We appreciate the time and efforts invested by the reviewers for examining our work and providing detailed comments that would help us in improving the overall quality of our paper. We acknowledge the general comments regarding the density of the technical content in the paper and welcome the constructive feedback on improving the clarity of the paper. We will improve the overall presentation and add more details for better accessibility. As NeurIPS allows an extra page in the final version if accepted, it would give us enough room to reorganize the material so as to alleviate the concerns of the reviewers. Having said that, we firmly believe that the current version of the paper is fairly self-contained and has all the necessary details pertinent to the significantly novel and unique contributions of our paper. We hope to convince the reviewers of the same through our clarifications below and request them to revisit their score of our work.

R1: Thank you for your review and clarifying questions. We respectfully disagree with your assessment that the paper needs significant expansion. The ingredients, that we leverage to propose our first of its kind contribution to learn network emergence games directly from the data, are well-established *individually* in respective game theory, graph learning and RL communities. Hence, we chose to introduce them in the main paper to expose them to the readers and expand them in Appendix A to the extent relevant to our work, but defer to the basic expertise of readers or relevant previous literature for deeper understanding of those basic concepts. **Datasets.** The use of Karate network follows standard evaluation approach from IRL literature where reward quality is tested using small example for tractability. For interpretability, we chose the bank dataset (c.f. [38] for motivation on its usefulness) as it is well known in game theory community and has been analyzed rigorously for strategic mechanisms. **Clarity.** The reward in (line 278) just represent one form of possible ground truth reward (c.f. [52] for motivation behind this form) that is used following standard evaluation procedure in IRL literature and it can take any other form with no effect on our approach. **Feedback.** Feature set description is provided in Appendix C.1. Line 215 is standard message passing architecture and we refer the reviewer to [49] for details of GNN. We use continuous actions and map output action vectors drawn from Gaussian policy to the discrete edge operations. Please refer to lines 300-325 in paper and also **R3**'s review for Figure 2(b,c). Circles in Figure 2(c) signify that Katz centrality for those agents was increased when perturbing structure for testing. You misunderstand line 299 as it exactly shows that learned objective is useful – the policy trained to optimize the learned objective, exhibit closer behavior to original network, which would not be the case if recovered objective was not useful. For transfer, the reward function remains fixed i.e. reward is not trained again on target set of agents.

R2: Thank you for providing your feedback in great detail that helped us to clearly understand your concerns which we address below. We wish to emphasize that, in addition to use of GNN for polices and reward function as you rightly point out, to the best of our knowledge, MINE is the first approach to formulate network emergence games as multi-agent RL problem and solve it using inverse learning framework to learn both payoffs and strategies directly from data (our most significant contribution). **Interpretability.** MINE enables discovering payoff functions from data. Simply put, interpretability analysis signifies that if a useful reward function is recovered, an agent trained to optimize such objective would receive higher reward value (utility) (during evaluation) for taking actions (achieving states) that align with real-world expectations, thereby demonstrating the agreement of learned reward with the generative mechanisms of observed data. As a case in point, for Katz centrality experiments, reward function computes utility of an agent for being in a specific structural configuration (state) at convergence. For instance, teller agents generally have low Katz centrality in real-world bank networks. So higher utility value obtained for a teller with low Katz in a given structural configuration, shows learning of useful reward function. **Evaluation.** It seems there is a misunderstanding on your part as connection order is not relevant for our evaluation. Link prediction only compares with ground truth links and qualitative experiments evaluate performance in terms of utility values, both does not need to infer connection order. **Correctness.** line 108 has gamma factor inside sum; line 162, equation continues on next line (with a multiplication between two terms); line 166, description of QRE is standard and correct but shortened. Please refer to the expansion in Appendix A and relevant papers we have cited for more details; line 186 again is a standard way to describe IRL procedure (c.f. [51]) – it aims to recover a function that rationalizes the expert behaviors with the least commitment. **Clarity.** For both the reward and policy function, we follow widely used parameter-sharing mechanism for MARL (c.f. [19]) which supports transfer. The input to these functions are agent-specific observations rather than global state (due to partial observable case) and agents are identified by their features, making transfer to different set of agents feasible. Computing reward using current state-action as input is equivalent to computing it using only next state (a standard practice in RL literature and c.f. [51] for our specific form). Please refer to our response to **R1** on **Datasets**. We provide more experimental details in Appendix C and we will address other notation issues in final version of the paper.

R3: Thank you for your positive reviews and accurate characterization of our contributions. For lines 31-32, we meant to say *underexplored* (typo). Thanks for interesting related works, we will add them in our discussions. For link recovery in qualitative analysis, our primary goal was to check how many strategic links are recoverable at convergence. But we agree with your overall point and we do report AUC for link prediction which would consider false positives and partially answers your question. We will also add your suggestion to qualitative analysis in final version.

R4: Thank you for your supportive feedback. We hope that above clarifications help to alleviate some of your related concerns. We wish to emphasize that link prediction experiments aim to show that the learned game is useful as a predictive model, a desirable but lacking property in classical game-theoretic approaches. However, note that we jointly learn network formation strategy and utility function from the data, unlike baselines that optimize task-specific hand-designed objective. Hence MINE's comparable prediction performance, along with the interpretability and transfer benefits, makes it a versatile contribution. In similar vein, we are not aware of any previous works on learning network emergence games that support transfer and interpretability dimensions, hence no baselines for qualitative analysis. The quality dimension corresponds to standard approach in IRL literature used to assess the quality of the learned reward.