Title of Thesis is here

by

Thiruvenkadam Sivaprakasam Radhakrishnan

THESIS

Submitted as partial fulfillment of the requirements
for the degree of Master's in Computer Science
in the Graduate College of the
University of Illinois at Chicago, 2023

Chicago, Illinois

Defense Committee:
Prof. Ian Kash, Chair and Advisor
Prof. Anastasios Sidiropoulos
Prof. Ugo Buy

ACKNOWLEDGMENTS

The thesis has been completed... (INSERT YOUR TEXTS)

YOUR INITIAL

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF NOTATIONS

θ	Parameters of a function approximator.

SUMMARY

Put your summary of thesis here.

CHAPTER 1

INTRODUCTION

Multi-agent reinforcement learning Two-player zero-sum games

CHAPTER 2

BACKGROUND

2.1 Game Theory

2.1.1 Problem Domains

2.1.2 Solution Concepts

2.1.3 Algorithms

2.2 Online Learning

2.2.1 Mirror Descent

CHAPTER 3

MAGNETIC MIRROR DESCENT

3.0.1 Mirror Descent

3.0.2 Magnetic Mirror Descent

Magnetic mirror descent (MMD) [1] is a last-iterate equilibrium approximation algorithm for two-player zero-sum games that is an extension of mirror descent with entropy regularization.

The idea behind MMD begins with the observation that solving for QRE in two-player zero-sum games can be reformulated as the solution to a negative entropy regularized saddle point problem as follows,

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \alpha g_1(x) + f(x, y) + \alpha g_2(y), \tag{3.1}$$

where $\mathcal{X} \subset \mathbb{R}^n$, $\mathcal{Y} \subset \mathbb{R}^m$ are closed and convex, and $g_1 : \mathbb{R}^n \to \mathbb{R}$, $g_2 : \mathbb{R}^m \to \mathbb{R}$, $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$.

The solution (x_*, y_*) to Equation 3.1, is a Nash equilibrium in the reguarlized game with the following best response conditions,

$$x_* \in \arg\min_{x \in \mathcal{X}} \alpha g_1(x) + f(x, y_*) \tag{3.2}$$

$$y_* \in \arg\min_{y \in \mathcal{Y}} \alpha g_2(y) + f(x_*, y) \tag{3.3}$$

Zero-sum games and Variational Inequalities

MMD reframes the solution to QRE as a variational inequality problem.

Definition 1 *Given $\mathcal{Z} \subseteq \mathbb{R}^n$ and mapping $G : \mathcal{Z} \to \mathbb{R}^n$, the variational inequality problem VI($\mathcal{Z}, G$) is to find $z_* \in \mathcal{Z}$ such that,*

$$\langle G(z_*, z - z_*) \rangle \geq 0 \quad \forall z \in \mathcal{Z}.$$

The optimality conditions are equivalent to VI($\mathcal{Z}, G$), where $G = F + \alpha \nabla g$, $\mathcal{Z} = \mathcal{X} \times \mathcal{Y}$, and $g : \mathcal{Z} \to \mathbb{R}$.

Now, the solution the the VI problem ($z_* = (x_*, y_*)$), corresponds to the solution of the saddle point problem stated in Equation 3.1, and satisfies the best response conditions Equation 3.2 and Equation 3.3.

We now restate the main algorithm as stated in [1](3.1),

AlgorithmMMD Starting with $z_1 \in \text{int dom } \psi \cap \mathcal{Z}$, at each iteration t do

$z_{t+1} = \arg\min_{z \in \mathcal{Z}} \eta(\langle F(z_t), z \rangle + \alpha g(z)) + B_\psi(z, z_t)$.

With the following assumptions, z_{t+1} is well defined:

- ψ is 1-strongly convex with respect to $\|.\|$ over $\mathcal{Z}$, and for any l, stepsize $\eta > 0$, $\alpha > 0$, $z_{t+1} = \arg\min_{z \in \mathcal{Z}} \eta(\langle l, z\rangle + \alpha g(z)) + B_\psi(z; z_t) \in \text{int dom } \psi$.
- F is monotone and L-smooth with respect to $\|.\|$ and g is 1-strongly convex relative to ψ over $\mathcal{Z}$ with g differentiable over int dom ψ.

Algorithm 1 provides the following convergence guarantees,

MMD attempts to unify Single-agent and Multi-agent RL by designing a single algorithm that performs competitively in both problem settings. In single-agent RL MMD's performance is competitive with that of PPO. And, in the multi-agent setting the performance of tabular MMD is as good as CFR, but worse than CFR+.

CHAPTER 4

EXTENDING MMD WITH NEURD FIX, EXTRAGRADIENT, AND OPTIMISM

4.1 Faster-MMD

4.2 Neural Replicator Dynamics

4.2.1 FMMD-N

4.3 Extragradient methods

4.3.1 FMMD-EG

4.4 Optimism

4.4.1 Optimistic Mirror Descent

4.4.2 OFMMD

APPENDICES

Appendix A

SOME ANCILLARY STUFF

Ancillary material should be put in appendices.

Appendix B

SOME MORE ANCILLARY STUFF

[2]

CITED LITERATURE

1. Sokota, S., D'Orazio, R., Kolter, J. Z., Loizou, N., Lanctot, M., Mitliagkas, I., Brown, N., and Kroer, C.: A Unified Approach to Reinforcement Learning, Quantal Response Equilibria, and Two-Player Zero-Sum Games. In *The Eleventh International Conference on Learning Representations* , February 2023.

2. Farine, D. R., Strandburg-Peshkin, A., Couzin, I. D., Berger-Wolf, T. Y., and Crofoot, M. C.: Individual variation in local interaction rules can explain emergent patterns of spatial organization in wild baboons. *Proceedings of the Royal Society of London B: Biological Sciences* , 284(1853), 2017.