

# Minimalist Receiver Functions

William Menke  
Lamont–Doherty Earth Observatory of Columbia University  
Draft of 02 April 2002

## Summary

We present a new method for constructing receiver functions and the impulse response functions that they are based upon. The method differs from others in common use in specifically seeking receiver functions with the minimal number of pulses that can satisfy the data up to its noise level, and is meant to be complementary to – as contrasted to a replacement for – them. No preconceptions about the underlying earth model are built into the procedure, other than the assumption that the pulses correspond to body wave arrivals, and thus cause in–phase motion on all seismogram components. By examining a sequence of receiver functions, with increasing number of pulses, we can investigate their overall information content and the robustness of their features. We apply the method to crustal structure at station PAL (Palisades, NY), and use it to constrain the depth and dip of the Moho beneath that station to  $33\pm 1$  ( $1\sigma$ ) km and  $8^\circ\pm 2^\circ$  ( $1\sigma$ ), striking  $N22^\circ E \pm 20^\circ$  ( $1\sigma$ ), respectively

## Introduction

Receiver Function (RF) analysis (Phinney 1964, Burdick and Langston 1977, Owens & Crosson 1988) has become a very popular and effective tool for investigating crustal and upper mantle structure. Its popularity is partly due to its being a passive method that requires only one receiver that has recorded teleseismic P waves. It can therefore be used to study structure in places where active–source (i.e. refraction)

experiments are impractical, and to depths below those that can be easily insonified through convention active sources (e.g. chemical explosives).

RF analysis has been very successfully applied to many aspects of near–receiver structure – too many to review fully here. But a representative sampling would include studies of the crust (e.g. Sheehan et al. 1995, Sandvol et al. 1998), subducting slabs (Regnier 1988, Li & Nabelek 1996), transition zone topography (e.g. Shen et al. 1996, Dueker & Sheehan 1997, Li et al. 1998), and magma chambers (Sheetz and Schlue, 1992).

The RF method relies upon the cascade of converted and multiply reflected compressional and shear waves that a teleseism excites in the structure beneath the receiver, and that sample the elastic properties of that part of the earth. Both the vertical and horizontal components of the seismogram contain these phases, which arrive seconds to tens of seconds after the P wave (Fig. 1). However, these secondary arrivals are usually difficult to discern in the raw seismograms, because of the typically long duration of the earthquake source time function. RF analysis is a way of removing the effect of the source time function, so that the secondary arrivals stand out.

Most commonly, receiver functions are interpreted in the context of an isotropic, plane–layered earth structure, although the effect of layer dip and velocity anisotropy (Levin & Park 1997) are sometimes modeled. We will refer to the converted and multiply–reflected phases in a plane–layered earth structure, generically, as

“reverberations”, in order to distinguish them from the more disorganized wavefield produced by “scattering” from three–dimensional heterogeneities. Both reverberations and scattering may occur beneath a station, with the later becoming more important at higher frequencies ( $>1$  Hz) (Abers 1998). Part of the emphasis of this paper is to develop RF interpretation methods that are not locked into any particular preconception about the earth structure, but rather which provide some insight on the data's ability to discriminate between them.

### The Receiver Function

Suppose that a seismic station detects teleseismic P phases from a variety of sources with different epicentral distances and azimuths. Let the vertical and radial–horizontal seismograms be denoted,  $V_i^{\text{obs}}(t)$ , and,  $R_i^{\text{obs}}(t)$ , respectively. Here the subscript,  $i$ , indexes the  $N$  phases and the superscript “obs” denotes “observed”. We assume that these seismograms have been windowed to isolate the P wave. We can represent each seismogram as the convolution of a “source” wavelet,  $s_i(t)$ , with the vertical and radial–horizontal impulse response functions,  $v_i(t)$  and  $r_i(t)$ , respectively. These response functions quantify the effect of near–receiver structure on the phase:

$$V_i^{\text{obs}}(t) = s_i^{\text{true}}(t) * v_i^{\text{true}}(t) \quad (\text{Eqn. 1})$$

$$R_i^{\text{obs}}(t) = s_i^{\text{true}}(t) * r_i^{\text{true}}(t) \quad (\text{Eqn. 2})$$

We use the superscript “true” to indicate that in the absence of noise the observed seismograms are related to the actual source wavelet and impulse responses. The term,  $s_i(t)$ , is a “source” wavelet in the sense that it characterizes the waveform of the phase before it interacts with the near–receiver structure. It describes both the effects

of the earthquake source itself and subsequent modification by structure far from the receiver. An important aspect of these equations is that the same source wavelet appears in both equations.

The RF method (Phinney 1964, Burdick and Langston 1977, Owens & Crosson 1988) is a strategy for removing the source wavelet from these equations, creating an observable quantity – the receiver function,  $F(t)$  – that depends exclusively on the near–receiver structure. Eqn. 1 is solved for the source wavelet:

$$s_i^{\text{true}}(t) = V_i^{\text{obs}}(t) * v_i^{-1\text{true}}(t) \quad (\text{Eqn. 3})$$

Here  $v_i^{-1\text{true}}(t)$  is the inverse function to  $v_i^{\text{true}}(t)$ , in the sense that  $v_i^{-1\text{true}}(t) * v_i^{\text{true}}(t) = \delta(t)$ , where  $\delta(t)$  is the Dirac delta function. Eqn. 3 is then inserted into Eqn.2, yielding:

$$R_i^{\text{obs}}(t) = [r_i^{\text{true}}(t) * v_i^{-1\text{true}}(t)] * V_i^{\text{obs}}(t) = F_i^{\text{true}}(t) * V_i^{\text{obs}}(t) \quad (\text{Eqn. 4})$$

Note that the “receiver function”,  $F_i^{\text{true}}(t) = r_i^{\text{true}}(t) * v_i^{-1\text{true}}(t)$ , involves only the response of the earth, and not the source wavelet. Finally, an “observed” receiver function,  $F_i^{\text{obs}}(t)$ , can be constructed by solving Eqn. 4 for  $F$ :

$$F_i^{\text{obs}}(t) = R_i^{\text{obs}}(t) * V_i^{-1\text{obs}}(t) \approx r_i^{\text{true}}(t) * v_i^{-1\text{true}}(t) = F_i^{\text{true}}(t) \quad (\text{Eqn. 5})$$

The observed receiver function,  $F_i^{\text{obs}}(t)$ , is constructed by deconvolving the observed vertical component seismogram from radial–horizontal component. Note that it contains information about the near–receiver structure, but not about the source

wavelet.

Information about the near-receiver structure can be gained by comparing the observed receiver function with a receiver function,  $F_i^{\text{pre}}(\mathbf{m}, t) = h_i^{\text{pre}}(\mathbf{m}, t) * v^{-1}_i^{\text{pre}}(\mathbf{m}, t)$ , predicted from some earth model,  $\mathbf{m}$  (see, for instance, Ammon et al. 1990). Here the vector,  $\mathbf{m}$ , represents whatever parameters are necessary to describe the structure of the earth near the receiver (e.g. thickness and velocities in a plane layered model, properties of discrete scatterers in heterogeneous models). Least-squares can be used to select a best-fitting model:

$$\begin{aligned} &\text{find the } \mathbf{m}^{\text{est}} \text{ that minimizes} \\ &E_i(\mathbf{m}) = \| F_i^{\text{obs}}(t) - F_i^{\text{pre}}(\mathbf{m}, t) \|^2 \\ &\text{with respect to } \mathbf{m} \end{aligned} \quad (\text{Eqn. 6})$$

Here the superscript, “est”, denotes “estimated” and  $\|.\|^2$  denotes some measure of the difference between two timeseries, say  $x(t) = F_i^{\text{obs}}(t)$  and  $y(t) = F_i^{\text{pre}}(\mathbf{m}, t)$ . One possible choice is:

$$\|x(t) - y(t)\|^2 = \int [x(t) - y(t)]^2 dt / [\int x^2(t) dt + \int y^2(t) dt] \quad (\text{Eqn. 7})$$

Eqn. 7 is based on the commonly-used  $L_2$  norm, and has been normalized so that the misfit,  $E_i(\mathbf{m})$ , does not depend upon the overall amplitude of the seismograms. The misfit is zero if all the receiver functions are exactly matched and is of order unity if all the fits are poor.

One undesirable feature of this fitting procedure arises from the need to perform deconvolutions of both the observed seismograms and the predicted responses. This

process often suffers from instability, especially when the observed timeseries are band-limited, or when the predicted vertical response has a spectral zero in the frequency band of interest (e.g. Clayton & Wiggins 1976, Ammon 1991). While this instability to some degree can be eliminated by using advanced signal processing techniques (e.g. Park & Levin, 2000), we suggest an alternate strategy that completely avoids the need to perform any deconvolutions. We first note that Eqn. 5 can be manipulated into the form:

$$\mathbf{R}_i^{\text{obs}}(t) * \mathbf{v}_i^{\text{true}}(t) \approx \mathbf{V}_i^{\text{obs}}(t) * \mathbf{r}_i^{\text{true}}(t) \quad (\text{Eqn. 7})$$

The predicted responses,  $\mathbf{v}_i^{\text{pre}}(\mathbf{m},t)$  and  $\mathbf{r}_i^{\text{pre}}(\mathbf{m},t)$ , will satisfy Eqn. 7 when they are equal to the true ones. A estimate of the model is thus given by least-squares:

$$\begin{aligned} &\text{find the } \mathbf{m}^{\text{est}} \text{ that minimizes} \\ &E_i(\mathbf{m}) = \|\mathbf{r}_i^{\text{pre}}(\mathbf{m},t) * \mathbf{V}_i^{\text{obs}}(t) - \mathbf{v}_i^{\text{pre}}(\mathbf{m},t) * \mathbf{R}_i^{\text{obs}}(t)\|^2 \\ &\text{with respect to } \mathbf{m} \end{aligned} \quad (\text{Eqn. 8})$$

Note that this procedure contains only convolutions of known (observed and modeled) quantities. An advantage of this procedure is that it avoids bias associated with smoothing (i.e. low pass filtering) that occurs during the computation of the receiver function, and therefore allows testing whether the data are compatible with the assumption of sharp layer interfaces or compact three-dimensional scatterers. Structural inversions that proceed from a smoothed estimate of the receiver function will of course be most compatible with a smooth velocity structure. But ruling out sharp interfaces or compact scatterers on that basis would be a fallacy.

## Minimalist Receiver Functions

Normal practice is to use a deconvolution technique to estimate the receiver function from the observed seismograms. The receiver function obtained in this way is a timeseries of the same length and sampling interval as the input seismograms. A preliminary interpretation can then be performed by finding pulses in these receiver functions that correspond to secondary arrivals associated with the compressional and shear waves scattered from heterogeneities at depth (e.g. the shear wave from the Moho interface). This preliminary interpretation can be used to develop a starting model that can subsequently improved by the least-squares method of Eqn. 6.

There are several unsatisfactory aspects to this approach.

First, there is a mismatch between the large number of parameters in the receiver function (equal to the number of samples,  $M$ , in the observed seismograms) and its actual information content, which is almost certainly much smaller. This low information content is due to both the insensitivity of low-frequency (typically  $<1$  Hz) teleseismic waves to small details in structure, and to our inability to model high frequencies even when they are available, owing to the large contribution from scattering (Abers 1998). We would like way of estimating how much information is actually available, in order to avoid over interpreting the data.

Second, if the receiver function is likely, at least to first order, to contain a series of discrete pulses, then its representation as a long time series only serves to complicate

detecting them and determining their arrival times and amplitudes.

We implement an alternate representation of the receiver function that assumes that it consists of a small number, say  $L$ , of discrete pulses:

$$F_i^{\text{est}}(t) = \sum_{j=1}^L c_j \delta(t-T_j) \quad (\text{Eqn. 9})$$

Here the  $c_j$  are the amplitudes of the pulses and  $T_j$  is their time of occurrence. (Note  $T_1$  can be taken to be zero). These parameters constitute unknown parameters, and are estimated using the least-squares procedure defined in Eqn. 6. Note that the  $c_j$  are not “velocity model” parameters, since they do not parameterize any particular near-receiver velocity model. This approach places no preconceptions on the processes that produce scattered compressional and shear waves. It works equally well for layered earth models and models with discrete, threedimensional heterogeneities. For this reason it is conceptually different from (an complementary to) methods, such as Sandvol et al.'s (1998), that fit the receiver function with an earth model containing only a few parameters.

We use a grid search to find the receiver function parameters that minimize the misfit between the observed and predicted radial–horizontal seismogram:

$$\begin{aligned} &\text{find the } \mathbf{m}^{\text{est}} \text{ that minimizes} \\ &E_i(\mathbf{m}) = \| \mathbf{R}_i^{\text{obs}}(t) - \mathbf{F}^{\text{est}}(\mathbf{m}, t) * \mathbf{Z}_i^{\text{obs}}(t) \|^2 \\ &\text{with respect to } \mathbf{m} \end{aligned} \quad (\text{Eqn. 10})$$

While the number of parameters is  $2L-1$  ( $L$  amplitudes and  $L-1$  pulse times), the grid



search need be performed only over the  $L-1$  pulse times, since the problem of determining the coefficients is completely linear when the pulse times are fixed. Eqn. 9 is inserted into Eqn. 4., yielding:

$$R_i^{obs}(t) = F_i^{est}(t) * V_i^{obs}(t) = \sum_{j=1}^L c_j V_i^{obs}(t-T_j) \quad (\text{Eqn.11})$$

This is a system of  $M$  linear algebraic equations for the  $L$  unknown amplitudes,  $c_j$ , and can be readily solved by least-squares (e.g. Menke 1989, Abers et al. 1995, Gurrola and Minster 1995). Our experience is that the grid search is an effective solution method when the number of pulses is less than about 6. Should receiver functions with more pulses be of interest, then non-linear minimization procedures, such as simulated annealing (Velis 2001) or genetic algorithms (Koper et al. 1999, Floyd 2001) need be applied to minimize the error,  $E$ , with respect to  $T_j$ .

An analogous method can be used to find spiky versions of the impulse response functions. We start with a spiky representation analogous to Eqn. 9:

$$\begin{aligned} v_i^{est}(t) &= \sum_{j=1}^L a_j \delta(t-T_j) \\ r_i^{est}(t) &= \sum_{j=1}^L b_j \delta(t-T_j) \end{aligned} \quad (\text{Eqn. 12})$$

Here  $a_j$  and  $b_j$  are the amplitudes of the pulses and  $T_j$  is their time of occurrence. Note that pulses are assumed to arrive at the same times on both components, as would be expected if they corresponded to seismic body waves, and that  $T_1$  can be taken to be zero. The minimization principle defined in Eqn. 8 can then be used as the basis for determining the unknown parameters. However, Eqn. 7 indicates that the impulse

responses can only be determined up to an arbitrary multiplicative constant, so some further constraint must be added to the minimization process. We use  $a_0=1$ . As in the receiver function case, the grid search need be performed only over the  $L-1$  pulse times, since for fixed pulse times the equation for the amplitudes is linear. Inserting Eqn. 12 into Eqn. 7 yields:

$$\sum_{j=2}^L a_j R_i^{\text{obs}}(t-T_j) - \sum_{j=1}^L b_j V_i^{\text{obs}}(t-T_j) = -R_1^{\text{obs}}(t-T_1)$$

(Eqn. 13)

This is a system of  $M$  linear algebraic equations for the  $2L-1$  unknown amplitudes,  $a_i$  and  $c_i$ .

Eqn. 7, which forms the basis of the minimization, would be satisfied even if  $v_i^{\text{est}}(t)$  and  $r_i^{\text{est}}(t)$  were both convolved by some arbitrary operator, say  $O(t)$ . One might suppose then that the estimated operators resulting from the grid search suffered from this same nonuniqueness. However, the condition that the responses contain exactly  $L$  pulses is very stringent. In general, if  $v_i^{\text{est}}(t)$  contains  $L$  pulses, then  $O(t)*v_i^{\text{est}}(t)$  will contain more than  $L$  pulses, so this nonuniqueness is usually not a problem. An exception is the degenerate case when some of the  $L$  pulses have zero amplitude.

Thus for example, if the minimal-error three-pulse operators,  $v_i^{\text{est}}(t)$  and  $r_i^{\text{est}}(t)$ , in fact have only two non-zero pulses, separated, say, by a time interval,  $\Delta t$ , then  $O(t)*v_i^{\text{est}}(t)$  and  $O(t)*r_i^{\text{est}}(t)$  will be minimal error with 3 pulses, for any two-pulse operator,  $O(t)$ , with a pulse separation of  $\Delta t$ .

An estimate of the receiver function can be formed by deconvolving the  $L$ -pulse

vertical operator from the radial–horizontal operator,  $F'^{\text{est}}(t) = r_1^{\text{est}}(t) * z_i^{-1\text{est}}(t)$ . Here we use the prime to distinguish this estimate from the previous estimate,  $F^{\text{est}}(t)$ , that is based on Eqn. 10. The two estimates will, in general, be different, and indeed  $F'^{\text{est}}(t)$  will, in general, have an infinite number of pulses. One might argue that for any fixed  $L$ ,  $F'^{\text{est}}(t)$  is a better estimate of the receiver function than is  $F^{\text{est}}(t)$ , since  $F'^{\text{est}}(t)$  is based on  $3L-2$  parameters while  $F^{\text{est}}(t)$  is based on only  $2L-1$ . Indeed,  $F'^{\text{est}}(t)$  will typically have smaller error than  $F^{\text{est}}(t)$ , as quantified by Eqn. 10. However, the process of deconvolution is non–trivial, and owing to spectral zeroes in  $z_i^{\text{est}}(t)$  can easily lead to an  $F'^{\text{est}}(t)$  that is ringy or otherwise unaesthetic.

#### Application to the PAL (Palisades, NY) Seismic Station

We selected 29 high quality teleseismic seismograms recorded by a broadband seismometer at station PAL, located about 20 km north of New York City in Palisades, NY. All earthquakes have magnitudes,  $m_b \geq 6$ , and epicentral distances between  $30^\circ$ – $90^\circ$ . A 40s segment of the P wave was extracted, low–pass filtered below 1 Hz, cosine tapered, and resampled with a interval of 0.2s. The median signal to noise ratio of these data, based on a comparison with the signal immediately preceding the P wave, is about 7:1 for the vertical component and about 5:1 for the radial–horizontal component. The actual “signal” relevant to estimates of the receiver function, however, is the mismatch between the radial–horizontal component and the vertical component, after an overall scale factor has been accounted for. (This is to say that near–receiver structure can only be detected insofar as the motion on the

radial–horizontal component is different than on the vertical component). The median signal to noise ratio of these data is considerably worse, about 2:1. The best ratio, 12:1, was for a  $m_b=7.8$  earthquake occurring only  $31^\circ$  away. Nevertheless, PAL has a fairly average noise level, when compared to other broadband stations operating in North America. (It is an intermediate noise station, when judged against Peterson's (1993) Low Noise Model (LNM), with a noise level that is about 5 dB higher than the LNM at a period of 10s).

A sequence of impulse responses,  $r_i^{\text{est}}(t)$  and  $z_i^{\text{est}}(t)$  and the receiver function,  $F^{\text{est}}(t)$  were estimated for each P wave, for  $L=1$  through  $L=5$ . The grid search allowed pulses to occur in the first 15s of the operators, with a time resolution equal to the 0.2s sampling interval of the seismograms. This time window is sufficient to include shear phased scattered from crust and uppermost mantle heterogeneity, but excludes shear waves scattered from the transition zone. The alternate receiver function,  $F'^{\text{est}}(t)$ , was also estimated from  $r_i^{\text{est}}(t)$  and  $v_i^{\text{est}}(t)$  using a standard deconvolution algorithm.

However, it needed to be low–pass filter below 1 Hz to suppress high–frequency ringing.

The estimated impulse responses and receiver functions (Fig. 2) evolve in an orderly way as the number of pulses in their representations increase (i.e. as  $L$  increases). A pulse with an arrival time of about 4–5s is present in nearly all cases, and is presumed to represent the Moho–converted shear wave,  $P_s$ . In all cases the logarithm of the misfit decreases monotonically and approximately linearly with  $L$  (Fig. 4). As expected, the misfit associated with the impulse responses,  $r_i^{\text{est}}(t)$  and  $z_i^{\text{est}}(t)$ ,

decreases faster with  $L$  than does the misfit associated with the receiver function,  $F^{\text{est}}(t)$ . The misfit is reduced, on average, by a factor of about 6 for the  $L=5$  impulse responses and about 3 for the  $L=5$  receiver functions. Since, as discussed above, the actual signal to noise ratio is only about 2:1, corresponding to a misfit of 4:1, the  $L=5$  impulse responses are probably slightly overfitting the data, while the  $L=5$  receiver functions are slightly underfitting it.

The  $F^{\text{est}}(t)$  and  $F'^{\text{est}}(t)$  receiver functions (Figs. 4 and 5) both indicate that the arrival time of  $P_s$  varies systematically with the arrival direction (i.e. backazimuth,  $\phi$ ) of the  $P$  wave.  $P_s$  arrives about 4s after  $P$  for southerly backazimuths, increasing to about 5s for northerly azimuths. We have inverted the three-pulse ( $L=3$ )  $P_s$  arrival times (Fig. 6) for crustal thickness, assuming a plane-layered structure and mean crustal compressional and shear wave velocities of 6.50 and 3.69 km/s, respectively (Levin et al. 1995). The mean crustal thickness of  $33 \pm 1$  km is consistent with what is known about the general thinning of the continental crust of the northern Appalachians, from 45–50 km at the center of the craton, to about 35 km at the coast, to 15–20 km on the continental shelf (Hughes and Luegert, 1992; Hennen et al., 1991; Keen and Barrett, 1981). The apparent dip of the Moho at Palisades, of about  $10^\circ$  downward in the direction  $N292^\circ E$  is about twice as large as might be expected from this regional crustal thinning, presuming it to be uniform. Station PAL is sited in the northeastern corner of the Newark basin, a half-graben bounded on its western margin by a steeply-dipping normal fault. The Newark basin is one of the Mesozoic rift basins associated with the opening of the Atlantic ocean. The inferred  $N22^\circ E \pm 20^\circ$  ( $1\sigma$ )

strike of the Moho is different than the N60°E strike of sediments within the basin (P. Olsen, personal communication, 2002), but both are generally consistent with the notion that the crust thickens towards the northwestern side of the basin. Other local irregularities in crustal thickness of similar magnitudes have been observed ~300 km away in New Hampshire (Shalev et al. 1991).

Additional support for the dipping Moho can be obtained by examining the apparent angle of incidence,  $\theta$ , of the P wave, as inferred from the direction of particle motion of the leading pulse in the L=3 impulse responses (i.e.  $\tan(\theta)=b_1/a_1$ ). We first consider the case of a horizontal Moho. Ignoring, for a moment, the perturbing effect of the free surface on particle motions, we would expect from Snell's law that the apparent angle of incidence is a function of the slowness,  $p$ , of the teleseism (as, inferred, say from the IASPEI tables (Kennett, 1991)) and the near-surface velocity,  $\alpha$ , of the crust:  $\sin(\theta)=\alpha p$ . Numerical calculations (not shown), based on Aki and Richard's (1991, p. 190) free-surface correction formula, indicate that even when the free surface is taken into account,  $\sin(\theta)$  still varies approximately linearly with  $p$ , although the proportionally constant is biased to values larger than  $\alpha$ . The apparent angle of incidence data (Fig. 7, top) do follow this rule, and scatter about a best fit line of slope,  $6.03\pm 0.7$  ( $1\sigma$ ) km/s. This slope implies a near-surface velocity of  $\alpha=5.69\pm 0.6$  ( $1\sigma$ ) km/s, after the free surface correction is applied. If the Moho is dipping, we would expect it to have a perturbing effect on the apparent angle of incidence, biasing it towards lower values when the P wave is propagating in the updip direction, and to higher values when it is propagating in the downdip direction.

This pattern is indeed detected in the data (Fig. 7, bottom). The overall amplitude of the sinusoidal variation is about  $\Delta \sin(\theta) = 0.04 \pm 0.01$  ( $1\sigma$ ), which is consistent with a dip of about  $8^\circ \pm 2^\circ$  ( $1\sigma$ ). The fact that the P wave samples a larger patch of Moho (radius of about 15 km) than does the Ps phase (Fig. 1, inset), yet both detect similar dip, indicates that the Moho topography has wavelength greater than this dimension.

## Conclusions

The analysis method that we present here enables the construction of a receiver function containing the minimal number of pulses needed to fit the observations to any desired level of accuracy. The methodology therefore allows one to assess the degree to which the data are compatible with the idea that the near-receiver response is dominated by the effect of just a few discrete interfaces or scatterers. The method is complimentary to deconvolution-based methods that produce receiver function estimates with many more degrees of freedom.

When applied to seismic data from station PAL (Palisades, NY), the method produces receiver functions that contain a clear Ps phase from a  $33 \pm 1$  ( $1\sigma$ ) km deep Moho that appears to dip  $8^\circ \pm 2^\circ$  ( $1\sigma$ ) downward to the northwest (strike  $N22^\circ E \pm 20^\circ$  ( $1\sigma$ )). Surprisingly few discrete pulses, usually no more than 4 or 5, are needed to match the data up to its noise level.

Acknowledgements. This research was supported by Columbia University. Lamont-Doherty Contribution No. 0000.





## References

- Abers, G. 1998. Array measurements for phases used in receiver function calculations: importance of scattering, *Bull. Seism. Soc. Am.* 87, 313–318.
- Abers, G., X. Hu and L. Sykes, 1995. Source–scaling of earthquakes in the Shumagin region, Alaska: time–domain inversions of regional waveforms, *Geophys. J. Int.* 123, 41–58.
- Aki, K. and P. Richards, 1980. *Quantitative Seismology, Theory and Methods*, Vol. 1., W.H. Freeman and Company, 557pp.
- Ammon, J., 1991. The isolation of receiver function effects from teleseismic P waveforms, *Bull. Seism. Soc. Am.* 81, 2504–2510.
- Ammon, C., G. Randall and G. Zandt, 1990. On the nonuniqueness of receiver function inversions, *J. Geophys. Res.* 95, 15,202–15,318.
- Burdick, L. and C. Lanston, 1977. Modeling crustal structure through the use of converted phases in teleseismic body–wave forms, *Bull. Seism. Soc. Am.* 67, 677–691.
- Clayton, R. and R. Wiggins, 1976. Source shape estimation and deconvolution of teleseismic body waves, *Geophys. J.R. Astr. Soc.* 47, 151–177.
- Dueker, K. and A. Sheehan, 1997. Mantle discontinuity structure from midpoint stacks of converted P to S waves across the Yellowstone hotspot track, *J. Geophys. Res.* 102, 8,313–8,327.
- Floyd, J., J. Mutter, A. Goodliffe and B. Taylor, 2001. Evidence for fault weakness and fluid flow within an active low–angle normal fault, *Nature* 411, 779–783.
- Gurrola, H. and J. Minster, 1988. Thickness estimates of the upper mantle transition zone from bootstrapped velocity spectrum stacks of receiver functions, *Geophys. J. Int.* 133, 31–43.
- Gurrola, H., F. Baker and J. Minster (1995). Simultaneous time–domain deconvolution with application to the computation of receiver functions, *J. Geophys. Res.* 120, 537–543.
- Hennet, C., J. Luetgert, and R. Phinney, 1999. The crustal structure in central Maine from coherency processed refraction data, *J. Geophys. Res.* 96, 12023–12037.
- Hughes S., and J. Luetgert, 1992. Crustal structure of the southeastern Grenville province, northern New York and Eastern Ontario, *J. Geophys. Res.* 97,

17455–17479.

Keen C., and D. Barrett, 1981. Thinned and subsided continental crust on the rifted margin of eastern Canada, crustal structure, thermal evolution and subsidence history, *Geophys. J. R. Astr. Soc.* 65, 443–465.

Kennett, B., ed., 1991. IASPEI 1991 Seismological Tables, Australian School of Earth Science, Canberra, Australia, 167 pp.

Koper, K., M. Wyssession and D. Wiens, 1999. Multimodal function optimization with a niching genetic algorithm; a seismological example, *Bull. Seism. Soc. Am.* 89, 978–988.

Levin, V. and J. Park, 1997, P–SH conversions in a flat–layered medium with anisotropy of arbitrary orientation, *Geophys. J. Int.* 131, 253–266.

Levin, V., W.Y. Kim and W. Menke, 1995. Seismic velocities in the shallow crust of western New England and northern New York, *Bull. Seism. Soc. Am.* 85, 207–219.

Li, A., K. Fischer, M. Wyssession and T. Clarke, 1998. Mantle discontinuities and temperature under the North American continental keel, *Nature* 395, 160–163.

Li, X.–Q. and J. Nabelek, 1996. Modeling the structure of the Cascadia subduction zone in Oregon using the receiver functions from a broadband seismometer array, *Eos, Trans. AGU* 77, Suppl., 656.

Menke, W., 1989. *Geophysical Data Analysis: Discrete Inverse Theory*, Revised Edition (textbook), Academic Press, Inc., New York, 285 pp.

Menke, W., 2000. SPLITTING MODELER Software, [ftp://ftp.ldeo.columbia.edu/pub/menke/SPLITTING\\_MODELER.tar.Z](ftp://ftp.ldeo.columbia.edu/pub/menke/SPLITTING_MODELER.tar.Z).

Owens, T. and R. Crosson, 1988. Shallow structure effects on broadband teleseismic waveforms, *Bull. Seism. Soc. Am.* 77, 96–108.

Park, J. and V. Levin, 2000. Receiver functions from multiple–taper spectral correlation estimate, *Bull. Seism. Soc. Am.* 90, 1507–1520.

Peterson, J., 1993. Observations and modeling of background seismic noise, U.G. Geological Survey Open File Report 93–322, Albuquerque, NM.

Phinney, R., 1964. Structure of the Earth's crust from spectral behavior of long–period body waves, *J. Geophys. Res.* 69, 2997–3017.

Regnier, M., 1988, Lateral variation of upper mantle structure beneath New Caledonia determined from P–wave receiver function; evidence for a fossil subduction zone,

Geophys. J. R. Astr. Soc. 95, 561–577.

Sandvol, E., D. Seeber, A. Calvert and M. Barazangi, 1998. Grid-search modeling of receiver functions: implications for crustal structure in the Middle East and North Africa, *J. Geophys. Res.* 103, 26,899–26,917.

Shalev, E., J. Park and A. Lerner-Lam, 1991. Crustal velocity and Moho topography in central New Hampshire, *J. Geophys. Res.* 96, 16,415–16,427.

Sheehan, A., G. Abers, C. Jones, and A. Lerner-Lam, 1995. Crustal thickness variations across the Colorado Rocky Mountains from teleseismic receiver functions, *J. Geophys. Res.* 100, 20,391–20,404.

Sheetz, K. and J. Schlue, 1992. Inferences for the Socorro magma body from teleseismic receiver functions, *Geophys. Res. Lett.* 19, 1,867–1,870.

Shen, Y., S. Solomon, I. Bjarnason, M. Purdy, 1996. Hot mantle transition zone beneath Iceland and the adjacent Mid-Atlantic Ridge inferred from P-to-S conversions at the 410- and 660-km discontinuities, *Geophys. Res. Lett.* 23, 3,527–3,530.

Velis, D., 2001. Traveltime inversion for 2-D anomaly structures, *Geophysics*. 66,1481–1487.

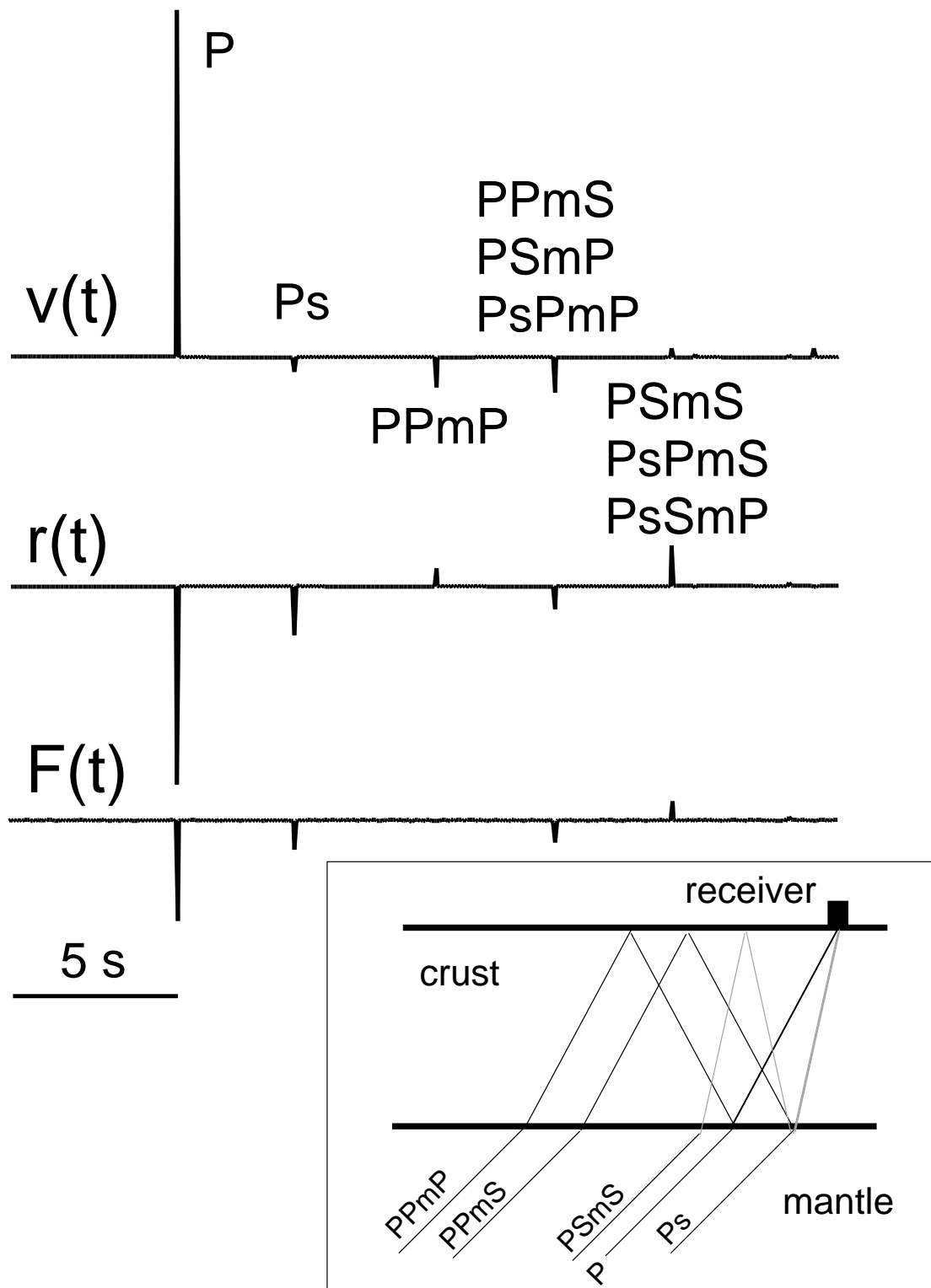


Fig. 1. Synthetic vertical,  $v(t)$ , and radial-horizonal,  $r(t)$ , impulse responses and receiver function,  $F(t)$ , for a simple model with a 30 km thick crust. Inset: Schematic of selected raypaths, with shear raypaths in gray. This example was computed using Menke's (2000) SPLITTING\_MODELLEER software.

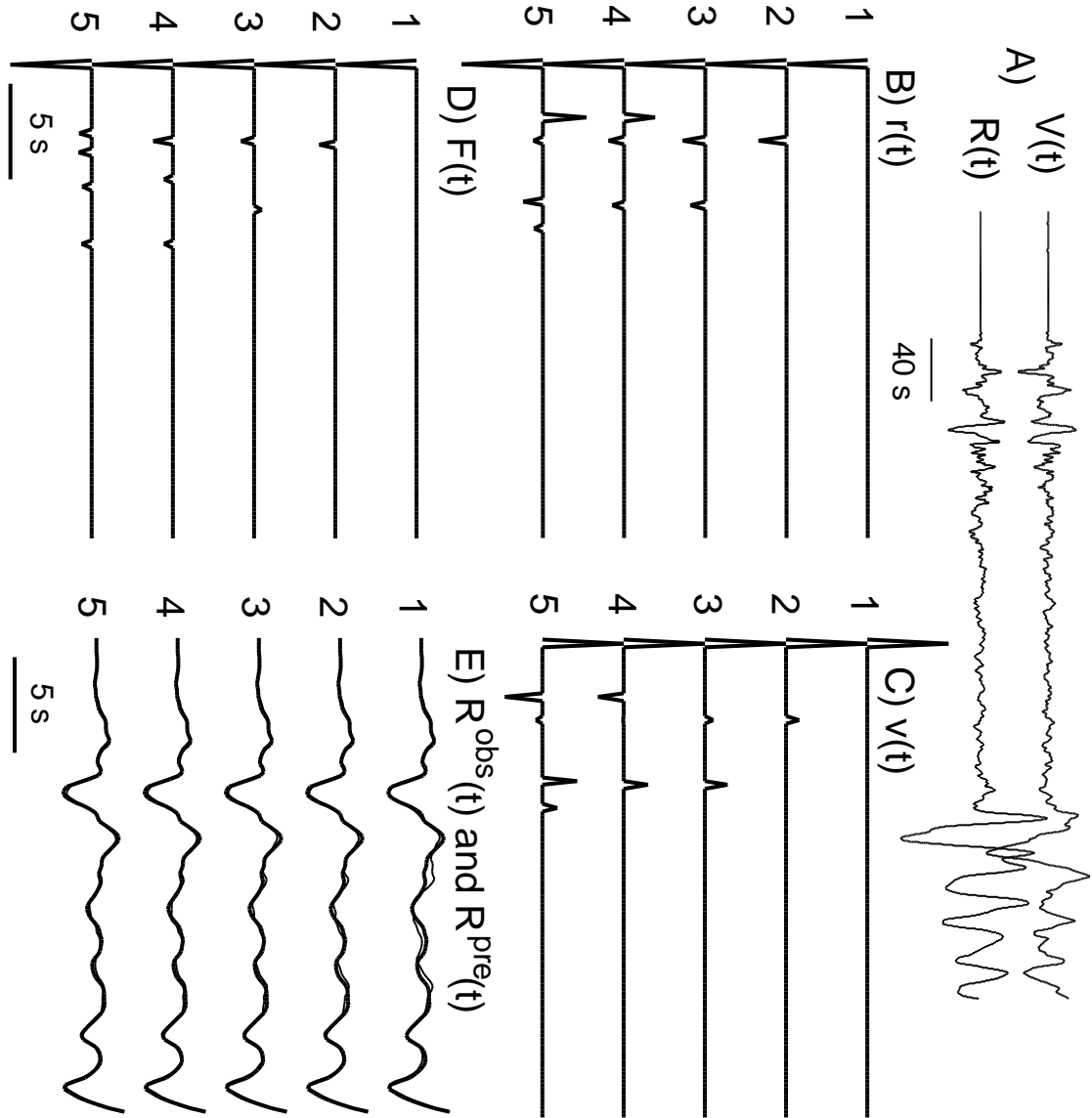


Fig. 2. Sample impulse responses,  $r_i^{\text{est}}(t)$ ,  $v_i^{\text{est}}(t)$  and the receiver functions,  $F_i^{\text{est}}(t)$  for an earthquake in El Salvador (range =  $30.7^\circ$ ) observed at PAL. A) Broadband vertical,  $V(t)$ , and radial–horizontal,  $R(t)$ , components. B) Sequence of  $v_i^{\text{est}}(t)$  for  $L$  increasing from 1 to 5. C) Sequence of  $r_i^{\text{est}}(t)$ . D) Sequence of  $F_i^{\text{est}}(t)$ . E) Sequence of observed (bold) and predicted (solid) radial–horizontal seismicograms,  $R(t)$ . The misfit, as defined by Eqn. 8, decreases from 0.034 for  $L=1$  to 0.003 for  $L=5$ , about an order of magnitude improvement.

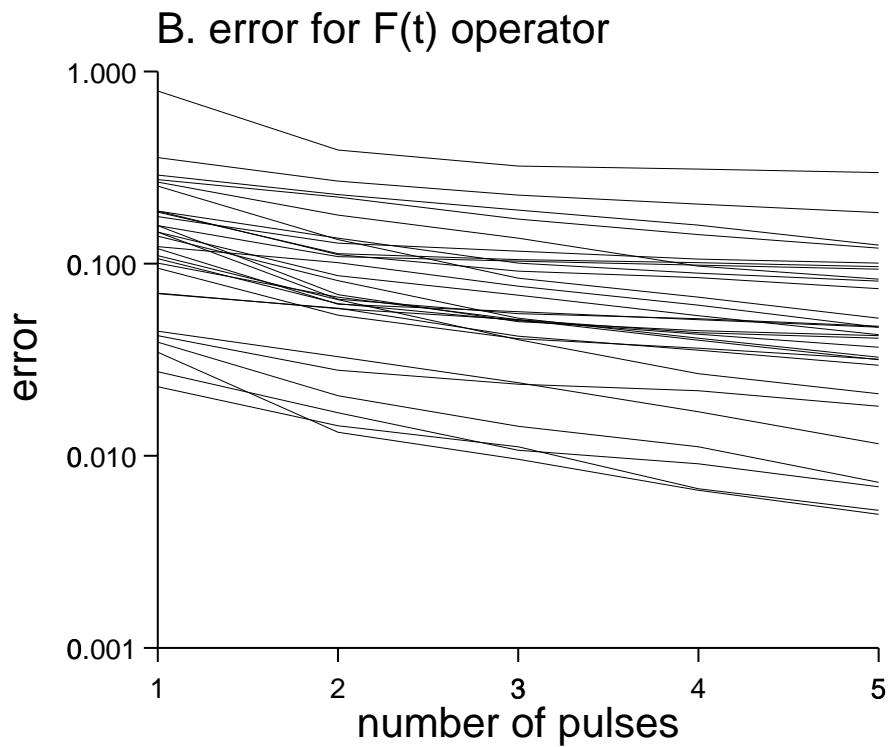
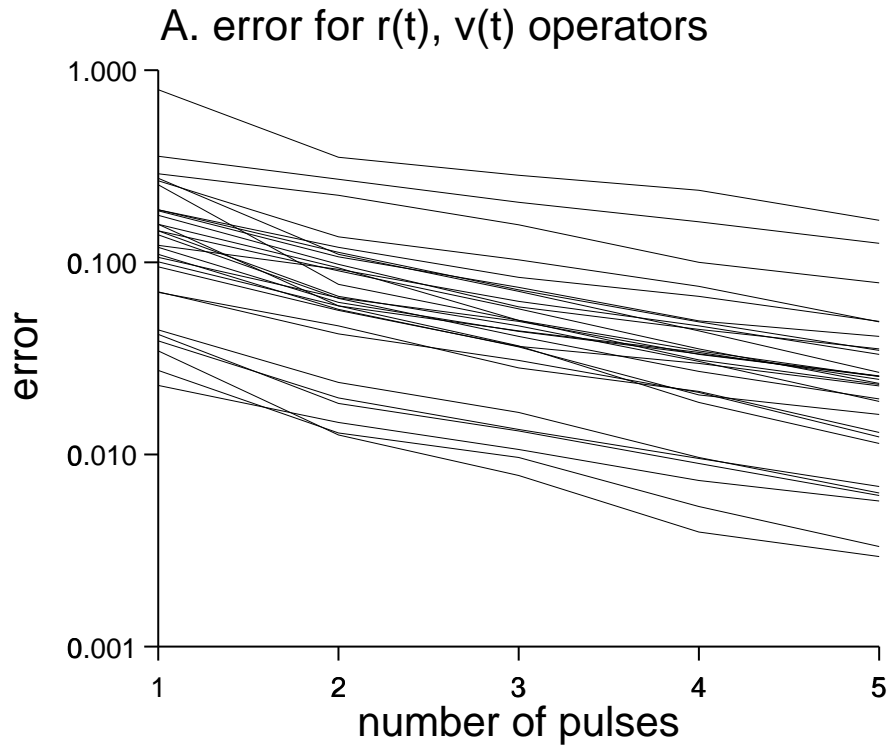


Fig. 3. A) Error, as defined by equation 8, of impulse response function determinations, as a function of the number of pulses,  $L$ . B) Error, as defined by equation 10, of receiver function determinations, as a function of the number of pulses,  $L$ .

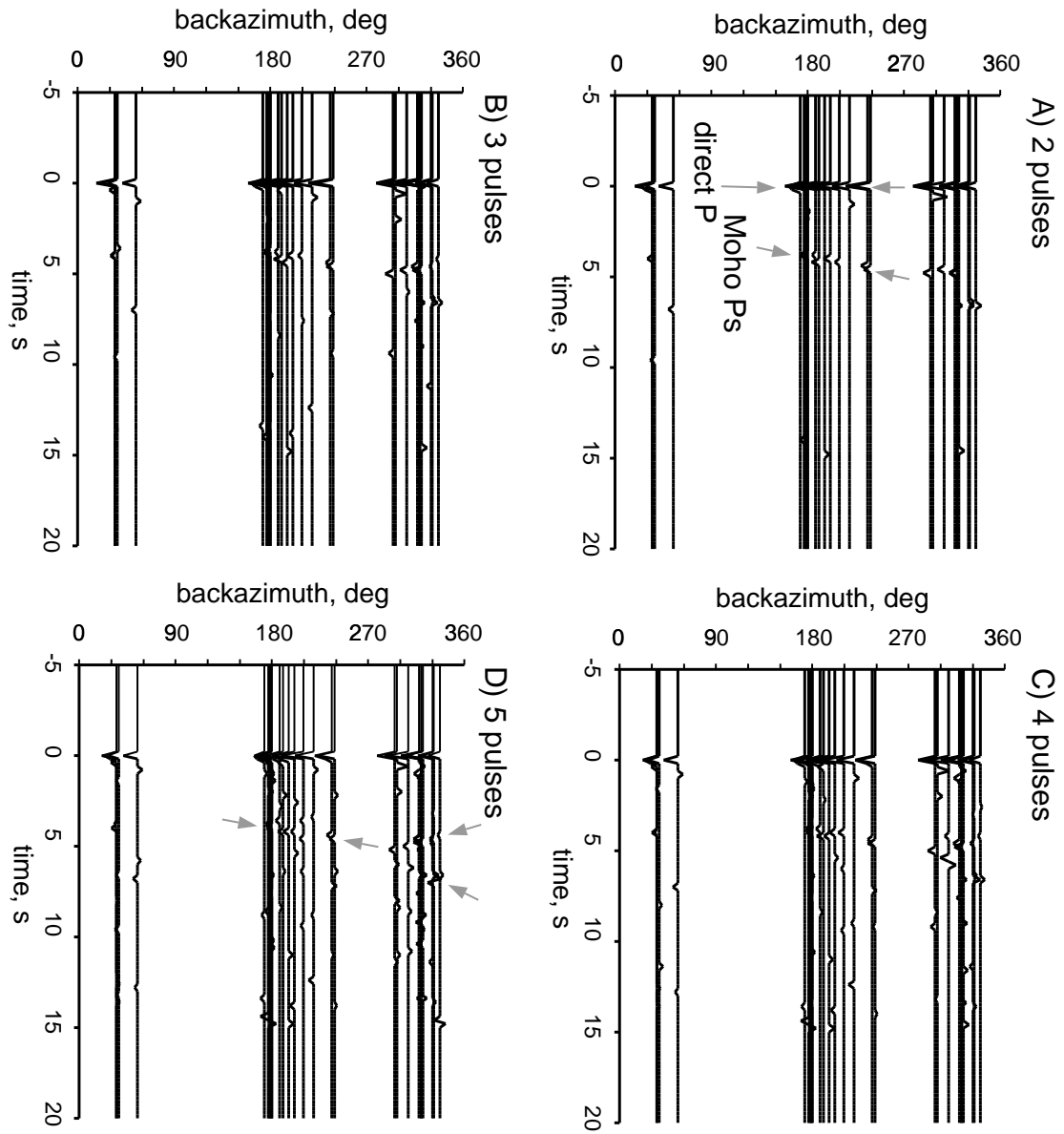


Fig. 4. Estimated receiver function,  $F^{\text{est}}(t)$ , plotted by backazimuth of the earthquake. A)  $L=2$ . B)  $L=3$ . C)  $L=4$ . D)  $L=5$ .

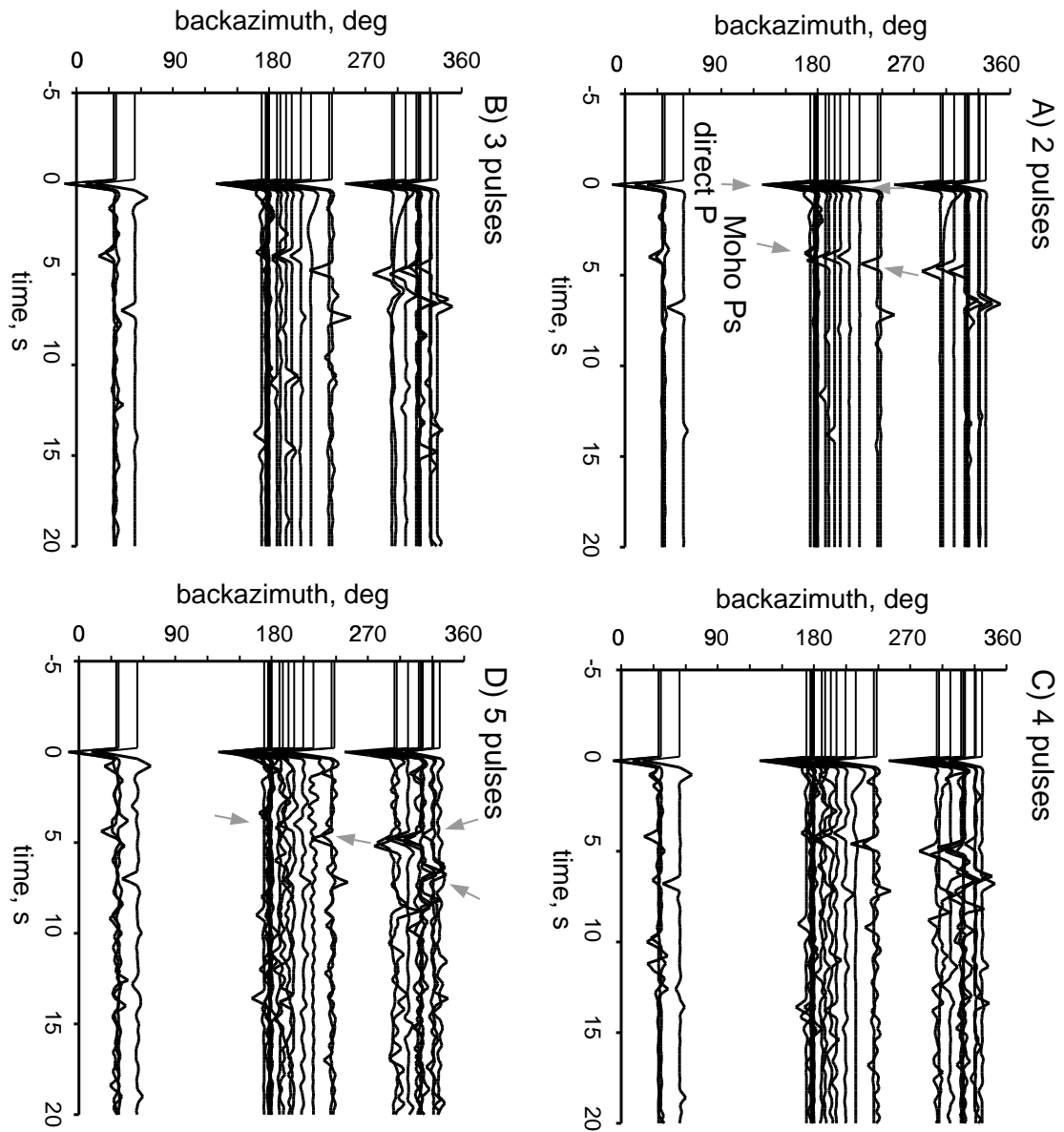


Fig 5. Estimated receiver function,  $F^{\text{est}}(t)$ , plotted by backazimuth of the earthquake. A)  $L=2$ . B)  $L=3$ . C)  $L=4$ . D)  $L=5$ .



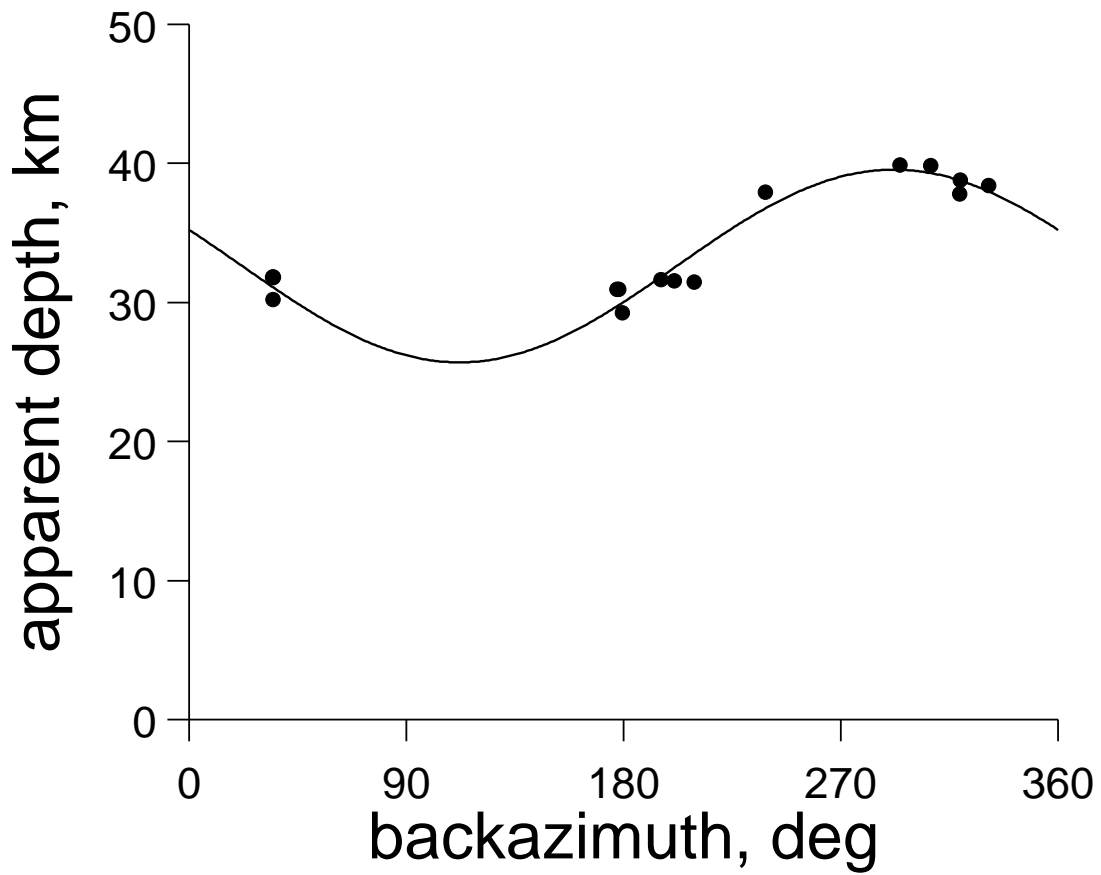


Fig. 6. Apparent Moho depth, plotted as a function of backazimuth,  $\phi$ . Circles: Depth estimates based on Ps arrivals from three-pulse calculation, calculated for a nominal mean crustal compressional and shear velocities of 6.50 and 3.69 km/s, respectively. (Solid line) Best fit sinusoid,  $32.62+6.93\cos(\phi-292^\circ)$ , indicating that the Moho dips down towards the northwest.

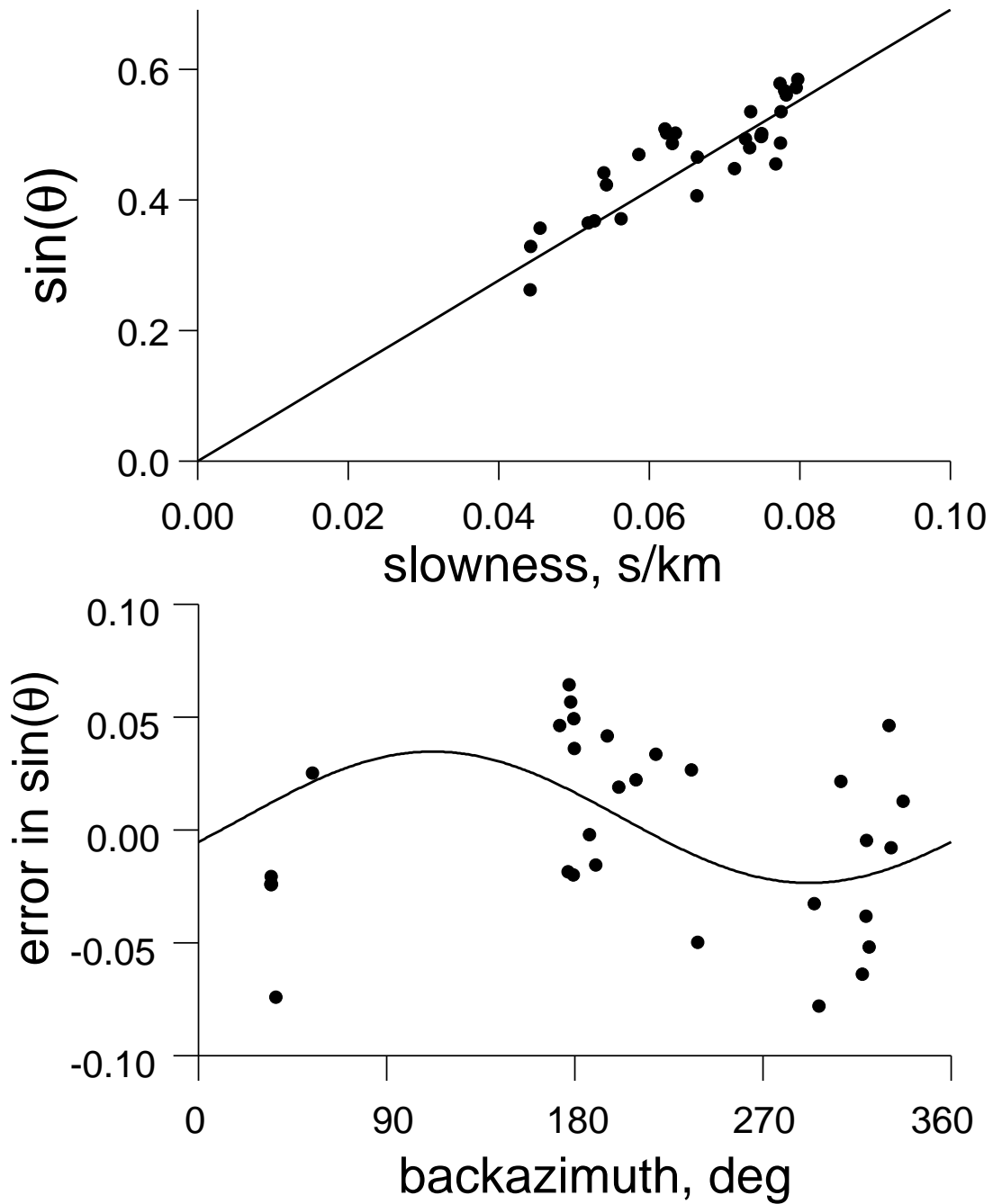


Fig. 7. (Top) Sine of the apparent angle of incidence,  $\theta$ , inferred from the amplitude of the P-wave pulse on the L=3 vertical and radial–horizontal impulse response functions, as a function of the slowness,  $p$ , of the teleseism as inferred from the IASPEI Tables (Kennett 1991). The data (circles) are compared with a best fit straight line with slope of 6.03 km/s (solid curve). (Bottom) Deviations of the  $\sin(\theta)$  data (circles) from the straight line fit as function of the backazimuth of the teleseism. Solid curve: best fit cosine.