Grid-based simulation program for gravitational wave interferometers with realistically imperfect optics

Brett Bochner†
Department of Physics and Astronomy, Hofstra University, Hempstead, NY 11549

Yaron Hefetz†
14 Shoshanim St., Herzliya, 46492, Israel

We describe an optical simulation program that models a complete, coupled-cavity interferometer like those used by the Laser Interferometer Gravitational-Wave Observatory (LIGO) Project. A wide variety of interferometer deformations can be modeled, including general surface roughness and substrate inhomogeneities, with no a priori symmetry assumptions about the nature of interferometer imperfections. Several important interferometer parameters are optimized automatically to achieve the best possible sensitivity for each new set of perturbed mirrors. The simulation output data set includes the circulating powers and electric fields at various points in the interferometer, both for the main carrier beam and for its signal-sideband auxiliary beams, allowing an explicit calculation of the shot-noise-limited gravitational-wave sensitivity of the interferometric detector to be performed. Here we present an overview of the physics simulated by the program, and demonstrate its use with a series of runs showing the degradation of LIGO performance caused by realistically-deformed mirror profiles. We then estimate the effect of this performance degradation upon the detectability of astrophysical sources of gravitational waves. We conclude by describing applications of the simulation program to LIGO research and development efforts.

PACS numbers: 04.80.Nn, 07.60.Ly, 95.55.Ym, 95.75.-z

I. INTRODUCTION

The LIGO Project † is part of the current initiative to detect gravitational radiation via its perturbing effects on resonant laser interferometers. This initiative involves several large collaborations around the world, including VIRGO ‡, GEO §, TAMA ‡, and ACIGA ‡. A paramount issue in all such projects is the maximization of interferometer sensitivity to detect the extremely weak signals that are expected from even the most powerful astrophysical sources [6]. To that end, coupled-cavity systems with multiple resonant stages will be used to maximize the shot-noise-limited signal-to-noise ratio of the gravitational wave (GW) signal readout. For these complex interferometer detectors, intensive modeling is necessary to estimate their performance in the presence of general optical imperfections.

A hierarchy of approaches has been used to estimate detector performance, each method negotiating the trade-off between accuracy and computational complexity. Analytical methods may suffice for the consideration of optical defects that can be treated as pure losses, or for some cases involving geometric § or randomized § mirror deformations. A matrix model that evaluates the coupling between the first few lowest-order TEM laser modes has been shown to be useful for the study of mirror tilts and beam displacements ‡; and another matrix model using discrete Hankel transforms exists for problems with axial symmetry §. Such models, which consider the exchange of power between a limited set of pre-specified modes, allow one to obtain fast results that can be used for predicting certain important interferometer behaviors (e.g., time-dependent detector responses [11, 12, 13]), at the expense of some sophistication in modeling the detailed interferometer steady-state power buildup. For the consideration of highly general optical imperfections, however, the most comprehensive method is the complete modeling of the transverse structure of the laser field wavefronts; this method simulates the electric fields on large grids, and uses (Fourier transform-based) numerical computations for the propagations of these laser beams through long cavities ‡. This technique is useful in the case of mirrors with complex deformations, and for mirrors with significant losses (due to absorption, scattering, and/or diffractive loss from finite-sized apertures), that violate the assumption of unitarity § for mirror operators in the matrix models, thus introducing complications into that approach [12].

Grid-based simulations of intra-cavity laser fields have a long history (e.g., [16]), and have been applied previously (e.g., [17, 18, 19]) to the study of the interferometric GW detectors now being implemented. But simplifications are generally imposed, such as restricting the optical deformations that are studied (e.g., considering only geometric imperfections, like tilt and curvature mismatch [17, 18]), and/or by modeling a relatively simple cavity system [19]. In this paper, we describe a simulation program (originally based upon the work of Vinet et al. [15]) that has been extended to efficiently model the complex fields that build up between realistically imperfect optical components, while resonating within complete, coupled-

†Electronic address: Brett.Bochner@alum.mit.edu, phybdb@hofstra.edu
‡Electronic address: Yaron.Hefetz@hotmail.com
cavity interferometers like those used for the LIGO detectors. Versions of our program have been used for a variety of applications by the gravitational wave community, including numerous design and performance estimation tasks conducted by the LIGO group itself (see Section VI), as well as for collaborative investigations between LIGO scientists and other groups — such as ACIGA-LIGO efforts to explore alternative interferometer length control schemes, and TAMA-LIGO efforts to estimate the effects of mirror imperfections and thermal lensing upon the performances of their future, large-scale interferometers.

A wide variety of interferometer imperfections can be modeled with our program, including mirror tilts and shifts, beam mismatch and/or misalignment, diffractive loss from finite mirror apertures, mirror surface figure and substrate inhomogeneity profiles — in particular, using deformation phase maps that are adapted from measurements of real mirror surfaces and substrates — as well as fluctuations of reflection and transmission intensity across the mirror profiles. In addition to the main (carrier frequency) laser field, auxiliary fields (i.e., radio frequency sidebands) for LIGO’s heterodyne signal detection scheme are also modeled, allowing us to give absolute numbers for the shot-noise-limited GW-sensitivity of a single LIGO interferometer. Furthermore, a number of active optimization procedures are performed continuously during code execution, guaranteeing that several key parameters in the LIGO configuration will be brought to their optimum values upon program completion: the GW-sensitivity is thus maximized for each specific “realistic” interferometer, given the particular imperfections being simulated in that run. Considering the time needed for program execution, we note that though our principal usage of the program has been on various supercomputing platforms, a full run of the simulation code (with all options included) can be performed in a non-prohibitive amount of time on a modest Sun SPARCstation, even for significantly (and non-symmetrically) deformed mirrors. This has been achieved through a combination of fast iteration techniques, efficient parameter optimization routines, and procedures designed to carefully choose the initial guesses for laser fields that are computed via relaxation. Lastly, we note that versions of the code are available for both the first-generation LIGO and advanced-LIGO (Dual Recycling) configurations, though we will focus primarily upon the former in this paper.

The discussion is organized as follows: in Section II we give a description of the physical system that is modeled here (i.e., a first-generation LIGO interferometer), and the computed electric fields that are necessary for the calculation of its shot-noise-limited GW-sensitivity function. In Section III we provide an overview of the technical details of the program’s operations, including the modeling of the optics, the iterative method used for computing the interferometer’s electric fields, and the various optimization procedures that are performed to maximize its sensitivity. In Section IV we present an array of runs with representative sets of mirror substrate and surface deformation maps adapted from real mirror measurements, in order to demonstrate how realistically-deformed mirrors will reduce the circulating power stored in a LIGO interferometer (as well as altering its modal structure), thus degrading the interferometer sensitivity and interfering with its control systems. In Section V we discuss the impact of optical imperfections upon LIGO science capabilities, by estimating how deformed-mirror effects will reduce LIGO’s ability to detect astrophysical sources of gravitational waves, such as non-axisymmetric pulsars and coalescing black hole binaries. In Section VI we conclude with a discussion of various LIGO research and development initiatives to which this grid-based simulation work has contributed.

II. THE PHYSICAL SYSTEM THAT IS MODELED

Figure II is a schematic diagram of the core optical configuration of a full-sized, first-generation LIGO interferometer. A summary of typical interferometer (and computational) parameters for this configuration is presented in Table I. These are the primary program input values that we will use for the sample runs to be presented in Section IV. Not shown (or modeled by us) are the mode cleaning, frequency stabilizing, and matching optics which prepare the laser light for the interferometer; also not modeled are the pickoffs, phase modulators, or the full control system apparatus that will be used in a real interferometer to read out its complete operational state.

The system depicted in Fig. II essentially a Michel-
TABLE I: Typical parameter values for a first-generation LIGO interferometer, including physical specifications and computational parameters required for the simulation program. The labeling of the optical elements are as depicted in Fig. 1. Some parameters are optimized during code execution, and are given here only as approximate ranges of values.

| Quantity                        | Value(s)   |
|---------------------------------|------------|
| Laser Wavelength                | 1.064 μm (Nd:YAG laser) |
| Modulation Frequency            | \( \nu_{\text{mod}} \approx 24.0 \text{ MHz} \) |
| Cavity Lengths                  |            |
| \( L_1 = 6.1 \text{ m} \)      |            |
| \( L_2 = 4.19 \text{ m} + L_{\text{asym}} \) |            |
| \( L_3 = 4.19 \text{ m} - L_{\text{asym}} \) |            |
| \( L_{\text{asym}} \approx 9 - 25 \text{ cm} \) |            |
| \( L_4 = L_5 = 4.0 \text{ km} \) |            |
| Mirror Curvature Radii          |            |
| \( R_{\text{curv,1}} = 10.0 \text{ km} \) |            |
| \( R_{\text{curv,2}} = R_{\text{curv,3}} = 14.6 \text{ km} \) |            |
| \( R_{\text{curv,4}} = R_{\text{curv,5}} = 7.4 \text{ km} \) |            |
| Mirror Intensity Reflectivities  |            |
| (Refl.-Side)                    |            |
| \( R_2 \approx 9861 - 9390 \)   |            |
| \( R_3 \approx 97 \)            |            |
| \( R_4 \approx 99994 \)         |            |
| \( R_{\text{BS}} = 49992 \)     |            |
| (AR-Side)                      |            |
| \( R_1 \approx 9861 - 9390 \)   |            |
| \( R_3 \approx 968817 \)        |            |
| \( R_{\text{BS}} = 49971 \)     |            |
| Mirror Intensity Transmissions  |            |
| (Both Sides)                   |            |
| \( T_1 \approx 01385 - 06995 \) |            |
| \( T_2 = T_3 = 02995 \)         |            |
| \( T_{\text{BS}} = 05003 \)     |            |
| Beam Waist Diameter             | 7.0 cm     |
| Mirror Aperture Diameters       | 24 cm (Circular Profiles), 24.4 × 17.2 cm (Beamsplitter at 45° w.r.t. Beam Axis) |
| Mirror Thicknesses              | Beamsplitter = 4 cm, All Others = 10 cm |
| (Perpendicular to Surface)      |            |
| Substrate Refraction Index \( n \) | 1.44963 |
| Calculational Window Size       | 70 cm × 70 cm (Square) |
| Gridling of Calc. Window        | 256 × 256 pixels |

son interferometer, converts the differential arm length changes caused by a gravitational wave (GW) into an oscillating output field amplitude at the exit port of the beamsplitter, where a carrier field dark-fringe would otherwise (ideally) have been maintained. The partially-transmitting input mirrors \( T_2, T_3 \) and highly-reflective end mirrors \( R_4, R_5 \) form Fabry-Perot (FP) arm cavities in the “inline” and “offline” arms, which amplify this effect; and the power recycling mirror \( R_1 \) provides a broadband amplification for the stored energy in the interferometer that is available for signal detection \[28\]. Coupling occurs between the fields from each of the different cavities in this optical configuration, and is responsible for the complex response behaviors of the system.

Radio-frequency modulation sidebands ("RF-sidebands") impressed upon the carrier-frequency input light serve as a “local oscillator” for a heterodyne detection scheme \[27\]. Substantial RF-sideband power is made to emerge from the beamsplitter exit/signal port, and it is added to the quantity of GW-induced carrier light that is escaping there, to ultimately produce an output signal linear in \( h \), the dimensionless amplitude of the GW. The carrier power is maximized by giving it a double-resonance: it is first resonant in (either of) the FP arm cavities, and again resonant in the “power recycling cavity” (PRC) formed by mirrors \( R_1, R_2 \), and \( R_3 \). Alternatively, the RF-sidebands (in this standard LIGO detection scheme) are only resonant in the PRC, to take advantage of the broadband amplification there; but they are far-off-resonance in the long FP-arms, so that they may serve as a stable reference which is unaffected by gravitational waves.

Assuming that the carrier is held to its double resonance (giving it a phase shift of \( \pi \) in resonant reflection from the FP-arms \[27\]), the RF-sidebands have their own resonance requirements, defined by the two conditions: 
\[
(2 \, k_{\text{mod}, L_{\text{arm}}} \approx N_{\text{odd}} \cdot \pi, \text{ and } (2 \, k_{\text{mod}, L_{\text{PRC}}} \approx N_{\text{odd}} \cdot \pi)
\]

Referring to Fig. 1, \( L_{\text{arm}} \) represents \( L_4 \) or \( L_5 \), \( L_{\text{PRC}} \equiv \frac{L_1 + \frac{L_2 + L_3}{2}}{2} \), and the RF-modulation frequency is defined according to 
\[
\nu_{\text{mod}} = 2 \pi |k_{\text{mod}}| \approx \frac{\pi}{2 |\nu_{\text{carr}} - \nu_{\text{SB}}|}
\]
The first (anti-resonance) condition is not enforced as a precise equality, in order to avoid the unwanted resonance of non-negligible second-order modulation sidebands at \( \nu_{\text{mod}} \) in the FP arm cavities. But the second (PRC-resonance) condition must be achieved to sub-wavelength tolerances, and requires careful fine-tuning corrections due to the effects of realistically-deformed optics (see Sec. III C 2).

The heterodyne GW-signal is obtained by interfering the GW-induced carrier beam at the beamsplitter exit/signal port with the emerging sideband beams, and demodulating the resultant photodiode output signal at \( \nu_{\text{mod}} \). Efficient coupling of the RF-sidebands to the signal port is achieved by incorporating a macroscopic length asymmetry between the two arms of the PRC, \( L_{\text{asym}} \equiv \frac{L_2 - L_3}{2} \), thus allowing an optimal fraction of sidband power to be extracted at the exit port during each round-trip through the PRC, even while the carrier is held to a dark-fringe there. This signal generation method (see Sec. III C 4) is referred to as the “Schnupp asymmetry scheme” \[29\].

To simplify our simulation task, we model only those aspects of the LIGO system which have a direct role in generating the GW-signals: the aforementioned carrier beam, and its RF-sidebands, resonating in the core optical system of Fig. 1. If we make the (good) approximation that most of the sensitivity-limiting shot noise at the beamsplitter exit port is due to these fields — some carrier field power emerging (primarily) because of imperfect dark-fringe contrast, and RF-sideband field power being maximally channeled to the exit port — and neglect contributions from various other fields used for interferometer control systems, or from other (controlled) noise couplings \[27\], then we can do a full calculation of the shot-noise-limited GW-sensitivity of the simulated interferometer. Beyond this, rather than explicitly emulating all of the control systems of the real LIGO, the simulation program uses alternative methods (see Sec. III C 4) for the numerous parameter adjustments (resonance finding,
etc.) that are needed to optimize interferometer performance; and it does so in a manner that reflects the behavior of the real system as accurately as possible.

In summary, the primary goal of our modeling efforts is to completely solve for the steady-state carrier and sideband fields that resonate in a LIGO interferometer with realistic optical imperfections, which is held to its proper operating point, and is optimally configured for maximal GW-sensitivity. The full effects of optical imperfections upon the sensitivity of interferometric detectors are then determined from these simulated cavity fields.

III. THE FUNDAMENTALS OF THE SIMULATION PROCESS

Here we present an overview of the program and its operations during execution, and the additional tasks which must be done during pre- and post-processing stages in the course of performing simulation experiments. The full technical details of our work can be found in [24].

In regards to computing platforms, the simulation program was originally written in the SPARCcompiler 3.0 version of Fortran 77. A complete run (including carrier and sideband frequencies, with all interferometer fields computed and all optimizations done) with a set of non-ideal mirrors, and with (for example) 128 × 128 pixelized grid maps used for the optical computations, takes less than a day on a 2-processor SPARCstation 20. This computation time is reduced to a couple of minutes when the program is executed on massively-parallel supercomputing platforms (e.g., the CACR Paragon machine Trex, a 512 (compute) node machine utilizing Intel i860 processors.

A. Basic Optical Operations

We use the customary approach for the grid-based modeling of the laser field wavefronts: a primary propagation direction is assumed for the collimated beam(s) in each part of the interferometer, and a perpendicular slice can be taken anywhere along the beam propagation axis, at locations of interest. Each of these beam slices is recorded on a two-dimensional (2D) grid, with a pixel entry representing the complex electric field ("e-field") amplitude at that transverse spatial position in the slice. Figure 2 is an example of such a grid-based electric field, in particular that of a Hermite-Gaussian TEM$_{10}$ mode. No polarization vector is currently recorded in our grids (we cannot, for example, model birefringence effects). The precision that can be achieved by the program depends upon how accurately it simulates the two basic physical processes which must be performed on the interferometer e-fields: propagations, and interactions with mirrors. We discuss propagations first.

The program utilizes Siegman’s method for the plane-to-plane propagation of light in the paraxial approximation, performed via a three step process: a Fourier transform and an inverse transform, sandwiched around a pixel-by-pixel multiplication step in spatial-frequency-space ("k-space"), in which the e-field slice is multiplied with a distance-dependent 2D propagator matrix of the same pixelization. For each curved, potentially jagged mirror profile to be simulated, a flat plane is defined near that side of the optic to serve as a reference plane. Propagations thus translate an e-field slice through the large distance from an initial reference plane to a destination reference plane near a mirror (or at any chosen position along the beam axis). The computationally intensive parts of this process (i.e., the transforms) can be performed rapidly by a Fast Fourier Transform (FFT) routine (e.g., [32]). All such macroscopic-distance propagations (i.e., $\Delta L \gg \lambda_{\text{car}}$) of e-fields are done in this manner.

A very important issue to deal with for propagations is the problem of aliasing, a common complication for calculational procedures that use discrete Fourier transform methods. Aliasing can create artifacts in the simulation results if (relatively) large-angle scattering sends power beyond the edge of the grid calculational window.
During the propagations. Such power automatically re-enters the calculation from the other side of the grid, and may (fraudulently) be Incorporated back in the physical simulation, instead of being filtered out, as it would be by absorbing baffles in a real interferometer. This problem is most significant for the long propagations through the 4 km FP arm cavities.

We can reduce (or eliminate) such aliasing by filtering the relevant $k$-space matrix that will multiply an e-field slice, in between the Fourier transform and inverse-transform steps of a propagation. The goal is to preserve as much “real power”, while eliminating as much “aliased power”, as possible; the distinction between them is that the former always stays within the apertures of the finite-sized mirrors, while the latter leaves the grid completely, but re-enters from the other side and comes far enough into the middle of the grid to fall once again within the mirror apertures.

Let $W$ be the physical side-length of the square calculational window, $A$ be the mirror aperture diameter, and $L$ be the propagation distance length. The cut-off between physically real power and “aliasing” power is a matter of propagation angle: power traveling at $\theta < \theta_a \equiv A/L$ may be able to stay within the mirrors, while power with $\theta > \theta_a \equiv (W - A)/L$ usually cannot, but will sometimes be able to leave the grid and return to erroneously re-enter the mirror apertures. For a discrete grid with $N_{\text{pix}}$ pixels on a side (numbered $n = -(N_{\text{pix}}/2 - 1) \ldots (N_{\text{pix}}/2)$), these angles correspond to (respectively) the pixel numbers $N_r = \text{Int}[A \cdot W/(L \cdot \lambda)]$ and $N_a = \text{Int}[(W - A) \cdot W/(L \cdot \lambda)]$, where $\lambda$ is the light wavelength. All of the aliasing power can be safely eliminated by nulling pixels with $|n| > N_a$ in the $k$-space propagator matrix, without removing any real power, as long as $N_a > N_r$. But if $N_a < N_r$, then one must choose some compromise between keeping real power and cutting out aliasing power for pixels $N_a < |n| < N_r$. One can, of course, force $N_a > N_r$ to be true by increasing the calculational window size to be much larger than the mirrors (i.e., $W > 2A$), but this requires more pixels to be used in the grid (e.g., $256 \times 256$), which increases the computational load. Making $W$ large without using an adequate number of pixels would lead to poor sampling of the laser beam itself [24], thus creating a new aliasing problem, due to the inadequate resolution of the discrete grids. In our runs, we use large calculational windows and many pixels whenever possible (such as for the runs presented below); but when necessary, the program has an option which applies an apodization scheme to “trim” the propagation operators in a way that enforces a graduated compromise between keeping real power and eliminating aliasing power.

Next, we consider mirror interactions. The program must carry out two basic mirror interaction operations: reflections ($r$) and transmissions ($t$). Mirrors will not be perfectly uniform or flat in the direction transverse to the beam; they will have inhomogeneities in the refraction index and/or thickness of their substrates, spatially-varying surface height profiles, variations in the quality of the reflective coatings, macroscopic curvatures, tilts, etc. This leads to optical path length variations (i.e., spatially-varying phase delays) across their profiles, as well as variations in the amplitudes of $r$ and $t$, the latter mostly due to variations in the reflective side or anti-reflective (AR) side coatings. These phase and amplitude effects are simulated by creating complex mirror maps for all $r$ and $t$ operators, which are also recorded on 2D pixelized grids. As demonstrated by Vinet et al. [12], it is a good (“short distance”) approximation to treat each pixel of the beam as an independent little plane wave, and reflect (or transmit) that piece of the e-field by multiplying that pixel in the e-field map by the corresponding pixel in the relevant mirror map, so that each e-field pixel interacts only with the mirror pixel located immediately in front of it. Thus each mirror reflection or transmission operation is reduced to a pixel-by-pixel multiplication step between a mirror operator map and the e-field slice on the reference plane near to it. The simulation program uses these mirror interaction operations, in combination with the propagation algorithm described above, to model the entire behavior of the laser fields in the interferometer.

There are some limitations in using this methodology for mirror interactions, besides the obvious one of restricting it to operations over short distances (i.e., in the near-field). Vinet et al. [12] derive a requirement which specifies how small the deviations of a mirror profile can be from its idealized shape, before this pixel-by-pixel multiplication method loses a large amount of accuracy compared to the exact calculation via the Huygens-Fresnel integral formulation in scalar wave theory [14]. Realistically imperfect (LIGO-quality) mirrors easily satisfy this requirement. Furthermore, as noted by Tridgell et al. [17], the difference in phase between one pixel in a mirror map and its neighbor must be smaller than $2\pi$ if a continuously-varying surface displacement on the mirror is to be adequately sampled by the grid. For our simulation parameters (cf. Table 1), this limits mirror tilts to $\theta < 10^{-4}$ radians (not including the base $45^\circ$ tilt of the beamsplitter, which is handled separately), and mirror curvature radii to $R_{\text{curv}} > 1.3$ km; both of these limitations are easily satisfied in our runs. Finally, we supplement these requirements with a general rule-of-thumb: a tiny Gaussian beam, with a waist size equal to the width of a pixel (and an initial propagation direction with respect to the beam axis that is defined by the e-field’s wavefront curvature at the location of that pixel), must not expand or shift over so much so that it would relocate a significant fraction of its power onto any neighboring pixels, during the entire course of the mirror interaction. If that rule is broken, then this method for mirror interactions will be inaccurate, and the pixel-by-pixel multiplication of maps will not be sufficient for modeling reflection or transmission operations.

To create a complete (but not over-determined) description of a mirror, we consider the full information...
expressed by mirror interactions, as well as physical sym-
metries and the conservation of energy. A 2-port mirror
(reflective- and AR-sides) must relate 2 complex input
fields to 2 complex output fields. Thus 4 complex (or
8 real) elements are needed, at each pixel location, to
specify that pixel of mirror completely. These 4 com-
plex elements correspond to the complex reflection and
transmission map and offline paths). The mirror description is thus reduced
to 3 complex elements per pixel, i.e., 3 complex mirror
dynamics of LIGO control systems do not necessarily
come into this calculation, since the mirrors act like
“free-masses” at the GW-frequencies that LIGO is most
sensitive to, and because the GW’s are presumably not
strong enough to interfere with the interferometer’s reso-
nant lock.) The primary task of the simulation program,
therefore, is to compute the relaxed, steady-state, reso-
nant electric fields that build up in each part of the in-
terferometer, when it is excited by a laser beam of fixed
amplitude, orientation, and frequency (at the nominal
carrier and sideband frequencies), entering through the
camera recycling mirror.

The multiply-coupled-cavity nature of the interferome-
ter, and the physical complexity of the simulation, means
that the program must iteratively solve for an irreducible
number of resonant e-fields (one per cavity) in the inter-
ferometer. For the first-generation LIGO configuration,
three e-fields must be computed via relaxation: one in the
PRC, and one in each of the FP arm cavities (though the
precise location of each relaxed e-field in its cavity can
be arbitrarily chosen). When the relaxation algorithm is
finished, these three principal e-fields can be propagated,
reflected, transmitted, and/or superposed together in or-
der to generate the complete steady-state e-fields every-
where in the system.

For each e-field that will be relaxed, there is a steady-
state equation of the form:

\[
\vec{E}_{\text{steady-state}} = \hat{\Lambda} \vec{E}_{\text{steady-state}} + \vec{E}_{\text{exc}} \quad (3)
\]

The left hand side of Eq. 3 is the e-field to be solved for;
we display it here as a vector because it has a propaga-
tion direction: either “forward” or “backward” along the
beam axis. The expression, \(\hat{\Lambda} \{\vec{E}_{\text{steady-state}}\}\), represents
this steady-state e-field after it has gone on a round-trip
through the interferometer, and has returned to its start-
ing point. (Note that this “round-trip” includes all possible closed-loop paths through the interferometer that do not ever pass through the same location — with the same propagation direction — of any of the other principal e-fields being relaxed.) Lastly, $\vec{E}_{\text{exc}}$ represents a composite excitation e-field, which may consist of both a primary excitation e-field (e.g., the input laser beam, for the e-field being relaxed in the PRC), plus any “leak” e-fields that arrive from the other principal fields being relaxed (e.g., fields leaking from the FP-arms into the PRC). It is this leak-field part of the excitation term which is the source of many coupled-cavity effects in the interferometer; and it is the presence of the $\Lambda \{ \vec{E}_{\text{exc}} \}$ term in Eq. (3) which makes $\vec{E}_{\text{exc}}$ “self-coupled”, thus requiring us [24] to use an iterative process to solve for it.

Specifying a relaxation convergence scheme is done by prescribing how the $(N+1)^{\text{th}}$ iterative guess for a given e-field is obtained from its $N^{\text{th}}$ iteration (and from the $N^{\text{th}}$ iteration of the other e-fields being relaxed simultaneously). Considering Eq. (3), the simplest possibility is to make the choice:

$$\vec{E}^{(N+1)} = \Lambda \{ \vec{E}^{(N)} \} + \vec{E}_{\text{exc}}^{(N)}.$$  (4)

This process can then be repeated for as many iterations as necessary, until the steady-state equation, Eq. (3), is satisfied by $\vec{E}^{(N)}$ to within a pre-specified threshold of accuracy (typically 1 part in $10^4$ in power, for our runs). This relaxation formula is guaranteed to succeed at smoothly converging an e-field to its correct steady-state form — barring complications which may arise from changes to the interferometer caused by the parameter optimization routines (see Sec. III C) — because it imitates how power actually does build up, through a sum of many bounces, in the cavity system of a real interferometer. This iteration process has been used successfully by previous researchers (e.g., [12]).

Since this method does model the true physical buildup of power, however, it requires a great many iterations to converge, especially for coupled-cavity systems with large Q-factors (i.e., long transient decay times). We have found this relaxation scheme to be forbiddingly slow for a full-LIGO simulation program.

Therefore, as first suggested (and implemented, for a simpler cavity arrangement) by a LIGO colleague [34], our simulation program uses a different approach. Instead of Eq. (4) for choosing the $(N+1)^{\text{th}}$ iteration, we use the following expression:

$$\vec{E}^{(N+1)} = a \cdot \vec{E}^{(N)} + b \cdot (\Lambda \{ \vec{E}^{(N)} \} + \vec{E}_{\text{exc}}^{(N)}) + c \cdot \vec{E}_{\text{exc}}^{(N)},$$  (5)

where $(a, b, c)$ are unknown, complex coefficients that are solved for by minimizing the error in what the steady-state equation will become in the next round, with $\vec{E}^{(N+1)}$ (as a function of these unknown coefficients) taking the place of $\vec{E}^{(N)}$ in Eq. (3), and $\vec{E}_{\text{exc}}^{(N)}$ taking the place of $\vec{E}_{\text{exc}}$ (noting that $\vec{E}_{\text{exc}}^{(N)}$ does not yet take into account the new $(a, b, c)$ coefficients that will be chosen for the other e-fields being iterated). The resulting expression for the steady-state error can be differentiated with respect to the six available degrees of freedom (the real and imaginary parts of $a$, $b$, and $c$), resulting in six simultaneous equations that are solved via matrix inversion. This relaxation method (which essentially reverts to the simpler method if we set $(a, b, c) = (0, 1, 0)$) gains a huge advantage compared to the method of Eq. (4), by having these additional useful degrees of freedom available for each iteration [34]. The number of iterations necessary to achieve convergence are greatly reduced (often by $\sim$1-2 orders of magnitude), resulting in much faster e-field relaxation.

This “abc” iteration algorithm does have certain drawbacks. It is more difficult to implement in code, and it requires more information to be stored (thus using more memory), and more propagations to be performed (despite what is stated in [34], because we use three unknown coefficients instead of two: $c \neq 0$), during each iteration. It is also somewhat less stable, because of the wide range of abc-coefficients that may be chosen by the error minimization algorithm; in fact, large excursions in the iterated e-field structure and power level appear to be required during (typically the first 50-100 iterations, in order for the accelerated convergence scheme to function properly. We have found that the best (straightforward) way to enhance the convergence stability of the “abc” relaxation algorithm, without slowing it down to the pace of the method of Eq. (4), is to place hard limits upon the allowed choices of the abc-coefficients during the early stages (the first 100-200 iterations) of runs. This allows the relaxation process to avoid failure during the initial, large e-field adjustments, after which it settles down to converge efficiently to the steady-state solution.

Lastly, we note that the choice of $(a, b, c)$ for any one field will affect future iterations of the other fields, because they are all coupled via the $\vec{E}_{\text{exc}}$ leak-fields, iterating them independently from one another ignores this coupling, and causes some slowing (and oscillation) of the relaxation process (we see a temporary “shifting” of power back and forth between the PRC field and the FP arm cavity fields). This problem can be eliminated by choosing the $(a, b, c)$ coefficients for all three relaxed e-fields simultaneously, by calculating the 18 unknown parameters (i.e., 9 complex “abc” coefficients) in order to minimize a specially-weighted sum of the iteration errors for all three relaxed fields. Such a “global abc” relaxation scheme is significantly more demanding in terms of coding difficulty and computer memory usage, but it is capable of making the relaxation process more stable, and further reduces the number of iterations needed to reach convergence. Though we have not yet implemented this global relaxation scheme into simulations of the first-generation LIGO interferometer, we have successfully incorporated global abc relaxation (with four relaxed e-fields needed, instead of three) into the version of our code used for the advanced-LIGO Dual Recycling
configuration, with good results [24].

In summary, the “abc” iteration method presented here is a sufficiently reliable and extremely fast relaxation scheme which greatly reduces program execution time, thus enabling us to perform full-LIGO simulation runs (with significant optical deformations) in a reasonable amount of time.

C. Parameter Optimizations

At the core of our efforts to make a realistic simulation of a LIGO interferometer are several procedures which bring the system into an “optimal” configuration for signal detection. The problem of optimization is a highly nontrivial matter: not only does the interferometer possess numerous degrees of freedom that must be optimized (e.g., all resonant cavity lengths, the Schnupp length asymmetry, etc.), but each evaluation of performance (i.e., GW-sensitivity) as a function of these optimizable parameters is extremely time consuming, since it requires the detailed computation of the carrier and sideband e-fields everywhere in the interferometer. It therefore appears infeasible to use the brute-force method of optimizing the interferometer’s GW-sensitivity function by evaluating it for a thorough sampling of points over the entire, multidimensional parameter space.

As an alternative, each key parameter is optimized separately in our program, using some error signal (to be brought to zero), or some function of merit (to optimize), which is strongly (and solely) dependent upon that individual parameter, and which is a true measure of when that parameter is well chosen for the maximization of GW-sensitivity. This strategy is aided by the fact that the interferometer’s GW-sensitivity is quite insensitive to particular combinations of parameter changes, such that if certain optimizable parameters are displaced from one apparently optimal point in the multidimensional parameter space, then the other parameters will adjust themselves (via our optimization procedures) to compensate with virtually no reduction in overall interferometer sensitivity. We note that these optimization routines in our program are performed concurrently with the e-field relaxations, so that the final results emerging from the iterative relaxation scheme are the steady-state e-fields of a fully-optimized interferometer.

The following subsections give an overview of the parameters that the program optimizes, the criteria for optimizing them, and the physical considerations which underlie their significance. Note that we have performed extensive empirical tests to verify that each of the optimization procedures discussed below works as desired.

1. Length Adjustments for Carrier Resonance

As discussed in Sec. III, the carrier frequency beam must have a double resonance in the system consisting of the power recycling cavity and Fabry-Perot arm cavities, while being held to a dark-fringe at the exit port of the beamsplitter. Our method for achieving these resonance (and dark-fringe) conditions is similar to that of Vinet et al. [15], in which we null the computed “phases” between certain specified cavity e-fields by adjusting the various cavity lengths. (A length change will alter the phase relationship between an e-field that has taken a round-trip through an adjusted path length, and one that has not — potentially bringing a cavity to resonance.) Alternatively, we have chosen not to use the method of McClelland et al. [18], in which the e-fields are re-relaxed for each trial set of lengths until the configuration for maximum power buildup is found, since that would involve a time-consuming search over a multi-dimensional phase space of independent, controllable cavity lengths.

The phase $\Phi$ between two e-fields is defined via an overlap integral (or more precisely, by a discrete sum over the pixelized $N \times N$ grids), as follows:

$$\Phi[\vec{E}_1, \vec{E}_2] \equiv \tan^{-1} \left[ \frac{\Im \langle \vec{E}_1 | \vec{E}_2 \rangle}{\Re \langle \vec{E}_1 | \vec{E}_2 \rangle} \right],$$

with:

$$\langle \vec{E}_1 | \vec{E}_2 \rangle \equiv \left( \frac{\text{Calc. Window size}}{N^2} \right) \sum_{i=1}^{N} \sum_{j=1}^{N} \vec{E}_1^{\ast}(i,j) \cdot \vec{E}_2(i,j).$$

The phase between an e-field and its round-tripped analogue can be driven to zero by a (sub-wavelength) path length change, computed via the formula: $\Delta L = -\Phi \cdot (-\lambda_{\text{laser}}/4\pi) \equiv \Phi/2k$ (an extra factor of $1/2$ is included here because the round-trip doubles the phase change caused by a cavity length adjustment). Note that this procedure cannot simply be performed once — these phases depend in detail upon the structure of the iterated e-fields, and they must be repeatedly measured and adjusted throughout the e-field relaxation process.

To achieve resonance in a particular cavity, the proper solution is to ensure that the fields $\vec{E}_{\text{steady-state}}$, $\hat{A}\{\vec{E}_{\text{steady-state}}\}$, and $\vec{E}_{\text{exc}}$ (as defined in Sec. III B) all
have zero phase between them. (This condition of mutually zero phase is given some theoretical justification in the literature [33], as well as being apparent from Eq. (3)). The phase-nulling procedure is performed in the FP arm cavities and in the PRC by performing microscopic adjustments to the three relevant cavity lengths here: \( L_4, \) \( L_5, \) and the “common-mode” recycling cavity length, \( L_{PRC}. \) Similarly, the dark-fringe condition at the beamsplitter exit port is achieved by setting the phase between the carrier e-fields coming from the in-line/offline recycling cavity arms to an odd multiple of \( \pi, \) by microscopically adjusting \( L_{asy} \equiv (L_2 - L_3)/2 \) via “differential-mode” corrections to \( L_2 \) and \( L_3. \)

Unlike the procedure discussed in [33], we do not choose any particular spatial mode (such as the lowest order, TEM\(_{00}\) mode) of the interferometer e-fields for calculating these phases. Rather, we use the entire e-fields, for two reasons: first, for the dark-fringe condition, one wishes to minimize the total dark-port power emerging from the beamsplitter exit port, all of which contributes to the shot noise; second, for the carrier resonance conditions, coupling between modes brings power back into the TEM\(_{00}\) mode from higher modes, so that bringing the total field (i.e., the “perturbed interferometer mode”) to resonance also maximizes the TEM\(_{00}\) power used in calculations of the GW-signal. In any case, we have observed that these two methods (i.e., with or without spatial mode selection for resonance-finding) typically give very similar results.

2. Sideband Frequency Fine-Tuning

The RF-sidebands must also satisfy important conditions, specifically resonance in the PRC, and near-antiresonance in the FP arm cavities. These conditions are affected by the optics, and the sidebands must also therefore be tuned for optimum performance.

As a first approximation, the cavity lengths and RF-modulation frequency are initialized through an analytic evaluation designed to simultaneously optimize carrier and RF-sideband performance. But finer adjustments are required for the PRC resonance condition, and since the cavity lengths are already fixed by the resonance requirements of the carrier beam (the carrier and sideband e-fields are relaxed in separate code executions, with the carrier first), the free parameter which remains to be adjusted is the sideband frequency. In a manner analogous to that specified for cavity length changes, sideband frequency adjustments for resonance are periodically computed as: \( \Delta \nu_{SB} = \Phi \cdot (c/\alpha L_{PRC}); \) though somewhat smaller frequency changes are actually made in practice, during each adjustment, to lessen the disturbances to the e-field relaxation process. The cumulative frequency change for typical runs is small (usually a few hundred Hz or less), but necessary to achieve sideband PRC resonance.

Lastly, we note that the overall results of runs for the two different RF-sidebands (i.e., the “upper” and “lower” sidebands, \( \pm \nu_{SB} = \nu_{car} \pm \nu_{mod} \) are usually fairly similar; we generally perform computations for only one sideband, and assume mirror-image results for the other.

3. Recycling Mirror Reflectivity Optimization

The level of power buildup in the interferometer, and hence its GW-response, depends upon the reflectivities of the mirrors, and is constrained by the mirror losses. In LIGO, the reflectivities of most of the mirrors have been predetermined according a variety of auxiliary physical requirements (e.g., [24]). But the principal criterion for specifying the reflectivity of the power recycling mirror \( R_1 \) is the maximization of the gain that is achievable with power recycling. The power recycling gain, in turn, depends critically upon the losses experienced in the imperfect interferometer, which cannot be precisely determined via analytical estimates. \( R_1 \) is therefore a parameter that can be optimized by the program during each run\(^3\), in order to determine the best achievable recycling gain, given the particular optical imperfections being studied. The reflectivity of a real mirror, of course, is not something that can ordinarily be adjusted on the fly; rather, the results of this optimization routine in our simulation program have been used to help determine appropriate “design values” of \( R_1 \), for power recycling mirrors that are procured by LIGO.

For interferometer losses that are small, it can be shown (e.g., [33]) that the optimal choice for the recycling mirror transmission is \( T_{1,\text{optim}} = A_{IFO} \), where \( A_{IFO} \) is the effective total loss in the full interferometer (including the loss in the power recycling mirror). The optimal reflectivity is therefore given by: 
\[
R_{1,\text{optim}} = 1 - T_{1,\text{optim}} = 1 - (A_1 - A_{IFO}).
\]

The value of \( A_{IFO} \), and thus of \( R_{1,\text{optim}}, \) will be strongly affected by optical deformations.

The quantity used for this optimization procedure is the real part of the integrated overlap between the immediate (“prompt”) reflection of the laser excitation beam from the AR-side of the power recycling mirror, with the complete, composite e-field that is ultimately reflected back from the interferometer. (The imaginary part of this overlap integral is related to interferometer resonance, and will essentially be zero if the cavity lengths are properly adjusted for carrier resonance in the interferometer, as per Sec. III C 1). The real part of the overlap constitutes an error signal\(^4\) that can be driven to zero.

\(^3\) \( R_1 \) optimization is performed specifically during carrier runs, since it is the carrier which requires a high PRC gain; the sidebands do not benefit from long PRC storage times, and would in fact suffer less degradation with a PRC storage time of zero, and immediate ejection at the beamsplitter exit port.

\(^4\) This error signal is only precisely valid when the power recycling mirror obeys the “lossless-mirror” Stokes condition, \( (\tau_{refl}/\tau_{AR}) = -(t/t^* ) \) (cf. Sec. III A). Our program automatically obeys the amplitude part of the condition (i.e., \( |\tau_{refl}| = \)
by changes to \( R_1 \). The magnitude of each change in the optimization process will be proportional to this error signal, though the proportionality constant (which we have selected through empirical tests) is not crucial, as long as it is large enough to achieve rapid optimization (and close convergence to the optimal value), while being small enough to ensure a stable optimization process.

This nulling of \( R < \hat{E}_{\text{prompt}}\hat{E}_{\text{total reflection}} > \) is equivalent to minimizing the total (“interferometer mode”) power that is reflected from the system, thus maximizing the power that is dissipated (and hence, that is circulating) inside of it. In practice, LIGO has chosen an initial value for \( R_1 \) that is slightly below such an “optimized” value, both to hedge (on the safer side) against uncertainties (and gradual increases with time) of the effective interferometer losses, and to provide some reflected light for length control signals. For the specific runs to be presented in this paper, the recycling mirror reflectivity has been driven all the way to \( R_1,\text{optim} \); though runs can also be performed with our program in which \( R_1 \) is held to any particular fixed value that may be preferred due to practical considerations.

4. Schnupp Length Asymmetry Optimization

Also noted in Sec. II was the incorporation of a macroscopic length asymmetry (~few tens of cm) between the inline and offline paths of the PRC, in order to maximally channel sideband power out through the beamsplitter exit port, for use as a local oscillator in the heterodyne GW-signal detection scheme. The maximization of this local oscillator light requires a careful balance between extracting sideband light from the interferometer promptly, before significant power is wasted due to mirror losses; yet leaving the sidebands in the interferometer long enough (i.e., for enough round-trip bounces) so that they can take full advantage of the broadband amplification provided by power recycling.

The total phase between the inline and offline RF-sideband fields, when they meet at the beamsplitter, is analytically given as: \( \Phi_{\text{asymm}} = 2 \times [-2 \text{ } k_{\text{mod}} \text{ } L_{\text{asymm}}] \), \( |r_{\text{BS}}| \) for the recycling mirror; and though the phase part of it is not mandated, we impose it for virtually all of our runs, including those described below in Sec. IV.\(^5\) Eq. (4) is actually a simplification of the proper formula; the simulation program uses a more complete expression \(^{24}\), which does not assume balance between the two paths or a 50%-50% beamsplitter, and which the program must solve numerically.

\[
\hat{M}\{E_{\text{las}} e^{i\omega t}\} \equiv E_{\text{las}} e^{i\omega t+\Gamma \sin \Sigma t} = E_{\text{las}} e^{i\omega t} \cdot \sum_{n=-\infty}^{\infty} J_n(\Gamma) e^{i(n\Sigma t)}
\]

\[
\approx E_{\text{las}} \cdot \{J_0(\Gamma) e^{i\omega t} + J_1(\Gamma) e^{i(\omega + \Sigma) t} + J_{-1}(\Gamma) e^{i(\omega - \Sigma) t}\},
\]

where: \( E_{\text{las}} \) is the (constant) laser field amplitude, \( \omega \equiv 2\pi \nu_{\text{carr}} \) is the carrier angular frequency, \( \Sigma \equiv 2\pi \nu_{\text{mod}} \) is the modulation angular frequency, \( \Gamma \) is the modulation

\[\text{arg}(\Phi_{\text{asymmetric}}) = 2\sqrt{R_1} \text{ } R_{\text{arm avg}} \text{ } R_{\text{BS}}, \]  

where \( R_{\text{arm avg}} \) is the average reflectivity experienced by the beam along its inline and offline paths (including all Fabry-Perot arm cavity effects), and where \( R_{\text{BS}} \approx T_{\text{BS}} \) for the beamsplitter has been assumed here.\(^5\)

The reflectivity values to be used in Eq. (7) (or in its generalized version for an unbalanced interferometer) cannot be analytically determined before the program is run, since \( R_{\text{arm avg}} \) and \( R_{\text{BS}} \) depend upon the losses due to optical imperfections, and \( R_1 \) is a variable that is optimized during the carrier run. The simulation program therefore uses the resonating powers at various interferometer locations (during the RF-sideband run) to determine “effective” reflectivities for use in Eq. (7), to estimate the optimal choice of \( \Phi_{\text{asymmetric}} \). Periodic changes to \( L_{\text{asymmetric}} \) (totaling a few cm, cumulatively) are performed during the sideband field relaxation iterations, until \( \Phi_{\text{asymmetric}} \) is brought to this desired value. These macroscopic length changes are applied antisymmetrically to the inline and offline paths, and in integral multiples of \( \lambda_{\text{carr}}/4 \), in order to avoid (as much as possible) any disruption to the carrier resonance/dark-fringe conditions.\(^6\) Lastly, we note that the program considers only the TEM\(_{00}\) components of the e-fields while evaluating \( \Phi_{\text{asymmetric}} \) and Eq. (4) (as opposed to considering the total e-fields for optimization, cf. Sec. III C\(^3\)), since this extracted RF-sideband light is to be used directly for generating the GW-signal.

5. Sideband Modulation Depth Optimization

The sideband fields are generated via electro-optic, radio frequency modulation of the initial, carrier-frequency laser field. This process is represented mathematically as follows:

The program also uses measured values of \( \Phi_{\text{asymmetric}} \), not this analytical one, in its \( L_{\text{asymmetric}} \)-optimization adjustments.\(^6\) We typically follow the initial carrier/sideband 2-run set with an additional carrier/sideband 2-run set, with no changes to \( L_{\text{asymmetric}} \) in the latter set, to verify that the proper conditions for the carrier have not been disturbed.
depth, \( J_n(\Gamma) = (-1)^n J_{-n}(\Gamma) \) is the Bessel function (with integer order \( n \)) as a function of \( \Gamma \), and where we have dropped the higher-order terms in the series because \( \Gamma \) will be kept small enough for them to be unimportant.

Specifying the modulation depth is a matter of dividing up the total available laser power between the carrier and its RF-sidebands. The carrier and sidebands both contribute to the GW-signal, and due to their (unavoidable for the carrier, and deliberate for the sideband) ejection at the beamsplitter exit port, they both contribute to the shot noise that competes with this signal. The optimal modulation depth is determined via maximization of the overall ratio of GW-signal to shot noise. This ratio (given explicitly in Sec. III D below) is dependent upon the buildup of the steady-state e-fields, and hence upon the specific interferometer deformations which exist; for a real LIGO detector, in particular, the best modulation depth (or the “best achievable” \( \Gamma \), given technical limitations) to impose on the carrier laser beam can only be precisely determined once the differing performances of the carrier and sideband fields in the interferometer are known.

In our modeling work, the modulation depth optimization is performed as a post-processing step, after the simulation runs for the carrier and sideband e-fields have been completed. Besides the straightforward optimization of the shot-noise-limited GW-sensitivity (see Eq. (11) below) with respect to \( \Gamma \), we are also careful to make sure of two things: that \( \Gamma \) is small enough so that the higher-order modulation terms are indeed unimportant as far as their potential buildup in the LIGO interferometer is concerned; and that the total amount of power (carrier plus sidebands) falling upon the GW-signal detection photodiode at the beamsplitter exit port is not excessively large.

6. Mirror Tilt Removal

The final optimization procedure discussed here is a pre-processing step that is performed upon the mirror profile maps before they are read in during execution of the simulation program. Specifically, it is the removal of any overall tilts due to mirror surface variations or substrate inhomogeneities. A “substrate tilt” is merely a mirror thickness wedge, which can be countered in a real interferometer by a small change to the reference axis of a cavity.

For irregular mirror profiles, which do not possess a unique definition of tilt, we choose the most useful definition for a LIGO interferometer: the tilt that is experienced (or “weighted”) by a Gaussian-profile incident beam. To first order, the effect of mirror tilts about the two axes perpendicular to a (predominantly Hermite-Gaussian TEM\(_{00}\)) incident beam is to generate power in the TEM\(_{01}\) and TEM\(_{01}\) modes, in a reflection from that mirror.20 These modes will have imaginary amplitudes (assuming a real incident beam amplitude) if the mirror has no overall “piston” displacement. Given a mirror surface/substrate “height” function \( Z(x, y) \), and defining \( M(x, y) \equiv \exp[-2ikZ(x, y)] \) (exponentiating each pixel individually, not the whole matrix), we can remove the beam-weighted tilts from a given mirror deformation map by performing small angular corrections that set \( \exists < \text{TEM}_{10} |M(x, y)|/\text{TEM}_{00} > = \exists < \text{TEM}_{01} |M(x, y)|/\text{TEM}_{00} > = 0 \). This is done after first removing any piston offset from the mirror, via uniform displacements that set \( \exists < \text{TEM}_{00} |M(x, y)|/\text{TEM}_{00} > \) to zero.

An important point about this procedure is that the appropriate beam spot size must be used for the modes in the overlap coefficients given above, in order for the “beam-weighting” of each mirror’s tilt to be correct; but since the beam spot size is different at different locations in the interferometer, one must therefore know which specific mirror a given deformation map will be used for, before its tilt can be properly removed. Also, we note that this tilt-removal process will actually give the full mirror a nonzero tilt, overall; but the center of the mirror (i.e., the part most sampled by the beam) will be essentially flat (on average) with respect to the plane transverse to the beam propagation axis. An example of such a tilt-removed surface deformation map is shown in Fig. 3.

D. Program Output Quantities and the GW-Sensitivity Function

A full set of program runs results in a detailed specification of the final, steady state of the interferometer. Of the large amount of output data available (either directly
or after some post-processing calculations), here are some of the most important quantities that we examine:

- The relaxed powers (for the carrier and RF-sideband fields) at several key locations in the interferometer.
- The complete, relaxed e-fields at all desired locations — available for graphical display, and/or for complete modal decomposition into Hermite-Gaussian TEM modes, which assists in the interpretation of interferometer conditions (e.g., TEM$_{10}$/TEM$_{01}$ power indicates residual tilts, TEM$_{20}$/TEM$_{02}$ power indicates beam mismatch, etc.).
- The carrier contrast defect, which quantifies how well a carrier dark-fringe was achieved at the beamsplitter exit port for the imperfect interferometer. Defining the carrier powers emerging from the relevant beamsplitter ports as $P_{\text{bright}}$ and $P_{\text{dark}}$, the contrast defect is given as:

$$1 - \text{Contrast} \equiv 1 - C = 1 - \frac{P_{\text{bright}} - P_{\text{dark}}}{P_{\text{bright}} + P_{\text{dark}}} \approx \frac{2P_{\text{dark}}}{P_{\text{bright}}}$$

A large contrast defect implies a substantial carrier power loss at the beamsplitter (and thus a broadband power loss in the system), as well as a large carrier contribution to the shot noise at the signal port photodetector, and an excess of raw power falling on that photodetector.

- The optimized interferometer parameters, as described above in Sec. IIIC. In addition to their role in optimizing the performance of the simulated interferometer, the computed values of these parameters often have intrinsic importance in terms of LIGO design considerations.

- The GW-strain-equivalent shot noise spectral density, $\hat{h}_{\text{SN}}(f)$, of a single LIGO interferometer. This crucial output function allows us to directly evaluate the sensitivity of the LIGO detector to astrophysical sources of gravitational waves, given interferometers with realistic optical deformations (and a realistic heterodyne GW-detection scheme).

Consider Fig. 4, which shows the three most significant noise sources for the first-generation LIGO interferometers: seismic, thermal, and shot noise. They are plotted versus GW-frequency $f$, as spectral densities expressed in terms of the gravitational-wave Fourier amplitudes, $\hat{h}(f)$, that would induce equivalent signals in a LIGO interferometer. These particular curves are the first-generation LIGO requirements for the maximum contributions that would be acceptable from each of these three main noise categories$^7$. The total noise envelope can be obtained from these individual curves by adding them together in quadrature (i.e., incoherent addition of uncorrelated noise is assumed).

Of these contributions, seismic and thermal noise are random forces which will push the LIGO mirrors around in imitation of GW’s. Photon shot noise, on the other hand, is a form of sensing noise, representing the quantum mechanical limit on the accuracy to which the mirror positions can be measured, given the finite amount of carrier power resonating in the FP arm cavities, and the finite amount of local oscillator sideband light available at the signal port. While the quality of interferometer optics has little effect upon the level of random force noise contributions (other than radiation pressure noise, which should be unimportant for the first-generation LIGO interferometers), the quality of the optics does have a direct impact upon the sensing noise limitation to LIGO’s GW-sensitivity. The presence of imperfect optics not only reduces the amount of circulating interferometer power available for sensing mirror positions (and thus GW-induced mirror motions), but also increases the amount of unwanted power at the beamsplitter exit port — such as carrier contrast defect power, and RF-sideband power in non-TEM$_{00}$ modes — which contribute to the shot noise, but not to the GW-signal. We therefore focus upon the shot-noise-limited region of the LIGO noise envelope in evaluating the effects of optical imperfections. For each set of output results, $\hat{h}_{\text{SN}}(f)$ can be computed, and compared either to the first-generation LIGO requirements (see Sec. IV below), or to astrophysical predictions (see Sec. V), in order to determine the effects of optical deformations upon LIGO’s ability to detect GW’s of reasonable, anticipated strengths.

A full derivation of the formula for $\hat{h}_{\text{SN}}(f)$ is given elsewhere$^{24}$. Here we present the resulting expression, in terms of the relevant output data from our simulation program. The shot noise sensitivity limit is expressed in

---

$^7$ Note that the thermal noise curve in Fig. 4 actually represents an approximate conglomeration of mirror internal vibration noise, suspension pendulum noise, and other technical noise sources.
and assume the interferometer performance to be fairly

\[ h^{TT}(f) = \sqrt{2} h \cos(2\pi ft + \phi_0) \]  

(10)

which is incident upon a single interferometric detector

\[ \tilde{h}_{SN}(f) \equiv \left( \frac{\text{Signal}(f)}{\text{Noise}} \right)^{-1} = \left( \sqrt{\eta \nu_carr P_{\text{las}} / h_{pl}} \right) J_0(\Gamma) J_1(\Gamma) \frac{4\sqrt{2\pi\tau_s} \cdot r_{FP \text{ arm back}} \cdot t_{FP \text{ arm input}}}{\sqrt{1 + (4\pi\tau_s)^2}} \times \sqrt{P^{00}_{SB, \text{exit port}} \cdot r_{BS} \cdot \sqrt{P^{00}_{carr, \text{inline FP-arm}} + r_{BS} \cdot \sqrt{P^{00}_{carr, \text{offline FP-arm}}}}} \]  

(11)

where: \( P_{\text{las}} \) is the total excitation laser power (before radio frequency modulation) in Watts, \( h_{pl} \) is Planck’s constant, \( \eta \) is the carrier frequency, \( \nu_carr \) is the quantum efficiency of the photodetector at the signal port, \( J_0(\Gamma) \) and \( J_1(\Gamma) \) represent the division of laser power between the carrier and either one of its RF-sidebands (cf. Eq. 5), and terms like “\( P_{carr, \text{exit port}}^{00} \)” etc., represent the dimensionless relaxed power value (in either the TEM\(_{00}\) mode or the total in all modes, as indicated) that is reported by the numerical simulation code for a carrier or sideband e-field in the indicated interferometer location. These “dimensionless” power values for the simulated e-fields are all normalized in the program to an input carrier/sideband laser beam power of 1 Watt. Note that we have not included both the upper and lower sidebands separately here; in many cases it is sufficient to plug the simulation results for either of them into Eq. 11, and assume the interferometer performance to be fairly symmetrical about the carrier frequency for these RF-modulated fields.

Some remaining quantities in Eq. (11) (e.g., \( r_{FP \text{ arm back}} \), \( t_{BS} \)) are estimated mirror (amplitude) reflection and transmission coefficients for the path that any GW-induced signal fields would take from the FP-arms (where they are physically generated) through the interferometer, until they exit at the beamsplitter signal port. Finally, the quantity \( \tau_s \) is the effective storage time of the GW-induced signal fields in the realistically-deformed FP arm cavities. Since the explicit simulation of the buildup of these signal e-fields in the arm cavities would involve an additional set of e-field relaxation procedures that are not performed by the program, the effective storage time of these GW-induced e-fields in the (imperfect) arm cavities must be approximated, as follows:

\[ \tau_s \approx \frac{L}{c} \cdot \frac{1}{\sqrt{P_{carr, \text{Recyc. Cav.}}^{00}}} \cdot \frac{1}{2} \left\{ \frac{\sqrt{P_{carr, \text{inline FP-arm}}^{00}}}{t_{BS} \cdot t_{FP \text{ arm input}}} + \frac{\sqrt{P_{carr, \text{offline FP-arm}}^{00}}}{r_{BS} \cdot t_{FP \text{ arm input}}} \right\} \]  

(12)

where \( L \) is the length of the FP arm cavities, \( c \) is the speed of light, and a properly-weighted average has been performed over the results for the inline and offline FP-arms. We note that the dependence of \( \tilde{h}_{SN}(f) \) upon the GW-frequency is contained within the expression \( \sqrt{1 + (4\pi\tau_s)^2} \), so that it has two basic regimes separated by the “knee” or pole frequency, \( f_{pole} = 1/(4\pi\tau_s) \), as follows:

\[ \tilde{h}_{SN}(f) \big|_{f \ll f_{pole}} \approx \text{constant} \quad \tilde{h}_{SN}(f) \big|_{f \gg f_{pole}} \propto f \]  

(13)

These two regimes are evident in the plot of \( \tilde{h}_{SN}(f) \) that is included in Fig. 4.

### IV. A SELECTION OF RESULTS OBTAINED WITH THE SIMULATION PROGRAM

In this section, we will demonstrate the code with a set of runs incorporating measurement maps made from very high quality mirrors. The only interferometer imperfec-
tions included are these mirror deformation maps, the finite sizes of the mirrors, and a small amount of pure loss specified for each mirror. The pre-specified mirror loss values represent absorption in the mirror substrates and coatings, as well as high-angle scattering due to roughness finer than the resolution of our grids (for which most of the scattered power is lost beyond the finite mirror apertures, particularly in the long FP-arms). For each individual run in this set, the simulated interferometer is completely optimized in the sense of Sec. III C. The program input parameters common to all runs are those listed in Table I.

First, however, we note that many tests of the simulation program have been performed [37] to ensure its validity as a realistic model of a LIGO interferometer. These tests include: comparisons with numerical simulation results in the literature [16, 38], to verify basic operations like beam propagation and diffractive loss from finite mirrors; comparisons against modal analysis methods [3] for simple interferometer imperfections, such as mirror tilts; and comparisons against analytical methods [59] for computing the effects of Zernike polynomial [10] mirror surface deformations. We have implemented anti-aliasing and energy-conservation procedures (cf. Sec. III A) to make sure that the physics of the interferometer is being properly simulated; and we have performed numerous “common-sense” tests to check that the program’s output results not only make physical sense, but also reasonably reflect the types of interferometer imperfections being modeled. Finally, direct comparisons of our simulation results with experimental measurements for large-scale interferometers are now becoming possible (e.g., [21, 22]), and have thus begun to provide useful mutual feedback for both experimental and modeling efforts.

A. The Realistically-Deformed Mirror Maps Used for these Runs

The mirror maps used in these simulation runs have been derived from two measurements of real optical components: the first one, obtained by LIGO from Hughes-Danbury Optical Systems, is a phase map of the reflection from the polished surface of the “Calflat” reference flat mirror used by the AXAF program (e.g., [11]) for the calibration of their extremely smooth, high-resolution conical mirrors; and the second one is a transmission phase map of a trial LIGO mirror substrate obtained from Corning. Both of these measurements were of uncoated, fused silica substrates. Measurement maps of fully-coated mirrors were not available for the runs in this study.

Each of these near-LIGO-quality mirror maps (one surface reflection map and one substrate transmission map) were extrapolated into an array of many maps, so that we could place deformed surfaces and substrates upon all of the interferometer mirrors simultaneously. One of us (Y. H.) used a three-step process for creating new mirror deformation maps: first, the surface (or substrate) source map was Fourier transformed into spatial-frequency space; then, the relative phases of its Fourier components were randomized; and finally, the application of an inverse Fourier transform created a new mirror with randomized features, but with the same power spectrum of deformations. This process was performed many times, to create 15 new surface maps out of the initial Calflat map, and 7 new substrate maps out of the initial Corning map. Test runs with various groupings of these mirrors have demonstrated to us that different members of a family of randomized mirrors have (as expected) very similar characteristics to one another in terms of their effects upon interferometer performance, and that swapping one for another has little effect upon the output results of the simulation program.

Further preparation steps for the mirror maps were taken to adapt them to the appropriate grid parameters and mirror aperture dimensions (cf. Table I), and also to supply related deformation maps for the beamsplitter, given its 45° tilt angle and consequently elliptical apertures; full details of all such preparations are given in [24]. Each of the resulting mirror maps were then tilt-removed (using appropriate beam spot sizes) with respect to normal-incidence laser beams, as per Sec. III C. An example of one of these surface deformation maps has been shown in Fig. 3.

As a final step, it was necessary to create several families of surface maps with different levels of deformations, in order to help the LIGO Project evaluate a range of mirror polishing specifications for the procurement of the core optics, as well as to make room for the not-fully-determined effects of mirror coating deformations. To accomplish this, the Calflat-derived surface deformation height maps were uniformly multiplied by scale factors to generate the new families. The original family of surfaces — which possess RMS deformations of about 0.6 nm when samples over their central 8 cm diameters — has been labeled “λ/1800” (with λ = λ_{YAG} = 1.064 μm). The scaled-up families are labeled λ/1200, λ/800, and λ/400, respectively. The mirror substrate maps (possessing RMS deformations of about 1.2 nm over their central 8 cm diameters), however, were considered likely to represent the best quality of fused silica substrates obtainable for the first-generation LIGO interferometers [42], and were not re-scaled; all of the deformed-mirror runs discussed below were done with this same family of substrate maps. Note that a direct comparison of the substrate RMS value with that of the surfaces is not informative, since the substrates are typically sampled less frequently by the electric fields (particularly for the carrier light resonating in

8 Although this process does not preserve coherent structures — e.g., spikes, etc., which might be systematic effects of mirror fabrication — such structures were manually added in by us for other tests, enabling us to place limits on their significance.
the FP arm cavities), and are therefore less significant (for the carrier fields, at least) in their effects.

Given these families of mirror deformation maps, it becomes possible to examine the overall performance of a coupled-cavity LIGO interferometer in the presence of “realistic” mirror deformations, to comprehensively estimate its true capabilities.

B. The Results of the Runs

This study contains results from five separate simulation runs: one run with perfectly smooth mirror substrates and surfaces, and four runs with: (i) deformed substrate maps for all of the mirrors, plus, (ii) deformed surface maps for all mirrors from, respectively, the λ/1800, λ/1200, λ/800, or λ/400 families. For all cases other than the “perfect mirrors” run, the transmission and (reflective-side) reflection maps for each 2-port mirror were constructed from one surface phase map and one substrate phase map; the program then derives the AR-side reflection map from the other maps via energy conservation (i.e., the lossless-mirror Stokes condition, cf. Sec. III A). The beamsplitter’s two reflection and two transmission operators were constructed from one surface map and two substrate maps, with the remaining map derived via its generalized energy conservation formula.

The results are summarized in Table II. Several of the quantities described in Sections III C and III D are included, as well as the true (“absolute”) carrier and sideband power exiting at the beamsplitter signal port. The DC values and pole frequencies of the GW-sensitivity curves calculated for each run are also given, from which one can construct $\tilde{h}_{SN}(f)$ for each case as follows (cf. Eq’s. (11)-(13)):

$$\tilde{h}_{SN}(f) = \tilde{h}_{SN}(0) \cdot \sqrt{1 + (f/f_{pole})^2} \quad \text{(14)}$$

Some of these quantities have also been re-computed in the hypothetical case of an idealized output mode cleaner functioning at the signal port, which would act to strip away all of the non-TEM$_{00}$ light (contributing only to noise) from the exiting beams, while passing all of the TEM$_{00}$ light (contributing all of the signal and some shot noise) through to the output photodetector. A photodetector quantum efficiency of $\eta = 0.8$ was assumed for computing the values of $\tilde{h}_{SN}(f)$, $\Gamma$, etc.

Utilizing this output data, one noteworthy result is that several effects of deformed optics have a quadratic dependence upon the deformation amplitudes$^9$. This is as expected, since power scattered out of a Gaussian beam by mirror roughness scales like the square of the roughness amplitude, even when the deformations are spread over a range of spatial frequencies$^1$. For example, the effects of imperfect mirrors on the FP arm cavity power buildup can be expressed in terms of an equivalent “effective mirror loss” that would analytically reproduce the same amount of circulating FP-arm power. Doing this for each of the runs, we obtain an effective loss function that increases quadratically (versus mirror surface RMS deformation amplitude) from the baseline value of ~50 parts per million of “absorptive” loss that is put in by hand for each mirror$^2$. Similarly, the contrast defect (1−C), which comes from power coupled into non-TEM$_{00}$ beam modes by (longer spatial wavelength) optical imperfections, is also well represented by a quadratic fit, as is shown in Fig. 5. Lastly, the optimized recycling mirror (power) reflectivity, $R_1$, should decrease counterintuitively.

$^9$ We note that the signal-generating sideband power at the beamsplitter exit port does not worsen quadratically with RMS deformation levels; and in fact, Table II shows an increase in exit-port sideband power, as the deformation levels get very large. This behavior is due to great effectiveness of the parameter optimizations (cf. Sec. III C) — specifically that of the Schnupp length asymmetry, which is forced to a larger value for highly deformed mirrors, in order to get the sideband fields out of the degraded interferometer as soon as possible. This type of behavior for the sideband fields in the presence of highly deformed mirrors, though technically correct, may be unduly optimistic when considered in the context of a real LIGO system, in which the many parameters are not as easily adjustable — if at all adjustable — in the experimental system, as they are in the simulation program.
noted), as well as a photodetector quantum efficiency of $\eta = 0.8$. 

The most fundamental requirement to be satisfied is that the shot-noise-limited sensitivity curve for an interferometer with deformed mirrors (cf. Eq. 14) should fall within the bounds of the $GW$-strain-equivalent noise envelope requirement, which is the official total-noise limit set for the (full 4 km baseline) first-generation LIGO interferometers [14]. In Fig. 6 the computed $h_{SN}(f)$ curves for these runs are plotted against the data points that define the LIGO requirement envelope, with all data given in terms of spectral densities. The seismic, thermal, and shot-noise-dominated regimes of the requirement envelope are apparent in the figure, and our $h_{SN}(f)$ curves can be compared to the shot-noise-dominated region. We conclude, once again, that all runs other than the $\lambda/400$ case succeed in meeting the initial-LIGO requirement. The overall conformity of quadratically from its “perfect mirrors” value, since $1 - R_{1,\text{optim}}$ is directly proportional to interferometer losses (cf. Sec. III.C.8), and the dominant losses (contrast defect loss and high-angle scattering in the FP-arms) are both quadratically dependent upon mirror deformation RMS. A fit of $R_{1,\text{optim}}$ vs. RMS [24] does indeed bear out this expected functional form (though not all the way to a quadratic from its “perfect mirrors” value, since 1 - $\eta$ [48] is directly proportional to interferometer losses $1 - \text{optim}$, as demonstrated in Table II and Fig. 5.

### Table II: Output results for the series of interferometer simulation runs performed using realistic deformation maps for the optical surfaces and substrates. A total (pre-modulation) laser input power of 6 Watts is assumed (except where otherwise noted), as well as a photodetector quantum efficiency of $\eta = 0.8$. 

| Quantity                                      | Values For Specified Run                                                                 |
|-----------------------------------------------|------------------------------------------------------------------------------------------|
| Deformed Surfaces (RMS in wavelengths)        |                                                                                          |
| Deformed Substrates (Y/N)                     |                                                                                          |
| Recycling Mirror Reflectivity                 |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Recycling Cavity |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Fabry-Perot Arm Avg. |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Exit Port b    |                                                                                          |
| Total Carrier Power, Exit Port b             |                                                                                          |
| Carrier Contrast Defect, 1 - C               |                                                                                          |
| TEM$\text{m}_0$ 1-Sideband Power, Recycling Cavity |                                                                                          |
| TEM$\text{m}_0$ 1-Sideband Power, Exit Port b |                                                                                          |
| Total 1-Sideband Power, Exit Port b          |                                                                                          |
| GW-Response Pole Frequency, $f_{\text{pole}}$ |                                                                                          |
| Modulation Depth, $\Gamma$                   |                                                                                          |
| Absolute Carrier Exit-Port Power (mW)        |                                                                                          |
| Absolute 2-Sideband Exit-Port Power (mW)     |                                                                                          |
| DC GW-Sensitivity, $h_{\text{SN}(0)}$        |                                                                                          |
| Re-Computed Values Assuming Ideal Output Mode Cleaner: |                                                                                          |
| Modulation Depth, $\Gamma$                   |                                                                                          |
| Absolute Carrier Exit-Port Power (mW)        |                                                                                          |
| Absolute 2-Sideband Exit-Port Power (mW)     |                                                                                          |
| DC GW-Sensitivity, $h_{\text{SN}(0)}$        |                                                                                          |

*Denotes parameter optimized by program, or during post-processing.

*Denotes quantity normalized to 1 Watt of carrier/sideband excitation light power.

---

Table II: Output results for the series of interferometer simulation runs performed using realistic deformation maps for the optical surfaces and substrates. A total (pre-modulation) laser input power of 6 Watts is assumed (except where otherwise noted), as well as a photodetector quantum efficiency of $\eta = 0.8$. 

| Quantity                                      | Values For Specified Run                                                                 |
|-----------------------------------------------|------------------------------------------------------------------------------------------|
| Deformed Surfaces (RMS in wavelengths)        |                                                                                          |
| Deformed Substrates (Y/N)                     |                                                                                          |
| Recycling Mirror Reflectivity                 |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Recycling Cavity |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Fabry-Perot Arm Avg. |                                                                                          |
| TEM$\text{m}_0$ Carrier Power, Exit Port b    |                                                                                          |
| Total Carrier Power, Exit Port b             |                                                                                          |
| Carrier Contrast Defect, 1 - C               |                                                                                          |
| TEM$\text{m}_0$ 1-Sideband Power, Recycling Cavity |                                                                                          |
| TEM$\text{m}_0$ 1-Sideband Power, Exit Port b |                                                                                          |
| Total 1-Sideband Power, Exit Port b          |                                                                                          |
| GW-Response Pole Frequency, $f_{\text{pole}}$ |                                                                                          |
| Modulation Depth, $\Gamma$                   |                                                                                          |
| Absolute Carrier Exit-Port Power (mW)        |                                                                                          |
| Absolute 2-Sideband Exit-Port Power (mW)     |                                                                                          |
| DC GW-Sensitivity, $h_{\text{SN}(0)}$        |                                                                                          |

*Denotes parameter optimized by program, or during post-processing.

*Denotes quantity normalized to 1 Watt of carrier/sideband excitation light power.

are operational. One caveat, however, is that although the contrast defect requirements are met for most of the runs, a large amount of total power (several hundred milliwatts) falls upon the output photodetector in all cases, especially for runs with highly deformed mirrors in which the optimized sideband modulation depth is large, in order to help the signal compete against increased shot noise. If a signal detection apparatus that could handle this large amount of power cannot be supplied, then an output mode cleaner may be needed; Table II shows that an output mode cleaner would greatly reduce the power that the photodetector must accommodate (while also improving the GW-sensitivity by ~15%), as long as it operates closely enough to an “idealized” performance, as described above.

The most fundamental requirement to be satisfied is that the shot-noise-limited sensitivity curve for an interferometer with deformed mirrors (cf. Eq. 14) should fall within the bounds of the $GW$-strain-equivalent noise envelope requirement, which is the official total-noise limit set for the (full 4 km baseline) first-generation LIGO interferometers [14].
Figure 6: Comparison of the shot noise curves, \( \tilde{h}_{SN}(f) \), computed for each of the interferometer simulation runs, against the official GW-strain-equivalent noise envelope requirement for the first-generation LIGO interferometers.

These results indicate that there is a very clear — and very strict, though achievable — quality level for the core optics that must be reached in order for the first-generation LIGO interferometers to achieve their target performances.

V. IMPACT OF OPTICAL DEFORMATIONS UPON LIGO SCIENCE CAPABILITIES

To place the results of our runs into a scientifically relevant perspective, we estimate the effects of optical deformations upon LIGO’s ability to detect gravitational waves from anticipated astrophysical events. Focusing here upon GW-sources that might be detectable by the first-generation LIGO interferometers, and considering cases for which an improvement in the shot noise limit may significantly increase the number of detection events (or enlarge the detectable range of some reasonably well-understood scientific parameter), we arrive at two good candidates for study: periodic GW’s from non-axisymmetric pulsars, and burst GW’s from the coalescence of black hole/black hole (BH/BH) binaries.

A. Non-Axisymmetric Pulsars

To facilitate comparisons against theoretical signal estimates, the output data from the simulation program are used to create representative interferometer noise curves. To that end, we form the quadrature sum (for each run) of three functions: the seismic and thermal noise requirement curves (from Fig. 4), added to the particular shot noise curve, \( \tilde{h}_{SN}(f) \), that is computed (cf. Eq. (11), with data from Table III for each of the simulation runs.

These summed, spectral density noise curves must be converted into useful signal-to-noise expressions. We follow the conventions of Thorne [3], in which each of these total-noise curves (\( \tilde{h}_{SN}(f) \)) is converted to a dimensionless expression (\( h_{3/yr} \)), and is compared to the “characteristic strength” (\( h_c \)) of a GW-source. For periodic sources, the condition \( h_c = h_{3/yr} \) means that after coincidence detection in two identical interferometers for one-third of a year of integration time, a source with strength \( h_c \) can be extracted from the Gaussian noise with a confidence level of 90%.

Averaging over all polarizations and orientations of the source on the sky, and treating the GW-frequency \( f \) (equal to twice the pulsar’s rotation frequency) and phase as known, Eq. 52a of [3] yields:

\[
\frac{h_{3/yr}}{f} \approx 1.7 \cdot \sqrt{5} \cdot 10^{-20} \cdot \frac{h(f)}{\text{Hz}}.
\]

For a pulsar with gravitational ellipticity \( \varepsilon \), a rotation-axis moment of inertia of \( 10^{45} \text{ g cm}^2 \), radiating at frequency \( f \) at a distance \( r \) from the earth, and averaging over orientation angles of the source, we have (from Eq. 55 of [3]):

\[
h_c \frac{f}{r} \approx 7.7 \times 10^{-20} \cdot \varepsilon \left( \frac{10 \text{ kpc}}{r} \right) \left( \frac{f}{1 \text{ kHz}} \right)^2.
\]

In Fig. 7, \( h_{3/yr} \) is plotted for all of the simulation runs\(^{10} \), and displayed with them are two \( h_c \) curves, each representing the locus of possible GW-strengths (plotted versus frequency, up to \( f \approx 2 \text{ kHz} \)) for a pulsar with given \( \varepsilon \) and \( r \). We have chosen \( \varepsilon = 10^{-6} \), which should be below the “breaking strain” of neutron star crusts [18], yet may produce a detectable signal. While this value is too large for millisecond pulsars given typical limits on their rates of GW-induced spin-down [18], it may not be unreasonable for newly-formed pulsars, of which it has been estimated [53] that there may be \( \sim 25 \) such “new” pulsars in the galaxy, or one every \( \sim \text{few kpc} \).

By setting \( h_c = h_{3/yr} \) at the frequency of peak sensitivity (\( f_{\text{peak}} \)) for a given \( h_{3/yr} \) curve, one obtains the rough estimate that in going from the \( \lambda/400 \) case to (or nearly to) the perfect mirrors case, the typical “lookout distance” to which such a pulsar is detectable increases from \( \sim 0.57 \) kpc to 2.0 kpc (while \( f_{\text{peak}} \) increases to \( \sim 200 \text{ Hz} \)).

This improvement roughly increases the expected number of pulsar detections (in a galactic disk distribution) by \( \sim (2.0/0.57)^2 \approx 12 \), and brings the lookout distance for the detection of even a single \( \varepsilon = 10^{-6} \) pulsar
with the first-generation LIGO interferometers to a more reasonable value. Alternatively, for a pulsar at a given distance and GW-emission frequency, it provides a factor of \( \approx (2.0/0.57) \approx 3.5 \) leeway in the smallest detectable value of \( \varepsilon \).

### B. Black Hole/Black Hole Binary Coalescences

A similar formulation is used for the analysis of burst sources. Each quadrature noise sum, \( \bar{h}(f) \), is again computed; and with appropriate angle averaging and assumptions for optimal filtering, we have (from Eq. 34 of [5]):

\[
h_{3/yr}|f| \approx \sqrt{5} \cdot \sqrt{\ln(f/10^{-8} \text{ Hz})} \cdot \sqrt{f \times \bar{h}(f)} . \quad (17)
\]

For bursts, \( h_c = h_{3/yr} \) means that after coincidence detection in two identical interferometers for one-third of a year of observation time, a detection of GW-strength \( h_c \) has a 90% probability of being a real signal, rather than an accidental conspiracy of the Gaussian noise in the detectors. (Coincidence operation between multiple interferometers should theoretically eliminate the false burst signals caused by non-Gaussian, uncorrelated noise [26], which we neglect here.)

We consider a BH/BH binary system with equal component masses, \( M_1 = M_2 = 10M_\odot \), located a distance \( r \) from the earth, and evolving in frequency through the insprial phase until it reaches the onset of the coalescence phase (i.e., merger and ringdown) at \( f_{\text{merger}} \sim 205 \text{ Hz} \cdot (20M_\odot/M_{\text{tot}}) \approx 200 \text{ Hz} \). For these parameters, Eq. 46b of [26] yields (with a cumulative factor of 2 adjustment from factor of 2 corrections to Eq. 29 and 44 of [26]):

\[
h_c|f=f_{\text{peak}} \approx 5.0 \times 10^{-21} \cdot \left(\frac{100 \text{ Mpc}}{r}\right) \cdot \left(\frac{100 \text{ Hz}}{f}\right)^{1/6} . \quad (18)
\]

The proper way to interpret this formula, is that if the frequency of peak detector sensitivity is \( f = f_{\text{peak}} \), then the total integrated insprial signal deposited into the detector (for comparison with \( h_{3/yr} \)) is determined by \( h_c \) evaluated specifically at that \( f_{\text{peak}} \).

In Fig. 8, we have plotted the \( h_{3/yr} \) noise curves (for burst searches) that are obtained for each of the simulation runs, along with two \( h_c \) curves (with arrows showing time evolution), for coalescence events that just manage to skirt the high-detection-confidence threshold of \( h_c = h_{3/yr} \) at \( f_{\text{peak}} \) during the course of their inspirels.

The first conclusion that one may draw from this plot is that LIGO’s sensitivity to these coalescence events (at least during inspirlal) is most strongly limited by the low frequency part of the noise curves, i.e., the seismic and thermal noise limits. Nevertheless, there is a measurable benefit from improving the shot noise limit: judging from the plot, it can be estimated that in going from the \( \lambda/400 \) case to the perfect mirrors case, the lookout distance is increased from \( \sim 125 \text{ Mpc} \) to \( \sim 195 \text{ Mpc} \); or equivalently, it increases the expected number of detectable events by the factor \( \sim (195/125)^3 \approx 3.8 \). The actual rate of BH/BH coalescence events is extremely uncertain (even their existence is uncertain), but good middle-of-the-road values that one could use as benchmarks are the “best estimates” that have been made by Phinney [52], and Narayan et al. [53], which are, respectively: \( \sim 3 \) per year out to 200 Mpc (assuming a Hubble constant of \( H_0 \approx 75 \text{ km s}^{-1} \text{ Mpc}^{-1} \)), and \( \sim 1 \) per year out to 200 Mpc \( \times (100 \text{ km s}^{-1} \text{ Mpc}^{-1}/H_0) \). Thus, the perfect mirrors run appears to put BH/BH binary coalescence events just within the conceivable reach of detection for the first-generation LIGO interferometers. Perhaps even more significantly, improving the shot noise limit could increase LIGO’s sensitivity to the onset of the merger phase of BH/BH binaries with masses like these; GW-emission during the actual merger is still poorly understood, but it may involve the most powerful radiation of detectable energy during the overall coalescence process [54].

We end this section by cautioning that the aforementioned numbers must be considered very rough estimates, given the highly simplified ways in which we have treated the noise curves of real interferometers, the sophisticated data analysis methods needed to extract signals from the noise, and the many inherent uncertainties about the GW-sources themselves. In fact, rather than interpreting the results of this section as indicating how LIGO optics would determine initial-LIGO physics, it would be more appropriate to interpret these results as a demonstration of how initial-LIGO physics places requirements on the optics. The firm scientific conclusion that one can draw, however, is that using optical components of the...

---

FIG. 7: Plots of characteristic GW-signal strength \( h_c \) versus frequency, for pulsars with specified ellipticity and distance from the earth (dashed lines), displayed against the dimensionless noise curves \( h_{3/yr} \) (for periodic searches) that are computed from the output results of the simulation runs (solid lines).
FIG. 8: Plots of characteristic GW-signal strength $h_c$ as a function of the detector’s peak sensitivity frequency, during the inspiral phase of $10 M_\odot$ black hole/black hole binaries (dashed lines), displayed against the dimensionless noise curves $h_{3/yr}$ (for burst searches) that are computed from the output results of the simulation runs (solid lines).

best achievable quality can indeed make a difference in whether or not the first-generation LIGO interferometers have a fighting chance to detect gravitational waves from these promising astrophysical sources.

VI. DISCUSSION AND CONCLUSIONS

We have shown our simulation program to be useful for gaining physical insight into interferometer behavior, and for roughly estimating the effects of optical deformations upon LIGO science capabilities. Due to the highly detailed nature of the model, it has been an effective tool for research and development in the LIGO Project. To date, it has been used to address several important design issues, providing support for technical initiatives such as: (i) Aiding LIGO in its transition from Argon-ion lasers ($\lambda = 514.5$ nm) to Nd:YAG lasers ($\lambda = 1.064 \mu m$) for the main carrier beam [52], and demonstrating that interferometer performance is less sensitive to mirror figure deformations (of a specified physical amplitude) when larger-wavelength laser beams are used; (ii) Assisting in the selection of the Schnupp Length Asymmetry scheme for GW-signal readout over an alternative, external modulation (Mach-Zehnder) scheme [53]; (iii) Helping model the performances (particularly the effects of optical deformations upon interferometer control systems) of the major LIGO prototype interferometers, including the Fixed Mass Interferometer (FMI) [54], and the Phase Noise Interferometer (PNI) [56] at MIT, and the 40-meter interferometer at Caltech [57]; (iv) Providing assistance in the selection of optical parameters for the long-baseline LIGO interferometers, such as beam spot sizes, mirror curvatures, and aperture sizes (particularly the perspective-foreshortened beamsplitter aperture) [58]; (v) Conducting preliminary studies of the usefulness of an output mode cleaner at the beamsplitter signal port (cf. Sec. IV B); (vi) Simulating the effects of refraction index variations due to thermal lensing (e.g., [59]) on interferometer performance, resulting in preliminary estimates of $\sim 15\%$ degradation to $h_{SN}(f)$ from mirror coating absorption values of 0.6 ppm, or (equivalently) from substrate bulk absorption values (in fused silica) of 5 ppm/cm [60].

Perhaps the most important use of the program has been its involvement in the LIGO “Pathfinder Project” [61], the initiative to set specifications and tolerances for LIGO’s core optical components, and to procure them through a cooperative effort of several vendors and optics metrology groups. Our program has also been used in conjunction with other modeling initiatives at LIGO [11, 12, 62], to create a broad-based interferometer simulation environment involving different algorithmic approaches and physical regimes of interest.

The latest versions of our program continue to be used to address important questions raised by the LIGO Project, such as estimating the performance of advanced-LIGO detectors [23, 31, 57] — i.e., interferometers incorporating Dual Recycling [23], or even Resonant Sideband Extraction [54] — in the presence of optical deformations, and participating in initiatives to set core optical specifications (and to design the control systems) for those advanced detectors [54, 13]. Many related issues will undoubtedly arise in the near future, as advanced interferometer configurations (and increasingly better optics) become available, for which this program can be used as a primary modeling tool for LIGO and its collaborating gravitational wave groups.

Acknowledgments

We are grateful to Jean-Yves Vinet and Patrice Hello of the VIRGO Project for supplying us with the original code that formed the early basis of our work; and to Hughes-Danbury and Corning for their generosity and spirit of research in sharing their data with us. We would like to thank D. Shoemaker and D. Sigg for their help in the early preparation of this manuscript. The development of our simulation program was principally supported by NSF Cooperative Agreement PHY-9210038.
We are grateful to E. D’Ambrosio, P. Fritschel, K. Ganezer, B. Kelis, N. Mavalvala, R. Savage, D. Shoemaker, D. Sigg, K. Sliwa, A. Weinstein, J. Weldon, S. Whitcomb, and H. Yamamoto for their help in the application of our simulation work to these and other LIGO-related issues.

[1] A. Abramovici et al., Science 256, 325 (1992).
[2] A. Giazotto, in First Edoardo Amaldi Conference on Gravitational Wave Experiments, edited by E. Coccia, G. Pizella, and F. Ronca (World Scientific, Singapore, 1995), p. 86.
[3] K. Danzmann, ibid., p. 100.
[4] K. Tsubono, ibid., p. 112.
[5] D. E. McClelland et al., Class. Quant. Grav. 18, No. 19, 4121 (2001); D. G. Blair, J. Munch, D. E. McClelland, and R. J. Sandeman, Australian Consortium for Interferometric Gravitational Astronomy, ARC Project, 1997 (unpublished).
[6] K. S. Thorne, in 300 Years of Gravitation, edited by S. Hawking and W. Israel (Cambridge University Press, Cambridge, England, 1987), p. 330.
[7] B. J. Meers and K. A. Strain, Phys. Rev. D 43, No. 10, 3117 (1991).
[8] J.-Y. Vinet, “Scattering of Light by Random Surfaces”, 1993 (unpublished); J.-Y. Vinet, V. Brisson, and S. Braccini, Phys. Rev. D 54, No. 2, 1276 (1996).
[9] Y. Hefetz, N. Mavalvala, and D. Sigg, J. Opt. Soc. Am. A 14, No. 7, 1597 (1997).
[10] J.-Y. Vinet and P. Hello, J. Mod. Opt. 40, No. 10, 1981 (1993).
[11] D. Sigg and N. Mavalvala, J. Opt. Soc. Am. A 17, No. 9, 1642 (2000).
[12] R. G. Beausoleil and D. Sigg, J. Opt. Soc. Am. A 16, No. 12, 2990 (1999).
[13] B. Caron et al., Astroparticle Phys. 10, 369 (1999).
[14] A. E. Siegman, Lasers (University Science Books, California, 1986).
[15] J.-Y. Vinet, P. Hello, C. N. Man, and A. Brillet, J. Phys. I (Paris) 2, No. 7, 1287 (1992).
[16] H. P. Kortz and H. Weber, Appl. Opt. 20, No. 11, 1936 (1981).
[17] A. J. Tridgell, D. E. McClelland, and C. M. Savage, in Gravitational Astronomy: Instrument Design and Astrophysical Prospects, edited by D. E. McClelland and H.-A. Bachor (World Scientific, Singapore, 1991), p. 222.
[18] D. E. McClelland, C. M. Savage, A. J. Tridgell, and R. Mavaddat, Phys. Rev. D 48, No. 12, 5475 (1993).
[19] P. Hello and J.-Y. Vinet, in Gravitational Astronomy: Instrument Design and Astrophysical Prospects, edited by D. E. McClelland and H.-A. Bachor. (World Scientific, Singapore, 1991), p. 1.
[20] D. E. McClelland et al., Opt. Lett. 24, No. 15, 1014 (1999).
[21] T. Tomaru et al., Appl. Opt.-OT 41, No. 28, 5913 (2002).
[22] T. Tomaru et al., Class. Quant. Grav. 19, No. 7, 2045 (2002).
[23] B. J. Meers Phys. Rev. D 38, No. 8, 2317 (1988).
[24] B. Bochner, Ph.D. thesis, Massachusetts Institute of Technology, 1998; B. Bochner, LIGO Tech. Doc. P980004-00-R, 1998 (unpublished).
[25] B. Bochner, Gen. Rel. Gravit. 35, No. 6, 1029 (2003).
[26] R. E. Vogt et al., “Proposal to the National Science Foundation: The Construction, Operation, and Support-
tors", presented at the Gravitational Wave Data Analysis Workshop (MIT, Cambridge, Massachusetts), 1996 (unpublished).

[50] E. E. Flanagan and S. A. Hughes, Phys. Rev. D 57, No. 8, 4535 (1998).

[51] K. Thorne (personal communication, 1996).

[52] E. S. Phinney, Astrophys. J. 380, L17 (1991).

[53] R. Narayan, T. Piran, and A. Shemi, Astrophys. J. 379, L17 (1991).

[54] A. Abramovici and D. Shoemaker, LIGO Tech. Doc. T950097-00-E, 1995 (unpublished).

[55] D. Shoemaker et al., “Comparison of 2 Fixed Mass Interferometer Testbeds and the Resulting Recommendation for the Initial LIGO Interferometer Design”, 1993 (unpublished).

[56] K. Sliwa and B. Bochner, “FFT model of the PNI”, 1997 (unpublished).

[57] K. Ganezer, LIGO Tech. Doc. G010119-00-Z, 2001 (unpublished).

[58] P. Fritschel, B. Kells (personal communications).

[59] J. Mizuno, Ph.D. thesis, Max-Planck-Institut für Quantenoptik, Garching (1995).

[60] B. Kells and J. Camp, LIGO Tech. Doc. T970097-01-D, 1997 (unpublished).

[61] S. Whitcomb et al., in Proceedings of the TAMA Workshop on Gravitational Wave Detection, edited by K. Tsubono (Universal Academic Press, Tokyo, 1997).

[62] R. Beausoleil et al., J. Opt. Soc. Am. B 20, No. 6, 1247 (2003).

[63] D. Coyne, LIGO Tech. Doc. G010019-00-D, 2001 (unpublished).