



**HAL**  
open science

# Nonlinear Anisotropic Diffusion Filters for FWI: Structure Preserving Smoothing and Data Low Frequency Enhancement

Ludovic Métivier, Romain Brossier

► **To cite this version:**

Ludovic Métivier, Romain Brossier. Nonlinear Anisotropic Diffusion Filters for FWI: Structure Preserving Smoothing and Data Low Frequency Enhancement. 83rd EAGE Annual Conference & Exhibition, Jun 2022, Madrid, Spain. pp.1-5, 10.3997/2214-4609.202210272 . hal-03852673

**HAL Id: hal-03852673**

**<https://hal.science/hal-03852673>**

Submitted on 15 Nov 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Nonlinear anisotropic diffusion filters for FWI: structure preserving smoothing and data low frequency enhancement

L. Métivier<sup>1,2</sup>, R. Brossier<sup>2</sup>,

<sup>1</sup> CNRS, Univ. Grenoble Alpes, LJK, F-38058 Grenoble, France

<sup>2</sup> Univ. Grenoble Alpes, ISTerre, F-38058 Grenoble, France

Thursday 13<sup>th</sup> January, 2022

## Main objectives

This abstract presents how a nonlinear anisotropic diffusion filtering technique, originally designed in the field of image processing, can be used for high resolution seismic imaging by full waveform inversion. We show how the application of this technique to full waveform inversion can be done at two (potentially complementary) levels: as a structure preserving smoothing technique, in the model space, and as a low frequency enhancing filtering technique, in the data space.

## New aspects covered

1. Introducing nonlinear anisotropic diffusion filtering to the community of seismic imaging.
2. Showing how this can be applied as a structure preserving/enhancing smoothing technique in the model space, i.e. to the gradient in the FWI workflow, or to an estimated subsurface model, or to a seismic image obtained by migration.
3. Showing how this can be applied as a low frequency enhancing filtering technique in the data space for FWI applications.

## **Summary (200 words)**

We present how filters developed in the field of image processing can be efficiently applied to full waveform inversion, both in model and data space. These filters are known as Nonlinear Anisotropic Diffusion Filters (NADF). They rely on the solution of nonlinear diffusion partial differential equations (PDE), where the diffusion operator is a tensor. Basing this diffusion tensor on the image structure tensor, it is possible to design coherence enhancing filters, with the ability to smooth an input image only along identified structures, and not across them. This has a direct usage as a structure preserving smoothing operator, applicable to the gradient in full waveform inversion, with the advantage, over other structure preserving smoothing techniques, that the structure is detected automatically in the gradient, and does not depend on a prior information. NADF can also be used in the data space as a low-frequency enhancement technique. In this context, we show how the PDE formalism on which NADF are based makes it possible to introduce them as a general modification of the FWI misfit function, and how it can help mitigating cycle skipping in FWI. An illustration of both strategies is provided on a realistic synthetic example.

## Nonlinear anisotropic diffusion filters for FWI: structure preserving smoothing and data low frequency enhancement

### Introduction

High resolution seismic imaging by full waveform inversion (FWI) is based on an iterative data matching optimization process (Lailly, 1983; Tarantola, 1984). It is now the state of the art for the interpretation of seismic data both in industry at the exploration scale and in academy at global and (deep) crustal scales. An overview of FWI and its applications is given for instance in Virieux et al. (2017).

FWI is basically an ill-posed inverse problem which requires regularization techniques to reduce the size of the solution space and accelerate the convergence to the solution. It is conventional to enforce a certain degree of smoothness of the estimated subsurface models. This can be done in several ways, from Tikhonov techniques to gradient filtering. The latter is commonly adopted for practical applications: at each iteration of the optimization process, the gradient is filtered to remove spurious small scale oscillations. Isotropic non-stationary Gaussian filtering is often employed (Operto et al., 2006), with coherence lengths deduced from a prior estimation of the local resolution based on a diffraction tomography analysis (Devaney, 1984). More recently, structure preserving smoothing has been promoted (Trinh et al., 2017). A prior information on the structure is required, and the smoothing is designed so as to follow the main orientations of this structure, to smooth the gradient along those structure but not across them.

Another, well documented source of ill-posedness for the FWI problem is the cycle skipping issue. In the absence of sufficiently low frequency, FWI convergence strongly relies on the quality of the initial model. The misfit function is non-convex and exhibit local minima in which local optimization methods can be trapped. Many efforts have been devoted to mitigate this issue, either reformulating the FWI problem with additional degrees of freedom, or modifying the function measuring the distance between observed and calculated data to make the FWI problem more convex. Another (complementary) take on this problem is to enhance the low-frequency content of the data through low-frequency extrapolation techniques (Li and Demanet, 2016).

In this abstract, we present how a filtering technique developed in the field of image processing, named Nonlinear Anisotropic Diffusion Filter (NADF, Weickert, 1998) can be used efficiently as a structure preserving smoothing technique in the model space, as well as a low frequency enhancement technique in the data space. We first introduce the mathematical formalism of NADF, and how it links with the conventional FWI formalism. Then we present numerical applications of this technique to a realistic synthetic data set, both in the model and the data space. One advantage of NADF in the model space is that contrary to previously developed structure preserving smoothing technique, it does not require a prior information on the structure before it is applied, and rather learn the structure automatically from the gradient itself. In the data space, the underlying PDE formalism of NADF makes it possible to combine it efficiently with the general FWI formalism, and to couple it efficiently with any misfit function modification designed to mitigate cycle skipping.

### Theory

#### *Nonlinear anisotropic diffusion filters*

Applying an NADF is done by solving nonlinear anisotropic diffusion equations. Given an input 2D function  $f(x)$ ,  $x \in \mathbb{R}^2$ , a filtered version of  $f$  denoted by  $\mathcal{F}_T(f)$  is obtained as  $\mathcal{F}_T(f) = u(x, T)$  where  $u$  is the solution of

$$\begin{cases} \partial_t u - \operatorname{div}(D(u)\nabla u) = 0, \\ u(x, 0) = f(x), \quad \langle D(u)\nabla u, n \rangle = 0. \end{cases} \quad (1)$$

In equation (1) the operator  $D(u)$  is a 2-by-2 matrix named diffusion tensor. The use of such diffusion tensor makes it possible to control the diffusion rate depending on the direction (hence the reference to anisotropy). The nonlinearity of the tensor (dependence on  $u$ ) makes it possible to adapt locally (in space) and progressively (in time) the tensor during the diffusion process.

In this study we are interested in a coherence enhancing filter, for which the diffusion tensor  $D(u)$  is obtained as follows. We first consider a locally averaged structure tensor

$$J_\rho(u_\sigma) = K_\rho * (\nabla u_\sigma \nabla u_\sigma^T), \quad u_\sigma := K_\sigma * u, \quad (2)$$

where  $K_a$  is a Gaussian kernel with standard deviation  $a$  and  $*$  (resp  $*$ .) denotes a spatial convolution (respectively a term-by-term spatial convolution for matrices). The parameters  $\sigma$  and  $\rho$  are correlation

lengths introduced respectively 1) for  $\sigma$ : to remove spurious oscillations from  $u$  2) for  $\rho$ : to define a characteristic window size over which coherent structure are to be extracted. Considering the singular value decomposition  $J_\rho(u_\sigma) = P^T(u)\Sigma(u)P(u)$ , where

$$\Sigma(u) = \text{diag}(\mu_1, \mu_2), \quad \mu_1 \geq \mu_2 \geq 0, \quad (3)$$

the diffusion tensor  $D(u)$  is defined as  $D(u) = P^T(u)\Lambda(u)P(u)$  with

$$\Lambda(u) = \text{diag}(\alpha, h(\mu_1, \mu_2)) \quad (4)$$

where

$$h(\mu_1, \mu_2) = \begin{cases} \alpha & \text{if } \mu_1 = \mu_2, \\ \alpha + (1 - \alpha) \exp\left(\frac{-C}{(\mu_1 - \mu_2)^{2m}}\right) & \text{else.} \end{cases} \quad (5)$$

The eigenvalues of  $J_\rho(u_\sigma)$  describe the average local contrast in the eigendirections. Denoting by  $v_1$  and  $v_2$  the corresponding eigenvectors, we observe that  $v_1$  is the orientation with the highest value fluctuations, while  $v_2$  corresponds to the direction with the smallest variations, the ‘‘coherence’’ direction. The eigenvalues  $\mu_1$  and  $\mu_2$  can thus be used to describe locally the image structure:  $\mu_1 = \mu_2 = 0$  correspond to zones with constant values,  $\mu_1 \gg \mu_2 = 0$  corresponds to straight edges,  $\mu_1 \geq \mu_2 \gg 0$  correspond to corners, and  $(\mu_1 - \mu_2)^2$  is a measure of local coherence, which becomes large for anisotropic structures.

In this study we set  $\alpha = 10^{-5}$ ,  $C = 10^{-8}$  and  $m = 1$ . We see that the function  $h(\mu_1, \mu_2)$  increases rapidly to 1 as soon as the measure of local coherence  $(\mu_1 - \mu_2)^2$  departs from 0. The coherence enhancing filter thus works as follows: as soon as a feature with local coherence is detected, the diffusion rate increases in the coherence direction  $v_2$ , while it remains weak in the orthogonal direction  $v_1$ . When the coherence is weak, the diffusion gets back to an isotropic diffusion. We refer the interested reader to [Weickert \(1998\)](#) for details on interpretation, discretization and implementation of NADF for image processing.

#### Coupling with FWI

We consider the FWI problem

$$\min_m f(m) = G(d_{cal}[m], d_{obs}), \quad (6)$$

where  $d_{obs}$  correspond to the observed seismic data,  $m$  is the parameter representing the subsurface mechanical properties,  $d_{cal}[m]$  is the calculated data,  $G(.,.)$  is a general function measuring the distance between two input dataset and  $f(m)$  the misfit function to be minimized. FWI is based on local optimization schemes for which an initial estimation  $m_0$  is updated following descent direction  $\Delta m$  based on the misfit function gradient  $\nabla f(m)$ . The first application of NADF to FWI consists in filtering  $\nabla f(m)$  at each iteration of such scheme following the process described from equations (1) to (5).

The second application of NADF to FWI consists in applying an NADF to both  $d_{cal}[m]$  and  $d_{obs}$  so as to define a new FWI problem

$$\min_m f_T(m) = G(\mathcal{F}_T(d_{cal}[m]), \mathcal{F}_T(d_{obs})), \quad (7)$$

Following the adjoint state strategy ([Plessix, 2006](#)), we demonstrate in [Métivier and Brossier \(2022\)](#) that the gradient of the misfit function  $f_T(m)$  is obtained as the correlation of one incident and one adjoint field (usual for FWI), where the source term  $s$  for the adjoint field is computed as

$$s = \mathcal{F}_T^*(\partial d_{cal}G(\mathcal{F}_T(d_{cal}[m]), \mathcal{F}_T(d_{obs}))), \quad (8)$$

where  $\mathcal{F}_T^*$  is a linearized anisotropic diffusion operator for which the diffusion operator depends on the solution of the PDE underlying the application of the filter  $\mathcal{F}_T$  on the calculated data  $d_{cal}[m]$ . For the sake of compactness, we do not detail here the PDE associated with this linearized diffusion process, and we refer instead the interested reader to [Métivier and Brossier \(2022\)](#) where the complete derivation is presented.

#### Numerical examples

We illustrate how these two strategies apply to a realistic synthetic data set based on the Marmousi II P-wave velocity model. We build a set of observed data using fine grid (5 m) P-wave velocity (Fig. 3a) and density models. We consider a fixed-spread acquisition with 128 sources and 169 receivers at 50 m depth in the water column, regularly spaced each 132 m and 100 m respectively. A Ricker source centered on 5 Hz, and filtered so as to remove energy below 2.5 Hz, is used to generate the data. An additive Gaussian noise, filtered in the frequency range of the data (0 to 12.5 Hz) is introduced, with a signal to noise ratio equal to 10.

The inversion is performed on a coarse (25 m) grid, for the P-wave velocity model only. The initial

P-wave velocity model is a smoothed version of the true model (isotropic Gaussian smoothing with 500 m correlation length, Fig. 3b). An initial density model is computed through the Gardner's law, and an initial source wavelet is estimated from a linear deconvolution.

We first illustrate how NADF apply in the model space in Figure 1. We present the FWI gradient before and after filtering, considering an NADF and a conventional non-stationary Gaussian filter based on an estimation of the local resolution. The level set lines help visualizing how the NADF preserves the main structure present in the gradient, while the Gaussian filter smooth simultaneously all the gradient features, especially at depth where the increasing velocity increases automatically the expected resolution hence the correlation length of the filter. The  $(k_x, k_z)$  spectrum shows how the spread of the gradient in the wavenumber space is reduced by the application of the Gaussian filter, especially in the vertical direction, compared to the output of the NADF.

Second, we illustrate how NADF apply in the data space in Figure 2. The leftmost shot gather is considered before and after NADF are applied with increasing diffusion times  $T$ . We see that the NADF enhances the low frequency content of the data. However, the coherence enhancing feature makes it possible to avoid the noise becomes dominant as would be the case were a conventional linear filter would be applied, as SNR tends to 0 in the 0 – 2.5 Hz frequency band.

Finally, we illustrate how these two strategies can collaborate constructively on an inversion example. Starting from model presented in figure 3a, conventional least-squares FWI converges toward a local minimum (Fig. 3b). Using an optimal transport based misfit function (GSOT approach, Métivier et al. (2019)) the inversion result is improved, however the deep part of the velocity model is not correctly reconstructed (Fig. 3c). Low-frequency enhancement through NADF improves this issue (Fig. 3d). Finally, complementing the latter strategy with an NADF structure-preserving gradient filtering provides a velocity estimation very close from the exact model (Fig. 3e). In the latter case, the convergence is obtained in 100 iterations.

## Conclusion

We show how a filtering procedure from image processing based on a nonlinear anisotropic diffusion equation can be used for structure preserving gradient smoothing, as well as for low-frequency data enhancement in FWI. We present how this can be formalized in a mathematically consistent way, and illustrate how this can be beneficial for FWI on a realistic synthetic example based on the Marmousi II model. Perspectives from this work encompass application to migrated images and results from high resolution FWI, generalization to 3D, and application to field data.

## Acknowledgements

This study has been partially funded by the SEISCOPE consortium (<http://seiscope2.osug.fr>), sponsored by AKERBP, CGG, CHEVRON, EQUINOR, EXXON-MOBIL, JGI, SHELL, SINOPEC, SISPROBE and TOTAL. This study was granted access to the HPC resources of CIMENT infrastructure (<https://ciment.ujf-grenoble.fr>), Cray Marketing Partner Network (<https://partners.cray.com>) and CINES/IDRIS/TGCC under the allocation 046091 made by GENCI.

## References

- Devaney, A. [1984] Geophysical diffraction tomography. *Geoscience and Remote Sensing, IEEE Transactions on*, **GE-22**(1), 3–13.
- Lailly, P. [1983] The seismic inverse problem as a sequence of before stack migrations. In: Bednar, R. and Weglein (Eds.) *Conference on Inverse Scattering, Theory and application, Society for Industrial and Applied Mathematics, Philadelphia*. 206–220.
- Li, Y.E. and Demanet, L. [2016] Full-waveform inversion with extrapolated low-frequency data. *GEOPHYSICS*, **81**(6), R339–R348.
- Métivier, L. and Brossier, R. [2022] On the use of nonlinear anisotropic diffusion filters for seismic imaging using the full waveform. *Inverse Problems*, **submitted**.
- Métivier, L., Brossier, R., Méritot, Q. and Oudet, E. [2019] A graph space optimal transport distance as a generalization of  $L^p$  distances: application to a seismic imaging inverse problem. *Inverse Problems*, **35**(8), 085001.
- Operto, S., Virieux, J., Dessa, J.X. and Pascal, G. [2006] Crustal imaging from multifold ocean bottom seismometers data by frequency-domain full-waveform tomography: application to the eastern Nankai trough. *Journal of Geophysical Research*, **111**(B09306), doi:10.1029/2005JB003835.
- Plessix, R.E. [2006] A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International*, **167**(2), 495–503.
- Tarantola, A. [1984] Inversion of seismic reflection data in the acoustic approximation. *Geophysics*, **49**(8), 1259–1266.
- Trinh, P.T., Brossier, R., Métivier, L., Virieux, J. and Wellington, P. [2017] Bessel smoothing filter for spectral element mesh. *Geophysical Journal International*, **209**(3), 1489–1512.
- Virieux, J., Asnaashari, A., Brossier, R., Métivier, L., Ribodetti, A. and Zhou, W. [2017] An introduction to Full Waveform Inversion. In: Grechka, V. and Wapenaar, K. (Eds.) *Encyclopedia of Exploration Geophysics*, Society of Exploration Geophysics, R1–1–R1–40.
- Weickert, J. [1998] *Anisotropic diffusion in image processing*, Treubner Verlag, Stuttgart. Treubner Verlag.

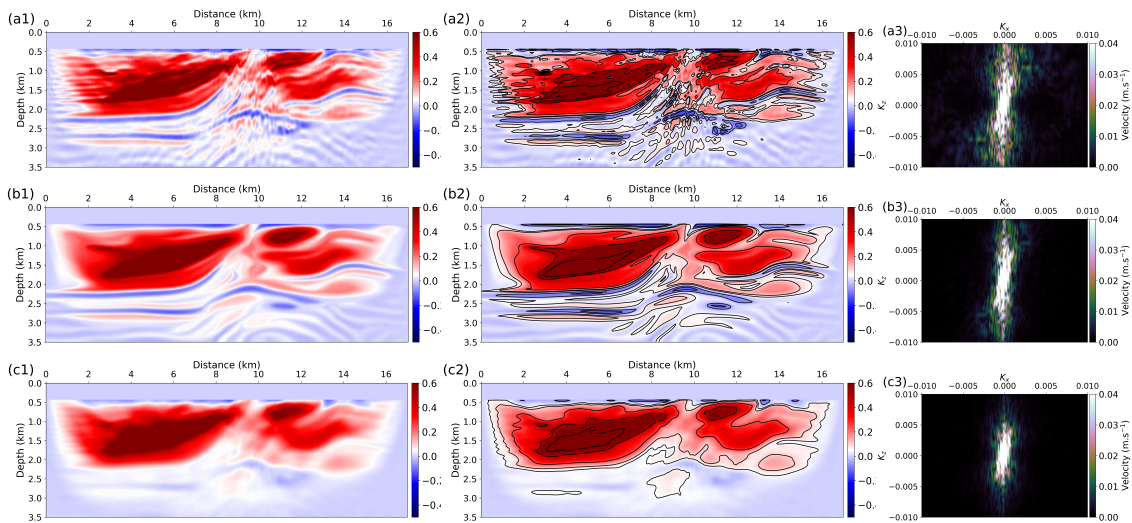


Figure 1: Gradient before filtering (a1-a3). Gradient after NADF (b1-b3), gradient after non-stationary isotropic Gaussian smoothing (c1-c3).

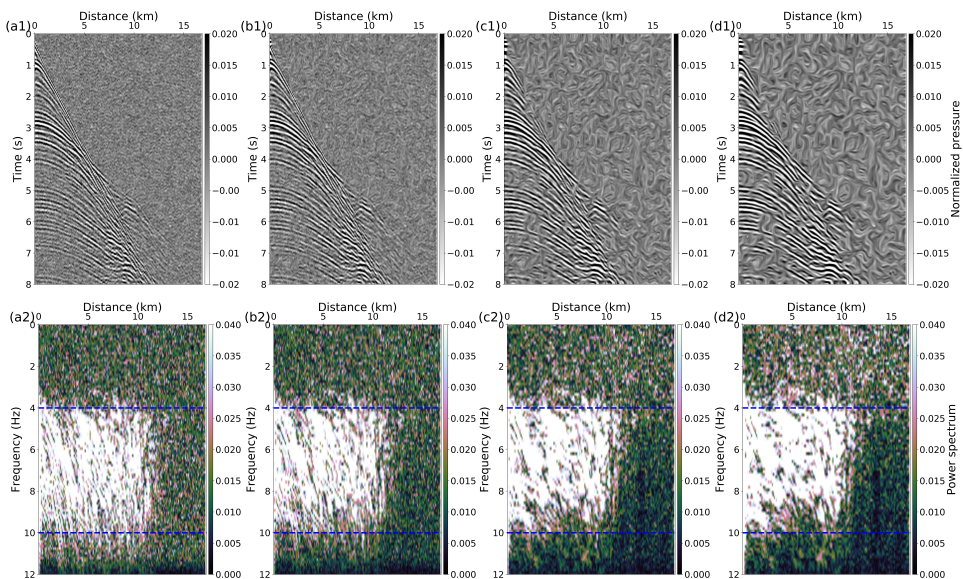


Figure 2: Shot gather before filtering (a1-a2), shot gather after NADF with  $T = 6.25$  (b1-b2), shot gather after NADF with  $T = 25$  (c1-c2), shot gather after NADF with  $T = 50$  (d1-d2).

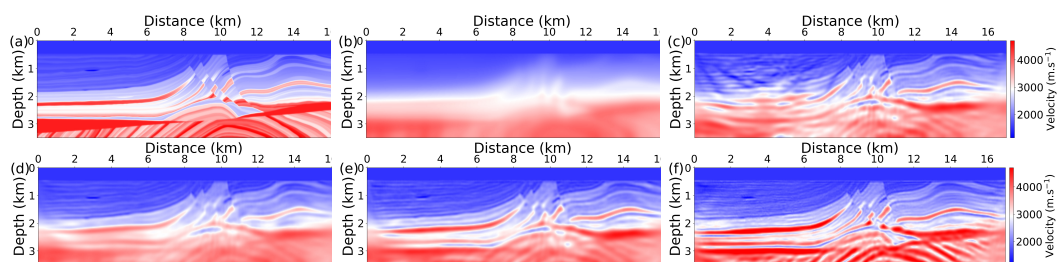


Figure 3: Exact fine grid model (a), initial model (b), final model with  $L^2$  FWI (c), final model with GSOT (Métivier et al., 2019) (d), final model with GSOT and NADF applied on the data (e), final model combining GSOT with NADF applied in data and model spaces (f).