Modelling the Performance of Interferometric Gravitational-Wave Detectors with Realistically Imperfect Optics

by

Brett Bochner

B.S. California Institute of Technology (1991)

Submitted to the Department of Physics in partial
fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1998

Author .....
Department of Physics
May, 1998

Certified by .....
Rainer Weiss
Professor of Physics
Thesis Supervisor

Accepted by .....
Professor Thomas J. Greytak
Associate Department Head for Education

Modelling the Performance of Interferometric Gravitational-Wave Detectors with Realistically Imperfect Optics

by

Brett Bochner

Submitted to the Department of Physics on May 13, 1998,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

Abstract

The LIGO project is part of a world-wide effort to detect the influx of Gravitational Waves upon the earth from astrophysical sources, via their interaction with laser beams in interferometric detectors that are designed for extraordinarily high sensitivity. Central to the successful performance of LIGO detectors is the quality of their optical components, and the efficient optimization of interferometer configuration parameters.

To predict LIGO performance with optics possessing realistic imperfections, we have developed a numerical simulation program to compute the steady-state electric fields of a complete, coupled-cavity LIGO interferometer. The program can model a wide variety of deformations, including laser beam mismatch and/or misalignment, finite mirror size, mirror tilts, curvature distortions, mirror surface roughness, and substrate inhomogeneities. Important interferometer parameters are automatically optimized during program execution to achieve the best possible sensitivity for each new set of perturbed mirrors. This thesis includes investigations of two interferometer designs: the initial LIGO system, and an advanced LIGO configuration called Dual Recycling.

For Initial-LIGO simulations, the program models carrier and sideband frequency beams to compute the explicit shot-noise-limited gravitational wave sensitivity of the interferometer. It is demonstrated that optics of exceptional quality (root-mean-square deformations of less than ~1 nm in the central mirror regions) are necessary to meet Initial-LIGO performance requirements, but that they can be feasibly met. It is also shown that improvements in mirror quality can substantially increase LIGO's sensitivity to selected astrophysical sources.

For Dual Recycling, the program models gravitational-wave-induced sidebands over a range of frequencies to demonstrate that the tuned and narrow-banded signal responses predicted for this configuration can be achieved with imperfect optics. Dual Recycling has lower losses at the interferometer signal port than the Initial-LIGO system, though not significantly improved tolerance to mirror roughness deformations in terms of maintaining high signals. Finally, it is shown that "Wavefront Healing", the claim that losses can be reinjected into the system to feed the gravitational wave signals, is successful in theory, but limited in practice for optics which cause large scattering losses.

Thesis Supervisor: Rainer Weiss
Title: Professor of Physics

Acknowledgments

There's no good way to begin giving thanks — it's just too big a job — so why don't I begin at the beginning . . .

I'd like to thank my father, for those long walks on the Coney Island boardwalk when I was young, in which we talked about every scientific idea under the sun, and every idea beyond the sun as well. During those days I learned that anything is possible, that I should always believe in myself, and that a person can only believe what they see with their own eyes, not what they hear with their ears. Without those lessons, and my Dad's constant support, I wouldn't have made it even this far.

Many times my mother and my uncles and aunts and cousins have asked me exactly what it is that I've been doing for a living, and I've done my best to share the excitement of gravitational wave science with them. I don't think I've gotten any converts to the physics community, but one thing I know is that they've always been there to help me out when I needed them, and that they've done a *great* job of keeping me humble when I needed it. To G.E.R.M.S. and Carvel Man and the gang, thanks for being there.

As an undergraduate and as a graduate student, I have been lucky in having some very true friends, people that I know I can always count on. Whether in Pasadena or Cambridge (and sometimes in both places!) it's nice to know that you guys are around. No need to name names, you know who you are. Maybe I'll even get to see all of you more often now that I've finally broken out of Shawshank.

Being a student in Gravitation and Cosmology at MIT is rare privilege, and I've tried to make the most of it, grabbing a little bit of Real Science^{TM} every now and then in between the time spent doing my technical research (i.e., sitting here and typing). I have learned a lot of lessons in this lab, not the least of which is self-reliance. But when that lesson wasn't enough to carry the day, I've had help from lots of kind folks, especially Yaron Hefetz, who took me under his wing when I was new to this place; and David Shoemaker, who has advised me on countless issues, and who has lent me his tact on occasion since I possess none of my own. And of course, I'm grateful to my comrades-in-arms in the Grad Wing (and elsewhere), who have waited this long to kick me out of the place. Keep the faith, guys: there really is a thesis out there for everybody.

And of course, I wish to thank Helen, for all those years of friendship and support. It's been a roller-coaster, but hasn't it been fun?

Have I forgotten anybody? Well maybe one guy, I think his name was Albert. I'd like to thank him too, because without him none of us would be employed right now, and there would be a lot less interesting stuff to think about and work on in the world of science.

Let's not forget what LIGO is all about, folks.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction to Gravitational Waves and interferometric detectors

A fundamental prediction of Einstein's General Theory of Relativity [1], as well as any other causal theory of gravitation, is the existence of *Gravitational Waves*. Gravitational Waves (GW's) carry energy and information away from strongly accelerating, massive systems, propagating away at the speed of light to update the resulting gravitational field structure in the surrounding universe. Gravitational Waves, almost completely unimpeded during propagation through intervening matter [2], are one of the best probes for examining the behavior of very distant and massive astrophysical systems.

The Laser Interferometer Gravitational-wave Observatory (LIGO) is one of a new breed of interferometric detectors of gravitational radiation [3]. LIGO will be a general-purpose observatory, designed to explore the universe in the "Gravitational-Wave band", as well as measuring the properties of the GW's themselves, thus testing gravitational theory in a fundamental way. In order to accomplish this goal, the LIGO Project is in the process of constructing several long-baseline interferometers, highly specialized systems with state-of-the-art optics and control systems. These interferometers must be exceptionally well isolated from all contributing noise sources, and must be held to exacting resonance conditions, in order to detect the extremely weak signals that are expected from even the most powerful astrophysical sources.

A number of LIGO prototypes have been constructed to test and refine various aspects of the final detectors, but it is very difficult to predict the behavior of a complete, LIGO-scale interferometer with *realistically imperfect* optical components, from analytically or experimentally simplified prototypes. It has therefore been necessary to construct a full-scale numerical model of a full-LIGO interferometer with precisely-defined, realistic optics. The thesis that is presented here will document the results of this LIGO simulation research.

1.1 The LIGO scientific mission

1.1.1 Gravitational Waves (GW's) and the principle of detection

According to Einstein's General Theory of Relativity, there is a precise mathematical relationship between the distribution of matter and energy in space, and the "curvature of spacetime" that it generates, and this spacetime curvature in turn embodies the action of

gravitational forces back upon that matter and energy [4]. Spacetime curvature is represented by the *metric tensor*, $g_{\mu\nu}$, which is said to be "flat" in the absence of nontrivial gravitational fields. The metric tensor is a measure of the relativistic "distance" (i.e., "invariant interval") between two points with coordinate separations dx^{μ} (in four dimensions), as follows:

$$Interval \equiv d\tau^2 = g_{\mu\nu} \cdot dx^{\mu} dx^{\nu} , \quad \text{for } \mu, \nu \text{ summed over } (x, y, z, t) \tag{1.1}$$

In Einstein's (linearized) theory, gravitational waves represent a oscillating perturbation to the Minkowski metric, $\eta_{\mu\nu}$, of flat-spacetime. In the so-called "transverse-traceless" (TT) gauge, a GW can be represented as follows [4]:

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{1.2}$$

with:

$$\eta_{\mu\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad h_{\mu\nu} = h\cos(\omega_{GW}\, t + \delta_{GW}) \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & \varepsilon_{+} & \varepsilon_{x} & 0 \\ 0 & \varepsilon_{x} & -\varepsilon_{+} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \tag{1.3}$$

where h is the dimensionless (peak-to-peak) amplitude of the GW (with $h \ll 1$), with the GW angular frequency being given by $\omega_{GW} = 2\pi\nu_{GW}$, and where δ_{GW} is the GW's initial phase in the plane of the detector (i.e., at z=0). This metric represents a GW propagating at the speed of light, c, along the z-axis with wavenumber $k_{GW} = \omega_{GW}/c$. There are two possible polarizations, called "+" ($\varepsilon_{+} = 1, \varepsilon_{x} = 0$) and "x" ($\varepsilon_{+} = 0, \varepsilon_{x} = 1$), which differ by a 45° rotation about the z-axis.

To visualize the effect of the GW upon "freely-falling" (i.e., unconstrained) masses, Figure 1.1 shows the motion of a ring of particles during one period of oscillation as a GW with the "+"-polarization travels perpendicular to the page:

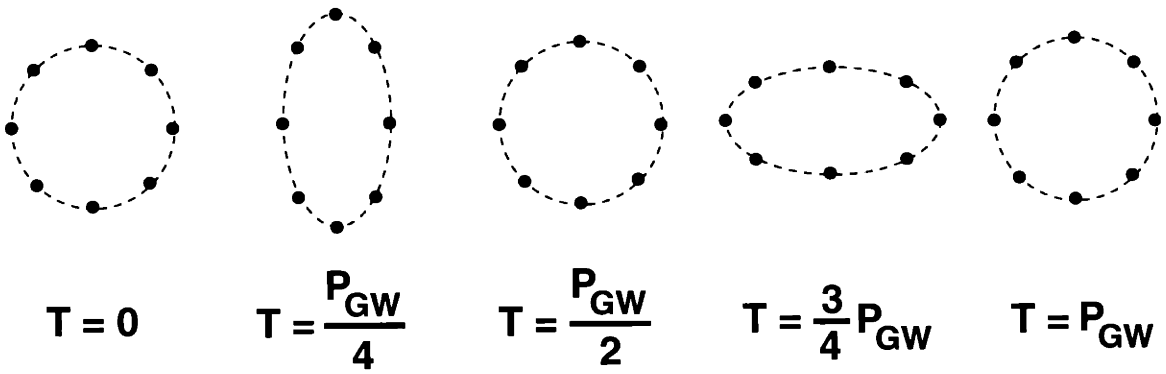

Figure 1.1: Action of a Gravitational Wave upon a ring of freely-falling (inertial) masses.

From this figure, it is apparent that a Michelson interferometer with the proper alignment (such as that shown in Figure 1.2) would be ideal for converting the GW forces into an oscillating output fringe, where the interferometer mirrors — not bolted to an optical table, but hanging as pendula — would act as freely-falling masses above the pendulum frequency of their suspensions. This is the fundamental principle of LIGO GW detection [5], where the additional optics in a LIGO interferometer (as will be discussed in detail in upcoming sections) serve only to amplify this signal or separate it from noise.

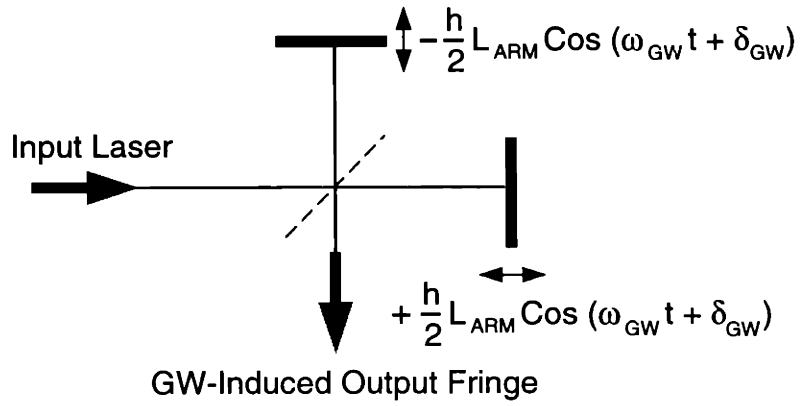

Figure 1.2: Measuring GW-forces with a simple Michelson interferometer.

Figure 1.2 depicts the action of a GW as a force which moves the mirrors back and forth in an opposite fashion, so that the round-trip time for the laser beam along the two paths is different, and thus a differential phase shift exists between the two beams for recombination at the beamsplitter. This phase shift causes output light to emerge from the exit port of the interferometer, which would otherwise have exhibited a dark fringe in the absence of GW's due to perfect destructive interference of the two beams.

This "forces on mirrors" or "phase shift" picture is a valid way to view the physical behavior of the interferometric system [5], but an equally valid perspective is the "GW-induced-sideband" picture, in which the mirrors are considered to be at rest with respect to coordinates that are "comoving" [4] with the GW, and for which the effect of the GW is to create sidebands on the laser light which are generated with opposite sign in the two Michelson arms, and thus emerge from the interferometer exit port to provide the GW-signal. Though these two pictures are completely equivalent (and equally valid), one is sometimes more useful than the other for visualization or calculational purposes. Both viewpoints will be used at appropriate times in this thesis.

Lastly, we note that the weak coupling of GW's to matter (i.e., infinitesimal h) which makes GW-detection so difficult, also has a beneficial effect for astrophysical observation: it makes the universe almost transparent to the waves. GW's experience virtually no absorption, scattering, or dispersion of any kind once they have left their initial "wave-

generation" region [2], and thus provide observers with a view of regions that are very optically dense (sometimes due to high concentrations of matter in the deep gravitational well of a GW-source), such as from the cores of supernovae, or from sources that are extremely distant, such as extragalactic coalescences of Black Hole binaries.

1.1.2 Gravitational-Wave sources and LIGO science goals

The emission of gravitational radiation is similar to that of electromagnetic radiation in principle, though it is so much smaller in practice because of the weakness of the gravitational force compared to electromagnetism ($\sim 10^{-39}$ times weaker for the force between a proton and an electron). Only very massive systems undergoing powerful accelerations, such as cataclysmic astrophysical events, will radiate detectable GW's. Observable gravitational radiation will therefore be emitted by masses undergoing coherent bulk motions, and will pass freely though space all the way from deep in their emission regions to terrestrial GW detectors. This is in contrast to electromagnetic emissions, which are formed from the incoherent sum of radiation from a great many particles, and which typically come to us from the surfaces of stars or from less optically thick regions such as stellar atmospheres and plasmas, and which typically suffer strong absorption, dispersion, or scattering along the way. Furthermore, typical GW-frequencies should be of order the (inverse of the) transit time of the system undergoing coherent bulk motions (i.e., $\nu_{GW} \sim 1/T$), so that $\nu_{GW} < 10^4$ Hz, unlike the very high frequencies of electromagnetic radiation which reflect the timescales of atomic transitions and/or thermal emission from high-temperature objects. The overall result of these differences is that Gravitational Wave observatories are expected to open up a completely different window on the universe and lead to a new revolution in astrophysical understanding, much as was achieved by the introduction of radio and x-ray astronomy earlier in the century, as compared to traditional optical astronomy [2].

Now consider the angular pattern of radiation which will by emitted by a typical source. As monopole radiation is forbidden in electromagnetism because of Gauss' Law and the conservation of charge, monopole radiation is also forbidden in gravitation because of Birkhoff's Theorem[1] [6] and the conservation of mass-energy. Unlike electromagnetism, however, dipole radiation in gravitation (both "electric" and "magnetic" dipole types) are forbidden by (respectively) linear and angular momentum conservation [4].

The energy in GW's should therefore be dominated by quadrupole radiation, which is proportional to the square of the third-derivative of the ("reduced") quadrupole moment. This third derivative is given as [4]:

---

1. The theorem states that, "a spherically-symmetric gravitational field in empty space must be static", even if measured from above a (fixed) quantity of mass-energy undergoing radial pulsations.

$$\dddot{I} \sim \frac{(\text{mass in motion}) \times (\text{system size})^2}{(\text{system transit time})^3} \sim \frac{MR^2}{T^3} \sim \frac{MV^2}{T} \sim \frac{(\text{Nonspherical Energy})}{T} \quad (1.4)$$

where it is seen that only the "nonspherical" part of the accelerating mass-energy contributes to the quadrupolar radiation. The power in a GW is proportional to the square of the GW-strength h times ν_{GW} [4], so taking into account the fact that the gravitational power must decrease as the inverse-square distance from the source in order to conserve energy (i.e., GW-power ~ GW-energy$/4\pi r^2$) we may write h as a function of emitted GW-energy, as follows:

$$h \sim \frac{\sqrt{\text{GW-power}}}{\nu_{GW}} \sim \frac{\sqrt{\text{GW-energy}/4\pi r^2}}{1/T} \sim \frac{T}{2\sqrt{\pi}\, r} \times \dddot{I} \sim \frac{MV^2}{3r} \quad (1.5)$$

By factoring in the appropriate constants (G, c) to make h dimensionless, we get:

$$h \sim \frac{GMV^2}{3c^4 r} \sim \frac{G}{c^4} \cdot \frac{(E^{\text{nonspherical}})}{r} \sim \text{few} \times 10^{-21} \left(\frac{E^{\text{nonspherical}}}{M_{\text{sol}}\, c^2} \right) \left(\frac{10\ \text{Mpc}}{r} \right) \quad (1.6)$$

This formula gives a (very rough) approximation of the GW emission from a "typical" astrophysical source, as well as a rough estimate (neglecting factors of order unity) of the measurable GW-induced strain caused by ~1 solar mass of energy undergoing nonspherical (primarily quadrupolar) motion at a distance from earth similar to that of the Virgo cluster of galaxies. Since even a fraction of a solar mass represents a tremendous amount of energy in coherent bulk motion, only very powerful sources will emit significant energy in the form of GW's. Such systems must be rare, and thus the nearest ones are likely to be extremely far away, requiring LIGO to construct interferometers of unprecedented sensitivity, measuring GW-strains of few $\times 10^{-21}$ or better.

Gravitational Wave sources are grouped into three categories: *periodic* sources (e.g., non-axisymmetric pulsars), *bursts* (e.g., supernovae, coalescing Black Hole binaries), and *stochastic* sources (e.g., primordial GW's from the Big Bang). Signal processing and estimation of signal-to-noise ratios is different for these three categories of sources [2], as will be seen during GW-signal calculations in Section 3.4; but for each of these source types one can compute a "characteristic" GW-strength, h_c, which can be used for comparison with detector sensitivity.

For each potentially important source of GW's, there are significant uncertainties and/or physics limitations which make it difficult to be certain of LIGO's ability to detect them. Pulsars, for example, are known to exist in large numbers, and the GW emission from such objects is fairly easy to compute, assuming a given quadrupole moment ("ellipticity") for the spinning neutron star [2, 7]; but there are stringent upper limits on their (static) ellipticities and on their overall energy emitted via GW's, from, respectively, estimates of the

breaking strain of their crusts, and from limits to the rate of slowdown in spin which they experience [7]. Similarly, the rate of occurrence of supernovae is fairly well known [8], but it is extremely difficult to predict the non-sphericity of the supernova core collapse, and thus the amount of energy that would be radiated away gravitationally [9]. Coalescences of compact-body binaries, however, have a fairly predictable gravitational waveform and radiation strength [9], but in the case of neutron star-neutron star (NS/NS) binaries they may be sufficiently weak, and such binaries may be sufficiently difficult to produce via stellar evolution [10, 11] (thus being rare and far away), so that they lie just below LIGO detectability. In the case of binaries with Black Holes (BH/NS or BH/BH), while their GW-emission will be much stronger (and are thus observable much farther out), their abundance in the universe — and in fact, their very existence — is extremely uncertain.

In the face of these uncertainties, LIGO has been designed with one very concrete goal in mind [12]: the Enhanced-LIGO interferometers[1] are designed to be *very likely* to detect the coalescence of NS/NS binaries, a fairly weak but potentially the most predictable source[2]. In addition, the Initial- and Enhanced-LIGO systems are also designed in order to have a good chance at detecting the other GW-sources mentioned above, as well as being broadband and flexible enough to detect the most important Gravitational Wave source of all: the Unknown.

Figure 1.3 shows the projected LIGO sensitivities for Initial and Enhanced (i.e., advanced-subsystem) detectors, along with theoretical "best estimates" of signal strengths (as a function of GW-frequency) from inspiraling compact binary sources (as they evolve in time), at various distances from the earth. The stippled areas represent a theoretical range of detector sensitivity depending upon source location on the sky and GW-polarization, as well as the signal-to-noise ratios required for detection; the h_{SB} ("sensitivity to bursts") curves represent randomized source direction plus the requirement of high-confidence, correlated detection in all LIGO interferometers, while the h_{rms} curves represent unity signal-to-noise detection for GW's with optimal source direction and GW-polarization. This figure, reproduced from a local LIGO presentation [15], demonstrates the estimated high-likelihood of compact-binary detection with planned LIGO systems.

If LIGO detectors succeed in observing these or other sources, then in addition to being an astrophysical observatory, LIGO would also test aspects of GW's such as their predicted quadrupolar nature, their spin (relativity predicts spin 2), and their propagation speed in vacuum (e.g., by comparing the arrival time of GW's from supernovae with the arrival time of its neutrinos), assumed to be at the speed of light. Information about the

---

1. An "Enhanced" LIGO interferometer refers to a system which integrates a number of advanced subsystems into the Initial-LIGO configuration.
2. The emission of gravitational radiation from a binary pulsar (NS/NS) system has in fact been demonstrated via observations by Hulse and Taylor [13] of PSR 1913+16, which was shown to lose orbital energy (presumably due to GW's) at a rate that matches the predictions of general relativity [14]; this finding has been the first (indirect) experimental proof of the existence of GW's.

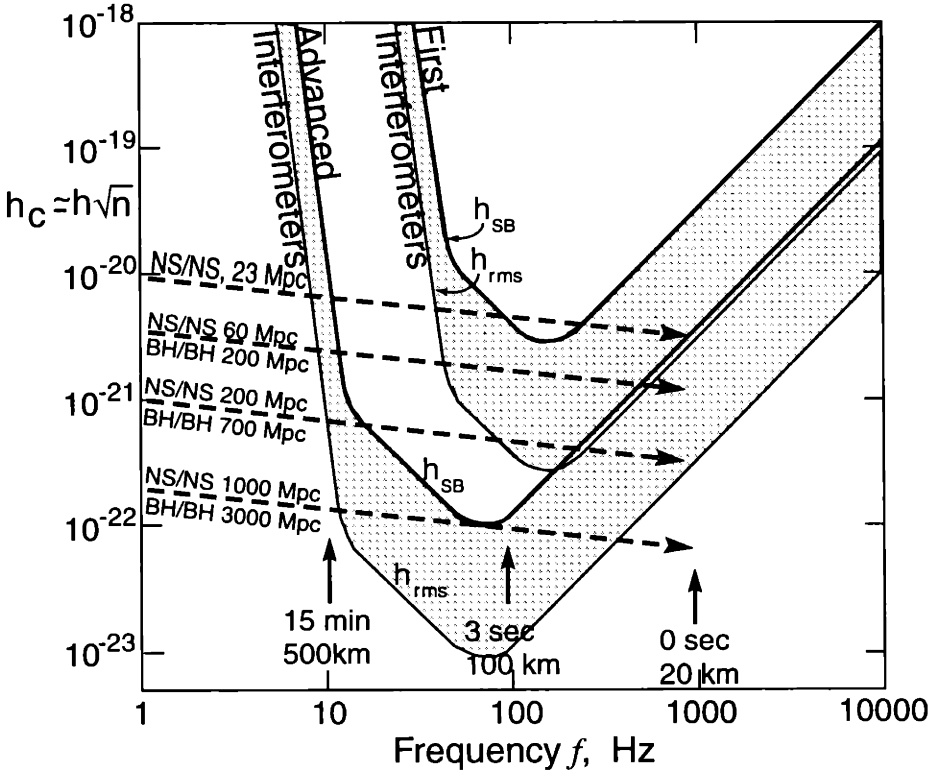

Figure 1.3: Target sensitivities of LIGO interferometers for inspiraling compact binaries.

physics of GW-generation would also be obtained, thus providing our first glimpse into the behavior of highly relativistic systems with strong, nonlinear gravitational fields.

1.2 Interferometer noise sources and the shot-noise sensitivity limit

The projected LIGO noise curves shown above in Figure 1.3 can be seen to possess some structure, including clearly-defined regimes of different functional behaviors. These noise curves actually represent the quadratic (i.e., incoherent) sum of a variety of anticipated noise sources. A plot of the individual noise contributions for the Initial-LIGO detector [15] are shown in Figure 1.4, in which certain "technical" noise sources (e.g., amplifier noise, and several narrow suspension wire resonances) have been omitted.

The "fundamental" noise sources which we consider generally fall into two categories: *sensing* (or *phase*) *noise*, and *random force noise*. In the language of our "forces on mirrors" interpretation of GW-action described in Sec. 1.1.1, we would say that sensing noise affects where we *measure* the interferometer mirrors to be, while random force noise actually *pushes them around*, as GW's would do. Examples of sensing noise are photon shot

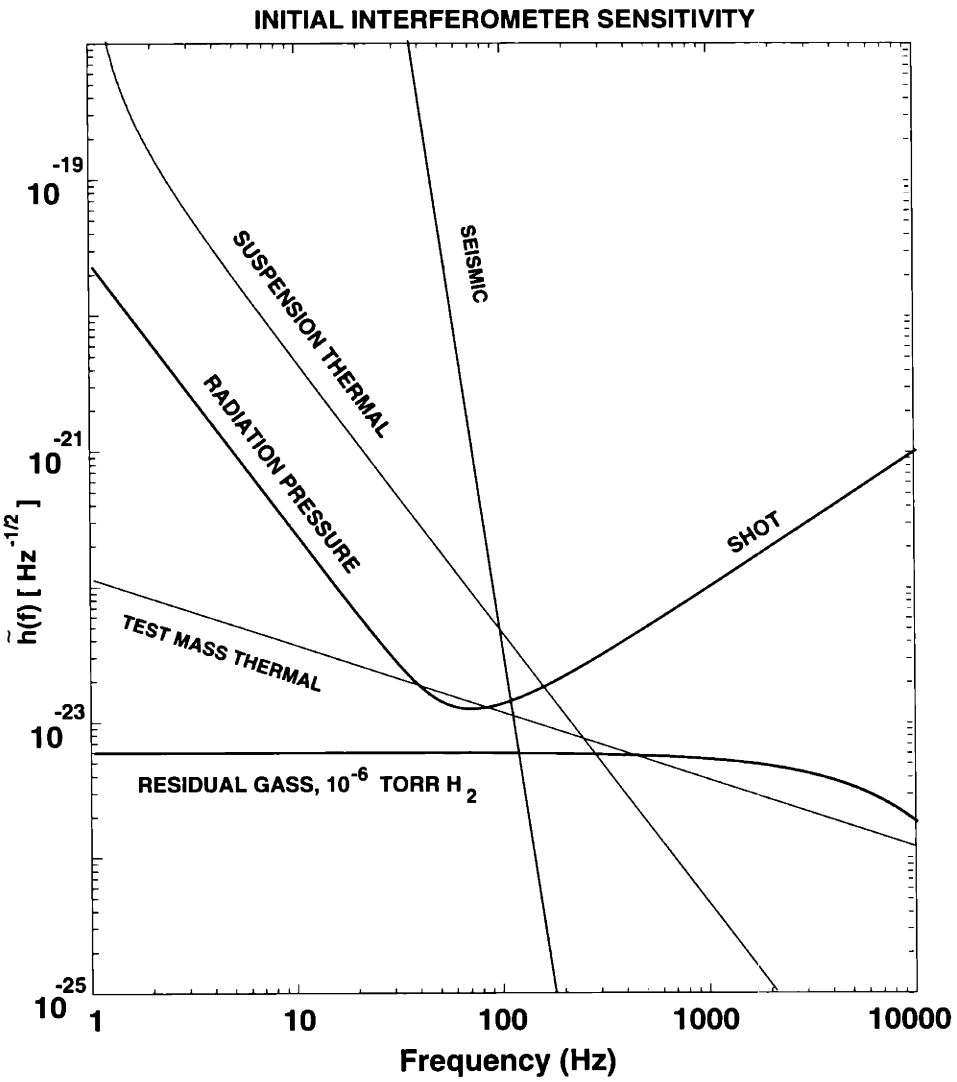

Figure 1.4: A summary of anticipated noise sources for the Initial-LIGO interferometers.

(i.e., counting) noise, phase shifts induced by residual gas in the beamtubes, and stray light pollution. Examples of random force noise are thermal vibrations (in the suspension wires and in internal vibrations of the mirrors), seismic motions, gravity-gradient-induced mirror motions, and radiation pressure fluctuations due to variations in circulating laser power.

The dominant noise sources for the Initial- and Enhanced-LIGO configurations are seismic, thermal, and shot noise (although radiation pressure noise may become important in Enhanced configurations with very large laser input powers, such as ~100 Watts). These

dominant contributions clearly define the overall noise estimates, $h(f)$ [1], shown in Fig. 1.3, and we will restrict our consideration to these noise sources in our study of interferometer behavior.

This thesis focuses on the effects of imperfect optics upon LIGO performance; in particular, imperfect optics reduces the amount of circulating interferometer power available for sensing mirror positions, and also increases the amount of scattered power in high transverse modes that appears at the exit (i.e., signal) port of the Michelson beamsplitter. The former effect reduces the statistics available for photon counting, while the latter provides stray light that adds to the shot noise but not to the GW-signal. The quality of interferometer optics therefore has a direct impact upon the interferometer shot noise curves, while it has little effect upon the level of random force noise contributions such as seismic or thermal noise.

In our interferometer simulation work, therefore, we 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, $\tilde{h}_{SN}(f)$ is computed, and can be compared to LIGO requirements. This shot noise function can be combined with the *expected* levels of seismic and thermal noise in order to represent the overall noise envelope, $\tilde{h}(f)$, of a LIGO interferometer with a particular set of imperfect optics. It can also be converted (such as in Chapter 3) into mathematical forms that are well-suited for comparison with astrophysical predictions, in order to determine the effects of optical deformations upon the capabilities of LIGO to detect gravitational waves of reasonable, anticipated strengths.

One can obtain a simple estimate of the shot-noise-limited sensitivity of a LIGO interferometer via the Heisenberg Uncertainty Principle for photon number and phase [16]:

$$\Delta N \cdot \Delta\phi \sim 1 \tag{1.7}$$

The power in a laser beam is given by $P = N \cdot h_{planck}\nu$, so for a coherent beam with $\Delta N \sim \sqrt{N}$, we have $\Delta N \sim \sqrt{P \cdot \tau_{int}/(h_{planck}\nu)}$ (for signal integration time τ_{int}). A phase shift is related to a length change according to $\Delta\phi = 2k\,\Delta L N_{bounces} = 4\pi\,\Delta L N_{bounces}/\lambda$. We can combine these results to re-express the uncertainty relationship as follows:

$$\Delta L \approx \frac{\lambda}{4\pi N_{bounces}}\sqrt{\frac{h_{planck}\nu}{P \cdot \tau_{int}}} = \frac{1}{4\pi N_{bounces}}\sqrt{\frac{h_{planck}c\lambda}{P \cdot \tau_{int}}} \tag{1.8}$$

With the change in the interferometer arm lengths from GW's given by $\Delta L \sim h \cdot L$ (where L for LIGO arms equals 4 km), and assuming $N_{bounces} \sim 130$ for the stored power in the *Fabry-Perot cavity* (see Fig. 1.5) arms, ~100 W of power encountering the Initial-LIGO beamsplitter, $\tau_{int} \sim 5 \times 10^{-3}$ s (for ν_{GW} ~ few x 100 Hz), and a factor of ~2 in sensitivity roll-off at the observation GW-frequency (say, ~150 Hz) this yields a GW-sensitivity of:

1. We define $\tilde{h}(f) \equiv \tilde{h}(\nu_{GW})$ as the GW-strain needed to produce a signal-to-noise ratio of 1, after a one second signal integration time.

$$h \approx \frac{2}{4\pi N_{bounces}\, L}\sqrt{\frac{h_{planck} c\lambda}{P \cdot \tau_{int}}} \approx \frac{1}{2\pi(4000\text{ m})(130\text{ bounces})} \times \sqrt{\frac{(6.626\times 10^{-34}\text{ J}\cdot\text{s})(3\times 10^{8}\text{ m/s})(1.064x10^{-6}\text{ m})}{100\text{ J/s}\cdot 5\times 10^{-3}\text{ s}}} \approx 2\times 10^{-22} \quad (1.9)$$

Comparing this approximate result with Eq. 1.6 for GW-emission strengths, and given the need for high signal-to-noise ratios (and considering non-isotropic detector sensitivity), we see that a significant amount of a solar mass' worth of energy undergoing nonspherical motion is necessary, for a source emitting GW's at a distance of 10 Mpc from the earth, for it to be within range of detection for an Initial-LIGO interferometer — *also* assuming that it is observed at GW-frequencies where noise sources other than shot noise are unimportant.

More sophisticated calculations of shot-noise-limited sensitivities will be done later on in this thesis to interpret the output results of our numerical interferometer simulations; but this order-of-magnitude calculation is sufficient to show that the detection of gravitational waves by LIGO is a viable prospect, and that reducing the shot noise level with high-quality optics will have a strong impact upon the detectability of GW-sources of reasonable strengths (provided that the levels of, e.g., thermal noise, are also low enough).

1.3 The purpose of this work

1.3.1 The effects of optical imperfections

Consider once more the equivalent phase shift that must be observed in a LIGO interferometer in order to detect GW's. For $\Delta L \sim h \cdot L \sim 8\times 10^{-19}$ m, we require:

$$\Delta\phi\,(\text{unity bandwidth}) = \sqrt{\tau_{int}}\cdot 4\pi\ \Delta L N_{bounces}/\lambda \sim 10^{-10}\ \text{radians}/\sqrt{\text{Hz}} \quad (1.10)$$

Given this extraordinarily strict phase requirement, and the complex interferometric apparatus that must be constructed to measure it, it is not sufficient to assume idealized optical elements; rather, it is necessary to perform a sophisticated evaluation of LIGO performance with "realistic" mirrors, and also to ensure that optics can be procured which are good enough to meet LIGO sensitivity goals.

The estimation of LIGO performance with realistically imperfect mirrors is a challenging task, because of the wide variety of physical effects that must be accounted for in a LIGO interferometer. One consideration, for example, is the finite size of the mirrors compared to the (Gaussian-profile) laser beam; because of the long interferometer arms (to make ΔL large for a given h), the beam spot size is large at various interferometer locations, and non-negligible power (up to ~1-2 parts per million per bounce) falls off the mirror edges, even for perfectly collimated Gaussian beams. Even more significant is the loss due to finite-size mirrors when realistic mirror deformations are taken into account, since

the fine-scale roughness of an imperfectly-polished mirror scatters power away at high angles, leading to very significant power loss (up to $\sim 1/2$ of *all* power dissipation in the interferometer in some cases) in the long Michelson arms. These losses due to mirror imperfections can cause a significant reduction of built-up power in the interferometer arms, thus degrading the shot-noise-limited phase sensitivity of the system and making it a poorer detector of GW's.

Besides the possibility of scattering losses, imperfect mirrors scatter light out of the fundamental ("TEM_{00}") Gaussian mode into higher modes (possessing different spatial profiles) that may *stay in the system*, thus corrupting the circulating laser field. These higher modes may leak out of the exit port of the Michelson beamsplitter where GW-signal detection is performed (see Fig. 1.2), thus increasing the shot noise at the photodetector while providing no additional GW-signal (in addition to acting as another source of loss). Even for very good optical surfaces with small coupling out of the fundamental laser mode, corruption of the circulating laser field can be very significant, because (as will be discussed shortly) the LIGO interferometer is not a simple Michelson interferometer, but rather it is a complex, resonant system of multiply-coupled cavities; and any higher mode that is generated by the optics and which is accidentally resonant in this coupled-cavity system can be significantly amplified, at the expense of the desired, fundamental-mode light.

Mirror deformations can be divided into a few different classes of interest. One class is that of "geometric" deformations, such as tilts or mirror curvature errors (i.e., mismatch of the beam into the interferometer), which must be controlled actively (via feedback systems) or passively (e.g., careful mirror polishing) to a very fine degree for the successful operation of LIGO interferometers. In the language of Hermite-Gaussian laser modes [17], one may say that geometric deformations are well-defined errors which convert TEM_{00} mode power into a very few specific higher modes; tilts, for example, cause the creation of TEM_{10} and TEM_{01} modes, while curvature errors generate power in the "donut" modes, TEM_{20} and TEM_{02}. For solely geometric deformations, therefore, it is often possible to determine interferometric beam behavior with analytical or semi-analytical methods. The analysis of interferometer laser power into this modal basis is discussed in Appendix A.

The situation is more complicated for non-"geometric" deformations such as either random or highly-structured mirror roughness, both of which would be expected to exist in significant amounts for any real-world optic. These deformations scatter a finite amount of power into *virtually all* modes, making an analytical solution nearly impossible, especially for (possibly non-invertible) systems with irreversible power losses such as those due to finite-aperture mirrors. The behavior of these modes can be quite complicated in a coupled-cavity system such as LIGO, since modes of different orders experience completely different ("Guoy") phase shifts during round trips through cavities where the beam experiences significant focusing [17]. Computing the behavior of laser fields circulating in a full-LIGO interferometer that possesses realistically-imperfect optics is therefore an extremely difficult problem; it is one which, we argue, requires a full-scale numerical model to

obtain a solution that one may have confidence in. Creating such an interferometer simulation program, and applying it to questions of expected LIGO performance — for both Initial- and Advanced-LIGO configurations — has been the focus of this thesis research.

1.3.2 A computational tool for complex interferometers

Prior to the creation of a full-scale numerical simulation program, a number of different approaches have been taken to estimate interferometric detector performance for simplified cases. Analytical methods suffice for the consideration of optical defects that can be treated as pure losses, such as absorption[1] within mirror substrates and reflective coatings, and very high spatial frequency surface figure deformations which scatter light power completely out of the interferometer apparatus. For the study of mirror tilts and beam displacements, a matrix model which evaluates the coupling between the first few lowest-order TEM laser modes has been shown to be useful [18], and a matrix model using discrete Hankel transforms exists for problems with axial symmetry [19]. For the consideration of more general optical imperfections, however, the most comprehensive (though computationally intensive) method is the complete, grid-based modelling of the transverse structure of the electric field wavefronts, while using discrete Fast Fourier transforms (FFT's) for the longitudinal propagations of these laser beams in the *paraxial beam* limit [17]. The physical principles and computational algorithms of such a simulation program will be discussed in detail as the subject of Chapter 2.

Grid-based simulations of the laser fields of optical cavities and/or interferometers has a long history [e.g., 20], and has had much application lately [21, 22] to the problem of interferometric GW-detection. The principle advance represented by the research presented here is the extensive modelling of *realistically deformed* optics in a *complete* (as defined below) LIGO interferometer, where additional laser fields needed for the GW-signal detection process are also modelled, and where the interferometers are *fully-optimized* for GW-detection for each different set of realistic optics that is included. This is also the first research performed for an Advanced-LIGO configuration known as "Dual Recycling" in which optics with realistic deformations have been modelled. All of these developments allow us to model first-generation and future-generation LIGO interferometers in unprecedented detail, which has enabled us to participate in a wide variety of research and development efforts (Section 3.5 and Chapter 4), and make practical recommendations (Chapter 5) to the LIGO Project regarding the design of interferometric GW detectors.

A brief overview of the current state-of-the-art of LIGO optics, and of the general Project requirements for optic polishing and coating, is given in Appendix B; the production of this first generation of LIGO mirrors and mirror specifications has been guided in part by the results of our numerical interferometer simulation research.

---

1. Absorption can be treated as pure loss as long as thermally-induced refractive effects due to mirror heating can be neglected.

1.3.3 Initial-LIGO configuration studies

A schematic diagram of the core optical configuration [5] of an Initial-LIGO interferometer is shown in Figure 1.5. While it is appears much more complex than the simple Michelson interferometer depicted in Fig. 1.2, the principle of GW-detection is the same: a GW will cause opposite motions for the mirrors in the two arms, leading to a differential phase shift at the beamsplitter, and an output fringe at the beamsplitter exit port that possesses the GW-signal. This output signal carries the time dependence of the GW, and is most effectively produced by a GW with optimal propagation direction (perpendicular to the page) and a polarization orientation aligned as in Fig. 1.1.

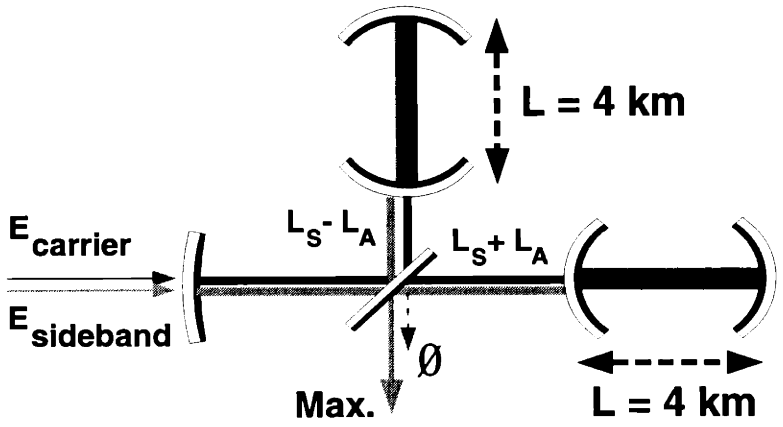

Figure 1.5: The core optical configuration of an Initial-LIGO interferometer.

The added mirrors in this interferometer configuration serve to amplify the GW-signal as much as possible. The long-baseline (4 km) arms increase the signal because the GW represents a *strain* in space, such that ΔL_{GW} is proportional to L. One can gain by increasing L all the way up to $L \sim \lambda_{GW}/2$, after which the round-trip travel time through the arms is so long that gains made during the first half-period of the GW are washed out during the second half period. For a highest (scientifically interesting) observation frequency of, say, $\nu_{GW} \sim 1$ kHz, the places a limit on the arm length of $L \sim 300$ km.

Given the infeasibility of constructing beam enclosures that are 300 km long, LIGO compensates by placing additional, semi-transparent mirrors in the arms, thus turning them into Fabry-Perot cavities. These cavities are multi-bounce systems which amplify (and narrow-band) the GW-response by bringing the arms up to the selected round-trip travel-time limit. The current choice of mirror reflectivity for the LIGO Fabry-Perot input mirrors brings the roll-off frequency for the GW-signal to ~90 Hz, after which the (shot-noise-limited) GW-sensitivity falls like $1/\nu_{GW}$ (c.f. Eq's. 3.3, 3.5, 3.6 in Section 3.1).

As depicted in the Fig. 1.5, LIGO is a multiple-frequency system: radio-frequency sidebands are impressed upon the carrier light in order to serve as a heterodyne detection scheme for GW's. The carrier light (in the absence of GW's) is brought, via control systems, as nearly as possible to a dark-fringe at the beamsplitter exit port, while the sidebands are maximally channelled through the exit port in order to superpose with any GW-induced light that emerges, so that the combined output beam can be demodulated to produce the GW-signal. These differing phase conditions for the carrier and sideband beams at the exit port are achieved via the implementation of an asymmetry, denoted in the figure by $L_s \pm L_a$, between the lengths of the small sections of the arms that lay between the beamsplitter and the Fabry-Perot cavities. The complete signal detection process, including the sideband frequencies and the length asymmetry scheme, is discussed further in Chapter 2, and full calculations are performed in the Appendices.

Lastly, in regards to Fig. 1.5, we note the presence of an input mirror to the entire interferometer, known as the *Power Recycling Mirror* [23]. This mirror is placed there to recycle all of the power emerging from the "bright-port" of the beamsplitter back into the system. In particular, this should include virtually all of the carrier light reflected from the Fabry-Perot arms (if the exit-port dark fringe is well-maintained), and by re-injecting this power into the interferometer (rather than absorbing or otherwise dissipating it), the Power Recycling Mirror increases the GW-signal by a broadband factor equal to the square root of the carrier power gain in the arms. Thus this interferometer configuration gets maximum use out of the power used to excite the system.

As introduced in the preceding sections, we have constructed a simulation program to model the behavior of this Initial-LIGO interferometric detector for optics possessing imperfections of many types, including deformations such as those present within real-world optics. The goals of this work are to help provide answers for the following questions:

- What are the effects of realistically imperfect optics upon interferometer performance and GW-sensitivity?
- What quality of mirrors are necessary in order to meet the Initial-LIGO performance requirements?
- Is it even possible to meet the Initial-LIGO requirements with optics possessing realistic deformations?
- How much can the scientific goals of LIGO benefit by using optics of the highest possible quality?

The thesis research to be presented here is intended to contribute significantly towards each of these fundamental lines of inquiry. The results, as will be seen in Chapter 3, portray an optimistic picture of predicted LIGO performance, as well as placing a premium upon the fabrication of very-high-quality optics.

1.3.4 Studies of an Advanced-LIGO configuration: "Dual Recycling"

In addition to characterizing the performance of Initial-LIGO interferometers, there has been a second initiative central to this research: investigating the performance of a promising Advanced-LIGO configuration, called Dual Recycling [24], with realistically imperfect optics. More than simply an Initial-LIGO configuration with enhanced subsystems, an *Advanced-LIGO* configuration such as this one incorporates additional optical components that change the GW-frequency response of the detector, as well as many aspects of the basic functioning of the interferometer. Dual Recycling, as a particular case, is fundamentally equivalent in response to many of the other Advanced-LIGO configurations that have been proposed (e.g., "Resonant Recycling" [23], "De-tuned Fabry-Perot Arms" [25]), and yet is both more convenient to use and better behaved than most [24]; we have therefore chosen to focus upon the Dual Recycling configuration for our research on Advanced-LIGO interferometers.

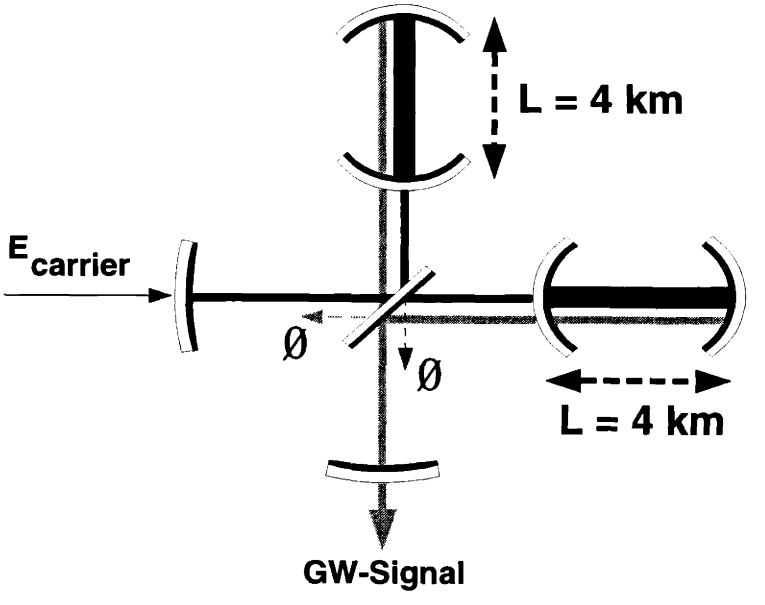

Figure 1.6: The core optical configuration of a Dual-Recycled LIGO interferometer.

A schematic diagram of an interferometer with Dual Recycling is shown in Figure 1.6. It is nearly identical to the Initial-LIGO configuration (c.f. Fig. 1.5), except for the addition of a mirror at the beamsplitter exit port, called the *Signal Recycling Mirror.* The function of this mirror is to reflect the power in the oscillating exit port fringe (which bears the GW-signal) back into the interferometer for resonant amplification. This "Signal Recycling", in conjunction to the Power Recycling which already exists, is what gives this configuration the name of "Dual Recycling".

For Dual Recycling, it is useful to consider the "GW-induced-sideband" picture (introduced in Sec. 1.1.1) as the appropriate basis for the analysis of the GW-detection process. In this picture, one views the action of GW as being the generation of GW-sidebands (at $\nu = \nu_{Laser} \pm \nu_{GW}$) in the Fabry-Perot arms. These GW-sidebands (depicted in Fig. 1.6) then experience a *double-resonance* in the coupled-cavity system of the Fabry-Perot arms and the *Signal Recycling Cavity* (SRC), where the SRC is defined as the path between the Fabry-Perot input mirrors and the Signal Recycling Mirror.

One difference between Figures 1.5 and 1.6 is the absence of radio-frequency sideband light (as well as absence of a length asymmetry) in the latter case. This is because the formulation of a detection scheme is more complicated for the Dual Recycling case, especially when the arms possess Fabry-Perot cavities, and also because the adoption of a detection scheme for Advanced-LIGO interferometers may depend upon practical considerations which are not yet clear at this time. Rather than modelling radio-frequency sidebands, we have chosen to retain full generality by simulating the GW-induced sideband beams themselves, from which we can obtain the GW-frequency dependent signal amplitude; and also by computing the shot noise level from just the exiting carrier-frequency light, assuming it to be the dominant source of power at the signal port.

The double-resonance experienced by the GW-sidebands in the coupled arm cavities/ SRC system can be used to both *narrowband* and *shift* the peak of the interferometer's shot-noise-limited GW-sensitivity function away from $\nu_{GW} = 0$. Further details of how these different response functions are produced are presented in Chapter 4; for now, we note that narrower sensitivity peaks are obtained with higher reflectivities of the Dual Recycling Mirror, and that shifting the GW-frequency of the peak is achieved via sub-wavelength-scale displacements of the Signal Recycling Mirror.

The benefits of altering the shot-noise curve with Dual Recycling are obvious from the rising shot noise level (versus frequency) shown in Figure 1.4. At low GW-frequencies, LIGO is dominated by a host of noise sources, especially seismic and thermal noise. But at higher frequencies (~100-200 Hz and above) shot noise alone dominates, so that it is beneficial to move the region of low shot noise away from $\nu_{GW} = 0$ to these frequencies. In addition, the range of ~100 Hz-1 kHz is a very important region of frequency space for astrophysical GW-sources that should be detectable to LIGO, such as periodic emissions from non-axisymmetric pulsars and the final merger/ringdown of Black Hole binaries [2]. Strong narrowbanding of the sensitivity curve at a chosen GW-frequency can, in fact, greatly enhance the detectability of a source which remains within a fixed range of ν_{GW} for many emission cycles, such as pulsars. Given the great potential usefulness of tuning the peak of the GW-sensitivity curve, Dual Recycling has an advantage over several other Advanced-LIGO interferometer configurations because of the ease with which it can be tuned: only microscopic adjustments of the Signal Recycling Mirror are required.

Dual Recycling is also believed to have other important benefits, especially in regards to tolerance for imperfections of the optics [24]. First of all, by placing a mirror in front of

the exit port of the beamsplitter, power that would leak out due to an imperfect carrier dark-fringe is inhibited from leaving, so that the output of undesirable beam modes (which may interfere with interferometer control systems or with GW-detection) is reduced, as well as their implicit contributions to the overall shot noise level. Perhaps more significantly, since exit-port losses would be reduced, the degradation of power buildup in the arms caused by mirror imperfections is less severe. Furthermore, the power returned ("recycled") to the interferometer arms may be re-combined with the resonant light that has remained there, thus increasing the total amount of power available for the creation of GW-sidebands. This re-injection of "noise power" into the arms to increase the amount of power available for signal detection is known as "Wavefront Healing", and this process is predicted to be a significant benefit of Dual Recycling interferometers [26].

Given the potential importance of Dual Recycling as a configuration for Advanced-LIGO interferometers, we have studied the performance of a Dual-Recycled interferometer with optics possessing realistic deformations. The focus has been to evaluate the various claims made by proponents of Dual Recycling, such as:

- Can frequency responses that are significantly tuned and/or narrowbanded be achieved with a Dual Recycling interferometer possessing realistically imperfect optics?
- How much does Dual Recycling improve signal-to-noise ratios and the overall GW-sensitivity curve of a LIGO interferometer?
- Is Dual Recycling really more tolerant of mirror imperfections than the Initial-LIGO configuration?
- Does "Wavefront Healing" exist, as predicted? How significant is it?

The results of this research, as will be shown in Chapter 4, are somewhat mixed: the desired sensitivity curves of Dual Recycling are indeed obtainable, but increased tolerance to mirror imperfections (and Wavefront Healing) is often less pronounced than hoped for, and it is only achieved when one is very careful about the design of interferometer optical parameters.

Chapter 2

A technical exposition of the interferometer simulation program

2.1 The physical systems to be modelled

2.1.1 Capabilities and assumptions of the model

We begin this chapter with an introduction to the range of physical properties that are modelled in this interferometer simulation initiative. Since a more faithful representation of the relevant physics will lead to a more accurate and predictive model, we have attempted to include as much realism as possible, while also retaining computational feasibility.

First, we note the variety of optical imperfections or features that can be modelled for each interferometer mirror. This list includes tilts, shifts, and mirror surface curvatures (used for modelling beam mismatch), and finite mirror sizes bounded by (fully absorbing) apertures. Also included are two-dimensional maps of both figure deformations and reflection/transmission amplitude variations, and separate maps are used for the independent operations of reflection (from either side of the mirror) and transmission. Pure losses are also included, representing power absorption and scattering at angles too high to be modelled, given our grid resolution (and also too high to be retained by an interferometer with finite-aperture optics). For almost all mirrors, the coefficients of reflection, transmission, and loss can be defined independently for each side of the mirror. The mirrors can be modelled with nonzero thicknesses and specifiable index of refraction, including the beamsplitter which is at 45° with respect to the incident beams (and thus applies lateral shifts to the beams which it transmits). Lastly, the intracavity distances and beam frequencies are handled very carefully in the program, so that beam focusing in all mirrors and during propagations, as well as all phase shifts (whether due to long-distance propagation, frequency differences between beams, or different round-trip phase shifts for different spatial beam modes) are handled in a physically correct and consistent fashion.

The laser beams themselves will be highly collimated, and it is therefore possible to consider the (folded) interferometer as a *pseudo-one-dimensional* system defined along a given propagation axis, where data for the electromagnetic fields (and mirrors) can be stored as 2-D slices sampled at various positions along the propagation axis, using two-dimensional grids. The physics of grid-based modelling will be discussed at length in Sec.

2.3; here we note that the information recorded about the propagating fields will be their electric field components (with the magnetic fields thus determined since the propagation is in vacuum), without polarization information, so that birefringence effects of the optics, for example, are not modelled.

Our simulation program models a *static* interferometer, one that is considered to be locked to the proper operating point by LIGO's tilt and displacement control systems. The output of our model will therefore be the *steady-state* electric fields that relax at various locations in the stationary interferometer. The problem of modelling dynamic interferometer performance during resonance acquisition and lock has been done to a limited degree within the LIGO Project [27, 28], but always for perfect or near-perfect mirrors, never for cases with spatially-complex mirror deformations. The more difficult problem of modelling dynamic performance *with* realistically imperfect optics would be very computationally intensive, because of the need for detailed temporal (as well as spatial) resolution, but it may turn out to be a feasible and valuable project for LIGO in the future.

The effects of a gravitational wave upon the interferometer laser fields, which generate the GW-signal, are also computed from the static interferometer model. In the case of the Initial-LIGO configuration, analytical calculations based upon the steady-state power buildup in the arms is sufficient for accurately calculating the response as a function of GW-frequency; and in the Dual Recycling model, where analytical calculation is inaccurate because the GW-induced signal sidebands are stored for long periods of time in cavities with imperfect optics (i.e., in the SRC and the Fabry-Perot arms), the steady-state behavior of the GW-induced-sideband at *each GW-frequency of interest* is individually modelled, so that the frequency response can be traced with any desired number of points. In either case, modelling a dynamic interferometer that oscillates with a GW is unnecessary; and though some dynamic simulations (with geometrical mirror deformations only, and without Fabry-Perot cavities in the Michelson arms) have been carried out in the literature [29, 30], we have chosen to use a static model because it takes far less time computationally (thus allowing us to explore a larger variety of runs), because the effects of steady-state field convergence need not be disentangled from GW-induced dynamic behavior, and because we can consider very small GW-amplitudes without having to relax the laser fields to correspondingly higher accuracy[1]. On the other hand, the principal drawback to a steady-state simulation is that we are unable to model the effects of dynamic noise sources, such as laser frequency- or amplitude-noise, unless they are specifically recast in a static form, such as by modelling them as noise sidebands imposed upon the carrier beam.

Finally, we note that although our program does not model a dynamic interferometer, it does perform a *dynamic optimization* of many interferometer parameters as the steady-

1. This is because we can assume arbitrary amplitudes for the GW-sidebands that are modelled, and apply the true GW-amplitude, h, to the results after the run has ended.

state fields are relaxed, so that the final result is an interferometer which is fully-optimized for GW-detection. The prototypical example of such an optimization task are *length adjustments* to obtain the proper resonance conditions for each of the interferometer cavities, a process which depends upon the exact nature of the steady-state cavity fields. This optimization task, as well as several others, will be the subject of Section 2.5. These optimizations are not done in the same way that they would be done in a real LIGO interferometer, since there are more convenient optimization techniques available for use in a simulation program, where any desired physical information is more easily available than in an experimental system. Our simulation program is therefore simplified by only having to model the optical components and beam frequencies that are needed for the direct computation of the GW-signal, such as the radio-frequency sidebands for the Initial-LIGO, or the GW-induced sidebands for the Dual-Recycled LIGO (along with the carrier-frequency beam for both cases). Other components for the control systems of a real interferometer can be neglected.

2.1.2 Interferometer specifications

Figure 2.1 is a diagram of the core optical configuration of a complete Initial-LIGO interferometer — with the additional mirror required for the Advanced-LIGO configuration of Dual Recycling shown in the dashed box — with components and cavity parameters labelled according to the scheme that will be used throughout this thesis. Not shown (or modelled by the simulation program) are the mode cleaning, frequency stabilizing, and matching optics which prepare the laser light for the interferometer; also not shown or modelled are the pickoffs, phase modulators, and control systems that will be used in a real interferometer to read out its operational state. As mentioned above, with the exception of the radio frequency (RF) sidebands for the Initial-LIGO case, the simulation program uses alternative optimization methods.

The mirror denoted by R_1, T_1 is the Power Recycling Mirror, and the folded cavity spanning L_1 and L_2/L_3 is called the *Power Recycling Cavity* (PRC); similarly, the mirror denoted by R_6, T_6 is called the Signal Recycling Mirror, with the folded cavity spanning L_6 and L_2/L_3 comprising the *Signal Recycling Cavity* (SRC). The PRC and SRC in a Dual Recycling interferometer are almost entirely decoupled because of the "dark-fringe" that is experienced in going from one to the other through the beamsplitter, and this decoupling simplifies the simulation process for Dual Recycling. For both configurations, there are two Fabry-Perot arm cavities, with the *inline* arm cavity being comprised of input mirror R_2, T_2 and back mirror R_4, and the *offline* arm cavity being comprised of input mirror R_3, T_3 and back mirror R_5.

The Dual Recycling configuration which we have investigated in this research is identical to the Initial-LIGO configuration except for the presence of the Signal Recycling Mirror, and the absence (for convenience and simplicity of results) of a length asymmetry

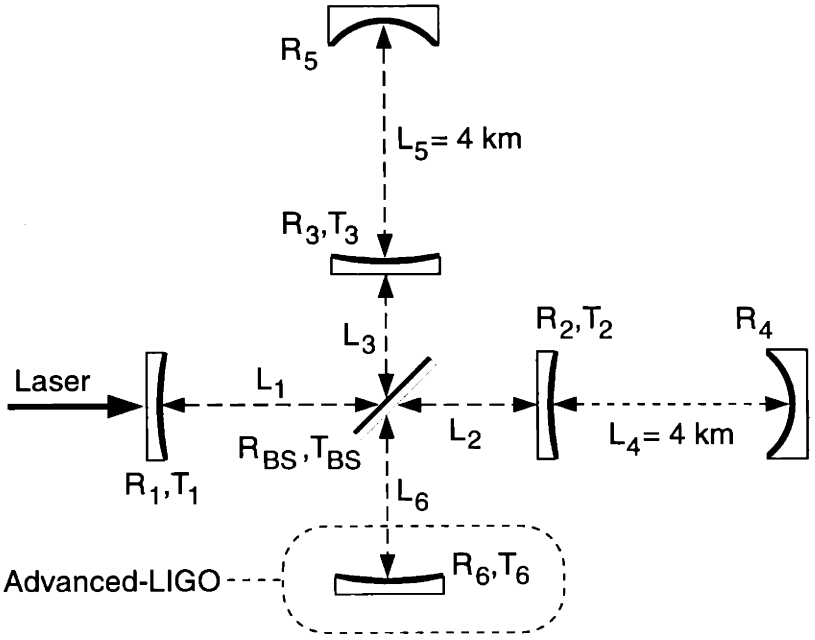

Figure 2.1: Interferometer components and component labels for the core optical configuration of an Initial- (or Advanced-) LIGO interferometer (not shown to scale).

between the interferometer arms. While the actual application of Dual Recycling in LIGO would likely come in an *enhanced* interferometer possessing updated interferometer parameters (e.g., lower mirror losses, higher input laser power, different Fabry-Perot cavity storage times), and differences in other relevant quantities (e.g., lower levels of thermal and/or seismic noise), we have opted to keep the interferometers parameters as similar as possible for these runs, in order to facilitate direct comparisons between the performance of the Initial and Advanced Configurations, given equivalent optics. A future goal of this simulation initiative may indeed be the exploration of Dual Recycling with fully-updated parameters, when accurate values for those parameters can be estimated with confidence.

A summary of values for the interferometer parameters shown in Fig. 2.1 (and other important quantities) is presented in Table 2.1. These are the main program input values which are used for all of the data runs to be presented below, barring exceptions that will be explicitly noted. In addition to physical parameters of the optical components, some computational parameters used by the simulation program are also included (the pixelization of the 2-D grids, etc.), which will be discussed in upcoming sections. Not included in Table 2.1 is the power of the excitation laser; in general, the powers of all relaxed interferometer fields are reported by the program as based upon 1 Watt of carrier and/or sideband power entering at the Power Recycling Mirror. When conversion to "true" power values is required, we re-scale the results to the estimated LIGO figure of 6 Watts of input power, divided as necessary (for the Initial-LIGO case) into the carrier and sideband beams.

Quantity	Value(s)
Laser Wavelength	1.064 μm (Nd:YAG light)
Sideband Modulation Frequency	ν_{mod} ~ -24.0 MHz (Initial-LIGO model only)
Cavity Lengths	L_1 = 5.0 m L_2 = 4.19 m + L_{asymm} L_3 = 4.19 m - L_{asymm} L_{asymm} ~ 9 – 25 cm (Zero for Dual Recycling) $L_4 = L_5$ = 4.0 km L_6 = 5 m (or 2004.19 m)
Mirror Curvature Radii	Rad_1 = 10.0 km (9.99 km for Dual Recycling) $Rad_2 = Rad_3$ = 14.56 km $Rad_4 = Rad_5$ = 7.4 km Rad_6 = 14.1 km
Mirror Intensity Reflectivities (Refl. Side)	R_1 ~ .9861 – .9390; R_6 varied from Zero – .99 $R_2 = R_3$ = .97 $R_4 = R_5$ = .99994 R_{bs} = .49992
Mirror Intensity Reflectivities (A.R. Side)	R_1, R_6 same as Reflective Side $R_2 = R_3$ = .968817 R_{bs} = .49971
Mirror Intensity Transmissions (Both Sides) (Pure Loss ≡ 1 – R – T)	T_1 ~ .01385 – .06095; T_6 = 1 - R_6 - 50 ppm loss $T_2 = T_3$ = .02995 T_{bs} = .50003
Beam Waist Diameter	7.02 cm
Mirror Aperture Diameters	24 cm (Circular Profile Mirrors), 24.4 x 17.2 cm (Beamsplitter at 45° w.r.t. Beam Axis)
Mirror Thicknesses (Perpendicular to Surface)	Beamsplitter = 4 cm All Others = 10 cm
Substrate Refraction Index	n = 1.44963
Calculational Window Size	70 cm x 70 cm (Square)
Gridding of Calculational Window	256 x 256 pixels
Max # of field relaxation steps; allowed errors	4000 iterations / 1 part in 10^{-4}-10^{-5} (varies)

Table 2.1: Typical parameter values for a LIGO interferometer, including both physical specifications and computational parameters. Some parameters are optimized during program execution, and are thus given only as approximate ranges of values here.

Since the program models the transverse spatial structure of the Gaussian-profile laser beams and the focusing properties of the mirrors, the simulated laser beam must be correctly mode-matched into the interferometer, or power will be scattered from the TEM_{00} mode into higher Hermite-Gaussian modes (TEM_{20} and TEM_{02}). The spot sizes and curvature radii of the beam everywhere in the interferometer can be fixed uniquely by specifying two beam parameters anywhere in the system; the two parameters used by the program are the *beam spot size* and *curvature radius* inside the inline Fabry-Perot arm cavity, next to its input mirror, M_2. The simulation program does not implement any form of automatic mode-matching to set these parameters, because the practical considerations for mode-matching in LIGO are complex (e.g., the nonzero length asymmetry leads to different curvature radii for M_2 and M_3, but same-curvature mirrors are easier to fabricate, so that mismatch for both arms must somehow be balanced). But a preliminary routine is implemented in the code which computes the focusing effects of finite-thickness mirrors, non-unity index of refraction, and free space propagations, and reports them before the electric field relaxations are done, so that a program user can perform trial-and-error searches for the two beam parameters which create the best match. For the data in Table 2.1, the best matching parameters we found were a spot size of 3.63 cm and a curvature radius of 14.56 km, at M_2.

The principal output of the simulation program, given all of the physical specifications and input data, are the relaxed steady-state electric fields which build up in the interferometer cavities. These fields are written out to data files for graphical and modal analysis, and it is possible to write out the steady-state fields at any interferometer location of interest. Figure 2.2, which concludes this section, shows the locations (and propagation directions) of the electric fields (carrier, RF-sideband, and/or GW-sideband fields) which we choose to generate in a typical run.

2.2 Computational specifications and facilities for program execution

The LIGO simulation program has been written in the SPARCcompiler version 3.0 of Fortran 77, and the local platforms for our runs are Sun SPARCstation 5 and 20 workstations. All variables are kept to double-precision. A complete simulation run for an Initial-LIGO interferometer (including carrier and sideband frequencies, with all interferometer fields computed and all optimizations done) with a set of non-ideal mirrors, and with a pixelization of 128x128 for the grids representing the electric fields and mirror maps, takes somewhat less than a day on a 2-processor SPARCstation 20. The run time varies depending upon the seriousness of the interferometer imperfections being studied, the number of parameters that are optimized during run-time, and the closeness of our initial guesses of these adjustable parameters to their final, optimized values.

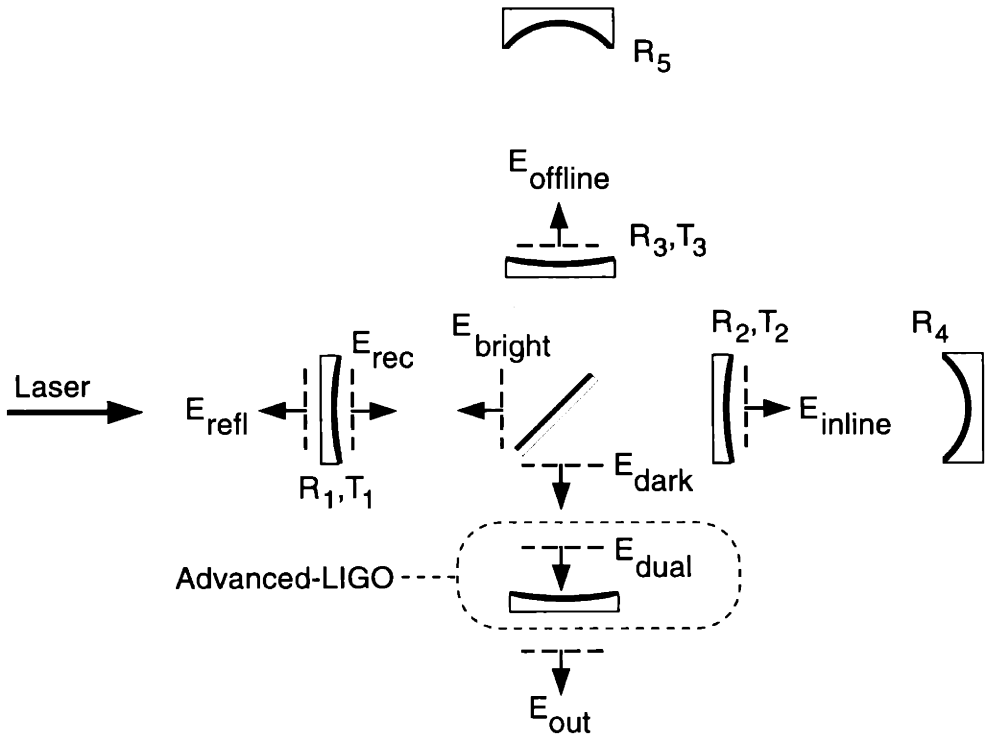

Figure 2.2: The steady-state electric fields computed by the LIGO simulation program.

While this achievement represents rapid execution times for simulations of full-LIGO interferometers with complex optical deformations, the necessity of going to 256x256 grids (Sec. 2.3.4.2), and the importance in Dual Recycling of modelling the GW-induced-sidebands for a large number of GW-frequencies (to obtain a frequency-response curve) requires a significant computational speed-up, especially if a large selection of runs are to be performed.

To that end, in collaboration with personnel of the Center for Advanced Computing Research (CACR) at the California Institute of Technology[1] we have converted the simulation code to parallel format for execution on the Paragon machine *Trex*, a 512 compute-node machine utilizing Intel i860 processors in each node. With Trex, we have achieved (using an appropriate number of parallel nodes) overall speedup factors of ~22 compared to the SPARCstation 20. The runs that will be presented in this paper have been performed on this parallel platform.

The number of parallel nodes to be used for a given simulation run depends upon the size of the simulation grids. Using more nodes means each node handles a smaller block of pixels, but it also creates more overhead in the form of message-passing between nodes. There is an optimum number beyond which the use of additional nodes has little effect on

1. Special thanks to Thanh Phung and Heidi Lorenz-Wirzba.

speedup, yet still increases the amount of node-hours (nodes used x run time) expended from LIGO's CACR account allotment. Table 2.2 shows the effect of varying the number of nodes for a sample 256x256 grid run, and demonstrates why the majority of our runs with 256x256 grids are performed with a compromise value of 64 nodes.

# of Trex nodes	Run Time (sec)	Run Time (re-normalized)	Allocation Time (nodes x time, re-normed)
32	7854	1.0	1.0
64	5021	.64	1.28
128	3835	.49	1.95

Table 2.2: Program run times versus number of parallel nodes used, for 256x256 grids.

2.3 The physics of grid-based modelling

This section includes many of the details and subtleties of modelling the physics of a LIGO interferometer on discrete, finite grids. Although it is a crucial component of the research that is presented in this thesis, it is possible to understand the results of our modelling initiative (particularly chapters 3 and 4) without the details presented in this section. A possible exception is Sec. 2.3.3.4, "Implementation of real-mirror measurement data".

2.3.1 Modelling laser beams with two-dimensional grids

As mentioned in Sec. 2.1.1, we use a pseudo-two-dimensional approach for the modelling of the laser field wavefronts in the interferometer. A primary propagation direction is taken for the (collimated) beam in each part of the interferometer, and a transverse slice of the propagating electromagnetic fields can be taken normal to the beam propagation axis, anywhere along the axis at an interferometer location of interest. Each of these slices is recorded on a two-dimensional grid, with a pixel entry representing the complex field amplitude at that transverse spatial position in the slice. No polarization vector information is recorded in our model, and the program utilizes scalar-wave theory, in particular the scalar Huygens-Fresnel integral formulation [17] for propagation. Only *electric* fields (e-fields) are computed and manipulated by the program, as the magnetic fields are assumed known, since the beam propagations occur in vacuum. Figure 2.3 is an example of such a grid-based electric field, in particular the real part of a Hermite-Gaussian TEM_{10} mode.

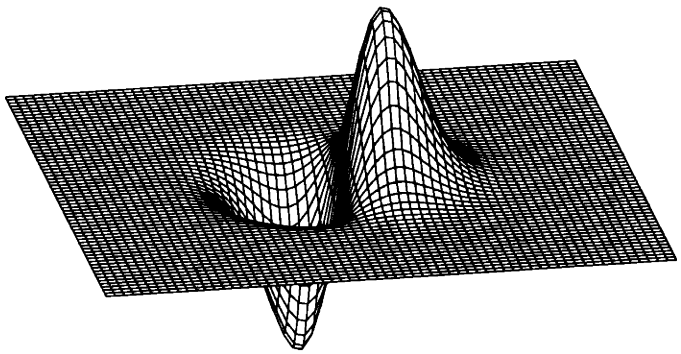

Figure 2.3: A transverse slice of a Hermite-Gaussian TEM_{10} mode (the real part is shown), taken at the waist plane of the beam, and recorded on a 64x64 pixelized grid.

An e-field slice taken at any interferometer location typically represents the superposition of *all* e-fields passing through that plane (i.e., the initial (excitation) field plus all reflections returning from other parts of the interferometer), so that a steady-state field summed from many bounces can be recorded upon a single 2-D grid. The summed e-field can thus be treated (for propagations and mirror interactions) as a single unit. Each e-field in the sum must be travelling in the same direction, however; if the counter-propagating field at the same interferometer location is desired, it must be recorded on a separate grid.

The realism and precision which can be achieved by the program ultimately depends upon how effectively it simulates the two basic physical processes that must be performed on the interferometer e-fields: *propagations*, and *interactions with mirrors*. We discuss propagations first.

2.3.2 The Paraxial Approximation and FFT-based propagation

For each curved, potentially jagged mirror profile, we define a flat plane in the region of that optic to serve as a reference plane[1]. Propagations translate an e-field slice through the macroscopically large distance from an initial reference plane to a destination reference plane, where either plane may be at a mirror location, or at an arbitrary position along the beam axis.

The LIGO interferometers will be excited by laser light that very well satisfies the *Paraxial Approximation*. A simple and intuitive statement of this approximation is that nearly all of the power propagating in the beam will be directed within a cone of half-angle θ about the main propagation axis, such that $\cos\theta \approx 1 - \theta^2/2$ is valid.

It is shown by Siegman [17] that once the paraxial (i.e., Fresnel) approximation is satisfied, plane-to-plane propagation of light can be performed via a three-step process

1. Each optic in fact has 2 reference planes, one on either side, separated by transmission through the mirror.

involving two (2-D) Fourier transforms, sandwiched around a pixel-by-pixel multiplication step in spatial-frequency-space ("k-space") of the e-field slice with a propagation operator matrix of the same dimensions. The computationally intensive steps in this propagation process are the forward and backward Fourier Transforms, which can be decimated by a Fast Fourier Transform (FFT) routine [e.g., 31]. The net result is an accurate and computationally efficient method for propagating the e-fields from one location to another in the simulated interferometer.

The propagation process, in detail, is as follows:

Step 1: At the initial plane, defined as z=0, perform an inverse Fourier transform of the electric field $\vec{E}(x, y, 0)$ into the spatial frequency domain:

$$\tilde{E}(k_x, k_y, 0) = \iint \vec{E}(x, y, 0) \times e^{i(k_x \cdot x + k_y \cdot y)} dx\, dy \tag{2.1}$$

Step 2: For each pixel in the spatial frequency domain (representing a single "plane wave" component of the original Gaussian beam), multiply by the appropriate phase factor representing a propagation to the far plane at distance z=L:

$$\tilde{E}(k_x, k_y, L) = \tilde{E}(k_x, k_y, 0) \times e^{-i(2\pi/\lambda)L + i\lambda(k_x^2 + k_y^2)L/4\pi} \tag{2.2}$$

Step 3: Finally, perform a forward Fourier transform to obtain the new position-space ("x-space") electric field on the destination plane, $\vec{E}(x, y, L)$:

$$\vec{E}(x, y, L) = \left(\frac{1}{2\pi}\right)^2 \iint \tilde{E}(k_x, k_y, L) \times e^{-i(k_x \cdot x + k_y \cdot y)} dk_x dk_y \tag{2.3}$$

where λ is the wavelength of the propagating light, and $i = \sqrt{-1}$.

The combination of these 3 steps (two transforms and a pixel-by-pixel multiplication of matrices) will be referred to as a "propagation". The integrals shown above are implemented in the simulation program as summations over the discrete grids, with an appropriate normalization factor for the 2-D sums [31]. Note that the use of FFT's requires us to set the number of pixels on a side of the (square) grids to be 2^N, for some integer N.

2.3.3 Interactions with spatially-complex mirrors

2.3.3.1 The physical representation of mirror maps

The interferometer simulation program must carry out two distinct mirror interaction operations: *reflections* (r) and *transmissions* (t). Mirrors will not be perfectly uniform or flat with respect to planes 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 their reflective coatings, significant curvatures, tilts, etc. This leads to optical path length variations (i.e., a spatially-varying *phase delay*) across their profiles, as well as variations in the *amplitudes* of r and t, the latter mostly due to variations in their reflective side or anti-reflective (A.R.) side coatings.

These phase and amplitude effects are simulated by creating complex mirror maps for all r and t operators, which are recorded (as the e-fields are) on 2-D pixelized grids. To make an e-field interact with one of these mirrors, it is first necessary to propagate that e-field, via the process described in the previous section, to a reference plane that is very close to the surface of the relevant mirror; one then is able to use a *short distance* algorithm to send the e-field from the reference plane to the (possibly irregular) reflective surface of the mirror, and then back again to the starting plane (or through the mirror to the reference plane on the other side, if it is for transmission rather than reflection).

As shown by Vinet, *et al.* [22], a very good short distance approximation to the exact (Huygens) integral calculation is to treat each pixel of the beam (in position-space) as an independent little plane wave segment, and to reflect (or transmit) that piece of the e-field by multiplying that pixel in the e-field map by the corresponding pixel in the (r or t) 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 single pixel-by-pixel multiplication step between a beam slice and a mirror operator map, a simple and computationally fast procedure.

To ascertain the mathematical form of the mirror map pixels, a mirror can be visualized as a surface of nonuniform deviation from a flat plane[1], with a "height function", $H(x,y)$, defined on the 2-D grid. An e-field pixel with laser wavenumber $k = 2\pi/\lambda$ would pick up a phase of:

$$\Phi_{\text{R.T.}}(x, y) = -2 \times k \times [-H(x, y)] \tag{2.4}$$

in a plane <—> mirror round-trip reflection[2], where we use the convention that a beam travelling a (shorter) *positive* distance picks up (less) *negative* phase. Note that reflections from the two different sides of a mirror are distinct and non-identical operations, which

1. Not all phase delays come from physical height variations, but this conceptualization is generalizable to other sources of mirror phase variations.
2. We use a similar formulation, without the round-trip factor of 2, for transmission operations.

differ by more than just a sign flip of the phase delay due to the height function, if the mirror has a nontrivial transmission map.

With the addition of an amplitude factor, a, which will equal the reflection or transmission coefficient of the mirror (and which is allowed to vary across the mirror profile), the spatial mirror map is defined in the plane transverse to the beam (z-)axis as follows:

$$M(x, y) = a(x, y)e^{i\Phi(x, y)} \tag{2.5}$$

Naturally, there are limitations to the mirror structure that can be represented with maps of these kind. These limitations, to be discussed in Sec. 2.3.4.1, generally occur when the laser wavelength is not significantly larger than the pixel-to-pixel height variations which one wishes to study. Nevertheless, this "independent little plane wave segment" formulation can be accurately used as a short distance approximation for interactions that extend over macroscopic distances, such as transmission through ~10 cm thick mirrors. The justification for this approach can be heuristically understood by considering the distance that a little beam with a waist that is *the size of a (half) pixel* would be able to travel before spreading out significantly onto its neighboring pixels[1]:

$$waist \approx \frac{1}{2} \cdot \left(\frac{\text{Calc. Window}}{\text{\# of Pixels}} \right) = \frac{1}{2} \cdot \left(\frac{.70 \text{ m}}{256} \right) = 1.37 \text{ mm} \tag{2.6}$$

$$\text{Rayleigh Range} \equiv Z_R = \frac{\pi (waist^2)}{\lambda} \approx \frac{\pi (1.37 \text{ mm})^2}{1.064 \, \mu\text{m}} \approx 5.5 \text{ m} \tag{2.7}$$

Since the Rayleigh Range, which is the distance a beam must propagate before spreading out to $\sqrt{2}$ times its waist size [17], is much bigger than the size of any mirror interaction region (even for the beamsplitter, a more complicated case to be discussed in Sec 2.3.3.3), the independent-pixel treatment of reflections and transmissions remains valid. In other words, even though an e-field pixel travels a distance equal to *many* wavelengths of light during propagation through a mirror, it does not experience any significant mixing with any other e-field pixels around it, and for this reason, representing mirror interactions via pixel-by-pixel multiplications of field maps with mirror maps is a very good approximation to the full 3-D physical process.

To sum up, we note that we can use these phase and amplitude maps to model a wide variety of mirror structure and/or imperfections, including — but not necessarily limited to — mirror tilts, various mirror curvatures or curvature distortions, finite apertures (Sec. 2.3.3.3), reflection and transmission maps taken from real, measured mirrors, or any specified mirror distortion function (e.g., Zernike polynomial distortions [32]) or surface roughness pattern which is desired.

---

1. This requirement is equivalent to the one given by Tridgell, *et al.* [29], in which a mirror profile must be in the "near field" for beam propagation angles supported by the grid (Sec. 2.3.4.2).

2.3.3.2 Enumeration of mirror maps and the conservation of energy

An important question which arises is how many mirror interaction maps will be needed to specify an entire mirror, for reflections and transmissions, from both sides. These maps will *not* be mutually independent, and in order to avoid simulations of an overdetermined and/or unphysical optic, one must account for the total information expressed by interactions with a mirror, as well as physical symmetries and overall energy conservation.

First, note that a 2-port mirror (reflective and A.R. sides) must relate 2 complex input e-fields to 2 complex output e-fields. Thus 4 complex (or 8 real) elements are needed, at each pixel location in the mirror grid map, to specify that pixel of mirror completely. These 4 complex elements correspond to the complex reflection and transmission operations performed from either side of the mirror. By time- or path-reversal symmetry — and by our assumption that each pixel of the beam expands negligibly during traversal of the mirror, so that the same small cross-section of mirror is sampled by a given beam pixel for transmission in either direction — the two transmission operations from either side must be the same: $t_{\text{refl.}} = t_{\text{A.R.}} \equiv t$. The mirror description is thus reduced to 3 complex elements per pixel, i.e., 3 complex mirror maps: the transmission map t, and reflection maps from either side, $r_{\text{refl.}}$, and $r_{\text{A.R.}}$[1].

Next, these maps must be specified carefully so that they conserve energy for all possible e-fields which may be incident upon them. We therefore have the necessary conditions (for each mirror pixel):

$$1 - |t|^2 - |r_{\text{refl.}}|^2 \equiv A_{\text{refl.}} \geq 0 \tag{2.8}$$

$$1 - |t|^2 - |r_{\text{A.R.}}|^2 \equiv A_{\text{A.R.}} \geq 0 \tag{2.9}$$

where $A_{\text{refl.}}$, $A_{\text{A.R.}}$ are the losses[2] experienced from either side (having $A \leq 0$ would represent *gain*, which we do not allow). But these conditions are not sufficient. Consider the complete e-fields which exist on either side of the mirror, both before and after the interaction (i.e., the incoming e-fields versus the outgoing ones):

---

1. The beamsplitter, which is a 4-port mirror, ends up requiring 4 complex mirror maps: $r_{\text{refl.}}$, $r_{\text{A.R.}}$, and *two distinct* transmission maps; one for the inline path, and one for the offline path. Also, for mirrors which need only be represented as 1-sided (e.g., the Fabry-Perot cavity back mirrors, which only encounter e-fields from the interferometer side), all A.R.-side and transmission information is neglected.
2. These loss coefficients may include contributions from several sources, such as absorption in reflective coatings and mirror substrates, as well as large-angle scattered power. Note that these coefficients only represent the *actual* power loss experienced by an e-field entering a given side of the mirror in the absence in interference effects, i.e., no if e-field is entering from the other side.

$$E_{\text{refl.}}^{\text{after}} = r_{\text{refl.}} \cdot E_{\text{refl.}}^{\text{before}} + t \cdot E_{\text{A.R.}}^{\text{before}} \tag{2.10}$$

$$E_{\text{A.R.}}^{\text{after}} = r_{\text{A.R.}} \cdot E_{\text{A.R.}}^{\text{before}} + t \cdot E_{\text{refl.}}^{\text{before}} \tag{2.11}$$

By requiring that the total power in all e-fields *does not increase* due to the mirror interaction, we obtain the inequality:

$$P^{\text{before}} - P^{\text{after}} = \left(\left|E_{\text{refl.}}^{\text{before}}\right|^2 + \left|E_{\text{A.R.}}^{\text{before}}\right|^2\right) - \left(\left|E_{\text{refl.}}^{\text{after}}\right|^2 + \left|E_{\text{A.R.}}^{\text{after}}\right|^2\right) \tag{2.12}$$

$$= A_{\text{refl.}} \cdot \left|E_{\text{refl.}}^{\text{before}}\right|^2 + A_{\text{A.R.}} \cdot \left|E_{\text{A.R.}}^{\text{before}}\right|^2$$

$$- 2\Re\left\{\left(t^* \cdot r_{\text{refl.}} + t \cdot r_{\text{A.R.}}{}^*\right) \cdot E_{\text{refl.}}^{\text{before}} \cdot \left(E_{\text{A.R.}}^{\text{before}}\right)^*\right\} \geq 0$$

The last term in this formula represents *interference* between the incoming e-fields from opposite sides of the mirror, and this interference term (which affects the amount of overall loss) must not be allowed to violate energy conservation.

To interpret Eq. 2.12, we first consider the simplest case: a loss-free mirror, with $A_{\text{refl.}} = A_{\text{A.R.}} = 0$, and $|r_{\text{refl.}}| = |r_{\text{A.R.}}| = \sqrt{1 - |t|^2} \equiv |r|$. Then, by constraining the mirror to satisfy energy conservation for *any* incoming e-fields, we must have the strict equality, $P^{\text{before}} - P^{\text{after}} = 0$, and hence:

$$\left(\frac{r_{\text{refl.}}}{r_{\text{A.R.}}{}^*}\right) = -\left(\frac{t}{t^*}\right) \tag{2.13}$$

Assuming the transmission coefficient, t, to be real, Eq. 2.13 generates the familiar possibilities for the signs (or complex phases) of the reflectivities on either side of a mirror: $\{r_{\text{refl.}}, r_{\text{A.R.}}\}/|r| = (\pm 1, \mp 1)$, $\{r_{\text{refl.}}, r_{\text{A.R.}}\}/|r| = (i, i)$, etc.

So far, this is simply equivalent to the relations derived by Stokes, by considering time-reversibility for the reflection of a light ray from a (single-layer, lossless) interface [33]. Extending our considerations to the more general case of a lossy mirror, we can convert Eq. 2.12 into a simpler prescription by varying the relative phases and amplitudes of the two input fields, $E_{\text{refl.}}^{\text{before}}$ and $E_{\text{A.R.}}^{\text{before}}$, in order to generate the *strictest possible* energy conservation condition; this condition becomes, after a little algebra[1]:

$$\frac{\sqrt{A_{\text{refl.}} \cdot A_{\text{A.R.}}}}{\left|t^* \cdot r_{\text{refl.}} + t \cdot r_{\text{A.R.}}{}^*\right|} \geq 1 \tag{2.14}$$

1. A similar energy-conservation condition can be derived for the 4-port beamsplitter.

This inequality is the most general condition that must be satisfied (at every pixel location) to guarantee conservation of energy in a mirror, given any e-fields which might be incident upon it. Our simulation program tests every mirror for this condition, and aborts the run if it is violated anywhere due to an unphysical combination of program input parameters and/or mirror deformation maps.

Now consider a mirror that is *nearly lossless*, such as a high-quality LIGO optic. In this case, Eq. 2.13 will be approximately true, and by using $|r_{\text{refl.}}| \approx |r_{\text{A.R.}}|$ (true since both of them $\approx \sqrt{1-|t|^2}$), one can extract the following phase condition:

$$\text{Phase}\{r_{\text{refl.}}\} + \text{Phase}\{r_{\text{A.R.}}\} \approx \pi + 2 \cdot \text{Phase}\{t\} \tag{2.15}$$

Plugging this formula into the general energy conservation equation, Eq. 2.14, one obtains the somewhat more intuitive result:

$$|r_{\text{refl.}}| - |r_{\text{A.R.}}| \leq \frac{\sqrt{A_{\text{refl.}} \cdot A_{\text{A.R.}}}}{|t|} \tag{2.16}$$

This formula merely states that $|r_{\text{refl.}}|$ and $|r_{\text{A.R.}}|$ cannot differ by very much — i.e., there cannot be much more loss on reflection from one side of the mirror than from the other side, such as the difference due to diversion of power by an A.R. side beam pickoff — unless the amount of loss common to both sides is high, or unless the overall transmission, t, is reduced.

In practice, the approximation specified by Eq. 2.15 is more than just a restriction upon the relationship between the phase maps for transmission and reflection from either side of the mirror. In fact, it can be used (if treated as a precise equality) to *derive* the remaining phase map, if two of the three are previously specified. We use this feature (available as a program option) often in our simulation work, since mirror losses are always small, in order to specify a complete mirror with only two input mirror maps: a "surface" and a "substrate" map.

2.3.3.3 Finite apertures and realistic-beamsplitter modelling

Another important aspect of constructing maps for the interferometer mirrors is the specification of their *finite apertures*. This is necessary because the mirrors are not infinitely large compared to the laser beam spot sizes, so that there is non-negligible diffractive loss past the mirror edges.

The simulation program thus multiplies the mirror function, $M(x,y)$ (c.f. Eq. 2.5), by a spatially-varying aperture function, $A(x,y)$. This function may either be graded or hard-edged, as desired (though Eq. 2.14 must always remain valid for all pixels, for energy conservation). For simplicity, in our runs we have implemented hard cutoffs, as follows:

$$A(x,y) = \begin{pmatrix} 1, & \sqrt{x^2+y^2} \le r_{\text{aper}} \\ 0, & \sqrt{x^2+y^2} > r_{\text{aper}} \end{pmatrix} \tag{2.17}$$

where r_{aper} is the radius of the circular aperture for each interferometer mirror. Standard aperture sizes for Initial-LIGO mirrors have been given in Table 2.1.

In this formulation, the power falling inside the mirror apertures remains unaffected by them, while 100% of the power falling outside of the apertures is removed from the system. Physical effects of this "lost" power, such as absorption-induced mirror heating, or diffracted light that bounces off the beamtubes and chances to return to the system (thus producing phase noise), are not currently modelled by the simulation program.

The beamsplitter, which will be a right-circular cylinder [34] tilted at 45° with respect to the propagating laser beams, will present an elliptical aperture to the e-fields rather than a circular one. The simulation program therefore allows independent specification of the major and minor axes of the elliptical beamsplitter aperture, where the short axis is in the plane of the interferometer (i.e., the plane within which the beamsplitter is tilted), and the long axis is out of the interferometer plane. The short axis of the beamsplitter is foreshortened by $\sqrt{2}$ compared to the circular diameter of the beamsplitter (at normal incidence), while the long axis is the same size as the circular diameter.

Because of the foreshortening of the beamsplitter in the plane of the interferometer, clipping of the beam by the reduced aperture is a significant concern, especially if there are position errors along the direction of the beamsplitter's short axis; in fact, the simulation program has been used as a tool to examine such effects [35]. Losses due to beamsplitter clipping are limited not only by keeping position errors low (~few mm), and by making a beamsplitter with a somewhat larger circular diameter than the other mirrors (see Table 2.1), but also by the predicted tendency (which we have observed in simulations) of laser fields in a resonant cavity to adapt their profiles to smaller widths, so that the dominant cavity mode is one with smaller diffraction losses [17].

The beamsplitter has an additional, important effect: because the beamsplitter has refraction index $\neq 1$, and because it exhibits a non-normal incidence angle for the laser beams which encounter it, the beams will experience a transverse "jog" when transmitted through it. Given a (perpendicular) beamsplitter thickness of T_{BS}, an incidence angle (from the normal) of θ, a mirror refraction index of n_{mir}, and a refraction index outside the optic of n_{ext}, a simple geometrical calculation using Snell's Law [33] shows that this transverse jog will be:

$$\Delta_{\text{jog}} = \frac{T_{\text{BS}} \cdot \sin^2\theta \cdot (n_{\text{ext}}/n_{\text{mir}})}{\sqrt{1 - [\sin\theta \times (n_{\text{ext}}/n_{\text{mir}})]^2}} \tag{2.18}$$

Note that Δ_{jog} is not defined as the *absolute* transverse beam displacement in space, but

rather it is the displacement of a beam from the center of the output face of the optic, assuming that in went in through the center of the input face of the optic. This definition is the appropriate one, since it reflects displacement with respect to the mirror-centered apertures.

Using $\theta = 45^\circ$, and $n_{ext} = n_{vacuum} = 1$, Eq. 2.18 simplifies to:

$$\Delta_{jog} = \frac{T_{BS}}{\sqrt{4 n_{mir} - 2}} \tag{2.19}$$

For the beamsplitter parameters shown in Table 2.1, this represents a (significant) jog of 1.8 cm.

When encountering either the reflective or A.R. side of the beamsplitter, the beam is multiplied (trimmed) by an aperture function. But because of the transverse jogs, each encountered aperture is shifted from the other one(s). Thus the net aperture acting on a beam cuts out all of the power that does not survive *every* aperture in the path.

Figure 2.4 shows the apertures (and the net aperture) for a beam traversing the different paths through the beamsplitter. Note that the two interfering beams emerging from any output port of the beamsplitter will have experienced different net apertures, so that constructive or destructive interference between them will not be perfect. In particular, the dark fringe at the beamsplitter exit port (i.e., destructive interference between offline transmission and A.R.-side reflection) cannot be perfectly cancelled out to zero power. This effect, which happens (fortunately) to be relatively small, will be demonstrated in Sec. 3.3.

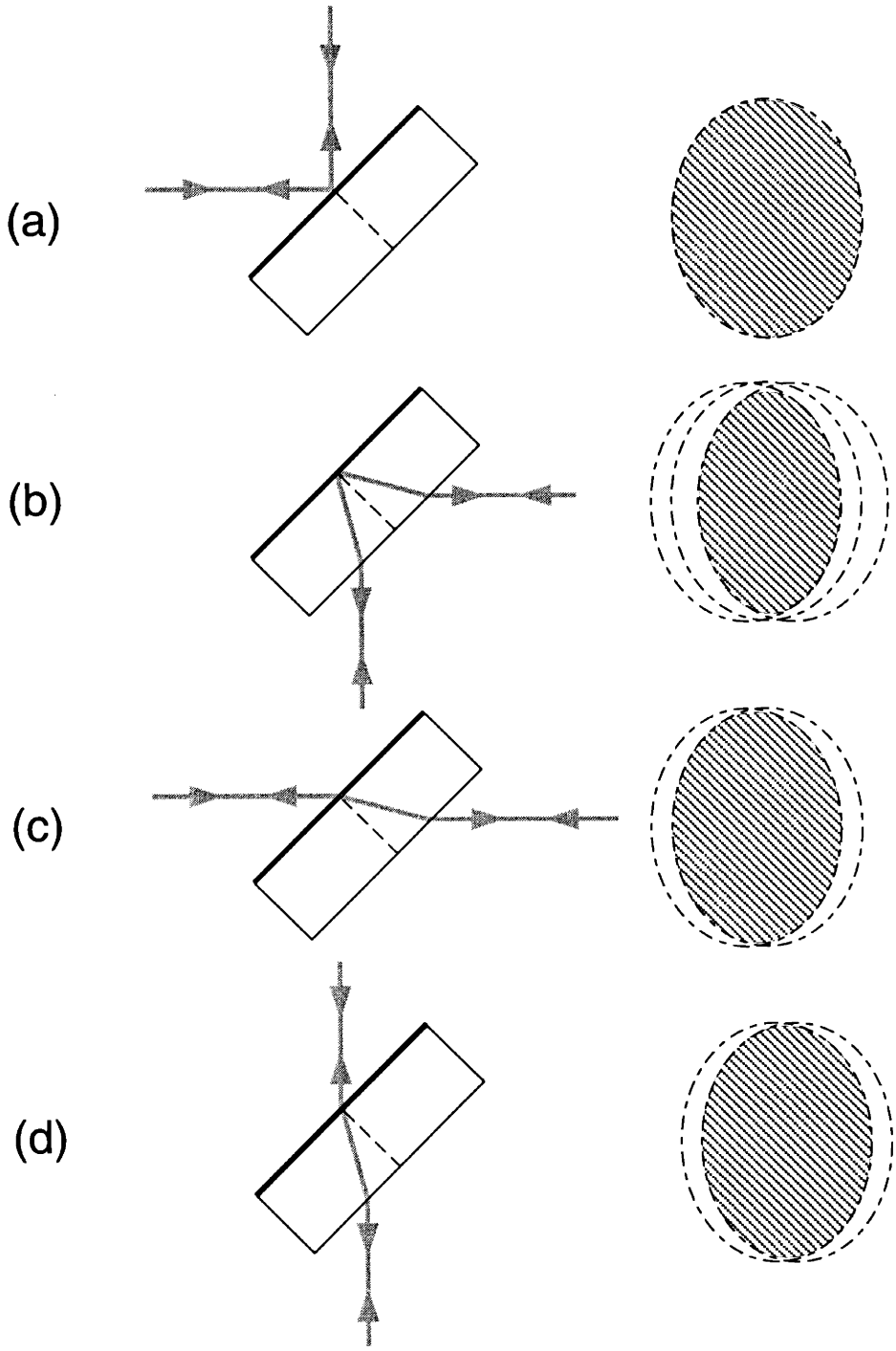

Figure 2.4: Different beam paths through the beamsplitter are shown in the left column, with their corresponding aperture contributions shown in the right column. The reflective surface of the beamsplitter is oriented to the left, facing the Power Recycling Mirror. The shaded areas are the net apertures for each case. The cases are: (a) reflective-side reflection, (b) A.R.-side reflection, (c) inline transmission, and, (d) offline transmission.

2.3.3.4 Implementation of real-mirror measurement data

One may divide complex mirror deformations up into two regimes: small spatial scale ("high spatial frequency") deformations, and large spatial scale ("low spatial frequency") deformations. The former regime of deformations tends to scatter power completely out of the interferometric detector, while the latter regime scatters power into high spatial modes that remain in the system, corrupting the resonating fields.

In the Fabry-Perot arm cavities (where the laser power is highest, and deformations are most significant), the boundary between the regimes is defined by a scattering angle of:

$$\theta \approx \frac{\text{Mirror diameter}}{\text{Arm Length}} = \frac{24 \text{ cm}}{4 \text{ km}} = 60 \ \mu\text{rad} \tag{2.20}$$

This scattering angle is associated with a spatial length scale given by:

$$\Delta x \sim \frac{\lambda}{2\theta} \approx 9 \text{ mm} \tag{2.21}$$

Since the pixel size in a typical run (c.f. Eq. 2.6) is 2.7 mm < 9 mm, the program grids have sufficient resolution to be used for detailed simulations of deformations in the low spatial frequency regime, and their effects upon the interferometer electric fields. The effects of significantly finer-scale deformations, on the other hand, must be dealt with as pure losses assigned to the mirrors — using "best guess" loss coefficients obtained from industrial mirror measurements — which is a good approximation for what actually happens to the large-angle scattered power, since it is completely lost from the system, barring significant back-scattering from baffles in the LIGO beamtubes.

We have obtained two measurements of low spatial frequency deformations in real optical components, in order to evaluate the behavior of a LIGO interferometer in the presence of realistic mirror deformations. 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 that was used by the AXAF program [e.g., 36] for calibration of the extremely smooth, high-resolution conical mirrors of their x-ray telescope. 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. Though initial studies are being done in LIGO of mirrors with a few layers of dielectric material (SiO_2 and Ta_2O_5) to be used for reflectivity and A.R. coatings [37], we do not currently have measurement maps of fully-coated mirrors available to us for study.

Since our resources of near-LIGO-quality mirror measurements were limited to one surface reflection map and one substrate transmission map, we proceeded to extend each of these maps into an array of many maps, in order to be able to place deformed surface and substrate maps on all of the interferometer mirrors simultaneously. With the help of former LIGO scientist Yaron Hefetz, a three-step process was used for the creation of

these new maps: (i) the surface (or substrate) source map was Fourier transformed into spatial-frequency space ("k-space"), (ii) the phases of the Fourier components were all *randomized*, and, (iii) an inverse Fourier transform was performed, which creates a new mirror with randomized features[1] but with the same power spectrum of deformations as the initial map. This process was performed many times, to create 15 new (128 x 128 pixel, 35 x 35 cm) surface maps out of the initial Calflat map, and 7 new substrate maps out of the initial Corning map. Several 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.

Mirror maps for the beamsplitter had to be specially developed. The beamsplitter can actually be modelled as a "flat" optic oriented perpendicular to the beam propagation axis (see Eq. 2.7 and subsequent discussion), as long as the deformation maps account for the perspective effect of the beamsplitter's 45° tilt angle with respect to its incident beams. This was done by squeezing them laterally by a factor of $\sqrt{2}$ in the appropriate transverse direction, and then by multiplying them by constant scalar factors in order to account for the altered optical path length for the reflection and transmission beam paths. For reflections, geometric considerations cause the surface to appear "smoother" by a factor of $\sqrt{2}$, while the adjustment factor for the beamsplitter transmission operator depends (via Snell's Law) upon the chosen substrate thickness and the index of refraction at the specified laser wavelength. The detailed computation of this (and all) mirror transmission optical path lengths are performed during program execution.

Next, to prepare these mirror maps for use in the simulation program, they had to be adapted to acceptable grid parameters. From the gridding requirements that will be derived in Sec. 2.3.4.2, an appropriate gridding scheme, given the parameters of the Gaussian laser beams, would be (at least) 128 x 128 pixels covering a 70 cm x 70 cm square calculational window[2]. But since the deformed mirror maps that we originally prepared for use were of the form 128 x 128 pixels over a 35 cm x 35 cm window, rather than lose fine mirror detail we chose to extend these maps to 256 x 256 pixels covering a 70 cm x 70 cm window by padding the mirror map edges with zeroes (all beyond the mirror apertures), in order to bring them up to the full grid size. We note that the generated surface deformation maps (as was the original map provided to us) are just about equal in size to the physical extent of the apertured mirrors; outside of the area with data we have filled the maps with a flat border region. The substrate maps, however, are slightly smaller than the

1. The process used here obviously does not preserve *coherent* structures in the deformation maps (spikes, excess mirror curvature, etc.), that could be important (and possibly systematic) side-effects of the mirror-making process. Concerns about such coherent structures are dealt with on an independent basis, by performing additional runs with such defects added in separately.
2. Note that the calculational window must actually be significantly larger than the diameters of the interferometer mirrors; the reason for this is made clear in Sections 2.3.4.2 and 2.3.4.3.

mirror apertures, but we have extended the area of useful substrate information as much as possible (via repositioning of data), in order to cover a little more of the mirror edges. In any case, full precision in the outer mirror regions is relatively unimportant because of the exponential fall-off in power far away from the center of the Gaussian-profile laser beams.

In a real LIGO interferometer, there will be an auto-alignment control system [38] that holds each of the mirrors in alignment with respect to the cavity laser fields to a very high degree of accuracy (typically $\theta < 10^{-8}$ radians [39]). Therefore, it is a useful procedure to remove the *inherent* tilts from each of the constructed mirror maps, so that it is possible to either perform a run with no mirror tilts at all (i.e., perfect auto-alignment), or with precisely-known tilts that are specified independently as input data, rather than simulating the system with the haphazard (and perhaps unmeasured) tilts present in the maps that have been created via Fourier component randomization. To that end, we have written a small Fortran program with which one can remove the tilts from all of these mirror maps as a pre-processing step, before they are used as input for the simulation program. This tilt-removal algorithm, outlined in Appendix C, uses Gaussian-profile beams to remove the *beam-weighted* tilts from each map, so that the LIGO laser beams will not suffer any effects of mirror tilt (i.e., generation of TEM_{10} or TEM_{01} mode power) while resonating in the interferometer. Thus even if a mirror would *appear* to have an overall tilt, if a straight average were performed (including the less important outlying areas of the mirror), a Gaussian-profile beam centered on the mirror would experience a surface that is as flat as possible given the presence of deformations in the mirror. Using our tilt-removal program, we have created an entire set of tilt-free surface and substrate[1] maps for use in the full simulation program.

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 effects of mirror *coating* deformations, which are not yet reliably estimated and may be significantly worse than the polished surface deformations. To accomplish this, the Calflat-derived surface maps (with deformation "heights" defined in meters) were uniformly multiplied by scale factors to generate the new families. The original family of surfaces, which possess root-mean-square (rms) deformations of ~.6 nm when sampled over regions of radius 4 cm, is labelled "λ/1800" (with $\lambda \equiv \lambda_{YAG} = 1.064\mu m$); the scaled-up families are λ/1200, λ/800, and λ/400. In contrast, the mirror substrate maps, which are likely to represent the best quality of fused silica substrates that will be obtainable for the first-generation LIGO interferometers [40], possess rms deformations of ~1.2 nm when sampled over regions of 4 cm radius, and were not renormalized; all of our deformed-sub-

1. For modelling purposes, "tilts" can also be removed from mirror *substrate* deformation maps, since those tilts really represent mirror thickness wedges, and the primary effect of these exceedingly small wedges is merely to alter the reference axes of the various cavities; this will have negligible effects because of automatic mirror angular alignment in LIGO interferometers.

strate runs have been done with these "best possible" substrate maps. Note that a direct comparison of the substrate rms value with that of the surfaces is not meaningful, since the substrates have less of an impact upon the interferometer electric fields than the surfaces do, especially in the case of the power which builds up between the reflective surfaces of the Fabry-Perot arm cavity mirrors. The substrates are therefore less significant in their effects, except possibly for the radio frequency sideband fields (used for signal detection in the Initial-LIGO configuration), which are non-resonant in the Fabry-Perot arm cavities but have significant interactions with mirror substrates in the Power Recycling Cavity, because of beamsplitter transmissions and reflections from the A.R.-sides of the Fabry-Perot input mirrors.

An example of a $\lambda/1800$ surface map that is used for runs of the simulation program is shown in Figure 2.5 (with its border region clipped, and with the mirror shown at an angle, for clarity). The vertical range of deformations in this mirror extends from -2 to +3.5 nanometers, and the map width is such that it approximately covers the 24 cm diameter mirrors.

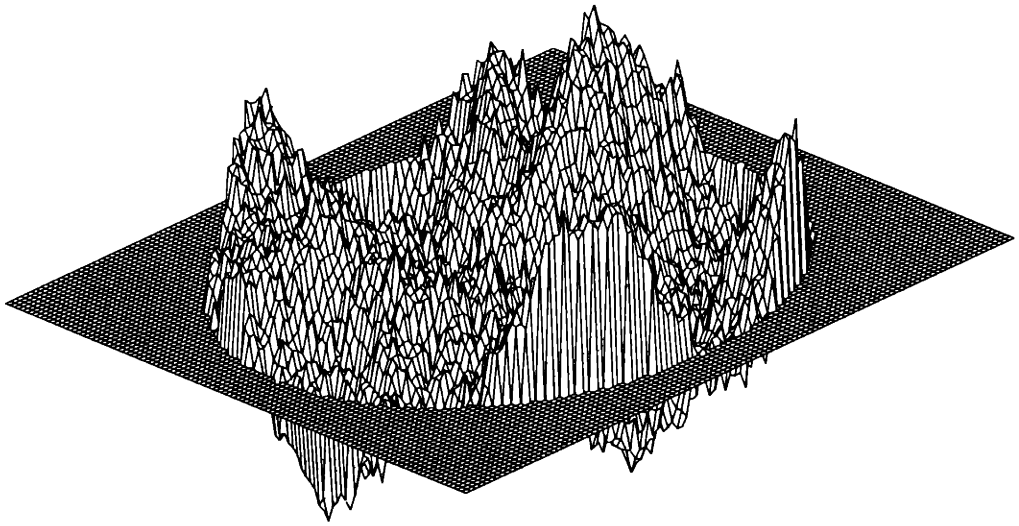

Figure 2.5: Map of a lambda/1800 mirror surface map with realistic deformations.

2.3.4 Fundamental limitations of discrete, finite grids

As might be expected, there are limitations to the physics that can be modelled in a grid-based optical simulation program. We break down these restrictions into three classes: limitations on the detail of the optics, gridding requirements to be met for representing Gaussian laser beams, and limitations on the information that can be carried by the finite, discrete grids. The following subsections describe each of these basic limitations.

2.3.4.1 Mirror tilt, curvature, and deformation limitations

Mirror structure is fundamentally represented as phase (and amplitude) maps on 2-D grids, where the phase part of a mirror map can be thought of being derived from a height

function, as per Eq. 2.4 and Eq. 2.5, and the related discussion. If the "height difference" between two neighboring pixels is $\lambda/2$, then the phase difference between those two points on the mirror map for a reflection[1] is given by:

$$|\Delta\Phi| = 2 \times k \times |\Delta H| = 2 \times \frac{2\pi}{\lambda} \times \frac{\lambda}{2} = 2\pi \tag{2.22}$$

As noted in [29], this height difference is unobservable, since $M(x, y)$ changes by the phase factor $e^{2\pi i} = 1$, and in fact any multiple of $\lambda/2$ added to the height difference between neighboring pixels is completely unmeasurable. It is therefore required that:

$$|\Delta H| < \frac{\lambda}{2} \tag{2.23}$$

must be obeyed from one pixel to the next by any deformation that is represented on the grids, in order to simulate the real mirror structure with sufficient spatial resolution.

Consider a mirror *tilt* deformation. A tilt about one mirror axis, of angle θ, creates a height difference of $\Delta H = \Delta x \cdot \sin\theta \approx \Delta x \cdot \theta$ for pixels that have a neighbor-to-neighbor separation of Δx. Inserting this into Eq. 2.23, we get the condition[2]:

$$\theta < \frac{\lambda}{2 \cdot \Delta x} = \frac{(1.064\ \mu\text{m})}{2 \cdot (.70\ \text{m}/256)} \approx 2 \times 10^{-4}\ \text{rad} \tag{2.24}$$

This tilt value is four orders of magnitude larger than the predicted angle errors for auto-aligned LIGO mirrors, and therefore this restriction does not place a serious constraint upon the study of realistic mirror tilts by the simulation program.

Similarly, there are restrictions upon the mirror curvature radii that can be modelled. The height function for a spherical mirror with curvature radius C_{rad} can be written as:

$$H(x, y) = C_{\text{rad}} - \sqrt{{C_{\text{rad}}}^2 - x^2 - y^2} \approx \frac{x^2 + y^2}{2C_{\text{rad}}} \tag{2.25}$$

The height differential between two neighboring pixels, for separations along the x-direction, is therefore equal to:

$$\Delta H = \frac{x}{C_{\text{rad}}} \cdot \Delta x \tag{2.26}$$

This ΔH is set to be less than $\lambda/2$ where the requirement is strictest — far away from the

1. *Reflections* give restrictions that are twice as conservative as those given by *transmissions*, because of the round-trip by the beam during reflection from the mirror phase profile.
2. The beamsplitter's base 45° angle with respect to the inline and offline beam propagation axes is handled separately (Sec. 2.333, 2.334) so that it does not fall under this tilt restriction; the beamsplitter, though, may be given additional (small) tilt which does.

center of the mirror. Choosing "x" to be at the edge of the aperture cutoffs for the interferometer mirrors with curvature (all have aperture radii of 12 cm), and Δx to be the inter-pixel separation of ~2.7 mm, we get the following restriction for mirror curvature radii in the program:

$$C_{\text{rad}} > 615 \text{ m} \tag{2.27}$$

Like the tilt restriction, this curvature restriction is not troublesome to satisfy for physically interesting runs.

Lastly, there are limitations on the spatially-fluctuating "realistic deformations" that can be represented on the grids. This limitation is also not very restrictive, since the rms fluctuation levels for the ("λ/1800", etc.) mirror deformation maps that we use are ~10^{-3}-10^{-4} times smaller than λ (and even the peak-to-peak variations are ~few nm), so that there is no danger of having a height differential of λ/2 for neighboring pixels.

2.3.4.2 Representing Gaussian beams

In this subsection, we consider the physics that can be represented with discrete, finite grids, and apply it to Gaussian-profile beams. In the end, the grids must satisfy two requirements: first, the grid-based e-fields must contain enough physical information to be realistic representations of the laser beams, and second, they must not introduce *false* information into program calculations to any significant degree. To that end, we will derive criteria that can be used to evaluate whether or not a given set of grid parameters is sufficiently well chosen.

We recall the well-known relationship between "position space" (x-space) and "momentum space" (k-space): large distances in position space represent small frequencies (or frequency differences) in momentum space, and small distances in position space represent large frequencies in momentum space. The prototypical example [41] is a Gaussian ("minimum uncertainty") wave packet for light or particles, in which a packet that is less physically spread out will have a greater momentum dispersion (e.g., strongly-focused laser beams spread faster), and vice-versa.

For grid-based modelling, this fundamental relationship can be summed up as follows:

Fine x-space sampling <—> High-momentum k-space information

Large x-space window <—> Fine k-space sampling

Because the Fourier transform based beam propagations (Sec. 2.3.2) move the laser fields back and forth between position and momentum space, the simulation grids must allow for a large enough range ("window") *and* for sufficiently fine sampling in both x-space and in k-space. The challenge is achieving all of these goals simultaneously.

For now, we will consider the k-space requirements. When an electric field is stored on a k-space grid during Step 2 of a propagation (Eq. 2.2), each pixel in the grid represents a certain spatial frequency value, k_x (or k_y), going up to some highest spatial frequency of k_{max}. This maximal value will be determined by the *Nyquist critical frequency* (f_c) for the grid, and we have [31]:

$$k_{max} = 2\pi f_c = \frac{\pi}{\Delta x} = \frac{\pi \cdot N_{pix}}{W} \tag{2.28}$$

where Δx is physical distance between neighboring pixels in the x-space grids, and W is the size of the calculational window spanned by the grid (with these 2-D grids having $N_{pix} \times N_{pix}$ pixels on a side).

In discretely sampling spatial frequencies $-k_{max}$ to k_{max} on a grid with this pixelization, the k-space grids contain the following spatial frequency components[1]:

$$k_n = n \cdot \Delta k\,, \quad n = -\left(\frac{N_{pix}}{2} - 1\right), \ldots, \frac{N_{pix}}{2} \tag{2.29}$$

$$\text{where:} \quad \Delta k = \frac{2\pi}{N_{pix}\,\Delta x} = \frac{2\pi}{W}$$

As is well known [31], discrete sampling of a function (or continuous data) which has power at spatial frequencies above the Nyquist frequency that is set by that sampling rate, causes *aliasing*. The effect of aliasing is take power at high spatial (or temporal) frequencies, and erroneously shift it to low frequencies, rather than recording it correctly or simply ignoring it. There are two principal ways to avoid aliasing: one is to *filter* all of the high-frequency information out of the function before it is sampled, and the other is to increase the sampling rate (thus increasing f_c, i.e., k_{max}) so that there is negligible power left in the function above the Nyquist frequency of the finer sampling pattern.

Aliased power can be present in the LIGO simulation program within maps of either the laser fields or the mirrors; but since the maps of "realistic" mirror measurement data are already discretized by the time they are read in by the program (so that the amount of aliased power in them — hopefully small — cannot be altered), the primary type of (k-space) aliasing that could be committed by the program during run-time would be due to inadequate sampling of the propagating e-fields on the grids.

The k-space requirements for adequate gridding can now be stated mathematically, and they are twofold: the grid pixelization must be fine enough so that there is negligible laser beam power above the highest grid spatial frequency, k_{max}; but there must also be enough pixels in the grid (large N_{pix} and x-space window W, i.e., small Δk) for there to be adequate k-space *resolution* for the beams' k-space information to be modelled with high detail, since having all of the beam power covered is not very useful if it is all lumped into,

1. Note that the $-k_{max}$ and k_{max} components are identical for the discrete representation.

say, one or two pixels of the k-space grids. Similarly, such resolution and large-range coverage requirements must also be satisfied in x-space for when the beam data is stored on position-space grids; this will be the subject of the following subsection.

To turn the k-space gridding requirements into practical specifications for LIGO beams, we consider Gaussian-profile laser modes, specifically the TEM_{00} mode, since the excitation laser beam — and nearly all of the circulating power — will be in that lowest mode[1]. An idealized TEM_{00} beam with spot size $w(z)$ and curvature radius $R(z)$ is given by the expression [17]:

$$\mathrm{TEM}_{00}(x, y, z) \equiv \sqrt{\frac{2}{\pi}} \times \frac{1}{w(z)} \times \mathrm{Exp}\left[-\frac{(x^2+y^2)}{w^2(z)} - \frac{ik(x^2+y^2)}{2R(z)}\right] \times \mathrm{Exp}[i(\psi(z) - kz)] \quad (2.30)$$

where $\psi(z) \equiv \tan^{-1}(z/Z_R)$ is the Guoy Phase, and $Z_R \equiv \pi w_0^2/\lambda$ is the Rayleigh Range.

The momentum-space distribution function of this beam is given by its inverse Fourier transform:

$$\mathrm{TEM}_{00}(k_x, k_y, z) = \sqrt{\frac{2}{\pi}} \times \frac{1}{w(z)} \times \mathrm{Exp}[i(\psi(z) - kz)] \quad (2.31)$$

$$\times \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathrm{Exp}\left[-\frac{(x^2+y^2)}{w^2(z)} - \frac{ik(x^2+y^2)}{2R(z)}\right] e^{i(k_x \cdot x + k_y \cdot y)}\, dx\, dy$$

Dropping the functions of z that are outside the integrals (and which only affect overall normalization), and using the fact that the x and y integrals are separable, we can extract the 1-dimensional distribution function for k_x (or, with equal ease, for k_y), as follows:

$$\mathrm{U}_0(k_x, z) \equiv \int_{-\infty}^{\infty} \mathrm{Exp}\left[-\left(\frac{1}{w^2(z)} - \frac{ik}{2R(z)}\right)x^2 + ik_x x\right] dx \quad (2.32)$$

$$= \sqrt{\frac{\pi}{\left(\frac{1}{w^2(z)} - \frac{ik}{2R(z)}\right)}}\, \mathrm{Exp}\left[-\frac{k_x^2}{4\left(\frac{1}{w^2(z)} - \frac{ik}{2R(z)}\right)}\right]$$

where the integration in the last step is computed with help from [42].

---

1. It is sufficient to perform this calculation for the lowest-order TEM_{00} mode, since that is the mode of the starting laser beam, and its generation at the beginning of the run is the only time when discrete sampling of a continuous Hermite-Gaussian function is performed to create an interferometer laser field. Subsequent interactions with the grid-based mirror maps cannot scatter power to values of spatial frequency that are inaccessible to those self-same grids.

Once again ignoring the normalization factor that depends only upon z, we examine the exponential factor. The argument in brackets can be separated into its real and imaginary parts; while the imaginary part only contributes a phase which is irrelevant for our current purposes, the real part of the argument represents the exponential tail-off of power in the momentum distribution function. This real-part factor can be written as follows:

$$\Re\left[-\frac{k_x^2}{4\left(\frac{1}{w^2(z)}-\frac{ik}{2R(z)}\right)}\right] = \Re\left[-\frac{k_x^2\cdot\left(\frac{1}{w^2(z)}+\frac{ik}{2R(z)}\right)}{\left(\frac{4}{w^4(z)}+\frac{k^2}{R^2(z)}\right)}\right] = -\frac{k_x^2}{\left(\frac{4}{w^2(z)}+\frac{k^2w^2(z)}{R^2(z)}\right)} \quad (2.33)$$

Using the formulas $w(z) = w_0\sqrt{1+(z/Z_R)^2}$ and $R(z) = z + Z_R^2/z$ from [17], along with $Z_R \equiv \pi w_0^2/\lambda$ and $k = 2\pi/\lambda$, a little bit of algebra yields the result:

$$-\frac{k_x^2}{\left(\frac{4}{w^2(z)}+\frac{k^2w^2(z)}{R^2(z)}\right)} = \ldots = -\frac{k_x^2 w_0^2}{4} \quad (2.34)$$

where w_0 is minimum spot size ("waist") of the beam in the focal plane (i.e., at $z = 0$).

The considerably simplified result of these calculations is thus:

$$\mathrm{TEM}_{00}(k_x, k_y, z) \propto \mathrm{U}_0(k_x, z)\cdot \mathrm{U}_0(k_y, z) \propto \mathrm{Exp}\left[-\frac{k_x^2 w_0^2}{4}\right]\cdot \mathrm{Exp}\left[-\frac{k_y^2 w_0^2}{4}\right] \quad (2.35)$$

Note that the distribution of power as a function of spatial frequency (k_x and k_y) depends *only* upon the beam's waist size, and not at all upon the spot size $w(z)$, or curvature radius $R(z)$, at any other location along the propagation (z) axis[1]. This result makes intuitive sense, when one considers the beam to be simply a superposition of plane waves with k-space distribution according to $\mathrm{U}_0(k_x, z)\cdot \mathrm{U}_0(k_y, z)$, since free-space propagation of the beam should not alter how much power there is among the various plane wave components, each with a given value of k_x and k_y. On the other hand, these calculations prove a somewhat non-intuitive point: the pixelization of the grids *everywhere in the interferome-*

1. This result may appear to violate the symmetry between x-space and p-space which is discussed at the beginning of this section, since the beam expands in x-space while its p-space distribution remains unaffected. The discrepancy in this symmetry may be due to the fact that it only takes *amplitudes* into account, and beam spot size expansion is an effect generated by differences in the *phases* picked up by the k-space components; or it may be due to the introduction of a new position-space scale length into the problem, namely z, the distance from the focus. Insofar as this result may go against prior expectations, it should be studied further for a satisfactory understanding to be achieved; but in any case, the mathematics presented here clearly leads to the result shown in Eq. 2.35, and interpreted thereafter.

ter must be capable of storing the k-space information present at the *waist* of the beam, in terms of adequate k-space resolution and high-spatial-frequency coverage. And if focusing due to the mirrors changes the beam to a new waist value (even if the focal plane with the adjusted waist size isn't located within the interferometer), then each mirror map must be gridded appropriately for the value of w_0 that the local beam which encounters it *would* have at its focal plane.

An example of a gridding procedure which is not valid, due to this requirement, would be the simple re-pixelization of the grids in order for them to be coarser when the beam spot size is larger, and finer when the beam is near its focus. Although this method would appear to be able to represent widely-varying beam sizes with a fixed number of pixels of varying size (as opposed to using a huge number of pixels and a large window size as a brute-force method of representing small or large spot sizes), a procedure of that type would lose significant high-spatial-frequency beam information while utilizing the coarser grids. Although there are certainly applications in existence which are designed to perform such "dynamic grid-sizing" (we must assume that they use clever routines in order to avoid the problem which we discuss here), our program lacks the capability to simulate a beam which undergoes dramatic spot-size changes due to strong focusing, without using a very large number of pixels (thus causing dramatic increases in run-time) in order to contain the entire beam when it is expanded, while also representing the k-space components of the beam as determined by its w_0 value.

Using Eq. 2.35 (in conjunction with Eq's. 2.28 and 2.29) we can quantify the requirements of adequate k-space resolution and high-spatial-frequency information storage. First, we re-define the "spatial frequency" value of a plane wave beam component as a wave with a propagation angle of θ with respect to the main propagation axis, which is given (to very good approximation for a paraxial beam) as follows[1]:

$$k_x = k\sin\theta \approx k\theta \equiv 2\pi\theta/\lambda \tag{2.36}$$

The power spectrum of the beam goes like the square of Eq. 2.35, and can be expressed as:

$$P_0(k_x) \propto \left|U_0(k_x, z)\right|^2 \propto \mathrm{Exp}[-k_x^2 w_0^2/2] = \mathrm{Exp}[-2(\pi w_0/\lambda)^2\theta^2] \equiv P_0(\theta) \tag{2.37}$$

We can set some approximate point, call it $P_0(\theta_{\text{die-off}}) \equiv e^{-2\Omega}$, to represent the propagation angle for which the power in the collimated laser beam is said to "die off". The constant Ω will be some number of order unity. The typical angle beyond which there is little power in the beam is therefore equal to:

$$\theta_{\text{die-off}} = \frac{\sqrt{\Omega}\lambda}{\pi w_0} \tag{2.38}$$

1. The discussion now considers only one dimension in k-space, e.g., k_x, without loss of generality.

Parenthetically, we note that the fundamental requirement of the Paraxial Approximation, $\sin\theta_{\text{die-off}} \approx \theta_{\text{die-off}}$, means that $\lambda \ll (\pi/\sqrt{\Omega})\, w_0$ must be true; i.e., the beam waist size must be significantly larger than the wavelength of light, an obvious condition to be met for a coherent laser beam.

Now recalling Eq's. 2.28 and 2.29 (and using Eq. 2.36), we can express the maximum propagation angle representable on the k-space grids, as well as $\Delta\theta$ from one pixel to the next, as follows:

$$\theta_{\text{max}} \approx \frac{k_{(x),\,\text{max}}}{k} = \frac{\lambda N_{\text{pix}}}{2W} \tag{2.39}$$

$$\Delta\theta_{\text{per pixel}} = \frac{\Delta k}{k} = \frac{\lambda}{W} \tag{2.40}$$

These values, determined by the pixelization and window sizes of the grids, must *bracket* the value of $\theta_{\text{die-off}}$ specified in Eq. 2.38, for the TEM_{00} beam in the LIGO interferometer. Specifically, for good *resolution* in k-space, we must have:

$$\frac{\theta_{\text{die-off}}}{\Delta\theta_{\text{per pixel}}} = \frac{\sqrt{\Omega}W}{\pi w_0} \gg 1 \tag{2.41}$$

But in order to have coverage for large spatial frequencies (i.e., a large "window" in k-space), we require:

$$\frac{\theta_{\text{die-off}}}{\theta_{\text{max}}} = \frac{\sqrt{\Omega}W}{\pi w_0} \cdot \frac{2}{N_{\text{pix}}} \ll 1 \tag{2.42}$$

These two conditions, manifestly in opposition to one another, can only be satisfied simultaneously if the grids contain a large number of pixels, $N_{\text{pix}} \gg 1$, and if one splits the difference to satisfy Eq's. 2.41 and 2.42 to an approximately equivalent degree.

These gridding requirements are reasonably well satisfied by the parameters given in Table 2.1. With a calculational window width of $W = 70$ cm, a beam waist *radius* of $w_0 \sim 3.5$ cm, $N_{\text{pix}} = 256$, and, say, $\Omega = 2$ (the choice of die-off parameter only affects the results as $\sim\sqrt{\Omega}$), we compute these two k-space ratios as $\theta_{\text{die-off}}/(\Delta\theta_{\text{per pixel}}) = 9 \gg 1$, and $\theta_{\text{die-off}}/\theta_{\text{max}} = .070 \approx 1/14 \ll 1$, thus demonstrating that both requirements are satisfied equivalently and with good margins for error. Also note that there is no significant focusing due to the optics of the interferometer[1], so that the value of w_0 used above is a good estimate of the waist parameter for the beam throughout the entire modelled system.

1. The only exception is in the case of Dual Recycling with a long (2 km) Signal Recycling Cavity, but even in that case, as will be seen, the focusing element used produces a similar waist size.

2.3.4.3 Suppressing "position-space" aliasing

We now turn our attention to the gridding requirements which must be satisfied for manipulating the beams in *position-space*. As before, there are two considerations: adequate resolution of the grids, and adequate x-space coverage with sufficiently large calculational windows.

For the former, a brief consideration shows that no more work is required to guarantee adequate x-space resolution. To achieve good resolution, the individual pixel size must be relatively small compared to waist size of the Gaussian beam, in order to map it out in detail. This criterion is expressed as follows, with Δx equal to the width of a pixel:

$$\frac{\Delta x}{w_0} = \frac{W/N_{\text{pix}}}{w_0} \ll 1 \tag{2.43}$$

But this is the same condition as Eq. 2.42, to within a factor of order unity ($2\sqrt{\Omega}/\pi \cong .9$). This result is no surprise, since the necessity for fine sampling in x-space in nothing more than the statement that high spatial frequency information in the sampled function must be adequately measured, in order to prevent the (k-space) aliasing which could occur when discrete samples are taken of a continuous function, like a beam profile.

In contrast, the one remaining criterion — that of having an x-space window large enough to contain most of the power in the Gaussian-profile beams — is a more complicated issue. Unlike the distribution of beam power in k-space, the x-space power distribution does *not* remain constant during free-space propagations: the beam *expands* as it propagates away from the focal plane. The requirement for making the physical window substantially larger than the beam would therefore involve the spot size at the plane of evaluation, rather than the beam waist size at the focus, as follows:

$$\frac{W}{w(z)} = \frac{W}{w_0\sqrt{1+(z/Z_R)^2}} \gg 1 \tag{2.44}$$

This requirement (minor numerical factors aside) is a more stringent one than its k-space counterpart, Eq. 2.41. It is therefore difficult to model beam with large expansion factors, since one must increase both W and N_{pix} to satisfy Eq's. 2.41–2.44 simultaneously, and increasing the number of pixels greatly increases the computational requirements of a run.

In order to understand the significance of violating such a requirement, consider again the implications of using discrete and finite grids to model the continuous physics of an unbounded beam. The problem of discreteness has been handled above; but because of the *finite number* of grid pixels, there is a finite window of x-space information, and all information outside of this calculational window is left indeterminate. It is *not* assumed to be zero, or tailing off exponentially to zero, as would be most accurate; rather, since beam information is stored on the momentum-space grids in the form of a finite superposition of

sinusoidal components, the default information outside the x-space window for this k-space representation turns out to be oscillatory. In other words, when a beam has been Fourier transformed into k-space for propagation, the lack of information outside the x-space window (equivalent to insufficient k-space resolution) is mathematically equivalent to a default pattern of an infinite series of "phantom" windows lined one after another next to the real window, with each of them being occupied by a phantom Gaussian beam that is a duplicate of the real beam. Propagations can therefore cause phantom power to manifest itself in the simulation results. For example, just as the real beam expands during propagation, and some of it may fall beyond the real window, the phantom beams may spill *onto* the real window, so that the propagation causes unphysical power to enter into the simulated laser beam. This false-power effect, which we will call "position-space aliasing", is depicted in Figure 2.6 as a consequence of Gaussian beam expansion.

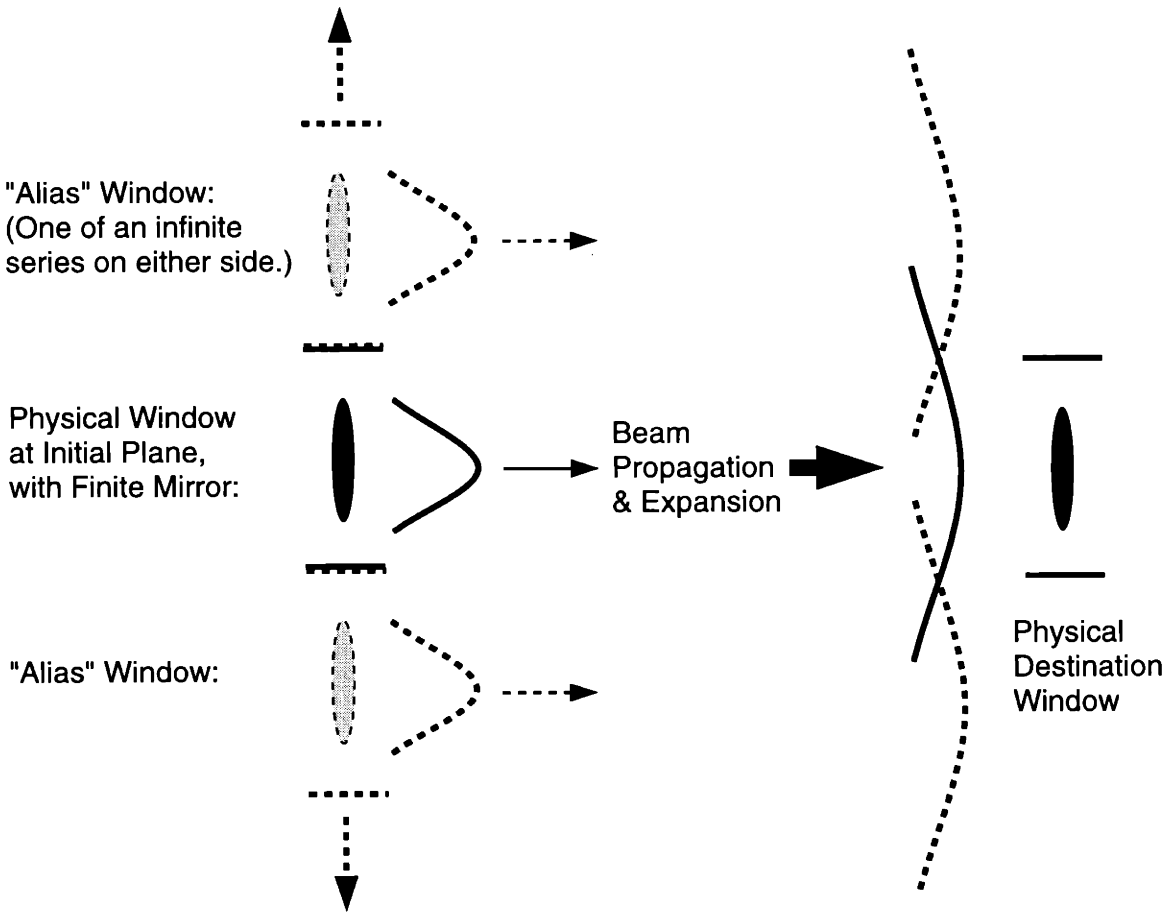

Figure 2.6: An infinite array of phantom windows, causing aliasing via beam expansion.

For the simulation runs that we perform, the largest spot sizes (located at the Fabry-Perot cavity back mirrors) of the beam are 3.638 cm in radius, only ~4% bigger than at the waist, so that position-space aliasing as depicted in Fig. 2.6 is not significant. In any case,

the apertures of the finite-sized mirrors are much smaller than the calculational window size, so that an expanding phantom beam will be eliminated by the apertures before undergoing such aliasing.

More relevant for our considerations, however, is aliasing due to large-angle scatter from high-spatial-frequency mirror deformations. The situation is once again unlike that in k-space: even the highest-spatial-frequency deformation, i.e., a fluctuation over one pixel, cannot scatter power to a k_x (or k_y) value that is higher than k_{max}; but such deformations easily scatter power out to very large angles, so that they pass x_{max} and exit the window during the propagation, while aliased power enters from the other side. It is even possible that the scattering angle is large enough such that the aliased power can make the trip from within the aperture of a phantom mirror to within the aperture of the real mirror, so that the finite apertures are not sufficient to remove this power from the system. This effect turns out to be very significant for our simulation runs, since mirrors with realistic deformations will have very fine structure, scattering power out to large angles, the effects of which we must simulate.

Position-space aliasing due to large-angle scattering is shown in Figure 2.7. The angles in the diagram are defined as follows: θ_p is the *largest angle* that physical information can travel at from anywhere in the initial (real) mirror aperture and still enter the (real) mirror aperture at the destination window; and θ_a is the *smallest angle* that alias information can travel at from anywhere in the initial (phantom) mirror aperture and still enter the (real) mirror aperture at the destination window. Obviously, one wishes to model the system such that as much as possible of the physical information, $\theta \leq \theta_p$, is retained in the model, while as much as possible of the aliased information, $\theta \geq \theta_a$, is eliminated or not modelled.

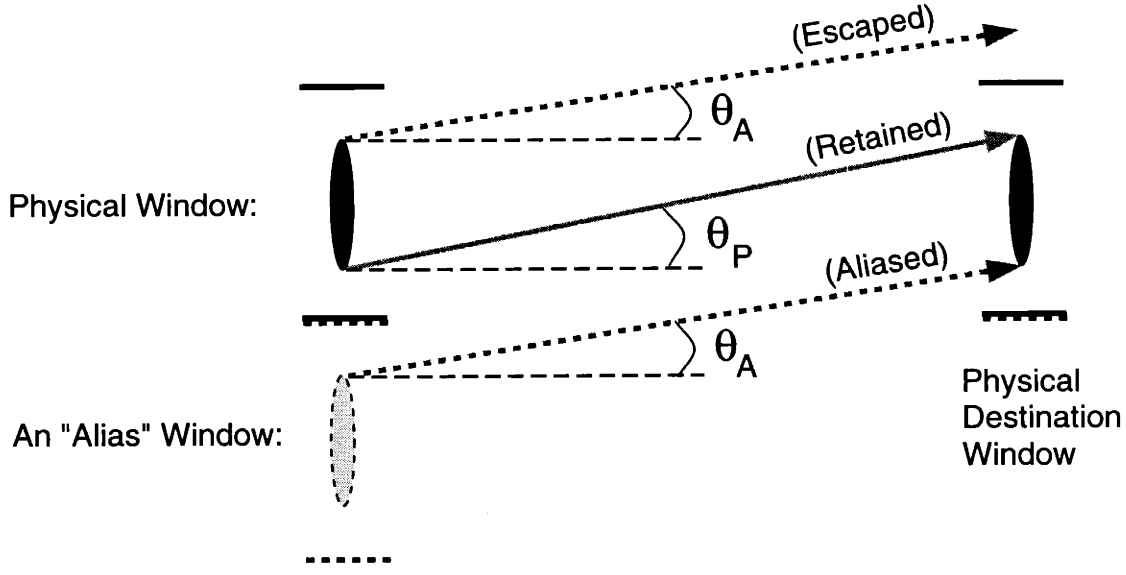

Figure 2.7: Aliasing due to large-angle scattering from fine-scale mirror deformations.

The angles θ_p and θ_a are determined by three parameters: the width of the calculational windows (W), the aperture diameter of the mirrors (A), and the length of the propagation (L) that separates the initial plane from the destination plane. Using small-angle approximations, we get:

$$\theta_p \approx \frac{A}{L} \tag{2.45}$$

$$\theta_a \approx \frac{W-A}{L} = \frac{W}{L} - \theta_p \tag{2.46}$$

In addition, recall Eq. 2.39 for the maximum propagation angle representable on the grids:

$$\theta_{max} \approx \frac{N_{pix}}{2} \cdot \frac{\lambda}{W} \tag{2.47}$$

In general, these three angles can have any relationship with respect to one another. If the inequality $\theta_a < \theta_{max}$ is true, then propagations are bound to introduce aliased power into the interferometer; and if $\theta_a < \theta_p < \theta_{max}$ is true, then it becomes impossible to completely separate the aliased power from the real power in the system. In any case, this will only be a problem in the Fabry-Perot arm cavities where L is large, since $\theta_{max} \ll \theta_a, \theta_p$ elsewhere.

Given fixed choices of the physical parameters A and L, one apparent method of enforcing $\theta_a > \theta_p, \theta_{max}$ is by making W large (while increasing N_{pix} to keep N_{pix}/W constant); if the calculational window is large enough, then aliased power will not be able to make it to within the real mirror aperture. To increase W, one adds a framework of pixels around the original grids, which is "dead space" because the power out there is eliminated every time an aperture is encountered. This method is called "zero-padding" [17], and theoretically it can be made 100% effective against position-space aliasing, as long as W is set large enough.

The serious drawback with zero-padding is that it is very expensive in time and computer memory. With FFT's computed at a speed proportional to $N_{pix}\log(N_{pix})$, increasing the 2-D grids by 2 along each dimension increases run time by a factor of more than 5; a day of running time becomes almost a week. Also, the amount of data stored in each grid would increase by a factor of 4, and the simulation program manipulates very many of these grids during a run, so that RAM limitations (on workstations) and disk space for storage of the computed e-fields become important factors. Due to the run time increase alone, zero-padding is not a method that is accessible to us for eliminating position-space aliasing from our runs.

Instead, we have implemented an alternative method for controlling aliasing: *momentum space beam filtering*. As its name implies, this method consists of taking the e-fields that have been transformed into k-space, and zeroing-out or reducing the large-spatial-frequency components which give rise to position-space aliasing. Essentially, this adds an intermediate step to the propagation procedure given in Sec. 2.3.2, as follows:

Step 2.5: Multiply each pixel in k-space by an aliasing-filter factor:

$$\tilde{E}(k_x, k_y, L)_{\text{Anti-Alias}} = \tilde{E}(k_x, k_y, L) \times F(k_x, k_y) \tag{2.48}$$

This filter function can be a simple function of zeroes and ones if $\theta_a > \theta_p$; but if $\theta_a < \theta_p < \theta_{max}$ is true, then a smoothing process must be done which makes a compromise between eliminating aliased power and retaining the real power.

From Eq. 2.29 and Eq. 2.36, one can express the propagation angle represented by a the n^{th} pixel in the k-space grids:

$$\theta_n \approx n \cdot \frac{\lambda}{W}, \qquad n = -\left(\frac{N_{pix}}{2} - 1\right), \ldots, \frac{N_{pix}}{2} \tag{2.49}$$

This can be combined with Eq's. 2.45 and 2.46 to derive the (integer) pixel numbers, N_p and N_a, which represent the angles, θ_p and θ_a, respectively:

$$N_p = \text{Int}\left[\frac{N_{pix}}{2} \cdot \frac{\theta_p}{\theta_{max}}\right] = \text{Int}\left[\frac{AW}{\lambda L}\right] \tag{2.50}$$

$$N_a = \text{Int}\left[\frac{N_{pix}}{2} \cdot \frac{\theta_a}{\theta_{max}}\right] = \text{Int}\left[\frac{(W-A)W}{\lambda L}\right] \tag{2.51}$$

The filter function of Eq. 2.48 is therefore designed to obey the following rules:

(i) For $n < N_a$, set $F(k_x, k_y) = 1$ for all pixels,

(ii) For $n > N_a, N_p$, set $F(k_x, k_y) = 0$ for all pixels,

(iii) For $N_p \geq n \geq N_a$ (the *overlap region*, if it exists), apply a smooth, asymptotic cut-off which removes more of the power as n, θ_n get larger.

In the runs which we demonstrate in this thesis, the typical parameters are A = .24 m, W = .70 m, λ = 1.064 μm, L = 4000 m, and N_{pix} = 256, which yields $N_p \sim 39$ and $N_a \sim 75$. In these runs, we are fortunate since $N_a > N_p$; but for earlier runs with Argon-ion laser light (λ = 514 nm) and 128 x 128 grids over a .35 x .35 m window, there was a significant overlap region within which we set $F(k_x, k_y)$ to be an exponentially-decreasing apodizing function bounded by values of 1 at $N = N_p$ and ~0 at $N = N_a$.

In the absence of such anti-aliasing procedures, the symptoms of position-space aliasing are the persistence of unphysical beam modes in the cavities of the interferometer. These modes, representing very high spatial frequencies and large-angle scattering, do not get apertured out of the system, but rather build up as background fields which corrupt the resonant beams and draw power from the lowest-order mode. While these background fields are small compared to the power of the resonant TEM_{00} mode in the Fabry-Perot

arms, they leak out easily from the dark port of the beamsplitter, sometimes nearly doubling the total power in the imperfectly-dark fringe, so that a simulation run without anti-aliasing procedures cannot be trusted to predict the true amount of power emerging from the interferometer exit port.

Alias modes are distinct in terms of their modal composition. When typical exit port fields *without* aliasing are decomposed into Hermite-Gaussian modes (as per Appendix A), most of the power is usually located in the first few low-order modes. But we have found that no matter how many modes one chooses to decompose a beam affected by aliasing into (up to a hundred or more), aliased power remains almost totally unaccounted for. This can be understood by looking at Figure 2.8, the real part of an e-field built-up in a Fabry-Perot arm cavity (during runs that will be discussed in Sec. 2.6.2), that is generated exclusively by the effects of position-space aliasing. This e-field is striking in appearance: wild (phase) fluctuations from pixel to pixel are apparent, thus implying structure of the highest spatial frequencies possible on the grids. In comparison to the smooth (TEM_{10}) field shown in Fig. 2.3, this is very clearly a mode that could not have been built up in a real cavity system with apertured mirrors.

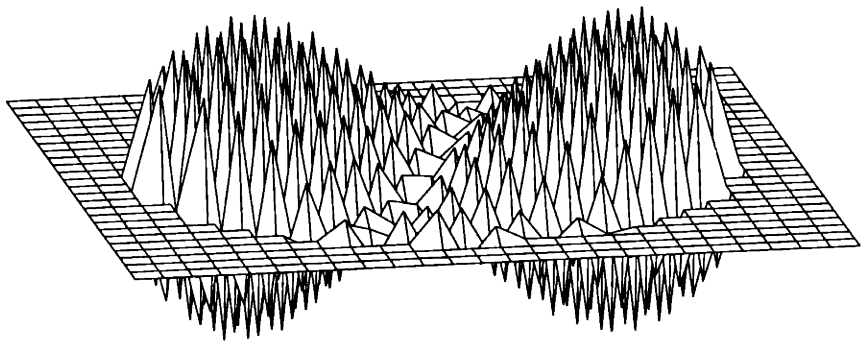

Figure 2.8: The real part of a Fabry-Perot cavity field generated via position-space aliasing. The outer portion of the map is clipped for visual clarity.

To test our anti-aliasing procedures, we performed runs with our old (Argon light) parameters for three cases: 128 x 128 grids over a .35 x .35 m window, with and without anti-aliasing, and for 256 x 256 grids over a .70 x .70 m window for which anti-aliasing was unnecessary (since $\theta_a > \theta_{max}$). All three runs were performed using a set of $\lambda_{(YAG)}/1800$ mirror surfaces and deformed substrates. For comparisons between the runs, we looked at the (average) TEM_{00} carrier power built up in the Fabry-Perot arms, the carrier Contrast Defect (a measure of how dark the dark-fringe really is), and the projected GW-sensitivity at 100 Hz, $\tilde{h}_{SN}(100)$, given the shot noise level at the dark- (signal-)port.

The results are presented in Table 2.3. These results clearly show that the anti-aliasing procedure, if performed with carefully chosen parameters, succeeds in nearly duplicating the (foolproof) effects of zero-padding, i.e., the 256 x 256 grid run. These results also show that the effects of position-space aliasing, if left unchecked, cause significant errors

Quantity	128 x 128 grid (no anti-aliasing)	128 x 128 grid (with anti-aliasing)	256 x 256 grid ("zero-padding")
Arm Cavity Power	1994.2	2011.3	2011.7
Contrast Defect	1.298×10^{-3}	7.944×10^{-4}	7.954×10^{-4}
GW-strain, $\tilde{h}_{SN}(100)$	9.192×10^{-24}	8.944×10^{-24}	8.943×10^{-24}

Table 2.3: Comparison of anti-aliasing and zero-padding runs versus an unmodified run.

in the simulation-based estimations of LIGO performance with realistically-deformed mirrors.

Since we have implemented anti-aliasing in our simulation runs, the existence of "unaccounted for" power in the dark-fringe carrier fields has been virtually eliminated for modal decompositions using a reasonable number of modes (e.g., the lowest 45 modes).

2.4 Relaxation methods for steady-state e-field solutions

As discussed in Sec. 2.1.1, the simulation program models a *static* LIGO interferometer, one which is assumed to be held to the correct operating point for an indefinite period of time. This assumption enormously simplifies the program and our overall simulation task, while still enabling us (see Appendix D) to compute the frequency-dependent GW-response of the interferometer. The primary task of the simulation program, therefore, is to compute the relaxed, steady-state, resonant electric fields that build up in each part of the interferometer, when it is excited by a laser beam of fixed amplitude, orientation, and frequency (at the nominal carrier & sideband frequencies) entering through the Power Recycling Mirror. This section will outline the process of e-field relaxation that is performed.

2.4.1 Steady-state equations for a system with cavities

Observe Figure 2.9, which is a diagram of the interferometer with many of the propagation, reflection, and transmission operators needed for e-field relaxation shown. The physics of these operators is as described above in Sections 2.3.2 and 2.3.3. Some e-fields are also shown, as in Fig. 2.2, but in this case only the fields necessary for computation of the interferometer steady-state behavior are included; as will be seen, all other e-fields of interest can be simply derived from them. A Dual-Recycled interferometer is depicted, which is the most general case; appropriate relaxation formulae for an Initial-LIGO system can be obtained from the general case by setting R_6 to zero and T_6 to unity, though there are some algorithmic differences between the way the two systems are handled in the simulation code.

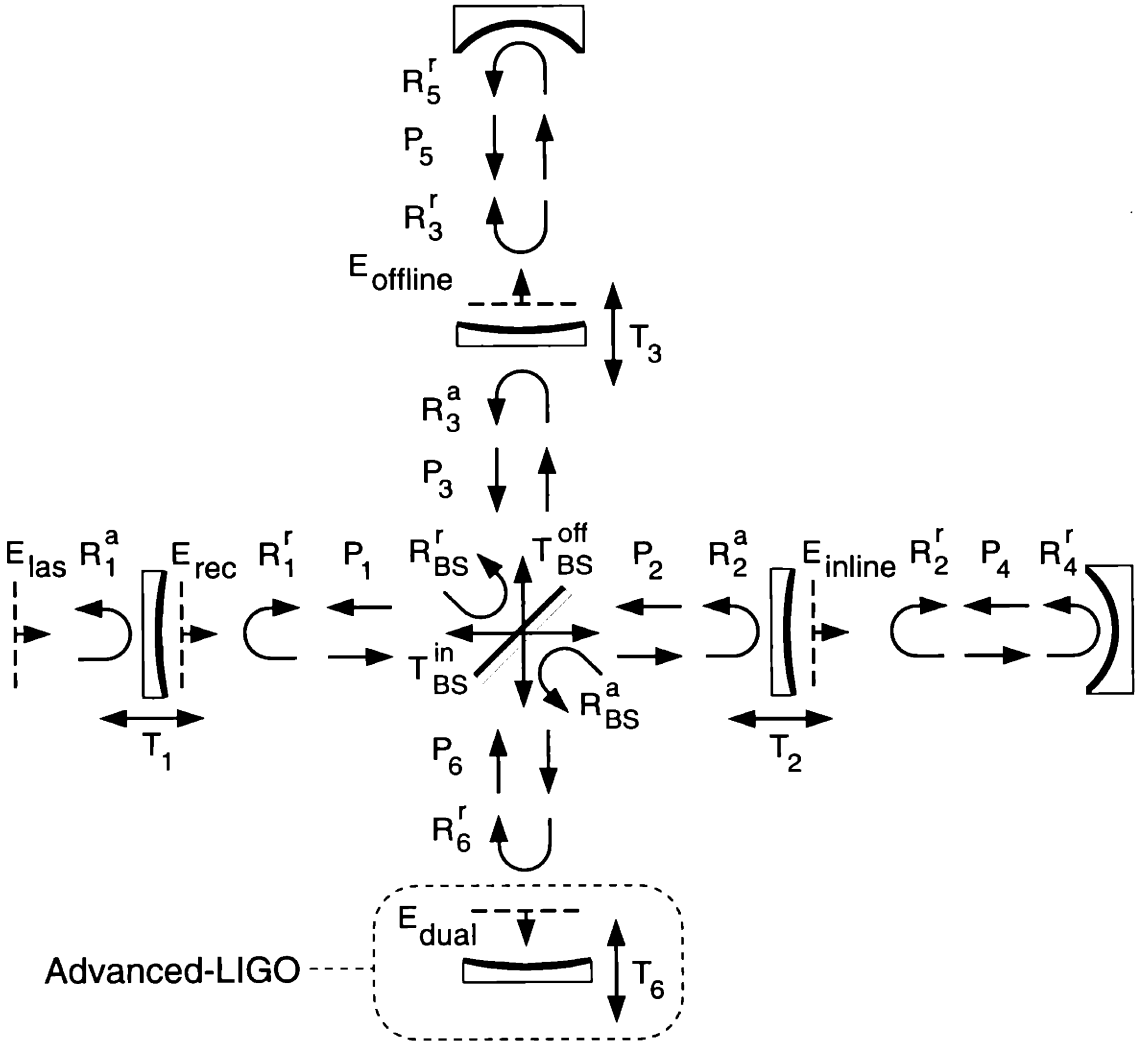

Figure 2.9: A (Dual-Recycled) LIGO interferometer shown with the propagation and mirror interaction operators that are used for computing the relaxed electric fields.

The full steady-state equations relating each of the e-fields in the interferometer to one another, which must be achieved by the e-field relaxation algorithm, is not presented here; but as an illustrative example, consider the steady-state equation for $\vec{\tilde{E}}_{\text{inline}}$:

$$\vec{\tilde{E}}_{\text{inline}} = (R_2^r \cdot P_4 \cdot R_4^r \cdot P_4) \cdot \vec{\tilde{E}}_{\text{inline}} \qquad (2.52)$$
$$+ (T_2 \cdot P_2 \cdot T_{\text{BS}}^{\text{in}} \cdot P_1) \cdot \vec{\tilde{E}}_{\text{rec}} + (T_2 \cdot P_2 \cdot R_{\text{BS}}^a \cdot P_6 \cdot R_6^r) \cdot \vec{\tilde{E}}_{\text{dual}}$$

(Note that these e-field operators are written out as working on the quantities to their right, so that right-to-left for each term represents the forward passage of time. Also note that the same propagation operator is used for bi-directional beam propagation: the propagator

only depends upon the distance travelled, not the direction.)

While Eq. 2.52 may appear mathematically complicated, the physical process that it represents is quite simple. The steady-state field $\hat{E}_{inline}$ must be equal to a superposition of three elements: the self-same field ($\hat{E}_{inline}$) sent on a round-trip through the arm cavity, added to the portions of $\hat{E}_{rec}$ and $\hat{E}_{dual}$ which are channelled into the inline arm cavity by the interferometer optical configuration, and thereby excite that arm cavity. The necessity of obeying an relationship like Eq. 2.52 is just an expression of the fact that an e-field at a given location must be *equal* to the sum of all (steady-state) fields entering that location, for an overall steady-state condition to exist.

This formula for $\hat{E}_{inline}$, as well as the steady-state equations for all other interferometer e-fields, can be recast into the following generic form:

$$\hat{E}_{steady\text{-}state} = R.T.\{\hat{E}_{steady\text{-}state}\} + \hat{E}_{exc} \tag{2.53}$$

Because $\hat{E}_{steady\text{-}state}$ is located within a cavity, it depends upon itself (i.e., it is "self-coupled") via some form of round-trip reflection, as well as depending upon[1] all of the fields which excite the cavity it resides in. For $\hat{E}_{inline}$ and $\hat{E}_{offline}$, $\hat{E}_{exc}$ has contributions from $\hat{E}_{rec}$ and $\hat{E}_{dual}$, while for $\hat{E}_{rec}$ and $\hat{E}_{dual}$, $\hat{E}_{exc}$ has contributions from $\hat{E}_{inline}$ and $\hat{E}_{offline}$. In addition, $\hat{E}_{rec}$ has a unique contribution to $\hat{E}_{exc}$ from the input laser field, $\hat{E}_{las}$. Note that $\hat{E}_{rec}$ and $\hat{E}_{dual}$ do not excite each other directly, nor do $\hat{E}_{inline}$ and $\hat{E}_{offline}$.

Because these four interferometer fields ($\hat{E}_{rec}$, $\hat{E}_{inline}$, $\hat{E}_{offline}$, and $\hat{E}_{dual}$) are self-coupled, they cannot simply be written down as functions of $\hat{E}_{las}$, or as a function of the other three fields; they are irreducible, and must (as we will show) be solved for via an iterative process. Generally speaking, there will be one self-coupled e-field to be iteratively solved for (i.e., relaxed) per cavity. In the Initial-LIGO interferometer, this means that a total of three e-fields must be relaxed: one in the Power Recycling Cavity, and one in each of the Fabry-Perot arm cavities. For a Dual-Recycled interferometer, a fourth e-field will be relaxed which must be located in the Signal Recycling Cavity.

Other than the requirement of having one relaxed e-field inside each cavity, the precise locations and propagation directions of those e-fields are arbitrary. For our simulations, we have chosen the relaxation fields to be $\hat{E}_{rec}$, $\hat{E}_{inline}$, $\hat{E}_{offline}$, and (for Dual Recycling) $\hat{E}_{dual}$, as depicted in Fig. 2.9. Once these e-fields are fully computed, all other interferometer e-fields (including the output field exiting the interferometer) can be obtained from them by propagating each of the relaxed e-fields to the interferometer location of interest, and then summing their contributions together to produce the total, steady-state e-field at the new location.

---

1. A field "$\hat{E}_A$" will be dependent upon field "$\hat{E}_B$" if there is a path to the location of $\hat{E}_A$ (bearing the same propagation direction as $\hat{E}_A$) that comes from the location of $\hat{E}_B$ (bearing the same propagation direction as $\hat{E}_B$), that does *not* pass through the location of any other field which $\hat{E}_A$ is said to depend upon.

Not discussed so far is how one solves the steady-state equation, Eq. 2.53, for each of the interferometer e-fields. Our procedure for iterative e-field relaxation will be the topic of Sections 2.4.3 and 2.4.4. First, however, we will justify the iterative approach.

2.4.2 Iterative vs. non-iterative solution methods

Technically speaking, all of the interferometer fields are physically determined by only a single electric field: the excitation laser field, $\vec{\mathrm{E}}_{\mathrm{las}}$. The discussion of Sec. 2.4.1 notwithstanding, it is mathematically possible to use non-iterative methods to solve for all of the steady-state e-fields. We discuss two such methods here, and our justification for not utilizing them.

The first non-iterative method is the most straightforward and brute-force approach: one simply propagates the laser beam into the interferometer, and then follows that beam — and all derivative beams generated from it by reflections and transmissions — as they propagate through the interferometer cavities, until they each eventually decay away to negligible power. One then sums over all such beam components (keeping track of their relative phases) in order to form the composite field at any given interferometer location, at a specifiable time.

Simulations of this kind have indeed been done successfully [43] for coupled-cavity interferometers (if not for the full Dual Recycling case with Fabry-Perot arm cavities), but not for models where each e-field contained the tremendous amount of spatial beam information that is necessary for simulating the effects of realistically imperfect optics. The brute-force procedure simply appears infeasible for the moderately-high finesse (i.e., many bounces before power decay) cavity system of the full-LIGO interferometers, given the fine spatial resolution for mirror and e-fields that our research requires.

The second non-iterative method is simply to invert the steady state equation, Eq. 2.53, as follows:

$$\vec{\mathrm{E}}_{\text{steady-state}} = [I - \mathrm{R.T.}]^{-1} \cdot \vec{\mathrm{E}}_{\mathrm{exc}} \tag{2.54}$$

where I is the identity operator, and R.T. is the composite "round-trip" operation that is undergone by this particular $\vec{\mathrm{E}}_{\text{steady-state}}$ in the interferometer. In the case of $\vec{\mathrm{E}}_{\mathrm{inline}}$, we can use Eq. 2.52 to write:

$$\begin{aligned} \vec{\mathrm{E}}_{\mathrm{inline}} &= [I - (R_2^{\mathrm{r}} \cdot P_4 \cdot R_4^{\mathrm{r}} \cdot P_4)]^{-1} \cdot \\ &[(T_2 \cdot P_2 \cdot T_{\mathrm{BS}}^{\mathrm{in}} \cdot P_1) \cdot \vec{\mathrm{E}}_{\mathrm{rec}} + (T_2 \cdot P_2 \cdot R_{\mathrm{BS}}^{\mathrm{a}} \cdot P_6 \cdot R_6^{\mathrm{r}}) \cdot \vec{\mathrm{E}}_{\mathrm{dual}}] \end{aligned} \tag{2.55}$$

This procedure defined by Eq. 2.54 has been demonstrated by Vinet and Hello [19], in which they transformed their e-fields into discrete arrays of Hankel coefficients, then found the transform-space propagation kernels, derived the complete matrix operator for a

Fabry-Perot cavity round-trip, and inverted it to solve for the steady-state cavity field. They found their matrix approach to be about a thousand times faster than iteratively solving for the cavity fields with their version [22] of the FFT-based e-field propagation program. To use their (1-D) Hankel transforms, however, they assumed circular symmetry throughout, thus reducing the number of dimensional indices (for physical space and the transform space) from 2 to1. Thus they required only rank-2 tensors to move between physical space and the transform space. In our simulation program, we make no *apriori* symmetry assumptions about the mirrors or fields, and so we require two-dimensional arrays for e-field information, and rank-4 tensors to move between physical space and momentum space. Inverting round-trip operators which are as large as rank-4 tensors, with many pixels on a side (typically as many as 256) may well be a less feasible approach than our current method of performing an iterative e-field relaxation scheme. Naturally, if some modal or functional representation is found [28] which encodes as much of the important beam and mirror structure information as our 2-D spatial grids do, while using considerably fewer than 256 x 256 coefficients, than an inversion scheme like Eq's. 2.54 and 2.55 might be more feasible; but such a representation remains to be demonstrated.

A potentially more serious problem with solving for the steady-state fields using this operator inversion method, is the presence of finite-sized mirror apertures in the real LIGO interferometer. Modelling such apertures is a necessary part of a LIGO simulation program, since finite apertures have the important effect of trimming high spatial modes and preventing them from accumulating in the interferometer. Light falling beyond these apertures is completely and irreversibly removed from the system; and once this information is "erased", it cannot be re-created by the inversion of any interferometer matrix operator. In the program, therefore, they represent a spatially-inhomogeneous loss of information which turns the method of e-field relaxation via operator inversion into a nonsingular problem.

2.4.3 Utilizing a fast iteration scheme

To specify a convergence scheme for e-field relaxation, one creates a prescription for how the N+1th iterative guess for a given e-field is obtained from its Nth iteration (and from the Nth iteration of the other e-fields being relaxed simultaneously). The simplest possibility is to make the choice:

$$\vec{\tilde{E}}^{(N+1)} = \text{R.T.}\{\vec{\tilde{E}}^{(N)}\} + \vec{\tilde{E}}^{(N)}_{\text{exc}} \tag{2.56}$$

This process is repeated for as many iterations as necessary, until the change in going from $\vec{\tilde{E}}^{(N)}$ to $\vec{\tilde{E}}^{(N+1)}$ — which is equal to the amount by which the steady-state equation, Eq. 2.5.3, is violated — is reduced to within a pre-specified threshold of accuracy (typically 1 part in 10^4 in power, for our runs). When the iteration process is complete, $\vec{\tilde{E}}^{(N+1)} \cong \vec{\tilde{E}}^{(N)} \equiv \vec{\tilde{E}}_{\text{steady-state}}$.

This relaxation formula is guaranteed to smoothly converge an e-field to its correct steady-state form[1], because it is little more than a simulated version of 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 to good effect by previous researchers [22].

Conversely, since the method of Eq. 2.56 does indeed model the true physical buildup of power in a system of cavities, it requires a great many iterations to converge, especially for coupled-cavity systems with Q-factors that are not small, since the number of iterations prior to convergence is directly related to how many reflections and round-trips transient e-fields can persist for, while retaining significant power, before dying away. In the case of a full-LIGO simulation program, the iteration scheme prescribed here represents a relaxation routine that would be overwhelmingly slow to complete.

Therefore, as first suggested (and implemented, for a simpler cavity arrangement) by a LIGO colleague [44], our simulation program uses a different approach. Instead of Eq. 2.56 for choosing the N+1th iteration, we use the following expression:

$$\vec{E}^{(N+1)} = a \cdot \vec{E}^{(N)} + b \cdot (\text{R.T.}\{\vec{E}^{(N)}\} + \vec{E}_{exc}^{(N)}) + c \cdot \vec{E}_{exc}^{(N)} \tag{2.57}$$

where (a, b, c) are unknown, complex coefficients[2] that are solved for by minimizing the error in what the steady-state equation will become in the *next* round:

$$\sigma^{(N+1)}(a, b, c) \equiv \left\| \vec{E}^{(N+1)} - (\text{R.T.}\{\vec{E}^{(N+1)}\} + \vec{E}_{exc}^{(N+1)}) \right\| \tag{2.58}$$

This 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 can be solved for via a matrix inversion operation. Compared to the simpler scheme of Eq. 2.56, the "abc" relaxation method derives a huge advantage from having these additional, useful degrees of freedom available for determining the next iteration. Many fewer iterations will be necessary to achieve convergence, therefore resulting in much faster e-field relaxation. The additional degrees of freedom are especially important for cases with significantly deformed optics, since the steady-state fields for those runs will differ significantly from the initial-guess ($\vec{E}^{(1)}$) fields[3], and thus substantial evolution of the e-field during relaxation is required for convergence to be achieved.

2.4.4 Computational and stability issues for fast relaxation

There are a number of important differences between the simple convergence scheme, and

1. Barring any complications which could arise from the dynamic interferometer parameter optimizations; see Sec. 2.5.
2. This method reduces to the simpler method of Eq. 2.56 for the special case of (a,b,c)=(0,1,0).
3. The initial-guess e-fields are pure TEM_{00} beams that are mode-matched into the interferometer, and multiplied by analytical estimates of cavity power gain, so that they are close to being solutions for an interferometer with perfect mirrors.

the fast ("abc") convergence scheme, which must be discussed. First of all, the abc algorithm requires much more information to be stored during each iteration, in terms of interim e-fields, to perform the error evaluation necessary for the optimization of the (a, b, c) coefficients; for example, Eq. 2.58 requires the computation of terms such as $\text{R.T.}\{\vec{\tilde{E}}^{(N+1)}\} = (\text{R.T.}\{\text{R.T.}\{\vec{\tilde{E}}^{(N)}\}\} + \ldots)$, etc. The abc method therefore uses more computer memory than the simple convergence scheme. In addition, each abc iteration requires more e-field propagations, so that it takes longer to perform than an iteration for the simpler relaxation algorithm. Nevertheless, this effect is not nearly severe enough to counter the overall speedup of doing a much smaller number of iterations.

A more significant difference for the abc convergence method is that it does not directly model the physical process of power build-up in the interferometer. As a result, there is some sacrifice of convergence stability, because the (a, b, c) coefficients may be poorly chosen for some iterations. In practice, this is due to the effects of finite numerical accuracy in the matrix-inversion step that is performed to solve for the minimization of the error function, $\sigma^{(N+1)}(a, b, c)$, and also due to the fact that large excursions from self-consistent physical states are possible because of the unconstrained way in which (a, b, c) are chosen. As a result, optics with large deformations, or even just significant errors in the initial power estimates that are made by the program in order to compute $\vec{\tilde{E}}^{(1)}$, can cause the abc relaxation process to diverge.

To enhance the stability of the abc relaxation algorithm, we have found that placing hard limits on the allowed choices of (a, b, c) during the early stages (e.g., first 75 iterations) of runs[1] allows the relaxation process to settle down, so that it can converge cleanly and quickly to the steady-state solution. These limits must not be so tight as to completely inhibit the fast relaxation process, which involves substantial changes to the electric fields (and often dramatic, temporary plunges in power during the early stages), but they must be tight enough to contain the process until it stabilizes. When the limits are well-chosen, and applied to each of the e-fields being relaxed in the interferometer[2], this is a very effective method for achieving convergence stability.

At least part of the stability problem mentioned here arises because abc-based iterations are performed simultaneously and independently for the three e-fields being relaxed in the (Initial-LIGO) interferometer. The next-guess field in one location is therefore unaware of the next-guess changes made in the other locations, and the changes to all fields are not mutually optimized. This tends to create a "power oscillation" between the Power Recycling Cavity and the Fabry-Perot arms during the convergence process, which usually decays away, but for runs with strongly deformed mirrors it may grow, leading to

1. These hard limits on (a,b,c) are useful for runs with significantly deformed mirrors, but are counter-productive for runs with perfect or nearly-perfect mirrors.
2. The radio-frequency sideband field in the Power Recycling Cavity does not require such limits, since it does not use the abc method anyway; we have seen that it is relaxed better with the simple convergence scheme, i.e., (a,b,c) = (0,1,0).

divergence. This instability was not relieved by simple expedient of turning off the abc process during alternate iterations, to allow the fields in different locations to be updated out-of-step with one another, thus delaying an e-field's iterative change until the next-guess information from the other e-fields can arrive. That approach merely slowed the convergence process down, and showed that a more general solution was needed to completely ensure stability of the abc method.

In response, former LIGO scientist Yaron Hefetz suggested a "global" convergence process, in which the (a, b, c) coefficients for all relaxed fields are chosen simultaneously by minimizing a global error function that is a weighted average of $\sigma^{(N+1)}(a, b, c)$ for each of the relaxed e-fields. This global convergence method is a more complex process than the original abc algorithm, with many more propagations and more interim e-fields being stored than before (so that even more memory is required); but in addition to enhancing stability (and obviating the need for limits on the (a, b, c) coefficients), it further speeds up the relaxation process by causing all interferometer e-fields to converge together in a more coherent process. We have implemented the global convergence algorithm in our code for Dual Recycling simulations, and it works as well as was expected.

In summary, despite the caveats presented here, our fast, "abc" relaxation scheme is a sufficiently reliable and *extremely fast* convergence algorithm that greatly reduces program execution time, thus enabling us to perform full-LIGO simulation runs with realistic optics over a wide space of physical and computational parameters.

2.5 Parameter adjustments for sensitivity optimization

At the heart of our efforts to make a realistic simulation of a LIGO interferometer are several routines which bring the system into an "optimal" configuration for signal detection. The problem of optimization is a highly nontrivial matter, due to the complex nature of the system being modelled. Not only does the interferometer possess many, often inter-dependent degrees of freedom which must be optimized, but each evaluation of performance (i.e., GW-sensitivity) as a function of these optimizable parameters is extremely time consuming, since it involves the complete (or near-complete) relaxation of the carrier and sideband electric fields everywhere in the interferometer. It is therefore 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, our program implements a strategy very similar to that which will be used for the real LIGO: each parameter is optimized separately, using some error signal (that is brought to zero), or some function of merit (to optimize), that is strongly depen-

dent upon that one parameter only, and which is a true measure of when the parameter is well-chosen for the maximization of GW-sensitivity.

We have found this to be a successful strategy, and it has been aided by the remarkable fact that the interferometer's (shot-noise-limited) GW-sensitivity function is quite insensitive to particular *combinations* of parameter changes, such that if one of the optimizable parameters are displaced from the "global" maximum in the multidimensional parameter space, then other parameters will adjust themselves (via our optimization procedures) to compensate with virtually no reduction in interferometer sensitivity. This effect (an example of which will be given in Sec. 3.3) is in fact of great use to LIGO itself, since it implies that LIGO will suffer less from interferometer degradations than might otherwise be expected; this is because combinations of parameters can be adjusted in concert to make the system less vulnerable to a particular optical imperfection that may be strongly present in a particular interferometer. Observing that this compensation effect actually occurs, and is pronounced, has been an important success of our LIGO modelling initiative [45].

In some sense, the complex task of parameter optimization arises from the physical realism of the simulation program. For example, rather than specifying "by proclamation" that the cavities are sitting exactly on their various, desired resonance conditions, the program employs real, physical cavity lengths and real beam frequencies, and must then tune the cavity lengths and/or frequencies to achieve these resonance conditions. Thus the use of effective (and fast) parameter optimization routines allows our program to model a more detailed and realistic system.

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. These optimizations are done at the same time that the electric field relaxations are occurring, so that the final product which emerges from the iterative relaxation process are the steady-state electric fields of a *fully-optimized* interferometer.

2.5.1 Interferometer laser frequencies and resonance conditions

We begin the discussion of parameter optimizations with a fuller introduction to the control systems of the LIGO interferometers. The Initial-LIGO and Dual-Recycled LIGO systems to be discussed here are as depicted (respectively) in Figures 1.5 and 1.6; and the variables representing physical quantities or interferometer parameters are as shown previously in Fig. 2.1. Fuller mathematical descriptions of the signal detection process and the coupled-cavity behavior of these interferometers are given elsewhere [e.g., 5, 71], but here we will utilize simplifications that are useful in providing an understanding of each individual optimization routine that is carried out during the interferometer simulation runs.

First, we discuss the Initial-LIGO interferometers. For this configuration, electro-optic modulation of the carrier frequency beam is carried out prior to injection of the laser into the interferometer, to create radio frequency (RF) sidebands which will serve as a local oscillator for a heterodyne detection scheme that produces an output signal linear in h, the

dimensionless amplitude of the GW. The modulation frequency (~few tens of MHz) is high enough to be above various technical noises which dominate at low frequencies (e.g., laser amplitude noise), and the final signal which is obtained via demodulation at the signal port will be retrieved free of those low-frequency noise sources [46].

In the detection system modelled here, the carrier light has a two-stage resonance in the Power Recycling Cavity (i.e., the PRC, consisting of L_1 plus L_2/L_3) and in the Fabry-Perot arm cavities. The RF sidebands, on the other hand, are resonant only in the PRC (to maximize the power available for the heterodyne detection scheme), and will be far off from resonance in the arm cavities so that they may serve as a stable reference which is unaffected by gravitational waves[1]. Assuming that the carrier is held to its double resonance, the RF-modulation frequency must therefore satisfy the relations:

$$2\,k_{\mathrm{Mod}}\,L_{\mathrm{Arm}} \approx N_{\mathrm{Odd}} \cdot \pi \tag{2.59}$$

$$2\,k_{\mathrm{Mod}}\,L_{\mathrm{PRC}} = N_{\mathrm{Odd}} \cdot \pi \tag{2.60}$$

where L_{Arm} is the length of either arm cavity (L_4 or L_5), L_{PRC} is the length of the PRC *averaged* between the inline and offline arms (L_2 and L_3 differ; see below), and the modulation wavenumber is given by:

$$k_{\mathrm{Mod}} = \frac{2\pi}{c} \cdot \nu_{\mathrm{Mod}} \equiv \frac{2\pi}{c} \cdot \left|\nu_{\mathrm{Carr}} - \nu_{\mathrm{SB}}\right| \tag{2.61}$$

Some comments about these resonance conditions are necessary. First, note that Eq. 2.59 should *not* be a precise equality for the real-world LIGO (though it can be modelled in our program as an equality), since that would lead to unwanted resonance of the non-negligible second-order modulation sidebands (separated from ν_{Carr} by $2\nu_{\mathrm{Mod}}$) in the Fabry-Perot arm cavities. Second, Eq. 2.60, which specifies sideband resonance in the PRC (where the carrier is also resonant), requires an *odd* multiple of π rather than an even one, because of the well-known fact [e.g., 5] that reflection from a resonant cavity (e.g., carrier reflection from the arms) picks up a phase shift of $\sim\pi$ with respect to reflection from a significantly off-resonant cavity (e.g., sideband reflection from the arms). Lastly, note that these formulas only represent approximate relationships that are analytically specifiable just for perfect interferometers; fine-tuning for interferometers with imperfections is required, which is where the active optimization routines of the simulation program come into play.

The heterodyne GW-signal will be obtained by interfering the output signal beam (generated by GW's from the carrier light in the arms) with the RF-sideband local oscillator beams[2], followed by demodulation of the resulting photodiode output signal at ν_{Mod}.

1. Since the GW forces are dimensionless strains that cause cavity length changes proportional to the unperturbed cavity lengths, only the *long-baseline* Fabry-Perot arm cavities (and the fields that circulate for long durations within them due to resonant buildup) are significantly affected by the incoming GW's.
2. All of these fields emerge — and are coherently summed — at the beamsplitter signal/exit port.

To optimize the GW-signal, the RF-sidebands must be *maximally channeled* through the signal port[1] while the carrier light is nominally held to a dark-fringe there. This is achieved by incorporating a macroscopic (~few tens of cm) length asymmetry:

$$L_{\text{Asymm}} \equiv \frac{L_2 - L_3}{2} \tag{2.62}$$

between the two arms of the PRC, so that the sideband beams returning from the inline and offline paths pick up significant (differential) phase with respect to one another, because of the sidebands' frequency offset from the carrier. Total destructive interference (as happens for the carrier light) is thus avoided, and a finite fraction of the resonating sideband power is extracted from the exit port during each pass of the beams through the interferometer. The resulting amount of local oscillator light can be maximized by an optimal choice of L_{asymm}, and the process of such optimization will be discussed in Section 2.5.5. This heterodyne GW-signal generation method is called the "Schnupp Asymmetry Scheme"[2] [47], and it will be used in the first generation of LIGO interferometers.

A large number of signals will have to be read out in a full-fledged LIGO interferometer in order to lock all of the lengths to their proper resonance conditions, as well as keeping all of the mirrors properly aligned, stabilizing the laser frequency, and so on. This sophisticated complex of control systems [48, 38] will involve the carrier and RF-sideband frequency beams (and beams at additional frequencies), sampled with pickoffs at various locations in the interferometer, and including the use of several modulation and demodulation operations, feedback loops, etc. In order to simplify our simulation task, however, we model only those elements which have a direct role in generating the gravitational-wave signals: the beams at the carrier and (upper and lower) RF sideband frequencies[3], oscillating within the core optical configuration of the interferometer (not including, for example, preliminary optics for mode-cleaning and frequency stabilization). Besides signal generation, these e-fields will also be the most (or only) significant sources of light power emerging from the beamsplitter exit port — RF-sideband light emerging intentionally for signal detection purposes, and carrier light emerging due to imperfect interferometer contrast — and thus will be the dominant sources of sensitivity-limiting shot noise. The simplified interferometer thus modelled by the simulation program, therefore, is sufficient for calculating signal-to-noise ratios and GW-sensitivity curves without any important loss of generality.

---

1. This is actually the *fundamental* requirement for the sideband fields, which the resonance conditions, Eq. 2.59 and 2.60, are merely meant to assist.
2. After the late Lise Schnupp.
3. In practice, we model only one of the two RF-sidebands (the lower one, $\nu_{SB-} \equiv \nu_{Carr} - \nu_{Mod}$, has been arbitrarily chosen), and use those numbers for the other sideband as well; this is justified since the two sidebands behave very similarly in the interferometer, with power differences that are either negligible or are only a few percent.

As discussed previously in Sec. 1.3.4, our Dual Recycling simulations are different from the Initial-LIGO case in that no RF-sideband beams are modelled. This removes the need for tuning RF-sidebands to resonance conditions, as well as making a Schnupp Length Asymmetry optimizer unnecessary (the Schnupp Asymmetry scheme might in fact not even be practical for Dual-Recycled LIGO interferometers). For convenience, therefore, we set L_{Asymm} to zero in the Dual Recycling runs.

Nevertheless, the adjustment of cavity lengths for achieving carrier resonance conditions is still necessary, and the Dual Recycling configuration has an additional length to adjust: L_6, the length of the Signal Recycling Cavity. The method for the adjustment of this parameter, and all other optimizable cavity lengths, is the subject of the following subsection.

2.5.2 Length-tuning for cavity resonance

The *macroscopic* lengths of the cavities in the LIGO interferometer (the 4 km long Fabry-Perot arms, etc.) are specified by the LIGO project for various scientific, technical, and economic reasons. But the *microscopic* lengths of these cavities must be adjusted (via sub-wavelength scale changes) to achieve the required interferometer resonance conditions. This "locking" procedure is done in our program for carrier beam simulations (and utilizing GW-induced sideband beams, for the Dual Recycling case), and the final, optimized lengths are exported to serve as input for the subsequent RF-sideband simulation runs[1].

The Initial- (Advanced-) LIGO configuration has four (five) independent lengths to be controlled. These microscopic lengths, obtained by taking the modulus of the relevant macroscopic lengths, are chosen to be:

$$L^{\text{micro}}_{\text{inline arm}} \equiv \text{Mod}\left\{L_4, \frac{\lambda_{\text{laser}}}{2}\right\} \tag{2.63}$$

$$L^{\text{micro}}_{\text{offline arm}} \equiv \text{Mod}\left\{L_5, \frac{\lambda_{\text{laser}}}{2}\right\} \tag{2.64}$$

$$L^{\text{micro}}_{\text{PRC}} \equiv \text{Mod}\left\{L_1 + \frac{(L_2 + L_3)}{2}, \frac{\lambda_{\text{laser}}}{2}\right\} \tag{2.65}$$

$$L^{\text{micro}}_{\text{SRC}} \equiv \text{Mod}\left\{L_6 + \frac{(L_2 + L_3)}{2}, \frac{\lambda_{\text{laser}}}{2}\right\} \tag{2.66}$$

1. The sideband resonance conditions are obtained via alternative procedures (see Sec. 2.5.3), not through length changes.

$$L_{\text{dark-port}}^{\text{micro}} \equiv \text{Mod}\left\{\frac{(L_2 - L_3)}{2}, \frac{\lambda_{\text{laser}}}{2}\right\} \tag{2.67}$$

Note that changes to $\{L_1 - (L_2 + L_3)/2\}$ (or $\{L_6 - (L_2 + L_3)/2\}$) which keep the lengths in Eq's. 2.63–2.67 constant have no effect on the interferometer resonance conditions. Also note that $\lambda_{\text{laser}}/2$ is used in these formulas instead of λ_{laser}, because the e-fields pick up twice as much phase during *round-trips* through these lengths.

The goals of length tuning are to set $L_{\text{inline arm}}^{\text{micro}}$, $L_{\text{offline arm}}^{\text{micro}}$, and $L_{\text{dark-port}}^{\text{micro}}$ to zero (the former two for Fabry-Perot arm resonances, the latter for a beamsplitter exit port dark-fringe), while setting $L_{\text{PRC}}^{\text{micro}}$ to $\lambda_{\text{laser}}/2$ in order to counter the π phase shift picked up in reflection from the arm cavities, so that a double resonance can be achieved for the carrier beam. For Dual Recycling runs, $L_{\text{SRC}}^{\text{micro}}$ must be adjusted to whatever value is necessary, given the GW-frequency that one wishes to optimize the (possibly narrowbanded) sensitivity peak for.

These purely analytical specifications must be translated into a prescription for resonance finding given nontrivial e-field wavefronts interacting with imperfect mirrors. Our method for accomplishing this is similar to that of Vinet, *et al.* [22], in that we adjust the various cavity lengths to optimize the computed "phases" (defined below) between certain specified cavity e-fields, which results in bringing the cavities to their particular resonant conditions. Alternatively, we have chosen not to use the method of McClelland, *et al.* [49], 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 an excessively time-consuming search over the multi-dimensional phase space of independent, controllable lengths.

The phase, Φ, between two e-fields is defined via an overlap integral (computed as a discrete sum over the pixelized grids), as follows:

$$\langle \hat{\mathrm{E}}_1 | \hat{\mathrm{E}}_2 \rangle \equiv \frac{(\text{Calc. Window size})^2}{N^2} \cdot \sum_{i=1}^{N} \sum_{j=1}^{N} \hat{\mathrm{E}}_1(i, j) \cdot \hat{\mathrm{E}}_2{}^*(i, j) \tag{2.68}$$

$$\text{with:} \qquad \Phi[\hat{\mathrm{E}}_1, \hat{\mathrm{E}}_2] \equiv \text{Tan}^{-1}\left[\frac{\text{Im}\langle \hat{\mathrm{E}}_1 | \hat{\mathrm{E}}_2 \rangle}{\text{Re}\langle \hat{\mathrm{E}}_1 | \hat{\mathrm{E}}_2 \rangle}\right]$$

This phase will be eventually driven to zero (or to its desired value) after several best-guess length changes, which are computed as:

$$\Delta L = -\Phi \cdot \left(\frac{-\lambda_{\text{laser}}}{4\pi}\right) = \frac{\Phi}{2k} \tag{2.69}$$

where a factor of $1/2$ has again been included because of the round-trips through the length-corrected beam paths.

Considering the steady-state equation, Eq. 2.53, it is clear that a cavity field will be unchanging and maximized (i.e., resonant) when $\vec{\tilde{E}}_{\text{steady-state}}$, R.T.$\{\vec{\tilde{E}}_{\text{steady-state}}\}$, and $\vec{\tilde{E}}_{\text{exc}}$ all have zero phase between them. There are several, related ways to bring this condition about. One can drive $\Phi[\vec{\tilde{E}}^{(N)}, \text{R.T.}\{\vec{\tilde{E}}^{(N)}\}] \rightarrow 0$, and allow the e-field relaxation algorithm to achieve a steady-state by zeroing $\Phi[\vec{\tilde{E}}_{\text{exc}}, \vec{\tilde{E}}^{(N)}]$. Alternatively, driving $\Phi[\vec{\tilde{E}}_{\text{exc}}, \text{R.T.}\{\vec{\tilde{E}}^{(N)}\}] \rightarrow 0$ will make the relaxer zero out $\Phi[\vec{\tilde{E}}^{(N)}, \text{R.T.}\{\vec{\tilde{E}}^{(N)}\}] \rightarrow 0$. The same result can be achieved all together by driving $\Phi[\vec{\tilde{E}}^{(N)}, (\text{R.T.}\{\vec{\tilde{E}}^{(N)}\} + \vec{\tilde{E}}_{\text{exc}})] \rightarrow 0$. Each of these methods are equivalent, and different ones are used in our program where convenient. In all cases, each of the e-fields used by the relaxation algorithm are multiplied by phases which reflect the length changes that have been performed (to smooth over the changes for the e-field relaxation algorithm), and the propagation operators are updated to represent the new cavity lengths.

To achieve the dark-fringe condition (by adjusting $L^{\text{micro}}_{\text{dark-port}}$), the program measures the phase at the beamsplitter exit port between the carrier e-fields coming from the inline and offline arms, and makes length adjustments by taking the length change prescribed by Eq. 2.69. Now, though, it sets Φ to π instead of to zero, and applies antisymmetric ("differential-mode") corrections of $\Delta L/2$ to L_2, and $-\Delta L/2$ to L_3.

The optimization of $L^{\text{micro}}_{\text{SRC}}$ in Dual Recycling runs requires an additional procedure. The program generates GW-induced sidebands in the Fabry-Perot arm cavities from the resonant carrier fields that are stored there, at the GW-frequency equal to the one which the interferometer is meant to be optimized for: $\nu^{\text{optim}}_{\text{GW}}$. The GW-induced sidebands are then (partially) relaxed in the interferometer, and made resonant in the SRC via adjustments of the position of the Signal Recycling Mirror; these adjustments are computed by nulling the phase for round-trips of the $\nu^{\text{optim}}_{\text{GW}}$ beam through the SRC. This method, involving the generation and relaxation of GW-induced sideband fields, is computationally expensive, but it is the only accurate method for tuning the SRC to the right GW-frequency for an interferometer with highly deformed mirrors.

The overall length control procedure that we use differs from the one in [22], in that 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, but rather the entire e-fields are used. This is done for a couple of reasons. First, for the dark-fringe condition, one actually wants to minimize the *total* dark-port power emerging from the beamsplitter exit port (all of which contributes to shot noise), not just the power in any particular mode. Second, for the carrier resonance conditions, coupling between modes brings power back into the TEM_{00} mode from higher modes[1], so that maximizing the total field also maximizes the particular (lowest) mode which we will want to consider in calculations of the GW-signal.

---

1. Equivalently, one can say the spatial beam mode oscillating in the interferometer will not be a pure TEM_{00} mode as estimated analytically from baseline optical parameters, but will be some perturbed "interferometer mode"; and that it is this perturbed mode which should be made resonant.

In either case, we have observed that both methods (i.e., with or without spatial mode selection) give virtually identical results for runs with realistically-deformed mirrors.

To satisfy ourselves that the length optimization scheme did indeed maximize the resonant powers (as well as keep the dark fringe dark), we performed runs in which we forced the lengths to be slightly off from the values that the optimizer had chosen, and observed that the optimizer had indeed found the best length settings of all those tested.

Lastly, we note that since most of the field propagations necessary for length adjustments must be performed anyway for the e-field relaxation process, these length optimizations generate little computational overhead (except for the $L_{\mathrm{SRC}}^{\mathrm{micro}}$ optimization), and also interfere minimally with the e-field relaxation process.

2.5.3 Sideband frequency fine-tuning

After the carrier run for an Initial-LIGO simulation is completed, a simulation run for one of the two RF-sidebands is performed. The sideband fields must be resonant in the PRC, as per Eq. 2.60, and nearly-anti-resonant in the Fabry-Perot arm cavities, as per Eq. 2.59. The sideband frequency can be roughly chosen by analytical calculation to approximately satisfy those relationships, as follows[1]:

$$\nu_{\mathrm{Mod}} = \left(k + \frac{1}{2}\right) \cdot \frac{c}{2L_{\mathrm{arm}}} \quad , \quad \mathrm{k} = \text{large integer} \tag{2.70}$$

$$\nu_{\mathrm{Mod}} = \left(n + \frac{1}{2}\right) \cdot \frac{c}{2L_{\mathrm{PRC}}} \quad , \quad n = 0, 1, 2\ldots \tag{2.71}$$

where c is the speed of light, $L_{\mathrm{PRC}} \equiv L_1 + (L_2 + L_3)/2$ is the approximate macroscopic length[2] of the PRC, averaged between the inline and offline paths, and where $L_{\mathrm{arm}} = L_4 = L_5 \approx 4\,\mathrm{km}$. Given the technical limitations of electro-optic modulation, a good value for n in Eq. 2.71 is 1, yielding a modulation frequency of $\nu_{\mathrm{Mod}} \cong 24\ \mathrm{MHz}$.

The approximate satisfaction of Eq's. 2.70 and 2.71 is good enough for the arm cavity condition (the variation in power buildup around an anti-resonance is very flat), but a more precise, dynamic process of parameter fine-tuning during the run is necessary to achieve a sharp resonance in the PRC, since runs involving optical deformations cause degraded and imperfectly-balanced sideband power buildup in the interferometer.

---

1. To satisfy both relationships simultaneously, the macroscopic interferometer cavity lengths cannot be arbitrarily specified; changes of ~few cm for the PRC, and ~few m for the Fabry-Perot arms (thus L_{Arm} is not *exactly* 4 km) must generally be applied to make it work. The carrier run must be done taking these restrictions into account.
2. Note that in calculating L_{PRC} we implicitly include the optical path length through the substrates of the finite-thickness mirrors; the relevant substrates here are those of the input mirrors to the Fabry-Perot arm cavities, and the beamsplitter.

Since the cavity lengths are completely fixed by the carrier run, the sideband resonance condition can only be achieved by small adjustments to the sideband modulation frequency, ν_{Mod}. In the same sense as in the previous section on length adjustments, a phase can be computed between $\vec{\tilde{E}}^{(N)}$ and R.T.$\{\vec{\tilde{E}}^{(N)}\}$ in the PRC, and resonance can be achieved by performing a series of frequency corrections which converges after many iterations. These corrections are given by the formula[1]:

$$\Delta\nu_{Mod} = \Phi \cdot \frac{c}{4\pi L_{PRC}} \tag{2.72}$$

This prescription for $\Delta\nu_{Mod}$ need not, in fact, be exactly followed; in practice, the program makes individual changes that are smaller than this by about an order of magnitude (and the size of $\Delta\nu_{Mod}$ is also given a hard maximum limit), in order to avoid overshoot and achieve a more stable optimization process. In total, the cumulative frequency change for runs with deformed mirrors and interferometer parameters such as those in Table 2.1, is small, typically a few hundred Hz or less.

As an example of how optical imperfections can cause the need for sideband frequency adjustments, consider the effects of a beamsplitter for which reflection and transmission is not exactly 50%-50%. In this case, the inline and offline paths in the PRC are not sampled equally by RF-sideband power, and hence the simple mean of L_2 and L_3 (which differ because of the Schnupp Length Asymmetry) is not the proper weighting[2] for computing an average L_{PRC}. This alters the ν_{Mod} computed via Eq. 2.71, by an amount that can be tens of kHz for a 51%-49% beamsplitter. That kind of frequency change is far too large to be performed by ν_{Mod} optimization during the simulation run, because it is comparable to the free spectral range of the Fabry-Perot arms, and the arm cavity near-anti-resonance would be spoiled by the optimization process. That particular, large adjustment is one that we must account for analytically, prior to the run; but smaller effects, such as those caused by realistic deformations in mirror surfaces and substrates, are successfully dealt with by the sideband frequency fine-tuning algorithm in the program.

In addition to the resonance/anti-resonance conditions discussed above, the RF-sidebands must also be maximally channeled out through the beamsplitter exit port. This is done via Schnupp Length Asymmetry optimization, performed simultaneously with sideband frequency optimization, and which will be discussed later in Sec. 2.5.5.

---

1. The resonance condition for the *upper* RF-sideband is generally the mirror-image of that for the *lower* sideband, so that the overall correction to ν_{mod} would be virtually the same (though sign-reversed) regardless of which of the two sidebands were simulated.
2. The correct weighting of L_2 and L_3 goes like the *square* of the amplitude r and t values for the beamsplitter, because the beamsplitter is encountered twice during a round-trip through the PRC.

2.5.4 Power-recycling mirror reflectivity optimization

The amount of power buildup in the interferometer, and the frequency dependence of its response, depend upon the reflectivities of the mirrors and are constrained by the mirror losses. In LIGO, the losses will be as low as can be feasibly achieved, and the mirror reflectivities will be determined by a variety of physical requirements. The Fabry-Perot arm cavity back mirrors will be as reflective as possible, since any light that transmits through them is wasted[1]. The arm cavity input mirrors have their reflectivities specified as $R \approx 0.97$, in order to set the "pole" or "roll-off" frequency of the arm cavity response to ~90 Hz, which in turn will put the frequency for highest LIGO GW-sensitivity (including total noise contributed from all sources) at a few tens of Hz above that pole frequency. The beamsplitter, of course, will be as close to 50%-50% as possible; and the reflectivity of the Dual Recycling (if present) is not fixed to any value, but can be chosen differently for different GW-response functions.

The reflectivity of the Power Recycling Mirror is unique, however, both because it has a definite optimal value, and because that value cannot be accurately predetermined via analytical considerations of an ideal interferometer. The criterion for choosing R_1 is that of maximizing the carrier power amplification that is achieved by Power Recycling; but this amplification depends critically upon the losses experienced in the interferometer, which may be radically affected by the presence of imperfect optics. This reflectivity must therefore be fixed by a dynamic optimization algorithm that adjusts R_1 during the course of a carrier[2] simulation run. And even though a real LIGO interferometer would not possess the capability of changing the Power Recycling Mirror reflectivity in a dynamic or continuous way, the R_1 optimization routine in the simulation program is a resource for helping *prescribe* the value of R_1 for the recycling mirrors that LIGO will procure.

First of all, consider the simplified example of a two-mirror cavity with input mirror parameters $t_1 = \sqrt{T_1}$, $r_1 = \sqrt{R_1}$, and a back mirror amplitude reflectivity of $r_2 = \sqrt{R_2}$. The input mirror here represents the Power Recycling Mirror of LIGO, while r_2 represents the overall reflectivity of the rest of the interferometer. The mirror losses in the simple cavity are given by $A_1 = 1 - (R_1 + T_1)$ and $A_2 = 1 - R_2$, where any back mirror transmission has been lumped into A_2.

If this cavity is held perfectly resonant for an excitation laser field, $\tilde{\mathrm{E}}_{\mathrm{las}}$, then the resonant, steady-state cavity field will be given by:

$$\tilde{\mathrm{E}}_1 = t_1 \tilde{\mathrm{E}}_{\mathrm{las}} \cdot \{1 + r_1 r_2 + (r_1 r_2)^2 + \ldots\} = \tilde{\mathrm{E}}_{\mathrm{las}} \cdot \left(\frac{t_1}{1 - r_1 r_2}\right) \tag{2.73}$$

1. Except for a *small* transmission, ~10 parts per million, to allow some light to escape for control system purposes.
2. Further justification for optimizing the Power Recycling Mirror reflectivity during the carrier run is given in the discussion at the end of Section 3.3.

The total e-field reflected from the cavity is formed by summing the "resonant" reflection leaking from the cavity with the "prompt" laser reflection, as follows:

$$\begin{aligned}\vec{\mathrm{E}}_{\mathrm{ref,\,tot}} &= \vec{\mathrm{E}}_{\mathrm{ref,\,res}} + \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} = \{\vec{\mathrm{E}}_1 \cdot r_2 t_1\} + \{\vec{\mathrm{E}}_{\mathrm{las}} \cdot -r_1\} \\ &= \vec{\mathrm{E}}_{\mathrm{las}} \cdot \left(\frac{(t_1^2 + r_1^2) \cdot r_2 - r_1}{1 - r_1 r_2}\right) = \vec{\mathrm{E}}_{\mathrm{las}} \cdot \left(\frac{(1 - A_1) \cdot r_2 - r_1}{1 - r_1 r_2}\right)\end{aligned} \tag{2.74}$$

All of the power from the excitation laser must either be lost due to dissipation by the cavity mirrors (via absorption and/or scattering losses), or reflected back from the cavity. Since a higher dissipation inside the cavity means a larger amount of stored power, the way to maximize $\vec{\mathrm{E}}_1$ is by zeroing $\vec{\mathrm{E}}_{\mathrm{ref,\,tot}}$. This is accomplished by the choice:

$$\begin{aligned}R_{1,\,\mathrm{optim.}} &= r_{1,\,\mathrm{optim.}}^2 = R_2(1 - A_1)^2 = (1 - A_2)(1 - A_1)^2 \approx 1 - 2A_1 - A_2 \\ \Rightarrow T_{1,\,\mathrm{optim.}} &= 1 - A_1 - R_{1,\,\mathrm{optim.}} = A_1 + A_2\end{aligned} \tag{2.75}$$

where higher powers and/or cross multiplications of the (very small) loss terms have been dropped. To summarize the result of Eq. 2.75, one says that the optimal choice of transmission for the Power Recycling Mirror is equal to the total losses experienced by the resonant beam in the interferometer — and that this optimum choice of $T_{1,\,\mathrm{optim.}}$ (and thus $R_{1,\,\mathrm{optim.}}$) causes the total reflected e-field to be zeroed out.

This prescription must be generalized to the case of an interferometer with imperfect optics. To do so, we consider the overlap integral between $\vec{\mathrm{E}}_{\mathrm{ref,\,prompt}}$ and $\vec{\mathrm{E}}_{\mathrm{ref,\,tot}}$:

$$\begin{aligned}&\langle \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} | \vec{\mathrm{E}}_{\mathrm{ref,\,tot}} \rangle = \\ &(\text{real factor}) \times \langle \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} | \vec{\mathrm{E}}_{\mathrm{ref,\,res}} \rangle + (\text{real factor}) \times \left| \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} \right|^2\end{aligned} \tag{2.76}$$

For a perfectly resonant Power Recycling Cavity (a state that will be enforced as well as possible by the length adjustment procedures of Sec. 2.5.2), $\vec{\mathrm{E}}_{\mathrm{ref,\,prompt}}$ and $\vec{\mathrm{E}}_{\mathrm{ref,\,res}}$ will be exactly out of phase (a condition apparent from Eq's. 2.73 and 2.74), so that $\mathrm{Im}\langle \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} | \vec{\mathrm{E}}_{\mathrm{ref,\,tot}} \rangle = 0$ is already satisfied. The minimization of $\vec{\mathrm{E}}_{\mathrm{ref,\,tot}}$, and hence the optimization of R_1, thus becomes the task of obtaining:

$$\mathrm{Re}\langle \vec{\mathrm{E}}_{\mathrm{ref,\,prompt}} | \vec{\mathrm{E}}_{\mathrm{ref,\,tot}} \rangle \rightarrow 0 \tag{2.77}$$

The actual zeroing of $\vec{\mathrm{E}}_{\mathrm{ref,\,tot}}$ is impossible for a realistically imperfect interferometer, since $\vec{\mathrm{E}}_{\mathrm{ref,\,prompt}}$ and $\vec{\mathrm{E}}_{\mathrm{ref,\,res}}$ will have different contributions from high spatial modes, so that complete cancellation is impossible. But Eq. 2.77 is a condition that is achievable for any modal structure that the e-fields may possess, and it brings $\vec{\mathrm{E}}_{\mathrm{ref,\,tot}}$ as close to zero as

is possible for the imperfect interferometer. Eq. 2.77 also provides an error signal which allows the program to compute the (approximate) magnitude and *sign* of the change to be made to R_1 during an optimization step.

For the reflectivity optimization algorithm in the simulation program, the magnitude of ΔR_1 for a given $\mathrm{Re}\langle \tilde{\vec{E}}_{\mathrm{ref,\,prompt}} | \tilde{\vec{E}}_{\mathrm{ref,\,tot}} \rangle$ has been determined by trial-and-error (though with some guidance from Eq. 2.74, for typical LIGO parameters), using the results from a variety of runs which run the gamut from perfect optics to highly deformed optics. With the additional restriction of upper and lower bounds on R_1 to maintain energy conservation[1] (as discussed in Sec. 2.3.3.2), this optimization algorithm is a highly effective procedure for maximizing the resonating power (and thus the broadband GW-signal gain) of the Power Recycling Cavity.

In the real LIGO system, the appropriate value for R_1 will actually be slightly *lower* than $R_{1,\mathrm{optim.}}$, both to hedge (on the safer side[2]) against uncertainties (and gradual increases with time, due to contamination) of the true interferometer losses, and also to provide some finite amount of reflected light for the cavity length control signals. In the runs that we present in this thesis, R_1 will have been driven exactly to $R_{1,\mathrm{optim}}$, in order for us to quantify the theoretically "optimal" performance of an interferometer with deformed mirrors; but when it has been needed for LIGO research and development, the simulation program has been run with a value of R_1 that is fixed slightly below the computed $R_{1,\mathrm{optim.}}$, once $R_{1,\mathrm{optim.}}$ has been determined via a preceding run.

Lastly, we note that optimal value of R_1 for the carrier e-fields is *not necessarily* the optimal value of R_1 for the RF-sidebands, because the losses experienced by the carrier and sidebands in the interferometer are different. Specifically, since the carrier is resonantly amplified in the Fabry-Perot arm cavities, its losses due to dissipation in the arm cavities are much larger than the losses of the sidebands are there. On the other hand, the sideband power channeled out of the interferometer exit port (to serve as a local oscillator for heterodyne GW-signal detection) can be thought of as a "loss" in the interferometer, and one which will vary as the Schnupp Length Asymmetry is varied. The sequence of optimizations in the program therefore computes $R_{1,\mathrm{optim.}}$ from the carrier run alone, and then adjusts L_{Asymm} in the subsequent sideband run to adapt to this reflectivity value. The optimization process for L_{Asymm} is the subject of the upcoming subsection.

---

1. Some additional procedures are also performed to ensure energy conservation in mirror maps with variations in the *amplitudes* of R and T across their profiles (as opposed to deformations that are pure phase fluctuations, as with our $\lambda/1800$, etc., surface maps).
2. The power buildup in a cavity drops much faster for an R_1 that is too large, than for one that is too small; c.f. Eq. 2.73. In addition, any transitions to the former ("undercoupled") case from the latter ("overcoupled") case, due to the use of a mirror reflectivity exactly at (or too close to) $R_{1,\mathrm{optim.}}$, would cause a sudden phase shift of π in the total reflected e-field (c.f. Eq. 2.74), which would interfere with the acquisition of stable lock for beam resonances in the PRC.

2.5.5 Optimization of the "Schnupp" Length Asymmetry

As described above, the Schnupp Length Asymmetry Scheme in the Initial-LIGO configuration is designed to maximize the amount of RF-sideband power emerging from the beamsplitter exit (signal) port. These exiting sideband fields, by being added to the carrier fields which exit through the signal port due to GW-perturbations of the arms, serve as a local oscillator for heterodyne detection of the gravitational waves.

Contrary to naive expectations, the way to maximize this emerging local oscillator power is *not* by making the sideband beams coming from the inline and offline interferometer arms experience total constructive interference at the exit port; this would cause all of the sideband power to exit the interferometer after only a single pass through the system. As will be seen, the optimum choice for L_{Asymm} is the one which extracts a carefully specified fraction of the sideband power during each round-trip pass, such that a balance is struck between allowing the sidebands to take advantage of the gain due to Power Recycling, while not keeping them in the system so long that too much of the sideband power is dissipated as losses in the interferometer.

For the following analysis, let $t_1 = \sqrt{T_1}$, $r_1 = \sqrt{R_1}$ be the parameters of the Power Recycling Mirror, and let the reflection and transmission amplitudes of the (possibly non-50%-50%) beamsplitter from the various directions be, as per Fig. 2.9: $r_{\mathrm{BS,r}}$, $r_{\mathrm{BS,a}}$, $t_{\mathrm{BS,in}}$, and $t_{\mathrm{BS,off}}$. Furthermore, let the *total* reflection amplitudes from the inline and offline paths (including the *resonant reflections* from the Fabry-Perot arm cavities) be represented algebraically by the complex amplitudes: $r_2 \cdot \mathrm{Exp}[\Phi_{\mathrm{Symm}} + \Phi_{\mathrm{Asymm}}]$ and $r_3 \cdot \mathrm{Exp}[\Phi_{\mathrm{Symm}} - \Phi_{\mathrm{Asymm}}]$. Here, Φ_{Symm} represents all phase shifts that are common to both paths (including, for this calculation, the phase of propagation from the Power Recycling Mirror to the beamsplitter), while Φ_{Asymm} represents (half of) the net phase difference between the paths. For perfectly balanced arm cavities, in which all of the differential phase shift comes from the Schnupp length asymmetry, we would have (for round-trips along $L_{\mathrm{Symm}} \pm L_{\mathrm{Asymm}}$):

$$\Phi_{\mathrm{Asymm}} = -2\, k_{\mathrm{Mod}}\, L_{\mathrm{Asymm}} \tag{2.78}$$

With these definitions, the resonant e-field built up in the PRC can be calculated as:

$$\begin{aligned} \tilde{\mathrm{E}}_1 &= t_1 \tilde{\mathrm{E}}_{\mathrm{las}} \cdot \Big\{ 1 + [r_1 \cdot (r_2\, t_{\mathrm{BS,in}}^2\, e^{i(\Phi_{\mathrm{Symm}} + \Phi_{\mathrm{Asymm}})} + r_3\, r_{\mathrm{BS,r}}^2\, e^{i(\Phi_{\mathrm{Symm}} - \Phi_{\mathrm{Asymm}})})] \\ &\qquad + [\text{""}]^2 + \ldots \Big\} \\ &= t_1 \tilde{\mathrm{E}}_{\mathrm{las}} \cdot \Big\{ 1 - [r_1 \cdot (r_2\, t_{\mathrm{BS,in}}^2\, e^{i(\Phi_{\mathrm{Symm}} + \Phi_{\mathrm{Asymm}})} + r_3\, r_{\mathrm{BS,r}}^2\, e^{i(\Phi_{\mathrm{Symm}} - \Phi_{\mathrm{Asymm}})})] \Big\}^{-1} \end{aligned} \tag{2.79}$$

This result can then be used to compute the amplitude of the sideband field exiting from the signal port, as follows:

$$\vec{E}_{\text{Exit-Port}} = \vec{E}_1 \cdot \left[-r_2\, t_{\text{BS,in}}\, r_{\text{BS, a}}\, e^{i\left(\frac{\Phi_{\text{Symm}}}{2} + \Phi_{\text{Asymm}}\right)} + r_3\, r_{\text{BS, r}}\, t_{\text{BS,off}}\, e^{i\left(\frac{\Phi_{\text{Symm}}}{2} - \Phi_{\text{Asymm}}\right)} \right] \quad (2.80)$$

$$= \frac{t_1 \vec{E}_{\text{las}} \cdot e^{i\left(\frac{\Phi_{\text{Symm}}}{2}\right)} \cdot \left\{ r_3\, r_{\text{BS, r}}\, t_{\text{BS,off}}\, e^{-i(\Phi_{\text{Asymm}})} - r_2\, t_{\text{BS,in}}\, r_{\text{BS, a}}\, e^{i(\Phi_{\text{Asymm}})} \right\}}{\left\{ 1 - [r_1 \cdot (r_2\, t_{\text{BS,in}}^2\, e^{i(\Phi_{\text{Symm}} + \Phi_{\text{Asymm}})} + r_3\, r_{\text{BS, r}}^2\, e^{i(\Phi_{\text{Symm}} - \Phi_{\text{Asymm}})})] \right\}}$$

While this formula for the e-field to be maximized is quite complicated, the interpretation is simple, and proves our earlier claim: the numerator term in braces represents the amount of local oscillator light extracted per interferometer round-trip, the denominator term in braces represents the gain due to Power Recycling (given the "loss" due to the extraction), and it is the ratio of these two effects which must be optimized with respect to Φ_{Asymm}, in order to maximize the exit port sideband power.

It is possible to find a simple, closed-form solution to the maximization problem if we make the approximations $r_2 \approx r_3 \equiv r_{\text{Arm}}$, $r_{\text{BS, r}} \approx r_{\text{BS, a}} \equiv r_{\text{BS}}$, and $t_{\text{BS,in}} \approx t_{\text{BS,off}} \equiv t_{\text{BS}}$. The resonance condition for the PRC then becomes $\text{Exp}[\Phi_{\text{Symm}}] = 0$, and we get:

$$\vec{E}_{\text{Exit-Port}} \approx \frac{[\text{factors independent of } \Phi_{\text{Asymm}}] \times \sin(\Phi_{\text{Asymm}})}{\left\{ 1 - r_1 \cdot r_{\text{Arm}} \cdot [t_{\text{BS}}^2\, e^{i(\Phi_{\text{Asymm}})} + r_{\text{BS}}^2\, e^{-i(\Phi_{\text{Asymm}})}] \right\}} \quad (2.81)$$

It is not difficult to show that $P_{\text{Exit-Port}} = \left|\vec{E}_{\text{Exit-Port}}\right|^2$ is now maximized by the choice:

$$\cos(\Phi_{\text{Asymm}}) \equiv \cos[-2 k_{\text{Mod}} L_{\text{Asymm}}] = 2\, r_1\, r_{\text{Arm}} (t_{\text{BS}}^2 + r_{\text{BS}}^2) \quad (2.82)$$

Assuming for now that the approximations used to derive Eq. 2.82 are valid, we must prescribe a method by which this relationship can be achieved for the e-fields in the simulation program.

First of all, the *r* and *t* values that are to be plugged into Eq. 2.82 (or into an equivalent formula derived without approximations) must correctly represent the action of the realistically imperfect mirrors upon the interferometer fields. To that end, the program calculates "effective" reflectivities and transmissions by comparing the power in an e-field *before* vs. *after* its interaction with a mirror or arm cavity. By using e-fields that are in the

process of being relaxed to steady-state for constructing these effective r and t values, one takes into account how the mirrors truly shape interferometer power buildup. In addition, the program only considers beam power that is in the TEM_{00} spatial mode for these calculations, since the generated local oscillator light must have the same modal structure as the GW-induced signal light in order to contribute to the GW-sensitivity. This GW-induced signal light is coming from the Fabry-Perot arm cavities (which resonantly amplify the TEM_{00} mode, and suppress the higher modes), and it will therefore be predominantly in that lowest order mode.

The program measures Φ_{Asymm} by taking the overlap integral (c.f. Eq. 2.68) between the RF-sideband e-fields approaching the beamsplitter exit port from the inline and offline paths of the interferometer. To bring Φ_{Asymm} to a value which satisfies Eq. 2.82, *macroscopic* adjustments (up to ~1 mm per step) are made to L_{Asymm} during the sideband simulation run as the interferometer fields are being relaxed. These adjustments are differential length changes: $\Delta L_{Asymm}/2$ is added to L_2 and subtracted from L_3. Since the ambiguity $\cos(\Phi_{Asymm}) = \cos(-\Phi_{Asymm})$ exists, the program makes sure that the *sign* of ΔL_{Asymm} is chosen in order to pick the smaller total adjustment that would be needed to satisfy Eq. 2.82. Furthermore, in order to minimize any change that the carrier beam would experience because of the altered Schnupp asymmetry length (especially differences in the carrier dark-fringe at the beamsplitter exit port), the change to each arm path, $\Delta L_{Asymm}/2$, is restricted to be an integral multiple of $\lambda_{laser}/4$. In any case, though the change in carrier behavior due to Schnupp Length Asymmetry optimization is generally negligible, the program is set up so that both the carrier and sideband runs may be re-performed with the optimized (and now fixed) L_{Asymm}.

Regarding the optimization formula, Eq. 2.82, it is apparent that the best value for L_{Asymm} has a direct dependence upon the precise sideband frequency, ν_{mod} (via k_{mod}). This fact has two important implications: first, the adjustments to L_{Asymm} cannot begin until enough iterations have passed so that the sideband frequency optimization process is *mostly* complete, and ν_{mod} is near its final value; and second, the two optimization procedures must not be active at the same time, since they interfere with one another. In practice, adjustments to ν_{mod} occur in two stages: before the L_{Asymm} optimization has begun (to start out by bringing the sideband frequency to the approximately correct value), and again after the L_{Asymm} optimization has finished (to refine ν_{mod} for the conclusion of the run). Even during the active part of the L_{Asymm} optimization process, changes to the asymmetry length are few and far between (only once every ~200 iterations), and are in fact suppressed when the desired ΔL_{Asymm} reaches below a certain (very generous) threshold, because changes to L_{Asymm} cause a large disruption for the e-field relaxation process. If multiple changes are made to L_{Asymm} without allowing the e-fields to re-relax sufficiently between adjustments, then the Schnupp asymmetry length does not settle down to a convergent value.

We wrap up this subsection by deriving the L_{Asymm} optimization requirement for the case where the approximations leading to Eq. 2.82 do not hold, i.e., an interferometer with inline and offline beam paths that are unbalanced, possibly due to optical imperfections. In this case, the power in the full $\tilde{\vec{E}}_{\text{Exit-Port}}$, as expressed in Eq. 2.80, must be maximized.

First, we consider the requirement for sideband resonance in the PRC, which is now more complicated then before. PRC resonance requires that the term that is subtracted from unity in the denominator of Eq. 2.80 must be maximized, which is done by making it purely *real* and *positive*. Zeroing out the imaginary component yields the result:

$$\tan(\Phi_{\text{Symm}}) = \tan(\Phi_{\text{Asymm}}) \cdot \left[\frac{r_3 \, r_{\text{BS, r}}^2 - r_2 \, t_{\text{BS,in}}^2}{r_3 \, r_{\text{BS, r}}^2 + r_2 \, t_{\text{BS,in}}^2} \right] \tag{2.83}$$

It is clear in Eq. 2.83 that optimization of the RF-sideband frequency for PRC resonance is directly affected by imbalance between the interferometer arms, and that this effect is weighted by $r_{\text{BS, r}}^2$ and $t_{\text{BS,in}}^2$, as has been asserted previously in Sec. 2.5.3.

By inserting Eq. 2.83 into Eq. 2.80, and using the fact that the denominator is now purely real, we derive (after some algebra) the quantity to be maximized by L_{Asymm} optimization, as follows:

$$P_{\text{Exit-Port}}\Big|_{\text{Optimum } \Phi_{\text{Symm}}} = \left|\tilde{\vec{E}}_{\text{Exit-Port}}\right|^2\Big|_{\text{Optimum } \Phi_{\text{Symm}}} \tag{2.84}$$

$$\propto \frac{(r_2 \, t_{\text{BS,in}} \, r_{\text{BS, a}})^2 + (r_3 \, t_{\text{BS,off}} \, r_{\text{BS, r}})^2 + 2 \, r_2 \, r_3 \, t_{\text{BS,in}} \, r_{\text{BS, a}} \, t_{\text{BS,off}} \, r_{\text{BS, r}} \cos(\Phi_{\text{Asymm}})}{\left[1 - r_1 \sqrt{(r_{\text{BS, r}}^4 \, r_3^2 + t_{\text{BS,in}}^4 \, r_2^2) + 2 \, r_2 \, r_3 \, r_{\text{BS, r}}^2 \, t_{\text{BS,in}}^2 \cos(2\Phi_{\text{Asymm}})}\right]^2}$$

A closed-form solution for the maximum of Eq. 2.84 with respect to Φ_{Asymm} is not only somewhat difficult to obtain, but would also be complicated enough (and would rely upon very delicate cancellations for setting the derivative equal to zero) so that it would be difficult to trust its validity, given the fact that the simulation program must use "effective" r and t values rather than precise analytical quantities for this optimization process. As an alternative, therefore, we have put Eq. 2.84 itself directly into the program, and each adjustment to L_{Asymm} is computed by using a numerical search routine (a "golden section" search from Numerical Recipes in Fortran [31]) to determine the value of Φ_{Asymm} that maximizes Eq. 2.84, given the values of r and t computed from the interferometer e-fields.

The result of the procedures discussed in this subsection, is an effective and reliable algorithm for maximizing the RF-sideband local oscillator power in an interferometer with significant optical imperfections, and also one that operates well in conjunction with all of the other optimization procedures that are implemented by the simulation program.

2.5.6 Sideband modulation depth optimization

All of the optimization procedures discussed so far are limited in the sense that they deal with *either* the carrier or the sideband — never with both. Thus they do nothing to strike a balance between the needs of the carrier and those of the RF-sideband(s), in an interferometer with deformations. Furthermore, none of these optimizers are *directly* linked to maximizing the shot-noise-limited sensitivity of the system, $\tilde{h}_{SN}(f)$. These considerations will therefore be dealt with by one more optimization procedure that is used in our simulation efforts: optimization of the RF-sideband *modulation depth.*

When a run is performed by the simulation program, the power of the (carrier or sideband) beam that excites the interferometer at the Power Recycling Mirror is specified as 1 Watt. In the real LIGO, the useful excitation laser power will be larger than this (perhaps 6 Watts for the first-generation interferometers), and will represent a fixed amount of total power that must be divided between the carrier and sideband beams by electro-optic phase modulation at radio frequencies (~24 MHz). The division of laser power between the carrier and its sidebands is determined by the modulation depth, Γ, and the proper selection of this parameter (determined by optimizing $\tilde{h}_{SN}(f)$) can only be determined after the simulation runs are complete; that is, once the behavior of the carrier and sideband fields in the realistically imperfect interferometer are computed. The optimization of Γ is therefore done as a *post-processing* step, which is handled during the analysis that follows the finished runs. At that point, it is a simple matter to multiply all of the simulated e-field powers by the true normalization of the carrier and sideband excitation beams. As inferred above, the other optimization routines that are active within the program may harm the sidebands (i.e., reduce their useful power) in order to help the carrier, or vice-versa; but the optimization of the modulation depth addresses this problem by striking the right balance between the carrier and sideband requirements.

The process of electro-optic modulation is mathematically represented as follows:

$$\mathrm{Mod}\{\vec{\tilde{E}}_{las}\, e^{i\omega t}\} \equiv \vec{\tilde{E}}_{las} e^{i\omega t + i\Gamma \sin \Sigma t} \tag{2.85}$$

where $\omega \equiv 2\pi\nu_{Laser}$ is the carrier angular frequency, $\Sigma \equiv 2\pi\nu_{Mod}$ is the modulation angular frequency for the sidebands, and Γ, of course, is the modulation depth. This expression for the phase-modulated beam can be expanded in terms of an infinite series of Bessel functions of the first kind, $J_n(x)$, using generating functions [50], via the result:

$$e^{i\Gamma\sin\Sigma t} = \cos(\Gamma\sin\Sigma t) + i\sin(\Gamma\sin\Sigma t) \tag{2.86}$$

$$= \left\{ J_0(\Gamma) + 2\sum_{n=1}^{\infty} J_{2n}(\Gamma)\cos(2n\Sigma t) \right\} + i\left\{ 2\sum_{n=0}^{\infty} J_{2n+1}(\Gamma)\sin[(2n+1)\Sigma t] \right\}$$

$$= \left\{ J_0(\Gamma) + \sum_{n=1}^{\infty} J_{2n}(\Gamma)\,[e^{i2n\Sigma t} + e^{-i2n\Sigma t}] \right\} + \left\{ \sum_{n=0}^{\infty} J_{2n+1}(\Gamma)\,[e^{i(2n+1)\Sigma t} - e^{-i(2n+1)\Sigma t}] \right\}$$

$$= J_0(\Gamma) + \sum_{n=1}^{\infty} J_n(\Gamma)\,[e^{in\Sigma t} + (-1)^n e^{-in\Sigma t}]$$

$$= \sum_{n=-\infty}^{\infty} J_n(\Gamma)\, e^{in\Sigma t}$$

where the last step was accomplished by using the property $J_{-n}(\Gamma) = (-1)^n J_n(\Gamma)$, for integral n.

Because Γ will be small, the higher-order terms in this series will be unimportant — assuming that they are non-resonant in the Fabry-Perot arm cavity, c.f. Eq. 2.59 and ensuing discussion — and may be dropped, thus yielding the following expression for the modulated laser field:

$$\mathrm{Mod}\{\vec{\tilde{E}}_{las}\, e^{i\omega t}\} = \vec{\tilde{E}}_{las}\, e^{i\omega t} \cdot \sum_{n=-\infty}^{\infty} J_n(\Gamma)\, e^{in\Sigma t} \tag{2.87}$$

$$\approx \vec{\tilde{E}}_{las} \cdot \{ J_0(\Gamma)\, e^{i\omega t} + J_1(\Gamma)\, e^{i(\omega+\Sigma)t} + J_{-1}(\Gamma)\, e^{i(\omega-\Sigma)t} + \ldots \}$$

The first term on the bottom line of Eq. 2.87 represents the carrier beam that remains after modulation, and the other two terms represent the first-order, upper and lower ("+/-") sidebands. These are the three laser frequencies which are modelled by the program for Initial-LIGO simulations. The remaining terms in the Bessel function series, which could lead to unpredictable control system behavior in LIGO if they were not very small, are not simulated; but Bessel functions with $|n| > 1$ are quite small for Γ much less than 1, which is a condition that our optimized choices of modulation depth always satisfy.

The carrier and its RF-sidebands both contribute to the GW-signal, and because of their ejection at the beamsplitter exit port — unavoidable ejection for the carrier due to imperfect dark-fringe contrast, but deliberate ejection for the sideband in order to produce local oscillator light — they both contribute to the shot noise that competes with this signal, as well. The optimal modulation depth is therefore determined via maximization of the ratio of GW-signal to shot noise. This ratio, which will be given explicitly in Eq. 3.3 of Section 3.1, and which is computed in Appendix D, is optimized with respect to Γ by running a small program in Mathematica [51] which we have written for this purpose.

In addition to the straightforward optimization of the shot-noise-limited sensitivity, we also keep an eye on the *total* amount of power (carrier and sideband) which falls upon the signal detection photodiode at the beamsplitter exit port. This quantity is also dependent upon Γ, and it is important to make sure that it is not excessive.

2.6 Verifications of the simulation program

In order for the LIGO simulation program to be a trusted resource for the research and development of gravitational-wave interferometers, it must be demonstrated to be a highly accurate and sufficiently complete model of reality. We have therefore performed rigorous tests of program algorithms and results to confirm that the outputs of our simulation runs are self-consistent, intuitively sensible, and in accordance with theoretical predictions.

The most conclusive test possible would naturally be the direct comparison of experimental and simulated results for a cavity system possessing optics that have been precisely measured, with those measurements being incorporated into the program as maps for the modelled mirrors. Unfortunately, due to the difference in *scale* between a table-top experiment that could be done, and a full-LIGO interferometer, it is not possible to perform such a comparison in a way that would produce meaningful results. This is mainly due to the size of a focused LIGO beam, which is large in order to prevent significant expansion of the beam spot size during propagation through the four-kilometer long Fabry-Perot arm cavities[1].

This large beam size has many important ramifications that would complicate such an experimental comparison. First of all, unlike table-top experiments where the beams are typically much smaller than the sizes of the optics, the LIGO beams effectively fill the mirrors, so that ~1-2 parts per million ("ppm") of power are lost from the edges of the mirror for each reflection or transmission. This diffractive loss (which would be insignificant for a table-top setup) is unavoidable, since it is impossible to maintain LIGO-required mirror quality over the entire surface of an exceptionally large optic. In addition, the larger beam size means that the typical spatial wavelengths of deformations that it samples are larger [52]; and since the power spectrum of deformations rises for larger wavelengths/ smaller wavenumbers (as shown in Appendix B), this means that LIGO is sampling larger-amplitude mirror deformations than a table-top experiment would be doing. Furthermore, since the mirror deformations relevant for the smaller (table-top) beam size would be much lower in amplitude, they would be extremely difficult to measure accurately enough in order to be reliably incorporated into the simulation program as mirror maps. For all of these reasons, it is not possible to perform realistic simulations of LIGO beam behavior without modelling a LIGO-sized beam. Even the possibility of constructing a table-top

1. The existence of the long-baseline arms themselves makes comparisons with a table-top experiment difficult, both because of high scattering losses due to the large propagation lengths and small mirror sizes, and — assuming a simple table-top experiment without arm cavities — because of the effects of these cavities upon the resonance properties of the coupled-cavity system (see, e.g., Sec. 4.3.2).

experiment with a strong focusing element in order to produce a large beam size would be an inadequate solution, since then it is the focusing element which would dominate the behavior of the system, and our grid-based optical simulation program is ill-equipped (see Sec. 2.3) to model a system where the size of the beam (and hence the grid pixelization) changes dramatically from one location to another, such as what happens when a strong focusing element is included.

In lieu of the ability to make direct experimental comparisons, we have performed a number of comparisons with the theoretical and computational work of other researchers, in order to individually check various aspects of our program's optical operations. We have also implemented many "sanity checks" in the code, and have monitored the results the ensure that they are sensible and theoretically justifiable. This process has led to many refinements of the program that have been documented earlier in this chapter, and has produced results that have given us confidence in the output of our simulation runs.

The rest of this section is devoted to giving a description of the checks and verification tests that we have performed in order to guarantee the realism and accuracy of the model.

2.6.1 Analytical, empirical, and algorithmic verifications

First of all, we discuss the common-sense tests, or "sanity checks", which we have done as a matter of course to make sure that the program does not behave in an overtly unphysical manner. The most basic of these tests are checks of the results of runs performed with perfectly smooth (but finite-sized) mirror maps. The built-up cavity powers for these runs precisely match the predictions of analytical resonance formulas for Initial- and Advanced-LIGO runs; for example, the full GW-frequency response curves obtained for Dual Recycling runs (see Figure 4.4 of Sec. 4.2.3) are precisely in accord with theoretical predictions. These theoretical agreements are obtained by factoring in the effects of the ~1-2 ppm of diffractive loss due to the apertures of the Fabry-Perot arm cavity mirrors. On the whole, when testing a run, we can meticulously account for all of the power that has been injected into the interferometer, whether it has been lost by supporting cavity resonance (i.e., absorbed by the mirrors after many bounces), leaked out through the exit-port of the system, leaked back to the laser in reflection from the Power Recycling Mirror, or scattered out beyond the mirror edges (which we can quantify by measuring what is absent after one propagation, and/or measuring the power eliminated by a finite mirror aperture). The total always comes to an amount acceptably close (given uncertainties) to the 1 Watt of carrier and/or sideband power that is exciting the system.

For runs with deformed mirrors, the modal structure of the interferometer electric fields remains consistent with theoretical expectations. Mirror tilts primarily create TEM_{10} and/or TEM_{01} modes, with the ratio of the two depending upon the orientation of the tilt axis. Mirror curvature mismatch leads to power in the "donut" modes, TEM_{20} & TEM_{02}, as expected. Polynomial deformations (Zernike polynomials) with even parity

lead to the presence of even-index beam deformation modes, and odd-parity Zernikes lead to odd-index deformation modes. Perturbations that are anti-symmetrically applied to the inline and offline interferometer arms cause power to leak out of the beamsplitter exit port, while symmetrically-applied perturbations cause power to leak out in reflection from the Power Recycling Mirror (though both types of losses are suppressed as much as possible by interferometer optimization procedures). Direct visual inspection of the relaxed interferometer fields often shows clear signs of the type(s) of mirror imperfections which have been included. And the behavioral trends of the cavity e-fields are reasonable for even highly and non-symmetrically deformed mirrors: the losses experienced by the e-fields have a fairly good *quadratic* dependence upon the amplitude of the deformations ([52], and see Sec. 3.3), and the GW-sensitivity curves for Dual Recycling runs consistently retain their basic theoretical shape, essentially just decreasing in overall amplitude (see Sec. 4.2.3). Even the modal distributions of the deformed cavity e-fields show a clear dependence upon *not only* the deformations that are prevalent in the particular mirror maps being used, but also upon the exact resonance conditions of the interferometer cavities — e.g., TEM_{mn} modes with $m + n = 3$ are analytically predicted to be a *little* more resonant in the arm cavities than the other non-TEM_{00} modes, and the program does indeed find them to be somewhat more prevalent when a variety of modes are excited by mirror imperfections.

The accuracy of the program is also guaranteed by several algorithms that have been discussed previously, which enforce physical correctness. An algorithm that enforces energy conservation for every pixel of the constructed mirror maps has been described above in Sec. 2.3.3.2. Section 2.3.4 has elucidated how the pixelized grids are designed to provide a correct representation of the optics, including the adequate sampling of the Gaussian-profile laser beams, a method of propagation that is specially modified in order to curb "position-space" aliasing, and a careful examination of the physical limitations of deformations that can be modelled with our grid parameters. Section 2.3.3.4 has shown how measurement maps of real, deformed mirrors are carefully prepared for accurate representation in the simulation runs; and Sec. 2.4 shows how all of the electric fields in the interferometer are relaxed to their mathematically-specified steady-state configurations. Even the *relevance* of the program's output results for LIGO is made manifest by the numerous optimization procedures described in Sec. 2.5, which tailor the simulated interferometer to achieve an optimum sensitivity to gravitational waves. As a whole, the program has been designed with all of these detailed algorithms for verifying the physics of the model, as well as many internal cross-checks, in order to ensure that it correctly simulates physical reality.

Nevertheless, what is more important than self-consistency algorithms and checks are *external comparisons* of our work with results obtained by other researches, using different methods. Each of the following subsections is therefore dedicated to an external comparison which verifies an important aspect of our program's simulational method.

2.6.2 Testing the propagation engine and diffractive beam behavior

The first external comparison test was designed to check the core mechanisms of our program: beam propagations, and interactions with finite-sized mirrors. This test verifies the fundamental validity of the grid-based representation of the optics, the FFT-based propagators, the reflection and transmission operations for smooth ("perfect") mirrors, and the effects of diffractive loss due to the mirror apertures.

Our simulation-based "experiment" was the following: after injecting an initial excitation beam into a Fabry-Perot cavity, the cavity was sealed off, allowing the cavity e-field to bounce back and forth between the mirrors until fading away via transient decay. Each of the mirrors were ideal, 100% reflective surfaces, with no losses *except* for the cut-off of all light falling outside of the mirror apertures. The test was to determine the final decay rate of the oscillating beam as it settled into the spatial mode structure which was most persistent against dying away.

Though this cavity does not have resonant modes, per se, because it is not fed by a renewing excitation beam, it can be said to have eigenmodes of a sort in the sense that the *shape* of an "eigenmode" would be unchanged after a round-trip through the cavity, even though the amplitude of the mode would decrease during each apertured mirror reflection. These transient modes would be similar, but not identical to, the typical Hermite-Gaussian modes of a perfect cavity, with the differences being most pronounced at the edges of the beam near the mirror apertures. The initial excitation beam that is injected into to the system — mostly Hermite-Gaussian TEM_{00}, but with a little randomized field added to speed the "convergence" to the asymptotic e-field state — will excite all of the transient cavity modes to some greater or lesser degree, up to the limits of the discrete, finite gridding.

Each of the transient modes can be assigned an accompanying "eigenvalue" for the act of round-trip propagation through the cavity, as follows:

$$\text{R.T.}\{\vec{\Phi}_{mn}\} = \lambda_{mn} \cdot \vec{\Phi}_{mn} \tag{2.88}$$

These eigenvalues would not only contain the (usual) round-trip phases for each cavity mode, but they would also contain information about the round-trip diffractive loss experienced by that mode; unlike normal eigenvalues, therefore, these λ_{mn} have magnitudes less than unity. We can define a round-trip loss fraction for each mode as follows:

$$L_{mn} = 100\% \times (1 - |\lambda_{mn}|^2) \tag{2.89}$$

Since one mode (assuming system nondegeneracy) must have a lower round-trip loss fraction (i.e., a larger eigenvalue) than all the other modes, this "principle cavity mode" will decay away more slowly than all others, and the oscillating cavity e-field will eventually asymptote to this mode. Once the convergence to the lowest-loss mode has occurred, its loss fraction, L_{00}, is a good parameter for evaluating the accuracy of the numerical cavity

simulation.

Finite-mirror cavity simulations of this kind have been done previously by Li [53] and by Kortz and Weber [20]. Li used the scalar formulation of Huygens' principle to generate Fresnel-Kirchoff integral equations[1] for the round-trip propagation of light within the cavity, and numerically solved them for the two lowest cavity modes. The loss values for these modes, L_{00} and L_{10}, were plotted versus mirror aperture size, for a few cavity geometries or *g-factors*. Kortz and Weber used a somewhat more modern approach, which treated cavities in a more general way, to obtain similar results, for a (discrete) selection of aperture sizes, with L_{00} plotted against a continuum of g-factors. By comparing our L_{00} results to those of these other researchers, we can provide a program verification which not only spans different investigators, but a completely different method of optical propagations as well.

In Figure 2.10, we reproduce Fig. 3 from [20], containing the results of Li's and Kortz and Weber's results; to this figure which we have added our results (marked by X's) for the asymptotic loss fraction (presumably L_{00}) of the cavity e-field. As Kortz & Weber have done, our simulation was of a symmetrical resonator with mirrors having radii of curvature $R_1 = R_2 \equiv R$, bounded by "normalized" aperture sizes $s_1 = s_2 \equiv s$, which are defined according to:

$$s \equiv \frac{\alpha}{\omega} \tag{2.90}$$

with α being the physical mirror aperture radius, and ω being the spot-size radius (at the mirrors) of an idealized TEM_{00} beam that is matched to the cavity. This parameter s is related to the Fresnel Number [17], N_f, of the resonator by the formula $s = \sqrt{\pi N_f}$. The g-factor for this symmetric cavity (of length L) is given by:

$$g = 1 - \frac{L}{R} \tag{2.91}$$

In [20] it is stated that the diffraction loss problem is re-scalable, so that it can be completely characterized by s and g, regardless of different individual values of α, ω, L, and R. The result comparisons presented here will therefore be uniquely parameterized by s and g.

We have generated eight data points, using s = 1.0, 1.2, 1.3, and 1.5, each for two different g-factors: $1/3$ and $1/2$. For the record, all runs were performed with mirror curvature radii of R = 6000 m, where the different g-factors were produced by cavity lengths of, respectively, L = 4000 m and L = 3000 m. Argon-ion laser light was used for these

1. These integral equations are what we have replaced with FFT-based propagations, by using the paraxial beam approximation.

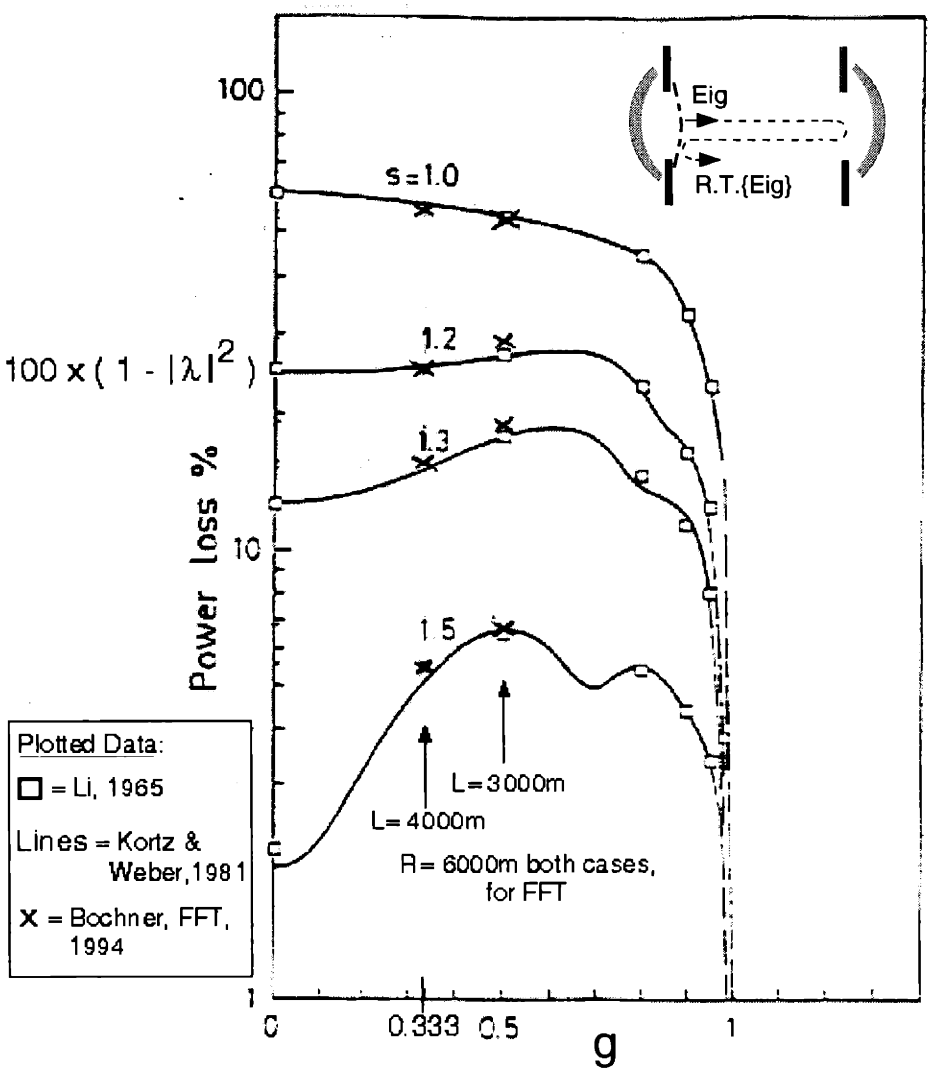

Figure 2.10: Power loss fraction versus g-factor for electric field round-trips through a cavity de-excited by diffraction losses from finite-sized mirror apertures.

runs, with a wavelength of $\lambda = 514$ nm. The mirror spot size radii are obtained via $\omega = \lambda L/(\pi\sqrt{1-g^2})$, and the physical aperture sizes can be computed via Eq. 2.90.

Our data points match up very well in Fig. 2.10 with the comparison data; especially given the unknown computational parameters used by the previous authors, such as gridding resolution, numerical accuracy used, their requirements for "convergence" to the final asymptotic state, and so on. We can therefore consider this to be a successful test of the fundamental engine for the physics of beam optics that we use in our code.

In addition to the level of program confidence that is achieved by this test, there was one other important result of these comparison runs: they provided our first direct detection of the effects of *position-space aliasing*. As described in Sec. 2.3.4.3, this form of aliasing occurs when a mode with very high spatial frequency (i.e., large propagation angle with respect to the main beam axis) propagates beyond the calculational window on one side, and then re-appears from the other side, coming far enough back in to land within the aperture of the destination mirror. In these diffraction-loss test runs, we have indeed found that there was an alias mode with a lower loss fraction than L_{00}; and even though it began by possessing a very low amplitude in the initial beam, after very many round-trips it was able to pass by the lowest-order mode and become the dominant mode in the system[1]. This aliasing-induced transition was marked by a sudden spike that interrupted the decreasing trend of "convergence errors", by the switch of the relaxing round-trip loss fraction from a good value of L_{00} to an unreasonably low value, and by the predominance of an extremely "ragged" mode in the cavity (the mode shown previously in Fig. 2.8), which replaced the smooth, TEM_{00}-like mode that dominated the asymptotic behavior of the system when the round-trip loss fraction was still equal to its correct L_{00} value.

The implementation of our anti-aliasing procedure from Sec. 2.3.4.3, however, has been completely successful at eliminating the effects of aliasing in these diffraction-loss test runs, and has led to the results shown above in Fig. 2.10. The enactment of this program test has therefore not only verified the current correctness of the algorithm we use for propagation, but has also been responsible for it making it correct in the first place.

2.6.3 Verification for geometric deformations: mirror tilts

The second program verification test that we performed involved runs with "imperfect" mirrors. The imperfections used were *geometric deformations*, i.e., tilts or mirror curvature errors; in this case, tilts were used. We compared our results to those generated by LIGO colleagues using a "Modal Model" program that they have developed [18], in order to compute the distribution of e-field power between a few low order modes (TEM_{00}, TEM_{10}, TEM_{01}, ...) for interferometer optics with geometric deformations. The Modal

1. This was possible because there was no permanent TEM_{00} excitation beam to continuously feed the cavity with modes that dwarfed the alias mode.

Model program involves direct analytical calculations for small matrices representing the modes in their model, and therefore serves as a good source of comparison for us, since it utilizes a completely different algorithm than the FFT-based propagations used by our program.

With the assistance of D. Sigg and N. Mavalvala, comparison runs were performed for the carrier and RF-sideband frequency beams, using decompositions of the resulting steady-state e-fields into their Hermite-Gaussian TEM_{00}, TEM_{10}, and TEM_{01} components. As depicted in Figure 2.11, two tilts were applied in the Initial-LIGO interferometer: a tilt causing a phase gradient along the y-axis (TEM_{01}-axis) for the inline arm cavity input mirror, and a tilt causing an x-axis (TEM_{10}-axis) phase gradient for the back mirror of the Fabry-Perot arm cavity. The e-fields were analyzed (modally decomposed) at the five locations indicated in the figure: inside each of the Fabry-Perot arm cavities, in the Power Recycling Cavity, at the exit (dark) port of the beamsplitter, and in reflection from the interferometer.

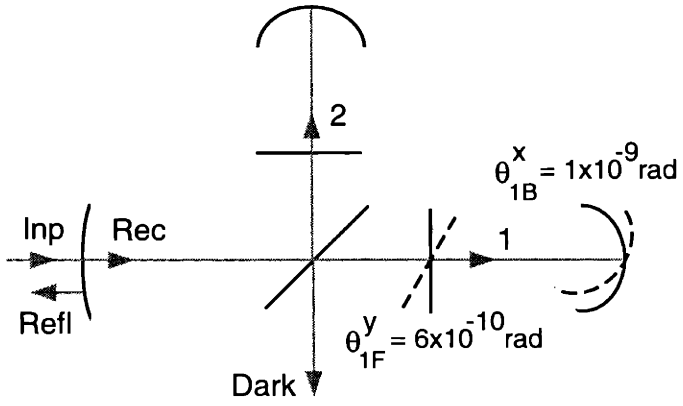

Figure 2.11: Simulation of an interferometer with tilted mirrors, for program verification.

Table 2.4 shows the comparison between the results of our simulation program ("FFT"), and those of the Modal Model ("M.M."), with all power values normalized to 1 Watt of carrier/sideband excitation laser light. The agreement is profound, with very few discrepancies of note (and always in modes with very small total powers). Given the fact that these two methods of performing interferometer simulations were developed in complete isolation from one another, with different algorithms and different basic capabilities, this comparison constitutes a strong verification of both of these LIGO simulation tools.

Carrier E-Field/ Position	TEM_{00} Power	TEM_{10} Power	TEM_{01} Power
Arm 1 (FFT)	2462.26	1.89859×10^{-4}	2.27572×10^{-5}
Arm 1 (M.M.)	2462.26	1.89859×10^{-4}	2.27572×10^{-5}
Arm 2 (FFT)	2462.26	4.23609×10^{-9}	3.25929×10^{-10}
Arm 2 (M.M.)	2462.26	4.23618×10^{-9}	3.25904×10^{-10}
Rec (FFT)	37.8732	7.44302×10^{-7}	5.72611×10^{-8}
Rec (M.M.)	37.8732	7.44284×10^{-7}	5.72659×10^{-8}
Refl (FFT)	9.18050×10^{-5}	2.05475×10^{-8}	1.55318×10^{-9}
Refl (M.M.)	9.18084×10^{-5}	2.05469×10^{-8}	1.55335×10^{-9}
Dark (FFT)	6.403×10^{-14}	2.90692×10^{-6}	1.14517×10^{-7}
Dark (M.M.)	2.241×10^{-12}	2.90686×10^{-6}	1.14536×10^{-7}
Sideband E-Field/ Position	TEM_{00} Power	TEM_{10} Power	TEM_{01} Power
Arm 1 (FFT)	0.140946	1.12104×10^{-6}	2.12684×10^{-7}
Arm 1 (M.M.)	0.140946	1.09742×10^{-6}	2.17620×10^{-7}
Arm 2 (FFT)	0.140946	8.39382×10^{-7}	1.97733×10^{-7}
Arm 2 (M.M.)	0.140946	8.16650×10^{-7}	2.02886×10^{-7}
Rec (FFT)	37.1454	7.16416×10^{-5}	1.78193×10^{-5}
Rec (M.M.)	37.1455	7.17492×10^{-5}	1.78251×10^{-5}
Refl (FFT)	1.11418×10^{-7}	1.98492×10^{-6}	4.92624×10^{-7}
Refl (M.M.)	8.88168×10^{-8}	1.98365×10^{-6}	4.92794×10^{-7}
Dark (FFT)	0.985081	1.93300×10^{-6}	1.01603×10^{-7}
Dark (M.M.)	0.985082	1.92905×10^{-6}	1.01376×10^{-7}

Table 2.4: Comparison of results between the Modal Model (M.M.) and our simulation program (FFT), for electric fields at the locations labelled in Figure 2.11.

2.6.4 More complex mirror structure: Zernike deformations

The third and last program verification test to be presented here involves runs with mirror deformations that are more complex than geometric deformations. Specifically, we used deformations in the form of various *Zernike polynomials* [32], functions that are typically used to characterize mirror structure because they are conveniently designed to model common types of phase distortion profiles, and on finite-sized mirrors. Though not as fully general as "real mirror" maps would be, Zernike-shaped deformations are complex enough to represent a good test of our program because they create nontrivial e-field structure, while still remaining simple enough for their effects to be calculated analytically, in order to present useful data for the comparison of results against those obtained by the simulation program.

The results of our program were compared to analytical work done by other LIGO colleagues [54], using a perturbative approach which finds (up to second order) the *new principle modes* for the light in a cavity with deformed mirrors, in terms of the unperturbed modes of the "perfect" system (i.e., perfectly smooth profiles on infinite-sized mirrors). An important approximation that is assumed in this method (and one which is easily satisfied), is that the mirror deformations are small enough so that after a single round-trip of any given mode, the amount of power that remains in the initial mode is much greater than the total power that has been converted into all the other modes.

The optical system for which this test was done is simpler than the complete LIGO interferometer[1]: it consisted of one (perturbed) Fabry-Perot arm cavity possessing a Zernike-deformed surface on either its front or back mirror, and one (idealized) Fabry-Perot cavity with perfectly smooth mirrors, with the (carrier frequency) light reflected from the two arm cavities getting re-combined by an idealized beamsplitter. For this modelled system, there were two figures of merit for comparison between the perturbative calculations and the numerical simulation program. The first one, α, is related to the degradation of power built-up inside the perturbed cavity due to the deformed surface, compared to the power inside the perfect arm cavity:

$$\alpha \equiv 1 - \frac{Power_{\text{perturbed}}}{Power_{\text{perfect}}} \tag{2.92}$$

The second figure of merit is the *Contrast Defect* (C. D.) at the idealized beamsplitter, which characterizes limitations to the totality of the dark-fringe that is achieved at the beamsplitter exit port. It is computed from the amount of power emerging from the "bright" and "dark" beamsplitter ports, as follows:

1. The perturbative computations, and the FFT-based simulation runs done to compare with them, were performed by P. Saha, utilizing a somewhat older version of the numerical simulation code.

$$\text{C. D.} \equiv 1 - \textit{Contrast} = 1 - \frac{P_{\text{bright}} - P_{\text{dark}}}{P_{\text{bright}} + P_{\text{dark}}} \approx \frac{2\,P_{\text{dark}}}{P_{\text{bright}}} \tag{2.93}$$

The Zernike functions (enumerated by indices n, l) that we used for these tests are derived from those detailed in [32], which were defined over the unit circle and normalized to give unity rms when integrated over that area:

$$\frac{1}{\pi} \int_0^1 r\,dr \int_0^{2\pi} [Z_{nl}(r, \theta)]^2 \, d\theta = 1 \tag{2.94}$$

To prepare these functions for the comparison runs, we "squeezed" these Zernikes laterally (without amplitude renormalization), in order to fit them within our finite-sized mirrors, which possessed aperture radii of 25 cm. Following that, the squeezed Zernike functions were multiplied by an overall scale factor of $\lambda_{\text{Argon}}/300$ (with $\lambda_{\text{Argon}} = 514$ nm), in order to simulate a "realistic" level of deformations for a very good mirror.

A comparison between the results of the perturbative calculations and those from runs of the numerical simulation program, using a variety of these modified Zernikes on either the front (i.e., input) or back mirror of the perturbed Fabry-Perot arm cavity, is presented in Table 2.5. Once again, we see that there is a very good agreement between the two methods used for most of these generated output values. The only substantial differences are for α in cases of even-index Zernikes, and those discrepancies were later attributed to a small degree of mis-centering of the mirror maps in the simulation program, a program which was fixed at an early stage of our research.

Front Mirror Perturbations				
Zernike	**Contrast Defect (C.D.)**		**Arm Power Decrease (α)**	
Z (n,l)	**Numerical (FFT)**	**Analytical (Perturbative)**	**Numerical (FFT)**	**Analytical (Perturbative)**
***Z* (5,1)**	$1.36\text{x}10^{-4}$	$1.35\text{x}10^{-4}$	$1.64\text{x}10^{-4}$	$1.68\text{x}10^{-4}$
***Z* (6,0)**	$2.54\text{x}10^{-5}$	$2.60\text{x}10^{-5}$	$4.01\text{x}10^{-5}$	$7.71\text{x}10^{-5}$
***Z* (7,1)**	$4.02\text{x}10^{-4}$	$4.10\text{x}10^{-4}$	$4.05\text{x}10^{-5}$	$3.79\text{x}10^{-4}$
***Z* (8,0)**	$9.39\text{x}10^{-5}$	$9.74\text{x}10^{-5}$	$1.99\text{x}10^{-4}$	$1.17\text{x}10^{-4}$
***Z* (9,1)**	$1.01\text{x}10^{-3}$	$9.80\text{x}10^{-4}$	$8.77\text{x}10^{-4}$	$7.70\text{x}10^{-4}$
Back Mirror Perturbations				
Zernike	**Contrast Defect (C.D.)**		**Arm Power Decrease (α)**	
Z (n,l)	**Numerical (FFT)**	**Analytical (Perturbative)**	**Numerical (FFT)**	**Analytical (Perturbative)**
***Z* (5,1)**	$1.11\text{x}10^{-3}$	$1.07\text{x}10^{-3}$	$6.14\text{x}10^{-4}$	$5.93\text{x}10^{-4}$
***Z* (6,0)**	$4.71\text{x}10^{-4}$	$4.54\text{x}10^{-4}$	$4.17\text{x}10^{-4}$	$2.12\text{x}10^{-4}$
***Z* (7,1)**	$2.18\text{x}10^{-3}$	$2.18\text{x}10^{-3}$	$1.11\text{x}10^{-3}$	$1.76\text{x}10^{-3}$
***Z* (8,0)**	$9.56\text{x}10^{-4}$	$9.28\text{x}10^{-4}$	$5.65\text{x}10^{-4}$	$7.16\text{x}10^{-4}$
***Z* (9,1)**	$2.69\text{x}10^{-3}$	$2.63\text{x}10^{-3}$	$3.42\text{x}10^{-3}$	$3.77\text{x}10^{-3}$

Table 2.5: Comparison of results between perturbative calculations and our numerical simulation program for Fabry-Perot cavity mirrors with Zernike polynomial deformations.

With the completion of this final verification test, we conclude the technical exposition of our interferometer simulation program, and proceed during the next two chapters to present the results of our research into the Initial- and Advanced-LIGO configurations.

Chapter 3

Simulations of an Initial-LIGO interferometer with optical deformations

In this chapter, we demonstrate the code with a set of runs incorporating measurement maps made from excellent quality mirrors. The simulations done were of an interferometer possessing the Initial-LIGO configuration, as depicted in Figure 1.5, and as discussed in Section 2.1.2. For these runs, the only interferometer imperfections used were the aforementioned measurement-derived deformation maps, the finite sizes of the mirrors, and a small amount of pure loss (50 parts per million = 50 "ppm") for each optic. This small loss component is included to account for the effects of power absorption in the mirror substrates and coatings, as well as high-angle scattering losses that are due to roughness finer than the resolution of our grids, which we cannot directly model[1]. The back mirrors of the arm cavities are each given an extra 10 ppm of loss to represent the finite transmission out of the interferometer that is allowed for control system purposes. These and all other interferometer parameters have been listed previously in Table 2.1.

A set of five runs with different mirror deformation maps has been performed with the simulation program. The first run was done with perfectly smooth mirror substrates and surfaces (a "perfect interferometer"), but the other four runs were done with: (i) deformed substrate maps for all of the mirrors, and, (ii) deformed surface maps for all mirrors from, respectively, the $\lambda/1800$, $\lambda/1200$, $\lambda/800$, or $\lambda/400$ families. These surface and substrate ("real mirror") deformation maps have been generated via the method discussed in Section 2.3.3.4, with $\lambda \equiv \lambda_{YAG} = 1.064\mu m$, and with the $\lambda/1800$ surface maps having *rms* deformations of ~.6 nm when sampled over regions of radius 4 cm, in the central portion of the optic. The four runs with their different families of surfaces all use the same mirror maps in the same places, with the only difference between the runs being a straight scale factor by which the normalization of the surface deformation amplitudes differ.

For all cases other than the "perfect mirrors" run, each (2-port) mirror was constructed from one surface phase map and one substrate phase map, from which the program derives mirror reflection and transmission operators via the method described in Sections 2.3.3.1–2.3.3.3. The beamsplitter is constructed from one input surface map and *two* input substrate maps (representing the inline and offline beamsplitter transmission paths, which dif-

1. This power, however, scatters completely out of the system in the long-baseline Fabry-Perot arm cavities, and can be treated as pure loss.

fer from one another), with its reflection and transmission operators being constructed by a method that is similar in principle to the one used for 2-port mirrors, but that is generalized to be appropriate for 4 input/output ports.

For each run in this set, the simulated interferometer is completely optimized in the sense of Section 2.5, with the final values of the optimized parameters being dependent upon the particular severity of mirror deformations that are used for that run. The results of these parameter optimizations will be displayed alongside the output powers of the simulated cavity e-fields, to be discussed in Sections 3.2 and 3.3, below.

Before proceeding on to the discussion of these runs, we first stress the importance of studying *realistically-deformed* mirror maps in simulations of the full, coupled-cavity LIGO system. Models which only consider simpler optical imperfections, such as geometric deformations (e.g., tilts, mirror curvature mismatch), may not capture the essential qualities of operating a gravitational-wave interferometer with imperfect optics. An example demonstrating this is the McClelland, *et al.*, study [49] of a Dual-Recycled interferometer[1], which employs Power Recycling, as does the Initial-LIGO configuration (though they modelled single-pass arms instead of Fabry-Perot arm cavities). Contrary to prior expectations [26], what they initially found in their study was that mirror curvature mismatch errors led to *more* signal degradation for their (supposedly deformation-tolerant) Dual Recycling configuration than for the configuration without Dual Recycling. The cause of this lowered tolerance to curvature error was the accidental near-resonance (in the additional, "Signal Recycling" cavity of the Dual-Recycled system) of the Hermite-Gaussian TEM_{20} and TEM_{02} modes that were created by the curvature mismatch. The expected behavior of improved tolerance to deformations for Dual Recycling was recovered by changing the average curvature of the Michelson arm mirrors, thus pushing the TEM_{20} and TEM_{02} modes away from resonance. A solution of that kind, however, would be much more difficult to implement in an interferometer with realistic mirror deformations, in which *every* mode is excited to some greater or lesser degree, and any of them could be accidentally resonant somewhere within the system of cavities. It therefore seems clear that the performance of a complex, coupled-cavity system such as a LIGO interferometer — either the Dual-Recycled version *or* the Initial-LIGO configuration with its three coupled cavities, one of them (the Power Recycling Cavity) being degenerately resonant for all modes — must be examined in the "full jeopardy" of realistic mirror deformations, in order to properly estimate its true capabilities.

The ensuing chapter is organized as follows: Section 3.1 begins with a general survey of the output information that is created by the simulation program, in order to provide a framework for understanding the results that are to be subsequently presented. Sections 3.2 and 3.3 show the output data from the five sample runs that we have performed, and

1. A Dual-Recycled interferometer has been depicted in Figure 1.6, and simulations with Dual Recycling will be discussed in Chapter 4.

they provide both a quantitative comparison of these results with the performance requirements of the Initial-LIGO interferometers, and a qualitative analysis of what is revealed about typical interferometer behavior. Section 3.4 extends the analysis to estimate the effects of "realistic" mirror deformations upon LIGO science capabilities; and Section 3.5 concludes with a general discussion of the investigations in which our simulation program has been used as a wide-ranging research and development tool within the LIGO Project.

3.1 Figures of merit for an Initial-LIGO detector

A complete set of runs with the simulation program — including the (separately done) carrier and sideband runs, with all desired optimizations included — results in a complete specification, to a significant level of realism, of the final, steady-state of the interferometer. Of the vast amount of output data available, here are the primary quantities which we can examine, either as raw output data or after some degree of post-processing:

- The *relaxed powers* at any desired location in the interferometer, both the total power in the electric fields and the power in any specified subset of beam modes (we choose to compute the TEM_{00}, TEM_{10}, and TEM_{01} components). We have currently set up the program to compute the e-fields and report the e-field powers at the locations indicated previously in Fig. 2.2. These powers are labelled P_{rec} (in the Power Recycling Cavity, i.e., "PRC"), P_{inline} and $P_{offline}$ (in the Fabry-Perot arm cavities), P_{bright} and P_{dark} (at the output ports of the beamsplitter), P_{refl} (in reflection from the interferometer), and P_{out} (or "$P_{exit\text{-}port}$", at the interferometer exit/signal port). For the Initial-LIGO configuration, $P_{dark} \equiv P_{out}$; but for the Advanced-LIGO configuration of Dual Recycling, they are distinct, and the power at an additional location, P_{dual}, is recorded.

- The *complete relaxed e-fields* at all desired locations, available either for viewing with a graphics program [e.g., 56], or for *modal decomposition* into many Hermite-Gaussian TEM beam modes, via a small program which we have written for that purpose (see Appendix A). Modal decomposition greatly assists in the interpretation of interferometer conditions; for example, residual mirror tilts are denoted by the presence of significant TEM_{10} or TEM_{01} power, and beam mismatch is indicated by significant power in the TEM_{20} and TEM_{02} modes. Most importantly, the gravitational wave signal itself is derived solely from the TEM_{00} components of the relevant e-fields.

- The *carrier contrast defect*, which quantifies how well or how poorly a carrier dark-fringe at the beamsplitter exit port was achieved for the imperfect interferometer. Defining P_{bright} as the power re-entering the Power Recycling Cavity (PRC) from the

beamsplitter, and P_{dark} exiting the interferometer, the contrast defect is given as:

$$1 - \text{Contrast} \equiv 1 - C \equiv \text{C. D.} = 1 - \frac{P_{bright} - P_{dark}}{P_{bright} + P_{dark}} \approx \frac{2\,P_{dark}}{P_{bright}} \tag{3.1}$$

A large contrast defect implies a substantial carrier power loss at the beamsplitter (and thus a reduced power buildup in the Fabry-Perot arm cavities), 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*, including all cavity lengths, the exact sideband modulation frequency, the reflectivity of the Power Recycling Mirror, the Schnupp asymmetry length, and (after post-processing optimization) the sideband modulation depth. In addition to guaranteeing the optimized performance of the simulated interferometer, knowledge of these parameters often has intrinsic value in terms of LIGO design considerations.

- The *GW-strain-equivalent shot noise spectral density*, $\tilde{h}_{SN}(f) \equiv \tilde{h}_{SN}(\nu_{GW})$, of a single LIGO interferometer. This most significant output quantity allows us to directly evaluate the sensitivity of the LIGO detector to astrophysical sources of gravitational waves, given interferometers with realistic optical deformations, and including full consideration of LIGO's heterodyne gravitational-wave (GW) detection scheme.

As discussed in Section 1.2, the dominant noise sources for the Initial-LIGO interferometers are expected to be seismic, thermal, and photon shot noise; and insofar as optical deformations principally affect the *shot noise* part of the detection limit[1], it is the computation of $\tilde{h}_{SN}(f)$ which we focus upon. Nevertheless, this shot noise function can be combined with the *expected* levels of seismic and thermal noise in order to represent the overall noise envelope, $\tilde{h}(f)$, of a LIGO interferometer with a particular set of imperfect optics. It can also be converted (see Sec. 3.4) into mathematical forms which are well-suited for comparison with astrophysical predictions, in order to determine the effects of optical deformations upon the capabilities of the Initial-LIGO to detect gravitational waves of reasonable, anticipated strengths.

Figure 3.1 displays seismic, thermal, and shot noise curves that are plotted versus GW-frequency, f, as noise spectral densities expressed in terms of the gravitational-wave Fourier amplitudes, $\tilde{h}(f)$, that would induce an equivalent signal response in a LIGO interfer-

1. The quality of the optics — and hence the quantity of circulating interferometer power — does indeed determine the level of "radiation pressure" noise in the interferometer (which is a *random force* noise source like seismic or thermal noise); but the power in the interferometer excitation laser beam will not be high enough for this radiation pressure fluctuation noise to be significant for the Initial-LIGO detectors (though it would be for "Advanced-Subsystem" LIGO detectors using ~100 W of excitation laser power [12]), and thus we can neglect this noise source in our studies.

ometer. These particular curves are the Initial-LIGO requirements [55] for the maximum contributions that would be acceptable from each of these different noise sources. The total noise envelope is obtained from these individual noise curves by adding them together in quadrature, i.e., incoherent addition of uncorrelated noise is assumed.

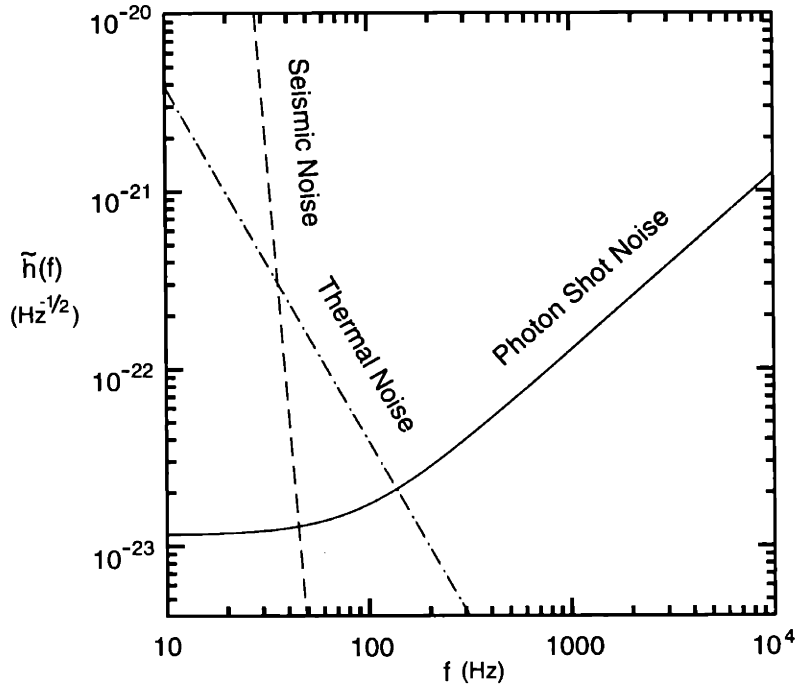

Figure 3.1: Requirement curves for the primary noise sources that are expected to limit the gravitational-wave sensitivity of the Initial-LIGO interferometers.

The full calculation of $\tilde{h}_{SN}(f)$ is given in Appendix D. Here we present the resulting expression, in terms of quantities obtainable from the output results of our numerical simulations. The shot noise sensitivity limit is expressed as equivalent to a monochromatic, "transverse-traceless" (TT) gravitational wave of the form[1] [2]:

$$h^{TT}(f) = \sqrt{2}h\cos(2\pi ft + \phi_0) \tag{3.2}$$

that is incident upon a single interferometric detector with optimal incidence angle and polarization, which would produce a unity signal-to-shot-noise (S/N) ratio when sampled over unity bandwidth (i.e., 1 second integration time). With these definitions, we have:

1. This definition differs by $\sqrt{2}$ from the one given in Eq. 1.3, because here (and below) h is the root-mean-square amplitude of the gravitational wave, rather than the peak amplitude.

$$\tilde{h}_{SN}(f) \equiv \left\{ \frac{S(f)/h}{N} \right\}^{-1} = \tag{3.3}$$

$$\left\{ \sqrt{\eta\, \nu_{Carr} P_{las} / h_{pl}}\; J_0(\Gamma)\, J_1(\Gamma)\, \frac{4\sqrt{2}\pi\tau_s \cdot r_{FP\ arm\ back} \cdot t_{FP\ arm\ input}}{\sqrt{1 + (4\pi\tau_s f)^2}} \right.$$

$$\left. \times \frac{\sqrt{P^{00}_{SB,\ exit\ port}}\ (r_{BS}^{A.R.\text{-}side} \cdot \sqrt{P^{00}_{Carr,\ inline\ FP\text{-}arm}} + t_{BS}^{offline} \cdot \sqrt{P^{00}_{Carr,\ offline\ FP\text{-}arm}})}{\sqrt{J_0(\Gamma)^2\, P^{tot}_{Carr,\ exit\ port} + 3\, J_1(\Gamma)^2\, P^{tot}_{SB,\ exit\ port}}} \right\}^{-1}$$

where: P_{las} is the total excitation laser power (before radio frequency modulation) in Watts, h_{pl} is Planck's constant, ν_{Carr} is the carrier laser frequency, η is the quantum efficiency of the photodetector at the signal port, $J_0(\Gamma)$ and $J_1(\Gamma)$ represent, respectively, the division of laser power between the carrier and sideband fields[1], and terms like $P^{tot}_{Carr,\ exit\ port}$, 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 program 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 in Eq. 3.3 we do not distinguish between the upper and lower radio-frequency sidebands, but (as discussed in previous sections) we simulate one of the two, and use those power results for both.

The remaining variables in Eq. 3.3 — namely $r_{FP\ arm\ back}$, $t_{BS}^{offline}$, etc. — are approximate, analytical values for the (amplitude) reflection or transmission coefficients of the interferometer mirrors that are indicated by each of their subscripts (with "FP" = Fabry-Perot arm cavity). These variables appear because the e-fields inside the various interferometer cavities must all be brought to the signal port for the determination of the output response of the heterodyne detection scheme to a gravitational wave. They each replace, to very good approximation, only single bounces (or transmissions) of an e-field with a mirror on the way to this destination.

Finally, the quantity τ_s is the *effective storage time* of the GW-induced signal sidebands in the realistically-deformed Fabry-Perot arm cavities. Since the explicit calculation of the buildup of these GW-induced signal fields in the arm cavities would involve an additional set of e-field relaxation procedures that are not performed by the (Initial-LIGO)

1. The division of power between the carrier and its RF-sidebands are as defined by Eq. 2.87, in Section 2.5.6. The above shot noise formula, Eq. 3.3, is the expression that we optimize with respect to Γ, the sideband modulation depth, to maximize the GW-sensitivity of the interferometer.

simulation program, the effective storage time in the Fabry-Perot arm cavities must be approximated, as follows:

$$\tau_s \cong \frac{L}{c} \cdot \frac{1}{\sqrt{P^{00}_{\text{Carr, Recyc. Cav.}}}} \cdot \frac{1}{2} \left\{ \frac{\sqrt{P^{00}_{\text{Carr, inline FP-arm}}}}{t_{BS}^{\text{inline}} \cdot t_{\text{FP arm input}}} + \frac{\sqrt{P^{00}_{\text{Carr, offline FP-arm}}}}{r_{BS}^{\text{ref-side}} \cdot t_{\text{FP arm input}}} \right\} \quad (3.4)$$

where L is the length of each arm cavity, c is the speed of light, and where we have performed a properly-weighted average over the results for the inline and offline arm cavities.

We note that the dependence of $\tilde{h}_{SN}(f)$ upon the GW-frequency is entirely contained within the expression $\sqrt{1 + (4\pi\tau_s f)^2}$, so that it has two basic regimes which are separated by a "knee" or pole frequency, $f_{pole} = 1/(4\pi\tau_s)$, such that:

$$\tilde{h}_{SN}(f)\Big|_{f \ll f_{pole}} \propto \text{constant} \quad (3.5)$$
$$\tilde{h}_{SN}(f)\Big|_{f \gg f_{pole}} \propto f$$

These two regimes are apparent in the shot noise curve that has been shown in Fig. 3.1. (Note that this simple functional behavior is only accurate for $f \ll f_{FSR,\ \text{arm cavities}} \equiv 2\,L/c$.)

3.2 Comparison of results with LIGO Project requirements

The results of our runs with the perfect, λ/1800, λ/1200, λ/800, and λ/400 mirror families are summarized in Table 3.1. Many of the output quantities described in the previous section are given here, including: (i) the relaxed carrier and sideband powers (total and/or in the TEM_{00} mode) at various locations in the interferometer, (ii) the *actual* carrier and sideband powers (the latter summed from *both* RF-sidebands) in Watts, that emerges from the beamsplitter signal port, given the total excitation laser power (6 Watts) and the optimized modulation depth, and, (c) the DC value and the effective pole frequency of the gravitational-wave sensitivity curve, with which one can heuristically construct the full shot-noise-limited GW-sensitivity function, $\tilde{h}_{SN}(f)$, as follows (c.f. Eq's. 3.3, 3.5):

$$\tilde{h}_{SN}(f) \equiv \tilde{h}_{SN}(0) \cdot \sqrt{1 + (f/f_{pole})^2} \quad (3.6)$$

A photodetector *quantum efficiency* of $\eta = 0.8$ has been assumed in all cases, in order to compute the optimized values of Γ, $\tilde{h}_{SN}(0)$, etc.

Additionally, some of these tabulated quantities are also given in the case of a *perfect output mode cleaner* operating at the signal port, that (ideally) acts to strip away all of the non-TEM_{00} light (which contributes only to shot noise) from the exiting beams, while passing all of the TEM_{00} light (which provides all of the gravitational-wave signal, though it contributes the rest of the shot noise) through to the output photodetector.

Quantity	Values For Specified Run				
Deformed Surfaces (RMS in wavelengths)	Zero	$\frac{\lambda_{YAG}}{1800}$	$\frac{\lambda_{YAG}}{1200}$	$\frac{\lambda_{YAG}}{800}$	$\frac{\lambda_{YAG}}{400}$
Deformed Substrates (Y/N)	No	Yes	Yes	Yes	Yes
Recycling Mirror Reflectivity *	98.61%	98.37%	98.07%	97.39%	93.90%
Schnupp Length Asymmetry (cm) *	9.0	12.3	13.5	15.9	24.7
TEM_{00} Carrier Power, Recycling Cavity †	72.40	61.54	51.84	38.41	16.37
TEM_{00} Carrier Power, Fabry-Perot Arm Avg. †	4726.7	4012.0	3374.3	2491.7	1042.4
TEM_{00} Carrier Power, Exit Port †	2.70×10^{-6}	9.29×10^{-6}	1.94×10^{-5}	4.56×10^{-5}	1.89×10^{-4}
Total Carrier Power, Exit Port †	2.15×10^{-3}	8.53×10^{-3}	1.43×10^{-2}	2.25×10^{-2}	3.65×10^{-2}
Carrier Contrast Defect, 1-C	6.02×10^{-5}	2.82×10^{-4}	5.62×10^{-4}	1.20×10^{-3}	4.73×10^{-3}
TEM_{00} 1-Sideband Power, Recycling Cavity †	59.08	28.32	25.01	20.17	10.66
TEM_{00} 1-Sideband Power, Exit Port †	.9067	.6745	.6955	.7344	.8196
Total 1-Sideband Power, Exit Port †	.9071	.7590	.7761	.8082	.8795
GW-Response Pole Frequency, f_{pole} (Hz)	90.32	90.38	90.45	90.61	91.45
Modulation Depth, Γ *	0.279	0.405	0.455	0.501	0.549
Absolute Carrier Exit-Port Power (mW)	12.41	47.1	77.2	118.6	187.8
Absolute 2-Sideband Exit-Port Power (mW)	207.1	358.9	458.3	571.3	737.7
DC GW-Sensitivity, $h_{SN}(0)$ *	4.79×10^{-24}	5.76×10^{-24}	6.41×10^{-24}	7.59×10^{-24}	1.20×10^{-23}
Re-Computed Values Assuming Ideal Output Mode Cleaner:					
Modulation Depth, Γ *	0.053	0.078	0.093	0.113	0.156
Absolute Carrier Exit-Port Power (mW)	1.6×10^{-3}	5.6×10^{-3}	0.12	0.27	1.10
Absolute 2-Sideband Exit-Port Power (mW)	7.7	12.2	17.9	28.1	59.6
DC GW-Sensitivity, $h_{SN}(0)$ *	4.61×10^{-24}	5.01×10^{-24}	5.48×10^{-24}	6.40×10^{-24}	1.00×10^{-23}

* Denotes Parameter Optimized by Program or During Post-Processing

† Denotes Quantity Normalized to 1 Watt of Carrier/Sideband Excitation Light Power

Table 3.1: Output results for the set of Initial-LIGO simulation runs that were performed using realistic deformation maps for the mirror surfaces and substrates. Except where otherwise noted, a total (pre-modulation) laser input power of 6 Watts is assumed, as well as a photodetector quantum efficiency of eta = 0.8.

The most significant issue to be addressed is whether or not an Initial-LIGO interferometer can perform according to Project requirements [57], given mirrors with realistic levels of optical deformation. That question is definitively answered, and in the affirmative, for all cases except for that of the worst surfaces which we simulated. First of all, Initial-LIGO interferometers are required to have a carrier gain of at least 30 in the Power Recycling Cavity; this target is achieved in all runs except for the (ultra-conservative) run performed with the $\lambda/400$ surface maps. In addition, the Initial-LIGO interferometers must have a contrast defect upper limit of $1 - C < 3 \times 10^{-3}$, a requirement that is also satisfied by all runs except for the $\lambda/400$ case (and anything better than ~$\lambda/500$ would suffice). Furthermore, it has been stated [12] that *Enhanced*-LIGO interferometers[1] should meet the more stringent requirement of $1 - C < 1 \times 10^{-3}$; this tougher requirement is itself satisfied by three of our five simulation runs (and anything better than ~$\lambda/900$ would suffice), which is very good considering the possibility of further improvements to mirror surface polishing by the time that Enhanced-LIGO interferometers are operational.

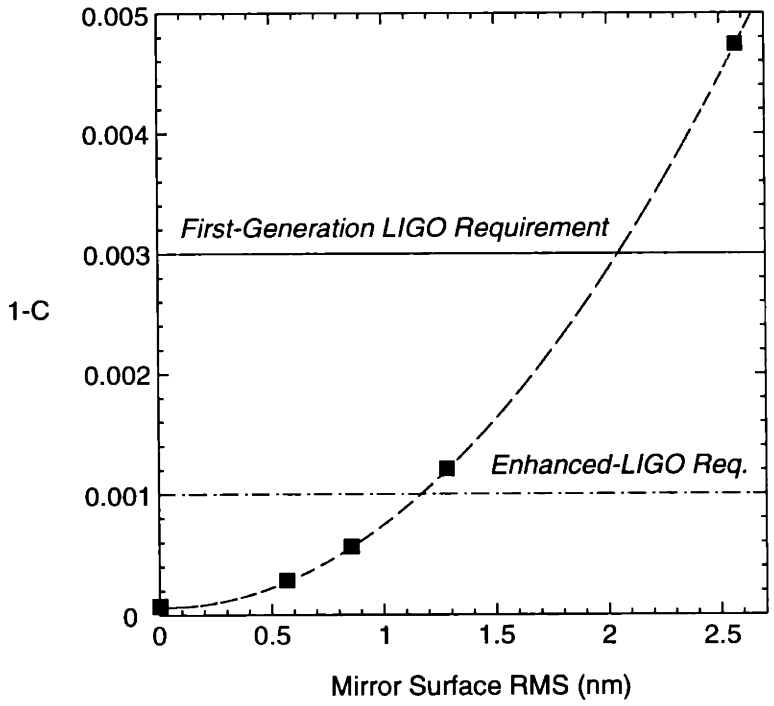

Figure 3.2: The carrier contrast defect, 1-C, plotted versus rms mirror surface deformations. The dashed line connecting the data points is a quadratic fit, and the points represent, respectively, the run with "perfect mirrors", then the $\lambda/1800$, $\lambda/1200$, $\lambda/800$, and $\lambda/400$ runs. The horizontal lines are the upper limits on contrast defect that are allowed for the "First-Generation" (Initial-) LIGO and Enhanced-LIGO interferometers, respectively.

1. "Enhanced" (i.e., "Advanced-Subsystem") LIGO refers to the implementation of better individual subsystems for an interferometer with an Initial-LIGO configuration; it is *not* the same as (though it may be the forerunner of) an Advanced-LIGO system, such as Dual Recycling.

The requirement upper limits for the contrast defect, along with a depiction of how the results of our runs compare with them, are shown in Figure 3.2. One note of caution 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, and this exiting power is especially high for runs with highly deformed mirrors, in which the sideband modulation depth must be increased (to increase the available local oscillator power) in order to help the GW-signal overcome the increased shot noise. Despite this potential problem, even if an appropriate signal detection apparatus that could handle this large amount of power cannot be supplied, then Table 3.1 shows that the addition of 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 close enough to the expressed ideal of stripping away the undesirable, non-TEM_{00}-mode light, without causing too much loss of the GW-signal-generating light.

The most fundamental Initial-LIGO performance requirement that we investigate as part of our research, is the necessity that the shot-noise-limited sensitivity curve ($\tilde{h}_{SN}(f)$) for an interferometer with deformed mirrors should fall within the bounds of the *GW-strain-equivalent noise envelope requirement*, which represents the official total-noise limit ($\tilde{h}(f)$) that is set for the full, 4 km baseline, Initial-LIGO interferometers [57]. In Figure 3.3, the $\tilde{h}_{SN}(f)$ curves that have been computed for the set of runs we performed

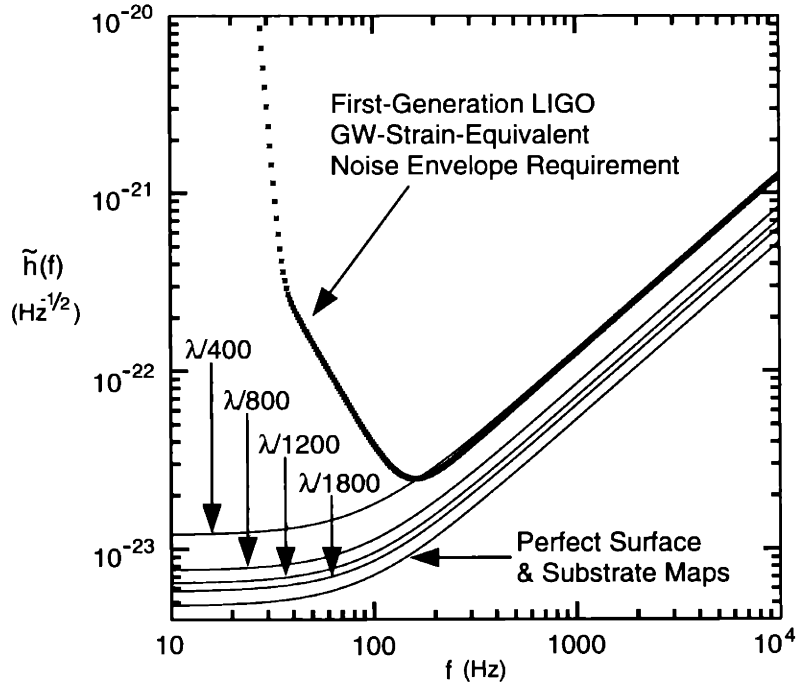

Figure 3.3: Comparison of the shot-noise-limited GW-sensitivity curves that are computed for each of the interferometer simulation runs, with the official, *GW-strain-equivalent* noise envelope requirement that is specified for Initial-LIGO interferometers.

are plotted against the data points that define the LIGO noise envelope requirement. All functions are shown as spectral densities. The roughly-defined regimes of seismic, thermal, and shot noise domination are apparent in the figure, and our $\tilde{h}_{SN}(f)$ curves can be compared to the shot-noise-dominated region of the total noise requirement. From this figure we can conclude, as was also found above, that all runs other than the $\lambda/400$ case succeed in meeting the LIGO shot noise requirement.

The overall conformity of the results presented in the section indicates that there is a very clear — and very strict, though realistically achievable — quality level for the polished (and eventually, coated) core optical components that must be reached in order to achieve the target performance of the Initial-LIGO interferometers.

3.3 Qualitative analysis of results

In this section, we use the data from Table 3.1 to explore the results from these simulation runs in a qualitative, though detailed manner. First of all, we note the fact that several of the effects caused by deformed optics possess a quadratic dependence upon the rms value of the deformation amplitudes. This is an expected result, since power scattered out of a Gaussian beam by mirror roughness scales like the *square* of the roughness amplitude, a functionality which is valid even when the deformations are spread over a range of spatial frequencies [52]. For example, the plot of the contrast defect (1-C) shown in Fig. 3.2, which is a form of loss that is generated by the coupling of power from the TEM_{00} beam mode into non-TEM_{00} modes by (long spatial wavelength) optical imperfections, was fit very well by a purely quadratic function of mirror deformation rms.

One can also determine the functional dependence of the cavity losses due to *high-angle scattering* from short spatial wavelength deformations, by computing an "effective loss" parameter to quantify the total scattered power (which vanishes entirely from the system, because of the finite mirror apertures) that is lost in the Fabry-Perot arm cavities. The effective loss for an arm cavity can be computed by setting the transmission of the cavity input mirror equal to its analytically specified value, while solving for the input and back mirror reflectivities (and hence, their losses) that would account, via Eq. 2.73, for the observed amount of power resonating in the deformed-mirror cavity. Doing this for each of the runs displayed in Table 3.1, the result is an effective loss function that (as expected) increases quadratically with rms from a calculated baseline value of ~51-52 ppm of loss for the case of perfect surfaces and substrates[1]; and this baseline value can clearly be interpreted as the sum of the 50 ppm of ("absorptive") loss that is put in analytically for each

1. For the worst case, $\lambda/400$ mirrors, the effective loss for the input and back arm cavity mirrors becomes as large as ~250 ppm per bounce; and even this effective loss does not encompass the total effect of how the perturbation of cavity e-fields due to mirror deformations can degrade interferometer performance.

simulated mirror, plus the ~1-2 ppm of diffractive loss occurring at the mirror edges for the Gaussian-profile beams.

Another parameter expected to have the same functional dependence on rms is the optimized reflectivity of the Power Recycling Mirror (Sec. 2.5.4), which should decrease quadratically from its value for the "perfect mirrors" run, because $1\text{-}R_{1,\mathrm{optim}}$ (or equivalently, $T_{1,\mathrm{optim}}$) is directly proportional (c.f. Eq. 2.75) to the total losses sustained (from scattering, absorption, imperfect contrast, etc.) in the interferometer. A fit of $R_{1,\mathrm{optim}}$ versus mirror deformation rms (not shown here) does indeed bear out this expected functional form[1].

Other than this quadratic dependence on rms for interferometer loss effects, additional insights can be gained through a close inspection of the results of these runs. For example, consider once more the behavior of the carrier contrast defect. It does not go *exactly* to zero, even for perfectly smooth mirror surfaces and substrates, but instead it levels off to a noticeably non-zero value of $\sim 6 \times 10^{-5}$. This effect can be explained by a careful consideration of the interferometer configuration, in combination with a visual inspection of the electric field that emerges from the exit port for the "perfect mirrors" case. This electric field is depicted[2] in Figure 3.4.

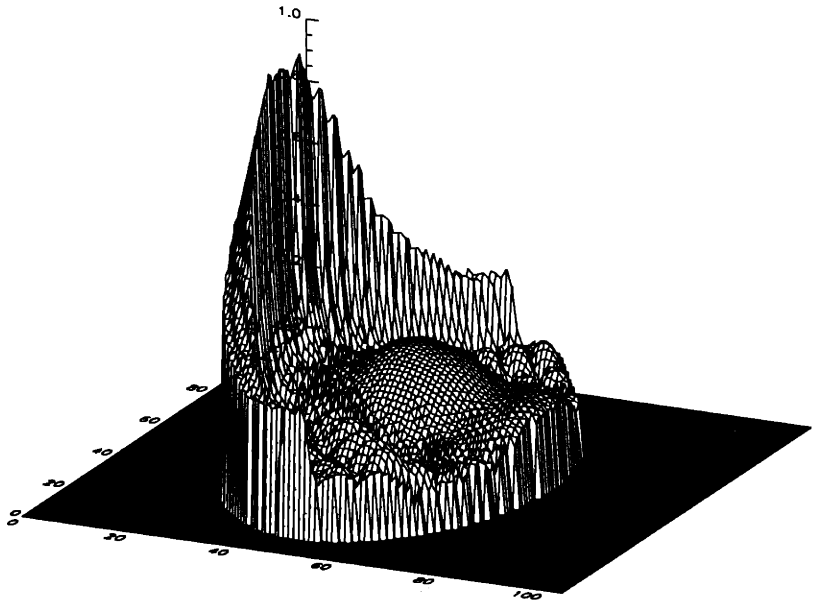

Figure 3.4: Plot of (intensity)$^{1/8}$ for the carrier electric field emerging from the exit port of the Initial-LIGO beamsplitter, for the run with perfect mirror surfaces and substrates.

1. We note, however, that *small* departures from pure quadratic forms for these quantities do appear as the mirror surface rms goes to zero, since other sources of loss (such as the presence of deformed substrates, which are constant and do not scale with rms value) become significant compared to the surface deformations. Nevertheless, the quadratic behavior remains very good for all cases.
2. More specifically, what is actually depicted is a close-up of the exit port e-field intensity distribution, raised to the one-eighth power in order to highlight its transverse structure.

Its structure tells the story: first, the generally oval (i.e., laterally "squashed") shape is due to the elliptical profile of the beamsplitter for the incident beams, caused by the 45° tilt of the circular beamsplitter with respect to inline and offline propagation axes (Sec. 2.3.3.3). Next, the round dimple in the middle of the e-field (composed of TEM_{20} and TEM_{02} mode power) is due to beam mismatch between the inline and offline arms, because of the macroscopic Schnupp asymmetry length (Sec's. 2.5.1, 2.5.5). And lastly, the intensity "wing" along one edge of the field, containing the vast majority of this electric field's power, is due to the fact that the inline and offline interferometer e-fields which pass through the thick, right-circular-cylinder beamsplitter do not see the same apertures. Rather, the effective aperture experienced by a beam during reflection from the anti-reflective side (i.e., the inline arm side) of the beamsplitter is *clipped* and *shifted* with respect to the aperture for transmission along the offline path, a fact which is due to the transverse displacement imparted to a beam passing through a substrate (with refraction index $\neq 1$) at non-normal incidence. This difference between the apertures that are experienced by beam components travelling along different paths (an effect that has been depicted in Figure 2.4, and quantified in Sec. 2.3.3.3), causes uncancelled power to leak out on one side[1] of the beamsplitter aperture, and leads to the telltale result that is observed in Fig. 3.4.

Although the net result of the three effects just mentioned is only to cause a small-amplitude e-field to leak out of the beamsplitter exit port, thus leading to a non-zero (though not very significant) contrast defect for even the perfect mirrors run, the important point to note is the relative ease and effectiveness with which the subtle behaviors of the interferometer laser fields (which in some cases may turn out to be highly significant) can be determined by analyzing the detailed results of our grid-based simulation program for the full-LIGO interferometer.

We wrap up this section by discussing an interesting (and quite counter-intuitive) effect apparent in the results shown in Table 3.1, which provides insight into the interplay between the parameter optimization procedures which have been discussed in Sec. 2.5, and the trade-offs that are made in optimizing the configuration of a full-LIGO interferometer.

To appreciate this effect, note the unusual behavior of the sideband power emerging from the beamsplitter exit port, as a function of mirror quality. This quantity represents the interferometer's efficiency at converting RF-sideband laser light that is introduced at the Power Recycling Mirror, into useful (i.e., TEM_{00}) power at the beamsplitter exit port to serve as local oscillator light for the heterodyne GW-detection scheme. In first going from

1. Alternatively, the beamsplitter could be given a position adjustment (transverse to the beam propagation axis) in order to *center* the exit port better, and thus reposition the lopsided "wing" of uncancelled power into two smaller wings on either side of the mirror that possess less total integrated power. This adjustment would, however, offset the center of the aperture for reflections at the reflective-side of the beamsplitter, thus affecting the electric fields resonating within the interferometer. In the end, some optimum position could be chosen that balances the effects to the e-fields both *inside of* and *emerging from* the interferometer.

the "perfect mirrors" case to the $\lambda/1800$ case, the amount of exiting TEM_{00} sideband power, $P^{00}{}_{SB,\,exit\,port}$, plummets — as expected — but after that it *increases* as the mirror surface deformations are made *worse*, all the way to the $\lambda/400$ case. This "improved" behavior of the sidebands for worse optics is explained as follows. First, the initial plummeting of $P^{00}{}_{SB,\,exit\,port}$ is due to the implementation of deformed substrates, which the sideband fields are far more sensitive to than the carrier fields are, since the sidebands are principally resonant in the Power Recycling Cavity (rather than in the Fabry-Perot arm cavities), and thus experience their most significant perturbation effects while passing repeatedly through the substrates of the Fabry-Perot input mirrors (which have their reflectivity coatings on the arm cavity sides of the mirrors) and the beamsplitter. Second, the use of increasing levels of *surface* deformation causes much more loss for the carrier beam (which undergoes many bounces between the reflective surfaces of the arm cavity mirrors) than for the sideband beam, and this has a crucial side-effect: the elevated carrier losses in the arm cavities cause a significant *decrease* in the optimized reflectivity of the Power Recycling Mirror (c.f. Eq. 2.75). As per the discussion of Sec. 2.5.5, the lessened amplification ability of the Power Recycling Mirror causes a shift towards a *larger* Schnupp asymmetry length via the L_{Asymm} optimization process performed for the sideband run (c.f. Eq. 2.82, or 2.84), which is tantamount to having a smaller storage time for the sideband light in the interferometer before it emerges at the exit/signal port, thus implying that there is less time for the sidebands to lose power due to interactions with both deformed surfaces *and* substrates[1]. The net result is an improvement in $P^{00}{}_{SB,\,exit\,port}$ for increasingly bad mirror surface deformations, as long as the substrate deformation levels are unchanged. A corollary effect is the reduction (or at least, the slower increase) of the modulation depth (Γ) as a function of mirror surface rms, which prevents *too much* of the excitation laser power from being modulated away into the sideband beams, and thus helps offset the decline of carrier power resonating within the interferometer arm cavities due to increasingly deformed surfaces.

At first glance, the effect discussed here may appear to be merely a quirk of our optimization procedures; but more fundamentally, it is a result of the dual nature of the LIGO interferometer: it attempts to simultaneously achieve optimal performance for both the carrier and sideband beams, although the two are optimized according to very different criteria — that is, the carrier must be optimally built-up *within* the interferometer, while the sidebands must be optimally *extracted* from it. Therefore, as discussed at the beginning of Sec. 2.5, the prescription for an "optimal" interferometer is not always as clear as one would be expect; and also, the (beneficial) interactions between the optimizable inter-

1. This, incidentally, is why we must optimize the Power Recycling Mirror reflectivity during the *carrier* run, and not during the sideband run; if $R_{1,\,optim.}$ were optimized for the benefit of the sidebands, then it would be driven immediately to zero (thus devastating the build-up of carrier power in the interferometer), followed by an optimization of the Schnupp length asymmetry which would cause the phase in Eq. 2.78 to become exactly equal to $\pi/2$.

ferometer parameters can lead to unanticipated behaviors which moderate the degradation effects of imperfections in the optics.

3.4 Impact of optical deformations upon LIGO science capabilities

In order to place the results of our runs into a scientifically relevant perspective, we will estimate here the effects of optical deformations upon LIGO's ability to detect gravitational waves from likely astrophysical events. Since the version of the simulation program discussed in this chapter was explicitly designed to model the Initial-LIGO interferometers, and may not be a good representation of an Advanced-LIGO system (due to changes in basic interferometer parameters, possible changes in configuration, etc.), we will focus upon GW sources that might be detectable with the Initial-LIGO interferometers.

Gravitational waves, which may be produced by a variety of astrophysical sources, are generally divided into 3 basic types: *bursts*, *periodic waves*, and *stochastic waves* [2, 5]. By restricting ourselves to sources for which an improvement in the Initial-LIGO shot noise sensitivity limit is likely to produce a significant increase in the number of detection events, or in the detectable range of some reasonably well-understood scientific parameter, we arrive at two good possibilities for study: *periodic* GW's that are emitted continuously from spinning Neutron Stars (i.e., Pulsars) which possess small deviations from axial symmetry, and *burst* GW's from coalescences of Black Hole/Black Hole orbital binaries.

Gravitational waves can approach LIGO's interferometric detectors from any angular location ("orientation") on the sky, as well as being in any superposition of its two possible (according to General Relativity) polarization states, as defined in Eq. 1.3. The analysis of GW-signals may involve correlations between several interferometers, as well as long integrations over time (for periodic sources), to obtain high signal-to-noise (S/N) ratios. However, the GW-strain-equivalent, shot-noise-limited sensitivity curves that we have quoted so far, $\tilde{h}_{\mathrm{SN}}(f)$, each refer to a 1-second integration with a single interferometer, for a periodic (monochromatic) GW that arrives at the interferometer with optimal orientation and polarization, and that is computed by designating a S/N of 1. In order to convert these simplified curves into formulas reflecting the effects of GW orientation, etc., for the fuller analysis below, we will use the formulations given by Thorne [2] for the generalized analysis of interferometer noise curves, which include distinct treatments for periodic and burst gravitational waves. While this analysis is very general and likely to be fairly accurate, it obviously does not include the effects of detailed "masks" of anticipated gravitational waveforms that would be derived from sophisticated modelling of the astrophysical GW sources themselves. We proceed, therefore, with the caveat that our analyses are necessarily ones involving rough estimates; not only because of the many inherent uncertainties about these GW sources, but also due to the highly idealized ways in which the theoretical

noise curves of a real interferometer have been formulated, as well as due to the simplifying assumptions which we adopt in regards to the complex process of data analysis.

On the other hand, though we must be cautious not to take the exact shot noise curves presented here too literally, we recognize that there are benefits that come with improved optics which are not apparent in our simple bottom line of $\tilde{h}_{SN}(f)$. These include, for example, reduced backscatter of "lost" power from the surrounding beamtubes back into the main optical pathway (which would cause phase noise at the signal detection port) [5], and smaller differences between reflections from the inline and offline arm cavities (which, if severe, would make the carrier contrast defect at the beamsplitter exit port worse, and thus worsen the common-mode noise rejection of the system as well as increasing the shot noise level). Thus if anything, the system performance benefits due to improvements in the optics that are shown in this chapter are *underestimates.*

3.4.1 Non-Axisymmetric Pulsars

To produce noise curves that represent the *total noise*, $\tilde{h}(f)$, of an interferometer, we form the quadrature sum of three functions: the seismic and thermal noise *requirement* curves (as shown in Fig. 3.1), along with the shot noise curve, $\tilde{h}_{SN}(f)$, which is computed for each of the runs by applying Eq. 3.3 to the data in Table 3.1.

These summed, spectral density noise curves must be converted into useful expressions for S/N. As stated above, we follow the conventions of Thorne [2], in which each of these total noise curves is converted to a dimensionless expression, $h_{3/yr}$, which can be compared to a formula for the "characteristic strength", h_c, of a GW source. For periodic sources, the condition $h_c = h_{3/yr}$ means that after one-third of a year of signal 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 frequency of the (monochromatic) GW, f (equal to twice the neutron star rotation frequency), and phase as known, Eq. 52a of [2] yields the following result for periodic-GW detection:

$$h_{3/yr}\big|_f \approx 1.7 \cdot \sqrt{5} \cdot \sqrt{10^{-7}\ \text{Hz}} \times \tilde{h}(f) \tag{3.7}$$

For a pulsar with *gravitational ellipticity* ε (representing the fractional deviation of the pulsar from perfect axisymmetry), and a rotation-axis moment of inertia of $10^{45}\ \text{g} \cdot \text{cm}^2$, which is radiating at frequency f at a distance r from the earth — and averaging over all possible inclination angles of the source with respect to the line of sight — we have (from Eq. 55 of [2]):

$$h_c\big|_f \approx 7.7 \times 10^{-20} \cdot \varepsilon \cdot \left(\frac{10\ \text{kpc}}{r}\right) \cdot \left(\frac{f}{1\ \text{kHz}}\right)^2 \tag{3.8}$$

In Figure 3.5 [1], $h_{3/yr}$ is plotted for each of the simulation runs, and displayed with them are two h_c curves, each one representing a *locus* of possible GW-strengths (plotted versus frequency, up to $f \sim 2$ kHz) for a pulsar with given values of ε and r. We have chosen $\varepsilon = 10^{-6}$, which should be below the limiting value determined by the "breaking strain" of their crusts [7], yet may be significant enough produce a detectable signal. While this value of ε is too large for millisecond pulsars, given typical limits on their rates of GW-induced spin-down [7], it may not be unreasonable for newly-formed pulsars, of which it has been estimated [60] that there may be ~25 such "new" pulsars in the galaxy. If they are assumed to be evenly distributed throughout the galactic disk, then there would be roughly an average of ~10 kpc of linear separation between neighboring ones; in other words, ~5 kpc could be the worst-case distance between the earth and the nearest such GW-source, if we are located directly in between the two pulsars closest to us.

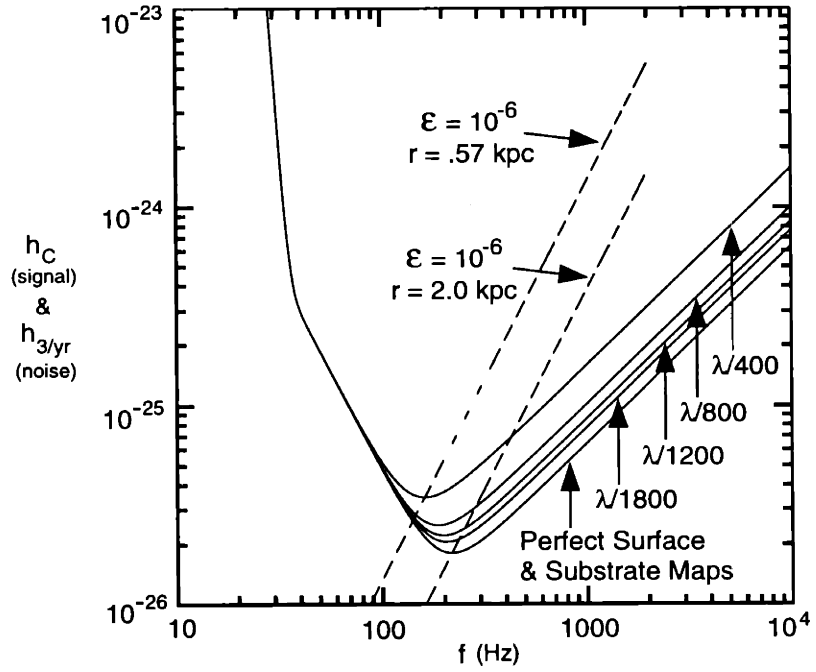

Figure 3.5: Plots of characteristic gravitational wave signal strength, h_c, versus GW-frequency for pulsars with specified ellipticity and distance from the earth (dashed lines), displayed against the dimensionless noise curves, $h_{3/yr}$, for periodic searches, that are computed from the output results of the simulation program runs (solid lines).

By setting $h_c = h_{3/yr}$ at the frequency of peak sensitivity (f_{peak}) for a given $h_{3/yr}$ curve, one obtains the rough estimate that in going from the $\lambda/400$ case to the "perfect mirrors" case, the *lookout distance* to which such a pulsar is detectable increases from ~.57

1. This figure differs slightly from the corresponding one (Fig. 2) in [58], because a somewhat older formulation was used there for the thermal noise requirement (specifically Eq's. 12 and 13 of [59], each with factor of 2 corrections). As a consequence, the quantitative estimates quoted in the ensuing discussion also differ slightly from those in [58], though the overall conclusions remain the same.

kpc to 2.0 kpc (this estimation also roughly maintains the "area under the curve" between h_c and $h_{3/yr}$, as depicted in Fig. 3.5, though it obviously ignores the frequency shift of the interferometer's sensitivity peak). This improvement increases the expected number of pulsar detections (for galactic disk pulsars) by $\sim(2.0/.57)^2 \sim 12$, or equivalently, it brings the lookout distance to a reasonable value for the detection of even a *single* $\varepsilon = 10^{-6}$ pulsar with the Initial-LIGO interferometers. Alternatively, for a pulsar at a given distance and GW-emission frequency, it provides a factor of $\sim(2.0/.57) \sim 3.5$ leeway in the smallest detectable value of ε. Even with the caveat that the existence of a strongly-emitting pulsar this close to the earth is unlikely (thus making the discovery of such pulsars somewhat of a longshot for Initial-LIGO detectors, even with improved optics), this improvement in sensitivity is still advantageous, in terms of making these discoveries more likely as the Initial-LIGO interferometers are gradually improved (e.g., as the laser power is increased).

3.4.2 Black Hole/Black Hole Binary Coalescences

A similar formulation is used for the analysis of burst GW's. Once again a quadrature noise sum, $\tilde{h}(f)$, is formed for each run, and with appropriate parameter averaging and optimal filtering assumptions, we have (from Eq. 34 of [2]):

$$h_{3/yr}\big|_f \approx \sqrt{5} \cdot \sqrt{\ln\left(f/10^{-8}\ \text{Hz}\right)} \cdot \sqrt{f} \times \tilde{h}(f) \tag{3.9}$$

For bursts, $h_c = h_{3/yr}$ means that after *coincidence detection* in two identical interferometers[1] 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 Gaussian noise[2] in the two detectors.

Consider a Black Hole/Black Hole (BH/BH) binary system with equal masses for the two objects, $M_1 = M_2 = 10M_{sol}$ (where M_{sol} = 1 Solar Mass = 2×10^{33} g), that is located a distance r from the earth, and evolves in frequency through the inspiral phase until it reaches the onset of the coalescence phase (i.e., merger and ringdown) at $f_{Merger} \sim 205\ \text{Hz}\,(20M_{sol}/M_{tot}) \approx 200$ Hz [61]. For these parameters, Eq. 46b of [2] yields (with a cumulative factor of 2 adjustment from factor of 2 corrections to Eq's. 29 and 44 of [2]):

$$h_c\big|_{f = f_{peak}} \approx 5.0 \times 10^{-21} \cdot \left(\frac{100\ \text{Mpc}}{r}\right) \cdot \left(\frac{100\ \text{Hz}}{f}\right)^{1/6} \tag{3.10}$$

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

1. The Initial-LIGO "detector" will involve the simultaneous operation of two full-length, 4 km Fabry-Perot arm cavity interferometers (plus one half-length, 2 km interferometer). Our calculations here are done assuming two identical, full-length interferometers working in coincidence.
2. Coincidence operation should theoretically eliminate the false signals caused by *non-Gaussian* noise.

In Figure 3.6, 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 (for the λ/400 run and for the "perfect mirrors" run) to skirt the high-detection-confidence threshold of $h_c = h_{3/yr}$ at f_{peak}, during the course of their inspirals. The first conclusion which we may draw from this diagram is that the best way to improve LIGO's sensitivity to these coalescence events is to improve the *low frequency* part of the noise spectrum, i.e., the seismic and thermal noise limits. Nevertheless, there is a measurable benefit from improving the shot noise limit: judging from this plot, one may estimate that in going from the λ/400 case to the "perfect mirrors" case, the lookout distance for detecting these coalescences is increased from ~125 Mpc to ~195 Mpc, or equivalently, it increases the expected number of detected events at a given distance r by a factor $\sim(195/125)^3 \sim 3.5$. The actual rate of black hole/black hole coalescence events is extremely uncertain (even their existence in uncertain), but good middle-of-the-road values that we can use as benchmarks are the "best estimates" that have been made by Phinney [10], and Narayan, *et al.* [11], which are, respectively: ~3 per year out to 200 Mpc (assuming a Hubble constant of $H_0 \sim 75$ km s^{-1} Mpc^{-1}), and ~1 per year out to 200 Mpc × (100 km s^{-1} Mpc^{-1}/H_0). Thus, the "perfect mirrors" run appears to put the detection of Black Hole/Black Hole binary coalescence events just within conceivable reach of detection for the Initial-LIGO interferometers.

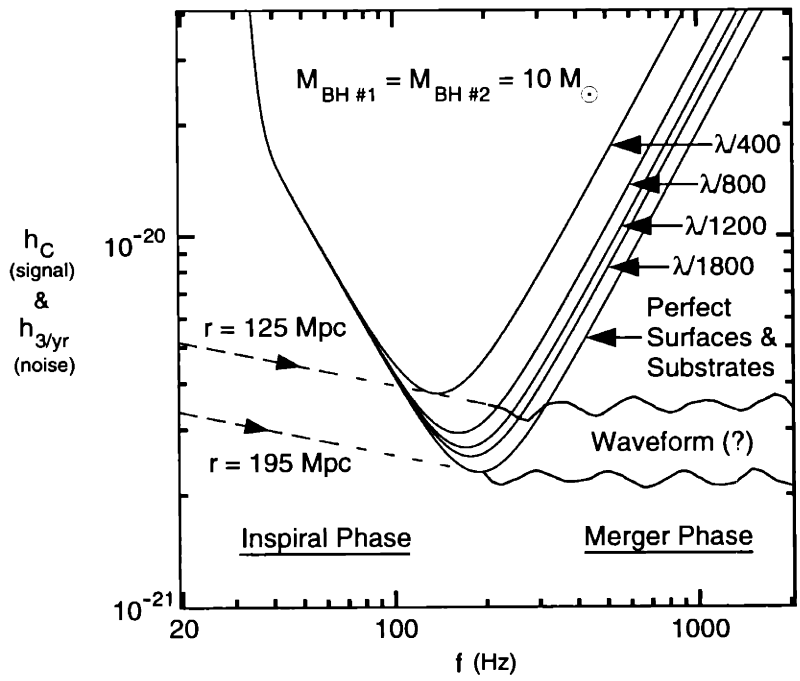

Figure 3.6: Plots of characteristic gravitational wave signal strength, h_c, as a function of the detector's peak sensitivity frequency, during the inspiral phase of 10 M_{sol} Black Hole/Black Hole binaries (dashed lines), displayed against the dimensionless noise curves, $h_{3/yr}$ (for burst searches), that are computed from the results of the simulation runs (solid lines).

Once again, we caution that these numbers must be treated as very rough estimates. There are many significant uncertainties in the above estimation process, such as: the poorly understood transition from the inspiral phase of the BH/BH binary to the merger part of the coalescence phase, which seems to fall near f_{peak} for the "perfect mirrors" run; the extreme sensitivity of these numerical estimates to the thermal noise edge, which may turn out to be very different from our adopted idealizations; the simplifications made in the signal-to-noise and data analysis calculations, and so on. The firm conclusion that one can draw, however, is that using optical components of the best achievable quality can indeed make a difference in whether or not the Initial-LIGO interferometers have a fighting chance to detect gravitational waves from these most promising astrophysical processes.

3.5 Project-wide uses of the simulation program

We have shown our simulation program to be useful for gaining physical insight into interferometer behavior, for proving that the Initial-LIGO performance requirements can be met with realistically obtainable optics, and for estimating the effects of optical deformations upon LIGO science capabilities. Due to the highly detailed nature of the model, it is also an effective and flexible tool for research and development in the LIGO Project. To date, it has been used to address several important design issues, such as[1]:

- Providing technical support and assurances to LIGO (by demonstrating robustness in the presence of optical deformations) for its selection of the Schnupp Length Asymmetry scheme for gravitational wave signal readout, over an alternative, external modulation (Mach-Zehnder) scheme [62].

- Providing technical support and assurances to LIGO in its transition from Argon-ion lasers ($\lambda = 514.5$ nm) to Nd:YAG lasers ($\lambda = 1.064\,\mu$m) for the interferometer excitation beam [63]. Demonstrating that interferometer performance is less sensitive to mirror distortions of a given level, when larger-wavelength laser beams are used[2].

- Collaborating with LIGO colleagues in the adaptation of the simulation program to the major LIGO prototype interferometers, including the Fixed Mass Interferometer (FMI) [3] and the Phase Noise Interferometer (PNI) [64] at MIT, and (currently) to the 40-meter interferometer at Caltech. In addition, providing assistance in the selection of optical parameters for the long-baseline LIGO interferometers [65], such as mirror

1. We are grateful to R. Savage, H. Yamamoto, B. Kells, K. Sliwa, S. Whitcomb, and D. Shoemaker for their help in the application of our simulation work to these LIGO-related issues.
2. Note, though, that $\tilde{h}_{SN}(f)$ does have an intrinsic overall proportionality to $\sqrt{1/\nu} = \sqrt{\lambda}$, as per Eq. 3.3.
3. To study the effects of optical deformations upon mirror alignment control systems.

curvatures, beam spot sizes, and mirror aperture sizes[1].

- Participating in the *Pathfinder Project* [37], the LIGO initiative charged to set the specifications for the optics that are being procured by LIGO through a cooperative effort of several vendors, followed by intermediate and final evaluation of the obtained core optics. Runs of the simulation program have provided information for several specifications that have been made, including tolerances for curvature errors[2] and surface deformations[3] in the polished (though yet uncoated) mirrors. In addition, evaluations of the mirror reflective-side and A.R.-side coatings are currently underway.

- Demonstrating that the inclusion of an (idealized) output mode cleaner would be beneficial, not only due to a modest improvement (~15%) in the shot-noise-limited GW-sensitivity, but also (and more significantly) due to the fact that an output mode cleaner would greatly reduce the total amount of power falling on the signal photodetector.

- Simulating the effects of refraction index variation due to thermal lensing [e.g., 66] on interferometer performance, resulting in rough preliminary estimates of ~15% degradation to the shot-noise-limited GW-sensitivity from either 0.6 parts per million of absorption in the mirror coatings, or 5 ppm of absorption per cm of thickness in the bulk (fused silica) substrate material [67].

Undoubtedly, many more issues such as these will arise in the near future, for which our simulation program can be a primary tool in addressing the important questions that are raised by the LIGO Project. Suggestions for future research directions with the LIGO simulation program, as well as general conclusions on what the results presented here mean for the importance of obtaining highest-quality optics for LIGO interferometers, will be discussed in Chapter 5.

---

1. Sufficiently large apertures is a particular concern for the perspective-foreshortened beamsplitter, for which limitations on minimum size and maximum allowable mirror position error (transverse to the beam) had to be specified. In these studies, the well-known [17] tendency of resonant laser fields to adapt their size to smaller profiles, in order to reduce diffraction losses from small mirrors, was observed.
2. To meet a criterion of < 5% degradation to the shot-noise limited GW-sensitivity, $\bar{h}_{SN}(f)$, runs of the simulation program for the Pathfinder Project have resulted in limits of +5%/-1% for the error in the curvature radius of the Power Recycling Mirror, +1%/-7% for the common-mode curvature errors of the Fabry-Perot arm cavity input mirrors, and ±1.5% for differential-mode curvature errors of the arm cavity back mirrors; all other combinations of curvature errors for the interferometer mirrors are less crucial.
3. Pathfinder polishing specifications require the rms surface figure deviations due to "roughness" (i.e., deformations in the spatial frequency band of ~.025 - 4.3 cm^{-1}) to be < 0.8 nm when sampled within the central 8 cm diameter of each fused-silica mirror blank, and < 1.6 nm within the central 20 cm diameters.

Chapter 4

Simulations of a Dual-Recycled interferometer with optical deformations

Having finished our studies of the Initial-LIGO interferometer configuration, we now proceed to the second area of inquiry that is central to this thesis — studies of an Advanced-LIGO configuration called *Dual Recycling* [24], for which optical elements possessing "realistic" surface and substrate deformations are included. In this chapter, our Dual Recycling studies will predominately be proof-of-principle tests, where for the first time, the wide variety of analytical and qualitative claims that have been made about the advantages of Dual Recycling will be put to the test by a quantitative numerical model, one that computes the effects of complex mirror deformations upon a fully-configured, Dual-Recycled LIGO interferometer.

This chapter is organized primarily into three sections. Section 4.1 discusses the theoretical benefits of Dual Recycling, including its ability to create a narrowbanded (shot-noise-limited) sensitivity function, as well as the predicted ability of Dual Recycling to both reduce the losses from the interferometer exit port, and to recycle such "lost" power for injection back into the system, so that it once again circulates in the Fabry-Perot arm cavities for the purpose of detecting GW-induced mirror motions. Section 4.2 discusses various demonstrations of Dual Recycling, including experimental and numerical work from the literature, as well as including as the main presentation of Dual Recycling simulation results that we have generated for this thesis. Section 4.3 continues with a discussion of some of the finer points of operating a Dual-Recycled interferometer with realistically imperfect optical elements, including the importance of using moderately-sized mirror apertures, and the necessity of implementing a *nondegenerate* "Signal Recycling Cavity" (c.f. Sec. 2.1.2) when tuning the system for peak response at high GW-frequencies. Finally, a summary of the conclusions from our simulation-based studies of the Dual Recycling configuration will be presented, along with the conclusions from our Initial-LIGO simulations, in Section 5.1 of the next (and final) chapter of this thesis.

4.1 The theory of Dual Recycling

4.1.1 The Dual Recycling optical configuration and tailoring of the Gravitational-Wave frequency response

4.1.1.1 Dual Recycling and the distribution of total interferometer sensitivity

The essence of the Dual-Recycling configuration is as shown in Figure 1.6, and again in Figure 2.1. It is the same as the Initial-LIGO configuration, except for an additional mirror (M_6, or equivalently M_{dual}) placed at the exit port of the beamsplitter, in order to *recycle* the Gravitational-Wave- (GW-) induced signal back into the interferometer before it can exit the system for GW-detection purposes. As with Power Recycling, this "Signal Recycling"[1] produces an overall amplification of the eventual GW-signal that exits through M_{dual}, but *only* if the recycled signal power is at the right GW-frequency to be properly resonant in the recycled system; otherwise, the presence of the large mirror reflectivity, R_{dual}, reduces the amplitude of the escaping GW-signal. Unlike Power Recycling, therefore, Signal Recycling is not a frequency-independent (broadband) signal amplification process.

The basic concept of what Signal Recycling achieves can be understood in terms of the "Sensitivity Theorem" presented in the thesis of Jun Mizuno [66]. This theorem states that for an interferometric GW-detector behaving at optimum efficiency, the *sensitivity* times the *detector bandwidth* is a constant that will be proportional to the *energy stored in the detector* for interaction with the incoming GW's. This sensitivity-bandwidth product will be the same for any configuration of detector, as long as the stored energy is the same. In view of the Sensitivity Theorem, we can define Dual Recycling as a configuration which narrows (and shifts) the bandwidth of an Advanced-LIGO interferometer in order to enhance the interferometer sensitivity, by moving most of it into a more scientifically and technically relevant range of GW-frequencies. Dual Recycling, as will be seen shortly, is a particularly convenient configuration for response curve narrowing and tuning; it is also a detector specifically designed to minimize losses caused by imperfect optics, thus keeping the energy stored in the system large, to maintain high sensitivity for even severely deformed mirrors.

4.1.1.2 The formulation and implementation of Dual Recycling

As discussed in Section 1.3.4, it is advantageous to use the "GW-induced-sideband" picture of GW-signal generation for Dual Recycling, rather than to think in terms of arm length changes that impose time-dependent phase shifts on the carrier e-fields resonating in the interferometer. In the GW-sideband picture, it is possible to view the carrier frequency beams and the GW-induced sideband beams as separate, non-interacting entities (with the obvious exception of the initial generation of the GW-sidebands via parametric conversion). Our numerical simulation program can therefore begin by relaxing the carrier e-fields in a "static" interferometer[2], and then model the GW-signal fields later. These GW-sidebands are treated as independent propagating entities which possess frequencies that

1. By adding Signal Recycling to the pre-existing system of Power Recycling, one produces an interferometer with "Dual Recycling".
2. As noted in Section 2.5.2, however, the program does model a (semi-relaxed) GW-sideband e-field at pre-specified frequency, during the carrier run, in order to optimally tune the Dual Recycling sensitivity to that frequency.

are offset from the carrier laser frequency ($\nu = \nu_{Laser} \pm \nu_{GW}$), and that are generated (with *time-independent* amplitudes) from the carrier fields which are stored in the Fabry-Perot arms. These GW-sidebands then proceed to build up in the interferometer, in a way which we can compute via an iterative relaxation scheme (c.f. Sec. 2.4), ultimately receiving a doubly-resonant boost due to the Fabry-Perot arms and M_{dual}, so that they emerge from the interferometer signal port with enhanced sensitivity and altered response bandwidth.

We note here that it is sufficient to model all of the Gravitational Wave effects via the generation of GW-sidebands in the Fabry-Perot arm cavities, followed by *free propagation* of these sidebands throughout the interferometer, for two reasons. First of all, the arm cavities are the only interferometer locations possessing both very large energy densities and very large storage times, which cause them to be the sites of nearly all significant interaction between the carrier laser power and any incident GW's. And second of all, any GW-perturbations to the already-generated GW-sidebands can be ignored, since they are proportional to the square of the dimensionless GW-strain amplitude, h, and are thus completely negligible.

The amplitudes of the "plus" and "minus" GW-sidebands generated in the Fabry-Perot arms are as expressed at the beginning of Appendix D, the same as for an Initial-LIGO interferometer. The important differences for the Advanced-LIGO case have to do with what occurs when these GW-sidebands first exit the arm cavities, and subsequently build up within the coupled-cavity system of the entire Dual-Recycled interferometer. In the Dual Recycling version of our simulation program, the ultimately relaxed GW-sideband fields are computed for every relevant location in the interferometer, and the calculated GW-signal will be proportional to the sum of the (absolute value of the) amplitudes of the plus and minus GW-sideband fields[1] that emerge from the exit/signal port.

In order to obtain a qualitative understanding of the behavior of a Dual-Recycled interferometer, we divide the interferometer up into a few coupled (or decoupled) cavities, and consider their interrelated behaviors for the GW-sideband beams. These cavities are as defined in Section 2.1.2 (as per Fig. 2.1): the folded cavity spanning L_1 and L_2/L_3 is called the Power Recycling Cavity (PRC), and the folded cavity spanning L_6 and L_2/L_3 is called the Signal Recycling Cavity (SRC). Because of the dark fringe for the carrier beam at the beamsplitter, the GW-sideband fields are almost completely decoupled[2] from the PRC — the "bright" and "dark" beamsplitter ports are actually *reversed* for the GW-sidebands, compared to the carrier fields — and hence their qualitative behavior can be understood in terms of how they resonate in the "three mirror cavity" system of the SRC coupled to the two Fabry-Perot arm cavities, with the arm cavities conceptually "folded" into a single effective cavity.

Tuning the response peak of the Dual-Recycled system to a frequency different from

---

1. More specifically, the TEM_{00} components of these e-fields.
2. The GW-sidebands would not be significantly coupled to the PRC by the presence of a Schnupp Length Asymmetry (c.f. Sec. 2.5.5), even if such an asymmetry were to be included for the Dual-Recycled case, since the GW-sidebands are much closer to the frequency of the carrier beam than the radio-frequency (RF) sidebands used for the Initial-LIGO configuration are, and hence pick up negligible asymmetric phase shift while propagating through the Michelson arms (i.e., L_2 and L_3).

ν_{Laser} is achieved by the very simple mechanism of displacing M_{dual} along the beam propagation axis by a fraction of the laser wavelength. This gives the GW-sideband fields an overall phase for round-trips through the SRC, which must be cancelled out by an opposing phase picked up in the long-baseline, multi-pass Fabry-Perot arm cavities; but these arm cavities (which are held resonant for the carrier beam) will only impose this phase cancellation upon by a beam that is displaced from the carrier frequency by the proper amount. Thus the particular GW-frequency for which a beam undergoes double-resonance in the SRC/arm cavity system is directly specified via microscopic motions of M_{dual}.

This GW-frequency "tuning" mechanism is a highly nonlinear process: when the system is first tuned away from the carrier frequency, the GW-sidebands resonant in the effective "three-mirror-cavity" system are virtually resonant in the arm cavities themselves, so that small frequency changes lead to *large* phase changes in reflection from the arms, and hence require *large* displacements of M_{dual} are necessary for small changes to the "optimized" GW-sideband frequency, $\nu_{\text{GW}}^{\text{optim.}}$. But as $\nu_{\text{GW}}^{\text{optim.}}$ gets large, the arm cavities become non-resonant for beams at that frequency, and large changes in tuning the value of $\nu_{\text{GW}}^{\text{optim.}}$ are obtained with much smaller displacements of M_{dual}.

Analytically computed curves[1] for the GW-frequency-response of Dual-Recycled interferometers are shown in Figures 4.1 and 4.2. In Figure 4.1, we hold the tuning of the SRC fixed at $\nu_{\text{GW}}^{\text{optim.}} = 500$ Hz, and vary the value of the power reflectivity of M_{dual} (which is labelled R_{dual}); while in Fig. 4.2, we hold R_{dual} fixed at 0.9, and vary the value of $\nu_{\text{GW}}^{\text{optim.}}$. Each of these curves have been generated using the interferometer parameters listed in Table 2.1 (though with no Schnupp Length Asymmetry, c.f. Sec. 2.1.2), assuming an optimized Power Recycling Mirror reflectivity for the carrier fields (as per Eq. 2.75), and no crossover into the PRC by the GW-sideband beams. The curves depicted all represent *unnormalized* GW-signal amplitudes (i.e., not multiplied by the amplitudes of local oscillator fields needed for detection, and the GW-strain amplitude, h, is simply set to unity), but their comparative amplitudes are shown correctly to scale.

Note that the curve in Fig. 4.1 with $R_{\text{dual}} = 0$ represents the "Initial-LIGO" case, and that the *shape* of the "shot-noise-limited sensitivity" curve for each value of R_{dual} (such as $\tilde{h}_{\text{SN}}(f)$ depicted in Figures 1.4, 3.1, and 3.3 for the Initial-LIGO configuration), is obtained by simply inverting these GW-signal sensitivity curves. This is because the measurable GW-strain of the detector is inversely proportional to the amplitude of the GW-signal, and because shot noise is essentially "white" noise (see Appendix D), and does not alter the shape of the S/N curve.

Dual Recycling is said to operate in two different modes: the "Broadband" and "Tuned" ("Narrowbanded") modes. Broadband Dual Recycling has $\nu_{\text{GW}}^{\text{optim.}} = 0$ Hz, and

---

1. These curves are computed by a small Fortran program entitled, "dual_recyc_IFO_V-M_GW-signal_simulator.f", which we have written to simulate the response of an interferometer with perfect mirrors. Power Recycling, with an *optimized* mirror reflectivity, is included. The mathematical formalism used in this code is based on that of Vinet, Meers, Man, and Brillet [25].

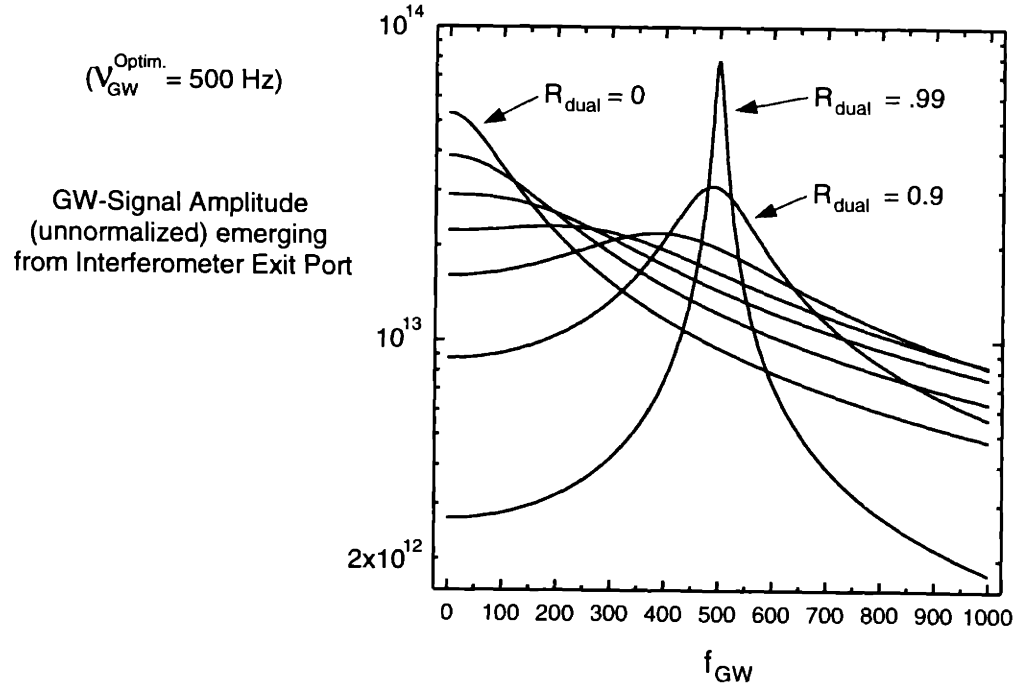

Figure 4.1: The GW-signal amplitude, proportional to the summed amplitudes of the plus and minus GW-induced sideband fields emerging from the interferometer exit port, is plotted versus GW-frequency for curves representing different Signal Recycling Mirror reflectivities. In order of increasing magnitude and sharpness at the optimization GW-frequency (200 Hz), the curves are for R_{dual} values of 0.0, 0.1, 0.3, 0.5, 0.7, 0.9, and 0.99.

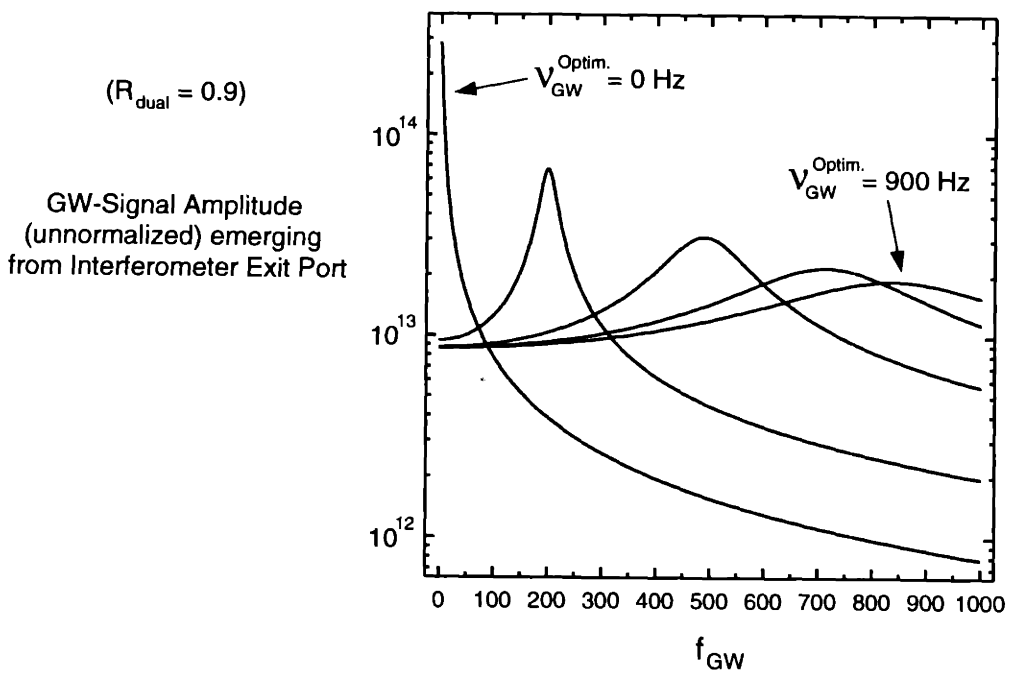

Figure 4.2: The GW-signal amplitude is plotted versus GW-frequency for a fixed value of R_{dual}, for curves representing different SRC-tuning optimization frequencies. From left to right, the curves are for optimization frequencies (in Hz) of 0, 200, 500, 750, and 900.

often uses small values of R_{dual}, more in order to reduce losses (see Sections 4.1.2 and 4.1.3) than to sharpen the GW-frequency response curve. But Tuned Dual Recycling can be used to offset the sensitivity peak from zero by large amounts (~few hundred Hz to 1 kHz or more), and may employ values of R_{dual} that are much closer to unity, in order to obtain a very deep sensitivity peak over a narrowly-specified frequency range. Such a narrowbanded sensitivity curve would be very useful for detecting periodic GW-emissions from objects such as pulsars, or for optimal sensitivity for [68] and even tracking of [69] the "chirp" of a coalescing Black Hole/Black Hole binary object as the binary's orbit decays, and the primary frequency of radiated GW's slowly climbs until merger ultimately occurs. Strong narrowbanding with Tuned Dual Recycling should also help a great deal in the detection of stochastic background sources (e.g., GW radiation from "cosmic strings"), since the sensitivity only scales as the one-fourth power of the bandwidth [2]. For Broadband Dual Recycling, however, obtaining very deep sensitivity peaks is not particularly useful (as well as making the term "Broadband" a misnomer), because at low GW-frequencies the principle limitations aren't from sensing (i.e., shot) noise at all, rather they are from *random force* noise such as seismic and thermal noise (c.f. Fig. 1.4), and in that region higher sensitivity does not improve the actual detection of gravitational waves. In our research, therefore, we have principally focused upon simulations of Tuned Dual Recycling with $\nu_{GW}^{optim.}$ offset significantly from zero.

An important point to note, which is apparent in Fig. 4.2, is that the frequency response peak does not maintain its shape as the tuning frequency is adjusted; rather, it spreads and flattens out as $\nu_{GW}^{optim.}$ becomes large. This effect is explained (as was the non-linearity of GW-frequency tuning) by the fact that the Fabry-Perot arm cavities become non-resonant as $\nu_{GW}^{optim.}$ is increased, and hence the double-resonance of the SRC/arm cavity system weakens until it becomes a single-resonance of the SRC alone. This means that a more reflective M_{dual} is necessary at higher tuning frequencies in order to obtain the same sensitivity curve as was achieved for lower tuning frequencies. While this does imply some limitations to the vaunted ease [24] of frequency tuning for the Dual Recycling configuration (which requires only microscopic adjustments of a single mirror), compared to more complicated procedures needed for other configurations such as Resonant Recycling [23], the tuning of $\nu_{GW}^{optim.}$ via microscopic displacements of M_{dual} will still be useful as long as R_{dual} is initially set high enough, or if the desired tuning range is small enough, so that the spreading of the sensitivity peak does not completely dilute the GW-response gain due to Signal Recycling.

An additional point, apparent in Fig. 4.1 (and to a lesser degree, in Fig. 4.2), is that the frequency of the sensitivity peak is not precisely equal to the GW-sensitivity optimization frequency, i.e. $\nu_{GW}^{optim.} \neq \nu_{GW}^{peak}$. In fact, $\nu_{GW}^{peak} \rightarrow 0$ as $R_{dual} \rightarrow 0$, even though $\nu_{GW}^{optim.}$ remains fixed (at 500 Hz, in this case). In general, the GW-response at a given frequency is not maximized by putting the peak at that frequency, but rather somewhat below it; tuning the response higher does indeed move the peak closer to $\nu_{GW}^{optim.}$, but it flattens the peak out so much that the actual response at $\nu_{GW}^{optim.}$ is diminished. This effect is due to the fact that

the Fabry-Perot arm cavities are still tuned to the carrier frequency, and hence provide a bigger arm cavity gain for smaller GW-frequencies; thus the peak of the compounded double-resonance occurs at a frequency, ν_{GW}^{peak}, which is dragged below $\nu_{GW}^{optim.}$. It would be possible to make the peak frequency *equal* to the optimization frequency (and perhaps gain a little bit in terms of the double-resonance peak) by "de-tuning" the arm cavities, hence giving us two parameters (L_{Arm}, L_{SRC}) for length adjustments instead of just one (i.e., L_{SRC} alone). That possibility, while intriguing, would lead to the difficult technical problem of maintaining an arm cavity "resonance" condition in the case where the carrier beam is not resonant in the arms.

In the simulation program (via methods such as those discussed in Sec. 2.5.2), and in our analytical calculations, the length of the SRC is tuned *specifically* so that the Signal-Recycling gain at $\nu_{GW}^{optim.}$ due to the SRC/arm cavity double-resonance is as made large as it can be. Although this does not put the sensitivity peak there, this is the most effective definition of SRC optimization that we can use, since it maximizes the interferometer output signal at the GW-frequency of interest[1].

Once the simulation program has successfully tuned the SRC length — via adjustments made during the process of relaxing interferometer e-fields for a GW-sideband beam at frequency $\nu_{GW}^{optim.}$ — the program must generate a full GW-frequency response curve, in order to compare the results for runs that employ imperfect optics with the analytical predictions made for perfect mirrors. The program does so by performing a number of independent runs (within the same, optimized interferometer) for a list of GW-frequencies that is specified by one of its input files. Though computationally intensive, this procedure allows the program to use the results for a series of discrete points along the GW-frequency axis for comparison with a continuous sensitivity curve that is calculated analytically. Obtaining a precise tracing of the continuous curves requires a large number of data points (often ~50 or more), and thus rather long run times (up to ~8-16 hours on the parallel *Trex* computer system mentioned in Sec. 2.2); therefore, the list of simulated GW-frequencies is carefully chosen to highlight whatever GW-signal peaks or other physics is expected to be most important in the run. Examples of such comparison plots will be shown in Section 4.2.3.

We wrap up this section on the GW-frequency response of a Dual-Recycled interferometer with the comment that for a Signal Recycling Cavity which is at (or near) "Broadband" tuning, the SRC — when considered as an isolated system — will be at (near) *anti-resonance* for the GW-sideband beam at the optimized, (near-) zero frequency. Since the GW-sideband is largely resonant in the arm cavities when $\nu_{GW}^{optim.}$ is small, and (as mentioned in previous chapters) a phase of π is picked up in reflection from a resonant cavity,

1. The only caveat to this procedure, is that the simulation program performs its L_{SRC} optimization for the "plus" GW-frequency alone. When the plus and minus GW-sideband amplitudes are summed together to produce the final ("single-sided") sensitivity curve, the tuning for $\nu_{GW}^{optim.}$ may be a little different from optimal; but this difference is very small, since the minus GW-sideband is always far off from (double-)resonance in the SRC/arm cavity system.

the anti-resonance in the SRC is necessary to counter the arm reflection phase and thus obtain the desired double-resonance in the SRC/arm cavity system. A corollary of this, is that as $\nu_{GW}^{optim.}$ is tuned to higher frequencies, and the arm cavities become off-resonant, the tuning adjustments to the SRC make it move from anti-resonance to near-resonance, for any and all power which enters it (since it is a "degenerate" cavity). These properties of SRC-tuning for different values of $\nu_{GW}^{optim.}$ become quite important in terms of losses due to power scattered into the SRC by deformations of the optics, and the resulting effects will be discussed in Section 4.1.2 (for TEM_{00} mode losses), and in Section 4.3.2 (for losses due to higher beam modes).

4.1.2 Noise reduction properties of Dual Recycling

As described earlier in Section 2.5.4, the mirror reflectivity and transmission parameters of a LIGO interferometer are chosen very carefully in order to minimize power output leaks (with the exception of small amounts needed for control system purposes), and to keep losses as low as can be achieved with real mirrors. Thus between the optimized Power Recycling Mirror and the high-reflectivity arm cavity back mirrors, it is very hard for light that is perturbed by mirror deformations to escape the interferometer — but for two important exceptions. The first is high-angle scattered power (created by fine-scale mirror roughness), which escapes during propagation through the Fabry-Perot arm cavities due to the finite size of the mirrors; and the second is power escaping from beamsplitter exit port due to an imperfect dark-fringe, which (for the Initial-LIGO) has nothing to stop it from leaving the system entirely. This "lost" exit-port power may be absorbed by the signal photodetector, thus increasing the interferometer shot noise level which competes with the GW-signals, as well as causing a power drain which reduces the power available in the arm cavities for generating those signals in the first place. The problem of high-angle scattering (which will be discussed in upcoming sections) is largely intractable, without improvements in the polishing and/or coating of the mirrors; but for losses at the beamsplitter exit port, which are caused by any noise or imperfection properties of the interferometer that operate differentially for the inline and offline arms, the presence of a Dual Recycling Mirror is intended to be a key factor in blocking such exit-port power from escaping the interferometer.

In several papers introducing and analyzing Dual Recycling [e.g., 24, 26], Meers, *et al.*, claim that the presence of a Signal Recycling Mirror[1] (with power transmission $T_{SRC} \equiv T_6$ from Fig. 2.1) will reduce the leaking exit-port power by a factor $\sim 1/T_{SRC}$. Note that this noise reduction factor is only correct for a distortion mode (and beam frequency) that is *significantly off resonance* in the SRC. If the distortion mode is resonant (or nearly resonant) in the SRC, then it will be amplified by the SRC resonance (or by the SRC/arm cavity double-resonance), just as the GW-signal is amplified by it, and the implementation of

1. In this thesis we will use the terms "Dual Recycling Mirror" and "Signal Recycling Mirror" interchangeably.

Dual Recycling would in that case cause the exit-port losses to be much more severe.

In particular, consider Broadband Dual Recycling, for which $\nu_{GW}^{optim.} = 0$ Hz. In this case, any power in the TEM_{00} mode that is not nulled by the beamsplitter "dark-fringe", and thus makes it into the SRC, will satisfy *exactly* the same resonance conditions everywhere that the GW-sideband light generated in the arms would satisfy; and the amplification of such losses could devastate TEM_{00} mode buildup in the interferometer. This is a well-known problem for Broadband Dual Recycling, and it is therefore considered to be a system effective at loss suppression for *small amplitude* deformations only, for which the primary effects of mirror deformations are the creation of power in higher-order beam modes (which are non-resonant, and are thus prevented from exiting by the Signal Recycling Mirror). Since the amplitude of TEM_{00} mode light leaking out of the dark port because of imbalances between the inline and offline arms is proportional to the *square* of the mirror deformation amplitude in the one of the arms — instead of linear in deformation amplitude, as is the creation of higher modes — large deformation amplitudes are necessary to cause significant leaks of "uncancelled" TEM_{00} light from the beamsplitter dark port into the latter half of the SRC [26]. Broadband Dual Recycling therefore becomes a very lossy system when the mirror deformation amplitudes are large. Using mirror tilts as an example, it has been found (analytically, experimentally, and numerically [26, 70]) that Broadband Dual Recycling functions well at reducing exit-port power losses for small tilts, but that it fails quickly as the condition $\theta \ll \theta_c$ is violated, where the "characteristic angle", θ_c, is given by $\theta_c \equiv \sqrt{2}\lambda/(\pi \cdot \text{beam spot size}) \propto \lambda/(\text{cavity length})$.

In contrast to this problem for TEM_{00} mode losses, Broadband Dual Recycling will be effective at suppressing losses due to higher spatial beam modes, because these modes are not resonant in the system. This off-resonance is due to the *Guoy Phase* (see Sec. 2.3.4.2) that the higher modes pick up during round-trips through the Fabry-Perot arm cavities, because of the focusing effects of propagations through the long-baseline arms. The off-resonance in the arms means that the effects of the arm cavities are negligible (and even suppressive) for these modes, and that only the SRC matters. But since the SRC is a short cavity (~few meters) with negligible focusing, it is a *degenerate* cavity, so that all modes will experience the same round-trip propagation phases there. As discussed at the end of the previous subsection, the spatial mode making up the GW-signal (that is, the TEM_{00} mode) is anti-resonant in the SRC for Broadband Dual Recycling, and hence *all modes* will be anti-resonant in the degenerate SRC, which means that higher-mode exit port losses will not be resonant in the arm cavities nor in the SRC, are thus they are doubly-suppressed from leaking out in the case of Broadband SRC tuning.

The situation, however, is very different for Tuned Dual Recycling, especially when $\nu_{GW}^{optim.}$ is large, and the SRC achieves near-resonance for all deformation modes[1]. In that

1. Note, however, that since the TEM_{00} mode for the carrier frequency beam remains resonant in the arm cavities regardless of the SRC tuning used, Tuned Dual Recycling does *not* amplify TEM_{00} mode losses, but rather it "flattens out" its resonant arm cavity buildup, lowering the TEM_{00} losses.

case, the implementation of Dual Recycling can greatly amplify non-TEM_{00} mode losses at the exit port. Tuned Dual Recycling using a degenerate SRC should therefore be *more vulnerable* to most optical deformations than the Initial-LIGO configuration (i.e., without Dual Recycling) would be. This problem will be observed in the results shown later in this chapter, and possible solutions will be discussed (and demonstrated) in Sec. 4.3.2.

In spite of these limitations, as long as beam modes created by mirror deformations are not resonant in the SRC/arm cavity system — and as long as the deformation-induced modes aren't completely lost due to scattering from very fine mirror roughness — the implementation of Dual Recycling should be an effective system for reducing exit-port power loss, thus increasing the tolerance of interferometers to deformations such as tilts, curvature errors, and other smooth mirror imperfections.

4.1.3 Reducing signal degradation via "Wavefront Healing"

When Dual Recycling is used to control the exit-port power losses, the reduction of their contribution to interferometer shot noise is only half of the story. The other, and perhaps more important aspect to consider, is the re-introduction of this salvaged power into the normal beam modes of the interferometer. By taking "waste" power and reclaiming it as useful light in the Fabry-Perot arm cavities, it is theoretically possible to boost the GW-signal back nearly to where it would have been for an interferometer with perfect mirrors, possessing no surface or substrate deformations. This process of repairing the effects of mirror deformations upon the GW-signal has been studied in the Dual Recycling literature [e.g., 26], and in this thesis we will refer to the process by one expression that has been commonly used for it: "Wavefront Healing".

Wavefront healing relies upon the ability of the reclaimed exit-port light to re-integrate itself among the stored, resonant light within the interferometer arms. When this reclaimed light consists mostly of TEM_{00} mode power, then it is easy to see that the process should be an effective one (though this light does indeed have to get back into the PRC part of the interferometer from the SRC part of it, which is equivalent to crossing the beamsplitter dark-fringe). But when the reclaimed power consists entirely of higher mode light, then some care must be taken to explain why (and if) this higher mode light should contribute successfully to the GW-signal, rather than simply being irrelevant, or even corrupting the TEM_{00} light that would have remained in the arms without Dual Recycling.

The configuration considered primarily by Meers and Strain [26] was that of a Dual-Recycling interferometer with *single-pass* Michelson arms instead of arm cavities; that is, the input mirrors to the arms — M_2 and M_3 of Figure 2.1 — would be nonexistent (i.e., unity transmission), so that the arms would be in direct contact with the beamsplitter, and would not be separated off into isolated cavities. Furthermore, their signal detection scheme did not rely (as does the Initial-LIGO configuration) on implementing a Schnupp Length Asymmetry and having radio-frequency sideband beams resonant in the PRC; rather, they assumed the use of an "external modulation" scheme [e.g., 71] for GW-signal

detection, for which the local oscillator light needed for heterodyne detection was drawn directly from the e-fields resonating in the Michelson arms[1].

For such a system, Meers and Strain argued that the reclaimed exit-port power *does not have to* re-enter the original TEM_{00} mode to contribute to the creation of GW-signals; rather, this higher-mode power would participate in the formation of a "new fundamental mode" (TEM_{00}') of the Dual-Recycled interferometer, that resonates effectively in the system. And since the local oscillator light is drawn directly from the Michelson arm e-fields that oscillate in this new fundamental mode, the overlap between the GW-induced sidebands and the local oscillator light (both generated from the TEM_{00}' mode carrier light) would be as complete as if the modal structure of the interferometer had never been perturbed by optical deformations. The process of wavefront healing is therefore expected to be almost completely effective for this type of interferometer configuration.

The answer is less clear, however, for systems that *do* possess Fabry-Perot cavities in the arms, such as Initial- or Advanced-LIGO interferometers. In such cases, the e-fields in the arm cavities will almost entirely consist of the TEM_{00} mode (due to resonant suppression of higher modes), and will be very different from the e-fields in the PRC and SRC, which they are separated from by the arm cavity input mirrors. In any case, the local oscillator light may not even be drawn from the carrier light resonating in the Michelson arms, since some advanced version of the Schnupp Asymmetry Scheme with RF-sidebands may be used, so that the "new fundamental mode" of the interferometer may be completely irrelevant. If, as we will assume in this thesis, all of the GW-signal comes from the TEM_{00} components of the GW-sideband and local oscillator e-fields (because of the TEM_{00} dominance in the arm cavities), the question to be answered is this: can re-injected exit-port power consisting of higher modes be converted back into TEM_{00} power in the arm cavities? Only then could we say that wavefront healing is effective for a Dual-Recycled LIGO interferometer.

The empirical answer to this question, as determined by the runs of the simulation program which we will discuss in Section 4.2.3, is *yes*. Non-TEM_{00} mode power that is injected into the arm cavities is very efficiently converted back into resonant TEM_{00} mode power. Despite the fact that the coefficient (nonzero due to mirror deformations) which couples power from a given non-TEM_{00} mode (call it TEM_{XY}) into TEM_{00} must be the same as the one that couples TEM_{00} power into TEM_{XY}, the resonance condition of the arm cavities forces almost all power injected into them to re-enter and remain within the original, unperturbed TEM_{00} mode.

This argument can be quantified as follows. Consider a simple cavity with input mirror (amplitude) parameters t_1 and r_1, and a back mirror amplitude reflectivity of r_2. If the mode-mixing coefficient is given by ε, and the one-way propagation (Guoy) phase for the

1. Specifically, it is drawn from the small amount of power that is reflected from the anti-reflective coating of the beamsplitter, which is encountered during the process of A.R.-side reflection from the optic (i.e., operator R_{BS}^{a} from Figure 2.9). This extracted power is then RF-modulated *externally* to the interferometer, which makes this detection scheme a heterodyne algorithm.

TEM_{XY} mode in the arm cavity is given by ϕ (with the resonant TEM_{00} mode having "zero" phase), then the round-trip matrix operator for the cavity is written as:

$$\begin{bmatrix} TEM_{00} \rightarrow TEM_{00} & TEM_{XY} \rightarrow TEM_{00} \\ TEM_{00} \rightarrow TEM_{XY} & TEM_{XY} \rightarrow TEM_{XY} \end{bmatrix}_{\text{Round-Trip}} = r_1 r_2 \begin{bmatrix} 1-\frac{\varepsilon^2}{2} & \varepsilon e^{i\phi} \\ -\varepsilon e^{i\phi} & \left(1-\frac{\varepsilon^2}{2}\right)e^{2i\phi} \end{bmatrix} \quad (4.1)$$

With this cavity round-trip operator, we can use Eq. 2.54 to solve for the total operator representing the action of the cavity on an excitation e-field, thus getting the steady-state field that builds up inside the cavity, given an arbitrary mixture of TEM_{00} and TEM_{XY} in the excitation beam. The resulting cavity operator is equal to:

$$\begin{bmatrix} TEM_{00} \rightarrow TEM_{00} & TEM_{XY} \rightarrow TEM_{00} \\ TEM_{00} \rightarrow TEM_{XY} & TEM_{XY} \rightarrow TEM_{XY} \end{bmatrix}_{\text{Cavity}} \quad (4.2)$$

$$= \frac{t_1}{Resfactor} \begin{bmatrix} 1-r_1 r_2\left(1-\frac{\varepsilon^2}{2}\right)e^{2i\phi} & \varepsilon r_1 r_2 e^{i\phi} \\ -\varepsilon r_1 r_2 e^{i\phi} & 1-r_1 r_2\left(1-\frac{\varepsilon^2}{2}\right) \end{bmatrix}$$

where:

$$Resfactor = \left[1-r_1 r_2\left(1-\frac{\varepsilon^2}{2}\right)\right] \times \left[1-r_1 r_2\left(1-\frac{\varepsilon^2}{2}\right)e^{2i\phi}\right] + \left(\varepsilon r_1 r_2 e^{i\phi}\right)^2 \quad (4.3)$$

From the above formulas, the power inside the cavity for either the TEM_{00} or TEM_{XY} mode can be computed (by taking the squared magnitude of the appropriate matrix element), and can be plotted versus ε (or ϕ) in order to quantify the effects of either TEM_{00} power degradation due to mirror deformations of amplitude ε, or the converse effect of wavefront healing via conversion of TEM_{XY} power to TEM_{00} power within the cavity.

Though such plots are not shown here, the pertinent results can be quickly summarized, for the choice of a "far off-resonant" TEM_{XY} mode (i.e., $0 \ll \phi \ll \pi$), and using LIGO-like parameters of $r_1 = \sqrt{.97}$, $t_2 = \sqrt{.02995}$, and $r_2 = \sqrt{.99995}$. For a substantial (~50%) reduction of TEM_{00} power in the cavity *without* recycling of the "lost" (TEM_{XY}) power, it is required that $\varepsilon \sim .05$ - $.07$. For a deformation amplitude this large (and given a far off-resonant TEM_{XY} mode), it turns out that injecting 1 Watt of pure-TEM_{XY} light into the cavity causes ~4 Watts of TEM_{00} power to resonate inside of it; this is ~10 times more cavity power than is generated in the TEM_{XY} mode by this pure-TEM_{XY} beam, and this

"healing efficiency factor" can go as high as ~30 for larger deformation amplitudes. Finally, considering the case of a pure-TEM_{00} excitation beam, we can compare what happens if power initially converted from $TEM_{00} \rightarrow TEM_{XY}$ is retained in the cavity so that it can eventually return via $TEM_{XY} \rightarrow TEM_{00}$ during a wavefront healing process, versus what would happen if any created TEM_{XY} power is immediately lost from the system. The result is that the amount of TEM_{00} power resonating within the cavity is fully %15 higher if the TEM_{XY} mode is retained in the system, rather than being thrown away, which illustrates the wavefront healing capabilities of a single cavity.

For a full Dual-Recycled LIGO interferometer, the impact of wavefront healing should be even more significant than the results quoted above, for several reasons. First, the SRC/arm cavity system results in a *double-suppression* of TEM_{XY} light escaping through the exit port, a situation that was not treated in the above example since the TEM_{XY} power that is reflected from the single cavity does not get recycled back to it indefinitely, as an off-resonant mode in a Signal Recycling Cavity would be indefinitely recycled. Second, the TEM_{00} double-resonance supported by the SRC/arm cavity system (it will be a double resonance as long as $\nu_{GW}^{optim.}$ is not tuned too high) enhances the conversion of $TEM_{XY} \rightarrow TEM_{00}$ more effectively than the resonance of the single cavity does.

A third reason, is that for a "fully sealed" Dual Recycling interferometer like the configuration of the proposed Advanced-LIGO system, there is nowhere for leaking TEM_{XY} power to go — assuming that it is not immediately lost to high-angle scattering during its initial generation by interaction of the TEM_{00} beam with mirror deformations — in order for it to effectively dissipate itself into the environment. This is in contrast to TEM_{00} light, which has a very effective means of dissipation: it builds up resonantly in the system, and causes a great deal of power to be absorbed (or scattered) by repeated interactions with the mirrors via their small, ~50 parts per million of losses for each bounce. But TEM_{XY} mode power, which is non-resonant everywhere, does not get dissipated effectively by interactions with mirrors. In short, if power is converted into the TEM_{00} mode, it becomes resonant and dissipates very quickly; but if TEM_{XY} power is created, then it stays in the system for a very long time, giving it many opportunities to convert back into the TEM_{00} mode — after which it quickly vanishes.

Fourth, and last, is the fact that when power is reclaimed and sent back into the arm cavities, any of it which gets converted back into TEM_{00} power can combine coherently (cavity length adjustments will match the phases) with the *ambient* TEM_{00} light already in the arm. Thus if a non-healed arm cavity has amplitude A of TEM_{00} light already resonating within it, and amplitude δ of both TEM_{00} light and TEM_{XY} light added to it because of the Signal Recycling Mirror, then the TEM_{00} power will increase by the amount $[(A+\delta)^2 - A^2] \cong 2\delta A$, while the TEM_{XY} power increases only by δ^2 — a very significant different difference when computing the TEM_{00}-based GW-signal benefits of Dual Recycling.

Indeed, in the runs to be presented in Section 4.2.3, it is generally found that less than one percent of the increased power in the arm cavities due to signal recycling remains in non-TEM_{00} modes; almost all of the extra power, due to the wavefront healing effect for resonant arm cavities, is converted back into the TEM_{00} mode.

This process of wavefront healing for arm cavity systems may seem somewhat non-intuitive, since if "entropy in increased" by scattering TEM_{00} power into a great many non-TEM_{00} modes, then doesn't entropy get *reduced* during healing, since power in all of these modes are converted uniquely back into the TEM_{00} mode? The answer is that it does, and no physical paradox is created by this effect, because it is always possible to reduce entropy within a closed system by *consuming work*. Achieving the wavefront healing effect is completely dependent upon the ability to lock the interferometer cavities to the correct resonance conditions. This requires the expenditure of work, the application of external forces upon the mirrors, and even the (entropy-generating) absorption of some of the interferometer laser power by the length control systems in order to measure the mirror positions and prescribe the correct adjustments. In the end, wavefront healing can be a successful method for re-generating power in the arm cavities that is useful for GW-signal creation, because the extremely complex array of interferometer control systems that operate within a LIGO interferometer ensure that it must be successful.

4.2 Demonstrations of Dual Recycling

4.2.1 Results from the literature: experiment and simulations

The predicted benefits of Dual Recycling have already been demonstrated, on a limited basis, via experimental and numerical tests of simplified interferometers. Here we give a quick summary of the principal verifications that are relevant to this thesis, not only in order to demonstrate the promise which exists with a properly functioning Dual-Recycled interferometer, but also to point out the complexities of a real-LIGO Dual Recycling system which must not be ignored when predicting the ultimate quality of its performance.

A table-top Dual Recycling experiment was performed by Meers and Strain [26, 72] in order to test the claims made by Meers (and discussed above in Sec. 4.1) about the robustness of Dual Recycling for interferometers possessing geometric deformations. Their setup was a Dual-Recycled interferometer using optics with 10% (power) transmission for the Power Recycling and Dual Recycling mirrors, and with single-pass Michelson arms (no Fabry-Perot arm cavities). The lengths of the arms were ~half of a meter, and they used an Argon-ion laser ($\lambda = 5.14$ nm) for their "carrier" beam.

The optimistic predictions were borne out by their results. They found that Dual Recycling reduced the fractional power leaking out of the interferometer exit port from 4×10^{-4} to 2.5×10^{-5}, leading to a substantial (~20%) decrease in the overall shot noise level. They also found that the presence of the Dual Recycling mirror increased the power stored in

the Michelson arms, *as well as* the useful GW-signal that could be extracted from it, the latter being a result which indicates the action of "Wavefront Healing"[1]. In addition, these beneficial effects of Dual Recycling went away for large-amplitude tilt deformations, for which the presence of deformation-induced TEM_{00} mode in the Signal Recycling Cavity began (as predicted) to make losses worse for their Broadband Dual Recycling scheme, than they were in the absence of a Signal Recycling Mirror. Finally, their analytical predictions as a function of mirror tilt matched both the measured exit port losses (with the first few lowest-order modes dominating), as well as agreeing with the measured tuning curve about the signal response peak of the Broadband system.

Despite these successes, there are a number of simplifications that had to be made for their experiment to be possible, such as:

- They tested a Broadband system only. Because of the shortness of their Michelson arms (and the lack of Fabry-Perot arm cavities), construction of a Tuned (Narrow-banded) Dual Recycling interferometer was not possible.
- Their Signal Recycling Cavity was completely nondegenerate. Again, this is unlike the case with Fabry-Perot arms, in which the SRC is a short (and degenerate) system because it is separate from the nondegenerate (and for LIGO, long-baseline) arms by the arm cavity input mirrors.
- Scattering was a less important source of loss for their system than for a LIGO interferometer, because their arms were short compared to the long-baseline LIGO arms.
- Their beams (with waist size ~.2 mm) were likely very small compared to the sizes of their mirrors, thus rendering negligible the effects of edge-diffraction losses and power falling off of the mirrors due to moderate-angle scattering from mirror deformations.
- They only examined geometrical deformations (and in particular, only tilts); there was no real study of the effects of significant "realistic" (see next point) deformations upon interferometer power buildup.
- Their small spot-size beams would not have sampled the range (or amplitude, as per the discussion at the beginning of Sec. 2.6) of mirror deformations that will be important for real-LIGO interferometers.

Each of the properties mentioned above for a "real-LIGO" Dual-Recycled interferometer, that could not be tested in their experimental apparatus, can be modelled by our simulation program. And, as will be seen in the following sections, every one of these properties is

1. In particular, the kind of wavefront healing which occurs in the absence of Fabry-Perot arm cavities or a Schnupp Length Asymmetry scheme, as was discussed in Sec. 4.1.3.

responsible for creating dramatic and negative (though sometimes fixable) problems in a real Dual Recycling interferometer.

At this point we note that our work is not the first simulation-based research done for Dual-Recycled interferometers. Investigations have been performed by McClelland and colleagues [29, 49, 70, 73, 74] for the performance of Dual Recycling for full-scale interferometers (i.e., ~3-4 km long Michelson arms). This simulation work was able to generalize the verification of Dual Recycling beyond the table-top experimental results, to the extent of addressing some of the concerns listed above; this has included simulations of Broadband and Tuned Dual Recycling systems, for realistic beam and mirror sizes. But by and large, only systems with single pass Michelson arms were included, instead of Fabry-Perot arm cavities[1], and thus the behaviors (such as the possible degeneracy, and the SRC/arm cavity double-resonance) of an SRC separated from interferometer arms by mirrors were not considered. More importantly, even though the list of deformations that were studied was expanded to mirror curvature errors as well as tilts, these are still only "geometric" deformations, which do not capture all of the effects of realistically deformed optics upon the interferometer beams. Alleviating this lack is the primary function of our Advanced-LIGO simulation program, with which, due to computational speedup and efficient program algorithms (as discussed in Sections 2.2, 2.4, and 2.5), we are able to model the effects of realistically deformed optics — i.e., "rough" optics, possessing realistic surface and substrate power spectra — for a full-LIGO Dual-Recycled interferometer. And the program also possesses the ability to produce a reasonably thorough sampling in GW-frequency space, for the construction of simulated interferometer response curves.

We will not discuss in further detail the results of Dual Recycling numerical simulation work done by previous researchers, only to say that the results were very good and in accordance with the predictions of the robust behavior of Dual Recycling in the presence of geometric mirror deformations. There is one point of note, however, mentioned in [49, 70], which states that Dual Recycling initially made exit port losses *worse* for curvature mismatch errors, because of the accidental resonance of the TEM_{20} and TEM_{02} modes in the SRC. As has been mentioned earlier in this thesis at the beginning of Chapter 3, this problem was removed by changing the mirror curvatures — and hence, the round-trip Guoy phases for modes in the SRC — so that these curvature-error-induced modes became non-resonant. But in the case of realistically deformed mirrors (which excite *every* beam mode to some degree), and for a Dual Recycling interferometer with Fabry-Perot arms (which thus has a short, degenerate SRC that can be resonant for every non-TEM_{00} mode simultaneously), it is simply not feasible to adjust the mirror curvatures to make every single one of an infinite number of modes non-resonant. Demonstrating the effects of this problem, as well the demonstration of an achievable solution (SRC degeneracy-breaking, along with finite-sized mirror apertures) will be discussed in Section 4.3.

1. The only published simulation results with Fabry-Perot arms in a Dual Recycling interferometer were shown in [74], based upon the (unpublished) thesis work of B. D. Stanley.

First, however, the remainder of Section 4.2 will be devoted to demonstrating the (moderately) successful and well-predicted performance of Dual Recycling in the presence of realistically imperfect optics, for both the Broadband case and the Narrowbanded case with low tuning frequencies.

4.2.2 A quick look at Broadband Dual Recycling

In our studies of Dual Recycling with realistically imperfect mirrors, we consider interferometers with essentially the same configuration as the Initial-LIGO case, as defined in Table 2.1 (with relevant modifications listed for the Dual Recycling configuration). We also use the same surface and substrate deformation maps that have been discussed in Section 2.3.3.4, and used for the Initial-LIGO characterization runs in Chapter 3. To this configuration we add a Signal Recycling Mirror with its own surface and substrate maps (from the same families as the maps for the other mirrors), and use this mirror's reflectivity, as well as the SRC tuning frequency for GW-signal optimization ($\nu_{GW}^{optim.}$), as parameters to be varied from run to run. In addition, note that the runs in the remainder of Section 4.2 will use a "short" length, 5 m, for the distance between the beamsplitter and the Dual Recycling Mirror, thus implying a degenerate SRC.

As discussed above in Section 4.1.1.2, Broadband Dual Recycling will likely be of less interest than Tuned Dual Recycling for Advanced-LIGO systems (and more generally, for ground-based interferometers with mirrors possessing realistic Q-factors), because the large levels of seismic and thermal noise at low GW-frequencies reduces the usefulness of improvements to the shot noise sensitivity limit for Broadband or near-Broadband SRC tunings. Our primary results in this thesis will therefore be for Narrowband cases. Nevertheless, a small set of Broadband runs happens to be useful as a clear-cut verification of the predicted behavior of interferometer e-fields in the presence of optical imperfections, in a special case where the results for realistic mirror deformations behave in a way that largely obeys the intuition which has been developed from the study of geometric deformations.

In Figure 4.3, the amount of power lost through the interferometer exit port is shown (normalized to 1 Watt of carrier power for the interferometer excitation laser), for runs possessing deformed substrates and $\lambda/800$ surfaces for all of the mirrors. The results are plotted versus the (power) reflectivity of the Signal Recycling Mirror, which is varied from zero (i.e., complete absence of the mirror and all of its effects) to $R_{dual} = 0.9$.

Both the total exit power and the power in the TEM_{00} mode alone are shown, and the results bear out the standard predictions for Broadband Dual Recycling. As R_{dual} is increased, the exit-port power lost in the TEM_{00} mode is made larger, because it is doubly-resonant in the SRC/arm cavity system; but the total, summed power lost in *all modes* is made smaller for increasing R_{dual}, because all of the non-TEM_{00} modes are non-resonant to various degrees in the Fabry-Perot arm cavities, and thus they are all anti-resonant in the

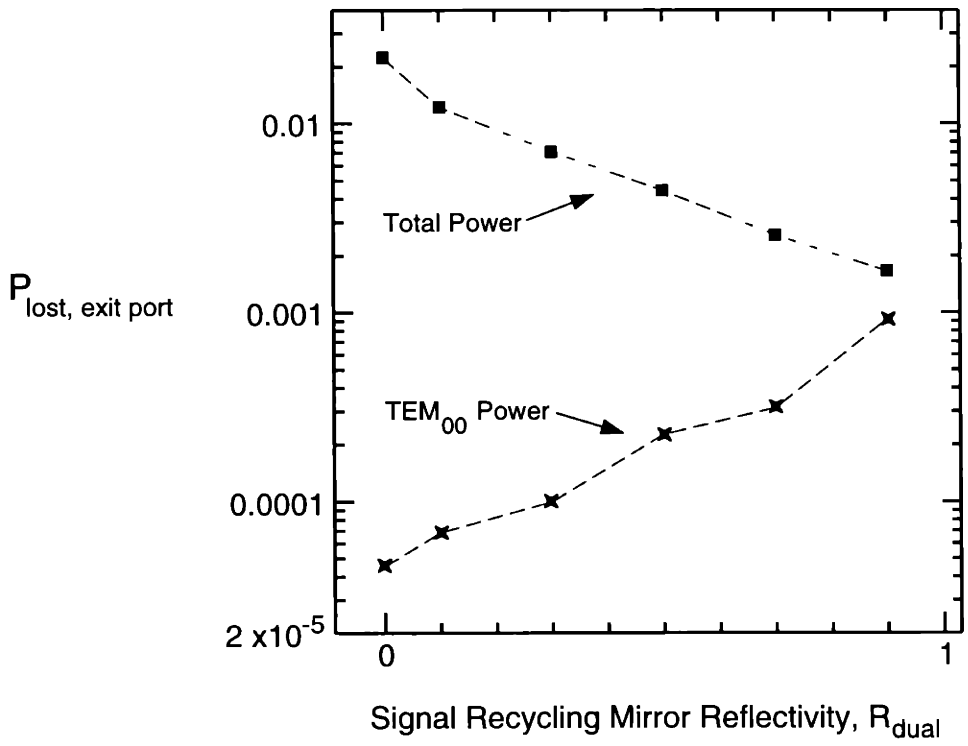

Figure 4.3: Fraction of interferometer excitation laser power lost through the exit port, plotted versus the power reflectivity of the Signal Recycling Mirror. These results are for Broadband Dual Recycling, i.e., optimized for a GW-frequency of zero.

degenerate SRC. The net effect for these runs — which have a mirror deformation level that is effectively quite large (~45% reduction in stored arm cavity power!), and which sends power into the SRC in all different modes, TEM_{00} included — is that there is an overall *reduction* in loss as R_{dual} is increased, because the losses due to non-TEM_{00} modes (which dominate the total loss for low R_{dual} values) are reduced. But the TEM_{00} mode already comprises more than half of the lost power at $R_{dual} = 0.9$, and it will dominate at even higher R_{dual}, so that above a certain threshold the power loss will begin *increasing* with increasing R_{dual}. This harmful effect of Broadband Dual Recycling, discussed in Sec. 4.1.2 and well known in the literature [26, 49, 70], will begin to degrade interferometer performance at even lower values of R_{dual}, when the system has optical deformations that are more effective than our mirror surface and substrate deformations (which mostly scatter power into very high modes) are at sending TEM_{00} power into the SRC to be amplified and lost. Thus we see that Broadband Dual Recycling cannot be used with large values of R_{dual} for significant levels of optical deformations in the interferometer. As stated in Sec. 4.1.1.2, however, this is not generally a strong limitation on the usefulness of Broadband Dual Recycling, since it is limited anyway to small values of R_{dual} because of the domination of random-force noise sources at low GW-frequencies.

As a last comment, we note that it is actually good to have a *degenerate* SRC for Broadband Dual Recycling, since that allows all non-TEM_{00} modes to be simultaneously anti-resonant there. A non-degenerate SRC would shift the modes from anti-resonance to various states of non-resonance, and perhaps a few of the modes would become (nearly-) resonant, thus greatly amplifying the losses in those modes.

4.2.3 A Narrowband study with GW-frequency tuned to 200 Hz

In this section, we will consider results from simulation runs with a "Narrowband" Dual Recycling interferometer, in particular one that is tuned to a GW-frequency of 200 Hz. Tuning $\nu_{GW}^{optim.}$ to ~few hundred Hz (or ~1-2 kHz) is useful not only because it moves the region of maximum detector sensitivity away from low-frequency noise sources, but also because it facilitates the detection of GW-sources which radiate within that general frequency band.

For coalescing Black Hole/Black Hole (BH/BH) binaries, it is possible that the poorly-understood "merger" stage of the coalescence, during which the two bodies plummet in free-fall until they collide and begin relaxing into a single object, may be the most significant epoch of GW-radiation for the BH/BH system. The merger phase (shown as wavy lines in the characteristic GW-signal strength curves from Figure 3.6) occurs at the end of the "inspiral" phase, and before the final "ringdown" phase. Flanagan and Hughes have estimated [61] that the merger phase begins when the GW-emission frequency is given by:

$$f_{\text{Merger}} \approx 205 \text{ Hz} \left(\frac{20 M_{\text{sol}}}{M_{\text{tot}}} \right) = 205 \text{ Hz} \tag{4.4}$$

where the latter equality was obtained assuming (for good overlap of GW-emission with the LIGO sensitivity band) a system of two identical, ten-solar-mass black holes. They also estimate the upper-bound ("shut-off") frequency for merger GW-emissions to be at:

$$f_{\text{high}} \approx 1430 \text{ Hz} \left(\frac{20 M_{\text{sol}}}{M_{\text{tot}}} \right) = 1430 \text{ Hz} \tag{4.5}$$

Equations 4.4 and 4.5 imply an approximate frequency range of interest for Narrowband Dual Recycling, and similar ranges can be inferred from the physics of other promising sources. The GW-emission from non-axisymmetric pulsars (which has a frequency of twice the pulsar rotation frequency) can be estimated [7] to lay within a typical range of ~20 – 1000 Hz, and radiation during "centrifugal hangup" or "neutron star boiling" behavior in the cores of supernovae have also been estimated [9] to lay within a similar frequency range.

Given these bounds, we have performed a comprehensive study of a Dual Recycling interferometer that is tuned to $\nu_{GW}^{optim.} = 200$ Hz, a value just high enough to sample important physics (e.g., at the boundary between the inspiral and merger phases of BH/BH

binaries), while still being low enough to have to consider the importance of seismic and thermal noise. This study will be followed up in Sec. 4.3.2 by a smaller selection of runs both with this tuning frequency and with $\nu_{GW}^{optim.} = 1000$ Hz, in order to study some of the problems with Narrowband Dual Recycling at moderate and high tuning frequencies, and in order to discuss (as well as demonstrate) some possible solutions.

The runs to be discussed here have been done with deformed substrates on all of the mirrors (including the Signal Recycling Mirror), and with deformed surface maps from the λ/1800, λ/1200, λ/800, or λ/400 families; in addition to these runs, simulations were also done with "perfect" mirror surfaces and substrates. For each run with the SRC tuned to an optimization frequency of 200 Hz, the ("plus" and "minus") GW-sidebands at many frequencies were simulated in order to produce discrete (but well-represented) frequency-response curves. The (unnormalized) GW-signal at each frequency, f_{GW}, is taken to be the sum of the absolute values of the interferometer signal-port e-fields at $\pm f_{GW}$, as in Section II.A. of [25]. The principle parameters which we vary in order to generate distinct runs includes the family of surface deformations that is used for a given run, as well as the power reflection coefficient of the Signal Recycling Mirror, R_{dual}.

In Figure 4.4, several plots are shown (each for a different value of R_{dual}) which depict the frequency response curves obtained for runs with different mirror maps. From the top, the dotted lines (each dot representing a frequency for which the plus and minus GW-sidebands were simulated) represent, respectively, the "perfect mirrors" run, the λ/1800 run, and the λ/800 run. Each plot is shown along with a solid line representing the analytically-calculated response curve of the Dual-Recycled interferometer, a function which has been discussed (with examples shown) in Sec. 4.1.1.2. These curves represent theoretical predictions of the GW-response using all of the physical parameters of the (perfect) interferometer, as per Table 2.1, *in addition* to the implementation of 2 parts per million ("ppm") of extra loss for each arm cavity back mirror (where the beam spot size is the largest, and the circulating power is greatest), to account for diffraction losses which predominantly occur at these finite-sized arm cavity back mirrors[1]. We can draw several important conclusions from these plots. First of all, we note that the "perfect mirrors" runs match the theoretical predictions very well, thus providing some evidence that the Dual-Recycled system is being modelled correctly by the simulation code. Second, the use of maps of realistically deformed mirrors does not reduce the overall GW-sensitivity by an overwhelming factor. Lastly, we note that the *shape* of the frequency response curves do not

1. It is easy to show [e.g., 17] that the power cut immediately out of the beam by the apertures of the arm cavity back mirrors, given the interferometer parameters which we have used, is ~1.0 ppm per reflection. This value has been doubled to 2.0 ppm in our calculations because of diffractive effects; specifically, if the beam encountering the back mirror is in the fundamental mode (or any other given mode), then a simple overlap calculation shows that the amount of power taken out of that mode by a hard-edged apertured is exactly *twice* the amount of power that is immediately lost at the aperture. This "extra loss" is actually power which has been shifted to higher modes, much of which will get cut out by the apertures during the next reflection at the arm cavity back mirror.

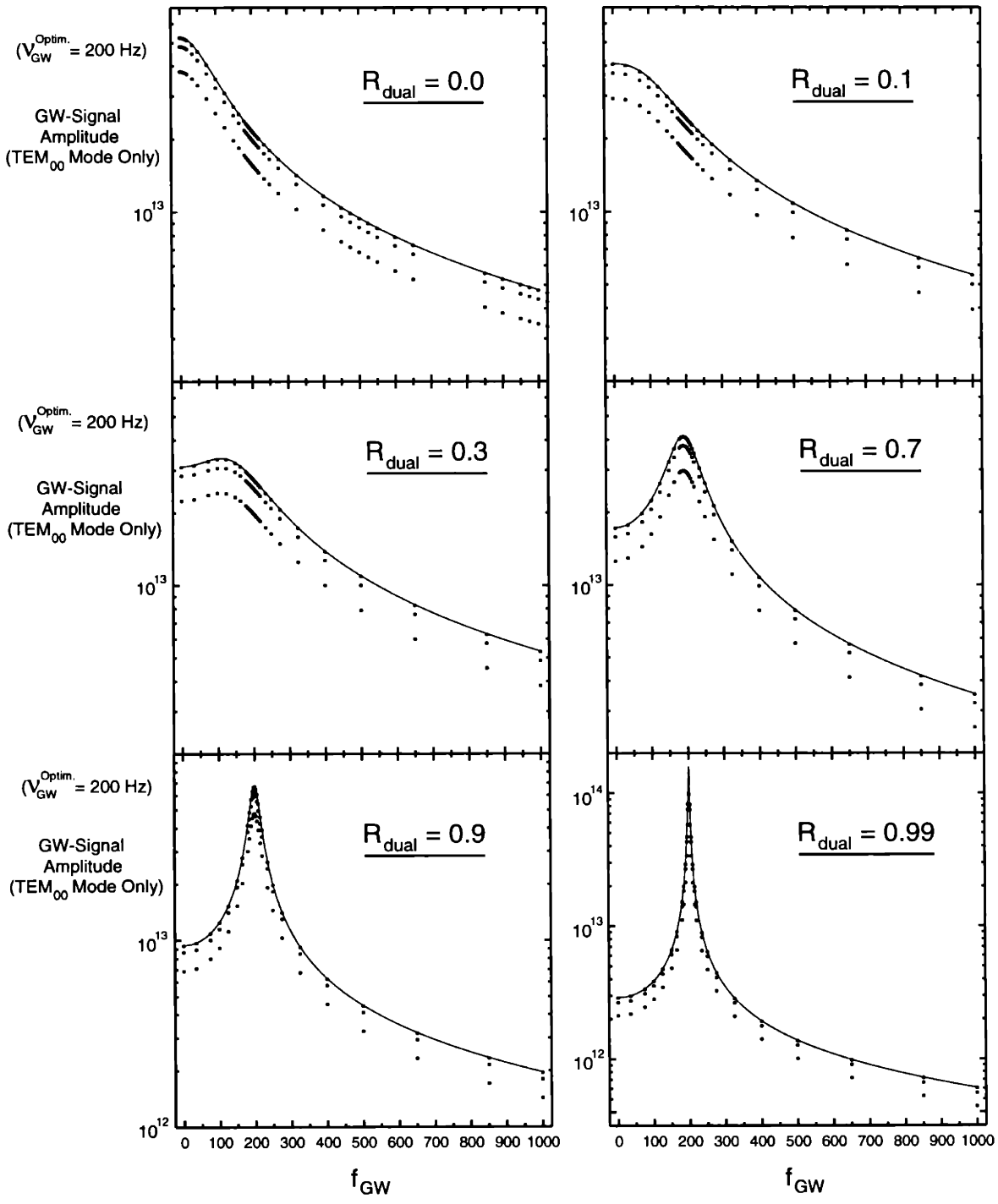

Figure 4.4: The GW-signal amplitude is plotted versus GW-frequency for several values of the Signal Recycling Mirror reflectivity, for a tuning optimization frequency of 200 Hz. Each plot compares the theoretically-calculated response curves (solid lines) with results from runs of our simulation program (dotted curves) using, in order of decreasing overall magnitude: (i) perfect mirrors, (ii) lambda/1800 mirrors, and, (iii) lambda/800 mirrors.

change due to the presence of deformed mirrors. In particular, they do not spread the sensitivity peaks out significantly (which would destroy the narrowbanding effect which is sought by this configuration), but rather appear to just reduce the response curves by an overall, broadband factor. This effect is due to fact that the predominant losses are due to scattering of carrier power (and to some extent, GW-sideband power) out of the system in the Fabry-Perot arm cavities, which essentially reduces the "reservoir" of power stored in the arms for GW-detection, thus reducing the overall magnitude of the GW-generated fields that could be extracted for signal detection.

Since a highly reflective Signal Recycling Mirror implies that the GW-sidebands must be stored in the SRC/arm cavity system for a long time in order to produce a deep, narrowbanded resonance, it is clear that a large (i.e., near-unity) value of R_{dual} will cause the optical deformations in the mirrors to be sampled more extensively, thus leading to higher losses in the system[1]. As with the optimization of the reflectivity of the Power Recycling Mirror (Sec. 2.5.4), the highest gain that can be achieved with a resonant cavity (or double, coupled-cavity) system is limited by the cavity losses. A question which must be addressed, therefore, is how high a value of R_{dual} can be usefully applied for an interferometer with deformed mirrors, and whether or not mirrors with very high (but "realistic") levels of deformation would make the successful use of Tuned Dual Recycling impossible.

This question is addressed, at least in preliminary fashion, in Figure 4.5. In this figure, each line connecting the data points shows (for a given value of R_{dual}) the fractional reduction of the (TEM_{00}-based) GW-signal at ν = 200 Hz, as the mirrors are changed from the perfect case to the $\lambda/400$ case. What can be concluded is that making R_{dual} as high as 0.9 causes little harm to the GW-signal amplitude, compared to not having Dual Recycling at all; but going as high as R_{dual} = 0.99 causes nearly one-third of the signal that had remained for the $\lambda/400$ case to be lost. In other words, an R_{dual} of 0.0 with $\lambda/400$ mirrors retains ~47% of the GW-signal that it had for the perfect mirrors case, while an R_{dual} of 0.9 with $\lambda/400$ mirrors retains ~44% of the GW-signal that it had for perfect mirrors, but an R_{dual} of 0.99 with $\lambda/400$ mirrors retains only ~30% of the GW-signal that it had for perfect mirrors. Clearly, driving R_{dual} up to values in the vicinity of 0.99 causes serious deformation-induced losses to occur for the GW-signal resonating in the arm cavities and in the SRC.

The significance of these results lay in the detector bandwidths that are represented by these choices of Signal Recycling Mirror reflectivity. At $\nu_{GW}^{optim.}$ = 200 Hz, using R_{dual} = 0.9 produces a GW-signal peak with a full-width at half-maximum (FWHM)

1. This effect of increased losses for a higher R_{dual} is not related to the "Wavefront Healing" effect that is discussed later in this section, and in fact it acts in an opposing fashion. Wavefront healing enhances the GW-signal by catching carrier light before it escapes through the exit port, and recycles it back into the system to increase the power available for interaction with GW's. But for very high values of R_{dual}, the GW-induced signal sidebands *themselves* cannot exit the system for detection without first experiencing large losses (due to high-angle scattering and mirror absorption), because of their long storage in the double-cavity system of the SRC and Fabry-Perot arm cavities.

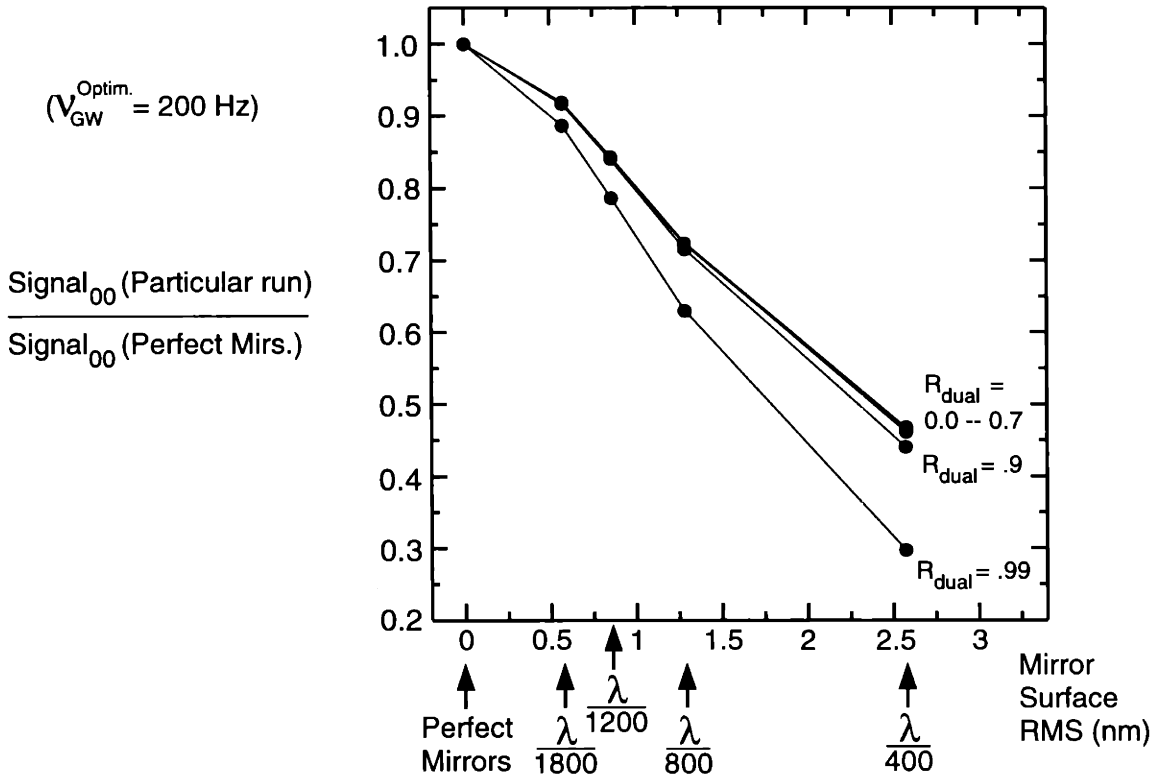

Figure 4.5: Fundamental-mode GW-signal ratios comparing runs with particular mirror deformation maps versus their "perfect mirrors" cases. Each solid line connects the points that share a given value of R_{dual}. The values used here are 0, .1, .3, .5, .7, .9, and .99.

of ~50 Hz (still not extremely "narrowband"), while using $R_{dual} = 0.99$ produces a FWHM of ~8 Hz. In a paper by Krolak, *et al.* [68], which analyzes the optimization of the interferometer sensitivity curve for the detection of coalescing binaries, the FWHM values which they discuss are in the range ~5–8 Hz, with the optimal value being 6 Hz (and they state the optimum bandwidth for periodic sources to be merely half of this, at 3 Hz). Though these optimum bandwidths were quoted for a setting of $\nu_{GW}^{optim.} \approx 110$ Hz, for which it is a little easier (i.e., lower R_{dual} values can be used) to obtain narrow sensitivity peaks, the main result from Fig. 4.5 still stands: a very deep narrowbanding cannot be done without additional loss of signal in cases with highly deformed mirrors, since the run with $\lambda/400$ mirror surfaces and $R_{dual} = 0.99$ has already lost ~70% of its amplitude in the peak of the response, and much of that loss came in the last step of narrowbanding from a FWHM of ~50 Hz to ~8 Hz. And for tuning $\nu_{GW}^{optim.}$ to values of up around ~1–2 kHz, where values of R_{dual} which are much closer to unity are necessary to truly narrowband the signal, this problem will likely be even worse. It therefore appears that mirror quality which is significantly better than the $\lambda/400$ mirror family would be recommended in order to achieve a truly narrowbanded GW-response function for Tuned Dual Recycling.

Moving beyond the problem of (non-exit-port) losses causes by ultra-high-finesse Signal Recycling Cavities, we can consider how the presence of moderate to high values of R_{dual} can successfully prevent unwanted (i.e., non-signal-bearing) power from escaping through the interferometer exit port. The results are shown in Figure 4.6, in which the exit port lost power (in all modes) is shown as a function of mirror deformation amplitude, for several curves with distinct values of R_{dual}. With the exception[1] of a couple of small-but-nonzero values of R_{dual} (0.1 and 0.3), it is clear that increasing R_{dual} *really does* decrease the exit port losses — thus implying that this recycled power is sent back into the system, possibly to undergo wavefront healing and thus contribute to the GW-signal — and for values of higher than 0.9 or so, increasing the amplitude of mirror deformations has almost no effect at all in terms of increasing the exit-port power losses. Thus this particular prediction of Dual Recycling is verified for this SRC-tuning case, and higher R_{dual} values can indeed prevent most of the SRC power from escaping to the exit (i.e., signal) port photodetectors, thus reducing the shot noise level of the interferometer output signal.

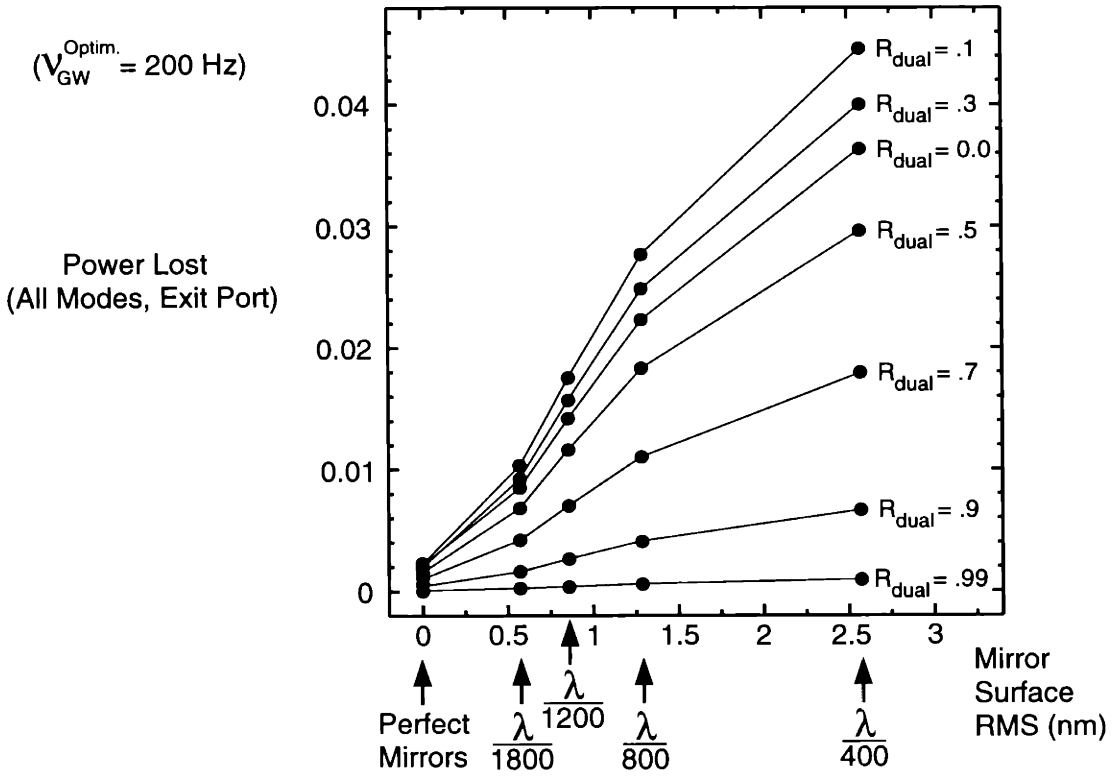

Figure 4.6: Interferometer exit-port power losses as a function of rms mirror (surface) deformation amplitude, for different values of the Signal Recycling Mirror reflectivity.

1. The fact that a "little" Dual Recycling (i.e., R_{dual} = 0.1 or 0.3) is worse than no Dual Recycling at all is due to SRC-degeneracy effects, which will be discussed more fully in Section 4.3.2.

Though this prediction of Dual Recycling is correct, the relevance of it may be doubted, since the exit-port power losses are never very high (always less than ~5% of the excitation laser power, as shown in Fig. 4.6). Thus any local oscillator power at the signal port (which we do not directly model) will contribute shot noise that will swamp that contributed by the carrier e-fields, *in addition* to the fact that any GW-signal gained due to wavefront healing will be extremely minor. This is due to the fact that the principle source of loss for "realistic", high-spatial-frequency mirror deformations is due to *scattering losses* in the Fabry-Perot arm cavities, rather than due to exit port losses (which are really more important for "geometric" mirror deformations such as tilts or curvature errors). In fact, for runs with $\lambda/800$ mirrors we estimate that *nearly half* of all of the power in the original excitation laser beam is lost due to scattering from these simulated mirror deformations (with the other half of the power being lost "normally", from the 50 ppm of loss per interaction of the highly resonant cavity fields with the mirrors). These high scattering losses thus reduce the amplitudes of the GW-induced fields generated from the carrier beam in the arm cavities by (in this case) $\sim\sqrt{2}$; and virtually none of that signal is reclaimable by Dual Recycling, since to quote Meers and Strain [26], "light which is scattered at high angles is gone for good". The exit port losses, by comparison, are not a significant factor.

These conclusions are demonstrated by the data in Table 4.1, which is also depicted in Figure 4.7. The data in the table does indeed show that the carrier power stored in the arm cavities (specifically quoting the results for the inline arm cavity) generally *goes up* for increasing R_{dual}, especially for large mirror surface deformations[1], and thus wavefront healing *works*, as predicted. But because of the presence of fine-scale mirror roughness deformations in "real" mirrors, the small amount of wavefront healing that can happen is not significant. This is apparent in Fig. 4.7, in which it is seen that the curves representing all of the different values used for R_{dual} lie right on top of one another, and thus have no real effect (only a ~3-4% increase, at most) on the overall arm cavity power. This is in contrast to the way that the stored arm carrier power *plummets* due to the increasing amplitude of the mirror surface deformations. Thus the overall conclusion which we can draw seems to be clear: in order to maintain high carrier power in the arm cavities for GW-signal generation, the implementation of Dual Recycling is not a sufficient solution; rather, the best solution is to make better (especially smoother) mirrors for the interferometer. Reading this result as severely as possible, the true importance of Dual Recycling (for "realistically deformed" mirrors, at any rate) appears to lay more significantly in its GW-signal curve adjustment capabilities, than in any ability to maintain high carrier power in the arm cavities. Never-

1. Increasing R_{dual} does not help in the case of perfect mirrors, because there the exit-port power consists of modes created by distortions in the *Michelson part* of the interferometer (due to the Schnupp Length Asymmetry, clipping by the beamsplitter apertures, etc.; c.f. Figure 3.4), and thus the Fabry-Perot arm cavities do not have the proper "matrix elements" (i.e., deformation patterns) to send this recycling power back into the TEM_{00} mode. Thus wavefront healing cannot occur.

	Mirror	Surface	Roughness	Family	(RMS)
R_{dual}	"Perfect"	$\lambda_{YAG}/1800$	$\lambda_{YAG}/1200$	$\lambda_{YAG}/800$	$\lambda_{YAG}/400$
0.0	4727.1	3997.6	3363.0	2484.1	1040.7
0.1	4722.4	3989.5	3352.6	2473.1	1034.8
0.3	4719.0	3991.8	3358.0	2480.8	1041.0
0.5	4716.9	3999.0	3370.2	2496.3	1052.3
0.7	4715.3	4007.0	3383.7	2513.2	1064.4
0.9	4714.1	4014.5	3396.1	2528.8	1075.7
0.99	4713.8	4017.4	3400.9	2534.9	1080.3

Table 4.1: Power resonating in the inline Fabry-Perot arm cavity (normalized to 1 Watt of carrier excitation power), as a function of Signal Recycling Mirror reflectivity and mirror deformation amplitude. A small amount of "wavefront healing" is demonstrated.

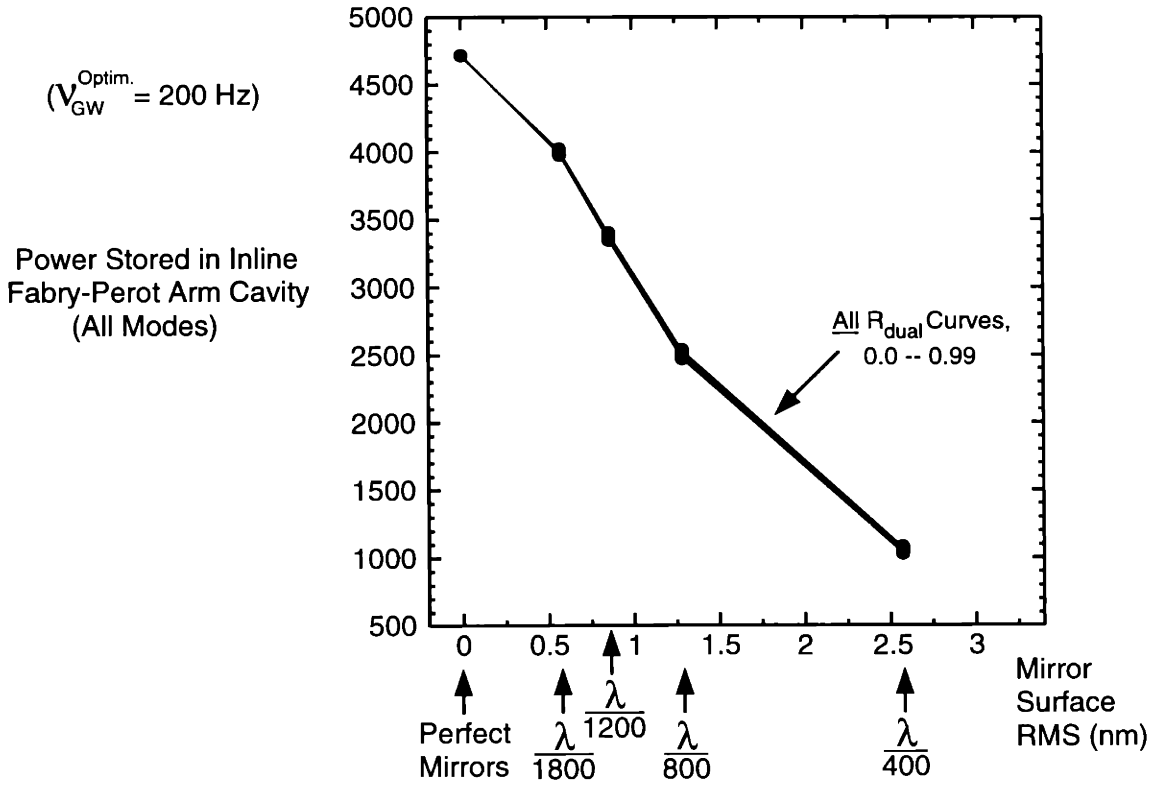

Figure 4.7: Resonant power buildup in the inline Fabry-Perot arm cavity, plotted versus mirror deformation amplitude. The curves representing different Signal Recycling Mirror reflectivities are virtually indistinguishable, despite the effects of Wavefront Healing.

theless, if the exit-port carrier power losses are indeed deemed to be significantly higher than desired (though perhaps small in an "absolute" sense), then Dual Recycling has been shown in the above plots to be successful at bringing these losses down.

In order to test that our program could correctly simulate the effects of wavefront healing in cases when they are truly significant, we performed runs for a Dual Recycling interferometer with perfect surfaces and substrates, but for which a "geometric deformation" has been added: *tilts* have been placed (antisymmetrically, to maximize exit port losses) on the two back mirrors of the Fabry-Perot arm cavities. The four tilt cases which we ran involved tilt angles (on each back mirror) of $\theta = 0$, 0.32 μrad, 0.475 μrad, and 0.95 μrad (for comparison, the "characteristic angle" at the back mirrors is given by $\theta_c \equiv \sqrt{2}\lambda/(\pi \times \text{spot size}) = 10.5\ \mu\text{rad}$). Note that these angles have been specifically chosen to cause the same degradation in the computed arm cavity powers (for an R_{dual} value of *zero*) as the realistically deformed mirror runs caused when using, respectively, the perfect, λ/1200, λ/800, and λ/400 mirror families.

The results of these runs are shown in Figure 4.8, where this time we have plotted the stored carrier power in the (inline) arm cavity as a function of R_{dual}, for curves representing different tilt values. The results shown here are very clear: increasing the reflectivity of

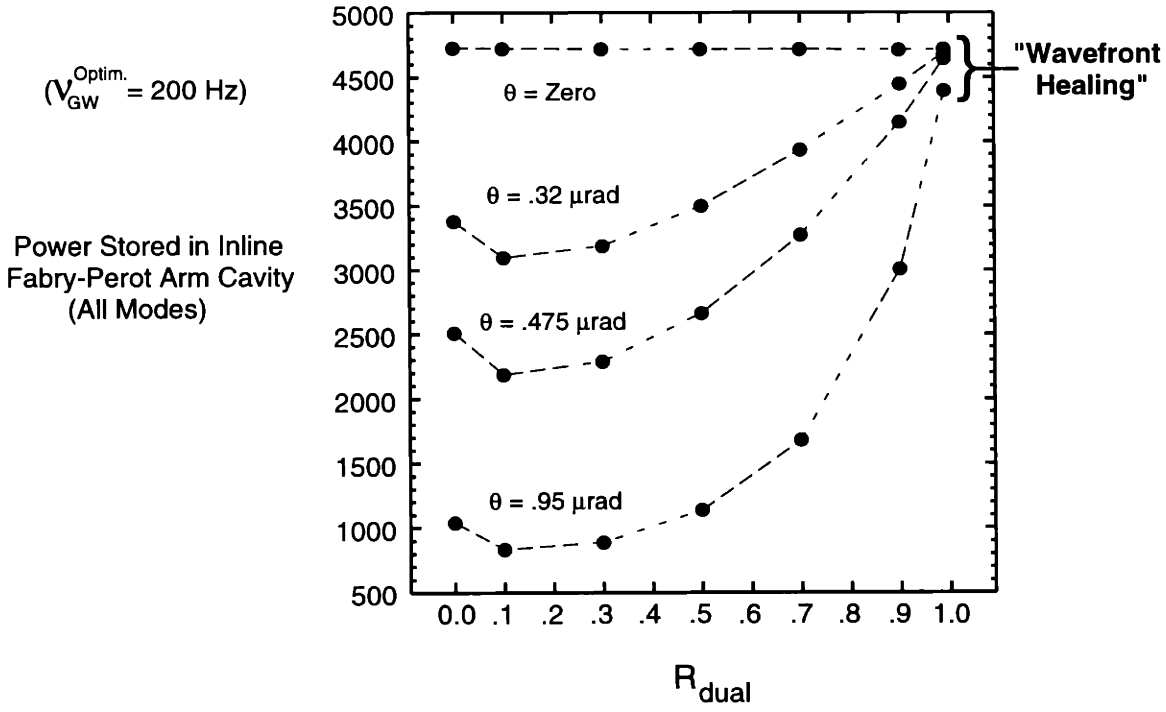

Figure 4.8: Resonant power buildup in the inline Fabry-Perot arm cavity, parameterized by arm cavity back mirror tilt values, and plotted versus Signal Recycling Mirror reflectivity. The effects of Wavefront Healing for high reflectivity values are evident.

the Signal Recycling Mirror strongly reduces the exit port losses[1], and hence restores the power in the arm cavities, which remains almost entirely TEM_{00} mode power. Using $R_{dual} = 0.99$, in fact, practically eliminated the adverse effects of the largest tilts that we modelled. This, then, is the true verification of wavefront healing, for a case where it is truly an important effect. So if this healing cannot be used to recycle the light lost due to *scattering* from "realistic" mirror deformations, then at least it can relax the tolerances on tilts and curvature mismatch errors of the interferometer optics. Still, we should note that the auto-alignment systems of LIGO are *already* intended to control mirror tilts to within ~.01 μrad [39] with or without the help of Dual Recycling, which is much too small to cause losses that are even remotely close to those which we have computed for any of the "realistically deformed" mirror cases.

To wrap up this section, we take the demonstrated effects of dual recycling — including signal enhancement and narrowbanding, noise reduction, and whatever wavefront healing effects have been achievable with realistically deformed mirrors — and combine them to produce an example of the overall benefits that the implementation of Dual Recycling could have for a LIGO interferometer. The shot-noise-limited part of the sensitivity curve must be computed, and its alteration due to Dual Recycling must be placed in perspective to the other dominant noise sources of the (Initial-) LIGO interferometer: seismic and thermal noise.

First, we compute the shot noise contribution via a straightforward, order-of-magnitude procedure. The GW-signal is produced by beating together the "local oscillator" light (i.e., radio-frequency sidebands) with the GW-induced sideband light, each of which exits the system at the interferometer signal (exit) port. The signal amplitude will thus be proportional to the product of these e-field amplitudes. Similarly, the shot noise will be generated by the power exiting the system there, including any carrier power which has not been prevented from escaping through the exit port, in addition to the local oscillator RF-sideband light which exits there intentionally for signal detection purposes. The (amplitude) signal-to-shot-noise ratio of the interferometer therefore obeys the proportionality:

$$\frac{S}{N}(f_{GW}) \propto \frac{h\sqrt{P_{GW\text{-}sideband}(f_{GW}) \cdot P_{Loc.\ Osc.}}}{\sqrt{P_{Carrier} + P_{Loc.\ Osc.}}} \tag{4.6}$$

Note that the Gravitational-Wave amplitude, h, has been factored out of the numerator so that we may use values of $P_{GW\text{-}sideband}$ (as depicted in Figures 4.1, 4.2, and 4.4) in this formula that are not normalized by any particular GW strength. Also note that both $P_{GW\text{-}sideband}$ and $P_{Loc.\ Osc.}$ represent the sums of the amplitudes of plus and minus sideband fields.

1. Once again, we note that small-but-nonzero values R_{dual} such as 0.1 or 0.3 make the losses worse, instead of better, due to the SRC-degeneracy effect which will be studied in Sec. 4.3.2.

The square of this ratio must then be multiplied by the efficiency of the output photodetector, η; and the shot noise power in the denominator — which is a measure of the uncertainty of the *number density* of photons striking the output photodetector at a given time — must be multiplied by Planck's constant (h_{pl}) and the frequency of the (carrier) laser beam, ν_{Las} (c.f. Eq. 1.7 and ensuing discussion). By setting the S/N ratio obtained in this way to unity, and solving for the resulting "shot-noise-limited" detectable value of h, we get:

$$\tilde{h}_{SN}(f_{GW}) = \sqrt{\frac{h_{pl}\,\nu_{Las}}{\eta}} \cdot \frac{\sqrt{P_{Carrier} + P_{Loc.\ Osc.}}}{\sqrt{P_{GW\text{-}sideband}(f_{GW}) \cdot P_{Loc.\ Osc.}}} \tag{4.7}$$

To convert this formula into a numerical result, we must set values for various quantities. The (Nd:YAG) laser frequency can be determined from the light wavelength of 1.064 microns, and as usual we assume a photodetector efficiency of $\eta = 0.8$. The total excitation laser power used for generating the carrier and RF-sidebands (and ultimately, the GW-sidebands as well) will be specified as 6 Watts, the typical power level that is assumed for the Initial-LIGO lasers.

Now, because RF-sidebands are not modelled in our Dual Recycling code, neither the amount of local oscillator power emerging from the signal port, nor the *division* of laser excitation power between the carrier and sideband frequencies via RF-modulation, can be numerically determined. In lieu of this, we assume a modulation depth that is "typical" of what we found for simulations of the Initial-LIGO interferometer; by examination of Table 3.1 for the results with $\lambda/800$ mirrors (and no output mode cleaner), we choose a modulation depth of $\Gamma = 0.5$, which means that there is initially ~5.3 Watts of carrier power and ~0.7 Watts of RF-sideband power initially entering the interferometer. We then make the rough approximation that ~90% of this RF-sideband power eventually makes it to the exit port, in order to serve as local oscillator light for GW-detection. The net result of these steps is that for the various power terms in Eq. 4.7, $P_{GW\text{-}sideband}(f_{GW})$ and $P_{Carrier}$ are multiplied by the factor 5.3, and $P_{Loc.\ Osc.}$ is multiplied by the factor 0.63.

Lastly, to compute the shot noise level implied by Eq. 4.7, we must specify the carrier power which is leaking out of the exit port, $P_{Carrier}$. As our example case, we will use runs with $\lambda/800$ mirror surface deformations (along with deformed substrate maps), for R_{dual} values of zero (i.e., "no Signal Recycling Mirror") and 0.9. For those runs, the exit port carrier power is, respectively, 2.234×10^{-2} and 4.115×10^{-3} (*before* multiplication by 5.3, to obtain their values in Watts).

Once the shot noise curves for the two runs have been thus determined, a quadrature sum of all contributing noise sources must be made to obtain the *total noise curve* in each case, much as was done previously in Sections 3.4.1 and 3.4.2 for runs with the Initial-LIGO configuration. Here, in order to represent the seismic noise curve, we will use the

same formula as was used above and depicted in Figure 3.1. This formula is the Initial-LIGO requirement [55] for the seismic noise level, and it is given as:

$$\tilde{h}_{\text{Seismic}}(f_{\text{GW}}) = 1.6 \times f_{\text{GW}}^{-14} \tag{4.8}$$

For the thermal noise, however, it is not appropriate to use the same noise curve as used previously (which was $3.8 \times 10^{-19} \cdot f_{\text{GW}}^{-2}$), because the thermal noise requirement curve for the Initial-LIGO interferometers is actually a *composite* curve meant to capture the effects of two distinct thermal noise contributions: internal vibrational noise within the mirrors, and pendulum-motion noise by the mirror suspensions. Internal thermal noise and suspension thermal noise have different functional dependencies upon GW-frequency, though, and for a configuration like Dual Recycling which has the capacity for deep narrowbanding to obtain sharp sensitivity peaks, it is not sufficient to approximate the total thermal noise curve well over a particular (possibly small) region; each contribution must be individually modelled.

Theoretical models for the suspension thermal [75, 76] and internal thermal [77] noise of LIGO mirrors have been studied by Gillespie and Raab, and formulas for the estimated mirror displacement errors thus generated were presented in a LIGO internal document on suspension design requirements [78]. In each case there is a (frequency-independent) "structural loss" value determining the amplitude and/or frequency dependence of the noise; we use recent estimates of $\varphi_{\text{suspension}} = 6.3 \times 10^{-6}$ and $\varphi_{\text{internal}} = 3.7 \times 10^{-7}$ from an internal LIGO document on the suspension preliminary design [79], and a factor of $4000/\sqrt{2}$ to convert from displacement noise to an equivalent rms Gravitational Wave strength. This yields the following expressions for the thermal noise components:

$$\tilde{h}_{\text{Thermal, Suspension}}(f_{\text{GW}}) = 1.16 \times 10^{-18} \cdot f_{\text{GW}}^{-5/2} \tag{4.9}$$

$$\tilde{h}_{\text{Thermal, Internal}}(f_{\text{GW}}) = 2.02 \times 10^{-22} \cdot f_{\text{GW}}^{-1/2} \tag{4.10}$$

Using Eq's. 4.7 – 4.10, we can then compute the total noise curve as follows:

$$\tilde{h}_{\text{Total}}(f_{\text{GW}}) = \sqrt{\tilde{h}^2_{\text{Seismic}} + \tilde{h}^2_{\text{Thermal, Suspension}} + \tilde{h}^2_{\text{Thermal, Internal}} + \tilde{h}^2_{\text{SN}}} \tag{4.11}$$

The total noise curves for the runs with and without Dual Recycling, along with each of the individual noise components that comprise these curves, are shown in Figure 4.9.

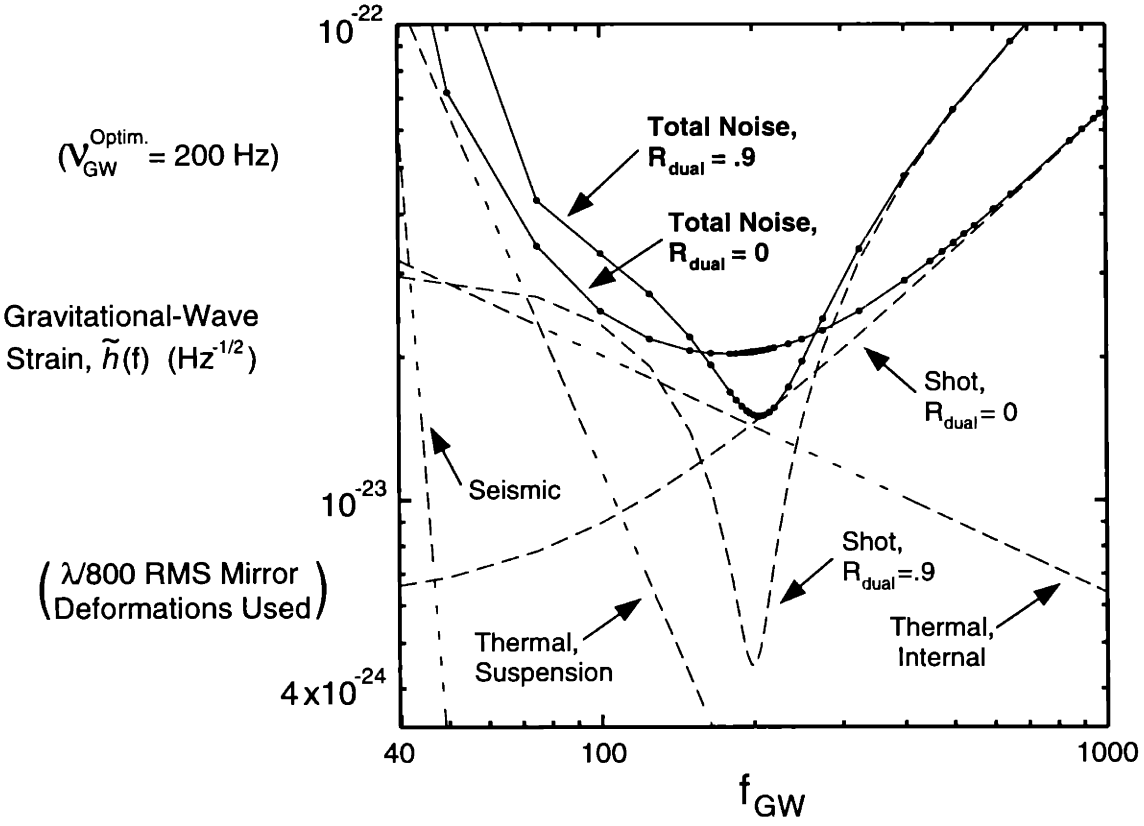

Figure 4.9: The Gravitational-Wave strain sensitivity of a simulated LIGO interferometer is shown, as a sum of all contributing noise sources. The narrowbanding and peak enhancing effects with Signal Recycling ($R_{dual} = .9$) are compared to the frequency response without Dual Recycling ($R_{dual} = 0$), for runs using lambda/800 mirror deformation maps. Note that the gain is peak sensitivity is limited by thermal noise.

This figure shows that Dual Recycling can indeed improve the sensitivity over a narrow band in frequency space, and hence produce a "notch" in sensitivity within some region of interest (e.g., to search for a pulsar of known rotation frequency that is believed to emit significant power in GW's). On the other hand, it is also shown how most of the deep, shot-noise-limited sensitivity enhancement achieved by a high Dual Recycling factor is rendered useless, due to the presence of an internal thermal noise "barrier", which decreases only as the square root of the GW-frequency. Thus much of the good of Signal Recycling for large values of R_{dual} would be undone without an improvement in the mirror internal thermal noise, and in fact the loss of sensitivity away from the peak turns the use of high R_{dual} into a more questionable trade-off. This result stresses the importance of improving the internal thermal noise curve, such as through the use of lower-mechanical-dissipation materials for mirror substrates (e.g., sapphire instead of fused silica) [12]. If the Q-factors of the materials used for the mirrors could be improved by something like

~10–100, then the amplitude of the $\tilde{h}_{\text{Thermal, Internal}}$ curve could be reduced by an overall factor of $\sqrt{Q}$ ~3–10.

It is clear that more of the sensitivity peak becomes useful as the optimization frequency of the SRC, $\nu_{GW}^{optim.}$, is tuned higher. This will indeed allow a sharper sensitivity peak to be created, but at the cost of moving away from frequencies like 200 Hz, where significant GW-emission events such as the onset of the "merger" phase for 10 Solar Mass Black Hole/Black Hole binaries (as well as the heart of the merger phase for more massive binaries) are expected. And even going all the way to a tuning of $\nu_{GW}^{optim.} = 1000$ Hz improves the internal thermal noise limit by only a factor of $\sqrt{5}$.

To sum up the results of this section, we conclude that Dual Recycling does indeed work — to a limited extent. It does reduce the exit port carrier power loss, and thus the shot noise, to a significant degree (if the total shot noise level is not completely dominated by contributions from local oscillator light), and a narrowband sensitivity peak can indeed be achieved even in the case of realistically deformed mirrors. In Fig. 4.9, particularly, the sensitivity was improved by a factor of ~1.36 (%26 noise reduction) at the optimization frequency of 200 Hz, leading to a projected event rate enhancement of ~$(1.36)^3$ ~2.5. Nevertheless, it is clear that the most benefit can be gained out of Dual Recycling if *better mirrors are made*. Not only would a decrease in the mirror internal thermal noise level allow more of the Dual Recycling sensitivity peak to be usefully employed, but mirrors with better high-spatial-frequency polishing and good coatings would reduce the high-angle scattering losses which dominate the system, and thus allow the benefits of "wavefront healing" to become a significant factor in maintain high Gravitational Wave signals.

4.3 Problems and solutions for Dual Recycling

4.3.1 Large-aperture mirrors and scattering losses

As noted above, the primary reason why wavefront healing is not very successful for recycling "lost" power into the interferometer for the regeneration of arm cavity power, is that most of the losses due to mirror deformations consist of power that is scattered beyond the apertures of the finite-sized mirrors. A possible solution to this problem becomes obvious: perhaps the mirrors can be made larger, this enabling the capture and recycling of more of the mirror-perturbed light power.

To address this question, we have performed four sets of simulation runs: both with ($R_{\text{dual}} = 0.9$) and without Signal Recycling, for both the perfect mirrors case and the $\lambda/800$ mirrors case. The goal is to see how close Dual Recycling can bring the stored arm cavity power back to its value for the perfect mirrors case, but also to determine how much of the "healing" (if any) which happens would be produced by the presence of the larger mirrors even in the absence of Dual Recycling.

For each of these four cases, we performed runs using mirror aperture radii of 12, 15, 18, 21, 24, 30, and 35 centimeters. In each run, *all* of the mirrors were given the same radius, including the beamsplitter, which was made circular[1] and equal in size to the other mirrors in order to avoid any complications in the results that could otherwise occur. The largest mirrors used in these runs take up a substantial fraction of the entire width of the LIGO beamtube, which must in fact house several different LIGO detection beams simultaneously [e.g., 3]. To numerically model these large mirror sizes, it was necessary (c.f. Sec. 2.3.4) to go to larger calculational window sizes. We used calculational windows spanning 1.4 x 1.4 m, thus requiring (to maintain the same pixel density) 512 x 512 grids.

The results of these runs are shown in Figure 4.10. It is clear that larger mirrors do increase the success of wavefront healing — at first. Up to radii of 21 cm, the power in the (e.g., inline) arm cavity does substantially increase with mirror size, regaining ~40% of the power that was lost due to the inclusion of deformed mirror maps. And this power gain

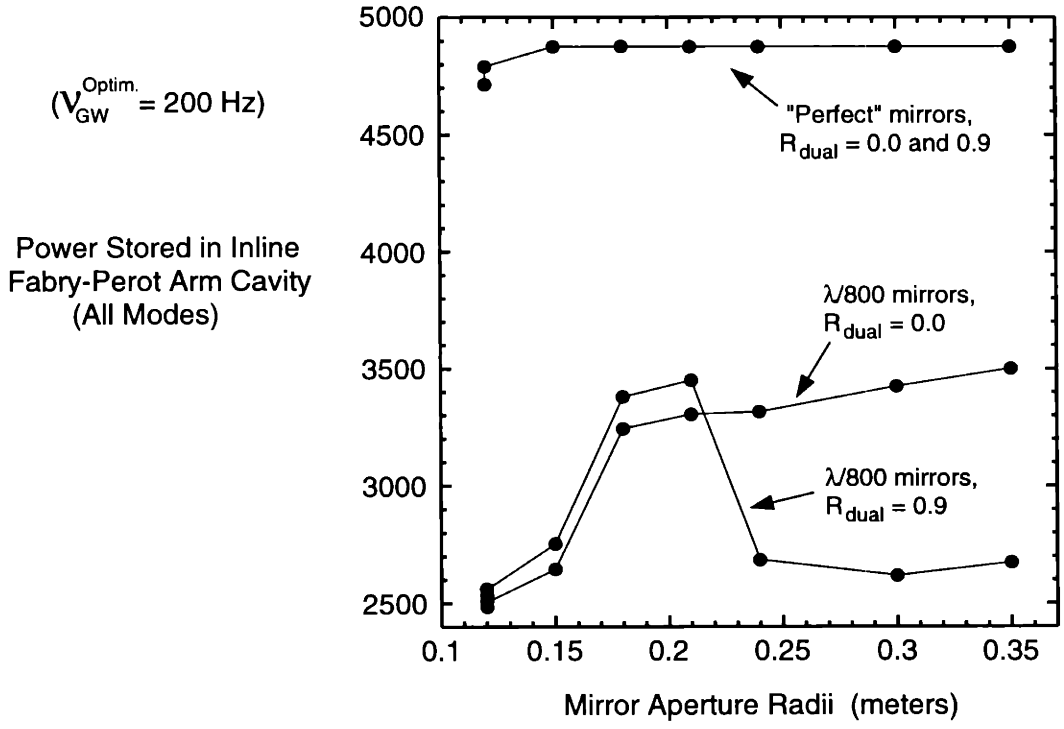

Figure 4.10: Resonant power buildup in the inline Fabry-Perot arm cavity, plotted versus the aperture radii of all interferometer mirrors. Curves for 4 cases are shown: "perfect" and "deformed" mirror maps, with and without a Signal Recycling Mirror.

1. A run was also done with a "normal", elliptical beamsplitter for the first run with 12 cm mirror radii, in order to see that any effects due to the difference were reasonably small; as is shown in Figure 4.10, the change in arm cavity power between the circular and elliptical beamsplitter cases is not large, for these smallest mirror runs (which have the most significant beam clipping effects).

is almost entirely in the TEM_{00} mode (all modes increased by about the same multiplicative factor, but non-TEM_{00} modes were negligible in the arm cavity to begin with), validating the arguments presented in Sec. 4.1.3. Both the Dual Recycling configuration and the ordinary Power Recycling configuration benefited from these healing effects, though the implementation of Dual Recycling does increase wavefront healing a little bit.

This result changes sharply, however, when the radii of the mirrors are increased to 24 cm and beyond. The rapid restoration of power ceases for both configurations at that point, and for Dual Recycling in particular, the arm cavity power plummets back down again. A corresponding plot (not shown) of the power leaking from the exit port shows that a massive increase in lost power also occurs for Dual Recycling at this point, which kills off the process of wavefront healing.

A modal analysis of the exit-port power reveals a tremendous increase in the leakage of power in *all* TEM_{mn} *modes* with $m + n = 23$. All of these modes experience the same Guoy phase for propagations through the long-baseline arms (since ϕ_{Guoy} is proportional to $m + n$), and could thus be simultaneously resonant in the arm cavities (and in the SRC, which is degenerate anyway). The interpretation of the power drop-off effect is therefore clear: the modes with indices summing to 23 are accidentally resonant within the interferometer — as is the case for the TEM_{00} loss mode in Broadband Dual Recycling — and thus they are highly amplified in their leakage out of the exit port, as soon as the mirrors are large enough to support their oscillation in the system.

For a given mode with index m (in one dimension), with spot size ω_z, the half-width of that mode can be estimated as [17]:

$$X_m \approx \sqrt{m} \cdot \omega_z \tag{4.12}$$

At the arm cavity back mirrors, $\omega_z = 4.565$ cm, which implies that for an X_m of 21 cm, $m \approx 21$, while for an X_m of 24 cm, $m \approx 27$. This result agrees nicely with the results of Figure 4.1: for radii of ~20 cm, several modes with $m + n = 23$ (but with neither index being much larger than the other) are resonantly amplified as losses, and thus wavefront healing is slowed; and as the mirror aperture radii become as large as 24 cm, *all* of these modes can survive in the arm cavity, and hence the stored power in the arms finally plummets, undoing the healing effects of the making the mirrors larger, in addition to creating a lot of $m + n = 23$ mode power which oscillates in the arm cavities. Direct visual inspection (not shown here) of the exit port e-field also shows a dramatic increase in its transverse size (as well as in its amplitude and nodal complexity), nearly doubling in transverse size when going from the 21 cm mirrors run to the 24 cm mirrors run. It is clear that a few spatially large modes have been enabled to resonate because of the larger interferometer mirrors.

The effects for the non-Dual-Recycled case are similar, though less severe, most likely because it does not have a Signal Recycling Mirror to support the full double resonance

for the $m + n = 23$ modes which occurs in the coupled SRC/arm cavity system. In this non-Dual-Recycled configuration, there is a little bit of wavefront healing still occurring even for runs with the largest mirrors, because the $m + n = 23$ mode losses are only *comparable* to the losses in other modes, not dominant. But for the Dual Recycling interferometer, the $m + n = 23$ mode losses become ~20 times larger than the next-largest loss modes (typically low-index modes) for the run with 24 cm mirror radii, so that it represents a major drain upon the circulating interferometer power. Fully ~15–20 mW *per mode* (when normalized to 1 Watt of total carrier excitation power) is lost for several of them, and the total exit port loss for the 24 cm mirrors run is adds up to ~170 mW (and this total loss actually grows as large as ~300 mW, for 35 cm mirrors).

It is not clear if there is a solution to this problem, which is most severe for the Dual Recycling system, and prevents any real benefits from being gained via wavefront healing for runs with realistically deformed mirrors. Making the mirrors larger always allows more and more modes to oscillate within the system, and it is impossible to ensure that *all* of them will be simultaneously non-resonant (though the *degree* of resonant loss amplification may be quite dependent upon the precise configuration of each particular system, e.g., whether or not the SRC is degenerate, etc.). Indeed, using apertured mirrors is a primary means of *clipping away* higher modes in order to prevent them from contaminating the circulating interferometer light. In the case of the 40 meter LIGO prototype interferometer at Caltech, for example, a similar case of resonant buildup of unwanted higher modes was seen to happen, for which at least part of the solution was the tighter clipping of the beam via smaller mirror apertures [80]. It is certainly possible to increase the mirror radii for the Dual Recycling configuration up to a carefully specified point, and then make sure that all of the modes that fit within that given mirror radius are non-resonant; but such calculations are analytically difficult for a system with many mirrors and complicated beam focusing behavior, and they must be increasingly accurate as the mode number — and thus the Guoy phase — becomes extremely large, so that even a small error could lead to the round-trip phase of a mode coming too close to an integral multiple of 2π.

Therefore, we conclude that making the mirrors larger is a strategy that is risky at best for augmenting the effects of wavefront healing, and at worst it can devastate the performance of a Dual-Recycled system. Once again, it appears that making the mirror surfaces (and substrates) better is a superior method — and perhaps no more costly than making very large mirrors with even *equivalently* well polished surfaces — for achieving the goal of better arm cavity power storage.

4.3.2 Degeneracy and degeneracy-breaking with a long-baseline SRC: The significance for runs with GW-frequency tuned to 1 kHz

The last topic which we will discuss in this chapter on Dual Recycling will be the effects of beam mode *degeneracy* in the Signal Recycling Cavity, the significance of which was

alluded to in Section 4.1.2, and the effects of which have been noticeable (albeit small) in the results of the Section 4.2.3. We mention again the fact that this SRC degeneracy is exclusively a property of a Dual-Recycled system which includes Fabry-Perot arm cavities, since it is the arm cavity input mirrors which cause the SRC to be a separate system of short length (and negligible beam focusing), so that the various beam modes pick up negligible mode-dependent Guoy phase during round-trip SRC propagations.

Consider once more the plot of exit-port power loss (for realistic mirror deformations) in Figure 4.6, and the plot of resonant inline arm cavity power (for tilt deformations) in Figure 4.8. In both cases, it is apparent that a *small* value of R_{dual} is worse than either a large value or zero. Thus a little bit of Dual Recycling actually increases the exit port losses, which in turn causes the inline arm cavity power to decrease by a little bit; and as we have determined from a modal decomposition of the exit port field for the run with realistic mirror deformations, this increased power loss occurs for *nearly all modes*. The sole significant exception to this behavior is, interestingly, the TEM_{00} mode, which always decreases with increasing R_{dual} for these $\nu_{GW}^{optim.} = 200$ Hz runs.

This behavior can be understood through consideration of the transmissive behavior of a simple two-mirror cavity, in this case representing the SRC, as its resonant condition is varied. For input mirror parameters $t_1 = \sqrt{T_1}$, $r_1 = \sqrt{R_1}$, back (output) mirror parameters $t_2 = \sqrt{T_2}$, $r_2 = \sqrt{R_2}$, and a (potentially mode-dependent) round-trip propagation phase of ϕ, the power in the transmitted e-field will be:

$$\left|\vec{E}_{transmitted}\right|^2 = \left|\vec{E}_{incident}\right|^2 \cdot \left(\frac{t_1^2\, t_2^2}{1 + r_1^2\, r_2^2 - 2\, r_1\, r_2 \cos\phi} \right) \tag{4.13}$$

Treating r_1, t_1 and ϕ as given, we consider what value of r_2 (i.e., $\sqrt{R_{dual}}$) would maximize the transmitted power — in other words, maximize the exit port losses.

First of all, for a system that is closer to anti-resonance than resonance, in which $\cos\phi < 0$, it is clear that any increase in r_2 (with the corresponding decrease in t_2), will decrease the transmitted power. Thus for Broadband SRC tuning, any finite value of R_{dual} will suppress the losses for all non-TEM_{00} modes, while for Tuned Dual Recycling with $\nu_{GW}^{optim.}$ set to a high value, the losses for the TEM_{00} mode (which picks up a π phase shift in reflection from the resonant arm cavities) will also be reduced by Dual Recycling.

But as the system becomes semi-resonant ($\cos\phi > 0$) and then eventually resonant ($\cos\phi = 1$), the negative term in the denominator of Eq. 4.13 becomes important, and implementing a finite value of R_{dual} can actually *increase* exit port losses; an example of the resonant case is the increase of TEM_{00} losses for Broadband Dual Recycling.

The same effect happens for non-TEM_{00} mode losses, also, but in reverse fashion: it is for higher tunings of $\nu_{GW}^{optim.}$ that implementing Signal Recycling can make the transmitted losses worse. For the case of a semi-resonant SRC, such as $\nu_{GW}^{optim.} = 200$ Hz, the non-TEM_{00} mode losses are actually maximized for *intermediate* values of R_{dual}. This is, in

fact, the explanation of the aforementioned loss behavior seen in Figures 4.6 and 4.8. This effect will only get worse for higher SRC tunings, where the loss-maximizing value of R_{dual} will be even closer to unity (thus harming any attempt to produce a sharp GW-signal peak for those cases), while also making the resonant amplification of losses more and more pronounced at the loss-maximizing point.

The source of these problems is the degeneracy of the SRC, since it is impossible to tune the system for TEM_{00} mode GW-signals at a given value of $\nu_{GW}^{optim.}$ without making all of the non-TEM00 modes simultaneously share the same resonance condition. Only by breaking this degeneracy can the different non-TEM_{00} modes fall under various conditions of semi- or non-resonance, while the TEM_{00} mode is properly resonant in the SRC for large $\nu_{GW}^{optim.}$ settings.

There are two fundamental ways of breaking the SRC degeneracy (barring the addition of yet another SRC mirror, such as in "Compound Dual Recycling" [26]). The beam can either be made to undergo strong focusing during propagation through the ~few-meters-long SRC (which would give the beam a very small waist size), or it can undergo the gradual focusing effect of normal propagation through a *much longer* SRC. The former option is very difficult to accomplish with a grid-based simulation program, because the small beam waist would require the use of much finer grids (c.f. Sec. 2.3.4.2), and the necessity of keeping the total size of the calculational windows the same (c.f. Sec. 2.3.4.3) would mean that many more pixels would need to be used in the 2-D grids, leading to much longer program run-times and memory (and disk storage) requirements. For our tests of degeneracy-breaking, therefore, we have opted to use a long-baseline SRC for the runs[1]. The program will thus have to implement a weak-focusing lens at the beamsplitter dark port in order to match the beam from the main interferometer to the Signal Recycling Mirror at the back of the SRC, and to ensure that it passes through a (not too narrow) focus so that each mode will pick up appreciable (and different quantities of) Guoy phase during SRC propagations.

SRC lengths anywhere from hundreds of meters to a ~few km would serve our purpose for these runs; for the sake of relevance to the LIGO project, we restrict our considerations to two of the more realizable possibilities: placing the Signal Recycling Mirror at either the *mid-station* of the interferometer ($L_{SRC} \approx 2$ km), or at the *end station* of the interferometer ($L_{SRC} \approx 4$ km).

For two reasons, both related to maintaining compatibility to previous results, we have chosen to put the Signal Recycling Mirror at the mid-station. First of all, it is convenient to keep all of the beam spot sizes and curvature radii the same everywhere as for the short-

1. The question of which method of degeneracy-breaking (strong focusing or a long-baseline SRC) is actually "better", is a technical decision which cannot be made on the basis of a superficial theoretical analysis. Nevertheless, as will be discussed below, we can state that the use of a long SRC will lead to some additional scattering losses, though it may have the benefit of allowing *both* the "plus" and "minus" GW-induced sidebands to be resonant in the system simultaneously.

SRC case, so that none of the optical parameters will change. This is important for maintaining the same modal resonance conditions of the system (that is, the same everywhere except in the SRC), the same scattering losses due to the apertures of the finite-sizes mirrors (*including* at the Signal Recycling Mirror), and so on. As it happens, given the particular beam parameters at the beamsplitter dark port for the LIGO-like system which we have been simulating, the length for a *confocal cavity* (which is the longest symmetric cavity possible, once the beginning and ending spot sizes have been fixed) is ~3900 m. Thus a 4 km long SRC would be near-confocal, therefore defeating the purpose of the exercise, since a confocal cavity has the property that all round-trip Guoy phases are equal to π, so that every TEM_{mn} mode with indices that sum to an even value will be simultaneously resonant in the SRC along with the TEM_{00} mode.

Second of all, one must consider the fact that the *frequency response curve* of the system changes for a long-baseline SRC. On the one hand, there a slight narrowing of the peak width for a long SRC; but this effect is relativity minor. Much more important is what happens because of the "minus" GW-induced sideband field, which is normally negligible in its effects. For the normal case of a short SRC, the system can only be optimized for one of the two GW-sidebands, not both; hence Tuned Dual Recycling immediately loses a factor of two in GW-signal amplitude compared to Broadband (or the absence of) Dual Recycling, which it must make up for by sharp narrowbanding. But a physically long SRC can be made resonant for *both* GW-sidebands simultaneously [e.g., 74], for the following reason. The minus GW-sideband creates a response peak that is similar to the one created by the plus GW-sideband, except that it is located at very high (or equivalently, at very "negative") frequencies; and as L_{SRC} and/or $\nu_{GW}^{optim.}$ are increased, the sensitivity peak created by the resonance of the minus GW-sideband migrates down until it occupies the same general region in GW-frequency space as the plus sideband does. In particular, when $\nu_{GW}^{optim.} = 1000$ Hz and $L_{SRC} = 4$ km, the two peaks are practically on top of one another, *nearly doubling* the sensitivity of the interferometer at 1000 Hz compared to the short-SRC case. As important as this result is for the real LIGO, however — and in fact, it provides an independent argument for the usefulness of long-baseline SRC lengths — it does not satisfy our stated goal of simulating a generally unchanged interferometer. For this reason, in addition to the confocal cavity problem, we have chosen to model the 2 km length SRC because it still has a very similar frequency response to the short-SRC case for the highest tuning frequency that we will consider in this thesis (i.e., 1000 Hz).

For this (as we shall see, initial) realization of a long-baseline SRC, we placed an idealized lens — which had to satisfy the grid-based curvature limitations discusses in Sec. 2.3.4.1 — at the output of the beamsplitter dark port, in order to convert the beam from a curvature radius of +10.02 km to -14.18 km. This led to a waist size of 3.506 cm in the SRC (which is very similar to the beam waist size in the arms), and a spot size at the Signal Recycling Mirror which was exactly equal to that for the short-SRC case. The beam curvature there was a bit flatter than for a short-SRC, but since the Signal Recycling Mir-

ror's radius of curvature was adjusted to be mode-matched to the beam, the difference mattered little. New, non-trivial *anti-aliasing* parameters (c.f. Sec. 2.3.4.3) had to be introduced for SRC propagations because of the SRC's length, and they were specified as $N_p = 81$ and $N_a = 148$. The actual length of the SRC was a little bit longer than 2 km (the distance between the beamsplitter and the Signal Recycling Mirror was 2004.19 m), because the path from the Signal Recycling Mirror to the mid-station has to first backtrack though the Michelson arms of the interferometer.

Initially, we studied the case of broken SRC-degeneracy for $\nu_{GW}^{optim.} = 200$ Hz, for comparison with the results presented in Sec. 4.2.3. All of the runs were conducted with the $\lambda/800$ family of mirror deformation maps, and the value of R_{dual} was varied to produce a set of runs. The resulting exit-port power loss as a function of R_{dual} is shown in Figure 4.11, which includes some of the same data shown previously in Fig. 4.6, to be used as a source of comparison for the long-SRC runs shown here.

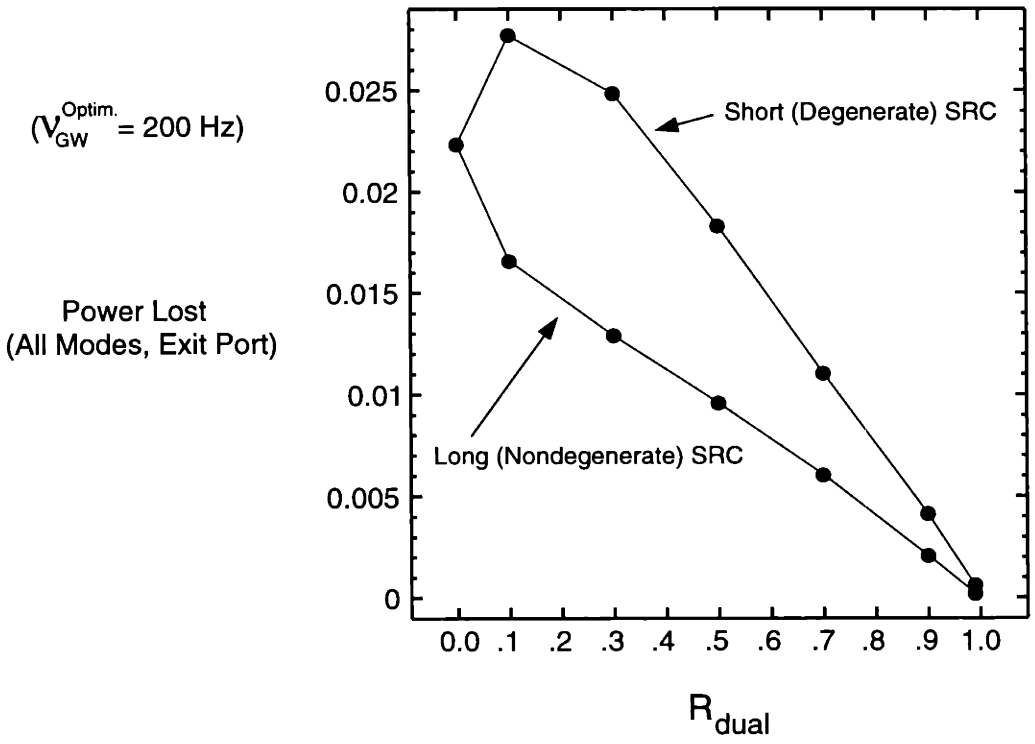

Figure 4.11: Interferometer exit-port power losses are plotted versus Signal Recycling Mirror reflectivity, both for the normal ("short") SRC, and for a long (~2 km) SRC designed for degeneracy-breaking. Lambda/800 mirror deformations have been used.

For the case of a degenerate SRC, note that the power loss *rises* as R_{dual} is initially increased from zero: this leads to a drop in resonating power in the inline arm cavity, which is analogous to the power dip seen for the curves in Figure 4.8 (except that these exit port losses for the realistic mirror case are less significant a source of loss than they

are for the runs with mirror tilt deformations). In any case, once R_{dual} is brought high enough so that the merely *semi*-resonant SRC no longer effectively channels non-TEM_{00} power out of the interferometer, then increasing the reflectivity of the Signal Recycling Mirror reverts to providing a suppression of the exit-port losses. But for intermediate values of R_{dual}, the performance with Dual Recycling is clearly degraded compared to the non-Dual-Recycled case.

By breaking the degeneracy with a long-baseline SRC, however, this effect disappears, and the power lost from the exit port drops monotonically with increasing R_{dual}. Similarly, the power resonating in the (e.g., inline) arm cavity also goes up monotonically with R_{dual}, though the results for the non-degenerate and degenerate cases eventually meet up for reflectivities very close to unity[1].

Despite this proof-of-principle result that degeneracy-breaking has solved this somewhat subtle problem, we must ask how relevant it is; after all, all of the exit-port power loss values shown in Figure 4.11 are fairly small, the increase in the losses for moderate R_{dual} due to the SRC degeneracy is relatively minor, and the total variation of the inline arm cavity power among all of these runs is only ~2.5%. The answer to this question, however, is that this effect *is* important, because of the fact (as alluded to above) that the problem of SRC degeneracy gets significantly worse as the SRC is tuned to higher and higher GW-frequencies.

To demonstrate this, we have performed similar runs for $\nu_{GW}^{optim.} = 1000$ Hz. In these runs, however, there was at first a small catch: it turned out that TEM_{mn} modes with $m + n = 6$ were very close to resonance in the SRC/arm cavity system, and that the exit port losses were dominated by the TEM_{06} mode — which was the most strongly produced, and the least clipped by the beamsplitter of these near-resonance modes — for large R_{dual}.

It was therefore necessary to alter the optical parameters slightly, in order to push these modes off resonance, much as McClelland, *et al.* [49], did for the TEM_{20} and TEM_{02} modes in Dual Recycling runs with mismatch due to mirror curvature errors. We implemented this change by altering the idealized lens at the beamsplitter dark port to give the departing beam a curvature radius of -12.0 km instead of -14.18 km, which changed the beam spot sizes and curvatures by a ~few percent everywhere, but those differences were nevertheless very minor. From these changes, however, the round-trip Guoy phases were altered just enough so that none of the modes generated in the interferometer were particularly close to resonance[2].

---

1. Actually, the inline arm cavity power is slightly (~.4%) lower for the non-degenerate SRC case when $R_{dual} = 0.99$, because of *scattering losses* in the long-baseline, moderately-high-finesse SRC.
2. There will of course be an "infinite" number of modes that fall close to resonance in the system, when arbitrarily high order modes are considered; but clipping by the mirror apertures was able to ensure that all such modes were prevented from resonating in the Dual-Recycled interferometer.

The results for the exit-port power loss for the $\nu_{GW}^{optim.} = 1000$ Hz tuning condition are shown in Figure 4.12, for three cases: the short (degenerate) SRC, the first (TEM_{06}-dominated) version of the long SRC, and the updated, second version of the long (non-degenerate) SRC configuration.

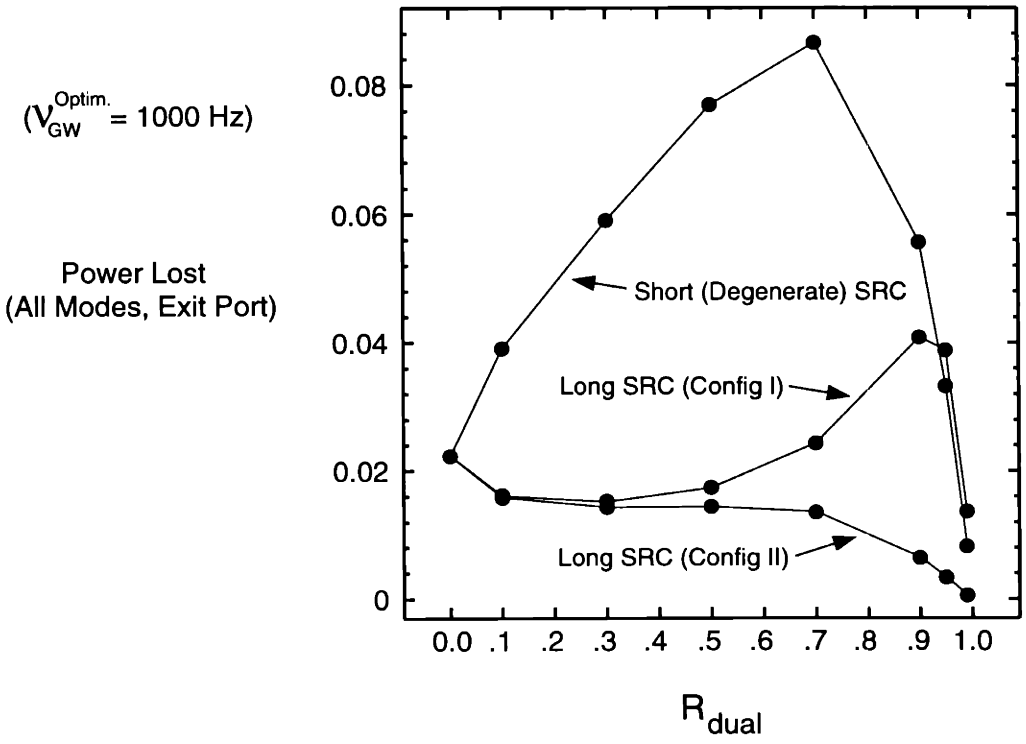

Figure 4.12: Interferometer exit-port power losses are plotted versus Signal Recycling Mirror reflectivity, for different configurations of the Signal Recycling Cavity. The lambda/800 family of deformed mirror maps has been used. The effects of SRC degeneracy, and degeneracy-breaking via long SRC round-trip lengths, are shown.

In accordance with the predictions made above, which are based upon Eq. 4.13, the degenerate system suffers much more of an increase in overall power loss with respect to the non-degenerate case than it did for prior runs with $\nu_{GW}^{optim.} = 200$ Hz; and the worst increase in losses (representing nearly a %400 jump) is encountered for a higher value of the Signal Recycling Mirror reflectivity than before, $R_{dual} \approx 0.7$. The first version of the non-degenerate SRC configuration alleviated this problem until the TEM_{06} mode began getting pumped out of the system for $R_{dual} \geq 0.5$, while the updated non-degenerate SRC configuration did even better, helping reduce the losses in monotonically improving fashion for *all* values of R_{dual} (with a factor of ~33 reduction for the $R_{dual} = 0.99$ case). Not only does this figure show that dramatic improvements can be made by the careful implementation of SRC degeneracy-breaking, but also that these improvements are not negligible, since the resonantly-amplified losses can become large — the highest exit-port loss

for this case was nearly 9% of all laser excitation power, and it would likely become worse for runs with higher GW-tuning frequencies.

The effects of this loss behavior upon the stored arm cavity power is shown in Figure 4.13. When the SRC-degeneracy is broken via version II of long-SRC configuration, we see that the normally expected wavefront healing effect does occur, ultimately increasing the resonant power in the (inline) arm cavity by a modest factor, ~2%. But for the fully-degenerate, short-SRC configuration, a significant *"Wavefront Harming"* effect occurs which is more that three times bigger than the healing that was seen with the non-degenerate case, which leads to a decrease in the arm cavity power of ~6% at its lowest point. Thus we see, from this and from the previous results in this chapter, that Dual Recycling can very easily have the opposite effect from which it is intended — if not used carefully, it can harm instead of heal.

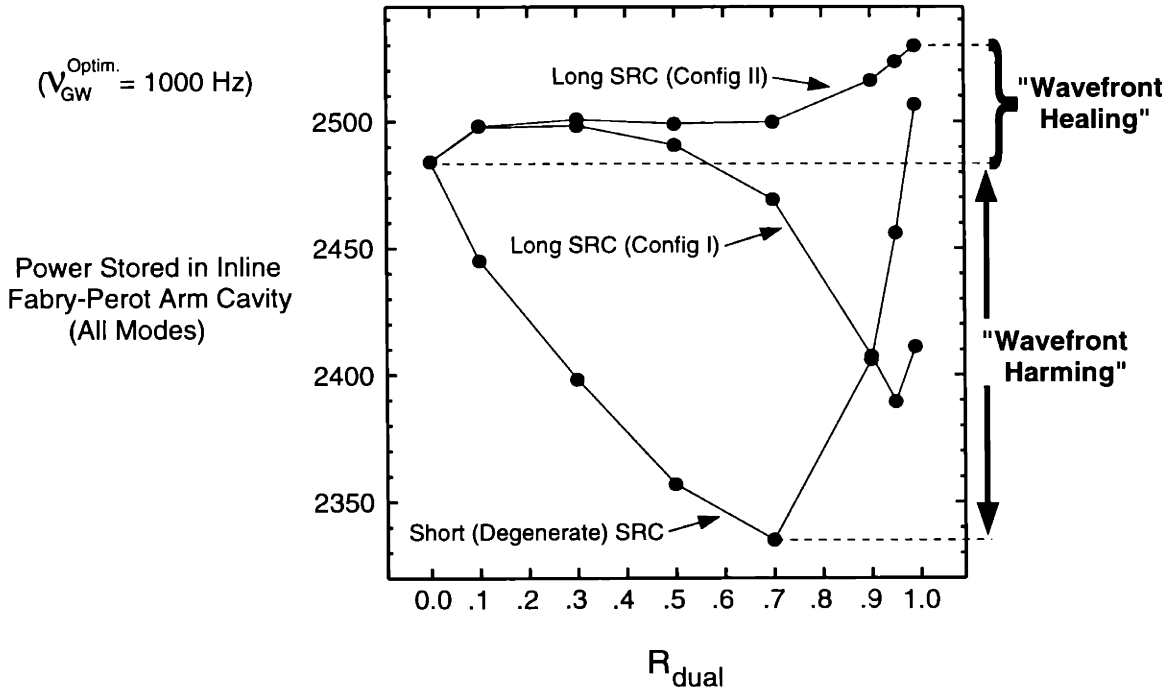

Figure 4.13: Resonant power buildup in the inline Fabry-Perot arm cavity is plotted vs. Signal Recycling Mirror reflectivity, for different SRC configurations. The lambda/800 family of deformed mirror maps has been used. The effects of "Wavefront Healing" (for nondegenerate SRC) and "Wavefront Harming" (for degenerate SRC) are shown.

As a concluding result for these investigations, it is useful to examine how the signal-to-noise ratio of the interferometer is affected by these degeneracy effects, both in terms of the healing (or harming) that occurs for the resonant power stored in the arm cavities, which affects the GW-signal, in addition to how it reduces (or increases) the power lost through the exit port, which affects the interferometer shot noise level.

In Figure 4.14, the ratio of the *square roots* of the carrier TEM_{00} mode power in the (inline) arm cavity, to the power emerging (in all modes) from the exit port, is shown; this ratio will be proportional to the amplitude S/N ratio. This particular expression (which we have normalized to 1 for the short-SRC case with $R_{dual} = 0$) is useful because the signal gain due to the SRC is not incorporated into it, so that only variations with respect to R_{dual} which directly demonstrate changes in the behavior of wavefront healing/harming are factored into the results. One note of caution is that since no local oscillator power has been put into this formula, this S/N estimate only represents a rough approximation to the true signal-to-noise ratio, and the estimate of the shot noise level in particular will be underestimated, especially for runs with very low carrier exit-port loss. On the other hand, large exit-port losses often mean that large local oscillator output powers are necessary for raising the signal back above the noise (c.f. Table 3.1), implying that the shot noise contributions of the local oscillator beams and the carrier beam would increase in roughly similar fashion, so that the ratio shown in Fig. 4.14 may indeed be a good predictor of the true signal-to-noise ratio for a Dual-Recycled system with RF-sidebands.

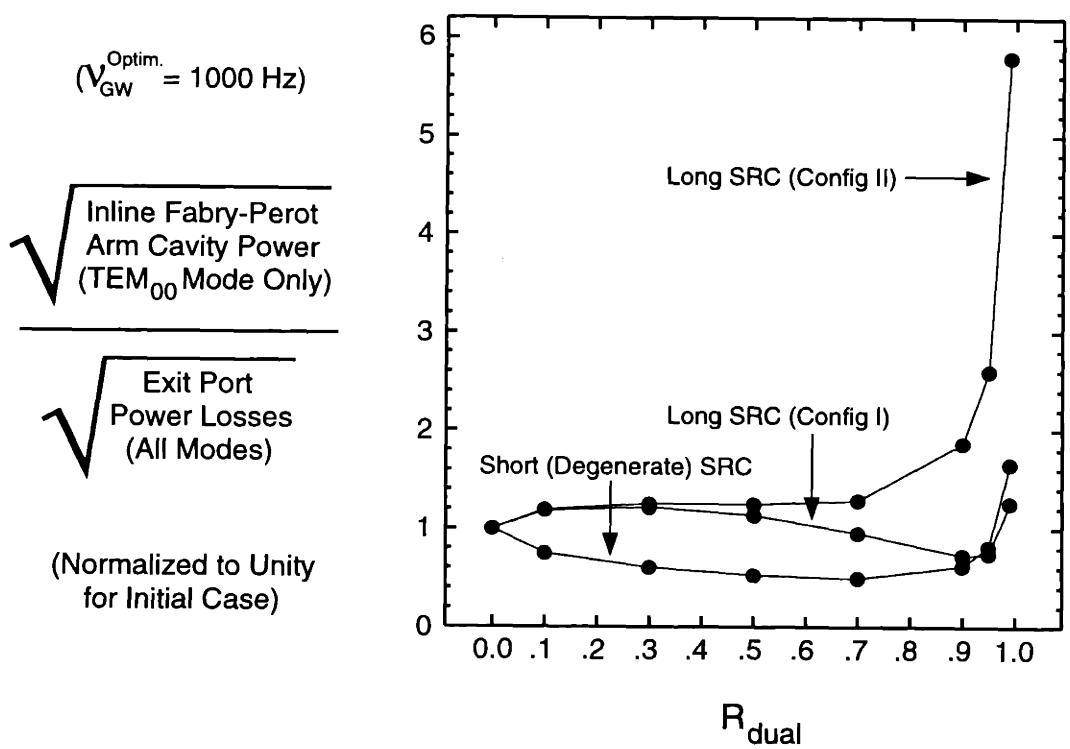

Figure 4.14: The ratio of (TEM_{00} mode) e-field amplitude in the inline Fabry-Perot arm cavity to the exit port e-field amplitude (in all modes), is plotted versus Signal Recycling Mirror reflectivity, for different SRC configurations. The lambda/800 family of deformed mirror maps has been used. This figure demonstrates the behavior of a function which is proportional to the signal-to-shot-noise of the interferometer, for different SRC-degeneracy conditions.

The principle result from Figure 4.14 is quite clear: the wavefront healing effects of Dual Recycling can indeed increase the sensitivity of the detector by a substantial factor — mostly due to the reduction of shot noise that is produced by the exit-port power — but *only* when the Signal Recycling Cavity has its degeneracy broken, and in careful fashion (i.e., avoiding the accidental resonance of individual modes). This conclusion, plus the possibility mentioned above for making both the plus *and* minus GW-sidebands simultaneously resonant within an SRC that is sufficiently long, and which has a sufficiently-high tuning frequency, makes a long-baseline Signal Recycling Cavity appear to be a promising form of Dual Recycling for future use in Advanced-LIGO interferometers.

Chapter 5

Conclusions and future directions

We now conclude this thesis by highlighting the principle results of our research, by providing general recommendations based on these results, and by proposing a number of options for future research with our simulation program, either as-is or with useful extensions of the main Fortran codes.

5.1 Summary of Initial- and Advanced-LIGO configuration results, and general recommendations

First, we consider the creation of the code itself, and the results that were obtained for simulations of the Initial-LIGO interferometer configuration:

- We have succeeded in constructing a simulation program for a full-LIGO interferometer that possesses optics with highly-detailed, "realistic" mirror deformations. Many different interferometer parameters or imperfections can be modelled, and the program computes the relaxed electric fields everywhere in a static interferometer which is fully-optimized for Gravitational Wave detection, in an affordable amount of run-time.
- Mirror surface roughness deformations have the most dramatic effect upon interferometer power buildup and performance (compared to, e.g., substrate inhomogeneities, "geometric" imperfections such as tilts or curvature errors, or straight losses such as absorption) because of loss and high-mode creation due to large-angle beam scattering.
- The Initial-LIGO performance requirements appear to be achievable with realistically producible optics. Our simulation runs with mirror surface deformations that possess rms values of $\lambda_{YAG}/800$ or better succeeded at passing all of the quantitative tests that we have considered, and the production of mirrors with rms values (in the center of the optic) of $\lambda_{YAG}/1200$ or better is feasible with current mirror-making technology [81]. This provides validation for the mirror specifications from the "Pathfinder Project" [37].
- Using optics with polished (and coated) surfaces of the highest possible quality does indeed yield a significant improvement in LIGO science capabilities, and would perhaps make the difference in giving interferometers with the Initial-LIGO configuration the ability to detect even a few of the most favorable (for LIGO) Gravitational Wave candidates, such as coalescing Black Hole/Black Hole binaries, and non-axisymmetric Pulsars (as well as increasing the detection rate with *Advanced*-LIGO interferometer configurations).

In summary, the outlook for LIGO performance with realistically achievable optics is very optimistic, and further improvements in the production of LIGO mirrors could improve the science output of LIGO even further.

Now, considering what our results have demonstrated about the Advanced-LIGO configuration known as Dual Recycling:

- Gravitational Wave response functions with moderate narrowbanding (e.g., $\Delta\nu_{FWHM}$ = 50 Hz for a Dual-Recycled system tuned to 200 Hz) can be obtained even with high levels of optical deformations, though ultra-narrowbanding (e.g., $\Delta\nu_{FWHM}$ = 8 Hz for a system tuned to 200 Hz), does significantly increase the loss of sensitivity for these deformed-mirror runs.

- Tuned Dual Recycling with a high-finesse Signal Recycling Cavity can successfully produce a deep "notch" in the interferometer's shot-noise-limited sensitivity curve; though much of this narrowbanded improvement in sensitivity is blunted by an underlying shelf of mirror internal thermal noise, especially so for a system specified to be optimized for low or moderate Gravitational Wave tuning frequencies.

- The incorporation of a Signal Recycling Mirror does (barring exceptions to be discussed below) significantly reduce the amount of power lost through the interferometer exit port, thus making it more tolerant for deformations which predominantly act by channelling power in perturbed beam modes out through the beamsplitter dark fringe. This property, however, is less significant for "realistic" mirror deformations than for "geometric" deformations (e.g., mirror tilts, curvature mismatch errors), since realistic deformations predominantly act by scattering power completely out of the system.

- The phenomenon known as "Wavefront Healing" is effective at recycling beamsplitter dark-fringe light back into the system, for re-integration as useful (i.e., TEM_{00} mode) power in the Fabry-Perot arm cavities, then to be used for measuring the effects of Gravitational Waves. This has very little effect upon improving the interferometer sensitivity for runs with realistically deformed optics, however, since exit-port losses are insignificant compared to high-angle scattering losses during arm cavity propagations.

- The problem of high-angle scattering losses (and the corresponding lack of effectiveness of wavefront healing) is not helped by the expedient of making the aperture sizes of the finite-mirrors larger, since this results in resonantly-enhanced exit port losses for high-index beam modes which are accidentally resonant in the system, and whose large spatial profiles can now be supported by the larger interferometer mirrors.

- Broadband Dual Recycling (i.e., a system tuned to Gravitational Wave signals at DC) is, as predicted, effective at the suppression of exit-port power loss in all modes except for the TEM_{00} mode, for which it amplifies the losses. Conversely, Narrowband Dual Recycling set to high tuning frequencies (~1 kHz or higher) is effective at the suppression of TEM_{00} mode losses, but losses in all other modes (assuming their non-resonance in the Fabry-Perot arm cavities) are resonantly amplified by the Signal Recycling Mirror. The across-the-board, non-TEM_{00}-mode effects for both of these systems (which we call "Wavefront Harming" for the Tuned Dual Recycling case) occurs because of the modal degeneracy of the Signal Recycling Cavity.

- Wavefront Harming is a deleterious and much more significant effect than the beneficial behavior known as Wavefront Healing, and this Harming effect therefore has to be eliminated by breaking the degeneracy of the Signal Recycling Cavity[1]. We have demonstrated the success of degeneracy-breaking via the implementation of a long-baseline Signal Recycling Cavity, which has the possible added benefit of extra sensitivity compared to the short-baseline case, due to the simultaneous resonance of both "plus" and "minus" Gravitational-Wave-induced sidebands in the Dual-Recycled system.

In summary, Dual Recycling can indeed be an effective system for obtaining sharp sensitivity peaks over narrow ranges in Gravitational Wave frequency space, as well as for significantly reducing the losses through the exit port of the interferometer; both of these properties, however, require the breaking of the Signal Recycling Cavity degeneracy.

On the other hand, the benefits of Dual Recycling are severely constrained by limitations of the optics. The barrier caused by mirror internal thermal noise renders much of the sensitivity peak of Narrowbanded Dual Recycling system useless, and large-angle scattering from high-spatial-frequency mirror surface deformations creates large losses that cannot be repaired significantly by Wavefront Healing. We are therefore left to conclude that Dual Recycling does *not* alleviate the need for the highest quality mirrors that can be produced; on the contrary, it is a configuration for which excellent quality mirrors are even more important, in order to have a system that performs up to its full potential. The strongest recommendation of this thesis must therefore be the following: the best path for obtaining superior interferometer (shot-noise-limited and overall) performance is to *make better mirrors*, especially in terms of fine-scale (i.e., high-spatial-frequency) deformations, and in terms of optics with higher mechanical Q-factors for lower internal thermal noise.

Despite these caveats, Dual Recycling remains a very promising configuration for implementation in Advanced-LIGO interferometers, especially due to its ability to narrowband and tune the Gravitational Wave response of these future-generation detectors.

1. In fact, it is possible that breaking the degeneracy in the Power Recycling Cavity (not a currently planned feature of the Initial-LIGO detector) may be an equally important step for preventing the oscillation of non-TEM_{00} modes in the Michelson (pre-cavity) arms of the interferometers.

5.2 Directions for future simulation-based research

Given the physical realism of the interferometer simulation program, along with the effectiveness of its electric field relaxation and interferometer optimization algorithms, there are a great many research questions that one may address with it. Certain investigations (such as the addition of more mirrors to the system) would naturally require modifications or extensions to the Fortran code, but on the whole, many topics of interest in the research and development of interferometers for Gravitational Wave detection should be reasonably amenable for study with this simulational tool.

We therefore end the body of this thesis with a list of possibilities for future numerical investigations using current or updated versions of our LIGO simulation program:

- Evaluations of system performance with the actual optics that will be purchased by the LIGO Project, using measurement maps of the deformations of the delivered mirrors. This should be a major part of the optics validation program for the LIGO Project.
- Development and evaluation of a Gravitational Wave signal measurement and control system for the full-LIGO, Dual-Recycled interferometer configuration.
- Studies of thermal lensing [82] due to absorption-induced mirror heating, perhaps including the nonlinear feedback effects of lensing upon the level of resonating interferometer power, and this power (and thus, absorption) level back upon the lensing.
- Studies of Resonant Sideband Extraction (RSE) [83], a configuration equivalent to Dual Recycling, though with an exactly resonant Signal Recycling Cavity. RSE lowers the storage time of Gravitational Wave signals, thus allowing the use of higher-finesse Fabry-Perot arm cavities, and lowering the power level in the Power Recycling Cavity. This reduces the power passing through the beamsplitter, reducing its thermal lensing.
- Evaluations of other promising Advanced-LIGO configurations, such as Compound Dual Recycling [74], or the implementation of a Sagnac interferometer [84].
- Simulations with realistically deformed mirrors in support of other Gravitational Wave detection collaborations around the world, such as GEO-600 [85], which plans to use Dual Recycling (without Fabry-Perot arm cavities) in their initial-stage detectors.
- Any investigations for the LIGO Project, involving a variety of optical parameters and/or optical imperfections, which may become high-priority issues for the LIGO mission.

In conclusion, the LIGO simulation program that we introduce in this thesis has been designed to be a modelling tool that is powerful, physically realistic, and as flexible as possible, given the highly detailed nature of its computational tasks.

Appendix A

Modal analysis of interferometer laser fields

An important part of the analysis of interferometer simulation results involves the *modal decomposition* of the electric fields that are computed by the main simulation program. In addition to direct reports by the simulation program giving a little bit of information about the division of power for each simulated e-field (such as the total e-field power, the power in the first one or two transverse beam modes, etc.), complete 2-D "slices" (each transverse to the beam propagation axis) representing the structure of several interferometer e-fields of interest are written out to files. These files can be read in for post-processing analysis by a small Fortran program that we have written, called "modal_decomp.f" (or some variant of this name), which analyzes each e-field into a large number of independent beam modes. This modal decomposition program is also grid-based, as is the main simulation program, and represents its electric fields on complex, pixellized grids in the same fashion.

The modes which we use for the tilt removal process itself are the Hermite-Gaussian TEM modes, defined as follows [17]:

$$\mathrm{TEM}_{MN}(x,y,z) \equiv \sqrt{\frac{1}{2^{M+N-1}\pi M!N!}} \times \frac{1}{w(z)} \times H_M\left(\frac{\sqrt{2}x}{w(z)}\right) H_N\left(\frac{\sqrt{2}y}{w(z)}\right) \qquad \text{(A.1)}$$

$$\times \mathrm{Exp}\left[-\frac{(x^2+y^2)}{w^2(z)} - \frac{ik(x^2+y^2)}{2R(z)}\right] \times \mathrm{Exp}[-ikz + i(M+N+1)\psi(z)]$$

where z is the longitudinal distance (along the propagation axis) from the beam waist position, and the x and y axes are transverse to the beam (i.e., they define the plane of our 2-D beam slices). The laser wavelength is λ, and the wavenumber is $k = 2\pi/\lambda$. The beam spot size at a given point along the propagation axis is equal to $w(z) = w_0\sqrt{1+(z/Z_R)^2}$, with the "Rayleigh Range" being defined as $Z_R \equiv \pi w_0^2/\lambda$, and with the beam waist (i.e., the spot size at the focus) being w_0. The spot size determines the Gaussian fall-off of power away from the center axis of the beam, and affects the overall beam normalization, as well as the spatial scale of the structure (including the spacing of beam nodes) for each mode. This spatial structure (other than the Gaussian envelope) is determined by the factors $H_M(f[x])$, $H_N(f[y])$ in Eq. A.1, where $H_M(x)$ is the M^{th} Hermite polynomial, with the first few such polynomials being:

$$H_0(x) = 1, \quad H_1(x) = 2x, \quad H_2(x) = 4x^2 - 2, \ \ldots \qquad \text{(A.2)}$$

The beam radius of curvature is given as $R(z) = z + Z_R^2/z$, and this curvature of the

beam at a given interferometer location must match the curvature of the mirror at that location, or some conversion of power from the fundamental mode (TEM_{00}) to higher modes (especially TEM_{20} and TEM_{02}) will occur, and the presence of these modes will be detected by the modal decomposition program. Lastly, the "Guoy Phase" for the MN^{th} mode is given by $(M + N + 1/2)\,\psi(z) \equiv (M + N + 1)\,\mathrm{atan}(z/Z_R)$, and this is the phase which causes different modes to experience different resonance conditions (c.f. Ch. 4, especially Sec. 4.3.2) in a cavity with strong focusing — i.e., a "non-degenerate" cavity.

The Hermite-Gaussian modes that have been defined here are orthogonal (and more precisely, orthonormal) to one another, due to the mutual orthogonality (with the weighting function, $\mathrm{Exp}[-2(x^2 + y^2)/w^2(z)]$) of the Hermite polynomials shown in Eq. A.1 [50].

The modal decomposition program operates by reading in an electric field, $\vec{E}$, as well as the beam spot size and radius of curvature at that e-field's interferometer location (plus a flag indicating whether the beam is approaching or moving away from a focus). It then takes the overlap integral (as per Eq. 2.68) of that e-field with a selected TEM mode, as follows:

$$A_{MN} = \langle \vec{E} \,|\, TEM_{MN}(x, y, z) \rangle \Big|_{z\,=\,z_{\text{E-field}}} \tag{A.3}$$

This procedure is repeated for each mode on a list that is read in by the modal decomposition program. The real and imaginary parts of the overlap coefficient defined by Eq. A.3 for each mode are reported as the output of the decomposition program, as is the total power in the mode, given by $|A_{MN}|^2$. The program also reports the total power (in all modes) of the e-field, the two highest-power modes in it, and one more quantity of importance: the left-over power that is *unaccounted for* by any of the modes in the list for decomposition. This "unaccounted for" power is a key sign of the existence of *position-space aliasing*, as described in Section 2.3.4.3.

To conclude this discussion, we note that just as the discrete grid in the main simulation program must have sufficiently high resolution to sample the Gaussian laser beam properly (c.f. Sec. 2.3.4.2) — that is, without spatial-frequency-space aliasing — the modal decomposition program must also have a high enough resolution to model the highest modes in the decomposition list (since higher-index Hermite polynomials cause finer beam amplitude fluctuations). If the mathematical functions representing these highest modes are not sampled properly, then the decomposition modes will be aliased, causing a breakdown of orthogonality between discretized modes, and the miscounting (and multiple-counting) of power in the various modes. To avoid these problems, we have performed orthogonality tests to verify that for the gridding resolution currently used in the decomposition program, and given the highest modal indices that we have coded for use in the program (specifically, M=N=15), that every mode which can be activated for e-field decomposition is represented accurately on the grid.

Appendix B

Interferometer optics: mirrors for simulation runs, and LIGO's state of the art

In Section 2.3.3.4, we described the creation of "realistic deformation" mirror maps from measurements of the reflection from a polished surface ("Calflat"), and the transmission through a fused-silica substrate ("Corning"). These initial surface (and substrate) measurement maps were used as sources for the generation of a large number of maps with the same deformation power spectra. The surface maps were then re-scaled to produce several families of mirrors, where each family is labelled by its rms deformation amplitude (expressed roughly in wavelengths of Nd:YAG laser light, i.e., 1.064 μm) when sampled over regions of fixed size near the center of the optic. These new families of mirror surfaces — called λ/1800, λ/1200, λ/800 and λ/400 — were used in conjunction with the substrate deformation maps to comprise a *standard set* of mirrors for all of our runs, including the Initial-LIGO simulation runs described in Chapter 3, and the Dual Recycling runs described in Chapter 4.

We must address the question of whether or not these mirror maps (most importantly, the Calflat mirror surfaces) are good representations of the mirrors that LIGO will eventually procure from industry; not to mention whether or not mirrors currently being obtained by LIGO appear to be adequate, given what we have learned via our simulation runs.

Some liberties have clearly been taken to create our standard set of mirror maps, since real mirrors will differ in their power spectra, but all of our surface maps share the same power spectrum (as do all of the substrate maps, respectively). Regarding that issue, the evaluation of interferometer performance for mirrors possessing different power spectra has not been a significant part of this thesis.

On the whole, however, we can say that the lessons learned from the runs performed for this thesis are fairly general in scope, because the Calflat (and Corning) power spectra possess three qualities which make them representative of "real LIGO optics": their rms deformation levels are of the right approximate magnitude, their power spectra decrease fairly steadily with increasing spatial frequency k (roughly like $1/k$, bumps and wiggles in the spectra aside), yet there is significant deformation power all the way up to the *highest spatial frequencies* representable on our two-dimensional grids (c.f. Sec. 2.3.4.2). This last point is especially important, since it means that the *scattering losses* in the long-baseline arm cavities have been the most significant effect of the incorporation of the Calflat and Corning maps, which will likely be true for real optics that possess deformations at all spatial frequencies, with the effects of the fine-scale roughness dominating system behavior and losses.

We can quantify the comparison of the polished surfaces of the optics that have recently been procured by the LIGO project, with the requirements listed in the Design Requirements Review ("DRR") for the Core Optical Components [86], as well as with the specifications that were given for polishing. Consolidating the results for several "End Test Masses" (i.e., Fabry-Perot arm cavity back mirrors), we get (Billingsley, *et al.*, [81]):

Deformations Sampled (spatial frequencies given by $\lambda_s = 2\pi/k$)	DRR Requirements	Polishing Specifications	Results to date
RMS surface errors for beam waist > λ_s > 2.3 mm, out to r = beam waist radius	$\lambda_{YAG}/1200$ (0.89 nm)	0.8 nm	0.5 nm maximum
RMS surface errors for 2 x beam waist > λ_s > 2.3 mm, out to r = 2 x beam waist radius	$\lambda_{YAG}/600$ (1.77 nm)	0.8 nm	0.5 – 0.8 nm
RMS surface errors for λ_s > 3–4 x beam waist	$\lambda_{YAG}/160$ (6.65 nm)	1.6 nm	0.9 nm

Table B.1: Comparing the surface deformations of procured LIGO optics with Project requirements and polishing specifications, for different spatial frequency regimes and mirror sampling regions.

The results in Table B.1 show that the polished mirror surfaces do satisfy these important surface roughness requirements (in fact, they are significantly better), as well as demonstrating that our Calflat-derived simulation maps have the right order-of-magnitude deformations for relevance to the performance of real LIGO optics.

This success of having met Project requirements with the polished surfaces, however, must be tempered with the recognition that the application of mirror coatings may make the results worse. In particular, for two arm cavity back mirrors labelled "p05ca" and "p06ca", it is found that the coatings increase the rms (over the full 14.6 cm mirror apertures) from 0.38 nm to 0.51 nm for p05ca, and from 0.37 nm to 0.67 nm for p06ca. These rms increases are in themselves significant (though not devastating), but the dependence of the addition deformations upon spatial frequency is an even more important issue. Figure B.1 shows the power spectrum of one of these polished mirrors before (p05ca) and after (p05cac) being coated. In this figure, the power spectrum (in "waves" of 6328 Angstroms rather than meters, per unit wavenumber) is plotted versus the log of the spatial-frequency wavenumber in inverse centimeters. On this abscissa, the size of one pixel (2.77 mm) corresponds to a value of 0.56, while the full mirror diameter of 24 cm corresponds to a value of -1.4. The addition of mirror deformation structure on p05ca due to its reflectivity coat-

ing is clearly concentrated in the region of the highest spatial frequencies that are representable on the grid, which as we have seen are very important because they lead to large scattering losses in the long-baseline arm cavities. Even if this is just a quirk of this particular optic, and not a general property of mirror coatings (indeed, such a prominent bulge at these spatial frequencies does not happen for the coated version of p06ca), we see that one must be very cautious in the process of deciding when LIGO optics are declared to be within their specified tolerances, and that the full simulation-based modelling of interferometer performance with measurements of the final, coated mirrors will be a very important procedure for LIGO optics validation.

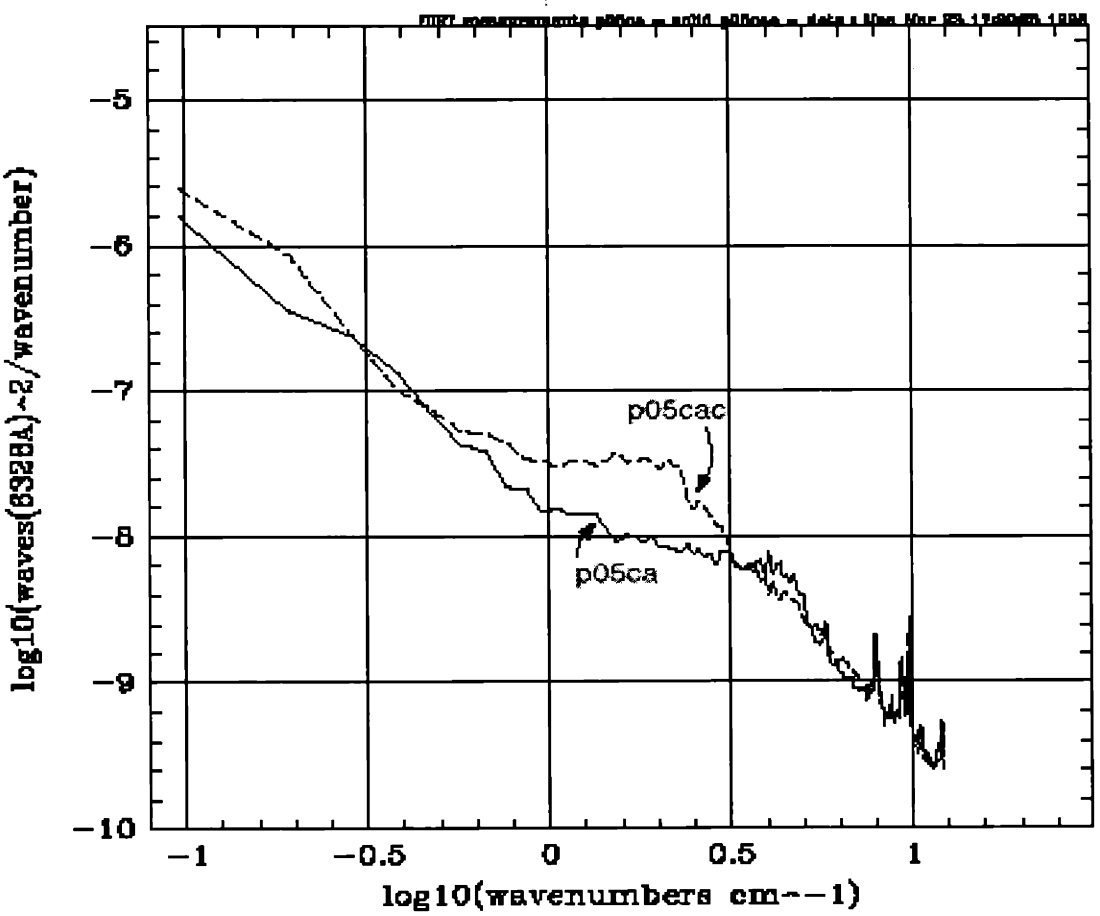

Figure B.1: Power spectra for a polished LIGO mirror with (p05cac) and without (p05ca) a high-reflectivity coating.

Appendix C

Tilt-removal for mirrors with realistic deformations

In this appendix, we discuss a pre-processing mirror "optimization" step that is performed upon the mirror profile maps (by a small Fortran program, "mirror_bounce.f", which we have written) before they are read in during execution of the simulation program[1]. This procedure, referred to earlier in the section on the implementation of real-world mirror measurement data (Sec. 2.3.3.4), is the removal of the inherent tilts from each of the mirror deformation profile maps[2].

For perfectly smooth optical profiles, the question of alignment is uniquely defined, and we have the ability to perfectly designate the misalignment values of a given mirror with respect to the laser beam propagation axis. But for mirrors possessing surface or substrate deformations, it is more difficult to specify the normal axis for the mirror surface (or substrate[3]) maps, and we must find a good definition for the orientation that is equivalent to "zero" tilt.

For irregular surface profiles, which do not possess a unique definition of "perfect" or "tilted" mirror alignment, we choose the definition which is most useful for a LIGO interferometer: it is the tilt which is experienced (or "weighted") by a Gaussian-profile incident beam. It can be shown [18] that to first order, for an incident beam which is predominantly in the Hermite-Gaussian [17] TEM_{00} mode, the effect of tilts about the two axes perpendicular to the beam is to generate TEM_{10} and TEM_{01} modes (with pure imaginary coefficients, for a purely real, incident TEM_{00} beam[4]) in a reflection from that mirror. We can therefore remove the beam-weighted tilts from a given mirror deformation map (which possesses "height" function $z = M(x, y)$; c.f. Sections 2.3.3.1 and 2.3.4.1), by performing small angular corrections that null the imaginary parts of the TEM_{10} and TEM_{01} overlap coefficients, as follows:

---

1. The procedure outlined here is not performed actively during program run-time (as are, e.g., many of the optimizations discussed in Sec. 2.5), since too many simultaneous optimizers would interfere with one another, and could lead to instability in the overall convergence process.
2. We refer here to the removal of the small tilts that are initially present in profiles of irregular or deformed mirrors. During a run, any amount of precisely specified tilt can be added as input data, if the effects of mirror tilt on interferometer performance are specifically being studied.
3. "Tilts" for mirror *substrate* deformation maps really represent mirror thickness wedges, and the primary effect of these exceedingly small wedges is merely to alter the reference axes of the various cavities, an effect which is nulled in real LIGO interferometers via automatic mirror angular alignment [38].
4. Before tilt-removal a mirror must first be "piston-removed", in order to make the incident TEM_{00} beam purely real. This is accomplished by substituting TEM_{00} for TEM_{10} or TEM_{01} in Eq. C.1, and by performing mirror map position adjustments parallel to the laser beam axis.

$$\mathrm{Im}\langle \mathrm{TEM}_{10} | \{\mathrm{Exp}[-2ikM(x,y)]\} | \mathrm{TEM}_{00}\rangle, \qquad \text{(C.1)}$$
$$\mathrm{Im}\langle \mathrm{TEM}_{01} | \{\mathrm{Exp}[-2ikM(x,y)]\} | \mathrm{TEM}_{00}\rangle \rightarrow 0$$

In this formula, the exponential refers the value of each pixel, not to an exponentiation of the matrix; and the overlap integrals (actually discrete sums) are as defined in Eq. 2.68.

Physically, Eq. C.1 is equivalent to taking the initial beam, a pure TEM_{00}, and reflecting it from a map of the deformed mirror being studied; this reflection map is subsequently overlapped against the TEM_{10} and TEM_{01} modes, in order to measure the beam-weighted tilts which must be removed by applying permanent phase-gradient corrections to the map about the axis of each measured tilt angle.

We note here that by using Gaussian functions for a weighted tilt-removal procedure, the mirror map *as a whole* will be given an overall tilt by the prescription of Eq. C.1. But the more important fact is that the *center* of the mirror (i.e., the part most sampled by the collimated, Gaussian beam) is made to be as flat as possible with respect to the plane transverse to the beam propagation axis.

In addition, we note that in a technical sense, tilts about different axes do not "commute", meaning that correcting the tilt about, e.g., the x-axis, alters the remaining amount of tilt about the y-axis. That is, if a pure tilt of angle θ about the y-axis is represented by the height function $M(x, y) \approx x\sin\theta$, and a pure tilt of angle ϕ about the x-axis is represented by $M(x, y) \approx y\sin\phi$, then the x-axis tilt *followed* by the y-axis tilt would produce the height function:

$$M(x, y) \approx x\sin\theta + y\sin\phi\cos\theta \qquad \text{(C.2)}$$

This cross-talk behavior between tilts is very small, however, since it multiplies the second tilt by the cosine of the (always very small) first tilt adjustment; and the effect could therefore either be neglected. Alternatively, as is done in our tilt-removal code, it can be taken into account very simply by using the mirror overlap integral results indicated in Eq. C.1 to first solve for θ, and then to use θ (and those overlap coefficients) to solve for ϕ. The two tilts would then be removed sequentially, in the proper order.

The sole relevant drawback to this general procedure is a minor matter of inconvenience: since particular modal functions are being used for mirror tilt removal, a spot size[1] must be chosen for the TEM modes used in Eq. C.1 that adequately represents the e-fields which will interact with that mirror. Therefore, if one wants to use a given deformation map on a new mirror in a different interferometer location (where the beam spot size is different), or if the interferometer configuration parameters are changed in a way that significantly alters the beam spot size at the location of that same mirror, then the tilt-removal process has to be repeated for the deformation map, using the new spot size.

1. Unlike the spot size, it is not necessary to specify the *radius of curvature* (see Appendix A) of the mirror that uses a particular deformation map, since it is assumed that the beam will be properly matched into the interferometer, and will always have the same curvature as the local mirror.

Appendix D

Calculation of the shot noise sensitivity limit for an Initial-LIGO interferometer

Here we present the full calculation of the shot-noise-limited sensitivity curve of a LIGO interferometer. As shown below, these results are not just for use in cases with idealized interferometers and perfect mirrors, but can also be used to compute the noise curve for runs with optical deformations, by incorporating the results from our FFT-based numerical simulation program.

D.1 Deriving the Gravitational Wave signal

For these calculations, we will consider a monochromatic, transverse-traceless (TT) Gravitational Wave (GW) arriving with optimal direction and polarization with respect to the interferometric detector. If the inline and offline interferometer arms are aligned (respectively) along the x- and y-axes, then the GW is a perturbation to the flat-spacetime Minkowski metric of the form [2, 25]:

$$[h^{TT}_{\mu\nu}(t, x, y, z)]_{z=0} = \sqrt{2}\, h_{ij} \cos(\Omega t + \phi_0) \tag{D.1}$$

where $h_{ij} = diag(0, h, -h, 0)$, and $|h|$ is the rms amplitude of the GW. The GW angular frequency is given by $\Omega = 2\pi f$, and ϕ_0 is its initial phase at t=0, z=0.

The entire effect of the GW is treated as a modulation of the electric fields (e-fields) resonating in the Fabry-Perot arm cavities, where a large amount of carrier power interacts with the GW's over a long round-trip path length (8 km). We follow the formalism of Vinet, *et al.* [25], and compute the creation of GW-induced sidebands at $\omega \pm \Omega$ from the circulating carrier frequency e-fields, originally at $\omega = 2\pi\nu_{Carr}$. If the initial e-field amplitudes exciting each of the (exactly resonant) arms is given by $E_0\, e^{-i\omega t}$, then to first order in h, the total e-fields reflected from the arm cavities will be[1]:

$$E_R = E_0\, e^{-i\omega t} \times \left\{ A_0 + h A_1\, e^{i(\Omega t + \phi_0)} + h A_2\, e^{-i(\Omega t + \phi_0)} \right\} \tag{D.2}$$

where:

---

1. For this calculation we use Vinet's sign convention of, e.g., $e^{-i\omega t}$ for temporal phase factors, in place of our regular convention of, e.g., $e^{+i\omega t}$.

$$A_0 = \frac{-(T_1 + R_1)\, r_2 + r_1}{1 - r_1 r_2} \tag{D.3}$$

$$A_1 = \frac{i}{\sqrt{2}} \varepsilon T_1 r_2 \left(\frac{\omega}{\Omega}\right) \sin\left(\frac{\Omega L}{c}\right) \times \frac{e^{-i\Omega L/c}}{(1 - r_1 r_2)(1 - r_1 r_2 e^{-2i\Omega L/c})} \tag{D.4}$$

$$A_2 = \frac{i}{\sqrt{2}} \varepsilon T_1 r_2 \left(\frac{\omega}{\Omega}\right) \sin\left(\frac{\Omega L}{c}\right) \times \frac{e^{i\Omega L/c}}{(1 - r_1 r_2)(1 - r_1 r_2 e^{2i\Omega L/c})} \tag{D.5}$$

In these formulas, $R_1 = r_1^2$ and $T_1 = t_1^2$ are, respectively, the power reflectivity and transmission of the arm cavity input mirrors, $R_2 = r_2^2$ is the power reflectivity of the arm cavity back mirrors, and L is the unperturbed length of the arms (4 km). Note that A_0 in Eq. D.3 is just the normal coefficient of reflection from a resonant cavity, while A_1 and A_2 (when multiplied by $h \cdot E_0$) represent the GW-induced signal e-fields of interest. The sign factor of $\varepsilon = +1$ (-1) for the inline (offline) arm path ensures that these signal e-fields will constructively interfere at the beamsplitter exit port (where the carrier is at a dark fringe), and maximally emerge from the interferometer to generate the GW-signal.

We now make the following approximation, often made for interferometric GW detection:

$$\frac{\Omega L}{c} \ll 1 \tag{D.6}$$

so that $\sin(\Omega L/c) \approx \Omega L/c$, and $\cos(2\Omega L/c) \approx 1 - (2\Omega L/c)^2/2$. This approximation with be very good for GW-frequencies of interest ($f \approx 10^2 \rightarrow$ few $\times\, 10^3$ Hz). Furthermore, we define an (analytical) cavity storage time, as follows:

$$\tau_s^{(a)} \equiv \frac{\sqrt{r_1 r_2}}{1 - r_1 r_2} \left(\frac{L}{c}\right) \tag{D.7}$$

Finally, taking the absolute values of the GW-induced e-field amplitudes, we use Eq's. D.6 and D.7 in Eq's. D.4, D.5 to get:

$$|A_1| = |A_2| \approx \frac{T_1 \sqrt{r_2/r_1}}{1 - r_1 r_2} \frac{\sqrt{2}\, \pi \nu_{Carr} \tau_s^{(a)}}{\sqrt{1 + (4\pi \tau_s^{(a)} f)^2}} \tag{D.8}$$

Thus, the GW-frequency response of the Fabry-Perot arm cavities is that of a simple pole with $f_{pole} = 1/(4\pi\tau_s^{(a)})$, and with an overall amplitude determined by the optical parameters of the system, such as the carrier frequency, mirror reflectivities, and the

resulting arm cavity storage time (though note that the direct dependence of $|A_1|$ and $|A_2|$ on $\tau_s^{(a)}$ cancels out for $f \gg f_{pole}$).

So far, these formulas have been dependent upon analytical values for several optical parameters of the system (r_1, r_2, etc.). This is not satisfactory for modelling the buildup of power in cavities with deformed mirrors, and the results of our simulation program must therefore be inserted in their place to model the impact of the various imperfections being studied. First of all, we note that the amount of power resonating in the Fabry-Perot arm cavities can be written as:

$$|E_{Carr,\,FP\text{-}arm}| = |E_0| \cdot \frac{t_1}{1 - r_1 r_2} \tag{D.9}$$

We can therefore re-write Eq. D.8 (now including the multiplication factor, $|E_0|$, from Eq. D.2), as:

$$|E_0| \cdot |A_1| = |E_0| \cdot |A_2| \approx t_1 \sqrt{\frac{r_2}{r_1}} \cdot |E_{Carr,\,FP\text{-}arm}| \frac{\sqrt{2}\,\pi \nu_{Carr} \tau_s^{(a)}}{\sqrt{1 + (4\pi\tau_s^{(a)} f)^2}} \tag{D.10}$$

Next, we replace the analytical cavity storage time with a more realistic, "effective" storage time, τ_s, that is derived from the amplitudes of the program's simulated e-fields. Starting from the definition in Eq. D.7, and utilizing Eq. D.9, we may write:

$$\tau_s = \frac{\sqrt{r_1 r_2}}{1 - r_1 r_2}\left(\frac{L}{c}\right) \approx \frac{1}{1 - r_1 r_2}\left(\frac{L}{c}\right) = \frac{|E_{Carr,\,FP\text{-}arm}|}{t_1 \cdot |E_0|}\left(\frac{L}{c}\right) \tag{D.11}$$

where we have used the approximation:

$$\sqrt{r_1 r_2} \approx 1 \tag{D.12}$$

which is good to ~98-99% accuracy, as long as we don't use it in the denominator of, e.g., Eq. D.9, where small inaccuracies cause large errors in the calculation of resonant power buildup.

We now simplify Eq. D.11 by eliminating $|E_0|$, as follows:

$$\begin{aligned} |E_0| &= |E_{Carr,\,Recyc.\,Cav.}| \cdot t_{BS}^{inline} \quad \text{(inline Fabry-Perot arm)}, \\ |E_0| &= |E_{Carr,\,Recyc.\,Cav.}| \cdot r_{BS}^{ref\text{-}side} \quad \text{(offline Fabry-Perot arm)} \end{aligned} \tag{D.13}$$

where $r_{BS}^{ref\text{-}side} = \sqrt{R_{BS}^{ref\text{-}side}}$ and $t_{BS}^{inline} = \sqrt{T_{BS}^{inline}}$ are the amplitude reflectivity and transmission values for the beamsplitter, with the "ref-side" and "inline" superscripts being chosen in accordance with the particular beamsplitter ports that the relevant e-field has interacted with, as per Figure 2.9. Plugging D.13 into D.11, and *averaging* the results for

the two arms (which may differ slightly due to unbalanced imperfections), we get:

$$\tau_s \cong \frac{L}{c} \cdot \frac{1}{\left|E_{\text{Carr, Recyc. Cav.}}\right|} \cdot \frac{1}{2}\left\{\frac{\left|E_{\text{Carr, inline FP-arm}}\right|}{t_{BS}^{\text{inline}} \cdot t_1} + \frac{\left|E_{\text{Carr, offline FP-arm}}\right|}{r_{BS}^{\text{ref-side}} \cdot t_1}\right\} \qquad \text{(D.14)}$$

Note that we use — with acceptable accuracy — analytical parameters such as t_1, etc., in these formulas when only *one trip* is represented, such as one reflection or one transmission through a mirror. Only when "many bounces" are involved, such as for the power buildup in a cavity, do the effects of mirror imperfections become very important, requiring us to use numerically simulated numbers.

We recall that Eq. D.10 represents the GW-induced signal e-fields that are generated *in reflection* from the Fabry-Perot arm cavities, which are being perturbed at GW-frequency *f*; this can be made more apparent if we restore the factor of $\sqrt{r_1 r_2}$ that was dropped from the overall amplitudes of the e-fields by approximation D.12, and rename r_2 and t_1 as, respectively, $r_{\text{FP arm back}}$ and $t_{\text{FP arm input}}$. Then Eq. D.10 becomes:

$$\left|E_0\right| \cdot \left|A_1\right| = \left|E_0\right| \cdot \left|A_2\right| \approx \qquad \text{(D.15)}$$

$$\left(\left|E_{\text{Carr, FP-arm}}\right| \cdot r_{\text{FP arm back}} \cdot t_{\text{FP arm input}}\right) \times \frac{\sqrt{2}\,\pi\,\nu_{\text{Carr}}\tau_s}{\sqrt{1 + (4\pi\tau_s f)^2}}$$

Intuitively speaking, this formula clearly represents the extraction of signal from the "reservoirs" of power in the arm cavities by the perturbing force of the Gravitational Waves, with the effects of signal amplification for long arm cavity storage times and narrowbanding due to the cavity pole frequency being quite apparent.

These (upper & lower GW-sideband, inline & offline arm) reflection e-fields must be brought to the exit port of the interferometer beamsplitter, so that they can be beaten against the local oscillator (radio frequency) sideband fields in order to generate the GW-signal. This entails, for the one-time trip through the beamsplitter, a multiplication by the factor $r_{BS}^{\text{A.R.-side}}$ (t_{BS}^{offline}) for the inline (offline) arm e-fields, followed by their constructive interference at the exit port. We then extract the TEM_{00} components from these fields (the non-TEM_{00} components will contribute insignificantly, and in a difficult-to-calculate way, to the total GW-signal), and re-write the field amplitudes (such as $\left|E_{\text{Carr, FP-arm}}\right|$) in terms of square-roots of the output powers generated by our simulation program. Putting this all together, we can write the total GW-induced signal e-field at the beamsplitter exit port as:

$$E_{\text{exit}}^{GW} \cong h e^{-i\omega t} \times \left\{e^{i(\Omega t + \phi_1)} + e^{-i(\Omega t + \phi_2)}\right\} \times \frac{\sqrt{2}\pi\,\nu_{\text{Carr}}\tau_s \cdot r_{\text{FP arm back}} \cdot t_{\text{FP arm input}}}{\sqrt{1 + (4\pi\tau_s f)^2}} \qquad \text{(D.16)}$$

$$\times \left(r_{BS}^{\text{A.R.-side}} \cdot \sqrt{P^{00}_{\text{Carr, inline FP-arm}}} + t_{BS}^{\text{offline}} \cdot \sqrt{P^{00}_{\text{Carr, offline FP-arm}}}\right)$$

In addition, we re-write τ_s from Eq. D.14 as (c.f. Eq. 3.4 in Sec. 3.1):

$$\tau_s \cong \frac{L}{c} \cdot \frac{1}{\sqrt{P^{00}_{\text{Carr, Recyc. Cav.}}}} \cdot \frac{1}{2} \left\{ \frac{\sqrt{P^{00}_{\text{Carr, inline FP-arm}}}}{t_{BS}^{\text{inline}} \cdot t_{\text{FP arm input}}} + \frac{\sqrt{P^{00}_{\text{Carr, offline FP-arm}}}}{r_{BS}^{\text{ref-side}} \cdot t_{\text{FP arm input}}} \right\} \quad \text{(D.17)}$$

In Eq. D.16, ϕ_1, ϕ_2 represent any (ultimately arbitrary) accumulated phases for the two GW-induced sidebands which we have lumped together during the above calculations.

The GW-signal e-field, E_{exit}^{GW}, gets added to the "quiescent" e-fields which exist at the exit port in the absence of gravitational waves. This quiescent power includes both carrier light, which emerges from the beamsplitter due to imperfect contrast, and radio frequency (RF) sideband light, which serves as a local oscillator for heterodyning. Thus we have:

$$E_{exit}^{tot} = E_{exit}^{qui} + E_{exit}^{GW} \quad \text{(D.18)}$$

where:

$$E_{exit}^{qui} = E_{exit,\,carr}^{qui}\, e^{-i\omega t} + E_{exit,\,+SB}^{qui}\, e^{-i(\omega+\Sigma)t} + E_{exit,\,-SB}^{qui}\, e^{-i(\omega-\Sigma)t} \quad \text{(D.19)}$$

with $\Sigma = 2\pi\nu_{Mod}$ being the RF-modulation angular frequency.

The power at the output photodetector, $P_{exit}^{tot} = \left|E_{exit}^{tot}\right|^2$, will have components at several frequencies, including terms at DC, Σ, 2Σ, Ω, 2Ω, and $\Sigma \pm \Omega$. But only those terms at frequencies Σ and $\Sigma \pm \Omega$ will be kept by the detection process; the rest will be eliminated via demodulation with a mixer frequency of Σ, followed by low-pass filtering with a pole frequency that is chosen such that $\Omega \ll \omega_{\text{filter pole}} \ll \Sigma$. Of the remaining terms that survive, we first consider those at frequency Σ, which are computed as:

$$\left|E_{exit}^{tot}\right|^2\Big|_{\Sigma} = \cos(\Sigma t) \cdot \left[2\,\text{Re}\left\{ E_{exit,\,carr}^{qui\,*} \cdot (E_{exit,\,+SB}^{qui} + E_{exit,\,-SB}^{qui}) \right\} \right] \quad \text{(D.20)}$$

$$+ \sin(\Sigma t) \cdot \left[2\,\text{Im}\left\{ E_{exit,\,carr}^{qui\,*} \cdot (E_{exit,\,+SB}^{qui} - E_{exit,\,-SB}^{qui}) \right\} \right]$$

Careful consideration [87] of the relevant phases between the upper and lower RF-sidebands, plus the fairly accurate approximation that $\left|E_{exit,\,+SB}^{qui}\right| \approx \left|E_{exit,\,-SB}^{qui}\right|$, leads to the result that both of the terms on the right hand side of Eq. D.20 predominantly cancel out, leaving only small remainders. These quantities are further reduced due to the poor spatial overlap between the e-fields being multiplied by one another; in particular, there should be very little power in the TEM_{00} mode for $E_{exit,\,carr}^{qui}$, while $E_{exit,\,+SB}^{qui}$ and $E_{exit,\,-SB}^{qui}$ should consist primarily of TEM_{00} mode power. In any case, the terms in Eq. D.20 represent (after demodulation and low-pass filtering) only a small, DC offset to the final GW-signal, existing because of the imperfect carrier dark-fringe at the beamsplitter

exit port, but which can presumably be subtracted out *because* it is at DC. As long as there are no interferometer imperfections (e.g., large-amplitude mirror tilt noise at GW-frequencies) which not only cause significant unbalances between the plus and minus RF-sidebands, but also represent very fast fluctuations, then we can safely ignore the terms in Eq. D.20 from our considerations of the GW-induced signal-to-noise ratio.

Next, we consider the terms at frequencies $\Sigma \pm \Omega$:

$$\left|E^{tot}_{exit}\right|^2\Big|_{\Sigma\pm\Omega} = h\left[\frac{\sqrt{2}\pi\nu_{Carr}\tau_s \cdot r_{FP\ arm\ back} \cdot t_{FP\ arm\ input}}{\sqrt{1+(4\pi\tau_s f)^2}}\right. \tag{D.21}$$

$$\left. \cdot \left(r^{A.R.\text{-}side}_{BS} \cdot \sqrt{P^{00}_{Carr,\ inline\ FP\text{-}arm}} + t^{offline}_{BS} \cdot \sqrt{P^{00}_{Carr,\ offline\ FP\text{-}arm}}\right)\right]$$

$$\times\left\{\left[\left(e^{-i(\Omega t+\phi_1)} + e^{i(\Omega t+\phi_2)}\right)\cdot\left(E^{qui}_{exit,\ +\ SB}\, e^{-i\Sigma t} + E^{qui}_{exit,\ -\ SB}\, e^{i\Sigma t}\right)\right] + C.C.\right\}$$

where "C.C", as usual, means "complex conjugate".

We can simplify this expression with a few approximations. First of all, we once again treat the amplitudes of the upper and lower RF-sideband exit port e-fields as nearly equal:

$$\left|E^{qui}_{exit,\ +\ SB}\right| \approx \left|E^{qui}_{exit,\ -\ SB}\right| \equiv \sqrt{P^{00}_{SB,\ exit\ port}} \tag{D.22}$$

We can then write, after some elementary algebra:

$$\left\{\left[\left(e^{-i(\Omega t+\phi_1)} + e^{i(\Omega t+\phi_2)}\right)\cdot\left(E^{qui}_{exit,\ +\ SB}\, e^{-i\Sigma t} + E^{qui}_{exit,\ -\ SB}\, e^{i\Sigma t}\right)\right] + C.C.\right\} \tag{D.23}$$

$$= 8\cos(\Omega t+\phi_3)\cos(\Sigma t+\phi_4)\cos(\phi_5)\sqrt{P^{00}_{SB,\ exit\ port}}$$

where $\phi_3 = (\phi_1+\phi_2)/2$ is the overall phase delay (plus ϕ_0) of the response of the interferometer to the gravitational wave, and ϕ_4 is whatever "differential" phase the RF-sidebands have picked up in propagating to the beamsplitter exit port, with respect to the GW-induced signal e-fields generated from the carrier that have arrived there. Lastly, the phase ϕ_5 includes whatever "common mode" phase the upper and lower RF-sidebands pick up with respect to the GW-induced signal fields. This phase, ϕ_5, includes a $-\pi/2$ from the initial creation of the RF-sidebands from the carrier beam (e.g., Eq. 2.87, using $J_{-1}(\Gamma) = -J_1(\Gamma)$), and a $\pi/2$ due to the factors of i preceding A_1 and A_2 in Eq's. D.4, D.5. If we make the assumption that no other significant amount of "common mode" phase is picked up by the sidebands in their propagations through the system (particularly reasonable for the sideband powers in the TEM_{00} mode, which remain near-anti-resonant

in the Fabry-Perot arm cavities, and therefore pick up phase symmetrically about the carrier's phase in reflection from the arms), then we can conclude that $\phi_5 \approx 0$, and that $\cos(\phi_5) \approx 1$.

Now we consider the demodulation and low-pass filtering process. Multiplying $\left|E_{exit}^{tot}\right|^2$ by a demodulation function with the proper phase — i.e., $\cos(\Sigma t + \phi_4)$ — and integrating, we get:

$$\frac{1}{T}\int_0^T \cos(\Sigma t + \phi_4) \cdot \left|E_{exit}^{tot}\right|^2 dt \cong (\text{Ampl.}) \times \frac{1}{T}\int_0^T [\cos(\Sigma t + \phi_4) \cdot \cos(\Sigma t + \phi_4)]\, dt \qquad (D.24)$$
$$= \frac{1}{2}(\text{Ampl.})$$

where "Ampl." is the overall coefficient of $\cos(\Sigma t + \phi_4)$ in $\left|E_{exit}^{tot}\right|^2$, and where the integration time, T, for the low-pass filtering step obeys $\Omega \ll 2\pi/T \ll \Sigma$.

Finally, we must convert these power values as reported by the program (i.e., pre-normalized to 1 Watt each of carrier and upper and lower RF-sideband power) into real power values, in Watts, that will be detected by the output photodetector. If we define the total (pre-modulation) laser power to be P_{las}, and the modulation depth (as per Eq. 2.87) to be Γ, then this entails multiplying the carrier e-fields by $\sqrt{P_{las}} \cdot J_0(\Gamma)$, multiplying the RF-sideband e-fields by $\sqrt{P_{las}} \cdot J_1(\Gamma)$, and multiplying the detected signal by the finite *quantum efficiency*, η, of the output photodetector. Putting all of this together with Eq's. D.21–D.24, we can at last write the amplitude (multiplying $\cos(\Omega t + \phi_3)$) of the GW-signal, divided by h, as follows:

$$\frac{S(f)}{h} = P_{las}\eta\, J_0(\Gamma)\, J_1(\Gamma) \frac{4\sqrt{2}\pi \nu_{Carr}\tau_s \cdot r_{FP\ arm\ back} \cdot t_{FP\ arm\ input}}{\sqrt{1 + (4\pi\tau_s f)^2}} \qquad (D.25)$$
$$\cdot \sqrt{P^{00}_{SB,\ exit\ port}}\ (r_{BS}^{A.R.\text{-}side} \cdot \sqrt{P^{00}_{Carr,\ inline\ FP\text{-}arm}} + t_{BS}^{offline} \cdot \sqrt{P^{00}_{Carr,\ offline\ FP\text{-}arm}})$$

D.2 Deriving the interferometer shot noise level

To compute the amount of photon shot noise interfering with the GW-signal detection process, we first consider the shot noise created by a constant amount of power, P, falling on the output photodetector (noting that the presence or absence of gravitational waves has a completely negligible effect upon this total amount of power). The number of photons detected, with detector quantum efficiency η, is equal to:

$$\#\,\text{Photons} = \eta \cdot \frac{P}{h_{pl}\nu_{Carr}} \qquad (D.26)$$

where h_{pl} is Planck's constant, and where we make the (very accurate) approximation that $\nu \equiv \nu_{Carr} \approx \nu_{SB}$ for counting the photons in both the carrier and RF-sideband e-fields

emerging from the beamsplitter exit port.

Assuming that the number of photons in the laser beam obeys a Poisson probability distribution [e.g., 88], the shot noise will be *white* with an rms amplitude equal to:

$$\Delta(\#\,\text{Photons}) \;=\; \sqrt{\#\,\text{Photons}} \;=\; \sqrt{\frac{\eta\, P}{h_{pl}\,\nu_{Carr}}} \tag{D.27}$$

To convert this expression into rms fluctuations in detected power, we multiply it by $h_{pl}\,\nu_{Carr}$, thus yielding a formula for the rms noise amplitude:

$$N \;=\; h_{pl}\,\nu_{Carr} \times \Delta(\#\,\text{Photons}) \;=\; \sqrt{\eta\, h_{pl}\,\nu_{Carr} P} \tag{D.28}$$

The bandwidth of contributing shot noise is typically given [e.g., 25] as the inverse of the GW-signal integration time, i.e., $N^2 \propto \delta\nu \propto 1/T_{int}$. In computing the spectral density of shot noise, we assume an integration over unit bandwidth, so that $\delta\nu = 1$, thus multiplying Eq. D.28 by unity. By counting the noise contributions at both "positive" and "negative" frequencies (or equivalently, the "sine" and "cosine" terms), we (incoherently) add together those contributions from $(\pm\,\nu_{Mod} \pm f)$ (c.f. the left hand side of Eq. D.23). This multiplies N^2 by 4, or N by 2. On the other hand, the demodulation step shown in Eq. D.24 reduces the noise amplitude N by $1/2$, just as it reduced the amplitude of S by half, so that these factors of 2 and $1/2$ cancel, and Eq. D.28 remains unchanged.

Lastly, we must determine what power level to use as "P" in the shot noise formula, Eq. D.28. Neglecting E^{GW}_{exit} in Eq's. D.18, D.19, we get:

$$\begin{aligned}
\left|E^{tot}_{exit}\right|^2 \approx \left|E^{qui}_{exit}\right|^2 \;&=\; \left|E^{qui}_{exit,\,carr}\right|^2 + \left|E^{qui}_{exit,\,+\,SB}\right|^2 + \left|E^{qui}_{exit,\,-\,SB}\right|^2 \\
&+ 2\,[E^{qui}_{exit,\,+\,SB}{}^{*} \cdot E^{qui}_{exit,\,carr} + E^{qui}_{exit,\,carr}{}^{*} \cdot E^{qui}_{exit,\,-\,SB}]\cos(\Sigma t) \\
&+ 2\,[E^{qui}_{exit,\,+\,SB}{}^{*} \cdot E^{qui}_{exit,\,-\,SB}]\cos(2\Sigma t) \\
&\approx P^{tot}{}_{Carr,\,exit\,port} + \{2P^{tot}{}_{SB,\,exit\,port} \cdot [1 + \{Phase\} \times \cos(2\Sigma t)]\} + \{Ampl.\} \times \cos(\Sigma t)
\end{aligned} \tag{D.29}$$

where the last step was performed via the approximation of Eq. D.22.

In Eq. D.29, we see that the power falling on the output photodetector is clearly not constant as was considered originally, but has variability at frequencies Σ and 2Σ. The effects of this temporal modulation of the output power on the measured level of shot noise was first discussed in the context of Gravitational Wave detection by Schnupp [89]. These effects were later calculated in detail and/or measured in the laboratory by several researchers [46, 90, 91, 92]. The resulting shot noise *after demodulation* is the sum of a white, "broadband" component, plus sharp components at DC and at harmonics of the RF-modulation frequency [92]. The noise terms at harmonics of Σ can be dropped due to the

low-pass filtering in Eq. D.24, and the DC term (produced from the $\cos(\Sigma t)$ term in Eq. D.29) is merely a DC offset, and in fact is very small (as per the discussion following Eq. D.20, above). But the broadband component, on the other hand, is what limits the interferometer sensitivity; and its amplitude due to the exit port RF-sideband power (assuming sinusoidal modulation and demodulation) is *bigger* than that which would be predicted from just the time-averaged sideband power — i.e., if we neglected the $\cos(2\Sigma t)$ term in Eq. D.29 — by a factor of $3/2$. Thus for the value of P in Eq. D.28, we must write:

$$P \rightarrow P^{tot}{}_{Carr,\,exit\,port} + \left(\frac{3}{2}\right) \cdot 2P^{tot}{}_{SB,\,exit\,port} = P^{tot}{}_{Carr,\,exit\,port} + 3\,P^{tot}{}_{SB,\,exit\,port} \qquad (D.30)$$

Finally, by inserting the necessary factors of P_{las}, $J_0(\Gamma)$, and $J_1(\Gamma)$ where appropriate in Eq. D.30 to convert the results into real power values (in Watts), we compute the final expression for the shot-noise as:

$$N = \sqrt{\eta\, h_{pl}\, \nu_{Carr} P_{las} \cdot [J_0(\Gamma)^2 P^{tot}{}_{Carr,\,exit\,port} + 3\, J_1(\Gamma)^2 P^{tot}{}_{SB,\,exit\,port}]} \qquad (D.31)$$

D.3 Putting it all together: S/N and the shot-noise-limited Gravitational Wave sensitivity function

The signal-to-shot-noise ratio of the interferometric detector (for a specified value of GW-strength, h, and GW-frequency, f) can be computed by dividing S, as given in Eq. D.25, by the formula for N, as given in Eq. D.31. If we define the shot noise sensitivity limit, $\tilde{h}_{SN}(f)$, as the value of h — assuming a monochromatic GW of optimal incidence angle and polarization — for which S/N is unity after a 1 second integration time, then we have:

$$\tilde{h}_{SN}(f) \equiv \left\{ \frac{S(f)/h}{N} \right\}^{-1} \qquad (D.32)$$

The computation of this final expression results in Eq. 3.3 of Sec. 3.1, as promised.

D.4 Other interferometer sensitivity expressions

The shot-noise-limited sensitivity expression which has been computed during the previous sections is not the only mathematical way in which the noise level in a Gravitational Wave detector is reported. The expression implied by Eq. D.32 is a generic expression, representing the raw spectral density of noise for a single LIGO interferometer; and while this formula represents maximum flexibility and universality, it is often useful to convert it into more complex expressions (as we have done in Sec. 3.4, for example) which quantify

relevant properties of the detectors (e.g., multi-detector correlations), or of the GW-sources (e.g., signal processing differences for "burst", "periodic", and "stochastic" sources [2].)

For this reason, we conclude this appendix by giving a brief definition of some of the commonly used GW-signal and noise expressions (for *total* noise, not just shot noise) that are considered in Gravitational Wave detection [12]:

- $\tilde{h}(f)$, the *noise spectrum*, which is the square root of the spectral density of a single interferometer's noise floor, converted into an effective GW-strain amplitude. It has dimensions of $1/\sqrt{\text{Hz}}$. The comparison strain amplitude for this expression represents a monochromatic Gravitational Wave of frequency f, impinging upon the interferometric detector with optimal polarization and source orientation on the sky, which produces a unity signal-to-noise ratio when integrated for a period, T_{int}, of 1 second[1]. A computed shot noise curve, $\tilde{h}_{SN}(f)$, as defined by Eq. D.32 represents only the shot noise component of this total noise spectrum, which must be added to other noise components incoherently (i.e., in quadrature) to form the total noise spectrum.

- $h_{rms}(f) \equiv \sqrt{f}\,\tilde{h}(f)$, the *rms strain amplitude noise*, in a bandwidth equal to frequency (i.e., this formula assumes that a broadband GW-source will spend time $T \approx 1/\Delta f \approx 1/f$ in the vicinity of frequency f). This brings the noise spectrum into dimensionless form, for ready comparison with the dimensionless "characteristic amplitude" (see below) of GW's from a broadband source.

- $h_c(f) \equiv \tilde{h}(f)\sqrt{n_{cycle}}$, the *characteristic amplitude of the GW's*, where n_{cycle} is the approximate number of GW oscillation periods which the GW spends near frequency f. Thus, for GW's with optimal source direction and polarization, the amplitude signal-to-noise ratio is given by $S/N \equiv h_c(f)/h_{rms}(f)$. Random source orientation and polarization reduces this S/N by a factor of $\sqrt{5}$.

- $h_{SB}(f) = 5\sqrt{5}\,h_{rms}(f)$, the *sensitivity to GW bursts*. The characteristic GW-source amplitude necessary for it to be jointly detected by *all* of LIGO's interferometers with high confidence ($S/N \geq 5$ in the 4 km interferometers at the Louisiana and Washington sites, and $S/N \geq 2.5$ in the 2 km interferometer at the Washington site), assuming random orientation and polarization, and no prior information for source triggering.

1. The integrated GW-signal goes up like T_{int}, while the integrated (incoherent) shot noise goes up like $\sqrt{T_{int}}$, so that for signal integration times other than 1 second, $\tilde{h}(f)$ goes down (i.e., is *improved*) like $(S/N)^{-1} \propto 1/\sqrt{T_{int}}$.

Bibliography

[1] A. Einstein, *Preuss. Akad. Wiss. Berlin, Sitzungsberichte der Physikalisch-mathematischen Klasse*, 154 (1918).

[2] K. S. Thorne, *"Gravitational Radiation"* in *300 Years of Gravitation*, S. W. Hawking and W. Israel (eds.), Cambridge University Press, Cambridge, 330 (1987).

[3] A. Abramovici, W. E. Althouse, R. W. P. Drever, Y. Gürsel, S. Kawamura, F. J. Raab, D. Shoemaker, L. Sievers, R. E. Spero, K. S. Thorne, R. E. Vogt, R. Weiss, S. E. Whitcomb, and M. E. Zucker, Science **256**, 325 (1992).

[4] C. W. Misner, K. S. Thorne, and J. A. Wheeler, *Gravitation*, W. H. Freeman and Co., N. Y., 1971.

[5] R. E. Vogt, R. W. P. Drever, F. J. Raab, K. S. Thorne, and R. Weiss, *"The Construction, Operation, and Supporting Research and Development of a Laser Interferometer Gravitational-Wave Observatory", Proposal to the National Science Foundation*, California Institute of Technology, 1989.

[6] S. Weinberg, *Gravitation and Cosmology*, John Wiley & Sons, N. Y., 1972.

[7] K. C. B. New, *et al.*, Astrophys. J. **450**, 757 (1995).

[8] S. van den Bergh and G. A. Tamman, Annu. Rev. Astron. Astrophys. **29**, 363 (1991).

[9] K. S. Thorne, *"Gravitational Waves"* in *Proceedings of the Snowmass 95 Summer Study on Particle and Nuclear Astrophysics and Cosmology*, E. W. Kolb and R. Peccei (eds.), World Scientific, Singapore, 160 (1995).

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

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

[12] B. Barish, *et al.*, *"Proposal for a Research and Development Program for Advanced LIGO Detectors by the LIGO MIT/Caltech Groups", Proposal to the National Science Foundation*, California Institute of Technology, 1996.

[13] R. A. Hulse and J. H. Taylor, Astrophys. J. **324**, 355 (1975).

[14] J. H. Taylor, Rev. Mod. Phys. **66**, 711 (1994).

[15] M. Zucker, *"Laser Interferometer Gravitational-Wave Observatory"*, presented at *Massachusetts Institute of Technology Astrophysics Colloquium*, September, 1997.

[16] R. Loudon, *The Quantum Theory of Light*, Clarendon Press, Oxford, 1983.

[17] A. E. Siegman, *Lasers,* University Science Books, California, 1986.

[18] Y. Hefetz, N. Mavalvala, D. Sigg, J. Opt. Soc. Am. B **14**, No. 7, 1597 (1997).

[19] J.-Y. Vinet and P. Hello, J. Mod. Opt. **40**, No. 10, 1981 (1993).

[20] H. P. Kortz and H. Weber, Appl. Opt. **20**, No. 11, 1936 (1981).

[21] P. Hello and J.-Y. Vinet, *"Modelling of the VIRGO Interferometer"* in *Gravitational Astronomy: Instrument Design and Astrophysical Prospects*, D. E. McClelland and H.-A. Bachor (eds.), World Scientific, Singapore, 1 (1991).

[22] J.-Y. Vinet, P. Hello, C. N. Man, and A. Brillet, J. Phys. I (Paris) **2**, No. 7, 1287 (1992).

[23] R. W. P. Drever, *"Interferometric Detectors for Gravitational Radiation"* in *Gravitational Radiation*, N. Deruelle and T. Piran (eds.), North-Holland Publishing Co., Amsterdam, 321 (1983).

[24] B. J. Meers, Phys. Rev. D **38**, No. 8, 2317 (1988).

[25] J.-Y. Vinet, B. J. Meers, C. N. Man, and A. Brillet, Phys. Rev. D **38**, No. 2, 433 (1988).

[26] B. J. Meers and K. A. Strain, Phys. Rev. D **43**, No. 10, 3117 (1991).

[27] D. Redding, M. Regehr, and L. Sievers, *"Dynamic Models of Fabry-Perot Interferometers"*, submitted to Appl. Opt., 1997.

[28] R. G. Beausoleil, *"Spatiotemporal Model of the LIGO Interferometer"*, Contract No. PC188097 Interim Report, Massachusetts Institute of Technology, June, 1996.

[29] A. J. Tridgell, D. E. McClelland, and C. M. Savage, *"Numerically Modelling a Dual Recycling Interferometric Gravitational Wave Detector"* in *Gravitational Astronomy: Instrument Design and Astrophysical Prospects*, D. E. McClelland and H.-A. Bachor (eds.), World Scientific, Singapore, 222 (1991).

[30] B. J. Meers, *"Some Recent Developments in Interferometric Gravitational-Wave Detection"* in *Gravitational Astronomy: Instrument Design and Astrophysical Prospects*, D. E. McClelland and H.-A. Bachor (eds.), World Scientific, Singapore, 281 (1991).

[31] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in Fortran: The Art of Scientific Computing*, Cambridge University Press, Cambridge, England, 1992.

[32] J. Y. Wang and D. E. Silva, Appl. Opt. **19**, No. 9, 1510 (1980).

[33] F. A. Jenkins and H. E. White, *Fundamentals of Optics*, McGraw-Hill, New York, 1976.

[34] B. Kells, LIGO Technical Note, LIGO-E950099-04 (1995).

[35] B. Kells, personal communication, 1995.

[36] M. C. Weisskopf, Astro. Lett. and Communications **26**, 1 (1987).

[37] S. Whitcomb, G. Billingsley, J. Carri, A. Golovitser, D. Jungwirth, W. Kells, H. Yamamoto, B. Bochner, Y. Hefetz, P. Saha, and R. Weiss, *"Optics Development for LIGO"* in *Proceedings of the TAMA Workshop on Gravitational Wave Detection*, K. Tsubono (ed.), Universal Academic Press, Tokyo, 1997.

[38] P. Fritschel, *et al.*, *"Alignment of an Interferometric Gravitational Wave Detector"*, to be submitted to Phys. Rev. D, 1998.

[39] N. Mavalvala, personal communication, 1996.

[40] P. Fritschel, personal communication, 1996.

[41] R. L. Liboff, *Introductory Quantum Mechanics*, Addison-Wesley Pub. Co., Reading, MA, 1980.

[42] *CRC Standard Mathematical Tables*, W. H. Beyer (ed.), CRC Press Inc., Boca Raton, 1988.

[43] B. Bhawal, Appl. Opt. **35**, No. 7, 1041 (1996).

[44] P. Saha, J. Opt. Soc. Am. A **14**, No. 9, 2195 (1997).

[45] R. Weiss, D. Shoemaker, personal communications.

[46] M. B. Gray, A. J. Stevenson, H.-A. Bachor, and D. E. McClelland, Optics Lett. **18**, No. 10, 759 (1993).

[47] L. Schnupp, Max Planck Institute for Quantum Optics, Garching, Germany, unpublished (1986); R. Drever, California Institute of Technology, Pasadena, unpublished (1991); M. W. Regehr, F. J. Raab, and S. E. Whitcomb, LIGO Technical Note, LIGO-P950001-00-R (1995).

[48] D. Sigg, *et al.*, *"Signal Extraction in a Power-Recycled Michelson Interferometer with Fabry-Perot Arm Cavities Using a Multiple-carrier Frontal Modulation Scheme"*, to be submitted to Appl. Opt., 1998.

[49] D. E. McClelland, C. M. Savage, A. J. Tridgell, and R. Mavaddat, Phys. Rev. D **48**, No. 12, 5475 (1993).

[50] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, Inc., San Diego, 1985.

[51] S. Wolfram, *Mathematica: A System for Doing Mathematics by Computer*, Addison-Wesley Publishing Company, Inc., Redwood City, CA, 1991.,

[52] W. Winkler, *et al.*, Appl. Opt. **33**, No. 31, 7547 (1994).

[53] T. Li, Bell Syst. Tech. J. **44**, 917 (1965).

[54] P. Saha and R. Weiss, personal communications; P. Saha, *"Perturbation Analysis"*, Massachusetts Institute of Technology, Cambridge, Mass., unpublished (1995).

[55] D. Shoemaker, LIGO Technical Note, LIGO-E960112-06-D (1996).

[56] *PV-WAVE*, Visual Numerics, ©1988–1992.

[57] A. Lazzarini and R. Weiss, LIGO Technical Note, LIGO-E950018-02-E (1996).

[58] B. Bochner, *"Modelling the Performance of an Initial-LIGO Detector with Realistically Imperfect Optics"*, in *Proceedings of the 18th Texas Symposium on Relativistic Astrophysics*, A. Olinto, J. Frieman, and D. Schramm (eds.), World Scientific, Singapore (1997).

[59] A. Abramovici, *et al.*, *"Gravitational Wave Astrophysics"*, in *Proceedings of the Snowmass 95 Summer Study on Particle and Nuclear Astrophysics and Cosmology*, E. W. Kolb and R. Peccei (eds.), World Scientific, Singapore, 398 (1995).

[60] C. Cutler, *"The CW Problem for Interferometric Detectors"*, presented at the *Gravitational Wave Data Analysis Workshop*, Massachusetts Institute of Technology, Cambridge, Massachusetts, unpublished (1996).

[61] E. E. Flanagan and S. Hughes, *"Measuring gravitational waves from binary black hole coalescences I. Signal to noise for inspiral, merger, and ringdown"*, astrophysics preprint gr-qc/9701039, submitted to Phys. Rev. D, 1997.

[62] D. Shoemaker, *et al.*, *"Comparison of 2 Fixed Mass Interferometer Testbeds and the Resulting Recommendation for the Initial LIGO Interferometer Design"*, Massachusetts Institute of Technology, Cambridge, Mass., unpublished (1993).

[63] A. Abramovici and D. H. Shoemaker, LIGO Technical Note, LIGO-T950047-01-E (1995).

[64] K. Sliwa and B. Bochner, *"FFT model of the PNI"*, LIGO Technical Note, in preparation.

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

[66] J. Mizuno, *"Comparison of optical configurations for laser-interferometric gravitational-wave detectors"*, Ph.D. thesis, Max-Planck-Institut für Quantenoptik, Garching (1995).

[67] B. Kells and J. Camp, LIGO Technical Note, LIGO-T970097-00-D (1997).

[68] A. Krolak, J. A. Lobo, B. J. Meers, Phys. Rev. D **43**, No. 8, 2470 (1991).

[69] B. J. Meers, A. Krolak, J. A. Lobo, Phys. Rev. D **47**, No. 6, 2184 (1993).

[70] A. J. Tridgell, D. E. McClelland, C. M. Savage, and B. J. Meers, *"A numerical model of a dual recycling laser interferometer"*, in *Proceedings of the Sixth Marcel Grossman Meeting on General Relativity*, H. Sato and T. Nakamura (eds.), World Scientific, Singapore, 218 (1992).

[71] P. K. Fritschel, *"Techniques for Laser Interferometer Gravitational Wave Detectors"*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts (1992).

[72] K. A. Strain and B. J. Meers, Phys. Rev. Lett. **66**, No. 11, 1391 (1991).

[73] R. Mavaddat, D. E. McClelland, P. Hello, and J.-Y. Vinet, J. Optics (Paris) **26**, No. 4, 145 (1995).

[74] D. E. McClelland, Aust. J. Phys. **48**, 953 (1995).

[75] A. Gillespie and F. Raab, Phys. Lett. A **178**, 357 (1993).

[76] A. Gillespie and F. Raab, Phys. Lett. A **190**, 213 (1994).

[77] A. Gillespie and F. Raab, Phys. Rev. D **52**, No. 2, 577 (1995).

[78] S. Kawamura and F. Raab, LIGO Technical Note, LIGO-T950011-14-D (1996).

[79] S. Kawamura, J. Hazel, and F. Raab, LIGO Technical Note, LIGO-T960074-07-D (1996).

[80] M. Zucker, personal communication, 1997.

[81] B. Kells, personal communication, 1998; G. Billingsley, *et al.*, LIGO Technical Note, LIGO-E980061-00-D (1998); G. Billingsley, *"Optical Development of Test Masses"*, California Institute of Technology, Pasadena, unpublished (1997).

[82] P. Hello and J.-Y. Vinet, Phys. Lett. A **178**, 351 (1993).

[83] J. Mizuno, *et al.*, Phys. Lett. A **175**, No. 5, 273 (1993).

[84] K.-X. Sun, *et al.*, Phys. Rev. Lett. **76**, No. 17, 3053 (1996).

[85] J. Hough, *et al.*, *"Proposal for a 600 m Laser-Interferometric Gravitational Wave Antenna — GEO 600"*, unpublished (1994).

[86] M. Zucker, *et al.*, LIGO Technical Note, LIGO-E960027-00-D (1996).

[87] B. Bochner, *"A Simulation Program for a First-Generation LIGO Gravitational-Wave Interferometer"*, unpublished (1996).

[88] D. Marcuse, *Principles of Quantum Electronics,* Academic Press, Inc., New York, 1980.

[89] L. Schnupp, *"The effect of modulation on the shot noise limited sensitivity of the gravitational wave detector"*, unpublished (1989).

[90] T. M. Niebauer, *et al.*, Phys. Rev. A **43**, No. 9, 5022 (1991).

[91] B. J. Meers and K. A. Strain, Phys. Rev. A **44**, No. 7, 4693 (1991).

[92] M. W. Regehr, *"Signal Extraction and Control for an Interferometric Gravitational Wave Detector"*, Ph. D. thesis, California Institute of Technology, Pasadena (1995).