Long gravitational-wave transients and associated detection strategies for a network of terrestrial interferometers

Eric Thrane, Shivaraj Kandhasamy, Christian D. Ott, Warren G. Anderson, Nelson L. Christensen, Michael W. Coughlin, Steven Dorsher, Stefanos Giampanis, Vuk Mandic, Antonis Mytidis, Tanner Prestegard, Peter Raffai, and Bernard Whiting

Long gravitational-wave transients and associated detection strategies for a network of terrestrial interferometers

Eric Thrane,[1, a] Shivaraj Kandhasamy,[1] Christian D. Ott,[2] Warren G. Anderson,[3]
Nelson L. Christensen,[4] Michael W. Coughlin,[4] Steven Dorsher,[1] Stefanos Giampanis,[5, 3]
Vuk Mandic,[1] Antonis Mytidis,[6] Tanner Prestegard,[1] Peter Raffai,[7] and Bernard Whiting[6]

[1]*School of Physics and Astronomy, University of Minnesota, Minneapolis, MN 55455, USA*
[2]*TAPIR, Caltech, Pasadena, California 91125*
[3]*University of Wisconsin-Milwaukee, Milwaukee, Wisconsin 53201*
[4]*Physics and Astronomy, Carleton College, Northfield, Minnesota 55057*
[5]*Albert-Einstein-Institut, Max-Planck-Institut für Gravitationsphysik, D-30167 Hannover, Germany.*
[6]*Department of Physics, University of Florida, Gainsville, FL 32611, USA*
[7]*Institute of Physics, Eötvös University, 1117 Budapest, Hungary*

Searches for gravitational waves (GWs) traditionally focus on persistent sources (e.g., pulsars or the stochastic background) or on transients sources (e.g., compact binary inspirals or core-collapse supernovae), which last for timescales of milliseconds to seconds. We explore the possibility of *long* GW transients with unknown waveforms lasting from many seconds to weeks. We propose a novel analysis technique to bridge the gap between short $\mathcal{O}$(s) "burst" analyses and persistent stochastic analyses. Our technique utilizes frequency-time maps of GW strain cross-power between two spatially separated terrestrial GW detectors. The application of our cross-power statistic to searches for GW transients is framed as a pattern recognition problem, and we discuss several pattern-recognition techniques. We demonstrate these techniques by recovering simulated GW signals in simulated detector noise. We also recover environmental noise artifacts, thereby demonstrating a novel technique for the identification of such artifacts in GW interferometers. We compare the efficiency of this framework to other techniques such as matched filtering.

PACS numbers: 95.55.Ym

I. INTRODUCTION

Historically, searches for gravitational-wave (GW) transients fall into one of two categories: searches for "bursts" whose precise waveforms we cannot predict and searches for compact binary coalescences, whose waveforms can be predicted (at least for the inspiral part). Typically, burst searches focus on events with $\lesssim 1\,$s durations and, indeed, there are many compelling models for short GW transients (see [1] and references therein).

In this paper, we put the spotlight on *long* GW transients whose durations may range from many seconds to weeks. Astrophysical GW emission scenarios for long transients exist (e.g., [2–5]), but their characteristics have not previously been broadly addressed and no data-analysis strategy has been proposed for such events until now. (In addition to this work, see recent developments in [6].) Most of the GW emission models we consider are burst-like in the sense that the signal evolution cannot be precisely predicted, however, we refer to them as "transients" to avoid connoting that they are short-duration.

In Sec. II, we survey a range of mechanisms for GW emission that may lead to long transients. These include long-lived turbulent convection in protoneutron stars (PNSs), rotational instabilities in rapidly spinning PNSs and in double neutron-star merger remnants, mag-netoturbulence and gravitational instabilities in gamma-ray burst (GRB) accretion torii, r-modes associated with accreting and newborn neutron stars, as well as, perhaps more speculatively, pulsar glitches and soft-gamma-repeater (SGR) outbursts.

In Sec. III, we introduce an analysis framework utilizing frequency-time (ft)-maps of GW strain cross-power created using data from two or more spatially separated detectors. The framework is extended to include multiple detectors, and we show that it is a generalization of the GW radiometer algorithm [7]. In Sec. IV, we compare ft-cross-power maps of GW data (time-shifted to remove astrophysical content) with Monte Carlo simulations of idealized detector noise. We shall see that GW interferometer data is well-behaved enough that thresholds for candidate events can be estimated analytically (in at least one case).

In Sec. V, we use ft-cross-power maps to cast the search for long GW transients as a pattern recognition problem. For the sake of concreteness, we consider two algorithms: a "box search" [8] and a Radon algorithm [9]. In Sec. VI, we demonstrate the Radon algorithm (as well as the "locust" and Hough algorithms [10]) to identify environmental noise artifacts in LIGO environmental monitoring channels—a novel technique for the identification of such artifacts in GW interferometers. In Sec. VII, we describe how our framework is related to other detection strategies such as matched filtering. Concluding remarks are given in Sec. VIII.

[a]Electronic address: `ethrane@physics.umn.edu`

model	source
PNS convection	core collapse
rot. instabilities	BNS coalescence, core collapse, isolated NS
r-modes	core collapse, isolated NS
disk instabilities	BNS coalescence, core collapse
high-ϵ BH binaries	BBH coalescence
pulsar glitches	isolated NS
SGR flares	isolated NS

TABLE I: Models of long GW transients with associated sources. BNS and BBH stand for "binary neutron star" and "binary black hole" respectively.

II. ASTROPHYSICAL SOURCES OF LONG GW TRANSIENTS

In this section, we review a variety of emission mechanisms for long GW transients. Most of the mechanisms we consider (summarized in Tab. I) are associated with one or more of three types of objects: core-collapse supernovae (CCSNe), compact binary inspirals, or isolated neutron stars.

A. Core-collapse supernovae and long gamma-ray bursts

There is tremendous electromagnetic observational evidence connecting both CCSNe and long gamma-ray bursts (GRBs) to the core-collapse death of massive stars (see, e.g., [11]). Both are ultimately powered by the release of gravitational energy, but the precise mechanism by which gravitational energy is converted into energy of ejecta and radiation is uncertain in both phenomena (see, e.g., [11–13] and references therein). However, all modern models of CCSN and long-GRB central engines involve violent non-spherical dynamics, making both systems prodigious emitters of GWs.

The GW signature of CCSNe (recently reviewed in [2]) may be composed of contributions from rotating collapse and core bounce [14], post-bounce protoneutron star (PNS) convection [2, 15, 16], neutrino-driven convection and the standing-accretion-shock instability (SASI) [17–19], PNS pulsations [20], nonaxisymmetric rotational instabilities (both dynamical and secular) [21, 22], asymmetric neutrino emission [2, 18, 19], aspherical outflows [23–27], magnetic stresses [26, 27], and r-mode pulsations in rotating PNSs (see, e.g., [28, 29]). Depending on the particular CCSN mechanism in operation, some emission processes may dominate while others are suppressed [13].

Currently, there are two favored scenarios for the long-GRB central engine. In the collapsar scenario [30], massive stars collapse to black holes either without an initial CCSN exploding or via fallback accretion after a successful, but weak explosion. The millisecond-protomagnetar scenario [31, 32] relies on highly magnetized, nascent neutron stars. In both cases, any long-GRB activity is preceded by stellar collapse and a post-bounce phase during which a PNS exists and GW emission occurs in very similar fashion to regular CCSNe. In the collapsar scenario, a black hole with an accretion disk forms. Magnetohydrodynamical processes and/or neutrino pair annihilation powered by accretion and/or by the extraction of black hole spin energy eventually launch the GRB jet. GWs may be emitted by disk turbulence and disk instabilities that may lead to clumping or disk fragmentation [4, 5]. In the millisecond-magnetar scenario, a successful magneto-rotational CCSN explosion (see, e.g., [33, 34]) occurs, after which a high-Lorentz-factor outflow is driven by the millisecond protomagnetar. GWs may be emitted by convective/meridional currents and dynamical and secular nonaxisymmetric rotational instabilities in the protomagnetar [3, 35].

In CCSNe and in the CCSN-phase of long GRBs most GW emission processes last until the onset of the CCSN explosion or until PNS-collapse to a black hole, and hence they have a short duration of order $\lesssim 1 - 2\,\text{s}$ [36, 37]. Exceptions are PNS convection, secular rotational instabilities including r-modes and long-GRB disk/torus instabilities. We discuss these below and provide order-of-magnitude estimates of their emission characteristics in the time and frequency domains.

1. Protoneutron star convection

If a CCSN explosion occurs, a stable PNS is left behind and will cool on a Kelvin-Helmholtz timescale (see, e.g., [38]). Fallback accretion [39], or, perhaps, a late-time hadron-quark phase transition (e.g., [40]) may lead to PNS collapse and black hole formation. If the PNS survives, a powerful convective engine, driven by thermal and lepton gradients may continue to operate for possibly tens of seconds in the cooling phase [15, 41–43], making it a long GW transient source.

GW emission estimates from PNS convection are based on results of simulations that track only the early phase ($\lesssim 1\,\text{s}$ after core bounce) [2, 16, 44], yet they have found a number of robust features that translate to later times. PNS convection occurs at moderate to high Reynolds numbers, hence, is turbulent and leads to an incoherent, virtually stochastic GW signal. Its polarization is random in the nonrotating or slowly rotating case, but may assume specific polarization due to axisymmetric rotationally-driven meridional currents in rapidly spinning PNSs (an effect that remains to be studied in computational models). In the phase covered by current models, typical GW strains are $h \sim 3\times10^{-23}$ at a galactic distance of 10 kpc [2, 16]. ("Strain" refers to the strain measured at Earth; strain amplitude scales like the inverse of the distance from the source.)

While on short timescales, the GW signal of PNS convection will appear almost as a white-noise burst, its time-frequency structure is non-trivial, exhibiting a

broad spectral peak at $\mathcal{O}(100\,\mathrm{Hz})$, which shifts to higher frequencies over the course of the first second after core bounce [2, 16]. This chirp-like trend is likely to continue for seconds afterward as the PNS becomes more compact. It should be nearly independent of the size of the convectively unstable region, since the eddy size will be set by the local pressure scale height (see, e.g., [45]).

Based on the $\sim 1.2\,\mathrm{s}$ evolution of a PNS model by [16, 41], we expect a total emitted energy of $\sim 1.6 \times 10^{-10}\,M_\odot c^2$. Assuming that convective GW emission continues with comparable vigor for tens of seconds, we can scale this to $E_{\mathrm{GW}} \sim 4 \times 10^{-9}\,(\Delta t/30\,\mathrm{s})\,M_\odot c^2$.

2. Rotational instabilities

Most massive stars ($\gtrsim 98\%$ or so, see [46] and references therein) are likely to be slow rotators, making PNSs with birth spin periods of $\sim 10 - 100\,\mathrm{ms}$. GRB progenitors, however, are most likely rapidly spinning, leading to PNSs with birth spins of $\mathcal{O}(1\,\mathrm{ms})$ and rotational kinetic energy of up to $10^{52}\,\mathrm{erg}$ [33], enough to power a long GRB through protomagnetar spin-down as suggested by the protomagnetar model [31, 32].

PNSs, like any self-gravitating (rotating or nonrotating) fluid body, tend to evolve toward a state of minimal total energy. PNSs are most likely born with an inner core in solid-body rotation and an outer region that is strongly differentially rotating [46]. Magnetorotational instabilities (see, e.g., [47]) and/or hydrodynamic shear instabilities (see, e.g., [48]) will act to redistribute angular momentum toward uniform rotation (the lowest-energy state). The latter type of instability may lead to significant, though short-term $\tau \lesssim 1\,\mathrm{s}$, nonaxisymmetric deformation of parts of the PNS and, as a consequence, to significant GW emission [2, 21, 49]. PNSs in near solid-body rotation that exceed certain values of the ratio of rotational kinetic to gravitational energy, $T/|W|$, may deform from an axisymmetric shape to assume more energetically-favorable triaxial shape of lowest-order $l = m = 2$, corresponding to a spinning bar. Such a bar is a copious emitter of GWs. At $T/|W| \gtrsim 0.27$, nonaxisymmetric deformation occurs dynamically, but will not last longer than a few dynamical times of $\mathcal{O}(\mathrm{ms})$ (see, e.g., [50, 51]) due to rapid angular momentum redistribution, and hence we do not consider its GW emission in this study.

At $T/|W| \gtrsim 0.14$, a secular gravitational-radiation reaction or viscosity-driven instability may set in, also leading to nonaxisymmetric deformation. The timescale for this depends on the detailed PNS dynamics as well as the details and strength of the viscosity in the PNS. It is estimated to be $\mathcal{O}(1\,\mathrm{s})$ for both driving agents, but the expectation is that gravitational-radiation reaction dominates over viscosity [52, 53]. The secular instability has the potential of lasting for $\sim 10 - 100\,\mathrm{s}$ [3, 52], and hence it is of particular interest for our present study.

Once the gravitational-radiation reaction instability sets in, the initially axisymmetric PNS slowly deforms into $l = m = 2$ bar shape and, in the ideal Dedekind ellipsoid limit, evolves toward zero pattern speed (angular velocity $\Omega = 0$) with its remaining rotational energy being stored as motion of the fluid in highly noncircular orbits inside the bar [22, 52]. GW emission occurs throughout the secular evolution with strain amplitudes h proportional to Ω^2 and to the ellipticity ϵ, characterizing the magnitude of the bar deformation, leading to an initial rise of the characteristic strain followed by slow decay as Ω decreases [22, 52]. We expect characteristic strain amplitudes, defined as $h_c \equiv \tau_{\mathrm{GW}} f_{\mathrm{GW}} h$, of $h_c = \mathcal{O}(10^{-23} \sim 10^{-22}\,\mathrm{Hz}^{-1/2})$ for a source located at $100\,\mathrm{Mpc}$ and the emission is expected to last at that level for $\mathcal{O}(100\,\mathrm{s})$ [3, 52]. The emitted GWs will be elliptically polarized.

3. R-modes

R-modes are quasi-toroidal oscillations that have the Coriolis force as their restoring force. It was shown in [28, 54] that r-modes in neutron stars are unstable to growth at all rotation rates by gravitational-radiation reaction via the secular Chandrasekhar-Friedman-Schutz instability [55, 56]. R-modes emit (at lowest order) current-quadrupole GWs with $f_{\mathrm{GW}} = 4/3\,(\Omega_{\mathrm{NS}}/2\pi)$ and typical strain amplitudes $h \sim 4.4 \times 10^{-24}\alpha(\Omega_{\mathrm{NS}}/\sqrt{\pi G\bar\rho})^3(20\,\mathrm{Mpc}/D)$ [57], where Ω_{NS} is the NS angular velocity, D is the distance to the source and $\bar\rho$ is the mean neutron star density. The parameter $\alpha \in [0, 1]$ is the dimensionless saturation amplitude of the r-modes and its true value has been the topic of much debate. Most recent work suggests (see [29, 58] and references therein) that $\alpha \ll 0.1$ and, perhaps, does not exceed $\sim 10^{-5}$ due to non-linear mode coupling effects [59]. Generally, r-modes are expected to be a source of very-long-lasting quasi-continuous GW emission, though long GW transients may be possible in the case of high saturation amplitudes (e.g., [60]).

Potential astrophysical sources of GWs from r-modes are accreting neutron stars in low-mass X-ray binaries (e.g., [58, 61, 62]) and, more relevant in the present context, newborn, rapidly spinning neutron stars [28, 29, 57]. In the latter, r-modes may play an important role in the early spin evolution [29, 63].

4. Accretion disk instabilities

In the long GRB collapsar scenario, the central engine consists of a black hole surrounded by an accretion disk/torus [11, 30]. The inner part of the disk is likely to be sufficiently hot to be neutrino cooled and thin [64] while the outer regions with radius $r \gtrsim 50R_{\mathrm{S}} = 100\,GM_{\mathrm{BH}}c^{-2}$ are cooled inefficiently and form a thick accretion torus [4, 64]. A variety of (magneto)-hydrodynamic instabilities may occur in the disk/torus

leading to the emission of GWs.

Piro and Pfahl [4] considered gravitational instability of the outer torus leading to fragmentation facilitated by efficient cooling through helium photodisintegration. Multiple fragments may collapse to a single big dense fragment of up to $\sim 1 M_\odot$ that then travels inward either by means of effective viscosity and/or GW emission. In both cases, the inspiral will last $\mathcal{O}(10 - 100\,\mathrm{s})$, making it a prime candidate source for a long GW transient. Piro and Pfahl predict maximum dimensionless strain amplitudes $|h| \sim 2 \times 10^{-23} (f_{\mathrm{GW}}/1000\,\mathrm{Hz})^{2/3}$ for emission of a system with a fragment mass of $1 M_\odot$, a black hole mass of $8 M_\odot$, and a source distance of $100\,\mathrm{Mpc}$. The frequency will slowly increase over the emission interval, making the emission quasi-periodic and, thus, increasing its detectability by increasing its characteristic strain h_c up to $\mathcal{O}(\times 10^{-22})$ at $f_{\mathrm{GW}} \sim 100\,\mathrm{Hz}$ and a source distance of $100\,\mathrm{Mpc}$.

Van Putten, in a series of articles (see, e.g., [5, 65, 66] and references therein), has proposed an extreme "suspended–accretion" scenario in which the central black hole and the accretion torus are dynamically linked by strong magnetic fields. In this picture, black hole spin-down drives both the actual GRB central engine and strong magnetoturbulence in the torus, leading to a time-varying mass quadrupole moment and, thus, to the emission of GWs. Van Putten postulated, based on a simple energy argument, branching ratios of emitted GW energy to electromagnetic energy of $E_{\mathrm{GW}}/E_{\mathrm{EM}} \gtrsim 100$ and thus, $E_{\mathrm{GW}} \sim$ few $\times 10^{53}\,\mathrm{erg}$. These numbers are perhaps unlikely to obtain in nature, but the overall concept of driven magnetoturbulence is worth considering.

In van Putten's theory, the magnetoturbulent torus excitations produce narrowband elliptically polarized GWs with a frequency between $(1 \sim 2\,\mathrm{kHz})(1 + z)$ for a GRB at a redshift of z [65]. The frequency is predicted to vary with time such that $df/dt = \mathrm{const}$ [65]. GW emission continues for approximately the same duration as the electromagnetic emission, lasting typically for seconds to minutes [11]. With van Putten's energetics, a long GRB at a distance of $100\,\mathrm{Mpc}$ is predicted to produce a strain of $h \sim 10^{-23}$, which is comparable to the expected Advanced LIGO noise at $1000\,\mathrm{Hz}$ [65]. Integrating many seconds of data, such a loud signal should stand out above the Advanced LIGO noise, making it likely that a strong statement can be made about this model in the advanced detector era.

B. Postmerger evolution of double neutron-star coalescence

In Subsec. II A, we discussed a variety of scenarios for long GW transients in the context of PNS and black-hole – accretion-disk systems left in the wake of CCSNe and in collapsars. A similar situation is likely to arise in the postmerger stage of double neutron-star coalescence. The initial remnant will be a hot supermassive neutron star that, depending on the mass of the binary constituents and on the stiffness of the nuclear equation of state, may survive for hundreds of milliseconds (e.g., [67] and references therein). In these systems, many of the GW emission mechanisms discussed in the stellar collapse scenario may be active. Hence, it may be fruitful to search for long GW transients following observed inspiral events as well as following short GRBs, (which are expected to be associated with binary inspirals).

Inspiral events, however, need not invoke the formation of a PNS in order to produce a long GW transient. Highly eccentric black hole binary inspirals are expected to produce complicated waveforms that are difficult to model with matched filtering and may persist for hundreds of seconds [68, 69], and thus they are suitable candidates for long GW transient searches. According to some models [70], a significant fraction of BBH form dynamically with high eccentricities ($\epsilon > 0.9$) leading to an Advanced LIGO event rate of $\sim 1 - 100\,\mathrm{yr}^{-1}$. Given such high rates, it is highly likely that these models can be thoroughly probed in the advanced detector era.

C. Isolated neutron stars

Isolated neutron stars are another potential source of long GW transients. In the following, we discuss pulsar glitches and soft-gamma repeater flares as potential sources of GWs.

Pulsar glitches are sudden speed-ups in the rotation of pulsing neutron stars observed by radio and X-ray observatories. The fractional change in rotational frequency ranges from $10^{-10} < \Delta f/f < 5 \times 10^{-6}$, corresponding to rotational energy changes of $\lesssim 10^{43}\,\mathrm{erg}$ [71, 72]. The speed-up, which takes place in $< 2\,\mathrm{min}$, is followed by a period of relaxation (typically weeks) during which the pulsar slows to its pre-glitch frequency [73].

The mechanism by which pulsar glitches occur is a matter of ongoing research [74–78], and the extent to which they emit GWs is unknown. We therefore follow Andersson et al. [79] and assume that the emitted energy in GWs is comparable to the change in rotational energy. Given these energetics, and assuming a simple exponentially-decaying damped waveform, a nearby ($d = 1\,\mathrm{kpc}$) glitch can produce, e.g., a $\mathcal{O}(10\,\mathrm{s})$ quadrupole excitation with a strain of $h \sim 8 \times 10^{-24}$ at $3.8\,\mathrm{kHz}$ [79]. This is about six times below the Advanced LIGO noise floor, which effectively rules out the possibility of detection. A long GW event measured by Advanced LIGO and coincident with a pulsar glitch would therefore suggest a radically different glitch mechanism than the one considered in [79].

Flares from soft-gamma repeaters (SGRs) and anomalous X-ray pulsars (AXPs), which may be caused by seismic events in the crusts of magnetars, have also been proposed as sources of GWs. Recent searches by LIGO have set limits on lowest-order quadrupole ringdowns in SGR storms [80] and in single-SGR events [81]. SGR

giant flares are associated with huge amounts of electromagnetic energy ($10^{44} - 10^{46}$ erg), and they are followed by a $\mathcal{O}(100\,\mathrm{s})$-long tail characterized by quasi-periodic oscillations (see, e.g., [82]). It is hypothesized that quasi-periodic oscillations following SGR flares may emit GWs through the excitation of torsional modes [83].

Current models of GW from SGRs/AXPs [84–91] are very preliminary, but even if we assume that only 0.1% of the 10^{46} erg of electromagnetic energy in a nearby SGR flare is converted into GWs, then SGRs are nonetheless attractive targets in the advanced detector era. Current experiments have set limits on the emission of GW energy from SGRs (in the form of short bursts) at a level of $E_{\mathrm{GW}} \lesssim 10^{45}$ erg [80, 81] ($\approx 10\%$ of the electromagnetic energy of a giant flare depending on the waveform type). Since sensitivity to $E_{\mathrm{GW}} \propto h^2$, it is likely that we can probe interesting energy scales ($E_{\mathrm{GW}} \approx 0.1\% \, E_{\mathrm{EM}}$) in the advanced detector era.

III. AN EXCESS CROSS-POWER STATISTIC

A. Definitions and conventions

Our present goal is to develop a statistic $\hat{\mathcal{Y}}_\Gamma$ which can be used to estimate the GW power H_Γ (or the related quantities of GW fluence F_Γ and energy E_Γ) associated with a long GW transient event confined to some set of discrete frequencies and times Γ. In order to define GW power, we first note the general form of a point-source GW field in the transverse-traceless gauge:

$$h_{ab}(t, \vec{x}) = \sum_A \int_{-\infty}^{\infty} df \, e_{ab}^A(\hat{\Omega}) \tilde{h}_A(f) \, e^{2\pi i f(t + \hat{\Omega} \cdot \vec{x}/c)}. \quad (3.1)$$

Here $\hat{\Omega}$ is the direction to the source, A is the polarization state and $\{e_{ab}^A\}$ are the GW polarization tensors with Cartesian indices ab, (see App. A 1 for additional details). Since Eq. 3.1 describes an astrophysical source, the Fourier transform of the strain $\tilde{h}(f)$ is defined in the continuum limit.

We now, however, consider a discrete measurement on the interval between t and $t + T$ measured with a sampling frequency of f_s, which corresponds to $N_s = f_s T$ independent measurements. We utilize a discrete Fourier transformations denoted with tildes:

$$\tilde{q}_k \equiv \frac{1}{N_s} \sum_{n=0}^{N_s-1} q_n e^{-2\pi i n k / N_s}$$

$$q_n \equiv \sum_{k=0}^{N_s-1} \tilde{q}_k e^{2\pi i n k / N_s}. \quad (3.2)$$

The variable t—separated from other arguments with a semicolon—refers to the segment start time, as opposed to individual sampling times, denoted by t with no semicolon, e.g., $s(t)$. Variables associated with discrete measurements are summarized in Tab. II.

f_s	sampling frequency
$\delta t \equiv 1/f_s$	sampling time
t	segment start time
T	segment duration
δf	frequency resolution
N_s	number of sampling points in one segment

TABLE II: Variables describing discrete measurements.

The GW strain power spectrum (measured between t and $t + T$ with a sampling frequency f_s in a frequency band between f and $f + \delta f$) is:

$$H_{AA'}(t; f) = 2\langle \tilde{h}_A^*(t; f) \tilde{h}_{A'}(t; f) \rangle. \quad (3.3)$$

The factor of 2 comes from the fact that $H_{AA'}(t; f)$ is the one-sided power spectrum.

It is convenient to characterize the source with a single spectrum that includes contributions from both $+$ and $\times$ polarizations. We therefore define

$$H(t; f) \equiv \mathrm{Tr}\left[H_{AA'}(t; f)\right], \quad (3.4)$$

so as to be invariant under change of polarization bases. This definition is a generalization of the one-sided power spectrum for unpolarized sources found in [7, 92, 93].

Our estimator $\hat{\mathcal{Y}}_\Gamma(\Omega)$ utilizes frequency-time (ft)-maps: arrays of pixels each with a duration determined by the length of a data segment T and by the frequency resolution δf. In Subsec. III D, we describe how $\hat{\mathcal{Y}}_\Gamma$ can be constructed by combining *clusters* of ft-map pixels. We thereby extend the stochastic-search formalism developed in [7, 92, 93] beyond models of persistent unpolarized sources to include polarized and unpolarized transient sources. In doing so, we endeavor to bridge the gap between searches for short $\mathcal{O}(\mathrm{s})$ signals and stochastic searches for persistent GWs. We begin by considering just one pixel in the ft-map.

B. A single ft-map pixel

In App. A 2, we derive the form of an estimator $\hat{Y}$ for GW power $H(t; f)$ in a single ft-pixel by cross-correlating the strain time series $s_I(t)$ and $s_J(t)$ from two spatially separated detectors, I and J, for a source at a sky position $\hat{\Omega}$ [118]. We find that

$$\hat{Y}(t; f, \hat{\Omega}) \equiv \mathrm{Re}\left[\tilde{Q}_{IJ}(t; f, \hat{\Omega}) \, C_{IJ}(t; f)\right]. \quad (3.5)$$

Here, $C_{IJ}(t; f)$ is the one-sided cross-power spectrum

$$C_{IJ}(t; f) \equiv 2 \, \tilde{s}_I^*(t; f) \tilde{s}_J(t; f). \quad (3.6)$$

Meanwhile, $\tilde{Q}_{IJ}(t; f, \hat{\Omega})$ is a filter function, which depends, among other things, on the source direction and

polarization. For unpolarized sources (see App. A 2),

$$\tilde{Q}_{IJ}(t; f, \hat{\Omega}) = \frac{1}{\epsilon_{IJ}(t; \hat{\Omega})} e^{2\pi i f \hat{\Omega} \cdot \Delta \vec{x}_{IJ}/c}. \qquad (3.7)$$

where $\epsilon_{IJ}(t; \hat{\Omega}) \in [0, 1]$, the "pair efficiency," is

$$\epsilon_{IJ}(t; \hat{\Omega}) \equiv \frac{1}{2} \sum_A F_I^A(t; \hat{\Omega}) F_J^A(t; \hat{\Omega}). \qquad (3.8)$$

Here $F_I^A(t; \hat{\Omega})$ is the "antenna factor" for detector I and $\Delta \vec{x}_{IJ} \equiv \vec{x}_I - \vec{x}_J$ is the difference in position vectors of detectors I and J; (see App. A 1). Pair efficiency is defined such that a GW with power H will induce IJ strain cross-power given by $\epsilon_{IJ} H$. It is unity only in the case where both interferometers are optimally oriented so that the change in arm length is equal to the strain amplitude. For additional details (including a derivation of the pair efficiency for polarized sources) see Apps. A 1, A 2 and A 5.

The variance of $\hat{Y}$ is calculated in App. A 3. Then in App. A 4, we show that the following expression for $\hat{\sigma}_Y^2(t; f, \hat{\Omega})$ (motivated by analogy with stochastic analyses [92]) is an estimator for the variance of $\hat{Y}$,

$$\hat{\sigma}_Y^2(t; f, \hat{\Omega}) = \frac{1}{2} |\tilde{Q}_{IJ}(t; f, \hat{\Omega})|^2 P_I^{\text{adj}}(t; f) P_J^{\text{adj}}(t; f), \quad (3.9)$$

where P_I^{adj} is the average one-sided auto-power spectrum in neighboring pixels,

$$P_I^{\text{adj}}(t; f) \equiv 2 \overline{|\tilde{s}_I(t; f)|^2}. \qquad (3.10)$$

The overline denotes an average over neighboring pixels [119].

From Eqs. 3.9 and 3.5, we define the signal to noise ratio $\text{SNR}(t; f, \hat{\Omega})$ for a single ft-map pixel:

$$\text{SNR}(t; f, \hat{\Omega}) \equiv \hat{Y}(t; f, \hat{\Omega})/\hat{\sigma}_Y(t; f, \hat{\Omega})$$
$$= \text{Re} \left[\frac{\tilde{Q}_{IJ}(t; f, \hat{\Omega})}{|\tilde{Q}_{IJ}(t; f, \hat{\Omega})|} \frac{C_{IJ}(t; f)}{\sqrt{\frac{1}{2} P_I^{\text{adj}} P_J^{\text{adj}}}} \right] \quad (3.11)$$

It depends on the phase of $\tilde{Q}_{IJ}(t; f, \hat{\Omega})$, but not on the magnitude. Thus, a single ft-pixel taken by itself contains no information about the polarization properties of the source, since the polarization does not affect the phase of $\tilde{Q}$. This degeneracy is broken when we combine ft-pixels from different times or from different detector pairs.

C. Energy, fluence and power

One of the most interesting intrinsic properties of a transient source of GWs is the total energy emitted in gravitational radiation, E_{GW}. By measuring E_{GW} (and,

when possible, comparing it to the observed electromagnetic energy, E_{EM}), we can make and test hypotheses about the total energy associated with the event as well as constrain models of GW production. Thus, it is useful to relate $\hat{Y}(t; f, \hat{\Omega})$ to E_{GW} and the related quantity of fluence. If the GW energy is emitted isotropically (in general it is not) then [94],

$$E_{\text{GW}} = 4\pi R^2 \frac{c^3}{16\pi G} \int dt \left(\dot{h}_+^2(t) + \dot{h}_\times^2(t) \right), \qquad (3.12)$$

where R is the distance to the source. It follows that the equivalent isotropic energy is related to our cross-power estimator as follows:

$$\hat{E}_{\text{GW}}(t; f, \hat{\Omega}) = 4\pi R^2 \frac{\pi c^3}{4G} \left(T f^2 \right) \hat{Y}(t; f, \hat{\Omega}). \qquad (3.13)$$

$\hat{E}_{\text{GW}}(t; f, \hat{\Omega})$ may contain significant uncertainty about the distance to the source or the isotropy of the GW emission. It is therefore useful to define a statistic that contains only uncertainty associated with the strain measurement. The natural solution is to construct a statistic for GW fluence, $\hat{F}_{\text{GW}}(t; f, \hat{\Omega})$, which is given by

$$\hat{F}_{\text{GW}}(t; f, \hat{\Omega}) = \frac{\hat{E}_{\text{GW}}(t; f\hat{\Omega})}{4\pi R^2}$$
$$= T f^2 \left(\frac{\pi c^3}{4G} \right) \hat{Y}(t; f, \hat{\Omega}). \qquad (3.14)$$

In the subsequent section, we show how multiple pixels can be combined to calculate the average power inside some set of pixels. The same calculation can be straightforwardly extended to calculate the total fluence. This is done by reweighting $\hat{Y}(t; f, \hat{\Omega})$ and $\hat{\sigma}(t; f, \hat{\Omega})$ by $(\pi c^3/4G)(T f^2)$. Also Eqs. 3.16 and 3.18 must be scaled by the number of pixels in a set, N; (otherwise we obtain average fluence instead of total fluence).

D. Multi-pixel statistic

We now generalize from our single-pixel statistic to accommodate transients persisting over N pixels in some set of pixels, Γ. We define H_Γ to be the average power inside Γ,

$$H_\Gamma \equiv \frac{1}{N} \sum_{t; f \in \Gamma} H(t; f). \qquad (3.15)$$

A minimum-variance estimator for the GW power in Γ can be straightforwardly constructed from a weighted sum of $\hat{Y}(t; f, \hat{\Omega})$ for each pixel in Γ,

$$\hat{\mathcal{Y}}_\Gamma(\hat{\Omega}) = \frac{\sum_{t; f \in \Gamma} \hat{Y}(t; f, \hat{\Omega}) \hat{\sigma}_Y(t; f, \hat{\Omega})^{-2}}{\sum_{t; f \in \Gamma} \hat{\sigma}_Y(t; f, \hat{\Omega})^{-2}}. \qquad (3.16)$$

Here we assume that the power is either evenly or randomly distributed inside Γ, which is to say $\langle H(t; f) \rangle =$

$\langle H(t'; f') \rangle \equiv H_0$ and so $\langle H_\Gamma \rangle = H_0$. Thus,

$$
\begin{aligned}
\langle \hat{\mathcal{Y}}_\Gamma(\hat{\Omega}) \rangle &= \left\langle \frac{\sum_{t;f\in\Gamma} \hat{Y}(t;f,\hat{\Omega})\,\hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}}{\sum_{t;f\in\Gamma} \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}} \right\rangle \\
&= \frac{\sum_{t;f\in\Gamma} \langle \hat{Y}(t;f,\hat{\Omega}) \rangle\, \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}}{\sum_{t;f\in\Gamma} \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}} \\
&= H_0 \left(\frac{\sum_{t;f\in\Gamma} \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}}{\sum_{t;f\in\Gamma} \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2}} \right) = \langle H_\Gamma \rangle.
\end{aligned}
\tag{3.17}
$$

Here we have additionally assumed that there are no correlations between $\hat{Y}(t;f,\hat{\Omega})$ in different pixels. If the GW signal in different pixels is correlated, then the $\{\hat{Y}(t;f,\hat{\Omega})\}$ are correlated and Eq. 3.16 should, in theory, be modified to include covariances between different pixels. In practice, however, the covariance matrix is not known, and so we must settle for this approximation, which gives the estimator a higher variance than could be achieved if the covariance matrix was known.

The associated estimator for the uncertainty is

$$
\hat{\sigma}_\Gamma(\hat{\Omega}) = \left(\sum_{t;f\in\Gamma} \hat{\sigma}_Y(t;f,\hat{\Omega})^{-2} \right)^{-1/2}.
\tag{3.18}
$$

The choice of the set of pixels Γ to include in the sum in Eq. 3.16 is determined by the signal model. For example, a slowly varying narrowband signal can be modeled as a line of pixels on the ft-map. We explore this and other choices for Γ in greater detail in Sec. V.

The SNR for given a set of pixels Γ is given by

$$
\text{SNR}_\Gamma(\hat{\Omega}) = \frac{\hat{\mathcal{Y}}_\Gamma(\hat{\Omega})}{\hat{\sigma}_\Gamma(\hat{\Omega})}.
\tag{3.19}
$$

Since SNR_Γ is the weighted sum of many independent measurements, we expect, due to the central limit theorem, that the distribution of SNR_Γ will be increasingly well-approximated by a normal distribution as the volume of Γ increases and more pixels are included in the sum [120].

E. Multi-detector statistic

It is straightforward to generalize $\hat{\mathcal{Y}}_\Gamma$ for a detector network $\mathcal{N}$ consisting of $n \geq 2$ spatially separated detectors. First, we generate $n(n-1)/2$ ft-maps for each pair of interferometers. Then we extend the sum over pixels in Eq. 3.16 to include a sum over unique detector pairs $p(I,J)$:

$$
\hat{\mathcal{Y}}_\Gamma^{\mathcal{N}}(\hat{\Omega}) = \frac{\sum_{p(I,J)} \sum_{t;f\in\Gamma} \hat{Y}_{IJ}(t;f,\hat{\Omega})\hat{\sigma}_{IJ}(t;f,\hat{\Omega})^{-2}}{\sum_{p(I,J)} \sum_{t;f\in\Gamma} \hat{\sigma}_{IJ}(t;f,\hat{\Omega})^{-2}}.
\tag{3.20}
$$

By construction, the expectation value is

$$
\langle \hat{\mathcal{Y}}_\Gamma^{\mathcal{N}} \rangle = H_\Gamma.
\tag{3.21}
$$

The associated uncertainty is

$$
\hat{\sigma}_{\mathcal{Y}}^{\mathcal{N}}(\hat{\Omega}) = \left(\sum_{p(I,J)} \sum_{ft} \hat{\sigma}_{IJ}(t;f,\hat{\Omega})^{-2} \right)^{-1/2}.
\tag{3.22}
$$

Adding new detectors to the network improves the statistic by mitigating degeneracies in sky direction and polarization parameters and also by improving sensitivity to H_Γ by increasing the number of pixels contributing to $\hat{\mathcal{Y}}_\Gamma^{\mathcal{N}}$.

F. Relationship to the GW radiometer

The multi-pixel statistic $\hat{\mathcal{Y}}_\Gamma$ is straightforwardly related to the GW radiometer technique, which has been used to look for GWs from neutron stars in low-mass X-ray binaries [7]. By constructing a rectangular set of pixels consisting of one or more frequency bins and lasting the entire duration of a science run, we recover the radiometer statistic as a special case.

It is instructive to compare the unpolarized radiometer statistic [7] with our $\hat{\mathcal{Y}}_\Gamma$:

$$
\hat{Y}^{\text{rad}}(t;f,\hat{\Omega}) \equiv \int_{-\infty}^{\infty} df \tilde{Q}_{IJ}^{\text{rad}}(t;f,\hat{\Omega}) \\
\tilde{s}_I^\star(t;f)\tilde{s}_J(t;f)
\tag{3.23}
$$

$$
\tilde{Q}_{IJ}^{\text{rad}}(t;f,\hat{\Omega}) \equiv \lambda_t \frac{\gamma(t;f,\hat{\Omega})\bar{H}(f)}{P_I(f)P_J(f)}
\tag{3.24}
$$

$$
\gamma_{IJ}(t;f,\hat{\Omega}) \equiv \epsilon_{IJ}(t;\hat{\Omega})\, e^{2\pi i f \hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c}.
\tag{3.25}
$$

Here $\gamma_{IJ}(t;f,\hat{\Omega})$ is the so-called overlap reduction factor, λ_t is a normalization factor and $\epsilon_{IJ}(t;\hat{\Omega})$ is the pair efficiency, which we define in Eq. A20 and Eq. 3.8.

There are two things worth noting here. First, the extra factor of $H(f)/P_I(f)P_J(f)$ in the expression for $\tilde{Q}_{IJ}^{\text{rad}}$ does not appear in our expression for $\tilde{Q}_{IJ}$ (see Eq. A19). The factor of $1/P_I(f)P_J(f)$ is proportional to $\sigma(f)^{-2}$, and so it is analogous to the weighting factors in Eq. 3.16. The difference is that $\hat{Y}^{\text{rad}}$ builds this weighting into the filter function whereas we opt to carry out the weighting when combining pixels. The factor of $\bar{H}(f)$ in $\tilde{Q}_{IJ}^{\text{rad}}$ is the expected source power spectrum. When we choose a set of pixels Γ, we effectively define $H(f)$ such that $H(f) = \text{const}$ inside Γ and $H(f) = 0$ outside Γ.

Second, we note that apparently $\tilde{Q}_{IJ}^{\text{rad}} \propto \epsilon$ whereas our filter scales like $\tilde{Q}_{IJ} \propto 1/\epsilon$. It turns out that both filters scale like $1/\epsilon$ because the radiometer normalization factor $\lambda \propto \epsilon^{-2}$. The historical reason for this is that the radiometer analysis was developed by analogy with isotropic analyses [92], which includes an integral over all sky directions. The inclusion of $\gamma(t;f,\hat{\Omega})$ in the expression for $\tilde{Q}_{IJ}^{\text{rad}}$ serves to weight different directions as more or less important just like the factor of

$1/P_I(f)P_J(f)$ weights different frequencies. We opt to remove γ in favor of ϵ, which deemphasizes the analogy with the isotropic analysis in order to highlight the network sensitivity, which is characterized by ϵ.

G. Relation to other search frameworks

This is not the first time that ft-maps of data have been proposed to search for GWs. The literature on this subject is extensive and diverse. We concentrate on comparison with "excess power" methods, (see e.g., [8, 95, 96]). The key difference between our framework and others is that we cross-correlate data from two interferometers *before* they are rendered as ft-maps. Previous implementations such as [95] and [96] instead form ft-maps by auto-correlating data from each interferometer individually and then correlating regions of significance in these maps. For Gaussian noise, neither of these ways of combining data from different detectors is optimal. Instead, the optimal multi-detector method incorporates both autocorrelated and cross-correlated components [8]. Real interferometric GW data, however, is not Gaussian. Rather, there is an underlying Gaussian component with frequent non-Gaussian bursts called "glitches." For situations of this type, our approach has two advantages.

First, noise bursts in both detectors that coincide in time and frequency increase the false-alarm rate for statistics with auto-correlated components, but are suppressed in our cross-correlation analysis unless the waveforms of the burst themselves are correlated in phase like a true GW. Second, even when noise bursts are present, the pixel values in an ft-map of cross-correlated data are well approximated by a simple model. This is unlike ft-maps with auto-correlated components, for which there is a no simple description. Thus, while our statistic is suboptimal for Gaussian data, we expect it to perform well for real interferometer data. Moreover, even in the case of Gaussian noise, we do not sacrifice much sensitivity compared to the optimal excess-power statistic, or even to matched filtering, as demonstrated in Sec. VII. We compare the sensitivity of the cross-correlation statistic to other methods in Sec. VII.

IV. DISTRIBUTION OF SIGNAL AND BACKGROUND

In order to determine if a candidate event warrants further examination, it is necessary to determine the threshold above which an event is elevated to a GW candidate. This threshold is usually phrased in terms of a false-alarm rate (FAR). In Sec. III, we argued that ft-maps of cross-power provide a convenient starting point for searches for long transients because cross-correlation yields a reasonably well-behaved $\mathrm{SNR}(t; f, \hat{\Omega})$ statistic whose probability density function (PDF) we can model numerically, thus allowing straightforward calculation of a nominal detection threshold in the presence of Gaussian noise. We now assess this claim quantitatively.

We consider $52\,\mathrm{s} \times 0.25\,\mathrm{Hz}$ pixels (created through a coarse-graining procedure described below), which are the intermediate data product in stochastic analyses such as LIGO's recent isotropic result [97]. There are, of course, other choices of pixel resolution, and different sources call for different resolutions. Typically one must balance concerns about the signal duration, the signal bandwidth and the stationarity of the detector noise. The PDF of $\mathrm{SNR}(t; f, \hat{\Omega})$ for a single pixel crucially depends on details of the pixel size. E.g., the PDF of $\mathrm{SNR}(t; f, \hat{\Omega})$ for coarse-grained pixels (described in App. B 1) is more nearly Gaussian-distributed since coarse-grained pixels are created by averaging over more than one frequency. Our goal here, therefore, is not an exhaustive treatment. Rather, we aim to assess the agreement of data with our model using one pixel size, and in doing so, demonstrate how this assessment can be carried out in general.

Our noise model assumes Gaussian strain noise, uncorrelated between detectors I and J,

$$2\langle \tilde{n}_I(t; f)\tilde{n}_J(t; f)\rangle = \delta_{IJ}N_I(t; f), \qquad (4.1)$$

where $N_I(t; f)$ is the one-sided noise power spectrum and $\tilde{n}_I(t; f)$ is the discrete Fourier transform of the noise strain time series in detector I. Although we are dealing with Gaussian noise, i.e., $\tilde{n}(t; f)$ is normally distributed, the associated PDF for $\mathrm{SNR}(t; f, \hat{\Omega})$ is *not* expected to be normally distributed. It is more peaked than a normal distribution and it has broader tails (see Fig. 1).

Fig. 1 shows a comparison of $\mathrm{SNR}(t; f, \hat{\Omega})$ for data and Monte Carlo. The data corresponds to approximately a third of a day of at the beginning of LIGO's S5 science run using the Hanford H1 and Livingston L1 detectors. We introduce an unphysical time-shift between the two data streams to remove all astrophysical content. Additional data processing details are described in Apps. B 1 and B 2, as the precise shape of the PDF for $\mathrm{SNR}(t; f, \hat{\Omega})$ depends crucially on details of how time series data is processed.

The Monte Carlo histogram is scaled by a normalization factor (derived analytically in App. B 2), which takes into account data processing not included in our Monte Carlo simulation, e.g., coarse-graining. After applying this normalization factor, we find that the standard deviation of the data and Monte Carlo distributions agree to better than four significant digits. We conclude that data and Monte Carlo are in qualitative agreement. Thus we expect that the data are well-behaved enough that we can use a Gaussian noise model to assign a detection candidate threshold for SNR_Γ, at least for this choice of pixel size.
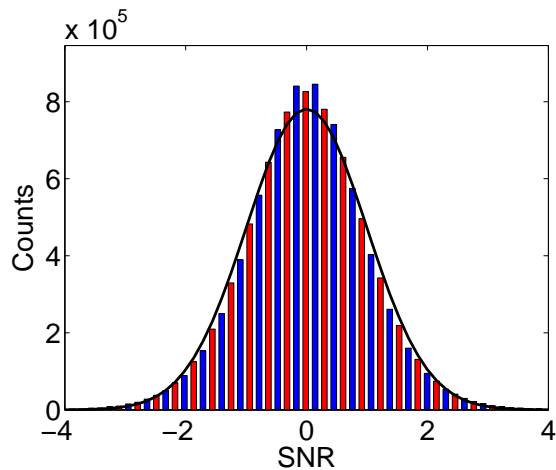
FIG. 1: Histogram of $\mathrm{SNR}(t; f, \hat{\Omega})$ using $52\,\mathrm{s} \times 0.25\,\mathrm{Hz}$ pixels comparing S5 data with an unphysical time-shift (blue) to Monte Carlo data (red) and a normal distribution with mean $= 0$ and $\sigma = 1$ (black). Error bars are too small to see.

FIG. 2: GW strain amplitude spectrum due to PNS convection in an axisymmetric PNS model at a typical galactic distance of $10\,\mathrm{kpc}$ [2]. This plot was generated using the data simulated in [2] available at [98].

V. PATTERN RECOGNITION

In this section we showcase the cross-power statistic developed in Sec. III using two different implementations (the box search and the Radon search) designed to address two different type of astrophysical scenarios (broadband signals and narrowband signals).

A. Broadband box search

We demonstrate how a box-shaped set of pixels can be used to search for a broadband GW transient source. For illustrative purposes, we consider a toy model based on protoneutron star (PNS) convection with a spectrum produced in an axisymmetric PNS model assuming a non-rotating, $15\,M_{\odot}$ progenitor [2] (see Fig. 2). We simulate a $d = 4.5\,\mathrm{kpc}$ source in the direction of ra $= 17\,\mathrm{hrs}$, decl $= 30°$ at 00:00 GMST on top of simulated detector noise comparable to the design sensitivity for initial LIGO. We calculate the cross-power statistic $\hat{\mathcal{Y}}_{\Gamma}$ utilizing a $200\,\mathrm{Hz} \times 16\,\mathrm{s}$ box constructed with the H1L1 detector network. We use $4\,\mathrm{s} \times 0.25\,\mathrm{Hz}$ pixels, and for each pixel we use 20 adjacent segments to calculate $\hat{\sigma}(t; f, \hat{\Omega})$, (10 on each side). We tile the ft-map and record the $\hat{\mathcal{Y}}_{\Gamma}$ within each box. We find that the signal can be recovered with $\mathrm{SNR}(t; f, \hat{\Omega}) = 8$. The results are summarized in Fig. 3.

B. Radon algorithm

Radon transforms are regularly used in imaging problems in order to identify line-like features in 2D maps [9]. This makes the Radon algorithm useful for looking for narrowband GW 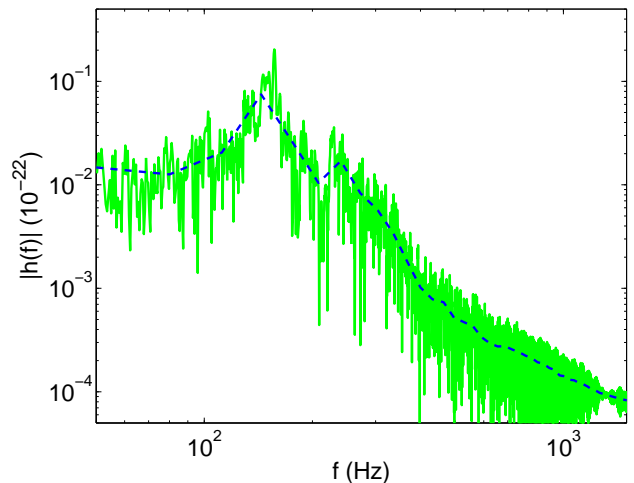tracks in ft maps, (see, e.g., [99]). By converting from the coordinates $(t; f)$ to impact parameter b and angle θ, a line-like cluster in ft-space is converted to a peak in Radon space. Thus, this algorithm provides a convenient way to search for GW signals that manifest themselves as line-like tracks in ft-space.

For continuous variables, the Radon transform of some function $g(t; f)$ is defined as [9]:

$$\mathcal{R}(b, \theta)[g(t; f)] \equiv \int df \int dt\, g(t; f)\delta(b - t\cos\theta - f\sin\theta). \tag{5.1}$$

For discrete variables, the Radon transform becomes

$$\mathcal{R}(b, \theta)[g(t; f)] = \sum_{t;f} w_{t;f}^{b\theta}\, g(t; f). \tag{5.2}$$

The weight factors $w_{t;f}^{b\theta}$ describe how close a line, parameterized by (b, θ), passes to the center of each ft-pixel. We use a (modified) Radon transform algorithm from [100], which is one of many possible implementations of the discrete Radon transform.

The estimator for the cross power in a pixel set described by (b, θ) can be written entirely in terms of Radon transforms:

$$\hat{\mathcal{Y}}(\hat{\Omega}, b, \theta) = \frac{\mathcal{R}[\hat{Y}(t; f, \hat{\Omega})\hat{\sigma}(t; f, \hat{\Omega})^{-2}]}{\mathcal{R}\left[\hat{\sigma}(t; f, \hat{\Omega})^{-2}\right]} \tag{5.3}$$

The associated variance is

$$\hat{\sigma}_{\mathcal{Y}}(\hat{\Omega}, b, \theta)^2 = \frac{\sum_{t;f}(w_{t;f}^{\theta b})^2\, \hat{\sigma}(t; f, \hat{\Omega})^{-2}}{\left(\mathcal{R}[\hat{\sigma}(t; f, \hat{\Omega})^{-2}]\right)^2} \tag{5.4}$$

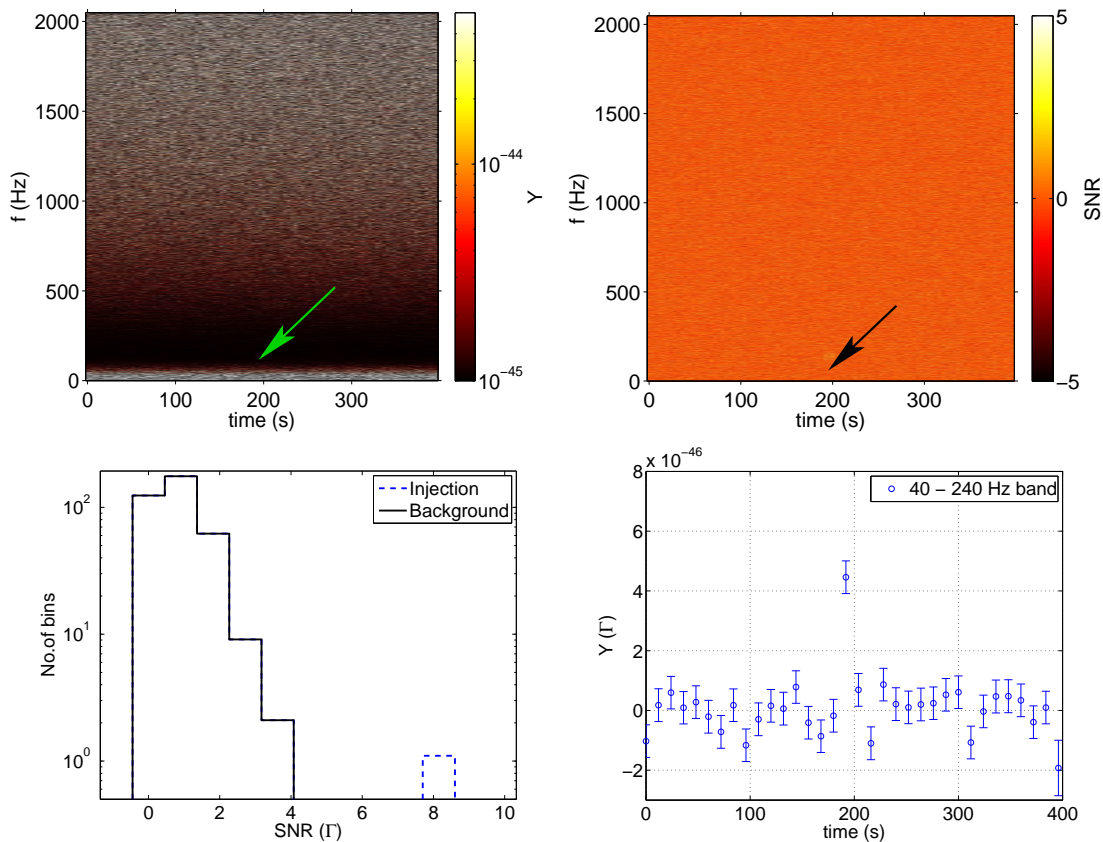We now consider a toy model of torus excitations from long GRBs [5], which are expected to produce line-like

FIG. 3: Injection recovery with the box-search algorithm. Top-left: an ft-map of $\hat{Y}(t; f, \hat{\Omega})$. The injected signal (not visible by eye) is indicated with a green arrow. Top-right: an ft-map of SNR$(t; f, \hat{\Omega})$. The injected signal (not visible by eye) is indicated with a black arrow. Bottom-left: a histogram of SNR$_\Gamma$ for a $200\,\text{Hz} \times 12\,\text{s}$ box. The blue dashed corresponds to the injection. Though the signal is weak in each pixel, the signal obtained by combining every pixel in Γ is large. Bottom-right: $\hat{\mathcal{Y}}_\Gamma$ as a function of time.

clusters in ft-space with durations of $2 \sim 200\,\text{s}$. Since we are dealing with an elliptically polarized source, $\hat{\mathcal{Y}}(t; f, \hat{\Omega})$ also depends on inclination angle ι and polarization angle ψ, (see App. A 5). For the sake of simplicity, however, we use an unpolarized filter, which has been shown to do a reasonably good job recovering elliptically polarized sources [101]. We simulate an elliptically polarized waveform (see Tab. III) on top of simulated detector noise comparable to design sensitivity for initial LIGO. Once again, we use $4\,\text{s} \times 0.25\,\text{Hz}$ pixels, and for each pixel we use 18 adjacent segments to calculate $\sigma(t; f, \hat{\Omega})$ (9 from each side). The resulting maps of $\hat{Y}(t; f, \hat{\Omega})$ and SNR$_\Gamma$ are shown in Fig. 4. For this toy-model, the signal was recovered with SNR$_\Gamma = 8.1$ at $d = 1.7\,\text{Mpc}$.

parameter	value
strain amplitude	$h_{\rm rms} = 1.2 \times 10^{-21}$
search duration	$100\,\text{s}$
(ra, dec)	$(17\,\text{hr}, 30°)$
(ι, ψ)	$(0°, 0°)$
injection start time	00:00 GMST
injection duration	$40\,\text{s}$
distance to GRB	$1.7\,\text{Mpc}$
$(f_0, df/dt)$	$(998\,\text{Hz}, 0.03\,\text{Hz/s})$

TABLE III: Search and injection parameters used for Fig. 4.

C. Other algorithms

In this section we have, for illustrative purposes, presented two of the many pattern recognition algorithms that may be applied to the problem of looking for features in ft-maps of cross-power. There is a diverse and extensive literature devoted to the study of cluster iden-

tification, (see, e.g., [102, 103]). In the next section we apply the Radon algorithm to non-GW channels in order to look for environmental noise artifacts that are qualitatively similar to our long-GRB toy model. For comparison, we also make use of locust and Hough algorithms [10], which have been proposed as a method of identifying long-GRB events in GW data.
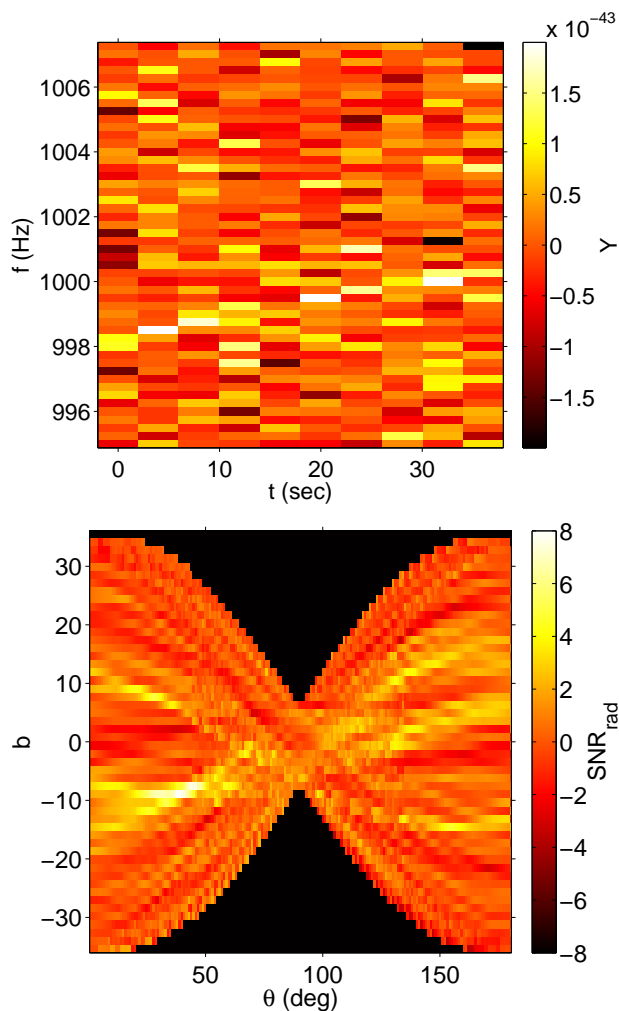
FIG. 4: Top: an ft-map of $\mathrm{SNR}(t; f, \hat{\Omega})$. The long-GRB track is a barely discernible diagonal line. Bottom: a map of $\mathrm{SNR}_{\Gamma}(\hat{\Omega}, b, \theta)$. The signal shows up as a hot spot at $(\theta, b) = (39°, -8)$ with $\mathrm{SNR}_{\Gamma} = 8.1$. The sinusoidal patterns are (expected) covariances between Radon-map pixels.

VI. APPLICATION TO ENVIRONMENTAL NOISE IDENTIFICATION

A. Environmental noise in GW interferometers

While our discussion until now has been focused on the detection of GW transients, the same formalism can be applied to look for structure in ft-maps of cross-power between any two data channels. In particular, it is illuminating to study the cross-power between an interferometer's GW-strain channel, (which we denote s_{GW}) and a physical environmental monitoring (PEM) channel such as a seismometer or a magnetometer channel located near the interferometer. Since PEM channels are not sensitive to GWs, statistically significant features in an ft-map of PEM-s_{GW} cross-power are likely due to environmentally-induced noise artifacts.

Transient artifacts are called "glitches" whereas persistent narrowband features are often called "lines" or "wandering lines" when the frequency slowly changes over time. Glitches and wandering lines can be problematic for searches for bursts / compact binary coalescences and for pulsars respectively, see, e.g., [104–107]. (They also produce non-Gaussian noise for our cross-power statistic.) It is thus desirable to identify and when possible mitigate these noise features.

In this section we show how the formalism we have developed to search for long GW transients can also be used to identify glitches and wandering lines in PEM-s_{GW} cross-power maps. There are two points we hope to make with this digression. First, we shall see that PEM-s_{GW} ft-maps are useful for identifying, characterizing (and in some cases eliminating) environmentally-induced noise. Second, we show that some environmentally-induced noise in PEM-s_{GW} ft-maps is qualitatively similar to the GW-transient signature in cross-power maps between two s_{GW} channels. Thus, these noise artifacts provide a convenient dataset to demonstrate our search algorithms with environmental noise events.

B. Environmental channels at LIGO and Virgo

In order to facilitate the detection of transient GWs, it is necessary to monitor and characterize glitches and lines. Efforts to identify and document noise artifacts are a major task of the Detector Characterization and Glitch groups within the LIGO Scientific Collaboration [105, 108, 109] and the Virgo Collaboration. To assist in this effort, each LIGO/Virgo observatory is supplemented with hundreds of sensors that monitor the local environment. (For an overview of the LIGO and Virgo interferometers, see [110] and [111] respectively.)

Accelerometers measure vibrations such as the motion of the beam tubes and of the optical tables that house photodiodes; microphones monitor acoustic noise at critical locations; magnetometers monitor magnetic fields that could couple to the detector; radio receivers monitor radio frequency power around the laser modulation frequencies; and voltage line monitors record fluctuations in the AC power. The PEMs are placed at strategic locations around the observatory, especially near the corner and ends of the interferometer where important laser, optical and suspension systems reside in addition to the test masses themselves.

For illustrative purposes, we consider a special class of noise artifacts induced by passing airplanes. These "airplane events," have attractive properties for our purposes. First, airplane events are relatively well understood. The existing LIGO airplane veto system (called `planemon`) has been shown to flag airplanes observed in microphone channels, and these flags have been shown to agree with airplane flight data [112, 113], though the existing `planemon` algorithm does not determine if the passing airplane affects $s_{\mathrm{GW}}(t)$. Since we already under-

stand a lot about airplane events, it is straightforward to assess if our algorithms are consistent with what we already know.

Second, we shall see that airplanes produce a slightly curved narrowband ft-map track lasting tens of seconds. These tracks are qualitatively similar to GW transient scenarios such as van Putten's long-GRBs model, and thus they provide an opportunity for us to demonstrate our search algorithms on a distribution of signals that resemble unmodeled GW transients; we have a qualitative picture of the signal, but it is impractical to model the space of all possible signals with a matched filter template bank.

C. Airplane noise identification

In order to identify airplane events, ft-maps correlating the GW channel with acoustic channels are computed in $400\,\mathrm{s}$ blocks. Then we take the absolute value of the $\mathrm{SNR}(t; f)$ map. This step is not necessary for GW studies because GW transients can only produce positive-definite $\mathrm{SNR}(t; f)$ once the phase delay between two sites has been taken into account. Transient noise artifacts, on the other hand, can produce complex (not positive-definite) $\mathrm{SNR}(t; f)$ since the cross-power phase depends on the coupling of the environmental noise into s_{GW}. By taking the absolute value of $\mathrm{SNR}(t; f)$, the PDF of $\mathrm{SNR}(t; f)$ changes from the description in Subsec IV, so we estimate it semi-empirically with the assistance of simulation (see below).

We create a Radon transform of each ft-map of $|\mathrm{SNR}(t; f)|$, an example of which can be seen in Fig. 5. We are presently interested in qualitative aspects of the airplane track, and so we make the simplifying assumption that $\sigma(t; f) = \mathrm{const}$. Although airplane tracks are slightly curved, the approximation of the tracks as lines is suitable for a simple identification. We record the brightest spot, i.e., the maximum SNR_Γ, on each Radon map. The maps are then ordered according to their maximum SNR_Γ. The ft-maps with the highest SNR_Γ are checked by eye for airplane tracks. In Fig. 6, we show a histogram of SNR_Γ. The red entries are for ft-maps with unambiguous airplane tracks. The green entries are possible airplane tracks, yet a visual inspection of these events was inconclusive. All other entries (due to non-airplane background) are blue. Using Fig. 6, we set a threshold of $\mathrm{SNR}_\Gamma > 16$ for an event to be tagged as an unambiguous airplane.

In Tab. IV, we compare the Radon triggers to the `planemon` triggers. During a span of $3\,\mathrm{days}$, the Radon algorithm found 37 unambiguous airplanes, 26 of which were also seen by `planemon`. The 11 events that were not confirmed by `planemon` were confirmed by eye as airplane-like. It is possible that they failed to create a `planemon` trigger because the airplane flight path did not trigger enough microphone channels. During this same period, `planemon` found 54 unique events, 28 of which

	`planemon`-flagged	not `planemon`-flagged
Radon-flagged	26	11
not Radon-flagged	28	583

TABLE IV: Number of airplane event triggers identified by the Radon and `planemon` algorithms out of 648 total ft-maps.

were not found by the Radon algorithm as the tracks were not coherent in s_{GW}.

We conclude that the Radon flags appear to be roughly consistent with the planemon flags. However, the Radon flag differs in two useful ways. First, Radon flags require coherence between microphone channels and s_{GW}, and so our algorithm only flags airplanes that contaminate $s_{\mathrm{GW}}(t)$. Second, the Radon flag does not need confirmation between more than one microphone channel so long as the signal from one is coherent with s_{GW}.

For each map deemed to contain an airplane, the inverse Radon transform is performed on the brightest spot, and the airplane track is reconstructed as an ft-map line as in Fig. 5. The start and stop times of the airplane noise in this frequency band are estimated to be the times at which the reconstructed track intersects the edge of the ft-map. This routine was run on the LIGO and Virgo data during recent science runs, correlating the GW channel with about 10 microphones at each detector. The algorithm identified $\sim 10 - 15$ airplane events at an observatory each day.

PEM-s_{GW} ft-maps created with real data contain additional environmental noise artifacts besides airplanes. In order to estimate the FAR for airplane tracks in "idealized" noise (where no other environmental artifacts are present), we perform Monte Carlo pseudo-experiments by scrambling ft-map pixels so as to wash out clustering while preserving the empirically observed distributions of $Y(t; f)$ and $\sigma(t; f)$. Using the weighted Radon algorithm, (where $\sigma(t; f) \neq \mathrm{const}$), we estimate that, on average, $0.4\,\mathrm{events/month}$ are falsely identified as airplanes in idealized noise.

D. Other algorithms

Having demonstrated how the Radon algorithm can recover airplane events in PEM-s_{GW} ft-maps, we now demonstrate two additional pattern recognition algorithms. Our point is to convey the wide array of tools available to solve the problem of pattern recognition in ft-maps of cross-power. There are typically both advantages and disadvantages associated with each algorithm, which means that each one lends itself to different applications, though extensive discussion of the merits of different techniques is beyond our current scope.

In particular, we consider the locust and Hough algorithms from [10], both of which are well-suited for narrow-band sources. The Hough algorithm is similar to
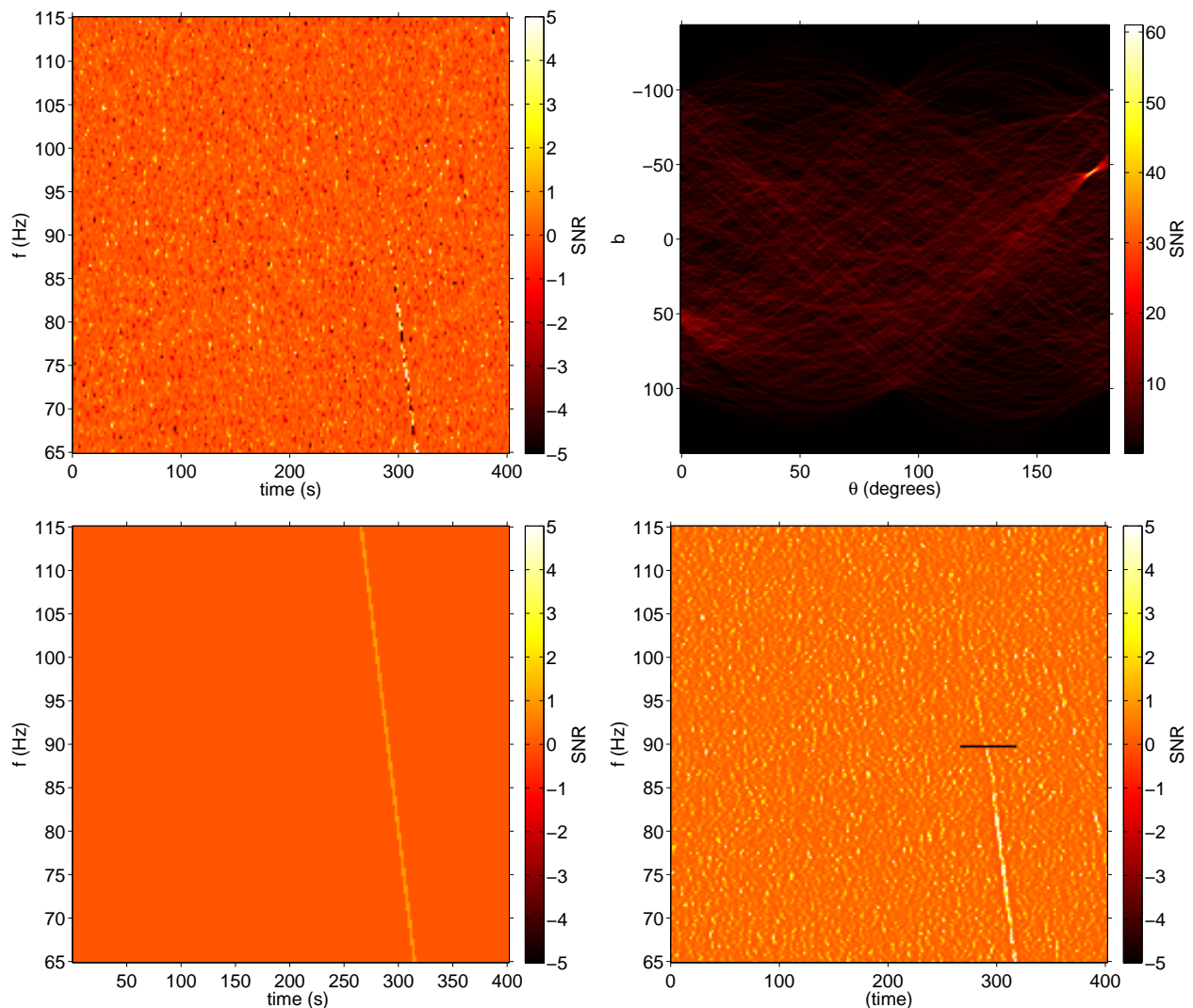
FIG. 5: Top-left: a 400 s-duration map of SNR$(t; f)$ (see Eq. 3.11) created with 4 s × 0.25 Hz pixels and using s_{GW} cross-correlated with a microphone. The slightly curved track on the right side of the plot is caused by the Doppler-shifted acoustic signal from a passing airplane. Top-right: the associated Radon map. Note the bright spot on the mid-right corresponding to the airplane track. Bottom-left: ft-map of the reconstructed track using the maximum SNR$(t; f)$ pixel in Radon space. Bottom-right: ft-map of the magnitude of SNR$(t; f)$ including a black line corresponding to the veto window. These data are from the beginning of LIGO's S5 science run.

the Radon algorithm, except that it can be extended to fit tracks described by arbitrarily high-order polynomials. By introducing additional fit parameters, the tracks tend to be reconstructed more accurately. However, by adding more parameters, the significance of a line-like event with little or no curvature can be less than the value obtained by the Radon algorithm.

The locust algorithm is a local wandering algorithm, which integrates the ft-map along a chain of local maxima. This algorithm has the advantage that it can reconstruct arbitrary-shaped tracks without large numbers of free parameters. Since it relies on local maxima, however, the Radon and Hough algorithms are more robust if the GW power is spread diffusely over many pixels. Both the locust and Hough algorithms produce a statis-

tic, which is the integral of cross-power along a track. We estimate significance by performing Monte Carlo pseudo-experiments in which we randomly scramble the ft-map pixels.

Applying the locust and Hough algorithms to an unambiguous airplane event, we obtain the reconstruction plots shown in Fig. 7. We determine that both the locust and Hough algorithms detect the event with a FAR no more than 0.04% per 400 s map in idealized noise.
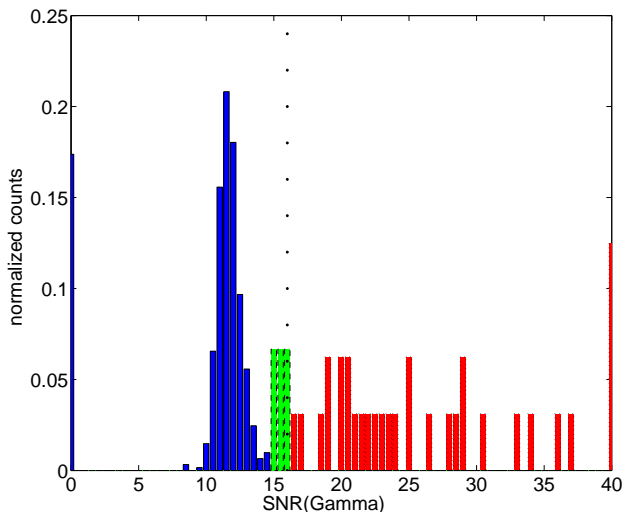
FIG. 6: A histogram corresponding to SNR_Γ for above-threshold airplane events (in red, dotted outline), near-threshold events (green dashed outline) and below-threshold background events (in blue, solid outline) in three days of LIGO Hanford Observatory data. The dotted black line corresponds to the chosen threshold of 16. The red, green and blue distributions are separately normalized to unity for the purpose of plotting. Entries with $\mathrm{SNR}_\Gamma = 0$ record the maps which were excluded from this study for failing a "glitch cut." Entries with $\mathrm{SNR}_\Gamma = 40$ record the maps with $\mathrm{SNR}_\Gamma \geq 40$.

VII. COMPARISON WITH OTHER TECHNIQUES

In this section, we compare the proposed excess cross-power statistic to matched filtering and to the general excess-power statistic from [8]. It is impractical to actually carry out a matched filtering analysis since we do not have a template bank from which we can construct an arbitrary transient signal. Nevertheless, it is possible to perform analytical estimates.

We consider a signal space characterized by an ft-volume V spanned by N_{eff} independent matched filtering templates. The ft-volume is simply the number of pixels in our pixel set Γ. We endeavor to address the following question: given a false-alarm probability (FAP) and a false-dismissal probability (FDP), what is the minimum signal amplitude detectable by either method? Following [8], we respectively define thresholds $A_{\min}^{\mathrm{CP}}$, $A_{\min}^{\mathrm{EP}}$ and $A_{\min}^{\mathrm{MF}}$ as the minimum detectable amplitudes by our cross-power (CP) search, by the optimal total excess-power (EP) search [8], (which includes both cross-power and auto-power terms) and by a matched filter search (MF). The ratio of the corresponding amplitudes is (by definition) the *efficiency* of the excess power statistic compared to matched filtering:

$$\eta_{\mathrm{EPMF}}(\mathrm{FAP}, \mathrm{FDP}, N_{\mathrm{eff}}, V) = A_{\min}^{\mathrm{MF}}/A_{\min}^{\mathrm{EP}} \quad (7.1)$$

$$\eta_{\mathrm{CPEP}}(\mathrm{FAP}, \mathrm{FDP}, V) = A_{\min}^{\mathrm{EP}}/A_{\min}^{\mathrm{CP}}. \quad (7.2)$$

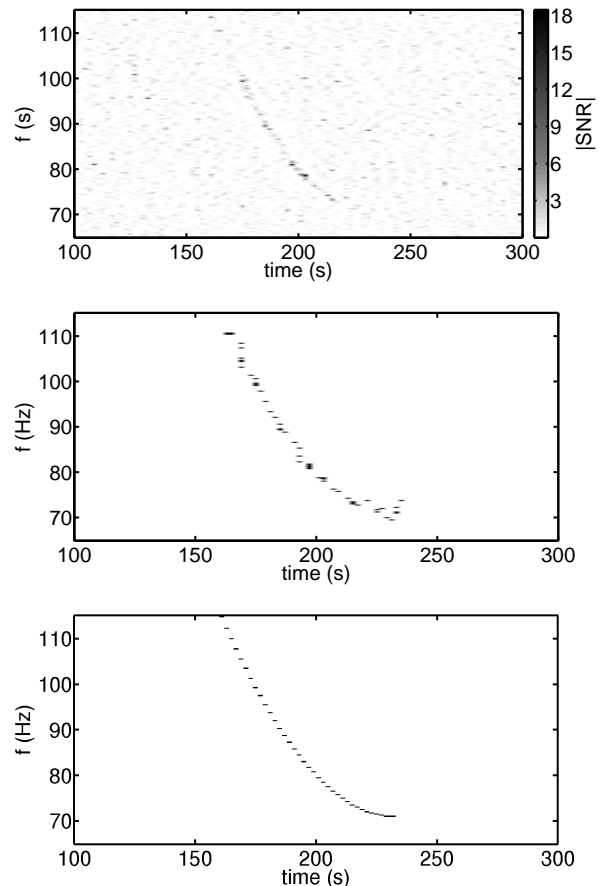To calculate these thresholds, signal and noise distri-

FIG. 7: Upper panel: an airplane track in a map of $\mathrm{SNR}(t; f, \hat{\Omega})$. Middle panel: the reconstructed track as determined by the locust algorithm. Bottom panel: the reconstructed track as determined by the Hough algorithm using a second-order polynomial.

butions for CP and EP are generated using Monte Carlo simulations. Throughout this section we assume stationary Gaussian white noise; simulated signals are characterized only by their amplitude, polarization and ft-volume. Other characteristics such as frequency content, evolution with time, etc. are not relevant for this white-noise calculation. Following [8], we approximate the MF threshold as

$$A_{\min}^{\mathrm{MF}} \approx A_{\min}^{\mathrm{EP}}(\mathrm{FAP}/N_{\mathrm{eff}}, \mathrm{FDP}, 1/2). \quad (7.3)$$

In Fig. 8, we plot η_{EPMF} as a function of V and N_{eff}. We see that $\eta_{\mathrm{EPMF}} \gtrsim 50\%$ over the range of parameter space considered.

To compare the CP method to the EP method, we calculate η_{CPEP} for the Hanford-Livingston and Hanford-Virgo networks averaging over an isotropically distributed population of unpolarized GW sources. In Fig. 9 we plot η_{CPEP} as a function of ft-volume V. The CP technique is highly efficient ($\eta_{\mathrm{CPEP}} \sim 77\%$) even at small values of V, but it becomes increasingly more efficient at higher V.
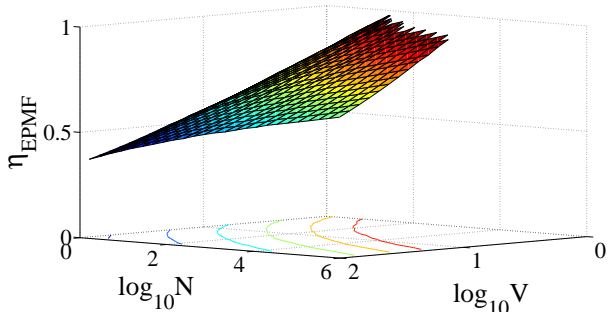
FIG. 8: The relative detection efficiency η_{EPMF} of excess-power compared to matched filtering as a function of ft-volume V and the number of effective templates N_{eff} (for FAP $= 10^{-2}$ and FDP $= 10^{-1}$). We assume Gaussian white noise. For a given value of V, there is maximum number for N_{eff} above which the templates will not be independent. Therefore, at high values of N_{eff} and small values of V, there is an unphysical region where $\eta > 1$. A realistic excess-power (and hence cross-power) search might have $V \sim \mathcal{O}(100)$, which corresponds to an efficiency of $\sim 50\%$ compared to matched filtering.
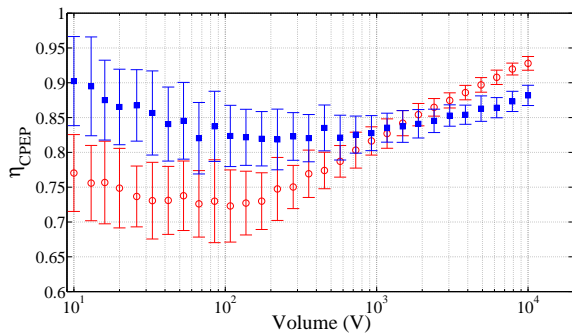
FIG. 9: The relative detection efficiency comparing the general excess power and cross-power methods for FAP $= 10^{-2}$ and FDP $= 0.1$ in simulated Gaussian white noise. The red circles are for the H1V1 detector pair and the blue squares are for the more nearly aligned H1L1 detector pair. The error bars represent one standard deviation using 100 trials. The cross-power method has a relative detection efficiency of at least 70% with respect to the excess power statistic and the efficiency increases with V.

VIII. CONCLUSIONS

We reviewed a variety of compelling scenarios for long-GW transients including protoneutron star convection, accretion disk fragmentation/excitations, rotational instabilities in neutron stars, r-modes, pulsar glitches and soft gamma repeater flares. Many of the models we considered predict strain amplitudes detectable in the advanced-detector era.

Next, inspired by stochastic analyses, we introduced a novel framework, which can be used to look for GW transients on timescales of seconds to weeks. This framework, which is a generalization of the GW radiometer [7], utilizes ft-maps of GW-strain cross-power using two or more spatially separated interferometers in order to look for statistically significant clustering. A comparison of simulated detector noise with time-shifted data revealed that ft-cross-power-maps made with real interferometer data are well-behaved (for at least one pixel size) suggesting that the threshold for candidate events can be determined analytically.

We illustrated how different pattern-recognition techniques can be used to identify GW signatures in ft-maps. We demonstrated some of these techniques using ft-maps generated using a GW-strain channel cross-correlated with a LIGO microphone channel and we presented a novel technique for the identification of environmental noise transients in GW interferometers. We assessed the effectiveness of our proposed statistic compared to matched filtering and other excess-power strategies. We found that our strategy is highly effective when looking for long GW transients whose precise waveforms are not known.

Many of the sources we considered here are plausible targets for the advanced detector era. If long GW transients are, in fact, detected, we shall gain invaluable information about objects and processes, for which we currently possess only preliminary models, e.g., long gamma-ray bursts. If, on the other hand, no long GW transients are detected, we expect that some models predicting relatively large strain amplitudes (e.g., [5]) may be ruled out or constrained. Third-generation detectors such as the proposed Einstein Telescope [114] can apply our long GW transient algorithm to probe still fainter sources.

Appendix A: Derivations

We describe how a GW source can be characterized by its power spectrum $H(t; f)$, we construct an estimator $\hat{Y}(t; f)$ for $H(t; f)$ and calculate the associated variance. We construct an estimator for the variance. Finally, we construct the filter function for an elliptically polarized source.

1. Introduction and notation

Working in the transverse-traceless gauge we write down the general form of a GW field, which can depend on direction $\hat{\Omega}$, polarization state A and frequency f (see, e.g., [7, 92, 93]):

$$h_{ab}(t, \vec{x}) = \sum_A \int_{-\infty}^{\infty} df \int_{S^2} d\hat{\Omega}\, e_{ab}^A(\hat{\Omega}) \tilde{h}_A(f, \hat{\Omega}) \\ e^{2\pi i f(t + \hat{\Omega} \cdot \vec{x}/c)}. \tag{A1}$$

Here $\vec{x}$ and $\hat{\Omega}$ are defined in the reference frame with the origin fixed at the center of Earth, but not rotating with the Earth. The indices ab run over a Cartesian coordinate system. We define unit vectors

$$\hat{\theta} = \cos\theta\cos\phi\,\hat{x} + \cos\theta\sin\phi\,\hat{y} - \sin\theta\,\hat{z} \tag{A2}$$

$$\hat{\phi} = -\sin\phi\,\hat{x} + \cos\phi\,\hat{y} \tag{A3}$$

$$\hat{\Omega} = \sin\theta\cos\phi\,\hat{x} + \sin\theta\sin\phi\,\hat{y} + \cos\theta\,\hat{z} \tag{A4}$$

such that $\{\hat{\theta}, \hat{\phi}, \hat{\Omega}\}$ form a right-handed coordinate system and the rotational axis of the Earth points along $\hat{z}$. The two GW polarization tensors can be written as (see, e.g., [7, 92, 93]):

$$e_{ab}^+(\hat{\Omega}) = \hat{\theta} \otimes \hat{\theta} - \hat{\phi} \otimes \hat{\phi} \tag{A5}$$

$$e_{ab}^\times(\hat{\Omega}) = \hat{\theta} \otimes \hat{\phi} + \hat{\phi} \otimes \hat{\theta}. \tag{A6}$$

Since we are looking for GW transients, we restrict our attention to *point sources* for which $h_A(f, \hat{\Omega}) = h_A(f)\delta(\hat{\Omega} - \hat{\Omega}_0)$. We perform the integral over $\hat{\Omega}$ and obtain:

$$h_{ab}(t, \vec{x}) = \sum_A \int_{-\infty}^{\infty} df\, e_{ab}^A(\hat{\Omega}_0) \tilde{h}_A(f)\, e^{2\pi i f(t + \hat{\Omega}_0 \cdot \vec{x}/c)}. \tag{A7}$$

For simplicity, we henceforth replace $\hat{\Omega}_0$ with $\hat{\Omega}$. It follows that the GW strain in detector I is given by

$$h_I(t) = \sum_A \int_{-\infty}^{\infty} df\, \tilde{h}_A(f, \hat{\Omega})\, e^{2\pi i f(t + \hat{\Omega} \cdot \vec{x}_I/c)}\, e_{ab}^A(\hat{\Omega})\, d_I^{ab}(t) \tag{A8}$$

where $d_I^{ab}(t)$ is the detector response tensor at time t:

$$d_I(t) \equiv \frac{1}{2} \left(\hat{X}_I(t) \otimes \hat{X}_I(t) - \hat{Y}_I(t) \otimes \hat{Y}_I(t) \right). \tag{A9}$$

Here, the two detector arms lie along the $\hat{X}(t)$ and $\hat{Y}(t)$ axes, which are time-dependent due to the rotation of the Earth.

We now consider a finite stretch of $h_I(t)$ and take the discrete Fourier transform of Eq.A8 to obtain:

$$\tilde{h}_I(t; f) = \sum_A \tilde{h}_A(t; f, \hat{\Omega})\, e^{2\pi i f \hat{\Omega} \cdot \vec{x}_I/c}\, F_I^A(t; \hat{\Omega}). \tag{A10}$$

where we define the "antenna factors" (see e.g., [92]) to be

$$F_I^A(t; \hat{\Omega}) \equiv e_{ab}^A(\hat{\Omega})\, d_I^{ab}(t). \tag{A11}$$

We define the GW strain power spectrum to be

$$\langle \tilde{h}_A^*(t; f) \tilde{h}_{A'}(t; f) \rangle = \frac{1}{2} H_{AA'}(t; f), \tag{A12}$$

where the factor $1/2$ comes from the fact that $H_{AA'}(t; f)$ is the one-sided power spectrum. Here we use the discrete Fourier transform defined in Eq. 3.2; (see also Tab. II).

It is convenient to characterize the source with a single spectrum that includes contributions from both $+$ and $\times$ polarizations. We therefore define

$$H(t; f) \equiv \text{Tr}\left[H_{AA'}(t; f) \right], \tag{A13}$$

so as to be invariant under change of polarization bases. This definition is a generalization of the one-sided power spectrum for unpolarized sources found in [7, 92, 93]. Our goal now is to derive an estimator for $H(t; f)$ in a data segment over which it is presumed to be constant.

2. Derivation of $\hat{Y}$

Let $s_I(t) = h_I(t) + n_I(t)$ be the strain time series from detector I, where $h_I(t)$ is the GW strain and $n_I(t)$ is the detector noise. Following [7, 92, 93], we combine the strain time series from two spatially separated detectors, $s_I(t), s_J(t)$, to construct an estimator for GW-power $H(t; f)$ for a point source at a sky position $\hat{\Omega}$,

$$\hat{Y}(t; f, \hat{\Omega}) \equiv 2\,\text{Re}\left[\tilde{Q}_{IJ}(t; f, \hat{\Omega}) \tilde{s}_I^*(t; f) \tilde{s}_J(t; f) \right] \tag{A14}$$

where $\tilde{Q}_{IJ}(t; f, \hat{\Omega})$ is some filter function to be determined below. We take the real part to ensure physicality of the estimator. The expectation value of $\hat{Y}(t; f, \hat{\Omega})$ is given by

$$\langle \hat{Y}(t; f, \tilde{\Omega}) \rangle = 2\,\text{Re}\left[\tilde{Q}_{IJ}(t; f, \hat{\Omega}) \langle \tilde{h}_I^*(t; f)\, \tilde{h}_J(t; f) \rangle \right], \tag{A15}$$

since, by assumption, there is no correlation between signal and noise and also no correlation between noise in two spatially separated detectors.

Combining Eqs. A10, A15 and A12 we get

$$\langle \hat{Y}(t; f, \hat{\Omega}) \rangle = 2\,\text{Re}\Big[\tilde{Q}_{IJ}(t; f, \hat{\Omega}) \sum_{AA'} \frac{1}{2} H_{AA'}(t; f) \\ e^{-2\pi i f(\hat{\Omega} \cdot (\vec{x}_I - \vec{x}_J)/c)}\, F_I^A(t; \hat{\Omega})\, F_J^{A'}(t; \hat{\Omega}) \Big] \tag{A16}$$

In order to simplify the form of $H_{AA'}(t; f)$ we now consider *unpolarized* sources, for which

$$H_{AA'}(t; f) = \frac{1}{2} H(t; f)\delta_{AA'}. \tag{A17}$$

(We consider the case of polarized sources in Subsec. A 5.) For unpolarized sources,

$$\langle \hat{Y}(t;f,\hat{\Omega})\rangle = \frac{1}{2}\mathrm{Re}\Big[\tilde{Q}_{IJ}(t;f,\hat{\Omega})\,H(t;f)\,e^{-2\pi i f\hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c} \\ \sum_A F_I^A(t;\hat{\Omega})\,F_J^A(t;\hat{\Omega})\Big], \tag{A18}$$

where we have defined $\Delta\vec{x}_{IJ} \equiv \vec{x}_I - \vec{x}_J$.

We desire that $\langle \hat{Y}\rangle = H(t;f)$, which implies:

$$\tilde{Q}_{IJ}(t;f,\hat{\Omega}) = \frac{2\,e^{2\pi i f\hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c}}{\sum_A F_I^A(t;\hat{\Omega})\,F_J^A(t;\hat{\Omega})}. \tag{A19}$$

By setting $Q_{IJ}(t;f,\hat{\Omega})$ thusly, we account for the phase difference between detectors I and J ensuring that the bracketed quantity in Eq. A18 is real. We also account for the detector pair efficiency.

Finally, we define (unpolarized) *pair efficiency* as

$$\epsilon_{IJ}(t;\hat{\Omega}) \equiv \frac{1}{2}\sum_A F_I^A(t;\hat{\Omega})F_J^A(t;\hat{\Omega}), \tag{A20}$$

which enables us to rewrite the filter function as

$$\tilde{Q}_{IJ}(t;f,\hat{\Omega}) = \frac{1}{\epsilon_{IJ}(t;\hat{\Omega})}e^{2\pi i f\hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c}. \tag{A21}$$

Since $\hat{Y}(t;f,\hat{\Omega}) \propto \tilde{Q}(t;f,\hat{\Omega})$ and $\tilde{Q} \propto 1/\epsilon_{IJ}(t;f,\hat{\Omega})$, it follows that $\hat{Y}(t;f,\hat{\Omega}) \propto 1/\epsilon_{IJ}(t;\hat{\Omega})$. This can be understood as follows. If we observe a modest value of strain power from a direction associated with low efficiency, we may infer (if the signal is statistically significant) that the true source power is much higher because the network only "sees" some fraction of the true GW power.

3. Variance of the estimator

We derive an expression for the variance of $\hat{Y}(t;f,\hat{\Omega})$, $\sigma_Y(t;f,\hat{\Omega})^2 \equiv \langle\hat{Y}(t;f,\hat{\Omega})^2\rangle - \langle\hat{Y}(t;f,\hat{\Omega})\rangle^2$. In searches for persistent stochastic GWs, the second term is usually omitted and the first term is simplified by assuming that signal in each pixel is small compared to the noise. Such small signals are extracted by averaging over a very large number of segments (see, e.g., [92]). Since we are dealing with transients, however, the signal may be comparable to the noise and so we can not neglect any terms in our calculation of σ_Y^2.

To begin we define a new (complex-valued) estimator that will be handy in our derivation of σ_Y^2:

$$\hat{W}(t;f,\hat{\Omega}) \equiv 2\,\tilde{Q}_{IJ}(t;f,\hat{\Omega})\tilde{s}_I^\star(t;f)\tilde{s}_J(t;f). \tag{A22}$$

Our GW power estimator $\hat{Y}(t;f,\hat{\Omega})$ is simply the real part of $\hat{W}(t;f,\hat{\Omega})$:

$$\hat{Y}(t;f,\hat{\Omega}) = \frac{1}{2}\left(\hat{W}(t;f,\hat{\Omega}) + \hat{W}(t;f,\hat{\Omega})^\star\right). \tag{A23}$$

For notational compactness, we shall omit the arguments of $\hat{W}(t;f,\hat{\Omega})$ in the remainder of this derivation. It follows that the variance of $\hat{Y}(t;f,\hat{\Omega})$ can be written as

$$\sigma_Y^2 = \frac{1}{4}\Big[\left(\langle\hat{W}^2\rangle - \langle\hat{W}\rangle^2\right) + \left(\langle\hat{W}^{\star 2}\rangle - \langle\hat{W}^\star\rangle^2\right) + \\ 2\,\sigma_W^2\Big], \tag{A24}$$

where

$$\sigma_W^2 \equiv \langle|\hat{W}|^2\rangle - |\langle\hat{W}\rangle|^2. \tag{A25}$$

Now we evaluate the three terms in Eq. A24 beginning with σ_W^2. We obtain

$$\sigma_W^2(t;f,\hat{\Omega}) = 4\Big[\langle\tilde{s}_I^*(t;f)\,\tilde{s}_J(t;f)\,\tilde{s}_I(t;f)\,\tilde{s}_J^*(t;f)\rangle \\ -\langle\tilde{s}_I^*(t;f)\,\tilde{s}_J(t;f)\rangle\langle\tilde{s}_I(t;f)\,\tilde{s}_J^*(t;f)\rangle\Big] \\ \left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2. \tag{A26}$$

For mean-zero Gaussian random variables, we can expand the four-point correlation into a sum of products of two-point correlations. We substitute $s = h + n$ and set signal-noise cross terms to zero along with noise-noise cross terms from different detectors. The variance becomes

$$\sigma_W^2(t;f,\hat{\Omega}) = 4\Big[\langle\tilde{h}_I^*(t;f)\tilde{h}_I(t;f)\rangle\,\langle\tilde{h}_J(t;f)\tilde{h}_J^*(t;f)\rangle + \\ \langle\tilde{h}_I^*(t;f)\tilde{h}_I(t;f)\rangle\,\langle\tilde{n}_J(t;f)\tilde{n}_J^*(t;f)\rangle + \\ \langle\tilde{h}_J(t;f)\tilde{h}_J^*(t;f)\rangle\,\langle\tilde{n}_I^*(t;f)\tilde{n}_I(t;f)\rangle + \\ \langle\tilde{n}_I^*(t;f)\tilde{n}_I(t;f)\rangle\,\langle\tilde{n}_J(t;f)\tilde{n}_J^*(t;f)\rangle\Big] \\ \left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2, \tag{A27}$$

Evaluating the four terms in Eq. A27, we obtain

$$\sigma_W^2(t;f,\hat{\Omega}) = \Big[\epsilon_{II}(t;\hat{\Omega})\epsilon_{JJ}(t;\hat{\Omega})H(t;f)^2 + \\ H(t;f)\left(\epsilon_{II}(t;\hat{\Omega})N_J(t;f) + \epsilon_{JJ}(t;\hat{\Omega})N_I(t;f)\right) \\ + N_I(t;f,)N_J(t;f)\Big]\left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2, \tag{A28}$$

where ϵ is defined in Eq. A20 and where $N_I(t;f)$ is the one-sided noise-power spectra:

$$N_I(t;f) \equiv 2\,|\tilde{n}_I(t;f)|^2. \tag{A29}$$

Using the same line of reasoning, we calculate the remaining terms in Eq. A24:

$$\langle\hat{W}^2\rangle - \langle\hat{W}\rangle^2 = \langle\hat{W}^{\star 2}\rangle - \langle\hat{W}^\star\rangle^2 = H(t;f)^2. \tag{A30}$$

Combining Eqs. A24 and A30, we conclude that

$$\sigma_Y^2 = \frac{1}{2}\left[\sigma_W^2 + H(t;f)^2\right]. \qquad (A31)$$

The factor of 1/2 comes about from the fact that $\hat{Y}(t;f,\hat{\Omega})$ is real whereas $\hat{W}(t;f,\hat{\Omega})$ is complex. We note that in the small-signal limit $H(f) \to 0$ and the variance reduces to the canonical stochastic result [92]:

$$\sigma_Y^2 \to \frac{1}{2}\left(N_I(t;f)N_J(t;f)\left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2\right), \qquad (A32)$$

4. Expectation value of $\hat{\sigma}_Y^2$

Our estimator for the variance of $\hat{Y}$ is given by

$$\hat{\sigma}_Y^2(t;f,\hat{\Omega}) = \frac{1}{2}\left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2 P_I^{\text{adj}}(f)P_J^{\text{adj}}(f), \quad (A33)$$

where P_I is the average auto-power in neighboring pixels:

$$P_I^{\text{adj}}(f) \equiv 2\overline{|\tilde{s}_I(f)|^2}. \qquad (A34)$$

The overline denotes an average over neighboring pixels. By averaging over neighboring pixels, we assume that the detector noise in any given pixel can be characterized by looking at its neighbors. This assumption is discussed below.

Now we calculate the expectation value of our estimator for variance $\hat{\sigma}_Y^2$ given in Eq. A33 in order to compare it to the theoretical variance given in Eqs. A31 and A28. Eqs. 3.9 and 3.10 together imply

$$\langle\hat{\sigma}_Y^2(t;f,\hat{\Omega})\rangle = 2\left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2 \\ \langle s_I^{*\text{adj}}(f)s_I^{\text{adj}}(f)s_J^{*\text{adj}}(f)s_J^{\text{adj}}(f)\rangle. \qquad (A35)$$

Using Equation A26 to write the expectation value of $\hat{\sigma}_Y^2$ in terms of the theoretical value of σ_W^2, we find

$$\langle\hat{\sigma}_Y^2(t;f,\hat{\Omega})\rangle = \frac{1}{2}\left[\sigma_W^2(t;f,\hat{\Omega}) + 4\left|\tilde{Q}_{IJ}(t;f,\hat{\Omega})\right|^2\right.$$
$$\left.\langle\tilde{s}_I^{\text{adj}*}(t;f)\tilde{s}_J^{\text{adj}}(t;f)\rangle\langle\tilde{s}_I^{\text{adj}}(t;f)\tilde{s}_J^{\text{adj}*}(t;f)\rangle\right]$$
$$= \frac{1}{2}\left[\sigma_W^2(t;f,\hat{\Omega}) + \left|\langle\hat{W}\rangle\right|^2\right]$$
$$= \frac{1}{2}\left[\sigma_W^2(t;f,\hat{\Omega}) + H(t;f)^2\right] \qquad (A36)$$

Since this is the theoretical variance from A31, we conclude that $\langle\hat{\sigma}_Y^2\rangle = \sigma_Y^2$. Thus, Eq. 3.9 provides an unbiased estimator for σ_Y^2. Here we have assumed that the noise and signal are comparable in neighboring segments. This assumption can fail for rapidly changing, high-SNR signals and also for highly non-stationary noise, and so additional work may be required to estimate σ in these situations.

5. Elliptically polarized sources

A variety of long-transient GW sources are expected to be elliptically polarized (e.g., long GRBs [4, 5] and pulsar glitches [115]). Elliptically polarized sources are parameterized by two angles. The inclination angle ι is the angle between the rotational axis of the source and the observer's line of sight and the polarization angle ψ describes the orientation of the rotational axis in the plane perpendicular to the line of sight (see, e.g., [116]).

Following [117], we characterize an elliptically polarized source with the so-called canonical amplitudes:

$$\mathcal{A}_\mu \equiv \begin{pmatrix} A_+ \cos 2\psi \\ A_+ \sin 2\psi \\ -A_\times \sin 2\psi \\ A_\times \cos 2\psi, \end{pmatrix} \qquad (A37)$$

where

$$A_+ \equiv (h_0/2)\left(1 + \cos^2\iota\right) \qquad (A38)$$
$$A_\times \equiv h_0 \cos\iota \qquad (A39)$$

We have set the initial phase $\phi_0 = 0$ for the sake of simplicity. (Ultimately, we are concerned with the average cross-power over many cycles and so the initial phase is unimportant.) Here h_0 is the strain amplitude.

Next, we define the tensor [117]

$$h_{ab}^\mu(t) \equiv \begin{pmatrix} e_{ab}^+ \cos\left[\phi(t)\right] \\ e_{ab}^\times \cos\left[\phi(t)\right] \\ e_{ab}^+ \sin\left[\phi(t)\right] \\ e_{ab}^\times \sin\left[\phi(t)\right] \end{pmatrix}, \qquad (A40)$$

where $\phi(t)$ describes the phase evolution of the signal [121]. Now we write the GW equation as [117]

$$h_{ab}(t) \equiv \sum_\mu \mathcal{A}_\mu h_{ab}^\mu(t) =$$
$$(A_+ \cos(2\psi)\cos\left[\phi(t)\right] - A_\times \sin(2\psi)\sin\left[\phi(t)\right])e_{ab}^+ +$$
$$(A_+ \sin(2\psi)\cos\left[\phi(t)\right] + A_\times \cos(2\psi)\sin\left[\phi(t)\right])e_{ab}^\times. \qquad (A41)$$

We Fourier transform a finite stretch of GW signal to obtain the coefficients $\tilde{h}_A(f)$ (see Eq. A1):

$$\tilde{h}_+(t;f) = \frac{[A_+ \cos(2\psi) + iA_\times \sin(2\psi)]\,\delta_{ff_0}}{2N_s} \qquad (A42)$$
$$\tilde{h}_\times(t;f) = \frac{[A_+ \sin(2\psi) - iA_\times \cos(2\psi)]\,\delta_{ff_0}}{2N_s}. \qquad (A43)$$

We have ignored negative frequencies, which will play no part in our subsequent calculation of one-sided GW power. Also, we have expanded $\phi(t)$ in a Taylor series,

$$\phi(t) = \phi_0 + 2\pi\left(f_0 t + \frac{1}{2}\dot{f}t^2\right) + \mathcal{O}(t^3). \qquad (A44)$$

We assume that $\dot{f}$ is approximately zero over the segment duration, T. In principle, this formulation could be extended to broadband sources, but we expect elliptically polarized sources (associated with spinning objects and a binary objects) to be narrowband.

Combining Eqs. A12 and A42, we calculate the one-sided power spectrum:

$$H_{AA'}(t;f) = \frac{\delta_{ff_0}}{2N_s^2} \left[\begin{pmatrix} A_+^2 \cos(2\psi)^2 + A_\times^2 \sin(2\psi)^2 & (A_+^2 - A_\times^2)\cos(2\psi)\sin(2\psi) \\ (A_+^2 - A_\times^2)\cos(2\psi)\sin(2\psi) & A_+^2 \sin(2\psi)^2 + A_\times^2 \cos(2\psi)^2 \end{pmatrix} + i \begin{pmatrix} 0 & -A_+A_\times \\ A_+A_\times & 0 \end{pmatrix} \right] \quad \text{(A45)}$$

The imaginary off-diagonal terms corresponds to the phase delay between $+$ and $\times$ polarization states.

Combining Eqs. A13 and A45, we obtain

$$H(t;f) = \frac{1}{2N_s^2}\left(A_+^2 + A_\times^2\right)\delta_{ff_0}. \quad \text{(A46)}$$

Our goal, once again, is to find an estimator for $H(t;f)$. We assume that the estimator for $H(t;f)$ can be con-structed from the cross power spectrum of two GW-strain channels multiplied by an appropriate filter function, $\tilde{Q}_{IJ}(t;f,\hat{\Omega},\iota,\psi)$:

$$\hat{Y}(t;f,\hat{\Omega},\iota,\psi) = 2\,\text{Re}\left[\tilde{s}_I^\star(t;f)\tilde{s}_J(t;f)\,\tilde{Q}_{IJ}(t;f,\hat{\Omega},\iota,\psi)\right]. \quad \text{(A47)}$$

Plugging Eq. A45 into Eq. A16, we obtain

$$\left\langle \hat{Y}(t;f,\hat{\Omega},\iota,\psi) \right\rangle = 2\,\text{Re}\left[\frac{1}{2}\frac{\delta_{ff_0}}{2N_s^2}\left(F_I^+ F_J^+ \left[A_+^2 \cos(2\psi)^2 + A_\times^2 \sin(2\psi)^2\right]\right.\right.$$
$$+ F_I^+ F_J^\times \left[(A_+^2 - A_\times^2)\cos(2\psi)\sin(2\psi) - iA_+A_\times\right] + F_I^\times F_J^+ \left[(A_+^2 - A_\times^2)\cos(2\psi)\sin(2\psi) + iA_+A_\times\right]$$
$$\left.\left. + F_I^\times F_J^\times \left[A_+^2 \sin(2\psi)^2 + A_\times^2 \cos(2\psi)^2\right]\right)e^{-2\pi i f(\hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c)}\,\tilde{Q}_{IJ}(t;f,\hat{\Omega},\iota,\psi)\right]. \quad \text{(A48)}$$

We write the polarized filter function as:

$$\tilde{Q}_{IJ}(t;f,\hat{\Omega},\iota,\psi) = e^{2\pi i f\hat{\Omega}\cdot\Delta\vec{x}_{IJ}/c + i\eta}/\epsilon(\hat{\Omega},\iota,\psi), \quad \text{(A49)}$$

where η is an angle arising from from the phase delay between $+$ and $\times$ polarizations. By requiring that $\langle \hat{Y}(t;f,\hat{\Omega},\iota,\psi)\rangle = H(t;f)$, it follows that

$$\epsilon(\hat{\Omega},\iota,\psi) = \left| F_I^+ F_J^+ [a_+^2 \cos(2\psi)^2 + a_\times^2 \sin(2\psi)^2] + \right.$$
$$(F_I^+ F_J^\times + F_I^\times F_J^+)(a_+^2 - a_\times^2)\cos(2\psi)\sin(2\psi) +$$
$$F_I^\times F_J^\times (a_+^2 \sin(2\psi)^2 + a_\times^2 \cos(2\psi)^2)$$
$$\left. + ia_+a_\times(F_I^\times F_J^+ - F_I^+ F_J^\times)\right|$$
$$/(a_+^2 + a_\times^2), \quad \text{(A50)}$$

and

$$\eta(\hat{\Omega},\iota,\psi) = -\text{phase}\left(\left[F_I^+ F_J^+ [a_+^2 \cos(2\psi)^2 + a_\times^2 \sin(2\psi)^2] + \right.\right.$$
$$(F_I^+ F_J^\times + F_I^\times F_J^+)(a_+^2 - a_\times^2)\cos(2\psi)\sin(2\psi) +$$
$$F_I^\times F_J^\times (a_+^2 \sin(2\psi)^2 + a_\times^2 \cos(2\psi)^2)$$
$$\left. + ia_+a_\times(F_I^\times F_J^+ - F_I^+ F_J^\times)\right]$$
$$\left./(a_+^2 + a_\times^2)\right), \quad \text{(A51)}$$

where we have defined

$$a_+ \equiv A_+/h_0 \quad \text{(A52)}$$
$$a_\times \equiv A_\times/h_0. \quad \text{(A53)}$$

As a sanity check, we note that an elliptically polarized source with pure $+$-polarization yields a sensible efficiency

$$\epsilon(\hat{\Omega},\iota=90°,\psi=0) = F_I^+ F_J^+. \quad \text{(A54)}$$

Appendix B: Data processing details

1. Overview

In this we describe the detailed procedure for the creation of the $52\,\text{s} \times 0.25\,\text{Hz}$ coarse-grained data in Fig. 1.

- Since $n(t)$ is drawn from a zero-mean, unit-variance normal distribution, the frequency domain data will be zero-mean with a standard deviation scaled by $\sqrt{N/2}$ for the real and imaginary parts separately where N is the number of sampling points.

- To calculate the cross-spectral density, CSD, data are first Hann windowed and zero padded in the time domain. Discrete Fourier transforms are multiplied to form cross-power, then summed (with a window that tapers at the ends) to yield a CSD with 0.25 Hz resolution.

- To calculate each $P_I(f)$, data are broken into 4 s blocks with 50% overlap. Each block is Hann windowed, and the resulting power samples from all blocks are summed so that $P_I(f)$ also has 0.25 Hz resolution.

- Hann windowing reduces the overall power by a factor of 3/8. To compensate both the CSD and $P_I(f)$ are corrected for this factor.

- Without windowing or overlap each power estimate from each data block will be χ^2-distributed with 2 degrees of freedom and scaled by $N/2$, and so the average of M blocks are χ^2-distributed with $2M$ degrees of freedom and scaled by $N/2M$.

- Overlap changes the effective number of degrees of freedom in a way that can be calculated numerically (see App. B 2).

- Without windowing, each cross-power product will have zero-mean and a standard deviation scaled by $N/\sqrt{2}$. Summing M values gives each CSD estimate a standard deviation scaled by $N\sqrt{M/2}$.

- These scaling behaviors indicate how we could expect SNR to behave, except for the non-trivial effects of windowing, zero-padding and weighted averaging. These somewhat subtle effects are evaluated explicitly in App. B 2.

2. Calculation of the normalization factor

Here we derive the stretch factor applied to the signal to noise ratio (SNR) histogram in Fig. 1 generated using simulated Gaussian noise. Let $\{x_n\}$ and $\{y_m\}$ be real, discrete, time sequences corresponding to two independent data streams with each x_n and y_m sampled from a (0,1)-Gaussian distribution. The time-domain $\{x_n\}$

and $\{y_m\}$ can be transformed into frequency-domain sequences of N complex numbers $\{X_p\}$ and $\{Y_q\}$ using the discrete Fourier transform (DFT), defined here as:

$$X_p = \Delta t \sum_{n=0}^{N-1} \exp\left(\frac{-2\pi i p n}{N}\right) x_n \quad \text{and}$$
$$Y_q = \Delta t \sum_{m=0}^{N-1} \exp\left(\frac{-2\pi i q m}{N}\right) y_m \tag{B1}$$

The inverse DFT can be used to express the original discrete-time signal as

$$x_n = \Delta f \sum_{p=0}^{N-1} \exp\left(\frac{2\pi i p n}{N}\right) X_p \quad \text{and}$$
$$y_m = \Delta f \sum_{q=0}^{N-1} \exp\left(\frac{2\pi i q m}{N}\right) Y_q \tag{B2}$$

To see the effects of zero padding, we define the real, 2N-periodic discrete-time signal $\tilde{x} \equiv \{x_n, 0_n\}$ where 0_n is a sequence of N-many zeros. Then

$$\tilde{X}_p = \Delta t \sum_{n=0}^{2N-1} \exp\left(\frac{-2\pi i p n}{2N}\right) \tilde{x}_n$$
$$= \Delta t \sum_{n=0}^{N-1} \exp\left(\frac{-2\pi i p n}{2N}\right) x_n \tag{B3}$$

where

$$\tilde{x}_n = \begin{cases} x_n = \frac{\Delta f}{2} \sum_{p=0}^{2N-1} \exp\left(\frac{2\pi i p n}{2N}\right) \tilde{X}_p \\ \\ 0_n = \frac{\Delta f}{2} \sum_{p=0}^{2N-1} (-1)^p \exp\left(\frac{2\pi i p n}{2N}\right) \tilde{X}_p \end{cases} \tag{B4}$$

with $\quad 0 \leq n \leq N-1$

From Eq. B4 we conclude that the X_p with $0 \leq p \leq 2N-1$ are not linearly independent. However, the X_{2p} $(0 \leq p \leq N-1)$ and X_{2p+1} $(0 \leq p \leq N-1)$ are separately linearly independent. In the calculations below, we use only the even terms to express both the even and odd X_k terms. Therefore, taking into account the zero-padding, we can write

$$x_n = \Delta f \sum_{p=0}^{N-1} \exp\left(\frac{2\pi i (2p) n}{2N}\right) \tilde{X}_{2p}$$
$$= \Delta f \sum_{p=0}^{N-1} \exp\left(\frac{2\pi i p n}{N}\right) \tilde{X}_{2p} \tag{B5}$$

By comparing x_n in Eqs. B2 B5 we see that

$$\tilde{X}_{2p} = X_p \tag{B6}$$

In addition to the zero padding, the windowing effect can be expressed by defining $\tilde{u} \equiv \{u_n, 0_n\} \equiv \{x_n f(n), 0_n\}$

with the discrete Fourier transform given by

$$
\begin{aligned}
U_k &= \Delta t \sum_{n=0}^{2N-1} \exp\left(\frac{-2\pi ikn}{2N}\right) \tilde{u}_n \\
&= \Delta t \sum_{n=0}^{N-1} \exp\left(\frac{-2\pi ikn}{2N}\right) u_n \\
&= \Delta t \sum_{n=0}^{N-1} \exp\left(\frac{-2\pi ikn}{2N}\right) f(n)x_n
\end{aligned}
\tag{B7}
$$

where x_n is given by Eq. B5 and $f(n)$ is the Hann window defined by

$$
\begin{aligned}
f(n) &\equiv \frac{1}{2}\left[1 - \cos\left(\frac{2\pi n}{N}\right)\right] \\
&= \sum_{a=-1}^{1} \frac{1}{2}\left(\frac{-1}{2}\right)^{|a|} \exp\left(\frac{2\pi ina}{N}\right)
\end{aligned}
\tag{B8}
$$

with $\quad 0 \le n \le N-1$

After some algebra Eq. B5, Eq. B6, Eq. B7 and Eq. B8 give

$$
\begin{aligned}
U_k = \frac{\Delta f \Delta t}{2} \sum_{n=0}^{N-1}\sum_{p=0}^{N-1}\sum_{a=-1}^{1} X_p \\
\exp\left(\frac{-2\pi in(k/2-p-a)}{N}\right)\left(-\frac{1}{2}\right)^{|a|}
\end{aligned}
\tag{B9}
$$

Similarly, we can express the second data stream by $\tilde{v} \equiv \{v_m, 0_m\} \equiv \{y_m f(m), 0_m\}$ with $0 \le m \le N-1$. In analogy with B9 the discrete Fourier transform is given by

$$
\begin{aligned}
V_k = \frac{\Delta f \Delta t}{2} \sum_{m=0}^{N-1}\sum_{q=0}^{N-1}\sum_{b=-1}^{1} Y_q \\
\exp\left(\frac{-2\pi im(k/2-q-b)}{N}\right)\left(-\frac{1}{2}\right)^{|b|}
\end{aligned}
\tag{B10}
$$

The sequences $\{\tilde{u}_n\}$ and $\{\tilde{v}_n\}$ are the real, 2N-periodic, discrete-time signals that are used in the code for the calculation of the contribution of the (CSD) to the (SNR).

From the definition

$$
\begin{aligned}
SNR(t;f) = \mathrm{Re}\Bigg| \frac{\mathrm{CSD}_{xy}(t;f)}{\sqrt{\mathrm{PSD}_x(t;f)\mathrm{PSD}_y(t;f)}} \\
\mathrm{phase}(Q(t;f,\hat{\Omega}))\Bigg|,
\end{aligned}
\tag{B11}
$$

where $Q(t;f,\hat{\Omega})$ is the filter function, we observe that

in order to explain the SNR distribution obtained from the processed data we need to understand the statistical consequences of the code that calculates the power spectral density PSD_x (or PSD_y) and the CSD_{xy}.

Since the CSD is the Fourier transform of the cross-covariance function, in order to characterize the numerator of SNR we need to study terms of the form $X_k^* Y_k$. Without windowing and without zero-padding the $X_k^* Y_k$ term depends only on the frequency k as the following result shows (obtained using Eq. B1 and Eq. B2).

$$
\begin{aligned}
X_k^* Y_k &= \sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1} X_p^* Y_q \exp\left(\frac{2\pi in(k-p)}{N}\right) \\
&\quad \exp\left(\frac{-2\pi im(k-q)}{N}\right) \\
&= \sum_{p=0}^{N-1}\sum_{q=0}^{N-1} \delta_{pk}\delta_{qk} X_p^* Y_q
\end{aligned}
\tag{B12}
$$

However, when we take into account the effects of windowing and zero-padding, the $U_k^* V_k$ term, contrary to the $X_k^* Y_k$ term, has contributions from a wider part of the frequency domain according to the following result

$$
\begin{aligned}
U_k^* V_k &= \frac{1}{4} \sum_{n=0}^{N-1}\sum_{m=0}^{N-1}\sum_{p=0}^{N-1}\sum_{q=0}^{N-1}\sum_{a=-1}^{1}\sum_{b=-1}^{1} X_p^* Y_q \left(-\frac{1}{2}\right)^{|a|+|b|} \\
&\quad \exp\left(\frac{2\pi in(k/2-p-a)}{N}\right)\exp\left(\frac{-2\pi im(k/2-q-b)}{N}\right)
\end{aligned}
\tag{B13}
$$

To imitate the way the code is calculating the CSD, we define

$$
Z_j \equiv \sum_{k=-13}^{+13} w(k) U_{j+k}^* V_{j+k}
\tag{B14}
$$

where

$$
w(k) \equiv (\frac{1}{2}, \underbrace{1, 1, ..., 1}_{25-\text{times}}, \frac{1}{2})
\tag{B15}
$$

It is easy to see that the expectation value of Z_j is zero, since the X_q and Y_q are statistically independent with mean zero. To evaluate the standard deviation of $Z_j/\sqrt{26}$, where the 26 is a normalization constant due to averaging, we then need to evaluate $\langle \sqrt{Z_j^* Z_j/26} \rangle$ (where there is no sum over j). After some calculations, the final expression for $\langle Z_j^* Z_j \rangle$ becomes

$$\langle Z_j^* Z_j \rangle = \sum_{k,k',\alpha,\alpha',\beta,\beta'} w(k)w(k')\Phi^*(\alpha,\alpha')\Phi(\beta,\beta')$$

$$\left(\frac{1}{4}\right)^2 \left(-\frac{1}{2}\right)^{(|a|+|b|+|a'|+|b'|)} \tag{B16}$$

where

$$\Phi(\alpha,\alpha') \equiv \frac{\sin\left[(\pi/2)(k-k'+2\alpha-2\alpha')\right]}{\sin\left[(\pi/2N)(k-k'+2\alpha-2\alpha')\right]}$$
$$\exp\left[-i\pi(\alpha-\alpha')\left(1-\frac{1}{N}\right)\right] \tag{B17}$$

Using $N = 212,992$, corresponding to $52\,\mathrm{s}$ of data sampled at $4,096\,\mathrm{Hz}$, the numerical result turned out to be

$$\sqrt{\frac{1}{26}\langle Z_j^* Z_j \rangle} = 0.7208. \tag{B18}$$

Multiplying this by the normalization factor of 8/3 due to Hann window we get 1.922—the normalization factor applied to MC in Fig 1.

[1] J. Abadie et al. (The LIGO Scientific Collaboration), Phys. Rev. D **81**, 102001 (2010).
[2] C. D. Ott, Class. Quantum Grav. **26**, 063001 (2009).
[3] A. Corsi and P. Mészáros, Astrophys. J. **702**, 1171 (2009).
[4] A. L. Piro and E. Pfahl, Astrophys. J. **658**, 1173 (2007).
[5] M. van Putten, Astrophys. J. Lett. **575**, 71 (2002).
[6] R. Prix, S. Giampanis, and C. Messenger, *Search method for long-duration gravitational-wave transients from neutron stars*, In preparation, LIGO-P1100002-v1 (2011).
[7] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **76**, 082003 (2007), astro-ph/0703234.
[8] W. G. Anderson, P. R. Brady, J. D. E. Creighton, and É. É. Flanagan, Phys. Rev. D **63**, 042003 (2001).
[9] S. R. Deans, *The Radon Transform and Some of Its Applications* (Dover Publications, Inc., 2007), 2nd ed.
[10] P. Raffai et al., Class. Quantum Grav. **24**, S457 (2007).
[11] S. E. Woosley and J. S. Bloom, Ann. Rev. Astron. Astrophys. **44**, 507 (2006).
[12] H.-T. Janka, K. Langanke, A. Marek, G. Martínez-Pinedo, and B. Müller, Phys. Rep. **442**, 38 (2007).
[13] C. D. Ott, Class. Quantum Grav. **26**, 204015 (2009).
[14] H. Dimmelmeier, C. D. Ott, A. Marek, and H.-T. Janka, Phys. Rev. D **78**, 064056 (2008).
[15] L. Dessart, A. Burrows, E. Livne, and C. D. Ott, Astrophys. J. **645**, 534 (2006).
[16] E. Müller, M. Rampp, R. Buras, H.-T. Janka, and D. H. Shoemaker, Astrophys. J. **603**, 221 (2004).
[17] J. W. Murphy, C. D. Ott, and A. Burrows, Astrophys. J. **707**, 1173 (2009).
[18] A. Marek, H.-T. Janka, and E. Müller, Astron. Astrophys. **496**, 475 (2009).
[19] K. Kotake, W. Iwakami, N. Ohnishi, and S. Yamada, Astrophys. J. Lett. **697**, 133 (2009).
[20] C. D. Ott, A. Burrows, L. Dessart, and E. Livne, Phys. Rev. Lett. **96**, 201102 (2006).
[21] C. D. Ott, H. Dimmelmeier, A. Marek, H.-T. Janka, I. Hawke, B. Zink, and E. Schnetter, Phys. Rev. Lett. **98**, 261101 (2007).
[22] S. Ou, J. E. Tohline, and L. Lindblom, Astrophys. J. **617**, 490 (2004).
[23] A. Burrows and J. Hayes, Phys. Rev. Lett. **76**, 352 (1996).
[24] C. L. Fryer, D. E. Holz, and S. A. Hughes, Astrophys. J. **609**, 288 (2004).
[25] M. Shibata, Y. T. Liu, S. L. Shapiro, and B. C. Stephens, Phys. Rev. D **74**, 104026 (2006).
[26] T. Takiwaki and K. Kotake, submitted to Astrophys. J. ArXiv:1004.2896 (2010).
[27] M. Obergaulinger, M. A. Aloy, H. Dimmelmeier, and E. Müller, Astron. Astrophys. **457**, 209 (2006).
[28] L. Lindblom, B. J. Owen, and S. M. Morsink, Phys. Rev. Lett. **80**, 4843 (1998).
[29] R. Bondarescu, S. A. Teukolsky, and I. Wasserman, Phys. Rev. D **79**, 104003 (2009).
[30] S. E. Woosley, Astrophys. J. **405**, 273 (1993).
[31] N. Bucciantini, E. Quataert, B. D. Metzger, T. A. Thompson, J. Arons, and L. Del Zanna, Mon. Not. R. Ast. Soc. **396**, 2038 (2009).
[32] T. A. Thompson, P. Chang, and E. Quataert, Astrophys. J. **611**, 380 (2004).
[33] A. Burrows, L. Dessart, E. Livne, C. D. Ott, and J. Murphy, Astrophys. J. **664**, 416 (2007).
[34] L. Dessart, C. D. Ott, A. Burrows, S. Rosswog, and E. Livne, Astrophys. J. **690**, 1681 (2009).
[35] A. Corsi and P. Mészáros, Class. Quantum Grav. **26**, 204016 (2009).
[36] T. Fischer, S. C. Whitehouse, A. Mezzacappa, F.-K. Thielemann, and M. Liebendörfer, Astron. Astrophys. **499**, 1 (2009).
[37] E. O'Connor and C. D. Ott, Submitted to the Astrophys. J., arxiv:1010.5550 (2010).
[38] A. Burrows and J. M. Lattimer, Astrophys. J. **307**, 178 (1986).
[39] W. Zhang, S. E. Woosley, and A. Heger, Astrophys. J. **679**, 639 (2008).
[40] K. Sumiyoshi, C. Ishizuka, A. Ohnishi, S. Yamada, and H. Suzuki, Astrophys. J. Lett. **690**, 43 (2009).
[41] W. Keil, H.-T. Janka, and E. Müller, Astrophys. J. Lett. **473**, 111 (1996).
[42] J. A. Miralles, J. A. Pons, and V. A. Urpin, Astrophys. J. **543**, 1001 (2000).
[43] J. A. Miralles, J. A. Pons, and V. Urpin, Astron. Astrophys. **420**, 245 (2004).
[44] E. Müller and H.-T. Janka, Astron. Astrophys. **317**, 140 (1997).
[45] C. Thompson and R. C. Duncan, Astrophys. J. **408**, 194 (1993).
[46] C. D. Ott, A. Burrows, T. A. Thompson, E. Livne, and

R. Walder, Astrophys. J. Suppl. Ser. **164**, 130 (2006).

[47] S. Akiyama, J. C. Wheeler, D. L. Meier, and I. Lichtenstadt, Astrophys. J. **584**, 954 (2003).

[48] A. L. Watts, N. Andersson, and D. I. Jones, Astrophys. J. Lett. **618**, 37 (2005).

[49] S. Scheidegger, R. Käppeli, S. C. Whitehouse, T. Fischer, and M. Liebendörfer, Astron. Astrophys. **514**, A51 (2010).

[50] L. Baiotti, R. De Pietri, G. M. Manca, and L. Rezzolla, Phys. Rev. D **75**, 044023 (2007).

[51] M. Shibata and Y.-I. Sekiguchi, Phys. Rev. D **71**, 024014 (2005).

[52] D. Lai and S. L. Shapiro, Astrophys. J. **442**, 259 (1995).

[53] D. Lai, in *AIP Conf. Proc. 575: Astrophysical Sources for Ground-Based Gravitational Wave Detectors*, edited by J. M. Centrella (2001), p. 246.

[54] N. Andersson, Astrophys. J. **502**, 708 (1998).

[55] S. Chandrasekhar, Astrophys. J. **161**, 561 (1970).

[56] J. L. Friedman and B. F. Schutz, Astrophys. J. **222**, 281 (1978).

[57] B. J. Owen, L. Lindblom, C. Cutler, B. F. Schutz, A. Vecchio, and N. Andersson, Phys. Rev. D **58**, 084020 (1998).

[58] R. Bondarescu, S. A. Teukolsky, and I. Wasserman, Phys. Rev. D **76**, 064019 (2007).

[59] J. Brink, S. A. Teukolsky, and I. Wasserman, Phys. Rev. D **70**, 121501 (2004).

[60] B. Owen and L. Lindblom, Class. Quantum Grav. **19**, 1247 (2002).

[61] L. Bildsten, Astrophys. J. Lett. **501**, 89 (1998).

[62] N. Andersson, K. Kokkotas, and N. Stergioulas, Astrophys. J. **516**, 307 (1999).

[63] N. Andersson, K. Kokkotas, and B. F. Schutz, Astrophys. J. **510**, 846 (1999).

[64] R. Popham, S. E. Woosley, and C. Fryer, Astrophys. J. **518**, 356 (1999).

[65] M. van Putten, Phys. Rev. Lett. **87**, 091101 (2001).

[66] M. van Putten, Astrophys. J. Lett. **684**, 91 (2008).

[67] K. Kiuchi, Y. Sekiguchi, M. Shibata, and K. Taniguchi, Phys. Rev. D **80**, 064037 (2009).

[68] B. Vaishnav, I. Hinder, D. Shoemaker, and F. Herrmann, Class. Quantum Grav. **26**, 204008 (2009).

[69] J. Levin and H. Contreras, *Inspiral of generic black hole binaries: Spin, precession, and eccentricity*, http://arxiv.org/abs/1009.2533 (2010).

[70] R. M. O'Leary, B. Kocsis, and A. Loeb, Mon. Not. R. Ast. Soc. **395**, 2127 (2009).

[71] A. Krawczyk, A. G. Lyne, J. A. Gil, and B. C. Joshi, Mon. Not. R. Ast. Soc. **340**, 1087 (2003).

[72] P. D. Morley and R. García-Pelayo, Europhys. Lett. **23**, 185 (1993).

[73] P. M. McCulloch, P. A. Hamilton, D. McConnell, and E. A. King, Nature **346**, 822 (1990).

[74] G. Baym, C. Pethick, D. Pines, and M. Ruderman, Nature **224**, 872 (1969).

[75] P. W. Anderson and N. Itoh, Nature **256**, 25 (1975).

[76] M. Ruderman, T. Zhu, and K. Chen, Astrophys. J. **492**, 267 (1998).

[77] T. Takatsuka and R. Tamagaki, Prog. Theor. Phys. **79**, 274 (1988).

[78] Y. Mochizuki, T. Izuyama, and I. Tanihata, Astrophys. J. **521**, 281 (1999).

[79] N. Andersson, G. L. Comer, and D. Langlois, Phys. Rev. D **66**, 104002 (2002).

[80] B. Abbott et al. (LIGO Scientific Collaboration), Astrophys. J. Lett. **701**, 68 (2009).

[81] B. P. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. Lett. **101**, 211102 (2008).

[82] K. D. Kokkotas, E. Gaertig, and A. Colaidua, J. Phys.: Conf. Ser. **222**, 012031 (2010).

[83] B. Abbott, R. Abbott, R. Adhikari, J. Agresti, P. Ajith, B. Allen, R. Amin, S. B. Anderson, W. G. Anderson, M. Arain, et al., Phys. Rev. D **76**, 062003 (2007).

[84] K. Glampedakis, L. Samuelsson, and N. Andersson, Mon. Not. R. Ast. Soc. Lett. **371**, 74 (2006).

[85] L. Samuelsson and N. Andersson, Mon. Not. R. Ast. Soc. **374**, 256 (2007).

[86] Y. Levin, Mon. Not. R. Ast. Soc. Lett. **368**, 35 (2006).

[87] H. Sotani, K. D. Kokkotas, and N. Stergioulas, Mon. Not. R. Ast. Soc. Lett. **385**, 5 (2008).

[88] B. J. Owen, PRL **95**, 211101 (2005).

[89] J. E. Horvath, Modern Physics Lett. A **20**, 2799 (2005).

[90] J. A. de Freitas Pacheco, A&A **336**, 397 (1998).

[91] K. Ioka, Mon. Not. R. Ast. Soc. **327**, 639 (2001).

[92] B. Allen and J. D. Romano, Phys. Rev. D **59**, 102001 (1999).

[93] E. Thrane et al., Phys. Rev. D **80**, 122002 (2009).

[94] S. Shapiro and S. Teukolsky, *Black Holes, White Dwarfs, and Neutron Stars* (Wiley, 1983).

[95] J. Sylvestre, Phys. Rev. D **68**, 102005 (2003).

[96] P. M. Kalmus, Ph.D. thesis, Columbia University (2008).

[97] B. Abbott et al. (LIGO Scientific Collaboration), Nature **460**, 990 (2009).

[98] *Gravitational wave signal catalog*, http://stellarcollapse.org/gwcatalog/ott2009 (2009).

[99] J. R. Gair, I. Mandel, and L. Wen, Class. Quantum Grav. **25**, 184031 (2008).

[100] C. Høilund, *The radon transform*, http://www.cvmt.dk/education/teaching/e07/MED3/IP/Carsten_Hoilund_-_Radon_Transform.pdf (2007).

[101] S. Ballmer, Ph.D. thesis, Massachusetts Institute of Technology (2006).

[102] R. Honda, S. Yamagishi, and N. Kanda (TAMA), Class. Quantum Grav. **25**, 184035 (2008).

[103] R. Khan and S. Chatterji, Class. Quantum Grav. **26**, 155009 (2009).

[104] J. Slutsky et al., Class. Quantum Grav. **27**, 165023 (2010).

[105] N. Christensen (LIGO Scientific Collaboration and the Virgo Collaboration), Submitted to GWDAW14 proceedings, LIGO Document P1000045 (2010).

[106] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. D **79**, 022001 (2009).

[107] B. Abbott et al. (LIGO Scientific Collaboration), Phys. Rev. Lett. **102**, 111102 (2009), URL http://link.aps.org/doi/10.1103/PhysRevLett.102.111102.

[108] L. Blackburn et al. (LIGO Scientific Collaboration), Class. Quantum Grav. **25**, 184004 (2008), URL http://www.iop.org/EJ/abstract/0264-9381/25/18/184004/.

[109] M. Coughlin (LIGO Scientific Collaboration and the Virgo Collaboration), Submitted to GWDAW14 proceedings, LIGO Document P1000044 (2010), URL https://dcc.ligo.org/cgi-bin/private/DocDB/ShowDocument?docid=10764.

[110] B. Abbott et al. (LIGO Scientific Collaboration), Rep. Prog. Phys. **72**, 076901 (2009).

[111] F. Acernese et al. (VIRGO Scientific), Class. Quantum Grav. **25**, 114045 (2008).

[112] E. Goetz and K. Riles, LIGO DCC T050174-00-D (2010).

[113] K. Riles, personal communication (2010).

[114] A. Freise et al., CQG **26**, 085012 (2009).

[115] C. A. van Eysden and A. Melatos, Class. Quantum Grav. **25**, 225020 (2008).

[116] R. Prix and B. Krishnan, Class. Quantum Grav. **26**, 204013 (2009).

[117] J. T. Whelan, R. Prix, and D. Khurana, Class. Quantum Grav. **25**, 184029 (2009).

[118] For the sake of simplicity, the calculations in this section ignore effects from windowing and the use of overlapping segments.

[119] In order to chose a suitable number of neighboring pixels to average over, one must typically take into account the stationarity of the detector noise. This discussion, however, is outside our current scope.

[120] Here we also assume that the probability density function for each pixel is the same, which is to say that the noise and signal are approximately stationary.

[121] The reader is advised to use caution when comparing our formulas with those in [117]. Following [7, 92, 93], we use e_{ab}^{A} to denote the "detector polarization basis" defined by the ecliptic plane. In [117], however, this quantity is denoted ϵ_{ab}^{A} and e_{ab}^{A} is the basis in which the plus polarization is maximal.