Replies to reviewers' comments.

Reviewer #1:

The article presents an extension to the theory on moment conditions for finite differences to multipole point sources. The moment conditions are first derived in a continuous setting followed by the derivation of a discrete set of moment conditions. Further, a convergence theory for staggered grid finite-difference schemes on a family of first order systems of hyperbolic PDEs with multipole source terms is presented. The authors conclude the article by testing the convergence by numerical experiments on the acoustic wave equation.

In the introduction the authors claim that the presented work is an extension of the theory on point source discretization in finite differences. The moment conditions in the first section appears as a valuable extension to current theory, but the convergence analysis is restricted to a specific formulation of hyperbolic PDEs solved with staggered grid finite differences. My interpretation is that these results are not general for all stable finite difference discretisations. If the theory is more general than my interpretation, this must be stated more clearly. Otherwise, the reason for the restriction to staggered grid-finite differences should be explained.

I think the for the most part, the paper does a good job in specifying that the analysis and results are limited to staggered-grid FD schemes. Hopefully this is more apparent now that the theory section has been cleaned up. I added some extra sentences in the conclusion section to emphasize the integral components of the theory. The motivation for using staggered-grid methods was briefly mentioned in the intro (page 4, bottom paragraph). Not sure if it is worth mentioning again. Would it be a good idea to "staggered-grid" somewhere in the title?

In segments of the article, the line of argument is difficult to follow. For example, the subsection "Examples" on p. 23 appears out of context and the subsection is not well written. In the introduction of this subsection numerical convergence tests in the "next section" are mentioned but the next section does not contain such tests. Figure references are also missing in this section. This subsection must be rewritten and perhaps moved to the section "numerical tests and results".

This section has been cleaned up, and figure references added. See page 23.

The structure of the convergence theory is difficult to follow since the more general system which is actually analyzed (21) is mixed with an example (23). I think the example causes confusion about what is actually proven. I believe that the example would fit better in the section "numerical tests and results".

The convergence theory section has been reworked, starting page 24. All references to the acoustic case have been moved from the theory section to a subsection titled "Acoustic Case" at the end of the respective section, page 37.

The way of presenting the convergence result is non-standard. I suggest that you perform a convergence study with several grid refinements presented similar to figure 7 in Petterson et al. 2016. If if this is not added, a motivation of the current form of convergence test is needed.

Convergence rates are presented in this manner to demonstrate their spatial dependence, in particular, relative to source location in 2D space. For example, this presentation highlights the degradation in convergence rate far from the source location when using source approximations with the wrong approximation order. The convergence rates reported here are equivalent to the slopes of typical convergence figures in log-log plots.

In figure 6 (a) a green line which corresponds to third order convergence appears vertically in the figure. In figure 7 (a) one can also detect some green areas surrounding the source. Can you comment on this locally lower convergence rate?

In progress. I think the author mean figures 6(c) and 7(c). The green line in fig 6(c) coincides with the nodal plane of the dipole.

Reviewer #3:

Your analysis is restricted to the Cauchy problem (periodic case), but that is not stated anywhere (please state). Many applications require treatment of point sources on the boundary. In this case, the moment conditions need to be modified to account for the quadrature rule used by the difference scheme, as shown in Petersson 2016.

Boundary conditions are not required in the analysis, except implicitly in satisfying the skew-adjoint relations in the continuum and discrete systems. In the numerical examples, we use PML boundaries to simulate an unbounded domain, though we implement reflective boundary conditions at the computational domain, consistent with the theory. Source discretization near boundaries is not considered here. Not sure what could be added to the paper.

Before (16) on page 19, invervals -> intervals
On page 33 "corollary above", what corollary? do you mean theorem?
Page 36, did you mean $\tilde{u}$ and not $\tilde{p}$?

Fixed.

Page 37, does $\pm$ indicate addition and subtraction?

Yes. This was done to shorten the algebra in the proof.
Dr. Symes, should I make any changes to this?

You rely on $P_h = -P_h^T$ for your proofs, but this property is never stated (only for the continuous problem).

It is now mentioned explicitly in the presentation of the general continuum problem and its discretization by the skew-adjoint relations (23) and (25) respectively.

Page 38, How can you bound $\mathcal{e}_2$ using L^2-error estimates of finite difference solutions for smooth problems, when the solution to the problem is not smooth? At no point in your proof do you make use of the fact that you have a staggered finite difference discretization. However, if the discretization is replaced by the collocated grid, central finite discretizations (e.g., -1/2h, 0,1/2h), then smoothness conditions are necessary for convergence. To prove convergence without imposing smoothness conditions, I believe, at some point you have to invoke the property that the Nyquist mode $(-1)^j$ does not lie in the nullspace of the staggered grid difference operators.

We can apply L_2-error estimates to \mathbcal{E}_2 since \delta{\tilde u} and \delta{\tilde v} refer to the finite difference error for a problem with smooth sources (\tilde f, and \tilde g), i.e., see equation (35).

We did most of the heavy lifting for the proof of theorem 8 (weak convergence) a bit ahead of the actual theorem. May have contributed to some confusion for the reviewer?

Page 42, how do you define the minimum number of grid points per wavelength? Usually, the frequency at 5% of peak amplitude in the spectral domain is defined as f_{max} and then the minimum wavelength is estimated by $\lambda_{min} = c_{min}/f_{max}$, where c_{min} is the minimum wavespeed. For Ricker wavelets, one obtains $f_{max} = 2.5 f_p$ where f_p is the peak frequency.

Dr. Symes, not sure how to handle this? Would it be sufficient to reference a paper stating the rule of thumb of 10 grid points per wavelength for 2-2 and 5 grid points per wavelength for 2-4? If so, could you remind me of the reference?

Page 43, You state "... our narrow source approximations seem twice as wide and therefore smoother, from the point of view of the difference operator P_h. For this reason, smoothness conditions are not necessary for staggered-grid finite difference schemes". I do not follow this argument. Why would stencil width imply smoothness? As previously mentioned, smoothness conditions are not necessary for convergence when the Nyquist mode does not lie in the nullspace of the difference operator.

The stencil width and smoothness comment was referring to smoothness in Fourier space. For example, the Fourier transform of a wider the hat function is a narrower sinc^2 function, thus the contribution from spurious modes is diminished. Incidentally, source discretizations presented by Petersson 2016 that satisfy m-moment and s-smoothness conditions (with m=s) coincide with approximations that satisfy m-moment conditions and have two times the stencil width relative to the finite difference operator. These are the same source approximations used here.

Page 44, Instead of saying convergence theory "applicable to a large set of wave propagation problems", why not simply say "wave equations"? In other places you say "symmetric hyperbolic systems, which also can change to "wave equations" to be more specific.

Dr. Symes, any thoughts on this?