First-arrival Travel-Time Tomography using Second Generation Wavelets

Matthieu Delost, J. Virieux, Stéphane Operto

First-arrival Travel-Time Tomography using Second Generation Wavelets

Matthieu Delost, Jean Virieux and Stéphane Operto

Géosciences Azur, CNRS and UNSA , Sophia-Antipolis, France

October 31, 2007

Abstract

Wavelet decomposition of the slowness model has been proposed as a multiscale strategy for seismic first-arrival time tomography. We propose the introduction of so-called second generation wavelets which could be used for any mesh structure and does not need a number of samples as the power of two in each direction. Moreover, one can handle easily boundary effects. A linearized procedure for inverting delayed travel-times considering either slowness coefficients or wavelet coefficients. The ray tracing is solved at each iteration through an eikonal solver while the linear system to be solved at each iteration goes through an iterative solver as $LSQR$ algorithm. We develop wavelet decomposition over constant patches (Haar wavelet) or over linear patches (Battle-Lemarie wavelet) of coefficients at different scales. This decomposition is introduced in the linear system to be solved and wavelet coefficients are considered as unknowns to be inverted. Synthetic examples show that the inversion behaves in a better way as wavelet decomposition seems to act as a preconditioner of the linear system. Local discretisation is possible but requires additional implementation as artefacs once built inside the model description never disappear because of the linearized approach. A binary mask operator is designed for each scale grid and could be applied locally leading to quite different spatial resolution depending on the analysis we could perform of the expected resolution at a given position of the medium. We show that indeed it is possible to design this binary operator and we apply it to synthetic examples as a crosswell experiment inside the Marmousi model. An application to a surface-surface experiment has been performed and the waveled decomposition shows that indeed we may recover detailed features nearby the free surface while preventing imprints of ray coverage at greater depths giving us smooth features at that depths. In spite of the increase demand of computer resources, the wavelet decomposition seems to be a rather promising alternative for controlling the resolution variation of seismic first-arrival tomography.

1 introduction

First-arrival delayed traveltime tomography (FADTT) applied to wide aperture seismic data has proved to be an interesting tool to investigate the Earth structures (Aki & Lee, 1976, Spakman & Nolet, 1988, Hole, 1992, Zelt & Smith, 1992, Benz et al., 1996, Zhang et al., 1998). The resolution of FADTT is controlled both by the theoretical resolution imposed by the relation between velocity structure and first-arrival travel-times and by the data acquisition configuration as well as the structure itself, which may lead to uneven ray coverage. The theoretical resolution power of FADTT is limited by the size of the first Fresnel zone for each ray. There is an antinomy between the limited resolution power of FADTT and the numerical sensitivity kernel of FADTT corresponding to rays. Rays have no thickness and, therefore, have an intrinsic infinite sensitivity although in practise discretization of the model space always smoothes out this feature (Dahlen, 2004). Indeed, one raypath induces high frequency information along the ray, even at locations where resolution of FADTT is expected to be low. Inversion can build up models which have a high degree of data fitting but which are far from the true model containing artefacts. Accounting for the limited resolution of FADTT goes through regularized tomographic inversion (Farra & Madariaga, 1988, Delprat-Jannaud & Lailly, 1993, Zhang et al., 1998) often based on the Tikhonov regularization approach (Tikhonov & Arsenin, 1977). The introduction of the covariance matrix for the model as well as for the data (Tarantola & Valette, 1982, Tarantola, 1987) has introduced a precise theoretical background although estimation of those matrices is still a difficult task (Monteiller et al., 2005). Another alternative regularization is based on smoothing constraints easier to implement as the laplacian (Menke, 1984, Zhang et al., 1998). Whatever we choose, regularization strategies lack spatial adaptivity. Frequently, when the smoothing is tuned to remove the footprints of raypaths in poorly illuminated areas, short-scale features that could be resolved in other areas of the model according to the resolution of FADTT are smoothed simultaneously, leading to a loss of information in the well-resolved areas.

The resolution of FADTT generally varies locally. For example, considering surface acquisition geometries, the shallow structures will be sampled both by short- and long-offset rays running across the heterogeneities with different azimuths while the deeper structures will be sampled by long-offset rays only with a much narrower range of azimuths. The resolution of FADTT decreases quite significantly with respect to depth for this geometry. Two other factors creating non-uniform ray coverage are the presence of low velocity anomalies which induces shadow zones in the ray coverage and non-uniform source and/or receiver arrays especially in

the case of passive tomography.

In order to consider this spatially varying resolution power, one could introduce adaptive parametrizations based on non-structured grids (Michelini, 1995, Spakman & Bijwaard, 1998, Spakman et al., 1998, Bohm et al., 2000, Trinks et al., 2005, Zhang & Thurber, 2005). The size of the elementary cells are adapted to the local ray coverage such that the amount of control for each model parameter tends to be uniform. More promizing is the multigrid approach where the inversion procedure recovers values on different grids depending on the resolution power (Madych, 1999, Chiao & Kuo, 2001, Chiao & Liang, 2003, Zhou, 2003) based on wavelet tools (Mallat, 1999). In this article, we shall propose an further investigation of this multigrid approach using new wavelet transform named as second-generation wavelets (Sweldens, 1997).

2   Wavelet transformation

Over the last years, many constructions of wavelets have been introduced both in mathematical analysis and in the signal processing literature. In mathematical analysis, wavelets were originally constructed to analyse and describe geophysical signals among other signals using translations and dilations of one specified function called mother wavelet. A mathematical framework has been set up by the "French school" (Daubechie et al., 1986, Grossmann & J.Morlet, 1984). In signal processing, wavelets found their ways in the context of quadrature mirror filters (Mintzer, 1985, Vetterli, 1986, Nguyen & Vaidyannathan, 1989, Vaidyannathan, 1993, Woods & O'neil, 1986). The connection between the two approaches has been performed through the introduction of multiresolution analysis and the fast wavelet transform by Mallat (1989a,b), Meyer (1999). An important step has been achieved with the construction of orthogonal, compactly supported wavelets by Daubechies (1988). Since then, several generalizations to the biorthogonal or semi-orthogonal cases were presented (Cohen et al., 1992, Vetterli & Herley, 1989). Biorthogonality allows the contruction of symmetric wavelets and thus the use of linear phase filters.

For seismic tomography, wavelets can be used for description of the model at various scales (Chiao & Kuo, 2001, Chiao & Liang, 2003, Zhou, 2003). The slowness model is expanded on a wavelet basis and the tomographic system is solved for the wavelet coefficients. The wavelet transform of the model provides a compact multigrid representation of the model thanks to orthogonal or bi-orthogonal basis functions built by translating (localization property) and dilating (scaling property) a mother wavelet and a scaling function. Those functions are basis of functional spaces generaly noted V_i for the scaling functions et W_i for the wavelet functions. Those spaces verify certain properties and constitute what is called a multiresolution analysis (Mallat, 1999). Each

level i corresponds to a grid size. Generally, the resolution of each grid decreases by a factor 2 and the mesh spacing is set accordingly when the level i increases. The starting point is the slowness model on the finest grid of the multiresolution representation (generally V_0). The wavelet transform applies a cascade of orthogonal restrictions of the slowness model over different approximation grids. The restrictions of the slowness model on a coarse grid is encoded by the scaling coefficients. Those coefficients are the components along the scaling functions (generally noted ϕ_i), basis functions of space V_i. At each step, the increment of information lost during the restriction on the coarser grid is encoded by the wavelet coefficients before proceeding to the next grid. Those coefficients are associated to wavelet functions (generally noted ψ_i), the basis functions of the spaces W_i. At the last iteration of the orthogonal transform, we end up with the wavelet coefficients on each grid plus the scaling coefficients on the coarser grid (see figure 1 for a simple example with a 1D signal). The inverse wavelet transform proceeds in the other direction from the coarse grids to the finer ones.

Multidimensional wavelet basis can be built by a tensor product of 1D basis. The decomposition of the 2D slowness model of a wavelet basis can be written as

$$
\begin{aligned}
u(x,z) = &\sum_{j=1}^{J} \sum_{n=-\infty}^{+\infty} \sum_{i=1}^{I} \sum_{m=-\infty}^{+\infty} c_{j,n,i,m}^{ww} \psi_{j,n}(x)\psi_{i,m}(z) \\
&+ \sum_{j=1}^{J} \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} c_{j,n,I,m}^{ws} \psi_{j,n}(x)\phi_{I,m}(z) \\
&+ \sum_{n=-\infty}^{+\infty} \sum_{i=1}^{I} \sum_{m=-\infty}^{+\infty} c_{J,n,i,m}^{sw} \phi_{J,n}(x)\psi_{i,m}(z) \\
&+ \sum_{n=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} c_{n,I,m,J}^{ss} \phi_{J,n}(x)\phi_{I,m}(z),
\end{aligned}
\tag{1}
$$

where c^{ww}, c^{ws}, c^{sw} and c^{ss} are the wavelet and scaling coefficients. The indices i,j and n,m are the scale and localization indexes respectively. The indices I and J denotes the coarsest levels of the multiresolution approximation. A compact representation of the wavelet coefficients of a 2D signal has been proposed by Mallat (Figure 2).

As Fourier driven transformation, wavelet transformations encounter some relevant drawbacks: the number of samples has to be a power of two, samples have to be regularly spaced and one must deal periodicity which complicates boundary conditions when working on edges or surfaces.

For overcoming these difficulties, the concept of biorthogonal wavelets constructed by lifting scheme has been introduced (Sweldens, 1994, Sweldens & Schroder, 1995, Sweldens, 1997), leading to the so-called second generation wavelets. Wavelets are not necessarily translations and dilations of a defined function. Such wavelets can be adapted to curved edges and surfaces delimited quite complex domains. They are concerned by weighted and irregular samples. The lifting scheme leads to a faster, in-place calculation of the wavelet

transform.

In term of signal processing, projections on spaces V and W can be seen as respectively high-pass and low-pass filtering operators. A procedure of direct and inverse transforms can then be interpreted as a perfect reconstruction filter bank (h_1, g_1, h_2, g_2), where (h_1, g_1) are the filters for the direct transform and (h_2, g_2) are the filters for the inverse transform. In the orthonormal transform, $h_1 = h_2 = h$ and $g_1 = g_2 = g$ and (h, g) are called conjugate mirror filters. In the biorthogonal case, relaxing the orthogonality condition implies working with larger spaces as $\tilde{V}$ and $\tilde{W}$ spaces, respectively the dual spaces of V and W. Dual spaces are used for the direct transform and normal spaces are used for the inverse transform (or *vice-et-versa*). The associated perfect reconstruction filter bank is denoted $(h, g, \tilde{h}, \tilde{g})$. The lifting scheme is based on the very simple property demonstrated by Sweldens (1994):

Given a biorthogonal filter, another one can be built as follows

$$(h, g, \tilde{h}, \tilde{g}) \Rightarrow (h, g_1, \tilde{h}_1, \tilde{g})$$

$$\begin{cases} \tilde{H}_1(w) = \tilde{H}(w) + S(2w)\tilde{G}(w), \\ G_1(w) = G(w) - S(2w)H(w), \end{cases}$$

where capital letters stand for the trigonometric polynomials associated with script letters filters and S is a simple trigonometric polynomial.

This polynomial is chosen so that the wavelet and dual scaling functions satisfy desired properties (as the number $\tilde{N}$ of vanishing moments of the dual wavelet for example). Of course, a dual lifting can also be defined in a similar way:

$$(h, g, \tilde{h}, \tilde{g}) \Rightarrow (h_1, g, \tilde{h}, \tilde{g}_1)$$

$$\begin{cases} H_1(w) = H(w) + \tilde{S}(2w)G(w), \\ \tilde{G}_1(w) = \tilde{G}(w) - \tilde{S}(2w)\tilde{H}(w). \end{cases}$$

With this operation, the real scaling function and the dual wavelet functions can then also be tuned, so they have a requested number N of vanishing moments.

These properties assert that, given a biorthogonal filter, one can construct another one satisfying chosen properties. Starting from a simple biorthogonal filter, it is then possible to bootstrap one's way up to a multiresolution analysis with desired properties on primal and dual wavelets.

The so-called "lifting scheme" proceeds with this strategy: one starts with the straightforward biorthogonal filter that does nothing but set aside odd an even indexed coefficients (the so-called Lazy Wavelet). Then one first does a dual lifting operation by predicting odd indexed coefficients using even indexed ones (we use the spatial correlation between neighbours). We thus ensure that polynomials of degree up to $N-1$ are well predicted. In a second time, we update the even indexed coefficients using new odd coefficients, so that $\tilde{N}$ moments of the signal are conserved. The algorithm can simply be written as

$$\begin{cases} (odd_{i+1}, even_{i+1}) = Lazy(\lambda_i), \\ odd_{i+1} - = P(even_i), \\ even_{i+1} + = U(odd_i). \end{cases}$$

The inverse transform can be written similarly as

$$\begin{cases} even_{i+1} - = U(odd_i), \\ odd_{i+1} + = P(even_i), \\ \lambda_i = Merge(odd_{i+1}, even_{i+1}). \end{cases}$$

This numerical flowchart can be extended because the predict stage is a simple polynomial interpolation and the update stage is all but the resolution of linear systems: boundary conditions and number of samples are no more a limitation of the wavelet decomposition.

3 Travel-time tomography

Once rays have been computed between each source and each receiver seeing the source through an eikonal solver (Podvin & Lecomte, 1991), the delayed traveltime tomographic system relating slowness anomaly Δu to traveltime residuals Δt written as

$$\Delta t = A\Delta u, \tag{2}$$

can be transformed into the new system

$$\Delta t = AW^t \Delta c, \tag{3}$$

where the A matrix is the sensitivity or Fréchet matrix and the equation $\Delta c = W\Delta u$ denotes the transformed slowness model through the transform matrix W. The matrix AW^t can be built by computing the 2D wavelet transform of each row of the original matrix A. A row of the sensitivity matrix has $n_1 \times n_2$ number of model parameters in a 2D geometry where the number of horizontal grid points is n_1 and the number of vertical grid points is n_2: it contains the contribution of one ray to the Fréchet derivatives. After transformation, we have the same number of unknowns but they are in the wavelet domain and partial derivatives are for wavelet coefficients. Once the system equation (3) has been solved through an iterative procedure as the LSQR algorithm (Paige & Saunders, 1982), the perturbation coefficients Δu model in the slowness space are

obtained by the inverse wavelet transform of $\Delta c = W\Delta u$. We shall plot velocity models from these slowness values.

Traveltime tomography in the wavelet domain has an interesting feature: as coefficients are associated to levels of the wavelet transform and, consequently, related to scales and local resolutions, adaptivity strategy is much easier to handle through scale hierarchie. The decomposition of a simple circular anomaly leads to a serie of grids for each scale level using square wavelet basis functions as shown in the figure 3. Controlling incremental perturbations through wavelet coefficients between these different description of the true model taken as the one in the V_0 space is an hierarchical procedure.

As we estimate imprints of rays on different coarser grids through estimations of the Fréchet derivatives, we extend equivalently its area of sensitivity as shown in figure 4. The sensitivity of a ray is related to the spatial discretisation of the slowness model at the given wavelet scale.

The adaptivity of the parameterization can be implemented by applying a mask to each row of the sensitivity matrix AW^t, zeroing coefficients in areas of the model where a poor resolution is expected et a given scale. This adaptivity selection can be implemented via an linear operator M applied to the Fréchet derivatives in the wavelet domain through the following extended linear system

$$\Delta t = AMW^t\Delta c. \tag{4}$$

Designing this operator will depend on local resolution analysis. For surface acquisition, one could, for example, define an empirical rule relating the resolution of the model at a given depth to the minimum source-receiver offset of the rays sampling this depth. The source-receiver offset would provide an estimate of the width of the first Fresnel zone at a given depth and, hence, of the expected resolution of the tomography varying with depth. According to the sampling theorem, only grids with mesh spacing twice smaller than the width of the first Fresnel zone, would be locally involved in the inversion: a simple rule for construction the M operator. We will discuss later how we can propose a more general strategy for building this operator. Let us remark that the wavelet transform has also the capacity to automatically adapt the parameterization to the resolution in the case where the real resolution is lower than the intrinsic one such as in the case of a shadow zone. In areas where no rays illuminate the medium, partial derivatives of the traveltime time with respect to the wavelet coefficients at the finest scales will be zero, thanks to the second generation property. On the contrary, the partial derivative of the traveltime time with respect to the wavelet coefficients of the coarsest

7

scales will be non-zero due to the spatially extended sensitivity of the coarse grids representations of the rays. Hence, only the coarse-scale wavelet coefficients of the slowness model will be non-zero after inversion and will automatically lead to a smooth representation of the true model.

4   Synthetic examples

We present synthetic examples to illustrate how the wavelet decomposition enables us to reconstruct the P-wave velocity model. First, we focus our attention to the wavelet coefficients transformation. We shall show that similar results are found when compared with those of the standard tomography using slowness coefficients. Finally, we illustrate the efficiency of scales selection using the simple binary operator M. Designing an objective criterion for the spatial definition of this binary operator will be based on resolution matrix analysis.

A densely illuminated model with positive anomalies

As a first example, we design an acquisition system for a complete illumination of the simple circular positive velocity anomaly of $0.5km/s$ we want to reconstruct (Figure 5) from a background model of constant velocity of $3km/s$. This background velocity model will be our starting model.

For each inversion, we estimate the lowest rms error from the misfit function as it flattens with linearized iterations. Moreover, we display horizontal and vertical sections for a better appreciation of the spatial frequency contents.

First, the model is reconstructed using slowness parameters in the inversion over a grid of $5\ m$ step size. No damping is applied to the over-determined system. The convergence is obtained after four iterations and the anomaly is accurately recovered although with high frequency artefacts (Figure 6). Other artefacts are related with uneven ray sampling in the center of the anomaly. How better will be a reconstruction in the wavelet domain ?

When considering wavelet decomposition, we select wavelets with two vanishing moments. We consider two and four levels for the scale sampling. This means that the coarser grid size will have respectively a $10m$ resolution and a $40m$ resolution. The anomaly is well recovered but the convergence to the flat level of the misfit function is slightly slower than before (Figures 7 and 8). We need respectively eight and thirteen iterations for reaching convergence. We may underline that the rms final error decreases as we increase the number of levels although the decrease is quite small. For real applications with errors in the picked traveltime, this may not be detected. More importantly, the high frequency artefacts are less noticeable providing a better fit of the model. As we increase the number of levels, we can see the decrease of the amplitude of the high

frequency artefacts. In other words, when coarser levels are introduced, the reconstruction is driven by a better spatial coherence, essentially for broad features. We may say that the parametrization into the wavelet domain provides a better conditionning of the system to be inverted, since it is less sensitive to high frequency perturbations.

A densely illuminated model with negative anomalies

We may proceed similarly when considering a negative velocity anomaly. The velocity amplitude will $-0.5\ km/s$ and the background velocity model will be the starting model. The reconstruction is more difficult as rays tend to avoid low-velocity zones creating shadow zones which are not sampled by rays (Figure 9). Velocity anomaly should ne be reconstructed in the center of the medium as rays are missing there. Therefore, reconstructed anomaly comes from the linearized iteration procedure where rays are present in the middle at the early stages of linearized. We may say that values are with constraints and come frome the null space contribution. One must underline that this behaviour is typical of linearized procedure.

The results with the negative velocity anomaly are very similar with those for a positive velocity anomaly, although fondamental differences could be observed (Figure 10). When considering a slowness description of the medium, we observe a poor reconstructed structure in the center. Rays do not sample this zone at the end of the procedure, leaving uncorrected slowness obtained at the initial stages of the inversion. Because of this poor sampling, edge anomalies are observed for correcting this imperfect reconstruction in the middle as few rays of different azimuths are emitted nearby sources and receivers. These edge anomalies are hardly seen for the positive anomaly.

By introducing wavelet coefficients with two vanishing moments, we recover the negative anomaly with an increased number of iterations (Figures 11 and 12). The final rms residue is smaller although the decrease is not significant. Because the anomaly we want to recover has a smooth wavelength content, the result is better when considering four levels of discretisation. Still, partial reconstruction introduces artefacts which spread inside the medium from the edges as we move to the four levels description. We can see that wavelet coefficients could not overcome automatically aliasing in the reconstruction essentially based on the linearized procedure. Therefore, we need the operator M for tuning the reconstruction on the specific level we expect in different zones of the medium.

In conclusion of these simple synthetic tests, the wavelet transformation has allowed a better preconditionning of the linearized system to be solved: high frequency perturbations are less noticeable and scale decomposition increases the stability of the reconstructed system at the expense of the number of iterations.

Of course, this analysis is for an ideal coverage by rays and the experiment has been conducted without any noise in the picked times. In spite of this ideal configuration, tuning the parametrization description with the help of expected varying resolution should be added to the automatic wavelet coefficient reconstruction.

A typical surface-surface synthetic example with two different anomalies

Moreover, dealing with uneven ray coverage is a crucial issue in tomography: how the wavelet decomposition performs when illumation is far from an ideal one. The experimental data acquisition as well as the structure itself always creates areas of limited ray coverage. Hence parts of the model are expected to be known precisely, while other parts will suffer from low resolution. Wavelets deal with the multiscale nature of the information contained in the data but the inversion is not able to focus automatically anomalies on related scales. Working in the wavelet domain makes easy the selection scales of the parameter model by a simply predefined operator. Zeroing wavelet coefficients corresponding to scales finer than the expected local resolution will help spatial aliasing of the inversion procedure. By applying a mask operator to each row of the sensitivity matrix written in the wavelet domain, we may successfully distinguish well-resolved and poorly-resolved zones in our reconstruction.

Applying a mask does not increase the computational need as we perform a multiplication by zero or one for each term of the matrix. Defining the mask may be quite demanding although pre-tabulated masks should be deviced for typical data acquisition configuration.

The effect of this simple operation is illustrated by the figure 13. In this example, we use two levels of wavelets. the model is a simple constant background model with two heterogeneities a positive and a negative one, enlightened from both sides. For the coarsest level, we keep all the coefficients (the upper left black dial of the mask). For finer levels, we only keep coefficients in a rectangle localized in the center of the model. Every boundary fine coefficients are scaled down to zero. The resulting model is similar to the model constructed whithout mask in the rectangle in the center. On the boundaries, only coarser coefficients have been kept, thus a coarser model coming from the poor ray sampling nearby sources and receivers.

A strongly spatially varying illumination related to a poor ray coverage comes out for a surface acquisition configuration. In this case, the theoretical resolution is expected to be heterogeneous: in the shallow part, both long and short rays spanning across a large range of azimuths make the resolution quite accurate. On the contrary, the deepest parts of the model are sampled and enlightened only by longer rays with a mainly horizontal azimuth leading to low resolution.

Because of the linearized formulation, we thus need a method that can adapt the resolution to the theoretical

one aside the natural wavelet reconstruction. Wavelet paramatrization makes it easy to adapt resolution locally, by applying a simple mask. By using an adequate mask, two velocity heterogeneities, a small one in the shallow part and a bigger one in the deep part, are reconstructed according to the expected resolution. The starting model is a constant velocity model without anomalies (Figure 14).

The figure 15 shows the reconstruction through slowness coefficients without any regularization. One can easily see the smearing related with ray sampling for the broad deepest anomaly. The small heterogeneity is well reconstructed (see horizontal and vertical sections A and C of the figure 15) but the model contains high frequency perturbations. The big one is not well fitted, due to enlightening which is not as dense as in the shallow part. Still results contain a reasonable amount of low frequency information. Nevertheless high frequency information are quite strong and should be removed somehow.

Dealing with the under-determination of the linearized system because of the poor ray coverage could be achieved throuhg smoothing gaussian operators, although these operators are hard to define when working with slowness coefficients. The standard encountered drawback is the over-smoothing by cutting down too small eigenvalues of the linearized system. The level of smoothing needed to remove high frequency perturbation in low resolution zone leads to a loss of high frequency in better defined parts of the model (see Figure 16).

Wavelet decomposition allows different resolutions at different places in the model, as the traveltime tomography predicts. Once the Fréchet derivative is computed, we applied the M operator defined as follows. From the surface down to the depth of 800 m, wavelet coefficents are kept at all scales. From the depth of 800 m down to the bottom, only the fourth level of wavelet coefficients is kept, corresponding to a resolution of 400 m, coefficients of the finer other three scales are zeroed out. The inversion provides a reconstruction which varies in resolution as we move to deeper depths (Figure 17). We have chosen the mask operator in order to separate zones of the two targets. This illustrates the local resolution with the wavelet decomposition: a high resolution in the shallow part and a broader resolution in depth. One can see than the section A is quite similar to the same section in the figure 15. The vertical section C displays the same profile for the shallow part as the figure 15 while the deepest part is smoother and similar to the same section of the figure 16. Similar features are displayed for the section D: the shallow part still contains high frequency information, as all wavelet coefficients have been kept, whereas the deep part is a smooth version of the same section of the figure 15. The high frequency artefacts have also been removed from the horizontal section B of the largest heterogeneity, thanks to the wavelet decomposition and the mask application.

We may conclude that wavelets enable us to adapt locally the resolution to the theoretical resolution of traveltime tomography. This adaptation is done with a multiplication between wavelet coefficients and the mask values, as the Fréchet derivative matrix is calculated in the wavelet domain. How to choose the mask operator is a problem we face?

5 Resolution analysis for defining the binary M operator

For better focusing of the reconstruction through different scale levels, one has to proceed through a resolution analysis in order to overcome difficulties coming from the linearized inversion. A more exhaustive exploration of the parameter space may be of great interest but it is beyond the scope of this paper. We assume that a model has been designed for which one can perform a resolution analysis. It could be the initial model or any model obtained through previous attempts of inversion. Starting with this model, we may want to have an rough estimation of the resolution matrix for better performance of the inversion (Nolet et al., 1999). Of course, once done, we may end up with another medium we shall use again for a new resolution analysis and so on. We cannot rule out this loop over adapted models.

For small systems, one can use a direct solver and computing the resolution matrix is quite fast. For important systems, we may rely on numerical iterative techniques as proposed by Latorre et al. (2004). As shown in the figure 18, a small local velocity perturbation is introduced with an amplitude of 0.5 km/s, making a spike with a given spatial extension. This amplitude is such that we have a perturbation of travel times above the numerical noise and that rays are kept nearly undeformed. Therefore, we may apply a linear procedure for the reconstruction of this spike. We keep rays and Fréchet derivatives computed in the reference model speeding up dramatically the spike reconstruction as we perform it in different zones of the medium for different spatial extensions of the spike. At a given scale, a resolution matrix estimation as shown in the figure 18 will drive us in the reconstruction of the mask at that scale. We zero out Fréchet derivative when less than 95 % of the wavelet amplitude is recovered at the central position of the spike. This decrease of the amplitude is related to the spreading of the reconstructed spike. Hence, we analyze only the amplitude right at the position where is the spike. Because the system is linear, positive or negative amplitudes give the same resolution matrix extimation.

6   A realistic crosswell experiment: application to the marmousi model

This procedure has been applied to a synthetic crosswell experiment performed inside the marmousi model (Bourgeois et al., 1991, Versteeg, 1994). Between two wells of 3 km depths at a distance of 1 km (Figure 19), ray tracing has been performed for computing synthetic travel times. A smooth model has been constructed and a resolution analysis performed on this smooth model leading to an estimation of a resolution map with a 95 % threshold level (Figure 19). This resolution analysis has led us to zero out the finest fifth level of Haar coefficients we are using in our wavelet decomposition as well as specific zones of other scale levels (Figure 20). In other words, we may start from a very fine grid and the resolution analysis will tell us if this grid could be used for inversion. Inversion is performed using either slowness coefficients or wavelet coefficients with the mask application (Figure 20). One can see that, where the ray sampling is not fine enough especially at the bottom of the two wells, a coarser description of the reconstructed image has been designed through the resolution analysis. Wavelet procedure seems to perform better and provides nicer features than the slowness reconstruction.

7   Application to a real data set

We have analyzed a real data set with a global offset configuration using a surface-surface geometry. A impressive number of shots as high as 702 shots with a regular spacing of 25 m have been recorded by 1440 receivers along a 18 km 2D profile with a spacing of 12.5 m. We have considering half the sampling over shots. First-arrival travel-times for 301 shots have been picked automatically with an industrial software for the entire receiver line and a tentative macromodel of 18 km long and 4 km deep should be reconstructed using these travel-times. The figure 21 displays one shot gather with initial picked time curve and computed time curve. Our slowness inversion grid has 169 vertical nodes and 725 horizontal nodes with a spatial step of 25 m. We design an initial model for the tomography by fitting a mean traveltime curve with a 1D velocity model. We apply this 1D model below the given topography typical of a foothill structure (Figure 22) using a spatial discretization of 100 m on both directions. The model through FADTT procedure has to fit as much as possible the data and it also has to be smooth especially for deep zones as the ray sampling becomes quite poor.

By a proper gaussian smoothing strategy with depth, a 2D velocity structure has been constructed through slowness coefficients (figure 23) with the same spatial discretization as the initial model. No damping is applied

except the internal damping of the LSQR procedure. The number of linearized inversion is five and the rms final error is 26 ms. This model seems to take into account spatial low-frequency contents while the high-frequency ones are missing.

This model is used for a resolution analysis (Figure 18) in order to build up the M mask operator which is displayed in the figure 24. We have selected three levels of the wavelet decomposition starting from 100 m spatial spacing. The coarsest grid is therefore with a spatial discretization of 400 m. Using these three characteristic lengths, we have performed resolution analysis and we have adopted a criterion of a 95 % amplitude reconstruction at the center of the spike for keeping Fréchet derivatives coefficients at that level (Figure 24). One can see that the mask is the identity at the scale of 400 m.

By using linear interpolation wavelet coefficients with two vanishing moments, after thirty iterations, a model with an higher spatial frequency content nearby the free surface has been obtaines (Figure 25). Patterns related with the ray sampling have kept under control and do not overprint the reconstructed image at depth. The rms final error does not show an significant decrease for assessing the quality of the recovered image through this simple quality criterion. By displaying residual times between sources and receivers (Figure 26), one may notice the difficult fitting at near offsets related both to a difficult picking as different phases cross each other and to the model description which might not be accurate enough for ther local variations often represented by source and receivers statics. During this work, we have not used these quantities known to be quite efficient for describing features under the mesh size (Kissling et al., 1994).

One way of assessing the resolution variation could be through a checker-board test applied to the final best resolved model. As shown by the figure 27, different scales are better reconstructed by wavelet decomposition than by slowness decomposition. The checker-board over the smoothest patch with a characteristic length of 2 km provides similar results when considering slowness or wavelet decompositions. Other scales are significantly better recovered using the wavelet decomposition, showing us the interest of this decomposition.

8   Conclusion

Seismic tomography of first-arrival delayed times could be performed using a wavelet decomposition of the model space. By doing so, one is able to control locally the expected resolution of the inversion. Using the linearized procedure does not automatically inject slowness anomalies at the appropriate scale as the resolution vary during the linearized iterations. Consequently, one must introduce an additional operator in order to prevent spatial aliasing which could lock the inversion procedure. The introduction of an binary operator

over scale grids will forbid wavelet coefficients perturbation at a scale when we do not expect resolution. This operator M is a local operator: it will allow us reconstruction of various features as detailed ones nearby the free surface simultaneously with coarse ones at depth for a surface-surface geometry. An illustration on a synthetic crosswell experiment should similar features.

Improvement of data misfits is not significant and may be quite below the noise level which leaves us without clear quality control criterion. We are left with a strategy based on the model space representation and the wavelet decomposition leads to similar results as a local covariance analysis.

Application to a real data set has been performed and shows that one can indeed perform seismic tomography using wavelet formulation over a quite important industrial 2D line.

Of course, the attention is put back into picking strategy as we have found that industrial softwares might misdetect picked times and one has to find alternative strategies for better picking as we have now accurate tools for better reconstruction of detailed features through travel-time tomography.

Extension to 3D configuration should be performed and combining this wavelet decomposition and the finite frequency tomography (Montelli et al., 2004) will certainly improve in the future local resolution of seismic imaging before tackling amplitude analysis of seismic traces (Pratt et al., 1998, Operto et al., 2004).

References

Aki, K. & Lee, W., 1976. Determination of three-dimentional velocity anomalies under a seismic array using first P-arrival times from local earthquakes, *J. Geophys. Res.*, **81**, 4,381–4,399.

Benz, H. M., Chouet, S. A., Dawson, P. B., Lahr, J. C., Page, R. A., & Hole, J. A., 1996. Three-dimentional P and S wave velocity structure of Redoubt Volcano, Alaska, *J. Geophys. Res.*, **101**, 8111–8128.

Bohm, G., Galuppo, P., & Vesnaver, A., 2000. 3d adaptative tomography using delauney triangles and voronóı polygons, *Geop. Prosp.*, **48**, 723–744.

Bourgeois, A., Bourget, M., Lailly, P., Poulet, M., Ricarte, P., & Versteeg, R., 1991. Marmousi, model and data, in *The Marmousi Experience*, pp. 5–16, Eur. Ass. Expl. Geophys.

Chiao, L.-Y. & Kuo, B.-Y., 2001. Multiscale seismic tomography, *Geophys. J. Int.*, **145**, 517–527.

Chiao, L.-Y. & Liang, W.-T., 2003. Multiresolution parametrization for geophysical inverse problems, *Geophysics*, **68**, 199–209.

Cohen, A., Daubechies, I., & Feauveau, J., 1992. Biorthogonal basis of compactly supported wavelets, *Comm. Pure Appl. Math.*, **45**, 485–560.

Dahlen, F., 2004. Resolution limit of traveltime tomography, *Geophys. J. Int.*, **157**, 315–331.

Daubechie, I., Grossmann, A., & Meyer, Y., 1986. Painless nonorthogonal expansions, *J. Math Phys.*, **27**, 1271–1283.

Daubechies, I., 1988. Orthonormal bases of compactly supported wavelets, *Comm. Pure Appl. Math.*, **41**, 909–996.

Delprat-Jannaud, F. & Lailly, P., 1993. Ill posed and well posed formulation of the reflection traveltime tomography problem, *J. Geophys. Res.*, **98**, 6589–6605.

Farra, V. & Madariaga, R., 1988. Non-linear reflection tomography, *Geophysical Journal*, **95**, 135–147.

Grossmann, A. & J.Morlet, 1984. Decomposition of hardy functions into square integrable wavelets of constant shape, *SIAM, J. Math. Anal.*, **15**, 723–736.

Hole, J. A., 1992. Nonlinear high-resolution three-dimentional seismic travel time tomography, *J. Geophys. Res.*, **97**, 6553–6562.

Kissling, E., Ellsworth, W. L., Eberhart-Phillips, D., & Kradolfer, U., 1994. Initial reference models in local earthquake tomography, *J. Geophys. Res.*, **99**, 19,635–19,646.

Latorre, D., Virieux, J., Monfret, T., Monteiller, V., Vanorio, T., Got, J. L., & Lyon-Caen, H., 2004. A new seismic tomography of Aigion area (Gulf of Corinth, Greece) from the 1991 data set, *Geophysical Journal International*, **159**, 1013–1031.

Madych, W., 1999. Tomography, approximate reconstruction, and continuous wavelet transforms, *Applied and computational harmonic analysis*, **7**, 54–100.

Mallat, S., 1989. Multifrequency channel decompositions of images and wavelet models, *IEEE Trans. Acoust. Speech signal Process.*, **37**, 2091–2110.

Mallat, S., 1989. Multifrequency approximations and wavelet orthonormal bases of $l_2(\mathbf{r})$, *Trans. Amer. Math. Soc.*, **315**, 69–87.

Mallat, S., 1999. *A wavelet tour of signal processing*, Academic Press, New York.

Menke, W., 1984. *Geophysical Data Analysis: Discrete Inverse Theory*, Academic Press, Inc., Orlando, USA.

Meyer, Y., 1999. *Ondelettes et Opérateurs, I:Ondelettes, II:Opérateurs de Calderon-Zygmund, III:(with R. Coifman), Opérateurs multilinéaires*, Academic Press, New York.

Michelini, A., 1995. An adaptative-grid formalism for traveltime tomography, *Geophys. J. Int.*, **121**, 489–510.

Mintzer, F., 1985. Filter for distorsion-free multirate filter banks, *IEEE Trans Acoust. Speech Signal Process.*, **33**, 626–630.

Monteiller, V., Got, J.-L., Virieux, J., & Okubo, P., 2005. An efficient algorithm for double-difference tomography and location in heterogeneous media, with an application to kilauea volcano, *J. Geophys. Res.*, **110**, B12306, doi:10.1029/2004JB003466.

Montelli, R., Nolet, G., Dahlen, F. A., Masters, G., Engdahl, E. R., & Hung, S. H., 2004. Finite-frequency tomography reveals a variety of plumes in the mantle, *Science*, **303**, 338–343.

Nguyen, T. & Vaidyannathan, P., 1989. Two-channnel perfect reconstruction fir qmf structures which yield linear phase analysis ans synthesis filters, *IEEE Trans Acoust. Speech Signal Process.*, **37**, 676–690.

Nolet, G., Montelli, R., & Vireux, J., 1999. Explicit, approximate expressions for the resolution and *a posteriori* covariance of massive tomographic systems, *Geophys. J. Int.*, **138**, 36–44.

Operto, S., Ravaut, C., Improta, L., Virieux, J., Herrero, A., & Dell'Aversana, P., 2004. Quantitative imaging of complex structures from multi-fold wide aperture seismic data, *Geophysical Prospecting*, **52**, 625–651.

Paige, C. C. & Saunders, M. A., 1982. ALGORITHM 583 LSQR : Sparse linear equations and least squares problems, *ACM Trans. math. Soft.*, **8**(2), 195–209.

Podvin, P. & Lecomte, I., 1991. Finite difference computation of traveltimes in very contrasted velocity model : a massively parallel approach and its associated tools, *Geophys. J. Int.*, **105**, 271–284.

Pratt, R. G., Shin, C., & Hicks, G. J., 1998. Gauss-newton and full newton methods in frequency-space seismic waveform inversion, *Geophysical Journal International*, **133**, 341–362.

Spakman, W. & Bijwaard, H., 1998. Irregular cell parametrization of tomographic problems, *Ann. Geophys.*, **16**, 28.

Spakman, W. & Nolet, G., 1988. Imaging algorithms, accuracy and resolution, in *Mathematical geophysics*, edited by N. Vlaar, pp. 155–187, D. Reidel Publ. Co., Dordrecht, Netherlands.

Spakman, W., Bijwaard, H., & Engdahl, E., 1998. Closing the gap between regional and global travel time tomography, *J. Geophys. Res.*, **103**, 30055–30078.

Sweldens, W., 1994. The lifting scheme, a custom design construction of biorthogonal wavelets, Tech. rep., Technical Report 1994:7, Indust. Math. Init., Dep. of Math., Univ. South Carol.

Sweldens, W., 1997. The lifting scheme, a construction of second generation wavelets, *SIAM J. Math. Anal.*, **29**, 511–546.

Sweldens, W. & Schroder, P., 1995. Buidling your own wavelets at home, Tech. rep., Technical Report 1995:5, Indust. Math. Init., Dep. of Math., Univ. South Carol.

Tarantola, A., 1987. *Inverse problem theory: methods for data fitting and model parameter estimation*, Elsevier, New York.

Tarantola, A. & Valette, B., 1982. Generalized nonlinear inverse problems solved using the least square criterion, *Reviews of Geophys. and Space Phys.*, **20**, 219–232.

Tikhonov, A. & Arsenin, V., 1977. *Solution of ill-posed problems*, Winston, Washington, DC.

Trinks, I., Singh, S., Chapman, C., Barton, P., Bosch, M., & Cherrett, A., 2005. Adaptive traveltime tomography of densely sampled seismic data, *Geophys. J. Int.*, **160**, 925–938.

Vaidyannathan, P., 1993. Theory an design of m-channel maximally decimated quadrature mirror filters with arbitrary m, having perfect reconstruction property, *SIAM J. Math. Anal.*, **30**, 141–162.

Versteeg, R., 1994. The marmousi experience: Velocity model determination on a synthetic complex data set, *The Leading Edge*, **13**, 927–936.

Vetterli, M., 1986. Filter banks allowing prefect reconstruction, *Signal Process.*, **10**, 219–244.

Vetterli, M. & Herley, C., 1989. Wavelet and filter banks : theory and design, *IEEE Trans Acoust. Speech Signal Process.*, **37**, 1057–1071.

Woods, J. & O'neil, S., 1986. Subband coding of images, *IEEE Trans Acoust. Speech Signal Process.*, **34**(5), 1278–1288.

Zelt, C. & Smith, R. B., 1992. Seismic traveltime inversion for 2-D crustal velocity structure, *Geophys. J. Int.*, **108**, 16–34.

Zhang, H. & Thurber, C., 2005. Adaptative mesh seismic tomography based on tetrahedral and voronoi diagrams: application to parkfield, california, *J. Geophys. Res.*, **110**, B04303,doi:10.1029/2004JB003186.

Zhang, J., ten Brink, U., & Toks'oz, M., 1998. Nonlinear refraction and reflection travel time tomography, *J. Geophys. Res.*, **103**, 29743–29757.

Zhou, H., 2003. Multiscale traveltime tomography, *Geophysics*, **68**, 1639–1649.
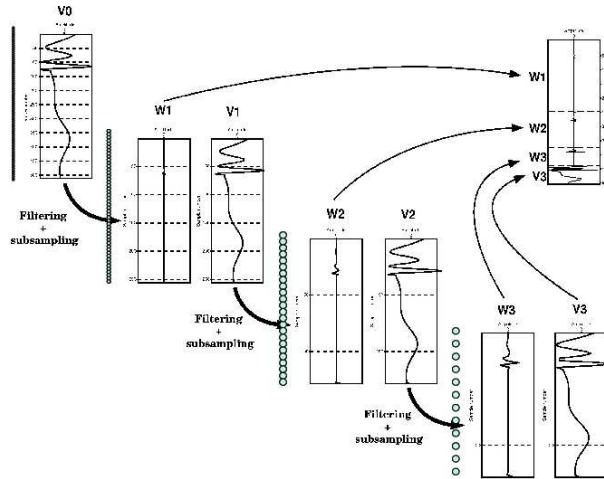
Figure 1: Projection of a 1D signal over 3 levels of wavelets
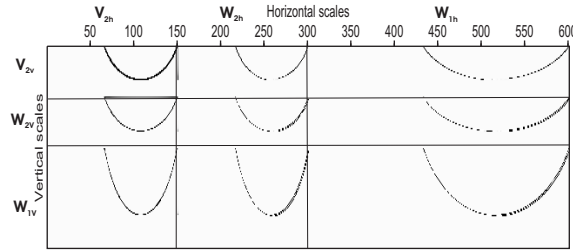
Figure 2: 2D Mallat representation of a 2D signal after 2 2D-wavelet transforms.Only the coarsest level of space V has been kept while wavelet coefficients and scale coefficients are deployed inside the same matrix as the one previously used for the slowness description.
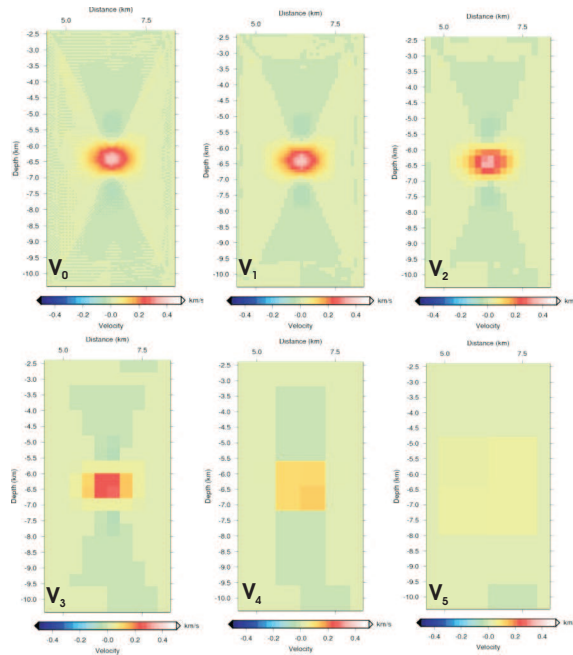
Figure 3: Decomposition of a V_0 slowness model for a crosswell experiment onto the V_i spaces using 5 levels of wavelets
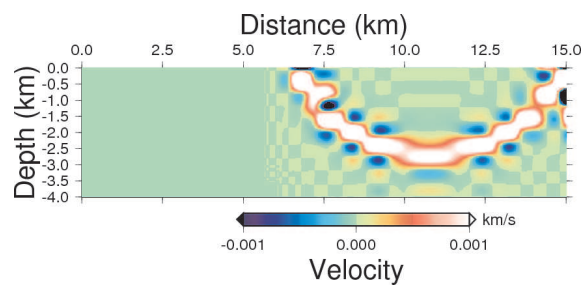
Figure 4: Non-zero Fréchet coefficients of a selected ray for the V_4 grid corresponding to the fourth level of wavelet decomposition.
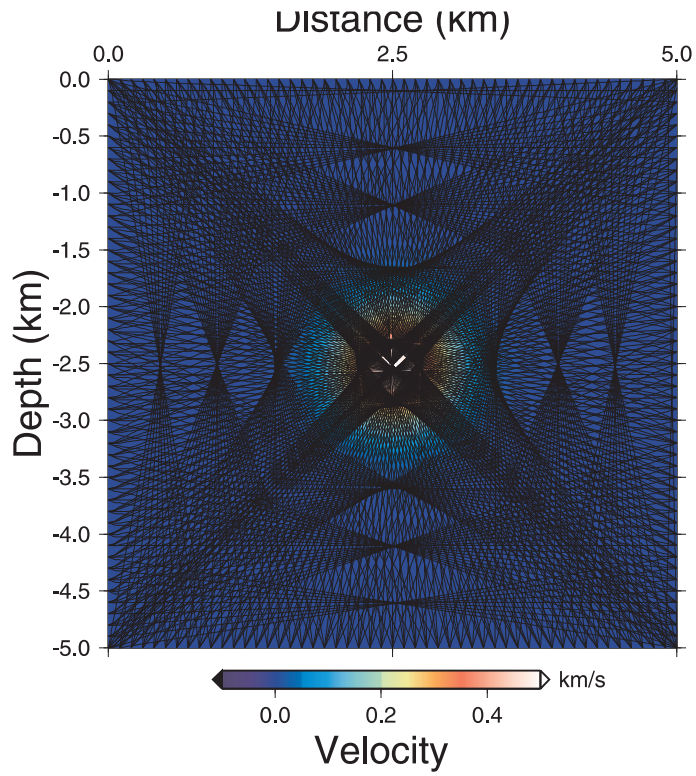
Figure 5: Synthetic positive anomaly to be reconstructed: the perturbation model and the ray coverage
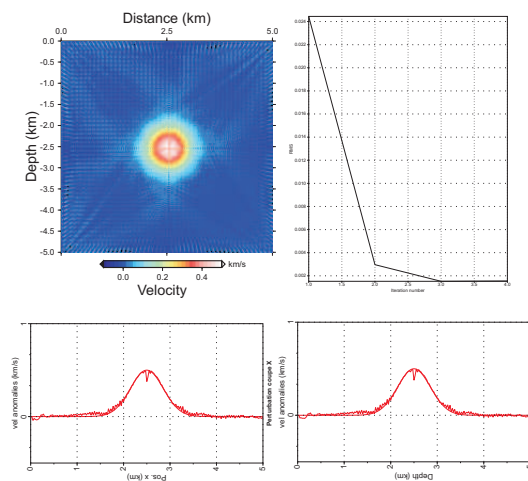
Figure 6: Tomographic inversion without any regularization or damping. On the top left panel, the reconstruction anomalu and the ray sampling. On the top right panel, the misfit behaviour leading to a minimal rms error of $1.52\ ms$. Four iterations are enough for convergence. Please note high frequency noise and the glitch along geometrical axes related to poor ray coverage
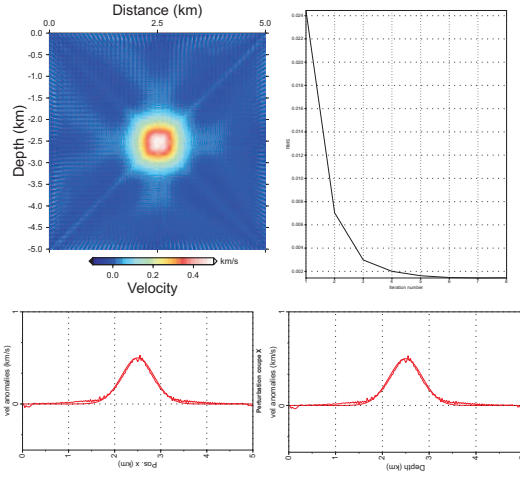
Figure 7: Tomographic inversion using wavelet with two vanishing moments and two levels. The convergence is obtained with an rms error of 1.43 ms after eight iterations. See figure 6 for other comments
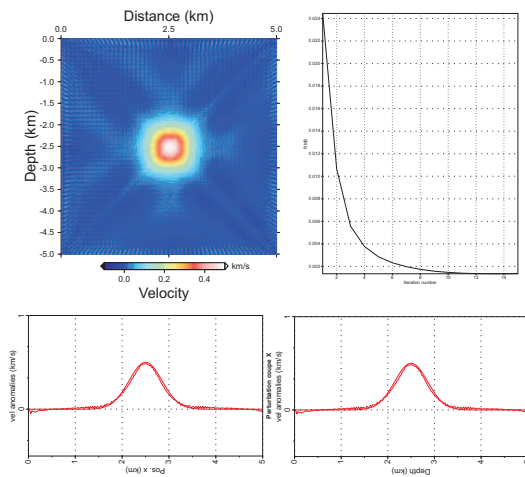

Figure 8: Tomographic inversion using wavelet with two vanishing moments and four levels. The convergence is obtained with an rms error of 1.35 ms after thirteen iterations. See figure 6 for other comments.
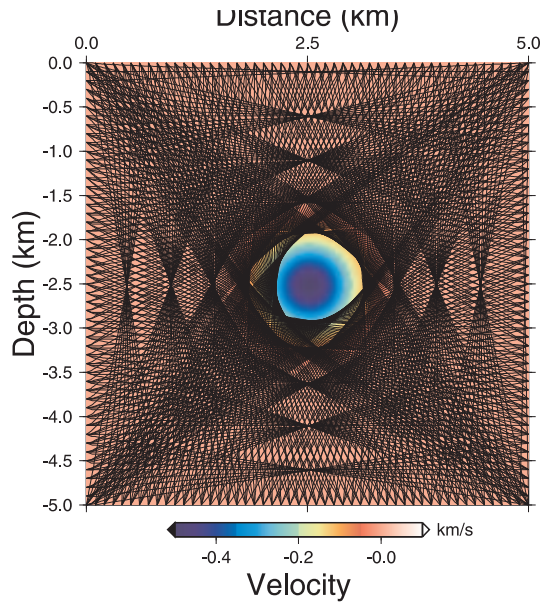
Figure 9: Synthetic example: a circular negative anomaly on top of an homogeneous model. Please, note the different ray coverage as for a positive anomaly essentially in the center of the box. The value should not be recovered in the center and only the parametrization makes this part of the model sensitive to travel times.
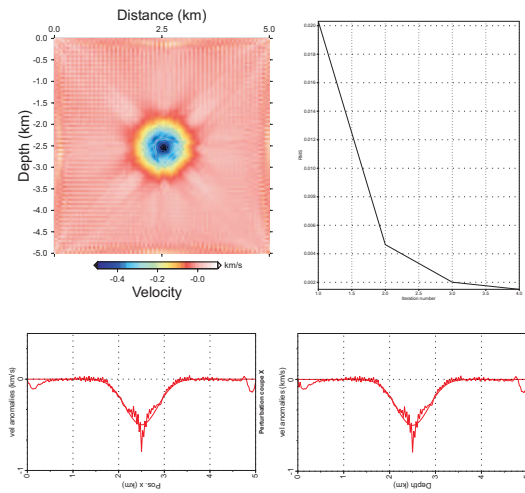
Figure 10: Tomographic inversion without any regularization or damping. On the top left panel, the reconstruction anomaly and the ray sampling. On the top right panel, the misfit behaviour leading to a minimal rms error of $1.50\ ms$. Four iterations are enough for convergence. Please note high frequency noise and the glitch along geometrical axes related to poor ray coverage as well as edge effects for similar reasons.
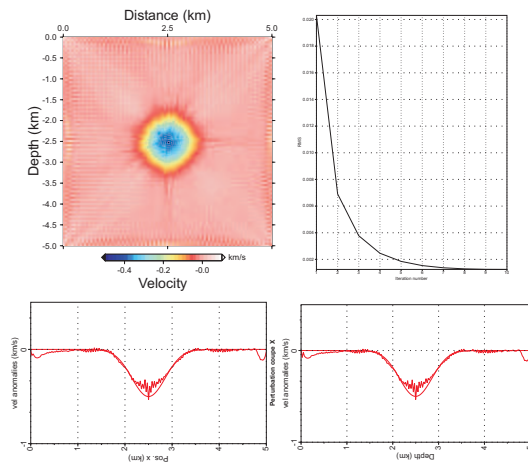
Figure 11: Tomographic inversion using wavelet with two vanishing moments and two levels. The convergence is obtained with an rms error of $1.24\ ms$ after ten iterations. See figure 10 for other comments

24
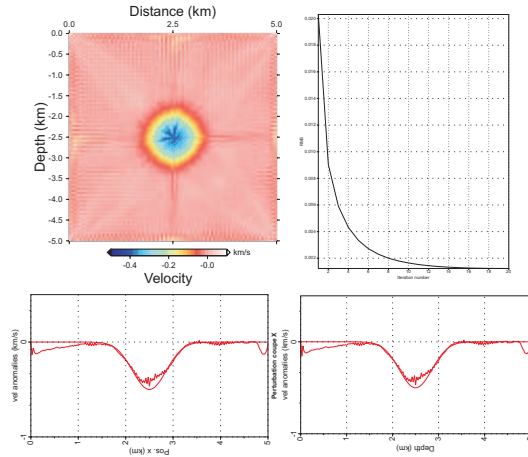
Figure 12: Tomographic inversion using wavelet with two vanishing moments and four levels. The convergence is obtained with an rms error of 1.23 ms after twenty iterations. See figure 10 for other comments
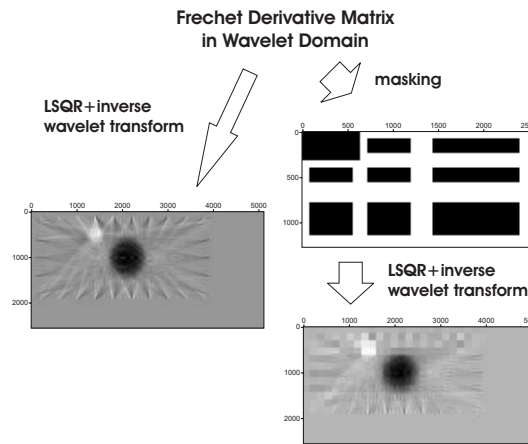
Figure 13: Application of masks for a better introduction of scales in the inversion. On the left panel is performed the inversion without the mask while we have applied the mask (top right panel) before performing the inversion (bottom right panel). The mask is a simple 2D $n_1 \times n_2$ matrix with zeroes (white) where Fréchet derivatives are cancelled and with ones (black) where Fréchet derivatives are preserved. Please note that the coarsest level of wavelet coefficients have been kept.
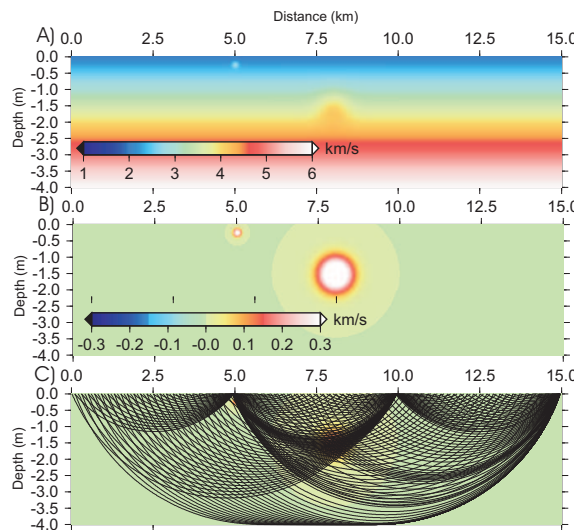
Figure 14: The synthetic model with a gradient in velocity and two circular anomalies on top panel (one is positive and the other negative). The model panel shows only these two anomalies with a central amplitude of 0.5 km/s while the bottom panel displays the ray coverage in the true synthetic model.

25
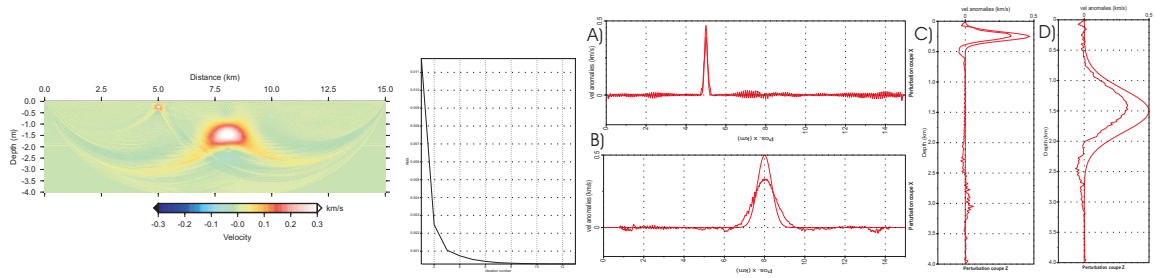
Figure 15: Reconstruction in the slowness domain whithout any regularization. Please note the ray sampling imprint into the reconstructed image as we want high resolution for the reconstruction of the small superficial anomaly
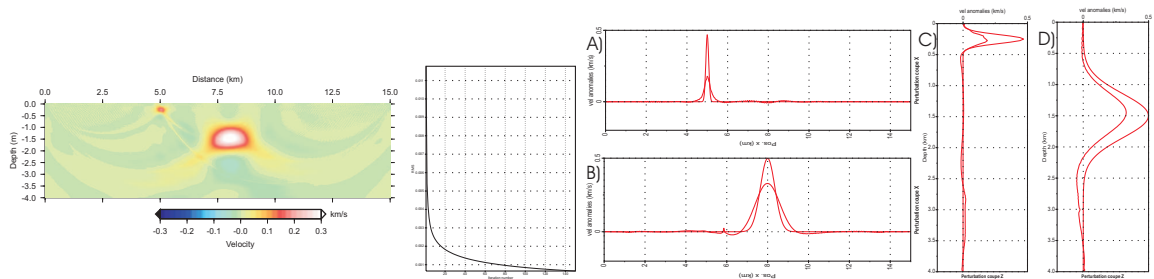
Figure 16: Gaussian regularization inversion: correlation length is taken as the value of 500 m. The deepest heterogeneity is smoother than in the figure 15, but the shallower one registers a loss of high frequencies and thus is not well reconstructed.
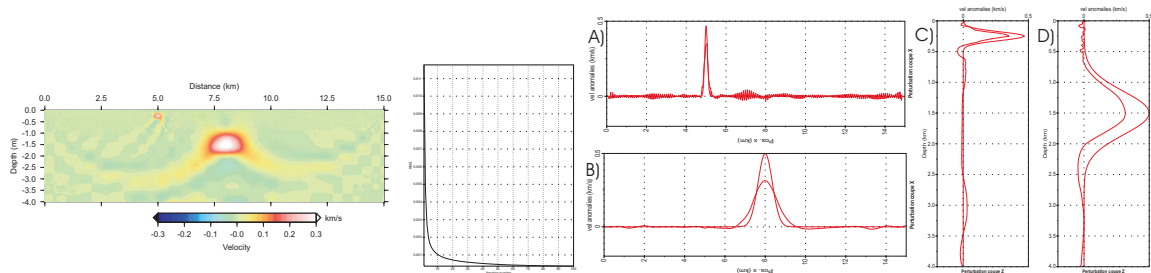
Figure 17: Reconstruction using wavelets with four vanishing moments and four levels: a mask ensures that, under the depth of 800 m, the fourth level of wavelet is the only one kept corresponding to a resolution of 400 m. The large heterogeneity is kept smooth and the shallow heterogeneity is well reconstructed.
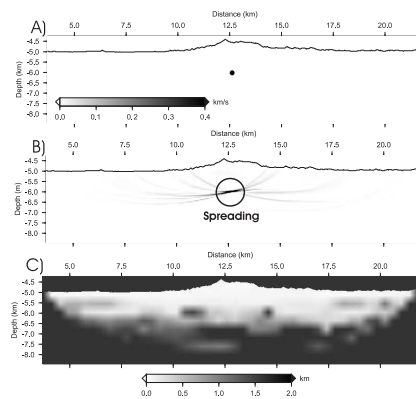
Figure 18: Resolution analysis: top panel displays the spike anomaly we want to reconstruct, the medium panel the spreading of the amplitude reducing it at the center and the bottom panel the mask operator as the percentage of the recovered amplitude.
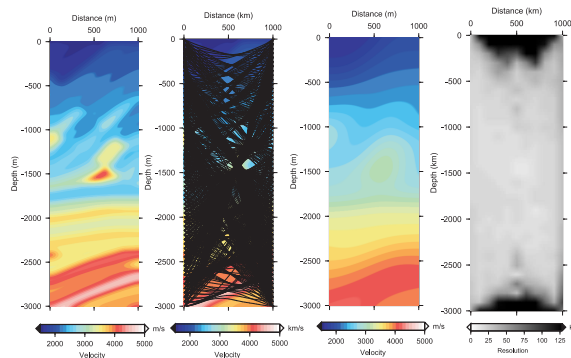
Figure 19: Resolution analysis: both two panels on the left display the true model between two hypothetical wells sampling this zone of the Marmousi model with quite dense ray coverage, the third panel shows the initial smooth model on which the resolution analysis is performed while the fourth panel shows resolution map as percentage of the recovered spike amplitude.

27
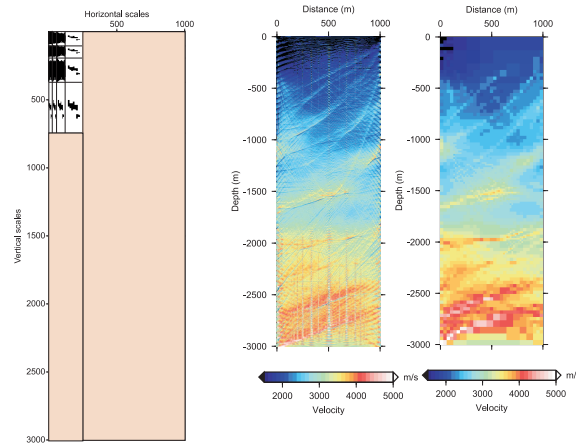
Figure 20: Tomographic inversion: the left panel displays the mask operator over the five scale levels from the finest one equivalent to the sampling of the synthetic Marmousi model to the coarsest one, the middle panel the inverted model using slowness coefficients without any damping and the right panel the inverted model using Haar wavelet coefficients over five scales. Please, note that cancelling the entire finest scale level does not harm the inversion procedure.
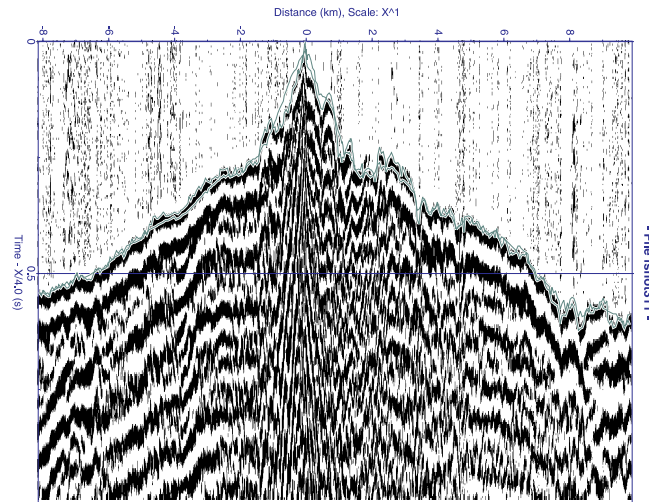

Figure 21: First-arrival picked time using an industrial software for the gather of shot 311; synthetic travel-times are displayed as well showing capacities of interpretation through the recovered velocity model
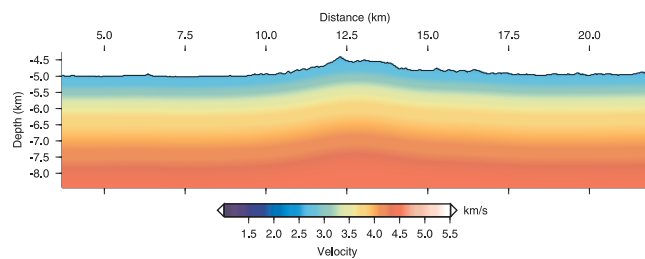
Figure 22: Initial 1D model for the travel-time inversion: this model starts exactly at the free surface following a complex topography.
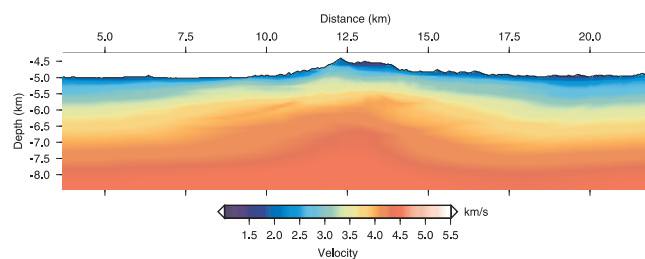

Figure 23: Reconstruction through slowness coefficients inversion: smooth gaussian regularization has been applied, preventing details to be exhibited as well as ray smearing features.
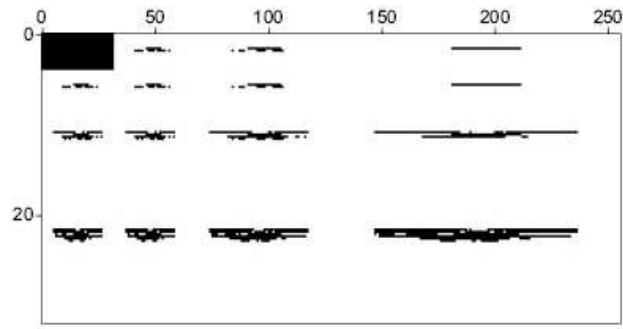
Figure 24: Mask construction over the already constructed smooth medium. Please note that the finest scales are only active nearby the free surface
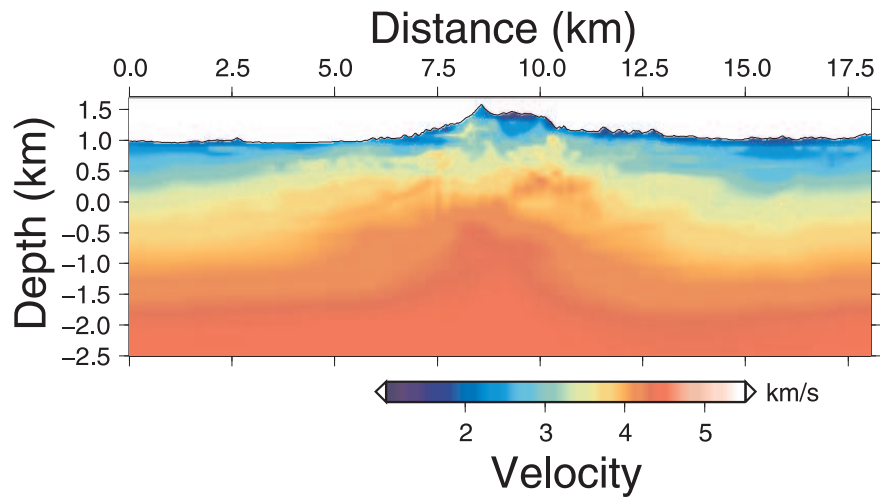

Figure 25: Reconstruction velocity model using wavelet decomposition: two vanishing moments have been requested and the rms final error is 23 ms. Please note the finest details nearby the free surface when comparing with the model of the figure 23.
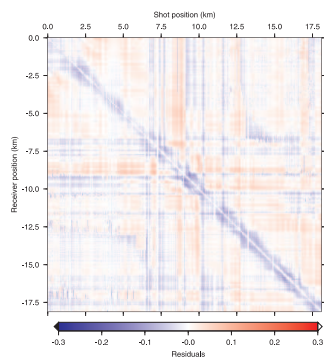

Figure 26: Final time residuals after thirty iterations. Please note the still too fast velocity structure nearby the free surface. Very local description of a low-velocity zone should be given and could be performed through statics evaluation.
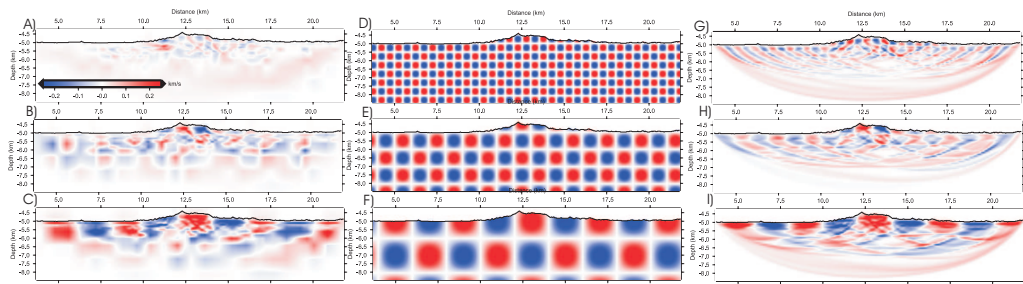
Figure 27: Checker-board test: the middle panel shows different wavelengths ranging from top to bottom as 500 m, 1 km and 2 km, the left panel is for slowness coefficients while the right panel is for wavelet coefficients. Please note the better behaviour through wavelet coefficients inversion.