Referee's report on
WEIGHT-ADJUSTED DISCONTINUOUS GALERKIN METHODS: MATRIX-VALUED WEIGHTS AND ELASTIC WAVE PROPAGATION IN HETEROGENEOUS MEDIA
by JESSE CHAN

In the paper, the author applies the weight-adjusted discontinuous Galerkin (WADG) method, his previous work with others [1,2], to elastic wave propagation in heterogeneous media. WADG is a good method for elastic wave propagation in heterogeneous media as it concerned with developing a method that reduces runtime by reducing the amount of memory required for heterogeneous media. In addition to proposing a new method of elastic wave propagation, the author discusses energy stability and convergence of the semi-discrete method. The paper concludes with numerical examples that demonstrate the convergence and applicability of the proposed method.

In its whole, this paper has almost all of the hallmarks of a good SISC paper, new method, analysis, and numerical results. It is only missing performance results. I do think this is a good method for simulating elastic wave propagation in heterogeneous media, especially on GPUs where the amount of memory is limited. My major issue with this work is that it is iterative, as most of the results follow almost directly from the author's previous works, which do not seem to be published yet although can be found on the arXiv. It seems (as one might expect) the differences between acoustic wave propagation and elastic wave propagation are minor. My concern is compounded by the fact that there are many typos and the presentation is unclear in some places. Although it is iterative, I do believe it may warrant publication after the author addresses the issues below. For this reason I am recommending the paper be accepted with the minor revisions listed below.

First, I would suggest that the author proofreads the paper again as there are many typographical errors. Below contains a small sample of typographical errors and is not meant as an exhaustive list. I think the paper would be better if the author addresses the following concerns:

1. *typo on page 2:* "These equations can be written a first order. . . " should be something like "These equations can be written as a first order. . . ".

2. The paper begins with $\Omega \in \mathbb{R}^d$ and shows results for two and three dimensions. But in some places the author assume $d = 3$ without warning. For example, at the beginning of section 3, coordinates are give as if $d = 3$ yet the polynomials spaces later on the page are given both for $d = 2$ and $d = 3$. This issue comes up again when the multi-index notation is defined in section 4.1.

3. There is inconsistent use (or non use) of differentials in the paper. Sometimes differentials appear in integrals and sometimes not. Sometimes the variable is bold and sometimes it is not. I suggest the author picks a convention and sticks with it throughout the paper.

4. In the last equation on page 3 there seems to be an errand integral sign.
5. On page 4, the component wise definition of the jumps for matrices is missing an index.
6. The first set of equations in section 3.2 seem to be missing the sum over the elements.
7. The symbol Π_N is not defined, I assume it is the L^2 projection. Further, it would help, assuming this is correct, if a for all $v \in P^N(D^k)$ be added to the T_w^{-1} operator definition.
8. In the proof of theorem 4: "The proof is the similar" should probably be "The proof is similar".
9. When I checked the arXiv I didn't find a Theorem 3.1 in [1]. Should the reference for Theorem 3 be Theorem 4 in [1]?
10. In section 6.2.1, the boundary conditions used should be noted. Also a time of $T = 1$ seems rather short assuming periodic boundary conditions are used. Simulating for more periods would show another advantage of the current method.
11. I was disappointed not to see performance results for the three dimensional example to help support why someone would want to use the proposed method. It would be nice to see memory bandwidth and flop count measurements for the example in section 6.4.