We thank all the reviewers for their comments on our work as well as for their suggestions on how it can be improved. We especially want to thank reviewer 3 for their detailed and insightful review, and for highlighting specific and important shortcomings of our original submission. We hope that the following will address all major concerns raised.

Additional analysis on NeurComm (R1, R2, R3): Almost all the reviewers requested that we look at the data more closely and attempt to understand why NeurComm performed the way it did. Although the primary goal of the EGTA was to highlight the differences in solution concepts between algorithms, we agree that understanding the reason why these differences exist is also important and will improve the quality of our paper. In this spirit, we have taken the advice of R2, and computed the metrics from [1], namely, the Utilitarian, Equality and Sustainability metric for each algorithm for the two cases where (1) all agents cooperate and (2) agents play their equilibrium strategy, shown in Fig. 1A-C. When all agents are cooperating, the Equality for NeurComm is lower (i.e. the distribution of reward among agents is less uniform) than that of the other communicating algorithms, however, it is still able to achieve both a higher score for the group (more utilitarian) and have higher levels of sustainability. This provides some supporting evidence to our claim that NeurComm is perhaps better able to learn how to coordinate agents effectively. We have included this analysis in our paper, with additional discussions and interpretations of the results.

Connections to prior game-theory work (R1, R3): We thank R1 for pointing us towards the valuable literature on network games (e.g. [2]). We will include a thorough discussion of this line of work in the context of our own. Furthermore, we thank R3 for suggesting that we connect our analysis to the underlying Markov game. Specifically, we used the finite sample analysis in [3] combined with bootstrap resampling (please see Fig. 1D for an example), to improve our estimates, obtain confidence intervals and ensure a large enough sample size to tightly bound the difference between our empirical estimates and the original game. We have included this analysis in our work.

Introduction of NSIs (R1, R3): The terms used in the literature on SSDs have a strong association with analysing human-like agents. For example, two of the conditions characterising a social dilemma are often referred to, respectively, as "fear" and "greed" [4], which in the context of an engineering system would seem an anthropomorphisation of the incentive to defect. Our aim with introducing the NSI was to remove this additional layer of association. Therefore, we combined the fear and greed conditions to form C_3 in our paper and opted for terminology such as "inefficiency" instead of "social dilemma". We have taken care to add the above motivation to the paper and improve the clarity behind the term. However, we are aware that social dilemmas are long established concepts in game theory and if a final suggestion is given that we remain with the SSD terminology to make the paper more accessible, we will do so.

R3 *(Schelling plots and meaning of ovals; NSI definition and computation)*: We have changed our Schelling diagrams to use the $|C| - 1$ parameterisation to improve readability and clarified the meaning of the shaded ovals. As R3 points out, these are meant to refer to the payoffs obtained by agents playing a strategy profile in Nash equilibrium. Furthermore, to compute the value in the NSI indicator related to C_3, we used the value that is the maximum of $R_d(n-1) - R_c(n)$ over the range specified in the paper. We have noted the calculation explicitly in the paper for the new parameterisation. We have also updated our stated goal surrounding the identification of NSIs. We agree with R3's suggestion that it makes more sense to refer to NSIs as situations where there *exists* at least one inefficient equilibrium. For example, in our original definition, IA2C would not have strictly been classified as an NSI, which is inconsistent with the indicator. Finally, we have added a more detailed explanation regarding the last two columns of Fig. 4A (also requested by R2), which is in line with R3's interpretation.

Misc: We have made the following corrections/changes as suggested: correctly labelling the equilibrium in Fig. 4D (R1); correcting all spelling, enlarging Fig. 4 for improved readability (R2); using the correct term for defecting agents, adding an explanation of the NSI indicator in Fig. 3D and numerical values, changing the order of the plots in Fig. 4B-H to match the order in Fig. 4A and removing the subjective language in the conclusion (R3). We greatly thank the reviewers for improving our work with these suggestions.

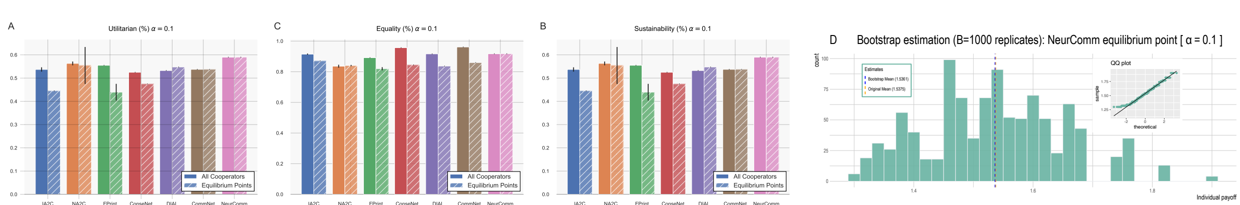

Figure 1: *Social metric and finite sample analysis (best viewed zoomed in).* **(A)** Utilitarian **(B)** Equality **(C)** Sustainability **(D)** Bootstrap estimation.

[1] J. Perolat, J. Z. Leibo, V. Zambaldi, C. Beattie, K. Tuyls, and T. Graepel, "A MARL model of common-pool resource appropriation," NeurIPS, 2017.
[2] Jackson, Matthew O., and Yves Zenou. "Games on networks." In Handbook of game theory with economic applications, 2015.
[3] Tuyls, Karl, Julien Perolat, Marc Lanctot, Joel Z. Leibo, and Thore Graepel. "A generalised method for empirical game theoretic analysis." AAMAS, 2018.
[4] Hughes, Edward, Joel Z. Leibo, Matthew Phillips, Karl Tuyls, Edgar Dueñez-Guzman, Antonio García Castañeda, Iain Dunning et al. "Inequity aversion improves cooperation in intertemporal social dilemmas." NeurIPS, 2018.