The authors present an interesting study for comparing different types of absorbing boundary conditions (ABCs) for the numerical implementations of the full waveform inversion (FWI) problems. They have carefully derived the ABCs in their adjoint form, and have designed and performed numerical examples that demonstrated their conclusions. The paper is overall written, and should be of interest to a wide audience in the seismology community.

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.

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

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?
- Line 18, “psychical” Please correct this typo, I suppose it should be physical
- 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 being the norm-2 of the waveform difference.
- 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 relatively-scale problem will not save the entire wavefield history in the memory, instead, a in-the-fly wavefield reconstruction technique is usually used.