolin, Phys. Rev. D 82, 124065 (2010).

[12] J. Mizuno, K. Strain, P. Nelson, J. Chen, R. Schilling, A. Rüdiger, W. Winkler, and K. Danzmann, Physics Letters A 175, 273 (1993).

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

[14] In the Hanford detector, a slight optical antispring is observed, with $\xi^2 \simeq -(7\,\mathrm{Hz})^2$ or $\phi \simeq 90.5°$, but this is not considered for the simulations in this work.

[15] D. V. Martynov, E. D. Hall, B. P. Abbott, R. Abbott, T. D. Abbott, C. Adams, R. X. Adhikari, R. A. Anderson, S. B. Anderson,

K. Arai, *et al.*, Phys. Rev. D **93**, 112004 (2016).

[16] In this work, we use the convention that the servo goes unstable if $G(f) = +1$. If instead the opposite convention is used [instability occurs for $G(f) = -1$], then the response function $R(f)$ has the form $[1 + G(f)]/C(f)$.

[17] U. von Toussaint, Rev. Mod. Phys. **83**, 943 (2011).

[18] J. D. E. Creighton and W. G. Anderson, *Gravitational-Wave Physics and Astronomy* (Wiley-VCH, 2011).

[19] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. J. Forteza, and A. Bohé, Phys. Rev. D **93**, 044006 (2016).

[20] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. J. Forteza, and A. Bohé, Phys. Rev. D **93**, 044007 (2016).

[21] C. M. Will, Phys. Rev. D **57**, 2061 (1998).

[22] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. X **6**, 041015 (2016).

[23] Note that in the noiseless limit ($\rho \rightarrow \infty$), $\Sigma_{m_G m_G} \rightarrow 0$ but $\Sigma_{\Lambda\Lambda} \rightarrow \infty$.

[24] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 241102 (2016).

[25] M. Punturo, M. Abernathy, F. Acernese, B. Allen, N. Andersson, K. Arun, F. Barone, B. Barr, M. Barsuglia, M. Beker, *et al.*, Class. and Quantum Grav. **27**, 194002 (2010).

[26] B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, K. Ackley, C. Adams, P. Addesso, R. X. Adhikari, V. B. Adya, C. Affeldt, *et al.*, Classical and Quantum Gravity **34**, 044001 (2017).