LETTER TO REFEREE RC1

Manuscript "Effectiveness and computational efficiency of absorbing boundary conditions for full-waveform inversion",
by D. I. Dolci, F. A. G. Silva, P. S. Peixoto and E. V. Volpe,
submitted to Geoscientific Model Development

We would like to thank the reviewer for his/her comments on the manuscript. Our responses to the issues raised in the review are presented below.

- **High-level questions:**

    - Could you add a clear "Contributions" section to the paper that highlights (bullet points) what is being claimed as new in this paper?

      Response: Following the referee's suggestion, contributions of the current manuscript are highlighted in the introduction section.

      "In summary, the contributions of the current work are highlighted as follows.

        - *New implementations of ABCs in Devito, openly available to the scientific and industry communities;*
        - *Theoretical and numerical study of the effects that the ABCs may have upon the adjoint problem;*
        - *Detailed comparison of several widely used ABCs in FWI, which analyses both effectiveness and computational efficiency.*
        - *The proposition of a HABC approach based on the Higdon method for FWI, which has shown to be more effective, and computationally more efficient, than the well-known PML method.* "

    - Can you comment how the choice of velocity models affects the results presented? e.g. why does the SEG/EAGE model behave so differently from the others? Or why all ABCs behave so similarly in the adjoint stage?

      Response: The velocity model represents a medium where the wave propagates. Thus, the velocity field affects the angle of incidence at which the wave reaches the truncated boundaries. The ABCs have a good performance when the incidence angle is closer to the normal direction to the truncated boundary, but lose their effectiveness at glancing angles of incidence, i.e. closer to 90 degrees (Gao et al. (2015)). In principle, then, any particular model poses its own set of challenges to those techniques. Here, we consider widely used models, such as SEG/EAGE and Marmousi model, as examples of realistic models (Chi et al. (2014), Ben-Hadj-Ali et al. (2009), Su et al. (2020), Zhu et al. (2022), Buchatsky & Eran (2021)).

      As seen in this work, the self-adjoint wave equations were achieved in all ABCs methods. Also, the forward and adjoint systems have the same parameters (velocity model) and the boundary conditions. Under such circumstances, similarities should be expected in most cases (between forward and adjoint problems). However, we did observe significant differences between the ABCs' performance both in the forward and the adjoint solvers. We verified that CPML was less effective in the adjoint than in the forward problem, especially for the part of the Marmousi velocity model, where CPML errors were the highest for higher frequency peaks. That can be explained by the only significant difference between the forward and adjoint wave equations, which is the non-homogeneous terms. Thus, even with this difference, the HABCs and PML method kept their performance to attenuate the artificial reflections. For the different velocity models, and also in the forward and adjoint solvers, the common results showed that PML and HABC-Higdon have been more effective in attenuating reflections that originate from truncated boundaries. The novelty is that the HABC-Higdon requires less memory usage and wall-clock time than the well-known PML method.

Additional comments were inserted in the new version of the manuscript to clarify these questions.

Papers cited in this response:

Gao et al. (2015): Yingjie Gao, Hanjie Song, Jinhai Zhang, Zhenxing Yao. *Comparison of artificial absorbing boundaries for acoustic wave equation modelling.* Exploration Geophysics, 2015.

Chi et al. (2014): Chi, Benxin, Liangguo Dong, and Yuzhu Liu. "Full waveform inversion method using envelope objective function without low frequency data." *Journal of Applied Geophysics* 109, 2014.

Su et al. (2020): Sun, Hongyu, and Laurent Demanet. "Extrapolated full-waveform inversion with deep learningEFWI-CNN." *Geophysics* 85.3 (2020): R275-R288.

Zhu et al. (2022): Zhu, W., Xu, K., Darve, E., Biondi, B., & Beroza, G. C. (2022). Integrating deep neural networks with full-waveform inversion: Reparameterization, regularization, and uncertainty quantification. *Geophysics*, *87*(1), R93-R109.

Buchatsky & Eran (2021): Buchatsky, Sagi, and Eran Treister. "Full waveform inversion using extended and simultaneous sources." *SIAM Journal on Scientific Computing* 43.5 (2021): S862-S883.

○ Can you comment on how different physics (e.g. elastic/viscouacoustic) might affect your results?

Response: Thank you for pointing this out. We have added in the new version of the manuscript a paragraph explaining the generality of the results with respect to different physics. The following paragraph was included in the conclusions:

“Since the ABC formulations presented here are available for general wave equations (e.g. elastic/viscouacoustic), the methods can be applied for different physics problems of wave propagation. Komatitsch and Tromp (2003) verified that PML condition is efficient for both pressure (P) and shear (S) waves. In an anisotropic medium, DIMITRI (2007) showed that the CPML methods were efficient to absorb the quasi-pressure wave and the quasi-shear wave. The HABC-A1 and HABC-Higdon are based on A1 Clayton’s and Higdon conditions, respectively. Engquist and Majda (1977) and Higdon (1991) evaluated the effectiveness of the ABCs methods for P and S waves. So, while not shown here, the ABCs presented in this work should be able to attenuate the spurious reflections generated in the truncated boundaries for other physics problems. In addition, previous works (Gao et al., 2017; Engquist and Majda, 1977; Higdon, 1991) indicated that the angle of incidence in which the waves reach the truncated boundaries has more effect on the ABCs’ performance than the wave propagation properties. Therefore, we expect that, overall, our results should hold for different physics, as long as they still rely on wave propagation physics.”

<u>Papers cited in this response:</u>

Gao et al. (2015): Yingjie Gao, Hanjie Song, Jinhai Zhang, Zhenxing Yao. *Comparison of artificial absorbing boundaries for acoustic wave equation modelling.* Exploration Geophysics, 2015.

Komatitsch & Tromp (2003): Komatitsch, Dimitri, and Jeroen Tromp. "A perfectly matched layer absorbing boundary condition for the second-order seismic wave equation." *Geophysical Journal International* 154.1 (2003): 146-153.

Komatitsch & Roland (2007): Komatitsch, Dimitri, and Roland Martin. "An unsplit convolutional perfectly matched layer improved at grazing incidence for the seismic wave equation." *Geophysics* 72.5 (2007): SM155-SM167.

Clayton & Engquist (1977): Clayton, R. and Engquist, B.: Absorbing boundary conditions for acoustic and elastic wave equations, Bulletin of the seismological society of America, 67, 1529–1540, 1977.

Higdon (1991): Higdon, Robert L. "Absorbing boundary conditions for elastic waves." *Geophysics* 56.2 (1991): 231-241.

- Does the adjoint field include reflection errors from the forward propagation? i.e. was the d_sim used the reference one or the one that potentially had reflection errors from the forward prop?

  Response: The adjoint problem experiences its own spurious boundary reflections, as well as those from the forward solution, which are carried over into the adjoint solution via the external forcing term. The latter depends on the forward wave solution (***u)*** stored in the receivers, as can be verified in the adjoint equation shown in eq. 29. The new version of the manuscript now includes a comment at the end of section 6.2 to make sure this is clear to the reader.

- Are there limitations in the computational implementations (in your/Devito's code) that affect any of the results presented here? e.g. All boundary conditions augment the PDE. Are you solving the same augmented PDE in the entire domain (instead of just the boundary region)? If you could solve a simpler PDE in the physical domain and only solve the augmented PDE in the boundary regions, is it possible that the increase in computational time becomes negligible for all boundary conditions? Similarly for memory - for the fields added because of the ABCs, do you allocate memory over the entire domain?

  Response: As written in Lines 421 and 422:

  "An important resource in Devito structure is the possibility of partitioning the domain of interest into subdomains, to which distinct attributes can be assigned." Therefore, additional PDEs originating from the ABCs methods (as PML, CPML and HABC's) were solved only in the extended domain as exemplified in Figure 3, considering their particular structure. In the case of Damping, PML, and CPML methods the pure Acoustic Wave Equation was solved in the main domain, and the modified Acoustic Wave Equation, that is, the equation that includes auxiliary functions and/or damping functions, as well as the equations for the auxiliary functions, are solved only in the extension (i.e. excluding the main domain) where they really are required. In the case of HABC's methods, as we showed before, we solved the pure Acoustic Wave Equation over all the domains (including the main and extended domains), but the additional equations for the boundaries are solved only in the extension, to be later combined with the original solution

through a convex procedure, as was previously described. The possibility of creating a specific subdomain allows us to save time and memory for the computations. This is now clarified in the new version of the manuscript.

- If I understand correctly, the reference problems in Sections 6.1 and 6.2 are using a domain much bigger than the versions with ABCs, in order to get a version with no reflections.

  Response: The referee is correct. We have done that to ensure the reference field would be free of reflections from the truncated boundaries returning to the physical domain.

- If this same reference problem is used as a baseline for Section 6.3, this would be an unfair comparison. I don't think you're doing that.

  Response: They are indeed not exactly the same reference. In all cases, the errors were computed with respect to the reference field as is described in section 6 and shown in Figures 5 and 6 for the forward and adjoint problems, respectively. However, in Section 6.3 the growth of wall-clock time and the memory usage for different domain extensions are measured with respect to the reference wall-clock time and memory usage reached in the physical domain only, without any extensions, ensuring a fair comparison. Such a case is henceforth referred to as the no-ABC case.

- However, if the reference/baseline problem has changed between sections, could you please make that more clear. e.g. by giving it another name - computational reference problem/reference problem B.

  Response: To clarify the reader's understanding, we defined the reference field used to compute the numerical errors as the *accuracy reference. Computational reference* was used to refer to the benchmark used to evaluate the growth of the memory usage and wall-clock time of the cases that are subjected to the various ABC methods.

  In section 6, the fourth paragraph named the reference solution used in the error measures. The text is shown below:

  “The reference solutions used to evaluate the ABCs' effectiveness in attenuating the reflections is referred to here as accuracy reference and will be used to compute the quantitative error…”

  The second paragraph of section 6.3 describes the reference case used to measure the growth of memory usage and wall clock time. The text is shown below:

“Table 2 shows the average wall-clock runtime of simulations and the memory usage of a computational reference case, which used homogeneous Dirichlet boundary conditions (A3), and no ABCs were employed. That is, the tests were performed in the physical domain Ω0 only, without any extensions. Such a case is henceforth referred to as the no-ABC case, and it is used as a computational reference to evaluate the growth of the memory usage and wall-clock time of the cases that are subjected to the various ABC methods. “

- You are using subsampling in your gradient calculations for FWI. Could you justify the use of subsampling, as well as the chosen subsampling factors?

  Response: The subsampling approach was employed to avoid any problem with memory allocation. We could also employ the checkpointing approach that may save more memory usage than the subsampling with factor r=5. However, in accounting for the RAM available on the computer, we decided to use the subsampling approach, since the execution is faster in computational time than that required by the checkpointing. Regarding the sampling factor, we used the Nyquist criterion, which was attended for every Ricker peak of frequency (f=7Hz and f=15Hz) and the time-step of 0.001 (seconds). The time-step data was added to the text to make it clear that the sub-sampling attended the Nyquist criterion.

- Minor:

  - Line 7: DSL is written as DLS (also Line 600)

    Response: Corrected in the new version of the manuscript.

  - Line 8: Devito is primarily used for seismic modelling problems. Maybe not best to say it targets them.

    Response: Following the referee's suggestion, in Line 8 the text was modified from *“Devito, which targets seismic modeling problems*” to “*Devito, which is primarily used for seismic modelling problems*”.

  - Line 18: Psychical instead of physical

    Response: Corrected in the new version of the manuscript.

  - Line 47: Devito provides simple examples. Not appropriate to call them defaults.

    Response: Following the referee's suggestion, in Lines 47 and 48 the text was replaced from *“In Devito, Sochaki's type of Damping Boundary Layer is the default method to reduce the spurious reflections.”*

to *“In Devito, the seismic modeling examples have used Sochaki's type of Damping Boundary Layer method to reduce the spurious reflections.”*

- Line 208-209: Combination of sponge layer and ABC is ambiguous since ABC is defined here to encompass all methods discussed, and all methods discussed so far in the paper use sponge layers.

  Response: Indeed, the reviewer is right to point out an ambiguity. The term ABC is used in two senses, without proper distinction: the ABC can be thought of as a simple pointwise different boundary condition, such as done in A1 or Higdon's original propositions; or, the ABC can be thought of as a sponge layer, as in Damping, PML, and CPML, in which case the actual outer boundary condition is not relevant as no waves are expected to reach them, so null Dirichlet boundary conditions can be used. In the Hybrid approach, these 2 senses are merged, as pointwise (A1, Higdon) absorbing boundary conditions are imposed in the PDEs' boundaries alongside with a series of layers (which here could be called sponge layers). We have tried to avoid this ambiguity in the new version of the manuscript and re-written the terminology where we thought it was required.

- Line 269-270: The paper mixes the use of dampening and damping. Please double-check that it means what you want it to mean.

  Response: In Line 271 the word *dampening* was replaced by *damping, and we have checked the rest of the text to ensure a single use of the term.*

- Line 357: Missing reference

  Response: Corrected in the new version of the manuscript.

- Figure 9: b and c look blank. Are these supposed to be initialised at constant values of 2.5 and 3? Could you choose a colour scheme that makes this less ambiguous?

  Response: Indeed, b and c refer to constant initial models, and this is now clarified in the figure caption. We think that, despite being "blank", this figure is valuable as a way to illustrate the domain and value used.

- Line 412: underling -> underlying

  Response: Corrected in the new version of the manuscript.

- Line 437: Could you highlight which C compiler was used?
  Response: C compiler was GCC 8.3. This information was added in the new version of the manuscript.

LETTER TO REFEREE RC2

Manuscript "Effectiveness and computational efficiency of absorbing boundary conditions for full-waveform inversion",
by D. I. Dolci, F. A. G. Silva, P. S. Peixoto and E. V. Volpe,
submitted to Geoscientific Model Development

We would like to thank the reviewer for his/her comments on the manuscript. Our responses to the issues raised in the review are presented below.

Main points:

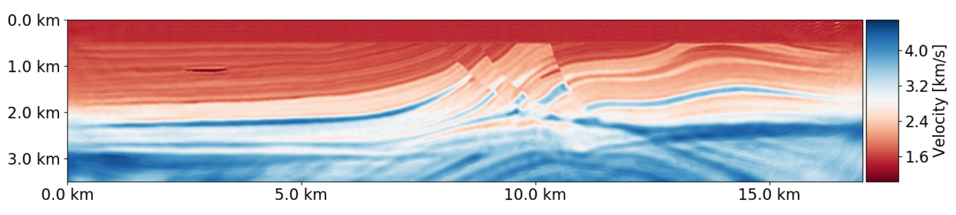
1. One of my main concerns is that, in my humble opinion, this paper in its current form contains too many equations (a total of 51). I understand that it is important and necessary to have many of the equations. On the other hand, I think the derivations of the ABCs of the adjoint form can be put into an appendix without damaging the main message of this paper.

Response: Following the referee's suggestion, the formulation of the adjoint equation is now put in Appendix A.

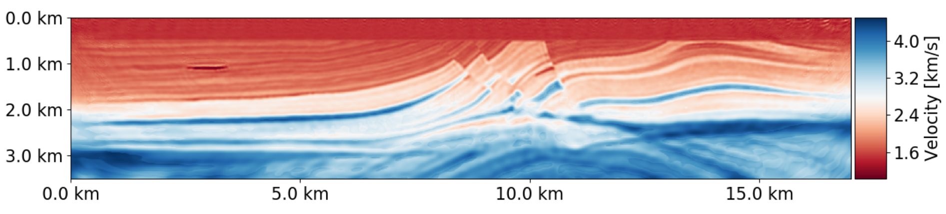
2. I would like to see FWI results for the Marmousi model for all the ABCs that this paper has compared.

Response: Computed models are shown as follows. The results used the settings presented in Table 7, in the manuscript text.

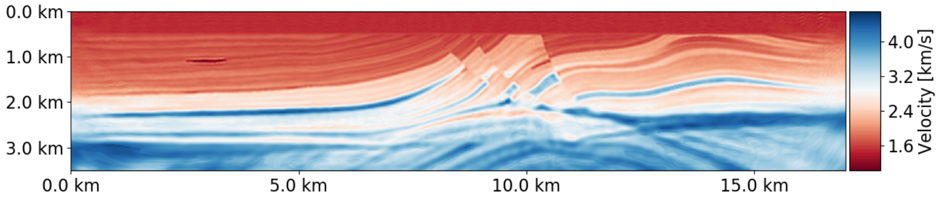
- PML



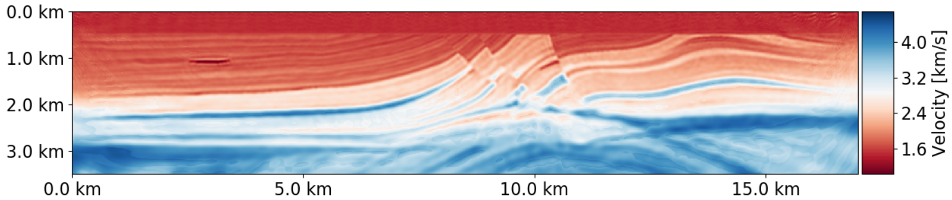
- CPML



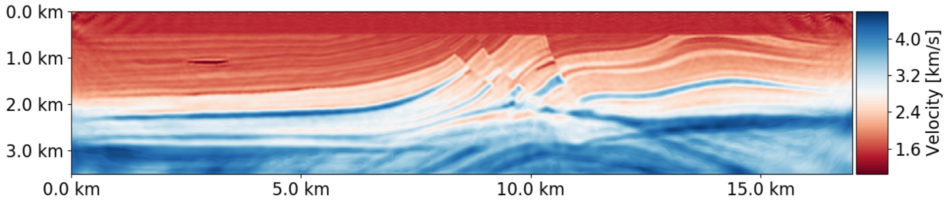
- HABC-Higdon



- HABC-A1



- Damping



The qualitative comparisons show that the computed velocity models are similar mainly when PML, CPML, and HABCs methods were applied. For the Damping method, one verifies the consequence of the spurious reflections closer to the corner of the free surface and the truncated boundary. Quantitative comparisons are presented below by evaluating the misfit values and the errors of the computed velocity models with respect to the true model, which is given by Marmousi.

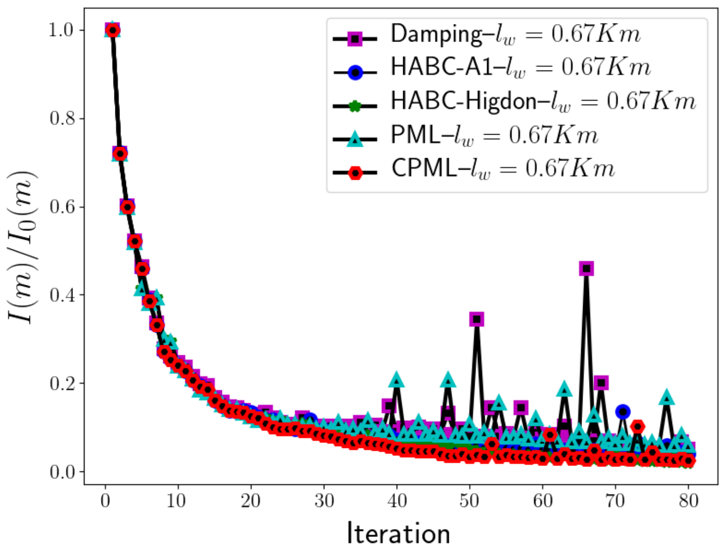


Figure above shows the misfit as given in eq. (1) with respect to the number of iteration. The misfit was normalized by the value at the first iteration, $I_0(m)$.

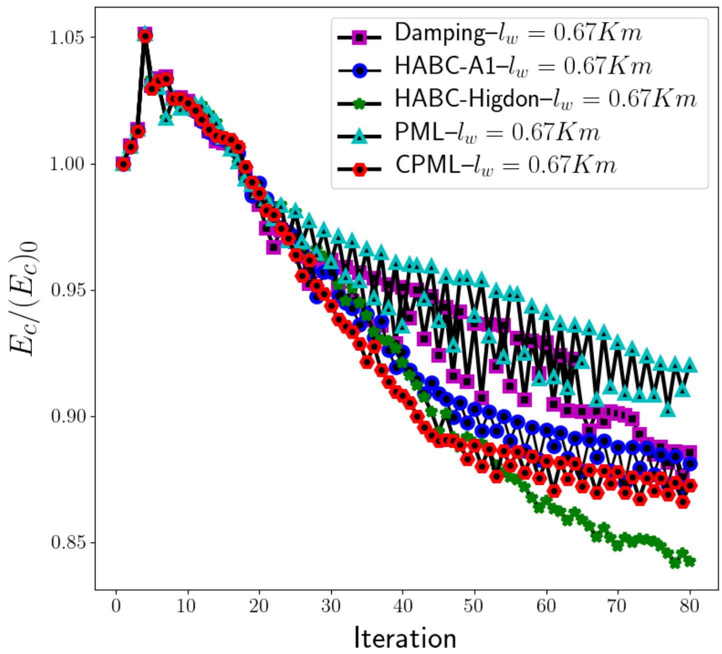


Figure above shows the velocity model error with respect to the number of iterations. The error was normalized by the value at the first iteration $(E_c)_0$. The error is computed by the expression

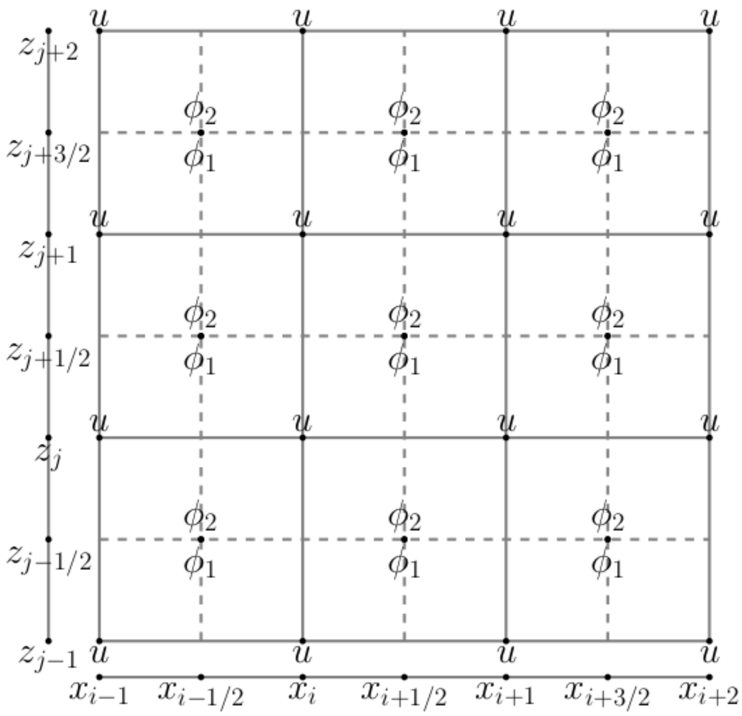
$$E_c = \frac{||c_{true} - c_{comp.}||_2}{||c_{true}||_2}$$

Where c_{true} is the Marmousi velocity model, c_{comp} is the velocity model computed by the FWI execution. To evaluate E_c , the velocity models c_{true} and c_{comp} were both defined in the same mesh setup. According to the error measures, it is verified that the error related to the HABC-Higdon is the smallest.

3. Line 168-178. The authors here defined staggered grid in the space and temporal domain. The staggered grid has been widely used for the numerical implementation of FWI, but can be tricky to understand for people without the background of seismic simulation. Can you please add a figure here showing the locations of the staggered grid for different variables in your equation?

Response:

The new version of the manuscript now has the following figure, and further information about the staggering used.



Here we have a representation about the variable staggering in the PML scheme. The auxiliary variables \phi_{1} and \phi_{2} are located in the "half positions" in both axes, while the variable of interest, u, is in the integer positions. Linear interpolations in time and space are used to infer values in shifted positions, as described in the manuscript.

4. Line 18, “psychical” Please correct this typo, I suppose it should be physical

Response: Corrected in the new version of the manuscript.

5. Line 27: “where the difference between the observed and synthetic data is back propagated in time from the receivers to the source of the waves”. The text here is related to the definition of the so-called “adjoint source” for the adjoint-state method. The adjoint source is not necessarily the difference between the observed data and synthetic data, but depending on what the misfit function is. This statement is only true when the misfit function is the norm-2 of the waveform difference.

6. Lines 28-29: “The back propagation requires saving the wave equation solution in every computational time step, thus meaning a high memory usage to solve a FWI problem. “ This statement is simply wrong. Any FWI software for a relatively-scale problem will not save the entire wavefield history in the memory, instead, an in-the-fly wavefield reconstruction technique is usually used.

Response: The referee is right regarding the points 2 and 3. Therefore, on account of the referee evaluation, the text was modified.

The text in Line 27 given by:

““where the difference between the observed and synthetic data is back propagated in time from the receivers to the source of the waves”

was replaced by the following text:

“for a least-square misfit function, the difference between the observed and synthetic data is back propagated in time from the receivers to the source of the waves.”

In addition, the text in Line 28-29 given by:

“The back propagation requires saving the wave equation solution in every computational time step, thus meaning a high memory usage to solve a FWI problem. “

was rewritten as:

“The back propagation requires saving the data of the wave equation solution, thus meaning a high memory usage to solve a FWI problem.”