Towards Fast, Flexible, and Data-Efficient Algorithms for Multi-Agent Interaction

Lasse Peters
Delft University of Technology

Robots that are capable of interacting with others—humans and robots alike—have the potential to revolutionize many aspects of our daily lives through applications such as autonomous driving, collaborative assembly, or autonomous construction. Recent research has made great advances in the *single-agent* domain, enabling impressive results for complex tasks such as manipulation of diverse objects [1–4] and mobile robotics [5–7]. Despite this remarkable progress, for *multi-agent* interactions—in which a robot must interact with other, potentially uncontrolled or even untrusted agents—similar results remain out of reach. A key challenge in solving complex multi-agent tasks is the scarcity of data in these settings, which can be attributed to two main issues: (i) the product space of states and actions grows exponentially with the number of players and (ii) the cost of acquiring data for multi-agent systems is inevitably higher than for single-agent systems. Therefore, rather than hoping that we can collect enough data to cover the entire space of multi-agent interactions, I am interested in exploiting the *game-theoretic structure* of multi-agent interaction to enable learning from limited data. *This demands new algorithms that encode game-theoretic priors for multi-agent interaction in a formulation compatible with online and offline learning from diverse data sources—including single- and multi-agent experience—and techniques for dealing with the remaining uncertainty.*

I. Past Research

Throughout my PhD, I have focused on three key areas to build towards my vision of efficient algorithms for multi-agent interaction: (i) a learning-based solver for certainty-equivalent games, (ii) algorithms for learning game-theoretic models from online and offline data, and (iii) techniques for planning under uncertainty in multi-agent interaction.

Solving games efficiently. Even when a game-theoretic model is known without uncertainty, solving such problems poses a computational challenge [8]. Prior works seek to solve games numerically through methods such as iterated best response [9, 10], iterative linear-quadratic games [11, 12], sequential quadratic programming [13], or augmented Lagrangian methods [14]. When the strategy space becomes high-dimensional—e.g., due to the number of agents or the need to model probabilistic (mixed) strategies—numerical online methods become impractical. To address this, my paper published at RSS [15] proposes a game solver that trains neural network policies by decomposing the mixed-strategy multi-agent problem into parallelizable pure-strategy single-agent optimization problems whose solutions are combined via a discrete game solver to approximate the original game solution. By recovering the gradients of the discrete game solver and the single-agent optimization problems via implicit differentiation, our method can train the neural network policy end-to-end. *In summary, my work contributed a fast and flexible, learning-enabled game solver that has since been used in various applications, including collision avoidance with up to 10 players [16], and head-to-head drone racing [17].*

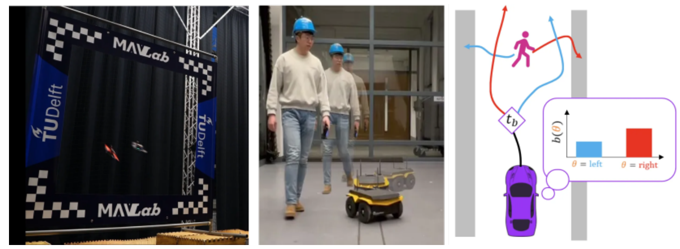

Fig. 1: Research overview. My research explores efficient algorithms for multi-agent interaction. I ground my work in applications such as head-to-head drone-racing (left), interaction with humans (center), and autonomous driving (right).

Learning game-theoretic models: online and offline. In real-world deployment, robots cannot assume access to a complete model of the world a priori. Instead, it is more practical to pre-train on offline data and refine the interaction model based on online observations. To address this issue, several works explored solving the *inverse* game problem; i.e. the problem of inferring unknown game parameters from *observed* interactions [18–21]. However, traditionally, these methods relied on the strong assumption that the robot can observe others' actions directly. In my experiments, I observed that this assumption impedes accurate intent inference in settings where robots have only partial and potentially noisy observations of the world. In my RSS paper [22], I therefore proposed a new formulation of inverse games as a bi-level problem: maximum-likelihood estimation with equilibrium constraints. By casting intent inference and state estimation in this framework, our method reliably recovers more accurate estimates of states *and* intents than prior approaches. In a follow-up IJRR work [23], I extended this approach to the online setting, allowing robots to adapt their beliefs about others' intents as they interact. Beyond the intent estimate, this approach also computes the corresponding best response strategy for the robot as the solution of a single optimization problem. Finally, to reduce the computational cost of solving these challenging bi-level programs online, we proposed a new technique that leverages *implicit* differentiation of the inner equilibrium problem to compute gradients of the observation likelihood. This implicit differentiation technique allows the use of game solvers, including those discussed above, as implicit neural network layers, enabling robots to learn a neural network policy that *directly predicts* the inverse game solution without online optimization [24, 25] (Figure 1, center). *In summary, my work contributed versatile tools to infer the intents of other agents from short interaction histories which other community members have extended to tackle various applications, including reasoning about occluded pedestrians in traffic [26] and adaptive autonomous head-to-head racing [27].*

Planning with uncertainty in multi-agent interaction. Even with a good base model identified from offline data, the information available online may not suffice to estimate an unambiguous interaction model. For example, in autonomous driving, a short interaction history may be insufficient to determine a unique goal position of a pedestrian. Two distinct approaches to deal with this issue dominate prior work: (i) selecting the most likely intent and optimizing under a certainty-equivalent model [24, 28, 29], or (ii) generating an ego-plan that minimizes expected cost under the distribution of other agents' intents [30–32]. However, the former approach leads to overly optimistic—or even unsafe—behavior, while the latter approach can be overly conservative since it ignores the fact that more information will be available as the interaction unfolds. In my paper published at RA-L [33], I formalized a middle-ground between these two extremes by casting multi-agent interaction under uncertainty as game-theoretic contingency planning. By solving the resulting contingency game, a robot jointly recovers a multi-hypothesis prediction of others and a corresponding conditional plan for itself, anticipating future information gains while accounting for the interdependence of actions between agents (Figure 1, right). Our results demonstrate that this approach matches the more conservative approaches in terms of safety while outperforming them in terms of interaction efficiency. To scale this paradigm to settings with a large number of hypotheses, I have closely collaborated with international partners to develop a specialized ADMM solver that parallelizes computation across hypotheses [34]. Finally, to further facilitate this research direction, I also guided a junior PhD student in developing a learning-based variational inference approach [35] that leverages our differentiable game solver [24] to efficiently estimate distributional beliefs of game-theoretic models and plans with the resulting uncertainty using the contingency games framework. *In summary, my work contributed a tractable method for planning under uncertainty in multi-agent interaction, demonstrated to integrate well with learning-based inference techniques. It attracted the interest of the autonomous delivery company Nuro, leading to an invited talk at their research seminar [36].*

II. Research Vision

I envision robots interacting with humans and other robots in diverse scenarios beyond autonomous driving, including collaborative assembly, construction, and household tasks. Toward this goal, I am excited about developing interaction algorithms that can handle high-dimensional observations, incomplete information, and complex dynamics while training on both single-agent and multi-agent data efficiently.

Leveraging single-agent data for multi-agent interaction. My previous works have largely focused on multi-agent interaction scenarios in which a compact state representation is available, and where coordination challenges primarily lie in avoiding collisions with other agents. Many tasks faced by general-purpose robots—e.g., preparing a complex dish in a shared kitchen—require a richer state representation and a more nuanced understanding of interaction that goes beyond collision avoidance. In the *single-agent* domain, generative models for imitation learning have shown promise for complex manipulation tasks [1–5]. However, multi-agent instantiations of this idea remain challenging since covering the product space of possible interactions between multiple agents with expert demonstrations is impractical. Therefore, we need to rethink how to learn complex multi-agent tasks in a data-efficient manner. To this end, I aim to develop algorithms that leverage *single-agent* expert demonstrations to aid learning of *multi-agent* policies—an idea closely related to multi-agent transfer learning (MATL) [37]. In contrast to prior work on MATL, which focuses on transfer between similar *multi-agent tasks*, I seek to use *single-agent* source tasks to tap into the wealth of single-agent data. Future work includes investigating how to efficiently combine single-agent pre-training, game-theoretic online reasoning, and limited multi-agent data to bridge the gap between single-agent behavior and tightly coupled multi-agent interaction.

Game-theoretic planning with multi-agent world models. Foundation models pre-trained on internet-scale data have shown great success in natural language processing [38–40] and beyond [41–46]. My research visit at CMU exposed me to one of these tools, DINO-WM [43], and the performance that we obtained for a complex real-world single-agent manipulation task [47] convinced me that learned world models are a key ingredient also for multi-agent interactions. However, when other agents act in the same environment, existing world models treat other agents implicitly as part of the environment dynamics, effectively relying on an implicit theory of mind that must be learned from data. Instead, I am interested in building world models that explicitly distinguish between *environment state dynamics* and each *agents' belief and decisions*. This game-theoretic structure serves as an inductive bias that makes agents' theory of mind explicit, reducing data requirements while enabling (i) strategic reasoning about interactions and (ii) more data-efficient training through the imposed structure. Future research should explore model structures and training objectives that are informed by game-theoretic insights, e.g. to capture the incomplete and time-varying nature of information available to each player. Based on such models, I aim to explore how to plan in the latent space of these multi-agent world models and exploit synergies between offline training and online reasoning.

REFERENCES

[1] Tony Z Zhao, Jonathan Tompson, Danny Driess, Pete Florence, Kamyar Ghasemipour, Chelsea Finn, and Ayzaan Wahid. Aloha unleashed: A simple recipe for robot dexterity. *arXiv preprint*, 2410.13126, 2024.

[2] Vitalis Vosylius and Edward Johns. Instant policy: In-context imitation learning via graph diffusion. *International Conference on Learning Representations (ICLR)*, 2025.

[3] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion. *The International Journal of Robotics Research (IJRR)*, 2023.

[4] Letian Fu, Huang Huang, Gaurav Datta, Lawrence Yunliang Chen, William Chung-Ho Panitch, Fangchen Liu, Hui Li, and Ken Goldberg. In-context imitation learning via next-token prediction. *arXiv preprint*, 2408.15980, 2024.

[5] Zipeng Fu, Tony Z Zhao, and Chelsea Finn. Mobile aloha: Learning bimanual mobile manipulation with low-cost whole-body teleoperation. *Conference on Robot Learning (CoRL)*, 2024.

[6] Kevin Black, Noah Brown, Danny Driess, Adnan Esmail, Michael Equi, Chelsea Finn, Niccolo Fusai, Lachy Groom, Karol Hausman, Brian Ichter, et al. π_0: A vision-language-action flow model for general robot control. *arXiv preprint*, 2410.24164, 2024.

[7] Karl Pertsch, Kyle Stachowicz, Brian Ichter, Danny Driess, Suraj Nair, Quan Vuong, Oier Mees, Chelsea Finn, and Sergey Levine. Fast: Efficient action tokenization for vision-language-action models. *arXiv preprint*, 2501.09747, 2025.

[8] Tamer Başar and Geert Jan Olsder. *Dynamic noncooperative game theory*. Society for Industrial and Applied Mathematics (SIAM), 2. edition, 1999.

[9] Riccardo Spica, Eric Cristofalo, Zijian Wang, Eduardo Montijano, and Mac Schwager. A real-time game theoretic planner for autonomous two-player drone racing. *IEEE Trans. on Robotics (TRO)*, 36(5), 2020.

[10] Zijian Wang, Riccardo Spica, and Mac Schwager. Game theoretic motion planning for multi-robot racing. *Distributed Autonomous Robotic Systems*, 2019.

[11] David Fridovich-Keil, Ellis Ratner, Lasse Peters, Anca D Dragan, and Claire J Tomlin. Efficient iterative linear-quadratic approximations for nonlinear multi-player general-sum differential games. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*. IEEE, 2020.

[12] Justin Lidard, Haimin Hu, Asher Hancock, Zixu Zhang, Albert Gimó Contreras, Vikash Modi, Jonathan DeCastro, Deepak Gopinath, Guy Rosman, Naomi Ehrich Leonard, et al. Blending data-driven priors in dynamic games. *Proc. of Robotics: Science and Systems (RSS)*, 2024.

[13] Edward L Zhu and Francesco Borrelli. A sequential quadratic programming approach to the solution of open-loop generalized nash equilibria for autonomous racing. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*. IEEE, 2023.

[14] Simon Le Cleac'h, Mac Schwager, and Zachary Manchester. ALGAMES: A fast solver for constrained dynamic games. In *Robotics: Science and Systems (RSS)*, 2020.

[15] Lasse Peters, David Fridovich-Keil, Laura Ferranti, Cyrill Stachniss, Javier Alonso-Mora, and Forrest Laine. Learning mixed strategies in trajectory games. *Proc. of Robotics: Science and Systems (RSS)*, 2022.

[16] Maximilian Schmidt. Game-theoretic motion planning on a real-world hardware platform. Technical report, Hamburg University of Technology, 2022.

[17] Andrei-Carlo Papuc. Interaction-aware autonomous drone racing. Master's thesis, Delft University of Technology, 2025.

[18] Kevin Waugh, Brian D Ziebart, and J Andrew Bagnell. Computational rationalization: The inverse equilibrium problem. In *Proc. of the Int. Conf. on Machine Learning (ICML)*, 2011.

[19] Florian Köpf, Jairo Inga, Simon Rothfuß, Michael Flad, and Sören Hohmann. Inverse reinforcement learning for identification in linear-quadratic dynamic games. *IFAC-PapersOnLine*, 50(1), 2017.

[20] Simon Rothfuß, Jairo Inga, Florian Köpf, Michael Flad, and Sören Hohmann. Inverse optimal control for identification in non-cooperative differential games. *IFAC-PapersOnLine*, 50(1), 2017.

[21] Chaitanya Awasthi and Andrew Lamperski. Inverse differential games with mixed inequality constraints. In *Proc. of the IEEE American Control Conference (ACC)*. IEEE, 2020.

[22] Lasse Peters, David Fridovich-Keil, Vicenç Rubies-Royo, Claire J Tomlin, and Cyrill Stachniss. Inferring objectives in continuous dynamic games from noise-corrupted partial state observations. *Proc. of Robotics: Science and Systems (RSS)*, 2021.

[23] Lasse Peters, Vicenç Rubies-Royo, Claire J Tomlin, Laura Ferranti, Javier Alonso-Mora, Cyrill Stachniss, and David Fridovich-Keil. Online and offline learning of player objectives from partial observations in dynamic games. *The International Journal of Robotics Research (IJRR)*, 42(10), 2023.

[24] Xinjie Liu*, Lasse Peters*, and Javier Alonso-Mora. Learning to play trajectory games against opponents with unknown objectives. *IEEE Robotics and Automation Letters (RA-L)*, 8(7), 2023.

[25] Jingqi Li, Chih-Yuan Chiu, Lasse Peters, Somayeh Sojoudi, Claire Tomlin, and David Fridovich-Keil. Cost inference for feedback dynamic games from noisy partial state observations and incomplete trajectories. *Autonomous Agents and Multi-Agent Systems (AAMAS)*, 2023.

[26] Tianyu Qiu and David Fridovich-Keil. Inferring occluded

agent behavior in dynamic games from noise corrupted observations. *IEEE Robotics and Automation Letters (RA-L)*, 2024.

[27] Haimin Hu, Jonathan DeCastro, Deepak Gopinath, Guy Rosman, Naomi Ehrich Leonard, and Jaime Fernández Fisac. Think deep and fast: Learning neural nonlinear opinion dynamics from inverse dynamic games for split-second interactions. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*. IEEE, 2025.

[28] Dorsa Sadigh, Shankar Sastry, Sanjit A Seshia, and Anca D Dragan. Planning for autonomous cars that leverage effects on human actions. In *Proc. of Robotics: Science and Systems (RSS)*, 2016.

[29] Negar Mehr, Mingyu Wang, Maulik Bhatt, and Mac Schwager. Maximum-entropy multi-agent dynamic games: Forward and inverse solutions. *IEEE Trans. on Robotics (TRO)*, 2023.

[30] Simon Le Cleac'h, Mac Schwager, and Zachary Manchester. LUCIDGames: Online unscented inverse dynamic games for adaptive trajectory prediction and planning. *IEEE Robotics and Automation Letters (RA-L)*, 2021.

[31] Forrest Laine, David Fridovich-Keil, Chih-Yuan Chiu, and Claire Tomlin. Multi-Hypothesis Interactions in Game-Theoretic Motion Planning. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, 2021.

[32] Ran Tian, Liting Sun, Andrea Bajcsy, Masayoshi Tomizuka, and Anca D Dragan. Safety assurances for human-robot interaction via confidence-aware game-theoretic human models. In *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, 2022.

[33] Lasse Peters, Andrea Bajcsy, Chih-Yuan Chiu, David Fridovich-Keil, Forrest Laine, Laura Ferranti, and Javier Alonso-Mora. Contingency games for multi-agent interaction. *IEEE Robotics and Automation Letters (RA-L)*, 2024.

[34] Jingqi Li, Chih-Yuan Chiu, Lasse Peters, Fernando Palafox, Mustafa Karabag, Javier Alonso-Mora, Somayeh Sojoudi, Claire Tomlin, and David Fridovich-Keil. Scenario-Game ADMM: A parallelized scenario-based solver for stochastic noncooperative games. In *Proc. of the IEEE Conference on Decision and Control (CDC)*. IEEE, 2023.

[35] Xinjie Liu*, Lasse Peters*, Javier Alonso-Mora, Ufuk Topcu, and David Fridovich-Keil. Auto-encoding bayesian inverse games. *The International Workshop on the Algorithmic Foundations of Robotics (WAFR)*, 2024.

[36] Lasse Peters. Contingency games for multi-agent interaction. https://www.youtube.com/watch?v=WVzK7yBejlU, 2024. Research seminar at Nuro. Last accessed: March 26, 2024.

[37] Felipe Leno Da Silva and Anna Helena Reali Costa. A survey on transfer learning for multiagent reinforcement learning systems. *Journal of Artificial Intelligence Research*, 64:645–703, 2019.

[38] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. In Jill Burstein, Christy Doran, and Thamar Solorio, editors, *Proc. of the Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (ACL)*, Minneapolis, Minnesota, June 2019. Association for Computational Linguistics.

[39] Hugo Touvron, Thibaut Lavril, Gautier Izacard, Xavier Martinet, Marie-Anne Lachaux, Timothée Lacroix, Baptiste Rozière, Naman Goyal, Eric Hambro, Faisal Azhar, et al. Llama: Open and efficient foundation language models. *arXiv preprint*, 2302.13971, 2023.

[40] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altenschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint*, 2303.08774, 2023.

[41] Alec Radford, Jong Wook Kim, Chris Hallacy, Aditya Ramesh, Gabriel Goh, Sandhini Agarwal, Girish Sastry, Amanda Askell, Pamela Mishkin, Jack Clark, Gretchen Krueger, and Ilya Sutskever. Learning transferable visual models from natural language supervision. In Marina Meila and Tong Zhang, editors, *Proc. of the Int. Conf. on Machine Learning (ICML)*, volume 139 of *Proceedings of Machine Learning Research*. PMLR, 18–24 Jul 2021.

[42] Maxime Oquab, Timothée Darcet, Théo Moutakanni, Huy V. Vo, Marc Szafraniec, Vasil Khalidov, Pierre Fernandez, Daniel HAZIZA, Francisco Massa, Alaaeldin El-Nouby, Mido Assran, Nicolas Ballas, Wojciech Galuba, Russell Howes, Po-Yao Huang, Shang-Wen Li, Ishan Misra, Michael Rabbat, Vasu Sharma, Gabriel Synnaeve, Hu Xu, Herve Jegou, Julien Mairal, Patrick Labatut, Armand Joulin, and Piotr Bojanowski. DINOv2: Learning robust visual features without supervision. *Transactions on Machine Learning Research (TMLR)*, 2024.

[43] Gaoyue Zhou, Hengkai Pan, Yann LeCun, and Lerrel Pinto. Dino-wm: World models on pre-trained visual features enable zero-shot planning. *arXiv preprint*, 2411.04983, 2024.

[44] Wenli Xiao, Haoru Xue, Tony Tao, Dvij Kalaria, John M Dolan, and Guanya Shi. Anycar to anywhere: Learning universal dynamics model for agile and adaptive mobility. *Proc. of the IEEE Intl. Conf. on Robotics & Automation (ICRA)*, 2025.

[45] Jason Baldridge, Jakob Bauer, Mukul Bhutani, Nicole Brichtova, Andrew Bunner, Kelvin Chan, Yichang Chen, Sander Dieleman, Yuqing Du, Zach Eaton-Rosen, et al. Imagen 3. *arXiv preprint*, 2408.07009, 2024.

[46] Jake Bruce, Michael D Dennis, Ashley Edwards, Jack Parker-Holder, Yuge Shi, Edward Hughes, Matthew Lai, Aditi Mavalankar, Richie Steigerwald, Chris Apps, et al. Genie: Generative interactive environments. In *Proc. of the Int. Conf. on Machine Learning (ICML)*, 2024.

[47] Kensuke Nakamura, Lasse Peters, and Andrea Bajcsy. Generalizing safety beyond collision-avoidance via latent-space reachability analysis. *Proc. of Robotics: Science and Systems (RSS)*, 2025.