

1 We thank all reviewers for careful reviews and many positive comments, including **R1**: "a highly original manuscript  
2 which goes beyond using known algorithms to study technically challenging problems and instead attempts to explain  
3 mechanisms of fundamental neuro-systems", "...providing an important critical analysis for gaps in the neuroscience  
4 literature", "could have impact on future work both theoretical and applied"; **R3**: "original and interesting and I believe  
5 relevant to NeurIPS", "...clearly written. The figures are well-thought-out" **R4**: "take[s] [previous] efforts significantly  
6 forward", "the authors have done an admirable job". We now respond to the other reviewer feedback:

7 **Response to Reviewer 3:** We thank **R3** for pointing out a pair of previous works deriving grid cells from a normative  
8 perspective. These are nice papers and we will add them to the many normative grid cell papers we already cited. We  
9 note however significant differences in focus between these previous works and ours. For example Whittington et. al.  
10 2018 derived grid cells by learning basis functions for local transition operators, but this approach yielded square, not  
11 hexagonal grid cells. Mathis et. al. eLife 2015 demonstrated that the Fisher information (FI) of independent Poisson  
12 firing hexagonal grid cells is larger than that of square grid cells by a factor of 1.15 under certain assumptions about  
13 firing field width. However, they did not address at all the question of how a neural network could convert velocity  
14 signals into spatial signals, and under what conditions such a neural network would specifically use intermediate  
15 representations resembling hexagonal grid cells. They simply *assumed* the existence of hexagonal firing fields and  
16 computed their FI. It is this latter question about neural networks that is the primary focus of our work; we feel it  
17 is important because previous results exploring how neural networks can convert velocity signals to spatial signals,  
18 yielded conflicting results: i.e. our Ref. 10 yielded square grid cells, while our Ref. 11, published in Nature, yields  
19 very noisy hexagonal cells, but as we demonstrate in this paper, yield very robust and clean *square* grid cells if *one*  
20 parameter (batch size) is changed. Indeed this in and of itself is a new result in our paper. Our main contribution, which  
21 has never been done in any prior work, involves connecting neural network training to the theory of pattern formation,  
22 thereby providing a unifying framework to reconcile these conflicting results as well as yield, to our knowledge for the  
23 first time, a robust method for generating *hexagonal* grid cells in a square box as intermediate representations in neural  
24 networks that convert velocity signals to spatial signals. We do however, take **R3**'s excellent suggestions to heart, and  
25 we will expand our introduction to discuss more carefully, and in a more toned down manner, the diverse and interesting  
26 contributions of previous normative works and explain how our work relates to and is different from this previous  
27 work. We will also provide full training details and code to ensure re-reproducibility of our work.

28 **R3** also raises an interesting point that hexagonal grid cells have higher FI than square ones. We do not see appreciable  
29 differences in positional coding between our square and hexagonal grid cell networks. There are two good theoretical  
30 reasons to believe such differences, if any, should be small. First, based on considerations of Mathis et. al. eLife 2015  
31 the FI of *both* hexagonal and square grid cell networks is  $O(N)$  where  $N$  is the number of grid cells. Thus, for large  
32  $N$ , the spatial resolution (inverse square root of FI) of both codes will be much more similar to each other relative to  
33 the scale of either the box or the mouse itself (i.e compare  $O(1/\sqrt{N})$  and  $O(1/\sqrt{1.15N})$  resolution for square and  
34 hexagonal grids respectively to the  $O(1)$  spatial scale of the mouse. Second, the dominant source of inaccuracies in  
35 the estimate of position in path integrator neural networks arises from imperfections or noise in the mechanism that  
36 integrates velocity to construct position. For example, Hardcastle et. al. Neuron 2015 showed errors in decoded position  
37 from both mouse and rat entorhinal cortex rose rapidly with time since a border encounter (which can correct position  
38 estimates). Thus the FI based theory of Mathis et. al. eLife 2015 cannot address this dominant source of error, as it  
39 simply *assumes* the existence of perfect hexagonal firing fields, without ever considering additional errors arising from  
40 imperfections in the very neural network mechanisms required to construct such firing fields from velocity signals. So  
41 based on these two good reasons, our preliminary analysis that square and hexagonal grid cells perform similarly is not  
42 inconsistent with Mathis et. al. eLife 2015. However, before publication in NeurIPS, we will follow up much more  
43 carefully on **R3**'s very interesting suggestion by doing extensive analyses to detect  $O(1/\sqrt{N})$  differences in position  
44 estimation between square and hexagonal grid networks and we will report what we will find.

45 **Response to Reviewer 4:** We very much appreciate **R4**'s general concern about the field that grid cells may not be  
46 learned from environmental statistics but rather could be innate before spatial experience. In fact, our prejudice is  
47 with the latter viewpoint, in agreement with the reviewer. We emphasize, that this paper, as we have written it, is  
48 completely agnostic to this issue, as we do not view our backpropagation training as a model of the *learning* process in  
49 the actual mouse. Rather, backpropagation to us (as in many other works applying task based neural network training to  
50 neuroscience) is simply a method of exploring the space of neural network solutions to the problem of path integration  
51 (converting velocity to spatial representations) to achieve theoretical clarity on when hexagonal grid cell representations  
52 emerge. We find it exciting that they emerge very naturally, simply when firing rates are positive. We make no claims  
53 of course, that the mouse learns them this way. And indeed, given the prevalence of grid cell solutions found by our  
54 path integrator networks, our work suggests innate formation of grid cell representations before spatial experience may  
55 actually facilitate path integration during spatial experience.

56 **Response to Reviewer 2:** We thank you for uniformly positive comments and we will provide code upon publication.