2 Science case

2.1 Einstein Telescope Science Goals

Some three hundred years after Galileo observed the Jovian satellites, the twentieth century heralded a new era in observational astronomy with radio and microwave antennas, gamma- and X-ray detectors, which revolutionized astronomy and opened the entire electromagnetic window for observing and understanding the Universe. A remarkable revelation coming from these observations is that about 96 percent of our Universe is invisible and that gravitational interaction powers the most luminous and spectacular objects and phenomena such as quasars, gamma-ray bursts, ultra luminous X-ray sources, pulsars, and the evolution of the early Universe.

Einstein's theory of gravity predicted that dynamical systems in strong gravitational fields will release vast amounts of energy in the form of gravitational radiation. This radiation has the potential to open a new window on the Universe, complementing the electromagnetic window. Russell Hulse and Joe Taylor were awarded the 1993 Nobel Prize in Physics for the discovery of a binary consisting of two neutron stars in close orbit in which indirect evidence for the emission of gravitational waves was found.

Interferometric gravitational wave (GW) detectors that are currently taking data and advanced detectors that will be built over the next ten years will be the first steps in establishing the field of gravitational astronomy through their detection of the most luminous sources such as the merger of binary neutron stars and black holes. Einstein Telescope will make it possible to observe a greater variety of phenomena, and provide a new tool for expanding our knowledge of fundamental physics, cosmology and relativistic astrophysics. Is the nature of gravitational radiation as predicted by Einstein's theory? Are black hole spacetimes uniquely given by the Kerr geometry? Do event horizons always form around gravitationally collapsing matter? How did the black holes at galactic nuclei form? What were the physical conditions in the very early Universe? What is the nature of quantum gravity and what is the origin of space and time? Are there really ten spatial dimensions? These are some key questions at the forefront of physics that future GW observations might shed some light on.

2.1.1 Fundamental physics

Astronomical sources of gravitational waves are essentially systems where gravity is extremely strong and often characterized by relativistic bulk motion of massive objects. The emitted radiation carries an uncorrupted signature of the nature of the space-time geometry and is therefore an invaluable tool to understand the behaviour of matter and geometry in extreme conditions of density, temperature, magnetic fields and relativistic motion. Here are some examples of how GW observations can impact fundamental physics.

In Einstein's theory, gravitational radiation travels at the speed of light and has two polarization states. In alternative theories of gravity one or both of these properties might not hold, owing to the presence of massive gravitons, or a scalar field mediating gravity in addition to the tensor field. Experimental tests of gravity, as well those afforded by the data from the Hulse-Taylor binary, are consistent with both Einstein's theory and one of its alternatives called the Brans-Dicke theory. Gravitational wave detectors will bring these theories vis-a-vis observations that could decisively rule out one or the other.

According to Einstein's gravity the space-time in the vicinity of black holes is described by a unique geometry called the Kerr solution. Observation of the radiation from the in-fall of stellar-mass black holes into intermediate-mass black holes will make it possible to test such uniqueness theorems. X-ray astronomy has provided firm indirect evidence that intense sources of x-rays may well host a black hole. An unambiguous signature of the black hole geometry, however, could eventually be provided by the detection of black hole quasi-normal modes: gravitational radiation with characteristic frequencies and decay times. Failure to detect such radiation from, for example, a newly formed black hole would mean that gravity is more exotic than what we currently believe, and might reveal new phases of matter at extremely high densities.

The most attractive versions of string theory require a ten- or eleven-dimensional space-time, far larger than what we perceive. In certain phenomenological models at the interface of string theory and cosmology, what

we perceive as a four-dimensional Universe could be one part, or "brane", within a higher dimensional "bulk" Universe. The extra spatial dimensions may be compact and sub-millimetre-scale, or even macroscopically large, if their geometry has properties known as "warping". The key feature of brane-world theories is that gravitational interactions, and in particular gravitational waves, propagate in the bulk, while other interactions are restricted to the brane, which partly explains why gravity is so weak.

2.1.2 Relativistic Astrophysics

Astronomy has revealed a Universe full of diverse and exotic phenomena which remain enigmas decades after their discovery. Supernovae are the end-states of stellar evolution, resulting in gravitational collapse followed by a huge explosion of in-falling matter. Gamma-ray bursts are intense sources of gamma radiation that last only a few seconds to minutes yet emit more energy than a star does in its entire lifetime. Radio pulsars are compact objects as massive as the Sun but only about 10 km in size, and the regularity of their radio pulses rivals the best atomic clocks in the world. Transient radio sources thousands of light years away are associated with magnetic fields so strong that the emitted radiation could breakdown terrestrial radio stations. For each of these objects the source is believed to be couched in dense environs and strong gravitational fields and, therefore, is a potential source of gravitational radiation. For example, gamma-ray bursts could be produced by colliding neutron stars which are electromagnetically invisible for most of their lives but are very powerful emitters of GW. Transient radio sources could be the result of quakes in neutron stars with concomitant emission of GW. Observing such 'multi-messengers' (sources that are strong emitters of both EM and GW radiation) will help understand phenomena that have remained puzzles for decades.

The centre of every galaxy is believed to host a compact object a million to a billion times as massive as our Sun, a powerful emitter of optical, radio and other radiation. What is the nature of this object? How and when it form? Did it form from small 100 solar mass seeds and then grow by accreting gas and other compact objects? What is its relation to the size of the galaxy as a whole? These are some of the questions which a model of the formation of structure in the Universe must answer. While electromagnetic observations have provided valuable data, GW observations can help address some of the key questions on the formation and nature of these objects.

Future gravitational wave detectors will also be sensitive to a population of sources at very high red-shifts, helping us study cosmological evolution of sources, the history of star formation and its dependence on the matter content of the Universe, and development of large-scale structure in the Universe.

2.1.3 Cosmology

The twentieth century was the golden age of cosmology. With the advent of radio and microwave astronomy it was possible to finally address key questions about the origin of the Universe. The cosmic microwave background (CMB) is a relic radiation from the hot "Big Bang" that is believed to have been the initial condition for primordial nucleosynthesis. Since the early Universe was very dense, this radiation was in thermal equilibrium with matter for about 380,000 years after the Big Bang and cannot directly reveal the conditions in the very early phases of the Universe's history. The most direct way of observing the primaeval Universe is via the gravitational window with a network of two or more detectors. From fairly general assumptions one can predict the production of a stochastic background of GW in the early Universe, which travel to us unscathed as a consequence of their weak coupling to matter.

The most amazing aspect of the Universe is that only about 4% of its energy density appears in the form of visible matter, the rest being dark matter and dark energy. In order to understand the behaviour of these 'dark' contents it is necessary to have a standard candle – a population of sources whose distance from Earth can be inferred from their luminosity. Compact binaries are an astronomer's ideal standard candle: By measuring the signature of the gravitational radiation they emit, it is possible to infer their intrinsic parameters (e.g. the masses and spins of the component objects) and accurately deduce their luminosity distance. In fact, compact binaries eliminate the need to build a cosmic distance ladder – the process by which standard candles at different distances are calibrated in astronomy since there is no source that is suitable at all distances.

The synergy of multi-messenger astronomy is nowhere more apparent than in the use of standard sirens of gravity to test the concordance model of cosmology. ET might detect several hundred compact binary coalescence events each year in coincidence with short-hard gamma-ray bursts, provided, of course, the two are related. While gravitational observations would provide an unambiguous measure of the luminosity distance, the host galaxy of the GRB could be used to measure the redshift. By fitting the observed population to a cosmological model, it will be possible to measure the Hubble parameter, dark matter and dark energy densities, as well as the dark energy equation-of-state parameter.

The early history of the Universe may have witnessed several phase transitions as the temperature decreased through the energy scales of a Grand Unified Theory (GUT) and electroweak symmetry-breaking, and eventually to the current state in which we see four different fundamental interactions. During some phase transitions, cosmic strings form as one-dimensional topological defects at the boundaries of different phases. Vibrations of these strings at the speed of light can sometimes form a kink which emits a burst of gravitational radiation. The spectrum of such radiation has a unique signature which can help us detect cosmic strings and measure their properties, and thus provide a glimpse of the Universe as it underwent phase transitions.

Perhaps the most exciting discovery of the new window will be none of the above. If the astronomical legacy is any example, gravitational astronomy should unveil phenomena and sources never imagined in the wildest theories – a possibility of any new observational tool.

2.2 Sources of gravitational waves in ET

The goal of this Section is to give an overview of the sources expected to be observed by ET and problems to be addressed in the context of the Design Study. A brief introduction to gravitational waves is given in Box 2.1 and then go on to describe the sources, their properties and the problems that need to be addressed over the next two years.

2.1: Gravitational Waves

Gravitational waves are described by a second rank tensor $h_{\alpha\beta}$, which, in a suitable coordinate system and gauge, has only two independent components h_+ and $h_\times$, $h_{xx} = -h_{yy} = h_+$, $h_{xy} = h_{yx} = h_\times$, all other components being zero. A detector measures only a certain linear combination of the two components, called the *response* $h(t)$, given by

$$h(t) = F_+(\theta, \varphi, \psi)h_+(t) + F_\times(\theta, \varphi, \psi)h_\times(t), \quad (1)$$

where F_+ and $F_\times$ are the detector antenna pattern functions, ψ is the polarization angle, and (θ, φ) are angles describing the location of the source on the sky. The angles are all assumed to be constant for a transient source but time-dependent for sources that last long enough so that the Doppler modulation of the signal due to the relative motion of the source and detector cannot be neglected.

2.2.1 Compact binary coalescences

A compact binary, consisting of neutron stars (NS) and/or black holes (BH), evolves by emitting gravitational radiation which extracts the rotational energy and angular momentum from the system, thereby leading to an inspiral of the two bodies towards each other. The dynamics of a compact binary consists of three phases: (i) The *early inspiral phase* in which the system spends 100's of millions of years and the luminosity in GW is rather low and the dynamics can be solved using approximation methods - the most popular being the post-Newtonian (PN) approximation (see Box 2.2). The inspiral signal has a characteristic shape, with slowly increasing amplitude and frequency and is called a *chirp* waveform. A binary signal that chirps (i.e. its frequency

changes perceptibly during the course of observation) is an astronomer's *standard candle* [5] (see below) and by observing the radiation from a *chirping* binary we can measure the luminosity distance to the source. (ii) The *plunge* phase when the two stars are moving at a third of the speed of light and experiencing strong gravitational fields with the gravitational potential being $\varphi = GM/Rc^2 \sim 0.1$. This phase requires the full non-linear structure of Einstein's equations as the problem involves strong relativistic gravity, tidal deformation (in the case of BH-BH or BH-NS) and disruption (in the case of BH-NS and NS-NS) and has only recently been solved by numerical relativists (see below). Analytical solutions based on resummation of the PN series have been very successful in describing the merger phase. (iii) The *merger,* or *ringdown,* phase when the two systems have merged to form either a NS or BH, settling down to a quiescent state by radiating the deformations inherited during the merger. The emitted radiation can be computed using perturbation theory and gives the quasi-normal modes (QNM) of BH and NS. The QNM carry a unique signature that depends only on the mass and spin angular momentum in the case of BH, but depends also on the equation-of-state (EOS) of the material in the case of NS.

2.2: Post-Newtonian Description of the Inspiral Signal

The adiabatic evolution of a compact binary, during which the emission of gravitational waves causes the component stars of the system to *slowly* spiral-in towards each other, can be computed very accurately using the post-Newtonian (PN) expansion of the Einstein equations. Currently, the dissipative dynamics is known to order $O(v^7/c^7)$, where v is the characteristic velocity in the system.

For a binary consisting of two stars of masses m_1 and m_2 (total mass $M \equiv m_1 + m_2$ and symmetric mass ratio $\nu \equiv m_1 m_2/M^2$) and located at a distance $D_{\rm L}$, the dominant gravitational wave amplitudes are

$$h_+(t) = \frac{2\nu M}{D_{\rm L}}(1+\cos^2\iota)\,[M\omega(t;t_0,M,\nu)]^{\frac{2}{3}}\cos\left[2\Phi(t;t_0,M,\nu)+\Phi_0\right], \quad (2)$$

$$h_\times(t) = \frac{2\nu M}{D_{\rm L}}2\cos\iota\,[M\omega(t;t_0,M,\nu)]^{\frac{2}{3}}\sin\left[2\Phi(t;t_0,M,\nu)+\Phi_0\right], \quad (3)$$

where ι is the angle of inclination of the binary's orbital angular momentum with the line-of-sight, $\omega(t)$ is the angular velocity of the equivalent one-body system around the binary's centre-of-mass and $\Phi(t;\,t_0,M,\nu)$ is the corresponding orbital phase. Parameters t_0 and Φ_0 are constants giving the epoch of merger and the orbital phase of the binary at that epoch, respectively.

The above expressions for h_+ and $h_\times$ are the dominant terms in what is essentially a PN perturbative series, an approximation technique that is used in solving the Einstein equations as applied to a compact binary. This dominant amplitude consists of only twice the orbital frequency. Higher order amplitude corrections contain other harmonics (i.e., phase terms consisting of $k\,\Phi(t)$, $k = 1, 3, 4, \ldots$). Also, the above expressions are written down for a system consisting of non-spinning components on a quasi-circular orbit. In reality, we can assume neither to be true. Waveforms for binaries on an eccentric inspiral orbit are known as are those with spin effects but we shall not discuss them here.

The adiabatic inspiral, during which the signal is approximated by (3), is followed by the merger of the compact objects, leading to the formation of a single black hole. This black hole then undergoes a rapid 'ringdown' as it settles down to a quiescent state. Following breakthroughs in 2005 [6–8], it is now possible to numerically solve the full Einstein equations for the last orbits, merger and ringdown of comparable mass black-hole-binary systems, and to calculate the emitted GW signal. Subsequent dramatic progress has lead both to simulations of rapidly increasing numerical accuracy and physical fidelity, and to the inclusion of larger numbers of GW cycles before merger, allowing full GR waveforms to be in principle useful for searches of black-hole binaries of ever lower mass; see Fig. 3 in [9]. The inclusion of merger and ringdown dramatically increases the signal-to-noise ratio, leading to a much larger distance reach than one would have with the inspiral signal alone. As we shall

see below, having observational access to these later stages of the coalescence process will lead to key insights into the structure of neutron stars; in the case of black holes it will open up the genuinely strong-field dynamics of spacetime.

Standard Sirens of Gravity Cosmologists have long sought for standard candles that can work on large distance scales without being dependent on the lower rungs of cosmic distance ladder. In 1986, Schutz [5] pointed out that gravitational astronomy can provide such a candle, or, more appropriately, a *standard siren*, in the form of a chirping signal from the coalescence of compact stars in a binary. The basic reason for this is that the gravitational-wave amplitude depends only on the ratio of a certain combination of the binary masses and the luminosity distance. For chirping signals observations can measure both the amplitude of the signal and the masses very accurately and hence infer the luminosity distance.

The detector response depends only on a small number of signal parameters, which can all be measured either directly or indirectly. The signal is insensitive to the composition of the component stars and there is no complicated modelling that involves the structure of the stars or their environments. Consequently, the measurement of the luminosity distance is precise, except for statistical errors, whose magnitude depends on the signal-to-noise ratio (SNR), and systematic errors due to weak gravitational lensing. We will discuss the relative magnitude of these errors later on.

Substituting the expressions given in Eq. (3) for h_+ and $h_\times$ in Eq. (1), we get

$$h(t) = \frac{2\nu M}{D_{\rm eff}} [M\omega(t)]^{\frac{2}{3}} \cos[2\Phi(t) + \Phi_0']. \tag{4}$$

Here $D_{\rm eff}$ is the effective distance to the binary, which is a combination of the true luminosity distance and the antenna pattern functions, and Φ_0' is a constant phase involving the various angles,

$$D_{\rm eff} \equiv \frac{D_{\rm L}}{\left[F_+^2(1+\cos^2\iota)^2 + 4F_\times^2\cos^2\iota\right]^{1/2}}, \quad \Phi_0' \equiv \Phi_0 + \arctan\left[-\frac{2F_\times\cos\iota}{F_+(1+\cos^2\iota)}\right]. \tag{5}$$

Note that $D_{\rm eff} \geq D_{\rm L}$. In the case of non-spinning binaries on a quasi-circular orbit, therefore, the signal is characterized by nine parameters in all, $(M, \nu, t_0, \Phi_0, \theta, \varphi, \psi, \iota, D_{\rm L})$.

Since the phase $\Phi(t)$ of the signal is known to a high order in PN theory, one employs matched filtering to extract the signal and in the process measures the two mass parameters (M, ν) (parameters that completely determine the phase evolution) and the two fiducial parameters (t_0, Φ_0). In general, the response of a single interferometer will not be sufficient to disentangle the luminosity distance from the angular parameters. However, EM identification (i.e., electromagnetic, especially optical, identification) of the source will determine the direction to the source, still leaving three unknown parameters $(\psi, \iota, D_{\rm L})$. If the signal is a transient, as would be the case in ground-based detectors, a network of three interferometers will be required to measure all the unknown parameters and extract the luminosity distance.

Although the inspiral signal from a compact binary is a standard siren, there is no way of inferring from it the red-shift to a source. The mappings $M \to (1+z)M$, $\omega \to \omega/(1+z)$, and $D_{\rm L} \to (1+z)D_{\rm L}$, in Eq. (3), leave the signal invariant. Note that a source with an intrinsic (i.e., physical) total mass $M_{\rm phys.}$ at a red-shift z will appear to an observer to be a binary of total mass $M_{\rm obs.} = (1+z)M_{\rm phys.}$. One must optically identify the host galaxy to measure its red-shift. Thus, there is synergy in GW and EM observations which can make precision cosmography possible, without the need to build a cosmic distance ladder. Later in this document we will see how to exploit compact binaries for fundamental physics and cosmography.

Cosmological evolution of compact object populations The calculation of the coalescence rate as a function of the redshift must take into account the following factors: the star formation rate history $SFR(z)$, the binary fraction $f_b(z)$, the formation efficiency of a given type of binary, i.e. the fraction of number of binaries that lead to formation of coalescing compact object binary, and their distribution of merger times.

These quantities may depend on redshift since the stellar populations evolve with cosmic time. Let us examine the effects of evolution of each of these factors.

The star formation rate is known to increase strongly to the redshift $z = 2$, and there is a debate about its behavior for higher redshifts. At redshift $z = 2$, the star formation is estimated to be a factor of 10 larger than the present value at $z = 0$.

The distribution of merger times can be estimated either by analyzing the present population of compact objects binaries or by involving the population synthesis. The first approach is limited to deal with the double neutron star binaries, and suffers from small number statistics. The second involves several uncertainties due to parametrization of binary evolution. However the two approaches yield similar results. The distribution of merger times for the double neutron star binaries can be well approximated by a distribution $\propto t^{-1}$. The lower cutoff for the DNS systems lies somewhere between 10 and 100 Myrs. The population synthesis leads to similar conclusions about the distribution of merger times for BHNS and BBH systems, however the low time cutoff may probably lie higher.

The evolution of the properties of binaries with cosmic time. The main factor that may affect the evolution of the binaries as a function of redshift are the changes in the distribution of metallicity. Metallicity affects strongly the mass loss rate in stars, and hence has a strong influence on the masses spectrum of compact objects. The lower the metallicity the higher the maximum mass of a black hole that may be formed in the course of stellar evolution. This leads to to stabilization of mass transfers and therefore to increase in the formation rate of compact object binaries.

Taking together the above factors we see that there are several reasons why the coalescence rate should increase strongly as we go to redshifts of $z = 1$–2. First the local star formation rate increases and the overall number of binary formation is larger. Second, the typical delay times for the DNS systems are low therefore their merger rate density will roughly follow that of the SFR. In the case of BHNS or BBH systems he typical delay times between formation and coalescence may be as large as 1–3 Gyrs. This delays the peak of coalescence rate density with respect to the star formation rate. Thus the delays are significant but not crucial. Third, the metallicity evolution may lead to higher compact object formation rate for high redshifts, and formation of larger number of massive BBH binaries.

This consideration can be put into detailed numerical codes to yield predictions about the rates. However even without such strong numerical support one can readily estimate with the back of the envelope calculation that the ratio of the coalescence rate (per unit volume per unti time) to the local one should be at least a few. The local coalescence rate can only be estimated with observations since neither the observational not the indirect approach mentioned earlier can yield the estimate of the rate with the accuracy better that plus minus an order of magnitude.

The Einstein telescope will provide a large sample of coalescences with the precise measurement of their masses and redshifts. This will be an extremely valuable tool for analysis of the cosmic compact object formation history. The measurement of their masses will yield information on the metallicity evolution as well as evolution of most massive stars. The Einstein telescope will yield a cosmic compact object census up to redshift $z = 2$, and will yield information about black holes and neutron star formed even at earlier epochs because of the delays between formation and coalescence.

There are two distinct routes to form BH binary. The first, conventional way, is to start with binary system of two main sequence stars and trace their evolution. There are several big uncertainties in this process. The first one is the initial mass ratio function: what is the distribution of the mass ratio in the binary of two main sequence stars, how it depends on the metallicity and spectral type. The second, and probably the biggest uncertainty, is related to the "common envelope" evolution, where the NS (or BH) and Helium core are emerged and evolve in the gaseous environment of the star. In this stage the NS/BH could merge with Helium core and binary is not formed. The third uncertainty is related to the direction and magnitude of the kick exerted on the newly born BH from the assymetric supernova explosion. All the above is reflected in the uncertainties on the rate of such binaries [10, 11].

The BH binaries could also be formed in the dense environment such as galactic nuclei. In the galaxies with

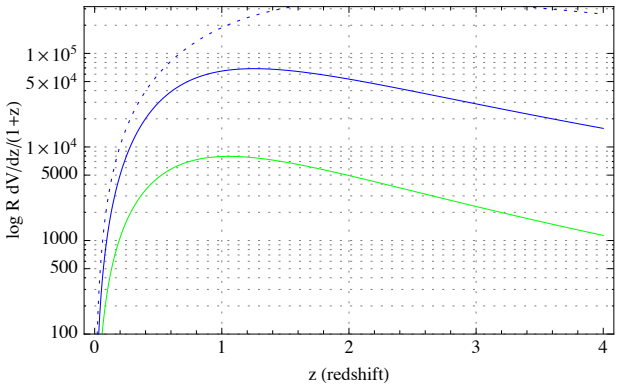


Figure 5: The merger rate dR/dz on our past light cone versus redshift. A thin dotted blue line extrapolates the Milky Way double neutron star merger rate to the universe using Eq. 6, which assumes that merger rates trace the star formation history [12] and that the Milky Way forms stars at a rate $dM/dt_{MW} \simeq \dot{\rho}_{MW}/n_{mw} = 3M_\odot$ yr^{-1}. More detailed calculations that account for the finite delay between binary birth and merger are shown in the two sold lines for NS-NS (blue) and BH-NS (green) binaries. These delays insure ET will probe the redshift region where most binaries merge.

SMBH ($M < 10^7\,M_\odot$), the relaxation time is less than a Hubble time, and a steep cusp of stars and stellar mass BHs can be formed. BHs as more massive and compact objects will segregate into central ≈ 1 pc region. Other two dense regions are massive globular clusters and nuclear star clusters in the centers of low-mass galaxies which may not have SMBH. The densities in those regions are high enough to have multiple encounters with formation and/or hardening of the BH binaries.

Expected coalescence rates Black holes or neutron stars are expected to form after Type II supernovae, which occur roughly once a century in galaxies like our own. Most stars seem to form in binaries binaries; a fraction of compact binary progenitors will survive the kicks that supernovae impart; and roughly half of the remaining low-mass binaries (BH-NS; NS-NS) will inspiral and eventually merge through the gradual emission of radiation. With roughly $n_{mw} = 0.01$ Milky Way-like galaxy per Mpc^3, we anticipate a rate per comoving volume ρ_c large enough to permit many detections even for advanced-LIGO scale detectors (Table 2). For example, the binary pulsar population in the Milky Way implies a local NS-NS merger rate $\rho_c^{(NS-NS)} \simeq 0.2 - 6$ Myr^{-1} Mpc^{-3} [13–15].

With its vastly greater sensitivity, the Einstein Telescope will reach deep back into the universe. Due to an enhanced the star formation rate between $z \simeq 1 - 3$ [12], ET will probe a regime of possibly significantly enhanced compact object merger rates [16–18]. By way of illustration, because double neutron stars have a relatively short delay time, their formation rate roughly traces the star formation rate of the universe [16]. For example, assuming all gas forms stars similar to the present day Milky Way, the current Milky Way compact object merger rate ($R_{MW} = \rho_c n_{MW}$) and star formation rate per volume ($\dot{\rho}_{MW}$) allow us rescale the *total* star formation rate of the universe ($\dot{\rho}_{SFR}$) into an instantaneous merger rate per unit volume ($R_{mw} n_{MW} \dot{\rho}_{SFR}/\dot{\rho}_{MW}$). The total merger rate follows by adding up all contributions on the past light cone out to ET's sensitivity limit, via an estimate for the star formation rate $\dot{\rho}_{SFR}$ [12, 17] (cf. Figure 5, which adopts a more realistic model):

$$\frac{dR}{dz} = \frac{dV_c}{dz}\frac{\mathcal{R}(t)}{1+z} \simeq \frac{dV_c}{dz}\rho_c^{(NS-NS)}\frac{\dot{\rho}_{SFR}/\dot{\rho}_{MW}}{1+z}. \quad (6)$$

Depending on the target sensitivity and beampattern of the ET network, the expected detection rate is roughly proportional to the integral of this rate up to some peak redshift. For most target ET sensitivities this limiting redshift z_{max} is greater than 1 for double neutron stars, suggesting $O(10^6)$ detections per year. The enormous collections of events that ET-scale instruments will provide permit high-precision modeling inaccessible with the sparse statistics available to smaller detectors.

Table 2: This Table gives the expected coalescence rates per Mpc^3 Myr in the local universe ($z \simeq 0$). Also shown are predicted event rates in Advanced LIGO (AL) and Einstein Telescope (ET).

Rate/Events	BNS	NS-BH	BBH
Rate (Mpc^{-1} Myr^{-1})	0.1-6	0.01-0.3	$2 \times 10^{-3} - 0.04$
Event Rate (yr^{-1} in AL)	0.4-400	0.2-300	2-4000
Event Rate (yr^{-1} in ET)	$O(10^3$-$10^7)$	$O(10^3)$-10^7	$O(10^4$-$10^8)$

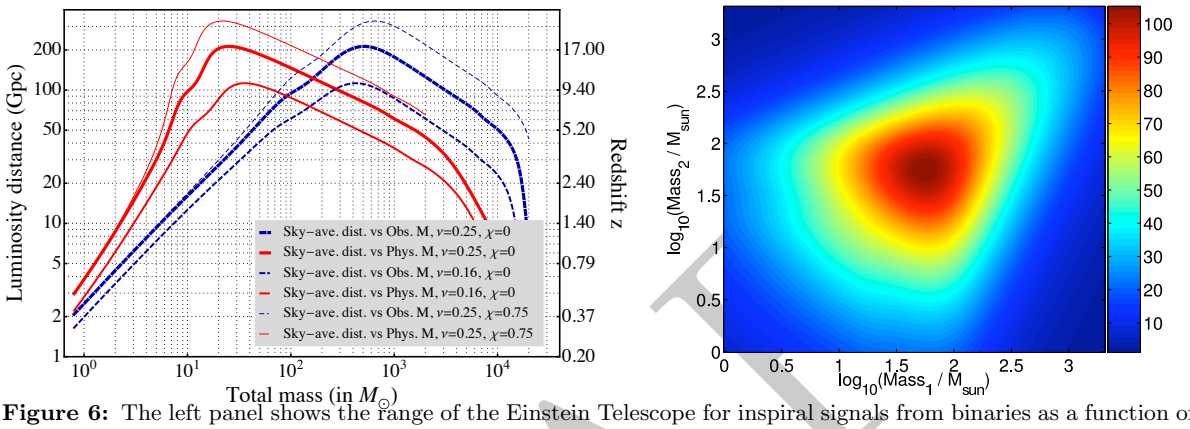


Figure 6: The left panel shows the range of the Einstein Telescope for inspiral signals from binaries as a function of the *intrinsic* (red solid line) and *observed* (blue dashed line) total mass. We assume that a source is visible if it produces an SNR of at least 8 in ET. The right panel shows in the plane of component masses the SNR for binaries at a distance of 3 Gpc.

Lacking direct observational input, predictions for BH-BH and BH-NS binaries rely entirely on theory. Studies of isolated binary evolution in the Milky Way [19–22] and local universe [17] produce event rates roughly in the ranges shown in Table 2, depending on the assumptions adopted in the model. As with the NS-NS rate, the BH-NS merger rate is roughly proportional to the star formation rate [18] and therefore also increases substantially with redshift [Figure 5]; many detections are expected.

The BH-BH merger rate is much less certain. First, long expected delays between BH-BH birth and merger imply black holes born in the early universe could merge now [17]. Second, BH masses depend strongly on the metallicity of the gas from which the progenitor star forms, as stellar winds operate much less efficiently [23]; so does the BH-BH binary formation rate [24]. In other words, low metallicity environments form both more binaries and binaries that can be detected farther away. Even restricting attention to the local universe, low-metallicity environments should be significantly over-represented in the present-day detection rate [25]. For example, the nearby BH-BH progenitor binary IC 10 X-1 both lies in a low metallicity environment and suggests a high BH-BH detection rate for initial LIGO ($O(0.5)$yr, strongly dependent on survey selection effects; see [26]). Further, in the early universe, where fewer generations of stars have produced metals, very massive binaries could form very frequently [24]. Third, being the most massive compact objects, black holes can *mass segregate* in interacting protoclusters. If enough protoclusters persist long enough for this process to occur, the BH-BH binary merger rate could be vastly enhanced [27–29]. As a practical matter, theory provides no useful upper bound; for example, the local BH-BH rate per mass bin is constrained only by existing gravitational wave measurements.

Expected distance reach and mass range A standard measure of the reach of a detector is the horizon distance D_h, defined as the distance at which a detector measures an SNR = 8 for an optimally-oriented and optimally-located binary, i.e. an overhead, face-on orbit. Suboptimally located and oriented sources are detected with SNR = 8 at closer distances. The sky-position averaged distance up to which a 3-detector ET network might detect inspiral signals from coalescing binaries with an SNR of 8 is shown in left panel of Fig. 6. We

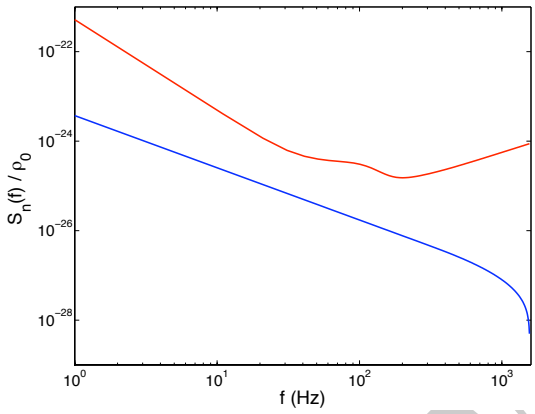


Figure 7: Spectral energy density of the background produced by the coalescence of double neutron stars, compared to the planned sensitivity of Einstein Telescope.

plot the range both as a function of the intrinsic (red solid lines) and observed (blue dashed lines) total mass. Those are related by the redshift function $z(d)$, which we compute according to the standard ΛCDM universe with parameters given by the first five years of the WMAP sky survey [30]. The binary systems are modeled by the phenomenological waveforms of [31] which comprise the inspiral, merger and ringdown stages of the coalescence. In left panel of Fig. 6 we show the reach associated to three physical configurations of the binary: equal-mass, non-spinning; non-spinning with mass ratio 1 : 4; and equal-mass, spin-aligned configuration with spins $\chi_1 = \chi_2 = \chi = 0.75$. Right panel displays the SNR for binaries located at 3 Gpc versus the component masses of the system. Our results show that a binary comprising two $1.4\,M_\odot$-neutron stars (BNS) can be observed by ET from a red-shift of $z \simeq 2$, and that comprising a $1.4\,M_\odot$-neutron star and a $10\,M_\odot$-black hole (NS-BH) from $z \simeq 4$. Binaries formed by stellar-mass black holes will be visible at much larger distances, allowing the ET to explore cosmological distances of $z \simeq 10$ and further.

Contribution of intermediate-mass black holes Globular clusters may host intermediate-mass black holes (IMBHs) with masses in the $\sim 100 - 1000\ M_\odot$ range (see [32, 33] for reviews on IMBHs, and [34] for an announcement of a recently discovered ultra-luminous X-ray source that represents a possible IMBH detection). These may contribute to binary merger rates observable by the ET in two ways.

Since an IMBH will be the most massive object in the cluster, it will readily sink to the center and substitute into a binary with a compact-object companion. The binary will then harden through three-body interactions and eventually merge via an intermediate-mass-ratio inspiral (IMRI) on timescales of less than one billion years [35]. The number of detectable mergers depends on the unknown distribution of IMBH masses and their typical companions. According to [36], 300 events could be detected to $z = 1.5$ for 100-solar-mass (redshifted) primaries and 10-solar-mass (redshifted) secondaries, but the range and rates drop for higher-mass primaries and lower-mass secondaries.

If the stellar binary fraction in a globular cluster is sufficiently high, two or more IMBHs can form [37]. These IMBHs then sink to the center in a few million years, where they form a binary and merge via three-body interactions with cluster stars followed by gravitational radiation reaction (see [37, 38] for more details). Then the ET could detect $2000\left(\frac{g}{0.1}\right)\left(\frac{g_{\rm cl}}{0.1}\right)$ mergers per year, where g is the fraction of all globular clusters hosting pairs of IMBHs, and $g_{\rm cl}$ is the fraction of star formation occurring in clusters. Mergers between pairs of globular clusters containing IMBHs can increase this rate by up to a factor of ~ 2 [39].

2.2.2 Continuous wave sources

The kinds of sources we consider in this section are ones which last for at least a few weeks or years, whose amplitude is constant (or at least roughly constant), and whose frequency varies relatively slowly over the observation time. These signals are expected to be produced by rapidly rotating non-axisymmetric neutron stars which are either isolated or in binary systems.

The waveforms for the two polarizations are taken to be

$$h_+(t) = A_+ \cos\Phi(t)\,, \qquad h_\times(t) = A_\times \sin\Phi(t) \tag{7}$$

where t is the time in the frame of the moving, accelerating detector, $\Phi(t)$ is the phase of the gravitational wave and $A_{+,\times}$ are the amplitudes; $A_{+,\times}$ are constant in time and depend on the other pulsar parameters such as its rotational frequency, moments of inertia, the orientation of its rotation axis, its distance from Earth etc. The phase $\Phi(t)$ takes its simplest form when the time coordinate used is τ, the proper time in the rest frame of the neutron star:

$$\Phi(\tau) = \phi_0 + 2\pi \sum_{n=0}^{s} \frac{f_{(n)}}{(n+1)!}\tau^{n+1} \tag{8}$$

where ϕ_0, $f_{(0)}$ and $f_{(n)}$ $(n \geq 1)$ are respectively the phase, instantaneous frequency and the spin-down parameters in the rest frame of the star at the fiducial start time $\tau = 0$, and s is the number of spin-down parameters included in our search. If ι is the angle between the line of sight to the star and its axis, then it is useful to write the amplitudes $A_{+,\times}$ in terms of a single number h_0

$$A_+ = \frac{1}{2}h_0(1+\cos^2\iota)\,, \qquad A_\times = h_0\cos\iota\,. \tag{9}$$

There are a number of mechanisms which may cause the star to be emitting gravitational waves. These include deformations of the neutron star crust, precession, magnetic fields, internal oscillation modes of the neutron star fluid etc.

Isolated neutron stars There are at present hundreds of pulsars known from either radio or X-ray observations. The parameters of many of these systems, i.e. the sky location and frequency evolution, have been accurately measured. We assume the GW phase evolution to be tightly correlated with the rotational phase as inferred from electromagnetic observations. For gravitational wave emission due to a non-negligible ellipticity, the GW emission occurs at twice the rotational frequency of the star. These two assumptions constrain the expected gravitational waveform upto an unknown initial phase ϕ_0, amplitudes $A_{+,times}$ and polarization angle ψ. It is then easy to search over these unknown parameters [40] and to either measure the amplitude h_0, or in the case that no signals are detected, to set upper limits on it.

The benchmark for these searches is the indirect upper bound on h_0 set by assuming that all of the kinetic energy of the star lost in the spindown is channeled into gravitational radiation. A straightforward calculation leads to the so-called spindown limit h_0^{sd}:

$$h_0^{sd} = 8.06\times10^{-19}\frac{I_{38}}{d_{kpc}}\sqrt{\frac{|\dot{\nu}|}{\nu}} \tag{10}$$

where $I_{38} = I/10^{38}\text{kg-m}^2$, d_{kpc} is the distance to the star in kpc, $\dot{\nu}$ is the spindown rate and ν is the spin frequency. This assumption is not expected to hold for any of the known pulsars where electromagnetic braking explains most of the spindown. Nevertheless, the spindown limit still a very useful benchmark for quantifying the astrophysical relevance potential targets and search results.

This procedure has been carried out for a number of known pulsars using data from the LIGO, GEO and Virgo detectors [41–43]. One highlight from these results is beating the spindown limit for the Crab pulsar [43] where the gravitational wave luminosity os constrained to be less than 6% of the spindown luminosity.

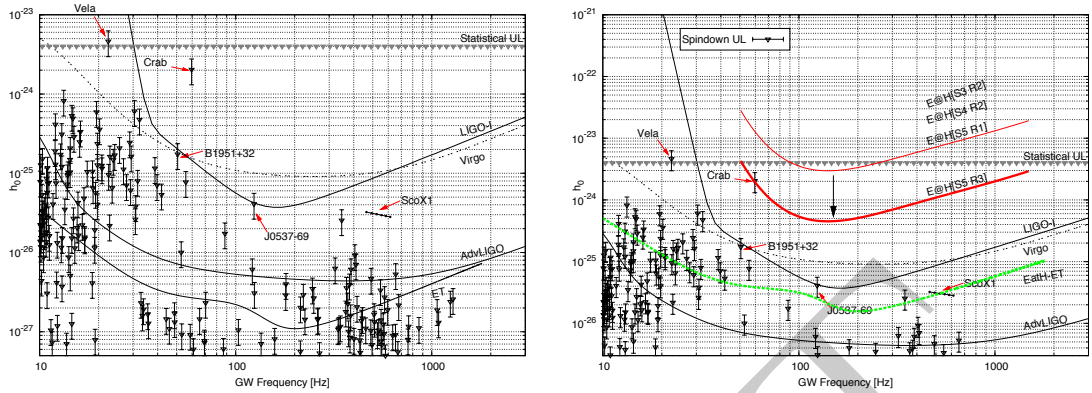


Figure 8: The left hand plot shows upper limits and the spindown limits for the known pulsars. Adapted from R.Prix, 2006. The right hand plot shows expected sensitivity for blind searches.

Let us now consider design noise curves for various detectors, including ET, and compare detectable values of h_0 with the spindown limits for a number of known pulsars. A useful benchmark for the detectability is given by

$$h_0 = 11.4\sqrt{\frac{S_n(f)}{DT_{obs}}} \tag{11}$$

where $S_n(f)$ is the detector noise power-spectral density at a frequency f, T_{obs} is the observation time, and D is the number of detectors. The factor of 11.4 corresponds to a false alarm rate of 1% and a false dismissal rate of 10%. Figure 8 shows (left panel) the detectable amplitude for Initial and Advanced LIGO, Virgo and ET, and spindown limits for various known pulsars.

Let us turn now to the wide parameter space searches. Here we don't target a known pulsar but rather, as an example, one whose radio pulse is not beamed towards; such a neutron star might still be visible in the GW sky. These searches are computationally limited because the number of templates increases much faster than linearly with the observation time T_{obs}. The large number of templates affects the search sensitivity in three basic ways. The first and most obvious one is simply the discreteness of the template grid. Secondly, it also leads to a large number of statistical trials which increases the false alarm rate and thus leads to a larger effective threshold. Finally, and most importantly, it limits the largest observation time that we can consider; even given the increases in computer power following Moore's law, this will most likely still be true in the ET era.

The problem of computational cost is addressed by the so-called *semi-coherent* methods. These rely on breaking up the full data set into shorter segments, analyzing the segments coherently and combining the power from the different segments incoherently; there are a number of different techniques available for performing the incoherent combination. For these searches, the sensitivity, incorporating all the effects mentioned above is typically given by

$$h_0 \approx \frac{25}{N^{1/4}}\sqrt{\frac{S_n(f)}{DT_{coh}}}\,. \tag{12}$$

This has been found to be a fairly good estimate (within $\sim 20\%$) of previous semi-coherent searches (see e.g. [44]).

Two sensitivity curves of ET have been proposed by S. Hild (the "pink" and "green" curves in his presentation at the Cardiff WP4 Meeting:
https://workarea.et-gw.eu/et/WG4-Astrophysics/meetings/cardiff-090325/
ET Sensitivity News, slide 10). One is the *ET-B* curve and the other, that we call it ET-B2, has a much worse sensitivity below $\sim$ 3 Hz but better around 10 Hz. Let us see what changes in the two cases for the search of GW from both known and unknown neutron stars. In Fig. (9) the top plot shows the minimum detectable

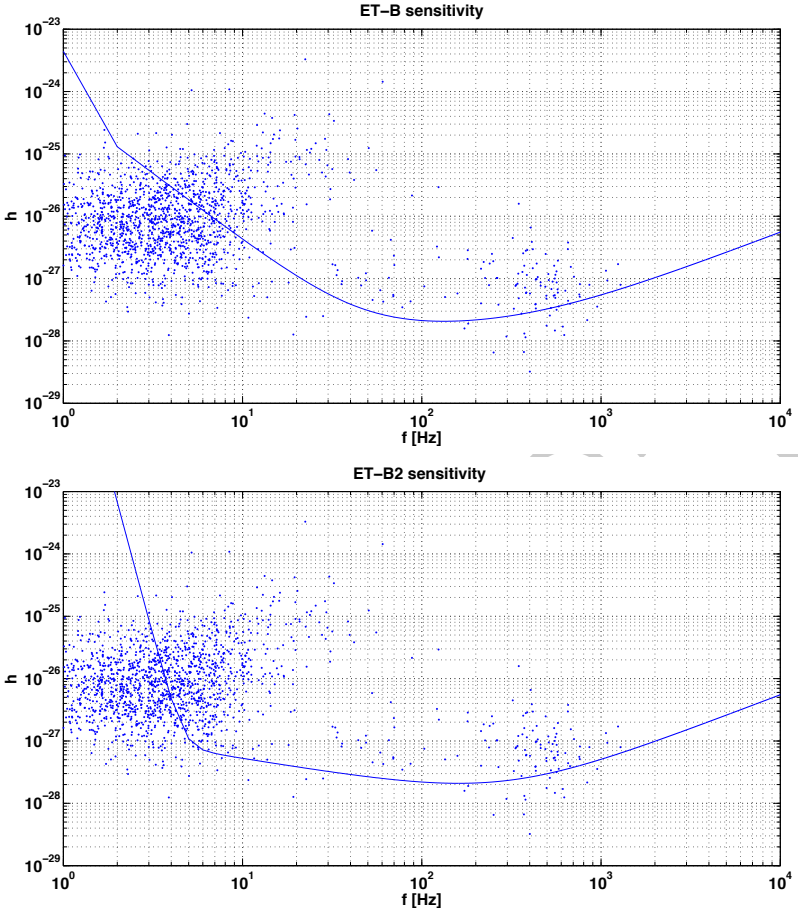


Figure 9: Minimum detectable amplitude with *ET-B* sensitivity (top) and *ET-B2* (bottom). An observation time $T_{obs} = 5\,\mathrm{yr}$, a false alarm probability of 1% and a false dismissal probability of 10% are assumed.

amplitude, assuming *ET-B* sensitivity, an observation time $T_{obs} = 5\,\mathrm{yr}$, a false alarm probability of 1% and a false dismissal probability of 10%, see Eq. (11), versus the spin-down limit of the known pulsars (taken from the ATNF Catalogue: http://www.atnf.csiro.au/research/pulsar/psrcat/). The bottom plot is done using (an approximate fit of) the *ET-B2* sensitivity curve. The important point is that no known pulsar (up to now) could emit a detectable signal with frequency below $\sim 2.5\,\mathrm{Hz}$. This means that there is no gain in having a good sensitivity at extremely low frequencies. On the other hand, having a better sensitivity around 10 Hz impacts positively on the possibility of detection.

This can be seen also in Fig. (10) where the ellipticity corresponding to the minimum detectable amplitude is plotted, only for the sources for which the spin-down limit can be beaten in the given observation time T_{obs}. Not only the number of pulsars for which the spin-down limit can be beaten is larger for the *ET-B2* curve (774 vs. 444) but, more important, the minimum ellipticy needed to produce a detectable signal is ~ 1 order of magnitude lower in the 10 Hz range. For instance, with *ET-B* we typically need ϵ in the range $0.1 - 5 \cdot 10^{-4}$ range for pulsars emitting around 10 Hz, while $\epsilon \sim 0.1 - 1 \cdot 10^{-5}$ is enough with *ET-B2*. The very few pulsars at frequencies below $\sim 3\,\mathrm{Hz}$ for which the spin-down limit could be beaten with *ET-B*, but not with *ET-B2*, correspond to ellipticity in the 10^{-2} range, a value difficult to reach also assuming an exotic equation of state for neutron star matter. We must however keep in mind that the number of pulsars increases with decreasing frequency, and so also the probability that an extremely deformed, EM-dim, neutron star exists, provided such large deformations are really attainable in Nature.

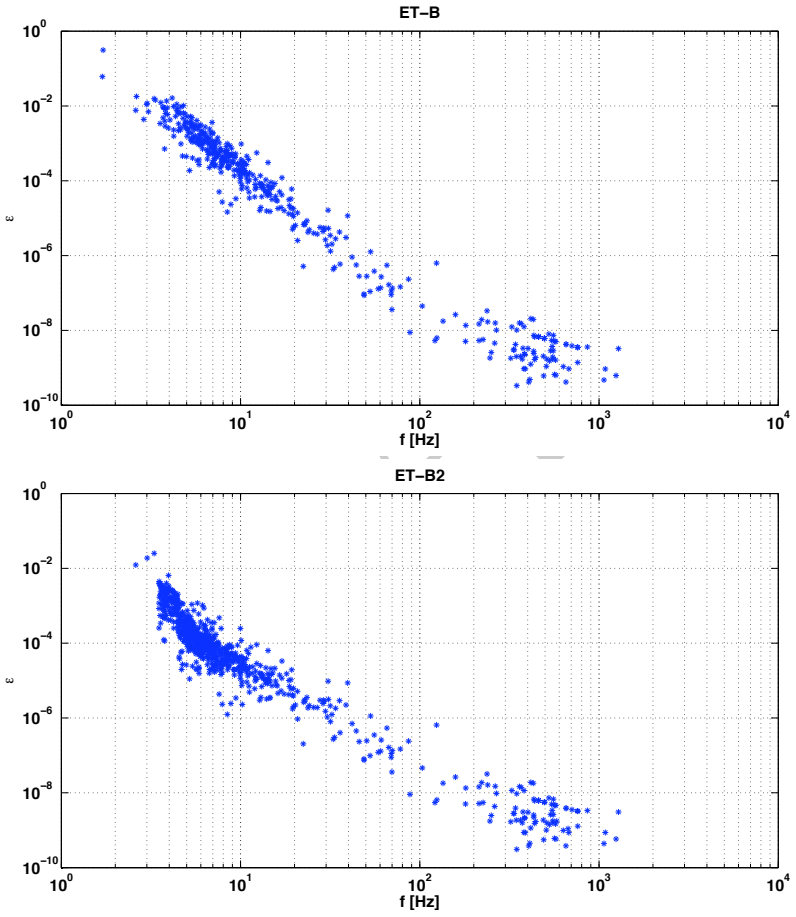


Figure 10: Minimum detectable ellipticity for known pulsars assuming *ET-B* sensitivity (top) and *ET-B2* sensitivity (bottom). The search parameters are the same as for Fig. (9).

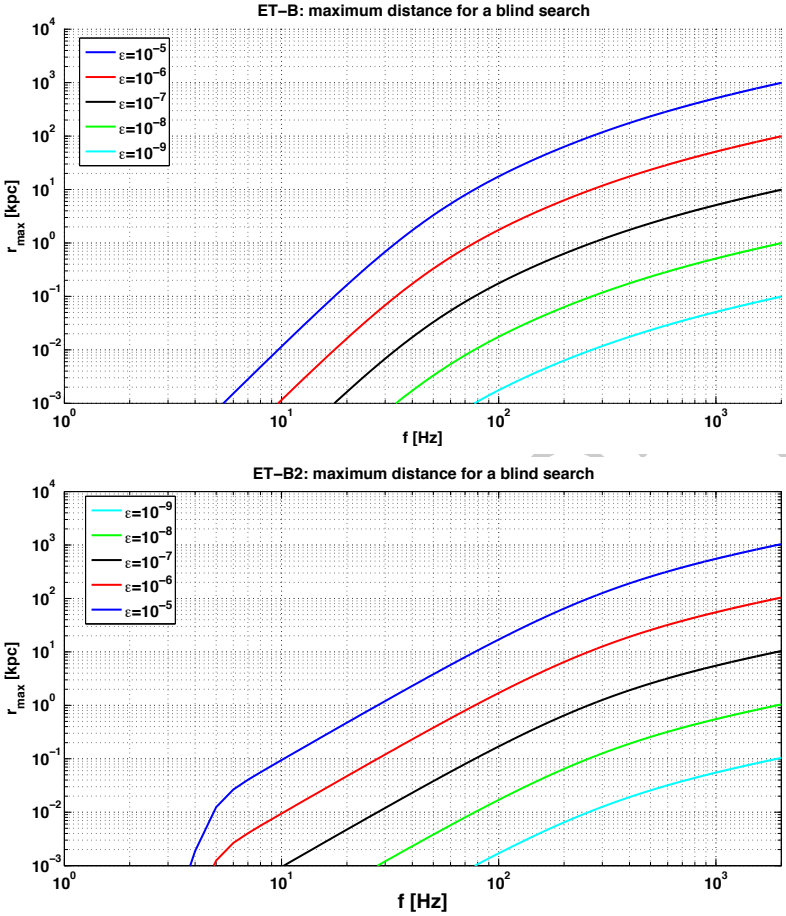


Figure 11: Maximum distance of an unknown source in order to be selected among the candidates of an all-sky search with *ET-B* sensitivity (top) and *ET-B2* sensitivity (bottom). Search parameters are given in the text.

Let us now consider the *blind* search for unknown neutron stars. For this case we plot in Fig. (11) the maximum distance of a source to be selected among the candidates of an all-sky incoherent or semi-coherent search, for different values of the neutron star ellipticity. An observation time $T_{obs} = 5\ yr$ and an FFT duration $T_{FFT} \simeq \frac{1.1 \cdot 10^5}{\sqrt{f}}\ s$, such that the Doppler effect does not spread the signal power outside a frequency bin, are assumed. Moreover, the threshold for the selection of candidates is chosen in order to have 10^9 candidates.

In practice, we do not expect detections for signal frequencies below $\sim 10\,\mathrm{Hz}$ for $\epsilon < 10^{-5}$ (the corresponding r_{max} becomes unrealistically small). And also considering extremely deformed neutron star ($\epsilon > 10^{-5}$) signal frequencies below $\sim 3\,\mathrm{Hz}$ are basically excluded. Then, having a better sensitivity at very low frequencies gives basically no gain. On the other hand, having a better sensitivity around 10 Hz somewhat increases the possibility of detection: for instance, assuming $\epsilon = 10^{-5}$, the maximum distance that a search can reach goes from $\sim 10\,\mathrm{pc}$ with *ET-B* to $\sim 100\,\mathrm{pc}$ with *ET-B2* at 10 Hz, while it goes from $\sim 150\,\mathrm{pc}$ to $\sim 500\,\mathrm{pc}$ at 20 Hz.

This conclusion does not change even assuming a long coherent step (compatible with the computing power we can think will be available in the ET era), because the sensitivity increases only as $T_{coh}^{1/4}$, see Eq. (12).

Low-mass X-ray binaries Observations of accreting neutron stars lead to perhaps the most important reason why, irrespective of the mechanism at work, at least some neutron stars might be actally emitting detectable

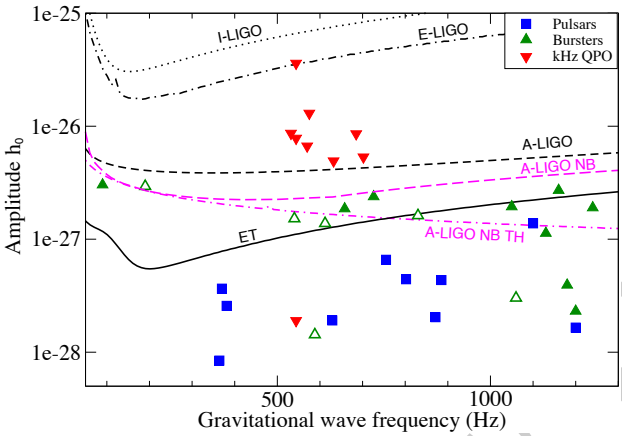


Figure 12: Sensitivity and the spin-balance limit for the accreting neutron stars.

gravitational waves. This is the observation that even the fastest accreting neutron stars spin at rates much lower than the expected break-up frequency. The current record is 716 Hz, while the theoretically expected upper limit is more than 1 kHz. Following a suggestion by Bildsten [45], it is possible that this limit occurs because of the balance between the spin-up torque due to the accreting matter, and the spindown torque due to gravitational wave emission. A short calculation assuming a link between the observed X-ray luminosity with the accretion rate, and taking the mountain scenario for the emission mechanism leads to the following estimate of the GW amplitude:

$$h_0 = 3 \times 10^{-27} F_{-8}^{1/2} \left(\frac{R}{10\mathrm{km}}\right)^{3/4} \left(\frac{1.4 M_\odot}{M}\right)^{1/4} \left(\frac{1\ \mathrm{kHz}}{\nu_s}\right)^{1/2}. \tag{13}$$

This is seen to be depend on frequency: $h_0 \propto \nu_s^{-1/2}$.

2.2.3 Stochastic background

The superposition of a large number of unresolved sources of gravitational waves produces a stochastic background, which could be detected by cross-correlating two (or more) detectors [46]. We can distinguish between two contributions: a background of cosmological origin, a memory of the early stages of the Universe (see Section 2.5.5), and a background of astrophysical origin, a memory of the evolution of the galaxies and star formation.

The astrophysical contribution is important for at least two reasons. On the one hand, it may mask the cosmological background in some frequency windows; on the other hand, its detection would put strong constraints on the physical properties of compact objects and their evolution with redshift, such as the mass of neutron stars or black holes, the ellipticity and the magnetic field of neutron stars or the rate of compact binaries. What is particularly interesting is that using stochastic searches, we are able to put constraints on the mean values and not on the properties of the brightest sources, more likely in the tail of the distributions.

The spectrum of the gravitational stochastic background is usually characterized by the dimensionless energy parameter [46]:

$$\Omega_{gw}(\nu_o) = \frac{1}{\rho_c} \frac{d\rho_{gw}}{d \ln \nu_o} \tag{14}$$

where ρ_{gw} is the gravitational energy density, ν_o the frequency in the observer frame and $\rho_c = 3H_0^2/(8\pi G)$ the critical energy of the Universe. For a stochastic background of astrophysical origin:

$$\Omega_{gw}(\nu_o) = 5.7 \times 10^{-56} \nu_o \int_{z_{\min}}^{z_{\max}} \frac{\dot{\rho}^o(z)}{(1+z)E(z)} \frac{dE_{gw}}{d\nu}(\nu_o) dz \tag{15}$$

where $\dot{\rho}^o(z)$ is the number of events in an element of comoving volume and interval of time in the observer frame, $\frac{dE_{gw}}{d\nu}$ the typical spectral energy density of a single source and $E(z)$ a function that depends on the cosmology. We assume $h_0 = 0.7$ for the Hubble parameter and a flat Universe with 70% of dark energy.

Binary Neutron Stars Double neutron star coalescences, which may radiate about 10^{53} erg in the last seconds of their inspiral trajectory, up to $1.4 - 1.6$ kHz, may be the most important contribution in the ET frequency range [47–50]. In the quadrupolar approximation, the GW energy spectrum emitted by a binary system, which inspirals in a circular orbit is given by:

$$dE_{gw}/d\nu = \frac{(G\pi)^{2/3}}{3} \frac{m_1 m_2}{(m_1 + m_2)^{1/3}} \nu^{-1/3} \tag{16}$$

Assuming $m_1 = m_2 = 1.4$ $M_\odot$ for the star masses, the energy density increases as $\nu_o^{2/3}$ before it reaches a maximum of $\Omega_{gw} \sim 2 \times 10^{-9} \dot{\rho}_0$ at around 600 Hz, where $\dot{\rho}_0$ is the local rate in My^{-3} Mpc^{-3} (about 0.01 times the galactic rate). This means that ET should be able to detect the background from binaries even for the most pessimistic predictions of the coalescence rate, down to $\dot{\rho}_0 \sim 0.035$ (roughly equivalent to a galactic rate of 3 My^{-1}), for a signal-to-noise ratio of 3, after one year of observation.

Rotating neutron stars: tri-axial emission Rotating neutron stars with a triaxial shape may have a time varying quadrupole moment and hence radiate GWs at twice the rotational frequency. The total spectral gravitational energy emitted by a NS born with a rotational period P_0, and which decelerates through magnetic dipole torques and GW emission, is given by:

$$\frac{dE_{gw}}{d\nu} = K\nu^3 \left(1 + \frac{K}{\pi^2 I_{zz}} \nu^2\right)^{-1} \quad \text{with } \nu \in [0 - 2/\mathrm{P}_0], \tag{17}$$

where

$$K = \frac{192\pi^4 G I_{zz}^3}{5c^5 R^6} \frac{\varepsilon^2}{B^2 \sin^2 \alpha}. \tag{18}$$

Here R is the radius of the star, $\varepsilon = (I_{xx} - I_{yy})/I_{zz}$ the ellipticity, I_{ij} the principal moment of inertia, B the magnetic field and α the angle between the rotation and the dipole axis.

The majority of neutron stars are born with magnetic fields of the order of $10^{12} - 10^{13}$ G and rotational periods of the order of tens or hundreds of millisecond [51–53], and very likely don't contribute very much to the stochastic background. But the population of newborn magnetars in which super-strong crustal magnetic fields ($B \sim 10^{14} - 10^{16}$ G) may have been formed by dynamo action in a proto-neutron star with very small rotational period (of the order of 1 ms) [54, 55], may produce a stochastic background detectable by ET [56].

For these highly magnetized neutron stars, the distortion induced by the magnetic torque becomes significant, overwhelming the deformation due to the fast rotation. When the deformation of the star is small ($K >> \pi^2 I \nu^{-2}$), the spindown is dominated by the magnetic torque but as the ellipticity increases GW emission may become the most important process. Taking $R = 10$ km for the radius, $I_{zz} = 10^{45}$ g cm^2 for the moment of inertia, and assuming that magnetars represent 10% of the population of NSs, we find that the stochastic signal is detectable with ET after an observation time $T = 1$ yr and with a signal to noise ratio of 3 when $\frac{\varepsilon}{B} > 1.5 \times 10^{-18}$. In the saturation regime where the spindown is purely gravitational, the energy density increases as ν_o^2 at low frequencies and reaches a maximum of $\Omega_{gw} \sim 1.3 \times 10^{-8}$ around 1600 Hz, giving a signal detectable by ET with a signal-to-noise ratio of 45.

Rotating neutron stars: initial instabilities The gravitational wave background signal from core collapse supernovae could be enhanced by a number of proposed post-collapse emission mechanisms. One intriguing mechanism is the bar-mode dynamical instability associated with neutron star formation. These instabilities derive their name from the 'bar-like' deformation they induce, transforming a disk-like body into an elongated

bar that tumbles end-over-end. The resulting highly non-axisymmetric structure resulting from a compact astrophysical object encountering this instability makes such an object a potentially strong source of gravitational radiation and have been the subject of a number of numerical studies [57–61]. Howell et al. have calculated the background signal from this emission process using simulated energy spectra data, $dE_{gw}/d\nu$, from [62], who performed the first three dimensional hydrodynamic simulations for stellar core collapse in full general relativity. Assuming a 20% occurrence of this instability, the authors find that the resulting background reaches a maximum of $\Omega_{gw} \sim 4\times10^{-10}$ around 2000 Hz would be detectable with a SNR of 3 after one year of integration. The optimistic event rate considered here is supported by suggestions that post collapse neutrino emission by the proto-neutron stars can induce contraction through cooling. This leads to increased spins though conservation of angular momentum [62]. The implication here is that the instability can set in tens of milliseconds post collapse, increasing the rate of occurrence.

The stochastic background from r-modes was first investigated by [63] and then reviewed by [64]. These estimates are based on the initial model of [65], which does not account for dissipation mechanisms such as the effect of the solid crust or the magnetic field, which may significantly reduce the gravitational instability. The spectral energy density of a single source is given by:

$$\frac{dE_{gw}}{d\nu} = \frac{2E_o}{\nu_{\rm sup}^2}\nu \ \text{ with } \ \nu \in [0-\nu_{\rm sup}], \tag{19}$$

where $\nu_{\rm sup}$ is 4/3 of the initial rotational frequency and E_0 is the rotational energy lost within the instability window. For neutron stars with radius $R = 10$ km and mass $M = 1.4$ M$_\odot$ the spectrum evolves as $\Omega_{gw} \sim 10^{-12}\xi\nu_o^3$ where ξ is the fraction of NS stars born near the keplerian velocity and which enter the instability window, until it reaches a maximum at 900 Hz. ET may be able to detect this signal with a SNR > 3 and $T = 1$ yr if $\xi > 0.23\%$. We obtain similar constraints with the secular bar mode instability at the transition between Maclaurin and Dedekind configurations [66].

Core collapse The GW background from core collapse supernovas that result in the formation of black holes was first calculated in [69] using the relativistic numerical simulations of [70] and later by [71] who found similar results assuming that all the energy goes into the ringdown of the $l = m = 2$ dominant quasi normal mode. The frequency ν_* of this mode is given by [72]:

$$\nu_* \approx \frac{c^3}{2\pi G}(1-0.63(1-a)^{0.3})\frac{1}{M(\mathrm{M}_\odot)} \tag{20}$$

where the mass of the BH is a fraction α of the mass of the progenitor and where a is the dimensionless spin factor ranging from 0 for a Schwarzschild BH to 1 in the extreme Kerr limit. The spectral energy distribution can be written as:

$$\frac{dE_{gw}}{d\nu} = \varepsilon M_{\rm bh}c^2\delta(\nu-\nu_*(M)) \tag{21}$$

where ε is the efficiency coefficient. Previous numerical simulations of [70] gave an upper limit of $\varepsilon \sim 7\times10^{-4}$ for an axisymmetric collapse, but accounting for more realistic scenarios, in particular the pressure reduction that triggers the collapse, [**?**] obtained an efficiency of the order of $10^{-7}-10^{-6}$ so $2-3$ orders of magnitude smaller. Assuming that stars in the range $30-100$ M$_\odot$ can produce a BH, taking $\alpha = 10\%$, and $a = 0.6$ we find that the energy density ranges between $0.25-5.6$ kHz, with a maximum of $\Omega_{gw} \sim \varepsilon\times10^{-8}$ around 1650 Hz, which means that an efficiency $> 2\times10^{-3}$ would give a signal detectable with a signal to noise ratio of 3 after one year of observation. Taking $\alpha = 20\%$, we find that the signal is detectable for efficiencies as small as 0.01%.

New estimates of the GW backgrounds generated by Pop III and Pop II sources have been recently published by [73]. These authors use the results of a numerical simulation by Tornatore et al.(2007) which follows the evolution, metal enrichment and energy deposition of both Population III and Population II stars. They predict the redshift dependence of the formation rate of black hole remnants of Population III stars with masses $100-500M_\odot$ and of neutron stars (black holes) remnants of Population II stars with masses $8-20M_\odot$ $(20-40M_\odot)$. In order to characterize the single source emission, the most appropriate signals available in the literature have been adopted, namely:

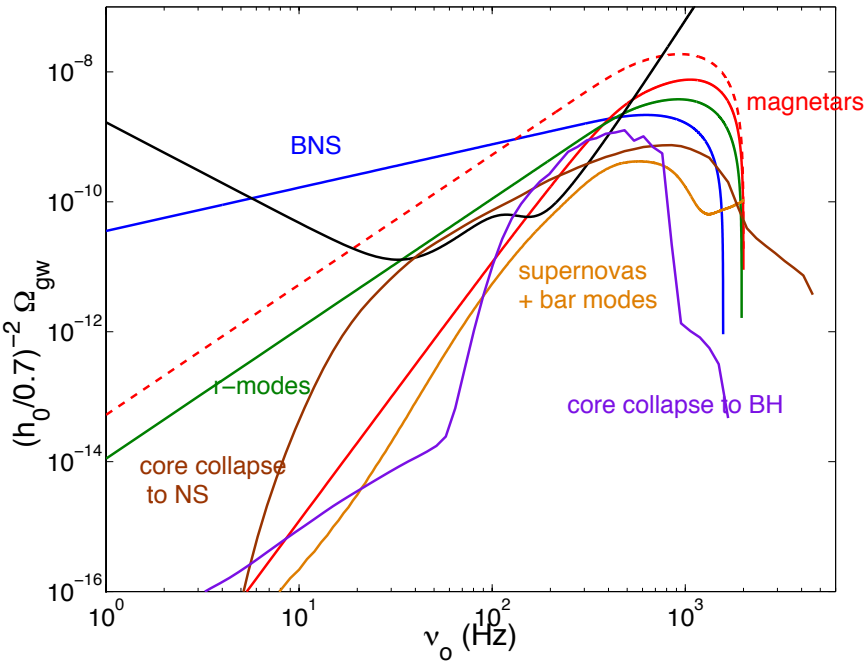


Figure 13: Energy density of the different contributions to the astrophysical background discussed in the text: magnetars (minimal detectable prediction in continuous red and model when the spindown is purely gravitational in dashed red), binary neutron stars in blue, dynamical bar modes in proto NS in yellow, r-modes assuming that 1% of newborn neutron stars cross the instability window in green, core collapse to black holes assuming an efficiency of 1% , a mass fraction of the progenitor of 10% and an angular parameter of 0.6 in purple, Pop II core collapse to NS (model of [67]) in brown and to BH (model D5a of [68]) in brown.

- For Pop III stellar collapse, the waveform recently obtained by [74] using a 2D numerical simulation which follows the entire evolution of a zero metallicity $300M_\odot$ star, taking the effects of General Relativity. and neutrino transport into account
- For Pop II progenitors with masses in the range $8-20M_\odot$ the emission spectrum corresponding to the model labelled as s15r in [75]; as an extreme and promising possibility, the gravitational wave emission produced by the excitation of the g-modes has also been considered using the template spectrum of [67].
- For Pop II progenitors with masses in the range $20-100M_\odot$, the gravitational wave spectra from a set of models (D5a,D5d,A5b) obtained by [68] using numerical simulations in full General Relativity.

The background is out of reach for Pop III stellar collapse, but could be be detected for Pop II progenitors. The authors found that the energy density reaches a maximum of $\Omega_{gw} \sim 10^{-9}$ around 1000 Hz for the collapse to NS model of Ott (2005), giving a signal to noise ratio of 8.2, and of $\Omega_{gw} \sim 4-7\times10^{-10}$ around 500 Hz for the collapse to BH model of [68], giving a signal to noise ratio of between $1.6-7.1$.

The sources discussed in this section are summarized in Figure 13: here the ET sensitivity curve is estimated for two co-located or minimally separated interferometers with opening angle 60° and relative rotation angle 120°, an integration time of 1 year and a detection SNR of 2.56 [46]. However, a search for stochastic background with co-located detectors could encounter difficulty in separating signal from correlated noise sources: different ET topologies and detection strategies are under consideration.

2.2.4 Probing Core-Collapse Supernova Physics

Stellar collapse is the most energetic event in the Universe, releasing $\sim10^{53}$ erg of gravitational energy in the compression of a massive star's iron core to a neutron star. Most of this energy ($\sim 99\%$) is emitted in neutrinos and only about 10^{51} erg go into energy of the core-collapse supernova (CC-SN) explosion. CC-SNe (SN types II, Ib, Ic) are $\sim$10 times more frequent than thermonuclear type-Ia SNe. A SN explosion pollutes the interstellar medium with the nucleosynthetic products of stellar evolution (CC-SNe are the Universe's primary source of oxygen) and enriches via the r-process the universe with rare heavy isotopes. The perturbation caused by an SN in its vicinity can trigger the formation of stellar systems and stellar collapse and CC-SNe are the birth sites of neutron stars (NSs) and stellar-mass black holes (BHs).

The Supernova Problem and GW observations The precise mechanism of explosion operating in CC-SNe is unknown [76–78]. When the inner part of the collapsing iron core reaches densities close to those in atomic nuclei, the strong force leads to a stiffening of the nuclear equation of state (EOS), resulting in *core bounce* of the inner core into the still infalling outer core. A shock wave is formed that propagates outward in mass and radius, but quickly loses energy due to the breakup of heavy nuclei and neutrinos that carry away energy from the postshock layer. The shock stalls, turns into an accretion shock and must be *revived* to drive a CC-SN explosion. If this does not happen, a BH will form on an accretion timescale of $\sim 2\,$s. *What is the mechanism of shock revival?* This is the fundamental question and primary unsolved problem of CC-SN theory. Indications are strong that the CC-SN mechanism involves a multitude of multi-dimensional processes, including rotation, convection/turbulence, and various hydrodynamic instabilities of the stalled shock and in the proto-NS. This opens up the possibility of probing the supernova mechanism with gravitational waves (GWs). GWs, even more so than neutrinos, carry direct dynamical information from the supernova engine deep inside a dying massive star, a region generally inaccessible by the traditional means of observational astronomy. GWs form a core-collapse event have the potential of putting very strong constraints on the CC-SN mechanism [78, 79]. With initial and certainly second-generation interferometric GW detectors, this should be possible for an event in the Milky Way ($D\sim10-15\,$kpc) and the Magellanic Clouds [78] ($D\sim50-70\,$kpc), but even optimistic estimates of the CC-SN rate in this region do not predict more than $\sim 1-2$ events per century. This number roughly doubles if one includes the entire local group ($D\sim1\,$Mpc). In the region from $3-5\,$Mpc a number of starburst galaxies increase the predicted and observed integrate SN rate to $\sim 0.5\,\mathrm{yr}^{-1}$. At $D\sim10\,$Mpc it is $\gtrsim 1\,\mathrm{yr}^{-1}$.

Supernova Science with ET The GW emission processes in a CC event emit GW strains h in the range $10^{-24}-10^{-22}\,(D/1\,\mathrm{Mpc})$ and most of the emission takes place at frequencies of $\sim 200-1000\,$Hz, but the

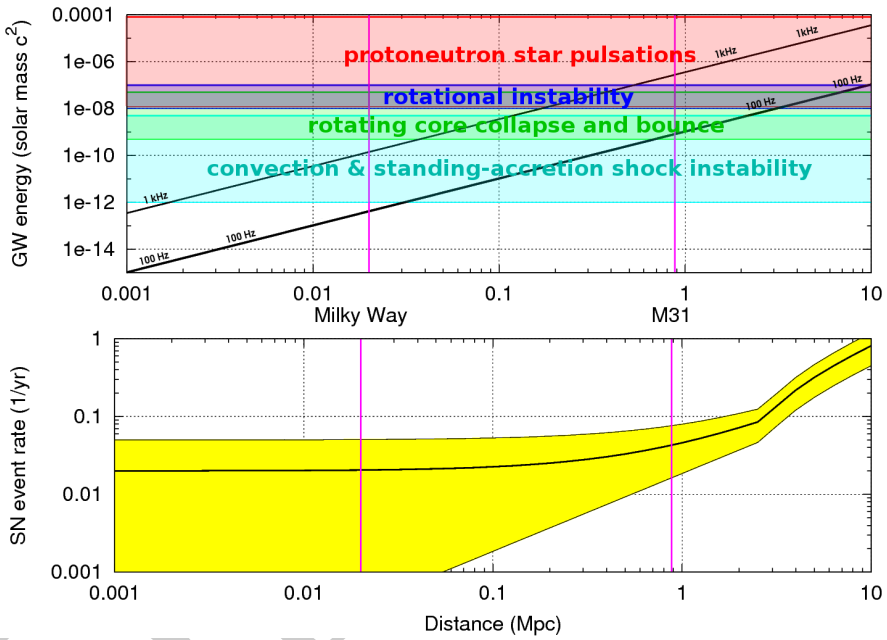


Figure 14: The upper plot displays the minimum GW energy that a supernova core collapse is required to radiate in order to be detectable by the Einstein telescope. We give two estimates assuming that the GW signature is described by a sine-Gaussian burst waveform. We consider two cases with a low-frequency $f = 100\,\mathrm{Hz}$ and with high-frequency $f = 1\,\mathrm{kHz}$ content. The minimum GW energy is given as a function of the source distance. We also indicate the expected range of radiated GW energy for several processes [78]. The lower plot shows an estimate of the cumulative event rate (with error bars) obtained from the star formation rate computed over a catalog of nearby galaxies [80].

various explosion scenarios exhibit unique spectral distributions and vary in total emitted energies [78, 79]. In addition, there is likely to be a low-frequency GW-memory-type component with large h up to 10^{-22} $(D/1\,\mathrm{Mpc})$ at $0-20\,\mathrm{Hz}$. ET as currently envisioned [81] is sufficiently sensitive to detect GWs from various CC-SN scenarios out to $2-4\,\mathrm{Mpc}$. If the high-f sensitivity was increased by a factor of $\sim 2-3$, detection out to $\sim 10\,\mathrm{Mpc}$ may be possible. Figure 14 summarizes the ET observational capabilities and examines each of the main generation processes of gravitational waves.

Even without this improvement, ET may see multiple CC-SNe during its lifetime and would have the power to provide strong hints for a particular SN mechanism and/or smoking-gun evidence against another – crucial astrophysics information that is unlikely to be attainable in other ways. At ET's implementation, megaton-class neutrino detectors will be operative and, having range similar to ET, will be able to provide coincident observations, narrowing down the time of the GW emission to $\sim 1\,\mathrm{ms}$. In addition, deep high-cadence optical transient surveys will be operative and targeting near-universe transients, providing additional coincident data as well as additional astrophysics output (progenitor type/mass, explosion morphology/energy etc.).

Impact Constraining the CC-SN mechanism will mean a breakthrough in our understanding of the large range of phenomena associated with stellar collapse, CC-SNe, BH and NS formation, and gamma-ray bursts (GRBs). However, the astrophysics and physics information provided by GWs observed from a CC event with ET goes beyond this: These GWs carry also information on the high-density nuclear EOS, explosion asymmetries and pulsar kicks, the formation of a BH in a failing CC-SN, and can help uncover rare events such as the accretion-induced collapse of a white dwarf to a NS or weak or failing CC-SNe that have very weak or absent EM signatures.

2.3 Strong field tests of GR and fundamental physics

The rich variety of sources and phenomena observed by gravitational wave detectors can be potentially used to address outstanding questions in fundamental physics. The sources in question will be in dense environs of ultra-strong gravity and thereby provide a cosmic laboratory for understanding phenomena and matter in extreme conditions of density, temperature, magnetic field, etc. Moreover, black hole binaries are fundamentally geometric objects whose interaction close to merger will provide insights into the nature of black hole space times and of gravity in ultra-strong fields. In this Section we will discuss what fundamental physics questions and strong-field tests of gravity could be addressed by 3G detectors.

2.3.1 Speed of gravitational waves and mass of the graviton

In Einstein's theory gravitational waves travel with the speed of light. This means that gravitons, particle analogs of gravitational waves, are massless particles. Although there is currently no strong motivation to consider massive graviton theories from an experimental point of view, they are natural extensions of Einstein's theory. In a massive graviton theory, gravitational waves would not travel at the speed of light and this can be tested by observation of gravitational-wave sources at very great distances. To do so we would need a source which emits at the same time both gravitational waves and electromagnetic radiation. By measuring the difference in their arrival times we could measure or constrain the speed of gravitational waves.

Supernovae in the local Universe and double neutron star and neutron star-black hole binaries are sources that are expected to exhibit after glows in electro-magnetic radiation soon after they emit a burst of of gravitational waves. If the source is near enough (a few Mpc in the case of supernovae and red-shifts of a few in the case of coalescing binaries) and the event is well-localized on the sky (fraction of a degree depending on the distance to the source), then it could be observed in coincidence as a transient EM and a GW event.

Current theories of supernovae and coalescing binaries cannot accurately predict how promptly after the collapse (in the case of SN) or merger (in the case of binaries) EM radiation will follow. However, the expected delay is no more than one second. If gravitational waves arrive a time Δt after EM waves, then the fractional difference

in their speeds is given by

$$\frac{|\Delta v|}{c} = 3.2 \times 10^{-18} \left(\frac{|\Delta t|}{1\,\mathrm{s}}\right) \left(\frac{3\,\mathrm{Gpc}}{D}\right) \tag{22}$$

where we have assumed that the source is at a distance of $D = 3\,\mathrm{Gpc}$.

2.3.2 Limiting the mass of the graviton

Observations of inspiralling compact binaries (neutron stars or black holes) can be used to put bounds on the mass of the graviton, or equivalently the compton wavelength of the graviton [82]. These bounds do *not* require the detection of an electromagnetic counterpart associated with the GW signal.

The basic idea is simple: if there is a mass associated with the propagation of gravitational waves ("a massive graviton"), then the speed of propagation will depend on wavelength in the form $v_g \approx 1 - (\lambda/\lambda_g)^2$, where λ_g is the Compton wavelength of the graviton, in the limit where $\lambda \ll \lambda_g$. Irrespective of the nature of the alternative theory that predicts a massive graviton , it is reasonable to expect the differences between such a hypothetical theory and general relativity in the predictions for the evolution of massive compact binaries to be of order $(\lambda/\lambda_g)^2$, and therefore to be very small, given that $\lambda \sim 10^3$ km for stellar mass inspirals and $\sim 10^8$ km for massive black hole inspirals.

As a result, the gravitational waveform seen by an observer close to the source will be very close to that predicted by general relativity. However, as seen by a detector at a distance D, hundreds to thousands of Mpc away, the phasing of the signal will be distorted because of the shifted times of arrival, $\Delta t \sim D(\lambda/\lambda_g)^2$ of waves emitted with different wavelengths during the inspiral. In addition to measuring the astrophysical parameters of the system, such as masses and spins, the matched filtering technique permits one to estimate or bound such effects.

Here we examine the bounds possible from the observations of binary black holes by ET [83]. As our waveform model we begin with amplitude-corrected, general relativistic waveforms which are 3PN accurate in amplitude [84–87] and 3.5PN accurate in phasing [88–93]. We ignore the spins of the bodies in the binary system. Previous calculations used waveforms which are of Newtonian order in amplitude and 2PN order in phase. As opposed to the Newtonian waveforms, the 3PN amplitude-corrected waveforms contain all harmonics from Ψ up to $8\,\Psi$, where Ψ is the orbital phase (the leading quadrupole component is at 2Ψ).

The effect of a massive graviton is included in the expression for the orbital phase following Ref. [82]. The wavelength-dependent propagation speed changes the arrival time t_a of a wave of a given emitted frequency f_e relative to that for a signal that propagates at the speed of light; that time is given, modulo constants,by

$$t_a = (1+Z)\left[t_e + \frac{D}{2\lambda_g^2 f_e^2}\right], \tag{23}$$

where f_e and t_e are the wave frequency and time of emission as measured at the emitter, respectively, Z is the cosmological redshift, and

$$D \equiv \frac{(1+Z)}{a_0}\int_{t_e}^{t_a} a(t)dt\,, \tag{24}$$

where $a_0 = a(t_a)$ is the present value of the scale factor (note that D is not exactly the luminosity distance [1]). This affects the phase of the wave accordingly. In the frequency domain, this adds a term to the phase $\psi(f)$ of the Fourier transform of the waveform given by $\Delta\psi(f) = -\pi D/f_e\lambda_g^2$. Then, for each harmonic of the waveform with index k, one adds the term

$$\Delta\psi_k(f) = \frac{k}{2}\Delta\psi(2f/k) = -\frac{k^2}{4}\pi D/f_e\lambda_g^2\,. \tag{25}$$

Here $k = 2$ denotes the dominant quadrupole term, with phase 2Ψ, $k = 1$ denotes the term with phase Ψ, $k = 3$ denotes the term with phase 3Ψ, and so on.

[1] For $Z \ll 1$, D is roughly equal to luminosity distance D_L. Hence we have assumed $D \simeq D_L$ in the case of ground based detectors for which we consider sources at 100 Mpc. For LISA, we have carefully accounted for this difference.

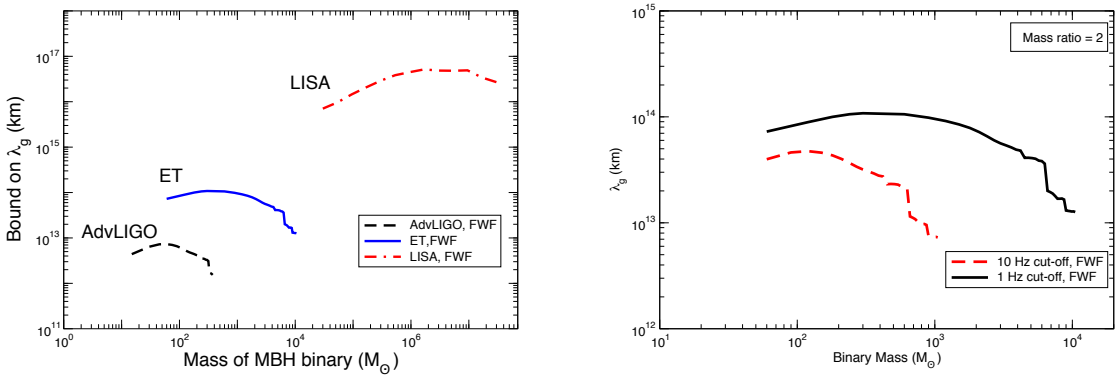


Figure 15: Left panel: Bounds on the graviton Compton wavelength that can be deduced from AdvLIGO, Einstein Telescope and LISA. The mass ratio is 2. The distance to the source is assumed to be 100 Mpc for AdvLIGO and ET, and 3 Gpc for LISA. Right panel: Possible bounds from ET when 1 Hz and 10 Hz are used as seismic cut-offs.

This is an adhoc procedure because a massive graviton theory will undoubtedly deviate from GR not just in the propagation effect, but also in the way gravitational wave damping affects the phase, as well as in in the amplitudes of the gravitational waveform. If, for example, such a theory introduces a leading correction to the quadrupole phasing $\psi_{\rm quad} \sim (\pi\mathcal{M}f_e)^{-5/3}$ of order $(\lambda/\lambda_g)^2 \times (\pi\mathcal{M}f_e)^{-5/3}$, where $\mathcal{M}$ is the chirp mass, then the propagation induced phasing term (25) will be larger than this correction term by a factor of order $k^2(D/\mathcal{M})(\pi\mathcal{M}f_e)^{8/3} \sim (D/\mathcal{M})v^8$. Since $v \sim 0.1$ for the important part of the binary inspiral, and $D \sim$ hundreds to thousands of Mpc, it is clear that the propagation term will dominate. In any case, given the fact that there is no generic theory of a massive graviton, we have no choice but to omit these unknown contributions.

Our estimate of the bounds on the massive graviton parameter is based on the Fisher matrix formalism. We construct the Fisher matrix for the different detector noise PSDs using the amplitude corrected PN waveform model described earlier, converted to the Fourier domain using the stationary phase approximation. We use a six-dimensional parameter space consisting of the time and phase (t_c, ϕ_c) of coalescence, the chirp mass $\mathcal{M}$, the mass asymmetry parameter $\delta = |m_1 - m_2|/(m_1 + m_2)$, the massive graviton parameter $\beta_g = \pi^2 D\mathcal{M}/\lambda_g^2(1+Z)$, and the luminosity distance D_L. We fix the three angles, θ, ϕ and ψ which appear in the antenna pattern functions to be $\pi/3$, $\pi/6$ and $\pi/4$ respectively and the inclination angle of the binary to be $\iota = \pi/3$. Details of the Fisher matrix approach as applied to the compact binary coalescence signals can be found in Refs. [94–96], and more recently in Ref. [97] ,which critically reexamines the caveats involved in using the Fisher matrix formalism to deduce error bounds for various gravitational wave detector configurations.

The square root of each of the diagonal entries in the inverse of the Fisher matrix gives a lower bound on the error covariance of any unbiased estimator. Our focus here is solely on the diagonal element corresponding to the massive graviton parameter. The $1-\sigma$ error bar on β_g can be translated into a bound on the Compton wavelength using $\Delta\beta_g = \beta_g$, and this is the quantity that we use in the plots as well as in the discussions.

The results are shown in Figure 15. On the left panel, the bounds on the compton wavelength of the graviton achievable with the second generation detector AdvLIGO and the space based LISA are compared against those possible from ET. The typical bounds from ET could be an order of magnitude better than from AdvLIGO, but worse than LISA. The observable mass range is also much larger for ET in comparison with AdvLIGO and extends up to about $10^4 M_\odot$. In the right panel, we compare the effect of the seismic cut-off on the massive graviton bounds by considering 1 Hz and 10 Hz cut-offs. As one might expect, there is improvement in the accessible mass range by almost an order of magnitude when 1 Hz cut-off is used as cut-off as opposed to 10 Hz.

Note that though we have considered the sources for ET to be at 100 Mpc, the bounds in principle are more or

less independent of the distance because in the definition of λ_g, there is a distance scale present. However, for very large distances the SNR may not be high enough and Fisher matrix estimate may not be reliable.

2.3.3 Bounds on Brans-Dicke parameter using ET

The Brans-Dicke (BD) theory of gravity [98] is an alternative theory of gravity which has an additional scalar field, which couples to matter, apart from the tensor field of general relativity. The coupling of the scalar field is described by a constant parameter ω_{BD}; in the limit of GR, $\omega_{BD} \to \infty$. Since scalar-tensor field theories predict dipolar gravitational radation, this parameter is also a measure of the dipolar GW content.

The best bound on this parameter so far has come from the solar system experiment *Cassini*, by measuring the frequency shift of radio signas to and from the spacecraft as it orbited near the sun [99]. The resulting lower limit on ωBD is about 4×10^4.

Gravitational wave observations can also put interesting bounds on ω_{BD} [100, 101]. This is possible because the GW phasing formula for the BD case is same as that of GR except for an additional dipolar term proportional to ω_{BD}^{-1}. Hence it is possible to measure or bound this quantity from GW observations.

The dipolar GW content also depends on the internal structure of the compact body via a quantity called "sensitivity" s_A (see Sec. 3.3 of [102]).

$$\left(\frac{dE}{dt}\right)_{\rm dipole} \propto \frac{\mathcal{S}^2}{\omega_{BD}}, \qquad \mathcal{S} = s_1 - s_2 \tag{26}$$

where s_1 and s_2 are the sensitivities of the binary constituents. For binary neutron stars $\mathcal{S} \sim 0.05 - 0.1$, for NS-BH binaries $\mathcal{S} \sim 0.3$ and for a binary BH $\mathcal{S} = 0$. Therefore, for bounding BD theories one of the components of the binary should be a NS. The bound is also very sensitive to the asymmetry of the binary: the more asymmetric the binary, the worse is the bound. Due to these factors, GW bounds on the ω_{BD} are very weak (~ 5000 at best [100]). However, we point out that if ET has very good low frequency sensitivity (seismic

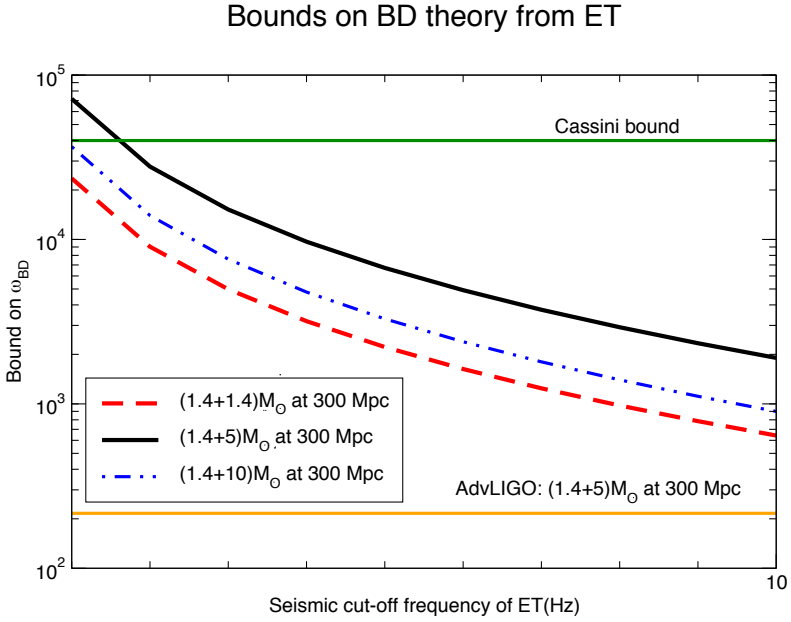


Figure 16: Bounds on Brans-Dicke parameter (ω_{BD}) from ET as a function of seismic cut-off frequency of ET. The existing bound from the Cassini experiment and the possible bounds from AdvLIGO are also shown.

cut-off frequency between $1-10$ Hz), the bounds from ET can beat the solar system bounds. Figure 16 shows the bound on ω_{BD} for different types of sources as a function of the seismic cut-off frequency of ET. The NS mass is assumed to be $1.4M_\odot$ and that of the BH to be either $5M_\odot$ or $10M_\odot$. The factor $\mathcal{S}$ of NS is assumed to be 0.1 and that of the BH is assumed to be 0.3. The best bounds would come from the observations of NS-BH binaries with the BH mass between $4-10M_\odot$ at 300 Mpc. If ET has a low frequency cut-off of 1Hz, then the bounds on ω_{BD} could be as high as $\sim 10^5$.

These bounds are likely to be the best possible by GW observations because estimates for LISA (a proposed space mission sensitive to low frequency GWs) for NS-BH binaries of total mass $\sim 10^3$ can beat the solar system bounds only if it observes a binary within 20-50 Mpc during its mission lifetime [103–105].

2.3.4 A more general test of GR and the effect of low frequency sensitivity

It is also possible to test for violations of GR without assuming a particular alternative model. One such a test was proposed by Arun et al. [106–108] and was based on the post-Newtonian expansion for the phase of an inspiral signal in the frequency domain:

$$\Psi(f) = -\phi_c + \sum_{j=0}^{7} \left[\psi_j + \psi_{jl}\ln(f)\right] f^{(j-5)/3}, \tag{27}$$

where the expressions for the coefficients ψ_j and ψ_{jl}, $j = 0,\ldots,7$ can be found in [108]. Under the simplifying assumption that spins are zero, all these coefficients only depend on the two component masses m_1, m_2. Hence only two of the ψ_j, ψ_{jl} are independent, and a possible test of post-Newtonian theory (and hence of GR) is to check for consistency between any three of them. Particular attention has been given to ψ_3 and ψ_{5l}:

- ψ_3 is the lowest-order coefficient which gets contributions from scattering of gravitational waves off the spacetime near the binary: the so-called tail terms; hence it encapsulates the non-linear character of GR.
- ψ_{5l} is the lowest-order coefficient of a logarithmic term in (27). General relativity is not consistent with a simple Taylor expansion.

Fig. 17 illustrates the proposed test of GR for the triplets of coefficients (ψ_0, ψ_2, ψ_3) and $(\psi_0, \psi_2, \psi_{5l})$.

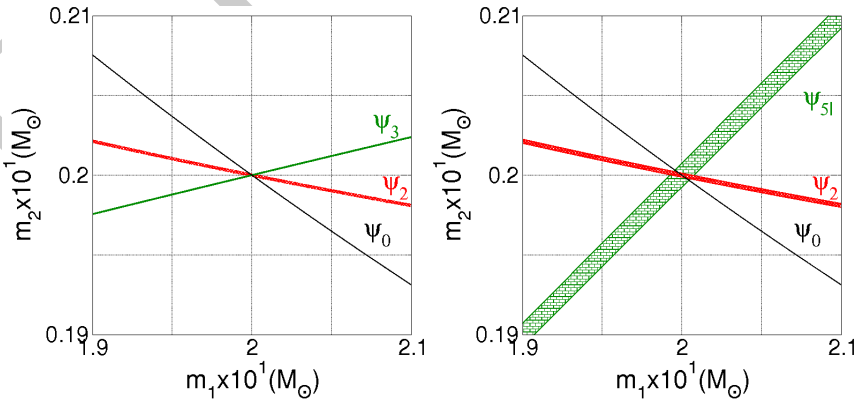


Figure 17: Plots showing the regions in the (m_1, m_2) plane corresponding to $1-\sigma$ uncertainties in (ψ_0, ψ_2, ψ_3) (left) and $(\psi_0, \psi_2, \psi_{5l})$ (right), for a $(2, 20)\,M_\odot$ binary black hole at 300 Mpc observed by ET. The lower cut-off frequency was assumed to be 1 Hz.

The sensitivity at low frequency can have a dramatic effect on the ability to test GR, as shown in Fig. 18.

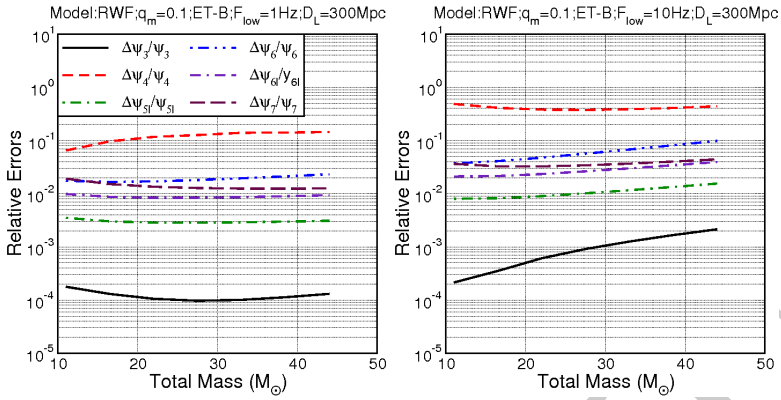


Figure 18: Varying the lower frequency cut-off can affect the accuracy in measuring phase coefficients by factors of several to an order of magnitude. On the left are relative errors with a 1 Hz cut-off, on the right with a 10 Hz cut-off.

2.3.5 Measuring the dark energy equation of state and its variation with z

Over the past decade, evidence has emerged suggesting that the expansion of the Universe is accelerating. Possible explanations include a failure of general relativity at large length scales, a cosmological constant in the Einstein equations, or a new contributor to the mass/energy content of the Universe called dark energy (see [109] for a review). Assuming a homogeneous and isotropic Universe, dark energy can be characterized by an equation of state of the form $p_{\rm DE} = w(z)\rho_{\rm DE}$, where $p_{\rm DE} < 0$ and $\rho_{\rm DE} > 0$ are the pressure and density, respectively. If the equation of state parameter $w(z)$ is constant and equal to -1 then this corresponds to having a positive cosmological constant in the gravitational field equations. Current constraints allow for this possibility, but other possibilities are not ruled out. The five year WMAP data combined with supernovae measurements and baryon acoustic oscillations in the galaxy distribution lead to the constraint $-1.11 < w < -0.86$ at the 95% confidence level [30].

The gravitational wave signal from inspiraling compact binaries (neutron stars and black holes) is particularly "clean" and well-understood. Consequently, as suggested by Schutz, one can think of using inspiral events as "standard sirens", much in the way Type Ia supernovae have been used as standard candles [5]. From the gravitational wave signal itself the luminosity distance $D_{\rm L}$ can be inferred, but not the redshift. However, if a particular compact binary coalescence event is accompanied by a sufficiently distinct electromagnetic counterpart, then it will be possible to find its position in the sky, identify the host galaxy, and obtain the redshift z. The relationship $D_{\rm L}(z)$ depends sensitively on cosmological parameters such as the Hubble constant at the current epoch H_0, the normalized matter and dark energy densities $\Omega_{\rm M}$ and $\Omega_{\rm DE}$, and the dark energy equation of state parameter w. For example, in a spatially flat FLRW Universe and assuming a constant w,

$$D_{\rm L}(H_0, \Omega_{\rm M}, \Omega_{\rm DE}, w; z) = (1+z) \int_0^z \frac{dz'}{H_0 \left[\Omega_{\rm M}(1+z')^3 + \Omega_{\rm DE}(1+z')^{3(1+w)}\right]^{1/2}}. \quad (28)$$

The intrinsic luminosity, and hence the luminosity distance, of an inspiral gravitational wave event can be inferred directly from their amplitude and from the component masses, which govern the structure of the signal. Thus, unlike Type Ia supernovae, their calibration does not depend on the brightness of other sources. Thus gravitational wave astronomy opens up the possibility of cosmography *without having to rely on the lower rungs of the cosmic distance ladder.*

Compact binary coalescences that involve a neutron star are assumed to have strong electromagnetic counterparts, mostly in the form of strongly beamed gamma radiation directed perpendicularly to the plane of the inspiral. Such events are believed to be the progenitors of short, hard Gamma Ray Bursts (GRBs): if the beam roughly points towards Earth then a flash of gamma radiation is seen, followed by an afterglow in the

lower-frequency electromagnetic spectrum. This would then allow us to identify the host galaxy and obtain a redshift.

The gravitational wave signal from a NS-BH coalescence will be visible out to $z = 3.5$. Within the corresponding volume, it is reasonable to expect $\sim 10^4$ or more such coalescences per year, but depending on the opening angle of the gamma ray beam only a few percent of these will be visible as a GRB. Hence we should have a few hundred sources at our disposal for which the redshift can be measured. The uncertainty on z will be negligibly small, while $D_{\rm L}$ will be measurable with $\sim 3\%$ inaccuracy at $z = 1$, rising to $\sim 10\%$ at $z = 3.5$. Fitting the measured values of $D_{\rm L}$ against redshift by varying H_0, $\Omega_{\rm M}$, $\Omega_{\rm DE}$, and w in the relationship (28) should then allow for the determination of these cosmological parameters with uncertainties of 5% or better, as discussed further in Section 2.5.4.

2.3.6 Testing the uniqueness theorem of black hole spacetimes

It is generally accepted that the massive compact objects observed in the centres of most galaxies are massive, rotating black holes described by the Kerr metric of General Relativity. This belief comes in part from the uniqueness theorem, which is the result that the Kerr metric is the unique endstate of gravitational collapse [110]. However, this theorem is based on several assumptions – the spacetime is vacuum, axisymmetric and stationary; there is a horizon in the spacetime; and there are no closed timelike curves. If one of these assumptions were violated, then objects that deviate from the Kerr metric could exist.

In black hole binary systems where the mass of one object is much bigger than the other, many gravitational wave cycles are emitted while the smaller object is in the strong field region close to the larger object. These gravitational waves encode a map of the spacetime structure in the vicinity of the large black hole, which can be used to measure properties of the central object [111]. Using such observations to measure spacetime structure has been explored extensively in the context of extreme-mass-ratio inspirals (binaries of $\sim 10 M_\odot$ objects with $\sim 10^6 M_\odot$ objects) for LISA (see [38] and references therein). There is an analogous source for ground based detectors, namely the inspiral of a $\sim 1 M_\odot$ object into a $\sim 100 M_\odot$ black hole. We refer to these as intermediate-mass-ratio inspirals (IMRIs) [112].

Explicit calculations have not yet been done for ET, but they do exist for Advanced LIGO [35, 112]. Extrapolating from those results, ET could see IMRI events out to a redshift $z \sim 3$ and could detect as many as several hundred events per year, although a few to a few tens is more likely. ET will observe these events for more cycles than Advanced LIGO, due to its better low-frequency performance, which is very important for the precision of spacetime mapping measurements.

Testing the black hole no-hair theorem The uniqueness of Kerr black holes as the endstate of collapse is sometimes referred to as the "no-hair theorem". A Kerr black hole has "no-hair" since the entire spacetime structure, characterised by "multipole moments", is determined by just two parameters, the black hole mass, M, and spin, S. It has been demonstrated that gravitational wave observations can measure the multipole moments independently of one another [111]. We can therefore directly verify that they satisfy the Kerr relationship

$$M_l + \mathrm{i} S_l = M(\mathrm{i}S/M)^l \tag{29}$$

We need only to measure three multipole moments to rule out an object as a Kerr black hole.

It has been shown that IMRI observations with Advanced LIGO could detect an $O(1)$ deviation in the quadrupole moment of an object [112]. The precision achievable with ET should be at least a factor of 10 better than this due to the improved low-frequency performance. To put this in perspective, one alternative to black holes, boson stars, have quadrupole moments two orders of magnitude bigger than black holes of the same mass and spin [113].

Any deviations from the no-hair theorem that are detected will have profound implications for our understanding of relativity and of black holes. Persistent deviations from the theory may lead to important insights in the search for a fundamental theory that unifies all four forces of nature.

Are there naked singularities? One of the assumptions of the uniqueness theorem is that a horizon exists in the spacetime. This arises from a belief embodied by the "Cosmic Censorship Hypothesis" [114] (CCH), which states that any singularity will be enclosed by a horizon. The CCH arises from a desire for predictability in the Universe — when Physics breaks down at a singularity, we do not want information from that to propagate into the rest of the Universe. However, the CCH is unproven and therefore "naked" singularities not enclosed within a horizon may still exist. Gravitational wave observations provide a unique way to look for these exotic objects. Observations may be indirect, via detection of a violation to the "no-hair" theorem. However, they may also be direct — if a horizon is not present in the spacetime, the gravitational waves will not cut-off when the object crosses the horizon [115], which will be a clear smoking gun signature for the absence of a clothing horizon in the system.

The Einstein Telescope will provide much more stringent constraints on potential violations of the CCH than are possible with Advanced LIGO. ET observations will therefore play an important role in answering the question as to whether naked singularities exist, which could have profound implications for our understanding of various aspects of the theory of relativity.

2.3.7 Limit on the maximum mass of "neutron stars"

It is generally believed that neutron stars have masses between $\sim 1.3\,M_\odot$ and $\sim 2\,M_\odot$, but such statements rely on guesses regarding the equation of state of dense nuclear matter. Above $2\,M_\odot$ a quark star might be created, or some other exotic object. Apart from the existence of such objects and their properties, an interesting question is how massive a star can be while still being stable. The left panel of Fig. 19 shows the maximum distance to which NS-BH binary inspirals can be seen in initial and advanced LIGO and Virgo, and in ET. The latter would have access to sources out to genuinely cosmological distances, up to redshifts of several, with an expected detection rate in the order of $10^6\,\mathrm{yr}^{-1}$. The right panel shows how accurately one would be able to measure the mass of a neutron star in an NS-BH inspiral process, as a function of the mass of the black hole as well as redshift. For a black hole mass $\gtrsim 5\,M_\odot$, the neutron star's mass can be inferred to a fraction of a percent out to $z = 1$, and up to a few percent out to a redshift of 2 or 3. This would enable us not only to establish the mass distribution of neutron stars and dense exotic objects, but also the evolution of this distribution over cosmological timescales.

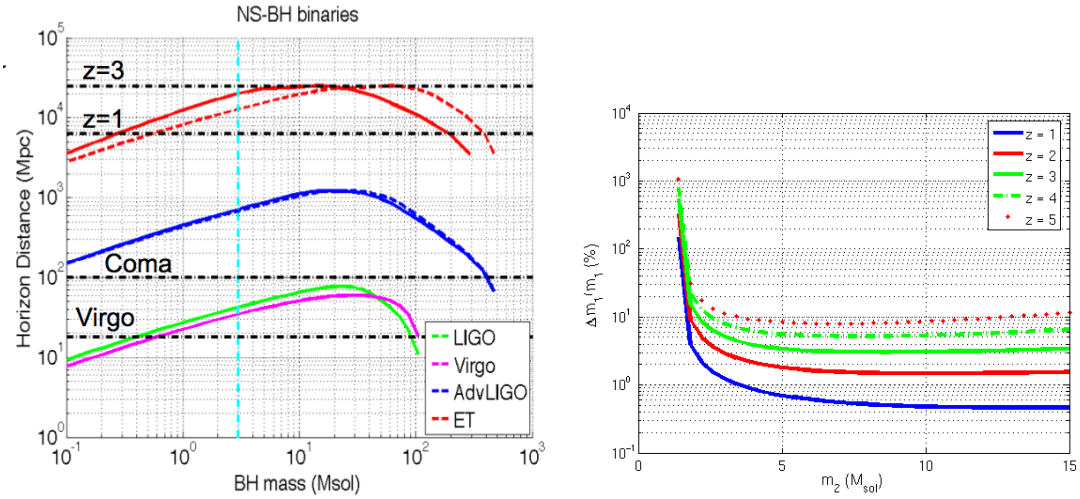


Figure 19: The plot on the left shows the horizon distance for neutron star-black hole binaries against black hole mass for initial and advanced LIGO, Virgo and ET. The solid coloured curves show the horizon distance against the *instrinsic* (non-redshifted) mass of the source. The dashed coloured curves show the horizon distance against the *observed* (redshifted by a factor of $1 + z$) mass of the source. The plot on the right shows the accuracy with which the mass of a neutron star can be determined from the inspiral signal, as a function of the mass of the companion object.

2.4 Astrophysics

2.4.1 Equation-of-state of neutron stars from binary coalescence

Binary neutron stars are known to exist and for some of the systems in the Galaxy general-relativistic effects in the binary orbit have been measured to high precision. The inspiral and merger of two neutron stars in binary orbit is the inevitable fate of close-binary evolution, whose main dissipation mechanism is the emission of gravitational waves. The detection of gravitational waves from neutron stars binaries will provide a wide variety of physical information on the component stars, including their mass, spin, radius and equation of state (EOS). The central densities of isolated neutron stars, in fact, can range up to ten times nuclear saturation density, and during the merger and coalescence of two neutron stars the maximum density will rise even further before the remnant object collapses to a black hole. The behaviour of bulk matter at these densities is not well understood, and measurements of gravitational wave signals from neutron star sources can usefully constrain the EOS at the these densities.

Quantum chromodynamics is expected to be a complete description of matter at these energies; the uncertainty in theoretical understanding comes from the many-body problem with strong interactions. The description of bulk neutral matter in terms of hadrons such as protons and neutrons may need to be expanded to accomodate new particles that are formed at these energies, such as hyperons, pions, and kaons. In fact the appropriate degrees of freedom describing cold matter at very high density may no longer be hadrons but the quarks and gluons themselves, in some form of quark matter.

While isolated or inspiralling neutron stars are well described by the ground state of matter, i.e. with a "cold" EOS, the temperatures reached in the coalescence as a result of the strong shocks, will be significant and of the order of $\sim 10^{10} - 10^{12}$ K. Yet, just as measurements of the hot out-of-equilibrium ion collisons in RHIC constrain the ground state of dense nuclear matter, characteristics of the collisions neutron stars may be able to constrain the ground state of dense neutral matter.

Reviews of the current range of candidate equations of state and their constraint using astrophysical and heavy ion collision experiments can be found in [116–118].

The signature of the neutron star EOS can be found in almost any neutron-star sourced gravitational wave; in the peak frequencies of supernova waveforms [119, 120], in the possibility of accretion-induced crust mountains [121, 122], and in the astroseismology of glitches and other oscillation mode excitations. Studies which have specifically explored the effect of varying EOS (or varying compactness for a given mass, which implies EOS variation) on gravitational wave spectra include [123–127] for binary neutron star inspiral, [128–135] for binary neutron star coalescence, and [116, 136–138] for mixed binaries. The gravitational waves from binary inspiral and merger are expected to be frequantly measured, and the predicted signals have several interesting EOS-dependent features.

This is illustrated as a representative example in Figure 20 and which reports the simulations of [134]. In particular, the left panel on the top row shows the comparison in retarded-time evolution of the real part of the $\ell = m = 2$ component of the gravitational-wave signal $r\Psi_4$ for a high-mass binary (i.e. with a total mass of $3.2\,M_\odot$) when evolved with the a "cold" EOS (i.e. a polytropic one) or with a "hot" EOS (i.e. the ideal-fluid one). Similarly, the right panel on the top row shows the same comparison but for a low-mass binary (i.e. with a total mass of $2.9\,M_\odot$). Finally, the bottom row offers a representation of the same data but in frequency space. While none of the two EOSs considered here is realistic, they span in some sense the extrems of the range of possibilities. Most importantly they show that the gravitational-wave signal will be very sensitive on the mass of the stars and on the EOS.

While this is especially true in the post-merger phase, also the inspiral phase will provide important information on the EOS. Indeed, although for most of the inspiral of a binary neutron star system the stars are well-modeled as point particles, as they approach each other, an EOS dependent tidal deformation modifies their orbits, changing the late inspiral waveform. The measurability of this effect in gravitational wave detectors can be estimated using both post-Newtonian tidal deformation calculations and full numerical simulations of binary neutron stars with varying EOS.

The set of numerical simulations of [127] using an ET proposed noise curve give, for a $1.35M_\odot$–$1.35M_\odot$ double neutron star binary at 100 Mpc in ET, estimates of the measurement uncertainty δR of isolated neutron star are of the order ± 0.5–$1.0\,\mathrm{km}$. This compares favorably to the range in predicted radius of roughly 9–16 km. Parameterizing the variation by the pressure at $5 \times 10^{14}\,\mathrm{g\,cm^{-3}}$, as in [126], gives estimates of δp_1, where p_1 is $\log(p/c^2)$, of order ± 0.05–0.10, compared to a range of roughly 13.0–13.8 for realistic EOS.

Estimates of a post-Newtonian tidal deformation following [125] give for a binary at 50 Mpc

- $\Delta\tilde{\lambda} \sim 1.22 \times 10^{36}$ for $1.35 - 1.35 M_\odot$
- $\Delta\tilde{\lambda} \sim 1.6 \times 10^{36}$ for $1.45 - 1.45 M_\odot$
- $\Delta\tilde{\lambda} \sim 1.85 \times 10^{36}$ for $1.35 - 1.7 M_\odot$

for inspiral *below 400 Hz*, where $\tilde{\lambda}(m_1, m_2)$ is a measure of tidal deformability in a given binary which ranges between 0.5×10^{36} g^2 cm^2 and 10×10^{36} g^2 cm^2 for realistic EOS.

The advantage of ET from the perspective of EOS understanding is not necessarily the larger number of detections possible with increased sensitivity, although information about mass distribuition of neutron star populations can also be useful for EOS constraint. Instead, ET will provide very strong signals at reasonable rates; for example two $1.4\,M_\odot$ neutron stars inspiralling towards each other within an effective distance of 100 Mpc, which is expected roughly once a year, would give a SNR in ET of over 900. This makes possible for the precise measurement of masses in early inspiral, the detection of small departures from point particle behaviour at moderate frequencies, and discrimination between merger and post-merger signals from different models at high frequencies.

An interesting feature that has emerged from studies of binary neutron star coalescences is the post-merger formation, in some cases, of a hyper-massive remnant object which oscillates and emits gravitational waveforms on fairly long timescales. The presence or absence of such post merger oscillations, as well as their characteristic frequency and duration, varies with the cold EOS. However, they are additionally sensitive to many physical effects from thermal properties, magnetic fields, and particle production, and so forth. The precise details of the signal, similarly to those from supernovae, are not easy to predict. However, the signal from such a post-merger

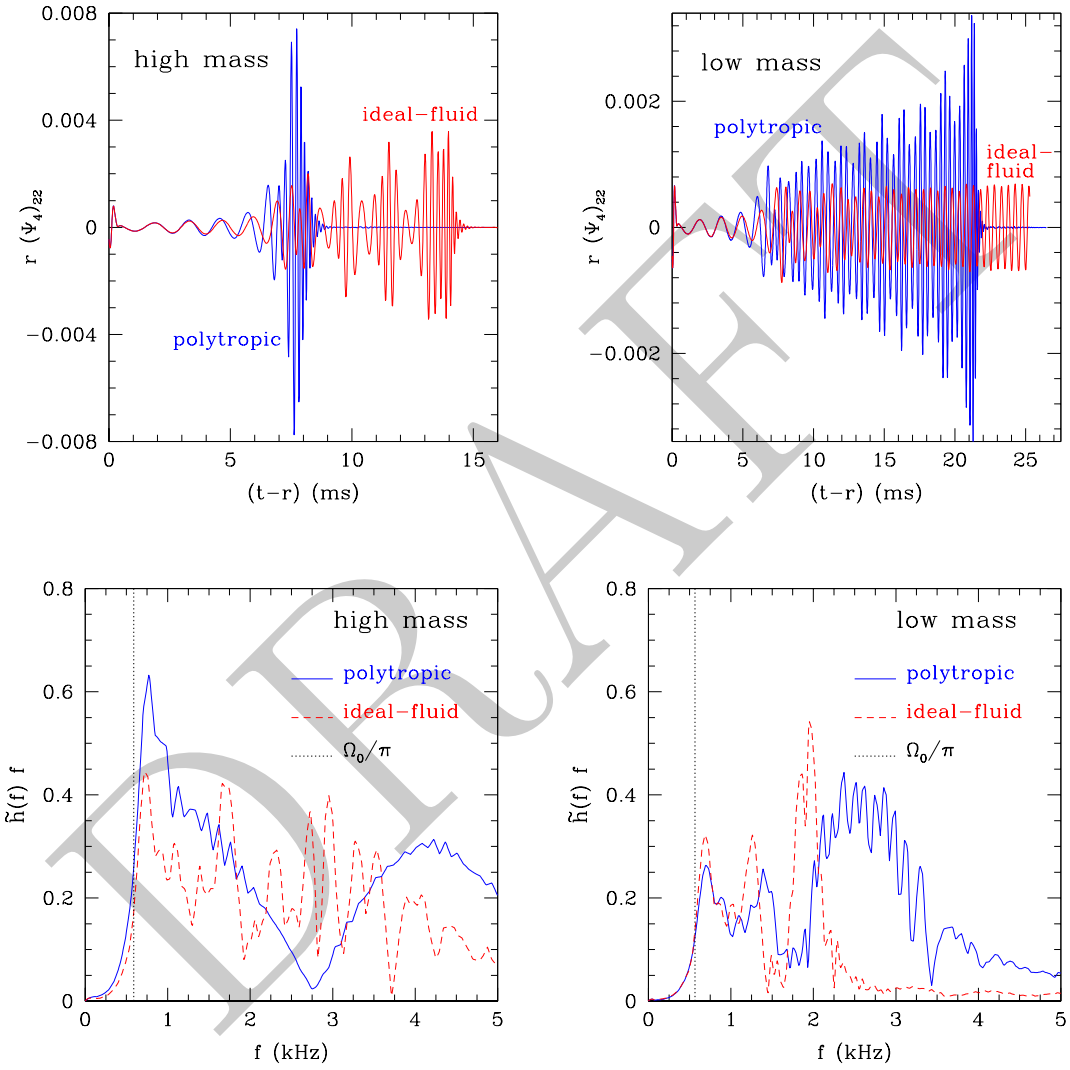


Figure 20: *Top row, left panel:* Comparison in retarded-time evolution of the real part of the $\ell = m = 2$ component of $r\Psi_4$ for a high-mass binary (i.e. with a total mass of $3.2\,M_\odot$) when evolved with the a "cold" EOS (i.e. a polytropic one) or with a "hot" EOS (i.e. the ideal-fluid one). *Top row, right panel:* The same as in the left panel but for a low-mass binary (i.e. with a total mass of $2.9\,M_\odot$). *Bottom row:* The same as in the top row but with the comparison being made in frequency space. Indicated with a vertical long-dashed line is twice the initial orbital frequency.

oscillation is potentially visible with advanced detectors [130], and analysis of the signal following a measured inspiral may provide useful constraints on the underlying astrophysics.

It has recently become possible to compute the first complete and accurate simulations of the merger of a neutron stars binary through to the delayed formation of a BH and to its ringdown [134, 135]. By computing the complete gravitational wave signal produced in the process it was possible to show that the gravitational waves are strongly correlated to the properties of the sources emitting them. Differences in the EOS or in the initial mass of the system produced different signals with different power spectra and different durations.

Furthermore, magnetic fields (MFs) are commonly present in neutron stars and their possible impact on the dynamics of binary neutron stars has only begun to be examined. The `Whisky` code has been recently used to investigate the effect that MFs have on the gravitational wave emission produced during the inspiral and merger of magnetized neutron stars [139]. In particular it has been shown that MFs do have an impact after the merger (for initial MFs $B_0 \gtrsim 10^{12}$ G), but also during the inspiral (for sufficiently strong initial MFs with $B_0 \gtrsim 10^{16}$ G). These results, are quantified by computing the overlap between the waveforms produced during the inspiral by magnetized and unmagnetized binaries. Moreover, through the inclusion of more realistic equations of state and of a radiation transport scheme, it will be possible to increase considerably our level of understanding of these objects.

2.4.2 Equation-of-state of neutron stars from pulsar glitches

Many radio pulsars exhibit glitches, sudden spin-up events followed by a relaxation period towards stable secular spin-down. Pulsar glitches have a long observational history (beginning shortly after the discovery of the first pulsar) and so far over a hundred pulsars are known to have glitched at least once. Glitches have also been observed in magnetars. The archetypal glitching pulsar is Vela, which exhibits regular large glitches with an amplitude corresponding to a fractional spin frequency jump of the order of 10^{-6}.

Despite the wealth of observational data, the phenomenon of glitches remains an enigma from the theoretical point of view. It is widely believed that glitches are related with the existence of supefluids in the interior of mature neutron stars and that they involve a transfer of angular momentum from a superfluid component to the rest of the star, which includes the crust (to which the pulsar mechanism is presumed to be rigidly attached) and the charged matter in the core. A superfluid rotates by forming an array of quantised vortices and it can spin down provided the vortices can move outwards. If the vortex migration is impeded by 'pinning' to the other component then the superfluid cannot keep up with the spin-down due to electromagnetic braking. As a result a rotational lag develops between the two components until some critical level is reached at which the vortices unpin and transfer angular momentum to the rest of the star, and the two components are driven to corotation.

The nature of the instability causing vortex unpinning and the subsequent stage of relaxation of the system is poorly understood. One might hope that gravitational radiation detected by a glitch event could help unveil the key physics associated with this enigmatic phenomenon. It is likely that a glitch event involves the excitation of some of the inertial modes of the two-component system and the post-glitch relaxation is governed by the coupling of the two components through the vortex-mediated mutual friction force and the magnetic field. In fact, as recent work suggests, the glitch trigger-mechanism may be the result of a superfluid 'two-stream' instability setting in through the inertial modes of the system.

An instrument like ET would be the ideal tool for detecting a gravitational wave signal in the 10-100 Hz band which is the relevant one for the inertial modes of a Vela-like pulsar. The detection of gravitational wave signals from glitching pulsars would provide a tool for probing the interior matter of neutron stars and supplement the existing and future electromagnetic observations. The realisation of this exciting prospect will require (as in the case of other potential sources of gravitational radiation) the input of theoretical waveform templates. These waveforms should be computed using detailed multifluid hydrodynamical models for superfluid neutron stars, accounting for effects like vortex mutual friction and pinning.

2.4.3 Understanding relativistic instabilities in neutron stars via observation of r-modes

Just after the catastrophic implosion of a core-collapse supernova, gravitational waves could be emitted by the newborn, hot, rapidly rotating neutron star. In particular, this may occur due to the excitation of *r-modes* [63], non-radial pulsation modes of rotating stars for which the Coriolis force acts as the restoring force. The characteristic frequency of the GW is comparable to the rotation speed of the star. r-modes have an instability driven by gravitational radiation [140], inducing differential rotation at second order in the amplitude of the modes. The second order perturbation plays an important role in the nonlinear evolution of the r-mode instability, which makes detection of gravitational waves more difficult.

GW detection and measurement depends strongly on the r-mode saturation amplitude, which is estimated as much lower than unity by recent simulations and theory. Thus detection of such gravitational waves is more difficult than initially supposed. In this section we consider both the r-modes and another nonlinear effect, the differential rotation induced by r-modes [141]. The amount of such differential rotation is described by a parameter K, which may varying in the range $[-5/4, 10^{13}]$ depending on the model and initial conditions.

Such signals are an important opportunity for ET science, as they could provide a deep probe of aspects of neutron star formation and nuclear physics.

The strength of gravitational waves from r-modes The detectability of GW produced by r-modes depends on the amount of angular momentum that they carry away. As described in [142], for $K = 0$, the total angular momentum of the star decreases to 65% of its initial value, and part of the initial angular momentum of the star, about 58%, is transferred to the r-mode as a consequence of the rapid increase of the average differential rotation. Therefore the initial angular momentum carried away by gravitational waves is about 35%. This result is strongly dependent on the value of K: for higher K the amount of angular momentum carried away by gravitational waves may even fall below 1%.

From the model in [142] the frequency f of the gravitational wave depends on the star angular velocity Ω by $f = 2\Omega/(3\pi)$. The frequency range is estimated as follows:

- $f_{min} \simeq [77 - 80]\,\mathrm{Hz}$, depending on the final value of the angular velocity $\Omega(t_f)$ and K;
- $f_{max} \simeq 1200\,\mathrm{Hz}$, depending on the initial value of the angular velocity Ω_0.

The amplitude in the frequency domain is given by:

$$H(f) = \frac{4.6 \times 10^{-25}}{\sqrt{2+K}} \sqrt{\frac{f_{max}}{f}} \frac{20\,\mathrm{Mpc}}{D}\,\mathrm{Hz}^{-1}$$

where D is the source distance, f the GW signal frequency and $f_{max} = 1191\,\mathrm{Hz}$ is its maximum frequency.

We may estimate the signal to noise ratio (SNR) at ET by adapting a calculation made for Advanced LIGO in [142].[2] The optimal SNR is given by

$$\frac{S}{N} = \frac{250}{\sqrt{(2+K)}} \frac{20\,\mathrm{Mpc}}{D}.$$

The strong dependence on the unknown parameter K is clear, as shown in Figure 21 where we consider an SNR of 20, arbitrarily chosen for a confident observation of the signal. It is possible to conclude that the range of distances on which gravitational signal could be visible is really large. Considering the optimistic case when $K = 0$, we obtain a sight distance for an optimally oriented source of 175 Mpc; while considering the pessimistic case when $K = [10^5 - 10^6]$, the sight distance falls down to less than 1 Mpc (galactic sources).

[2] The sensitivity curve used for our estimation is ET-B.

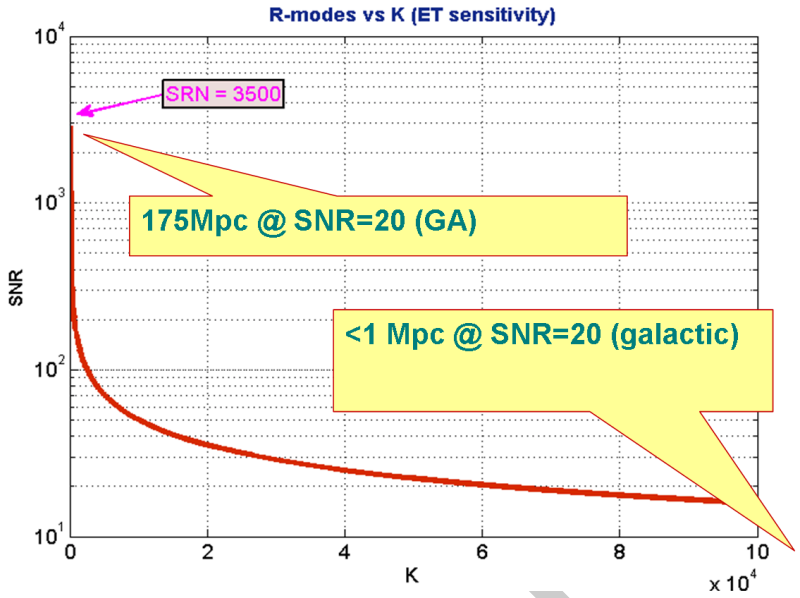


Figure 21: R-modes expected in Einstein Telescope.

R-modes and ET science goals A significant motivation for studying gravitational waves from the r-mode instability at ET is the opportunity to obtain a unique correlation with the nuclear physics and formation processes of neutron stars. Such a signal could yield a fundamental probe of neutron star dynamics. Some possible implications are given below:

- In principle from gravitational wave signals and signal models it is possible to trace and quantify the initial conditions of the new born star such as initial frequency, initial temperature and others. This could lead to confirm or exclude a set of supernova models, star formation processes and NS models that agree with these experimental data.
- Another implication concerns neutron star nuclear physics models and the equation of state (EOS), potentially opening a window on the core and crust physics, especially superfluid aspects.
- The phenomenon of neutron star cooling due to neutron emission also interacts with the r-mode instability. In this case gravitational waves could provide information about the cooling rate and cooling model, which at the moment is assumed to follow the modified URCA process.

Hence, a possible observation of this signal by Einstein Telescope can put constraints on a set of theories concerning neutron star formation processes, nuclear physics models and the EOS.

Gravitational waves generated by the r-mode instability are a subject of considerable recent interest. We use here the results of considering the r-mode with $l = 2$ [142] , but more generally there are efforts to better understand the full dynamics of the newborn star considering higher numbers and different types of modes [143]. It is important to investigate fully such models and simulations as a science opportunity for ET.

2.4.4 Solving the enigma of GRB progenitors

Gamma-ray bursts (GRBs) are the most luminous explosions in the EM spectrum occurring in the universe. Through observations made by satellite-based gamma-ray observatories it was found that the the duration of the GRBs follows a bimodal distribution [144]. We classify GRBs either as *short-hard* or *long-soft* bursts depending on their duration and spectra. Through follow-up observations of the x-ray, optical and radio afterglow emission of GRBs it is possible to determine their sky-location, redshift and host galaxy.

Long GRBs are always associated with late-type star-forming host galaxies [145]. A handful of long GRBs have also been associated with supernovae [146–149]. It is therefore thought that core-collapse supernovae are the progenitor of long GRBs [150, 151].

Short GRBs are observed at lower redshifts than long GRBs are associated with a variety of galaxy types including early-type elliptical and lenticular galaxies without active star forming regions [152]. Currently, it is widely thought that merger of neutron star binaries or neutron star-black hole binaries (NS-BH) are the progenitors of most short-hard GRBs [153]. Some small fraction of short GRBs (less than 15% of known short GRBs) may be caused by soft gamma repeater flares (SGRs) [154, 155]. SGRs are described further in Section 2.4.4. Accurate predictions for the gravitational wave emission of the inspiral, merger and ringdown of compact binaries are possible through post-Newtonian approximations to Einstein's equations or through numerical relativity simulations. Searches for gravitational waves from inspiralling binaries using matched-filtering on data from initial interferometers (LIGO, Virgo, GEO) are underway. A search for the emission of GRB 070201 whose sky-locatiion error box overlaps the spiral arms of M31 was carried out in Ref. [156]. The matched-filter analysis excluded an inspiral progenitor for GRB 070201 if it was indeed located in M31 with a confidences of 99%.

It is typical to characterise the sensitivity of a gravitational wave observatory to inspiral distances by the horizon distance. The horizon distance is the distance at which we would measure a matched-filter signal-to-noise ratio of 8 for an optimally oriented (i.e., face-on) and overhead source. Figure 19, left plot, shows the horizon distance achieved by initial and advanced LIGO, Virgo and ET for NS-BH binaries with $m_{\rm NS} = 1.4 M_\odot$ and a range of BH masses.

Predicting the gravitational wave emission of core-collapse supernovae associated with long GRBs is more difficult and involves modelling the complicated internal dynamics of the collapsing star, see e.g., [78]. Searches for unmodelled gravitational emission from GRBs on data from initial interferometers (LIGO, Virgo, GEO) are underway [156, 157].

From gravitational wave searches using coherent analysis techniques [157] we find that in general the 90%-confidence upper limit on $h_{\rm rss}$ gravitational wave amplitude is around an order of magnitude above the amplitude spectrum of the interferometer, i.e., $\sim 10 \times S_h(f)^{0.5}$. For narrow-band burst signals we can use the following approximation

$$E_{\rm GW}^{\rm iso} \simeq \frac{\pi^2 c^3}{G} D^2 f_0^2 h_{rss}^2 \tag{30}$$

where $E_{\rm GW}^{\rm iso}$ is the isotropic energy emission in gravitational waves, D is the distance of the source, f_0 is the central frequency, and h_{rss} is the root-sum-square amplitude of the gravitational wave:

$$h_{\rm rss} = \sqrt{\int \left(|h_+(t)|^2 + |h_\times(t)|^2\right)\, dt}\ . \tag{31}$$

From Eqn. 31 we can calculate a lower limit on source distance from our amplitude upper limit for a given assumption of $E_{\rm GW}^{\rm iso}$. For long GRBs the energy of emission in gravitational waves is not well known but has been estimated to be as high as $0.2 M_\odot c^2$ in the LIGO-Virgo frequency band of good sensitivity [158]. In Fig. 22 we estimate the distance to which various detectors are sensitive to a narrow-band burst of gravitational waves assuming $E_{\rm GW}^{\rm iso} = 0.05 M_\odot c^2$.

Soft Gamma Repeater Flares As described in section 2.4.4, a significant fraction, up to 15%, of short, hard γ-ray bursts may be associated with flaring activity in soft γ-repeaters (SGRs). These sources often undergo sporadic periods of activity which last from days to months where they emit short bursts of hard X-rays and soft γ-rays with luminosities $L \sim 10^{41}\,{\rm erg\,s^{-1}}$ and photon energies in the range 10-30 keV. Much more occassionally, they exhibit enormous, giant flares with luminosities as large as $10^{47}\,{\rm erg\,s^{-1}}$. There exist 4 known soft γ-repeaters, 3 in the milky way and 1 in the Large Magellanic cloud. It is generally believed that SGRs belong to a class of neutron star, magnetars, with extraordinarily large magnetic fields in the the range 10^{14}-10^{15} G where the flaring activity is due to sudden, violent reconfigurations of complex magnetic field topologies.

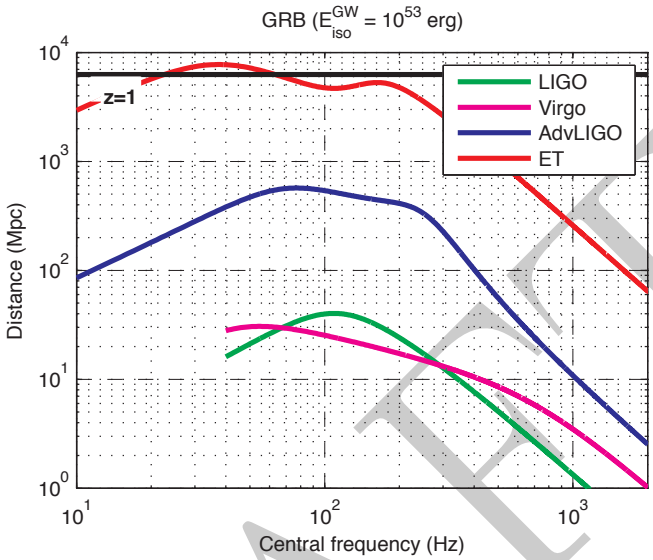


Figure 22: 90%-confidence lower limit on distance for GRB burst sources assuming a GRB energy emission of $E_{\rm GW}^{\rm iso} = 0.05 M_\odot c^2 \sim 9 \times 10^{52}$ergs. The solid horizontal black line near the top of the figure shows a redshift $z = 1$.

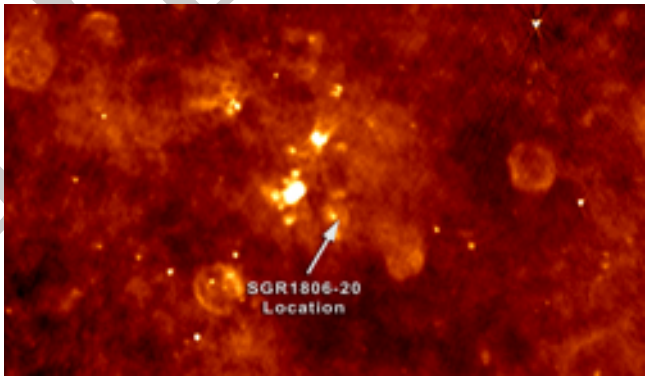


Figure 23: A high resolution, wide-field image of the area around SGR1806-20 as seen in radio wavelength. SGR1806-20 can not be seen in this image generated from earlier radio data taken when SGR1806-20 was ˘201cradio quiet˘201d. The arrow locates the position of SGR1806-20 within the image. Credit: University of Hawaii.

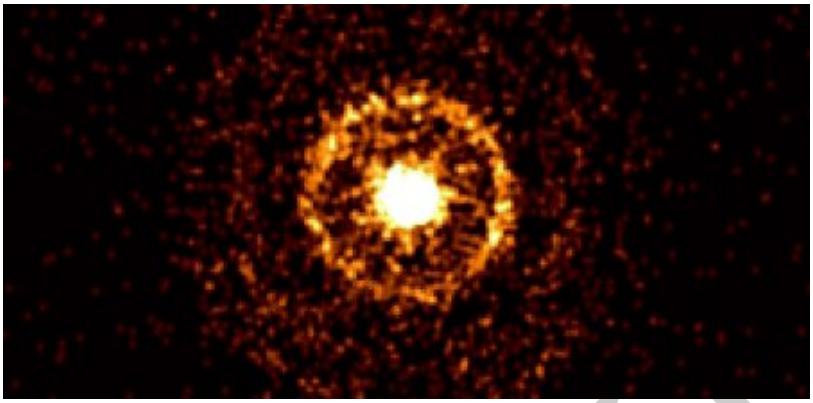


Figure 24: Swift's X-Ray Telescope (XRT) captured an apparent expanding halo around the flaring neutron star SGR J1550-5418. The halo formed as X-rays from the brightest flares scattered off of intervening dust clouds. Credit: NASA/Swift/Jules Halpern, Columbia University

The hardness of their spectra and the enormous luminosities involved mean that giant flares from nearby SGRs, such as that of SGR 1806-20 [159], represent an intriguing candidate progenitor scenario for some short duration γ-ray bursts. Indeed, in [160], the authors report a correlation between the positions of some short GRBs with those of low redshift galaxies, suggesting that $10-25\%$ of short GRBs occur in the local universe and, therefore, are likely to be associated with giant SGR flares. Furthermore, evidence for the existence of two classes of progenitors for short GRBs is provided in [155]. Here, it is found that a bimodal luminosity function, representing a dual-population of short GRB progenitors with low and high luminosities, is required to reproduce the observed distributions of short GRB luminosities. As well as statistical evidence, there have been observations of at least three individual short GRBs which present candidates for extragalactic SGR flares. Optical and infrared observations [161] of GRB 050906 suggest a tentative association with the local, fairly massive ($M \sim 10^{11} M_{\odot}$) starburst galaxy IC328 which lies at a redshift of $z = 0.031$. If GRB 050906 had indeed originated in IC328, the isotropic equivalent energy would be $E_{\rm ISO} \sim 1.5 \times 10^{46}$ erg in the 15-150 keV range. The giant flare from SGR 1806-20, by comparision, emitted $E_{\rm ISO} \sim 4 \times 10^{46}$ erg with photon energies > 30 keV. As well as the potential similarity in the energetics of this burst, the association with a starburst galaxy, where young, shortly lived magnetars are believed to be most prevalent, corroborates the SGR progenitor scenario. Two other short GRB-SGR flare candidates, GRB 051103 and GRB 070201, were detected by the Konus-Wind GRB spectrometer [162, 163]. The localisation area of GRB 051103 was found to lie near M81 (D=3.6 Mpc), suggesting an isotropic equivalent energy $E_{\rm ISO} = 7 \times 10^{46}$ erg. As remarked in [162], if GRB 051103 was *not* related to an SGR flare, we would expect an optical and/or transient in the localisation area, which has not been observed. Finally, the localisation area of GRB 070201 was found to overlap with the spiral arms of M31 (D=0.78 Mpc), leading to an estimate of $E_{\rm ISO} \sim 1.5 \times 10^{45}$ erg under the SGR flare scenario and, again, comparable to the giant flare from SGR 1806-20.

In addition to these types of arguments related to the energetics of the electromagnetic emission, gravitational wave observations can provide an extremely powerful tool to identify SGRs as sGRB progenitors. First, we note that the failure to detect the signature of a compact binary coalescence from sGRBs at distances where such a signal is expected can provide compelling evidence for the SGR progenitor scenario alone. Indeed, observations by the initial LIGO instruements recently excluded the coalescence of a binary neutron star system within M31 at more than 99% confidence as the progenitor for GRB 070201 [156]. Furthermore, a binary neutron star merger is excluded at distances less than 3.5 Mpc with 90% confidence. If, however, the progenitor had been an SGR flare, the LIGO observations imply an upper bound on the isotropic energy released as an unmodeled gravitational wave burst of $E_{\rm ISO}^{\rm GW} < 7.5 \times 10^{50}$ erg, within the bounds permitted by existing models.

The non-detection of an expected inspiral gravitational wave signature, however, is not the only way that an instrument like the Einstein telescope can provide evidence for the SGR progenitor scenario. In section **??**, we discuss giant SGR flares as a source of quasi-periodic oscillations, with quadrupolar components in the

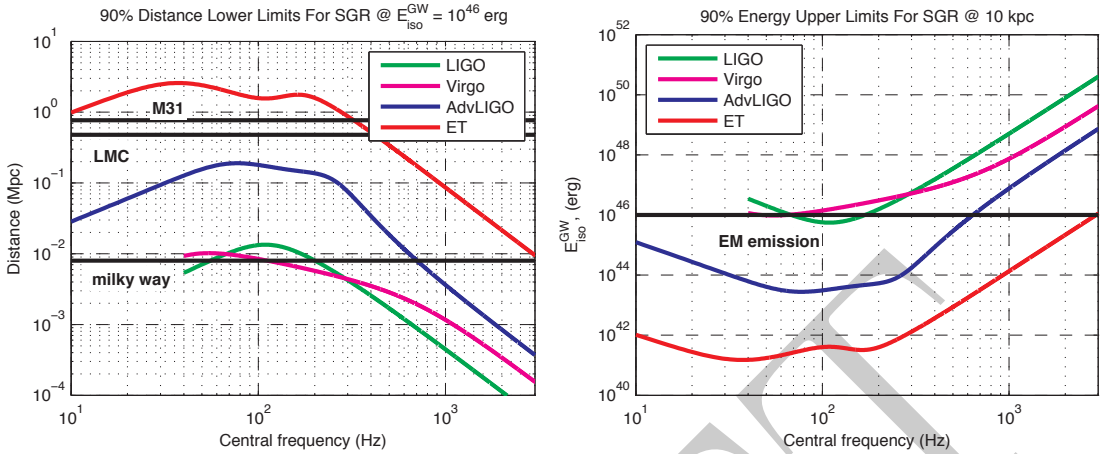


Figure 25: *Left panel*: 90%-confidence lower limit on distance for burst sources assuming $E_{\rm GW}^{\rm iso} = 10^{46}$ergs for an SGR progenitor scenario. Starting from the lower edge of the figure, the solid horizontal black lines show the distances to center of our galaxy, the distance to the large Magellanic cloud and the distance to M31 in Andromeda. *Right panel*: predicted 90% upper limits on isotropically emitted gravitational wave energy from a galactic SGR flare (i.e., distance of 10 kpc). The solid black horizontal line shows the expected upper limit of 10^{46} erg from energetic arguments alone.

$\sim 10-40$ Hz range. Observations of these shear mode oscillations in gravitational waves, with no accompanying inspiral signal, would only be explicable under the SGR scenario. It is also possible that non-radial oscillatory modes would become excited by tectonic activity associated with a giant SGR flare [164]. These modes will then be damped by gravitational wave emission, resulting in a characteristic ring-down signal [165]. Various families of oscillatory modes, such as fluid (f), pressure (p) and purely space-time (w) modes may be excited and simultaneous gravitational wave observations of all three of these families can be used to place tight constraints on the neutron star equation of state [166]. The p and w modes, however, tend to have frequencies well above 4 kHz making the f-mode, with frequencies expected in the range $1-3$ kHz [167], the most accessible to currently planned gravitational wave observations. Again, gravitational wave observations of f-mode ring-downs associated with sGRBs, where there is no accompanying inspiral signal, would point directly to an SGR giant flare as the progenitor.

Current models for SGRs [121, 164, 168] indicate that they will emit less than 10^{46}ergs in gravitational waves. In the left panel of Fig. 25 we show 90%-confidence lower limits on the distances to which various gravitational wave detectors will be sensitive to gravitational wave bursts with this energy and we see that, in their most sensitive frequencies, the current generation of interferometers are just able to probe our own galaxy. While advanced LIGO improves this reach substantially, it is really only with the Einstein telescope that observations of gravitational waves associated with extra-galactic SGR flares become possible. Figure 25 shows the complementary plot of 90% energy upper limits obtainable by the various instruments for a galactic SGR. We typically take this to mean a distance of ~ 10 kpc. Again, it is only with an instrument like the Einstein telescope that we are able to probe interesting energy regimes across the entire frequency spectrum one might reasonably expect for gravitational wave emission associated with SGR flares.

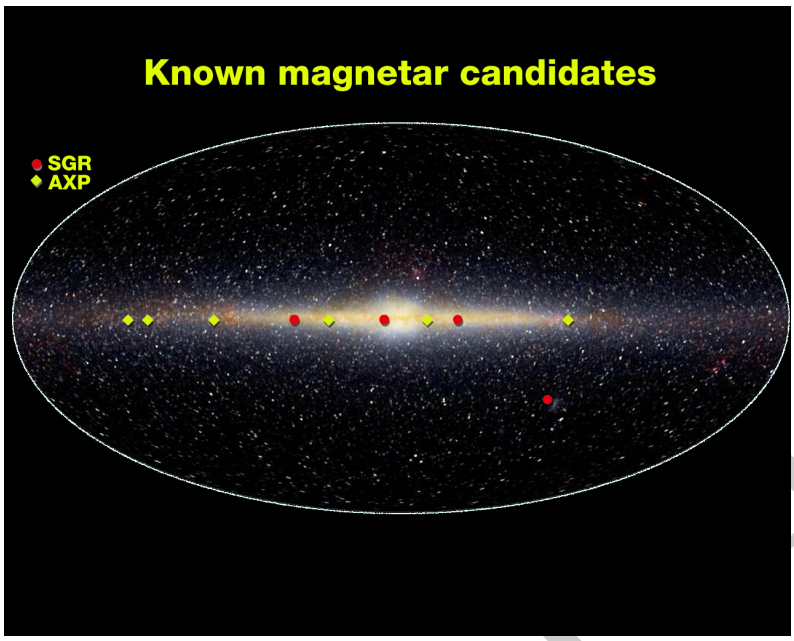


Figure 26: The locations of known magnetar candidates (Soft Gamma-ray Repeaters and Anomalous X-ray Pulsars) in the Milky Way. Credit: NASA/Marshall Space Flight Center

2.4.5 Understanding supernova cores

Stellar collapse is the most energetic event in the Universe, releasing $\sim 10^{53}$ erg of gravitational energy in the compression of a massive star's iron core to a neutron star. Most of this energy ($\sim 99\%$) is emitted in neutrinos and only about 10^{51} erg go into energy of the core-collapse supernova (CC-SN) explosion. CC-SNe (SN types II, Ib, Ic) are $\sim$10 times more frequent than thermonuclear type-Ia SNe. A SN explosion pollutes the interstellar medium with the nucleosynthetic products of stellar evolution (CC-SNe are the Universe's primary source of oxygen) and enriches via the r-process the universe with rare heavy isotopes. The perturbation caused by an SN in its vicinity can trigger the formation of stellar systems and stellar collapse and CC-SNe are the birth sites of neutron stars (NSs) and stellar-mass black holes (BHs).

The Supernova Problem and GW observations The precise mechanism of explosion operating in CC-SNe is unknown [76–78]. When the inner part of the collapsing iron core reaches densities close to those in atomic nuclei, the strong force leads to a stiffening of the nuclear equation of state (EOS), resulting in *core bounce* of the inner core into the still infalling outer core. A shock wave is formed that propagates outward in mass and radius, but quickly loses energy due to the breakup of heavy nuclei and neutrinos that carry away energy from the postshock layer. The shock stalls, turns into an accretion shock and must be *revived* to drive a CC-SN explosion. If this does not happen, a BH will form on an accretion timescale of $\sim 2\,$s. *What is the mechanism of shock revival?* This is the fundamental question and primary unsolved problem of CC-SN theory. Indications are strong that the CC-SN mechanism involves a multitude of multi-dimensional processes, including rotation, convection/turbulence, and various hydrodynamic instabilities of the stalled shock and in the proto-NS. This opens up the possibility of probing the supernova mechanism with gravitational waves (GWs). GWs, even more so than neutrinos, carry direct dynamical information from the supernova engine deep inside a dying massive star, a region generally inaccessible by the traditional means of observational astronomy. GWs form a core-collapse event have the potential of putting very strong constraints on the CC-SN mechanism [78, 79]. With initial and certainly second-generation interferometric GW detectors, this should be possible for an event in the Milky Way ($D \sim 10 - 15\,$kpc) and the Magellanic Clouds [78] ($D \sim 50 - 70\,$kpc), but even optimistic estimates of the CC-SN rate in this region do not predict more than $\sim 1 - 2$ events per century. This number roughly doubles if one includes the entire local group ($D \sim 1\,$Mpc). In the region from $3 - 5\,$Mpc a number of starburst

galaxies increase the predicted and observed integrate SN rate to $\sim 0.5\,\mathrm{yr}^{-1}$. At $D \sim 10\,\mathrm{Mpc}$ it is $\gtrsim 1\,\mathrm{yr}^{-1}$.

Supernova Science with ET The GW emission processes in a CC event emit GW strains h in the range $10^{-24} - 10^{-22}\,(D/1\,\mathrm{Mpc})$ and most of the emission takes place at frequencies of $\sim 200 - 1000\,\mathrm{Hz}$, but the various explosion scenarios exhibit unique spectral distributions and vary in total emitted energies [78, 79]. In addition, there is likely to be a low-frequency GW-memory-type component with large h up to $10^{-22}\,(D/1\,\mathrm{Mpc})$ at $0-20\,\mathrm{Hz}$. ET as currently envisioned [81] is sufficiently sensitive to detect GWs from various CC-SN scenarios out to $2-4\,\mathrm{Mpc}$. If the high-f sensitivity was increased by a factor of $\sim 2-3$, detection out to $\sim 10\,\mathrm{Mpc}$ may be possible. Figure 14 summarizes the ET observational capabilities and examines each of the main generation processes of gravitational waves.

Even without this improvement, ET may see multiple CC-SNe during its lifetime and would have the power to provide strong hints for a particular SN mechanism and/or smoking-gun evidence against another – crucial astrophysics information that is unlikely to be attainable in other ways. At ET's implementation, megaton-class neutrino detectors will be operative and, having range similar to ET, will be able to provide coincident observations, narrowing down the time of the GW emission to $\sim 1\,\mathrm{ms}$. In addition, deep high-cadence optical transient surveys will be operative and targeting near-universe transients, providing additional coincident data as well as additional astrophysics output (progenitor type/mass, explosion morphology/energy etc.).

Impact Constraining the CC-SN mechanism will mean a breakthrough in our understanding of the large range of phenomena associated with stellar collapse, CC-SNe, BH and NS formation, and gamma-ray bursts (GRBs). However, the astrophysics and physics information provided by GWs observed from a CC event with ET goes beyond this: These GWs carry also information on the high-density nuclear EOS, explosion asymmetries and pulsar kicks, the formation of a BH in a failing CC-SN, and can help uncover rare events such as the accretion-induced collapse of a white dwarf to a NS or weak or failing CC-SNe that have very weak or absent EM signatures.

2.4.6 Explaining neutron star spin frequencies in low-mass X-ray binaries

Observations of accreting neutron stars lead to perhaps the most important reason why, irrespective of the mechanism at work, at least some neutron stars might be actally emitting detectable gravitational waves. This is the observation that even the fastest accreting neutron stars spin at rates much lower than the expected break-up frequency. The current record is 716 Hz, while the theoretically expected upper limit is more than 1 kHz. Following a suggestion by Bildsten [45], it is possible that this limit occurs because of the balance between the spin-up torque due to the accreting matter, and the spindown torque due to gravitational wave emission. A short calculation assuming a link between the observed X-ray luminosity with the accretion rate, and taking the mountain scenario for the emission mechanism leads to the following estimate of the GW amplitude:

$$h_0 = 3 \times 10^{-27} F_{-8}^{1/2} \left(\frac{R}{10\mathrm{km}}\right)^{3/4} \left(\frac{1.4 M_\odot}{M}\right)^{1/4} \left(\frac{1\ \mathrm{kHz}}{\nu_s}\right)^{1/2} . \tag{32}$$

This is seen to be dependent on frequency: $h_0 \propto \nu_s^{-1/2}$.

2.4.7 Intermediate mass black holes

The existence of intermediate mass black holes (IMBHs) with masses in the range $10^2 - 10^4\,M_\odot$ has not yet been corroborated observationally, but these objects are of high interest for astrophysics. Our understanding of the formation and evolution of supermassive black holes, as well as galaxy evolution modeling and cosmography would dramatically change if an IMBH were to be observed. From the point of view of traditional electromagnetic astronomy, which relies on the monitoring of stellar kinematics, the direct detection of an IMBH seems to be rather far in the future. However, the prospect of the detection and characterization of an IMBH has good chances in lower-frequency GW astrophysics, in particular with ET. The detection and characterization of a binary containing an IMBH would corroborate the existence of such systems and provide a robust test of general relativity through tests of the black hole uniqueness theorem (see subsection 2.3.6).

Signals from IMBH binaries can start in the band of LISA, and sweep through to the ET band, allowing us to observe different aspects of the coalescence event, as illustrated in Fig. 27. For ET, a lower cut-off frequency of 1 Hz was assumed.

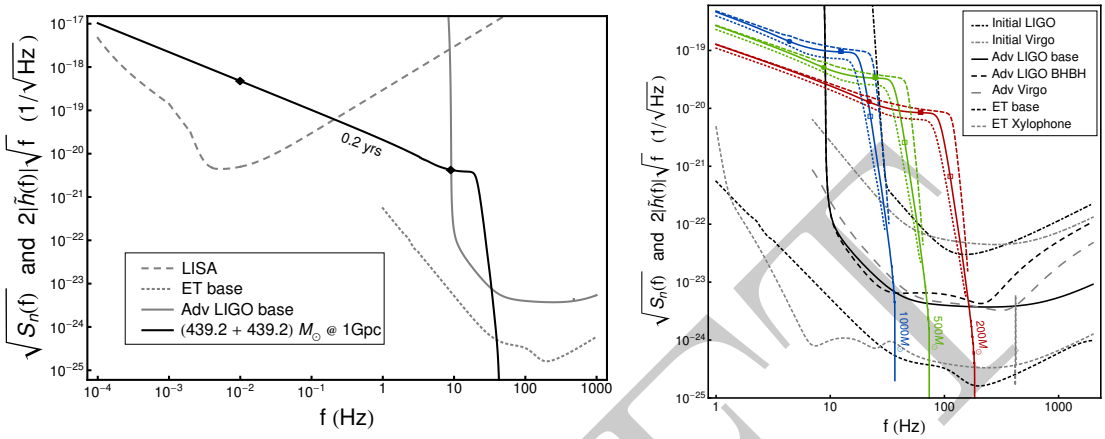


Figure 27: Left: different aspects of the inspiral, merger and ringdown signal from an intermediate mass black hole binary with masses $(439.2, 439.2)\, M_\odot$ at 1 Gpc can be studied as it passes from the LISA band into the ET band. The later, more interesting part of the evolution is seen in ET. Right: signals from IMBH binaries with different total masses. The solid lines are for equal mass, non-spinning binaries; the dashed lines for binaries with equal masses and dimensionless spins $\chi = 0.75$; and the dotted lines for zero-spin binaries with $m_2/m_1 = 3$.

2.5 Cosmology and Cosmography

2.5.1 Reconstruction of the evolution of compact binary coalescence rates by ET

The rate at which neutron stars and black holes coalesce at different redshifts can provide indirect but extremely valuable insights into star formation rate (SFR). We now consider how accurately ET would be able to distinguish between coalescence rate predictions from different SFR models. Considering that BNS coalescences are expected to be the most abundant, they are the best "trackers" of SFR , and these are the events we will focus on.

The rate per unit of time and per unit of comoving volume at which BNS systems are observed to coalesce at redshift z can be written as

$$\dot{\rho}_c^0(z) = \dot{\rho}_c^0(0)\frac{\dot{\rho}_{*,c}(z)}{\dot{\rho}_{*,c}(0)}. \tag{33}$$

Here $\dot{\rho}_c^0(0)$ is the coalescence rate at the current epoch, and $\dot{\rho}_{*,c}$ relates the past star formation rate to the rate of coalescence. One has

$$\dot{\rho}_{*,c}(z) = \int \frac{\dot{\rho}_*(z_f)}{(1+z_f)} P(t_d) dt_d, \tag{34}$$

where $\dot{\rho}_*$ is the SFR itself, z is the redshift at which the binary coalesces, z_f is the redshift at which the progenitor binary formed, and $P(t_d)$ is the probability distribution of the delay time t_d between the formation of the progenitor and coalescence. $P(t_d)$ has been estimated as

$$P(t_d) \propto \frac{1}{t_d} \quad \text{for } t_d > \tau_0, \tag{35}$$

where τ_0 is some minimum delay time. For more details we refer to [16] and references therein. As in that paper, for BNS coalescence we will assume $\tau_0 \sim 20$ Myr. The coalescence rate per unit redshift as observed in

our local Universe is found by multiplying $\dot{\rho}_c^0$ by the gradient of comoving volume:

$$\frac{dR_c^0}{dz} = \dot{\rho}_c^0(z)\frac{dV_c}{dz}(z). \tag{36}$$

We may now ask how well ET will be able to discriminate between different SFR models $\dot{\rho}_*(z)$ through differences in the resulting observed BNS coalescence rates $\dot{\rho}_c^0(z)$. To this end we consider four different models:

- Hopkins and Beacom [12]: An update of an observational compilation by Hopkins [**?**], placing lower bounds using the evolution of stellar mass density, metal mass density, and supernova rate density, and an upper bound using Super-Kamiokande results for the electron antineutrino flux from core-collapse supernovae;
- Nagamine et al. [169]: An approach comparing and combining results from direct observations, a model using local fossil evidence at $z \sim 0$, and theoretical *ab initio* models;
- Fardal et al. [170]: A model involving a new proposal for the initial mass function with a view on reconciling SFR predictions with the total extragalactic background radiation;
- Wilkins et al. [171]: Based on stellar mass density measurements together with a new ansatz for the initial mass function.

We place simulated BNS sources according to the coalescence rates $\dot{\rho}_c^0(z)$ inferred from these four models using their proposed $\dot{\rho}_*(z)$, with a minimum delay time $\tau_0 = 20$ Myr. Sources are positioned uniformly in the sky; an SNR cut is imposed such that a source is disregarded unless $\rho > 8$. We (very) conservatively set the coalescence rate at the current epoch to $\dot{\rho}_c^0 = 0.03\,\mathrm{Mpc}^{-3}\mathrm{Myr}^{-1}$, in which case between $\sim 150,000$ and $\sim 275,000$ sources survive the SNR cut, depending on the SFR model. Note that this local coalescence rate will most likely already be measured by Advanced LIGO and Virgo, so that we may consider it a known quantity in the context of ET.

Given a fiducial cosmological model, one can associate a luminosity distance $D_\mathrm{L}(z)$ with each simulated source. In a spatially flat Friedman-Robertson-Walker universe, this relationship takes the form

$$D_\mathrm{L}(H_0, \Omega_\mathrm{M}, \Omega_\Lambda, w; z) = \frac{1+z}{H_0}\int_0^z \frac{dz'}{\left[\Omega_\mathrm{M}(1+z')^3 + \Omega_\Lambda(1+z)^{3(1+w)}\right]^{1/2}}, \tag{37}$$

where H_0 is the Hubble parameter at the current era, Ω_M is the density of matter normalized by the critical density, Ω_Λ the density of dark energy (similarly normalized), and w is the equation-of-state parameter of dark energy. For definiteness, we choose our fiducial cosmological model such that $H_0 = 70\,\mathrm{km s^{-1} Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.27$, $\Omega_\Lambda = 0.73$, and $w = -1$.

The distance measured from the gravitational wave signal, $\hat{D}_\mathrm{L}$, will be different from the true distance due to (i) the noise in the detector, and (ii) (de)magnification as a result of weak lensing. Thus, with each source we associate a "measured" luminosity distance

$$\hat{D}_\mathrm{L}(z) = D_\mathrm{L}(z) + \delta D_\mathrm{L}(z), \tag{38}$$

where $\delta D_\mathrm{L}(z)$ is drawn at random from a Gaussian distribution with a spread given by

$$\Delta D_\mathrm{L} = (\sigma_\mathrm{ET}^2 + \sigma_\mathrm{WL}^2)^{1/2}, \tag{39}$$

The contribution σ_ET is due to detector noise while σ_WL results from weak lensing. Note that ΔD_L will depend not only on redshift but also on sky position and the orientation of the orbital plane. As a rule of thumb, one can take $\sigma_\mathrm{ET}/D_\mathrm{L} \simeq 1/\rho$, and for the weak lensing error we assume $\sigma_\mathrm{WL}/D_\mathrm{L} = 0.05z$.

Having distributed sources as described above and associated measured luminosity distances to them, we have a simulated "catalog" of detected inspiral events. Using the fiducial cosmological model, the measured luminosity distances $\hat{D}_\mathrm{L}(z)$ of Eq. (38) can be inverted to obtain measured redshifts $\hat{z}$. These recovered redshifts are then binned to obtain a recovered rate distribution $d\hat{R}_c^0/dz$. The measured luminosity distances and recovered

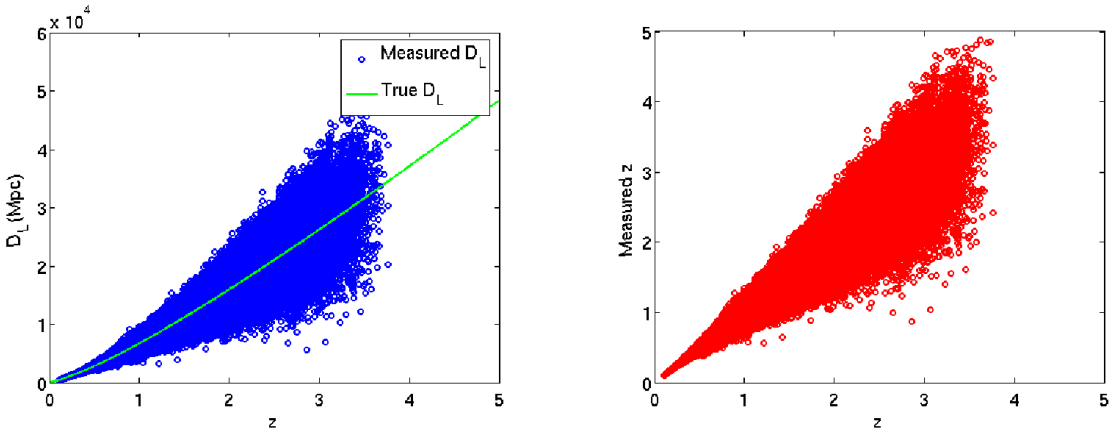


Figure 28: "Measured" luminosity distances (left) and inferred redshifts (right) for sources in a simulated catalog of BNS inspiral events.

redshift for one such catalog are shown in Fig. 28. By doing this for a large number of different simulated catalogs (say, 1000 catalogs), one can compute at an average and a 1-sigma spread for the number of sources in each recovered redshift bin.

To check how well ET will be able to distinguish between different coalescence rate predictions, we can fold in the anticipated efficiency $\epsilon(z)$, i.e., the fraction of coalescences at a given redshift that survive the SNR cut $\rho > 8$. The left panel of Fig. 29 shows efficiency as a function of redshift; it is essentially 1 up to $z \simeq 0.7$, after which it starts to drop rather quickly. Beyond $z \simeq 3.5$ no signals can be seen even when optimally positioned and oriented. The efficiency can be folded into the recovered rate distribution:

$$\left[\frac{dR_c^0(z)}{dz}\right]_{\rm recovered} = \epsilon(z)^{-1}\frac{dR_c'(z)}{dz}, \tag{40}$$

where dR_c'/dz is the distribution inferred from binning the measured redshifts.

The right hand panel of Fig. 29 shows both the underlying rates $dR_c^0(z)/dz$ and the recovered rates, with 1-sigma spreads, for the four SFR models we are considering. First we note some systematic effects due to uncertainties in the redshift measurements:

- At small reshifts ($z \lesssim 1.5$), the recovered rate distribution is shifted very slightly to the left with respect to the underlying distribution. This is because of higher-redshift events ending up in lower redshift bins due to measurement errors. The effect is not compensated by lower redshift events ending up in higher redshift bins, because at lower redshifts the spread in measured redshift is smaller;
- At intermediate redshifts ($1.5 \lesssim 3$) the true coalescence rate is being underestimated (despite having folded in efficiency loss) because of events ending up in both higher and lower measured redshift bins;
- Beyond $z \simeq 3.5$ the recovered rate diverges, because there are still *measured* redshift values there, but the efficiency $\epsilon(z) \to 0$.

We see that ET can easily distinguish between the four models we took from recent literature. Generally, two models for BNS coalescence rates can be distinguished from each other if over at least one $\Delta z = 0.1$ redshift bin at $z \lesssim 1.5$, the number of sources in the bin differs by more than a few percent.

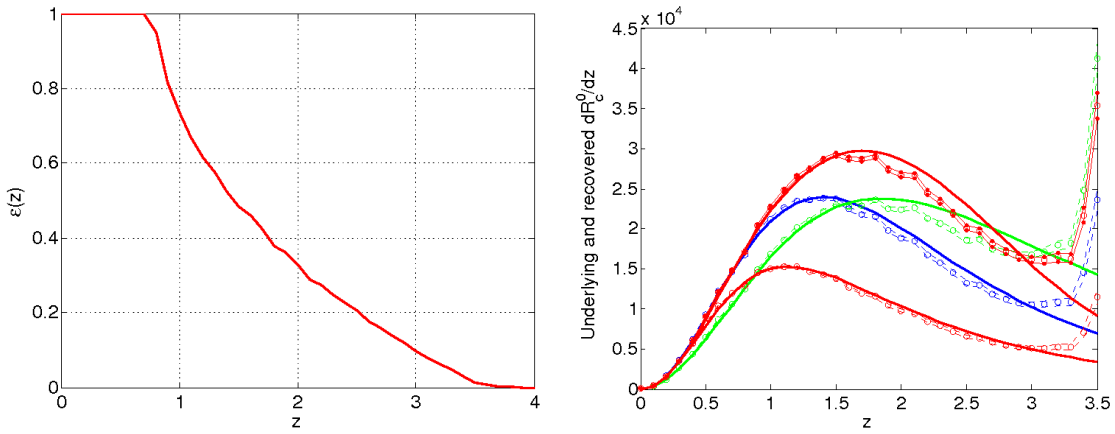


Figure 29: Left: the fraction of found versus missed signals as a function of redshift. Right: underlying and recovered coalescence rates (taking into account detection efficiency) for the models of Hopkins and Beacom [12] (top, red), Fardal et al. [170] (blue), Wilkins et al. [171] (bottom, red), and Nagamine et al. [169] (green). The solid lines are the true rates, the circles give the number of measured coalescences in a redshift bin, and the dashed lines give a 1-sigma spread in recovered redshifts.

2.5.2 Cosmological evolution of compact object populations

The calculation of the coalescence rate as a function of the redshift must take into account the following factors: the star formation rate history $SFR(z)$, the binary fraction $f_b(z)$, the formation efficiency of a given type of binary, i.e. the fraction of number of binaries that lead to formation of coalescing compact object binary, and their distribution of merger times. These quantities may depend on redshift since the stellar populations evolve with cosmic time. Let us examine the effects of evolution of each of these factors.

The star formation rate is known to increase strongly to the redshift $z = 2$, and there is a debate about its behavior for higher redshifts. At redshift $z = 2$, the star formation is estimated to be a factor of 10 larger than the present value at $z = 0$.

The distribution of merger times can be estimated either by analyzing the present population of compact objects binaries or by involving the population synthesis. The first approach is limited to deal with the double neutron star binaries, and suffers from small number statistics. The second involves several uncertainties due to parametrization of binary evolution. However the two approaches yield similar results. The distribution of merger times for the double neutron star binaries can be well approximated by a distribution $\propto t^{-1}$. The lower cutoff for the DNS systems lies somewhere between 10 and 100 Myrs. The population synthesis leads to similar conclusions about the distribution of merger times for BHNS and BBH systems, however the low time cutoff may probably lie higher.

The evolution of the properties of binaries with cosmic time. The main factor that may affect the evolution of the binaries as a function of redshift are the changes in the distribution of metallicity. Metallicity affects strongly the mass loss rate in stars, and hence has a strong influence on the masses spectrum of compact objects. The lower the metallicity the higher the maximum mass of a black hole that may be formed in the course of stellar evolution. This leads to to stabilization of mass transfers and therefore to increase in the formation rate of compact object binaries.

Taking together the above factors we see that there are several reasons why the coalescence rate should increase strongly as we go to redshifts of $z = 1$–2. First the local star formation rate increases and the overall number of binary formation is larger. Second, the typical delay times for the DNS systems are low therefore their merger rate density will roughly follow that of the SFR. In the case of BHNS or BBH systems he typical delay times between formation and coalescence may be as large as 1–3 Gyrs. This delays the peak of coalescence rate density

with respect to the star formation rate. Thus the delays are significant but not crucial. Third, the metallicity evolution may lead to higher compact object formation rate for high redshifts, and formation of larger number of massive BBH binaries.

This consideration can be put into detailed numerical codes to yield predictions about the rates. However even without such strong numerical support one can readily estimate with the back of the envelope calculation that the ratio of the coalescence rate (per unit volume per unti time) to the local one should be at least a few. The local coalescence rate can only be estimated with observations since neither the observational not the indirect approach mentioned earlier can yield the estimate of the rate with the accuracy better that plus minus an order of magnitude.

The Einstein telescope will provide a large sample of coalescences with the precise measurement of their masses and redshifts. This will be an extremely valuable tool for analysis of the cosmic compact object formation history. The measurement of their masses will yield information on the metallicity evolution as well as evolution of most massive stars. The Einstein telescope will yield a cosmic compact object census up to redshift $z = 2$, and will yield information about black holes and neutron star formed even at earlier epochs because of the delays between formation and coalescence.

There are two distinct routes to form BH binary. The first, conventional way, is to start with binary system of two main sequence stars and trace their evolution. There are several big uncertainties in this process. The first one is the initial mass ratio function: what is the distribution of the mass ratio in the binary of two main sequence stars, how it depends on the metallicity and spectral type. The second, and probably the biggest uncertainty, is related to the "common envelope" evolution, where the NS (or BH) and Helium core are emerged and evolve in the gaseous environment of the star. In this stage the NS/BH could merge with Helium core and binary is not formed. The third uncertainty is related to the direction and magnitude of the kick exerted on the newly born BH from the assymetric supernova explosion. All the above is reflected in the uncertainties on the rate of such binaries [10, 11].

The BH binaries could also be formed in the dense environment such as galactic nuclei. In the galaxies with SMBH ($M < 10^7\,M_\odot$), the relaxation time is less than a Hubble time, and a steep cusp of stars and stellar mass BHs can be formed. BHs as more massive and compact objects will segregate into central ≈ 1 pc region. Other two dense regions are massive globular clusters and nuclear star clusters in the centers of low-mass galaxies which may not have SMBH. The densities in those regions are high enough to have multiple encounters with formation and/or hardening of the BH binaries.

2.5.3 Intermediate mass black holes as seeds of galaxy formation

It is widely accepted that the massive black holes (MBHs) found in the centres of many galaxies grow from initial seeds through the processes of accretion and mergers following mergers between their host dark matter halos. However, little is known about the seeds from which these black holes grow. Open questions include — how and when did they form? What are their masses? Where are they? Current observations are consistent with both *light seed* scenarios, in which $\sim 100M_\odot$ black hole seeds form at redshift $z \approx 20$ from the collapse of Population III stars [172, 173], and *heavy seed* scenarios, in which black holes of mass $\sim 10^5 M_\odot$ form from direct collapse of dust clouds [174, 175]. Mergers between MBHs in merging dark matter halos will generate gravitational waves. These are a major source for LISA [176], but LISA will only see mergers with total mass $\gtrsim 10^3 M_\odot$. LISA can therefore probe black hole seeds only in the heavy seed scenario and does not have the power to discriminate between the light and heavy scenarios.

The Einstein Telescope will have sensitivity in the 1–50Hz band in which gravitational waves from mergers involving $\sim$ 10–$100M_\odot$ black holes will lie. It will therefore provide complementary information to LISA and could directly observe the first epoch of mergers between light seeds. Present estimates, based on Monte Carlo simulations of galaxy merger trees [177, 178], suggest the Einstein Telescope could detect between a few and a few tens of seed black hole merger events over three years of operation [179]. Several of these events will be at high redshift, $z \sim 5$, by which time it is unlikely that $100M_\odot$ black holes could have formed by other routes. ET and LISA in conjunction probe the whole merger history of dark matter halos containing black

holes in the 10–$10^6 M_\odot$ range, which will provide detailed information on the hierarchical assembly of galaxies. ET on its own is not able to measure the distance to a gravitational wave source, but provided one additional, non-colocated, interferometer is in operation concurrently with at least two detectors at the ET site, the network will be able to determine the luminosity distance of a source to $\sim 40\%$ precision and the redshifted total mass of the system to $\lesssim 1\%$. Using a concordance cosmology, this distance estimate can be used to estimate redshift with comparable accuracy and so it should be possible to say that the $z \approx 5$ events are of *low mass* and at *high redshift*, and therefore are convincing candidates as Pop III seed mergers.

Just one detection by ET will rule out the heavy seed model. With several detections, we will be able to make statements about pop III seed black hole properties, such as their mass distribution, their early accretion history etc. [179]. These observations cannot be made be any other existing or proposed detector — it is science that is unique to ET. Such observations will be vital to our understanding of the assembly of structure in the Universe, and of the close link between black holes residing in the centres of galaxies and their hosts [172].

2.5.4 Cosmography with a population of standard sirens

The goal of modern cosmology is to measure the geometrical and dynamical properties of the Universe by projecting the observed parameters onto a cosmological model. The Universe has a lot of structure on small scales, but on a scale of about 100 Mpc the distribution of both baryonic (inferred from the electromagnetic radiation they emit) and dark matter (inferred from large scale streaming motion of galaxies) components is quite smooth. It is, therefore, quite natural to assume that the Universe is homogeneous and isotropic while describing its large-scale properties. In such a model, the scale factor $a(t)$, which essentially gives the proper distance between comoving coordinates, and curvature of spatial sections k, are the only quantities that are needed to fully characterize the properties of the Universe. The metric of a smooth homogeneous and isotropic spacetime is

$$ds^2 = -dt^2 + a^2(t)\frac{d\sigma^2}{1-k\sigma^2} + \sigma^2\left(d\theta^2 + \sin^2\theta\, d\varphi^2\right),$$

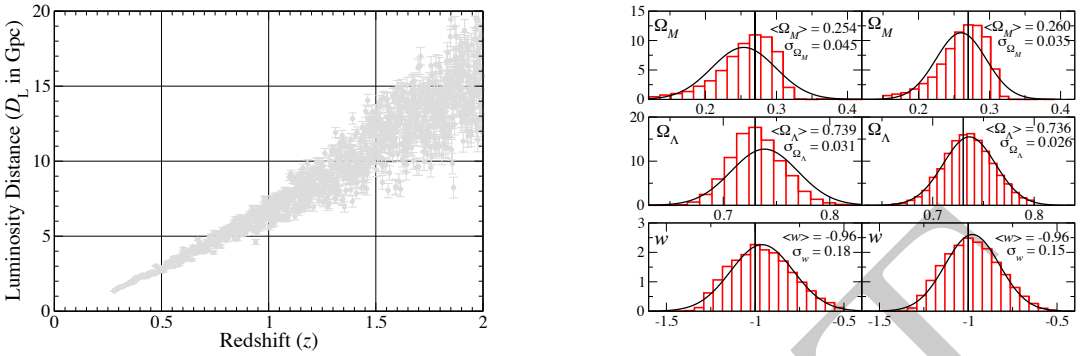
where t is the cosmic time-coordinate, $(\sigma, \theta, \varphi)$ are the comoving spatial coordinates, and k is a parameter describing the curvature of the t = const. spatial slices. $k = 0, \pm 1$, for flat, positively and negatively curved slices, respectively. The evolution of $a(t)$, of course, depends on the parameter k, as well as the "matter" content of the Universe. The latter could consist of radiation, baryons, dark matter (DM), dark energy (DE), and everything else that contributes to the energy-momentum tensor.

The Friedman equation, which is one of two Einstein equations describing the dynamics of an isotropic and homogeneous Universe, relates the cosmic scale factor $a(t)$ to the energy content of the Universe through

$$H(t) = H_0\left[\hat{\Omega}_{\rm M}(t) - \frac{k}{H_0^2 a^2} + \hat{\Omega}_\Lambda(t)\right]^{1/2}, \tag{41}$$

where $H(t) \equiv \dot{a}(t)/a(t)$ is the Hubble parameter ($H_0 = H(t_P)$ being its value at the present epoch t_P), $\hat{\Omega}_{\rm M}(t)$ and $\hat{\Omega}_\Lambda(t)$ are the (dimensionless) energy densities of the DM and DE, respectively. The above equation has to be supplemented with the equation-of-state of DM, assumed to be pressure-less fluid $p = 0$ [$\hat{\Omega}_{\rm M}(t) = \Omega_M(1+z)^3$, $\Omega_{\rm M} = \hat{\Omega}_{\rm M}(t_P)$] and of DE, assumed to be of the form $p = w\rho_\Lambda$ [$\hat{\Omega}_\Lambda(t) = \Omega_\Lambda(1+z)^{3(1+w)}$, where $\Omega_\Lambda = \Omega_\Lambda(t_P)$], with $w = -1$ corresponding to a cosmological constant. The goal of cosmography is to measure $(H_0, \Omega_{\rm M}, \Omega_\Lambda, w, k, \ldots)$, which essentially determine the large-scale geometry and dynamics of the Universe. In the rest of this paper we shall assume that the spatial slices are flat (i.e., $k = 0$).

Measuring the cosmological parameters Astronomers use "standard candles" to measure the geometry of the Universe and the various cosmological parameters. A standard candle is a source whose intrinsic luminosity L can be inferred from the observed properties (such as the spectral content, time-variability of the flux of radiation, etc.). Since the observations also measure the apparent luminosity F, one can deduce the luminosity distance $D_{\rm L}$ to a standard candle from $D_{\rm L} = \sqrt{L/(4\pi F)}$. In addition, if the red-shift z to the source is known



ee

Figure 30: The plot on the left shows one realization of a catalogue of binary neutron star (BNS) events that might be observed by ET. The plot on the right shows the distribution of errors in $\Omega_{\rm M}$, Ω_Λ and w obtained by fitting 5,190 realizations of a catalogue of BNS merger events to a cosmological model of the type given in Eq. (42), with three free parameters. The fractional 1-σ width of the distributions $\sigma_{\Omega_{\rm M}}/\Omega_{\rm M}$, $\sigma_{\Omega_\Lambda}/\Omega_\Lambda$, and $\sigma_w/|w|$, are 18%, 4.2% and 18% (with weak lensing errors in $D_{\rm L}$, left panels) and 14%, 3.5% and 15% (if weak lensing errors can be corrected, right panels).

then by observing a population of such sources it will be possible to measure the various cosmological parameters since the luminosity distance is related, when $k = 0$, to the red-shift via

$$D_{\rm L} = \frac{c(1+z)}{H_0}\int_0^z \frac{dz'}{\left[\Omega_{\rm M}(1+z')^3 + \Omega_\Lambda(1+z')^{3(1+w)}\right]^{1/2}}. \tag{42}$$

There is no unique standard candle in astronomy that works on all distance scales. An astronomer, therefore, builds the distance scale by using several steps, each of which works over a limited range of the distance. For instance, the method of parallax can determine distances to a few kpc, Cepheid variables up to 10 Mpc, the Tully-Fisher relation works for several tens of Mpc, the D_n-σ relation up to hundreds of Mpc and Type Ia supernovae up to red-shifts of a few. This way of building the distance scale has been referred to as the *cosmic distance ladder*. For cosmography, a proper calibration of the distance to high red-shift galaxies is based on the mutual agreement between different rungs of this ladder. It is critical that each of the rungs is calibrated with as little an error as possible.

Fitting a cosmological model to a CBC population The expected rate of coalescences per year within the horizon of ET is $\sim$ several $\times\, 10^5$ for BNS and NS-BH. Such a large population of events to which luminosity distances are known pretty accurately, would be very useful for measuring cosmological parameters. If, as suspected, BNS and NS-BH are progenitors of short-hard gamma-ray bursts (GRBs) [154], then it might be possible to make a coincident detection of a significant subset of the events in GW and EM windows and obtain both the luminosity distance to and red-shift of the source.

Since GRBs are believed to be beamed with beaming angles of order 40°, we assume that only a small fraction ($\sim 10^{-3}$) of binary coalescence events will have GRB or other EM afterglows that will help us to locate the source on the sky and measure its red-shift z. Eventually, we will be limited by the number of short-hard GRBs observed by detectors that might be operating at the time. As a conservative estimate, we assume that about $1,000$ BNS and NS-BH mergers will have EM counterparts over a three-year period. For definiteness we consider only BNS mergers and take these to have component masses of $(1.4, 1.4)M_\odot$.

How well would we measure cosmological parameters with a catalogue of such sources? To answer this question we simulated 5,190 realizations of the catalogue containing 1,000 BNS coalescences with known red-shift and sky location, but the luminosity distance subject to statistical errors from GW observation and weak lensing. One

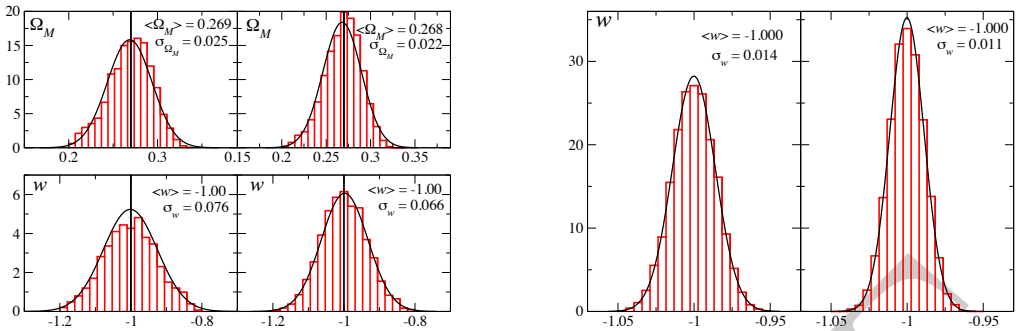


Figure 31: Same as the right plot of Fig. 30 except that one or more of the cosmological parameters are assumed to be known. The plot on the left assumes that Ω_Λ is known to be $\Omega_\Lambda = 0.73$, and fits the "data" to the model with two free parameters. The fractional 1-σ widths in the distribution $\sigma_{\Omega_M}/\Omega_M$ and and $\sigma_w/|w|$, are 9.4% and 7.6% (with weak lensing errors in D_L, left panels) and 8.1% and 6.6% (if weak lensing errors can be corrected, right panels). The plot on the right is the same but assuming that w is the only unknown parameter. The fractional 1-σ width of the distribution $\sigma_w/|w|$ is 1.4% (with weak lensing errors in D_L, left panel) and 1.1% (if lensing errors can be corrected, right panel).

such realization is shown in Fig. 6 (right panel). We assumed that the sources were all in the red-shift range $0 \le z \le 3.5$, distributed uniformly (i.e., with constant comoving number density) throughout this red-shift range. The luminosity distance to the source was computed by assuming an FRW cosmological model with $H_0 = 70\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_M = 0.27$, $\Omega_\Lambda = 0.73$, and $w = -1$, but the *measured* distance was drawn from a Gaussian distribution whose width σ_{D_L} was determined by the quadrature sum of the errors due to weak lensing and GW observation. Weak lensing error in D_L was assumed to be 5% at $z = 1$ and linearly extrapolated to other red-shifts. GW observational error was estimated from the covariance matrix C_{km} of the five-dimensional parameter space of the unknown signal parameters $p_k = (M, \nu, t_0, \Phi_0, D_L)$:

$$C_{km} = \Lambda^{-1}_{km}, \quad \Lambda_{km} = \langle h_k,\, h_m \rangle\,, \quad h_k = \frac{\partial h}{\partial p_k}. \tag{43}$$

Here the angular brackets denote the scalar product, which, for any two functions $a(t)$ and $b(t)$, is defined as

$$\langle a,\, b \rangle = 4\Re \int_0^\infty \frac{\mathrm{d}f}{S_h(f)} A(f)\, B^*(f) \tag{44}$$

where A and B are the Fourier transforms of the functions $a(t)$ and $b(t)$, respectively, and $S_h(f)$ is the ET noise power spectral density. Note that since GRBs are expected to be strongly beamed, we did not take the angles (ι, ψ) associated with the unit normal to the plane of the inspiral as unknown variables. This assumption is justified: even if the opening angle of a GRB beam is as large as 40°, the unit normal to the plane of the inspiral would still be confined to only 3% of the area of a unit sphere. Averaging errors over (ι, ψ) with the constraint $\iota < 20°$ would then be little different from taking $\iota = 0°$. We did, however, average the errors over the sky position angles (θ, ϕ). We then fitted each realization of the source catalogue to the cosmological model given in Eq. (42), using the Levenberg-Marquardt algorithm [180, 181], in order to find a set of best fit parameters. It turns out that a catalogue of 1,000 sources is not quite enough for an accurate determination of all the parameters. However, assuming that H_0 is known accurately, the algorithm gave the best fit parameters in $(\Omega_M, \Omega_\Lambda, w)$ for each of the 5,190 realizations.

Accurate measurement of the Hubble constant at low redshifts ($z \simeq 0.01 - 0.5$) Taking H_0 to be the only free parameter and with only 50 sources up to a redshift of $z = 0.5$, ET itself would achieve an accuracy of 0.55% (using the kind of estimation explained here). For sufficiently low-redshift sources, the rest of cosmology is not very important. Having determined H_0 in this way, the higher-redshift sources can then be used to explore

the cosmological parameter space more fully. In this way, GW astronomy will provide an independent measure of cosmography.

Thus, ET itself can use sources in the low-redshift Universe to measure H_0 with an error that is negligible compared to the uncertainties on the other parameters obtained from all sources combined, as we shall see below. In the rest of the discussion we will consider H_0 essentially known.

Measuring matter and dark energy densities The distributions $\mathcal{P}$ of the parameters obtained in the above way are shown in Fig. 30, where the vertical line is at the true value of the relevant parameter. The relative 1-σ errors in Ω_Λ, Ω_M and w, are 4.2%, 18% and 18% (with weak lensing, left panels) and 3.5%, 14% and 15% (with weak lensing errors corrected, right panels). Although $\mathcal{P}(w)$ is quite symmetric, $\mathcal{P}(\Omega_\mathrm{M})$ and $\mathcal{P}(\Omega_\Lambda)$ are both skewed and their mean values are slightly off the true values. However, the medians are mostly coincident with the true values.

In addition to H_0 if Ω_Λ is also known (or, equivalently, if $\Omega_\mathrm{M}+\Omega_\Lambda = 1$), then one can estimate the pair (Ω_M, w) more accurately, with the distributions as shown in Fig. 30 with greatly reduced skewness and 1-σ errors in Ω_M and w, of 9.4% and 7.6% (with weak lensing) and 8.1% and 6.6% (with lensing errors corrected). Finally, if w is the only parameter unknown, it can be measured to an even greater accuracy as shown in Fig. 30 with 1-σ errors of 1.4% (with weak lensing) and 1.0% (with lensing errors corrected).

Effect of unknown orientation and polarization In the previous section our study neglected the effect of different inclinations of the orbit to the line-of-sight. Varying the inclination has two distinct effects: On the one hand, as noted in Ref. [182], due to the strong correlation between the luminosity distance and inclination, the estimation of luminosity distance could get corrupted. On the other hand, binaries that are not face-on are, in general, elliptically polarized and have a non-zero polarization angle. Since polarization angle is correlated with the luminosity distance, there could be further degradation in the estimation of the luminosity distance.

In this section we will relax the condition that the inclination of the orbit is precisely known. However, we shall restrict the inclination of the binary's angular momentum with the line-of-sight to be within 20 degrees. We shall also assume that the radiation is described by an arbitrary polarization angle. Since the sky position is still assumed known, this gives us a 7×7 covariance matrix with a revised estimate for the error in the luminosity distance. As before, we construct catalogs of binary coalescence events but with the luminosity distance now drawn from a Gaussian distribution with revised widths. We fit each catalog to a cosmological model and then repeat the exercise 5,190 times to estimate the accuracy with which the various cosmological parameters can be measured.

As expected, the parameter measurments get worse if we assume two or more parameters to be unknown. For instance, errors in the estimation of Ω_M, Ω_Λ and w, are, respectively, $> 100\%$, 24% and 47% with weak lensing and $> 100\%$, 21% and 43%, if weak lensing can be corrected. Similarly, if Ω_Λ is assumed to be known then the errors in the estimation of Ω_M and w are, respectively, 12% and 9.5% if weak lensing is uncorrected for and 11% and 9.2% if weak lensing can be corrected. However, the results are more or less the same if dark energy parameter w is the only unknown quantity. Even when the inclination and polarization angles are taken as a free parameters, but inclination angle is restricted to within 20 degrees, the error in the estimation of w is 1.4% with weak lensing and 1.3% if weak lensing can be corrected.

Variation of dark energy with redshift We have seen that the Hubble constant H_0 can be measured with high accuracy using low-redshift sources, after which parameters like Ω_M, Ω_Λ, and the dark energy equation-of-state parameter w can be determined, where so far we have assumed that the latter is constant. One could go one step further and use prior information from, e.g., future Planck CMB measurements to get high-accuracy values for $(\Omega_\mathrm{M}, \Omega_\Lambda)$, and then measure the variation of w with time. Since CMB data would have a very wide prior on w and its first time derivative, this would constitute an independent measurement of the latter variables.

In practice, rather than looking at variation with time, it is more convenient to consider variation with scale factor or redshift:

$$\begin{aligned} w(z) &= w_0 + w_a(1-a) + \mathcal{O}\left[(1-a)^2\right] \\ &\simeq w_0 + w_a \frac{z}{1+z}. \end{aligned} \tag{45}$$

Since here we are mostly interested in the later stages of the universe's evolution, higher order terms will be ignored. With a redshift dependent $w(z)$, the Hubble parameter as a function of redshift becomes

$$H(z) = H_0 \left[\Omega_{\rm M}(1+z)^3 + \Omega_\Lambda (1+z)^{3(1+w_0+w_a)} e^{-3w_a z/(1+z)}\right], \tag{46}$$

and luminosity distance is still related to the Hubble parameter as

$$D_{\rm L}(z) = c(1+z)\int_0^z \frac{dz'}{H(z')}. \tag{47}$$

To estimate the constraints on cosmological parameters from future Planck CMB data, one can consider the Fisher matrix

$$F_{ij}^{\rm CMB} = \sum_{\ell=2}^{\ell_{max}} \sum_{XX',YY'} \frac{\partial C_\ell^{XX'}}{\partial p_i} {\rm Cov}^{-1}(D_\ell^{XX'}, D_\ell^{YY'}) \frac{\partial C_\ell^{YY'}}{\partial p_j}, \tag{48}$$

wherer p_i are the cosmological parameters to be evaluated; $C_\ell^{XX'}$ are the CMB power spectra and $D_\ell^{XX'}$ their estimates; and ${\rm Cov}^{-1}$ is the inverse of the covariance matrix at given angular size ℓ and channels X, X' (T for temperature, E for polarization). For a detailed discussion we refer to [183].

Given a collection of inspiral events, we can also associate a Fisher matrix to the set of events:

$$F_{ij}^{\rm GW} = \sum_k \frac{\partial_i(\ln D_{\rm L}(z_k))\partial_j(\ln D_{rmL}(z_k))}{\Delta \ln D_{\rm L}(z_k)}, \tag{49}$$

where ∂_i, ∂_j are partial derivatives with respect to cosmological parameters, and the k refer to individual events. This Fisher matrix can be combined with the one for the CMB as

$$F_{ij}^{\rm combined} = F_{ij}^{\rm GW} + F_{ij}^{\rm CMB}. \tag{50}$$

Allowing for GRB beaming angles of 40°, one then finds $\Delta w_0 = 0.096$ and $\Delta w_a = 0.30$, which is comparable to projections from both the SNAP Type Ia supernova and the JDEM Baryon Acoustic Oscillations projects [183]; see Fig. 32. However, we stress that GW standard sirens are *self-calibrating* and have no dependence on a cosmic distance ladder.

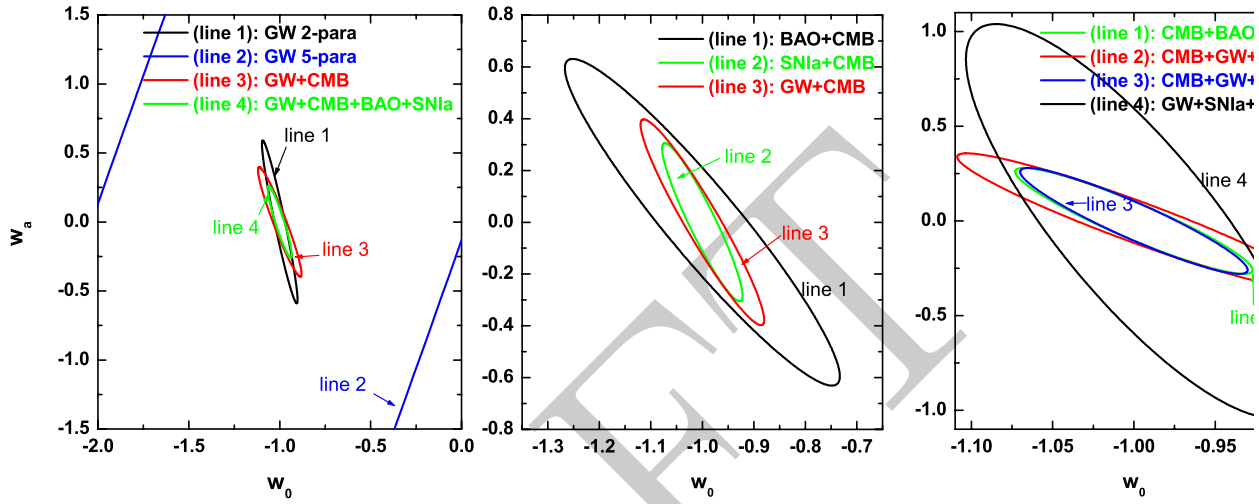


Figure 32: Investigating dark matter with GW standard sirens compared with other methods. Left: The accuracy in (w_0, w_a) with GW standard sirens, using projected Planck CMB accuracies as a prior for the other cosmological parameters (line 3); line 4 shows the minor improvements when also incorporating BAO and SNIa supernovae results. Middle: Assuming a Planck CMB prior throughout, one can compare accuracies from Baryon Acoustic Oscillations (BAO), SNIa supernovae, and GW standard sirens; they are comparable, but it should be noted that GW standard sirens are *self-calibrating*. Right: accuracies for various combinations of the GW, SNIa, BAO, and CMB methods.

2.5.5 Primordial gravitational waves

The cosmological stochastic background of GW [184] is an unique window on the very early Universe, as gravitational radiation propagates uninterrupted to us from cosmic events at the highest temperatures and densities, potentially up to the GUT scale 10^{16} GeV. The detection of any such background would have huge consequences for fundamental physics, possibly giving us direct indications of inflation, phase transitions or formation of topological defects. As shown in Figure 33, many types of cosmological stochastic background are potentially above the ET sensitivity curve. It may also be possible to extract detailed information about the cosmological events that produced GW if the spectrum has some characteristic shape.

Cosmological backgrounds are broadly of two types: wide-band, where $\Omega_{gw}(f)$ is approximately constant over a large range of frequency; and peaked, where Ω_{gw} varies strongly over f reaching a maximum at $f_{\rm peak}$. Wide-band sources are processes that extend over a large range of the cosmological scale factor $a(t)$, such as inflation and cosmic string evolution. Both these sources depend on unknown fundamental physics, and also have an approximate scaling symmetry. The detectability of a "flat" background spectrum depends only on the value of Ω_{gw} at a nominal frequency of 10 Hz.

Wide-band sources: inflation and strings There are many models of inflation, but they share a few essential features: exponentially expanding the scale factor in a short time; sourcing primordial density perturbations with amplitude $\mathcal{O}(10^{-5})$ and an approximately Harrison-Zeldovich ($n_s \simeq 1$) spectrum; and finally reheating the

Universe to at least the temperature required for primordial nucleosynthesis (order of 10 MeV).

Tensor perturbations sourced during inflation can be described by an amplitude and spectral index n_t. Since they evolve similarly to the scalar perturbations, often only the *tensor-to-scalar ratio* r is considered. This is sensitive to the "energy scale of inflation" $V^{1/4}$ as $V^{1/4} \simeq (1.8 \times 10^{16}\,\mathrm{GeV})(r/0.07)^{1/4}$ [185]. In single-field models, the value of r also indicates the minimum distance the field travels during inflation via $\Delta\phi/M_P \simeq 0.46(r/0.07)^{1/2}$.

For a scale-invariant spectrum ($n_s = 1$, $n_t = 0$) the CMB determination of the scalar amplitude $S \sim 10^{-10}$ together with the current bound $r \lesssim 0.2$ translates to a very small value $\Omega_{gw}(f) \lesssim 10^{-15}$ for all frequencies accessible to interferometers [186, 187]. However, since the CMB bounds apply at $f \sim 10^{-18}$ Hz, a positive spectral index n_t could change the picture [188]: we have $\Omega_{gw}(10\,\mathrm{Hz}) = (10^{19})^{\bar{n}_t}\Omega_{gw}(10^{-18}\,\mathrm{Hz})$, where $\bar{n}_t$ is the averaged spectral index between these two frequencies [189]. In Fig. 33 we plot the possible signal for $r = 0.15$, $n_t = 0.2$.[3]

Alternatives to inflation There are two recognized alternatives to exponential inflation as a source of primordial perturbations: the 'pre-big-bang' scenario in string cosmology and the "ekpyrotic" and "cyclic" models involving a contracting phase and subsequent brane collision. For both, it is currently debated whether a realistic spectrum of scalar density perturbations can be achieved. The tensor or gravitational wave amplitude is known to be undetectably small in the "cyclic" model [190].

A potentially more interesting case arises from pre-big-bang scenarios in string cosmology [**?** **?**]. According to these models, the standard radiation-dominated and matter-dominated eras were preceded by phases in which the Universe was first large and shrinking (inflaton phase) and then characterized by a high curvature (stringy phase). The GW spectrum produced at the transition between the stringy phase and the RD era is described as $\Omega_{gw}(f) \sim f^3$ for $f < f_s$ and $\Omega_{gw}(f) \sim f^{3-2\mu}$ for $f_s < f < f_1$ [**?** **?**]. The cutoff frequency f_1, which depends on string related parameters, has a typical value of 4.3×10^{10} Hz. An upper limit on Ω_{gw} is imposed by the Big Bang Nucleosynthesis (BBN) bound down to 10^{-10} Hz, corresponding to the horizon size at the time of BBN. Measurements of the light element abundances combined with the WMAP data gives $N_\nu < 4.4$ [**?**], which translates to $\Omega_{gw} < 1.5 \times 10^{-5}$. Recent measurements of CMB anisotropy spectrum, galaxy power spectrum and of the Lyman-α forest give a bound of similar amplitude which extends down to 10^{-15} Hz, corresponding to the horizon size at the time of CMB decoupling [**?**].

However, attempts to reach a more realistic spectrum of scalar perturbations by considering specific forms of pre-big-bang cosmological evolution may lead to much smaller values of Ω_{gw} at frequencies accessible to ground-based detectors [191]. The spectrum and amplitude of primordial GW in such scenarios is strongly model-dependent.

Cosmic string evolution Cosmic strings (see *e.g.* [192]) in field theory are extended topological defects formed in phase transitions. Fundamental strings may also result from cosmological evolution, for instance in "brane inflation" models [193] strings are formed at a brane collision near the end of inflation [194]. After formation strings evolve by reconnection and oscillation of the resulting loops, which emit gravitational radiation (and possibly other quanta) and gradually shrink. The evolution is believed to have a scaling property and produces an almost flat spectrum in Ω_{gw} across frequencies accessible to interferometers.

The most important parameter is the string tension or energy per unit length $G\mu$, determined by the energy scale of the phase transition; in brane inflation this may take values from 10^{-6} down to 10^{-11}. The current limit from CMB and other cosmological probes is a few times 10^{-7} [195–197]. Fundamental strings also have a "reconnection probability" p significantly smaller than unity (the value for field-theoretic strings). The properties of both field-theory and fundamental string networks can be summarized by parameters $\alpha < 1$ (size of newly-created loops relative to the Hubble horizon) and $\Gamma \sim 50$ (gravitational-wave luminosity of string loops). These are subject to uncertainty from numerical simulations. There are two limits: the "large loop" case where α is

[3] Such optimistic values are however not consistent with most scalar field models.

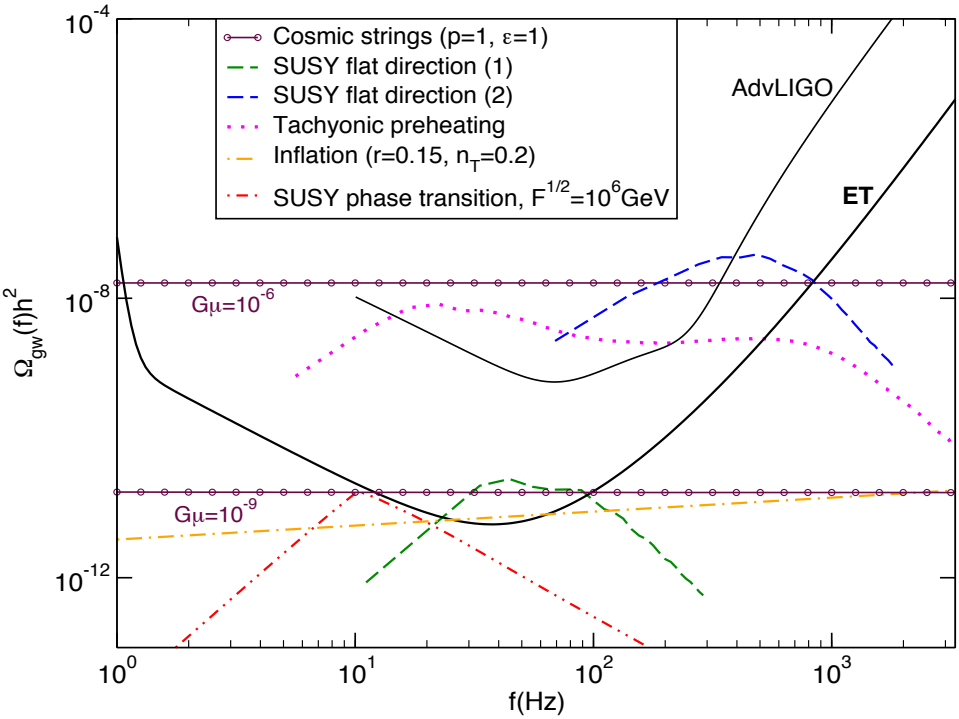


Figure 33: Cosmological stochastic GW backgrounds at Advanced LIGO and Einstein Telescope. The sensitivity curves correspond to an observation time of 1 year, S/N of 2.56, and co-located but not necessarily coaligned detectors (see end of Section 2.2.3). Models and parameter values are described in the main text. Data for tachyonic preheating and decay of SUSY flat directions were provided by J.-F. Dufaux; for phase transitions between metastable SUSY vacua, by N. J. Craig; the cosmic string GW spectra are based on a calculation of X. Siemens *et al.* [200].

comparable to unity, for which the "plateau" value of $\Omega_{gw}(f)$ may be estimated [198] as

$$\Omega_{gw}(f) \sim 10^{-8}(G\mu/10^{-9})^{1/2}p^{-1}(0.2\Gamma\alpha)^{1/2}. \tag{51}$$

The "small loop" case is motivated if the size of loops is determined by gravitational backreaction, giving $\alpha \simeq \Gamma G\mu \ll 1$; deviations from this value [199] are parameterized by a factor ϵ. In Fig. 33 we use a recent evaluation [200] of the GW spectrum for $p = \epsilon = 1$.

Peaked sources: phase transitions and reheating Peaked sources of stochastic background result from an event localized in cosmic time, typically a phase transition or reheating after inflation. There are many candidates arising from models of high-energy physics. Their detectability depends on the value of $f_{\rm peak}$ as well as the amplitude $\Omega_{gw}(f_{\rm peak})$; either side of $f_{\rm peak}$ the spectrum will decline as a power law [201]. The present-day frequency f is related to the frequency at the time of production f_* via

$$f = f_*\frac{a_*}{a_0} \approx (6\times10^{-8}\,{\rm Hz})\frac{f_*}{H_*}\frac{T_*}{1\,{\rm GeV}}\left(\frac{g_*}{100}\right), \tag{52}$$

(see *e.g.* [202]), where a is the scale factor, T the temperature, g the number of relativistic degrees of freedom and H the Hubble rate, the suffixes "0" and "$*$" denoting the present time and time of production respectively.

Phase transitions and colliding bubbles First order phase transitions proceed by the nucleation of spherical bubbles in a "false vacuum" with latent heat (energy density) ϵ. The bubbles grow rapidly and may collide;

after collision the bubble walls have a nonzero, rapidly-varying quadrupole moment and radiate gravitational waves. In the latter stages of the transition, gravitational waves may also be sourced by turbulence [203] as the energy difference ϵ is converted into heat.

The transition is characterized by the temperature T_* at which bubble nucleation occurs and the duration or characteristic timescale β^{-1}, assumed much shorter than a Hubble time H_*^{-1}. The present peak frequency and amplitude of GW are estimated as [204]

$$\begin{aligned} f_{\rm peak} &\simeq (5.2\times10^{-8}\,{\rm Hz})\frac{\beta}{H_*}\cdot\frac{T_*}{1\,{\rm GeV}}\left(\frac{g_*}{100}\right)^{1/6}, \\ \Omega_{gw}(f_{\rm peak})h_{100}^2 &\simeq (1.1\times10^{-6})\kappa\left(\frac{\alpha}{1+\alpha}\right)^2\left(\frac{v^3}{0.24+v^3}\right)\left(\frac{H_*}{\beta}\right)^2\left(\frac{100}{g_*}\right)^{1/3}, \end{aligned} \tag{53}$$

where α is a measure of the strength of the phase transition, κ is an "efficiency factor" for conversion of false vacuum energy to kinetic energy, and v is the speed of expansion of the bubbles. In the limit of a strongly first-order transition $\alpha \gg 1$, $v \to 1$ and $\kappa \to 1$, while β/H_* is expected to be of order 10^2. Turbulent plasma motion leads to similar values.

The electroweak phase transition with $T_* \sim 100\,{\rm GeV}$ is likely to produce GW with frequencies at or below the milliHz range accessible to LISA [205]. Similarly in the "Randall-Sundrum 1 model" there may be a first-order phase transition in the warped extra-dimensional geometry at temperatures of around $10^3\,{\rm GeV}$ [206]. A transition temperature of 10^6–$10^7\,{\rm GeV}$ corresponds to the sensitive range of ET [202]: this could be achieved for phase transitions between metastable SUSY-breaking vacua [207]. In Fig. 33 we plot one scenario with a hidden sector SUSY-breaking scale $\sqrt{F} = 10^6\,{\rm GeV}$.

Reheating and related phenomena At the end of inflation, the Universe is reheated by converting the inflationary energy density into radiation. In the "preheating" scenario the fluctuations of fields coupled to the inflaton grow exponentially rapidly via parametric resonance. The stochastic GW spectrum produced from preheating after chaotic inflation has a peak value $\Omega_{gw}(f_{\rm peak}) \gtrsim 10^{-11}$ [208, 209], however the peak frequency is well above the range of interferometers, unless coupling constants in the model take fine-tuned (extremely small) values.

Hybrid inflation ends due to the presence of a *tachyon* or "Higgs" field whose value sits at the top of a hill-shaped potential. The field decays by spinodal instability with a characteristic spectrum, giving rise to bubble-like regions which collide, fragment and finally thermalize. The GW spectrum resembles that of a phase transition, and for specific parameter values may be accessible to ET [210, 211]; its peak frequency and amplitude are estimated as

$$\begin{aligned} f_{\rm peak} &\simeq (6\times10^{10}\,{\rm Hz})Cg\lambda^{1/4}, \\ \Omega_{gw}(f_{\rm peak})h_{100}^2 &\simeq (2\times10^{-6})\left(\frac{v}{M_P}\right)^2(Cg)^{-2}, \end{aligned} \tag{54}$$

for $g^2 \gtrsim \lambda$, where g is the coupling of the "Higgs" field to the inflaton, λ is its self-coupling and v its expectation value after symmetry-breaking; C is a constant determined by numerical simulation. When $g^2 \ll \lambda$ the relevant formulae are

$$\begin{aligned} f_{\rm peak} &\simeq (3\times10^{10}\,{\rm Hz})\frac{g}{\sqrt{\lambda}}\lambda^{1/4}, \\ \Omega_{gw}(f_{\rm peak})h_{100}^2 &\simeq (8\times10^{-6})\left(\frac{v}{M_P}\right)^2\frac{\lambda}{g^2}. \end{aligned} \tag{55}$$

In Fig. 33 we plot the spectrum for the parameter values $\lambda = 2g^2 = 10^{-14}$, $v = 3\times10^{-7}M_P$.

The rapid decay of "flat directions" (scalar degrees of freedom) in supersymmetric models after inflation is a similar potential source of stochastic GW [212]. The characteristic momentum of fluctuations is of order the

SUSY-breaking mass scale $m \sim$TeV giving a present-day frequency of 10^2–10^3 Hz. We plot in Fig. 33 the spectra for two choices of parameter values. The lower curve has $m = 100\,\mathrm{GeV}$, reheating temperature $T_R = 10^8\,\mathrm{GeV}$ and initial field value $\Phi_i = 2 \times 10^{18}\,\mathrm{GeV}$; for the upper curve, $m = 1\,\mathrm{TeV}$, $T_R \geq 10^9\,\mathrm{GeV}$ and $\Phi_i = 10^{18}\,\mathrm{GeV}$.

Conclusion Many diverse and exciting phenomena in the physics and cosmology of the early Universe may be probed by ET via the stochastic GW background, either almost immediately if a signal is well above the detection threshold, or with an extended observation period. It may be possible to estimate the parameters (*e.g.* mass scales and couplings) of the new physics responsible: observational evidence for a cosmological phase transition, or of the temperature of reheating, would be an epoch-making result.

Author(s): S. Aoudia, L. Bosi, L. Storchi,
...

2.6 Computing:Status of Art and evolutive processes

Computing has made great strides in recent years in order to provide each year even more fast processors. Since the invention of the integrated circuit in 1958, the number of transistors on a given chip double (roughly) every two years, as states the first Moore's law [213]. This exponential growth has allowed computers to get both cheaper and more powerful at the same time. In order to mantain this trend for the next decade the next technological steps need some rethinking of basic foundations such as:

- Process technology
- Hardware Architecture
- Software Software

In this section we report about processors, computing infrastructure and architectures status of art, highligthing technical and physical problems that are leading major vendors to change Hardware and Software architectural paradigma.

2.6.1 Moore's law

"Moore's Law is a violation of Murphy's Law. Everything gets better and better" [**?**] this is how Gordon Moore commented the law, that bears his name, in 2005. Moore's law describes trend in the history of computing hardware. The law has been originally thought to describes the number of transistors that can be placed inexpensively on an integrated circuit. But now we see this law can be applied also to the capabilities of many digital electronic devices, such as memory capacity, sensors and even even the number and size of pixels in digital cameras. There are in fact many other laws related to the Moore's one. Other laws prediction for example hard disk storage cost per unit of information, or network capacity, or pixels per dollar, and more.

Gordon E. Moore formulated the law by a simple observation. In 1965 he noted that number of components in integrated circuits had doubled every two years from the invention of the integrated circuit in 1958. Thus, he predicted that the trend would continue "for at least ten years". Years after the law has been reformulated to take into account an higher growth, and the final formulation state that integrated circuits would double in performance every 18 months. Thus "Moore's first law predict" an exponential rates for the transistor counts in a microprocessor: $P_n = P_0 \cdot 2^n$, where P_n is the predicted computer processing power in future years, P_0 is the actual computer processing power, and n the number of years divided by the doubling period, expressed in years. i.e. if we consider the transistor count the doubling factor is 2 (every 2 years), while if we consider the processors speed the doubling factor is 1.5 (every 18 months).

Moore's first law can be viewed just as an observation, but maybe there is even more. Maybe behind it there is a more deeper law, a law driving evolution of information and technology, of which the Moore's law is just a

consequence. But up to now what it is clear is that this law has been widely accepted, and is used as a goal for both marketing and engineering departments of semiconductor manufacturers.

2.6.2 Main difficulites on computing power increase

Quoting Mr Moore recent statment issued during an interview "..by about 2020, my law would come up against a rather intractable stumbling block: the laws of physics.." [214] we may be led to think that the Moore's Law is at the end. Today the integration scale (the typical size for the CMOS realization) is about 32-22nm that is comparable to few hundreds of atomic radii. It is so clear that one of the main limitation on continuing the actual processors renewal strategies is posed by the atomic limit. The difficulty on even more reducing the integration scale was evident and concrete already during the last 10 years. In fact major manufacturers introduced several technological innovations and hardware paradigma in order to provide even more fast processor, limiting the integration scale reduction and CPU frequencies increase.

We report on Computing power increase issue, introducing briefly the top500 and the last 10 years technological innovations.

The Top500 project [215] goal is to generate, twice a year, a list of the 500 most powerful computers in the world. The list is manteined by Hans Meuer of the University of Mannheim, Erich Strohmaier and Horst Simon of Lawrence Berkeley National Laboratory, and Jack Dongarra of the University of Tennessee. Dongarra invented Linpack benchmark [216] many years ago, used to rank the system within the list.

The Top 500 ranking has always been a good overview of the actual technology trends, and along the last ten/twenty years the list has showed an evident trends toward parallel and massively parallel systems. This trend has been confirmed by the appearance in the home computing system of superscalar and pipilened CPU firstly, and multi-core CPU after.

The actual top500 list emphasizes the following evolutive characteristics:

- Natural evolution from multi-core to many-core era
- Many-core architecture (i.e. GPU) together with multi-core X64 processors are in current use
- Moore's Law is alive and still works well
- Development of faster and more integrated interconnects as an obvious consequence of the increase in number nodes and cores

To stress those aspects it is essentials to cite the first ranked system on November 2010: the Tianhe-1A. This supercomputer is stationed at the National Supercomputer Center in Tianjin, it has achieved a performance level of 2.57 petaflop/s. This system equipped with 29376 GB of memory, is based on Nvidia graphics processing units (GPUs) as well as Intel Xeon 5600 series processors. The system collects together both the trends, fast interconnection (i.e. the Tianhe-1A is able to handle data at about twice the speed of InfiniBand), and many-core (i.e. the system is capable of such performance essentially thanks to the the acceleration given by Nvidia Fermi GPU processors). It is clear that a new programming paradigms and new algorithms are needed to get top performance from these computing infrastructure.

Looking for high performances, one has also to deal with **power consumption**. The power consumption of an HPC resource is a fundamental aspect on computing facilities setup. This requires an optimization in terms of efficiency, cost and resources reliability. In order to stress how this is today an extremely important and sensible theme, the Top500 has started to collect data related to power consumption of the 500 most powerful computer in the world.

To achieve greater performance per compute node, vendors have increased not just transistors and speed but consequently also the power density. Thus, a large-scale high-performance computing system requires continual cooling in a large machine room, or even a new building in order to work properly. So to achieve greater performances one has to consider direct and indirect (i.e. cooling system) costs. We have to remember the HPC systems failure rate increase is directly related to the working temperature.

Many-core seems to provide good performances also in terms of power consumption. In fact making the assumtion that The Tianhe-1A 2.5 PFlops system was built entirely with CPUs, it would consume more than **12 MW**. Thanks to the use of GPUs in a heterogeneous computing environment, Tianhe-1A consumes only **4 MW**, making it **3 times more power efficient**.

It is importaint to compare the power consumption of Tianhe-1A and Jaguar Cray XT5-HE, that is the second ranked system in top 500. In fact the Jaguar use 5-10 MW to achieve 1.7 PFlops while Tianhe-1A use 4 MW to achieve 2.5 PFlops. **This is a success of the many-core architecture providing 3-4 time more computing power per Watt!**

This enforce the evidence that computing hardware architecture are moving toward the many-core era direction.

2.6.3 20 years of Parallelization

In this section we introduce more details about architecture innovation introduced along last 20 years by hardware manufacturers. We intend to underline as the implicit and explicit parallelization concept has been used as a way to go around the miniaturization limitations and frequency increase.

A scalar processor is the simplest CPU that can considered. Is is capable of executing a single instruction per clock cycle and to manipulate one or two data items at a time.

A superscalar processor is instead capable of intrinsic parallelism. Each instruction processes one data item, but multiple instructions and data are processed concurrently, having multiple redundant functional units within each CPU. In fact modern superscalar processors includes multiple ALUs, multiple FPUs. Thus the dispatcher of the CPU reads instructions from memory and decides which ones can be run in parallel. An important step forward has been the introduction (around 1998/1999) of one or more **SIMD** units by AMD and Intel. Those units are used through the AMD's 3DNow and the Intel's Streaming SIMD Extensions (SSE) instruction set extension to perform basic vector operations (i.e. adding two vectors of float, in one step).

The capability of executing more than one instruction per clock cycle is another level of parallelism introduced into superscalar CPU. The basic idea is to split each instruction into several micro-instructions, each executed by an indipendent functional unit of the pipeline.

This approach permits a natural parallelism. In fact usually when there are several instructions to be executed, as soon as the first functional unit has finished the execution of the first micro-instruction, this is sent to the second unit. So the first functional unit of the pipeline is free to start the execution of the second instruction, and so on. Given a starting latency to fill the pipeline, the CPU reach a steady state where N instructions are executed together for each clock cycle, where N is number of functional units (so called depth of the pipeline).

Another step on improving the efficiency of CPUs, has been the introduction of **Simultaneous multithreading (SMT)**, roughly about 2003-2004. Maybe one of the most famous implementation of this technique is the Intel's Hyper-Threading Technology. The HT, or HTT, works duplicating some sections of the processor pipeline. In this way the hyper-threading processor appears as two "logical" processors to the host operating system. This allows the operating system to schedule two threads or processes simultaneously.

Starting from 2005 multi-core CPU have been introduced in the everyday computing architecture, both in the embedded and standard systems. This solution implements multiprocessing in a single physical package, namely the full processor, replicating the whole computing core. Different cores may or not share caches, and may implement message passing or shared-memory inter-core communication. The actual multi-core CPU implements up to four/six cores per package. In case of the multi-core CPU the performance gain is strictly related to the efficiency of the parallelized software. About that we have cite Amdahl's law[**?**], that connect the parallelization efficiency with the fraction of the software that can be parallelized to run on multiple cores simultaneously.

2.6.4 Emerging technologies for distributed computing

In the this section we want to briefly review some new and emerging technologies like grid and cloud computing. We will try to briefly introduce booth the technologies, and explain similarities, differences, and issues to consider in grid and cloud computing.

As the complexity of a computational problem grows, it is crucial, in order to continue being able to solve it, to design a computational system flexible enough to grow comparably in complexity. Solutions can be worked out only by gathering sufficient computing power and the appropriate know-how. know-how and computing power must be collected and deployed in an efficient and appropriate manner, so that anyone is enabled to access transparently and simply any needed information and resources. What we have just pointed out is the general idea of the grid or, perhaps more precisely, of a computational grid environment. This new approach to network computing is known by several names such as: meta computing, scalable computing, global computing, internet computing, and, more recently, peer-to-peer computing.

One of the most famous definitions of the grid perfectly describes such design: "the grid is a flexible, secure, coordinated resource sharing among dynamic collections of individuals, institutions, and resources - what we refer to as virtual organizations." foster at al. [217]. more precisely the grid can be thought as a distributed system, where heterogeneous resources are geographically dispersed and connected by a network. so grids represent a form of distributed computing facilities, where a "super virtual computer" is composed by many interconnected computing resources, like clusters, single workstations and traditional single super computers. The middleware (i.e. a collection of software libraries), like the operating system in a pc, gives to the user all the necessary instruments to use the grid in a transparent and secure way, it provides uniformity through a standard set of interfaces to the underlying resources. It is a layer of software between the network and the applications that provides services such as identification, authentication, authorization, directories, and security.

A working example of what we briefly described is the Worldwide LHC Computing Grid project (WLCG) [218] currently operates the world's largest scientific Grid, with over 140 computing centers in 34 countries contributing resources, including more than 10,000 CPUs and several Petabytes of storage.

2.3: Worldwide LHC Computing Grid (WLCG)

WLCG is the computational solution adopted by the LHC to store and analyze the expected Petabytes per year of data generated by the four main LHC experiments (ALICE, ATLAS, CMS and LHCb). It consists of a hierarchical network of computational centers distributed in 33 countries. The zeroth level (TIER0) is the CERN Computer Centre. It is responsible for the safe-keeping of the raw data and its distribution to the first level sites (TIER1). These are the principal centers dedicated to store and analyze all the data uploaded by the LHCb detector. They are hosted by : Canada, France, Germany, Italy, Netherlands, Nordic countries, Spain, Taiwan, United Kingdom, USA. The second level sites (TIER2), which are more numerous but less important, store temporarily a subset of data loaded from the TIER1 sites. In addition to these computational centers, each collaborating institution could have its own cluster (TIER3) with regulated access [219, 220]. WLCG achieved in 2011 the following capacities [221] :

- 87,089 physical CPUs,
- 280,510 logical CPUs,
- 158,526,879 GB of total online storage capacity,
- 104,588,138 GB of total near-line storage capacity,
- more than 2,000 TFlops of processing power.

Cloud computing goes a step further in the direction of separating the user from the computing resources. This new computing paradigm represents also an improvement in the direction of the on-demand resource provisioning. Cloud computing generally means a collection of technologies enabling the final user to benefit of wide set of hardware and or software remotely distributed. We can compares the Cloud with the electric

power network: when we switch on a light, or we plug an electrical device into a wall socket, we are not aware from where the power comes from, and generally we do not care much about such details. Now we are thinking about a world in which computer power, storage and software capabilities are as easily accessible as the electric power.

Wen can try to briefly summarize the Cloud computing architecture as follow. The final user simply uses a specific service provided by a customer administrator. The administrator uses some interfaces to select a specific service (i.e. a virtual server or just some storage) and to administer the service (i.e. configure it, activate or deactivate the service, or maybe to ask for more computing power or storage). The service provider is the one that physically owns the real server, and it is the one in charge for providing some transparent interface to manage the resources.

There are up to now several examples of working cloud infrastructures, here we report just a short list:

- Amazon Elastic Compute Cloud (EC2), that allows users to rent and use virtual server for the scalable deployment of applications.
- Amazon S3 (Simple Storage Service), an online storage web service
- Google App Engine (GAE), a platform for developing and hosting web applications in Google-managed data centers

2.4: Einstein@Home

Einstein@Home [222] is a volunteer distributed computing program that uses more than half a million personal computer's idle[a] time (400 TFlops [223]) to search for gravitational waves from isolated pulsars or radio pulsars in binaries using data from the LIGO gravitational wave detector and the Arecibo Observatory in Puerto Rico [224]. The project is hosted by the University of Wisconsin-Milwaukee (USA) and the Max Planck Institute for Gravitational Physics (Albert Einstein Institute, Hanover, Germany).

[a] A computer processor is described as idle when it is not being used by any program.

2.6.5 Manycore

The transition to many-cores systems seems to be the natural evolution of computational architecture. Many-core processoris have a larger number of cores respect to traditional multi-processor, roughly in the range of several tens of cores. The actual state of art in many-core architecture is represented by GPU processors, where in a single package hundreds of computing cores are implemented.

In the next part of the section, we would like to provide some information about major vendors technological decisions and roadmaps in order to give the feeling about many-core technological trends.

All major CPU processors manufactures, such as Intel and AMD, are researching and developing new innovative solutions in order to bypass the even more stringent technological challenges. In our work we have collected several information about major manufacturer production roadmaps and production planning.

As been previously reported in some sense GPU are the precursor of many-core architecture with several already marketed and used hardware devices. Even if these have been developed specifically for computer graphics this hardware is now widely used in other computing fields, proving a resounding success.

Moreover during 2010, Intel and AMD have published and made official communication about subsequent CPU generation. Also if in different way and implementation they report technological solution that are following the path of increasing number of computing core elements per CPU.

In detail AMD has declared to be close to release a new processors family based on Fusion [225]. AMD Fusion is the codename for next-generation microprocessor design and a product merging together AMD and ATI. Where AMD brings knowledge about CPU technology and ATI its own knowledge about Graphic Progessing Units, combining general processor execution as well as 3D geometry processing and other functions of modern GPUs into a single package. The core of this new architecture are the APU (Accelerated Processing Units). This technology is expected to debut in the first half of 2011.

Intel has recently declared during the New Orleans Supercomputing Conference its approach to the High Performance Computing, introducing the MIC (Many Integrated Core) solution [226], known with the name Knights Ferry. The Intel MIC architecture is derived from several Intel projects, including "Larrabee" [227] and such Intel Labs research projects as the Single-chip Cloud Computer [228, 229]. The architecture is based on chip containing 32 cores x86 at 22nm. Moreover Intel has declared for 2012 the production of an higher solution based on 50 core 4 hyper-threading processor. Obviously this solution can be compared directly with GPGPU, having an equivalent high number of cores. One of the key point of the Intel solution is the code portability, being a x86 compatible architecture. Moreover comparing Knights Ferry with GPU solution we have to remark that a CPU core is much more complex than a GPU core, providing for example SSE4, permitting 8 single precision operations per cycle per core.

In this context it is interesting to report also other experiences, such as Tilera products [230]. It provided the first innovative many-core solution based on x86 architecture. For example the 64-core TILEPro64 integrates 64 identical cores, having a complete full-featured processor, including cache memories and more.

The previous statements indicate a clear direction about new processor products: **CPU are evolving toward the direction of the many-core computing**. Each vendor is traducing this concept on different shapes (i.e. "homogeneous collections of general-purpose cores" rather than "computers using a variety of specialty cores"), but all agree on the need of increase significantly the parallelization level [231].

A so deep changes in the hardware architecture will require also a deep changes on software side and about the way of thinking algorithms. Without this effort it is impossible to extract the real power of these new computing resources.

2.6.6 Manycore as real scenario for future computing infrastructure

In order to understand the real capabilities of these new architectures, in Perugia group we started to explore the status of art of manycore devices. In particular we perform some test with NVIDIA C1060 (i.e. GT200) and with NVIDIA C2050 (i.e. the brend new Fermi GPU).

Our tests are based on a fully Multi-GPU implementation of a Coalescing Binaries Detection pipeline. This software includes specifically an input data conditioning, signal Post Newtonian generator up to PN 3.5 [**?**] and a complete matched fitlering procedure with colored noise [**?**].

Respect to CPU implementation of the same algorithms, results show an average gain factor (normalized by price) of about 50, using a single C1060 GPU. This can be translated in a number of applyied matched filtering per seconds of about **30**. Obviously this number depends on vector size. This numbers are about length of 2^23 samples. Using shorter vector it is possible to achieve higher gains. Performing the same test with new NVIDIA Fermi GPU (i.e. the Tesla C2050), this number increase up to **120 templates per second**.

In Perugia we implemented also a Multi-GPU version of the pipeline, which gives another increasing factor of 3.5 using 4 GPU, bringing us to an impressive result of about **400 templates per second** processed [**?**].

Another interesting gain factor is about FFT algorithm. Several benchmarks reports a gain factor of 50-80 using GPU respect to CPU architecture. This value has been renormalized by device prices.

Thus, we can state that using the already available many-core thecnologies the gain factor respect to the standard single core architecture is, conservatively speaking, about 100. Obviously an exhaustive analysis of the gain factor needs to deal also with power consumption and costs.

Author(s): S. Aoudia, L. Bosi, L. Storchi,
...

2.7 Computing Challenges

In this section we try to use our trend analysis and our simulations in order to foresee for the next 12-14 years the computing solutions and power gain.

What about future trends? As we are writing, also within computer industry technology's road map "believes Moore's Law will continue to hold good through 2029", citing Pat Gelsinger, SVP and co-GM of Intel's Digital Enterprise Group (DEG) [232].

Although specialists agree that by 2019 the current strategy of ever-finer photolithography will probably have run its course, it is likely that some new type of technology will be needed to replace the current integrated-circuit production process, like new materials, optical or quantum computers.

In any case is not a simple exercise to translate transistor growth into practical CPU performances. This is particularly true for recent multi and many core systems. In this case it is always needed a great work on software side in order to take advantage of the modern multi and many core CPU. Often it is needed to substantially think back the software implementation.

There are other factors limiting the possibility of taking full advantage of the modern CPU, such as internal bandwidth and storage speed. In other words memory and disk access speeds have failed to keep up respect the CPU. In fact to reduce the impact due to these problems, different solutions been introduced in processor and software design, like: out-of-order execution, caching and prefetching strategies. This also means that there is still a big optimization margin on other computer components.

2.7.1 Future Possible Computing Scenarios

Next decades main key actions can be summarized as :

- Chip-Level Multiprocessing (CMP)- increasing parallelism for increased performance
- Special Purpose Hardware - Embed important functions relegated to software and specialized chips inside the microprocessor itself.
- Large Memory Subsystems - Memory access is a main bottleneck. In order to keep many high-performing cores, it is important to have a large quantity of memory on-chip and close to the cores.
- Microkernel - Microprocessors will need a sizable integrated intelligence, in order to coordinate all this complexity: assigning tasks to cores, powering up and powering down cores as needed
- Virtualization - Future microprocessors will need several levels of virtualization. virtualization is needed to hide the complexity of the hardware from the overlying software. The OS, kernel and software should not have to deal with the intricacies of many cores, specialized execution hardware, multiple caches, reconfiguration and so on.

We can assume that during next decades, these innovations will contribute on keep alive the Moore's law, that we can use to forecast the future computing power. On figure [**?**] we report the expected computing power gain respect to the epoch of evaluation. Thus, if we consider the actual computing power of a typical CPU, Moore's law predicts a gain factor of about 200 in 12 years and 700 in 14 years.

At least up today, this behavior is experimentally proved also by top500 performances graph shown in figure 35. As previously noted, the Linpack benchmark is used to rank the system within the list. Thus, the values reported in the graph can be considered a good estimation of the real power (i.e. sustained performances) of the 500 most powerful computer in the world.

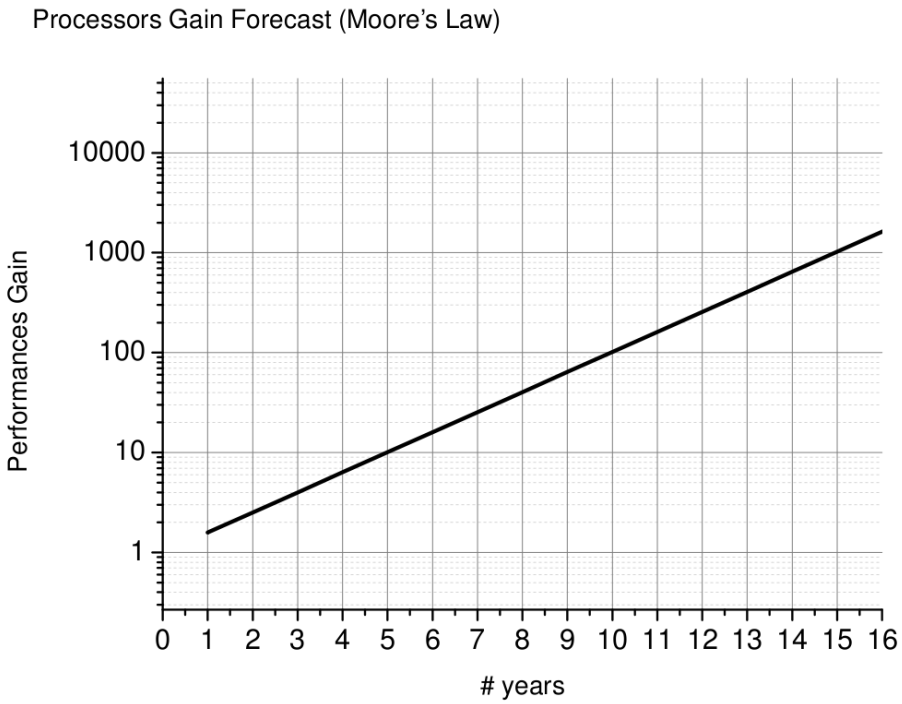


Figure 34: Moore's law projection of expected computing power gain respect to the epoch of evaluation.

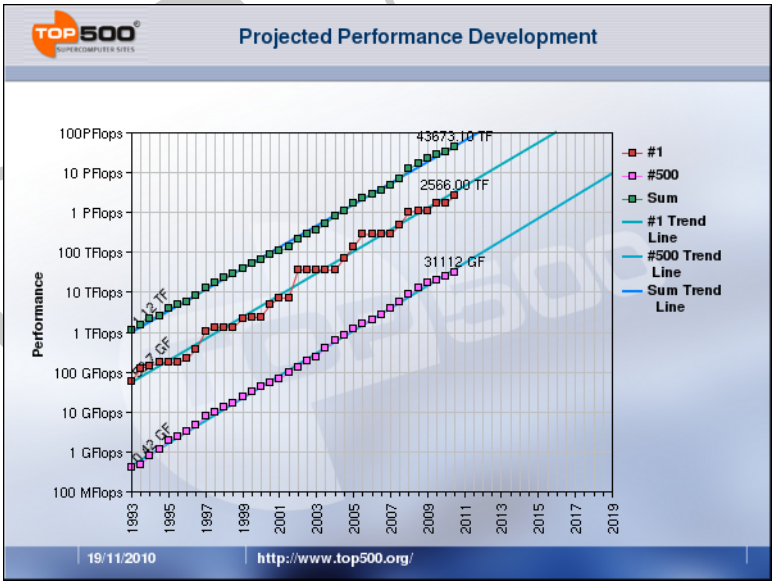


Figure 35: The Top 500 projected performance development.

In the figure 35 are reported three different sequences. The first one, labeled with **Sum**, represents the total amount of computing power of which all the 500 supercomputers in the list are capable. The second sequence, the one labeled with **#1**, is the power of the first ranked system. The last one, the sequence labeled with **#500**, is instead the computing power of the 500-th system in the list.

A simple execise to show to the reader how the first Moore's law works. If we consider the computing power of the last ranked system in the list ten year ago, this gives a number of approximately 100 GFlops. Today the 500-th system, namely a Xeon Cluster, has a power of more the 10 TFlops. This gives a gain factor of about 100, as the Moore's first law states. So, as it may seem incredible, experimental data confirms that in ten years we had and increase of a factor 100.

These information and in particular about GFlops are closely related with computational problem to be faced. In the ET case we can try to select some reference core algorithms and use the our many-core / Moore's law forecast factor to predict the realistic computing power in 10 years from now. These results can be used to understand the future investigation capabilities and limitation about some of the main ET data analysis area of interest.

A fundamental point that distinguish first and second generation gravitational wave detectors respect to ET science is the nature of the experiment. ET has to possess enough sensitivity and computing power to achieve its main goal, the GW observation and not detection. This means real-timand on-time analysis that will stress the computing infrastructure, requiering the handling of the two main data analysis aspects: detection and parameter recostruction problems.

2.7.2 Detection Stategies

Binary systems

As shown in the previous sections, binary systems, formed by pairs of neutron stars (NS) or an intermediate massive black hole (IMBH) with another NS or IMBH, are the most important sources for ET [233]. In fact, the detection and the study of IMBH is an astrophysical challenge in itself.

In the context of full General Relativity, the exact equation describing the gravitational waves emitted by the motion of two compact objects is not really known. The only solution for that problem comes from the Post-Newtonian theory. In this context, a significant progress had been made in the last few years and now we are able to have waveforms at the 3.5PN order in phase, 2.5PN order in spin-orbit (SO) effect and 2PN order for the spin-spin (SS) effects [234].

Two sets of parameters characterize waveforms emitted by quasi-cicular coalescing binaries : intrinsic and extrinsic ones. The extrinsic or kinematic parameters are all the parameters incorporated in the detection statistic. In our case, they are defined by the inclination i, the polarisation ψ of the source, the fiducial time of coalescence of the binary t_c, and phase at coalescence Φ_c. The intrinsic or dynamic parameters are the reduced mass η, the chirp mass M_c, the algebraic values of the spins χ_1, χ_2 and the sky location of the source θ, ϕ (here we consider the algebraic values of the spins since we will limit ourselves to parallel and ant-parallel spins only without including the time evolution of spin directions which will be considered as irrelevant for the purpose of our study).

In addition of these parameters, an other effect which will change the shape of the detected waveform is the Doppler shift effect of the frequency. The detector's motion can no longer be considered as negligible since the observation time could reach few days. This also suggests that a single ET detector could be able to determine the sky location of the source (θ and ϕ) even with a moderate precision.

As the model waveform depends on a number of parameters, the most adequate way to search for the signal is to filter the data through a template bank covering the astrophysically interesting region of the parameter space[235, 236]. It represents a grid of theoretical waveforms (templates) placed, according to a fixed value of the minimal match MM, on the parameter space. Each point is associated with a template built from a specific values of the parameters. By matched filtering technique, one finds the maximal value of the likelihood which corresponds to the best mimicking template.

The problem with such technique of search is the large number of templates needed. For an ET-D configuration with a minimum match MM of 95% and for a total mass M range $[1M_o \;\; 300M_o]$, the number of templates roughly ranges from 10^9 (for templates built using only two parameters : η, M_c) and 10^{18} (for spinning binaries templates wrought by six parametrs : two masses, two angles and two algebraic values of spin), respectively.

In the Table-??, we summarized the computational time needed for flat search with different number of parameters.

Computers		GPU C2050	E@H	WLCG	Tianhe-1A	GPUs C2050 @ 2023
Processing power (TFlops DP)		0.5	200	2000	2567	100
Nbr of para	Nbr of templ	Computational time needed				
2	10^9	64 days	4 hrs	24 min	18 min	7 hrs
4	10^{13}	1760 yrs	5 yrs	182 days	128 days	9 yrs
6	10^{18}	179 10^6 yrs	5 10^5 yrs	5 10^4 yrs	4 10^4 yrs	9 10^5 yrs

Table 3: This table gives the computational time needed for a flat search with different number of parameters using : 1 GPU C2050, E@H??, WLCG??, the supercomputer Tianhe-1A and using the future performance of a single GPUs C2050 at 2023, respectively. The results show that using matched filtering search with six parameters template bank or more is not a realistic choice.

For low SNRs, one could suggest a hierarchical search, looking for a detection at first with a coarse grid and reconstructing the parameters around the detected template with a fine grid template bank in the second pass. However, this could be problematic if we have great SNRs (as in the case of ET). Many templates could have the same maximum value of the likelihood even with disparate values of the parameters (multi-sources, degenerescence). This suggests to use a large parameter space ranges for the second step of search giving a large number of templates, or/and to use a multi-templates search technique (overlap of many sources) where also a huge number of templates is needed.

The large SNR for ET and the relatively great number of overlapping expected sources make the most used search techniques, in the case of ground-based detectors, obsolete when the number of parameters $N > 4$). Thus, may be the most promising outcome is to adapt the algorithms used for the future space detector LISA such as MCMC, MHMC or Generics Algorithms (GA) to our purpose [237].

Another possible approach can be a two steps multi-band analysis. In this case the analysis is made first performing a detection phase with templates starting from higher cut-off frequency and assuming a small snr-lost, .e.g. of about 5%. In fact the template length is reduced exponentially starting from highest frequency. Starting for example from 10Hz we loose few percent in terms of signal-to-noise but we gain propotionally in terms of computing time. Thus, the first step can be considered similarly to a trigger zero, used for candidates selection. Then, in the second step, the whole template is analyzed for parameters reconstruction and observationi, using a finer template bank grid. Another possible optimization could be achieved using the Stationary Phase Approximation for template production of phase 1 and than the PN approximation or others during the phase2. This schema can be used also for other kind of generators.

Stochastic gravitational wave background

The stochastic gravitational wave background could be of cosmological (isotropic - like cosmic microwave background) or of astrophysical nature (anisotropic following the spatial distribution of the sources).

It could be interpreted as the superposition of a number of plane waves having a frequency f and a certain direction Ω. We assume that the stochastic background is unpolarized, Gaussian, stationary but we allow an anisotropy in its distribution. The anisotropy of the signal will affect the statistical properties of data with the presence of a signature in the output of the detector.

The search methods for each nature of sources should depend on its angular distribution and its spectral properties.

The search for isotropic sources, called **isotropic analysis**, is based on the use of a network of detectors [238, 239] (the radiometer analysis is almost based on the same framework [240]). For each detectors pair, first, one searches the times when both the detectors were taking science data, then split that time up into segments

of length T (usually T 60 seconds). Let's label the interferometer pair with an index I=1, 2, ... , N, and for that pair, each segment is labelled by an index J = 1, 2, ... $n_{\rm segs}$. So that, each of the segments analysed can be identified by the unique (I, J) pair. Since the segments will be Hann-windowed, only segments that overlap by 50% are used.

The bulk of the processing time is taken up by reading the data and finding the frequency dependent cross-correlation spectrum (the integrand of the cross-correlation estimator statistic, Y_{IJ}), and sensitivity integrand (the integrand of the theoretical error bar, σ_{IJ}) in each unique segment. The steps are:

- read the data for the segment for each IFO,
- downsample (to 2048 or 4096 Hz),
- high-pass filter to remove low-frequency components,
- put the data through a standard PSD estimation function,
- multiply the data with a Hann window,
- Fourier Transform the data,
- decimate the fourier transform to the same frequency resolution as the PSD (usually 0.25 Hz),
- calculate the overlap reduction function at the same resolution,
- calculate the optimal filter function,
- calculate the cross-correlation spectrum and sensitivity integrand,
- write some or all of these quantities to file.

Once this has been done one get the cross-correlation spectrum and sensitivity integrand for each unique segment given by the indexes (I,J). Then, the 'post-processing' could be started. The last is often repeated several times, without having to re-run the calculation of the spectra.

For each unique segment (I, J), the cross-correlation statistic, Y_{IJ} and theoretical error bar, σ_{IJ}, are calculated.

For each pair, one finds the weighted sum of the cross-correlation statistics and theoretical error bars to give combined results Y_I and σ_I.

One then finds a weighted sum of the results from the individual pairs, to give an overall statistic Y and error bar σ. (Note, that normalization is done such that $Y = \Omega T$, where Ω is the amplitude of the GW spectrum [238]:

$$\Omega = \frac{f}{\rho_c}\frac{\mathrm{d}\rho}{\mathrm{d}f} \tag{56}$$

where $\mathrm{d}\rho$ is the energy density of gravitational radiation contained in the frequency range f to $f + \mathrm{d}f$, ρ_c is the critical energy density of the Universe, and f is frequency.

The final step is to use this statistic to find a Bayesian posterior PDF on the GW amplitude Ω, since this is a one-parameter problem (One marginalizes over any other unknown parameters, such as the calibration factors of the IFOs), one can simply calculate the posterior on a grid

Parallelization : One can split the data from each IFO pair into 'jobs' which contain several segments (anything up to a few hundred segments). Then, run the cross-correlation estimation (the items bit) for all the segments in one job on a single node, and these jobs are all run at the same time. So, the analysis which would take 300 hours of CPU time will usually complete within a few hours on the caltech cluster or on atlas. Once all the jobs are completed, they carry out the post-processing (very cheap computationally).

- forecast: Stoc. problem size and Computing implications - for 300h x 2months 2pear of interferometers –> ET –> Isotropic/Radiomether –> 30Hour –> 1 Months –> IO Limit - 1h CPU time X 1 Day Knowing the actual

detection algorithm parameterization, it is possible to estimate the computational cost and foresee the future capabilities give by computing infrastructure improvements. In particular we can evaluate the cross correlation problem. In this case we know that today the typical analysis is made in the frequency band 600-10kHz with 0.25 Hz of frequency resolution. With this information we can deduce the need to use a 20kHz as sample frequency and in order to achieve a frequency resolution of 0.25 [Hz] we need at least a data vector length of 1/0,25 s = 4 s. With these information we can estimate the typical data vector length equal to L= 4 x 20000 samples = 80000 samples.

If we consider the cross correlation algorithm and 1 Day of data, this computational problem is equal to perform roughly 43000 analysis cycle x cross correlated channels pair. The total ammount of processed data is 43000 x 8 * 80000 = 26GB per channels pair. We know that for 1 Day of data we need 1 h of CPU time, thus the estimated IO throughput is about 10MB/s per channels pair.

Moreover, We can try to estimate the GFlops considering a complexity manly due to FFT algorithm. The FFT complexity is $N\dot{l}og_2(N)$, under these hypothesis during a day we have a total value: $Nc\,N\,log_2(N)$. Considering as FFT reference library the fftw [241],it is possible to show that this algorithm expresses roghly 6 GFlops for vector data of 80000 sample.

Now in order to understand if this is a CPU or IO bounded problem we can test the two hypotesis. Under the hypotesis of a CPU bounded problem, we know from the previous evaluation made using real information that now we achieve 1MFlops x pairs channels. This data is 3 order of magnitude smaller than real one.

Usually the number of cross correlated channels is of the order of some tens, by that seem clear that this kind of analysis seems to be characterized by limitation due to data Input/Output from/to storage. Moreover the CPU power seems to be already enough to process data "in-time". A faster analysis could be realizing High throughput SAN (Storage Attached Network)

The sky decomposition method is used for anisotropic background, the method provides maximum likelihood estimates of the gravitational wave distribution $P(\Omega)$ decompesed with respect to some set of basis functions on the sky [242]. This basis could be pixel basis or spherical harmonic basis defined with respect to the Earth's rotational axis.

For point sources the best choice is the pixel basis. However, for a diffuse background dominated by dipolar or quadrupolar distributions the best choice is the spherical harmonic basis. Note that for spherical harmonic decomposition and in the case of an isotropic signal, only the monopole moment contributes to the stochastic gravitational wave background strength function Ω_{GW}.

The likelihood estimator for anisotropic background gravitational wave signal is the quantity :

$$P_\alpha = (\Gamma_{\alpha\beta})^{-1} X_\beta \tag{57}$$

X_β is the components of the dirty map and $(\Gamma_{\alpha\beta})^{-1}$ are the components of the inverse of the beam pattern matrix $\Gamma_{\alpha\beta}$. It plies a role similar to that of the Fisher matrix. For spherical harmonic decomposition, the index α corresponds to the couple l, m.

$$X_\beta = \sum_t \sum_f \gamma^*_\beta(f,t) \frac{\tilde{H}(f)}{P_1(f,t)P_2(f,t)} C(f,t) \tag{58}$$

$$\Gamma_{\alpha\beta} = \sum_t \sum_f \gamma^*_\alpha(f,t) \frac{\tilde{H}(f)}{P_1(f,t)P_2(f,t)} \gamma_\beta(f,t) \tag{59}$$

The different functions appearing in the above formulas are defined and computed as following [242]:

- The input $\tilde{H}(f) = (f/f_R)^\beta$ is the assumed spectral shape of the gravitational wave background. It is fixed at the start of the analysis by initializing the value of the reference frequency f_R (the most sensitive

frequency) of the detector and the spectral index β for the power-law behavior of the gravitational wave spectrum (note that $\beta = 0$ corresponds to a constant gravitational strain power.

- $C(f,t)$ is the cross-spectra of the data calculated as the SFT of the time-series output data of the two detectors :
  - down-sample output data to few kHz.
  - high-pass above few Hz to reduce contamination from seismic noise.
  - windowed.
  - discrete Fourier transformed to frequency domain.

Note that the frequency resolution (0.01 Hz) for $C(f,t)$ is much smaller than that one for $\tilde{H}(f,t)$ or for the spectral density function $P_i(f,t)$ $(i = 1,\ 2)$ and $\gamma_\alpha(f,t)$ which is fixed to 0.25 Hz. This is used to avoid unnecessary frequency resolution for these last functions, by averaging several frequency bins of $C(f,t)$ before to match Δf.

- $P_i(f,t)$ $(i = 1, 2)$ are the power spectra associated with the individual detectors computed according to Welch's modified periodogram using segments of 4-sec long (0.25 Hz) and 50% overlapping.
- $\gamma_\alpha(f,t)$ are the components of the overlap factor $\gamma(\Omega, f, t)$. they encode informations about the relative separation and orientation of the detectors as expressed by their beam pattern functions. They are analytically computed in the case of spherical harmonics decomposition.
- Choose a cutoff for the SVD regularization (fixe minimum eigenvalues for which the Fisher matrix is reversible).
- Choose $l_{\max}$. It consists of tradeoff between increasing the number of parameters to fit the data and minimizing the uncertainties.

The spherical harmonic analysis method can successfully recover simulated signals injected into simulated noise for several different types of stochastic gravitational wave backgrounds, for examples, isotropic sources, dipole sources, point sources, diffuse sources, etc.

However this method is more computationally intensive than isotropic and radiometer methdos. The computation time will be proportional to some function of the maximum value $l_{\max}$, but roughly estimates show that the spherical harmonic search takes 10 times as much CPU time as the isotropic/radiometer searches.

Continuos Waves Analysis

We clasify as continuous waves (CW) a gravitational wave signal with duration longer than the typical observation time of a detector. The signal of these sources could be affected by a various processes. (i) By spin-down : the rotation frequency and then the emitted signal frequency slowly decreases due to the energy loss of the source (EM, GW...). (ii) The signal is frequency modulated by the Doppler effect associated to the relative source-detector motion. (iii) Signal is amplitude modulated due to the non-uniform antenna sensitivity pattern of the detector. (iv) There are also two further complications that can affect the GW signal from real sources : glitches which are brief spin-up due to star-quakes or interactions between crust and superfluid and timing noise which is random fluctuations of the rotation frequency or phase.

The analysis methods for CW can be divided among coherent, i.e. that take into account the signal phase when known (matched filtering, cross-correlation, MCMC) and incoherent, where the signal phase is discarded (periodogram, Hough transform). Typically, incoherent methods are more robust and less computationally demanding even less sensitive.

The choice of a method (or combination of methods) depends closely on the prior information we have on the signal we are searching and on the available computing power. We can distinguish between two kinds of cases :

- Targeted search : cases where knowledges about the source parameters (position, frequency, spin-down) are available from electromagnetic observations of the source.

- Blind search : cases where all the source parameters are given within large intervals based on the available theoretical understanding of the astrophysical scenarios.

The targeted search could be schematized as following : (1) Extract the band of interest around the emission frequency : due to the Doppler effect, the amplitude modulation and the spin-down, the received GW is no longer monochromatic and covers a small frequency band (fraction of Hz) around the known emission frequency. (2) Correct Doppler and spin-down effect : if the spin-down parameters of a source are known with high accuracy we can remove both the Doppler and spin-down by multiplying data by an appropriate function. (3) Matched filtering : apply matched filter of the new slowly varying signal (demodulated signal) over the unknown nuisance parameters (amplitude of signal as received on Earth, polarization and inclination angles, initial phase...) and over the uncertainty range of the known parameters (position angles, frequency, spin-down). An alternative to explicitly apply a matched filter for the nuisance parameters is the use of the so-called F-statistics [243]. (4) Select maximum of the values of the unknown parameters at filter output. (5) Compare it with noise distribution and claim detection or set an upper limit.

The computational cost for targeted searches relies on the accurate knowledge of the source parameters: position, frequency, spin-down. Considering that the frequency is very accurately known within a small bandwidth, one can substantially down-sample the data stream around the expected source frequency and the data analysis cost for such sources is therefore minimal (few templates). However, even for sources observed in the electromagnetic domain, these parameters are known with finite accuracy and this can lead to a loss of sensitivity, especially for very long observation times. This will also increase the computational cost.

For NSs about which no informations about the spin-domw parameters is known, the computational cost is more important because, in this case, it is not possible to set an upper-limit on the signal strain. The number of templates is in the range $10^2 - 10^{10}$, for a year's worth of integration [244]. Therefore, even if we perform the search in a narrow frequency band around twice the radio pulsar frequency, the computational task is tricky. It is also obvious that the lack of information about the higher spin-down parameters will highly increase the computational cost, since one would also have to scan the parameter space of all these spin-down parameters. It is clear that more accurate data from radio observations are necessary in order to narrow down the uncertainty in the source parameters for making such searches feasible.

These results suggest that applying filtering method for a blind search is computationally not possible, since the minimal range for the source parameters is important (whole sky for angular position, whole bandwidth for the signal's frequency[4]...). Even for the more optimistic cases, the computational power required for such search is far from our actual computational capabilities (10^17 GFlops for circular orbit binaries) . A different approach with even a small loss of sensitivity but which ensures to reduce significantly the computational power is then needed.

One of the solution is the so-called hierarchical search (**stack-slide search** [245]) where coherent and incoherent steps are alternated. In the incoherent step a rough exploration of the parameter space using short data segments is done with a low threshold which allows for many false alarms (some candidates are selected). In the coherent step each candidate is followed with a more refined search using long data segments but searching the parameter space only in the vicinity of the candidate detections of the first step.

Both search steps considered above share the following scheme [246]:

- The output data are divided into short segments, called stacks.
- Each segment is phase corrected using an appropriate mesh of correction points to confine a putative signal in one frequency bin in each stack.
- Fourier Transform the phase corrected stacks
- Correct the individual power spectra for residual frequency drift using a finer mesh which allows to remove phase modulations over the entire data stream.
- Sum the corrected power spectra.

[4]The computational power is proportional to the maximum frequency of the detector to the fourth power.[244]

- Search for candidates exceeding some fixed thresholds.

i

From the computational point of view, the interesting feature of stack-slide method is the fact that it considers the available computational power as an initial parameter which fixes the values of the other parameters of the scheme. In fact, before the search begins, one has to specify the size of the parameter space to be searched (choose maximum frequency, region of the sky...), the available computational power to do the data analysis, and an acceptable false alarm probability. From these, one can fixe optimal values for the maximum mismatch for a patch, the number and the length of stacks. Optimization consists of maximizing the sensitivity function over these parameters given the definition of the total computational power as constraint.

The advantages of a hierarchical search are, from one hand the fact that the low threshold on the first step allows detection of low-amplitude signals which would otherwise be rejected, and from the other hand the second step can search longer data stretches on a limited computing budget, because of the reduced parameter space being searched, thus excluding false positives from the first pass. For given computational resources, this technique achieves the best sensitivity if the thresholds and mesh points are optimally chosen between the first and second step of search.

Wide-area searches for continuous gravitational signals emitted by rotating neutron stars is an hard task and it cannot be addressed using optimal data analysis methods, due to the huge computational power needed. A possible alternative approach is e.g. proposed by cite(PHYSICAL REVIEW D, VOLUME 70, 082001)(Quantum Grav. 22 (2005) S1255.S1264), where thay developed a hierarchical method allowing a cut of data analysis computational requirement. This method is sub-optimal and the processing gain is paied by a small reduction in sensitivity. The method consists mainly on alternating coherent steps,based on FFT, and incoherent steps based on the Hough transform. The Hough transform is a feature extraction technique used typicaly in image analysis, computer vision, and digital image processing. The purpose of the technique is to find imperfect instances of objects within a certain class of shapes by a voting procedure. This voting procedure is carried out in a parameter space, from which object candidates are obtained as local maxima in a so-called accumulator space that is explicitly constructed by the algorithm for computing the Hough transform. In paper (Quantum Grav. 22 (2005) S1255.S1264) the computational requirements for a search over these many templates is also estimated. Analyzing the data in roughly real time requiresa computational power 10^8GFlops, moreover is reported the performances of the fastest supercomputers ca. 2010, 510^4GFlops.

Considering a foreseen factor of 200 we can expect in 2023 that the fastest super computer is capable of 10^7GFlops. This number is still an order of magnitude lower and we can conlude that with type of analysis the online will be still not achievable.

Unmodelled sources

[R-Modes] After the supernova event, the newborn neutron star may spin down during up to one year due to R-modes instability (first investigated by Owen at al.[1998]) [**?**]. R-modes are non-radial pulsation modes of rotating stars that have the Coriolis force as restoring force with a characteristic frequency comparable to star rotation speed. These modes are driven unstable by gravitational radiation, inducing a differential rotation at second order in mode's amplitude, which leads the non linear evolution of the r-mode instability. This process generates Gravitational waves that are very difficult to be observed.

r-modes gravitational waves are interesting for ET Science, because could provide very foundamentals information about Formation Processes, such as initial condition, Neutron Star model and Equation of State and others like Neutrino emission models confirmations. In some sense R-mode gravitational wave observation can produce important and unique correlation with the nuclear physics of the neutron star and formation processes.

Using information reported in (Sï¿½ and Tomï¿½ model - Astrophys Space Sci (2007) 308: 557.561) we can acquire some characterizing data about this type of signals. The frequency domain of these waves is related to the angular velocity W by: $f = 2W/(3p)$. The frequency bound is given

- fmax ï¿½ 1200 Hz, it depends on the initial value of the angular velocity $W0$

- fmin ï¿½ (77-80) Hz, it depends on the final value of the angular velocity $W(tf)$ and K
- The GW duration is roughly $tf = (3.67.1)10^6 s$

In ((Sï¿½ and Tomï¿½ model)) is also reported the Gravitational waves strain h(f) generated by R-mode: (formula)

We estimate the optimal signal-to-noise ratio for Einstein Telescope case, and it is given by the formula:

$$\frac{S}{N} = \frac{200}{\sqrt{2+K}} \frac{20Mpc}{D}$$

Given these results, we can make hypotesis about the ET horizon respect to rmodes gravitational waves. On douing that we consider the best and pessimistic cases realted to a very high K. Thus, if a NS borns with substantial differential rotation of $K = (10^5 - 10^6)$ the ovserved optimal SNR is about $[4 - 12]@1Mpc$. While if $K = 0$ we achieve $120Mpc$ close to GA.

Burst Analysis Currently, there are two successful methods for searching burst events: Omega and coherent wave burst (cWB). With cWB, the basic idea is to combine coherently time-frequency (t-f) maps from several detectors [247]. The sky position is encoded in the time delays between interferometers. If there was no time delay, then excess of power in some pixels in t-f map should be similar (modulo the antenna pattern). So one can determine the sky location of the signal from the antenna pattern and by doing multiple time shifts of the data corresponding to different sky positions.

The number of detectors does not affect the computational cost since we do not analyze data from each detector separately but generate triggers on combined data stream. However, if we do joint analysis with LIGO and Virgo detectors, there is no point to combine the data if the sensitivity is very uneven (the weakest detector will bring noise to the analysis). In this case, it only makes sense to use the bandwidth where all detectors are reasonably sensitive. This would probably reduce the bandwidth of ET. At least, we can subdivide bandwidth into regions with comparable sensitivities and analyze each separately. If ET sees something but LIGO and Virgo do not, it would effectively be 1-detector analysis with no way to find the sky coordinates.

Recent use of that method in case of signal given by noise and glitches only (no gravitational waves), the computational cost is trivial. To process two years of S5 for 3 detectors with triple coincidence duty cycle of 40%, it would just take one week on Atlas cluster (80 TFlops[5]) if it is not too loaded or broken. Also, during S6, burst search was run online for H1,L1,V1 producing coherent triggers about 5 minutes behind the real time doing computation on 1 dedicated computer. However, in the ET era there might be such a rate of events that it becomes much more computationally intensive. Also if each event is long duration or wide bandwidth, it might consists of millions pixels. Processing such huge clusters might be quite computationally intensive, however still feasible with the future computational configurations.

Finally, one can emphasis that the computational cost for cWB search depends on which detectors, what frequency band one want to explore and based on this one could estimate how many sky locations he can/should consider. But practically, the most important factor is the large false alarm rate that could not be predicted before the detector is switched on. If some detector is glitchy, that might increase computational cost a lot since one would have to consider a lot of huge glitches before finding that they are inconsistent in some way.

Note that for these kinds of search only sky location and the gravitational strength of the gravitational burst could be recovered, however, the likelihood method used in the search offers a convenient framework for introduction of constraints arising from the source models [247]. In fact one of the constraints is related to the different polarization states of the GW signals. For example, linear or random polarization is predicted for some of core collapse models [248]. Merging compact binaries produce, in principal, elliptically polarized gravitational wave signals [249]. Also relativistic jets are emitted along the rotational axis of binary systems during the merger of neutron star, and where circular polarization of the gravitational waves is expected [250]. The cWB algorithm allows searches with several types of the polarization constraints : circular, linear, elliptical and

[5]Atlas is composed by 1680 2.4GHz 64bit quad core CPUs which sums up to 64TFlops and 132 Tesla graphic cards which gives us a theoretical peak performance of 132TFlops. The real power processing of Atlas is less than the half of its peak value.

random, or un-modeled search[247]. Moreover, joining these results with neutrino observations will probably give a powerful tool to probe the physical scenarios behind such kind of sources.

For **modeled bursts** like GRBs from both core collapse, supernovae and some of mergers, the search is based on a matched filtering variant so-called time sliced matched filtering [251]. The application of matched filtering depends crucially on phase coherence in the true signal, in correlating it to a model template. To circumvent phase incoherence on long time scales created by turbulence in the inner disk or torus of magneto-hydrodynamical systems powered by Kerr black hole, the template is sliced into several segments on intermediate time scales, for which phase-coherence may be sustained. Matched filtering is then applied using each slice, by correlating each template slices with the detector output with arbitrary offset in time. The relevant parameters in this algorithm are three: a choice of coherence time which can be course grained, fine grained time of onset(s) of the (slices of the) burst and the mass of the black hole. The mass is equivalent, for all practical purposes, to the duration of the burst. Changing the black hole mass changes the duration and also the strength of the signal, but the latter is automatically absorbed by the algorithm and requires no adjustments. For this kind of search, a computational cost can be extrapolated. For a single black hole mass parameter, it is about 6 hours of CPU time (core 2 duo) per 1 hour of detector data (sampled at 20 kHz). For a range of black hole parameters with, e.g., 0.1% partition, a full parameter search amounts to a few thousand hours per 1 hour of detector data. In this case we can expect for about 2013 a gain factor of 200 considering the same architecture. Implementing Many-core tecnology factor we can consider a conservative 6 extra-factor (1200). The full parameter search will be roughly 10 times far from the "in-time" constrain using a "CPU" model, while trying a many-core implementation few hours of data will be analyzed in one hour. Thus for about 2014 the "in-time" analysis of full parameter search could be reasonable.

[**Reconstruction**] The genetic algorithm (GA) is a search technique that mimics the process of natural evolution based on the positive mutation principle. Initially, a group of organisms (templates) is chosen, each organism is characterized by a different set of genes (parameters). Then, the quality (log of the likelihood) of each organism is evaluated. Based on the value of their quality (templates with higher likelihood), a set of pairs (parents) are selected and their genotypes are combined to produce children (new templates). In the last step, one could, with a controlled probability, allow random mutations in the children's genes (by randomly changing the parameters of the new template to explore a large area of the parameter space). The new children will form the new parent and the procedure is repeated until one reaches steady state (maximum in the log likelihood). From the computational point of view and for the case of coalescing binaries, one could be able to recover the values of the coalescing time and the chirp mass in 1 hour of time with 1 CPU. For the other parameters except the direction of the spins 10 hours of time is needed with 100 CPU. An all parameters recovery will need few days of time with 100 CPU. The cost is purely dominated by the computation (time domain + FFT) of the templates (1 to 2 sec with 1 CPU for a waveform lasting 2 years).

2.7.3 Conclusions

The most plausible computing scenario of the near future is a combination of CPU and GPU technologies, an evolution of most recent AMD's APU or Intel's Knight ferry technologies. We already showed that it is possible to exploit GPU power in Coalescing Binaries Detection. That means that manycore programming will not be a choice, because will be the status of art at the ET era. This consideration has permitted to add and extra gain factor to the Moore's law expected performances. In general we can consider a base gain factor of 200 by 2023 due to Moore's law and a manycore extra factor that can oscillate, depending on the analysis that can be posed conservatively to 10. This means a potential gain of more than 3 orders of magnitudes.

Given the computing power forecast, we believe that, if we discard a flat search approach with more than 4 parameters and the search for CW with great number of spin-down parameters (where more sophisticated improvements should be done), the combination of existing search methods and the future improvement of computing resources will allow ET to fully take advantage of its technological design and to play its real role as telescope giving the scientist community the expected information to explore the universe with new eyes.

We should also mention emerging technologies for distributed computing, such as: Grid and Cloud comput-

ing facilities. Here the Seti@H, E@H, LHC@home experiments are a great example. These will follow the evolutionary trend and will provide an important CPU power containers for off-line analysis.

We would like to conclude remarking that, technological breakthroughs are taking place. The computing infrastructure trend is toward manycore solutions, as show by top500 and manufacturers road-maps. The possibility to address future computing power for ET science needs will be proportional on how we will be able to use these new architectures. The doubling of performances each 18 months is no more for free. A change on programming paradigm and coding will be mandatory.

3 Site and infrastructures

3.1 Description

Responsible: WP1 coordinator, J.v.d.Brand

Interferometric gravitational wave detectors are large and complex and the selection of their site is an issue of great importance. The selected site should allow the highest possible level of scientific productivity at reasonable cost of construction and operation, and at minimal risk. Of paramount importance are the selection criteria that impact the scientific potential of ET. These include natural and anthropogenically generated seismicity and site geological constrains that affect critical parameters such as interferometer arm lengths. The first section of this chapter provides the requirements for the site and infrastructure of Einstein Telescope. An important aspect within these requirements is the allowed seismic motion, which is addressed according to source frequency. Seismic sources include the ambient seismic background, microseisms, meteorologically generated seismic noise, and cultural seismicity from anthropogenic activity.

The second section describes the background on one of the fundamental infrastructure limitations called Newtonian noise. Newtonian noise originates from fluctuations in the surround geologic and atmospheric density, causing a variation in the Newtonian gravitational field. New analytical formalisms and finite element models are presented for subterranean detectors giving an estimate for NN in ET for different geologies. Using these models we show that it is in principle possible to deploy seismic sensor arrays that monitor seismic displacements and filter the detector data using Wiener or Kalman filters.

The third section discusses the ET site selection. As part of the site selection and infra-structure program for ET, 11 European sites were systemically characterised to catalogue regions within Europe that would comply with the ET site demands. Among the data that were logged were the local seismic activity, already existing local infra-structure, and population density.

Finally, we discuss the subterranean infrastructure and cost aspects for caverns, tunnels, vacuum, and cryogenics for the ET site.

3.2 Executive Summary

Responsible: WP1 coordinator, J.v.d.Brand

Throughout the ET design study, the site selection and infrastructure design working group has seen an opulent evolution from a site study at 11 locations in 9 different countries, to a completely designed underground detector infrastructure. The site study has revealed several promising underground sites that comply with ET low seismic background performance requirements. In order to ascertain all site characterisation procedures were carried according to seismologic standards, measurements and data collection was carried out in collaboration with the Observatories and Research Facilities for European Seismology (ORFEUS) which is maintained by seismology department of the Dutch bureau of Meteorology.

The infrastructure definition contained several aspects such as the vacuum envelope, cryogenic infrastructure, tunnels, and caverns that are transverse through the detector optical configuration and suspension working groups. With the choice of combining a triple triangular detector with a xylophone detector topology, ET amalgamates an optimised planned disbursement for construction with a realistic proposal for a robust, highly sensitive, wide-band gravitational wave observatory. Through the design process, the vacuum system, caverns, and tunnel diameter are optimised such that the total infrastructure can be used for decades after its construction. The excavation of underground tunnels, caverns, and halls will occur over a period of approximately two years, where the ET observatory will be built in stages; the first construction stage containing a single 10 km xylophone detector. Later the second stage will incorporate the second and third 10 km xylophone detector.