

Optimized data masks for focussed solar tomography: background and artificial diagnostic experiments

S. J. Hughes^{1,2,*}, F. P. Pijpers¹, and M. J. Thompson²

¹ Space & Atmospheric Physics Group, Imperial College London, Exhibition Road, London, SW7 2AZ, UK

² Department of Applied Mathematics, University of Sheffield, Hounsfield Road, Sheffield S3 7RH, UK

Received 20 April 2006 / Accepted 19 March 2007

ABSTRACT

Context. The use of solar tomography for detecting subsurface features in the Sun is now well established. It customarily proceeds from an analysis of data on the solar surface given weightings in a predetermined geometric (e.g. centre-annulus) configuration.

Aims. We seek to improve these weightings by developing a scheme for choosing optimal combinations of data that maximise the contribution from signal at a desired horizontal and depth location.

Methods. We employ a subtractive optimally localized averaging (SOLA) scheme to pick weights for particular data points and analyse the quality of the results.

Results. We show in this work that particularly by using wave kernels instead of rays for modelling purposes, one can do well at localizing a measurement through our techniques while achieving desirable error-magnification properties.

Key words. Sun: activity – Sun: helioseismology – Sun: interior

1. Background

The power of solar tomography, also known as time-distance helioseismology (Duvall et al. 1993), for making inferences about solar structures is that it has the ability to study the solar subsurface regions locally at high resolution. Moreover it is highly adaptable and intuitive. The approach has been very successful for analysis of near-surface features, including sunspots (Kosovichev et al. 2000; Zhao & Kosovichev 2003; Hughes et al. 2005) and large-scale flows (Giles 2000). Application of the technique to see deeper into the convection zone is made harder by problems of obtaining quality data spanning large enough distances on the Sun but avoiding unsuitable measurement geometries; the sensitivity of the measurements also naturally decreases with depth because of the greater sound speed deeper beneath the surface.

In time-distance helioseismology one seeks to link a surface measurement of travel times of acoustic waves to the solar properties beneath, e.g. the sound speed or the presence of a bulk flow. This is conveniently cast in terms of a linearization around a known reference model, typically a time-independent model where the stratification is a function only of depth. Then the differences $\delta\tau$ between observed travel times and those of the reference model can be related to differences between properties of the Sun and model as

$$\delta\tau = \sum_{\alpha} \int_{\odot} dr \delta q_{\alpha}(r) K^{\alpha}(r) \quad (1)$$

(Gizon & Birch 2002), where each q_{α} is a particular property of the interior, via the *sensitivity kernel* K^{α} , which describes the response of the travel time to a perturbation δq_{α} of property q_{α}

of the medium. Inversion techniques can be used to make inferences about the properties δq_{α} from the measurements.

To obtain good results it is useful to understand the form of the sensitivity kernels and to attempt to choose groups of measurements that optimize the dependence on a particular part of the solar sub surface; by choosing the points on the surface correctly the spatial distribution of the total sensitivity can be distributed as desired. Conventional measurement arrangements (Duvall et al. 1995; Hughes et al. 2005) such as centre-annulus, where the measured signal at a central point is cross-correlated with the signal in an annular region centred on it, result naturally in an enhanced sensitivity in the surface layers particularly beneath the central point. Jensen (2002) examined the concept of *deep-focussing* (Duvall et al. 1995; Kosovichev et al. 2001) in the sense of an arrangement that would provide enhancement deeper in the convection zone. They encountered problems with the quality of their solution recovering a shell of sensitivity surrounding a target location rather than focussed on it.

We suggest that it is possible to use an optimisation technique to design observational masks for specific subsurface target locations. In general one would start with a full set of observational data covering a significant region of the solar surface and attempt, via the method outlined in this paper, to compute a weighting for each data point which emphasises sensitivity to a particular target region of the subsurface. In practice the computational task is considerably simplified by starting with only a subset of these data which we guess a priori to contain the bulk of the relevant signal. In this work we choose a naïve subset of the data; a rectangular region containing the pixels at which measurements are made, and use an algorithm to find the weights that will provide the enhanced sensitivity.

In practice it is possible to apply weightings either directly to Doppler velocity (or intensity) measurements at the pixels, or to the inferred travel times after these have been obtained from the measurements. The former necessarily results

* Present address: AWE Aldermaston, Reading, Berkshire, RG7 4PR, UK.

in a single travel time that is related to concentrated sensitivity (a “single-measurement” scheme), the latter involves a weighted average of a set of travel times to achieve a similar focussing (a “multiple-measurement” scheme). In this paper we seek to outline a general scheme for these two approaches and to provide some diagnostics for preliminary evaluation of their usefulness and applicability. An exploration with real data is left for future work.

2. Statement of the problem

Our suggestion for construction of a focussing scheme rests on building optimized masks to compute a weighted average of data; an analogous forward model must be constructed to link this data with the perturbation of interest as in Eq. (1). To set the idea of these masks in context it is useful to discuss first the pertinent details of the measurements made at the solar surface and the relationship with the construction and application of the sensitivity kernels.

2.1. The form of the data

In time-distance helioseismology one uses cross-correlation to recover relevant observables, typically travel times, that describe local wave-propagation in the Sun. Gizon & Birch (2002) suggest a general framework for the consideration of travel times, the basic formalism of which we paraphrase here.

The cross-correlation of two oscillation signals at separate locations can be written:

$$C(\mathbf{1}, \mathbf{2}, t) = \frac{1}{T} \int_{-\infty}^{\infty} dt' \phi(\mathbf{1}, t') \phi(\mathbf{2}, t' + t), \quad (2)$$

where $\phi(\mathbf{1}, t)$ is the signal at time t at point (or region) $\mathbf{1}$ on the surface and $\phi(\mathbf{2}, t)$ is the signal at point (or region) $\mathbf{2}$. The integration is performed over time T representing the total period of observation. By varying the separation of observing points $\Delta = \|\mathbf{1} - \mathbf{2}\|$ and the time-lag t , one can produce solar seismograms that exhibit maxima corresponding to wave packets travelling between $\mathbf{1}$ and $\mathbf{2}$. In practice, this convolution can be efficiently computed in the Fourier domain.

Since a cross-correlation picks up the wave packet, we can hope to measure the time of travel of that wave packet between points $\mathbf{1}$ and $\mathbf{2}$. It is possible to consider the effect of cross-correlation on a superposition of global normal modes analytically (Kosovichev & Duvall 1997; Giles 2000; Jensen 2002), the procedure then is to identify features of the measured wave packet with the group and phase travel-times between $\mathbf{1}$ and $\mathbf{2}$ and to pick one of these as representative of a travel-time. But it is also useful to define travel time in a different way, as we shall now discuss.

We expect, with the addition of some local disturbance, that the result of the cross-correlation will shift away from some reference correlation i.e. that obtained in quiet Sun. The least-squares difference between the measured and this reference cross-correlations is then:

$$X_{\pm}(\mathbf{1}, \mathbf{2}, t) = \int_{-\infty}^{\infty} dt' f(\pm t') [C(\mathbf{1}, \mathbf{2}, t') - C^{\text{ref}}(\mathbf{1}, \mathbf{2}, t' \mp t)]^2, \quad (3)$$

where $C^{\text{ref}}(\mathbf{1}, \mathbf{2}, t' \mp t)$ is the reference and $C(\mathbf{1}, \mathbf{2}, t')$ is a measured correlation. Using Eq. (2) one can detect waves travelling both forward ($\mathbf{1} \rightarrow \mathbf{2}$) or backward ($\mathbf{2} \rightarrow \mathbf{1}$) between the measurement points which correspond to positive and negative time

lags respectively; $f(\pm t')$ is a window function chosen to isolate particular parts of these two branches. Finding the minimum of Eq. (3):

$$\tau_{\pm}(\mathbf{1}, \mathbf{2}) = \arg \min_t \{X_{\pm}(\mathbf{1}, \mathbf{2}, t)\}, \quad (4)$$

then yields our travel time τ_{\pm} . Gizon & Birch (2002, 2004) give the solution to this problem:

$$\tau_{\pm}(\mathbf{1}, \mathbf{2}) = \mp \frac{\int_{-\infty}^{\infty} dt f(\pm t) \dot{C}^{\text{ref}}(\Delta, t) \times [C(\mathbf{1}, \mathbf{2}, t) - C^{\text{ref}}(\Delta, t)]}{\int_{-\infty}^{\infty} dt' f(\pm t') [\dot{C}^{\text{ref}}(\Delta, t')]^2}, \quad (5)$$

where \dot{C}^{ref} denotes the time derivative of the reference correlation. For a specific analytic form of C^{ref} , e.g. the Gabor wavelet, it is then possible to calculate travel times directly from this linear combination.

To first order, isotropic and anisotropic contributions to τ can be separated by summing or taking differences between the positive and negative time-lags. It can therefore be argued that the sound speed and material flows can be isolated because they are isotropic and anisotropic effects respectively. In this work we concentrate on the anisotropic component:

$$\tau_{\text{diff}} = \tau_{+} - \tau_{-}, \quad (6)$$

which we consider for the purposes of this work as being related only to material flows, through a choice of solar model.

The raw data are observations of line-of-sight Doppler velocities made for example by the Michelson Doppler Imager (MDI) (Scherrer et al. 1995) or by the Global Oscillations Network Group (GONG) (Harvey et al. 1996), which are first filtered and tracked, as described for example in Giles (2000) and Rajaguru et al. (2004), to ensure sufficient quality of surface oscillation measurements that cross-correlations can be usefully interpreted. This process aims to remove unwanted contributions from super-granulation and rotation for example. The fluctuations are then cross-correlated to recover travel times of waves between points on the solar surface.

Generally, spatial averaging over data pixels is also employed, to produce a cross-correlation with sufficiently high signal-to-noise for a reliable measurement. This averaging usually has a particular geometric form, for instance a central point is correlated with a surrounding annulus, as this will constructively add the signals of waves travelling a particular horizontal distance to or from the central point; such a choice of geometry leads to an averaging kernel which is not optimal, as described briefly in Sect. 1, which one can anticipate undermining attempts at focussing on a specific point of choice. In the future, using the method of computing travel times laid out in Eq. (2) to Eq. (5) after Gizon & Birch (2002, 2004), and with the improved quality of data from the Helioseismic and Magnetic Imager (HMI) scheduled for launch in 2008, it may be possible to use single pairs of pixels for recovering travel times satisfactorily, allowing greater control over the form of the sensitivity kernels.

2.2. The forward model

With travel-time data in hand, one seeks to determine subsurface conditions in the Sun. Until recently, the most common approach has been to use the high-frequency ray approximation, in which the sensitivity of the data to subsurface parameters is presumed to be confined to an infinitely thin ray-path linking the measurement points. We use this approach in this work as a preliminary

step, but note the findings of Bogdan (1997) (for instance) indicate that the radial order of modes in the Sun is too low to justify such an approximation and the waves will in fact sample a much larger subsurface region. The use of wave-kernels generated in the Born approximation (Birch & Kosovichev 2000; Gizon & Birch 2002) or Rytov approximation (Marquering et al. 1999; Jensen & Pijpers 2003) is preferable, and this approach (or a similar wave-based conceptualisation) should be used for an operational scheme. We shall analyse both the ray and the wave-kernel approach despite the incorrectness of the ray approach; we seek to show that it is always worthwhile to optimise the weighting mask to get the best possible localisation.

As we can separate flow perturbations by a combination of travel times (Eq. (6)), the expression in Eq. (1) can be written as a convolution of a flow-perturbation to a one-dimensional background model of the Sun with a consistent forward model determined in e.g. the ray approximation. By discretising this problem onto a mesh it becomes possible to use an inversion method such as LSQR Paige & Saunders (1982) to solve the problem

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \epsilon. \quad (7)$$

The rows of matrix \mathbf{A} then represent the sensitivity kernels in Eq. (1) to perturbations \mathbf{x} spread over some volume of the Sun. This results in the travel times \mathbf{b} which will have some associated measurement error ϵ .

2.3. The relationship between grids

Several different grids become important in understanding and refining the problems of optimal masks, largely because of the constraints of observations. Here we introduce some distinct terms to avoid a confusion of terminology.

We consider a two-dimensional *data grid* similar to that shown in Fig. 1 constituting a horizontal map of notional pixels (discussed in Sect. 2.1), each value of which corresponds to either an individual data-point, or to a weighted average of these primitive data-points (a “super-pixel”). Cross-correlation of all the measurements in these pixels provides a set of travel times which can be summed to give a final measurement. The details of this process will be revisited in Sect. 3.1 as they are important to the practicality of making travel-time measurements.

The *model grid* is the grid on which the three-dimensional discrete model is computed. It has been chosen in this work to give a reasonable size of problem whilst providing sufficient resolution. For a particular pair of pixels or super-pixels the forward model is computed on this grid. For simplicity we choose to align our model grid and data grid so that the centres of the data pixels (or super-pixels) are aligned vertically with the centres of the model grid cells.

3. Application of SOLA to the problem

We know in principle how one moves from a set of observations \mathbf{b} , through a set of sensitivity kernels \mathbf{K} , to recover our solution the unknown perturbation \mathbf{x} . The challenge is to construct a framework, which is computationally practical, for choosing a combination of these kernels that focuses the sensitivity in a desired location. Later, we also need to understand what this corresponds to in terms of making and understanding actual observations.

The method of subtractive optimally localized averages (SOLA) due to Pijpers & Thompson (1992) is similar in spirit to the optimally localized averages (OLA) method

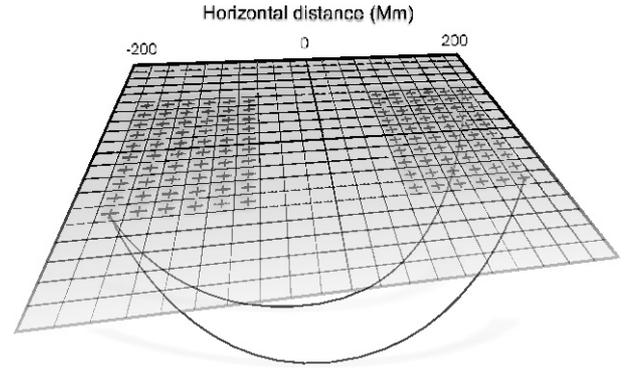


Fig. 1. A schematic of the basic pixel framework for optimal mask generation showing how the measurement points lie in relation to each other at the surface of the Sun. One can then calculate the forward model for each pair of measurement points (+) based on this geometry. For the ray approximation one can easily visualise the rays linking any two points and two of the 3600 possible rays have been illustrated. The region is subdivided with black line corresponding to the horizontal scale of the model grid cells, which in this case are chosen so that the grid cells align precisely with the measurement points.

(Christensen-Dalsgaard et al. 1990) based on the work of Backus & Gilbert (1968). In OLA a set of coefficients $c_i(\mathbf{r}_0)$ is constructed so that the *averaging kernel* defined by

$$\mathcal{K}(\mathbf{r}_0, \mathbf{r}) \equiv \sum_{i \in \Gamma} c_i(\mathbf{r}_0) K_i(\mathbf{r}), \quad (8)$$

in the continuous limit is peaked around $\mathbf{r} = \mathbf{r}_0$ and is small elsewhere. Here the $K_i(\mathbf{r})$ are the sensitivity kernels that constitute the rows of the matrix \mathbf{A} in our forward problem (Eq. (7)). These individual sensitivity kernels form the set of forward models Γ that link pairs of notional pixels from the measurements, and the averaging kernel is a linear combination of these components. This is subject to further normalisation constraint that:

$$\int \mathcal{K}(\mathbf{r}_0, \mathbf{r}) d\mathbf{r} = 1. \quad (9)$$

The averaging kernel describes how the estimated solution $\bar{\mathbf{x}}(\mathbf{r}_0)$ at a particular point \mathbf{r}_0 is a combination of the elements of the true solution $\mathbf{x}(\mathbf{r})$ at all points \mathbf{r} :

$$\bar{\mathbf{x}}(\mathbf{r}_0) = \int \mathcal{K}(\mathbf{r}_0, \mathbf{r}) \mathbf{x}(\mathbf{r}) d\mathbf{r}. \quad (10)$$

By making a good choice for the c_i we can limit the spread of the dependence of our solution in the true solution.

How should the coefficients c_i be picked? In SOLA they are chosen so as to minimise

$$\int [\mathcal{K}(\mathbf{r}_0, \mathbf{r}) - \mathcal{T}(\mathbf{r}_0, \mathbf{r})]^2 d\mathbf{r} + \lambda \sum_{ij} c_i \mathbf{E}_{ij} c_j, \quad (11)$$

(Pijpers & Thompson 1994) subject to the constraint in Eq. (9), where \mathbf{E} is the variance-covariance matrix of the errors in the data \mathbf{b} , $\mathcal{T}(\mathbf{r}_0, \mathbf{r})$ is the target function and λ is a trade-off parameter with some chosen value. Note that λ is included implicitly in the numerical schemes employed in our solution scheme and is dropped from the following discussion. A good solution of Eq. (11) would produce a \mathcal{K} close to \mathcal{T} .

Ideally $\mathcal{T}(\mathbf{r}_0, \mathbf{r})$ would be non-zero only at the point of interest i.e.

$$\mathcal{T}(\mathbf{r}_0, \mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0); \quad (12)$$

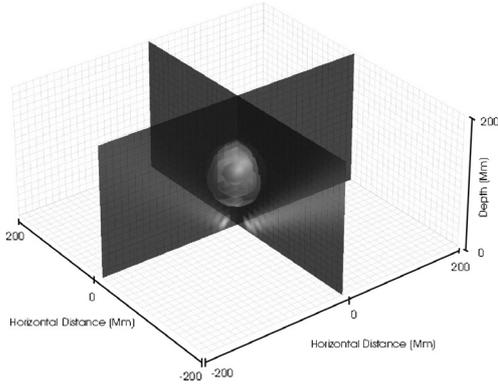


Fig. 2. The Gaussian target function used for the SOLA method, although we show it here with unit amplitude rather than normalised over all space. The two horizontal dimensions indicate the x - and y -directions with distances in megametres, and the vertical dimension is increasing depth also in megametres. The target is horizontally centred on the middle of the model grid as indicated by the locations of the half-slices. The central depth of the Gaussian is 90 Mm. Isosurfaces are shown at values of 0.66, 0.2 and 0.05 indicating the spatial extent of the function.

however such an extreme localization is unrealistic and tends to give poorly localized averaging kernels (Pijpers & Thompson 1992). Instead it is more common to adopt a target with a spatial Gaussian profile:

$$\mathcal{T}(\mathbf{r}_0, \mathbf{r}) = \frac{1}{N} \exp \left\{ - \left(\frac{\mathbf{r} - \mathbf{r}_0}{\delta \mathbf{r}} \right)^2 \right\}, \quad (13)$$

where $\delta \mathbf{r}$ is the FWHM of the function and N is the normalisation coefficient such that the integral of $\mathcal{T}(\mathbf{r}_0, \mathbf{r})$ over all space is unity. Our target function is shown in Fig. 2.

For this SOLA case we then seek solutions to the discrete problem:

$$\bar{\mathbf{c}} = \arg \min_{\mathbf{c}} \{ \| \mathbf{K} \mathbf{c} - \mathbf{T} \|^2 \}, \quad (14)$$

where \mathbf{K} is the combination of discrete sensitivity kernels, \mathbf{c} is the set of coefficients for which $\bar{\mathbf{c}}$ is the optimized set, and \mathbf{T} is the target function on the model grid. By “discrete” we mean calculated on the model grid outlined in Sect. 2.3. The matrix \mathbf{K} will be $m \times n$ in size set by the total number of model grid cells m and by the number of measurement points n i.e. the number of individual sensitivity kernels required to link pairs of points.

One can solve the problem using the LSQR algorithm to recover an optimal set of coefficients. One can then recover a single travel time:

$$t = (\mathbf{K} \bar{\mathbf{c}}) \cdot \mathbf{u}. \quad (15)$$

which is a measurement that is hopefully primarily due to flows \mathbf{u} in the target area. Note that both sides of Eq. (9) are dimensionless, but this yields an averaging kernel with dimensions L^{-3} from Eq. (8) the inversion coefficients are no longer dimensionless. Moreover, the solution t on the left-hand side of Eq. (15) no longer has dimensions of time; rather its dimensions are the same as those of \mathbf{u} .

3.1. A single-measurement approach: *n*SOLA

Thus far we have associated individual weights \mathbf{c} with individual data (i.e. travel times) and hence with *pairs* of points cross-correlated at the solar surface. We may instead wish an algorithm to select weights to be applied to individual pixels prior

to computing the cross-correlations, i.e. treating the data not as the travel times but as the Doppler signal at individual pixels. In that case the weights will be associated with single pixels at the surface. The condition that a pixel has a single associated weighting can be enforced by rewriting the averaging kernel defined in Eq. (8) as

$$\mathcal{K}(\mathbf{r}_0, \mathbf{r}) \equiv c_1(\mathbf{r}_0) \mathbf{K}_{1,2}(\mathbf{r}) c_2(\mathbf{r}_0), \quad (16)$$

where we now have two sets of coefficients c_1 and c_2 corresponding to the two regions, named 1 and 2, that we are cross-correlating, and the kernel $\mathbf{K}_{1,2}$ has been reformulated to represent the sensitivity linking the particular pair of pixels in the two regions.

One can make use of the relationship between the coefficient groups c_1 and c_2 to reduce the problem. It can be assumed that lines of symmetry will be present in the model and any grid used (although \mathbf{K} itself will not be symmetric for an arbitrary choice of pixels in regions 1 and 2) so one can suggest the case $c_1 = c_2$ for the indexing scheme discussed in Sect. 2.3 and shown in Fig. 1. Appendix B demonstrates that one can always find a symmetric set that is at least as good as an asymmetric set. Moreover the symmetric solution has as good or better error propagation properties. One can therefore exclude non-symmetric solutions without loss of quality of the inversion, and by explicitly formulating the problem as symmetric one obtains a numerically more robust solution.

Using Eq. (16) in Eq. (11) we must find the minimum of the quartic problem:

$$\bar{\mathbf{c}} = \arg \min_{\mathbf{c}} \{ \| \mathbf{c} \mathbf{K} \mathbf{c} - \mathbf{T} \|^2 \}, \quad (17)$$

which is more complex to solve as it is a non-linear problem. Here \mathbf{c} corresponds to the set of coefficients for the two regions of pixels and \mathbf{K} to the model. Hence:

$$\mathbf{c} = \mathbf{c}_1 = \mathbf{c}_2. \quad (18)$$

In this approach \mathbf{K} has size $l \times l \times m$ with m retaining its meaning as in Sect. 3, and l being the number of measurement points in each measurement region. A particular choice of row and column of \mathbf{K} represents a pair of measurement pixels; the matrix contains at this location an individual sensitivity kernel linking the two points.

A solution is possible here using a non-linear curve-fitting algorithm; we use the MATLAB routine `lsqnonlin`. A travel-time difference is then associated with:

$$\delta t = (\bar{\mathbf{c}} \mathbf{K} \bar{\mathbf{c}}) \cdot \mathbf{u}. \quad (19)$$

3.2. The spectral components

Filtering in frequency and in phase speed is used to improve the cross-correlation signal in practical time-distance applications. The filtering discards many of the modes that are present in the Sun in order to isolate, for instance, waves of a particular horizontal wave-speed. Birch (2004) and Birch et al. (2004) computed wave kernels constructed with this in mind, yielding sensitivity kernels significantly different from those corresponding to unfiltered data, or corresponding to monochromatic signals depending on the filtering that was applied.

Formally the wavelet constructed by the various possible filtering schemes can be expressed in a normal mode expansion and an associated spatial sensitivity kernel would then correspond to a sum of individual normal mode kernels with

weights f_p in which the index p runs over all normal modes. Constructing kernels from specific modal components introduces an extra set of weightings f_p that can be included in the optimisation i.e.

$$\mathbf{K} = \sum_p \mathbf{K}_p f_p, \quad (20)$$

where each \mathbf{K}_p is a sensitivity kernel for a mode p . The interference of the \mathbf{K}_p for the relevant modes can significantly change the character of the final sum in the case of wave kernels. The extra control may help with the localization of the signal but will increase the number of coefficients to be determined and add an extra layer to both the single- and multiple-measurement schemes. Although a possible future development, the data does not currently warrant this added degree of complexity.

For example, minimizing the problem (Eq. (17)) yields

$$(\bar{\mathbf{c}}, \bar{\mathbf{f}}_p) = \arg \min_{\mathbf{c}} \{ |\mathbf{c} \mathbf{K} \mathbf{f} \mathbf{c} - \mathbf{T}|^2 \}, \quad (21)$$

where both the coefficients \mathbf{c} and the filter weights \mathbf{f} have to be chosen by the algorithm.

In this work we assume that the weights have a Gaussian spectrum (as is approximately the case for the Sun) instead of solving for them as part of the optimization problem. We can then compute the kernels more efficiently and solve a simpler optimisation problem. This is a preliminary study in which we wish to lay out a general framework for the optimal masks technique; the extra complication that proper inclusion of the modes introduces is left for a later refinement, which should in principle be a straightforward extension of the calculation in the Rytov approximation following Jensen & Pijpers (2003).

3.3. Making measurements

We can now estimate a travel time optimised to provide custom sensitivity kernels using either of two approaches, motivated by the measurement scheme to be employed on real data. In the multiple-measurement paradigm we allow each component sensitivity kernel to have an independent weighting, yet this means we must make one measurement for every kernel (i.e. each pair of measurement points) that we include in the problem. The travel times τ computed for all the measurements can then be combined in a weighted average:

$$t = \mathbf{c} \tau. \quad (22)$$

Here τ is the set of measurements obtained for each measurement configuration and \mathbf{c} are the coefficients from Eq. (14). Whilst the associated problem Eq. (14) for finding \mathbf{c} is straightforward, use of real data requires a large number of measurements to be made by an observer.

In contrast the single-measurement approach demands more work to obtain the relevant coefficients but fewer actual observations. In principle this simplifies the task of the observer, but the measurement process requires further clarification.

The multiple measurement approach required a linear combination of travel times as obtained in Eq. (4). Instead we now require to carry out averaging on the basic oscillation data prior to cross-correlation to permit a single measurement to be made of a travel time directly. Combining signals in this way will produce a cross-correlation over the pixels of interest, and we can no longer necessarily expect this to correspond to the reference

correlation of a single feature, but the sum of cross-correlations:

$$\begin{aligned} \bar{C}(t) &= \sum_{i,j} \frac{1}{T} \int_{-\infty}^{\infty} dt' a_i \phi_i(t') a_j \phi_j(t' + t) \\ &= \sum_{i,j} a_i c_{ij} a_j, \end{aligned} \quad (23)$$

where we have introduced the constants a_i to modify the signal, which we seek to relate to the c_i (Eq. (16)). It is evident that the resulting cross-correlation is a linear combination of the individual cross-correlations. For application of nSOLA we further require that the resulting travel time estimate from fitting this cross-correlation is a linear combination of the travel times that would be obtained from fitting the individual cross-correlations. We see from Eq. (5) that the relationship will indeed be linear (and the a_i equivalent to the c_i) provided the same reference cross-correlation C^{ref} and window function $f(t)$ are used. This would be the case for example in centre-annulus geometries where the distances between all the pixels being cross-correlated are the same. It may not be desirable for arbitrary geometries. We propose that nSOLA can be applied to geometries (obtained perhaps by breaking the problem at hand into a series of sub-geometries) in which one can reasonably apply the same C^{ref} and $f(t)$ to all the pixel pairs involved.

3.4. Error magnification

If we construct an optimized mask using the SOLA or nSOLA case what effect will this have on the error? It will alter the magnitude of the noise that enters the solution since this will enter as a weighted root-mean square (rms) term. In the SOLA case this will look like:

$$\sigma(\bar{\mathbf{r}}_0)^2 = \sum_i \mathbf{c}_i(\mathbf{r}_0)^2 \sigma(\mathbf{b}_i)^2, \quad (24)$$

so the variance of the solution at a target cell \mathbf{r}_0 is given by the sum of squares of the averaging kernel coefficients \mathbf{c} of that cell, multiplied by the variances of the associated data $\sigma(\mathbf{b})$. Hence we recover the error magnification, or the degree to which our optimization exacerbates the problem of noise:

$$\Lambda(\bar{\mathbf{r}}_0)^2 = \sum_i \mathbf{c}_i(\mathbf{r}_0)^2. \quad (25)$$

This will allow an estimation of this magnification to be associated with an artificial measurement. Quantitative use of this information is only possible if the $\sigma_i(\mathbf{b})$ are known.

4. Averaging kernels

We now apply the ideas in Sect. 2 to generate coefficient sets for both SOLA and nSOLA cases, and examine the averaging kernels produced to ascertain the quality of the localization. We employ the data-sampling method described in Sect. 2.3 and we seek initially to compare the ray results for the SOLA and nSOLA cases before examining the use of wave-based kernels.

4.1. SOLA analysis

To determine the effect of altering the number of measurement points in the problem we used a constant number in the x -direction but varied the number in the y -direction, arranged

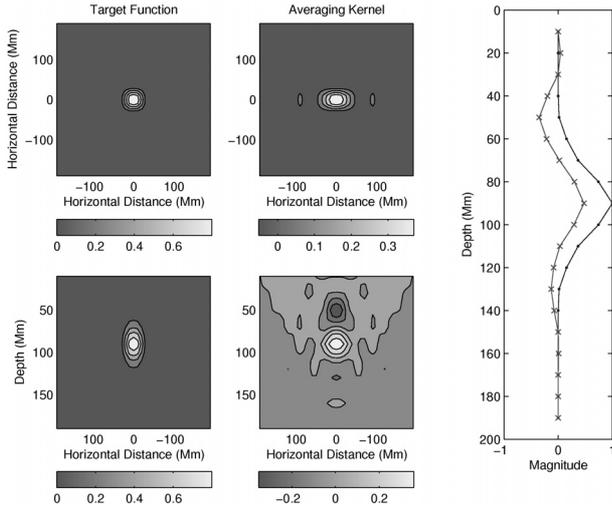


Fig. 3. Contour plots of horizontal (*top row*) and vertical (*bottom row*) slices comparing the target T (*left column*) with the solution \bar{K} (*right column*) computed via standard SOLA. On the far right hand side the plot shows magnitude at the horizontal centre for the model (dotted) and solution (crossed) as a function of depth. This example is from a ray theoretic calculation from LSQR after 20 iterations with 6 pixels in the x -direction and 2 in the y -direction yielding 144 possible ray paths. The target and the solution have been divided by the maximum value of the target function for comparison.

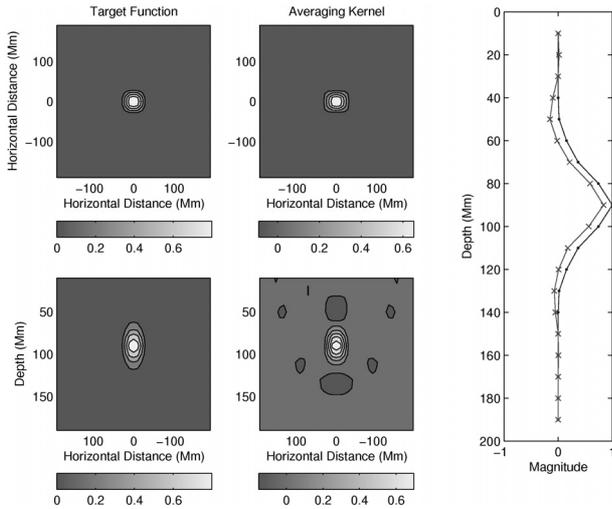


Fig. 4. As Fig. 3 but for 6 pixels in the y -direction yielding 1296 possible ray paths.

symmetrically with 2, 6 and 10 pixels. An example geometry is illustrated schematically in Fig. 1 with ten points in the y -direction.

The examples in Figs. 3–5 show the averaging kernels resulting from SOLA analysis of the three cases respectively, using the ray approximation, in comparison with the target function. The expected result that the localization of the kernels is improved by the addition of more data points and associated ray paths appears to be evident. The evolution of the averaging kernel is very encouraging as one sees that the structures surrounding the peak disappear and that the magnitude plots become more similar.

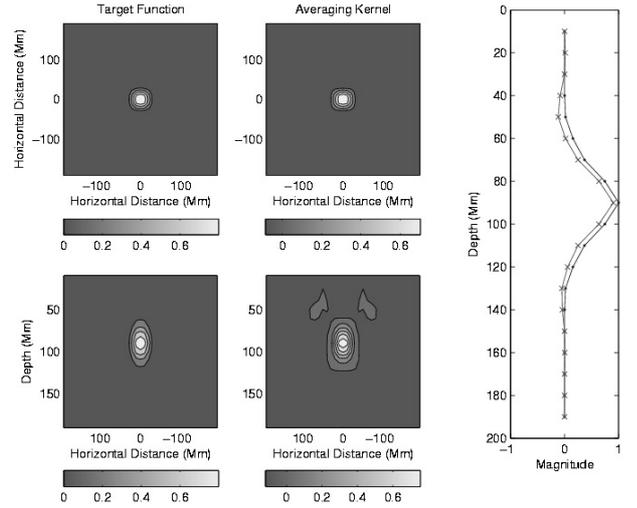


Fig. 5. As Fig. 3 but for 10 pixels in the y -direction yielding 3600 possible ray paths.

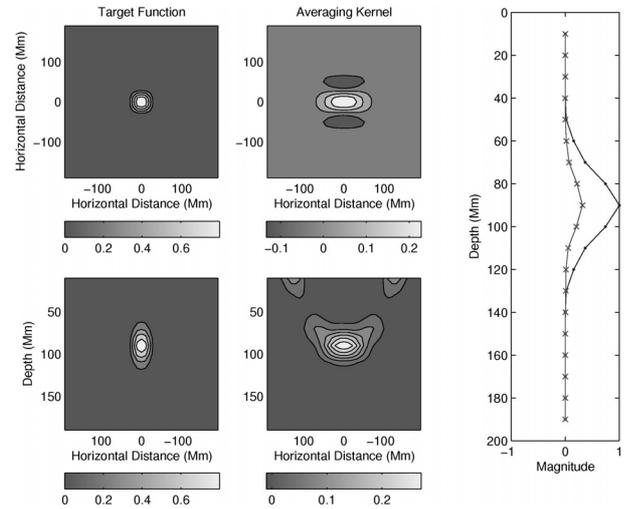


Fig. 6. Contour plots of horizontal (*top row*) and vertical (*bottom row*) slices comparing the target T (*left column*) with the nSOLA solution K^* (*right column*) computed via the MATLAB non-linear fitting algorithm `lsqnonlin`. On the far right hand side the plot shows magnitude at the horizontal centre for the model (dotted) and solution (crossed) as a function of depth. This example is from a ray theoretic calculation from with 6 pixels in the x -direction and 10 in the y -direction yielding 3600 possible ray paths. The target and the solution have been divided by the maximum value of the target function for comparison.

4.2. nSOLA analysis

A similar analysis to that in Sect. 4.1 was then carried out using the nSOLA analysis. Figure 6 shows the corresponding averaging kernel for the 6×10 case. In this case it is clear that we do much worse in localizing the signal, particularly near the measurement points themselves: this is because of the strong sensitivity here in the ray paradigm. This would mean strong surface components would completely dominate the resulting travel-time measurement. Whilst such a result is unsurprising given the reduced degrees of freedom in the optimization, things again improve with an increase in the number of data-points.

Coefficients recovered from this analysis are easy to display as shown in Figs. 7 and 8. In the first plot the set from the 6×2 case reveal a maximum centred on points that sample the target region directly, surrounded by negative regions that compensate

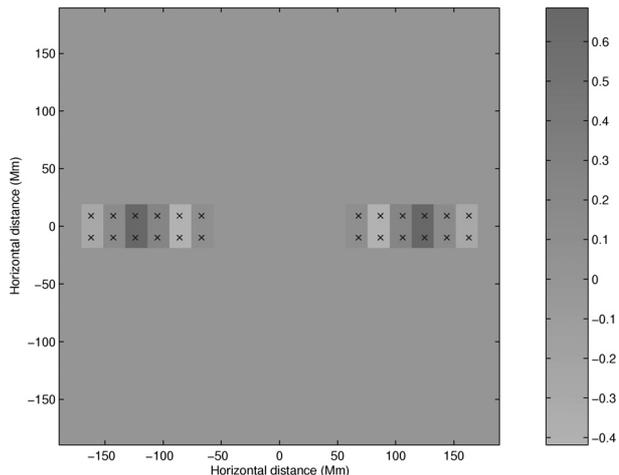


Fig. 7. The normalised coefficients c for the nSOLA analysis of the 6×2 pixel case. The ray footprints are indicated with a cross.

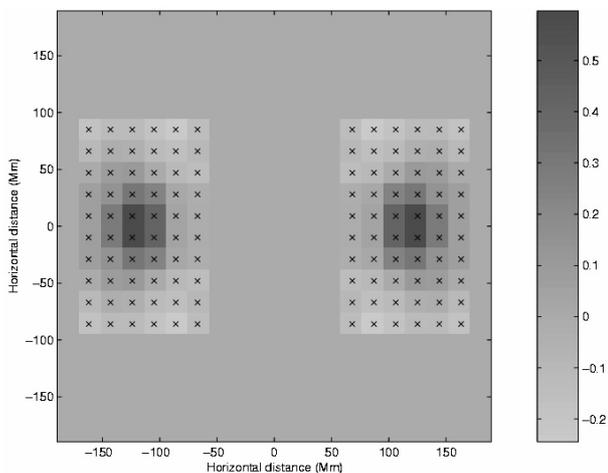


Fig. 8. The normalised coefficients c for the nSOLA analysis of the 6×10 pixel case in Fig. 6. The ray footprints are indicated with a cross.

for the sensitivity to the near surface measurements, although as we have seen this is not satisfactorily achieved.

The expectation with the move to the 6×10 case might have been that arc structures would be seen as these coefficients would be the ones most sensitive to the target. Instead one sees a central oval of positive values surrounded by a halo of negative coefficients. However, if one thinks of these pixels as linked to rays that are largely insensitive to the target region, then it is possible to imagine that they are used instead to reduce signal in other areas, such as near the surface. This would account for the presence of the negative values.

The nSOLA scheme is not required operationally, since a measurement can itself be an average of several measurements as mentioned in Sect. 2 made in order to obtain a good travel time. It seems from the results presented in this section that SOLA provides a much better localization, and so we shall proceed using this approach but using the super-pixels discussed in Sect. 2.3.

4.3. Using a Rytov model

One can employ the Rytov approximation in exactly the same framework, but in this case the sensitivity of the measurements is dependent on a finite (and quite significant) volume of the

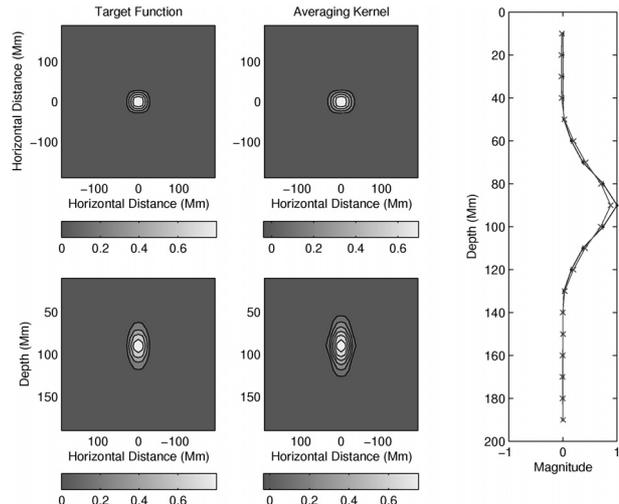


Fig. 9. Contour plots of horizontal (*top row*) and vertical (*bottom row*) slices comparing the target T (*left column*) with the solution \bar{K} (*right column*) computed via standard SOLA. On the far right hand side the plot shows magnitude at the horizontal centre for the model (dotted) and solution (crossed) as a function of depth. This example is from a Rytov model calculation from LSQR after 20 iterations with 6 pixels in the x -direction and 10 in the y -direction yielding 3600 possible ray paths. The target and the solution have been divided by the maximum value of the target function for easy comparison.

solar region of interest. This may significantly alter the qualities of the averaging kernel produced by optimization.

We begin, as with the ray case, looking at the SOLA case. The results shown in Fig. 9 are for the 6×10 case, a visual comparison with Fig. 5 indicates that we are doing even better with the use of the Rytov approximation than when using rays. We interpret this as being a benefit of the extended spatial structure of such a model, which allows for better cancellation of signal away from the target location.

Jensen (2002) showed that Rytov kernels have a richer structure, but if applied naively have a poorer sensitivity at the target location. With an nSOLA analysis it can be seen that the richer structure can in fact be exploited. Examining the coefficients generated from an nSOLA analysis we see that the optimized choice uses the sense of the spatial structure of the Rytov models to overcome this difficulty. There are positive and negative components in such wave-based models (Birch & Kosovichev 2000; Jensen & Pijpers 2003) and these can be combined to produce positive interference at the target and negative interference elsewhere. It would be difficult to construct such a scheme by hand.

One problem with such a set of coefficients is that the rms value may be much larger because of the contribution of the negative terms. This is a potentially serious problem that could undermine measurements with real data. We shall attempt to quantify this problem in experiment by analysing the error magnification described by Eq. (25).

An improvement on the error analysis would be obtained by incorporating a full error (co-)variance matrix in Eq. (24). Techniques for obtaining this matrix have been shown in Jensen et al. (2003) and Gizon & Birch (2004).

5. Synthetic experiments

The next aim of development is to design an operational SOLA Rytov scheme. Note that we are not concerning ourselves with a

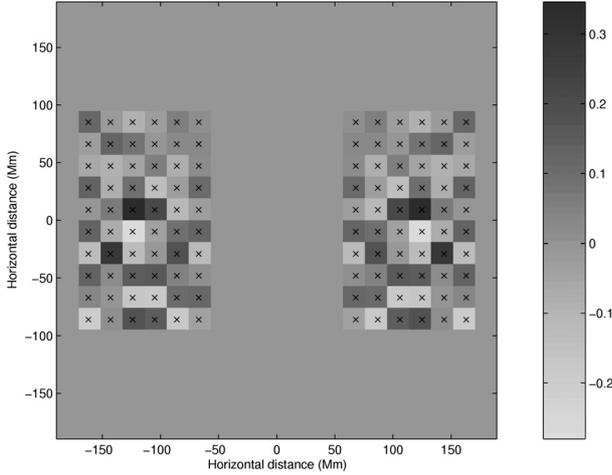


Fig. 10. The normalised coefficients c for the nSOLA analysis of the 6×10 pixels using a Rytov forward model. The footpoints are again indicated with a cross.

proper treatment of noise here: the artificial data have not been perturbed with random errors. We proceed with an arc-to-arc geometry such that the model represents averaging over two arcs. These arcs are each Gaussian in polar coordinates locally centred on a notional measurement point:

$$f(R, \phi) \propto \exp\left\{\frac{R - R_0}{\delta R^2}\right\} \exp\left\{\frac{\phi - \phi_0}{\delta \phi^2}\right\}, \quad (26)$$

lying on a common circle of radius R_0 equivalent in magnitude to the separation of these notional points. The two values of $0 \leq \phi_0 < 2\pi$ for each arc are separated by π . The FWHM in R is $\delta R = 1$ Mm and in azimuth ϕ is $\delta \phi = 1/32$ of a radian. Each arc is then equivalent to a super-pixel and an arc-to-arc model must be constructed for every pair of these notional measurement locations as shown in Fig. 1 for the 6×10 case. This is a time-consuming computation and it is important to efficiently calculate the model by determining in advance any repetition in R_0 and in ϕ_0 .

A simple synthetic meridional circulation was devised of the form:

$$u(z) = A \left(\frac{B - z}{200}\right)^3 \cos\{\pi z/C\}, \quad (27)$$

with z being depth into the Sun in mega-meters, A , B and C are coefficients set at 20 ms^{-1} , 200 Mm and 150 Mm respectively. This is not intended to accurately reproduce the flow in the Sun; our aim here is to see if we can detect the deep sign-change of the flow. For validation purposes we will also consider a modified flow where u is set to exactly zero below 80 Mm.

We proceed by computing averaging kernels for different target depths and multiply this result by the flow-field to obtain a representative travel-time measurement for the optimized mask.

This is plotted in Fig. 11 for a range of target depths to “sample” the mock flow. It is encouraging that this result clearly recovers the shape of the flow configuration between about 40 Mm and 140 Mm. There is some smearing in depth but our target function (Fig. 2) is quite coarse in depth; it may be possible to reduce this effect by decreasing the extent of the target function in depth. However, the smearing is also probably due to the extended structure of the Rytov model, and more investigation of this point will be needed in the future to refine the technique.

The same basic set of sensitivity kernels has been used to solve the SOLA problem for all choices of target depth, although

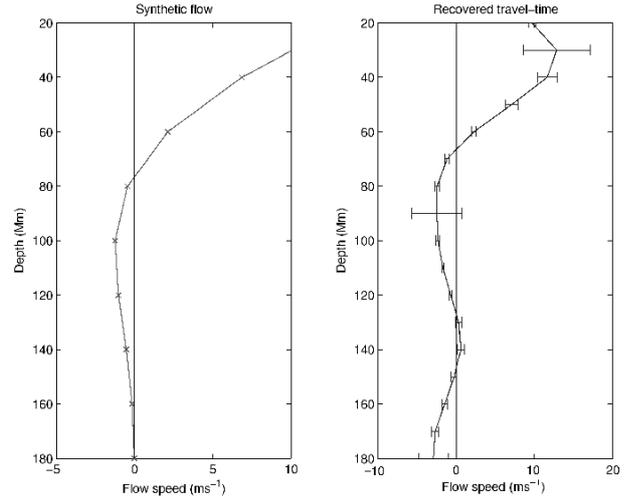


Fig. 11. Optimal-mask data recovery of possible meridional flow using the Rytov paradigm. The left-hand panel shows the original model and the right-hand the recovered measurements.

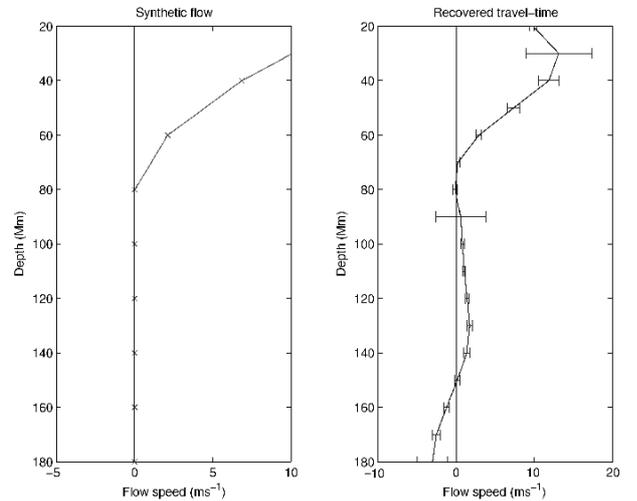


Fig. 12. As in Fig. 11 but with a modified meridional flow.

in practice one would ideally pick the kernels considering the specific target function. When we change the depth of the target we move away from the principal region of sensitivity at 90 Mm depth. In this example the model is consequently poor near the surface and at depth hence the estimate of the flow becomes worse. Another possibility could be to increase the number of data points in the model and hence the number of free parameters. We have arbitrarily chosen to solve an almost square problem with 4000 model cells and 3600 data configurations, yet we could investigate whether anything is to be gained by moving to a significantly over- or under-determined problem.

The error bars shown in Fig. 11 show the magnification as in Eq. (25). Generally the values are moderate despite the concerns that the presence of negative coefficients might inflate this quantity (cf. Fig. 13). Whilst this is an encouraging result there are very large anomalies at 90 Mm and in the shallow and deep layers. We suggest that this is understandable if one considers how the averaging kernel will be constructed. As mentioned, using uniform weighting with the Rytov case (Jensen 2002), one generates a hole in the averaging kernel at the region where maximum signal would be expected. To overcome this issue the optimization uses the sidebands of the Rytov model to interfere

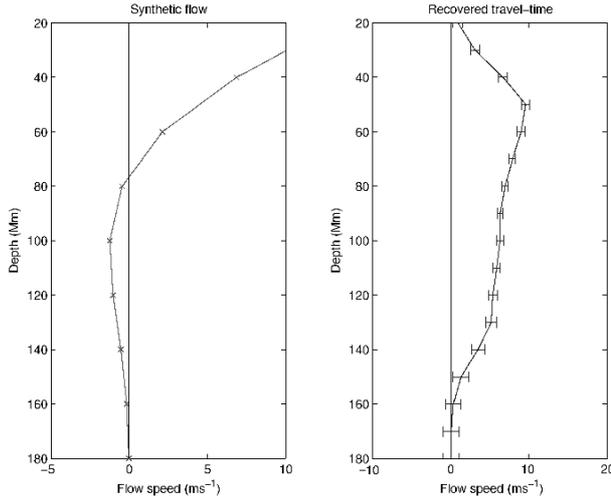


Fig. 13. As in Fig. 11 but using a ray approach.

constructively at the target, using a significant range of positive and negative values to achieve this. This can lead to a significant rms value in regions where the forward model has a maximum coincident with the target; we have built our model to be focussed at 90 Mm and so counter-intuitively we suffer from enhanced magnification most strongly with the target location at the same depth. At shallow and large depths, where the target is far from the maximum, this problem returns.

In Fig. 13 the results are shown from the same experiment made with a ray model. In this case we fail to pick up the signal of the mock flow convincingly, and the strong shallow component appears to dominate. This is strong, potentially disquieting evidence that the ray approximation can be dominated by shallow features even with a highly localized averaging kernel. The error magnification is larger away from the optimal target location as could be expected, but note we do not see the anomalies noted in the Rylov case.

Finally, as a check on the results in Fig. 11 the modified meridional flow (where $u = 0$ below 80 Mm) was used to double-check the measurement of the flip in sign at 80 Mm in the original mock flow. The results in Fig. 12 reveal that one recovers the correct form of the flow in the trustworthy region.

6. Discussion

It is clear from the results presented here that this approach for focussing signal via the use of optimized masks could prove an extremely valuable tool in probing the solar interior. Our aim here was to present a framework for an operational scheme aimed at recovering a flow structure in the Sun. Our examination of the optimization results and the outcome of some simple tests have proved extremely favourable for the basic SOLA scheme described. The next tasks are to refine this scheme and to apply it to real data.

In terms of refining the approach presented here, there are areas of the optimization process that themselves need to be optimized: the distribution of pixels used in this study is not necessarily the most suitable; the target function could be further constrained in depth, and it would be preferable to have an accurate model for the correlation in the data errors – in particular for pixels that are only separated by a single wavelength.

Another area of improvement is proper consideration of the spectrum of modes considered in the optimisation. We have used a Gaussian spectrum here to outline a general approach, but real

data is filtered to exclude unwanted spectral components such as supergranulation and surface modes; it would be important to include the modal components correctly. As outlined in this work, this will provide an extra set of variables to the problem, which may prove very interesting in terms of changes to the quality of the localization.

The approach outlined may have other interesting applications in addition to detecting the return flow. Once a mask has been constructed it can be used very quickly and efficiently to produce synoptic maps of data, hence output of solar subsurface weather maps would be very much enhanced using this technique. These measurements could then be refined with full inversion. In Hughes et al. (2006) we shall present such a scheme and attempt measurements with real data.

Acknowledgements. S.J.H. was supported by a studentship from the UK Particle Physics and Astronomy Research Council (PPARC) and this work was supported in part by PPARC grant PP/X501812/1. We would also wish to extend our thanks to our referee, who provided a very thorough and constructive review of our initial submission.

Appendix A: coefficient asymmetries in optimised masks

The computational step of SOLA, obtaining the linear coefficients, requires solving a large set of linear equations, i.e. performing a matrix factorisation of a large matrix. It is known that matrix factorisation algorithms for symmetric matrices suffer much less from problems of stability and accuracy than do algorithms designed for non-symmetric matrices. It is therefore of interest to demonstrate that in the application of SOLA in this paper, where the sets of coefficients for the two patches to be correlated are not necessarily constrained to be symmetric, the problem can be formulated in terms of a symmetric matrix factorisation.

The forward model that constitutes K is not symmetric i.e.:

$$K_{ij} \neq K_{ji}. \quad (\text{A.1})$$

By splitting up the coefficients into symmetric (c_s) and anti-symmetric (c_a) parts one can write:

$$\begin{aligned} c_s &= \frac{1}{2}(c_1 + c_2) \\ c_a &= \frac{1}{2}(c_1 - c_2). \end{aligned} \quad (\text{A.2})$$

When substituted into the integrand of Eq. (11), the resulting equation is:

$$\begin{aligned} c_1^T K c_2 - \mathcal{T} &= (c_s + c_a)^T K (c_s - c_a) - \mathcal{T} \\ &= c_s^T K c_s + c_a^T K c_s - c_s^T K c_a - c_a^T K c_a - \mathcal{T}. \end{aligned} \quad (\text{A.3})$$

For any matrix A it is true that:

$$c_s^T A c_s = (c_s^T A c_s)^T = c_s^T A^T c_s \quad (\text{A.4})$$

and similarly that:

$$\begin{aligned} c_a^T A c_a &= c_a^T A^T c_a \\ c_s^T A c_a &= c_a^T A^T c_s. \end{aligned} \quad (\text{A.5})$$

In a similar fashion to the coefficients, one can also split up the matrix K into explicitly symmetric and anti-symmetric components:

$$\begin{aligned} K_{s,ij} &= K_{s,ji} \equiv \frac{1}{2}(K_{ij} + K_{ji}) \\ K_{a,ij} &= -K_{a,ji} \equiv \frac{1}{2}(K_{ij} - K_{ji}). \end{aligned} \quad (\text{A.6})$$

If for the matrix A in Eq. (A.4) and Eq. (A.5) the anti-symmetric matrix K_a is substituted, it is clear that it must render the combinations with equal coefficients equal to 0. Equally, substituting for A the symmetric matrix K_s must make the final equation of Eq. (A.5) disappear. Hence the terms in Eq. (A.3) can be re-written:

$$\begin{aligned} c_s^T K c_s &= c_s^T K_s c_s \\ c_a^T K c_a &= c_a^T K_s c_a \\ c_a^T K c_s - c_s^T K c_a &= 2c_a^T K_a c_s. \end{aligned} \quad (\text{A.7})$$

One can now define an explicitly symmetric matrix A and a vector s as:

$$A \equiv \begin{pmatrix} K_s & K_a^T \\ K_a & -K_s \end{pmatrix}, \quad (\text{A.8})$$

$$s \equiv \begin{pmatrix} c_s \\ c_a \end{pmatrix}, \quad (\text{A.9})$$

and with these definitions re-write Eq. (A.3) as:

$$\begin{aligned} c_1^T K c_2 - \mathcal{T} &= c_s^T K_s c_s - c_a^T K_s c_a + 2c_a^T K_a c_s - \mathcal{T} \\ &= s^T A s - \mathcal{T}. \end{aligned} \quad (\text{A.10})$$

This demonstrates that an asymmetric coefficient set can be obtained using numerically stable (and faster) algorithms for factorisation of symmetric matrices.

Appendix B: Superior localisation and error propagation for symmetric coefficient sets

In this appendix it is demonstrated that one should expect that SOLA coefficients that are symmetric between the two patches perform at least as well or better than sets of coefficients that are not symmetric, both in terms of localization and in terms of error propagation.

The starting point is Eq. (A.10):

$$c_1^T K c_2 - T = c_s^T K_s c_s - c_a^T K_s c_a - T_s + 2c_a^T K_a c_s - T_a. \quad (\text{B.1})$$

The functional to be minimised for the weights is:

$$\begin{aligned} &\int (c_1^T K c_2 - T)^2 dx = \\ &\int [(c_s^T K_s c_s - c_a^T K_s c_a - T_s) + (2c_a^T K_a c_s - T_a)]^2 dx \\ &= \int (c_s^T K_s c_s - c_a^T K_s c_a - T_s)^2 dx + \int (2c_a^T K_a c_s - T_a)^2 dx \\ &\quad + \int 2(c_s^T K_s c_s - c_a^T K_s c_a - T_s)(2c_a^T K_a c_s - T_a) dx. \end{aligned} \quad (\text{B.2})$$

The third of the three integrals in Eq. (B.2) is the integral of the product of a function that is symmetric with respect to the midplane of the patches, and one that is antisymmetric:

$$\begin{aligned} K_{s,ij}(\mathbf{r}) &= K_{s,ij}(\mathbf{r} - 2(\mathbf{r} \cdot \mathbf{n})\mathbf{n}) \\ K_{a,ij}(\mathbf{r}) &= -K_{a,ij}(\mathbf{r} - 2(\mathbf{r} \cdot \mathbf{n})\mathbf{n}) \\ T_s &= \frac{1}{2} (T(\mathbf{r}) + T(\mathbf{r} - 2(\mathbf{r} \cdot \mathbf{n})\mathbf{n})) \\ T_a &= \frac{1}{2} (T(\mathbf{r}) - T(\mathbf{r} - 2(\mathbf{r} \cdot \mathbf{n})\mathbf{n})). \end{aligned} \quad (\text{B.3})$$

Integrating that product over the entire volume yields by definition 0. If furthermore the target function is itself symmetric then

the antisymmetric part of it T_a is 0. Equation (B.2) then reduces to minimising:

$$\int (c_s^T K_s c_s - c_a^T K_s c_a - T_s) 2dx + \int (2c_a^T K_a c_s) 2dx. \quad (\text{B.4})$$

The second term is clearly minimal if all antisymmetric coefficients $c_{a,i} = 0$, which also satisfies the constraint that $\sum c_{a,i} = 0$. Since K_s is symmetric it can be written as $K_s = U^T D U$ where U is a unitary matrix and D is diagonal. Define α as follows:

$$\alpha = \min_i |D_i| - \min_i D_i = \begin{cases} 0 & \text{if } \min_i D_i > 0 \\ 2 \min_i |D_i| & \text{if } \min_i D_i < 0. \end{cases}$$

With this definition the matrix defined by:

$$D + \alpha I \equiv P P \quad (\text{B.5})$$

is diagonal and positive definite, and therefore so is its ‘‘square root’’ matrix P . Define a vector w by:

$$w \equiv \begin{pmatrix} w_s \\ w_a \end{pmatrix} = \begin{pmatrix} P & 0 \\ 0 & P \end{pmatrix} \begin{pmatrix} U & 0 \\ 0 & U \end{pmatrix} \begin{pmatrix} c_s \\ c_a \end{pmatrix}, \quad (\text{B.6})$$

it is clear that:

$$\begin{aligned} c_s^T K_s c_s - c_a^T K_s c_a - T_s &= (c_s^T c_a^T) \begin{pmatrix} K & 0 \\ 0 & -K \end{pmatrix} \begin{pmatrix} c_s \\ c_a \end{pmatrix} - T_s \\ &= (w_s^T w_a^T) \begin{pmatrix} D P^{-2} & 0 \\ 0 & -D P^{-2} \end{pmatrix} \begin{pmatrix} w_s \\ w_a \end{pmatrix} - T_s \end{aligned} \quad (\text{B.7})$$

which can be written out in individual elements as:

$$\sum_i \beta_i (w_{s,i}^2 - w_{a,i}^2) - T_s \quad (\text{B.8})$$

in which:

$$\beta_i \equiv d_i / p_i^2. \quad (\text{B.9})$$

Now assume that an acceptable solution w has been found, minimising the first term of Eq. (B.4). Clearly a solution with $w_{a,i} = 0$ would provide the absolute minimal value for the second term, but it is assumed here merely that the minimisation of the first term has yielded a solution which satisfies the weaker constraint that:

$$|w_{a,i}| < |w_{s,i}| \quad \forall i. \quad (\text{B.10})$$

Now construct a $w_{s,i}^*$ such that:

$$\begin{aligned} w_{s,i}^* &= w_{s,i} \cosh \alpha_i - w_{a,i} \sinh \alpha_i \\ w_{a,i}^* &= -w_{s,i} \sinh \alpha_i + w_{a,i} \cosh \alpha_i = 0 \end{aligned} \quad (\text{B.11})$$

which implies that α_i satisfies:

$$\tanh \alpha_i = \frac{w_{a,i}}{w_{s,i}}. \quad (\text{B.12})$$

With these definitions it is true that:

$$\begin{aligned} (w_{s,i}^{*2} - w_{a,i}^{*2}) &= w_{s,i}^{*2} \\ &= w_{s,i}^2 \cosh 2\alpha_i + 2w_{s,i} w_{a,i} \cosh \alpha_i \sinh \alpha_i \\ &\quad + w_{a,i}^2 \sinh 2\alpha_i \end{aligned} \quad (\text{B.13})$$

$$= w_{s,i}^2 - w_{a,i}^2. \quad (\text{B.14})$$

This process can be carried out for each index i independently. Therefore, given a non-symmetric solution (w_s, w_a) of the minimisation of the first part of Eq. (B.4), that satisfies Eq. (B.10),

by the above procedure one can always construct an equally satisfactory symmetric solution ($w_s^*, 0$) of it, i.e. with all $w_{a,i}^* = 0$ and therefore all $c_{a,i} = 0$. Since this also minimises the second part the symmetric solution must be superior. There is therefore no need to consider asymmetric weighting of the patches if the target function T is symmetric.

Assume the data errors for each pixel are translationally invariant and uncorrelated between the first and second patch of pixels, so that the variance-covariance matrices $\Sigma_{1,2}$ satisfy:

$$\Sigma_1 = \Sigma_2 \equiv \Sigma. \quad (\text{B.15})$$

Then the propagated error for the asymmetric solution is:

$$\begin{aligned} \sigma_{p2} &= c_1^T \Sigma_1 c_1 + c_2^T \Sigma_2 c_2 \\ &= (c_1^T \Sigma c_1 + c_2^T \Sigma c_2) \\ &= ((c_s + c_a)^T \Sigma (c_s + c_a) + (c_s - c_a)^T \Sigma (c_s - c_a)) \\ &= (c_s^T \Sigma c_s + c_s^T \Sigma c_a + c_a^T \Sigma c_s + c_a^T \Sigma c_a \\ &\quad + c_s^T \Sigma c_s - c_s^T \Sigma c_a - c_a^T \Sigma c_s + c_a^T \Sigma c_a) \\ &= 2(c_s^T \Sigma c_s + c_a^T \Sigma c_a) \\ &= 2(w_s^T P^{-1} U \Sigma U^T P^{-1} w_s + w_a^T P^{-1} U \Sigma U^T P^{-1} w_a) \\ &= 2w_s^{*T} (C P^{-1} U \Sigma U^T P^{-1} C + S P^{-1} U \Sigma U^T P^{-1} S) w_s^*. \quad (\text{B.16}) \end{aligned}$$

In which C is a diagonal matrix with elements $\cosh \alpha_i$ and S is a diagonal matrix with elements $\sinh \alpha_i$. The propagated error for the symmetric solution is :

$$\sigma_{p2} = 2w_s^{*T} P^{-1} U \Sigma U^T P^{-1} w_s^*. \quad (\text{B.17})$$

Since the determinants satisfy :

$$|C P^{-1} U \Sigma U^T P^{-1} C + S P^{-1} U \Sigma U^T P^{-1} S| \geq |P^{-1} U \Sigma U^T P^{-1}| \quad (\text{B.18})$$

the propagated error is smaller for the symmetric solution compared to the asymmetric one.

References

- Backus, G. E., & Gilbert, J. F. 1968, *Geophys. J.*, 16, 169
 Birch, A. C., & Kosovichev, A. G. 2000, *Sol. Phys.*, 192, 193
 Birch, A. C. 2004, in *Proc. SOHO/GONG 2004, Helio- and Asteroseismology: Towards a Golden Future* ed D. Danesy, 138
 Birch, A. C., Kosovichev, A. G., & Duvall, T. L., Jr. 2004, *AJ*, 608, 580
 Bogdan, T. J. 1997, *AJ*, 477, 475
 Christensen-Dalsgaard, J., Schou, J., & Thompson, M. J. 1990, *MNRAS*, 242, 353
 Duvall, Jr., T. L. 1995, in *Proc. GONG 1994, Helio- and Asteroseismology*, ed. R. K. Ulrich, E. J. Rhodes Jr., & W. Dappen, ASP Conf. Ser., 76, 465
 Duvall, Jr., T. L., Jefferies, S. M., Harvey, J. W., & Pomerantz, M. A. 1993, *Nature*, 362, 430
 Giles, P. 2000, *Time-distance measurements of large-scale flows in the solar convection zone*, Ph.D. Thesis, Stanford University
 Gizon, L., & Birch, A. C. 2002, *AJ*, 571, 966
 Gizon, L., & Birch, A. C. 2004, *AJ*, 614, 472
 Harvey, J. W., Hill, F., Hubbard, R. P., et al. 1996, *Science*, 272, 1284
 Hughes, S. J., Rajaguru, S. P., & Thompson, M. J. 2005, *AJ*, 627, 1040
 Hughes, S. J., Pijpers, F. P., & Thompson, M. J. 2006, in preparation
 Jensen, J. M. 2002, *Helioseismic Time-Distance Inversion*, Ph.D. Thesis, Aarhus University
 Jensen, J. M., & Pijpers, F. P. 2003, *A&A*, 412, 257
 Jensen, J. M., Duvall, Jr., T. L., & Jacobson, B. H. 2003, in *Proc. SOHO 12 / GONG+ 2002: Local and Global Helioseismology: the Present and Future*, ed. H. Sawaya-Lacoste, ESA SP-517, 315
 Kosovichev, A. G., & Duvall, Jr., T. L. 1997, in *Proc. SCORE 96: Solar Convection and Oscillations and their Relationship*, ed F. P. Pijpers, J. Christensen-Dalsgaard, & C. S. Rosenthal (Kluwer Academic Publishers), *Astrophysics and Space Science Library*, 225, 241
 Kosovichev, A. G., Duvall, Jr., T. L., & Scherrer, P. H. 2000, *Sol. Phys.*, 192, 159
 Kosovichev, A. G., Duvall, Jr., T. L., Birch, A. C., et al. 2001, in *Helio- and Asteroseismology at the Dawn of the Millennium*, ESA SP-464, 701
 Marquering, H., Dahlen, F. A., & Nolet, G. 1999, *Geophys. J. Int.*, 137, 805
 Paige, C. C., & Saunders, M. A. 1982, *ACM Trans. Math. Soft.*, 8, 43
 Pijpers, F. P., & Thompson, M. J. 1992, *A&A*, 262, L33
 Pijpers, F. P., & Thompson, M. J. 1994, *A&A*, 281, 231
 Rajaguru, S. P., Hughes, S. J., & Thompson, M. J. 2004, *Sol. Phys.*, 220, 381
 Scherrer, P. H., Bogart, R. S., Bush, R. J., et al. 1995, *Sol. Phys.*, 162, 129
 Zhao, J., & Kosovichev, A. G. 2003, *AJ*, 591, 446