[19] Yuma Fujimoto and Kunihiko Kaneko. Functional dynamic by intention recognition in iterated games. *New Journal of Physics*, 21(2):023025, 2019.

[20] Robert Axelrod and William D Hamilton. The evolution of cooperation. *Science*, 211(4489):1390–1396, 1981.

[21] Martin Nowak and Karl Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(6432):56–58, 1993.

[22] Drew Fudenberg and Eric Maskin. The folk theorem in repeated games with discounting or with incomplete information. In *A long-run collaboration on long-run games*, pages 209–230. World Scientific, 2009.

[23] Wolfram Barfuss, Jonathan F Donges, and Jürgen Kurths. Deterministic limit of temporal difference reinforcement learning for stochastic games. *Physical Review E*, 99(4):043305, 2019.

[24] Wolfram Barfuss. Reinforcement learning dynamics in the infinite memory limit. In *AAMAS*, pages 1768–1770, 2020.

[25] Janusz M Meylahn, Lars Janssen, et al. Limiting dynamics for q-learning with memory one in symmetric two-player, two-action games. *Complexity*, 2022, 2022.

[26] Yuma Fujimoto and Kunihiko Kaneko. Emergence of exploitation as symmetry breaking in iterated prisoner's dilemma. *Physical Review Research*, 1(3):033077, 2019.

[27] Yuma Fujimoto and Kunihiko Kaneko. Exploitation by asymmetry of information reference in coevolutionary learning in prisoner's dilemma game. *Journal of Physics: Complexity*, 2(4):045007, 2021.

[28] Lloyd S Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953.

[29] Michael L Littman. Markov games as a framework for multi-agent reinforcement learning. In *ICML*, pages 157–163, 1994.

[30] Josef Hofbauer. Evolutionary dynamics for bimatrix games: A hamiltonian system? *Journal of Mathematical Biology*, 34(5):675–688, 1996.

[31] Georgios Piliouras, Carlos Nieto-Granda, Henrik I Christensen, and Jeff S Shamma. Persistent patterns: Multi-agent learning beyond equilibrium and utility. In *AAMAS*, pages 181–188, 2014.

[32] Constantinos Daskalakis, Andrew Ilyas, Vasilis Syrgkanis, and Haoyang Zeng. Training GANs with optimism. In *ICLR*, 2018.

[33] Constantinos Daskalakis and Ioannis Panageas. Last-iterate convergence: Zero-sum games and constrained min-max optimization. In *ITCS*, pages 27:1–27:18, 2019.

[34] Panayotis Mertikopoulos, Bruno Lecouat, Houssam Zenati, Chuan-Sheng Foo, Vijay Chandrasekhar, and Georgios Piliouras. Optimistic mirror descent in saddle-point problems: Going the extra(-gradient) mile. In *ICLR*, 2019.

[35] Noah Golowich, Sarath Pattathil, and Constantinos Daskalakis. Tight last-iterate convergence rates for no-regret learning in multi-player games. In *NeurIPS*, pages 20766–20778, 2020.

[36] Chen-Yu Wei, Chung-Wei Lee, Mengxiao Zhang, and Haipeng Luo. Linear last-iterate convergence in constrained saddle-point optimization. In *ICLR*, 2021.

[37] Qi Lei, Sai Ganesh Nagarajan, Ioannis Panageas, et al. Last iterate convergence in no-regret learning: constrained min-max optimization for convex-concave landscapes. In *AISTATS*, pages 1441–1449, 2021.

[38] Kenshi Abe, Mitsuki Sakamoto, and Atsushi Iwasaki. Mutation-driven follow the regularized leader for last-iterate convergence in zero-sum games. In *UAI*, pages 1–10, 2022.