

TIME-DISTANCE HELIOSEISMOLOGY: INVERSION OF NOISY CORRELATED DATA

S. COUVIDAT,¹ L. GIZON,¹ A. C. BIRCH,² R. M. LARSEN,¹ AND A. G. KOSOVICHEV¹

Received 2004 December 17; accepted 2005 March 20

ABSTRACT

In time-distance helioseismology most inversion procedures ignore the correlations in the data errors. Here we simulate the travel-time perturbations of wavepackets that result from known distributions of sound speed inhomogeneities. The forward and inverse problems are carried out using recently developed Born approximation sensitivity kernels. A realistic solar noise component, with the correct statistics, is added to the data. We then apply a three-dimensional inversion procedure based on an improved multichannel deconvolution algorithm that includes the full covariance matrix of the simulated data and constrains the solution both in the vertical and horizontal directions. The validation of the inversion is achieved through comparison of the inferred sound speed distributions with the exact solutions. We show that including the covariance matrix matters for sound speed inhomogeneities varying on a length scale smaller than the correlation length. We also find that the inversion procedure is improved by adding horizontal regularization.

Subject headings: methods: data analysis — Sun: helioseismology

1. INTRODUCTION

Solar acoustic waves give access to local properties of the plasma below the visible surface of the Sun. Time-distance analysis (Duvall et al. 1993) uses observations of the solar surface line-of-sight velocity to provide measurements of the wave travel-time perturbations caused by buried inhomogeneities. Here we consider sound speed inhomogeneities only. For the sake of simplicity a plane-parallel geometry is assumed, with \mathbf{r} the horizontal coordinate vector and z the vertical coordinate (height). First, the velocity signal is filtered in Fourier space to select different parts of the wave propagation diagram. Then, travel times are obtained from temporal cross-covariances between the velocity signal measured at two different locations on the solar surface ($z = 0$). In a standard averaging scheme (Duvall et al. 1997), a cross-covariance function is computed between the velocity signal at a central point, \mathbf{r} , and the signal averaged over an annulus at a distance Δ from \mathbf{r} . Travel times, averaged over acoustic waves travelling outward from and inward to the center of the annulus, are measured from fits to the cross-covariance function and denoted by $\tau(\mathbf{r}, \Delta)$. In this paper we consider travel-time perturbations, $\delta\tau(\mathbf{r}, \Delta)$, caused by a given distribution of steady sound speed inhomogeneities, $\delta c(\mathbf{r}, z)$. All perturbations are defined with respect to a reference solar model that is invariant by horizontal translation.

The dominant source of noise in time-distance helioseismology is realization noise because of the stochastic nature of the excitation mechanism of solar oscillations (Jensen et al. 2003; Gizon & Birch 2004). In a first approximation, we assume that signal and noise separate:

$$\delta\tau = \overline{\delta\tau} + \tau_n, \quad (1)$$

where $\overline{\delta\tau} = E[\delta\tau]$ is the expectation value of the travel-time perturbation (the “signal”) and τ_n is the random noise component. We ignore noise perturbations arising from changes in solar structure. In particular, noise is presumed to be spatially

homogeneous. Such an assumption is likely to be acceptable for quiet-Sun data only. We note that the variance of the noise scales like $1/T$, where T is the integration time over which the cross-covariances are computed (Gizon & Birch 2004). Thus, $\overline{\delta\tau}$ is the travel-time perturbation that would be measured in the limit $T \rightarrow \infty$.

The linear forward problem is computing $\overline{\delta\tau}$ given small steady perturbations $\delta c(\mathbf{r}, z)$. This problem was solved by Birch et al. (2004) in the single-scattering Born approximation, following the method described by Gizon & Birch (2002). In plane-parallel geometry, the travel-time perturbations can be written as

$$\overline{\delta\tau}(\mathbf{r}, \Delta) = \int \int d\mathbf{r}' \int_{-\infty}^0 dz K(\mathbf{r} - \mathbf{r}', z; \Delta) \frac{\delta c^2}{c^2}(\mathbf{r}', z), \quad (2)$$

where the function K , which gives the linear sensitivity of the travel times to relative perturbations in the squared sound speed, depends on the details of the measurement procedure. We caution the reader that perturbations are large in the sunspots and that $\delta\tau$ is not necessarily linear in $\delta c^2/c^2$.

The random noise τ_n has zero mean and is characterized by the cross-covariance “matrix”:

$$\Lambda_{ij}(\mathbf{r}) = E[\tau_n(\mathbf{r}', \Delta_i)\tau_n(\mathbf{r}' + \mathbf{r}, \Delta_j)]. \quad (3)$$

An expression for the noise covariance can be estimated directly from real travel-time data (Jensen et al. 2003) or calculated from a model (Gizon & Birch 2004). The model of Gizon & Birch (2004), which is used in this paper, assumes that solar oscillations are stationary and homogeneous on the solar surface. The cross-covariance matrix Λ_{ij} depends implicitly on observation time, T .

Our goal is to solve equation (2) for $\delta c^2/c^2$ inside the Sun, given a set of travel-time measurements, $\delta\tau$. Such inverse problems were solved without taking into account the noise covariance by, e.g., Jensen et al. (1998), Giles (2000), and Hughes et al. (2004). We use a technique developed by Jensen et al. (1998) and named multichannel deconvolution (MCD). In general, the solution of the inverse problem depends on the noise properties specified by equation (3). The importance of noise correlations for two-dimensional linear inversions of global helioseismic data

¹ W. W. Hansen Experimental Physics Laboratory, Stanford University, Stanford, CA 94305-4085.

² Colorado Research Associates, NorthWest Research Associates, Inc., 3380 Mitchell Lane, Boulder, CO 80301.

TABLE 1
ANNULI AND PHASE SPEED FILTER PARAMETERS

Index	Mean Δ (Mm)	Δ (Mm)	v (km s ⁻¹)	δv (km s ⁻¹)	t_0 (minutes)
1.....	6.20	03.7, 04.95, 06.20, 07.45, 08.7	12.77	2.63	16.1, 17.8, 18.9, 20.2, 21.4
2.....	8.70	06.2, 07.45, 08.70, 09.95, 11.2	14.87	2.63	18.9, 20.2, 21.4, 23.0, 24.1
3.....	11.60	08.7, 10.15, 11.60, 13.05, 14.5	17.49	2.63	21.4, 23.2, 24.4, 25.6, 27.0
4.....	16.95	14.5, 15.72, 16.95, 18.17, 19.4	25.82	3.86	27.0, 27.9, 28.7, 29.6, 30.6
5.....	24.35	19.4, 21.87, 24.35, 26.82, 29.3	35.46	5.25	30.6, 32.2, 33.5, 34.6, 35.7
6.....	30.55	26.0, 28.27, 30.55, 32.82, 35.1	39.71	3.05	34.2, 35.2, 36.3, 37.2, 38.1
7.....	36.75	31.8, 34.27, 36.75, 39.22, 41.7	43.29	3.15	36.8, 37.8, 38.7, 39.5, 40.4
8.....	42.95	38.4, 40.67, 42.95, 45.22, 47.5	47.67	3.57	39.3, 40.0, 40.8, 41.6, 42.3
9.....	49.15	44.2, 46.67, 49.15, 51.62, 54.1	52.26	4.46	41.3, 42.0, 42.8, 43.6, 44.3
10.....	55.35	50.8, 53.07, 55.35, 57.62, 59.9	57.16	3.78	43.3, 44.0, 44.7, 45.4, 46.1
11.....	61.65	56.6, 59.12, 61.65, 64.18, 66.7	61.13	3.41	45.1, 45.9, 46.6, 47.3, 48.1

NOTES.—Eleven filters of mean phase speed v and dispersion δv are used for different ranges of annulus radii Δ . The first column gives the annulus index, the last column gives the center of the window function $f(t)$ used to measure first-bounce travel times (see text). We caution the reader that these windows are not perfectly centered on the first-bounce ridge.

has been stressed by Gough & Sekii (2002). The effect of noise in three-dimensional inversions of time-distance was first investigated by Jensen et al. (2003). They studied the propagation of noise through the inversions. However, they did not consider the correlations between travel times with different Δ -values, nor did they explicitly include the noise covariance matrix in their inversion algorithm. Here we improve the original inversion code of Couvidat et al. (2004) by including the full noise covariance matrix and by adding horizontal regularization to the MCD algorithm.

This paper is organized as follows. In § 2 we generate realizations of filtered power spectra of solar oscillations for a horizontally homogeneous solar model. From these data, we compute noise travel times and derive the noise cross-covariance matrix. In § 3 we choose a distribution of sound speed inhomogeneities from which signal travel-time perturbations are computed. In § 4 we present a modified MCD algorithm for solving the inverse problem; this algorithm includes the full noise covariance matrix obtained in § 2.3. The results of the inversion are discussed in § 5. We show that taking into account the correlation in the data errors matters when sound speed perturbations vary on a scale that is smaller than the correlation length of the noise. We conclude in § 6.

2. NOISE PROPERTIES

2.1. Simulation of Artificial Data

In order to compute the noise covariance, we first produce many realizations of travel-time maps $\delta\tau(\mathbf{r}, \Delta)$ for various distances Δ . We generate artificial data cubes of simulated Doppler velocities $\phi(\mathbf{r}, t)$ for a reference solar model following the procedure of Gizon & Birch (2004). This procedure consists in simulating the observable in the discrete Fourier domain:

$$\phi(\mathbf{k}, \omega) = \sqrt{P(\mathbf{k}, \omega)} \mathcal{N}(\mathbf{k}, \omega), \quad (4)$$

where \mathbf{k} is the horizontal wavevector and ω is the angular frequency. Here $P = E[|\phi|^2]$ is the expectation value of the solar oscillation power spectrum and \mathcal{N} is a complex Gaussian random variable with independent real and imaginary parts, zero mean, and unit standard deviation. The random variables $\mathcal{N}(\mathbf{k}, \omega)$ are independent at different grid points in the \mathbf{k} - ω domain, except for the fact that $\mathcal{N}(\mathbf{k}, \omega) = \mathcal{N}^*(-\mathbf{k}, -\omega)$ since $\phi(\mathbf{r}, t)$ has no imaginary part. We take a spatial sampling of $h_x = 1.652$ Mm, corresponding to four high-resolution pixels of the Michelson-

Doppler Imager instrument (MDI; Scherrer et al. 1995). The temporal sampling is $h_t = 1$ minute. Each data cube $\phi(\mathbf{r}, t)$ contains $220 \times 220 \times 512$ grid nodes, so that the computation domain is $L = 362$ Mm on the side and the time duration is $T = 512$ minutes. The sampling in Fourier space is $h_k = 2\pi/L = 12.1/R_\odot$ and $h_\omega = 2\pi/T = 0.20$ mrad s⁻¹. For $P(\mathbf{k}, \omega)$ we use the model power spectrum of solar oscillations introduced by Birch et al. (2004). This theoretical power spectrum was obtained for a plane-parallel background solar model, which is invariant by translation and isotropic, i.e., P only depends on $k = \|\mathbf{k}\|$ and ω . It spans the following ranges: $77 < kR_\odot < 2074$, and $1.8 \text{ mHz} < \omega/2\pi < 6.5 \text{ mHz}$. This model is a fair representation of observed average quiet-Sun power spectra. For more detail, we refer the reader to the paper of Birch et al. (2004).

In time-distance analysis, data cubes are filtered in Fourier space before cross-covariance functions and travel-time maps are computed. This filtering uses Gaussian phase speed filters (Duvall et al. 1997). The rationale for this choice is that waves with similar horizontal phase speeds follow approximately the same path as they travel inside the Sun. For any given target distance, Δ , corresponds a phase speed filter, F . The filtered signal, Φ , is obtained by multiplication in Fourier space:

$$\Phi(\mathbf{k}, \omega) = F(k, \omega; \Delta) \phi(\mathbf{k}, \omega), \quad (5)$$

where F is a phase speed filter of the form

$$F(k, \omega; \Delta) = \exp[-(\omega/k - v)^2/2\delta v^2]. \quad (6)$$

The mean phase speed, v , and the filter width, δv , both depend on Δ . Here we use 11 filters provided by T. L. Duvall (2003, private communication) and listed in Table 1. While v can be derived from the solar model, there is no obvious rule for choosing the width δv . Rather, δv was chosen empirically to yield a good signal-to-noise ratio for the travel-time measurements.

2.2. Measurements of Travel Times

For each 512 minute data cube, we compute maps of point-to-annulus temporal cross-covariance functions, $C(\mathbf{r}, \Delta, t)$. We derive these cross-covariances for the 55 distances Δ listed in Table 1, corresponding to 55 one-pixel wide annuli. These cross-covariances are computed from 11 filtered data cubes: a particular filter is applied to five consecutive distances. Eventually, the 55 travel-time maps are averaged over five consecutive distances.

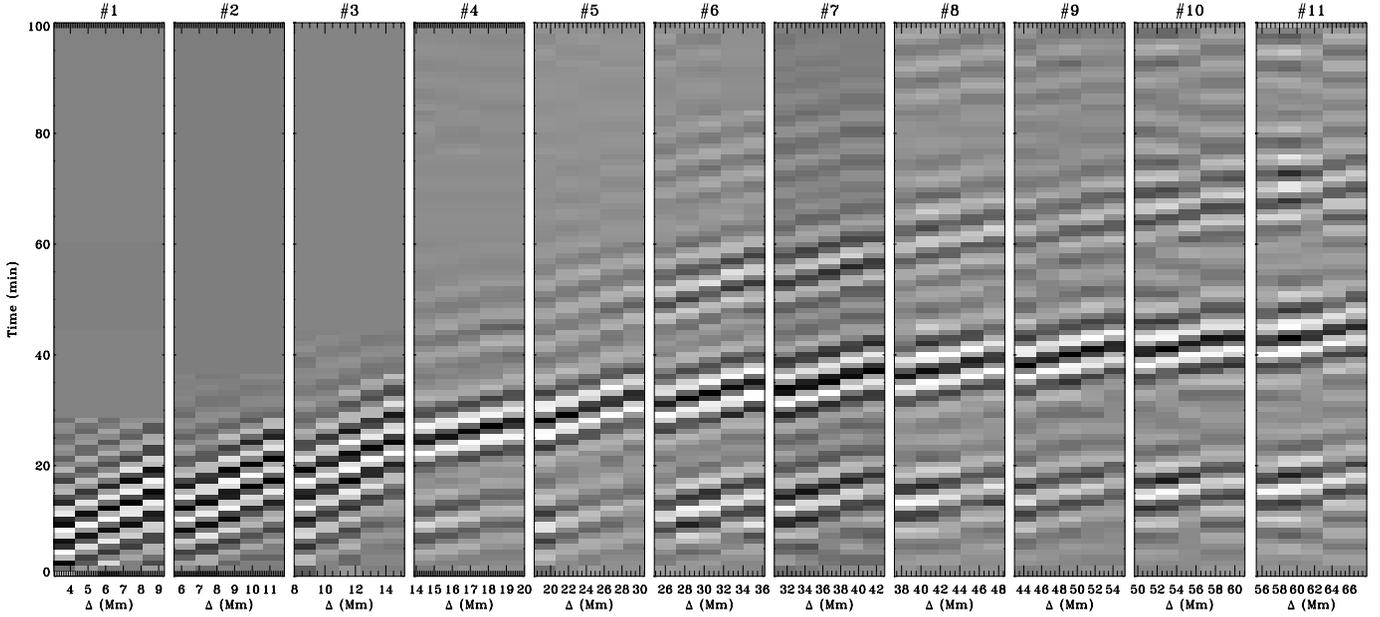


FIG. 1.—Average cross-covariance function $\bar{C}(\Delta, t)$ obtained for the 55 distances Δ and 11 phase speed filters listed in Table 1. Shown are the 11 groups of five consecutive distances. The first-bounce travel times are measured from the most prominent ridge. The ridge at small times is an artifact caused by the phase speed filters.

The temporal cross-covariance functions are computed according to the procedure of Gizon & Birch (2004). At fixed target distance Δ , we first compute maps of point-to-point cross-covariance functions,

$$C(\mathbf{r}|\mathbf{r}', t) = \frac{h_t}{T - |t|} \sum_i \Phi(\mathbf{r}, t_i) \Phi(\mathbf{r}', t_i + t), \quad (7)$$

where Φ is the oscillation signal filtered with the appropriate function $F(k, \omega; \Delta)$. Point-to-point cross-correlations are then averaged over an annulus:

$$C(\mathbf{r}, \Delta, t) = \frac{\sum_i A(\|\mathbf{r}_i - \mathbf{r}\| - \Delta) C(\mathbf{r}|\mathbf{r}_i, t)}{\sum_j A(\|\mathbf{r}_j - \mathbf{r}\| - \Delta)}. \quad (8)$$

The spatial weight function A , defined by

$$A(\mathbf{r} - \Delta) = e^{-(\mathbf{r} - \Delta)^2 / 2\sigma^2}, \quad (9)$$

selects an annulus of radius Δ and thickness $\sigma = h_x/2$. A mean cross-covariance function, $\bar{C}(\Delta, t)$, is obtained by averaging $C(\mathbf{r}, \Delta, t)$ over \mathbf{r} and over all realizations. It is an estimate of the expectation value of the cross-covariance function for our background solar model:

$$\bar{C}(\Delta, t) = E[C(\mathbf{r}, \Delta, t)]. \quad (10)$$

Figure 1 shows \bar{C} as a function of Δ and $t > 0$ (it is symmetric in time by construction). The function \bar{C} is used as a reference to measure individual travel-time perturbations.

We use the definition provided by Gizon & Birch (2004) to measure these travel-time perturbations, averaged over incoming and outgoing waves:

$$\delta\tau(\mathbf{r}, \Delta) = \sum_i W(\Delta, t_i) [C(\mathbf{r}, \Delta, t_i) - \bar{C}(\Delta, t_i)]. \quad (11)$$

The weight function W is given by

$$W(\Delta, t) = -\frac{[f(t) - f(-t)] \partial_t \bar{C}(\Delta, t)}{2 \sum_i f(t_i) [\partial_t \bar{C}(\Delta, t_i)]^2}, \quad (12)$$

where $\partial_t \bar{C}$ is the partial derivative of \bar{C} with respect to time. In the above definition, the function $f(\pm t)$ is a window function that selects a time interval around the first-bounce arrival time of the wavepacket. The functions we use are rectangular windows such that $f(t) = 1$ for $|t - t_0| \leq 10$ minutes and $f(t) = 0$ otherwise. The central times t_0 are listed in Table 1 as a function of annulus radius, Δ . The travel-time maps obtained by this procedure provide us with an estimate of realization noise on travel-time perturbation measurements, i.e., the term $\tau_n(\Delta, t)$ in equation (1).

2.3. Noise Covariance Matrix

The noise covariance matrix $\Lambda_{ij}(\mathbf{r})$ (eq. [3]) is obtained from spatial/ensemble averaging over 20 realizations of travel-time maps. Purely horizontal correlations correspond to the case $\Lambda_{ii}(\mathbf{r})$, while purely “vertical” correlations correspond to $\Lambda_{ij}(\mathbf{r} = 0)$. The term “vertical” is used because the larger the travel distance, the deeper the sensitivity to sound speed perturbations.

At fixed distance Δ_i , horizontal correlations are equivalently specified by the power spectrum of the travel times $\Lambda_{ij}(\mathbf{k}) = E[|\tau_n(\mathbf{k}, \Delta_i)|^2]$ (see Gizon & Birch 2004). Further averaging is performed over the direction of \mathbf{k} , since travel-time maps are isotropic. Figure 2 shows the spatial power spectra of τ_n as a function of k for three particular values of Δ_i . Similar power spectra were previously measured by Jensen et al. (2003) and Gizon & Birch (2004). In particular, we observe a sudden drop in power beyond a characteristic wavenumber, k_c . This cutoff wavenumber decreases with Δ and is related to the dominant wavenumber of the wave packets, \bar{k} , according to $k_c \sim 2\bar{k}$. The horizontal correlation length is at most 6 Mm for the largest annulus. For the small annuli there is almost no horizontal correlations (correlation length less than $h_x = 1.652$ Mm).

The vertical noise covariance is given by the 11×11 matrix $\Lambda_{ij}(\mathbf{r} = 0)$. Figure 3 shows these vertical correlations averaged

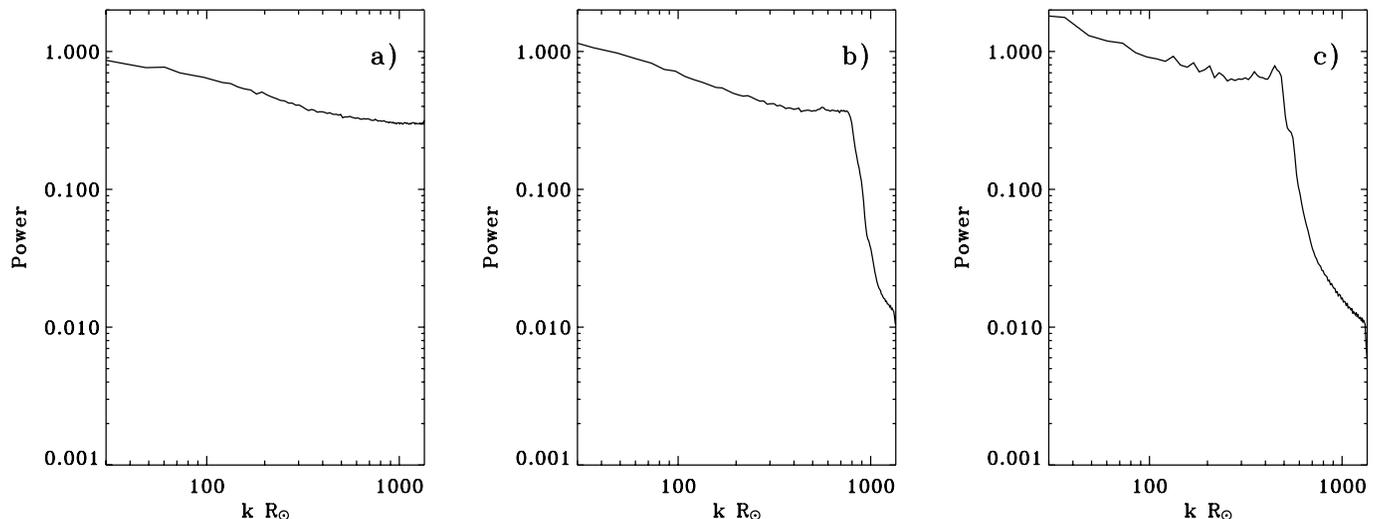


FIG. 2.—Spatial power spectra of travel time maps, $|\tau_n(k, \Delta)|^2$, for different distances Δ . These plots are averaged over 20 noise realizations. The Nyquist wavenumber is at $kR_\odot = 1323$. (a) $\Delta = 6.2$ Mm, (b) $\Delta = 30.55$ Mm, and (c) $\Delta = 61.65$ Mm.

over all the simulated data, for any set (Δ_i, Δ_j) . We see that the correlation falls off quite fast away from the diagonal $i = j$. Travel times for radii Δ_i and Δ_j are weakly correlated for $|i - j| \geq 2$. The vertical correlation is controlled by the overlap in phase speed between the different wavepackets that contribute to the travel times. In particular, the correlation cannot be very important when the phase speed filters do not overlap significantly (see Gizon & Birch 2004 for details). To illustrate this point, we have computed the vertical correlations in the case when no phase speed filtering is applied. Figure 4 shows the vertical noise correlation matrix without filtering, which displays larger vertical off-diagonal correlations (in absolute value). When no phase speed filter is applied, there is almost no impact on the horizontal correlations except for $\Delta \lesssim 20$ Mm: the correlation length is increased but remains below 10 Mm.

3. SOUND SPEED PERTURBATIONS

We wish to produce realizations of travel-time maps, $\delta\tau(\mathbf{r}, \Delta)$, when the sound speed is not uniform inside the Sun. To do this,

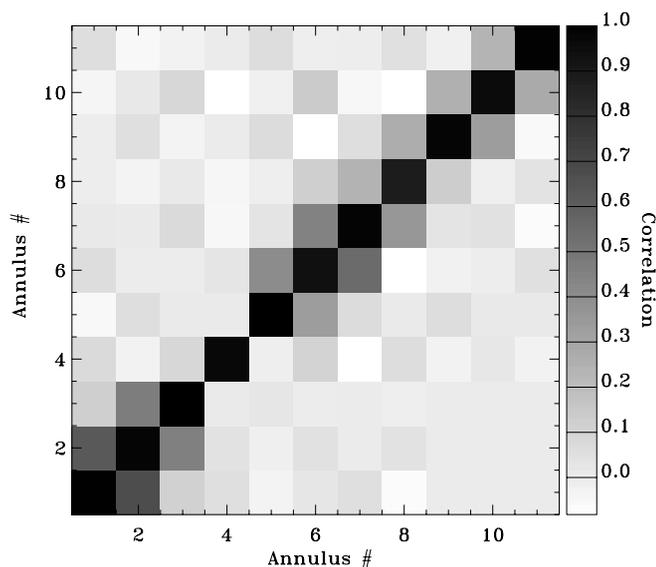


FIG. 3.—Vertical correlation matrix for the phase speed filters listed in Table 1. Each pixel gives the correlation in the travel-time noise for any two annulus radii Δ_i and Δ_j .

we choose a three-dimensional sound speed distribution, $\delta c^2/c^2$. A vertical cut through this distribution is shown in Figure 5. These perturbations were chosen because they correspond (remotely) to sunspot-like structures, with the difference that the sound speed perturbation amplitude is sometimes reversed compared to real sunspots. Using equation (2) we obtain 11 maps of travel-time perturbations, $\delta\tau(\mathbf{r}, \Delta)$, one for each annulus radius. We then add the noise realizations $\tau_n(\mathbf{r}, \Delta)$ computed above for $T = 512$ minutes to obtain realizations of $\delta\tau(\mathbf{r}, \Delta)$. The S/N ratio on these maps, defined here as the maximum value of $|\delta\tau|$ divided by the noise dispersion is close to what is actually measured from MDI data. In this paper, this S/N ratio varies between 5 and 10, depending on the value of Δ .

The linear forward problem is solved using the Born-approximation sensitivity kernels computed by Birch et al. (2004). The Born kernels take into account the sensitivity of wavepackets to the sound speed perturbations off the ray paths. They depend on the phase speed filters and travel-time definition used to derive them. The set of kernels we work with were computed with

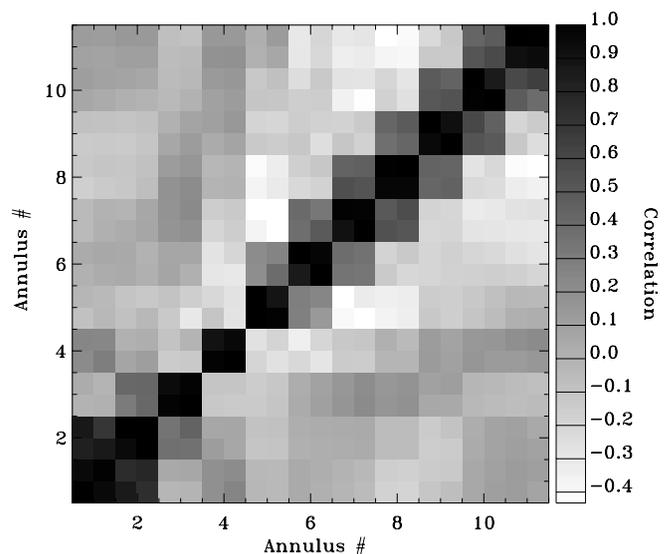


FIG. 4.—Vertical correlation matrix obtained when no filtering is applied to the velocity signal (except exclusion of f modes).

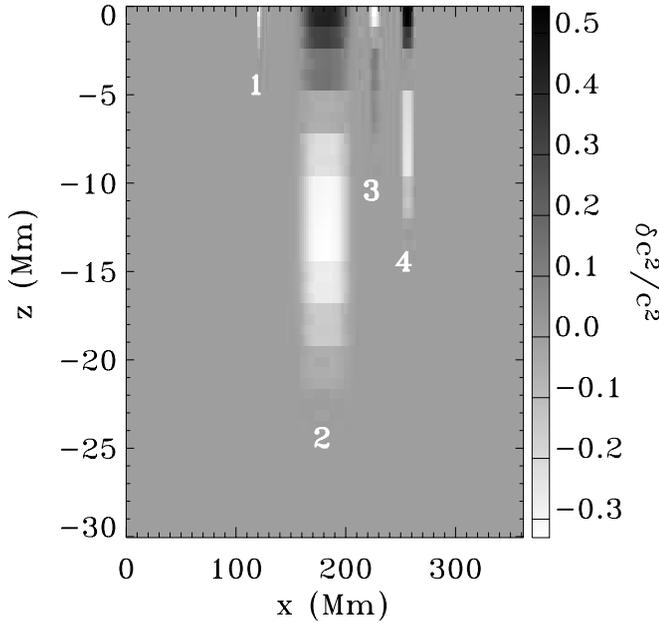


FIG. 5.—Vertical cut through the input sound speed perturbations, $\delta c^2/c^2$. Each structure is axisymmetrical: the axis is vertical and is in the plane of the figure. The numbers below the structures are used as labels in Figs. 7 and 9.

the same Gaussian filters and the same travel-time definition defined above. This self consistency in the analysis is important and was often neglected in previous studies (e.g., Couvidat et al. 2004).

4. INVERSION PROCEDURE

4.1. MCD Algorithm with Horizontal Regularization

The inversion of the travel times is done with a regularized least-squares method through a modified MCD algorithm. The MCD is based on the horizontal translational invariance of the sensitivity kernels that makes equation (2) a convolution product in the horizontal plane. In the Fourier domain, this convolution product becomes a multiplication. It is advantageous to invert the travel times in the \mathbf{k} plane. Instead of solving a huge three-dimensional inverse problem, we solve many one-dimensional problems in the vertical (z) direction: one inversion for each \mathbf{k} . Because the transformation to the Fourier domain can take advantage of the fast Fourier transform, the MCD algorithm is typically several orders of magnitude faster than a conventional regularized least-squares approach. Our inversion code also returns the averaging kernels, which give an estimate of the effective horizontal and vertical resolutions reached in the inversion results (Hansen 1997).

In order to formulate the problem in Fourier space we introduce the following vectors (Jensen 2001):

$$d_i = \delta\tau(\mathbf{k}, \Delta_i), \quad (13)$$

$$n_i = \tau_n(\mathbf{k}, \Delta_i), \quad (14)$$

$$G_{i\alpha} = K(\mathbf{k}, z_\alpha; \Delta_i), \quad (15)$$

$$m_\alpha = \frac{\delta c^2}{c^2}(\mathbf{k}, z_\alpha), \quad (16)$$

where all functions of \mathbf{k} are obtained by two-dimensional spatial Fourier transformation. For each \mathbf{k} we want to minimize the misfit to the data and solve (in matrix notation)

$$\min\{(\mathbf{d} - \mathbf{G}\mathbf{m})^H \Lambda^{-1}(\mathbf{d} - \mathbf{G}\mathbf{m})\}, \quad (17)$$

where

$$\Lambda \equiv E[\mathbf{nn}^H] \quad (18)$$

is the noise covariance matrix. For computational expediency we rewrite equation (17) as an equivalent linear least-squares problem:

$$\min\{\|\Sigma^{-1}(\mathbf{d} - \mathbf{G}\mathbf{m})\|_2^2\}, \quad (19)$$

where

$$\Lambda = \Sigma\Sigma^T \quad (20)$$

is the Cholesky decomposition (Golub & Van Loan 1996) of Λ with Σ being a lower triangular matrix, and $\|\dots\|_2$ is the 2-norm. Because the data errors can be assumed to be Gaussian, the solution \mathbf{d} to equation (19) is the best unbiased linear estimate, for the component with wavenumber \mathbf{k} , of the solution to equation (2). In practice, however, the matrix \mathbf{G} is highly ill-conditioned and to obtain a well-behaved solution we need to add a regularization term to equation (19):

$$\min\{\|\Sigma^{-1}(\mathbf{d} - \mathbf{G}\mathbf{m})\|_2^2 + \lambda^2(k)\|\mathbf{L}\mathbf{m}\|_2^2\}, \quad (21)$$

where \mathbf{L} is a regularization operator, and $\lambda(k)$ is the regularization parameter; together these determine the effect of the regularization and must be chosen carefully to obtain a reliable and effective inversion method. It is natural to choose \mathbf{L} so that $\|\mathbf{L}\mathbf{m}\|_2^2$ is a discrete approximation to $\int (\delta c^2/c^2)^2(\mathbf{k}, z) dz$. Consequently, \mathbf{L} is a diagonal matrix whose elements are the inverse of the square root of the spatial sampling Δz at each depth. Such a weighting is necessary because our grid in z for $\delta c^2/c^2$ is roughly uniform in acoustic depth, which means the spatial sampling of deep layers is larger than the sampling of surface layers. Without this weighting, deep layers would be more constrained than surface ones.

Unlike the standard MCD algorithm, our regularization parameter λ varies with k , and takes the form $\lambda^2(k) = \lambda_v^2 + \lambda_h^2(k)$, where λ_v is a vertical regularization parameter, and λ_h is a horizontal one. In the standard MCD there is no regularization in the horizontal direction: $\lambda_h = 0$, and because the deconvolution is intrinsically ill-conditioned, MCD may give spurious inversion results by emphasizing the high-wavenumber components of the results in the horizontal direction. The result is a large amount of noise with high horizontal wavenumbers in the inverted $\delta c^2/c^2$. To counter this effect we add a second regularization parameter λ_h , and we change a singly constrained regularization algorithm into a multiply constrained one. In light of the problems in the regular MCD algorithm one would suggest that λ_h increases with k , such that high horizontal wavenumbers are regularized more than low wavenumbers. This procedure is commonly used in image restoration applications (Kang 1998). We advocate choosing either $\lambda_h = \lambda_2 k$ or $\lambda_h = \lambda_2 k^2$, with λ_2 constant, which means that our regularization term is proportional to the 2-norm of the first or second horizontal derivative of the solution—a very commonly used regularization operator. Since we invert in the Fourier domain, $k\|\mathbf{L}\mathbf{m}\|_2$ is the norm of the Fourier transform of the first horizontal derivatives of $\delta c^2/c^2(\mathbf{r}, z)$, while $k^2\|\mathbf{L}\mathbf{m}\|_2$ is the norm of the second horizontal derivatives. Owing to Parseval's theorem, the 2-norm of a function is equal to the 2-norm of its Fourier transform. Consequently, with our choice of λ_h the regularization applied in the MCD algorithm is the same as if we had explicitly regularized by the norm of the first or second derivatives of the solution in the horizontal direction. This makes the inversion procedure more complex than with the standard MCD, especially when it comes to choosing the best regularization parameter values, but it gives significantly better inversion results.

When inverting travel times corresponding to a known solution, the improvement due to the horizontal regularization can be quantified by computing the mean squared error (MSE):

$$\text{MSE}(\lambda_v, \lambda_2) = \frac{1}{N} \sum_{\mathbf{r}, z} \left[\frac{\delta c^2}{c^2}(\mathbf{r}, z) - \left(\frac{\delta c^2}{c^2} \right)_{(\lambda_v, \lambda_2)}(\mathbf{r}, z) \right]^2, \quad (22)$$

where N is the total number of grid nodes, $\delta c^2/c^2_{(\lambda_v, \lambda_2)}(\mathbf{r}, z)$ is the regularized inverse solution, and $\delta c^2/c^2(\mathbf{r}, z)$ is the true solution. For our specific $\delta c^2/c^2$ perturbation—sunspot-like structures—the minimum MSE calculated in a small region around the “sunspots” location is reduced by 30% when we add horizontal regularization (see Fig. 6). The corresponding averaging kernels are also slightly better localized in the vertical direction. As expected, in the horizontal direction these kernels are slightly broader, which explains the smoother aspect of the solution. The spurious high-frequency features have been reduced significantly.

In the following, all the inversion results we show are obtained by regularizing vertically with the norm of $\delta c^2/c^2$ and horizontally with the norm of the second derivatives of $\delta c^2/c^2$. To speed up the inversion procedure we do not invert the travel-time maps for $kR_\odot \geq 665$, which is the cutoff wavenumber on the power spectrum of $\tau_n(\mathbf{r}, \Delta)$ for the largest Δ . We assume that the inversion result is zero at higher wavenumbers. We invert for 13 layers in depth for $\delta c^2/c^2$: the center of the deepest layer is located 28.5 Mm below the solar surface.

4.2. Choosing the Regularization Parameters

In the test cases we know the exact inverse solution $\delta c^2/c^2(\mathbf{r}, z)$. Therefore, we use a least-squares (LS) estimator to get the “best” regularization parameters. The LS estimator merely returns the (λ_v, λ_2) pair corresponding to the minimum of the MSE. Larger values of the regularization parameters give larger bias and smaller variance of the estimated inversion result. The inversion is always biased because we minimize the norm of the solution in the vertical direction, meaning that the solution is biased toward zero. The LS estimator constraints a minimum variance more than a minimum bias. Thus, it returns as best (λ_v, λ_2) pair one that systematically overregularizes the solution: the “best” regularization parameters found are too large, and the inverted $\delta c^2/c^2$ amplitudes are too small compared to the exact solution. The signal we try to invert is located in a small region of the $\delta c^2/c^2(\mathbf{r}, z)$ slab, the main part of this slab is noise only. The LS estimator systematically overregularizes because it does not discriminate between useful signal and useless noise. It tries to find the best inverse solution over the entire slab, while a human being knowing where the signal is located focuses on this region and prefers a better solution there at the expense of the rest of the slab.

To test the impact of the way we pick (λ_v, λ_2) on the inversion results, we invert our 20 realizations of travel-time maps $\delta\tau(r, \Delta)$, with $\overline{\delta\tau}(\mathbf{r}, \Delta)$ based on the sunspot-like perturbations. We invert these maps using either (λ_v, λ_2) returned by the LS estimator over the entire $\delta c^2/c^2$ slab, or (λ_v, λ_2) returned by the LS estimator applied to a small region around the sunspot-like structures. We then average the 20 inversion results obtained in each case. Figure 7 shows that applying the LS criterion to the entire slab produces an unsatisfactory result. Applying it to a small region surrounding the structures we want to retrieve gives better results (Figs. 8 and 9). In this paper the latter method is the preferred one to pick the best regularization parameters.

When working with the actual time-distance data obtained from MDI, λ_v and λ_2 may be chosen through the so-called L -surface criterion (e.g., Hansen 1997; Belge et al. 2002). The L -surface is a

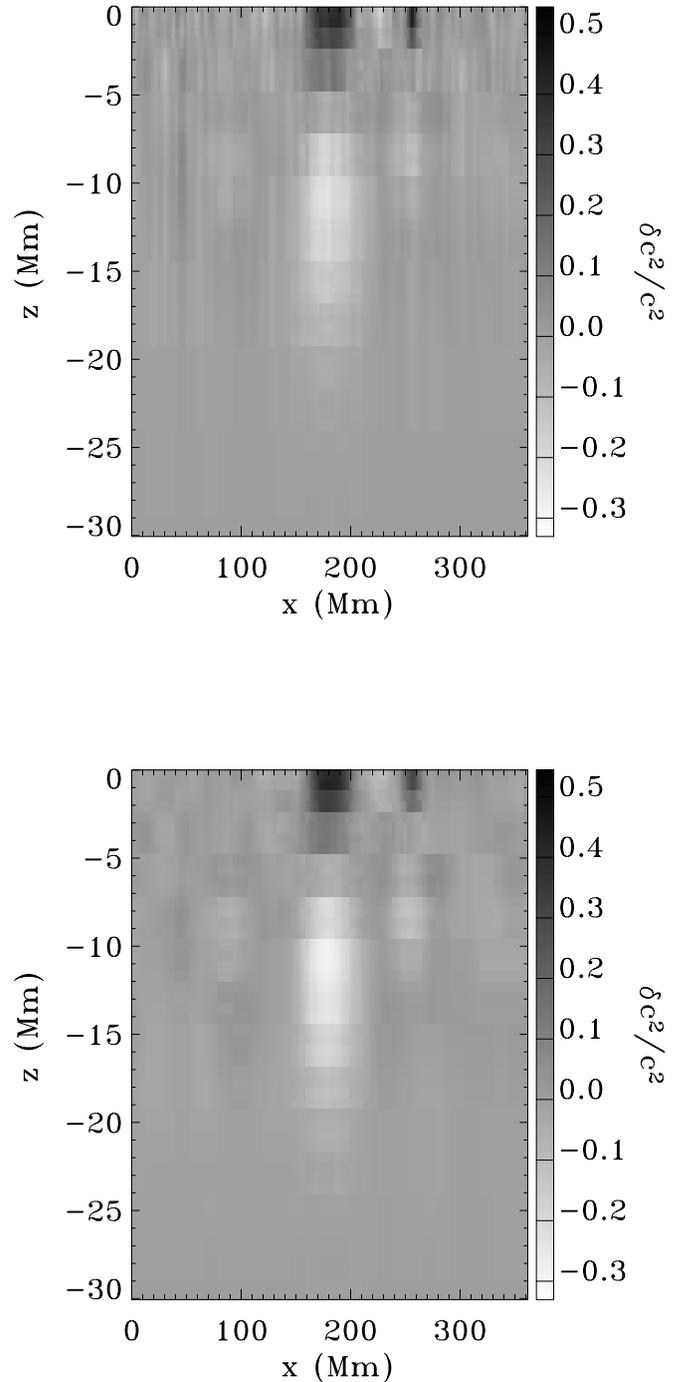


FIG. 6.—Vertical cuts in the inverted $\delta c^2/c^2$ slabs. *Upper panel*: inversion result for a specific noise realization and with no horizontal regularization (standard MCD). *Lower panel*: inversion with horizontal regularization (modified MCD). The minimum MSE has been reduced by 30% by adding horizontal regularization.

plot of the residual $\|d - Gm\|_2$ as a function of the norms $\|Lm\|_2$ and $k^2\|Lm\|_2$. This criterion is a generalization of the L -curve criterion and consists in choosing as best regularization parameters the (λ_v, λ_2) pair corresponding to the maximum Gaussian curvature of the L -surface. According to Belge et al. (1998), the regularization parameters producing the maximum curvature and the parameters producing the minimum MSE are very close: L -surface criterion and LS estimator are equivalent. Therefore, the L -surface criterion should also overregularize when it is applied to the entire slab, and a solution might be to compute the residual and norms only on the region where the signal is located. Such a

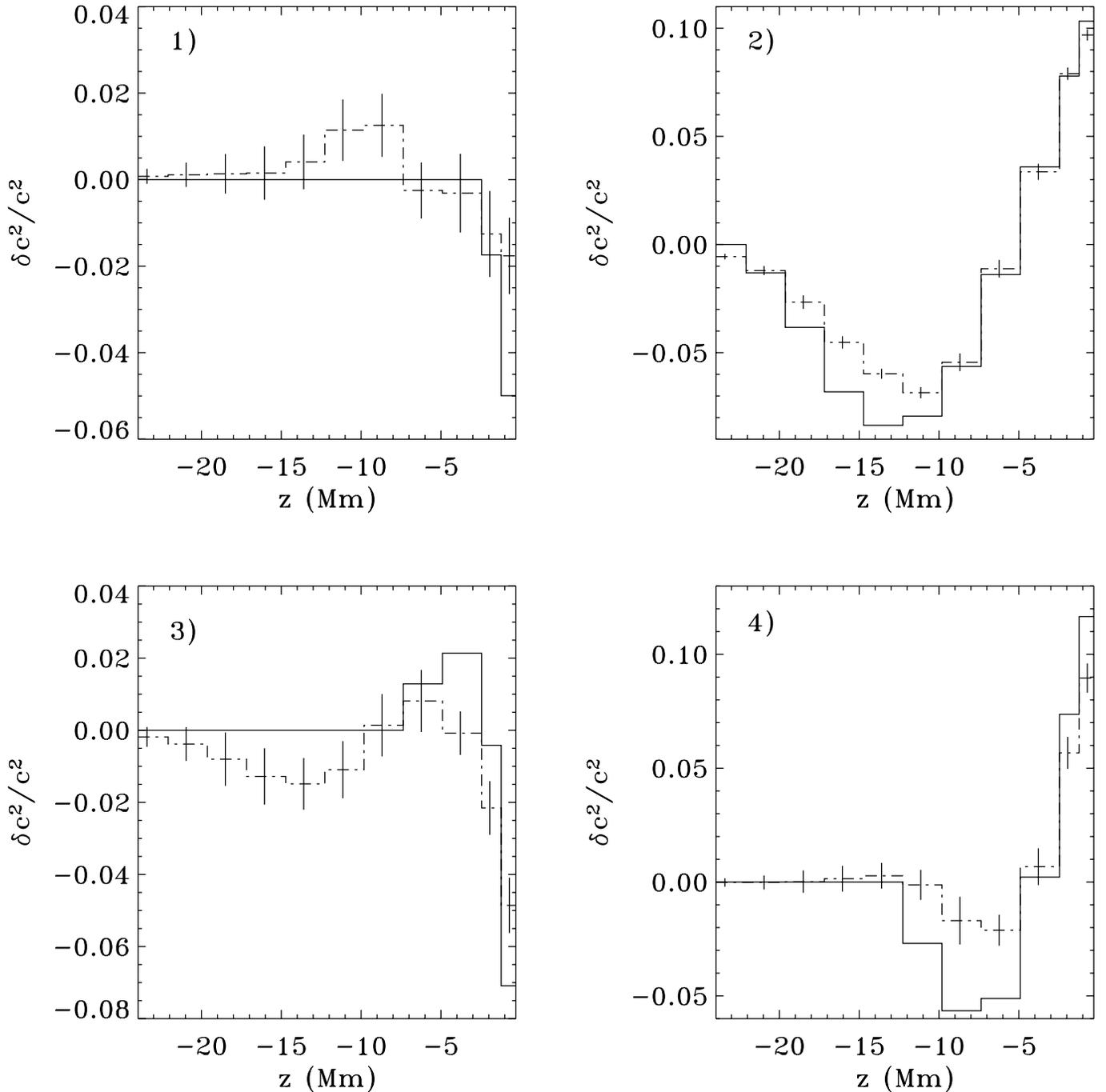


FIG. 7.—Vertical profiles of the four inverted sunspot-like structures shown in Fig. 5. We averaged $\delta c^2/c^2$ over the volumes roughly occupied by the different structures. The regularization parameters used for the inversion are returned by the LS estimator applied to the entire $\delta c^2/c^2$ slab. The solid line is the exact solution, the dot-dashed line is the inversion with noise covariance matrices. We also show the corresponding 1σ vertical error bars. These plots are an average over 20 realizations.

process could be applied to the travel-time maps obtained from MDI: when we invert the sound speed below an active region, it is easy to locate on the maps where the sunspots are located and derive optimal regularization for this area. However, this approach may not resolve structures outside the spot.

5. INVERSION RESULTS

5.1. Inversion of Travel Times Obtained with Sunspot-like Perturbations

To properly compare the inversion results when equation (21) is, or not, scaled by Σ^{-1} , we average the inversion results for

$20 \delta\tau(r, \Delta)$ realizations and for both cases, and we compare these averages. The regularization parameters λ_v and λ_2 are not the same for inversions done with and without Σ^{-1} . In each case and for each of the 20 sets of travel-time maps, we pick the best regularization parameters, and then we average these 20 parameters. We use these mean parameters to invert the $\delta\tau(r, \Delta)$ maps. The averaged inversion results are shown on Figures 8 and 9. The solutions seem too much regularized, and indeed the regularization parameters chosen pragmatically, with no automatic program, give slightly better results. This confirms that when it comes to choosing regularization parameters human beings are sometimes more efficient than numerical algorithms.

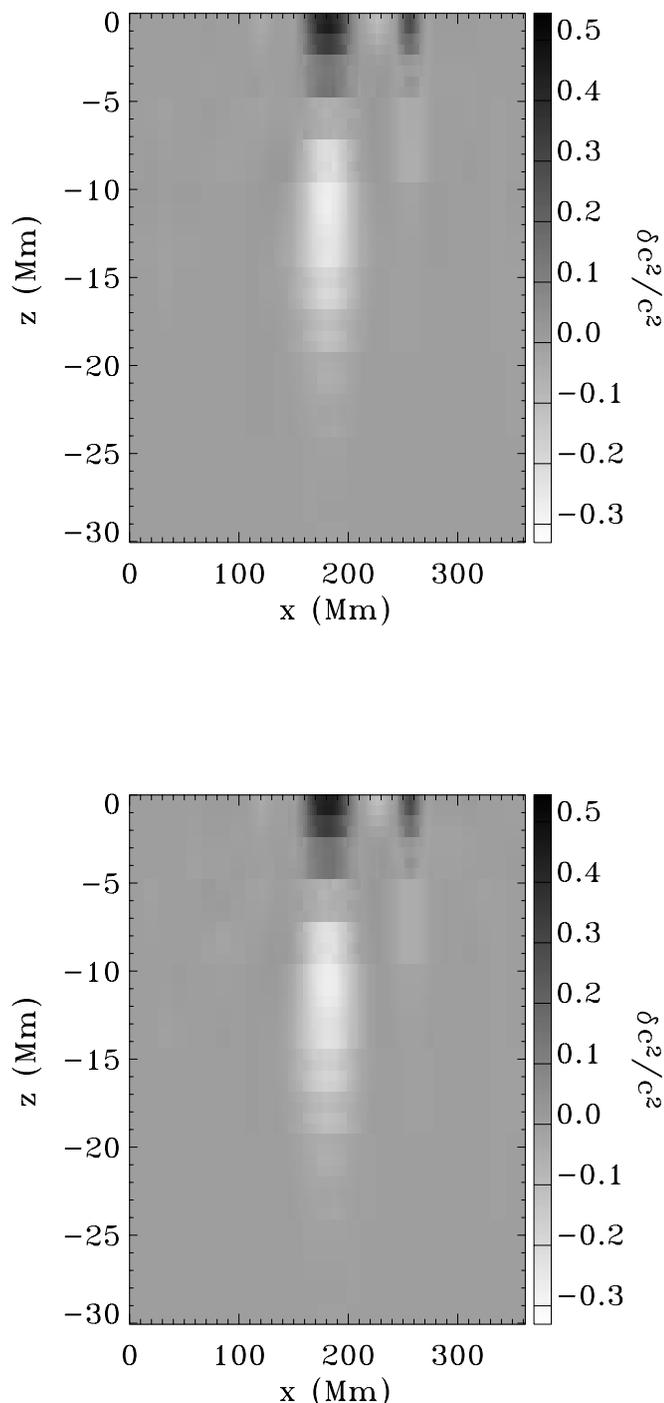


FIG. 8.—Vertical cuts in the inverted $\delta c^2/c^2$. *Upper panel*: inversion results averaged over 20 realizations without including the noise covariance matrices. *Lower panel*: averaged inversion results with the noise covariance matrices.

Although on an inversion of travel-time maps obtained with a specific noise realization we may notice differences between the use or not of noise covariance matrices, on the averaged inversion results any difference is hardly noticeable. That was expected considering the low off-diagonal “vertical” correlations and the very low horizontal correlations that our set of phase speed filters and 11 Δ distances produce. To test the impact on the inversion results of the signal-to-noise ratio of the travel-time maps, we produce noise realizations with a dispersion twice as large as previously, and we invert the new $\delta\tau(\mathbf{r}, \Delta)$:

again, taking into account the noise covariance does not affect the averaged inversion results. The $\delta c^2/c^2$ perturbation we try to retrieve is too smooth in both the horizontal and vertical directions: the numerical code has no difficulty to invert the corresponding maps even without including the Σ^{-1} matrices. $\delta\tau(\mathbf{r}, \Delta)$ is clearly separated from $\tau_n(\mathbf{r}, \Delta)$ in the Fourier space and is only located in the small wavenumbers domain of the spectrum, which makes it easy for the regularization to suppress the noise. The characteristic horizontal sizes of the four “sunspots” are larger than the largest horizontal correlation length on the noise travel-time maps. Actually in this case the use of Σ^{-1} even slightly degrades the inversion results according to the LS criterion: the minimum MSE increases by a few percent. The reason may be twofold. First, there is a lack of accuracy in the covariance matrices determination: these matrices are obtained from 20 noise realizations only and are assumed to be isotropic. However, the use of the fast Fourier transform produces edge effects: horizontal and vertical stripes on the two-dimensional power spectra of travel-time maps make the noise slightly anisotropic. Moreover, we average over k , and the precision on Λ is better for high k than for low k (because there are more points to average): 20 noise realizations may be too small a number to have precise enough Λ at low k . Indeed, we recompute the covariance matrices over 40 travel-time noise realizations, and the minimum MSE of the inversion result obtained with these matrices is closer to the minimum MSE of the inversion results without covariance matrices. However, this minimum MSE is still larger than when no Σ^{-1} is included. A second reason to explain the apparent lack of improvement with Σ^{-1} is the negative slope of the power spectra of the τ_n maps (Fig. 2), which means that the inversion code regularizes less the intermediate and high wavenumbers when the Σ^{-1} matrices are included. For a smooth sound speed perturbation with no signal at intermediate and high k , that means a slightly larger noise variance on the inversion results, explaining the small increase in the minimum MSE. This larger variance makes the uncertainties on the surface layers larger for inversions with covariance matrices. On Figure 9 the error bars are the dispersion of the $\delta c^2/c^2$ results at each layer amongst the 20 inversions. For deepest layers, the uncertainty seems to be reduced when covariance matrices are included, probably because the horizontal correlation length increases with Δ and becomes important enough for the covariance matrices to make a slight difference. The averaging kernels corresponding to these inversions are shown in Figures 10, 11, and 12. On average, the kernels obtained with Σ^{-1} are less oscillatory and have smaller negative sidelobes, but reduced amplitudes. Including the covariance matrices has no positive impact on the vertical and horizontal resolutions of the inversion results, and even degrade them.

With the specific phase speed filters and Δ distances we use, for the signal-to-noise ratio we have on the travel-time maps, and above all for the smooth perturbation $\delta c^2/c^2$ we invert, taking into account the Σ^{-1} matrices in the inversion procedure does not improve the inversion results. The small correlations in the data errors and the smoothness of the sound speed perturbation explain this result.

5.2. Fast-Varying Sound Speed Perturbation/Increased Noise Correlations

To further study the role of correlations in the inversion procedure we use another artificial $\delta c^2/c^2$ perturbation: this one looks like a checkerboard and has sharp and rapidly

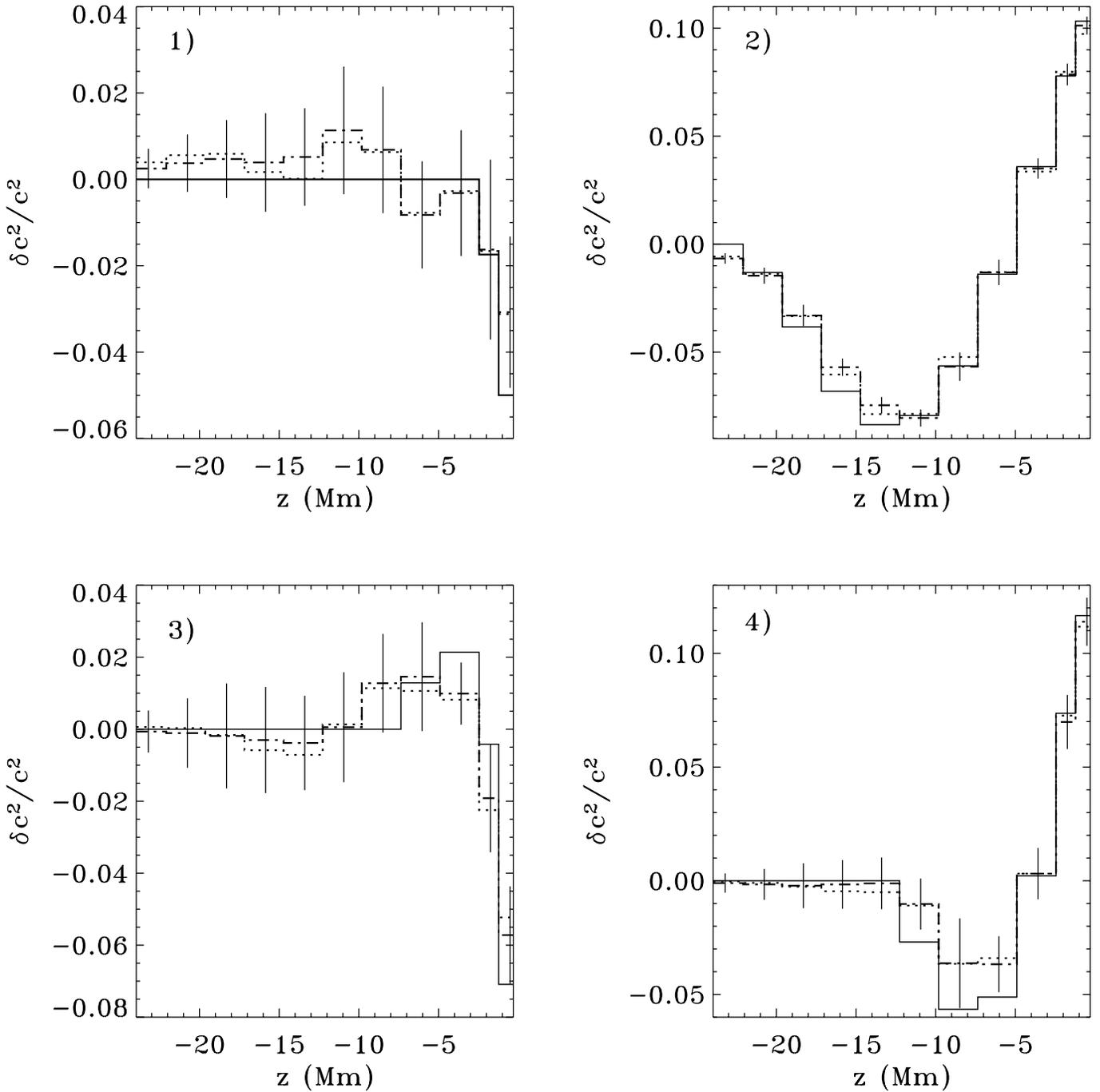


FIG. 9.—Vertical profiles of the four inverted sunspot-like structures shown in Fig. 5. We averaged $\delta c^2/c^2$ over the volumes roughly occupied by the different structures. The regularization parameters used in the inversion are returned by the LS estimator applied to a small region of the $\delta c^2/c^2$ slab, where the signal is located. The solid line is the exact solution. The dot-dashed curve is the inversion result when we include the covariance matrices, the dotted curve is when we invert without the noise covariance matrices. We also show the corresponding 1σ vertical error bars for the case with covariance matrices. These plots are an average over 20 realizations.

changing features, which make it more difficult to invert (Fig. 13). Owing to the small-scale features of the new $\delta c^2/c^2$ perturbation, the power is spread over a larger wavenumber range than with the sunspot-like perturbations. On the corresponding travel-time maps, it is less easy to separate noise from signal.

We invert these maps obtained for the 11 phase speed filters listed in Table 1 and for a specific noise realization (no averaging). The result confirms that the shape of the sound speed perturbation matters: for the checkerboard-like perturbation, taking

into account the noise correlation does improve the result, according to the LS criterion. The minimum MSE is reduced by 9% when we include Σ^{-1} , even though this is hardly visible on the upper panels of Figure 14. The inversion code regularizes relatively too much the intermediate and high wavenumbers when no Σ^{-1} is provided. Therefore, the inverted signal at these wavenumbers is also damped.

To test the case of large off-diagonal vertical correlations, we compute the travel-time maps related to the new $\delta c^2/c^2$ perturbation and we add extra correlations between these different

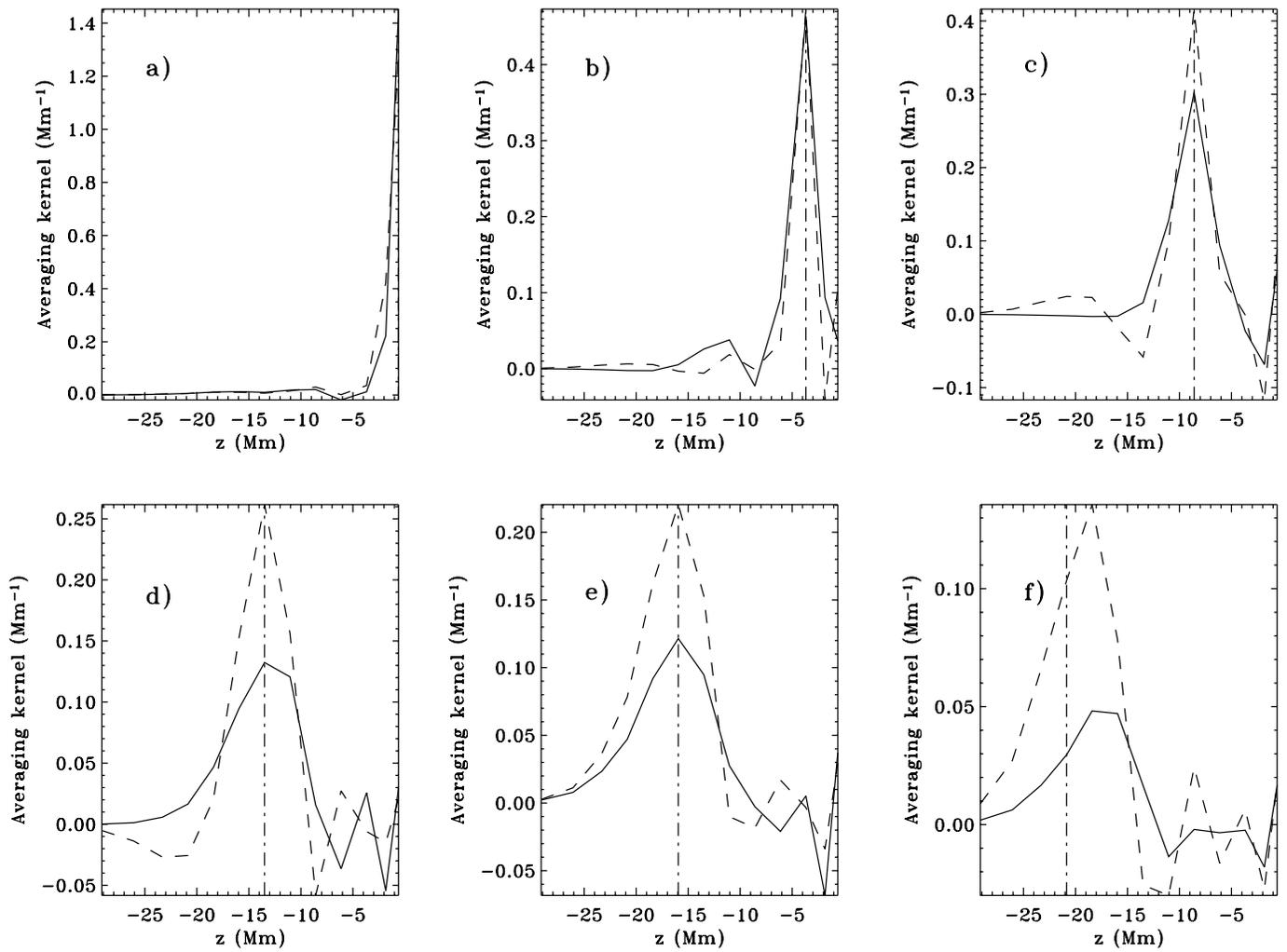


FIG. 10.—Profiles of the averaging kernels corresponding to the inversion results of Figs. 8 and 9. *Solid lines*: inversions done with noise covariance matrices. *Dashed lines*: inversion done without noise covariance matrices. We show the kernels for six target depths z indicated by vertical dot-dashed lines, and for $(x = 0 \text{ Mm}, y = 0 \text{ Mm})$. (a) $z = -0.6 \text{ Mm}$, (b) $z = -3.7 \text{ Mm}$, (c) $z = -8.6 \text{ Mm}$, (d) $z = -13.5 \text{ Mm}$, (e) $z = -15.9 \text{ Mm}$, and (f) $z = -20.8 \text{ Mm}$. These profiles are an average over x and y .

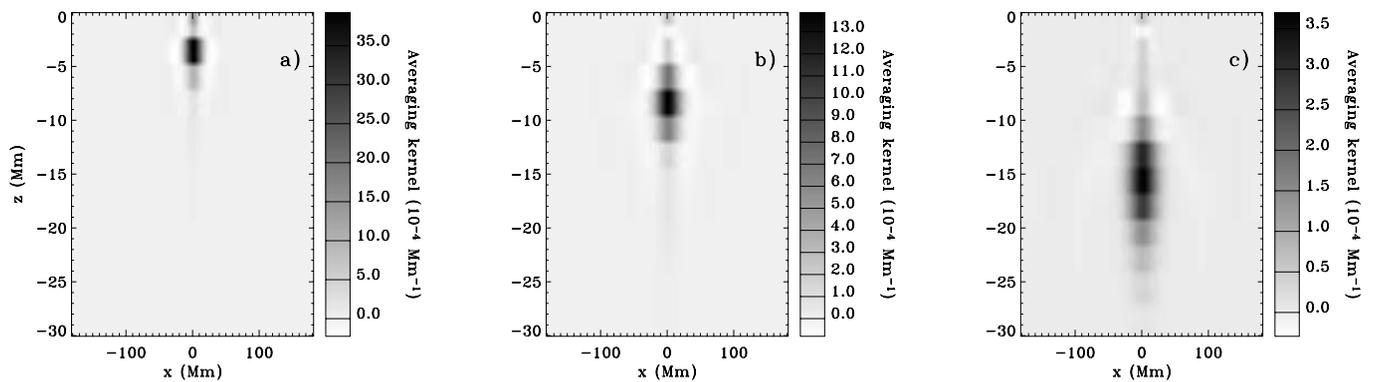


FIG. 11.—Vertical cuts in the averaging kernels for the targets $(x = 0 \text{ Mm}, y = 0 \text{ Mm})$ and (a) $z = -3.7 \text{ Mm}$, (b) $z = -8.6 \text{ Mm}$, and (c) $z = -15.9 \text{ Mm}$. Inversion done with noise covariance matrices.

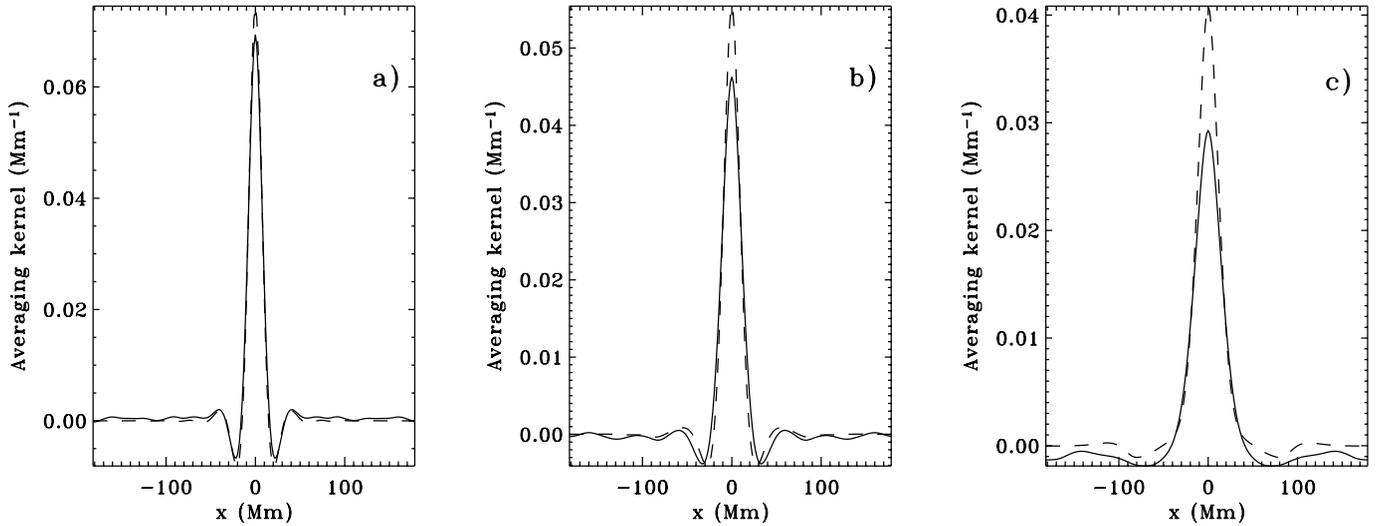


FIG. 12.—Profiles of the averaging kernels for the targets ($x = 0$ Mm, $y = 0$ Mm) and (a) $z = -3.7$ Mm, (b) $z = -8.6$ Mm, and (c) $z = -15.9$ Mm. *Solid lines*: inversion done with noise covariance matrices. *Dashed lines*: inversion done without noise covariance matrices. These profiles are an average over z and y .

maps by combining them through a specific matrix. On Figure 15 we show the corresponding vertical correlation matrix. There are strong off-diagonal correlations and negative values are present instead of only positive values. We combine our Born sensitivity kernels the same way before applying the forward and then inverse problems. The inversion results with and without including the noise covariance matrices are clearly different (Fig. 14, *middle panels*): the use of Σ^{-1} reduces the minimum MSE by 68%. Not surprisingly, this confirms the strong need to include Σ^{-1} in the inversion procedure when travel-time maps are strongly correlated. However, the amount of extra correlation we added here is clearly unrealistic. We also add extra horizontal correlations to the original (no vertical correlation added) travel-time maps: we convolve these maps with $f(r) = 2 \cos(2\pi r/T) \exp(-r^2/\sigma^2)$, where $T = 58$ Mm, and $\sigma = 33$ Mm. We convolve the sensitiv-

ity kernels using this function. The characteristic horizontal correlation length on the travel-time maps is now about 25 Mm, larger than the horizontal length scale of the checkerboard-like structure. As shown on the lower panels of Figure 14, the use of noise covariance matrices again improves the inversion results: the minimum MSE is reduced by 55% and the upper part of the checkerboard-like structure becomes visible, which is not the case when we do not include Σ^{-1} . The near surface layers are especially poorly resolved with no Σ^{-1} owing to the horizontal length scale of the perturbation: this length scale is well below the correlation length. The deeper layers, where the horizontal size of the perturbation is larger, are better retrieved. This example shows the interest of using the noise covariance matrices in the inversion procedure when there are large horizontal correlations.

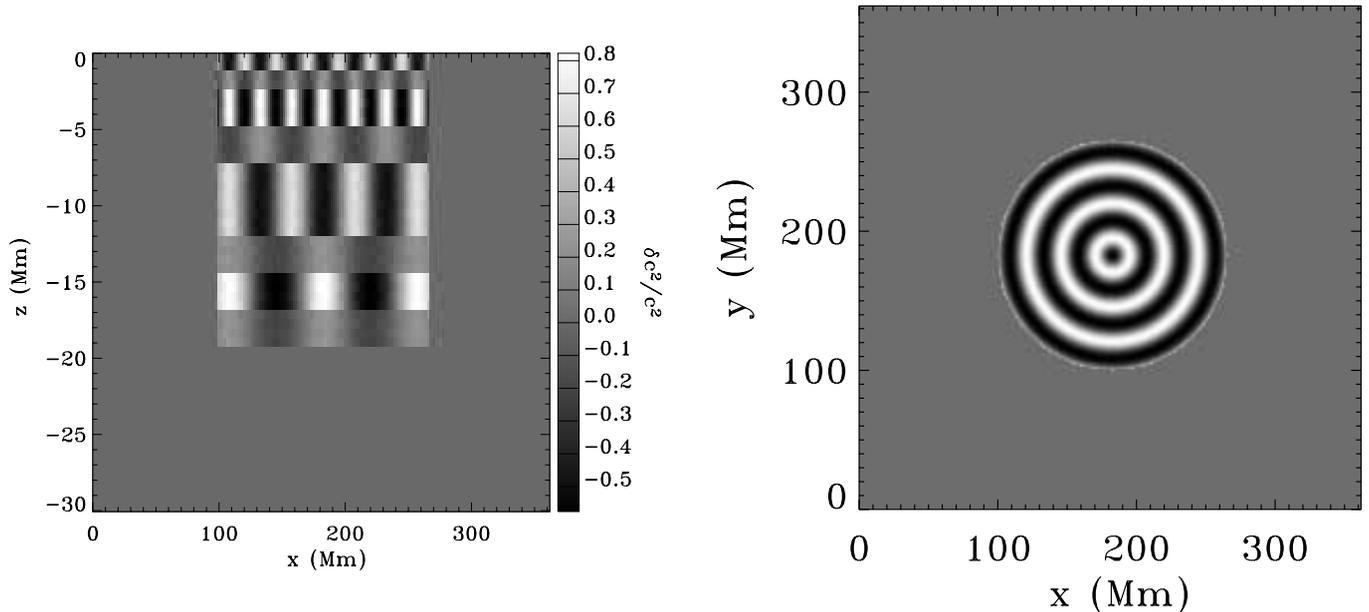


FIG. 13.—*Left panel*: vertical cut through the input sound speed perturbation, $\delta c^2/c^2$. This perturbation is a checkerboard-like structure. *Right panel*: horizontal cut at the surface.

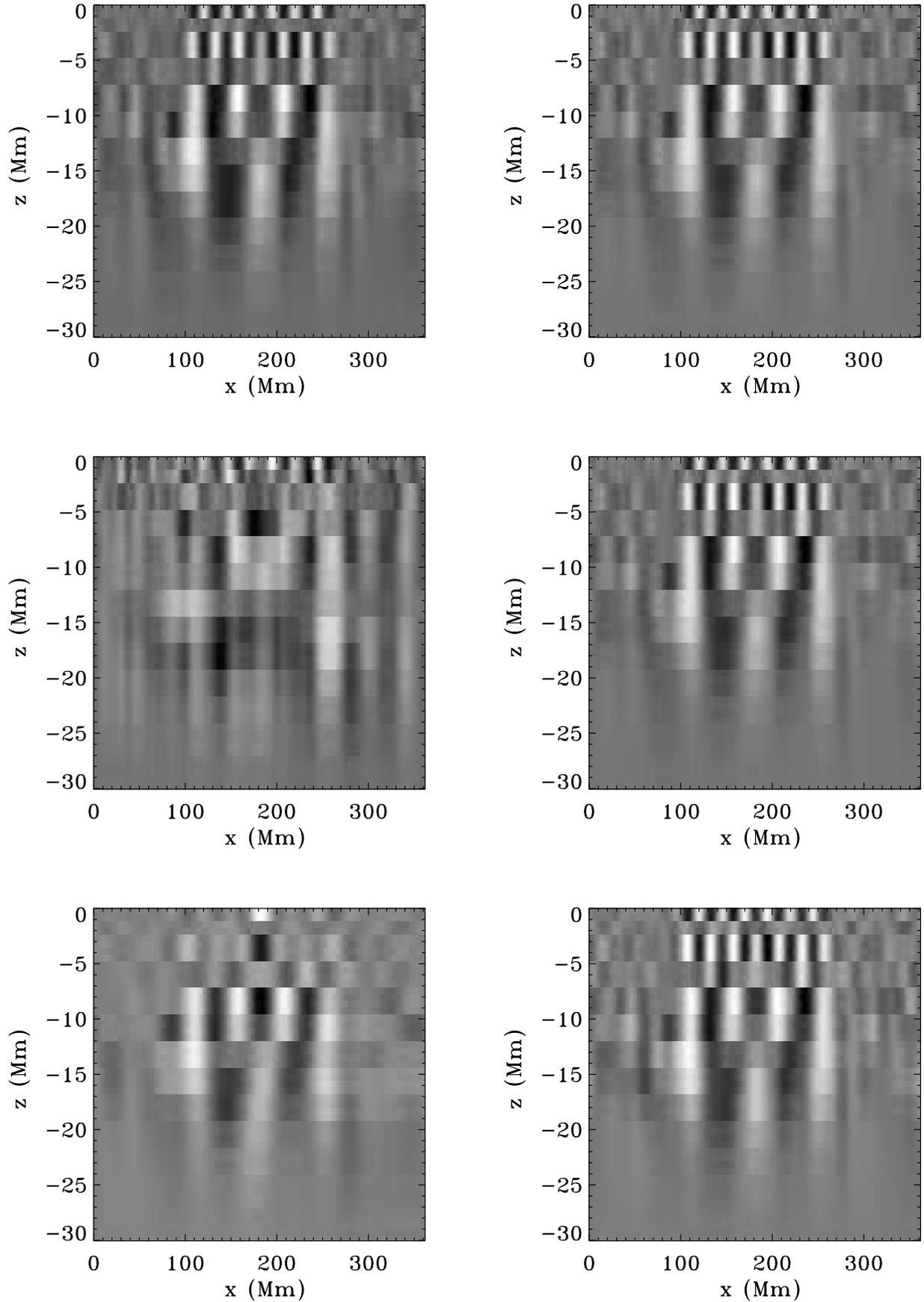


FIG. 14.—Vertical cuts in the inversion results for the perturbation simulating a checkerboard-like structure (Fig. 13). The left column shows the results obtained without including the noise covariance matrices in the inversion procedure, the right column shows the results obtained with the noise covariance matrices. We show the inversion results derived with the 11 mean Δ -values and phase speed filters listed in Table 1 (*upper panel*), with additional vertical correlations (*middle panel*), or with additional horizontal correlations (*lower panel*). The gray scale is the same as on Fig. 13.

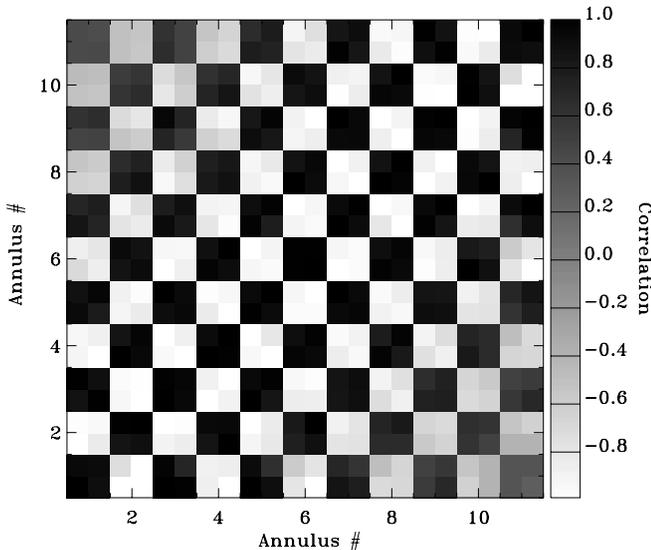


FIG. 15.—Modified vertical correlation matrix obtained for the phase speed filters listed in Table 1: we added extra vertical correlations between the travel-time maps.

The averaging kernels for this section are much better localized when the Σ^{-1} are included, which means a finer horizontal and vertical resolution on the inversion results.

6. CONCLUSION

Although inverse algorithms of time-distance data already reach a good level of accuracy, they still need to be refined to fully take advantage of the data that will be provided by the future Helioseismic and Magnetic Imager (HMI) instrument on board the *Solar Dynamics Observatory* satellite. This refinement implies better sensitivity kernels, faster, less memory-consuming, and more accurate algorithms, and taking into account the noise covariance. In this paper, we applied the recently developed Born approximation kernels to the forward and inverse problems. These kernels introduced in Birch et al. (2004) take into account the sensitivity of waves to perturbations off the ray path. We showed that inversions can be made with these kernels that are expected to derive more realistic sound speed perturbations. To improve the inversion we introduced horizontal regularization in the multichannel deconvolution algorithm. This allows us to regularize with the norm of the first or second derivatives of the solution in the horizontal direction. The inversion results are significantly improved, if we base on the minimum mean squared error. Unfortunately, this makes it more time-consuming to find the best regularization parameters, and work needs to be done to determine the best estimator to pick these parameters.

We suggest that when inverting the sound speed perturbations below localized structures, applying the L -surface criterion to a small region of the $\delta c^2/c^2$ slab where the signal is located should yield better inversion results than applying this criterion to the entire slab.

The main results concern the noise correlations and their impact on the inverted sound speed perturbations. We managed to include the full noise covariance matrix in the inversion code. This does improve the results, provided that the sound speed perturbation changes rapidly on small scales. If the perturbation is very smooth, meaning that the vertical and/or horizontal length scale of variation is larger than the characteristic correlation length, the covariance matrix has—at best—no impact on the inversion result. It may even degrade these results by increasing the noise variance. However, when inverting actual time-distance data from MDI, we do not know the vertical shape of the sound speed perturbations, which makes it safer to include the Σ^{-1} matrices in the inversion procedure: if the $\delta c^2/c^2$ perturbation presents small-scale features, including these matrices definitely helps to retrieve the correct perturbation shape and amplitude. Also, it is worth mentioning that when working on point-to-point instead of point-to-annulus travel times, the spatial correlations are expected to be larger, making the use of Σ^{-1} a necessity. The computation of Σ^{-1} is very time-consuming, especially if we want these matrices to be accurate enough, but it needs only to be done once for a specific set of Δ distances and phase speed filters. Moreover, the inversion procedure is not slowed down when the noise covariance is included.

We emphasize that the phase speed filters greatly impact on the travel-time noise correlations. With the filters we used, these correlations are very low horizontally and low vertically. Hence, their limited impact on the inversion results except for very sharp sound speed perturbations. Using narrow phase speed filters keeps the correlations low on the travel-time maps, while producing temporal cross-covariance functions with an appropriate noise level. However, there is a trade-off between noise level in the travel-time measurements and filter broadness. A good compromise seems to be attained with filters comparable to the ones listed in Table 1.

Finally, our inversion procedure appears to give satisfying results, and in a next step we shall apply this procedure to real quiet-Sun data. We shall also develop a similar inversion tool for the retrieval of flow velocities.

This work was supported by NASA grants NAG5-12452 and NAG5-13261. A. C. B. is supported by NASA contract NNH04CC05C.

REFERENCES

- Belge, M., Kilmer, M. E., & Miller, E. L. 1998, *Proc. SPIE*, 3459, 328
 ———. 2002, *Inverse Problems*, 18, 1161
 Birch, A. C., Kosovichev, A. G., & Duvall, T. L., Jr. 2004, *ApJ*, 608, 580
 Couvidat, S., Birch, A. C., Kosovichev, A. G., & Zhao, J. 2004, *ApJ*, 607, 554
 Duvall, T. L., Jr., Jefferies, S. M., Harvey, J. W., & Pomerantz, M. A. 1993, *Nature*, 362, 430
 Duvall, T. L., et al. 1997, *Sol. Phys.*, 170, 63
 Giles, P. M. 2000, Ph.D. thesis, Stanford Univ.
 Gizon, L., & Birch, A. C. 2002, *ApJ*, 571, 966
 ———. 2004, *ApJ*, 614, 472
 Golub, G. H., & Van Loan, C. F. 1996, *Matrix Computations* (2nd ed.; Baltimore: Johns Hopkins Univ. Press)
 Gough, D. O., & Sekii, T. 2002, *MNRAS*, 335, 170
 Hansen, P. C. 1997, *Rank Deficient and Discrete Ill-Posed Problems* (Philadelphia: SIAM Publications)
 Hughes, S. J., Rajaguru, S. P., & Thompson, M. J. 2004, in *Proc. SOHO 14 / GONG 2004 Workshop*, ed. D. Danesy (ESA SP-559) (Noordwijk: ESA), 493
 Jensen, J. M. 2001, Ph.D. thesis, Univ. Aarhus, Denmark
 Jensen, J. M., Duvall, T. L., Jr., & Jacobsen, B. H. 2003, in *Proc. SOHO 12/ GONG+ 2002 Workshop*, ed. H. Sawaya-Lacoste (ESA SP-517) (Noordwijk: ESA), 315
 Jensen, J. M., Jacobsen, B. H., & Christensen-Dalsgaard, J. 1998, in *Proc. SOHO 6/GONG 98 Workshop*, ed. S. Korzenik & A. Wilson (ESA SP-418) (Noordwijk: ESA), 635
 Kang, M. G. 1998, *Opt. Eng.*, 37, 2953
 Scherrer, P. H., et al. 1995, *Sol. Phys.*, 162, 129