

UNIVERSITÀ DEGLI STUDI DI PADOVA

Dipartimento di Fisica e Astronomia “Galileo Galilei”

Master Degree in Astrophysics and Cosmology

Final Dissertation

Ultralight bosons and supermassive black holes:
new fundamental physics from GW observations

Thesis supervisor

Prof. Jean-Pierre Zendri

Thesis co-supervisor

Prof. Andrea Maselli

Candidate

Giovanni Benetti

Academic Year 2025/26

Contents

Abstract	vii
1 Gravitational waves, LISA and Extreme Mass Ratio Inspirals	1
1.1 Gravitational waves: the latest window on our universe	1
1.1.a An introductory overview	1
1.1.b Gravitational Wave stress-energy tensor	3
1.1.c Gravitational Waves experimental observations	4
1.2 Detection and parameter estimation with gravitational waves	5
1.2.a Characterization of detector noise	5
1.2.b A whisper in the crowd: matched filtering and parameter estimation	7
1.2.c Parameter estimation forecasts: the Fisher matrix technique	10
Geometrical interpretation of the Fisher matrix	12
Limitations of Fisher matrix technique	12
1.3 Towards the next generation of detectors	13
1.3.a The audio band: Einstein Telescope and Cosmic Explorer	13
1.3.b The deciHertz window: LGWA and DECIGO	14
1.3.c The milliHertz band: Laser Interferometer Space Antenna	15
1.4 Extreme Mass Ratio Inspirals	16
1.4.a The Kerr metric and geodesic motion	17
1.4.b Astrophysical relevance	19
2 Theories of modified gravity and massive scalar fields	23
2.1 Motivations and examples	23
2.1.a Beyond General Relativity: a classification of the alternatives	23
2.1.b An overview of scalar-tensor theories	24
2.1.c Current estimations and upper limits	26
Theory-agnostic tests of GR	26
Scalar-tensor specific tests	28
2.2 Massive scalar fields around EMRIs	28
2.2.a Gravitational and massive scalar field equations of motion	28
2.2.b Leading and subleading order terms in mass ratio	31
3 Solutions to the field equations in a curved background	35
3.1 Teukolsky formalism for massless fields	35
3.2 Teukolsky equation for massive scalar fields	39
3.2.a Asymptotic behavior of the master equation	41
3.2.b Green's function method for second order ODE	42
3.3 Energy and angular momentum scalar fluxes	43
3.3.a Source term computation	44
Specialization to equatorial circular orbits	45
4 Finding the homogeneous solution	47

4.1	MST method for a massless scalar field	47
4.1.a	The hypergeometric function	47
4.2	MST method for the massive scalar field	49
4.2.a	Recurrence relation for the term $(-x)(1-x)p$	52
	Contiguous relations	52
	Second approach: redefinition of radial function	54
	Behavior for $\eta = 1$	56
	Behavior for $\eta > 1$	56
4.3	Upgoing solution: possible paths for a generalized MST method	57
4.3.a	Direct approach	57
4.3.b	Usage of the double-confluent Heun Equation	58
4.4	Upgoing solution in terms of a Frobenius expansion	59
4.4.a	Implementation of the massive MST method	59
4.4.b	The Frobenius expansion	59
4.5	Numerical integration of the Teukolsky equation	61
	Wronskian computation	62
4.5.a	Spin-weighted spheroidal harmonics computation	62
5	Results: waveforms and observability forecasts	65
5.1	Field solution and scalar fluxes	65
5.2	Energy fluxes, grid sampling and interpolation	66
5.2.a	Energy scalar fluxes with STORM	66
	Data quality and completeness	68
5.2.b	Interpolation	69
5.3	Results: energy scalar fluxes for circular equatorial orbits	69
5.3.a	Theoretical explanation of energy flux features	69
5.3.b	An unidimensional case study	72
5.4	Results: dephasing	75
5.4.a	Definition and implementation	75
5.4.b	Observational thresholds of the dephasing	77
5.4.c	dephasing observability	80
5.5	Results: Fisher matrix forecasts for parameter estimation	82
	Case study 1	85
	Case study 2	87
	Case study 3	89
6	Superradiance and floating orbits	93
6.1	Superradiance: a general introduction	93
	The Klein paradox	93
	Superradiance in fluidodynamics	94
6.1.a	Superradiance for a rotating BH	95
	Superradiance from BH thermodynamics	95
	Superradiance from computed energy fluxes	96
6.1.b	Superradiant resonances	96
6.2	Resonance phenomenology	97
6.2.a	Resonance shape and parametrization	97
6.2.b	Variation under change of eccentricity or inclination	100
6.2.c	Resonance parameters for generic orbits	100
6.3	Dynamical evolution of resonant orbits	103

6.3.a	Resonant but not floating: the uncomfortable situation of generic orbits	103
6.3.b	Methods and qualitative phenomenology of generic resonant orbits	106
6.4	A theoretical explanation for the resonance shape	108
6.4.a	Resonance interference with a wide-band background	108
6.4.b	Field amplitude related to the Green's function method	112
Conclusions		117
A Coefficients of the Frobenius series for ingoing solution		119
Bibliography		123
Acknowledgments		131

Abstract

Gravitational waves (GW) are opening the possibility of probing otherwise inaccessible physical scenarios, with a transformative impact on our understanding of fundamental physics. One of the most important science cases for GW physics is the discovery of an extension of General Relativity (GR); for example, scalar-tensor theories of gravity beyond GR predict the existence of a scalar mediator of gravity other than the usual tensor interaction. The existence of such scalar fields will have possibly observable effects in Extreme Mass Ratio Inspirals (EMRI), highly-informative astrophysical systems which will be observed by the future GW detector LISA.

In this Thesis, I explore the phenomenology of a massive scalar field affecting the EMRI orbital evolution in the perturbative regime, characterizing systematically the effects of the field dynamics using custom analytical and numerical methods. Based on a classical field-theoretical approach, I provide an efficient numerical implementation to compute the field dynamics and its back-reaction on the secondary. I use these results to assess the observability of such effects and forecast the LISA capabilities for estimating the field's parameters. Particular attention is given to the description, modeling, and interpretation of effects specific to massive fields, such as field confinement, global superradiance, and superradiant resonances; for the latter, a thorough investigation of resonant and floating orbits is performed.

Table of acronyms (by alphabetical order):

Denomination	Acronym
BBH	Binary Black Hole
BD	Brans-Dicke (theories)
BH	Black Hole
BNS	Binary Neutron Star(s)
CBC	Compact Binary Coalescence
CE	Cosmic Explorer
DECIGO	Deci-hertz Interferometer Gravitational-wave Observatory
DHCE	Double-Confluent Heun Equation
dHz	deciHertz
DoF	Degrees of Freedom
EM	Electromagnetic
EMRI	Extreme Mass Ratio Inspiral(s)
EoM	Equations of Motion
ET	Einstein Telescope
FAR	False Alarm Rate
FEW	Fast EMRI Waveforms (software package)
FWHM	Full Width at Half Maximum
GR	General Relativity
GW	Gravitational Wave(s)
HD	Hellings & Downs (curve)
IGWN	International Gravitational-Wave Network
IMRI	Intermediate Mass Ratio Inspiral(s)
IPTA	International Pulsar Timing Array
ISCO	Innermost Stable Circular Orbit
KG	Klein-Gordon
LGWA	Lunar Gravitational Wave Antenna
LHS / RHS	Left-Hand Side / Right-Hand Side
LIGO	Laser Interferometer Gravitational-Wave Observatory
LISA	Laser Interferometer Space Antenna
MEW	Modified EMRI Waveforms (software package)
mHz	milliHertz
MST	Mano–Suzuki–Takasugi (method)
NS	Neutron Star
ODE	Ordinary Differential Equation
PPN	Parametrized Post-Newtonian
PSD	Power Spectral Density
PTA	Pulsar Timing Array
QNM	Quasi-Normal Modes
sGB	scalar Gauss-Bonnet
SMBH	Supermassive Black Hole
SNR	Signal to Noise Ratio
STORM	Scalar Tensor Orbital Radiation from EMRIs (software package)
TDE	Tidal Disruption Event(s)
ToA	Time of Arrival
TT	Transverse-Traceless
WD	White Dwarf
WEP	Weak Equivalence Principle

Gravitational waves, LISA and Extreme Mass Ratio Inspirals

1.1 Gravitational waves: the latest window on our universe

1.1.a An introductory overview

Gravitational waves (GW) are a general prediction of a wide class of theories of gravity (Arun et al., 2013; Berti et al., 2015) including General Relativity (GR) (Einstein, 1916; Einstein, 1918; Misner et al., 1973; Maggiore, 2007). They are defined as wave solutions (propagating oscillations) of the field equations describing gravity: in the context of General Relativity, gravity is described by the metric $g_{\mu\nu}$ of a 4-dimensional pseudo-Riemannian spacetime manifold $\mathcal{M}$ through the Einstein Field Equations¹

$$G_{\mu\nu} \equiv R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi T_{\mu\nu} \quad (1.1)$$

where $G_{\mu\nu}$ is the Einstein tensor, $R_{\mu\nu} \equiv R_{\mu\alpha\nu}^{\alpha}$ is the contraction of the Riemann tensor $R_{\nu\alpha\beta}^{\mu} \equiv \partial_{\alpha}\Gamma_{\nu\beta}^{\mu} - \partial_{\beta}\Gamma_{\nu\alpha}^{\mu} + \Gamma_{\sigma\alpha}^{\mu}\Gamma_{\nu\beta}^{\sigma} - \Gamma_{\sigma\beta}^{\mu}\Gamma_{\nu\alpha}^{\sigma}$ built from the Christoffel symbols $\Gamma_{\nu\alpha}^{\mu} \equiv \frac{1}{2}g^{\mu\sigma}(\partial_{\nu}g_{\alpha\sigma} + \partial_{\alpha}g_{\sigma\nu} - \partial_{\sigma}g_{\nu\alpha})$, and $T_{\mu\nu}$ is the stress-energy tensor that describes the source of the field in terms of matter/energy density, momentum density, and energy/momentum fluxes. Eq. (1.1) corresponds to a set of ten second-order coupled differential equations in the metric components $g_{\mu\nu}$, which in general are very difficult, if not impossible, to solve. The GW solution, however, arises in a particularly simple way from the linearization of Eq. (1.1) in vacuum, $T_{\mu\nu} = 0$, around the trivial solution of Minkowski spacetime. Let $h_{\mu\nu}$ be a small perturbation of the metric $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, and consider the tensor $\bar{h}_{\mu\nu}$ such that $h_{\mu\nu} = \bar{h}_{\mu\nu} - \frac{1}{2}\bar{h}\eta_{\mu\nu}$. The linearized Einstein Equations lead to a wave equation $\square\bar{h}_{\mu\nu} = 0$, which in general describes the propagation of a wave packet (superposition of monochromatic waves) that travels with the speed of light independently of the frequency (non-dispersive propagation). The tensor $h_{\mu\nu}$ originally has 10 degrees of freedom (DoF), being a symmetric 4×4 matrix; not all of these describe physical degrees of freedom, since a gauge choice is still possible. A customary and useful choice is the *Transverse Traceless* (TT) gauge, in which $h_{\mu 0} = 0$ (the perturbation is purely spatial), $\partial_{\alpha}\bar{h}^{\alpha\beta} = 0$ (transverseness) and $\bar{h}^{\alpha}_{\alpha} = 0$ (tracelessness: this has the effect that $\bar{h}_{\mu\nu} = h_{\mu\nu}$, thus the wave equation involves directly the perturbation $h_{\mu\nu}$). These three conditions make use, respectively, of 4, 3 and 1 DoF, thus leaving only 2 physical DoF, corresponding to the two possible independent polarization states of GW. The metric perturbation can thus

¹throughout this Thesis, if not otherwise specified I will use geometrized units $c = G = \hbar = 1$ and mostly-positive metric signature $(-, +, +, +)$.

be written as the sum of two amplitudes $h_\lambda(t)$, with index λ taking only 2 values, each multiplied by a polarization tensor $e_{\mu\nu}^\lambda$:

$$h_{\mu\nu}(t) = h_\lambda(t)e_{\mu\nu}^\lambda \quad e_{\lambda,\mu\nu}e_{\lambda'}^{\mu\nu} = 2\delta_{\lambda\lambda'}$$

A convenient choice of the polarization tensors is given by $\lambda = +, \times$ as

$$e_{ij}^+ = \mathbf{e}_x \otimes \mathbf{e}_x - \mathbf{e}_y \otimes \mathbf{e}_y = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad e_{ij}^\times = \mathbf{e}_x \otimes \mathbf{e}_y + \mathbf{e}_y \otimes \mathbf{e}_x = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

which describe the “plus” and “cross” polarizations of a quadrupolar transverse wave, and have the effect of deforming a circle of free test particles into an oscillating ellipse with the axes oriented along the reference frame’s axis direction (+) or with an inclination of 45° ($\times$). This is the working mechanism for the detection of GW with laser interferometers (Sect. 1.1.c).

Since GWs are usually small perturbations of the background spacetime, their production is often described by linearizing the Einstein Equations and expressing the stress-energy tensor (the source term) explicitly to describe the physical setup. The classical treatise to obtain the GW amplitudes far from the source consists in expanding the metric tensor $g_{\mu\nu}$ around the Minkowski metric, eventually obtaining (Maggiore, 2007)

$$h_{\mu\nu}(t, \vec{x}) = \frac{4G}{c^4} \int d^3y \frac{T_{\mu\nu}(t - |\vec{x} - \vec{y}|/c, \vec{y})}{|\vec{x} - \vec{y}|}$$

namely a *retarded potential*, since the information about the source at position $\vec{y}$ propagates with finite speed c and reaches the observer at position $\vec{x}$ at time $\tau = |\vec{x} - \vec{y}|/c$, with a magnitude that decreases like $1/|\vec{x} - \vec{y}| = 1/D_L$, where D_L is the luminosity distance from the source. This is a very important feature of GWs: detection involves measuring the gravitational *potential*, differently from electromagnetic waves, for which we can measure only the intensity of the *energy flux*, which scales as $1/D_L^2$ due to energy conservation. As a result, although GW are intrinsically extremely weak due to the numerically small prefactor G/c^4 , the decay with distance is less severe than the one experienced by the energy flux: this will enable GW detectors to reach detectability horizons that would not be possible for electromagnetic (EM) observations. For example, for suitable mass configurations, the GW detector LISA (see Sec. 1.3.c) will reach redshift $z \sim 20$ with signal-to-noise ratio (SNR) $\rho \sim 200$ (Bogdanovic et al., 2022), and much further, $z \sim 100$, at lower but still observable SNR, practically covering the whole observable Universe and probing the electromagnetically inaccessible era before reionization.

The standard post-Minkowskian expansion presented here applies to low-compactness systems (roughly speaking, the condition $(m_1 + m_2)/r \ll 1$ holds). This is not necessarily true for the systems considered in this Thesis, which can exhibit high compactness in the last phases of their evolution. In Chapter 3 the Teukolsky formalism is presented, in which the fundamental simplification is given by the expansion around a Kerr metric background instead (Sec. 1.4.a), exploiting the small mass ratio q .

1.1.b Gravitational Wave stress-energy tensor

Gravitational waves as described in the TT gauge do not have any effect on the coordinates of single test particles, namely every particle does not perceive an acceleration. This however does not imply that GWs cannot transfer energy to an extended system of particles, since the proper distance between the particles themselves change accordingly to the metric perturbation. For this reason, in principle if a generically dissipative mechanism is present (e.g. a dissipative spring between two test masses) energy and momentum can be extracted from a passing GW. This implies that the *perturbation of spacetime itself* is carrying energy and momentum, which are represented by a GW stress-energy tensor.

The most direct way of obtaining such a tensor is by using a field-theoretical approach (Maggiore, 2007), namely, to consider the metric perturbation $h_{\mu\nu}^{\text{TT}}$ as a classical tensorial field defined on a Minkowski (flat) spacetime. The definition of such tensor presents some conceptual difficulties: for example, it cannot be defined locally, since the spacetime is Minkowskian point-by-point (morally, the GW effect is manifest only on *finite-size* systems). For this reason, the GW stress-energy tensor must be defined as a finite-size spatial average of the standard stress-energy tensor found with the Noether currents associated with space-time symmetries. The equation of motion of the field is given by the Einstein-Hilbert action

$$S_{\text{EH}} = \int d^4x \frac{\sqrt{-g}}{16\pi} R \quad (1.2)$$

Where R is the Ricci scalar, $R = R_{\alpha}^{\alpha} = R_{\alpha\gamma\beta}^{\gamma} g^{\alpha\beta}$, and $g = \det g_{\mu\nu}$. At the first relevant order in h , which is the second since $R \sim h^2$, the action becomes

$$S_{\text{GW}} = -\frac{1}{64\pi} \int d^4x \left[\partial_{\mu} h_{\alpha\beta} \partial^{\mu} h^{\alpha\beta} - \partial_{\mu} h \partial^{\mu} h + 2\partial_{\mu} h^{\mu\nu} \partial_{\nu} h - 2\partial_{\mu} h^{\mu\nu} \partial_{\rho} h_{\nu}^{\rho} \right]$$

The Lagrangian is then $\mathcal{L} = -\frac{1}{64\pi} \left[\partial_{\mu} h_{\alpha\beta} \partial^{\mu} h^{\alpha\beta} - \partial_{\mu} h \partial^{\mu} h + 2\partial_{\mu} h^{\mu\nu} \partial_{\nu} h - 2\partial_{\mu} h^{\mu\nu} \partial_{\rho} h_{\nu}^{\rho} \right]$ and the stress-energy tensor (interpreted as a collection of four Noether currents) can be obtained as

$$t^{\mu\nu} = \left\langle \eta^{\mu\nu} \mathcal{L} - \frac{\partial \mathcal{L}}{\partial (\partial_{\mu} h_{\alpha\beta})} \partial^{\nu} h_{\alpha\beta} \right\rangle$$

where $\langle \star \rangle$ represents a spatial average over a region with typical size $L \gg \lambda$, where λ is the wavelength of the GW. Note that the application of the TT gauge must be performed *after* the computation of the derivative with respect to $\partial_{\mu} h_{\alpha\beta}$: the application of the gauge choice corresponds to choosing a univocal representation of h among a class of equivalence in a function space (e.g. L_2); the choice then corresponds to the reduction to a submanifold, which is possibly curve: while in the general description ∂ is computed as a simple partial derivative, if the gauge choice is explicitated before the application of the differential operator the correct expression would feature Christoffel coefficients. Instead, the derivative is performed in the general (flat) space of functions, and the gauge choice is specified at the end.

The final expression for the stress-energy tensor in the TT gauge and with explicit dependence from the constants c , G is eventually given by

$$t^{\mu\nu} = \frac{c^4}{32\pi G} \left\langle \partial^{\mu} h_{\alpha\beta} \partial^{\nu} h^{\alpha\beta} \right\rangle \quad (1.3)$$

This expression is crucial for the adiabatic treatment of gravitating binary systems, which is the fundamental theoretical tool used in Chapter 5 and 6 to evaluate the dynamical evolution of Extreme Mass Ratio Inspirals (EMRI) under the effect of standard GW energy and angular momentum dissipation and an additional scalar field dissipative channel (see Chapters 2, 3).

1.1.c Gravitational Waves experimental observations

Currently, the detection of deterministic GW signals from individual astrophysical sources is performed using interferometric techniques. Although other approaches have been explored in the past, such as resonant bars (Weber, 1968; Cerdonio et al., 1997; Astone et al., 1991) and planetary resonant modes, which are now being taken again in consideration for lunar GW detection (Ajith et al., 2025), and more are being considered, such as atom interferometry (Dimopoulos et al., 2009; Abe et al., 2021; Norcia et al., 2017), the first direct GW detection, followed by an extremely rich series of other events, was performed by the interferometers LIGO (Laser Interferometer Gravitational-Wave Observatory) and Virgo (Abbott et al., 2016; Aasi et al., 2015; Acernese et al., 2015). The scientific output of these detectors had a transformative impact on physics and astrophysics; the instruments are constantly subject to progressive improvement, leading to increasingly informative observations, larger detection volume and probed binary compact objects population (A. G. Abac et al., 2026a; N. Abac et al., 2026).

Although a detailed description of an interferometer is not necessary for the following, the physical principle that enables GW measurement is briefly summarized. As seen in the previous Section, the transit of a GW induces a quadrupolar deformation of proper distances between free test masses, with opposite expansion/contraction along the main polarization axes, which are positioned at 90° in the plane transverse to the GW propagation direction. Broadly speaking, the goal of an interferometer is to monitor the relative position of a set of test masses and detect the anti-correlated oscillations along the polarization's axes. This is achieved by comparing the phase of a laser linkages between the masses, in such a way that a small proper distance shift will induce a measurable dephasing between distinct linkages. In the case of ground-based interferometers, the test masses are heavy mirrors that constitute the termination of a resonant cavity, through which small length changes induce a much more pronounced dephasing with respect to a simple Michelson-Morley interferometric setup. Note, however, that the given broad definition also includes different setups used for space-borne detectors: for example, the Laser Interferometer Space Antenna (LISA) detector will use metallic test masses embedded in a constellation of three spacecrafts, the laser linkage and phase locking will be provided by a system of telescopes and transponders, and the test mass relative position monitoring will be mediated by individual measurements of test mass-spacecraft, spacecraft-spacecraft, and spacecraft-test mass distances. The dephasing measurement does not directly provide the proper distance, since the latter is an isochronous modification of spacetime, while the interferometric measurement is extended in time due to the finite speed of light. As a result, while longer interferometer arms are desirable, since the proper distance change is proportional to the baseline length, the time of flight increases, and higher frequency signals are averaged out, decreasing the instrument's response. For this reason, space interferometers like LISA can exploit an extremely long baseline, but the frequency range is forced to shift towards smaller frequencies, in the milliHertz band

for this example.

The only other technique that has already produced a direct observation of GW is the pulsar timing array (PTA), which monitors the correlations between the time of arrival (ToA) of radio pulses from an extended network of millisecond pulsars (Hobbs et al., 2010). The target of the measurement is the GW spectrum in the nanoHertz frequency band, which mainly consists of a stochastic background producing a characteristic averaged angular distance correlation between ToAs, namely the Hellings & Downs (HD) curve (Hellings et al., 1983). The International PTA (IPTA) revealed the presence of a stochastic GW background (Verbiest et al., 2016); the individuation of resolved sources or anisotropies, which did not yet occur, is expected to produce statistically relevant departures of the angular correlation function from the HD curve.

1.2 Detection and parameter estimation with gravitational waves

1.2.a Characterization of detector noise

The frequency window spanned by a detector, as well as its sensitivity in the frequency domain, is determined by its intrinsic noise. The output datastream of a detector can be decomposed as a signal + noise,

$$d(t) = s(t) + n(t) \quad (1.4)$$

where the magnitude of the noise can potentially be dominant with respect to the signal that we want to detect and characterize. For this reason, a rigorous treatment of noise properties is essential to correctly recover the very faint signal. Note that in Eq. (1.4) the noise is *referred to the input*: we consider the noise as it was produced by the same physical mechanism that we are probing, in this case GW, to be able to compare different sources of noise directly with the physical GW signal; morally, even if the noise is produced later in the detection apparatus, we always treat $n(t)$ as the intensity that a GW signal should have had to produce the particular realization of the total datastream $d(t)$. The GW *strain* $s(t)$ corresponds to the projection of the GW tensor h_{ij} through an *antenna pattern* tensor F^{ij} , which depends on the detector configuration and the orientation of the source, parametrized by the right ascension and declination α, δ ; in terms of the $+$ and $\times$ polarization amplitudes, $s(t) = F_+(\alpha, \delta; t)h_+(t) + F_\times(\alpha, \delta; t)h_\times(t)$; it is customary to express the GW amplitude as the *complex strain* $\mathbb{h} = h_+ - ih_\times$, such that the antenna pattern expressed as a complex function $\mathbb{F} = F_+ + iF_\times$ gives the strain $h(t)$ simply through $h(t) = \mathbb{h} \cdot \mathbb{F}$. Note that in general the antenna pattern can be a function of time, especially for long-duration signals where the detector's motion becomes relevant.

Noise can be defined as *any source of stochastic unwanted signal in the detector datastream*. It can be produced by the detector itself, the environment, or even be of the same physical origin of the “true” signal. For example, typical sources of noise are seismic vibrations, which in current interferometers are damped by a cascade of passive and active seismic insulators (standard and inverted pendula, feedback mechanisms) to which the reflective masses are attached, the readout noise caused by the physical mechanism that converts the interferometer's laser intensity to an electric signal (transducer), the thermal noise that is typical of finite-temperature objects as the mirrors (which for this reason will be maintained in a cryogenic environment for the Einstein Telescope), the quantum noise

caused by the momentum transfer of single photons impinging on the mirrors (shot noise); the Newtonian noise consists in local gradients of the local gravitational field, and it is much more difficult to avoid since it consists of the very same nature of the signal that we want to isolate; even worse, the Laser Interferometer Space Antenna (see Sec. 1.3.c) will be limited in the mHz band by a stochastic background of unresolvable galactic double white dwarfs: in this case, not only this contribution has the same nature of the signal it covers (for example, a very distant EMRI or massive symmetric binary), but also comes from an astrophysical sources, and can be considered as a signal or as a noise depending on the topic of interest. This brief list showcases the variety and great difference in the possible noise sources, and motivates the development of a unified formalism to take into account all these factors through a set of common theoretical tools.

Noise is a stochastic process, thus it can be characterized only through its statistical properties. In order to develop operative data analysis methods, three fundamental assumptions are made (Maggiore, 2007):

- **Ergodicity:** the statistical properties of a set of multiple noise realizations at fixed time t coincide with the statistical properties of a single noise realization at different times, namely ensemble properties can be obtained from time-averaged properties.
- **Stationarity:** the statistical properties of the noise are independent of the time of measurement. This hypothesis must hold for at least the duration of the signal.
- **Gaussianity:** at fixed time, the probability distribution for the noise realization is gaussian; this is a consequence of the Central Limit Theorem, thanks to the fact that the noise results from a large set of stochastic processes (which are not necessarily gaussian).

These hypotheses constitute the foundation of the current data analysis techniques; while I will always assume their validity, which is strongly motivated and validated in current interferometers, it is important to notice that they will be violated in the context of next-generation detectors. The stationarity condition is ensured for current observations thanks to their brevity, $\Delta t \lesssim 128$ s (A. G. Abac et al., 2026a), however this will not be the case for next-generation interferometers, where the signal can last $\mathcal{O}(10^3)$ s for solar-mass binary black holes (BH) in the audio band (in the case of Einstein Telescope and Cosmic Explorer), up to years in the case of mid- to low-frequencies as for the Lunar Gravitational Wave Antenna (LGWA, Ajith et al., 2025) and the Laser Interferometer Space Antenna (LISA, Amaro-Seoane et al., 2017), for which the signal can be present in the observational band for the whole duration of the mission. In these cases, variations of the statistical characteristics of noise are expected during the evolution of the signal (Edwards et al., 2020). Regarding Gaussianity, the prevalence of a single non-Gaussian noise source could prevent the application of the central limit theorem; this will be the case for LISA, when considering the stochastic noise produced by the galactic double-white dwarfs background (Buscicchio et al., 2025).

The main statistical property of the detector's noise is the autocorrelation function, which is a measurement of the expected correlation of the noise value at two different times:

$$\vartheta^{(2)}(t_1, t_2) \equiv \langle n(t_1)n^*(t_2) \rangle$$

It is more convenient to work with another quantity, namely the power spectral density (PSD) of the signal. The PSD is defined as the average “power” contained in the signal at

a fixed frequency, namely $S_n \equiv \langle |n(f)|^2 \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |n_t(f)|^2 dt$. The PSD and the 2-point correlation function are related by the Wiener-Kinchin theorem under the assumption of wide-sense stationarity, a weaker requirement of the general stationarity. Under this hypothesis, the 2-point CF corresponds to the inverse Fourier transform of the PSD:

$$\vartheta^{(2)}(t) = \int df e^{ift} S_n(f)$$

This easily defines a handy expression for the power spectrum:

$$\langle n(f_1) n^*(f_2) \rangle = \frac{1}{2} S_n(f_1) \delta(f_1 - f_2)$$

This is easily shown as follows, rewriting the 2-point correlation function in Fourier domain and verifying that the use of the previous expression (passage marked as $\star$) leads to the statement of the Wiener-Kinchin theorem:

$$\begin{aligned} \vartheta^{(2)}(\tau) &= \langle n(t + \tau) n^*(t) \rangle \\ &= \left\langle \int df_1 \int df_2 e^{if_1(t+\tau)} e^{-if_2 t} n(f_1) n^*(f_2) \right\rangle \\ &= \int df_1 \int df_2 e^{if_1(t+\tau)} e^{-if_2 t} \langle n(f_1) n^*(f_2) \rangle \\ &\stackrel{\star}{=} \int df e^{ift} \frac{1}{2} S_n(f) \\ &= \int_0^\infty df e^{ift} S_n(f) \end{aligned}$$

Note that the PSD is dimensionful and has the dimensions of the inverse of the measure df in order to give an adimensional correlation function; thus $[S_n] = \text{Hz}^{-1}$. Other representations of the detector noise are the *spectral amplitude* $\sqrt{S_n(f)}$, which is motivated by the fact that S_n is quadratic in the noise amplitude, thus the spectral amplitude becomes linear, and the *characteristic strain* $h_n = \sqrt{f S_n(f)}$, which is adimensional and allows for an intuitive graphical evaluation of the detectability of a signal (Moore et al., 2015). In Fig. 1.1 the characteristic strain noise curves of some present and future detectors are shown, together with the typical strain amplitude of relevant GW sources.

1.2.b A whisper in the crowd: matched filtering and parameter estimation

The detection of GW is by definition a high-precision measurement process, in the sense that the amplitude of the noise is larger than the amplitude of the signal present in the datastream, namely $\|n(t)\|^2 \gg \|s(t)\|^2$; to be able to extract relevant information, it is therefore necessary to develop efficient and robust data analysis techniques, the most important of which is *match filtering*, which addresses transient and modeled signals. In this Section I will introduce the match filtering, the signal-to-noise ratio (SNR) and the Wiener dyadic product, general tools that will be useful in Chapter 5 to assess detectability and parameter estimation (Maggiore, 2007).

The **Wiener dyadic product** between two signals $a(f)$ and $b(f)$ in the frequency domain

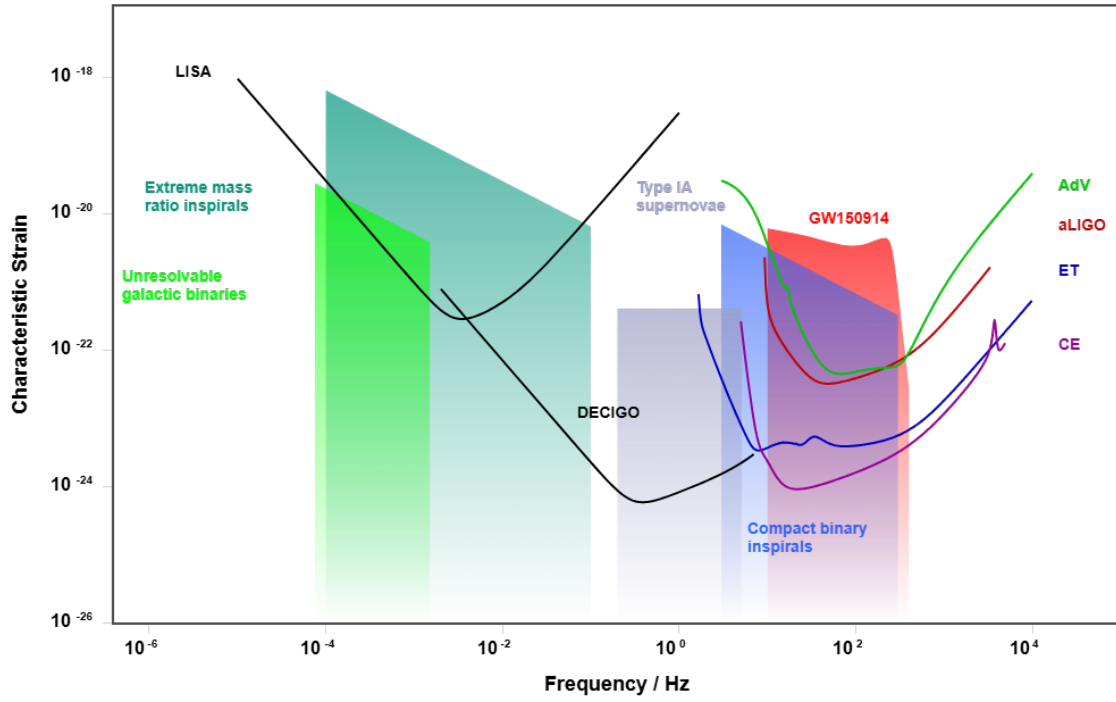


Figure 1.1: The present and future landscape of GW detection, with some of the most relevant detectors and sources. Refer to Sec. 1.3 for a description of the future detectors. In the frequency vs characteristic strain space, from left to right: LISA will observe sources in the mHz band, such as Extreme Mass Ratio Inspirals; the unresolved galactic binaries background in green showcases the possibility of a GW source effectively acting as a noise feature. In the deci-Hertz band is represented DECIGO, and the representative source of supernovae type Ia resulting from double-white dwarfs merging. In the audio band, the current configurations of Advanced LIGO and Advanced Virgo (red and green respectively), and the future ground-based interferometers Einstein Telescope and Cosmic Explorer, mainly probing compact binary coalescences of solar-mass compact objects. Plot realized with the GWPLOTTER tool (Moore et al., 2015).

is defined as

$$(a|b) \equiv \Re \int_{-\infty}^{+\infty} df \frac{a^*(f)b(f)}{S_n(f)/2} = 4\Re \int_0^{+\infty} df \frac{a^*(f)b(f)}{S_n(f)} \quad (1.5)$$

where the second equation is motivated by the fact that since $a(t), b(t)$ are real signals in time, $a^*(f) = a(-f), b^*(f) = b(-f)$ and the dyadic product involves only the real part of the integral. The dyadic product is real, thus symmetry $(a|b) = (b|a)^* = (b|a)$ is implied, depends on the specific properties of the detector's noise and has the property of acting like a standard L_2 product between functions of one variable (frequency) with the property of weighting more the product in correspondence of frequencies where the detector is more sensitive. In particular, following the definition of the power spectrum, the product is directly comparing *the product of the two signals* with *the average square amplitude of the noise*. Note that $(a|b)$ is a well-defined scalar product in a space of functions: this implies the existence of a corresponding metric. This geometric interpretation will be useful when introducing Fisher matrices (Sec. 1.2.c).

A generic **linear filter** consists in the convolution of a signal $h(t)$ with a *kernel function* $K(t)$, to obtain a filtered time series $S(t)$:

$$S(t) = \int_{-\infty}^{+\infty} d\tau h(\tau) K(\tau + t) \quad (1.6)$$

let's consider a noisy signal $h(t) = s(t) + n(t)$, with $s(t) = \alpha \tilde{s}(t)$, $\|s\|_{L_2}^2 = 1$ such that α represents an overall amplitude of the signal present in the datastream. Then, if the kernel is exactly $\tilde{s}(t)$, the expectancy value of the filtered signal at $t = 0$ will read

$$\langle S \rangle = \left\langle \int_{-\infty}^{+\infty} d\tau h(\tau) \tilde{s}(\tau) \right\rangle = \int_{-\infty}^{+\infty} d\tau \alpha \tilde{s}^*(\tau) \tilde{s}(\tau) = \alpha \quad (1.7)$$

since $\langle n \rangle = 0$. This is the fundamental concept of match filtering: *if the functional form of a signal is known, it is possible to extract its amplitude from a noisy datastream*. This explains why for GW searches it is of primary importance to *know in advance* the functional form of the target signal up to a certain accuracy, both for detection and parameter estimation. The biased knowledge of a signal will result in a non-detection, or in a systematically biased parameter inference (Benetti et al., 2026a; Dhani et al., 2025).

Of course, even if $\langle n \rangle = 0$ for an infinite amount of time, the specific realization of the noise for a finite-duration signal (or observational period, or even data analysis window) will introduce a finite statistical bias on the signal amplitude $S(t)$. For this reason, we can estimate the square amplitude of the noise through the same kernel:

$$\begin{aligned} N^2 &= \langle n^2 \rangle - \langle n \rangle^2 = \langle n^2 \rangle = \int_{-\infty}^{+\infty} dt \int_{-\infty}^{+\infty} dt' K(t) K(t') \langle n(t) n(t') \rangle \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} df |K(f)|^2 S_n(f) \end{aligned}$$

using a double Fourier transform and the definition of the spectral noise. Thus, for a generic kernel K and a signal $s(t)$ present in the datastream it is possible to define a **signal**

to noise ratio (SNR)

$$\frac{S}{N} = \frac{\int_{-\infty}^{+\infty} df K^*(f) s(f)}{\sqrt{\int_{-\infty}^{+\infty} df |K(f)|^2 S_n(f)/2}}$$

The **optimal matched filter** corresponds to the choice of the kernel K that maximizes the SNR once the signal and noise are fixed. It can be shown that the optimal kernel is given by $K(f) = s(f)/S_n(f)$. In this case, the SNR is called *the optimal SNR* and takes the value $\rho = \sqrt{(s|s)}$, where $(\star|\star)$ is the previously defined Wiener dyadic product.

The operative implementation of matched filtering consists in a continuous evaluation of the SNR statistics over a extensive bank of waveform templates, such that every signal is well approximated by at least one template. An event is further considered as a possible detection if the SNR exceeds a given threshold (currently typically 8, for LISA the detection threshold will be $\rho = 20$), such that the false alarm rate (FAR) caused by the statistical behavior of the noise (which will display a stochastic behavior with occasional peaks of relatively high SNR) is sufficiently low to avoid an unmanageable excess of false positives.

1.2.c Parameter estimation forecasts: the Fisher matrix technique

Once a signal is detected through matched filtering, the goal of a parameter estimation procedure is to *assign a joint probability density to specific values of a parameter set describing the system*. To do so, we consider the *likelihood function* $\mathcal{L}$ associated with the noise statistics: since the noise is gaussian and stationary, it can be written as a product of single gaussian likelihoods for every noise sample:

$$\mathcal{L}_n(d; S_n) \propto \exp \left\{ -\frac{1}{2} \sum_i \frac{|n_i|^2}{S_n(f_i)T} \right\}$$

Note that the expression is already written in frequency domain, where $S_n(f)$ acts like a frequency-dependent variance, and includes the dependence from the observational time T : in particular, the noise floor decreases as $T^{-1/2}$ as expected.

If a deterministic signal $h(f; \vec{\theta})$ which is a function of the frequency and a finite set of parameters $\vec{\theta}$ is present in the datastream $d(f)$, the likelihood of the parameter set $\vec{\theta}$ will be inherited from the noise likelihood via

$$\mathcal{L}(d; S_n, \vec{\theta}) \propto \exp \left\{ -\frac{1}{2} \sum_i \frac{|d_i - h(f_i; \vec{\theta})|^2}{S_n(f_i)T} \right\}$$

in other words, we are evaluating *how probable it is that the residuals are exclusively described by noise*. The maximum-likelihood parameters will indicate the optimal set of parameters such that the datastream minus the template generated with that parameter set “looks like” a realization of gaussian stationary noise. Moving away from the ideal parameter set will result in a increasingly relevant residual, which will lower the likelihood since a true noise realization with such residuals is progressively less probable.

The Bayes theorem allows to compute the *posterior distribution*, namely the proper probability density of the parameter set, as the normalized product between the likelihood and

the *prior distribution*, which is a probability density describing the probability of an event *a priori* of the detection. For example, the prior distribution as a function of the luminosity distance D_L will scale approximately as D_L^2 , reflecting the fact that a spherical shell located at radius D_L and with fixed thickness has a volume that scales as D_L^2 ; but priors are used also to exclude unphysical combinations of parameters (for example, a negative mass of a binary system's component). In general, the priors are chosen to be the *less* informative possible, since a biased prior can severely bias the posterior; for example, a prior on the sky location will be completely homogeneous on the S^2 sphere. At the same time, if strongly physically-motivated prior distributions are available, they can help constraining the posterior also starting from a loosely-constrained likelihood.

The **Fisher matrix technique** is a gaussian approximation of the likelihood function in the parameter space. In particular, we write

$$\mathcal{L} \approx \exp \left\{ -\frac{1}{2} \Delta \theta^i \mathcal{F}_{ij} \Delta \theta^j \right\} \quad (1.8)$$

To find the explicit expression of the Fisher matrix $\mathcal{F}_{ij}$, let's expand the logarithm of the likelihood around its maximum, defined by the parameter set $\vec{\theta} = \vec{\theta}_0$, where $\partial_i \ln \mathcal{L} = 0$:

$$\begin{aligned} \ln \mathcal{L}(d; S_n, \vec{\theta}) &= \ln \mathcal{L}|_{\theta=\theta_0} + \frac{1}{2} \Delta \theta^i \Delta \theta^j \partial_i \partial_j \left[-\frac{1}{2} (d - h(\vec{\theta}) | d - h(\vec{\theta})) \right] \Big|_{\theta=\theta_0} + \mathcal{O}(\Delta \theta^3) \\ &\approx \ln \mathcal{L}|_{\theta=\theta_0} + \frac{1}{2} \Delta \theta^i \Delta \theta^j \left[-\frac{1}{2} \left[2(\partial_i \partial_j h(\vec{\theta}) | h(\vec{\theta}) - d) + 2(\partial_i h(\vec{\theta}) | \partial_j h(\vec{\theta})) \right] \right] \Big|_{\theta=\theta_0} \end{aligned}$$

Where in the second line the approximation is due to dropping the residual $\mathcal{O}(\Delta \theta^3)$ terms. Finally, consider the expectation value of the noise $\langle n \rangle = 0$ (ensemble average over multiple realizations), such that $\langle d - h(\vec{\theta}) \rangle = 0$:

$$\ln \mathcal{L}(d; S_n, \vec{\theta}) = \ln \mathcal{L}|_{\theta=\theta_0} + \Delta \theta^i \Delta \theta^j \left[-\frac{1}{2} (\partial_i h(\vec{\theta}) | \partial_j h(\vec{\theta})) \right] \Big|_{\theta=\theta_0} + \mathcal{O}(\Delta \theta^3)$$

Thus, comparing this expanded expression with Eq. (1.8) we found that the Fisher matrix can be computed as

$$\mathcal{F}_{ij} = (\partial_i h | \partial_j h) \quad (1.9)$$

From the gaussian approximation, Eq. (1.8), it can be seen that the Fisher matrix is the inverse of the covariance matrix $\mathcal{C}_{ij}$: a generic Gaussian distribution is indeed defined by $\mathcal{G}(\delta \theta) \propto \exp \left\{ -\frac{1}{2} \delta \theta^i \mathcal{C}_{ij}^{-1} \delta \theta^j \right\}$. Thus the inverse of $\mathcal{F}_{ij}$ directly provides the information on the joint and marginalized probability densities in the parameter space. Eq. (1.9) is a simple expression that allows fast and reliable forecasts for parameter estimation with future detectors, in the limit of high SNR (where the application of the expectancy value over an ensemble of noise realizations is actually valid). The ingredients are the noise power spectrum of the detector, which is necessary to compute the dyadic product, and a *template* of the waveform, $h(\vec{\theta})$, that can be generated with different parameter values in order to compute the partial derivatives. This approximates well the likelihood (and usually also the posterior distribution, since at high SNR the generic priors are often nearly

constant; however that it not always the case, see Sec. 5.5) under the assumptions that the noise spectrum and the signal template correctly represent the respective physical phenomena.

Geometrical interpretation of the Fisher matrix

The choice to parametrize every possible waveform² with a finite set of parameters, corresponding to a D -dimensional vector $\vec{\theta}$, corresponds to a dimensionality reduction from a ∞ -dimensional space of functions for the signal h to a D -dimensional submanifold $\mathfrak{M}$. Since the Wiener scalar product $(h_1|h_2)$ induces a metric $d^2w = (dh|dh)$, the submanifold metric is inherited through the projection $d^2\tilde{w} = (\partial_i h|\partial_j h)d\theta^i d\theta^j = \mathcal{F}_{ij}d\theta^i d\theta^j$. Thus, the submanifold $\mathfrak{M}$ is locally homeomorphic to $\mathbb{R}^D$ with a metric given by the Fisher matrix³. With this interpretation, the distance between two waveforms (defined as vectors in the manifold $\mathfrak{M}$) provides a direct *distinguishability criterion* through the standard Gaussian compatibility, which will be used to obtain an estimation for the observable dephasing threshold in Sec. 5.4.b.

Limitations of Fisher matrix technique

In Chapter 5 the Fisher matrix technique is applied to assess the detectability and measurability of EMRI parameters, including those describing an additional massive scalar field that mediates gravity (Chapter 2). It is important to note that, while the method offers a cheap and fast tool for evaluating measurement errors, it is limited by a series of factors. These limitations usually arise in the low-SNR limit and could become important for the EMRI population in LISA, and consist mainly of (Marsat et al., 2021; Piro et al., 2023):

- **Non-gaussian posteriors:** the posterior distribution could present non-gaussianities in the general shape; typically, posterior probability contours can assume curved shapes in the parameter space, especially when the errors are comparable with the absolute value of the considered parameter. For example, wide sky locations involving right ascension and declination α and δ often display complex shapes; more importantly, consider luminosity distance and inclination, which are naturally degenerate: the degeneracy provides the elongation of the posterior in the parameter space, and non-cartesian nature of the parameter space contributes to non-gaussian deformations of such elongated posterior.
- **Multi-modality:** especially for the LISA detector, which will probe the phenomena described in this Thesis, posteriors with multiple peaks are possible; this usually arises due to approximate discrete symmetries in the detector, causing different sky positions to be nearly equivalent. In particular, the true sky location is often mirrored in every octant of the sky (in the LISA reference frame), obtaining not one but eight local maxima of the likelihood function. If strong covariance with other parameters is present, multi-modality can propagate to such parameters.
- **Single-event fit:** while the previous effects are at least partly mitigated by the high SNR considered, a big uncertainty comes from the fact that data analysis for LISA

²within a certain class of phenomena: for example, a binary black hole (BBH) coalescence is usually defined by a set of 15 parameters, 8 intrinsic (component masses m_1 , m_2 and spins $\vec{\chi}_1$, $\vec{\chi}_2$) and 7 extrinsic (inclination ι , phase φ , sky location α , δ , polarization angle ψ , luminosity distance D_L and time of arrival t_c).

³Note that $\mathfrak{M}$ is not globally homeomorphic to $\mathbb{R}^D$ since some variables are defined periodically on a limited domain, for example the angular coordinates.

will be completely different from the currently used techniques. Stationarity and gaussianity of the noise will not be satisfied, multiple signals from different sources will overlap simultaneously: for this reason, it is likely that a *global fit* will be performed, estimating the parameters of all the present sources at once. There are multiple proposals on how to perform a global fit (Toubiana et al., 2026; Criswell et al., 2026; Srinivasan et al., 2025; De Santi et al., 2026), each with a focus on different aspects. However, a common aspect is that the global fit could introduce covariance between parameters of *distinct* sources: the covariance would be incidental and not physical, namely purely due to the particular way in which different sources are overlapping. In this sense, the Fisher matrix technique could underestimate the statistical uncertainties on the parameters' posteriors, since it would represent a *conditioned* probability over the rest of the other sources' parameters, and not a *marginalized* distribution.

Clarified the possible limitations of this method, it is however evident the immediate practical value to approximately assess at first order the measurability of the source parameters. Errors obtained in this way should thus be regarded as an indication of the error magnitude, rather than a necessarily accurate forecast. The other alternative is the full Bayesian analysis of an injected signal, which, however, is computationally prohibitively expensive for signals with a mission lifetime duration, such as the ones considered in this work.

1.3 Towards the next generation of detectors

In this section I will give an overview of the next-generation GW detectors that will be relevant for assessing the occurrence of the phenomena predicted by beyond-GR theories, and in particular the existence of a ultralight scalar mediator of gravity, which is the topic of this Thesis. Although many different detectors are being proposed, the fundamental distinction is the frequency band; I will then follow a frequency order, starting from the audio band (≈ 10 Hz to $\mathcal{O}(1000)$ Hz), passing through the deciHertz (dHz) window down to the milliHertz (mHz) band, which will be of primary interest in the next chapters and is described more thoroughly.

1.3.a The audio band: Einstein Telescope and Cosmic Explorer

The audio band, spanning from few Hz to the kHz, has been proven to be extremely rich of information and compact binary coalescence (CBC) events (N. Abac et al., 2026; A. G. Abac et al., 2026a). In light of this high informativeness, next-generation ground-based interferometers are currently being designed.

Einstein Telescope (ET) is a European project for an underground, cryogenic laser interferometer (Punturo et al., 2010; A. Abac et al., 2026), consisting in either a triangular 10-km arms configuration or two separate L-shaped, 15-km arms configurations (Branchesi et al., 2023). Cosmic Explorer (CE) is the American successor of the LIGO interferometers, consisting in two L-shaped surface interferometers with 20- or 40-km arms (Dwyer et al., 2015; Reitze et al., 2019). Despite differences in the configurations and instrumental choices, the sensitivity curves are comparable, with ET reaching lower frequency thanks to the reduced seismic noise in the underground environment, and CE being slightly more sensitive at higher frequencies due to larger arm lengths. For this reason, the science cases

of these detectors are the same, and they are often treated as a unique next-generation GW detection network.

The ET-CE science cases include the observation of BBH, with a merger rate of $\sim 10^4 \text{ yr}^{-1}$ within redshift $z = 1$, and 10 times more in the whole detection volume, with $\sim 10^4$ events per year having $\text{SNR } \rho > 100$ (Borhanian et al., 2024). This will enable extremely precise characterization of BBH populations up to high redshifts, strong-gravity dynamics and tests of GR, precision cosmology via the dark siren method (Maggiore et al., 2020; A. Abac et al., 2026). Binary neutron stars (NS) and BH-NS systems will also manifest with high SNR, with a rate of $\sim 10^5 \text{ yr}^{-1}$ within $z = 1$ at $\text{SNR } \rho > 10$ (Borhanian et al., 2024), enabling an accurate reconstruction of the equation of state in the NS core and probing dense-matter phenomena that would be otherwise impossible to observe. ET and CE will also open the possibility of observing a stochastic GW background in the audio band, either of cosmological or astrophysical origin, monochromatic signals from fast-spinning neutron stars, and possibly other unmodeled phenomena such as supernovae explosions. The science cases will be sometimes observable in lower GW frequency bands, enabling a multi-band approach between different detectors such as LISA and LGWA (following Sections), or by electromagnetic facilities as for the multimessenger event GW170817 (Abbott et al., 2017), or possibly by neutrino detection; the multimessenger observation of an event would provide different sets of information, enabling a more complete description and comprehension of such phenomena.

1.3.b The deciHertz window: LGWA and DECIGO

Towards lower frequencies, the dHz frequency band offers unique case studies of primary physical importance. Among these, particularly important will be massive and intermediate-mass BBH (higher mass shifts the waveform towards lower frequencies, enabling to study a different subpopulation with respect to the one currently observed by LIGO/Virgo), enabling population studies, tests of GR and multibanding with audio-frequency detectors, binary NS mergers, cosmology and cosmography (the localization capabilities of dHz detectors allow to precisely locate the GW source, linking it to the galaxy host without a prompt counterpart; this can be used to calibrate the Hubble law) (Berti et al., 2026). Many other sources are specific to this narrow band: double-white-dwarfs mergers and the associated counterparts (possibly supernovae type Ia) enabling multimessenger astronomy (Benetti et al., 2026b), tidal disruption events (TDE), in which a compact secondary consisting in a neutron star or a white dwarf is tidally disrupted when approaching a primary Super Massive Black Hole (SMBH), possibly providing an observable EM transient (Sesana et al., 2008), and most importantly for the content of this thesis, Intermediate Mass Ratio Inspirals (IMRI), which are asymmetric binary systems with mass ratio $q \equiv m_{\text{sec}}/M \sim 10^{-3}$. Like EMRIs (see Sec. 1.4), they are of particular interest since the secondary can spend hundreds of orbits in the observable band, enabling a precise inference of intrinsic parameters and the potential presence of smaller effects, such as environmental effects or modifications to GR.

The proposals for the dHz band detectors, unlike for the other frequencies, are more variegated. The Lunar Gravitational Wave Antenna (LGWA, Ajith et al., 2025) would exploit the interaction between GWs and the Moon as a resonant body to measure the gravitational signal, following the concept of resonant bars detectors proposed by Weber. In particular, the tidal forces associated with a passing GW would excite quadrupolar nor-

mal modes of the Moon, which can be measured by a network of four high-precision seismometers (Heijningen et al., 2023) as a low-frequency seismic signal.

The Deci-hertz Interferometer Gravitational Wave Observatory (DECIGO) is a proposed space-based detector operating between the dHz and the low audio band. The configuration is analogous to LISA (Sec. 1.3.c), and consisting in possibly multiple independent clusters of 3 satellites linked by a 1000 km laser linkage.

Other proposed detectors are the Lunar Interferometer Laser Antenna (Jani et al., 2025), consisting in a standard interferometric setup brought to the Moon to exploit the lack of seismic noise, or atom interferometry, both on Earth and in space (Abe et al., 2021; Norcia et al., 2017).

1.3.c The milliHertz band: Laser Interferometer Space Antenna

The Laser Interferometer Space Antenna (LISA) is a space-based future interferometer operating from $\sim 10^{-4}$ to 1 Hz (Amaro-Seoane et al., 2017). After the success of the technology demonstrator LISA Pathfinder (Armano et al., 2016) the LISA mission is currently under development, with scheduled launch in 2035. The mission duration will be nominally 4 years, with a probable extension to 10 years, given the transformative impact of such detector. LISA will consist in a constellation of three spacecrafts carrying inertial test masses; the relative position of the test masses is monitored through a laser linkage between the three spacecrafts, each located at ~ 2.5 million kilometers from each other. The characteristic noise curve can be found in Fig. 1.1, and results from the combined effect of *stray acceleration* of the test masses at low frequency (imperfect cancellation of the residual acceleration due to solar wind, radiation pressure, electrostatic charges on the test masses), *displacement noise* between pairs of test masses, and increased *strain noise* at high frequencies (above 30 mHz) due to partial cancellation of the GW signal after the violation of the long-wave approximation (for such frequencies, $\lambda_{\text{GW}} < L_{\text{arm}}$). These few sources of noise define an optimal frequency range around few mHz. In addition, it is expected the presence of a relevant unresolved background generated by galactic binary white dwarfs; despite the genuine GW nature of such background, it is customary to consider it as a noise source, since it effectively prevents the detection of fainter signals. The galactic white dwarfs (WD) background is mostly present between 0.5 and 2 mHz, causing a small bump in the sensitivity curve. An important feature of LISA is a time-changing antenna pattern, which will help localizing long-lived sources by providing a known modulation to the signal depending on the sky location. The sampling rate of 3.33 Hz determines the maximum reachable frequency of ≈ 1 Hz through the Nyquist sampling theorem; for the sources treated in this Thesis (namely Extreme Mass Ratio Inspirals) the maximum frequency before plunging is lower, making sufficient even smaller sampling rates.

The most important astrophysical science cases for LISA (Barausse et al., 2020) will be the study of supermassive BH binary mergers, with total mass $M \sim 10^5 - 10^7 M_{\odot}$, extreme- and intermediate- mass ratio inspirals, which are treated in the next Section, early stages of stellar-mass BBH enabling multiband observations with the dHz and audio band detectors, galactic WD and NS binaries; cosmological breakthroughs could come from the detection of a cosmological stochastic GW background, and the accurate sky location reconstruction of high-redshift sources will improve cosmography and the calibration of the Hubble law. Regarding fundamental physics, LISA will enable precision

tests of GW propagation, generic tests of GR and strong-gravity regimes, which are one of the topics of this Thesis: in particular, massive fields potentially acting as dark matter candidates would become observable in EMRI dynamics, affecting the evolution through additional dissipative channels (Chapter 5) or very distinctive features such as superradiance and floating orbits (Chapter 6).

Other similar concepts are being proposed, as the detectors TianQuin (Luo et al., 2016) and Taiji (Hu et al., 2017), with respectively arm lengths of 3×10^5 and 10^6 km, which would locate the optimal frequency window at slightly higher frequency with respect to LISA. In Chapter 5, the detection forecasts for EMRIs with modified gravity are exclusively referred to the LISA detector, for a signal duration of 2 years; mission extensions, noise curve upgrades and orbital configuration should be taken into account for future estimations of GW parameter inference.

1.4 Extreme Mass Ratio Inspirals

Extreme Mass Ratio Inspirals (EMRI) are a class of GW sources consisting of a binary of compact objects with mass ratio $q = m_{\text{sec}}/M \ll 1$, usually of $\approx \mathcal{O}(10^{-6})$. The primary must be a super-massive BH with $M \approx \mathcal{O}(10^6 - 10^8 M_\odot)$, while the secondary is usually a solar-mass BH with $m_{\text{sec}} \approx \mathcal{O}(10 M_\odot)$ or a neutron star, $m_{\text{NS}} \approx 3 M_\odot$, or even a white dwarf ($m_{\text{WD}} \approx 1.4 M_\odot$). Given the extreme mass difference, the dynamics of the system can be solved with a perturbative approach using the mass ratio q as an expansion parameter.

The approximate frequency at which EMRIs emit can be found by considering a typical orbit, say a circular orbit at $r \sim 10M \left(\cdot \frac{G}{c^2} \right)$. Approximating the orbital dynamics with

Newtonian gravity, we have that for a $10^6 M_\odot$ primary, $f = \frac{1}{2\pi} \sqrt{\frac{GM}{r^3}} \sim \frac{1}{2\pi} \sqrt{\frac{GMc^6}{(10MG)^3}} \sim$

$\frac{1}{2\pi MG} \sqrt{\frac{c^6}{1000}} = 10^{-3}$ Hz. Thus we see that EMRIs will be primary targets for the LISA detector, which operates from $\sim 10^{-4}$ Hz to ~ 0.1 Hz. Other objects of interest are the Intermediate Mass Ratio Inspirals (IMRI), which exhibit $q \sim 10^{-2}$ and emit in the dHz band. These can be treated with the same formalism (although reaching a lower accuracy with a truncated perturbative approach) and will be relevant for LGWA. The length of the inspiral can be computed using the formula for the time to coalescence in a quasi-adiabatic inspiral (Maggiore, 2007),

$$\tau = \frac{5}{256} (\pi f_{\text{start}})^{-\frac{8}{3}} \left(\frac{c^3}{G\mathcal{M}} \right)^{\frac{5}{3}} \quad (1.10)$$

where $\mathcal{M} = M \left(\frac{q^3}{1+q} \right)^{\frac{1}{5}} \approx Mq^{\frac{3}{5}}$ is the chirp mass expressed as a function of M and q and approximated for small q in the second passage. For an EMRI defined by a starting frequency of 1 mHz, $M = 10^6 M_\odot$ and $m_{\text{sec}} = 10 M_\odot$, we obtain $\tau \sim 10^3$ years: for many systems, the permanence in the observable frequency is set by the duration of the mission, and not by the duration of the inspiral. This will result in an extremely long integration of the signal, enabling high-precision parameter estimation and a large detection volume.

In this Section I will present the general phenomenology associated with a EMRI constituted of a primary Kerr BH and a secondary point-like particle; as detailed in the following, this case is not only of physical relevance for the standard general relativity treatise,

but also for a wide class of theories of modified gravity.

1.4.a The Kerr metric and geodesic motion

The Kerr metric (Kerr, 1963; Saul A. Teukolsky, 2015) represents a vacuum solution of the Einstein field equations around a massive rotating object, and corresponds to the metric around every supermassive black hole. I will use a Boyer-Lindquist coordinate system, which enables to represent the space with variables t, r, θ, ϕ that resemble the spherical coordinates of Minkowski spacetime. The relation between B-L coordinates and usual Cartesian coordinates is given by

$$\begin{aligned}x &= \sqrt{r^2 + a^2} \sin \theta \cos \phi \\y &= \sqrt{r^2 + a^2} \sin \theta \sin \phi \\z &= r \cos \theta\end{aligned}$$

where a is a rotation parameter that will be characterized later; for now it simply represents an oblation of the spherical coordinates.

The components of the metric are the following:

$$\begin{aligned}g_{tt} &= -\left(1 - \frac{2Mr}{\Sigma}\right) & g_{rr} &= \frac{\Sigma}{\Delta} & g_{\theta\theta} &= \Sigma \\g_{\phi\phi} &= \left(r^2 + a^2 + \frac{2Mra^2 \sin^2 \theta}{\Sigma}\right) \sin^2 \theta & g_{t\phi} &= -\frac{2Mra \sin^2 \theta}{\Sigma}\end{aligned}$$

where $a = J/(Mc^2) \in [0, M)$ is the BH spin parameter, proportional to the BH angular momentum J , and $\Sigma \equiv r^2 + a^2 \cos^2 \theta$, $\Delta \equiv r^2 - 2rM + a^2 = (r - r_+)(r - r_-)$, where $r_{\pm} = M \pm \sqrt{M^2 - a^2}$. The metric becomes singular at $r_{\pm}$, but as in the Schwarzschild case these singularities are only a pathology of the metric and not physical ones; they define an outer (+) and inner (-) horizon; in the following, only the outer horizon is considered.

Before specializing the dynamics to the Kerr metric, recall that in General Relativity the trajectory of a test particle (i.e. a mass that is sufficiently small not to perturb the background metric) is described by geodesics on the spacetime manifold; in particular, once a metric $g_{\mu\nu}$ is defined, the equation of motion will correspond to the Geodesics Equation

$$\frac{dp^\mu}{d\tau} + \Gamma_{\nu\rho}^\mu p^\nu p^\rho = 0$$

where $p^\mu \equiv \frac{dx^\mu}{d\tau}$ is the contravariant 4-momentum of the massive particle, τ is the proper time of the test mass and the $\Gamma_{\nu\rho}^\mu$ are the Christoffel symbols associated with the metric $g_{\mu\nu}$. Despite being the most used formulation, in the context of trajectory integration it's often more useful to use the completely equivalent Hamiltonian formalism. In fact, it can be shown that the geodesics equation can be obtained directly from the least action principle once the correct Lagrangian is adopted:

$$\mathcal{L}(x, \dot{x}) = \frac{1}{2} g_{\mu\nu}(x) \dot{x}^\mu \dot{x}^\nu \quad (1.11)$$

To pass to the Hamiltonian formalism, the conjugate momenta are defined:

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{x}^\alpha} = g_{\alpha\nu} \dot{x}^\nu \quad (1.12)$$

and finally the Hamiltonian $\mathcal{H}$, with $\dot{x}^\alpha$ to be intended as a function of p_β via Eq. (1.12):

$$\mathcal{H} = p_\mu \dot{x}^\mu - \mathcal{L}(x, \dot{x}) = \frac{1}{2} g^{\mu\nu} p_\mu p_\nu = \frac{1}{2} g_{\mu\nu} \dot{x}^\mu \dot{x}^\nu \quad (1.13)$$

The Equations of Motion (EoM) are then given by the usual set of Hamilton equations

$$\dot{x}^\mu = \frac{d\mathcal{H}}{dp_\mu} \quad (1.14)$$

$$\dot{p}_\mu = -\frac{d\mathcal{H}}{dx^\mu} \quad (1.15)$$

This is a useful formulation to highlight the conserved quantities due to the presence of symmetries: if $\mathcal{H}$ is independent from a particular coordinate x^σ (and in this case x^σ is said to be *cyclic*), then from Eq. (1.15) it follows that $\dot{p}_\sigma = 0$, namely p_σ is a conserved quantity that can be exploited to separate the equations for different variables and disentangle the system.

The Kerr metric is stationary and axisymmetric; hence, since the coordinates t and ϕ are cyclic, two conserved quantities p_t and p_ϕ can be defined:

$$\begin{aligned} -E = p_t &= g_{t\mu} \dot{x}^\mu = g_{t\phi} \dot{\phi} + g_{tt} \dot{t} \\ L_z = p_\phi &= g_{\phi\mu} \dot{x}^\mu = g_{\phi\phi} \dot{\phi} + g_{\phi t} \dot{t} \end{aligned}$$

in addition, it presents another non-trivial symmetry other than time translation and rotation around z axis, which results in an additional constant of motion, namely the Carter constant (Carter, 1968):

$$Q = p_\theta^2 + \cos^2 \theta [a^2 (m_p^2 - E^2) + L_z^2 / \sin^2 \theta] \quad (1.16)$$

The conservation of the Carter constant allows for the complete decoupling of the EoM for r and θ , and thus the integrability of the system; an Hamiltonian system is *integrable* if there exist a variable change $(\mathbf{p}, \mathbf{q}) \mapsto (\mathbf{I}_p, \varphi_p)$ (action-angle variables) such that

$$\begin{aligned} \dot{\varphi} &= \frac{\partial H}{\partial I} \equiv \omega \\ \dot{I} &= -\frac{\partial H}{\partial \varphi} = 0 \end{aligned}$$

I is a prime integral of the motion (conserved quantity along the orbit) and φ is a vector of angular variables that describe the oscillatory motion and evolve linearly with time.

In practice, the existence of n prime integrals enable the separation of the EoM into a set of independent 1-dimensional systems, which can be brought to the form of n harmonic oscillators via the definition of proper angle variables. As a result, the motion occurs on the direct product of each variable's periodic domain, namely a n -torus T^n immersed in the $2n$ -dimensional phase space. If all the ratios between the frequencies are irrational,

the orbit covers densely the torus; to be more precise, the dimensionality of the (hyper-) surface densely covered is equal to the number of *rationaly independent angular velocities* of the system. In the case in which two or more frequencies are commensurate, the orbit closes and the torus hosts an infinite number of copies, distinguished only by an initial offset.

In the Kerr metric, four independent action variables which are constants of motion can be defined: energy and angular momentum along the z -axis from the two cyclic coordinates t and ϕ , the Carter constant Q from a tensorial Killing field, and finally the Lagrangian itself, which based on the normalization corresponds equivalently to the mass of the secondary (as the norm of the 4-momentum) or simply to -1 (as the modulus of 4-velocity). As a result, the system is integrable and can be solved separately for each coordinate.

Although the set of action variables (m_p, E, L_z, Q) uniquely define a trajectory, it is customary to work with a different set of coordinates which provide a more intuitive description of the characteristics of the orbit, namely the *semilatus rectum* p , the *orbital eccentricity* a and the *orbital inclination* θ_{inc} with respect to the equatorial plane. The relation between the two set of coordinates are algebraically involved, and reported in Gliorio et al. (2026). The implementation of geodesic motion in a Kerr background is provided by the code STORM (Scalar Tensor Orbital Radiation from EMRIs, Gliorio et al. (2026)), which will provide the foundation for the computation of the modified dynamics in the next chapters.

1.4.b Astrophysical relevance

EMRIs represent an unprecedented opportunity for probing with high precision new physics, strong gravity regimes, BH environments, formation channels and populations for SMBH, cosmology, physics of condensed matter and nuclear physics, and many more (Amaro-Seoane et al., 2015).

It is established that most of the galaxies, if not all, host a massive ($M \approx 10^6 M_\odot$) BH in the nucleus (Eisenhauer et al., 2005; Ghez et al., 2005; Gillessen et al., 2009; Magorrian et al., 1998; Heckman et al., 2014) that can act as the extremely massive primary, and a significant concentration of stellar-mass compact objects (Hailey et al., 2018; Miralda-Escude et al., 2000), which are candidates for the inspiralling secondary. Such compact objects can potentially undergo an inspiral around the primary, radiating energy and angular momentum through GW emission and reaching the innermost region, where strong gravity effects become relevant. This possibility is due to the relatively weak tidal forces across the secondary, thanks to the compactness and the extremely small mass ratio q , which prevent either the tidal disruption of the secondary (in principle possible for non-BH compact objects) or the substantial modification of the metric before the plunge (which is the case for the colliding solar-mass BH observed until now).

EMRIs will be used to precisely map environmental effects around SMBHs, such as the presence of gaseous disks (Dittmann et al., 2026; Kocsis et al., 2011; Yunes et al., 2011; Bonvin et al., 2017) which can modify the inspiral dynamics from pure-GR evolution, driven by GW emission, to a friction-dominated dynamics where the dissipative mechanism is provided by the interaction with the gaseous environment. In this case, the modifications to the waveform can provide valuable information on the density profile of the disk. Similarly, the inspiral can be modified by the presence of a dark matter cloud (Mitra et al.,

2025) which can act on the secondary either by dynamical friction (Vicente et al., 2022) as it was the case for the gaseous disk, or/and by deformations of the metric, which ceases to be described by the Kerr metric tensor since the energy-momentum tensor is not null anymore. This second case is of great observational relevance, since even little perturbations of the Kerr metric are expected to produce relevant modifications of the dynamics in correspondence to resonant configurations of the system (Pan et al., 2023; Brink et al., 2015b; Destounis et al., 2021; Destounis et al., 2023; Brink et al., 2015a) if the perturbation is non-integrable, namely the dynamical system does not admit anymore exactly 4 constants of motion (as it is the case for the Kerr metric, see previous Section). The relevant case for astrophysical systems, which are typically stationary and axisymmetric and thus conserve the action variables E and L_z due to the cyclic coordinates t and ϕ , is when the Carter symmetry is broken by the non-integrable perturbation; in this case, the Kolmogorov-Arnold-Moser theory (Arnold, 1989) predicts that for non-resonant orbits, namely configurations for which the characteristic frequencies associated to the radial (r) and angular (ϕ) action-angle variables are not even approximately rational dependent, the dynamics is only a small deviation of the unperturbed system, and the motion in the action-angle variable space continues to be foliated in nested invariant tori (as the integrable case). However, when the frequencies are resonant, the invariant tori are destroyed and a Birkhoff chain of stable/unstable *resonant islands* is formed. Across these resonant islands, the *rotational number* $\kappa = \omega_r/\omega_\phi$ is constant, leading to clear observable consequences in the GW waveform (*glitches* in the GW phase, and sharp features in the GW amplitude in correspondence with the resonant frequency). The resonant islands are embedded in a thin chaotic layer where positive Lyapunov exponents can influence the secular evolution of the system, amplifying small initial coordinate displacements that would have been otherwise not measurable; the very thin nature of this boundary layer could also allow for direct probing of the properties of the secondary, which is usually treated as a point-mass.

Other ways in which the GR dynamics is affected although the background is described by the Kerr metric involve dissipative mechanisms related to GR or to modified theories of gravity. Regarding the former possibility, the presence of the secondary can excite quasinormal modes of the primary (Santos et al., 2026), initiating a very complex evolution determined by the non-trivial interplay of different modes; the internal structure of the secondary, for example given by a close compact objects binary instead of a single mass particle, can also produce a wide variety of observable effects (Santos et al., 2025). The last aspect, namely testing the validity of General Relativity, represents an extensive field of research including the topics explored in the next Chapters of this Thesis. The possibility of following for a long time a GW signal tracing the evolution of a system dominated by gravity and reaching the strong field regime opens the possibility of testing the accuracy of GR predictions, and discover possible deviations hinting at new fundamental physics.

Among the proposed theories of modified gravity, some of which are presented in the next Chapter, the scalar-tensor theories involving the presence of a scalar field as an additional force mediator in addition to the standard graviton are being considered as potentially relevant in the context of EMRI dynamics. The relevance of new fields detection is not limited to astrophysics, but concerns also fundamental physics and cosmology (Clifton et al., 2012). The possible effects on the dynamical evolution are given by a variety of phenomena:

- **Additional radiative channel** provided by the scalar field, equivalent to a further polarization mode, which can speed up or even slow down the inspiral as I will motivate in the following. The energy, angular momentum and Carter constant fluxes due to scalar emission would be usually subdominant with respect to the GW fluxes (otherwise we would have already observed these effects in BBH mergers, A. G. Abac et al., 2026b), thus the main effect will consist in a little dephasing of the GW waveform (Chapter 5); however, under circumstances provided by EMRIs, these fluxes can substantially change the dynamics of the system giving rise to the unusual phenomenon of *floating orbits*, discussed in Chapter 6.
- **Scalar cloud formation** is another possibility, in which some scalar field modes are excited around the primary, and form a boson cloud (Hod, 2014; Benone et al., 2014) which can affect the evolution via dynamical friction (Tomaselli et al., 2023), by direct coupling with the secondary (Baumann et al., 2022b; Baumann et al., 2022a) or by deformations of the metric, as discussed before for ordinary matter; in this sense, scalar clouds would constitute a form of dark matter that would become detectable through GW observations.

The specificity of EMRIs as ideal probes for ultralight scalar fields with mass $m_s < 10^{-16}$ eV comes from the typical frequency of these objects, $f \lesssim 10^{-2} \text{ s}^{-1}$, which in natural units corresponds to a mass of about 10^{-17} eV. Massive field interaction are limited by a Yukawa damping factor $e^{-m_s r}$, thus the transition across this characteristic range will highlight the presence of an ultralight boson.

Theories of modified gravity and massive scalar fields

Since its first formulation, General Relativity (GR) has been confirmed on a wide set of gravitational phenomena (Bhattacharyya, 2019), from the deflection of light rays (Dyson et al., 1920) to extremely recent tests of its validity using GW observations of merging BBH (A. G. Abac et al., 2026b). Nevertheless, GR displays theoretical issues, such as the presence of singularities, as well as indirect observational problems, for example the late-time accelerated expansion of the Universe, which can be explained within the GR framework only by invoking the presence of a still unknown dark energy. Although GR is a sufficiently good description of gravity at least at the mesoscale level (Murata et al., 2015), namely between the laboratory scale, $\sim 10^{-4}$ m, and $\approx 10^{15}$ m (~ 0.1 pc), there is still the possibility of substantial modifications acting in the strong-field regime or at the long range. The former are motivated by the arising of singularities and quantum gravity phenomena on small scales, and would be necessary, for example, to describe primordial black holes (Hawking, 1971), the latter by the search for a purely gravitational explanation of dark energy. From an experimental point of view, a wide range of GR modifications is still possible, since their effect would be smaller than that of the systems currently probed (Bergé, 2017); note that outside the mentioned scales, GR is still very poorly constrained. From a purely theoretical point of view, GR is not the unique theory of gravity since other *self-consistent theories* proposing little modifications to GR are possible.

In this Chapter I will give a general overview of modified theories of gravity mainly based on Berti et al., 2015, with a specific focus on scalar-tensor theories, which represent the main topic (Sec. 2.1). In Sec. 2.2 I will describe the field equations for an additional massive scalar mediator of gravity in the astrophysical environment of a EMRI (see Sec. 1.4), motivating the usage of the field equations with a first-order mass ratio expansion.

2.1 Motivations and examples

2.1.a Beyond General Relativity: a classification of the alternatives

General Relativity is in fact well motivated by the Lovelock's theorem, which states that (Berti et al., 2015; Lovelock, 1971):

In four spacetime dimensions the only divergence-free symmetric rank-2 tensor constructed solely from the metric $g_{\mu\nu}$ and its derivatives up to second differential order, and preserving diffeomorphism invariance, is the Einstein tensor plus a cosmological term.

This implies that GR, defined by the Einstein Equations, Eq. (1.1), is the only formulation of a theory of gravity compatible with the Weak Equivalence Principle (through the divergence-free requirement) and diffeomorphism invariance (namely general covariance), with the only exception of the cosmological term, which is not included in Eq. 1.1. Despite being seemingly very general, it is possible to circumvent the theorem by violating the hypothesis in four different ways. (1) The theorem is valid only for a four-dimensional spacetime: if extra “hidden” dimensions are present, GR is not unique anymore; this is the case for Kaluza-Klein gravity (Overduin et al., 1997), in which both gravity and electromagnetism arise from a unified theory in five dimensions. (2) Weak Equivalence Principle (WEP) violation is obtained if the divergence-free hypothesis is dropped, leading to more general theories (Capozziello et al., 2011), which, however, are well constrained by strong experimental evidence for the validity of the WEP (Schlamminger et al., 2008). (3) Diffeomorphism invariance violation: dropping the requirement of general covariance (but usually not Lorentz covariance in Minkowski spacetime, which is experimentally verified with great accuracy) is possible and generates viable theories. (4) As the most relevant approach for this Thesis, the Lovelock theorem restricts a priori the gravitational interaction to the metric field: thus, adding additional fields acting as mediators of gravity will produce a viable extension of GR. Equivalently, it can be shown that enabling higher differential order in the metric derivatives is equivalent to asking second-order derivatives but introducing additional fields.

The “additional (scalar) field”, “broken diffeomorphism invariance” and “higher dimension” approaches predict the presence of still unknown scalar fields, the latter arising from the dimensionality reduction to the observable four dimensions. In other words, for most of the possible modifications of GR, despite the fundamental mechanisms being very different, the discrepancy will manifest *effectively* as the presence of an additional field. I will restrict to such scenarios, in which the standard gravity is complemented by the presence of a generic additional scalar field; this field can be either massless or massive.

Scalar-tensor theories could explain the inflationary dynamics of the primordial Universe, with the additional scalar driving inflation given a suitable slow-roll potential; similarly, the presence of a scalar field could explain the late-time cosmic acceleration, effectively acting as a dark energy component. In the following, I will formalize the presence of an additional boson within a classical field-theoretical approach.

2.1.b An overview of scalar-tensor theories

Scalar-tensor theories of gravity, in which the gravitational interaction is mediated by the graviton and one additional scalar field (generalizable to multiple scalar fields), can be described by the general action (Bergmann, 1968; Wagoner, 1970; Berti et al., 2015)

$$S = \frac{1}{16\pi} \int d^4x \sqrt{-g} \left[\phi R - \frac{\omega(\phi)}{\phi} \partial_\mu \phi \partial^\mu \phi - U(\phi) \right] + S_M(\psi; g_{\mu\nu}) \quad (2.1)$$

where ϕ is the scalar field, $\omega(\phi)$ and $U(\phi)$ are generic scalar functions of the field, and $S_M(\psi; g_{\mu\nu})$ is the action of the matter fields ψ ; note that the dependence from ϕ is entirely contained in the gravitational action, which acts as a modified Einstein-Hilbert action. This description, in which modified gravity and matter actions are independent, is called “Jordan frame”. However, it is possible to perform a conformal transformation of the

metric, $\tilde{g}_{\mu\nu} \equiv A^{-2}g_{\mu\nu}$, in such a way to recover a gravitational action that separates the scalar field contribution from the standard GR Lagrangian $\mathcal{L}_{\text{HE}} \propto \sqrt{-g}R$; in addition, a redefinition of the scalar field $\varphi = \varphi(\phi)$ is performed to reabsorb the function $\omega(\phi)$. Then, the transformation corresponds to $A(\varphi) = \phi^{-1/2}$, where $\phi = \phi(\varphi)$ through the inverse transformations, and the action becomes

$$S = \frac{1}{16\pi} \int d^4x \sqrt{-\tilde{g}} \left[\tilde{R} - \frac{1}{2} \partial_\mu \varphi \tilde{\partial}^\mu \varphi - V(\varphi) \right] + S_M(\psi; A^2(\varphi) \tilde{g}_{\mu\nu}) \quad (2.2)$$

where the quantities with the tilde are calculated from the transformed metric $\tilde{g}_{\mu\nu}$, and the potential V is defined by $V(\phi) = A^4(\varphi)U(\phi)$. This is known as the ‘‘Einstein frame’’, and exhibits a minimal coupling of the gravitational sector (the pure Hilbert-Einstein term is not multiplied by the field; instead, there is a standard Lagrangian for gravity and a standard Lagrangian for the scalar field, composed of a kinetic and a potential term in the usual form), with a non-minimal coupling of the matter sector: the field φ is affecting the transformed metric inside the matter action.

Note that the physical description of the dynamics is provided by the Jordan frame; however, since the results after an inverse transformation are the same, the Einstein frame can be used to describe the dynamics more easily, especially when matter is not present: in this case, $S_M = 0$ and the equations of motion reduce to a minimally coupled scalar field in a GR metric background.

A specific scalar-tensor theory is defined by the choice of the potential $U(\phi)$ or equivalently $V(\varphi)$, and the function $\omega(\phi)$. The potential term possibly encodes a mass term, crucial in the following, and self-interactions of the field, which are not explored in this work. The function $\omega(\phi)$ can be mapped in the Einstein frame to a *scalar coupling* $\alpha(\varphi) \equiv d(\ln A(\varphi))/d\varphi$ through $3+2\omega(\phi) = \alpha^{-2}(\varphi)$ (this is the explicit transformation that absorbs the function $\omega(\phi)/\phi$ in the definition of the new field φ); note that this quantity acts only through S_M , since the only non-minimal coupling is through the matter field.

The prototype of scalar-tensor theories defined by the action reported in Eq. (2.2), (2.1) are Brans–Dicke (BD) theories, where $\omega(\phi) = \omega_{\text{BD}}$ is constant and the potential is generic (for this reason, the choice actually corresponds to a wide class of theories). Also, theories that do not explicitly impose the existence of a scalar field can be reformulated as BD theories: for example, $f(R)$ theories in which the Ricci scalar in the Hilbert-Einstein action is replaced by a function of R , in such a way that modifications in specific curvature regimes can be applied. Indeed, it is possible to reformulate the $f(R)$ action $S_f = \frac{1}{16\pi} \int d^4x \sqrt{-g} f(R)$ to the equivalent $S_f = \frac{1}{16\pi} \int d^4x \sqrt{-g} [\varphi R - f(\phi) + \phi f'(\phi)]$ where again $\phi = \phi(\varphi)$, thus in the form of Eq. 2.1 with a potential determined by the functional form chosen for $f(R)$.

More general alternatives include a second-order coupling of the scalar field with the metric (Barsanti, 2024): for example, scalar Gauss-Bonnet (sGB) gravity is defined by the previously specified action including geometry and standard scalar field kinetic and potential term, plus an additional coupling $\hat{S}_c = \frac{\alpha_{\text{sGB}}}{4} \frac{1}{16\pi} \int d^4x \sqrt{-g} f(\varphi) \mathcal{G}$,¹ where $\mathcal{G} \equiv R^2 - 4R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\alpha\beta}R^{\mu\nu\alpha\beta}$ is a second-order scalar built from the Riemann tensor $R_{\mu\nu\alpha\beta}$, and $f(\varphi)$ is as usual a scalar function of the field. In particular, $f(\varphi) = \exp \varphi$ generates the

¹The coupling action is written with a hat ($\hat{S}$) since in the following I will consider separately $S_c = \hat{S}_c/\alpha$.

Einstein-dilaton scalar Gauss-Bonnet gravity, and $f(\varphi) = \varphi$ generates *shift-symmetric* gravity. These particular choices are motivated by string theory and are possible candidates for driving inflation (Odintsov et al., 2020; Oikonomou, 2021). Another way of coupling the scalar field to a quadratic combination of the Riemann tensor is the dynamical Chern-Simons gravity, in which $\hat{S}_c = \alpha_{\text{CS}} \frac{1}{16\pi} \int d^4x \sqrt{-g} \varphi \mathcal{P}$, where $\mathcal{P} = \frac{1}{2} R_{\mu\nu\alpha\beta} \epsilon^{\nu\mu\gamma\sigma} R^{\alpha\beta}_{\gamma\sigma}$; the structure is formally equal to scalar Gauss-Bonnet, with a different quadratic coupling scalar. In the following, theories with a quadratic term are included in the treatise, provided that the physical dimension of the coupling α is a mass to the power of n , with $n \geq 2$; the motivation is given in Sec. 2.2.b, but note that the theories presented here satisfy this requirement since $[\alpha_{\text{sGB/CS}}] = \text{Mass}^2$, which can be directly found by dimensional analysis of the action, that in geometrized units is dimensionless.

In conclusion, each theory is defined by specific physical motivations, but many of these mathematically reduce to an action involving a scalar field and possibly a higher-order coupling mediated by a coupling constant α ; the mapping between the theory parameters and the general description can be recovered after a general treatise, both theoretically and observationally. For this reason, in the following I will not refer to a specific theory but simply obtain the dynamical signatures of a scalar field using the general description of scalar-tensor theories in the Einstein frame.

2.1.c Current estimations and upper limits

Quantitative tests of GR can be performed in a wide range of varieties; in particular, it is possible to perform *agnostic* or *theory-specific* tests. The former challenge a particular aspect of GR, verifying at which extent it holds, and in particular if it is possible to falsify the null hypothesis $\Theta_{\text{experiment}} - \Theta_{\text{GR}} = 0$ in an experimental setting, where $\Theta_{\text{experiment}}$ are the results of some experiment and Θ_{GR} is the corresponding prediction of GR. The latter aim to find the values of the characteristic parameters of a specific theory assuming its validity; this is a correct approach since modified gravity theories are continuously connected to GR, thus if the full theory is not describing the results the test will simply return the GR limit. Note that agnostic tests are generally less powerful in constraining deviations, since they are more general by design. Theory-specific tests can constrain the values of the single parameters, although their specificity is an intrinsic limitation. Note that theory-agnostic tests can always be converted to theory-specific tests just by considering the effect that a particular theory would have in that particular context, although resulting in looser constraints. Although deviations from GR have not been observed until now, the tests that were performed pose upper limits to the strength of the additional interactions under consideration. In the following, an overview of the various tests is provided.

Theory-agnostic tests of GR

A detailed explanation of agnostic tests is presented in Clifton et al. 2012; they are mainly based on the parametrized post-Newtonian (PPN) formalism, a flexible theoretical setup that allows for a wide class of modification of GR in the weak field regime. The approach is parametric in the sense that metric deviations from GR are phenomenologically described by a (wide) set of parameters, each carrying a physical significance. Although the exact

definition of such parameters is not relevant here, I briefly mention some of the tests that were used to constrain them.

- **Light deflection and Shapiro delay:** using the Sun's gravitational potential, in terms of the PPN parameters the angular deflection of light is given by $\theta = \theta_{\text{GR}}(1 + \gamma)/2$, where $\gamma_{\text{GR}} = 1$. Analogously, the Shapiro delay Δt_s is corrected by the same prefactor. Bounds in γ can be obtained experimentally, resulting in $\gamma - 1 = (-1.7 \pm 4.5) \cdot 10^{-4}$ for light deflection and $\gamma - 1 = (2.1 \pm 2.3) \cdot 10^{-5}$ for the Shapiro delay Clifton et al. 2012.
- **Nordtvedt effect and lunar laser ranging:** tests of the strong equivalence principle for the Moon-Earth system in the Sun's gravitational field. Again, no statistically significant discrepancies from GR are found.
- **Torsion-balance pendulum:** used to test the weak equivalence principle, measuring the difference in the gravitational acceleration of distinct materials with the same inertial mass. Accelerations are verified to be the same up to 1 part of 10^{13} .
- **Tests of Yukawa potential:** more important for our case of study, namely massive fields, tests of a Yukawa-type gravitational potential with a component behaving as $V \sim \alpha e^{-r/\lambda}/r$ characterized by a characteristic scale λ dependent on the mass of the mediator and a coupling constant α . The determination of an upper limit for α is dependent on the chosen scale λ , thus a scan of the scales is usually performed with a variety of techniques. The results range from $\alpha \lesssim 10^{-8}$ on the planetary scales (10^8 – 10^{12} m) down to $\alpha \lesssim 10^{-3}$ on the sub-meter lab scales.
- **GW residual tests:** we recently entered the era of precision GW astrophysics, and accurate waveform models are available, reproducing with sufficient accuracy the GR predictions. This means that when a GW template with the correct reconstructed parameters is subtracted from the datastream of a detected CBC event, the residuals can act as a generic test of GR: if the residual statistic is not compatible with only noise, then the null hypothesis can be rejected. At the moment, the observed signals confirm the GR predictions (A. G. Abac et al., 2026b).
- **Extra polarizations:** the existence of additional scalar / vector GW modes associated with a spin-0 or spin-1 mediator of gravity would have the effect of introducing additional polarization modes, which in principle are detectable by the current interferometers. Using the GWTC-5 catalog (A. G. Abac et al., 2026b) it is possible to assign a Bayes factor of -19.5 (extremely strongly disfavoured) to the scalar-only hypothesis (as expected), and a completely neutral Bayes factor of 0.01 ± 0.7 to a generic scalar-tensor theory. This means that based on the observations it is not yet possible to distinguish the presence of a scalar field, if it is actually present. Similar results hold for vector, vector-tensor and scalar-vector-tensor theories, with strongly disfavoured hypotheses lacking a tensor component and inconclusive results for the presence of a further mediator beyond the standard graviton.
- **Ringdown / quasi-normal modes:** after merging, two BH produce a single rotating BH with strong perturbations, which are radiated away via quasi-normal modes (QNM), corresponding to damped oscillations of a discrete set of resonant frequencies associated to the Kerr metric. Once the system parameters are estimated from the inspiraling/merging phase, the QNM spectrum can be determined through an analytic treatise and compared to the observed data, providing a test for the Kerr metric itself. In particular, we can expect QNM to be modified by the presence of a BH "hair", namely an additional characteristic beyond mass and spin. This is particularly interesting for scalar-tensor theories such as sGB, since they predict the

possibility for a BH to acquire a charge. At present, no anomaly in the QNM spectrum has been found yet (A. G. Abac et al., 2026b).

Scalar-tensor specific tests

Specific tests are obtained either from predicting the phenomenology of a theory in the context of theory-agnostic test setup or from designing specific fits. Some examples are given by:

- **Bounds on massive BD theories:** Alsing et al. 2012 applies existing agnostic tests to the specific case of massive BD scalar-tensor theories introduced before, obtaining lower bounds on ω_{BD} as a function of the scalar mass m_s . The theory is extremely poorly constrained for scalar masses larger than 10^{-17} eV, with the constraint $\omega_{\text{BD}} > 50$. Tighter bounds apply for smaller masses; this strongly motivates the search for ultralight bosons through EMRIs dynamics, which, as motivated above, would be sensitive to scalar masses of about 10^{-16} eV.
- **Bounds on GB gravity:** the EdGB gravity effects become relevant when the coupling term becomes $\mathcal{O}(1)$. Approximately, the coupling behaves as $S_c \sim \alpha R \sim \alpha/L^2$, where L is the typical scale of the spacetime curvature. Thus, tests of GR on a reference scale L produce a bound on α_{GB} . On the Solar System scales (low curvature), the coupling can be loosely constrained as $\sqrt{\alpha_{\text{GB}}} \lesssim 10^{13}$ cm, while strong gravity regimes such as those of binary neutron stars it is possible to measure $\sqrt{\alpha_{\text{GB}}} \lesssim 5 \cdot 10^6$ cm (Berti et al., 2015). For BBH tests of GR, the curvature being $L \sim R_{\text{Sch}} \sim M$ we can find $\alpha_{\text{GB}}/M^2 \lesssim 0.7$. The GWTC-5 data release constrained in this way $\sqrt{\alpha_{\text{EdGB}}} \lesssim 1.32$ km (A. G. Abac et al., 2026b), but a previous tighter bounds existed due to the GW event GW230529, which resulted in the tighter constraint $\sqrt{\alpha_{\text{EdGB}}} \lesssim 0.28$ km (Sänger et al., 2026). In the Bergmann-Wagoner formalism, this would correspond to a scalar charge $d \lesssim 0.02$ (Speri et al., 2024). Note, however, that this is valid for a massless scalar field: introducing a scalar mass has the effect of constraining less the charge, since mass and charge are highly degenerate; for this reason, in the following even larger scalar charges are considered, still representing a realistic scenario compatible with current bounds.

2.2 Massive scalar fields around EMRIs

2.2.a Gravitational and massive scalar field equations of motion

In the following, I will adopt the Einstein frame description. This choice, already motivated previously, allows for a simpler treatment of the scalar field equations, since the field is minimally coupled to gravity. The price is that the matter action also depends on the field; however, since the only source to be included in the stress-energy tensor is the point-like particle representing the secondary, the adaptation is very easy and can be addressed by matching the expansion of the field in a neighborhood of the source with the expected Yukawa potential.

The dynamics of the whole problem, consisting of the metric field and the scalar field, is thus set by the action:

$$S(\mathbf{g}, \phi, \psi) = S_0(\mathbf{g}, \phi) + \alpha S_c(\mathbf{g}, \phi) + S_M(\mathbf{g}, \phi, \psi)$$

where S_0 encodes geometry, the usual kinetic term and mass term of the scalar field, thus all the gravitational interactions, S_c the coupling between the scalar and the tensor mediators, and S_M is the matter action accounting for the secondary.

These contributions explicitly read:

- Modified Einstein-Hilbert action, including scalar field contribution:

$$S_0 = \int d^4x \frac{\sqrt{-g}}{16\pi} \left(R - \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{1}{2} \mu_s^2 \phi^2 \right)$$

Note: the global rescaling of the scalar field is arbitrary and is set by this convention for the mass: the choice of this action (in particular the normalizing constant inherited from the GR Einstein-Hilbert action) fixes the mass definition. Note that since geometrized units $G = c = 1$ are used, the length has the physical dimensions of a mass. This means that $\mu_s \equiv m_s/\hbar$ has the dimension of the inverse of a mass. Choosing this action selects a massive scalar field without self-interactions; this can be considered either as an exact formulation of a free field theory or an approximation for a very weakly self-interacting field. This second hypothesis is motivated also by the fact that the interaction between the graviton and the boson, encoded in the coupling action, will be discarded thanks to the q -expansion (see Sec. 2.2.b).

- Matter action: in GR, $S_p = -m \int d\tau = -m \int d\lambda \sqrt{g_{\alpha\beta} \dot{y}^\alpha \dot{y}^\beta}$. Here in general we assume a dependence of the mass from the scalar field (the only way to couple ϕ with the matter Lagrangian):

$$S_M = S_p = - \int m(\phi) \sqrt{g_{\alpha\beta} \frac{dy_p^\alpha}{d\lambda} \frac{dy_p^\beta}{d\lambda}} d\lambda = - \int d^4x \int d\tau m(\phi) \delta^{(4)}(x - y_p(\tau))$$

In the last step I introduced a $\delta^{(4)}$ to write the integral over d^4x , allowing later the application of a functional derivative. In addition, I prefer to explicitly indicate the affine parameter as the proper time τ of the particle, instead of a generic affine parameterization λ , just for clarity.

We obtain the field equations through the Euler-Lagrange equations, or equivalently requiring null functional derivative with respect to the scalar field ($\rightarrow$ scalar field equations) and to the metric tensor ($\rightarrow$ Einstein Equations with source term given by both the secondary and the field). The functional derivative with respect to a generic tensorial field $\chi_{\mu_1, \dots, \mu_n}$ is given by applying to the integrand under $\int d^4x$ the following differential operator:

$$\frac{\delta}{\delta \chi_{\mu_1 \dots \mu_n}} \equiv \frac{\partial}{\partial \chi_{\mu_1 \dots \mu_n}} - \partial_\mu \frac{\partial}{\partial \chi_{\mu_1 \dots \mu_n} \partial_\mu}$$

Field equations:

$$\begin{aligned} 0 = \frac{\delta S}{\delta \phi} &= \frac{\sqrt{-g}}{16\pi} \left(-\mu_s^2 + \partial_\mu \partial^\mu \right) \phi - \int \frac{\partial m(\phi)}{\partial \phi} \delta^{(4)}(x - y_p(\tau)) d\tau \\ &= \frac{\sqrt{-g}}{16\pi} \left(\square - \mu_s^2 \right) \phi - \int m'(\phi) \delta^{(4)}(x - y_p(\tau)) d\tau \end{aligned}$$

In the second passage, the notation is simplified, thus leading to the Klein-Gordon-like equation

$$(\square - \mu_s^2)\phi = 16\pi \int m'(\phi) \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} d\tau$$

Note that I can bring the square root of the metric determinant under integration due to the presence of the Dirac $\delta^{(4)}$. The same goes for the metric tensor field equations (analogous to Einstein equations): let's compute all the different terms:

- Einstein-Hilbert GR action, involving only R : this gives rise to a term

$$\frac{\delta S_{\text{EH}}}{\delta g_{\mu\nu}} = \frac{\sqrt{-g}}{16\pi} G_{\mu\nu}$$

Since $\sqrt{-g} \neq 0$ always, this reduces to the vacuum Einstein Equations. This is the classical derivation and follows directly from expliciting R as a function of $g_{\mu\nu}$ and its derivatives.

- Scalar field kinetic and mass terms: we make use of $\frac{\partial \sqrt{-g}}{\partial g^{\mu\nu}} = -\frac{1}{2}g_{\mu\nu}\sqrt{-g}$ and open the term $\partial_\mu \phi \partial^\mu \phi = \partial_\mu \phi g^{\mu\nu} \partial_\nu \phi$ to make explicit the dependence from the metric tensor, finally obtaining

$$\frac{\delta S_{\text{scalar}}}{\delta g^{\mu\nu}} = -\frac{\sqrt{-g}}{16\pi} \left[-\frac{1}{4}g_{\mu\nu} \partial_\alpha \phi \partial^\alpha \phi - \frac{1}{4}g_{\mu\nu} \mu_s^2 \phi^2 + \frac{1}{2} \partial_\mu \phi \partial_\nu \phi \right]$$

This will later correspond to a stress-energy tensor; since we want to write the EE in a form like $G_{\mu\nu} = 8\pi T_{\mu\nu}$, it is sufficient to multiply by a factor $-\frac{16\pi}{\sqrt{-g}} \frac{1}{8\pi} = -\frac{2}{\sqrt{-g}}$. The stress-energy tensor for the scalar field takes the following form:

$$T_{\mu\nu}^{\text{scalar}} = \frac{1}{16\pi} \left[-\frac{1}{2}g_{\mu\nu} \partial_\alpha \phi \partial^\alpha \phi - \frac{1}{2}g_{\mu\nu} \mu_s^2 \phi^2 + \partial_\mu \phi \partial_\nu \phi \right] \quad (2.3)$$

- Mass term: the dependence of the matter action from $g_{\mu\nu}$ is only through the term $\sqrt{g_{\alpha\beta} \frac{dy_p^\alpha}{d\lambda} \frac{dy_p^\beta}{d\lambda}}$, and there is no dependence from $\partial_\alpha g_{\mu\nu}$. Thus:

$$\begin{aligned} \frac{\delta S_p}{\delta g_{\mu\nu}} &= - \int m(\phi) \delta^{(4)}(x - y_p(\lambda)) \frac{1}{2} \frac{1}{\sqrt{g_{\alpha\beta} \frac{dy_p^\alpha}{d\lambda} \frac{dy_p^\beta}{d\lambda}}} \frac{dy_p^\alpha}{d\lambda} \frac{dy_p^\beta}{d\lambda} d\lambda \\ &= - \frac{1}{2} \int m(\phi) \delta^{(4)}(x - y_p(\tau)) \frac{dy_p^\alpha}{d\tau} \frac{dy_p^\beta}{d\tau} d\tau \end{aligned}$$

Again, we want to express this in terms of a stress-energy tensor by multiplying by $-\frac{2}{\sqrt{-g}}$, thus obtaining

$$T_{\mu\nu}^p = \int m(\phi) \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} \frac{dy_p^\alpha}{d\tau} \frac{dy_p^\beta}{d\tau} d\tau$$

We still have an unknown dependence of the mass from the field, $m(\phi)$. We approxi-

mate this dependence with a linear expansion, $m(\phi) \cong m_p + \frac{\partial m}{\partial \phi} \phi$, since $\phi \sim \mathcal{O}(q)$. We find the derivative using the weak field approximation: outside the particle's trajectory, the field equation reads $(\square - \mu_s^2)\phi = 0$, which is solved (in flat spacetime, which we can assume in a small neighborhood of the secondary in a free-falling frame) by the Yukawa potential

$$\phi \cong \frac{m_p d}{\tilde{r}} e^{-\mu_s \tilde{r}}$$

where $m_p d$ is the scalar charge of the secondary, with m_p the gravitational charge (mass). Applying the Klein-Gordon (KG) operator in a local neighborhood of the secondary results in

$$(\square - \mu_s^2)\phi = -4\pi m_p d \delta^{(3)}(x - y_p)$$

Compared to the source term in the field equation, we obtain $m'(\phi) = -\frac{m_p d}{4}$, thus the final field equations read:

$$(\square - \mu_s^2)\phi = -4\pi m_p d \int \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} d\tau \quad (2.4)$$

Note that we must use the linear expansion since the source term is proportional to $m'(\phi)$, which happens to be proportional to m_p , resulting in the first-order approximation in q .

The matter stress-energy tensor instead depends directly on $m(\phi)$, and thus it is sufficient to use the first-order term m_p , hence

$$G^{\mu\nu} = 8\pi m_p \int \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} \frac{dy_p^\mu}{d\tau} \frac{dy_p^\nu}{d\tau} d\tau$$

2.2.b Leading and subleading order terms in mass ratio

Throughout this work, a perturbative approach in the expansion parameter q is adopted: the zero-order solution would be a stationary Kerr BH, since this would correspond to setting $q = 0$; the relevant effects can be seen at first order in q , which is the fundamental choice for the rest of the theoretical approach. This allows for a simplified treatise of the system, that becomes analytically and computationally approachable thanks to linearity; for this reason, it is important to clarify the expansion order of the fundamental quantities and their combinations, and develop the theory discarding contributions beyond linear order. In the following, I will specify and comment the expansion order of the main quantities.

- **Background metric:** if the no-hair theorem (Bekenstein, 1972; Saa, 1996; Banerjee et al., 1998) is valid, the background must be simply described by the Kerr metric. It is however possible to evade no-hair theorems (Sotiriou et al., 2014b; Sotiriou et al., 2014a); in this case, let's consider a coupling constant α has dimensions of mass^n with $n \geq 2$ (because S_c is in general not dimensionless: for example, scalar Gauss-Bonnet gravity can have $\mathcal{L} \propto \phi$ which is dimensionless, thus taking into account the integrand measure d^4x , it must be that $[\alpha] = \text{mass}^2$). The corrections to the Kerr metric must then be described by an adimensional parameter $\zeta \equiv \alpha/M^n \equiv q^n \zeta_p$, where $\zeta_p \equiv \alpha/m_p^n$. Observationally $\zeta_p < 1$ (see Sec. 2.1.c), so we can consider the

correction to be $\mathcal{O}(q^n)$, and since $n \geq 2$ we can ignore it. Although it is important to notice that this is actually a restriction of the possible theories that can be probed with this theoretical framework, the condition is actually satisfied for the most relevant scalar-tensor theories of gravity, such as the previously commented scalar Gauss-Bonnet. This restriction allows for the usage of known analytical results relative to the Kerr metric, which enjoys particularly convenient characteristics such as symmetries and integrability, and for which computational infrastructures are already available.

Result: the background metric is exactly the Kerr metric, without corrections (**order 0 in q**)

- **Scalar field ϕ :** the scalar field is dimensionless, as can be seen from the modified Hilbert-Einstein action S_0 : morally, “ $R \sim \partial\partial g$ ”, where the metric $g_{\mu\nu}$ is dimensionless, thus comparing this term with the kinetic and mass terms, it is evident that the field itself is dimensionless. We write $\phi = \phi_0 + \phi_1$, where ϕ_0 is a constant background field ($\partial_\mu \phi_0 = 0$) and ϕ_1 is a perturbation, which is of our interest. For the case of a massless field, it is easy to check that $\phi_1 \sim \mathcal{O}(q)$, since the KG equation corresponds to $\square\phi = \square(\phi_0 + \phi_1) = \square\phi_1 = \mathcal{O}(q)$ because the source term being proportional to m_p . For the massive case, the calculation is more involved: the KG equation without the $\mathcal{O}(q)$ source reads $(\square - \mu_s^2)(\phi_0 + \phi_1) = 0 \implies (\square - \mu_s^2)\phi_1 = \mu_s^2\phi_0$ which is KG equation for ϕ_1 with a constant source term (potential). If we assume $\phi_0 \neq 0$, the Green’s function method can be used to compute ϕ_1 . the Green’s function for the KG operator is $\phi_g(x) = \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ikx}}{-k^2 - \mu_s^2}$ and the general solution is given by

$$\begin{aligned} \phi_1 &= \int d^4y \mu_s^2 \phi_0 \phi_g(x-y) = \int d^4y \mu_s^2 \phi_0 \int \frac{d^4k}{(2\pi)^4} \frac{e^{-ik(x-y)}}{-k^2 - \mu_s^2} \\ &= \mu_s^2 \phi_0 \int d^4k \frac{e^{-ikx}}{-k^2 - \mu_s^2} \delta^{(4)}(k) \\ &= -\phi_0 \end{aligned}$$

This means that $\phi = 0$ everywhere without a source term. If we introduce a source as explained previously, we obtain proportionally $\phi \sim \mathcal{O}(q)$, which directly follows from the Green’s function method with a $\mathcal{O}(q)$ source term.

Result: for a massive scalar field, $\phi \sim \mathcal{O}(q)$

- **Scalar field stress-energy tensor:** the explicit form has been derived previously and always involves the field amplitude ϕ quadratically, meaning that $T_{\mu\nu}^{\text{scalar}} \sim \mathcal{O}(q^2)$ and thus can be ignored at linear order. It is important to note that, differently from the massless case, the stress-energy tensor for the massive field features a term $g_{\mu\nu} \mu_s^2 \phi^2$: thus it was important to check that the whole ϕ was actually $\mathcal{O}(q)$, and not only the non-constant part as in the massless case, since this term does not involve derivatives that remove a constant component $\phi_0 \sim \mathcal{O}(1)$. This is an additional independent condition needed for stating that the background metric is the Kerr metric: while the first point was evaluating the possible effects of a scalar charge of the primary on the metric and found that in the described restrictive hypothesis the metric continues to be Kerr at least at the first order, here we assess that the scalar field modes excited by the secondary will not contribute, at first order, to the stress-energy tensor, which would affect metric perturbations through the Einstein Equations.

Result: for a scalar field, $T_{\mu\nu}^{\text{scalar}} \sim \mathcal{O}(q^2)$

- **Coupling term in the global action:** The argument is analogous to the one presented for the background metric: since $[S_c] = \text{mass}^{-n}$ due to the coupling constant α , for a dimensional scaling argument we expect $S_c \sim M^{-n} S_0$, so that the coupling action $\alpha S_c \sim \frac{\alpha}{M^n} S_0 \ll S_0$ and in particular is at least $\mathcal{O}(q^2)$.
Result: the coupling terms can be ignored since $\alpha T_{\mu\nu}^c \sim \mathcal{O}(q^{n \geq 2})$
- **Matter stress-energy tensor:** $S_M \propto m_p$, thus this results in a first-order effect in the mass ratio q . This is by definition and constitutes the foundation of the first-order post-adiabatic approach.

In conclusion, the first-order approach is based on a rather simple and ordered set of “motivated prescriptions”:

- The secondary follows a trajectory determined by geodesics in Kerr spacetime; the orbits are described by the usual action variables (E, L_z, Q) which undergo a secular evolution due to the GW and scalar fluxes that are found at the first post-adiabatic order.
- The scalar field is described by a massive KG equation in a curved spacetime, without back-reaction on the spacetime itself due to the field stress-energy tensor; the source term of the field is given by the adiabatic description (0 order in q of the secondary trajectory), which, as seen before, is a geodesic in Kerr spacetime.
- The perturbations of the gravitational field, corresponding to GWs, are conceptually equivalent to the scalar field: they can be treated as a separate (tensorial) field on a Kerr background and governed by the linearized Einstein equations with a source term given by the secondary.
- The post-adiabatic evolution of the system is obtained by evolving the action variables accordingly to the relative GW and scalar fluxes. The evolution happens on secular timescales, namely the dynamical orbit can be considered as a series of slowly-changing Kerr geodesics (which are not dissipative in the test-mass approximation).

This clearly defines a program to compute the dynamics of massive scalar fields: given an orbital configuration and using the KG equation it is possible to find the solution of the field amplitude on the entire spacetime domain; this is an “adiabatic” solution, in the sense that it describes a stationary configuration of the system; we will exploit the stationarity to decompose the system in a continuous energy spectrum. Once the field is known, the integration of the stress-energy tensor on the boundary surfaces, one at the BH horizon and one at spatial infinity, will provide the energy, angular momentum and Carter constant fluxes which, together with the corresponding GW solution and fluxes, determine the evolution of the secondary. Since the computation of the fluxes is adiabatic, it depends only on the current state of the system and not on the past evolution; this means that the flux evaluation and the subsequent evolution can be treated subsequently and not simultaneously, since the evolution has no back-action on the flux computation. This would not be true for an exact treatise of the dynamics.

Physically, this treatment is motivated by the evolution of the system on two very different timescales, the *orbital* timescale and the *secular evolution* timescale; since the two are clearly separated, it is possible to consider phenomena that evolve secularly as stationary over the orbital timescale, and phenomena that happen on the orbital timescale can be averaged over a large number of cycles when considering secular effects.

3

Solutions to the field equations in a curved background

This Chapter is entirely dedicated to the development of the theoretical foundation for the numerical solution of the scalar field and its characterization in terms of energy and mode decomposition, and scalar fluxes (of energy and angular momentum) at spatial infinity and at the horizon. I will begin with a review of the Teukolsky formalism that is commonly used to treat small perturbation of massless fields in the Kerr spacetime, and later adapt it to treat a massive scalar field. Finally, considering the source term and the homogeneous solutions to the massive master equation, it is possible to compute the field magnitude on the entire spatial and energy domains, resulting in the computation of the scalar fluxes.

3.1 Teukolsky formalism for massless fields

For a massless scalar field, the master equation is set by $\square\phi = T^{(0)}$, where $T^{(0)}$ is a source term (cf. Eq. 2.4) and the Dalemberertian is computed in a curved spacetime manifold. The same procedure can be applied to spinors ($s = 1/2$, Dirac equation), vectors ($s = 1$, Proca equation) and tensors $s = 2$, particularly relevant here since they represent gravitons, or gravitational wave perturbation in the classical field approach. Very conveniently, the explicit equations of motion specialized to the Kerr metric (which enters through the differential operators) for different spins can be written as a unified master equation, called the Teukolsky equation (Saul A. Teukolsky, 1973; S. A. Teukolsky et al., 1974). Let's define the Teukolsky operator $\mathcal{T}^{(s)}$ for a spin- s perturbation as:

$$\begin{aligned} \mathcal{T}^{(s)} \equiv & \left[\frac{(r^2 + a^2)^2}{\Delta} - a^2 \sin^2 \theta \right] \partial_0^2 + \frac{4Mar}{\Delta} \partial_0 \partial_\phi + \left[\frac{a^2}{\Delta} - \frac{1}{\sin^2 \theta} \right] \partial_\phi^2 - \\ & - \Delta^{-s} \partial_r \Delta^{s+1} \partial_r - \frac{1}{\sin \theta} \partial_\theta \sin \theta \partial_\theta - \\ & - 2s \left[\frac{a(r - M)}{\Delta} + i \frac{\cos \theta}{\sin^2 \theta} \right] \partial_\phi - 2s \left[\frac{M(r^2 - a^2)}{\Delta} - r - ia \cos \theta \right] \partial_0 + (s^2 \cot^2 \theta - s) \end{aligned}$$

Where the functions $\Delta(r, a)$ and $\Sigma(r, a, \theta)$ are defined in Sec. 1.4.a and inherited from the Kerr metric through the explicit definition of the GR Delambertian operator. The Teukolsky equation for a perturbation field ψ simply reads

$$\mathcal{T}^{(s)}\psi = 4\pi\Sigma T^{(s)}$$

where $T^{(s)}$ is the source term inherited from the field equations. For the scalar I showed that $T^{(0)} = -m_p d \int \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} d\tau$, and analogue expressions are valid for $s = 1/2, 1, 2$. This allows to write different spin perturbations with a unified master equation, but having custom expressions for each source term and field representation, which is a great advantage when solving the whole differential equation with the Green's function method.

The Teukolsky equation can be solved more easily by decomposing the operator into a series of multipoles, each having a factorized angular and radial part, each expressed in frequency domain through a Fourier transform. This decomposition must be done simultaneously for the Teukolsky operator and the source term:

$$\begin{aligned}\psi^{(s)} &= \int d\omega \sum_{\ell, m} R_{\ell, m}^{(s)}(r, \omega) S_{\ell, m}^{(s)}(\theta, \omega) e^{im\phi} e^{-i\omega t} \\ 4\pi \Sigma T^{(s)} &= \int d\omega \sum_{\ell, m} J_{\ell, m}^{(s)}(r, \omega) S_{\ell, m}^{(s)}(\theta, \omega) e^{im\phi} e^{-i\omega t}\end{aligned}$$

Some observations on this passage:

- This is possible since the Teukolsky operator is linear in the perturbation; this is enforced by the linear treatment of the perturbation problem.
- Here we indicate the *spin-weighted spheroidal harmonics* $S_{\ell, m}^{(s)}(\theta)$ as a real function of the angular coordinate θ ; the dependence from the angular coordinate ϕ is factorized as a complex phase, similarly to what is done for spherical harmonics.
- The spheroidal harmonics are orthonormal in the usual sense, and differently from the simpler spherical harmonics they account for the dragging terms in the Kerr metric. These two properties will be evident from the definition of the angular part of the Teukolsky equation; for now they represent very general requirements of a mode decomposition.
- The partial derivatives in t and ϕ within $\mathcal{T}^{(s)}$ become $\partial_0 \mapsto -i\omega$ and $\partial_\phi \mapsto im$. Thus formally $\mathcal{T}^{(s)} = \mathcal{T}^{(s)}(r, \theta, \partial_r, \partial_\theta)$.

Now we can separate the angular and radial contributions: starting from the homogeneous equation $\mathcal{T}^{(s)}\phi = 0$, the Teukolsky operator can be schematically separated as follows

$$\mathcal{T}^{(s)} R(r) S(\theta) = [f_1(r) + f_2(\theta)] R(r) S(\theta) + R(r) \tilde{D}_\theta S(\theta) + S(\theta) \tilde{D}_r R(r) = 0$$

Where $\tilde{D}_r$ and $\tilde{D}_\theta$ are differential operators in the respective variables and derivatives. This can be separated as:

$$R(r) [\tilde{D}_\theta + f_2(\theta)] S(\theta) = -S(\theta) [\tilde{D}_r + f_1(r)] R(r)$$

If either R or S are zero, the equation becomes trivial. If not, we can separate the equation in radial and angular part, which must be equal to a constant A since they involve different independent variables. The final equations take the form of an eigenproblem:

$$\begin{aligned} \left[\tilde{D}_\theta + f_2(\theta) \right] S(\theta) &= -AS(\theta) \\ \left[\tilde{D}_r + f_1(r) \right] R(r) &= AR(r) \end{aligned}$$

inserting the actual expressions for the differential and algebraic operators, we obtain the angular equation (defining the *obliteness* $\gamma \equiv a\omega$)

$$\begin{aligned} \left[\frac{1}{\sin \theta} \partial_\theta \sin \theta \partial_\theta - \gamma^2 \sin^2 \theta - \frac{m^2}{\sin^2 \theta} - 2sm \frac{\cos \theta}{\sin^2 \theta} - 2s\gamma \cos \theta - s^2 \cot^2 \theta + s + A \right] S(\theta) &= 0 \\ \left[\frac{1}{\sin \theta} \partial_\theta \sin \theta \partial_\theta - \gamma^2 \sin^2 \theta - \frac{(m + s \cos \theta)^2}{\sin^2 \theta} - 2s\gamma \cos \theta + s + A \right] S(\theta) &= 0 \end{aligned}$$

Note that since the angular decomposition is the same at RHS and LHS of the Teukolsky equation, spin-weighted spheroidal harmonics are universal and independent from the realization of the problem; they depend only the field's spin and the obliteness parameter, which results from the BH angular momentum and the considered energy.

The radial part of the master equation after some algebra becomes:

$$\left[\Delta^{-s} \partial_r \Delta^{s+1} \partial_r + \frac{K^2 - 2is(r - M)K}{\Delta} + 4is\omega r + 2m\gamma - A \right] R(r) = 0$$

Where $K = (r^2 + a^2)\omega - ma$. We define the *angular eigenvalue* as ${}_s\lambda_{\ell m} \equiv A - 2m\gamma$. This definition, instead of considering the eigenvalue simply as A , comes from the fact that in the angular equation the eigenvalue can be chosen with an arbitrary offset without changing much the expression, while relatively simplifying the radial equation (which will be the main problem to solve).

Note: in the original papers (Saul A. Teukolsky, 1973; S. A. Teukolsky et al., 1974) the definitions are slightly different, due to the fact that Teukolsky prefers to write a term $\gamma^2 \cos^2 \theta$ instead of $-\gamma^2 \sin^2 \theta$, with a consequent extra term γ^2 within the definition of the angular eigenvalue.

The decoupled homogeneous Teukolsky equations thus become:

$$\begin{aligned} \left[\frac{1}{\sin \theta} \partial_\theta \sin \theta \partial_\theta - \gamma^2 \sin^2 \theta - \frac{(m + s \cos \theta)^2}{\sin^2 \theta} - 2s\gamma \cos \theta + s + 2m\gamma + {}_s\lambda_{\ell m} \right] S_{\ell m}(\theta) &= 0 \\ \left[\Delta^{-s} \partial_r \Delta^{s+1} \partial_r + \frac{K^2 - 2is(r - M)K}{\Delta} + 4is\omega r - {}_s\lambda_{\ell m} \right] R_{\ell m}(r) &= 0 \end{aligned}$$

The solution to the inhomogeneous equation (with a source term represented by a contribution $J(r)S(\theta)$ at RHS) correspond to the homogeneous case for the angular dependence (the angular function is still the spheroidal harmonics), and to a inhomogeneous equation for the radial functions:

$$\left[\Delta^{-s} \partial_r \Delta^{s+1} \partial_r + \frac{K^2 - 2is(r - M)K}{\Delta} + 4is\omega r - {}_s\lambda_{\ell m} \right] R_{\ell m}(r) = J_{\ell m}(r)$$

Now the equation is in a form $[\tilde{D}_r + f(r)]R(r) = J(r)$, where $\tilde{D}_r = \Delta^{-s}\partial_r\Delta^{s+1}\partial_r$ is a differential operator containing two derivatives and $f(r)$ is a function of the radial coordinate only, including the eigenvalue. We want to recast this expression in a form that resembles a Schrodinger equation, namely

$$\frac{d^2 Y(r_*)}{dr_*^2} + V(r_*)Y(r_*) = J_Y(r_*)$$

expliciting all the dependences. Define $Y(r) \equiv \Delta^{s/2}\sqrt{r^2 + a^2}R(r)$ and make use of the tortoise radial coordinate r_* defined by $\frac{dr_*}{dr} = \frac{r^2 + a^2}{\Delta}$. The radial equation becomes

$$\frac{d^2}{dr_*^2}Y + V(\omega)Y = J_Y$$

With a source term $J_Y \equiv J \frac{\Delta}{(r^2 + a^2)^{\frac{3}{2}}}$. The potential is given by

$$V(\omega) = \frac{K^2(r, \omega) - \Delta\lambda_{\ell m}}{(r^2 + a^2)^2} - G^2(r) - \frac{dG(r)}{dr_*}$$

where $G(r) = \frac{s(r - M)}{r^2 + a^2} + \frac{r\Delta}{(r^2 + a^2)^2}$ and $K = (r^2 + a^2)\omega - ma$.

In order to compute the energy flux, we are interested in the asymptotic behavior of the field solution at the horizon and at spatial infinity, where it will define the stress-energy tensor encoding the fluxes. In the case of interest of a scalar field, $s = 0$, we observe that:

- At the horizon, $\Delta = 0$ (by definition of the horizon, or equivalently due to the value of the radius of the outer horizon $r_+ = M + \sqrt{M^2 - a^2}$), so $G = \partial_r G = 0$, $r_+^2 + a^2 = 2Mr_+$ and the potential V reduces to $k^2 \equiv \omega - m\Omega_+$ with $\Omega_+ \equiv \frac{a}{2Mr_+}$. Note that this is not valid for tensor ($s=2$) perturbations, as $G \rightarrow 0$.
- At $r \rightarrow \infty$, the asymptotic behaviors are $\Delta \sim r^2$, $K \sim r^2$, $G \sim r^{-1}$, thus the leading term is $K^2/r^4 \sim \omega^2$, so $V \sim \omega^2$.

This results in the asymptotic homogeneous equations:

$$\begin{aligned} \frac{d^2}{dr_*^2}Y + k^2Y &= 0 & \text{for } r \rightarrow r_+ \\ \frac{d^2}{dr_*^2}Y + \omega^2Y &= 0 & \text{for } r \rightarrow \infty \end{aligned}$$

The complete master equation can be solved by specifying two homogeneous solutions and building the total solution as a superposition using the Green's function method. The choice of the two independent homogeneous solutions can greatly simplify the treatise; in particular, since the fluxes are extracted in correspondence with the asymptotic limits here described, we select a purely outgoing solution at ∞ and a purely ingoing solution at the horizon. Since the solution to each equation is a sum of just two complex exponentials $e^{\pm ikr_*}$ and $e^{\pm i\omega r_*}$ respectively, a combination of these two purely oscillatory behaviors will entirely describe the field oscillations on the other side:

Purely ingoing at horizon:

Purely outgoing at infinity:

$$\begin{aligned} Y_- &= e^{-ikr_*} & \text{for } r \rightarrow r_+ & Y_+ = B_{\text{in}} e^{-ikr_*} + B_{\text{out}} e^{+ikr_*} & \text{for } r \rightarrow r_+ \\ Y_- &= A_{\text{in}} e^{-i\omega r_*} + A_{\text{out}} e^{+i\omega r_*} & \text{for } r \rightarrow \infty & Y_- = e^{+i\omega r_*} & \text{for } r \rightarrow \infty \end{aligned}$$

The effect of the field dynamics on the Kerr metric (represented by the effective potential $V(\omega)$) between horizon and infinity is uniquely encoded in the coefficients A_{in} , A_{out} , B_{in} , B_{out} . Note that we already implicitly choose a normalization of the two homogeneous solutions, which is completely arbitrary since there is no source term to compare. Once the homogeneous solutions are known on the whole range (for example, by integrating this oscillatory limit through the potential at finite values of the radial coordinate), they can be used to find the particular solution using the Green's function method.

After presenting the standard treatise for a massless scalar field let's adapt it to a massive field, having already clear the theoretical meaning of the used procedure.

3.2 Teukolsky equation for massive scalar fields

In this section, the formalism presented before is adapted to the massive field and specialized to a scalar. The master equation is not given anymore by the Teukolsky operator, since that is valid only for massless fields. In order to recover the correspondice master equation, I restart from the field equations, namely a Klein-Gordon with a source term:

$$(\square - \mu_s^2)\varphi = -4\pi T \quad (3.1)$$

where the Dalemberertian operator is

$$\square = \nabla_\mu \nabla^\mu = g^{\mu\nu} \nabla_\mu \nabla_\nu = \frac{1}{\sqrt{-g}} \partial_\mu \sqrt{-g} \partial^\mu \quad (3.2)$$

This expression can be specialized to the Kerr metric, for which $\sqrt{-g} = \Sigma \sin \theta$, obtaining

$$\square = \frac{1}{\Sigma \sin \theta} \partial_\mu \Sigma \sin \theta g^{\mu\nu} \partial_\nu \quad (3.3)$$

The standard Teukolsky operator $\mathcal{T}^{(0)}$ is obtained by expliciting this operator and multiplying by $-\Sigma$ (it is evident for example in the $\tilde{D}_\theta$ term); thus the massive Teukolsky operator is

$$\begin{aligned} -\Sigma(\square - \mu_s^2) &= \mathcal{T}^{(0)} + \Sigma \mu_s^2 \\ &= \mathcal{T}^{(0)} + r^2 \mu_s^2 + a^2 \mu_s^2 - a^2 \sin^2 \theta \mu_s^2 \\ &= \left[-\frac{(r^2 + a^2)^2}{\Delta} (\omega^2 - \mu_s^2) + \frac{4Mar}{\Delta} \omega m - \frac{2Mr}{\Delta} (r^2 + a^2) \mu_s^2 - \frac{m^2 a^2}{\Delta} - \partial_r \Delta \partial_r \right] + \\ &\quad + \left[a^2 (\omega^2 - \mu_s^2) \sin^2 \theta + \frac{m^2}{\sin^2 \theta} - \frac{1}{\sin \theta} \partial_\theta \sin \theta \partial_\theta \right] \end{aligned}$$

In the last passage I already separated the radial (angular) operators in the upper (lower) line. I also already used the usual Fourier-radial-angular decomposition, which enables

writing ∂_t and ∂_ϕ as $-i\omega$ and im respectively. As for the massless case, this equation can be separated into an angular and a radial part.

The decoupled homogeneous massive Teukolsky equations thus become:

$$\left[\frac{1}{\sin\theta} \partial_\theta \sin\theta \partial_\theta - a^2(\omega^2 - \mu_s^2) \sin^2\theta - \frac{m^2}{\sin^2\theta} + A \right] S_{\ell m}(\theta) = 0$$

$$\left[\partial_r \Delta \partial_r + \frac{(r^2 + a^2)^2(\omega^2 - \mu_s^2) - 4Mar\omega m + 2Mr(r^2 + a^2)\mu_s^2 + m^2 a^2}{\Delta} - A \right] R_{\ell m}(r) = 0$$

where A is the separation constant. I define as before the angular eigenvalue $\lambda = A - 2ma\omega$, obtaining the massive radial master equation

$$\Delta R'' + 2(r - M)R' + \left[\frac{K^2}{\Delta} - \mu_s^2(r^2 + a^2) - \lambda \right] R = 0 \quad (3.4)$$

Note that in this case it is not possible to write the potential as the square of a function K , since an additional term is introduced in the double product. Nevertheless, the correction is very minimal and simply corresponds to adding a linear term $-\mu_s^2 \rho^2 R$. Note that this is the massive scalar equivalent of Eq. 1.2 of Mano et al., 1996.

Introducing the Tortoise coordinate r_* and the auxiliary radial function $Y = \sqrt{r^2 + a^2} R = \rho R$, we obtain the Schrodinger-like equation for the radial function:

$$\frac{d^2 Y}{dr_*^2} + V(r)Y = J_Y \quad (3.5)$$

with the effective potential given by

$$V(r) = \left(\omega - \frac{am}{\rho^2} \right)^2 - \frac{\Delta}{\rho^8} \left[\lambda \rho^4 + 2Mr^3 + a^2(\Delta - 2Mr) + \mu_s^2 \rho^6 \right] \quad (3.6)$$

Some observations on the separated massive Teukolsky equation:

- **Angular equation:** adding a mass to the field has effects on the definition of the spheroidal harmonics, although it only consists in a simple reparametrization of the ω : the massive spheroidal harmonics are the massless functions evaluated in $\tilde{\omega} = \sqrt{\omega^2 - \mu_s^2}$, namely the “bare” energy stripped from the mass contribution. However, the eigenvalue λ reported in the effective potential still corresponds to $A - 2ma\omega$, thus it is equal to $\lambda_{\text{massless}}(\tilde{\omega}) + 2ma\tilde{\omega} - 2ma\omega$. Note that if $\mu_s \neq 0$, then $\tilde{\omega}$ becomes purely imaginary for sufficiently low energy. This is not problematic for the definition of the spheroidal harmonics and will become extremely relevant for massive fields. Note that even if $\tilde{\omega}$ is purely imaginary, the angular equation remains real, and thus also $S_{\ell m}^{(0)}$. This is true only for scalars, $s = 0$.
- **Radial equation:** the effect of a mass is to add a contribution $-\frac{\Delta}{\rho^2} \mu_s^2$ to the massless potential. This goes to $-\mu_s^2$ at $r \rightarrow \infty$ and to 0 for $r \rightarrow r_+$. The interpretation at ∞ is clear, since we recover the standard behavior of a massive field in a flat spacetime. The interpretation at the horizon is more interesting: the massive field *is behaving like a massless field* in the strong potential; morally, for a particle which at infinity has a finite energy from a mass contribution and a kinetic contribution, near the horizon

the mass is progressively less relevant with respect to the energy that is accumulated by the descent in the potential, becoming irrelevant exactly at the horizon; another way of seeing this is that in Kerr coordinates, as for the Schwarzschild coordinates, lightcones accumulate on the horizon becoming more and more squeezed as they approach it; this means that at the horizon the inwards and outwards light-like surfaces (populated by massless particles) coincide, and with them everything that was in between (massive particles): in this sense, at the horizon massless and massive particle behave in the same way. Of course, this serves only as an intuitive description of the phenomenon, but it will be useful to recall this particular feature in the following.

3.2.a Asymptotic behavior of the master equation

In order to compute the behavior of the field in the limits $r_* \rightarrow \pm\infty$, it would be tempting to simply evaluate the effective potential in the two limits and plug the result into the master equation. This, however, would be wrong, since we must explicitly consider the dependence of $V(r)$ on terms of order $\mathcal{O}(r_*^{-2\alpha})$ with $\alpha < 1$. This comes from the following argument: let's consider a master equation

$$\frac{d^2 Y}{dx^2} + V(x)Y = 0$$

with $V(x) \sim \varepsilon^2 x^{-2\alpha}$ in the limit (without loss of generality) $x \rightarrow \infty$. Then for $x \rightarrow \infty$ we can decompose the master equation in the form

$$\left(\frac{d}{dx} + i\varepsilon x^{-\alpha} \right) \left(\frac{d}{dx} - i\varepsilon x^{-\alpha} \right) Y = 0$$

The two linearly independent solutions to each of the two operators can be easily obtained by separating the variables, and are

$$\begin{aligned} Y &\propto \exp \left[\pm i\varepsilon \frac{x^{1-\alpha}}{1-\alpha} \right] && \text{if } \alpha \neq 1 \\ Y &\propto x^{\pm i\varepsilon} && \text{if } \alpha = 1 \end{aligned}$$

This solution is non-trivial (it does not reduce to a constant, which implies a null solution) only if $1 - \alpha \geq 0$, namely $\alpha \leq 1$.

Let's apply this to the effective potential: for $r \rightarrow r_h$ we have that ρ assumes a finite value, while $\Delta \rightarrow 0$ linearly in r (it has a finite derivative in r_h), and thus is exponentially suppressed in r_* . Since the exponential suppression is faster than any polynomial, this is a term that does not contribute to the asymptotic limit of Y^- . In conclusion, at the horizon we have

$$\frac{d^2 Y^-}{dr_*^2} + k_-^2 Y^- = 0, \quad \text{with} \quad k_- \equiv \omega - m\Omega_h$$

with purely ingoing asymptotic (normalized) solution

$$Y^-(r_* \rightarrow -\infty) = e^{-ik_- r_*}$$

Again, this can be interpreted with the fact that at the horizon, massive fields behave as massless. Now let's consider the limit $r_* \sim r \rightarrow +\infty$, expanding the effective potential in powers of $1/r$, keeping only constants and $\mathcal{O}(1/r)$ and writing it in the form used above:

$$\begin{aligned} V(r) &= \omega^2 - \frac{\Delta}{\rho^2} \mu_s^2 + \mathcal{O}\left(\frac{1}{r^2}\right) \\ &= \omega^2 - \mu_s^2 + 2\frac{M\mu_s^2}{r} + \mathcal{O}\left(\frac{1}{r^2}\right) \\ &= \left(\sqrt{\omega^2 - \mu_s^2} + \frac{1}{r} \frac{M\mu_s^2}{\sqrt{\omega^2 - \mu_s^2}} + \mathcal{O}\left(\frac{1}{r^2}\right) \right)^2 \end{aligned}$$

Thus, the asymptotic solution will be

$$Y^+(r_* \rightarrow +\infty) = \exp \left[i r_* \sqrt{\omega^2 - \mu_s^2} \right] r^{i \frac{\mu_s^2 M}{\sqrt{\omega^2 - \mu_s^2}}} = e^{i k_+ r_*} r^{i \frac{\mu_s^2 M}{\sqrt{\omega^2 - \mu_s^2}}}$$

with $k_+ = \sqrt{\omega^2 - \mu_s^2}$. Note that in this case the asymptotic behavior is qualitatively (and quantitatively) different from that expected by simply performing the limit of the effective potential for $r \rightarrow \infty$, which would give only the dependence $e^{i k_+ r_*}$.

3.2.b Green's function method for second order ODE

Given a second order Ordinary Differential Equation (ODE) in the form $y'' + p(x)y' + q(x)y = f(x)$, the solution is $y_h + y_p$, where y_h is the solution to the homogeneous problem and y_p is the particular solution. The homogeneous is given by the sum of two linearly independent solutions y_1 and y_2 , namely $y_h = a_1 y_1 + a_2 y_2$. The particular solution is given by

$$y_p = y_1(x)U_1(x) + y_2(x)U_2(x)$$

where the functions U_1 and U_2 are given by

$$U_1(x) = - \int_{x_0}^x \frac{y_2(z)f(z)}{W(z)} dz \quad U_2(x) = \int_{x_0}^x \frac{y_1(z)f(z)}{W(z)} dz$$

And W is the Wronskian, $W = y_1 y_2' - y_1' y_2$. Note that since W is the determinant of the matrix

$$\begin{bmatrix} y_1 & y_2 \\ y_1' & y_2' \end{bmatrix}$$

and y_1, y_2 are linearly independent, it is surely non-zero¹.

Note that this corresponds to taking the Greens function $G(x, \tau) = \frac{y_2(x)y_1(\tau) - y_1(x)y_2(\tau)}{W(\tau)}$.

In our case it is better to maintain this separate form, since we have boundary conditions at different spatial coordinate for the 2 homogeneous solutions ($r_{*0} \rightarrow \pm\infty$ respectively).

Let's apply the general Green's function method to the specific case of the scalar field's radial master equation: let $Y^\pm(r_*)$ be the homogeneous solutions with purely ingoing/out-

¹In reality, this condition is violated by quasi-normal modes (see Chapter 6), for which however the zero of the Wronskian is located in the complex plane of the variable x away from the real axis, thus not introducing domain incompleteness when evaluated in $x \in \mathbb{R}$.

going boundary conditions at horizon/infinity; then the general solution will be

$$Y = Y^+ \int_{-\infty}^{r_*} \frac{Y^- J_Y}{W} dr_* + Y^- \int_{r_*}^{+\infty} \frac{Y^+ J_Y}{W} dr_* \quad (3.7)$$

Note that the minus sing present in the general form is reabsorbed by exchanging the integration boundaries when integrating R^+ . This also gives a more natural result in terms of the integrating direction. In addition, we are always working with the tortoise coordinate, since the original ODE is expressed using r_* .

The application of the method requires the knowledge of the homogeneous solutions. We obtain the homogeneous using a Frobenius expansion ansatz, exploiting the asymptotic dependences that we found before:

$$Y^- = e^{-ik_- r_*(r)} \sum_{n=0}^{n_h} a_n (r - r_h)^n \quad Y^+ = e^{+ik_+ r_*(r)} r^{\frac{i\mu_s^2 M}{\sqrt{\omega^2 - \mu_s^2}}} \sum_{n=0}^{n_\infty} \frac{b_n}{r^n}$$

In fact, choosing $a_0 = b_0 = 1$ correctly gives the correct asymptotic solutions. In practice, the sum is truncated at $n_h^{\max} = n_\infty^{\max} = 10$, which gives an accuracy of ≈ 1 part on 10^{30} . The coefficients a_n, b_n are found by solving the master equation iteratively with the Frobenius ansatz. Refer to Sec. 4.4 for the implementation and details.

3.3 Energy and angular momentum scalar fluxes

At this point, we have the solution φ to the field equations and we can compute the energy emission. This can be done considering the stress-energy tensor of the scalar field: the component T_{tr} encodes the flux of energy (selected by the time component) along the radial coordinate. Thus, it is sufficient to integrate this element over a surface at ∞ (on the horizon) to obtain the total outgoing (ingoing) energy flux. Since the stress-energy tensor component we are considering is out of the diagonal, the matrix element simply corresponds to (cf. Eq. (2.3))

$$T_{rt} = \frac{1}{16\pi} \partial_r \varphi \partial_t \varphi^*$$

so that the total energy flux is

$$\dot{E}^\pm = \pm \Delta \int_{\Omega} T_{rt} d\Omega$$

The time derivative of the field is simply $\partial_t \varphi = -i\omega \varphi$, while the radial derivative can be computed directly from the asymptotic limits, giving $\partial_r \varphi = \pm ik_\pm \varphi$ depending on the evaluation surface (horizon or infinity). The angular-radial decomposition ensures that the angular integral can be separated and integrated to unity by definition of the spheroidal eigenvalues. After some calculation (remember that we are using the Y radial function, thus in the derivatives some additional terms are present; however, the prefactor Δ at the horizon and the asymptotic behavior at infinity allow us to simplify the limits) the energy fluxes are:

$$\dot{E}^\pm = \frac{1}{16\pi} \sum_{\ell, m} \omega k_\pm |Y_{\ell m}(r_* \rightarrow \pm\infty)|^2$$

We define the amplitudes

$$Z_{\ell m}^{\pm} \equiv Y_{\ell m}^{\pm}(r_* \rightarrow \pm\infty) \int_{-\infty}^{+\infty} \frac{Y_{\ell m}^{\mp} J_{Y, \ell m}}{W} dr_*$$

which correspond to the asymptotic limit of the general solution to the field equation. Thus, the energy fluxes are simply

$$\dot{L}^{\pm} = \frac{1}{16\pi} \sum_{\ell, m} \omega k_{\pm} |Z_{\ell m}^{\pm}|^2 \quad (3.8)$$

Similarly for the angular momentum flux, which is obtained by integrating $T_{\phi t}$ on a suitable spherical surface and is given by

$$\dot{E}^{\pm} = \frac{1}{16\pi} \sum_{\ell, m} m k_{\pm} |Z_{\ell m}^{\pm}|^2 = \sum_{\ell, m} \frac{m}{\omega} \dot{E}_{\ell m} \quad (3.9)$$

thus it is possible to obtain, mode-by-mode, the flux of angular momentum as a function of the energy flux. Note that for circular equatorial orbits $\omega = m\Omega_p$, thus the prefactor is factorized out of the sum and the energy and angular momentum fluxes are linearly dependent: morally, the circular equatorial orbital evolution is controlled by only the energy flux.

3.3.a Source term computation

The Teukolsky equation has a source term $4\pi\Sigma T^{(0)}$, where $T^{(0)}$ is the source in the Klein Gordon equation. Thus

$$\begin{aligned} 4\pi\Sigma T^{(0)} &= -4\pi\Sigma dm_p \int \frac{\delta^{(4)}(x - y_p(\tau))}{\sqrt{-g}} d\tau \\ &= -4\pi\Sigma dm_p \int \frac{\delta(t - t_p(\tau))\delta(r - r_p(\tau))\delta(\theta - \theta_p(\tau))\delta(\phi - \phi_p(\tau))}{\Sigma \sin \theta} \frac{d\tau}{dt} dt \\ &= -4\pi dm_p \frac{\delta(r - r_p(t))\delta(\theta - \theta_p(t))\delta(\phi - \phi_p(t))}{\dot{t} \sin \theta} \end{aligned}$$

where I used $\sqrt{-g} = \Sigma \sin \theta$, and the conversion between the proper time τ and the coordinate time t to apply the Dirac δ along the particle trajectory. The term J_Y used in the radial equation corresponds to the radial part of the angular-radial decomposition of the whole source term; to obtain J , I must project the whole term on $S_{\ell m}(\theta)$ and take the anti-Fourier transform (we are working in frequency domain, not time):

$$\begin{aligned} J_{\ell m} &= \frac{1}{2\pi} \int \int_{\Omega} 4\pi\Sigma T^{(0)} S_{\ell m}^*(\theta, \omega) e^{-im\phi} e^{i\omega t} d\Omega dt \\ &= -\frac{1}{2\pi} \int \int_{\Omega} 4\pi dm_p \frac{\delta(r - r_p(t))\delta(\theta - \theta_p(t))\delta(\phi - \phi_p(t))}{\dot{t} \sin \theta} S_{\ell m}^*(\theta, \omega) e^{-im\phi} e^{i\omega t} \sin \theta d\theta d\phi dt \\ &= -\frac{1}{2\pi} \int \int_{(\phi)} 4\pi dm_p \frac{\delta(r - r_p(t))\delta(\phi - \phi_p(t))}{\dot{t}} S_{\ell m}^*(\theta_p(t), \tilde{\omega}) e^{-im\phi} e^{i\omega t} d\phi dt \\ &= -4\pi dm_p \frac{1}{2\pi} \int \frac{\delta(r - r_p(t))}{\dot{t}} S_{\ell m}^*(\theta_p(t), \omega) e^{i[\omega t - m\phi_p(t)]} dt \end{aligned}$$

The projection on the angular part is obtained through the usual dyadic product, namely integrating the function times the complex conjugate of $S(\theta)e^{im\phi}$ over $d\Omega$. Recall that we separated the simple dependence on ϕ with the exponential, which must be included (and conjugated) in the true entire spheroidal harmonic function.

Note that this is the maximal simplification that can be done to the general expression of $J_{\ell m}$: all the quantities under integration are in general functions of the time coordinate. The general solution to this problem is obtained by using the explicit behavior of the Kerr geodesics; this is described in Gliorio et al. (2026), and involves a further Fourier series decomposition of the frequencies, resulting in a summation over the additional indexes n, k of the frequency spectrum given by the energies $\omega_{mkn} = m\Omega_\phi + n\Omega_r + k\Omega_\theta$, where the Ω_i are the angular velocities associated to the angle variables of the action-angle variable decomposition of integrable Hamiltonian systems (see Sec. 1.4.a).

Specialization to equatorial circular orbits

For equatorial circular orbits, both $r = r_p$, $\dot{t}$ and $\theta_p(t) = \pi/2$ do not depend on the time coordinate. The angular coordinate ϕ is linear in time: $\phi(t) = \Omega_p t$, such that

$$\begin{aligned} J_{\ell m} &= -4\pi d m_p \frac{\delta(r - r_p) S_{\ell m}^*(\pi/2, \omega)}{\dot{t}} \frac{1}{2\pi} \int e^{i(\omega - m\Omega_p)t} dt \\ &= -4\pi d m_p \frac{\delta(r - r_p) S_{\ell m}^*(\pi/2, \omega)}{\dot{t}} \delta(\omega - m\Omega_p) \\ &= -4\pi d m_p \frac{\delta(r - r_p) S_{\ell m}^*(\pi/2, m\Omega_p)}{\dot{t}} \delta(\omega - m\Omega_p) \end{aligned}$$

Note that in the Schrodinger-like equation we use the modified source term $J_Y \equiv J \frac{\Delta}{(r^2 + a^2)^{\frac{3}{2}}}$

Finding the homogeneous solution

In the previous Chapter we saw that the scalar field solution can be found using the Green's function method, once the homogeneous solutions of the massive Teukolsky equation are known; in addition, we made the choice for these oscillatory homogeneous solutions to be purely ingoing at the outer horizon and purely outgoing at spatial infinity, in order to compute more easily the total energy and angular momentum fluxes. This Chapter is dedicated to the methods that are implemented to find the homogeneous solutions.

A solution of a second-order ODE such as the Teukolsky Equation can be found numerically, once initial conditions are given at the boundaries; very interestingly however, semi-analytical solutions exist, such as the ones provided by Mano et al. (1996). These solutions are very valuable for finding accurate boundary conditions, from which is then more convenient to numerically integrate in a finite radial interval. The code STORM, later used for the numerical evaluation of massive scalar fluxes, already implements this method for massless fields. In Sec. 4.1 I will explain and comment the original Mano-Suzuki-Takasugi (MST) method for a massless scalar field; In Sec. 4.2 I will present original results on the generalization of the MST method to a massive scalar field; in Sec. 4.4, a complete different approach is discussed, in terms of a Frobenius series expansion at the horizon and at infinity. Finally, in Sec. 4.5 I will explain the integration scheme used to compute the homogeneous solution on a finite spatial (radial + angular) region.

4.1 MST method for a massless scalar field

The MST method (Mano et al., 1996) provides an analytical solution to the massless Teukolsky equation in terms of a sum of hypergeometric functions. We saw the radial homogeneous Teukolsky equation (without the tortoise radial variable redefinition):

$$\Delta \partial_r^2 R + 2(r - M)(s + 1) \partial_r R + \left[\frac{K^2 - 2is(r - M)K}{\Delta} + 4is\omega r - \lambda \right] R = 0 \quad (4.1)$$

4.1.a The hypergeometric function

Definition: the hypergeometric function ${}_2F_1(a, b; c; z)$, for brevity $F(a, b; c; z)$, is defined by the power series:

$$F(a, b; c; z) \equiv \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!} \quad (4.2)$$

where the Pochhammer symbol $(a)_n$ is defined as $(a)_0 = 1$, and for $n > 0$

$$(a)_n \equiv \frac{(a+n-1)!}{(a-1)!} = a(a+1) \dots (a+n-1) \quad (4.3)$$

Warning: the definition holds for $|z| < 1$, but F can be analytically continued (with some care, elaborate later if needed).

Lemma: The hypergeometric function $F(a, b; c; z)$ is the solution to the following second-order differential equation:

$$z(1-z) \frac{d^2 F}{dz^2} + [c - (a+b+1)z] \frac{dF}{dz} - abF = 0 \quad (4.4)$$

Lemma: The derivatives of the hypergeometric function are

$$\frac{d^n}{dz^n} F(a, b; c; z) = \frac{(a)_n (b)_n}{(c)_n} F(a+n, b+n; c+n; z)$$

Observation: even if we can use more general theorems about regular points, it is evident that the structure of the radial Teukolsky equation, Eq. (4.1), closely resembles the structure of the Euler hypergeometric differential equation, Eq. (4.4), in particular regarding the second-power (inhomogeneous) coefficient of the second-order term and the first-power (inhomogeneous) coefficient of first-order term. We want to make some variable and radial function redefinition in order to express the result as a sum of hypergeometric functions. Note that the zeroth-order term of Eq. (4.1) is not a constant, as the ab term featured in Eq. (4.4); this means that the solution will correspond to the Euler's equation with a non-zero RHS featuring both the variable and the radial function. The trick is that the RHS will be of order ϵ , enabling for a recursive solution. In particular, the solution will be obtained as an expansion in the parameter $\epsilon = 2M\omega$ (low-energy solution, for series coefficients with exponent $n > 0$).

Defined the auxiliary quantities

$$\epsilon = 2M\omega, \quad q = a/M, \quad \kappa = \sqrt{1-q^2}, \quad \tau = (\epsilon - mq)/\kappa,$$

I make the following redefinitions of the working variable/functions:

$$\begin{aligned} x &\equiv \omega(r_+ - r)/\epsilon\kappa \\ R_{\text{in}}^\nu &= e^{i\epsilon\kappa x} (-x)^{-s-\frac{i}{2}(\epsilon+\tau)} (1-x)^{\frac{i}{2}(\epsilon-\tau)} p_{\text{in}}^\nu(\tau) \\ R_{\text{out}}^\nu &= e^{i\epsilon\kappa x} (-x)^{\frac{i}{2}(\epsilon+\tau)} (1-x)^{-s-\frac{i}{2}(\epsilon-\tau)} p_{\text{out}}^\nu(\tau) \end{aligned}$$

Let's consider the purely ingoing solution at the horizon. The radial equation becomes

$$\begin{aligned} x(1-x) \frac{d^2 p_{\text{in}}^\nu}{dx^2} + \left[1 - s - i\epsilon - i\tau - (2 - 2i\tau)x \right] \frac{dp_{\text{in}}^\nu}{dx} + (\nu + i\tau)(\nu + 1 - i\tau) p_{\text{in}}^\nu = \\ = 2i\epsilon\kappa \left[-x(1-x) \frac{dp_{\text{in}}^\nu}{dx} + (1 - s + i\epsilon - i\tau)x p_{\text{in}}^\nu \right] + \\ + \left[-\lambda - s(s+1) + \nu(\nu+1) + \epsilon^2 - i\epsilon\kappa(1-2s) \right] p_{\text{in}}^\nu \end{aligned}$$

- **Role of ν :** note that the terms in the parameter ν exactly vanish at LHS and RHS; they are added with the aim of making the LHS resemble Euler's equation. For this reason, this is in principle a free parameter, which however will be later fixed in order to allow a recursive solution.
- **The RHS is $\mathcal{O}(\epsilon)$:** we want an ϵ -expansion of the solution, which at 0 order is the hypergeometric function. For this reason, we require the RHS to be $\mathcal{O}(\epsilon)$: we can use the fact that ν is a free parameter and that the spheroidal eigenvalue has an explicit expression as $\lambda = E - s(s+1) - m \frac{a}{M} \epsilon + \left(\frac{a}{M} \epsilon\right)^2$, with

$$E = \ell(\ell+1) - \frac{s^2 m}{\ell(\ell+1)} \frac{a}{M} \epsilon + \left[H(\ell+1) + H(\ell) - 1 \right] \left(\frac{a}{M} \epsilon\right)^2 + \mathcal{O}(\epsilon^3)$$

where $H(\ell) = \frac{2(\ell^2 - m^2)(\ell^2 - s^2)^2}{(2\ell - 1)\ell^3(2\ell + 1)}$ so that $\lambda = \ell(\ell+1) - s(s+1) + \mathcal{O}(\epsilon)$. This means that if $\nu = \ell + \mathcal{O}(\epsilon)$, the RHS is globally $\mathcal{O}(\epsilon)$. This constraint, however, still leaves the degree of freedom of choosing ν from first order on.

- **Homogeneous solved by hypergeometric function:** the homogeneous solution is the hypergeometric function with $a = -\nu - i\tau$, $b = \nu + 1 - i\tau$, $c = 1 - s - i\epsilon - i\tau$, namely $p_\nu(x) \equiv F(\nu + 1 - i\tau, -\nu - i\tau; 1 - s - i\epsilon - i\tau; x)$.
- **General solution given by sum of hypergeometric functions:** we write the solution as $p_{\text{in}}^\nu(x) = \sum_{n=-\infty}^{\infty} a_n^\nu p_{n+\nu}(x)$, where $p_{n+\nu}$ is simply p_ν with the substitution $\nu \mapsto \nu + n$. This is an ansatz that will finally prove to give a convergent solution.

In addition, we can write the terms $x p_{n+\nu}$ and $x(1-x)p'_{n+\nu}$ present at RHS as linear combinations of $p_{n+\nu-1}$, $p_{n+\nu}$ and $p_{n+\nu+1}$. By plugging the series expansion ansatz into the radial equation we can solve the coefficients a_n order-by-order in n . In particular, the coefficients must obey a three-term recurrence relation

$$\alpha_n^\nu a_{n+1}^\nu + \beta_n^\nu a_n^\nu + \gamma_n^\nu a_{n-1}^\nu = 0 \quad (4.5)$$

with $\alpha_n, \beta_n, \gamma_n$ constants dependent on the usual set of parameters and the index n . The three-term recurrence can be rewritten in terms of continued fraction: we define the right- and left-continued fractions as

$$R_n(\nu) = \frac{a_n^\nu}{a_{n-1}^\nu} = -\frac{\gamma_n^\nu}{\beta_n^\nu + \alpha_n^\nu R_{n+1}(\nu)} \quad L_n(\nu) = \frac{a_n^\nu}{a_{n-1}^\nu} = -\frac{\alpha_n^\nu}{\beta_n^\nu + \gamma_n^\nu L_{n-1}(\nu)} \quad (4.6)$$

The first expression explains the “right” and “left” denominations, while the second explains why they are called continued fractions. Either L or R allow to build a ladder of coefficients; in order for these to be consistent, we ask $R_n(\nu)L_{n-1}(\nu) = 1$, namely one step to the right and one to the left is equivalent to the identity. This fixes the value of the renormalized angular momentum ν .

4.2 MST method for the massive scalar field

In this section, I will generalize the MST method by closely following the approach used in Mano et al., 1996. I previously found the radial master equation Eq. (3.4), which differs

from the massless one for the term proportional to μ_s^2 . With the formalism outlined in App. B of Mano et al., 1996, where $y \equiv \omega r$ and thus $y_{\pm} = \frac{\epsilon}{2}(1 \pm \kappa)$, this corresponds to the following linear term in R (the ω^2 comes from the variable change $r \mapsto y = \omega r$):

$$\begin{aligned} -\frac{\mu_s^2 \rho^2}{\Delta \omega^2} &= -\frac{\mu_s^2}{\omega^2} \frac{(r^2 + a^2)\omega^2}{(y - y_+)(y - y_-)} \\ &= -\frac{\mu_s^2}{\omega^2} \left[1 + \frac{\epsilon}{y - y_{\mp}} + \frac{\epsilon}{(y - y_+)(y - y_-)} \frac{\epsilon}{2}(1 \pm \kappa) \right] \\ &= -\frac{\mu_s^2}{\omega^2} \left[1 + \frac{\epsilon/2}{y - y_+} + \frac{\epsilon/2}{y - y_-} + \frac{1}{(y - y_+)(y - y_-)} \frac{\epsilon^2}{4}(1 + \kappa) + \frac{1}{(y - y_+)(y - y_-)} \frac{\epsilon^2}{4}(1 - \kappa) \right] \\ &= -\frac{\mu_s^2}{\omega^2} \left[1 + \frac{\epsilon/2}{y - y_+} + \frac{\epsilon/2}{y - y_-} + \frac{1}{(y - y_+)(y - y_-)} \frac{\epsilon^2}{2} \right] \end{aligned}$$

Note that in the third line I choose to maintain symmetry by summing half of each possibility in line 2 (there is the choice of the $\pm$). This allows the simplification in line 4, in which the κ dependence vanishes.

Before continuing, let's define an additional adimensional parameter $\eta \equiv \mu_s/\omega$. This is needed if the additional independent dimensional parameter μ_s is introduced, and the choice of this adimensionalization seems quite straightforward. The other possible option, the commonly used adimensional parameter $\mu_s M$ which compares the typical range of the scalar with the gravitational radius of the primary, produces more complicated factors that ultimately lead to a more natural definition of the same parameter η . The complete radial equation becomes (equivalent to Eq. B1 of Mano et al., 1996):

$$\begin{aligned} \frac{d^2 R}{dy^2} + \left(\frac{1}{y - y_+} + \frac{1}{y - y_-} \right) \frac{dR}{dy} + \\ + \left[1 - \eta^2 + \frac{1}{y - y_+} \epsilon \left(1 - \frac{\eta^2}{2} + \frac{1}{\kappa} \right) + \frac{1}{y - y_-} \epsilon \left(1 - \frac{\eta^2}{2} - \frac{1}{\kappa} \right) + \right. \\ + \frac{1}{(y - y_+)^2} \frac{(\epsilon + \tau)^2}{4} + \frac{1}{(y - y_-)^2} \frac{(\epsilon - \tau)^2}{4} + \\ \left. + \frac{1}{(y - y_+)(y - y_-)} \left(-\lambda + \frac{\epsilon^2}{2}(1 - \eta^2) - \frac{\tau^2}{2} - \epsilon m q \right) \right] R = 0 \end{aligned}$$

Note that this contribution to the potential does not include terms in $(y - y_{\pm})^{-2}$, namely those in the third line; this is very important, since it implies that in the redefinition of the radial variable via $y - y_+ = \epsilon \kappa(-x)$ and $y - y_- = \epsilon \kappa(1 - x)$, and the field $R = (-x)^{\alpha}(1 - x)^{\beta} \tilde{R}$, in order to eliminate the coefficients of $1/x^2$ and $1/(1 - x)^2$, the exponents α and β will take the same values as the massless case:

$$\alpha_{\pm} = \pm \frac{i}{2}(\epsilon + \tau) \qquad \beta_{\pm} = \pm \frac{i}{2}(\epsilon - \tau)$$

These exponents are easily found algebraically; since the differential equation is at sec-

ond order, the solutions for α, β will be (for each) a pair of solutions of a corresponding quadratic equation. The $\pm$ choice for each is a "true choice", namely the corresponding solutions will be different (obviously) and linearly independent when taken at pairs. Note that the choice of α determines the asymptotic behavior of the solution at the outer horizon, corresponding to purely ingoing/outgoing waves.

Another redefinition is applied, namely $\tilde{R} = e^{i\epsilon\kappa x}p$, which leads to the radial equation presented in the previous section. Note that this corresponds to isolating an amplitude p from an oscillating part $e^{i\epsilon\kappa x} = e^{-i\omega(r-r_+)}$; However, this is not the complete oscillating asymptotic behavior that we expect at the outer horizon, since the frame-dragging term involves the index m through some sort of k_- , which is not present at exponential, having instead simply ω .

Let's now compute the additional terms in the master equation (the one that at 0th order in ϵ has the hypergeometric function as solution): in the massless case, this equation is obtained by the substitution of the new function p and the variable x , and finally multiplying by $(-x)^{1-\alpha}(1-x)^{1-\beta}e^{-i\epsilon\kappa x}\epsilon^2\kappa^2$ in order to have the simple coefficient $(-x)(1-x)$ in front of the second derivative. Thus it is sufficient to compute the contributions from the 4 additional terms in the previous equation, and then multiply by this additional coefficient. The terms are:

- Zero-order term in the radial variable:

$$-\eta^2 R = -\eta^2 (-x)^\alpha (1-x)^\beta e^{i\epsilon\kappa x} p \quad \mapsto \quad -\eta^2 (-x)(1-x)\epsilon^2\kappa^2 p \quad (4.7)$$

This is a second-order term in the expansion parameter ϵ , and is proportional to $(-x)(1-x)p$; thus, it will need an additional recurrence relation, which is not provided in Mano et al., 1996.

- first-order terms in $(y - y_\pm)^{-1}$:

$$-\frac{\eta^2\epsilon}{2} \frac{1}{y - y_+} R = -\frac{\eta^2\epsilon}{2\epsilon\kappa} (-x)^{\alpha-1} (1-x)^\beta e^{i\epsilon\kappa x} p \quad \mapsto \quad -\frac{\eta^2}{2} (1-x)\epsilon^2\kappa p \quad (4.8)$$

$$-\frac{\eta^2\epsilon}{2} \frac{1}{y - y_-} R = -\frac{\eta^2\epsilon}{2\epsilon\kappa} (-x)^\alpha (1-x)^{\beta-1} e^{i\epsilon\kappa x} p \quad \mapsto \quad -\frac{\eta^2}{2} (-x)\epsilon^2\kappa p \quad (4.9)$$

Also, these terms are second-order in ϵ . For these, the recurrence relations are already known.

- second order term in $[(y - y_+)(y - y_-)]^{-1}$:

$$-\frac{\eta^2\epsilon^2}{2} \frac{1}{(y - y_+)(y - y_-)} R = -\frac{\eta^2\epsilon^2}{2\epsilon^2\kappa^2} (-x)^{\alpha-1} (1-x)^{\beta-1} e^{i\epsilon\kappa x} p \quad \mapsto \quad \frac{-\eta^2}{2} \epsilon^2 p \quad (4.10)$$

And also this term is $\propto \epsilon^2$, and linear in p .

Leaving aside the first term for now, the other three can be summed and separated again in algebraically distinct contributions, giving:

- $+\eta^2\epsilon^2\kappa xp$: the recurrence relation for xp is known, and this term can be integrated in the master equation as $2i\epsilon\kappa(-i\epsilon\eta^2/2)xp$ (the term in front is gathered to resemble the corresponding term in the original master equation).
- $-\frac{\eta^2}{2}\epsilon^2(1+\kappa)p$: this term is linear in p and thus does not require a recurrence relation. It can be included in the master equation as it is.

The master equation RHS thus becomes

$$2i\epsilon\kappa \left[(-x)(1-x)p'_{\text{in}} + i\epsilon\kappa \frac{\eta^2}{2} (-x)(1-x)p + \left(1 + i\epsilon \left(1 - \frac{\eta^2}{2} \right) - i\tau \right) x p_{\text{in}} \right] + \\ + \left[-\lambda + \nu(\nu+1) + \epsilon^2 - i\epsilon\kappa - \frac{\eta^2}{2} \epsilon^2 (1+\kappa) \right] p$$

4.2.a Recurrence relation for the term $(-x)(1-x)p$

Since we want to express every term as a linear combination of $p_{\nu+n+1}$, $p_{\nu+n-1}$, $p_{\nu+n}$, we have to find a recurrence relation for the term $(-x)(1-x)p_{\nu+n}$. Note that with the standard notation, if $p_{\nu+n} = F(a, b; c; x)$, then $p_{\nu+n+1} = F(a+1, b-1; c; x)$ and $p_{\nu+n-1} = F(a-1, b+1; c; x)$. This means that we have to find some coefficients A, B, C , algebraic functions of a, b, c as previously defined as functions of the dimensionless parameters, such that

$$(-x)(1-x)F(a, b; c; x) = AF(a+1, b-1; c; x) + BF(a, b; c; x) + CF(a-1, b+1; c; x) \quad (4.11)$$

In order to do so, we have to express the RHS in terms of only 2 hypergeometric functions (this is always possible and it is the best one can do, see next paragraph), and impose the equality order-by-order in x for the terms proportional to $F(a, b; c; x)$ and that the additional contiguous function vanishes. This is a total of 4 constraints with only 3 DoF, so a solution may not exist. We can also try adding two terms $DF'(a, b; c; x)$, $ExF'(a, b; c; x)$ (but not terms in $(-x)(1-x)F'$ or xF , since for these the contiguous relations exist and thus would be already accounted for in the form above), since these can be reabsorbed in the LHS, only contributing to a redefinition of the a, b, c parameters of the 0-order equation.

In the next paragraph, the recurrence relations for $F(a\pm, b\mp)$ and $F(a+, b+; c+) = c/(ab)F'(a, b; c)$ are computed.

Contiguous relations Two hypergeometric functions are said to be contiguous if their parameters a, b, c and a', b', c' differ by a triplet of integers l, m, k . Gauss proved that any hypergeometric function $F(a, b; c; x)$ can be written as a linear combination of any pair of contiguous hypergeometric functions $F(a+l, b+m; c+k; x)$, $F(a+l', b+m'; c+k'; x)$ through coefficients P and Q dependent on $a, b, c, l, m, k, l', m', k', x$. This linear combination is called *contiguous relation*. Any contiguous relation can be obtained by combining

the following elementary relations (*Gauss contiguous relations*, Gcr):

$$(c-a)F(a-1, b; c; x) + (2a-c+(b-a)x)F(a, b; c; x) + a(x-1)F(a+1, b; c; x) = 0 \quad (\text{Gcr1})$$

$$(b-a)F(a, b; c; x) + aF(a+1, b; c; x) - bF(a, b+1; c; x) = 0 \quad (\text{Gcr2})$$

$$(c-a-b)F(a, b; c; x) + a(1-x)F(a+1, b; c; x) - (c-b)F(a, b-1; c; x) = 0 \quad (\text{Gcr3})$$

$$c(a+(b-c)x)F(a, b; c; x) - ac(1-x)F(a+1, b; c; x) + (c-a)(c-b)xF(a, b; c+1; x) = 0 \quad (\text{Gcr4})$$

$$(c-a-1)F(a, b; c; x) + aF(a+1, b; c; x) - (c-1)F(a, b; c-1; x) = 0 \quad (\text{Gcr5})$$

$$c(1-x)F(a, b; c; x) - cF(a-1, b; c; x) + (c-b)xF(a, b; c+1; x) = 0 \quad (\text{Gcr6})$$

$$cF(a, b; c; x) - cF(a-1, b; c; x) - bxF(a, b+1; c+1; x) = 0 \quad (\text{Gcr7})$$

$$(a-1+(b+1-c)x)F(a, b; c; x) + (c-a)F(a-1, b; c; x) - (c-1)(1-x)F(a, b; c-1; x) = 0 \quad (\text{Gcr8})$$

$$c(c-1)(x-1)F(a, b; c-1; x) + c(c-1-(2c-a-b-1)x)F(a, b; c; x) + (c-a)(c-b)xF(a, b; c+1; x) = 0 \quad (\text{Gcr9})$$

and their *dual*, namely the same equation with a and b exchanged. Note that some equations are the dual of themselves (e.g. [Gcr2](#)). With a intuitive notation, the contiguous relations can be found as:

- Contiguous relation for $F(a\pm, b\mp)$:

$$(1^\dagger + 3^\dagger \rightsquigarrow 10)(10 + 2(b-) \rightsquigarrow 11)(11 + 2^\dagger(a-) \rightsquigarrow \star)$$

$$\left[(b-a)(1-x) + \frac{(c-b)(b-1)}{b-1-a} - \frac{(c-a)(a-1)}{a-1-b} \right] F(a, b; c) - \frac{(c-b)a}{b-1-a} F(a+, b-; c) - \frac{(c-a)b}{a-1-b} F(a-, b+; c) = 0$$

- Contiguous relation for $F(a+, b+, c+)$:

$$(2(b-) + 3 \rightsquigarrow 12)(12 + 7(a+) \rightsquigarrow \star)$$

$$[b-c+(a-b+1)x]F(a, b; c) + \frac{b}{c}(a-b+1)(-x)(1-x)F(a+1, b+1; c+1) + (c-b)F(a+1, b-1; c) = 0$$

Note that the Gcr for $F(a\pm, b\mp)$ has only terms at first order in x in the coefficient of F ; this means that Eq. (4.11) cannot be satisfied. It does not help adding terms in F' or xF' , since the Gcr for $F(a+, b+, c+)$ would involve a term $x(1-x)$ at the denominator. Note that the second Gcr, properly symmetrized by using the respective dual, corresponds to the second recurrence relation provided by Mano et al., 1996.

Conclusion: the massless reparametrization of the radial function would need a contiguous relation that **does not exist** in order to obtain a recurrence relation in the expansion of hypergeometric functions.

Second approach: redefinition of radial function

In the previous approach, I maintained the transformation $R = (-x)^\alpha(1-x)^\beta e^{i\epsilon\kappa x} p(x)$, obtaining a problematic term $-\eta^2(-x)(1-x)\epsilon^2\kappa^2 p$ in the final equation, which cannot be expressed through a recurrence relation. Thus, I want to change this redefinition of the function in order to get rid of this term.

Note that choosing $R = (-x)^\alpha(1-x)^\beta e^{i\epsilon\kappa\sqrt{1-\eta^2}x} p(x)$ produces exactly this term among those of d^2R/dx^2 , with opposite sign. It can be shown that this is also the only possible option. The complete calculation of the master equation as a function of the radial variable x and the new function p (remembering the mass contribution inside the potential) leads to:

$$\begin{aligned} x(1-x)p'' + \left[(2\alpha+1) - 2(\alpha+\beta+1)x \right] p' - abp = \\ = 2i\epsilon\kappa\sqrt{1-\eta^2} \left[(-x)(1-x)p' + \left(\alpha+\beta+1 + i\epsilon \left(1 - \frac{\eta^2}{2\sqrt{1-\eta^2}} \right) \right) xp \right] + \\ + \left[-\lambda + 2\alpha\beta + (\alpha+\beta) - i\epsilon\kappa\sqrt{1-\eta^2}(2\alpha+1) + \kappa\epsilon^2 \left(1 - \frac{\eta^2}{2} \right) + \right. \\ \left. + \epsilon^2 \left(\frac{3}{2} - \frac{\eta^2}{2} \right) - \frac{1}{2}\tau^2 - \epsilon m q - ab \right] p \end{aligned}$$

The coefficients a, b (which are added both at RHS and LHS, thus they are, in principle, free parameters) are chosen in such a way that $a+b = 2(\alpha+\beta)+1$, in order to form the Euler hypergeometric equation at LHS. Note that the choice of α, β and thus a, b is exactly the same of the massless case! So choose $\alpha_- = -\frac{i}{2}(\epsilon+\tau)$ to obtain an ingoing solution, $\beta_+ = \frac{i}{2}(\epsilon-\tau)$, $a = -\nu - i\tau$, $b = \nu + 1 - i\tau$, obtaining

$$\begin{aligned} x(1-x)p'' + \left[(1-i\epsilon-i\tau) - 2(1-i\tau)x \right] p' - (\nu+i\tau)(\nu+1-i\tau)p = \\ = 2i\epsilon\kappa\sqrt{1-\eta^2} \left[(-x)(1-x)p' + \left(1-i\tau + i\epsilon \left(1 - \frac{\eta^2}{2\sqrt{1-\eta^2}} \right) \right) xp \right] + \\ + \left[-\lambda + \nu(\nu+1) + \epsilon^2 \left(1 - \frac{\eta^2(1+\kappa)}{2} \right) - i\epsilon\kappa\sqrt{1-\eta^2} + \epsilon\kappa(\epsilon+\tau) \left(1 - \sqrt{1-\eta^2} \right) \right] p \end{aligned}$$

Now we make the ansatz $p_{\text{in}}^\nu = \sum_{n=-\infty}^{+\infty} a_n p_{n+\nu}(x)$ with $p_{n+\nu} = F(a+n, b-n; c; x)$, $a = \nu+1-i\tau$, $b = -\nu-i\tau$, $c = 1-i\epsilon-i\tau$. If I insert the ansatz into the radial equation, the LHS can vanish by adding and subtracting $(a+n)(b-n)$ to complete the Euler's differential equation for $p_{n+\nu}$. This produces an additional term $(2n\nu+n(n+1))p_{n+\nu}$ at RHS. Now we use the recursion relations for xp and $x(1-x)p'$: we have that

$$xp_{n+\nu} = \gamma_1 p_{n+\nu+1} + \gamma_0 p_{n+\nu} + \gamma_{-1} p_{n+\nu-1} \quad x(1-x)p'_{n+\nu} = \zeta_1 p_{n+\nu+1} + \zeta_0 p_{n+\nu} + \zeta_{-1} p_{n+\nu-1}$$

with

$$\begin{aligned}
 \gamma_1(n) &= -\frac{(n+\nu+1-i\epsilon)(n+\nu+1-i\tau)}{2(n+\nu+1)(2n+2\nu+1)} \\
 \gamma_0(n) &= \frac{1}{2} \left[1 - \frac{\tau\epsilon}{(n+\nu)(n+\nu+1)} \right] \\
 \gamma_{-1}(n) &= -\frac{(n+\nu+i\epsilon)(n+\nu+i\tau)}{2(n+\nu)(2n+2\nu+1)} \\
 \zeta_1(n) &= \frac{(n+\nu+i\tau)(n+\nu+1-i\tau)(n+\nu+1-i\epsilon)}{2(n+\nu+1)(2n+2\nu+1)} \\
 \zeta_0(n) &= \frac{i\epsilon}{2} \left[1 + \frac{i\tau(1-i\tau)}{(n+\nu)(n+\nu+1)} \right] \\
 \zeta_{-1}(n) &= -\frac{(n+\nu+i\tau)(n+\nu+1-i\tau)(n+\nu+i\epsilon)}{2(n+\nu)(2n+2\nu+1)}
 \end{aligned}$$

Thus, we require that

$$\begin{aligned}
 0 = \sum_{n=-\infty}^{+\infty} a_n & \left\{ 2i\epsilon\kappa\sqrt{1-\eta^2} \left[-\zeta_1 p_{n+\nu+1} - \zeta_0 p_{n+\nu} - \zeta_{-1} p_{n+\nu-1} + \right. \right. \\
 & + \left. \left(1 - i\tau + i\epsilon \left(1 - \frac{\eta^2}{2\sqrt{1-\eta^2}} \right) \right) (\gamma_1 p_{n+\nu+1} + \gamma_0 p_{n+\nu} + \gamma_{-1} p_{n+\nu-1}) \right] + \\
 & + \left[-\lambda + \nu(\nu+1) + 2n\nu + n(n+1) + \epsilon^2 \left(1 - \frac{\eta^2(1+\kappa)}{2} \right) - \right. \\
 & \left. \left. - i\epsilon\kappa\sqrt{1-\eta^2} + \epsilon\kappa(\epsilon+\tau) \left(1 - \sqrt{1-\eta^2} \right) \right] p_{n+\nu} \right\}
 \end{aligned}$$

The equality must hold for every $p_{n+\nu}$, thus we have a three-term relation $\alpha_n a_{n-1} + \beta_n a_n + \gamma_n a_{n+1} = 0$, with

$$\begin{aligned}
 \alpha_n &= \frac{i\epsilon\kappa\sqrt{1-\eta^2}(n+\nu+1+i\epsilon)(n+\nu+1+i\tau)(n+\nu+1-i\epsilon(1-\eta^2/(2\sqrt{1-\eta^2})))}{(n+\nu+1)(2n+2\nu+3)} \\
 \beta_n &= -\lambda + (\nu+n)(\nu+n+1) + \epsilon^2(1-\eta^2/2) + \epsilon\tau\kappa + \epsilon^2\kappa \left(1 - \sqrt{1-\eta^2} \right) + \\
 & \quad + \frac{\epsilon^3\tau\kappa}{(n+\nu)(n+\nu+1)} \left(\sqrt{1-\eta^2} - \frac{\eta^2}{2} \right) \\
 \gamma_n &= -\frac{i\epsilon\kappa\sqrt{1-\eta^2}(n+\nu-i\epsilon)(n+\nu-i\tau)(n+\nu+i\epsilon(1-\eta^2/(2\sqrt{1-\eta^2})))}{(n+\nu)(2n+2\nu-1)}
 \end{aligned}$$

Convergence: analogously to massless case, if $R_n L_{n-1} = 1$ is satisfied (we impose that a_n is a minimal succession for both $n \rightarrow \pm\infty$) we have

$$\lim_{n \rightarrow \infty} n \frac{a_n}{a_{n-1}} = - \lim_{n \rightarrow -\infty} n \frac{a_n}{a_{n+1}} = \frac{i\epsilon\kappa}{2} \sqrt{1-\eta^2}$$

thus

$$\lim_{n \rightarrow \pm\infty} n \frac{a_n p_{n+\nu}(x)}{a_{n\mp 1} p_{n+\nu\mp 1}(x)} = \pm \frac{i\epsilon\kappa}{2} \sqrt{1-\eta^2} \left[1 - 2x + \sqrt{(1-2x)^2 - 1} \right]$$

the convergence is thus similar to the massless case, namely the series converges for all $|x|$ except $|x| = +\infty$.

It is also useful to express $R_{\text{in}}^\nu = R_0^\nu + R_0^{-\nu-1}$, with $R_0^\nu = \dots$, which follows directly from the properties of the hypergeometric function. Since the only modifications to the massless case are the coefficient series a_n and the global oscillatory term in the reparametrization of R (which becomes exponential for $\eta > 1$), this expression is also easily generalized to $R_0^\nu = e^{i\epsilon\kappa\sqrt{1-\eta^2}x} \dots \sum_{n=-\infty}^{\infty} a_n^{\text{massive}} \dots$, namely the formalism is exactly the same, only changing the coefficient series and the oscillating factor.

Behavior for $\eta = 1$

I want to discuss the special case $\eta = 1$, for which the previous formulas present some apparent singularities, having the factor $\sqrt{1-\eta^2}$ at denominator. These obviously aren't real singularities, since they are multiplied by the same factor in front (see, for example, the master equation); however, the asymptotic analysis differs, since the final factors are multiplied by $\sqrt{1-\eta^2}$, making the leading order identically zero. In particular, the three-term relation has the following coefficients:

$$\begin{aligned} \alpha_n^{\eta=1} &= \frac{-\epsilon^2\kappa(n+\nu+1+i\epsilon)(n+\nu+1+i\tau)}{2(n+\nu+1)(2n+2\nu+3)} \sim -\frac{\epsilon^2\kappa}{4} \\ \beta_n^{\eta=1} &= -\lambda + (\nu+n)(\nu+n+1) + \epsilon^2/2 + \epsilon\tau\kappa + \epsilon^2\kappa - \frac{\epsilon^3\tau\kappa}{2(n+\nu)(n+\nu+1)} \sim n^2 \\ \gamma_n^{\eta=1} &= -\frac{\epsilon^2\kappa(n+\nu-i\epsilon)(n+\nu-i\tau)}{2(n+\nu)(2n+2\nu-1)} \sim -\frac{\epsilon^2\kappa}{4} \end{aligned}$$

so that

$$\lim_{n \rightarrow \infty} n^2 \frac{a_n(\eta=1)}{a_{n-1}(\eta=1)} = \lim_{n \rightarrow -\infty} n^2 \frac{a_n}{a_{n+1}} = \frac{\epsilon^2\kappa}{4}$$

Note that in this case, the signs are the same for both the $r_* \rightarrow \pm\infty$ limits. All of this is not in contradiction with the case $\eta \neq 1$ since these terms have a different dependence with respect to n : with $\eta \neq 1$, $L/R \sim \pm \frac{i\epsilon\kappa}{2} \sqrt{1-\eta^2} \frac{1}{n}$, which is dominant with respect to $\frac{\epsilon^2\kappa}{4} \frac{1}{n^2}$ unless we have explicitly $\eta = 1$.

In conclusion, differently from what appears from the general asymptotic analysis, also for $\eta = 1$ we have a valid full three-term recurrence relation, the solution is non-trivial and converges even faster.

Behavior for $\eta > 1$

Here we truly expect that the solution is fundamentally different from the massless case. In the radial function reparametrization we have an exponential term $e^{i\epsilon\kappa\sqrt{1-\eta^2}x} = e^{\pm\epsilon\kappa\sqrt{|1-\eta^2|x}}$ which is either divergent or goes exponentially to 0. We expect the second to be the physical solution, since the field is confined within a gravitational potential pit, behaving like

an evanescent wave far from the BH.

To obtain the exponentially damped behavior, I have to choose the branch $\sqrt{1 - \eta^2} = +i\sqrt{|1 - \eta^2|}$. This is in fact a choice, but it is forced by vanishing boundary conditions, so it does not correspond to a real choice of the solution as it was in the case of the α, β exponents.

4.3 Upgoing solution: possible paths for a generalized MST method

While the ingoing homogeneous solution at the horizon can be obtained by means of a modified MST method as described previously, the MST procedure does not generalize to the massive case for the upgoing solution at spatial infinity. In this Section I will motivate this fundamental difference, and despite not providing a definitive solution I will propose a possible path which could lead to a massive MST method at spatial infinity; due to the complexity of the analytical treatise and the fact that it is not strictly necessary to find such method for the further computation of the modified EMRI phenomenology, this investigation will be conducted in a future study.

The massless MST method at infinity is based on a series expansion of *Coulomb functions*, similarly to the role that the hypergeometric function covered for the solution at the horizon. These functions are defined using the *confluent hypergeometric function*, which in terms of a series is defined as

$$\Phi(a; b; x) \equiv \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{x^n}{n!} \quad (4.12)$$

and is a solution of the Kummer's confluent hypergeometric equation

$$zf'' + (c - z)f' - af = 0 \quad (4.13)$$

The Coulomb wave function is

$$z^2 f'' + (z^2 - 2\xi z - L(L + 1))f = 0 \quad (4.14)$$

and its solution is the previously mentioned Coulomb function

$$F_L(\xi, z) = 2^L z^{L+1} e^{-iz} e^{-\pi\xi/2} \frac{\Gamma(L + 1 + i\xi)}{\Gamma(2L + 2)} \Phi(L + 1 - i\xi, 2L + 2, 2iz) \quad (4.15)$$

Note that (Mano et al., 1996) use the unnormalized Coulomb function, where the factor $e^{-\pi\xi/2}$ is omitted. Effectively, this is only a multiplication constant that is irrelevant in the definition of the homogeneous solution, since the normalization is arbitrary.

4.3.a Direct approach

Motivated by the success of the direct approach for the solution at the horizon, namely following the MST treatise including scalar mass terms in the ϵ -perturbative RHS, I adapt the traditional approach by using a proper corrective factor to eliminate the problematic massive terms like $-\eta^2$. I first try with a corrective coefficient $e^{-\eta z}$ other than the usual reparametrization done for the Coulomb solution, obtaining

$$\begin{aligned}
 z^2 f'' + [z^2 + 2\epsilon z - \nu(\nu + 1)]f &= \epsilon \kappa z(f'' + f) + [\epsilon \kappa(-1 + i\epsilon + i\tau) + 2\eta z(z - \epsilon \kappa)]f' - \\
 &\quad - \frac{\epsilon}{z} \kappa(i + \epsilon)(i + \tau)f + \\
 &\quad + (-\nu(\nu + 1) + \lambda - 2\epsilon^2 \left(1 - \frac{\eta^2}{4}(1 - \kappa) + \frac{i\eta\kappa}{2}\right) + \\
 &\quad + \epsilon m q + \kappa \epsilon^2 + \epsilon \eta(z\eta + \kappa - i\kappa\tau))f
 \end{aligned}$$

The blue terms contain the mass corrections, in order to follow more simply the modifications to the massless MST. This master equation can be expressed with a RHS in an ϵ -expanded form as:

$$\begin{aligned}
 z^2 f'' + [z^2 + 2\epsilon z \left(1 - \frac{\eta^2}{2}\right) - \nu(\nu + 1)]f &= \epsilon \kappa z(f'' + f) + [\epsilon \kappa(-1 + i\epsilon + i\tau) + 2\eta z(z - \epsilon \kappa)]f' - \\
 &\quad - \frac{\epsilon}{z} \kappa(i + \epsilon)(i + \tau)f + \\
 &\quad + \left[-\nu(\nu + 1) + \lambda - 2\epsilon^2 \left(1 - \frac{\eta^2}{4}(1 - \kappa) + \frac{i\eta\kappa}{2}\right) + \right. \\
 &\quad \left. + \epsilon m q + \kappa \epsilon^2 - i\eta\epsilon\kappa(i + \tau) \right]f
 \end{aligned}$$

This second form involves corrective factors such as $1 - \eta^2/2$, which were already present in the hypergeometric expansion. However, in this case the LHS operator is modified, meaning that the massless Coulomb functions themselves cannot serve as a 0-order solution corresponding to the massless case (differently from the hypergeometric expansion at the horizon).

Not only a Coulomb function with massless parameters cannot serve as a basis for the expansion, but also a generic Coulomb function with modified parameters. This follows from the following observations:

- It is not possible to eliminate the extra coefficient $z(z - \epsilon\kappa)f'$ by multiplying the reparameterization for a generic function $w(z)$: the solution is $w = e^{\eta z}$ which would compensate for the $e^{-\eta z}$ factor previously introduced.
- It is not possible to have any effect on the terms containing f' by considering $\tilde{f} = f^\alpha$ with α free, since it does not have any effect on the coefficient of f' .
- $\implies$ the 0-order solution cannot be a "modified Coulomb function", namely a function proportional to the confluent hypergeometric function via some prefactor.

4.3.b Usage of the double-confluent Heun Equation

The double-confluent Heun equation (DHCE) is given by (Jaccoud El-Jaick et al., 2008):

$$z^2 f'' + (B_1 + B_2 z)f' + (B_3 - 2\xi\sigma z + \sigma^2 z^2)f = 0 \quad (4.16)$$

This exactly resembles the functional form of the master equation, including the problematic term $\eta z(z - \epsilon\kappa)f'$, with the coefficients mapped in the following quantities:

$$\begin{aligned} B_1 &= 2\eta\epsilon\kappa \\ B_2 &= -2\eta \\ B_3 &= -\nu(\nu + 1) \\ \xi &= \mp\epsilon(1 - \eta^2/2) \\ \sigma &= \pm 1 \end{aligned}$$

Note that $B_1 \neq 0$ is required: this implies that the behavior of the Coulomb function is not immediately recovered in the massless case $\eta = 0$, or at least a limit treatment is necessary. The solution of the DHCE, Eq. (4.16), is in the form of an infinite sum of confluent hypergeometric functions (either regular or irregular, corresponding to the two independent solutions of the 2nd order differential equation) characterized by a three-term recurrence relation. As a result, it could be feasible to use this solution (which is itself an infinite sum of hypergeometric functions determined by a specific three-term recurrence relation) as a 0-order solution of the master equation, and successively find the ϵ -expansion by means of an infinite sum with a different characteristic three-term recurrence relation. Note that this treatment becomes more involved than the MST method, since two contiguous relations should be found for the 0-order solution $U(z)$, one for $z^{-1}U(z)$ and one for $U'(z)$, which could be possible but surely complicated, since the series coefficients are not explicit.

4.4 Upgoing solution in terms of a Frobenius expansion

4.4.a Implementation of the massive MST method

The massive MST method to compute the ingoing solution at the horizon, as described in the previous Sections, has been implemented in the code STORM. Despite the theoretical self-consistency of the theory, its physically-motivated features and multiply-checked derivations, I found a systematic non-constant discrepancy of the order of 1% – 2% with the fluxes predicted by Barsanti (2024) and Barsanti et al. (2023). Given the conceptual and practical complexity and length of the MST implementation in the code, it is very likely that an unknown (and very tiny) implementation error is causing the observed discrepancy. Not having found the error in a reasonable time, I resorted to a method which is conceptually simpler, even though much more computationally expensive, namely the expansion of the homogeneous solution in terms of a Frobenius series, presented in the following. Still lacking a massive MST series at spatial infinity, the Frobenius series method would have been used in any case for the upgoing solution, even in the presence of a working version of the massive MST at horizon.

4.4.b The Frobenius expansion

The method described in this section has been used in Barsanti et al. (2023) and Barsanti (2024) to find the homogeneous solutions of the Teukolsky equation. It is based on the

physically-motivated asymptotic behavior of the solutions, found in 3.2.a directly from the massive scalar field master equation. The fundamental ansatz is then obtained by multiplying the asymptotic solution by a Taylor/Laurent series expansion, which will capture the behavior of the solution in a small neighborhood of the horizon and infinity. The expansion at infinity reads:

$$u_{\text{up}}(r) = e^{i\sqrt{\omega^2 - \mu^2}r_*} r^{\frac{iM\mu^2}{\sqrt{\omega^2 - \mu^2}}} \sum_{n=0}^{\infty} \frac{b_n}{r^n} \quad (4.17)$$

and at the outer horizon:

$$u_{\text{in}}(r) = e^{-ik-r_*} \sum_{n=0}^{n_h} a_n (r - r_h)^n \quad (4.18)$$

We want to iteratively solve the coefficients a_n, b_n ; The symbolic expression of the coefficients in terms of the usual adimensional parameters is found using WOLFRAM MATHEMATICA, and subsequently translated to C++ for the STORM implementation. The coefficients b_n that solve the ansatz inside the master equation are listed in App. A; the coefficients a_n are too long to be included.

This will provide the boundary conditions, meaning that we also need to compute the coefficients for u' and u'' . I found:

$$\begin{aligned} u'_{\text{up}}(r) &= e^{i\sqrt{\omega^2 - \mu^2}r_*} r^{\frac{iM\mu^2}{\sqrt{\omega^2 - \mu^2}}} \sum_{n=0}^{\infty} \frac{b_n}{r^n} \left[i\sqrt{\omega^2 - \mu^2} r'_* + \frac{iM\mu^2}{r\sqrt{\omega^2 - \mu^2}} - \frac{n}{r} \right] \\ u''_{\text{up}}(r) &= e^{i\sqrt{\omega^2 - \mu^2}r_*} r^{\frac{iM\mu^2}{\sqrt{\omega^2 - \mu^2}}} \sum_{n=0}^{\infty} \frac{b_n}{r^n} \left[\frac{n(n+1)}{r^2} - \frac{M^2\mu^4}{r^2(\omega^2 - \mu^2)} - \frac{iM\mu^2(2n+1)}{r^2\sqrt{\omega^2 - \mu^2}} - \right. \\ &\quad \left. - \frac{2r'_*(M\mu^2 + in\sqrt{\omega^2 - \mu^2})}{r} - (\omega^2 - \mu^2)r_*'^2 + i\sqrt{\omega^2 - \mu^2}r_*'' \right] \end{aligned}$$

where the Tortoise coordinate and its derivatives are explicitly:

$$\begin{aligned} r_*(r) &= r + \frac{2M(M + \sqrt{M^2 - a^2})}{2\sqrt{M^2 - a^2}} \log \frac{r - M - \sqrt{M^2 - a^2}}{2M} - \\ &\quad - \frac{2M(M - \sqrt{M^2 - a^2})}{2\sqrt{M^2 - a^2}} \log \frac{r - M + \sqrt{M^2 - a^2}}{2M} \\ &= r + \left(1 + \frac{1}{\sqrt{1 - a^2}}\right) \log \frac{r - r_+}{2} + \left(1 - \frac{1}{\sqrt{1 - a^2}}\right) \log \frac{r - r_-}{2} \quad (\text{normalized } M=1) \\ r'_*(r) &= \frac{r^2 + a^2}{r^2 - 2Mr + a^2} \\ r''_*(r) &= -\frac{2M(r^2 - a^2)}{(r^2 - 2Mr + a^2)^2} \end{aligned}$$

4.5 Numerical integration of the Teukolsky equation

In order to obtain the solution to the Teukolsky field equation on a finite radial interval, it is more computationally efficient to integrate the field equation rather than computing the MST semianalytical solution point-by-point. In addition, the small convergence radius of the Frobenius expansion is not in general sufficient to provide a series expansion of the field at the radius of the secondary, where the evaluation is necessary for the flux computation. The MST method and the Frobenius expansions are thus used to provide accurate fast and accurate boundary conditions, from which is then possible to start the numerical integration. Note that in the case of circular orbits the solution must be evaluated at a single point rather than a finite interval; for this reason, we can directly use the MST method for the ingoing solution, but the inwards integration of the upgoing solution is still needed due to the limited Frobenius series radius of convergence.

A transformation of the Teukolsky equation is usually performed in order to deal with a short-range potential that decays as r^{-2} :

$$R(r) = r^{-1} \Delta^{-s} e^{i\omega h + im\tilde{\phi}} \psi(r)$$

where $h(r)$ is a function (*height function*) such that its derivative with respect to the tortoise coordinate, $H \equiv dh/dr_*$ has the following properties:

$$\lim_{r_* \rightarrow \pm\infty} H = \pm 1 \qquad |H| \leq 1 \qquad \lim_{r_* \rightarrow \pm\infty} H' = 0$$

and the new angular coordinate $\tilde{\phi}$ is defined as

$$d\tilde{\phi} = d\phi + \frac{a}{r^2 + a^2} dr_*$$

this choice of the angular coordinate, together with the tortoise radial coordinate, regularizes the behavior of the field equation near the horizon. Morally, this angular coordinate is “undoing”, “unwinding” the effect of the frame dragging caused by the BH rotation.

With this reparameterization, the Teukolsky equation for a massive field takes the form

$$\frac{(r^2 + a^2)^2}{r^4} \frac{d^2 \psi}{dr_*^2} - \frac{2\tilde{G}}{r^5} \frac{d\psi}{dr_*} + \frac{\tilde{U}}{r^6} \psi = 0$$

where

$$\begin{aligned} \tilde{G} &= a^2 \Delta + (r^2 + a^2)(rs(r - M) - ir((r^2 + a^2)\omega H + ma)) \\ \tilde{U} &= 2is\omega r^2(r\Delta(1 - H) - M(r^2 - a^2)(1 + H)) - 2iar\Delta(m + a\omega H) + \Delta(2a^2 - r^2\lambda - 2Mr(s + 1)) \\ &\quad - 2ma\omega r^2(r^2 + a^2)(1 + H) + r^2(r^2 + a^2)^2(\omega^2(1 - H^2) + i\omega H' - \Delta\mu^2) \end{aligned}$$

Note that $\tilde{G}$ is independent of the mass, whereas $\tilde{U}$ has a relevant term $-r^2(r^2 + a^2)^2\Delta\mu^2$. With $\mu = 0$, the effective potential $\tilde{U}r^{-6}$ is short-ranged (it goes like r^{-2} at ∞), but with a non-vanishing scalar mass the potential assumes an asymptotical constant value. It is impossible to find a reparameterization that brings the massive equation in the form of a short-ranged potential: multiplying the new function for a generic factor $e^{A(r)}$ has the effect of producing a massive term in $\tilde{U}$ that contains factors of $O(r^6)$, $O(r^5)$ with some

prefactors that depend on A' only and A' , A'' respectively. However, this means that we have only one degree of freedom (namely the choice of A') to satisfy the condition that the prefactor of r^6 should behave like r^{-2} , namely we have 2 constraints, one for the term of order r^0 and one for the term r^{-1} , which cannot be satisfied at the same time. The best we can do is to find a potential that goes like r^{-1} . Not having the real advantage of a short-ranged potential, I prefer to maintain the standard redefinition and integrate the equation with asymptotically constant potential. This has the advantage of needing a minimal change to the code already used for massless perturbations.

The choice of the height function $h(r)$ is the minimal one, namely $H(r)$ is fixed at ± 1 for upgoing/incoming solutions, respectively. Although this choice does not in principle satisfy the asymptotic limits (e.g., for the ingoing solution $H = -1$ we still expect that $H \rightarrow +1$ as $r \rightarrow +\infty$), the integration is actually performed from $r_+ \sim \pm\infty$ to a finite value, so that we can imagine H detaching from the constant value and approaching the correct asymptotic limit only in the interval where the integration is not performed. With this choice, $h(r) = r_*$ up to an integration constant.

Wronskian computation

Since we do not have an analytic solution for the upgoing solution, the Wronskian must be computed numerically. This does not imply the need for an integration up to the radius of the particle, because the quantity $\Delta \times W$ is constant over the radial coordinate. This can be seen starting from the radial field equation

$$\frac{d^2 R}{dr^2} + \frac{\Delta'}{\Delta} \frac{dR}{dr} + V_{\text{eff}} R = 0$$

Using the Abel's identity for second order ODEs, the evolution of the Wronskian is given by

$$W(r) = C e^{-\int \frac{\Delta'}{\Delta} dr} = C e^{-\ln \Delta} = C \Delta^{-1}$$

where C is a constant.

Since the integration of the upgoing solution is necessary anyway, the Wronskian is computed at a fixed radius in the proximity of the horizon, where the Frobenius solution converges fast. In practice, $r = r_+ + 2 \cdot 10^{-5}$ is used to compute the boundary conditions, and then the field solutions are integrated to $r = r_p$. Afterwards, the constant combination $W \Delta$ is used on the entire range.

4.5.a Spin-weighted spheroidal harmonics computation

The spin-weighted spheroidal harmonics ${}_s S_{\ell m} : \mathcal{S}^1 \mapsto \mathbb{C}$ are the solutions of the equation

$$\left[\frac{1}{\sin \theta} \frac{d}{d\theta} \left(\sin \theta \frac{d}{d\theta} \right) - \zeta^2 \sin^2 \theta - \frac{(m + s \cos \theta)^2}{\sin^2 \theta} - 2\zeta \cos \theta + s + 2m\zeta + {}_s \lambda_{\ell m} \right] {}_s S_{\ell m} = 0 \quad (4.19)$$

Note that the dependence from the longitudinal variable ϕ is completely factorized out of this definition into the Fourier transform previously defined, as a factor $e^{im\phi}$. As a result, ${}_s S_{\ell m} = {}_s S_{\ell m}(\theta)$. The parameter $\zeta \in \mathbb{C}$ is the *spheroidicity*, also called *oblateness* or *prolateness* in the specific cases in which it takes purely real or purely imaginary values, respectively.

The calculation of the functions and the corresponding eigenvalues are originally performed by the STORM code in the case of oblated harmonics, $\zeta = a\omega$. The differential equation is real, and thus both the eigenvalues and eigenfunctions are real. This is not the general case, since for $\eta > 1$ we have $\zeta = a\omega\sqrt{1-\eta^2}$ purely imaginary. However, for a massive scalar field $s = 0$, and the linear contributions in ζ vanish, since Eq. (4.19) remains hermitian (and in particular real).

The generalization in the code is obtained by promoting the variables involving ζ from real to complex types. The method used to compute the eigenfunctions, the Leaver's method, works also for prolate spheroidicity, thus the available code has been used with minimal changes.

Results: waveforms and observability forecasts

5.1 Field solution and scalar fluxes

The previous Chapter was dedicated to finding the homogeneous solutions for a massive scalar field with dynamics set by the homogeneous Teukolsky Equations. Once the homogeneous is known, the integration scheme that implements the Green's function method is completely equivalent to the one that is already implemented in the code STORM (Gliorio et al., 2026) for a massless scalar field in a generic orbital configuration. The only significant differences are the expressions of the energy and momentum fluxes at spatial infinity, which read

$$\dot{E}_s^+ = \frac{1}{16\pi} \sum_{\ell,m} \omega k_+ |Z_{\ell m}^+|^2 \quad (5.1)$$

$$\dot{L}_s^+ = \frac{1}{16\pi} \sum_{\ell,m} m k_+ |Z_{\ell m}^+|^2 \quad (5.2)$$

where $k_+ = \sqrt{\omega^2 - \mu_s^2}$ contains a further dependence on the scalar mass, other than the obvious dependence inside the field solution, which is implicitly encoded in $Z_{\ell m}^+$. Once these changes are correctly implemented, the STORM code is able to produce the corresponding energy fluxes. In this and the following Chapters, I will use a version of STORM I specifically modified as detailed in the previous Chapters to include the effects of the massive scalar field. The difference from previous works (Barsanti et al., 2023; Barsanti, 2024) is that the new code efficiency and generality (it naturally implements generic orbits, namely eccentric and inclined, with, however, the possibility of selecting the special cases of circular equatorial and circular eccentric orbits) allows for a relatively fast and systematic computation of the scalar fluxes and the phenomenology that follows.

The computations were executed on the GWAVE cluster at the Pennsylvania State University, a Tier-2 center of the LIGO Data Grid and part of the International Gravitational-Wave Network (IGWN) Computing Grid. The availability of a cluster infrastructure enabled parallelization of the tasks, making it possible to compute the energy fluxes on large grids later used for interpolation.

Due to limited time, mainly the circular equatorial case was systematically explored: the computational cost of fluxes evaluation on a grid increments steeply when adding eccentricity and inclination, since not only it is necessary to add the corresponding grid dimensions (and thus a power scaling with an exponent that grows like the dimensionality), but also a sum over the additional mode numbers n, k , for which a finite-size spec-

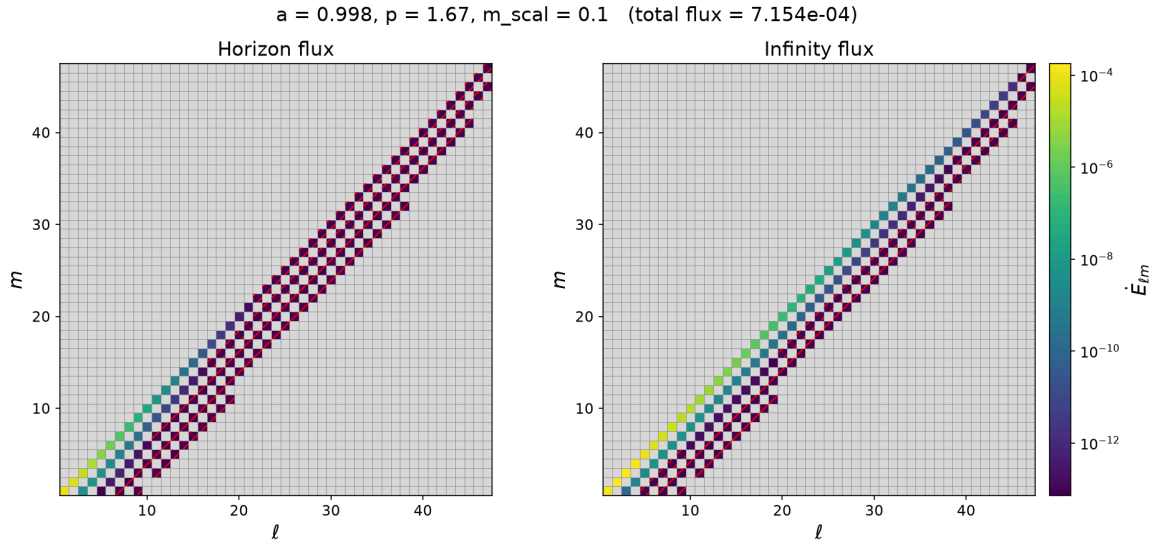
trum appears as respectively eccentricity and inclination are added (corresponding to the characteristic frequencies Ω_r, Ω_θ of the angle variables). At the end of this Chapter I will briefly present the results on a single eccentric case study, for which a local grid was computed only in the neighborhood of a 2-dimensional manifold spanning only eccentricity and semilatus rectum. The availability of an efficient code and a computational infrastructure will allow for a complete investigation of the fluxes as a follow-up study of the work presented in this Thesis.

5.2 Energy fluxes, grid sampling and interpolation

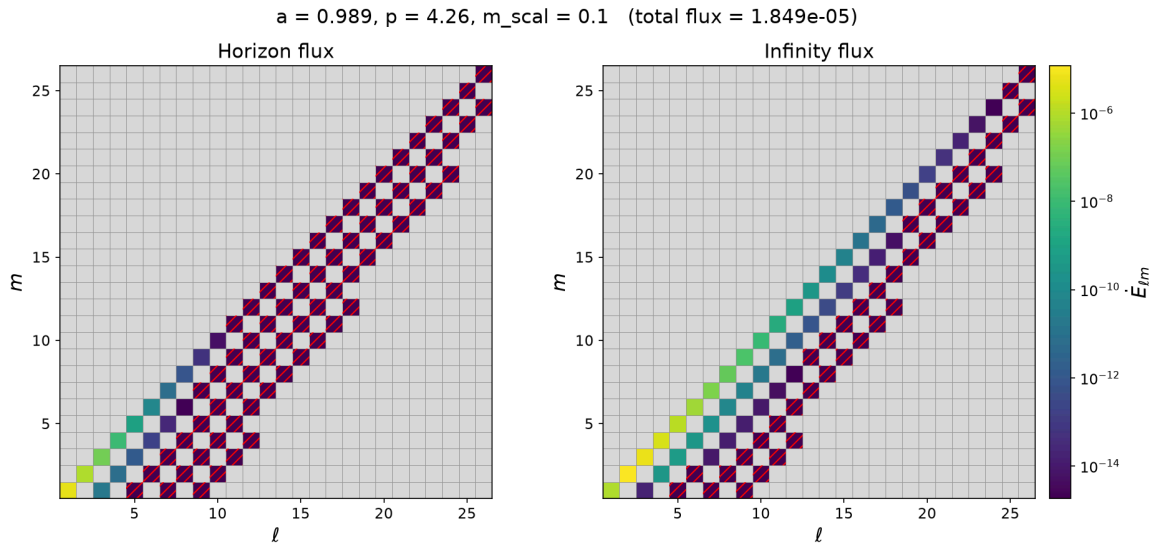
5.2.a Energy scalar fluxes with STORM

The calculation of the scalar flux for a given set of parameters consists of a summation of ℓ, m modes ($n = k = 0$ for circular equatorial orbits). The computation is simplified by the symmetry property $\dot{E}_{\ell,m} = \dot{E}_{\ell,-m}$ and by the selection rule $\dot{E}_{\ell,m} = 0$ if $\ell + m$ is odd (ℓ and m have different parity). The summation proceeds from $\ell = 1$ to a maximum cap $\ell_{\max} = 50$, and is interrupted before if two consecutive ℓ -modes bring a energy flux contribution that is less than 10^{-10} of the total flux accumulated until then. This threshold is chosen such that mode truncation does not lead to systematic biases in the binary orbital dynamics. Nested inside the ℓ -model computation, the m -modes summation follows the exact same logic, starting from $m = \ell$ and going down at steps of 2 (for the selection rule), and stopping when the relative energy contribution condition is satisfied. Three typical examples of mode hierarchy at extremal spin and different radii are shown in Fig. 5.1; the computed modes are shown as colored (note the logarithmic color scale for the absolute value of the energy flux), and those that are below the energy threshold are hatched. In Fig. 5.1a the modes corresponding to an orbit near the Innermost Stable Circular Orbit (ISCO) are plotted: the mode structure is very rich, especially for the flux at spatial infinity, reaching $\ell = 47$; in addition, the mode that contributes more to the flux is not the lowest $\ell = m = 1$, but instead $\ell = m = 3$. Incrementing the radius, in Fig. 5.1b we can observe a more typical behavior, with $\ell_{\max} \sim 20$ and still a prevalent flux at infinity. Finally, if the radius is incremented further (Fig. 5.1c) we can enter the field confinement region in which $m\Omega_\phi < \mu_s$: the flux at infinity is identically zero, and superradiance can occur (note the negative total energy flux). Recall, however, that superradiance and field confinement are two independent conditions: in fact, for the displayed configurations, every mode at the horizon is superradiant, thus the energy flux at the horizon is always negative (globally and mode-wise). A net superradiance is observed only in the last configuration because the flux at infinity, which is usually dominant, is instead suppressed by the massive field confinement.

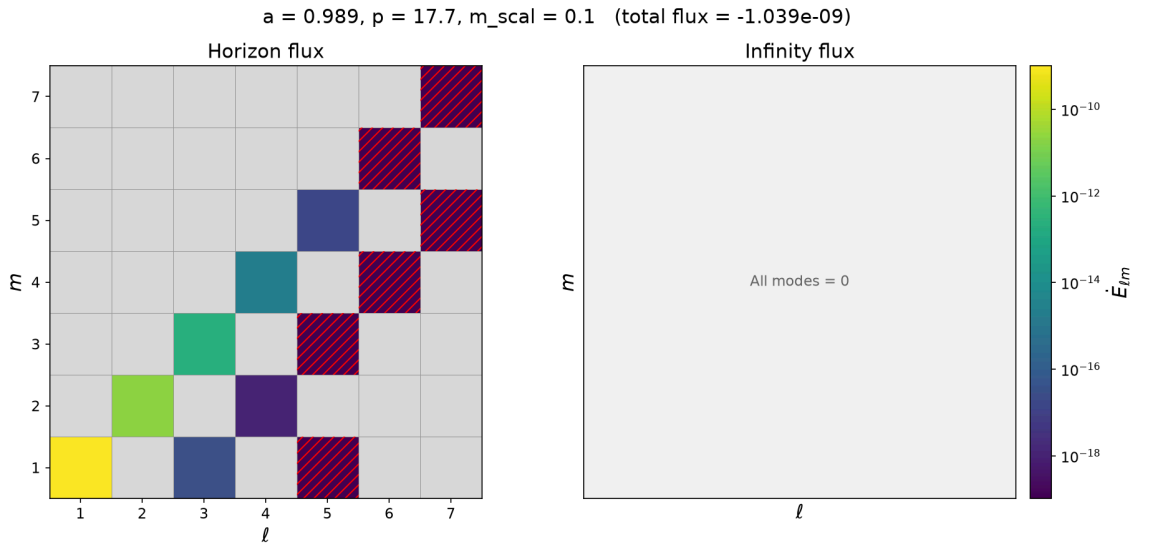
For circular equatorial orbits, I used a grid that covers the entire physically relevant space in the primary spin a and the semilatus rectum p (which reduces to the orbital radius), following the grid building prescription of Chapman-Bird et al. (2025). The grid is sampled taking into account the fact that the innermost stable circular orbit (ISCO) prevents circular orbits to exist under a given radius for each value of the spin a ; as a result, the spin a is sampled more finely near the extremal configuration $a \sim 0.99$; for each value of the spin, the radius is then sampled from ISCO to $p_{\max} = 200$ with 195 values, more finely in an inner zone and coarser in the outer zone. For details about the



(a) Near-ISCO orbit for extremal Kerr. The system is superradiant at horizon and radiative at infinity, with a net positive energy flux distributed over a large spectrum in ℓ .



(b) Middle-region orbit, still globally radiative but with a narrower ℓ -spectrum.



(c) High-radius orbit, superradiant and confined by the low-energy configuration $m\Omega_\phi < \mu_s \forall m < \ell_{\text{max}}$.

Figure 5.1: Typical energy flux $\dot{E}$ modes hierarchy for extremal spin $a = 0.998, 0.989$, circular equatorial orbits with orbital radius $p = 1.67, 4.26, 17.7$. Fluxes at the horizon and at spatial infinity are shown separately but with common colorscale within the same subfigure. The colorscale refers to the absolute value of the energy flux; the horizon flux is always superradiant, namely for all the three configurations every mode at horizon has negative energy flux.

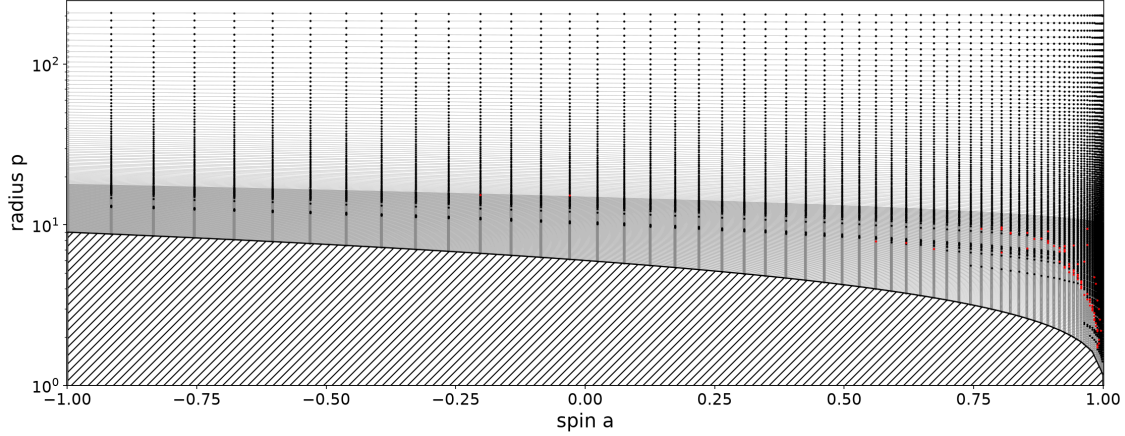


Figure 5.2: Computation grid for global spin-radius coverage. The hatched area corresponds to radii below the ISCO, black markers to configurations for which the scalar fluxes have been computed, red markers for incomplete (ongoing) computation, and hollow markers for configurations yet to be computed, referred to the case study $\mu_s = 0.1$ (and scalar mass values nearby).

grid sampling, refer to App. B of Chapman-Bird et al. (2025).

In order to compute the Fisher matrices, which are built from the partial derivatives of the waveform with respect to a single parameter at the time, in addition to the grid sampling it is necessary to sample a small neighborhood of the scalar mass parameter. In this case, the two case studies $\mu_s = 0.1$ and 0.01 have been chosen, and a total of 9 scalar mass values centered around the reference μ_s are considered.

In Fig. 5.2 is represented the sampling grid (hollow markers) in the $a - p$ parameter space, with the actual completion of the flux calculation (black markers, still incomplete). The parallel computation of fluxes is carried out at constant radial index, namely the parameter space scan starts from extremal values of $a = 0.99$ and proceeds towards lower values.

Data quality and completeness

As it can be seen from Fig. 5.2, the scan is still incomplete since low radial values need much higher modes contributions to converge, lengthening the computational time. In addition, for some configurations a failure to compute single, isolated modes has occurred; these configurations have been excluded from the interpolation later, since the missing modes can produce severe artifacts if used in the interpolation. The cause of the mode failure has been detected and removed for the other evaluations¹, however the data quality is sufficient to obtain relevant and reliable results, under a correct treatment of these isolated defects.

¹The error was occasionally produced by the functions implementing the MST method calculations; even if the massive scalar field code does not depend on the evaluation of such quantities, they were computed as part of the whole `ScalarFluxMode` class initialization. For such reason, the sporadic occurrence of errors during the usage of the class is not a symptom of numerical/physical issues in the computation of the massive field fluxes, but nevertheless used to compromise the application of the class when defined with certain parameter sets. A further refinement of the implementation, which bypasses the computation of the MST quantities, completely eliminated these failures and further improved the evaluation speed.

5.2.b Interpolation

The energy fluxes grid described in the previous section were used to build an interpolator (based on `REGULARGRIDINTERPOLATOR` of the `SCIPLY` suite (Virtanen et al., 2020)) with proper masking to account for mode failure defects in the grid. Visual inspection led to the choice of the `pchip` method which, despite being very slow to evaluate, did not produce oscillating artifacts near mode failure defects, as it was the case for faster methods like `cubic`.

In Fig. 5.3 are plotted the interpolated energy fluxes on the valid domain (left column) and the derivative of the scalar fluxes with respect to the scalar mass (right column), obtained with a finite-size difference approach through the interpolator (2-points symmetric half-step stencil with $\delta\mu_s = 10^{-8}$ chosen for stability and reduced production of artifacts). Fig. 5.3a shows the results for $\mu_s = 0.1$, the lower plot Fig. 5.3b the results for $\mu_s = 0.01$. The coverage of the two cases is very similar. The physical features of the fluxes and the derivatives are discussed in the next Section. Here we notice only the presence of small and isolated artifacts, limited to the derivatives plots (not on the energy fluxes), for both scalar masses in correspondence of $a \sim 0.8$, $p \sim 5$, likely due to mode failure defects.

5.3 Results: energy scalar fluxes for circular equatorial orbits

5.3.a Theoretical explanation of energy flux features

The main qualitative features of the energy flux are determined by two main phenomena: *superradiance* and *field confinement*. Let's briefly review the expected properties and definitions:

- **Superradiance:** for a deep understanding and treatise of superradiant phenomena in their generality and the application to massive scalar fields around Kerr BHs, see the dedicated Chapter 6. For now, the term *superradiance* refers to the extraction of energy from the primary BH through the scalar field, which is possible mode-wise when $k_- = \omega - m\Omega_h < 0$: when this condition is met, it immediately follows from Eq. (3.8) that the energy flux at the horizon is negative, namely energy is being delivered from the primary to the secondary through the scalar field. Remembering that $\Omega_h = a/(2Mr_+)$, where $r_+ = M + \sqrt{M^2 - a^2}$ is the outer horizon radius, and that for circular equatorial orbits $\omega = m\Omega_\phi$ with $\Omega_\phi = \frac{M^{1/2}}{r^{3/2} + aM^{1/2}}$ for prograde orbits, the superradiant condition becomes independent of the index m^2 , and in particular is given by

$$\frac{r}{M} > \left[\frac{2(M + \sqrt{M^2 - a^2})}{a} - \frac{a}{M} \right]^{2/3} \quad (5.3)$$

This defines a region of the $p - a$ plane where the flux at horizon is superradiant (mode-by-mode, since the condition is independent from m). Note also that this

²Note that for $m = 0$ we cannot divide the superradiant condition by m to show that it's m -independent, however for $m = 0$ the angular momentum (and thus the energy) flux is manifestly 0, thus it doesn't contribute. Equivalently, for $m = 0$ we have $k_- = 0$, namely the mode sits exactly between superradiance and standard radiative regime, thus $\dot{E}_{m=0}^- = 0$.

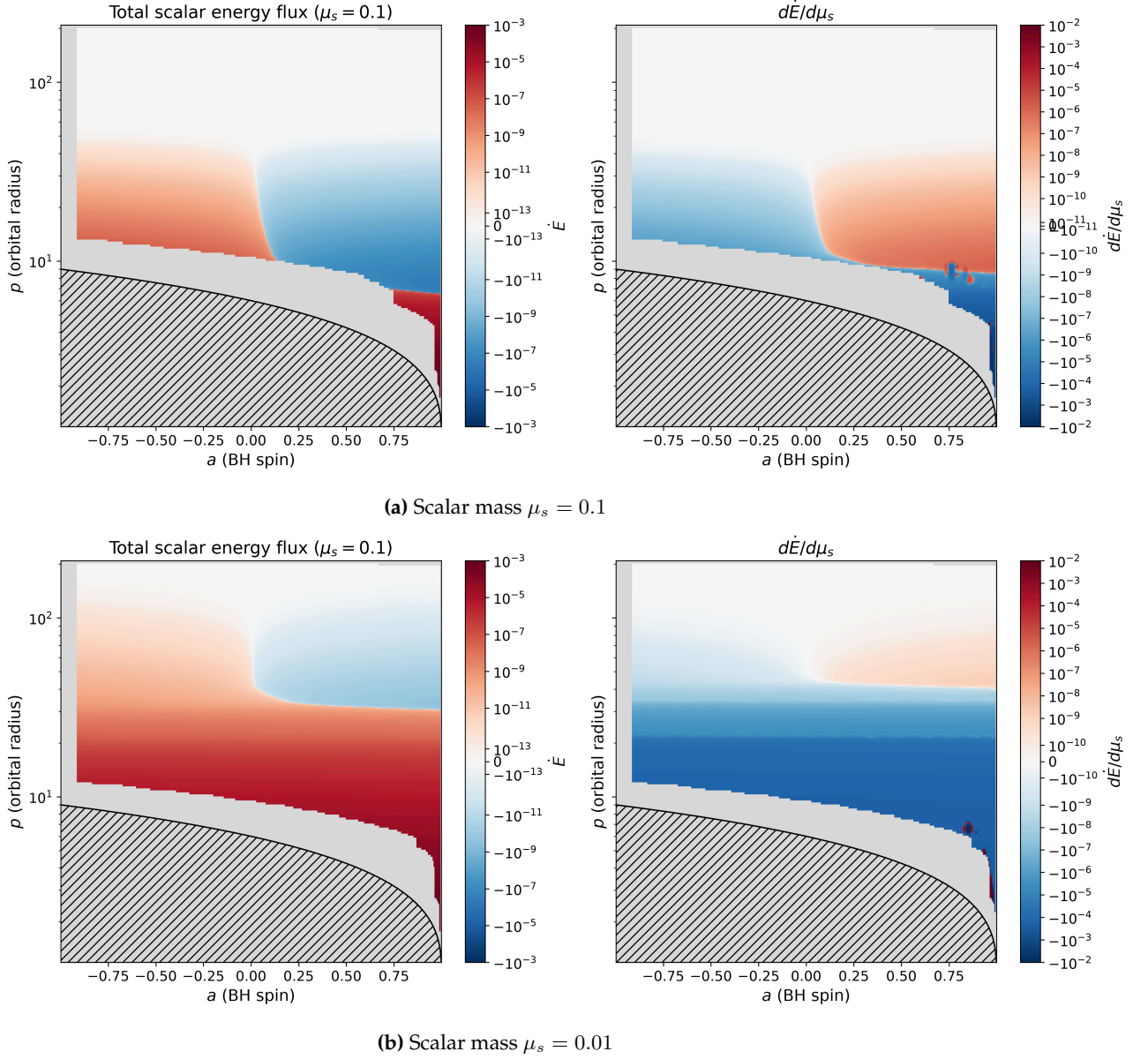


Figure 5.3: Interpolated total energy fluxes due to the scalar field (left column) and their derivative with respect to the scalar mass (right column). Scalar mass μ_s is equal to 0.1 for Fig. 5.3a and to 0.01 for Fig. 5.3b. The hatched region represents the zone inside the ISCO. Refer to Sec. 5.2.b, 5.3 for details and interpretation.

condition is not exclusive to massive fields; in fact, the scalar mass has no role in defining the occurrence of superradiance.

Comparing Eq. 5.3 with the ISCO we obtain that, for circular equatorial orbits, superradiance occurs for every value of the orbital radius if $a \gtrsim 0.4$.

- **Massive field confinement** Referring again to Eq. (3.8), we observe that if $\omega < \mu_s$, then k_+ becomes purely imaginary. Since the energy flux is a real quantity, the field amplitude contribution $|Z_{\ell,m}|^2$ must then vanish. This is what can be called *field confinement*: for a massive scalar field, orbital frequencies $\omega = m\Omega_\phi$ that are less than the scalar mass cannot produce radiative modes at infinity; thus, the energy flux at spatial infinity (depending on the specific mode choice) will be null. This can be seen in Fig. 5.1c, where the radius is large enough to produce an orbital frequency that cannot radiate energy for at least $m < m_{\max} = 7$ in this specific example. Note that higher-order modes emission is possible, but the mode hierarchy implies that these higher modes will contribute with a negligible energy flux. The energy confinement condition is the following: a mode is confined if $\omega < \mu_s$, namely

$$\frac{r}{M} > \left[\frac{m}{M\mu_s} - \frac{a}{M} \right]^{3/2} \quad (5.4)$$

In this case, the condition depends on the mode index m ; this means that the interplay between this effect and superradiance cannot be predicted *a priori* without knowing the flux amplitudes. Note that the dependence on the spin parameter a is subdominant with respect to little μ_s and large m : with the chosen scalar masses $\mu_s = 0.1, 0.01$ the dependence on the primary BH spin is practically lost.

These two main effects are graphically represented in Fig. 5.4, overlapped to the previously presented interpolated energy fluxes for each of the two considered scalar masses. The superradiance condition, affecting every and only the horizon flux modes, is represented with the red line, marking the equality of Eq. (5.3), and is common between Fig. 5.4a and 5.4b. Above the line superradiance occurs under the line the horizon is absorbing energy (“standard” radiative behavior towards the primary’s horizon). The black lines mark the equality of Eq. (5.4), the field confinement condition, which affects radiative modes at spatial infinity and depend both on the m index (plotted $m = 1, \dots, 20$) and on the scalar mass (the set of lines is different between the two plots).

Let’s consider $\mu_s = 0.1$, namely Fig. 5.4a. The theoretical prediction for the occurrence of superradiance matches very well the boundary between net radiance and superradiance at high orbital radius. This confirms the fact that the field confinement, which in this region applies to several values of m , limits the effect of radiative modes on the total flux, which is then determined purely by the flux at horizon. Radiative modes become dominant only at small p (lower right region of the plot), where the low- m modes are no longer confined. We observe in particular that the most relevant mode is $m = 2$ (in principle receiving contributions from every ℓ with even parity, but practically dominated by $\ell = 2$), which displays a sufficient magnitude to return to globally radiative behavior as soon as the corresponding confinement radius is crossed.

Similarly for $\mu_s = 0.01$, where the superradiant region is not interrupted by grid incompleteness due to a lower scalar mass that shifts the confinement radii towards a higher p ; as before, the relevant index m for obtaining a net positive energy flux in the superradiant region corresponds to $m = 2$. This is in accordance with the mode hierarchy found

in Sec. 5.2.a and Fig. 5.1, where it can be seen that $\ell = 2$ is prominent at a small orbital radius. In this second example, it is also visually evident the quick appearance of single m -modes contributions when the confinement radius is reached: near the confinement boundary, the magnitude of the energy flux is dominated by the $k_+ = \sqrt{m^2\Omega_\phi^2 - \mu_s^2}$, which decreases sharply as $m\Omega_\phi$ approaches μ_s (with vertical tangent at confinement radius, given by the characteristic behavior of the square root). These horizontal features (due to the weak dependence of the mode confinement radius from the spin a) are somewhat visible in the energy flux plots (see also the interpolated energy fluxes of Fig. 5.3, where the mode confinement boundary lines do not visually interfere) even if they are “diluted” by the presence of other modes with comparable magnitude; they become extremely well visible when considering the derivative of the energy flux with respect to the scalar mass, right panels of Fig. 5.3: since the flux at the horizon depends weakly on the scalar mass (morally: the effective potential reduces exactly to the massless case at the horizon, namely *the massive field behaves as massless in strong potentials*, refer to Chapter 3) and in particular the energy flux does not exhibit the strong dependence of k_+ on μ_s , the derivative will isolate and highlight the k_+ derivative features in the proximity of the mode confinement boundaries. This is exactly what can be clearly seen in the right panel of Fig. 5.3b, where progressive “layers” of radiative modes are switched on as the orbital radius approaches the ISCO.

5.3.b An unidimensional case study

After understanding the global features of the energy flux, let’s consider a single case study, namely a slice of the parameter space with reference parameters $a = 0.9$ and $\mu_s = 0.1$. The data used in this isolated case are computed independently from the previously presented data set. The radial coordinate is still sampled over the entire range, choosing the radial sampling provided by the 2-dimensional grid; to compute the derivatives, 5 values of scalar mass and BH spin are also sampled around the reference configuration. The final coverage, excluding datapoints that suffered mode failure defects, is $\approx 95\%$.

This case study, which is more complete also towards smaller radial values, gives the opportunity of assessing the role of each mode contribution to the energy flux and its derivative with respect to the scalar mass. In Fig. 5.5a the energy fluxes are plotted for each m , at horizon and infinity separately and then summed. In the lower semiplane the dotted lines report the horizon fluxes: note that since for $a = 0.9$ the flux at horizon is always superradiant. In contrast, the upper semiplane dotted lines refer to the spatial infinity flux; the total fluxes, mode-by-mode and summed, are plotted with a solid line. To achieve a single visualization of negative and positive energy fluxes, a small region with linear y -axis scaling is interposed between $\dot{E} = \pm 10^{-10}$. Mode confinement radii are reported for $m = 1, 2, 3, 4$.

Starting from high radius (from right to left of the plot 5.5a): the prevalent mode is the superradiant $\ell = m = 1$ energy flux at the horizon: the flux at infinity is zero by field confinement and far from the strong potential region the modes follow a monotonic ℓ -hierarchy. At $p \approx 10$, the envelope of the energy flux at infinity (dotted black line) begins to become relevant: for some modes (e.g. $m = 3$, blue colored) radiative modes become dominant on superradiance, creating a sharp “dip feature”. The quick dominance of radiative modes is due to the steep functional dependence of k_+ from p in correspondence

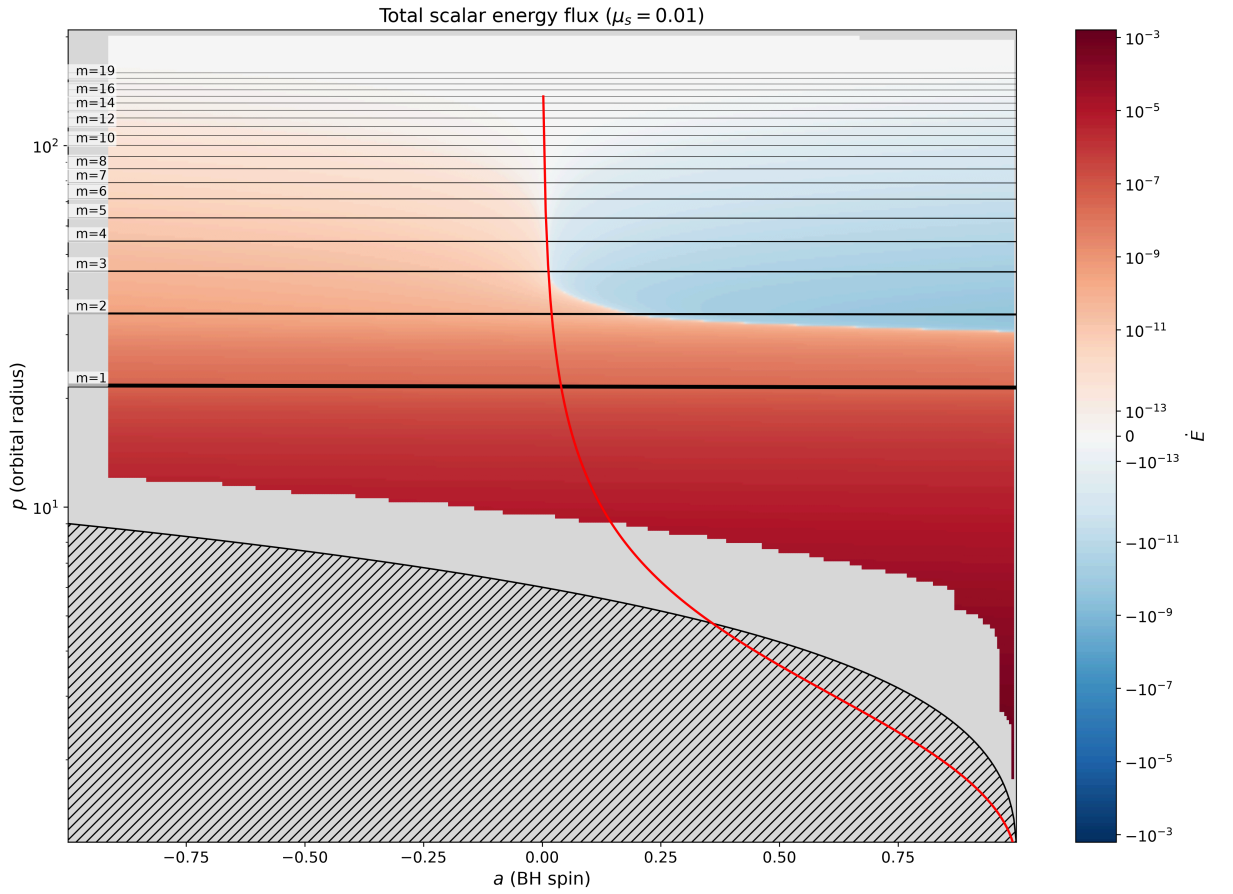
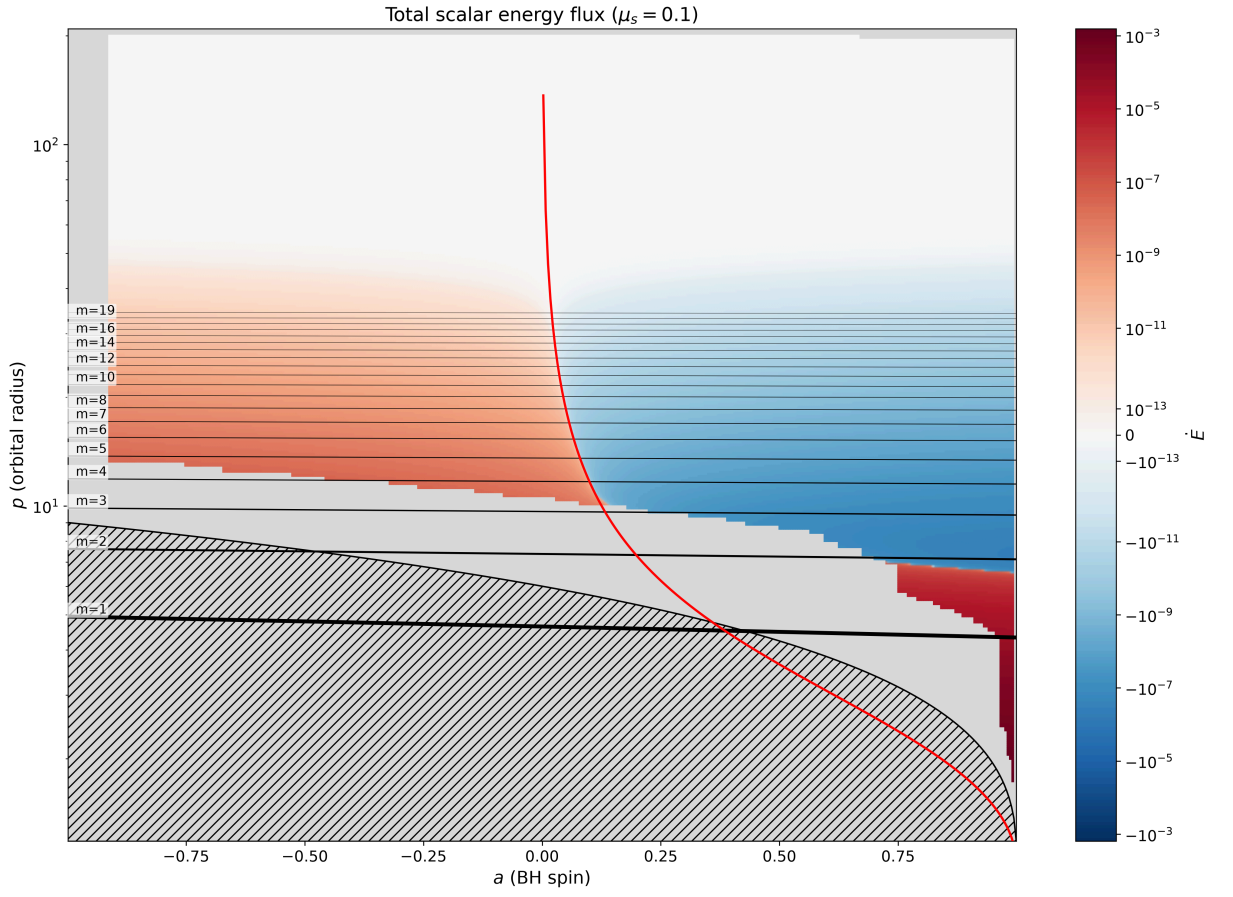
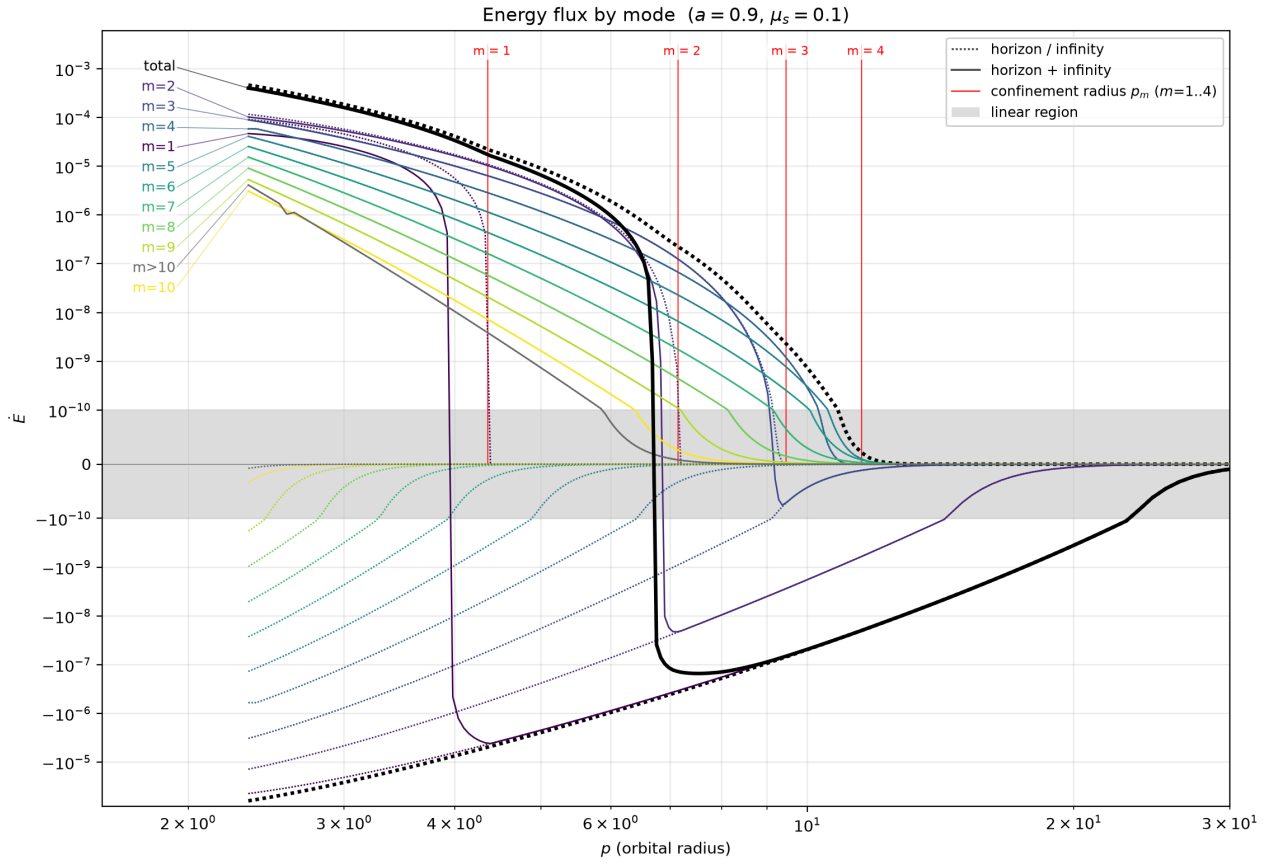
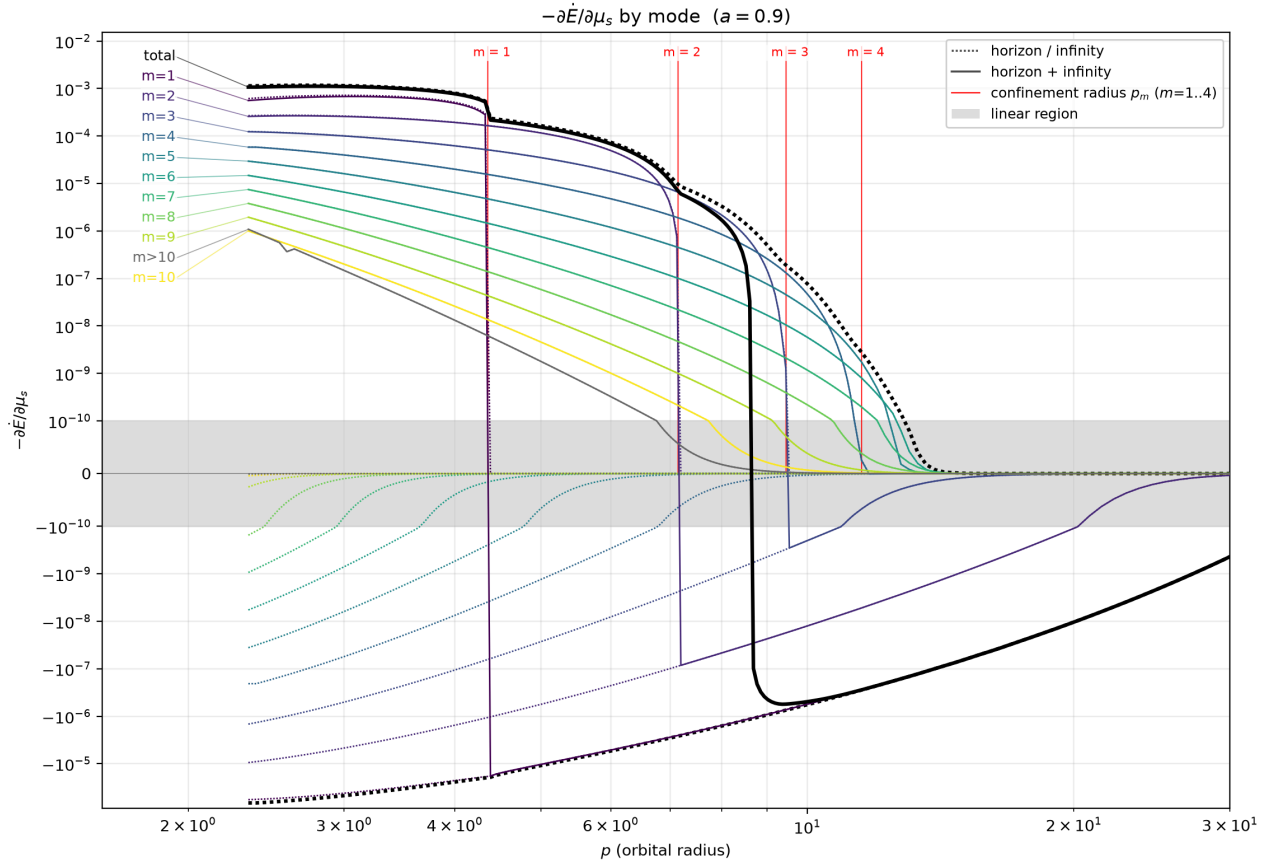


Figure 5.4: Energy fluxes structure interpreted with superradiance (red curve) and mode confinement (black lines, with index $m = 1, \dots, 20$), as discussed in Sec. 5.3.



(a) Energy flux dependence, mode-by-mode, at horizon/infinity and total.



(b) Derivative of energy flux with respect to μ_s , mode-by-mode, at horizon/infinity and total.

Figure 5.5: Mode comparison, for fixed a and μ_s , of $\dot{E}_s(p)$ (5.5a) and $-\partial\dot{E}_s/\partial\mu_s$ (5.5b) at the horizon (dotted lines on the negative semiplane) at infinity (dotted lines on the positive semiplane), and the sum of the two (solid lines), for each value of $m = 1, \dots, 8$ (summed over ℓ , colored), and for the remaining modes $m > 8$ (gray). The total (horizon/infinity/total) is plotted in black. The y-axis is logarithmic until $\pm 10^{-10}$, and linear around 0 (shaded in gray). Mode confinement radii (Eq. 5.4) for $m = 1, 2, 3, 4$ are shown with vertical red lines.

to the confinement radius. As observed in the previous Section, the relevant contribution to pass from globally superradiant to globally radiative behavior is the appearance of the $\ell = m = 2$ radiative mode at $p \approx 7$: the global energy flux detaches from the $\ell = m = 1$ mode, dominant until now, to follow more closely the $m = 2$ line. The net flux $\dot{E}^{\text{tot}} = 0$ is not immediately attained, since the radiative mode must increase enough to compensate for the sum of the strongly superradiant modes $m = 1$ and $m = 2$ at the horizon: it is thus explained the small superradiant strip still visible in Fig. 5.4a and 5.4b immediately below the confinement radius for $m = 2$. At low radii $p \lesssim 6$, even though the prominent mode is still $m = 2$, as observed in Sec. 5.2.a the spectrum becomes very rich in the sense that the higher- m modes become relevant, acquiring magnitudes that are comparable to the low- m modes: already at $p = 5$, the total flux is clearly distinct from the flux for $m = 2$. As a last point-like feature, the appearance of the radiative $m = 1$ mode causes a small bump in the total flux, without however becoming dominant again. At the smallest sampled radius, $p = 2.34$, not only the plotted modes have similar magnitude, but for at least $\tilde{m} < 10$, the sum of all the modes with $m > \tilde{m}$ produces a larger energy flux than the single mode $\tilde{m}$. In addition, at low radii the flux at horizon, despite continuing to grow in modulus, is essentially irrelevant on the total flux, which is dominated by radiative modes.

An analog description can be made for the derivative of the energy flux with respect to the scalar mass: in Fig. 5.5b is plotted $-\partial \dot{E}_s / \partial \mu_s$, with the same conventions as in the previously seen Fig. 5.5a. The choice of the $-$ sign for visualization purposes is to better relate these results with the simple energy flux $\dot{E}_s$: the inverted sign allows to have once again the modes at the horizon in the lower semiplane, the modes at infinity in the upper semiplane and the total sum that follows the same qualitative features of the $\dot{E}_s^{\text{tot}}$ curve. It is here evident the “step behavior” of the derivative, that mode-by-mode is completely dominated by the radiative component at spatial infinity, as argued in the previous Section: for every mode, the total curve is overlapped to the mode at infinity, and abruptly changes to the superradiant mode only when the radius is larger than the confinement radius. However, the modes are not equivalent, and the hierarchy between them forces the total curve to coincide with the dominant $\ell = m = 1$ mode for large p , at which the mode is purely superradiant. In this case, for small p the mode hierarchy continues to follow the m -order, since the prevalent functional dependence is

given by k_+ and $\partial k_+ / \partial \mu_s = -\sqrt{\frac{m^2 \Omega_\phi^2}{\mu_s^2} - 1}^{-1} \approx -\frac{\mu_s}{m \Omega_\phi}$ for $\omega \gg \mu_s$, which means that $\frac{\partial}{\partial \mu_s} \dot{E}_{\ell m}^+ \approx -\mu_s |Z_{\ell m}^+|^2$ follows the mode hierarchy of the square amplitudes $|Z_{\ell m}^+|^2$ (which is m -suppressed compared to the energy hierarchy by a factor $1/m^2$, see Eq. 3.8).

5.4 Results: dephasing

5.4.a Definition and implementation

The detection of EMRIs and the characterization of the orbits will be performed by next-generation GW detectors; in order to assess the detectability of a massive scalar field and our ability to measure its physical properties (the scalar mass μ_s and the scalar charge of the secondary), the computation of the energy fluxes is not directly indicative. It is necessary to evaluate the effects of energy dissipation on the dynamics of the system,

which is then characterized by the GW signal, from now on called waveform to distinguish the main *dynamical evolution* mechanism (dissipation of energy through emission of GW radiation, dominant with respect to scalar field dissipation) from the *detection* signal (the same GW radiation that arrives at the detector and can be measured): despite being the same phenomenon, they are treated with different methods and “serve” different scopes in our treatise.

A first-order assessment of the scalar field energy dissipation on the observable waveform is the waveform *dephasing* $\Delta\psi$, defined as

$$\Delta\psi(T) = 2 \int_0^T \Delta\Omega_\phi(t) dt \quad (5.5)$$

where the orbital angular velocity difference is $\Delta\Omega_\phi = \Omega_\phi^{\text{GR} + \text{scalar}}(t) - \Omega_\phi^{\text{GR}}(t)$, given identical initial orbital configurations at $t = 0$. The multiplicative factor 2 is due to the fact that GW radiation, which is quadrupolar, is produced at a frequency that is double the orbital frequency. This means that the integration of $2\Omega_\phi$ actually gives the GW dephasing, while the simple integration of Ω_ϕ would return the *orbital dephasing*, which observationally is not relevant. Note that the phase difference is given by the dynamical evolution in the first-order adiabatic approximation, not by a modified structure of geodesics, thus it is necessary to evolve the system using the previously computed energy fluxes. The configuration of a circular equatorial orbit is uniquely determined by the orbital energy E (see Sec. 1.4.a) which will evolve following $\dot{E} = -\dot{E}_g - \dot{E}_s$; the energy fluxes associated to GW dissipation are available through the python package Fast EMRI Waveforms (FEW, Chua et al. (2021), Katz et al. (2021), Speri et al. (2024), and Chapman-Bird et al. (2025)). The package provides an integrator that for circular equatorial orbits integrates

$$\dot{p} = \frac{dE}{dt} \left(\frac{dE}{dp} \right)^{-1} = \dot{E} \left[\frac{8ap^{1/2} + p(p-6) - 3a^2}{2(2ap + (p-3)p^{3/2})(2ap^{3/2} + (p-3)p^2)^{1/2}} \right]^{-1} \quad (5.6)$$

thus, the inclusion of the scalar energy flux can be performed by adding the term $\dot{p}_s = -\dot{E}_s \left(\frac{dE}{dp} \right)^{-1}$ to the integration variable $\dot{p}$ in the integrator. This is implemented by a Python class `KerrCircEqFluxScalarMassive` built upon the `KerrCircEqFluxScalar` class present in the additional package Modified EMRI Waveforms (MEW, (Gliorio et al., 2026), originally written to interface the scalar flux produced by STORM with the robust implementation of EMRI dynamics in FEW).

The numerical results for the dephasing over a 4-year evolution period with primary mass $M = 10^5 M_\odot$, secondary mass $m_{\text{sec}} = 10 M_\odot$, scalar charge $d = 0.01$, and scalar masses $\mu_s = 0.1, 0.01$ are plotted respectively in Figs. 5.6a and 5.6b. The plots show the dephasing accumulated in 4 years, namely the nominal mission duration for LISA, for a system that at the end of its evolution has reached the orbital radius shown in the figure (thus, that started the evolution at a suitably larger p). The net superradiant zone, obtained directly from the interpolated energy fluxes (negative $\dot{E}$ in Fig. 5.4) is hatched for reference in Fig. 5.6a. Note that the sign of the energy flux directly determines the dephasing sign only for large p , where the frequency evolution is negligible; at low radial coordinate (e.g. lower right zone of the $\mu_s = 0.1$ dephasing plot). In the $\mu_s = 0.01$ plot, the negative dephasing corresponds nearly exactly to the superradiant zone, since the system

is nearly monochromatic throughout the observational period; it is also still possible to see the step structure inherited from the mode confinement mechanism acting on the energy flux at spatial infinity.

5.4.b Observational thresholds of the dephasing

The modified dynamics will thus produce a dephasing that changes with time in the GW waveform (note that $\Delta\Omega_\phi$ is not positive- or negative- defined over the whole range of orbital evolution; thus, in general the dephasing is not strictly monotonically increasing/decreasing with time). We are interested in computing an indicative dephasing threshold over which the modified waveform can be unambiguously distinguished from the GR waveform. To do so, I adopt the formalism presented in Benetti et al. (2026a) to assess the measurability of little waveform discrepancies; while in Benetti et al. (2026a) the formalism is applied to unwanted waveform systematics, here the discrepancy is the goal of the measurement.

Consider a GR waveform $h = h(\mu_s = 0)$, and the modified gravity waveform $h^* = h + \delta h$, where δh is a small deviation caused by the scalar flux. Note that in this case the deviation will cause exclusively a dephasing, without modifying the GW amplitude, since in both cases the trajectories correspond to Kerr geodesics. The difference between the waveforms can be globally quantified with the *faithfulness* (Apostolatos, 1995)

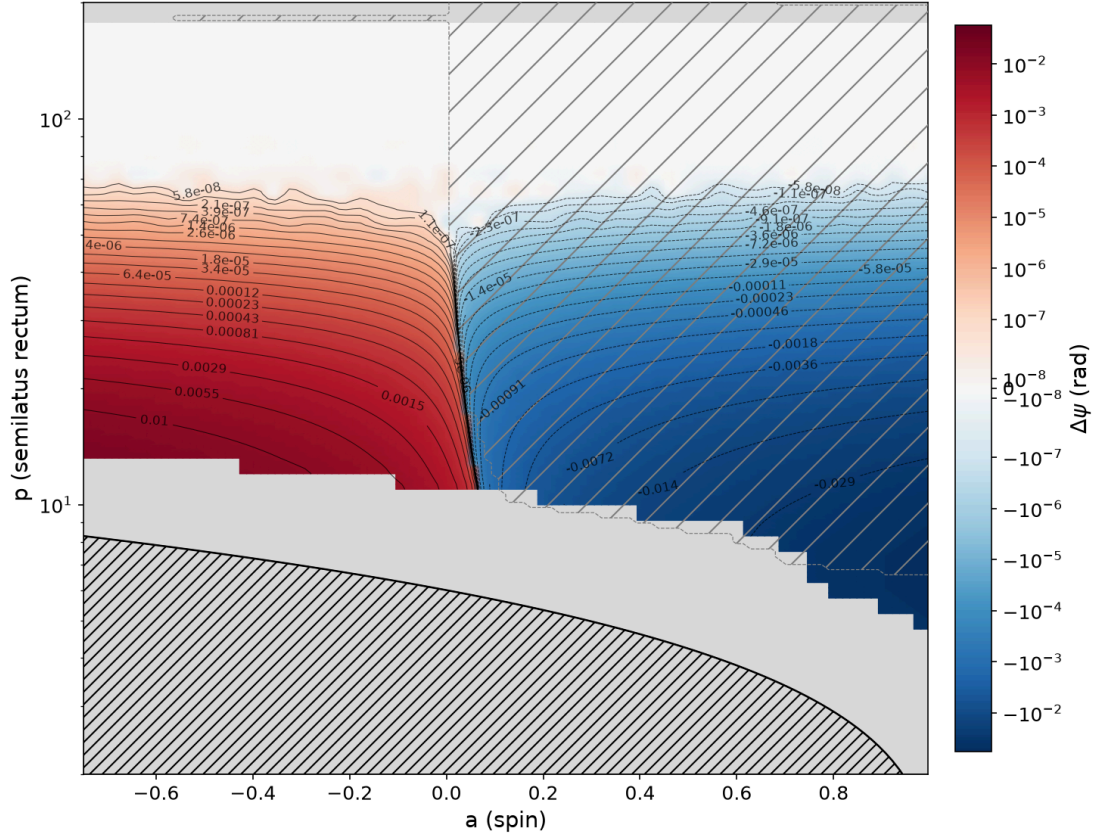
$$\mathcal{F} = \max_{\Lambda} \frac{(h | h^*)}{\sqrt{(h | h)(h^* | h^*)}}, \quad (5.7)$$

where the maximization is performed over a subset of the parameters that describe the system. In our case, adopting the modeller's perspective described in Benetti et al. (2026a), the extremization set corresponds to the extrinsic parameters, thus the faithfulness is a function of the intrinsic parameters only, and let D be the number of such parameters. Expanding at linear order the modified waveform around the GR waveform (this operation is valid in the high-SNR limit, as it is done in the Fisher matrix approach) we obtain $h^* = h + \Delta\lambda^i \partial_i h + \dots$, where $\Delta\lambda^i$ is a small variation of the parameters; in this case, since we want to distinguish the modified waveform (scalar charge $d \neq 0$ but nevertheless small) from the GR waveform ($d = 0$), the $\Delta\lambda^i$ will be exactly the scalar charge. There is however a problem: the discrepancy caused by the presence of a scalar charge could be masked by simultaneous variation (biasing) of the other intrinsic parameters: this behavior is well described by the Cutler-Vallisneri formula

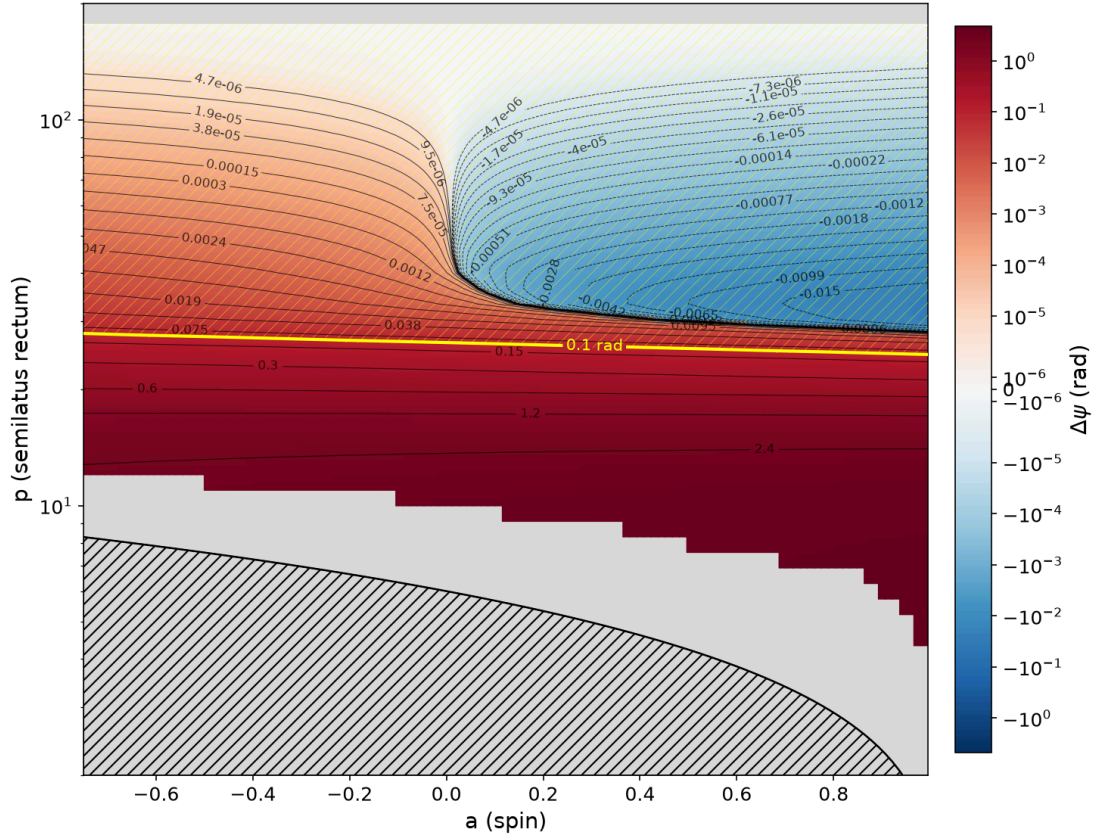
$$\delta\lambda^i = \Gamma^{ij-1}(\partial_j h | \delta h) \quad (5.8)$$

which is found in the linear signal approximation within the Fisher matrix formalism. From this, it can be seen that a waveform discrepancy caused by the variation of a single parameter can, in fact, affect the measurement of the others, propagating through the covariance matrix Γ^{ij-1} . For this reason, a global estimation of the induced bias is necessary. Following the approach of Chatziioannou et al. (2017), the faithfulness can be expanded at the first relevant order (the second) to

$$F \approx 1 - \frac{1}{2} g_{ij} \Delta\lambda^i \Delta\lambda^j + \dots \quad (5.9)$$



(a) Scalar mass $\mu_s = 0.1$. The coarse-hatched region corresponds to the globally superradiant regime $\dot{E}_s^{\text{tot}} < 0$.



(b) Scalar mass $\mu_s = 0.01$. The hatched region corresponds to a dephasing $\Delta\psi < 0.1$ rad.

Figure 5.6: GW dephasing, obtained as the difference between purely-GR phase and GR + massive scalar phase, for $\mu_s = 0.1$ (upper figure, 5.6a) and $\mu_s = 0.01$ (lower figure, 5.6b). The other parameters are $T = 4$ yrs, $M = 10^5 M_\odot$, $m_s = 10 M_\odot$, $a = 0.9$, $d = 0.01$.

where

$$g_{ij} = \frac{(\partial_i h | \partial_j h)(h | h) - (h | \partial_i h)(h | \partial_j h)}{(h | h)^2} \quad (5.10)$$

Let's introduce the *unfaithfulness* $\bar{F} = 1 - F$. Then the expectancy value of $\bar{F}$ can be computed as:

$$\begin{aligned} \mathbb{E}[\bar{F}] &= \mathbb{E} \left[\frac{1}{2} \frac{\mathcal{F}_{ij} \rho^2 - (h | \partial_i h)(h | \partial_j h)}{\rho^4} \Delta \lambda^i \Delta \lambda^j \right] \\ &= \frac{1}{2} \frac{D \rho^2 - \rho^2}{\rho^4} \\ &= \frac{D - 1}{2 \rho^2} \end{aligned}$$

where I used the definition of the SNR $\rho^2 = (h | h)$, the Fisher matrix $\mathcal{F}_{ij} = (\partial_i h | \partial_j h)$, the expectation value for the product of the parameter biases $\mathbb{E}[\Delta \lambda^i \Delta \lambda^j] = C^{ij} = \mathcal{F}_{ij}^{-1}$ (such that in the first term $\mathcal{F}_{ij} C^{ij} = \delta_j^j = D$, since we are considering only intrinsic parameters following the modeller's perspective). This corresponds to the geometric interpretation of the indistinguishability criterion provided in Eq. A4 of Benetti et al. (2026a). Finally, we write the unfaithfulness in terms of phase and amplitude discrepancies, $\delta\phi$ and δA ; we can expand $h^* = h + \delta h = h e^{\delta A + i\delta\phi} \approx h(\delta A + i\delta\phi)$, so that $(\delta h | \delta h) \approx \rho^2(\delta\phi^2 + \delta A^2)$. Comparing this approximation with the unfaithfulness definition, to practically distinguish a *physical difference* between waveforms from a *statistical fluctuation* of the unfaithfulness, due to a particular realization of the noise, it is required that

$$\delta\phi^2 + \delta A^2 > \frac{D - 1}{2 \rho^2} \quad (5.11)$$

in our case as I mentioned above $\delta A = 0$, thus this translates to a simple **distinguishability condition** regarding exclusively the dephasing:

$$\Delta\psi(T) = \delta\phi > \frac{\sqrt{D - 1}}{\rho\sqrt{2}} \quad (5.12)$$

Let's apply this condition to the EMRI system previously parametrized: since we are considering circular equatorial orbits, the parameters of the system are the primary mass M , the secondary mass m_p , the BH spin a , initial orbital radius p_i , initial phase ϕ_i , the scalar charge d and scalar mass μ_s for a total of 7 parameters; to be rigorous, this is a specific configuration of zero measure in the parameter space of generic orbits, thus in principle also orbital eccentricity and inclination should be added, taking the count to 9 parameters. In addition, since the final goal is not to assess the *intrinsic* difference between waveforms but the *detectability* of small parameter variations, the correct diagnostic would also include the extrinsic parameters, in the so-called "data-analyst perspective" presented in Benetti et al. (2026a), taking the total to $D = 15$ (adding sky location, luminosity distance, inclination and polarization angles, time of arrival). Since the square root weakens the dependence on the number of parameters for large D , let's adopt the conservative criterion with $D = 15$ and an average SNR of $\rho \approx 30$. This results in a dephasing threshold of $\delta\phi > 0.088 \text{ rad} \sim 0.1 \text{ rad}$. Note that this is an indicative and conservative estimate for the minimum dephasing needed for the detection of the scalar field, and also that SNR of the

Variable	Reference (run 1)	Range (run 1)	Reference (run 2)	Range (run 2)
$M (M_\odot)$	10^5	$[10^5 - 10^7]$	$3 \cdot 10^6$	$[10^5 - 10^7]$
$m_s (M_\odot)$	10	$[1 - 100]$	3	$[1 - 100]$
d	0.01	$[10^{-3} - 1]$	0.01	$[10^{-3} - 1]$
T (yrs)	4	$[1 - 4]$	4	$[0 - 4]$
p/M	2.4	$[2.4 - 20]$	3	$[2.4 - 20]$

Table 5.1: Reference values and range intervals for the intrinsic parameters considered in Sec. 5.4.c.

order of $\sim 10^2$ are expected, leading to lower dephasing thresholds of ~ 0.01 rad or less for a small number of particularly favorable sources.

Following this prescription, in Fig. 5.6b the $\Delta\psi = 0.1$ rad level line is highlighted, and the area with $|\Delta\psi| < 0.1$ rad is hatched. For $\mu_s = 0.1$ (Fig. 5.6a) the modulus of the dephasing is always under this threshold. Note that for low primary spin $a \sim 0$ and large orbital radius, since both the modes at horizon and infinity are suppressed the detectability is generally worse. Observationally, the most convenient cases are extreme-Kerr, for both a approaching ± 1 .

5.4.c dephasing observability

In this Section I will make use of the isolated case study presented in Sec. 5.3.b to assess the observability of the dephasing for a wide range of configurations, taking advantage of the excellent completeness of such a configuration also toward the ISCO. It is indeed clear from Fig. 5.6 that the confinement effect of the scalar mass will strongly suppress the dephasing, leading to observable effects only near the ISCO, where the modes at infinity are radiative. In fact, the evolution through the purely superradiant regime and then across the radiative regime can have the effect of eliding the total dephasing (this effect does not occur for a massless field, for which the radiative modes are always dominant on the superradiant behavior and the dephasing is monotonically increasing. This means that lower masses are needed to distinguish the presence of a generic scalar field, but this could prevent the observability of a scalar mass unambiguously different from 0. Note that a final dephasing which is under the conservative threshold indication $\Delta\psi \sim 0.1$ rad computed in the previous Section does not imply the non-observability of the scalar field effects, since this is only a sufficient condition to claim the observability of a measurable discrepancy from the expected GR waveform. This is especially true for massive scalar fields where, as argued, the dephasing can grow and then diminish under the observability threshold; this behavior could be detected by a complete signal analysis (as the Fisher matrix technique presented in the following), even if the dephasing condition is not met.

Under these premises, I used the case study with fixed $\mu_s = 0.1$, despite this being a disfavorable choice of the scalar mass, as can be seen in Fig. 5.6, to probe the dephasing across different values of the remaining intrinsic parameters (primary mass M , secondary mass m_p , scalar charge d , observational time T , final orbital radius p). A reference value for each of these is fixed while 2 parameters per time are let varying on a chosen range. The reference parameters and the corresponding ranges are listed in Tab. 5.1.

Two runs are performed, considering different reference parameters. The first (Run 1 in Tab. 5.1) adopts the standard parameters that were considered in the previous Sections;

the 2D dephasing distributions resulting from the variations of parameter couples are plotted in Fig. 5.7.

before commenting the plots, let's describe the dependence of the energy fluxes from the intrinsic parameters. As for the massless scalar field, both the scalar and GW energy fluxes scale as q^2 : this follows from the fact that the source terms are proportional to m_s , and the energy fluxes are a sum over the square modulus of the fields' modes. In addition, the scalar flux is proportional to the square of the scalar charge d , which follows from the same argument. The scaling with respect to the total mass of the system (which is M for an EMRI) follows the typical GR rescaling of mass, length and time: since p is the orbital radius normalized to the mass ($p = p_{\text{physical}}/M$), and analogously for the spin ($a = a_{\text{physical}}/M$), and the energy flux depends on the mass ratio q uniquely, the only remaining relevant scaling is the time dependence which is proportional to the total mass: the dephasing at physical time T will thus scale as $1/M$. In the massive scalar field case however the GR global rescaling symmetry is broken, since a *physical reference mass* is introduced, the mass of the scalar field; in light of the two main mechanisms that determine the scalar flux, we can verify that superradiance, being independent from the scalar mass (Eq. 5.3) is not affected; however, massive field confinement is affected by the choice of the scalar mass,

The first plot (p vs T , upper left) shows that, at the chosen reference values, the final radius that provides the best observability of the dephasing is not the innermost one (as it is the case for a massless field), instead it is $p_{\text{opt}} \sim 6$, for which the accumulated dephasing is negative and due to the prevalence of superradiance. Shifting toward smaller radii will actually decrease the dephasing (and thus the observability of a deviation from GR) since the radiative modes will compensate for the effect of the superradiant modes; in particular, the evolution in the radiative regime is too fast to accumulate a net positive dephasing (lower left corner of the plot: a positive dephasing is attained at low radii for a short observational period, over which the initial radius is too deep in the superradiant zone and the net dephasing is negative). The second plot (M vs T , upper right), shows a non-trivial interplay between observational time and the mass of the primary: at smaller M , like the one chosen for reference, the orbital evolution is faster and allows the secondary to span a larger radial interval, averaging radiative and superradiant configurations to a net unobservable dephasing; moving to larger M , the slower evolution allows for a consistent accumulation of dephasing only in the radiative zone (where the reference final configuration lies; a similar argument holds for the superradiant zone, although the net energy flux is smaller with respect to radiative modes, leading to a potentially more difficult detection, even if more informative about the mass of the scalar field). On the other hand, towards larger M the slower evolution will accumulate a smaller dephasing for a quasi-monochromatic system in a fixed physical time interval. As a result of these two competing phenomena, for a certain reference set of parameters, there exists an optimal M such that the dephasing is maximized; for the parameter set of Run 1, this mass corresponds to approximately $10^6 M_{\odot}$, and an observable dephasing becomes visible already after 1 year of observation. Another important observation is that, in general, the time dependence becomes weak after ≈ 2 years of observation (almost vertical level lines).

The middle plot of Fig. 5.7 shows the dephasing dependence from the primary and secondary masses. The primary mass dependence is analogous to what has already been described; at fixed M , the choice of m_s determines the mass ratio q and thus the global

rescaling of both GW and scalar fluxes; thus larger m_s implies a faster evolution and more radiative-superradiant averaging across the field confinement boundary, as it was the case for the evolution time. In this plot, however, it is also possible to note that for the reference primary mass, the dephasing is not monotonically decreasing with increasing secondary mass after reaching a (negative) minimum around $m_s \sim 10 M_\odot$, since with larger secondary mass most of the evolution will happen at even larger radial coordinate, where also superradiant energy fluxes are suppressed (see Fig. 5.3, 5.4), and thus the evolution is completely dominated by standard GR, namely dephasing is not accumulated. A similar behavior is expected for the time dependence, but in this case it cannot be seen due to the limited mission duration (also note that the m_s axis is logarithmic, the time axis is linear).

The last two plots (at the bottom of the figure) show the dephasing dependence with the scalar charge d and the previously considered primary mass M and observational time T . As expected, the modulus of the dephasing is monotonically increasing with d (since it represents a simple ratio between GW and scalar emissions): the level curves are all equal and simply shifted depending on d , since d^2 is a mere multiplicative factor for the scalar energy flux. As a result, features such as net-0 dephasing time (left plot) or primary mass (right plot) or optimal primary mass are independent of the choice of d .

Motivated by these qualitative dependences, a second set of reference values (Run 2, see Tab. 5.1) is evaluated with similar ranges, and presented with the same conventions in Fig. 5.8. In particular, I choose a smaller secondary mass ($m_s = 3 M_\odot$, compatible with a neutron star as the secondary) and a larger primary ($M = 3 \cdot 10^6 M_\odot$) to achieve a slower evolution in the radiative zone and a potentially observable dephasing. The qualitative explanation of the dephasing plots is analogous to the one provided for Run 1; the notable difference is that with such low q the system is practically monochromatic even near the smallest radius considered, leading to a time-independent separation of radiative and superradiant behavior in the p vs T plot (upper left) and thus the absence of an initially superradiant phase (for example, in the T vs d plot, lower left). Note that with such a small secondary mass, the optimal primary mass for detection is actually located at a much smaller $M_{\text{opt}} \sim 5 \cdot 10^5 M_\odot$, thus nearly 1 order of magnitude less than the considered primary mass.

5.5 Results: Fisher matrix forecasts for parameter estimation

The most informative diagnostic of the observability of a parametrized phenomenon is the Fisher matrix. With such long signals, a full Bayesian analysis is prohibitively computationally expensive with current tools; however, in the high-SNR limit the Fisher approximation provides an excellent representation of the posterior distribution for the parameter set used to describe the system, at a reasonable computational cost and under the theoretical limitations presented in Sec. 1.2.c. This set of parameter consists of the primary mass $m_1 = M$, secondary mass $m_2 = m_{\text{sec}}$, the spin parameter a , initial orbital radius p_0 and orbital phase $\Phi_{\phi 0}$, source luminosity distance D_L , the sky location polar and azimuthal angles q_s, ϕ_s , the spin vector orientation polar and azimuthal angles q_k, ϕ_k , and the additional parameters describing the scalar field, namely the scalar charge d and the scalar mass μ_s . Note that since $h \propto D_L^{-1}$ and $\mathcal{F}_{ij} = (\partial_i h | \partial_j h)$, the matrix elements are inversely proportional to the square of the luminosity distance; it follows that the covariance matrix

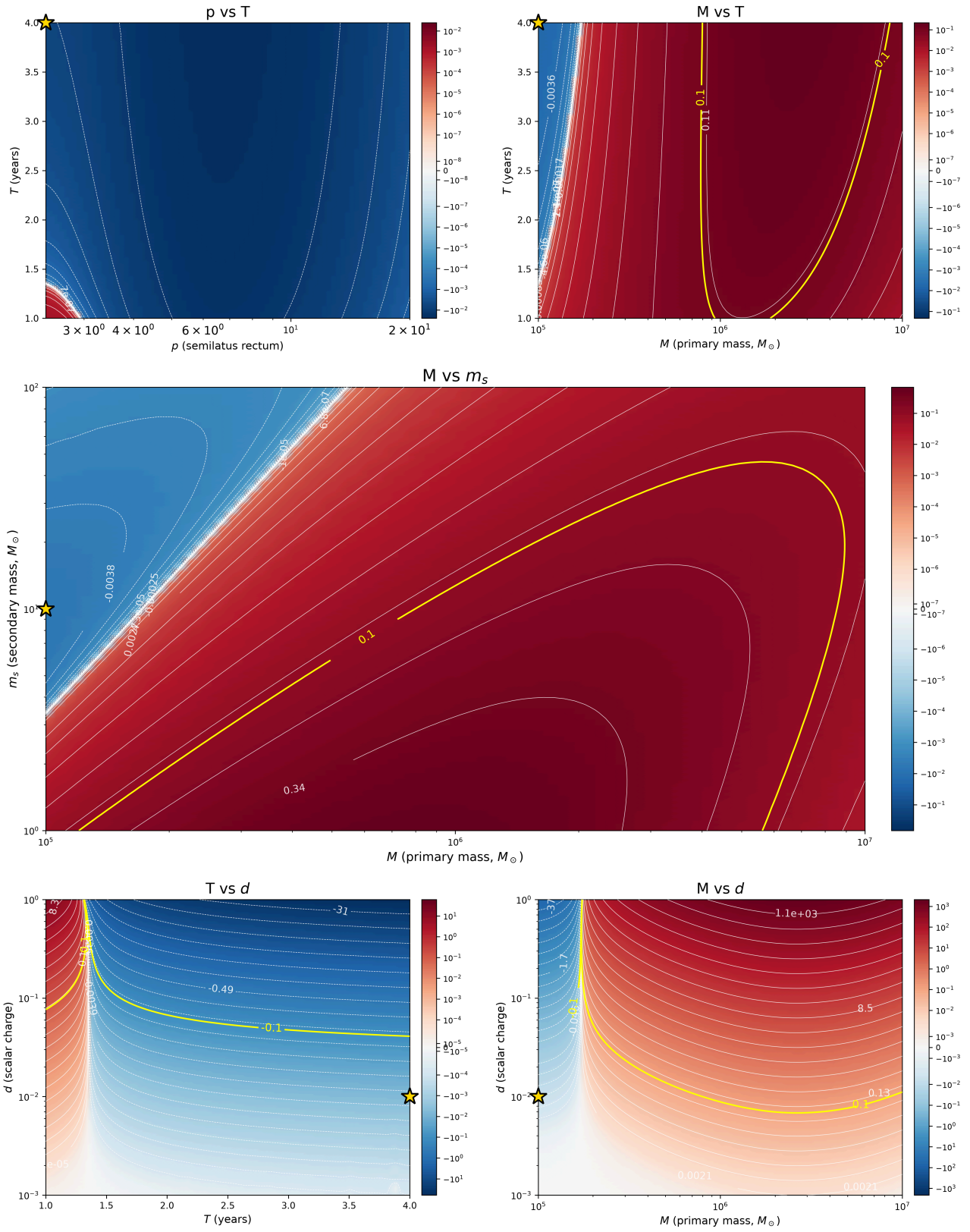


Figure 5.7: Run 1 for the dephasing dependence on the intrinsic parameters M , m_s , d , T , p . Two parameters are varied on a range while the other are fixed to the reference values (see Tab. 5.1); the reference values are here marked with a star. All the colorbars refer to the dephasing $\Delta\psi$ in radians.

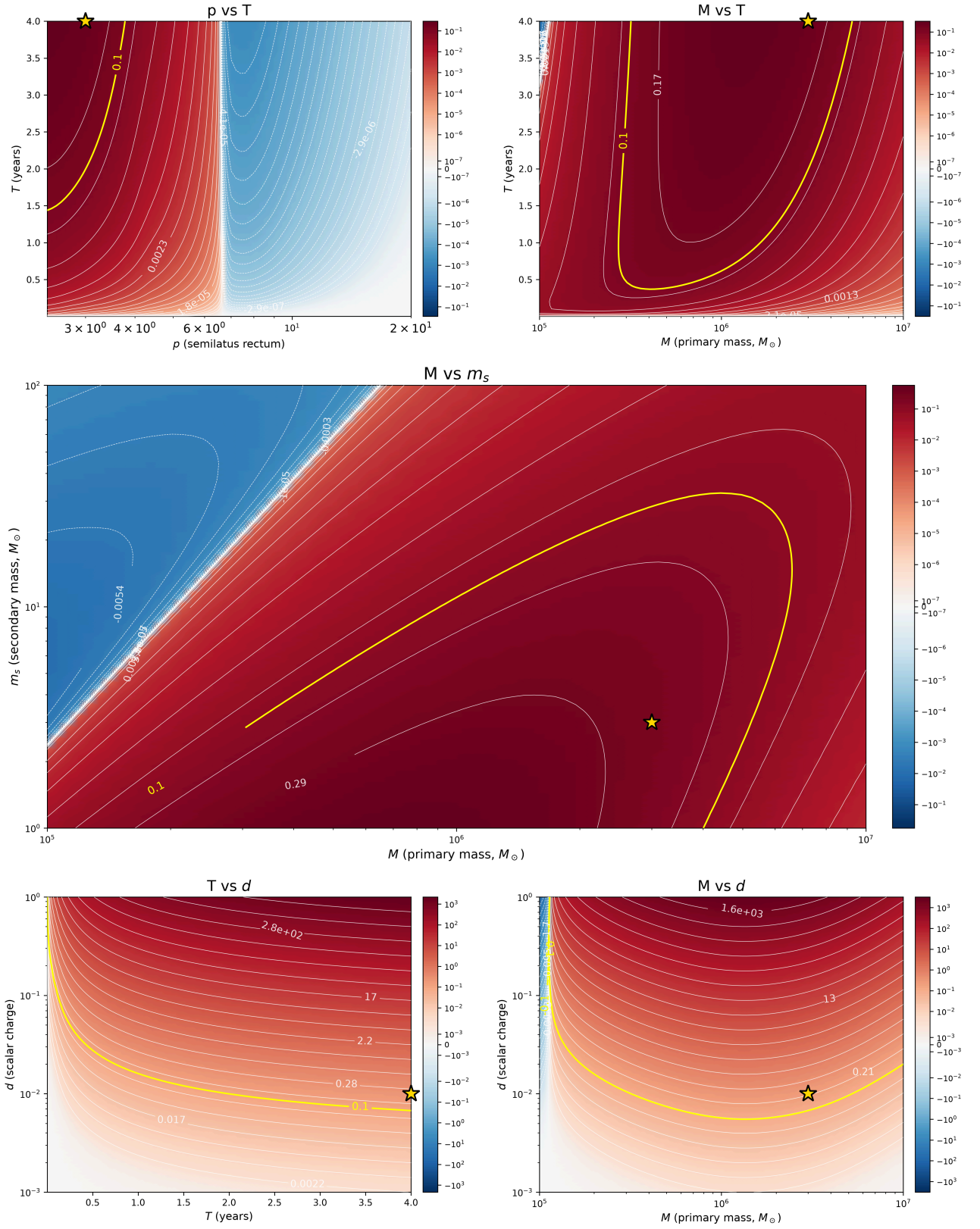


Figure 5.8: Run 2 for the dephasing dependence on the intrinsic parameters M , m_s , d , T , p (see Tab. 5.1). Refer to Fig. 5.7

$\mathcal{C}_{ij} = \mathcal{F}_{ij}^{-1}$ is directly proportional to D_L^2 , or more informatively inversely proportional to the squared SNR ρ , since $\rho^2 \equiv (h|h) \propto D_L^{-2}$.

The computation of the Fisher matrix is performed by the Python package STABLEMRIFISHER (Kejriwal et al., 2026), which computes stable waveform derivative with a finite-difference approach, automatically selecting an optimal stencil to ensure robustness of the results and invertibility of the Fisher matrix; this is a general problem for Fisher matrix computation, where the eigenvalues can span several orders of magnitude, thus being easily subject to numerical instability. The Fisher matrix computation utilizes the waveform generators from FEW, previously generalized to include the contribution from scalar field energy fluxes through the custom class `KerrCircEqFluxScalarMassive` previously introduced. The STABLEMRIFISHER is used without GPU acceleration and with sky-average sensitivity, thus not directly implementing the time-changing antenna pattern and response functions of LISA. This choice is motivated by the fact that the main interest is to assess detectability of *intrinsic parameters* for an “average source” without introducing sky-location dependencies. Furthermore, I verified that the Fisher matrix is block-diagonal with two main blocks, one corresponding to the angles q_s, ϕ_s, q_k, ϕ_k , which, as expected, are decoupled from the rest of the parameters. For this reason (and not considering an antenna response, which leads to unphysically poorly constrained orientation angles), these parameters are excluded from the Fisher matrix computation. Note that this is a physically (parameter independence), numerically (the cross-parameter correlations are verified to be small), and statistically (since the covariance is block-diagonal, the marginalized posteriors of the parameters left are unchanged) well motivated restriction, which will not affect the rest of the estimations.

I start by using the unidimensional case study grid presented in Sec. 5.3.b, since its completeness facilitates stability especially regarding the scalar mass parameter (as we saw in Fig. 5.3, incompleteness can cause artifacts in the mass derivative plot). In the following, I will present and comment different configurations.

Case study 1

Motivated by the observable dephasing observed for the Run 2 configuration (see Tab. 5.1), I start with a system defined by $M = 3 \cdot 10^6 M_\odot$, $m_{\text{sec}} = 3 M_\odot$, $p_{\text{final}} = 4$, $T_{\text{obs}} = 2$ yrs, with SNR $\rho = 24$. Note that the corresponding dephasing is below the threshold of 0.1 rad; however, this is only an explorative run with a radius that allows stability of the Fisher matrix (smaller radii were producing instabilities), and a shorter signal duration reduces the computational time. The resulting Fisher and covariance matrices are shown in Fig. 5.9; the values that are plotted are normalized as $\tilde{\mathcal{M}}_{ij} = \mathcal{M}_{ij} / \sqrt{\mathcal{M}_{ii}\mathcal{M}_{jj}}$, where $\mathcal{M}$ is either the Fisher matrix $\mathcal{F}$ or the covariance matrix $\mathcal{C} = \mathcal{F}^{-1}$, in such a way that the covariance is normalized to the standard deviation of the respective parameters; this allows an immediate comparison of the degeneracy between parameters, easily found by searching for off-diagonal elements $\tilde{\mathcal{C}}_{ij} \sim \pm 1$.

The figure highlights a crucial feature of the system, namely the strong degeneracy between multiple parameters; this follows from considering quasi-monochromatic circular equatorial orbits. A circular equatorial orbit is defined exclusively by the orbital frequency Ω_ϕ , which is a function of M , p_0 and a simultaneously: this means that for a monochromatic source these 3 parameters are completely degenerate. The parameter evolution can help to break the degeneracy; however, the setup was deliberately chosen for its limited

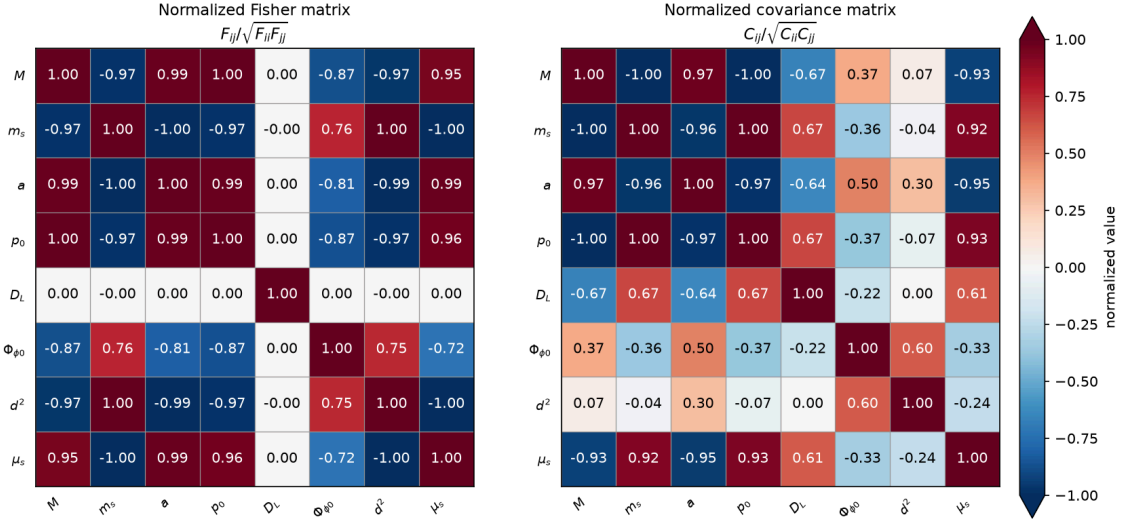


Figure 5.9: Normalized Fisher and covariance matrices for $M = 3 \cdot 10^6 M_\odot$, $m_{\text{sec}} = 3 M_\odot$, $p_{\text{final}} = 4$, $T_{\text{obs}} = 2 \text{ yrs}$, $d = 0.01$, $\mu_s = 0.1$. $p_i \approx 4.5$ follows from the chosen parameters.

evolution, which led to a more substantial accumulation of dephasing: the initial radius was only $p_i \approx 4.5$, to be compared with the final radius $p_f = 4$. This means that only a small portion of the orbital dynamics is observed, which breaks the degeneracy only slightly: all the matrix elements in the upper-left submatrix 4×4 are all very close to 1. In addition, the classical EMRI intrinsic parameters in the monochromatic regime are in general degenerate with the scalar mass: at linear order, the scalar flux can be linearly expanded around the true value of μ_s ; a small shift in μ_s will cause a proportionally small change of the energy flux. However, the same operation can be performed on the GW energy flux, which depends only on the orbital parameters; then, at linear order a small shift in *any* of the intrinsic parameters of the system is equivalent to a certain shift in μ_s . This degeneracy can be broken only by higher-order derivatives in the local expansion, and in particular as many as the considered parameters. In other words, *to disentangle parameter degeneracy, it is necessary to probe a finite (large) radial interval*. This goes in the opposite direction of the dephasing criterion, since exploring a large radial interval implies averaging over superradiance and radiative behavior; note that the dephasing criterion is not incorrect: it simply assumes that the other intrinsic parameters are correctly estimated, which becomes operationally impossible when a complete degeneracy is present. It is thus likely that there will be an optimal trade-off between the two tendencies, between fast-evolving systems that cannot show a relevant phase mismatch and slowly-evolving systems that are completely degenerate in the parameters. In particular, for the considered system the *marginalized* 1σ statistical uncertainties are far from sufficient to distinguish either μ_s or d^2 from 0 (null hypothesis, GR), with estimated $\sigma_{\mu_s} \sim 25$ completely losing physical meaning.

Note that, independently from the scalar field, adding eccentricity and inclination to the dynamics will break the degeneracy between M, m_s, a, p_0 (at the cost of adding the additional parameters e and θ_{inc}). In this sense, the circular equatorial case is more problematic since its symmetry directly causes degeneracy.

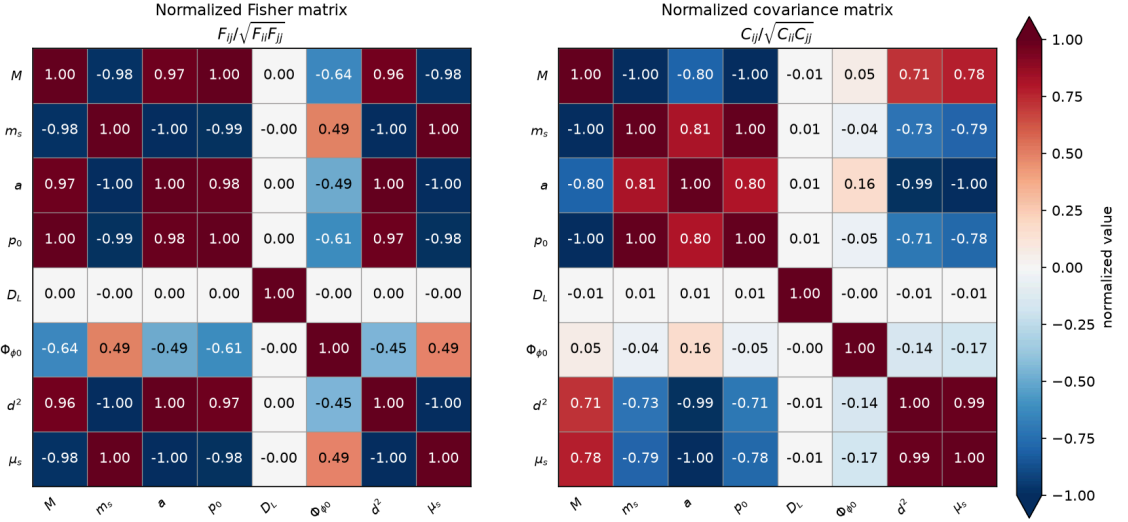


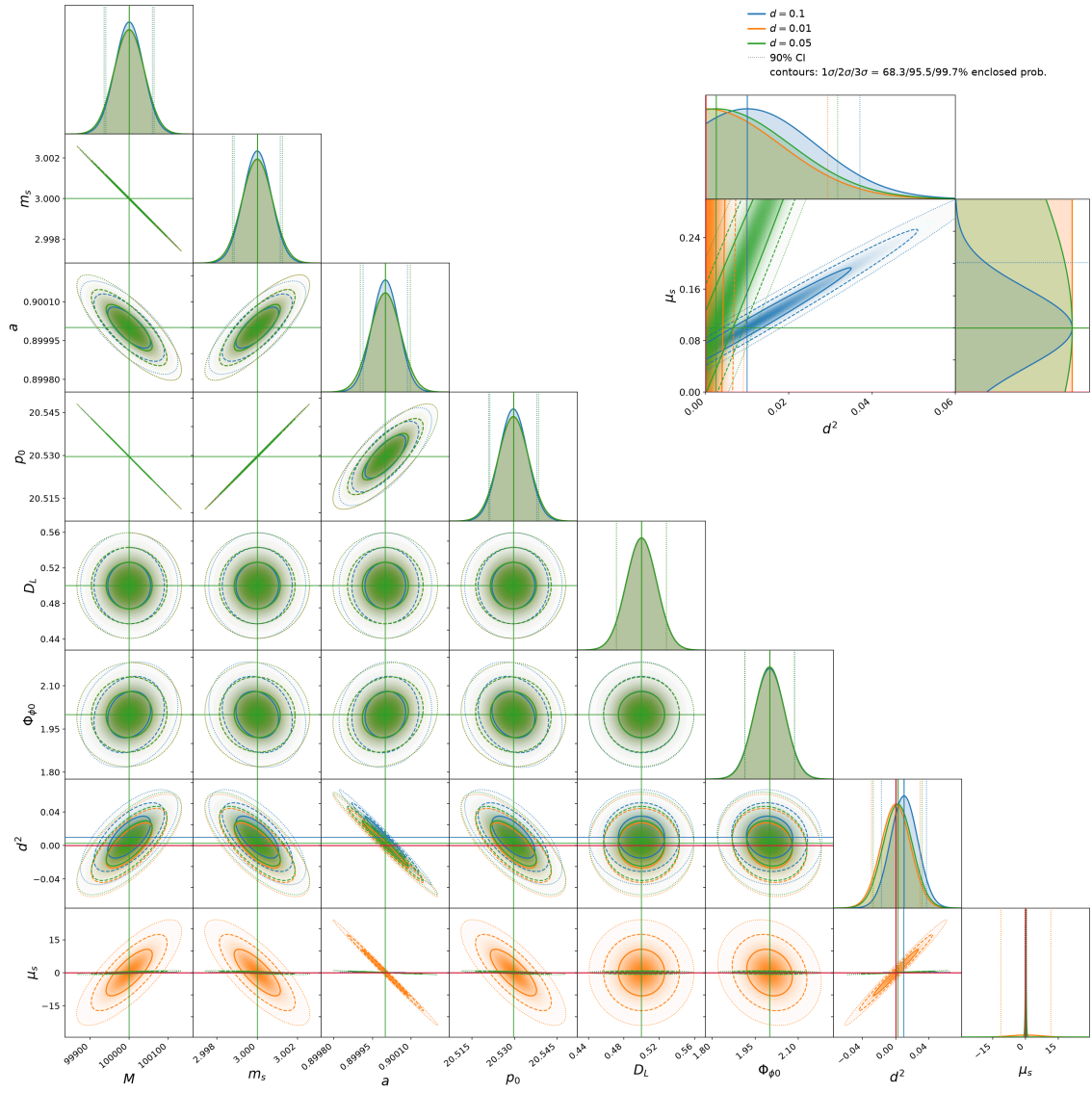
Figure 5.10: Normalized Fisher and covariance matrices for $M = 10^5 M_\odot$, $m_{\text{sec}} = 3 M_\odot$, $p_{\text{final}} = 4$, $T_{\text{obs}} = 2 \text{ yrs}$, $d = 0.01$, $\mu_s = 0.1$. $p_i \approx 20$ follows from the chosen parameters.

Case study 2

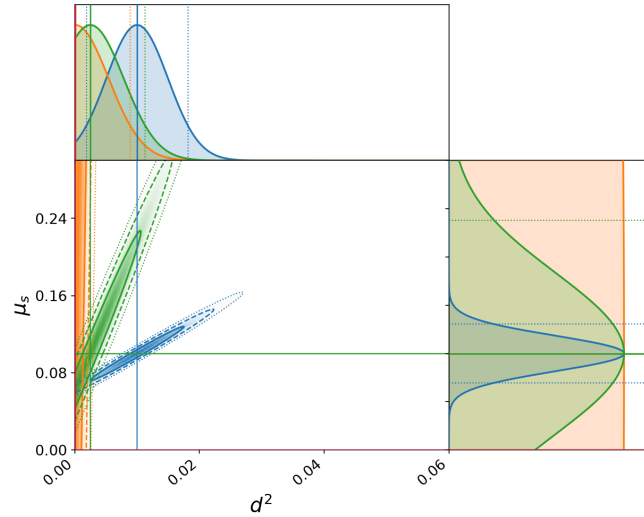
In order to avoid degeneracy, I then consider a system that covers a wider radial range, with parameters $p_0 = 4$, $M = 10^5 M_\odot$, $m_s = 3 M_\odot$. The other parameters are the reference ones, and the luminosity distance has been chosen to have $\text{SNR } \rho = 30$ or 100 with a observational period of 2 years. With this set of parameters, the initial orbital radius is $p_i = 20$: from Fig. 5.5a we can see that the orbit spans the superradiant region, the field confinement boundaries for $m \geq 2$ and lands in the radiative zone just before the $m = 1$ radiative mode becomes active. Thus it is expected that all this structure can help to break the degeneracies between EMRI parameters and the scalar field mass μ_s .

In Fig. 5.10 the Fisher and covariance matrices for a scalar charge $d = 0.01$ are graphically reported; the qualitative features of the normalized Fisher matrix do not substantially change with the scalar charge. In this case, a strong covariance between scalar mass and scalar charge is observed, hinting at the fact that radiative/superradiant dephasing cancellation in different parts of the orbit, which are controlled by the scalar mass, have a comparable effect with a global rescaling of the total flux through d^2 . In particular, increasing the scalar mass μ_s will damp the flux by reducing the innermost radiative zone, thus the waveform will be “minimally changed” if this is accompanied by a larger scalar charge which compensates for the missing radiative zone, resulting in a strong positive correlation.

Thus also in this case the system is not very informative: despite spanning a wide radial interval, the majority of the time the secondary orbits in a very weakly superradiant zone, and the transition between superradiant and radiative behavior happens too fast to produce a strong signature in the waveform. Keeping the other parameters fixed, I performed 3 runs for $d_1 = 0.01$ (shown in Fig. 5.10), $d_s = 0.05$ and $d_3 = 0.1$. The corresponding corner plots of the approximated gaussian posterior distributions are shown in Fig. 5.11 for the two choices of SNR (Fig. 5.11a and 5.11b for SNR 30 and 100 respectively). It is evident the importance of the scalar charge d which, entering the scalar flux as d^2 , if increased has the double effect of reducing the marginalized uncertainty on μ_s (but, note,



(a) Corner plot for SNR $\rho = 30$



(b) Sub-corner plot for SNR $\rho = 100$

Figure 5.11: Corner plot of the Gaussian posterior distributions for scalar charge $d = 0.01$ (orange), $d = 0.05$ (green), $d = 0.1$ (blue) and $M = 10^5 M_\odot$, $m_{\text{sec}} = 3 M_\odot$, $p_{\text{final}} = 4$, $T_{\text{obs}} = 2$ yrs, $\mu_s = 0.1$. The plot shows the 1, 2, 3 σ contours, and the null hypothesis $d = 0$, $\mu_s = 0$ for the scalar field parameters. Upper plot SNR $\rho = 30$, lower plot SNR $\rho = 100$.

without relevant effects on the width of the marginalized d^2 posterior) and detaching d^2 from 0, thus making simpler the rejection of the null hypothesis $d = 0$.

The upper-right panel of Fig. 5.11a and the whole Fig. 5.11b show the μ_s, d corner subplots with marginalized distributions normalized at the peak and meaningful axis limits: for both μ_s and d^2 , the axis stops at 0 since negative values are unphysical; in a realistic data analysis process, these values would have been ruled out by the choice of prior distributions. With $d = 0.01$ the joint posterior is completely consistent with $d = 0$ and a practically unconstrained μ_s on a large interval (namely, without a charge it is not important if a field exists at all, thus the mass can *in principle* assume any value; this explanation is violated considering the posterior at sufficiently high d^2 due to strong covariance). With $d = 0.05$, it is possible to reject the null hypothesis $d = 0 = \mu_s$ (plot origin) at 1σ confidence level, however still not being able to constrain significantly further charge and mass. A good result is attained at $d = 0.1$, where not only the null hypothesis is rejected at $> 3\sigma$, but also physically meaningful marginalized posteriors for d, μ_s can be extracted. The effect on the remaining intrinsic parameters is practically negligible.

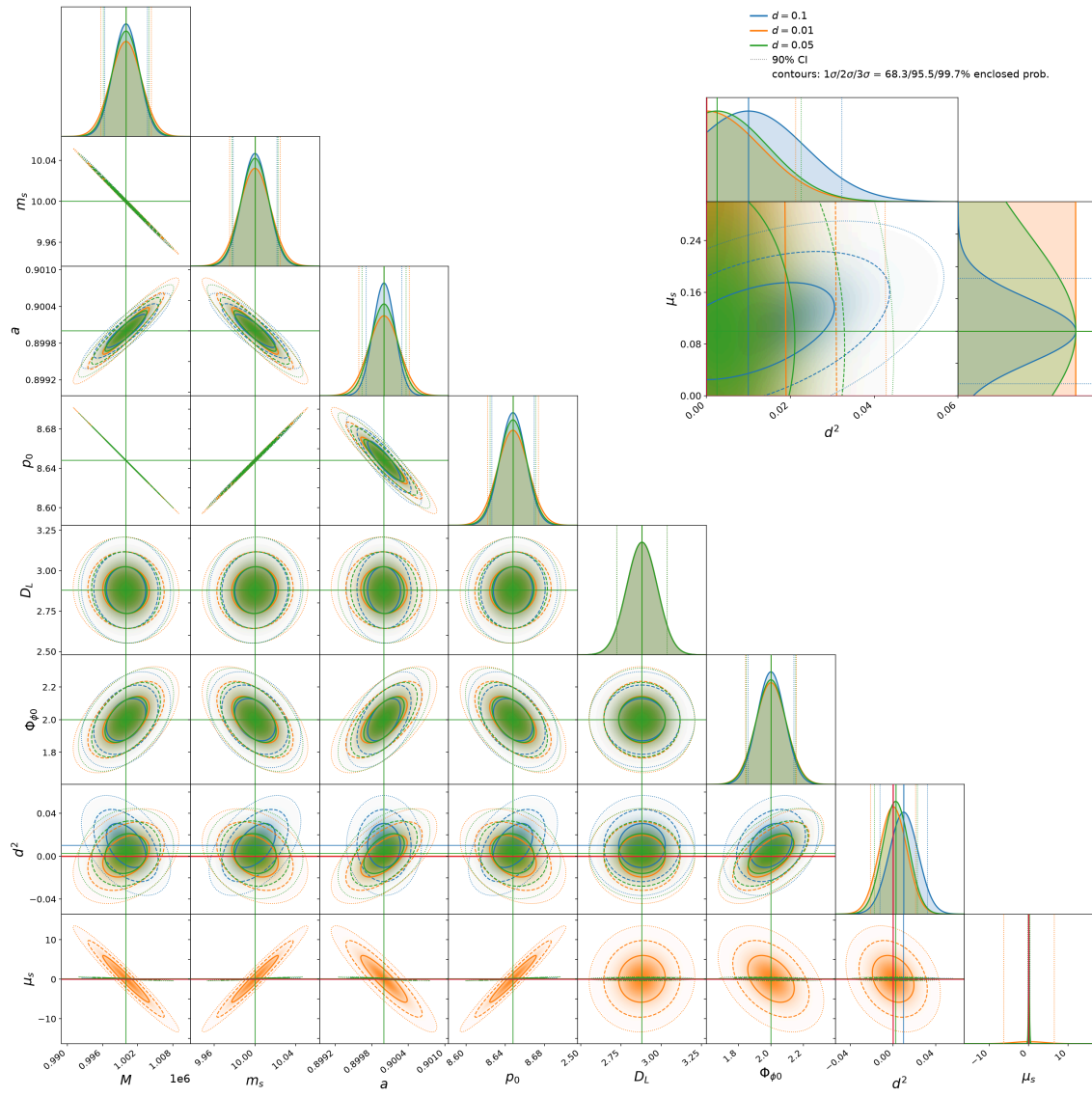
For higher SNR $\rho = 100$ (Fig. 5.11b) a relevant measurement of the scalar field parameters is already possible at $d = 0.05$. In fact, since the covariance matrix scales as $1/\rho^2$, the simplest way to obtain tighter bounds is to just consider higher-SNR events. This approach is not a trivial simplification of the data analysis process, instead is motivated by the nature of the detection: the goal is to find a system which has *the optimal configuration* together with a sufficiently high SNR to enable the detection of the field, not *assessing the presence of the field* on a large sample of objects with average SNR.

Case study 3

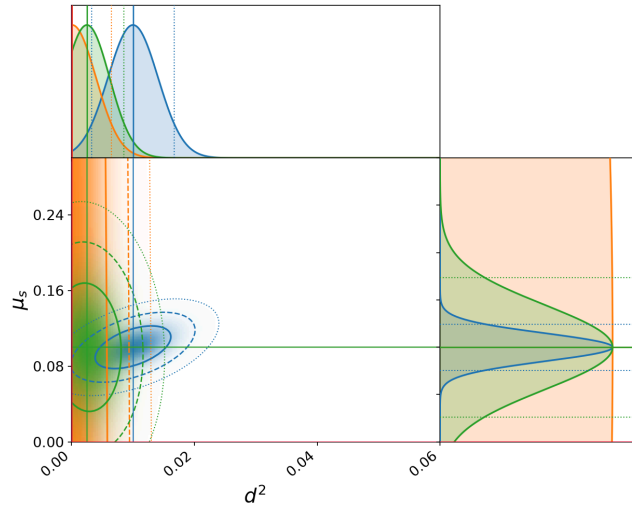
The previous two cases showed that, for the specific task of measuring the properties of a massive scalar field with relatively high mass μ_s , the optimal cases at fixed SNR do not correspond neither to quasi-monochromatic systems, which consistently accumulate dephasing but are highly degenerate, nor for fast-evolving systems, for which the radiative evolution is too short to produce a relevant signature in the GW signal. This suggests that the optimal system be a configuration starting in the strong superradiant regime and transitioning to the radiative regime, accumulating significant SNR in both the configurations: in this way, the *presence* of a scalar field is assessed through the detection of a consistent and long radiative phase, while the *massive nature* of the field can be inferred from the location of the sudden switch between superradiant and radiative regimes. The change should occur slowly, in such a way to accurately resolve the transitioning frequency.

To attain this specific dynamics, I used a third configuration consisting of a primary mass $M = 10^6 M_\odot$, secondary mass $m_s = 10 M_\odot$, as always $a = 0.9$ and $\mu_s = 0.1$ are fixed by the case study unidimensional grid, the observational time is kept at $T = 2$ years and the scalar charge varied across the three reference values $d = 0.01, 0.05, 0.1$. The luminosity distance is normalized to have the reference SNR $\rho = 30$ and 100 as for the previous case: in this way, the global scaling of the Fisher matrix does not mask the specific effect of the choice of the intrinsic parameters. The posterior probability representation is plotted in Fig. 5.12 with the same conventions of Fig. 5.11.

As expected, the dynamics through the superradiant-radiative boundary breaks the degeneracy between μ_s and d^2 , leaving a mild normalized covariance $-0.3 < \mathcal{C}_{d^2\mu_s} < 0.5$ across the three considered values of d^2 ; for $d = 0.05$ the covariance is almost zero. The



(a) Corner plot for SNR $\rho = 30$



(b) Sub-corner plot for SNR $\rho = 100$

Figure 5.12: Corner plot of the Gaussian posterior distributions for scalar charge $d = 0.01$ (orange), $d = 0.05$ (green), $d = 0.1$ (blue) and $M = 10^6 M_\odot$, $m_{\text{sec}} = 10 M_\odot$, $p_{\text{final}} = 4$, $T_{\text{obs}} = 2$ yrs, $\mu_s = 0.1$. The plot shows the 1, 2, 3 σ contours, and the null hypothesis $d = 0$, $\mu_s = 0$ for the scalar field parameters. Upper plot SNR $\rho = 30$, lower plot SNR $\rho = 100$.

marginalized posteriors are moderately tighter, with a substantial improvement in the measurement of μ_s and a moderate improvement for d^2 . This is expected, since the configuration was chosen mainly for probing the superradiant to radiative transition, which is regulated by μ_s , and not necessarily to maximize the flux, which would have resulted in a better measurement of d^2 (but affected by the degeneracy problem, see **Case study 1**: the measurement is actually better only if the scalar mass is independently known, namely the posterior is not marginalized over the scalar mass).

6.1 Superradiance: a general introduction

The term *Superradiance* indicates phenomena in which some type of incoming radiation is coherently amplified by a collective behavior of the system (Dicke, 1954; Brito et al., 2015).

The definition is very general and qualitative; let us analyze the schematic definition “**amplification of radiation by collective phenomena**”:

- **Amplification:** to observe superradiance, an incoming “seed” is needed, which is then amplified with a larger outgoing amplitude. In general, the seed can be classical (incoming radiation) or quantum (fluctuations of the vacuum for fields or the metric itself).
- **Radiation:** in a wide sense, radiation can be intended as any type of oscillatory phenomenon (acoustic waves in a fluid, oscillatory fields, either classical or quantum, scalar, vector, tensor) and the corresponding particle interpretation. For example, superradiance can manifest itself in electromagnetic radiation (photons), gravitational waves (gravitons), and exotic fields such as beyond-standard model bosons. The latter will be the case of interest, given the topic of a massive scalar field that mediates gravity.
- **Collective phenomena:** the amplification mechanism is usually macroscopic and acts as a virtually infinite energy reservoir for the amplification of the radiation. For example, such collective systems can be a supersonic flow of a fluid for acoustic waves, a strong potential barrier for charged particles, a rotating dielectric cylinder for electromagnetic radiation, a strong gravitational tide or spacetime dragging effect for GW or any field that is defined on such spacetime.

Before treating superradiance in the specific context of a massive scalar field around a black hole, I want to introduce a classical example in which superradiance occurs, namely the Klein Paradox, in order to build an intuitive and general comprehension of such a wide class of phenomena. Many other cases not regarding BH physics are thoroughly treated in Brito et al. (2015).

The Klein paradox

The Klein paradox is a first intuitive occurrence of superradiance. Let’s consider a massless scalar field with charge q , moving in a potential $A_\mu = (V_0, 0, 0, 0)$. The dynamics is set by the Klein-Gordon equation $\square\phi = \phi_{;\mu}{}^{;\mu} = 0$, where the covariant derivative (with

respect to the quadripotential A_μ) is defined as $\phi_{;\mu} = (\partial_\mu - iqA_\mu)\phi$. The problem can be solved by decomposing the radiation in an incoming, reflected and transmitted oscillating components, the first two with a wave vector $|\vec{k}| = \sqrt{\omega^2 - \mu_s^2}$, the latter with $|\vec{k}'| = \sqrt{(\omega - V_0)^2 - \mu_s^2}$. The magnitudes of the reflected and transmitted components are characterized by a reflection and transmission coefficient that quantifies the flux associated with the charge conservation. These coefficients can be expressed as

$$\mathcal{R} = \left| \frac{\vec{k} - \vec{k}'}{\vec{k} + \vec{k}'} \right|^2$$

$$\mathcal{T} = \frac{4|\vec{k}|\Re(\vec{k}')}{|\vec{k} + \vec{k}'|^2} e^{i(\vec{k} + \vec{k}') \cdot \vec{x}}$$

While in the weak and intermediate potential regimes, respectively $V_0 < \omega - \mu_s$ and $\omega - \mu_s < V_0 < \omega + \mu_s$, the reflection and transmission coefficients behave in a classical way, with both in the $[0, 1]$ interval and $\mathcal{R} + \mathcal{T} = 1$, in the strong potential regime they become $\mathcal{T} < 0$, $\mathcal{R} > 1$ and still $\mathcal{T} + \mathcal{R} = 1$. Classically, this would be a paradoxical and unphysical situation, in which the flux of a transmitted field is negative and the reflected field is more intense than the incoming.

The interpretation of the paradox is that particle/antiparticle pairs are created at the interface in such a way that the reflected radiation is amplified ($R > I$). We recognize the elements that define superradiance: 1) a *seed*, provided by the incoming radiation, without which the reflected radiation would be zero; note that at classical level although the interface presents a potential difference that would be enough to create the pair, it is necessary to have an incoming particle to actually produce it; 2) a propagating oscillatory behavior given by the Klein-Gordon equation, and 3) a macroscopic collective phenomenon, the potential that needs to be sustained for example by a battery: the creation of new particles or equivalently the amplification of incoming radiation comes at expense of this battery, which needs to do some work to maintain the potential barrier at a constant level.

We can observe two distinctive characteristics of a superradiant system, which will be present in the following examples:

- Superradiance involves “low-energy” radiation (ω small enough) and a “high-energy” amplification system (V big enough); in particular, in this case the relation $\omega \ll qV$ is required.
- Presence of a dissipative mechanism: the potential barrier (and the battery that keeps it at a constant value) can act as a dissipative system, in the sense that it has the capacity to absorb momentum from an incoming charged particle (and this has the effect of charging the battery). The dissipation in a wide sense allows a coupling between the reservoir and the radiation, enabling superradiant dynamics to occur.

Superradiance in fluidodynamics

Another case in which superradiance occurs is at the interface of two masses of fluids moving with supersonic relative velocity (Brito et al., 2015). This case is of special interest for the BH superradiance modelization since the linearized Navier-Stokes equations

for the motion of the fluid are uniquely mapped to the field equations on a curved background; in this sense, it is possible to perform fluid-mechanics experiments that can be directly mapped to the corresponding GR realization. For the specific case of superradiant instability for a Kerr BH (formally defined in Sec. 6.1.b), Švančara et al. 2026 were able to recreate a fluidodynamical experiment that maps the dynamics of a bosonic field in the Kerr metric, offering an experimental probe to investigate in a laboratory setting extreme physics and non-linear effects that are not captured by a first-order analytical analysis.

In this case, the energy needed to amplify an incoming signal is provided by the kinetic energy of the supersonic mass of fluid, that acts as a large battery. Mapped to the case of a Kerr black hole, the “supersonic mass of fluid” is represented by the interior of the ergosphere, which is supersonic in the sense that the dragging effect forces every worldline to co-rotate with the horizon (the speed of sound in GR corresponds to the speed of light c , which establishes causality), thus the “battery” is represented by the angular momentum of the Kerr BH.

6.1.a Superradiance for a rotating BH

Superradiance from BH thermodynamics

Very interestingly, the possibility of occurrence of superradiant phenomena for a rotating BH can be obtained directly from the “laws of BH dynamics”, which closely resemble the ones of thermodynamics:

- **First law: conservation of total energy** $\delta M = \frac{k}{8\pi} \delta A_H + \Omega_H \delta J \quad (+\Phi_H \delta Q)$
- **Second law:** $\delta A_H \geq 0$.

Where M is the gravitational mass of a BH system, A_H the total area of the horizon(s), k the surface gravity, Ω_H the angular velocity of the horizon as previously introduced and J the BH’s angular momentum. The complete formulation of the first law also includes the electric charge Q and the surface electric field Φ_H , but for the astrophysical case this is irrelevant.

Now we ask, given a field around the BH, if it can be used to extract energy from the BH itself: following the usual Fourier transform and spherical decomposition (or equivalently, considering a monochromatic wave for a specific spheroidal harmonic mode with azimuthal number m) for a scalar field $\varphi = f(r, \theta)e^{-i\omega t + im\phi}$ having a stress-energy tensor $T_{\mu\nu} = \varphi_{,\mu}\varphi_{,\nu} - \frac{1}{2}g_{\mu\nu}\varphi_{,\alpha}\varphi^{,\alpha}$, it is possible to relate the flux of angular momentum to the flux of energy, as previously detailed:

$$\frac{\delta L}{\delta E} = -\frac{T_{\phi}^r}{T_t^r} = \frac{m}{\omega}$$

this ratio corresponds to the ratio between the angular momentum and the energy (mass) lost by the BH itself. Note that the mass can, in fact, decrease, as long as sufficient angular momentum is extracted, in accordance to the first law with the constraint of the second law. Thus $\delta J = \frac{m}{\omega} \delta M$, and substituting this in the first law, we obtain

$$\delta M = \frac{\omega k}{8\pi} \frac{\delta A_H}{\omega - m\Omega_H}.$$

Thus, since $\delta A_H \geq 0$, the field extracts mass (namely $\delta M < 0$) from the BH if $\omega < m\Omega_H$. Note the close analogy to the Klein paradox case: superradiance occurs for small energy ω compared with a “potential” $m\Omega_H$ provided by the rotation of the BH. The angular momentum loss via superradiance slows the rotation of the BH, diminishing the potential, in the same way as the production of particle pairs in the Klein-Gordon paradox setup discharges the battery.

Superradiance from computed energy fluxes

In Sec. 3.3 I presented the general expression for the energy flux at horizon and infinity due to the scalar field radiation, namely (Eq. 3.8):

$$\dot{E}^\pm = \frac{1}{16\pi} \sum_{\ell, m} \omega k_\pm |Z_{\ell m}^\pm|^2$$

with $k_+ = \sqrt{\omega^2 - \mu_s^2}$, $k_- = \omega - m\Omega_H$. Once again, it is immediately evident that the energy flux at horizon can be negative provided that $k_- < 0$, thus obtaining the same result.

6.1.b Superradiant resonances

The superradiant regime for $\omega < m\Omega_H$ enables the amplification of incoming radiation at the expense of the BH rotational energy; if a feedback mechanism is provided, in which the amplified radiation is redirected cyclically to the superradiant region, an instability will develop. Let's consider a simplified model of a BH enclosed in a perfectly reflecting mirror. Without the BH, inside the closed mirror *normal modes* of a field are defined with a discrete spectrum analogously to a wave function confined inside a deep potential well; the spectrum is given by wavelengths $\lambda \sim 2R/n$, where R is the radius of the mirror. After factorizing the time dependence with an oscillating term $e^{i\omega t}$, equivalent to performing a Fourier decomposition, further radial-angular separation leads to a eigenmodes structure with the quantitative details (energy levels, mode decomposition) determined by the boundary conditions and the functional form of the potential well (“perfect mirror” $\sim$ square well, “massive field in a weak gravitational potential” $\sim$ Coulomb well, ecc.). However, if no dissipative mechanism is provided, the system will always be decomposed into a discrete energy spectrum of eigenfunctions, called *normal modes*. This is the case, for example, for the hydrogen atom: the electron is confined in a Coulomb potential well, and despite the singularity of the potential at $r = 0$ regularity conditions of the field can be imposed, eventually obtaining the well-known atomic structure organized in orbitals with a discrete spectrum indexed by the principal number n , azimuthal number ℓ and magnetic number m_ℓ , which define the angular decomposition in spherical harmonics. Note the great similarity with the gravitational case: a well that, at least in the weak-field approximation, resembles the Coulomb potential, a confined field (thanks to the massive field confinement mechanism) and the Teukolsky radial-angular decomposition, with the usage of spin-weighted spheroidal harmonics to include the Kerr frame-dragging effects. If the presence of the horizon is ignored (which modifies the regularity/boundary conditions at the “bottom” of the well), the eigenfunction problem for a massive admits a set of normal modes at given frequencies, which are determined by the massive Teukolsky equation. However, the BH inside the well has

a relevant effect on the system, that is adding a *dissipation mechanism*: an initially normalized normal mode can lose energy via absorption at the BH horizon (the field “falls inside”, a non-rigorous expression that, however, captures the essence of the mechanism). The overall amplitude of the field will thus vary in time (the eigenstates do not conserve normalization): after a time t , the trapped particle will have interacted $N \sim tv_g/L$ times with the central BH (v_g is the particle’s velocity), with an absorption probability $\mathcal{A}$. The amplitude between consecutive scatterings varies as $A_{n+1} = (1 - |\mathcal{A}|^2)A_n$, thus $A(t) = (1 - |\mathcal{A}|^2)^N A_0 \sim (1 - N|\mathcal{A}|^2)A_0 \sim (1 - t|\mathcal{A}|^2/L)A_0$. This can be easily described by a complex eigenfrequency $\omega = \omega_R + \omega_I$, with $\omega_I \sim |\mathcal{A}|^2/L$: the originally purely oscillating factor $e^{i\omega t}$ now displays a decreasing component $e^{-\frac{t|\mathcal{A}|^2}{L}}$ which dampens the oscillations (implementing the standard behavior for an oscillating system with a weak dissipation mechanism). For the case of quasi-normal modes in superradiant regime, the cross section $|\mathcal{A}|^2$ is *negative* (corresponding to the negative transmission coefficient in the Klein paradox), since the radiation is not absorbed but amplified instead: this gives rise to an unstable growth of the resonant mode, which forms a *bosonic cloud*, analogous to an electronic orbital in the hydrogen structure; for this reason, such a solution is called a *gravitational atom*. However, we are interested in the homogeneous solution of the field equations (with fixed boundary conditions) and not in the cloud itself and its spontaneous formation even in the absence of a charged source, which could indeed be astrophysically relevant for mainly two reasons: 1) direct GW observations: boson clouds are expected to radiate away the accumulated energy via monochromatic GW emission of a discrete spectrum, encoding the zero-order, fine and hyperfine structure of the gravitational atom’s energy levels (Bošković et al., 2026), and 2) for population properties, since superradiant extraction of angular momentum would modify the spin distribution of observed BHs. The presence of an ultralight bosonic field extracting angular momentum from rapidly-spinning BHs via superradiant instabilities is consistent with the currently observed spin distribution inferred from BH mergers (Kou et al., 2026).

Nevertheless, the presence of a quasi-resonant mode, corresponding to a pole of the Green’s function, will produce a resonant feature in the homogeneous solution field’s amplitude, as detailed in Sec. 6.4.b. Note that the pole is not on the real axis (the corresponding frequency has a little imaginary part); thus the field amplitude and the energy fluxes are always well-behaved.

6.2 Resonance phenomenology

6.2.a Resonance shape and parametrization

I thoroughly study the main resonance $n = 0$ for the $\ell = 1, m = 1$ mode. For a circular equatorial orbit and at first order in the $\mu_s M \ll 1$ parameter, the position of the resonance is given by the approximated resonant frequency (Brito et al., 2015; Detweiler, 1980)

$$\omega_{\text{res}}^2 = \mu_s^2 \left[1 - \left(\frac{\mu_s M}{\ell + 1 + n_{\text{res}}} \right)^2 \right] \quad (6.1)$$

with $n_{\text{res}} = 0, 1, 2, \dots$, analogous to the principal number of the hydrogen atom orbital structure.

The resonance shape follows a “resonance-antiresonance profile”, namely

$$-\frac{dE^s}{dt} = \dot{E}_0 \frac{1 + \left(\frac{p-(p_0-q)}{\Gamma/2}\right)^2}{1 + \left(\frac{p-p_0}{\Gamma/2}\right)^2} = \dot{E}_0 \frac{\left(\frac{\Gamma}{2}\right)^2 + (p - (p_0 - q))^2}{\left(\frac{\Gamma}{2}\right)^2 + (p - p_0)^2} \quad (6.2)$$

where p_0 is the semilatus rectum of the positive resonance peak and $p_0 - q$ is the semilatus rectum of the negative resonance peak (thus q is the radial distance between the resonance and antiresonance’s peaks), $\dot{E}_0$ is the baseline energy flux and Γ is the width of the resonances. In total we have 4 parameters, meaning that with 4 samples of the energy flux the whole shape of the resonance can be easily reconstructed.

It is important to make clear that this is a *phenomenological* parametrization, namely I found that this function describes extremely well the true resonance curve. This phenomenological function, however, hints at the underlying physics: for example, the widths of the resonance and the antiresonance are the same, and this fact is valid across the entire considered parameter space (also in the following). The parametrization is thus *minimal*, in the sense that it is not possible to describe the system with less degrees of freedom, and *physical*, in the sense that the parameters represent physical quantities and are not a generic parametrization (as it could have been a Taylor expansion around the peak). A physical motivation for the shape of the resonance is provided in Sec. 6.4.

The presence of an antiresonance peak at $p_0 - q$ could seem irrelevant for the dynamics of the secondary around the resonance; however, the value of the shift q has a direct impact on the main resonance:

- **Peak energy flux:** the energy flux at $p = p_0$ reads $E_p = E_0 \left(1 + (2q/\Gamma)^2\right)$, namely for larger separation the peak reaches higher values (becoming potentially more relevant).
- **Full Width at Half Maximum (FWHM):** the FWHM of the main peak reads $\Delta p_{\text{FWHM}} = \Gamma \sqrt{1 + (\Gamma/q)^2}$: in accordance to the previous result, a higher shift q produces a narrower resonance. The simple Breit-Wigner width Γ is recovered for $q \rightarrow \pm\infty$, namely the antiresonance is moved far away from the resonance; note, however, that the peak energy in this case diverges to $+\infty$.

For these reasons, the shift parameter q must be considered as a crucial quantity of the same importance as Γ when treating the magnitude and width of the main resonance peak. Unlike a simple resonance curve, the interplay between Γ and q determines the quantitative shape of the physically relevant energy flux peak.

In general, the width is extremely small for small scalar masses, and thus it is difficult to resolve numerically. For this reason, I wrote a code that follows the evolution in parameter space of the resonance to find where it falls with generic parameters. Briefly, the code samples the resonance in 4 points to find the 4 parameters that describe it, estimates the position of the peak with a variation of a certain generic parameter $\delta\vartheta$ using a quadratic interpolant extrapolated from the previous two iterations (a quadratic interpolant already displays a very good accuracy for the predicted position at a step size that allows for a complete coverage of the parameter space with $\mathcal{O}(10^3)$ iterations), jumps to the new semilatus rectum, and repeats the measurement of the actual peak using again the 4-point sampling method. The step is adaptive and is chosen to have an accuracy in finding

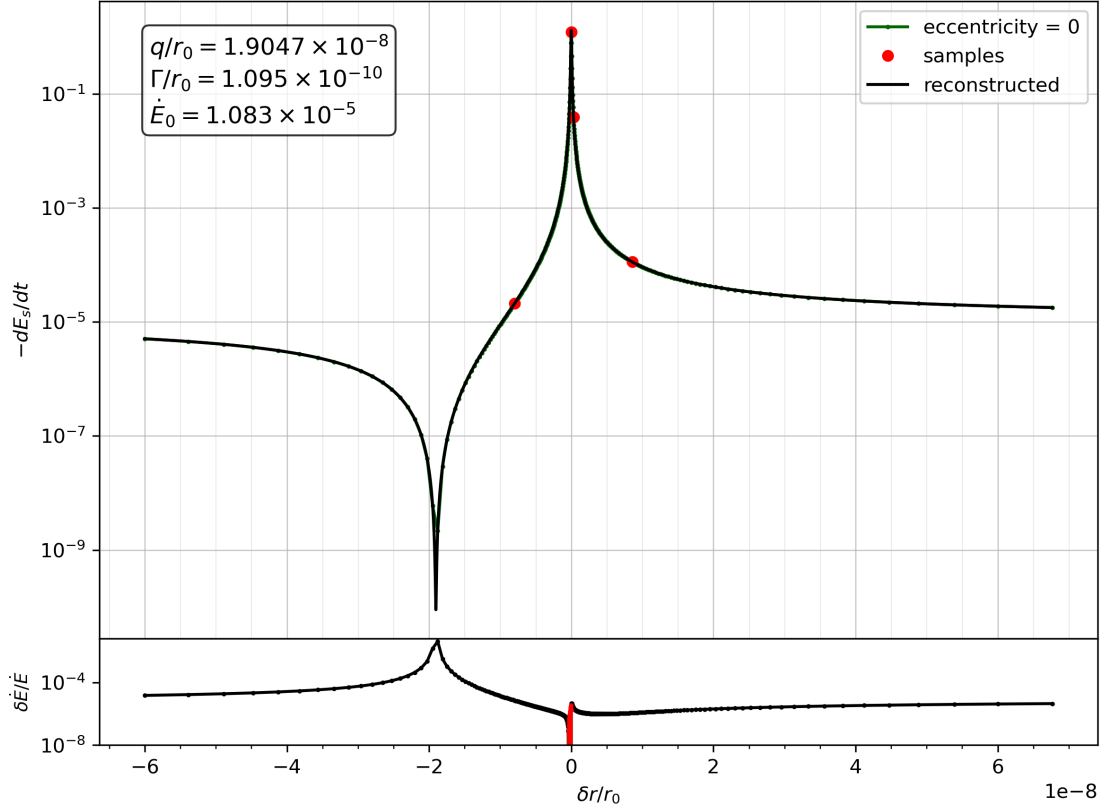


Figure 6.1: Main resonance in the circular equatorial case, with reconstructed resonance shape and residuals.

the peak of $\sim q/8$, for which the reconstruction of the peak is still possible and accurate. The step is subject to additional constraints and checks, mainly because the variation of the parameters is approximately quadratic, $\delta\vartheta$ tends to become large very quickly; on the other side, at a certain point the quadratic behavior changes abruptly, and if $\delta\vartheta$ was big at this point, the resonance is simply lost.

The resonance corresponding to circular equatorial orbits for $\mu_s = 0.1$, $a = 0.99$, $\ell = m = 1$ and $n_{\text{res}} = 0$ is shown in Fig. 6.1, together with sampled points (in red) for parameter reconstruction, reconstructed curve (in black, overlapped with computed fluxes), and residuals on the bottom plot. The sampling distance between the 400 samples is chosen to be finer near the resonance, where the function is steeper. I found that the shape is better reconstructed if the samples are taken around the positive peak, where larger fluxes enable a more precise reconstruction. In particular, the estimation of Γ is sensitive to having a sample near the peak of the resonance, and it is poorly estimated when the peak is not sampled. For this reason, the code that follows the resonance computes a fine-tuned estimation of the parameters before proceeding with the next step, in order to evaluate with accuracy also the Γ by centering a sample point exactly on the first estimation of the peak's radius (which in contrast is already very accurate).

In the circular equatorial case, the presence of this superradiant resonance implies the existence of a *floating orbit*: the total energy flux $\dot{E}^{\text{tot}} = \dot{E}^g + \dot{E}^s = 0$ is null together with the total angular momentum flux in correspondence with 2 values of the radius, namely on the right and on the left shoulders of the positive resonance peak. The equilibrium

on the left shoulder is always unstable, so it will not be considered. The equilibrium on the right shoulder is stable: at higher radii, the GW emission will dissipate energy shrinking the orbit, and at smaller radii the superradiant resonance will transfer energy from the BH to the secondary, rising the orbit. The stable equilibrium is thus completely stalled, in the sense that at first order in the adiabatic approximation no evolution occurs; to see an evolution, one must consider the variation of the BH parameters caused by the energy extraction through the resonance, and how the resonance profile is affected. Being a second-order effect, a first-order treatise would be inconsistent to solve this change of the background metric.

6.2.b Variation under change of eccentricity or inclination

I begin by exploring the effects of eccentricity on equatorial orbits ($\theta = 0$). Starting from the previously commented resonance (lowest-order $n_{\text{res}} = 0$, $\ell = m = 1$, $\mu_s = 0.1$ and extremal spin $a = 0.99$), the eccentricity is increased until the resonance is “lost”, in the sense that $\dot{E}_0$ becomes 0 and the inversion of Eq. (6.2) becomes impossible (some parameters are singular): in particular, we observe that the peak shift q exhibits an asymptote. The resonance parameters related to the circular equatorial values are plotted in Fig. 6.2; the left column shows the behavior for small eccentricity values, while the right column contains the entire eccentricity interval. The width parameter Γ is plotted separately in the upper row, since the variation is much smaller compared to the variation of p_0 , q , E_0 . I verified that the functional form of the resonance, Eq. (6.2), appropriately represents the resonance profile also at high eccentricity (for example, Fig. 6.4).

The qualitative behavior inferred from Fig. 6.2 implies that the resonance also survives at high eccentricity without substantial modifications of width and peak energy (plotted separately in Fig. 6.3). However, a significant shift in the position of the semilatus rectum is observed (relevant with respect to the resonance width and antiresonance shift); this could lead to important modifications of orbital dynamics, treated in more generality in Sec. 6.2.c.

Analogously, spherical orbits characterized by null eccentricity but non-zero inclination are considered in Fig. 6.5 and 6.6.

For both the cases, the peak energy is decreasing when departing from circular equatorial orbits; in case floating orbits are not stable under orbital perturbations (Sec. ??), this would provide a first-order exit mechanism from the resonant orbit in the adiabatic approximation, without the need for a second-order treatise: if the orbit is dynamically shifted away from the circular equatorial configuration, the resonance weakens until it cannot sustain a “quasi-floating” configuration; at this point the usual adiabatic evolution resumes, and the resonance is crossed. The net effect of this crossing would be a discontinuity of the orbital parameters, since to overcome the resonance the orbit must be pushed either to high eccentricity or high inclination.

6.2.c Resonance parameters for generic orbits

A complete grid in the eccentricity vs radius parameter space is built to study the resonance for a generic orbit (eccentric and inclined). The parameters are evaluated for every inclination value between 0° and 20° at intervals of 0.1° with the procedure used in Sec. 6.2.b (non-eccentric orbits), and starting from each of these values the eccentricity has

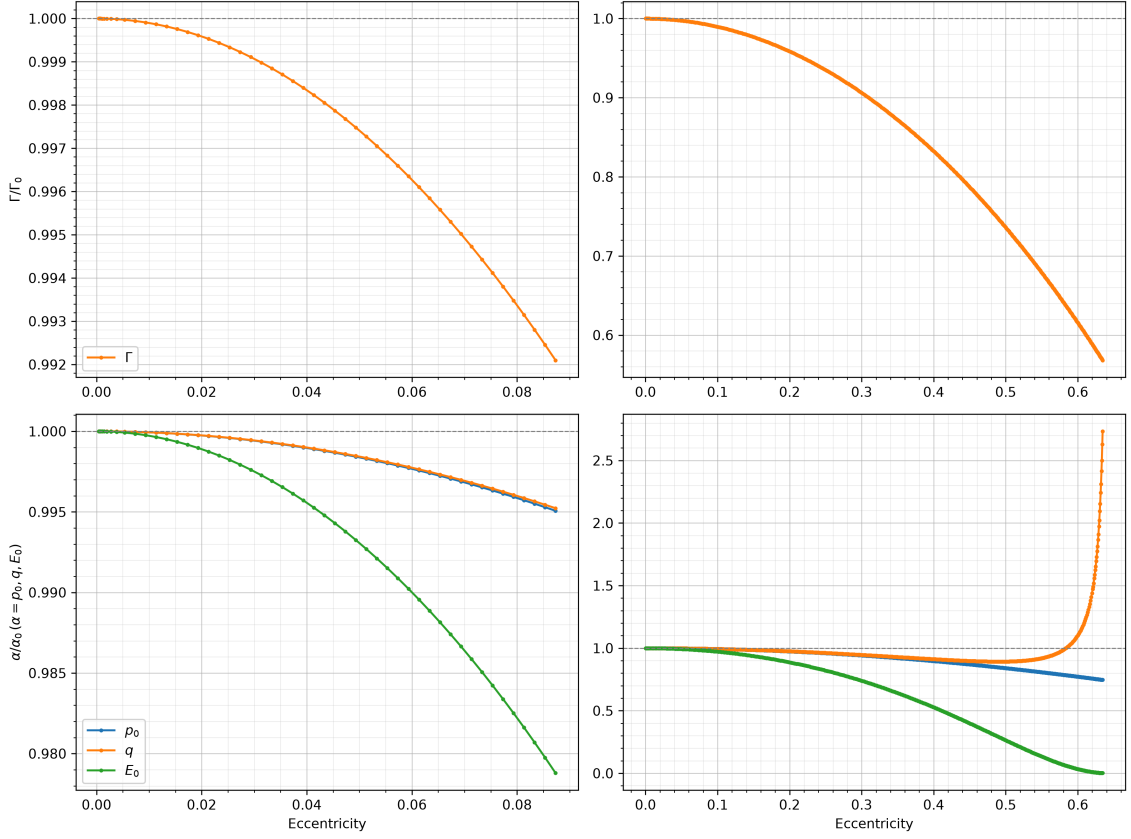


Figure 6.2: Behavior of the parameters that describe the resonance with varying eccentricity (equatorial case). The values are normalized to the circular equatorial parameters. Upper plots: resonance width Γ (isolated from the others to appreciate the change from 1, otherwise not visible); Lower plots: resonance radius p_0 , peaks separation q , energy flux baseline E_0 . The two plots on the left represent the behavior for small values of eccentricity, up to ≈ 0.1 , while the plots on the right contain the entire eccentricity range.

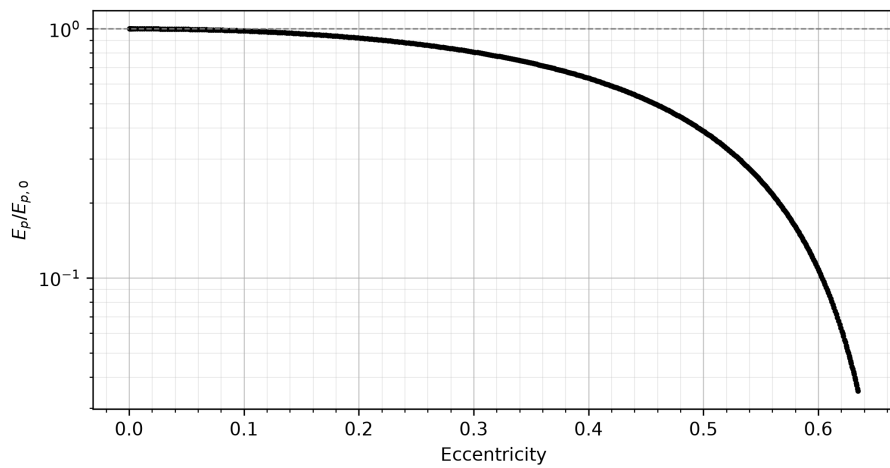


Figure 6.3: Peak energy flux (related to the circular equatorial resonant orbit) as a function of the eccentricity, for equatorial orbits. Before the parameters become singular, the peak energy decreases, in such a way that eventually it cannot sustain a resonant orbit. Note the logarithmic y -scale.

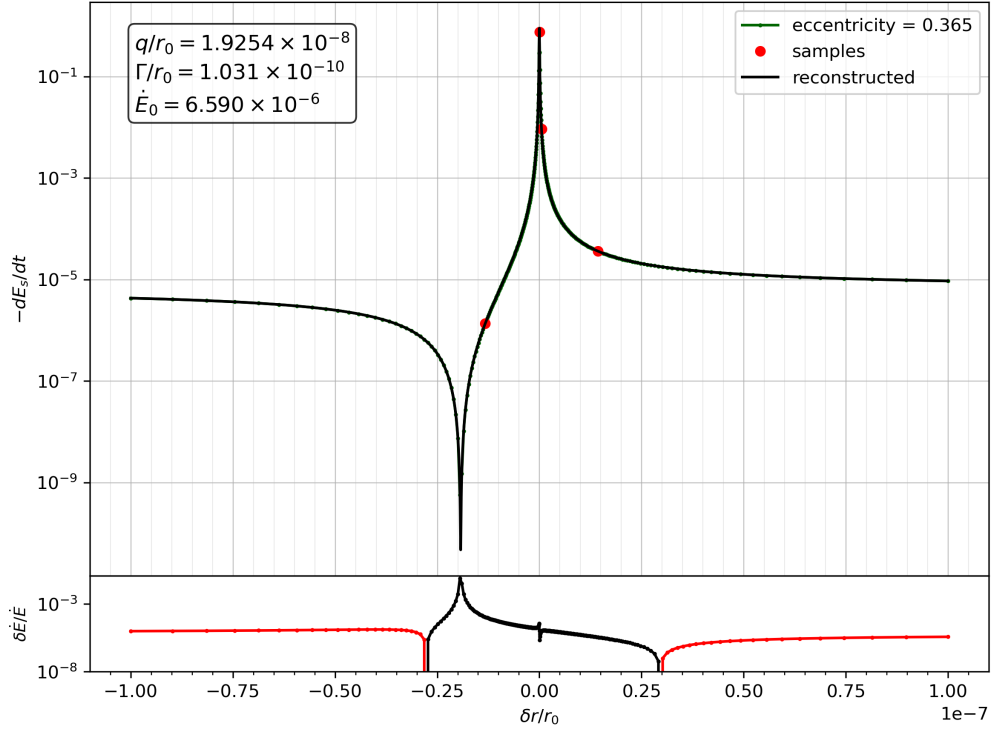


Figure 6.4: Resonant profile for highly eccentric equatorial orbit, analog to Fig. 6.1. The resonance is preserved with small residuals.

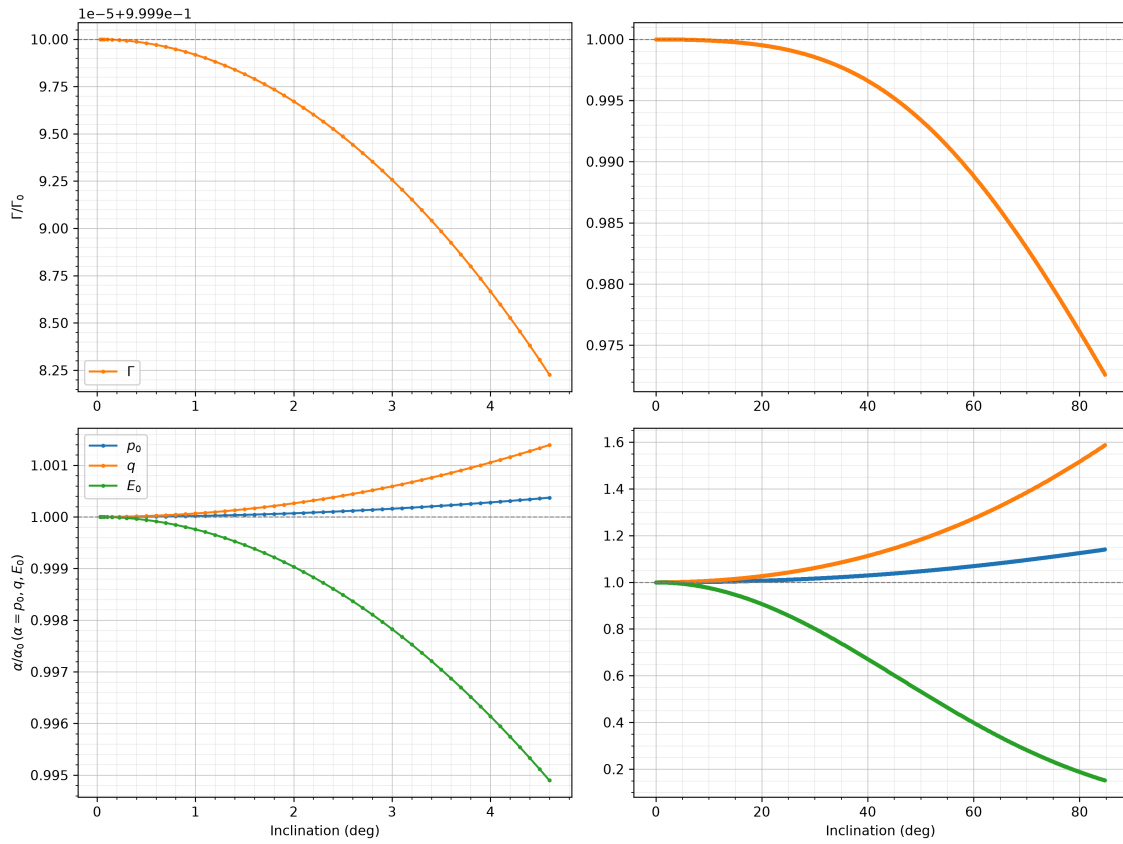


Figure 6.5: Behavior of the parameters that describe the resonance with varying orbital inclination (non-eccentric case). The representation is analogous to Fig. 6.2.

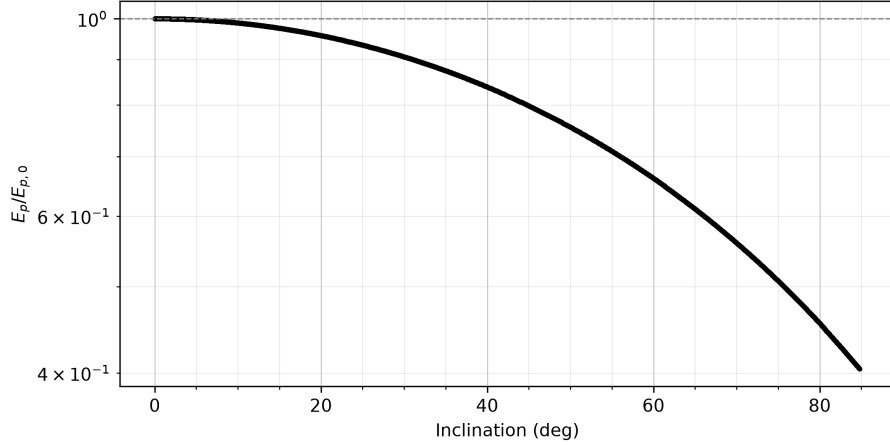


Figure 6.6: Peak energy flux (related to the circular equatorial resonant orbit) as a function of the inclination, for circular inclined orbits. The representation is analogous to Fig. 6.3.

been varied analogously to Sec. 6.2.b.

The resulting set of parameters that characterize the resonance is represented in Fig. 6.7, together with the relevant quantity E_p , the modulus of the energy flux at the peak, $E_p = E_0 \left(1 + (2q/\Gamma)^2 \right)$, as previously reported. Levels are also plotted to facilitate reading.

A general consequence of symmetry with respect to inclination and eccentricity (together with a regularity condition of at least differentiability) is that the level lines $\mathcal{I}(\alpha_*)$ for a certain parameter α with a fixed value α_* cross the two axes perpendicularly, namely at first order $\alpha(\epsilon, \tilde{\theta}) = \alpha(\epsilon, \theta = 0)$ for small inclinations $\tilde{\theta}$, and viceversa with small eccentricity. This is due to the fact that an eccentricity $-\epsilon$ is equivalent to ϵ , and analogously for the inclination. Specialized to the origin of the eccentricity-inclination plane, the circular orbit can be either a maximum/minimum or a saddle point for a chosen parameter; the derived quantities, as the aforementioned peak energy and FWHM, will inherit this structure, with the nature of the origin determined by the functional form of the derived quantity.

6.3 Dynamical evolution of resonant orbits

In the previous Section, I obtained a wide (and generalizable to different $\ell, m, k, n, n_{\text{res}}$ modes) description of the scalar fluxes in a neighborhood of a superradiant resonance. In this Section I will describe the dynamical evolution of such orbits; indeed, a *resonant* orbit is not necessarily a *floating orbit*, for the reason explained in Sec. 6.3.a. Afterwards, a characterization of such orbits is provided, following the two different perspectives of stability theory and full dynamical evolution.

6.3.a Resonant but not floating: the uncomfortable situation of generic orbits

A *floating orbit* is defined as a configuration in which the dynamics is completely stalled, i.e. both the fluxes of energy and angular momentum are null. For circular equatorial orbits, the energy condition and the angular momentum condition are dependent, namely the net zero energy flux condition implies null angular momentum flux. This can be seen through the mode-by-mode relation (valid for both scalar and gravitational emis-

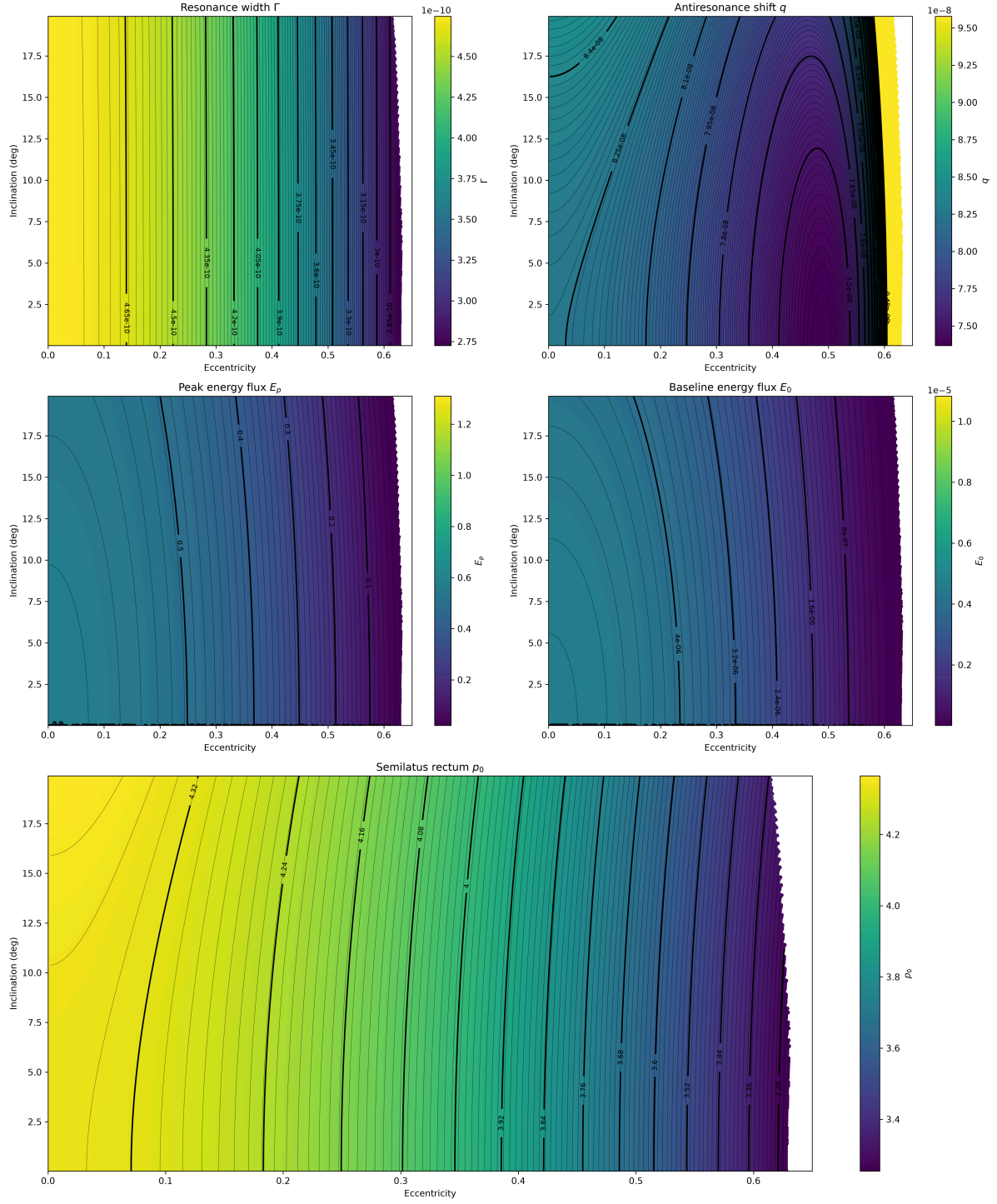


Figure 6.7: Behavior of the parameters that describe the resonance with varying orbital inclination and eccentricity (generic orbital configuration).

sion) $\dot{L}_{\ell mnk} = \frac{m}{\omega_{mnk}} \dot{E}_{\ell mnk}$ as found before, with $\omega_{mnk} = m\Omega_\phi + n\Omega_r + k\Omega_\theta$; for circular equatorial orbits, $n = k = 0$, thus $\omega_{mnk} = m\Omega_\phi$ and the ratio m/ω_{mnk} becomes simply $1/\Omega_\phi$, which can be brought outside the series when resumming the modes; thus, globally $\dot{L} = \dot{E}/\Omega_\phi$.

This is not true for a generic orbit: the condition of a null balance of energy fluxes is:

$$0 = \dot{E}^{\text{tot}} = \sum_{\ell mnk} \left[\dot{E}_{s,\ell mnk}^- + \dot{E}_{g,\ell mnk}^- + \dot{E}_{g,\ell mnk}^+ \right] \quad (6.3)$$

and using $\dot{L}_{\ell mnk} = \frac{m}{\omega_{\ell mnk}} \dot{E}_{\ell mnk}$ as before, we obtain

$$0 = \dot{E}^{\text{tot}} = \sum_{\ell mnk} \frac{\omega_{\ell mnk}}{m} \left[\dot{L}_{s,\ell mnk}^- + \dot{L}_{g,\ell mnk}^- + \dot{L}_{g,\ell mnk}^+ \right] \quad (6.4)$$

This is in general not equivalent to an angular momentum balance $\dot{L}_{\ell mnk}^{\text{tot}} = 0$, meaning that even if the energy balance condition is met, the angular momentum flux is still making the orbit evolve. This represents a “functional independence” argument, in the sense that the non-zero angular momentum balance is due to the angular momentum independence from energy in the case of eccentric orbits. A similar consideration is valid also for the Carter constant, even if this has a more involved algebraic expression. For now, let’s reason only on the eccentricity / angular momentum evolution.

A more rigorous perspective comes from considering energy and angular momentum fluxes simultaneously, without giving a privileged role to the energy flux as implicitly assumed before; the orbital evolution is given by the orbital parameters p, e , which can be seen as a coordinate transformation of the action variables $E(p, e)$, $L_z(p, e)$; this means that the orbital evolution is determined equivalently by

$$\begin{aligned} \frac{dp}{dt} &= \frac{\partial p}{\partial E} \frac{dE}{dt} + \frac{\partial p}{\partial L_z} \frac{dL_z}{dt} \\ \frac{de}{dt} &= \frac{\partial e}{\partial E} \frac{dE}{dt} + \frac{\partial e}{\partial L_z} \frac{dL_z}{dt} \end{aligned}$$

which in matricial form can be written as

$$\frac{d}{dt} \begin{pmatrix} p \\ e \end{pmatrix} = \begin{pmatrix} \frac{\partial p}{\partial E} & \frac{\partial p}{\partial L_z} \\ \frac{\partial e}{\partial E} & \frac{\partial e}{\partial L_z} \end{pmatrix} \frac{d}{dt} \begin{pmatrix} E \\ L_z \end{pmatrix} \quad (6.5)$$

Since the transformation is invertible (the determinant of the Jacobian is finite for generic orbits and becomes singular only when the orbit is circular, in which case the system reduces to a single equation since $\dot{E}$ and $\dot{L}_z$ are not independent anymore), the orbital parameters are floating ($\dot{p} = \dot{e} = 0$) if and only if also $\dot{E} = \dot{L}_z = 0$. This is easily generalizable to the three-dimensional case with Carter constant and orbital inclination. It is then clear that the floating orbit configurations are uniquely determined by the simultaneous condition $(\dot{E}, \dot{L}_z, \dot{Q}) = 0$.

In the first-order adiabatic approximation, a change in the angular momentum of the particle will change the value of eccentricity. Note that it is not obvious *a priori* in which direction: it could be a circularization as well as a “eccentricization”, depending on the interplay between the scalar modes m, n, k , the resonance order n_{res} and the GW modes:

since the scalar mode is superradiant, it is possible that a net angular momentum at fixed energy is acquired rather than radiated away. This will affect the resonance parameters: the resonance profile will thus evolve, again enabling a net energy flux. In the hypothesis that the resonance is very peaked, two scenarios are possible:

- The stable equilibrium moves inward (this is equivalent to the eccentricity *increasing* or the inclination *decreasing*, see Fig. 6.2, 6.5, 6.7): in this case, the orbit evolves more similarly to the standard GW dissipation, following the equilibrium inwards.
- The stable equilibrium moves outward (viceversa, eccentricity *decreasing* or inclination *increasing*): in this case, whether the resonance will push outward the orbit, rising the semilatus rectum, or if the resonance shifts faster than the orbital radius, enabling the orbit to “jump” over the resonance, depends on the specific parameters such as the height and the width of the resonant peak.

Clarified the fact that general resonant orbits evolve at first adiabatic order and that the evolution is non-trivial, in the following Sections I will develop a qualitative (but strongly physically motivated) description of the resonant dynamics.

6.3.b Methods and qualitative phenomenology of generic resonant orbits

The next step is to evaluate the dynamics of an orbit near a resonance; this can be done at two different levels:

- Study of the stability properties of a configuration which is slightly displaced from the equilibrium position; this “linearized” approach gives a quantitative description of quasi-equatorial, quasi-circular orbits and a good qualitative description of generic orbits, thanks to the global regularity of the resonance’s parameters (which from Fig. 6.7 clearly follow a functional behavior that resemble globally the local properties of stationary points, described by stability theory).
- Full evolution of a resonant, non-floating orbit: this approach involves the explicit integration of the energy, angular momentum and Carter constant fluxes. While giving the exact dynamics (and waveforms), it is computationally expensive and, as today, difficult to achieve. For this reason, this approach will be replaced here by a demonstrative toy model.

As we saw, resonant orbits are not generally floating and will thus undergo a dynamical evolution at the first adiabatic order, which however is completely different from the usual adiabatic inspiral due to the presence of the scalar resonance. An important consideration for assessing stability concerns the dimensionality of the problem:

- If orbits are parametrized just by the radial coordinate, we obtain circular equatorial geodesics. As explained before, with this only dimension, the floating orbit is actually a stable equilibrium.
- Adding a second coordinate, either the eccentricity e or the inclination $x \equiv \cos \theta$, the equilibrium can lose its stability properties: adding a second principal axis to the matrix that describes the linearized flux around the floating equilibrium (previously purely attractive) can render it unstable if the new direction is repulsive. As argued before, it is not given *a priori* whether the new direction will be attractive or repulsive with respect to the equilibrium; however, it is clear that the principal axis is

aligned with the new coordinate axis. This is very important for the following, and a consequence of the symmetries of the system (and the fact that the chosen modes, defined by ℓ, m, k, n , are eigenfunctions of the corresponding action variables).

- Adding the last dimension, the third, is equivalent to what has been said for the second: another possibility of changing the stability of the floating orbit from stable to unstable.

This encodes more formally the “functional independence” argument presented in Sec. 6.3.a: in the $(p - p_{ce}, e, \cos \theta)$ parameter space (p_{ce} is the radius of a circular equatorial resonant orbit, which is truly floating), every balance condition defines a 2-dimensional manifold (surface) passing through the origin (the representation of the null energy flux manifold corresponds approximately to the lower panel of Fig. 6.7; only approximately since the net zero energy condition is satisfied not at the resonance peak, but slightly nearby, and it depends on the GW flux); the intersection of these manifolds thus is 1-dimensional, which is already geometrically not general, but enforced by the functional dependence of the three conditions in the origin.

It is then possible to predict with more accuracy the qualitative evolution of a floating orbit based on the stability properties of the floating orbit:

- **Stable equilibrium:** if all the three principal axes of the linearized system are attractive (that is, the two corresponding to eccentricity and inclination), the orbit will evolve towards a circular equatorial configuration. To reach it, the action variables E, L_z, Q can vary in *any* direction (the orbit can gain or lose energy, ecc), but with the constraint that the three functions $\delta p \equiv |p - p_{ce}|, e$ and $|\theta|$ are monotonically decreasing toward 0. Once the floating orbit is “reached” (inappropriate term, since for existence and unicity of the solution the actual equilibrium is never reached), the evolution will be determined by second-order effects in the adiabatic approximation, thus it is not possible anymore to predict what will happen.
- **Unstable equilibrium:** if even one of the e, x coordinates has repulsive character, then the equilibrium is unstable and the configuration will drift away from the floating orbit. As can be seen in Figs. 6.3 and 6.6, the peak energy flux of the resonance diminishes as the eccentricity and/or the inclination increase. This means that the orbit will reach a regime in which the resonance is not strong enough to maintain a net zero energy flux, and at this point the trajectory can pass over the resonance and resume the standard inspiral; observationally, this would correspond to a “glitch” in which the orbital parameters change between two phases of quasi-adiabatic inspiral (how fast this change happens is a matter of a quantitative analysis). In principle, another possibility is that the resonance is strong enough to bring the trajectory near the separatrix (which is the reason why the resonance parameters become singular, see Fig. 6.2) before disappearing, thus having the potential effect of causing a premature plunge.

Quite conveniently, the study of the equilibrium properties can be performed with a point-like linearized analysis on the principal axes of the flux. Since we know that these axes correspond to the $e = 0$ (inclined circular) and $\theta = 0$ (eccentric equatorial) configurations, this reduces to the computation of the GW modes in these two isolated cases.

For the case $\ell = 1$, floating orbits have been proved to be stable against small perturbations of eccentricity, namely the orbit tends to circularize, by Cardoso et al., 2011. This is

not necessarily the case for generic ℓ ; in addition, stability under perturbations of the inclination has not been assessed yet. A possible hint for inclination instability comes from the fact that the resonance's semilatus rectum exhibits opposite trends under change of eccentricity and inclination: in the former case, increasing the eccentricity reduces p_{res} , while in the latter increasing the inclination increases p_{res} . A quantitative assessment of this stability is left for a future study.

6.4 A theoretical explanation for the resonance shape

6.4.a Resonance interference with a wide-band background

As argued before, the resonant profile in the horizon energy flux arises from the squared amplitude at r_p of the field component R^+ , namely the part of the homogeneous solution that is purely outgoing at infinity. Near the resonance frequency determined by the existence of an unstable quasi-normal mode of the field, the total field is composed of *two different field amplitudes* which are then summed: one is the *baseline amplitude*, which is smooth on wide energy scales (small derivative with respect to energy everywhere; again, *smooth* not in a differentiability sense, just as lack of short-scale features) and is represented by the approximately constant $\dot{E}$ that enters the phenomenological parameterization of the resonance curve. Over this is added the *resonant amplitude* stemming from the presence of the quasi-normal mode of the scalar field. This comes from a different mechanism (not a homogeneous solution for a continuous energy spectrum but a feature arising from the discrete spectrum of eigenmodes, which are approximately stationary *even without the need for an oscillating boundary condition*, which corresponds to a forcing mechanism. In other words, a source (the charged inspiraling secondary) is not needed to sustain the discrete oscillation spectrum of the field, which is present regardless. This does not mean that the properties of the secondary have no effect on the energy flux: if the secondary is in the resonant regime, it couples with the discrete mode through its scalar charge, producing the resonant flux. As an extreme case, if $d = 0$ in generic conditions the wide-band solution of the field is *not* produced since it is not actively excited by the secondary, and thus no scalar energy flux occurs; the discrete spectrum, by contrast, *is* produced regardless of the *absence* of a source, but if the secondary has null charge it cannot couple with the field and exchange energy, thus again $\dot{E}_s = 0$. Despite this case being physically not informative, it makes clear the conceptual distinction between the wide-band solution and the resonance at the *field* level, while at the *energy flux* level the charge plays a crucial role for both the components.

This observation can also be seen from the computed fluxes with a data-drive approach: if exclusively a resonant amplitude were present, the flux would appear as a Bode plot with different slopes in the log-log plane for the tails of the resonance, which is not the case; instead, the energy follows a consistent baseline at both sides of the resonance.

The total field can thus be written exactly as

$$\varphi^+(\omega(r_p)) = \psi_r(\omega)e^{i(\omega t + \Phi_r(\omega))} + \psi_b(\omega)e^{i(\omega t + \Phi_b(\omega))} \quad (6.6)$$

where $\psi_r(\omega)$ is the *resonant amplitude*, which depends rapidly on ω and is parametrized by the resonant frequency $\omega_0 = \omega_R$ and the peak width $\Gamma = |\omega_I|$ through a standard

resonance curve

$$\psi_r(\omega) = \alpha \frac{\omega_0^2}{\sqrt{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2}} \quad (6.7)$$

and $\Phi_r(\omega)$ is the associated standard resonance phase, also this is rapidly dependent on ω and given by

$$\Phi_r(\omega) = \arctan \left[\frac{\Gamma\omega}{\omega^2 - \omega_0^2} \right] \quad (6.8)$$

such that, using $\cos \arctan x = 1/\sqrt{1+x^2}$ and $\sin \arctan x = x/\sqrt{1+x^2}$, the sin and cos of the phase can be written as

$$\cos \Phi_r = \frac{\omega^2 - \omega_0^2}{\omega_0^2 \alpha} \psi_r(\omega) \quad \sin \Phi_r = \frac{\Gamma\omega}{\omega_0^2 \alpha} \psi_r(\omega)$$

These are well-known results that are common to every physical system behaving like a resonant harmonic oscillator. On the other hand, the background field contribution is weakly dependent on ω on the considered energy interval (of the order of Γ , which is small compared to ω_0 since the cross-section is extremely small: in other words, the characteristic time for decay/explosion is very long), thus approximately $\psi_r \sim \text{constant}$, $\Phi_r \sim \text{constant}$ in ω . In addition, the background field dominates outside the resonance, namely $\xi \equiv \alpha/\psi_b \ll 1$. We then see that there are two small expansion parameters, namely the ratio ξ between the baselines of the fields for $\omega < \omega_0$ and the energy shifts $\delta \equiv \delta\omega/\omega_0 \sim \Gamma/\omega_0$. I will use the properties $\xi \ll 1$ and $\delta \ll 1$ (independently) to describe the resulting amplitude.

The energy flux is computed from $|Z^+|^2 \propto \varphi^+(\varphi^+)^*$, thus let's compute the square modulus:

$$\begin{aligned} |\varphi^+|^2 &= \left(\psi_r(\omega)e^{i(\omega t + \Phi_r(\omega))} + \psi_b(\omega)e^{i(\omega t + \Phi_b(\omega))} \right) \left(\psi_r(\omega)e^{-i(\omega t + \Phi_r(\omega))} + \psi_b(\omega)e^{-i(\omega t + \Phi_b(\omega))} \right) \\ &= \psi_r^2(\omega) + \psi_b^2 + 2\psi_r(\omega)\psi_b \cos(\Phi_r(\omega) - \Phi_b) \\ &= \psi_b^2 \left[1 + \left(\frac{\psi_r(\omega)}{\psi_b} \right)^2 + 2 \frac{\psi_r(\omega)}{\psi_b} \left(\cos \Phi_r \cos \Phi_b + \sin \Phi_r \sin \Phi_b \right) \right] \\ &= \psi_b^2 \left[1 + \xi \frac{\omega_0^4}{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2} \left[\xi + \frac{2}{\omega_0^2} \left((\omega^2 - \omega_0^2) \cos \Phi_b + (\Gamma\omega) \sin \Phi_b \right) \right] \right] \\ &= \psi_b^2 \frac{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2 + \xi^2 \omega_0^4 + 2\xi\omega_0^2 \left[(\omega^2 - \omega_0^2) \cos \Phi_b + (\Gamma\omega) \sin \Phi_b \right]}{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2} \end{aligned}$$

where I used the previously found trigonometric relations for the resonant phase and started to organize the terms in order to reproduce the phenomenological description used previously, featuring a multiplicative amplitude that corresponds to the baseline solution, a positive peak due to the denominator and an antiresonance due to the numerator. Note that until this point everything is exact, namely no approximation (linearization in the small parameters) has been adopted. The amplitude resonance correctly displays a positive peak corresponding to the resonant frequency ω_0 due to the denominator; the numerator contains additional terms. In order to reproduce a shifted antiresonance, let's

add and subtract a term $(\omega^2 - (\omega_0 - q)^2)^2$ to the numerator. In this moment, q is a free parameter, in the sense that its numerical choice is arbitrary. This results in the resonance-antiresonance shape with a remainder:

$$|\varphi^+|^2 = \psi_b^2 \frac{\left\{ (\omega^2 - (\omega_0 - q)^2)^2 + (\omega\Gamma)^2 \right\} + G(p)}{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2}$$

let's consider the polynomial in q outside curl brackets in the numerator, called $G(p)$:

$$G(p) = (\omega^2 - \omega_0^2)^2 - (\omega^2 - (\omega_0 - q)^2)^2 + \xi^2 \omega_0^4 + 2\xi \omega_0^2 \left[(\omega^2 - \omega_0^2) \cos \Phi_b + (\Gamma \omega) \sin \Phi_b \right]$$

Note that the first relevant order in the small parameter $\delta \equiv \delta\omega/\omega_0 \sim \Gamma/\omega_0$ for both the numerator and the denominator of $|\varphi^+|^2$ is the second: indeed, expanding the denominator leads to $(\omega^2 - \omega_0^2)^2 + \omega^2 \Gamma^2 \propto (\omega - \omega_0)^2 + (\Gamma/2)^2 + \mathcal{O}(\delta\omega^3)$, which is exactly the form of the phenomenological parameterization. We thus learned that the parametrization is actually an expansion (extremely well motivated a posteriori) of the exact resonance shape. As this contribution is at second order in $\delta\omega \sim \Gamma$, let's expand $G(q)$ at the same order in ξ and $\delta\omega \sim \Gamma \sim q$. Note that the last requirement $q \sim \delta\omega$ needs to be verified *a posteriori*. We thus obtain:

$$\begin{aligned} G(q) &\sim -4\omega_0^2(2\delta\omega q + q^2) + \xi^2 \omega_0^4 + 2\xi \omega_0^2 \left[2\omega_0 \delta\omega \cos \Phi_b + (\Gamma \omega_0) \sin \Phi_b \right] \\ &= -4\omega_0^2 \left\{ q^2 + 2q\delta\omega - \frac{\omega_0^2}{4} \left[\xi^2 + 2\xi \left(2\frac{\delta\omega}{\omega_0} \cos \Phi_b + \frac{\Gamma}{\omega_0} \sin \Phi_b \right) \right] \right\} \end{aligned}$$

as argued before, since q is a free parameter, we can substitute it with a root $\tilde{q}$ of this second-degree polynomial. The roots are:

$$\tilde{q}_{\pm} = -\delta\omega \pm \sqrt{\delta\omega^2 + \frac{\omega_0^2}{4} \left[\xi^2 + 2\xi \left(2\frac{\delta\omega}{\omega_0} \cos \Phi_b + \frac{\Gamma}{\omega_0} \sin \Phi_b \right) \right]}$$

Under this choice of $\tilde{q}$, the square amplitude profile reduces to the desired resonance-antiresonance curve:

$$|\varphi^+|^2 = \psi_b^2 \frac{(\omega^2 - (\omega_0 - \tilde{q})^2)^2 + (\omega\Gamma)^2}{(\omega^2 - \omega_0^2)^2 + (\omega\Gamma)^2} \quad (6.9)$$

with a parameter $\tilde{q}$ that is now a function of the energy shift from the resonance. Which branch of $\tilde{q}_{\pm}$ to choose? One will be physical and one unphysical. Note that if the resonance is not present, namely $\alpha = 0 \implies \xi = 0$, we must recover the simple background amplitude. With $\xi = 0$, $\tilde{q}_{\pm} = -\delta\omega \pm \delta\omega$: if I choose the sign $+$, I obtain $q = 0$, thus simplifying the square amplitude function numerator and denominator: only the background amplitude remains and no resonant features survive. Thus, the physical solution is $\tilde{q}_+$. From now on, let us simplify the notation redefining $q = \tilde{q}_+$, where it is understood that q is not anymore a free parameter, but the $+$ branch of the approximated roots $G(q)$.

The antiresonance peak condition is $\omega = \omega_0 - \tilde{q}(\delta\omega)$, namely $\tilde{q} = -\delta\omega$. This is satisfied

when the argument of the square root is zero:

$$\kappa(\delta\omega) \equiv \delta\omega^2 + \frac{\omega_0^2}{4} \left[\xi^2 + 2\xi \left(2\frac{\delta\omega}{\omega_0} \cos \Phi_b + \frac{\Gamma}{\omega_0} \sin \Phi_b \right) \right] \stackrel{!}{=} 0$$

the roots of this second-degree polynomial are (define the adimensional width $\gamma \equiv \Gamma/\omega_0$):

$$\begin{aligned} \frac{\delta\omega_{\pm}}{\omega_0} &= -\frac{1}{2}\xi \cos \Phi_b \pm \frac{1}{2}\sqrt{-\xi^2 \sin^2 \Phi_b - 2\xi\gamma \sin \Phi_b} \\ &= -\frac{1}{2} \left[\xi \cos \Phi_b \mp \sqrt{\gamma^2 - (\gamma + \xi \sin \Phi_b)^2} \right] \end{aligned}$$

There are here several options, which are specialized to the physical system that we are studying:

- **Zero background dephasing:** if $\Phi_b = 0 \pmod{\pi}$, then only one antiresonance peak appears, since the two branches coincide. In this case, the antiresonance peak is located at

$$\delta\omega_{\text{anti}} = -\frac{1}{2}\xi\omega_0 \quad (6.10)$$

This would provide a direct, simple and very informative information on the relation between the background flux and the resonant amplitude. In addition, another extremely desirable property is attained: $q(\delta\omega)$ becomes a constant equal to

$$q = \frac{1}{2}\omega_0\xi$$

thus, reducing the expression exactly to the phenomenological parametrization, where the shift q was assumed to be constant, with the only approximation of the linear expansion in the parameter δ .

Two “fine-tuning conditions” are necessary: the first is $\Phi_b = 0$; rather than being a fine-tuning, this is more a symptom of a physical mechanism that leads to resonant phase to be defined *as related to the homogeneous solution phase*. The second is the requirement $\xi \sim \gamma$, namely all the expansion parameters are similar in magnitude; note however that is not *strictly* required, it is only something that is incidentally seen in our dataset for a specific resonance; in fact, for the $\ell = m = 1$, $n_{\text{res}} = 0$ resonance, these are different by 2 orders of magnitude. Hence, this condition is quite loose and does not need a physical explanation. If a physical explanation exists for the reason why $\xi \sim \gamma$ loosely, it must come from the fact that all the quantities are related to the same set of physical parameters, and in particular they will be regulated by $\mu_s M \ll 1$, which will act as some global scaling of the expansion parameters.

- **Generic background dephasing:** for a general Φ_b , the antiresonance locations will be either 2 or none depending on the sign of $\gamma \sin \Phi_b$ (+ or - respectively; remember that $\gamma < 0$ since we are in the superradiant regime, but this is a generic treatment). In both cases, the function $q(\delta\omega)$ will distort the amplitude profile, but causing a double dip in the first case, or a potentially mild distortion in the second.

The computed fluxes qualitatively and quantitatively support the first option. The generic occurrence of the second (namely, without requiring special conditions) can be directly discarded from the simple observation that only one clear antiresonance peak is *always*

present, hinting at the fact that the zero background dephasing required by the first option is actually a structural feature. As a strong quantitative confirmation of the first hypothesis, the residuals of the fit obtained from only 4 points are always of the order of $\delta|A|^2/|A|^2 \ll 10^3$, often reaching the precision of 1 part to 10^8 , with these small discrepancies that are imputable to the linear expansion approximation.

In fact, an option is still open for the second hypothesis in the limit $\xi \gg \gamma$, which is actually satisfied in our case study. In this case, real solutions for $\tilde{\omega}_{\pm}$ exist only if $\sin \Phi_b < 2|\gamma|/\xi$ and $\sin \Phi_b > 0$, thus practically an extremely fine-tuned condition. If instead the roots are complex, they will be (for a non-fine-tuned background phase) $\tilde{\omega}_{\pm} \sim -\omega_0\xi/2(\cos \Phi_b \mp i \sin \Phi_b)$, which would show as a “antiresonance ghost”, namely a blunt deformation of the amplitude profile, near $\delta\omega = -\delta\omega_{\text{anti}} \cos \Phi_b$. This is clearly not observed in the produced data; thus, the only option would be a fine-tuned Φ_b , which is less likely than a structurally motivated equality $\Phi_b = 0$.

6.4.b Field amplitude related to the Green’s function method

While the previous approach is completely general, in the sense that it does not require an explicit solution for the field, a more practical approach comes from the direct inspection of the field amplitude at the horizon, which determines the only active energy flux $\dot{E}^- = \omega k_- |Z_{\ell m}^-|^2$, since the flux at infinity is quenched. ω and k_- are “wide-band” functions, thus the resonance comes from $|Z_{\ell m}^-|^2$. This is expected, since resonance profiles often arise from the square modulus of a field amplitude. Let’s trace back $Z^- = R^-(r^* \rightarrow -\infty) \int_{r_+}^{\infty} \frac{R^+(r)J(r)}{W(r)\Delta} dr$. The $R^-(r^* \rightarrow -\infty)$ contribution is normalized by definition (the homogeneous R^- is purely ingoing with unit amplitude at the horizon), thus it does not contribute to the square modulus of Z^- . The resonant amplitude then comes from $R^+(r_p)/W(r_p)$, namely the purely outgoing solution at infinity evaluated at the particle radius r_p , divided by the Wronskian.

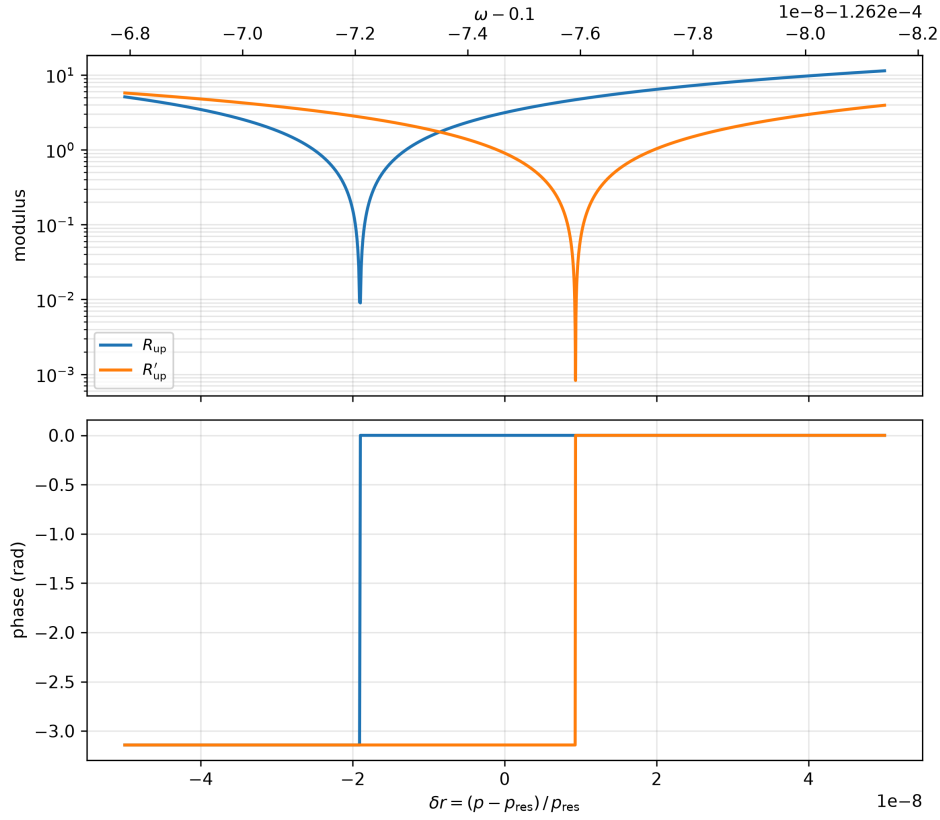
Intuitively, this can be interpreted as follows: we are considering homogeneous solutions with normalized boundary conditions; fixing a normalized boundary condition at infinity (although exponentially declining, since the field is confined) has the effect of quenching the field modulus at finite radial coordinate in correspondence to the resonant frequency. This is because a fixed amplitude at finite radius produces a resonantly enhanced amplitude at the boundary; in this case, however, the boundary conditions are normalized, thus the amplitude at finite radius is *divided* by the standard resonant amplitude function, namely Eq. (6.7) modulated by the resonant phase Eq. (6.8). This produces the “dip structure” observed in this work. At the same time, the Wronskian exhibits a similar dip, but *not at the same radial coordinate*: the physical origin is the same dip that modulates $R_+(r)$, however, the (anti-)resonance here is shifted by the simultaneous effect of the interplay with amplitude and phases of R'_+ , R_- , R'_- , which also enter the Wronskian formula. The Wronskian follows an inverse resonance profile as a consequence of the fact that quasi-normal modes are defined by the poles of the Green’s function in the complex frequency plane, where the two homogeneous solutions become linearly dependent and thus the Wronskian is null. Around the pole frequency $\omega_p = \omega_R + i\omega_I$ the Wronskian can be expanded in a Taylor series as $W \approx W_{,R}(\Re(\omega) - \omega_R) + W_{,I}(\Im(\omega) - \omega_I) = W_{,R}(\omega - \omega_R) - W_{,I}\omega_I$ since the physical ω is real. This implies that on the real axis, near the

pole, $\sqrt{|W|^2} \approx \left[|W_{,R}|^2(\omega - \omega_R)^2 + |W_{,I}|^2\omega_I^2 + 2(\omega - \omega_R)\omega_I \Re(W_{,R}W_{,I}) \right]^{1/2}$, which since in the energy flux is at denominator generates the positive resonance peak. The peak shift of ω_0 from ω_R in this case is due to the extra $(\omega - \omega_R)$ term which is proportional to $\Re(W_{,R}W_{,I})$.

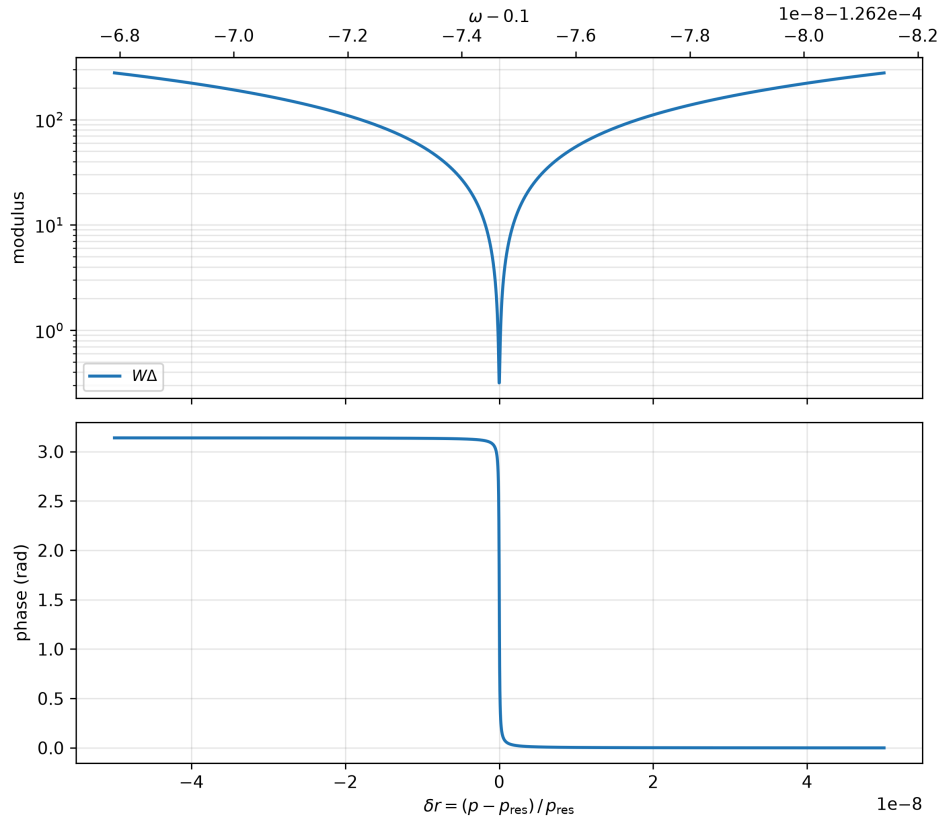
I use the modified version of STORM described in the previous Chapters to numerically visualize the phenomenology described above. In particular, I focus on the magnitude of the scalar field R_{up} and the Wronskian, considering the quantity $W\Delta$ which, as shown in Sec. 4.5, is constant over the radial coordinate at fixed energy ω (thus considering ω as a free parameter, and not related to p through $\Omega_p(p) = \omega$). Fixing $\mu_s = 0.1$, I perform a scan of R_{up} and $W\Delta$ in the radial coordinate p considering, point by point, the energy given by the orbital frequency $\omega = \Omega_p$; the scans of R_{up} , its radial derivative R'_{up} and $W\Delta$ in a tight interval around the resonance are shown, respectively, in Fig. 6.8a and 6.8b. The scalar field and its derivative display a resonantly damped behavior as motivated before, centered in different radial coordinates; the interaction of these resonances with the nearly constant R_{in} , R'_{in} finally produces a Wronskian with a shifted antiresonance; finally, the amplitude proportional to $R_{\text{up}}/(W\Delta)$ inherits the antiresonance from R_{up} and the resonance from W^{-1} .

A wider scan, with otherwise the same plotting conventions, is presented in Fig. 6.9. Here is well visible the radiative zone, where $\omega - \mu_s > 0$, on the left of the vertical green line. Regarding the phase, it has been obtained by unwrapping the dephasing between consecutive points of the radial sampling: for this reason, around $\omega = \mu_s$, where the magnitude and phase evolution becomes singular from both sides, the sampling interval is not sufficient to correctly unwrap the phase, leading to artifacts. In addition, since the phases of the radiative and confined zones are separated by such singularity, at the two sides of $\omega = \mu_s$ the phase is always referred to the data points at the extremes of the whole sampling interval. In the region $\omega < \mu_s$ it is well visible the resonant structure: starting from lower energy (higher radius), the field amplitude and phase is determined by the series of resonances (the theoretical prediction for the position of the resonances, Eq. (6.1), are shown with red lines for $n_{\text{res}} = 0, \dots, 6$), which has the effect of increasing the slope of the field magnitude towards higher energies (analogously to a Bode plot, and taking into account that the field is described by the inverse of a resonant profile), and the phase accumulates a π dephasing in correspondence of each resonance. The same qualitative structure characterizes R'_{up} and is inherited by $W\Delta$.

Note that the perspective presented in this Section is conceptually different (and independent) from the one presented before: for the interpretation presented previously, the structure is motivated by an overlap of two distinct phenomena (a wide-band solution and a resonant discrete spectrum) regardless of the particular mathematical treatise of the energy flux; here instead we are considering the explicit energy flux formula, and tracing the origin of the resonance's features to the amplitude given by the Green's function method, which incorporates the field amplitudes in a non-trivial way. Both approaches give a correct interpretation of how the resonant structure arises based on different physical details of the problem.



(a) Scalar field R_{up} and its radial derivative R'_{up} .



(b) The quantity $W\Delta$, alias for the Wronskian W .

Figure 6.8: Scalar field (upper figure) and Wronskian $W\Delta$ (lower figure) for $\mu_s = 0.1$ around the resonance $n_{\text{res}} = 0$ for $\ell = m = 1$, represented with amplitude and phase. The horizontal axis represents either the radial coordinate, centered and normalized to p_{res} , the resonance peak coordinate, or the energy $\omega = \Omega_p(p)$, referenced to the scalar mass μ_s .

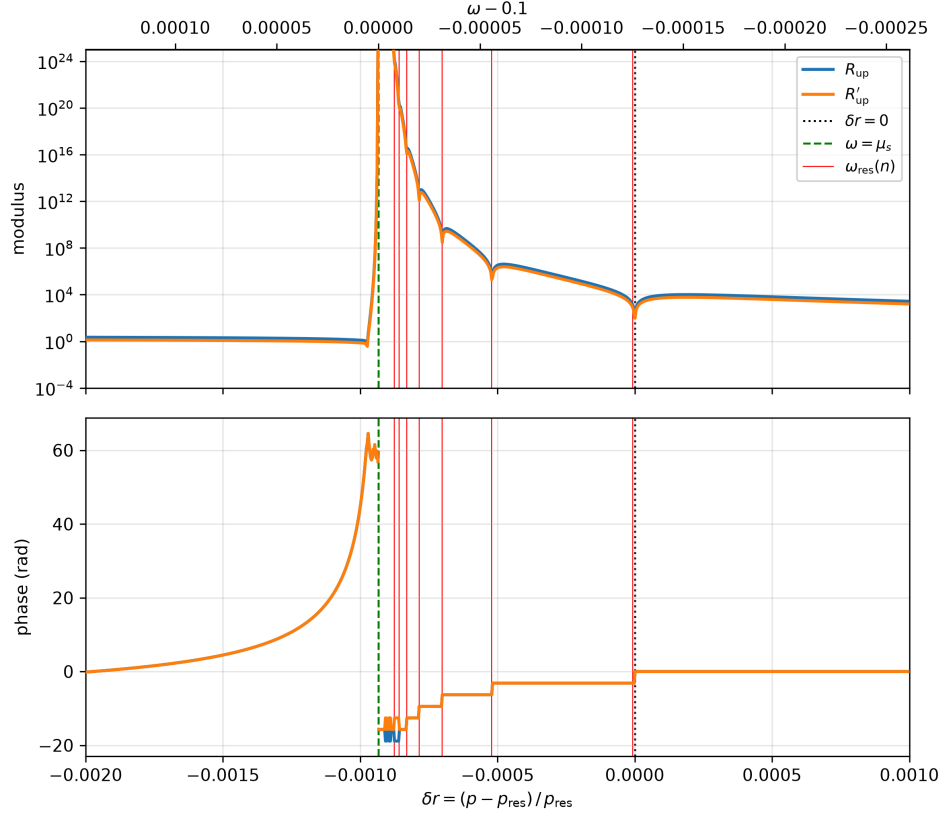
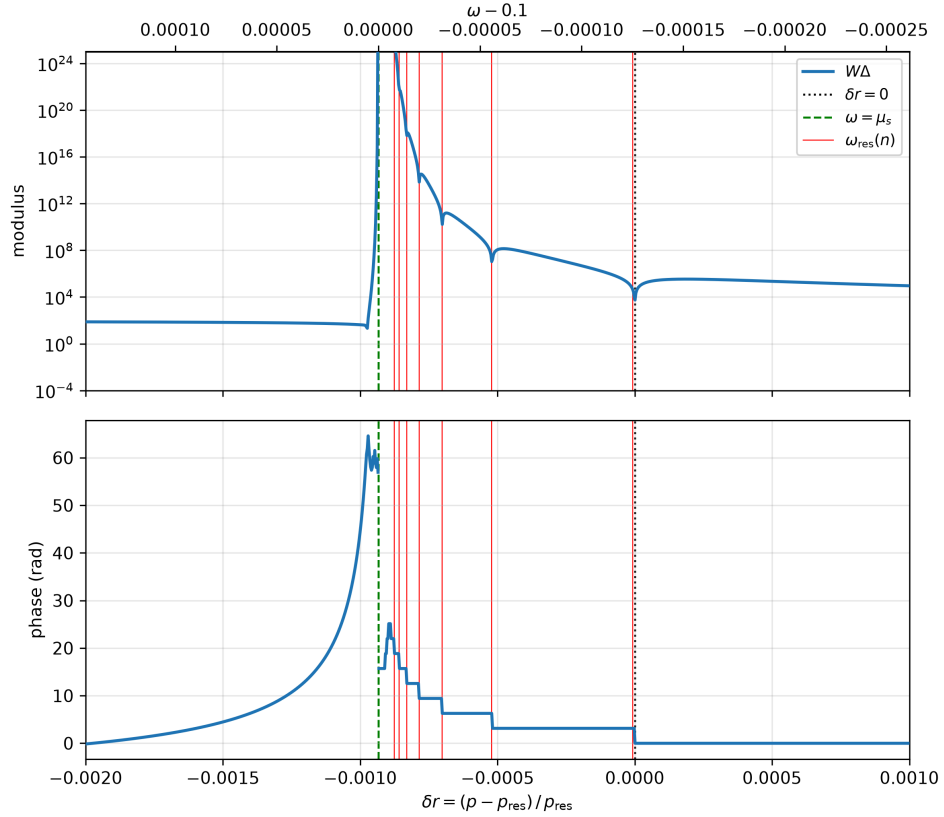

 (a) Scalar field R_{up} and its radial derivative R'_{up} .

 (b) The quantity $W\Delta$, alias for the Wronskian W .

Figure 6.9: Same of Fig. 6.8 for a wider radial interval, across radiative regime ($\omega - \mu_s > 0$, at the left of the vertical green line), and the series of discrete resonances (theoretical prediction Eq. 6.1 marked with vertical red lines).

Conclusions

In this Thesis I developed theoretical and numerical tools to characterize the phenomenology associated with the presence of a free massive scalar field coupled with the secondary compact object of an EMRI. The general treatise is based on the adiabatic approximation with a first order expansion in the mass ratio q . The scalar field's equation of motion is obtained from the action and is represented by a Klein-Gordon equation in a curved (Kerr) background, with a point-like source term corresponding to the secondary; it can be solved using the Green's function method, once the homogeneous solutions are known.

Although not exploited in the numerical implementation of the results, I presented a generalization of the semi-analytical MST method to compute the homogeneous solution defined by purely ingoing boundary conditions at the horizon. The solution consists in a generalization of the original MST solution, obtained through a summation of hypergeometric functions weighted through a series of coefficients determined by a three-term recurrence relation. I proved that a similar minimal modification to the massless MST does not exist for the homogeneous solution with purely outgoing boundary conditions at spatial infinity, and proposed a promising path for a non-minimal generalization. The development of such a solution and a working implementation of the massive MST method will be conducted in future work.

I produced a modified version of the code STORM to compute the solution of the massive field equations for a generic orbital configuration and obtain, mode-by-mode, the energy fluxes at horizon and spatial infinity. This implementation of the code, easily parallelizable and executed on the LIGO/IGWN GWave Cluster, enabled a systematic characterization of the scalar field phenomenology. For circular equatorial orbits, I computed the energy fluxes, GW waveform dephasing, and Fisher matrix observability estimations with the LISA GW detector, and interpreted the results in light of superradiance and massive field confinement. Covariance between the scalar charge and the scalar mass is observed, making it difficult to constrain the parameters of the scalar field; however, through a series of physically motivated case study examples, I found the concrete possibility of precisely measuring such parameters under the condition that the observable evolution happens sufficiently slowly across the transition between superradiant and radiative behavior. The availability of a accurate and efficient computational infrastructure will enable a wider exploration of the phenomenology for generic orbits, which will break the degeneracies and facilitate parameter estimation.

Finally, I explored the phenomenology of superradiant resonances and floating orbits associated with the massive nature of the scalar field. For the first time, a complete mapping of the main resonance was performed for high values of eccentricity and inclina-

tion, which were considered both separately and simultaneously. This opens the possibility of testing for the instability of floating orbits against small orbital perturbations, which could produce drastic changes in the evolution, leading to unmistakable observational signatures for the presence of a massive scalar field. Furthermore, I identified and parametrized a relevant antiresonant feature in the resonance shape, and explained its physical and mathematical origin.

In conclusion, EMRIs will represent an unprecedented opportunity to probe scalar fields and in particular ultralight scalar mediators of gravity predicted by many beyond-GR theories. The observational signatures of a massive scalar field could be detectable via GW observations in the LISA era through slight modifications of the orbital dynamics due to the presence of the additional dissipative (or superradiant) channel continuously acting throughout the entire evolution, and/or discrete features, such as floating orbits, which affect the dynamics with punctual but extremely impactful modifications to the pure-GR evolution. In this Thesis I provided an efficient computational infrastructure to investigate quantitatively these effects, uncovering a rich phenomenology, expanding the present knowledge of the possible observables and interpreting them with analytical methods or, when it was not possible, heuristic but physically motivated arguments. This work opens many avenues for future research, from the most mathematically-oriented ones such as the search for a complete massive MST solution to the more observationally motivated ones, such as the construction of wide and dense grids of energy fluxes across the whole parameter space, parameter estimation forecasts for generic orbits, dynamical evolution of resonant orbits and the study of the corresponding GW waveforms.

A

Coefficients of the Frobenius series for ingoing solution

I report the explicit expressions to compute the coefficients of the Frobenius series for outgoing solution up to order 10. Here $\sqrt{\omega^2 - \mu^2} \equiv \sigma$. The coefficients for the ingoing solution at the horizon are not reported, being much longer.

$$d_0 = 1 \quad (\text{A.1})$$

$$d_1 = \frac{i}{2\sigma^3} \left(\lambda(\omega^2 - \mu^2) + iM\mu^2\sigma + 2am\omega(\omega^2 - \mu^2) - M^2(-5\mu^4 + 4\mu^2\omega^2) \right) \quad (\text{A.2})$$

$$d_2 = \frac{i}{4\sigma^3} \left(-4M^3\mu^4 + (-2 + \lambda + 2am\omega)(\omega^2 - \mu^2)d_1 + M\sigma(2\sigma - 2\lambda\sigma + 2a^2\mu^2\sigma + 7i\mu^2d_1 - 4i\omega^2d_1) - M^2\mu^2(6i\sigma + (4\omega^2 - 5\mu^2)d_1) \right) \quad (\text{A.3})$$

$$d_3 = \frac{i}{6\sigma^3} \left(2a^3m(\mu - \omega)\omega(\mu + \omega) + 4M^2((1 + 2iM\sigma)\mu^2 + M^2\mu^4 - \omega^2) - 2M((6 + 7iM\sigma - \lambda)\mu^2 + 2M^2\mu^4 + (-6 + \lambda)\omega^2)d_1 - a^2(\mu - \omega)(\mu + \omega)(1 - m^2 - \lambda + 2M\mu^2(2M + d_1)) + 2am\omega(-\mu + \omega)(\mu + \omega)d_2 + (\mu^2(6 + 13iM\sigma - \lambda + 5M^2\mu^2) + (-6 + \lambda - 4M(2i\sigma + M\mu^2))\omega^2)d_2 \right) \quad (\text{A.4})$$

$$d_4 = \frac{i}{8\sigma^3} \left(2a^4M\mu^2(\mu - \omega)(\mu + \omega) + 2a^3m(\mu - \omega)\omega(\mu + \omega)d_1 + 4M^2((4 + 4iM\sigma)\mu^2 + M^2\mu^4 - 4\omega^2)d_1 - 2M((15 + 11iM\sigma - \lambda)\mu^2 + 2M^2\mu^4 + (-15 + \lambda)\omega^2)d_2 + a^2(4M^3\mu^4 + (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_1 + 2M^2\mu^2(5i\sigma + 2(-\mu^2 + \omega^2)d_1) - 2M(\mu - \omega)(\mu + \omega)(-6 + 2\lambda + 2i\sigma d_1 + \mu^2d_2)) + 2am\omega(-\mu + \omega)(\mu + \omega)d_3 + (\mu^2(12 + 19iM\sigma - \lambda + 5M^2\mu^2) + (-12 + \lambda - 4M(3i\sigma + M\mu^2))\omega^2)d_3 \right) \quad (\text{A.5})$$

$$d_5 = -\frac{i}{10\sigma^3} \left(2a^5m(\mu - \omega)\omega(\mu + \omega) - a^4(\mu - \omega)(\mu + \omega)(2 - 2m^2 - \lambda + 2M\mu^2(2M + d_1)) + 2a^3m\omega(-\mu + \omega)(\mu + \omega)d_2 + 2M(-2M((9 + 6iM\sigma)\mu^2 + M^2\mu^4 - 9\omega^2)d_2 + ((28 + 15iM\sigma - \lambda)\mu^2 + 2M^2\mu^4 + (-28 + \lambda)\omega^2)d_3) + a^2(8M^4\mu^4 + M^3(24iM\sigma\mu^2 - 4\mu^4d_1) - (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_2 + 2M^2(-12\omega^2 + \mu^2(12 - 9i\sigma d_1 + 2(-\mu^2 + \omega^2)d_2)) + 2M(\mu - \omega)(\mu + \omega)((-13 + 2\lambda)d_1 + 4i\sigma d_2 + \mu^2d_3)) + 2am(\mu - \omega)\omega(\mu + \omega)d_4 + (\mu^2(-20 + \lambda - 5M(5i\sigma + M\mu^2)) + (20 - \lambda + 4M(4i\sigma + M\mu^2))\omega^2)d_4 \right) \quad (\text{A.6})$$

$$\begin{aligned}
 d_6 = & -\frac{i}{12\sigma^3} \left(2a^6 M \mu^2 (\mu - \omega)(\mu + \omega) + 2a^5 m (\mu - \omega) \omega (\mu + \omega) d_1 \right. \\
 & + a^4 (4M^3 \mu^4 + (-2 + 2m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_1 + 2M^2 \mu^2 (7i\sigma + 2(-\mu^2 + \omega^2) d_1) \\
 & - 2M(\mu - \omega)(\mu + \omega)(-15 + 3\lambda + 2i\sigma d_1 + \mu^2 d_2)) \\
 & + 2a^3 m \omega (-\mu + \omega)(\mu + \omega) d_3 \\
 & + 2M(-2M(8(2 + iM\sigma)\mu^2 + M^2 \mu^4 - 16\omega^2) d_3 + ((45 + 19iM\sigma - \lambda)\mu^2 + 2M^2 \mu^4 + (-45 + \lambda)\omega^2) d_4) \\
 & + a^2 (8M^4 \mu^4 d_1 + M^3 (40iM\sigma \mu^2 d_1 - 4\mu^4 d_2) - (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_3 \\
 & + M^2 (-56\omega^2 d_1 + 4\mu^4 d_3 + \mu^2 (56d_1 - 26i\sigma d_2 - 4\omega^2 d_3)) \\
 & + 2M(\mu - \omega)(\mu + \omega)(2(-12 + \lambda) d_2 + 6i\sigma d_3 + \mu^2 d_4)) \\
 & + 2am(\mu - \omega) \omega (\mu + \omega) d_5 \\
 & \left. + (\mu^2 (-30 - 31iM\sigma + \lambda - 5M^2 \mu^2) + (30 - \lambda + 4M(5i\sigma + M\mu^2)) \omega^2) d_5 \right) \quad (A.7)
 \end{aligned}$$

$$\begin{aligned}
 d_7 = & \frac{i}{14\sigma^3} \left(2a^7 m (\mu - \omega) \omega (\mu + \omega) - a^6 (\mu - \omega)(\mu + \omega)(3 - 3m^2 - \lambda + 2M\mu^2 (2M + d_1)) \right. \\
 & + 2a^5 m \omega (-\mu + \omega)(\mu + \omega) d_2 \\
 & + a^4 (12M^4 \mu^4 + M^3 (48i\sigma \mu^2 - 4\mu^4 d_1) - (-2 + 2m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_2 \\
 & + M^2 (-72\omega^2 + 2\mu^2 (36 - 11i\sigma d_1 + 2(\mu - \omega)(\mu + \omega) d_2)) \\
 & + 2M(\mu - \omega)(\mu + \omega)(3(-8 + \lambda) d_1 + 4i\sigma d_2 + \mu^2 d_3)) \\
 & + 2a^3 m (\mu - \omega) \omega (\mu + \omega) d_4 + 4M^2 (5(5 + 2iM\sigma) \mu^2 + M^2 \mu^4 - 25\omega^2) d_4 \\
 & - 2M((66 + 23iM\sigma - \lambda) \mu^2 + 2M^2 \mu^4 + (-66 + \lambda) \omega^2) d_5 \\
 & + a^2 (-8M^4 \mu^4 d_2 + 4M^3 (-14i\sigma \mu^2 d_2 + \mu^4 d_3) + (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_4 \\
 & + 2M^2 (52\omega^2 d_2 - 2\mu^4 d_4 + \mu^2 (-52d_2 + 17i\sigma d_3 + 2\omega^2 d_4)) \\
 & - 2M(\mu - \omega)(\mu + \omega)((-39 + 2\lambda) d_3 + 8i\sigma d_4 + \mu^2 d_5)) + 2am\omega(-\mu + \omega)(\mu + \omega) d_6 \\
 & \left. + (\mu^2 (42 + 37iM\sigma - \lambda + 5M^2 \mu^2) + (-42 + \lambda - 4M(6i\sigma + M\mu^2)) \omega^2) d_6 \right) \quad (A.8)
 \end{aligned}$$

$$\begin{aligned}
 d_8 = & \frac{i}{16\sigma^3} \left(2a^8 M \mu^2 (\mu - \omega)(\mu + \omega) + 2a^7 m (\mu - \omega) \omega (\mu + \omega) d_1 \right. \\
 & + a^6 (4M^3 \mu^4 + (-3 + 3m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_1 + 2M^2 \mu^2 (9i\sigma + 2(-\mu^2 + \omega^2) d_1) \\
 & - 2M(\mu - \omega)(\mu + \omega)(-28 + 4\lambda + 2i\sigma d_1 + \mu^2 d_2)) \\
 & + 2a^5 m \omega (-\mu + \omega)(\mu + \omega) d_3 \\
 & + a^4 (12M^4 \mu^4 d_1 + M^3 (72i\sigma \mu^2 d_1 - 4\mu^4 d_2) - (-2 + 2m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_3 \\
 & + 2M^2 (-66\omega^2 d_1 + 2\mu^4 d_3 + \mu^2 (66d_1 - 15i\sigma d_2 - 2\omega^2 d_3)) \\
 & + 2M(\mu - \omega)(\mu + \omega)((-37 + 3\lambda) d_2 + 6i\sigma d_3 + \mu^2 d_4)) \\
 & + 2a^3 m (\mu - \omega) \omega (\mu + \omega) d_5 + 4M^2 (12(3 + iM\sigma) \mu^2 + M^2 \mu^4 - 36\omega^2) d_5 \\
 & - 2M((91 + 27iM\sigma - \lambda) \mu^2 + 2M^2 \mu^4 + (-91 + \lambda) \omega^2) d_6 \\
 & + a^2 (-8M^4 \mu^4 d_3 + 4M^3 (-18i\sigma \mu^2 d_3 + \mu^4 d_4) + (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega) d_5 \\
 & + 2M^2 (84\omega^2 d_3 - 2\mu^4 d_5 + \mu^2 (-84d_3 + 21i\sigma d_4 + 2\omega^2 d_5)) \\
 & - 2M(\mu - \omega)(\mu + \omega)(2(-29 + \lambda) d_4 + 10i\sigma d_5 + \mu^2 d_6)) + 2am\omega(-\mu + \omega)(\mu + \omega) d_7 \\
 & \left. + (\mu^2 (56 + 43iM\sigma - \lambda + 5M^2 \mu^2) + (-56 + \lambda - 4M(7i\sigma + M\mu^2)) \omega^2) d_7 \right) \quad (A.9)
 \end{aligned}$$

$$\begin{aligned}
d_9 = & \frac{i}{18\sigma^3} \left(2a^9 m\omega(-\mu^2 + \omega^2) + a^8(\mu - \omega)(\mu + \omega)(4 - 4m^2 - \lambda + 2M\mu^2(2M + d_1)) \right. \\
& + 2a^7 m(\mu - \omega)\omega(\mu + \omega)d_2 \\
& + a^6(-16M^4\mu^4 + 4M^3(-20i\sigma\mu^2 + \mu^4d_1) + (-3 + 3m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_2 \\
& + 2M^2(80\omega^2 + \mu^2(-80 + 13i\sigma d_1 + 2(-\mu^2 + \omega^2)d_2)) \\
& - 2M(\mu - \omega)(\mu + \omega)((-39 + 4\lambda)d_1 + 4i\sigma d_2 + \mu^2d_3)) + 2a^5 m\omega(-\mu + \omega)(\mu + \omega)d_4 \\
& + a^4(12M^4\mu^4d_2 + M^3(96i\sigma\mu^2d_2 - 4\mu^4d_3) - (-2 + 2m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_4 \\
& + M^2(-216\omega^2d_2 + 4\mu^4d_4 + \mu^2(216d_2 - 38i\sigma d_3 - 4\omega^2d_4)) \\
& + 2M(\mu - \omega)(\mu + \omega)(3(-18 + \lambda)d_3 + 8i\sigma d_4 + \mu^2d_5)) + 2a^3 m(\mu - \omega)\omega(\mu + \omega)d_6 \\
& + 4M^2(7(7 + 2iM\sigma)\mu^2 + M^2\mu^4 - 49\omega^2)d_6 \\
& - 2M((120 + 31iM\sigma - \lambda)\mu^2 + 2M^2\mu^4 + (-120 + \lambda)\omega^2)d_7 \\
& + a^2(-8M^4\mu^4d_4 + 4M^3(-22i\sigma\mu^2d_4 + \mu^4d_5) + (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_6 \\
& + M^2(248\omega^2d_4 - 4\mu^4d_6 + \mu^2(-248d_4 + 50i\sigma d_5 + 4\omega^2d_6)) \\
& - 2M(\mu - \omega)(\mu + \omega)((-81 + 2\lambda)d_5 + 12i\sigma d_6 + \mu^2d_7)) + 2am\omega(-\mu + \omega)(\mu + \omega)d_8 \\
& \left. + (\mu^2(72 + 49iM\sigma - \lambda + 5M^2\mu^2) + (-72 + \lambda - 4M(8i\sigma + M\mu^2))\omega^2)d_8 \right) \quad (A.10)
\end{aligned}$$

$$\begin{aligned}
d_{10} = & \frac{i}{20\sigma^3} \left(2a^{10} M\mu^2(\omega^2 - \mu^2) + 2a^9 m\omega(\omega^2 - \mu^2)d_1 \right. \\
& - a^8(4M^3\mu^4 + (-4 + 4m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_1 + 2M^2\mu^2(11i\sigma + 2(\omega^2 - \mu^2)d_1) \\
& + 2M(\omega^2 - \mu^2)(-45 + 5\lambda + 2i\sigma d_1 + \mu^2d_2)) + 2a^7 m(\mu - \omega)\omega(\mu + \omega)d_3 \\
& + a^6(-16M^4\mu^4d_1 + 4M^3(-28i\sigma\mu^2d_1 + \mu^4d_2) + (-3 + 3m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_3 \\
& + M^2(256\omega^2d_1 - 4\mu^4d_3 + \mu^2(-256d_1 + 34i\sigma d_2 + 4\omega^2d_3)) \\
& + 2M(\omega^2 - \mu^2)((-54 + 4\lambda)d_2 + 6i\sigma d_3 + \mu^2d_4)) + 2a^5 m\omega(\omega^2 - \mu^2)d_5 \\
& + a^4(12M^4\mu^4d_3 + M^3(120i\sigma\mu^2d_3 - 4\mu^4d_4) - (-2 + 2m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_5 \\
& + M^2(-324\omega^2d_3 + 4\mu^4d_5 + \mu^2(324d_3 - 46i\sigma d_4 - 4\omega^2d_5)) \\
& + 2M(\mu - \omega)(\mu + \omega)(3(-25 + \lambda)d_4 + 10i\sigma d_5 + \mu^2d_6)) \\
& + 2a^3 m(\mu - \omega)\omega(\mu + \omega)d_7 + 4M^2(16(4 + iM\sigma)\mu^2 + M^2\mu^4 - 64\omega^2)d_7 \\
& - 2M((153 + 35iM\sigma - \lambda)\mu^2 + 2M^2\mu^4 + (-153 + \lambda)\omega^2)d_8 \\
& + a^2(-8M^4\mu^4d_5 + 4M^3(-26i\sigma\mu^2d_5 + \mu^4d_6) + (-1 + m^2 + \lambda)(\mu - \omega)(\mu + \omega)d_7 \\
& + M^2(344\omega^2d_5 - 4\mu^4d_7 + \mu^2(-344d_5 + 58i\sigma d_6 + 4\omega^2d_7)) \\
& - 2M(\mu - \omega)(\mu + \omega)(2(-54 + \lambda)d_6 + 14i\sigma d_7 + \mu^2d_8)) + 2am\omega(-\mu + \omega)(\mu + \omega)d_9 \\
& \left. + (\mu^2(90 + 55iM\sigma - \lambda + 5M^2\mu^2) + (-90 + \lambda - 4M(9i\sigma + M\mu^2))\omega^2)d_9 \right) \quad (A.11)
\end{aligned}$$

Bibliography

- Aasi, J. et al. (2015). “Advanced LIGO”. In: *Class. Quant. Grav.* 32, p. 074001. DOI: [10.1088/0264-9381/32/7/074001](#). arXiv: [1411.4547 \[gr-qc\]](#).
- Abac, A. G. et al. (May 2026a). “GWTC-5.0: Methods for Identifying and Characterizing Gravitational-wave Transients”. In: arXiv: [2605.27224 \[gr-qc\]](#).
- (July 2026b). “GWTC-5.0: Tests of General Relativity”. In: arXiv: [2607.19293 \[gr-qc\]](#).
- Abac, Adrian et al. (2026). “The Science of the Einstein Telescope”. In: *JCAP* 03, p. 081. DOI: [10.1088/1475-7516/2026/03/081](#). arXiv: [2503.12263 \[gr-qc\]](#).
- Abac, None et al. (May 2026). “GWTC-5.0: An Introduction to Version 5.0 of the Gravitational-Wave Transient Catalog”. In: arXiv: [2605.27223 \[gr-qc\]](#).
- Abbott, B. P. et al. (2016). “Observation of Gravitational Waves from a Binary Black Hole Merger”. In: *Phys. Rev. Lett.* 116.6, p. 061102. DOI: [10.1103/PhysRevLett.116.061102](#). arXiv: [1602.03837 \[gr-qc\]](#).
- (2017). “GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral”. In: *Phys. Rev. Lett.* 119.16, p. 161101. DOI: [10.1103/PhysRevLett.119.161101](#). arXiv: [1710.05832 \[gr-qc\]](#).
- Abe, Mahiro et al. (2021). “Matter-wave Atomic Gradiometer Interferometric Sensor (MAGIS-100)”. In: *Quantum Sci. Technol.* 6.4, p. 044003. DOI: [10.1088/2058-9565/abf719](#). arXiv: [2104.02835 \[physics.atom-ph\]](#).
- Acernese, F. et al. (2015). “Advanced Virgo: a second-generation interferometric gravitational wave detector”. In: *Class. Quant. Grav.* 32.2, p. 024001. DOI: [10.1088/0264-9381/32/2/024001](#). arXiv: [1408.3978 \[gr-qc\]](#).
- Ajith, Parameswaran et al. (2025). “The Lunar Gravitational-wave Antenna: mission studies and science case”. In: *JCAP* 01, p. 108. DOI: [10.1088/1475-7516/2025/01/108](#). arXiv: [2404.09181 \[gr-qc\]](#).
- Alsing, Justin et al. (2012). “Gravitational radiation from compact binary systems in the massive Brans-Dicke theory of gravity”. In: *Phys. Rev. D* 85, p. 064041. DOI: [10.1103/PhysRevD.85.064041](#). arXiv: [1112.4903 \[gr-qc\]](#).
- Amaro-Seoane, Pau et al. (Feb. 2017). “Laser Interferometer Space Antenna”. In: arXiv: [1702.00786 \[astro-ph.IM\]](#).
- Amaro-Seoane, Pau et al. (2015). “Research Update on Extreme-Mass-Ratio Inspirals”. In: *J. Phys. Conf. Ser.* 610.1. Ed. by Giacomo Ciani, John W. Conklin, and Guido Mueller, p. 012002. DOI: [10.1088/1742-6596/610/1/012002](#). arXiv: [1410.0958 \[astro-ph.CO\]](#).
- Apostolatos, T. A. (1995). “Search templates for gravitational waves from precessing, inspiraling binaries”. In: *Phys. Rev. D* 52, pp. 605–620. DOI: [10.1103/PhysRevD.52.605](#).

- Armano, M. et al. (2016). “Sub-Femto- g Free Fall for Space-Based Gravitational Wave Observatories: LISA Pathfinder Results”. In: *Phys. Rev. Lett.* 116.23, p. 231101. doi: [10.1103/PhysRevLett.116.231101](#).
- Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics*. Vol. 60. Graduate Texts in Mathematics. Springer. doi: [10.1007/978-1-4757-2063-1](#).
- Arun, K G and Archana Pai (2013). “Tests of General Relativity and Alternative theories of gravity using Gravitational Wave observations”. In: *Int. J. Mod. Phys. D* 22, p. 1341012. doi: [10.1142/S0218271813410125](#). arXiv: [1302.2198 \[gr-qc\]](#).
- Astone, P. et al. (1991). “First cooling below 0.1-K of the new gravitational wave antenna ‘Nautilus’ of the Rome group”. In: *EPL* 16, pp. 231–235. doi: [10.1209/0295-5075/16/3/002](#).
- Banerjee, N. and S. Sen (1998). “No scalar hair theorem for a charged spherical black hole”. In: *Phys. Rev. D* 58, p. 104024. doi: [10.1103/PhysRevD.58.104024](#). arXiv: [gr-qc/9803019](#).
- Barausse, Enrico et al. (2020). “Prospects for Fundamental Physics with LISA”. In: *Gen. Rel. Grav.* 52.8, p. 81. doi: [10.1007/s10714-020-02691-1](#). arXiv: [2001.09793 \[gr-qc\]](#).
- Barsanti, Susanna (Jan. 2024). “Testing general relativity with gravitational waves from extreme mass-ratio inspirals”. PhD thesis. Università degli Studi di Roma “La Sapienza”, Italy, Rome U.
- Barsanti, Susanna et al. (2023). “Detecting Massive Scalar Fields with Extreme Mass-Ratio Inspirals”. In: *Phys. Rev. Lett.* 131.5, p. 051401. doi: [10.1103/PhysRevLett.131.051401](#). arXiv: [2212.03888 \[gr-qc\]](#).
- Baumann, Daniel et al. (2022a). “Ionization of gravitational atoms”. In: *Phys. Rev. D* 105.11, p. 115036. doi: [10.1103/PhysRevD.105.115036](#). arXiv: [2112.14777 \[gr-qc\]](#).
- (2022b). “Sharp Signals of Boson Clouds in Black Hole Binary Inspirals”. In: *Phys. Rev. Lett.* 128.22, p. 221102. doi: [10.1103/PhysRevLett.128.221102](#). arXiv: [2206.01212 \[gr-qc\]](#).
- Bekenstein, J. D. (1972). “Nonexistence of baryon number for black holes. ii”. In: *Phys. Rev. D* 5, pp. 2403–2412. doi: [10.1103/PhysRevD.5.2403](#).
- Benetti, Giovanni, Koustav Chandra, and Bangalore S. Sathyaprakash (Feb. 2026a). “Tarnished by Tools: Cost of Systematics in Golden Dark Siren Cosmology”. In: arXiv: [2602.14898 \[gr-qc\]](#).
- Benetti, Giovanni et al. (2026b). “Observing double white dwarfs with the Lunar GW Antenna”. In: *Astron. Astrophys.* 708, A277. doi: [10.1051/0004-6361/202556940](#). arXiv: [2509.07849 \[gr-qc\]](#).
- Benone, Carolina L. et al. (2014). “Kerr-Newman scalar clouds”. In: *Phys. Rev. D* 90.10, p. 104024. doi: [10.1103/PhysRevD.90.104024](#). arXiv: [1409.1593 \[gr-qc\]](#).
- Bergé, Joel (2017). “The Inverse Square Law And Newtonian Dynamics space explorer (ISLAND)”. In: *52nd Rencontres de Moriond on Gravitation*, pp. 191–198. arXiv: [1809.00698 \[gr-qc\]](#).
- Bergmann, Peter G. (1968). “Comments on the scalar tensor theory”. In: *Int. J. Theor. Phys.* 1, pp. 25–36. doi: [10.1007/BF00668828](#).
- Berti, Emanuele et al. (2015). “Testing General Relativity with Present and Future Astrophysical Observations”. In: *Class. Quant. Grav.* 32, p. 243001. doi: [10.1088/0264-9381/32/24/243001](#). arXiv: [1501.07274 \[gr-qc\]](#).

- (Feb. 2026). “deci-Hz Gravitational Wave Observations on the Moon and Beyond”. In: arXiv: 2602.05923 [gr-qc].
- Bhattacharyya, Soham (2019). “Distinguishing general relativity and modified theories of gravity using quasinormal modes”. PhD thesis. IISER, Trivandrum.
- Bogdanovic, Tamara, M. Coleman Miller, and Laura Blecha (2022). “Electromagnetic counterparts to massive black-hole mergers”. In: *Living Rev. Rel.* 25.1, p. 3. DOI: 10.1007/s41114-022-00037-8. arXiv: 2109.03262 [astro-ph.HE].
- Bonvin, Camille et al. (2017). “Effect of matter structure on the gravitational waveform”. In: *Phys. Rev. D* 95.4, p. 044029. DOI: 10.1103/PhysRevD.95.044029. arXiv: 1609.08093 [astro-ph.CO].
- Borhanian, Ssohrab and B. S. Sathyaprakash (2024). “Listening to the Universe with next generation ground-based gravitational-wave detectors”. In: *Phys. Rev. D* 110.8, p. 083040. DOI: 10.1103/PhysRevD.110.083040. arXiv: 2202.11048 [gr-qc].
- Bošković, Mateja and Nikola Savić (July 2026). “Perturbing Gravitational Atoms: Negative Love, Resonant Tides and Shifted Resonances”. In: arXiv: 2607.27174 [gr-qc].
- Branchesi, Marica et al. (2023). “Science with the Einstein Telescope: a comparison of different designs”. In: *JCAP* 07, p. 068. DOI: 10.1088/1475-7516/2023/07/068. arXiv: 2303.15923 [gr-qc].
- Brink, Jeandrew, Marisa Geyer, and Tanja Hinderer (2015a). “Astrophysics of resonant orbits in the Kerr metric”. In: *Phys. Rev. D* 91.8, p. 083001. DOI: 10.1103/PhysRevD.91.083001. arXiv: 1501.07728 [gr-qc].
- (2015b). “Orbital resonances around Black holes”. In: *Phys. Rev. Lett.* 114.8, p. 081102. DOI: 10.1103/PhysRevLett.114.081102. arXiv: 1304.0330 [gr-qc].
- Brito, Richard, Vitor Cardoso, and Paolo Pani (2015). “Superradiance: New Frontiers in Black Hole Physics”. In: *Lect. Notes Phys.* 906, pp.1–237. DOI: 10.1007/978-3-319-19000-6. arXiv: 1501.06570 [gr-qc].
- Buscicchio, Riccardo et al. (2025). “Test for LISA foreground Gaussianity and stationarity: galactic white-dwarf binaries”. In: *Eur. Phys. J. C* 85.8, p. 887. DOI: 10.1140/epjc/s10052-025-14616-w. arXiv: 2410.08263 [astro-ph.HE].
- Capozziello, Salvatore and Mariafelicia De Laurentis (2011). “Extended Theories of Gravity”. In: *Phys. Rept.* 509, pp. 167–321. DOI: 10.1016/j.physrep.2011.09.003. arXiv: 1108.6266 [gr-qc].
- Cardoso, Vitor et al. (2011). “Floating and sinking: The Imprint of massive scalars around rotating black holes”. In: *Phys. Rev. Lett.* 107, p. 241101. DOI: 10.1103/PhysRevLett.107.241101. arXiv: 1109.6021 [gr-qc].
- Carter, Brandon (1968). “Global structure of the Kerr family of gravitational fields”. In: *Phys. Rev.* 174, pp. 1559–1571. DOI: 10.1103/PhysRev.174.1559.
- Cerdonio, M. et al. (1997). “The Ultracryogenic gravitational wave detector AURIGA”. In: *Class. Quant. Grav.* 14, pp. 1491–1494. DOI: 10.1088/0264-9381/14/6/016.
- Chapman-Bird, Christian E. A. et al. (2025). “Efficient waveforms for asymmetric-mass eccentric equatorial inspirals into rapidly spinning black holes”. In: *Phys. Rev. D* 112.10, p. 104023. DOI: 10.1103/scbp-75pf. arXiv: 2506.09470 [gr-qc].
- Chatziioannou, Katerina et al. (2017). “Constructing Gravitational Waves from Generic Spin-Precessing Compact Binary Inspirals”. In: *Phys. Rev. D* 95.10, p. 104004. DOI: 10.1103/PhysRevD.95.104004. arXiv: 1703.03967 [gr-qc].

- Chua, Alvin J. K. et al. (2021). “Rapid generation of fully relativistic extreme-mass-ratio-inspiral waveform templates for LISA data analysis”. In: *Phys. Rev. Lett.* 126.5, p. 051102. DOI: [10.1103/PhysRevLett.126.051102](https://doi.org/10.1103/PhysRevLett.126.051102). arXiv: [2008.06071](https://arxiv.org/abs/2008.06071) [gr-qc].
- Clifton, Timothy et al. (2012). “Modified Gravity and Cosmology”. In: *Phys. Rept.* 513, pp. 1–189. DOI: [10.1016/j.physrep.2012.01.001](https://doi.org/10.1016/j.physrep.2012.01.001). arXiv: [1106.2476](https://arxiv.org/abs/1106.2476) [astro-ph.CO].
- Criswell, Alexander W. et al. (Apr. 2026). “A Foundation for Gravitational-Wave Population Inference within the LISA Global Fit”. In: arXiv: [2604.03390](https://arxiv.org/abs/2604.03390) [astro-ph.IM].
- De Santi, Federico et al. (2026). “Inferring the population properties of Galactic binaries from LISA’s stochastic foreground”. In: *Phys. Rev. D* 113.12, p. 123006. DOI: [10.1103/91c5-vgw5](https://doi.org/10.1103/91c5-vgw5). arXiv: [2602.18560](https://arxiv.org/abs/2602.18560) [astro-ph.HE].
- Destounis, Kyriakos and Kostas D. Kokkotas (2021). “Gravitational-wave glitches: Resonant islands and frequency jumps in nonintegrable extreme-mass-ratio inspirals”. In: *Phys. Rev. D* 104.6, p. 064023. DOI: [10.1103/PhysRevD.104.064023](https://doi.org/10.1103/PhysRevD.104.064023). arXiv: [2108.02782](https://arxiv.org/abs/2108.02782) [gr-qc].
- Destounis, Kyriakos et al. (2023). “Extreme-mass-ratio inspirals into rotating boson stars: Nonintegrability, chaos, and transient resonances”. In: *Phys. Rev. D* 108.8, p. 084062. DOI: [10.1103/PhysRevD.108.084062](https://doi.org/10.1103/PhysRevD.108.084062). arXiv: [2305.05691](https://arxiv.org/abs/2305.05691) [gr-qc].
- Detweiler, Steven L. (1980). “KLEIN-GORDON EQUATION AND ROTATING BLACK HOLES”. In: *Phys. Rev. D* 22, pp. 2323–2326. DOI: [10.1103/PhysRevD.22.2323](https://doi.org/10.1103/PhysRevD.22.2323).
- Dhani, Arnab, Jonathan Gair, and Alessandra Buonanno (July 2025). “The fault in our sirens: Hierarchical diagnosis of waveform systematics in Hubble-Lemaître constant measurements”. In: arXiv: [2507.11278](https://arxiv.org/abs/2507.11278) [gr-qc].
- Dicke, R. H. (1954). “Coherence in Spontaneous Radiation Processes”. In: *Phys. Rev.* 93, pp. 99–110. DOI: [10.1103/PhysRev.93.99](https://doi.org/10.1103/PhysRev.93.99).
- Dimopoulos, Savas et al. (2009). “Gravitational Wave Detection with Atom Interferometry”. In: *Phys. Lett. B* 678, pp. 37–40. DOI: [10.1016/j.physletb.2009.06.011](https://doi.org/10.1016/j.physletb.2009.06.011). arXiv: [0712.1250](https://arxiv.org/abs/0712.1250) [gr-qc].
- Dittmann, Alexander J., Abhishek Hegade K. R., and Callum W. Fairbairn (Aug. 2026). “Extreme-Mass-Ratio Inspirals in Gaseous Disks”. In: arXiv: [2608.17003](https://arxiv.org/abs/2608.17003) [astro-ph.HE].
- Dwyer, Sheila et al. (2015). “Gravitational wave detector with cosmological reach”. In: *Phys. Rev. D* 91.8, p. 082001. DOI: [10.1103/PhysRevD.91.082001](https://doi.org/10.1103/PhysRevD.91.082001). arXiv: [1410.0612](https://arxiv.org/abs/1410.0612) [astro-ph.IM].
- Dyson, F. W., A. S. Eddington, and C. Davidson (1920). “A Determination of the Deflection of Light by the Sun’s Gravitational Field, from Observations Made at the Total Eclipse of May 29, 1919”. In: *Phil. Trans. Roy. Soc. Lond. A* 220, pp. 291–333. DOI: [10.1098/rsta.1920.0009](https://doi.org/10.1098/rsta.1920.0009).
- Edwards, Matthew C. et al. (2020). “Identifying and Addressing Nonstationary LISA Noise”. In: *Phys. Rev. D* 102.8, p. 084062. DOI: [10.1103/PhysRevD.102.084062](https://doi.org/10.1103/PhysRevD.102.084062). arXiv: [2004.07515](https://arxiv.org/abs/2004.07515) [gr-qc].
- Einstein, Albert (1916). “Approximative Integration of the Field Equations of Gravitation”. In: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1916, pp. 688–696.
- (1918). “Über Gravitationswellen”. In: *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* 1918, pp. 154–167.
- Eisenhauer, Frank et al. (2005). “SINFONI in the Galactic Center: Young stars and IR flares in the central light month”. In: *Astrophys. J.* 628, pp. 246–259. DOI: [10.1086/430667](https://doi.org/10.1086/430667). arXiv: [astro-ph/0502129](https://arxiv.org/abs/astro-ph/0502129).

- Ghez, A. M. et al. (2005). "Stellar orbits around the galactic center black hole". In: *Astrophys. J.* 620, pp. 744–757. doi: [10.1086/427175](https://doi.org/10.1086/427175). arXiv: [astro-ph/0306130](https://arxiv.org/abs/astro-ph/0306130).
- Gillessen, S. et al. (2009). "Monitoring stellar orbits around the Massive Black Hole in the Galactic Center". In: *Astrophys. J.* 692, pp. 1075–1109. doi: [10.1088/0004-637X/692/2/1075](https://doi.org/10.1088/0004-637X/692/2/1075). arXiv: [0810.4674](https://arxiv.org/abs/0810.4674) [[astro-ph](#)].
- Glorio, Sara et al. (Mar. 2026). "Adiabatic evolution of asymmetric binaries on generic orbits with new fundamental fields I: characterization of gravitational wave fluxes". In: arXiv: [2603.10116](https://arxiv.org/abs/2603.10116) [[gr-qc](#)].
- Hailey, Charles J. et al. (2018). "A density cusp of quiescent X-ray binaries in the central parsec of the Galaxy". In: *Nature* 556.7699, pp. 70–73. doi: [10.1038/nature25029](https://doi.org/10.1038/nature25029).
- Hawking, Stephen (1971). "Gravitationally collapsed objects of very low mass". In: *Mon. Not. Roy. Astron. Soc.* 152, p. 75. doi: [10.1093/mnras/152.1.75](https://doi.org/10.1093/mnras/152.1.75).
- Heckman, Timothy M. and Philip Best (2014). "The Coevolution of Galaxies and Supermassive Black Holes: Insights from Surveys of the Contemporary Universe". In: *Ann. Rev. Astron. Astrophys.* 52, pp. 589–660. doi: [10.1146/annurev-astro-081913-035722](https://doi.org/10.1146/annurev-astro-081913-035722). arXiv: [1403.4620](https://arxiv.org/abs/1403.4620) [[astro-ph.GA](#)].
- Heijningen, Joris van et al. (2023). "The payload of the Lunar Gravitational-wave Antenna". In: *J. Appl. Phys.* 133.24, p. 244501. doi: [10.1063/5.0144687](https://doi.org/10.1063/5.0144687). arXiv: [2301.13685](https://arxiv.org/abs/2301.13685) [[gr-qc](#)].
- Hellings, R. w. and G. s. Downs (1983). "UPPER LIMITS ON THE ISOTROPIC GRAVITATIONAL RADIATION BACKGROUND FROM PULSAR TIMING ANALYSIS". In: *Astrophys. J. Lett.* 265, pp. L39–L42. doi: [10.1086/183954](https://doi.org/10.1086/183954).
- Hobbs, G. et al. (2010). "The international pulsar timing array project: using pulsars as a gravitational wave detector". In: *Class. Quant. Grav.* 27. Ed. by Zsuzsa Marka and Szabolcs Marka, p. 084013. doi: [10.1088/0264-9381/27/8/084013](https://doi.org/10.1088/0264-9381/27/8/084013). arXiv: [0911.5206](https://arxiv.org/abs/0911.5206) [[astro-ph.SR](#)].
- Hod, Shahar (2014). "Kerr-Newman black holes with stationary charged scalar clouds". In: *Phys. Rev. D* 90.2, p. 024051. doi: [10.1103/PhysRevD.90.024051](https://doi.org/10.1103/PhysRevD.90.024051). arXiv: [1406.1179](https://arxiv.org/abs/1406.1179) [[gr-qc](#)].
- Hu, Wen-Rui and Yue-Liang Wu (2017). "The Taiji Program in Space for gravitational wave physics and the nature of gravity". In: *Natl. Sci. Rev.* 4.5, pp. 685–686. doi: [10.1093/nsr/nwx116](https://doi.org/10.1093/nsr/nwx116).
- Jaccoud El-Jaick, Lea and Bartolomeu D. B. Figueiredo (July 2008). "On Certain Solutions for Confluent and Double-Confluent Heun Equations". In: *arXiv e-prints*, arXiv:0807.2219, arXiv:0807.2219. doi: [10.48550/arXiv.0807.2219](https://doi.org/10.48550/arXiv.0807.2219). arXiv: [0807.2219](https://arxiv.org/abs/0807.2219) [[math-ph](#)].
- Jani, Karan et al. (Aug. 2025). "Laser Interferometer Lunar Antenna (LILA): Advancing the U.S. Priorities in Gravitational-wave and Lunar Science". In: arXiv: [2508.11631](https://arxiv.org/abs/2508.11631) [[gr-qc](#)].
- Katz, Michael L. et al. (2021). "Fast extreme-mass-ratio-inspiral waveforms: New tools for millihertz gravitational-wave data analysis". In: *Phys. Rev. D* 104.6, p. 064047. doi: [10.1103/PhysRevD.104.064047](https://doi.org/10.1103/PhysRevD.104.064047). arXiv: [2104.04582](https://arxiv.org/abs/2104.04582) [[gr-qc](#)].
- Kejriwal, Shubham et al. (2026). *StableEMRIFisher*. Version 2.0.0. GitHub repository: <https://github.com/perturber/StableEMRIFisher>. doi: [10.5281/zenodo.18778938](https://doi.org/10.5281/zenodo.18778938).
- Kerr, Roy P. (1963). "Gravitational field of a spinning mass as an example of algebraically special metrics". In: *Phys. Rev. Lett.* 11, pp. 237–238. doi: [10.1103/PhysRevLett.11.237](https://doi.org/10.1103/PhysRevLett.11.237).

- Kocsis, Bence, Nicolas Yunes, and Abraham Loeb (2011). “Observable Signatures of EMRI Black Hole Binaries Embedded in Thin Accretion Disks”. In: *Phys. Rev. D* 84, p. 024032. DOI: [10.1103/PhysRevD.86.049907](#). arXiv: [1104.2322 \[astro-ph.GA\]](#).
- Kou, Xiao-Xiao et al. (Sept. 2026). “Ultralight Bosons Explain the Mass-Spin Correlations in the Merging Binary Black Hole Population”. In: arXiv: [2609.02678 \[gr-qc\]](#).
- Lovelock, D. (1971). “The Einstein tensor and its generalizations”. In: *J. Math. Phys.* 12, pp. 498–501. DOI: [10.1063/1.1665613](#).
- Luo, Jun et al. (2016). “TianQin: a space-borne gravitational wave detector”. In: *Class. Quant. Grav.* 33.3, p. 035010. DOI: [10.1088/0264-9381/33/3/035010](#). arXiv: [1512.02076 \[astro-ph.IM\]](#).
- Maggiore, Michele (2007). *Gravitational Waves. Vol. 1: Theory and Experiments*. Oxford University Press. ISBN: 978-0-19-171766-6, 978-0-19-852074-0. DOI: [10.1093/acprof:oso/9780198570745.001.0001](#).
- Maggiore, Michele et al. (2020). “Science Case for the Einstein Telescope”. In: *JCAP* 03, p. 050. DOI: [10.1088/1475-7516/2020/03/050](#). arXiv: [1912.02622 \[astro-ph.CO\]](#).
- Magorrian, John et al. (1998). “The Demography of massive dark objects in galaxy centers”. In: *Astron. J.* 115, p. 2285. DOI: [10.1086/300353](#). arXiv: [astro-ph/9708072](#).
- Mano, Shuhei, Hisao Suzuki, and Eiichi Takasugi (1996). “Analytic solutions of the Teukolsky equation and their low frequency expansions”. In: *Prog. Theor. Phys.* 95, pp. 1079–1096. DOI: [10.1143/PTP.95.1079](#). arXiv: [gr-qc/9603020](#).
- Marsat, Sylvain, John G. Baker, and Tito Dal Canton (2021). “Exploring the Bayesian parameter estimation of binary black holes with LISA”. In: *Phys. Rev. D* 103.8, p. 083011. DOI: [10.1103/PhysRevD.103.083011](#). arXiv: [2003.00357 \[gr-qc\]](#).
- Miralda-Escude, Jordi and Andrew Gould (2000). “A cluster of black holes at the galactic center”. In: *Astrophys. J.* 545, p. 847. DOI: [10.1086/317837](#). arXiv: [astro-ph/0003269](#).
- Misner, Charles W., K. S. Thorne, and J. A. Wheeler (1973). *Gravitation*. San Francisco: W. H. Freeman. ISBN: 978-0-7167-0344-0, 978-0-691-17779-3.
- Mitra, Soumodeep et al. (2025). “Extreme mass ratio inspirals in rotating dark matter spikes”. In: *Phys. Rev. D* 112.4, p. 044030. DOI: [10.1103/yc11-kx7d](#). arXiv: [2505.04697 \[gr-qc\]](#).
- Moore, C. J., R. H. Cole, and C. P. L. Berry (2015). “Gravitational-wave sensitivity curves”. In: *Class. Quant. Grav.* 32.1, p. 015014. DOI: [10.1088/0264-9381/32/1/015014](#). arXiv: [1408.0740 \[gr-qc\]](#).
- Murata, Jiro and Saki Tanaka (2015). “A review of short-range gravity experiments in the LHC era”. In: *Class. Quant. Grav.* 32.3, p. 033001. DOI: [10.1088/0264-9381/32/3/033001](#). arXiv: [1408.3588 \[hep-ex\]](#).
- Norcia, Matthew A., Julia R. K. Cline, and James K. Thompson (2017). “Role of atoms in atomic gravitational-wave detectors”. In: *Phys. Rev. A* 96.4, p. 042118. DOI: [10.1103/PhysRevA.96.042118](#). arXiv: [1707.04571 \[physics.atom-ph\]](#).
- Odintsov, S. D., V. K. Oikonomou, and F. P. Fronimos (2020). “Rectifying Einstein-Gauss-Bonnet Inflation in View of GW170817”. In: *Nucl. Phys. B* 958, p. 115135. DOI: [10.1016/j.nuclphysb.2020.115135](#). arXiv: [2003.13724 \[gr-qc\]](#).
- Oikonomou, V. K. (2021). “A refined Einstein–Gauss–Bonnet inflationary theoretical framework”. In: *Class. Quant. Grav.* 38.19, p. 195025. DOI: [10.1088/1361-6382/ac2168](#). arXiv: [2108.10460 \[gr-qc\]](#).

- Overduin, J. M. and P. S. Wesson (1997). “Kaluza-Klein gravity”. In: *Phys. Rept.* 283, pp. 303–380. DOI: [10.1016/S0370-1573\(96\)00046-4](#). arXiv: [gr-qc/9805018](#).
- Pan, Zhen et al. (2023). “Resonant dynamics of extreme mass-ratio inspirals in a perturbed Kerr spacetime”. In: *Phys. Rev. D* 108.10, p. 104026. DOI: [10.1103/PhysRevD.108.104026](#). arXiv: [2306.06576 \[gr-qc\]](#).
- Piro, L. et al. (2023). “Chasing Super-Massive Black Hole merging events with *Athena* and LISA”. In: *Mon. Not. Roy. Astron. Soc.* 521.2, pp. 2577–2592. DOI: [10.1093/mnras/stad659](#). arXiv: [2211.13759 \[astro-ph.HE\]](#).
- Punturo, M. et al. (2010). “The Einstein Telescope: A third-generation gravitational wave observatory”. In: *Class. Quant. Grav.* 27. Ed. by Fulvio Ricci, p. 194002. DOI: [10.1088/0264-9381/27/19/194002](#).
- Reitze, David et al. (2019). “Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO”. In: *Bull. Am. Astron. Soc.* 51.7, p. 035. arXiv: [1907.04833 \[astro-ph.IM\]](#).
- Saa, Alberto (1996). “New no scalar hair theorem for black holes”. In: *J. Math. Phys.* 37, pp. 2346–2351. DOI: [10.1063/1.531513](#). arXiv: [gr-qc/9601021](#).
- Sänger, Elise M. et al. (2026). “Tests of general relativity with GW230529: A neutron star merging with a lower mass-gap compact object”. In: *Phys. Rev. D* 113.8, p. 084070. DOI: [10.1103/r43k-51yq](#). arXiv: [2406.03568 \[gr-qc\]](#).
- Santos, João S. et al. (2025). “Gravitational Waves from Binary Extreme Mass Ratio Inspirals: Doppler Shift and Beaming, Resonant Excitation, Helicity Oscillations, and Self-Lensing”. In: *Phys. Rev. Lett.* 135.21, p. 211402. DOI: [10.1103/qq3m-6phh](#). arXiv: [2506.14868 \[gr-qc\]](#).
- Santos, João S. et al. (2026). “Resonances in binary extreme mass ratio inspirals”. In: *Phys. Rev. D* 113.6, p. 064025. DOI: [10.1103/7qjn-j14n](#). arXiv: [2601.02468 \[gr-qc\]](#).
- Schlamming, Stephan et al. (2008). “Test of the equivalence principle using a rotating torsion balance”. In: *Phys. Rev. Lett.* 100, p. 041101. DOI: [10.1103/PhysRevLett.100.041101](#). arXiv: [0712.0607 \[gr-qc\]](#).
- Sesana, A. et al. (2008). “Observing white dwarfs orbiting massive black holes in the gravitational wave and electro-magnetic window”. In: *Mon. Not. Roy. Astron. Soc.* 391, pp. 718–726. DOI: [10.1111/j.1365-2966.2008.13904.x](#). arXiv: [0806.0624 \[astro-ph\]](#).
- Sotiriou, Thomas P. and Shuang-Yong Zhou (2014a). “Black hole hair in generalized scalar-tensor gravity”. In: *Phys. Rev. Lett.* 112, p. 251102. DOI: [10.1103/PhysRevLett.112.251102](#). arXiv: [1312.3622 \[gr-qc\]](#).
- (2014b). “Black hole hair in generalized scalar-tensor gravity: An explicit example”. In: *Phys. Rev. D* 90, p. 124063. DOI: [10.1103/PhysRevD.90.124063](#). arXiv: [1408.1698 \[gr-qc\]](#).
- Speri, Lorenzo et al. (2024). “Fast and Fourier: Extreme Mass Ratio Inspiral Waveforms in the Frequency Domain”. In: *Front. Appl. Math. Stat.* 9. DOI: [10.3389/fams.2023.1266739](#). arXiv: [2307.12585 \[gr-qc\]](#).
- Srinivasan, Rahul et al. (2025). “Simulation-based population inference of LISA’s Galactic binaries: Bypassing the global fit”. In: *Phys. Rev. D* 112.10, p. 103043. DOI: [10.1103/shym-w46f](#). arXiv: [2506.22543 \[astro-ph.GA\]](#).
- Švančara, Patrik et al. (Sept. 2026). “End state of the experimental black-hole bomb”. In: arXiv: [2609.05595 \[gr-qc\]](#).

- Teukolsky, S. A. and W. H. Press (1974). “Perturbations of a rotating black hole. III - Interaction of the hole with gravitational and electromagnetic radiation”. In: *Astrophys. J.* 193, pp. 443–461. DOI: [10.1086/153180](https://doi.org/10.1086/153180).
- Teukolsky, Saul A. (1973). “Perturbations of a rotating black hole. 1. Fundamental equations for gravitational electromagnetic and neutrino field perturbations”. In: *Astrophys. J.* 185, pp. 635–647. DOI: [10.1086/152444](https://doi.org/10.1086/152444).
- (2015). “The Kerr Metric”. In: *Class. Quant. Grav.* 32.12, p. 124006. DOI: [10.1088/0264-9381/32/12/124006](https://doi.org/10.1088/0264-9381/32/12/124006). arXiv: [1410.2130](https://arxiv.org/abs/1410.2130) [gr-qc].
- Tomaselli, Giovanni Maria, Thomas F. M. Spieksma, and Gianfranco Bertone (2023). “Dynamical friction in gravitational atoms”. In: *JCAP* 07, p. 070. DOI: [10.1088/1475-7516/2023/07/070](https://doi.org/10.1088/1475-7516/2023/07/070). arXiv: [2305.15460](https://arxiv.org/abs/2305.15460) [gr-qc].
- Toubiana, Alexandre and Jonathan R. Gair (2026). “A framework for LISA population inference”. In: *JCAP* 07, p. 025. DOI: [10.1088/1475-7516/2026/07/025](https://doi.org/10.1088/1475-7516/2026/07/025). arXiv: [2601.04168](https://arxiv.org/abs/2601.04168) [gr-qc].
- Verbiest, J. P. W. et al. (2016). “The International Pulsar Timing Array: First Data Release”. In: *Mon. Not. Roy. Astron. Soc.* 458.2, pp. 1267–1288. DOI: [10.1093/mnras/stw347](https://doi.org/10.1093/mnras/stw347). arXiv: [1602.03640](https://arxiv.org/abs/1602.03640) [astro-ph.IM].
- Vicente, Rodrigo and Vitor Cardoso (2022). “Dynamical friction of black holes in ultralight dark matter”. In: *Phys. Rev. D* 105.8, p. 083008. DOI: [10.1103/PhysRevD.105.083008](https://doi.org/10.1103/PhysRevD.105.083008). arXiv: [2201.08854](https://arxiv.org/abs/2201.08854) [gr-qc].
- Virtanen, Pauli et al. (2020). “SciPy 1.0–Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Meth.* 17, p. 261. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2). arXiv: [1907.10121](https://arxiv.org/abs/1907.10121) [cs.MS].
- Wagoner, Robert V. (1970). “Scalar tensor theory and gravitational waves”. In: *Phys. Rev. D* 1, pp. 3209–3216. DOI: [10.1103/PhysRevD.1.3209](https://doi.org/10.1103/PhysRevD.1.3209).
- Weber, J. (1968). “Gravitational-Wave-Detector Events”. In: *Phys. Rev. Lett.* 20.23, p. 1307. DOI: [10.1103/PhysRevLett.20.1307](https://doi.org/10.1103/PhysRevLett.20.1307).
- Yunes, Nicolas et al. (2011). “Imprint of Accretion Disk-Induced Migration on Gravitational Waves from Extreme Mass Ratio Inspirals”. In: *Phys. Rev. Lett.* 107, p. 171103. DOI: [10.1103/PhysRevLett.107.171103](https://doi.org/10.1103/PhysRevLett.107.171103). arXiv: [1103.4609](https://arxiv.org/abs/1103.4609) [astro-ph.CO].



Acknowledgments

I would like to sincerely thank my supervisors, Prof. Andrea Maselli and Prof. Jean-Pierre Zendri, for their constant support, guidance, and encouragement throughout the writing of this Thesis. I am particularly grateful for their availability and for the opportunity to learn from their expertise.

I gratefully acknowledge the financial support provided by the internship programme “Probing the Existence of New Fundamental Fields with Gravitational-Wave Observations” at the Gran Sasso Science Institute (GSSI).

I acknowledge the use of the Penn State GWave cluster (part of the LIGO Data Grid/IGWN Computing Grid) for computational resources. This work used the Open Science Grid, which is supported by the National Science Foundation and other federal partners.