

Time-frequency analysis of gravitational wave data

Neil J. Cornish 

*eXtreme Gravity Institute, Department of Physics, Montana State University,
Bozeman, Montana 59717, USA*



(Received 28 August 2020; accepted 17 November 2020; published 16 December 2020)

Data from gravitational wave detectors are recorded as time series that include contributions from myriad noise sources in addition to any gravitational wave signals. When regularly sampled data are available, such as for ground based and future space based interferometers, analyses are typically performed in the frequency domain, where stationary (time invariant) noise processes can be modeled very efficiently. In reality, detector noise is not stationary due to a combination of short duration noise transients and longer duration drifts in the power spectrum. This nonstationarity produces correlations across samples at different frequencies, obviating the main advantage of a frequency domain analysis. Here an alternative time-frequency approach to gravitational wave data analysis is proposed that uses discrete, orthogonal wavelet wave packets. The time domain data is mapped onto a uniform grid of time-frequency pixels. For locally stationary noise—that is, noise with an adiabatically varying spectrum—the time-frequency pixels are uncorrelated, which greatly simplifies the calculation of quantities such as the likelihood. Moreover, the gravitational wave signals from binary systems can be compactly represented as a collection of lines in time-frequency space, resulting in a computational cost for computing waveforms and likelihoods that scales as the square root of the number of time samples, as opposed to the linear scaling for time or frequency based analyses. Key to this approach is having fast methods for computing binary signals directly in the wavelet domain. Multiple fast transform methods are developed in detail.

DOI: [10.1103/PhysRevD.102.124038](https://doi.org/10.1103/PhysRevD.102.124038)

I. INTRODUCTION

Gravitational wave data analysis employs many of the standard tools of time-series analysis. With some exceptions, the majority of gravitational wave data analysis is performed in the frequency domain. Notable exceptions include the time domain analysis of pulsar timing array data [1], time domain low-latency LIGO/Virgo searches [2,3], analysis of black hole ringdowns [4,5], and wavelet domain searches for gravitational wave transients [6,7]. The primacy of frequency domain analyses is due to the advantages it confers for modeling detector noise, under the assumption that the noise properties are at least approximately stationary [8]. Since most analyses are performed in the frequency domain, significant effort has gone into developing frequency domain waveform models [9–17].

As ground based interferometers continue to improve in sensitivity, especially at low frequencies, and with the advent of future space based interferometers, the duration that gravitational wave signals spend in the sensitive band of the detectors will increase. The assumption that the noise can be treated as approximately stationary will break down for these long duration signals. It is time for gravitational wave data analysis to leave the frequency domain.

The wavelet domain provides a time-frequency representation of that data that is well suited for modeling nonstationary noise processes. Indeed, for certain discrete

wavelet bases, and for noise with statistical properties that vary slowly in time, the noise correlation matrices are diagonal in the wavelet domain, with diagonal entries given by the dynamic, or evolutionary power spectrum [18] of the noise, $S(f, t)$. The Wilson-Daubechies-Meyer (WDM) wavelet basis [19] employed by the coherent WaveBurst algorithm [6,7] is a good choice for more general gravitational wave analyses. The WDM wavelet basis provides excellent frequency separation and a uniform tiling in time and frequency. The coherent WaveBurst algorithm uses the WDM basis to represent both data and signals—all data conditioning, power spectral estimation and likelihood calculations are carried out in the wavelet domain. In contrast to the coherent WaveBurst algorithm, which reconstructs signals as a collection of WDM wavelets, here the WDM basis simply replaces the Fourier basis—any signal model, including waveform templates derived from general relativity, can be mapped to the WDM basis. A key ingredient for the wavelet domain analysis is having waveform models that can be computed directly and efficiently in the wavelet domain. For the most commonly encountered type of gravitational wave signal—produced by the merger of compact binary stars on quasicircular orbits—it turns out that the waveforms can be computed much more efficiently in the wavelet domain than in the time or frequency domains. The computational cost in the

wavelet domain scales as the square root of the observation time, as opposed to linearly with time as is the case in the time and frequency domains. Methods are introduced for computing binary inspiral-merger-ringdown waveforms directly and efficiently in the wavelet domain.

II. MOTIVATION

Gravitational wave data is recorded as a collection of time series from a network of detectors. Evenly spaced data from interferometric detectors are amenable to the standard techniques of time series analysis, such as linear filtering and spectral analysis [20,21]. A key quantity that appears in both frequentist and Bayesian approaches to time series analysis is the noise weighted inner product, $(a|b)$, of time series a and b , given by the expression

$$(a|b) = \mathbf{a}^\dagger \cdot \mathbf{C}^{-1} \cdot \mathbf{b} = a_i C_{ij}^{-1} b^j, \quad (1)$$

where C_{ij} is the noise correlation matrix. Here the sums run over the individual data samples from each detector in the array. Important examples where this inner product is used include the match-filter signal-to-noise ratio

$$\rho(h) = \frac{(d|h)}{\sqrt{(h|h)}}, \quad (2)$$

where d is the data and h is a waveform model for the signal, and the likelihood function for Gaussian noise

$$p(d|h) = \frac{1}{\sqrt{\det(2\pi\mathbf{C})}} e^{-\frac{1}{2}(\mathbf{d}-\mathbf{h})^\dagger \cdot \mathbf{C}^{-1} \cdot (\mathbf{d}-\mathbf{h})}. \quad (3)$$

Several challenges present themselves when computing these quantities. The first is that the noise correlation matrix is not known in advance and has to be estimated from the data. The second is the computation cost of evaluating the inverse of the noise correlation matrix and the cost of evaluating the sum that appears in (1). Interferometer noise is usually highly colored, with both red and blue components, leading to strong correlations between data samples in the time domain. With N data samples, the cost of inverting the noise correlation matrix in the time domain scales as N^3 , while the cost of the sums scale as N^2 . It is also very difficult to estimate the N^2 elements of the correlation matrix from the N observed data points. One solution to these problems is to apply a linear transform to the time series: $\tilde{\mathbf{a}} = \mathbf{Q} \cdot \mathbf{a}$ that diagonalizes the noise correlation matrix:

$$(a|b) = \tilde{\mathbf{a}}^\dagger \cdot (\mathbf{Q}^\dagger \mathbf{C}^{-1} \mathbf{Q}) \cdot \tilde{\mathbf{b}} = \frac{\tilde{a}_i \tilde{b}_i}{\sigma_i^2}, \quad (4)$$

where σ_i^2 are the diagonal entries of the transformed noise correlation matrix. The computation cost of the matrix

inversion and sum now scales linearly with the number of data points N . The cost of the matrix inversion has been traded for the cost of finding the linear transform \mathbf{Q} . If the data happen to be stationary, that is, if the correlations between the time domain samples only depends on the time interval between samples, and not the sample times themselves, then \mathbf{Q} can be found using several different methods. One method is to fit an auto regressive (AR) or auto regressive moving average model to the data, then use the inverse of this model to prewhiten the data [22,23]. A simpler method is to Fourier transform the data, since for stationary time series the noise correlation matrix is diagonal in the frequency domain, with $\tilde{C}_{ij} = S(f_i) \delta_{ij}$, where $S(f)$ is the Fourier power spectrum. The power spectrum can be estimated by a variety of methods, including classical methods such as Welch averaging, or Bayesian methods that use parametric [24] or semiparametric [25] modeling.

In reality, the noise encountered in gravitational wave data analysis is nonstationary [26–28], which significantly complicates the estimation of the noise correlation matrix and the calculation of noise weighted inner products. If the noise properties vary slowly in time, and are approximately constant over the time a signal is in the sensitive band of the detectors, then it is possible to treat the data as stationary in the short data segment containing the signal. The data is then analyzed in short chunks, with the power spectral density (PSD) recomputed for each chunk. This approach becomes increasingly inaccurate as the duration of the signals increases. For example, binary neutron star signals will be in-band for minutes or hours as the low frequency sensitivity of ground based detectors improves. For space based detectors such as LISA, the signals will be in band for months or years. Even if the LISA instrument noise is perfectly stationary (unlikely!), the residual signal from unresolved galactic binaries will be highly nonstationary [29].

The usual frequency domain approach fails for nonstationary data as the noise correlation matrix is no longer diagonal. When the variations in the noise properties are small, the frequency domain noise correlation matrix becomes band diagonal. A $N \times N$ band diagonal matrix with bandwidth n can be decomposed into a product of upper and lower triangular matrices in order $n^2 N$ operations [30]. These triangular matrices can be inverted in order N^2 operations, resulting in the total cost of the inverse scaling as N^2 . Since the inverse of a band diagonal matrix is generally not band diagonal, the cost of the sum and the inverse in (1) will both scale as N^2 . Added to this increased computational cost is the challenge of estimating the $\mathcal{O}(nN)$ off-diagonal elements of the noise correlation matrix.

Nonstationary noise has been shown to impact frequency domain matched filter searches for gravitational waves, resulting in significant losses of efficiency for long duration signals [31,32]. It was found the loss of efficiency could be

mitigated to some extent by applying PSD drift corrections [31] or by reranking the search statistic [32]. These workarounds avoid the more costly step of using a nondiagonal noise correlation matrix, but it is unclear if these approaches could be adapted for use in other settings, such as Bayesian parameter estimation.

An alternative approach that has been proposed is to generalize the autoregressive moving average prewhitening to allow for nonstationary noise [33,34]. It is unclear how the computational cost of such an approach will scale given the added complexity of the nonstationarity, but it is likely to be significantly more expensive than the standard frequency space analysis. As a point of comparison, fitting an $AR(p)$ model to a stationary time series requires order pN operations to compute the p lagged autocorrelation functions, and order p^2 operations to solve the Yule-Walker equations to find the model coefficients. For typical LIGO-Virgo spectra the model order p will be order N . Note that both the data and the waveform templates have to be filtered by the model at a cost that scales as pN . The computational cost for nonstationary time series will likely be significantly higher.

Given that the noise properties vary with time, it is natural to consider wavelet based time-frequency methods. A discrete wavelet (or wavelet wave packet) transformation provides a linear mapping from the time domain to the two-dimensional time-frequency domain indexed by time n and frequency m . With an appropriate wavelet basis, and for a wide array of locally nonstationary processes [35,36], the wavelet noise correlation matrix is diagonal:

$$C_{(nm)(n'm')} = \delta_{nn'}\delta_{mm'}S(t_n, f_m), \quad (5)$$

where $S(t, f)$ is the dynamic, or evolutionary, power spectrum [18]. Given the diversity of nonstationary processes, there is no general proof that the wavelet correlation matrix will always be diagonal, but proofs exist for certain cases. To understand, at least heuristically, why the wavelet noise correlation matrix might be diagonal, it is helpful to recall why the Fourier domain noise correlation matrix is nondiagonal for nonstationary processes. In the Fourier domain, information about when an event occurred is encoded in the Fourier phase, not the Fourier spectrum $S(f)$. For stationary processes, the phase in each Fourier bin is uncorrelated, and the noise correlation matrix is diagonal. For nonstationary processes the time dependence leads to correlations in the Fourier phases, and this gets encoded in the off-diagonal elements of the noise correlation matrix. In contrast, the wavelet domain has an evolutionary spectrum $S(t, f)$ that can account for certain types of nonstationarity, so the time dependence does not have to be encoded in the off-diagonal elements of the noise correlation matrix. The noise correlation matrix has been shown to be orthogonal for waveletlike bases, such as certain types of windowed Fourier transforms [37,38]. The

WDM wavelet family [19] used in the analyses in this paper can be computed using an FFT with a specially chosen window function. Additionally, discrete wavelet bases that are well localized in frequency, such as the ones used here, have been shown to yield approximately diagonal correlation matrices for locally stationary red noise processes [39]. Turning things around, the condition (5) has been used to define a class of locally stationary wavelet processes [40,41] that can be used to simulate and model nonstationary time series.

III. DISCRETE WAVELET WAVE PACKETS

Time-frequency representations of gravitational wave data using oversampled continuous wavelets (e.g., Q scans [42]) provide nice visualizations of the gravitational wave signals, and are also widely used to identify noise transients. The oversampling makes for nice smooth images, but the resulting correlations between pixels excludes these representations from being used in quantitative analyses. In contrast, orthonormal discrete wavelet transform produce poor visualizations, but provide a critically sampled representation of the data with a lossless inverse and appealing statistical properties for handling colored, nonstationary noise.

While there are many choices of discrete wavelet bases, there is one that is ideally suited for gravitational wave data analysis: the WDM wavelet wave packets introduced by Nuccia *et al.* [19]. The WDM representation produces a uniform time-frequency grid, as opposed to the more common Dyadic representations which shrink in time as they go up in frequency. The WDM wavelets provide excellent spectral separation, with a bandpass that maintains the same shape across all frequencies. Unlike typical wavelet transform that are computed using repeated application of filter function, the WDM transform can be computed using windowed fast Fourier transforms. Indeed, the WDM transform can be thought of as a short-time Fourier transform (SFT) with a well-designed window function. Because of the close relationship to SFTs, the WDM transform can be computed efficiently using FFTs, and the time translations can be incorporated using Fourier domain phase shifts [19].

The WDM wavelet wave packets form a complete, orthogonal basis that can be used to faithfully represent any time series:

$$x[k] = \sum_{n=0}^{N_t-1} \sum_{m=0}^{N_f-1} w_{nm} g_{nm}[k]. \quad (6)$$

A time series of duration T sampled at a cadence $\Delta t = T/N$ can be represented by a rectangular grid with N_t time slices of width ΔT and N_f frequency slices of width ΔF :

$$\begin{aligned} \Delta T &= N_f \Delta t, \\ \Delta F &= \frac{1}{2\Delta t N_f}. \end{aligned} \quad (7)$$

Each of the $N = N_t N_f$ cells has area $\Delta T \Delta F = 1/2$. The time and frequency resolution can be varied by changing N_f , such that as $N_f \rightarrow N/2$ the expansion approaches a Fourier series, and as $N_f \rightarrow 1$ the expansion approaches the original time series. The orthogonality condition

$$\sum_{k=0}^{N-1} g_{nm}[k] g_{pq}[k] = \delta_{np} \delta_{mq} \quad (8)$$

can be used to find an expression for the expansion coefficients:

$$w_{nm} = \sum_{k=1}^N x[k] g_{nm}[k]. \quad (9)$$

The WDM wavelets are actually a wavelet family, with their precise shapes controlled by three parameters, the half-width of the flat response in frequency, A , the width of the frequency roll-off B , and the steepness of the frequency roll-off d , subject to the constraint $2A + B = \Delta\Omega$, where $\Delta\Omega = 2\pi\Delta F$. Accordingly the wavelets are defined in the frequency domain:

$$\tilde{g}_{nm}(\omega) = e^{-in\omega\Delta T} (C_{nm} \Phi(\omega - m\Delta\Omega) + C_{nm}^* \Phi(\omega + m\Delta\Omega)), \quad (10)$$

where $C_{nm} = 1$ for $(n + m)$ even and $C_{nm} = i$ for $(n + m)$ odd. Note that the $m = 0$ frequency band needs to be handled a little differently from the others—see Ref. [19] for details. The Meyer [43] window function $\Phi(\omega)$ is defined:

$$\Phi(\omega) = \begin{cases} \frac{1}{\sqrt{\Delta\Omega}}, & |\omega| < A \\ \frac{1}{\sqrt{\Delta\Omega}} \cos \left[\nu_d \left(\frac{\pi}{2} \frac{|\omega| - A}{B} \right) \right], & A \leq |\omega| \leq A + B \end{cases}, \quad (11)$$

where $\nu_d(x)$ is the normalized incomplete Beta function

$$\nu_d(x) = \frac{\int_0^x y^{d-1} (1-y)^{d-1} dy}{\int_0^1 y^{d-1} (1-y)^{d-1} dy}. \quad (12)$$

Figure 1 shows the window function in the time domain, $\phi(t)$, and frequency domain $\Phi(f)$ for $d = 4$, $A = \Delta\Omega/4$ and $B = \Delta\Omega/2$. Formally $\phi(t)$ extends indefinitely in time, but in practice it can be truncated without incurring a significant error. For the example shown in Fig. 1, the time domain window is truncated beyond $\pm q\Delta T$ with $q = 16$. The usual rules for time-frequency localization pertain—a well-localized frequency response (large d and small B) comes at the cost of a longer window in the time domain.

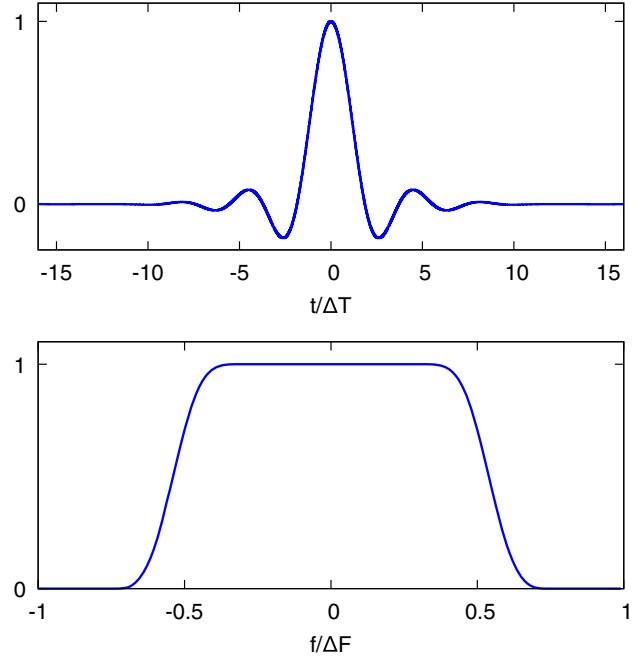


FIG. 1. The window function for the WDM wavelets for the choice $d = 4$, $A = \Delta\Omega/4$ and $B = \Delta\Omega/2$. For this choice of window parameters the wavelets are better localized in frequency than they are in time. Here the overall normalization is arbitrary.

Using the definition of the WDM wavelets (10) and the expression (9), the expansion coefficients are given by

$$w_{nm} = \sqrt{2\Delta t} \Re C_{nm} \sum_{k=-K/2}^{K/2-1} e^{i\pi k m / N_f} x[nN_f + k] \phi[k]. \quad (13)$$

Here $K = 2qN_f$ is the width of the window function, where $q > 1$ is a multiplier that sets the size of the window relative to the temporal extent of the time-frequency pixels. Defining the windowed discrete Fourier transform of $x[nN_f + k]$ as

$$X_n[j] = \sum_{k=-K/2}^{K/2-1} e^{2\pi i k j / K} x[nN_f + k] \phi[k], \quad (14)$$

we have

$$w_{nm} = \sqrt{2\Delta t} \Re C_{nm} X_n[mq]. \quad (15)$$

In other words, the WDM transform of a time series can be computed using N_t windowed short Fourier transforms of length K , then downsampling to extract every q th coefficient. The cost of these operations scale as $N_t K \ln(K) = 2qN \ln(2qN_f)$, as compared to the $N \ln(N)$ cost of a standard fast Fourier transform of the data. For the window parameters used in Fig. 1, $q = 16$ and the WDM transform is more than an order of magnitude slower than a straight FFT.

The additional cost of the time-domain WDM transform is due to the length of the time domain window function.

Since the window function is more compact in the frequency domain, it turns out to be faster to compute the transform using the discrete Fourier transform of the data $X[l]$:

$$w_{nm} = \sqrt{2}(-1)^{nm} \Re C_{nm} x_m[n], \quad (16)$$

where

$$x_m[n] = \sum_{l=-N_t/2}^{N_t/2-1} e^{-2\pi i l n / N_t} X[l + mN_t/2] \Phi[l]. \quad (17)$$

Note that the discrete Fourier samples are evaluated at $f = l\Delta f$, where $\Delta f = 1/T$, and $T = N\Delta t$ is the total span of the data. The $x_m[n]$ can be computed using a FFT of length N_t for each m . The total computational cost of the frequency domain version of the transform is then $N_f N_t \ln(N_t) = N \ln(N_t)$. This is less than the cost of the FFT of the full time series that has to be done before computing the frequency domain WDM transform. Performed this way, the total cost of the WDM transform is less than twice the cost of the full FFT.

While the gravitational wave data has to be transformed directly using the above methods, gravitational wave signals for which analytic models exist can be evaluated directly in the wavelet domain at a much reduced computational cost using techniques described in later sections.

A. Likelihood, Fisher matrix and searches

The Whittle log likelihood can be evaluated in the wavelet domain:

$$\ln p(d|h) = -\frac{1}{2} \sum_{nm} \left(\ln(2\pi S_{nm}) + \frac{(d_{nm} - h_{nm})^2}{S_{nm}} \right) \quad (18)$$

where S_{nm} is the evolutionary power spectrum. For stationary noise, the evolutionary power spectrum can be related to the usual power spectral density $S(f)$ using Eqs. (16) and (17):

$$S_{nm} = \sum_{l=-N_t/2}^{N_t/2-1} S[l + mN_t/2] \Phi[l]^2 \approx S(f_m) \Delta F, \quad (19)$$

where $f_m = mN_t/2$ is the central frequency of the wavelet band, and ΔF is the width of the band. Note that the S_{nm} are dimensionless.

The Fisher information matrix, defined as $\Gamma_{ij} = -\partial_i \partial_j \ln p(d|h)|_{ML}$, is given by

$$\Gamma_{ij} = \sum_{nm} \frac{h_{nm,i} h_{nm,j}}{S_{nm}}. \quad (20)$$

The wavelet domain Fisher matrix avoids taking derivatives of rapidly oscillating waveforms and is thus more stable

than versions based on direct numerical evaluation in the frequency domain or time domain.

The signal-to-noise statistic ρ is given by

$$\rho(h) = \sum_{nm} \frac{\hat{h}_{nm} d_{nm}}{S_{nm}}, \quad (21)$$

where $\hat{h}_{nm} = h_{nm}/(h|h)^{1/2}$ is a unit normalized waveform template. The waveform templates typically depend on multiple parameters, including intrinsic parameters such as the masses and spins of a binary system, in addition to extrinsic parameters such as the overall amplitude, phase and reference time. The later three parameters are usually analytically maximized over in a search [8] resulting in considerable computational savings. The maximization over time exploits the fact that time shifts δt appear as linear phase shifts $\exp(2\pi i f \delta t)$ in the frequency domain. Similar tricks can be used in the wavelet domain, after taking care of some complications caused by the evolutionary power spectrum S_{nm} . When the noise is nonstationary $S_{nm} \neq S_{lm}$ for $n \neq l$, and the time-translation invariance of the template normalization is broken. The time shifted normalization,

$$(h|h)_k = \sum_{nm} \frac{h_{nm} h_{nm}}{S_{(n+k)m}}, \quad (22)$$

is typically a slowly varying quantity that can be computed for a small number of time shifts k and interpolated for intermediate values. The computational cost of evaluating $(h|h)_k$ for a binary system is low since the h_{nm} are typically only nonzero for order \sqrt{N} points (more precisely, for a single harmonic the waveforms are nonzero for between $\sim N_t$ and $\sim qN_f$ points). The total cost to compute $(h|h)_k$ is less than N for waveforms with a single dominant harmonic. The inner product $(h|d)$ can be maximized with respect to the overall phase ϕ_0 using the usual method of quadrature phase waveforms $h(\phi_0 = 0)$ and $h(\phi_0 = \pi/2)$. The maximization over the time shift δt can be performed in a number of ways. The simplest and fastest method is to overwhiten the data using the evolutionary spectrum, $\bar{d}_{nm} = d_{nm}/S_{nm}$, then inverse wavelet transform the data to the frequency domain $\bar{d}(f)$. The time maximization can then be done by inverse Fourier transforming the quantity

$$\tilde{z} = 4\tilde{h}(f)\bar{d}(f)\Theta(f), \quad (23)$$

where $\Theta(f)$ is the Heaviside step function. The signal-to-noise statistic is then given by $\rho(\delta t) = |z(\delta t)|/(h|h)_k^{1/2}$ with $k = [\delta t/\Delta T]$. The cost of this procedure is roughly two to three times that of a purely frequency domain implementation due to the extra Fourier transforms involved in computing d_{nm} then $\bar{d}(f)$, and the additional cost of computing the time-shift template normalization

$(h|h)_k$. On the other hand, if the data is nonstationary on the time scale of the signal, the wavelet based search statistic will be more robust. This method shares some similarities to the PSD drift correction approach [31].

The approach described above can be incorporated in existing search pipelines with minimal changes to the existing codes. The frequency domain waveforms that are currently being used can be transformed to the wavelet domain using Eq. (16). While not as fast as the direct wavelet domain methods described in later sections, the cost is only a factor of two larger than the standard frequency domain calculation, and requires no changes to the waveform libraries that have been developed over many years. The main new ingredient is the wavelet domain whitening and associated estimation of the evolutionary spectrum S_{nm} . Methods for implementing these steps are described in the next section.

The time-shift maximization can also be performed directly in the wavelet domain by treating h_{nm} and \bar{d}_{nm} as time series in n and applying an FFT to each:

$$H_m[k] = \sum_{n=0}^{N_t-1} e^{2\pi i n k / N_t} h_{nm}, \quad (24)$$

and similarly for \bar{d}_{nm} . The inverse FFT, $z[p]$ of the quantity

$$Z[k] = \sum_{m=0}^{N_f-1} H_m[k] \bar{D}_m^*[k] \quad (25)$$

yields $\rho(p\Delta T) = |z[p]|$. The cost of the N_f FFTs used to compute H_{km} and \bar{D}_{km} scale as $N_t \ln N_t$, so the total cost of computing Z_k scales as $N \ln N_t$. The limitations of this approach are that the time maximization is quantized in units of $\Delta T = N_f \Delta t$, which is much coarser than the Δt resolution of the usual time maximization (23). The time resolution can be improved by generating p reference templates spaced by $\Delta T/p$ and repeating the maximization for each shifted template. The cost of achieving the full Δt resolution is then N_f times greater than the usual Fourier domain approach, though using lower time resolutions may be acceptable. With WDM wavelets the time maximization yields a maximization across frequency bands that may be useful for computing banded time-frequency statistics [44]. Using Eqs. (16) and (17), the signal-to-noise statistic becomes

$$\rho(h) = \frac{N_t}{2} \sum_{m=0}^{N_f} \sum_{l=-N_t/2}^{N_t/2-1} (H[l + mN_t/2] \bar{D}^*[l + mN_t/2] + H^*[l + mN_t/2] \bar{D}[l + mN_t/2]) \Phi[l]^2, \quad (26)$$

where $H[j]$ denotes the DFT of h and $\bar{D}[j]$ denotes the DFT of \bar{d} . The above expression follows from performing the sum over the time index n and using the identity

$$\sum_{n=N_t/2}^{N_t/2-1} e^{-2\pi i (l-k)n/N_t} = N_t \delta_{lk}. \quad (27)$$

Introducing the time-shifted template

$$H_{\delta t}[j] = H[j] e^{-2\pi i j \delta t / T}, \quad (28)$$

and setting $\delta t = p\Delta T$, we have

$$\rho(p\Delta T) = \sum_{m=0}^{N_f} (-1)^m u_m[p], \quad (29)$$

where $u_m[p]$ is given by the inverse FFT of

$$U_m[l] = H[l + mN_t/2] \bar{D}^*[l + mN_t/2] \Phi[l]^2. \quad (30)$$

The cost of the banded time maximization scales as $N_f N_t \ln N_t$. The time resolution of the maximization can be increased by setting $\delta t = p\Delta T/a$ for some integer a , and extending the sum over l in (26) to run from $-aN_t/2$ to $aN_t/2 - 1$ and writing

$$\rho(p\Delta T) = \sum_{m=0}^{N_f} e^{-\pi i p m / a} u_m[p]. \quad (31)$$

Since $\Phi[l]$ is zero for $|l| > N_t/2$ the $U_m[l]$ are simply being zero padded to increase the temporal resolution. Note that for $a = N_f$ the temporal resolution is equal to the original sample cadence Δt . Using a standard FFT, the cost of the banded time maximization scales as $aN \ln(aN_t)$, but it may be possible to achieve faster speeds using sparse FFTs that are optimized for almost empty arrays.

Figure 2 compares the standard Fourier domain time maximization to the banded WDM time maximization. The

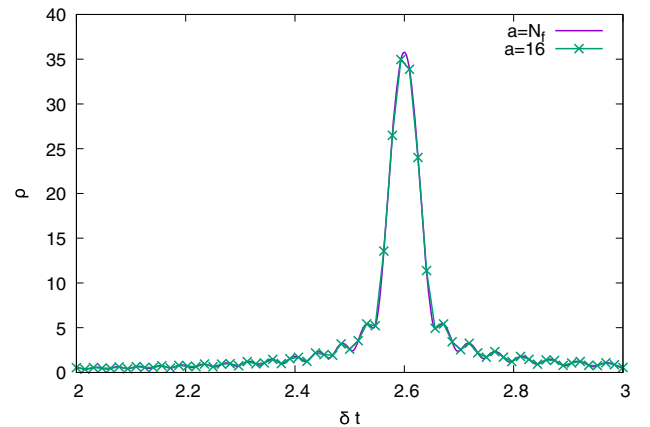


FIG. 2. The signal-to-noise time series $\rho(\delta t)$ for a chirplet waveform with $t_p = 2.6$ s. The curve labeled $a = N_f$ used the standard Fourier domain time maximization with time resolution $\Delta t = 1/1024$ s, while the curve labeled $a = 16$ used the WDM banded maximization with resolution $\Delta T/16 = 1/60$ s.

simulation used a chirplet waveform [defined in Eq. (55) of Sec. V C] with central time $t_p = 2.6$ s, central frequency $f_p = 416$ Hz, duration $\tau = 3.23$ s, and frequency spread $\gamma = 13$ Hz, sampled at cadence $\Delta t = 1/1024$ s. The WDM transform used $N_f = 256$, $N_t = 32$ and $\Delta T = 0.25$ s.

IV. GLITCHES, GAPS AND OTHER GREMLINS

“We don’t need your fancy statistics, we just need to build better detectors”—Rainer Weiss, to the author, at the tenth Harvard-Smithsonian Conference on Theoretical Astrophysics, May 2018.

The noise in an ideal detector would be stationary and Gaussian, with a level set by irreducible quantum uncertainty. In reality, the noise encountered in contemporary gravitational wave detectors is neither stationary nor Gaussian due to frequent short duration noise transients (glitches), in addition to longer duration drifts in the system [8]. The data also has many gaps, and the dynamic noise spectra are littered with dozens of sharp spectral lines that wander around in frequency and vary in amplitude. Performing analyses in the wavelet domain helps reduce the impact of many of these issues, but the main benefit is in dealing with relatively slow changes in the noise properties. Methods such as Bayesian data augmentation, edge wavelets and glitch subtraction are still needed to mitigate the impact of gaps and short duration noise transients. The coherent WaveBurst algorithm [6,7] uses the WDM transform for noise modeling, including for power spectrum estimation and spectral line removal [45].

The Heisenberg-Gabor uncertainty relation [46] (which is saturated for discrete wavelets: $\Delta T \Delta F = 1/2$) implies a trade-off between spectral and temporal resolution in the wavelet transform. The time resolution ΔT should be short compared to the timescale of the nonstationarity of the data, while the frequency resolution ΔF should be small compared to the scale on which the average power spectrum $S(f)$ varies. If the data are highly non-stationary and also contains sharp spectral features it can be difficult to satisfy both criteria simultaneously with a single choice of resolution, and multi-resolution approaches may be required (e.g., first removing spectral lines using a transform with large ΔT , then taking care of the nonstationarity using a transform with a smaller ΔT).

A. Drifts in the noise spectrum

Nonstationary noise comes in myriad forms. Stationary noise, heuristically speaking, is generated by a random process with statistical properties that are unchanging in time. Wide sense stationary (WSS) noise (sometimes called weak sense stationary) requires that the first (mean) and second (covariance) moments of the random process are unchanging in time. For WSS processes the autocorrelation between data at times t_1 and t_2 depends only on the lag $\tau = |t_1 - t_2|$. In practice, noise is never exactly WSS. A

more realistic approximation is the concept of locally stationary noise (LS), where the statistical properties of the noise process vary slowly with time [35,47]. The autocorrelation properties of LS processes can be described in terms of dynamic, or evolutionary, spectra $S(f, t)$ [18]. In contrast to stationary processes, where the frequency domain autocorrelation (power spectrum) is nonzero, nonstationary processes produce nonvanishing cross-correlations between frequencies [37]. The traditional way to estimate $S(f, t)$ is using periodograms of short time Fourier transforms. Scalograms computed from discrete wavelet transforms are a powerful tool for estimating evolutionary spectra.

As discussed earlier, a major advantage of the WDM wavelet basis over a Fourier basis is that it yields a diagonal noise correlation for a range of nonstationary processes. As a concrete example of a nonstationary process with an evolutionary spectra, consider the $AR(N)$ model with time varying variance:

$$\sum_{j=0}^N a_j x[i-j] = (1 + \beta e^{-\frac{25(i-N/2)^2}{2N^2}}) \delta_i, \quad (32)$$

where δ_i is a zero norm, unit variance Gaussian random variate, $\delta_i \sim \mathcal{N}(0, 1)$, and N is the total number of time samples $x[i]$. Setting the coefficients a_j such that

$$a_j = \left(\frac{2j-2-\alpha}{2j} \right) a_{j-1} \quad (33)$$

with $a_0 = 1$ generates colored noise with a $1/f^\alpha$ spectrum [48]. When β is nonzero the process is nonstationary. Figure 3 shows one realization of this $AR(N)$ process with $N = 2^{20}$, $\alpha = 2$ and $\beta = 9$. The nonstationarity is evident in the increased variance of the process for samples near the midpoint. The nonstationarity can also be seen in the Fourier domain correlation coefficients $c_k = C_k/C_0$, where

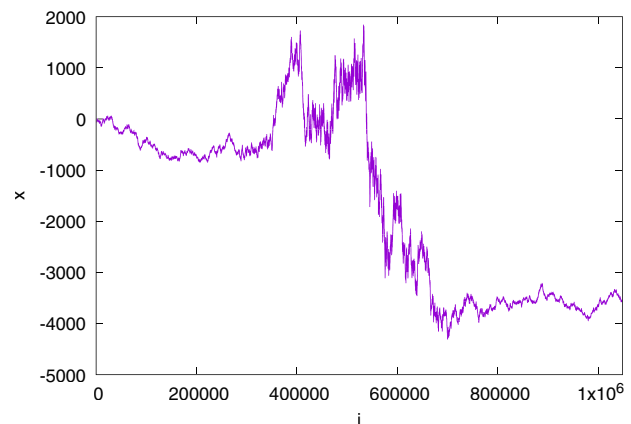


FIG. 3. A realization of the $AR(N)$ process described by Eq. (32) with $N = 2^{20}$, $\alpha = 2$ and $\beta = 9$.

$$C_k = \frac{1}{N-k} \sum_{i=0}^{N-1-k} (\bar{X}[i]\bar{X}^*[i+k] + \bar{X}^*[i]\bar{X}[i+k]), \quad (34)$$

and $\bar{X}[i]$ are the whitened Fourier coefficients. Here these are found by taking a FFT of the time domain data and dividing the Fourier coefficients by the square root of the theoretical $1/f^\alpha$ spectrum. To reduce spectral leakage, a Tukey filter that is flat across 95% of the data was applied to the time domain data before transforming to the frequency domain. Figure 4 shows the first 100 Fourier domain correlation coefficients for stationary ($\beta = 0$) and nonstationary ($\beta = 9$) realizations of the $AR(N)$ process. Strong correlations between Fourier bins are evident for nonstationary case. The correlation coefficients for the stationary case deviate slightly from zero due to the application of the Tukey window. As discussed in more detail later, apodizing windows introduce nonstationarity to otherwise stationary time series.

The evolutionary spectrum $S(t, f)$ can be estimated in many different ways. One method is to introduce a parametric model for the evolutionary spectra, $S_\theta(t, f)$, and solve for the parameters of the model, θ . Under the assumption that the noise is Gaussian distributed, the model parameters can be estimated by maximizing the Whittle log likelihood

$$\ln p(d|\theta) = -\frac{1}{2} \sum_{n,m} \left(\ln(2\pi S_\theta(t_n, f_m)) + \frac{w_{nm}^2}{S_\theta(t_n, f_m)} \right). \quad (35)$$

Alternatively, the Whittle likelihood can be used to infer the model parameters in a Bayesian setting. To avoid overparameterizing the model, transdimensional methods are preferable. For example, the Bayesline spectral estimation algorithm [24] can be generalized to use a bicubic spline model, where the number of spline points and their

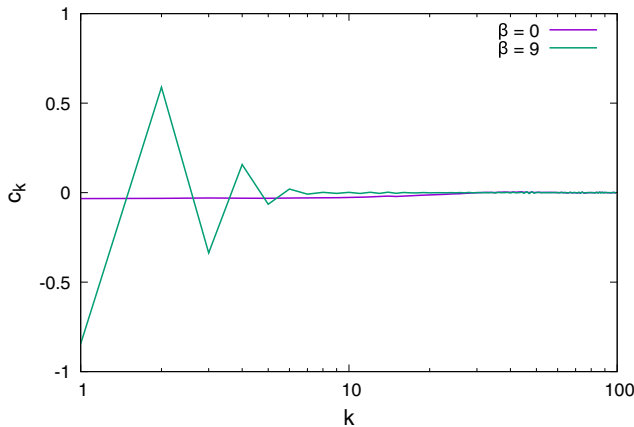


FIG. 4. Fourier domain correlation coefficients c_k for the $AR(N)$ process with $\alpha = 2$ and $\beta = 0$ (stationary case) and $\beta = 9$ (nonstationary). The index k corresponds to the number of frequency bins between the data samples used to compute the correlation.

amplitudes are allowed to vary. A much simpler, though less accurate estimate for the evolutionary spectra can be found using a fixed grid of bicubic spline points and estimating their amplitude from the median of the points in the scalogram, w_{nm}^2 , surrounding the spline control point. The median is more robust to outliers than the mean. The median of the scalogram has to be divided by a factor of $\bar{\chi}_1^2/\bar{\chi}_1^2$ to yield an estimate of $S(t, f)$ under the assumption that the scalogram is chi-squared distributed with one degree of freedom. Here $\bar{\chi}_1^2 = 1$ and $\bar{\chi}_1^2/\bar{\chi}_1^2 = 0.4549364231\dots$ are the mean and median of the chi-squared distribution. The median is found from the transcendental equation $\Gamma[\frac{1}{2}, \frac{1}{2}\bar{\chi}_1^2] = \frac{1}{2}\Gamma[\frac{1}{2}]$, where $\Gamma[a, b]$ and $\Gamma[a]$ are the incomplete and complete Gamma functions, respectively.

Figure 5 shows a bicubic spline fit to the evolutionary spectrum for the $AR(N)$ process (32) with $\alpha = 2$ and $\beta = 9$. The spectrum varies over seven decades in magnitude. The time series was generated with $N = 2^{20}$ points and the WDM transform was computed on a grid with $N_t = N_f = 1024$ points. The bicubic spline control points were spaced at intervals of 32 points in the both time and frequency directions. The whitened wavelet amplitudes, $W_{nm} = w_{nm}/\sqrt{S_\theta(t_n, f_m)}$ follow a $\mathcal{N}(0, 1)$ unit normal distribution (verified using an Anderson-Darling test).

Under the hypothesis that the wavelet noise correlation matrix is diagonal, the products $W_{nm}W_{n'm'}$ for $n \neq n'$ and/or $m \neq m'$ should follow a product normal distribution, $p(x) = K_0(|x|)/\pi$, where $K_0(x)$ is a modified Bessel function of the second kind. A histogram of the products confirms this expectation. A more quantitative test is to look at estimates of the whitened noise correlation matrix

$$c_{ij} = E[W_{nm}W_{n+i, m+j}] \approx \frac{1}{N} \sum_{n,m} W_{nm}W_{n+i, m+j}. \quad (36)$$

In the limit that $N \rightarrow \infty$, the off-diagonal elements of c_{ij} should vanish. To allow for the finite number of samples,

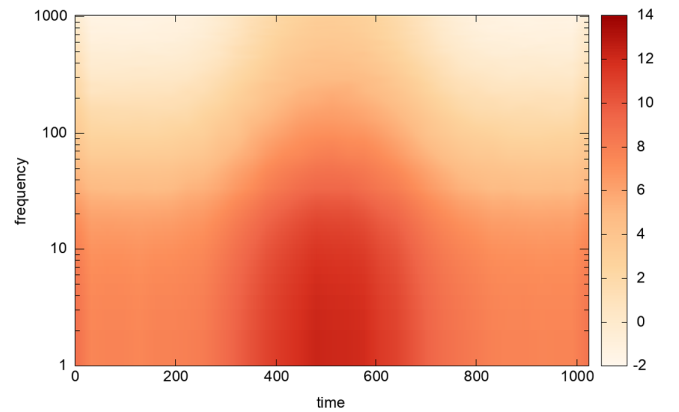


FIG. 5. A bicubic spline fit to logarithm of the evolutionary spectrum, $\ln S(f, t)$, for the nonstationary $AR(N)$ model described by Eq. (32) with $\alpha = 2$ and $\beta = 9$.

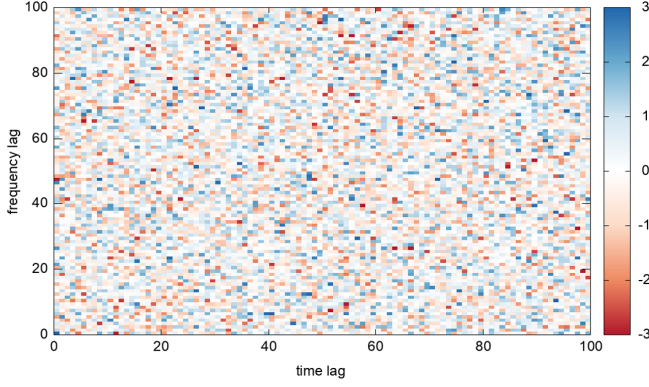


FIG. 6. The scaled off-diagonal noise correlation coefficients \bar{c}_{ij} of the WDM wavelet transform for the nonstationary $AR(N)$ model described by Eq. (32) with $\alpha = 2$ and $\beta = 9$.

the c_{ij} can be scaled by the standard deviation of the sample mean, which is given by $\sigma^2 = \text{Var}[W_{nm} W_{n+i,m+j}]/N$. The scaled correlations, $\bar{c}_{ij} = c_{ij}/\sigma$ should follow a $\mathcal{N}(0, 1)$ unit normal distribution for $i \neq j$. Figure 6 shows the first 10000 \bar{c}_{ij} for the data shown in Fig. 3. The distribution appears to be consistent with unit variance white noise, with no discernible correlation pattern. Figure 7 shows a histogram and quantile-quantile plot of the coefficients, and confirms that they follow a $\mathcal{N}(0, 1)$ unit normal distribution.

Similar tests were applied to a variety of noise processes, including stationary white noise, stationary red noise,

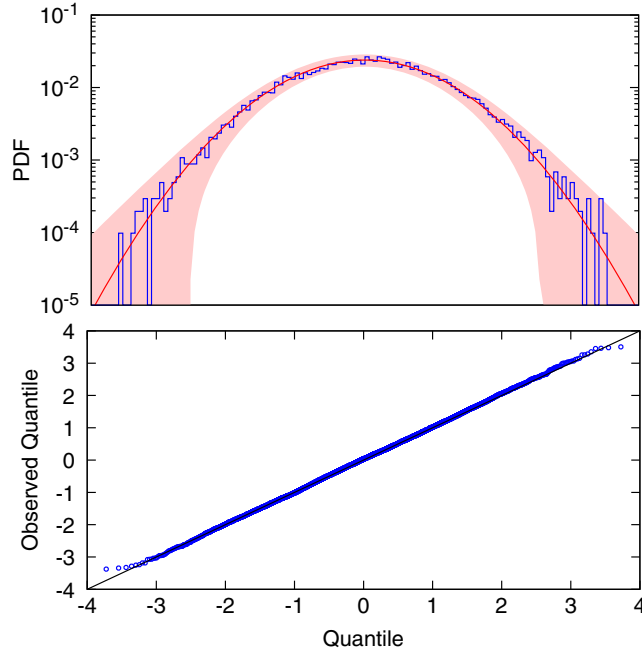


FIG. 7. Histogram (upper panel) and quantile-quantile plot (lower panel) for the scaled off-diagonal noise correlation coefficients \bar{c}_{ij} shown in Fig. 6. The \bar{c}_{ij} follow a $\mathcal{N}(0, 1)$ unit normal distribution, confirming that the WDM noise correlation matrix is diagonal.

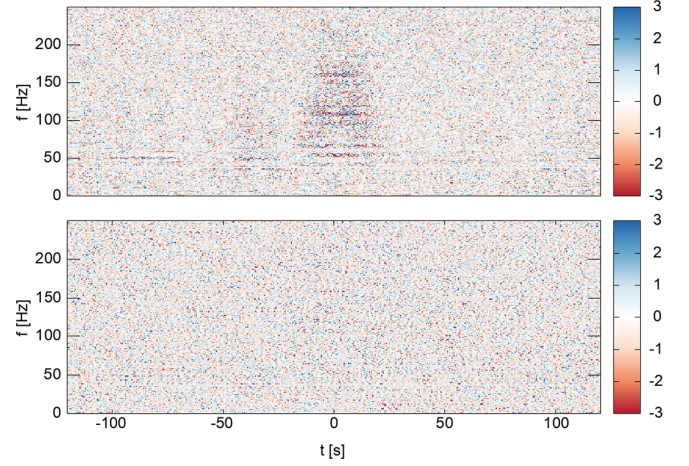


FIG. 8. The upper panel shows the WDM transform of 256 seconds of data from the LIGO Hanford detector centered on GPS time 1165067917. The data has been whitened using the median wavelet PSD $S(f)$. The lower panel shows the same data whitened and made stationary by the evolutionary spectrum PSD $S(f, t)$.

cyclo-stationary noise, nonstationary white noise with variance that increased linearly with time, and nonstationary red noise generated by the $AR(N)$ process with different values of α and β . In all cases the WDM wavelet noise correlation matrix was found to be diagonal.

A more relevant question is whether the WDM transform yields a diagonal noise correlation matrix when applied to real gravitational wave data. The LIGO-Virgo data analysis guide paper [8] investigated a particular stretch of LIGO Hanford data that exhibited nonstationary noise. The WDM scalogram of the same 256 second stretch of data studied in Ref. [8] is shown in Fig. 8, focusing on the frequency band below 250 Hz where the nonstationarity is most pronounced. The WDM transform had a time resolution of $\Delta T = 0.5$ s and a frequency resolution of $\Delta F = 1$ Hz. The upper panel in Fig. 8 shows the whitened data, using an estimate for $S(f)$ based on the median of the scalogram in each frequency band. The lower panel in Fig. 8 shows the same data, now whitened with the evolutionary spectrum $S(f, t)$. Applying the same analysis to the off-diagonal noise correlation coefficients, c_{ij} , as was done previously for the nonstationary $AR(N)$ model yields the distribution shown in Fig. 9. The scaled coefficients \bar{c}_{ij} follow a $\mathcal{N}(0, 1)$ unit normal distribution, confirming that the WDM noise correlation matrix for this stretch of nonstationary LIGO Hanford data is diagonal.

B. Gaps and edge effects

It is well known that time domain data with edges and gaps lead to spectral leakage in the Fourier domain. The spectral leakage can be mitigated using window functions, but at the cost of lost data, and artificial nonstationarity. The latter point is not widely appreciated. Even using a Tukey

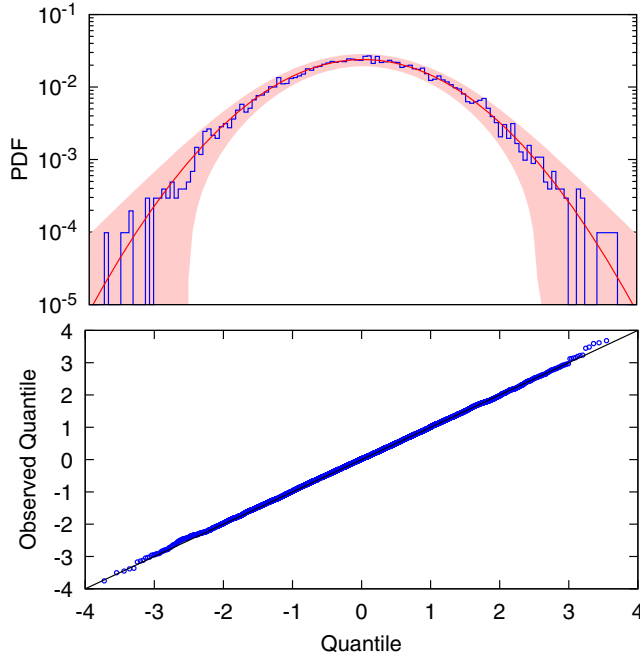


FIG. 9. Histogram (upper panel) and quantile-quantile plot (lower panel) for the scaled off-diagonal noise correlation coefficients \bar{c}_{ij} for the LIGO Hanford data shown in Fig. 8. The \bar{c}_{ij} follow a $\mathcal{N}(0, 1)$ unit normal distribution, confirming that the WDM noise correlation matrix is diagonal.

window that is flat across a large fraction on the data results in a nondiagonal noise correlation matrix in the Fourier domain. Figure 10 shows the Fourier domain noise correlation function c_k for 16 seconds of stationary white noise, $x[i]$, sampled at 4096 Hz, with a Tukey window applied. The Tukey window had a rise and decay time of 0.4 s to mimic the settings in typical LIGO analyses [8]. The window is flat across $1 - \alpha$ of the data, with

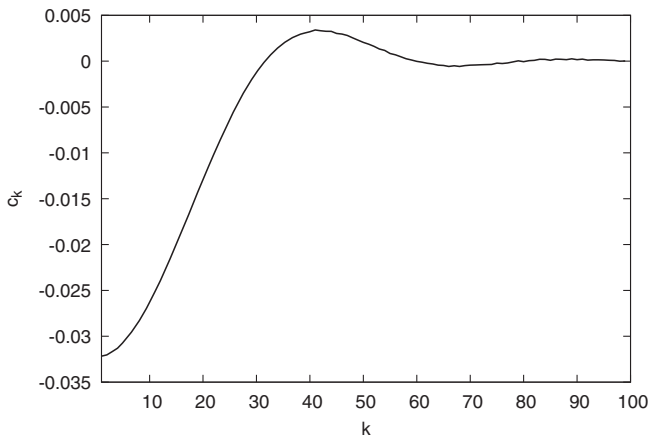


FIG. 10. The off diagonal elements of the Fourier domain noise correlation function c_k for stationary white noise with a Tukey window function applied. The index k corresponds to the number of frequency bins between the data samples used to compute the correlation.

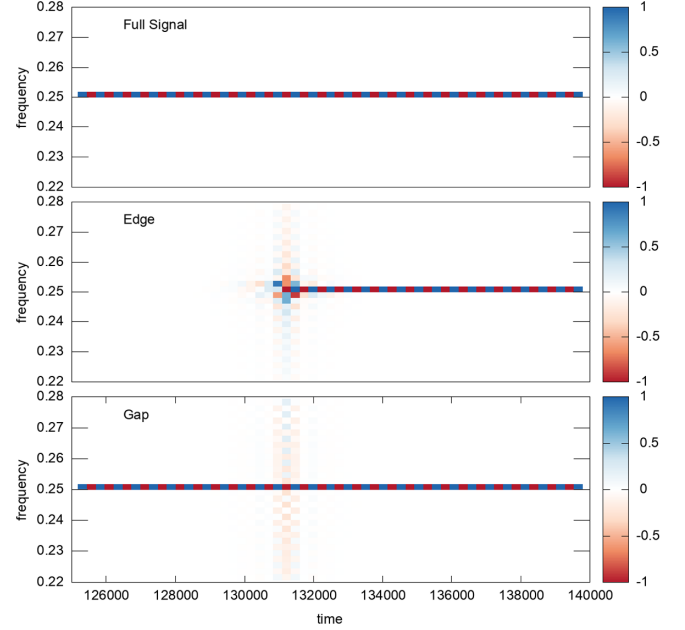


FIG. 11. The WDM transform of a sine wave (upper panel), a sine wave with an edge at time 131072 (middle panel), and a sine wave with a gap of width $\Delta T = 1/32$ (lower panel). The amplitude of the sine wave is 8, but the color scale has been capped between $[-1, 1]$ to better show the leakage of power caused by the edge and gap.

$\alpha = 0.8/16 = 0.05$. The magnitude of the maximum correlation is of order α .

Wavelet transforms are not immune to the edge effects and gaps, though the impact is more localized. Edges and gaps introduce spurious sharp features in the data that cause Gibbs oscillation and spectral leakage, as also happens with a Fourier transform. Another way of looking at it is that the edges and gaps render the wavelet basis functions non-orthogonal, resulting in leakage of power between wavelet pixels. The leakage is confined to wavelets whose time domain filters overlap with the edge or gap. Recall that the time domain window functions are always wider than the wavelet pixels time span ΔT , typically by factors of order ten. Figure 11 illustrates the effects of edges and short duration gaps (less than the width of wavelet pixel) on the WDM transform of a sine wave. The time domain window function for the WDM transform had width $q\Delta T$ with $q = 16$. The power leakage due to the signal having a sharp edge extends for order $16\Delta T$ in time and $32\Delta F$ in frequency. The short data gap leads to a similar pattern of leakage, with a slightly smaller extent in time and slightly larger extent in frequency.

Methods for mitigating the effects of edges and gaps that are used in frequency domain analyses, such as apodizing the edges with a window function, and filling the gaps with fake data using Bayesian data augmentation [49] or otherwise [31], can also be used with wavelet transforms. An alternative is to introduce specialized boundary

wavelets and gap wavelets that maintain orthogonality and eliminate leakage [50–52], but at the cost of added computational complexity, and a decrease in frequency separation in the regions covered by these special wavelets.

C. Glitches and spectral lines

The data from ground based interferometers exhibit sharp spectral features and suffer from frequent noise transients, or glitches [8]. Figure 12 compares Fourier domain and wavelet domain estimates of the power spectral density of the 256 seconds of LIGO Hanford data used to produce Fig. 8. The Fourier domain estimate used a running median of the Fourier periodogram combined with an outlier identification to catch the spectral lines. The wavelet domain estimate used the median of the wavelet scalogram across each frequency band. The spectrum contains numerous spectral lines. The weaker lines are absent in the wavelet estimate as the contribution from the weaker lines is a small contribution to the total integrated power across a wavelet. The wavelet estimate is more robust against the nonstationary excess below 250 Hz, which is why the wavelet estimate sits below the Fourier domain estimate at low frequencies.

Spectral lines can be both very narrow in frequency and very large in amplitude. Strong lines can saturate an entire wavelet frequency band, rendering the data useless in that band. The fraction of otherwise useable data lost to strong lines scales as the ratio of the width of the line δf to the bandwidth of the wavelet layer, ΔF . This suggests choosing a small ΔF to avoid losing information. However, small ΔF implies large ΔT , and transforms with large ΔT are less effective in dealing with nonstationarity. A way around this problem is to first transform the data using a small value for

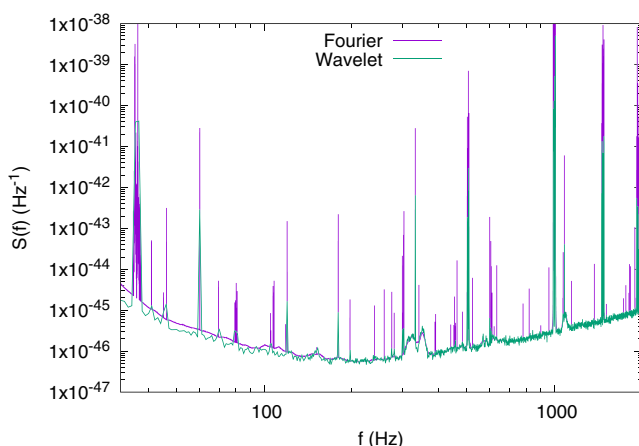


FIG. 12. Estimates of the power spectral density of the LIGO Hanford data shown in Fig. 8 using Fourier domain and wavelet domain methods. The spectrum exhibits many sharp spectral lines. Some of the weaker lines are smoothed away in the wavelet domain analysis.

ΔF , and use this transform to estimate the power spectrum $S(f)$, ignoring any nonstationarity for now. The data can then be whitened using this estimate, returned to the time or frequency domain, then wavelet transforming a second time using a larger ΔF . The evolutionary power spectrum of the retransformed data, $\bar{S}(t, f)$, can then be computed. The full dynamic power spectrum is given by $S(f, t) = S(f)\bar{S}(t, f)$.

In addition to slow drifts in the evolutionary power spectrum, such as those seen in the LIGO Hanford data in Fig. 8, interferometer data also manifest frequent short duration noise transients or glitches. In contrast to long term drifts, which can be modeled as Gaussian noise with a time varying power spectrum, glitches are short duration, coherently structured non-Gaussian features. A better term for them might be “instrument generated signals.” Figure 13 shows the WDM transform of 16 seconds of LIGO Livingston data surrounding the binary neutron star merger GW170817 [53]. A very loud glitch can be seen roughly one second prior to merger. Whitening the data with the evolutionary wavelet spectrum helps mask the effect of the glitch, but a much better approach is to coherently model and remove the glitch from the data, as was done in the original analysis of GW170817 [53], where the BayesWave algorithm [54] was used to subtract the glitch from the data. Treating glitches as instrument signals and removing them from the data is superior to gating the affected data or whitening the data with the evolutionary spectrum, as the later approaches lead to a loss of information, while the former approach preserves all of the data [55].

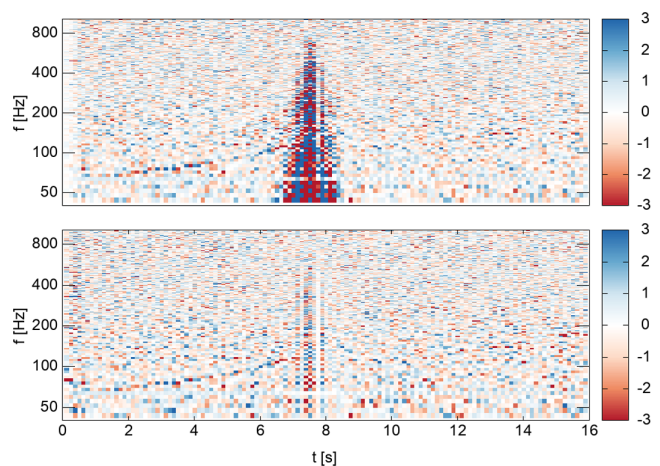


FIG. 13. The upper panel shows the WDM transform of 16 seconds of data from the LIGO Livingston detector surrounding the BNS merger GW170817. The data has been whitened using the median wavelet PSD $S(f)$. The lower panel shows the same data whitened and made stationary by the evolutionary spectrum $S(f, t)$. The loud noise transient is significantly reduced by the evolutionary PSD. The inspiral track of the GW signal is visible in lower left of both panels.

V. WAVELET DOMAIN WAVEFORMS

Gravitational wave data analysis often requires millions of likelihood calculations, each involving one or more waveform evaluations. While the computational cost of the WDM wavelet transform is comparable to a standard FFT based frequency domain analysis, the cost can be prohibitive for long-lived signals. To reduce the computational cost of standard frequency domain analyses, significant effort has gone into producing waveform models that can be evaluated directly in the frequency domain [9–17]. The goal here is to develop fast wavelet domain waveforms using time domain or frequency domain waveforms as a starting point. Remarkably, it turns out that waveforms can be generated more efficiently in the wavelet domain: the computational cost scales as the square root of the number of data points, as opposed to linearly, or worse, for the time and frequency domains.

Let us assume that the waveform templates for binary systems can be expressed in the form

$$h(t) = \Re \sum_k A_k(t) e^{i\Psi_k(t)}, \quad (37)$$

or equivalently

$$\tilde{h}(f) = \sum_k \mathcal{A}_k(f) e^{i\Theta_k(f)}. \quad (38)$$

Here the amplitude and phase are taken to include contributions from the instrument response in addition to contributions from the evolution of the binary system. The series does not have to be harmonic, though that will often be the case [56]. The fast wavelet transform is applied term by term in the sums. In some cases only a few terms are needed to accurately model the signal, such as for comparable mass spin-aligned binaries on quasicircular orbits. In other cases a large number of terms will be needed, such as for system on highly eccentric orbits with large mass ratios. The efficiency of the rapid wavelet transforms degrades as the number of terms in the sum increases, eventually reaching the point where there are no savings to be gained over directly transforming the full signal. Similar considerations apply when computing frequency domain waveforms for eccentric binaries [17].

Figure 14 illustrates the time-frequency evolution of one term in the sums (37), (38). During the early inspiral (horizontal shaded region) the signal evolves slowly in time, and fast time-domain transforms are most efficient. During the late inspiral, merger and ringdown (vertical shaded region) the signal evolves rapidly in frequency, and fast-frequency domain transforms are most efficient. In each region there are (at least) two fast methods for computing the wavelet transform. The first method, sparse sampling, involves no approximations and is typically one or two orders of magnitude faster than the direct transform.

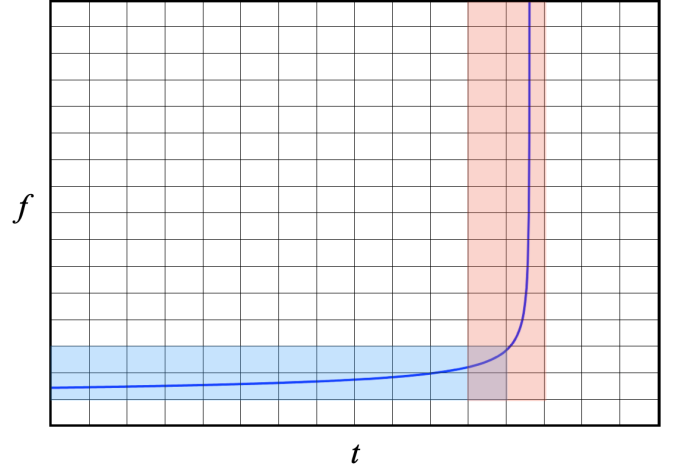


FIG. 14. An illustration of the time-frequency decomposition of a single term in the expansion of the signal from a binary merger. The horizontal (light blue) shaded region indicates where fast time-domain transforms are most efficient, while the vertical (light red) shaded region indicates where fast frequency-domain transforms are most efficient.

The second method employs a local Taylor series expansion of the amplitude and phase, and interpolation over a look-up table. The second method is typically two to three orders of magnitude faster than the direct transform. The computational savings are greatest for long-lived signals, such as binary Neutron star mergers for ground based detectors and systems with masses below $10^5 M_\odot$ for space based detectors.

A. Fast wavelet waveforms: Sparse sampling

The sparse sampling technique employs a stationary-phase-inspired time-frequency mapping defined:

$$\begin{aligned} t_k(f) &= \frac{1}{2\pi} \frac{d\Theta_k(f)}{df}, \\ f_k(t) &= \frac{1}{2\pi} \frac{d\Psi_k(t)}{dt}, \end{aligned} \quad (39)$$

and similarly for the time and frequency derivatives:

$$\begin{aligned} t'_k(f) &= \frac{1}{2\pi} \frac{d^2\Theta_k(f)}{d^2f}, \\ \dot{f}_k(t) &= \frac{1}{2\pi} \frac{d^2\Psi_k(t)}{d^2t}. \end{aligned} \quad (40)$$

The sparse sampling approach can be understood as a form of heterodyning, or equivalently, a stationary phase evaluation. At time $t_k = n\Delta T$ the signal has a frequency $f_k \approx m\Delta F$ for some specific m . Alternatively, the signal has frequency $f_k = m\Delta F$ at some time $t_k \approx n\Delta T$ for some specific n . At these instances, the oscillatory factors in Eqs. (14) and (17) largely cancel the oscillations due to the

phase terms $e^{i\Psi_k(t)}$ or $e^{i\Theta_k(f)}$. Since the summands are then slowly varying, the sums can be evaluated using a much sparser sampling in time or frequency, thereby significantly reducing the computational cost. Time-domain sparse sampling is most efficient in regions where the signal is slowly evolving in frequency: $\dot{f}_k T_w < \Delta F$, where T_w is the duration of the wavelet window function. Frequency-domain sampling is most efficient in regions where the signal is rapidly evolving in frequency: $t'_k F_w < \Delta T$, where F_w is the bandwidth of the wavelet window function.

Sparse sampling in the time domain works by reducing the sample cadence such that the time samples are spaced by $p\Delta t$, where p is some power of two. The larger the p the greater the savings in computational cost. The downsampling is limited by how rapidly the amplitude and heterodyned phase vary. Significant savings ($p = 256$ or larger) are only possible when the amplitude and heterodyned phase are evolving relatively slowly in time. Considering a single harmonic of the waveform, and dropping the subscripts to simplify the notation, the wavelet coefficients are given by

$$w_{nm} = \sqrt{2\Delta t} \Re C_{nm} S_n[(m - m_c)q], \quad (41)$$

where $m_c = [f(n\Delta T)/\Delta F]$ is the central frequency band and

$$S_n[l] = \sum_{j=-L/2}^{L/2-1} e^{2\pi i j l / L} A[nN_f/p] \phi[j] e^{i(\Psi[nN_f/p] - 2\pi m_c q j / L)}, \quad (42)$$

with $L = K/p$. The S_{nl} can be computed using a FFT at cost $L \ln L$ for each n , yielding a total cost that scales as $N_f L \ln(L)$. This is more than factor of p faster than the full time domain transform. The sparse transform is exact so long as the summand in (42) is adequately sampled. The downsampling factor p can be dynamically adjusted according to how fast the signal is evolving.

Sparse sampling in the frequency domain works by decreasing the time span $T \rightarrow T/r$ such that the frequency samples are spaced by r/T , where r is some power of two. The time span must be equal to or greater than the window size $T_w = K\Delta t$, so that $r \leq N/K$. Since the frequency domain method is most useful when the signals are rapidly evolving, we saturate the bound and set $r = N_t/(2q)$. The wavelet coefficients are then given by

$$w_{nm} = \sqrt{2}(-1)^{nm} \Re C_{nm} s_m[n - n_c], \quad (43)$$

where $n_c = [t(m\Delta F)/\Delta T]$ is the central time slice and

$$s_m[j] = \sum_{l=-q}^{q-1} e^{-\pi i l j / q} \mathcal{A}[l + mq] \Phi[l] e^{i(\Theta[l + mq] + \pi l n_c / q)}. \quad (44)$$

Evaluating (44) for each m using a FFT yields a total cost of $2N_f q \ln(2q)$, which is a factor of $r = N_t/(2q)$ less than the full frequency domain transform. The limitation of using a time span equal to the window size is that the time is periodic with period T_w , meaning that the transformation becomes multivalued for slowly evolving signals. For (44) to be single valued we require that the signal evolves to occupy a new frequency layer in the time it takes to cross the window function. That is, we require $\dot{f} T_w > \Delta F$. When this condition is met the sparse frequency domain transform is exact.

B. Fast wavelet waveforms: Taylor expansion

The second fast transform takes advantage of the fact that the individual harmonics of the gravitational wave signal from a binary system go from evolving slowly in frequency to evolving slowly in time (see Fig. 14). During the early inspiral the signals evolve slowly in time, and the amplitude and frequency are approximately constant across an individual wavelet. Consider a wavelet centered at $t_n = n\Delta T$. The gravitational wave amplitude and phase can be Taylor expanded:

$$\begin{aligned} \Psi(t) &= \Psi(t_n) + 2\pi(t - t_n)f(t_n) + \pi(t - t_n)^2 \dot{f}(t_n) + \dots, \\ A(t) &= A(t_n) + (t - t_n)\dot{A}(t_n) + \dots, \end{aligned} \quad (45)$$

where the expressions for $f(t)$ and $\dot{f}(t)$ are given in Eqs. (39) and (40). To achieve high accuracy these expansions can be continued to higher order, but it is usually enough to stop at second order in the phase and zeroth order in the amplitude. The wavelet coefficients

$$w_{nm} = \sum_{k=1}^N g_{nm}[k] A(t) e^{i\Psi(t)} \quad (46)$$

are then given by

$$\begin{aligned} w_{nm} &= A(t_n)(c_{nm}(f(t_n), \dot{f}(t_n)) \cos \Psi(t_n) \\ &\quad - s_{nm}(f(t_n), \dot{f}(t_n)) \sin \Psi(t_n)), \end{aligned} \quad (47)$$

where

$$c_{nm}(f, \dot{f}) = \int dt g_{nm}(t) \cos(2\pi(t - t_n)f + \pi(t - t_n)^2 \dot{f}) \quad (48)$$

and similarly for $s_{nm}(f, \dot{f})$ with the cosine replaced by a sine. The c_{nm} and s_{nm} can be precomputed on a grid in (f, \dot{f}) and interpolated. This procedure can be used for any wavelet family. The WDM wavelets offer an additional advantage in that they have a uniform window in frequency, so that the coefficients, computed relative to the central

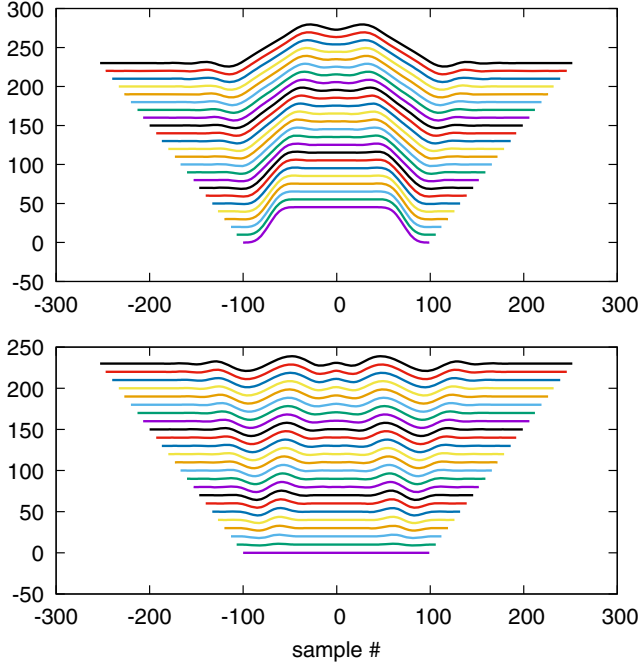


FIG. 15. An illustration of the lookup table for the cosine (upper) and sine (lower) coefficients used to compute the time domain expansion coefficients $c_{nm}(f, \dot{f})$ and $s_{nm}(f, \dot{f})$. The horizontal axis shows the successive frequency samples, while the vertically offset lines show increments of the frequency derivative.

frequency of one layer, are the same for all layers. This dramatically shrinks the size of the lookup table since it only has to be computed for a single frequency layer. The lookup table is generated on a grid spaced by $\delta f = \epsilon \Delta F$ with $\epsilon \approx 0.01$ and $\delta \dot{f} = \epsilon \Delta F / T_w$ with $\epsilon \approx 0.1$. The number of frequency derivative samples depends on the maximum frequency we wish to cover—typically $\dot{f}_{\max} = 8\Delta F / T_w$ is sufficient. The number of frequency samples required depends on the filter bandwidth $(A+B)/\pi$ and the frequency evolution of the signal, which increases with increasing \dot{f} . The total number of frequency samples scales as $((A+B)/\pi + \dot{f}_{\max} T_w) / \delta f$. The full lookup table typically has order 10^5 entries. For very slowly evolving signals, such as LISA galactic binaries, the number is far less—of order 100. Note that since $((A+B)/\pi + \dot{f} T_w) > \Delta F$, the wavelet coefficients at each time sample n often span several frequency layers m . Figure 15 illustrates the time domain expansion coefficients $c_{nm}(f, \dot{f})$, $s_{nm}(f, \dot{f})$ as a function of the instantaneous frequency (across the horizontal) and frequency derivative (vertically offset lines) of the signal. The WDM wavelets used to generate this figure used the same settings as those used in Fig. 1, namely $d = 4$, $A = \Delta\Omega/4$, $B = \Delta\Omega/2$ and $q = 16$.

As a binary system approaches merger, the frequency evolution accelerates rapidly and the time domain approach becomes inefficient as it has to cover a larger and larger

range of frequency derivatives. Rather than expanding the phase in df/dt , it becomes more efficient to switch to the frequency domain and expand the phase in dt/df :

$$\begin{aligned}\Theta(f) &= \Theta(f_m) + 2\pi(f - f_m)t(f_m) \\ &\quad + \pi(f - f_m)^2 t'(f_m) + \dots, \\ \mathcal{A}(f) &= \mathcal{A}(f_m) + (f - f_m)\mathcal{A}'(f_m) + \dots,\end{aligned}\quad (49)$$

where $f_m = m\Delta F$, and the expressions for $t(f)$ and $t'(f)$ are given in Eqs. (39) and (40). To leading order in the amplitude and second order in the phase expansion, the wavelet expansion coefficients are given by

$$\begin{aligned}w_{nm} &= \mathcal{A}(f_m)(c_{nm}(t(f_m), t'(f_m)) \cos \Theta(f_m) \\ &\quad - s_{nm}(t(f_m), t'(f_m)) \sin \Theta(f_m)),\end{aligned}\quad (50)$$

where

$$c_{nm}(t, t') = \int df \tilde{g}_{nm}(f) \cos(2\pi(f - f_m)t + \pi(f - f_m)^2 t') \quad (51)$$

and similarly for $s_{nm}(t, t')$ with cosine replaced by sine. Since the WDM wavelets are uniform in time and frequency, the coefficients need only be computed for a single reference frequency band and time slice. The lookup table is generated on a grid spaced by $\delta t = T_w/M$ with $M \approx 400$, and $\delta t' = \epsilon/(\Delta F^2)$ with $\epsilon \approx 0.1$. Figure 16 illustrates the frequency domain expansion coefficients $c_{nm}(t, t')$, $s_{nm}(t, t')$ as a function of the instantaneous time (across the horizontal) and time derivative (vertically offset lines) of the signal. The number of time derivatives in the table is chosen so as to cover a maximum time derivative t'_{\max} that overlaps with the time domain expansion. Writing $t'_{\max} = a/\Delta F^2$, $\dot{f}_{\max} = b\Delta F/T_w$ and demanding that $\dot{f}_{\max} \geq 1/t'_{\max}$ to ensure overlap of the two expansions, leads to the requirement $ab \geq \Delta F T_w = q$. The lookup tables used here have $q = 16$, $a = 8$ and $b = 8$. The number of time samples required depends on the filter duration, T_w , and time evolution of the signal which increases with increasing t' . The total number of time samples scales as $(T_w + t'_{\max} \Delta F) / \delta t$. Since $(T_w + t' \Delta F) > \Delta T$, the wavelet coefficients for frequency band m span at least $2q$ time slices.

C. Waveform test case

To test the fast wavelet domain transforms it is helpful to work with a simple, analytically tractable waveform model. A good example is a “chirplet”—Gaussian enveloped sinusoid with linearly evolving frequency [57,58]. In the frequency domain, a chirplet with central time t_p , central frequency f_p , time extent τ and frequency extent γ have frequency domain phase and amplitude

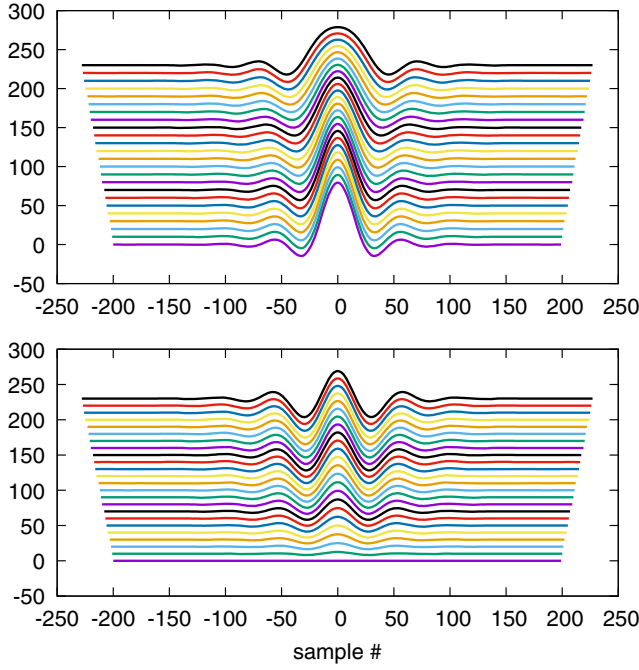


FIG. 16. An illustration of the lookup table for the cosine (upper) and sine (lower) coefficients used to compute the frequency domain expansion coefficients $c_{nm}(t, t')$ and $s_{nm}(t, t')$. The horizontal axis shows the successive time samples, while the vertically offset lines show increments of the time derivative.

$$\begin{aligned}\Theta(f) &= \Theta_0 + 2\pi f t_p + \pi \frac{\tau}{\gamma} (f - f_p)^2, \\ \mathcal{A}(f) &= \mathcal{A}_0 e^{-(f-f_p)^2/(2\gamma^2)}.\end{aligned}\quad (52)$$

The time-frequency mapping follows from Eq. (39):

$$\frac{t - t_p}{\tau} = \frac{f - f_p}{\gamma}. \quad (53)$$

The time domain signal can be found by using the inverse stationary phase approximation:

$$h(t) = \mathcal{A}(f(t)) \sqrt{\frac{2\pi}{|\Theta''(f(t))|}} e^{i(2\pi f(t)t + \Theta(f(t)) + \pi/4)}. \quad (54)$$

The time domain amplitude and phase evaluate to

$$\begin{aligned}\Psi(t) &= 2\pi f_p(t - t_p) + \pi \frac{\gamma}{\tau} (t - t_p)^2 + \Theta_0 + \pi/4, \\ A(t) &= \sqrt{\frac{\gamma}{\tau}} \mathcal{A}_0 e^{-(t-t_p)^2/(2\tau^2)}.\end{aligned}\quad (55)$$

The expression is exact up to fractional corrections of order $|\mathcal{A}''(f)/(\mathcal{A}(f)\Theta''(f))| \simeq 1/(4\pi\gamma\tau)$. The quantity $\gamma\tau$ is proportional to the time-frequency volume of the chirplet,

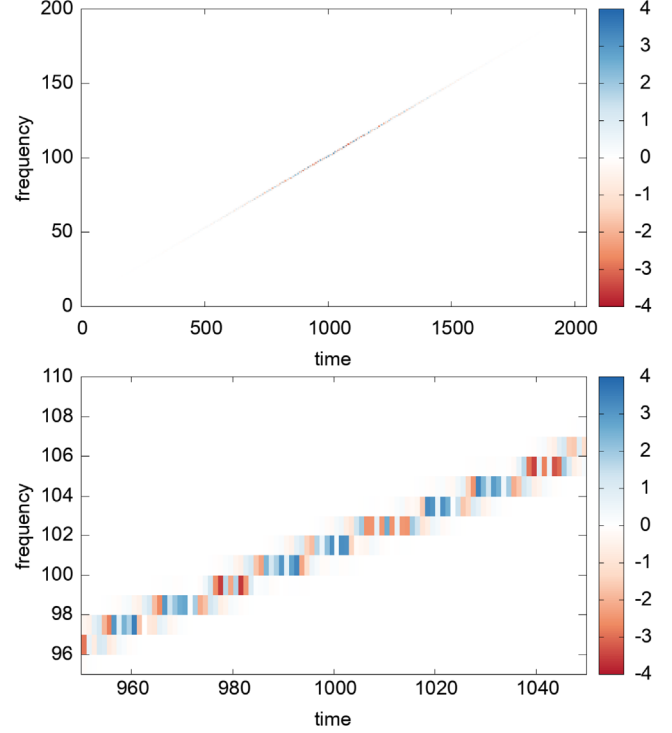


FIG. 17. The WDM transform of a chirplet. Time is in units of ΔT and frequency is in units of ΔF . The upper panel shows the full signal, while the lower panel enlarges a smaller segment. The color bar indicates the wavelet amplitudes.

indicating that the stationary phase approximation improves as the time-frequency volume increases.

Figure 17 shows the WDM transform of a chirplet with parameter $t_p = 1024\Delta T$, $f_p = 103.4\Delta F$, $\mathcal{A}_0 = 10000$, $\Theta_0 = 0.0$, $\dot{f} = \gamma/\tau = 3.1\Delta F/T_w$ and $\gamma = f_p/4$. The wavelet transform has $N_t = 2048$ time slices and $N_f = 2048$ frequency bands. Only the frequency bands occupied by the signal are shown in Fig. 17. The unweighted match between two wavelet transform w_{nm} , u_{nm} can be defined as

$$M = \frac{(w|u)}{\sqrt{(w|w)(u|u)}}, \quad (56)$$

where

$$(w|u) = \sum_{n,m} w_{nm} u_{nm}. \quad (57)$$

The mismatch $MM = 1 - M$ between the direct WDM transform and the various fast approximations was $MM = 6 \times 10^{-10}$ for the sparse time domain transform, $MM = 4 \times 10^{-6}$ for the sparse frequency domain transform, $MM = 3 \times 10^{-6}$ for the Taylor expanded time domain transform and $MM = 3 \times 10^{-5}$ for the Taylor expanded frequency domain transform. In this instance the fast frequency domain transforms are slightly less accurate

than the fast time domain transforms since the data was generated using the time domain expression (55), while the frequency domain transforms use the frequency domain expression (52). These differ slightly since only the lowest order stationary phase approximation was used to map between the time and frequency domain expressions. If necessary, the accuracy of the Taylor expanded WDM transforms can be improved by including the leading order amplitude evolution. Using a single 2.9 GHz CPU core, the direct time domain transform took 5.6 s, and the direct frequency domain transform took 0.9 s. These numbers do not include the 4.1 s it took to generate the full time domain waveform. The sparse time domain transform took 15 ms, the sparse frequency domain transform took 2 ms, the Taylor expanded time domain transform took 0.7 ms and the Taylor expanded frequency domain transform took 0.9 ms.

D. Fast inspiral-merger-ringdown transform

The time domain and frequency domain fast wavelet transform techniques can be combined together to yield a fast wavelet transform for inspiral-merger-ringdown models. As an illustrative example, consider the frequency domain “PhenomD” phenomenological model [14], which describes the dominant $\ell = |m| = 2$ mode of a quasicircular, spin-aligned binary system using a combination of the TaylorF2 post-Newtonian model for the inspiral and a black hole perturbation theory model for the ringdown, bridged by a fit to a suite of numerical relativity waveforms describing the final inspiral and merger. The PhenomD model returns expression for the frequency domain amplitude and phase, $\mathcal{A}(f)$, $\Theta(f)$. A numerical implementation of the stationary phase approximation (SPA) can be used to find the time domain amplitude and phase $A(t)$, $\Psi(t)$. The SPA breaks down at merger, but this is not a limitation since the time domain expressions are only used at early times where the signal is slowly evolving. To generate a reference time domain signal to be used with the direct wavelet transform the SPA waveform is spliced together with the inverse FFT of the frequency domain signal, which has to be smoothly tapered to zero at low frequencies to avoid Gibbs oscillations in the time domain. The splicing procedure is illustrated in Fig. 18. Note that the SPA and the FFT match perfectly in amplitude and phase at the transition point.

A fast wavelet transform of the PhenomD waveform can be constructed by combining the time domain Taylor expansion for early times and the sparse frequency domain transform for late times. The transition time is chosen such that the times t_m , t_{m+1} corresponding to successive frequency bands m , $m+1$ are separated by no more than half-width of the sparse transform: $t_{m+1} - t_m < q\Delta T$. This condition ensures that the sparse frequency domain transform covers the time slices occupied by the signal. The binary black hole systems shown in Fig. 18 had masses $m_1 = 35 M_\odot$, $m_2 = 30 M_\odot$, dimensionless spins $\chi_1 = 0.3$,

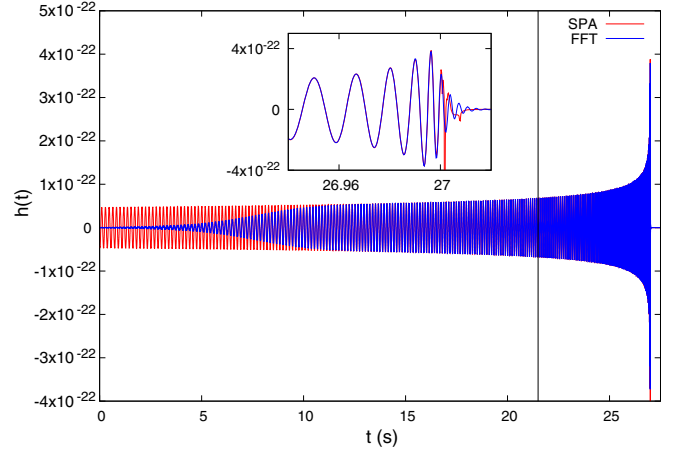


FIG. 18. Construction of the reference time-domain signal for a binary black hole merger. At early times the signal is computed using the stationary phase approximation, while at late times the inverse Fourier transform is used. The transition between the two regimes is marked by a vertical black line. The inset shows the breakdown of the SPA at merger.

$\chi_2 = 0.1$ and merger time $t_c = 27$ s. For this system the transition from the time domain transform to the frequency domain transform occurred at $t_* = 21.5$ s. The transform was computed using $N_f = 1024$, $N_t = 128$, $\Delta t = 1/4096$ and covered 32 s. The wavelet transform of the signal is shown in Fig. 19. The mismatch between the direct wavelet transform of the signal and the fast wavelet transform was $MM = 1.05 \times 10^{-4}$. The match could be improved by including time evolution of the amplitude in the time domain Taylor expansion. The fast wavelet transform took 1.3 ms. This is a significant saving compared to the native frequency domain waveform, which took 8.5 ms to generate. The wavelet transform covered $N_p = 9339 \approx 26\sqrt{N}$ pixels.

The computational savings increase for lower mass systems. For example, consider a system with neutron star

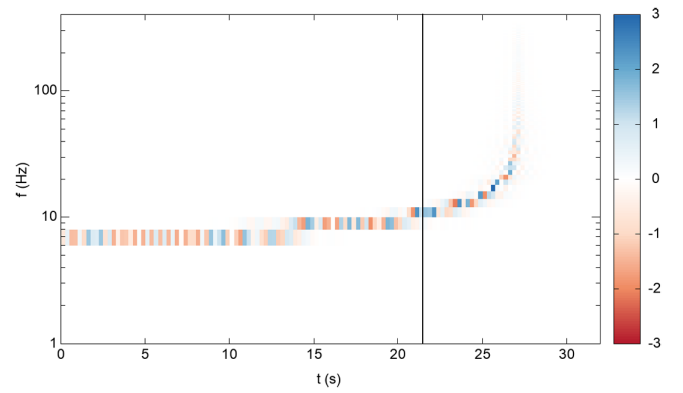


FIG. 19. The WDM wavelet transform of the binary black hole signal shown in Fig. 18. The vertical black line marks the transition from the time domain Taylor transform to the sparse frequency domain transform.

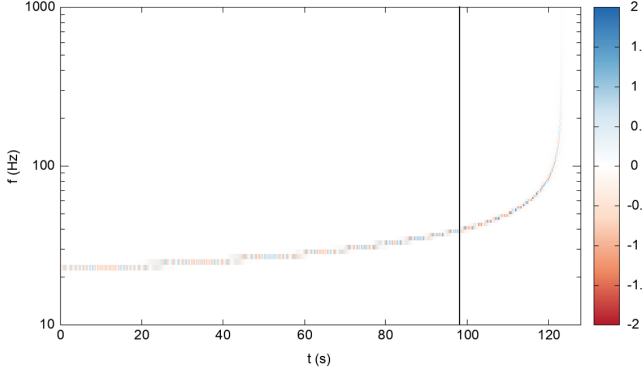


FIG. 20. The WDM wavelet transform of a binary system with neutron star masses and spins. The vertical black line marks the transition from the time domain Taylor transform to the sparse frequency domain transform.

masses and spins: $m_1 = 1.6 M_\odot$, $m_2 = 1.4 M_\odot$, $\chi_1 = 0.02$, $\chi_2 = 0.03$ and merger time $t_c = 123$ s. The WDM transform of this signal is shown in Fig. 20 for $N_f = 1024$, $N_t = 512$, $\Delta t = 1/4096$ covering 128 s. In this case the transition between the time domain Taylor expansion and the sparse frequency domain transform occurred at $t_* = 98.25$ s. The mismatch between the direct wavelet transform of the signal and the fast wavelet transform was $\text{MM} = 3.6 \times 10^{-6}$. The fast wavelet transform took 4.0 ms, which is significantly less than the 45 ms for the native frequency domain signal. The wavelet transform covered $N_p = 32843 \approx 45\sqrt{N}$ pixels.

E. Fast LISA ultra compact binary transforms

Ultra compact galactic binaries (UCBs) are expected to be the most common source detected by future space based gravitational wave detectors operating in the mHz regime [59]. Most UCBs are expected to be on slowly evolving, quasicircular orbits, producing signals that can be accurately modeled using leading order post-Newtonian waveforms. For detectors such as LISA, which orbit the Sun and have time-dependent response functions, the UCB signals pick up amplitude and frequency modulations that vary on the orbital timescale. Wavelet transforms in the mHz regime will typically have pixels that are several hours in extent and hundredths of a mHz in bandwidth. The modulated UCB signals are effectively constant across the wavelet pixels, and can be accurately modeled using the leading order fast time domain Taylor transform. The A , E , T time delay interferometry channels can each be expressed in the form $h_i(t) = A_i(t) \cos \Psi_i(t)$. In the low frequency limit the expressions have the form [60]

$$\begin{aligned} A_I(t) &= [A_+^2(t)F_{I+}^2(t) + A_\times^2(t)F_{I\times}^2(t)]^{1/2}, \\ \Psi_I(t) &= \int 2\pi f(t)dt + \Psi_{Ip}(t) + \Psi_{Id}(t), \\ f(t) &= f_0 + \dot{f}_0 t + \frac{1}{2}\ddot{f}_0 t^2, \end{aligned} \quad (58)$$

with amplitudes

$$\begin{aligned} A_+(t) &= \frac{4\mathcal{M}^{5/3}(\pi f(t))^{2/3}}{D_L} \left(\frac{1 + \cos^2 \iota}{2} \right), \\ A_\times(t) &= \frac{4\mathcal{M}^{5/3}(\pi f(t))^{2/3}}{D_L} \cos \iota, \end{aligned} \quad (59)$$

and polarization and Doppler phases

$$\begin{aligned} \Psi_{Ip}(t) &= \text{atan} \left(\frac{-A_{I\times}(t)F_{I\times}}{A_+(t)F_{I+}(t)} \right), \\ \Psi_{Id}(t) &= 2\pi f(t) \text{ AU} \sin \theta \cos(2\pi t/\text{yr} + \phi_0 - \phi). \end{aligned} \quad (60)$$

Here \mathcal{M} is the chirp mass, D_L is the luminosity distance, ι is the orbital inclination relative to the line of sight and θ, ϕ are the sky location of the source in ecliptic coordinates. The time dependent antenna patterns $F_{I+}(t)$, $F_{I\times}(t)$ depend on the sky location and polarization angle of the source. Similar, yet more complicated expressions for the time dependent amplitude and phase can be found for the full response function, taking into account finite armlength effects—see Appendix B of Ref. [61] for details. The results shown here utilize the full LISA instrument response.

To leading order, the intrinsic frequency change for a UCB scales as

$$\dot{f} = 3.255 \times 10^{-16} \left(\frac{f}{3 \text{ mHz}} \right)^{11/3} \left(\frac{\mathcal{M}}{1 M_\odot} \right)^{5/3} \text{ Hz s}^{-1}, \quad (61)$$

while the magnitude of the maximum Doppler shift scales as

$$\dot{f}_d = 5.93 \times 10^{-14} \left(\frac{f}{3 \text{ mHz}} \right) \sin \theta \text{ Hz s}^{-1}. \quad (62)$$

For typical choices of $\Delta T \sim 10^5 \rightarrow 10^6$ s and $\Delta F \sim 1 \rightarrow 10 \mu\text{Hz}$, the fractional frequency change across the window function, $\dot{f}T_w/\Delta F$, is of order 10^{-2} or less. Consequently, the lookup table for UCBs only has to cover zeroth order in frequency change across the filter.

Figure 21 shows the WDM transform of the LISA response to a galactic binary with $f_0 = 3$ mHz, $\dot{f}_0 = 2.584 \times 10^{-16} \text{ Hz s}^{-1}$, sky location $\cos \theta = 0.2$, $\phi = 1.0$, orientation $\cos \iota = -0.3$, $\psi = 0.8$, initial phase $\Psi_0 = 1.2$ and amplitude $A_0 = 1.96 \times 10^{-21}$. These parameters are encoded into the “bar-code” structure of the WDM transform. Each galactic binary will have a unique bar code, and even though many thousands of binaries will occupy a given frequency band of the transform, their unique encodings allow the individual signals to be extracted, much like a television cable box decodes the multiplexed signals carried on a single coaxial cable. The transform used $N_f = N_t = 2048$, and the data had a sample cadence of $\Delta t = 15$ s and a total duration of roughly two years. The wavelet pixels have duration $\Delta T = 30720$ s and bandwidth $\Delta F = 1.627604 \times 10^{-2} \text{ mHz}$. Generating both channels

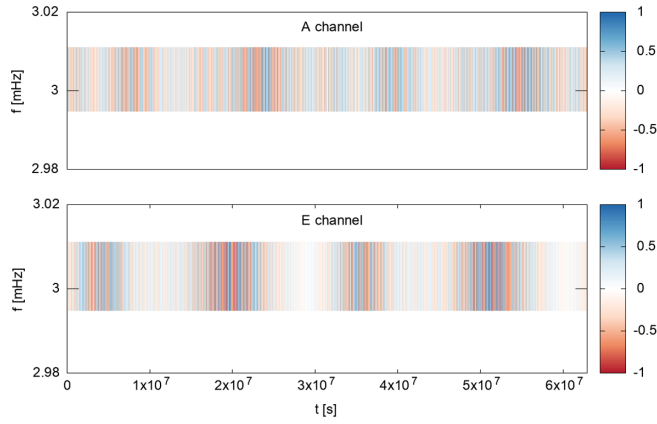


FIG. 21. The WDM wavelet transform of the LISA response to a galactic binary in the TDI A and E channels. The color scale is in units of 10^{-20} .

took a total of 1.3 ms, which is slightly faster than the fast frequency domain transform [62] that is currently used in LISA data analysis studies [63]. The mismatch between the fast and full transforms was $MM = 4.6 \times 10^{-7}$ for the A channel and $MM = 1.1 \times 10^{-6}$ for the E channel.

The UCB transform can be extended to include stellar origin black holes (SOBHs) by using a high order post-Newtonian inspiral model, and by including frequency derivatives in the lookup table. Using the PhenomD waveform model, and considering the same SOBH system as before, but starting three years before merger, yields the LISA signal shown in Fig. 22. To make the signal track more visible, only the final four months of the full two years of simulated LISA data are shown. The fast WDM waveform generation took 4.6 ms. This is a significant saving compared to the 6.8 seconds it takes to compute the waveform in the frequency domain. The mismatch between the fast and full transforms was $MM = 8.1 \times 10^{-6}$ for the A channel and $MM = 1.6 \times 10^{-5}$ for the E channel. Here only

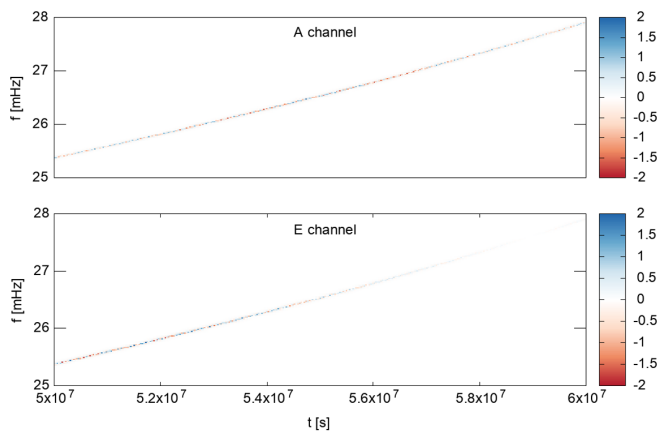


FIG. 22. The WDM wavelet transform of the LISA response to a stellar origin black hole binary in the TDI A and E channels. The color scale is in units of 10^{-22} .

the time domain Taylor expansion was used. To achieve higher accuracy, and to cover systems with larger chirp masses and/or shorter times until merger, it is necessary to switch to the sparse frequency domain response once the frequency evolution is much larger than $\Delta F/T_w$, as was done for the BH mergers described in § V D.

VI. SUMMARY

A time-frequency approach to gravitational wave data analysis using discrete wavelet transforms offers many advantages over the traditional time or frequency domain approaches. Wavelet based analyses are well suited for modeling nonstationary instrument noise. Wavelet domain models of binary merger signals are often significantly faster to compute than time or frequency domain models. Together these advantages make a strong case for moving gravitational wave data analysis to the wavelet domain.

To facilitate the transition, a simple translator is being developed for converting existing waveforms to the wavelet domain. Ready to use algorithms for estimating the evolutionary power spectra are also being developed. The translator takes as input the amplitude and phase of each harmonic, either in the time domain or the frequency domain, then based on the frequency evolution, determines which fast transform to apply. The transform returns the nonzero wavelet coefficients as a sparse array with an indexing table to facilitate fast likelihood calculations. A range of tools for estimating the evolutionary power spectrum are also being developed, including simple smoothed median estimators like the ones described here, in addition to more sophisticated adaptive (transdimensional) Bayesian methods.

Codes for computing the WDM transform, with examples of the fast signal transform methods can be found at https://github.com/eXtremeGravityInstitute/WDM_Transform.

ACKNOWLEDGMENTS

This work was supported by the NASA LISA foundation Science Grant No. 80NSSC19K0320 and NSF Awards No. PHY1607343 and No. PHY1912053. This work was initiated while the author was on sabbatical at the Observatoire de la Côte d’Azur, kindly hosted by Nelson Christensen, and supported in part by the Centre National d’Études Spatiales. The author benefited from discussions with Will Farr, Sergey Klimenko and Guy Nason. This research has made use of data obtained from the Gravitational Wave Open Science Center [64], a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

- [1] R. van Haasteren, Y. Levin, P. McDonald, and T. Lu, *Mon. Not. R. Astron. Soc.* **395**, 1005 (2009).
- [2] C. Messick, K. Blackburn, P. Brady, P. Brockill, K. Cannon, R. Cariou, S. Caudill, S. J. Chamberlin, J. D. Creighton, R. Everett *et al.*, *Phys. Rev. D* **95**, 042001 (2017).
- [3] S. Sachdev *et al.*, [arXiv:1901.08580](#).
- [4] G. Carullo, W. Del Pozzo, and J. Veitch, *Phys. Rev. D* **99**, 123029 (2019); **100**, 089903(E) (2019).
- [5] M. Isi, M. Giesler, W. M. Farr, M. A. Scheel, and S. A. Teukolsky, *Phys. Rev. Lett.* **123**, 111102 (2019).
- [6] S. Klimenko, S. Mohanty, M. Rakhmanov, and G. Mitselmakher, *Phys. Rev. D* **72**, 122002 (2005).
- [7] S. Klimenko *et al.*, *Phys. Rev. D* **93**, 042004 (2016).
- [8] B. P. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **37**, 055002 (2020).
- [9] A. Buonanno, B. Iyer, E. Ochsner, Y. Pan, and B. Sathyaprakash, *Phys. Rev. D* **80**, 084043 (2009).
- [10] P. Ajith, S. Babak, Y. Chen, M. Hewitson, B. Krishnan, A. M. Sintes, J. T. Whelan, B. Brügmann, P. Diener, N. Dorband *et al.*, *Phys. Rev. D* **77**, 104017 (2008).
- [11] P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Brügmann, N. Dorband, D. Müller, F. Ohme, D. Pollney, C. Reisswig *et al.*, *Phys. Rev. Lett.* **106**, 241101 (2011).
- [12] L. Santamaria *et al.*, *Phys. Rev. D* **82**, 064016 (2010).
- [13] M. Hannam, P. Schmidt, A. Boh, L. Haegel, S. Husa, F. Ohme, G. Pratten, and M. Prer, *Phys. Rev. Lett.* **113**, 151101 (2014).
- [14] S. Khan, S. Husa, M. Hannam, F. Ohme, M. Prer, X. J. Forteza, and A. Boh, *Phys. Rev. D* **93**, 044007 (2016).
- [15] C. Garca-Quirs, M. Colleoni, S. Husa, H. Estells, G. Pratten, A. Ramos-Buades, M. Mateu-Lucena, and R. Jaume, *Phys. Rev. D* **102**, 064002 (2020).
- [16] K. Chatziioannou, A. Klein, N. Yunes, and N. Cornish, *Phys. Rev. D* **95**, 104004 (2017).
- [17] B. Moore, T. Robson, N. Loutrel, and N. Yunes, *Classical Quantum Gravity* **35**, 235006 (2018).
- [18] M. B. Priestley, *J. R. Stat. Soc. Ser. B* **27**, 204 (1965).
- [19] V. Nacula, S. Klimenko, and G. Mitselmakher, *J. Phys. Conf. Ser.* **363**, 012032 (2012).
- [20] L. S. Finn, *Phys. Rev. D* **46**, 5236 (1992).
- [21] B. Allen, W. G. Anderson, P. R. Brady, D. A. Brown, and J. D. E. Creighton, *Phys. Rev. D* **85**, 122006 (2012).
- [22] E. Cuoco, G. Losurdo, G. Calamai, L. Fabbri, M. Mazzoni, R. Stanga, G. Guidi, and F. Vetrano, *Phys. Rev. D* **64**, 122002 (2001).
- [23] E. Cuoco, L. Fabbri, M. Mazzoni, R. Stanga, G. Calamai, G. Losurdo, and F. Vetrano, *Classical Quantum Gravity* **18**, 1727 (2001).
- [24] T. B. Littenberg and N. J. Cornish, *Phys. Rev. D* **91**, 084034 (2015).
- [25] M. C. Edwards, R. Meyer, and N. Christensen, *Phys. Rev. D* **92**, 064011 (2015).
- [26] J. Aasi *et al.* (VIRGO Collaboration), *Classical Quantum Gravity* **29**, 155002 (2012).
- [27] J. Aasi *et al.* (LIGO Scientific and VIRGO Collaborations), *Classical Quantum Gravity* **32**, 115012 (2015).
- [28] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Classical Quantum Gravity* **33**, 134001 (2016).
- [29] M. R. Adams and N. J. Cornish, *Phys. Rev. D* **89**, 022001 (2014).
- [30] E. Kl and P. Stanica, *J. Comput. Appl. Math.* **237**, 126 (2013).
- [31] B. Zackay, T. Venumadhav, J. Roulet, L. Dai, and M. Zaldarriaga, [arXiv:1908.05644](#).
- [32] S. Mozzon, L. Nuttall, A. Lundgren, T. Dent, S. Kumar, and A. H. Nitz, *Classical Quantum Gravity* **37**, 215014 (2020).
- [33] E. Cuoco, G. Cella, and G. Guidi, *Classical Quantum Gravity* **21**, S801 (2004).
- [34] R. Dahlhaus, in *Time Series Analysis: Methods and Applications, Vol. 30 of Handbook of Statistics*, edited by T. Subba Rao, S. Subba Rao, and C. Rao (Elsevier, New York, 2012), pp. 351–413.
- [35] R. Dahlhaus, *Ann. Stat.* **25**, 1 (1997).
- [36] B. W. Silverman, J. C. Vassilicos, G. P. Nason, and R. v. Sachs, *Phil. Trans. R. Soc. A* **357**, 2511 (1999).
- [37] S. Mallat, G. Papanicolaou, and Z. Zhang, *Ann. Stat.* **26**, 1 (1998).
- [38] H. C. Ombao, J. A. Raz, R. von Sachs, and B. A. Malow, *J. Am. Stat. Assoc.* **96**, 543 (2001).
- [39] J. F. Macias-Perez and A. Bourrachot (ARCHEOPS Collaboration), *Astron. Astrophys.* **459**, 987 (2006).
- [40] G. P. Nason, R. Von Sachs, and G. Kroisandt, *J. R. Stat. Soc. Ser. B* **62**, 271 (2000).
- [41] A. Cardinali and G. P. Nason, *J. Time Ser. Anal.* **38**, 151 (2017).
- [42] S. Chatterji, L. Blackburn, G. Martin, and E. Katsavounidis, *Classical Quantum Gravity* **21**, S1809 (2004).
- [43] Y. Meyer, *Ondelettes et Opérateurs: Ondelettes*, Actualités Mathématiques (Hermann, Paris, 1990).
- [44] B. Allen, *Phys. Rev. D* **71**, 062001 (2005).
- [45] V. Tiwari *et al.*, *Classical Quantum Gravity* **32**, 165014 (2015).
- [46] D. Gabor, *J. Instit. Electr. Eng. III* **93**, 429 (1946).
- [47] M. Vogt, *Ann. Stat.* **40**, 2601 (2012).
- [48] N. J. Kasdin, *Proc. IEEE* **83**, 802 (1995).
- [49] Q. Baghi, I. Thorpe, J. Slutsky, J. Baker, T. Dal Canton, N. Korsakova, and N. Karnesis, *Phys. Rev. D* **100**, 022003 (2019).
- [50] L. Andersson, N. Hall, B. Jawerth, and G. Peters, in *Topics in the Theory and Applications of Wavelets*, edited by L. L. Schumaker (Academic Press, New York, 1995), pp. 1–61.
- [51] A. Cohen, I. Daubechies, and P. Vial, *Applied and Computational Harmonic Analysis* (1993), <https://hal.archives-ouvertes.fr/hal-01311753>.
- [52] J. R. Williams and K. Amaratunga, *J. Fourier Anal. Appl.* **3**, 435 (1997).
- [53] B. Abbott *et al.* (LIGO Scientific and Virgo Collaborations), *Phys. Rev. Lett.* **119**, 161101 (2017).
- [54] N. J. Cornish and T. B. Littenberg, *Classical Quantum Gravity* **32**, 135012 (2015).
- [55] C. Pankow *et al.*, *Phys. Rev. D* **98**, 084016 (2018).
- [56] S. Roy, A. S. Sengupta, and K. Arun, [arXiv:1910.04565](#).
- [57] S. Mann and S. Haykin, The Chirplet Transform: A Generalization of Gabor's Logon Transform, in *Vision Interface '91* (Canadian Image Processing and Pattern Recognition Society, Calgary, Alberta, 1991), pp. 205–212.
- [58] M. Millhouse, N. J. Cornish, and T. Littenberg, *Phys. Rev. D* **97**, 104057 (2018).

- [59] D. Hils, P. L. Bender, and R. F. Webbink, *Astrophys. J.* **360**, 75 (1990).
- [60] C. Cutler, *Phys. Rev. D* **57**, 7089 (1998).
- [61] N. J. Cornish and K. Shuman, *Phys. Rev. D* **101**, 124008 (2020).
- [62] N. J. Cornish and T. B. Littenberg, *Phys. Rev. D* **76**, 083006 (2007).
- [63] T. Littenberg, N. Cornish, K. Lackeos, and T. Robson, *Phys. Rev. D* **101**, 123021 (2020).
- [64] <https://www.gw-openscience.org>.