Game-Family Learning for Simulation-Based Games

Doctoral Consortium

Madelyn Gatchel
University of Michigan
Ann Arbor, MI, USA
gatchel@umich.edu

ABSTRACT

My Ph.D. research introduces an approach to learning families of parametrically related symmetric game instances in both normal-form and Bayesian settings. This approach eliminates the need to model each game instance separately, improving data efficiency and enabling broader exploration of the parameter space. Game-family models support more comprehensive empirical mechanism design, and facilitate iterative generation of piecewise best-response strategies in the Bayesian setting. Overall, my work aims to expand model expressiveness while prioritizing compactness and data efficiency in simulation-based game environments, promoting diverse real-world insights.

KEYWORDS

Computational Game Theory; Simulation-Based Games; Game-Model Learning; Equilibrium Computation

ACM Reference Format:
Madelyn Gatchel. 2025. Game-Family Learning for Simulation-Based Games: Doctoral Consortium. In *Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Detroit, Michigan, USA, May 19 – 23, 2025*, IFAAMAS, 3 pages.

1 INTRODUCTION

Building a game-theoretic model of a multi-agent strategic environment requires making assumptions about what information is strategically relevant and how agents use that information to make decisions. Outcomes depend on agents' joint strategies and on parameters of the environment, where each parameter setting induces a different ***game instance***. My work is motivated by the setting of ***simulation-based games***, where payoff data comes from an agent-based simulator [9–11]. Given limited modeling resources, analysts must decide in advance which parameter settings to explore (and at what granularity) and which strategies to include, typically selecting only a small subset of all possible strategies.

These assumptions, while necessary for tractability, may limit the scope and applicability of the game-theoretic analysis—for example, if an omitted parameter setting exhibits different or desirable equilibrium behavior, or if the strategies are too narrowly defined. To address these limitations, my dissertation introduces methods for ***game-family learning***, in which a single neural-network model is trained to represent a family of game instances related by a common environment parameter. These methods also leverage other environment structure, such as symmetry or agent type, to further promote data efficiency, which is advantageous for simulation-based game environments.

Proc. of the 24th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2025), Y. Vorobeychik, S. Das, A. Nowé (eds.), May 19 – 23, 2025, Detroit, Michigan, USA. © 2025 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

2 LEARNING GAME FAMILIES

Normal-Form Game Families. The normal-form representation is the most basic game-theoretic model of incentives, specifying each player's strategy set and a payoff function that measures their satisfaction with each outcome. Sokota et al. [7] showed that learning the ***deviation payoff function***—the expected payoff for a unilateral deviator—enables equilibrium computation in symmetric normal-form game instances without the combinatorial mixture summations required with direct payoff functions.

A key hypothesis of my research is that game instances related by a common environment parameter also have related (deviation) payoff functions. Accordingly, we proposed augmenting the neural-network model to also input the environment parameter, and evaluated this approach on two classes of random games, where the number of symmetric players serves as the environment parameter [3]. We compared a single model trained across the full parameter space with several fixed-parameter learning models (FPL) [7], one trained for each game instance, given the same overall training data. Figure 1 shows that the game-family learning approach (VPL) has lower payoff error than FPL given the same amount of data, and remains more accurate even with less. These results suggest that payoff data from neighboring parameter settings aids game-family learning, improving approximation of deviation payoffs.

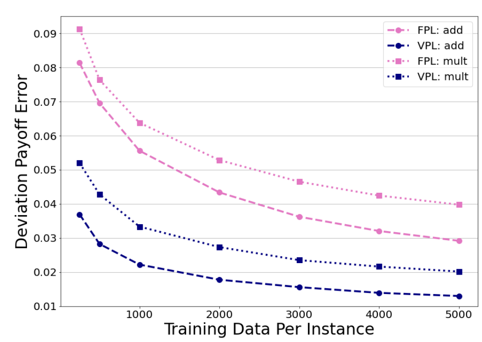

Figure 1: A game-family model (VPL) achieves lower payoff error than separately trained models for each parameter setting (FPL) on two random game classes ("add" and "mult").

2.0.1 Bayesian Game Families. Bayesian games build upon the normal form by introducing ***types***, private information players use to make decisions. In a Bayesian game, strategies map types to actions, and payoff functions map joint strategy profiles and type profiles to payoffs. We aim to approximate an ***ε*-Bayes-Nash equilibrium** (ε-BNE), where no player can gain more than ε payoff by unilaterally deviating *in expectation over player types*. The ***ex ante deviation payoff function*** computes this expected payoff across all possible player types, and is equivalent to the normal-form payoff function when types are abstracted away. The ***interim deviation payoff function*** represents a player's deviation payoff given their type, in expectation over all possible opponent types. Marginalizing over the deviator's type then yields the ex ante deviation payoff.

We proposed learning the interim deviation payoff function for a Bayesian game family to exploit the type-conditional structure of strategies in Bayesian games [2]. We evaluated both interim and ex ante (baseline) game-family learning approaches on a dynamic sponsored search auction game family, with reserve price as the environment parameter. Figure 2 shows the learning performance of each approach on the trained parameter range, $(0, 8]$, and beyond $(8, 15]$. Within the trained range, interim models with sufficient marginalization samples have errors comparable to ex ante, and all errors are small relative to the payoff scale. Further, all marginalized interim models extrapolate well, while the ex ante model exhibits a noticeable increase in error when extrapolating. As the parameter range covered by the training data must be set prior to any game-theoretic analysis, there is a risk that it may be too narrow. A model that can extrapolate is therefore valuable for applications like empirical mechanism design.

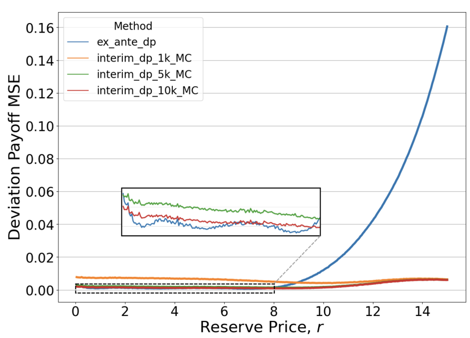

Figure 2: With enough marginalization samples, interim model accuracy matches ex ante on the trained range, $(0, 8]$. Interim, but not ex ante, models extrapolate well.

3 EMPIRICAL MECHANISM DESIGN

In ***empirical mechanism design*** (EMD), a designer sets or influences an environment parameter, each value of which results in a different Bayesian game instance induced from simulator data. The goal is to find the parameter setting that optimizes a relevant objective, such as social welfare or revenue. In past EMD studies [1, 4, 8], researchers selected a limited set of mechanism settings, separately modeling and analyzing each game instance. Once trained, a game-family model can instead evaluate any game instance within the trained range (and often beyond), supporting a more granular parameter search. When used in an optimization algorithm, it eliminates the need to train separate models at each iteration, reducing algorithm dependence on the sampling budget.

In our experiments, we also found that the interim approach had equal or lower absolute error on candidate BNE regret (Figure 3), where a candidate BNE has predicted regret below ε. Consequently, the expected revenue curve in equilibrium with the interim approach exhibits far fewer holes—where no candidate equilibria have true regret below ε—a property crucial for effective EMD.

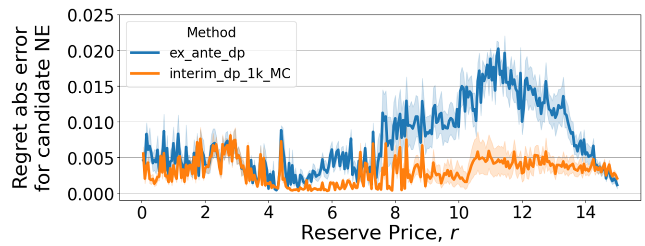

Figure 3: Candidate equilibria from the interim approach have equal or lower regret error compared to the ex ante approach.

4 ITERATIVE BEST-RESPONSE GENERATION

The interim learning approach enables us to iteratively grow the strategy set with more sophisticated piecewise best response strategies. In an ε-BNE no player can gain by deviating to another strategy *in expectation over player types*. However, conditional on its own type, a player can often benefit by deviating to a strategy not in the modeled strategy set. We exploit these gains by introducing new strategies which explicitly condition on type and map to actions that are more strategically relevant.

For a given ε-BNE, we can compute a piecewise best response strategy: for each interval in a partitioned type space, we use the learned model to predict interim deviation payoffs for many randomly generated types in the interval, and find the average deviation payoff for each strategy. The best response for that interval is the strategy with the largest interim deviation payoff. We find that a player can indeed gain additional payoff by deviating to a piecewise best response when other players play the ε-BNE.

We can further integrate these piecewise strategies into an expanded model trained without requiring any additional simulation samples. These operations enable an iterative procedure that expands the game-family model from an initial set of atomic strategies through a double oracle [6] approach: repeated generation of new (piecewise) strategies that best-respond to an equilibrium of the previous configuration. By choosing an equilibrium from the game instance that optimizes the environment parameter, this becomes an iterative method for EMD. In applying this method to the dynamic search auction, our results suggest even a couple of iterations refines the model to produce decisions that improve revenue.

5 FUTURE WORK

In the future, I would like to extend the game-family learning approach to extensive-form game environments. I am also interested in incorporating game-family learning into PSRO [5].

REFERENCES

[1] Erik Brinkman and Michael P. Wellman. 2017. Empirical Mechanism Design for Optimizing Clearing Interval in Frequent Call Markets. In *18th ACM Conference on Economics and Computation*. 205–221.

[2] Madelyn Gatchel and Michael P. Wellman. 2025. Learning Bayesian Game Families, with Application to Mechanism Design (Extended Abstract). In *24th International Conference on Autonomous Agents and Multiagent Systems*.

[3] Madelyn Gatchel and Bryce Wiedenbeck. 2023. Learning Parameterized Families of Games. In *22nd International Conference on Autonomous Agents and Multiagent Systems*. 1044–1052.

[4] Patrick R. Jordan, Michael P. Wellman, and Guha Balakrishnan. 2010. Strategy and Mechanism Lessons from the First Ad Auctions Trading Agent Competition. In *11th ACM Conference on Electronic Commerce*. 287–296.

[5] Marc Lanctot, Vinicius Zambaldi, Audrunas Gruslys, Angeliki Lazaridou, Karl Tuyls, Julien Perolat, David Silver, and Thore Graepel. 2017. A Unified Game-Theoretic Approach to Multiagent Reinforcement Learning. In *31st Conference on Neural Information Processing Systems*. 4193–4206.

[6] H. Brendan McMahan, Geoffrey J. Gordon, and Avrim Blum. 2003. Planning in the presence of cost functions controlled by an adversary. In *20th International Conference on Machine Learning* (Washington, DC). 536–543.

[7] Samuel Sokota, Caleb Ho, and Bryce Wiedenbeck. 2019. Learning Deviation Payoffs in Simulation-Based Games. In *33rd AAAI Conference on Artificial Intelligence*. 2173–2180.

[8] Yevgeniy Vorobeychik, Christopher Kiekintveld, and Michael P. Wellman. 2006. Empirical Mechanism Design: Methods, with Application to a Supply-Chain Scenario. In *7th ACM Conference on Electronic Commerce*. 306–315.

[9] Michael P Wellman. 2006. Methods for Empirical Game-Theoretic Analysis. In *21st National Conference on Artificial Intelligence*. 1152–1155.

[10] Michael P. Wellman. 2020. Economic Reasoning from Simulation-Based Game Models. *Œconomia. History, Methodology, Philosophy* 10, 2 (2020), 257–278.

[11] Michael P. Wellman, Karl Tuyls, and Amy Greenwald. 2025. Empirical Game-Theoretic Analysis: A Survey. *Journal of Artificial Intelligence Research* 82 (2025).