Supplementary information for *Wavelet–based tools to analyze, filter, and reconstruct transient gravitational–wave signals*

Andrea Virtuoso, Edoardo Milotti

A The non–standard inversion formula

A.1 Inversion formula

Since the wavelet Qp–transform extends the wavelet Q–transform, we prove the inversion formula for the more general case of the Qp–transform – the inversion formula for the Q–transform follows as a special case with $p = 0$.

We start from the definition of the wavelet Qp–transform in the frequency domain

$$T(\tau,\nu,Q,p) = \left(\frac{2}{\pi Q^2}\right)^{1/4} \int_{-\infty}^{+\infty} df\, \tilde{s}(f) \frac{1}{\sqrt{2\pi\nu}} \sqrt{\pi} \frac{Q}{\sqrt{1+2iQp}} e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2} e^{2\pi i f\tau}, \tag{A.1}$$

where $\tilde{s}(f) = \int_{-\infty}^{+\infty} dt\, s(t) e^{-2\pi i f t}$ is the Fourier transform of $s(t)$. Next, we take the derivative with respect to the time parameter τ

$$\frac{\partial}{\partial\tau} T(\tau,\nu,Q,p) = \left(\frac{2}{\pi Q^2}\right)^{1/4} \int_{-\infty}^{+\infty} df\, \tilde{s}(f) 2\pi i f \frac{1}{\sqrt{2\pi\nu}} \sqrt{\pi} \frac{Q}{\sqrt{1+2iQp}} e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2} e^{2\pi i f\tau} \tag{A.2}$$

and divide by

$$\left(\frac{2}{\pi Q^2}\right)^{1/4} i\sqrt{2\pi\nu}\, 2\pi\nu$$

obtaining

$$\left(\frac{2}{\pi Q^2}\right)^{-1/4} \frac{1}{i\sqrt{2\pi\nu} 2\pi\nu} \frac{\partial}{\partial\tau} T(\tau,\nu,Q,p) = \int_{-\infty}^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau} \frac{f}{\nu^2} \frac{1}{\sqrt{\pi}} \frac{Q}{2\sqrt{1+2iQp}} e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2}. \tag{A.3}$$

Note that with the original Q–transform of eq. (6a) in the main text, this result would be different, the following calculations could not be performed analytically, and it would not be possible to find an analytical inversion formula.

Integrating over the frequency parameter ν, we find

$$\begin{aligned}
&\left(\frac{2}{\pi Q^2}\right)^{-1/4} \int_0^{+\infty} d\nu \frac{1}{i\sqrt{2\pi\nu}\, 2\pi\nu} \frac{\partial}{\partial\tau} T(\tau,\nu,Q,p) \\
&= \int_{-\infty}^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau} \left[\int_0^{+\infty} d\nu \frac{f}{\nu^2} \frac{1}{\sqrt{\pi}} \frac{Q}{2\sqrt{1+2iQp}} e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2} \right] \\
&= \int_{-\infty}^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau} I(f,Q,p)
\end{aligned} \tag{A.4}$$

where the integral $I(f,Q,p)$ is defined by the expression

$$I(f,Q,p) \equiv \int_0^{+\infty} d\nu \frac{f}{\nu^2} \frac{1}{\sqrt{\pi}} \frac{Q}{2\sqrt{1+2iQp}} e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2} \tag{A.5}$$

which includes all the ν-dependent terms. This integral can be computed using the change of variable

$$\gamma = f/\nu \Rightarrow d\gamma = -\frac{f}{\nu^2} d\nu. \tag{A.6}$$

Here, it is necessary to distinguish between positive and negative frequencies f

$$f > 0 \Rightarrow I(f,Q,p) = \frac{1}{\sqrt{\pi}} \frac{Q}{2\sqrt{1+2iQp}} \int_0^{+\infty} d\gamma e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}(\gamma-1)\right)^2} \tag{A.7a}$$

$$f < 0 \Rightarrow I(f,Q,p) = \frac{1}{\sqrt{\pi}} \frac{Q}{2\sqrt{1+2iQp}} \int_0^{-\infty} d\gamma e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}(\gamma-1)\right)^2}. \tag{A.7b}$$

Then, recalling the definition of error function

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x dt\, e^{-t^2},$$

eqs. (A.7a) and (A.7b) become

$$f > 0 \Rightarrow I(f,Q,p) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right) + 1\right] \tag{A.8a}$$

$$f < 0 \Rightarrow I(f,Q,p) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right) - 1\right]. \tag{A.8b}$$

It is important to remark that the analytical calculation of the integrals $I(f,Q,p)$ is only possible thanks to the chosen parameterization of the frequency derivative: indeed, as discussed above, taking a "resolution-dependent" frequency derivative – i.e. depending on σ_τ and σ_ν, and therefore on both ν and Q – allows the factorization of the exponential part of the frequency derivative. Trivially, this result also holds in the case of $p = 0$ (wavelet Q–transform). Plugging this result into eq. (A.4), we find

$$\begin{aligned}
&\left(\frac{2}{\pi Q^2}\right)^{-1/4} \int_0^{+\infty} d\nu \frac{1}{i\sqrt{2\pi\nu}\, 2\pi\nu} \frac{\partial}{\partial\tau} T(\tau,\nu,Q,p) \\
&= \int_{-\infty}^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau} I(f,Q,p) \\
&= \int_0^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau} I(f>0,Q,p) + \tilde{s}^*(f) e^{-2\pi i f\tau} I(f<0,Q,p) \\
&= \frac{1}{2}\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right) \int_0^{+\infty} df\, \left[\tilde{s}(f) e^{2\pi i f\tau} + \tilde{s}^*(f) e^{-2\pi i f\tau}\right] \\
&+ \frac{1}{2}\int_0^{+\infty} df\, \left[\tilde{s}(f) e^{2\pi i f\tau} - \tilde{s}^*(f) e^{-2\pi i f\tau}\right] \\
&= \frac{1}{2}\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right) s(\tau) + i\,\mathrm{Im}\left[\int_0^{+\infty} df\, \tilde{s}(f) e^{2\pi i f\tau}\right].
\end{aligned} \tag{A.9}$$

Taking the real part of (A.9) we find the general inversion formula, which holds for both the wavelet Qp–transform and for the wavelet Q–transform (setting $p = 0$):

$$s(\tau) = \frac{2}{\left(\frac{2}{\pi Q^2}\right)^{1/4} \mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right)\right]} \mathrm{Re}\left[\int_0^{+\infty} d\nu\, \frac{1}{i\sqrt{2\pi\nu}\, 2\pi\nu} \frac{\partial}{\partial\tau} T(\tau,\nu,Q,p)\right]. \tag{A.10}$$

A.2 Denoising formula

We turn now to the time-frequency denoising formulas of eqs. (21) and (22) in the main text. In practice, denoising amounts to carrying out the following substitution in the reconstruction formula of eq. (A.10)

$$\frac{\partial}{\partial\tau} T(\tau,\nu,Q) \;\rightarrow\; \mathbf{1}_C(\tau,\nu)\, \frac{\partial}{\partial\tau} T(\tau,\nu,Q) \tag{A.11}$$

where $\mathbf{1}_C(\tau,\nu)$ is the indicator function introduced in the main text. Then eq. (A.10) becomes

$$
\begin{aligned}
s_D(\tau) &= \frac{2}{\left(\frac{2}{\pi Q^2}\right)^{1/4}\mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+i2Qp}}\right)\right]}\,\mathrm{Re}\left[\int_0^{+\infty} d\nu\,\frac{1}{i\sqrt{2\pi\nu}2\pi\nu}\frac{\partial}{\partial\tau}T(\tau,\nu)\mathbf{1}_C(\tau,\nu)\right] \\
&= \frac{2}{\mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+i2Qp}}\right)\right]}\mathrm{Re}\left[\int_{-\infty}^{+\infty} df\,\tilde{s}(f)e^{2\pi i f\tau}\left(\int_0^{+\infty} d\nu\frac{f}{\nu^2}\frac{1}{\sqrt{\pi}}\frac{Q}{2\sqrt{1+i2Qp}}e^{-\left(\frac{Q}{2\sqrt{1+i2Qp}}\frac{f-\nu}{\nu}\right)^2}\mathbf{1}_C(\tau,\nu)\right)\right] \\
&= \mathrm{Re}\left[\int_{-\infty}^{+\infty} df\,\tilde{s}(f)e^{2\pi i f\tau}w(\tau,f,Q,p)\right] \qquad \text{(A.12)}
\end{aligned}
$$

where $s_D(\tau)$ is the denoised signal and where we have introduced the time-frequency denoising window $w(\tau,f,Q,p)$.

Now, we fix Q and p, and for each τ we define a generic indicator function $\mathbf{1}_C(\tau,\nu)$ which selects a (finite) set of intervals $\left[\nu_l^{\mathrm{low}}(\tau),\nu_l^{\mathrm{high}}(\tau)\right]$ with $l=1,\ldots,L$, so that the denoising time-frequency window becomes:

$$
\begin{aligned}
w(\tau,f,Q,p) &= \frac{2}{\mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right)\right]}\int_0^{+\infty} d\nu\,\frac{f}{\nu^2}\frac{1}{\sqrt{\pi}}\frac{Q}{2\sqrt{1+2iQp}}e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2}\mathbf{1}_C(\tau,\nu) \qquad \text{(A.13)} \\
&= \frac{2}{\mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right)\right]}\sum_k\int_{\nu_k^{\mathrm{low}}(\tau)}^{\nu_k^{\mathrm{high}}(\tau)} d\nu\,\frac{f}{\nu^2}\frac{1}{\sqrt{\pi}}\frac{Q}{2\sqrt{1+2iQp}}e^{-\left(\frac{Q}{2\sqrt{1+2iQp}}\frac{f-\nu}{\nu}\right)^2}. \qquad \text{(A.14)}
\end{aligned}
$$

Finally, the compact representation of the denoising formula is

$$
s_D(\tau) = \mathrm{Re}\left[\int_{-\infty}^{+\infty} df\,\tilde{s}(f)e^{2\pi i f\tau}w(\tau,f,Q,p)\right] \qquad \text{(A.15)}
$$

with the denoising window

$$
w(\tau,f,Q,p) = \frac{1}{\mathrm{Re}\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\right)\right]}\sum_l\left[\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\left(\frac{f-\nu_l^{\mathrm{low}}(\tau)}{\nu_l^{\mathrm{low}}(\tau)}\right)\right)-\mathrm{erf}\left(\frac{Q}{2\sqrt{1+2iQp}}\left(\frac{f-\nu_l^{\mathrm{high}}(\tau)}{\nu_l^{\mathrm{high}}(\tau)}\right)\right)\right]. \qquad \text{(A.16)}
$$

B The pdf of $|T(\tau,\nu,Q,p)|^2$ for a Gaussian white noise background

Following the standard numerical implementations of the Fourier transform, we obtain a straightforward discretization of eq. (19c) as follows

$$T(\tau,\nu,Q,p) = \frac{f_s}{N}\sum_m \tilde{s}(f_m)\psi^*(f_m;\tau,\nu,Q,p), \tag{B.1}$$

As already mentioned in Section III C of the main text, in addition to the standard version of eq. (19c) our code also implements a scaled (dimensionless) version of the wavelet Q– and Qp–transform, which is

$$T_{\rm nd}(\tau,\nu,Q,p) = \frac{\sqrt{f_s}}{N}\sum_m \tilde{s}(f_m)\psi^*(f_m;\tau,\nu,Q,p), \tag{B.2}$$

or, in the time domain,

$$T_{\rm nd}(\tau,\nu,Q,p) = \frac{1}{\sqrt{f_s}}\sum_k s(t_k)\psi^*(t_k;\tau,\nu,Q,p), \tag{B.3}$$

where $\tilde{s}(f_m) = \sum_k s(t_k)e^{-2\pi i f_m t_k}$, $m = -N/2+1, ..., +N/2$ and $k = 0, ..., N-1$.

The scaled version of the transform is particularly well–suited to the analysis of the fluctuations of stochastic signals such as GW interferometer noise. To see why this is so, we start from a discretized version of the original transform, eq. (19a)

$$T(\tau,\nu,Q,p) = \frac{1}{f_s}\sum_k n(t_k)\psi^*(t_k;\tau,\nu,Q,p), \tag{B.4}$$

where $n(t_k)$ represents a zero-mean Gaussian white noise with $\langle n(t_k)n(t_l)\rangle = \sigma^2\ \delta_{kl}$, where the angle brackets indicate the ensemble average.

We start by writing the Qp–wavelet as follows

$$\psi^*(t_k;\tau,\nu,Q,p) = g(t_k;\tau,\nu,Q)e^{-2\pi i\left[\nu+\frac{1}{2\pi}\left(\frac{2\pi\nu}{Q}\right)^2 2pQ(t_k-\tau)\right](t_k-\tau)} \tag{B.5a}$$

$$g(t_k;\tau,\nu,Q) = \left(\frac{8\pi\nu^2}{Q^2}\right)^{1/4} e^{-\left(\frac{2\pi\nu(t_k-\tau)}{Q}\right)^2}. \tag{B.5b}$$

where g is the Gaussian window function used in the definition of the wavelet Q–transform. Next, we split $T(\tau,\nu,Q,p)$ into real and imaginary parts, respectively a and $-b$, as follows

$$T(\tau,\nu,Q,p) = a - ib \tag{B.6a}$$

$$a = \frac{1}{\sqrt{f_s}}\sum_k n(t_k)g^*(t_k;\tau,\nu,Q,p)\cos\left\{Q\left[\frac{2\pi\nu(t_k-\tau)}{Q} + 2p\left(\frac{2\pi\nu(t_k-\tau)}{Q}\right)^2\right]\right\}, \tag{B.6b}$$

$$b = \frac{1}{\sqrt{f_s}}\sum_l n(t_l)g^*(t_l;\tau,\nu,Q,p)\sin\left\{Q\left[\frac{2\pi\nu(t_l-\tau)}{Q} + 2p\left(\frac{2\pi\nu(t_l-\tau)}{Q}\right)^2\right]\right\}. \tag{B.6c}$$

We find that both a and b have zero mean that and asymptotically they are nearly statistically independent from each other. Indeed, for a large number ($N \gg 1$) of data samples, and using the shorthand notation $x_k = 2\pi\nu(t_k-\tau)/Q$ and $g(t_k;\tau,\nu,Q) = g(x_k)$, we find

$$\langle a\rangle = \frac{1}{f_s}\sum_k g(x_k)\cos\left[Q(x_k+2px_k^2)\right]\underbrace{\langle n(t_k)\rangle}_{=0} = 0 \tag{B.7a}$$

$$\langle b\rangle = \frac{1}{f_s}\sum_l g(x_l)\sin\left[Q(x_l+2px_l^2)\right]\underbrace{\langle n(t_l)\rangle}_{=0} = 0 \tag{B.7b}$$

$$\langle ab\rangle = \frac{1}{f_s^2}\sum_{k,l} g(x_k)g(x_l)\underbrace{\langle n(t_k)n(t_l)\rangle}_{=\delta_{kl}\sigma^2}\cdot\cos\left[Q(x_k+2px_k^2)\right]\sin\left[Q(x_l+2px_l^2)\right] \tag{B.7c}$$

$$= \frac{\sigma^2}{f_s^2}\sum_k g(x_k)^2\cos\left[Q(x_k+2px_k^2)\right]\sin\left(Q(x_k+2px_k^2)\right) \tag{B.7d}$$

$$\simeq \frac{\sigma^2}{f_s}\sqrt{\frac{2}{\pi}}\int_{-\infty}^{+\infty} e^{-2x^2}\cos\left[Q(x+2px^2)\right]\sin\left[Q(x+2px^2)\right]dx \tag{B.7e}$$

$$= -\frac{\sigma^2}{f_s}\frac{1}{8}\left(\frac{1}{1+(2pQ)^2}\right)^{3/4}\exp\left(-\frac{1}{2}\frac{Q^2}{1+(2pQ)^2}\right)\cdot$$
$$\cdot\left[3\left(\sin(z_1+z_2)-2pQ\cos(z_1+z_2)\right)+\sqrt{1+(2pQ)^2}\sin(z_1-z_2)\right]\approx 0. \tag{B.7f}$$

where

$$z_1 = pQ\frac{Q^2}{1+(2pQ)^2} \tag{B.8}$$

$$z_2 = \frac{1}{2}\arctan(2pQ). \tag{B.9}$$

Eq. (B.7f) is exact for $p=0$, i.e., for the case of the wavelet Q–transform, and maintains very high precision for all practical choices of Q and p (i.e., as explained in Section III in the main text, $Q \gtrsim 2\pi$ and $p \lesssim 1/Q$): even if the result of the integral is not monotone in Q and p, it is well approximated by the exponential term, which for $Q > 2\pi$ and $p < 0.1$ is less than 4.7×10^{-4}, and, therefore, a and b have almost vanishing correlation when compared to the the σ^2/f_s prefactor.

We find similar expressions for the variances of a and b

$$\sigma_a^2 = \langle a^2\rangle = \frac{1}{f_s^2}\sum_{k,l} g(x_k)g(x_l)\underbrace{\langle n(t_k)n(t_l)\rangle}_{=\delta_{kl}\sigma^2}\cdot\cos\left(Q(x_k+2px_k^2)\right)\cos\left(Q(x_l+2px_l^2)\right) \tag{B.10a}$$

$$= \frac{\sigma^2}{f_s^2}\sum_k g(x_k)^2\cos^2\left[Q(x_k+2px_k^2)\right] \tag{B.10b}$$

$$\simeq \frac{\sigma^2}{f_s}\sqrt{\frac{2}{\pi}}\int_{-\infty}^{+\infty} e^{-2x^2}\cos^2\left[Q(x+2px^2)\right]dx \tag{B.10c}$$

$$= \frac{\sigma^2}{f_s}\frac{1}{2}\left\{1+\left[1+(2pQ)^2\right]^{-1/4} e^{-\frac{1}{2}\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2}\cos\left[pQ\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2-\frac{1}{2}\arctan(2pQ)\right]\right\}, \tag{B.10d}$$

$$\sigma_b^2 = \langle b^2\rangle = \frac{1}{f_s^2}\sum_{k,l} g(x_k)g(x_l)\underbrace{\langle n(t_k)n(t_l)\rangle}_{=\delta_{kl}\sigma^2}\cdot\sin\left(Q(x_k+2px_k^2)\right)\sin\left(Q(x_l+2px_l^2)\right) \tag{B.11a}$$

$$= \frac{\sigma^2}{f_s^2}\sum_k g(x_k)^2\sin^2\left[Q(x_k+2px_k^2)\right] \tag{B.11b}$$

$$\simeq \frac{\sigma^2}{f_s}\sqrt{\frac{2}{\pi}}\int_{-\infty}^{+\infty} e^{-2x^2}\sin^2\Big[Q(x+2px^2)\Big]dx \tag{B.11c}$$

$$= \frac{\sigma^2}{f_s}\frac{1}{2}\left\{1-\Big[1+(2pQ)^2\Big]^{-1/4} e^{-\frac{1}{2}\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2}\cos\left[pQ\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2-\frac{1}{2}\arctan(2pQ)\right]\right\}, \tag{B.11d}$$

The above formulas are simpler if we use the scaled version of the discretized transform in the analysis of the noise background, because the dependence on sampling rate disappears. Note that this dependence does not show up in the analysis of deterministic signals, where $\langle s(t_k)s(t_l)\rangle = s(t_k)s(t_l)$.

The corresponding formulas for the scaled transform are

$$(\sigma_a^2)_{\rm nd} = \frac{1}{2}\left\{1+\Big[1+(2pQ)^2\Big]^{-1/4} e^{-\frac{1}{2}\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2}\cos\left[pQ\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2-\frac{1}{2}\arctan(2pQ)\right]\right\} \tag{B.12a}$$

$$(\sigma_b^2)_{\rm nd} = \frac{1}{2}\left\{1-\Big[1+(2pQ)^2\Big]^{-1/4} e^{-\frac{1}{2}\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2}\cos\left[pQ\left(\frac{Q}{\sqrt{1+(2pQ)^2}}\right)^2-\frac{1}{2}\arctan(2pQ)\right]\right\}, \tag{B.12b}$$

which become even simpler in the case of the Q–transform ($p = 0$)

$$(\sigma_a^2)_{\rm nd} = (1+e^{-Q^2/2})/2 \tag{B.13a}$$

$$(\sigma_b^2)_{\rm nd} = (1-e^{-Q^2/2})/2 \tag{B.13b}$$

For $Q > 2\pi$, eqs. (B.13a) and (B.13b) differ by less than about 3×10^{-9}. The difference between eqs. (B.12a) and (B.12b) is larger, but still small in absolute terms, e.g., for $Q > 2\pi$ and $p = 0.1$ it is smaller than $3.7\times10{-}4$. This means that, for all practical values of Q and p, $(\sigma_a^2)_{\rm nd} \simeq (\sigma_b^2)_{\rm nd} \simeq 1/2$.

Combining these results, we find that we can treat a and b as independent, identically distributed random variables. Moreover, being $n(t_k)$ a Gaussian variate, a and b are Gaussian as well with nearly the same variance $(\sigma_a^2)_{\rm nd} \approx (\sigma_b^2)_{\rm nd} \approx (\sigma^2)_{\rm nd} \equiv 1/2$. Using these approximations, it is easy to show that $|T_{\rm nd}(\tau,\nu,Q,p)|^2 = |a|^2+|b|^2$ has a χ^2 probability density function (pdf) with 2 degrees of freedom[1]

$$P(|T_{\rm nd}(\tau,\nu,Q,p)|^2) = \frac{1}{2\cdot(\sigma^2)_{\rm nd}} e^{-\frac{|T_{\rm nd}(\tau,\nu,Q,p)|^2}{2\cdot(\sigma^2)_{\rm nd}}} = e^{-|T_{\rm nd}(\tau,\nu,Q,p)|^2}. \tag{B.14}$$

We use this pdf to establish the statistical significance of the energy thresholds used in the filtering process described in the main text. For example, only 0.1% of time-frequency regions due to Gaussian noise are above the energy threshold $E_{\rm thr} = 7$ used in this paper. As explained in the paper, this threshold requires a careful setting to avoid rejecting weak signals. Clearly, lowering the energy threshold increases the Gaussian noise contribution: as an example, 0.7% of time-frequency regions due to Gaussian noise are above the energy threshold $E_{\rm thr} = 5$ used for GW190521 Virgo data.

Finally, it is also interesting to note that this probability density function does not depend on the values of either Q or p (for reasonable values of both).

[1] For a proof see, e.g., [4], Section 6.1.3, Theorem 6.1.1, which deals with the very similar case of the power spectral density estimated with the discrete Fourier transform

C Two–point correlation function of the wavelet Qp–transform in the time–frequency domain

It is well–known that continuous wavelet transforms are highly redundant (see, e.g., [34]), and our sampled version of Section III C is no exception. Here, assuming Gaussian noise, i.e., $\sigma^2 = 1$, we derive an explicit expression for the statistical two-point correlation function of Qp–wavelets that represents a mathematical statement of the sampled wavelet redundancy.

Consider the discretized transform, evaluated at two different points in the time–frequency plane

$$T(\tau_0, \nu_0, Q, p) = \frac{1}{f_s} \sum_k n(t_k) \psi^*(t_k; \tau_0, \nu_0, Q, p) \tag{C.1a}$$

$$T(\tau, \nu, Q, p) = \frac{1}{f_s} \sum_l n(t_l) \psi^*(t_l; \tau, \nu, Q, p). \tag{C.1b}$$

and, just as in the previous section, define the following variables

$$x_k = \frac{2\pi\nu_0(t_k - \tau_0)}{Q} \tag{C.2a}$$

$$y_l = \frac{2\pi\nu(t_l - \tau)}{Q} \equiv \beta x_l + \gamma \tag{C.2b}$$

with $\beta = \nu/\nu_0$ and $\gamma = [2\pi\nu(\tau_0 - \tau)]/Q$. With these definitions, the wavelets in eqs. (C.1) can be written as

$$\psi^*(t_k; \tau_0, \nu_0, Q, p) = \left(\frac{8\pi\nu_0^2}{Q^2}\right)^{1/4} e^{-x_k^2 - iQx_k(1+2px_k)} \equiv \psi^*(x_k) \tag{C.3a}$$

$$\psi^*(t_l; \tau, \nu, Q, p) = \left(\frac{8\pi\nu^2}{Q^2}\right)^{1/4} e^{-y_l^2 - iQy_l(1+2py_l)} \equiv \psi^*(y_l). \tag{C.3b}$$

We can now compute the correlation between the transform at two different points (τ_0, ν_0) and (τ, ν)

$$\langle T^*(\tau_0, \nu_0, Q, p) T(\tau, \nu, Q, p) \rangle = \frac{1}{f_s^2} \sum_{k,l} \underbrace{\langle n(t_k) n(t_l) \rangle}_{= \sigma^2 \delta_{kl}} \psi(x_k) \psi^*(y_l) \tag{C.4a}$$

$$= \frac{\sigma^2}{f_s} \frac{1}{f_s} \sum_k \sqrt{\frac{8\pi\nu_0^2}{Q^2}} \sqrt{\beta} e^{-(x_k^2 + y_k^2) + iQ[x_k(1+2px_k) - y_k(1+2py_k)]} \tag{C.4b}$$

$$\simeq \frac{\sigma^2}{f_s} \sqrt{\frac{2\beta}{\pi}} \int_{-\infty}^{+\infty} e^{-(x^2 + (\beta x + \gamma)^2)) + iQ\left[x(1+2px) - (\beta x + \gamma)\cdot\left(1+2p(\beta x+\gamma)\right)\right]} dx \tag{C.4c}$$

$$= \frac{\sigma^2}{f_s} \sqrt{\frac{2\beta}{1 + \beta^2 + i2pQ(\beta^2 - 1)}} \exp\left[- \frac{\gamma^2 + \left(\frac{(1-\beta)Q}{2} + 2pQ\gamma\right)^2 + i(1+\beta)\gamma Q}{1 + \beta^2 + i2pQ(\beta^2 - 1)} \right] \tag{C.4d}$$

Note that the dependence on sampling rate disappears if we consider the same correlation for the scaled version of the transform

$$\langle T^*_{\rm nd}(\tau_0, \nu_0, Q, p) T_{\rm nd}(\tau, \nu, Q, p) \rangle = \sqrt{\frac{2\beta}{1 + \beta^2 + i2pQ(\beta^2 - 1)}} \exp\left[- \frac{\gamma^2 + \left(\frac{(1-\beta)Q}{2} + 2pQ\gamma\right)^2 + i(1+\beta)\gamma Q}{1 + \beta^2 + i2pQ(\beta^2 - 1)} \right]. \tag{C.5}$$

Figure 1 shows a few selected examples of the module of the correlation function for the scaled version of the transform, i.e. $|\langle T^*_{\rm nd}(\tau_0, \nu_0, Q, p) T_{\rm nd}(\tau, \nu, Q, p) \rangle|$, for different values of Q and p.

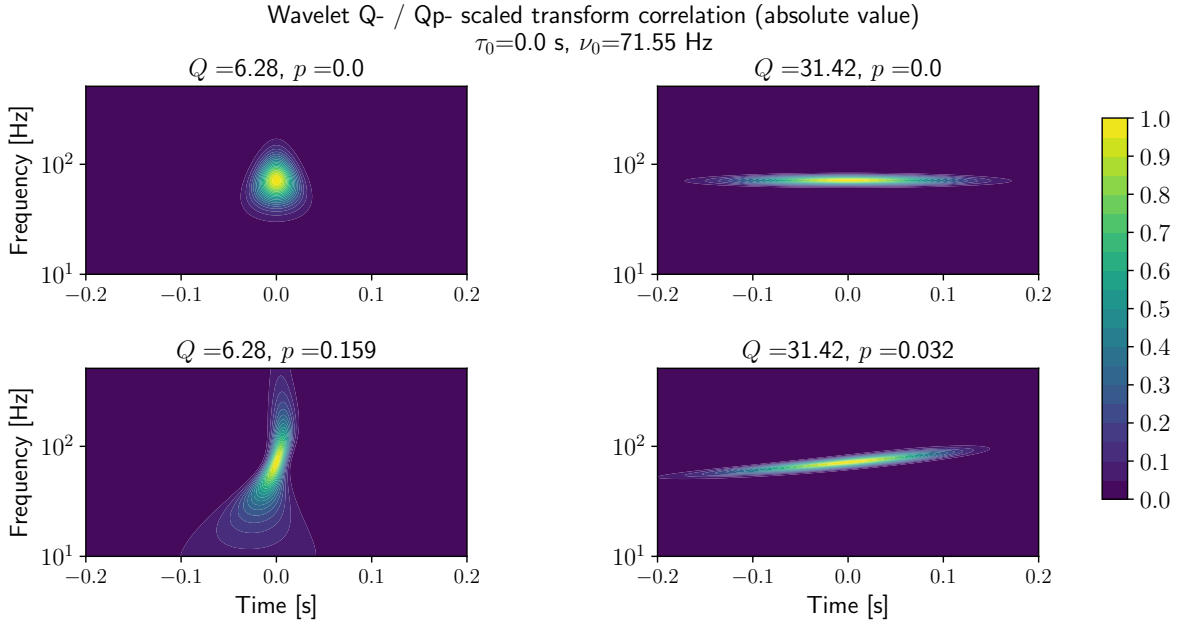

Figure 1: The module of the expected value for the correlation between $T^*_{\rm nd}(\tau_0, \nu_0, Q, p)$ and $T_{\rm nd}(\tau, \nu, Q, p)$, i.e. $|\langle T^*_{\rm nd}(\tau_0, \nu_0, Q, p) T_{\rm nd}(\tau, \nu, Q, p)\rangle|$, evaluated as a function of τ and ν, assuming Gaussian noise.

D Tables and Figures

In this section we report tables and figures referenced in Section IV in the main text.

GW150914						
	L1			H1		
	Q	Qp	%	Q	Qp	%
Energy peak	43.33	68.25	57.51%	80.31	106.79	32.97%
Energy density	18.30	23.79	29.98%	27.02	33.93	25.56%
TF area	4.84	3.89	−19.71%	6.97	5.73	−17.73%
Overlap with LAL	0.933	0.934	0.0%	0.967	0.972	0.52%
Residuals std dev	0.966	0.960	-	1.035	1.033	-

Table 1: Energy peak, energy density, TF area and overlap with LALInference waveform obtained for GW150914, for the L1 and H1 interferometers. The energy peak, energy density and the TF area are reported for both transforms, as well as the fractional difference between the two. The overlap has been calculated for both transforms with the waveforms obtained using the denoising formula of eqs. (21) and (22) or (12) and (13) and selecting the TF regions where the transform energy is larger than 7. Finally, the standard deviations of the respective residuals are also reported for both transforms.

GW190521									
	L1			H1			V1		
	Q	Qp	%	Q	Qp	%	Q	Qp	%
Energy peak	67.15	71.58	6.59%	26.59	32.80	23.33%	6.44	7.18	11.4%
Energy density	25.37	26.40	4.06%	14.47	15.59	7.76%	5.68	5.91	3.95%
TF area	2.76	2.67	−3.11%	1.84	1.77	−4.08%	0.27	0.36	31.69%
Overlap with LAL	0.943	0.943	0.0%	0.939	0.944	0.10%	0.539	0.692	28.39%
Residuals std dev	0.988	0.988	-	1.040	1.039	-	0.973	0.971	-

Table 2: Energy peak, energy density, TF area and overlap with LALInference waveform obtained for GW190521, for the L1, H1 and V1 interferometers. The energy peak, energy density and the TF area are reported for both transforms, as well as the fractional difference between the two. The overlap has been calculated for both transforms with the waveforms obtained using the denoising formula of eqs. (21) and (22) or (12) and (13) and selecting the TF regions where the transform energy is larger than 7 for L1 and H1 and larger than 5 for V1. Finally, the standard deviations of the respective residuals are also reported for both transforms.

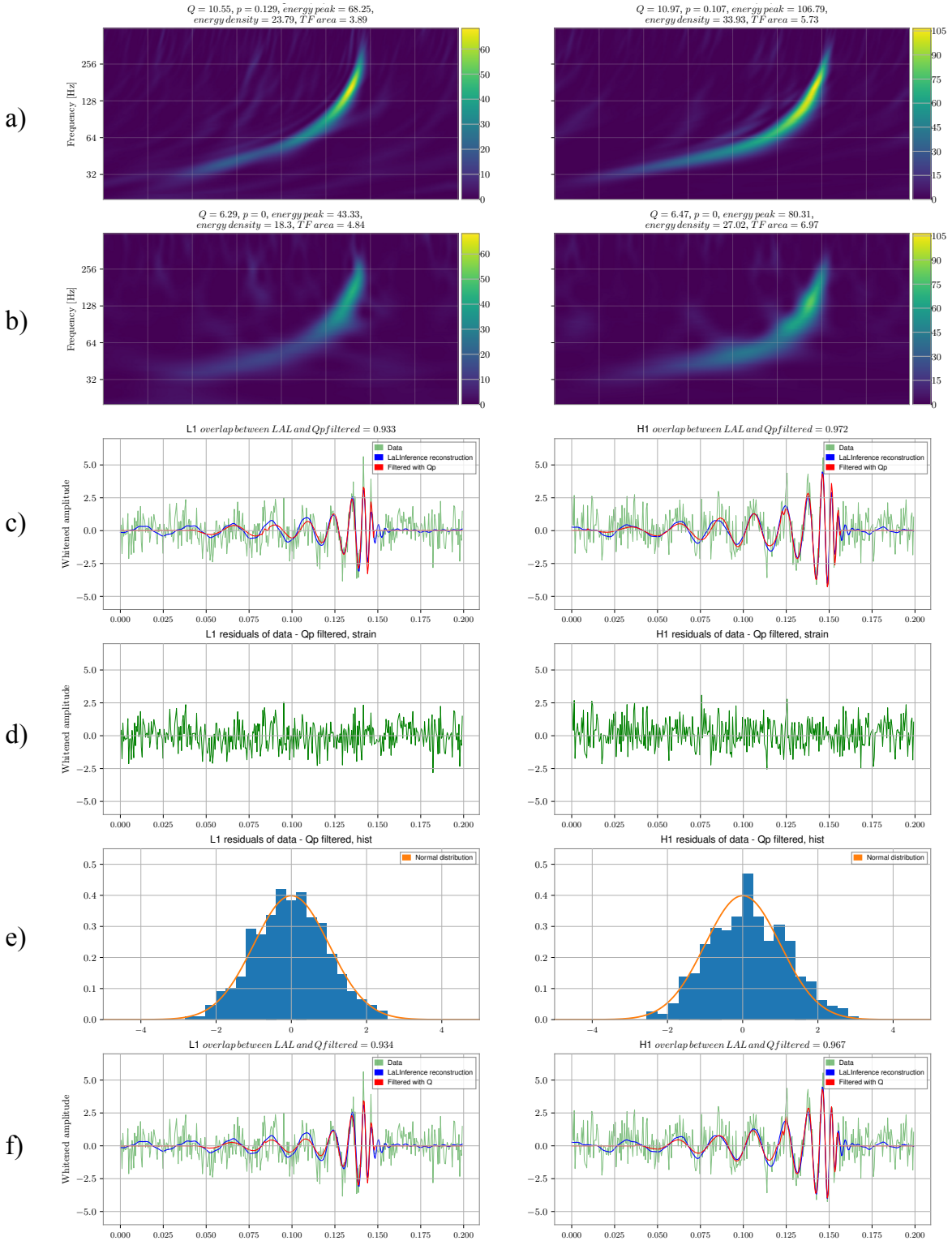

Figure 2: Q- and Qp-reconstruction of the first event GW150914, using data from L1 (left column) and H1 (right column). Rows **a)** and **b)** show, respectively, the wavelet Qp–transform and the wavelet Q–transform; the colorbars display the energy scale, i.e., the value of $|T_{\rm nd}(\tau,\nu,Q,p)|^2$. Row **c)** displays the original strain data (green), the waveform reconstructed with LALInference (blue) reported in [5] and the waveform obtained with the denoising eqs. (21) and (22) for the wavelet Qp–transform (red). Row **d)** shows the residuals corresponding to the wavelet Qp–transform reconstruction and row **e)** shows empirical distribution of these residuals. Row **f)** shows the same as row **c)** but for wavelet Q–transform, using the corresponding denoising eqs. (12) and (13).

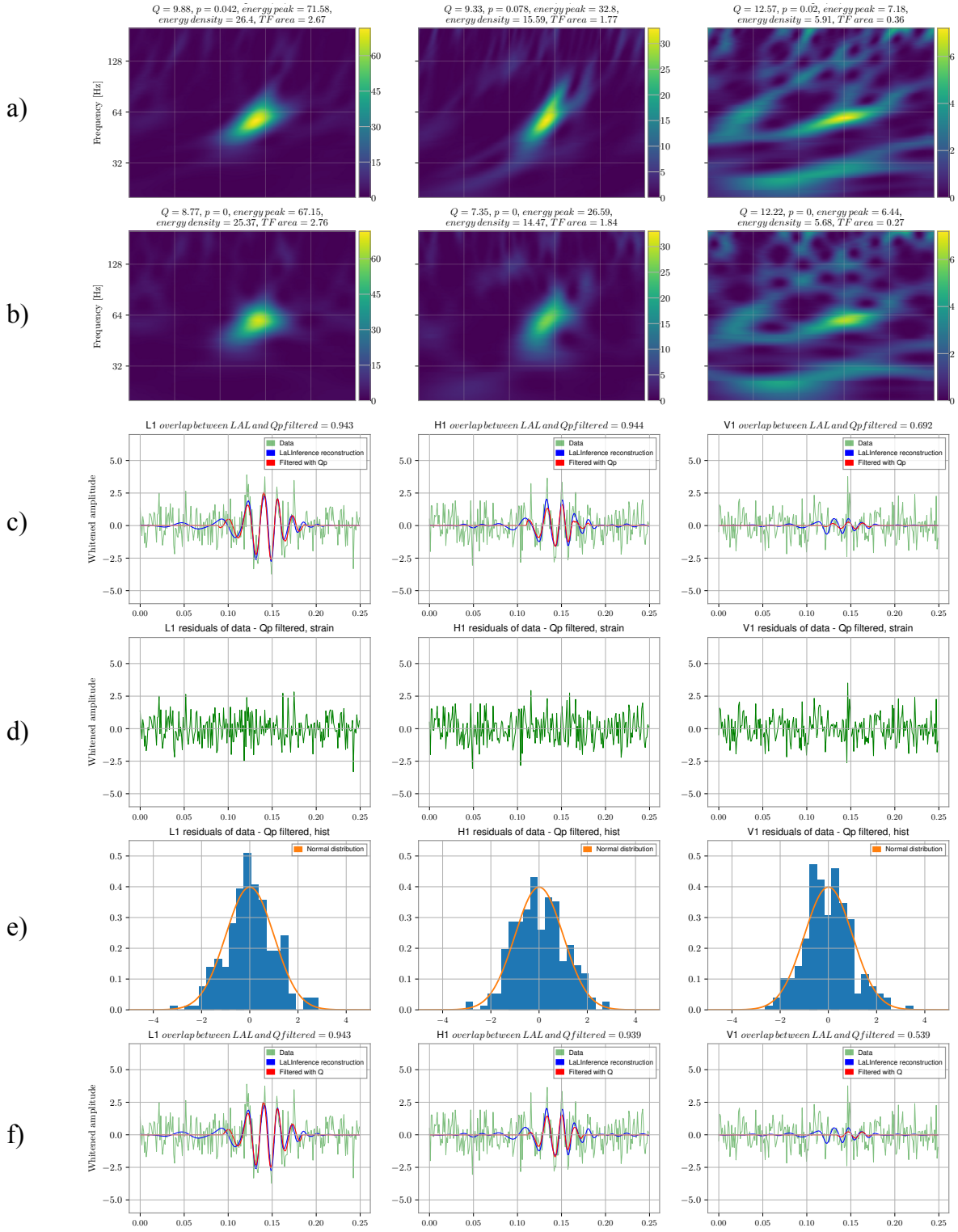

Figure 3: Q- and Qp-reconstruction of the first event GW190521, using data from L1 (left column), H1 (central column), and V1 (right column). Rows **a)** and **b)** show, respectively, the wavelet Qp–transform and the wavelet Q–transform; the color bars display the energy scale, i.e., the value of $|T_{\mathrm{nd}}(\tau, \nu, Q, p)|^2$. Row **c)** displays the original strain data (green), the waveform reconstructed with LALInference (blue) reported in [11] and the waveform obtained with the denoising eqs. (21) and (22) for the wavelet Qp–transform (red). Row **d)** shows the residuals corresponding to the wavelet Qp–transform reconstruction and row **e)** shows empirical distribution of these residuals. Row **f)** shows the same as row **c)** but for wavelet Q–transform, using the corresponding denoising eqs. (12) and (13).

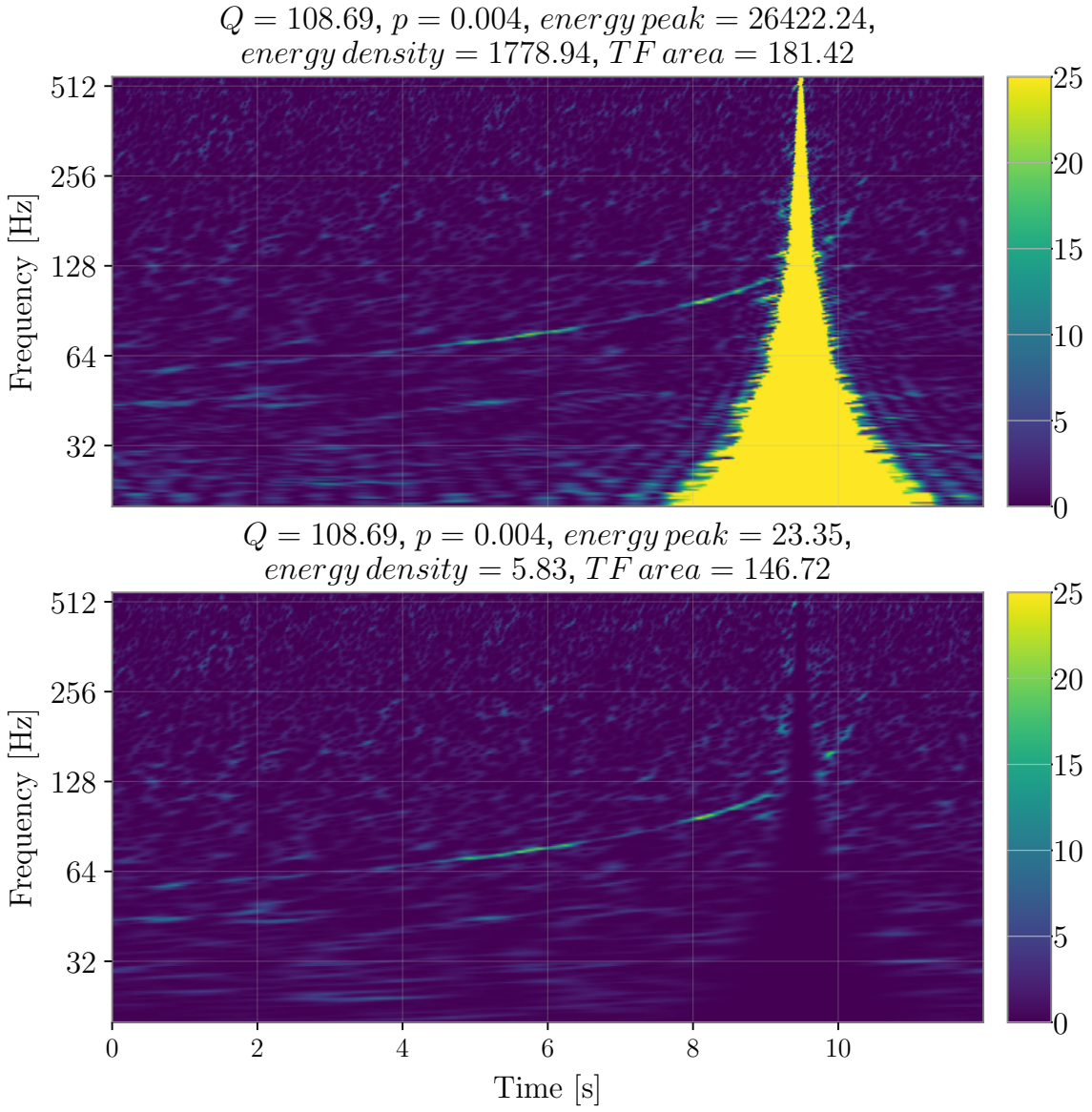

Figure 4: TF representation of the wavelet Qp–transform of the L1 data for GW170817. The first row represents the original data including a strong glitch; the color bar saturates at $E = 25$. The second row shows the result of a simple filtering operation where we have set to zero all the values of $|T_{\mathrm{nd}}(\tau, \nu, Q, p)|^2$ with an energy greater than 25. Finally, the glitch is almost completely removed.

References

[1] B. P. Abbott, R. Abbott, T. Abbott, M. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. Adhikari, et al. Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102, 2016.

[2] R. Abbott, T. Abbott, S. Abraham, F. Acernese, K. Ackley, C. Adams, R. Adhikari, V. Adya, C. Affeldt, M. Agathos, et al. GW190521: a binary black hole merger with a total mass of 150 $M_{\odot}$. *Physical review letters*, 125(10):101102, 2020.

[3] S. Mallat. *A Wavelet Tour of Signal Processing: The Sparse Way.* Academic Press, Inc., 2008.

[4] M. Priestley. *Spectral Analysis and Time Series, Volume 1.* Academic Press, 1981.