On the Detection of Upper Mantle Discontinuities with Radon-Transformed Ps Receiver Functions (CRISP-RF)

Tolulope Olugboji[1]

Ziqi Zhang[1]

Steve Carr[1]

Canberk Ekmekci[2]

Mujdat Cetin[2,3]

[1]Department of Earth and Environmental Sciences, University of Rochester, Rochester, NY 14627, USA
[2]Department of Electrical and Computer Engineering, University of Rochester, Rochester, NY 14627, USA
[3]Georgen Institute of Data Sciences, University of Rochester, Rochester, NY 14627, USA

Key Points:

- CRISP-RF workflow is proposed for computing multiple-free and denoised P-to-S receiver functions using the sparse Radon transform
- The effectiveness of CRISP-RF is demonstrated by synthetic experiment and real data examples from single stations and common conversion point stack
- CRISP-RF enables detection of clear signals of the MLD beneath the Superior and Yilgarn cratons and the X-discontinuity beneath the SW Pacific

Corresponding author: Tolulope Olugboji, `tolulope.olugboji@rochester.edu`

Abstract

Seismic interrogation of the upper mantle from the base of the crust to the top of the mantle transition zone has revealed discontinuities that are variable in space, depth, lateral extent, amplitude, and lack a unified explanation for their origin. Improved constraints on the detectability and properties of mantle discontinuities can be obtained with P-to-S receiver function (Ps-RF) where energy scatters from P to S as seismic waves propagate across discontinuities of interest. However, due to the interference of crustal multiples, uppermost mantle discontinuities are more commonly imaged with lower resolution S-to-P receiver function (Sp-RF). In this study, a new method called CRISP-RF (Clean Receiver-function Imaging using SParse Radon Filters) is proposed, which incorporates ideas from compressive sensing and model-based image reconstruction. The central idea involves applying a sparse Radon transform to effectively decompose the Ps-RF into its underlying wavefield contributions, i.e., direct conversions, multiples, and noise, based on the phase moveout and coherence. A masking filter is then designed and applied to create a multiple-free and denoised Ps-RF. We demonstrate, using synthetic experiment, that our implementation of the Radon transform using a sparsity-promoting regularization outperforms the conventional least-squares methods and can effectively isolate direct Ps conversions. We further apply the CRISP-RF workflow on real data, including single station data on cratons, common-conversion-point (CCP) stack at continental margins, and seismic data from ocean islands. The application of CRISP-RF to global datasets will advance our understanding of the enigmatic origins of the upper mantle discontinuities like the ubiquitous Mid-Lithospheric Discontinuity (MLD) and the elusive X-discontinuity.

1 Introduction

Global seismic imaging has produced maps of upper mantle layering that have important implications for mantle thermo-chemical heterogeneity, rheology, and dynamics (Deuss, 2009; Fischer et al., 2020; Karato et al., 2015; Karato & Park, 2018; Schmerr, 2015; Shearer, 2000; Tharimena et al., 2017). A few examples include the detection of a ubiquitous middle-lithosphere discontinuity (MLD) (Abt et al., 2010; Hopper & Fischer, 2018; Krueger et al., 2021), the global lithosphere-asthenosphere system (Kind et al., 2020; Liu & Shearer, 2021; Mancinelli et al., 2017; Rychert et al., 2005), the Lehmann discontinuity (Deuss & Woodhouse, 2004; Karato, 1992), and the X-discontinuity (Pugh et al., 2021, 2023; Schmerr, 2015; Srinu et al., 2021). Each of these layers can be explained by invoking some combination of partial-melting, phase-changes, chemical stratification, variable anisotropy, and elastically-accommodated grain-boundary sliding (Beghein et al., 2014; Karato et al., 2015; Olugboji et al., 2013; Rader et al., 2015; Rychert et al., 2020; Schmerr, 2015; Selway et al., 2015). Improved resolution of the depth, amplitude of velocity change, and sharpness (i.e., the depth interval of the velocity gradient) is important for discriminating between proposed models for the various types of upper mantle layering (Benz & Vidale, 1993; Fischer et al., 2020; Karato et al., 2015; Kawakatsu et al., 2009; Mancinelli et al., 2017; Petersen et al., 1993; Rychert et al., 2005).

While the Moho and the mantle transition-zone discontinuities are generally global, relatively sharp, consistently marked by a velocity increase, and widely accepted to be caused by changes in rock composition and mineral phase transformations, other upper-mantle discontinuities are often sporadic, inconsistent in amplitude and polarity (Abt et al., 2010; Krueger et al., 2021; Revenaugh & Jordan, 1991), variably gradational (Eaton et al., 2009; Liu & Shearer, 2021; Sun, Kennett, et al., 2018), and lacking an agreed-upon explanation for their origins (Aulbach, 2018; Karato & Park, 2018; Krueger et al., 2021). As a result, these discontinuities are typically better detected by high-resolution reflectivity techniques that use reflected and converted waves with or without earthquake-source deconvolution (Kind et al., 2020; Kind & Yuan, 2018; Liu & Shearer, 2021; Tauzin et al., 2019).

Amongst the different types of imaging methods based on body-wave reflectivity, e.g., top-side S-reflections (Buehler & Shearer, 2017; Liu & Shearer, 2021; Schutt et al., 2018), seismic daylight imaging (Sun, Kennett, et al., 2018; Sun & Kennett, 2017), and earthquake or noise correlation (Gómez-García et al., 2022; Kennett, 2015; Poli et al., 2012; Sun & Kennett, 2016), the receiver function technique has seen the widest application for upper mantle discontinuity imaging (Birkey et al., 2021; Fischer, 2015; Ford et al., 2010; Hopper & Fischer, 2018; Kind & Yuan, 2018; Rychert et al., 2020). This is because receiver functions target receiver-side structure after the source and path have been deconvolved. These source-deconvolved seismograms aid in detecting discontinuities either using shear-to-compressional converted waves (Sp-RFs) or compressional-to-shear converted waves (Ps-RFs) (Rychert et al., 2005, 2007; Rychert & Shearer, 2009). The S-to-P receiver function (Sp-RF) approach is most commonly used for mantle-discontinuity imaging because it is not affected by interference from crustal reverberations (Kind & Yuan, 2018; Kumar et al., 2012). However, it is well known that its spatial and depth-resolution is not comparable to the P-to-S receiver function (Ps-RF) due to it being observed at limited epicentral distances, having poorer signal-to-noise quality and containing longer period signals (Kind et al., 2020; Kind & Yuan, 2018; Lekić & Fischer, 2017; Shearer & Buehler, 2019). By contrast, the Ps-RF technique, which has been widely successful for crustal imaging (Bostock, 2004; Olugboji & Park, 2016; Zhu & Kanamori, 2000), is higher resolution, but has seen limited use in continental-scale lithospheric imaging primarily due to signal distortion caused by the overprinting of crustal reverberations, i.e., wave echoes trapped in the crustal column (Figure 1).

Here, we describe a new methodology called the CRISP-RF, an acronym that stands for the signal processing workflow that promotes '**C**lean **R**eceiver-function **I**maging (from a noisy one) using a **SP**arse **R**adon **F**ilter'. This approach addresses some of the limitations of the more traditional Ps-RFs by developing a sparse Radon transform to model the observed data, and a masking filter to suppress the effects of crustal reverberations that overprint the Ps-RF traces. The Radon transform is widely applied in medical imaging, radar astronomy, and material science (Deans, 2007). In global geophysics, the Radon transform has been widely used for noise suppression when interpreting mantle discontinuities imaged with bottom-side reflections, e.g., SS, PP, or P'P' (An et al., 2007; Gu & Sacchi, 2009; Schultz & Gu, 2013; Schultz & Jeffrey Gu, 2013). However, much of the initial development and current advances have been focused in the field of exploration geophysics (Hampson, 1986; Sacchi & Ulrych, 1995; Trad, 2003), with a few recent applications in Ps-RF imaging (Aharchaou & Levander, 2016; Chen et al., 2022; Dokht et al., 2016; Gu et al., 2015; Wilson & Guitton, 2007; Q. Zhang et al., 2022). In our extension of the Radon transform to high-resolution upper mantle imaging, we borrow from recent advances in the fields of compressed sensing and low-dimensional model-based image reconstructions (Candès & Wakin, 2008; Geng et al., 2022; Trad et al., 2003; Wright & Ma, 2022) with the goal of attenuating crustal multiples that interfere with upper mantle discontinuities (Figure 1).

2 CRISP-RF: Methodological Overview

In traditional processing, the observed Ps-RF is obtained by source deconvolution (Abt et al., 2010; Bostock, 2004):

$$\mathbf{d}(t,p) = \mathcal{F}^{-1}[\frac{U^r(\omega,p) * U^r(\omega,p)}{U^z(\omega,p) * U^z(\omega,p) + \zeta)}] \tag{1}$$

where $\mathcal{F}^{-1}$ is the inverse Fourier transform, $U^r(\omega,p)$ and $U^z(\omega,p)$ are the Fourier transformed radial and vertical seismograms for each recorded earthquake propagating with slowness p, and ζ is a damping factor. The Ps-RF data is a 2-D matrix in which each row represents one trace of Ps-RF (time series) with a distinct slowness, and each column a discrete-time sample. Depending on the data distribution, for a given ray parameter at a given time, the observed Ps-RF in (equation 1) can be modelled as arrivals with amplitudes

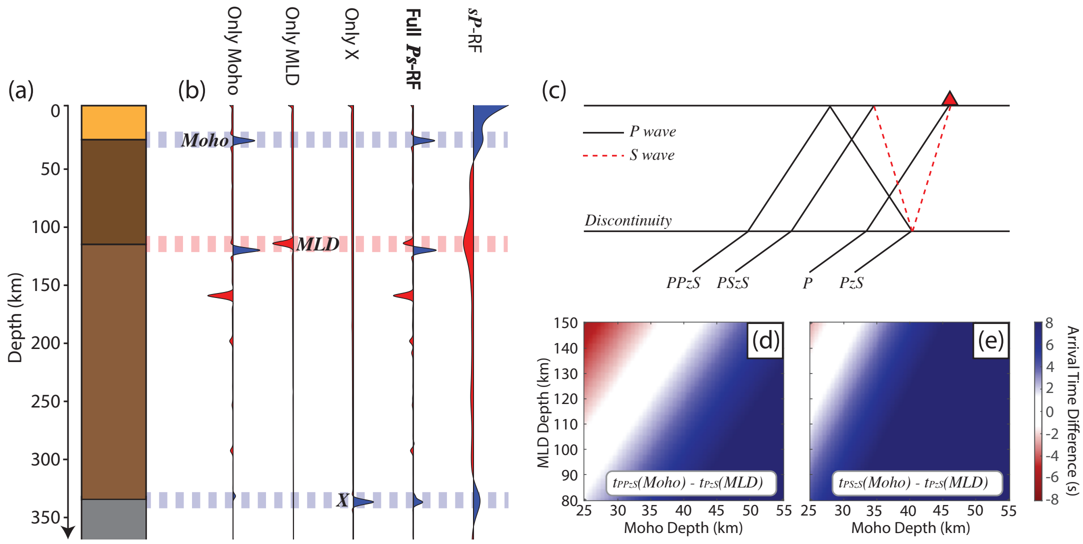

Figure 1. Synthetic example illustrating the challenge of upper mantle imaging using Ps-RF: the interference (and aliasing) of crustal multiples with (as) conversions from upper mantle discontinuities. (a) A representative earth model showing crust and upper mantle discontinuities, including the Moho, MLD, and X. (b) The synthetic single-event Ps-RF trace assuming a single-interface model (Moho, MLD, or X), compared with the synthetic Ps- and Sp-RF from the full model shown in (a). (c) Wave propagation of direct conversions, PzS, and multiples, PPzS and PSzS, associated with a layer at depth z. (d) An interference diagram showing which crustal models creates a PPzS multiple that coincides with the direct conversion (PzS) from a mid-lithosphere discontinuity (MLD). (e) Similar interference diagram but for the later arriving PSzS multiple that interferes with the direct conversion (PzS) of the MLD and, being the same polarity, can alias as an MLD. See also Figure S1 in Supporting Information.

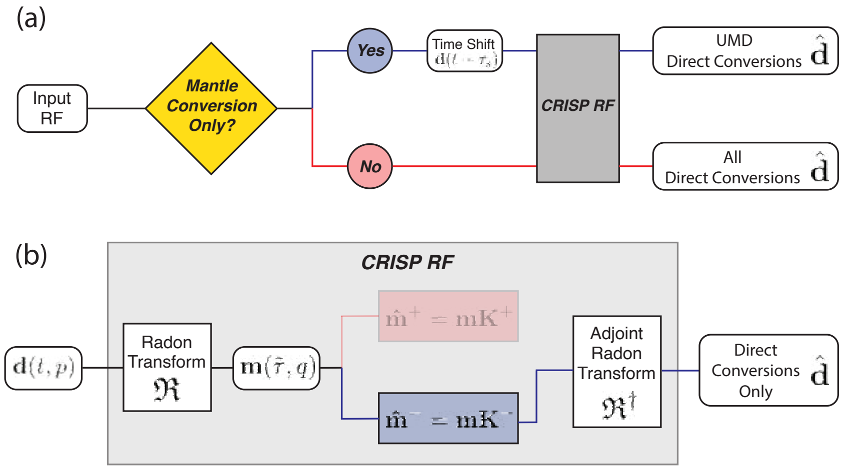

Figure 2. The CRISP-RF signal processing workflow for targeting mantle discontinuities by removing crustal multiples. (a) The workflow can be preceded by an optional pre-processing step that applies a time shift operator, τ_s, to the input data. In this mode, all crustal conversions and reverberations are removed and only late-arriving mantle conversions are targeted, e.g., X-discontinuity. (b) CRISP-RF workflow: The first step is a Radon transform of the original Ps-RF generating a model in the intercept-time curvature domain (see section 3). The second step applies a masking filter on the Radon model to remove negative curvatures (K^-). If isolating crustal multiples is a goal, then they can be separated, instead of being removed, by using a masking filter on positive curvatures (K^+). The last step returns the filtered Ps-RF to the time-slowness domain by using an adjoint Radon transform on the filtered Radon model.

corresponding to direct conversions (k=1) and their respective multiples (k >1) within j layers (Galetti & Curtis, 2012; Tauzin et al., 2019):

$$\mathbf{d}(t,p) = \underbrace{\sum_j \sum_{k>1} \mathbf{a}^{jk}\delta(t-\tau^{jk}) + \mathbf{n}(t)}_{\text{multiples + noise (remove)}} + \underbrace{\mathbf{a}^{j1}\delta(t-\tau^{j1})}_{\text{direct conversions, k=1 (keep)}} \tag{2}$$

where $\mathbf{a}^{jk}$ are the amplitudes of the converted and reflected waves, τ^{jk} are their arrival times with respect to the direct P arrival, and δ is the Dirac-delta function (assuming an impulsive source). In the CRISP-RF workflow, the goal is to remove(separate) the unwanted wavefield contributions, i.e., multiples and incoherent noise, $\mathbf{n}(t)$, from the target arrival (direct Ps conversions, k=1, generated by upper mantle discontinuities: $j > n_c$ where n_c is the number of crustal layers) in the original Ps-RF data (Figure 2 and Equation 1).

Our approach involves three steps: (1) transforming the time-slowness domain Ps-RF (input data $\mathbf{d}$) into a intercept-time-curvature domain Radon image (intermediate model $\mathbf{m}$), using a sparsity-promoting Radon transform (curvature is the moveout of the arriving phase), $\mathfrak{R}$: $\mathbf{d} \xrightarrow{\mathfrak{R}} \mathbf{m}$; (2) applying a masking filter $\mathbf{K}$ in the Radon model that suppresses multiples and noise such that $\hat{\mathbf{m}} = \mathbf{m}\mathbf{K}$; and finally (3) obtaining a filtered Ps-RF output, $\hat{\mathbf{d}}$, after transforming back into the time-slowness domain, using the adjoint Radon transform,

$\mathfrak{R}^\dagger$: $\hat{\mathbf{m}} \xrightarrow{\mathfrak{R}^\dagger} \hat{\mathbf{d}}$. In cases where the target upper mantle discontinuities arrive much later than crustal multiples (e.g., > 250 km), the CRISP-RF workflow can be preceded by an optional step that implements a moving-window time-shifting algorithm that targets sub-Moho conversions (Helffrich, 2006; Shibutani et al., 2008; Park & Levin, 2016)(Figure 2a). Applying this step improves the detection of low-amplitude arrivals that convert in the mantle since the Ps-RF amplitude is not overwhelmed by the stronger coherent phases (the Moho and its multiples). In the following sections, we illustrate each processing step of the CRISP-RF workflow, explaining how they produce the desired effect of high-resolution imaging of mantle discontinuities with Ps-RFs following the removal of noise and crustal multiples.

3 Denoising and Attenuating Undesired Multiples in Ps-RFs Using Radon Transform

The Radon transform, like most other transforms, allows us to represent data, i.e., the Ps-RF data $\mathbf{d}$, by a sparse model-set, $\mathbf{m}$ (Beylkin, 1987; Ö. Yilmaz, 2015):

$$\mathbf{d}(t,p) = \mathfrak{R}^\dagger\{\mathbf{m}(\tilde{\tau},q)\} \triangleq \sum_{i=1}^{N_q} \mathbf{m}(\tilde{\tau} = t - q_i p^2, q_i) \tag{3}$$

where $\mathbf{d}(t,p)$ is the Ps-RF data in time-slowness domain, $\mathbf{m}(\tau,q)$ is the Radon model in intercept-time-curvature domain, and $\mathfrak{R}^\dagger$ is the adjoint Radon transform. Ideally, the Radon model should be sparse and only has non-zero amplitudes ($\mathbf{a}^{jk}$) at intercept-times ($\tilde{\tau}^{jk}$), i.e., zero-slowness arrival times of coherent phases (direct conversions, PzS, and multiples, PPzS and PSzS), and curvatures (q^{jk}), i.e., the extent of the moveout of the phases (e.g., Figure 1a). The adjoint Radon transform, $\mathfrak{R}^\dagger$, reconstructs the Ps-RF ($\mathbf{d}$) by summing the amplitudes of the Radon model at all curvature (q_i) along each slowness (p).

The Radon model reconstructs each wavefield contribution at the required slowness p with the correct time-shift $q_i p^2$, which is parabolic in slowness with the curvature q as the coefficient. To better understand this approximation and why it can separate direct conversions from multiples, consider the Taylor expansion of the arrival time for each wavefield contribution given a single-layer model with thickness h, compressional velocity α, and shear velocity β (Ryberg & Weber, 2000; Shi et al., 2020):

Direct conversions (PzS, k=1):

$$\begin{aligned} &\tau^{j1} = \tilde{\tau}^{j1} + q^{j1}p^2 \\ &\tilde{\tau}^{j1}_{PzS} \approx h(1/\alpha - 1/\beta) \qquad q^{j1} \approx +\frac{h(\alpha-\beta)}{2} \end{aligned} \tag{4a}$$

Multiples (PPzS and PSzS, k>1):

$$\begin{aligned} &\tau^{jk} = \tilde{\tau}^{jk} + q^{jk}p^2 \\ &\tilde{\tau}^{j2}_{PPzS} \approx h(1/\alpha + 1/\beta) \qquad q^{j2} \approx -\frac{h(\alpha+\beta)}{2} \\ &\tilde{\tau}^{j3}_{PSzS} \approx \frac{2h}{\beta} \qquad q^{j3} \approx -h\beta \end{aligned} \tag{4b}$$

Since the direct Ps conversions have a positive curvature, while the multiples, typically from reflections in the overlying crustal layer, have a negative curvature, the wavefield contributions of a conversion from a mantle discontinuity can be separated from the interfering crustal multiples. The adjoint Radon transform can be written in matrix form by applying a Fourier transform to both sides of Equation 3 (Gu & Sacchi, 2009; Ö. Yilmaz, 2015):

$$\mathbf{D}(\omega,p) = \sum_{i=1}^{N_q} \mathbf{M}(\omega,q_i) e^{(-i\omega q_i p^2)} \tag{5}$$

where $\mathbf{D}$ and $\mathbf{M}$ are obtained from the Fourier transform of $\mathbf{d}$ and $\mathbf{m}$. The moveout is then modeled as a phase-shift term in the frequency domain, $\omega q_i p^2$, and allows for timing corrections that are not integer multiples of the sampling interval of the data.

We illustrate the properties of sparsity and curvature-based mode separation by generating a synthetic radon model, $\mathbf{m_s}$, for a layered model with a two-layer crust (intra-crustal boundary, ICB, and Moho), a mantle discontinuity (MD) and a half-space (Figure 3a). The relative amplitudes ($\mathbf{a}^{jk}$) are derived from the reflection and transmission coefficients, while the intercept-time ($\tilde{\tau}^{jk}$) and curvature (q_{jk}) are estimated analytically from Equation 4 (Figure 3b). The adjoint radon transform of $\mathbf{m_s}$ produces a synthetic Ps-RF data, $\mathbf{d_s}$, comparable to that generated by reflectivity synthetics (Figure 3c). The interference problem is clearly observed, as the Ps conversion from the mantle discontinuity (120 km) arrives at the same time ($\sim$12 s) as the multiples from the shallow crustal discontinuity (24 km), and when processed without a radon filter, produces a stack that is difficult to interpret (Figure 3c). This is corrected by applying a masking filter in the radon model that sets all amplitudes with negative curvatures to zero and only keeps amplitudes at positive curvatures, followed by the adjoint of the radon transform: $\hat{\mathbf{d}}_\mathbf{s} = \mathfrak{R}^\dagger(\mathbf{m_s K^-})$, where $\mathbf{K}^-$ denotes the masking filter and $\hat{\mathbf{d}}_\mathbf{s}$ denotes the output clean Ps-RF. After this treatment, only direct conversions can be observed in the Ps-RFs, and the consequent average stack clearly shows all discontinuities (ICB, Moho, and MD) (compare Figure 3d and 3c).

4 The Sparsity-promoting Radon Transform: Algorithms & Synthetic Examples

As elaborated above, in the Radon domain, mode-conversions are clearly separated from multiple reflections within the crust and can be removed by a masking filter, $\mathbf{K}$, that eliminates the unwanted multiples (Figure 3). In practice, however, the challenge is not in designing the masking filter but in the first step of the CRISP-RF workflow, which involves obtaining the Radon model by computing a stable Radon transform of the Ps-RF (Figure 2). This is because, while the adjoint Radon transform ($\mathfrak{R}^\dagger$) is unique and easily computable, its forward transform ($\mathfrak{R}$) is non-unique, difficult to estimate, and requires finding the inverse of the following equation:

$$\mathcal{F}(\mathbf{d}) = \mathbf{L}\, \mathcal{F}(\mathbf{m}) \tag{6a}$$

$$\begin{bmatrix} D_{p_1} \\ D_{p_2} \\ \vdots \\ D_{p_n} \end{bmatrix} = \begin{bmatrix} e^{(-i\omega q_1 p_1^2)} & e^{(-i\omega q_2 p_1^2)} & \dots & e^{(-i\omega q_m p_1^2)} \\ e^{(-i\omega q_1 p_2^2)} & e^{(-i\omega q_2 p_2^2)} & \dots & e^{(-i\omega q_m p_2^2)} \\ \vdots & \ddots & \ddots & \vdots \\ e^{(-i\omega q_1 p_n^2)} & e^{(-i\omega q_2 p_n^2)} & \dots & e^{(-i\omega q_m p_n^2)} \end{bmatrix} \begin{bmatrix} M_{q_1} \\ M_{q_2} \\ \vdots \\ M_{q_m} \end{bmatrix} \tag{6b}$$

where $\mathcal{F}$ is the one-dimensional Fourier transform operator, $\mathbf{L}$ is the matrix operator that implements the adjoint transform from the Fourier-transformed Radon model ($\mathbf{M}$) to the Fourier-transformed Ps-RF data ($\mathbf{D}$), for each frequency (ω), curvature (q), and slowness (p). Because $\mathbf{D}$ is noisy and sparsely sampled, $\mathbf{L}$ is underdetermined, non-orthogonal, and does not have a true inverse (Menke, 2012; Sacchi & Ulrych, 1995). The most popular solution to this general inverse solution is the damped least-squares approach (An et al., 2007; Aster et al., 2018; Menke, 2012) and it defines the least-squares Radon transform:

$$\mathfrak{R}_{\text{ls}}(\mathcal{F}(\mathbf{d})) = \underset{\mathbf{m}}{\operatorname{argmin}}\{||\mathbf{L}\mathcal{F}\{\mathbf{m}\} - \mathcal{F}\{\mathbf{d}\}||_2^2 + \mu||\mathcal{F}\{\mathbf{m}\}||_2^2\} \tag{7}$$

It obtains the Radon model of the Fourier-transformed Ps-RF data by imposing a least-squares, ℓ_2, error (first term on the right-hand side of Equation 7) subject to the regularization term that minimizes the ℓ_2-norm on the model size (second term on the

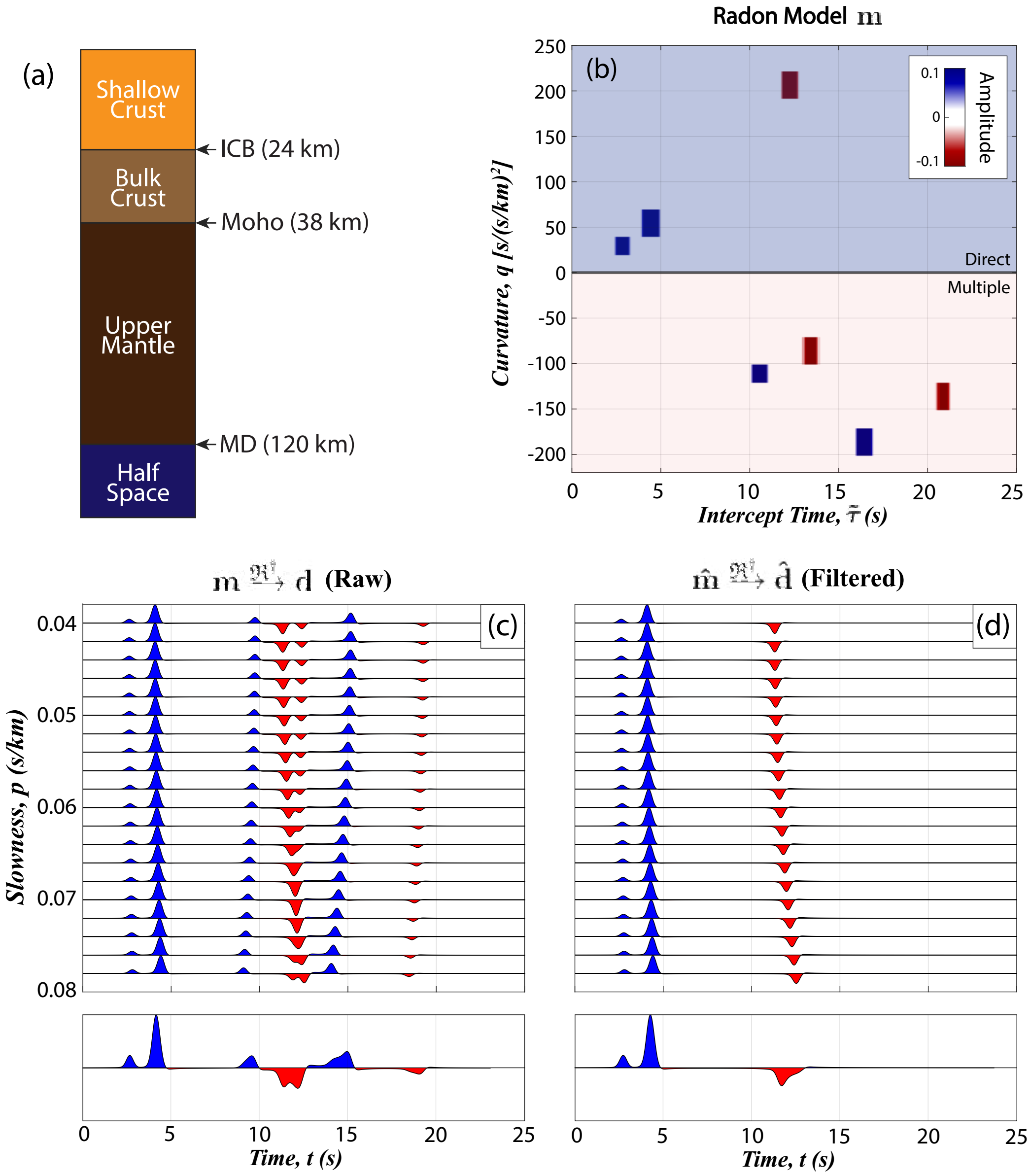

Figure 3. Filtering of crustal multiples in the synthetic radon model by masking arrivals with negative curvature. (a) The Earth model used to calculate the synthetic radon model. (b) Synthetic radon model showing sparse representation of Ps-RFs: direct conversions (blue shading) and multiples (red shading). (c) Ps-RF traces and the average stack calculated from the adjoint of radon transform of the radon model shown in (b), showing robust identification of the upper mantle discontinuity even in the presence of multiples which arrive at similar times. (d) Same as (c) but with a filter masking all negative curvatures in the radon model, showing improved detection of the conversion for the mantle discontinuity with multiples eliminated.

right-hand side of Equation 7). We demonstrate the behavior of the damped least-squares inverse solution to the Radon transform, $\mathfrak{R}_{\mathbf{ls}}$, using a noisy synthetic Ps-RF that mimics the behavior of realistic data:

$$\mathbf{d}^* = \mathfrak{R}^\dagger(\mathbf{m_s}) + \mathbf{n}(t, \eta_1, \eta_2, n_r) \tag{8}$$

We use a realistic noise model, where η_1 is the signal-to-noise ratio (SNR) of the background noise applied to all traces, η_2 is the SNR of the noisiest traces ($\eta_2 << \eta_1$) and η_2 is applied to n_r traces chosen at random from the set of all traces in the noise-free synthetic: $\eta_1, \eta_2, n_r = (1, 0.5, 10\%)$ (Figure 4a and b). We compute the Radon transform of this noisy synthetic Ps-RF using the damped least-squares inversion, $\mathfrak{R}_{\mathbf{ls}}$ (Figure 4c).

We observe that this approach to computing the Radon transform introduces artifacts (streaking and low-amplitude errors) that lead to filtering errors when using this to compute the filtered Ps-RF. Improvement to the filtered Ps-RF requires suppressing artifacts that are caused by the damped least-squares process: (1) requiring a Radon model whose amplitudes are better resolved along the intercept time-curvature axes (reduced streaks seen in Figure 4c), especially for crustal phases and multiples; (2) suppressing background noise in the Radon model that maps into the Ps-RF data as spurious phases; and (3) improved regularization using information on noise gleaned from data. A variety of techniques have been proposed for achieving the goals of higher resolution and they reduce to enforcing sparsity on the recovered Radon model by modifying the regression problem with an ℓ_1-norm constraint or its equivalent (Ji, 2006; Luo et al., 2008; Sacchi & Ulrych, 1995; Thorson & Claerbout, 1985; Trad et al., 2003):

$$\mathfrak{R}_{\text{sp}}(\mathcal{F}(\mathbf{d})) = \underset{\mathbf{m}}{\text{argmin}}\{||\mathbf{L}\mathcal{F}\{\mathbf{m}\} - \mathcal{F}\{\mathbf{d}\}||_2^2 + \lambda||\mathcal{F}\{\mathbf{m}\}||_1\} \tag{9a}$$

This formulation is the popular frequency-domain sparse Radon transform which enforces sparsity along the curvature axis, but may still retain spurious artifacts in the time axis due to the frequency-time coupling of noise present in a few high-energy traces. The proposed solution to this problem requires implementing a mixed frequency-time sparse Radon transform that imposes sparsity along both the time-and-curvature axis:

$$\mathfrak{R}_{\text{sp}}(\mathbf{d}) = \underset{\mathbf{m}}{\text{argmin}}\left\{\frac{1}{2}\|\mathcal{F}^{-1}\{\mathbf{L}\mathcal{F}\{\mathbf{m}\}\} - \mathbf{d}\|_2^2 + \lambda\psi(\mathbf{m})\right\} \tag{9b}$$

where the sparsity-promoting regularizers could either be the ℓ_1-norm regularization ψ_1 or the mixed $\ell_1 - \ell_2$ regularization ψ_2, defined by $\psi_1(\mathbf{m}) = \|\mathbf{m}\|_1$ and $\psi_2(\mathbf{m}) = \|\mathbf{m}\|_1 - \beta\|\mathbf{m}\|_2$, where $\beta \geq 0$ is an additional regularization parameter that needs to be tuned.

The form of this restated problem describing the sparse Radon transform has been well studied in the field of optimization and compressed sensing and several methods have been proposed to solve such problems. Examples include, but are not limited to, alternating direction method of multipliers (ADMM) (Boyd, 2010), proximal gradient descent methods (Parikh, 2014), and iterative shrinkage algorithms (Beck & Teboulle, 2009). We explore three different iterative algorithms for computing the sparse Radon transform: (1) SRTIS: the iterative 2D model shrinkage-based sparse inverse Radon transform(Gong et al., 2016; Lu, 2013); (2) SRTFISTA: the fast iterative shrinkage-thresholding algorithm-based sparse inverse Radon transform(Beck & Teboulle, 2009; Gong et al., 2016); and (3) SRTL$_{1-2}$: the mixed-norm sparse Radon transform(Geng et al., 2022; Tao & An, 1998).

In the most general case (SRTIS and SRTFISTA), the algorithms follow a variation of the following steps in the accelerated proximal gradient methods (Wright & Ma, 2022):

(0) Initialize a Radon-model, $\mathbf{m_0}$, and $\mathbf{s_1} = \mathbf{m_1} \leftarrow \mathbf{m_0}$;

(1) compute auxiliary point $\mathbf{s_{i+1}} = \mathbf{m_i} + \beta_\mathbf{i}(\mathbf{m_i} - \mathbf{m_{i-1}})$;

(2) descend from this point using gradient: $\mathbf{z_{i+1}} = \mathbf{s_{i+1}} - \gamma\mathbf{A}^*(\mathbf{A}\mathbf{s_{i+1}} - \mathbf{d})$, where $\mathbf{A}^*$ is adjoint of operator $\mathbf{A} = \mathcal{F}^{-1}\mathbf{L}\mathcal{F}$ in Equation 9b: $\mathbf{d} = \mathbf{Am} + \mathbf{n}$;

(3) apply a thresholding function to promote sparsity: $\mathbf{m}_{i+1} = \mathcal{S}_\gamma(\mathbf{z}_{i+1}, \gamma\lambda)$.

Set $i = i + 1$, and repeat steps 1 to 3 for I_t times until convergence (see Text S1 and Figures S2-S6 in Supporting Information for details of each specific algorithm and the solution of SRTL_{1-2} using the ADMM algorithm). We illustrate the performance of the sparse Radon transform, $\mathfrak{R}_{\mathrm{sp}}(\mathbf{d})$, by comparing it to the damped least-squares solution, $\mathfrak{R}_{\mathrm{ls}}(\mathbf{d})$(compare Figure 4f and 4c). The sparse Radon solution, obtained using the SRTIS algorithm and initialized using the damped least-squares solution, is a higher-resolution Radon model with most of the artifacts from the least-squares process removed. This sparse Radon model is then used to compute a filtered Ps-RF after applying a diagonal masking filter: $\hat{\mathbf{d}}*$ (Figure 4d). The masking filter is obtained by predicting the $\tilde{\tau} - q$ and associated bounds (dashed lines in Figure 4f) through the reference Earth model in Figure 3a. A comparison of the input and output model to the CRISP-RF workflow (Figure 4e and 4b) shows that the sparse Radon transform has successfully denoised and attenuated the multiple reflections in the crust isolating the direct mantle conversions.

5 Application to Real Data: Single Station and CCP Ps-RFs

We now present four exemplary Ps-RFs to further illustrate the utility of the CRISP-RF methodology for upper mantle imaging using real data (Figure 5): (1) from a station located above the Superior Craton (CN.ULM), (2) from a station located on the Yilgarn Craton in Western Australia (AU.KMBL), (3) a common-conversion point example with grid-center located near the passive continental margin in Massachusetts, USA (IU.HRV), and (4) a final example obtained from the Samoa ocean island station (IU.AFI). In the first three examples, we illustrate the denoising and attenuation of crustal multiples interfering with the MLD, and in the last example, we show improved resolution and sharpness of the X- discontinuity. The stations are all selected based on previous detections of the target upper mantle discontinuities with other imaging approaches, e.g., Sp-RF or autocorrelation analysis (Abt et al., 2010; Birkey et al., 2021; Ford et al., 2010; Pugh et al., 2021; Sun, Kennett, et al., 2018).

At each of these stations we choose earthquakes with the best signal-to-noise ratio (> 2.0 on radial components) and moment magnitude > M_w 5.0 located at epicentral distances between 30° and 90° (p = 0.04 s/km to 0.08 s/km). We calculate radial Ps-RF traces, and uncertainties, using the moving-window migration multi-taper correlation (MWM-MTC) approach (Park & Levin, 2000, 2016). This involves time-shifting and tapering seismograms with κ Slepian windows, W_κ, before spectral estimation and source deconvolution:

$$\tilde{U}_\kappa^{z,r}(\omega, p) = W_\kappa * [U_\kappa^{z,r}(\omega, p)e^{(i\omega\tau_s)}] \tag{10a}$$

$$\tilde{\mathbf{D}}(\omega, p) = [\frac{\sum_{\kappa=0}^{\kappa-1} \tilde{U}_\kappa^r(\omega, p) * \tilde{U}_\kappa^r(\omega, p)}{\sum_{\kappa=0}^{\kappa-1} \tilde{U}_\kappa^z(\omega, p) * \tilde{U}_\kappa^z(\omega, p) + \delta}] \tag{10b}$$

When targeting the MLD (100 - 200 km), we eliminate the direct P arrival by applying a small time shift: τ_s = 1.0 s, to the radial seismograms, and when imaging the deeper X-discontinuity (250 - 400 km), a longer time shift of τ_s = 15.0 s is applied (see Figure 2a). The final Ps-RF data, $\mathbf{D}$, is a less-noisy, low-dimension filtered copy of $\tilde{\mathbf{D}}$:

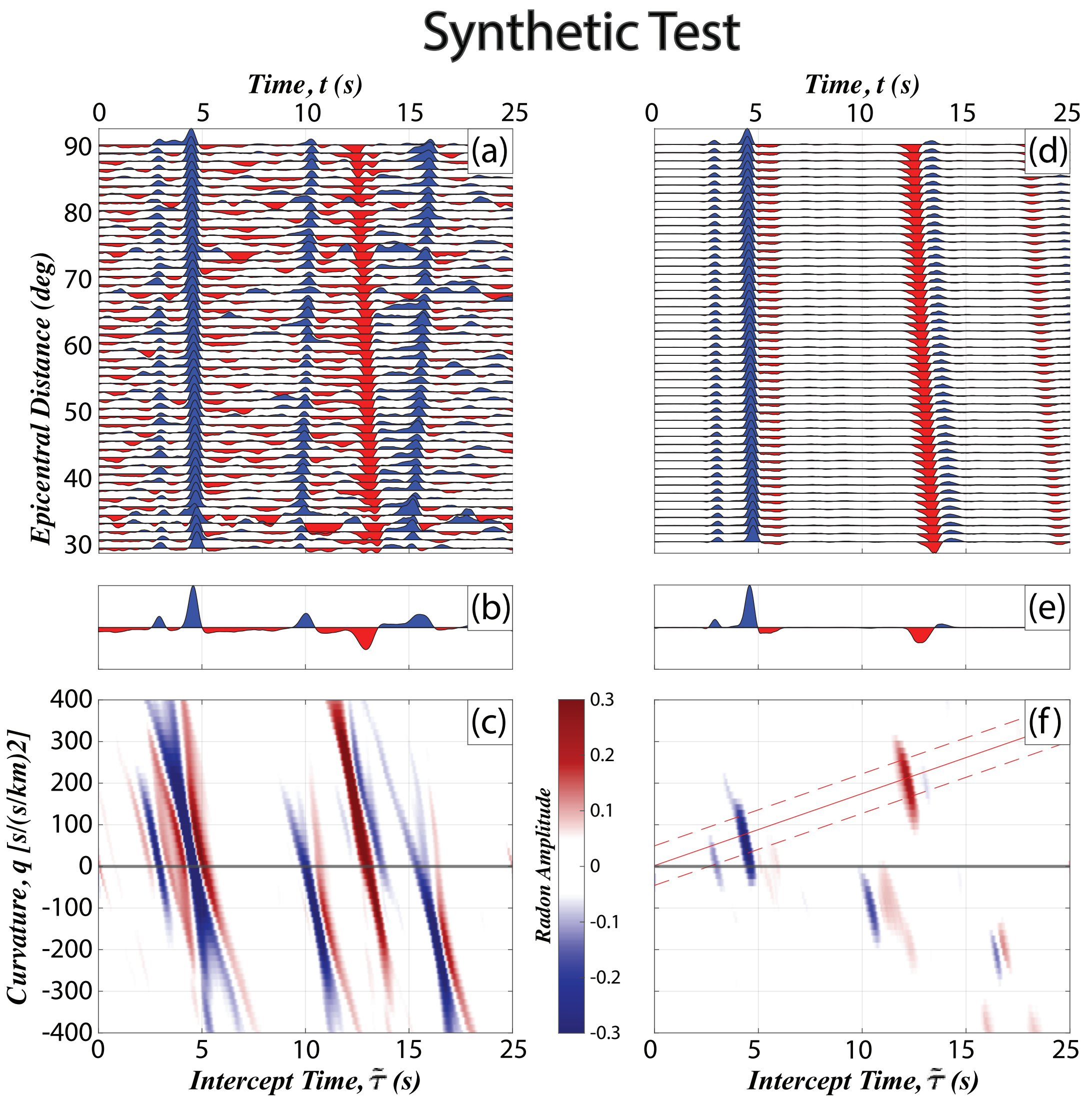

Figure 4. Synthetic example of multiple removal and noise attenuation using the CRISP-RF signal processing workflow. (a) A noisy synthetic Ps-RF data obtained using a realistic noise model. (b) Average stack of the unfiltered synthetic Ps-RF. (c) Radon model obtained from a damped least-squares inverse in the first step of the CRISP-RF workflow. (d) The filtered Ps-RF using the sparse radon model in (f) below with only direct Ps conversion phases visible. (e) Average stack of filtered Ps-RF showing elimination of all unwanted signals. (f) The sparse radon model overlaid with a K-diagonal filter (red lines). The sparse radon model is initialized with (c) and obtained after 30 iterations. A noise-free version of this synthetic test can be found in Figure S7 in Supporting Information.

$$\mathbf{D}(\omega, p_s) = \frac{\sum_{l=0}^{n_p} 1/\sigma^2 \tilde{\mathbf{D}}(\omega, p_l)}{\sum_{l=0}^{n_p} (1/\sigma^2)} \tag{11a}$$

$$\sigma^2(\omega, p) = \frac{1 - C_{ZR}^2}{(\kappa - 1)C_{ZR}^2} |\mathbf{D}(\omega, p)|^2 \tag{11b}$$

with stacking weights, $1/\sigma^2$, prescribed by the frequency-dependent variance estimates, $\sigma^2(\omega, p)$, and obtained from the coherence, C_{ZR}^2 , between the vertical and radial seismograms. The slowness dimension is reduced by averaging of $\tilde{\mathbf{D}}$ along the slowness axis after discretization into n_p equally-spaced slowness p_s. The optimal discretization of the slowness bins is chosen by trial and error. Since earthquake data is band-limited, we apply a low-pass cosine filter with a cut-off frequency of 2.0 Hz to the first three examples targeting the MLD (CN.ULM, AU.KMBL, and CCP-IU.HRV), and set the cut-off frequency to 0.6 Hz for the last example targeting the X-discontinuity (IU.AFI).

5.1 MLD beneath Superior Craton (CN.ULM)

We select Ps-RF traces located at epicentral distances between 30° and 84° and stack them every 1° with 10° overlapping bins at station CN.ULM, which has previously been studied for MLD detection (Abt et al., 2010; Fischer et al., 2010; Karato et al., 2015; Selway et al., 2015). The direct conversion (∼4 s) and multiples (∼15 s and ∼18 s) from the Moho are clearly visible in the obtained Ps-RF and the average stack (Figure 6a-b). However, it is hard to judge the presence and depth of the MLD solely from this stacked Ps-RF since there are two major negative phases (15 - 20 s). We initialize the Radon model using the least-squares optimization (Figure 6c) and then apply the SRTIS algorithm with 30 iterations to obtain a sparse Radon model (Figure 6f). All three phases associated with the Moho can be clearly observed in the obtained sparse radon image, including the direct conversion (blue phase at ∼4 s at positive curvature) and two multiples (blue phase at ∼15 s and red phase at ∼18 s at negative curvature) (Figure 6f). In addition, a direct conversion phase is clearly visible at ∼15 s with a bigger curvature than the Moho (Figure 6f), indicating the presence of the MLD. Note that this phase arrives between the two Moho multiples, but can be well separated using the moveout (curvature) information and retrievable by the sparse Radon model. Compared to the sparse radon image, the least-squares Radon model of the same data is harder to interpret with a lot more artifacts and amplitudes that are smeared across the curvature axis (Figure 6c).

We then apply the adjoint Radon transform on the sparse Radon model after applying the diagonal masking filter (red lines in Figure 6f). The resulting filtered Ps-RF (Figure 6d) and the final migrated and phase-weighted stack (Figure 6e) show only two major direct conversion phases. A comparison of this final stack with the simple average stack of the unfiltered Ps-RF further reinforces the performance of the entire CRISP-RF workflow (compare Figure 6e and 6b). In the traditional average stack of the unfiltered Ps-RF, it is difficult to distinguish between multiples, direct conversions, and other incoherent arrivals. However, for the CRISP-RF migrated and phase-weighted stack, only the clear arrivals, i.e., Moho and MLD, are visible (Figure 6e). Interpretation is therefore unambiguous.

5.2 MLD beneath the Yilgarn Craton (AU.KMBL)

A second example is from the long-running station located on the Yilgarn craton (AU.KMBL), where a mid-lithosphere discontinuity has previously been detected using the Sp-RF and autocorrelation approach (Birkey et al., 2021; Ford et al., 2010; Kennett, 2015; Sun, Fu, et al., 2018; Sun, Kennett, et al., 2018; Sun & Kennett, 2016). None of these previous observations apply the Ps-RF technique since published results observe a sharp

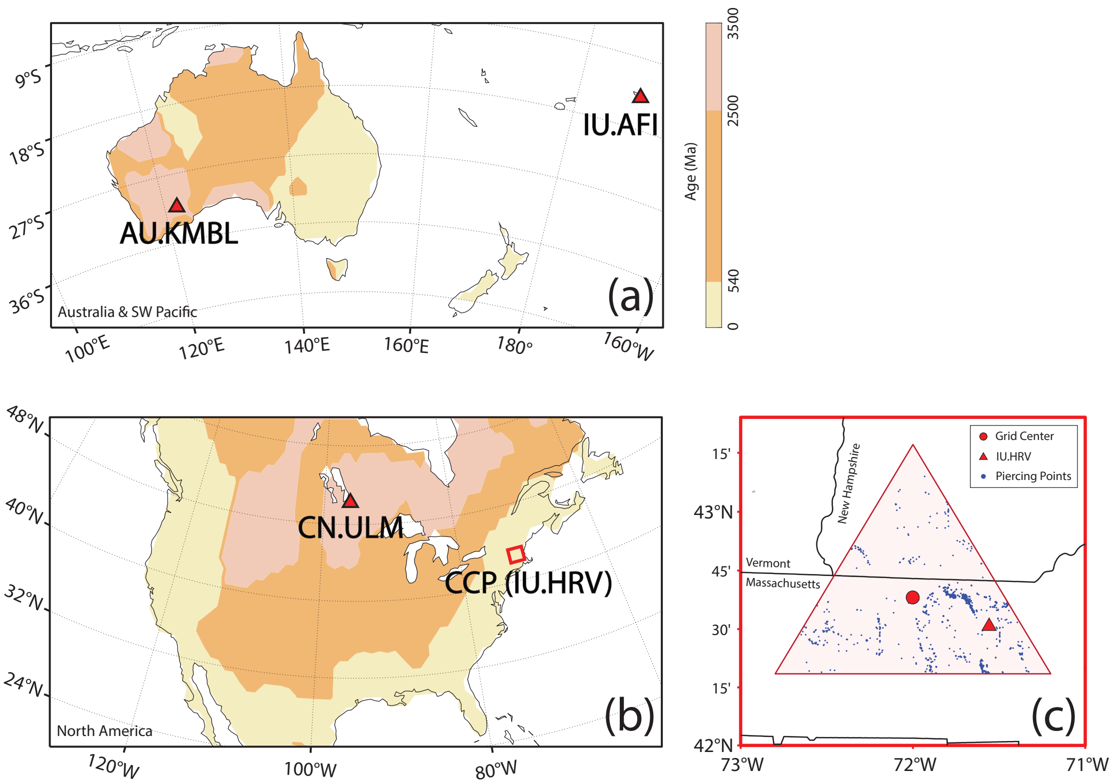

Figure 5. The location of four long-running seismic stations used in computing receiver function stacks. (a) Single-station analysis of the Yilgarn (AU.KMBL) and Samoa ocean island station (IU.AFI) (b) The Superior craton station (CN.ULM) and a virtual station located near the passive continental margin in the eastern US. This virtual station (CCP-IU.HRV) is processed using the common conversion point (CCP) analysis and located near long running station IU.HRV. (c) A close-up of the grid-center of the virtual station (red dot), the location of IU.HRV (red-triangle), and all the earthquake pierce-points at 50-km (blue dots). For a full azimuthal equidistant plot of all the earthquakes at each station see Figure S8 in Supporting Information.

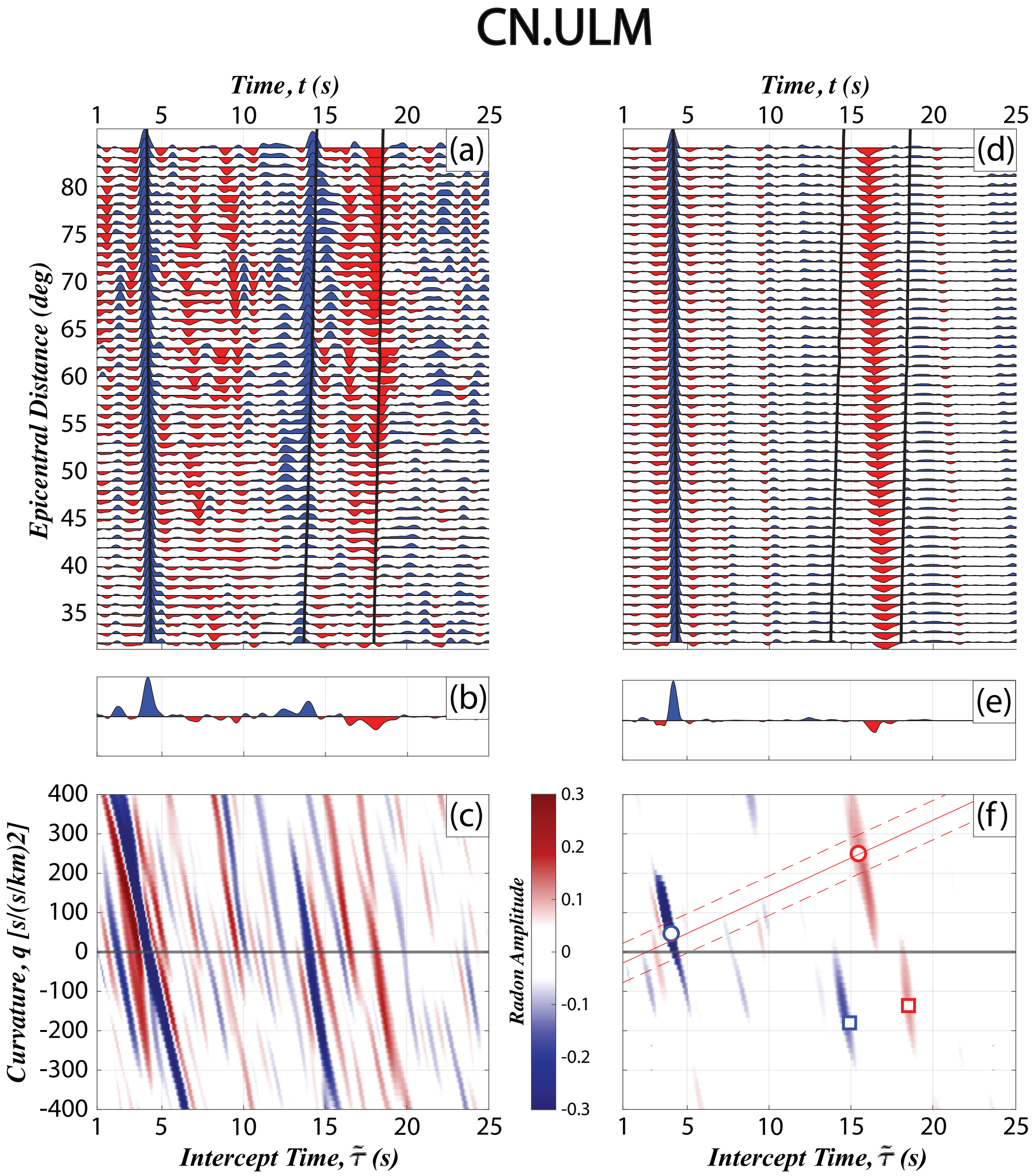

Figure 6. Improved detection of the mid-lithosphere discontinuity (MLD) beneath station CN.ULM using the CRISP-RF workflow. (a) The input Ps-RF data computed from the MTC algorithm. (b) The traditional average stack of the Ps-RF shown in (a). (c)The initial Radon model obtained from the least-squares optimization and used to initialize the SRTIS algorithm. (d) The filtered Ps-RF data obtained from the sparse radon model in (f) after filtering. (e) The final migrated and phase-weighted stack of the filtered Ps-RF in (d). (f) The sparse Radon model obtained after 30 iterations of the SRTIS algorithm.

Moho that generates very prominent crustal multiples which interfere with the MLD arrivals (H. Yuan, 2015). At this station, we select Ps-RF traces located at epicentral distances between 40° and 80° and stack them every 0.5° with 8° overlapping bins. Our Ps-RF confirms the crustal studies and highlights the difficulty of detecting upper mantle discontinuities when crustal multiples are present (Figure 7a).

The obtained Ps-RF and its average stack show clear Moho arrival and two prominent multiples at the predicted arrival times, ∼4 s, ∼15 s, and ∼19 s, calculated using the AuSREM crustal reference model (Kennett et al., 2023; Salmon et al., 2012), and therefore an attempt to visually identify the MLD phase in the stacked, unfiltered, and unmigrated Ps-RFs is very challenging (Figure 7a-b). There are a few arrivals between the Moho and its multiples, but it is hard to judge which ones have the correct move-out and coherence to be identified as the MLD. This is overcome by transforming the Ps-RF into a sparse Radon model (Figure 7f) using similar processing steps described earlier, i.e., 30 iterations of the SRTIS algorithm initialized from the damped least-squares solution (Figure 7c). The sparse Radon image clearly shows the Moho (blue circle) and its multiples (blue and red squares) at the appropriate intercept-time and curvature, and a coherent MLD phase (red circle) at ∼12 s (∼80 km). The other phases between the Moho and the multiples (∼4 - 15 s) observed on the Ps-RF (Figure 7a), though prominent, are incoherent, and do not map into either half of the sparse Radon model (Figure 7f). After appropriate filtering with a diagonal masking filter that eliminates all arrivals except those that follow the predicted curvature (red lines in Figure 7f), the resulting Ps-RF from the adjoint Radon transform (Figure 7d) and the final migrated and phase-weighted stack (Figure 7e) show only two major direct conversion phases. A comparison of the filtered and unfiltered stack makes it clear which of the arrivals is a coherent MLD with the appropriate moveout and phase-coherence (compare Figure 7e and 7f).

5.3 MLD beneath a Passive Continental Margin with CCP (CCP-IU.HRV)

Although receiver functions are sometimes estimated beneath single stations, the advent of large arrays makes it more likely that they will be processed beneath a virtual station using a common conversion point (CCP) scheme (Dueker, 1997; Rondenay, 2009). In this example, we show that the CRISP-RF workflow can be applied to Ps-RFs obtained using such a scheme. We calculate Ps-RFs using a virtual station with coordinates located close to the station IU.HRV, which is sited on a passive continental margin (Figure 5c). Previous Sp- and Ps-RF imaging at this location suggests that the crust and upper mantle structure is not laterally heterogeneous at the scale of the CCP-grid-size chosen for our analysis (Abt et al., 2010; Rychert et al., 2007). We select Ps-RF traces located at epicentral distances between 30° and 80° and stack them every 1° with 10° overlapping bins. The Ps-RFs we obstain are similar to those observed by the earlier studies (Figure 8a). However, without CRISP-RF processing, the Ps-RFs and its average stack are hard to interpret, with many coherent phases being visible making it is difficult to determine, by visual inspection alone, which of the coherent phases is from an upper mantle discontinuity (Figure 8a-b). After applying the CRISP-RF processing steps, the resulting sparse Radon model (Figure 8d) shows clear direct arrivals of the Moho (∼3 s) and the MLD (∼7 s), as well as the two crustal multiples (∼12 s and ∼16 s). The multiples are visibly attenuated in the final migrated and phase-weighted stack of Ps-RF, along with most of the noisy and some of the coherent phases (Figure 8e). The coherent phases that are being eliminated are those with a move-out that does not follow the predicted curvature for direct conversions; only the direct conversions with the move-out correctly modeled by the analytical equations (Equation 4) are retained.

5.4 X-Discontinuity beneath Samoa (IU.AFI)

Our final example targets the detection of the X-discontinuity, which is a deeper upper mantle discontinuity marked by a sharp velocity increase and is generally located at the

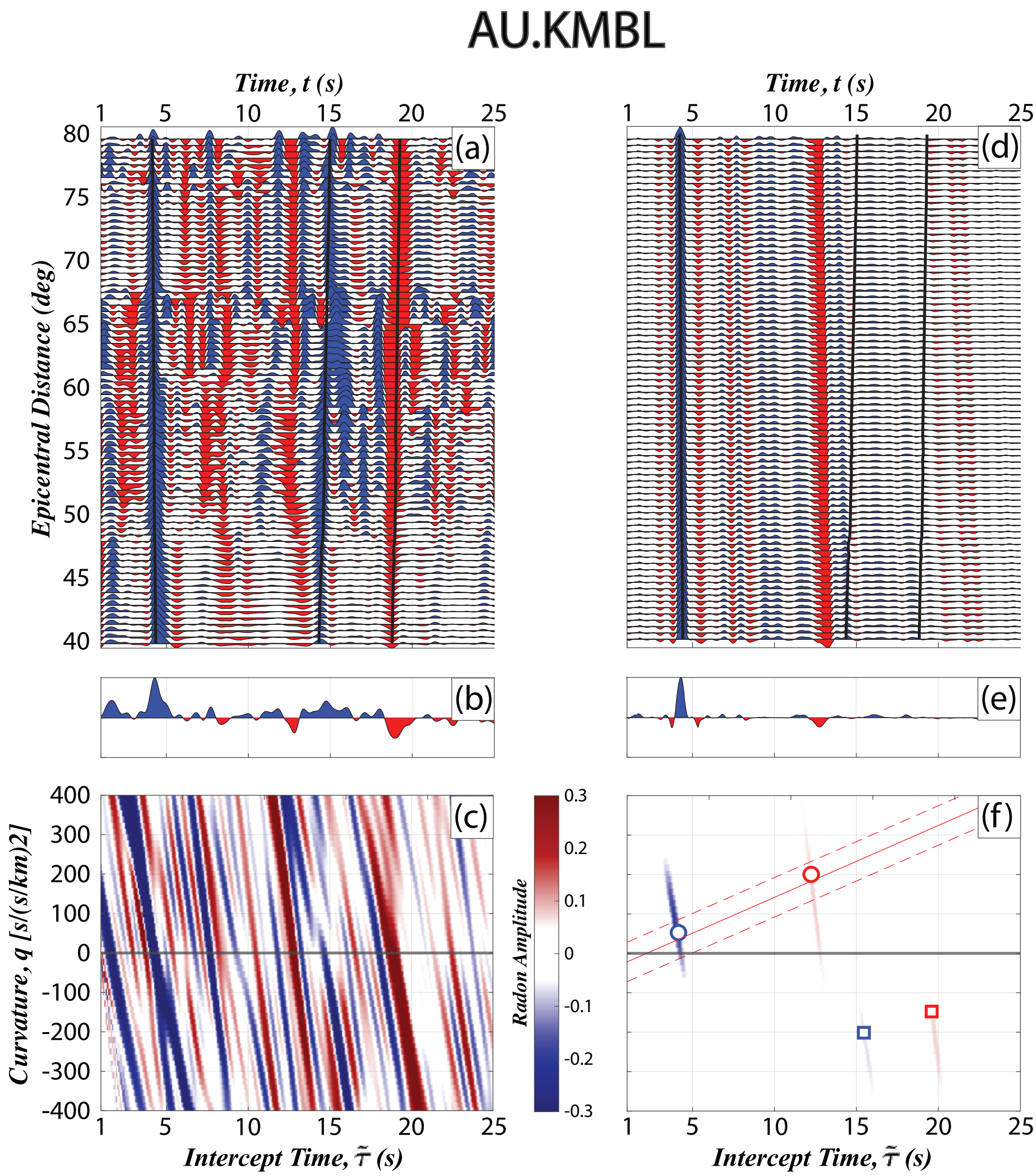

Figure 7. Same as Figure 6 but for station AU.KMBL.

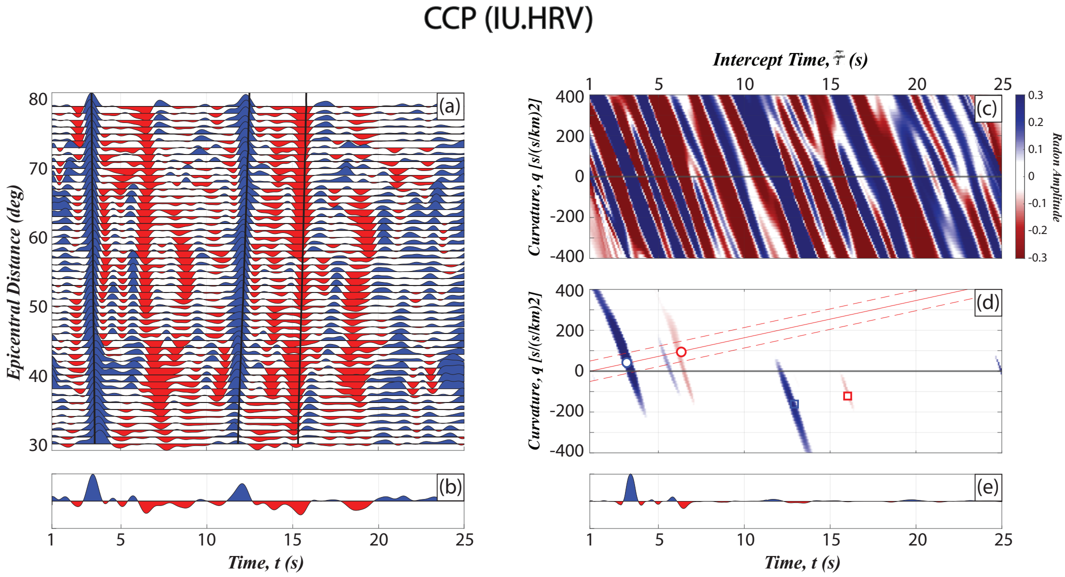

Figure 8. Improved detection of the MLD near the passive continental margin station (IU.HRV) using the CRISP-RF workflow applied to a CCP Ps-RF. (a) The input Ps-RF computed from the MTC algorithm using all earthquakes passing through the CCP grid shown in Figure 5c. (b) The traditional average stack of the Ps-RF shown in (a). (c) The initial Radon model obtained from the least-squares optimization and used to initialize the SRTIS algorithm. (d) The sparse Radon model obtained after 30 iterations of the SRTIS algorithm. The sparse Radon model shows the separation of direct conversions (open circles) from crustal multiples (open squares). The diagonal masking filter (red lines) is used to eliminate crustal multiples and retains the direct phases. (e) The final migrated and phase-weighted stack of the Ps-RF obtained from the adjoint Radon transform of the sparse Radon model shown in (d).

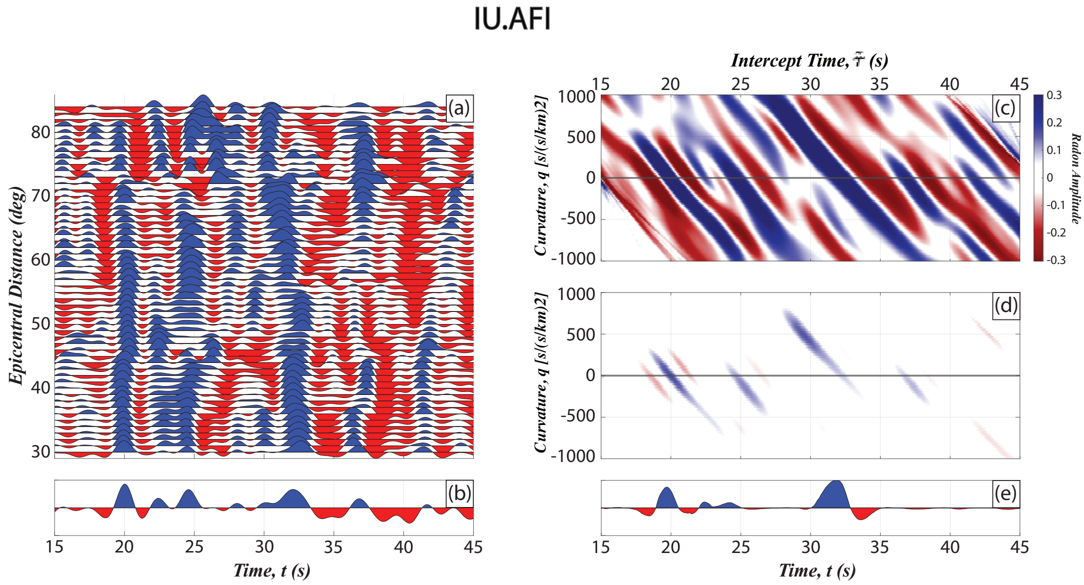

Figure 9. Same as Figure 8 but for station IU.AFI. Note that the Ps-RFs are shifted 30 s to target deeper upper mantle discontinuities.

depth range of 230 to 350 km (Pugh et al., 2021; Schmerr, 2015; Srinu et al., 2021). In this data example, we use teleseismic data from a permanent GSN station (IU.AFI) located at Samoa near the convergent boundary between the Pacific and Australian Plates. We apply a time-shift of 30 s to the radial seismogram and calculate the Ps-RF using the MTC algorithm at a cutoff frequency of 0.6 Hz (Frazer & Park, 2021; Park & Levin, 2016). This pre-processing step eliminates all crustal conversions, their multiples, and shallow upper mantle discontinuities that arrive earlier, and targets only deeper upper mantle discontinuities (see Figure 2). We select Ps-RF traces located at epicentral distances between 30° and 84° and stack them every 1° with 10° overlapping bins. The unfiltered Ps-RFs and the average stack show multiple positive phases from 18 s to 40 s (Figure 9a-b), making it hard to judge which are from mantle discontinuities. After applying the CRISP-RF methodology (Figure 9c-d), the final migrated and phase-weighted stack of the Ps-RF shows clear arrivals with positive curvature at ∼20 s and ∼32 s which we interpret as the Lehmann discontinuity and the X-discontinuity (Figure 9e).

6 Discussion

6.1 Comparing the Sparse Non-linear Radon Filters and Vespagrams

The sparse non-linear Radon filter we have implemented here bears some resemblance to other stacking techniques widely used for the global detection of upper mantle discontinuities. For example, the Radon transform (sparse or otherwise; Equations 7 and 8) is a high-resolution generalization of the time-domain delay-sum algorithm, which is a central idea in array-based seismology and is used to improve the detection of low-amplitude phases buried in random stochastic noise (Chapman, 1981; Gu & Sacchi, 2009; Krüger et al., 1993; Rost, 2002). In the slowness slant-stack analysis, also called a vespagram, the time delay of the different phases is a linear function of slowness (or ray parameter) and the delay-sum is calculated for different ray parameters effectively transforming the data into a $\tau - p$ Radon model (similar to our $\tilde{\tau} - q$). This approach has been widely used for imaging discontinuities within the mantle and across the core-mantle boundary, by improving the detection of precursors to global body-wave phases: P'P', PP, SS, and S-P converted waves. The detection of the weak precursor phases over and above other global seismic phases arriving within the same time window is improved by stacking with the appropriate time-slowness move-out (Davies et al., 1971; Deuss, 2009; Kawakatsu & Niu, 1994; Rost, 2002; Rost & Thomas, 2009; Schultz & Gu, 2013; Waszek et al., 2021). Based on the time-slowness move-out, the precursor phases are separable from that of other global phases (e.g., SS, PP, etc.) because the rays follow different paths and travel at different speeds (slownesses) through the mantle from source to receiver. Our implementation here can be viewed as a curvature slant-stack, where the Ps-RF is Radon-transformed using the time-curvature move-out, separating the direct Ps conversions from crustal multiples because the rays, with the same slowness, follow different paths only at the receiver-side (Figure 1 and Equation 4).

Recent extensions of the slowness slant stack methodology for global body-wave imaging improve resolution by incorporating the notion of the time-and-space locality as well as phase-coherence before stacking (Ventosa et al., 2012; Ventosa & Romanowicz, 2015a, 2015b; Zheng et al., 2015). Similar ideas have been applied to Ps-RFs in many variations (Guan & Niu, 2017; Gurrola et al., 1994; Shi et al., 2020), all borrowing slightly from exploration seismology, where velocity spectral analysis is used to disentangle phases, given a known earth model (O. Yilmaz, 1987). What distinguishes our approach is that, unlike the slowness slant stack technique which is a time-domain approach, the frequency-domain or mixed time-frequency Radon transform method is invertible, band-limited, and leads to higher-resolution Radon models (An et al., 2007; Gu et al., 2009; Gu & Sacchi, 2009; Schultz & Gu, 2013; Schultz & Jeffrey Gu, 2013; Wilson & Guitton, 2007) that improve the detection and isolation of direct phases buried within multiple reflected phases at the receiver side. The slowness slant stack approach, unlike the Radon transform, implements the essential delay-sum step in the time domain. In contrast, the Radon transform implements the time-shift delay step in the frequency domain using operators that benefit from the phase-shift property of the Fourier transform (Equations 5 and 9).

Implementing the delay-sum in the frequency domain provides two key advantages: (1) improved time resolution through frequency domain interpolation for time-shifts that are non-integer multiples of the sampling interval, and (2) taking advantage of frequency-dependence of the signal-to-noise and variance estimates useful in data preconditioning and regularization (Park & Levin, 2000, 2016). Our application of the Radon transform to receiver-side imaging with converted teleseismic waves is similar to ideas proposed by other authors (Aharchaou & Levander, 2016; Chen et al., 2022; Gu et al., 2015; Wilson & Guitton, 2007; Q. Zhang et al., 2021, 2022). However, these implementations differ from ours in some key aspects: (1) they often implement a low-resolution least-squares Radon solution (Gu et al., 2015; Q. Zhang et al., 2022), (2) focus on removing random incoherent noise either in the raw seismogram or the post-processed receiver function traces (Aharchaou & Levander, 2016; Q. Zhang et al., 2021, 2022), or (3) are applied solely as an aid to migration and data interpolation (Gu et al., 2015). In the study closest to ours and dedicated solely to multiple attenuation, the least-squares parabolic Radon transform is the recommended algorithm (Chen et al., 2022). In our treatment here, we have shown that a sparse high-resolution Radon transform, implemented using recent advances in optimization theory, is preferable, and is able to improve the detection of upper mantle discontinuities, especially in the presence of complex noise models (Figure 4). Additionally, the sparse Radon transform we have developed, sits within an end-to-end CRISP-RF signal processing workflow that exclusively targets mantle conversion and can achieve all our stated goals: sparse-recovery for slowness-interpolation, sub-crustal imaging, multiple removal, and denoising using selective masking filters that are informed using suitable reference earth models (Figure 2).

6.2 Benefits of CRISP-RF for Imaging Sharp UMDs: MLD, Lehmann, X, Melt

We advocate the use of the sparse Radon transform when high-resolution Ps-RF imaging of a sharp upper mantle discontinuity is required. Our analysis suggests that by passing the Ps-RF through the CRISP-RF workflow, multiples, generated at shallow interfaces, which mask the target upper mantle discontinuities, can be attenuated without compromising on signal quality and spatial resolution of structural features. This addresses the main disadvantage of Ps-RFs compared to Sp-RFs (Kind et al., 2012; Kind & Yuan, 2018; X. Yuan et al., 2006) and makes it possible to use both techniques in a joint-inversion scheme for investigating the sharpness of upper mantle discontinuities (Olugboji et al., 2013; Rychert et al., 2005, 2007). As a comparison, the CRISP-RF performs the task of removing multiples in the crust using a sparse Radon transform while a recently developed technique, FADER (Fast Automated Detection and Elimination of Echoes and Reverberations), removes repeating echoes in the shallow reverberant layers (sediments, oceans, or glaciers) using a homomorphic transform (Z. Zhang & Olugboji, 2021, 2023). Both techniques model the behavior of reverberations using appropriate transforms that separate the unwanted wavefield contribution from the signal of interest: crustal multiples (single echoes) are separated in a Radon-transformed domain while the reverberations in resonant layers (repeating echoes) are separated in a homomorphic-transformed domain. In practice, the interference of shallow crustal multiples is most severe when applying Ps-RFs to upper mantle imaging in the depth-range of the mid-lithosphere discontinuity ($\sim$60 - 170 km) as it is strongly overprinted by crustal multiples from a sharp Moho or intra-crustal boundary (Figures 1d-e and 3). When interpreting Ps-RFs for structural features at mid-lithosphere depths (Ford et al., 2016; Luo et al., 2021; Wirth & Long, 2014), confusion can be avoided by passing the single station or CCP Ps-RFs through a Radon transform like ours, or a slowness-weighted stack (Guan & Niu, 2017; Pugh et al., 2021, 2023) before interpretation.

For the other upper mantle discontinuities, e.g., Lehmann and X-discontinuity, due to their later arrival times, it is less likely that the Ps-RF will suffer interference from crustal or shallow lithospheric multiples. In this case, the CRISP-RF workflow can be beneficial to improving robust detection of discontinuities by serving as a denoiser and aiding in sparse signal recovery (Figure 9). We point out that in most recent applications of time-domain slowness slant stack in Ps-RF imaging, the linear moveout is assumed instead of the parabolic equations used in our implementation (Guan & Niu, 2017; Pugh et al., 2021, 2023; Srinu et al., 2021). To the best of our understanding, our implementation of the sparse Radon transform with the mixed time-frequency iterative solvers, using a suite of modern compressive sensing algorithms, is the most complete treatment of this problem for global upper mantle imaging with Ps-RFs.

6.3 Current Limitations of CRISP-RF and Future Work

The most challenging part of the CRISP-RF workflow is in the selection and tuning of the algorithms that implement the sparse Radon transform. Until now, we have been agnostic about which algorithm to use and have presented, in the Supporting Information, a detailed comparison of three different methods that can be utilized to compute the sparse Radon transform. Our comparison includes the investigation of the visual quality of the final Radon model, examination of the convergence behavior of each algorithm, a comparison of their wall-clock run time, and a discussion on the parameter tuning problem. Based on our comprehensive analysis, both on synthetic and real data, we have made several key observations that might be valuable in practice: (1) we observe that all three of the methods provide visually appealing sparse Radon models, with the difference that the output of each method is slightly different than the others while sharing some common structures; (2) we observe that all of the methods converge to a fixed point within a moderate number of iterations; however, employing an early stopping is needed to achieve fixed point convergence for the SRTIS algorithm due to its heuristic nature; (3) we observe that the run time of

SRTFISTA is significantly less than that of the SRTIS and $SRTL_{1-2}$ algorithms, making SRTFISTA the preferable method under run-time constraints if there is any; (4) to make the algorithms work in practice, we need to tune their parameters carefully.

If one assumes that the number of iterations for each algorithm is fixed, then for SRTIS, SRTFISTA, and $SRTL_{1-2}$, we have to adjust 2, 1, and 3 parameters, respectively. We observe that having only one parameter to tune makes SRTFISTA desirable when trying several different parameter combinations is computationally prohibitive, e.g., in the case of large data arrays with limited computational resources. For instance, in our experiments on real data, we observe that performing a 20-point grid search for SRTFISTA is almost 20 times faster than performing a 5-point grid search for SRTIS and $SRTL_{1-2}$ algorithms. Based on our observations, we suggest that, in practice, any of these methods can be used if the computational budget is not an issue and if there are no run-time constraints. On the other hand, if there is a strict computational budget or a certain run-time requirement, we suggest the use of SRTFISTA.

Through extensive experiments, we observe that incorporating the notion of sparsity into the reconstruction and filtering problem of Ps-RF imaging has led to significant improvements over the traditional methods. We believe that principled utilization of machine learning methods can further advance the state-of-the-art. For example, machine learning methods such as reinforcement learning (Sutton & Barto, 2018) can be used to tune the parameters of the iterative reconstruction algorithms automatically, which may accelerate the processing of large datasets since human expert involvement will be minimized. In the future, instead of being limited to using simple, analytic regularizers such as ℓ_1-regularization, one may be able to use deep learning techniques such as deep algorithmic unrolling (Monga et al., 2021), to learn more complex regularizers from an ensemble of large datasets. Moreover, conceiving the sparse Radon transform as an optimization problem, we may be able to use some of the state-of-the-art deep neural networks designed to perform regression, thereby reducing the run-time of the reconstruction. In the age of increasing computational power and parallelization of modern GPUs we may also be able to learn robust uncertainty information from applying generative models to the denoising and attenuation problem (Bond-Taylor et al., 2022). In follow-on studies, we envision that these new ideas will enable high-resolution Ps-RF imaging of the upper mantle using large Ps-RF datasets obtained from large seismic arrays, e.g., in North America (Long et al., 2014; Shearer & Buehler, 2019) and in Africa (Olugboji & Xue, 2022).

7 Conclusions

We have developed a novel method, CRISP-RF, for generating clean Ps-RFs free of unwanted interferences, i.e., waves reverberating in the crust and incoherent noise. We show, using synthetic and real data examples, how the high-resolution sparse Radon transform facilitates the successful elimination of the unwanted signals. We review different methods for solving the Radon transform, and show that sparse recovery of the Radon model using the iterative shrinkage algorithm is preferred and outperforms the conventional least-squares approach. Higher resolution denoised Ps-RF imaging with the crustal multiples removed will result in a more accurate characterization of upper mantle structure. This improved imaging capability sets a new standard for seismic studies, with future applications in regional and large-scale array configurations. We anticipate future application the CRISP-RF philosophy to imaging using the full body-wave field, including top-side and bottom-side reflections (e.g., SS and PP precursors, and SsdS reflections), which will extend the range of mantle imaging through to the mid- and lower-mantle.

Data Availability Statement

All seismic data used in this study can be obtained from the IRIS Data management center (https://ds.iris.edu/ds). The code used for receiver function deconvolution and the

CRISP-RF workflow can be retrieved from the open source repository at doi: 10.5281/zenodo.7996504.

Acknowledgments

This work was made possible by a seed grant from the University of Rochester's Goergen Institute for Data Science and support from the National Science Foundation under grant number: 1818654. The authors acknowledge the use of the BlueHive Linux cluster at the University of Rochester's Center for Integrated Research Computing, CIRC (https://www.circ.rochester.edu/). The authors acknowledge many helpful discussions with Lara Wagner, Baowei Liu, Jeffrey Park, Shun-ichiro Karato, and Gary Egbert.

References

Abt, D. L., Fischer, K. M., French, S. W., Ford, H. a., Yuan, H., & Romanowicz, B. (2010, September). North american lithospheric discontinuity structure imaged by ps and sp receiver functions. *Journal of geophysical research*, *115*(B9), 1–24. doi: 10.1029/2009JB006914

Aharchaou, M., & Levander, A. (2016, August). A compressive sensing approach to the high-resolution linear radon transform: application on teleseismic wavefields. *Geophysical Journal International*, *207*(2), 811–822. doi: 10.1093/gji/ggw307

An, Y., Gu, Y. J., & Sacchi, M. D. (2007, October). Imaging mantle discontinuities using least squares radon transform. *Journal of geophysical research*, *112*(B10). doi: 10.1029/2007jb005009

Aster, R. C., Borchers, B., & Thurber, C. H. (2018). *Parameter estimation and inverse problems*. Elsevier.

Aulbach, S. (2018). *Cratonic lithosphere discontinuities*. doi: 10.1002/9781119249740.ch10

Beck, A., & Teboulle, M. (2009, January). A fast iterative Shrinkage-Thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, *2*(1), 183–202. doi: 10.1137/080716542

Beghein, C., Yuan, K., Schmerr, N., & Xing, Z. (2014, March). Changes in seismic anisotropy shed light on the nature of the gutenberg discontinuity. *Science*, *343*(6176), 1237–1240. doi: 10.1126/science.1246724

Benz, H. M., & Vidale, J. E. (1993, September). Sharpness of upper-mantle discontinuities determined from high-frequency reflections. *Nature*, *365*(6442), 147–150. doi: 10.1038/365147a0

Beylkin, G. (1987, February). Discrete radon transform. *IEEE transactions on acoustics, speech, and signal processing*, *35*(2), 162–172. doi: 10.1109/TASSP.1987.1165108

Birkey, A., Ford, H. A., Dabney, P., & Goldhagen, G. (2021, April). The lithospheric architecture of australia from seismic receiver functions. *Journal of Geophysical Research, [Solid Earth]*, *126*(4). doi: 10.1029/2020jb020999

Bond-Taylor, S., Leach, A., Long, Y., & Willcocks, C. G. (2022, November). Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, Energy-Based and autoregressive models. *IEEE transactions on pattern analysis and machine intelligence*, *44*(11), 7327–7347. doi: 10.1109/TPAMI.2021.3116668

Bostock, M. G. (2004). Theory and observations -teleseismic Body-Wave scattering and Receiver-Side structure. *Treatise in Geophysics*, 235–246.

Boyd, S. (2010). Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, *3*(1), 1–122. doi: 10.1561/2200000016

Buehler, J. S., & Shearer, P. M. (2017). Uppermost mantle seismic velocity structure beneath USArray. *Journal of Geophysical Research, [Solid Earth]*, *122*(1), 436–448.

Candès, E. J., & Wakin, M. B. (2008, March). An introduction to compressive sampling. *IEEE Signal Processing Magazine*, *25*(2), 21–30. doi: 10.1109/MSP.2007.914731

Chapman, C. H. (1981, August). Generalized radon transforms and slant stacks. *Geophysical Journal International*, *66*(2), 445–453. doi: 10.1111/j.1365-246X.1981.tb05966.x

Chen, Y., Chen, J., Guo, B., Li, S., Li, Y., Qi, S., & Zhao, P. (2022, March). The removal of multiple reflection waves in the P receiver function through parabolic radon transformation. *Geophysical Journal International*, *230*(2), 1052–1064. doi: 10.1093/gji/ggac097

Davies, D., Kelly, E. J., & Filson, J. R. (1971, July). Vespa process for analysis of seismic signals. *Nature Physical Science*, *232*(27), 8–13. doi: 10.1038/physci232008a0

Deans, S. R. (2007). *The radon transform and some of its applications*. Courier Corporation.

Deuss, A. (2009, October). Global observations of mantle discontinuities using SS and PP precursors. *Surveys in Geophysics*, *30*(4), 301–326. doi: 10.1007/s10712-009-9078-y

Deuss, A., & Woodhouse, J. H. (2004, September). The nature of the lehmann discontinuity from its seismological clapeyron slopes. *Earth and planetary science letters*, *225*(3), 295–304. doi: 10.1016/j.epsl.2004.06.021

Dokht, R. M. H., Gu, Y. J., & Sacchi, M. D. (2016, December). Waveform inversion of SS precursors: An investigation of the northwestern pacific subduction zones and intraplate volcanoes in china. *Gondwana Research*, *40*, 77–90. doi: 10.1016/j.gr.2016.07.006

Dueker, K. G. (1997). *Mantle discontinuity structure from midpoint stacks of converted P to S waves across the yellowstone hotspot track*. American Geophysical Union.

Eaton, D. W., Darbyshire, F., Evans, R. L., Grütter, H., Jones, A. G., & Yuan, X. (2009, April). The elusive lithosphere–asthenosphere boundary (LAB) beneath cratons. *Lithos*, *109*(1-2), 1–22. doi: 10.1016/j.lithos.2008.05.009

Fischer, K. M. (2015, January). 1.18 - crust and lithospheric structure - seismological constraints on the Lithosphere-Asthenosphere boundary. In G. Schubert (Ed.), *Treatise on geophysics (second edition)* (pp. 587–612). Oxford: Elsevier. doi: 10.1016/B978-0-444-53802-4.00026-9

Fischer, K. M., Ford, H. A., Abt, D. L., & Rychert, C. A. (2010, April). The Lithosphere-Asthenosphere boundary. *Annual review of earth and planetary sciences*, *38*(1), 551–575. doi: 10.1146/annurev-earth-040809-152438

Fischer, K. M., Rychert, C. A., Dalton, C. A., Miller, M. S., & others. (2020). A comparison of oceanic and continental mantle lithosphere. *Physics of the Earth and Planetary Interiors*.

Ford, H. A., Fischer, K. M., Abt, D. L., Rychert, C. A., & Elkins-Tanton, L. T. (2010, December). The lithosphere–asthenosphere boundary and cratonic lithospheric layering beneath australia from sp wave imaging. *Earth and planetary science letters*, *300*(3), 299–310. doi: 10.1016/j.epsl.2010.10.007

Ford, H. A., Long, M. D., & Wirth, E. A. (2016, September). Midlithospheric discontinuities and complex anisotropic layering in the mantle lithosphere beneath the wyoming and superior provinces. *Journal of Geophysical Research, [Solid Earth]*, *121*(9), 6675–6697. doi: 10.1002/2016JB012978

Frazer, W. D., & Park, J. (2021, October). Seismic evidence of mid-mantle water transport beneath the yellowstone region. *Geophysical research letters*, *48*(20). doi: 10.1029/2021gl095838

Galetti, E., & Curtis, A. (2012, April). Generalised receiver functions and seismic interferometry. *Tectonophysics*, *532-535*, 1–26. doi: 10.1016/j.tecto.2011.12.004

Geng, W., Chen, X., Li, J., Ma, J., Tang, W., & Wu, F. (2022, September). Sparse radon transform in the mixed frequency-time domain with 1-2 minimization. *Geophysics*, *87*(5), V545–V558. doi: 10.1190/geo2021-0343.1

Gómez-García, C., Lebedev, S., Meier, T., Xu, Y., Le Pape, F., & Wiesenberg, L. (2022, December). Ambient noise autocorrelation scheme for imaging the p-wave reflectivity of the lithosphere. *Geophysical Journal International*, *233*(3), 1671–1693. doi: 10.1093/gji/ggac509

Gong, X., Yu, S., & Wang, S. (2016, August). Prestack seismic data regularization using a time-variant anisotropic radon transform. *Journal of Geophysics and Engineering*, *13*(4), 462–469. doi: 10.1088/1742-2132/13/4/462

Gu, Y. J., An, Y., Sacchi, M., Schultz, R., & others. (2009). Mantle reflectivity structure beneath oceanic hotspots. *Geophysical Journal*.

Gu, Y. J., & Sacchi, M. (2009, October). Radon transform methods and their applications in mapping mantle reflectivity structure. *Surveys in Geophysics*, *30*(4), 327–354. doi: 10.1007/s10712-009-9076-0

Gu, Y. J., Zhang, Y., Sacchi, M. D., Chen, Y., & others. (2015). Sharp mantle transition from cratons to cordillera in southwestern canada. *Journal of geophysical research*. doi: 10.1002/2014JB011802

Guan, Z., & Niu, F. (2017, June). An investigation on slowness-weighted CCP stacking and its application to receiver function imaging: Slowness-Weighted CCP stacking. *Geophysical research letters*, *44*(12), 6030–6038. doi: 10.1002/2017gl073755

Gurrola, H., Minster, J. B., & Owens, T. (1994, May). The use of velocity spectrum for stacking receiver functions and imaging upper mantle discontinuities. *Geophysical Journal International*, *117*(2), 427–440. doi: 10.1111/j.1365-246X.1994.tb03942.x

Hampson, D. (1986, January). Inverse velocity stacking for multiple elimination. In *SEG technical program expanded abstracts 1986.* Society of Exploration Geophysicists. doi: 10.1190/1.1893060

Helffrich, G. (2006, February). Extended-Time multitaper frequency domain Cross-Correlation Receiver-Function estimation. *Bulletin of the Seismological Society of America*, *96*(1), 344–347. doi: 10.1785/0120050098

Hopper, E., & Fischer, K. M. (2018). The changing face of the lithosphere-asthenosphere boundary: Imaging continental scale patterns in upper mantle structure across the contiguous US with sp converted waves. *Geochemistry, Geophysics, Geosystems*, *19*(8), 2593–2614.

Ji, J. (2006, July). CGG method for robust inversion and its application to velocity-stack inversion. *Geophysics*, *71*(4), R59–R67. doi: 10.1190/1.2209547

Karato, S.-I. (1992, November). On the lehmann discontinuity. *Geophysical research letters*, *19*(22), 2255–2258. doi: 10.1029/92GL02603

Karato, S.-I., Olugboji, T., & Park, J. (2015, June). Mechanisms and geologic significance of the mid-lithosphere discontinuity in the continents. *Nature geoscience*, *8*(7), 509–514. doi: 10.1038/ngeo2462

Karato, S.-I., & Park, J. (2018, November). On the origin of the upper mantle seismic discontinuities. In H. Yuan & B. Romanowicz (Eds.), *Lithospheric discontinuities* (Vol. 115, pp. 5–34). Hoboken, NJ, USA: John Wiley & Sons, Inc. doi: 10.1002/9781119249740.ch1

Kawakatsu, H., Kumar, P., Takei, Y., Shinohara, M., Kanazawa, T., Araki, E., & Suyehiro, K. (2009, April). Seismic evidence for sharp lithosphere-asthenosphere boundaries of oceanic plates. *Science*, *324*(5926), 499–502. doi: 10.1126/science.1169499

Kawakatsu, H., & Niu, F. (1994, September). Seismic evidence for a 920-km discontinuity in the mantle. *Nature*, *371*(6495), 301–305. doi: 10.1038/371301a0

Kennett, B. L. N. (2015, December). Lithosphere–asthenosphere p-wave reflectivity across australia. *Earth and planetary science letters*, *431*, 225–235. doi: 10.1016/j.epsl.2015.09.039

Kennett, B. L. N., Gorbatov, A., Yuan, H., Agrawal, S., Murdie, R., Doublier, M. P., . . . Gessner, K. (2023, January). Refining the moho across the australian continent. *Geophysical Journal International*, *233*(3), 1863–1877. doi: 10.1093/gji/ggad035

Kind, R., Mooney, W. D., & Yuan, X. (2020, April). New insights into the structural elements of the upper mantle beneath the contiguous united states from S-to-P converted seismic waves. *Geophysical Journal International*. doi: 10.1093/gji/ggaa203

Kind, R., & Yuan, X. (2018). Perspectives of the S-Receiver-Function method to image upper mantle discontinuities. *Lithospheric Discontinuities*, 139–154.

Kind, R., Yuan, X., & Kumar, P. (2012). *Seismic receiver functions and the lithosphere–asthenosphere boundary* (Vol. 536-537). doi: 10.1016/j.tecto.2012.03.005

Krueger, H. E., Gama, I., & Fischer, K. M. (2021, June). Global patterns in cratonic mid-lithospheric discontinuities from sp receiver functions. *Geochemistry, Geophysics, Geosystems*, *22*(6). doi: 10.1029/2021gc009819

Krüger, F., Weber, M., Scherbaum, F., & Schlittenhardt, J. (1993, July). Double beam analysis of anomalies in the core-mantle boundary region. *Geophysical research letters*, *20*(14), 1475–1478. doi: 10.1029/93GL01311

Kumar, P., Yuan, X., Kind, R., & Mechie, J. (2012, May). The lithosphere-asthenosphere boundary observed with USArray receiver functions. *Solid Earth*, *3*(1), 149–159. doi: 10.5194/se-3-149-2012

Lekić, V., & Fischer, K. M. (2017, August). Interpreting spatially stacked sp receiver functions. *Geophysical Journal International*, *210*(2), 874–886. doi: 10.1093/gji/ggx206

Liu, T., & Shearer, P. M. (2021, May). Complicated lithospheric structure beneath the contiguous US revealed by teleseismic s-reflections. *Journal of Geophysical Research, [Solid Earth]*, *126*(5). doi: 10.1029/2020jb021624

Long, M. D., Levander, A., & Shearer, P. M. (2014, September). An introduction to the special issue of earth and planetary science letters on USArray science. *Earth and planetary science letters*, *402*, 1–5. doi: 10.1016/j.epsl.2014.06.016

Lu, W. (2013, July). An accelerated sparse time-invariant radon transform in the mixed frequency-time domain based on iterative 2D model shrinkage. *Geophysics*, *78*(4), V147–V155. doi: 10.1190/geo2012-0439.1

Luo, Y., Long, M. D., Karabinos, P., Kuiper, Y. D., Rondenay, S., Aragon, J. C., ... Makus, P. (2021, July). High-resolution PS receiver function imaging of the crust and mantle lithosphere beneath southern new england and tectonic implications. *Journal of Geophysical Research, [Solid Earth]*, *126*(7). doi: 10.1029/2021jb022170

Luo, Y., Xia, J., Miller, R. D., Xu, Y., Liu, J., & Liu, Q. (2008, May). Rayleigh-Wave dispersive energy imaging using a High-Resolution linear radon transform. *Pure and Applied Geophysics*, *165*(5), 903–922. doi: 10.1007/s00024-008-0338-4

Mancinelli, N. J., Fischer, K. M., & Dalton, C. A. (2017). How sharp is the cratonic Lithosphere-Asthenosphere transition? *Geophysical research letters*, *44*(20), 10,189–10,197. doi: 10.1002/2017GL074518

Menke, W. (2012). Describing inverse problems. In *Geophysical data analysis: Discrete inverse theory* (pp. 1–14). Elsevier. doi: 10.1016/B978-0-12-397160-9.00001-1

Monga, V., Li, Y., & Eldar, Y. C. (2021, March). Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing. *IEEE Signal Processing Magazine*, *38*(2), 18–44. doi: 10.1109/msp.2020.3016905

Olugboji, T., Karato, S., & Park, J. (2013, April). Structures of the oceanic lithosphere-asthenosphere boundary: Mineral-physics modeling and seismological signatures. *Geochemistry, Geophysics, Geosystems*, *14*(4), 880–901. doi: 10.1002/ggge.20086

Olugboji, T., & Park, J. (2016, March). Crustal anisotropy beneath pacific Ocean-Islands from harmonic decomposition of receiver functions. *Geochemistry, Geophysics, Geosystems*, *17*(3), 810–832. doi: 10.1002/2015GC006166

Olugboji, T., & Xue, S. (2022, May). A short-period surface-wave dispersion dataset for model assessment of africa's crust: ADAMA. *Seismological Research Letters*, *93*(3), 1943–1959. doi: 10.1785/0220210355

Parikh, N. (2014). Proximal algorithms. *Foundations and trends® in optimization*, *1*(3), 127–239. doi: 10.1561/2400000003

Park, J., & Levin, V. (2000, December). Receiver functions from Multiple-Taper spectral correlation estimates. *Bulletin of the Seismological Society of America*, *90*(6), 1507–1520. doi: 10.1785/0119990122

Park, J., & Levin, V. (2016, October). Statistics and frequency-domain moveout for multiple-taper receiver functions. *Geophysical Journal International*, *207*(1), 512–527. doi: 10.1093/gji/ggw291

Petersen, N., Vinnik, L., Kosarev, G., Kind, R., & others. (1993). Sharpness of the mantle discontinuities. *Geophysical*. doi: 10.1029/93GL00684

Poli, P., Campillo, M., Pedersen, H., & LAPNET Working Group. (2012, November). Body-wave imaging of earth's mantle discontinuities from ambient seismic noise. *Science*, *338*(6110), 1063–1065. doi: 10.1126/science.1228194

Pugh, S., Boyce, A., Bastow, I. D., Ebinger, C. J., & Cottaar, S. (2023, March). Multigenetic origin of the x-discontinuity below continents: Insights from african receiver functions. *Geochemistry, Geophysics, Geosystems*, *24*(3). doi: 10.1029/2022gc010782

Pugh, S., Jenkins, J., Boyce, A., & Cottaar, S. (2021, May). Global receiver function observations of the x-discontinuity reveal recycled basalt beneath hotspots. *Earth and planetary science letters*, *561*, 116813. doi: 10.1016/j.epsl.2021.116813

Rader, E., Emry, E., Schmerr, N., Frost, D., Cheng, C., Menard, J., ... Geist, D. (2015, October). Characterization and petrological constraints of the midlithospheric discontinuity. *Geochemistry, Geophysics, Geosystems*, *16*(10), 3484–3504. doi: 10.1002/2015GC005943

Revenaugh, J., & Jordan, T. H. (1991, November). Mantle layering fromScSreverberations: 3. the upper mantle. *Journal of geophysical research*, *96*(B12), 19781–19810. doi: 10.1029/91jb01487

Rondenay, S. (2009, October). Upper mantle imaging with array recordings of converted and scattered teleseismic waves. *Surveys in Geophysics*, *30*(4), 377–405. doi: 10.1007/s10712-009-9071-5

Rost, S. (2002). Array seismology: Methods and applications. *Reviews of geophysics*, *40*(3), 305. doi: 10.1029/2000RG000100

Rost, S., & Thomas, C. (2009, October). Improving seismic resolution through array processing techniques. *Surveys in Geophysics*, *30*(4), 271–299. doi: 10.1007/s10712-009-9070-6

Ryberg, T., & Weber, M. (2000, April). Receiver function arrays: a reflection seismic approach. *Geophysical Journal International*, *141*(1), 1–11. doi: 10.1046/j.1365-246X.2000.00077.x

Rychert, C. A., Fischer, K. M., & Rondenay, S. (2005, July). A sharp lithosphere–asthenosphere boundary imaged beneath eastern north america. *Nature*, *436*(7050), 542–545. doi: 10.1038/nature03904

Rychert, C. A., Harmon, N., Constable, S., & Wang, S. (2020, September). The nature of the Lithosphere-Asthenosphere boundary. *Journal of Geophysical Research, [Solid Earth]*. doi: 10.1029/2018JB016463

Rychert, C. A., Rondenay, S., & Fischer, K. M. (2007, August). P -to- S and S -to- P imaging of a sharp lithosphere-asthenosphere boundary beneath eastern north america. *Journal of geophysical research*, *112*(B8), B08314. doi: 10.1029/2006JB004619

Rychert, C. A., & Shearer, P. M. (2009, April). A global view of the lithosphere-asthenosphere boundary. *Science*, *324*(5926), 495–498. doi: 10.1126/science.1169754

Sacchi, M. D., & Ulrych, T. J. (1995, July). High-resolution velocity gathers and offset space reconstruction. *Geophysics*, *60*(4), 1169–1177. doi: 10.1190/1.1443845

Salmon, M., Kennett, B. L. N., & Saygin, E. (2012, November). Australian seismological reference model (AuSREM): crustal component. *Geophysical Journal International*, *192*(1), 190–206. doi: 10.1093/gji/ggs004

Schmerr, N. (2015). Imaging mantle heterogeneity with upper mantle seismic discontinuities. In A. Khan & F. Deschamps (Eds.), *The earth's heterogeneous mantle: A geophysical, geodynamical, and geochemical perspective* (pp. 79–104). Cham: Springer International Publishing. doi: 10.1007/978-3-319-15627-9_3

Schultz, R., & Gu, Y. J. (2013, August). Multiresolution imaging of mantle reflectivity structure using SS and PP precursors. *Geophysical Journal International*, *195*(1), 668–683. doi: 10.1093/gji/ggt266

Schultz, R., & Jeffrey Gu, Y. (2013, March). Flexible, inversion-based matlab implementation of the radon transform. *Computers & geosciences*, *52*, 437–442. doi: 10.1016/j.cageo.2012.08.013

Schutt, D. L., Lowry, A. R., & Buehler, J. S. (2018, March). Moho temperature and mobility of lower crust in the western united states. *Geology*, *46*(3), 219–222. doi: 10.1130/G39507.1

Selway, K., Ford, H., & Kelemen, P. (2015). The seismic mid-lithosphere discontinuity. *Earth and planetary science letters*, *414*, 45–57. doi: 10.1016/j.epsl.2014.12.029

Shearer, P. M. (2000). Upper mantle seismic discontinuities. *GEOPHYSICAL MONOGRAPH-AMERICAN GEOPHYSICAL UNION*, *117*, 115–132.

Shearer, P. M., & Buehler, J. (2019, September). Imaging Upper-Mantle structure under USArray using Long-Period reflection seismology. *Journal of Geophysical Research, [Solid Earth]*, *124*(9), 9638–9652. doi: 10.1029/2019JB017326

Shi, J., Wang, T., & Chen, L. (2020, July). Receiver function velocity analysis technique and its application to remove multiples. *Journal of Geophysical Research, [Solid Earth]*. doi: 10.1029/2020JB019420

Shibutani, T., Ueno, T., & Hirahara, K. (2008). Improvement in the extended-time multitaper receiver function estimation technique. *Bulletin of the Seismological Society of America*, *98*(2), 812–816.

Srinu, U., Kumar, P., Haldar, C., Kumar, M. R., & others. (2021). X-Discontinuity beneath the indian Shield—Evidence for remnant tethyan oceanic lithosphere in the mantle. *Journal of*.

Sun, W., Fu, L.-Y., Saygin, E., & Zhao, L. (2018, February). Insights into layering in the cratonic lithosphere beneath western australia. *Journal of Geophysical Research, [Solid Earth]*, *123*(2), 1405–1418. doi: 10.1002/2017JB014904

Sun, W., Kennett, B., Zhao, L., & Fu, L.-Y. (2018). Continental lithospheric layering beneath stable, modified, and destroyed cratons from seismic daylight imaging. *Lithospheric Discontinuities*, 155–176.

Sun, W., & Kennett, B. L. N. (2016, June). Receiver structure from teleseisms: Autocorrelation and cross correlation. *Geophysical research letters*, *43*(12), 6234–6242. doi: 10.1002/2016gl069564

Sun, W., & Kennett, B. L. N. (2017, February). Mid-lithosphere discontinuities beneath the western and central north china craton. *Geophysical research letters*, *44*(3), 1302–1310. doi: 10.1002/2016gl071840

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning, second edition: An introduction.* MIT Press.

Tao, P. D., & An, L. T. H. (1998, May). A D.C. optimization algorithm for solving the Trust-Region subproblem. *SIAM journal on optimization: a publication of the Society for Industrial and Applied Mathematics*, *8*(2), 476–505. doi: 10.1137/S1052623494274313

Tauzin, B., Pham, T.-S., & Tkalčić, H. (2019, April). Receiver functions from seismic interferometry: a practical guide. *Geophysical Journal International*, *217*(1), 1–24. doi: 10.1093/gji/ggz002

Tharimena, S., Rychert, C., & Harmon, N. (2017). A unified continental thickness from seismology and diamonds suggests a melt-defined plate. *Science*, *357*(6351), 580–583.

Thorson, J. R., & Claerbout, J. F. (1985, December). Velocity-stack and slant-stack stochastic inversion. *Geophysics*, *50*(12), 2727–2741. doi: 10.1190/1.1441893

Trad, D. (2003, November). Interpolation and multiple attenuation with migration operators. *Geophysics*, *68*(6), 2043–2054. doi: 10.1190/1.1635058

Trad, D., Ulrych, T., & Sacchi, M. (2003, January). Latest views of the sparse radon transform. *Geophysics*, *68*(1), 386–399. doi: 10.1190/1.1543224

Ventosa, S., & Romanowicz, B. (2015a, October). Extraction of weak PcP phases using the slant-stacklet transform – II: constraints on lateral variations of structure near the core–mantle boundary. *Geophysical Journal International*, *203*(2), 1227–1245. doi: 10.1093/gji/ggv364

Ventosa, S., & Romanowicz, B. (2015b, February). Extraction of weak PcP phases using the slant-stacklet transform — i: method and examples. *Geophysical Journal International*, *201*(1), 207–223. doi: 10.1093/gji/ggv010

Ventosa, S., Simon, C., & Schimmel, M. (2012, March). Window length selection for optimum slowness resolution of the local-slant-stack transform. *Geophysics*, *77*(2), V31–V40. doi: 10.1190/geo2010-0326.1

Waszek, L., Tauzin, B., Schmerr, N. C., Ballmer, M. D., & Afonso, J. C. (2021, November).

A poorly mixed mantle transition zone and its thermal state inferred from seismic waves. *Nature geoscience*, *14*(12), 949–955. doi: 10.1038/s41561-021-00850-w

Wilson, C., & Guitton, A. (2007, January). Teleseismic wavefield interpolation and signal extraction using high-resolution linear radon transforms. *Geophysical Journal International*, *168*(1), 171–181. doi: 10.1111/j.1365-246X.2006.03163.x

Wirth, E. A., & Long, M. D. (2014). A contrast in anisotropy across mid-lithospheric discontinuities beneath the central united States—A relic of craton formation. *Geology*, 1–4.

Wright, J., & Ma, Y. (2022). *High-Dimensional data analysis with Low-Dimensional models: Principles, computation, and applications.* Cambridge University Press.

Yilmaz, O. (1987). Seismic data processing: Investigations in geophysics. *Society of Exploration Geophysicists*, *2*, 526.

Yilmaz, Ö. (2015, January). *References.* Society of Exploration Geophysicists. doi: 10.1190/1.9781560803300.refs

Yuan, H. (2015). Secular change in archaean crust formation recorded in western australia. *Nature geoscience*, *8*(10), 808–813. doi: 10.1038/ngeo2521

Yuan, X., Kind, R., Li, X., & Wang, R. (2006, May). The S receiver functions: synthetics and data example. *Geophysical Journal International*, *165*(2), 555–564. doi: 10.1111/j.1365-246X.2006.02885.x

Zhang, Q., Chen, Y., Zhang, F., & Chen, Y. (2022, March). Improving receiver function imaging with high-resolution radon transform. *Geophysical Journal International*, *230*(2), 1292–1304. doi: 10.1093/gji/ggac116

Zhang, Q., Wang, H., Chen, W., & Huang, G. (2021, January). A robust method for random noise suppression based on the radon transform. *Journal of Applied Geophysics*, *184*, 104183. doi: 10.1016/j.jappgeo.2020.104183

Zhang, Z., & Olugboji, T. (2021, May). The signature and elimination of sediment reverberations on submarine receiver functions. *Journal of Geophysical Research, [Solid Earth]*, *126*(5). doi: 10.1029/2020jb021567

Zhang, Z., & Olugboji, T. (2023, May). Lithospheric imaging through reverberant layers: Sediments, oceans, and glaciers. *Journal of Geophysical Research, [Solid Earth]*, *128*(5). doi: 10.1029/2022jb026348

Zheng, Z., Ventosa, S., & Romanowicz, B. (2015, April). High resolution upper mantle discontinuity images across the pacific ocean from SS precursors using local slant stack filters. *Geophysical Journal International*, *202*(1), 175–189. doi: 10.1093/gji/ggv118

Zhu, L., & Kanamori, H. (2000, February). Moho depth variation in southern california from teleseismic receiver functions. *Journal of geophysical research*, *105*(B2), 2969–2980. doi: 10.1029/1999JB900322