Deep–Frequency Modulation Interferometry for Gravitational Wave Detectors

Dissertation
zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
Fachbereich Physik
der Universität Hamburg

vorgelegt von

Tobias Benjamin Eckhardt

Hamburg

2025

Abstract

This thesis presents my research on a new type of local displacement sensors proposed to be used in gravitational wave detectors to improve upon the alignment and control noise in the low frequency regime around 3 Hz.

Gravitational wave astronomy is an emerging field using large interferometers with km long arms to measure small displacements of less than $10^{-20}\,\mathrm{m}/\sqrt{\mathrm{Hz}}$ caused by passing gravitational waves. The currently running LIGO and Virgo detectors are limited in the low-frequency region (below 30 Hz) by alignment and control noises of their suspended optics. As a possible path to improve upon these noise limitations, the use of more precise local sensors to measure the local displacement of the optics, counteract their motion and improve upon their alignment has been proposed. This thesis presents my research on local displacement sensors based on "deep–frequency modulation interferometry (DFMI)" which allows for precise measurements down to displacements in the order of $\sim 100\mathrm{fm}/\sqrt{\mathrm{Hz}}$ while simultaneously providing an absolute displacement readout allowing for measurements over a large dynamic range over millimeter and centimeter, necessary for some of the suspended optics in gravitational wave detectors. I present an "analytic readout algorithm" developed to extract the displacement parameters of interest from a measured DFMI signal and I conduct a thorough analysis of the achievable precision limits of DFMI using the "Cramér-Rao bound". The readout algorithm I present allows for a faster readout than previous experiments using DFMI, significantly increasing the control bandwidth and the number of sensor channels that can be processed. My noise analysis proves that the readout algorithm runs close to optimal precision (given by the Cramér-Rao bound) and provides a displacement readout similar to other interferometry–based displacement sensing techniques. These results show that using "deep–frequency modulation interferometry (DFMI)" based local displacement sensors can help current and future planned ground–based gravitational wave detectors to reach their target/design sensitivity in the low frequencies around $10^{-21}\,\mathrm{m}/\sqrt{\mathrm{Hz}}$.

Zusammenfassung

In dieser Arbeit wird mein Forschungsbeitrag zu einem neuartigen Typ lokaler Längensensoren vorgestellt, mit dem Ziel solchen lokalen Sensoren in zukünftigen Gravitationswellendetektoren einzusetzen, um dort die Ausrichtung der Optiken zu verbessern und das Rauschen im Niederfrequenzbereich um 3 Hz zu reduzieren. Die Gravitationswellenastronomie ist ein aufstrebendes Gebiet, bei dem große Interferometer mit kilometerlangen Armen eingesetzt werden, um kleinste Längenänderungen von weniger als $10^{-20}\,\mathrm{m}/\sqrt{\mathrm{Hz}}$, die durch vorbeiziehende Gravitationswellen verursacht werden, zu messen. Die derzeit laufenden LIGO- und Virgo-Detektoren sind im Niederfrequenzbereich (unter 30 Hz) durch Rauschen in der Ausrichtung und der Regelungstechnik ihrer aufgehängten Optiken limitiert. Als möglicher Weg zur Verbesserung dieser Rauschlimitierungen wurde die Verwendung präziserer lokaler Sensoren vorgeschlagen, um die lokale Verschiebung der Optiken zu messen, ihrer Bewegung entgegenzuwirken und ihre Ausrichtung zu verbessern. In dieser Arbeit werden meine Forschungen zu lokalen Längensensoren auf der Grundlage von „acrfulldfmi" vorgestellt. DFMI ermöglicht präzise Messungen in Größenordnung von bis zu $\sim 100\mathrm{fm}/\sqrt{\mathrm{Hz}}$, und gleichzeitig eine absolute Distanzmessung über einen großen dynamischen Bereich von Millimetern und Zentimetern, was für einige der aufgehängten Optiken in Gravitationswellendetektoren erforderlich ist. Ich stelle einen „analytic readout algorithm" vor, der entwickelt wurde, um die relevanten Parameter aus einem gemessenen DFMI-Signal zu extrahieren, und ich führe eine umfangreiche Analyse der erreichbaren Genauigkeitsgrenzen von DFMI unter Verwendung des „Cramér-Rao bound" durch. Der von mir vorgestellte Auslesealgorithmus ermöglicht eine schnellere Auslesung als frühere DFMI Experimente, wodurch die Bandbreite von verbundenen Regelschleifen und die Anzahl der zu verarbeitenden Sensorkanäle deutlich erhöht werden können. Meine Rauschanalyse beweist, dass der Auslesealgorithmus nahe an der optimalen Genauigkeit (gegeben durch den Cramér-Rao bound) arbeitet und eine Messung der Distanz ermöglicht, die mit anderen interferometrischen Messtechniken vergleichbar ist. Diese Ergebnisse zeigen, dass die Verwendung von lokalen Längensensoren, die auf „acrfulldfmi" basieren, aktuellen und zukünftig geplanten bodengestützten Gravitationswellendetektoren helfen kann, ihre Ziel–Empfindlichkeit in den niedrigen Frequenzen um $10^{-21}\,\mathrm{m}/\sqrt{\mathrm{Hz}}$ zu erreichen.

Contents

Acronyms

ADC   Analog-to-Digital Converter.
ARA   analytic readout algorithm.
ASD   amplitude spectral density.

BH   black hole.
BOSEM   Birmingham Optical Sensor and Electromagnetic actuator.

CDS   Control and Design System.
COBRI   Compact Balanced Readout Interferometer.
CRB   Cramér-Rao bound.

DESY   Deutsches Elektronen-Synchrotron.
DFM   deep–frequency modulation.
DFMI   deep–frequency modulation interferometry.
DPM   deep–phase modulation.
DPMI   deep–phase modulation interferometry.

EFE   Einstein's field equations.
EM   electromagnetic.
ESD   energy spectral density.
ET   Einstein Telescope.
European XFEL   European X-Ray Free-Electron Laser Facility.

FPGA   Field Programmable Gate Array.

GWD   gravitational wave detector.

HEPI   Hydraulic External Pre-Isolator.
HoQI   Homodyne quadrature Interferometer.

ISI   Internal Seismic Isolation.

KAGRA   Kamioka Gravitational Wave Detector.

LIGO   Laser Interferometer Gravitational-Wave Observatory.

LISA	Laser Interferometer Space Antenna.
LTI	linear time-invariant system.
LUT	lookup table.
MTCA	Micro Telecommunication and Computing Architecture.
NCO	numerically controlled oscillator.
NS	neutron star.
OSEM	optical (shadow) sensor and electromagnetic actuator.
PD	photodiode.
PSD	power spectral density.
PUM	penultimate mass.
QMC	quasi monolithic component.
ReDFMI	resonantly enhanced deep–frequency modulation interferometry.
RMS	root-mean-square.
SNR	signal-to-noise ratio.
SoC	System-on-a-Chip.
SQL	standard quantum limit.
THD	total harmonic distortion.
VHDL	VHSIC Hardware Description Language.

Thesis structure

The following thesis is structured into 4 main parts:

Part I serves as basic introduction to the field of gravitational wave detectors and presents the main motivation for my later work: the improvement of current and future ground–based gravitational wave detectors. I give a brief rundown of the history of gravitational waves and their detectors, starting with the earliest attempts to detect them, up to the currently running, and future planned detectors. Although I specifically highlight the Einstein Telescope – a potential future detector that my work may support – a more detailed analysis of how local displacement sensors can enhance the sensitivity of gravitational wave detectors in general is provided in the chapter thereafter, using the currently operational LIGO detector as an example.

Part II then provides a "What is ... ?" summary of the deep–frequency modulation interferometry (DFMI) technique (the main topic of this thesis) and an in-depth explanation of the "analytic readout algorithm (ARA)" I developed for DFMI. As I consider this readout algorithm one of the main results of my Ph.D. studies, I am providing an extensive guide (corresponding to chapters 4 and 5) that one could use as reference for the DFMI readout in the future. To keep the overall structure concise, this part also contains some references to an "optimal precision" which is only explained later and a summary of the "ReDFMI" technique which is similarly motivated by my noise analysis of the next part, but included here as part of the basic summary of DFMI related techniques.

Part III then contains an in-depth mathematical analysis of the noise sources limiting the DFMI readout. As I was aiming to provide a mathematical rigorous formalism to calculate the precision limits for different noise sources, this part also includes recapitulations of common terms of (mathematical) statistics and signal analysis and how they relate to one another, relevant for the later explanations. While not strictly necessary for understanding the other parts of this thesis, it might be interesting for readers already well–versed in the field to look at some commonly used expressions from a sightly different perspective. The main takeaway for the reader here is to show how a "power spectral density (PSD)," which is often used to express measured signals, relates to the probability density function of a signal and its noise, which is the key component when calculating the formal noise limits using the "Cramér-Rao bound." The Cramér-Rao bound and the precision limits for deep–frequency modulation interferometry signals derived from it are then the main subjects of this part and present the second major result of my Ph.D. studies. While the noise limits given by the Cramér-Rao bound presented there only apply to some very specific noise sources affecting DFMI sensors, I believe that the formal use of the Cramér-Rao bound can be extended to other (non–white) noise sources in the future as well.

The final **Part IV** of this thesis shows an experimental DFMI setup implementing the analytic readout algorithm and an example noise analysis of the previously derived readout limits. It serves mainly as example application of the theoretical concepts presented earlier and adds some minor practical considerations.

Part I

Gravitational waves and their detectors

Chapter 1

A short history of gravitational waves and their detectors

The first part of this thesis is inspired by an article from Sormani et al. [1] about the mathematical history of gravitational waves. As it is often helpful for understanding a topic to follow its historic development, this chapter provides some historic context to the research of gravitational waves and their detectors. For any interested reader, I recommend having a look at Sormani's article, as it briefly mentions these contributions and adds further references for reading.

1.1 Einstein's first prediction of gravitational waves

Nov 1915
General relativity

Einstein published his field equations in November 1915 [2], describing a connection between physical quantities like masses and geometric quantities like the metric- and curvature-tensors that describe the shape of space-time. A common way to write Einstein's field equations (EFE) is:

$$R_{\mu\nu} - \frac{R}{2}g_{\mu\nu} + \Lambda g_{\mu\nu} = T_{\mu\nu} \qquad (1.1)$$

with $R_{\mu\nu}$ as Ricci curvature tensor, R as scalar curvature, $g_{\mu\nu}$ as metric tensor, Λ as cosmological constant and $T_{\mu\nu}$ as stress-energy-tensor with $\mu, \nu \in \{0, 1, 2, 3\}$ as indices.

1916 & 1918
Einstein's first prediction of linear gravitational waves

Einstein himself also studied (and published in 1916 and 1918 [2]) the case of small perturbations of the metric tensor for empty space-time (setting $T_{\mu\nu} \equiv 0$), which yields a wave equation ($\Box h_{\mu\nu} = 0$) for small metric perturbations $h_{\mu\nu}$ which he interpreted as gravitational waves.

As the solutions to this linearized wave equations do not solve the full (non–linear) field equations, Einstein was uncertain if physical gravitational waves that solve the full EFE exist.

For completeness and later reference, the calculations leading to the theory of "linearized gravity" and the resulting "plus" and "cross" polarized gravitational waves are given here briefly:

1.1.1 Definition : Linear gravitational waves

The first step is to decompose the gravitational field given by the metric tensor $g_{\mu\nu}$ into a static (and flat) background given by the Minkowski-Metric ($\delta_{\mu\nu} = \text{diag}(-1,1,1,1)$) and a perturbation (the gravitational wave) given by $\varepsilon h_{\mu\nu}$ with $\varepsilon \ll 1$ as small scaling factor:

$$g_{\mu\nu} = \delta_{\mu\nu} + \varepsilon h_{\mu\nu} \, . \tag{1.2}$$

Inserting (1.2) into the EFE and neglecting all higher order terms of ϵ ($\varepsilon^2 \approx 0$) yields as "linearized" EFE for empty space (and $\Lambda = 0$):

$$-\frac{1}{2}\Box h_{\mu\nu} + \frac{1}{2}\delta_{\mu\nu}\Box h + \frac{1}{2}\left(\partial^\alpha\partial_\mu h_{\alpha\nu} + \partial^\alpha\partial_\nu h_{\alpha\mu} - \delta_{\mu\nu}\partial^\alpha\partial^\beta h_{\alpha\beta} - \partial_\mu\partial_\nu h\right) = 0 \tag{1.3}$$

This approximation is also referred to as "linearized gravity" [3]. Imposing the "harmonic gauge" condition ($g^{\mu\nu}\Gamma^\alpha_{\mu\nu} = 0 \overset{\text{here}}{\Longrightarrow} \partial^\nu h_{\alpha\nu} = \frac{1}{2}\partial^\alpha h$) simplifies the term inside the brackets in (1.3) to ($-\frac{1}{2}\delta_{\mu\nu}\Box h$). Using this gauge and multiplying with -2, (1.3) becomes

$$\Box\left(h_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}h\right) = 0 \, . \tag{1.4}$$

Further restricting the choice of coordinates ($h_{\mu\nu} \mapsto \bar{h}_{\mu\nu}$) by scaling the ratio between time and space coordinates such that the trace of the metric perturbation vanishes ($\bar{h}^a_a = \bar{h} = 0$, the "traceless" condition), simplifies (1.4) further to a wave equation for the metric perturbation

$$\Box\bar{h}_{\mu\nu} = 0 \, . \tag{1.5}$$

When writing the solutions down explicitly, the coordinate choice is usually restricted further by setting the time coordinate orthogonal to the spatial coordinates ($\bar{h}_{0i} = 0$) and choosing a (light-like) observer travelling along the gravitational wave in the time coordinate ($\bar{h}_{00} = 0$). This, together with the traceless condition (which yields together with the "harmonic gauge" $\implies \partial^\alpha\bar{h}_{\mu\alpha} = 0$) is also referred to as "transverse-traceless (TT)" gauge which can be defined directly as:

$$\bar{h}^{\text{TT}}_{0\mu} = 0, \quad \bar{h}^{\text{TT}\,\alpha}_{\quad\alpha} = 0, \quad \partial^\alpha\bar{h}^{\text{TT}}_{\mu\alpha} = 0 \tag{1.6}$$

The wave equation (1.5) in this TT-gauge is solved by

$$\bar{h}^{\text{TT}}_{\mu\nu} = A_{\mu\nu}\cos(k_\beta x^\beta) \, . \tag{1.7}$$

The TT-gauge conditions translate here to conditions for the amplitude $A_{\mu\nu}$ and the wave-vector k_β given by:

$$\bar{h}^{\text{TT}}_{0\mu} = 0 \implies A_{0\mu} = 0 \tag{1.8}$$

$$\bar{h}^{\text{TT}\,\alpha}_{\phantom{\text{TT}}\alpha} = 0 \implies A_\alpha^{\,\alpha} = 0 \tag{1.9}$$

$$\partial^\alpha \bar{h}^{\text{TT}}_{\mu\alpha} = 0 \implies A_{\mu\alpha} k^\alpha = 0 \tag{1.10}$$

$$\left(\text{and}\quad \Box \bar{h}^{\text{TT}}_{\mu\nu} = 0 \implies k^\alpha k_\alpha = 0\right) \tag{1.11}$$

For a wave travelling along the $x_3 = z$ (spatial) direction with frequency ω (such that $k_0 = \omega/c$) fixes the wave vector to $k_\mu = (\omega/c, 0, 0, -\omega/c)$ and allows us to write the metric perturbation / the gravitational wave down as

$$\bar{h}^{\text{TT}}_{\mu\nu} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & h_+ & h_\times & 0 \\ 0 & h_\times & -h_+ & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cos(\omega(z/c - t)), \tag{1.12}$$

with the free parameters h_+ and $h_\times$ being referred to as "plus-" and "cross-polarization" amplitudes of the gravitational waves, and $x_0 = t$ as time coordinate.

1937
Einstein's (and Rosen's) first non–linear solution and his doubts on the existence of gravitational waves

After his initial presentation of linearized GWs, Einstein together with Nathan Rosen continued to work on finding a solution that would satisfy the full (non–linear) theory. By 1937, Rosen and Einstein [4] were somewhat successful and even published a paper ("On gravitational waves") presenting an exact solution for what can be considered cylindrical gravitational waves.

Despite this success, Einstein remained highly doubtful that such mathematical solutions correspond to real physical phenomena. E.g., his cylindrical solution contained coordinate singularities which he did not recognize as such at the time, and even with a better choice of coordinates they contain a singularity along the rotational symmetry axis and the wavefront being infinitely stretched, which lead Einstein to believe that such mathematical solution were non–physical and gravitational waves may not exist [1]. With Einstein adapting a critical opinion about gravitational waves the search for them did not make significant progress for about 20 years until new researchers began to work on this problem again.

Chapel Hill Conference in 1957

The first singularity free rigorous solution of gravitational waves and answers to the questions of whether GWs carry energy and could physically exist were answered by the combined works of Herman Bondi, Ivor Robinson, Felix Pirani and Andrzej Trautman, shortly after they met at the Chapel Hill Conference in 1957.

Priani 1957

Pirani, Bondi and Robinson tried to find a proper definition of gravitational waves by drawing analogies to electromagnetic waves in 4-d spacetime and found additional solutions that fulfilled their newly defined conditions for gravitational waves [5, 6, 7].

Trautman 1958

The most mathematically rigorous definition came however in 1958 by Andrzej Trautman, who published two papers about conditions for metrics to be associated with gravitational radiation independent of the choice of coordinates [8, 9].

Bondi (1957), Pirani & Robinson (1959)

Robinson & Trautman 1962

His findings also revealed that the known linearized waves can be seen as the asymptotic limit of exact gravitational waves emitted by some bounded sources. Later Trautman and Robinson additionally found a large class of gravitational waves solutions that satisfied Trautman's rigorous conditions, carry energy and have a closed surface as wavefront (corresponding to a localized source) [10, 11].

1.2 Modern approaches to calculating GW waveforms

Today, a set of different techniques is used to calculate waveforms of gravitational waves which can be compared to measured data. Paradoxically, the mathematically most rigorous calculations are done by numerical simulations. While this can lead to numerical errors, it usually involves the full, non–linear EFE to calculate the metric, unlike the finite or weak-field approximations. Prof. Alessandra Buonanno from the Max Planck Institute for Gravitational Physics in Potsdam once[1] referred to the numerical method as being "the slow but exact method" compared to the analytic "fast but approximate" way of calculating gravitational waves. Most other methods to calculate waveforms that are used and explained in literature like [12, 13, 14] usually involve linearized gravity or some other kind of (post-Newtonian) series approximation. Calculating waveforms for comparison with measurable signals is still an active research field where new models and techniques with varying precision, computational time and ease of use are being investigated.

One way to model a waveform that could be measured by gravitational wave detector is by using the linearized EFE (also referred to as the weak-field approximation). In this approximation, the linearized EFE given by

$$\Box \bar{h}_{\mu\nu} = -\frac{16\pi G}{c^4} \bar{T}_{\mu\nu} \tag{1.13}$$

are solved using the retarded Green's function $G(\underline{x} - \underline{x}') = -1/4\pi |x - x'| \cdot \delta(t - t')$ of the Laplace operator. The resulting weak-field (gravitational wave) for a given dynamic mass distribution $\bar{T}_{\mu\nu}$ can be calculated by convolution with the Green's

---

[1]During a lecture that I visited during my Ph.D. studies in 2022.

function (with a more detailed explanation for the Green's function in section 7.3) via

$$\bar{h}_{\mu\nu}(t, x) = \frac{4G}{c^4} \int d^3x' \frac{1}{|x - x'|} \bar{T}_{\mu\nu}\left(t - \frac{|x - x'|}{c}, x'\right) . \qquad (1.14)$$

The $\bar{T}_{\mu\nu}$ describing the dynamics of a physical system (the moving masses) are generally unknown and what physicists try to deduce from (some) measured gravitational wave signals. For practical purposes, the expression inside the integral (1.14) is expanded into a series of spherical harmonics (different Bessel functions) and by applying additional gauges and some further approximations, the integral is reduced to leave only a small set of parameters (like masses, initial/current separation distance and speeds). For a given set of these parameters, the linearized gravitational waves received by a far away observer can then be calculated. A detailed calculation of this can be found in [12, Chapter 3]. Since the initial assumptions for linear gravity assume only a small disturbance (a "weak" field), waveforms from sources with strong dynamic fields like two coalescing binary stars are typically calculated using more elaborate techniques like the "post-Newtonian" expansion where the metric tensor is expanded further into additional coefficients beyond the linear expansion. [12, Chapter 5].

1.3 Typical sources of gravitational waves

The first indirect detection of gravitational waves was done by observing the periodic bursts of a binary pulsar system discovered by Hulse and Taylor in 1974 [15]. The decay in frequency matched the theoretical prediction of the energy emitted by gravitational waves [16] for such a binary mass system, giving evidence for the existence of gravitational waves. The discovery of the pulsar lead to Hulse and Taylor being awarded the Nobel Prize in Physics in 1993. The first direct detection of gravitational waves was made on the 14th September 2015 by the LIGO detectors with the signal named GW150914. Related to this, in 2017, Weiss, Barish and Throne received a Nobel Prize "for decisive contributions to the LIGO detectors and the observation of gravitational waves."

Besides measuring and thus proving the existence of gravitational waves a more detailed analysis of their exact shape can yield new insight on their sources, the space-time between sources and detector, and the equations and fundamental physics describing it. Figure 1.1 shows the amplitude spectral density (ASD) of some of these expected sources and proposed detectors that are briefly described below:

Coalescing binaries

Binary systems of dense and heavy objects like black holes (BH) or neutron stars (NS) are the most common and best understood sources of gravitational waves [17, Section 3.4.1]. Especially black holes, which are almost completely characterized by their mass and angular momentum offer a good theoretical reference without additional internal dynamics. Figure 1.2 shows the shape of the characteristic "chirp" signal that such a system would emit around the time of its coalescence. In Figure 1.1, these chirps lead to the marked "compact binary inspirals" region. Any binary system, like the binary pulsar Hulse and Taylor discovered, emit gravitational waves,
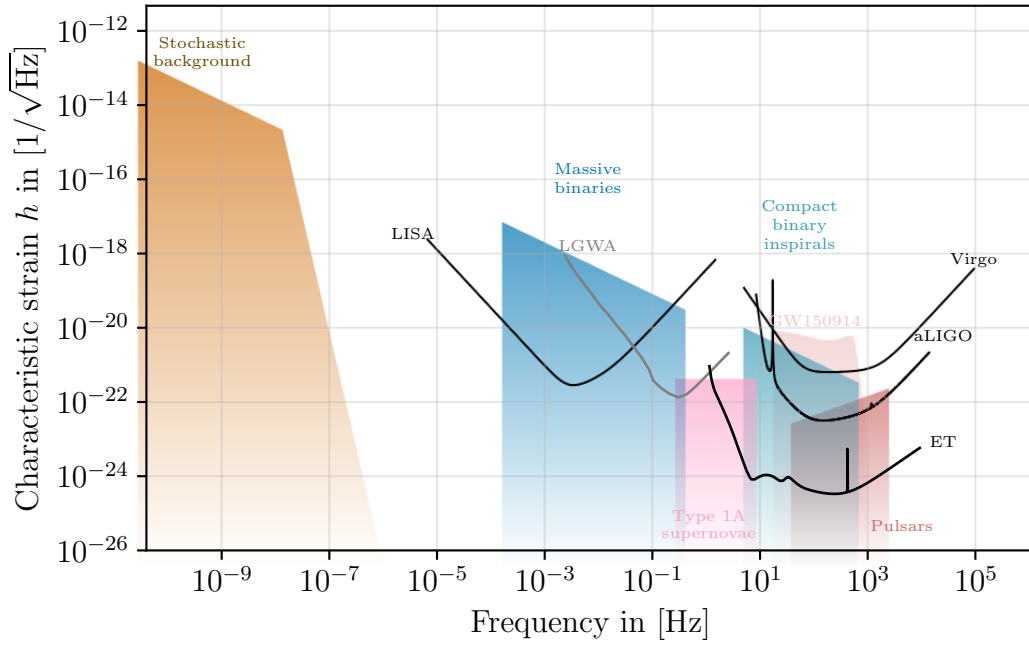
Figure 1.1: ASD plot of some expected gravitational wave signals and sensitivity curves of gravitational wave detectors. Depending on the publication, the exact values for these sensitivity curves can vary by over an order of magnitude, with the values shown here corresponding to the highest design sensitivity proposed. LIGO is an exception here as its most recent observation run is better by an order of magnitude compared to aLIGO's initial design sensitivity.

not only shortly before their coalescence. These signals are however much weaker due to the larger distances between the masses and at much lower frequencies in the order of the orbital periods. For very large and (or) massive binaries, some of these signal are expected to become measurable in the mHz region marked as "Massive binaries" in Figure 1.1.
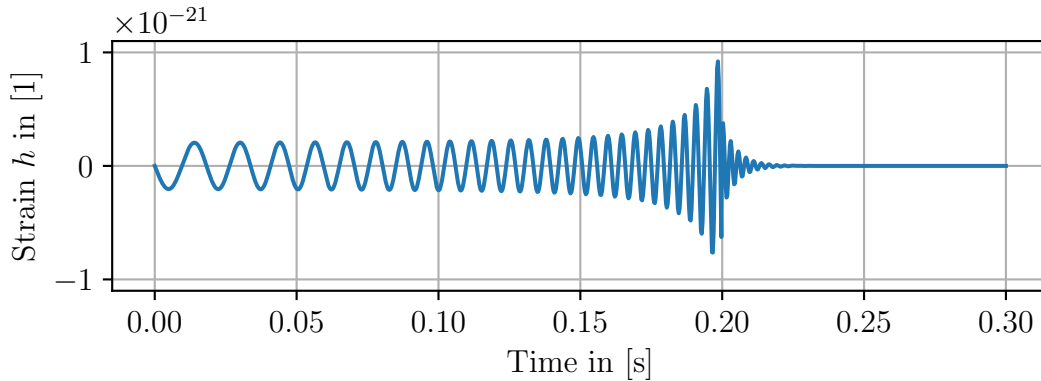
Figure 1.2: Example sketch of a characteristic signal. The signal in this plot was not calculated from a proper multipole formula but serves only as reference of the shape of a binary inspiral GW signal. Up to the merge event at around 0.2 seconds, the frequency and amplitude of the emitted waves increase. From the merge onward, there is a characteristic "ring-down" of the newly formed single massive object.

Asymmetric rotating neutron stars

Rotating neutron stars have also been suggested as sources of continuous gravitational waves. An individual, fast rotating neutron stars that is not perfectly spherically symmetric, would lead to non–vanishing higher order multipole moments in the metric. The frequency of the emitted gravitational waves would then be proportional to the rotational frequency of the star, leading to almost continuous gravitational waves in a frequency range of up to 1 kHz. By observing the gravitational waves of a pulsar together with the received electromagnetic (EM) signal, which also helps to pinpoint the source, physicists hope to gain insight into the internal structure of neutron stars and their corresponding equations of state [17].

Supernova explosions

Supernova explosions are another potential candidate for significant gravitational waves emissions that might be detectable in the future and would also produce an electromagnetic signal that could be observed with optical telescopes on Earth. As a relatively sudden and fast process (involving large accelerations), where large masses move in a non-spherical-symmetric way, the multipole expansion of the metric (as given in (1.14)) would yield significant contribution from the quadrupole- and higher order moments which would lead to strong detectable gravitational waves. Calculations of supernova explosions are however difficult, and it is uncertain just how "asymmetric" the process actually is and how strong the possible gravitational waves might be [17, Section 3.4.2]. The emitted gravitational wave would yield another way, independent of the measured EM emissions, to verify the mass and distance of the supernova and its use as standard siren in astronomy.

Stochastic background

Besides individually resolvable signals, different kinds of sources are expected to produce stochastic backgrounds, creating regions where individual signals cannot be isolated from the abundance of overlapping signals. The low frequency region labelled "stochastic background" in Figure 1.1 corresponds to the expected background signals caused by the hypothesized supermassive black holes in the center of galaxies which could create signals with a large strain at frequencies of $\sim 1\,\mathrm{yr}^{-1}$ [18]. With a much smaller strain, smaller binaries also create a "background" noise which could at some point limit the detection of individual signals. Mapping the exact shape of this background does however still yield insights into the population and distribution of these sources. Besides an overabundance of known signals causing a stochastic background, there are also more exotic processes theorized to have caused a cosmological stochastic background (in analogy to the cosmological microwave background) which is assumed to be in a similar region as the aforementioned supermassive black hole stochastic background.

Other cosmological sources

Besides these cases of large moving masses, there are also theorized processes like i.e. a 1st order phase transition in the early universe that could have emitted gravitational waves. The scales and frequencies for these events is expected to be very

different from the more localized sources above and their measurement could yield insight into new physics [19].

1.4   Gravitational wave detectors

First experimental GW Detectors: Joseph Weber and his bars (1960s)

The beginning of the experimental search for gravitational waves and the first gravitational wave detector is generally attributed to Joseph Weber in the 1960s [20]. Weber's concept was to use "resonant bar detectors," large rigid aluminum bars as mechanical resonators for passing gravitational waves. A passing gravitational wave would, in principle, excite the mechanical modes of the resonator and a bonded piezo-crystal would convert this stress into measurable current. The geometry of these bars (and their elastic modulus) set the resonance frequency around which the conversion between gravitational energy to electric signal would be strongest. In his laboratory, Weber used a set of $1.53\,\mathrm{m}$ long aluminum cylinders, with resonance frequencies around $1660\,\mathrm{Hz}$, where he expected to see transient signals from waves generated during supernova explosions [20]. The search for gravitational waves using this kind of mechanical oscillators remained however unsuccessful [19] (claims of detected GW signals by Weber could not be verified by other research groups doing similar experiments).

Planing of interferometric detectors (1970s)

From the 1970s onward, researchers began to consider alternative concepts like laser interferometers as detectors for gravitational waves. The basic principle is that a passing gravitational wave stretches and contracts the distance between two test–masses (two mirrors) and by interfering a light beam traveling along two different paths, this change can be made visible. One general advantage of using interferometers compared to Weber-type bar antennas for detection is their inherent broadband detection bandwidth (compared to the narrow detection bandwidth close to the resonance frequency of the bar antennas). Some of the first experimental interferometer setups to detect gravitational waves were done in München (Germany) starting from the 1970s. Ranging from table-top setups, and going over a 3-meter prototype up to a 30-meter prototype in the 1980s [19].

1990s – today

The work of this early research on gravitational wave detectors culminated in the planning and eventual building of multiple larger scale interferometers starting with German-British gravitational wave observatory GEO600, and leading to the LIGO dectector in the USA, the Virgo detector in Italy and the KAGRA detector in Japan that operate today.

1.4.1 : Laser interferometers as gravitational wave detectors

The basic principle of a Michelson Interferometer as gravitational wave detector is straight forward: The light emitted by the laser ($\propto A\sin(\omega_0 t - \vec{k}\vec{x})$) travels along (null-) geodesics within both arms of the interferometer and recombines at the beam-splitter. The light measured at the photodiode is proportional to $\sim \cos(\varphi)$ with $\varphi = 2\pi(L_2 - L_1)/\lambda_0$ with L_1 and L_2 as proper distances between the beam-splitter and the end-mirrors of both arms (and λ_0 as the light's wavelength). A passing gravitational wave leads to changes of this proper length and thus to a change of measurable power proportional to the phase φ.
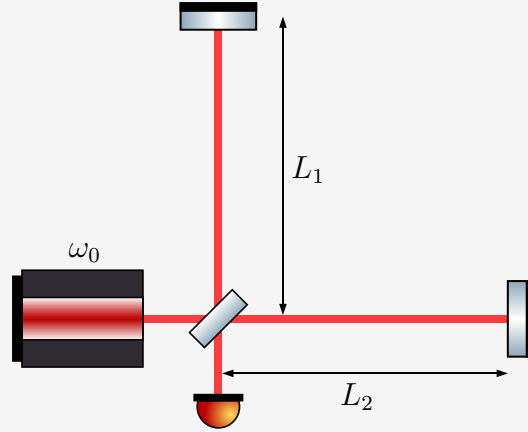
Figure 1.3: Sketch of a Michelson Interferometer with two perpendicular arms of length L_1 and L_2. For GWD, the interferometer arms are keep close to the same length $L_0 \approx L_1 \approx L_2$ to reduce certain types of noise (laser frequency noise and shot noise) while passing gravitational waves cause a measurable arm length change $\delta L := L_2 - L_1$.

While the effect of the gravitational wave is measured as differential length by the detector, the signal and the detectors' sensitivity are usually expressed as dimensionless "**strain**" factor $h := \delta L/L_0$ which is the ratio of the differential length $\delta L = (L_2 - L_1)$ relative to the initial absolute total length L_0. As long as the wavelength of a passing gravitational wave is larger than the arm length, the "stretching" of space will sum up along the interferometer arms to a larger absolute length increase, which is why longer interferometer arms directly improve the sensitivity of the interferometer for gravitational waves.
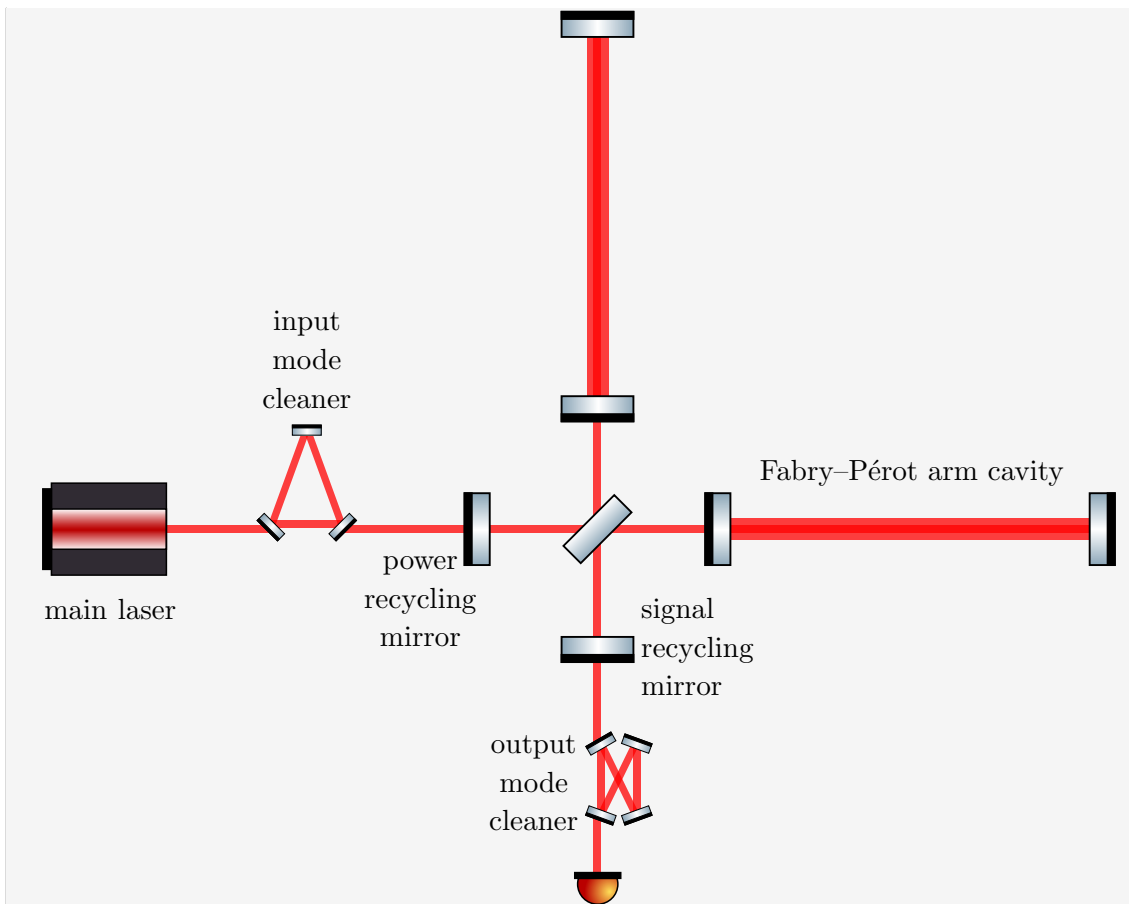
Figure 1.4: Schematic overview of the interferometer setup used in current ground–based gravitational wave detectors like LIGO. Some additional cavities and optical setups i.e. related to the generation of the squeezed light, while crucial for achieving a high performance of the detector, are not shown here.

Current GWDs employ more elaborate optical setups shown in Figure 1.4, where the arms themselves act as (Fabry–Pérot) cavities, and additional mirrors at the input and output port of the interferometer, used for "power-" and "signal-recycling," similarly increases the sensitivity of the whole detector. The additional cavities before the input and after the output are used for preparation and filtering of the laser light to remove control signals and unwanted noise (i.e. higher order laser modes) from the signal. The use of so-called "squeezed light" explained further below, further reduces noise from certain quantum effects of the light.

1.4.1   The first large interferometric detector: GEO600

The earliest (but also the smallest) of the large scale interferometric GW detectors was the German-British gravitational wave observatory named GEO600. While initially planned to be larger, due to cost reasons, its final design was reduced to 600 m long arms whose construction started in 1995 and finished in 2002 [21]. While the detector is too small to detect common GW signals directly (with an optimistic design sensitivity of $h \sim \mathcal{O}(10^{-20})$), its main objective was to advance the instrument development and improve the techniques for future and potentially

larger GWDs [19]. E.g. GEO600 pioneered techniques like "squeezed light" [22] to improve the sensitivity beyond the "standard quantum limit" (mentioned in the next section) or the use of a signal recycling cavity which were first tested in GEO600 before they were implemented in the larger detectors like LIGO [23].

GEO600 also first employed and tested an elaborate seismic isolation system which isolated the optical components from the seismic motion of the ground [21], which is sketched out in Figure 1.5. It consists of active and passive noise filters in form of geophones (ground motion velocity sensors) and Piezo actuators as active part, and a multi-stage pendulum suspension which passively dampens the test–mass motion above its resonance frequencies. (The same suspension design tested here was later used for the two LIGO detectors in their initial setup [24]). The basic principle of such seismic isolation systems is also explained in more detail below in section 2.2.
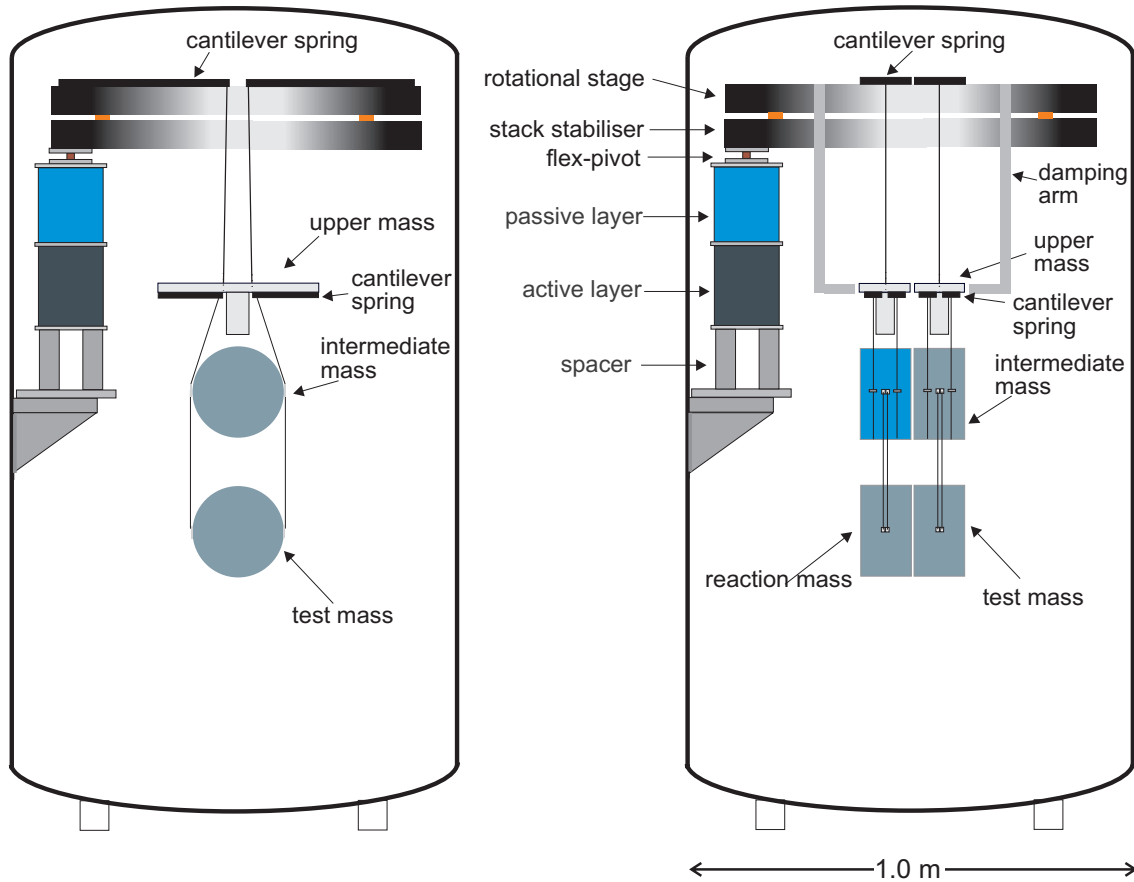
Figure 1.5: The GEO600 main optics suspension (Beamsplitter and End-Mirrors of the Michelson setup.) [21]

1.4.2   LIGO

While the GEO600 detector project had to struggle with a lack of funding in the early 1990s, the US American NSF approved the Laser Interferometer Gravitational-Wave Observatory (LIGO) project proposal in 1991. The LIGO project consists of two identical detectors build in the USA: LIGO Hanford in Washington and LIGO Livinigston in Louisiana. The LIGO detectors have an arm length of 4 km and are the largest gravitational wave detector build to date. The LIGO project was planned with the intention to start using known and proven techniques for the initial LIGO

design and improve upon them over time with several major updates. After the initial LIGO measurement campaign ended in 2007, the LIGO detectors received a first upgrade called "enhanced LIGO" from 2007 onward (improving the laser source and changing to a DC readout system) and a second upgrade from 2011-2014 [23] called "advanced LIGO" (aLIGO) which culminated in the world's first detection of a gravitational wave signal in 2015 [25].

Compared to the smaller and simpler GEO600 setup, advanced LIGO runs as a Fabry–Pérot-Michelson Interferometer with the arms operated as Fabry–Pérot Cavities, leading to a higher laser power in the arms and higher sensitivity. Currently, at the "advanced LIGO plus" (or A+) stage, LIGO continues to run with ongoing improvements and is in its 4th observation run planned from 2023 to 2025. In the recent "O4a" observation part from 2023 to 2024, LIGO detected an average of ~ 100 signal alerts per year. Most of these signals are attributed to transient events like BB, BN or NN inspiral merging events (with B: black hole and N: neutron star) [26].

Besides the two LIGO detectors, there are two more "second generation" gravitational waves detector: The "Virgo interferometer" built in Italy (3km long arms) and the "Kamioka Gravitational Wave Detector (KAGRA)" built in Japan (also with 3km long arms). There are also several other GWD concepts in various implementation stages.

1.4.3   The Virgo detector

The Virgo detector was the first larger (3km arm length) detector built in Europe and started out as a joint project between the French Conseil Nationale de la Recherche Scientifique (CNRS) and the Italian Istituto Nazionale di Fisica Nucleare (INFN)[27]. It evolved at the same time as the LIGO project, with Virgo's initial construction finished in 2003. It joined during advanced LIGO's second observation run [28] and it adds to the global network of GWDs, allowing for a precise triangulation of detected signals by being in a different geographical location compared to the two LIGO detectors (which are, on a global scale, relatively close to one-another).

1.4.4   The KAGRA detector

The Kamioka Gravitational Wave Detector (KAGRA) is the most recent of the 2nd generation gravitational wave detectors and became first operational in 2020. Compared to LIGO and Virgo, it has some novel features like being built underground, and it employs cryogenic cooling of its optics to reduce thermal noise. Because of the new technology development in KAGRA, it is sometimes also referred of being a 2.5 generation detector. The KAGRA has however not yet reached its design sensitivity [29] and its new technology is still being worked on.

1.4.5   The Laser Interferometer Space Antenna (LISA)

Since before the construction of the first 2nd generation ground-based gravitational wave detector even began, the concept of a space-based interferometer setup to similarly measure the effects of gravitational waves has been worked on. Submitted to and approved by the European Space Agency, this work resulted in the "Laser

Interferometer Space Antenna (LISA)" mission which is currently being worked on and planned to launch in 2035.

The LISA mission consists of 3 satellites (sketched in Figure 1.6) that will make up a triangular interferometer with 2.5 million km arm length. The long arm length and the nonexistence of any seismic and Newtonian noise promise a very high sensitivity, especially in the low frequency region that is impossible to achieve for a ground–based detector. The working principle of LISA differs however from ground–based detectors in several key points. Unlike ground–based interferometers, the arm length of LISA cannot be held stable as the satellites drift and move with speeds up to ~ 50km/h. Instead of a constant signal (in case of no gravitational wave), LISA employs a heterodyne interferometer scheme where beat–frequency is measured, and only relative length changes are resolved (in the frequency domain). To deal with ~10 orders of magnitude of laser frequency noise, a core feature of LISA is the so-called "time-delay interferometry" [30] to achieve a high precision readout.
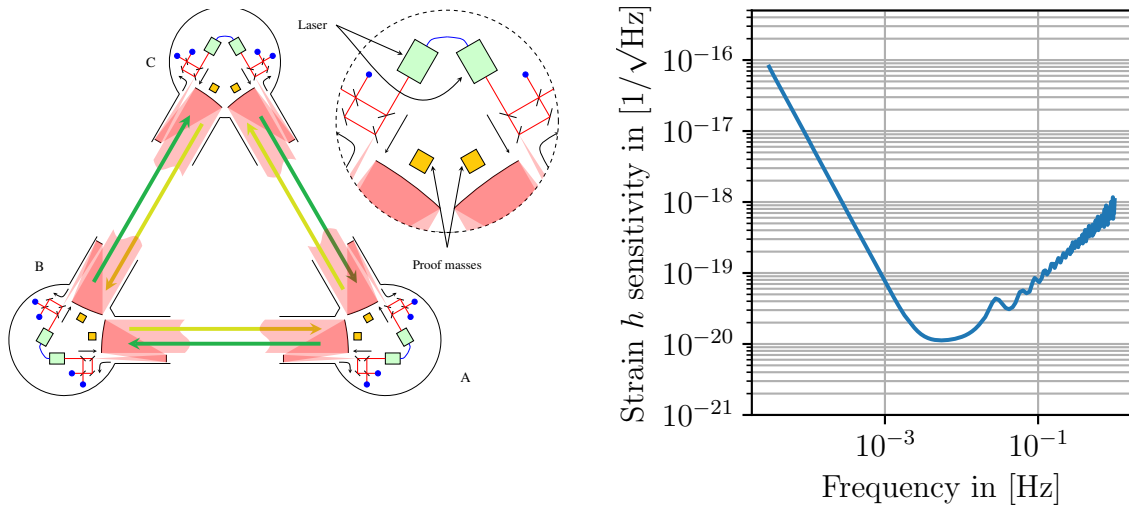
Figure 1.6: Left: Sketch of the LISA satellites. Right: Design sensitivity of LISA. Source: [31]

1.4.6 Pulsar timing array

Not a singular detector, but a collaboration that tries to detect gravitational waves by measuring the timing of radio signals emitted by pulsars is the "International Pulsar Timing Array" project [32]. Pulsars are rotating neutron stars that emit a periodic signal when observed from the earth. (I.e., when the rotational axis differs from the magnetic field axis and earth lies within the resulting radiation cone.) The timing of these pulses is very uniform (with variations below a µs). In theory, a gravitational wave passing between the pulsar and earth would lead to variations of the path length and consequently in the timing between pulses. However, since space is expected to be filled with gravitational waves from numerous sources with varying strength, directions and points of origin, no direct detection of a single timing delay attributed to a single gravitational wave can be made. Instead, the project measures

the seemingly irregular (noisy) time delays from many known pulsars and correlate the measured signals. The idea is that signals travelling through the same space also experience the same gravitational wave background noise which would lead to a seemingly noisy, but correlated measured signal.

In 2023, the American "NANOGrav" collaboration announced the existence of a gravitational wave background noise measured with a 3σ significance from a dataset of 67 pulsars over a period of 15 years [33]. Further measurements with more data are expected to improve the significance and the detailed results of this in the future.

1.4.7   The Lunar gravitational wave antenna concept

Another (not necessarily interferometric) concept is the 2021 published "Lunar gravitational wave antenna" by Harms et al. [34, 35]. It is, in some sense, an evolution of an earlier concept of Weber as it intends to use the moon as a large resonant body (with much less seismic noise than on earth) to detect gravitational waves. The concept suggests that placing an array of very precise seismometers or gravimeters on the moon's surface could detect relative changes when a passing gravitational wave induces seismic waves (similar to the resonant bar antennas from Weber). An interesting feature of the concept is that the expected sensitivity could bridge the gap between space–based detectors (LISA) and ground–based detectors and provide a higher sensitivity in the region of 0.1 – 1 Hz. So far the project is in an early conceptual stage and has not yet been adopted by any space Agency.

1.4.8   3rd generation ground–based gravitational wave detectors

While the second generation ground–based gravitational wave detectors are currently running and regularly detecting new events, the planning for future "3rd generation" detectors is already underway with technical designs being specified and potential sites being scouted. On the US American side there is the concept of the "Cosmic Explorer" [36] which is planned to be 10x larger than LIGO with the initial setup taking a conservative approach of using already tested technologies and potentially employing new techniques (like different laser wavelength or cryogenic cooling) at later expansion stages like LIGO did with enhanced- and advanced LIGO.

On the European side, there are plans for the so-called "Einstein Telescope" [37, 38], an underground, 10 km long detector employing several new techniques right from its initial design to improve upon the limits set by currently used techniques.

1.5   The Einstein Telescope

The Einstein Telescope (ET) is a planned 3rd generation gravitational wave detector to be built in Europe. While it follows the same basic working principle as the current 2nd generation detectors, it aims to be more sensitive by at least one order of magnitude (and several orders of magnitude at low frequencies $\sim 3\,\mathrm{Hz}$) and to have a wider bandwidth (a larger frequency range) to detect signals, compared to the current most sensitive detectors.
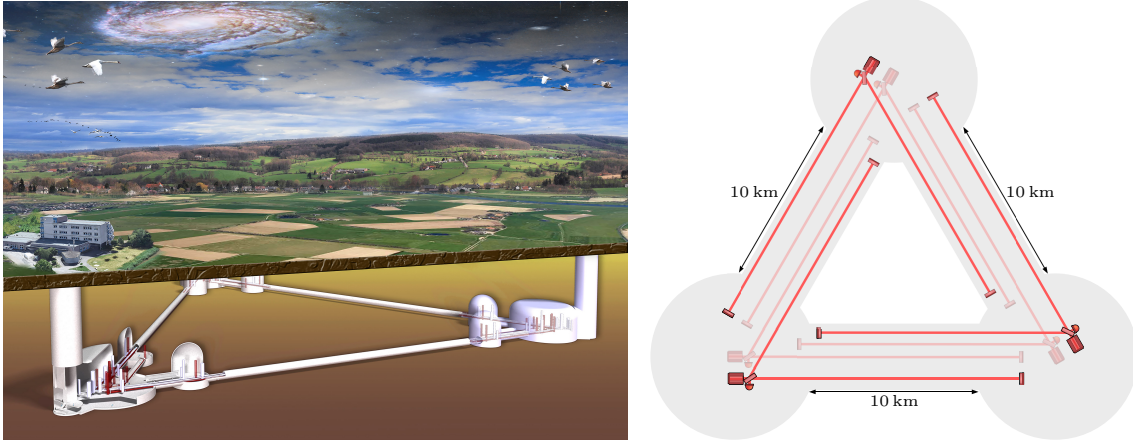
Figure 1.7: Left: Artistic image of ET built in its proposed triangular geometry underground. Right: Simplified sketch of the interferometer topology of the proposed triangular geometry. Each corner station supports two interferometers, one cryogenic for "low" frequencies and one at room temperature for "high" frequencies leading to a total of 6 interferometers as sketched in the figure. The real proposed setup is however more complex and includes additional optical paths inside the tunnels corresponding to different cavities which are not sketched here.

Key features

From its initial design, ET is planned to include several novel features that have not yet been realized in the current ground–based gravitational wave detectors. Besides the techniques tested and integrated in the most recent 2nd generation detectors upgrades, ET's new features would be:

1. **Underground**
   ET is designed to be built a few hundred meters underground to reduce the effect of (surface) seismic waves, the seismic disturbance from anthropogenic/human influence (which is more significant as ET would be built in the rather densely populated Europe compared to the remote LIGO locations), and the correlated Newtonian noise which is generally weaker in the homogeneous (and stiff) environment underground.

2. **Cryogenic + Xylophone design**
   To reduce thermal noise (from mechanical suspensions and light absorption / mirror coatings) that is expected to limit future GWD sensitivity, parts of ET are planned to be cooled down to cryogenic temperatures [38]. Besides the
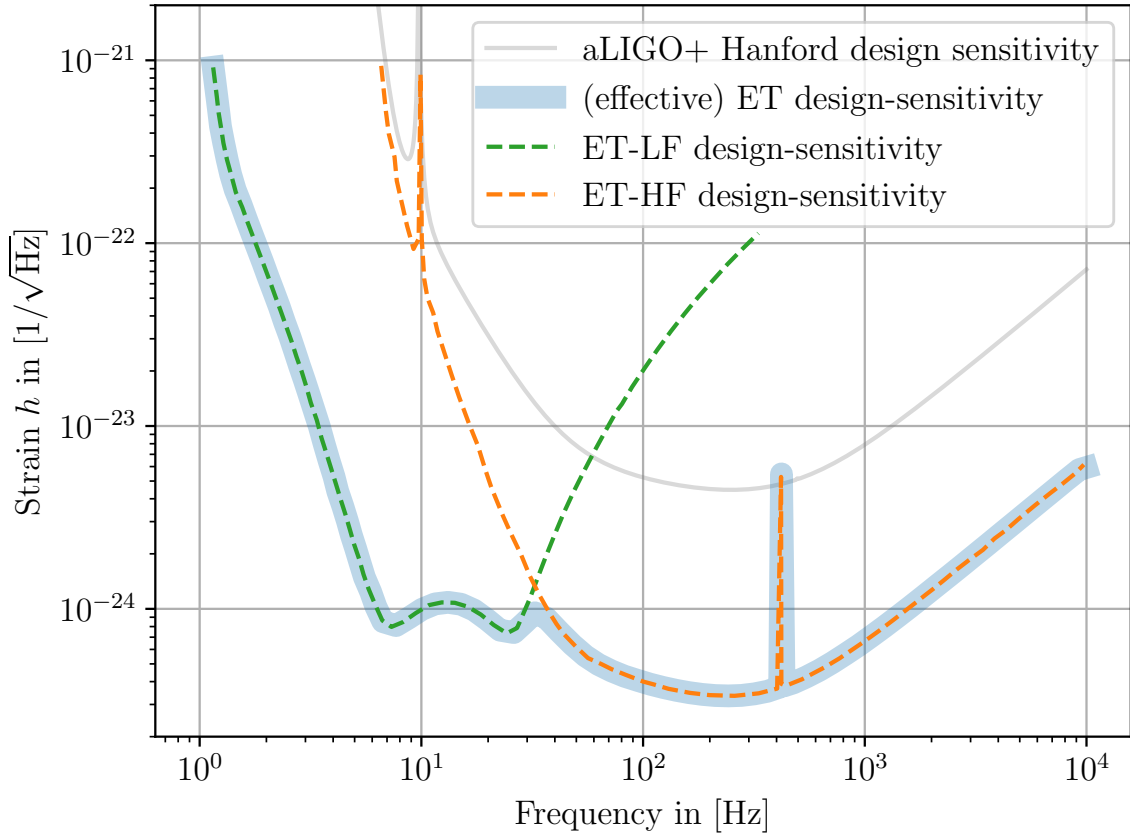
Figure 1.8: Plot of ET's design sensitivity. ET is planned to operate two separate 'detectors, one optimized for low-frequencies (ET-LF, dashed green line) and one for high-frequencies (ET-HF, dashed yellow line). The visible peaks in the sensitivity curve of ET-HF are projected resonance frequencies of the mechanical suspension (including "thermal modes") based on Virgo's design. The actual resonance and control signal peaks will likely change for the final design. Source: [38]

cooling infrastructure, running cryogenic will require new, actively cooled suspensions systems to hold the optics. Due to the interferometer arms effectively being Fabry–Pérot cavities, current detectors operate with relatively high optical power in their arms (~ 200 kW in LIGO [39]) and even the currently best available low absorption coatings would cause the mirrors (and suspension) to significantly heat up when running cryogenic. To mitigate this problem, ET will be split into two interferometer "parts" being referred to as "Xylophone" design. One interferometer (for low frequencies) will run cryogenic / being cooled and only has a moderate optical power increase in the interferometer arms. And a second interferometer (for high frequencies) will run at room temperature like the current 2nd generation detectors.

3. **Multi–interferometer Detector**
   The currently favored design for ET features a geometry of an equilateral triangle instead of the "L" shape of the 2nd generation detectors. A sketch of this geometry can be seen in Figure 1.7. ET would then operate 2 interferometers (a high frequency one and a low frequency one) for each of the 3 corners of the geometry leading to a total of 6 interferometers. The main argument for this geometry is twofold: As drilling the tunnels is the main cost factor for ET, it

would allow for 6 interferometers in total, compared to only 2 interferometers in an "L" shape geometry, for the cost of only 3 tunnels (compared to 6 tunnels when building 3 L's). Secondly, by having multiple interferometers, the detector would have additional redundancy of the signal, a better sky localization of the signal and no, or at least less of a "blind spot" compared to a single interferometer. It should be noted however that the triangular geometry is not yet fully decided and alternative concepts of multiple individual "L" shaped detectors are still being discussed. [38]

4. **New seismic isolation systems for sensitivity at a few Hz**
   One of the design goals of ET is to be sensitive in the region from $\sim 1\,\mathrm{Hz}$ to $\sim 30\,\mathrm{Hz}$, which the 2nd generation (and current plans for the initial Cosmic Explorer [36]) are not. To achieve this, ET will employ new suspensions to reduce the seismic and control noise and new techniques to remove the Newtonian noise from the measured signal.

The science case of ET

Being more sensitive by an order of magnitude in the high frequencies and many orders of magnitude at low frequencies (meaning a wider frequency bandwidth) as shown in Figure 1.8, ET will be able to detect more signals with a higher signal-to-noise ratio (SNR) allowing for new discoveries in astrophysics, cosmology and fundamental physics.

On the astrophysical side, the higher sensitivity means a farther reach / a larger detection volume for events, and generally more detections of events. With an expected $\mathcal{O}(10^6)$ [40] detections per year (compared to a current rate of $\mathcal{O}(10^2)$), ET enables the statistical analysis of known GW sources and comparison of astrophysical models with the measured population of binary BH/NS merger events. The higher SNR will allow for a more in-depth analysis of the properties of black holes and neutron stars. Especially in the low frequency regime (shown in 1.8), the higher SNR will allow for an early detection of events (whose signal frequency increases over time) and a much longer observation time. By detecting such events early, ET will enable researchers to align other (directional) telescopes to these sources and perform multi-messenger astronomy by observing not only the emitted gravitational wave but also any electromagnetic signals, neutrinos or cosmic rays emitted by the same source. Besides detecting more of already known types of GW sources with greater accuracy, there are also new types of signals expected to become visible with ET like transient signals from "core collapse supernovae" or isolated rotating neutron stars. Beyond the large number of distinguishable detection, ET could even reach the "stochastic background" of some coalescing binary gravitational wave signal frequencies, which cannot be isolated/distinguished from one another anymore. There is also hope to detect more exotic (stochastic) signals from cosmic inflation, phase transition, or cosmic strings [40] that go beyond standard cosmological models.

Chapter 2

Current limitations of gravitational wave detectors shown by example of LIGO

As this thesis focuses on the development of local displacement sensors as components to improve the sensitivity of current and future gravitational wave detectors, this chapter gives a brief overview of the noise contributions in GWDs (taking LIGO as example), and shows how local displacement sensors can help to improve upon current limitations.

2.1 LIGO's noise budget from the third observation run (O3)

For the currently running second generation GWDs, different noise sources limit the sensitivity at different frequencies. At low frequencies (below ~ 10 Hz), "control noise" from the alignment and length control systems limit the sensitivity for aLIGO. And above that, thermal noise and "quantum noise" (shot noise and radiation pressure noise) are more fundamental limits of the detector's current sensitivity. Figure 2.1 shows the "noise budget" of aLIGO Hanford with all known and measured noise sources affecting its sensitivity. As a short summary, the shown noise contributions are:

1. **Quantum noise**
   "Quantum noise" is the term for the sum of "shot noise" and "radiation pressure noise." By Heisenberg's uncertainty principle, the quantum nature of the photons leads to an uncertainty of their position and momentum. The position uncertainty translates to a varying number of photons (the measured signal) arriving at the photodiode at a given time interval which leads to so-called "shot noise" [17, Chapter 5]. The momentum uncertainty leads to a variation of the radiation pressure acting upon the suspended optics. The resulting force moves the optics and causes a change of the path length, also coupling into the measured readout. The sum of these two unmitigated noise contributions is called the "standard quantum limit." Currently, the detectors operate close to the "dark-fringe" to reduce the effect of this shot noise and employ a tech-
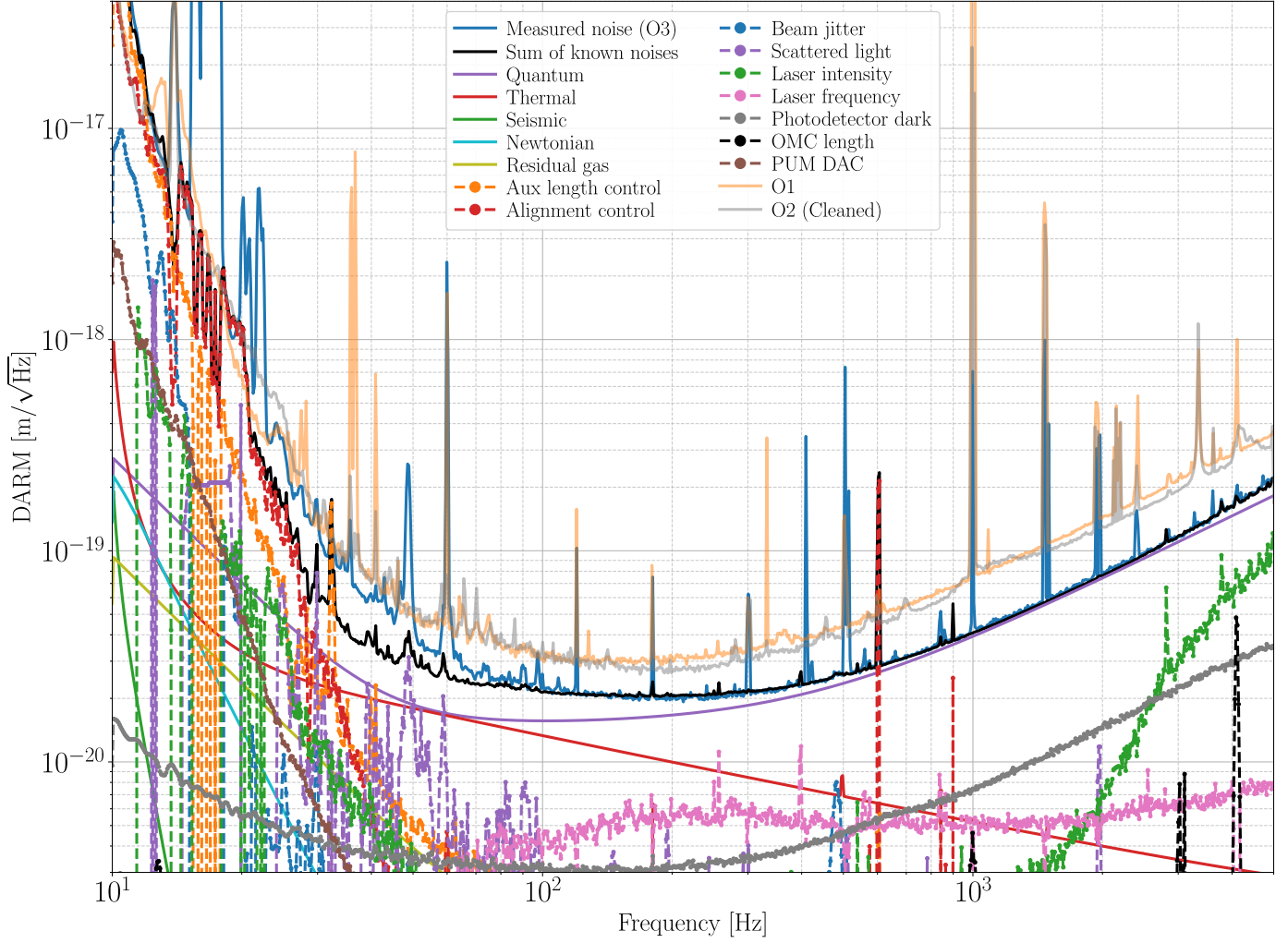
Figure 2.1: Measured detector noise sources during the O3 run of LIGO. The current O4 observation run includes some technical improvements and can be assumed to be slightly better. Source: Cahillane's interactive noise budget [41]

nique called "squeezed light" to reduce this quantum noise below the standard quantum limit (SQL)[22].

2. **Thermal noise**
   The relevant thermal noise contributions in LIGO comes from the Brownian motion of the mirror coatings and the suspension mechanics, which couples as phase noise into the laser beam and limit the sensitivity in the frequency band between 40 Hz and 100 Hz. Besides laser light absorption from the coating, the mirrors are also (thermally) coupled to the mechanical suspension holding them, contributing to the bulk thermal noise of the mirror substrates. Special materials (i.e. fused silica wires) are used to reduce this thermal noise contribution but it remains one of the main limiting noise sources.

3. **Seismic noise**
   Seismic motion couples into the readout since a moving ground will lead to changes of the mirror positions (which stand on / are coupled to the ground)

and thus change the optical path length of the interferometer arms. To block this induced motion and to mitigate the effect of seismic waves, the optics of the interferometer are suspended by special suspensions systems to dampen the induced motion by many orders of magnitude [42]. While the dampening of the seismic motion can generally be improved by scaling these suspensions (making them larger / adding additional suspension stages), the suspension system also adds additional "control noise" and thermal noise which limits arbitrary scaling.

4. **Newtonian noise**
   Newtonian noise is the effective noise coming from the (local) gravitational potential that couples, same as the gravitational wave, into the readout. I.e., a passing seismic wave corresponds to a tiny motion of the ground which in turn causes a small change in the gravitational pull from the ground to the mirrors. This type of noise cannot be shielded against in ground–based GWDs since it acts on the detector same as the signal itself by gravitational interaction. Schemes to reduce the Newtonian noise aim to measure the seismic motion of the ground with high precision and try to subtract the expected local gravitational signal from the full measured signal [43].

5. **Residual gas noise**
   While the vacuum chambers of detectors like LIGO and Virgo are pumped down to a pressure of $\mathcal{O}(10^{-9}\,\mathrm{mbar})$ [44], the remaining (residual) gas particles have multiple effects on the readout described by an effective low frequency (residual gas) noise. The movement of the vacuum chamber walls (which is mainly due to seismic motion) couples to the residual gas, resulting in acoustic waves inside the vacuum chamber. At the suspended optics, this causes a force on the suspended optics and a small displacement / path length change. In between the optics, the pressure fluctuations cause the refractive index to fluctuate which causes the light's optical path length to change resulting in a phase noise for the light itself. (Additionally, the residual gas causes a small viscous damping of the suspension's mechanics slightly reducing their efficiency.)

6. **Aux. length and Alignment control noise**
   A precise alignment of the optics is necessary to archive a high optical contrast, minimize stray light, achieve good mode-matching / mode-filtering and high finesse for the optical cavities. This is achieved by the length and alignment control system(s), consisting of: (a) local sensors, which measure the precise position and alignment of the optics (b) controllers/systems, which take the sensors signals (input) and derive a control signal from it (output), and (c) motors/actuators which act on the optics and "force" them into position. The length and alignment control noise is effectively the self-noise of the used sensor and actuators of these control systems. At low frequencies, these are currently one of the limiting noise sources in LIGO. These noise sources are discussed in more detail in section 2.2.

7. **Beam jitter noise**
   Beam jitter noise is the result of fluctuations of the laser beam position (align-

ment fluctuation) on the optics, where small inhomogenities (point absorbers) on the mirrors cause an effective jitter noise of the beam.

8. **Scattered light noise**
   Scattered light noise is mainly caused by motion of the optics (i.e. due to seismic noise) leading to scattering of small parts of the main beam that couple back into the main beam. It is mainly mitigated by using baffles to capture stray beams and by improving the suspensions to reduce the overall motion of the optics and variations of the stray light coupled to it.

9. **Laser intensity and frequency noise**
   Extensive laser preparations are done to reduce the laser noise as much as possible, i.e., by using input filter cavities to filter higher modes, and control loops to stabilize the laser. The remaining intensity and frequency noise is currently about an order of magnitude below the other limiting noise sources.

10. **Photodiode dark-noise**
    Even when no light hits the photodiode, thermal fluctuations in the photodiode cause a (temperature dependent) noise current. Since it is present without light (when it's dark), this temperature- (and bias voltage-) dependent noise is called photodiode dark-noise.

11. **OMC length and PUM DAC noise**
    Output mode cleaner length noise (OMC length) is the noise coming from side-bands and higher order laser modes that are not properly filtered out by the output mode cleaner cavity. The efficiency of the filter is given by the precise length of the OMC and variation of its length translate to more noise in the output beam of the cavity. PUM DAC noise is the DAC noise from the actuators at the penultimate mass (PUM) position of the mechanical suspensions.

While a longer arm length and other scaling factor are expected to improve the sensitivity by a factor of $\sim 10 - 100$ in the high frequency range for the future (larger) 3rd generation GWDs, additional effort is necessary to bridge the gap of nearly 6 orders of magnitude difference in the low frequency region between the currently most sensitive running aLIGO+ detector and the planned Einstein Telescope, as shown in Figure 2.2.

At frequencies below 20 Hz, LIGO is currently limited by "technical" noise coming from the (global) alignment control and auxiliary length control. More precisely, the self-noise of the local displacement sensors and actuators used for the alignment (and seismic isolation) of the suspended optics limits these systems. New sensors have been suggested to replace the currently used "Birmingham Optical Sensor and Electromagnetic actuator (BOSEM)," which have an accuracy of $\mathcal{O}(10^2 \, \text{pm})$ [47], with more precise (interferometric) displacement sensors. I.e., van Dongen et al. [48] have shown that they could reduce the control noise at LIGOs main optics suspension by replacing the currently used optical shadow sensors (component of the BOSEMs) with their own interferometric sensors called HoQI for the local displacement readout.
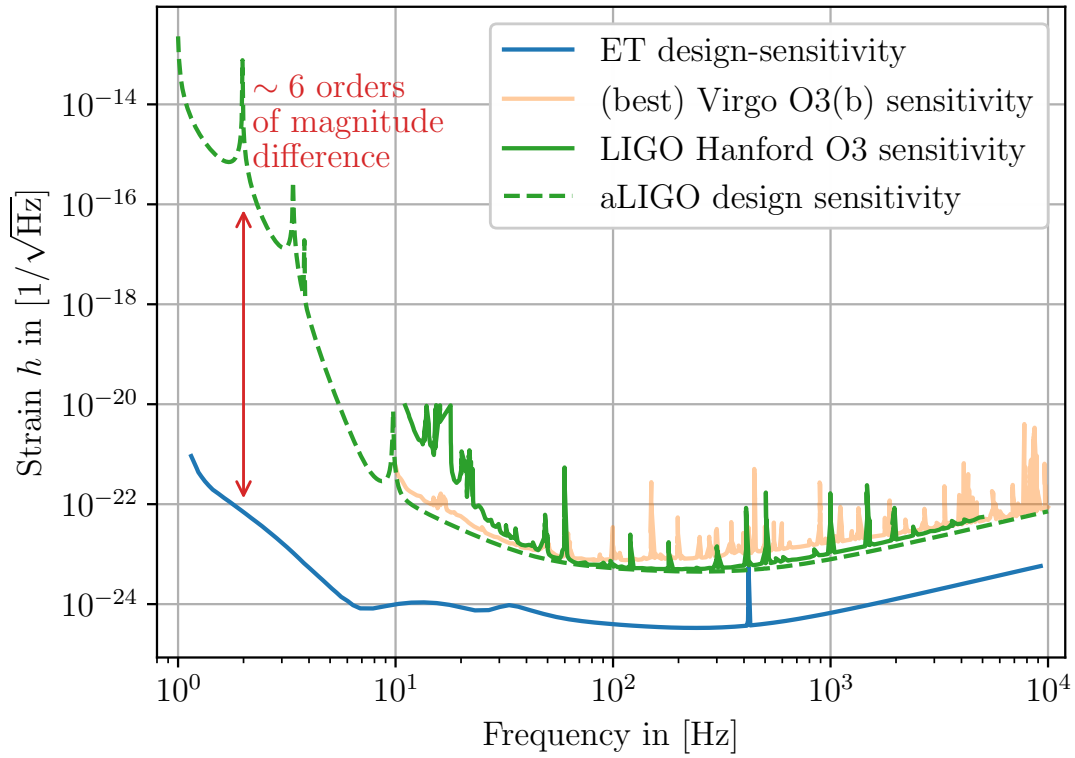
Figure 2.2: Measured detector noise of LIGO Hanford and Virgo during the third observation run (O3) compared to the ET design sensitivity. Source: [45, 38, 46, 41]

Besides the HoQI, there are currently a number of different interferometric local displacement sensors being developed [49], and this thesis focuses on another specific type of sensor concept called "DFMI." To better understand the mode of operation and the challenges of the alignment control and the seismic isolation, a brief overview of their working principle is given in the following section.

2.2   Seismic isolation and local control systems

To isolate the optics in GWDs from the ground motion, a mix of passive and active systems is used. Figure 2.3 shows the schematic display of aLIGO's "Internal Seismic Isolation (ISI)" system, which is a more advanced version of the initial GEO600 main suspension system mentioned earlier.

LIGO's seismic isolation system combines multiple active and passive layers to isolate the suspended optics from the ground motion. On the ground stand seismometers measuring the seismic motion and a "support pier" holding the seismic isolation system. The first isolation layer is the active "Hydraulic External Pre-Isolator" (HEPI) on top of the pier outside the vacuum chamber. It acts as first filter in the frequency range from 0.1 Hz to 10 Hz. Connected to it is the vacuum chamber and in it is the "Internal Seismic Isolation" (ISI) system. The ISI is the second layer of (active) isolation. Its three substages of stiff mechanical structures contain cantilever blade springs, and an array of seismometers, geophones, capacitive position
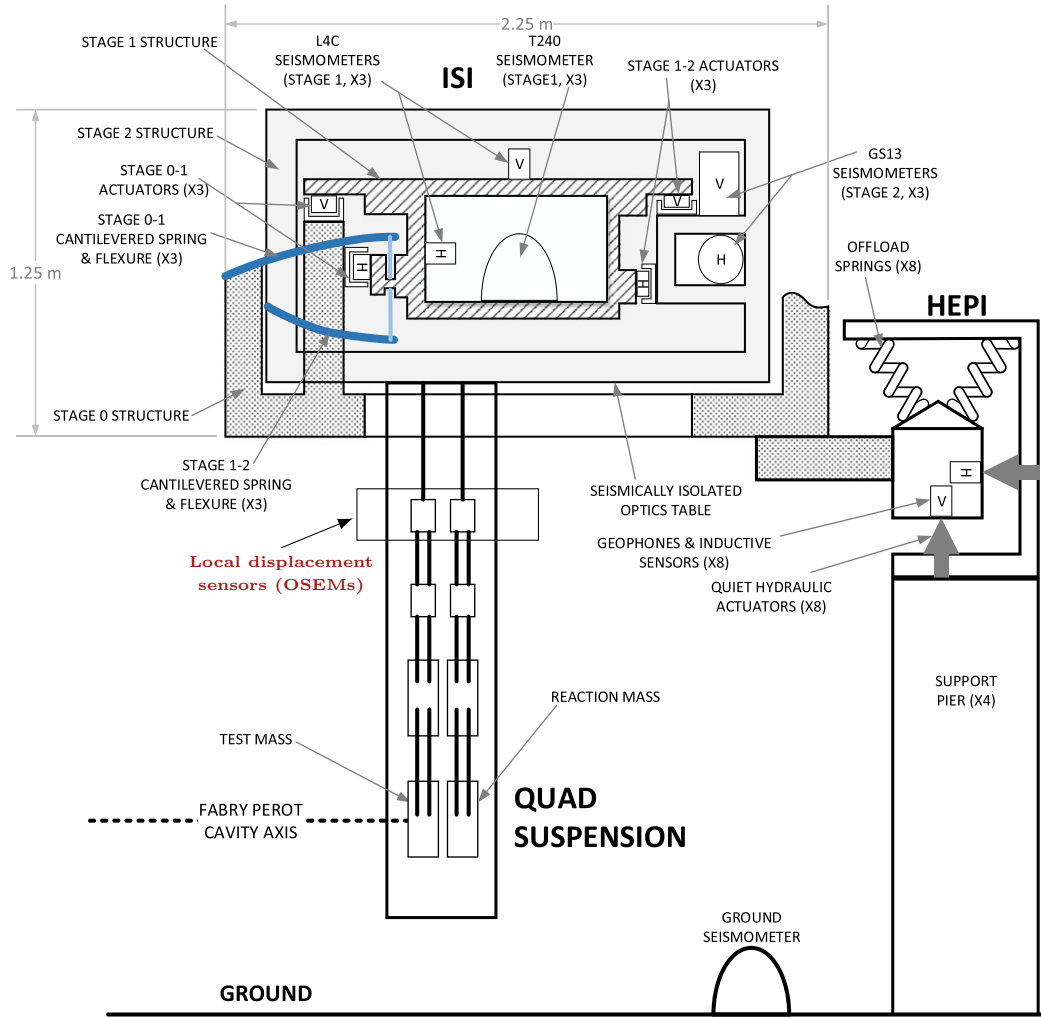
Figure 2.3: aLIGO seismic isolation system for the test–mass optics. Source: [23]

sensors and electromagnetic actuators to further isolate the suspended masses in the frequency range from $0.1\,\mathrm{Hz}$ to $\sim 100\,\mathrm{Hz}$ and help with the rough alignment of the suspension. The third layer of seismic isolation system is the actual pendulum suspension. It consists of two, 4-stage pendulum suspensions, positioned closely next to each other. The frontal mirror at the bottom of the first suspension (the one that is hit by the interferometers laser beam) is also referred to as "test–mass" The mass closely behind it at the bottom of the second suspension is called the "reaction mass." Due to an electrostatic coupling between the test–mass and the reaction mass, the second suspension with the reaction mass, allows for further electrostatic dampening and thermal control through the "compensation plate" mounted to the reaction mass. Keeping test–mass and reaction mass close to each other also leads (together with the residual gas in the chamber) to so called "proximity-enhanced gas damping." At the upper stages of the "quad suspension," so-called "optical (shadow) sensor and electromagnetic actuator" (OSEMs) [23, 24] are used as final active dampening system. Due to the OSEM's control noise, they can however not be used to further alignment, currently limiting the precision of the alignment and leading to added "alignment control noise" from the test–mass tilt-to-length coupling as explained further below.

2.2.1 Passive seismic isolation using pendulum suspensions

The basic principle of pendulum suspensions as "passive" motion isolators is that of the driven harmonic oscillator. I.e., the equation of motion of some suspended mass m at position $x(t) = l \cdot \sin \varphi(t) \approx l \cdot \varphi(t)$ (using the small angle approximation) would be given by

$$m \, l \, \frac{\partial^2}{\partial t^2} \varphi(t) + \beta \frac{\partial}{\partial t} \varphi(t) + m \, g \cdot \varphi(t) = F_{\text{ext}}(t) \tag{2.1}$$

with F_{ext} as external Force, l as fixed length wire, g as gravitational acceleration (causing the linear restoring force), and β as (velocity) dampening factor. Expressing the external force by a dimensionless angular acceleration $F_{\text{ext}} =: m \, g \, \alpha_{\text{ext}}$ allows one to write

$$\implies \quad \tilde{\varphi}(\omega) = \frac{\frac{g}{l}}{\left(\frac{g}{l} + i \frac{\beta}{ml} \omega - \omega^2 \right)} \, \tilde{\alpha}_{\text{ext}}(\omega) \tag{2.2}$$

$$=: \frac{\omega_{\text{res}}^2}{\left(\omega_{\text{res}}^2 + i \frac{\omega_{\text{res}}}{Q} \omega - \omega^2 \right)} \, \tilde{\alpha}_{\text{ext}}(\omega) \tag{2.3}$$

with $\tilde{x}, \tilde{\alpha}_{\text{ext}}$ as Fourier-Transform of x, α_{ext}. The external motion given by (the angular acceleration) α_{ext} is reduced by the transfer function $\tilde{T}(\omega) := \omega_{\text{res}}^2 / (\omega_{\text{res}}^2 + i\omega_{\text{res}}/Q - \omega^2)$, with $\omega_{\text{res}} := \sqrt{g/l}$ as "resonance frequency" and $Q = m\sqrt{gl}/\beta$ as "quality factor." Figure 2.4 shows the transfer function of a harmonic oscillator with a resonance frequency of $10\,\text{Hz}$.

(a) **(b)**
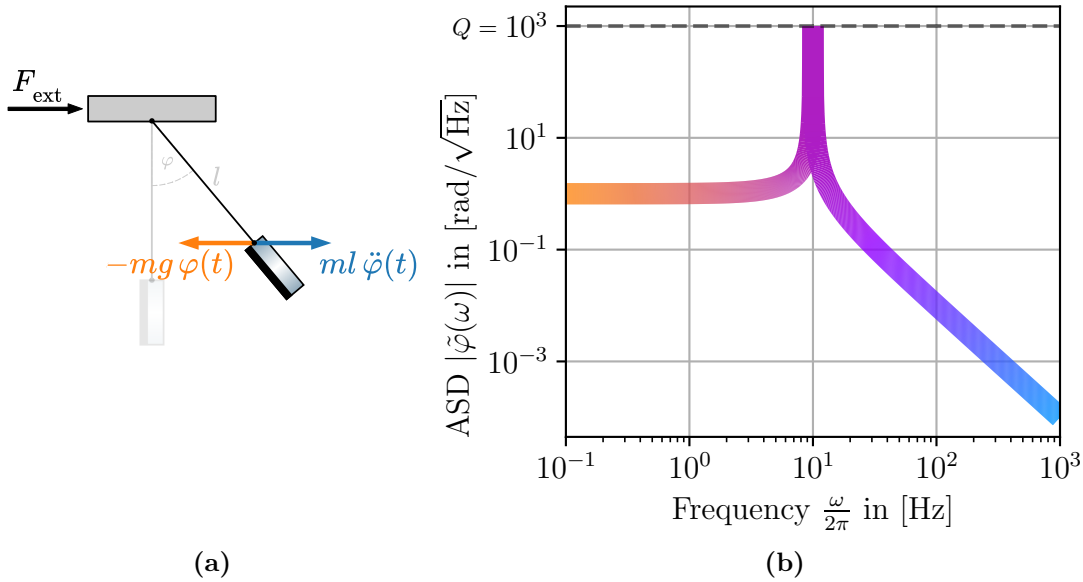
Figure 2.4: (a): Sketch of the forces acting on a suspended mass (mirror) in a pendulum suspension (when using the small angle approximation $\sin(\varphi) \approx \varphi$). (b): Approximate ASD-Plot of the transfer function of a harmonic oscillator with a resonance frequency of $10\,\text{Hz}$ and a quality factor of $Q = 1000$. The color of the plotted line is used as additional indicator of the frequency (x-axis) as used below.

By chaining multiple harmonic oscillators together one can achieve a higher suppression of the ground motion (above the resonance frequencies). I.e., when chaining

(a) Low frequency   **(b)** At resonance frequency   **(c)** High frequency
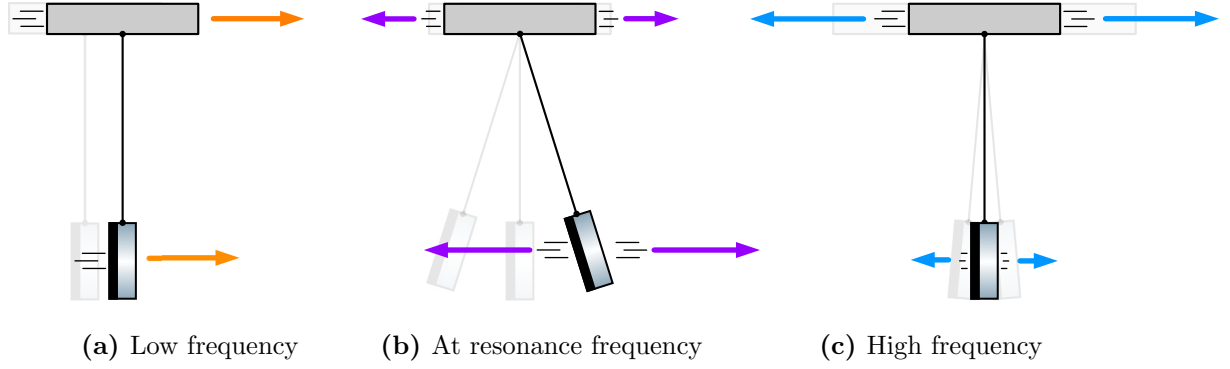
Figure 2.5: Sketch of the induced movement of a suspended mass for different input frequencies. For low frequencies (below the resonance frequency), the suspended mass simply follows the induced movement ($\tilde{T} \propto 1$). Around the resonance frequency, the transfer function (the gain $|\tilde{T}|$) becomes large ($|\tilde{T}| \gg 1$). The energy of the external movement couples strongly into the system and the induced kinetic energy accumulates. The motion of the suspended mass can exceed the external motion. At high frequencies (above the resonance frequency), the system is damped as the external motion does not couple well into the system ($|\tilde{T}| < 1$). The induced motion of the suspended mass is smaller than the external motion.

multiple linear systems together, the total transfer functions would just be the product

$$\tilde{T}_{\text{multi-stage}} \approx \frac{\omega_{\text{res},1}^2}{\left(\omega_{\text{res},1}^2 + i\frac{\omega_{\text{res},1}}{Q_1}\omega - \omega^2\right)} \cdot \frac{\omega_{\text{res},2}^2}{\left(\omega_{\text{res},2}^2 + i\frac{\omega_{\text{res},2}}{Q_2}\omega - \omega^2\right)} \cdot \ldots \qquad (2.4)$$

with $\omega_{\text{res},1}, \omega_{\text{res},2}, \ldots$ as resonance frequencies of the individual harmonic oscillators / stages. As the exact equation of motion of a single pendulum stage is however not perfectly linear, the linearized models of multi-stage pendulums used in GWDs are more complex and contains additional terms to account for these deviations.

LIGO's seismic isolation includes a 4-stage pendulum suspension as well as blade springs (with a linear restoring force described by their Young's modulus), which can be well modelled as such chained harmonic oscillators.
Besides any losses, modelled by more elaborate transfer functions / differential equations than (2.3), active isolation schemes are required to deal with the passive systems' resonance frequencies around which noise couples well into system and can accumulate to potentially large (RMS) motion of the suspended masses.

2.2.2 Active seismic isolation using local sensors and actuators

In active isolation schemes, local sensors are used to measure the (seismically induced) motion of a local mass and counteract this motion with actuators. The advantage of active control systems is that they allow for an (almost) freely designable frequency response up to the precision limits of the used sensors and actuators.
Figure 2.7 shows the schematics of an active control loop in a seismic isolation system. Here the position of a suspended mass is measured with an optical sensor
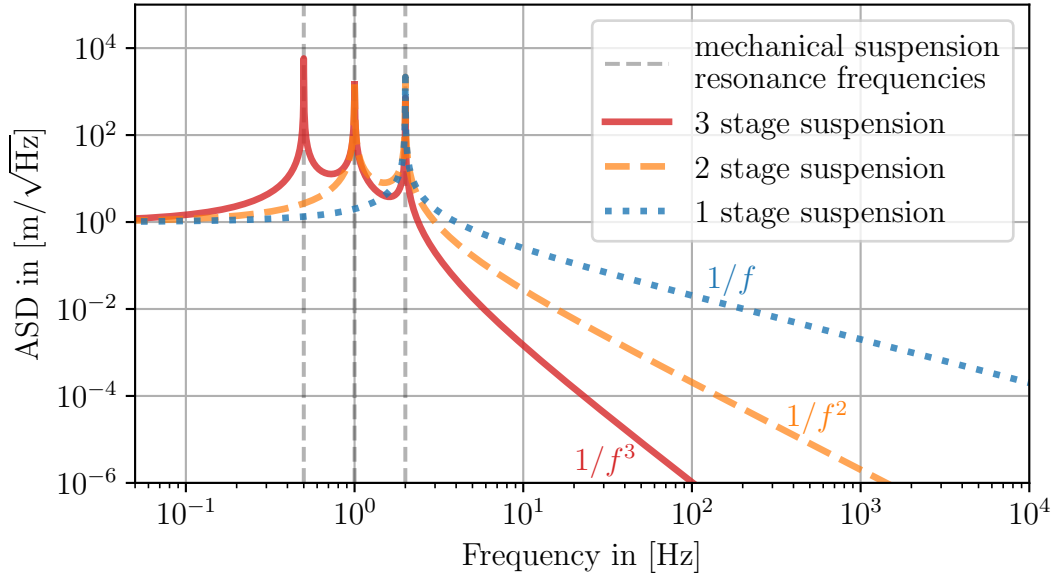
Figure 2.6: ASD plot for the horizontal transfer function for different (simplified) multi-stage pendulum examples.
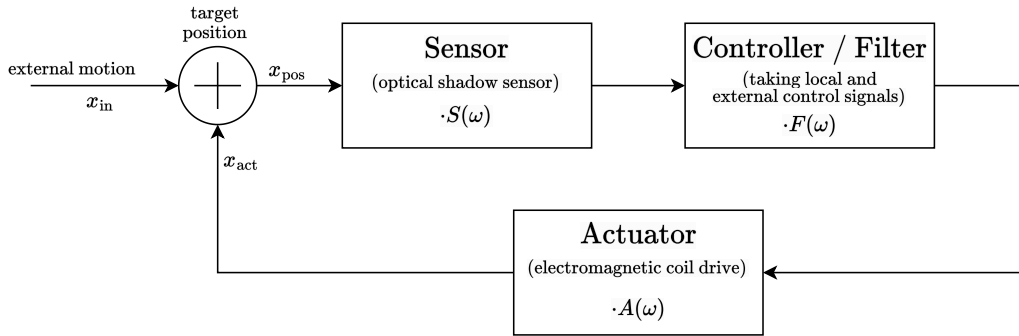
Figure 2.7: Simple block diagram of an active control plant. In LIGO's suspension system, optical shadow sensors are used as "sensors" to measure the position of a target. The "controller" takes the sensors readout and applies a custom filter / transfer function to the signal. The "actuator" is realized by an electromagnetic coil drive acting upon the target.

(described by $S(\omega)$) and fed to a controller which "filters" the signal according to the set design $F(\omega)$. The output of the controller (the "control signal") is then fed to a motor / an actuator which acts upon the system. In a real system, sensor $S(\omega)$ and actuator $A(\omega)$ will at least have a time-delay between their input and output and add additional noise, specific to the device used. The main purpose of the controller for the case of some suspended optics is so filter out the passive isolations resonances and help with the alignment of the optics.

The effect of such an active control loop on the movement of a suspended mass would look similar to the "effective suspension TF" shown in Figure 2.8. Assuming a white frequency input (like an instantaneous kick of the suspension), the lines shown there would correspond to the frequency spectrum of the induced motion of the suspended mass. By designing the controller to have a "zero" $F(\omega) \sim (\omega - \omega_{\mathrm{res}})$ at the passive systems' resonance frequency, the motion of the test–mass around
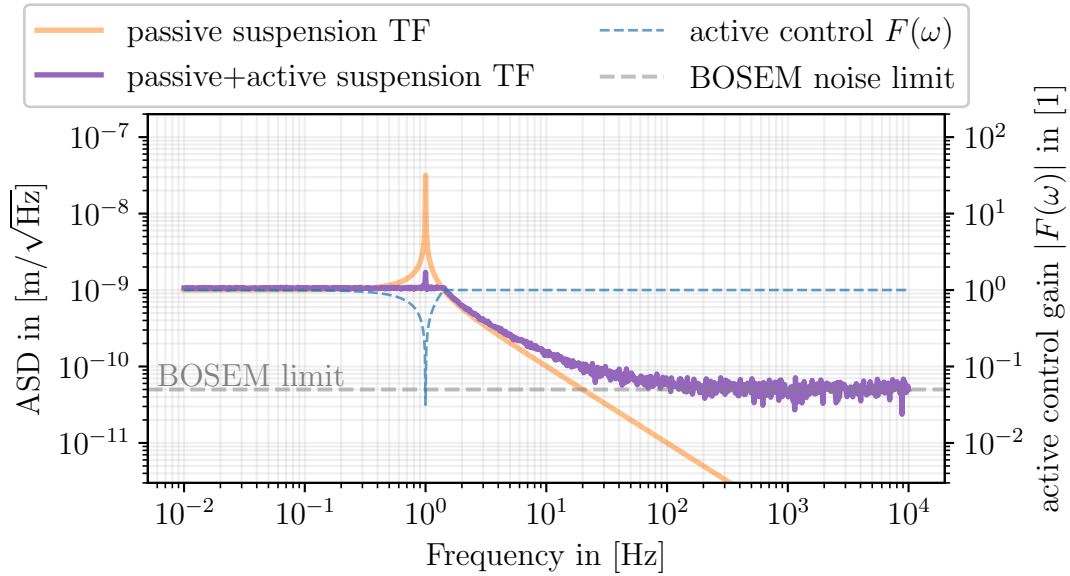
Figure 2.8: Plot of a model single-stage suspension transfer function (yellow) with a resonance frequency at $1\,\text{Hz}$, an active control / dampening (blue) around the resonance, and the combined transfer function of the passive + active system (purple). The gray dashed line marks the precision limit of the used sensor (within LIGO's suspension system given by the used BOSEMs), which adds in as noise limiting the effective dampening of the suspended mass.

this resonance can be reduced significantly. The trade off is the additional noise introduced by the used sensors and actuator. I.e., sensors have a limit in the precision they can reach and introduce sensor noise even when the position is exact. Actuators similarly have a limited precision (step-size) and introduce actuator noise even with a constant error signal.

While the currently used BOSEMs in LIGO's suspension system can significantly reduce the motion at the resonance frequencies of the passive suspensions, they introduce noise at higher frequencies leading to more overall or "root-mean-square (RMS)" motion, compared to only the passive isolation. This is one of the reasons why better local displacement sensors can help reduce the RMS motion of the target and improve upon the sensor's sensitivity at low frequencies.

2.2.3   Active alignment control and tilt-to-length coupling

Besides the dampening of the passive isolation systems resonance frequencies, local sensors and actuators are also used for the precise alignment of the optics. A misalignment in beam position and tilt leads to additional noise in the readout [50]. Changes in the beam position (together with inhomogenities of the optics) cause beam jitter noise. Variation of the tilt couple into the laser beams path length as shown in Figure 2.9. Additionally, offsets in both beam position and tilt of one mirror lead to the beam not hitting the other mirror centrally and add a torque to it (via the beams' radiation pressure). The resulting motion and the corresponding noise is also correlated to the misalignment of both mirrors.

Since the primary mirrors in GWDs are separated by large distances $\mathcal{O}(10^4\,\text{m})$, small
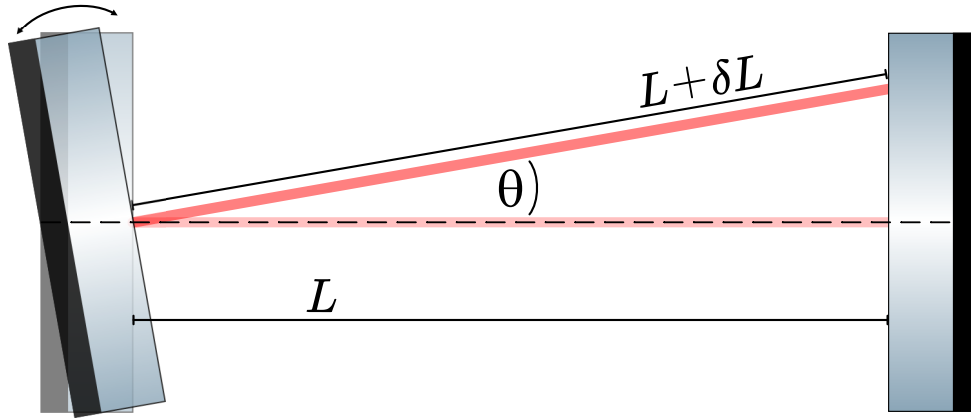
Figure 2.9: Sketch showing the simplified tilt-to-length coupling between two mirrors. By tilting the left mirror by an angle θ, the optical path length of the laser light changes from L to $L + \delta L$ with $\delta L = L \cdot (1 - \cos\theta) \approx L/2 \cdot \theta^2$. The actual coupling of the mirrors is however more complex as the radiation pressure also causes a coupling of the mirrors motion. In the shown picture the right mirror would also experience a torque tilting it, which leads to additional mechanical ("Sigg-Sidle") modes, requiring a more complex (global) control scheme.

angular displacements of $\sim 1\,\mu$rad already lead to length errors of $\delta L \approx 5\,\text{nm} \sim h \approx 10^{-13}$ which is many orders of magnitude larger than the length changes due to gravitational waves ($h < 10^{-21}$).

2.3 A path to overcome the "seismic wall": improvements on local displacement sensors

At frequencies below 30 Hz, LIGO is currently limited by "technical" noise coming from the alignment control and auxiliary length control. Maggiore et al. [50] have shown that the angular control noise that is present in both LIGO and Virgo would potentially also limit the performance of the planned Einstein Telescope when using a suspension design similar to either LIGO or Virgo without any improvements. The (internal) limits to the angular-alignment control systems are the self-noise of the used sensors (BOSEMs) and actuators.

The used BOSEMs operate by shining a collimated light beam directly into a photodiode with a 'flag' mounted at the suspended target blocking this beam partially. When the target moves, more or less light is blocked leading to a change of the brightness at the photodiode. While their dynamic range, given by the width of the beam, is sufficient, their precision limit in the order of ≈ 0.3 nm/Hz at 1 Hz (as shown in Figure 2.10) is one of the currently limiting noise sources in the alignment and control system of the suspensions in GWDs.

New sensors have been suggested to replace the currently used BOSEMs with more precise (interferometric) displacement sensors. I.e., van Dongen et al. [48] have shown that they could reduce the control noise at LIGO's main optics suspension by replacing the currently used optical shadow sensors with their own interferometric sensors called "HoQI." Other approaches suggest building dedicated tilt sensors
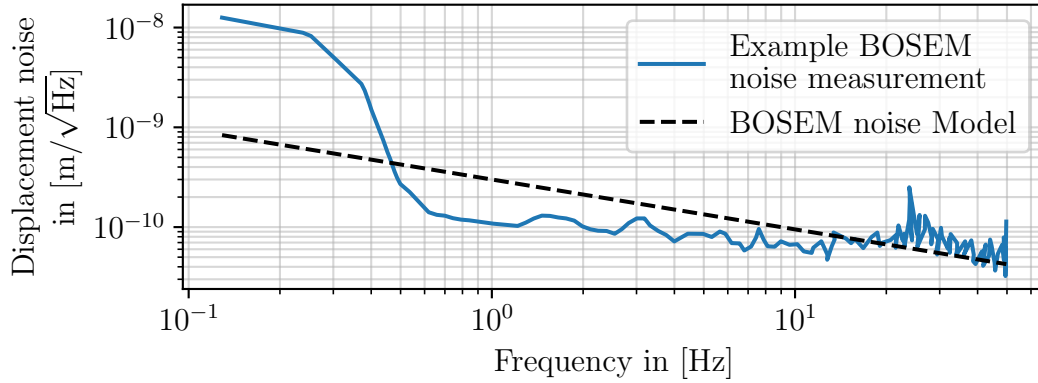
Figure 2.10: Example sensitivity of the shadow sensor part of a BOSEM (blue line). The sensitivity can vary slightly between different BOSEMs due to small variations in quality of its components. BOSEM sensor noise is usually modelled by $\sim 0.3\,\mathrm{nm}/\sqrt{f}$ (black dashed line) with a limit at around $\sim 50\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$.

to reduce the noise introduced by the tilt-to-length coupling [51, 52]. Reducing the overall / RMS motion of the optics could also help to reduce the amount of scattered light which has been suggested to be the cause of some of the unknown noise contributions at low frequencies.

While interferometric sensors are generally very precise being able to resolve displacements δL of a fraction of the laser's wavelength λ_0, their dynamic range is usually also limited to less than a wavelength as the measured signal i.e. for homodyne interferometry is periodic with one wavelength $\sim \sin(2\pi \cdot \delta L/\lambda_0)$, which is not enough to measure the sometimes larger motion of the suspended optics in GWDs.

The core of my Ph.D. thesis focused using a new kind of interferometric sensors called "DFMI" as local displacement- (and possibly tilt-) sensors intended to supersede the currently used optical shadow sensors. DFMI sensors provide a more precise displacement readout and would improve the alignment, leading to less noise from the "tilt-to-length" coupling described above. Unlike the previously tested HoQI, which uses two phase-offset signals and tries to measure large displacements by tracking of the changing phase, DFMI sensors inherently provide an absolute distance readout of a dynamic range of $\sim 1\,\mathrm{m}$ and have achieved a precision of down to $\sim 230\,\mathrm{fm}/\sqrt{\mathrm{Hz}}$.

The next part of this thesis explains what DFMI sensors are, how they work, and my contributions to their development.

Part II

Deep–Frequency Modulation Interferometry

Chapter 3

History and definition of DFMI

Deep–frequency modulation interferometer (DFMI) is an interferometry technique first developed by O. Gerberding [53] as derivative of an earlier deep–phase modulation interferometry (DPMI) developed by Heinzel et al. [54].

In a deep–phase modulation setup by Heinzel, as sketched in Figure 3.1, a sinusoidal phase modulation is added to one of two interfering laser-beams. This creates a characteristic signal consisting of several harmonic frequencies (shown in Figure 3.4), different from conventional single-carrier / sideband signals. With their setup, Heinzel et al. managed to reach picometer precision and track the distance of a target mirror, encoded in the signals phase, over several fringes.
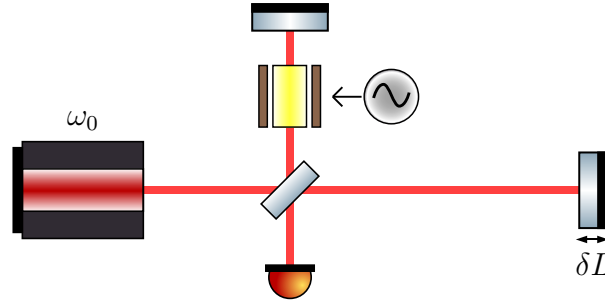
Figure 3.1: Sketch of an example Deep-Phase modulation setup. Here an electro–optic modulator is placed in one arm and modulated the laser beam with "a fixed modulation index" (explained in the next section).

Inspired by deep–phase modulation interferometry, Gerberding [53] first researched deep–frequency modulation interferometry as a more compact multi–fringe displacement sensing technique, reusing the DPM readout algorithm. In deep–frequency modulation interferometry (DFMI), the frequency of the main beam itself is modulated, leading to less optical components (required for the phase modulation) compared to DPMI. Instead of being only applied to one beam, in DFMI all beams carry the frequency modulation which simplifies the setup. The resulting interference signal measured on a photodiode is then nearly identical to the signal measured in DPMI. A novelty of DFMI is that it allows for an absolute path-length difference readout. More precisely, the DFMI readout yields two displacement parameters: a microscopic displacement derived from the interferometric phase, similar to the DPM readout, and a macroscopic, absolute path-length difference independent of

the used laser's mean wavelength and unique to DFMI.

The "deep" in the DFM name refers to the strength of the frequency modulation. Modulating the laser frequency (the "carrier frequency" ω_0) by some sinusoidal signal ($\omega_0 \mapsto \omega_0 + \Delta\omega \sin(\omega_m t)$) is, to some degree, common practice in physics and engineering. For small enough values of $\Delta\omega$, the resulting signal can be approximated by the carrier frequency and two so–called "sidebands" of frequency $\omega_0 + \omega_m$ and $\omega_0 - \omega_m$, i.e.

$$\sin(\omega_0 t + \Delta\omega \sin(\omega_m t)t) \approx A\sin(\omega_0 t) + B\sin((\omega_0 + \omega_m)t) + B\sin((\omega_0 - \omega_m)t) + r(t)\,,$$
(3.1)

with some amplitudes A and B and some rest term $r(t)$ which is considered unwanted noise. I.e., for the PDH locking technique, such a sinusoidal modulation is used to create these single sidebands which are then used to read out the cavities' length. The exact/full signal of a sinusoidal modulation contains, however, an infinite amount of sidebands (in $r(t)$) with their signal power depending on $\Delta\omega$. For the approximation (3.1) to hold (and for $r(t)$ to remain small) the modulation depth $\Delta\omega$ needs to be sufficiently small. In deep–phase and deep–frequency modulation interferometry this same modulation depth is chosen to be large or "deep" enough to have significant signal energy in many (usually > 10) of such sidebands. Instead of approximating the signal like above, the full analytic signal is then used to derive the length information.

Classic homodyne interferometry only operates within a single 'fringe'/ wavelength and while homodyne and heterodyne schemes exists that allow for larger displacements by tracking the phase of the signal, none of them provide an absolute displacement readout. For Michelson Interferometers using such techniques, the two interferometer arms are usually kept at (close to) the same length so that the absolute arm length difference is as small as possible. Fluctuations of the laser frequency then cancel mostly out when interfering the two beams. In DFMI, the two interferometer arms intentionally have different lengths to create a macroscopic arm length difference that is measured. For only a single DFMI setup this leads to a higher impact of laser frequency noise. For the intended use case of using multiple DFMI sensors for a local displacement readout, one additional sensor would be used in a fixed setup, to lock the laser frequency to this fixed length and suppress the laser frequency noise for all other setups.

3.1 Definition of DFMI

3.1.1 Definition : Deep–Frequency Modulation Interferometry

Deep-**F**requency **M**odulation **I**nterferometry (short: **DFMI**) is an interferometry technique where the frequency of the laser beam is modulated by adding a sinusoidal signal. The laser frequency is given by

$$\omega_{\text{DFM}}(t) = \omega_0 + \Delta\omega \cdot \sin(\omega_m t + \psi) \ . \tag{3.2}$$

The idealized (plane wave) electric field of a DFMI laser beam (at a fixed position) can be written as $\vec{E}_{\text{DFMI}}(t,\tau) = \vec{E}_0 \cdot \cos(\omega_{\text{DFM}}(t) \cdot \tau)$ with parameters:

ω_0	mean laser frequency
ω_m	modulation frequency
$\Delta\omega$	modulation depth
ψ	modulation phase
t	time of emission
τ	time of propagation
E_0	electric field amplitude .
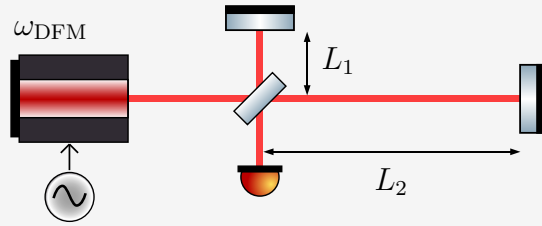
Figure 3.2: Typical DFMI setup with its unequal arm lengths L_1 and L_2.

Figure 3.2 shows a typical DFMI setup. The propagation-time difference of $\delta\tau := 2(L_2 - L_1)/c$ between the two beams in the setup leads to a measurable signal on the photodiodes of[a]:

$$s_{\text{DFMI}}(t,\tau) = B + A \cdot \cos(\underbrace{\omega_0\delta\tau}_{=:\varphi} + \underbrace{\Delta\omega\delta\tau}_{=:m} \cdot \sin(\omega_m t + \psi)) \tag{3.3}$$

$$= B + A \cdot \cos\left(\varphi + m \cdot \sin(\omega_m t + \psi)\right) \tag{3.4}$$

with

B	constant offset	A	signal (AC) amplitude
φ	interferometric phase	m	modulation index
ψ	modulation phase		

as main signal parameters. Any signal with a from as in (3.4) is referred to as a "DFMI signal" in this thesis.
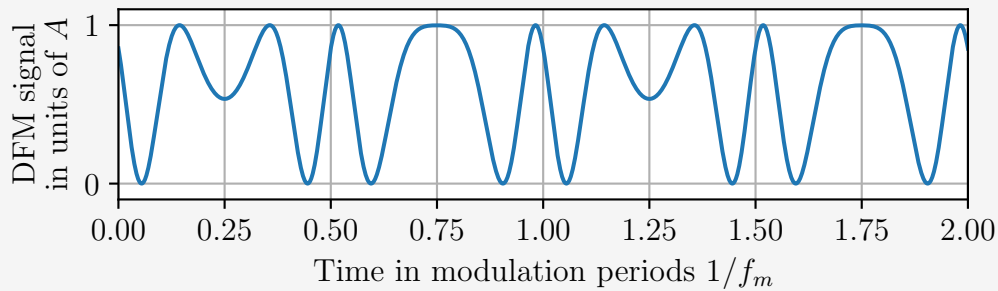
Figure 3.3: Plot an example time series of a DFMI signal with $m = 7$, $\varphi = \pi/4$ and $B = A$ over two modulation periods $2 \cdot 1/f_m$.

---

[a]The exact calculation can be found in appendix A.

It is important to note that B is not the constant DC part of the signal. The $\cos(\varphi + m \cdot \sin(\omega_m t + \psi)) = A J_0(m) \cos\varphi + \mathcal{O}(t)$ part adds another constant part of $A J_0(m) \cos\varphi$ dependent on the signal parameters as seen in (3.7).

The displacement information, e.g. the arm length difference ΔL is encoded twice in the signal, once in the modulation index $m = \Delta\omega/c \cdot \Delta L$ (with c as speed of light), encoding the absolute path-length difference, and once in the interferometric phase $\varphi = 2\pi/\lambda \cdot (\Delta L \mod \lambda)$, encoding the relative displacement.

One of the key features of DFMI signals is that they can be written as a somewhat compact Fourier-Series:

3.1.2 Definition : DFM signal harmonics

Using the Jacobi-Anger identity

$$e^{im\sin\theta} = \sum_{n\in\mathbb{Z}} J_n(m) \, e^{in\theta} \tag{3.5}$$

with J_n as Bessel functions of the first kind, a DFMI signal (3.4) can also be written as Fourier series with integer $n \in \mathbb{Z}$ as

$$s(t) = B + A\sum_{n\in\mathbb{Z}} J_n(m) \, \cos\left(n(\omega_m t + \psi) + \varphi\right). \tag{3.6}$$

Using the symmetry of the Bessel functions $J_{-n}(x) = (-1)^n J_n(x)$ for $n \in \mathbb{Z}$, the series can also be written with a summation index $n \in \mathbb{N}$ as

$$s(t) = B + A J_0(m) \cos\varphi$$
$$+ 2A\sum_{n=1}^{\infty}\left[J_{2n}(m) \cdot \cos\varphi \cdot \cos\left(2n(\omega_m t + \psi)\right) \right.$$
$$\left. - J_{2n-1}(m) \cdot \sin\varphi \cdot \sin\left((2n-1)(\omega_m t + \psi)\right) \right] \tag{3.7}$$

Further expanding the sine and cosine terms allows one to write this Fourier series with factors of $\sin(n\omega_m t)$ and $\cos(n\omega_m t)$ as

$$s(t) = B + A J_0(m) \cos\varphi \tag{3.8}$$
$$+ \sum_{n=1}^{\infty} 2A J_n(m) \times \begin{cases} -\sin(\varphi)\cos(n\psi)\cdot\sin(n\omega_m t) - \sin(\varphi)\sin(n\psi)\cdot\cos(n\omega_m t) & \text{for } n \text{ odd} \\ +\cos(\varphi)\cos(n\psi)\cdot\cos(n\omega_m t) - \cos(\varphi)\sin(n\psi)\cdot\sin(n\omega_m t) & \text{for } n \text{ even} \end{cases}$$

Throughout this thesis I refer to the individual parts of the Fourier series (3.6) - (3.8) as **signal harmonics** and integer multiples of ω_m as **harmonic frequencies**.

Figure 3.4 shows a time series and the power spectral density of an example signal
with these signal harmonics clearly visible as 'peaks' in the spectrum.
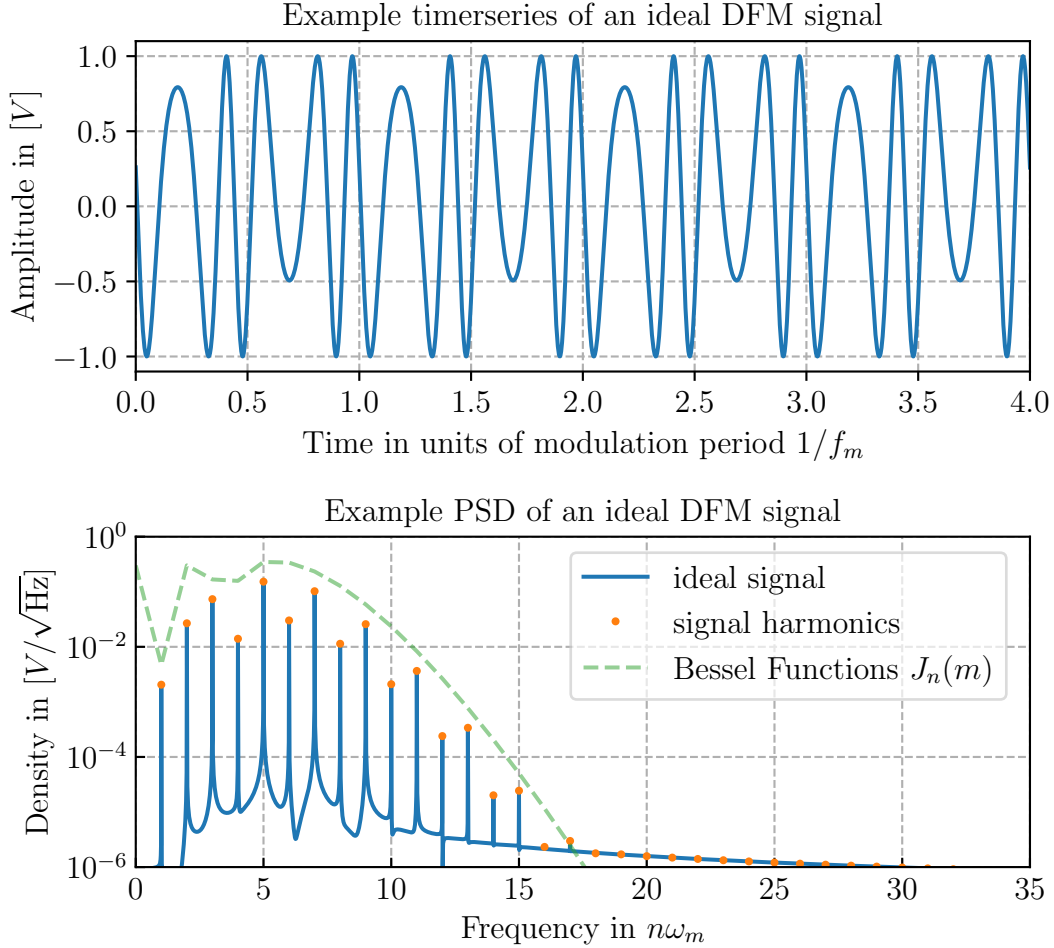
Figure 3.4: Plot of an example DFMI signal with $B = 0$, $A = 1\,V$, $m = 7$, $\varphi = 0.2$,
$\psi = \pi/8$ and $\omega_m = 2\pi \cdot 1\,\text{kHz}$. The upper plot shows a segment of the time series
with its characteristic shape that is periodic with the modulation frequency ω_m. The
lower plot shows the ASD of the signal with the signal harmonics being clearly visible as
"peaks" in the frequency spectrum. The harmonics are enveloped by the Bessel functions
$|J_n(m)| = |J_{\left(\frac{\omega}{\omega_m}\right)}(m)|$, with $m = \text{const}$.

The naming conventions used here and throughout this thesis were taken over from
the initial publications from Heinzel and Gerberding et al. Switching between this
thesis and the cited publications should be mostly seamless. In Heinzel's initial pub-
lication and the names "modulation depth" and "modulation index" are sometimes
used interchangeably both of $\Delta\omega$ and $m = \Delta\omega\delta\tau$. In the initial DFM publication
m is referred to as "effective modulation depth." Other publications by Gerberding
& Isleif also refer to m just as "modulation depth." In this thesis I refer to $\Delta\omega$ as
"modulation depth" and m as "modulation index."

3.2   Previously done research on DFMI

The following section gives a brief overview of the publication history of DFMI.

**1987 – 2010
Precurser to DFMI**

The origin of DFMI ranges back to several publications in the 1990's where researchers added a sinusoidal phase modulation to a laser causing a DFM-like signal being measured [55, 56, 57]. The DFM parameter estimation was done there by either approximating and simplifying the signal [55, 56, 57], or the use of a numeric, non–linear fit algorithm [54]. The signal was generally treated as having a fixed modulation index such that only the interferometric phase was used for the displacement readout. The direct precurser for DFMI was the deep–phase modulation interferometry publication bei Heinzel et al. [54] with its fitting readout algorithm being reused later for DFMI.

**Gerberding 2015
First DFMI study**

Gerberding's initial DFM publication [53] presents the concept of Deep–Frequency modulation interferometry as compact multi–fringe measurement technique that also allows for an absolute length readout via the modulation index (m) parameter. The two major arguments made for DFMI here are the absolute ranging capabilities and the greatly compactified setup compared to other heterodyne setups. The concept for implementing DFMI sensors is presented and analyzed, and their feasibility is demonstrated by simulating a DFMI signal with different types of noise and, using Heinzel's DPM readout algorithm, the signal parameters are calculated, showing promising performance comparable to other tabletop interferometer setups.

**Kissinger et al. 2015
An alternative
DFMI approach**

Besides Heinzel and Gerberding, Kissinger et al. also published an interferometer concept [58] which uses a DFMI signal to measure the displacement to a target. One of the key features of Kissinger's concept is to use a single modulated laser source, splitting the beam multiple times and sending them over different optical path (with different lengths), but recombining them all at a single photodiode, and then extracting the displacement information of the different paths from the measured signal. For the parameter calculation, Kissinger uses a special filtering technique which essentially isolates different signal harmonics which can be attributed mostly to a single optical path (a specific modulation index m) which needs to be sufficiently different from all other optical paths encoded in the signal. While Kissinger's alternative readout algorithm alone is
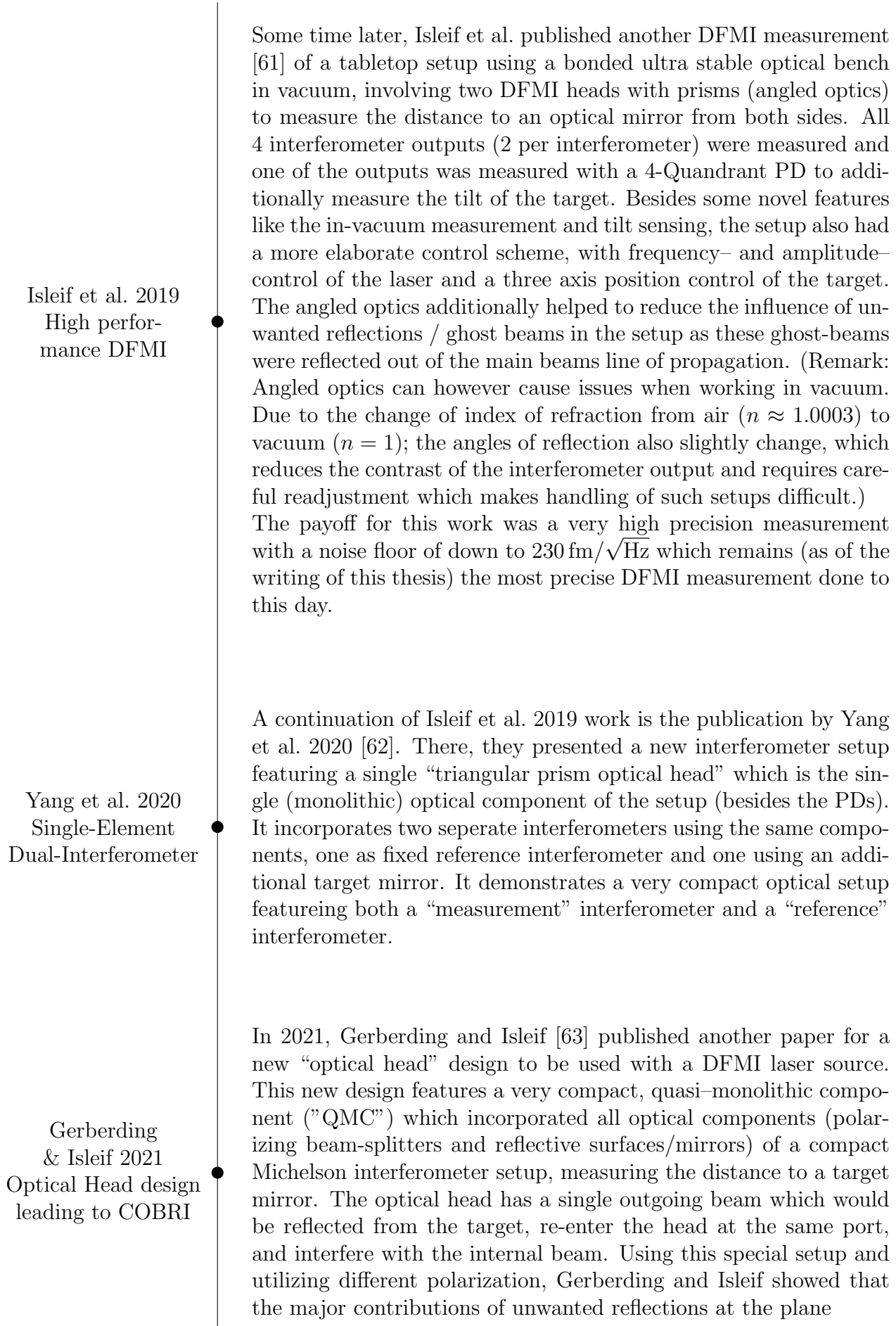
less well suited for a high-precision readout, it demonstrates how DFMI signals with different modulation indices can be (mostly) separated within a measured time series. This suggests that DFMI might be able to mitigate parasitic/ghost beams that travel along a sufficiently different path better than other interferometry techniques. (This is also briefly discussed in section 9.3.4).

Isleif & Gerberding et al. 2016 First experimental demonstration of DFMI

Shortly after the initial DFM paper, Isleif & Gerberding et al. performed a first experimental demonstration of DFMI [59]. They set up an in-air experiment with a Mach–Zehnder interferometer setup and a second Michelson interferometer, both fed with the same modulated laser source. The readout was done, same as the first DFMI publication, using Heinzel's DPM algorithm. By comparing both interferometers, and by splitting some of the outputs multiple times and feeding it to additional ADC channels, they could perform a detailed noise analysis. I.e., they managed to isolate and attribute thermal noise (and air density fluctuations), amplitude and shot-noise of the laser beam and digitization noise of the ADCs. With this first experiment, they reached a performance level of down to $250\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ as first ever high precision DFMI measurement published.

Schwarze 2018 Kalman filtering as DFMI readout algorithm

As part of his Ph.D. thesis in 2018 [60], Thomas Schwarze also presented a slightly different phase extraction scheme for DFMI by using a Kalman filter to extract the signal parameters. In his implementation, a Kalman filter uses the anayltic description of the DFMI signal and interativly models a DFMI signal using the measured signal as live feedback. As the DFMI signal is non–linear in most of its parameters (m, φ and ψ), the Kalman filter operation is somewhat comparable to a gradient descent where (derivatives of) the analytic model yield the gradient and the difference between the last filter result and the new data yield the step size for the filters iterative steps. In Schwarze's experiment, the readout of the Kalman filter reaches almost (within $\times 2$) the same precision as the DPM fitting algorithm while using significantly less computational resources.

Some time later, Isleif et al. published another DFMI measurement [61] of a tabletop setup using a bonded ultra stable optical bench in vacuum, involving two DFMI heads with prisms (angled optics) to measure the distance to an optical mirror from both sides. All 4 interferometer outputs (2 per interferometer) were measured and one of the outputs was measured with a 4-Quandrant PD to additionally measure the tilt of the target. Besides some novel features like the in-vacuum measurement and tilt sensing, the setup also had a more elaborate control scheme, with frequency– and amplitude–control of the laser and a three axis position control of the target. The angled optics additionally helped to reduce the influence of unwanted reflections / ghost beams in the setup as these ghost-beams were reflected out of the main beams line of propagation. (Remark: Angled optics can however cause issues when working in vacuum. Due to the change of index of refraction from air ($n \approx 1.0003$) to vacuum ($n = 1$); the angles of reflection also slightly change, which reduces the contrast of the interferometer output and requires careful readjustment which makes handling of such setups difficult.)

Isleif et al. 2019 High performance DFMI

The payoff for this work was a very high precision measurement with a noise floor of down to $230 \, \text{fm}/\sqrt{\text{Hz}}$ which remains (as of the writing of this thesis) the most precise DFMI measurement done to this day.

A continuation of Isleif et al. 2019 work is the publication by Yang et al. 2020 [62]. There, they presented a new interferometer setup featuring a single "triangular prism optical head" which is the single (monolithic) optical component of the setup (besides the PDs). It incorporates two seperate interferometers using the same components, one as fixed reference interferometer and one using an additional target mirror. It demonstrates a very compact optical setup featureing both a "measurement" interferometer and a "reference" interferometer.

Yang et al. 2020 Single-Element Dual-Interferometer

In 2021, Gerberding and Isleif [63] published another paper for a new "optical head" design to be used with a DFMI laser source. This new design features a very compact, quasi–monolithic component ("QMC") which incorporated all optical components (polarizing beam-splitters and reflective surfaces/mirrors) of a compact Michelson interferometer setup, measuring the distance to a target mirror. The optical head has a single outgoing beam which would be reflected from the target, re-enter the head at the same port, and interfere with the internal beam. Using this special setup and utilizing different polarization, Gerberding and Isleif showed that the major contributions of unwanted reflections at the plane

Gerberding & Isleif 2021 Optical Head design leading to COBRI

surfaces of the QMC, referred to as "ghost-beams," can be removed from the readout by using balanced detection, i.e., by adding and subtracting the different measured ports. By measuring both outgoing ports (using all of the outgoing signal energy), the setup also enables one to reach the maximal readout precision, as shown in more detail in this thesis in chapter 8. Due to its straight design without any angled optics, this optical head would be well-suited to be used in vacuum setups as it is the case for the local displacement sensors used in gravitational wave detectors. In later works, the here designed optical head has been given the name "**Co**mpact **B**alanced **R**eadout **I**nterferometer" Short: COBRI. [64]

Smetana et al. 2022 DFMI sensor head design made from commercial components

In 2022 another group at the University of Birmingham tested DFMI based sensors and designed an optical head from purely commercially available components [65]. Their readout scheme is however again different from the others used. Here, the modulation depth is tuned such that the amplitude of a selected pair of signal harmonics becomes equal and only a single pair of signal harmonics is used for the distance readout. By keeping the modulation index roughly fixed, their readout does not provide the macroscopic distance information generally encoded in a DFMI signal, but only the phase, which is then tracked to cover a range larger than one wavelength to achieve a displacement readout beyond a single fringe. While this approach showed a more easily reproducible way to build a DFMI sensor (without the need for specialized optical components), it did not archive the same performance level as previous setups which was later addressed and analyzed in more detail by the authors in [66].

Eckhardt & Gerberding 2022 DFMI Noise analysis

In 2022, Gerberding and I published a survey and analytical study of the readout noise limits of various interferometer types [67] with a focus on DFMI sensors. Since the distance information is encoded in a non–linear way into a measured DFM type signal, it is not immediately clear how different noise sources couple into the readout and how precise a DFM interferometer can be compared to e.g. a more simple homodyne interferometer setup. This paper aims at giving an answer to this question, provides some predicted noise curves for a measured DFM readout and shows that DFMI sensor can in principle reach (close to) the same performance as other interferometry techniques. The primary results of this paper are are incorporated in part III of this thesis.

Eckhardt & Gerberding 2023 Analytic readout algorithm

Up to this point, most of the previously done DFMI experiments used variants of Heinzel's DPM readout algorithm, which uses at its core a non–linear (Levenberg-Marquard) fitting algorithm, or other types of approxmations and fitting routines to find values for the m parameter encoded in the non–elemental and non–invertible Bessel-function $J_n(m)$ appearing in the measured signal. In my second paper [68], I present a new readout algorithm that only uses analytical operations to find values for all signal parameters without the need for any approximations or iterative fitting. Here, the Bessel-functions are eliminated by use of a recurrence relation. I consider this new algorithm, which we called "Analytic Readout Algorithm" as my biggest contribution toward DFM–interferometry. As a major part of my work, it is presented in chapter 4 of this thesis.

3.3   Planned DFMI applications

The primary planned use case for DFMI sensors is as local displacement sensors in gravitational wave detectors.

A key features of DFMI is that it allows for a very compact optical setup. O. Gerberding once stated that "DFMI takes the complexity from the optical setup [compared to other multi–fringe capable interferometry setups] and puts it into the digital [signal processing] domain." Having a simple and compact optical setup is highly beneficial for the integration into the mechanical suspensions in GWDs. And its high precision, multi–fringe readout allows for displacement sensing beyond currently used sensor capabilities. A potential candidate that might be used for such displacement sensing is the "COBRI" design.

3.3.1   Compact Balanced Readout Interferometer – "COBRI"

As mentioned in the previous section, Gerberding and Isleif developed the COBRI with an optical layout as shown in Figure 3.5. By using polarizing beamsplitters and a quarter–wave plate, the reflected beam can be separated from the incoming beam while using the same optical path in this miniature Michelson Interferometer setup. By having one of the two Interferometer arms (the internal one) very short, the measured DFMI signal is also very close to the ideal one written in (3.4) as the additional rest terms $\mathcal{O}(\psi_L)$ as written in Appendix A become very small. In its mechanical design, the optical components are fused together into a quasi mono-lithic component (QMC). The measured arm length difference is almost exactly the distance between the exit point of the COBRIs QMC and the target mirror. The design notably features plain optics (perpendicular to the beam's path) to intentionally avoid angles between the reflective surfaces and the laser beam, which would change in vacuum. When aligning an optical setup at normal pressure and subsequently switching to a vacuum, the reflective index changes from $n_{\text{air}} \approx 1.0003$ to
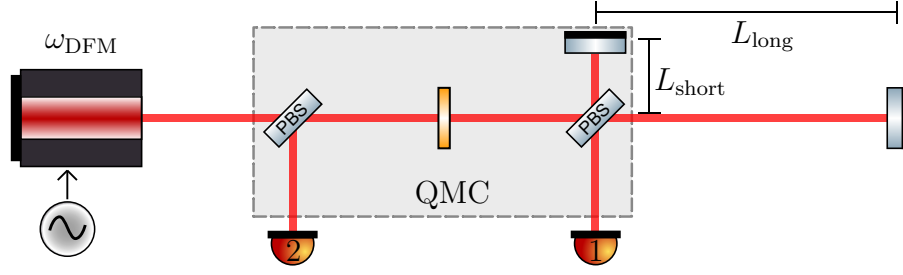
Figure 3.5: Simplified sketch of the optical setup of the COBRI design by Gerberding & Isleif [63]. By using of the polarization of the beam, the incoming and outgoing beam can be separated, and all reflected light is measured in photodiodes 1 and 2. In the COBRI design, all optical parts are fused together in a "quasi monolithic component (QMC)."

$n_{\mathrm{vacuum}} = 1$. While this change is not large, it does affect the non–zero angles of reflection at every air/vacuum surface which can reduce the contrast and lower the readout precision of the setup. Plain aligned optics prevented this, and allow for the use of retroreflectors on the target which minimize the loss of contrast due to tilts. The downside of such a setup is however that every surface causes a small amount of unwanted reflections of the passing beam called "ghost beams." For angled optics, these ghost beams usually diverge away from the main beam's path of propagation. For plain aligned optics, they do couple back strongly into the main beam. To deal with these additional unwanted signals, Gerberding and Isleif developed a scheme to remove the major ghost beam in their COBRI setup by calculating different linear combinations of the two (or more) COBRI outputs. The COBRI design also features
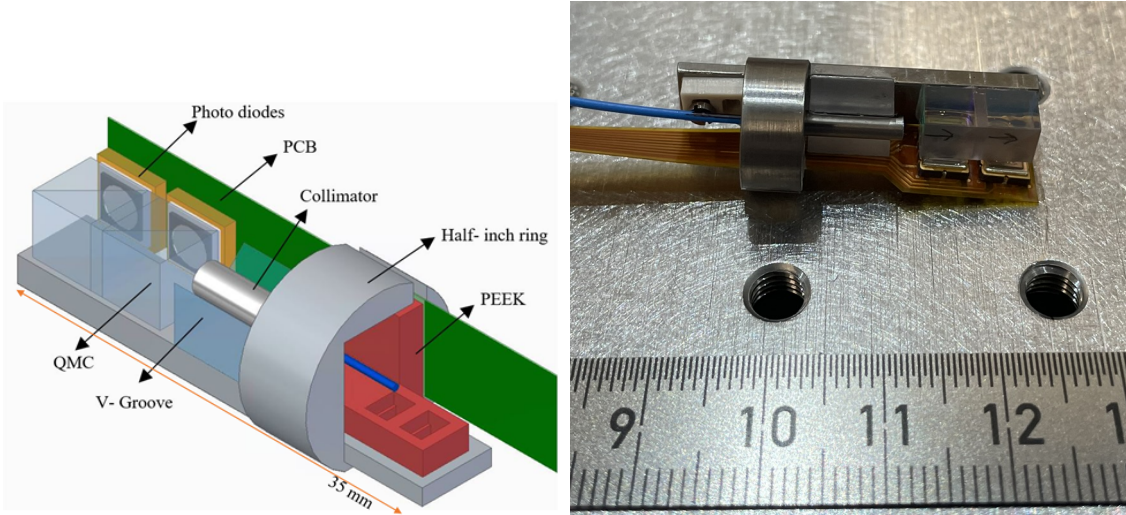
(a) Mechanical CAD drawing of the COBRI by [M. Mahesh]

(b) Picture of the used COBRI prototype build by M. Mahesh. The scale shows a total length of $\sim 3\,\mathrm{cm}$ length of the COBRI prototype.

the option to use a quadrant photodiode at e.g. the "direct port" to read out the tilts (yaw and pitch) of the target. By measuring the phase differences between the different quadrants of the photodiode, tilts of the target down to $\approx 20\,\mathrm{nrad}/\sqrt{\mathrm{Hz}}$ have been measured in previous DFM experiments [61].

3.3.2 DFMI sensor network in GWDs

A compact and precise interferometric sensor, like the COBRI presented in the last section, would enable a precise positioning and local displacement readout for the optics in a suspension chain used in GWDs. The sensor readout would be used to improve the active damping and to support the global interferometer control. Figure 3.7 shows a sketch of where an array of compact COBRI–type sensors (and additional target mirrors) would be added to the suspension to measure all degrees of motion of a suspension stage. A single modulated laser source would supply
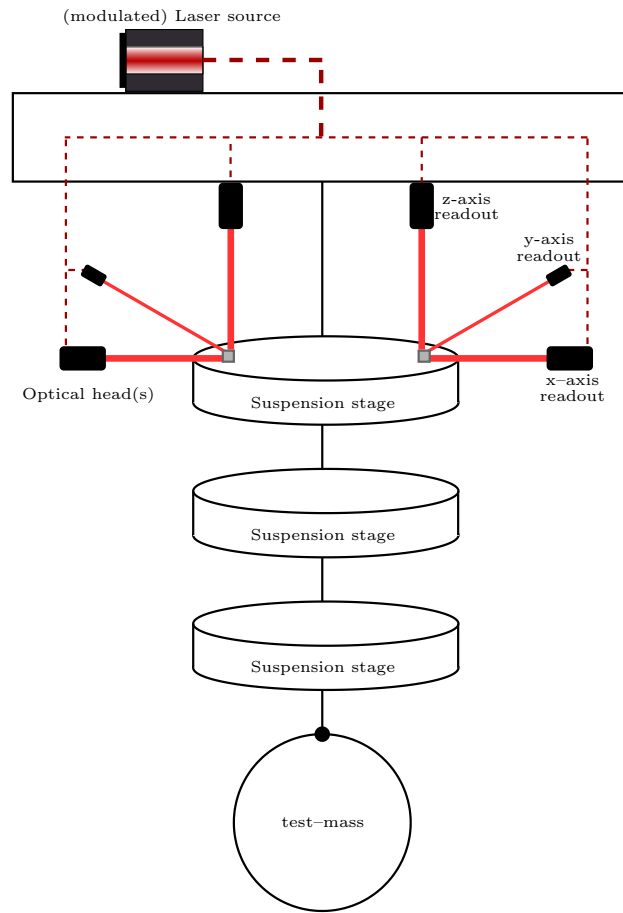
Figure 3.7: Simplified sketch of sensor network of multiple COBRI–type sensors at a test mass suspension.

a potentially large number of optical heads / COBRIs, where each COBRI would measure displacements in one dimension. Three COBRIs at different positions could measure the target displacement in every direction and additional 3 COBRIs targeting another point of the same object would allow measurements of tilt, yaw and pitch.

3.3.3 DFM parameters for local displacement readout in GWDs

In such a mechanical suspension, the DFMI sensor would be mounted closely ($L_0 \approx$ 5 cm) to the masses and would need to provide a multi–fringe readout with a dynamic range in the order of $\delta L \approx 1$ cm. Aiming for a typical modulation depth of $m = 7$, which was used in previous DFMI experiments that reached a high precision, leads to a modulation depth of $\Delta \omega = 7 \cdot c_0/L_0 \approx 2\pi \cdot 3.37$ GHz. Orienting myself on the 1550 nm external cavity laser I used for my experiments, I recommend a modulation frequency of around $f_m = 500$ Hz which should be as high as possible but below typical resonance frequencies of the Piezos used to modulate the laser frequency, and small enough to not be limited by non–linear (frequency) behavior of the laser. For the ADC, I consider a fully differential input (so that B can be zero) with 1 MHz sampling frequency. In Summary, what I consider the default DFMI parameters for a local test–mass readout, leading to a signal as shown in Figure 3.8, are:

$$f_S = 1\,\text{MHz} \qquad \Delta f = 3.37\,\text{GHz} \qquad f_m = 500\,\text{Hz}$$

$$T = 16\,\text{ms} \qquad B = 0V \qquad A = 5V$$

$$\lambda_0 = 1550\,\text{nm} \qquad \psi = 0.1$$

$$\delta L = (10 \pm 1)\,\text{cm} \implies m = (7 \pm 0.77)\,\text{rad} \quad \varphi = (0.81 \pm 445904)\,\text{rad}$$

Table 3.1: Typical DFM parameters for a local test–mass readout. The laser and modulation parameters are based on the laser I used, and the measurement time is based on readout considerations presented in the next chapter.
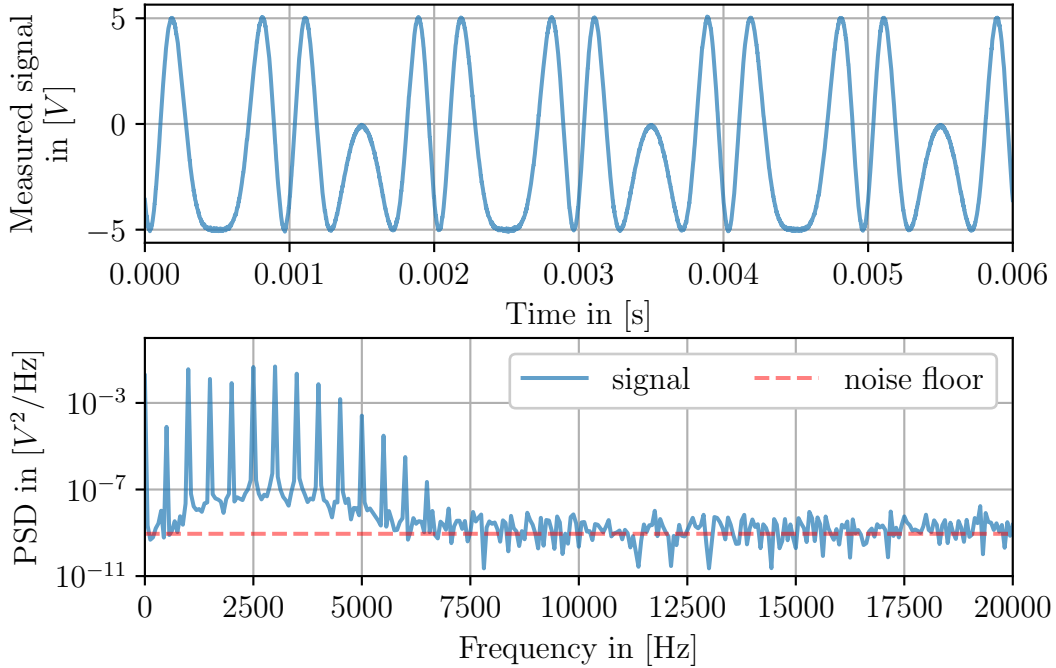

Figure 3.8: Plot of a DFM signal for a local test–mass readout with the DFM signal parameters as given in table 3.1.

3.3.4   Alternative signal parameters for different scenarios

While the parameters shown above correspond to a common use–case scenario, DFMI could be used in a wide range of applications for distance measurements with very different signal parameters. I.e. while ADC sampling frequencies of $1 - 2$ MHz are not uncommon, the setup above does not fully utilize the frequency range provided by such a fast sampling rate compared to the modulation frequency (assuming that the analog electronics in front of the ADC do not add significant noise at high frequencies). This leaves room to either increase the absolute distance or to increase the readout speed. I.e. by increasing the absolute distance, leading to a larger modulation index, the higher order harmonics (at higher frequencies) contain more signal energy and become relevant for the readout. Table 3.2 shows DFMI parameters for an alternative setup with its signal plotted in Figure 3.9. Such a setup could measure the absolute range and interferometric displacements in picometer precision over a range of ~ 10 m. The shorter measurement time of only 2 ms would lead to a readout frequency of $\sim 1/2$ ms $= 0.5$ kHz, which is useful when tracking (relatively) fast moving targets.

$$f_S = 2\,\mathrm{MHz} \quad \Delta f = 1\,\mathrm{GHz} \quad f_m = 4\,\mathrm{kHz}$$

$$T = 2\,\mathrm{ms} \quad\quad B = 2.5V \quad\quad A = 2.5V$$

$$\delta L = 10\,\mathrm{m} \quad \lambda_0 = 1064\,\mathrm{nm} \quad\quad \psi = 0$$

$$\implies \quad\quad m = 210\,\mathrm{rad}$$

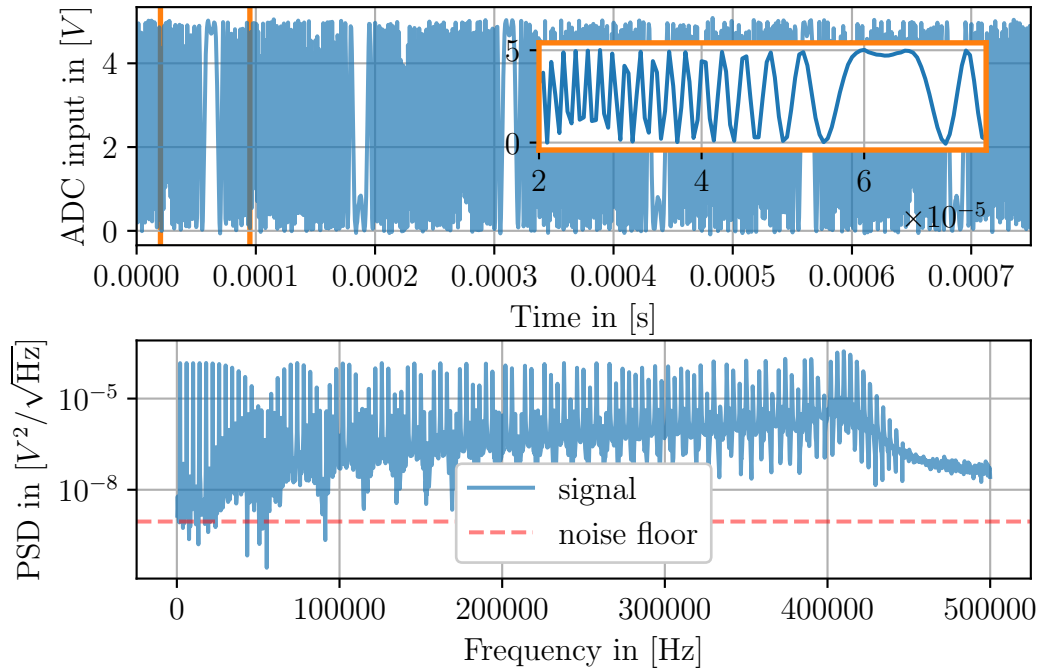**Table 3.2:** DFM parameters with a focus on a long absolute distance readout.

Figure 3.9: Time series and signal spectrum of a DFM signal for a "high-speed readout." Here, the DFM signal is distributed over ~ 110 relevant signal harmonics. The Region in the orange rectangles is zoomed in for a smaller window for better visibility of the signal.

# Chapter 4

Analytic Readout Algorithm

Most of the previously done DFMI experiments used of Heinzel's DPM readout algorithm [54]. It makes use of the Fourier series form (3.6) of the signal which includes Bessel-functions of the first kind $J_n(m)$ which are non–linear (in m) and non–invertible. To be able so solve all parameters; the DPM readout algorithm uses an iterative (Levenberg–Marquard) fit algorithm. Two disadvantages of this approach are that (a) the algorithm needs sufficiently good starting values to find precise results for all parameters and (b) it is computationally expensive and limits the speed of the readout (and requires sufficiently potent hardware to run). I.e. in Isleif et al. [61], the readout algorithm required a ($\sim$2019 era) 8-core desktop CPU to exclusively run Heinzel's readout algorithm on 8 channels allowing for a max speed of $\sim 100\,\mathrm{Hz}$ with almost full CPU utilization.

In this chapter I present an alternative method to calculate the signal parameters in a non–iterative way that does not require initial values. This new algorithm was named **analytic readout algorithm** and published in [68].
The analytic readout algorithm operates on a set of signal harmonics, which are demodulated individually, and from the resulting Fourier coefficients, the signal parameters are calculated by arithmetic operations. The non–linear Bessel functions that appear in the signal's Fourier coefficients are eliminated by use of a recurrence relation without any approximations or restrictions to the signal parameters.

A detailed overview of the algorithm and its individual steps can be seen as flowchart in Figure 4.1. And a more compact overview, that will be used to guide the reader, is shown at the beginning of every of the following sections, with each section expanding one of the colored algorithm steps. Throughout this chapter, algorithm results are sometimes compared to an "optimal result" or "optimal readout performance." This refers to the CRB of the discussed parameter which is explained in detail in the later part of this thesis.

Section
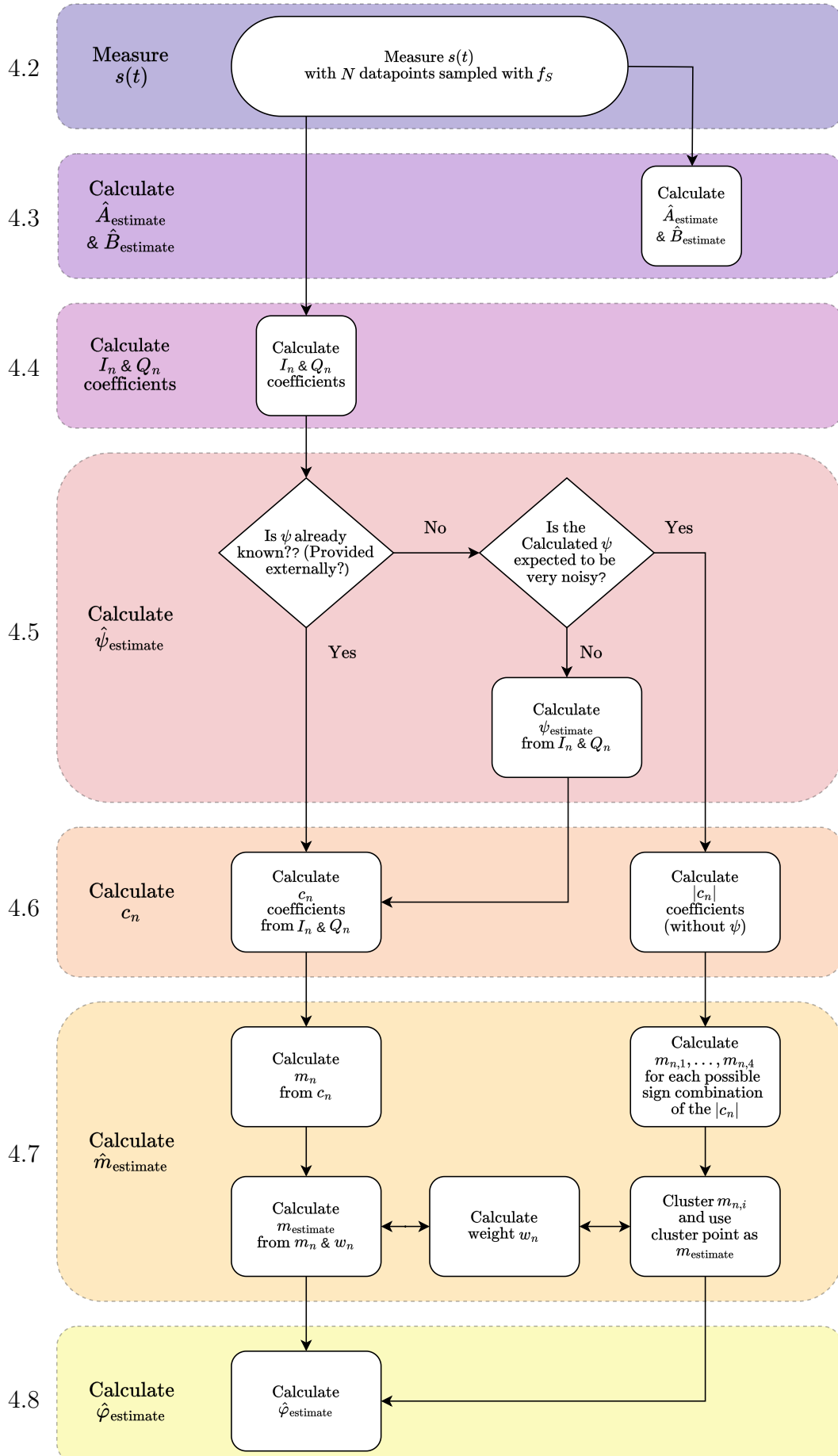
4.1   The algorithm

The ideal DFMI signal (as written in 3.6) is given by

$$s(t) = B + A \cdot \cos \left(\varphi + m \cdot \sin(\omega_m t + \psi) \right) \tag{4.1}$$

$$= B + A \cdot \sum_{n \in \mathbb{Z}} J_n(m) \, \cos \left(n\omega_m t + n\psi + \varphi \right) . \tag{4.2}$$

In general this signal has 5 free parameters (B, A, m, φ, ψ) with t as time and ω_m as fixed parameter. While the ω_m parameter could also be derived from the measured signals simply by finding the (frequency) position of the harmonics in the power spectrum; in a DFMI experiment, the laser modulation (including ω_m) is a set / fixed parameter and not an unknown.

4.1.1   Example Case (Definition of $s_{\mathbf{example}}$)

For a better understanding of the algorithm, the following algorithm steps also include calculations and results for a simulated example signal $s_{\text{example}}(t)$ which uses the signal parameters of a DFMI signal for a local displacement readout, as presented in section 3.3.3, with the signal parameters as written in Table 3.1. I.e.,

$$s_{\text{example,ideal}}(t) = B_{\text{ex}} + A_{\text{ex}} \cdot \cos \left(\varphi_{\text{ex}} + m_{\text{ex}} \cdot \sin(\omega_{m,\text{ex}} t + \psi_{\text{ex}}) \right) + n(t) \tag{4.3}$$

with signal parameters as written in Table 4.1 and $n(t) \sim \mathcal{N}(\mu = 0, \sigma^2 = 10^{-3} \, \text{V}^2)$ as random, additive, Gaussian distributed noise with zero mean ($\mu = 0$) and a variance of $\sigma^2 = 10^{-3} \, V^2$.

$B_{\text{ex}} = 0V$	$A_{\text{ex}} = 5V$	$\varphi_{\text{ex}} = 0.81 \, \text{rad}$	$f_S = 1 \, \text{MHz}$
$m_{\text{ex}} = 7 \, \text{rad}$	$\omega_{m,\text{ex}} = 2\pi \cdot 500 \, \text{Hz}$	$\psi_{\text{ex}} = 0$	$\sigma^2 = 10^{-3} \, \text{V}^2$

Table 4.1: Table of the DFM parameters used for the simulation of s_{example}.

4.2 The digitized signal

Measure $s(t)$ → Calculate $\hat{A}_{\text{estimate}}$ & $\hat{B}_{\text{estimate}}$ → Calculate I_n & Q_n coefficients → Calculate $\hat{\psi}_{\text{estimate}}$ → Calculate c_n → Calculate $\hat{m}_{\text{estimate}}$ → Calculate $\hat{\varphi}_{\text{estimate}}$
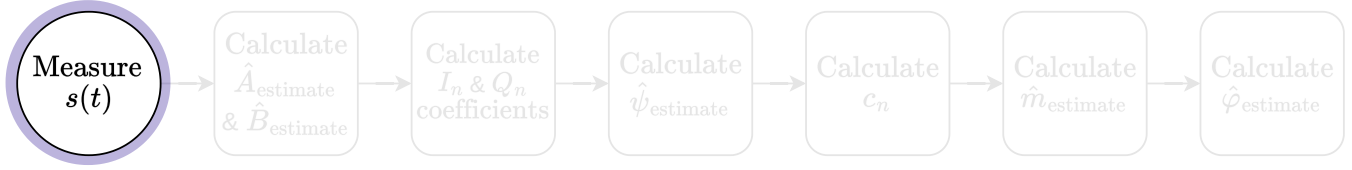
As a "zero'th" step, the signal is measured over a time period T and sampled (and digitized) with a sampling frequency of f_S, leading to $N = T \cdot f_S$ data points. The measured signal is then given as discrete time series $\underline{s} = \{s(t = \frac{0}{f_S}), s(t = \frac{1}{f_S}) \cdots, s(t = \frac{N-1}{f_S})\} =: \{s_0, \cdots, s_{N-1}\}$. Analytically, the finite measurement time can be expressed by multiplying the ideal signal with (4.2) with a rectangular window function $\text{rect}_T(t) = \text{rect}((t - T/2)/T)$ such that:

$$s_{\text{measured}}(t) = s(t) \cdot \text{rect}_T(t) \quad \text{with} \quad \text{rect}_T(t) = \begin{cases} 1 \text{ for } 0 < t < T \\ 0 \text{ else} \end{cases} . \quad (4.4)$$

While not strictly necessary, it is beneficial (as explained later in section 4.4.2) to choose a measurement time T close to an integer multiple of the modulation period $2\pi/\omega_m$ to improve the readout's precision. For the example signal we assume at this stage a sampling frequency of $f_S = 1\,\text{MHz}$ and a measurement period of $T_{\text{ex}} = 16\,\text{ms} = 8 \cdot 2\pi/\omega_m$. This leads to a total of $N = 16000$ data points. The analytic description of the example signal then becomes: $s_{\text{example}}(t) := s_{\text{example,ideal}}(t) \cdot \text{rect}_{T_{\text{ex}}}(t)$.
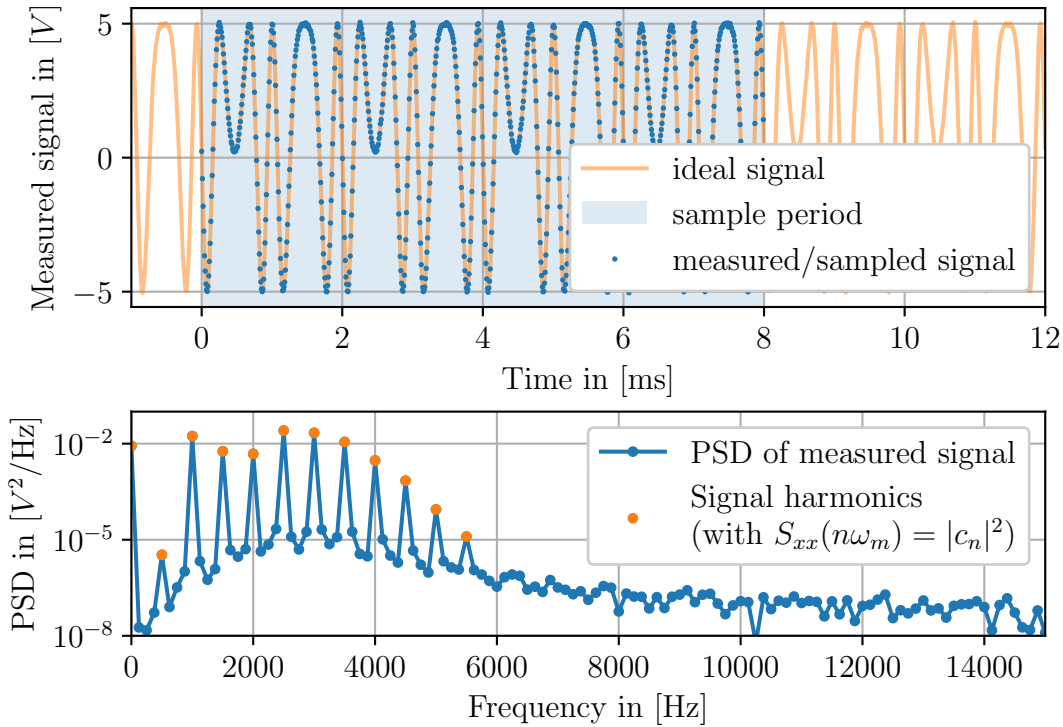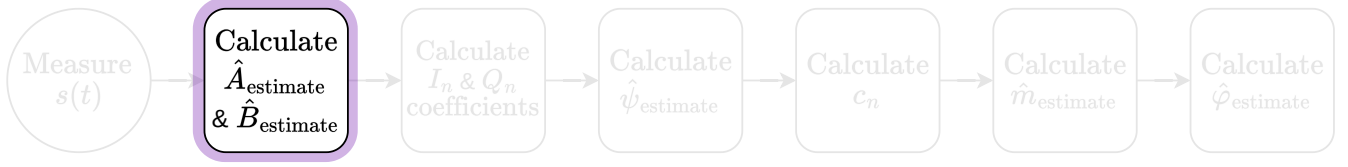

Figure 4.2: Example sketch of the ideal signal $s_{\text{example,ideal}}$ (yellow line), the time of measurement $[0, T]$ (light blue background) and the resulting sampled data of s_{example} (blue dots). For visibility, in this plot only 4 signal periods (so $T_{\text{ex}}/2$) are displayed with a slower sampling frequency of only $0.1\,\text{MHz}$.

4.3 Calculating the amplitude A and offset B

As first step of the analytic readout algorithm we calculate the constant B coefficient and the amplitude of the signal A. For typical DFM applications these two parameters are however only of minor interest. B is used in the next (demodulation) step as it adds to the zeroth harmonic $n = 0$. Since the DC part of the measured signal is however usually more noisy than the other signal harmonics, I generally recommend not using it all together. The amplitude A can be used to measure effects which reduce the optical power, like a loss of optical contrast or as error signal to correct power fluctuations of the laser. The precision of the B and A parameters do however not directly influence the precision of the length / displacement measurement.

Assuming that the modulation index is larger than $m \geq \pi$ immediately shows that

$$s(t) = B + A \cos(\varphi + m \cdot \sin(\omega_m t + \psi)) \tag{4.5}$$

$$\implies \begin{array}{l} \min\{s(t)\} = B - A \\ \max\{s(t)\} = B + A \,. \end{array} \tag{4.6}$$

From the maximal and minimal value of the measured signal we calculate the parameter estimates via:

$$\hat{B} := \tfrac{1}{2}\left(\max\{s(t)\} + \min\{s(t)\}\right) \tag{4.7}$$

$$\hat{A} := \tfrac{1}{2}\left(\max\{s(t)\} - \min\{s(t)\}\right) \tag{4.8}$$

A feature of the DFM signal is that "DC" or mean signal measured over an interval also depends on the signal parameters and is given by:

$$s_{\mathrm{DC}} := \int_0^T s(t)\,dt = B + A \cdot J_0(m) \cdot \cos\varphi \tag{4.9}$$

This is a key difference to other typical heterodyne signals which only carry signal at frequencies $\omega > 0$ and where the DC value of the signal can simply be estimated by calculating the mean as in (4.9). The downside of having to use (4.7) and (4.8) as estimators for B and A is that for most (symmetric) additive noise terms, picking the absolute minimum and maximum of the measured signal will lead to a slightly larger error ($\sim \sigma \cdot \sqrt{2 \ln(T f_S)}$, compared to calculating a mean).

4.4   Demodulation and definitions for I_n, Q_n & $|c_n|^2$

The second step in the analytic readout algorithm is (similar to [54, 53]) to calculate the coefficients of the Fourier series of the signal. Experimentally this is done with the so-called "I/Q demodulation" technique.

4.4.1   Definition of I_n and Q_n

When writing the DFMI signal as Fourier series

$$s(t) =: \sum_{n=0}^{\infty} I_n \, \cos(n\omega_m t) + Q_n \, \sin(n\omega_m t) \ , \tag{4.10}$$

I define the I_n and Q_n as the Fourier coefficients of the signal harmonics $\sin(n\omega_m t)$ and $\cos(n\omega_m t)$ as written in Table 4.2.

	n even	n odd
I_n	$A \cdot J_n(m) \cdot \cos\varphi \cdot \cos n\psi + B\delta_{n0}$	$-A \cdot J_n(m) \cdot \sin\varphi \cdot \sin n\psi$
Q_n	$-A \cdot J_n(m) \cdot \cos\varphi \cdot \sin n\psi$	$-A \cdot J_n(m) \cdot \sin\varphi \cdot \cos n\psi$

Table 4.2:   Table of the (ideal) I_n and Q_n coefficients for a DFMI signal without any noise.

Practically, we calculate these coefficients by "I/Q demodulation," i.e., by calculating:

$$I_n = \frac{1}{T} \int_0^T s(t) \cdot \cos(n\omega_m t) \cdot W_T(t)\, dt \quad \text{and} \quad Q_n = \frac{1}{T} \int_0^T s(t) \cdot \sin(n\omega_m t) \cdot W_T(t)\, dt \tag{4.11}$$

$$\approx \frac{1}{N} \sum_{k=0}^{N} s(t_k) \cdot \cos(n\omega_m t_k) \cdot W_T(t_k) \qquad \approx \frac{1}{N} \sum_{k=0}^{N} s(t_k) \cdot \sin(n\omega_m t_k) \cdot W_T(t_k), \tag{4.12}$$

with $W_T(t)$ as window function (4.15) as explained in detail below.
Inserting the ideal Fourier series (4.10) into (4.11), using n' as summation index, choosing a measurement time of an integer multiple of the modulation period i.e., $T = 2\pi/\omega_m$, and setting $W_T \equiv 1$, causes to all other parts (with $n' \neq n$) of the sum in (4.10) to vanish and leaves only the n'th Fourier coefficient leading to the values as written in Table 4.2. In practice, the time T and the frequency ω_m are however not exact and the other parts of the sum of (4.10) will also slightly contribute to the calculation of the I_n and Q_n. To mitigate this effect I employ an additional filter in form the window function $W_T(t)$.

4.4.2 Filtering of the demodulated signal with the W_T window function

Only when integrating in (4.11) over exactly an integer multiple of $2\pi/\omega_m$ do the other parts of the sum in (4.10) vanish. This can also be visualized by looking at the signal in Fourier domain as shown in Figure 4.3 where the finite window of length T leads to zeros in the spectrum at exactly integer multiples of $2\pi/T$. Analytically,
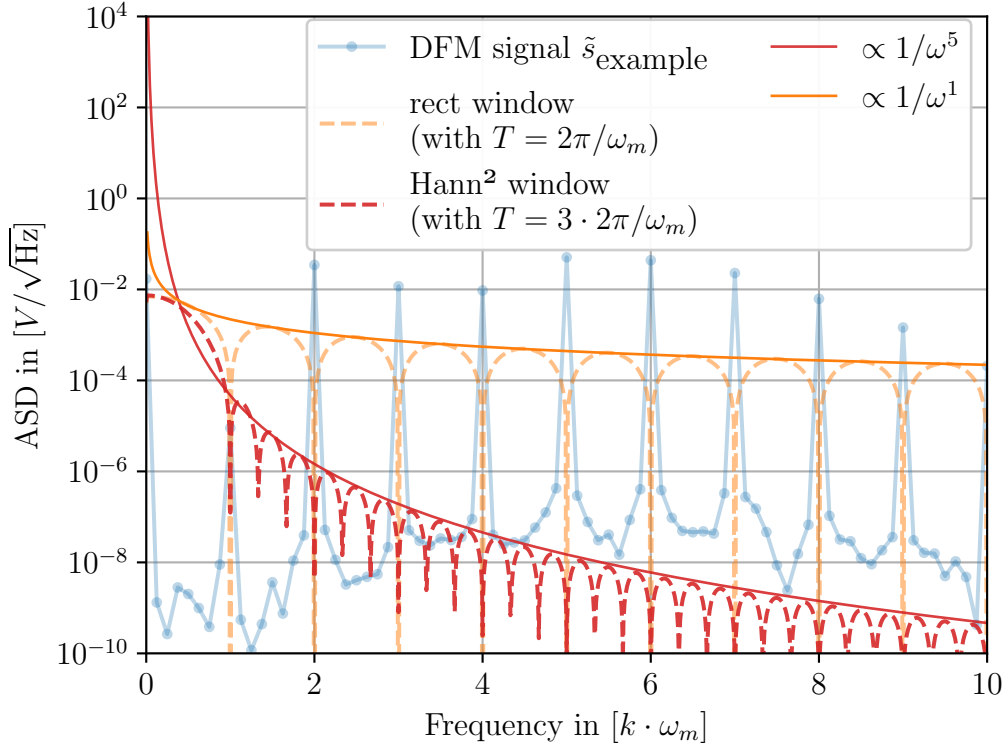

Figure 4.3: Plot of an example DFMI signal spectrum and two window functions (a rectangular window (yellow) and a squared Hann window (red)) with Window length / measurement time of $T = 2\pi/\omega_m$.

the I_n and Q_n coefficients in (4.11) can also be written as function of the Fourier transformed[1] of the signal (with W_T and s real and $*$ as convolution operation) as

$$I_n = \text{Real}\left\{\left(\tilde{s} * \tfrac{1}{T}\tilde{W}_T\right)\Big|_{\omega=n\omega_m}\right\} \qquad \text{and} \qquad Q_n = \text{Imag}\left\{\left(\tilde{s} * \tfrac{1}{T}\tilde{W}_T\right)\Big|_{\omega=n\omega_m}\right\}. \tag{4.13}$$

For a plain rectangular window, the Fourier transformed is the sinc function given by

$$\mathcal{F}\left\{\frac{1}{T}\text{rect}_T(t)\right\} = e^{-i\frac{T}{2}\omega} \cdot \text{sinc}\left(\frac{\omega}{\left(\frac{2\pi}{T}\right)}\right) \qquad \sim \frac{\left(\frac{2\pi}{T}\right)}{\omega} \text{ for large } \omega . \tag{4.14}$$

I.e. it has zeros at integer multiples of $2\pi/T$ (i.e., with $T = 2\pi/\omega_m \implies$ zeros at all harmonic frequencies $n\omega_m$ with $n \neq 0$) and falls off asymptotically with $1/\omega$ (the yellow dashed line in Figure 4.3). The convolution operation in (4.13) causes

---

[1] For the convention of the used Fourier transform, signified by the tilde ˜, see section 7.1.

only the same harmonic $\tilde{s}(n\omega_m)$ to contribute to I_n (or Q_n) as all other harmonics ($\tilde{s}(n'\omega_m)$ with $n' \neq n$) are multiplied with the zeros of the window function.

In a real experiments, the signal is always measured with a finite precision and the T and ω_m can be imprecise (i.e., due to timing jitter noise). In this case, the other harmonics are shifted slightly off the zeros of the window function, which leads to errors in the calculated I_n and Q_n coefficients. For a rectangular window, these errors are attenuated by a factor of $\propto 1/\omega$ (visible in Figure 4.3). Choosing another
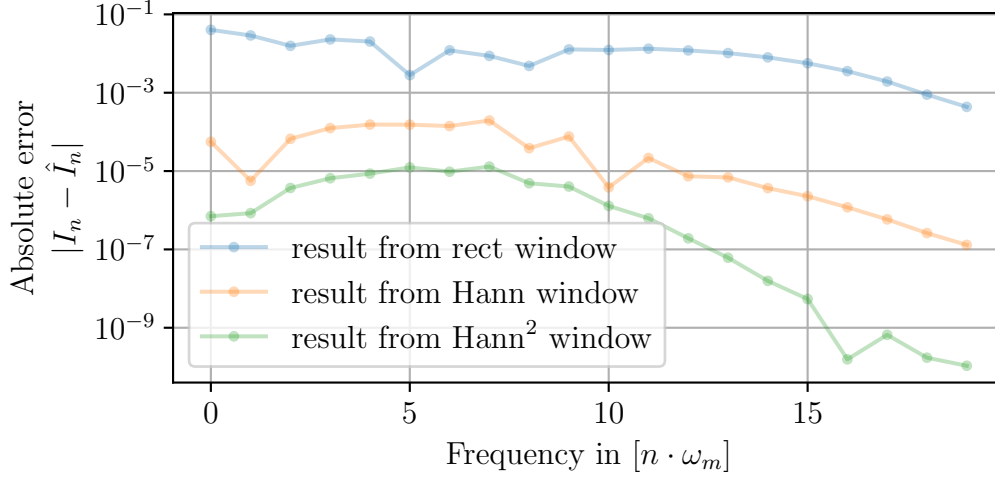
Figure 4.4: Plot showing the absolute error $\Delta I_n = |I_n - \hat{I}_n|$ of the calculated I_n coefficients calculated from s_{example} when plugged into (4.12) for different window functions $W(t)$. To showcase the filtering effect of the window, the measurement time was slightly offset ($1.01 \cdot T_{\text{ex}}$) from an exact integer multiple of the modulation period $2\pi/\omega_m$ and no additional noise was added ($\sigma^2 = 0$), leaving only the truncation noise from the finite digital precision. $\Delta I_n = I_n - \hat{I}_n$ is the difference between the exact values I_n written in the second column (calculated from Table 4.2), and the values calculated $\hat{I}_n$ from (4.12).

window function with a steeper fall-off like a Hann window ($\propto 1/f^3$) or a squared Hann window ($\propto 1/f^5$) can further mitigate these errors. Figure 4.4 shows the error of some calculated I_n coefficients for different window function, showcasing this effect. In my algorithm implementation, I use a squared Hann-Window given by:

$$W_T(t) := \frac{4}{3T^2} \sin^2(\pi t/T) \qquad \text{with } \tilde{W}_T(\omega) \sim \frac{1}{\omega^5} \qquad (4.15)$$

Keeping the total measurement time close to an integer multiple of $2\pi/\omega_m$ (I choose a measurement time of 8 modulation periods), will keep the zeros at the other harmonic frequencies when calculating I_n and Q_n, and the steeper falloff from using (4.15) as window will help mitigate timing jitter noise. The integration operation (or sum) in (4.11) also factors in as another $1/f$ attenuation factor in frequency domain.

Remark 1: Choosing a squared Hann window has some subtle implications for the required measurement time T. While a rectangular window has zeros (in the frequency domain) at all frequencies $n \cdot 2\pi/T$ with $n \neq 0$; the Hann window has "only" zeros at all $n \notin \{-1, 0, 1\}$, and the squared Hann for all $n \notin \{-2, -1, 0, 1, 2\}$. Figure 4.5 illustrates this graphically as the different window function behave slightly
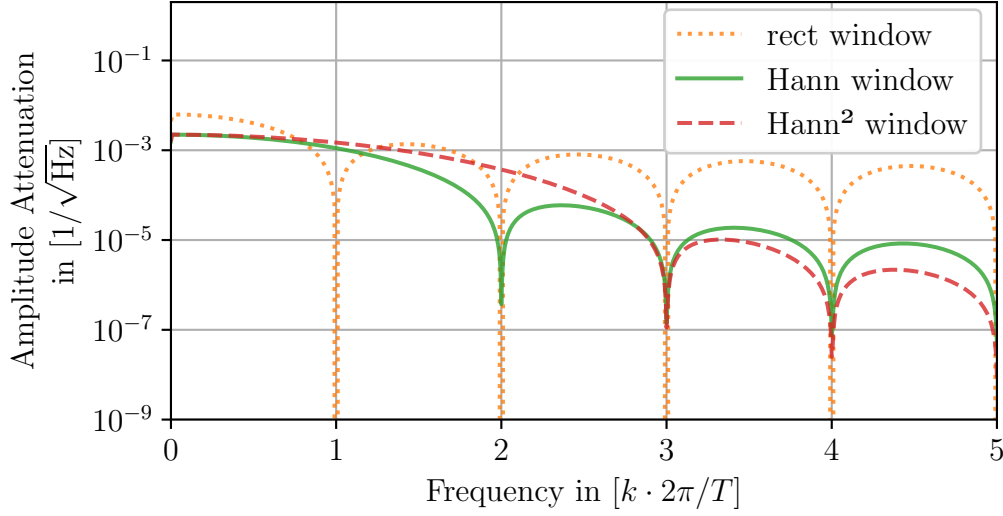
Figure 4.5: Plot of the ASD of a rectangular window, a Hann-window and a squared Hann window function, all with the same length T. While the Hann and squared Hann window have a steeper asymptotic falloff than just the rectangular window, the tradeoff can be a slightly higher contribution at low frequencies. While the rectangular window has its lowest zero at $\omega = 1 \cdot 2\pi/T$, the Hann window has a first zero at $\omega = 2 \cdot 2\pi/T$ and the squared Hann at $\omega = 3 \cdot 2\pi/T$.

differently at low frequencies. This means that with a Hann window, the measurement time needs to be at least 2 modulation periods long; and for the squared Hann window at least three modulation period; to have again a zero at $1 \cdot \omega_m$ from the window function.

Remark 2: In the general I/Q demodulation scheme, this filtering would be part of the "low-pass" filter stage which does not necessarily need to be realized by such a window function.

4.4.3 $|c_n|^2$ as SNR of the signal harmonics

From just the here calculated I_n and Q_n, one can already eliminate the ψ dependency directly by calculating

$$|c_n| := \sqrt{I_n^2 + Q_n^2} = \begin{cases} |A \cdot J_n(m) \cdot \cos\varphi| & n \text{ even} \\ |A \cdot J_n(m) \cdot \sin\varphi| & n \text{ odd} \end{cases} \quad . \tag{4.16}$$

For the later algorithm steps it turned out, however, that the missing sign information of these absolute $|c_n|$ is required to calculate exact values of the other signal parameters. Nevertheless, are the absolute $|c_n|^2$ as defined here still useful as they correspond to the signal harmonic's SNR. More specifically: Calculating the $|c_n|$ using the I_n and Q_n as defined in (4.11) is exactly the definition of the spectral density as written in (7.32), meaning that $|c_n|^2 = S_{ss}(n\omega_m)$.

 For a DFMI signal, the distribution of the signal energy over the DFM harmonics depends on, and changes with, the signal parameters. This is shown in Figure 4.6 by example of two DFM signals which are identical up to a different phase φ. At $\varphi = 0$ (like the blue curve in Figure 4.6), the signal energy is distributed mostly at
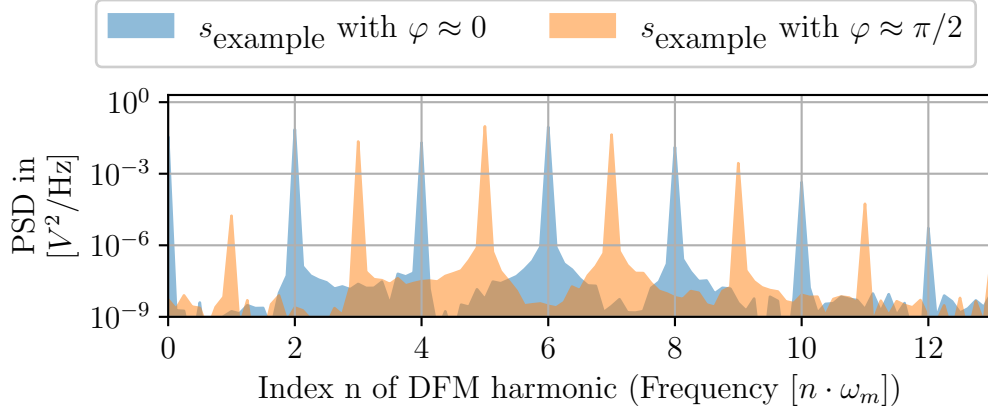
Figure 4.6: PSD of two DFMI signal with parameters identical to s_{example} as given in (4.1) but one time with $\varphi = 0$ and one time with $\varphi = \pi/2$. For these values of φ, either the even numbered or the odd numbered harmonics at frequencies $n \cdot \omega_m$ vanish.

the harmonics with n even ($n \in \{0, 2, 4, \ldots\}$) and no signal energy at the harmonics with n odd. With any kind of noise present, the calculated I_n, Q_n and $|c_n|$ coefficients for n odd are dominated by noise and have a high relative error.

We can see this also from the exact values in Table 4.2 as for e.g. $\varphi = 0$, all the odd I_n and Q_n becomes zero and the calculated values would in this case be only noise. For the case of $\varphi = \pi/2$ (the yellow curve in Figure 4.6), it is exactly the other way around with all the even harmonics being zero and all the odd harmonics carrying all the signal energy. Values of φ in-between will lead to different distributions of the signal energy over the "even" and "odd" harmonics.

To mitigate the noise introduced by these relative errors of the calculated I_n, Q_n coefficients, the $|c_n|^2$ ($\sim$ harmonic SNR) will be used to weight the individual coefficients.

4.5 Calculating the modulation phase ψ

The calculation of the modulation phase ψ is done similar to the algorithm by Heinzel et al. [54]. By calculating the arctan of the ratio I_n/Q_n for the measured n'th harmonics and unwrapping the resulting values one obtains ψ.

4.5.0.1 Definition of ψ_{estimate}

As first step we calculate

$$\Psi_n := \begin{cases} \arctan(-Q_n/I_n) & \text{n even} \\ \arctan(I_n/Q_n) & \text{n odd.} \end{cases} = n\psi \bmod \pi \qquad (4.17)$$

and unwrap the resulting list by adding $+\pi$ at "jumps" where $\Psi_{n+1} < \Psi_n$. Averaging over the gradient of this unwrapped list would then yield an estimate for ψ.

As shown above in section 4.4.3, either even or odd harmonics can be very noise and add relatively larger errors to a ψ_{estimate} calculated this way. As small adjustment, I split the list of Ψ_n into even and odd harmonics $\implies \Psi_{\text{even}} = \{\Psi_0, \Psi_2, \Psi_4, \dots\}$ and $\Psi_{\text{odd}} = \{\Psi_1, \Psi_3, \Psi_5, \dots\}$ and unwrap the resulting lists separately (by adding $+\pi$ at jumps where $\Psi_{n+2} < \Psi_n$). Then, we can calculate the gradient of these lists by using some finite difference technique and calculate a weighted average, using $|c_n|^2$ as weight to account for the individual harmonic's SNR. The result of this calculation is ψ_{estimate}.

For a ψ_{estimate} calculated this way there remains, however, an ambiguity of $\psi_{\text{estimate}} \approx \psi \bmod \pi$. Due to the arctan mapping only into an interval of $[-\pi/2, +\pi/2]$, a phase shift of $+\pi$ yields the exact same Ψ_n in (4.17) and subsequently the same ψ_{estimate}. In other phase unwrapping application, this is commonly overcome by using an "arctan2" routine which also compares the sign of the sin and cos terms. In our case this is not possible since the J_n and $\sin\varphi/\cos\varphi$ factors affect the signs of the I_n & Q_n as well.
While this algorithm cannot calculate an unambiguous estimate of ψ, the ambiguity of ψ_{estimate} does not hinder the algorithm to get accurate (and unambiguous) estimates of the other signal parameters as we will later see.

4.5.1 Ψ_n as data quality check

In my experimental setups the laser modulation signal $\omega_m t + \psi$ is a set parameter which can be measured independently of the DFMI signal. Even when the ψ is given or can be measured directly, the here present Ψ_n and ψ_{estimate} are, however, still useful to estimate the noise at exactly the signal harmonics. E.g. Figure 4.7 shows the calculated Ψ_n and ψ_{estimate} for two signals. In one case (left side of the Figure) there is very high level of white (frequency independent) noise, which is

clearly visible in the time series; but has only a small effect at the DFM harmonics, leading to still accurate Ψ_n estimates. The lower part of Figure 4.7 seems, in its time series, to be less noisy; calculating the Ψ_n reveals however more noise at exactly the DFM harmonics as the calculated Ψ_n values do not "line up" nicely (on the line given by $n \cdot \psi_{\text{estimate}}$). This way, calculating the difference of $|\Psi_n - n \cdot \psi_{\text{estimate}}|$ can still be used as a measure of noise present at the DFM harmonics and is an indicator for the quality of the data even before calculating the remaining signal parameters.
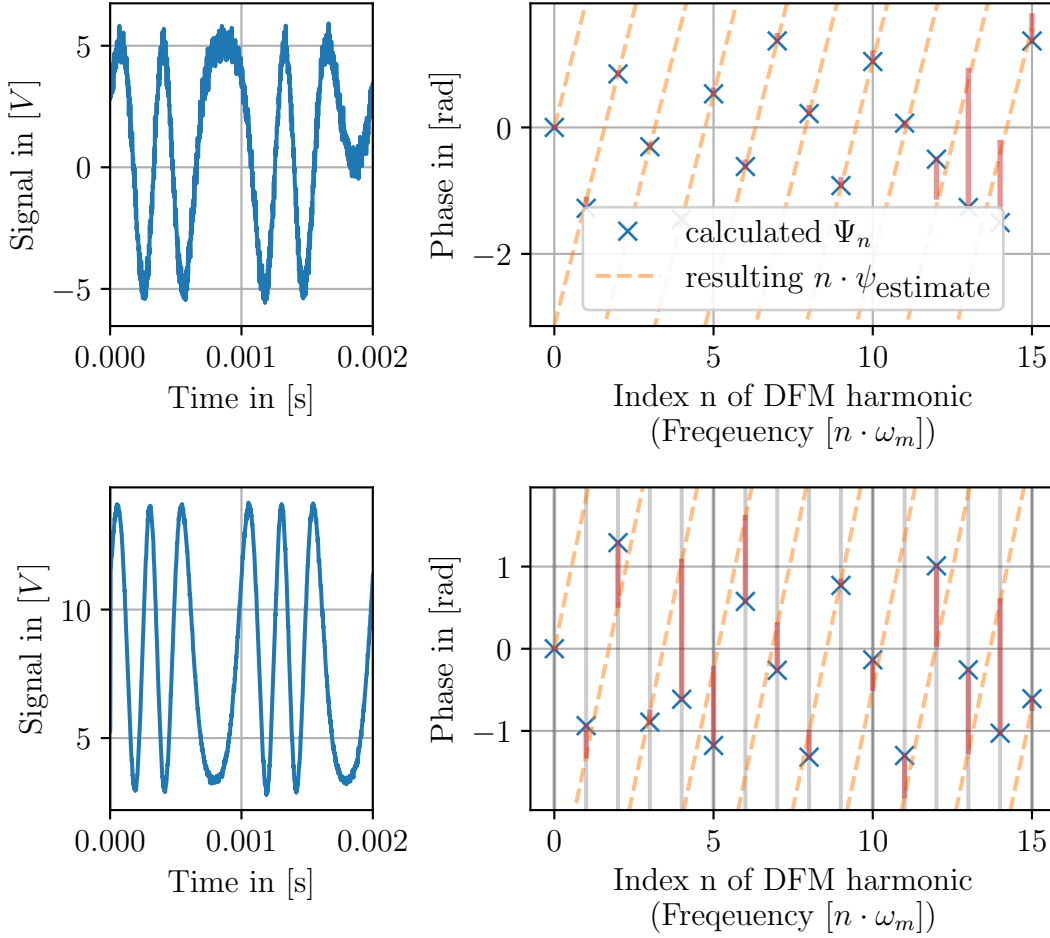
Figure 4.7: Plots of two DFM signals and the calculated Ψ_n and ψ parameters. The upper two plots show a part of the time series of the measured signals and the lower plot shows the calculated Ψ_n (as blue crosses) and the calculated ψ_{estimate} (and dashed yellow line with $y = n \cdot (\psi_{\text{estimate}} + \pi \cdot \mathbb{N})$) from the same signals. The upper two plots correspond to the simulated signal s_{example} but with $\times 10$ more white noise and a larger modulation phase of $\psi \approx 2$. The lower plots are from real measured DFM data from a tabletop DFMI experiment. While the upper signal looks much nosier (and has more 'white' noise), the calculation of the Ψ_n (and all following parameters) yield much better results compared to the lower signal. The reason in this case is that the lower signal has much more (non–Gaussian) noise located at exactly the signal harmonics (i.e. due to ghost beams), which is not obvious from looking at the shown time series alone.

4.6   Definition of the c_n coefficient
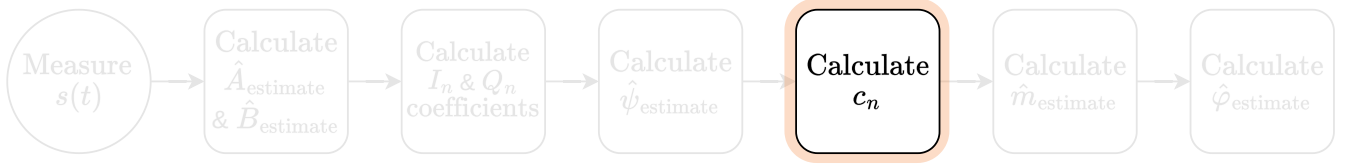
The next step in the algorithm is to eliminate the ψ dependency from the calculated coefficients. While the ψ dependency can be eliminated directly from the I_n and Q_n coefficients by simply summing their squares as done in section 4.4.3, the resulting absolute value is insufficient for the later calculation steps where the sign will play a significant role. Instead, I will use either a given, or the previously calculated ψ value, to eliminate it from the I_n and Q_n.

4.6.1   Calculation of the (signed) c_n

By using either the previously calculated ψ_{estimate} or an externally provided ψ, I define the c_n coefficients as:

$$c_n := \begin{cases} +I_n \cdot \cos n\psi - Q_n \cdot \sin n\psi & \text{n even} \\ -I_n \cdot \sin n\psi - Q_n \cdot \cos n\psi & \text{n odd} \end{cases} = \begin{cases} A \cdot J_n(m) \cdot \cos \varphi & \text{n even} \\ A \cdot J_n(m) \cdot \sin \varphi & \text{n odd} \end{cases}.$$

$$(4.18)$$

Using $\psi_{\text{estimate}} = (\psi \bmod \pi)$ in place of ψ here can still lead to a wrong sign of all c_n, as replacing $\psi \mapsto \psi + \pi$ in equation (4.18) will yield $-c_n$ instead of $+c_n$. This "error" is however consistent for all c_n, meaning they <u>all</u> have a flipped sign which will cancel out later on and not be an issue for further calculations.

4.7   Calculating the modulation index m
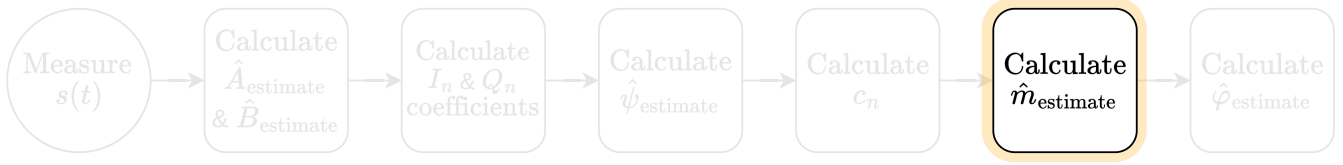
4.7.1   Definition of m_n

The crucial step in the algorithm is the estimation of the modulation index m. The recurrence relation (4.19) for Bessel-functions of the first kind can be rearranged to yield m m:

$$J_{n-1}(m) + J_{n+1}(m) = \frac{2n}{m}J_n(m) \implies m = \frac{2nJ_n(m)}{J_{n-1}(m) + J_{n+1}(m)} \quad \text{for } m \in \mathbb{Z}\,.$$

(4.19)

Since the c_n coefficients contain a $\sin\varphi$ factor for n even and $\cos\varphi$ factors for n odd, one cannot directly insert them in (4.19). However, applying the recursive relation twice yields

$$m = \sqrt{\frac{4n(n-1)(n+1)J_n(m)}{2nJ_n(m) + (n+1)J_{n-2}(m) + (n-1)J_{n+2}(m)}}\,.$$

(4.20)

Inserting the c_n coefficients obtained from (4.18) into this relation in place of the J_n, the A and φ factors cancel out and only the $J_n(m)$ factors remain such that

$$m_n := \sqrt{\frac{4n(n-1)(n+1)c_n}{2nc_n + (n+1)c_{n-2} + (n-1)c_{n+2}}} = m \quad \text{for } \forall n\,.$$

(4.21)

These m_n correspond to an estimate of m for each set of 3 harmonics with indices $(n-2, n, n+2)$. Figure 4.8 shows these calculated m_n values for an example DFMI signal. The final step would now be to average the m_n from the sets of 3 harmonics to calculate an m estimate.

Remark: For noise dominated harmonics, leading to highly erroneous c_n, the value under the square-root of (4.21) becomes also highly erroneous and can even become negative. To avoid errors in the implementation, use the absolute value of the expression under the square-root and effectively discard the erroneous m_n later by using a weight that takes the harmonics individual signal-to-noise ratio into account.

Without any prior knowledge, i.e., when running the algorithm on a first set of data, use the power of the harmonics $|c_n|^2$ as weight when averaging over the m_n values. To achieve the highest possible precision it is however necessary to use a specific weighting function as shown in the next section.

4.7.2   Averaging over individual harmonics

In the algorithm, every harmonic yields a c_n coefficient which are used to calculate estimates for m and later φ. When using just the harmonics signal powers $\sim |c_n|^2$
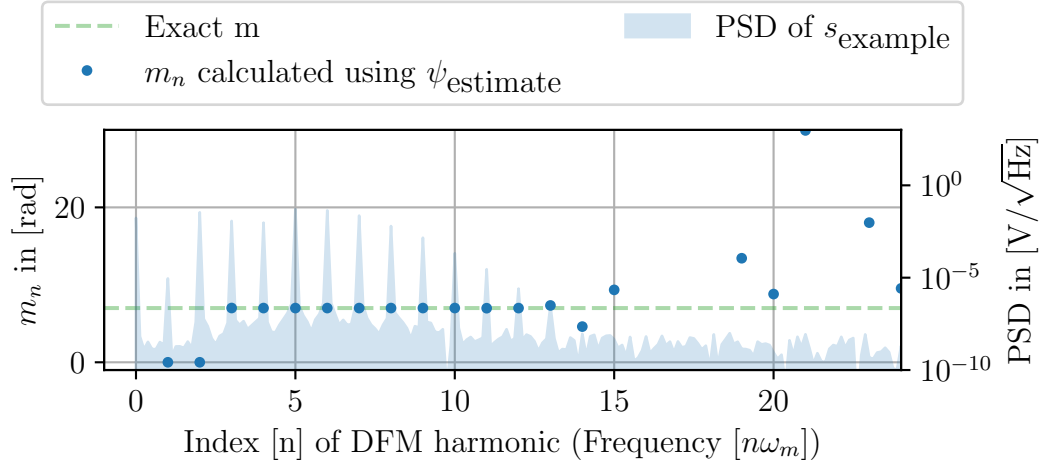
Figure 4.8: Plot of calculated m_n values for s_{example}. The blue dots correspond to the calculated m_n values with the dashed green line as exact m, both scaled to the left y-axis. The filled light-blue area corresponds to the PSD using the right y-axis. Comparing both shows that at harmonics with a relatively high power, the calculated m_n are close to the exact value while the higher harmonics are 'hidden' below the white noise floor leading to erroneous m_n values. Since 2 more harmonics to either side are needed to calculate a value of m_n, the lowest two used harmonics yield no values here (i.e., $m_1 = m_2 = 0$).

as weight, the resulting parameter estimate can still become highly erroneous and many order of magnitude above their ideal precision[2]. Figure 4.9 shows the error of the calculated m_n coefficients for two example DFMIs signal with two different modulation indices. For the particular case of the 100 used harmonics in the right plot of Figure 4.9, the difference between smallest to the largest error of the m_n is over 4 orders of magnitude.

Due to the linearity of the equations, any white noise in the signal will lead to the same white noise level for the I_n, Q_n and c_n coefficients. In an attempt to estimate the error of the calculated m_n I calculate the linear error propagation of equation (4.21) for variations of the input c_n coefficients. E.g. adding the errors $(\delta c_{n-2}, \delta c_n, \delta c_{n+2})$ to (4.21) yields:

$$m_n = m_n(J_{n-2}(m) + \delta c_{n-2}, J_n(m) + \delta c_n, J_{n+2}(m) + \delta c_{n+2}) \tag{4.22}$$

$$\approx m + \sum_{k \in \{n-2,n,n+2\}} \frac{\partial m_n}{\partial c_k} \cdot \delta c_k + \mathcal{O}(\delta c^2) \tag{4.23}$$

$$\overset{(4.21)}{=} m + \frac{m}{8J_n(m)}\left[4\delta c_n - \frac{m^2}{n(n^2-1)}\big((n-1)\delta c_{n+2} + 2n\delta c_n + (n+1)\delta c_{n-2}\big)\right] \tag{4.24}$$

$$\implies m_n - m = \frac{m}{8J_n(m)}\left[4\delta c_n - \frac{m^2}{n(n^2-1)}\big((n-1)\delta c_{n+2} + 2n\delta c_n + (n+1)\delta c_{n-2}\big)\right]. \tag{4.25}$$

---

[2]Given by the CRB explained in chapter 8
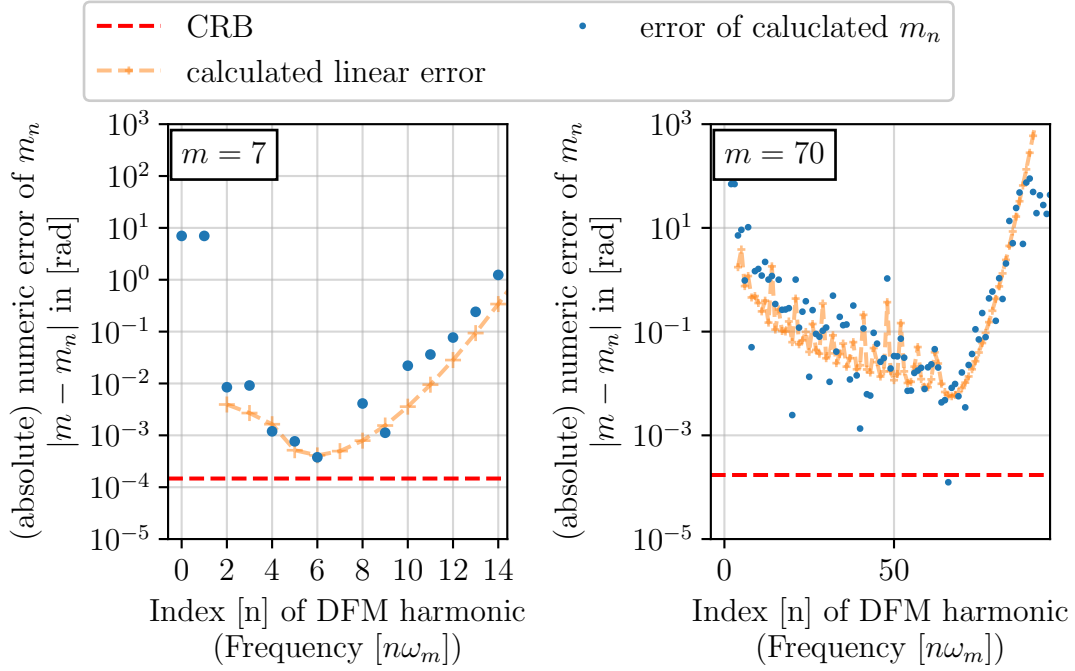
Figure 4.9: Absolute error $|m - m_n|$ of the calculated m_n values for s_{example} (left plot) and for an almost identical signal but with a $\times 10$ larger modulation index $m = 70$ (right plot). The right signal with its larger m value corresponds to an identical DFMI setup but a larger absolute distance ($\times 10 L$) being measured. In case of a small m (left plot), the signal energy is distributed between fewer harmonics compared to the case of a larger m (right plot), and the individual harmonics can have a smaller relative error close to the CRB. The average is however roughly the same for both cases given by the CRB (explained in the next part of this thesis). The errors at the edges of the spectrum (the two largest and the two smallest harmonics) are the result of equation (4.21) not being applicable as there are neighboring harmonics missing. The δc for the calculated linear error of (4.26) is here given by the σ parameter of the simulated Gaussian noise.

Approximating the $\delta c_{n+2}, \delta c_n, \delta c_{n-2}$ by some upper error δc (i.e., $\delta c_k \leq \delta c \; \forall k$) then yields:

$$(m_n - m) \lesssim \frac{m}{2|J_n(m)|} \left[1 + \frac{m^2 \left(1 + \frac{1}{2n}\right)}{(n^2 - 1)} \right] \cdot \delta c. \tag{4.26}$$

The result of this calculated error estimation is also shown in Figure 4.9 as "calculated linear error" where it aligns with the measured error of the m_n coefficients. To reach a high precision readout I therefore use the inverse of (4.26) as weight for the m_n coefficients and approximate the harmonics individual δc_n error by (the inverse of) their signal power $|c_n|^2$. The resulting weight is then given by

$$w_n = \frac{2|J_n(m)|}{m} \left(1 + \frac{m^2 \left(1 + \frac{1}{2n}\right)}{(n^2 - 1)} \right)^{-1} \cdot |c_n|^2. \tag{4.27}$$

A drawback is that this weight itself also depends on the modulation index m, making it unusable for the initial weighting. The algorithm therefore performs the weighting of the m_n twice: the first time, an m_{initial} estimate is calculated using the harmonics signal power $|c_n|^2$ as weight. This m_{initial} is then used in (4.27) to

calculate the more precise weight and average over the m_n a second time to get a high precision estimate for m.

$$m_{\text{initial}} := \frac{\sum_n m_n \cdot |c_n|^2}{\sum_n |c_n|^2} \quad \text{and} \quad \hat{m}_{\text{estimate}} := \frac{\sum_n m_n \cdot w_n(m_{\text{initial}})}{\sum_n w_n} \tag{4.28}$$

For an application with a continuous readout where the changes of the signal parameters are sufficiently small between measurements, it is also possible to simply use the m estimate of the previous measurement to calculate the w_n weights directly (and thus reducing the number of necessary calculations).

4.7.3 Alternative calculation for m_{estimate} using $|c_n|$

The advantage of calculating the absolute $|c_n|$ directly from

$$|c_n| = \sqrt{I_n^2 + Q_n^2}, \tag{4.29}$$

is that it eliminates the ψ dependency without needing any knowledge of ψ and without adding any errors from an imprecise ψ_{estimate}. Figure 4.10 shows an extreme case, where the calculated ψ_{estimate} is very erroneous, resulting in erroneous c_n and m_n coefficients. Calculating the m_n for the same noisy signal, but using the absolute $|c_n|$ from (4.29) and any possible combination of signs for the $\pm|c_n|$ in (4.21) yields however much better values for some m_n candidates.
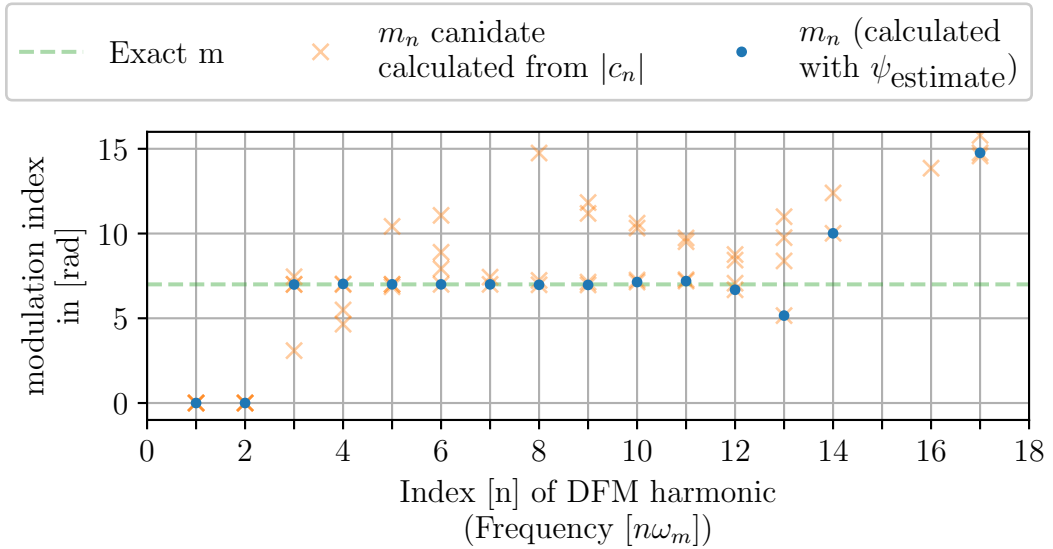

Figure 4.10: Plot of the calculated values for m_n for the example signal s_{example} but with $\times 10$ more white noise (which lead to an erroneous ψ_{estimate}). The dashed green line marks the exact value of m of the simulated signals. The blue dots mark the m_n calculated from (4.21) when using the c_n calculated with ψ_{estimate}. The yellow crosses mark the m_n calculated from the all possible sign combinations when using the absolute $|c_n|$ of (4.29).

Examining (4.21) again shows that there are 4 unique combination of signs for the three given c_n values as written in Table 4.3. Each of these 4 combinations yields a "candidate" for the correct m_n value as, plotted as yellow cross in Figure 4.10, but there is no way of knowing initially which is the correct one per harmonic.

Sign of the	c_n	c_{n+2}	c_{n-2}
	$+$	$+$	$+$
	$+$	$-$	$+$
	$+$	$+$	$-$
	$+$	$-$	$-$

Table 4.3: Combination of signs for the c_n leading to different results for m_n calculated from (4.21). While there are in total $2^3 = 8$ possible combinations, the symmetry of (4.21) reduces them to only 4.

My initial approach was to calculate all possible m_n candidates, 4 for each harmonic, and try to cluster the resulting values. The resulting most dense region should give a first approximate value for m and reveal the correct signs of the c_n, allowing for a proper calculating of an m estimate as weighted average over the harmonics. Practically, it turned out however that an m estimate calculated this way was prone to glitching specifically around phase values of $\varphi = 0$ or $\varphi = \pi/2$ where all of the signal energy was either at the even or odd harmonics and the accuracy of the resulting m estimate was in most scenarios worse than the ones calculated from the signed c_n when using ψ_{estimate}.

4.7.4 Absolute length calculated from m

In an DFMI context, the m parameter contains the information on the absolute path length difference via $m = \Delta\omega \cdot \tau$ with τ as propagation time (difference) of the light (compared to the second interferometer arm). The absolute path length is given by

$$L = \frac{c_0}{\Delta\omega} \cdot m_{\text{estimate}} \,. \tag{4.30}$$

4.8   Calculating the interferometric phase φ
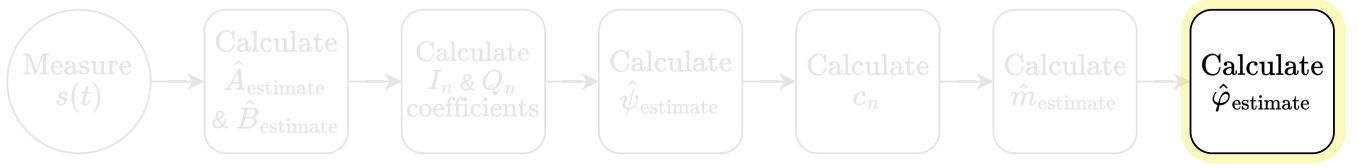
Since the summed up $|c_n|^2$ equal the total signal energy, the φ information is also entirely contained in the c_n coefficients. The remaining task would be to eliminate the A and m dependency without adding additional errors. While the amplitude A can be eliminated by calculation ratios of different c_n coefficients, eliminating the $J_n(m)$ is not so straight forward.

As the c_n scale differently for even and odd n (with $\sin/\cos \varphi$) the two quadratures have to be treated separately. Using the Jacobi-Anger identity and the recurrence relations of differences and sums of Bessel functions, I was able to find summation formulas for even and odd n:

$$\sum_{n=1}^{\infty} J_{2n}(m) = \frac{1}{2}(1 - J_0(m)) \tag{4.31}$$

$$\sum_{n=1}^{\infty} J_{2n-1}(m) = J_1(m) + \frac{1}{2}\int_0^m J_2(m')dm' \tag{4.32}$$

$$\sum_{n=1}^{\infty} J_{2n}^2(m) = \frac{1}{4}(1 - 2J_0^2(m) + J_0(2m)) \tag{4.33}$$

$$\sum_{n=1}^{\infty} J_{2n-1}^2(m) = \frac{1}{4}(1 - J_0(2m)) \tag{4.34}$$

Using these summation formulas; summing over even and odd c_n allows one to write either

$$\frac{\sum_{n=1}^{\infty} c_{2n-1}}{\sum_{n=1}^{\infty} c_{2n}} = \frac{2J_1(m) + \int_0^m J_2(m')dm'}{1 - J_0(m)} \cdot \tan \varphi \tag{4.35}$$

or

$$\frac{\sum_{n=1}^{\infty} c_{2n-1}^2}{\sum_{n=1}^{\infty} c_{2n}^2} = \left| \frac{1 - 2J_0^2(m) + J_0(2m)}{1 - J_0(2m)} \right| \cdot |\tan \varphi| \; . \tag{4.36}$$

To calculate φ, I use one of these summation formulas and remove the m dependent factor by multiplying with its inverse. The advantage of the second formula (4.36) is that it does not contain an integral and only depends on a single Bessel function J_0, and it also makes the calculation independent of the ψ estimation. The resulting $|\tan \varphi|$ has however a limited range of only $[0, \pi/2]$ compared to the full 2π range than can be gained from using an atan2 routine on the two quadratures of 4.35, which is why this first summation formula is preferred in this algorithm.

4.8.1 Definition of $\hat{\varphi}_{\text{estimate}}$

Summing over the even and odd harmonic amplitudes I define two phase quadratures as:

$$I_\varphi := \left(\frac{2}{1 - J_0(m)} \right) \cdot \sum_{n=1}^{\infty} c_{2n} = A \cos \varphi \qquad \text{and} \qquad (4.37)$$

$$Q_\varphi := \left(\frac{1}{J_1(m) + \frac{1}{2} \int_0^m J_2(m') dm'} \right) \cdot \sum_{n=1}^{\infty} c_{2n-1} = A \sin \varphi \qquad (4.38)$$

Similar to other interferometric phase readout procedures, we calculate the final estimate for φ from the two quadratures via

$$\hat{\varphi}_{\text{estimate}} := \operatorname{atan2}(Q_\varphi, I_\varphi) . \qquad (4.39)$$

4.8.2 Alternative calculation of $\hat{\varphi}_{\text{estimate}}$

An alternative and computationally more expensive way to calculate an estimate of φ would be to eliminate the m dependency individually per harmonic, where

$$\frac{c_n}{J_n(m)} = \begin{cases} A \cdot \cos \varphi & \text{for } n \text{ even} \\ A \cdot \sin \varphi & \text{for } n \text{ odd} \end{cases} . \qquad (4.40)$$

In this case the individual harmonic SNR can be accounted for by averaging over the harmonics as shown in section 4.7.2, which can improve the readout performance in some cases.

4.8.3 Relative displacement calculated from φ

From the calculated phase estimate, the 'microscopic' distance in a DFMI setup with mean wavelength λ_0 is then given by

$$\delta L = \frac{\lambda_0}{2\pi} \cdot \hat{\varphi}_{\text{estimate}} . \qquad (4.41)$$

4.9 Example readout performance

4.9.1 Readout precision for a large parameter range

Figure 4.11 shows the precision of the analytic readout algorithm for s_{example} with varying distance to the target. It shows that the analytic readout algorithm reaches the highest possible precision (for φ) for a large dynamic range starting from a few centimeters up a meter. (The details of the "highest possible precision," given by the Cramér-Rao bound, are explained and discussed in part III of this thesis).
Starting from a range of $\Delta L \approx 4\,\text{cm}$, the analytic readout algorithm consistently reaches the CRB for the φ readout for the entire simulated parameter range of up to $1\,\text{m}$. At smaller distances, the frequency spectrum does not contain enough harmonics above the noise level for the algorithm to work correctly leading to larger noise and glitches below. Since the algorithm requires at least three even or three
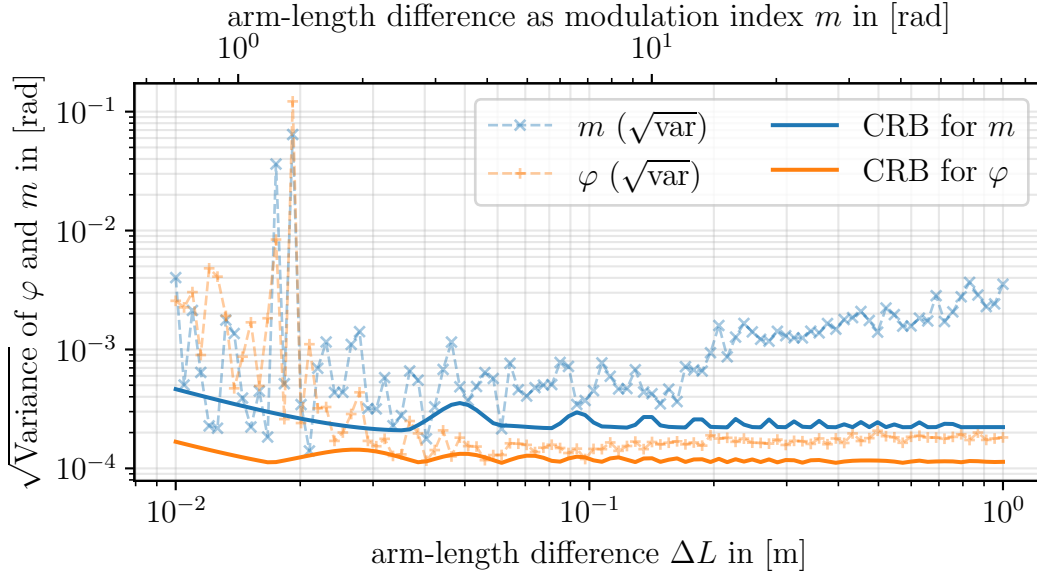
Figure 4.11: Plot showing the performance of the analytic readout algorithm for a simulated DFMI signal with the signal parameters as before (given in Table 3.1), with varying range / arm length difference ΔL.

odd harmonics to calculate estimate for m, if all the signal energy is concentrated in the first 4 (or less) harmonics (meaning 2 even and two odd), the calculated m and φ estimates will be dominated by noise.

The m calculation is generally less precise than the φ calculation. Around a target distance of $10 \pm 1\,\mathrm{cm}$, the modulation index too, reaches its highest precision, and is less optimal for larger distances. For a high precision absolute length measurement however, the modulation index only needs to reach a precision of roughly one wavelength to provide the distance information up to an optical fringe as the interferometric phase provides the high precision distance readout beyond that.

4.9.2 Readout precision for moving targets

Figure 4.12 shows the performance of the analytic readout algorithm for s_{example} and different scenarios of varying absolute distance. The algorithms results are displayed as distance to the target with the green curve corresponding to the macroscopic distance calculated from the modulation index via $L_{\mathrm{macro}} = \hat{m}_{\mathrm{estimate}} \cdot c_0 / \Delta\omega$ and the yellow curve as microscopic distance calculated from the unwrapped interferometric phase $L_{\mathrm{micro}} = \hat{\varphi}_{\mathrm{estimate}} \cdot c_0 / \omega_0$. Ideally, the macroscopic distance readout reaches a precision down to the wavelength (one fringe) of the used laser corresponding to 2π in interferometric phase. The distance readout is then truly absolute down to i.e. picometer precision and no tracking or unwrapping of the phase is necessary for a multi–fringe readout. (It is however not a 'hard' requirement for the functionality of the algorithm.)

For a still target (and only additive Gaussian noise with $\sigma^2 = 10^{-3}\,\mathrm{V}^2$) as displayed in Figure 4.12a, the analytic readout algorithm reaches $\approx 0.15\,\mu\mathrm{m}/\sqrt{\mathrm{Hz}}$ for the macroscopic distance and $\approx 1.1\,\mathrm{pm}/\sqrt{\mathrm{Hz}}$ for the microscopic distance.

Figures 4.12b and 4.12c show the readout performance for the same signal but for a moving target in which case $m(t)$ and $\varphi(t)$ change over the course of the measurement and are no longer constant as assumed so far. While the algorithm is able to follow the simulated movement (seen by the overlap between exact movement and algorithm readout), it is several orders of magnitude less precise compared to a still target. To first order, these examples show that the faster the target moves, the less precise the readout is.

In the context of a local DFMI displacement sensors in gravitational wave detectors, the readout algorithm does not need to reach the same (highest) precision for every scenario. It is sufficient if the readout can detect the movement precise enough for an actuator to slow the target down, leading again to a more precise readout and ultimately to a graceful degradation of the target's displacement (readout). By analyzing the changes of the signal in case of a moving target in more detail, I found that the algorithm can be extended to improve its precision for moving targets. The extension and some discussion of the changes of the signal are presented in the next chapter.

### 4.9.3   Readout speed using a python implementation

To run and test the analytic readout algorithm presented in this thesis, I created a reference implementation in python3 (available upon request). The python implementation does not implement any kind of parallelization to speed up the calculation of the individual harmonics and simply calculates the signal parameters as presented in this chapter for a given number of harmonics (usually from $n = 1$ to 16).

With all algorithm steps including the demodulation a single algorithm run, measuring the runtime of a single calculation (i.e. demodulation of 16 harmonics and calculation of the parameter estimates), using the pythons "timeit" module, leads to an average runtime of 11.13 ms ($\sim 89.8\,\text{Hz}$). Which is, for this unoptimized implementation, similar to the previously used fitting algorithm.

Implementing the demodulation of the harmonics into an FPGA and performing only the calculations using the I_n and Q_n coefficients (still not parallelized) reduces the average time of a single run to $\sim 0.5\,\text{ms}$ ($\sim 2\,\text{kHz}$), which corresponds to a speed improvement by a factor of $\times 20$. Further paralleling the calculations into different threads for each harmonic could potentially further improve the readout speed by a factor of the number of the used harmonics.
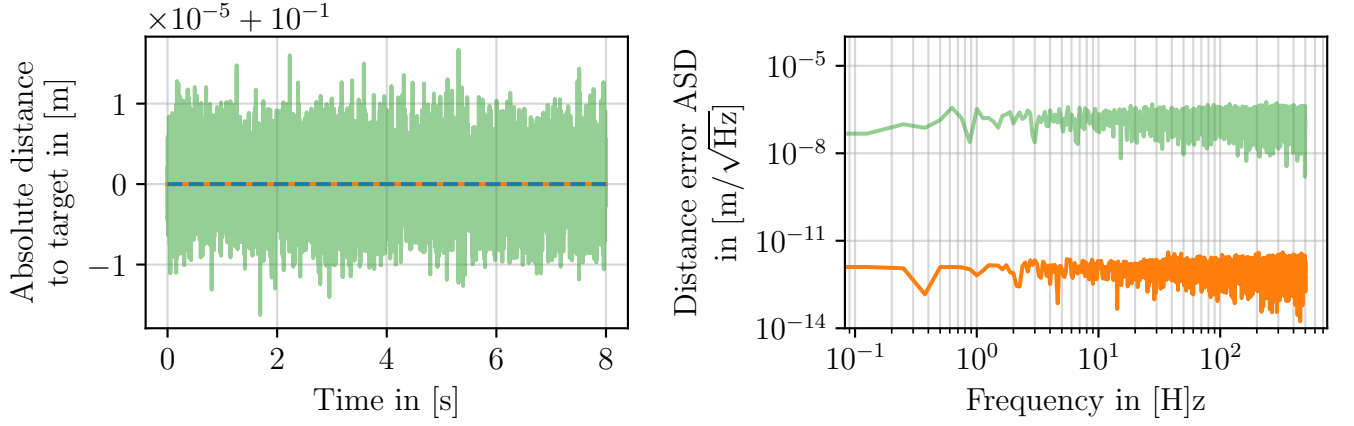
(a) Analytic Algorithm performance for a still target. To the left shown as time series, and to the right the resulting ASD.
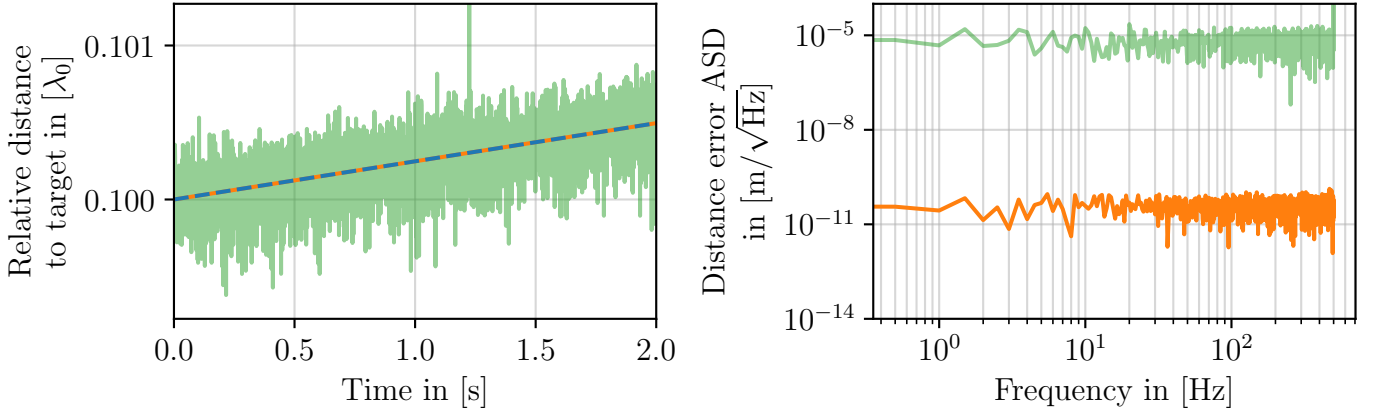
(b) Analytic Algorithm performance for a linearly moving target. While the macroscopic distance readout is too imprecise for a correct distance within a single fringe, unwrapping that phase readout (by adding $\pm 2\pi$ at jumps larger than $|\pi|$ between two consecutive measurements) still yields the correct absolute distance. The constant speed of the target is $\approx 0.25\,\mathrm{mm/s}$.
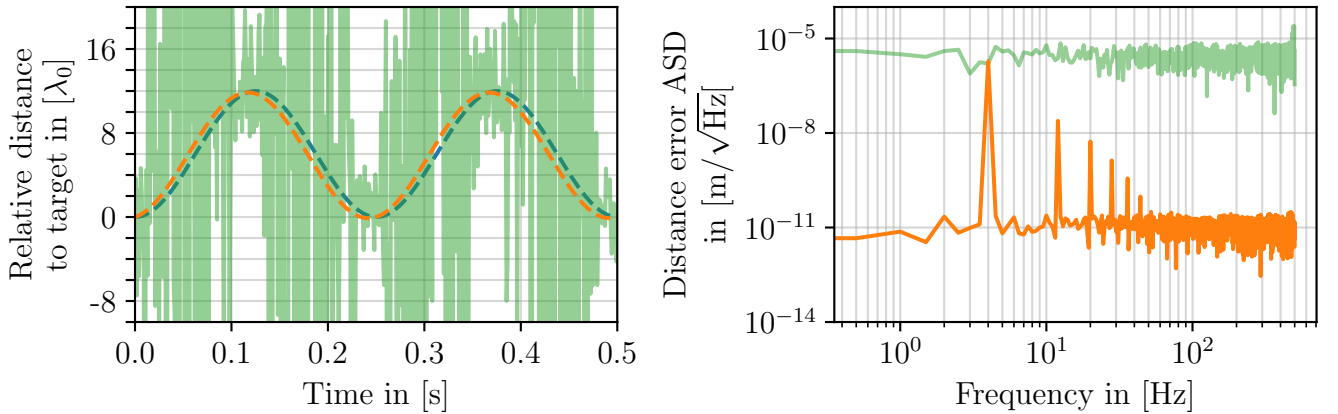
(c) Analytic Algorithm performance for a sinusoidal moving target. The phase is unwrapped as before and the maximum speed of the target is $\approx 0.23\,\mathrm{mm/s}$. While the absolute movement in smaller than the purely linear movement above, both microscopic and macroscopic distance readout reach roughly the same precision with the phase being less precise with significant errors at multiples of the movement frequenecy of $4\,\mathrm{Hz}$.

Figure 4.12: Output of the analytic readout algorithm for s_{example} for different distances $L(t)$ to the target. In all three cases the initial distance to the target at $t = 0$ is given by $L_0 = 10\,\mathrm{cm}$. In (a), this distance remains constant and only the noise varies between consecutive measurements. For (b) the distance is given by $L(t) = L_0 + 160\lambda_0 \cdot t$ and for (c) by $L(t) = L_0 + 6\lambda_0(1 - \cos(2\pi \cdot 4\,\mathrm{Hz} \cdot t))$.

Chapter 5

Dynamic DFMI signals

In DFMI, the length information ΔL in encoded in the modulation index $m = \Delta\omega \cdot \Delta L/c$ and the interferometric phase $\varphi = (\omega_0 \cdot \Delta L/c \mod 2\pi)$ with c as speed of light. So far we considered the phase φ and the modulation index m to be constant over the course of a DFMI measurement. When using DFMI sensors to measure the local displacement in gravitational wave detectors, we are particularly interested in the changes of the distance ΔL over time, where ΔL is not constant. In this chapter I will look into the effect of a moving target on the DFMI signal and present how the analytic readout algorithm can be adjusted to improve its performance for fast moving targets.

5.1 DFMI signals with linear (phase) term

We start by considering a DFMI sensor measuring a linear moving target. For the arm length difference of two interferometer arms we write

$$\Delta L_{\mathrm{lin}}(t) = \Delta L_0 + \delta v \cdot t + \mathcal{O}(t^2) \tag{5.1}$$

with δv as momentarily speed of the target. The corresponding DFM signal then becomes

$$s_{\mathrm{lin}}(t) := B + A\cos\left((\varphi_0 + \delta\omega \cdot t) + (m_0 + \delta m \cdot t) \cdot \sin\left(\omega_m t + \psi\right)\right) \tag{5.2}$$

$$= B + A\sum_{n\in\mathbb{Z}} J_n(m + \delta m \cdot t)\cos\left((n\omega_m + \delta\omega)t + n\psi + \varphi\right). \tag{5.3}$$

with

$$\delta\omega := \frac{\omega_0}{c}\delta v \qquad \text{and} \qquad \delta m := \frac{\Delta\omega}{c}\delta v \ . \tag{5.4}$$

Equation (5.3) shows that the DFMI signal for a linearly moving target consists of a (Doppler) frequency shift $\delta\omega$ of the harmonics and a shift of the modulation index δm.

For a still target with $\delta\omega = 0$, the parts in the sum of (5.3) with positive and negative n indices have the same frequencies, leading to overlapping signals for every $n\omega_m$ harmonic. For $\delta\omega \neq 0$ this is not the case as all of the harmonics are shifted in the direction of $\delta\omega$ and appear to split into a 'positive' and a 'negative' index part since

$\cos(n\omega_m t + \delta\omega t) \neq \cos(-n\omega_m t + \delta\omega t) = \cos(n\omega_m t - \delta\omega t)$. Figure 5.1 shows this behavior for a small $\delta\omega < \omega_m$ where instead of a single harmonic at exactly $n\omega_m$, two harmonics with $n\omega_m + \delta\omega$ and $n\omega_m - \delta\omega$ appear.
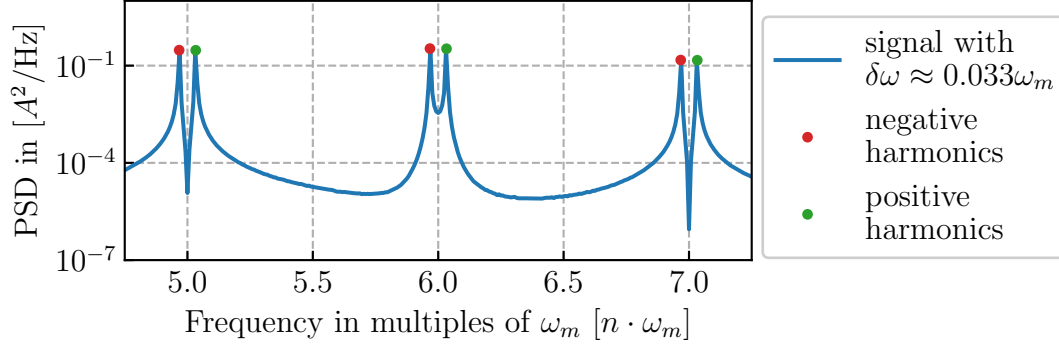

Figure 5.1: Example of a DFM signal with signal parameters as in Table 3.1 and a 'small' ($\delta\omega < \omega_m$) Doppler frequency shift. The positive and negative index parts of the signal harmonics become visible and do no longer overlap.

For large enough frequency shifts, the signal is shifted so much that positive and negative harmonics are no longer close to each other but clearly separated, and the signal appears symmetric around the Doppler frequency as seen in Figure 5.2.
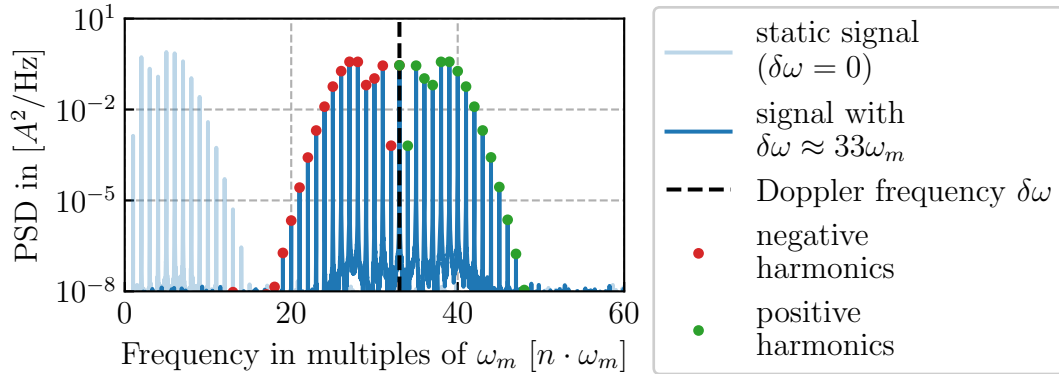

Figure 5.2: Example of a DFM signal with signal parameters as in Table 3.1 and a 'large' ($\delta\omega \gg \omega_m$) Doppler frequency shift. The DFM signal is shifted so far that all positive and negative index parts of the signal harmonics become separate and not even neighbor their counterpart.

If one would run an unmodified DFM readout algorithm for static DFM signals on such dynamic signals, two major errors would occur: First, the demodulation step would demodulate the wrong frequencies. Since the harmonics are shifted; they also need to be demodulated at their shifted frequencies. Secondly, the calculated phase of the demodulated harmonics picks up an additional phase ($\delta\omega \cdot T$) due to the Doppler shift of the carrier frequency. For the algorithm to work in the presence of a large Doppler shift, an estimator for $\delta\omega$ is needed to (a) demodulate the DFM harmonics at their correct frequencies ($n\omega_m \pm \delta\omega$) and (b) to account for the additional phase shift $\delta\omega T$ that is added to φ over the course of the measurement period T.

The effects of δm on the readout are mostly neglected in the following section and only treated as additional noise. In typical DFMI applications, where the modulation depth is much smaller than the mean laser frequency $\Delta\omega \ll \omega_0$, the error due to δm is similarly smaller $\delta m \ll \delta\omega$ and can be mostly neglected.

For a DFMI setup using a 1550 nm laser the plotted signals corresponds to an end-mirror moving with $\delta v \approx 0.26\,\mu\text{m/s}$ for the 'slow' case (Figure 5.1) and $\approx 26\,\text{mm/s}$ for the 'fast' case (Figure 5.2). At these velocities, the target already moves many wavelengths over the course of a single measurement and a high precision sub–micrometer readout is practically not necessary.

5.2 Dynamic Extension of the analytic readout algorithm

The simplest way to improve any DFMI readout precision for dynamic signals is to increase the speed of the readout. By shortening the measurement period, the absolute changes of the parameters also become smaller. Depending on the experimental setup, this is however not always possible.

For linearly changing distances, the analytic readout algorithm can be adjusted to compensate for most of the readout errors cause by the movement. For this case, we approximate the dynamic DFMI signal (5.3) further by dropping the δm, leaving

$$s_{\text{dynamic}}(t) = B + A\sum_{n\in\mathbb{Z}} J_n(m)\cos\left((n\omega_m + \delta\omega)t + n\psi + \varphi\right). \qquad (5.5)$$

The main idea to account for the additional Doppler frequency $\delta\omega$.
In the following I also restrict my analysis to speeds where the Doppler shift is smaller than the modulation frequency $\delta\omega \leq \omega_m$. While the extension presented here generally also works for larger Doppler shifts; if the Doppler shift is close to an integer multiple of the modulation frequency, some harmonics start to overlap and cannot be isolated/demodulated properly. While those harmonics could just be ignored for the readout; it introduces additional edge cases that will need additional steps for the algorithm to work, which I do not include here for simplicity and because they are outside the intended use case of local displacement sensing for seismic isolation.

5.2.1 Demodulation of 'positive' and 'negative' harmonics

Instead of a single demodulation for each harmonic we implement two for each of the split harmonics. Demodulating at exactly $(n\omega_m + \delta\omega)$ and $(n\omega_m - \delta\omega)$ by calculating

$$I_{n,\pm} := \frac{1}{T}\int_0^T dt\ s(t)\cdot\cos((n\omega_m \pm \delta\omega)t)\cdot W(t)$$

$$Q_{n,\pm} := \frac{1}{T}\int_0^T dt\ s(t)\cdot\sin((n\omega_m \pm \delta\omega)t)\cdot W(t) \qquad (5.6)$$

yields slightly different I and Q coefficients ($\mapsto I_{n,\pm}$ and $Q_{n,\pm}$) given by

$$Q_{n,+} = \frac{A}{2} \cdot J_n(m) \cdot \big[- (\sin\varphi - (-1)^n \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \cos(n\psi)$$
$$- (\cos\varphi + (-1)^n \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \sin(n\psi) \big]$$
$$(5.7)$$

$$Q_{n,-} = (-1)^n \cdot \frac{A}{2} \cdot J_n(m) \cdot \big[+ (\sin\varphi - (-1)^n \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \cos(n\psi)$$
$$- (\cos\varphi + (-1)^n \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \sin(n\psi) \big]$$
$$(5.8)$$

$$I_{n,+} = \frac{A}{2} \cdot J_n(m) \cdot \big[- (\sin\varphi - (-1)^n \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \sin(n\psi)$$
$$+ (\cos\varphi + (-1)^n \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \cos(n\psi) \big]$$
$$(5.9)$$

$$I_{n,-} = (-1)^n \cdot \frac{A}{2} \cdot J_n(m) \cdot \big[+ (\sin\varphi - (-1)^n \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \sin(n\psi)$$
$$+ (\cos\varphi + (-1)^n \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)) \cos(n\psi) \big]$$
$$(5.10)$$

In the limit of $\delta\omega \mapsto 0$, the $Q_{n,\pm}$ and $I_{n,\pm}$ coefficients converge to the previously introduced I_n and Q_n coefficients from Table 4.2. These $Q_{n,\pm}$ and $I_{n,\pm}$ can now be combined to factor out the $J_n(m)$ and ψ dependence written in Table 5.1.

	n even	n odd
$(Q_{n,+} + Q_{n,-})$	$-AJ_n(m)\Big(\cos\varphi + \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\sin(n\psi)$	$-AJ_n(m)\Big(\sin\varphi + \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\cos(n\psi)$
$(I_{n,+} + I_{n,-})$	$AJ_n(m)\Big(\cos\varphi + \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\cos(n\psi)$	$-AJ_n(m)\Big(\sin\varphi + \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\sin(n\psi)$
$(Q_{n,+} - Q_{n,-})$	$-AJ_n(m)\Big(\sin\varphi - \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\cos(n\psi)$	$-AJ_n(m)\Big(\cos\varphi - \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\sin(n\psi)$
$(I_{n,+} - I_{n,-})$	$-AJ_n(m)\Big(\sin\varphi - \sin(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\sin(n\psi)$	$AJ_n(m)\Big(\cos\varphi - \cos(\delta\omega T + \varphi)\mathrm{sinc}(\delta\omega T)\Big)\cos(n\psi)$

Table 5.1: Table of additional coefficients derived from equations (5.7) to (5.10).

5.2.2 Modified parameter estimation

By replacing the static I_n and Q_n with these new, dynamic $I_{n,\pm}$ and $Q_{n,\pm}$ coefficients the analytic readout algorithm works almost the same as before with only a few minor adjustments as follows:

5.2.2.1 Modified ψ estimation

Instead of calculating the Q/I, the ψ_{estimate} is now calculated via

$$\hat{\psi}_{\text{estimate}} := \arctan\left(\frac{Q_{n,+} + Q_{n,-}}{I_{n,+} + I_{n,-}}\right) \tag{5.11}$$

5.2.2.2 Calculation of the 'positive' and 'negative' $c_{n,\pm}$

In analogy to the previous c_n, I define $c_{n,+}$ and $c_{n,-}$ coefficients via

$$c_{n,+} = \begin{cases} -(Q_{n,+} + Q_{n,-})\sin(n\psi) + (I_{n,+} + I_{n,-})\cos(n\psi) & \text{for } n \text{ even} \\ -(Q_{n,+} + Q_{n,-})\cos(n\psi) - (I_{n,+} + I_{n,-})\sin(n\psi) & \text{for } n \text{ odd} \end{cases} \tag{5.12}$$

$$c_{n,-} = \begin{cases} -(Q_{n,+} - Q_{n,-})\cos(n\psi) - (I_{n,+} - I_{n,-})\sin(n\psi) & \text{for } n \text{ even} \\ -(Q_{n,+} - Q_{n,-})\sin(n\psi) + (I_{n,+} + I_{n,-})\cos(n\psi) & \text{for } n \text{ odd} \end{cases} \tag{5.13}$$

$$\tag{5.14}$$

The resulting $c_{n,+}$ and $c_{n,-}$ coefficients, as written in Table 5.2, are now independent of the ψ. In the limit of $\delta\omega \mapsto 0$, the $c_{n,-} \mapsto 0$ and $c_{n,+} \mapsto 2c_n$.

	n even	n odd
$c_{n,+}$	$AJ_k(m)\left(\cos\varphi + \cos(\delta\omega T + \varphi)\text{sinc}(\delta\omega T)\right)$	$AJ_k(m)\left(\sin\varphi + \sin(\delta\omega T + \varphi)\text{sinc}(\delta\omega T)\right)$
$c_{n,-}$	$AJ_k(m)\left(\sin\varphi - \sin(\delta\omega T + \varphi)\text{sinc}(\delta\omega T)\right)$	$AJ_k(m)\left(\cos\varphi - \cos(\delta\omega T + \varphi)\text{sinc}(\delta\omega T)\right)$

Table 5.2: Table of additional coefficients derived from equations (5.7) to (5.10).

5.2.2.3 Calculating m_n

With the $J_n(m)$ factored out, we can use either the $c_{n,+}$ or the $c_{n,-}$ in place of the c_n in equation (4.21) as before to calculate the m_n and average them exactly as done in section 4.7.

For small speeds (small $\delta\omega$) one should rather use the $c_{n,+}$ as the $c_{n,-}$ become small and will have more noise than the $c_{n,+}$.

5.2.2.4 Modified φ estimation

For a high precision phase readout, we first define an additional helper coefficient $d_\pm$ via:

$$d_{\pm,n} := \frac{c_{n,\pm}}{J_n(m)}, \tag{5.15}$$

which yields for even and odd n:

$$d_{+,n_{\text{even}}} = A(\cos\varphi + \cos(\delta\omega T + \varphi)\text{sinc}(\delta\omega T))$$
$$d_{-,n_{\text{even}}} = A(\sin\varphi - \sin(\delta\omega T + \varphi)\text{sinc}(\delta\omega T))$$
$$d_{+,n_{\text{odd}}} = A(\sin\varphi + \sin(\delta\omega T + \varphi)\text{sinc}(\delta\omega T))$$
$$d_{-,n_{\text{odd}}} = A(\cos\varphi - \cos(\delta\omega T + \varphi)\text{sinc}(\delta\omega T)) \tag{5.16}$$

Next, we average these coefficients over all even and odd harmonics using our custom weights, leading to exactly 4 averaged coefficients $d_{+,\text{even}}, d_{+,\text{odd}}, d_{-,\text{even}}, d_{-,\text{odd}}$. From these 4 coefficients we calculate the phase via:

$$\varphi := \arctan\left(\frac{d_{+,\text{odd}} + d_{-,\text{even}}}{d_{+,\text{even}} + d_{-,\text{odd}}}\right) \tag{5.17}$$

The calculated interferometric phase φ now corresponds to the distance of the target at the beginning of the measurement (at time $t = 0$). When converting this phase into a distance; one has to keep in mind that due to the target's movement, the position will be different and proportional to $\varphi + T\delta\omega$. A precise value for the Doppler frequency $\delta\omega$ is therefore crucial to archive a high precision readout.

5.2.3   Precision in case of a linear frequency shift $\delta\omega$

Figure 5.3 shows the performance of the dynamic extension of the analytic readout algorithm for simulated DFM signals with different linear frequency shifts $\delta\omega$ and additive white noise.
In the simulation, that lead to Figure 5.3 I used the 'full' DFM signal of a linearly moving target given by:

$$s_{\text{DFM, moving}}(t) = B + A\cos\left(m(t)\cdot\sin\left(\omega_m t + \psi\right) + \varphi + \delta\omega\, t\right) \tag{5.18}$$

$$= B + A\sum_{n\in\mathbb{Z}} J_n(m + \delta m\cdot t)\cdot\cos\left((n\omega_m + \delta\omega)t + n\psi + \varphi\right) \tag{5.19}$$

with $\delta m = \Delta\omega\frac{v}{c} \neq 0$. The "error" $|\varphi_{\text{ex}} - \varphi_{\text{estimate}}|$ plotted is saturated around 2π (in the y-axis), since the phase results of the readout algorithm are wrapped to the interval $\varphi_{\text{estimate}} \in [-\pi, \pi]$. An error of $\sim 2\pi$ simply means that the algorithm is unable to calculate the proper phase value and the interferometric phase can no longer be used to calculate a correct distance. Since I also only consider a linear moving target here; the improvements from the dynamic readout extension are maximized in this scenario and will be less (compared to the static algorithm) for other, more non–linearly moving targets.

Results for the "static" readout:

While the performance of the static readout algorithm looks much worse than the dynamic in Figure 5.3, it is not quite as imprecise as it seems to be at first glance. The error of the phase readout of the static algorithm below 30 Hz is dominated by the additional phase from the Doppler shift $T\delta\omega$. The output of the static algorithm corresponds in this case to something of an average phase (displacement) between the beginning of the measurement $L(t = 0) \sim \varphi_{\text{ex}}$ and the end $L(t = T_{\text{ex}}) \sim \varphi_{\text{ex}} + T_{\text{ex}}\delta\omega$. The static readout algorithm will still yield a relatively low noise floor in its frequency spectrum as demonstrated earlier in Figure 4.12b; it will however contain an "offset" or "averaged" phase values since the phase does not remain constant throughout the measurement. In Figure 5.3 this is demonstrated with the dotted brown line which corresponds to the results of the static algorithm minus a Doppler frequency dependent offset $T_{\text{ex}}\delta\omega/2$ (and minus the exact φ_{ex} value to get the "error"). From ~ 10 Hz on, the static algorithm does however become erroneous beyond this offset due to the imprecise demodulation of the shifted harmonics as explained before.
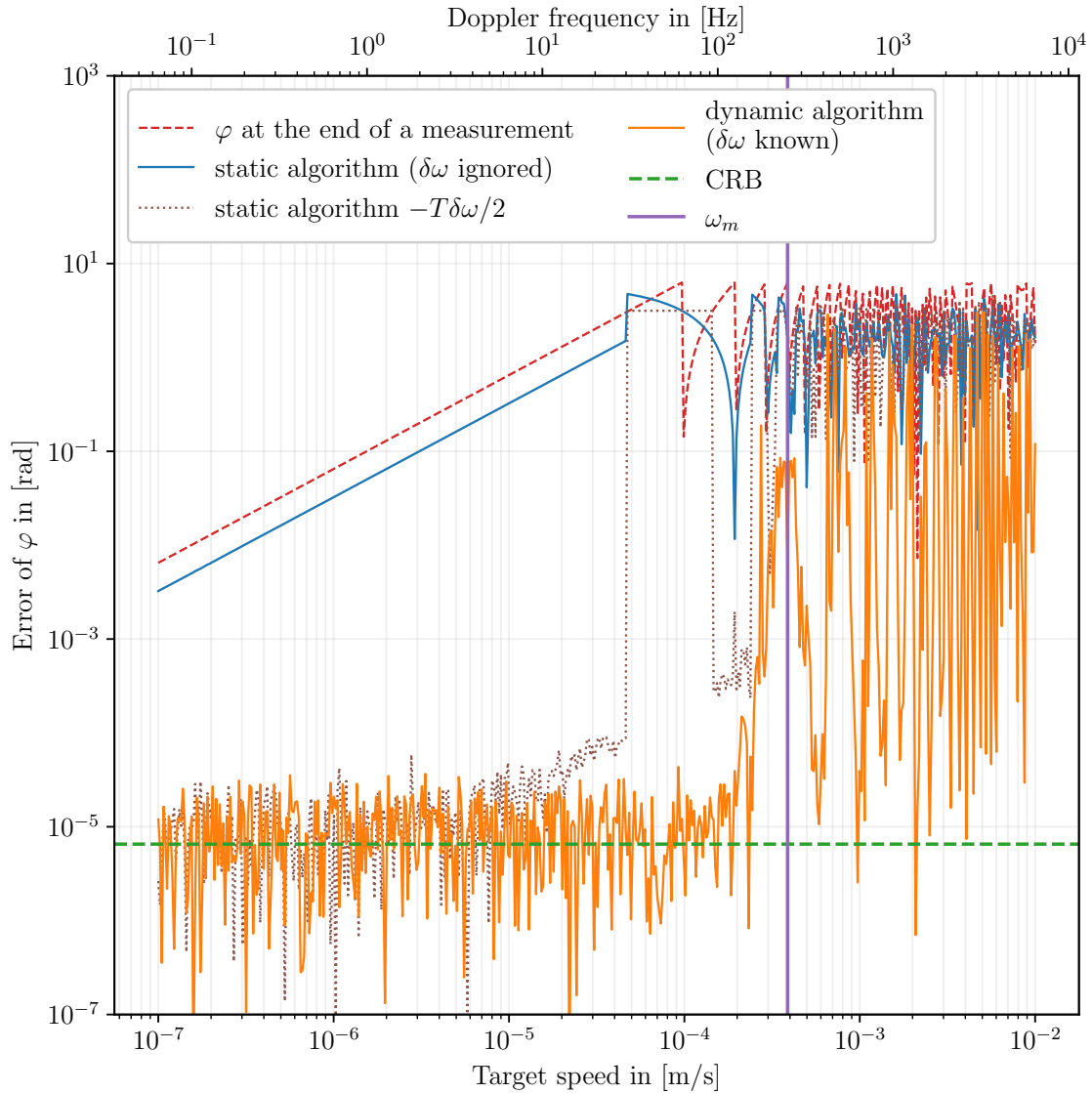
Figure 5.3: Error ($|\varphi_{\text{ex}} - \varphi_{\text{estimate}}|$) of the analytic readout algorithm for a simulated DFMI signal (s_{example} with parameters as given in Table 3.1, additive white Gaussian noise ($\sigma = 2 \cdot 10^{-7}$) and an initial of distance $L_0 = 20\,\text{cm}$) which changes of the course of the measurement with $L(t) = L_0 + \delta v\,t$, with δv as "target speed." The lower x–axis shows this target speed δv in m/s while the upper x–axis shows this speed converted into the Doppler frequency shift $\delta \omega = 2\pi/\lambda_0 \cdot \delta v$. For each point in the plot, I simulated a target moving with constant speed 100 times (but always starting at the same position L_0) and averaged the result. The blue line shows the readout algorithm result in case the Doppler shift is ignored, and the algorithm runs as described in chapter 8. The orange line shows the algorithms result in case the Doppler shift is known before and the algorithm can account for it as explained in this chapter. The dashed green line is the CRB of the phase showing the highest reachable precision (calculated in section 8.2.2). The dashed red line marks the additional phase from the Doppler shift $T\delta\omega$ and the vertical purple line marks a Doppler shift of $f_m/2$ (here at 250 Hz). At this point two neighboring harmonics (e.g. $n\omega_m + \delta\omega$ and $(n+1)\omega_m - \delta\omega$) overlap and the demodulation of a single isolated harmonic is not possible, leading to the algorithm yielding erroneous results at exactly multiples of $f_m/2$.

Results for the "dynamic" readout:

For Doppler shifts below $\approx 50\,\text{Hz}$ the target is moving so slow that the algorithm consistently reaches the lower theoretical limit (green dashed line, given by the CRB), showing that the dynamic readout extension shown here works and can be used to improve the readout precision. Comparison with the static algorithm and the explanation before, the greatest improvement from static to dynamic readout happens for targets speeds between $\delta v \approx 3 \cdot 10^{-5}\,\text{m/s}$ and $\approx 3 \cdot 10^{-4}\,\text{m/s}$. Below that, in many applications, the "averaged" static readout results will suffice for i.e., dampening the movement of the target with an actuator. In this specific speed range, the dynamic readout will yield however more precise displacement values useful for i.e., small but fast movements. The simulation also shows that within the reasonable target speeds assumed here, the error due to the δm term (that was present in the simulation) can be safely ignored as the other considered errors dominate the phase readout.

5.2.4   Example readout for an oscillating mass

As more realistic example of a measurement, Figure 5.4 shows the results of the readout algorithm for an oscillating target with the target's position given by: $L(t) = L_0 + 2\lambda_0 \cdot (1 - \cos(6\pi \cdot 4\,\text{Hz} \cdot t))$, which has a maximum speed of $\approx 233 \cdot 10^{-6}\,\text{m/s}$ at the slopes ($150\,\text{Hz}$ Doppler frequency). While Figure 5.3 shows ~ 3 orders of magnitude difference between the static and dynamic algorithm for such speeds, the difference in Figure 5.4 is only a factor of 10^1. The reason is that, unlike before, the target's speed varies over time and the movement contains not only linear but also higher order terms in t.
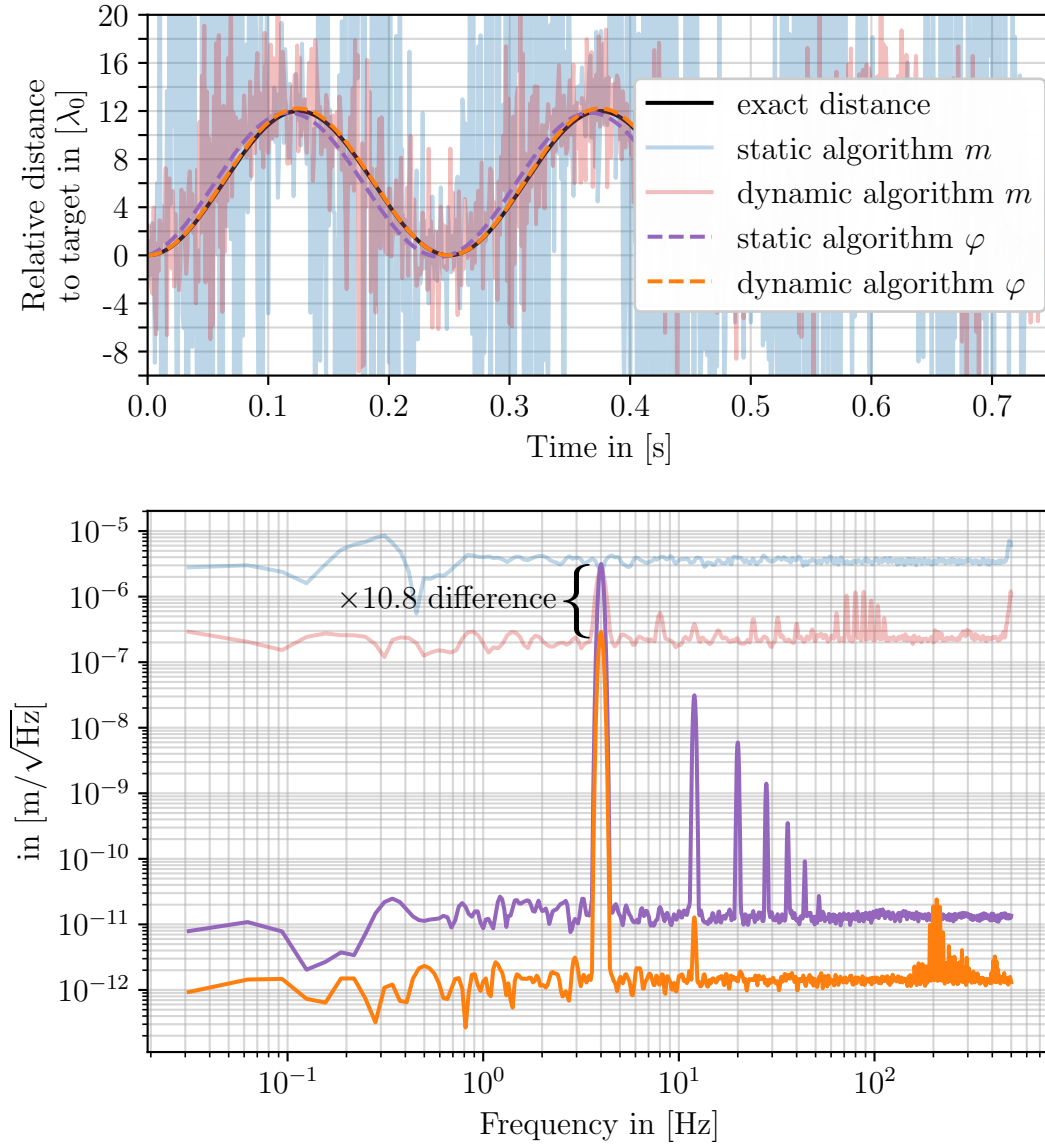
Figure 5.4: Algorithm results for a simulated moving target, oscillating with $L(t) = L_0 + 6\lambda_0(1 - \cos(2\pi \cdot 4\,\text{Hz} \cdot t))$, same as in Figure 4.12c before. The black line marks the exact phase of the target. The blue and purple lines are the (static) algorithm output (without the dynamic extension) and the red and orange line are the result when using the dynamic extension presented here. The upper plot shows three periods of the resulting time series and the lower plot shows the ASD of the error calculated with an LPSD algorithm [69].

Chapter 6

Resonantly enhanced DFMI

Another technique that derived from DFMI interferometry is what we call **Resonantly enhanced Deep–Frequency Modulation Interferometry**. Originally motivated by the studies of the Cramér-Rao bound of the readout of DFMI signals, early simulations have shown that combining cavities with a DFMI laser signal can lead to a more precise length readout compared to the regular DFMI phase readout. My studies of the Cramér-Rao bound for DFMI signals is discussed in the next part of this thesis and for ReDFMI specifically in section 8.5. This Chapter gives an overview of the technique and a series description of the expected signal. The ReDFMI technique itself is currently being investigated at the University of Hamburg and a first paper presenting the technique is being written.

6.1   Definition of ReDFMI

Resonantly **e**nhanced **D**eep-**F**requency **M**odulation **I**nterferometry (**ReDFMI**) is an interferometry technique where a sinusoidal modulated laser beam is sent into a cavity and the reflection and/or the transmitted beam from the cavity is measured to calculate the cavity length. Figure 6.1 shows an example ReDFMI setup.
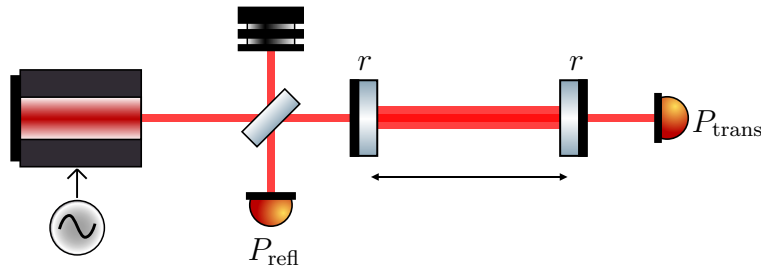

Figure 6.1: Sketch of a ReDFMI setup.

The response to a cavity for a plane wave is well understood and can be found in standard textbooks for optics as e.g. [70, Chapter 4.2]. To calculate the cavity response for the plane wave, all of the reflections inside the cavity are being added up,

resulting in an (infinite) converging geometric series yielding the closed expressions:

$$P_{\text{trans}} = P_{\text{in}} \left(\frac{(1-R)^2}{(1-R)^2 + 4R\sin^2(\varphi)} \right) = P_{\text{in}} \left(\frac{1}{1 + F\sin^2\left(\omega_0 \frac{2L}{c}\right)} \right) \quad (6.1)$$

$$\text{and} \quad P_{\text{refl}} = P_{\text{in}} - P_{\text{trans}} = P_{\text{in}} \left(1 - \frac{1}{1 + F\sin^2\left(\omega_0 \frac{2L}{c}\right)} \right) \quad (6.2)$$

for the reflected P_{refl} and transmitted P_{trans} signals, r as reflective coefficient identical for both mirrors and $F = 4r^2/(1-r^2)^2$ as Finesse coefficient. For a DFMI signal, the starting point of the calculation is similar, deriving a closed expression is however more complicated.

6.2 The Fourier Series of a ReDFMI signal

To simplify the calculation, we write the electric field of the DFM signal as (complex) phasor given by

$$E_{\text{DFM}}(t, \tau) = E_0 \; e^{i(\omega_0 \tau + \Delta\omega\tau \cdot \sin(\omega_m t + \psi))} . \quad (6.3)$$

In this case we use t as time of emission and τ as (fixed) propagation time. This corresponds to the photon field being measured at a fixed point in time (and space). When writing EM-waves, the t is usually used as propagation time variable, here this would be τ. Next we consider a simple Fabry–Pérot cavity as shown in Figure 6.2 with identical reflective coefficients of r for both mirrors (leading to reflectivity $R := r^2$ and transmissivity of $(1-R)$).

The ReDFMI signal transmitted out of the cavity would then be the sum over all reflections that exit the cavity at the same time given by the time of emission plus the propagation time $t+\tau$. The first beam emitted at time t, that would go "straight through" the cavity, would travel some time until it reaches the first mirror, travel through the cavity and exit at time $t+\tau_0$. The beam that is reflected twice (once at each mirror) before being transmitted would have been emitted at an earlier time $t - \tau_L$ with $\tau_L = 2L/c_0$ so that after a propagation time of $\tau_0 + \tau_L$ it would exit the cavity to the right at time $(t - \tau_L) + (\tau_0 + \tau_L) = t + \tau_0$, and so on and so forth.
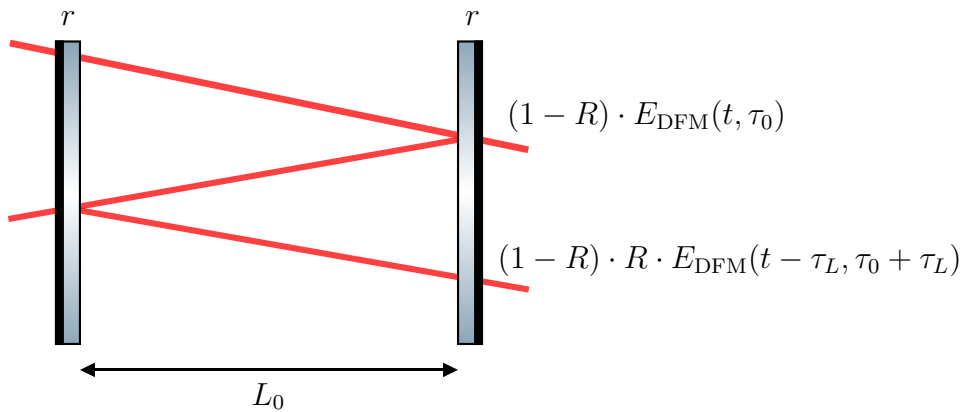

Figure 6.2: Sketch of a (Fabry–Pérot) cavity where both mirrors have the same reflective coefficient r.

As series this can be written down as

$$E_{\text{ReDFM, trans}}(t, \tau) = (1 - R) \sum_{k=0}^{\infty} R^k \, E_{\text{DFM}}(t - k\tau_L, \tau_0 + k\tau_L) \tag{6.4}$$

$$\overset{(6.3)}{=} (1 - R)E_0 \sum_{k=0}^{\infty} R^k e^{i(\omega_0\tau + k\omega_0\tau_L + \Delta\omega(\tau_0 + k\tau_L)\cdot\sin(\omega_m t - k\omega_m\tau_L + \psi_0))}$$

$$\tag{6.5}$$

$$=: (1 - R)E_0 \sum_{k=0}^{\infty} R^k e^{i(\varphi_0 + k\varphi_L + (m_0 + km_L)\cdot\sin(\omega_m t - k\psi_L + \psi_0))} \tag{6.6}$$

From (6.5) to (6.6), I introduced a set of composite parameters (also written in Table 6.1) to shorten the expression.

$$\omega_0 \cdot \tau_0 =: \varphi_0 \qquad \omega_0 \cdot \tau_L =: \varphi_L \qquad \omega_m \cdot \tau_0 =: \psi_L$$
$$\Delta\omega \cdot \tau_0 =: m_0 \qquad \Delta\omega \cdot \tau_L =: m_L$$

Table 6.1: Table of ReDFMI composite parameters

After more extensive calculations using multiple summation formulas for the Bessel functions and applying the residual theorem of complex analysis, shown in detail in appendix B, this expression can be rewritten as complex Fourier series of "plane wave" exponentials as:

$$E_{\text{ReDFM, trans}}(t, \tau_0) = \sum_{n \in \mathbb{Z}} C_n \cdot e^{in\omega_m t} \tag{6.7}$$

$$\text{with} \quad C_n = (1 - R)E_0 e^{i(\varphi_0 + n\psi_0)} \, \frac{2i}{m_L} \frac{\cosh\left(nx - m_0 \sinh(x)\right)}{\cosh x} \tag{6.8}$$

$$\text{and} \quad x := \text{arcsinh}\left(\frac{\ln R + i(\varphi_L - n\psi_L)}{m_L}\right) \tag{6.9}$$

This complex field description is not yet equivalent to the (real) measured current on a photodiode which corresponds to the energy flux / Poynting vector calculated from this electric field. The series written above can however be helpful for further calculations and research in this topic.

Part III

Noise analysis of a DFMI signal

One of the questions Oliver Gerberding and I discussed when working on DFM-Interferometry was: How precise can DFMI actually be? Isleif and Gerberding have conducted multiple DFMI tabletop experiments reaching sub-picometer precision [61] similar to other tabletop interferometry experiments. DFMI signals differ, however, in some aspects from other interferometry signals. I.e., homodyne interferometers operate near the dark fringe to mitigate the effects of shot noise. A DFMI signal varies between the light and dark fringe and would have a potentially higher shot noise. Heterodyne interferometry, on the other hand, is used with carrier frequencies up to the MHz range, which is much higher than the typical DFMI frequency range of a few kHz. Since many noise sources have a frequency dependence, they could affect DFMI signals differently than other heterodyne signals. As a DFMI signal is also spread over multiple frequencies, calculating an SNR is also not as straight forward when trying to assess the quality of a measured signal.

To answer the initial questions of "How precise can DFMI actually be?," I found a solution in the Cramér-Rao bound of statistics.

For a given noise model, it allows one to calculate how precise a parameter estimate would be. While terms like the signal-to-noise ratio (SNR) are commonly used in physics or engineering as indicator of how good a measured signal is, with a higher SNR correlating to a higher precision of the readout, the Cramér-Rao bound (and its inverse, the "Fisher Information") are the mathematical rigorous quantification of how precise a parameter estimates / the readout will be. For some common cases as e.g. in a heterodyne interferometer, the signal-to-noise ratio of the amplitude spectral density of a measured signal is exactly its Fisher Information.

The Cramér-Rao bound uses the probability distribution function of a random signal to calculate the precision limits for a given parameter. Outside quantum physics where i.e. the quantum wave function / a quantum field serves as probability distribution, this formalism is often not directly applicable. In the field of GWD physics for example, signals are more commonly described by their amplitude spectral density or power spectral density. To keep this thesis as much self–contained as possible, I will recapitulate these expressions briefly in the following part and show explicitly how a probability distribution relates i.e. to the amplitude spectral density of a measured signal. The next chapter will therefore mostly recapitulate many commonly used statistical expressions with their application on a DFMI signal being the novel case.

Chapter 7

Revision and clarifications on used terms from signal analysis

This chapter is mainly a reference for some common equations used in this thesis. I assume that most readers know the content presented here already sufficiently well (i.e. what a PSD is). There are, however, some subtleties that students new to this field may not know or that can be helpful to briefly refresh before delving into the more theoretical parts of this thesis.

When analyzing and discussing signals in the gravitational wave detector community, it is usually done by expressing the measurement as either "power spectral density" or "amplitude spectral density." While being defined by the Fourier transform of the autocorrelation function of stochastic signals, these expressions are very closely related to the Fourier transform[1] of the measured signal itself, and I want to highlight these similarities here once.

7.1   Used Conventions for the Fourier transform

7.1.1 Definition : Fourier transform $\mathcal{F}$

The **Fourier transform** $\tilde{f}(\omega)$ is a bijective integral transformation between two (complex valued) function spaces. For a given function $f(t)$ the Fourier transform (and its inverse) of the real parameters t to ω is defined by:

$$\mathcal{F}_{t \mapsto \omega}\{f\} := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(t) \cdot e^{-i\omega t} \, dt = \tilde{f}(\omega) \tag{7.1}$$

$$\mathcal{F}_{t \mapsto \omega}^{-1}\{\tilde{f}\} := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \tilde{f}(\omega) \cdot e^{i\omega t} \, dt = f(t) \,. \tag{7.2}$$

In this thesis, a ~ above a function refers to its Fourier transform as defined here. The image of $\mathcal{F}$ (or the domain of $\tilde{f}$) is sometimes referred to as "Fourier domain" with ω as frequency. While the signals we measure are usually real-valued functions, their Fourier transform is complex ($\sim \mathbb{C}$) and generally has non–trivial real and complex parts. In my case, where signals are measured over time t, the domain of $f(t)$ is also referred to as "time domain." Throughout my work I use the "unitary" Fourier transforms as defined above with its pre-factors of $1/\sqrt{2\pi}$. This way, the

---

[1]Or other closely related integral transformations like the Laplace Transform

Fourier transform, and its inverse are simply related by complex conjugation and don't differ by any other factors. In most cases I refer to ω as angular frequency, when writing or plotting discrete values I use however the ordinary frequency (like i.e. "$f_m = 500\,\text{Hz}$") related via $\omega = 2\pi \cdot f$. Any frequency variables in this thesis that appear as ω and f but share the same subscript refer to the same frequency and are only represented differently.

The Fourier transform plays a special role in mathematics for solving differential equations as it translates a large class of differential equations onto algebraic equations of rational functions. For systems described by a linear differential operator, like the equations of motion for a suspended mass in a GWD, the corresponding algebraic rational function is closely related to the so called "transfer function" of the system which I explain further in section 7.3. In engineering, there is often another integral transformation used to describe systems and simplify calculations: the Laplace transform.

7.1.2 Definition : Laplace transform $\mathcal{L}$

The **Laplace transform** $\mathcal{L}\{f\}(s)$ is a bijective integral transformation between two (complex valued) function spaces. For a given function $f(t)$ the Laplace transform (and its inverse) of the parameter t to s is defined by:

$$\mathcal{L}_{t \mapsto s}(f) := \int_0^\infty f(t) \cdot e^{-st}\, dt = F(s) \qquad \text{and} \tag{7.3}$$

$$\mathcal{L}_{t \mapsto s}^{-1}(F) := \frac{1}{\sqrt{2\pi i}} \lim_{T \to \infty} \int_{\gamma - iT}^{\gamma + iT} F(s) \cdot e^{st}\, ds = f(t)\,. \tag{7.4}$$

While the Fourier transform mostly considers the transformed variable ω real valued; the domain of a Laplace transformed function is complex valued ($s \in \mathbb{C}$).

The Laplace transformed of a function is generally different from its Fourier transformed. Both integral transformations are however closely related, and the Laplace transform of a function can be expressed as a Fourier transform of:

$$\mathcal{L}_{t \mapsto s}(f) = \underbrace{\int_0^\infty dt}_{= \int_\mathbb{R} dt\ \Theta(t)}\ f(t) \cdot e^{-st} =: F(s) \tag{7.5}$$

$$= \sqrt{2\pi}\ \mathcal{F}_{t \mapsto \omega}\Big\{ f(t) \cdot \Theta(t) \Big\}(\omega = -is) \tag{7.6}$$

with $\Theta(t)$ as Heaviside / unit-step function. Specifically for functions that are zero for $t < 0$, meaning when $f(t) \cdot \Theta(t) = f(t)$, Fourier transform and Laplace transform are identical up to a substitution of the Laplace variable s to the Fourier (frequency) variable $s \mapsto i\omega$. In experimental setups we often consider exactly such signals where a system is "at rest" before some event, and after switching some button, or when some external transient signal (like a passing seismic wave) reaches the setup, a measurement starts. For such signals, Fourier transform and Laplace transform are almost identical and while engineers will often describe their signals by its Laplace transform (with the Laplace variable s), I will use in almost all cases the Fourier transform to describe systems, their function and measured signals.

7.1.3 Definition : Convolution Operator $*$

The **convolution** between two functions $f(t)$ and $g(t)$ is defined by

$$f(t) * g(t) := \int_{\mathbb{R}} f(\tau) \cdot g(t - \tau) d\tau = (f * g)(t) \tag{7.7}$$

Under both Fourier $\mathcal{F}$ and Laplace $\mathcal{L}$ transformation, the convolution is mapped onto a simple multiplication of the transformed of the two functions. I.e.

$$\mathcal{F}_{t \to \omega} \{f * g\} = \tilde{f}(\omega) \cdot \tilde{g}(\omega) \qquad \text{and} \qquad \mathcal{L}_{t \to s} \{f * g\} = \mathcal{L}\{f\}(s) \cdot \mathcal{L}\{g\}(s) \tag{7.8}$$

$$\text{and} \qquad \mathcal{F}_{t \to \omega}^{-1} \left\{ \tilde{f} \cdot \tilde{g} \right\} = f(t) * g(t) = \mathcal{L}_{t \to s}^{-1} \{\mathcal{L}\{f\}(s) \cdot \mathcal{L}\{g\}(s)\} \tag{7.9}$$

And similarly, the n'th derivative of some function $f(t)$ mapped to Fourier or Laplace domain becomes:

$$\mathcal{F}_{t \to \omega} \left\{ f^{(n)}(t) \right\} = (i\omega)^n \tilde{f}(\omega) \tag{7.10}$$

$$\text{and} \qquad \mathcal{L}_{t \to s} \left\{ f^{(n)}(t) \right\} = s^n F(s) - \sum_{k=1}^{n} s^{n-l} \cdot \lim_{t_0 \to 0^-} f^{(k-1)}(t_0) \tag{7.11}$$

Here, $\lim_{t_0 \to 0^-}$ is the limit of t_0 going to 0 from the negative t plane. For signals that are $\equiv 0$ for $t < 0$ as mentioned above, all the coefficients of the sum in (7.11) become zero and the Laplace and Fourier transformed become identical up to a variable substitution.

7.2 The autocorrelation function and the "PSD"

In physics and engineering, measured signals are often expressed as power spectral density (PSD) or its square root, the amplitude spectral density (ASD). While the PSD is strictly defined as the Fourier transform of the autocorrelation function of a signal, for the practical calculation, it can often simply be regarded as an approximation of the absolute value of a signals Fourier transform. There are, however, some differences in the actual calculation of the PSD (or the autocorrelation function) for different types of signals, which are not always explained. Especially when comparing an analytic expression of an idealized signal, like the Cramér-Rao bound that I present later, to a measured or simulated noisy sample, it is important to understand which equation for the power spectral density (PSD) needs to be used, and how they relate to one another. For this reason I briefly explain the calculations leading to a proper PSD in a bit more detail and show how the different expressions relate to one another. For the names and conventions presented here I mostly follow the textbook for Oppenheim and Verghese [71] who discuss this topic in bit more detail.

7.2.1 Definitions for deterministic signals

For a square-integrable function[2] $(x \in L^2)$, with $x : \mathbb{R} \to \mathbb{C}$, the **deterministic autocorrelation** (marked with a bar $^-$) is defined as

$$\bar{R}_{xx}(\tau) := x(\tau) * x^*(-\tau) = \int_{\mathbb{R}} x(t) \cdot x^*(t - \tau) \, dt \tag{7.12}$$

---

[2] A function being called square–integrable is exactly when $\int_{\mathbb{R}} |x(t)|^2 dt < \infty$.

with $*$ as convolution operator. If the deterministic autocorrelation exists, then the **energy spectral density (ESD)** is defined as its Fourier transform:

$$\bar{S}_{xx}(\omega) := \mathcal{F}_{\tau \to \omega}\left\{\bar{R}_{xx}(\tau)\right\} = |\tilde{x}(\omega)|^2. \tag{7.13}$$

This (deterministic) energy spectral density is in fact just the squared absolute value of the signals Fourier transform. For $\tau = 0$, the right side integral of (7.12) also defines the so called "energy" of the signal. E.g.

$$E := \underbrace{\int_{\mathbb{R}} |x(t)|^2 \, dt}_{=\bar{R}_{xx}(0)} \overset{\overset{\text{Plancherel}}{\text{theorem}}}{=} \underbrace{\int_{\mathbb{R}} |\tilde{x}(\omega)|^2 \, d\omega}_{=\int_{\mathbb{R}} \bar{S}_{xx}(\omega) d\omega} . \tag{7.14}$$

For many practical applications these definitions are however not very useful as common analytic (not random) signals in physics and engineering, like a sine $\sin(\omega_0 t)$, are not square-integrable and the deterministic autocorrelation and the energy spectral density do not exist[3].

For a random signal x, the signal energy also relates to the mean μ and variance σ^2 of the random variable via

$$\sigma^2 := \text{Var}(x) = \mathbb{E}[(x - \mu)^2] = \underbrace{\mathbb{E}[x^2]}_{=R_{xx}(\tau=0)=E} - \underbrace{(\mathbb{E}[x])^2}_{=\mu^2} \tag{7.15}$$

$$\iff E = \sigma^2 + \mu^2 = \int_{\mathbb{R}} \bar{S}_{xx}(\omega) d\omega \tag{7.16}$$

(with $\mathbb{E}$ as expected value) which is a useful relation I will refer to later on.
Specifically when discussing the random motion of a suspended mass in a GWD, its movement is often characterized by its "RMS motion" which corresponds to the square root of the variance σ or signal energy $\sqrt{E}$ (when $\mu = 0$) of its position signal.

7.2.2   Strict definition of the autocorrelation and the PSD

The mathematically rigorous definition of the **autocorrelation** of a **random process** $x(t)$ (random variable x depending on some time parameter t) is given by

$$R_{xx}(t_1, t_2) := \mathbb{E}[x(t_1) x^*(t_2)] \tag{7.17}$$

with x^* as the complex-conjugate of x and $\mathbb{E}$ as expected value.
The expected value is calculated (for any function f depending on the random variable x) by the integral or sum of the probability density function ρ over the space Ω of all possible outcomes of x:

$$\mathbb{E}[f(x)] := \int_{\Omega} \rho(x, t) \cdot f(x) \, dx . \tag{7.18}$$

---

[3]In Fourier domain, many of these signals don't exist as regular functions, but can be expressed as distributions. In the space of distributions, multiplication "$\cdot$" is however not a proper operation and the product $|\tilde{x}(\omega)|^2 = \tilde{x}(\omega) \cdot \tilde{x}^*(\omega)$ may not exist.

Additionally, the signals (or rather the noises) are generally assumed to be **weak–sense stationary (WSS)**. For a weak–sense stationary process, (a) the mean value does not change over time ($\mathbb{E}[x(t)] \equiv$ const), (b) the variance is finite for all times and (c) the autocorrelation only depends on the time difference $\tau := t_1 - t_2$. I.e. for an arbitrary $t := t_1$ value, the autocorrelation can be written as

$$R_{xx}(t_1, t_2) = R_{xx}(t_1 - t_2, 0) = R_{xx}(\tau) = \mathbb{E}[x(t)x^*(t - \tau)]. \qquad (7.19)$$

Now, for such WSS processes, the **power spectral density (PSD)** is defined as the Fourier transform of the autocorrelation function with respect to the time delay τ, e.g.

$$S_{xx}(\omega) := \mathcal{F}_{\tau \to \omega}\left\{R_{xx}(\tau)\right\} \qquad (7.20)$$

It is important to note here that our measured signals themselves are generally not (weak–sense) stationary processes. In many cases, we can describe them however by an analytic (and maybe non–stationary) part and some random noise which often is weak–sense stationary or even fully stationary (meaning that it has no explicit dependence on time t).
Instead of calculating the PSD of a measured signal by the integral (7.18) over some probability space, it is usually calculated differently as an integral or sum some over some time interval. The measured signal is then (even if often not explicitly stated) considered to be "ergodic."

7.2.3   Autocorrelation for "ergodic" signals

To simplify stochastic calculations, signals (random processes) are often considered to be "ergodic." Oppenheim and Vergese state that "a process is simply termed ergodic if ensemble statistics can be replaced by temporal statistics on (almost) every particular realization" [71, p. 391]. I.e. it is implicitly assumed that the random variable x takes on every random realization from Ω in either a finite or infinite time interval $t \in [0, T]$ or $t \in \mathbb{R}$, and that the probability is uniformly distributed over this time (such that $\rho(x(t)) \equiv 1/T$). (A little more rigorous definition of signals being "ergodic in the wide sense" is also given in [71] but skipped here).

7.2.3.1   Autocorrelation for ergodic signals with finite signal energy

In its most general form, an ergodic signal is assumed to take on every realization of some random variable exactly once over infinite time $t \in \mathbb{R}$. The expected value of some function depending on x would then be written as

$$\mathbb{E}[f(x)] = \int_{\Omega} \rho(x, t) \cdot f(x)\, dx \approx \lim_{T \to \infty} \int_{-T}^{T} \frac{1}{2T} \cdot f(x(t))\, dt\,, \qquad (7.21)$$

and for a WSS process one finds the better known formula for the autocorrelation of:

$$R_{xx}(\tau) = \int_{\Omega} \rho(x, t) \cdot x(t)x^*(t - \tau)\, dx \approx \lim_{T \to \infty} \int_{-T}^{T} \frac{1}{2T} \cdot x(t)x^*(t - \tau)\, dt\,. \qquad (7.22)$$

If $x(t)$ is also a square-integrable function, the PSD for such an ergodic "over infinite time" signal is identical to the energy spectral density. I.e.

$$S_{xx}(\omega) = \mathcal{F}_{\tau \to \omega} \{R_{xx}(\tau)\} \tag{7.23}$$

$$\overset{(7.22)}{=} \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} x(t) \cdot \mathcal{F}_{\tau \to \omega} \{x^*(t - \tau)\} \, dt \tag{7.24}$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{\mathbb{R}} \mathrm{rect}\left(\tfrac{t}{2T}\right) \cdot x(t) \cdot e^{-i\omega t} \cdot \tilde{x}^*(\omega) \, dt \tag{7.25}$$

$$= \lim_{T \to \infty} \frac{1}{2T} \tilde{x}^*(\omega) \cdot \sqrt{2\pi} \, \mathcal{F}_{t \to \omega}\left\{x(t) \cdot \left(\lim_{T \to \infty} \tfrac{1}{2T}\mathrm{rect}\left(\tfrac{t}{2T}\right)\right)\right\} \tag{7.26}$$

$$\overset{(7.9)}{=} \tilde{x}^*(\omega) \cdot \left(\underbrace{\tilde{x}(\omega) * \left(\lim_{T \to \infty} \mathrm{sinc}\left(\tfrac{T\omega}{\pi}\right)\right)}_{\sim \tilde{x}(\omega) * \delta(\omega)} \right) \tag{7.27}$$

$$= |\tilde{x}(\omega)|^2 \,. \tag{7.28}$$

In other words, for finite signal energy, the PSD converges toward the ESD as $T \mapsto \infty$.

7.2.3.2   Autocorrelation for periodic signals (ergodic within a finite interval)

For periodic (and non–square integrable) signals with period T, the more restrictive assumption is usually made that the random signal takes on every realization within one (finite) signal period. In this case the autocorrelation becomes:

$$R_{xx}(\tau) = \int_{\Omega} \rho(x, t) \cdot x(t)x^*(t - \tau) \, dx \approx \int_{t_0}^{t_0 + T} \frac{1}{T} \cdot x(t)x^*(t - \tau) \, dt \tag{7.29}$$

and the PSD can in this case be expressed as

$$S_{xx}(\omega) = \tilde{x}^*(\omega) \cdot \left(\tilde{x}(\omega) * \tfrac{1}{\sqrt{2\pi}}\mathrm{sinc}\left(\tfrac{T\omega}{2\pi}\right) e^{-i\frac{T}{2}\omega} \right) \,. \tag{7.30}$$

Writing a periodic function $x(t)$ with period T as complex Fourier series, these expressions can also be written as:

$$x(t) = \sum_{n \in \mathbb{Z}} c_n \cdot e^{in\frac{2\pi}{T}t} \implies R_{xx}(\tau) = \sum_{n \in \mathbb{Z}} |c_n|^2 \cdot e^{in\frac{2\pi}{T}\tau} \,. \tag{7.31}$$

$$\text{and} \quad S_{xx}(\tau) = \sum_{n \in \mathbb{Z}} |c_n|^2 \cdot \sqrt{2\pi}\delta(\omega + n\tfrac{2\pi}{T}) \,. \tag{7.32}$$

The often not mentioned difference between the autocorrelation (and PSD) as written in (7.22) and (7.31) is that one signal is considered ergodic over the entire $t \in \mathbb{R}$ while the other is ergodic over a smaller interval $t \in [0, T]$.

7.3   Linear systems, the fundamental solution, greens functions and the transfer function

Another often unclear topic that one encounters when working in the intersection between physics, mathematics and engineering, is that researchers sometimes talk

about the same, or at least very similar objects, while using very different words and terminology. The prime example for this is the "transfer function" of a system. While it is defined straightforward as the ratio of the Laplace transformed output divided by the Laplace transformed input of a "system," there are very similar objects like the "fundamental solution" in mathematics and the "greens function" which are more commonly used in physics. While these three words refer, strictly, to different objects; they are very closely related to one another.

7.3.1 Linear time-invariant systems (LTI)

In mathematics and physics we often describe physical systems by linear differential equations. Written as operator $\hat{L}$, one would write i.e.

$$\hat{L}\, x(t) = 0 \qquad \text{homogeneous linear differential equation} \tag{7.33}$$

$$\hat{L}\, x(t) = f(t) \qquad \text{inhomogeneous linear differential equation} \tag{7.34}$$

with $\hat{L}$ linear meaning that $\hat{L}\,\{x_1(t) + x_2(t)\} = \hat{L}\,\{x_1\} + \hat{L}\,\{x_2\}$ for two solutions x_1, x_2 and $\hat{L}\,\{a \cdot x_1(t)\} = a \cdot \hat{L}\,\{x_1\}$ for a scalar a not dependent on t.

In engineering, slightly different terminology is used for similar objects.
A system described by a homogeneous differential operator $\hat{L}$, with no explicit time dependency, is here called a **linear time-invariant (LTI) system**. The signal that it acts upon is usually referred to as the **input signal** $x_{\text{in}}(t)$, and the inhomogeneous part of the differential equation is referred to as the **output signal** $\hat{L}\, x_{\text{in}}(t) =: x_{\text{out}}(t)$.

When discussing (linear) differential equations in mathematics, a common tool for solving these equations, that is related to the LTI systems transfer function, is the fundamental solution.

7.3.2 The fundamental solution

7.3.2.1 Definition: Fundamental solution

For a given linear differential operator $\hat{L}$ the **fundamental solution** x_{FS} is defined by

$$\hat{L}\, x_{FS}(t) = \delta(t) \qquad \text{with } t \in \mathbb{R} \tag{7.35}$$

with $\delta(t)$ as Dirac delta distribution.

Because linear differential operators commutate with the convolution operation $*$, e.g.

$$\hat{L}\,(f * g) = (\hat{L}f) * g = f * (\hat{L}g) \tag{7.36}$$

the solution $x(t)$ to any inhomogeneous linear differential equation $\hat{L}\, x(t) = f(t)$ can be calculated from the fundamental solution via:

$$\hat{L}\, x(t) = f(t) \qquad \text{and} \qquad \hat{L}\, x_{FS}(t) * f(t) = \delta(t) * f(t) = f(t) \tag{7.37}$$

$$\implies \quad x(t) = x_{FS}(t) * f(t) = \int_{\mathbb{R}} x_{FS}(t') \cdot f(t - t')\, dt'. \tag{7.38}$$

As a brief example, let us consider the fundamental solution for the harmonic oscillator:

7.3.2.2   Fundamental solution of the harmonic Oscillator

I.e. starting the differential operator describing the harmonic oscillator

$$\hat{L}_{\text{harmonic}} = \frac{\partial^2}{\partial t^2} + \omega_0^2 \,, \tag{7.39}$$

inserting the fundamental solution as defined in (7.35) and using the Fourier transform $\mathcal{F}$ to simplify the equation yields:

$$\hat{L}_{\text{harmonic}} x_{\text{FS}} = \delta(t) \quad \overset{\mathcal{F}}{\Longleftrightarrow} \quad \mathcal{F}\left\{\hat{L}_{\text{harmonic}} x_{\text{FS}}\right\} = \frac{1}{\sqrt{2\pi}} \tag{7.40}$$

$$\Longleftrightarrow (-\omega^2 + \omega_0^2)\tilde{x}_{\text{FS}} = \frac{1}{\sqrt{2\pi}} \tag{7.41}$$

$$\Longleftrightarrow \tilde{x}_{\text{FS}}(\omega) = -\frac{1}{\sqrt{2\pi}} \frac{1}{(\omega^2 - \omega_0^2)} \tag{7.42}$$

$$= \frac{1}{2\omega_0\sqrt{2\pi}} \left(\frac{1}{\omega + \omega_0} - \frac{1}{\omega - \omega_0}\right) \tag{7.43}$$

transforming back into time domain now yields the fundamental solution as

$$\overset{\mathcal{F}^{-1}}{\Longrightarrow} \quad x_{\text{FS,harmonic}}(t) = \frac{1}{2\omega_0} \sin(\omega_0 t)(2\Theta(t) - 1) \tag{7.44}$$

with $\Theta(t)$ as Heaviside / unit step function.

Fundamental solutions as tools to solve inhomogeneous differential equations are mostly encountered in mathematics, whereas in physics, the closely related "Green's functions" are often referred to instead.

7.3.3   Green's functions

Green's functions (here denoted by G) are similarly used to find a solution to an inhomogeneous linear differential operator $\hat{L}$. They are very closely related to fundamental solutions as they are generally defined by

$$\hat{L}G(x, x') = \delta(x - x') \qquad \text{for } x \in D \subseteq \mathbb{R}^n \tag{7.45}$$

$$\text{with } G(x, x') = [\text{some condition}] \qquad \text{for } x \in \partial D \,. \tag{7.46}$$

I.e a Green's function can be understood as the fundamental solutions of a specific linear differential operator within a finite domain D that fulfill a boundary condition on ∂D. They are used, identical to the fundamental solution, to calculate the solution of some inhomogeneous differential equation by convolution with the inhomogeneous part. Sometimes their name is also used interchangeably to express

a fundamental solution. Whereas fundamental solutions are usually defined for the entire domain of i.e. $\sim \mathbb{R}^n$ (or $\mathbb{C}$), Green's functions are only required to be valid within a (smaller) region and to satisfy some boundary condition [72].

Commonly used Green's function are for example the solution to the Laplacian $\Delta = \partial^2/\partial x^2 + \partial^2/\partial y^2 + \partial^2/\partial z^2$ with the Dirichlet condition (fixed boundary condition), which is solved by the Green's function $G(\underline{x} - \underline{x}') = -1/4\pi|\underline{x} - \underline{x}'|$, which was mentioned in the introductions as tool calculate solutions to the linearized field equations of gravity.

7.3.4   The transfer function

In engineering, a linear time-invariant system (LTI) is instead described by its "transfer function."

7.3.4.1   Definition: Transfer function

For a given LTI with input $x_{\text{in}}(t)$ and output $x_{\text{out}}(t)$ such that

$$\hat{L}_{\text{LTI}}\, x_{\text{in}} = x_{\text{out}}\,, \tag{7.47}$$

the transfer function $H(s)$ is defined by:

$$H(s) := \frac{\mathcal{L}\left\{x_{\text{out}}\right\}}{\mathcal{L}\left\{x_{\text{in}}\right\}}\,. \tag{7.48}$$

Using that $\delta(t) * x_{\text{in}}(t) = x_{\text{in}}(t)$, the Laplace transformed of (7.47) can also be written as

$$\mathcal{L}\left\{x_{\text{out}}\right\} = \mathcal{L}\left\{\hat{L}_{\text{LTI}}\, x_{\text{in}}\right\} = \mathcal{L}\left\{\hat{L}_{\text{LTI}}\,(\delta * x_{\text{in}})\right\} = \mathcal{L}\left\{\hat{L}\delta\right\} \cdot \mathcal{L}\left\{x_{\text{in}}\right\} \tag{7.49}$$

$$\iff H(s) = \mathcal{L}\left\{\hat{L}\delta\right\}\,. \tag{7.50}$$

Experimentally, (7.50) is used to measure the transfer function of a system by providing an approximate delta peak (a short, high pulse) as input and simply measuring the output. The Laplace transformed of this output (and using that $\mathcal{L}\left\{\delta\right\} = 1$) yields the systems transfer function.

Comparing this expression to the fundamental solution (or Green's function) above shows:

$$\underset{\text{def. of fundamental solution}}{\overset{}{\Longrightarrow}}\ \underbrace{\hat{L}\, x_{\text{FS}}}_{=\hat{L}\delta * x_{\text{FS}}} = \delta(t) \quad \overset{\mathcal{L}}{\mapsto} \quad \mathcal{L}\left\{\hat{L}\delta\right\} \cdot \mathcal{L}\left\{x_{\text{FS}}\right\} = 1 \tag{7.51}$$

$$\iff H(s) = \frac{1}{\mathcal{L}\left\{x_{\text{FS}}\right\}} \tag{7.52}$$

I.e. the transfer function of a system is the (inverse of the) Laplace transformed of the fundamental solution. Whether it is directly the Laplace transformed of the fundamental solution or its inverse depends on the "direction" of the signal as shown in the example below.

7.3.4.2   Transfer function of the driven harmonic Oscillator (Single stage pendulum)

When describing a system with its transfer function, the "direction" of the signal, i.e. what counts as "input" and as "output" slightly changes the way the transfer function is written. Let us consider for example the transfer function of a simple 1-stage pendulum described as a driven (undamped) harmonics oscillator. In this case, the inhomogeneous part of the equation would be the input driving the pendulum, such that

$$\hat{L}_{\text{LTI}} \, x_{\text{in}} = \hat{L}_{\text{harmonic}}^{-1} \, x_{\text{in}} = x_{\text{out}} \tag{7.53}$$

with $\hat{L}_{\text{harmonic}}$ as defined before in (7.39). Since I already wrote the fundamental solution for $\hat{L}_{\text{harmonic}}$ in (7.44), I can directly write the transfer function as:

$$H(s) = \mathcal{L}\left\{x_{\text{FS,harmonic}}\right\} = \frac{1}{s^2 + \omega_0^2}. \tag{7.54}$$

Similar to the fundamental solution or the Green's function, the transfer function of a system allows one to easily calculate the response of the system for arbitrary input signals.

Chapter 8

Limits of the readout : Fisher Information and the Cramér-Rao bound

8.1   The Cramér-Rao bound

In the following section I denote a measured noisy signal with $x(t, \theta)$, the ideal (non–random) signal with $s(t, \theta)$, and $n(t)$ as noise (random variable). Here t it the time of the measurement and θ some parameter of interest. Additionally, I introduce a "hat" $\hat{}$ above a variable as signifier for an estimator of the same parameter. An estimator is a function to calculate a parameter from a given sample. Common estimators are for example the "arithmetic mean" as estimator for the mean μ or the "sample variance" as estimator for the variance σ^2. For a random sample $\underline{x} = \{x_1, \ldots x_N\}$ of the random variable x, I would write these two estimators as:

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad \text{and} \qquad \widehat{\sigma^2} = \frac{1}{N-1} \sum_{i=1}^{N} (\hat{\mu} - x_i)^2 \,. \tag{8.1}$$

Other estimators are for example the phase $\hat{\varphi}$, that is calculated from a measurement of an interferometric signal, which is proportional to the distance of a target mirror in a GWD suspension. The Cramér-Rao bound states how precise such an estimator can be for a given known noise distribution.

8.1.1 Definition : Fisher Information and Cramér-Rao bound (CRB)

Let $\underline{x} = \{x(t_1), \ldots x(t_N)\}$ be a sample (random vector) representing a measurement with $x(t, \theta)$ as random variable, $\rho(\underline{x}, \theta)$ as probability distribution of $\underline{x}$, depending on the time t and a θ parameter. Let further $\hat{\theta}$ be an unbiased estimator of θ; meaning that $\hat{\theta}(\underline{x})$ is a function of the sample $\underline{x}$ (estimator) and its expected value is given by $\mathbb{E}[\hat{\theta}] = \theta$ (unbiased).

If $\rho(\underline{x}, \theta)$ fulfills the "regularity conditions" written in Appendix C, then the **Fisher Information** for θ (in $\underline{x}$) is defined as:

$$\mathrm{FI}_\theta := \mathbb{E}\left[\left(\frac{\partial(\ln \rho(\underline{x}, \theta))}{\partial \theta} \right)^2 \right] \,. \tag{8.2}$$

If the Fisher Information exists and is $\neq 0$, then the **Cramér-Rao bound (CRB)** is defined by [73, Sec 4.3]:

$$\mathrm{var}(\hat{\theta}) \geq \frac{1}{\mathbb{E}\left[\left(\frac{\partial(\ln \rho(\underline{x};\theta))}{\partial \theta}\right)^2\right]} = \frac{1}{\mathrm{FI}_\theta(\underline{x})}\,. \tag{8.3}$$

When constructing an estimator for a parameter of interest, the CRB gives us a lower limit of the variance of this calculated estimator. Or in other words, for a given measurement, the CRB tells us the lower limit of how precise the readout of a parameter (here: $\hat{\theta}$) could be. Requiring the probability distribution ρ to calculate (8.3) is a rather strong restriction on the usability of the CRB. Often we don't know the exact cause of the random noise that we see in a measurement, and we don't know exactly how it is distributed. We can however derive the CRB for common and known noise source and calculate a model CRB.

Since we usually express measured signals $x(t)$ (and also derived signals like the phase $\varphi(t)$) as spectral densities, I want to introduce a slightly more specialized variation of the CRB expressed as PSD.

8.1.2 Definition : Spectral density approximation of the CRB

In terms of signal analysis (for ergodic signals), the variance of some random process x is related to its PSD as written in (7.16) via:

$$\mathrm{var}(x) = \sigma_x^2 = \int_{\mathbb{R}} S_{xx}(\omega)\ d\omega - |\mu_x|^2\,. \tag{8.4}$$

For the variance of the difference between estimator and exact value $\Delta\theta = \hat{\theta} - \theta$ we can thus write

$$\mathrm{var}(\Delta\theta) = \int_{\mathbb{R}} S_{\Delta\theta\Delta\theta}(\omega)\ d\omega\,. \tag{8.5}$$

While the exact shape of the spectrum $S_{\Delta\theta\Delta\theta}(\omega)$ depends on the present noise and how it couples into θ, I can approximate the expression by assuming it is uniformly distributed over the measured frequency bandwidth such that $S_{\Delta\theta\Delta\theta}(\omega)$ is constant. With a readout frequency of f_R (corresponding to calculating a single estimate for the interval of length $1/f_R$) the Nyquist–Shannon theorem specifies the resolvable frequency bandwidth to $f \in [0, f_R/2]$ with higher frequencies being projected/aliased back into this finite interval. For the spectral density of the error $\Delta\hat{\theta} := \hat{\theta} - \theta$ of some unbiased estimator $\hat{\theta}$ this allows us to write:

$$S_{\Delta\hat{\theta}}^{\mathrm{CRB}}(\omega) := \frac{2}{f_R}\frac{1}{\mathrm{FI}_\theta(\underline{x})} = \mathrm{const.\ in\ } \omega \tag{8.6}$$

with $S_{\Delta\hat{\theta}}^{\mathrm{CRB}}$ as lower limit of the error of $\hat{\theta}$ expressed as spectral density.

8.2 The CRB for uncorrelated additive Gaussian noise

As first case, I consider additive Gaussian noise. Meaning that the measured signal

$$x(t) = s(t, \theta) + n(t) \tag{8.7}$$

can be expressed by an analytic part $s(t, \theta)$ and an added, WSS, Gaussian distributed noise $n(t)$ that does not explicitly depend on the θ parameter, but could depend explicitly on time t. For simplicity, I assume in the following the noise term to have zero mean, such that $n \sim \mathcal{N}(\mu = 0, \sigma^2)(t)$. For non–zero mean noise, the calculation is mostly identical and can be reduced to the one shown by shifting the ideal signal by this offset $s \mapsto s + \mu$. I also shorten the expression for a specific time point t_i from $x(t_i) = s(t_i, \theta) + n(t_i)$ to simply $x_i = s_i + n_i$ in the following section. For the Gaussian distribution of n, I can write the probability distribution for measuring a single point at fixed time t_i as

$$\rho_{\text{Gauss}}(x_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - s_i)^2}{2\sigma_i^2}\right) \tag{8.8}$$

For uncorrelated noise, the joint probability distribution $\rho(\underline{x})$ can simply be expressed as the product of the individual probability functions

$$x_i, x_j \text{ uncorrelated} \implies \rho(\underline{x}) = \prod_{i=1}^{N} \rho(x_i) \tag{8.9}$$

and the resulting joint distribution would again be Gaussian.

For correlated ("colored") noise, the joint probability distribution for $\underline{x}$ can however be more complex and cannot be generally expressed from only the probability of a single point as written above. While I have a few more considerations regarding additive correlated noise terms which I included in Appendix D, I was unable to find a closed, useful expression for the CRB for such correlated noise. In the following I will therefore assume the noise n to be uncorrelated such that its variance σ^2 is constant over time (leading to $\mathbb{E}[n_i n_j] = \sigma_i^2 \delta_{ij}$).

For uncorrelated Gaussian noise, the Fisher Information for some parameter θ can be calculated straight forward via:

$$\text{FI}_\theta = \mathbb{E}\left[\left(\frac{\partial}{\partial \theta} \ln \rho(\underline{x}, \theta)\right)^2\right] \tag{8.10}$$

$$\overset{(8.9)}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N} \frac{\partial}{\partial \theta} \ln \rho_i(x_i, \theta)\right)^2\right] \tag{8.11}$$

$$\overset{(8.8)}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N} \underbrace{\frac{\partial x_i}{\partial \theta}}_{=\frac{\partial s_i}{\partial \theta}} \cdot \left(\underbrace{\frac{\partial}{\partial x_i} \ln\left(\frac{1}{\sqrt{2\pi}\sigma_i}\right)}_{=0} - \underbrace{\frac{\partial}{\partial x_i}\left(\frac{(x_i - s_i)^2}{2\sigma^2}\right)}_{\frac{1}{\sigma_i^2}(x_i - s_i)}\right)\right)^2\right] \tag{8.12}$$

$$\overset{\text{Expanding the }(...)^2}{=} \sum_{i,j=1}^{N} \frac{\partial s_i}{\partial \theta} \frac{\partial s_j}{\partial \theta} \cdot \frac{1}{\sigma^4} \quad \underbrace{\mathbb{E}\left[(x_i - s_i)(x_j - s_j)\right]}_{=\mathbb{E}[n_i n_j]=\sigma^2 \delta_{ij} \text{ because uncorrelated}} \tag{8.13}$$

$$\implies \boxed{\mathrm{FI}_\theta = \frac{1}{\sigma^2} \sum_{i=1}^{N=f_S T} \left(\frac{\partial s(t_i,\theta)}{\partial \theta}\right)^2 \approx \frac{f_S}{\sigma^2} \int_0^T \left(\frac{\partial s(t_i,\theta)}{\partial \theta}\right)^2 dt} \tag{8.14}$$

for uncorrelated Gaussian noise .

Expressed as (single-sided) spectral density this yields:

$$S_{\Delta\hat\theta}^{\mathrm{min,gauss}}(\omega) \overset{(8.6)}{=} \frac{2\sigma^2}{f_R\, f_S\, \int_0^T \left(\frac{\partial s(t,\theta)}{\partial \theta}\right)^2 dt} \quad \text{for uncorrelated Gaussian noise ,} \tag{8.15}$$

with f_R as readout frequency (and $f_R/2$ as bandwidth of the measurement). The Fisher Information contained in a measure signal $s(t)$, with some uncorrelated additive Gaussian noise, therefore corresponds to how strongly the signal changes, with respect to the parameter ($\sim \partial s/\partial\theta$).

For the original equation for the Fisher Information, the noise is characterized by its variance σ^2 which is, for zero–mean noise, equal to is (signal) energy. For practical applications, this noise energy would be given in the same units as the signal s. I.e., for a signal measured in i.e. Volt, the additive white noise would also have to be given in Volt (squared). If the additive white noise is already given in form of some amplitude spectral density, the σ^2/f_S factor corresponds to this PSD value of the noise (with total noise energy σ^2 distributed over the frequency band $[-1/2f_S, +1/2f_S]$).

8.2.1 Example CRB for the phase information of a single carrier frequency with uncorrelated additive Gaussian noise

As the previously derived formulas for the CRB may be rather abstract, let us consider a more practical example of the calculated CRB and it's meaning. Let's assume we measure a signal given by:

$$x(t) = A\sin(\omega_0 t + \varphi) + n \tag{8.16}$$

with the phase φ as parameter of interest and n as zero–mean, uncorrelated Gaussian noise $n \sim \mathcal{N}(0, \sigma^2)$. Such signals appear for example when modulating a signal $\varphi(t)$ containing information onto a carrier frequency ω_0 resulting in the measured $x(t)$. In such applications, φ changes sufficiently slow and can be assumed to be approximately constant within a "short" interval $[0, T]$ for which we estimate φ. Calculating the Fisher Information contained in a measurement interval of exactly one signal period $t \in [0, T = 2\pi/\omega_0]$ and using (8.14) yields:

$$\mathrm{FI}_\varphi = \frac{1}{\sigma^2} \sum_{i=1}^{N} \left(\underbrace{\frac{\partial x(t_j, \varphi)}{\partial \theta}}_{=A\cos(\omega_0 t + \varphi)} \right)^2 \approx \frac{f_S}{\sigma^2} \int_0^T A^2 \underbrace{\cos^2(\omega_0 t + \varphi)}_{\frac{1}{2}(1+\cos(2\omega_0 t + 2\varphi))} = \frac{f_S T A^2}{2\sigma^2} \tag{8.17}$$

If only a single phase value is estimated for a measurement of length T (corresponding to a readout frequency of $f_R = 1/T$), the CRB written as PSD is:

$$S_{\Delta\hat\varphi}^{\text{min,gauss}}(\omega) = \frac{2T}{\text{FI}_\varphi} = \frac{4\sigma^2}{A^2 f_S} \tag{8.18}$$

The calculated CRB now tells one the lower limit of how accurate $\hat\varphi$ can calculated from the measured data; independent of how exactly this calculation looks like. Calculating alternatively the (classic) SNR from the PSD of the same measured signal yields

$$\text{SNR} = \frac{\overbrace{S_{xx}(\omega_0)}^{=(A^2T+2\sigma^2/f_S)\approx A^2 T}}{\underbrace{S_{xx}(\omega \neq \omega_0)}_{\approx 2\sigma^2/f_S}} \approx \frac{f_S T A^2}{2\sigma^2} \, . \tag{8.19}$$

Comparing this expression with the one above for the Fisher Informations shows that

$$\implies \text{SNR} \stackrel{\text{here}}{=} \text{FI}_\varphi \quad \text{and} \quad S_{\Delta\hat\varphi}^{\text{min,gauss}}(\omega) \approx \frac{2}{f_R} \frac{1}{\text{SNR}} \, , \tag{8.20}$$

which means that the SNR of the measured signal corresponds in this case to the Fisher Information for the phase φ. The inverse of the SNR therefore yields its CRB and tells us how precise an estimate for φ could be.

8.2.2   CRB for a DFMI signal with additive white Gaussian noise

My main motivation for the calculation of the CRB was initially to calculate the maximal reachable precision for DFMI. Using equation (8.14) similar to the example before, I can now calculate the CRB for all parameters of a DFMI signal. I.e. assuming an ideal signal

$$s(t) = B + A\cos(\varphi + m \cdot \sin(\omega_m t + \psi)) \tag{8.21}$$
$$= B + A\sum_{n\in\mathbb{Z}} J_n(m) \ \cos(n\omega_m t + n\psi + \varphi) \tag{8.22}$$

being measured over an exact integer multiple of the modulation period $2\pi/\omega_m$ and plugged into (8.14) yields for the Fisher Information of the DFMI signal parameters:

FI_B	$\frac{f_S T}{\sigma^2}$	$S_{\Delta\hat B}^{\text{min,gauss}}(\omega) = \frac{2\sigma^2}{f_R f_S T}$
FI_A	$\frac{f_S T}{\sigma^2} \frac{1}{2}\left(3 + 4J_0(m)\cos(\varphi) + J_0(2m)\cos(2\varphi)\right)$	$S_{\Delta\hat A}^{\text{min,gauss}}(\omega) = \frac{2\sigma^2}{f_R f_S T} \frac{2}{(3+4J_0(m)\cos(\varphi)+J_0(2m)\cos(2\varphi))}$
FI_m	$\frac{f_S T}{\sigma^2} A^2 \frac{1}{4}\left(1 - J_0(2m) + J_2(2m)\cos(2\varphi)\right)$	$S_{\Delta\hat m}^{\text{min,gauss}}(\omega) = \frac{2\sigma^2}{f_R f_S T A^2} \frac{4}{(1-J_0(2m)+J_2(2m)\cos(2\varphi))}$
FI_φ	$\frac{f_S T}{\sigma^2} A^2 \frac{1}{2}\left(1 - J_0(2m)\cos(2\varphi)\right)$	$S_{\Delta\hat\varphi}^{\text{min,gauss}}(\omega) = \frac{2\sigma^2}{f_R f_S T A^2} \frac{2}{(1-J_0(2m)\cos(2\varphi))}$
FI_ψ	$\frac{f_S T}{\sigma^2} A^2 \frac{1}{4}\left(m^2 - m\,J_1(2m)\cos(2\varphi)\right)$	$S_{\Delta\hat\psi}^{\text{min,gauss}}(\omega) = \frac{2\sigma^2}{f_R f_S T A^2} \frac{4}{(m^2-m\,J_1(2m)\cos(2\varphi))}$

Table 8.1: Table of Fischer-Information and the CRB of the DFMI parameters for additive uncorrelated Gaussian noise.

I present here only the results of the calculation. The full calculation leading to these parameters involves using multiple summation identities for the Bessel functions and the integrals shown in Appendix E to reduce the expressions to their compact form written here.

8.3 The CRB for uncorrelated, Poisson distributed noise (shot noise)

When measuring the electric current consisting of individual electrons, or the power of a laser beam consisting of individual photons, the probability of measuring a number of n identical particles within a fixed time period $1/f_S$ is described by the Poisson distribution [17, section 5.1]

$$\rho(n) = \frac{e^{-\nu}\nu^n}{n!} \,, \tag{8.23}$$

with ν as mean number of particles ($\mathbb{E}[n] = \nu$), which is identical to the variance for the Poisson distribution ($\mathbb{E}[n^2] = \nu$). The physical signal would be described by $x(t) = q \cdot N(t)$ with q as conversion factor and N as random number of particles. I.e. for an electric current measured as elemental charges e^- over the time $1/f_S$ the conversion factor would be given by $q = e^- \cdot f_S$. For the optical power of a laser beam consisting out of photons with the power of $\hbar\omega_0$ sampled with the same sampling frequency and converted into an electric current by a photodiode with some responsivity R_{PD} the factor would be $q = \hbar\omega f_S R_{\mathrm{PD}}$. In Interferometry, photon shot noise is one of the fundamental noise sources, which is not the result of some faulty or imperfect technical equipment, but stems from Heisenberg's uncertainty principle for the time of arrival of the photons hitting the photo diode.

With the mean number of particles ν of the signal signal s given by $\nu = s/q$, the joint probability distribution for a measurement $\underline{n}$ would be given by

$$\rho(\underline{n}) = \prod_{i=1}^{N} \frac{e^{-\nu}\nu_i^n}{n_i!} \,. \tag{8.24}$$

Plugged into (8.2) to calculate the resulting Fisher Information then yields:

$$\mathrm{FI}_\theta^{\mathrm{poisson}} = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln\rho(\underline{n},\theta)\right)^2\right] \tag{8.25}$$

$$= \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\sum_{i=1}^{N} -\nu_i + n_i \cdot \ln\nu_i - \ln n_i!\right)^2\right] \tag{8.26}$$

$$= \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial\nu_i}{\partial\theta}\left(\frac{n_i}{\nu_i}-1\right)\right)^2\right] \tag{8.27}$$

$$= \sum_{i,j=1}^{N}\frac{\partial\nu_i}{\partial\theta}\frac{\partial\nu_j}{\partial\theta}\frac{1}{\nu_i\nu_j}\cdot\underbrace{\mathbb{E}\left[(n_i-\nu_i)(n_j-\nu_j)\right]}_{=\nu_i\delta_{ij} \text{ for uncorr. Poisson dist. noise}} \tag{8.28}$$

$$= \sum_{i=1}^{N} \left(\frac{\partial \nu_i}{\partial \theta} \right)^2 \cdot \frac{1}{\nu_i} \tag{8.29}$$

$$= \sum_{i=1}^{N} \left(\frac{\partial s_i}{\partial \theta} \right)^2 \cdot \frac{1}{q \cdot s_i} \tag{8.30}$$

$$\approx \frac{f_S}{q} \int_0^T \frac{\left(\frac{\partial s(t,\theta)}{\partial \theta} \right)^2}{s(t,\theta)} \, dt \tag{8.31}$$

For a DFMI signal with optical power P_0 ($\sim s$) and Poisson distributed noise, calculating the CRB for the phase φ yields:

$$S_{\Delta\hat{\varphi}}^{\text{min,Poisson}}(\omega) = \frac{1}{(1 + J_0(m) \cos \varphi)} \cdot \frac{\lambda q_e}{\pi f_R R_{\text{PD}} \kappa P_0} \tag{8.32}$$

8.4   Digitization Noise and Dithering

Digitization noise is the result of the truncation of the analog signal into digital bits. For example: a 16-bit ADC and an input range of $\pm 10V$ will truncate a digitized signal to the next smallest integer multiple of LSB $= 20/2^{16}$ V ≈ 0.3 mV. Here, LSB denotes the voltage level of the least significant bit of the digitized signal. The noise of this truncation is the difference between exact and digitized (truncated) value and can be expressed as additive noise given by

$$r(t) := s(t) - \lfloor s(t) \rfloor \in [0, \text{LSB}] \,. \tag{8.33}$$

Respectively, the measured signal can be written as $\lfloor s(t) \rfloor = s(t) + r(t) \,.$

For an extended measurement this type of noise is at first however correlated to the exact signal. To mitigate the effects of this quantization noise, a technique called "dithering" is used. While the total noise power cannot be reduced, dithering aims to break the correlation between the noise and the signal to convert the correlated noise to white noise leading to effectively less noise within a finite measurement band. To achieve this, an additional (known) random signal is added, referred to as dither noise, which effectively shapes the quantization noise to an almost uniformly distributed white noise. While the uniform distribution does not satisfy the CRB regularity conditions [73, Section 4.3] (meaning that the CRB inequality cannot be applied directly for such (purely) uniform distributed noise), a CRB for similar white additive Gaussian noise with the same variance can be used as a good approximation.

8.4.1   Dithering

Let us first consider a simplified example where adding random 'noise' can improve a truncated signal's accuracy: Consider a constant signal of e.g. $s = 0.9$ which will be truncated during some processing step, down to the next smallest integer $\lfloor 0.9 \rfloor = 0$. Without dithering, this value will always remain at 0. Aiming to improve the accuracy, we add a random number x uniformly distributed between 0 and 0.9

in steps of 0.1. Ten consecutive measurements of $\lfloor s + x \rfloor$ now yield on average:

$$s + x \in \{0.9, 1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9\} \tag{8.34}$$

$$\implies \lfloor s + x \rfloor \in \{0, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\} \tag{8.35}$$

$$\implies \mathrm{mean}(\lfloor s + x \rfloor) = 0.9 \tag{8.36}$$

While the truncated value of $\lfloor s + x \rfloor$ remains inaccurate; the mean value of x over an extended measurement with dithering ($\mathrm{mean}(\lfloor s + x \rfloor) = 0.9$) is more accurate than the mean value without dithering ($\mathrm{mean}(\lfloor s \rfloor) = 0$).

8.4.1.1   Breaking noise correlation using dither noise

Truncating a digital signal s down to a smaller number of bits introduces 'truncation noise' as the difference between the original signal and the truncated one. Adding a uniformly distributed dither noise x to a signal s before truncation, with x in the range of 0 up to the least significant bit of the truncated signal, reduces the average error due to truncation noise to 0.

To show this formally: let s_{exact} be a digitized signal with bit length b and $s_{\mathrm{no\text{-}dither}} = \lfloor s \rfloor$ the truncated signal with fewer bits ($b - \Delta b$). The truncation operation rounds down to the nearest integer multiple of Δb. Additionally, we will assume that the signal is positive ($s_{\mathrm{exact}} \geq 0$). For arbitrary (positive and negative) values the following calculation is largely the same but requires additional case distinctions expanding the equations significantly which I omit here.

The rest term that is dropped by the truncation can be written as

$$r := s_{\mathrm{exact}} - s_{\mathrm{no\text{-}dither}} \quad \text{with} \quad 0 \leq r < \Delta b. \tag{8.37}$$

Without any dithering, the additive noise of the truncation operation is given by this r:

$$s_{\mathrm{no\text{-}dither}} = s_{\mathrm{exact}} - r. \tag{8.38}$$

Now let x be a random dither variable with $0 \leq x \leq \Delta b$ added to the signal s_{exact} before truncation. The truncated signal with dithering can then be written as

$$s_{\mathrm{dithered}} := \lfloor s_{\mathrm{exact}} + x \rfloor = s_{\mathrm{exact}} - r_{\mathrm{dithered}} \tag{8.39}$$

$$\text{with} \quad r_{\mathrm{dithered}} := s_{\mathrm{exact}} - \underbrace{\lfloor s_{\mathrm{exact}} + x \rfloor}_{\substack{= \lfloor s_{\mathrm{no\text{-}dither}} + r + x \rfloor \\ = s_{\mathrm{no\text{-}dither}} + \lfloor r + x \rfloor}} = r - \lfloor r + x \rfloor \tag{8.40}$$

The expected value of the (dithered) truncation noise can be calculated by:

$$\mathbb{E}\big[r_{\mathrm{dithered}}\big] = r - \mathbb{E}\big[\lfloor r + x \rfloor\big] = r - \int_0^{\Delta b} \rho(x) \cdot \lfloor r + x \rfloor \, dx. \tag{8.41}$$

Since $r < \Delta b$ and $x \leq \Delta b$ for the sum we find that $0 \leq r + x < 2\Delta b$. After truncation this leads to $\lfloor r + x \rfloor$ being either 0 or $1 \cdot \Delta b$. In other words I can write:

$$r_{\mathrm{dithered}} = \lfloor r + x \rfloor = \begin{cases} 0 & \text{if} \quad r + x < \Delta b \\ \Delta b & \text{if} \quad r + x \geq \Delta b \end{cases} = \Delta b \cdot \Theta(x + r - \Delta b) \tag{8.42}$$

with Θ as unit step function. For a uniformly distributed dither noise $\rho_{\mathrm{uni}}(x) = 1/\Delta b$, the expected error of the dithered and truncated signal thus becomes:

$$\mathbb{E}\big[r_{\mathrm{dithered}}\big] = r - \int_0^{\Delta b} \rho_{\mathrm{uni}}(x) \cdot \lfloor r + x \rfloor \, dx \tag{8.43}$$

$$= r - \int_0^{\Delta b} \frac{1}{\Delta b} \cdot \Delta b \cdot \Theta(x + r - \Delta b) \, dx \tag{8.44}$$

$$\stackrel{(x':=x+r-\Delta b)}{=} r - \int_{r-\Delta b}^r \Theta(x') \, dx' \tag{8.45}$$

$$\stackrel{(\Delta b \geq r)}{=} r - \int_0^r 1 \, dx' \tag{8.46}$$

$$= 0 \tag{8.47}$$

I.e. for a perfect uniformly distributed dither, the additive truncation noise (r_{dithered}) becomes zero–mean and since $\mathbb{E}\big[r_{\mathrm{dithered}} \cdot s_{\mathrm{exact}}\big] = \mathbb{E}\big[r_{\mathrm{dithered}}\big] \cdot s_{\mathrm{exact}} = 0 \cdot s_{\mathrm{exact}} = 0$, the (cross-) correlation between truncation noise and exact signal vanishes as well. The remaining noise level for this dither noise would be given by the variance of the uniform distribution (of length Δb) equal to $\Delta b^2/12$.

Practically, the dither noise x does not have infinite precision but would also be quantized to the LSB of s_{exact} which leaves a still correlated noise part in the truncation noise. (Is most cases this correlated noise part is however smaller than the white noise contribution from the dithering). Besides a uniformly distributed dither noise, differently shaped noise distributions like a triangular distribution or higher order distributions are also used to generate the dither noise. For these distributions (not shown explicitly here), the mean / the correlated noise part, and the white noise part vary in power and depending on the signal it can be useful to choose a distribution which causes a slightly higher correlated noise part if the resulting effective (white) noise may be lower compared to the uniform distribution.

8.4.1.2   Approximate CRB for the digitization noise of a dithered DFMI signal

For a perfect uniform dither distribution with variance $\sigma_{\mathrm{digit}}^2 = \mathrm{LSB}^2/12$, the CRB for additive Gaussian noise with the same power, would be

$$\mathrm{var}(\hat{\theta}) \geq \frac{\frac{\mathrm{LSB}^2}{12}}{f_S \int_0^T dt \left(\frac{\partial s(t,\theta)}{\partial \theta}\right)^2} \,, \tag{8.48}$$

which I later will use as approximate CRB for the digitization noise.

8.5   CRB for ReDFMI

When first analyzing the CRB for a given signal with additive noise; equation (8.14) showed analytically that the amount of Fisher Information (and therefore a higher possible precision) scaled with the slope / the derivative of the signal with respect to

the readout parameter. In Interferometry, signals from optical cavities on resonance are known to have such a steep slope for the phase, and they are commonly used to generate error–signals, highly sensitive to the ratio of laser frequency to optical cavity length. As mentioned in chapter 1, the arms of the laser interferometers of gravitational wave detectors are in fact set up as resonant cavities, as it improves the sensitivity and thus the precision of the readout. For a homodyne signal, this can directly be shown by calculating the CRB:

8.5.1 CRB for the cavity response to a homodyne signal

With a cavity response for a homodyne signal of the form:

$$P_{\text{trans}} = \frac{\kappa P_{\text{in}}}{1 + F \sin^2(\varphi)} \quad \text{and} \quad P_{\text{refl}} = P_{\text{in}} - P_{\text{trans}} , \tag{8.49}$$

the Fischer-Information for additive white Gaussian noise can be calculated via:

$$\text{FI}_\theta^{\text{Gaussian}} = \frac{f_S}{\sigma^2} \int_0^T dt \left(\frac{\partial P_{\text{trans}}}{\partial \varphi} \right)^2 = \frac{f_S}{\sigma^2} \int_0^T dt \left(\frac{\partial P_{\text{relf}}}{\partial \varphi} \right)^2 = \frac{f_S T}{\sigma^2} \frac{\kappa^2 P_{\text{in}}^2 F^2 \sin^2(2\varphi)}{(1 + F \sin^2(\varphi))^4} . \tag{8.50}$$

Close to the resonance (where $\varphi < \arcsin(\sqrt{1/F})$ and thus $F \sin^2(\varphi) < 1$ and $F^2 \sin^2(2\varphi) \sim F$), increasing the Finesse coefficient $F = 4R/(1 - R)^2$ of the cavity (meaning higher reflectivities R) also increases the amount of Information in the signal.

8.5.2 Cavity response to a frequency ramped signal

Next we consider a signal with a (slow) linear frequency shift where the phase would be given by $\varphi \mapsto \varphi + \Delta\Omega \cdot t$, leading to an approximate cavity response of

$$P_{\text{refl}} = P_{\text{in}} \left(1 - \kappa \frac{1}{1 + F \sin^2(\varphi + \Delta\Omega \cdot t)} \right) \tag{8.51}$$

$$\text{and} \quad P_{\text{trans}} = P_{\text{in}} \kappa \left(\frac{1}{1 + F \sin^2(\varphi + \Delta\Omega \cdot t)} \right) , \tag{8.52}$$

with the squared derivative:

$$\left(\frac{\partial P_{\text{refl}}}{\partial \varphi} \right)^2 = \left(\frac{\partial P_{\text{trans}}}{\partial \varphi} \right)^2 = \frac{4 P_{\text{in}}^2 \kappa^2 F^2 \sin^2(\varphi + \Delta\Omega \cdot t) \cos^2(\varphi + \Delta\Omega \cdot t)}{(1 + F \sin^2(\varphi + \Delta\Omega \cdot t))^4} \tag{8.53}$$

For the Gaussian noise limit, the Fischer Information for a measurement over the interval $[0, 2\pi/\Delta\Omega]$ is given (up to some constant coefficients) by:

$$\text{FI}_\varphi^{\text{Gauss}} = \frac{f_S}{\sigma^2} \int_0^{2\pi/\Delta\Omega} dt \left(\frac{\partial P_{\text{refl}}}{\partial \varphi} \right)^2 \tag{8.54}$$

Substituting $\sqrt{F} \sin(\varphi + \Delta\Omega \cdot t) =: y$ with $dt = \pm \frac{1}{\Delta\Omega \sqrt{F - y^2}} dy$ and using the symmetries of the $\sin^2$ over the 2π interval to rearrange the integration boundaries leads

to:

$$\int_0^{2\pi/\Delta\Omega} dt \ \left(\frac{\partial P_{\text{refl}}}{\partial \varphi}\right)^2 = 4 \int_0^{\sqrt{F}} dy \ \frac{4P_{\text{in}}^2 \kappa^2}{\Delta\Omega} \frac{y^2 \sqrt{F - y^2}}{(1 + y^2)^4} \tag{8.55}$$

$$= \frac{16P_{\text{in}}^2 \kappa^2}{\Delta\Omega} \frac{\pi F^3(F + 2)}{32(F + 1)^{5/2}} \tag{8.56}$$

$$\sim F^{3/2} \ . \tag{8.57}$$

(The integral on the right side can be solved by repeated integration by parts).

The resulting Fisher Information (and the corresponding CRB) scales with the Finesse coefficient $\sim F^{3/2}$, meaning that by increasing the Finesse of the cavity (using mirrors with a higher reflectivity), one could theoretically improve the readout precision not just for homodyne signals close to the cavity resonance, but for a frequency modulated signal as well.

While the signal used here is only an approximation, it is the first indication that a DFMI signal, in combination with a cavity, can also yield a higher precision phase readout compared to the DFMI signal alone. This is the main motivation for the concept of ReDFMI presented in chapter 6.

Chapter 9

Noise model of a measured DFMI signal

In this chapter the theoretical results of the previous chapter are applied to calculate the lower limit of the phase readout for the case of a measured DFMI signal.
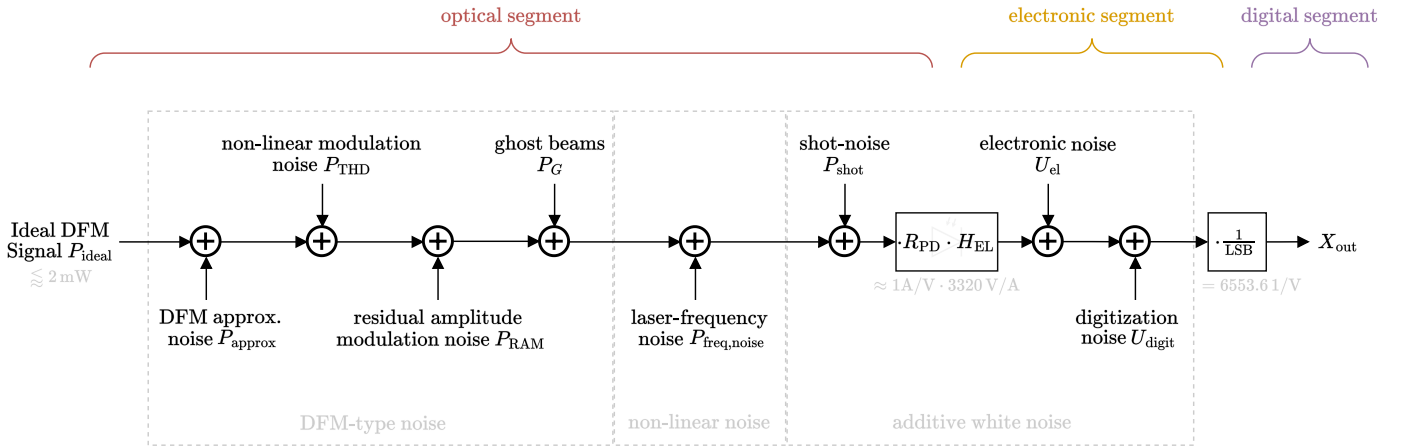
Figure 9.1: Schematic view of all 8 contributions to the noise model considered in this chapter

Figure 9.1 shows a block diagram of the readout 'chain' with the noise contributions considered here and the conversion factor appearing when converting the initial optical signal into bit values within the computational hardware.

For the initial optical signal P_{ideal}, I assume a signal power of $P_0 = 2\,\text{mW}$ and an optical contrast of $\kappa = 0.9$ which leads to $B = 1\,\text{mW}$ and $A = 0.9\,\text{mW}$. From left to right, the initial signal and the noise sources in the DFM interferometer are given as optical signals in units of [W]. The optical signals are converted by a photo diode into a small photo current in units of [A]. The proportionality factor is given by the photo diode responsivity $R_{\text{PD}} \approx 1\,\text{A/W}$. The electrical current is then amplified by a trans–impedance amplifier and filtered. The full effect of this electronic stage is modeled by the electronic transfer function $H_{\text{EL}}(\omega)$ which is for my experimental setup discussed in Appendix F. For the noise analysis here, I use a simplified model and approximate this transfer function only by an idealized amplifier gain $H_{\text{EL}}(\omega) \approx 3320\,\text{V/A}$ converting the current into a voltage signal (in units of [V]). In

the end, the voltage signal is digitized by a 16-bit ADC with a total range of 10 V leading to a least significant bit conversion factor of LSB $= 10\,\text{V}/2^{16} \approx 152.59\,\mu\text{V}$.

The noise sources considered here are sorted into 3 categories: (1) additive (uncorrelated) white noise sources, (2) additive "colored" noise sources and (3) DFMI–type noise sources. As DFMI–type noise sources I denote additional DFMI–type signals superposed on the main signal and with the same modulation frequency as the main signal. "Colored" noise source are all other noise contributions that have a frequency dependent shape, and white noise sources are the remaining uncorrelated noise contributions with a flat frequency spectrum.

9.1 White noise sources

For uncorrelated white noise sources, the previously calculated CRB as presented in (8.14) and (8.32) can be applied directly. Such noise terms are often expressed in one of three ways: as stochastic variance (or "RMS amplitude"), as time series / signal or already as ASD/PSD. When plugging the signal $s(t)$ and noise in the CRB equation, one has to make sure that both are given in the same units and applying additional factors of $R_{\text{PD}}, H_{\text{el}}$ and/or LSB to the signal may be necessary. A noise expressed as stochastic variance σ^2 can be plugged in directly in (8.14). A noise expressed as PSD equals the σ^2/f_S factor in (8.14). For a noise expressed as time series, the corresponding variance has to be calculated via $\sigma^2 = \int_0^T (x(t) - \overline{x})^2 dt$ (with $\overline{x}$ as mean value of $x(t)$, which is in all cases considered here zero).

9.1.1 Digitization (ADC) noise U_{digit}

For the technical specifications of the used ADC (1 MHz sampling frequency, 16-bit precision) the noise limit derived in section 8.4 leads to a limit due to digitization noise of

$$\delta\varphi_{\text{ADC}} < 1.52 \cdot 10^{-8}\frac{\text{rad}}{\sqrt{\text{Hz}}} \quad \sim \quad \delta L_{\text{ADC}} < 3.75 \cdot 10^{-15}\frac{\text{m}}{\sqrt{\text{Hz}}}. \tag{9.1}$$

9.1.2 Electronic noise and PD dark noise U_{el}

For the effective electric- and PD dark noise, I use the values measured from the electronics as shown in Figure 9.2. Without any signal on the photodiode, the combined electronic-amplifier and photodiode dark-noise lead roughly to a white noise floor of $\sigma_{\text{el}}/\sqrt{f_S} = 2\,\mu\text{V}/\sqrt{\text{Hz}}$. For frequencies above $\sim 50\,\text{kHz}$, the electronic circuit seems to oscillate / 'ring' at certain frequencies. This was also confirmed when measuring the electronic amplifiers output with a different ADC and is likely the caused by non–ideal and not properly damped OpAmps. The amplifier circuit is currently being worked on to remove these resonances.

The readout limit from the measured white-noise floor and calculated by (8.1) is:

$$\delta\varphi_{\text{el}} < 4.88 \cdot 10^{-7}\frac{\text{rad}}{\sqrt{\text{Hz}}} \quad \sim \quad \delta L_{\text{el}} > 1.20 \cdot 10^{-13}\frac{\text{m}}{\sqrt{\text{Hz}}}. \tag{9.2}$$
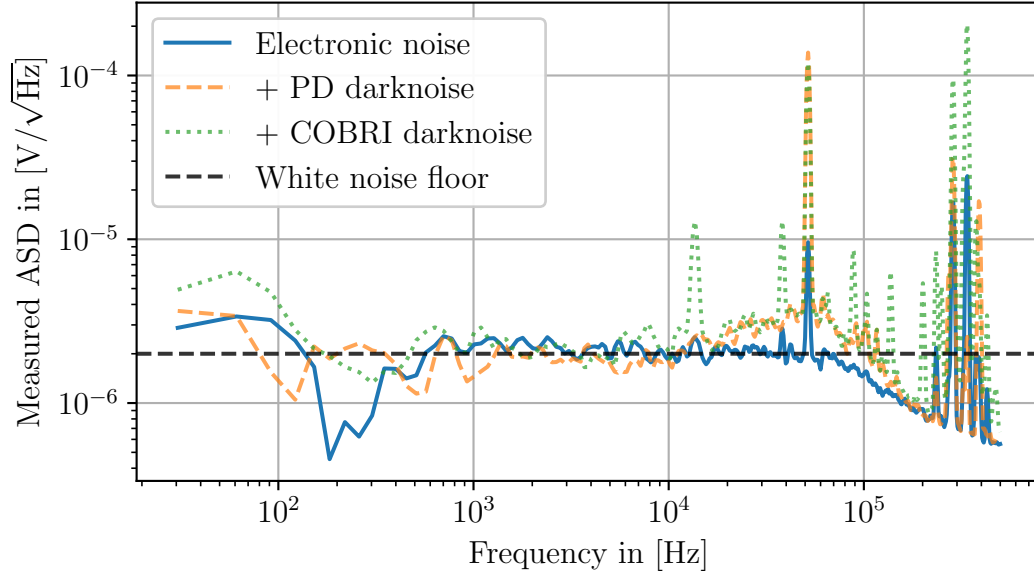
Figure 9.2: Electronic noise of the electronic amplifier circuit while the laser is switched off, measured with a 1 MHz ADC, downsampled to $\approx 4\,\text{kHz}$, and plotted with a LPSD plotting routine. The measured noise shown here is dominated by a white noise floor from the analog amplifier and filter electronics, even when it's not connected to any photo diode. When connecting a photodiode, some excess noise above 10 kHz appears. This additional noise visible as distinct "peaks"/resonances of the electronics are likely related to the additional capacitance of the connected cables and interaction with the OpAmps. The electronics board is currently being redesigned to fix this issue.

9.1.3 Shot noise P_{shot}

For the shot noise, described by a Poisson' instead of a Gaussian distribution, I derived the CRB in section 8.3 (and specifically in (8.32)) to be:

$$\delta\varphi_{\text{shot}} < \frac{1}{(1 + |J_0(m)|)} \cdot \frac{\lambda q_e}{\pi f_R R_{\text{PD}} \kappa P_0} \tag{9.3}$$

For an optical power of roughly $P_0 \approx 2\,\text{mW}$, an optical contrast of $\kappa \approx 0.9$ and a photodiode responsivity of $R_{\text{PD}} \approx 0.9$, the resulting shot noise limit is:

$$\delta\varphi_{\text{shot}} < 1.19 \cdot 10^{-8} \frac{\text{rad}}{\sqrt{\text{Hz}}} \quad \sim \quad \delta L_{\text{shot}} < 2.93 \cdot 10^{-15} \frac{\text{m}}{\sqrt{\text{Hz}}} \tag{9.4}$$

9.2 "Colored" noise sources

As "colored" noise sources I consider all independent noise sources that have a specific frequency-dependent shape but do not depend on the main DFMI signal. In my experiment, I only consider the laser frequency noise of this class.

9.2.1 Laser Frequency noise $P_{\text{freq,noise}}$

Independent of the applied frequency modulation, the laser frequency is not perfectly constant but noisy with a characteristic $1/f$ dependence. This noise couples directly into the measured phase (phase from laser frequency noise is indistinguishable from the signals phase) which allows me to write down the resulting phase

and length noise limit directly. For the laser I used, the frequency noise was measured by interfering two identical lasers and tracing the resulting "beat-frequency" (the varying frequency difference between the two lasers). This measurement was done by Leander Weickhardt of our working group who measured an approximate laser frequency noise of $\sim 10\,\text{kHz}/\sqrt{\text{Hz}}$ at $1\,\text{Hz}$ (with a mean laser frequency of $1550\,\text{nm} \sim f_0 = 193\,\text{THz}$). With a relative arm length difference of $\Delta L = 10\,\text{cm}$ between two interfering laser beams, the phase (and subsequent length) noise can be written as

$$\delta\varphi_{\text{low-f-laser}}(f) \approx 10^4 \frac{\text{Hz}}{\sqrt{\text{Hz}}} \cdot 2\pi \frac{\Delta L}{c_0} \cdot \frac{\text{Hz}}{f} = 2.10 \cdot 10^{-5} \frac{\text{rad}}{\sqrt{\text{Hz}}} \cdot \frac{\text{Hz}}{f} \qquad (9.5)$$

$$\text{and} \quad \delta L_{\text{low-f-laser}}(f) \approx 5.17 \cdot 10^{-12} \frac{\text{m}}{\sqrt{\text{Hz}}} \cdot \frac{\text{Hz}}{f} \qquad (9.6)$$

As this phase noise adds linearly to the measured phase, it can be subtracted by measuring a second, fixed reference signal sharing the same laser frequency noise and subtracting its measured phase from the main measurement.

9.3 DFMI–type noise sources

As DFMI–type noise sources I denote additive DFMI signals with DFMI parameters different from the main signal but sharing the same modulation frequency. These signal appear at the same harmonic frequencies as the main signal and are highly correlated to it.

My calculations using the CRB did not yield a prediction for the precision of correlated noise terms. For the precision limits caused by this noise I therefore use the following estimate: In a "worst case" scenario, a DFMI–type noise "signal" has the same modulation index as the main signal and only varies in phase and amplitude. In this case the DFMI–type noise has the highest correlation with the main signal as its signal power shares (almost) the same distribution over the signal harmonics. When calculating the phase within the analytic readout algorithm, the noise terms then add linearly to the quadratures of the harmonics such that for the I_φ and Q_φ from section 4.8 one finds:

$$I_\varphi \sim \cos\varphi + \epsilon\cos\varphi_{\text{noise}} \quad \text{and} \quad Q_\varphi \sim \sin\varphi + \epsilon\sin\varphi_{\text{noise}}\,, \qquad (9.7)$$

with φ_{noise} as deterministic phase of the DFMI–type noise "signal" (and not the resulting phase noise itself) and ϵ as amplitude ratio between the DFMI–type noise and the main signal $\epsilon \sim P_{\text{DFMI–type-noise}}/P_{\text{ideal}}$. Expanding the result of the subsequent arctan calculation yields

$$\varphi_{\text{estimate}} = \arctan\left(\frac{Q_\varphi}{I_\varphi}\right) = \varphi + \arctan\left(\frac{\epsilon\sin(\Delta\varphi_{\text{noise}})}{1 + \epsilon\cos(\Delta\varphi_{\text{noise}})}\right) \qquad (9.8)$$

$$\approx \varphi + \epsilon \cdot \underbrace{\sin(\Delta\varphi_{\text{noise}})}_{\approx \Delta\varphi_{\text{noise}}} + \mathcal{O}(\epsilon^2) \qquad (9.9)$$

with $\Delta\varphi_{\text{noise}} = \varphi - \varphi_{\text{noise}}$ as phase difference between the "noise signal" and the main signal. With $\delta\varphi_{\text{error}} = \varphi_{\text{estimate}} - \varphi \approx \epsilon \cdot \Delta\varphi_{\text{noise}}$, the effective noise ASD

would to first order be the ASD of the DFMI–type noise phase difference $S_{\Delta\varphi_{\text{noise}}}(\omega)$ attenuated by the amplitude ratio ϵ.

Some of DFMI–type noise sources presented below should only have a constant offset compared to the main signal, which would lead to effectively no contribution to the noise for $\omega > 0$. More generally, I consider a phase difference noise of $\Delta\varphi_{\text{noise}} \sim 10^{-3}\,\text{rad}/\sqrt{\text{Hz}}$ which roughly corresponds to some of the expected movement of a suspended mass in a GWD suspension. Practically, it should be possible to mitigate these noise contribution by subtracting the corresponding (deterministic) DFMI signals from the measured signal if their signal parameters are known.

9.3.1   DFMI approximation noise P_{approx}

As "DFMI approximation noise" I denote the terms promotional to $\psi_L = \omega_m\tau_L$ that are dropped when calculating the "ideal" DFMI signal as shown in Appendix A. To first order, (when dropping all $\mathcal{O}(\psi_L^2)$ terms), the exact DFM signal measured is given by

$$P_{\text{exact}}(t) = P_{\text{ideal}}(t) + P_{\text{approx}}(t) + \mathcal{O}(\psi_L^2) \tag{9.10}$$

with the DFMI approximation noise given by

$$P_{\text{approx}}(t) = \psi_L \cdot (m + m_0) \cdot \sin(\omega_m t + \psi) \cdot P_{\text{ideal}} \tag{9.11}$$

$$\implies \epsilon_{\text{approx}} = \max\left\{\frac{P_{\text{approx}}}{P_{\text{ideal}}}\right\} = \psi_L \cdot (m + m_0)\,. \tag{9.12}$$

The m_0 here corresponds to the total propagation time of the light since emission and not only the path length difference of the two interfering beams. For the Free-Beam setups in my experiment (shown in the next chapter), this m_0 is larger $\sim 4 \times m$ than the path length difference. Together with a modulation frequency of $500\,\text{Hz}$ and a path length difference of $\Delta L = 10\,\text{cm}$, this yields an "approximation phase" of $\psi_L = 1\,\mu\text{rad}$, which yields (with $m \approx 7$) for the error of the Free-Beam setup:

$$\delta\varphi_{\text{approx}} \approx 3.67 \cdot 10^{-8}\frac{\text{rad}}{\sqrt{\text{Hz}}} \tag{9.13}$$

$$\sim \quad \delta L_{\text{approx}} \approx 9.05 \cdot 10^{-15}\frac{\text{m}}{\sqrt{\text{Hz}}}\,. \tag{9.14}$$

For the COBRI, with a shorter second arm leading to $m_0 \sim 0.1 \times m$, the same approximation noise should be roughly a factor of ~ 5 smaller.

9.3.2   Non–linear laser modulation noise P_{THD}

The second DFMI–type noise considered here are additional terms caused by an imperfect sinusoidal frequency modulation.

A common class of frequency tunable lasers are "external-cavity diode laser." By tuning the length of the external cavity, the frequency of the light coming from the laser diode is filtered and only the specific frequency on resonance with the cavity is amplified and emitted. The 'tuning' of the external cavity length is usually done by a Piezo element which expands when a voltage is applied. In the experimental setup

presented in the next chapter, the modulation of the laser frequency is realized by applying a (sinusoidal) voltage $U_{\text{mod}} = U_{\text{mod},0} + U_\Delta \cdot \sin(\omega_m t)$ to the Piezo, changing the cavity length and with it, the laser output frequency.

In an ideal scenario, the Piezo length changes linearly with an applied voltage. In practice, the Piezo crystal has a hysteresis behavior [74] and heats up over time. Figure 9.3 displays an exaggerated example of such (non–linear) hysteresis behavior and how it would affect the laser output frequency. Both of these effects cause the Piezo displacement, and thus the outgoing laser frequency, to not scale perfectly linear with the input voltage. For the sinusoidal input voltage, this will lead to additional higher order terms ($+2\omega_m, +3\omega_m, \ldots$) of the laser frequency. The ratio of the summed up RMS amplitude of these higher order terms divided by the amplitude of the fundamental frequency (here: ω_m) defines the "harmonic distortion" in electrical engineering.
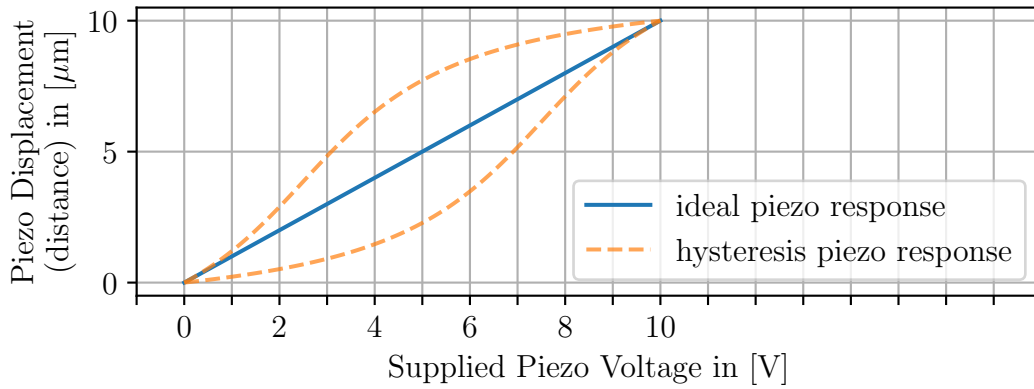
(a) "Displacement-response" of a model Piezo-crystal to a supplied input voltage.
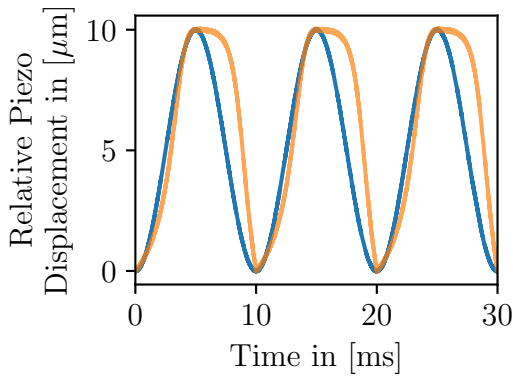
(b) Relative displacement / movement of the model Piezo for a sinusoidal input voltage.
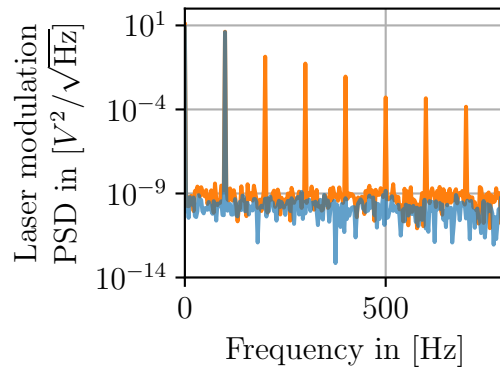
(c) Model PSD of the Piezo output (the measured modulation signal).

Figure 9.3: Plots of the (non–linear) behavior of an example Piezo. The blue line corresponds to an ideal, fully linear displacement response of a Piezo and the yellow line to a more realistic Piezo response that leads to the harmonic distortion visible in (c).

Assuming that most of this unwanted noise is concentrated in the second harmonic (closest to the main modulation frequency), the laser frequency could be written as:

$$\omega_{\text{DFM}}(t) \approx \omega_0 + \Delta\omega \cdot (\sin(\omega_m t + \psi) + \text{THD} \cdot \sin(2\omega_m t + \psi_{\text{THD}})) \tag{9.15}$$

with THD as "total harmonic distortion" coefficient.

A series expansion of the measured DFMI signal then yields:

$$P_{\text{measured}} = B + A\cos(\omega_{\text{DFM}}(t) \cdot \tau) \tag{9.16}$$

$$= B + A\cos\Big(\varphi + m \cdot \sin(\omega_m t + \psi)\Big)$$

$$- \underbrace{\text{THD} \cdot m \cdot \sin(2\omega_m t + \psi_{\text{THD}})A\sin\Big(\varphi + m \cdot \sin(\omega_m t + \psi)\Big)}_{=:P_{\text{THD}}(t)}$$

$$+ \mathcal{O}(\text{THD}^2) \tag{9.17}$$

$$\implies P_{\text{THD}}(t) = -\text{THD} \cdot m \cdot \sin(2\omega_m t + \psi_{\text{THD}}) \cdot P_{\text{ideal}}(t) \tag{9.18}$$

The THD noise as written in (9.18) can alternatively be expressed as amplitude noise where the main signals amplitude P_{ideal} is modulated by

$$P_{\text{ideal}} \mapsto (1 - \text{THD} \cdot m \cdot \sin(2\omega_m t + \psi_{\text{THD}})) \cdot P_{\text{ideal}}(t) \tag{9.19}$$

In previous experiments with a similar laser, Isleif & Gerberding measured a THD in the order of 5% of the main signal. For the resulting phase noise with $\epsilon_{\text{THD}} \approx \text{THD} \cdot m$, this yields:

$$\delta\varphi_{\text{THD}} \approx 3.5 \cdot 10^{-4} \frac{\text{rad}}{\sqrt{\text{Hz}}} \tag{9.20}$$

$$\sim \quad \delta L_{\text{THD}} \approx 8.63 \cdot 10^{-11} \frac{\text{m}}{\sqrt{\text{Hz}}} \, . \tag{9.21}$$

9.3.3　Residual amplitude modulation noise P_{RAM}

Besides the laser frequency, the laser output power also varies with the applied modulation signal. Figure 9.4 shows the laser power output, depending on the applied Piezo voltage. This amplitude modulation is however not completely random
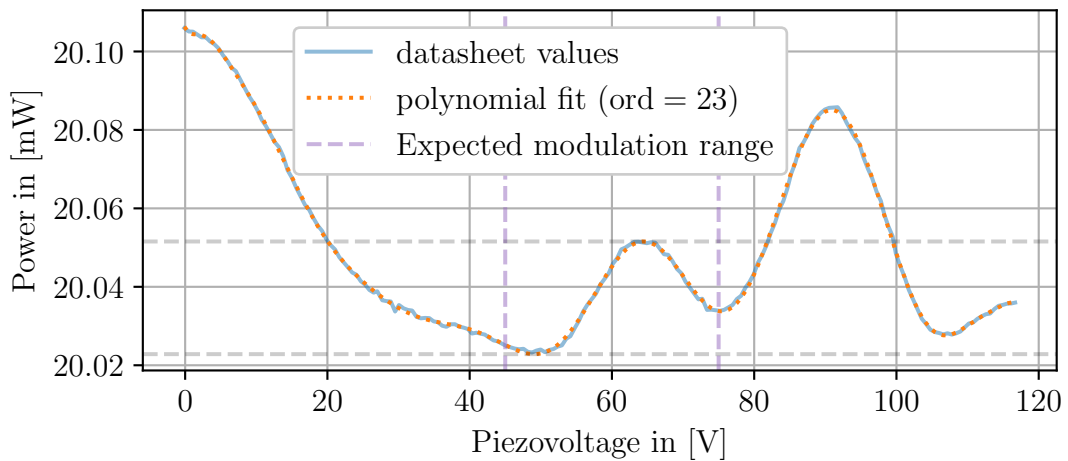
Figure 9.4: Picture of the Laser-manufacturers' test data sheet showing a measurement of the laser's output power depending on the applied Piezo voltage. If unaccounted for, a mean Piezo voltage of $U_0 = 60\,\text{V}$ and a modulation amplitude of $U_\Delta = 5.05\,\text{V}$ (leading to $\Delta\omega \approx 3.34\,\text{GHz}$) would lead to amplitude variations in the order of $\epsilon_{\text{RAM}} \approx 1.5 \cdot 10^{-3}$.

but a deterministic response of the laser to the frequency modulation signal. It can

be compensated by an amplitude stabilization control loop that measured the laser's outgoing power and compensates fluctuations by using a variable gain amplifier. If not stabilized, the expected modulation signal will lead to an output power (and thus the signal amplitude) fluctuation of a factor of $\epsilon_{\mathrm{RAM}} \approx 1.5 \cdot 10^{-3}$. For the resulting phase and length noise this would yield

$$\delta\varphi_{\mathrm{RAM}} \approx 1.5 \cdot 10^{-6} \frac{\mathrm{rad}}{\sqrt{\mathrm{Hz}}} \qquad (9.22)$$

$$\sim \quad \delta L_{\mathrm{RAM}} \approx 3.7 \cdot 10^{-13} \frac{\mathrm{m}}{\sqrt{\mathrm{Hz}}} \,. \qquad (9.23)$$

9.3.4 Stray light & ghost Beams P_{ghost}

The remaining noise that can be written exactly as DFMI–type noise are stray light or "ghost beams" of the optical setup. Small partial reflections of the main beam (ghost beams) that leave the main path of propagation but hit the photo diode again will interfere with the main beam leading to a DFMI–type noise. Gerberding and Isleif specifically discuss the effect of such beams traveling along the main beam in their COBRI setup in [63]. In their setup, they consider the ghost beams that are caused by small reflections at the various (plane) surfaces between two mediums with different refractive indices. Being reflected back and forth in the optical setup, these ghost beams can have a wide range of potentially very large modulation indices (corresponding to a large path length difference between the main beam and the ghost beam). Additionally, seismic motion slightly shaking the setup also causes some additional ghost beams that will correlate with the (low frequency) seismic noise. For the COBRI, the main ghost beam contributions are expected to come from some residual reflectivity of $R_{\mathrm{ghost}} < 0.25\%$ (assumed as upper limit by Gerberding & Isleif [63]). In the most extreme case this would lead to a readout limit of

$$\delta\varphi_{\mathrm{ghost}} \approx 1.0 \cdot 10^{-5} \frac{\mathrm{rad}}{\sqrt{\mathrm{Hz}}} \qquad (9.24)$$

$$\sim \quad \delta L_{\mathrm{ghost}} \approx 2.47 \cdot 10^{-12} \frac{\mathrm{m}}{\sqrt{\mathrm{Hz}}} \,. \qquad (9.25)$$

Realistically, these ghost beams vary in path length difference (and thus modulation index and phase) from the main beam. For a DFMI signal with a (much) larger modulation index, the signal energy is spread over (many) more harmonic frequencies and also distributed differently which would further reduce their impact on the harmonics of the main signal. Figure 9.5 shows a simulated example ASD of a main signal and two ghost beams with different modulation indices.
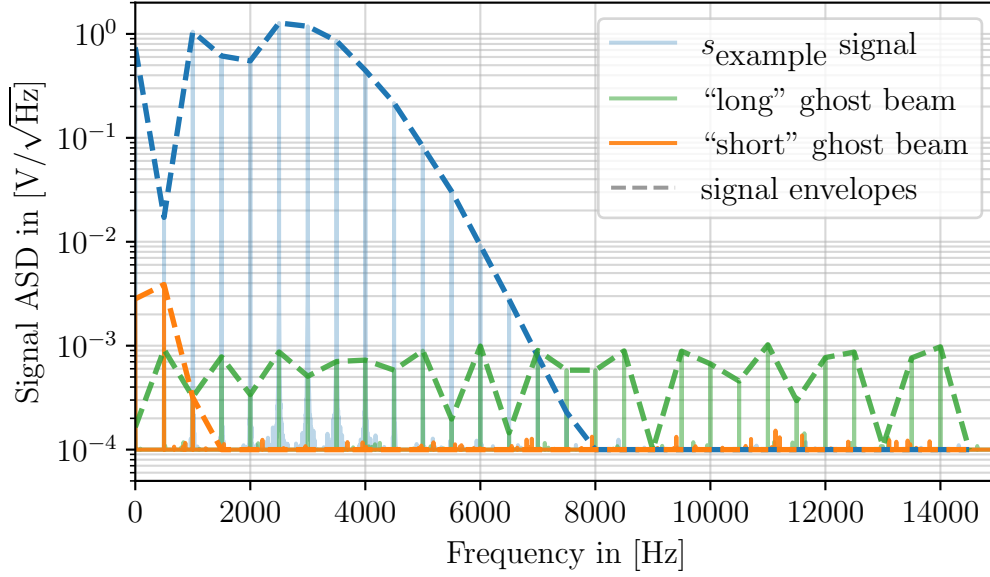
Figure 9.5: Plot of a example DFMI signal and two ghost beams with different modulation indices and random phases. The "short" ghost beam corresponds to a length (and modulation index) of $\times 0.1$ of the main signal and the "long" ghost beam corresponds to $\times 8$ the length of the main signal. For the "short" ghost beam all of its signal power is concentrated at the first two harmonics leading to a relatively large error there, but having almost no impact at all other harmonics. The "long" ghost beams signal power is distributed over more harmonics with a significant part of the ghost beam power localized at the higher harmonics where the main signal has almost no signal power.

For ghost beams only traveling within the fixed optical setup (and not to a potentially moving target), the additional ghost beam signal would remain constant and could in theory be subtracted from the main signal if the ghost beam parameters are known. In previous publications, Gerberding and Isleif modified their fitting algorithm also fit smaller DFMI–type signals besides the main signal which corresponds in parts to these ghost beams. Similarly, if the modulation index of the ghost beams is sufficiently different from the main signal, using only specific signal harmonics that carry a lot of energy of the ghost beam and very little of the main signal, the analytic readout algorithm can be used to calculate the ghost beam parameters. (For the signals shown in Figure 9.5, one could for example use only the harmonics around 10 kHz to derive the signal parameters for the "long" ghost beam). The ghost beam can then be subtracted from the measured signal before calculating the main signal parameters. The implementation of such a ghost beam mitigation scheme is currently being investigated and might be included into the analytic readout algorithm at a later stage.

9.4   Estimated phase readout limit

The results of the preciously estimated noise limits are shown together in Figure 9.6. It should be noted that the shown DFMI–type noise limits of P_{RAM}, P_{ghost}, P_{approx} and P_{THD} correspond to rough upper limits where their corresponding DFMI–type signals share the same parameters as the main signal and only vary in their (random) phase and amplitude. In this case they couple strongest into the main beam leading to the largest errors. The values presented here serve as guidelines

for i.e. approximating the noise contributions from different sources relative to each other. Their actual contribution to the overall readout is likely smaller than the here presented upper estimate.
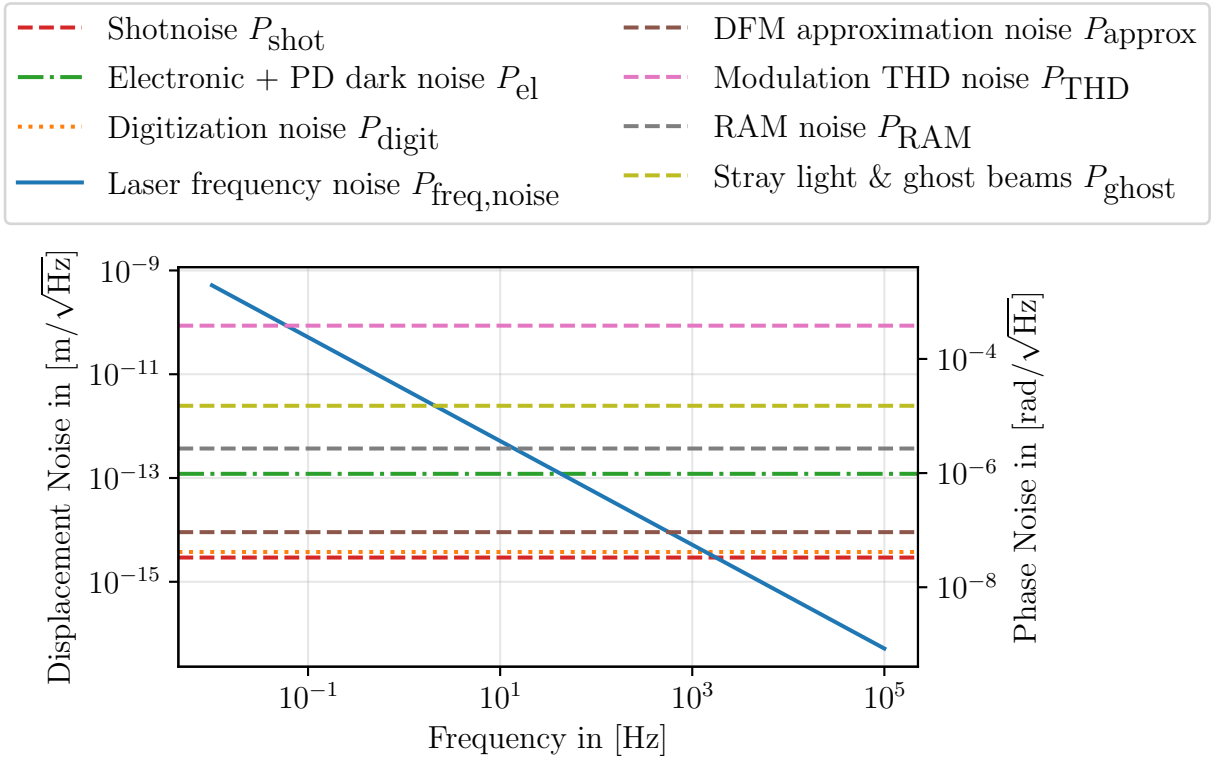

Figure 9.6: Plot of the estimated lower limit for the length (phase) readout of my experimental DFMI setup.

Part IV

Implementation of a DFMI phasemeter

Chapter 10

DFMI ARA–phasemeter setup

After testing the analytic readout algorithm for DFMI signals on simulated signals (shown in section 4.9), and reaching nearly optimal performance (given by the CRB derived in chapter 8), the next steps are the implementation of the analytic readout algorithm into modern hardware that can be used in gravitational wave detectors and to verify its performance in a physical setup.

As chapter overview, Figure 10.1 shows the main steps of the experimental setup I used to verify and test the readout algorithm.
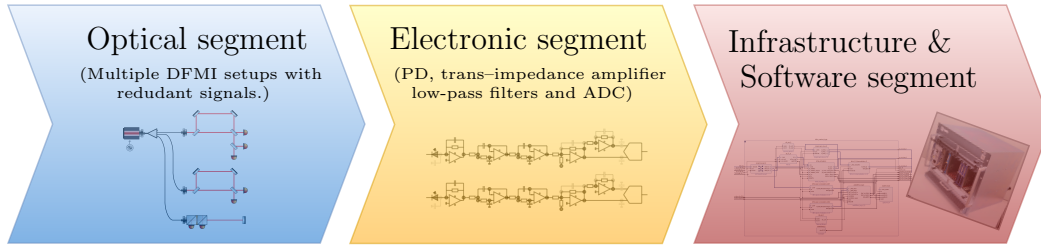
Figure 10.1: Sketch of the "readout chain" of the experimental setup discussed in this chapter.

10.1 Optical segment

Figure 10.2 shows an overview of the full optical setup used in my phasemeter experiment which was built up together with my colleague M. Mahesh. It consists of two Mach–Zehnder setups and one COBRI prototype with ideally all setups measuring the same distance of $\Delta L = 10\,\text{cm}$. This corresponds to the distance DFMI sensors and target mirrors could be installed inside the suspensions of GWDs, where they would act as local displacement sensors. The experiment aims to create a DFMI signal similar to one that would be measured there, allowing me to investigate the performance of the analytic readout algorithm for such cases in a more realistic environment. Instead of only a single interferometer, the setup includes three individual setups, supplied by the same laser source. This allows one to "lock" the laser frequency to the measured length of one of the two setups and additionally compare the COBRI prototype to a regular free beam setup. I.e. due to its short second arm, the COBRI signal is expected to show less "DFMI approximation noise," but more noise from ghost beams due to the number of additional plane surfaces.
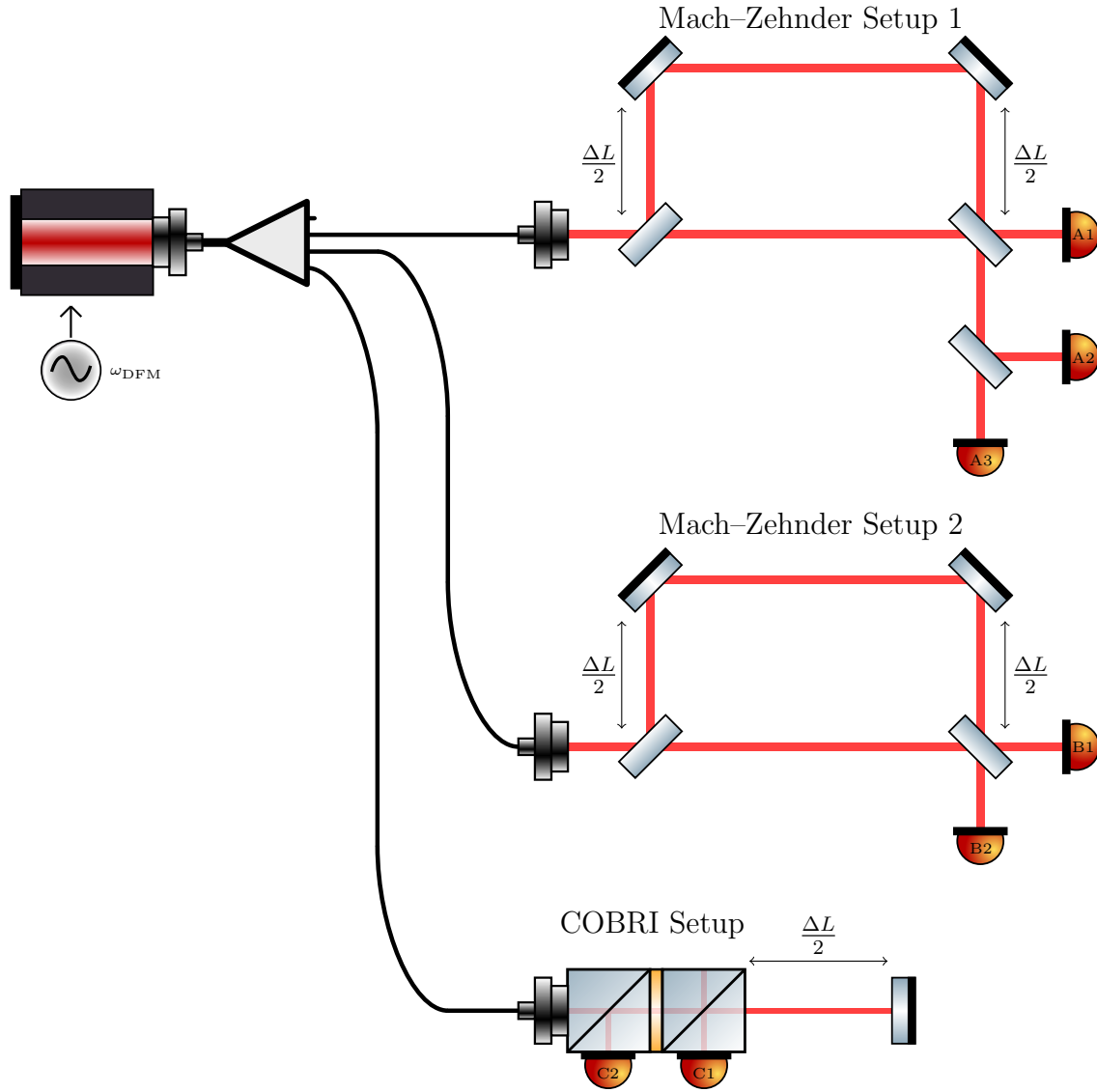
Figure 10.2: Schematic overview of the optical setup used for the DFMI phasemeter test. It consists of two Mach–Zehnder setups and a COBRI prototype all set up with roughly $\Delta L = 10\,\mathrm{cm}$ path-length difference between their two interfering beams. For the COBRI this corresponds to a target distance of $5\,\mathrm{cm}$ which leads for the beam traveling back and forth to an absolute path length difference of $10\,\mathrm{cm}$. All shown beam-splitter should split the beams 50/50 in both directions.

The laser is tuned to a modulation depth and modulation frequency of

$$\Delta\omega = 3.37\,\mathrm{GHz} \quad \text{and} \quad f_m = 500\,\mathrm{Hz}$$
$$\implies \omega_{\mathrm{DFM}}(t) = 3.37\,\mathrm{GHz} \cdot \sin(2\pi \cdot 500\,\mathrm{Hz} \cdot t)\,.$$

Together with the distance $\Delta L = 10\,\mathrm{cm}$ the DFMI parameters should be similar to the previous simulations using the same values as written earlier in Table 3.1.

Measuring all outgoing beams and also splitting one of the signals again allows for a more granular analysis of some of the occurring noise contributions. Taking the difference between the two quadratures of one setup will subtract the signal and all common noises sources in both signal quadratures and leave only the channel

specific noises (photodiode noise, electronic noise and algorithm noise). This is also referred to as zero–measurement as it removes the "signal" and the data will only show non–common noise contributions.

Similarly, splitting one of the quadratures again into two separate photodiodes (i.e. photodiodes A2 and A3 in Figure 10.2) allows one to isolate noise sources that depend on the signal power / amplitude (which is identical for PDs A2 and A3, but differs for A1 and A2). E.g. when the signal power in PD A1 goes high, it goes low in PDs A2 and A3.

The second Mach–Zehnder setup serves mainly as reference to remove the laser frequency. It does allow, however, to estimate errors due to misalignment and the influence of different components (like i.e., slight variations in the non–perfect 50/50 beam splitters).

The used COBRI (shown earlier in section 3.3.1) is a prototype developed in Hamburg and build by M. Mahesh using custom optics with a quasi monolithic component (QMC) size of 5mm x 5mm x 11mm. Its signal is expected to contain stronger ghost beams due to the plain aligned surfaces of its QMC, compared to a free beam DFMI setups. For the interferometric signal, this means that the short arm (L_{short} in Figure 3.5) has a length of $\sim 2.5\,\text{mm}$ which leads together with a modulation frequency of $f_m = 500\,\text{Hz}$ to $\psi_L \approx 10^{-6}\,\text{rad}$ for the DFMI approximation noise.

At the end of the optical segment, the interfering beams with a respective power of $\sim 2\,\text{mW}$ are converted into an electric current by photodiodes with a responsivity of $R_{\text{PD}} \approx 1\,\text{A/W}$ and continue in the electronic segment.

10.2 Electronic segment (Amplifier, Filter and ADCs)

In the electric segment the small currents coming from the PDs are amplified and filtered by a custom electronic circuit shown in Figures 10.3 and 10.4.
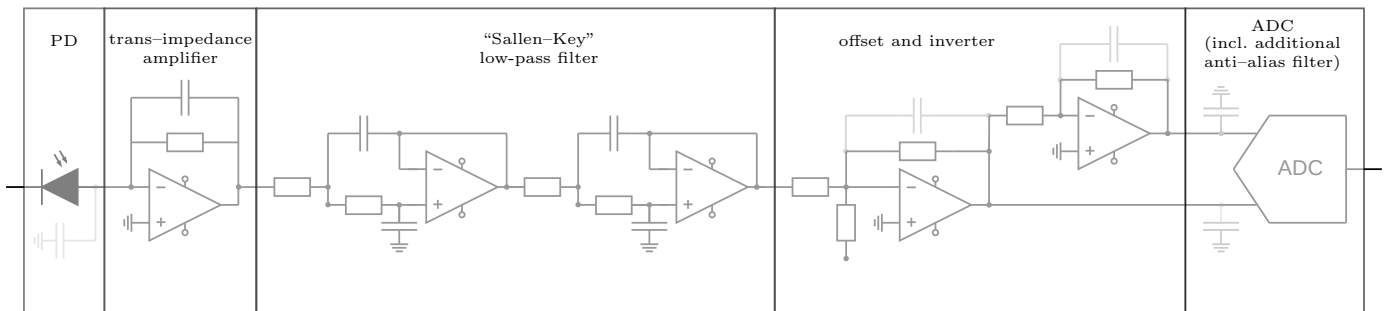
Figure 10.3: Overview a single channel of the electronic segment of the readout chain. The greyed out capacitance at the start and end of the circuit correspond to the connected cables.

A more detailed discussion of used PDs, the build electronics and their transfer function are shown in Appendix F. Its main purpose is to amplify the small photo current ($\sim 2\,\text{mA}$) and convert it into a differential signal measured by the used differential ADCs ($\sim \pm 5\,\text{V}$). The additional low-pass filter helps to mitigate parasitic
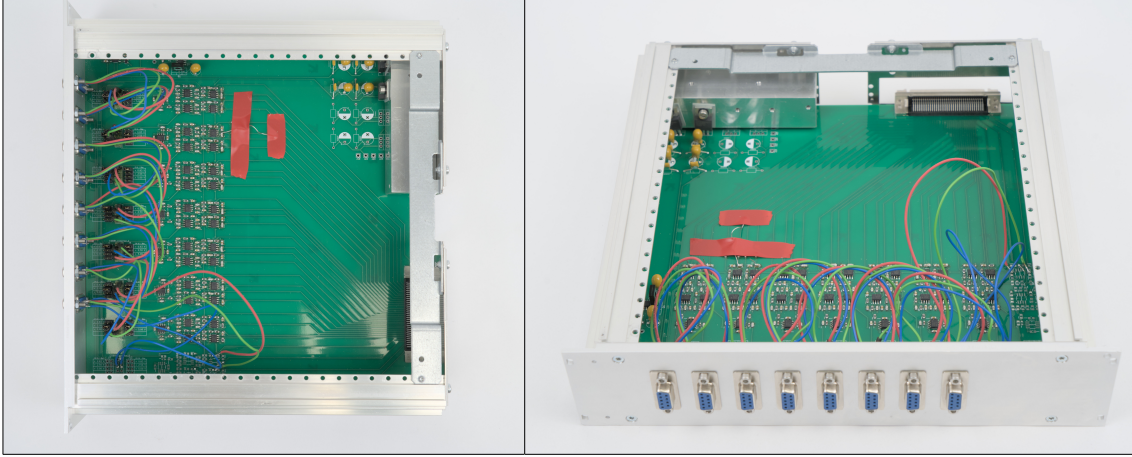
Figure 10.4: Picture of the electronic amplifier I build consisting of 16 channels with the schematic as shown in Figure 10.3. The front is made up of 8 D-Sub-9 Connectors, each of which connects to two channels. The back contains a power connector to the GEO infrastructure and a half-pin D-Sub-39 connector as connection to my ADC module (the "ACQ425ELF").

electronic noise and acts as further antialiasing filter for the ADC.

The measured transfer function of the amplifier and filter circuit can be seen in Figure 10.5. It shows both a phase delay and a small attenuation of the amplitude for different harmonic frequencies. To reach a high precision readout, this delay and attenuation must be compensated by applying the appropriate factors during demodulation and subtracting the delay when calculating the phase within the analytic readout algorithm.

I.e. the phase of the individual n'th harmonic as written in (4.2) is proportional to $\sim \varphi + n\psi + \varphi_{\text{el,delay}}(nf_m)$. The easiest way to remove the phase delay and get a high precision absolute phase readout, is to subtract the linear factor $(36.06\,\mu\text{rad} \times 500\,\text{Hz})$ from the ψ coefficient when calculating the c_n and subtract the constant offset $(214.95\,\mu\text{rad})$ from the calculated $\hat{\varphi}_{\text{estimate}}$. In Heinzel et al. [54] the phase delay from the electronics was similarly measured and subtracted from the harmonics. When using $\hat{\psi}_{\text{estimate}}$ as calculated in 4.5, it already includes the linear component of $\varphi_{\text{el,delay}}$ and no further corrections are necessary.

10.3   Hardware and Software infrastructure

At the end of the electronic segment, the signal is digitized by a 16-bit "LTC2380-16" ADC and further processed in a "Zynq7000" FPGA where my implementation of the analytic readout algorithm in VHDL and python calculates the DFMI signal parameters.

The ADCs and the computing hardware used are integrated components in a larger "MTCA" infrastructure. A significant part of my Ph.D. studies revolved around understanding and commissioning of this infrastructure, and specifically the usage of "FPGAs" and the implementation of the readout algorithm thereon. "Field Programmable Gate Array (FPGA)" are specialized chips emulating the behavior of
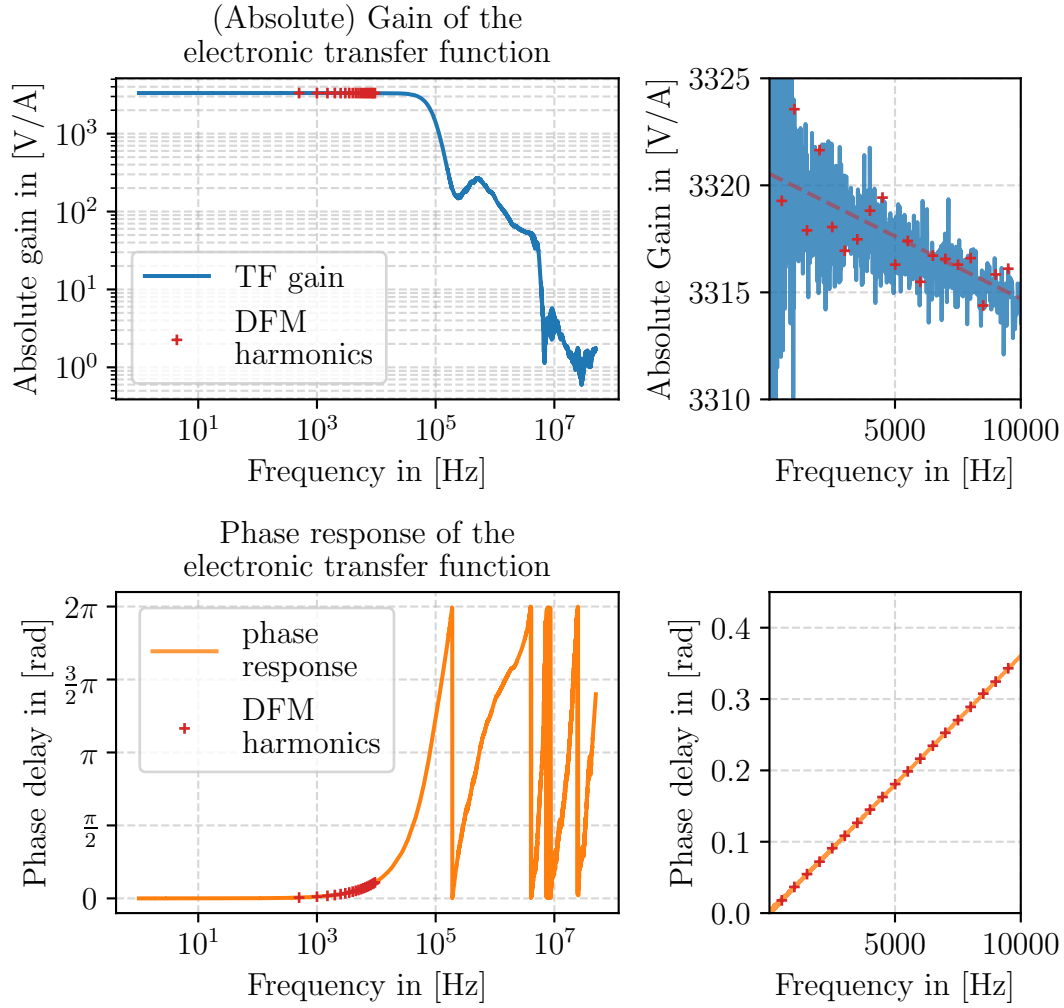
Figure 10.5: Transfer function of the complete electronic TIA-amplifier and filter segment, measured with a "Moku:Lab." The left (upper and lower) plots show the full measured transfer function gain and phase. The right plots show a "zoomed in" picture in the frequency range of the DFM harmonics. In the frequency region of the DFM harmonics, the transfer function H_{EL} of the circuit can be approximated by $|H_{EL}| \approx -5.8410^{-4} \cdot f/\text{Hz} + 3320.54$ and a phase delay of $\arg(H_{EL}) = \varphi_{\text{el,delay}}(f) \approx 36.06\,\mu\text{rad} \cdot f/\text{Hz} + 214.95\,\mu\text{rad}$, shown as dashed line.

logic gate level hardware which allows for the design of highly specialized ("hardware") function blocks yielding two main advantages for the readout algorithm:

1. **A high readout speed (due to parallelization and specialized hardware modules)**
   In the DFM readout algorithm a potentially large number of harmonics need to be demodulated and filtered. Especially the demodulation is a computationally expensive task for which many DSPs or MCUs have dedicated hardware modules. Implementing the demodulation and the filtering of the harmonics in an FPGA allows for a very efficient calculation of analytic readout algorithm within a few µs, faster than on comparable CPUs or DSPs.

2. **Stable synchronous timing for operating control systems**
   Due to the planned implementation of various control loops to regulate i.e.,

the laser frequency and amplitude modulation, using an FPGA as custom controller allows one to regulate multiple in- and outputs synchronous to the same clock, leading to short delays and a wide control bandwidth.

The infrastructure I used is based on the "Micro Telecommunication and Computing Architecture (MTCA)" open industry standard. MTCA is a modular Infrastructure standard specifying "mechanical, electrical, thermal and management [hardware properties]" used for data processing and control systems in large research facilities and industry [75]. It is partially developed by DESY and extensively used in i.e., their European XFEL. We also currently investigate the option to implement the "Control and Design System (CDS)" used in LIGO into our MTCA hardware which would greatly simplify the implementation of the DFMI readout into gravitational wave detectors that use LIGOs CDS. In my experiment, using MTCA hardware allows me to use different digital computing modules (a "AM-G6x/msd" (CPU) and a "Z7IO" (FPGA) module) and analog modules (a "ACQ420FMC" as ADC module and a "AO420FMC" as DAC module), from different manufacturers. A detailed overview and description of the components I used can be found in Appendix G.1 and a picture of the MTCA crate containing all of the mentioned modules is seen in Figure 10.6.

After digitization of the electronic signal by the "ACQ420FMC" module, the demodulation and calculation of the I_n and Q_n coefficients is done within the FPGA fabric of the "Z7IO" module. Figure 10.7 shows a schematic overview of the implemented I-Q-demodulation. Every channel (every PD signal digitized by a separate ADC) is fed to ~ 16 identical I-Q-demodulation blocks (one for each DFM harmonic) which calculate the I_n and Q_n coefficients in parallel (for $n = 1, \ldots, 16$). The operation of an individual I-Q-demodulation block corresponds to a regular "phase locked loop" with the modulation frequency as input parameter. Integrating the given frequency value, and applying a dither to mitigate truncation noise, yields a phase value $\sim (n\omega_m t)$. This phase value is used as input value for a lookup table (LUT) which contains sine and cosine values in a given (here: 16-bit) range, yielding output values $\sim \sin(n\omega_m t)$. (Together these individual components correspond to a numerically controlled oscillator (NCO) generating an output signal of $\sin(n\omega_m t)$). This signal is then mixed with the input signal and low-pass filtered (as described in section 4.4) to calculate the I_n and Q_n coefficients for all relevant harmonics.

The remaining computational parts of the analytic readout algorithm are currently implemented in a python script running on the "AMG6" CPU. A single run with the default parameters (Table 3.1) running exclusively on the CPU takes $\approx 10\,\text{ms}$, which would lead to a similar readout frequency of $\sim 100\,\text{Hz}$ of older publications by Isleif et al. using Heinzel's fitting algorithm [61]. Using the parallelized demodulation in the FPGA fabric, and the resulting I_n and Q_n coefficients, reduces the computational time to $0.55\,\text{ms}$ per run improving the readout speed and thus the corresponding control bandwidth by a factor ~ 20. Implementing the remaining calculation steps into the FPGA fabric as well should reduce the readout computation time even further.

Figure 10.6: Pictures of a partially filled "NATIVE-R9" (MTCA) Crate. The left picture shows the front side and the right the back side of the crate. From left to right of the front side one can see: the PSM, the MCH, a "CPU" AMC module, two empty slots, the "Z7IO" AMC module I use, 5 more open slots, and then 3 more Z7IO AMC modules. In the rear / back side, from left to right are: the PSM, the MCH, empty slots, and a single RTM connected to the Z7IO AMC that I use.
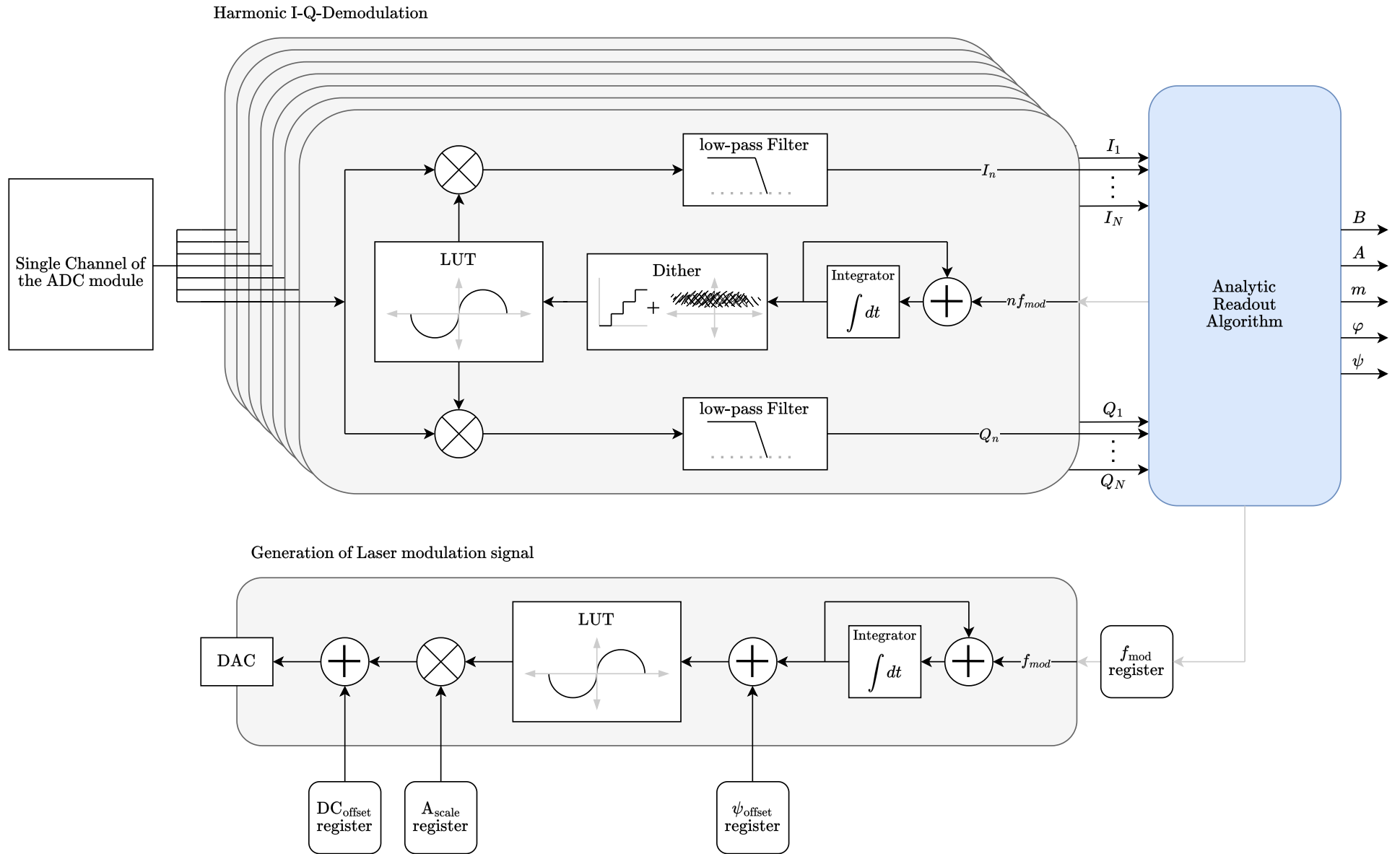
Harmonic I-Q-Demodulation

Generation of Laser modulation signal

Figure 10.7: Schematic overview of the currently implemented parts of the "Analytic Readout Algorithm" in the FPGA fabric.

10.4   ARA–Phasemeter performance
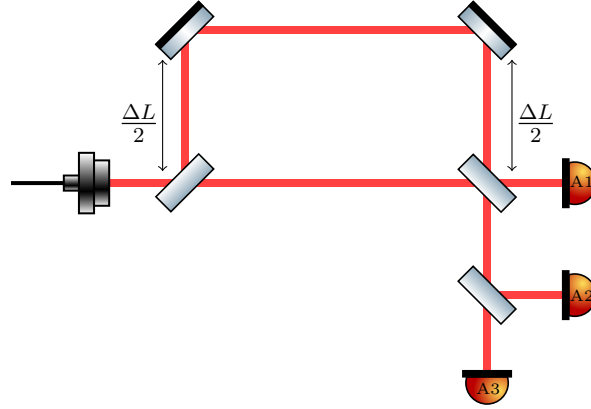
Figure 10.8: Sketch of the free beam Mach–Zehnder setup, operated in air, used to test the analytic readout algorithm.

Figure 10.9 shows the (ARA) φ phase readout for one of the free beam Mach–Zehnder setups (shown in Figure 10.8) with the three separate PD outputs A1, A2 & A3. PD A1 measures one of the two signal quadratures and PDs A2 and A3 measure the other quadrature, each one with their own separate amplifier electronics and readout channel.

The direct readout of A1, A2 and A3 is dominated by some kind of noise that is common to all channels which has yet to be determined. It could be related to real optical path length changes due to air fluctuations or ground motion coupling in the mechanical components moving them by $\mathcal{O}(100\,\mathrm{nm})$. As similar DFMI setups in air have achieved a significantly higher precision, there is likely another noise sources limiting the readout here, From the noise models derived in section 9.3, ghost beams, residual amplitude modulation, and the laser's harmonics distortion are all possible candidates that can be present at these levels and would be common to all channels.

The ASD of the phase differences $\varphi_{\mathrm{A1}} - \varphi_{\mathrm{A2}}$ (different quadratures, zero measurement) and $\varphi_{\mathrm{A2}} - \varphi_{\mathrm{A3}}$ (same quadrature, π measurement) removes this common noise and reveals a lower noise floor almost at the level of the laser frequency noise (up to 10 Hz) which couples directly into the phase readout. The white noise floor above 10 Hz as well as the slightly higher noise for the π measurement hint however at an additional noise term that limits the readout here. The main difference between zero and π measurement is that any amplitude noise would affect the zero measurement more strongly than the π measurement (where both signals experience the exact same amplitude noise). The fact that the noise floor of the π measurement is actually higher (and not flat) hints to an additional non–linear noise term where the smaller power of the A2 and A3 signals could possibly factor in the effective higher noise (as A2 and A3 only have around half the power of A1 signal since their quadrature is split an additional time).

The COBRI prototype yields a similar noise floor as the free beam setup for the "direct" single channel readout, as shown in Figure 10.10. Subtracting the two

COBRI channels from another leads however to a $\times 10$ higher noise compared to the free beam's zero measurement.
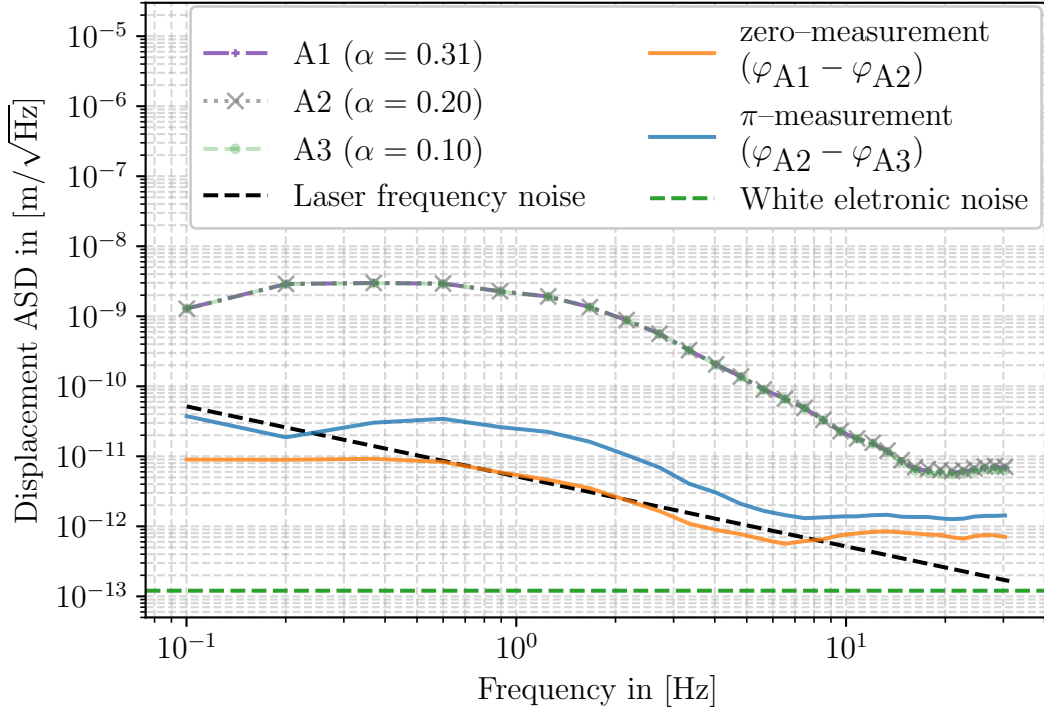

Figure 10.9: Phase Readout for the free beam Mach–Zehnder Setup with three PD outputs (A1, A2 & A3) and the difference between the different phase readouts. The $\alpha = \kappa A/\text{ADC-range}$ value mentioned in the legend corresponds to the "filling factor" of the ADC (κ as optical contrast).
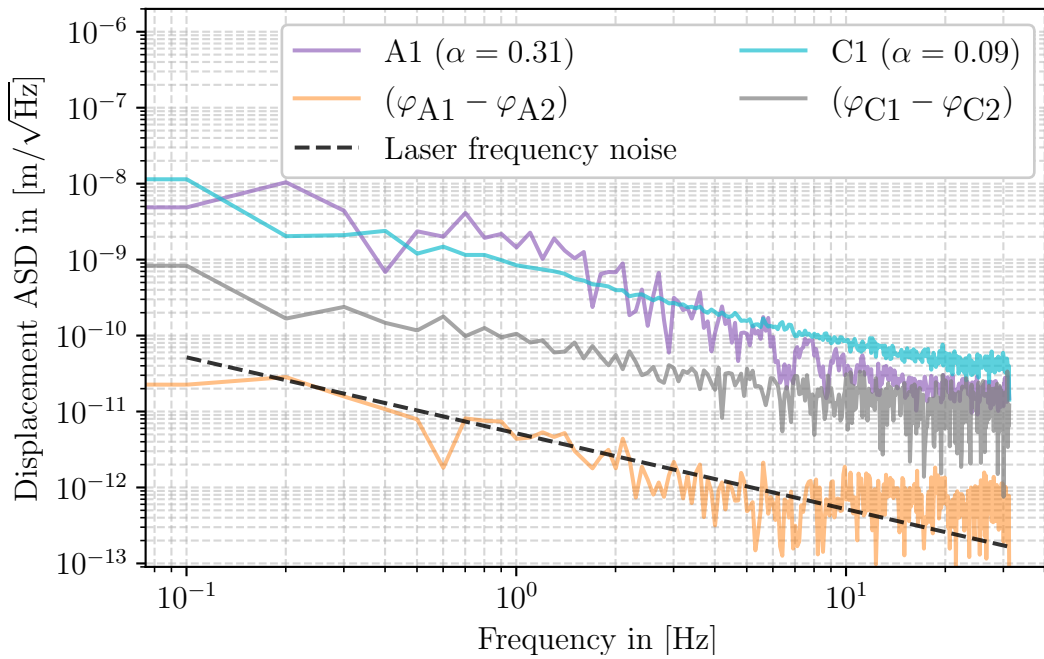

Figure 10.10: Phase Readout for the COBRI prototype setup.

Besides the worse contrast of the COBRI prototype in this particular setup, the signal ASD as shown in Figure 10.11 reveals significant signal energy in the higher harmonics ($> 15 \cdot f_m$) above the noise floor, with orders of magnitude more for the COBRI compared to the free beam signal. This indicates the presence of additional ghost beams which have a longer path length (larger modulation index) leading to more signal energy at higher frequency harmonics and which could explain parts of the excess noise of the COBRI measurement. Not specifically shown here is another effect of the used COBRI prototype: During its assembly and initial testing, M. Mahesh discovered a "signal envelope," a signal around the modulation frequency correlated to the polarization of the laser beam, that distorts the DFMI signal and creates a characteristic enveloped signal, different from other free beam DFMI setups tested so far. Its origin and effect are still being investigated, and we consider it another likely cause for the excess noise observed from the COBRI.
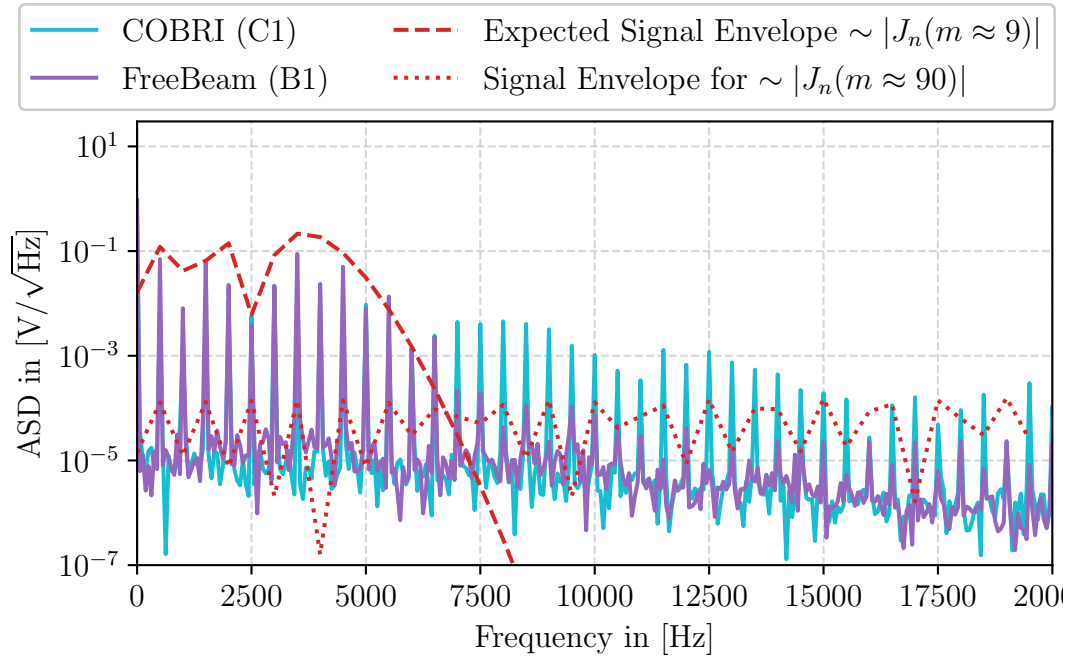
Figure 10.11: Signal ASD for a single FreeBeam and a single COBRI channel. Above frequencies of $15 \cdot f_m = 7500\,\text{Hz}$ there is significant signal energy above the expected level (red dashed line). An additional DFMI signal (ghost beam) with i.e. a $\times 10$ longer path length difference (dotted red line) could explain this excess signal.

As of now the setup is still several orders of magnitude less precise than the previously achieved noise-floor of $\sim 230\,\text{fm}/\sqrt{\text{Hz}}$. The next steps would be to improve upon the setup by increasing the amplifier gains and the optical contrast to improve the ADC filling-factor α, repeat the measurement in a vacuum environment to remove noise due to air fluctuations, and stabilize the laser frequency (and power) by locking it to one of the two free beam interferometers. For the COBRI, additional baffles and software schemes to subtract static ghost beams from the signal could also help to improve the readout. Newly ordered optics aimed to reduce the undesired polarization effects and the resulting signal envelope should also help to remove the excess noise in the next COBRI prototype.

Chapter 11

Conclusion

Motivated by a limited sensitivity of current ground based gravitational wave detectors around 3 Hz due to alignment and control noise, I present and discuss the use of deep–frequency modulation interferometry (DFMI) based sensors for the local displacement readout of the optics in these detectors to improve upon these limitations.

The readout algorithm (by Heinzel et al. [54]) used in previous DFMI experiments to calculate the distance information from a measured DFMI signal relied on a non–linear fit algorithm which was computationally expensive and limited the readout bandwidth to ~ 100 Hz. The two main results of my Ph.D. work presented in this thesis are a new and potentially faster analytic readout algorithm for DFMI signals presented in chapter 4, and the noise analysis of DFMI signals using the CRB presented in chapter 8.

The "analytic readout algorithm (ARA)" I developed does not iteratively fit a model signal but instead calculates the DFMI signal parameters by arithmetic calculations from the measured signals. While a first implementation has already shown an increase of the readout speed by a factor of ~ 20, its non–iterative structure allows for its implementation within an FPGA can yield an even higher readout speed up to the used sampling frequency of ~ 1 MHz.

Presented as an extension of the ARA, I also analyzed the changes of DFMI signals for a fast moving target (dynamic DFMI signals). By using the dynamic extension to the ARA presented here, I am able to improve the precision in some extreme (purely linear moving) cases by a factor of 10^3 compared to a readout where the changes to the ideal DFMI signal are ignored.

With my analysis of how noise affects the DFMI readout, I could show that the analytic readout algorithm performs close to optimal precision, given by the CRB, for the intended use case for a local displacement readout.

The analysis of the Cramér-Rao bound for DFMI signals presented in this thesis provides new insight into how additive white noise affects the precision of the different DFMI parameters and allows for an a priori prediction of the achievable precision for known noise sources in an experiment.

By calculating the CRB for the response of an optical cavity to a phase modulated signal, I was also able to show that the amount of Fisher information in the signal, and thus the precision of the length readout, can be significantly increased by com-

bining a DFMI signal with an optical cavity. This served as initial motivation for the presented ReDFMI concept, which allows for an even higher precision readout for local displacement sensors in the future.

Lastly, the experimental implementation of the ARA presented in Part III demonstrates the functionality of the algorithm in a real experiment, and it incorporates the first ever readout of a COBRI sensor prototype. A significant part of my Ph.D. work involved the commissioning of the hardware and computing infrastructure that was used to operate this first experiment, and while it is only briefly mentioned in this thesis (and mainly in Appendix G), it will help to pinpoint sources of excess noise and allow for several improvements to the setup in order to reach the theoretical optimum with sub-picometer precision in the future.

Besides these local improvements, steps are being taken to implement a LIGO compatible CDS into the here used (MTCA) hardware which would allow for an easy implementation of the COBRI and ARA into i.e. the LIGO detectors.
While the final decision, which local displacement sensor to use in LIGOs future has not been made yet, both COBRIs and the here developed ARA are major contenders for an implementation in LIGO post-O5 upgrades within this decade, which would significantly improve their sensitivity in the low frequency regime and thus allow for an earlier detection and a higher SNR of the detected gravitational waves.

Bibliography

[1] Christina Sormani. "A Two-Part Feature: The Mathematics of Gravitational Waves". In: *Notices of the American Mathematical Society* 64.7 (Aug. 1, 2017), pp. 684–685. ISSN: 0002-9920, 1088-9477. DOI: `10.1090/noti1551`.

[2] Wolfgang Trageser, ed. *Das Relativitätsprinzip: Eine Sammlung von Originalarbeiten zur Relativitätstheorie Einsteins.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2018. ISBN: 978-3-662-57410-2 978-3-662-57411-9. DOI: `10.1007/978-3-662-57411-9`.

[3] Subhendra Mohanty. "Field Theory of Linearised Gravity". In: *Gravitational Waves from a Quantum Field Theory Perspective.* Ed. by Subhendra Mohanty. Cham: Springer International Publishing, 2023, pp. 39–71. ISBN: 978-3-031-23770-6. DOI: `10.1007/978-3-031-23770-6_3`.

[4] A. Einstein and N. Rosen. "On gravitational waves". In: *Journal of the Franklin Institute* 223.1 (Jan. 1, 1937), pp. 43–54. ISSN: 0016-0032. DOI: `10.1016/S0016-0032(37)90583-0`.

[5] "Gravitational waves in general relativity III. Exact plane waves". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 251.1267 (June 23, 1959), pp. 519–533. ISSN: 2053-9169. DOI: `10.1098/rspa.1959.0124`.

[6] H. Bondi. "Gravitational Waves in General Relativity". In: *Nature* 186.4724 (May 1960). Publisher: Nature Publishing Group, pp. 535–535. ISSN: 1476-4687. DOI: `10.1038/186535a0`.

[7] H. Bondi et al. "Gravitational waves in general relativity III. Exact plane waves". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 251.1267 (Jan. 1997). Publisher: Royal Society, pp. 519–533. DOI: `10.1098/rspa.1959.0124`.

[8] Andrzej Trautman. "Radiation and Boundary Conditions in the Theory of Gravitation". In: arXiv:1604.03145 (Apr. 11, 2016). DOI: `10.48550/arXiv.1604.03145`.

[9] Andrzej Trautman. "Boundary Conditions at Infinity for Physical Theories". In: arXiv:1604.03144 (Apr. 11, 2016). DOI: `10.48550/arXiv.1604.03144`.

[10] F. A. E. Pirani. "Invariant Formulation of Gravitational Radiation Theory". In: *Physical Review* 105.3 (Feb. 1, 1957). Publisher: American Physical Society, pp. 1089–1099. DOI: `10.1103/PhysRev.105.1089`.

[11] Ivor Robinson and A. Trautman. "Spherical Gravitational Waves". In: *Physical Review Letters* 4.8 (Apr. 15, 1960), pp. 431–432. ISSN: 0031-9007. DOI: `10.1103/PhysRevLett.4.431`.

[12] Michele Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments.* Oxford University Press, Oct. 4, 2007. ISBN: 978-0-19-857074-5. DOI: `10.1093/acprof:oso/9780198570745.001.0001`.

[13] Cosimo Bambi, Stavros Katsanevas, and Konstantinos D. Kokkotas, eds. *Handbook of Gravitational Wave Astronomy.* Singapore: Springer Nature Singapore, 2022. ISBN: 9789811643057 9789811643064. DOI: `10.1007/978-981-16-4306-4`.

[14] Charles W. Misner, Kip S. Thorne, and John Archibald Wheeler. *Gravitation.* 27. printing. New York, NY: Freeman, 2008. 1279 pp. ISBN: 978-0-7167-0334-1 978-0-7167-0344-0.

[15] Russell A Hulse and Joseph H Taylor. "Discovery of a pulsar in a binary system". In: *The Astrophysical Journal* 195 (1975), pp. L51–L53.

[16] Joel M Weisberg, Joseph H Taylor, and Lee A Fowler. "Gravitational waves from an orbiting pulsar". In: *Scientific American* 245 (1981), pp. 74–82.

[17] Peter R. Saulson. *Fundamentals of Interferometric Gravitational Wave Detectors.* 2nd ed. Singapore: World Scientific Publishing Co. Pte. Ltd., Feb. 2017. 336 pp. ISBN: 978-981-314-307-4.

[18] R M Shannon. *Gravitational-Wave Limits from Pulsar Timing Constrain Supermassive Black Hole Evolution.* DOI: `10.1126/science.1238012`. URL: `https://www.science.org/doi/10.1126/science.1238012` (visited on 02/03/2025).

[19] Luisa Bonolis and Juan-Andres Leon. "Gravitational-Wave Research as an Emerging Field in the Max Planck Society: The Long Roots of GEO600 and of the Albert Einstein Institute". In: *The Renaissance of General Relativity in Context.* Ed. by Alexander S. Blum, Roberto Lalli, and Jürgen Renn. Einstein Studies. Cham: Springer International Publishing, 2020, pp. 285–361. ISBN: 978-3-030-50754-1. DOI: `10.1007/978-3-030-50754-1_9`.

[20] J. Weber. "Detection and Generation of Gravitational Waves". In: *Physical Review* 117.1 (Jan. 1, 1960), pp. 306–313. ISSN: 0031-899X. DOI: `10.1103/PhysRev.117.306`.

[21] Benno Willke et al. "The GEO 600 gravitational wave detector". In: *Classical and Quantum Gravity* 19.7 (2002), p. 1377.

[22] H. Grote et al. "First Long-Term Application of Squeezed States of Light in a Gravitational-Wave Observatory". In: *Physical Review Letters* 110.18 (May 1, 2013), p. 181101. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.110.181101`.

[23] J. Aasi et al. "Advanced LIGO". In: *Classical and Quantum Gravity* 32.7 (Mar. 2015). Publisher: IOP Publishing, p. 074001. ISSN: 0264-9381. DOI: `10.1088/0264-9381/32/7/074001`.

[24] S M Aston et al. "Update on quadruple suspension design for Advanced LIGO". In: *Classical and Quantum Gravity* 29.23 (Dec. 7, 2012), p. 235004. ISSN: 0264-9381, 1361-6382. DOI: `10.1088/0264-9381/29/23/235004`.

[25] B. P. Abbott et al. "Observation of Gravitational Waves from a Binary Black Hole Merger". In: *Physical Review Letters* 116.6 (Feb. 11, 2016), p. 061102. ISSN: 0031-9007, 1079-7114. DOI: `10.1103/PhysRevLett.116.061102`.

[26] S. Soni et al. *LIGO Detector Characterization in the first half of the fourth Observing run.* Sept. 4, 2024. DOI: `10.48550/arXiv.2409.02831`.

[27] Adele La Rana. "The Origins of Virgo and the Emergence of the International Gravitational Wave Community". In: *The Renaissance of General Relativity in Context.* Ed. by Alexander S. Blum, Roberto Lalli, and Jürgen Renn. Cham: Springer International Publishing, 2020, pp. 363–406. ISBN: 978-3-030-50754-1. DOI: `10.1007/978-3-030-50754-1_10`.

[28] Rich Abbott et al. "Open data from the first and second observing runs of Advanced LIGO and Advanced Virgo". In: *SoftwareX* 13 (Jan. 2021), p. 100658. ISSN: 23527110. DOI: `10.1016/j.softx.2021.100658`.

[29] T Akutsu et al. "Overview of KAGRA: Detector design and construction history". In: *Progress of Theoretical and Experimental Physics* 2021.5 (May 1, 2021). ISSN: 2050-3911. DOI: `10.1093/ptep/ptaa125`.

[30] Otto Markus. "Time-Delay Interferometry Simulations for the Laser Interferometer Space Antenna". In: (). DOI: `10.1088/0264-9381/29/20/205003`.

[31] LISA International Science Team. *LISA Unveiling a hidden Universe.* Assessment Study Report ESA/SRE(2011)3. ESA, 2011.

[32] G Hobbs et al. "The International Pulsar Timing Array project: using pulsars as a gravitational wave detector". In: *Classical and Quantum Gravity* 27.8 (2010). tex.ids: hobbs2010a, hobbs2010b publisher: IOP Publishing, p. 084013.

[33] Gabriella Agazie et al. "The NANOGrav 15 yr Data Set: Evidence for a Gravitational-wave Background". In: *The Astrophysical Journal Letters* (2023).

[34] Jan Harms et al. "Lunar Gravitational-wave Antenna". In: *The Astrophysical Journal* 910.1 (Mar. 2021). Publisher: American Astronomical Society, p. 1. ISSN: 0004-637X. DOI: `10.3847/1538-4357/abe5a7`.

[35] Parameswaran Ajith et al. "The Lunar Gravitational-wave Antenna: mission studies and science case". In: *Journal of Cosmology and Astroparticle Physics* 2025.1 (Jan. 1, 2025), p. 108. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2025/01/108`.

[36] B P Abbott et al. "Exploring the sensitivity of next generation gravitational wave detectors". In: *Classical and Quantum Gravity* 34.4 (Feb. 23, 2017), p. 044001. ISSN: 0264-9381, 1361-6382. DOI: `10.1088/1361-6382/aa51f4`.

[37] ET Steering Committee, ed. *Design Report Update 2020 for the Einstein Telescope.* 2020.

[38] M Punturo et al. "The Einstein Telescope: a third-generation gravitational wave observatory". In: *Classical and Quantum Gravity* 27.19 (2010), p. 194002.

[39] Craig Cahillane and Georgia Mansell. "Review of the Advanced LIGO Gravitational Wave Observatories Leading to Observing Run Four". In: *Galaxies* 10.1 (Feb. 2022). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, p. 36. ISSN: 2075-4434. DOI: `10.3390/galaxies10010036`.

[40] Michele Maggiore et al. "Science Case for the Einstein Telescope". In: *Journal of Cosmology and Astroparticle Physics* 2020.3 (Mar. 24, 2020), pp. 050–050. ISSN: 1475-7516. DOI: `10.1088/1475-7516/2020/03/050`.

[41] Craig Cahillane. *LIGO Hanford Noisebudget Interactive SVG*. URL: `https://ccahilla.github.io/lho_noisebudget.svg` (visited on 02/12/2025).

[42] Calum Iain Eachan Torrie. "DEVELOPMENT OF SUSPENSIONS FOR THE GEO 600 GRAVITATIONAL WAVE DETECTOR". PhD thesis. Glasgow: University of Glasgow, Nov. 17, 1999.

[43] Jennifer C. Driggers, Jan Harms, and Rana X. Adhikari. "Subtraction of Newtonian noise using optimized sensor arrays". In: *Physical Review D* 86.10 (Nov. 1, 2012). ISSN: 1550-7998, 1550-2368. DOI: `10.1103/PhysRevD.86.102001`.

[44] A. Buikema et al. "Sensitivity and performance of the Advanced LIGO detectors in the third observing run". In: *Physical Review D* 102.6 (Sept. 11, 2020). Publisher: American Physical Society, p. 062003. DOI: `10.1103/PhysRevD.102.062003`.

[45] Irene Fiori et al. "The Hunt for Environmental Noise in Virgo during the Third Observing Run". In: *Galaxies* 8.4 (Dec. 2020). Number: 4 Publisher: Multidisciplinary Digital Publishing Institute, p. 82. DOI: `10.3390/galaxies8040082`.

[46] *Gravitational Wave Interferometer Noise Calculator*. URL: `https://git.ligo.org/gwinc/pygwinc` (visited on 02/12/2025).

[47] Samuel James Cooper. "Breaking the Seismic Wall: How to Improve Gravitational Wave Detectors at Low Frequency". In: (), p. 207.

[48] J. van Dongen et al. "Reducing control noise in gravitational wave detectors with interferometric local damping of suspended optics". In: *Review of Scientific Instruments* 94.5 (May 23, 2023), p. 054501. ISSN: 0034-6748. DOI: `10.1063/5.0144865`.

[49] Jennifer Watchi et al. "Contributed Review: A review of compact interferometers". In: *Review of Scientific Instruments* 89.12 (Dec. 1, 2018). tex.ids= watchi2018a, watchi2018b publisher: American Institute of Physics, p. 121501. ISSN: 0034-6748. DOI: `10.1063/1.5052042`.

[50] Riccardo Maggiore et al. *Angular control noise in Advanced Virgo and implications for the Einstein Telescope*. Mar. 4, 2024.

[51] Hang Yu et al. "Prospects for Detecting Gravitational Waves at 5 Hz with Ground-Based Detectors". In: *Physical Review Letters* 120.14 (Apr. 6, 2018). tex.ids= yu2018 publisher: American Physical Society, p. 141102. DOI: `10.1103/PhysRevLett.120.141102`.

[52] Amit Singh Ubhi et al. "A compact six degree of freedom seismmometer with interferometric readout". In: *arXiv:2109.07880 [astro-ph, physics:physics]* (Sept. 16, 2021).

[53] Oliver Gerberding. "Deep frequency modulation interferometry". In: *Optics Express* 23.11 (2015). tex.ids: gerberding2015c, pp. 14753–14762. DOI: `10.1364/OE.23.014753`.

[54] Gerhard Heinzel et al. "Deep phase modulation interferometry". In: *Opt. Express* 18.18 (2010). tex.ids: heinzel2010a, pp. 19076–19086.

[55] VS Sudarshanam and K Srinivasan. "Linear readout of dynamic phase change in a fiber-optic homodyne interferometer". In: *Optics letters* 14.2 (1989), pp. 140–142.

[56] Osami Sasaki, Hirokazu Okazaki, and Makoto Sakai. "Sinusoidal phase modulating interferometer using the integrating-bucket method". In: *Applied optics* 26.6 (1987). tex.ids: sasaki1987a, sasaki1987b publisher: Optical Society of America, pp. 1089–1093.

[57] Peter J. de Groot. "Vibration in phase-shifting interferometry". In: *JOSA A* 12.2 (Feb. 1, 1995). Publisher: Optica Publishing Group, pp. 354–365. ISSN: 1520-8532. DOI: 10.1364/JOSAA.12.000354.

[58] Thomas Kissinger, Thomas OH Charrett, and Ralph P Tatam. "Range-resolved interferometric signal processing using sinusoidal optical frequency modulation". In: *Opt. Express* 23.7 (2015). tex.ids: kissinger2015a, pp. 9415–9431.

[59] Katharina-Sophie Isleif et al. "Experimental demonstration of deep frequency modulation interferometry". In: *Opt. Express* 24.2 (2016). tex.ids: isleif2016c, pp. 1676–1684. DOI: 10.1364/OE.24.001676.

[60] Thomas S. Schwarze. "Phase extraction for laser interferometry in space : phase readout schemes and optical testing". tex.ids: schwarze2018a. doctoralThesis. Hannover : Institutionelles Repositorium der Leibniz Universität Hannover, 2018. DOI: http://dx.doi.org/10.15488/4233.

[61] Katharina-Sophie Isleif et al. "Compact Multifringe Interferometry with Subpicometer Precision". In: *Physical Review Applied* 12.3 (Sept. 13, 2019), p. 034025. DOI: 10.1103/PhysRevApplied.12.034025.

[62] Yichao Yang et al. "Single-Element Dual-Interferometer for Precision Inertial Sensing". In: *Sensors* 20.17 (Jan. 2020). Number: 17 Publisher: Multidisciplinary Digital Publishing Institute, p. 4986. DOI: 10.3390/s20174986.

[63] Oliver Gerberding and Katharina-Sophie Isleif. "Ghost Beam Suppression in Deep Frequency Modulation Interferometry for Compact On-Axis Optical Heads". In: *Sensors* 21.5 (Jan. 2021). Number: 5 Publisher: Multidisciplinary Digital Publishing Institute, p. 1708. DOI: 10.3390/s21051708.

[64] Wanda Vossius. "Development of a Compact Balanced Readout Interferometer using Deep Frequency Modulation". Master's Thesis. DE: University of Hamburg, July 7, 2023. 82 pp.

[65] Jiri Smetana et al. "Compact Michelson interferometers with subpicometer sensitivity". In: (). tex.ids= smetana2022 publisher: American Physical Society, p. 5.

[66] Jiri Smetana et al. "Nonlinearities in Fringe-Counting Compact Michelson Interferometers". In: *Sensors* 23.17 (Jan. 2023). Number: 17 Publisher: Multidisciplinary Digital Publishing Institute, p. 7526. ISSN: 1424-8220. DOI: 10.3390/s23177526.

[67] Tobias Eckhardt and Oliver Gerberding. "Noise Limitations in Multi-Fringe Readout of Laser Interferometers and Resonators". In: *Metrology* 2.1 (Mar. 2022). Number: 1 Publisher: Multidisciplinary Digital Publishing Institute, pp. 98–113. ISSN: 2673-8244. DOI: 10.3390/metrology2010007.

[68] Tobias Eckhardt and Oliver Gerberding. "An analytic, efficient and optimal readout algorithm for compact interferometers based on deep frequency modulation". In: *Scientific Reports* 14.1 (Sept. 23, 2024). Publisher: Nature Publishing Group, p. 21988. ISSN: 2045-2322. DOI: 10.1038/s41598-024-70392-9.

[69] Mert Turkol. *mturkol/lpsd*. original-date: 2020-02-29T17:56:21Z. Sept. 16, 2024.

[70] Norman Hodgson and Horst Weber. *Laser resonators and beam propagation*. 2nd ed. Springer, 2005. ISBN: 0-387-25110-3.

[71] Alan V Oppenheim and George C Verghese. "Signals, Systems & Inference". In: ().

[72] Ulrich Hohenester. *Nano and Quantum Optics: An Introduction to Basic Principles and Theory*. Graduate Texts in Physics. Cham: Springer International Publishing, 2020. ISBN: 978-3-030-30503-1 978-3-030-30504-8. DOI: 10.1007/978-3-030-30504-8.

[73] Claudia Czado and Thorsten Schmidt. *Mathematische Statistik*. Google-Books-ID: fXEkBAAAQBAJ. Springer-Verlag, May 27, 2011. 266 pp. ISBN: 978-3-642-17261-8.

[74] Giorgio Bertotti and Isaak Mayergoyz. "The science of hysteresis. Vol. I. Mathematical modeling and applications". In: 1 (Jan. 1, 2006).

[75] *MicroTCA Technology Lab*. URL: https://innovation.desy.de/technologies/microtca/index_eng.html (visited on 02/12/2025).

[76] Hideharu Amano, ed. *Principles and Structures of FPGAs*. Singapore: Springer Singapore, 2018. ISBN: 9789811308239 9789811308246. DOI: 10.1007/978-981-13-0824-6.

Appendix A

Appendix

A   Exact DFM interference signal

In complex phasor notation, the electric field of laser signal can be written as

$$\vec{E}_{\text{DFMI}}(t, \tau) = \vec{E}_0 \cdot e^{i(\omega_0\tau + \Delta\omega\tau \cdot \sin(\omega_m t + \psi))}. \tag{A.1}$$
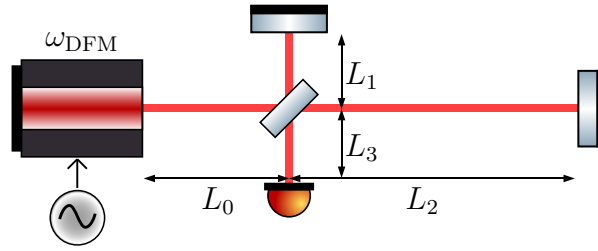
Figure A.1: Typical DFMI setup with laser frequency $\omega_{\text{DFM}} = \omega_0 + \Delta\omega\sin(\omega_m t + \psi)$ and its unequal arm-lengths L_1 and L_2.

For a DFM interferometer setup as shown in Figure A.1, the signal measured on the photo-diode (the time averaged Poynting vector) can be written as:

$$P_{\text{exact}}(t) = c_0\epsilon_0 R_{\text{PD}} \left\| \vec{E}_{\text{DFMI}}(t, \tau_0) + \vec{E}_{\text{DFMI}}(t - \tau_L, \tau_0 + \tau_L) \right\|^2 \tag{A.2}$$

with t being the time of measurement, $\tau_0 := (L_0 + 2L_1 + L_3)/c_0$ as propagation time of the beam traveling along the upper arm and $\tau_L := 2(L_2 - L_1)/c_0$ as propagation time difference between the two beams traveling through L_1 and L_2. Expanding (A.2) yields:

$$P_{\text{exact}}(t) = c_0\epsilon_0 R_{\text{PD}} \left\| \vec{E}_{\text{DFMI}}(t, \tau_0) + \vec{E}_{\text{DFMI}}(t - \tau_L, \tau_0 + \tau_L) \right\|^2 \tag{A.3}$$

$$= c_0\epsilon_0 R_{\text{PD}} E_0^2 \Bigg(2 + 2\underbrace{\kappa}_{\text{optical contrast}} \cos\Big(\omega_0\tau_L + \Delta\omega(\tau_0 + \tau_L)$$

$$\cdot \underbrace{\sin(\omega_m t + \psi - \omega_m\tau_L)}_{=\sin(\omega_m t+\psi)-\omega_m\tau_L\cos(\omega_m t+\psi)+\mathcal{O}((\omega_m\tau_L)^2)} - \Delta\omega\tau_0 \cdot \sin(\omega_m t + \psi)\Big)\Bigg) \tag{A.4}$$

$$= c_0 \epsilon_0 R_{\mathrm{PD}} E_0^2 2 \left(1 + \kappa \cos \left(\underbrace{\varphi}_{\omega_0 \tau_L} + \underbrace{m}_{\Delta\omega\tau_L} \cdot \sin(\omega_m t + \psi) - \underbrace{(m_0+m)}_{\Delta\omega\tau_0} \; \underbrace{\psi_L}_{\omega_m\tau_L} \cos(\omega_m t + \psi) + \mathcal{O}(\psi_L^2) \right) \right)$$

$$\tag{A.5}$$

$$= c_0 \epsilon_0 R_{\mathrm{PD}} E_0^2 2 \bigg(1 + \kappa \cos \Big(\varphi + m \cdot \sin(\omega_m t + \psi) \Big) \tag{A.6}$$

$$- \psi_L \kappa (m_0 + m) \cos(\omega_m t + \psi) \sin \Big(\varphi + m \cdot \sin(\omega_m t + \psi) \Big) + \mathcal{O}(\psi_L^2) \bigg)$$

$$=: P_{\mathrm{ideal}}(t) + P_{\mathrm{approx}}(t) + \mathcal{O}(\psi_L^2) \,. \tag{A.7}$$

Besides the ideal DFMI signal

$$P_{\mathrm{ideal}}(t) = \underbrace{2 c_0 \epsilon_0 R_{\mathrm{PD}} E_0^2}_{=:B} + \underbrace{2 c_0 \epsilon_0 R_{\mathrm{PD}} E_0^2 \kappa}_{=:A} \cdot \cos \Big(\underbrace{\omega_0 \tau_L}_{\varphi} + \underbrace{\Delta\omega\tau_L}_{m} \cdot \sin(\omega_m t + \psi) \Big), \quad \text{(A.8)}$$

the exact expression of the power of the DFMI laser beam also contains additional terms proportional to $\psi_L := \omega_m \tau_L$. Due to the relativly small modulation frequency ω_m these terms are generally small ($\psi_L \approx 10^{-6}$ in the experimental setups discussed in this thesis) and can be neglected. In leading order, the difference between exact and ideal signal can be expressed by P_{approx}, written as:

$$P_{\mathrm{approx}}(t) = \psi_L \kappa A \underbrace{(m_0 + m)}_{=\Delta\omega\tau_0 + \Delta\omega\tau_L} \cos(\omega_m t + \psi) \sin \Big(\varphi + m \cdot \sin(\omega_m t + \psi) \Big).$$

Using $\cos x \sin y = 1/2 \sin(x + y) - 1/2 \sin(x - y)$, shortening the expression by writing $x := (\omega_m t + \psi)$ and using the Jacobi-Anger identity this can be expressed as Fourier series via:

$$P_{\mathrm{approx}}(t) = \psi_L \cdot \kappa A (m_0 + m) \cdot \cos(x) \cdot \sin(\varphi + m \cdot \sin(x))$$

$$= \psi_L \cdot \frac{\kappa A}{2} (m_0 + m) \cdot \left(\sum_{n \in \mathbb{Z}} J_n(m) \sin\left(\varphi + (n + 1)x \right) - J_n(m) \sin\left(\varphi + (n - 1)x \right) \right).$$

Reordering the indices and applying the recurrence relations $J_{n-1}(m) - J_{n+1}(m) = 2 \, \partial J_n(m)/\partial m$ also allows one to write the approximation term as:

$$P_{\mathrm{approx}}(t) = \psi_L \cdot \frac{\kappa A}{2} (m_0 + m) \cdot \left(\sum_{n \in \mathbb{Z}} \underbrace{(J_{n-1}(m) - J_{n+1}(m))}_{=2\frac{\partial}{\partial m} J_n(m)} \sin\left(\varphi + nx \right) \right)$$

$$= \psi_L \cdot \kappa A (m_0 + m) \cdot \frac{\partial}{\partial m} \underbrace{\left(\sum_{n \in \mathbb{Z}} J_n(m) \sin\left(\varphi + nx \right) \right)}_{=\sin(\varphi + m \sin x)}$$

$$= \psi_L \cdot \kappa A (m_0 + m) \cdot \sin(x) \cos(\varphi + m \sin x)$$

$$= \psi_L \cdot (m_0 + m) \cdot \sin(\omega_m t + \psi) \cdot P_{\mathrm{ideal}}(t) \,.$$

B   Calculations of the exact ReDFMI signal

As already stated in section 6, we start from the complex Phasor description of the DFMI signal given by

$$E_{\mathrm{DFM}}(t,\tau) = E_0 \; e^{i(\omega_0\tau + \Delta\omega\tau\cdot\sin(\omega_m t + \psi))}. \tag{A.9}$$

with t as time of emission and τ as (fixed) propagation time. The signal transmitted from a Fabry-Perot cavity as shown in Figure 6.2 with identical reflective coefficients r for both mirrors (with reflectivity $R := r^2$ and transmissivity of $(1 - R)$) can be written as the sum over all reflections that exit the cavity at the same time given by the time of emission plus the propagation time $t + \tau$. The first beam emitted at time t that would go "straight through" the cavity would travel some time until it reaches the first mirror, travel through the cavity and exit at time $t + \tau_0$. The beam that is reflected twice (once at each mirror) before being transmitted would have been emitted at an earlier time $t - \tau_L$ with $\tau_L = 2L/c_0$ so that after a propagation time of $\tau_0 + \tau_L$ it would exit the cavity the right at time $(t - \tau_L) + (\tau_0 + \tau_L) = t + \tau_0$, and so on and so forth.
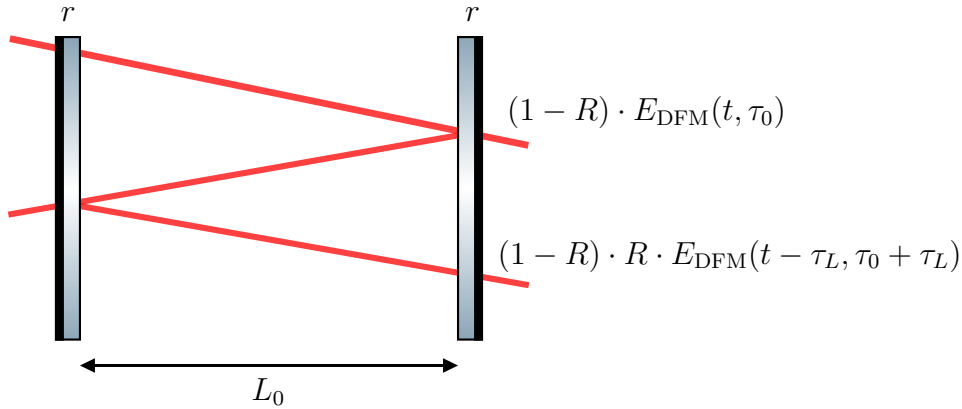
Figure B.1: Sketch of a Fabry-Perot-Cavity where both mirrors have the same reflective coefficient r.

As series this can be written down as

$$E_{\mathrm{ReDFM,\,trans}}(t,\tau) = (1 - R)\sum_{k=0}^{\infty} R^k \; E_{\mathrm{DFM}}(t - k\tau_L, \tau_0 + k\tau_L) \tag{A.10}$$

$$\stackrel{(A.9)}{=} (1 - R)E_0 \sum_{k=0}^{\infty} R^k e^{i(\omega_0\tau + k\omega_0\tau_L + \Delta\omega(\tau_0 + k\tau_L)\cdot\sin(\omega_m t - k\omega_m\tau_L + \psi_0))} \tag{A.11}$$

$$=: (1 - R)E_0 \sum_{k=0}^{\infty} R^k e^{i(\varphi_0 + k\varphi_L + (m_0 + km_L)\cdot\sin(\omega_m t - k\psi_L + \psi_0))} \tag{A.12}$$

with ReDFMi parameters

$$\begin{aligned} \omega_0 \cdot \tau_0 &=: \varphi_0 & \omega_0 \cdot \tau_L &=: \varphi_L & \omega_m \cdot \tau_0 &=: \psi_L \\ \Delta\omega \cdot \tau_0 &=: m_0 & \Delta\omega \cdot \tau_L &=: m_L & & . \end{aligned}$$

Next we use the Jacobi–Anger expansion

$$e^{im\sin(x)} = \sum_{n\in\mathbb{Z}} J_n(m)\, e^{inx} \tag{A.13}$$

to further rewrite the signal to

$$E_{\text{ReDFM, trans}}(t,\tau_0) = (1-R)E_0 \sum_{k=0}^{\infty} R^k e^{i(\varphi_0+k\varphi_L)} \sum_{n\in\mathbb{Z}} J_n(m_0+km_L)e^{in(\omega_m t-k\psi_L+\psi_0)} \tag{A.14}$$

$$= \sum_{n\in\mathbb{Z}} \underbrace{\left((1-R)E_0 e^{i(\varphi_0+n\psi_0)} \sum_{k=0}^{\infty} R^k e^{ik\varphi_L} J_n(m_0+km_L)e^{-ink\psi_L}\right)}_{=:\,C_n} \cdot e^{in\omega_m t} \tag{A.15}$$

This expression already has the general shape of a complex Fourier Series like

$$E_{\text{ReDFM, trans}}(t,\tau_0) = \sum_{n\in\mathbb{Z}} C_n \cdot e^{in\omega_m t} \tag{A.16}$$

with the fundamental frequency ω_m. What remains is to find a closed expression for the "C_n" term, e.g. the term inside the brackets in (A.15).

To find a closed expression for the C_n, we start by using the "Hansen-Bessel formula"

$$J_n(m) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{i(n\sigma - m\sin(\sigma))}\, d\sigma. \tag{A.17}$$

to replace the appearing Bessel function with an integral expression. Using it, we can write

$$C_n \overset{(A.15)}{:=} \left((1-R)E_0 e^{i(\varphi_0+n\psi_0)} \sum_{k=0}^{\infty} R^k e^{ik\varphi_L} J_n(m_0+km_L)e^{-ink\psi_L}\right) \tag{A.18}$$

$$\overset{(A.17)}{=} \frac{(1-R)E_0}{2\pi} e^{i(\varphi_0+n\psi_0)} \sum_{k=0}^{\infty} e^{k(\ln R+i\varphi_L)} \int_{-\pi}^{\pi} e^{i(n\sigma-m_0\sin\sigma-km_L\sin\sigma)}\, d\sigma\; e^{-ink\psi_L} \tag{A.19}$$

$$= \frac{(1-R)E_0}{2\pi} e^{i(\varphi_0+n\psi_0)} \int_{-\pi}^{\pi} d\sigma\; e^{i(n\sigma-m_0\sin\sigma)} \sum_{k=0}^{\infty} e^{k(\ln R+i(\varphi_L-n\psi_L-m_L\sin\sigma))} \tag{A.20}$$

The sum over k on the right corresponds to the geometric series which can be written as:

$$\sum_{k=0}^{\infty} e^{k(\ln R+i(\varphi_L-n\psi_L-m_L\sin\sigma))} = \sum_{k=0}^{\infty} \left(e^{\ln R+i(\varphi_L-n\psi_L-m_L\sin\sigma)}\right)^k \tag{A.21}$$

$$= \frac{1}{1-e^{\ln R+i(\varphi_L-n\psi_L-m_L\sin\sigma)}} \tag{A.22}$$

Plugged back in, this yields

$$C_n = \frac{(1-R)E_0}{2\pi} e^{i(\varphi_0+n\psi_0)} \int_{-\pi}^{\pi} d\sigma \, e^{i(n\sigma-m_0\sin\sigma)} \left(\frac{1}{1 - e^{\ln R + i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \right) \tag{A.23}$$

$$\stackrel{(A.17)}{=} \frac{(1-R)E_0}{2\pi} e^{i(\varphi_0+n\psi_0)} \left(\int_{-\pi}^{\pi} \frac{e^{i(n\sigma-m_0\sin\sigma)}}{1 - e^{\ln R + i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \right) \tag{A.24}$$

The final step in the calculation is to solve the integral in (A.24). To do this we rewrite the expression as integral in the complex plane and solve it by applying the residue theorem of complex analysis.

Starting with the integral and again using the Jacobi–Anger expansion we can write

$$\int_{-\pi}^{\pi} \frac{e^{i(n\sigma-m_0\sin\sigma)}}{1 - e^{\ln R + i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \stackrel{(A.13)}{=} \int_{-\pi}^{\pi} \frac{e^{in\sigma} \sum_{l\in\mathbb{Z}} J_l(m_0) e^{-il\sigma}}{1 - Re^{i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \tag{A.25}$$

$$= \sum_{l\in\mathbb{Z}} J_l(m_0) \int_{-\pi}^{\pi} \frac{e^{i(n-l)\sigma}}{1 - Re^{i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \tag{A.26}$$

Substituting $z := e^{i\sigma}$ this integral becomes an integral of z over the unit circle S_1 (1-Sphere) in the complex plane (with an additional $1/z$ factor from the substitution $d\sigma = dz/z$).

$$\int_{-\pi}^{\pi} \frac{e^{i(n\sigma-m_0\sin\sigma)}}{1 - e^{\ln R + i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \tag{A.27}$$

$$= \sum_{l\in\mathbb{Z}} J_l(m_0) \int_{-\pi}^{\pi} \frac{e^{i(n-l)\sigma}}{1 - Re^{i(\varphi_L - n\psi_L - m_L\sin\sigma)}} \, d\sigma \tag{A.28}$$

$$\stackrel{(z:=e^{i\sigma})}{=} \sum_{l\in\mathbb{Z}} J_l(m_0) \int_{S_1\subseteq\mathbb{C}} \frac{z^{(n-l)}}{\left(1 - Re^{i(\varphi_L-n\psi_L)} e^{im_L\sin(i\ln z)}\right)} \frac{dz}{z} \tag{A.29}$$

$$= \sum_{l\in\mathbb{Z}} J_l(m_0) \int_{S_1\subseteq\mathbb{C}} \frac{z^{(n-l-1)}}{\underbrace{\left(1 - Re^{i(\varphi_L-n\psi_L)} e^{-\frac{m_L}{2}\frac{(z^2-1)}{z}}\right)}_{=:\, f(z)}} \, dz \tag{A.30}$$

$$=: \sum_{l\in\mathbb{Z}} J_l(m_0) \int_{S_1\subseteq\mathbb{C}} f(z) \, dz \tag{A.31}$$

Since this is a closed integral in the complex plane we can apply the residual theorem to calculate the integral via

$$\int_{S_1\subseteq\mathbb{C}} f(z) \, dz = \sum_{z\in\text{Residuals in } S_1} 2\pi i \cdot \text{Res}(f,z). \tag{A.32}$$

Firstly, we can directly read off the singularities of the nominator of $f(z)$ in (A.30), which is a pole of order $(n-l-1)$ at $z=0$. The denominator however also diverges in the limit of $z \mapsto 0$ and scales with $\propto 1 - e^{1/z}$.

$$\lim_{z\to 0} \left(1 - Re^{i(\varphi_L-n\psi_L)} e^{-\frac{m_L}{2}\frac{(z^2-1)}{z}} \right) \mapsto \infty \tag{A.33}$$

Further calculation yields that the residual at $z = 0$ is in fact zero:

$$\text{Res}(f, z = 0) = \lim_{z \to 0} z \cdot f(z) = 0. \tag{A.34}$$

The other singularities would be given by the zeros of the denominator of $f(z)$ which requires a few more steps. First, we find the zeros of the denominator by calculating:

$$0 = 1 - R e^{i(\varphi_L - n\psi_L)} \, e^{-\frac{m_L}{2} \frac{(z^2-1)}{z}} \tag{A.35}$$

$$\Longleftrightarrow 0 = \ln R + i(\varphi_L - n\psi_L) - \frac{m_l}{2}\frac{(z_0^2 - 1)}{z_0} \tag{A.36}$$

$$\Longleftrightarrow 0 = z_0^2 + \frac{2}{m_L}(\ln R + i(\varphi_L - n\psi_L)) \cdot z_0 - 1 \tag{A.37}$$

$$\Longrightarrow z_{0,\pm} := \left(\pm\sqrt{\frac{(\ln R + i(\varphi_L - n\psi_L))^2}{m_L^2} + 1} - \frac{(\ln R + i(\varphi_L - n\psi_L))}{m_L} \right) \tag{A.38}$$

Thus we found that the denominator of $f(z)$ has two singularities, one at $z_{0,+}$ and one at $z_{0,-}$. At this point I also want to define an auxiliary variable of

$$\theta_n := \frac{\ln R + i(\varphi_L - n\psi_L)}{m_L} \tag{A.39}$$

which allows us to write the values for $z_{0,\pm}$ as:

$$z_{0,\pm} = -\theta_n \pm \sqrt{\theta_n^2 + 1} \tag{A.40}$$

which makes the following equations a little shorter. Since we derived these singularities from a quadratic equation, I start with the assumption that both singularities are 1st (or 2nd order) poles and try to calculate the residuals directly via:

$$\text{Res}(f, z_{0,\pm}) = \lim_{z \to z_{0,\pm}} (z - z_{0,\pm}) \cdot f(z) = \lim_{z \to z_{0,\pm}} \frac{\overbrace{(z - z_{0,\pm})z^{(n-l-1)}}^{=:\, u(z)}}{\underbrace{\left(1 - R\, e^{i(\varphi_L - n\psi_L)} e^{\frac{m_L}{2}\frac{(z^2-1)}{z}} \right)}_{=:\, v(z)}} \tag{A.41}$$

In the limit of $z \to z_{0,\pm}$ and using the two simplifications:

$$\frac{(z_{0,\pm}^2 - 1)}{z_{0,\pm}} \overset{(A.37)\&(A.39)}{=} -2\,\theta_n \qquad \text{and} \qquad e^{m_L \theta_n} = R e^{i(\varphi_L - n\psi_L)} \tag{A.42}$$

we find that both nominator and denominator of (A.41) become zero. To calculate

the limit we must apply L'Hoptial's rule which yields:

$$\lim_{z \to z_{0,\pm}} \frac{u(z)}{v(z)} = \lim_{z \to z_{0,\pm}} \frac{u'(z)}{v'(z)} = \lim_{z \to z_{0,\pm}} \frac{z^{(n-l-1)} + (n-l-1)(z - z_{0,\pm})z^{(n-l-2)}}{-e^{m_L \theta_n} \underbrace{\frac{\partial}{\partial z} e^{\frac{m_L}{2} \frac{(z^2-1)}{z}}}_{=e^{\frac{m_L}{2} \frac{(z^2-1)}{z}} \cdot \frac{m_L}{2} \frac{(z^2+1)}{z^2}}} \tag{A.43}$$

$$= \lim_{z \to z_{0,\pm}} \frac{z^{(n-l-1)} + (n-l-1)(z - z_{0,\pm})z^{(n-l-2)}}{e^{m_L \theta_n} \frac{m_L}{2} \frac{(z^2+1)}{z^2} \cdot e^{\frac{m_L}{2} \frac{(z^2-1)}{z}}} \tag{A.44}$$

$$= \frac{z_{0,\pm}^{(n-l-1)}}{e^{m_L \theta_n} \frac{m_L}{2} \frac{(z_{0,\pm}^2+1)}{z_{0,\pm}^2} \cdot e^{\frac{m_L}{2} \frac{(z_{0,\pm}^2-1)}{z_{0,\pm}}}} \tag{A.45}$$

$$\overset{(A.42)}{=} \frac{z_{0,\pm}^{(n-l-1)}}{e^{m_L \theta_n} \frac{m_L}{2} \frac{(z_{0,\pm}^2+1)}{z_{0,\pm}^2} \cdot e^{-m_L \theta_n}} \tag{A.46}$$

$$= \frac{2}{m_L} \frac{z_{0,\pm}^{(n-l+1)}}{(z_{0,\pm}^2 + 1)} \tag{A.47}$$

Now, applying the residual theorem to calculate the integral in (A.32) yields:

$$\int_{S_1 \subseteq \mathbb{C}} f(z) \, dz = \sum_{z \in \text{Residuals in } S_1} 2\pi i \cdot \text{Res}(f, z) \tag{A.48}$$

$$= \frac{4\pi i}{m_L} \left(\frac{z_{0,+}^{(n-l+1)}}{(z_{0,+}^2 + 1)} + \frac{z_{0,-}^{(n-l+1)}}{(z_{0,-}^2 + 1)} \right) \tag{A.49}$$

$$= \frac{4\pi i}{m_L} \frac{1}{(z_{0,+}^2 + 1)(z_{0,-}^2 + 1)} \left(z_{0,+}^{(n-l+1)}(z_{0,-}^2 + 1) + z_{0,-}^{(n-l+1)}(z_{0,+}^2 + 1) \right) \tag{A.50}$$

This expression can now be further simplified. By substituting a second time with

$$x := \text{arcsinh}(\theta_n) = \text{arcsinh}\left(\frac{\ln R + i(\varphi_L - n\psi_L)}{m_L} \right) \tag{A.51}$$

the the $z_{0,\pm}$ values can be written as:

$$z_{0,\pm} \overset{(A.40)\&(A.51)}{=} -\sinh(x) \pm \sqrt{\sinh^2(x) - 1} = -\sinh(x) \pm \cosh(x) \tag{A.52}$$

$$\implies z_{0,+} = e^{-x} \quad \text{and} \quad z_{0,-} = -e^x. \tag{A.53}$$

Using (A.53) in place of $z_{0,\pm}$ yields the relations

$$z_{0,+}^2 z_{0,-}^2 = e^{-2x} e^{2x} = 1 \tag{A.54}$$

$$\text{and} \quad (z_{0,+}^2 + 1)(z_{0,-}^2 + 1) = 4\cosh^2 x \tag{A.55}$$

Plugged back into (A.50) this yields

$$\int_{S_1 \subseteq \mathbb{C}} f(z) \, dz = \frac{\pi i}{m_L \cosh^2 x} \left(z_{0,+}^{(n-l+1)} + z_{0,-}^{(n-l+1)} + z_{0,+}^{(n-l-1)} + z_{0,-}^{(n-l-1)} \right) \tag{A.56}$$

$$= \frac{\pi i}{m_L \cosh^2 x} \left(e^{-(n-l)x} \underbrace{(e^{-x} + e^x)}_{=2\cosh x} + e^{(n-l)x} \underbrace{(-e^x - e^{-x})}_{=-2\cosh x} \right) \tag{A.57}$$

$$= \frac{2\pi i}{m_L \cosh x} \left(e^{-(n-l)x} + e^{(n-l)x} \right) \tag{A.58}$$

Using this, we can now write the earlier integral expression of (A.27) as

$$\int_{-\pi}^{\pi} \frac{e^{i(n\sigma - m_0 \sin \sigma)}}{1 - e^{\ln R + i(\varphi_L - n\psi_L - m_L \sin \sigma)}} \, d\sigma = \sum_{l \in \mathbb{Z}} J_l(m_0) \int_{S_1 \subseteq \mathbb{C}} f(z) \, dz$$

$$= \sum_{l \in \mathbb{Z}} J_l(m_0) \cdot \frac{2\pi i}{m_L \cosh x} \left(e^{-(n-l)x} + e^{(n-l)x} \right) \tag{A.59}$$

$$= \frac{2\pi i}{m_L \cosh x} \left(e^{-nx} \sum_{l \in \mathbb{Z}} J_l(m_0) e^{il(-ix)} + e^{nx} \sum_{l \in \mathbb{Z}} J_l(m_0) e^{il(ix)} \right) \tag{A.60}$$

$$= \frac{2\pi i}{m_L \cosh x} \left(e^{-nx} e^{im_0 \sin(-ix)} + e^{nx} e^{im_0 \sin(ix)} \right) \tag{A.61}$$

$$= \frac{2\pi i}{m_L \cosh x} \left(e^{-nx + m_0 \sinh(x)} + e^{nx - m_0 \sinh(x)} \right) \tag{A.62}$$

$$= \frac{4\pi i}{m_L \cosh x} \cosh(nx - m_0 \sinh(x)) \tag{A.63}$$

$$= \frac{4\pi i}{m_L} \frac{\cosh(n \operatorname{arcsinh} \theta_n - m_0 \theta_n)}{\sqrt{\theta_n^2 + 1}} \tag{A.64}$$

Plugging this expression for the integral in (A.24) yields for the C_n:

$$C_n = (1 - R) E_0 e^{i(\varphi_0 + n\psi_0)} \frac{2i}{m_L \cosh x} \cosh(nx - m_0 \sinh(x)) \tag{A.65}$$

This allows us to write the transmitted electric field phasor of the ReDFMI signal as:

$$E_{\text{ReDFM, trans}}(t, \tau_0) = (1 - R) E_0 \sum_{n \in \mathbb{Z}} \frac{2i}{m_L \cosh x} \cosh(nx - m_0 \sinh(x)) e^{i(\varphi_0 + n\psi_0)} \cdot e^{in\omega_m t} \tag{A.66}$$

The signal measured with a photodiode of such an incoherent EM-wave would be proportional to the (real valued) averaged Poynting-vector

$$\langle S \rangle = \tfrac{1}{2} \Re(\vec{E} \times \vec{H}^*), \tag{A.67}$$

which corresponds here to $\langle S \rangle \propto |\vec{E}_{\text{ReDFM, trans}}|^2$.

C   Regularity Conditions for the Cramér-Rao bound

The **Cramér-Rao regularity conditions** as written in [73], for a statistic model / estimator $T(x)$ and a random variable X with variable θ (such that the probability distribution of X is a function of $\rho(x, \theta)$), are:

1. The set $\Theta \subset \mathbb{R}$ is open.
   With $\theta \in \Theta$ as the set of possible values of θ. (If Θ were only half-open or closed, the derivative at the boundary of Θ would not exist).

2. $\Omega := \{x \in \mathbb{R}^n : \rho(x, \theta) > 0\}$ does not depend in θ, and the derivative $\partial/\partial\theta \ln \rho(x, \theta)$ exists and is finite for $\forall x \in \Omega, \forall \theta \in \Theta$.
   The derivative, being the Fisher-Information, is always defined.

3. For an estimator T with finite expected value $\mathbb{E}[T] < \infty$ for $\forall \theta \in \Theta$, the integraion over x and differention with respect to θ commute for the expected value $\mathbb{E}[T]$; i.e.

$$\frac{\partial}{\partial\theta} \int_\Omega \rho(x, \theta) T(x) dx = \int_\Omega \frac{\partial}{\partial\theta} \rho(x, \theta) T(x) dx \qquad (A.68)$$

   The probability distribution and the estimator are "sufficiently smooth".

D   Additional considerations for the CRB of correlated noise

For noise correlated to the main signal, the (joint) probability distribution of an extended measurement can not simply be written as product of the (not explicitly time dependent) probability distributions of a single measurement as done in (8.9). It might however still be possible to derive the CRB for such cases without actually knowing the precise joint probability distribution (and its behavior over time). Using the "chain rule of probability," we can write the joint probability distribution $\rho(\underline{x}) = \rho(x_1, \ldots, x_N)$ as

$$\rho(\underline{x}) = \rho_i(x_i) \cdot \int_{\Omega_i} \rho(x_1, \ldots, x_i', \ldots, x_N)\, dx_i' \,. \qquad (A.69)$$

Calculating the derivative $\partial/\partial x_i$ of (A.69) yields the later usefull equation:

$$\implies \frac{\partial}{\partial x_i}\rho(\underline{x}) = \frac{\partial}{\partial x_i}\rho_i(x_i) \cdot \int_{\Omega_i} \rho(x_1, \ldots, x_i', \ldots, x_N)\, dx_i' \qquad (A.70)$$

$$\implies \frac{\frac{\partial}{\partial x_i}\rho(\underline{x})}{\rho(\underline{x})} = \frac{\frac{\partial}{\partial x_i}\rho_i(x_i)}{\rho_i(x_i)} \qquad (A.71)$$

$$\implies \frac{\partial}{\partial x_i} \ln \rho(\underline{x}) = \frac{\partial}{\partial x_i} \ln \rho_i(x_i) \qquad (A.72)$$

I.e. the partial derivative of the logarithm of the full/joint probability ρ equals the partial derivative of the specific individual distribution ρ_i. Using this expression

when calculating the CRB then yields:

$$\text{FI}_\theta = \mathbb{E}\left[\left(\frac{\partial}{\partial\theta}\ln\rho(\underline{x},\theta)\right)^2\right] \tag{A.73}$$

$$\stackrel{\text{total derivative}}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial x_i}{\partial\theta}\cdot\frac{\partial}{\partial x_i}\ln\rho(\underline{x},\theta)\right)^2\right] \tag{A.74}$$

$$\stackrel{(A.72)}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial x_i}{\partial\theta}\cdot\frac{\partial}{\partial x_i}\ln\rho_i(x_i,\theta)\right)^2\right] \tag{A.75}$$

$$\stackrel{(8.8)}{=} \mathbb{E}\left[\left(\sum_{i=1}^{N}\frac{\partial s_i}{\partial\theta}\cdot\left(\underbrace{\frac{\partial}{\partial x_i}\ln\left(\frac{1}{\sqrt{2\pi}\sigma_i}\right)}_{=0}-\underbrace{\frac{\partial}{\partial x_i}\left(\frac{(x_i-s_i-\mu)^2}{2\sigma_i^2}\right)}_{\frac{1}{\sigma_i^2}(x_i-s_i-\mu)}\right)\right)^2\right] \tag{A.76}$$

$$\stackrel{\text{Expanding the }(\ldots)^2}{=} \sum_{i,j=1}^{N}\frac{\partial s_i}{\partial\theta}\frac{\partial s_j}{\partial\theta}\cdot\frac{1}{\sigma_i^2\sigma_j^2}\underbrace{\mathbb{E}\left[(x_i-(s_i+\mu))(x_j-(s_j+\mu))\right]}_{=R_{xx}(t_i,t_j)-(s_i+\mu)(s_j+\mu)} \tag{A.77}$$

$$\implies \boxed{\text{FI}_\theta = \sum_{i,j=1}^{N}\left(\frac{\partial s(t_i,\theta)}{\partial\theta}\right)\left(\frac{\partial s(t_j,\theta)}{\partial\theta}\right)\cdot\frac{1}{\sigma_i^2\sigma_j^2}\left(R_{xx}(t_i,t_j)-(s(t_i)+\mu)(s(t_j)+\mu)\right)} \tag{A.78}$$

I want to note here that for WSS noise n, the autocorrelation only depends on the time-difference $t_i - t_j$. The signal s is however generally <u>not</u> WSS and the full measured signal $x = s + n$ therefore also not.

While the autocorrelation of x can be decomposed into the auto- and cross-correlation of its constituents:

$$x = s + n \implies R_{xx}(t_i,t_j) = R_{ss}(t_i,t_j) + R_{sn}(t_i,t_j) + R_{sn}(t_j,t_i) + R_{nn}(t_i,t_j), \tag{A.79}$$

I was yet unable to find a useful expression from inserting this term back into (A.78). Maybe some additional assumptions as i.e. s being somehow WSS and the use of Plancherels theorem theorem might provide a more useful expression for the Fisher Information somehow derived from the the auto- and cross-spectral densities of the signal s and the noise n. For now, it remains however an open problem which I was unable to solve yet.

E   List of used integrals of DFMI signal derivatives

This section contains a selection of integrals of the squared derivative of the ideal DFMI with respect to different variables. These integral appear when calculating

the Fisher-Information or CRB of a DFMI signal for different signal parameters. The integration period is always assumed to be (exactly) an integer multiple of the modulation period i.e. $T = 2\pi/\omega_m \cdot \mathbb{N}$. Such that terms containing the modulation frequency exactly vanish, i.e.

$$\int_0^T \cos(\theta + k\omega_m t) = T\,\delta_{k,0}\,. \tag{A.80}$$

Integration of the ideal DFMI signal yields:

$$\int_0^T dt\ P_{\text{ideal}}(t) = \int_0^T dt A\left(1 + \cos(m \cdot \sin(\omega_m t + \psi) + \varphi)\right) \tag{A.81}$$

$$= \int_0^T dt\left(A + A\sum_{n\in\mathbb{Z}} J_n(m)\cos(n(\omega_m t + \psi) + \varphi)\right) \tag{A.82}$$

$$= TA + A\sum_{n\in\mathbb{Z}} J_n(m)\int_0^T dt\,\cos(n\omega_m t + n\psi + \varphi) \tag{A.83}$$

$$= TA + A\sum_{n\in\mathbb{Z}} J_n(m)\,\delta_{n0} T\cos(\varphi) \tag{A.84}$$

$$= TA\left(1 + J_0(m)\cos\varphi\right) \tag{A.85}$$

The integrals appearing when calculating the Fisher-Information for additive white gaussian noise for the different DFMi parameters are: For φ

$$\int_0^T dt\ \left(\frac{\partial P_{\text{DFM}}(t)}{\partial\varphi}\right)^2 = \int_0^T dt\ A^2\ \sin^2(m \cdot \sin(\omega_m t + \psi) + \varphi) \tag{A.86}$$

$$= \int_0^T dt\ A^2\ \frac{1}{2}\left[1 - \cos(2m \cdot \sin(\omega_m t + \psi) + 2\varphi)\right] \tag{A.87}$$

$$= T\frac{A^2}{2} - \frac{A^2}{2}\int_0^T dt\ \cos(2m \cdot \sin(\omega_m t + \psi) + 2\varphi) \tag{A.88}$$

$$= T\frac{A^2}{2} - T\frac{A^2}{2}J_0(2m)\cdot\cos(2\varphi) \tag{A.89}$$

$$= T\frac{A^2}{2}\left(1 - J_0(2m)\cdot\cos(2\varphi)\right) \tag{A.90}$$

$$\implies \text{var}(\hat\varphi) \geq \frac{2\sigma^2}{f_S T A^2\left(1 - J_0(2m)\cdot\cos(2\varphi)\right)} \tag{A.91}$$

For m

$$\int_0^T dt \ \left(\frac{\partial P_{\mathrm{DFM}}(t)}{\partial m}\right)^2 = \int_0^T dt \ A^2 \ \sin^2(\omega_m t + \psi) \ \sin^2(m \cdot \sin(\omega_m t + \psi) + \varphi) \tag{A.92}$$

$$= \int_0^T dt \ \frac{A^2}{4} \left[1 - \cos(\omega_m t + \psi) - \cos(2m \cdot \sin(\omega_m t + \psi)) \right.$$

$$+ \frac{1}{2} \cos(2m \cdot \sin(\omega_m t + \psi) + 2\varphi + 2(\omega_m t + \psi))$$

$$\left. + \frac{1}{2} \cos(2m \cdot \sin(\omega_m t + \psi) + 2\varphi - 2(\omega_m t + \psi)) \right] \tag{A.93}$$

$$= \frac{A^2}{4} \left[T - 0 - J_0(2m)T + \frac{1}{2} J_{-2}(2m) \cos(2\varphi)T + \frac{1}{2} J_2(2m) \cos(2\varphi)T \right] \tag{A.94}$$

$$= \frac{TA^2}{4} \left[1 - J_0(2m) + J_2(2m) \cos(2\varphi) \right] \tag{A.95}$$

$$\implies \mathrm{var}(\hat{m}) \geq \frac{4\sigma^2}{f_S T A^2 \left(1 - J_0(2m) + J_2(2m) \cos(2\varphi)\right)} \tag{A.96}$$

For A

$$\int_0^T dt \ \left(\frac{\partial P_{\mathrm{DFM}}(t)}{\partial \varphi}\right)^2 = \int_0^T dt \ (1 + \cos(m \cdot \sin(\omega_m t + \psi) + \varphi))^2 \tag{A.97}$$

$$= \int_0^T dt \frac{3}{2} + 2\cos(m \cdot \sin(\omega_m t + \psi) + \varphi) + \cos(2m \cdot \sin(\omega_m t + \psi) + 2\varphi) \tag{A.98}$$

$$= T \left(\frac{3}{2} + 2J_0(m) \cdot \cos(\varphi) + \frac{1}{2} J_0(2m) \cdot \cos(2\varphi) \right) \tag{A.99}$$

$$\implies \mathrm{var}(\hat{A}) \geq \frac{2\sigma^2}{f_S T \left(3 + 4J_0(m) \cos(\varphi) + J_0(2m) \cos(2\varphi)\right)} \tag{A.100}$$

For ψ (substituting $\omega_m t + \psi =: x$ for readability):

$$\int_0^T dt \, \left(\frac{\partial P_{\text{DFM}}(t)}{\partial \psi} \right)^2 = \int_0^T dt \, A^2 m^2 \, \cos^2(x) \sin^2(m \cdot \sin(x) + \varphi) \tag{A.101}$$

$$= \int_0^T dt \, \frac{A^2 m^2}{4} \Bigg[1 + \cos(2x) - \cos(2m \cdot \sin(x) + 2\varphi)$$

$$- \frac{1}{2} \cos(2m \cdot \sin(x) - 2x + 2\varphi) - \frac{1}{2} \cos(2m \cdot \sin(x) + 2x + 2\varphi) \Bigg] \tag{A.102}$$

$$= \frac{A^2 m^2}{4} \Bigg[T + 0 - T J_0(2m) \cos(2\varphi) - \frac{T}{2} J_2(2m) \cos(2\varphi)$$

$$- \frac{T}{2} J_{-2}(2m) \cos(2\varphi) \Bigg] \tag{A.103}$$

$$= \frac{T A^2 m^2}{4} \Bigg[1 - J_0(2m) \cos(2\varphi) - J_2(2m) \cos(2\varphi) \Bigg] \tag{A.104}$$

$$= \frac{T A^2 m^2}{4} \left(1 - \frac{J_1(2m)}{m} \cos(2\varphi) \right) \tag{A.105}$$

$$\implies \text{var}(\hat{\psi}) \geq \frac{4\sigma^2}{f_S T A^2 \left(m^2 - m \, J_1(2m) \cos(2\varphi) \right)} \tag{A.106}$$

In the expressions above, σ^2 is the variance of the gaussian noise given in the units of the measured signal. If one would want to calculate the CRB for a given experimental setup, appropriate conversion factor may need to be added (multiplied) either to the noise (σ) or the signal (A) to ensure that noise and signal are given in the same units.

F   The electronic segment in detail

The electronics segment of the DFMI phasemeter test-setup consists of 5 parts: The conversion of the optical signal into an electrical one by the PD. The amplification of the small photo current by a trans-impendace amplifier. The (low-pass) filtering of the amplified signal. The generation of an offset and a signal inverter to generate a fully differential output. And finally the digitization of the electronic signal by an ADC.

An overview of the used electric circuit is shown in Figure F.1 with its individual parts and the resulting transfer function discussed in the following subsections.
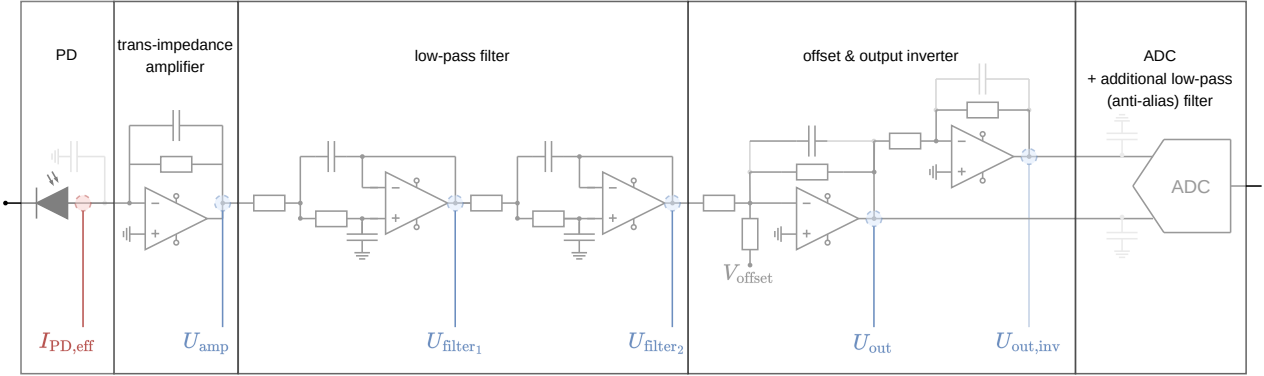
Figure F.1: Overview of the electronic segment of the readout chain

F.1   PD characterization

As Photodiodes we used LAPD-1-09-17-TO463 InGaAs photodiodes for the free-beam setup and two MTSM2601SMF2-150 InGaAs photodiodes glued directly to the QMC (direct and indirect output) as part of the COBRIs. Table F.1 shows some characteristic values of these two photo-diodes extracted from their data-sheets. In

	Bias Voltage U_{Bias}	Responsivity R_{PD}	Dark Capacitance
In-Air PD	$5\,V$	$\approx 1\frac{A}{W}$	$\approx 60 \leq 55\mathrm{pF} \leq 80$
COBRI PD	$0.83\,V$	$\approx 0.787\frac{A}{W}$	NA

Table F.1: Characteristic of the used photo-diodes extracted from their data-sheets by use of a Plot-digitizer application ("WebPlotDigitizer")

our configuration we set the photo-diodes in "reverse bias" mode. For an unbiased PD the photo-current flows in the direction of the PDs Anode (+) to Cathode (-) pole. When supplying a positive Bias voltage to the PDs cathode, the depletion area of the semiconductor becomes larger, the reaction time becomes shorter (the signal bandwidth becomes larger) and the photo-current will flow "away" from the PDs Anode. The direction of the PD in a circuit is in this case flipped or "reversed". This mode of operation is called "reverse bias" mode.

For the Free-Beam PD, the internal capacitance for unbiased operation is given by ≈ 140pF. With a bias voltage of 5V this capacitance if halved down to ≈ 70pF. While the dampening effect on the harmonics of the expected DFMI signals (with $\approx 500\,$Hz as modulation frequency) is relatively small, running the Photo-diode in revered bias mode lowers this small attenuation of the harmonics further by another factor of ~ 2. The trade-off for this setup is a higher parasitic capacitance and a slightly higher dark current (offset) of about $I_{\text{dark}} \approx 0.275\,$nA and a higher capacitance.
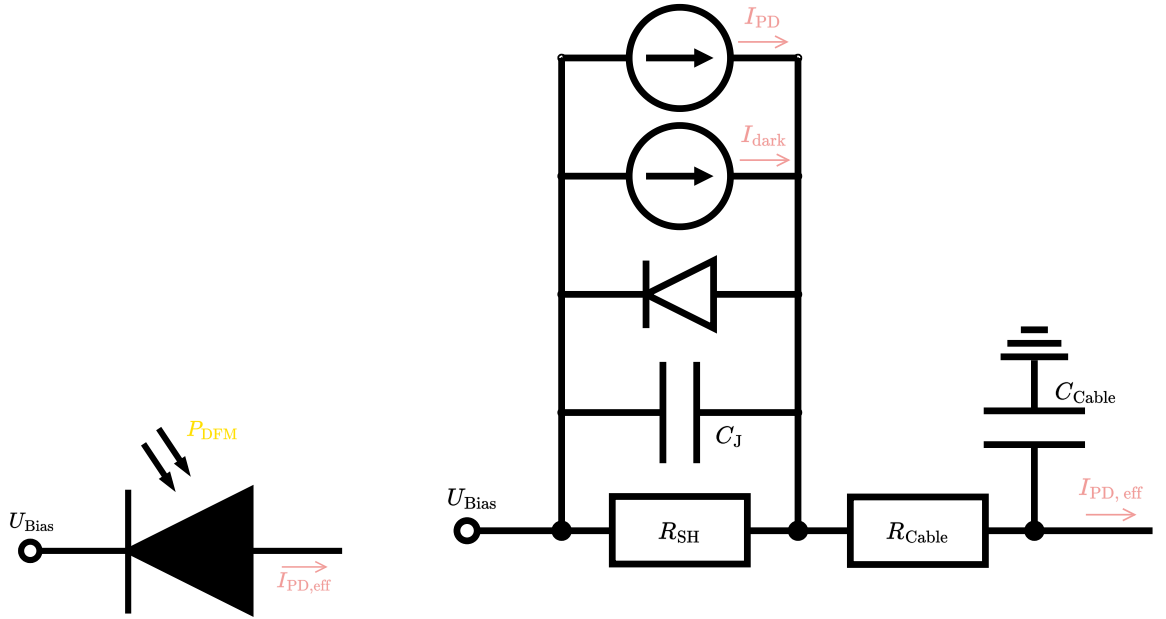
Figure F.2: Schematic symbol of a Photodiode. Ideally it only acts as a current source with an output current proportional to the incoming light P_{DFM}

Figure F.3: Circuit model of a Photodiode. Besides the ideal diode and the photo-current, there is an additional dark current I_{dark}, a shunt resistance R_{SH} and a junction capacitance C_{J}. This capacitance (and the dark current) depend also on the Bias Voltage U_{Bias}.

For an optical signal P_{DFM} on the photo diodes, the resulting current $I_{\text{PD,eff}}$ is given by

$$I_{\text{PD,eff}} = R_{\text{PD}} \cdot P_{\text{DFM}} + I_{\text{dark}} \,. \tag{A.107}$$

F.2 Trans-Impedance Amplifier (TIA) stage

The trans-impedance amplifier stage (sketched in Figure F.4) amplifies the small photo-current to a larger voltage by using an operational amplifier.
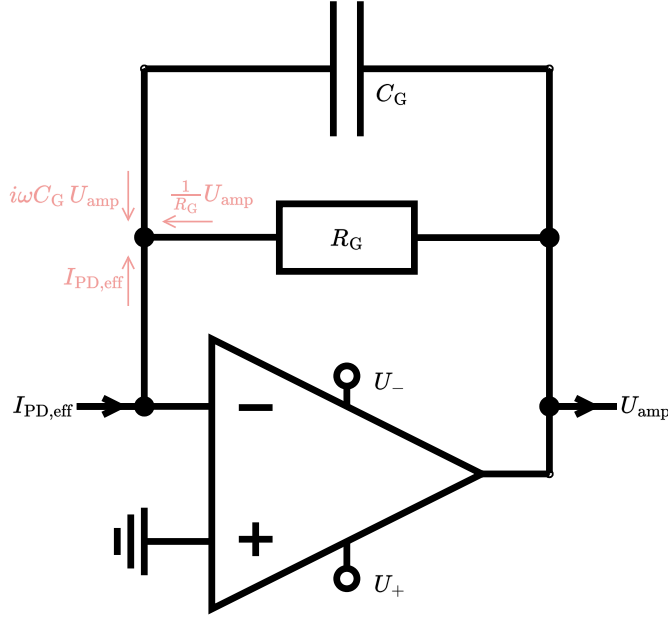
Figure F.4: Sketch of a simple trans-impedance amplifier with inverted output. The value of the resistor R_G corresponds to the gain of the amplifier. Connecting incoming signal $I_{PD,eff}$ and the feedback current $(\frac{1}{R_G} + i\omega C_G)U_{amp}$ to the negative port of the OpAmp leads to an inverted (negative) voltage output.

Applying Kirchhoffs law for the arising currents (shown in red in Figure F.4) allows on to calculate the transfer function directly as:

$$0 = I_{PD,eff} + \frac{1}{R_G}U_{amp} + i\omega C_G U_{amp} \iff \frac{-R_G}{1 + i\omega R_G C_G} \cdot I_{PD,eff} = U_{amp} \quad (A.108)$$

The "amplification" or gain of the amplifier circuit mainly scales with the value of the feedback resistor R_G. The feedback capacitance C_G main purpose is to dampen potential resonance frequencies of the OpAmp component. An ideal OpAmp acts as a "large gain" amplifier of the voltage difference between "+" and "-" port with a flat frequency response. Real OpAmps are however not completely flat and, together with the the capacitance of the connected cable, can have resonances (usually at high frequencies above $> 1\,MHz$). To damp them "away", a small capacitance is added.

In my amplifier circuit, the incoming photo-diode signal is in a range of $0 \leq I_{in} \leq 1\,\mu V$. With supply voltages of $V_{\pm} = \pm 5\,V$, I chose a value of $R_G = 3.32k\Omega$ and $C_G = 100\,nF$, which lead to an ideal gain of $3320\,V/A$.

F.3 "Sallen key" low-pass filter stage

The frequencies of interest for in our DFMI setup are relatively low (multiples of $500\,Hz$ for the DFM harmonics) compared to the signal sampling frequency of $f_S = 1\,MHz$. To reduce the readout noise we add a second order "Sallen-Key" low-pass filter to filter out additional high-frequency signals. Applying Kirchhoffs laws to the occurring currents and Voltages (ignoring the OpAmp for now) yields the 3
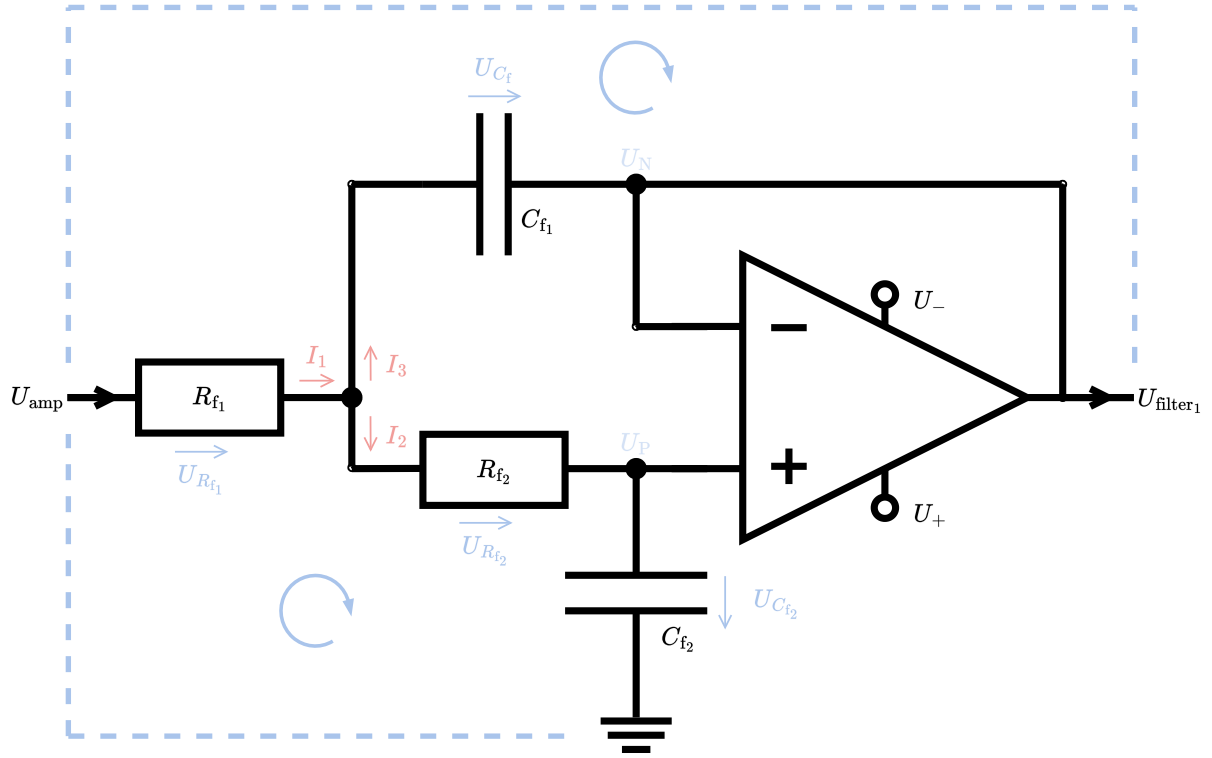
Figure F.5: Some text

equations:

$$U_{\text{amp}} - U_{\text{filter}_1} = R_{\text{f}_1} I_1 + \frac{1}{i\omega C_{\text{f}_1}} I_3 \tag{A.109}$$

$$U_{\text{amp}} = R_{\text{f}_1} I_1 + \left(R_{\text{f}_2} + \frac{1}{i\omega C_{\text{f}_2}} \right) I_2 \tag{A.110}$$

$$I_1 = I_2 + I_3 \tag{A.111}$$

In a stable state the Voltage level at the positive (U_{P}) and the negative input (U_{N}) of the OpAmp must be identical ($U_{\text{P}} = U_{\text{N}} = U_{\text{filter}_1}$) which yields an additional equation between the input voltage U_{amp} and the voltage the junction of U_{P}:

$$U_{\text{amp}} - U_{\text{P}} = U_{\text{amp}} - U_{\text{filter}_1} = R_{\text{f}_1} I_1 + R_{\text{f}_2} I_2 \tag{A.112}$$

Solve this set of 4 linear equations (and eliminating I_1, I_2 & I_3) yields the (well known) transfer function of this Sallen key topology:

$$\tilde{U}_{\text{filter}_1}(\omega) = \frac{1}{1 + i\omega C_{\text{f}_2}(R_{\text{f}_1} + R_{\text{f}_2}) - \omega^2 R_{\text{f}_1} R_{\text{f}_2} C_{\text{f}_1} C_{\text{f}_2}} \cdot \tilde{U}_{\text{amp}}(\omega) \tag{A.113}$$

$$= \frac{\omega_{\text{f}}^2}{\omega_{\text{f}}^2 + i\frac{\omega_{\text{f}}}{Q}\omega - \omega^2} \cdot \tilde{U}_{\text{amp}}(\omega) \tag{A.114}$$

$$= \frac{\omega_{\text{f}}^2}{\left(\frac{\omega_f}{2Q}(i - \sqrt{4Q^2 - 1}) + \omega \right) \left(\frac{\omega_f}{2Q}(i + \sqrt{4Q^2 - 1}) + \omega \right)} \cdot \tilde{U}_{\text{amp}}(\omega) \tag{A.115}$$

$$\tag{A.116}$$

with $\omega_{\mathrm{f}} = \sqrt{R_{\mathrm{f}_1} R_{\mathrm{f}_2} C_{\mathrm{f}_1} C_{\mathrm{f}_2}}$ and $Q_f = \frac{\sqrt{R_{\mathrm{f}_1} R_{\mathrm{f}_2} C_{\mathrm{f}_1} C_{\mathrm{f}_2}}}{C_{\mathrm{f}_2}(R_{\mathrm{f}_1} + R_{\mathrm{f}_2})}$. Figure F.6 shows the ideal and the measured transfer function (measured with a "MokuPro") for the amplifier circuit I build.

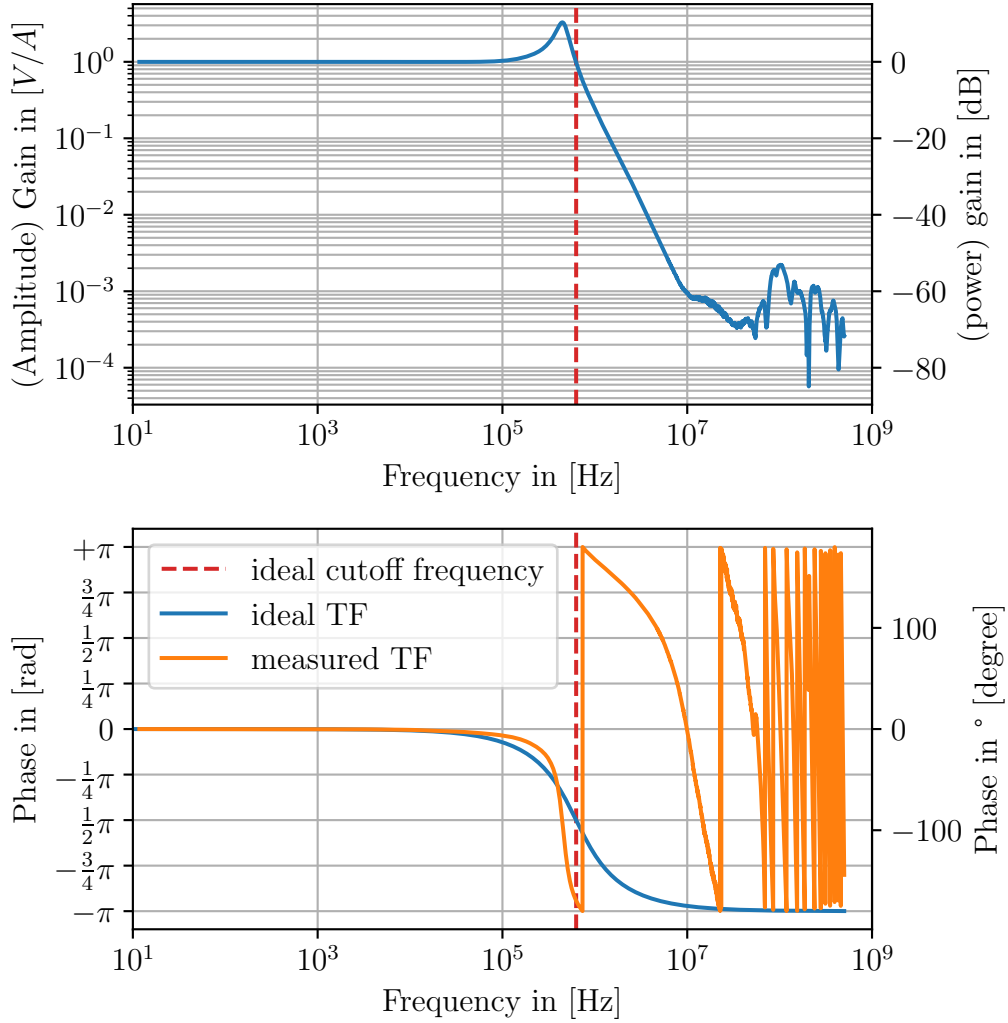## Bode-Plots of the transfer-function of the Sallen-Key Filter stage

Figure F.6: Plot of the ideal and the measured transfer function of the Sallen-key low-pass filter build. The circuit design is shown in Figure F.5 and was realized with the values $R_{\mathrm{f}_1} = 30\,\mathrm{k\Omega}, R_{\mathrm{f}_2} = 18\,\mathrm{k\Omega}, C_{\mathrm{f}_1} = 100\,\mathrm{pF}, C_{\mathrm{f}_2} = 47\,\mathrm{pF}$.

In the full electronics chain two identical Sullen-key low-pass filter are chained one after the other. The transfer function between $\tilde{U}_{\mathrm{amp}}$ and $\tilde{U}_{\mathrm{amp}} \mapsto \tilde{U}_{\mathrm{filter}_1} \mapsto \tilde{U}_{\mathrm{filter}_2}$ is then just the square of (A.116):

$$\tilde{U}_{\mathrm{filter}_2}(\omega) = \frac{\omega_{\mathrm{f}}^4}{\left(\frac{\omega_f}{2Q}(i - \sqrt{4Q^2 - 1}) + \omega\right)^2 \left(\frac{\omega_f}{2Q}(i + \sqrt{4Q^2 - 1}) + \omega\right)^2} \cdot \tilde{U}_{\mathrm{amp}}(\omega)$$

$$(A.117)$$

F.4 Offset stage and Inverter stage

The ADCs we use ("LTC2378-16") have a fully differential input with one positive and one negative pin and digitize the voltage difference ΔV between these to pins with values between $\Delta V \in [-5\,\text{V}, -5\,\text{V}]$. Since the initial PD signal only provides signals between $0\,\text{A}$ and $\approx 2\text{mA}$ and the following amplifier stage only multiplies this signal with some gain factor G, I needed to add an offset and convert the signal from single ended to differential in order to make use of the full $[-5\,\text{V}, -5\,\text{V}]$ Voltage range of the ADC.

At the end of the filter stage the signal scales between zero and the supply voltage so that $\tilde{U}_{\text{filter}_2}(\omega) \in [0, -5\,\text{V}]$. In the offset stage I add an (inverting) offset of $+2.5\,\text{V}$ (yielding $U_{\text{out}} \in [-2.5\,\text{V}, +2.5\,\text{V}]$) and invert the signal to get an additional output of $U_{\text{out, inv}} \in [+2.5\,\text{V}, -2.5\,\text{V}]$. The voltage difference that the ADC digitized is then in the range of $U_{\text{out}} - U_{\text{out,inv}} \in [-5\,\text{V}, +5\,\text{V}]$.

During the first measurements using the circuit presented here, an excess noise was detected which was ultimately the result of internal resonances of the non-ideal OpAmps. Additional capacitors (shown in grey in the following circuit diagrams) were added later to dampen them and reduce the excess electronic noise. For the idealized circuit design, I neglect their influence (and the non-ideal behavior of the OpAmp) in this section.

F.5 Offset stage

To add a stable offset voltage we used a specialized chip (a "MAX6126AASA50+") which provides a stable $+5\text{V}$ reference voltage. (For the COBRI PDs this voltage is regulated down with a very stable voltage divider to 0.83V). The voltage reference Chip can however only supply a very small current and since the same voltage is used as reference for up to 16 readout Channels, I use an additional Operational Amplifier as an (inverting) buffer to not draw too much from the (reference) offset supply. Figure F.7 shows the circuit diagram of this (inverting) offset stage. Like before, the transfer function can be easily read off by applying Kirchoffs current law:

$$\frac{1}{R_{\text{offset}}}U_{\text{filter}_2} + \frac{1}{R_{\text{offset}}}U_{\text{offset}} + \frac{1}{R_{\text{offset}}}U_{\text{out}} = 0 \quad \Longleftrightarrow \quad U_{\text{out}} = -\left(U_{\text{filter}_2} + U_{\text{offset}}\right) \tag{A.118}$$

F.6 Inverter stage

Finally, the ADC we use have a fully differential input with one positive and one negative pin and digitize the voltage difference ΔV between these two pins.

$$\frac{1}{R_{\text{inv}}}U_{\text{out}} + \frac{1}{R_{\text{inv}}}U_{\text{out,inv}} = 0 \quad \Longleftrightarrow \quad U_{\text{out,inv}} = -U_{\text{out}} \tag{A.119}$$

F.7 The used ADC

In the experimental setup, using the MTCA infrastructure discussed in the next chapter, an ACQ425ELF Board (from the company D-TACQ) was used to measure
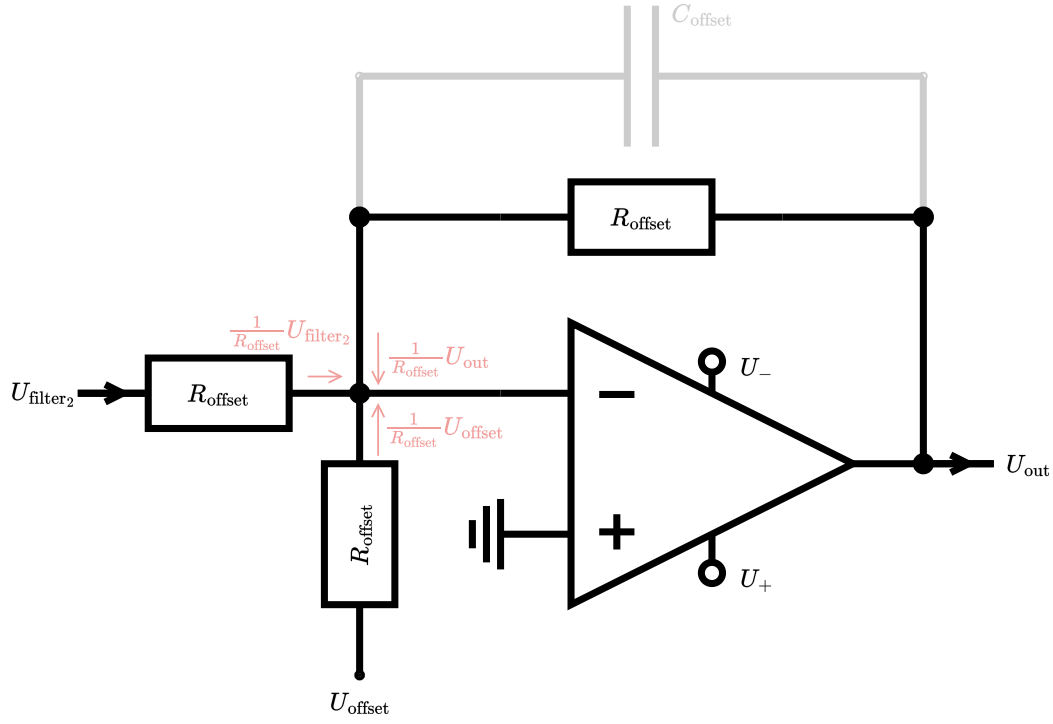
Figure F.7: Circuit Diagram of the "Offset stage". By using an OpAmp as inverting buffer we prevent drawing too much power from the Offset Voltage source.
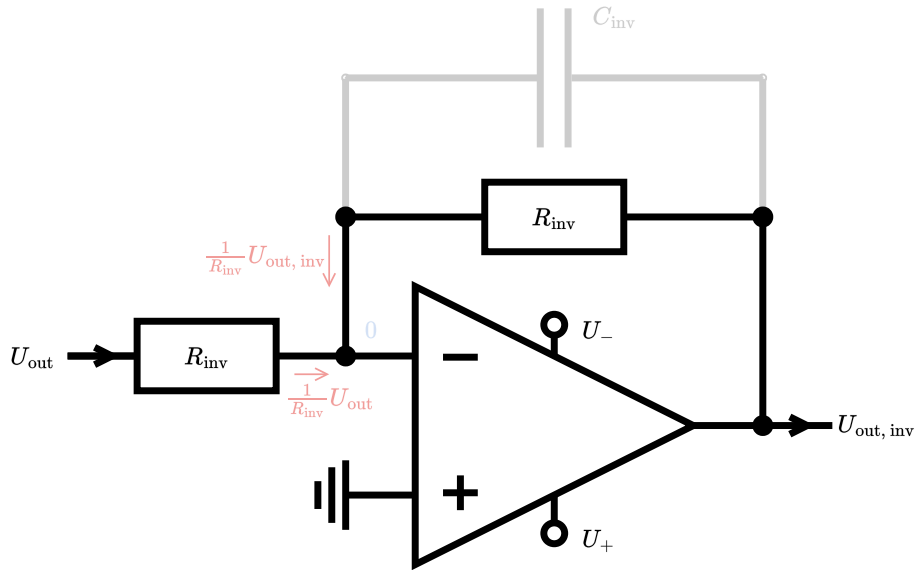

Figure F.8: Circuit Diagram of the "Inverter stage".

the signals. One the ACQ425ELF board are 16 "LTC2378-16" ADCs which are connected by a "VHDCI" (half-inch 68-pin D-Sub) connector to an electronic amplifier and filter board (Figure 10.4).

In its datasheet, the "LTC2378-16" ADC is specified to have a SNR of 97 dB and a total harmonic distortion of ≤ -103 dB for signals around 2 kHz. For an ideal harmonic with an (maximal) amplitude of ≈ 10 V (2×5 V at the end of the differential

amplifier), the given SNR translates to a white-noise floor of

$$\sigma_{\text{ADC-whitenoise}} \approx 0.14\,\text{mV} \quad \Longrightarrow \quad 1.2 \cdot 10^{-14} \frac{\text{m}}{\sqrt{\text{Hz}}} \tag{A.120}$$

and

$$P_{\text{THD}} \sim 0.27\% \, P_{\text{ideal}} \, . \tag{A.121}$$

With these values much smaller than the expected white noise from the remaining electronics and the harmonic distortion from the modulation signal (both discussed in chapter 9), the noise characteristic of the used ADCs won't be limiting the readout performance beyond the limitations set by the 16-bit precision.

F.8 Full electronic transfer function

The ideal transfer function between the Photo-current and the digitized signal $U_{\text{ADC}} = U_{\text{out}} - U_{\text{out, inv}}$ can now be written as:

$$U_{\text{ADC}} = \underbrace{\frac{\omega_{\text{f}}^4}{\left(\frac{\omega_f}{2Q}(i - \sqrt{4Q^2 - 1}) + \omega\right)^2 \left(\frac{\omega_f}{2Q}(i + \sqrt{4Q^2 - 1}) + \omega\right)^2} \frac{R_{\text{G}}}{(1 + i\omega R_{\text{G}} C_{\text{G}})}}_{=: H_{\text{el}}(\omega)} \cdot I_{\text{PD, eff}} - 2 U_{\text{offset}}$$

$$\tag{A.122}$$

This transfer function has both a specific gain and phase delay for different (harmonic) frequencies that distort the demodulated harmonics and must be corrected for, to reach a high precision readout.

While the ideal transfer function written above is a helpful for modeling the circuit and understanding its general behavior, to get the gain and phase factors of the harmonic frequencies, I use the values derived from the measured transfer function of the real electronics (which also includes the build tolerances).

G Hardware and Software infrastructure in detail

After amplifying and filtering the analog signal by the electronic circuit discussed in the previous chapter, the signal in form of an electric current is next digitized by an ADC and the analytic readout algorithm presented in chapter 4 runs on an (embedded) system.

The ADCs that I use are integrated components in a larger (MTCA) infrastructure system that also contains computational units with FPGAs and CPUs. A large part of my Ph.D. studies revolved around understanding and commissioning of this computing infrastructure, and specifically the usage of "FPGAs" and the implementation of the readout algorithm thereon. This chapter will give the reader an overview of the used hardware and explain their basic functionality. Together with the specific data-sheets of the individual components, it should be possible to recreate my setup with reasonable effort.
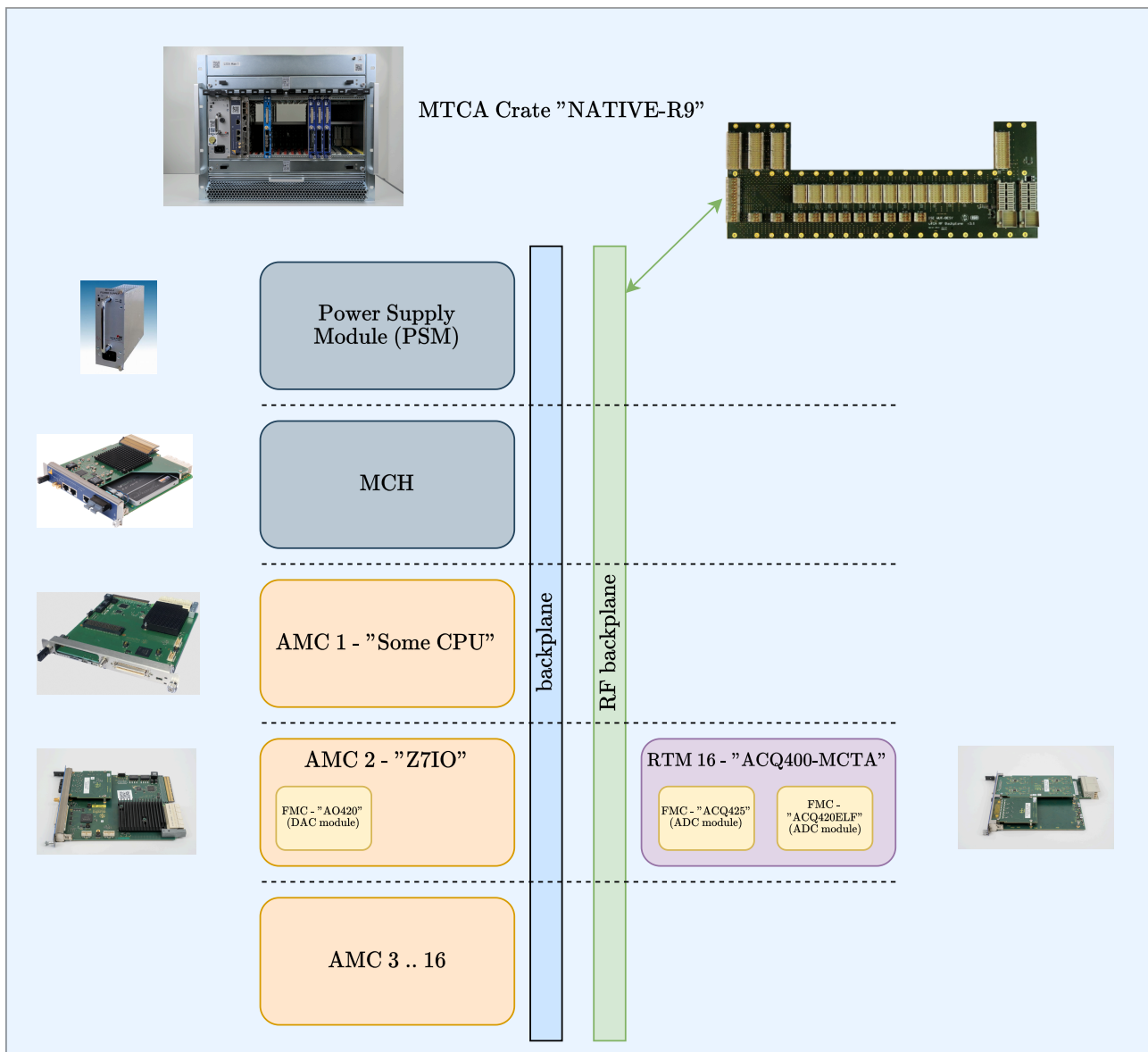
Figure G.1: Schematic overview of the MTCA hardware infrastructure that I use.

G.1   MTCA Infrastructure

The infrastructure I used is based on the "Micro Telecommunication and Computing Architecture (MTCA)" open industry standard. MTCA is a modular Infrastructure standard specifying "mechanical, electrical, thermal and management" properties of the hardware. The ADCs and the computing hardware that I use are (all) build by a company called "D-TACQ" following (mostly) the MTCA standard. Figure G.1 shows the schematic overview and rough hierarchical structure of this hardware. From top (MTCA Crate) to bottom (ADC module), the components that I use are:

1. MTCA Crate "NATIVE-R9" to as frame for the hardware

2. PSM & MCH for basic functionality

3. A "AM-G6x/msd" AMC board for data processing and as network (ethernet) server

4. A "Z7IO" AMC board for data acquisition, control of the hardware and running the readout algorithm

5. A "ACQ420MTCA-RTM" RTM board to as extension for the "Z7IO" with addition FMC slots

6. 3 FMC cards ("AO420FMC", "ACQ425FMC" and "ACQ420ELF") which contain the ADCs and DACs that I use.

G.1.1 MTCA Crate

The top layer of the MTCA infrastructure is the the MTCA Crate holding the MTCA components. In my experiments, I used a "NATIVE-R9 Crate" that is shown in Figure G.2. Each MTCA Crate has at least one dedicated slot for a "Power supply module" (PSM), one for an "MTCA Carrier Hub (MCH)" management board, a number of slots on the front and the back for additional modules following the "AMC" and "RTM" standard and a connection plane ("backplane") providing a physical (PCI) connection for all "AMC" cards. The power supply mod-

Figure G.2: Pictures of a partially filled "NATIVE-R9" (MTCA) Crate. The left picture shows the front side and the right the back side of the crate. From left to right of the front side one can see: the PSM, the MCH, an XYA AMC module, two empty slots, the "z7io" AMC module I use, 5 more open slots, and after 3 more "z7io" AMC modules. In the rear / back side, from left two right are: the PSM, the MCH, empty slots, and a single RTM connected to the z7io" AMC that I use.

ule provides power for the crate and all connected components. The MCH is the "management board" of the crate and allows among other things to monitor the individual components, regulate their power and configure the connections between all components. The backplane of the crate, controlled by the MCH, can be for example be set up such that two plugged modules are directly connected by some PCI connection (which is something I made use of during my measurements).

With the MTCA Crate, the PSM and the MCH provide the infrastructure basis to run any MTCA compliant component. More specific functionality for applications comes from the connected "AMC" (and "RTM") modules.

G.1.2 AMC modules

"AMC" stands for "Advanced Mezzanine Cards" which specifies i.e. the mechanical form factor (to fit into an MCTA Crate), the electrical properties including the "back" connector (which is connected to the backplane of the crate), power specifications, and a minimal management system in form of an "Intelligent Platform Management Bus" (IPMI) system which interfaces with, and can be accessed from the MCH. I.e. MTCA compliant components usually contain an onboard i2c interface and one or more memory chips containing information like i.e. name of the board, serial number, manufacturer etc. which can be read from the MCH (with its i2c controller).

During my work I made use of two specific AMC modules plugged into my MTCA Crate.

G.1.3 The Z7IO (my computing and FPGA unit)

The primary component that I worked with was an "Z7IO" AMC board which si shown in Figure G.3. It features a Xilinx 7-series SoC which combines an ARM CPU with an FPGA (see G.4.3 for details) which is the main computational unit I use to implement the analytic readout algorithm and control all other peripherals like the ADCs and DACs. Besides the SoC and all the MTCA specific components, it has one "FMC" slot connecting a plugged FMC module directly to the 7-series SoC. During my work with the Z7IO, I used a prepared Linux distribution provided
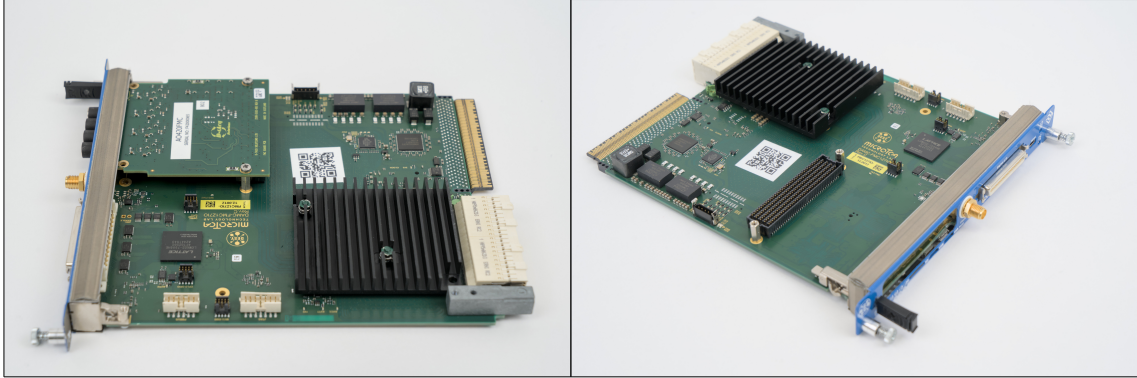

Figure G.3: Pictures of a "Z7IO" AMC board. The picture to the right shown the 'empty' Z7IO with a free FMC slot. In the left picture an FMC extension card is plugged into the board.

by the DESY-MTCALab (which originally developed the Z7IO), that ran on the ARM Core(s) of the SoC, which made it very convenient to use. I.e. I developed and tested most of my (python) scripts on my own PC and simply copied the script over to be able to run it on the SoC. While I needed to adjust (install or replace) specific software packages a few times, the (software) transition is almost seamless and an advantage of the setup. A minor downside is that certain computational task like i.e. computing graphical outputs are very inefficient on the ARM CPU compared to other systems with an integrated or a dedicated GPU.

G.1.4   A "AM-G6x/msd" (AMG6) (an additional CPU/Computer module)

Besides the Z7IO, I also made use of a second "AM-G6x/msd" (short: AMG6) AMC which contained a dedicated CPU featuring more performance than the Z7IOs CPU. I.e. I unloaded demanding computational task like the plotting of data to this CPU and tested the readout algorithm for large sets of data to get results faster / at a higher rate.

G.1.5   RTM modules

"Rear Transition Modules" (RTMs) are designed to be connected to an AMC module, as shown in Figure G.4. Where AMC modules are plugged into the "front" of an MTCA Crate, RTM modules are plugged the back. The purpose of RTMs is
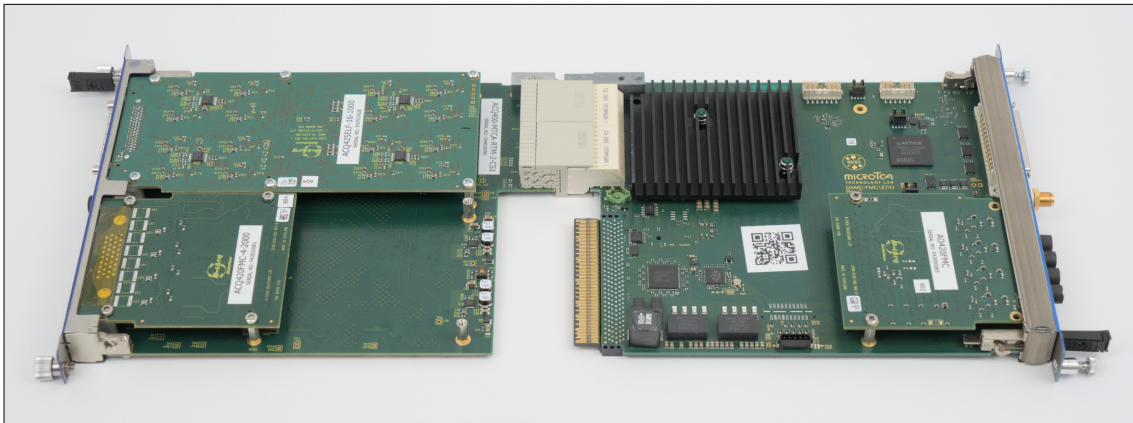

Figure G.4: Picture of an AMC module (a "Z7IO", right) connected to an RTM module (a "ACQ420MTCA-RTM", left), outside of an MTCA Crate.

to provide an extension platform for AMC modules and are used i.e. to separate sensitive electronic parts and provide an analog front end to the AMC. In my case I used a "ACQ420MTCA-RTM" RTM module (shown in Figure G.5) developed by D-TACQ. This RTM provides (almost) no functionality on its own but only serves as extension card providing two more FMC connector slots to the AMC. Besides the
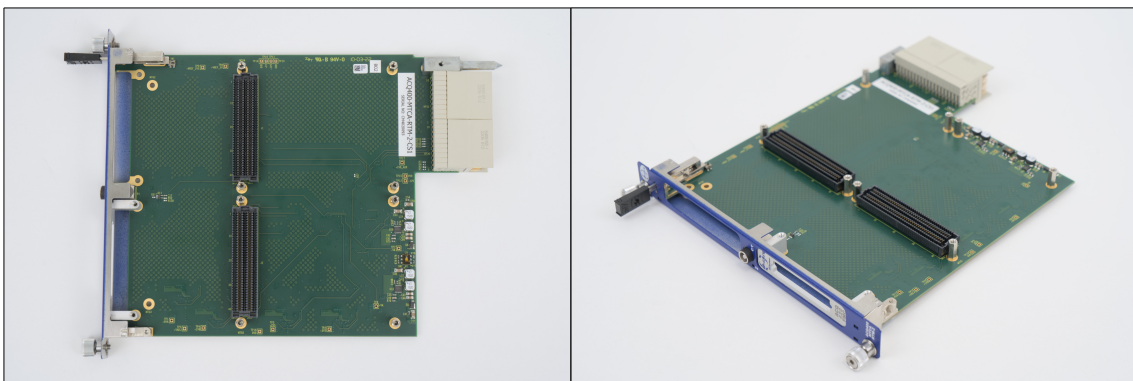

Figure G.5: Pictures of a "ACQ420MTCA-RTM" module from D-TACQ (Company).

these FMC slots the "ACQ420MTCA-RTM" contains a handful of microcontrollers for monitoring, for switching on some of the power lines of the FMC slots, and an

additional I2C and SMBus I/O Expander connected to some of the RTMs FMC pins to the RTM connector ("Zone-3" connector). As deviation of the MTCA standard, D-TACQ uses some signal lines of the RTM connector as additional power lines which makes it generally not compatible to all MTCA compliant AMCs.

G.1.6 FMCs (my ADC and DAC modules)

"FPGA Mezzanine Cards" (FMCs) are another, not MTCA specific, standard for a type of electronic (mezzanine) board and connector (the FMC connector), designed to be connected to an FPGA. I worked with 3 distinct FMC, cards which contained the ADCs and DACs I used, all of the connected to the same FPGA.

G.1.7 ACQ420FMC – A 4 Channel ADC board

The first FMC that I worked with the "ACQ420FMC" shown in Figure G.6. It contains 4x 16-bit ADCs ("LTC2380-16") and allows a sampling speed up to 2 MHz. The communication between the ADCs and the FPGA works via a custom SPI-like interface which I build into the FPGA fabric. It should be noted that this specific module has not a very robust timing and every channel needs to be adjusted with specific delays to achieve a continuous error free readout from the ADCs.
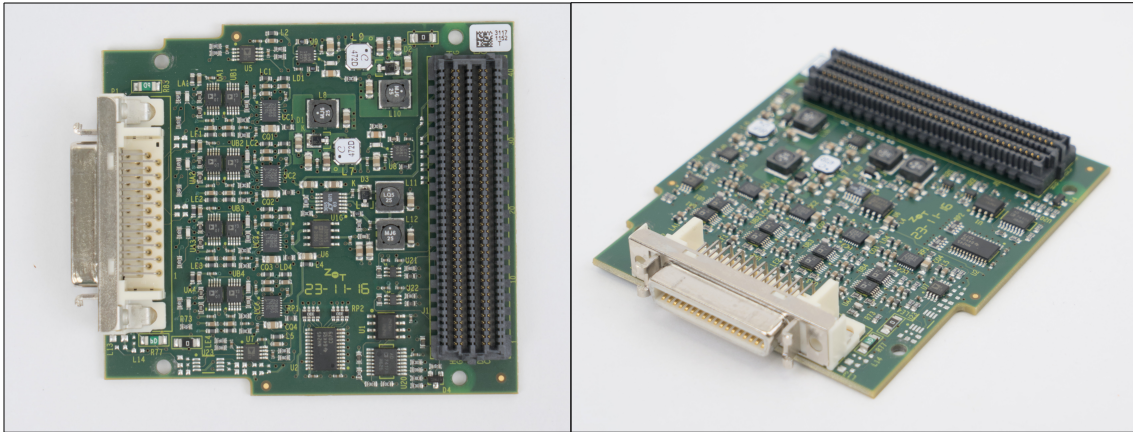

Figure G.6: Pictures of the "ACQ420FMC" FMC.

G.1.8 ACQ425ELF – A 16 Channel ADC board

The second and main FMC that I used was the "ACQ425ELF" board as shown in Figure G.7. ELF stands for "Electrically Extended FMC" which is a term used for a series of products from "D-TACQ" and it is not compatible to regular FMC modules. ELF modules use the same physical FMC connector and share many of the pins with FMC modules, but are generally not compatible to standard FMC connectors. The "ACQ425ELF" uses i.e. additional pins to power the board. While regular FMC modules (at least the other two mentioned here) run on "ELF" slots (like the two provided by the "ACQ420MTCA-RTM"), the inverse, ELF modules plugged into standard FMC slots, does not work. Function wise, the "ACQ425ELF" work very similar to the "ACQ420FMC", the key differences are that it provides 16 separate ADC Channels ("LTC2378-16") and has a maximum sampling speed of 1 MHz.
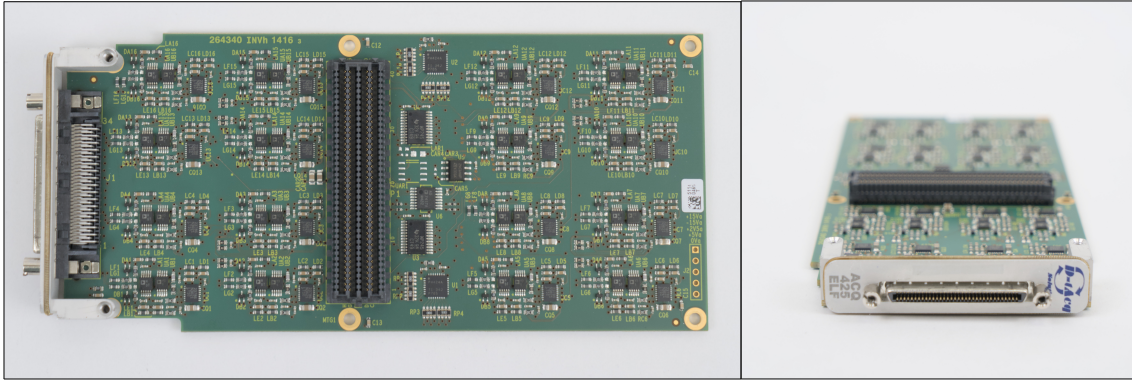
Figure G.7: Pictures of the "ACQ425ELF" module.

G.1.9 AO420FMC – A 4 Channel DAC board

The last FMC I worked with was the "AO420FMC" shown in Figure G.8. It is a 4 Channel DAC board that I plan to use as output for the DFMI laser modulation signal and possible control & error signals later on. Function wise the DACs have, like the ADCs before, an SPI interface for communication. They require however a more elaborate setup routine than the ADCs (which work with minimal configuration).
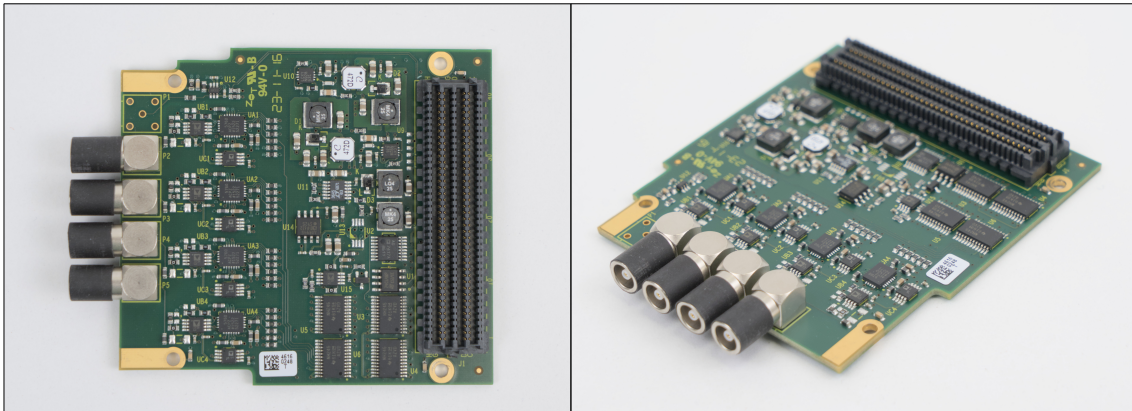

Figure G.8: Pictures of the "AO420FMC" FMC.

G.2 Field-Programmable Gate-Arrays (FPGAs)

While I started my initial FPGA work on a "ZC706" development board designed by Xilinx, I later switched to a "Z7IO"-AMC board designed by DESY which was designed in compliance with their MTCA.4 standard. Both boards contain a "7-series FPGA" from Xilinx (now AMD) of their ZYNQ7000 product series which is an SoC containing both an FPGA part and an ARM CPU (+Peripheral connections).

The two major reasons, and main motivation, to use an FPGA for my DFMI readout are:

1. **A high readout speed (due to parallelization and specialized hardware modules)**
   In the DFM readout algorithm as presented in chapter 4 a (potentially) large

number of harmonics need to be demodulated and filtered. Especially the
demodulation is a computationally expensive task for which many DSPs or
MCUs have dedicated harware modules. Implementing the demodulation and
the filtering of the harmonics in an FPGA allows potentially for a very efficient
calculation of analytic readout algorithm within a few ms.

2. **Stable synchronous timing for operating (multiple) control systems**
   Due to the planned implementation of various control loops to regulate i.e.
   the laser frequency and amplitude modulation, using an FPGA as custom
   controller is advantageous as it allows me to regulate multiple in- and output
   synchronous to the same clock.

Before diving into the details of the FPGA chip I used and my implementation,
I want to give a brief overview of what FPGA are briefly explain their working
principles.

G.3   Integrated Circuits as collection of logic gates

At their lowest hardware level, any integrated circuit (CPU, GPU, FPGA, MCU,
DSP and so on) is a collection of transistors and resistors form flip-flops that together
form logic "gates".

G.3.1   Logic gates

A logic gate is a device (circuit) that perform a single boolean (bit-wise) logic op-
eration like i.e. AND, OR, XAND, XOR etc. Figure G.9 shows an example "AND"
gate and its function commonly expressed as "truth table". Physically, these '0's
and '1's correspond to some chip specific internal voltage levels of e.g. 0V and 1.5V.
From these very basic logic Gates, more complex circuits are build performing larger,
more complex operations like the addition of two 64-bit integers or even the calcu-
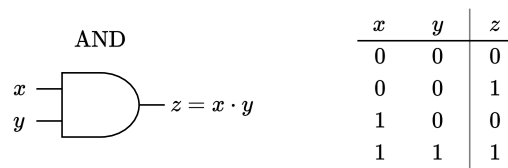lation of the FFT of an array of 32-bit numbers in specialized microchips.

x	y	z
0	0	0
0	0	1
1	0	0
1	1	1

AND

x —
y — $z = x \cdot y$

Figure G.9: Example of a single logic "AND" gate. The left shows the symbolic repre-
sentation of an AND gate and the right shows the "truth table" of the operand. Arith-
metically, this single AND gate can already be seen as a 1-bit multiplicator.

When designing larger function blocks like i.e. the 4-bit Adder as pictured in Figure
G.10, the hardware designer also needs to take timing considerations into account.
I.e. the output of a logic gate (or generally flip-flop) does not change instantaneously
with the input but has a short delay, the so-called "setup time". The setup time
is the time it takes between a stable input signal to lead to a stable output signal.
And after the output signal is stable, the minimum time that it remains valid is

called the "hold time". In the case of this 4-bit Adder example, the designer has to make sure that for synchronous input of 4 bits, all 4 bits of the output are valid at roughly the same time. Otherwise, the measured output bits would not correspond to a single specific input and this Adder would simply not work properly.
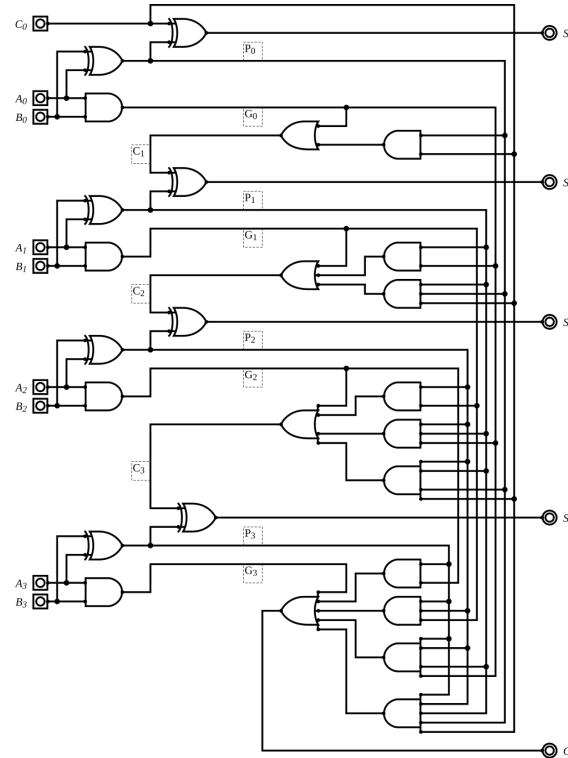

Figure G.10: Example of a 4-bit Adder "hardware module" build from logic gates. Source: Wikimedia

Some common examples for integrated circuits build from these logic gates are:

- **Hardware module or "IP Core"**
  Larger logic arrays performing a specific function are sometimes just referred to as "hardware modules" (like the Adder block mentioned before). In a more specific FPGA environment (explained below), the term "IP" or "IP Core" is also used to refer to individual design blocks build from "logic gates". IP stands here for "Intellectual Property" that is the specific hardware design.

- **MCU**
  Microcontroller unit (MCU) and Microprocessors is a broad category for usually small processors (smaller than a CPU or GPU) performing various tasks from managing a specific interface (like SPI, Ethernet, etc.) of a small microcontroller to more general purpose microprocessors running a supplied program within a control system.

- **DSP**
  Digital signal processors are microprocessors usually used for computational "heavy" task and often include i.e. FTT function blocks to speed up common calculation in signal processing which would take much more time on not specialized microprocessors.

- **GPU**
  Graphics processing unit are specialized for parallel calculations of arrays of (usually less precise) data like the (color) values of the number of pixels on a screen.

- **CPU**
  Central processing units are designed to be versatile ICs being able to perform various different operations depending on their input instructions. In some abstract sense CPUs are electric circuits that take an "instruction" in form of a specific input sequence of bits "0110101" and perform a predefined task like adding two 32-bit numbers (written in some specific location in memory). The sum of all these possible inputs (tasks) is called its "instruction set". When creating a computer program, the programmer writes code in a programming language of their choice at this code is translated by a compiler (another program) into "machine code" i.e. into the CPU specific instruction set. This instruction sets differ for different CPUs (and different generations) and even newly released CPUs today likely have a slightly different (extended) instruction set than CPUs from the same manufacturer from a few years ago. Some of the most used instruction set are the "x86-64" (or "AMD64") instruction set that most PCs use and the "ARM" instruction sets (used in many embedded and mobile CPUs), and a multiplicity of versions (i.e. ARMv1 to ARMv9) and extensions to these sets (i.e. the "Intel AVX-512" instructions for some "x86-64" CPUs from Intel). In hardware, a single instruction often corresponds to a specific hardware module performing the intended operation.

For CPUs or (micro)processors in general, the circuit made from such logic gates is created/burned fixed into some silicon substrate and does not change any more. The goal of an FPGA is to provide or at least emulate the behavior of an array of configurable logic gates.

G.4 FPGA definition

A Field-Programmable Gate-Array (FPGA) is a type of integrated circuit that emulates the behavior of a dynamically configurable array of logic gates. Unlike the "fixed" ICs mentioned before, where the logic gates are build fixed in form of physical present arrays of transistors; modern FPGA emulate the behavior of a these logic gates by a tiny piece of memory (SRAM) holding the truth table of the emulated logic gate[76]. For given inputs to the logic cell, the output will correspond to the logic operation written in its "truth table memory". While an idealized FPGA consists only of a number of identical logic cells and blocks dedicated to connecting cells and routing signals, modern FPGAs are more in-homogenious constructs with > 100 "design primitives" that can be used like, different kind of signal buffers (with different "timing" behavior), specialized routing blocks for "clock" signals, larger dedicated memory blocks, basic arithmetic function blocks and more. Theoretically, any kind of the previously mentioned ICs could be emulated with a large enough FPGA with "enough" configurable logic gates.

One of the two main advantages of an FPGA for me is that it allows me to 'build'/emulate a specialized chip performing exactly the calculations that I need

with the potential to run much faster than a comparable DSP or CPU could. E.g. a DSP may include a single FFT function block, but in case I need to calculate multiple FFTs for a single measurement; I will need to use the function block multiple times in sequence which will add a delay until the final calculation results are available. And while modern desktop CPUs are well optimized and fast (performing calculations within a few ns speed), their sequential mode of operation can still lead to slower speeds (and efficiency) compared to a dedicated IC, or a dedicated design realized in an FPGA.

G.4.1 "Programming" an FPGA

The "programming" of an FPGA, differs significantly from programming a "regular" CPU. While there are hardware agnostic languages designed to describe a logic level design, so-called "hardware description languages" (HDLs), programming an FPGA also includes a set of additional vendor specific tools to fully integrate a more abstract logic level design into a specific FPGA chip. "Programming" an FPGA specifically requires tools to manage the timing and delays of signals traveling physically through the IC, which does not exist for other types of programmable ICs where the hardware designer already ensure a proper propagation of signals within the chip. In my case of using a Xilinx/AMD FPGA, the vendor specific toolset comes in the form of the "Vivado" software suite.

HDL

As hardware description language I worked mostly with "VHDL" ("Very High Speed Integrated Circuit Program Hardware Description Language"). While (micro) processors usually some fixed length register (i.e. 8-bit or 64-bit) as smallest unit. Representing other objects like 'characters' or 'floats' with these fixed length registers, the natural unit in an FPGA is a single bit ('0' or '1') and all other types are derived from them. In VHDL this is presented by the "std_logic" type as single bit and "std_logic_vector" as array of bits, which would written in VHDL as:

```
signal my_enable : std_logic := '0';
signal my_data   : std_logic_vector(15 downto 0) := (others => '0');
```

While there exists additional type definitions in commonly used libraries like i.e. "integers" derived from std_logic_vectors, these types are merely renamed constructs with limited functionality. There is i.e. not a fixed length for an "integer" in VHDL and and only simple arithmetic operations like an addition between two integers are included in the standard numeric library. In what is called the "synthesis" step of the FPGA programming process, written HDL code is compiled/translated into the logic gate structure / the hardware primitives available for the used FPGA chip.

Timing constraints (with TCL)

The biggest difference between "programming" FPGAs and a processor lies however in the additional complexity for FPGAs in the signal timing. For processors, the engineer designing the circuit manages the runtime delays introduced by different logic gates or larger function blocks of the signal; making sure that the output of some function block (and by extension the entire processor) is available at a specific fixed time (given i.e. by a fixed number of processor clock cycles) later.

For FPGAs, it is the task of the "programmer" / hardware designer to manage the delays of signals such that all input of a given function block are available at the same time for processing. Within an FPGA, the individual logic cells are also "clocked", meaning that besides the signal inputs, they also have an addition input "clock" signal and the update of their output is triggered on a rising (or falling) edge of this clock signal and stable after a device specific setup time (as mentioned earlier). The signals inside the FPGA also do not travel instantaneously and take a certain time to travel to different (physical) locations on the chip.

Different setup and hold times of the logic cells and actual differently sized and structured chips are the reason why the programming of an FPGA is not entirely platform agnostic. The same hardware module written in HDL will have different signals timings in different physical chips.

During the "Implementation" step of FPGA programming process, the vendor specific FPGA programming tools physically place the logic gate structure into the given FPGA chip layout while trying to minimize the signal delays and keeping multiple signal used by individual function blocks synchronous. While these software tools take care of the majority of the signal timing, it is often nesseccary to adjust the timing of specific (critical) signals 'by hand'. In Vivado, this is done by specific keyword added to the HDL code and by additional "constraint" files written in the "Tool command language (TCL)" script language that includes Vendor specific commands.

In case Vivado is unable to successfully place a logic part so that all input signals arrive at the same time; a very simple solution would be to write a constraint file, "telling" Vivado to increase the time these signals are allowed to travel until they reach the critical logic block, and in VHDL, add a 'manual' counter to wait for multiple clock cycles until the signals are stable for certain.

G.4.2 The FPGA Programming Process

With the Vivado toolset for FPGAs from Xilinx/AMD, the "programming" of an FPGA is structures into multiple specific steps:

1. HDL source code & the Block design
   After selecting a specific chip, the initial step is to include the HDL coded source files of the project. Besides "hand written" HDL source code, Vivado uses additional "board design" files, which are managed via a GUI. Board design files allow a kind of "visual programming" using prepared IP cores as functional blocks and Xilinx strongly encurages using the Block design for all of their own IP cores which are not freely accessable as source code. From the block design, Vivado generates a wrapper written in the used HDL.

2. Simulation
   The next step is the hardware agnostic simulation of the written HDL source code. Here the basic functionallity and any possible bugs and logical errors can be revealed before continuing.

3. RTL Analysis
   The "Register Transfer Level" (RTL) Analysis does a first translation of the HDL source code into a logic gate structure and nets connecting them.

4. Synthesis (conversion into LUT blocks)
   In the Synthesis step, the logic gate structure is translated into the device specific hardware primitives. This can be seen as an equivalent of translating some hardware agnostic software code written e.g. in C into the chip specific instruction set / assembler code which can vary for different processors (like x86 or ARM).

5. Instantiation
   In the Instantiation step the placement of the Synthesised hardware design on the selected chip is done. Vivado calculated the timing delays of the signals and used hardware components and uses the specified timing constraints.

6. Bitstream generation and programming
   After a successful Instantiation, the design is finally compiled into a machine readable 'bitstream' file which programms the LUTs of the FPGA cells to to emulate the hardware design.

G.4.3   Xilinx (now AMD) 7 series FPGA

The SoCs that I worked with were all part of the "Zynq-7000" product line from (formerly "Xilinx") now "AMD Inc.", first released in 2012. The "Zynq-7000" product line is a family of SoCs that incorporate both a (configurable) FPGA fabric and an "ARM Cortex-A9" CPU-core (together with a set of additional peripheral microcontrollers for various interfaces like i.e. USB, SPI, I2C, etc.). In FPGA designs it is not uncommon to replicate a small CPU core within the FPGA fabric to perform various serial tasks. Xilinx themselves offer a so-called "MicroBlaze™ processor" as IP core for synthesis within their FPGA fabric. Having however a full ARM core connected to the FPGA has some advantages as it is generally faster and easier to program with many pre-made software libraries for the ARM architecture readily available compared to the smaller "MicroBlaze" architecture from Xilinx. Figure G.11 shows a schematic overview of the Zynq-7000 SoC that is the main Chip on the earlier mentioned Z7IO board that I used in my project.

G.4.4   Readout algorithm Implementation into FPGA

From the components shown in Figure G.11, I mostly used the "CPU" in form of the "Application Processing Unit", its "Direct Memory Access (DMA)" Channels and the "Programmable Logic to Memory Interconnect" which connected the CPU to the FPGA fabric denoted as "Programmable Logic".

As default interface within the FPGA fabric and to the DMA Channels, Xilinx Chips use the "Advanced eXtensible Interface (AXI)" which is a close-to-hardware, memory-mapped interface for which Xilinx offers many prebuild components and which I also implemented manually within my own VHDL code.

Figures G.12 shows an overview of the Board design I implemented within the SoCs FPGA fabric. The biggest part of my FPGA work so far was building the various interfaces to the difference MTCA components. In the lower picture of Figure G.12 they are incorporated into the "RTM_ACQ_400_0" IP core (which also holds the two ADC FMCs that I use) and the "AO420FMC_top_0" IP core for the DAC module.
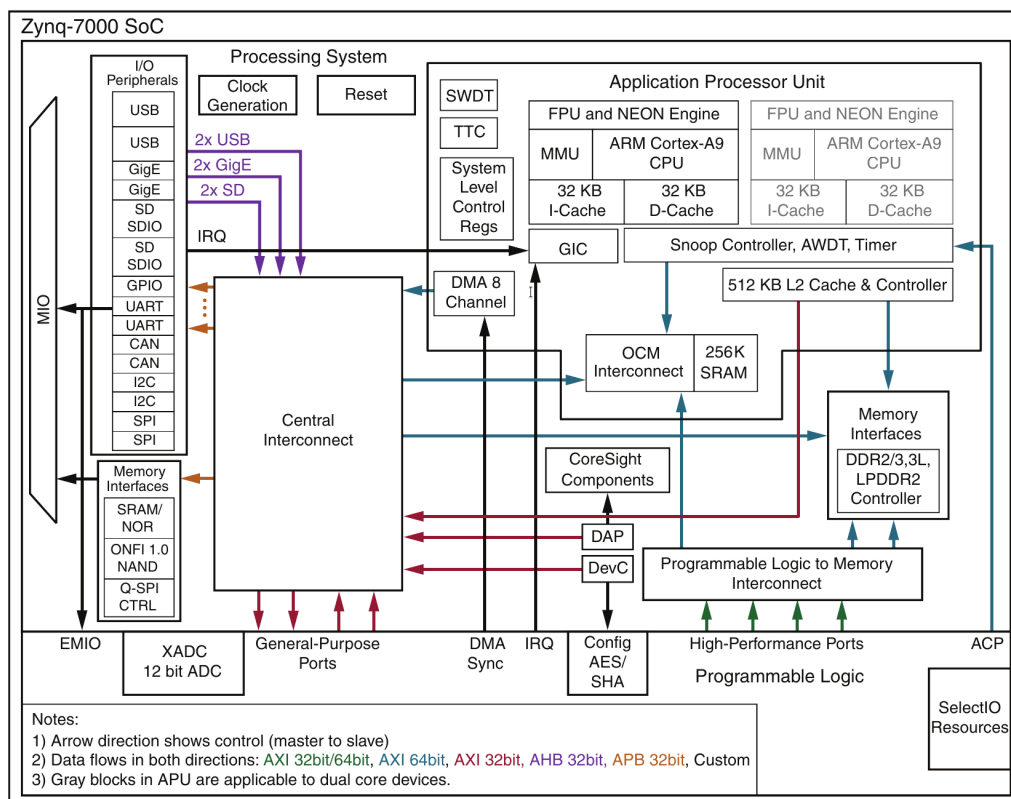
Figure G.11: Schematic overview of the "Zynq-7000" SoC from AMD as printed in the technical reference manual.
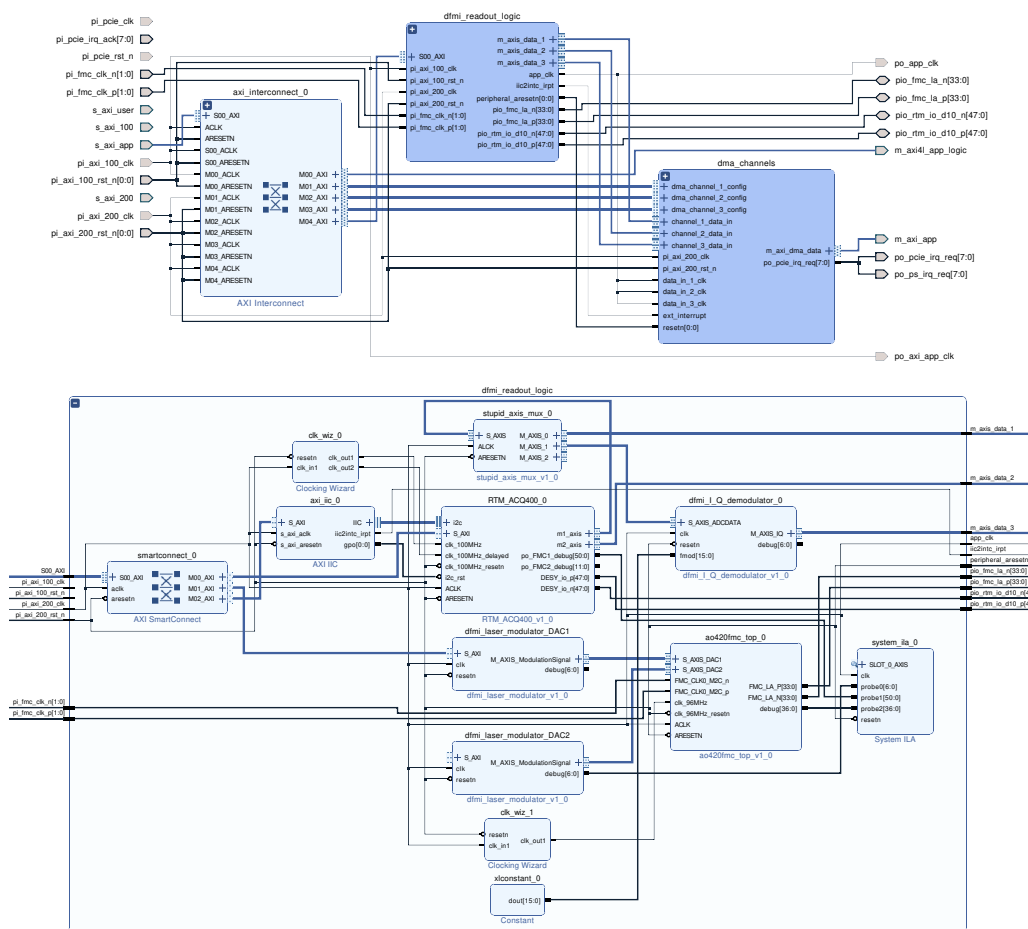
Figure G.12: Top Picture: Overview of the Block Design opened in Vivado used in my FPGA design. Bottom Picture: Zoomed in View of the current "Analytic Readout Algorithm" hardware implementation and the interface blocks to the various MTCA components I use.

Appendix B

Curriculum vitae

Personal information

Name:	Tobias Benjamin Eckhardt
Date of Birth:	19. April 1991
Birthplace:	Braunschweig
Nationality:	German

Education

Jan 2020 – Mar 2025	Studying towards Ph.D. in physics at the Institute of experimental physics of the Universität Hamburg.
Aug 2015 – July 2016	Studying abroad at the Department of Physics of the University of Stockholm as result of an ERASMUS Scholarship.
Oct 2010 – Apr 2017	Graduating with a B.Sc. (2013) & M.Sc. Degree (2017) at the Physics Department of the Technische Universität Braunschweig with focus on "condensed matter physics" and "astrophysics". I conducted my master's thesis in cooperation with the Stockholm University (Prof. Kjell Rosquist) with the thesis topic: "Interaction energies and initial conditions for time evolution in discrete cosmology with a $S^1 \times S^2$ spatial topology."
Aug 2010	Abitur at the IGS Franzsches Feld (Secondary school) in Braunschweig, Focus on Mathematics, English and Politics.

Appendix C

Publications

2022 "An analytic, efficient and optimal readout algorithm for compact interferometers based on deep frequency modulation" (Link: https://www.nature.com/articles/s41598-024-70392-9)
An article a novel readout algorithm for nested sinusodial signals as they appear i.e. in "Deep-Frequency modulation Interferometry".

2024 "Noise Limitations in Multi-Fringe Readout of Laser Interferometers and Resonators." (Link: https://www.mdpi.com/2673-8244/2/1/7)
An article about the limit (Cramér-Rao bound) for the readout of various interferometer techniques.

Appendix D

Eidestattliche Versicherung

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Hilfsmittel und Quellen benutzt zu haben.

Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

14-02-2025
_____   _____
Datum   Unterschrift der Doktorandin / des Doktoranden