

1 We would like to begin by thanking the reviewers for their time, their careful reading of our submission, and for the  
2 valuable feedback they have provided. We are particularly excited that all four reviewers agreed that the most important  
3 aspect of the work – the codebase itself – seemed like a valuable contribution to the *NeurIPS community*: R1: “Overall,  
4 the library seems to be a great tool for a larger number of communities, and I’m particularly excited about getting  
5 to know more about its capabilities to integrate physical simulations and GNNs.”, R2: “...inserting itself in an area  
6 that has recently been the subject of interest and many works. This differentiation capability... is also of interest to  
7 many subfields in machine learning.”, R3: “The presented software library seems to be well-engineered, and I am sure  
8 that it can be adopted by the ML and physics communities and lead to potentially impactful research.”, and R4: “As a  
9 practitioner, I can see myself using this code to do research. It is easy to use as well as having an intuitive interface that  
10 makes it easy to build on top.”

11 **Suitability of Venue.** R1: “Not sure, if NeurIPS accepts papers about frameworks” and R3: “I would suggest the  
12 authors submit this paper to a venue specialized in ML software packages...” While we agree that software packages  
13 are less common submissions, the NeurIPS call-for-papers includes “3. Data, Competitions, Implementations, and  
14 *Software*: Benchmarks; Competitions or Challenges; Data Sets or Data Repositories; *Software Toolkits*.”

15 **Performance.** All of the reviewers commented that our performance lags behind traditional MD software; e.g. R2:  
16 “...it suffers from some drawbacks in comparison to similar (though non-differentiable) high-performance frameworks,  
17 especially, for example, when running large simulations on the CPU.” and R4: “As the authors mention, there are certain  
18 limitations regarding deploying it on TPUs or by directly competing with more established MD codes (e.g. LAMMPS).”  
19 As we mention in the paper, we agree wholeheartedly that there is a lot of room for improvement, especially on CPU  
20 and TPU. We are actively working to improve performance across the board. We agree with R3 that our messaging is  
21 inconsistent (R3: “...revise the paper and homogenize the way they are talking about the performance of JAX MD.”).  
22 We will change the language to say that JAX MD is “fast enough to effectively do research with.” Since JAX MD’s  
23 GPU performance is significantly faster than LAMMPS / HOOMD Blue CPU, we feel this is an accurate phrasing.

24 **Citing prior work.** R3: “The related works section’s first paragraphs reads like an introduction and could actually help  
25 improving your introduction section, if you move it there. In the rest of the related works section, you just have 1 (one!)  
26 related work cited which is DiffTaichi (Hu et al 2019)...” We disagree on this point. We discuss a range of related work  
27 in the first paragraph, including a number of fluid dynamics papers dating back to 2005. We wanted to spend more  
28 time discussing DiffTaichi since, as far as we are aware, it is the most closely related software to our own. If there are  
29 citations that we have missed, we are more than happy to add them.

30 **Writing.** R3’s main issue with the paper was R3: “The main weaknesses are the lack of technical depth (both in the  
31 description of the work and the results) and the paper content not being in good shape in terms of language used and the  
32 writing style. In my assessment the paper reads like a workshop paper at best.” Although phrased differently, R1 seems  
33 to share a similar concern, R1: “A bit too fluid, and less structured (the second half is “just a bunch of examples”)”.

34 Our goal in writing the paper was to be accessible and interesting to the broad and interdisciplinary audience of NeurIPS.  
35 Indeed, we received feedback that a previous version of the paper was overly technical. However, we definitely do not  
36 want to come across as “marketing” the work and we appreciate the reviewer for pointing out that it came across this  
37 way. We are still working to strike a balance; since most of the reviewers seemed to view the exposition favorably it  
38 seems that the writing needs to be tweaked rather than overhauled. We agree with R3’s specific suggestions here and  
39 we will make these changes to the manuscript to sharpen the language. We also agree with R3 that we should state what  
40 MD stands for at the start of the paper and be more precise about the range of physics simulations we consider.

41 One particular point on writing style: in cases where JAX MD was an improvement over previous MD packages we  
42 intentionally refrained from making a juxtaposition (e.g. questions like R3: “What does “seamlessly integrating” mean?  
43 What is the alternative when it is not “seamless”?”). We have strong responses to these questions (e.g. previously people  
44 wrote bridges between ML software, like TF, and the C / CUDA internals of existing MD systems; here features still  
45 need to be differentiated *by hand*), but we thought the *average* NeurIPS reader would get more benefit from learning  
46 about what they can do with JAX MD than from a comparison with software packages that they may not have heard of  
47 before. R3, do you think the paper would be stronger if we make these comparisons? In either case we can make our  
48 use of language more precise.

49 **Technical Depth.** Above, R3 suggested that the paper lacked “technical depth”. Here we disagree with the reviewer.  
50 The paper’s contribution is principally the software library, which we feel offers ample technical depth for a NeurIPS  
51 paper. Moreover, we have included three detailed examples that can be run interactively in colab. While these examples  
52 are pedagogical, they are very close to current research (e.g. the GNN we train was used in several papers published  
53 this year). Perhaps the structure of the paper obscured the technical depth and we would appreciate feedback. Did the  
54 referee notice the colab notebooks or should we emphasize them more? If you did notice them, did you not find the  
55 style of conveying content effective?