KU LEUVEN

FACULTY OF
ENGINEERING SCIENCE

Unmodelled Detection of Overlapping Gravitational Wave Signals with the Einstein Telescope

Stijn Heyde

Thesis submitted for the degree of
Master of Science in
Electrical Engineering, option
Electronics and Chip Design

Supervisors
Prof. dr. Tjonnie Li
Prof. dr. ir. Marc Moonen

Assessors
Prof. dr. ir. Hugo Van hamme
Dr. Henri Inchauspé

Assistant-supervisors
Ir. M. Wils
MSc. F. Cireddu

Academic year 2025 – 2026

Published by Stijn Heyde,
Department of Electrical Engineering, Kasteelpark Arenberg 10 bus 2440, B-3001 Leuven

Preface

Over the past four months or so, this thesis grew from an unfamiliar subject into the work presented here. It did so with the help of many people. First and foremost, I would like to thank my assistant-supervisors, Milan Wils and Francesco Cireddu. They were the people I met with most often, always flexible about when we could sit down together, and they introduced me to a subject I knew little about at the start. Their proofreading improved this text considerably. I am also grateful to Professor Tjonnie Li, who helped me appreciate that a good thesis is not about including all the data one has, but about organising the figures around a story. My thanks go as well to everyone who reads and evaluates this work, and to anyone who may read it in the future. Last but not least, I thank my parents, whose support brought me to where I am today, and my friends and everyone close to me for their encouragement along the way.

Stijn Heyde

Contents

Abstract

The third generation of gravitational-wave observatories, including the Einstein Telescope, will detect hundreds of thousands of compact-binary coalescences per year. A rare Galactic core-collapse supernova signal arriving during such observations would land in a detector record likely already convoluted by binary inspirals, yet current burst pipelines assume one source per trigger and would blend the two. This thesis addresses the problem in three steps.

First, the time-frequency overlap probability between a Galactic core-collapse supernova and a realistic one-year population of binary inspirals is computed for the Einstein Telescope. At the low frequencies the detector reaches, the temporal overlap with binary neutron stars is essentially 100% and the time-frequency overlap reaches 87–91% for the heavier supernova progenitors, concentrated in the sub-16 Hz band where the supernova memory signal deposits most of its energy.

Second, the Modular Data Analysis Pipeline, an existing burst pipeline, is extended with a completely new sky-localisation module and restructured so that every stage from there onward carries a list of source hypotheses rather than a single one. New multi-source implementations of the likelihood, clustering and event-reconstruction modules are also developed, while the pipeline's modularity is preserved. An end-to-end test with a simultaneous binary black hole and supernova injection identifies, separates and reconstructs both sources in a single pass.

Third, this new module localises by extending the established single-source coherent-network statistic into a joint maximum-likelihood framework for two simultaneous sources, with a per-time-frequency-pixel test for how many sources are present. Validated on a network consisting of the Einstein Telescope plus Cosmic Explorer plus LIGO-India under realistic noise, it reaches single-digit-degree single-source accuracy at the reference distances of each source type and localises all three pairwise multi-source combinations.

List of Figures and Tables

List of Figures

List of Tables

List of abbreviations and symbols

Abbreviations

A1	LIGO-India (third-generation detector)
ASD	Amplitude spectral density
BBH	Binary black hole
BNS	Binary neutron star
CBC	Compact binary coalescence
CCSN	Core-collapse supernova
CE	Cosmic Explorer
CR	Credible region
cWB	Coherent WaveBurst
DWT	Discrete wavelet transform
ECEF	Earth-centred Earth-fixed (coordinate frame)
EM	Expectation-maximization
ET	Einstein Telescope
FFT	Fast Fourier transform
GPS	Global Positioning System
GW	Gravitational wave
HPD	Highest posterior density
LIGO	Laser Interferometer Gravitational-wave Observatory
LTF	Likelihood time-frequency
MDAP	Modular Data Analysis Pipeline
ML	Maximum likelihood
MUSIC	Multiple signal classification
PSD	Power spectral density
SASI	Standing accretion shock instability
SNR	Signal-to-noise ratio
STFT	Short-time Fourier transform
TF	Time-frequency
3G	Third generation (detector era)

Symbols

$h_+(t)$, $h_\times(t)$	Plus and cross GW polarization strain
$F_+^{(k)}$, $F_\times^{(k)}$	Antenna pattern functions of detector k
$\tau_k(\hat{n})$	Signal time delay at detector k from sky direction $\hat{n}$
$n_k(t)$	Noise in detector k
$\hat{n}$	Unit vector pointing toward the GW source on the sky
χ	Cross-polarization fraction
θ_{jn}	Inclination angle (orbital angular momentum to line of sight)
ψ	Polarization angle
K_{eff}	Number of independent GW-sensitive data streams
$S_n(f)$	One-sided noise power spectral density
E_{coh}	Coherent energy
E_{null}	Null-stream energy
E_{total}	Total incoherent energy
$\Lambda_{\mathrm{single}}$	Per-pixel single-source sky localisation statistic
$\Lambda_{\mathrm{total}}(\hat{n}_1, \hat{n}_2)$	Joint ML sky surface for two sources
$\eta(s)$	TF overlap fraction at time shift s
P_{ov}	Overlap probability
ξ	Null fraction classifier
$\varepsilon_{\mathrm{null}}$	Null fraction threshold
$\mathbf{a}_{\mathrm{eff}}$	Effective steering vector (CBC, polarization-reduced)
$A(\hat{n}, f)$	Steering matrix ($K \times 2$ per source)
M_1, M_2	Marginal sky maps for source 1 and 2
$M_\odot$	Solar mass

Chapter 1

Introduction

The third generation of gravitational-wave (GW) observatories will extend the network sensitivity by an order of magnitude across the current detection band and push the low-frequency cutoff down to a few hertz [1, 2]. With the Einstein Telescope (ET), Cosmic Explorer (CE), and LIGO-India (A1) operating jointly in the 2030s, the expected detection rate rises to $\mathcal{O}(10^5)$ binary neutron star (BNS) and $\mathcal{O}(10^6)$ binary black hole (BBH) coalescences per year [3], with individual BNS inspirals spending hours in the ET band. A once-per-$\sim$60-year Galactic core-collapse supernova (CCSN) [4] would therefore arrive in a detector record likely already convoluted by compact-binary coalescence (CBC) signals.

Coherent burst pipelines such as Coherent WaveBurst (cWB) [5] and the Modular Data Analysis Pipeline (MDAP) [6] are built around a single-source assumption. Each cluster of excess-power time-frequency pixels is treated as one event, and no representation of signal multiplicity is propagated downstream of the clustering stage. When two unrelated signals share the same time-frequency neighbourhood, the cluster blends them and the reconstruction is biased toward the more phase-coherent contribution. For a rare Galactic CCSN, occurring only once every few decades, a missed detection would forfeit a once-in-a-generation observation.

Research gaps

Three concrete gaps emerge from the literature reviewed in Chapter 2.

1. **Quantification.** The probability and time-frequency (TF) structure of CCSN–CBC overlap have not been quantified for the Einstein Telescope sensitivity against realistic population-level CBC rates.

2. **Pipeline support.** Burst pipelines lack a multi-source data model. Each cluster is propagated downstream as a single coherent event. The MDAP framework [6], developed in a recent master's thesis as a modular alternative to monolithic pipelines, still enforces the single-source assumption through the data products exchanged between its modules.

3. **Multi-source localisation.** Coherent sky-localisation methods assume one source per coherent trigger. Adaptations to the unmodelled-burst framework for jointly localising multiple simultaneous sources have not been sufficiently studied.

Thesis direction

This thesis addresses all three gaps. The approach throughout is to keep the coherent likelihood framework and relax the single-source assumption at each stage where it is currently embedded, rather than attempting blind source separation, which is not well-suited to multi-source GW data with a small detector network [7].

Chapter 2 reviews the literature on the astrophysical sources, the third-generation network, the coherent burst pipelines on which this work builds, and the prior treatment of overlapping GW signals.

Chapter 3 quantifies the TF overlap. The temporal and TF overlap probability between a Galactic CCSN and a realistic one-year CBC population (BNS and BBH) is computed for ET using both short-time Fourier and discrete wavelet transforms. The headline result is that overlap is essentially 100% at the low frequencies ET is designed to reach, which is also where the CCSN memory effect concentrates most of its energy [8]. This establishes the problem the rest of the thesis addresses.

Chapter 4 extends the MDAP architecture. The pipeline is reviewed with attention to how the single-source assumption is enforced through the data flow between modules. The adapted pipeline propagates a per-source descriptor from the sky-localisation stage through to the clustering output.

Chapter 5 develops the multi-source localiser. This chapter contains the main algorithmic contribution of the thesis. A maximum-likelihood algorithm is developed to determine the sky location of one or multiple simultaneous sources. The framework is validated on injections covering both the single-source case and multi-source combinations.

The conclusion summarises the contributions, identifies the remaining weaknesses, and outlines directions for future work.

Chapter 2

Literature Study

This chapter provides the background needed for the rest of the thesis: the astrophysical sources injected throughout (Section 2.2), the third-generation detector network (Section 2.3), the unmodelled burst pipelines on which this work builds (Section 2.4), and the prior treatment of overlapping signals (Section 2.5).

2.1 Gravitational-wave signals and detector response

A gravitational wave (GW) is a propagating transverse-traceless perturbation of spacetime, characterized by two polarization components $h_+(t)$ and $h_\times(t)$. The strain recorded by detector k for a source at sky direction $\hat{n}$ is

$$d_k(t) = F_+^{(k)}(\hat{n})\, h_+(t-\tau_k) + F_\times^{(k)}(\hat{n})\, h_\times(t-\tau_k) + n_k(t), \tag{2.1}$$

where $\tau_k(\hat{n})$ is the inter-detector time delay, $F_+^{(k)}, F_\times^{(k)}$ are the antenna pattern functions, and $n_k(t)$ is the detector noise [9].

For compact binaries the two polarizations are phase-locked by the orbital geometry. In the source frame, the frequency domain constraint reads $\tilde{h}_\times = \pm i\chi\tilde{h}_+$, where the cross-polarization fraction

$$\chi(\theta_{jn}) = \frac{2\cos\theta_{jn}}{1+\cos^2\theta_{jn}} \in [-1, 1] \tag{2.2}$$

is determined by the inclination angle θ_{jn} [9]. This polarization constraint reduces the per-source signal subspace from two dimensions to one, which is exploited in the multi-source sky-localisation framework of Chapter 5. CCSN signals do not satisfy this relation and require the full two-dimensional treatment.

2.2 Astrophysical sources

Two source classes drive this thesis: compact binary coalescences (well-modelled, deterministic, phase-coherent) and core-collapse supernovae (poorly-modelled, broadband, stochastic) [9]. The contrast in waveform predictability and polarization structure between them is what makes the CCSN–CBC overlap scenario qualitatively different from the CBC–CBC case studied in the literature (Section 2.5).

2.2.1 Compact binary coalescences

Compact binary coalescences (CBCs), pairs of neutron stars or black holes spiralling together, emit a coherent and deterministic gravitational-wave signal through inspiral, merger and ringdown [9]. The waveform is parametrised over mass, spin and orbital geometry, which makes matched filtering against template banks (e.g. PyCBC [10], GstLAL [11] pipelines) the optimal detection method when the signal is isolated in the data. The methods developed in this thesis do not depend on the specific waveform family beyond the polarization geometry of Eq. (2.2).

2.2.2 Core-collapse supernovae

Core-collapse supernova (CCSN) gravitational-wave emission arises from non-spherical, time-dependent mass motions during and after core bounce: proto-neutron-star oscillations, neutrino-driven convection behind the stalled shock, and large-scale asymmetric mass ejection [8]. The signal is a superposition of contributions from different regions of the star whose relative weight shifts over time, with high-frequency emission migrating inward through the post-bounce phase and a late-time low-frequency component from the outermost region that lies precisely in the band where ET's sensitivity advantage lies. Detailed morphology, amplitude and polarization structure are strongly model-dependent, so matched filtering is impractical and CCSN searches rely on unmodelled coherent excess-power methods [8].

2.3 Third-generation detector network

All analyses in this thesis use a third-generation (3G) ground-based detector network. The 3G era is defined by two flagship instruments that improve on Advanced LIGO/Virgo by roughly an order of magnitude in strain sensitivity and extend the observing band down to a few Hz [1]. The following network is the most realistic 3G scenario likely to operate during ET's observing run:

Einstein Telescope (ET). Two ET configurations are currently under consideration: a triangular underground European facility comprising three pairs of co-located, $\sim 10\,\mathrm{km}$ interferometers in a xylophone (low-frequency cryogenic + high-frequency room-temperature) configuration, and a 2L design with two L-shaped interferometers at separate European sites [2]. This thesis adopts the triangular design throughout. Its three interferometers share a common site, so the inter-detector timing baselines are effectively zero.

Cosmic Explorer (CE). A US-led 40 km L-shaped interferometer [12] with extreme sensitivity at intermediate frequencies, dominant over ET in much of the band relevant to compact binaries.

LIGO-India (A1). A 4 km LIGO-class interferometer [13] at the Aplus / A-sharp sensitivity level still operating in the 3G era. Its sensitivity is modest compared to ET and CE but its location adds a long timing baseline.

2.4 Existing burst pipelines

CCSN searches cannot rely on matched filtering and instead use unmodelled (burst) pipelines, which forgo templates and look for excess power consistent with a coherent GW signal arriving at all detectors from a common sky direction. Bursts are the standard tool for CCSN-like sources and underlie all analyses in this thesis.

2.4.1 Coherent WaveBurst (cWB)

Coherent WaveBurst (cWB) [5] is a frequentist coherent burst pipeline used by the LIGO–Virgo–KAGRA collaboration for unmodelled searches. It implements the coherent maximum-likelihood method of Chapter 5. Detector data are first whitened and decomposed into a wavelet time–frequency representation, pixel-level coherent likelihoods are evaluated locally, significant pixels are grouped into coherent clusters, and each cluster is ranked by a detection statistic combining its coherent likelihood with a network-correlation coefficient (next paragraph). The pipeline ends with empirical false-alarm calibration via time-slid background.

The cWB approximation introduces two design choices that are directly relevant to this thesis.

Coherence and consistency statistics. cWB combines the cluster-level $L_{\max}$ with a network correlation coefficient $c_c \in [0, 1]$ built from off-diagonal likelihood-matrix elements: high c_c marks coherent signals, low c_c marks glitch-dominated triggers. Chapter 5 uses the simpler per-pixel null fraction $E_{\text{null}}/E_{\text{total}}$ as a single/multi-source pixel classifier.

Single-source clustering. cWB groups significant TF pixels into clusters under the implicit assumption that all pixels in a cluster originate from one coherent source. Overlapping signals violate this assumption and bias the downstream likelihood.

2.4.2 Modular Data Analysis Pipeline

The Modular Data Analysis Pipeline (MDAP) of Mannaert [6] is a coherent burst pipeline organised around the same approximations as cWB but with the analysis stages (time–frequency transform, likelihood generation, clustering, event detection, reconstruction) implemented as independent modules exchanging standardised data products. This modularity is preserved throughout this thesis. However, the standardised data products themselves are organised around a single-source assumption, so the multi-source extension developed here goes beyond replacements of individual modules and modifies the data products exchanged between them in Chapter 4.

2.5 Overlapping gravitational-wave signals

2.5.1 Overlap in CBC searches

Relton et al. [14] investigated overlapping CBC signals in both the matched-filter pipeline PyCBC and the coherent burst pipeline cWB. For mergers separated by less than $\sim 1\,\mathrm{s}$, PyCBC's clustering routine retains only the louder of the two triggers and discards the quieter one. In cWB, the burst pipeline forms a single trigger and reconstructs the data under a single coherent-signal model, so the reconstruction represents both signals with one sky location and one pair of polarization waveforms. Part of the secondary signal leaks into the null subspace, degrading coherence statistics and reducing detection significance.

2.5.2 Why CCSN–CBC overlap is qualitatively different

The implications of these results are potentially more severe in CCSN–CBC overlap, where the CCSN memory band and the CBC inspiral band coincide at low frequencies (Section 2.2.2) and the two signals are highly likely to share TF pixels. When the burst pipeline merges them into a single trigger, the coherent reconstruction rewards phase coherence across detectors: the CBC's deterministic, phase-locked chirp aligns cleanly across the network while the CCSN's stochastic, weakly-polarized emission does not, so the reconstruction biases toward the CBC and the CCSN signal is absorbed into the incoherent residual.

2.5.3 Source separation and the limits of blind methods

Blind source-separation techniques from signal processing (independent component analysis, non-negative matrix factorisation, second-order blind identification, time–frequency methods, and neural networks) have been surveyed for the 3G overlap problem [7]. The review identifies three obstacles. First, the problem is underdetermined: ET's three co-located interferometers cannot separate more overlapping sources than available sensor channels. Second, GW signals have time-varying spectra (chirps) that violate the stationarity assumptions central to most blind methods. Third, detector noise is non-Gaussian and correlated across instruments, breaking the noise model these techniques rely on. The conclusion is that none of these techniques in their vanilla form is expected to outperform existing coherent burst pipelines on the 3G overlap problem.

An alternative is to keep the coherent likelihood framework and relax the single-source assumption built into clustering and reconstruction. Different sources generally correspond to distinct coherent-response subspaces of the network, so multi-source attribution remains possible even when individual source localisation is poor. Chapter 5 takes this route: it extends the per-pixel coherent statistic to a joint hypothesis. CCSN searches are background-dominated and typically supplemented by a multimessenger trigger (e.g. a neutrino detection within a narrow window around core collapse) [8]. This thesis does not develop that aspect, and false-alarm calibration for the multi-source framework is left to future work.

Chapter 3

Quantifying CBC–CCSN Overlap

Coherent burst searches like cWB [5] and the MDAP [6] detect gravitational-wave events by clustering excess-power pixels in the time–frequency (TF) plane before performing any source-specific reconstruction. This approach implicitly assumes that all clustered pixels originate from a single source. If a Galactic core-collapse supernova (CCSN) signal occurs while a compact binary coalescence (CBC) inspiral is already present in the detector band, their TF footprints may overlap, causing a burst algorithm to merge pixels from two distinct sources into one cluster. This chapter quantifies how often and to what extent this happens for the Einstein Telescope (ET).

The analysis thresholds both signal types into binary TF masks, assembles a day-long CBC mask from a realistic population, and slides the CCSN mask across it to build overlap statistics under two TF representations: the short-time Fourier transform (STFT) and the discrete wavelet transform (DWT).

3.1 Signal models

3.1.1 CBC signals

BNS. A one-day BNS population drawn from a realistic one-year BNS catalog [15], simulated with `gwmock` [16] using the IMRPhenomPv2_NRTidalv2 waveform model [17, 18], with each event at its own sky position and source parameters drawn from the population. Chapter 5's sky-localisation tests reuse the BNS waveform model.

BBH. A one-day BBH population drawn from a realistic one-year BBH catalog [15], simulated with `gwmock` [16] using the IMRPhenomXPHM waveform model [19]. BBH signals are shorter and merge at higher frequencies than BNS. Sky-localisation tests use a simpler aligned-spin waveform model IMRPhenomD [20, 21].

Caveat on the BBH population. The BBH catalog file used here was found, during the analysis, to underrepresent the realistic BBH event rate by roughly an order of magnitude. Curve shapes remain valid, but absolute probabilities are biased low. Since the true BBH rate exceeds the BNS rate, a correctly populated catalog would yield overlap that meets or exceeds the BNS values. BBH figures are therefore collected in Appendix A.3 and results are discussed qualitatively here.

3.1.2 CCSN waveforms

Five CCSN sources are studied, spanning two explosion mechanisms: the standard neutrino-driven case and the exotic magnetorotational case.

Mezzacappa 2023 (neutrino-driven, 3D) [22]. Three progenitor masses, $9.6\,M_\odot$ (D9.6), $15\,M_\odot$ (D15) and $25\,M_\odot$ (D25), including core bounce, prompt convection, SASI and neutrino-driven convection [23].

Powell (magnetorotational) [24]. A $39\,M_\odot$ progenitor (m39) at $B = 10^{10}$ G and $B = 10^{12}$ G.

Both sets are taken at two viewing directions, equatorial (X-axis, $\theta = \pi/2$) and polar (Z-axis, $\theta = 0$).

All CCSN signals are injected at 10 kpc, a standard reference distance for Galactic events and sampled at $f_s = 4096$ Hz. The D15-3D waveform also serves as the CCSN companion in the multi-source sky-localisation tests of Chapter 5.

3.2 Temporal overlap: a first check

A lightweight check first establishes temporal overlap as non-negligible. This serves as an upper bound on TF-overlap. The check uses the one-day CBC strain at a lower sampling rate. The CCSN is modelled as a plain time interval of duration T_{CCSN} slid across the day to calculate the probability of overlap if it were randomly placed in the day. The overlap probability is mapped over T_{CCSN} (1–10 s) and the CBC waveform starting frequency $f_{\text{CBC}}^{\text{min}}$ (2–30 Hz), the low-frequency cutoff at which the inspiral enters the detector band, shown for BNS in Figure 3.1.

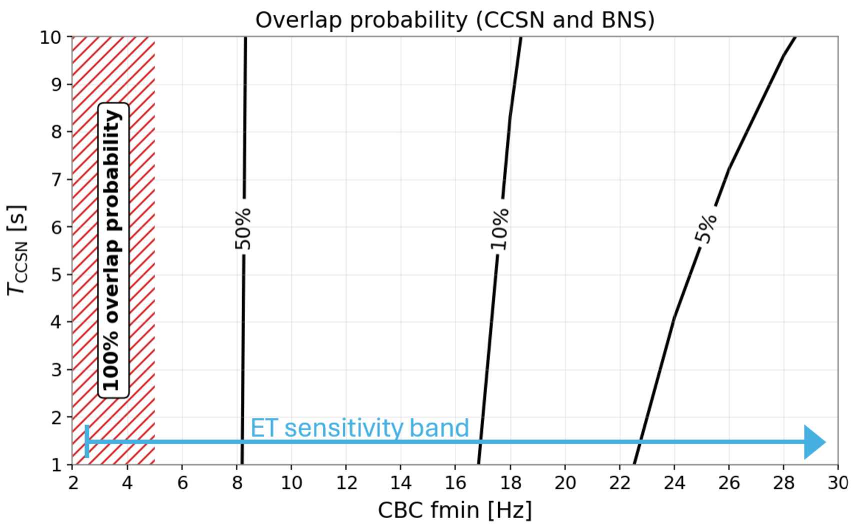

Figure 3.1: Temporal overlap probability between a CCSN time interval of duration T_{CCSN} and a one-day BNS population, as a function of $f_{\text{CBC}}^{\text{min}}$. The hatched band marks 100% overlap at the low frequencies (∼3 Hz) ET is designed to reach [1]. Because Galactic CCSN are rare (∼one per 60 years [4]), even the 5–10% contours are relevant and lie within the band of any current or planned detector.

At the low frequencies around 3 Hz that ET is designed to push for [1], CCSN–BNS temporal overlap is 100%: any CCSN within a day-long observation lands on top of a BNS signal. Since Galactic CCSN are rare, even the lower 5–10% overlap regime is already relevant, and Figure 3.1 shows it falls well inside the sensitivity band of every current or planned detector. BBH inspirals spend even more time in band, so the BBH overlap is expected to be at least as high, and continued investigation of both populations is warranted.

3.3 Time-frequency representation and thresholding

3.3.1 STFT

A pixel at (f, t) is declared active if its normalized squared magnitude exceeds threshold σ^2:

$$\rho^2(f,t) \equiv \frac{|Z(f,t)|^2}{S_n(f) \cdot \Delta f} > \sigma^2, \tag{3.1}$$

where $S_n(f)$ is the one-sided ET-D power spectral density [25] and Δf the frequency bin width. The CBC and CCSN STFTs use resolutions tuned to each waveform: $\Delta f = 4\,\mathrm{Hz}$, $\Delta t = 0.25\,\mathrm{s}$ for the slowly evolving CBC inspiral, and $\Delta f = 16\,\mathrm{Hz}$, $\Delta t = 0.0625\,\mathrm{s}$ for the sharper CCSN features. Both masks are then mapped onto the finest common grid by splitting each coarser pixel into its sub-pixels.

At $\Delta f = 16\,\mathrm{Hz}$ the CCSN has no bins between 0 and 16 Hz, so for this bin the CCSN is recomputed on the CBC grid, recovering dedicated bins at 4, 8 and 12 Hz that match the ET 4 Hz design floor [25].

3.3.2 Wavelet transform

The DWT is included as a second representation. Its dyadic tiling resolves slow low-frequency CBC inspirals and fast high-frequency merger in the same decomposition. Following Mannaert [6], the sym14 wavelet is used for its compaction and boundary behaviour. A pixel (l, k) is declared active if

$$\rho_{l,k}^2 \equiv \frac{c_{l,k}^2}{S_n(f_l^c) \cdot f_s/2} > \sigma^2. \tag{3.2}$$

3.3.3 Threshold choice

Both normalisations are chosen so that $\mathbb{E}[\rho^2] = 1$ in pure noise. A threshold of $\sigma = 1$ is applied consistently to both CBC and CCSN throughout, so pixels at or above the noise floor are considered part of the signal and thus capable of overlapping. Examples of the resulting masks are shown in Figure 3.2.

3.4 Sliding-window overlap statistic and CCSN mask structure

The CCSN mask is slid across it one time bin at a time, as illustrated for one of the CCSN in Figure 3.2. Each shift position represents one possible placement of the CCSN in time, so the fraction of positions with overlap directly estimates the probability of a random coincidence. For STFT the step is the common fine-grid step. For DWT it is the coarsest wavelet bin.

At each shift position s, the overlap fraction $\eta(s)$ is the fraction of CCSN active pixels that overlap the day mask. Averaging the indicator $I(s) = \mathbf{1}[\eta(s) > 0]$ over all N_{shift} positions gives the overlap probability P_{ov}.

To localise the overlap in frequency, overlap below a minimum frequency f_{min} is dropped and $P_{\text{ov}}(f_{\text{min}})$ is plotted. Per-interferometer and per-viewing-angle statistics are averaged into one curve per CCSN source.

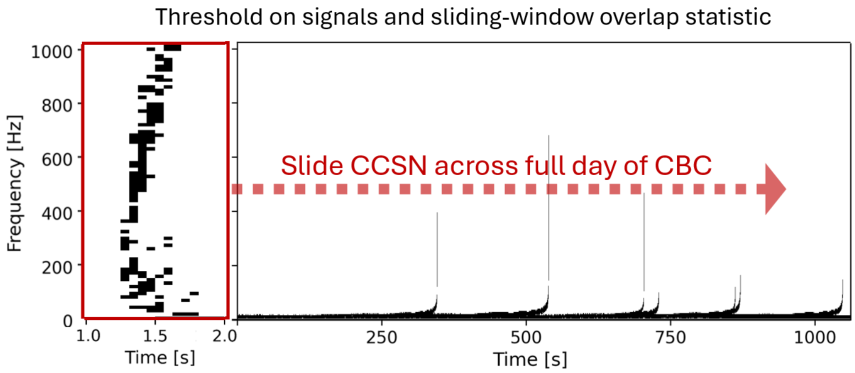

Figure 3.2: Post-threshold TF masks ($\sigma = 1$, ET-D, STFT) and the sliding-window construction. *Left*: CCSN mask (D15 shown as example, 10 kpc, equatorial viewing direction). *Right*: BNS day mask; the CCSN mask is slid across it one time bin at a time and each shift position yields an overlap fraction $\eta(s)$.

3.5 Results

Figure 3.3 shows $P_{\text{ov}}(f_{\text{min}})$ for the BNS day mask. The conditional η-distribution is in Appendix A.2, Figure A.2. At full ET bandwidth most models reach $P_{\text{ov}} \approx 87$–91%. D9.6's lower P_{ov} is a detector-sensitivity effect: one ET interferometer registers too few low-frequency CCSN pixels at $\sigma = 1$, pulling the per-interferometer average down. The mask is also sparse, selecting only ≈ 10 pixels, so coincident pixels make up a larger fraction of a small footprint, inflating η.

Above 64 Hz, P_{ov} drops to near zero, since only BNS mergers and ringdowns contribute at high frequencies, and these are short so the day mask is sparse and a coincident CCSN pixel is unlikely. When overlap does occur, η is typically 1–5% of CCSN pixels for D15, D25 and the Powell models (Appendix A.2). Combined with the sharp drop of P_{ov} once the sub-16 Hz bands are excluded, this places overlap firmly in the low-frequency regime, highly likely whenever a CCSN occurs but confined to a small but energy-dense fraction of the CCSN footprint. The hybrid STFT and the DWT agree qualitatively and quantitatively.

The BBH day mask reproduces the same qualitative TF structure as the BNS case (near-zero above 64 Hz). Absolute overlap is lower, but with the true BBH rate it would meet or exceed the BNS values.

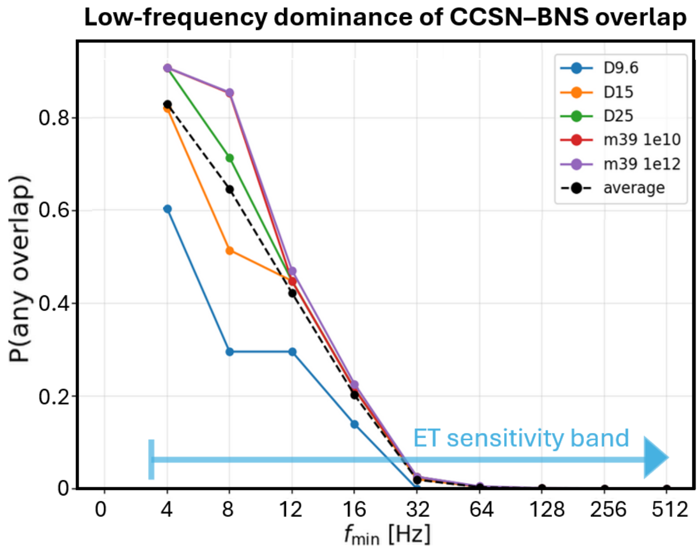

FIGURE 3.3: CCSN–BNS overlap probability vs. f_{min} (hybrid STFT, $\Delta f = 4/16$ Hz, $\sigma = 1$, five source curves and their mean (dashed), averaged over the three ET interferometers and two viewing orientations). At full bandwidth $P_{\text{ov}} \approx 87$–91% for D15, D25 and the Powell models ($\approx 55\%$ for D9.6), dropping sharply once the sub-16 Hz bands are excluded: overlap is highly likely whenever a CCSN occurs and is confined to the low-frequency regime.

3.6 Conclusion

CCSN–CBC overlap is highly likely whenever a Galactic CCSN occurs during ET operation: temporal overlap reaches 100% at ET-relevant low frequencies, and TF overlap reaches $P_{\text{ov}} \approx 87$–91% for the heavier CCSN progenitors, concentrated below 16 Hz, precisely the regime ET is being optimised for. The conditional fraction is small (2–10% of CCSN pixels), so each coincidence touches only a minor part of the CCSN footprint, but this part coincides with the low-frequency band where the CCSN memory effect carries much of the signal energy. The high probability means an overlapping CCSN cannot be dismissed as an unlikely scenario. The BBH catalog used here is underpopulated, so with the true rate BBH overlap would meet or exceed BNS and the BNS result is a lower bound for both populations. If CBC pixels merge with CCSN pixels during clustering, the resulting candidate no longer reflects the CCSN signal alone, potentially biasing detection and reconstruction. This motivates the development of clustering strategies that can handle overlapping astrophysical sources, addressed in the next chapters.

Chapter 4

Modifications to the Modular Data Analysis Pipeline

The Modular Data Analysis Pipeline (MDAP) developed by Mannaert [6] implements a coherent burst analysis chain for the Einstein Telescope through six stages: data generation, time–frequency transform, likelihood generation, clustering, event detection, and event reconstruction.

This chapter describes the modifications required to identify and separately reconstruct N overlapping gravitational-wave signals from a geographically separated detector network. Two limitations of the original pipeline must be addressed before multi-source analysis becomes possible:

Single-hypothesis data flow. As established in Chapter 3, the time–frequency footprints of CBC and CCSN signals overlap. The original pipeline is single-hypothesis throughout. The new implementation must propagate N distinct hypotheses through every stage downstream of the time–frequency transform, indexed by source $i \in \{1, \ldots, N\}$.

Sky-position degeneracy of the co-located ET network. The three ET interferometers are arranged with 60° angles at a common site. As follows from the triangular triple-Michelson null-stream relation of Freise et al. [26], the three antenna-pattern vectors sum to zero:

$$\sum_{k=1}^{3} F_{+}^{(k)}(\Omega) \;=\; \sum_{k=1}^{3} F_{\times}^{(k)}(\Omega) \;=\; 0 \qquad \forall\, \Omega. \tag{4.1}$$

The three detector streams are therefore spanned by two independent signal combinations plus a null combination that carries no gravitational-wave content from any direction. With only two effective streams constraining two polarization unknowns the system is exactly solvable for $(h_+, h_\times)$ at every sky direction, the residual vanishes, and the coherent likelihood is identical across the sky. Empirically, LTF maps as computed by Mannaert [6] at widely separated sky positions for co-located ET data are indistinguishable. The consequence for the multi-signal extension is that no sky-based separation of overlapping signals is possible with ET alone.

4.1 Pipeline architecture

The modified pipeline replaces the single-hypothesis data flow with one indexed by a per-source descriptor. Figure 4.1 shows the architecture. The time-domain detector strain and the time–frequency coefficients are unchanged. Starting from the sky-localisation stage, every downstream data product is a list of length N.

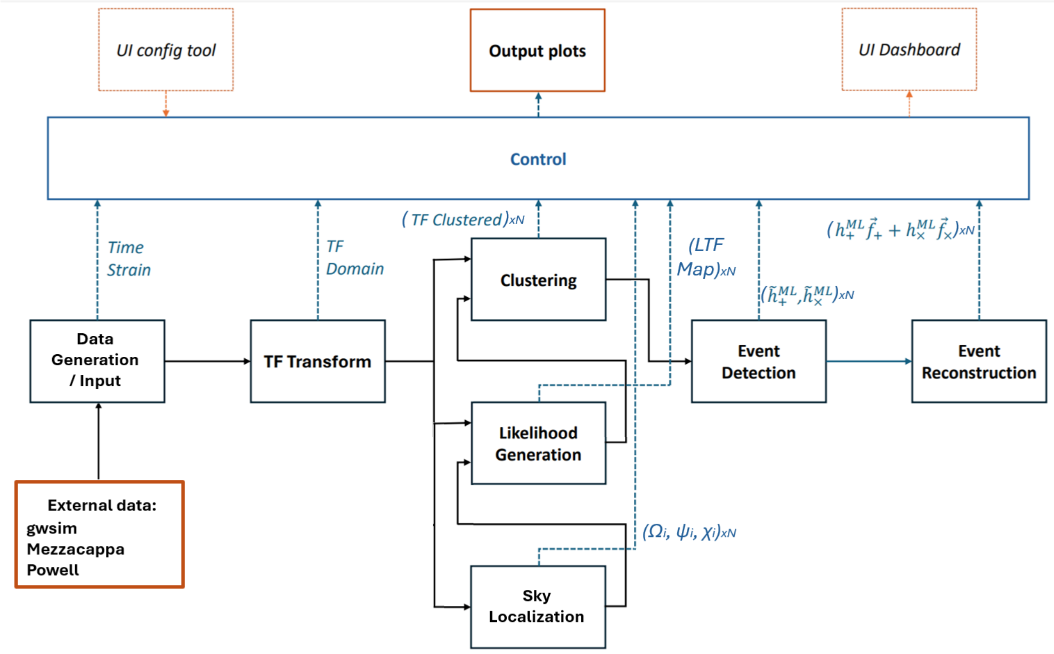

FIGURE 4.1: Modified pipeline architecture, adapted from [6]. Data products to the left of Sky Localisation are single objects (time-domain strain, TF coefficients). From Sky Localisation onwards the pipeline is N-indexed: each source i carries a descriptor $(\Omega_i, \psi_i, \chi_i, \text{model\_type}_i)$ that feeds the unified Likelihood Generation and the Clustering stage. Likelihood Generation emits per-source coherent maps $\{L_i\}$ and a per-pixel single/multi-source classification array ℓ; Clustering converts these into N pixel masks $\{\mathcal{C}_i\}$ along with the per-detector clustered TF coefficients. Event Detection and Reconstruction run the $\mathrm{K} \times 2$ ML solver once per source, using each source's pixel mask and sky position (Section 4.2.5).

4.1.1 Per-source descriptor

Each source i is described by a dictionary carrying

$$\left(\text{ra}_i,\ \text{dec}_i,\ t_{\text{geo},i},\ \psi_i,\ \chi_i,\ \text{model\_type}_i\right).$$

The first three entries fix the antenna pattern and inter-detector delays. The polarization parameters (ψ_i, χ_i) and the model type fix the source's signal model. Two values of model_type are supported:

1D-sources. The signal lies on the one-dimensional space spanned by the effective steering vector

$$a_{\text{eff}}^{(k)}(\Omega,\psi,\chi) = \left[\cos 2\psi - i\chi \sin 2\psi\right] F_{+}^{(k)}(\Omega) + \left[\sin 2\psi + i\chi \cos 2\psi\right] F_{\times}^{(k)}(\Omega), \tag{4.2}$$

parametrized by $\psi \in [0, \pi/2)$ and $\chi \in [-1, 1]$. Pure circular polarization corresponds to $\chi = \pm 1$ (in which case ψ is degenerate). The per-source response column at each detector is 1-dimensional. This model is used for CBC sources, with (ψ, χ) either fixed at $(0, 1)$ in a circular search or estimated by the localiser (Chapter 5).

2D-general sources. The two polarization components $(h_+, h_\times)$ have no fixed phase relation, so the per-source signal subspace is the full 2-dimensional space spanned by the antenna-pattern columns $\{F_+, F_\times\}$. This model is used for CCSN signals, where the polarization structure is not known a priori.

The descriptor is written once by the sky-localisation stage (Section 4.2.2) and used by every downstream module that needs to know the source's signal model. The clustering and event-detection stages need only the source's position. The likelihood stage additionally needs model_type and, for 1D-sources, (ψ, χ).

4.1.2 Detector network

Geographically separated detectors break the ET sky-position degeneracy through time-of-flight delays $\tau_k(\hat{n})$ (Equation 5.1, Chapter 5). The differential delay $\Delta\tau_{ij}$ introduces a frequency-dependent inter-site phase that depends strongly on Ω and cannot be absorbed into ψ. This is the physical basis on which the sky-localisation stage of the modified pipeline rests.

The pipeline supports an arbitrary network of detectors from the Bilby interferometer library [27]. The default configuration is the ET+CE+A1 3G network introduced in Section 2.3, providing $K_{\text{eff}} = 4$ independent gravitational-wave streams (two from co-located ET, one each from CE and A1). Time-of-flight delays are applied via the frequency-domain phase shift $\tilde{s}_k(f) = \tilde{s}_{\text{geo}}(f)\, e^{-i\,2\pi f\,\tau_k}$ followed by an inverse FFT. An integer-rounded sample count $n_\tau = \text{round}(\tau_k f_s)$ is used in parallel only to zero the wrap-around region from the periodic FFT.

4.2 Module implementations

Each new module follows the same three-class `SendData` → `Stage` → `GetData` adapter pattern used throughout the codebase, where `SendData` is the public interface invoked by the central `OutputControl`, `Stage` orchestrates the algorithmic work, and `GetData` fetches the upstream inputs.

4.2.1 Data input and multi-signal injection

A dedicated data-input module reads externally simulated signals and exposes them through the same interface as the internal Bilby waveform generator, so the downstream stages do not need to know which path produced the data. Two CCSN waveform families are supported: Mezzacappa [22] (source-frame $D \cdot h_+, D \cdot h_\times$ at known distance) and Powell [24] (3×3 strain tensor rotated to the observer frame). CBC strain time series from external simulators (e.g. `gwmock` [16]) are read and resampled if needed. The multi-signal extension allows injecting any combination of source types into a common detector data stream. For each signal s the geocenter-frame polarizations $h_+^{(s)}, h_\times^{(s)}$ are projected onto each detector using its antenna patterns at the source's sky position and shifted by the per-detector delay $\tau_k^{(s)}$, then summed:

$$d_k(t) = \sum_{s=1}^{N} \Big[F_+^{(k)}(\Omega_s)\, h_+^{(s)}(t) + F_\times^{(k)}(\Omega_s)\, h_\times^{(s)}(t)\Big]\big(t - \tau_k^{(s)}\big) + n_k(t). \tag{4.3}$$

Both noiseless and per-detector PSD-coloured Gaussian noise runs are supported.

4.2.2 Sky localisation

The sky-localisation stage is where the N-hypothesis structure emerges. The methods to localise sources are developed and verified in Chapter 5.

Two configuration modes are exposed:

- *Blind*: This is the case where the source positions are unknown. N candidate positions are determined and written into the per-source descriptor list.
- *Known*: the per-source descriptor is populated directly and the localisation is skipped. This is useful for both validation of later modules and realistic cases where external multimessenger information, e.g. a neutrino detection localising a CCSN [8], provides the sky position.

In addition to $(\mathrm{ra}_i, \mathrm{dec}_i)$, the sky-localisation stage writes $(\psi_i, \chi_i, \mathrm{model\_type}_i)$. Methods that estimate (ψ, χ) jointly with sky position write the recovered values for the CBC source and assign the 2D-general model to the CCSN. Methods that do not estimate polarization default to 2D-general for all sources.

4.2.3 Likelihood generation

The likelihood stage evaluates every time–frequency pixel against each of the N source hypotheses. The coherent sky statistic and its multi-source extension are derived in Chapter 5. What is new here is that the computation is applied uniformly to all pixels in the TF plane, using the per-source descriptor to set the signal model. The output for each pixel (f, t) is:

- N coherent-energy scalars $L_i(f, t)$, one per source, giving the data's consistency with source i's signal model at that pixel.
- A label $\ell(f, t) \in \{1, \ldots, N, \text{multi}\}$ indicating which source best explains the pixel, or that it is multi-source.

No energy threshold is applied at this stage. Every pixel is evaluated and passed to clustering.

4.2.4 Clustering

Clustering takes the N per-source likelihood maps $\{L_i\}$ and per-pixel labels ℓ from the likelihood stage and produces N pixel masks $\{\mathcal{C}_i\}$ together with the per-detector TF coefficients restricted to each mask.

The default strategy applies the standard cWB core+halo construction [28, 6] independently to each per-source map L_i, thresholding at a per-map percentile and dilating the resulting core pixels in time. Clusters are independent and may overlap: a pixel in exactly one cluster is a single-source pixel, a pixel in two or more is a multi-source pixel, the rest are background. This structural multi-source category is consistent with the likelihood-stage label $\ell(p) = \text{multi}$ when both thresholds are calibrated consistently.

4.2.5 Event detection and reconstruction

Given the per-source pixel masks from clustering, event detection produces N independent maximum-likelihood polarization estimates $\{\hat{h}_+^{(i)}, \hat{h}_\times^{(i)}\}$. Reconstruction then applies the inverse time–frequency transform to each set independently, yielding N reconstructed waveforms.

The ML solve distinguishes two pixel types. At a pixel assigned to exactly one source i, the standard $\mathrm{K} \times 2$ Cramer's-rule solve from [6] is applied using source i's antenna pattern. At overlap pixels attributed to two or more sources by independent masks, a joint $\mathrm{K} \times 2N$ least-squares solve is used instead: the noise-scaled antenna-pattern columns of all active sources are stacked into a single $K \times 2N$ response matrix A and the full polarization vector $[h_+^{(1)}, h_\times^{(1)}, \ldots, h_+^{(N)}, h_\times^{(N)}]^T$ is recovered in one step via $\hat{h} = (A^H A)^{-1} A^H d$. For the ET+CE+A1 network with $K_{\text{eff}} = 4$ and $N = 2$ the system is exactly determined.

Figure 4.2 shows a representative result for a simultaneous BBH and CCSN injection for ET1 at fairly high signal-to-noise ratio. The BBH chirp is recovered, the reconstructed waveform follows the injected signal. The CCSN reconstruction contains noticeably more noise. This is expected and reflects the nature of the source rather than a fundamental limitation of the framework. CCSN gravitational-wave emission is stochastic and spectrally broad, spreading signal power across a large number of time–frequency pixels. This diffuse footprint leads to extensive overlap with noise pixels and with the BBH cluster, making clean source separation harder. Clustering was not the primary focus of this work. The implementation presented here leaves substantial room for optimisation. Improved clustering rules tailored to extended bursts would reduce contamination and improve the CCSN reconstruction quality, which could in turn also be applied to benefit sky localisation for that source. Within these constraints the result still demonstrates that the MDAP pipeline is operationally ready for multi-signal analysis. Both sources are identified, separated, and reconstructed in a single pipeline pass, which is the primary goal of the modifications described in this chapter.

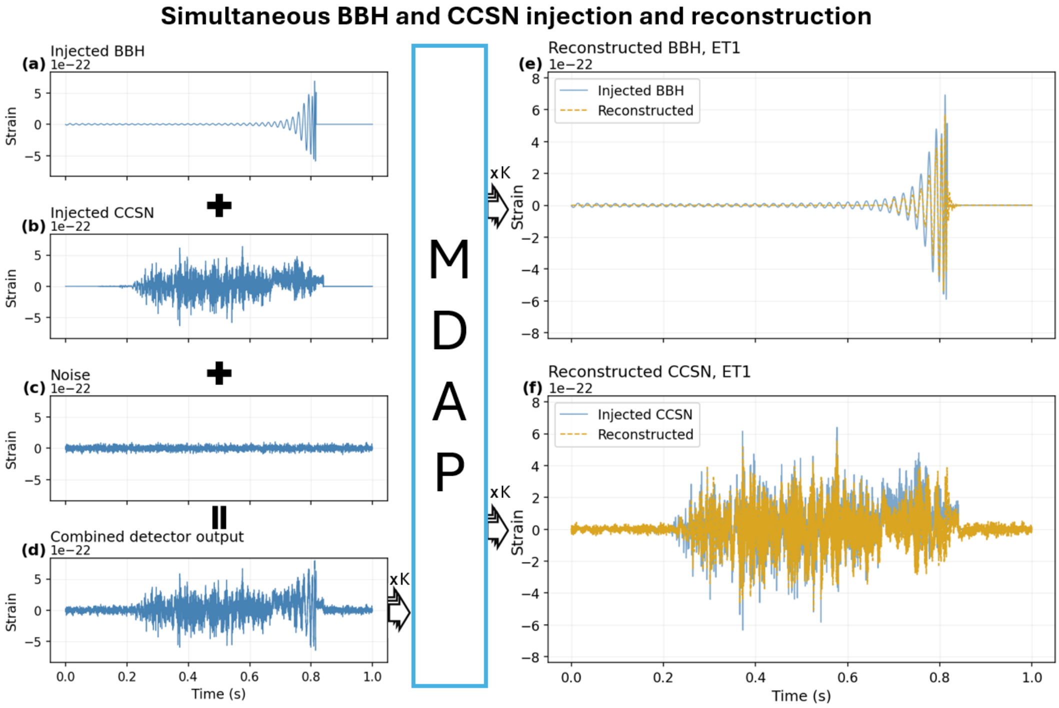

FIGURE 4.2: Simultaneous BBH and CCSN injection and reconstruction for one of K detectors (ET1) with realistic ET-D noise. *Left column:* injected signal components: BBH (a), CCSN (b), noise (c), and the combined strain (d). *Right column:* per-source reconstruction overlaid on the injected signal for BBH (e) and CCSN (f).

4.3 Conclusion

The modifications described in this chapter convert the original single-signal, single-site pipeline into a framework that propagates a per-source descriptor from sky localisation through every downstream module. A unified, model-aware likelihood stage produces N coherent maps and a per-pixel single-/multi-source classification, a multi-signal clustering stage converts these into N pixel masks, and event detection and reconstruction return N polarization estimates and the corresponding waveforms in a single pipeline pass. A geographically separated ET+CE+A1 detector network makes the sky-based separation of overlapping signals possible in the first place.

The implementation is a proof of concept rather than an optimally-tuned pipeline. Each new stage works end-to-end on multi-source injections, but individual choices in clustering rules, threshold calibrations, and model-type assignment leave substantial room for refinement. Clustering rules tailored to extended bursts in particular would reduce contamination of the CCSN cluster, improving both reconstruction quality and sky localisation. The sky-localisation methods that populate the per-source descriptor are developed in Chapter 5.

Chapter 5

Sky Localisation

This chapter develops the sky localisation module of the multi-signal pipeline. The analysis operates at two levels. At the *pixel level* (Section 5.1), individual time–frequency bins are assigned either a single-source coherent statistic or a joint maximum-likelihood statistic depending on whether one or two signals are present. At the segment level (Section 5.2), these per-pixel statistics are accumulated into a joint sky map for both sources simultaneously. Section 5.3 validates the methods and Section 5.4 summarizes the chapter.

5.1 Pixel-Level Localisation

The fundamental question at the pixel level is: given the data at a single time–frequency bin, which sky position or positions maximize the coherent energy? Three building blocks answer this: a single-source coherent statistic (Section 5.1.1), a joint maximum-likelihood statistic for pixels where two signals overlap (Section 5.1.2), and a null-energy classification criterion that selects between them (Section 5.1.3).

5.1.1 Single-Source Coherent Sky Map

The coherent network method phase-aligns all detectors to a common geocentric reference frame, compensating for the gravitational-wave time-of-flight delays introduced in Section 4.1.2, and projects the combined data onto the two-dimensional polarization subspace. For whitened Gaussian noise, this projection is the maximum-likelihood estimator of $(\tilde{h}_+, \tilde{h}_\times)$. It is the approach used in coherent WaveBurst [5] and X-Pipeline [29] and the primary single-source sky localisation method here.

A GW from sky direction $\hat{n}$ arrives at detector k at time

$$t_k = t_{\mathrm{geo}} + \frac{\mathbf{x}_k \cdot \hat{n}}{c}, \tag{5.1}$$

where t_{geo} is the geocentric arrival time and $\mathbf{x}_k$ is the detector position in Earth-Centred Earth-Fixed (ECEF) coordinates [30]. The differential delay between detectors i and j is $\Delta\tau_{ij}(\hat{n}) = \mathbf{b}_{ij} \cdot \hat{n}/c$, with baseline vector $\mathbf{b}_{ij} = \mathbf{x}_i - \mathbf{x}_j$, bounded by the light-travel time $|\mathbf{b}_{ij}|/c$ across the baseline. The sky position is parameterized by right ascension α and declination δ, the ECEF unit vector is

$$\hat{n}_{\mathrm{ECEF}} = \begin{pmatrix} \cos\delta\cos(\alpha-\theta_G) \\ \cos\delta\sin(\alpha-\theta_G) \\ \sin\delta \end{pmatrix}, \tag{5.2}$$

where θ_G is the Greenwich Mean Sidereal Time at the geocentric GPS time t_{geo}.

Antenna pattern functions. An interferometric detector responds to the two polarization states h_+ and $h_\times$ through the beam pattern functions [31]

$$F_+(\hat{n},\psi) = \tfrac{1}{2}(\hat{u}^a\hat{u}^b - \hat{v}^a\hat{v}^b)\, e^{+}_{ab}(\hat{n},\psi), \qquad F_\times(\hat{n},\psi) = \tfrac{1}{2}(\hat{u}^a\hat{u}^b - \hat{v}^a\hat{v}^b)\, e^{\times}_{ab}(\hat{n},\psi), \tag{5.3}$$

where $\hat{u}$, $\hat{v}$ are the detector arm unit vectors, $e^{+,\times}_{ab}$ are the polarization tensors, and ψ is the polarization angle that orients the $(+,\times)$ basis in the plane transverse to the propagation direction $\hat{n}$. In the frequency domain, the detector output is

$$\tilde{d}_k(f) = \left[F_+^{(k)}\tilde{h}_+(f) + F_\times^{(k)}\tilde{h}_\times(f)\right]e^{-i2\pi f t_k} + \tilde{n}_k(f). \tag{5.4}$$

The noise-normalized response vectors collect these pattern functions across all detectors, weighted by the detector ASD:

$$\tilde{\mathbf{F}}_+(f) = \left(\frac{F_+^{(1)}}{\mathrm{ASD}_1(f)}, \ldots, \frac{F_+^{(K)}}{\mathrm{ASD}_K(f)}\right)^\top, \quad \tilde{\mathbf{F}}_\times(f) = \left(\frac{F_\times^{(1)}}{\mathrm{ASD}_1(f)}, \ldots, \frac{F_\times^{(K)}}{\mathrm{ASD}_K(f)}\right)^\top. \tag{5.5}$$

The sky-dependent sensitivity of each detector and of the full network can be characterised by $S(\hat{n}) = \sqrt{\sum_k \int (F_+^{(k)\,2} + F_\times^{(k)\,2})/\mathrm{ASD}_k^2\,\mathrm{d}f}$ shown in Fig. 5.1.

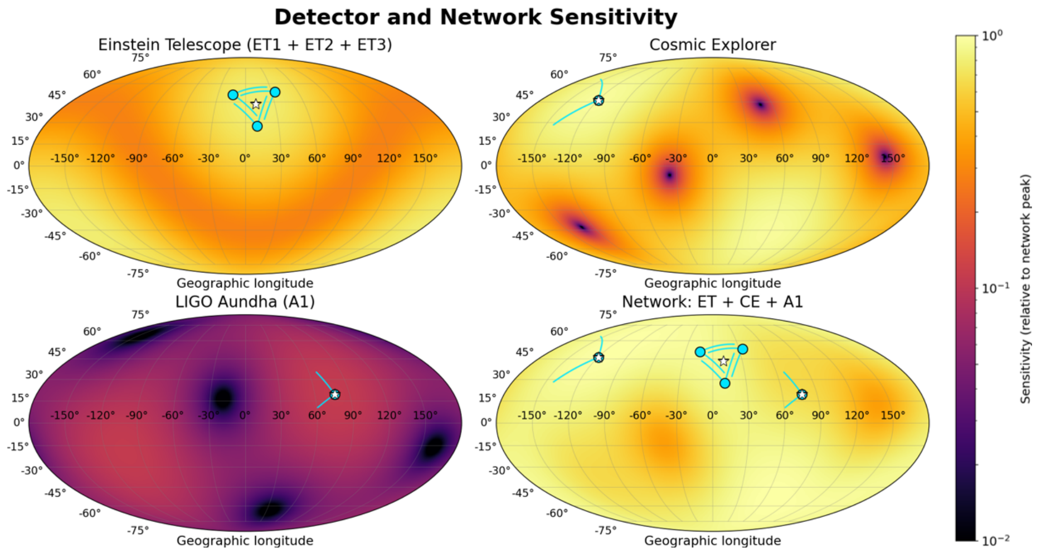

FIGURE 5.1: Band-integrated antenna sensitivity $S(\hat{n})$ (10–2048 Hz).

Stars mark detector sites, cyan lines show arm orientations. CE dominates the sensitivity, ET is comparable, and A1 (4 km arms) contributes at the $\sim$10% level.

Maximum-likelihood projection. Equation 5.4 shows that a signal from $\hat{n}$ enters the K-detector data only through the two antenna patterns $F_+^{(k)}, F_\times^{(k)}$. After phase-alignment, the signal therefore lives on a two-dimensional plane in the whitened detector space, spanned by $\tilde{\mathbf{F}}_+$ and $\tilde{\mathbf{F}}_\times$ (Eq. 5.5). Whitened detector noise is independent complex Gaussian, so the log-likelihood reduces to $-\|\tilde{\mathbf{D}} - \tilde{\mathbf{F}}_+\tilde{h}_+ - \tilde{\mathbf{F}}_\times\tilde{h}_\times\|^2$. Minimising this over $(\tilde{h}_+, \tilde{h}_\times)$ is a least-squares fit, whose solution is the orthogonal projection of $\tilde{\mathbf{D}}$ onto the polarization plane. The next paragraph carries out that projection explicitly.

Polarization basis and coherent beamforming. The detector STFT coefficients $S_k(f,t)$ are phase-aligned for a trial direction $\hat{n}$ and whitened by the ASD:

$$\tilde{\mathbf{D}}(f,t;\hat{n}) = \left(\tilde{D}_1, \ldots, \tilde{D}_K\right)^\top, \qquad \tilde{D}_k(f,t;\hat{n}) = \frac{S_k(f,t)}{\mathrm{ASD}_k(f)}\, e^{+i2\pi f\,\tau_k(\hat{n})}, \tag{5.6}$$

with $\tau_k(\hat{n}) = \mathbf{x}_k \cdot \hat{n}/c$ from Eq. 5.1. At a wrong trial direction the phase-aligned data is inconsistent with the signal model: destructive interference suppresses the projection onto the polarization plane, and residual energy spills into the null directions. The vectors $\tilde{\mathbf{F}}_+$ and $\tilde{\mathbf{F}}_\times$ (Eq. 5.5) span the two-dimensional polarization subspace of the whitened detector space, with each detector weighted inversely by its noise floor. Gram–Schmidt orthonormalization gives the basis $(\hat{\mathbf{e}}_+, \hat{\mathbf{e}}_\times)$, the coherent signal amplitudes are the projections of the phase-aligned data onto it, and the coherent energy per time–frequency bin is their summed squared magnitudes:

$$\begin{gathered}
\hat{\mathbf{e}}_+ = \frac{\tilde{\mathbf{F}}_+}{|\tilde{\mathbf{F}}_+|}, \quad \hat{\mathbf{e}}_\times = \frac{\tilde{\mathbf{F}}_\times - (\hat{\mathbf{e}}_+ \cdot \tilde{\mathbf{F}}_\times)\hat{\mathbf{e}}_+}{|\tilde{\mathbf{F}}_\times - (\hat{\mathbf{e}}_+ \cdot \tilde{\mathbf{F}}_\times)\hat{\mathbf{e}}_+|}, \\
b_+(f,t) = \hat{\mathbf{e}}_+^H\, \tilde{\mathbf{D}}(f,t), \quad b_\times(f,t) = \hat{\mathbf{e}}_\times^H\, \tilde{\mathbf{D}}(f,t), \\
E_{\mathrm{coh}}(f,t) = |b_+(f,t)|^2 + |b_\times(f,t)|^2.
\end{gathered} \tag{5.7}$$

If $\tilde{\mathbf{F}}_\times^\perp$ falls below a tolerance (degenerate sky position where only one polarization direction is accessible), the analysis proceeds with $\hat{\mathbf{e}}_+$ alone.

Null streams. With N detectors and a two-dimensional polarization subspace there are $N-2$ null-stream dimensions [32]. As established in Section 4, ET's three co-located interferometers contribute only two independent streams. Null streams are used implicitly via the null energy $E_{\mathrm{null}}(f,t) = E_{\mathrm{total}}(f,t) - E_{\mathrm{coh}}(f,t)$. At the true sky direction E_{null} stems from noise and is small, while at a wrong direction residual power remains in the null channels.

The coherent network statistic. Following the constrained maximum-likelihood framework of Klimenko et al. [33, 28, 5], the per-pixel sky-localisation statistic is

$$\boxed{\Lambda_{\mathrm{single}}(f,t;\hat{n}) = \frac{E_{\mathrm{coh}}^2(f,t;\hat{n})}{E_{\mathrm{total}}(f,t)}} \tag{5.8}$$

The statistic is dimensionless because the data are whitened. The factor $1/E_{\text{total}}$ weights E_{coh} by the coherent fraction $E_{\text{coh}}/E_{\text{total}} \in [0, 1]$, suppressing pixels dominated by incoherent (glitch-like) power. This places Eq. (5.8) in the cWB family of pixel statistics [33, 28, 5], though it is not algebraically identical to the cWB form of [28], which isolates the cross-detector coherent energy as a separate quality factor $c_c = E_c/(E_c + |E_n|)$. The two behave similarly for sky localisation on the simulated data used here. The cWB statistic is empirically tuned for the 2-L LIGO/Virgo network, so adopting it verbatim is not appropriate for the ET+CE+A1 3G network considered in this thesis.

5.1.2 Multi-Source Joint Maximum-Likelihood

When two signals are present, the single-source statistic of Section 5.1.1 fails in two ways. Evaluating Λ_{single} at the true direction of source 1 misaligns source 2, whose energy leaks into the null streams and suppresses the statistic even at the correct sky position. Using E_{coh} alone avoids this but creates false peaks: a wrong direction that partially aligns both sources can score higher than either true direction. A new formulation is needed that accommodates two sources without null-stream penalisation and without rewarding accidental combined-energy positions.

Statistic derivation and response matrix

Two sources cannot be simultaneously phase-aligned by a single trial direction, so the time-delay phase factor is carried by the steering vectors rather than the data. Let $\tilde{\mathbf{d}}(f, t) \in \mathbb{C}^K$ be the whitened (but not phase-aligned) data vector with components $\tilde{d}_k(f, t) = S_k(f, t)/\text{ASD}_k(f)$. For two sources $i = 1, 2$ at trial sky positions $\hat{n}_1$ and $\hat{n}_2$, the two-polarization signal model for whitened data is

$$\tilde{\mathbf{d}}(f, t) = A_1(\hat{n}_1, f)\,\mathbf{h}_1(f, t) + A_2(\hat{n}_2, f)\,\mathbf{h}_2(f, t) + \mathbf{n}(f, t), \tag{5.9}$$

where $\mathbf{h}_i(f, t) = (h_{+,i}(f, t),\ h_{\times,i}(f, t))^\top$ is the two-polarization vector of source i, and $A_i(\hat{n}_i, f) = [\mathbf{a}_+(\hat{n}_i, f),\ \mathbf{a}_\times(\hat{n}_i, f)]$ is the $K \times 2$ steering matrix whose columns are the whitened, time-delay-loaded antenna responses to the + and × polarizations (explicit form in Eq. (5.13)). Stacking the two per-source matrices gives the $K \times 4$ joint response matrix

$$A(\hat{n}_1, \hat{n}_2, f) = \big[A_1(\hat{n}_1, f),\ A_2(\hat{n}_2, f)\big]. \tag{5.10}$$

The maximum-likelihood estimate of the joint amplitude vector $\mathbf{h} = (\mathbf{h}_1^\top, \mathbf{h}_2^\top)^\top \in \mathbb{C}^4$ is the least-squares solution $\hat{\mathbf{h}}(f, t) = (A^H A)^{-1} A^H \tilde{\mathbf{d}}(f, t)$. The corresponding coherent energy is the squared norm of the projection of $\tilde{\mathbf{d}}$ onto the column space of A:

$$E_{\text{coh}}(\hat{n}_1, \hat{n}_2; f, t) = \tilde{\mathbf{d}}^H(f, t)\, A\, (A^H A)^{-1}\, A^H\, \tilde{\mathbf{d}}(f, t), \tag{5.11}$$

the direct $K \times 4$ analog of the single-source $K \times 2$ coherent energy (Eq. 5.7). The same cWB-style per-pixel statistic as in the single-source case (Eq. (5.8)) is used, now with this two-source E_{coh}:

$$\boxed{\Lambda_{\text{joint}}(\hat{n}_1, \hat{n}_2; f, t) = \frac{E^2_{\text{coh}}(\hat{n}_1, \hat{n}_2; f, t)}{E_{\text{total}}(f, t)}} \tag{5.12}$$

Signal subspace geometry and effective network size

For the most realistic 3G network ET+CE+A1, $K_{\text{eff}} = 4$ (Section 4.1.2). A general source contributes a 2-column response matrix $A_i = [\mathbf{a}_+, \mathbf{a}_\times]$ spanning a 2D signal subspace, so two simultaneous general sources would fully exhaust this signal space $(2 + 2 = 4)$ and render the joint problem degenerate. To restore a noise dimension, each CBC source can be constrained to its elliptical polarization ($\tilde{h}_\times = \pm i\chi\tilde{h}_+$, $\chi \in [0, 1]$), collapsing its 2D subspace onto the 1D space spanned by the effective steering vector

$$\mathbf{a}(\hat{n}, f) = \mathbf{a}_+(\hat{n}, f) \pm i\chi\, \mathbf{a}_\times(\hat{n}, f), \tag{5.13}$$

Face-on ($\theta_{jn} = 0°$) and edge-on ($\theta_{jn} = 90°$) orientations are special cases of Eq. (2.2) of Chapter 2, giving $\chi = 1$ (circular polarization) and $\chi = 0$ (linear polarization) respectively. The circular assumption sets $\chi = 1$ for all CBC sources, treating them as exactly face-on. Table 5.1 shows that on ET+CE+A1 with $K_{\text{eff}} = 4$, at least one source must have its polarization either assumed or estimated from the data, otherwise the joint problem is degenerate. Section 5.1.2 pursues both routes: it validates the circular assumption ($\chi = 1$) and tests whether (ψ, χ) can be estimated directly from the pixel data.

Table 5.1: Recovery outcomes for the primary source in two-source joint ML localisation, with a circular (face-on) secondary. The primary is either face-on ($\theta_{jn} = 0°$) or edge-on ($\theta_{jn} = 90°$). With ET+CE+A1 ($K_{\text{eff}} = 4$), the circular assumption recovers a face-on primary but fails for edge-on (wrong model), while free polarization fails in both cases because the $2 + 2 = 4$ steering columns exhaust the data space and lead to a perfectly flat map. Extending to a large network (ET+CE+A1+H1+L1+NEMO+V1+K1+GEO600) re-opens the noise subspace: circular still recovers only face-on, and free polarization recovers both. Figure 5.2 shows the second and third cases in the table.

Polarization model	Network	Face-on src	Edge-on src
Free	ET+CE+A1	Missed	Missed
Circular	ET+CE+A1	Found	Missed
Free	Large network	Found	Found
Circular	Large network	Found	Missed

Note that the large-network case (Figure 5.2, right panel) recovers the primary source even with free polarization for both signals, since the extra detectors add the noise dimensions needed to absorb the second 2D subspace. The right panel however shows lower contrast between peak and background than the left, because the older detectors contribute only weakly at this SNR. As the SNR drops, exactly the regime where dual-source localisation is hardest, those detectors contribute negligibly and the effective network reduces to ET+CE+A1 with $K_{\text{eff}} = 4$. Restricting the default network to ET+CE+A1 is therefore not a serious compromise.

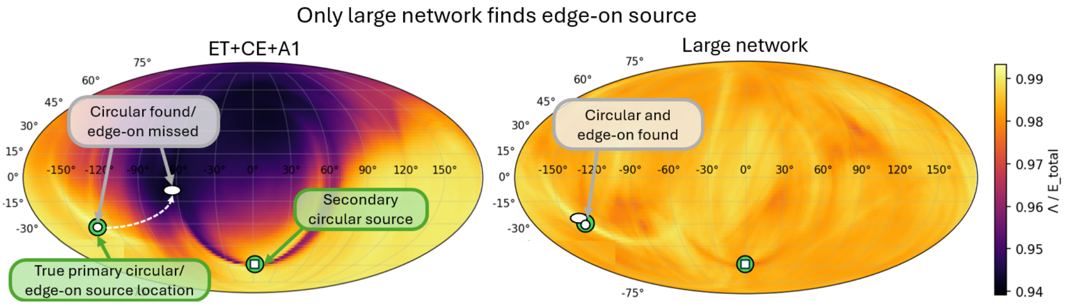

FIGURE 5.2: Marginal sky maps for the primary source in the presence of a circular secondary one, illustrating second and third cases in Table 5.1. *Left*: ET+CE+A1 with the circular assumption recovers a face-on secondary and misses an edge-on one. *Right*: large network with all detectors in Bilby (ET+CE+A1+H1+L1+NEMO+V1+K1+GEO600) without the circular assumption recovers an edge-on secondary as well.

Justifying the Circular Polarization Assumption

CCSN signals do not obey the quadrupole relation [23] and need to be modelled with the full two-column $\mathbf{A} = [\mathbf{a}_+, \mathbf{a}_\times]$, which together with a polarization-constrained CBC occupies $1 + 2 = 3$ of the four available signal dimensions.

For CBCs the polarization constraint collapses the signal onto the one-dimensional space spanned by the effective steering vector

$$\mathbf{a}_{\text{eff}}(\hat{n}, \psi, \chi) = (\cos 2\psi - i\chi \sin 2\psi)\, \mathbf{a}_+(\hat{n}) + (\sin 2\psi + i\chi \cos 2\psi)\, \mathbf{a}_\times(\hat{n}), \tag{5.14}$$

parametrised by $\psi \in [0, \pi/2)$ and $\chi \in [-1, 1]$. At $\psi = 0$ this reduces to Eq. (5.13) from the previous subsection. A single pixel admits a closed-form (ψ, χ) maximizer, solved as a 2×2 generalised eigenvalue problem. A single pixel is too noisy for a useful estimate though, and aggregating across pixels with a shared (ψ, χ) no longer admits a closed form. The implementation therefore maximises it by a brute-force scan over the (ψ, χ) grid. This is sufficient for a single source: every pixel contributes to the same (ψ, χ) likelihood and the maximum is well-defined. For two simultaneous sources the scan must also know which pixels belong to the CBC, which calls for a hard-EM coordinate ascent, the classification EM of Celeux and Govaert [34]. An E-step assigns each pixel to the model with the higher per-pixel likelihood (the same routing used at the segment level, Section 5.2.2), and an M-step re-runs the (ψ, χ) scan on the CBC-assigned pixels. The two steps alternate until the relative likelihood improvement falls below $\epsilon = 10^{-4}$. Iteration is monotone and the best-seen state is retained.

Three treatments are compared: *circular* ($\psi = 0, \chi = 1$), *oracle* (true ψ, χ), and *scan* (Λ grid-maximised at every sky pixel). CBC injections are placed at 100 random sky locations with random ψ across a range of θ_{jn}, on ET+CE+A1 with PSD-coloured detector noise. The multi-source extension adds a Galactic CCSN (Mezzacappa D15-3D [22], $d = 10\,\mathrm{kpc}$). For the single-source CBC, the polarization scan outperforms the circular assumption starting at $\theta_{jn} \approx 60°$, falling in between the circular and the oracle treatments, at distances out to around 10 Gpc for BBH and 1 Gpc for BNS. At twice those distances the (ψ, χ) estimates from the joint maximization become too noisy to be useful, and the scan no longer outperforms the circular assumption even at edge-on. Adding a Galactic CCSN at 10 kpc makes this worse still, possibly because the EM coordinate ascent gets stuck in local maxima of the joint objective. The rest of the thesis therefore uses the circular assumption for CBC signals, although the scan remains fully implemented in the MDAP.

5.1.3 Pixel Classification: Single vs. Multi-Source

Each selected pixel must be routed to the appropriate statistic: the single-source cWB statistic of Section 5.1.1 when one signal dominates, or the joint maximum-likelihood statistic of Section 5.1.2 when two signals overlap. A single-source pixel projects almost entirely onto the 2D polarization subspace at its true sky direction, so its null energy vanishes there. A multi-source pixel retains residual energy in the null channels at every trial direction, because no single 2D subspace can absorb both contributions. The minimum null-energy fraction over the sky is therefore a natural classifier. Subspace methods such as MUSIC were considered and ruled out for the ET+CE+A1 network (the noise subspace is too small to suppress the pseudospectrum at non-source directions).

The null fraction at the pixel's best single-source sky direction gives the classifier

$$\xi(f,t) = \frac{E_{\mathrm{null}}(f,t;\hat{n}^*)}{E_{\mathrm{total}}(f,t)} \begin{cases} < \varepsilon_{\mathrm{null}} & \to \text{single-source pixel,} \\ \geq \varepsilon_{\mathrm{null}} & \to \text{multi-source pixel.} \end{cases} \tag{5.15}$$

The threshold $\varepsilon_{\mathrm{null}}$ is calibrated on a noise-free coincident BNS+BBH injection: ground-truth labels are obtained by running each source separately, and the 90th percentile of the resulting single-source ξ distribution fixes $\varepsilon_{\mathrm{null}} = 0.072$ (Appendix B.1, Figure B.1).

5.2 Segment-Level Integration

The pixel-level methods of Section 5.1 score individual time–frequency bins, segment-level integration accumulates this evidence across the full (f,t) plane into a joint map for two sources.

5.2.1 STFT Windowing

The sky-localisation STFT is computed internally at a resolution tailored to the localisation problem. The window length T_w must satisfy $T_w \gg \tau_{\max}$, where $\tau_{\max} \approx$

37 ms is the longest inter-detector delay (CE–A1). When this holds, the inter-detector delay appears as a smooth phase offset $\exp(-i2\pi f\tau_k)$ within each window rather than a boundary crossing, making the phase alignment of Eq. (5.6) possible.

5.2.2 Per-Pixel Classification and Sky Map Accumulation

Pixel selection. The incoherent power $E_{\text{total}}(f,t) = \sum_k |d_k(f,t)|^2$ is the sum of squared whitened STFT coefficients. Pixels above the noise floor are retained for the sky scan.

Per-pixel sky scan and classification. For each selected pixel the single-source statistic $\Lambda_{\text{single}}(f,t;\hat{n})$ of Eq. (5.8) is evaluated across all sky positions on the grid using precomputed steering vectors. The best direction $\hat{n}^*(f,t)$ and the null fraction $\xi(f,t)$ are recorded, and the pixel is labelled single- or multi-source via Eq. (5.15).

Joint surface accumulation. The joint surface combines contributions from the single-source set $\mathcal{S}$ and the multi-source set $\mathcal{M}$:

$$\Lambda_{\text{total}}(\hat{n}_1,\hat{n}_2) = \sum_{(f,t)\in\mathcal{S}} \max\!\big(\Lambda_{\text{single}}(f,t;\hat{n}_1),\ \Lambda_{\text{single}}(f,t;\hat{n}_2)\big) + \sum_{(f,t)\in\mathcal{M}} \Lambda_{\text{joint}}(f,t;\hat{n}_1,\hat{n}_2), \tag{5.16}$$

with Λ_{joint} the per-pixel joint ML statistic of Eq. (5.12). The $\max(\cdot)$ on the single-source sum attributes the pixel's evidence to the better-fitting trial direction without double-counting. Replacing it with a sum inflates the diagonal $\hat{n}_1 = \hat{n}_2$, and it favours nearby pairs $(\hat{n}_1,\hat{n}_2)$ because pixels well explained by either direction contribute almost identically to both terms.

Multi-source pixels enter via the full joint ML expression. $(A^H A)^{-1}$ becomes singular when the steering vectors lose independence (at $\hat{n}_1 = \hat{n}_2$, or when the CBC direction lies in the CCSN subspace). The limit is well defined but the closed form is $0/0$ on the locus. As a vectorised proxy for a rank-drop, Tikhonov regularization [35] adds $\varepsilon \sim 10^{-6}\bar{g}$ to $A^H A$ (or as a floor on the Schur complement in the mixed-pol case), keeping the surface finite everywhere and recovering the single-source statistic at the locus.

5.2.3 Why Joint Optimization Is Necessary

A CLEAN-type strategy [36]: locate the maximum, subtract the corresponding beam, iterate, would cost only $O(N_{\text{grid}})$ per iteration rather than the $O(N_{\text{grid}}^2)$ of a joint scan, but is structurally unable to localise two overlapping sources, as the following ghost-direction argument shows.

Ghost-direction obstruction. Let the noiseless whitened pixel data be

$$\mathbf{d} = c_1\,\mathbf{a}(\hat{n}_1) + c_2\,\mathbf{a}(\hat{n}_2), \qquad c_1, c_2 \neq 0, \quad \mathbf{a}_1 \not\propto \mathbf{a}_2. \tag{5.17}$$

For a trial direction $\hat{n} \neq \hat{n}_1, \hat{n}_2$, the single-source coherent energy $E_{\mathrm{coh}}(\hat{n}) = |\hat{\mathbf{a}}^H(\hat{n})\mathbf{d}|^2$ obeys Cauchy–Schwarz with equality iff $\hat{\mathbf{a}}(\hat{n}) \propto \mathbf{d}$. The $\hat{n}$ at which this equality holds is neither $\hat{n}_1$ nor $\hat{n}_2$, so a ghost direction is found instead of either true source. In practice the bound is not exactly attained for the 3G network, since the inter-detector phases and magnitudes cannot simultaneously match the mixture at any single $\hat{n}^*$. However, in general, the ghost peak still lies above E_{coh} at the true directions $\hat{n}_1, \hat{n}_2$, where both phase and magnitude are misaligned with the mixture.

A CLEAN iteration therefore picks a ghost first, and subtracting its beam produces a new mixture pointing to a new ghost, never converging to $\hat{n}_1$ or $\hat{n}_2$. A correct algorithm must maintain the joint hypothesis $(\hat{n}_1, \hat{n}_2)$ from the start. The implementation evaluates $\Lambda_{\mathrm{total}}(\hat{n}_1, \hat{n}_2)$ directly on the equal-area grid. Coarse-then-fine alternating coordinate descent was implemented but found to miss low-SNR peaks.

5.2.4 Peak Finding and Marginal Sky Maps

The joint ML estimate is the global maximum of $\Lambda_{\mathrm{total}}(\hat{n}_1, \hat{n}_2)$. The surface is evaluated on an equal-area sky grid with 60 declination strips uniform in $\sin\delta$. Each strip i contains $\max(1, \lfloor 120\cos\delta_i \rceil)$ right-ascension points, so that all cells cover approximately the same solid angle. The total number of sky points is ≈ 5600 by default. Both resolution parameters are freely configurable. No angular-separation mask is needed: the Tikhonov regularization keeps the surface finite at $\hat{n}_1 = \hat{n}_2$, and the max structure of the single-source sum suppresses the diagonal relative to off-diagonal configurations.

Two marginal sky maps are read off at the found peak:

$$M_1(\hat{n}_1) = \Lambda_{\mathrm{total}}(\hat{n}_1, \hat{n}_2^*), \qquad M_2(\hat{n}_2) = \Lambda_{\mathrm{total}}(\hat{n}_1^*, \hat{n}_2). \tag{5.18}$$

Each is a slice of the 4D joint surface with the other source fixed at its found position, giving a per-source sky-map for visualization and credible-region computation. Figure 5.2 shows examples of such marginal maps.

5.2.5 Credible Regions

The STFT coherent statistic Λ is not a normalised posterior, so credible-region sizes cannot be fixed by a probability mass as in the standard HPD approach [37]. Here they are determined empirically from the injection population using greedy pixel sorting.

For each of N injection runs, the sky-map statistic is recorded across the sky ($\Lambda(\hat{n})$ for the single-source map, the marginal M_1 or M_2 from Eq. (5.18) for multi-source maps), and the fractional sky area whose statistic equals or exceeds the value at the true direction is computed. Across the N runs this fraction has an empirical distribution. The X-th percentile of that distribution is the $X\%$ credible-region area: by construction, $X\%$ of injections needed at most that fractional area (counted top-down by statistic) to enclose the true direction, so on a future injection a region of that size has an $X\%$ chance of containing the true source.

5.3 Validation

5.3.1 Single-source results

The single-source results establish the standalone localisation capability of the statistic and serve as the baseline for the multi-source tests. A 24×12 grid of injection positions (uniform in $\alpha \in [0, 2\pi)$ and $\delta \in [-85°, +85°]$, 15° spacing, 288 injections) is run in ET+CE+A1 with PSD-coloured Gaussian noise. Statistics are area-weighted by $\cos \delta_i$ so they are representative of a sky position drawn uniformly on the sphere.

Three source classes are tested. The standard reference distance for a Galactic core-collapse supernova is 10 kpc, so a Mezzacappa D15-3D CCSN [22] is injected at that distance and the CBC reference distances are picked to give comparable network SNR: 10 Gpc for a face-on 20+20 $M_\odot$ BBH (IMRPhenomD [20, 21], same masses as all multi-source tests) and 1 Gpc for a face-on 1.4+1.4 $M_\odot$ BNS. The BBH and BNS sweeps additionally include more demanding distances (20 and 30 Gpc for BBH, 2 Gpc for BNS) to probe the SNR-degradation regime. Aggregate performance is summarised in Table 5.2, and the angular-error histograms and Mollweide accuracy maps for all six configurations are in Appendix B.2 (Figures B.2, B.4 and B.3).

Table 5.2: Single-source localisation summary. $N = 288$ injections per column, areas in deg^2. All metrics degrade with distance: doubling the distance inflates the median error by a factor of ∼3–4 for every source type. BNS and BBH behave near-identically at matched SNR, consistent with their shared quadrupole-driven waveform morphology. The CCSN run keeps a small angular error (median 1.3°, 90th percentile 12.4°) but its 90% CR is exceptionally large (1463 deg^2 versus 35 for BBH/BNS at their reference distance), because the credible area is spread over disjoint clusters of pixels scattered across the sky rather than a single contiguous region.

	BBH (20+20 $M_\odot$)			BNS (1.4+1.4 $M_\odot$)		CCSN
	10 Gpc	20 Gpc	30 Gpc	1 Gpc	2 Gpc	10 kpc
Mean error	3.4°	14.4°	36.6°	2.5°	14.1°	11.8°
Median error	1.9°	5.8°	15.9°	1.3°	5.5°	1.3°
90th percentile	6.1°	21.3°	115°	5.5°	16.8°	12.4°
50% CR area	12	64	402	12	68	12
90% CR area	35	375	5359	35	631	1463

The sky-position dependence of localisation quality comes from the network antenna response: high-error sky positions fall inside the network's low-sensitivity regions, as shown in Figure 5.3 for the 20 Gpc BBH case as a representative example. The credible region for one example injection inflates from a compact patch at 10 Gpc to a broad region at 30 Gpc as the coherent-energy peak broadens with decreasing SNR (Figure 5.4).

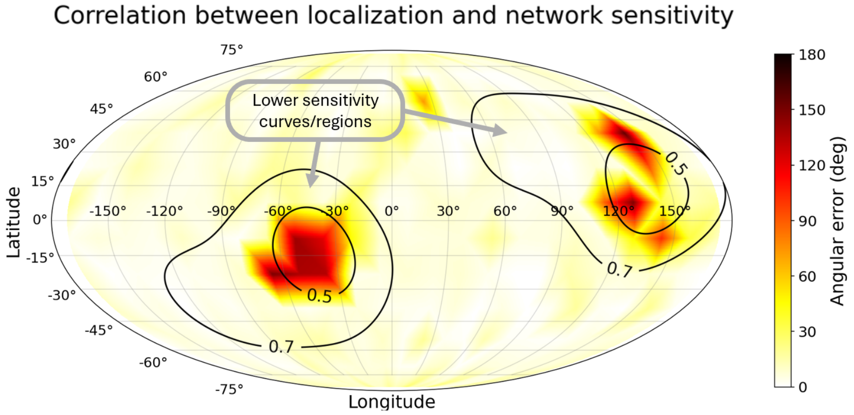

Figure 5.3: Single-source localisation vs. network sensitivity, shown for the 20 Gpc BBH sweep as a representative example. Filled hot colormap: per-injection angular error. Black iso-contours: band-integrated ET+CE+A1 network sensitivity normalised to the network peak. The two high-error clusters sit inside the 0.5 contour and concentrate near the 0.7 level, confirming that single-source localisation quality is governed by the network antenna response rather than by the signal itself.

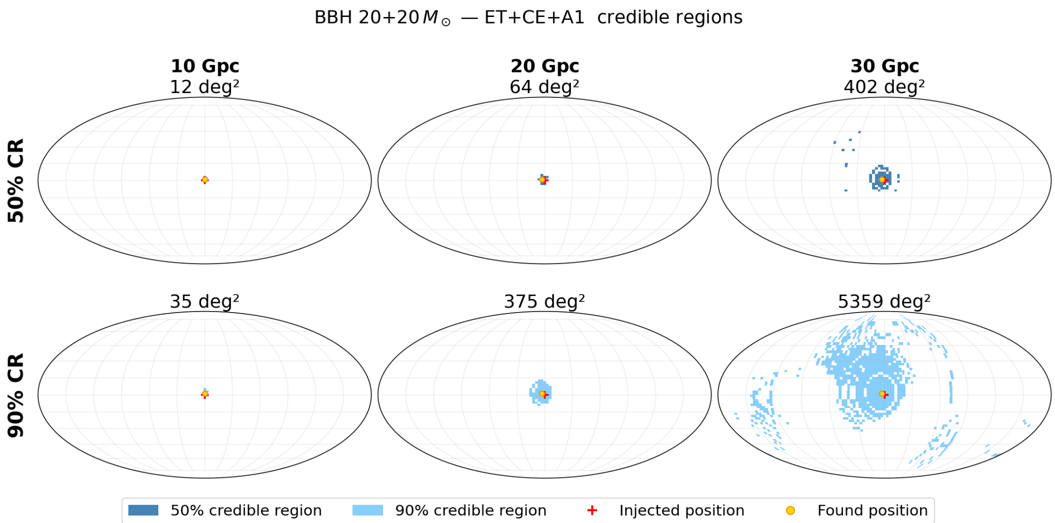

Figure 5.4: 50% (top) and 90% (bottom) credible regions based on greedy pixel sort statistics at a single fixed sky position at 10, 20, 30 Gpc (columns). Red cross = injection, gold circle = recovered peak. The regions grow with distance as the SNR drops. 90% CR grows proportionally faster with distance than 50% CR.

5.3.2 Multi-source overlapping events

The following sections validate the multi-source localisation. CBC–CBC scenarios draw both positions uniformly on the sphere per injection, the sampling already gives the correct area weighting, so statistics are unweighted.

CCSN scenarios require a different approach. ET detects CCSNs only at Galactic distances. The expected sky distribution follows the Galactic stellar mass, concentrated in the disk, peaked toward the Galactic centre. Following Adams et al. [38], the progenitor distribution is modelled as a pure exponential thin disk:

$$\rho(R, z) \propto \exp\left(-\frac{R}{R_d}\right) \exp\left(-\frac{|z|}{H}\right), \qquad R_d = 2.9\,\text{kpc},\ H = 0.095\,\text{kpc}, \tag{5.19}$$

truncated at $R_{\text{out}} = 15\,\text{kpc}$, with the Sun at $R_\odot = 8.7\,\text{kpc}$, $z_\odot = 24\,\text{pc}$. Positions are inverse-CDF sampled ($R \sim \text{Gamma}(2, R_d)$, $z \sim \text{Laplace}(0, H)$, $\phi \sim \mathcal{U}[0, 2\pi)$), converted to heliocentric Galactic (l, b, d) and then equatorial (α, δ), and clipped to $d < 15\,\text{kpc}$ ($\sim 85\%$ of the cumulative Galactic rate, Fig. 2 of [38]). Geocentric time is randomised over one sidereal day. The companion CBC is placed at a random sky position per injection.

Source type combinations

All three pairwise combinations of the single-source reference signals (Section 5.3.1) are validated in two-source mode: BNS+BBH ($N = 150$), BBH+CCSN ($N = 50$) and BNS+CCSN ($N = 50$). BBH and BNS sit at their single-source reference distances of 10 and 1 Gpc respectively, and the CCSN sky position is drawn from the Galactic distribution above. Sources are timed to maximise time–frequency overlap, the worst case for the joint localisation. Mixed polarization is used, where CBCs use the circular assumption from Section 5.1.2 and CCSN uses the general 2D steering basis. Aggregate performance is summarised in Table 5.3, and representative spectrograms and per-source Mollweide scatter and error histograms for all three combinations are in Appendix B.3.

TABLE 5.3: Per-source localisation summary across the three multi-source combinations. Areas in deg^2. Compared to the single-source baseline (Table 5.2), two degradation patterns emerge. CBC+CBC (BNS+BBH) is tail-heavy: the median roughly doubles but the 90% CR grows far more than the 50% CR compared to single-source performance, meaning a fraction of injections are lost to mutual interference while the rest retain near single-source accuracy. CBC+CCSN degrades the full distribution (BBH median ×4, BNS and CCSN median ×6 compared to single-source), reflecting the CCSN's wide time–frequency footprint interfering with the CBC.

	BNS+BBH		BBH+CCSN		BNS+CCSN	
	BNS	BBH	BBH	CCSN	BNS	CCSN
Mean error	5.0°	10.2°	22.3°	39.2°	30.7°	27.9°
Median error	2.2°	4.2°	7.7°	7.6°	8.2°	7.9°
90th percentile	7.2°	22.8°	73°	153°	84°	82°
50% CR area	17	23	41	221	88	1086
90% CR area	82	280	20936	32933	16845	38041

Distance dependence and localisability

The CCSN (D15-3D) is fixed at 10 kpc while the CBC companion is swept logarithmically over ten distances from 50 to 20 000 Mpc, from CBC-dominated to effectively CCSN-only data. Three configurations are compared: *joint ML with dual-source pixels* ($\varepsilon_{\mathrm{null}} = 0.072$, the full algorithm), *strict pixel assignment* (single-source statistic at every pixel via the max mixture), and *single-source mode* (returns only the dominant source).

Both BBH+CCSN (Figure 5.5) and BNS+CCSN (Appendix B.3, Figure B.9) show joint ML consistently outperforming strict assignment, which proves the value of admitting multi-source pixels in the joint statistic. The single-source transition from CBC to CCSN occurs at a smaller distance for BNS than for BBH, reflecting the lower BNS SNR. The CCSN is missed at the closest BBH distances in this run but recovered in the BNS run, more likely a coincidence of sky location than a signal-dependent effect.

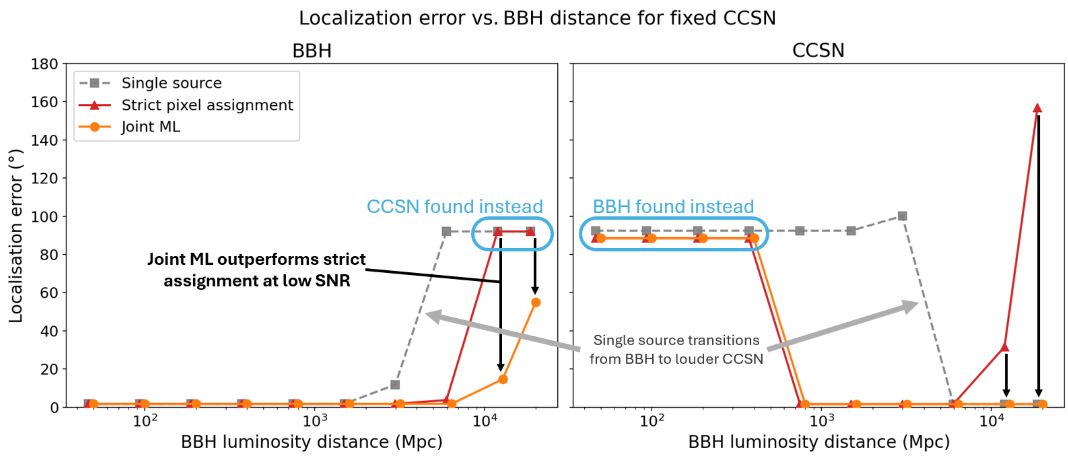

Figure 5.5: BBH+CCSN distance sweep, BBH error (left) and CCSN error (right) vs. BBH luminosity distance with the CCSN fixed at 10 kpc. Joint ML (orange), strict pixel assignment (red), single source (grey). At small BBH distances the BBH dominates and single-source mode locks onto it, missing the CCSN. At the largest distances the BBH becomes undetectable; strict assignment then corrupts the CCSN by forcing a two-source surface onto single-source data, while joint ML still recovers the CCSN accurately. Since the pipeline always reports two positions, the now-undetectable BBH is assigned a spurious location.

Single source in two-source mode

Figure 5.6 shows $N = 50$ injections of a single BBH at 10 Gpc run in two-source mode. The Higher-L position (higher per-source likelihood) concentrates near truth in every run while the Lower-L position is broadly scattered, confirming that the per-position likelihood is an unambiguous criterion for identifying the spurious second position. The Higher-L median is only marginally worse with a median angular error of 2.5° compared to the single-source baseline of 1.9° (Section 5.3.1), so running in two-source mode does not significantly harm localisation of the real source. Deciding one- versus two-source automatically via a likelihood ratio test is left as future work.

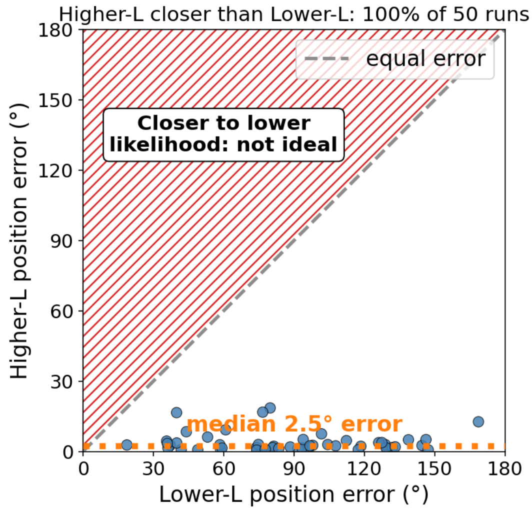

FIGURE 5.6: Single BBH at 10 Gpc run in two-source mode ($N = 50$). *Left:* angular errors of the Higher-L (steelblue) and Lower-L (firebrick) found positions, with median lines. *Right:* per-injection scatter, dashed diagonal = equal error. The Higher-L position always corresponds to the true source with median angular error of 2.5°, while the Lower-L position is essentially random.

5.4 Conclusion

This chapter developed the sky localisation module at two levels: a pixel-level coherent statistic (single-source) and a joint ML statistic (two-source) with a null-fraction classifier, accumulated into a segment-level 4D joint sky surface from which marginal maps and empirical credible regions are read at the joint peak. The circular polarization assumption for CBC sources was justified for $K_{\text{eff}} = 4$ and validated against an oracle and a (ψ, χ) scan.

Validation on PSD-coloured ET+CE+A1 noise gave single-digit-degree single-source localisation for BBH at 10 Gpc, BNS at 1 Gpc, and CCSN at 10 kpc. The pipeline also localises all three pairwise combinations (BBH+BNS, BBH+CCSN, BNS+CCSN), with the joint ML statistic consistently outperforming strict assignment. When only one source is present, the pipeline still reports two candidate positions and recovers the real source as the higher-likelihood of the two, with accuracy close to the single-source case, without ever deciding the source count.

There is however clear room for improvement in tuning parameters such as the STFT window length, the choice of likelihood statistic, the pixel-classification calibration, etc. Tuning of this kind requires iteration over computationally heavy injection runs and was left out of the thesis. Raising the sky-grid resolution would similarly improve localisation, but the $O(N_{\text{grid}}^2)$ scaling of the two-source scan made finer grids prohibitive within the practically available compute budget (my laptop).

Chapter 6

Conclusion

This thesis addressed the three gaps set out in the introduction: overlap quantification, pipeline support, and multi-source localisation.

Quantification (Chapter 3). The temporal and time-frequency overlap probability between a Galactic CCSN and a realistic one-year CBC population was computed for ET. At the few-hertz frequencies ET is designed to reach, CCSN–BNS temporal overlap is 100% and the time-frequency overlap probability reaches 87–91% for the heavier progenitors, concentrated in the sub-16 Hz regime where the CCSN memory effect leaves most of its energy. CCSN–BBH overlap is expected to be even higher than the BNS case given the larger expected BBH rate. Both CCSN–BNS and CCSN–BBH overlap are therefore relevant scenarios for the third-generation pipeline.

Pipeline support (Chapter 4). The MDAP [6] was extended with a completely new sky-localisation module and refactored at the data-flow level so that an N-indexed per-source descriptor is propagated from that module through to event reconstruction. The likelihood, clustering and event-reconstruction modules received new multi-source implementations: a model-aware likelihood stage emits per-source coherent-energy maps and a single- versus multi-source pixel classification, clustering produces per-source pixel masks, and reconstruction runs once per source. This is a structural change beyond replacing individual module implementations, and the pipeline's modular design is preserved so that future module implementations remain straightforward. An end-to-end test with a simultaneous CBC and CCSN injection identifies the sky positions, separates and reconstructs both sources in one pipeline pass.

Multi-source localisation (Chapter 5). The single-source coherent statistic of cWB [5] was extended into a joint maximum-likelihood framework for two simultaneous sources, with a null-fraction classifier to select the appropriate statistic for each pixel. The CBC polarization constraint collapses the per-source signal subspace from two dimensions to one, making the two-source problem well-posed on the 3G network. Validation on PSD-coloured noise gave single-digit-degree single-source localisation for BBH at 10 Gpc, BNS at 1 Gpc and CCSN at 10 kpc, and the same pipeline localises all three pairwise multi-source combinations.

Limitations and future work

Three limitations bound the present results. (i) The BBH population file used in Chapter 3 was found during this work to underrepresent the realistic rate by about an order of magnitude. The file was obtained externally and producing a corrected one was beyond the scope of the thesis, so the BNS result acts as a lower bound on CCSN–CBC overlap rather than an absolute number for CCSN–BBH as well. (ii) Clustering was newly implemented for the multi-source pipeline along the lines of the general cWB approach [5], but was not tuned further: the focus was on proving the multi-source concept rather than on fine-tuning individual modules, and the present rules in particular do not yet handle diffuse bursts optimally, limiting the CCSN reconstruction quality. (iii) False-alarm calibration of the single- and multi-source statistics on realistic non-stationary backgrounds is outside the present scope. The algorithmic performance is established, but the detection significance under glitchy backgrounds is not. A number of pipeline parameters (STFT window length, likelihood-statistic variants, convergence thresholds and sky-grid resolution) were also left at their initial values for the compute-budget reasons discussed in the conclusion of Chapter 5.

The pipeline's modular design (the M in MDAP) is explicitly built to absorb replacement and refinement of individual modules as research progresses, and further fine-tuning of the current module implementations along the lines of the three limitations above is possible for almost all modules.

Appendices

Appendix A

Chapter 3 Supplementary Figures

This appendix collects supplementary figures for Chapter 3.

A.1 CCSN source gallery

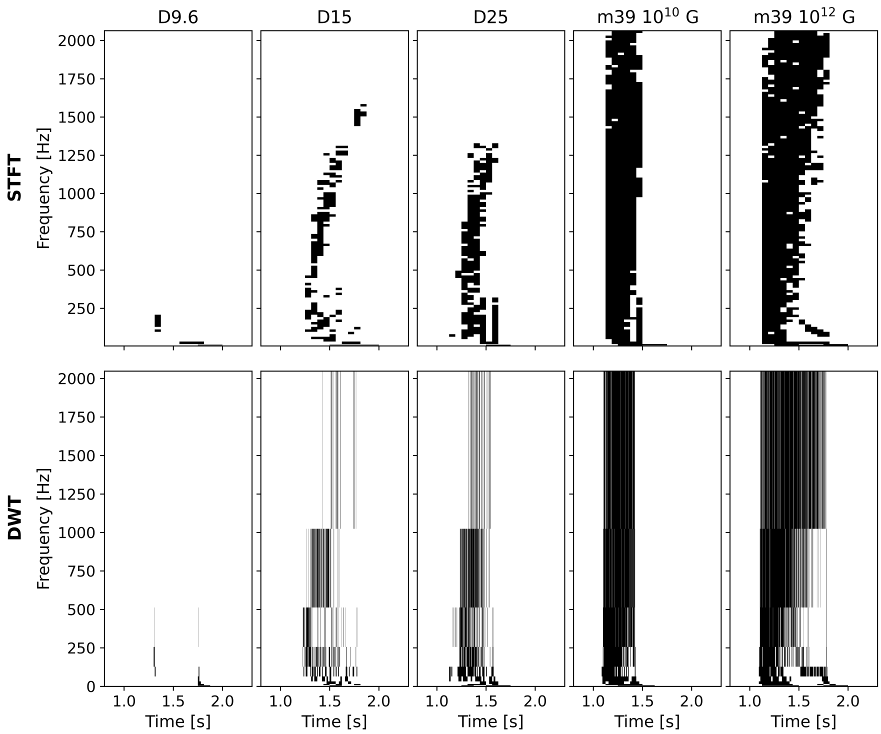

Figure A.1: TF masks ($\sigma = 1$, ET-D, 10 kpc, equatorial viewing) for the five CCSN sources. *Top row*: hybrid STFT. *Bottom row*: DWT (sym14, $L = 9$). Signal injected at $\approx$1.1 s, window [0.8, 2.3] s. D9.6, D15, D25: neutrino-driven models [22]; m39 (1e10), m39 (1e12): magnetorotational models [24].

A.2 BNS overlap-fraction distribution

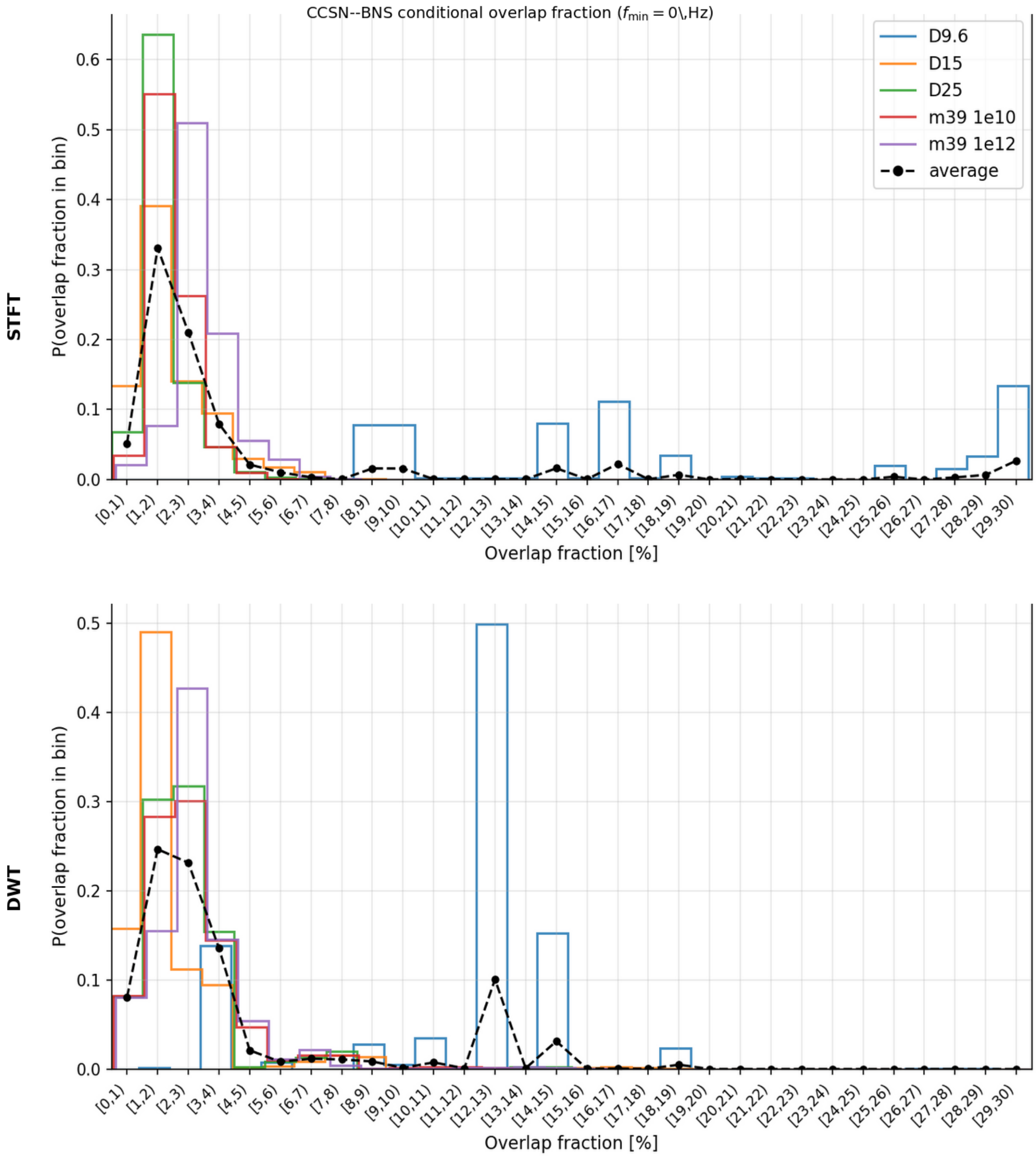

Figure A.2: Conditional distribution of η at $f_{\mathrm{min}} = 0\,\mathrm{Hz}$ for the BNS day mask ($\sigma = 1$, averaged over the three ET interferometers and two viewing orientations, zero-overlap shifts excluded). Top: hybrid STFT. Bottom: DWT.

A.3 CCSN–BBH overlap

As explained in the BBH population caveat (Section 3.1), the BBH catalog file used in this thesis underrepresents the realistic BBH event rate by roughly an order of magnitude. The absolute overlap probabilities shown below are therefore underestimates. The shape of the curves is unaffected, since it is set by per-event signal morphology rather than event count. Because the true BBH population outnumbers the BNS population, a correctly populated BBH catalog would yield overlap that meets or exceeds the BNS values reported in Chapter 3.

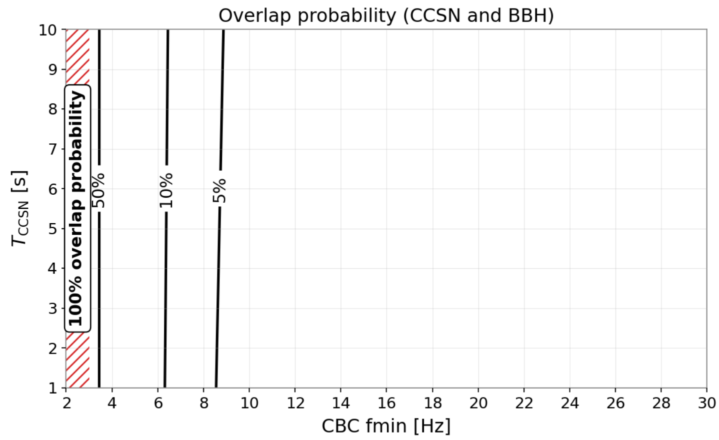

Figure A.3: Temporal overlap probability between a CCSN time interval of duration T_{CCSN} and a one-day BBH population, as a function of $f_{\mathrm{CBC}}^{\mathrm{min}}$. The hatched band marks 100% overlap at the low frequencies (∼3 Hz) ET is designed to reach [1]. Because Galactic CCSN are rare (∼one per 60 years [4]), even the 5–10% contours are operationally relevant and lie well within the band of any current or planned detector. Absolute probabilities here underestimate the realistic BBH rate (see preamble); with the true BBH population the overlap is expected to meet or exceed the BNS values in Chapter 3.

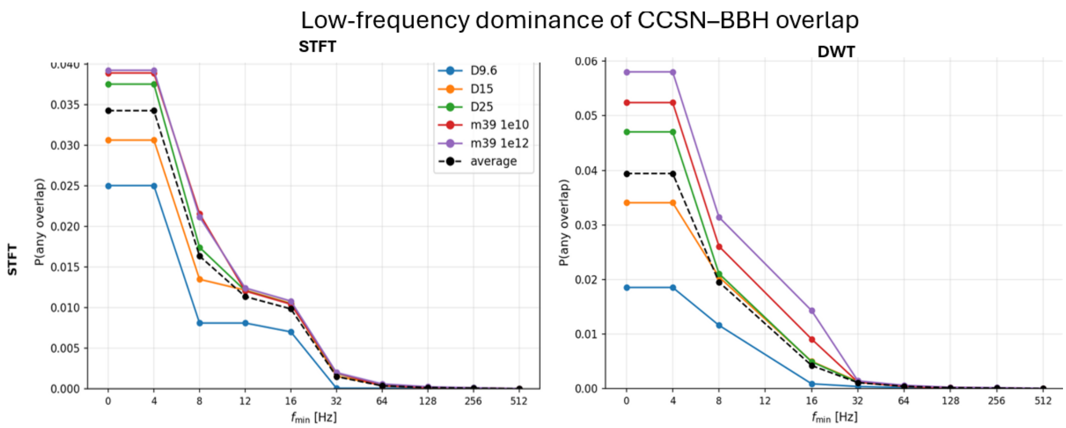

Figure A.4: CCSN–BBH overlap probability vs. f_{min} ($\sigma = 1$, five source curves and their mean (dashed), averaged over the three ET interferometers and two viewing orientations). Top: hybrid STFT. Bottom: DWT; the 12 Hz tick has no data point, reflecting the absence of a dedicated bin at this frequency in the dyadic decomposition.

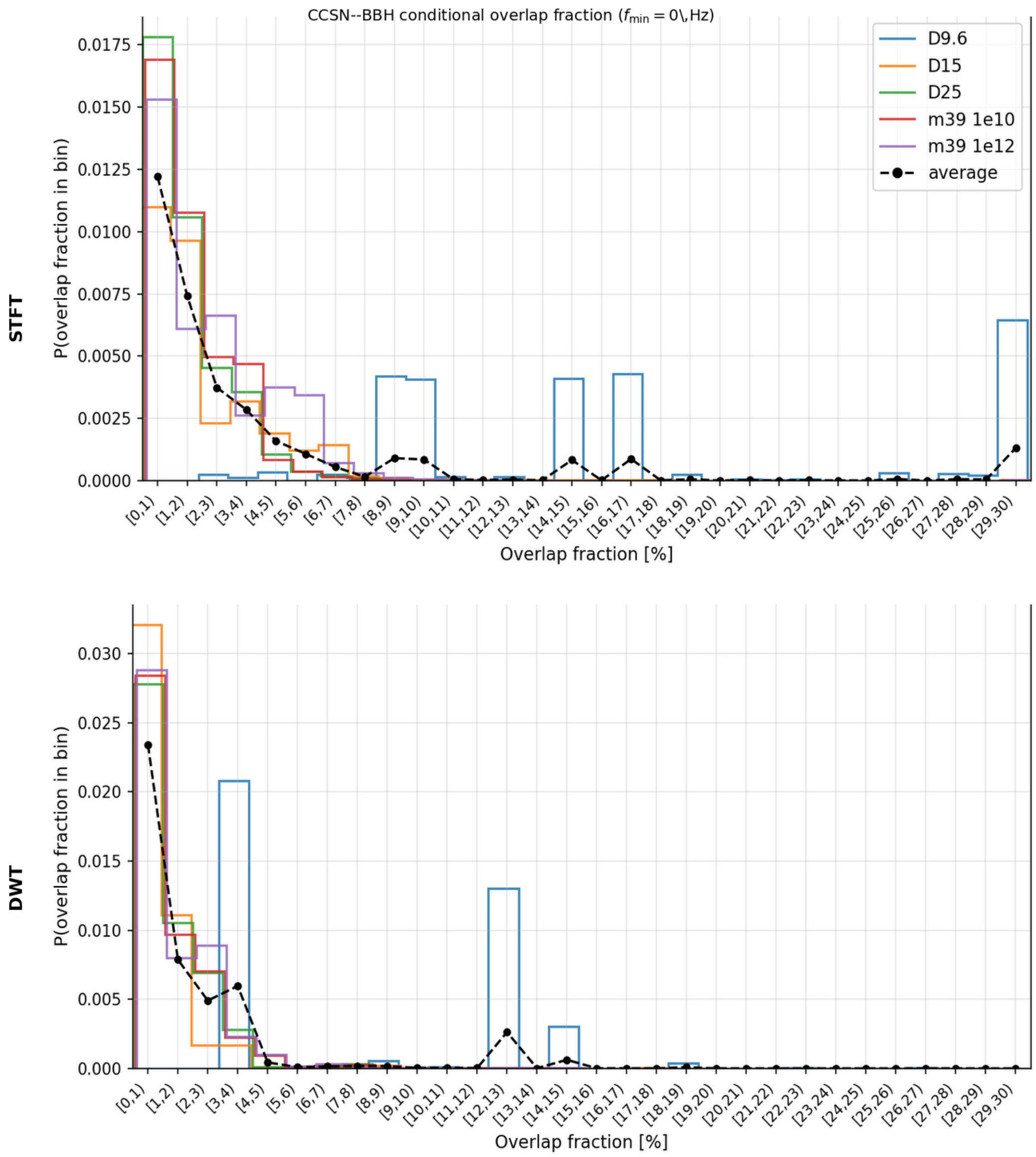

Figure A.5: Conditional distribution of η at $f_{\mathrm{min}} = 0\,\mathrm{Hz}$ for the BBH day mask ($\sigma = 1$, averaged over the three ET interferometers and two viewing orientations, zero-overlap shifts excluded). Top: hybrid STFT. Bottom: DWT.

Appendix B

Chapter 5 Supplementary Figures

This appendix collects supplementary figures for Chapter 5.

B.1 Null-fraction threshold calibration

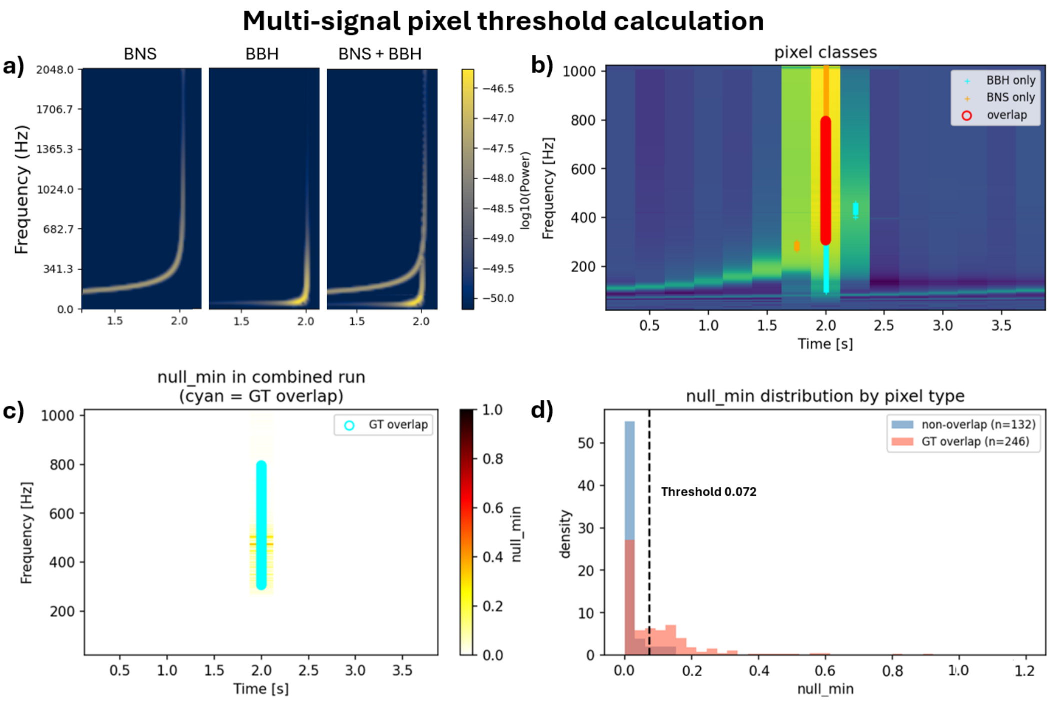

Figure B.1: Null-fraction threshold calibration on a noise-free maximum-TF-overlap BNS+BBH injection. *(a)* TF power maps (BNS-only, BBH-only, combined). *(b)* ground-truth pixel labels. *(c)* ξ per selected pixel of the combined run. *(d)* ξ distributions split by ground-truth class; the dashed line at $\varepsilon_{\text{null}} = 0.072$ is the 90th percentile of the single-source distribution.

B.2 Single-source localisation maps

The summary statistics for these runs are reported in Table 5.2 of Chapter 5. The sky-position accuracy maps are reproduced here for visual reference.

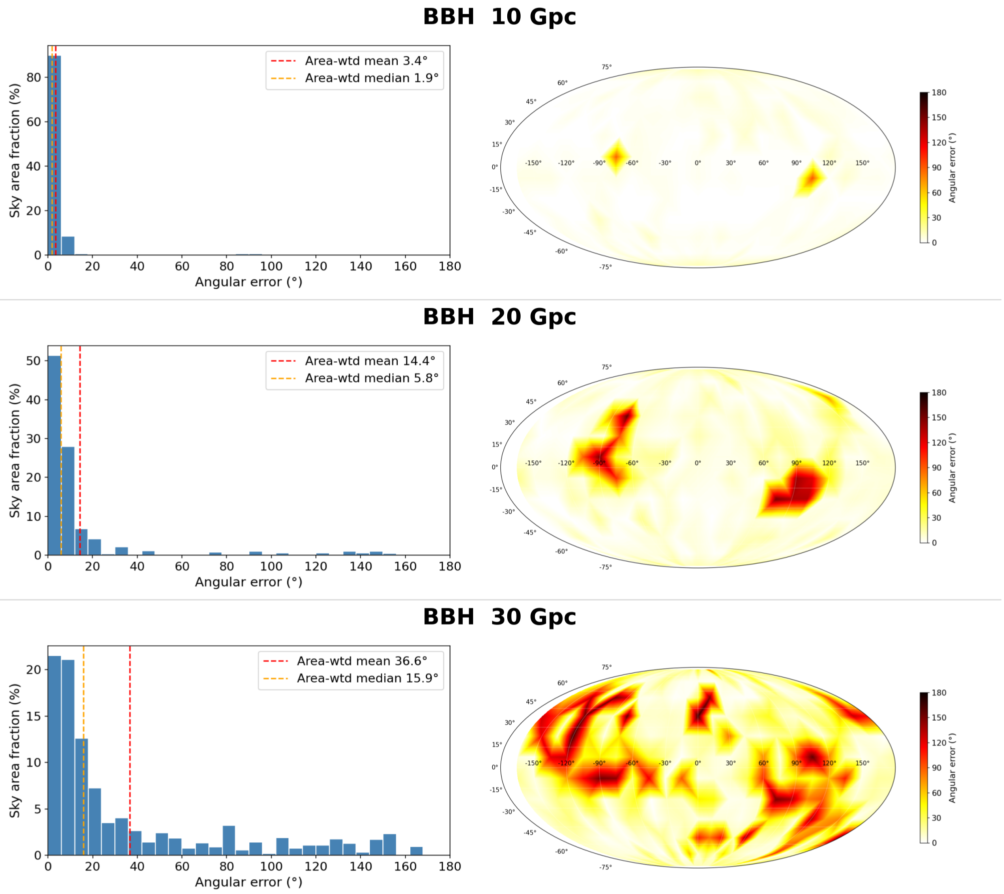

Figure B.2: BBH single-source localisation at 10, 20, 30 Gpc (rows): error histogram (left) and Mollweide error map (right).

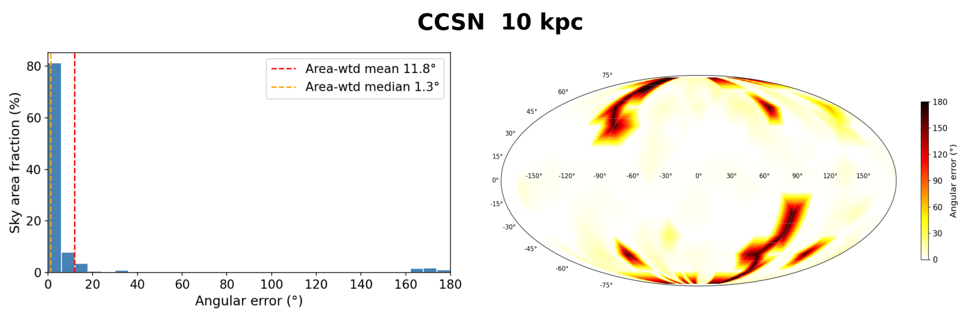

Figure B.3: CCSN single-source localisation (D15-3D, 10 kpc)

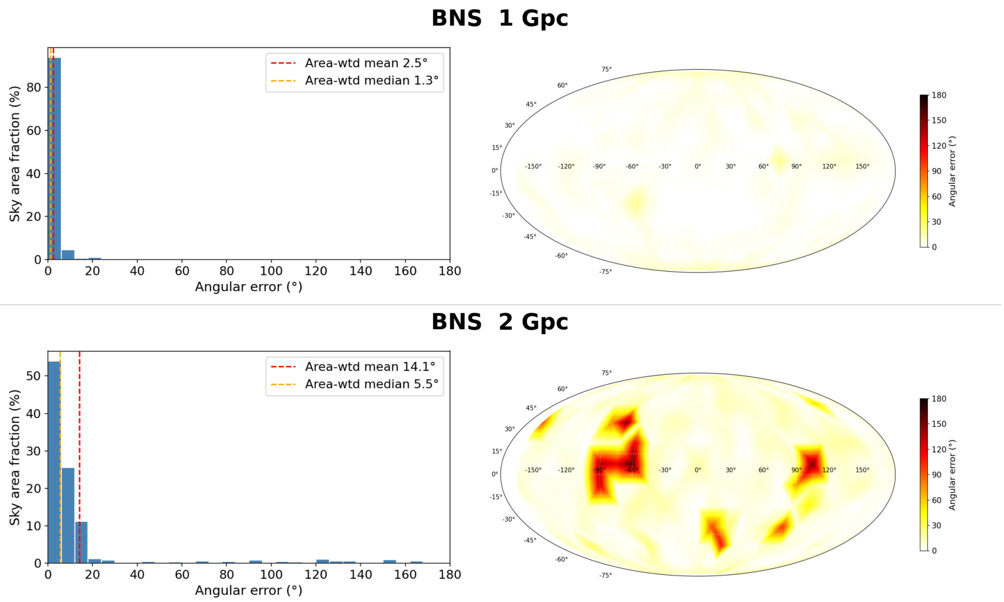

FIGURE B.4: BNS single-source localisation at 1 Gpc (top) and 2 Gpc (bottom)

B.3 Multi-source per-injection figures

Summary statistics are in Table 5.3 of Chapter 5. The figures below show a representative spectrogram and the per-source Mollweide scatter and error histograms.

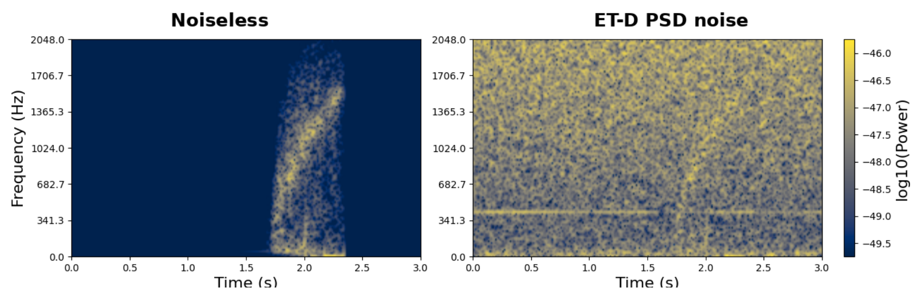

FIGURE B.5: Representative ET1 spectrogram of a BBH+CCSN injection

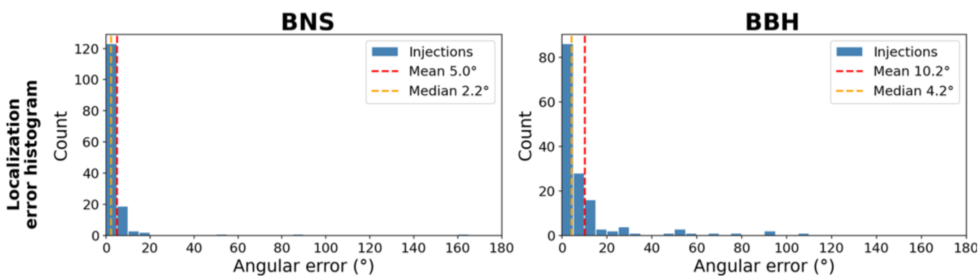

FIGURE B.6: BNS+BBH ($N = 150$) per-source error histograms (bottom).

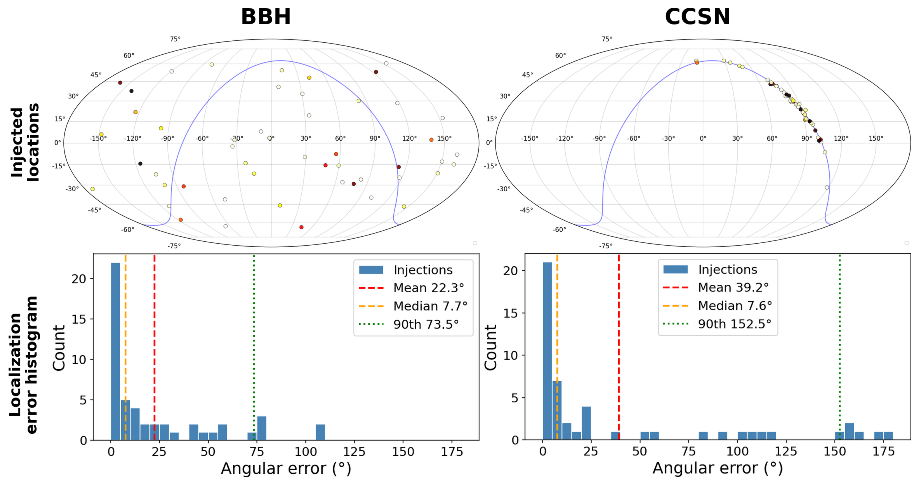

Figure B.7: BBH+CCSN ($N = 50$), layout as Figure B.6. CCSN positions trace the Galactic plane.

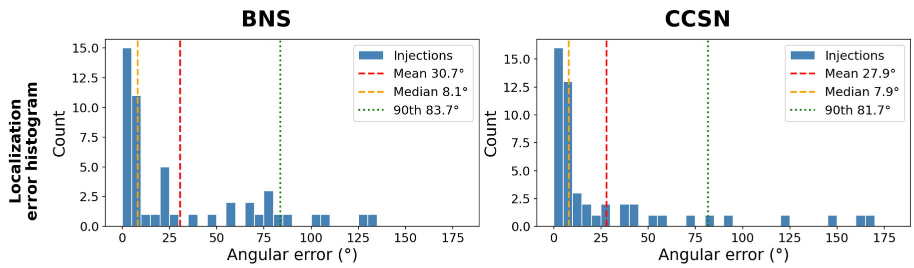

Figure B.8: BNS+CCSN ($N = 50$): per-source error histograms. The CCSN position scatter matches BBH+CCSN (Figure B.7) and is not reproduced.

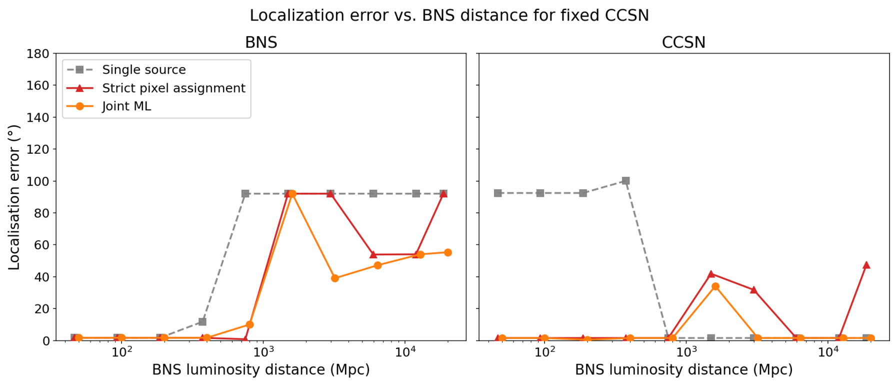

Figure B.9: BNS+CCSN distance sweep, layout as Figure 5.5 of Chapter 5. The single-source transition occurs at smaller distances than the BBH case, reflecting the lower BNS network SNR, and the BNS does not suppress the CCSN at close distances. At the largest distances strict assignment degrades the CCSN, while joint ML keeps sub-degree CCSN accuracy throughout.

Bibliography

[1] M. Maggiore, C. Van Den Broeck, N. Bartolo, E. Belgacem, D. Bertacca, M. A. Bizouard, M. Branchesi, S. Clesse, S. Foffa, J. García-Bellido, S. Grimm, J. Harms, T. Hinderer, S. Matarrese, C. Palomba, M. Peloso, A. Ricciardone, and M. Sakellariadou, “Science case for the Einstein Telescope,” *Journal of Cosmology and Astroparticle Physics*, vol. 2020, no. 03, p. 050, 2020.

[2] M. Branchesi *et al.*, “Science with the einstein telescope: a comparison of different designs,” *Journal of Cosmology and Astroparticle Physics*, vol. 2023, no. 07, p. 068, 2023.

[3] N. Singh, T. Bulik, K. Belczynski, and A. Askar, “Exploring compact binary populations with the Einstein Telescope,” *Astronomy & Astrophysics*, vol. 667, p. A2, 2022.

[4] K. Rozwadowska, F. Vissani, and E. Cappellaro, “On the rate of core collapse supernovae in the Milky Way,” *New Astronomy*, vol. 83, p. 101498, 2021.

[5] S. Klimenko, I. Yakushin, A. Mercer, and G. Mitselmakher, “Coherent method for detection of gravitational wave bursts,” *Classical and Quantum Gravity*, vol. 25, p. 114029, 2008.

[6] A. Mannaert, “A modular wavelet analysis pipeline for the detection of gravitational waves with the Einstein Telescope,” Master’s thesis, KU Leuven, 2025.

[7] F. Badaracco *et al.*, “Blind source separation in 3rd generation gravitational-wave detectors,” *New Astronomy Reviews*, vol. 99, p. 101707, 2024.

[8] A. Mezzacappa and M. Zanolin, “Gravitational waves from neutrino-driven core collapse supernovae: Predictions, detection, and parameter estimation,” *arXiv preprint*, 2024.

[9] A. Buonanno, “Gravitational waves,” *Les Houches Summer School Lecture Notes*, 2007, arXiv:0709.4682.

[10] S. A. Usman *et al.*, “The PyCBC search for gravitational waves from compact binary coalescence,” *Classical and Quantum Gravity*, vol. 33, no. 21, p. 215004, 2016.

[11] C. Messick *et al.*, "Analysis framework for the prompt discovery of compact binary mergers in gravitational-wave data," *Physical Review D*, vol. 95, p. 042001, 2017.

[12] M. Evans *et al.*, "A horizon study for cosmic explorer: Science, observatories, and community," Cosmic Explorer Horizon Study report, 2021.

[13] M. Saleem *et al.*, "The science case for LIGO-India," *Classical and Quantum Gravity*, vol. 39, no. 2, p. 025004, 2022.

[14] P. Relton, A. Virtuoso, S. Bini, V. Raymond, I. Harry, M. Drago, C. Lazzaro, A. Miani, and S. Tiwari, "Addressing the challenges of detecting time-overlapping compact binary coalescences," *Physical Review D*, vol. 106, p. 104045, 2022.

[15] F. Cireddu, "GW mock data challenge v1: Population files," Zenodo sandbox, https://sandbox.zenodo.org/records/413548, 2025, version 3. Files: `18321_1yrCatalogBNS.h5`, `18321_1yrCatalogBBH.h5`.

[16] I. C. F. Wong and F. Cireddu, "gwmock: Python package for creating mock gravitational-wave samples," https://pypi.org/project/gwmock/.

[17] M. Hannam, P. Schmidt, A. Bohé, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Pürrer, "Simple model of complete precessing black-hole-binary gravitational waveforms," *Physical Review Letters*, vol. 113, p. 151101, 2014.

[18] T. Dietrich, A. Samajdar, S. Khan, N. K. Johnson-McDaniel, R. Dudi, and W. Tichy, "Improving the NRTidal model for binary neutron star systems," *Physical Review D*, vol. 100, p. 044003, 2019.

[19] G. Pratten, C. García-Quirós, M. Colleoni, A. Ramos-Buades, H. Estellés, M. Mateu-Lucena, R. Jaume, M. Haney, D. Keitel, J. E. Thompson, and S. Husa, "Computationally efficient models for the dominant and subdominant harmonic modes of precessing binary black holes," *Physical Review D*, vol. 103, p. 104056, 2021.

[20] S. Husa, S. Khan, M. Hannam, M. Pürrer, F. Ohme, X. Jiménez Forteza, and A. Bohé, "Frequency-domain gravitational waves from nonprecessing black-hole binaries. I. New numerical waveforms and anatomy of the signal," *Physical Review D*, vol. 93, p. 044006, 2016.

[21] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Pürrer, X. Jiménez Forteza, and A. Bohé, "Frequency-domain gravitational waves from nonprecessing black-hole binaries. II. A phenomenological model for the advanced detector era," *Physical Review D*, vol. 93, p. 044007, 2016.

[22] A. Mezzacappa, P. Marronetti, R. E. Landfield, E. J. Lentz, W. R. Hix, J. A. Harris, S. W. Bruenn, J. M. Blondin, O. E. B. Messer, J. Casanova, and L. L.

Kronzer, "Core collapse supernova gravitational wave emission for progenitors of 9.6, 15, and 25 $m_{\odot}$," *Physical Review D*, vol. 107, p. 043008, 2023.

[23] E. Abdikamalov, G. Pagliaroli, and D. Radice, "Gravitational waves from core-collapse supernovae," in *Handbook of Gravitational Wave Astronomy.* Springer, 2022, pp. 21–40.

[24] J. Powell, B. Müller, D. R. Aguilera-Dena, and N. Langer, "Three dimensional magnetorotational core-collapse supernova explosions of a 39 solar mass progenitor star," *Monthly Notices of the Royal Astronomical Society*, 2023.

[25] S. Hild *et al.*, "Sensitivity studies for third-generation gravitational wave observatories," *Classical and Quantum Gravity*, vol. 28, no. 9, p. 094013, 2011.

[26] A. Freise, S. Chelkowski, S. Hild, W. Del Pozzo, A. Perreca, and A. Vecchio, "Triple Michelson interferometer for a third-generation gravitational wave detector," *Classical and Quantum Gravity*, vol. 26, no. 8, p. 085012, 2009, arXiv:0804.1036.

[27] G. Ashton *et al.*, "BILBY: A user-friendly Bayesian inference library for gravitational-wave astronomy," *The Astrophysical Journal Supplement Series*, vol. 241, no. 2, p. 27, 2019.

[28] S. Klimenko *et al.*, "Method for detection and reconstruction of gravitational wave transients with networks of advanced detectors," *Physical Review D*, vol. 93, p. 042004, 2016.

[29] P. J. Sutton *et al.*, "X-Pipeline: An analysis package for autonomous gravitational-wave burst searches," *New Journal of Physics*, vol. 12, p. 053034, 2010.

[30] W. G. Anderson, P. R. Brady, J. D. E. Creighton, and E. E. Flanagan, "Excess power statistic for detection of burst sources of gravitational radiation," *Physical Review D*, vol. 63, p. 042003, 2001.

[31] M. Drago, "Search for transient gravitational wave signals with unknown waveform in the LIGO-Virgo network of interferometric detectors using a fully coherent algorithm," Ph.D. dissertation, Università degli Studi di Padova, Padova, Italy, 2010.

[32] S. Chatterji, A. Lazzarini, L. Stein, P. J. Sutton, A. Stroeer, and M. Tinto, "Coherent network analysis technique for discriminating gravitational-wave bursts from instrumental noise," *Physical Review D*, vol. 74, p. 082005, 2006.

[33] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, "Constraint likelihood analysis for a network of gravitational wave detectors," *Physical Review D*, vol. 72, p. 122002, 2005.

[34] G. Celeux and G. Govaert, “A classification EM algorithm for clustering and two stochastic versions,” *Computational Statistics & Data Analysis*, vol. 14, no. 3, pp. 315–332, 1992.

[35] A. N. Tikhonov, “Solution of incorrectly formulated problems and the regularization method,” *Soviet Mathematics Doklady*, vol. 4, pp. 1035–1038, 1963.

[36] J. A. Högbom, “Aperture synthesis with a non-regular distribution of interferometer baselines,” *Astronomy and Astrophysics Supplement Series*, vol. 15, pp. 417–426, 1974.

[37] T. Sidery, B. Aylott, N. Christensen, B. Farr, W. Farr, F. Feroz, J. Gair, K. Grover, P. Lasky, I. Mandel, M. Pitkin, G. Pratten, V. Raymond, C. Rover, L. Singer, M. van der Sluys, J. Veitch, S. Vitale, and A. Zimmerman, “Reconstructing the sky location of gravitational-wave detected compact binary mergers: methodology for testing and comparison,” *Physical Review D*, vol. 89, no. 8, p. 084060, 2014.

[38] S. M. Adams, C. S. Kochanek, J. F. Beacom, M. R. Vagins, and K. Z. Stanek, “Observing the next galactic supernova,” *The Astrophysical Journal*, vol. 778, no. 2, p. 164, 2013.

Code of conduct and transparency statement on the use of GenAI for KU Leuven-students (academic year 2025–2026)

Generative AI (GenAI) assistance tools can be used to generate various types of content, including text, images, code, video, music, or combinations thereof. Common examples of such tools include ChatGPT, Google Gemini, Microsoft Copilot, Midjourney, Claude.ai, Perplexity.ai, and DALL · E, among others.

This code of conduct is a tool that helps students to be transparent about the use of GenAI and fits within the university's principles on academic integrity.

Important guidelines and remarks

- **Sensitive or personal data:** Some GenAI tools protect your input and sensitive or personal data better than others. There is often no transparency on what the owners of the AI applications do with the data entered. Therefore, **do not enter sensitive or personal data in free GenAI tools.** More info about what sensitive and personal data are, is described in the three confidentiality levels of the KU Leuven data classification model (non-confidential, confidential, strictly confidential). For confidential data you should preferably use **M365 Copilot Chat with your KU Leuven account**. If you use another GenAI tool, you must be **absolutely certain** it does **not store or reuse the data you enter. Try to avoid entering these data. Do so only if strictly necessary, and only to the extent required. For personal data, work with anonymous or pseudonymized data.** Additionally, in case of strictly confidential data this should first be discussed with the teaching staff of the course or your thesis supervisor.

- **Copyrighted materials:** For **lawfully obtained copyrighted material** you should preferably use M**365 Copilot Chat with your KU Leuven account.** If you use another GenAI tool, you must be **absolutely certain** it does n**ot store or reuse the data you enter.**

- GenAI assistance may not be used for data or topics covered by a Non-Disclosure Agreement (NDA). Even in the absence of an NDA, certain information may still need to be treated as confidential—for example, due to regulatory requirements or the risk of significant harm to the university if disclosed. In case of doubt, check with your teaching staff or supervisor.

- If your master's thesis is under embargo, you should first discuss with your supervisor whether the use of GenAI is permitted.

- Before using a GenAI tool, always consider whether its use is responsible, including from a sustainability perspective (e.g. when using GenAI as a search engine, language assistant,...).

- **Take a scientific and critical attitude** when interacting with GenAI assistance and interpreting its output, that may not always be correct.

- As a student you are responsible for complying with Article 84 of the Regulations on Education and Examinations: your report or thesis should reflect your own knowledge, understanding and skills. Be aware that plagiarism rules also apply to (work that is the result of) the use of GenAI assistance tools.
  Exam Regulations Article 84: *"Every conduct individual students display with which they (partially) inhibit or attempt to inhibit a correct judgement of their own knowledge, understanding and/or skills or those of other students, is considered an irregularity which may result in a suitable penalty. A special type of irregularity is plagiarism, i.e. copying the work (ideas, texts, structures, designs, images, plans, codes , ...) of others or prior personal work in an exact or slightly modified way without adequately acknowledging the sources. Every possession of prohibited resources during an examination (see article 65) is considered an irregularity."*

- In order to maintain academic integrity and avoid plagiarism, **more information about being transparent on the use of GenAI assistance and about correctly citing and referencing GenAI** can be found on this website for students (Dutch/English).

- **Additional reading: KU Leuven guidelines on responsible use of Generative AI tools, and other information** (Dutch/English)

A few final words

If you are uncertain whether or not you should declare your use of GenAI tools, we suggest that you discuss this with your instructor or supervisor. It is always safer to declare GenAI use, even when it is not strictly required. However, declaring GenAI use does not entail that its use is allowed; the right column in the table below provides more detailed instructions in this regard (code of conduct).

Moreover, advanced AI tools are evolving rapidly, and their capabilities have expanded significantly in a short period of time. As a result, we do not yet have all the answers about their responsible use. Finally, it is important to follow-up on the most recent evolutions in AI technologies, to have an open mind but also to be a bit cautious, to communicate with instructors, teaching assistants, supervisors and peers, to be as transparent as we can, and to learn together as we move along.

Student name: Stijn Heyde **Student number: R0887972**

Please indicate whether it relates to a course assignment or to the Bachelor's or Master's thesis:

This form is related to my **Master's thesis.**

Title Master's thesis:	Unmodelled detection of overlapping gravitational wave signals with the Einstein Telescope	**Supervisor:**	Prof. dr. Tjonnie Li and Prof. dr. ir. Marc Moonen

Daily supervisors: Milan Wils and Francesco Cireddu

I **did use** GenAI Assistance. In this case: ChatGPT, Claude.ai, Google Gemini

GenAI assistance used as/for:	**Name of the GenAI tool(s) used** If helpful, also describe in which way you were using GenAI related to what is specified as code of conduct.	**Code of conduct: For each of the categories below, always take into account the important guidelines and remarks mentioned above (e.g. copyrighted data, sensitive or personal data,. . .).**
As a language assistant for reviewing or improving texts I wrote myself	*ChatGPT, Claude.ai, Google Gemini used to make the text flow better after writing it myself*	This use is similar to using spelling and grammar check tools. In general, you do not have to refer to such kind of GenAI use in the text. **However, be careful:** • When using GenAI tools on texts you did not write yourself to improve the text, **you have to refer to the original source or author**, otherwise you are committing plagiarism and thus an irregularity.
As a paraphrasing tool	*ChatGPT only on text I wrote myself*	This use is allowed except when it is prohibited by the teacher or the program of study. You may paraphrase your own text or texts by an author other than yourself and take inspiration from what a GenAI tool or another tool suggests (unless it is not allowed). In general, you do not have to refer to such kind of GenAI use or other paraphrasing tools in the text. **However, be careful**: • If it entails text by an author other than yourself, **you are not allowed to include that paraphrased text without reference to the original source or author**. Without such reference, you would be committing plagiarism and thus an irregularity.

For translation aid to improve texts I wrote myself or to better understand text from others	/	This use is allowed except when it is prohibited by the teacher or the program of study. It is similar to using translation tools (Google translate, DeepL, ...). In general, you do not have to refer to such kind of GenAI use in the text. **However, be careful:** • **You are not allowed to include that translated text without reference to the original source or author.** Without such reference, you would be committing plagiarism and thus an irregularity. • Always check the translated text for correctness and meaning.
As a search engine to get information on a topic or to search for existing research on the topic	*Google Gemini to understand signal processing concepts*	This use is similar to e.g. a Google search or checking Wikipedia. If you write your own text based on this information, you do not have to refer to the use of GenAI in the text. **You only have to refer to the existing research and references you have checked and used** (without such references you would be committing plagiarism and thus committing an irregularity). **However, be careful:** • Be aware that the output of the GenAI tool cannot be guaranteed as a 100% reliable source of information. The output may not be entirely correct and/or be limited due to the databases it uses. Moreover, knowledge evolves and may change over time; therefore, the database of the GenAI tool may not be up to date. Therefore, **verify the information and do not just copy-paste it** as you should understand and critically process everything you are writing.

For literature search	*Claude.ai, ChatGPT, Google Gemini*	This use is comparable to e.g. a Google Scholar search. You do not have to refer to such kind of GenAI use; **you only have to refer to the literature references you have checked and used** (without such references you would be committing plagiarism and thus committing an irregularity). **However, be careful:** • Be aware that the search output is restricted to the database the GenAI tool is built on. After this initial search, look for scientific sources and conduct your own analysis of the source documents. **Interpret, analyse and process the information you obtained; verify it and do not just copy-paste it** as you should understand and critically process everything you are writing. • Be aware that some GenAI tools may **output no or wrong references.** As a student you are responsible for further checking and verifying the absence or correctness of references; do not just copy-paste it.
For generating programming code	*ChatGPT used for code to fix bugs / generate quick diagnostic plots and cleaning up / documenting my own and other's code*	Use of GenAI for coding is allowed except when it is prohibited by the teacher or the program of study. If used for coding, correctly mention the use of GenAI assistance in accordance with the instructions on the page on being transparant about the use of GenAI.

For generating new (research) ideas	/	This use of GenAI is allowed except when it is prohibited by the teacher or the program of study. Correctly mention the use of GenAI assistance in accordance with the instructions on the page on being transparent about the use of GenAI. **Be careful:** • Further verify in this case whether the idea is novel or not. It is likely that it is related to existing work. **If so, that existing work should be correctly referenced in the text** (without such reference you would be committing plagiarism and thus committing an irregularity).
For generating synthetic data	/	Use of GenAI for generating synthetic data is allowed, **provided that it is methodologically and ethically justifiable**, except when it is prohibited by the teacher or the program of study. Always correctly mention the use of GenAI assistance in accordance with the instructions on the page on being transparent about the use of GenAI. **Be careful:** • Always **carefully evaluate** the generated synthetic data for quality and possible bias since the output is highly dependent on the quality of the data on which the models are trained.

For generating blocks of text (other than the allowed use without referencing mentioned above)	/	According to Article 84 of the Regulations on Education and Examinations your text should allow to correctly and properly assess your own knowledge, understanding and skills. **Therefore, inserting blocks of text without quotes and a reference to GenAI assistance in your work is not allowed.** **Be careful:** • If it is really needed to insert a block of text from a GenAI tool, for instance because of the nature of your assignment, mention it as a citation by using quotes and correctly mention the use of GenAI assistance in accordance with the instructions page on being transparent about the use of GenAI. • However, in general, such GenAI use should be kept to an **absolute minimum**; you should always check the original sources.

For generating visuals, video or audio	/	Use of GenAI for generating visuals, audio or video is allowed except when it is prohibited by the teacher or the program of study. **Be careful:** • If used, refer to GenAI for visuals according to the usual referencing style, following the instructions on the page on referencing GenAI. • If you work with existing visuals, audio or video, **that existing work should be correctly referenced in the text** (without such reference you would be committing plagiarism and thus an irregularity). • Explain the usage in the methods section (if there is one) and optionally attach (or link to) the prompts with the full output (or history).
Other use (specify here; this may also include a combination of types of use mentioned above):	/	To make sure other use of GenAI is allowed within the course or thesis, and if so, the conditions that may apply, contact the teaching staff of the course or the supervisor of the thesis **beforehand** and explain the intended GenAI use. Also inform the programme director. Motivate how you would comply with Article 84 of the exam regulations. Explain the use and the added value of the AI tool you consider to use and how it is in accordance with the assignment or thesis and the KU Leuven guidelines on responsible use of Generative AI tools. Depending on the kind of GenAI use, it may be needed to properly reference it in the text, in accordance with the instructions page on being transparent about the use of GenAI.