Dear Associate Editor and reviewers:

Thanks for all comments and suggestions to the paper. The manuscript has been through a number of modifications, and benefited a lot from your reviews. Here I first outline the major changes that I made to the paper, then I provide a detailed response to Associate Editor and reviewers' comments and questions. In general, the revised manuscript includes a more complete reference list, and the introduction part has been rewritten to provide more detailed background on previous works of Hessian; the abstract and conclusions have also be rewritten and a discussion section has been added; the notation $'$ has been replaced with $*$ to denote complex conjugation; some equations have been rewritten to avoid unnecessary confusions; examples of Hessian computed with more random realizations have been included to demonstrate how the crosstalk is suppressed by stacking more random realizations. A detailed response to Associate Editor and each reviewer follows bellow.

Yaxun Tang

Associate Editor

1. More references have been included to provide a more detailed background and honor previous work.

2. The title has been changed to "Wave-equation least-squares imaging with phase-encoded Hessian".

3. A Discussion section has been included.

4. The Abstract and Conclusion have both been rewritten.

5. Tables have been put at the end of the MS.

6. Space has been added between number and unit, please see the revised manuscript.

7. Most of the equations are given with identifying phrases.

8. Changes have been made with respect to the style for terms.

9. All ft has been converted to SI-unit

10. I have included the Sigsbee2A velocity model.

11. The comments indicated by Associate Editor in the annotated manuscript were also taken into consideration and included in the revised version.

Reviewer 1

1. There should be a $\frac{1}{2}$ in equation 4, I have corrected this in the revised manuscript. I agree with the reviewer that a proper preconditioner can speed up the convergence of the gradient-based optimization, I have to admit the statement in the original manuscript is not very precise, though what I meant was the simplest conjugate-gradient method without any preconditioning. I delete the corresponding sentence. Thanks for pointing this out to me.

2. The reviewer gave very good suggestions on the equations, I changed the way presenting the theory of phase encoding, some equations have been modified to avoid confusions, please see the revised manuscript.

3. The number of wavefield propagations for receiver-side random-phase encoding should be $(1 + N_{\text{realize}})N_s N_\omega$, I agree with the reviewer on this point, however, I disagree with the cost comparison in his comments. The number of wavefield propagations for data-space inversion should be $N_{it} \times 2 \times (2N_s N_\omega) = 4N_{it} N_s N_\omega$, instead of $N_{it} N_s N_\omega$, the first 2 comes from the fact that each iteration costs two migrations (one for computing the data residual, the other for model updating), the second 2 comes from the fact that we have to propagate both source and receiver wavefields for each migration. N_{it} is the number of iterations for data-space minimization. On the other hand, the number of wavefield propagation for the phase-encoded Hessian is outlined in Table 1 and Tabel 2 in the revised manuscript for different acquisition geometries. One thing I want to emphasize is that the explicit Hessian has to be computed ONLY ONCE, then we use the explicit Hessian to iteratively reconstruct the model parameters. So it is not correct to multiply a N_{it} factor in front of the cost of Hessian computation. This point has been sufficiently elaborated in the Least-squares Hessian section (below equation 7). To make it even more clear to the reviewer, the total cost of model-space inversion would be

$$C_{\text{total}} = C_{\text{migration}} + C_{\text{Hessian}} + C_{\text{filtering}}, \tag{1}$$

where $C_{\text{migration}}$ is the cost for performing one migration to get $\mathbf{m}_{\text{mig}}$, C_{Hessian} is the cost for computing the Hessian, while $C_{\text{filtering}}$ is the cost for solving the following system which has a much smaller size than the original system, because we are doing inversion in the model space in a target-oriented fashion:

$$J(\mathbf{m}) = \frac{1}{2}||\mathbf{Hm} - \mathbf{m}_{\text{mig}}||_2. \tag{2}$$

With the reduced size of the problem, the cost of iterative inverse filtering is very small compared to the cost of Hessian computation, thus

$$C_{\text{total}} \approx C_{\text{migration}} + C_{\text{Hessian}}, \tag{3}$$

and $C_{\text{Hessian}} \propto (1 + N_{\text{realize}})N_s N_\omega$ for the case of receiver-side random phase encoding, and $C_{\text{migration}} \propto 2N_s N_\omega$. Therefore, the ratio of data-space inversion and model-space inversion is about $r \propto \frac{4N_{it}}{3+N_{\text{realize}}}$. My experience with N_{it} is about 20 to 30 iterations with a proper preconditioner, and N_{realize} can be knocked down to 1 if there are many sources in an acquisition geometry, as demonstrated in the revised manuscript. Thus we obtain $r \propto 20 \sim 30$. Even with a relatively big N_{realize}, say 20, the ratio is $r \propto 3.5 \sim 5.2$, which suggests that the model-space inversion is more efficient.

4. The symbol $'$ has been replaced with $*$ to denote complex conjugation. I changed N'_x, N'_y, N'_z into M_x, M_y, M_z.

5. This is a very good suggestion by reviewer 1. Examples with more random realization for random phase encoding have been added to the paper, I think these examples sufficiently justify the reason for just choosing one realization.

6. I think comparing with the migration weights presented in Plessix and Mulder, 2004 would be very interesting. However, the starting point of this paper is to provide an efficient method to compute the Hessian, and then use it to deblur/invert the migrated image. The diagonal of the Hessian is just a byproduct of this process, probably it would be more appropriate to compare the performance of different diagonal approximations to the Hessian in a separate paper.

Reviewer 2

1. More references have been added in both Introduction section and Least-squares Hessian section, please see the revised manuscript.

2. Symbol $'$ has been replaced with $*$ to denote complex conjugation.

3. The misprint $\mathbf{s}_s$ has been corrected to $\mathbf{x}_s$; "aslo" has been changed to "also", thanks for pointing them out.

4. I keep the Cost comparison section, since I think it is very important to demonstrate how much savings in cost could be gained by using the phase-encoding method. Though I did not show any 3-D examples, I think it would also be useful to project the cost in 3-D.

5. Unfortunately, I can not show any 3-D examples at this point. However, the theory described in the paper does not prevent it from being applied to 3-D, since the algorithm for phase-encoded Hessian is essentially no different to the migration algorithm. Computing 3-D phase-encoded Hessian would be as easy as performing a 3-D wave-equation migration.

6. Color bar have been added to make better comparisons.

7. I don not think the resolution is decreased after inversion, it is very obvious to me that the inverted image is much sharper and crisper than the migrated image, especially the faults, this suggests a higher resolution is obtained in the

inverted image. The boosted noise is discussed in the Discussion section in the revised manuscript. It is not a surprise to me that inversion boosts up some noise, this is partially because of our incomplete acquisition geometry which makes the image lack of low frequency (or more precisely, low spatial wavenumber), therefore, when inversion iterates, it tend to oscillate the result and consequently adding more noise, that is why regularization is required to stabilize inversion; another possible source would be the randomized crosstalk presented in the phase-encoded Hessian. The latter one is beyond the scope of the paper and should be investigated more.

Reviewer 3

1. Thanks for pointing out that I missed the paper by Xie and Wu, 2006, I have included it in the revised manuscript and briefly discussed its difference with the approach presented in this paper.

2. I think the integral formulation of the Born modeling equation is very well known, i.e., equation 1, and the reference is also given as suggested by one of the reviewers, so I do not find it is necessary to add an additional graph to demonstrate this equation.

3. I elaborated more on how the composite wavefields work in the revised manuscript. A more detailed explanation with some math follows below. Take the receiver-side encoding for example, the composite source function is obtained by linearly combining point sources as follows, or equation 10 in the revised manuscript,

$$f_c(x, y, \mathbf{x}_s, \mathbf{p}_r, \omega) = \sum_{\mathbf{x}_r} w(\mathbf{x}_r, \mathbf{x}_s)\delta(x - x_r, y - y_r)\alpha(\mathbf{x}_r, \mathbf{p}_r, \omega). \tag{4}$$

   If one-way wave equation is used, the corresponding composite receiver wavefield $R(\mathbf{x}, \mathbf{x}_s, \mathbf{p}_r, \omega)$ is obtained by solving the following equations:

$$\begin{cases} \left(\frac{\partial}{\partial z} - i\sqrt{\omega^2 s^2(\mathbf{x}) + \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}}\right) R(\mathbf{x}, \mathbf{x}_s, \mathbf{p}_r, \omega) = 0 \\ R(x, y, z = 0, \mathbf{x}_s, \mathbf{p}_r, \omega) = f_c(x, y, \mathbf{x}_s, \mathbf{p}_r, \omega) \end{cases}, \tag{5}$$

   where $s(\mathbf{x})$ is the slowness, $f_c(x, y, \mathbf{x}_s, \mathbf{p}_r, \omega)$ is used as the boundary condition for the above one-way wave equation system. We can also express $R(\mathbf{x}, \mathbf{x}_s, \mathbf{p}_r, \omega)$ in terms of Green's functions as follows:

$$R(\mathbf{x}, \mathbf{x}_s, \mathbf{p}_r, \omega) = \sum_{\mathbf{x}_r} w(\mathbf{x}_r, \mathbf{x}_s)G(\mathbf{x}_r, \mathbf{x}, \omega)\alpha(\mathbf{x}_r, \mathbf{p}_r, \omega), \tag{6}$$

   which is equivalent to equation 11 in the revised manuscript.

4. The suggested references by Lecomte et al. have be cited in the revised manuscript.

5. Color bars have been added to some figures.

6. The comments indicated by reviewer 3 in the annotated manuscript were also taken into consideration and included in the revised version.