A unified wavefield-partitioning approach for distributed acoustic sensing

James Atterholt, Zhongwen Zhan, Zhichao Shen and Zefeng Li[*]

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, USA. E-mail: atterholt@caltech.edu

SUMMARY

While distributed acoustic sensing (DAS) has been demonstrated to have great potential in seismology, DAS data often have much higher levels of stochastic and coherent noise (e.g. instrument noise, traffic vibrations) than data collected by traditional seismometers. The linearly, densely spaced nature of DAS arrays presents a suite of opportunities for more innovative processing techniques that can be used to address this issue. One way to take advantage of DAS's array architecture is through the use of curvelets. Curvelets have a non-uniform scaling property that makes them an excellent tool for representing images with discontinuities along piecewise, twice continuously differentiable curves. This anisotropic scaling property makes curvelets an ideal processing tool for DAS data, for which the measured wavefield can be represented as an image composed of curved features. Here, we use the curvelet frame as a tool for the manipulation of DAS signal and demonstrate how this manipulation can improve our ability to identify important features in DAS data sets. We use the curvelet representation to partition the measured wavefield using DAS data collected near Ridgecrest, CA, following the 2019 M_w7.1 Ridgecrest earthquake. Here, we isolate the earthquake-induced wavefield from coherent and stochastic noise using the curvelet frame in an effort to improve the results of template matching of the Ridgecrest aftershock sequence. We show that our wavefield-partitioning technique facilitates the identification of over 30 per cent more aftershocks and greatly reduces the magnitude of diurnal depressions in the aftershock catalogue due to cultural noise.

Key words: Instrumental noise; Image processing; Time-series analysis; Seismicity and tectonics.

1 INTRODUCTION

Distributed acoustic sensing (DAS) is a new tool in seismology that repurposes fibre optic cables as arrays of densely spaced strainmeters (see Hartog 2017, for a review). DAS employs a laser interrogator unit, which sends short laser pulses through optical fibres and performs interferometry to measure phase shifts in the Rayleigh backscattered light. These phase shifts are quasi-linearly proportional to strain or strain rate in the fibre. This strain is averaged over a 'gauge length' and sampled at discrete intervals along the fibre that constitute the so-called channels of the array. Laser pulses are sent at regular intervals to measure the temporal dependence of strain in the fibre.

DAS has the potential to transform seismology because of its dense spatial sampling and its capacity to transform *in situ* fibres

into seismic arrays (see Zhan 2019, for a review). DAS functions as a large-*N* array of instruments, and because of this high spatial sampling, DAS may allow seismologists to probe the subsurface at resolutions that have historically been limited to expensive exploration experiments (Dou *et al.* 2017; Ajo-Franklin *et al.* 2019). Pre-existing fibre optic cables in place for telecommunications can be easily re-purposed as DAS arrays. DAS's capacity to use pre-existing fibre allows seismologists to both deploy seismic arrays faster and extend seismology, particularly dense array seismology, to logistically challenging locations of immense societal and scientific interest (Lellouch *et al.* 2019; Lindsey *et al.* 2019; Sladen *et al.* 2019; Booth *et al.* 2020; Spica *et al.* 2020).

However, DAS remains an emerging technology, and there still exist fundamental challenges in the acquisition and analysis of DAS data. One such challenge is noise, both stochastic and coherent. Stochastic noise in DAS signals is likely dominated by instrumental deficiencies like sampling error and phase noise and is bounded by the Crámer–Rao lower bound (Costa *et al.* 2019). Coherent noise can be defined as the coherent components of the DAS wavefield

that are not interesting to the end user. In many cases, particularly for seismologists, traffic noise is a persistent source of coherent noise. Traffic noise is especially problematic for DAS because pre-existing fibre optic cables are often placed along major roads. Both stochastic and coherent noise are problematic because they can either mask or obscure relevant signal. This obfuscation is exemplified in Fig. 1(b), which shows a case in which cultural noise masks much of the signal generated by a local earthquake. Fig. 1(b) also shows that even for a relatively quiet station along the DAS array, stochastic noise is much stronger in the DAS station than in a nearby broad-band station. DAS stochastic noise is in large part instrument related, and the noise level in DAS will likely vary between generations and decrease as technology improves. However, since DAS signal degrades with distance, stochastic noise may persist as a challenge at the far end of the DAS array for many years to come.

Previous efforts to remove stochastic noise from DAS data have been varied and proposed both outside of and within seismology. Many of these efforts have successfully applied time–space analysis techniques from signal processing such as wavelet transforms (Qin *et al.* 2012), 2-D edge detection (Zhu *et al.* 2013), 2-D bilateral filters (He *et al.* 2017), empirical mode decomposition (Qin *et al.* 2017b) and principal component analysis (Ibrahim *et al.* 2020). In particular, Qin *et al.* (2017a) proposed an approach to remove random noise in the curvelet domain. They recognized that the curvelet domain, much like the wavelet domain, sparsely represents DAS data, and they perform thresholding of curvelet coefficients below a certain magnitude to remove random noise. We build on this approach in this study. More recently, van den Ende *et al.* (2021) proposed a deep learning approach for removing incoherent signal with a focus on DAS data.

There has been less discussion of removing coherent noise from DAS, partly because coherent noise is not well defined. The difference between coherent noise and the signal of interest depends on how we define coherent noise and what the signal of interest is. Usually, signals of different types can be distinguished by one or more physical characteristics. For the seismic wavefield, signals can be distinguished using the various velocity attributes of seismic waves. We can exploit this quality by transforming our DAS data from the spatiotemporal domain to the frequency–wavenumber domain, where our signals are localized by velocity. Williams *et al.* (2019) illustrate this point by localizing seismic waves from ocean waves recorded by a submarine DAS fibre using the frequency–wavenumber domain. It follows that if the coherent noise is localized by velocity, then muting the velocity range in the frequency–wavenumber domain associated with the coherent noise removes the coherent noise from the signal. This method of denoising has been used in exploration seismology for decades (Embree *et al.* 1963) and has been applied as a pre-processing step to DAS data (Fang *et al.* 2020). Though frequency–wavenumber domain filtering is commonly used, other tools, such as the continuous wavelet transform, have been used to diminish coherent noise in DAS data as well (Martin *et al.* 2018).

With DAS, we encounter stochastic noise and coherent noise together. We argue that a simple, unified approach to noise reduction in DAS would be useful. Moreover, given the large volumes of DAS data that seismologists are prone to process, this filtering approach must be scalable such that it can be applied to terabytes of data in a reasonable amount of time. In this paper, we propose combining some of the principles described in this introduction into a unified approach to isolate a signal of interest in seismological data by wavefield partitioning entirely under the curvelet transform. We then illustrate the efficacy of this approach by applying it as a

pre-processing step for template matching applied to a subset of DAS data recording the M_w7.1 Ridgecrest earthquake aftershock sequence.

2 METHODOLOGY

2.1 Curvelets

Curvelets were described in a continuous setting (Candés & Donoho 2004) as an almost optimal representation of images with discontinuities along twice continuously differentiable (C^2) curves. Candés *et al.* (2006) developed the fast discrete curvelet transform (FDCT), making curvelets easily applicable in image processing. This transform has since been used widely in a number of fields (see Ma & Plonka 2010, for examples). Curvelets have been evaluated by how well they represent a 2-D object, F, that has a discontinuity along a curve, but is otherwise smooth, using an m-term approximation of the object, F_m. It has been shown that for such objects the mean squared error of the curvelet frame representation decays with $m^{-2}[\log(m)]^3$, whereas the mean squared error of the wavelet representation decays with m^{-1} (Candés & Donoho 2004). We claim that DAS data sections measuring the seismic wavefield can be approximated as smooth images with discontinuities along C^2 curves because DAS data are marked by periods of quiescence (smoothness) interrupted by wavefronts (curvature). If this is a good approximation, the curvelet frame is the best available non-adaptive sparse representation of DAS data. In this study, we perform denoising of DAS data through applications of the FDCT on DAS record sections of finite length in time and space (e.g. Fig. 1b).

Following closely after Candés *et al.* (2006), curvelets are constructed by first creating a polar tiling in the frequency–wavenumber domain. Specifically, we take the 2-D Fourier transform of our function and compartmentalize the resultant function using special tiling geometry. Letting (ω, k) be a coordinate in the frequency–wavenumber domain and letting $r = \sqrt{\omega^2 + k^2}$ and $\theta = \arctan(\omega/k)$, we can construct the polar tiling using windowing functions:

$$U_j(r, \theta) = 2^{-\frac{3j}{4}} W\left(2^{-j}r\right) V\left(\frac{2^{\lfloor \frac{j}{2} \rfloor}\theta}{2\pi}\right), \tag{1}$$

where $U_j(r,\theta)$ is a polar wedge corresponding to radius r, orientation θ, and scale integer j. Here, W is a window function along the radius and V is a window function along the orientation. Importantly, the scale, given by 2^{-j}, is inversely related to the scale integer term. Note, then, that the radius is dilated by 2 for each jump in scale, and the number of wedges increases by a factor of 2 for every two jumps in scale. This prescribes a non-uniform scaling that results in curvelets becoming more needle-like with finer scales. The spacing of r and θ are thus dictated by the number of scales and the number of wedges at the coarsest scale. To be clear, tiles further away from the origin of the frequency–wavenumber domain correspond to smaller and more needle-like curvelets in the spatiotemporal domain. This relationship is illustrated in Fig. 2(b). The intuitive motivation for this non-uniform scaling rests in the observation that smooth curves appear more linear when viewed from up close, and thus needle-like objects can capture the sharp edges of smooth curves if the needles are sufficiently small.

Now, a so-called 'mother' curvelet, $\phi_j(x, t)$, with scale integer j and coordinate (x, t) in the space–time domain, is defined by taking the inverse Fourier transform of the polar wedge:

$$\phi_j(x, t) = \mathcal{F}^{-1}[U_j(\omega, k)], \tag{2}$$

where $U_j(\omega, k)$ is the wedge described in eq. (1) in Cartesian coordinates in the frequency–wavenumber domain. This term, $\phi_j(x, t)$, is called the mother curvelet because all curvelets of scale 2^{-j} can be defined as some rotation and translation of $\phi_j(x, t)$. Indeed, with a sequence of rotation angles defined as $\theta_l = 2\pi \cdot 2^{\lfloor -j/2 \rfloor} \cdot l$, where $l = 0, 1, 2, ...$ satisfying $\theta_l \in [0, 2\pi)$ and translation parameters defined as $k = (k_1, k_2)$, all curvelets can be defined as

$$\phi_{j,l,k}(x, t) = \phi_j(x, t) \cdot R_{\theta_l} \cdot \left((x, t) - (x_{k_1}, t_{k_2})^{(j,l)} \right), \qquad (3)$$

where R_{θ_l} is a standard rotation matrix and $(x_{k_1}, t_{k_2})^{(j,l)}$ is the position prescribed by $R_{\theta_l}^{-1}(k_1 \cdot 2^{-j}, k_2 \cdot 2^{-j})$. Finally, the curvelet coefficients, the values of which we will use in the subsequent denoising procedure, are defined using the scalar product of the curvelets with an arbitrary function $f \in L^2(\mathbf{R}^2)$:

$$c_{j,k,l} = \langle f, \phi_{j,k,l} \rangle = \iint f(x, t) \cdot \overline{\phi_{j,k,l}(x, t)} \mathrm{d}x \, \mathrm{d}t = \frac{1}{(2\pi)^2} \iint$$
$$f(\omega, k) \cdot \overline{\phi_{j,k,l}(\omega, k)} \mathrm{d}\omega \mathrm{d}k, \qquad (4)$$

where $\overline{\phi_{j,k,l}}$ is the conjugate of $\phi_{j,k,l}$.

In words, these curvelets form a set of basis functions that, when weighted by curvelet coefficients, can represent an arbitrary smooth function like the seismic wavefield in space and time. However, the antecedent equations describe the curvelet construction for the continuous case, and seismic data are collected at discrete intervals. So, in practice, a few modifications to this methodology are necessary to make this continuous transformation applicable to discrete data. One modification is that the polar wedges, defined in eq. (1), are necessarily computed as so-called Cartesian shears. An illustration of the compartmentalization of the frequency–wavenumber domain into Cartesian shears is given in Fig. 2(a). The construction of these shears requires thoughtful considerations of geometry and windowing that are given a more complete treatment in Ma & Plonka (2010). In short, these windows are real, positive, smooth functions that localize the frequency–wavenumber domain into a tiling of trapezoidal shears and allow the resultant curvelets to form a tight frame. Additionally, these shears are not rectangular, and
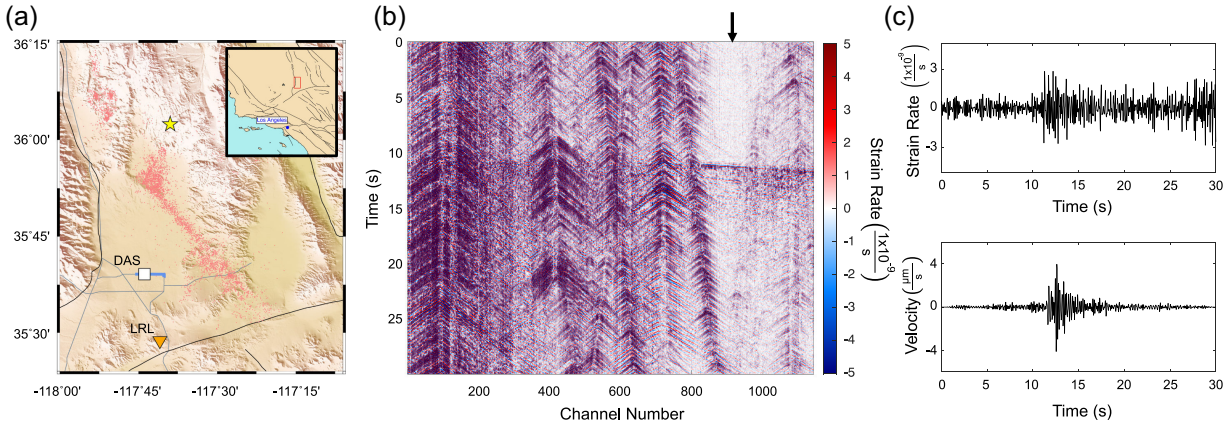

Figure 1. (a) Map showing the geographic setting for the data segments in this figure. The yellow star indicates the location of the M_l 1.79 earthquake recorded by the Ridgecrest array and the LRL station at distances of 44 km (specifically the station highlighted in panel c) and 62 km, respectively. The blue line indicates the location of the DAS array that recorded the time-series shown in the middle image of this figure. The white box indicates the approximate location of the DAS station whose individual station data are shown in this figure. The orange triangle shows the location of the broad-band station (LRL) whose station data are shown in this figure. The black and grey lines delineate faults and roads, respectively. The light red dots indicate the locations of aftershocks used as templates in this study. The inset gives regional context. (b) Full DAS record section at the time of a regional earthquake, the location of which is shown in the map in panel (a). The black arrow marks the location of the DAS channel given as an example in panel (c) (channel 899). (c) Comparison of station data from a single DAS station (top) with the east–west component of station data from the nearby LRL broad-band station (bottom). The DAS stations in panels (b) and (c) and the broad-band station in panel (c) are bandpass filtered from 2 to 8 Hz using a Butterworth filter.
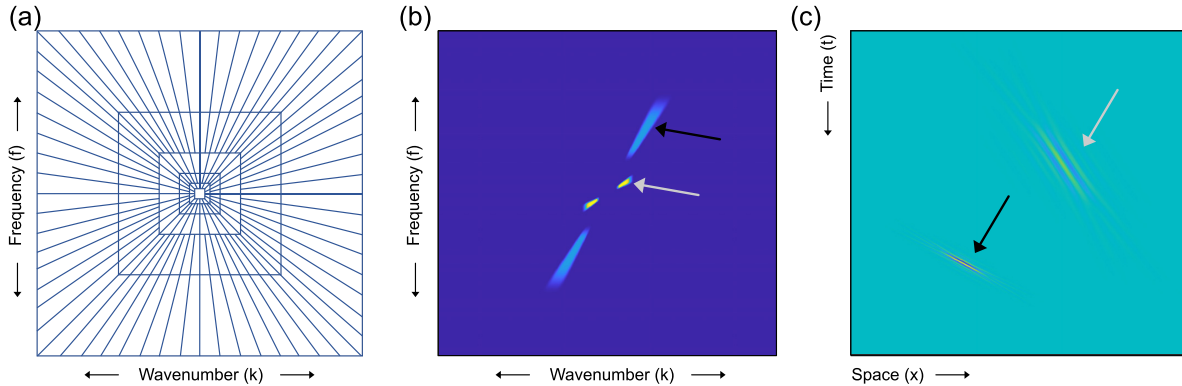

Figure 2. (a) Schematic illustration of the polar tiling used to construct curvelet coefficients. The wedges delineate the compartmentalization of the frequency–wavenumber domain. (b) Plot of select curvelets in the frequency–wavenumber domain. (c) Plot of the same curvelets as in panel (b) in the spatiotemporal domain. Colours of arrows correspond to equivalent curvelets between panels (b) and (c).

so the necessary inverse fast Fourier transform (FFT) for the discrete case is impossible to compute. To get around this problem, the FDCT employs a 'wrapping' operation; that is, the curvelets are duplicated many times and situated adjacent to each other over the frequency–wavenumber domain. Then, the inverse FFT is taken on a rectangular section centred on the origin in this domain (see section 3.3 of Candés *et al.* 2006, for details). Curvelets of different scales and orientations for the discrete case are shown in both the frequency–wavenumber and spatiotemporal domains in Fig. 2(b). As mentioned previously, the spacing of r and θ are determined by the number of scales and the number of tiles at the coarsest scale. These parameters are user-defined in the FDCT according to the data matrix sampling and size.

2.2 Stochastic noise removal

The curvelet frame has previously been used to remove random noise from seismograms collected using traditional seismic arrays and has even been extended to non-uniformly sampled arrays (Hennenfent & Herrmann 2006). Denoising with curvelets is often accomplished by producing white noise images with some assumed variance, performing FDCT transforms on these images to establish scale-dependent thresholds, and applying either hard or soft thresholding using these thresholds (Starck *et al.* 2002). This method is imperfect because it requires an estimation of the variance and assumes that the noise is white.

With DAS seismic data, we benefit from having continuously recorded data over long periods of time. In these long time-series, there are usually time windows in which there is minimal coherent signal. These time periods can be used as benchmarks for the level of incoherent noise throughout the time-series. So, in this study, we chose wedge-dependent denoising thresholds using the distribution of curvelet coefficient amplitudes in quiet sections of the time-series. Implicit in this selection of thresholds is the assumption that the stochastic noise levels are independent of time. That is, we assume that the noise in a quiet time segment will be representative of the noise throughout the data. This is not always true, as we expect the noise in DAS to fluctuate with time according to environmental factors like temperature variability. However, we still favour a uniform application of thresholding because non-uniform applications are generally less stable and can produce artefacts like step discontinuities in the noise floor of denoised data. Because we intend to apply this technique to more data than can be visually scrutinized, we consider these potential issues to be unacceptable in this workflow. We also choose thresholds that are not dependent on translational parameter k. This means that we ignore any spatial and temporal variability of the noise floor *within* the quiet time segment. This is a reasonable assumption because we consider the noise floor to be largely reflective of the DAS interrogation unit, which is shared by all channels in the array, and because including a temporal variability within each time segment would suggest that the noise floor is periodic according to the arbitrarily chosen time segment duration.

When choosing a thresholding technique, one often chooses between hard thresholding and soft thresholding. Here, we chose soft thresholding, because it minimizes noise that barely exceeds the threshold and prevents the generation of artefacts that result from preserving high-amplitude noise curvelet coefficients and zeroing their neighbours. Soft thresholding is defined as

$$c_{j,k,l} = \begin{cases} 0 & \text{if } |c_{j,k,l}| < \tau_{j,l} \\ c_{j,k,l} - \text{sign}\,[c_{j,k,l}] \cdot \tau_{j,l} & \text{if } |c_{j,k,l}| > \tau_{j,l} \end{cases} \qquad (5)$$

where $c_{j,k,l}$ is a given curvelet coefficient and $\tau_{j,l}$ is its associated threshold. We show an example of applying soft thresholding using empirical thresholds defined using a pure stochastic noise window in Fig. 3. Soft thresholding systematically diminishes the observed amplitude of the seismic wavefield. This is acceptable in the coming example, but for amplitude dependent analyses, hard thresholding may be preferable.

2.3 Coherent noise removal

As described earlier, coherent noise can often be localized from interesting signal using velocity contrasts in the frequency–wavenumber domain. The curvelet frame, by construction, creates a polar tiling in the frequency–wavenumber domain that finely compartmentalizes the wavefield by velocity and scale. Naturally, then, we can exploit this compartmentalization to filter out coherent noise under the curvelet transform. This can be done by simply muting the wedges under the curvelet transform that contain the part of the wavefield associated with the coherent noise. The curvelet transform has been used in array seismology for similar purposes in the past, namely to improve measurements of *SS* precursors by eliminating interfering phases (Yu *et al.* 2017) and to help isolate the scattered wavefield from teleseismic P waves (Zhang *et al.* 2020).

We can justify our decision to perform velocity filtering under the curvelet transform in a few ways. First, it is convenient to perform coherent noise filtering under the same transform with which we perform stochastic noise filtering. Additionally, this filtering procedure is straightforward and useful for experimentation. The Cartesian shear framework provides a convenient basis with which to identify and remove wavefield components in velocity-scale space. Yet another reason is that the windowing functions used by the FDCT employ tapers that are effective at removing large artefacts produced by the creation of discontinuities in the frequency–wavenumber domain. We show this in Fig. 4, where we evaluate the removal of coherent noise using real data. These tapers are also highly localized in the angular direction, yielding a precise separation of velocities.

2.4 Unified approach

Since the methods described for removing stochastic noise and coherent noise are both performed under the same transform, we can remove both of these noise sources in the same step. This procedure simply involves transforming from the spatiotemporal domain to the curvelet domain under the FDCT, performing thresholding using thresholds determined by a noise window to remove stochastic noise, muting the curvelet wedges that are associated with unwanted signal, and finally performing the inverse FDCT to return to the spatiotemporal domain. If the threshold values are fixed, then the order of the second and third steps does not matter. Performing these procedures together simplifies pre-processing and lowers computational costs by avoiding additional transforms. The simplicity argument for this approach should not be undervalued. Rather than requiring a set of arbitrary signal processing decisions, our methodology effectively offers a 'dial' to turn down stochastic noise and a 'switch' to turn off coherent signal by velocity and scale. The computational cost reduction may vary, but if one were to perform stochastic noise filtering under the curvelet transform and coherent noise filtering under a second transform, the cost saved would amount to the cost of the forward and inverse computations
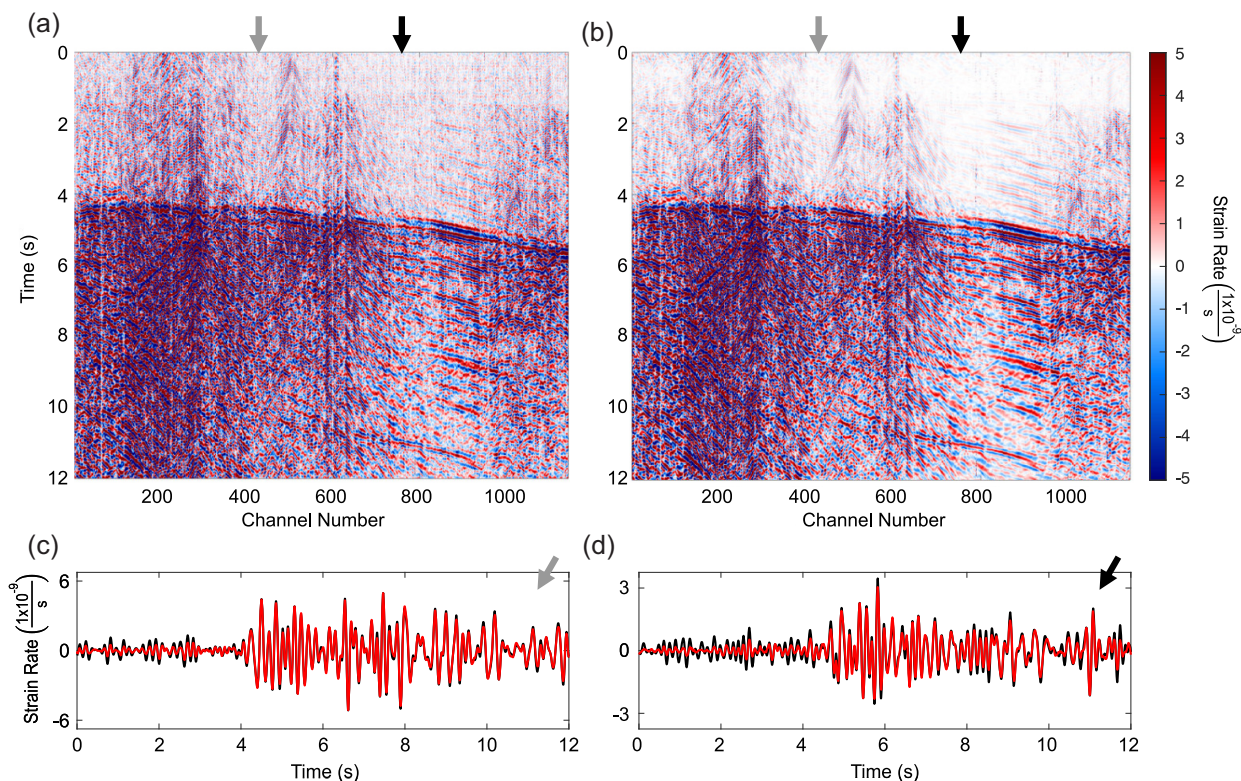
Figure 3. Illustration of stochastic denoising using soft-thresholding in the curvelet domain. Thresholds for this example were determined empirically using a window with little coherent signal from the same data set. (a) Unfiltered record sections for the entire DAS array. (b) Curvelet filtered record sections for the entire DAS array. (c) The individual DAS station data from the record sections in panels (a) (black) and (b) (red) for the DAS station marked by the grey (channel 422) arrows. (d) The individual DAS station data from the record sections in panels (a) (black) and (b) (red) for the DAS station marked by the black (channel 780) arrows.
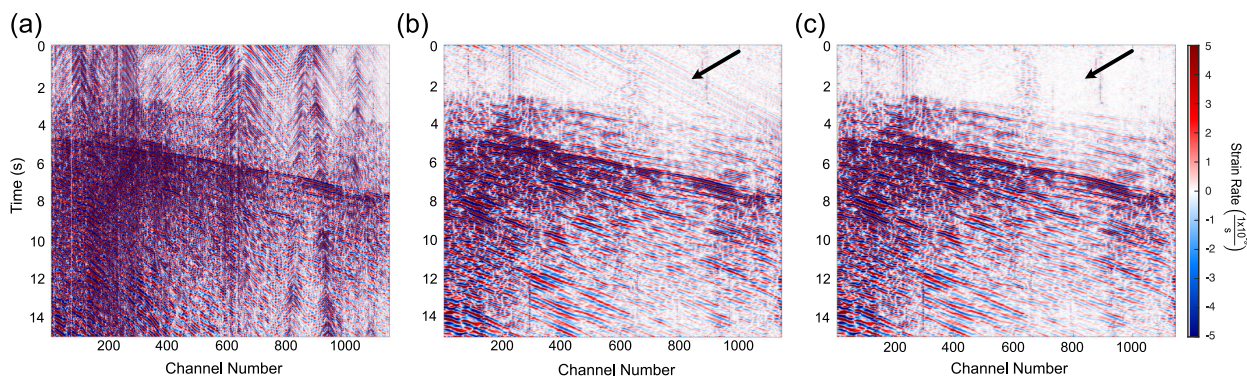

Figure 4. Illustration of coherent denoising using the polar-tiling framework of the curvelet construction. (a) Pseudo-synthetic noisy window made by superimposing a record section with high cultural noise onto a record section with an earthquake. (b) Time-series in which velocities in the range 0–1000 m s^{-1} were removed using an untapered FK filter. (c) Time-series in which velocities in the range 0–1000 m s^{-1} were removed using our curvelet windowing technique. The black arrows point to a region of prominent filter-generated artefacts.

3 APPLICATION

3.1 Data

Following the M_w 7.1 Ridgecrest earthquake on 2019 July 5, an OptaSense ODH3 interrogator unit was used to deploy a DAS array in the town of Ridgecrest, CA, near the epicentre of the event. This array samples at 250 Hz and has 1250 channels with a channel spacing of 8 m for an overall cable length of 10 km. The array continuously recorded much of the aftershock sequence. We choose this array both because it has recorded many earthquakes, and because it is linear. The linearity of the array ensures that the apparent velocities of incoming waves are mostly constant across the entire array. A linear array geometry is not strictly necessary to perform the curvelet filtering approach outlined above, but the linear

geometry ensures that the earthquake wavefield is more localized in the frequency–wavenumber domain. The location and aperture of the DAS array is shown in Fig. 1(a).

3.2 Template matching

In order to demonstrate the efficacy of our filtering approach, we apply it to the DAS data prior to a template matching procedure to better illuminate small earthquakes masked by noise. Template matching is a technique that correlates known events, known as templates, with continuous time-series to detect previously unknown events that are similar to the templates (Gibbons & Ringdal 2006). Li & Zhan (2018) showed that template matching could be successfully applied to DAS data. This technique has been modified to perform template matching using the Ridgecrest DAS array to supplement the Southern California Seismic Network's (SCSN) catalogue of Ridgecrest earthquake aftershocks (Li *et al.* 2021). Though this catalogue is successful in improving the number of catalogued events, the template matching catalogue contains diurnal depressions in the number of events due to cultural noise. Additionally, the completeness of this catalogue may be limited by the noise floor of the Ridgecrest DAS array. We seek to show that these limitations can be mitigated by applying our unified curvelet filtering framework to Ridgecrest DAS array data.

Our filtering framework includes a few pre-processing steps. We first remove the stations of the array that are either coiled or too close to the interrogator unit. We then segment these array data into 60-s time windows recorded at the original sampling rate by the 1142 DAS stations that remain after the station removal step. The temporal and spatial sampling rates dictate the velocities associated with Cartesian shears under the curvelet domain. The number of samples along the time and space axes control the frequency and wavenumber resolution respectively, which in turn dictates the precision of the velocity filtering. We segment these array data to ensure that each allocation of memory does not exceed a few gigabytes, but 60-s segments are long enough so as to ensure we have sufficient frequency resolution for velocity filtering. Segmenting these data also allows us to filter in parallel. We applied a Tukey window along the time axis to minimize artefacts due to discontinuities at the start and end of each segment. Time segments were staggered such that we could clip the tapered portions of the time segments when constructing the final time-series. For each time segment, we applied a median filter to remove a source of optical noise, not handled by the curvelet filter, that results in random, high-strain spikes in the data (Bakku 2015). Failure to remove these spikes before filtering results in the creation of star-like artefacts after filtering. Though median filtering mitigates this issue, these artefacts are often unavoidable. An adaptive spectrum screening algorithm can potentially help better separate earthquake signals from coherent noise, but for consistency with the original template matching framework, we bandpass filter each station between 2 and 8 Hz (Li *et al.* 2021).

For each time segment, we applied our unified curvelet filter procedure described above. For the stochastic noise removal, we used a representative noise window to establish thresholds for the entire time-series. We opted to use a single representative noise window rather than multiple noise windows throughout the time-series, because using multiple noise windows produces the undesirable side effect of discontinuities in the noise floor. After applying these thresholds, we muted wedges associated with apparent velocities between 0 and 1000 m s^{-1} in both the east-going and west-going

directions at scales corresponding to the seismic wavefield. We then took the inverse FDCT and removed the tapered portions of the time segment. We then placed the filtered time segment in its appropriate position in the final, filtered time-series.

The templates used in the template matching algorithm were taken directly from the data. So, once we had filtered the data, we had also filtered the corresponding templates. We found template candidates using the SCSN catalogue and determined which template candidates to keep using a criterion that requires at least a set number of stations to exceed a fixed signal-to-noise ratio (SNR). Because the curvelet denoising improves the SNR of the data, the curvelet filtered data produced many more templates that met our criteria. To be fair, we only used templates that the filtered and unfiltered data had in common. We then performed template matching on both our filtered and unfiltered time-series. Processing 192 hr of DAS data (2019 July 14–21) using our intentionally redundant methodology on 24 CPUs took just over 19 hr. Both the catalogue produced before filtering and the catalogue produced after filtering are plotted, by number of events per hour, in Fig. 5. The template matching catalogue corresponding to the unfiltered data contains 20 935 events while the template matching catalogue corresponding to the filtered data contains 28 044 events. There exists a diurnal pattern in the difference between the filtered and unfiltered catalogues, with the difference reaching peaks during the daytime when traffic noise is highest. In particular, we see a 56 per cent increase in number of events during the daytime (7 a.m. to 7 p.m.) and a 21 per cent increase in number of events during the night-time (7 p.m. to 7 a.m.).

To confidently identify the cause of this diurnal effect, we applied this template matching procedure to the same DAS data filtered using only the coherent noise filter and only the stochastic noise filter. These hourly catalogues are shown in Supporting Information Fig. S1. This diurnal pattern in the difference between the filtered and unfiltered catalogues is amplified when only the coherent noise is removed while this diurnal pattern is non-existent when only the stochastic noise is filtered. This observation suggests that the improvement made by the coherent velocity filter is largely due to the removal of traffic-generated surface waves. Interestingly, performing only the stochastic noise filtering results in a catalogue with slightly fewer events (19 826 events) than the catalogue produced using unfiltered data (20 935 events); incongruously, performing the stochastic and coherent noise filtering together results in a catalogue with many more events (28 044 events) than the catalogue produced using data filtered using only the coherent noise filtering (25 731 events). This observation may be explained by noting that the stochastic noise filter will amplify both the earthquake waveforms and the coherent noise (e.g. traffic noise), and so the stochastic noise filtering in the presence of a large amount of coherent noise may be detrimental to the template matching procedure and stochastic noise filtering in the presence of little coherent noise is beneficial to the template matching procedure.

4 DISCUSSION

Clearly, the curvelet filtering improved the performance of the template matching algorithm. One noteworthy feature of Fig. 5 is the reduction in the magnitude of the diurnal depressions in the aftershock catalogue. This is largely due to the removal of the cultural noise (mostly surface waves generated by vehicles), which is far more prevalent during the daytime, by the coherent noise filter. Because the apparent velocities of the earthquake-generated surface
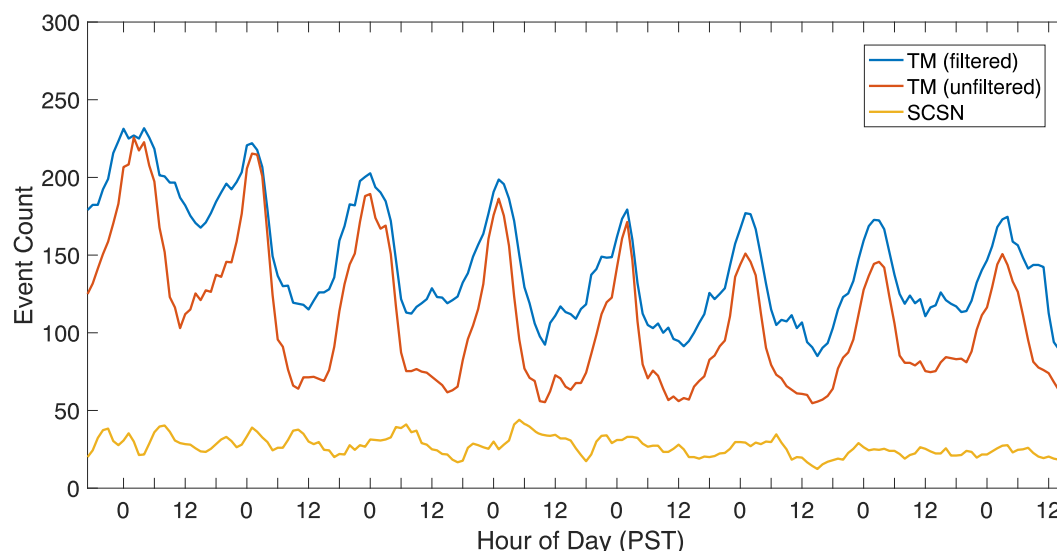
Figure 5. Hourly (PST) event counts of the catalogues created using template matching. A 3-hr moving average filter was applied to each of the time-series. The yellow line shows the hourly counts for the catalogue of templates made using the SCSN catalogue. The red line shows the hourly counts for the catalogue created using template matching (TM) on the unfiltered data. The blue line shows the hourly counts for the catalogue created using TM on the curvelet filtered data.

waves also often fall in our filtered velocity range, earthquake signals in our filtered time-series are often degenerate. We argue this loss of information is well compensated by the remarkable illumination of the earthquake-generated body waves. As is clear in Fig. 5, at no point did we see a decrease in the number of catalogued events due to filtering, suggesting that the information recovered by filtering outweighed the information lost. Given that DAS is more sensitive to the low-velocity wavefield, earthquake-generated surface waves may prove useful in the detection of earthquakes in the future (Lior *et al.* 2021). However, traffic noise is also composed of largely low-velocity surface waves and is thus especially prominent in DAS data, making the removal of this noise essential.

To evaluate the amount of overlap between catalogues, we compute the proportion of events in the original template matching catalogue that were also detected in the new template matching catalogue made using filtered data. Because the templates are altered in the filtering process, the same event may be matched with different templates between catalogues. This possibility adds significant uncertainty to any measure of overlap between catalogues. We choose to measure this overlap using the metric that if for any given event in the original catalogue, there is a corresponding event in the new catalogue with an origin time within 2.5 s of that of the given event, then the given event is represented in both catalogues. By this metric, 73 per cent of the events in the original catalogue were carried over to the new catalogue. Indeed, this metric is conservative; if we extend the time window about the origin to 5 s, then we find that 83 per cent of the events in the original catalogue are represented in the new catalogue. Additionally, the fact that some events are not represented in the new catalogue does not necessarily represent a failure of the filtering methodology, but could be indicative of a success. Cultural noise is a consistent source of coherent signal that produces observables that are frequently almost replicated at different times. This noise can potentially produce false detections in the original catalogue that will not be present in the new catalogue, for which cultural noise has been largely removed. This removal of false detections is partially evidenced by the diurnal pattern shown in the plot of the per cent of events retained between catalogues by hour

shown in Supporting Information Fig. S2, which suggests that the percentage of events retained between catalogues is lowest during a large portion of the population's morning commute. Further evidence that the detections lost from the original catalogue to the new catalogue are false detections is given by the observation that most of these detections were made using noisy templates as opposed to quiet templates. These noisy templates thus provide a substantial coherent noise wavefield that can produce cross-correlation peaks with the coherent noise in the data. An example of one of these suspected false detections is shown in Supporting Information Fig. S3.

Our results summarized in Fig. 5 and Supporting Information Fig. S1 suggest that, in combination with the coherent noise filtering, the stochastic noise filtering produced a considerable improvement in the number of catalogued events. This improvement likely results from the fact that for a template matching algorithm to label a detection, the cross-correlation between the template and the time-series must produce a peak that exceeds some significance threshold. It is easier, then, for an event's correlation to exceed a significance threshold when the baseline noise is lower. Additionally, in the absence of coherent noise filtering, the stochastic noise filtering produced a slight decrease in the number of catalogued events. This decrease may potentially be explained by the fact that the stochastic noise filtering increases the SNR of all types of coherent signal, thus increasing the prominence of coherent noise in the data and making event detection more challenging. We suggest that this explanation for the loss of detections is more plausible than an explanation of simple overfiltering, because there exists a diurnal pattern in the loss of detections (we see a 9 per cent decrease in the number of events between 7 a.m. and 7 p.m. and only a 3 per cent decrease in detections between 7 p.m. and 7 a.m.). This pattern runs counter to our expectation for overfiltering, for which we would expect overfiltering to present more of a problem during the nighttime, because the detection threshold for template matching is much lower when coherent noise is lower.

Though this curvelet filtering methodology performs very well for this problem and is easily generalizable to other problems, it

has some limitations. Many of these limitations are a result of the imperfect nature of DAS as a measuring instrument. Indeed, curvelets are exceptional at representing smooth discontinuities, but not rough discontinuities. In reality, DAS data exhibit many rough discontinuities such as random spikes of high strain that act as localized discontinuities. We also note that, for high-strain events like earthquakes and large vehicles driving near the fibre, strain rates are so high that they produce DAS phase errors that result in additional discontinuities. Optical fading, a spatially random effect that results from the destructive interference of scatterer-generated electric fields (Zhou *et al.* 2013), produces muted stations that, when near unmuted stations, act as discontinuities. A similar effect can be produced by variability in the degree of coupling between the DAS cable and the ground. These discontinuities are not well represented by curvelets and produce artefacts of varying severity under our filtering procedure.

We also note the importance of array geometry in the practicability of our methodology. The degree of localization of velocity dictates whether or not two signals can be separated on the basis of velocity. For example, a plane wave recorded by a complex DAS array geometry may be spread over a large range of velocities because of a high variability in the angle of incidence at different stations along the array. Traffic-generated surface waves are conducive to this technique, however, because DAS cables often run parallel to roads. In this case, most of the traffic-generated surface wave signal is recorded at near true velocity. This is the case in this study, but with a not insignificant caveat. Because the cable and the road are separated by a small distance, for a short section of the array directly next to passing vehicles, traffic-generated surface waves are recorded with very high velocities. This results in systematic imperfections in the coherent noise removal in this study.

In this study, we apply stochastic and coherent noise removal together to illustrate a unified approach to noise removal that is convenient, scalable and effective. However, there is no hinderance to applying either the stochastic or coherent noise filtering independently. Applying the stochastic noise filtering by itself is easily justifiable, because the curvelet representation is, to our knowledge, the best available, non-adaptive sparse representation of objects with smooth discontinuities. Applying the coherent noise filtering by itself is effective, but not more effective than traditional frequency–wavenumber filters with specialized tapers. The choice to perform only coherent noise filtering under the curvelet transform may be made out of convenience, but the flexibility of this filter is limited by the Cartesian coronae framework. Performing only coherent noise filtering using the curvelet domain can be made considerably faster by creating a frequency–wavenumber mask under the curvelet transform. Such a mask can be made by performing coherent noise filtering on a 2-D delta function under the curvelet transform.

5 CONCLUSIONS

Here, we proposed a unified wavefield-partitioning approach for DAS data in a seismological context using curvelets. One component of our technique involves the application of thresholding to curvelet coefficients using data-specific thresholds under the curvelet transformation. Another component of our technique requires the implementation of scale-dependent velocity filtering under the curvelet transform. By using these partitioning techniques together under the curvelet transform, we demonstrated that stochastic and coherent noise can be removed in a single step. We showed that this approach is effective and scalable by applying this filtering

technique to 8 d of DAS data collected following the M_w7.1 Ridgecrest earthquake. With a simple application of our filtering technique, we greatly improved a template-matching-generated earthquake catalogue, increasing the number of detected events by over 30 per cent and reducing the magnitude of diurnal depressions in the catalogue due to cultural noise.

With its high spatial sampling and logistical advantages, DAS presents numerous opportunities to advance the field of seismology. However, more so than for other seismic instrumentation, noise in DAS will persist in limiting how well we can seize these opportunities. In the coming years, we may expect improvements in DAS instrumentation to lower the instrument noise floor, but we should continue searching for pre-processing steps that help us get the most out of our data. Here, we have outlined a methodology to remove noise sources from DAS data in a seismological context using a basis that we think is particularly well suited for DAS data. However, our approach is non-adaptive, and we might expect to see significant improvements in coherent noise removal from adaptive techniques targeted to seismological applications. In particular, learning the structure of DAS data, using techniques like machine learning, will likely allow for more precise coherent noise removal. We see the framework of our methodology and the curvelet basis as potentially useful components of more effective adaptive filtering techniques that may be developed in the foreseeable future.

DATA AVAILABILITY

Earthquake record sections used in Figs 1, 3, and 4 are available for download at Caltech's research data repository (https://data.caltech.edu/records/1955). Code for performing this wavefield-partitioning technique with a working example is available on Github (https://github.com/atterholt/curvelet-denoising). This code makes use of the CurveLab toolbox that is available on the curvelet.org website (http://www.curvelet.org).

REFERENCES

Ajo-Franklin, J., Dou, S., Lindsey, N., Monga, I., Tracy, C., Robertson, M. & Li, X., 2019. Distributed acoustic sensing using dark fiber for near-surface characterization and broadband seismic event detection, *Sci. Rep.,* **9,** 1–14.

Bakku, S., 2015. Fracture characterization from seismic measurements in a borehole, *PhD thesis*, Massachusetts Institute of Technology.

Booth, A. *et al.*, 2020. Distributed acoustic sensing of seismic properties in a borehole drilled on a fast-flowing Greenlandic outlet glacier, *Geophys. Res. Lett.,* **47,** e2020GL088148, doi:10.1029/2020GL088148.

Candés, E. & Donoho, D., 2004. New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, *Commun. Pure Appl. Math.,* **28,** 219–266.

Candés, E., Demanet, L., Donoho, D. & Ying, L., 2006. Fast discrete curvelet transforms, *Multiscale Model. Simul.,* **5,** 861–899.

Costa, L., Martins, H., Martín-López, S., Fernández-Ruiz, M. & González-Herráez, M., 2019. Fully distributed optical fiber strain sensor With $10^{-12}\epsilon/\sqrt{}$ Hz sensitivity, *J. Lightwave Technol.,* **37,** 4487–4495.

Dou, S. *et al.*, 2017. Distributed acoustic sensing for seismic monitoring of the near surface: a traffic-noise interferometry case study, *Sci. Rep.,* **7,** 1–12.

Embree, P., Berg, J. & Backus, M., 1963. Wide-band velocity filtering—the pie-slice process, *Geophysics,* **28,** 948–974.

Fang, G., Li, Y., Zhao, Y. & Martin, E., 2020. Urban near-surface seismic monitoring using distributed acoustic sensing, *Geophys. Res. Lett.,* **47,** e2019GL086115, doi:10.1029/2019GL086115.

Gibbons, S. & Ringdal, F., 2006. The detection of low magnitude seismic events using array-based waveform correlation, *Geophys. J. Int.,* **165,** 149–166.

Hartog, H., 2017. *An Introduction to Distributed Optical Fibre Sensors,* CRC Press, Boca Raton, FL.

He, H., Li, H., Pan, W., Luo, B., Zou, X. & Yan, L., 2017. SNR enhancement in phase-sensitive OTDR with adaptive 2-D bilateral filtering algorithm, *IEEE Photonics J.,* **9,** 655–666.

Hennenfent, H. & Herrmann, F., 2006. Seismic denoising with nonuniformly sampled curvelets, *Comp. Sci. Eng.,* **8,** 16–25.

Ibrahim, A., Lin, S., Xiong, J., Jiang, J., Fu, Y. & Wang, Z., 2020. Integrated principal component analysis denoising technique for phase-sensitive optical time domain reflectometry vibration detection, *Appl. Opt.,* **59,** 669–675.

Lellouch, A., Yuan, S., Spica, Z., Biondi, B. & Ellsworth, W., 2019. Seismic velocity estimation using passive downhole distributed acoustic sensing records: examples from the San Andreas Fault Observatory at depth, *J. geophys. Res.,* **124,** 6931–6948.

Li, Z. & Zhan, Z., 2018. Pushing the limit of earthquake detection with distributed acoustic sensing and template matching: a case study at the Brady geothermal field, *Geophys. J. Int.,* **215,** 1583–1593.

Li, Z., Shen, Z., Yang, Y., Williams, E., Wang, X. & Zhan, Z., 2021. Rapid response to the 2019 Ridgecrest earthquake with distributed acoustic sensing, *AGU Adv.,* **2,** e2021AV000395, doi:10.1029/2021AV000395.

Lindsey, N., Dawe, T. & Ajo-Franklin, J., 2019. Illuminating seafloor faults and ocean dynamics with dark fiber distributed acoustic sensing, *Science,* **366,** 1103–1107.

Lior, I. *et al.*, 2021. On the detection capabilities of underwater distributed acoustic sensing, *J. geophys. Res.,* **126,** e2020JB020925, doi:10.1029/2020JB020925.

Ma, J. & Plonka, J., 2010. A review of curvelets and recent applications, *IEEE Signal Process. Mag.,* **27,** 118–133.

Martin, E., Huot, F., Ma, Y., Cieplicki, R., Cole, S., Karrenbach, M. & Biondi, B., 2018. A seismic shift in scalable acquisition demands new processing: fiber-optic seismic signal retrieval in urban areas with unsupervised learning for coherent noise removal, *IEEE Signal Process. Mag.,* **35,** 31–40.

Qin, Z., Chen, L. & Bao, X., 2012. Wavelet denoising method for improving detection performance of distributed vibration sensor, *IEEE Photonics Technol. Lett.,* **24,** 542–544.

Qin, Z., Chen, H. & Chang, J., 2017a. Detection performance improvement of distributed vibration sensor based on curvelet denoising method, *Sensors,* **17,** 1380–1388.

Qin, Z., Chen, H. & Chang, J., 2017b. Signal-to-noise ratio enhancement based on empirical mode decomposition in phase-sensitive optical time domain reflectometry systems, *Sensors,* **17,** 1870–1880.

Sladen, A., Rivet, D., Ampuero, J., De Barros, L., Hello, Y., Calbris, G. & Lamare, P., 2019. Distributed sensing of earthquakes and ocean-solid Earth interactions on seafloor telecom cables, *Nat. Commun.,* **10,** 1–8.

Spica, Z., Nishida, K., Akuhara, T., Pétrélis, F., Shinohara, M. & Yamaha, T., 2020. Marine sediment characterized by ocean bottom fiber-optic seismology, *Geophys. Res. Lett.,* **47,** e2020GL088360, doi:10.1029/2020GL088360.

Starck, J., Candés, E. & Donoho, D., 2002. The curvelet transform for image denoising, *IEEE Trans. Image Process.,* **11,** 670–684.

van den Ende, M., Lior, I., Ampuero, J., Sladen, A., Ferrari, A. & Richard, C., 2021. A self-supervised deep learning approach for blind denoising and waveform coherence enhancement in distributed acoustic sensing data, *EarthArXiv,*.

Williams, E., Fernández-Ruiz, M., Magalhaes, R., Vanthillo, R., Zhan, Z., González-Herráez, M. & Martins, H., 2019. Distributed sensing of microseisms and teleseisms with submarine dark fibers, *Nat. Commun.,* **10,** 1–11.

Yu, C., Day, E., Van de Hoop, M., Campillo, M. & Van der Hilst, R., 2017. Mapping mantle transition zone discontinuities beneath the Central Pacific with array processing of SS precursors, *J. geophys. Res.,* **122,** 364–378.

Zhan, Z., 2019. Distributed acoustic sensing turns fiber-optic cables into sensitive seismic antennas, *Seismol. Res. Lett.,* **91,** 1–15.

Zhang, J. & Langston, C., 2020. Separating the scattered wavefield from teleseismic P using curvelets on the long beach array data set, *Geophys. J. Int.,* **220,** 1112–1127.

Zhou, J., Pan, Z., Ye, Q., Cai, H., Qu, R. & Fang, Z., 2013. Characteristics and explanations of interference fading of a ϕ-OTDR with a multi-frequency source, *J. Lightwave Technol.,* **31,** 2947–2954.

Zhu, T., Xiao, X., He, Q. & Diao, D., 2013. Enhancement of SNR and spatial resolution in ϕ-OTDR system by using two-dimensional edge detection method, *J. Lightwave Technol.,* **31,** 2851–2856.

SUPPORTING INFORMATION

Supplementary data are available at *GJI* online.

Figure S1 Hourly (PST) event counts of the catalogues created using template matching. A 3-hr moving average filter was applied to each of the time-series. The yellow line shows the hourly counts for the catalogue of templates made using the SCSN catalogue. The red line shows the hourly counts for the catalogue created using the unfiltered data. The green line shows the hourly counts for the catalogue created using data filtered using only the coherent noise filter. The purple line shows the hourly counts for the catalogue created using data filtered using only the stochastic noise filter. The blue line shows the hourly counts for the catalogue created using data filtered with both the coherent noise and stochastic noise filters.

Figure S2 Hourly (PST) measure of the percentage of events from the catalogue produced using unfiltered data retained in the catalogue produced using data filtered for both stochastic and coherent noise. Events are considered retained if there exists a corresponding event in the catalogue produced using filtered data that has an origin time within 2.5 s of the event. A 3-hr moving average filter was applied to this time-series.

Figure S3 Example of a detection made in the original catalogue but not made in the new catalogue. (a) Template event taken from the unfiltered data. (b) Wavefield of the detection made using the template in panel (a). (c) Same template as in panel (a) but taken from the filtered data. (d) Same wavefield as in panel (b) but taken from the filtered data. Black arrows point to the earthquake-generated body waves of the template event.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the paper.