

A Systems Framework for Multi-Messenger Astronomy

by

Alexander P. Koenig

S.B., Massachusetts Institute of Technology (2022)

Submitted to the Department of Aeronautics and Astronautics
in partial fulfillment of the requirements for the degree of
Master of Science in Aeronautics and Astronautics
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2023

© Massachusetts Institute of Technology 2023. All rights reserved.

The author hereby grants to MIT a nonexclusive, worldwide,
irrevocable, royalty-free license to exercise any and all rights under
copyright, including to reproduce, preserve, distribute and publicly
display copies of the thesis, or release the thesis under an open-access
license.

Authored by: Alexander P. Koenig

Department of Aeronautics and Astronautics

May 22, 2023

Certified by: Olivier L. de Weck

Apollo Program Professor of Astronautics

Thesis Supervisor

Certified by: Jacqueline N. Hewitt

Julius A. Stratton Professor of Physics

Thesis Supervisor

Accepted by: Jonathan P. How

R. C. Maclaurin Professor of Astronautics and Aeronautics

Chair, Graduate Program Committee

A Systems Framework for Multi-Messenger Astronomy

by

Alexander P. Koenig

Submitted to the Department of Aeronautics and Astronautics
on May 22, 2023, in partial fulfillment of the
requirements for the degree of
Master of Science in Aeronautics and Astronautics

Abstract

Multi-messenger (and more broadly, panchromatic) astronomy regards the use of multimodal information — incident photons, gravitational waves, neutrinos, and cosmic rays — to form astrophysical inferences. Since each messenger interacts uniquely with the dynamics of the phenomena in question, drawing information from multiple messengers poses a more complete probe of the universe. However, the exact inference method is scenario-specific, and we lack a general means to design multi-messenger instrument networks to best formulate scientific knowledge. To this end, this thesis presents a framework using probabilistic graph models to simulate the performance of heterogeneous instrument networks, with applications to two case studies.

The first case study regards the measurement of the Hubble parameter, i.e. the rate of expansion of the universe, with joint gravitational-wave and electromagnetic detection of neutron star mergers — cosmological standard sirens. This case study predicts 45^{+68}_{-32} joint detections by the end of the 2020s, likely sufficient to measure the Hubble parameter with 4% uncertainty. Furthermore, $\mathcal{O}(10^5)$ instrument networks are simulated. The most promising configurations rely on a highly-sensitive set of ground-based interferometers with wide geographic distribution along with a set of narrow-field, large-aperture ground- or space-based telescopes.

The second case study regards using star tracker imagery from LEO satellite constellations to improve our knowledge of resident space objects — active satellites and debris. Traditionally, orbit determination relies on bespoke ground-based radar systems which are increasingly insufficient to meet the needs of LEO satellite operators. For two simulated objects, this case study shows star trackers could supplement but not replace radars to improve knowledge: including imagery from 10^3 satellites could reduce positional uncertainty by a factor of ~ 3 compared to a radar-only network.

Thesis Supervisor: Olivier L. de Weck
Title: Apollo Program Professor of Astronautics

Thesis Supervisor: Jacqueline N. Hewitt
Title: Julius A. Stratton Professor of Physics

Acknowledgments

This thesis could not have happened without the support and guidance of my advisor Prof. Olivier de Weck. Oli, thank you for encouraging ambition in my research. I believe what is special about MIT is that there is no ceiling — the limit of possibility is up to oneself. That road is infinitely long, but not only do you embody that trait best, you helped me inch towards it as well. I only hope to continue meandering forward on that path. I also thank my co-advisor Prof. Jacqueline Hewitt who provided an essential second eye. Jackie, it was from you I learned the physics of relativity and neutron stars in the first place years ago — it was deeply appropriate to have you on board.

I am immeasurably grateful for all those that helped me and believed in me along the way. That includes my teachers throughout school who inspired curiosity, my mentors in life and in industry who imbued a degree of grit, my friends who kept me at least somewhat sane, and my family for their continuous support. Thank you.

I would also be remiss to not mention the rest of the 10^{11} of us to have lived and died on this pale blue dot. It has been a pleasure contributing in my own small way to the expanse of knowledge set forth by others. I look forward to seeing our progress in that Sisyphean task.

Contents

1	Introduction	19
1.1	Towards θ -complete astronomy	20
1.2	Astronomical sensor fusion	23
1.3	Research goals	24
2	Systems framework for multi-messenger astronomy	25
2.1	Messengers	25
2.2	Detectors	28
2.2.1	Electromagnetic detectors	28
2.2.2	Gravitational-wave detectors	30
2.2.3	Particle detectors	36
2.3	Measurements	37
2.3.1	Measurement models	38
2.3.2	The limits of information	41
2.3.3	Composing sets of measurements	44
2.4	Method	46
3	Multi-messenger astronomy of binary neutron star mergers	55
3.1	Science background and motivation	55
3.1.1	Cosmological context for the Hubble parameter	58
3.1.2	Gravitational-wave standard sirens as a new measurement of the Hubble parameter	62
3.1.3	Compact binary coalescence (CBC) and EM/GW emissions .	64

3.2	Simulation methods	73
3.2.1	Assumed binary neutron star merger population	76
3.2.2	Gravitational wave signal modeling for matched filtering	79
3.2.3	Gravitational wave parameter inference	81
3.2.4	Electromagnetic signal modeling	88
3.2.5	Electromagnetic follow-up of the sky localization regions	91
3.2.6	Calculating the detection rate	100
3.3	Results	102
3.3.1	Joint EM/GW detection rates through the 2020s	102
3.3.2	Mapping future multi-messenger networks	107
3.4	Discussion	115
4	Joint radar and optical measurement of resident space objects	117
4.1	Object covariances	119
4.2	Simulation methods	122
4.2.1	System dynamics	123
4.2.2	Radar measurement model	125
4.2.3	Optical measurement model	128
4.2.4	State estimation	134
4.3	Results	137
5	Conclusion	141
5.1	Future work	142
5.1.1	Improvements to case study modeling fidelity	142
5.1.2	Further analyses to build upon case study results	144
5.1.3	Additional applications	144
5.1.4	Mathematical development of the systems framework	146
A	Tables of instruments for Chapter 3	149

List of Figures

1-1	All Hubble images mapped onto the celestial sphere	21
1-2	Electromagnetic opacity of Earth's atmosphere	22
2-1	A standard layout of a gravitational-wave Michelson interferometer .	31
2-2	The characteristic strain noise of past, present, and future GW interferometers	33
2-3	Sensitivity of GW observatories over time	34
2-4	GW interferometer antenna pattern	35
2-5	A color movie represented in θ -space is a hypercube of pixel measurements.	47
2-6	The Bayesian network of JWST characterizing early-universe galaxy formation	48
2-7	The Bayesian network of the EHT taking an image of a black hole accretion disk	49
2-8	The Bayesian network for the binary neutron star merger case study .	49
2-9	The Bayesian network for the resident space object case study . . .	50
2-10	Breakdown of messengers and frequency ranges by case study . . .	53
3-1	Joint gravitational-wave and electromagnetic detection of a binary neutron star merger	56
3-2	21st century measurements of H_0	61
3-3	The three dynamic phases of a binary merger	68
3-4	The gravitational wave chirp of GW170817 as viewed by different detectors	68

3-5	KN ejecta	71
3-6	Rapid reddening and dimming of AT 2017gfo in the days following the merger	73
3-7	Timeline of multi-messenger observations of GW170817/AT 2017gfo .	74
3-8	A block diagram of the simulation approach	76
3-9	Rate density of BNS mergers with comoving distance	80
3-10	GW parameter estimation software packages	85
3-11	The KN light curve modeling approach	89
3-12	An example light curve	90
3-13	The wide-field tiling approach used by DECam to successfully find the optical counterpart of GW170817	92
3-14	The narrow-field galaxy-targeting approach used by 1M2H to detect the optical counterpart of GW170817	93
3-15	Analytical detection probability calculation	96
3-16	Observability criteria for a hypothetical signal observed by a hypothetical ground-based observatory	98
3-17	IGWN observing runs through the 2020s	102
3-18	Instrument network used for the 2020s simulation	103
3-19	GW network SNR of simulated mergers during O2	104
3-20	Marginal distribution of merger SNRs in O2 for simulated events at 40.7 Mpc	105
3-21	Sky localization regions of detectable mergers during O2	105
3-22	Apparent magnitudes of the simulated merger population	106
3-23	Marginal distribution of peak apparent magnitudes at 40.7 Mpc	107
3-24	Predicted jointly-detected binary neutron star mergers for IGWN observing runs through the 2020s	108
3-25	Ground-based portion of the instruments considered	109
3-26	Performance of different multi-messenger networks	110
3-27	Network performance gain due to individual GW interferometers	111
3-28	Network performance gain due to individual EM telescopes	112

3-29 Comparing performance improvement with the WINTER instrument and the MASTER/LCO network	114
4-1 Monthly number of objects in Earth orbit greater than 10 cm in size .	118
4-2 A block diagram of the simulation approach.	122
4-3 An example orbit propagation	124
4-4 The part of the Bayesian network showing just the radar measurement model	126
4-5 Example orbit showing radar ground station passes	128
4-6 The part of the Bayesian network showing just the optical measurement model	129
4-7 Shift-and-add centroiding technique	130
4-8 Optical astrometric covariance	132
4-9 Solar phase angle φ of an object/observer pair	133
4-10 Star tracker FOVs	135
4-11 Joint tracking of a small debris fragment in equatorial LEO	139
4-12 Joint tracking of a smallsat in dawn/dusk sun-synchronous orbit . . .	140

List of Tables

2.1	Astrophysical [messengers] via the standard model of particle physics	27
2.2	Present and proposed gravitational-wave interferometers	34
2.3	Distributing uncertainty among measurement dimensions	44
3.1	WMAP Λ CDM parameters [85]	60
3.2	Qualitative observability of compact binary coalescence pairs	66
3.3	Binary neutron star parameters	67
3.4	Assumed binary neutron star parameter distributions	77
3.5	Comparison of GW parameter estimation software packages	86
A.1	Gravitational-wave interferometers	150
A.2	Electromagnetic telescopes	150

Nomenclature

Acronyms

Λ CDM Lambda cold dark matter cosmological standard model

BNS binary neutron star

CCD charge-coupled device (imaging array)

CDM conjunction data message

CMB cosmic microwave background

EKF extended Kalman filter

EM electromagnetic

FOV field of view

FWHM full width at half maximum

GW gravitational wave

HMNS hypermassive neutron star

IGWN International Gravitational-Wave Observatory Network

KN kilonova

LIGO Laser Interferometer Gravitational-Wave Observatory

PSD power spectral density

PSF point spread function

RSO resident space object

SSA space situational awareness

SSN Space Surveillance Network

STM space traffic management

Greek characters

α right ascension

δ declination

η quantum efficiency of a CCD

ι binary neutron star merger inclination

λ wavelength

ν frequency

ν neutrino

ϕ phase

ψ polarization angle

ρ signal-to-noise ratio (SNR)

σ standard deviation

θ measurement parameter space

ε polarization ellipticity

Roman characters

\hbar reduced Planck constant

\mathcal{C}	covariance matrix
\mathcal{I}	Fisher information matrix (FIM)
\mathcal{M}	binary neutron star chirp mass
$\mathcal{O}(\cdot)$	order of magnitude
A	amplitude
c	speed of light in vacuum
d_C	comoving distance
d_L	luminosity distance
f	focal ratio
h	gravitational wave strain
H_0	Hubble constant
X	dynamical state vector
z	redshift

Chapter 1

Introduction

Our windows into the universe have historically been significantly constrained in their extent. Until the first radio astronomy observation in 1932 [145], humanity only ever viewed the cosmos through the lens of the narrow optical portion of the electromagnetic (EM) spectrum. While in the 20th century we largely filled out the EM spectrum with panchromatic suites of instruments like those of the Great Observatories program [14], other windows into the universe have remained closed until recently — not just narrow or limited in extent, but completely inaccessible to measurement.¹ Three new spectra have garnered the attention of the astronomical community: the gravitational-wave (GW), neutrino (ν)², and cosmic ray spectra. In the same way that a color movie captures qualitatively different information about a scene compared to a black-and-white film, multi-messenger astronomy promises to unlock novel insights about the universe. To quote the Astronomy and Astrophysics 2020 Decadal Survey, multi-messenger observations “are the only path to disentangle and decipher the nature and history of our complex universe. The realization of that goal has been the basis of twenty-first-century space astrophysics. No single telescope alone can answer all of the most pressing questions in the field, from the nature of the Big Bang to the emergence of life on planets” [113].

The age of multi-messenger astronomy is still in its infancy. As of early 2023, only

¹Though arguably, multi-messenger astronomy predates panchromatic astronomy, as cosmic rays were first detected during a 1912 balloon flight [132].

²Not to be confused with frequency ν .

a handful of astronomical sources have been detected in more than one spectrum. Prominent examples include a binary neutron star merger (GW/EM) [51], the Sun (EM/ ν /cosmic rays), a supernova (EM/ ν) [86], a blazar (EM/ ν) [119], and a possible black hole tidal disruption event (EM/ ν) [67].

In the ideal limit of receiving maximum information about the universe, we would measure every messenger, in every frequency, from every incident direction, at all points in time, and with sensitivity constrained only by quantum uncertainty. Our windows into the cosmos would be maximally wide and maximally precise. Given that we are far from that limit and have finite effort to dedicate towards getting there, the natural question that follows is: which set of measurements should we prioritize first?

1.1 Towards θ -complete astronomy

Every astronomical detector measures one or more of the amplitude A , frequency ν , phase ϕ , polarization orientation ψ and ellipticity ε , directions α and δ , and timing t of incident waves, henceforth denoted $\theta \equiv [A, \nu, \phi, \psi, \varepsilon, \alpha, \delta, t]$. [95] refers to these as the *latent variables* of the measurement, since the measurement dimensions may not directly coincide with θ -space. For the most part, different types of detectors of the same messenger merely measure different sets of these same physical quantities. For example, a monochromatic optical telescope determines $\{A, \alpha, \delta\}$ of photons — the amplitude of incident light in a certain direction — whereas a spectrometer primarily infers $\{A, \nu\}$, sacrificing directionality uncertainty for frequency uncertainty.

From the very first telescopes in the 17th century all the way through the modern era, much of astronomical instrument design has focused on improving the informational quality of measurements of A , ν , α , and δ at specific points in θ -space.³ Larger telescopes with more sensitive detectors and more finely-tuned filters have progressively improved — by orders of magnitude! — at resolving smaller, dimmer sources at finer frequency resolutions. However, improvements in the *breadth* of detection

³Lots of progress has also been made in radio and optical polarization, though discussion of polarization more generally is left to future work.

capabilities have been largely bereft. Our ability to map the complete domain of the electromagnetic spectrum — the complete θ -space, the entirety of the luminous events in our sky — remains relatively limited.

As a somewhat hyperbolic point, contrast Hubble with Galileo’s first telescope: with a 26 mm aperture, Galileo’s telescope had a resolution limit of $\mathcal{O}(10^{-3})$ deg [117]. In comparison, Hubble’s 2.4 m aperture has a resolution limit of $\mathcal{O}(10^{-5})$ deg [2], representing a $\mathcal{O}(10^2)$ improvement in both directional uncertainty and sensitivity.⁴ Yet Hubble’s field of view, 0.02 deg², is even smaller than that of Galileo’s at roughly 0.05 deg² — if not for Hubble’s higher throughput which allows it to take more measurements in a shorter period of time, our quintessential 20th-century space telescope would be no better than 17th-century technology at achieving θ -complete astronomy. By one estimate [149], Hubble has only collectively imaged 0.8% of the entire sky, as depicted in Fig. 1-1.

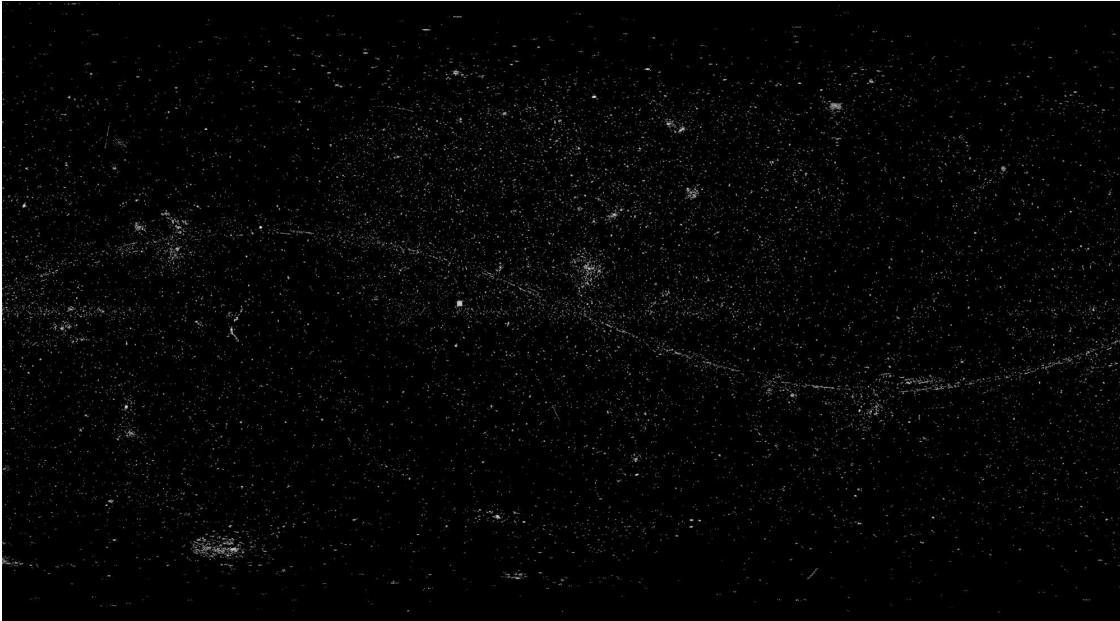


Figure 1-1: All Hubble images mapped onto the celestial sphere. These 1.4+ million images, taken during Hubble’s 3+ decade operational lifetime, collectively map a mere 0.8% of the 4π sr area of the sky. The sinusoidal feature is the ecliptic plane corresponding to imagery of solar system objects. Image credit [149].

In the late 20th century, we began the first steps towards θ -complete astronomy

⁴SNR scales with the square root of the signal, so a doubling of diameter does not mean a quadrupling of sensitivity.

by filling out broader portions of measurement-space via compositions of a suite of astronomical instruments. Particular attention was paid to ν -space by building instruments that could measure all frequencies of the electromagnetic spectrum. This transition towards panchromatic astronomy was enabled by access to space, as Earth's atmosphere is opaque to much of the electromagnetic spectrum (see Fig. 1-2). For a brief and by no means complete history of panchromatic astronomy:

- The first purpose-built radio telescope was built by amateur astronomer Grote Reber in 1937 [145].
- Uhuru, the first spaceborne X-ray observatory, was launched in 1970 [71].
- The first UV telescope was the International Ultraviolet Explorer (IUE), launched in 1978 [11].
- IRAS (aptly, the Infrared Astronomical Satellite) was the first spaceborne infrared telescope, flown in 1983 [114].

NASA's Great Observatories program, a major focus in astronomy with launches from 1990 to 2003, specifically focused on filling out the electromagnetic spectrum with exquisite astrophysical instruments: Hubble in the UV/optical/NIR, Spitzer in the IR, Chandra in X-rays, and Compton in γ -rays [14].

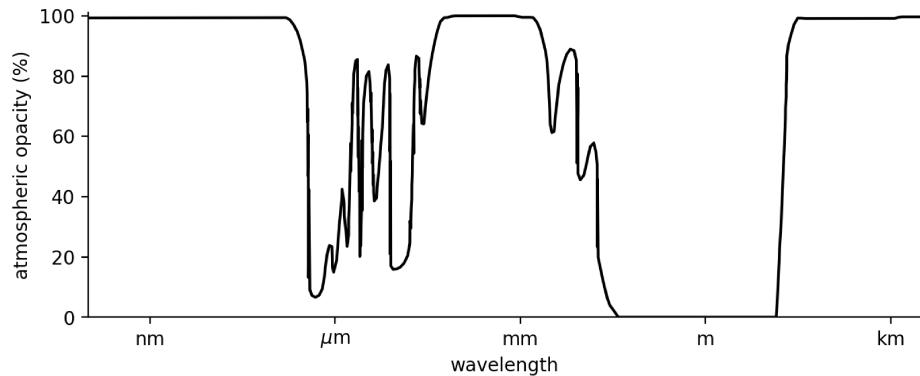


Figure 1-2: Electromagnetic opacity of Earth's atmosphere.

Access to space has enabled missions which can measure a broader range of EM frequencies, allowing us to achieve ν -complete astronomy. Adapted from [112].

More recent developments in astronomy have taken even greater strides towards θ -complete astronomy via time-domain, transient, and survey astronomy, which collectively address the t , α , and δ components of the θ -space of the EM spectrum. The Zwicky Transient Facility (ZTF) is a wide-field survey and transient observatory with a 47 deg^2 field of view, acquiring first light in 2018 and able to scan the entirety of the Northern sky on a nightly basis [55]. The Vera C. Rubin Observatory is a ground-based, $f/1.23$, 8.4 m diameter wide-field telescope that will image the accessible Southern sky every few nights via a 9.6 deg^2 field of view, generating 20 TB of data per night [136]. The Roman Space Telescope, planned for launch in 2026, promises a field of view 100 times larger than that of Hubble [53].

Still, the window of observations of our universe remains small. Even the Rubin Observatory addresses a mere 0.02% of $\alpha\delta t$ -space, or 0.0005% of $\alpha\delta t\nu$ -space.⁵ The extent of physical possibilities before us is vast, and we are far from the fundamental limit of measuring every incident photon, from every direction, at every frequency, and at all points in time — not even to mention the three other underexplored messengers.

1.2 Astronomical sensor fusion

Simply measuring the θ -space of messengers is a necessary but insufficient condition for achieving progress in astronomy. As with traditional astronomy, one must use the measurements to make hypotheses about how the world functions. However, in contrast with traditional astronomy, the distinguishing trait of multi-messenger astronomy is not just that it refers to multiple messengers. Rather, it involves making cohesive scientific inferences from a variety of disparate data sources. Multi-messenger astronomy is just one name used to refer to this process. Other examples of the same phenomenon include panchromatic astronomy, which refers to measurements of disparate frequencies or energy levels of a singular messenger. Similarly, time-domain astronomy refers to temporally-variable measurements.

⁵Where the considered frequency range is from long-wave radio to hard γ rays (10^0 to 10^{24} Hz) and the fraction of ν that is measured is calculated in log-space.

For lack of a better term, this work uses “multi-messenger” as an umbrella phrase to capture all aspects of this process of data fusion. In all such cases, sensor fusion cannot happen exclusively at the measurement-level, because the measurements involved are multimodal. Unlike measurements require feature-level fusion rather than measurement-level fusion (further described in Chapter 2). Essentially, one must assume a certain model of the phenomena in question — a specific mapping from measurements to phenomenological features — in order to perform multi-messenger astronomy at all.

1.3 Research goals

The central issues of multi-messenger astronomy regard (1) how to form scientific hypotheses given multimodal data, (2) how to determine where multi-messenger measurements can yield better information than traditional methods alone, and (3) how to optimally build a network of observatories given finite resources. This thesis presents a general framework for conceptualizing multi-messenger astronomy, drawing inspiration from the fields of information theory and sensor fusion. Chapter 2 further elaborates on the astrophysical messengers, their corresponding detectors, and fundamental limits of information and uncertainty. In the second half of the chapter, a general method is presented which characterizes the performance of multi-messenger observatory networks through the use of Bayesian networks. These networks track the flow of information from sensors to scientific hypotheses. Chapters 3 and 4 exemplify this method for two case studies which demonstrate different flavors of astronomical sensor fusion. The first case study, Chapter 3, involves using joint gravitational and electromagnetic detection of binary neutron star mergers to measure the Hubble parameter, the present expansion rate of the universe. The second case study, Chapter 4, involves performing orbit determination with asynchronous radar and optical measurements of artificial objects in Earth orbit for the purpose of space situational awareness. Chapter 5 concludes and recommends future work to further develop this conceptual framework.

Chapter 2

Systems framework for multi-messenger astronomy

2.1 Messengers

Given only *a posteriori*, the set of messengers of multi-messenger astronomy may appear somewhat arbitrary: why photons, gravitational waves, neutrinos, and cosmic rays specifically? Are there other messengers we might be interested in but which we have yet to measure? These four messengers can be understood in the lens of the standard model of particle physics and the fundamental forces, which describe the set of known particles and how they interact.

The four fundamental forces — electromagnetic, gravitational, weak, and strong — are generally mediated by elementary particles called gauge bosons, with the possible exception of the gravitational field [123]. Photons, the subject of traditional astronomy, are the familiar and well-studied bosons which mediate the electromagnetic field. Gravitons, the hypothesized boson mediator of the gravitational field, are not in the standard model of particle physics as a true quantum description of gravity is not presently known [43].¹ Nevertheless, gravitational waves are similarly detectable over significant distances, as the spatial distortions predicted by general

¹However, if the gravitational field were able to be expressed in the framework of a quantum field theory, the existence of such a graviton would presumably follow.

relativity can be directly measured. W and Z bosons, which transmit the weak force, have a half life of 10^{-25} s due to their large mass, and therefore quickly decay into other particles before any chance of an astrophysical detection [123]. Gluons, which mediate the strong force, do not decay but also do not separate from quarks due to the confinement property of the strong force. As a result, between the four fundamental forces, only EM and GW forces directly propagate over astrophysical distances [104].

However, the weak and strong forces can provide astrophysical insight through fermions, the non-bosonic half of the standard model consisting of quarks and leptons. Fermions are massive² particles which can propagate over astrophysical distances given sufficiently relativistic energies. Cosmic rays are combinations of quarks and leptons, consisting of baryons (individual protons and neutrons, as well as the atomic nuclei they compose) and electrons [132]. Cosmic rays correspond most closely to the strong force, which confines sets of quarks into protons and neutrons, the primary component of cosmic rays. Neutrinos are the other form of leptons — essentially, chargeless and less massive versions of electrons — which interact with other forms of matter primarily through the weak force.³ All types of fermions are included under the umbrellas of cosmic ray and neutrino messengers.

Therefore, the complete set of four astrophysical messengers arises fairly naturally from fundamental physics as per the four fundamental forces as well as the elementary particles of the standard model, summarized by Table 2.1. There is a close mapping between the four messengers and the four fundamental forces. Absent possible extensions to the standard model, some of which propose new particles — and some further subset of which may actually be detectable — these messengers are the exclusive means by which we can gain information about the universe from our vantage point in the solar system.

²Massive meaning “having mass,” not to be confused with “very heavy”.

³Neutrinos also interact with other matter through the gravitational field.

Table 2.1: Astrophysical messengers via the standard model of particle physics

type	particles/antiparticles	name	relation to forces:			
			EM	gravity	weak	strong
fermions	quarks	u/\bar{u} d/\bar{d} s/\bar{s} c/\bar{c} t/\bar{t} b/\bar{b}	up down strange charm top bottom	experience	experience	experience
	cosmic rays	e^-/e^+ μ^-/μ^+ τ^-/τ^+	electron	experiences	experience	experience
leptons	ν	$\nu_e/\bar{\nu}_e$ $\nu_\mu/\bar{\nu}_\mu$ $\nu_\tau/\bar{\nu}_\tau$	neutrino	experience	experience	experience
bosons	Higgs boson			unrelated; accounts for mass		
	gauge bosons	EM W^-/W^+ Z g	photon W boson Z boson gluon	mediate	mediate mediate	mediate
outside of the standard model...			GW [gravitational strain]	graviton?	mediate	

2.2 Detectors

Any astrophysical instrument is a physical instantiation of a device which can measure one of the above messengers. The particular forms of the implementation, i.e. the technologies used to construct the instrument, do not directly determine the information it intends to measure. However, the instrument’s function, i.e. its performance or ability to effectively measure the messenger, certainly influences what information can be known. As such, it is necessary to understand the operating principles of typical implementations of these measurement devices, particularly as they relate the form and function of the device. This thesis considers a variety of astrophysical instruments which take measurements in a range of spectra. In this section, we cover the form and function of electromagnetic instruments ranging from radars to γ -ray detectors, as well as gravitational-wave interferometers, neutrino detectors, and cosmic ray detectors.

2.2.1 Electromagnetic detectors

The first type of electromagnetic detectors considered are passive telescopes which gather and focus incident light throughout different portions of the electromagnetic spectrum. Traditional imagers measure focused light in pixels across a focal plane array, capturing tight directional constraints on the objects (particularly at higher wavelengths, due to the Rayleigh criterion) as well as measuring amplitudes, but without differentiating the particular frequency of the incident light. In terms of detecting distant objects, telescopes observe an inverse square law where the brightness of a signal decays as $1/r^2$. The exact form of the instrument depends on the application and the particular portion of the spectrum:

- Radio telescopes, which in the basic case may feature a single dish and antenna receiver, but can also be composed of many antennae in a phased array configuration
- Infrared telescopes, which can either feature thermal detectors (e.g., bolometers)

or photonic detectors (e.g., mercury-cadmium-telluride and indium-gallium-arsenide CCDs, which are relatively new within astronomy), and which typically require cooling to avoid self-emission of infrared radiation

- Optical telescopes, which reflect or refract photons onto a 2-dimensional CCD or CMOS sensor, but can also come in the form of a spectrometer or polarimeter to measure the frequency or polarization of light respectively
- Ultraviolet telescopes, similar in operation to optical telescopes but with different semiconductor bandgaps and often specialized coatings to enhance UV reflectivity
- X-ray telescopes, which use grazing-incidence optics to focus high-energy photons onto detectors such as CCDs, microcalorimeters, or proportional counters
- γ -ray telescopes, which do not use traditional focusing optics due to the extreme energies of gamma-ray photons,⁴ and instead employ designs like scintillation detectors, Compton telescopes, or pair-production telescopes to measure the frequency and direction of γ -rays

These detectors may also use filters to observe other properties of incident EM waves, including spectrographs to measure frequency and polarimeters to measure polarization.

Other electromagnetic detectors include active instruments such as radars and lidars — short for RAdio Detection And Ranging and LIght Detection and Ranging respectively — which both emit an electromagnetic signal and then receive the reflected signals off of objects within the beam. By determining the time delay between the emission and reception of the reflected signals, these instruments can determine the range, range-rate, and direction of objects. Active instruments suffer from a fundamental $1/r^4$ signal power loss due to the two-way free space loss (the $1/r^2$ spread) of emissions, and therefore are limited to observations of relatively nearby objects

⁴Recall the wave-particle distinction discussed in Sec. 2.3: at γ -ray energies, photons are not as amenable to manipulation via optics as they behave less like waves and more like individual particles.

even in high-power, high-gain scenarios. In astronomy, radars are primarily used for studying solar system objects, such as planets, moons, asteroids, and comets. They are also applied to space situational awareness to map artificial objects in Earth orbit, as analyzed in Chapter 4. Lidars, the optical counterpart to radars, have also been used in astronomy to measure the orbit of the moon [30].

2.2.2 Gravitational-wave detectors

The standard and (to date) most successful gravitational-wave instrument design is that of a Michelson interferometer, which measures changes in the length of two orthogonal laser cavities (interferometer “arms”). Incident gravitational waves lengthen and contract space transversely to the direction of propagation as per the amplitude and frequency of the waves. The actual signal one intends to measure is the GW strain $h(t)$, defined as:

$$h(t) \equiv \frac{\Delta L}{L}, \quad (2.1)$$

a unitless measure of the relative expansion and contraction of a reference length L . From Eqn. 2.1, one can see a predominant consideration of interferometer design. The best design maximizes L , the length of the interferometer arms, so as to minimize uncertainty in $h(t)$.⁵

To measure ΔL , a coherent laser beam is split and sent through each cavity, returned by suspended mirrors, and interferometrically rejoined at the beam splitter. Phase differences accumulated within each cavity result in changes in amplitude in the rejoined beam, which are recorded by a photodiode. As a result, what is measured is not exactly ΔL , but rather the difference in ΔL between the two interferometer arms. A diagram of this design is given in Fig. 2-1. At a higher level of abstraction, GW interferometers are in a sense active EM instruments — both generating and receiving the observed electromagnetic signal — which can map an observed photodiode amplitude over time back to $h(t)$. The length of the interferometer arms and the input laser power are two predominant design considerations, as those determine

⁵Presuming that ΔL is not proportional to L , which is sometimes the case.

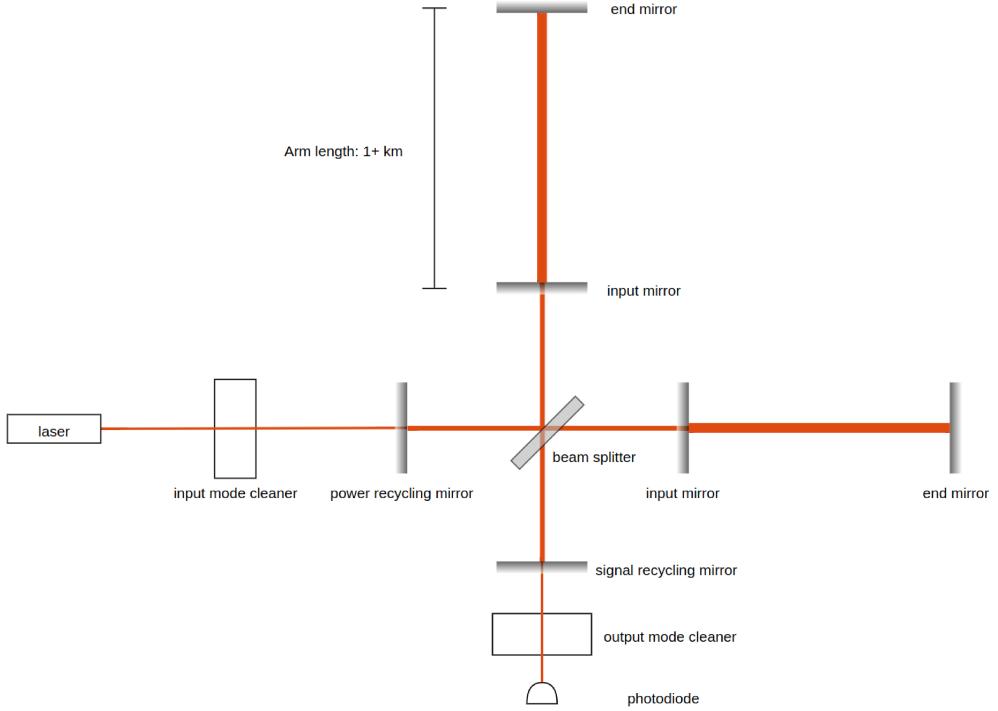


Figure 2-1: A standard layout of a gravitational-wave Michelson interferometer. A laser beam is generated, cleaned, and sent to two optical cavities. The beam is recycled (resonated) within these cavities and eventually returned to the beam splitter, where it is again cleaned and then measured by a photodiode. Variations in the arm lengths cause interference in the recombined beam, which can be measured to infer the gravitational wave strain.

the frequency and sensitivity of the design: longer arms can measure weaker, lower-frequency signals,⁶ and increased laser power can measure weaker, higher-frequency signals.⁷

This baseline design is the reference from which more advanced variations are constructed. One major sensitivity gain (also shown in Fig. 2-1) is created by passing the beam back and forth within the laser cavities many times before returning it to the beam splitter. This resonance boosts the effective length of the interferometer arms [25]. There are a number of ways to achieve this resonance, including power

⁶Higher L results in lower received power due to $1/r^2$ loss, which is one reason why high- L designs are sensitive to lower frequencies — essentially, receiving the same total wattage requires a longer integration time.

⁷Higher power reduces shot noise, the predominant noise contribution at high frequencies. However, the added photon pressure induces additional accelerations in the test masses, resulting in acceleration noise at low frequencies.

recycling (reflection on the laser side of the beam splitter), arm recycling (reflection within each cavity in advance of return to the beam splitter), and signal recycling (reflection on the photodiode side of the beam splitter), which are collectively combined in a format known as dual recycling [25]. Furthermore, input and output mode cleaners placed after the laser source and in advance of the photodiode respectively can improve the quality of the coherent laser beam. This is achieved by geometrically “cleaning” the beam, essentially reducing fluctuations in its amplitude, frequency, and pointing [56]. Another gain is created from the quantum uncertainty limit relating amplitude A and phase ϕ , given by $\sigma_A \sigma_\phi \geq \hbar/2$. Since the photodiode only measures output light amplitude and not phase, one can improve sensitivity by transferring the uncertainty from amplitude into phase (known as “quantum squeezing”) [29], a technique successfully demonstrated in the Laser Interferometer Gravitational-Wave Observatory (LIGO) detectors [4].

The performance of a GW interferometer can be uniquely identified by its power spectral density (PSD), or equivalently its strain sensitivity curve: a measure of the noise floor of the strain amplitude as a function of GW frequency. Fig. 2-2 shows these curves for a noncomprehensive but expansive list of GW interferometers described in Table 2.2. Various noise contributions exist, but the most dominant by far are acceleration disturbances in low-frequency bands (seismic and suspension noise for ground-based detectors, and acceleration noise for space-based detectors) and photon shot noise (i.e., quantum amplitude uncertainty) in high-frequency bands. Photon shot noise increases with frequency, whereas acceleration noise increases with lower frequency — as such, the sum of these primary contributions creates a characteristic “V” shape in the sensitivity curves, where the minimum noise corresponds to the point at which these two contributions are roughly equal. The absolute height of the strain sensitivity curve indicates the noise level, the width indicates the GW bandpass, and the horizontal location indicates the frequency sensitivity of the detector.

Geographic location⁸ and seismic/acceleration isolation improves low-frequency

⁸Specifically, placement away from seismically active areas (or strongly-perturbative regions in the case of space-based interferometers).

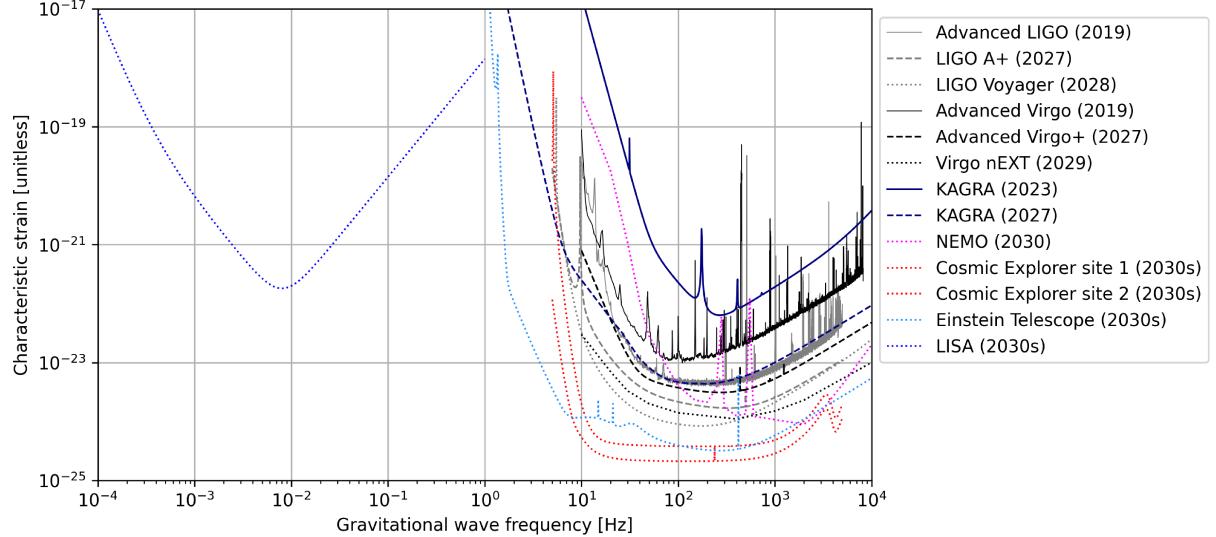


Figure 2-2: The characteristic strain noise of past, present, and future GW interferometers.

Lower strain noise indicates higher sensitivity. Data sources for this figure are given in Table 2.2.

sensitivity, whereas laser power, interferometer arm length, and quantum squeezing all affect high-frequency sensitivity. Note that the Laser Interferometer Space Antenna (LISA), the only space-based interferometer shown here, has a much lower frequency response as a result of its longer interferometric arms, which result in a reduction of observed laser power but also of acceleration noise [38].

GW interferometers are a relatively new technology, with first invention in 1972 [25] and first successful detection in 2015 [5]. As such, GW interferometers are experiencing continuous and rapid advancements in sensitivity, both through the design of new observatories as well as through upgrades to existing facilities. Though different interferometers may be tuned to different wavelengths, a figure of merit for performance can be defined as a frequency-normalized minimum sensitivity, which corresponds to the minimum of the strain sensitivity curve multiplied by the square root of the frequency at that minimum. Fig. 2-3 shows the trends in sensitivities for different variants of GW detectors. Note that pulsar timing arrays, not otherwise discussed, are a separate method of measuring particularly low-frequency GW signals by detecting variations across known millisecond pulsars [70].

Table 2.2: Present and proposed gravitational-wave interferometers

name	year	location(s)	countries	source
Advanced LIGO	2019	WA & LA	USA	[45]
LIGO A+	2027	WA & LA	USA	[6]
LIGO Voyager	2028?	WA & LA	USA	[66]
Advanced Virgo	2019	Italy	EU	[56]
Advanced Virgo+	2027	Italy	EU	[6]
Virgo nEXT	2029?	Italy	EU	[63]
KAGRA	2023	Japan	Japan	[6]
KAGRA (design)	2027	Japan	Japan	[6]
NEMO	2030?	Australia?	TBD	[58]
Cosmic Explorer	2030s	2x in USA?	USA	[54]
Einstein Telescope	2030s	Belgium or Germany?	EU	[60]
LISA	2030s	heliocentric orbit	ESA (+NASA)	[38]

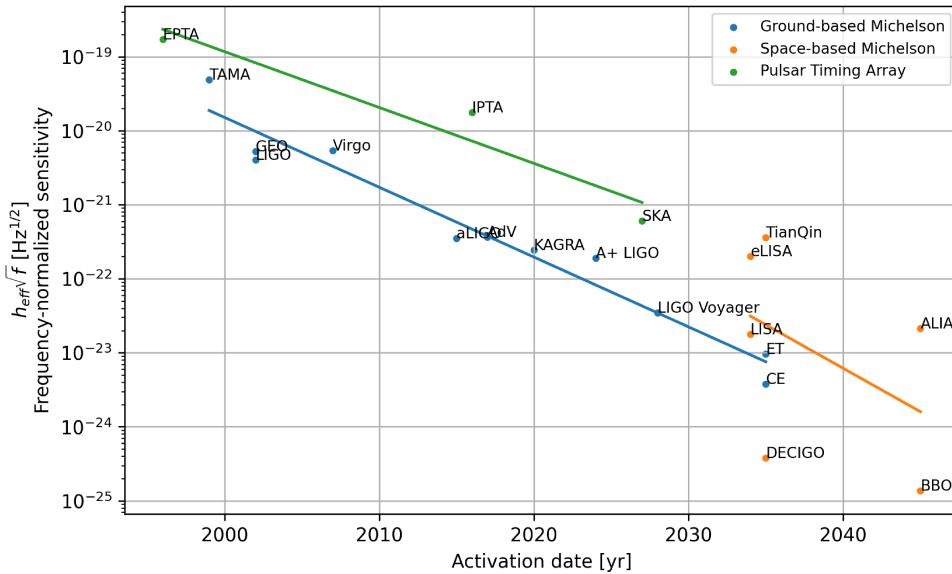


Figure 2-3: Sensitivity of GW observatories over time.
A factor of 10^6 improvement is anticipated in the first half of the 21st century. Data from [108].

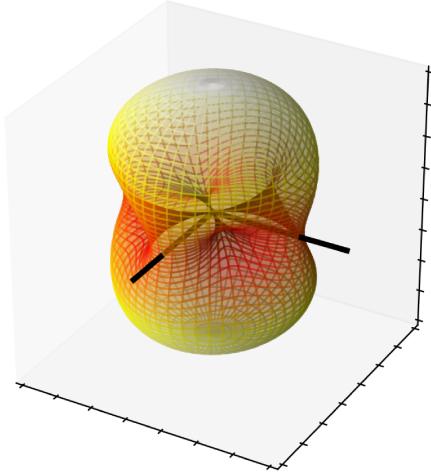


Figure 2-4: GW interferometer antenna pattern.

Shown here is the RMS sensitivity (both + and \times polarizations of incident waves), where arm orientations are indicated by the two black lines. Interferometers are most sensitive to sources at zenith and nadir, half as sensitive to sources along either arm's axis, and entirely insensitive to sources angularly halfway between the arms.

One simultaneous strength and shortcoming of GW interferometers is that they act as omnidirectional antennae, which enables them to observe large portions of the sky yet also renders them ineffectual at localizing detected sources. This antennae pattern is due to the fact that an interferometer measures the difference in ΔL between its laser cavities rather than the raw ΔL . Orientations of incident waves for which ΔL is equivalent between the arms result in no detection; orientations where an opposite effect is induced correspond to the most sensitive detection. The antenna pattern for an interferometer with two 90° arms is shown in Fig. 2-4.

In order to localize sources with such an omnidirectional antenna, an approximate origin can be inferred via the time delay in a signal observed by multiple geometrically diverse detectors. Secondarily, the observed amplitude can further indicate the orientation of the source relative to the interferometer antenna pattern when the source amplitude is predictable or otherwise known (e.g. by comparing it to the amplitude observed by other detectors).

2.2.3 Particle detectors

Particle detectors represent a third class of instruments which measure incident massive particles from astrophysical sources. Here we refer to fermion particles exclusively (thereby excluding bosons like photons), classifying the fermions separately as either cosmic rays or neutrinos (ν).

A “cosmic ray” is something of a catch-all term for baryons and leptons — primarily high-energy protons, neutrons, and atomic nuclei, but also electrons — which originate from astrophysical sources (e.g., supernovae, active galactic nuclei, and γ -ray bursts) with an appropriate energy range that permits them to propagate through the interstellar medium [132]. Cosmic ray detectors vary in design based on the energy (frequency) range of their detection, generally from \sim GeV at the low end to \sim TeV at the high end [104]:

- For detecting low-energy cosmic rays (up to a few GeV), solid-state detectors operate based on the interaction between incoming particles and the detector material, producing charge carriers that can be measured as an electrical signal. Gas-filled detectors such as drift chambers use ionization of gas within the detector to measure incoming particles.
- For middle-energy cosmic rays (a few GeV to several TeV), scintillation detectors emit light via incident ionizing radiation, which is converted to an electric signal through a photomultiplier tube. Cherenkov detectors sense radiation produced when a charged particle travels through a medium faster than the speed of light in that medium.
- For high-energy cosmic rays (above several TeV), air shower arrays and imaging atmospheric Cherenkov telescopes both detect cosmic rays indirectly by sensing secondary particles produced when cosmic rays interact with Earth’s atmosphere.

Neutrinos are other massive leptons like the electron/muon/tau trio of particles, but which are much lighter than their electron counterparts. Unlike electrons, neu-

trinos are not charged and therefore do not interact with other matter through the electromagnetic field, instead only weakly interacting with other matter through gravity and the weak force. As a result, detecting neutrino emissions from cosmic sources poses a major challenge, and requires amassing a significant amount of material for the neutrinos to interact with. One common design for neutrino detectors involves embedding photomultiplier tubes in a large volume of water or ice, such as the Ice-Cube Neutrino Observatory in Antarctica [62]. In the event a neutrino interacts with the material, it produces Cherenkov radiation which can be detected by the photomultiplier tubes. The nature of the measured Cherenkov radiation can then indicate the energy and direction of the original source.

2.3 Measurements

Recall the various dimensions of incident messenger waves defined in Chapter 1 as $\theta \equiv [A, \nu, \phi, \psi, \varepsilon, \alpha, \delta, t]$. In the ideal world, a measurement taken by a particular detector would describe the exact values of each of these dimensions. More realistically, any measurement a detector may have taken places some constraint on each of these dimensions. These constraints are described by the measurement model, which maps a measurement to a space of possible inputs that could have caused the measurement. Some components of the measurement model may be more or less sensitive than others, and some may be entirely insensitive to the range of inputs. Instrument design largely regards improving the measurement model, reducing its noise contributions, and tuning it to specific science goals or target sources. The measurement model is also the level at which sensor calibration occurs, either through adjusting or characterizing each component of the measurement model.⁹

Strictly speaking, not every astronomical instrument detects waves. Optical CCDs, for example, count individual photons, and neutrino detectors measure individual neutrino particles rather than the quantum waveform of the neutrinos. However,

⁹Since the measurement model is abstracted from the particular phenomena that are measured, we can know that a sensor calibrated via a well-known source will work equally well for lesser-known sources.

wave-particle duality indicates that every quantum entity is describable as either a particle or a wave, so greater specificity is not required here. Often times the distinction will be made based on energy, where sufficiently high-energy entities behave more like particles than waves due to their high frequency and corresponding small wavelength. The convention of a common θ -space among messengers is preferred in this section, so it is easier to think of messengers in a wave-like context.

2.3.1 Measurement models

A measurement model for a given detector maps the latent variables θ to the observables actually measured by the detector, θ_{measured} :

$$\theta_{\text{measured}} = F(\theta) \quad (2.2)$$

where $F(\theta)$ is the measurement model [95].¹⁰ Treating Eqn. 2.2 as an inverse problem allows us to infer θ — the properties of the incident waves — based on the measurement and the *inverse* of the measurement model, $F^{-1}(\theta_{\text{measured}})$. Therefore, the measurement model informs how well we are able to constrain θ based on any particular measurement, and thus how much information we gain.

For the sake of concreteness, consider the example of a single pixel within the CCD of a monochromatic optical telescope. This pixel exclusively produces a single output quantity — an electron count — which is essentially a measured amplitude of incident light, A_{measured} . A_{measured} depends on the θ of the incident EM wave via the measurement model, components of the measurement are described in the following paragraphs.

In the basic case, the amplitude dependence of the measurement model for a CCD pixel is simply $A_{\text{measured}} \propto A$. However, CCDs have a discretized electron counts and

¹⁰This equation is written as an equality, but is really just attributing a certain model to the process. Instrument noise and bias also contribute to the measurement, and the model may be incomplete or otherwise imperfect.

a finite well depth A_{\max} . Therefore the amplitude dependence is given by:

$$A_{\text{measured}} \propto \min(|A|, A_{\max}). \quad (2.3)$$

Note that this function alone is relatively easy to invert, as a given A_{measured} maps back to a small set of possible input A for most ranges of A . Assuming the pixel is not saturated, the measurement is able to tightly constrain A .

Consider next the time dependence. A measured photon may have arrived at a pixel on the detector at any time $t \in [t_i, t_f]$ corresponding to the start and end of the image exposure. Therefore, the time dependence of the pixel is:

$$A_{\text{measured}} \propto \begin{cases} 1 & \text{if } t \in [t_i, t_f] \\ 0 & \text{if } t \notin [t_i, t_f]. \end{cases} \quad (2.4)$$

In addition, the photon may have come from some direction corresponding to the directional dependence of the pixel, also known as the point spread function (PSF). An on-axis, in-focus point source measured through a single circular optic has an Airy disk as the PSF:

$$A_{\text{measured}} \propto \left(\frac{2J_1(\pi D_c \nu \sin \theta)}{\pi D_c \nu \sin \theta} \right)^2, \quad (2.5)$$

where $\theta \approx \sqrt{\Delta\alpha^2 + \Delta\delta^2}$, D is the circular aperture diameter, and J_1 is the Bessel function of the first kind of order one. The first zero of the Airy disk occurs at the well-known Rayleigh criterion $\theta = 1.22 \frac{c}{\nu D} = 1.22 \frac{\lambda}{D}$. We can begin to see that the measurement model defines the instrument performance, as it indicates how well a given measurement constrains the input parameters. In this case, the directional dependence defined by the PSF indicates the angular resolution — essentially, the angular uncertainty — of a source.

Furthermore, even though the telescope is monochromatic, it too measures frequency, as the incident light must be within the spectral response of the pixel in order to be detected. For a given CCD quantum efficiency $\eta(\nu)$, the frequency dependence

is modeled through

$$A_{\text{measured}} \propto \eta(\nu). \quad (2.6)$$

The remaining quantities ϕ , ψ , and ε are ones that an optical telescope is considered not to measure, because the measurement is uniform over the space of physically-possible values, so inverting the measurement back into θ -space does not further constrain ϕ , ψ , or ε :

$$A_{\text{measured}} \propto 1 \quad \text{if } \phi \in [0, 2\pi), \psi \in [0, 2\pi), \text{ and } \varepsilon \in [0, \infty). \quad (2.7)$$

These equations can be composed to define the overall measurement model:

$$A_{\text{measured}} = \underbrace{\min(\lfloor A \rfloor, A_{\max})}_{A} \underbrace{\begin{cases} 1 & \text{if } t \in [t_i, t_f] \\ 0 & \text{if } t \notin [t_i, t_f] \end{cases}}_t \underbrace{\left(\frac{2J_1(\pi D_c^\nu \sin \theta)}{\pi D_c^\nu \sin \theta} \right)^2}_{\alpha, \delta} \underbrace{\eta(\nu)}_\nu \underbrace{1}_{\phi, \psi, \varepsilon} \quad (2.8)$$

A “good” detector is one that has an easily-invertible measurement model. More specifically, each final measurement should map back to a small and specific set of possible input quantities so as to tightly constrain the estimate of the physical quantities of the source being measured. [95] refers to this trait as the precision of the measurement, not to be confused with the uncertainty of the measurement, which regards the confidence rather than uniqueness of the mapping. From Eqn. 2.8, we can see that a CCD pixel:

- tightly constrains A , because each A more or less uniquely maps to an output A_{measured} (as long as the pixel is not saturated beyond A_{\max});
- somewhat constrains t , but only if the exposure is short and the source actually emits photons during the exposure;
- tightly constrains θ if the aperture a is large compared to the wavelength $\lambda = c/\nu$;
- loosely constrains frequency based on detector quantum efficiency η ;

- and does not constrain ϕ , ψ , or ε at all.

$\{A, \alpha, \delta\}$ are the physical quantities which are best-constrained by any given CCD readout; the others are well-defined but imprecise (as per [95]’s definition). Notably, the measurement model still provides information about the source even in the event no detection is made: namely, that the physical quantities must be outside the addressable range of the measurement model. Either the source frequency was lower or higher than the detector response, or it emitted photons at a different point in time, or from a different direction, or below the detectable amplitude threshold.

2.3.2 The limits of information

Due to quantum uncertainty — specifically, Heisenberg’s uncertainty principle — there are fundamental physical limits on how well a given measurement can constrain each the physical quantities of incident information. The most famous of these is that the standard deviation in a wave/particle’s position (σ_x) and in its momentum (σ_p) cannot be lower than a certain bound:

$$\sigma_x \sigma_p \geq \frac{\hbar}{2} \quad (2.9)$$

Measuring a photon with a CCD, for example, necessarily involves gaining positional knowledge of the photon merely because it must have intercepted the optical system. Therefore, the inequality above dictates at least some degree of uncertainty in its momentum, including its directionality. In fact, this uncertainty can be used as an analogue to estimate a limit similar to the Rayleigh criterion. Suppose the uncertainty in photon position is $\sigma_x \approx D$, since we know the photon must have entered the telescope aperture. The uncertainty in the photon’s momentum can be expressed as an uncertainty in its incident angle via $\sigma_p = p\Delta\theta = h\nu\Delta\theta$ by assuming the frequency,

and therefore magnitude of the momentum, is known. Then:

$$\begin{aligned}\sigma_x \sigma_p &\approx D h \nu \Delta\theta \geq \frac{\hbar}{2} \\ D \frac{h}{\lambda} \Delta\theta &\geq \frac{\hbar}{2} \\ \Delta\theta &\geq \frac{1}{4\pi} \frac{\lambda}{D}\end{aligned}\tag{2.10}$$

Compare this to the classically-derived Rayleigh criterion, which states that:

$$\Delta\theta = 1.22 \frac{\lambda}{D}\tag{2.11}$$

The position-momentum uncertainty is one example of a broader class of inequalities relating canonically conjugate variables. One more of such an example is the time-energy uncertainty relation $\Delta t \Delta E \geq \frac{\hbar}{2}$, which can also be equated to a time-frequency uncertainty $\Delta t \Delta \nu \geq \frac{1}{4\pi}$.

Beyond quantum uncertainty, one can only gain so much information about the physical quantities $\theta = [A, \nu, \phi, \psi, \varepsilon, \alpha, \delta, t]$ based on the instrument observables θ_{measured} and the measurement model $\theta_{\text{measured}} = F(\theta)$. Many measurement models are degenerate for multiple reasons: firstly, the dimensionality of any individual measurement is often lower than the dimensionality of the physical quantities it depends upon. In the optical CCD pixel example, the measurement model maps all quantities to a single observable, the electron count. Secondly, the measurement model often maps large regions of θ -space to small regions of θ_{measured} -space, so inverting the measurement model leads to poor constraints (i.e., a high variance) on θ , particularly in the presence of noise.

Due to uncertainties induced via the measurement model as well as various noise contributions, any measurement of a physical quantity provides a possible distribution for that quantity — a posterior — rather than a single value. An appropriate way to consider this is via Bayes' rule:

$$\mathbb{P}(\theta | \theta_{\text{measured}}) = \frac{\mathbb{P}(\theta_{\text{measured}} | \theta) \mathbb{P}(\theta_{\text{measured}})}{\mathbb{P}(\theta)}.\tag{2.12}$$

Prior to some measurement θ_{measured} (possibly but not necessarily vectorial), the physical quantities θ are described by a prior distribution $\mathbb{P}(\theta)$, representing our knowledge or belief about these quantities before obtaining the measurement. This prior distribution could be informed by previous observations, theoretical models, or any other relevant information. Upon obtaining a measurement θ_{measured} , we update our knowledge of the physical quantities by combining the prior knowledge with the likelihood of the measurement given the physical quantities, $\mathbb{P}(\theta_{\text{measured}}|\theta)$. This distribution is calculated via $F^{-1}(\theta_{\text{measured}})$ along with some model of noise sources. The posterior is then $\mathbb{P}(\theta|\theta_{\text{measured}})$, the distribution of the physical quantities given the measurement. This process always represents the first step of the Bayesian network which maps measurements to latent variables to scientific hypotheses, and at a minimum, some approximation of this inference is required for every scientific measurement.

Instrument performance can be characterized by the degree to which its measurements constrain the posterior—e.g., the inverse of the variance of the posterior—as well as the range of inputs for which it is able to constrain the posterior. The former quantity is more formally defined as the Fisher information:

$$\mathcal{I}(\theta) = \int \left(\frac{\partial}{\partial \theta} \log \mathbb{P}(\theta_{\text{measured}}|\theta) \right)^2 \mathbb{P}(\theta_{\text{measured}}|\theta) d\theta_{\text{measured}}, \quad (2.13)$$

an average distribution “tightness” for each possible θ_{measured} at a particular θ . Here, $\frac{\partial}{\partial \theta} \log \mathbb{P}(\theta_{\text{measured}}|\theta)$ is known as the score of θ_{measured} , describing the sensitivity of the instrument to changes in θ at a particular value of θ . For illustration, in the case of a Gaussian with variance σ^2 , $\mathcal{I} = \frac{1}{\sigma^2}$; in the case of infinite variance, $\mathcal{I} = 0$.

Fisher information illustrates a fundamental tradeoff in the case of a fixed number of informational bits per $\theta_{\text{measurement}}$. Since the information content \mathcal{I} depends on the unknown variable θ , one has a choice between being able to make a good (high- \mathcal{I}) measurement of θ for a narrow range of θ versus being able to make a poor measurement of θ for a broad range of θ . For more concrete examples: a larger pixel will measure photons from a broader range of directions, but at the sacrifice of less effectively constraining the directionality of the incident photons. A longer exposure

Table 2.3: Distributing uncertainty among measurement dimensions

Observational strategy	A	ν	α, δ	t	ϕ
Longer exposures	green			red	red
Pixel binning	red	green	red		
Color filter	red	green	green		
Spectrometry		green	red	red	
	increased certainty		decreased certainty		

collects photons from across a broader range of times, but at the expense of worse timing knowledge. A wider spectral response $\eta(\nu)$ collects photons from a broader range of input frequencies, but at the expense of providing less knowledge about the input frequency. A greater electron well depth for each pixel extends the range of observable amplitudes, but at the expense of less constraint on the exact amplitude (due to larger ADU, the analogue-to-digital unit of CCD pixels).

While one can stave off this tradeoff by increasing the number of bits obtained by each measurement — switching from an 8-bit pixel to a 16-bit pixel, for example — the quantum uncertainties discussed above will always limit the amount of information one can gather about a single incident entity, e.g. a single photon. One can therefore never precisely know θ , but must choose how to distribute their uncertainty within each component dimension. For further illustration, Table 2.3 outlines common observational strategies for how these uncertainties are redistributed in practice (for an electromagnetic telescope in this example). Note that this table shows a limited number of such observational strategies, but redistribution among essentially arbitrary dimensions is enabled by combining multiple compatible strategies, or by using other strategies not shown.

2.3.3 Composing sets of measurements

A widely-applicable solution to the Fisher information tradeoff is to use tightly-constraining measurements — each of which may only address a small portion of

θ -space — but to take a set of many measurements which together address a wide breadth of θ , thus gathering more informational bits even at a fixed number of bits per measurement. Conveniently, the Fisher information conditioned on N independent but identically distributed measurements is simply the sum of the information from each measurement:

$$\mathcal{I}(\theta) = \sum_{i=1}^N \mathcal{I}_i(\theta) \quad (2.14)$$

Therefore, if the Fisher information is low for a particular portion of θ -space, one can simply add an additional measurement with high information content in that region. In essence, the quantity of recoverable data about a given astrophysical source depends on the quantity N of independent measurements of the source as well as the quality \mathcal{I} of each measurement.

Fisher information allows one to rederive a well-known result regarding the scaling of signal-to-noise ratio (SNR) due to taking a set of N identical (but independent) measurements, assuming Gaussian distributions of the measurements of θ :

$$\begin{aligned} \frac{1}{\sigma^2} &= \mathcal{I} = \sum_{i=1}^N \mathcal{I}_i = \frac{N}{\sigma_i^2} \\ \text{SNR} &= \frac{1}{\sigma} \propto \sqrt{N} \end{aligned} \quad (2.15)$$

This method of measurement composition relies on a degree of assumed symmetry among one or more of the dimensions of θ . This limitation arises from the fact that any source emits a finite quantity of informational bits. Where the detector measurement model is inefficient, more raw information must be averaged in order to recover a high-information measurement, which “blurs” the information content within the averaged dimension. In image stacking, for example — where SNR is increased by stacking images across time or across α, δ — one has to make the assumption that the signal is either somewhat atemporal or somewhat isotropic,¹¹ respectively. Other compositional methods include color imaging (discussed below, and which also assumes isotropy or atemporality); interferometry, which assumes isotropy; and astrometry,

¹¹Not entirely isotropic — the signal just has to look sufficiently similar as viewed from the vantage point of each detector. Of course, this assumption often applies for astronomical applications.

which assumes isotropy. Even just increasing the aperture size or exposure length of an instrument falls under this umbrella, as those respectively assume the source is isotropic across the width of the aperture and atemporal during the length of the exposure.

For a more advanced example of tiling measurements across θ -space, consider the case of capturing a color movie of a scene. First, an individual image is formed via a CCD comprised of many small pixels (implying a tight directionality constraint) which span a grid over a broader range of α and δ . This single monochromatic image actually represents a large number of individual measurements. The Fisher information of a particular pixel will be zero in most cases — where incident photons do not fall on that pixel in particular — but will be high for the photons that it does measure. In this way, the overall Fisher information is much higher than if the CCD were comprised of a single, larger pixel, which would have low but non-zero information in all cases across the CCD; therefore the directional uncertainty of components of the scene are reduced. Next, a color image is generated by using filters to create tighter distributions in ν , and multiple filters are used (e.g. a Bayer filter with r,g,b-sensitive pixels:  so as to still span a broad range of ν -space. Next, to create a movie, one can take many quasi-instantaneous (tightly-constrained in time) images spanning across t -space. In essence, a color movie is a multi-dimensional measurement set spanning across $\alpha\delta\nu t$ -space, depicted in Fig. 2-5. This example illustrates that one can get fundamentally new information merely by simply linearly stacking different measurements across a range of θ -space, even where each measurement has a low-dimensional readout.

2.4 Method

Astrophysical instrumentation design has historically focused on inventing and improving standalone instruments optimized for a particular portion of θ -space. However, multi-messenger astronomy requires a cohesive network of observatories, which have non-simple interactions.

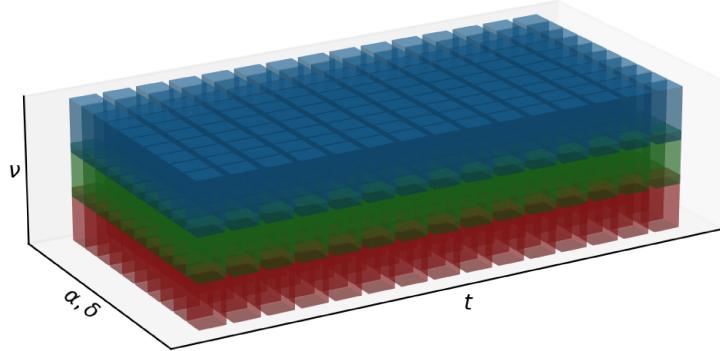


Figure 2-5: A color movie represented in θ -space is a hypercube of pixel measurements.

This problem is not unique to astronomy alone: the field of data fusion deals with incorporating multimodal data in many domains, such as remote sensing, robotics, image processing, and artificial intelligence [95].¹² Modality here refers to the degree of novelty between the dimensions along which different measurements occur, whether through heterogeneity of detectors, conditions under which the detections occur, or experiments conducted by the detectors. A key insight from the field of data fusion is that insights from multimodal measurements can often be greater than the sum of the parts, either through superlinear¹³ scaling of uncertainty with collected data or emergence of new insights made feasible only with such data fusion [95].

To better understand the difference between traditional and multi-messenger astronomy, we turn to probabilistic graph models known as Bayesian networks.¹⁴ In all inference problems, whether those conducted via the human brain using the five senses, or by autonomous robots with cameras and sensors, or scientific inference performed with astronomical instruments, there is a defined flow of information from *measurements* of the environment to *latent variables* describing the events that occur

¹²Even evolution has tackled this problem many times over. Consider humans' five senses, which are multi-messenger detection mechanisms that have been optimized for survival through an evolution-driven approximation of gradient descent. The constraints faced by evolution are not too different either: inference based on data fusion is energetically and computationally expensive for biological and computational intelligences alike.

¹³In contrast with Eqn. 2.15, which notes that for unimodal data, SNR scales sublinearly with the amount of collected data.

¹⁴Bayesian networks are probabilistic graph models that explicitly use Bayesian inference. Bayesian inference is not a prerequisite feature in this case: in fact, in many components of the network, the inference is strictly non-Bayesian. However, the latter term is used here as it is the more widespread convention.

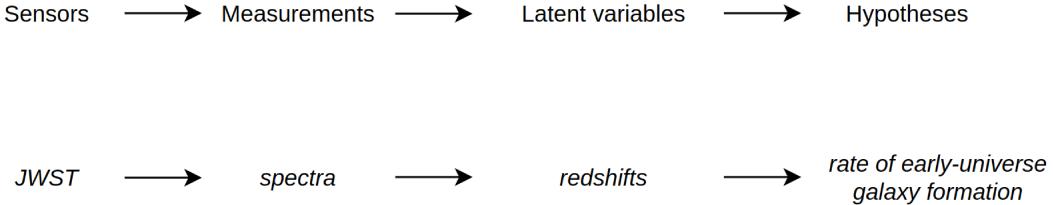


Figure 2-6: The Bayesian network of JWST characterizing early-universe galaxy formation.

to *hypotheses* about the underlying causal mechanisms. Bayesian networks can track this flow of information during scientific inference with multiple steps.

The structure of Bayesian networks employed in traditional astronomy are often simple in nature. Even for complicated scientific inquiries, there is a well-defined, linear flow from measurements to hypotheses. First, a particular instrument will take one measurement (or multiple similar measurements). These measurements then inform the observer of latent variables, which are the particular properties of the object(s) which were observed. The latent variables are then used by the observer to infer particular hypotheses about the phenomenon in question. Fig. 2-6 shows an example involving JWST’s science goal of measuring the rate of early-universe galaxy formation. JWST will use not just one but rather many spectra of different galaxies to determine the early-universe galactic population, but this particular Bayesian network is linear regardless, as the data fusion follows Eqn. 2.14. For a more advanced example from traditional astronomy, consider the Bayesian network for the black hole image created by the Event Horizon Telescope (EHT), shown in Fig. 2-7. The process of sensor fusion is much more involved than in the JWST example, requiring complex heterodyne interferometry techniques and significant compute [68].¹⁵ All the same, the information flow is highly linear, as all the measurements are of the same nature.

In contrast, Bayesian networks involved in multi-messenger astronomy are often complex and involve nonlinear steps such as sensor tasking or cueing. The networks

¹⁵This represents a case of measurement-level fusion. The same applies to optical interferometry, whether through heterodyne interferometry as mentioned or through direct detection where a beam combiner sends light from multiple apertures onto a single detector.

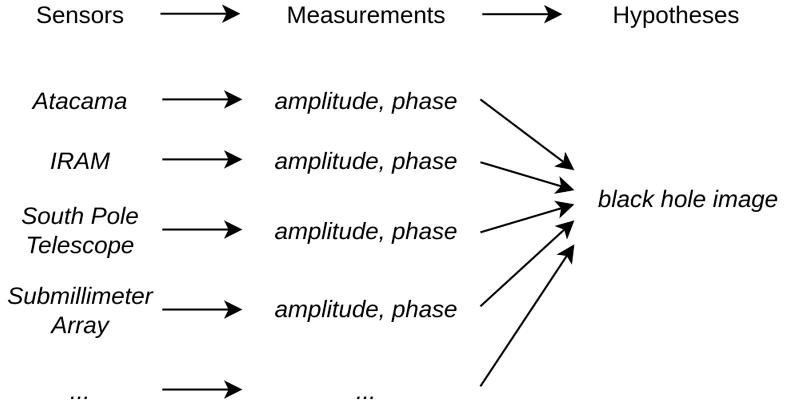


Figure 2-7: The Bayesian network of the EHT taking an image of a black hole accretion disk.

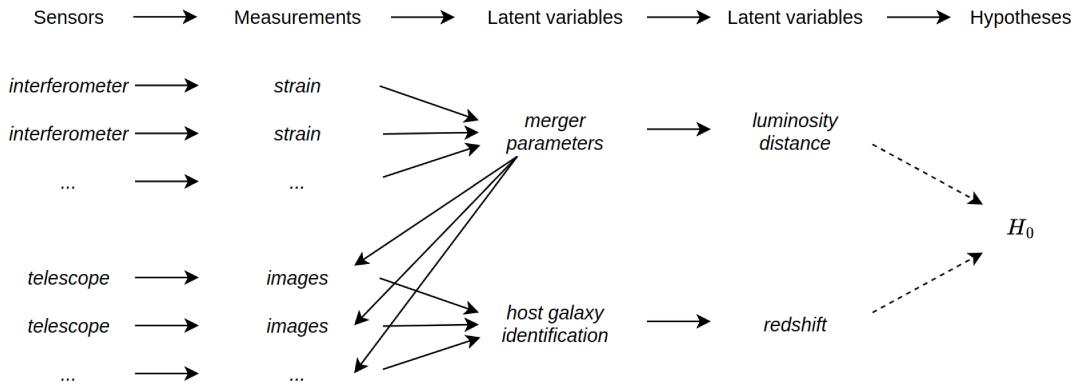


Figure 2-8: The Bayesian network for the binary neutron star merger case study. This case study regards measuring H_0 , the Hubble parameter, with joint GW/EM detections of binary neutron star mergers. In this case, the GW detections cue follow-up observations with EM telescopes.

for the case studies in this thesis are shown in Fig. 2-8 and 2-9. These networks cannot rely on Eqn. 2.14 for data fusion, because measurements of disparate types are not independently and identically distributed. Instead, the general form of arbitrary methods of data fusion follows the conditional information:

$$\mathcal{I}(\theta) = \mathcal{I}(\theta_i) + \sum_{i=2}^N \mathcal{I}(\theta_i | \theta_1, \dots, \theta_{i-1}) \quad (2.16)$$

where θ_i represents measurements from sensor i for $i \in [1..N]$ total measurements.

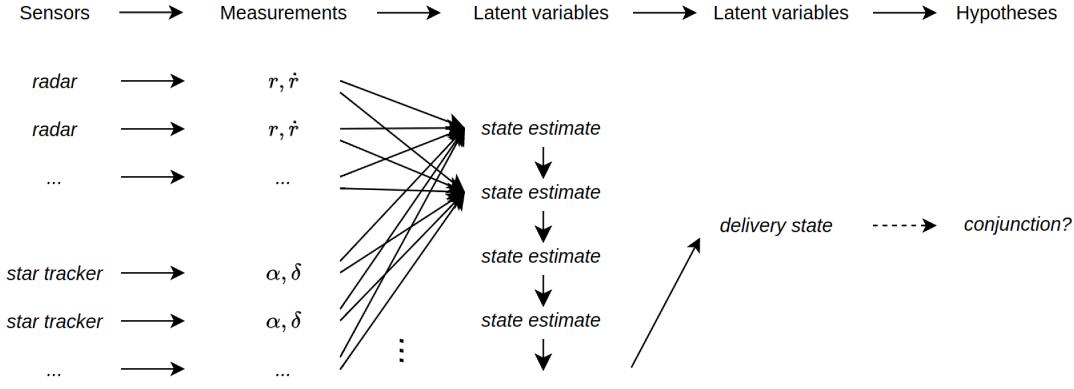


Figure 2-9: The Bayesian network for the resident space object case study.

This case study regards orbit determination of objects in LEO. In this case, a state estimation filter is used to perform inference with radar and star tracker observations to determine the object's state over time.

One can pose astrophysical instrumentation design as an optimization problem:

$$\begin{aligned}
 & \underset{\theta}{\text{maximize}} \quad \mathbb{E}(\mathcal{I}(\theta)) \\
 & \text{subject to} \quad g(\theta) \leq c,
 \end{aligned} \tag{2.17}$$

where one wishes to maximize the expected value of information \mathcal{I} from a set of planned measurements θ given resource constraints $g(\theta) \leq c$. In principle, this is how traditional instrumentation design occurs, where θ represents measurements from a single instrument rather than of many.¹⁶ The key difference lies in calculating $\mathcal{I}(\theta)$: essentially, in understanding how data fusion improves knowledge. For traditional or unimodal astronomy, data fusion is necessary but relatively straightforward, as fusion always occurs at the measurement level (and generally follows Eqn. 2.14). For multi-messenger astronomy, data fusion necessarily involves feature-level fusion, where disparate measurements or latent variables are combined into other specific latent variables or hypotheses. Feature-level fusion requires Eqn. 2.16 to track the flow of information, which is significantly more challenging to compute than Eqn. 2.14 due to the dependencies of conditional probabilities.

¹⁶It may not be typically expressed as an optimization problem, but optimization under constraints certainly describes most engineering efforts!

While one ultimately cares about the hypotheses that can be formed from various astronomical measurements, the networks representative of traditional astronomy allow the quality and quantity of measurements to pose a convenient proxy for value, specifically because Eqn. 2.14 is linear and can directly compare information from similar measurements. This proxy is especially helpful during the process of instrumentation design. It abstracts the design away from the particulars of the Bayesian network, which otherwise necessitates concrete implementation of models of the astrophysical phenomena in question. Instead, a science traceability matrix defines the flow-down from science goals to instrument requirements [162]. Rather than optimizing for scientific hypotheses directly, one can use the science traceability to identify a particularly promising window of θ -space to measure, and optimize for high-quality measurements in that window alone.

For multi-messenger astronomy, however, no such proxy is applicable both because Eqn. 2.16 is nonlinear and because the quality of disparate measurements cannot be directly compared. Instead, a Bayesian network is required, which necessarily involves an assumed model of the phenomena in question: the source must be modelable in order to determine how challenging it is to detect in a multi-messenger context.

Existing works in the literature have addressed using information theory via Bayesian networks for multi-sensor data fusion and for planning observations, particular in the contexts of astronomy and state estimation, as is relevant for the case studies considered in this thesis. Review papers including [95] and [165] provide a background to the central challenges of sensor fusion and the various tools that have been used to approach the task. [88] is a more applied example that uses a Bayesian network to optimize JWST exoplanet atmosphere observations by optimizing for mutual information per degree of freedom, a metric similar to the Fisher information considered in this work. [72] similarly uses Bayesian inference to identify gravitational-wave host galaxies with GW and EM observations, the application considered in the binary neutron star merger case study. It does not explicitly track information, but rather the expected probability of successful identification. [100] focuses more abstractly on state estimation from an information theory perspective, and describes how the

Kalman filter (the state estimation filter used in the resident space object case study) relates to the concept of mutual information.

Other works have investigated sensor optimization from information fusion more generally, and not from an astronomical context. [166] describes the same method of probabilistic graph models used in this work. Rather than optimizing sensors by explicitly performing inference with the Bayesian network, a proxy of “sensor synergy” is defined. Specifically, a synergy coefficient relates the information derived from each pair of sensors in the network compared to the information derived from either sensor alone. This proxy enables a computationally-efficient greedy approach to sensor optimization. The optimization then does not require new inference for each unique network combination, but rather approximates information using a synergy graph based on a sensor synergy matrix. One limitation of the work is that it assumes each sensor pairing is conditionally independent, which often will not apply. Similarly, [161] use an information-theoretic approach to optimize sensor placement for target localization and tracking, a “flat” network consisting of a single inference step. They employ a sensor selection heuristic which approximates the information content derived from multiple sensors.

The rest of this thesis demonstrates astronomical data fusion for two case studies. Both case studies rely on a statistical inference approach,¹⁷ but separately address the emergence and uncertainty scaling properties of data fusion. An outline of the instruments considered in each of these case studies are given in Fig. 2-10.

Case study 1 (Chapter 3) regards using joint electromagnetic and gravitational-wave measurements of binary neutron star mergers to infer the Hubble constant, the universe’s expansion rate. Since merger observations via either messenger alone do not permit a measurement of the Hubble constant, this represents a form of emergence via data fusion. In this case, it is shown that the inference computation can be sufficiently parallelized so as to simultaneously compare a wide variety of observatory networks. Rather than an $\mathcal{O}(N^2)$ scaling of inference computations with the number

¹⁷The particular inference approach is explicitly non-Bayesian, instead relying on linearizations of the posterior corresponding to the Fisher information matrix (FIM) method.

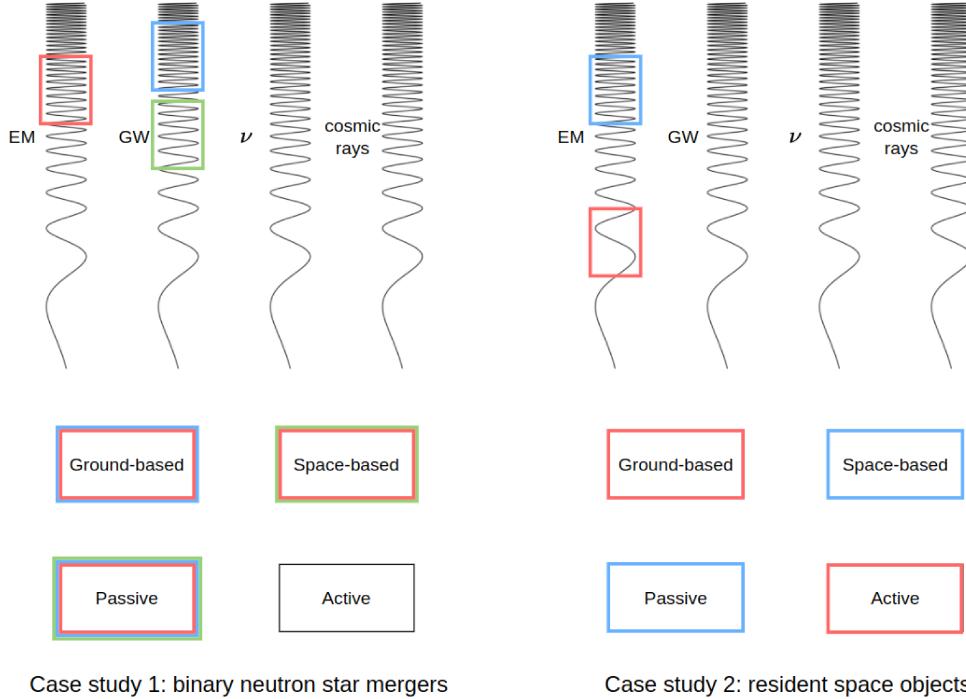


Figure 2-10: Breakdown of messengers and frequency ranges by case study.

Case study 1 uses passive ground- and space-based detectors in both the EM and GW spectra. Case study 2 uses active ground-based instruments and passive space-based optical instruments.

of considered instruments, as is typical, this work achieves $\mathcal{O}(N)$ scaling by redefining the network graph so as to not require explicit computation of conditional information.

Case study 2 (Chapter 4) regards determining the orbits of resident space objects (e.g., artificial satellites and on-orbit debris) using ground-based radar and space-based optical measurements. Though this case study only considers photons as a messenger, the concepts of data fusion involved in multi-messenger astronomy equally apply. In this case, one performs orbit determination regardless of whether the sensor network is homogeneous or heterogeneous. Instead of demonstrating emergence, this case study demonstrates uncertainty scaling with multimodal data.

Chapter 3

Multi-messenger astronomy of binary neutron star mergers

3.1 Science background and motivation

A binary neutron star (BNS) merger is a stellar collision that occurs as two neutron stars orbit about a combined center of mass and gradually inspiral due to energy loss to gravitational radiation. Along with gravitational wave spectra, these collisions produce electromagnetic emissions via kilonovae (KN) associated with their post-merger matter ejecta. BNS mergers occur rarely throughout the local universe; in a galaxy similar to the Milky Way in terms of mass, age and star content, such an event will only occur once every 10^6 years on average [160]. As a result, only two of such events in nearby galaxies have been detected via gravitational waves to date: GW170817 and its electromagnetic counterpart AT 2017gfo during LIGO/Virgo's¹ 2nd observing run in 2017, at a distance of 40^{+8}_{-14} Mpc [51], and GW190425 (not successfully detected in the electromagnetic spectrum) during LIGO/Virgo's 3rd observing run in 2019, at a distance of 159^{+69}_{-71} Mpc [8].

Multi-messenger studies of BNS mergers have the potential to provide insight towards a variety of open questions in stellar physics, particle physics, and cosmology,

¹LIGO, the U.S.'s Laser Interferometer Gravitational-Wave Observatory, and Virgo as the E.U.'s equivalent.

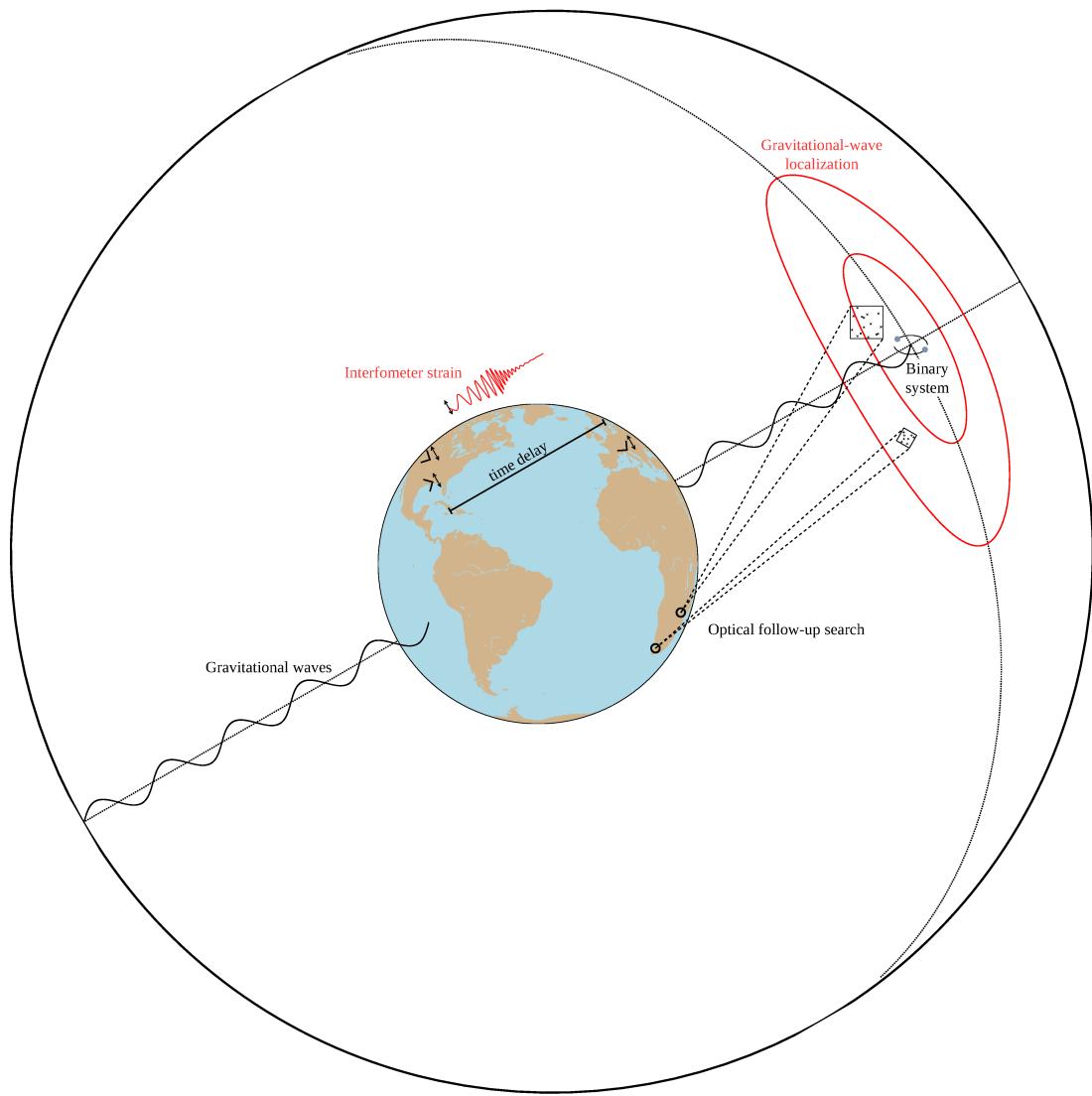


Figure 3-1: Joint gravitational-wave and electromagnetic detection of a binary neutron star merger.

Interferometers detect the gravitational “chirp” of the merging binary, which is used to approximately localize the source. Telescopes then take follow-up imagery of the region to identify the exact origin and determine properties of the corresponding kilonova.

and will serve as a further test of our fundamental theories such as general relativity. With regards to astrophysics, such observations can constrain the neutron star equation of state, determine the internal stellar structure of neutron stars [126], and describe the origin of most r-process elements² in the universe [75]. Neutron stars also provide an exotic particle physics laboratory, achieving temperatures and densities unobservable through Earth-based experiments alone; observations may therefore enable tests of the equation of state of supranuclear matter [90]. The cosmological applications of multi-messenger BNS study, however, are arguably the most widely relevant to physics as a whole. Gravitational-wave and electromagnetic observations provide independent measurements of the distance and redshift of each merger, respectively. These distance-redshift measurements can calibrate and greatly extend the range of the extragalactic distance scale (the “cosmic distance ladder”) [80], a widely-used tool across all of astronomy. In the local universe, sufficiently many of these data could directly resolve the present $4 - 6\sigma$ [153] disagreement in estimates of the Hubble constant (H_0), the universe’s current rate of expansion [134]. More sensitive observations conducted through a range of higher redshifts will permit the construction of a Hubble diagram from which parameters of the standard model of cosmology³ can be inferred, including the dark matter and energy content of the universe, as well as the dark energy equation of state [28]. These data may either validate the standard model or lead the way towards new physics.

In order to fully realize the scientific promise of multi-messenger observations of BNS mergers, we must have a capable observational network comprised of both gravitational-wave interferometers and electromagnetic telescopes which can collaboratively detect and follow-up on these short-lived⁴ transient events. The central challenge of detecting these events is two-fold: first, since gravitational wave interferometers act as omnidirectional antennae, the localization of detected sources is

²r-process elements, including roughly half those heavier than iron, are thought to be primarily or in some cases exclusively generated by BNS mergers [141].

³ Λ CDM, or the Lambda cold dark matter “standard model”, is further discussed in Sec. 3.1.1.

⁴While BNS emit gravitational waves throughout their lifetime, the most detectable waves, in terms of amplitude and frequency, occur in the seconds leading up to the merger itself. Kilonovae peak in brightness on the order of hours to days later, but afterwards quickly dim and redden [52].

poor, constrained only by the gravitational wave time delay between geographically diverse interferometers and differences in observed amplitude due to the orientation of the antenna patterns. Second, once a source is detected, the corresponding telescope network must have both sufficient sensitivity and overall throughput to quickly search the interferometer-based localization regions before the kilonovae dim beyond the point of feasible observability.

A chapter outline is as follows: Sec. 3.1.1 will further elaborate the background for scientific inferences enabled by BNS mergers, which will motivate a specific definition of network performance. Sec. 3.1.3 will then describe how BNS gravitational-wave and electromagnetic spectra are generated and observed. Sec. 3.2 will describe how the performance of arbitrary detector networks is simulated, starting from generating a population of random mergers from realistic parameter distributions (Sec. 3.2.1), calculating their orientation-dependent gravitational-wave and electromagnetic waveforms over time (Sec. 3.2.2, 3.2.4), determining the GW-derived sky localization regions (Sec. 3.2.3), tasking telescopes to follow up on gravitational-wave sky localization regions (Sec. 3.2.5), and then determining the overall likelihood and rates of merger detection (Sec. 3.2.6). The results and discussion sections, Sec. 3.3 and 3.4, will then detail and analyze simulation outputs.

3.1.1 Cosmological context for the Hubble parameter

One of the most intriguing applications of multi-messenger BNS observations is independently measuring the Hubble parameter. In this case, “independent” means that the measurement does not rely on the assumptions and types of observations tied to current measurement methods, which gives these observations the potential to resolve discrepancies across those methods.

The Hubble parameter H describes the relative rate of expansion of the Universe. Specifically, a comoving object with a proper distance r from a comoving observer will have a recession velocity given by Hubble’s law:

$$v = Hr, \quad (3.1)$$

where H has the SI unit of s^{-1} . It is more frequently expressed in units of $\text{km s}^{-1} \text{Mpc}^{-1}$: a velocity of expansion (km/s) at a certain distance from an observer (Mpc). The Hubble parameter is an outcome of a limited set of assumptions: (1) that the universe is homogenous, isotropic, and expanding, (2) that a cosmological equation of state $w = \frac{p}{\rho}, w \in \mathbb{R}$ applies⁵, where the pressure density p and mass density ρ of different energy components of the universe follow fixed ratios, and (3) that the formulation of general relativity applies. By deriving the Robertson-Walker metric from (1) and applying that and the cosmological equations of state (2) to Einstein's field equations $G_{\mu\nu} + \Lambda g_{\mu\nu} = \kappa T_{\mu\nu}$ (3), the Friedmann equations can be derived. These equations describe the dynamical evolution of the universe in terms of its scale factor a (the ratio of its proper size at a specific time to its proper size at an arbitrary reference time), mass density ρ , and pressure density p :

$$H^2 = \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G + \Lambda c^2}{3} \rho - \frac{kc^2}{a^2} \quad (3.2)$$

$$\dot{H} + H^2 = \frac{\ddot{a}}{a} = -\frac{4\pi G}{3} \left(\rho + \frac{3p}{c^2} \right) + \frac{\Lambda c^2}{3} \quad (3.3)$$

where G is Newton's gravitational constant, Λ is the cosmological constant, c is the speed of light, and k is a dimensionless quantity describing the spatial curvature of the universe (measured to be near-zero). As evidenced by the Friedmann equations, the Hubble parameter is homogeneous but varies with time; its present value is denoted as H_0 . Note that the scale factor a can be rewritten in terms of a redshift z as per the relation $1 + z = 1/a$.

The Friedmann equations have wide-reaching consequences for our understanding of the evolution of the universe, having implications for its age and ultimate fate, such as whether it collapses or expands indefinitely. These equations are also how we determine the composition of the universe in terms of its overall densities of ordinary

⁵An example cosmological equation of state is given by the ideal gas law $p = \rho RT$, which leads to $p/\rho \equiv w = RT \approx 0$ assuming a “cold” gas where $\sqrt{RT} \ll c$. Other cosmological equations of state apply for different energy components of the universe, such as radiation ($w = c^2/3$, or just $w = 1/3$ in natural units with $c = 1$), dark matter, and dark energy (generally assumed to be $w = -1$).

Table 3.1: WMAP Λ CDM parameters [85]

parameter	estimate
present Hubble parameter, H_0	$69.7 \pm 2.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$
age of the universe, t_0	$13.76 \pm 0.11 \text{ Gyr}$
baryon density / critical density, Ω_b	0.0463 ± 0.0024
cold dark matter density / critical density, Ω_c	0.233 ± 0.023
dark energy density / critical density, Ω_Λ	0.721 ± 0.025

matter, dark matter, and dark energy; it is how we know ordinary matter comprises only 5% of the energy density of the observable universe [85]. Measurements of the Hubble parameter, the primary dynamical outcome of the Friedmann equations, therefore have central importance to cosmology.

Our present most widely-used model of the universe is known as the flat Λ CDM model (Lambda cold dark matter), frequently referred to as the “standard model” due to its success in predicting the acceleration of the universe, the existence and isotropy of the cosmic microwave background (CMB), and the large-scale structure of galaxies. This model is derived from the Friedmann equations by assuming $k = 0$ and that the density ρ is comprised only of baryonic matter, cold dark matter, and dark energy. The rest of this thesis will assume the Λ CDM model for any cosmological calculations, relying in particular on the nine-year Wilkinson Microwave Anisotropy Probe (WMAP) 2013 measurements of Λ CDM parameters [85] outlined in Table 3.1.

Tension in the present measurements of H_0

Previous sections have alluded to the discrepancies between different methods of measuring H_0 , in particular a $4 - 6\sigma$ [153] difference between our two main methods: direct measurements from the cosmological distance ladder and indirect measurements from anisotropies in the CMB. From the mid 2010s onward, the distance ladder and CMB methods have converged to disagreeing values, particularly with the help of the WMAP and Planck, two spaceborne missions that have mapped the CMB. The evolution of these measurements is depicted in Fig. 3-2. A variety of hypotheses aiming to explain the measurement tension have been put forth, ranging from in-

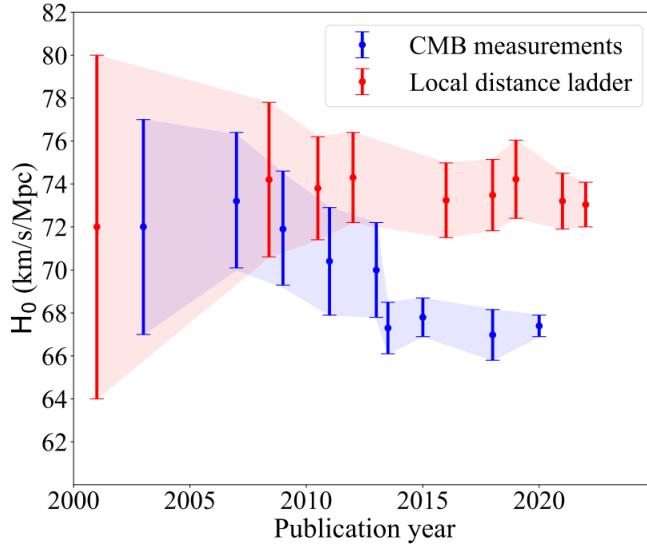


Figure 3-2: 21st century measurements of H_0 .

The discontinuity in the CMB measurement in 2013 corresponds with the first data release of the Planck mission, launched in 2009 [33]. Credit: [89].

vestigations of the systematic miscalibration of the local distance ladder [94], [44], systematic miscalibration of CMB measurements [9], to possible physics beyond the Λ CDM model [128], [121]. Currently, physics beyond Λ CDM is the preferred explanation, primarily because the size and consistency of the discrepancy seems to rule out any single systematic error as the root cause [158].

Present measurement methods

Using the distance ladder, directly measuring the Hubble parameter requires taking simultaneous measurements of the redshift of an object and its distance. Optical measurements of type Ia supernovae can provide these values for galaxies at various distances. The spectral lines of the supernovae indicate its redshift z . The brightness of the supernovae indicate the luminosity distance: the peak luminosity of type Ia supernovae is reasonably consistent, and so varies primarily with inverse square of the distance from the observer (and other confounding effects such as galactic extinction, which can be accounted for to some extent). This observational method can also be confounded by the peculiar velocity of the galaxies, since galaxies typically have

some non-zero velocity in the comoving frame due to local gravitational affects from nearby extragalactic structures. By assuming that galaxies have a peculiar velocity distribution with zero bias, taking enough of these measurements allows the effect of peculiar velocity to be averaged out. Our best type Ia supernovae measurement of H_0 is from the SH0ES (Supernovae H_0 for the Equation of State) Collaboration [46], providing a value of $H_0 = 73.04 \pm 1.04 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

The angular power spectrum of the CMB’s anisotropies determine the matter, dark matter, and dark energy densities of the early universe, along with its spatial curvature. These data do not directly determine H_0 , but H_0 can be inferred by assuming a cosmological model (such as ΛCDM). Therefore this method is highly dependent on the assumed cosmological model. The present best measurements of H_0 with this method are from the Planck mission [124], providing a value of $H_0 = 67.4 \pm 0.5 \text{ km s}^{-1} \text{ Mpc}^{-1}$.

3.1.2 Gravitational-wave standard sirens as a new measurement of the Hubble parameter

Another means of measuring the Hubble parameter is enabled by multi-messenger astronomy through joint gravitational-wave and electromagnetic detection of binary neutron star mergers [87]. This method has the potential to explain or resolve the Hubble tension, as it relies on neither a local distance ladder nor an assumed cosmological model, instead using mergers as gravitational-wave “standard sirens.” Gravitational-wave detection provides a direct measurement of the luminosity distance of merger events, and corresponding electromagnetic detection indicates the redshift of the host galaxies of the mergers.

First, inference on the gravitational wave signals from these mergers provides an estimate of their luminosity distance: the binary mass estimates define the amplitude of the gravitational wave strain, which falls with the inverse square of the luminosity distance. This inference step also determines the sky localization of the merger. Second, follow-up electromagnetic observations search the sky localization region to

identify potential transient signals that could be associated with the merger. Assuming false positives can be ruled out, the transient signal created by the merger allows it to be associated to a particular host galaxy. The redshift of the host galaxy (and therefore its merger) can be obtained.

The uncertainty of the Hubble constant measurement depends primarily on the quantity rather than quality of joint detections. On the EM side, the redshift measurement does not depend on the kilonova (KN) signal itself, only association with a particular host galaxy. On the GW side, in principle more sensitive detectors can better constrain the luminosity distance d_L of a given merger. As a first-order approximation one might expect $\Delta d_L/d_L \approx 1/\sqrt{\text{SNR}}$ [125]. However, in practice, the distance uncertainty is limited by the degeneracy between a merger's distance and inclination ι , which is challenging to resolve just based on the GW strain [32]:

$${}^6\text{GW STRAIN AMPLITUDE} \propto \frac{\cos \iota}{d_L} \Rightarrow d_L \propto \cos \iota. \quad (3.4)$$

One method to address this degeneracy is by using GW polarization to measure inclination; for GW170817, this method constrained inclination to $\cos \iota \leq -0.54$ [51]. Another means to determine the inclination is through fitting EM observations to a inclination-dependent KN emission model. However, [102] discounts this method as being a significant contributor to driving down H_0 uncertainty since it would be challenging to avoid introducing systematic bias through the KN emission model, and even in optimistic scenarios the fraction of observations where this method could apply is small.

Since H_0 uncertainty is dominated by the quantity rather than quality of detections, a natural performance measure of a GW/EM observatory network is the rate of BNS mergers successfully detected in both spectra. However, owing to the significant uncertainty in the rate density of BNS mergers, a simpler measure is the fraction of mergers within a certain distance that are detectable, or relatedly the horizon distance of the network. The horizon distance of a GW network is traditionally defined

⁶This equation shows the “cross” polarization of the strain amplitude; the “plus” polarization is proportional to $\frac{1+\cos^2 \iota}{d_L}$.

as the luminosity distance at which 50% of sources of a fixed mass are detectable, averaged across the sky since the directional sensitivity varies between networks. In this case, the horizon distance is defined as the 50% range for both GW and EM detection.

[102] suggests that ~ 50 joint GW and EM merger detections are required to derive a new measurement of H_0 within 4% (the degree of tension in existing measurements), though estimates range from $\mathcal{O}(10)$ [135] to $\mathcal{O}(100)$ [7]. Between O1-O3 (the 1st-3rd LIGO observing runs) which cumulatively lasted 1.5 years, only one of such detections was made (GW170817 / AT 2017gfo). Though this lone detection provided a measurement of $H_0 = 70.0_{-8.0}^{+12.0} \text{ km s}^{-1} \text{ Mpc}^{-1}$ [7], the current detection rate suggests that 4% uncertainty would not be reached before the end of the century absent of further network improvements, even with a 100% duty cycle⁷.

Beyond constraining the present value of the Hubble parameter, observations of BNS mergers at higher redshifts, which require more sensitive observatories than we have today, promise to measure other Λ CDM parameters by assessing the Hubble parameter over time. With BNS observations out to $z \sim 2$ specifically, the dark energy equation of state parameter w and the matter density of the universe Ω_M may be measured to percent-level uncertainty in the coming decades [164] [31], though in principle most if not all Λ CDM parameters are estimatable given enough observations across a range of sufficiently deep redshifts.

3.1.3 Compact binary coalescence (CBC) and EM/GW emissions

Compact binaries are systems comprised of two astrophysically small objects (radius $r \ll R_\odot$, the radius of the Sun) such as white dwarfs, neutron stars, and stellar-mass black holes, which orbit together around a shared center of mass. Owing to the

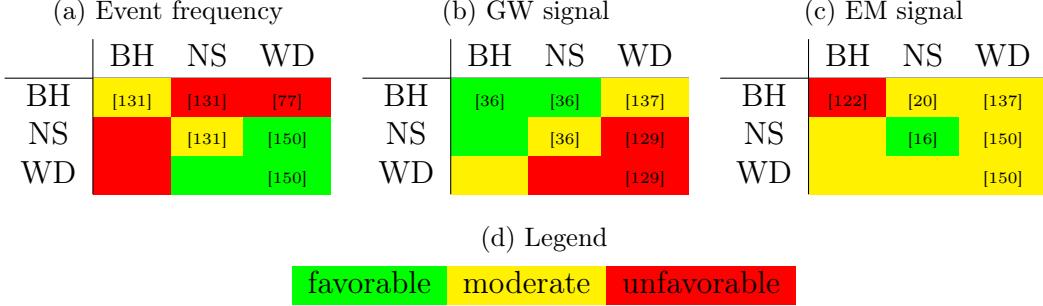
⁷The best duty cycle thus far was achieved by LLO in O3, at 77.0% [26]. However, offline time spent on detector upgrades between runs means that the overall duty cycle from 2015-2023 is only at $\sim 20\%$, which implies \sim four centuries before 50 sources are detected. Of course, continuous upgrades and the addition of new detectors to the network such as Japan's KAGRA will drive down the time considerably.

gravitational wave radiation of the combined system, these objects gradually inspiral as they lose orbital energy. Eventually, they coalesce and become another white dwarf, neutron star, or black hole, dependent primarily on their combined mass. CBCs are the primary targets for gravitational wave observatories. The first successful CBC observation, GW150914 — also the first-ever gravitational wave detection — was of a binary black hole in 2015 [5].

In principle, each combination of compact objects that can comprise these systems may provide the desired distance-redshift measurements; however, BNS (binary neutron star) mergers are the best candidates for multi-messenger detection owing to their relatively bright kilonovae emissions as well as their moderate mass range that create strong electromagnetic and gravitational-wave emission respectively. The factors that play into observability are outlined qualitatively in Table 3.2, including (a) the frequency of these events, in terms of a rate density typically expressed in units of $\# \text{ Gpc}^{-3} \text{ yr}^{-1}$; (b) the GW signal, in particular the amplitude of the strain and secondarily the GW frequency; and (c) the predicted or observed bolometric magnitude of the EM signal. To mention other top candidates: BBH (binary black hole) mergers are more massive than BNS mergers, and thus can be detected to a further distance by gravitational wave interferometers. However, owing to their event horizons, BBH create essentially no electromagnetic emissions apart from their relatively meager accretion disks [122]. Similarly, much of BH-NS (black hole / neutron star binary) post-merger mass is constrained within the event horizon, so these events are significantly less bright in the electromagnetic spectrum than BNS [20]. BH-NS are also anticipated to be less common than either BNS or BBH [131]. The combinations of binary systems involving white dwarfs (WDs) are predicted to be more common than any others [150], but their gravitational wave spectra are currently prohibitively difficult to detect owing to a low frequency and amplitude [137].

Note that these observability traits are coupled; for example, a low event frequency can be compensated for by a strong GW signal, which permits an extended detection range. Furthermore, as the detection range increases, the EM detection becomes correspondingly more challenging than the GW detection due to a $1/d_L^2$ scaling

Table 3.2: Qualitative observability of compact binary coalescence pairs



versus a $1/d_L$ scaling of amplitude. Regardless, this thesis will focus exclusively on BNS mergers owing to their favorable traits for multi-messenger detection. Even when other binary types may be detectable, the efficacy of an observatory network at discovering BNS mergers will represent a reasonable proxy for network performance; a network optimized for study of BNS mergers exclusively will also be capable for studying BH-NS mergers, for example.

Gravitational wave emissions of binary neutron star mergers

For the sake of fully characterizing BNS merger gravitational wave emissions, each merger is describable by a set of 15 intrinsic and extrinsic parameters which are detailed in Table 3.3. Intrinsic parameters belong to the state space of the binary system, as they determine the dynamical evolution of the binary. Extrinsic parameters are related to an external observer’s point of view, such as its distance from the observer and orientation in the observer’s sky. Both parameter sets influence observability. Intrinsic parameters determine the frequency and amplitude of gravitational wave and electromagnetic emissions. Extrinsic parameters such as redshift and luminosity distance then modulate the resulting signal.

Based on these parameters, gravitational waves are emitted from binary mergers across three dynamic phases, including the inspiral, merger, and ringdown, illustrated in Fig. 3-3. The inspiral emissions are known as the gravitational wave “chirp” due to the increasing frequency and amplitude during dynamic evolution; the chirp is

Table 3.3: Binary neutron star parameters

symbol	description	
m_1	mass of the first ^a binary object	intrinsic
m_2	mass of the second ^a binary object	
\vec{s}_1	spin vector of the first ^a object (3 parameters)	
\vec{s}_2	spin vector of the second ^a object (3 parameters)	
ϕ_c	phase at the time of coalescence	extrinsic
ψ	polarization angle	
ι	orbital inclination of the binary system ^b	
d_C	comoving distance	
α	right ascension	
δ	declination	
t_c	geocentric time of coalescence	

^a The designation of the first versus the second object is arbitrary, but general convention identifies them to satisfy $m_1 \geq m_2$.

^b Orbital inclination is always measured with respect to a specific reference plane. In this case, the reference plane is defined by a normal vector pointing along the observer line of sight to the binary.

generally the most detectable component of the emissions, the observable portion of which might last $\mathcal{O}(10)$ s [36]. Fig. 3-4 shows the Fourier transform of the chirp, clearly illustrating the rise in amplitude and frequency as the inspiral phase progresses.

While the exact waveform of the chirp in terms of its strain amplitude⁸ $h(t)$ and phase $\phi(t)$ is only numerically solvable in the framework of numerical general relativity (NR), analytical solutions to the chirp are derivable by assuming a post-Newtonian (PN) relativistic approximation where the equations of motion acquire a Taylor expansion in terms of v/c , valid for $v/c \ll 1$. In this case, the two neutron stars may be considered as Newtonian point particles in circular orbits with an orbital energy loss rate given by the Newtonian quadrupole formula as per [36]:

$$h = \frac{2G}{c^4} \frac{\ddot{Q}}{r}, \quad (3.5)$$

where r is the instantaneous separation distance and Q is the mass quadrupole moment tensor. To leading order, and using natural units of $G = c = 1$ for clarity, the strain amplitude h and phase ϕ of the gravitational wave are then derivable as

⁸The strain amplitude is the relative contraction or expansion of space, defined as $h(t) \equiv \frac{\Delta L}{L}$.

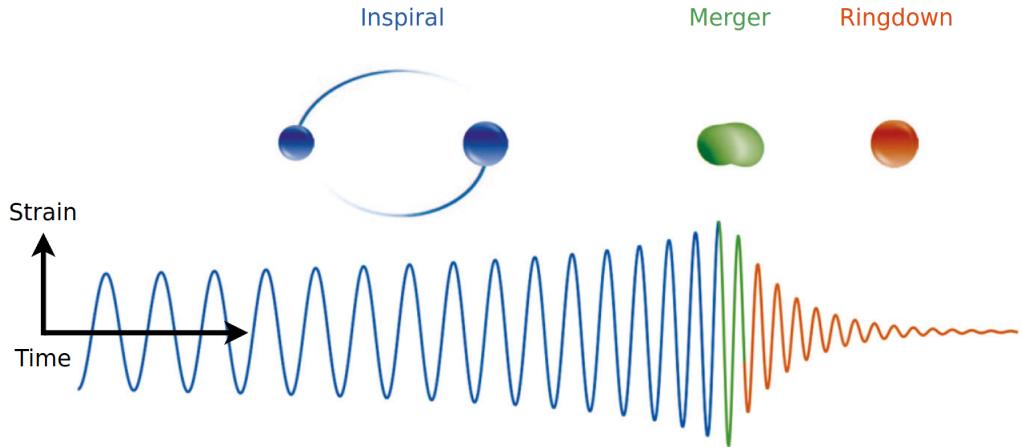


Figure 3-3: The three dynamic phases of a binary merger.
A hypothetical temporal strain curve induced at a particular detector is shown below the illustrations of each of the three phases. Adapted from [12].

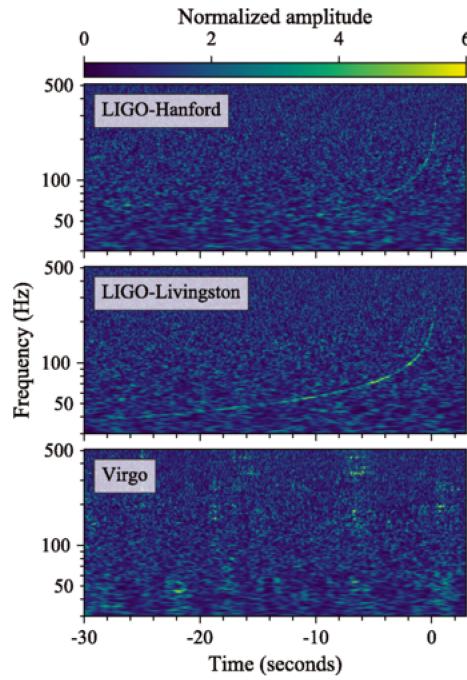


Figure 3-4: The gravitational wave chirp of GW170817 as viewed by different detectors. Credit: [51].

in [125]:

$$h(t) \propto \frac{1}{d_L} \mathcal{M}^{5/3} \dot{\phi}^{2/3} \cos \phi \cos \iota \quad (3.6)$$

$$\phi(t) \sim 2 \left(\frac{t}{\mathcal{M}} \right)^{5/8} \quad (3.7)$$

where $\mathcal{M} = (m_1 + m_2)\eta^{3/5}$ is the chirp mass, $\eta = \frac{m_1 m_2}{(m_1 + m_2)^2}$ is the symmetric mass ratio, and d_L is the source luminosity distance. Notably, $h \propto \frac{1}{d_L}$, compared to electromagnetic waves for which amplitude $\propto \frac{1}{d_L^2}$: the sensitivity of gravitational wave interferometers scales more favorably with distance than traditional telescopes.

As discussed in Sec. 3.1.1, the primary goal of the GW portion of the multi-messenger observatory network (with respect to cosmological measurements) is to localize GW sources so as to enable EM follow-up. As with traditional EM instruments, the resolution θ of a GW network can be crudely estimated with the Rayleigh criterion⁹:

$$\sin \theta \sim 1.22 \frac{\lambda}{D}, \quad (3.8)$$

where λ is the GW wavelength and D is the maximum separation between detectors. Consider the sky localization area — the region of sky that must be covered by follow-up EM observations — to be roughly θ^2 . Then:

$$\theta^2 \sim \arcsin^2 \left(1.22 \frac{\lambda}{D} \right) \quad (3.9)$$

$$\theta^2 \sim \frac{\lambda^2}{D^2} \quad (3.10)$$

For a principally ground-based network, $D \sim 10^4$ km, the diameter of Earth. For BNS, the maximum frequency is achieved during coalescence, roughly 10^2 to 10^3 Hz [51], or a wavelength of $10^{2.5}$ - $10^{3.5}$ km. This bounds the sky area θ^2 to within 10^{-3} - 10^{-1} sr, or $10^{0.5}$ - $10^{2.5}$ deg 2 . Comparing this figure to some well-known telescopes begins to illustrate the challenges of KN detection: Hubble's WFC3 has a FOV of $10^{-2.5}$ deg 2 [41], requiring 10^3 - 10^6 individual exposures to map the full sky localization

⁹The Rayleigh criterion is not a lower bound since super-resolution is possible through further signal modeling as well as observations collected at multiple points across time.

region, absent of galaxy-targeting strategies that effectively reduce the relevant sky localization region. Even a wide-field instrument like the Vera C. Rubin observatory with a field of view of 10 deg^2 [136] may require $10^{1.5}$ individual exposures.

In principle, all portions of the GW spectra of BNS mergers are detectable with the right suite of instruments, including the inspiral in the years leading up to the coalescence. These types of extended observations require the use of one or more space-based GW interferometers which can address a much lower-frequency portion of the GW spectra compared to ground-based interferometers, as appropriate for the early inspiral phase. While space-based interferometers can achieve higher separation distance than ground-based interferometers, the lower frequency and therefore longer wavelength means the space-based interferometers may not necessarily improve localization ability.

Electromagnetic emissions of binary neutron star mergers

Kilonovae (KN) are the transient events associated with the mass ejecta from BNS and BH-NS mergers, so named since they are $\mathcal{O}(1000)$ times brighter than traditional novae, though still orders of magnitude fainter than supernovae [105]. The time-varying anisotropic electromagnetic emissions of KN include contributions from a breadth of disparate emission mechanisms ranging from the radio to gamma portions of the spectrum. Prior to the observation of AT 2017gfo, the anticipated EM emissions were only predicted through radiation transfer and radioactive heating models such as [105], though AT 2017gfo proved to be consistent with the majority of model predictions [151]. Perhaps the most surprising discovery was that of GRB170817A, a short gamma-ray burst (sGRB) occurring 1.7 s after the GW170817 merger [49] and coincident with AT 2017gfo; the progenitors of sGRBs were uncertain prior to the discovery but are now thought to be primarily due to BNS mergers.

While the science of kilonovae emissions is perhaps the most complex portion of BNS observability, this section will do it brief justice; for further reading refer to [110] and [73]. The material outflow from BNS mergers, which comprises the bulk source of post-merger luminosity, can be grouped into two categories: sub-

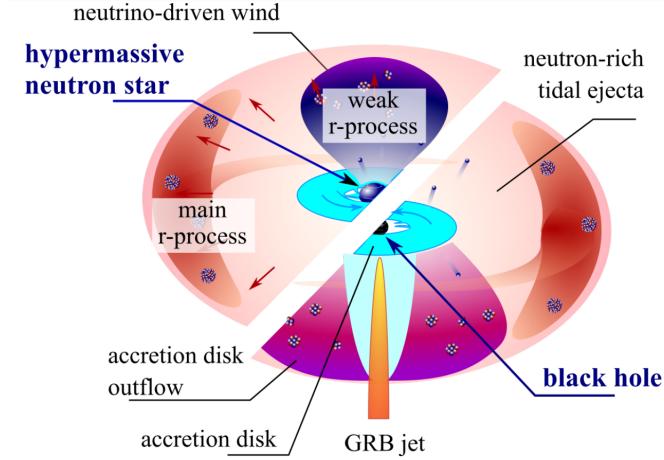


Figure 3-5: KN ejecta. Credit: [64].

relativistic outflow and relativistic outflow. Sub-relativistic components include the accretion disk surrounding the post-merger singular NS/BH remnant, shocks and winds of ejecta due to the collision, and energy imparted to the ejecta via the tidal forces in the final stages of inspiral and initial stages of the ringdown. Relativistic components include the “fast tail” of dynamical ejecta, which is the small portion of material outflow that reaches $>0.4c$; jets and/or structured jets, which are collimated, polar-oriented fast ejecta of material; and the interaction of the jet with the circum-merger medium, potentially including a jet “cocoon” and cocoon breakout depending on the jet-ejecta interactions. These components are summarized in Fig. 3-5, which depicts the scenario where the BNS merges into a hypermassive neutron star remnant (HMNS) in the top left, and the scenario where it merges into a black hole in the bottom right.

KN EM emissions can vary widely depending on challenging-to-predict dynamical outcomes of the merger, such as whether the BNS merges immediately into a HMNS, a black hole, or an unstable HMNS followed by collapse into a black hole; whether the jet cocoon successfully “breaks out” of the circum-merger medium or is “choked” and does not escape; the degree of weak and main r-process radioactive material ejected from the stars; the electron fraction and opacity of ejecta; etc. [110]. Some modeling approaches attempt a full dynamics and radiation transfer simulation [105] [130], which may be precise but also computationally burdensome; other

models such as [115] and [138] take a more phenomenological approach to generating KN light curves, using analytical approximations and fits to known data such as GW170817 and previously-detected sGRBs.

While all portions of the KN light curve may be detectable with the right set of ground- or space-based instruments, for simplicity we may discount portions of the spectra that are less relevant to the sky localization problem. Referencing Fig. 3-7, which shows the timeline of GW and EM observations of GW170817 and its EM counterpart AT 2017gfo, will be illustrative. The γ -ray emission was the first to be detected at 1.7 s post-merger. However, γ -ray instruments are generally not targeted, and so the sGRB observation did not further constrain the sky localization region beyond the initial 31 deg^2 GW-derived region [50]. Therefore γ -ray instruments are not immediately applicable for localization. The X-ray emissions are promising, as AT 2017gfo continues to be observable in X-rays even several years after the merger [82]. However, X-ray telescopes typically have limited FOV¹⁰ and are not suitable for initial localization; though the X-ray signal lasts for years, a strategy that relies on years of space-based X-ray observatory time searching for a single target is likely an inefficient one. Such a strategy would be more appropriate for optimizing the likelihood of detecting any single event rather than maximizing the total quantity of detected events. Next, while the radio portion of the light curve may last similarly as long as the X-ray portion [19], KN are generally quite dim in the radio bands: AT 2017gfo was only detected via radio waves 16 days post-merger despite that the host galaxy was already known by that point [103].

These arguments lead to the remaining candidates for initial sky localization region searches, which are the UV, optical, and IR portions of the EM spectra. Indeed, of the 6 independent discoveries of AT 2017gfo, all were ground-based telescopes that conducted their searches in either the optical or near-IR (NIR), including DECam [61], DLT40 [152], LCO [13], 1M2H [35], VISTA [147], and MASTER [69]. The light curve within this portion of the spectra is depicted for AT 2017gfo in Fig. 3-6, which shows how a typical KN will rapidly redden and dim in the days following the merger due to

¹⁰The Chandra X-ray Observatory, for example, has a $16' \times 16'$ FOV [163].

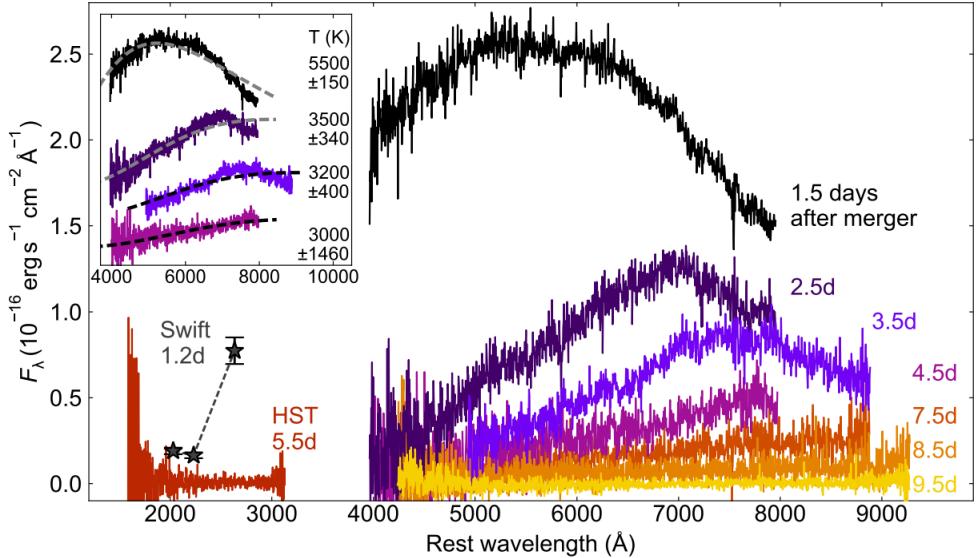


Figure 3-6: Rapid reddening and dimming of AT 2017gfo in the days following the merger. Credit: [59].

cooling of the ejecta material, rendering initial detection more difficult as time progresses. Aside from possible redshifting and other light curve modulations, most KN will reach their peak magnitude in the r band around 1 day post-merger [138] [152]. This renders detection beyond \sim 2 weeks prohibitively challenging and emphasizes the need for rapid follow-up sky searches.

3.2 Simulation methods

The goal of this section is to describe the process of calculating the performance of different possible networks of multi-messenger observatories. Based on the cosmology goals described in Sec. 3.1.1, the relevant science figure of merit for a given network is the quantity (or more straightforwardly, the rate) of BNS mergers it can detect in both the GW and EM spectra. To estimate this quantity, a simulation is set up following the block diagram shown in Fig. 3-8. First, a representative population of hypothetical BNS mergers is created by drawing parameters for each merger independently from realistic parameter distributions (Sec. 3.2.1). Then the GW and EM waveforms of each merger are simulated, taking into account the physical processes of CBC and

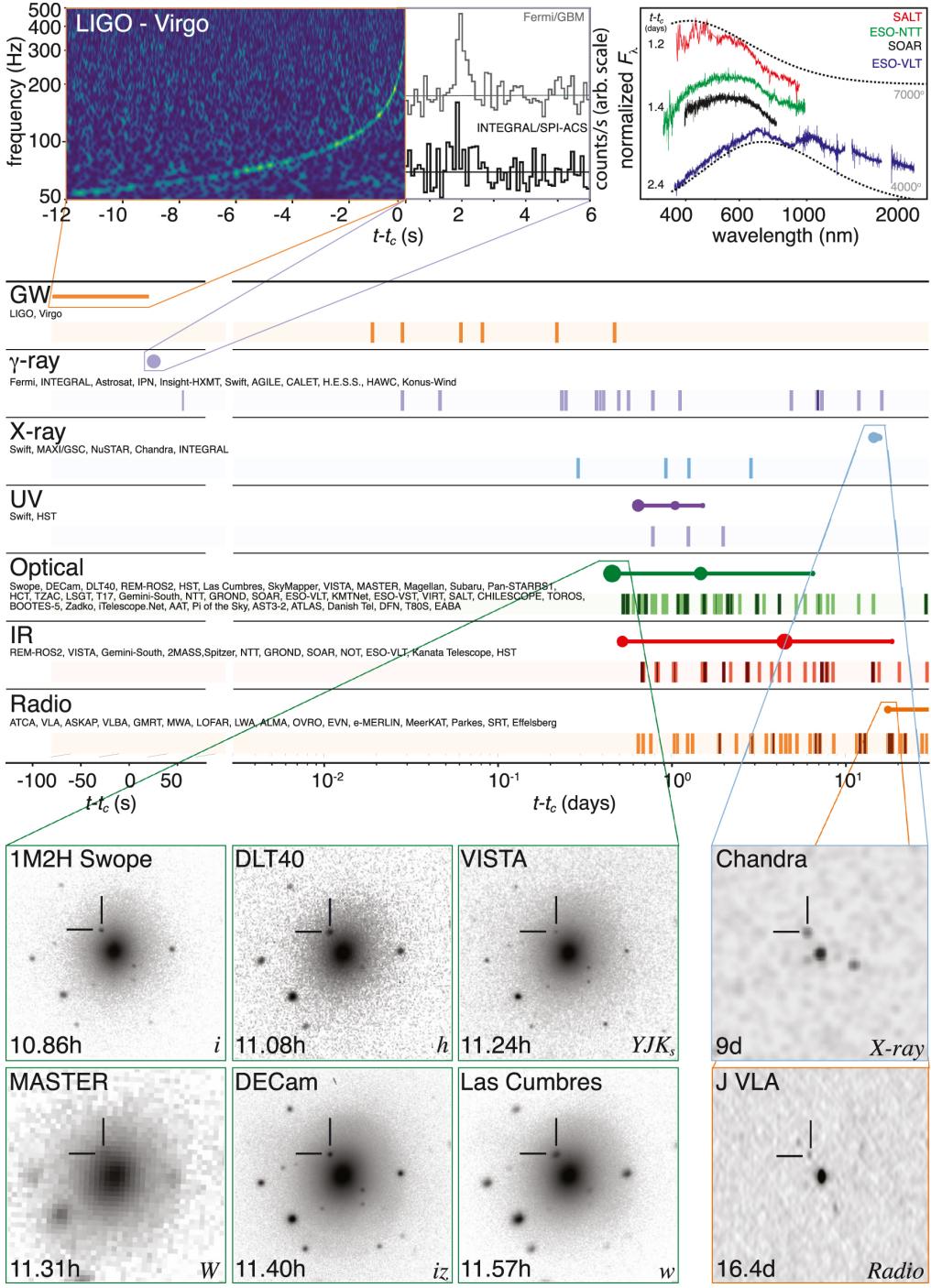


Figure 3-7: Timeline of multi-messenger observations of GW170817/AT 2017gfo. Observations range from ~ 1 min. pre-merger to ~ 1 mo. post-merger and from radio waves to γ rays. Credit: [50].

KN emission (Sec. 3.2.2 and 3.2.4 respectively). Based on the GW interferometer portion of the network, parameter inference determines whether the GW signals are observable (whether they exceed an SNR threshold) and outputs the GW network’s sky localization region for each observed signal (Sec. 3.2.3). The EM telescope portion of the network then attempts to search each sky localization region using a tiling strategy; if the network has sufficient optical depth (limiting magnitude) to observe the target, the detection probability is calculated based on the portion of the sky localization region observed before the KN dims (Sec. 3.2.5).¹¹ The performance of the network is then calculated as the expected fraction of detected events out to a certain distance (Sec. 3.2.6).

Prominently, the simulation calculates the performance of disparate networks of GW and EM instruments in parallel. As a result, more than $\mathcal{O}(10^5)$ possible networks can be analyzed overnight on a typical laptop. This contrasts to otherwise similar studies such as [116], which computes the performance of different GW/EM observatory networks given a simulated BNS/BH-NS merger population, but which considers only 24 different networks. Parallelization is enabled by a cornerstone of the simulation methodology: complexity in performance modeling is shifted away from the network as a whole and instead towards individual instruments. Specifically, modeling of instrument performance is allowed to be as detailed and computationally expensive as necessary to achieve a realistic result. The network performance, however, is defined such that it is straightforward and computationally efficient to determine from the performance of its set of individual detectors, as described by Sec. 3.2.6. Essentially, the conditional probabilities of the Bayesian network shown in Fig. 2-8 (as per Eqn. 2.16) are analytically computable from a more involved simulation of instrument/merger observations. Once each instrument’s observation of each signal is simulated, the result can be reused for arbitrary network configurations, and calculations for a large number of networks is as easy as calculations for a single network.

¹¹This strategy assumes that the sky area tiling is optimally timed: specifically, that the optical search begins the exact moment the merger is sufficiently bright so as to be observed. The search proceeds until either the merger is identified or more than 15 days have passed.

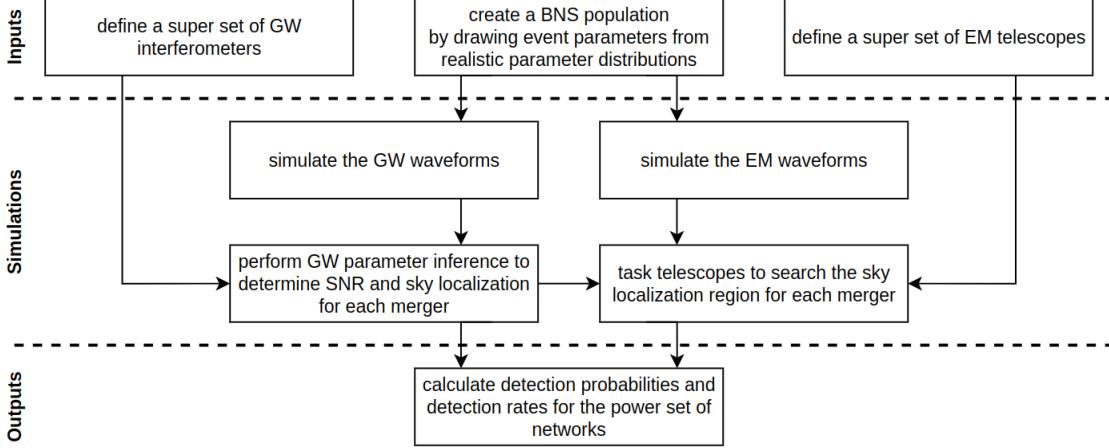


Figure 3-8: A block diagram of the simulation approach.

By this process, calculations of detector performance can be easily extended to the power set of observatories: the set of all possible subsets, i.e., all $2^N - 1$ network configurations it is possible to construct from a specific set of N instruments, excluding the empty set. This parallelization enables an $\mathcal{O}(N)$ computation of performance of the network power set. This compares favorably to an otherwise $\mathcal{O}(2^N)$ algorithm given by computing each configuration in a serial fashion, and allows considering both a large N set of instruments as well as a vast quantity of possible networks. In this way, simulations of individual instruments or network configurations do not have to have significant manual input to motivate promising candidates for future networks, but can instead rely on brute force permutation.

3.2.1 Assumed binary neutron star merger population

Each simulation begins by generating a random population of merger events out to a certain comoving distance and over a specific geocentric time range. To model a population, the individual parameters of each event are randomly sampled from the distributions defined in Table 3.4, which are assumed to be independent. Each merger is describable by a set of 15 intrinsic and extrinsic parameters as per Section 3.1.3. The intrinsic parameters follow from the physics of binary neutron star formation, and therefore require physical models and/or population studies to determine rea-

Table 3.4: Assumed binary neutron star parameter distributions

Parameter	distribution	assumptions
m_1	$\mathcal{N}(1.5, 0.15^2) M_{\odot}$	PEAK Gaussian model [96]
m_2	$\mathcal{N}(1.5, 0.15^2) M_{\odot}$	PEAK Gaussian model [96]
\vec{s}_1	$\mathbf{0} \text{ kg m}^2 \text{ s}^{-1}$	spin may be ignored for localization inference [116]
\vec{s}_2	$\mathbf{0} \text{ kg m}^2 \text{ s}^{-1}$	spin may be ignored for localization inference [116]
ϕ_c	$\mathcal{U}[0, 2\pi) \text{ rad}$	rotational symmetry of merger dynamics
ψ	$\mathcal{U}[0, 2\pi) \text{ rad}$	isotropic merger orientation
$\cos(\iota)^a$	$\mathcal{U}[-1, 1]$	isotropic merger orientation
d_C^3	$\mathcal{U}[0, d_{C,\max}^3] b \text{ Gpc}^3$	homogeneous rate density for $d_C \leq d_{C,\max}$
α	$\mathcal{U}[0, 2\pi) \text{ rad}$	isotropic spatial distribution
$\cos(\delta)^a$	$\mathcal{U}[-1, 1]$	isotropic spatial distribution
t_c	$\mathcal{U}[t_0, t_f]^c \text{ s}$	negligible cosmological time span is simulated
all		all parameters independently distributed

^a Given a point uniformly distributed on the surface of a sphere, $p \sim \mathcal{U}(\mathbb{S}^2)$, the cosine of latitude-related quantity, e.g. the declination or inclination, has a uniform distribution, whereas the longitude-related quantity is itself uniformly distributed.

^b $d_{C,\max}$ is the maximum comoving distance at which sources are randomly generated, chosen to be either slightly beyond the horizon distance or set at a maximum of 1.3 Gpc to minimize the impact of modeling assumptions regarding redshifting. Since mergers are assumed to occur uniformly in comoving volume, the cube of the comoving distance is uniform.

^c t_0 and t_f represent the GPS time of the beginning and end of the simulated time range, e.g. $\mathcal{U}[2208643218, 2240179218] \text{ s}$ for the 2050 calendar year.

sonable distributions. This work relies on studies performed by [96] and [116] for intrinsic parameter distributions. The extrinsic parameters follow from the distribution of events across space and time; for the purpose of this thesis, the assumptions of homogeneity and isotropy generally lead to uniform extrinsic parameter distributions. In reality, the local universe is not necessarily homogeneous nor isotropic due to random variation in galaxy clusters and matter density, but this simplifying assumption used to generate the extrinsic parameters should not significantly alter the calculated network performance.

During the population modeling, it is necessary to incorporate a specific cosmological model which converts comoving distance (which defines the merger distribution over volume) to both a redshift and a luminosity distance for each event. This step is required for both the gravitational wave and electromagnetic observations, as the frequency and amplitude of each wave as seen by an Earth-based observer are af-

fected by the redshift and luminosity distance respectively. Consistent with the rest of this thesis, a flat Λ CDM model `ASTROPY.cosmology.WMAP9` is used [85], along with the corresponding functions `z_at_value()` to root-find redshift z based on comoving distance d_C , and `luminosity_distance()` to then convert to d_L .

Beyond the event-level parameter distributions, population-level parameter distributions are modeled, including the rate density \mathcal{R}_{BNS} (a quantity of events per unit comoving volume per unit proper time) and a maximum event distance. The rate density is assumed to be uniform with comoving distance, which is only approximately true in the local universe; in reality the rate density varies with the age of the universe.¹² The difference between the assumed uniform distribution and the predicted distribution is depicted in Fig. 3-9., where the merger rate can be seen to be fairly uniform below $d_C \sim 1$ Gpc. [65] finds a value of $\mathcal{R}_{BNS} = 320^{+490}_{-240} \text{ Gpc}^{-3} \text{ yr}^{-1}$ based on the first and second LIGO–Virgo Gravitational-Wave Transient Catalog (GWTC-1 and GWTC-2), which this work adopts. For the sake of sufficient statistical sampling, the rate density is incorporated via statistical bootstrapping only after the complete simulation has been run; the simulation itself uses a (higher) fixed number of events, $N \approx 10^{3.5}$, chosen to balance sample size with overall compute. Conveniently, $10^{3.5}$ events permitted 6-12 hour runs which could be completed overnight.

The maximum distance is an aphysical variable chosen for statistical reasons. Since the number of events scales with distance cubed, it would be ideal to model events only out to the slightly beyond the horizon distance of a given network, else a majority of the compute would be spent on unobservable events. However, this method poses challenges when comparing networks with incomparable horizon distances. Using the same simulated population to calculate the performance of each network (enabling a fair comparison), the weaker network would only observe a small fraction of events in the population, whereas the stronger network would observe most of them. Therefore the statistics of nearby events observable by the weaker network would be poorly sampled, so the performance calculation would be less accurate and

¹²Consider that in the early universe (corresponding to high redshifts), when no neutron stars had yet to form in the first place, no mergers would be expected.

also inconsistent between runs.

An alternative method is to simulate populations out to different maximum distances for each network, but this poses several computational challenges. Most importantly, a consistent population across networks is taken advantage of in Sec. 3.2.6 to calculate the performance of many networks in parallel—switching to a serial approach would greatly constrain the number of networks that could be simulated. Secondarily, an initial simulation is required to determine the horizon distances, further hindering this approach.

Instead, the maximum distance is set uniformly as less than the horizon distance of the best networks. Then the performance can be calculated as the fraction of events observable within the maximum distance, scaled to 0 – 1. The downside of this approach is that comparisons between high-performance networks will be somewhat weakened due to performance saturation; e.g. if two networks can both observe 100% of events, the question of which is the better network will not be clarified. However, this approach permits comparison of a broad range of networks ranging from the weakest to the strongest, and as long as the performance does not fully saturate, comparisons will still be permitted between the high-performance networks.

For most of the simulations detailed in Sec. 3.3, the maximum distance is chosen as $d_C = 1.3$ Gpc, or $z \sim 0.3$. Since this distance is at a relatively low redshift it ensures a variety of simplifying assumptions hold true, such as the uniformity of BNS merger rate density. More ambitious cosmology goals requiring measurements out to higher redshifts will not be reflected in the network performance analysis, but the baseline measurement of H_0 will be well-addressed.

3.2.2 Gravitational wave signal modeling for matched filtering

A signal model is required in this work both in order to simulate the signal itself, as well as to perform parameter inference on the signal (discussed in Sec. 3.2.3) through matched filtering, the maximum likelihood technique for GW data analysis. Matched filtering is the process by which $\mathcal{O}(10^{4+})$ templates (possible signals that could be generated by a real signal) are checked against the measured strain so as to reveal

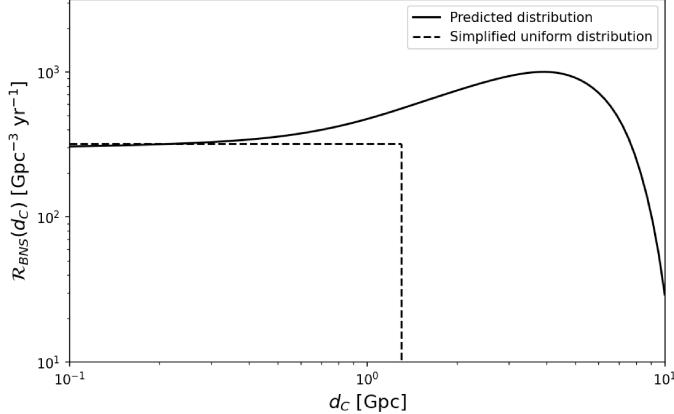


Figure 3-9: Rate density of BNS mergers with comoving distance. Data for the predicted distribution is from [18], which shows a spike in BNS mergers around cosmic noon ($z \sim 2$) corresponding to the peak star formation rate. The cutoff of the uniform distribution corresponds to $d_{C,max} = 1.3$ Gpc, the maximum simulated merger distance.

weak signals otherwise hidden by noise. The general goal is come up with a function μ that outputs the gravitational waveform (or alternatively, its Fourier transform) given a set of signal parameters θ :

$$h(t) = \mu_t(\theta), \quad (3.11)$$

OR $h(f) = \mathcal{F}\{h(t)\}(f) = \mu_f(\theta).$

Modeling the gravitational wave signal of a compact binary coalescence requires a waveform approximant since no exact analytical solutions exist. In general, one has a choice between using (1) a numerical relativity (NR) simulation, which solves the Einstein field equations for the metric tensor of the two orbiting masses and propagates their states forward in time [83]; (2) a post-Newtonian (PN) approximation which takes the form of a Taylor series expansion of orders $(\frac{v}{c})^2$ [24] [91], like the 1st-order case discussed in Sec. 3.1.3; or (3) a phenomenological waveform approximant that typically joins PN-derived inspiral and NR-derived merger/ringdown results to create analytical waveforms [39]. For this work, NR simulations are immediately not applicable; computational requirements dictate that they typically are not used for matched filtering [40]. The remaining choices is between a PN or phenomenological

method. A further choice is whether the signal is modeled in the frequency or time domain; here, a frequency-domain model is required for the Fisher information matrix method utilized in Sec. 3.2.3.

From these remaining options, two widely-used candidates are `TaylorF2` [27] for the PN waveform and `IMRPhenomXPHM` [57] for the phenomenological waveform. `TaylorF2` is a promising choice because it presents the simplest frequency domain approximant yet is a robust method used for e.g. initial inspiral searches. It takes advantage of the stationary phase approximation to analytically determine a 3.5-order frequency domain waveform exclusively of the inspiral portion of the coalescence. `IMRPhenomXPHM` is considered as it is a computationally efficient phenomenological method that models the entirety of the inspiral, merger, and ringdown. These latter portions of the signal may be important for localization due to the higher-frequency signal generated in the merger and ringdown phases (corresponding to a lower resolution as per the Rayleigh criterion, described in Sec. 3.1.3). The phenomenological waveform is derived by fitting the merger parameter space to hundreds of numerical relativity simulations, avoiding overfitting by using a hierarchical approach that models the contribution of each parameter in order of relative importance [93]. From initial tests, the computational runtime of `TaylorF2` versus `IMRPhenomXPHM` is negligible (a factor of ~ 2 in favor of `TaylorF2`), so `IMRPhenomXPHM` is used due to the completeness of the waveform. [42] suggests that `TaylorF2` is more than sufficient for BNS localization, but also that `IMRPhenomXPHM` is generally more accurate. Both waveforms are implemented in the `LALSuite` [99] and `GWFish` [42] software packages used in this work and therefore are relatively interchangeable.

3.2.3 Gravitational wave parameter inference

Parameter inference is required to extract information about each of the simulated signals—in particular, the SNR and localization—based on the measured strain. Standard methods rely on Bayesian inference with statistical sampling, specifically via Markov Chain Monte Carlo (MCMC) or nested sampling [148]. Broadly, these methods represent solving an inverse problem to construct a posterior distribution $p(\theta|d)$,

which is a probability density function for the set of 15 merger parameters θ given strain data d measured by the GW network:

$$p(\theta|d) = \frac{\mathcal{L}(d|\theta)\pi(\theta)}{\mathcal{Z}}. \quad (3.12)$$

The likelihood function \mathcal{L} is a measurement description that assumes a noise model. GW inference typically assumes Gaussian noise of the form

$$\mathcal{L}(d|\theta) \propto \exp\left(-\frac{1}{2} \frac{|d - \mu(\theta)|^2}{\sigma^2}\right) \quad (3.13)$$

where σ is the detector noise and $\mu(\theta)$ is the gravitational waveform template given by the signal model described in Sec. 3.2.2. Gaussian noise represents a reasonable approximation since noise contributions are generally random, numerous, and independently distributed, therefore satisfying the central limit theorem. In some works the noise model is also expanded to include non-Gaussian contributions like instrumental lines.

Note that in this application, the data d is represented as the FFT of the measured strain $d(t)$ since the waveform model $\mu_f(\theta)$ outputs in the frequency domain. Waveform approximants can alternatively output in the time domain, but detector noise is most easily modeled as a power spectral density (PSD) $S_n(f)$ which is defined as a function in the frequency domain—essentially, the frequency-dependent noise contributions. The strain data is composed of a true signal $h(f) = \mu_f(\theta)$ and noise source corresponding to $S_n(f)$, or in the time domain, the true signal $h(t)$ and noise $n(t)$:

$$d(t) = h(t) + n(t). \quad (3.14)$$

See [109] for further details regarding how the matched filtering process to Bayesian inference of GW signals; the overview given here primarily addresses the inference portion of the analysis.

Using Bayes' rule (Eqn. 3.12) the prior distributions for merger parameters θ are obtained from hypothesizing plausible distributions. This procedure is the same work

performed for the physically-motivated population distributions given in Sec. 3.2.1. The evidence \mathcal{Z} used to normalize the total probability is calculated as per

$$\mathcal{Z} \equiv \int \mathcal{L}(d|\theta) \pi(\theta) d\theta. \quad (3.15)$$

The posterior distribution is approximated through iterative sampling; with an MCMC algorithm, for example, the distribution is constructed as a histogram of points sampled from the Markov chain's stationary distribution. The stationary distribution is the converged distribution after sufficiently many iterative updates of the Markov chain. MCMC's strength is that it can model arbitrary distributions, including those that would otherwise be challenging to calculate directly [157].

An alternative approach for parameter inference is the Fisher information matrix (FIM) method, which analytically approximates the posterior $p(\theta|d)$ as a multivariate normal distribution. The Fisher information matrix for a specific detector l or network of N detectors is computed by:

$$\begin{aligned} \mathcal{I}_{ij,\text{detector } l}(\theta) &= \sum_{k=1}^N \langle \partial_{\theta_i} \mu^k(\theta) | \partial_{\theta_j} \mu^k(\theta) \rangle \\ \mathcal{I}_{ij,\text{network}}(\theta) &= \sum_{l=1}^N \mathcal{I}_{ij,\text{detector } l}(\theta), \end{aligned} \quad (3.16)$$

where μ^k is the 1st derivative of the waveform with respect to the k -th model parameter θ_k . The inner product is defined on a signal-model manifold as an integral involving the signal and the detector PSD $S_n(f)$; for further detail see [42]. Note that Eqn. 3.16 is mathematically equivalent to assuming a quadratic approximation to the likelihood \mathcal{L} :

$$\mathcal{L} \propto \exp \left(-\frac{1}{2} \Delta \theta_i \mathcal{I}_{ij}(\theta) \Delta \theta_j \right). \quad (3.17)$$

Importantly, the inverse of the Fisher matrix can be interpreted as the covariance of $p(\theta|d)$, i.e. the uncertainty in measurement of parameters θ :

$$\mathcal{C}(\theta) = \mathcal{I}^{-1}(\theta), \quad (3.18)$$

by assuming (1) stationary Gaussian noise, (2) uniform priors on the model parameters, and (3) that the waveform can be considered as a linear function of θ , which applies only in the high-SNR limit. The SNR measured by detector l or network of N detectors is defined as:

$$\rho_l \equiv \sqrt{\langle h|h \rangle}$$

$$\rho_{\text{network}} = \sqrt{\sum_{l=1}^N \rho_l^2} \quad (3.19)$$

where h is the frequency-domain GW strain, and the inner product is defined equivalently as the inner product seen in Eqn. 3.16 [92].

Due to the nature of inference as an inverse problem as well as the relative weakness of GW signals, degeneracies and nonlinearities in the signal parameter space are common. The Fisher-matrix formalism cannot capture these effects as it is fundamentally a linear method, and so is arguably somewhat inappropriately-suited for this application [156]. Additionally, ill-conditioned or singular Fisher matrices can arise in degenerative cases, which prohibit accurate calculation of the covariance $\mathcal{C}(\theta)$. Unfortunately, the geometry of GW detection dictates that the majority of signals will always be measured close to the SNR threshold due to d^3 distance distribution—i.e., most observed signals will be furthest from us and close to the edge of the observable volume of space—and therefore $\mathcal{I}^{-1}(\theta)$ does not always represent an accurate posterior. In this simulation, a higher-than-typical coherent SNR threshold ($\rho \geq 12$) was used to minimize this effect.

However, the major upside of the FIM method is that it is significantly more computationally economical than full Bayesian treatments, in part because the posterior is analytically computable rather than being based on iterative methods like MCMC or nested sampling. Moreover, Eqn. 3.16 represents a subtle but powerful approach to modeling a large suite of detector networks, as it enables parallel rather than serial processing of each network configuration. By calculating $\mathcal{I}_{\text{detector}}(\theta)$ for each considered detector, one can determine $\mathcal{C}_{\text{network}}(\theta)$ for *every possible* combination of detectors merely through an additional summation and inversion of \mathcal{I} for each unique combination. This represents an $\mathcal{O}(2^N)$ speedup for a set of N considered detectors

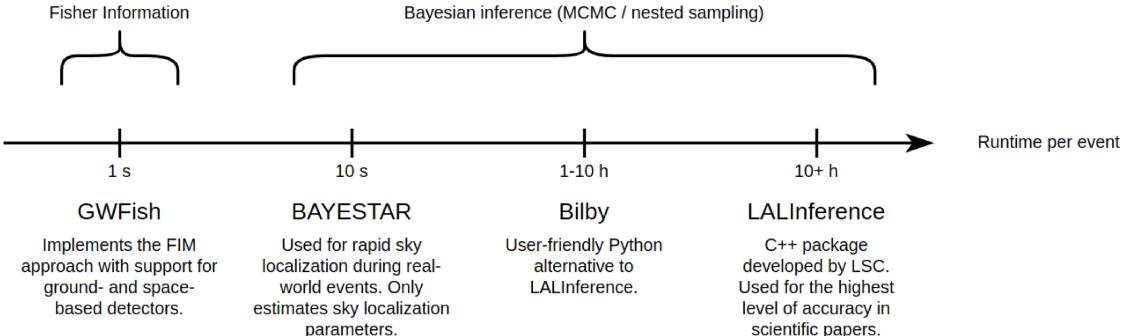


Figure 3-10: GW parameter estimation software packages.
Ordered by approximate runtime per event.

compared to serial Bayesian analysis—a simulation-enabling capability.

GW software packages

Bayesian inference algorithms and the FIM approximation are implemented in numerous software packages for GW data analysis. A few of the most popular and relevant options include `GWFish` [42], `BAYESTAR` [144], `Bilby` [17], and `LALInference` [99], described briefly in Fig. 3-10. Qualitative capabilities of each software package are given in Table 3.5, which demonstrates the tradeoff between accuracy and computational runtime. Runtime is the primary consideration in simulation design because this work simulates a population of $\mathcal{O}(10^{3.5})$ binaries, which is then analyzed for $\mathcal{O}(10^3)$ possible GW networks. The most accurate and widely-used Bayesian methods implemented by `Bilby` and `LALInference` are therefore disregarded due to their computational expense, requiring hours or more to perform inference for measurements of a single compact binary coalescence.

The remaining options are `GWFish` and `BAYESTAR`. `GWFish` is a FIM-based Python package intended for general population-level analysis of the science enabled by future GW detector networks. The software models the effects of detector motion on the posterior estimate, enabling analysis for space-based detectors. Conveniently, it also provides out-of-the-box support for defining custom detectors and implements all `LALSuite` waveform approximants including `IMRPhenomXPHM`, the approximant used in this work. In contrast, `BAYESTAR` is a Bayesian inference algorithm based on low-

Table 3.5: Comparison of GW parameter estimation software packages

	<code>GWFish</code> [42]	<code>BAYESTAR</code> [144]	<code>Bilby</code> [17]	<code>LALInference</code> [99]
Accuracy				
Runtime				
Readily provides out-of-the-box support for different types of detectors:				
Ground-based				
Space-based				
Custom				

order Gaussian quadrature, and is primarily used for real-world EM follow-up of GW events. It achieves accuracy close to that of full MCMC analyses and enables quasi-real-time inference by exclusively addressing the extrinsic parameters, ignoring the full parameter space of the binary system.

Ultimately `GWFish` is used in this work because the main tradeoff between `BAYESTAR` and a FIM-based approach such as `GWFish` comes down to computational cost. Beyond the roughly order-of-magnitude difference in raw runtime per event, the FIM method provides an enabling capability to simulate many GW networks in parallel, whereas otherwise the networks have to be processed in a serial fashion. As long as the limitations of the FIM method (inaccuracies in the event of degeneracies and when modeling low-detector-count networks) are kept close in mind, it represents a powerful and simple approach to GW analysis [156].

Simulation of the sky localization region

The way the network functionally localizes a source is somewhat obscured by the mathematical formalism provided above. Intuitively speaking, the means by which localization is enabled is through two physical mechanisms. Most importantly, depending on the source origin, there is a difference in signal arrival time (via a phase measurement) at each of the detectors corresponding to $\Delta t = \frac{1}{c} \vec{r}_{\text{detector}} \cdot \frac{\vec{r}_{\text{source}}}{|\vec{r}_{\text{source}}|}$, where \vec{r} is the position with respect to e.g. a geocentric origin. A minimum of four detectors are required for precise localization using this method exclusively; three detectors pinpoint two possible sky regions, two detectors determine a ring of possible locations,

and one detector provides no localization knowledge at all. This method relies on arrival time (phase) measurements which correlate with extrinsic parameters t_c , ϕ_c , and ψ . Secondarily, the differing orientations of interferometer antenna patterns influences the relative amplitude observed by each detector, further constraining the localization region. For example, Virgo’s *lack* of detection of GW170817 actually considerably constrained the localization region because the event must have resided in a location of the sky where Virgo was least sensitive [51]. This method relies on amplitude measurements which correlate with extrinsic parameters d_L and ι and less strongly with intrinsic parameters m_1 and m_2 .

Critically, estimates of sky localization parameters α and δ correlate strongly with other extrinsic parameters of the binary system and to a much lesser degree with the intrinsic parameters. BAYESTAR takes advantage of this distinction by assuming fixed intrinsic parameters when estimate the sky localization region, which provides a computational speedup due to the reduced dimensionality of the posterior. Initial testing within GWFish showed restricting the parameter subset merely to α and δ resulted in order-of-magnitude underestimates of σ_{sky} , the standard deviation of the “sky error”, and less significant underestimates when restricting merely to the set of extrinsic parameters. Personal communication with the GWFish team including Jan Harms and Jacopo Tissino suggested to not restrict the Fisher analysis. Instead, the full dimensionality of merger parameters θ is used, merely excluding the two NS spin vectors (which are assumed to be negligible). This results in a 9-dimensional covariance matrix compared to the complete 15-dimensional parameter space θ . This precaution ensures no spurious correlations between parameters are missed, all of which would increase σ_{sky} .

Presuming the network SNR exceeds the coherent SNR threshold ($\text{SNR}_{\text{network}} \geq 12$)¹³, the simulation proceeds to analytically calculate the uncertainty in measurement of the signal parameters for each BNS merger. These uncertainties are represented in the form of a covariance matrix as per the FIM method. The covariance

¹³This work utilizes a higher-than-typical coherent SNR threshold for GW detection in order to avoid degenerative cases where the FIM approach is ill-suited.

elements associated with right ascension α and declination δ give the sky localization as a bivariate normal distribution:

$$\mathcal{C}_{\alpha,\delta} = \begin{pmatrix} \sigma_{\alpha\alpha}^2 & \sigma_{\alpha\delta}^2 \\ \sigma_{\delta\alpha}^2 & \sigma_{\delta\delta}^2 \end{pmatrix}, \quad (3.20)$$

where the distribution is centered on the maximum likelihood estimate of the BNS merger sky position, the true sky position. The size of the sky localization region, the standard deviation of the sky error, is then given by:

$$\sigma_{\text{sky}} \equiv \sqrt{\text{trace}(\mathcal{C}_{\alpha,\delta})} = \sqrt{\sigma_{\alpha\alpha}^2 + \sigma_{\delta\delta}^2}. \quad (3.21)$$

This sky localization region is passed on to the EM follow-up portion of the simulation to determine what area of the sky the EM network needs to search to determine the precise BNS merger position.

3.2.4 Electromagnetic signal modeling

The complexity of the electromagnetic emissions of kilonovae, discussed in Sec. 3.1.3, renders waveform modeling a challenging exercise. Due to the quantity of mergers simulated in this work, it is important to select a modeling approach that is computationally lightweight and flexible. Studies such as [116] have defaulted to using a range of fixed absolute magnitudes M adjusting for apparent magnitude m based on luminosity distance as per:

$$m - M = 5 \log \frac{d_L}{10}. \quad (3.22)$$

Others use a preset light curve but with a simple parametrization for changes with viewing angle [107]. Now that we have better knowledge of KN emissions, we can better take into account the full range of variability in real KN. Rather than relying on a full hydrodynamic and radiation transfer simulation such as in [130], a simpler phenomenological model (that still takes into account time-variability, spectral emissivity, and anisotropy) is employed from [115]. Fig. 3-11 shows the structure of

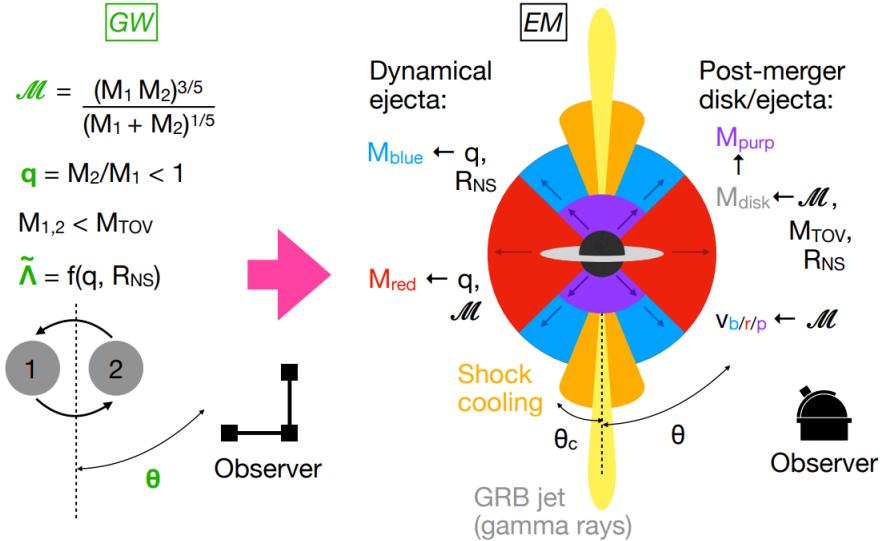


Figure 3-11: The KN light curve modeling approach taken by [115], used in this work. Credit: [115].

the modeling approach: it requires only a few input parameters, including the chirp mass \mathcal{M} , mass ratio q , observing angle θ (related to the inclination), and an assumed neutron star tidal deformability parameter $\tilde{\Lambda}$. While internal model parameters such as shock cooling and magnetic surface winds can be tuned to fit actual BNS observations, the baseline model predicts GW170817’s apparent magnitude accurately to within a factor of 2 ($\Delta m \approx 0.7$ magnitudes) according to the original study [115], so the baseline model is used. Model inputs also include the fraction of polar dynamical ejecta and fraction of the remnant accretion disk that is ejected; in this study, these fractions are fixed to 0% and 10% respectively.

The model outputs a set of absolute AB magnitudes in the KN rest frame for 17 photometric bands from UV to NIR in 0.1-day increments from 0 to 15 days post-merger, depicted for an example simulated merger in Fig. 3-12. Since the model ingests relatively few input parameters, for the sake of computational savings across runs, a database of $\mathcal{O}(10^{4.5})$ merger light curves for the full range of possible merger parameters is generated prior to any individual simulation. During simulations, light curves are determined by linearly interpolating between the nearest merger data points within the database, and the absolute magnitude is converted into an apparent magnitude based on the luminosity distance of the merger using Eqn. 3.22.

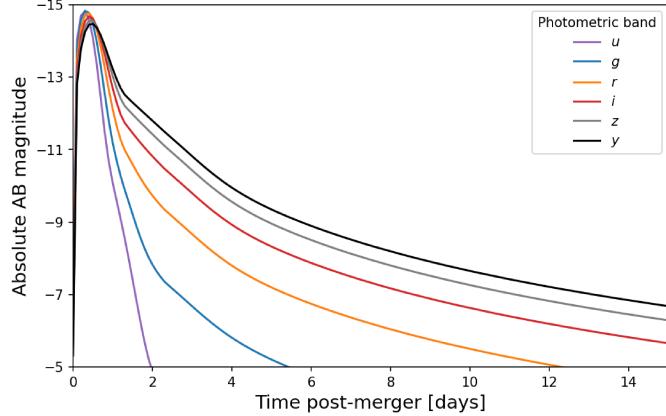


Figure 3-12: An example light curve.

This light curve is produced by the output of [115] for $\mathcal{M} = 1.7 \text{ M}_\odot$, $q = 0.9$, $\theta = 0$.

Other possible corrections include taking into account light curve redshifting (known as a K correction) and applying wavelength-dependent amplitude decreases based on atmospheric, galactic, and intergalactic extinction. For simplicity, none of these corrections are applied. Since the maximum simulation distance is relatively local at $z \approx 0.3$, redshifting and intergalactic extinction should be reasonably negligible. Atmospheric effects are taken into account primarily by assuming a certain sky background brightness as a noise source, discussed further in Sec. 3.2.5. Galactic extinction is chosen to be ignored because it applies equally to all mergers based only on galactic inclination, and so its incorporation would not necessarily provide further insight into comparing the relative performance of different multi-messenger network configurations. Plausibly, the main effect of these simplifications in the simulation output is that the performance of space-based telescopes will be underpredicted relative to ground-based telescopes, since they avoid atmospheric extinction and may employ IR imaging which is useful for higher-redshift observations. Regardless, this simulation accurately captures KN emission processes and is able to identify the key tradeoffs in EM network composition.

3.2.5 Electromagnetic follow-up of the sky localization regions

Targeting strategy

Once a GW-based sky localization region for a given merger has been determined, telescopes are tasked to search the region for potential transients to identify the KN host galaxy. Typically, telescopes either take a wide-field tiling strategy or a narrow-field galaxy-targeting strategy. Wide-field tiling approaches divide the localization region into a grid of FOVs which are systematically imaged by one or more telescopes. The grid is searched in order of decreasing integrated probability across the grid tiles—no attention is paid to other factors which may affect the detection probability such as the relative quantity of galaxies within each tile, in part because each tile contains a large number of galaxies. Tiles may be visited out-of-order based on observability constraints such as sky altitude and airmass, and may also be revisited as the kilonova evolves, in particular if the 90% confidence region (c.r.) is fully searched with no transients found. This approach tends to rely on large FOVs and instruments with high photon throughput that can quickly address large localization regions. During the search for the optical counterpart to GW170817, DECam [61], MASTER [69], and VISTA [147] all successfully employed a tiling approach. Fig. 3-13 depicts DECam’s GW170817 strategy, in which the sky tiles were determined algorithmically based on the published sky localization region and the resulting raw exposures checked by visual inspection for any new transients.

For narrow-field telescopes which may achieve greater optical sensitivities but at the expense of smaller FOVs, the tiling strategy is not optimal—constructing the same style of sky localization grid with smaller tiles would produce images with few or no galaxies in many of the FOVs, which would therefore not be relevant to determining the host galaxy of the KN. Instead, galaxy-targeting approaches rely on comparing the sky localization region to existing galaxy catalogues. They use a variety of bespoke ranking systems such as [15], [101], and [111] to determine likely galaxy candidates within the sky localization region (or localization volume, if they also consider the estimated merger distance). These galaxies are then followed up

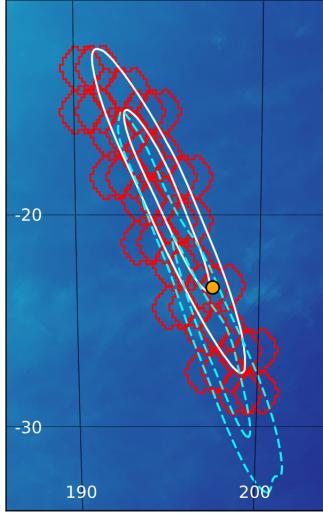


Figure 3-13: The wide-field tiling approach used by DECam to successfully find the optical counterpart of GW170817.

The red tiles show the FOVs of individual exposures; the contours show the initial (solid) and updated (dashed) 90% and 50% confidence regions published by LIGO at 4 hours post-merger and 12 hours post-merger respectively. Credit: [61].

on individually, or in some cases in parallel if multiple candidate galaxies fit in a single FOV. This search strategy was successfully used to find GW170817’s optical counterpart by 1M2H [35], DLT40 [152], and LCO [13]. In the case of 1M2H, depicted in Fig. 3-14, the ranking system calculated that the search had an estimated 95.3% chance of discovering the host galaxy despite only imaging 26.9% of the integrated raw sky localization probability—specifically, the search covered only 13.3 deg^2 of the 31 deg^2 90% c.r.—indicating the strength of this approach [35].

The simulation method used in this work assumes a wide-field tiling approach. Geometric arguments suggest that the strength of the galaxy-targeting approach relative to the tiling approach will diminish as the GW network performance improves, as it is expected to in the near future. Since the number of galaxies within a given area of the sky scales with d^3 , one can expect the performance gain from removing empty tiles to quickly fall off as the detection distance increases. If GW170817 were at 400 Mpc rather than 40 Mpc, for example, the same sky region would have contained $\mathcal{O}(10^5)$ galaxy candidates to get an integrated 95% detection probability rather than $\mathcal{O}(10^2)$. Galaxy catalogues are also significantly less complete at further distances: GLADE+

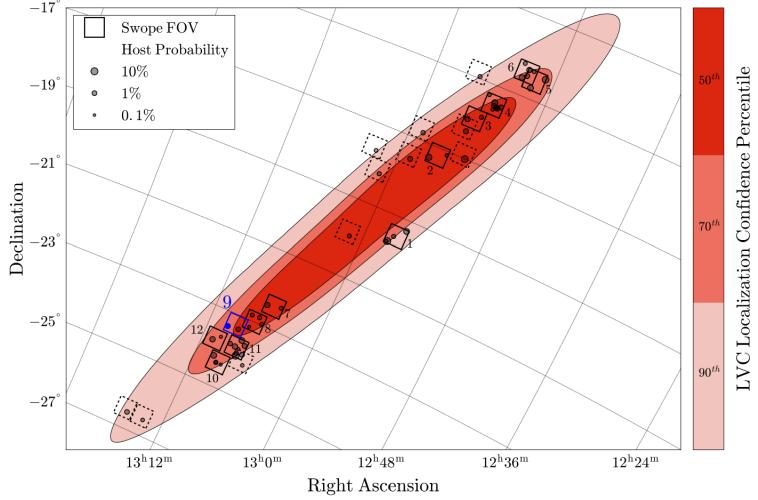


Figure 3-14: The narrow-field galaxy-targeting approach used by 1M2H to detect the optical counterpart of GW170817.

The boxes indicate exposure FOVs; the 9th exposure (blue) was the one containing AT 2017gfo, though 54 exposures were taken in total to rule out a false positive. Credit: [35].

(a catalogue specifically designed for follow-up searches of GW sources) is only 20% complete at $d_L = 800$ Mpc [37]. Some works have proposed a galaxy-survey-on-the-fly strategy, where relatively shallow observations from wide-field telescopes search for galaxies in the localization volume within 1-7 days of the merger, and then task narrow-field telescopes to follow up with deeper observations that could detect the kilonova itself [21]. Barring these more exotic modifications, the relative importance of galaxy-targeting approaches will lessen in the near future. Furthermore, simulation of the galaxy-targeting approach presents more of a challenge than a tiling approach, as it requires use of a galaxy catalogue and ranking algorithm—not to mention that the simulated BNS population would have to be generated based on the galaxy mass distribution from the catalogue rather than uniformly across the sky.

Tiling approaches have a suite of algorithmic methods to optimally map the sky localization region given arbitrary viewing constraints, localization region shapes, and FOV shapes. [143] and [79] together derive five different methods of generating a ranked tiling: [79] shows a fixed tiling approach that simply divides the localization region into a grid of squares based on the (square) telescope FOV, and a more so-

phisticated approach that then slides the tiling in right ascension and declination via iterative optimization. [143] assumes arbitrary telescopes and FOVs and uses different optimization methods including simulated annealing, greedy sort, and a noncoordinated approach to determine FOV placement and tile ranking. In principle, the tasking problem is simply a subset of well-known field of sensor tasking applicable to e.g. remote sensing [120] and space situational awareness [142], and condenses to a modeling problem with a corresponding global optimization algorithm.

In this work, an analytical approach is used which assumes optimal tiling is achieved, and simply calculates the detection probability based on the imaged area of sky A and the sky localization standard deviation σ_{sky} , as depicted in Fig. 3-15. This method is possible because the sky localization region is modeled as a bivariate Gaussian as per the FIM method output of Sec. 3.2.3. The $N\%$ c.r. corresponding to an area Ω of a bivariate gaussian is derived using the inverse survival function (isf) of the χ^2 distribution with $n = 2$ (two degrees of freedom) evaluated at $100\% - N\%$, or more simply $1 - \mathbb{P}$, where \mathbb{P} is the detection probability. The survival function (sf) is simply 1–cdf, the cumulative distribution function (the integral of the χ^2 probability density function, pdf):

$$f_{\chi^2\text{sf}}(x) = 1 - \int_0^x f_{\chi^2\text{pdf}}(\tilde{x})d\tilde{x} = 1 - \int_0^x \frac{1}{2}e^{-\tilde{x}/2}d\tilde{x} = e^{-x/2} \quad (3.23)$$

$$\Omega/\sigma_{\text{sky}} = f_{\chi^2\text{isf}}(x = 1 - \mathbb{P}) = -2 \ln x = -2 \ln(1 - \mathbb{P}) \quad (3.24)$$

$$\mathbb{P} = 1 - e^{-\frac{\Omega}{2\sigma_{\text{sky}}}} \quad (3.25)$$

The advantages of this approach are that it requires no optimization strategy for tiling placement and ranking, and it is flexible with respect to whether telescopes follow a coordinated or uncoordinated strategy (i.e., whether they cooperate to search the localization region and ensure their observations do not overlap, or whether they each

take an independently optimal approach). In the coordinated case:

$$\Omega = \sum_m \Omega_m \quad (3.26)$$

$$\mathbb{P} = 1 - e^{-\frac{\Omega}{2\sigma_{\text{sky}}}},$$

where Ω_m is the solid angle of sky successfully imaged by telescope m . In the uncoordinated case the detection probability is simply the maximum of the individual probabilities for each telescope:

$$\mathbb{P}_m = 1 - e^{-\frac{\Omega_m}{2\sigma_{\text{sky}}}} \quad (3.27)$$

$$\mathbb{P} = \max(\mathbb{P}_m).$$

This work assumes the uncoordinated case simply because that is the approach traditionally taken in practice, as evidenced by the fact that AT 2017gfo was discovered six independent times all within the span of ~ 1 hr [50]. However, it is noted that a coordinated approach would present a considerably more efficient use of telescope time; [143] predicts an average doubling of cumulative detection probability (across many sources) with a coordinated approach, though of course it depends on the exact EM network used.

Notably, both Eqn. 3.26 and 3.27 are constructed in such a way that enables parallel processing of EM network configurations, mirroring the parallelization of the FIM method for GW inference discussed in Sec. 3.2.3. Explicitly, this means that the performance of a network of telescopes is rapidly and analytically computable from the performance of each of its individual telescopes, which are themselves not necessarily analytical nor straightforward to compute. Therefore determining the performance of every telescope independently also determines the network performance in a parallelized fashion, providing a $\mathcal{O}(2^N)$ speedup for computing all network performances, where N is the number of all considered telescopes.

There are several assumptions implicit in this detection probability calculation. First, we assume the “discretization” of telescope FOVs is negligible—essentially, that the FOVs have no overlap and are small with respect to the localization region. [79]

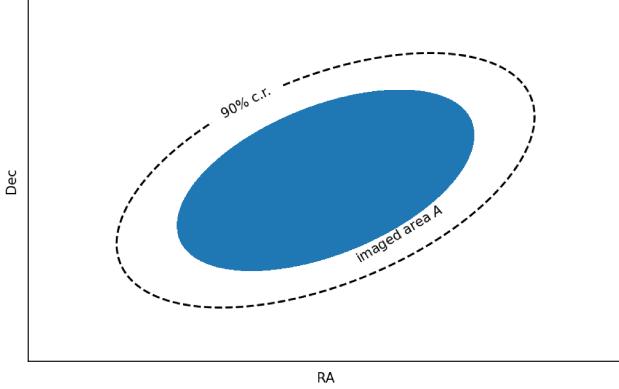


Figure 3-15: Analytical detection probability calculation.

This calculation assumes the imaged area is optimally distributed over the bivariate Gaussian sky localization region. In this example, the imaged area is smaller than the 90% c.r.; the detection probability is roughly 70%.

notes that this assumption can account for potentially a 10% difference in detection probability, since a large, monolithic square FOV is roughly 10% worse than many small, widely-distributed FOVs even for the same total area. Second, we assume each area of the localization region can be imaged equally as well; i.e. there are no differences in observability due to airmass or other constraints across the localization region. This assumption should hold well for regions less than $\mathcal{O}(10^2)$ deg 2 , where the angular separation between opposite sides of the region are less than $\mathcal{O}(10)$ deg apart. Lastly, it is assumed that false positives—transients identified in the localization region other than the KN source itself—either need not be ruled out by the initial sky search or are able to be ruled out based on the search images alone.

Observability criteria

The next step is to determine the area Ω_m that each telescope is capable of imaging in a search for the KN signal. This calculation uses the area of the FOV, A_{FOV} , and the number of exposures which meet observability criteria, n_{exp} , calculated by dividing the total duration the source is detectable with the duration dedicated to

each exposure:

$$\Omega_m = n_{\text{exp}} A_{\text{FOV}}$$

$$n_{\text{exp}} = \left\lfloor \frac{\int_{t=0}^{t=15 \text{ days}} \delta(\Theta) dt}{t_{\text{exp}} + \max(t_{\text{readout}}, t_{\text{slew}})} \right\rfloor, \quad (3.28)$$

where t_{exp} is the exposure time of each image, $\max(t_{\text{readout}}, t_{\text{slew}})$ is the time between successive exposures, and $\delta(\Theta)$ is a function that takes on the following values:

$$\delta(\Theta) = \begin{cases} 1 & \text{if all observability criteria are met} \\ 0 & \text{otherwise.} \end{cases} \quad (3.29)$$

Since $\delta(\Theta)$ depends on t_{exp} through a SNR criterion, an optimal follow-up strategy would maximize A_i by choosing the t_{exp} for each image that would globally maximize n_{exp} . This strategy would amount to constantly adjusting the exposure times so that each image just barely exceeds the minimum SNR threshold based on the predicted source brightness as well as the current observing conditions. In practice, most follow-up strategies (including all that successfully detected AT 2017gfo) use a fixed t_{exp} justified by adding a $\sim 1\text{-}3$ magnitudes margin to the predicted peak magnitude, such as [35]. Note that this strategy fails in cases where the KN source is dimmer than expected; [74] argues this fixed-exposure method was a probable cause of the failure to find the electromagnetic counterpart of GW190425. Nevertheless, following the lead of [116] and [81], this study adopts a fixed exposure time of $t_{\text{exp}} = 120$ s for ground-based telescopes and $t_{\text{exp}} = 2000$ s for space-based telescopes.

The full list of observability criteria include:

1. An SNR threshold $\rho \geq 10$
2. A minimum/maximum sun separation angle for space-based telescopes
3. Airmass and nighttime restriction for ground-based telescopes

An example calculation of $\delta(\Theta)$ with these criteria is depicted by Fig. 3-16. All conditions are individually checked for validity at each $(t_{\text{exp}} + \max(t_{\text{readout}}, t_{\text{slew}}))$ -duration interval. For computational efficiency, the SNR condition is checked first in

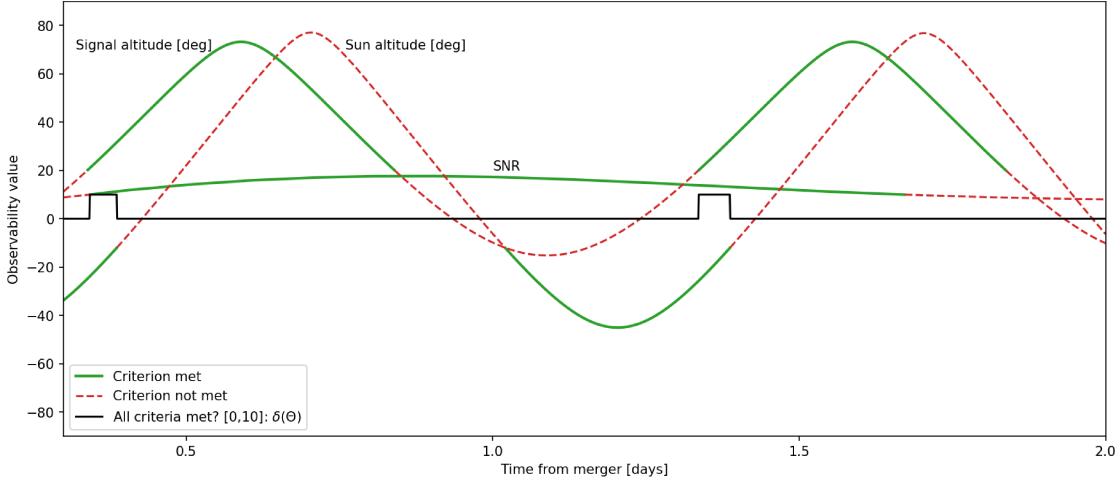


Figure 3-16: Observability criteria for a hypothetical signal observed by a hypothetical ground-based observatory.

In this example, the main constraint is that the signal and sun are nearby in the sky, so the signal is only observable for a brief period before sunrise. There are only two of such windows before the signal becomes too faint to successfully image based on the SNR threshold.

0.1 day increments, and then all conditions are only rechecked for the higher-fidelity set of times where the SNR threshold is met, adding additional margin to ensure the lower-fidelity first pass did not incorrectly exclude any portions of time. For the sake of modeling simplicity, other possible criteria that are notably excluded include cloud/weather constraints for ground-based telescopes, Earth eclipse or other Earth viewing constraints for space-based telescopes (which would apply mostly to observatories in LEO), lunar phase, duty cycle considerations, and observatory prioritization. The last point is perhaps the most noteworthy: this work assumes each observatory is available “on-demand” at any point in time for BNS follow-up and that it has no other higher-priority observations to attend to.

For (1), a minimum SNR of 10 must be met for the signal to be considered observable. Calculation of SNR is discussed in the subsequent portion of this section. For (2), values for the minimum sun separation angle are taken from the engineering documents for each observatory, such as 55 deg for Hubble (HST) [2]. Plausible values are used where otherwise unavailable, such as assuming JWST’s sun separation angle for LUVOIR-A and LUVOIR-B, which are anticipated to utilize a similar sunshield

design. Note that JWST additionally has a maximum sun separation angle, which is also taken into consideration in this work; this restriction typically does not apply for enclosed telescopes like HST. The sun separation angle calculation relies on the `AstroPy.coordinates` module [48] and assumes a geocentric observer, which applies well if the telescope were in Earth orbit, at ESL (Earth-Sun Lagrange) points 1 or 2, or otherwise nearby Earth. For (3), at the time and location of observation it must be astronomical twilight or darker (sun altitude < -12 deg) and the signal altitude must be > 20 deg such that airmass $< \sec(90^\circ - 20^\circ) \approx 3$, taken directly from the observing strategy used by [35] for GW170817 follow-up. These values are similarly computed using the `AstroPy.coordinates` module.

SNR calculation

A standard SNR calculation for a CCD-based UV/optical/NIR telescope is as follows:

$$\text{SNR} \equiv \frac{S}{N} = \frac{n_{\text{signal}}}{\sqrt{n_{\text{signal}} + \text{pix}(n_{\text{sky}} + n_{\text{dark}} + n_{\text{read}}^2)}}, \quad (3.30)$$

where n_{source} is the electron count per pixel, e^-/pix , from a particular noise or signal source. n_{sky} applies for ground-based telescopes but not space-based telescopes. Expressed in terms of a photon count rate ϕ and CCD quantum efficiency η , values are given by:

$$\begin{aligned} n_{\text{signal}} &= \phi_{\text{signal}} t_{\text{exp}} \eta \\ n_{\text{sky}} &= \phi_{\text{sky}} t_{\text{exp}} \\ n_{\text{dark}} &= \phi_{\text{dark}} t_{\text{exp}}. \end{aligned} \quad (3.31)$$

The read noise is not represented as a photon count rate because it applies on a per-readout basis rather than per unit time. pix is the number of pixels used for source photometry, which is determined via the Rayleigh criterion for space-based telescopes, and by the maximum of the Rayleigh criterion and the atmospheric seeing condition ($\sim 0.5^\circ$) for ground-based telescopes:

$$n_{\text{pix}} = \left\lceil \frac{\pi}{4} \left(\frac{\max(1.22 \frac{\lambda}{D}, 0.5^\circ)}{p/f} \right)^2 \right\rceil \quad (3.32)$$

Telescope parameters such as the read noise, dark noise, pixel pitch p , focal length f , and aperture diameter D are either taking from the observatory engineering documents or an assumed default value is used (e.g. a pixel pitch of 15 μm).

The photon count rates ϕ_{signal} and ϕ_{sky} are derived from their AB magnitudes in particular bandpasses by first converting to a photon flux with units of $\text{W m}^{-2} \text{ Hz}^{-1}$ and then to a photon count rate by using the mean photon energy $E = \frac{hc}{\lambda_{\text{mean}}}$ and bandwidth $\Delta f = f_{\text{max}} - f_{\text{min}}$. Quantities for the bandpasses, bandwidths, and sky background AB magnitudes by bandpass are used from [167].

3.2.6 Calculating the detection rate

To summarize, the intermediate output at this stage of the simulation is:

- ρ_{kl} , the SNR of GW interferometer l observing merger k ,
- $\mathcal{I}_{ij,kl}$, the ij components of the Fisher information matrix \mathcal{I} for interferometer l observing merger k , and
- Ω_{km} , the solid angle of sky telescope m is able to search to find merger k ,

computed $\forall k \in [1 \dots N_k]$ for N_k simulated BNS merger signals, $\forall l \in [1 \dots N_l]$ for N_l total considered interferometers, and $\forall m \in [1 \dots N_m]$ for N_m total considered telescopes. The total number of instruments is $N = N_l + N_m$, from which it is possible to construct $2^N - 1$ possible networks (excluding the empty set).

We wish to determine the performance of a particular network—a particular subset of the interferometers and telescopes—based on these quantities. Suppose our network is comprised of a set of interferometers $L \subseteq [1 \dots N_l]$ and a set of telescopes $M \subseteq [1 \dots N_m]$. The GW network SNR for observing merger k is given by Eqn. 3.19:

$$\rho_{kL} = \sqrt{\sum_{l \in L} \rho_{kl}}. \quad (3.33)$$

The localization error is given by Eqn. 3.16, 3.18, and 3.21:

$$\begin{aligned}\mathcal{C}_{ij,kL} &= \left(\sum_{l \in L} \mathcal{I}_{ij,kl} \right)^{-1} \\ \sigma_{kL} &= \sqrt{\mathcal{C}_{\alpha\alpha,kL} + \mathcal{C}_{\delta\delta,kL}}.\end{aligned}\tag{3.34}$$

Lastly, the probability of detection by the EM network is given by Eqn. 3.27:

$$\mathbb{P}_{kLM} = \begin{cases} \max_{m \in M} \left(1 - \exp \left(-\frac{\Omega_{km}}{2\sigma_{kL}} \right) \right) & \rho_{kL} \geq 12 \\ 0 & \text{otherwise.} \end{cases}\tag{3.35}$$

Recall that this probability assumes detection by the GW network, so indicates only joint detection of both GW and EM signals. The network performance—the expected value of the fraction of successfully-detected events—is calculated as

$$p = \frac{1}{N_k} \sum_k \mathbb{P}_{kLM},\tag{3.36}$$

where the detection rate is simply $r = \frac{4}{3}\pi d_{C,\max}^3 \mathcal{R}_{\text{BNS}} p$. The horizon distance d_h at which 50% of sources are detected is found via the solution to

$$0 = 0.5 - \frac{1}{|K|} \sum_{k \in K} \mathbb{P}_{kLM}\tag{3.37}$$

given $K \equiv \{k \in [1 \dots N_k] : \theta_{k,d_L} \approx d_h\}$, a subset of signals (recall, each with merger parameters θ_k and luminosity distance θ_{k,d_L}) that are approximately at the horizon distance.

Between these computations, Eqn. 3.34 and Eqn. 3.37 are the most expensive due to requiring a matrix inversion and root finder respectively. Otherwise, they are vectorizable and efficient to compute for a large number of network configurations.

3.3 Results

3.3.1 Joint EM/GW detection rates through the 2020s

As of 2023, the observing plan for the International Gravitational-Wave Observatory Network (IGWN) — comprised of the United States’ LIGO Hanford and LIGO Livingston, the European Union’s Virgo, and Japan’s KAGRA — is determined through the 2020s. Observing runs are set so that all detectors can be simultaneously online for collecting science data or offline to undergo planned upgrades. Fig. 3-17 shows the schedule for these observing runs along with the achieved or anticipated binary neutron star horizon distance for each instrument. This section will refer to observing runs O1 (2015) through O5 (2027-2029).

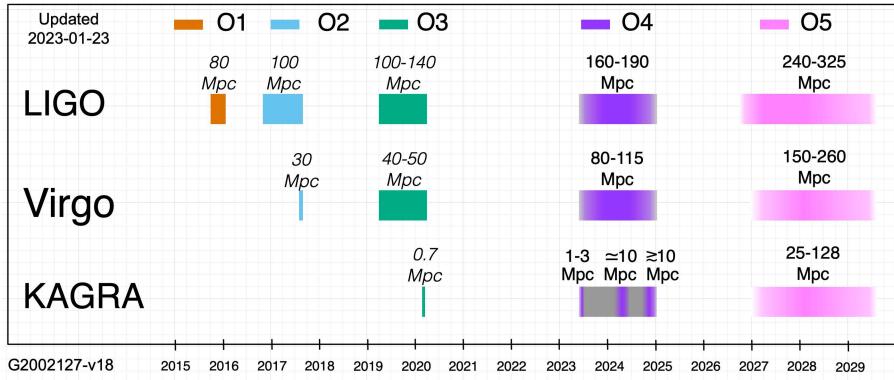


Figure 3-17: IGWN observing runs through the 2020s. Credit: [6].

An anticipated multi-messenger detection rate can be simulated by using the achieved or anticipated interferometer PSDs for each of these runs along with a corresponding EM network and a hypothetical (randomly-generated) neutron star merger population. This procedure will also help to validate the simulation by comparing against other detection rate predictions. Note that these simulations were run with $|K| = 3000$ mergers for sufficient statistical sampling as well as statistical bootstrapping of anticipated merger detection rates. The detection rate is then calculated based on the observed fraction of mergers and the actual merger rate density within the simulated volume. The network used for this simulation is shown in Fig. 3-18. The telescopes considered are all ground-based, and are comprised of instruments

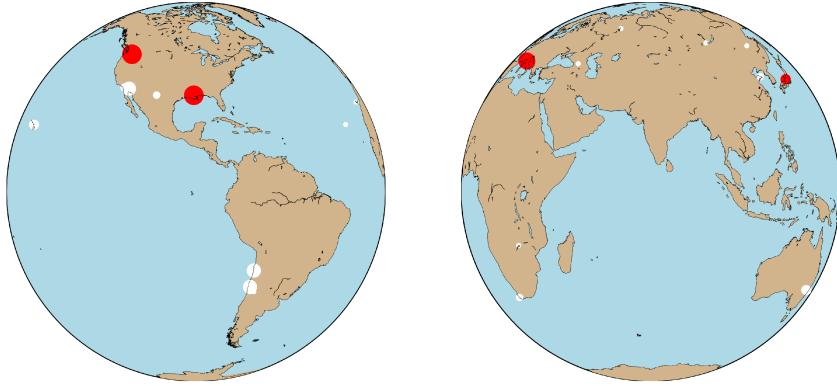


Figure 3-18: Instrument network used for the 2020s simulation. There are 24 instruments total including 4 GW interferometers (red points) and 20 EM telescopes (white points). Points are scaled by relative instrument sensitivity.

that were used to observe GW170817 and GW190425 along with several other that were more recently commissioned such as the Wide-Field Infrared Transient Explorer (WINTER) [76]. A comprehensive list of instruments is given in Tables A.1 and A.2 (which also lists those used in Sec. 3.3.2, and so includes more instruments than were considered in Sec. 3.3.1 alone).

First, we can examine the portion of mergers detectable by the GW network exclusively. Fig. 3-19 shows the detection predictions for O2 (observing run 2). GW170817, the BNS merger that was actually detected in O2, is also displayed for reference. This event was not explicitly written into this simulation’s merger population, but was run in a separate analysis also using the Fisher method. What this simulation confirms is that GW170817 was an outlier: based on the predicted merger rate density, it represents a once-in-a-decade occurrence at best in terms of its luminosity distance.

Note that while GW170817’s SNR is roughly within the predicted distribution relative to similarly-distanced events, its actual SNR was $\rho = 32.4$ [51] rather than a predicted $\rho = 39.0$ when simulating GW170817 with the Fisher method directly. This 20% discrepancy is quite reasonable: the Fisher analysis provides only an approximation of SNR and so will not be precisely accurate when analyzing any individual

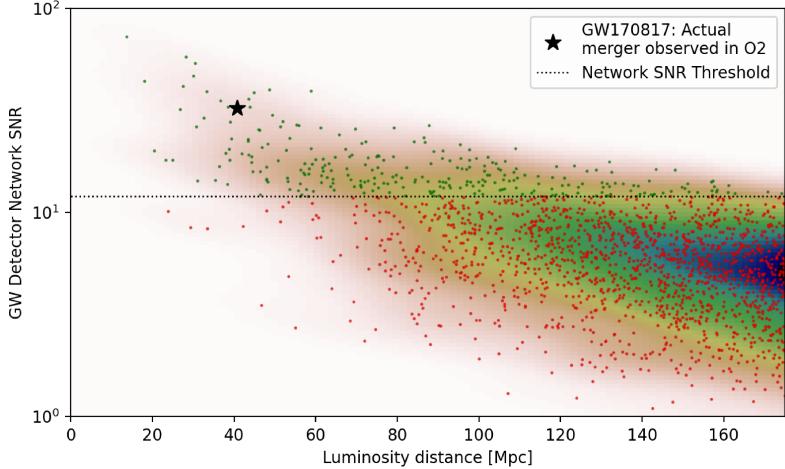


Figure 3-19: GW network SNR of simulated mergers during O2. The actual observation of GW170817, with a measured SNR of 32.4, is shown for reference.

event. The method is more suited to population statistics where estimate noise matters less than estimate bias. The marginalized predicted SNR at this distance can be computed via a Gaussian kernel density estimate on the simulation results from Fig. 3-19, as shown in Fig. 3-20. The median marginalized SNR, calculated to be $\rho = 19.3$, suggests GW170817’s SNR was quite favorable compared to other possible events at a similar luminosity distance. GW170817 was by all considerations a lucky event.

Next, we can determine the EM observability of these simulated events by calculating the sky localization regions. Note that these regions only apply where each merger met the GW network SNR threshold. Fig. 3-21 shows the area of the 90% confidence region of these regions. Again, running the simulation against GW170817 provides a useful point of validation. GW170817 had an initial sky localization region of 28 deg^2 after the first `BAYESTAR`-derived localization, which was later updated to 16 deg^2 in a more comprehensive `LALInference` analysis [51]. The 13 deg^2 region predicted by the Fisher analysis compares reasonably to the actual area, especially owing to the higher SNR predicted by the Fisher method.

Further necessary for determining EM observability are the kilonovae light curves, which are simulated as per [115]’s method described in Sec. 3.2.4. Fig. 3-22 shows

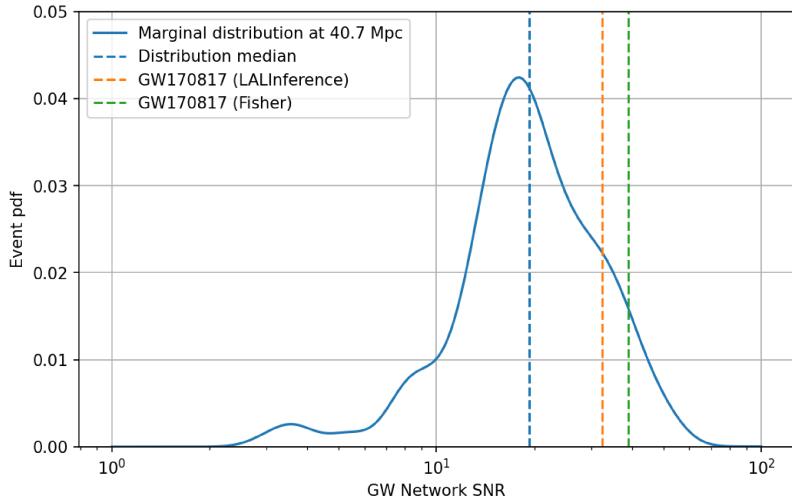


Figure 3-20: Marginal distribution of merger SNRs in O2 for simulated events at 40.7 Mpc.

40.7 Mpc is the estimated distance of GW170817. GW170817's actual SNR, whether via the official LALInference analysis [51] or the Fisher method used in this work, lies roughly 1σ above the median.

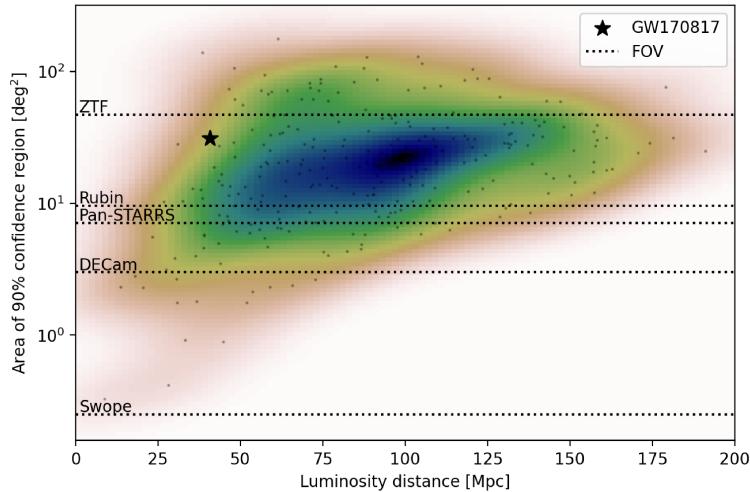


Figure 3-21: Sky localization regions of detectable mergers during O2. The initial 28 deg^2 sky localization of GW170817 is shown for comparison.

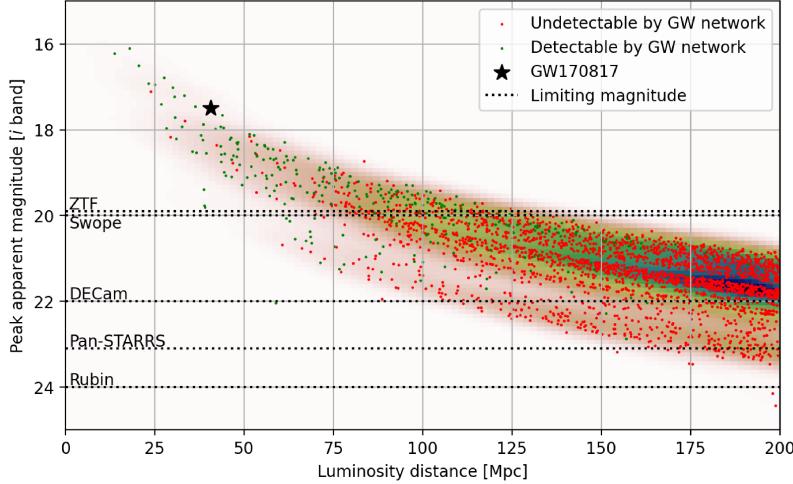


Figure 3-22: Apparent magnitudes of the simulated merger population. The distribution bimodality is due to divergent kilonovae evolution; some kilonovae immediately collapse into black holes, whereas others form a semistable hypermassive neutron star.

these apparent magnitudes, along with which portion of the mergers are detectable by the O2 GW network, and therefore which are eligible for EM follow-up. Similarly, the marginal distribution of peak magnitudes at 40.7 Mpc can be computed with a Gaussian kernel density estimate, shown in Fig. 3-23. This marginal distribution compares the known GW170817 peak i band magnitude (17.5 [35]) with the simulation results (median peak magnitude of 18.4). Directly simulating GW170817 gives a peak magnitude of 18.2. It is thought that GW170817 was a brighter-than-typical event [115], but the more significant explanation for this discrepancy is that the light curve simulation conservatively estimates kilonovae brightness.

By performing the same procedure for the baseline GW and EM network in O3, O4, and O5, we can predict the expected quantity of jointly-detected BNS mergers through the rest of the 2020s. Fig. 3-24 shows the cumulative number of detections by observing run, starting from O3 (which has already occurred, but is included retroactively in this analysis for the sake of validation). The 68% confidence region ($\pm 1\sigma$) is included for (1) the underlying merger rate density of all events and (2) the merger sampling, i.e. variations in events actually detected. Poor constraints on the merger rate density produces a significant uncertainty contribution as the rate

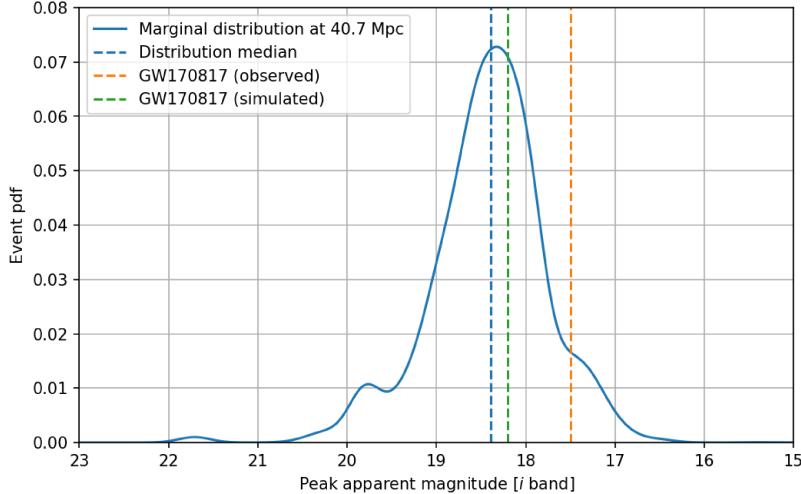


Figure 3-23: Marginal distribution of peak apparent magnitudes at 40.7 Mpc. While GW170817 is in-distribution when simulated with the same light curve model, its actual observed magnitude is in the 90th percentile of kilonovae. This discrepancy is likely due to simulation conservatism.

density is only known within $^{+153\%}_{-75\%}$ 1σ bounds [65]. All the same, the zero¹⁴ BNS mergers actually jointly detected in O3 is within the predicted 2.0^{+3}_{-2} bounds. By the end of O4, 9.9^{+16}_{-8} cumulative merger detections are anticipated. The end of O5 may see 46.1^{+71}_{-34} total observations.

Notably, from O4 onward (including the single joint detection so far, GW170817, and excluding the O3 prediction of ~ 2 joint detections) we anticipate 45^{+68}_{-32} jointly-detected BNS mergers by the end of the decade. The median scenario of 45 detections will be just barely short of the ~ 50 source threshold required for a 4% measurement of H_0 [102], the degree sufficient to resolve present tensions in the measurements of H_0 from the CMB and Type Ia supernovae.

3.3.2 Mapping future multi-messenger networks

In addition to investigating near-term modifications and upgrades to the present GW/EM network, we can also use the method outlined in this chapter to investigate the science enabled by next-generation instruments and beyond. In a second simu-

¹⁴GW190425 was a BNS merger observed in O3. However, it was only observed in the GW spectrum, not the EM spectrum.

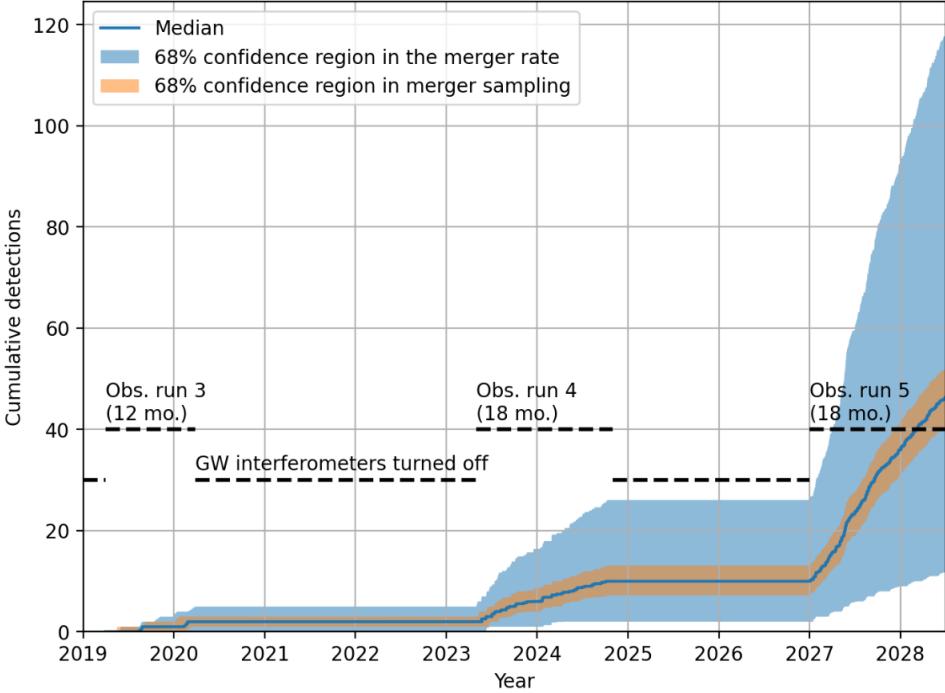


Figure 3-24: Predicted jointly-detected binary neutron star mergers for IGWN observing runs through the 2020s.

45^{+68}_{-32} BNS merger detections are anticipated by the end of the decade.

lation, we consider a suite of 33 ground- and 4 space-based telescopes along with 14 next-generation interferometers (including one space-based interferometer). Fig. 3-25 shows a map of the ground-based portion of this instrument portfolio. Tables A.1 and A.2 give a complete list of the instruments.

The number of networks it is possible to construct with this set of 51 instruments is $2^{51} \approx 10^{15}$, which becomes prohibitively challenging to visualize, let alone analyze or use to derive useful insights. Instead of using all possible networks via the power set of these instruments, only a subset of the networks are considered. Subsets of instruments are further partitioned in **AND** groups and **OR** groups. In **AND** groups, that subset of instruments is either entirely within the network or entirely excluded from it. This grouping is useful for observatories or observatory proposals which collectively operate a set of instruments, such as the MASTER Global Robotic Network of telescopes or the One-Meter Two-Hemispheres (1M2H) telescopes. In **OR** groups, one can have at most one instrument from the grouping. This category is

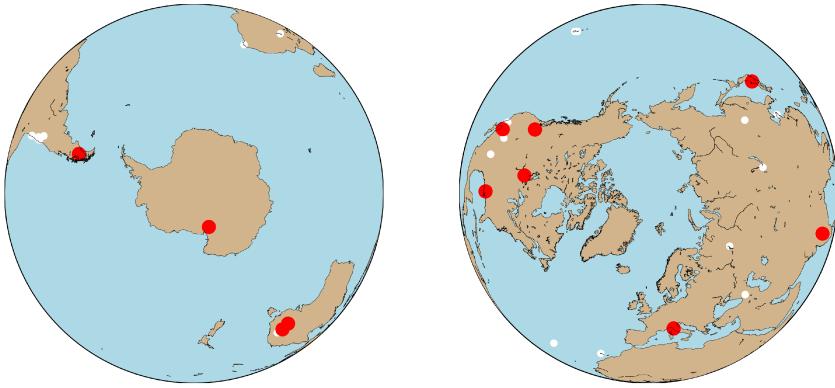


Figure 3-25: Ground-based portion of the instruments considered. Red points indicate interferometers; white points indicate telescopes.

useful where certain instrument upgrade paths or proposals are incompatible with each other, such as downscoped/upscoped versions of what is otherwise the same instrument. This strategy of grouping allows us to arrive at 647 unique EM networks and 647 unique GW networks (excluding the empty set), for a total of $647^2 \approx 4 \times 10^5$ possible multi-messenger networks.

The simulation is used to compute the performance of each unique network in a parallelized fashion. In this case, only BNS mergers at a redshift lower than $z = 0.3$ are considered, as the kilonovae light curve computation did not take into account a redshift correction factor. Fig. 3-26 shows the performance for each of these unique networks. On the left, unique combinations are not sorted, but are arranged by an increasing quantity of instruments (from top-left to the bottom-right). On the right, the performance is sorted by the median performance along that ‘slice’ (the median along the column for sorting interferometer combinations, and the median along the row for sorting telescope combinations).

While Fig. 3-26 shows a general comparison of network performance, it is not directly useful without understanding which instruments grant the network the best improvement in performance. This performance delta can be computed from the data above by considering for each instrument: (1) the performance of each network which

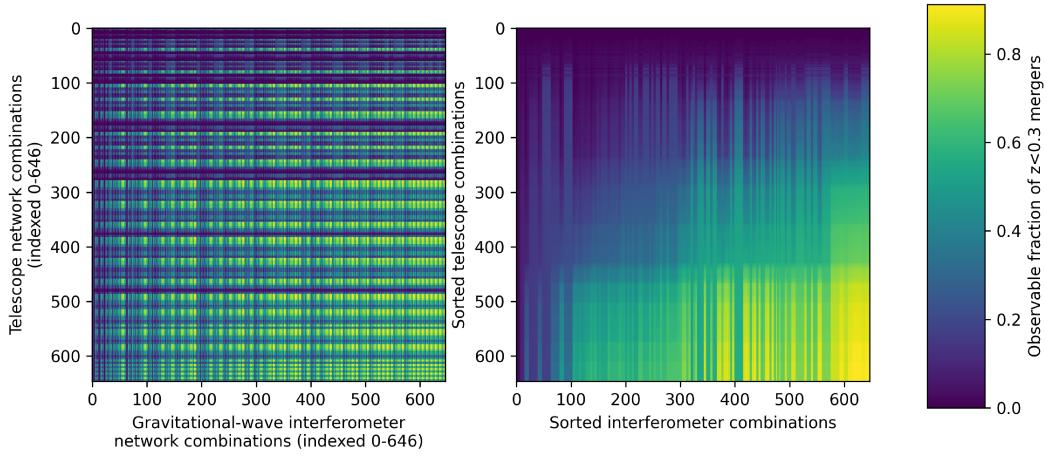


Figure 3-26: Performance of different multi-messenger networks. Each pixel represents a unique combination of instruments.

does not have that instrument, and (2) the performance of each network which does have the instrument. There is a one-to-one mapping between set (1) and set (2), as every network “pair” can be subjected to either an addition or removal of the considered instrument. Since the performance delta is dependent on the particular network configuration, it is most useful to not exclusively consider the mean performance between set (1) and (2), but rather the individual changes.

Fig. 3-27 and Fig. 3-28 show these data as scatter plots where the abscissa indicates the network performance without the considered detector and the ordinate indicates performance when including the detector. Here, performance is again defined as fraction of jointly-observable $z \leq 0.3$ mergers. A scatter point on the diagonal identifies that performance does not change when adding that particular detector to the mix; a scatter point far above the diagonal identifies a large performance improvement. As a useful point of validation, no points lie below the diagonal, beyond a select few which are due to numerical imprecision incurred when inverting Fisher information matrices.

The right-most subplot shows the full case of every network pair: each scatter point represents a unique network composition both with and without a given instru-

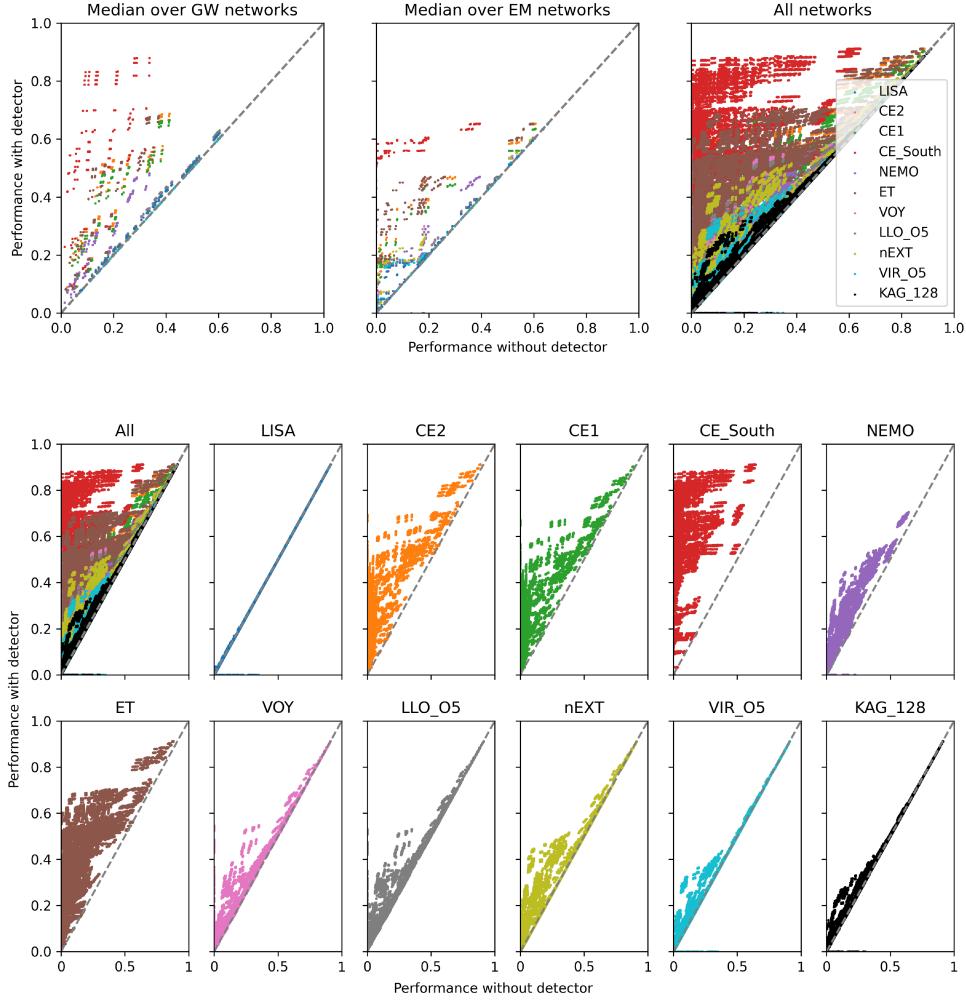


Figure 3-27: Network performance gain due to individual GW interferometers.

ment. The central subplot shows the performance of these pairs where the median performance is taken over all combinations of EM networks. For a particular instrument, each scatter point will therefore exclusively refer to separate compositions of the GW network, and any concurrent changes in the EM network are factored out. The left-most subplot performs this equivalent operation but for the median GW network performance instead.

By following the curves for each instrument in the separate plots, one can infer how the performance delta of an instrument changes as the underlying network composition is altered. For example, consider Zwicky Transient Facility (ZTF) and the

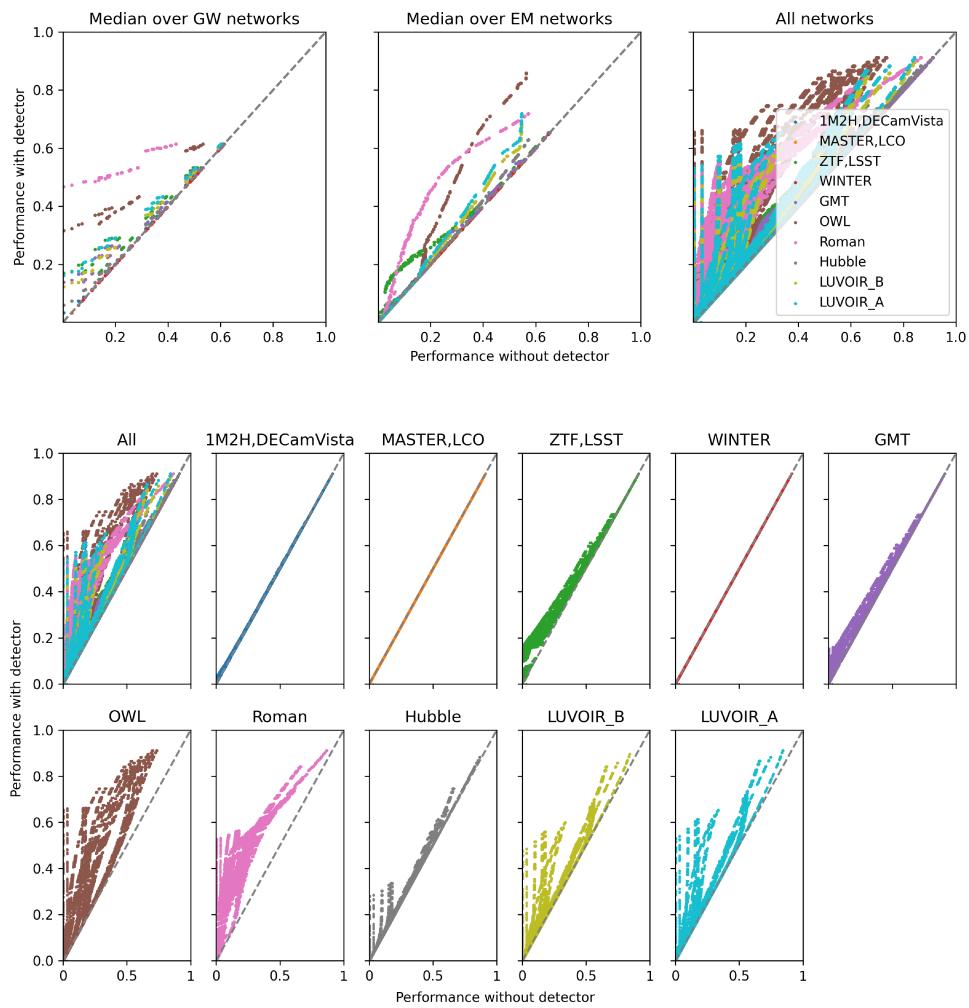


Figure 3-28: Network performance gain due to individual EM telescopes.

Rubin Observatory (LSST) from Fig. 3-28. The central subplot suggests that when one has either a very low-performing or high-performing GW network, little to no performance is gained by adding these instruments to the multi-messenger network. There is some intermediate composition where the most performance is gained. While unintuitive at first, this factor can be explained by the directionality of the observations: the GW network cues EM follow-up searches. For low-performing networks, the network will be constrained by the quantity of initial cues, and EM observations will saturate the number of events that one can detect. Conversely, at some point of GW network performance, most of the new sources cued will be from beyond the horizon of the EM instrument. At this point, additional cues to that instrument do not grant more performance. These factors result in a general bell curve for performance delta from telescopes: each telescope is tuned to a specific performance range based on its limiting magnitude and FOV.

As another example, consider the case of the Neutron star Extreme Matter Observatory (NEMO) interferometer shown in the central subplot of Fig. 3-27. NEMO represents a promising 2.5-generation interferometer, one intended to bridge the present gap between the end of O5 in 2029 and the commissioning of 3rd-gen detectors such as CE and ET sometime in the 2030s. NEMO is designed to have an unusually high-frequency sensitivity. This tuning is particularly useful for localization due to the same factors that influence the Rayleigh criterion for traditional telescopes. In terms of performance scaling, NEMO demonstrates that as one improves the overall GW network composition, the performance gain from a particular interferometer trends towards zero.

We can also look at other individual instruments more closely. Fig. 3-29 compares the WINTER telescope to the MASTER/LCO (Las Cumbres Observatory) networks of instruments. The WINTER telescope is able to match or outperform the MASTER/LCO network owing to its relatively wide field of view along with its NIR sensitivity. Kilonovae emissions are generally sustained longer in the infrared bands than optical or UV bands, which permits more time for a follow-up sky search; see Fig. 3-12. On average, adding WINTER to any given network improves the BNS

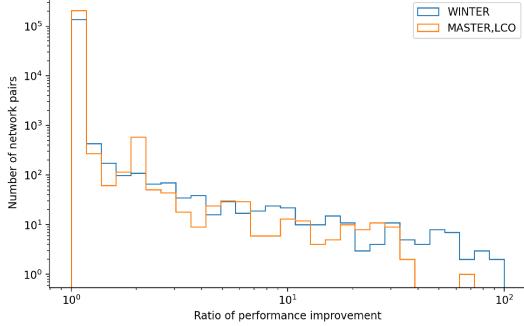


Figure 3-29: Comparing performance improvement with the WINTER instrument and the MASTER/LCO network.

The ratio shown compares the performance with and without the particular set of instruments.

detection rate by 83%.

Overall, the Roman Space Telescope stood out as a clear outlier in terms of individual telescope performance. As a caveat, this analysis assumes that Roman would be available “on-call” to perform follow-up of BNS mergers no matter the observing priority relative to other possible targets. Nevertheless, this analysis strongly suggests that a wide-field space telescope for BNS follow-up would strongly boost the quantity and depth of mergers that we can access. This performance increase would occur largely irrespective of the underlying network composition, also indicating a longevity to Roman’s useful lifetime.

Wide-field instruments such as ZTF and Rubin currently dominate the GW follow-up landscape owing to their considerable optical throughput which permits them to quickly search large areas of the sky. However, this simulation shows that as the GW horizon distance expands, we can expect a trend towards more sensitive, smaller-FOV instruments. For example, LUVOIR A and B — two proposals for a Large UV/Optical/IR space telescope — were some of the highest-performing instruments despite a meager 0.0017 deg^2 field of view. This shift in telescope capabilities is due to the reduced sky localization regions generated by higher-performing GW networks. This shift may be somewhat counterintuitive: as GW sensitivities improve, most events are still at the network SNR threshold since mergers are distributed by volume rather than by distance, so one might expect the sky localization regions to similarly remain

the same. Eqn. 3.10, however, shows that the regions also scale with the observed frequency. Here, more sensitive detectors also means a higher maximum observed frequency, which permits better localization.

On the interferometer side, the most promising instrument is a hypothetical variation of the Cosmic Explorer interferometers, here named CE_South. CE_South moves the two proposed Cosmic Explorer sites from the U.S. to the Southern hemisphere, placing one at McMurdo Station in Antarctica¹⁵ and another in Western Australia at a candidate interferometer site [58]. In addition to the fact that this pairing represents a strong 3rd-generation set of interferometers, placement in the Southern hemisphere greatly boosts the geographic diversity of the interferometer network. This geometric factor permits more tightly constrained sky localization regions across the entire sky.

3.4 Discussion

This chapter presented a simulation of joint gravitational-wave and electromagnetic observations of binary neutron star mergers. The quantity of these joint observations is itself an important metric for a variety of open questions in astrophysics and cosmology, in particular directly determining the uncertainty of a gravitational-wave standard siren measurement of the Hubble parameter. The simulation undertaken in this chapter suggests that a 4% measurement of H_0 is possible by the end of the decade — a level of precision sufficient to resolve the $4 - 6\sigma$ tension in existing measurement methods — which may open the path to physics beyond the standard model. This prediction assumes that the IGWN undergoes planned upgrades, meets or exceeds performance targets, and that the present maximum likelihood estimate of the merger rate density is close to the truth.

Furthermore, performance of over 10^5 multi-messenger observatory network configurations was modeled. This data set shows the scaling laws for performance gains from

¹⁵The engineering merit of building an interferometer in Antarctica is not discussed here, though it is noted that such a detector would presumably find great comfort among multi-messenger brethren such as the IceCube Neutrino Observatory, which is also stationed at the South Pole.

individual telescopes and interferometers, and can be used for generating actionable insights about future multi-messenger networks. In particular, geographic diversity and high-frequency sensitivity are two important factors for the gravitational-wave network composition. As interferometer sensitivities improve, we can expect a trend from wide-field follow-up telescopes to highly-sensitive narrow-field telescopes. The best configurations of networks utilize a highly-sensitive and widely-distributed set of ground-based interferometers in tangent with narrow-field, large-aperture ground-and/or space-based telescopes.

Chapter 4

Joint radar and optical measurement of resident space objects

Resident space objects (RSOs) are artificial or natural objects in orbit around a given body, in this case the Earth. These objects may include both active and inoperative satellites, other debris generated from human activity in space, and natural objects like micrometeoroids. With the ever-increasing use of low-Earth orbit (LEO) for commercial purposes, the concurrent growth of the space debris population poses a heightened threat to the safety and operations of spacecraft in orbit. As such, tracking of these objects is required to maintain a clear and actionable understanding of the environment, in particular for the purposes of collision avoidance. This ability to monitor and predict the behavior of RSOs is known as space situational awareness (SSA).¹ Fig. 4-1 depicts the significant recent growth in Earth RSOs.

The standard orbit determination method is sustained through range and range-rate measurements via ground-based radars like those of the Space Fence, which is part of the U.S. Space Surveillance Network (SSN) maintained and operated by the US Department of Defense. In some cases, ground-based optical telescopes like GEODSS (Ground-Based Electro-Optical Deep Space Surveillance) are also used for surveillance of objects in medium-Earth orbit (MEO) and beyond [84]. The Space Surveillance Network maintains a catalogue of $\mathcal{O}(10^{4.5})$ objects over 10 cm in size,

¹Or more recently, space domain awareness (SDA), though here the former acronym is used.

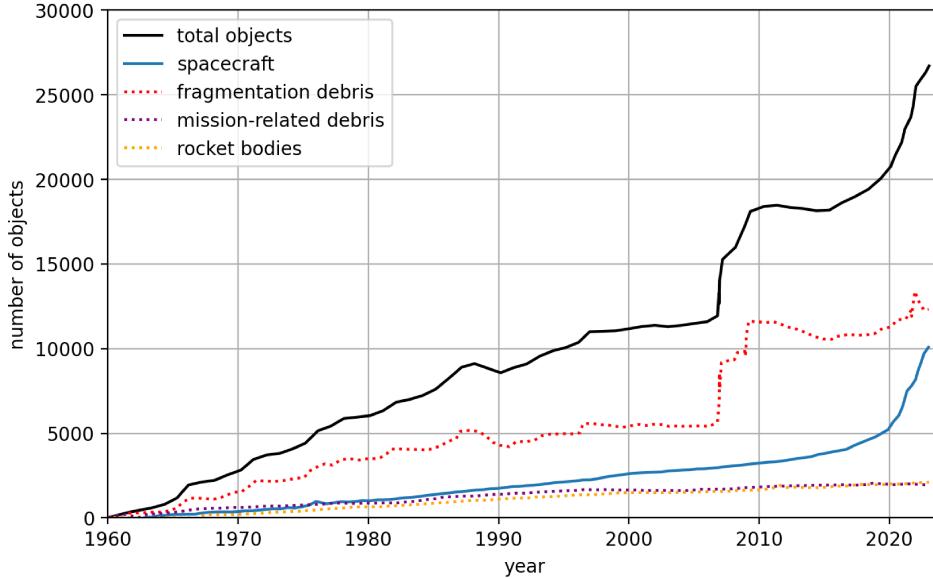


Figure 4-1: Monthly number of objects in Earth orbit greater than 10 cm in size. These include all objects tracked by the SSN. Large spikes in fragmentation debris in 2007 and 2009 correspond to a Chinese ASAT test and accidental collision of the defunct Cosmos 2251 and active Iridium 33 satellites, respectively. The 2020-onward growth in active spacecraft is largely attributable to the deployment of the Starlink constellation. Adapted from [118].

representing only a small portion of the $\mathcal{O}(10^6)$ pieces of debris over 1 cm in size, all of which may be lethal to an active spacecraft in the event of a collision [10]. Furthermore, even if an object is known, its orbital state is not always known with sufficient certainty for a satellite operator to feasibly maneuver to avoid a potential collision [140].

There is a growing understanding in the global spaceflight community that existing SSA capabilities are insufficient to meet present and future needs of satellite operators in LEO. In order to mitigate the effect of space debris, a clearer understanding of the environment is needed. Simultaneously, the Department of Defense generally intends to offload the work of SSA provision onto more commercially-focused entities in accordance with Space Policy Directive-3 (SPD-3) [1]. SPD-3, issued in 2018, calls for a new method of space traffic management (STM) and designates the US Office of Space Commerce as the responsible agency for future STM development.

One method proposed in a DARPA solicitation called Space-WATCH addresses

both of these concerns simultaneously. The method calls for leveraging pre-existing star trackers on commercial satellites in LEO to monitor RSOs [3]. Though these star trackers may not actively track RSOs by gimbaling to known targets, they are likely to coincidentally image RSOs during routine operation, which represents potentially useful data which would otherwise be discarded as noise. These on-orbit star trackers which generally operate in the optical portion of the EM spectrum could complement ground-based radar to sustain broad and actionable SSA in LEO.

This case study concerns the efficacy of a cohesive network of ground-based radar and untasked on-orbit optical sensors to track RSOs in LEO. Though these sensors do not strictly concern multiple astrophysical messengers, radar and optical observations are sufficiently multimodal² so as to represent a worthy case study. In particular, the photons measured by each sensor are unique in their origin, frequency, and behavior. While radars actively emit and receive their own low-frequency photons reflected off the targets of interest, optical sensors observe relatively high-frequency photons from the Sun which are reflected off the source, or in some cases photons self-emitted by the target objects.

4.1 Object covariances

A primary use of orbit determination data output is to alert satellite operators to potential collisions (“conjunctions”). In the event of a conjunction, operators receive conjunction data messages (CDMs) through the 18th Space Defense Squadron or commercial providers like LeoLabs via conjunction screening services, which informs them of details such as object states, covariances at closest approach (Eqn. 4.1), and calculated collision probability. Based on these data, operators may choose to maneuver to avoid the collision to reduce residual risk, or may choose to simply accrue the risk if the CDM is deemed inactionable.

Data actionability is a holistic measure which depends on the reliability, accuracy, and timeliness of the underlying measurements. The most important factors for

²Multimodal in a data fusion sense, as discussed in Sec. 2.4.

actionability are the size and realism of the object positional covariances at the time of conjunction. The covariance matrix \mathcal{C} of an RSO indicates the uncertainty of knowledge in each element of its state vector $X = [x, y, z, \dot{x}, \dot{y}, \dot{z}]$. The 3×3 upper-left portion of the covariance matrix gives the positional uncertainty:

$$\mathcal{C}_{\vec{r}} = \begin{pmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 & \sigma_{xz}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 & \sigma_{yz}^2 \\ \sigma_{zx}^2 & \sigma_{zy}^2 & \sigma_{zz}^2 \end{pmatrix}, \quad (4.1)$$

essentially a 3-dimensional uncertainty ellipsoid at a certain point of time, where the size of the positional uncertainty is calculated as:

$$\sigma_{\vec{r}} = \sqrt{\text{trace}(\mathcal{C}_{\vec{r}})}, \quad (4.2)$$

largely similar to Eqn. 3.21. Even in situations with a low $\sigma_{\vec{r}}$, another important factor is the realism of the covariance, which regards how accurately a given covariance ellipsoid represents the true uncertainty of the underlying phenomena. One measure of covariance realism is the Mahalanobis distance, essentially a multivariate extension of the concept of a statistical standard score (“Z-score”), which informs how far out-of-distribution an object is measured from its predicted position (the center of the covariance ellipsoid) [155].

Covariance sizes and realism are an important performance measure for SSA networks because poor covariances are widely detrimental to effectual on-orbit collision avoidance. If the state covariances at close approach are poor, for example, wider maneuvers are required to avoid the conjunction. With finite resources to dedicate towards collision avoidance, this results in less-actionable data. Secondly, unrealistic covariances inflate the rate of false positive and false negative maneuver decisions: essentially, they cause unnecessary maneuvers, and conceal situations where maneuvers may actually be warranted. For stationkeeping satellites, large covariances imply a need to exit the nominal stationkeeping orbit to avoid a conjunction, which may introduce additional traffic management complications. Even in the event of maneuvering

for essentially every conjunction, poor covariance uncertainty and realism result in a degradation of collision avoidance efficacy, which causes an unavoidable residual rate of risk accrual. No explicit requirements exist for covariance sizes and realism; instead there is a three-way tradeoff between prediction accuracy, quantity of collision avoidance maneuvers performed, and operator risk posture. Improving predictive information of objects in orbit lessens the burden on the latter two aspects.

For the most part, satellite operators choose to maneuver for conjunctions within 72 hours ahead of the actual close approach. As such, one performance metric for an SSA network is the size of the covariances (say, the 90th percentile size for conservatism) for a population of objects at 72 hours into the future. However, propagation of covariances into the future can be somewhat subjective, as the rate of growth uncertainty depends on the dynamical noise (“process noise”) assumed by the propagator. This requires knowledge of the degree of dynamical mismodeling within the propagator, either due to unincorporated or unpredictable dynamical inputs (space weather, satellite maneuvers, rotations or deployments affecting solar radiation pressure and drag, outgassing, etc.). From a simulation perspective, this poses a challenge as one knows precisely the dynamical inputs used within the simulation.

While one can attempt to estimate these factors or assume a blanket acceleration noise, a simpler approach is to define a distinct but related performance metric which is independent of process noise assumptions. Instead of covariance propagation 72 hours into the future, the performance metric can be defined as the covariance size at any given point in time. This alternate definition essentially equates to the typical knowledge you have about an object’s state at any point in the present time. This alternative also permits improved statistical sampling without requiring additional computation, as one can sample the covariance size throughout all time steps within the simulation rather than exclusively propagating from certain points onward.

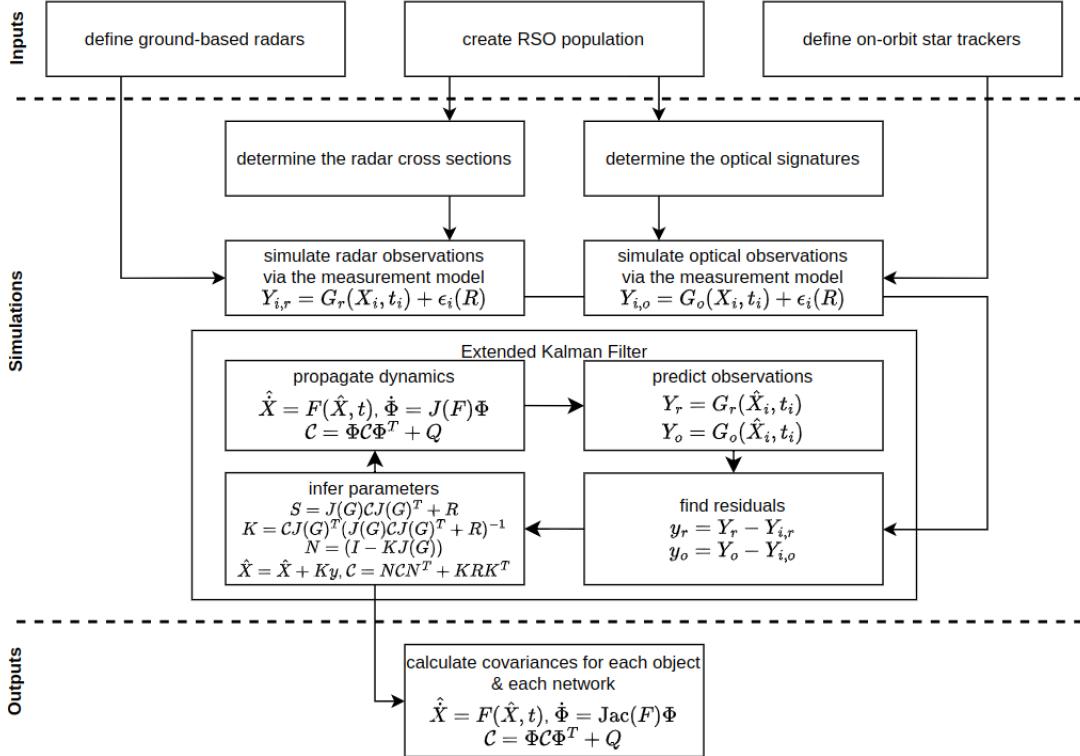


Figure 4-2: A block diagram of the simulation approach.

4.2 Simulation methods

The simulation performed in this chapter largely mirrors that of Chapter 3: we define the instruments, generate a hypothetical population of objects which may be observed, determine the observable properties of those objects, and simulate network performance. In this case, the statistical inference performed within the simulation is not a static measure of object properties (inverting from measurement space into object parameter space), but rather a dynamical state estimation (inverting from measurement space into state space at various times). Fig. 4-2 outlines the approach, which shows this recursive state estimation process. Another way this simulation differs from that in Chapter 3 is that network performance is not arbitrarily computable from the performance of individual instruments, but requires individual simulation. This difference is due to the state estimation process and is unavoidable regardless of the estimation process that is used (in this case, an extended Kalman filter).

First, the simulation inputs are defined, including a set of radars, star trackers, and

resident space objects (along with their corresponding properties such as radar cross-sections and optical signatures). Ahead of the state estimation filter, simulated data is generated with noise by running the measurement model against the ground truth object states (determined with a custom dynamical propagator). The measurement model is run at 60 s intervals and data is generated only where observing constraints such as line of sight are met. Then the state estimation filter is run with a perturbed initial state along with the noisy data, which returns a refined estimate of the satellite state vectors and covariances over time. The simulation output is a performance metric based on the covariance time series data.

Note a caveat of the simulation is that it does not take into account object association. That is: when one takes a measurement of an RSO, it is not always the case that one can associate that measurement with a known object in the database. If the object is not within the database, one may not be able to associate the measurement with other similarly unassociated measurements of that object. This confusion is particularly strong in the case of poorly-tracked objects, tightly-clustered objects, and either recently-deployed or not-recently-observed objects. Since one can generally associate objects given sufficiently precise data over a long enough period of time, this simulation assumes a steady-state scenario where all objects are already associated and measurements can be accurately attributed to their causal objects.

4.2.1 System dynamics

The dynamical model of a system relates its state derivative $\dot{X} = [\dot{x}, \dot{y}, \dot{z}, \ddot{x}, \ddot{y}, \ddot{z}]$ with its state $X = [x, y, z, \dot{x}, \dot{y}, \dot{z}]$ at time t :

$$\dot{X} = F(t, X). \quad (4.3)$$

The simplest dynamical model for orbital mechanics assumes only two-body gravity. In this work, a custom higher-fidelity model is used which incorporates several perturbative forces:

- Earth gravity field given by the 1996 Earth gravitational model (EGM96) [98],

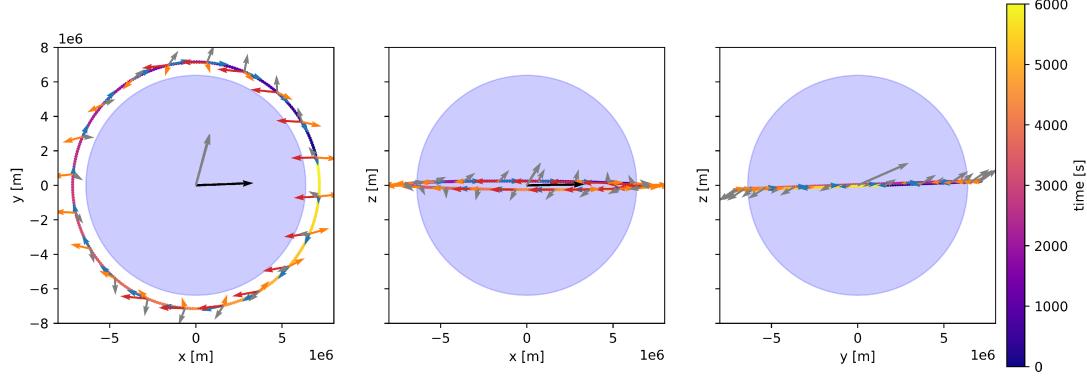


Figure 4-3: An example orbit propagation.

Orbit-centric arrows indicate perturbations including [solar pressure](#), [solar tidal gravity](#), [lunar tidal gravity](#), and [drag](#). The geocentric arrows show sun and moon directions. The orbit is shown in the ECI frame.

which includes significant terms such as two-body gravity, latitude-dependent zonal perturbations such as J_2 , J_3 , etc., and longitude-dependent tesseral/sectoral perturbations up to degree 360×360

- Lunisolar gravitational (tidal) perturbations
- Solar radiation pressure
- Atmospheric drag

For derivations and further detail regarding these perturbative forces, see texts [154] and [133]. For orbital propagation, this work relies on an adaptive step size 4th-order Runge-Kutta integrator (RK4) implemented in `scipy.integrate.solve_ivp` [159] which is known to be stable in orbital dynamics propagation.

We additionally require modeling of process noise, which characterizes uncertainty in the system dynamics such as unmodeled perturbations or unpredictable dynamical inputs such as satellite maneuvers and space weather effects. In the state estimator, process noise is implemented as a 6×6 covariance matrix Q representing dynamical uncertainty along each state vector dimension. Q is generally defined through physical insights of the unmodeled dynamical inputs, tuning to adjust the size of the final state covariances, or adaptive techniques that adjust Q within the state estimation filter. Here, we define Q based on an assumed acceleration noise σ_X , σ_Y , and σ_Z along the

three spatial dimensions:

$$Q(\Delta t) = (\Delta t)^2 \begin{pmatrix} \frac{(\Delta t)^4}{4} \sigma_X^2 & 0 & 0 & \frac{(\Delta t)^2}{2} \sigma_X^2 & 0 & 0 \\ 0 & \frac{(\Delta t)^4}{4} \sigma_Y^2 & 0 & 0 & \frac{(\Delta t)^2}{2} \sigma_Y^2 & 0 \\ 0 & 0 & \frac{(\Delta t)^4}{4} \sigma_Z^2 & 0 & 0 & \frac{(\Delta t)^2}{2} \sigma_Z^2 \\ \frac{(\Delta t)^2}{2} \sigma_X^2 & 0 & 0 & \sigma_X^2 & 0 & 0 \\ 0 & \frac{(\Delta t)^2}{2} \sigma_Y^2 & 0 & 0 & \sigma_Y^2 & 0 \\ 0 & 0 & \frac{(\Delta t)^2}{2} \sigma_Z^2 & 0 & 0 & \sigma_Z^2 \end{pmatrix} \quad (4.4)$$

which assumes no cross-correlation terms between unique spatial dimensions. Note that the 3×3 upper-left portion relates spatial uncertainty to acceleration noise, hence the additional $(\Delta t)^n$ terms due to integration of acceleration.

4.2.2 Radar measurement model

Two radar simulation steps are required in this work: a measurement model to generate simulated radar data given object position and radar noise, and the inverse of the above process for the state estimation filter to map observations into object state space. A measurement model here differs slightly from the definition given in Eqn. 2.2. That measurement model maps an observed *signal* — the actual digital signal output by the radar receiver — to the object observables of range (r) and range-rate (\dot{r}).³ The measurement model discussed here further maps between range and range-rate and the object state space X . Here, rather than simulating noise in the radar signal directly, noise is modeled as a covariance matrix in range and range-rate as per Eqn. 4.11 and 4.12. This simplifies part of the probability graph model from Fig. 2-9, the relevant portion of which is shown in Fig. 4-4. This distinction touches upon a central point: measurement models are chained together to map signal \leftrightarrow measurement \leftrightarrow observables \leftrightarrow more abstracted model properties. It is this chain of functions which informs the efficacy of a given instrument or network of instruments.

³Range-rate is the line-of-sight component of the object velocity relative to the radar.

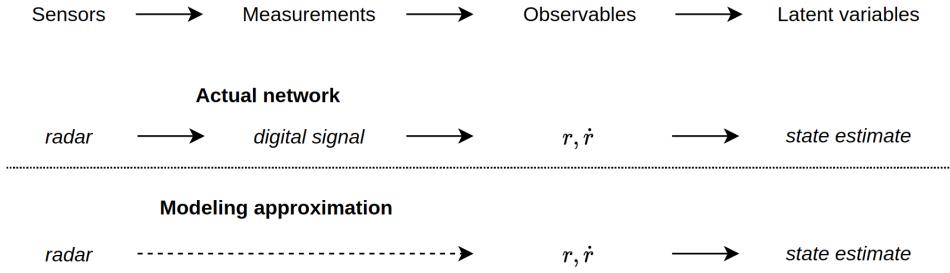


Figure 4-4: The part of the Bayesian network (Fig. 2-9) showing just the radar measurement model.

Rather than having measurement models chaining the measured radar signal to the object observables to the object state estimate, the network is approximated as having a single measurement model converting between observables and the state. Uncertainty is calculated directly in terms of the observables using radar performance (Eqn. 4.8) rather than from simulated noise in the measured signal.

The radar measurement model G_r is given as:

$$[r, \dot{r}] = G_r(\vec{r}_{\text{object}}, \vec{v}_{\text{object}}, \vec{r}_{\text{radar}}, \vec{v}_{\text{radar}}) \quad (4.5)$$

$$r = |\vec{r}_{\text{object}} - \vec{r}_{\text{radar}}| \quad (4.6)$$

$$\dot{r} = \frac{\vec{r}_{\text{object}} - \vec{r}_{\text{radar}}}{r} \cdot (\vec{v}_{\text{object}} - \vec{v}_{\text{radar}}) \quad (4.7)$$

Since this function is linearized in the state estimation process, inverting it analytically is not required in order to map observation residuals back into object state space.

The computation of G_r performed in this work is somewhat more complicated than given above, as radar positions are defined in the ECEF (Earth-centered Earth-fixed) frame, but the object states are propagated in the ECI (Earth-centered inertial) frame. Therefore G_r also performs a coordinate transformation. The ECEF↔ECI transformation used throughout this chapter is based on the IAU-76/FK5 reduction of position, in particular relying on the IAU-1976 precession model and IAU-1980 theory of nutation [154]. Furthermore, a light time correction is employed which uses the radar state at the given measurement time but the satellite state at the light-corrected time, which takes into account the light travel delay $\Delta t = r/c$ between the

satellite and radar.

G_r can be used directly to simulate radar data ahead of the state estimation process, along with the addition of noise σ . To determine radar noise, we use the radar power equation:

$$\frac{1}{\sigma^2} \propto P_r = \left(\frac{P_t G_t A_r}{(4\pi)^2} \right) \frac{d}{r^4}, \quad (4.8)$$

where P_r is received power, P_t is transmitted power, G_t is the gain of the transmitting antenna, A_r is the effective area of the receiver, and d is the object's radar cross-section (RCS). Rather than completely simulating the processing to transform the observed radar signal and noise into a range and range-rate, we use an assumed baseline radar performance, i.e. a baseline range and range-rate noise σ_r and $\sigma_{\dot{r}}$, and vary σ based only on d and r . Then data is generated via:

$$Y_r = [r, \dot{r}] = G_r(\cdot) + \epsilon_r \quad (4.9)$$

$$\epsilon_r \sim \left[\mathcal{N}\left(0, \frac{\sigma_r^2 r^4}{d}\right), \mathcal{N}\left(0, \frac{\sigma_{\dot{r}}^2 r^4}{d}\right) \right] \quad (4.10)$$

where noise is Gaussian-distributed with no bias (i.e., a mean of 0). Note that range and range-rate noise are assumed to be independent. Written more succinctly, the measurement noise $\epsilon_r = [\epsilon_{\text{range}}, \epsilon_{\dot{r}}]$ is drawn from a bivariate Gaussian:

$$\epsilon_r \sim \mathcal{N}(\vec{0}, \mathcal{C}_r), \quad (4.11)$$

where the 2D noise covariance matrix \mathcal{C}_r is defined as:

$$\mathcal{C}_r = \begin{pmatrix} \frac{\sigma_r^2 r^4}{d} & 0 \\ 0 & \frac{\sigma_{\dot{r}}^2 r^4}{d} \end{pmatrix}. \quad (4.12)$$

The object RCS d is defined in the process of generating the simulated objects. Here, a spherical object is assumed where d is equivalent to the object's cross-sectional area. This assumption holds true where $\lambda \ll \text{radius}$, i.e. where the radar frequency is sufficiently high.

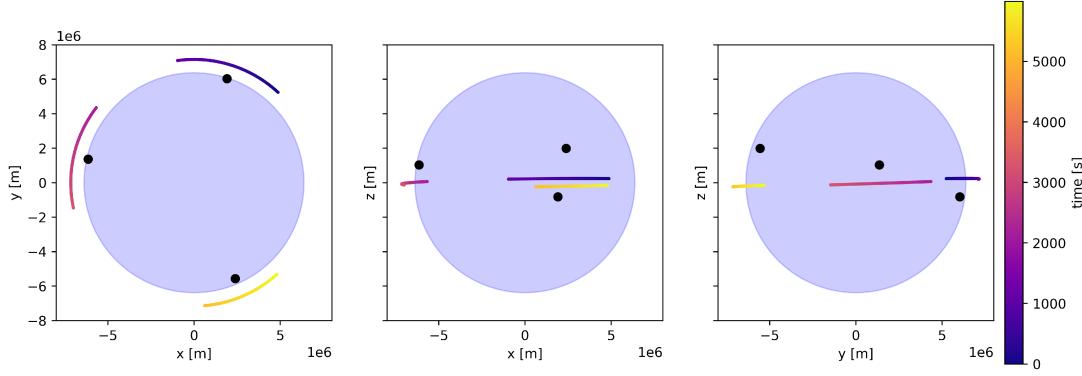


Figure 4-5: Example orbit showing radar ground station passes.
Only portions of the orbit where observability criteria are met are plotted. This plot uses the ECEF frame where ground stations (shown in black) are fixed.

There are three observability criteria applied to the generation of radar data:

- The observed object is above the horizon
- The sun separation angle is greater than 30°
- The lunar separation angle is greater than 30°

If all conditions are met at a given time step, simulated data is output for that time step; otherwise no measurement is performed. Fig. 4-5 shows an example orbit with these conditions applied.

4.2.3 Optical measurement model

The same process above applies to the optical measurement model: we use the model to both simulate noisy measurements as well as to estimate object states within the state estimation filter. Unlike radar, optical sensors measure the right ascension α and declination δ to a target, not the range and range-rate. The star tracker portion of the Bayesian network is shown in Fig. 4-6.

Star trackers generally use on-board star centroiding algorithms such as SExtractor (Source Extractor) [23] to convert a star field into pixel coordinates. Further software such as SCAMP (Software for Calibrating AstroMetry and Photometry) [22]

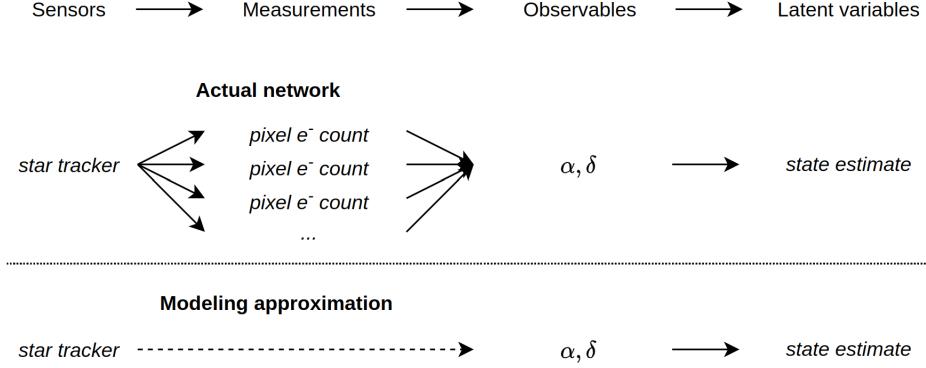


Figure 4-6: The part of the Bayesian network (Fig. 2-9) showing just the optical measurement model.

As with the radar model, the optical measurement model maps between object observables and the object state. Uncertainty is modeled not through simulated noise in pixel-level measurements, but rather as an uncertainty in observables based on star tracker performance (Eqn. 4.17 and 4.18).

matches the observed star field to a stellar catalogue in order to determine the orientation of the image FOV in terms of its right ascension, declination, and rotation. In addition to this typical workflow, astrometry of orbital objects requires further processing steps and is generally more involved than astrometry for static star fields. Rather than a coherent point source which can be centroided per the optics' point spread function (PSF), orbital objects tend to have a high angular rate within the star field and therefore are streaked across the FOV. In addition to reducing the observed SNR, this smearing poses the additional challenge of obscuring the object's right ascension and declination.

Two common procedures for high-angular-rate astrometry include streak detection and the “shift-and-add” technique [139]. Streak detection is akin to traditional centroiding but which fits both a streak length and width to the image rather than just a PSF. While simplest in approach, streak detection does not alleviate poor object SNR. Conversely, the shift-and-add technique does recover higher SNRs by *shifting* successive image frames based on an assumed object velocity vector, and *adding* the signals along the velocity vectors between each frame. One can use a priori knowledge or optimization techniques to determine the velocity vector, or can more simply brute force all plausible vectors in order to recover the source object. Note both the

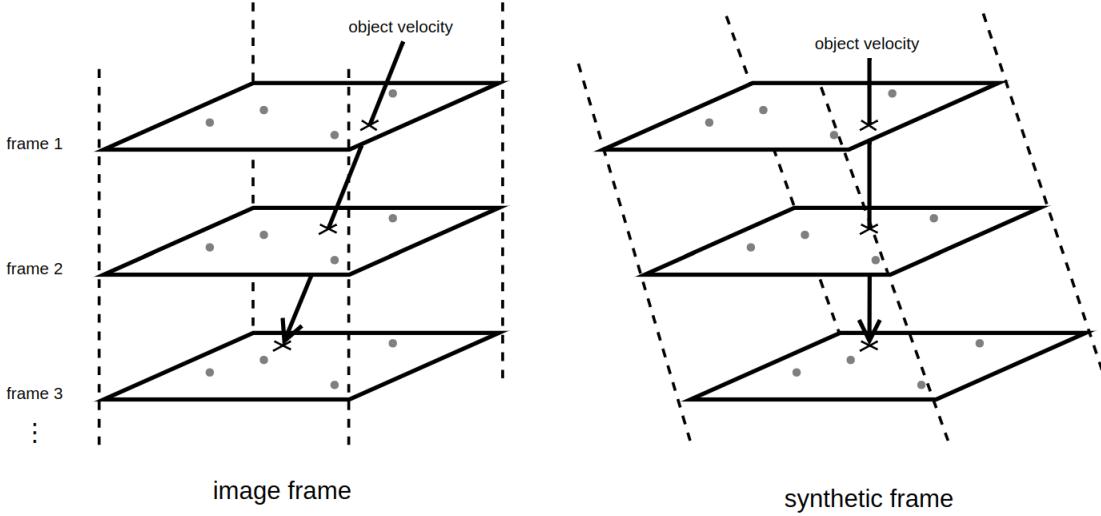


Figure 4-7: Shift-and-add centroiding technique.

This technique permits accurate astrometry of high-angular-rate orbital objects. Star field astrometry occurs in the image frame with classical techniques; object astrometry occurs in the synthetic frame, where frames are moved so as to co-add the object signal over time. The two astrometric solutions are compared to determine the object position and image-plane velocity.

streak and shift-and-add techniques can, in theory, estimate the image-plane portion of object velocity. For simplicity, this work assumes only a measurement of α and δ rather than the full measurement vector $[\alpha, \delta, \dot{\alpha}, \dot{\delta}]$.⁴

The optical measurement model $[\alpha, \delta] = G_o(\vec{r}_{\text{object}}, \vec{v}_{\text{object}}, \vec{r}_{\text{sat}}, \vec{v}_{\text{sat}})$ is calculated as follows:

$$\vec{u} = \frac{\vec{r}_{\text{object}} - \vec{r}_{\text{sat}}}{|\vec{r}_{\text{object}} - \vec{r}_{\text{sat}}|} \quad (4.13)$$

$$\alpha = \arctan \vec{u}_y / \vec{u}_x \quad (4.14)$$

$$\delta = \arccos \vec{u}_z \quad (4.15)$$

As with the radar measurement model, G_o is used to generate simulated optical data ahead of the state estimation process. Unlike range and range-rate noise, α and δ noise are correlated due to the image-plane velocity $[\dot{\alpha}, \dot{\delta}]$. Therefore the measurements are

⁴Uncertainties in image-plane velocity are likely to be much higher than astrometric uncertainties, so the full measurement vector may not significantly reduce the object state covariance regardless.

generated with noise $\epsilon_o = [\epsilon_\alpha, \epsilon_\delta]$ per:

$$Y_o = [\alpha, \delta] = G_o(\cdot) + \epsilon_o \quad (4.16)$$

$$\epsilon_o \sim \mathcal{N}(\vec{0}, \mathcal{C}_o) \quad (4.17)$$

$$\mathcal{C}_o = \begin{pmatrix} \sigma_{\alpha\alpha} & \sigma_{\alpha\delta} \\ \sigma_{\delta\alpha} & \sigma_{\delta\delta} \end{pmatrix} = R \begin{pmatrix} \frac{1}{\rho}(\theta^2 + t_{\text{exp}}\omega) & 0 \\ 0 & \frac{\theta^2}{\rho} \end{pmatrix}, \quad (4.18)$$

where R represents a rotation matrix from the synthetic frame to the image frame, θ is the Rayleigh criterion $\theta = 1.22\frac{\lambda}{D}$, t_{exp} is the exposure duration, ω is the angular rate of the object as measured by the star tracker, and ρ is the SNR.

The rotation $R \in SO(2)$ transforms the uncertainty in the synthetic frame to uncertainty in the global coordinates α and δ as shown in Fig. 4-8. The axes of the synthetic frame in which \mathcal{C} is defined are given by the object's image-plane velocity vector and a secondary orthogonal axis. In this frame, there is a variance in the direction of motion given by $\frac{1}{\rho}(\theta^2 + t_{\text{exp}}\omega)$: the θ^2 term corresponds to the optical PSF with FWHM (full-width half-maximum) angle θ , given by the Rayleigh criterion, and the $t_{\text{exp}}\omega$ term corresponds to uncertainty due to object smearing. In the orthogonal direction, the variance is only due to the optical PSF, as there is no relative motion along this axis. In the synthetic frame, noises are independently distributed, so there is no cross-correlation (off-diagonal covariance terms). After the rotation R there is a strong correlation in the α, δ frame, as in general the object motion is the dominant uncertainty contribution rather than the Rayleigh limit (i.e. $t_{\text{exp}}\omega \gg \theta^2$).

The SNR ρ is computed similarly to Eqn. 3.30:

$$\rho \equiv \frac{S}{N} = \frac{n_{\text{signal}}}{\sqrt{n_{\text{signal}} + \text{pix}(n_{\text{dark}} + n_{\text{read}}^2)}}, \quad (4.19)$$

where in this case the number of pixels over which the source is distributed is given by

$$\text{pix} = \left\lceil \frac{1.22\frac{\lambda}{D}t_{\text{exp}}\omega}{(p/f)^2} \right\rceil \quad (4.20)$$

where the denominator is the instantaneous (pixel) FOV angular area calculated with

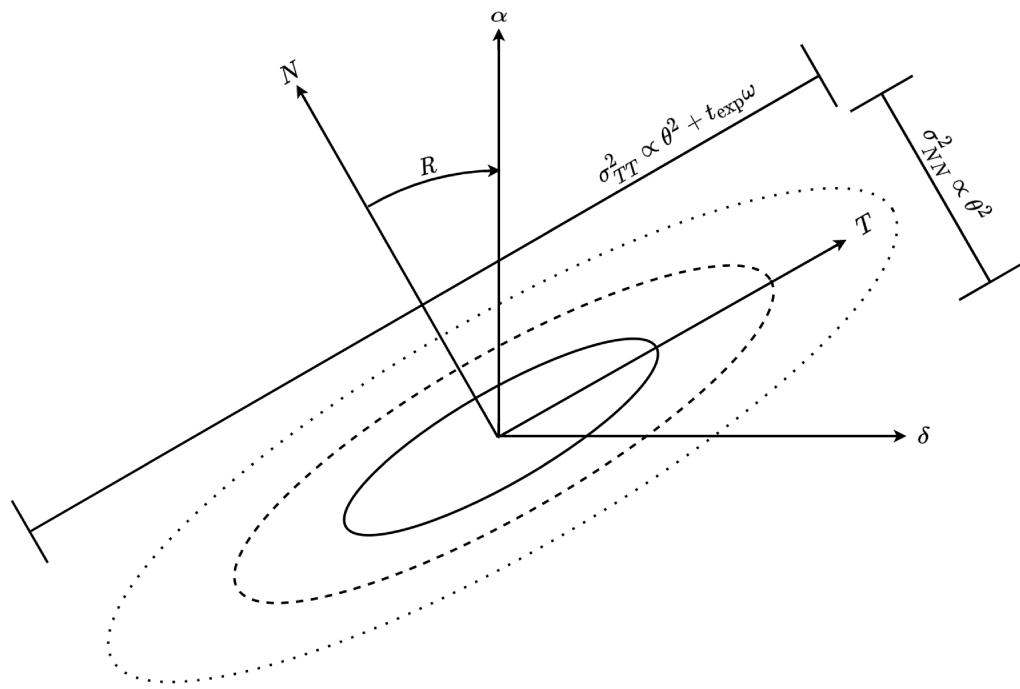


Figure 4-8: Optical astrometric covariance.

The astrometric covariance is defined as a diagonal matrix in the synthetic frame and is transformed to the global α, δ frame with rotation matrix R . The angular uncertainty is greatest in the direction of motion, axis T , though also has a component in the orthogonal direction due to the optical PSF.

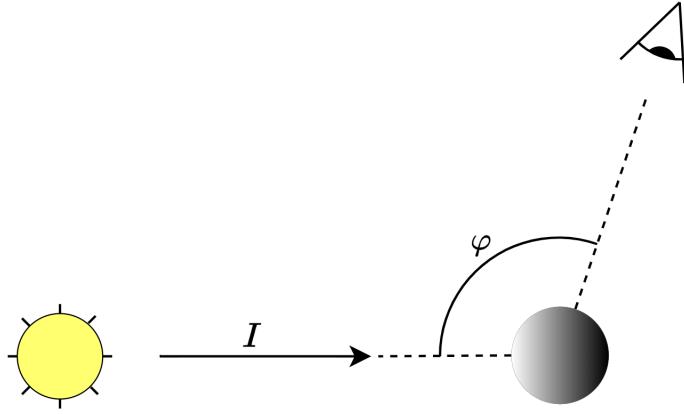


Figure 4-9: Solar phase angle φ of an object/observer pair.

the pixel pitch p and focal length f , and the numerator is the angular area the source covers during the exposure. The latter is calculated where the width of the streak is the FWHM of the PSF, and the length of the streak is the angular distance traveled by the source. Furthermore, the signal flux n_{signal} has to be defined.

The signal flux is calculated based on the area of the object and its reflective properties. Specifically, it is assumed the object is spherical with a cross-sectional area A , and is 0% absorptive/emissive, 50% specularly reflective, and 50% diffusely reflective. The equation for the received signal can then be derived as:

$$n_{\text{signal}} = I \frac{D^2}{4\pi} \frac{A}{2\pi r^2} \cos \varphi \quad (4.21)$$

where D is the optical aperture diameter, I is the intensity of the incident light from the Sun, and φ is the phase angle of the object as shown in Fig. 4-9.

Furthermore, observability criteria are applied to the optical measurement model which dictates whether or not the object can be observed at each point in time. Specifically, the observability criteria include that:

- the object must be within the star tracker FOV during the entirety of the exposure,
- the SNR threshold must be satisfied ($\rho \geq 10$),

- there must be a minimum sun separation angle of 50° , i.e. $\varphi < 130^\circ$,
- there must be a minimum lunar separation angle of 30° ,
- the line of sight axis must not intersect Earth,
- and the object must not be in Earth eclipse (within its umbra or penumbra).

Note that the measurement model simultaneously depends on the object state as well as the star tracker state. The star tracker host satellite state is similarly propagated at the beginning of the simulation in order to simulate the optical measurement data. This state propagation is reused within the state estimation filter to assess prediction residuals.

The first observability criterion regarding star tracker FOV addresses the fact that star tracker observations are coincidental rather than targeted. The star tracker orientations do not change in order to track targets, but are instead static. The FOVs are defined in the satellite Frenet frame⁵ with the tangential (velocity/in-track) axis T , normal (radial) axis N , and binormal (cross-track) axis B shown in Fig. 4-10. This definition assumes the satellite body frame always coincides with the Frenet frame, and that there are four star trackers colocated on each satellite, each coplanar but separated by 90° . The star trackers are individually granted $10^\circ \times 10^\circ$ FOVs.

4.2.4 State estimation

For the dynamical state estimation representing the core of the simulation, this work relies on an extended Kalman filter (EKF), a nonlinear variant of the Kalman filter which linearizes about the current mean (the estimated object state) and covariance [127]. The EKF is essentially a Bayesian inference engine applied to state estimation of nonlinear dynamical systems under the assumption of Gaussian noise. The EKF follows two main steps during the recursive estimation process. In the prediction step, the EKF predicts the next state of the system given its previous state and

⁵An equivalent and perhaps more familiar convention is the LVLH (local vertical, local horizontal) frame.

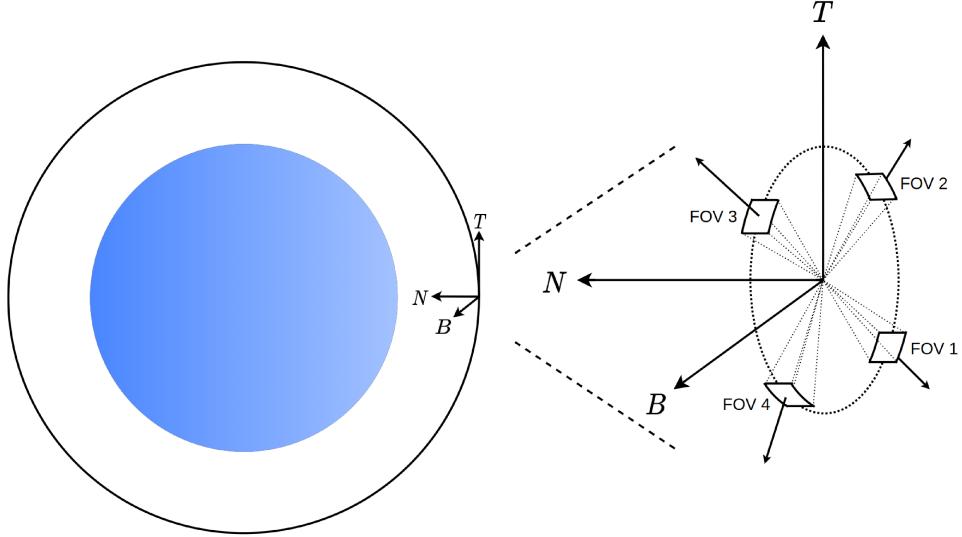


Figure 4-10: Star tracker FOVs.

Each satellite hosts four star trackers in the T - B (in-track/cross-track) plane which are individually separated by 90° .

a dynamical model, and computes predicted measurements based on the propagated state. During the update step, the EKF determines measurement residuals based on the observed versus predicted data, and infers a new state given measurement residuals and associated uncertainty. Essentially, the EKF inverts observations from measurement space into state space at each observation time in order to recursively estimate the state of the system.

The EKF differs from other nonlinear filters such as the unscented Kalman filter (UKF) and Monte Carlo filters (e.g., particle filters) in that it requires definition of the Jacobians of the dynamical function and related measurement functions in order to linearize the state estimation process. For systems with complicated dynamics and measurement models, deriving and implementing these Jacobians can represent a significant up-front cost. The upsides of the EKF rest in its computational efficiency and rigid definition: unlike the UKF or Monte Carlo filters, it requires no parameter tuning for the specific state estimation problem, and only requires propagation of one state forward in time rather than multiple test points or “particles”.

Convergence in this context refers to the ability of the state estimator to “find” the true mean of the system and accurately track it over time. In general, observational

data may be roughly compatible with a wide breadth of possible system states, so convergence onto the true state is not guaranteed. Convergence is aided by accurate priors, i.e. good initial state estimates, and precise modeling of system dynamics and measurement models. While the EKF may fail to converge in the case of highly nonlinear dynamics or poor initial state estimates, these failure modes are easily avoided in this work as initial state estimates can be arbitrarily chosen within the simulation.

The recursive EKF process is given explicitly in Algorithm 1. Understanding the internals of this algorithm requires precursory definitions and clarification of notation. To summarize, so far we have defined:

- the system state X ,
- the covariance \mathcal{C} of knowledge of the system state,
- the system dynamics $\dot{X} = F(t, X)$,
- the process (dynamical) noise $Q(\Delta t)$,
- the radar measurement model $Y_r = G_r(\cdot)$ and noise model $\epsilon_r \sim \mathcal{N}(\vec{0}, \mathcal{C}_r)$,
- and the optical measurement model $Y_o = G_o(\cdot)$ and noise model $\epsilon_o \sim \mathcal{N}(\vec{0}, \mathcal{C}_o)$.

We require a few further definitions. In particular, the EKF is a nonlinear state estimation filter which linearizes about the current state. As such, we must define the Jacobian of several functions:

- $A(X) \equiv \frac{\partial F}{\partial X}|_X$
- $H_r(X) \equiv \frac{\partial G_r}{\partial X}|_X$
- $H_o(X) \equiv \frac{\partial G_o}{\partial X}|_X$

H_r and H_o represent the sensitivity of estimates of each state dimension to each measurement dimension, where near-zero elements of H imply an inability to precisely constrain the corresponding state dimensions. These Jacobians are generally quite

complicated to derive, so are implemented symbolically in SymPy [106] using the `lambdify` method for computationally-efficient function evaluations.

The final definition required for understanding the EKF is the 6×6 state transition matrix $\Phi_{(t_f, t_i)}$. Φ maps the system state at time t_i to its state at time t_f , and is propagated similarly to state X via the relation $\dot{\Phi} = A(X)\Phi$.

4.3 Results

The analysis focuses on two example RSOs to demonstrate network performance in a couple of realistic object tracking scenarios. Case 1 shows a debris tracking scenario involving a small (10 cm) piece of debris in LEO. Case 2 shows tracking of a newly-deployed sun-synchronous orbit (SSO) smallsat. In the case of newly-deployed spacecraft, fast state convergence is often required to set up an initial communication link. In each case, the considered network configurations include:

- a ground-only radar network consisting of 0, 3, or 10 equally-sensitive radar sites, spaced roughly equally in longitude and distributed from -30° to $+30^\circ$ in latitude
- a space-only star tracker network consisting of 0, 100, or 1000 hypothetical satellites, all in a Walker-Delta configuration in 45° orbits at 500 km altitude
- hybrid networks consisting of combinations of the above radar and star tracker configurations

Scenario 1 is shown in Fig. 4-11 and scenario 2 is shown in Fig. 4-12. The plots indicate the size of the position covariance over a 24-hour tracking period for various network configurations. Interpreting the heatmaps on the right of Fig. 4-11 and 4-12 indicate the trade between building out additional radars versus incorporating star tracker imagery into existing space situational awareness pipelines. In either scenario, having zero radars but some star trackers grants worse predictions overall than having some radars and zero star trackers. In the first scenario, 3 radars and

Algorithm 1 Extended Kalman Filter (EKF)

```
1: procedure EKF( $X_0, t_0, \mathcal{C}_0$ , observations)
2:   // Initialize variables
3:    $\Phi \leftarrow I$ 
4:    $X_{i-} \leftarrow X_0$ 
5:    $t_{i-} \leftarrow t_0$ 
6:    $\mathcal{C}_{i-} \leftarrow \mathcal{C}_0$ 
7:   // Initialize output arrays
8:    $X \leftarrow \text{zeros}()$ 
9:    $\mathcal{C} \leftarrow \text{zeros}()$ 
10:  for  $Y_i, \mathcal{C}_{\text{data}}, t_i$  in observations do
11:    // Prediction step
12:    // Propagate state to time of observation
13:     $X_i, \Phi \leftarrow \text{propagate}(X_{i-}, t_{i-}, t_i)$ 
14:    // Update covariance based on propagation
15:     $\mathcal{C}_i \leftarrow \Phi \mathcal{C}_{i-} \Phi^T + Q(t_i - t_{i-})$ 
16:    // Determine measurement residuals
17:     $y \leftarrow Y_i - G(X_i)$ 
18:
19:    // Update step
20:    // Perform intermediate computations
21:     $S \leftarrow H(X_i) \mathcal{C}_i H(X_i)^T + \mathcal{C}_{\text{data}}$ 
22:     $K \leftarrow \mathcal{C}_i H(X_i)^T S^{-1}$ 
23:    // Update the state and covariance
24:     $X_i \leftarrow X_i + Ky$ 
25:     $C_i \leftarrow (I - KH(X)) \mathcal{C}_i (I - KH(X))^T + K \mathcal{C}_{\text{data}} K^T$ 
26:    // Store outputs at this time step
27:     $X[i] \leftarrow X_i$ 
28:     $\mathcal{C}[i] \leftarrow \mathcal{C}_i$ 
29:
30:    // Reinitialize variables for next observation
31:     $\Phi \leftarrow I$ 
32:     $X_{i-} \leftarrow X_i$ 
33:     $t_{i-} \leftarrow t_i$ 
34:     $\mathcal{C}_{i-} \leftarrow \mathcal{C}_i$ 
35:  end for
36:  return  $X, \mathcal{C}$ 
37: end procedure
```

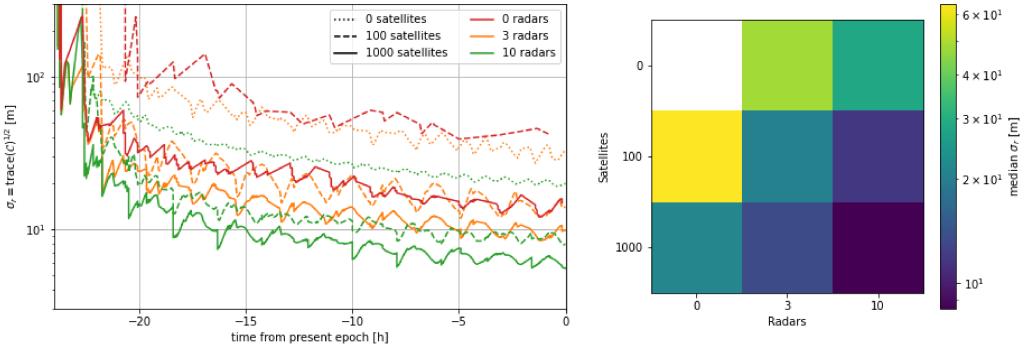


Figure 4-11: Joint tracking of a small debris fragment in equatorial LEO.

1000 satellites is roughly as good as 10 radars and only 100 satellites: it takes many star trackers have the equivalent information contribution of a single radar. The same applies to the second case, albeit somewhat more balanced between the two types of instruments, likely because the object is in a dawn/dusk SSO. In this orbit, the object is permanently lit and so is always observable by a star tracker given an appropriate viewing angle.

We find that the marginal contribution from adding an additional radar to the instrument network reduces covariance significantly more than adding additional star trackers, even in an ideal permantly-lit sun-synchronous case. However, the only requirement to use extant star tracker data is additional processing of pre-existing image frames. Assuming the issues of image processing and downlink can be overcome, along with the object association issue not considered in this work, incorporating star tracker measurements could be a viable approach to supplement (but not replace) ground-based radar for the purpose of space situational awareness.

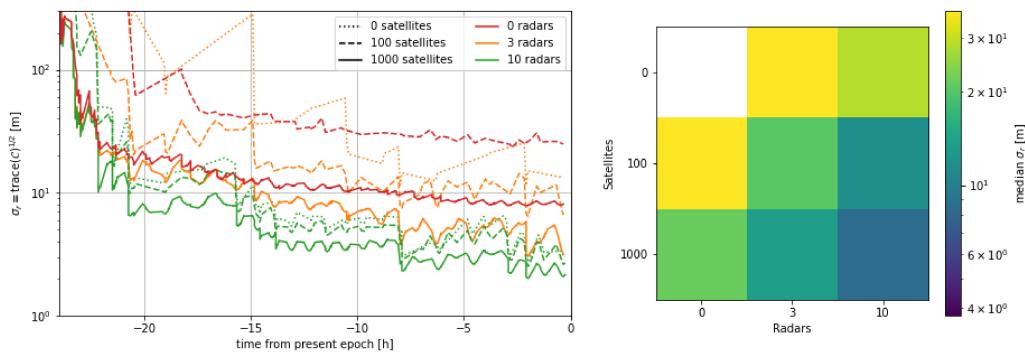


Figure 4-12: Joint tracking of a smallsat in dawn/dusk sun-synchronous orbit.

Chapter 5

Conclusion

This thesis presents a general framework for conceptualizing multi-messenger astronomy. The framework relies on probabilistic graph models¹ to trace the flow of information — here defined as Fisher information — from *measurements* of astrophysical sources to *latent variables* describing the phenomena which are measured to *scientific hypotheses* regarding the processes that define them. The fusion-based inference performed during this process is a prerequisite feature of multi-messenger astronomy, unlike traditional astronomy for which information at the measurement level can serve as a direct proxy to represent information at the hypothesis level. As a result, designing and deploying multi-messenger networks necessarily requires assumed models of the phenomena at hand, even if that understanding is rudimentary or incomplete.

This work demonstrates that adjacent fields of information theory and heterogeneous sensor fusion have deep relation and are readily applicable to modern problems faced by multi-messenger astronomy. Central challenges include how to perform sensor fusion, whether sensor fusion is useful for any given scenario, and how to design a network of instruments that best takes advantage of cross-domain informational content. All of these aspects are addressed in this work.

Chapter 2 discusses the fundamental limits of astrophysical information at the measurement level. These limitations occur due to a finite quantity of incident information, imperfect (lossy or noisy) measurement models of the instruments which

¹Also known more commonly (but more narrowly in definition) as belief or Bayesian networks.

access that information, and fundamental quantum uncertainty. Imperfections propagate through belief networks, thereby determining the quality of inferences that can be made with imperfect or incomplete measurements, as is always the case with real-world data. To best understand this effect in a multi-messenger context, Chapters 3 and 4 apply the framework to two case studies regarding cosmology with binary neutron star mergers and space situational awareness of resident space objects respectively. Both case studies demonstrate different aspects of multi-messenger astronomy: Chapter 3 shows that novel hypotheses can emerge from fusing multimodal information, whereas Chapter 4 shows that informational content scales with heterogeneous data sets, even where the same kinds of hypotheses are accessible by traditional measurements alone.

5.1 Future work

The multi-messenger framework presented in this work can be extended on numerous fronts.

5.1.1 Improvements to case study modeling fidelity

Most straightforwardly, the fidelity of each case study can be improved. In Chapter 3, this could include using a Bayesian treatment of gravitational-wave parameter inference (such as `BAYESTAR` or `LALInference`) rather than using the Fisher information method. In addition, instead of relying on fraction of events observed, network performance could instead be characterized more informatively as the horizon distance of the network. Furthermore, various simplifying assumptions could be revisited, including (but not limited to!):

- homogeneity of mergers, such as the assumptions of uniformity of rate density with redshift and lack of electromagnetic redshift in the EM measurement model,
- fixed image exposure duration, which could instead be adaptable based on the

predicted luminosity of the kilonovae,

- 100% duty cycle for every instrument, which is particularly unlikely for survey instruments or space-based instruments with predefined observing schedules,
- and a lower frequency cutoff for gravitational-wave chirps, which may have rendered space-based interferometry to appear less appropriate for this application than it otherwise could be.

In Chapter 4, similar fidelity limitations could be addressed. These include the assumption of a “steady-state” catalogue where all objects are already correlated, and where the task of object association introduces no additional uncertainty. In reality, object association is likely one of the central challenges that must be overcome in order to effectively repurpose star tracker imagery. Furthermore, it would be better to incorporate an improved radar measurement model which takes into account the actual radar properties and observed signal rather than an assumed baseline SNR that simply scales with range and radar cross-section.

Lastly, it would be more appropriate to address a broader and more realistic set of resident space objects rather than just the two hypothetical objects considered so far. While the same simulation method and probabilistic graph model would apply, simulation inputs would require additional development, and the simulation itself would need to be more computationally efficient (or simply run with more compute). To develop inputs, one could pull TLE data directly from the NORAD catalogue of presently-known resident space objects. While the optical properties and radar cross-section of the objects may not be exactly known (and are not included in TLE sets), sizes could be approximated using e.g. the ballistic coefficients. To make the simulation more efficient, a method similar to the one applied in Chapter 3 could be used to analytically recompute conditional probabilities for each network just using simulated performance of individual instruments. Absent further computational improvements, $\mathcal{O}(10^4)$ objects could be simulated within a few weeks on a typical laptop, or more realistically could be run in parallel with cloud computing services.

5.1.2 Further analyses to build upon case study results

Furthermore, in both cases, there is much room for additional analyses even exclusively based on the simulation data set already generated in this work. The simulation data set is vast and highly-dimensional, particular for the binary neutron star merger case study. Questions one could answer with these data include:

- identification of specific multi-messenger network configurations to recommend based on a balance of performance and resources,
- the merit of various engineering trades, such as in the quality versus quantity of detectors and the use of ground-based versus space-based sensors,
- and the design of purpose-built observatories specifically for measuring cosmological standard sirens, such as high-frequency space-based interferometers or ground-based detectors with better geographical distribution.

Some of these questions would require costing models for gravitational-wave interferometers. Present costing models even for well-known instruments like telescopes are an art form at best, so developing similar models for interferometry would likely necessitate further research.

For the resident space object case study, one could not only demonstrate the efficacy of star trackers, but also could optimize the placement of worldwide radar ground station networks. This extension could include arbitrary constraints such as using only geopolitically “friendly” territory and relatively radio-quiet zones.

5.1.3 Additional applications

This thesis considered only two narrow applications of the multi-messenger systems framework. A vast variety of other potential applications exist. To name just a few in different fields, focusing on but not limited to astronomy and space sciences:

- Planetary science poses an intriguing extension, as the messengers can extend to include direct physical contact with the phenomena in question. For example,

the Double Asteroid Redirection Test (DART) mission impacted the asteroid Dimorphos with a satellite probe [47]. DART’s LICIACube cubesat companion observed the collision itself in both the optical and infrared to study asteroid morphology and the mechanisms of impact plume generation. Ground- and space-based telescopes, including Hubble, performed follow-on observations of the ejecta evolution and changes to Dimorphos’ trajectory. Other planetary science applications also use gravitational measurements (via perturbations of spacecraft orbits) to map asteroid interiors [78].

- Heliophysics is a field rich with multimodal sensor fusion. Information sources range from EM imagery of the sun and its corona (including the full EM spectra along with different methods such as spectrometry and magnetography), neutrinos and cosmic rays (e.g. solar wind) also emitted by the sun, direct EM field measurements within the full volume of the heliosphere, to measurements of space weather interactions with various planetary bodies including Earth. In the future, heliophysics may even expand to include all four messengers²: [146] proposes a means to map the sun’s interior by observing background GW sources (pulsars) which are lensed by the sun.
- This framework could potentially apply to non-astronomical applications. A promising exercise is to use the method to determine how low-cost sensors can cue high-cost sensors in terrestrial applications which rely on heterogeneous sensor networks. For example, autonomous navigation systems often require a suite of heterogeneous sensing equipment such as cameras, lidars, and radars, and the field of neuromorphic computing deals with low-cost sensors that mimic biological neural systems. While the framework would require further development to address modeling of non-astronomical sources, in principle the method of probabilistic graph models equally applies.

²Arguably, all four messengers have already been utilized in this application: two of the first confirmations of general relativity came from measurements of the gravitational effect of the sun. These tests included gravitational deflection of light (in a way jointly gravitational/electromagnetic) and the precession of Mercury’s perihelion.

5.1.4 Mathematical development of the systems framework

Perhaps most importantly, the systems framework itself could be improved or built upon in a number of ways. Firstly, the mathematical rigor could be further developed. The informational content at each network node could be more explicitly tracked and shown in each case study. Information content is likely better represented by concepts including mutual information, entropy, or Kullback-Leibler divergence rather than Fisher information considered in this work. Mutual information, for example, quantifies the informational bits gained about one random variable through observing another random variable. Contrast that to Fisher information, which measures the amount of information that an observable random variable carries about an unknown parameter of a given distribution.

In addition, it has been proven that exact inference in Bayesian networks is NP-hard [34], meaning using this framework to optimize multi-messenger instrumentation networks could be computationally intractible in many scenarios. While this work placed restrictions on the conditional probabilities to stave off this effect (Chapter 3) or otherwise just reduced the scale of the problem considered (Chapter 4), much related progress has been performed regarding how to best approximate the inference process [97], which this research could draw from instead.

The Bayesian network itself merely computes the efficacy of a particular set of measurements at performing scientific inference. It does not directly address how to perform the optimization defined by Eqn. 2.17. In this thesis, this issue was essentially treated via a brute force approach by simulating the performance of a broad number of networks in parallel. It did not explicitly perform optimization with feedback based on the results of modeled networks, which is typically a more effective and fruitful approach. Connecting the output of such a Bayesian network with a true optimization method represents a promising branch of future work.

Lastly, a significant practical question remains to be addressed. This framework implies that design of observatory networks necessarily requires quantitative, physically-descriptive models of the phenomena that may be studied. However, even

just the two scientific goals addressed in this thesis posed a significant undertaking in terms of modeling time, effort, and computation. A necessary avenue for future work includes streamlining this process. While there is presumably a fundamental degree of computational irreducibility to modeling astrophysical phenomena, perhaps further abstraction or mathematical development could provide helpful simplification of the method.

Appendix A

Tables of instruments for Chapter 3

Tables A.1 and A.2 show the list of gravitational-wave interferometers and electromagnetic telescopes used in the binary neutron star merger case study. The `and` ID column identifies groupings of detectors; i.e., two detectors with the same ID must be either both included or both excluded from the network. The `or` ID column identifies groupings of detectors which were not simultaneously included in the network; i.e. only one detector (or one `and` grouping of detectors) can be included at a time. Data for these instruments were generally pulled from Wikipedia or other online sources. Some instruments are purely hypothetical in nature, but most are based on extant instruments, planned upgrades, or future proposals.

Table A.1: Gravitational-wave interferometers

name	abbreviation	location(s)	or ID	and ID
Laser Interferometer Space Antenna	LISA	heliocentric orbit		
Cosmic Explorer (1-site only)	CE1	US (CA)	0,1	
Cosmic Explorer (2-site)	CE2	US (CA and MT)	0,1	
Cosmic Explorer (1 site relocated)	CE_South	Argentina	1	1
Cosmic Explorer (1 site relocated)	CE_South	Australia	4	1
Cosmic Explorer (1 site relocated)	CE_South	Antarctica		1
Neutron Star Extreme Matter Observatory	NEMO	Australia	4	
Einstein Telescope	ET	Germany		
LIGO Livingston	LLO	US (LA)	2	2
LIGO Hansford	LHO	US (WA)		2
LIGO Voyager (LLO & LHO upgrade)	VOY	US (LA and WA)	2	
Virgo	VIR	Italy	3	
Virgo nEXT (Virgo upgrade)	nEXT	Italy	3	
Kamioka Gravitational Wave Detector	KAGRA	Japan		

Table A.2: Electromagnetic telescopes

name	abbreviation	location(s)	or ID	and ID	FOV [deg ²]	aperture [m]
One-Meter Two-Hemispheres	1M2H	29°S, 71°W 37°N, 122°W	0 0	0.25 0.25	1 1	
Dark Energy Camera	DECam	30°S, 71°W	0	3	4	
Distance Less Than 40 Survey	DLT40	30°S, 71°W	0	0.03	0.4	
Visible and IR Survey Telescope	VISTA	25°S, 70°W	0	0.6	4.1	
Wide-Field Infrared Transient Explorer	WINTER	33°N, 117°W		1	1	
Zwicky Transient Facility	ZTF	33°N, 117°W	1	47	1.2	
Rubin Observatory	LSST	30°S, 71°W	1	9.6	8.4	
MASTER Global Robotic Net	MASTER	50°N, 127°E 52°N, 103°E 59°N, 57°E 43°N, 44°E 21°N, 32°W 17°S, 28°E 32°S, 69°W	2 2 2 2 2 2 2	8 8 8 8 8 8 8	0.4 0.4 0.4 0.4 0.4 0.4 0.4	
Las Cumbres Observatory	LCO	31°N, 104°W 21°N, 156°W 28°N, 17°W 30°S, 71°W 31°S, 149°E 34°S, 18°E	3 2 2 2 2 2	0.15 0.15 0.15 0.15 0.15 0.15	1 1 1 1 1 1	
Extremely Large Telescope	ELT	25°S, 70°W		0.03	39.4	
Thirty Meter Telescope	TMT	20°N, 155°W	0	0.06	30	
Thirty Meter Telescope (in Arizona)	TMT_Az	35°N, 111°W	0	0.06	30	
Giant Magellan Telescope	GMT	29°S, 70°W	1	0.11	25.4	
Giant Magellan Telescope (in Arizona)	GMT_Az	35°N, 111°W	1	0.11	25.4	
Overwhelmingly Large Telescope	OWL	25°S, 70°W		0.003	100	
Hubble Space Telescope	Hubble	LEO	2	0.002	2.4	
Roman Space Telescope	Roman	Earth-Sun L2	2	0.28	2.4	
Large UV, Optical, and IR Telescope A	LUVOIR-A	Earth-Sun L2	3	0.002	15	
Large UV, Optical, and IR Telescope B	LUVOIR-B	Earth-Sun L2	3	0.002	8.1	

Bibliography

- [1] National space traffic management policy.
- [2] Hst cycle 27 primer pointing, orientation, and roll constraints, 2019.
- [3] Space-domain wide area tracking and characterization (space-watch), Jan 2023.
- [4] J. et al. Aasi. Enhanced sensitivity of the ligo gravitational wave detector by using squeezed states of light. *Nature Photonics*, 7(8):613–619, Aug 2013.
- [5] B. P. et al. Abbott. Observation of gravitational waves from a binary black hole merger. *Phys. Rev. Lett.*, 116:061102, Feb 2016.
- [6] B. P. et al. Abbott. Prospects for observing and localizing gravitational-wave transients with advanced ligo, advanced virgo and kagra. *Living Reviews in Relativity*, 23(1):3, Sep 2020.
- [7] B.P. et al. Abbott. A gravitational-wave standard siren measurement of the hubble constant. *Nature*, 551(7678):85–88, oct 2017.
- [8] B.P. et al. Abbott. GW190425: Observation of a compact binary coalescence with total mass $\sim 3.4 m_{\odot}$. *The Astrophysical Journal Letters*, 892(1):L3, mar 2020.
- [9] Graeme E. Addison. High h_0 values from CMB e-mode data: A clue for resolving the hubble tension? *The Astrophysical Journal Letters*, 912(1):L1, Apr 2021.
- [10] European Space Agency. About space debris.
- [11] National Air and Space Museum. International ultraviolet explorer (iue) satellite notebooks, 2012.
- [12] Javier M. Antelis and Claudia Moreno. Obtaining gravitational waves from inspiral binary systems using LIGO data. *The European Physical Journal Plus*, 132(1), jan 2017.
- [13] Iair Arcavi, Curtis McCully, Griffin Hosseinzadeh, D. Andrew Howell, Sergiy Vasylyev, Dovi Poznanski, Michael Zaltzman, Dan Maoz, Leo Singer, Stefano Valenti, Daniel Kasen, Jennifer Barnes, Tsvi Piran, and Wen fai Fong. Optical follow-up of gravitational-wave events with las cumbres observatory. *The Astrophysical Journal*, 848(2):L33, oct 2017.

- [14] L. Armus, S. T. Megeath, L. Corrales, M. Marengo, A. Kirkpatrick, J. D. Smith, M. Meyer, S. Gezari, R. P. Kraft, S. McCandliss, S. Tuttle, M. Elvis, M. Bentz, B. Binder, F. Civano, D. Dragomir, C. Espaillat, S. Finkelstein, D. B. Fox, M. Greenhouse, E. Hamden, J. Kauffmann, G. Khullar, J. Lazio, J. Lee, C. Lillie, P. Lightsey, R. Mushotzky, C. Scarlata, P. Scowen, G. R. Tremblay, Q. D. Wang, and S. Wolk. Great observatories: The past and future of panchromatic astrophysics, 2021.
- [15] M Celeste Artale, Yann Bouffanais, Michela Mapelli, Nicola Giacobbo, Nadeen B Sabha, Filippo Santoliquido, Mario Pasquato, and Mario Spera. An astrophysically motivated ranking criterion for low-latency electromagnetic follow-up of gravitational wave events. *Monthly Notices of the Royal Astronomical Society*, 495(2):1841–1852, may 2020.
- [16] Stefano Ascenzi, Gor Oganesyan, Marica Branchesi, and Riccardo Ciolfi. Electromagnetic counterparts of compact binary mergers. *Journal of Plasma Physics*, 87(1), feb 2021.
- [17] Gregory Ashton, Moritz Hübner, Paul D. Lasky, Colm Talbot, Kendall Ackley, Sylvia Biscoveanu, Qi Chu, Atul Divakarla, Paul J. Easter, Boris Goncharov, Francisco Hernandez Vivanco, Jan Harms, Marcus E. Lower, Grant D. Meadors, Denyz Melchor, Ethan Payne, Matthew D. Pitkin, Jade Powell, Nikhil Sarin, Rory J. E. Smith, and Eric Thrane. Bilby: A user-friendly bayesian inference library for gravitational-wave astronomy. *The Astrophysical Journal Supplement Series*, 241(2):27, Apr 2019.
- [18] Vishal Baibhav, Emanuele Berti, Davide Gerosa, Michela Mapelli, Nicola Giacobbo, Yann Bouffanais, and Ugo N. Di Carlo. Gravitational-wave detection rates for compact binaries formed in isolation: LIGO/virgo o3 and beyond. *Physical Review D*, 100(6), Sep 2019.
- [19] Arvind Balasubramanian, Alessandra Corsi, Kunal P. Mooley, Murray Brightman, Gregg Hallinan, Kenta Hotokezaka, David L. Kaplan, Davide Lazzati, and Eric J. Murphy. Continued Radio Observations of GW170817 3.5 yr Post-merger. *Astrophys. J. Lett.*, 914(1):L20, 2021.
- [20] C. Barbieri, O. S. Salafia, A. Perego, M. Colpi, and G. Ghirlanda. Electromagnetic counterparts of black hole–neutron star mergers: dependence on the neutron star properties. *The European Physical Journal A*, 56(1), jan 2020.
- [21] I. Bartos, A. P. S. Croots, and S. Má rka. GALAXY SURVEY ON THE FLY: PROSPECTS OF RAPID GALAXY CATALOGING TO AID THE ELECTROMAGNETIC FOLLOW-UP OF GRAVITATIONAL WAVE OBSERVATIONS. *The Astrophysical Journal*, 801(1):L1, Feb 2015.
- [22] E. Bertin. Automatic Astrometric and Photometric Calibration with SCAMP. In C. Gabriel, C. Arviset, D. Ponz, and S. Enrique, editors, *Astronomical Data*

Analysis Software and Systems XV, volume 351 of *Astronomical Society of the Pacific Conference Series*, page 112, July 2006.

- [23] Bertin, E. and Arnouts, S. Sextractor: Software for source extraction. *Astron. Astrophys. Suppl. Ser.*, 117(2):393–404, 1996.
- [24] Luc Blanchet. Gravitational radiation from post-newtonian sources and inspiralling compact binaries. *Living Reviews in Relativity*, 17(1), feb 2014.
- [25] Charlotte Bond, Daniel Brown, Andreas Freise, and Kenneth A. Strain. Interferometer techniques for gravitational-wave detection. *Living Reviews in Relativity*, 19(1):3, Feb 2017.
- [26] A. et al. Buikema. Sensitivity and performance of the advanced ligo detectors in the third observing run. *Phys. Rev. D*, 102:062003, Sep 2020.
- [27] Alessandra Buonanno, Bala R. Iyer, Evan Ochsner, Yi Pan, and B. S. Sathyaprakash. Comparison of post-newtonian templates for compact binary inspiral signals in gravitational-wave detectors. *Phys. Rev. D*, 80:084043, Oct 2009.
- [28] Eric Burns, Aaron Tohuvavohu, James Buckley, Tito Dal Canton, S. Brad Cenko, John W. Conklin, Filippo D’Ammando, David Eichler, Chris Fryer, Alexander J. van der Horst, Marc Kamionkowski, Mansi Kasliwal, Raffaella Margutti, Brian D. Metzger, Kohta Murase, Samaya Nissanke, David Radice, John Tomsick, Colleen A. Wilson-Hodge, and Bing Zhang. A summary of multimessenger science with neutron star mergers, 2019.
- [29] Carlton M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23:1693–1708, Apr 1981.
- [30] J. Chapront, M. Chapront-Touzé, and G. Francou. Determination of the lunar orbital and rotational parameters and of the ecliptic reference system orientation from LLR measurements and IERS data. , 343:624–633, March 1999.
- [31] Hsin-Yu Chen, Philip S. Cowperthwaite, Brian D. Metzger, and Edo Berger. A program for multimessenger standard siren cosmology in the era of LIGO a, rubin observatory, and beyond. *The Astrophysical Journal*, 908(1):L4, feb 2021.
- [32] Hsin-Yu Chen, Salvatore Vitale, and Ramesh Narayan. Viewing angle of binary neutron star mergers. *Physical Review X*, 9(3), aug 2019.
- [33] Planck Collaboration. Planck 2013 results. XVI. cosmological parameters. *Astronomy & Astrophysics*, 571:A16, oct 2014.
- [34] Gregory F. Cooper. The computational complexity of probabilistic inference using bayesian belief networks. *Artificial Intelligence*, 42(2):393–405, 1990.

- [35] D. A. Coulter, R. J. Foley, C. D. Kilpatrick, M. R. Drout, A. L. Piro, B. J. Shappee, M. R. Siebert, J. D. Simon, N. Ulloa, D. Kasen, B. F. Madore, A. Murguia-Berthier, Y.-C. Pan, J. X. Prochaska, E. Ramirez-Ruiz, A. Rest, and C. Rojas-Bravo. Swope supernova survey 2017a (SSS17a), the optical counterpart to a gravitational wave source. *Science*, 358(6370):1556–1558, dec 2017.
- [36] Curt Cutler and Éanna E. Flanagan. Gravitational waves from merging compact binaries: How accurately can one extract the binary’s parameters from the inspiral waveform? *Physical Review D*, 49(6):2658–2697, mar 1994.
- [37] G Dálya, R Díaz, F R Bouchet, Z Frei, J Jasche, G Lavaux, R Macas, S Mukherjee, M Pálfi, R S de Souza, B D Wandelt, M Bilicki, and P Raffai. GLADE : an extended galaxy catalogue for multimessenger searches with advanced gravitational-wave detectors. *Monthly Notices of the Royal Astronomical Society*, 514(1):1403–1411, May 2022.
- [38] K. Danzmann. Lisa mission overview. *Advances in Space Research*, 25(6):1129–1136, 2000. Fundamental Physics in Space.
- [39] Tim Dietrich, Tanja Hinderer, and Anuradha Samajdar. Interpreting binary neutron star mergers: describing the binary neutron star dynamics, modelling gravitational waveforms, and analyzing detections. *General Relativity and Gravitation*, 53(3):27, Mar 2021.
- [40] Tim Dietrich, Tanja Hinderer, and Anuradha Samajdar. Interpreting binary neutron star mergers: describing the binary neutron star dynamics, modelling gravitational waveforms, and analyzing detections. *General Relativity and Gravitation*, 53(3):27, Mar 2021.
- [41] Linda Dressel. WFC3 Instrument Handbook for Cycle 29 v. 13. In *WFC3 Instrument Handbook for Cycle 29 v. 13*, volume 13, page 13. 2021.
- [42] U. Dupletsas, J. Harms, B. Banerjee, M. Branchesi, B. Goncharov, A. Maselli, A. C. S. Oliveira, S. Ronchini, and J. Tissino. GWFISH: A simulation software to evaluate parameter-estimation capabilities of gravitational-wave detector networks. *Astronomy and Computing*, 42:100671, January 2023.
- [43] Freeman Dyson. Is a graviton detectable? *International Journal of Modern Physics A*, 28(25):1330041, 2013.
- [44] G. Efstathiou. A lockdown perspective on the hubble tension (with comments from the sh0es team), 2020.
- [45] A. Buikema et al. Sensitivity and performance of the advanced LIGO detectors in the third observing run. *Physical Review D*, 102(6), Sep 2020.
- [46] Adam G. Riess et al. A comprehensive measurement of the local value of the hubble constant with 1 km/s/mpc uncertainty from the hubble space telescope and the sh0es team. *The Astrophysical Journal Letters*, 934(1):L7, Jul 2022.

- [47] Andrew F. Cheng et al. Dart mission determination of momentum transfer: Model of ejecta plume observations. *Icarus*, 352:113989, 2020.
- [48] Astropy Collaboration et al. The Astropy Project: Sustaining and Growing a Community-oriented Open-source Project and the Latest Major Release (v5.0) of the Core Package. , 935(2):167, August 2022.
- [49] B. P. Abbott et al. Gravitational waves and gamma-rays from a binary neutron star merger: GW170817 and GRB 170817a. *The Astrophysical Journal*, 848(2):L13, oct 2017.
- [50] B. P. Abbott et al. Multi-messenger observations of a binary neutron star merger. *The Astrophysical Journal*, 848(2):L12, oct 2017.
- [51] B. P. Abbott et al. GW170817: Observation of gravitational waves from a binary neutron star inspiral. *Physical Review Letters*, 119(16), oct 2017.
- [52] Curtis McCully et al. The rapid reddening and featureless optical spectra of the optical counterpart of GW170817, AT 2017gfo, during the first four days. *The Astrophysical Journal*, 848(2):L32, oct 2017.
- [53] D. Spergel et al. Wfirst-2.4: What every astronomer should know, 2013.
- [54] David Reitze et al. Cosmic explorer: The u.s. contribution to gravitational-wave astronomy beyond ligo, 2019.
- [55] Eric C. Bellm et al. The zwicky transient facility: System overview, performance, and first results. *Publications of the Astronomical Society of the Pacific*, 131(995):018002, dec 2018.
- [56] F Acernese et al. Advanced virgo: a second-generation interferometric gravitational wave detector. *Classical and Quantum Gravity*, 32(2):024001, dec 2014.
- [57] Geraint Pratten et al. Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes. *Physical Review D*, 103(10), may 2021.
- [58] K. Ackley et al. Neutron star extreme matter observatory: A kilohertz-band gravitational-wave detector in the global network. *Publications of the Astronomical Society of Australia*, 37, 2020.
- [59] M. Nicholl et al. The electromagnetic counterpart of the binary neutron star merger LIGO/virgo GW170817. III. optical and UV spectra of a blue kilonova from fast polar ejecta. *The Astrophysical Journal*, 848(2):L18, oct 2017.
- [60] M Punturo et al. The einstein telescope: a third-generation gravitational wave observatory. *Classical and Quantum Gravity*, 27(19):194002, Sep 2010.

- [61] M. Soares-Santos et al. The electromagnetic counterpart of the binary neutron star merger LIGO/virgo GW170817. i. discovery of the optical counterpart using the dark energy camera. *The Astrophysical Journal*, 848(2):L16, oct 2017.
- [62] M.G. Aartsen et al. The icecube neutrino observatory: instrumentation and online systems. *Journal of Instrumentation*, 12(03):P03012, mar 2017.
- [63] N Arnaud et al. Virgo next: beyond the adv+ project. May 2022.
- [64] Oleg Korobkin et al. Axisymmetric radiative transfer models of kilonovae. *The Astrophysical Journal*, 910(2):116, Apr 2021.
- [65] R. Abbott et al. Population properties of compact objects from the second LIGO–virgo gravitational-wave transient catalog. *The Astrophysical Journal Letters*, 913(1):L7, may 2021.
- [66] R X Adhikari et al. A cryogenic silicon interferometer for gravitational-wave detection. *Classical and Quantum Gravity*, 37(16):165003, Jul 2020.
- [67] Simeon Reusch et al. Candidate tidal disruption event AT2019fdr coincident with a high-energy neutrino. *Physical Review Letters*, 128(22), Jun 2022.
- [68] The Event Horizon Telescope Collaboration et al. First m87 event horizon telescope results. iv. imaging the central supermassive black hole. *The Astrophysical Journal Letters*, 875(1):L4, apr 2019.
- [69] V. M. Lipunov et al. MASTER optical detection of the first LIGO/virgo neutron star binary merger GW170817. *The Astrophysical Journal*, 850(1):L1, nov 2017.
- [70] Zaven Arzoumanian et al. The NANOGrav 12.5 yr data set: Search for an isotropic stochastic gravitational-wave background. *The Astrophysical Journal Letters*, 905(2):L34, dec 2020.
- [71] A. C. Fabian. UHURU - the first X-ray astronomy satellite. *Journal of the British Interplanetary Society*, 28:343–346, May 1975.
- [72] XiLong Fan, Christopher Messenger, and Ik Siong Heng. A BAYESIAN APPROACH TO MULTI-MESSENGER ASTRONOMY: IDENTIFICATION OF GRAVITATIONAL-WAVE HOST GALAXIES. *The Astrophysical Journal*, 795(1):43, oct 2014.
- [73] Rodrigo Fernández and Brian D. Metzger. Electromagnetic signatures of neutron star mergers in the advanced LIGO era. *Annual Review of Nuclear and Particle Science*, 66(1):23–45, oct 2016.
- [74] Ryan J Foley, David A Coulter, Charles D Kilpatrick, Anthony L Piro, Enrico Ramirez-Ruiz, and Josiah Schwab. Updated parameter estimates for GW190425 using astrophysical arguments and implications for the electromagnetic counterpart. *Monthly Notices of the Royal Astronomical Society*, 494(1):190–198, mar 2020.

- [75] C. Freiburghaus, S. Rosswog, and F.-K. Thielemann. r-process in neutron star mergers. *The Astrophysical Journal*, 525(2):L121, oct 1999.
- [76] Danielle Frostig, Sylvia Biscoveanu, Geoffrey Mo, Viraj Karambelkar, Tito Dal Canton, Hsin-Yu Chen, Mansi Kasliwal, Erik Katsavounidis, Nathan P. Lourie, Robert A. Simcoe, and Salvatore Vitale. An infrared search for kilonovae with the WINTER telescope. i. binary neutron star mergers. *The Astrophysical Journal*, 926(2):152, feb 2022.
- [77] Chris L. Fryer, S. E. Woosley, Marc Herant, and Melvyn B. Davies. Merging white dwarf/black hole binaries and gamma-ray bursts. *The Astrophysical Journal*, 520(2):650–660, aug 1999.
- [78] Roberto Furfaro, Riccardo Barocco, Richard Linares, Francesco Topputo, Vishnu Reddy, Jules Simo, and Lucille Le Corre. Modeling irregular small bodies gravity field via extreme learning machines and bayesian optimization. *Advances in Space Research*, 67(1):617–638, 2021.
- [79] Shaon Ghosh, Steven Bloemen, Gijs Nelemans, Paul J. Groot, and Larry R. Price. Tiling strategies for optical follow-up of gravitational-wave triggers by telescopes with a wide field of view. *Astronomy & Astrophysics*, 592:A82, aug 2016.
- [80] Anuradha Gupta, Derek Fox, B. S. Sathyaprakash, and B. F. Schutz. Calibrating the cosmic distance ladder using gravitational-wave observations. *The Astrophysical Journal*, 886(1):71, nov 2019.
- [81] Ish Gupta, Ssohrab Borhanian, Arnab Dhani, Debatri Chattopadhyay, Rahul Kashyap, V. Ashley Villar, and B. S. Sathyaprakash. Neutron star-black hole mergers in next generation gravitational-wave observatories, 2023.
- [82] A. Hajela, R. Margutti, J. S. Bright, K. D. Alexander, B. D. Metzger, V. Nedora, A. Kathirgamaraju, B. Margalit, D. Radice, C. Guidorzi, E. Berger, A. MacFadyen, D. Giannios, R. Chornock, I. Heywood, L. Sironi, O. Gottlieb, D. Coppejans, T. Laskar, Y. Cendes, R. Barniol Duran, T. Eftekhar, W. Fong, A. McDowell, M. Nicholl, X. Xie, J. Zrake, S. Bernuzzi, F. S. Broekgaarden, C. D. Kilpatrick, G. Terreran, V. A. Villar, P. K. Blanchard, S. Gomez, G. Hosseinzadeh, D. J. Matthews, and J. C. Rastinejad. Evidence for x-ray emission in excess to the jet-afterglow decay 3.5 yr after the binary neutron star merger GW 170817: A new emission component. *The Astrophysical Journal Letters*, 927(1):L17, mar 2022.
- [83] Mark Hannam. Status of black-hole-binary simulations for gravitational-wave detection. *Classical and Quantum Gravity*, 26(11):114001, may 2009.
- [84] Karl G Henize, John F Stanley, Christine A O'Neill, and Barbara S Nowakowski. Detection of orbital debris with geodss telescopes. In *Space Debris Detection and Mitigation*, volume 1951, pages 76–84. SPIE, 1993.

- [85] G. Hinshaw, D. Larson, E. Komatsu, D. N. Spergel, C. L. Bennett, J. Dunkley, M. R. Nolta, M. Halpern, R. S. Hill, N. Odegard, L. Page, K. M. Smith, J. L. Weiland, B. Gold, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, G. S. Tucker, E. Wollack, and E. L. Wright. NINE-YEAR iWILKINSON MICROWAVE ANISOTROPY PROBE/i (iWMAP/i) OBSERVATIONS: COSMOLOGICAL PARAMETER RESULTS. *The Astrophysical Journal Supplement Series*, 208(2):19, Sep 2013.
- [86] K. Hirata, T. Kajita, M. Koshiba, M. Nakahata, Y. Oyama, N. Sato, A. Suzuki, M. Takita, Y. Totsuka, T. Kifune, T. Suda, K. Takahashi, T. Tanimori, K. Miyano, M. Yamada, E. W. Beier, L. R. Feldscher, S. B. Kim, A. K. Mann, F. M. Newcomer, R. Van, W. Zhang, and B. G. Cortez. Observation of a neutrino burst from the supernova sn1987a. *Phys. Rev. Lett.*, 58:1490–1493, Apr 1987.
- [87] N.B. Hogg, M. Martinelli, and S. Nesseris. Constraints on the distance duality relation with standard sirens. *Journal of Cosmology and Astroparticle Physics*, 2020(12):019–019, dec 2020.
- [88] Alex R. Howe, Adam Burrows, and Drake Deming. AN INFORMATION-THEORETIC APPROACH TO OPTIMIZEiJWST/iOBSERVATIONS AND RETRIEVALS OF TRANSITING EXOPLANET ATMOSPHERES. *The Astrophysical Journal*, 835(1):96, jan 2017.
- [89] Jian-Ping Hu and Fa-Yin Wang. Hubble tension: The evidence of new physics, 2023.
- [90] Sabrina et al. Huth. Constraining neutron-star matter with microscopic and macroscopic collisions. *Nature*, 606(7913):276–280, Jun 2022.
- [91] Soichiro Isoyama, Riccardo Sturani, and Hiroyuki Nakano. Post-newtonian templates for gravitational waves from compact binary inspirals. In *Handbook of Gravitational Wave Astronomy*, pages 1–49. Springer Singapore, 2021.
- [92] Piotr Jaranowski and Andrzej Królak. Gravitational-wave data analysis. formalism and sample applications: The gaussian case, 2007.
- [93] Xisco Jiménez-Forteza, David Keitel, Sascha Husa, Mark Hannam, Sebastian Khan, and Michael Pürller. Hierarchical data-driven approach to fitting numerical relativity data for nonprecessing binary black holes with an application to final spin and radiated energy. *Physical Review D*, 95(6), mar 2017.
- [94] W. D’Arcy Kenworthy, Dan Scolnic, and Adam Riess. The local perspective on the hubble tension: Local structure does not impact measurement of the hubble constant. *The Astrophysical Journal*, 875(2):145, Apr 2019.
- [95] Bahador Khaleghi, Alaa Khamis, Fakhreddine O. Karray, and Saiedeh N. Razavi. Multisensor data fusion: A review of the state-of-the-art. *Information Fusion*, 14(1):28–44, 2013.

- [96] Bulent Kiziltan, Athanasios Kottas, and Stephen E. Thorsett. The neutron star mass distribution, 2010.
- [97] Johan Kwisthout. Approximate inference in bayesian networks: Parameterized complexity results. *International Journal of Approximate Reasoning*, 93:119–131, 2018.
- [98] François et al. Lemoine. The development of the joint nasa gsfc and the national imagery and mapping agency (nima) geopotential model egm96. 08 1998.
- [99] LIGO Scientific Collaboration. LIGO Algorithm Library - LALSuite. free software (GPL), 2018.
- [100] Yarong Luo, Jianlang Hu, and Chi Guo. Kalman filter from the mutual information perspective, 2021.
- [101] Chris Messenger Martin Hendry Man Leong Chan, Yi-Ming Hu and Ik Siong Heng. Maximizing the detection probability of kilonovae associated with gravitational wave observations. *The Astrophysical Journal*, 834(1):84, Jan 2017.
- [102] Mastrogiovanni, S., Duque, R., Chassande-Mottin, E., Daigne, F., and Mochkovitch, R. The potential role of binary neutron star merger afterglows in multimessenger cosmology. *A&A*, 652:A1, 2021.
- [103] Andrea Melandri. The evolution of the x-ray and radio emission of gw 170817/grb 170817a. *Nuclear and Particle Physics Proceedings*, 306-308:50–52, 2019. CRIS 2018 "Entering the Era of Multi-Messenger Astronomy".
- [104] Péter Mészáros, Derek B. Fox, Chad Hanna, and Kohta Murase. Multi-messenger astrophysics. *Nature Reviews Physics*, 1(10):585–599, Oct 2019.
- [105] B. D. Metzger, G. Martínez-Pinedo, S. Darbha, E. Quataert, A. Arcones, D. Kasen, R. Thomas, P. Nugent, I. V. Panov, and N. T. Zinner. Electromagnetic counterparts of compact object mergers powered by the radioactive decay of r-process nuclei. *Monthly Notices of the Royal Astronomical Society*, 406(4):2650–2662, Jun 2010.
- [106] Aaron et al. Meurer. Sympy: symbolic computing in python. *PeerJ Computer Science*, 3:e103, January 2017.
- [107] R. Mochkovitch, F. Daigne, R. Duque, and H. Zitouni. Prospects for kilonova signals in the gravitational-wave era. *Astronomy & Astrophysics*, may 2021.
- [108] C J Moore, R H Cole, and C P L Berry. Gravitational-wave sensitivity curves. *Classical and Quantum Gravity*, 32(1):015014, dec 2014.
- [109] C J Moore, R H Cole, and C P L Berry. Gravitational-wave sensitivity curves. *Classical and Quantum Gravity*, 32(1):015014, dec 2014.

- [110] Ehud Nakar. The electromagnetic counterparts of compact binary mergers. *Physics Reports*, 886:1–84, nov 2020.
- [111] Tatsuya Narikawa, Masato Kaneyama, and Hideyuki Tagoshi. Optimal follow-up observations of gravitational wave events with small optical telescopes. *Phys. Rev. D*, 96:084067, Oct 2017.
- [112] NASA. Atmospheric electromagnetic transmittance or opacity.
- [113] Engineering National Academies of Sciences and Medicine. *Pathways to Discovery in Astronomy and Astrophysics for the 2020s*. The National Academies Press, Washington, DC, 2021.
- [114] G. et al. Neugebauer. The Infrared Astronomical Satellite (IRAS) mission. , 278:L1–L6, March 1984.
- [115] Matt Nicholl, Ben Margalit, Patricia Schmidt, Graham P Smith, Evan J Ridley, and James Nuttall. Tight multimessenger constraints on the neutron star equation of state from GW170817 and a forward model for kilonova light-curve synthesis. *Monthly Notices of the Royal Astronomical Society*, 505(2):3016–3032, may 2021.
- [116] Samaya Nissanke, Mansi Kasliwal, and Alexandra Georgieva. IDENTIFYING ELUSIVE ELECTROMAGNETIC COUNTERPARTS TO GRAVITATIONAL WAVE MERGERS: AN END-TO-END SIMULATION. *The Astrophysical Journal*, 767(2):124, Apr 2013.
- [117] Library of Congress. Galileo and the telescope. 2021.
- [118] NASA Orbital Debris Program Office. Orbital debris quarterly news, Mar 2023.
- [119] P Padovani, P Giommi, E Resconi, T Glauch, B Arsioli, N Sahakyan, and M Huber. Dissecting the region around IceCube-170922a: the blazar TXS 0506056 as the first cosmic neutrino source. *Monthly Notices of the Royal Astronomical Society*, 480(1):192–203, Jul 2018.
- [120] Joseph C Pemberton and LG Greenwald. On the need for dynamic scheduling of imaging satellites. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, 34(1):165–171, 2002.
- [121] L. Perivolaropoulos and F. Skara. Challenges for λ CDM: An update. *New Astronomy Reviews*, 95:101659, dec 2022.
- [122] Rosalba Perna, Davide Lazzati, and Will Farr. Limits on electromagnetic counterparts of gravitational-wave-detected binary black hole mergers. *The Astrophysical Journal*, 875(1):49, Apr 2019.
- [123] Abigail Pillitteri. Fundamental forces and particles. In *The Universe Untangled*, 2053–2571, pages 2–1 to 2–9. Morgan Claypool Publishers, 2017.

- [124] Planck Collaboration. Planck 2018 results - vi. cosmological parameters. *A&A*, 641:A6, 2020.
- [125] David Radice, Sebastiano Bernuzzi, and Albino Perego. The dynamics of binary neutron star mergers and GW170817. *Annual Review of Nuclear and Particle Science*, 70(1):95–119, oct 2020.
- [126] Jocelyn S. et al. Read. Measuring the neutron star equation of state with gravitational wave observations. *Phys. Rev. D*, 79:124033, Jun 2009.
- [127] Maria Isabel Ribeiro. Kalman and extended kalman filters: Concept, derivation and properties. *Institute for Systems and Robotics*, 43(46):3736–3741, 2004.
- [128] Adam G. Riess. The expansion of the universe is faster than expected. *Nature Reviews Physics*, 2(1):10–12, Jan 2020.
- [129] J.A. Rueda, R. Ruffini, Y. Wang, Y. Aimurato, U. Barres de Almeida, C.L. Bianco, Y.C. Chen, R.V. Lobato, C. Maia, D. Primorac, R. Moradi, and J.F. Rodriguez. GRB 170817a-GW170817-AT 2017gfo and the observations of NS-NS, NS-WD and WD-WD mergers. *Journal of Cosmology and Astroparticle Physics*, 2018(10):006–006, oct 2018.
- [130] M. Ruffert, H. Th. Janka, and G. Schaefer. Coalescing neutron stars – a step towards physical models. i. hydrodynamic evolution and gravitational-wave emission, 1995.
- [131] Filippo Santoliquido, Michela Mapelli, Nicola Giacobbo, Yann Bouffanais, and M Celeste Artale. The cosmic merger rate density of compact objects: impact of star formation, metallicity, initial mass function, and binary evolution. *Monthly Notices of the Royal Astronomical Society*, 502(4):4877–4889, feb 2021.
- [132] R. Schlickeiser. *Cosmic Ray Astrophysics*. Astronomy and Astrophysics Library. Springer Berlin Heidelberg, 2013.
- [133] B. Schutz, B. Tapley, and G.H. Born. *Statistical Orbit Determination*. Elsevier Science, 2004.
- [134] Bernard F. Schutz. Determining the hubble constant from gravitational wave observations. *Nature*, 323(6086):310–311, Sep 1986.
- [135] Bernard F. Schutz. Determining the hubble constant from gravitational wave observations. *Nature*, 323(6086):310–311, Sep 1986.
- [136] Lynn G. Seppala. Improved optical design for the Large Synoptic Survey Telescope (LSST). In J. Anthony Tyson and Sidney Wolff, editors, *Survey and Other Telescope Technologies and Discoveries*, volume 4836, pages 111 – 118. International Society for Optics and Photonics, SPIE, 2002.

- [137] A. Sesana, A. Vecchio, M. Eracleous, and S. Sigurdsson. Observing white dwarfs orbiting massive black holes in the gravitational wave and electro-magnetic window. *Monthly Notices of the Royal Astronomical Society*, 391(2):718–726, dec 2008.
- [138] Christian N Setzer, Hiranya V Peiris, Oleg Korobkin, and Stephan Rosswog. Modelling populations of kilonovae. *Monthly Notices of the Royal Astronomical Society*, 520(2):2829–2842, Jan 2023.
- [139] Michael et al. Shao. Finding very small near-earth asteroids using synthetic tracking. *The Astrophysical Journal*, 782, 09 2013.
- [140] David Sibert, David Borgeson, Glenn Peterson, Alan Jenkin, and Marlon Sorge. Operational impact of improved space tracking on collision avoidance in the future leo space debris environment. Technical report, SPACE AND MISSILE SYSTEMS CENTER LOS ANGELES AFB CA, 2010.
- [141] Daniel M. Siegel. Gw170817 –the first observed neutron star merger and its kilonova: Implications for the astrophysical site of the r-process. *The European Physical Journal A*, 55(11):203, Nov 2019.
- [142] Peng Mun Siew and Richard Linares. Optimal tasking of ground-based sensors for space situational awareness using deep reinforcement learning. *Sensors*, 22(20), 2022.
- [143] Leo Singer, Larry Price, and Antony Speranza. Optimizing optical follow-up of gravitational-wave candidates, 2012.
- [144] Leo P. Singer and Larry R. Price. Rapid bayesian position reconstruction for gravitational-wave transients. *Physical Review D*, 93(2), Jan 2016.
- [145] W. T. Sullivan. *The Early Years of Radio Astronomy: Reflections Fifty Years after Jansky’s Discovery*. Cambridge University Press, 1984.
- [146] Ryuichi Takahashi, Soichiro Morisaki, and Teruaki Suyama. Probing the solar interior with lensed gravitational waves from known pulsars, 2023.
- [147] N. R. Tanvir, A. J. Levan, and Vinrouge Collaboration. LIGO/Virgo G298048: VISTA/VIRCAM detection of candidate counterpart. *GRB Coordinates Network*, 21544:1, August 2017.
- [148] Eric Thrane and Colm Talbot. An introduction to bayesian inference in gravitational-wave astronomy: Parameter estimation, model selection, and hierarchical models. *Publications of the Astronomical Society of Australia*, 36, 2019.
- [149] Universe Today. All of hubble’s observations in one picture, 2022.

- [150] Toonen, S., Perets, H. B., Igoshev, A. P., Michaely, E., and Zenati, Y. The demographics of neutron star - white dwarf mergers - rates, delay-time distributions, and progenitors. *A&A*, 619:A53, 2018.
- [151] Stefano Valenti, J. Sand David, Sheng Yang, Enrico Cappellaro, Leonardo Tartaglia, Alessandra Corsi, Saurabh W. Jha, Daniel E. Reichart, Joshua Haislip, and Vladimir Kouprianov. The discovery of the electromagnetic counterpart of GW170817: Kilonova AT 2017gfo/DLT17ck. *The Astrophysical Journal*, 848(2):L24, oct 2017.
- [152] Stefano Valenti, J. Sand David, Sheng Yang, Enrico Cappellaro, Leonardo Tartaglia, Alessandra Corsi, Saurabh W. Jha, Daniel E. Reichart, Joshua Haislip, and Vladimir Kouprianov. The discovery of the electromagnetic counterpart of GW170817: Kilonova AT 2017gfo/DLT17ck. *The Astrophysical Journal*, 848(2):L24, oct 2017.
- [153] Eleonora Di Valentino, Olga Mena, Supriya Pan, Luca Visinelli, Weiqiang Yang, Alessandro Melchiorri, David F Mota, Adam G Riess, and Joseph Silk. In the realm of the hubble tension—a review of solutions. *Classical and Quantum Gravity*, 38(15):153001, Jul 2021.
- [154] D.A. Vallado and W.D. McClain. *Fundamentals of Astrodynamics and Applications*. Fundamentals of Astrodynamics and Applications. Microcosm Press, 2001.
- [155] David A Vallado and John H Seago. Covariance realism. *Advances in the Astronautical Sciences*, 135(1):49–67, 2009.
- [156] Michele Vallisneri. Use and abuse of the fisher information matrix in the assessment of gravitational-wave parameter-estimation prospects. *Physical Review D*, 77(4), feb 2008.
- [157] Don van Ravenzwaaij, Pete Cassey, and Scott D. Brown. A simple introduction to markov chain monte-carlo sampling. *Psychonomic Bulletin & Review*, 25(1):143–154, Feb 2018.
- [158] Licia Verde, Tommaso Treu, and Adam G. Riess. Tensions between the early and late universe. *Nature Astronomy*, 3(10):891–895, Sep 2019.
- [159] Pauli et al. Virtanen. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.
- [160] R. Voss and T. M. Tauris. Galactic distribution of merging neutron stars and black holes – prospects for short gamma-ray burst progenitors and LIGO/VIRGO. *Monthly Notices of the Royal Astronomical Society*, 342(4):1169–1184, 07 2003.

- [161] Hanbiao Wang, Kung Yao, and Deborah Estrin. Information-theoretic approaches for sensor selection and placement in sensor networks for target localization and tracking. *Journal of Communications and Networks*, 7(4):438–449, 2005.
- [162] James R. Weiss, William Smythe, and Wenwen Lu. Science traceability. *2005 IEEE Aerospace Conference*, pages 292–299, 2005.
- [163] Martin C. Weisskopf. Chandra x-ray optics. *Optical Engineering*, 51(1):011013, feb 2012.
- [164] Christine Ye and Maya Fishbach. Cosmology with standard sirens at cosmic noon. *Physical Review D*, 104(4), aug 2021.
- [165] Z. Yu, Zhang JiXian, and John Lodewijk Van Genderen. Comparison and analysis of remote sensing data fusion techniques at feature and decision levels. 2006.
- [166] Yongmian Zhang and Qiang Ji. Efficient sensor selection for active information fusion. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 40(3):719–728, 2010.
- [167] Željko Ivezić, Lynne Jones, and Robert H. Lupton. The lsst photon rates and snr calculations , v1.2. 2010.