Quadratic Speedup in Quantum Zero-Sum Games via Single-Call Mirror-Prox Matrix Methods

Francisca Vasconcelos[1], Emmanouil-Vasileios Vlatakis-Gkaragkounis[1]
Panayotis Mertikopoulos[2], Georgios Piliouras[3], and Michael I. Jordan[1]
[1]UC Berkeley
[2]Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, LIG, 38000 Grenoble, France
[3]Singapore University of Technology and Design
francisca@berkeley.edu, emvlatakis@berkeley.edu

Since its conception by Von Neumann in the early 20th century [VN28], game theory has become a staple of modern mathematics, economics, and computer science. Of great importance in non-cooperative games are Nash equilibria [Nas51], representing optimal play strategies, by which no player benefits from unilaterally changing their strategy. Despite their significance, determining strategies that achieve these equilibria remains a hard computational challenge. However, among the many game types, zero-sum games (where one player's win is the other's loss) have proven more computationally tractable, with many applications – including operational research for optimal resource allocation [PR16], politics [Bra75], and training AI models for games like chess or Go [SHM$^+$16] – underscoring their appeal in the realm of game theory.

In the quantum setting, zero-sum games have proven useful for various theoretical computational and cryptographic tasks, from non-local games [KLVY23, CHTW04] to multi-prover quantum interactive proofs [JJUW11] to quantum coin-flipping [Mil22, Amb01]. In Quantum Generative Adversarial Networks (QGANs) [CHL$^+$19] a generator and a discriminator quantum circuit engage in a complex zero-sum game to reach an equilibrium capable of faithfully replicating a target data distribution. Finally, following standard reductions, even constrained minimization problems like Quantum Support Vector Machines (QSVMs) [RML14] and quantum error correction methods [Ter15] can be formulated as quantum min-max optimization problems. In this work, we ask:

How can machine learning principles help us compute Nash equilibria in quantum zero-sum games using limited quantum states?

We revisit one of the original algorithmic proposals for quantum games. In 2008, Jain and Watrous proposed a parallel algorithm that determines Nash equilibria of non-interactive, non-collaborative zero-sum quantum games [JW09]. Within this setting, two competing players, Alice and Bob, each transmit a mixed quantum state to a referee (henceforth referred to as Roger), who performs a joint measurement on both states. Roger also posses a utility function that attributes payoffs to the players based on the measurement outcomes. Since it is a zero-sum game, Bob's payoff will be the negative value of Alice's – e.g. if Alice wins \$5, Bob will lose \$5.

Concretely, Alice and Bob, "play" the game by independently preparing multi-qubit quantum states α and β, respectively, to be sent to Roger. It is important to note that these are *mixed* states, corresponding to "mixed strategies," and that Alice and Bob do not share entanglement. Thus, Alice and Bob's set of possible "moves" is constrained to the set of density matrices lying in their respective *spectraplexes*. Denote Alice's spectraplex as $\mathcal{A} = \{\alpha \in \mathcal{H}_+^n : \mathrm{Tr}(\alpha) = 1\}$ and Bob's spectraplex as $\mathcal{B} = \{\beta \in \mathcal{H}_+^m : \mathrm{Tr}(\beta) = 1\}$. Denote the different concatenations of Alice and Bob's states as $\Psi^{\otimes} = \alpha \otimes \beta$ and $\Psi = (\alpha, \beta)^\top$, for all (α, β) lying in the joint space $\mathcal{C} = \mathcal{A} \times \mathcal{B}$.

Assuming Alice sends mixed state α and Bob sends β, Roger gives each player a payoff determined by the the bilinear mapping $\phi : \mathcal{H}_+^n \times \mathcal{H}_+^m \to \mathbb{R}$. Since the game is zero-sum, if Alice receives payoff $\phi(\alpha, \beta)$, Bob receives payoff $-\phi(\alpha, \beta)$. In playing the game, Alice and Bob's objectives are to maximize their payoff. Mathematically, Alice aims to produce a state $\alpha^*(\beta) = \arg\max_\alpha \phi(\alpha, \beta)$, $\forall \beta \in \mathcal{H}_+^m$, and Bob aims to produce a state $\beta^*(\alpha) = -\arg\max_\beta \phi(\alpha, \beta) = \arg\min_\beta \phi(\alpha, \beta)$, $\forall \alpha \in \mathcal{H}_+^n$. Note that $(\alpha^*(\beta^*), \beta^*(\alpha^*))$ correspond to a Nash equilibrium, which by the SDP version of

Manuscript submitted for review to the 7th International Conference on Quantum Techniques in Machine Learning (QTML 2023) .

Von-Neumann's Min Max Theorem admits unique game value $v \in \mathbb{R}$ [ITT22],

$$\max_{\alpha} \min_{\beta} \phi(\alpha, \beta) = v = \min_{\beta} \max_{\alpha} \phi(\alpha, \beta). \qquad \text{(Quantum Min-Max Theorem)}$$

Relating back to our algorithm design task, [JW09] employed the Matrix Multiplicative Weights Update (MMWU)[1] algorithm [TRW05, AHK12], to achieve an $O(\frac{1}{\sqrt{T}})$ convergence rate to such Nash equilibria when Roger returns to Alice and Bob respectively their full feedback –the payoff vector Feedback$(\Psi) = (\nabla_{\alpha^\top} \phi(\Psi), \nabla_{\beta^\top} \phi(\Psi))$.

However, even in classical games, the standard Multiplicative Weights Update algorithm fails to achieve the optimal convergence rate. Recently, [JPS22] introduced the notion of exponential quantum replicator dynamics (exp-QREP), a continuous-time quantum extension of the exponential formulation of replicator dynamics (referred to as just quantum replicator dynamics in [JPS22]). Importantly, [JPS22] demonstrate Poincaré recurrence of discretized MMWU in quantum zero-sum games for exp-QREP dynamics. In parallel, [LMB23] proposed a quantum games variant of the *folk Game Dynamics Theorem*, stating that *no-regret algorithms, like MMWU, only converge stably to* ***pure*** *quantum states*. In response to these intrinsic difficulties in successful last iterate convergence for quantum games, we ask:

Could there be a gradient-based method that not only mimics the ergodic behavior of MMWU, in quantum zero-sum games, but also outperforms it?

In classical literature regarding monotone variational inequalities with Lipschitz continuous operators, the optimal convergence rate was firmly established as $O(1/T)$ by Korpelevich's Extra-Gradient (EG) algorithm [Kor76] and its Bregman adaptation, the Mirror-Prox (MP) algorithm by Nemirovski [JNT11]. However, these methodologies have a notable drawback – they require not one, but two gradients per iteration. Since each gradient calculation in a quantum game requires preparation of a novel quantum state, which can be highly resource-demanding, decreasing the number of oracle calls could offer substantial practical advantages. Therefore, our work aims to: (i) devise a simplified approach that retains the performance of the Extra-Gradient algorithm, but uses **only one oracle call per iteration**, and (ii) quantify convergence rates for these **single-call matrix mirror-prox** (1-MMP) algorithms.

MATRIX DUAL AVERAGING (MDA) METHOD

We begin by presenting the Matrix Dual Averaging (MDA) method, whose instantiation is actually a variant of the MMWU method of [JW09]. In words, at each stage, both players play a "soft" version of the best mixed state response, regularized by a convex penalty (or "distance generating") function, and then accumulate Roger's feedback. For example, if the players' cumulative feedback is denoted by $C\mathcal{F}$ and the penalty function is the squared Frobenius norm $h(X) = (1/2)\|X\|_F^2$, their response would correspond to the orthogonal projector $\Pi_{\mathcal{X}}(C\mathcal{F}) = \arg\min_{X \in \mathcal{X}} \|X - C\mathcal{F}\|_F^2$. On the other hand, if players employ the negative Von Neumann entropy $h(X) = \mathrm{Tr}[XlogX]$ on the spectraplex, they induce logit map $\Lambda_{\mathcal{X}}(C\mathcal{F}) = \frac{\exp(C\mathcal{F})}{\mathrm{Tr}(\exp(C\mathcal{F}))}$ and, thus, the MMWU algoirhtm of [JW09]. For this family of algorithms (See Algorithm 1), it is shown in [JW09] that

Theorem 1 (MDA Rate). *MDA methods compute an ϵ-Nash equilibrium in $O(1/\epsilon^2)$ steps.*

SINGLE-CALL MATRIX MIRROR-PROX METHODS

Theorem 2 (Main Result). *1-MMP methods compute an ϵ-Nash equilibrium in $O(1/\epsilon)$ steps.*

In the context of classical games, the first Extra-Gradient variant with a single oracle call per iteration dates back to Popov [Pop80]. This algorithm was subsequently studied by, among others, Rakhlin

[1]Beyond quantum zero-sum games, MMWU has also been used for many important quantum applications such as solving SDPs [AK07], proving QIP=PSPACE [JJUW11], finding balanced separators [CKGS23] , and enhancing spectral sparsification [AZLO15].

and Sridharan [RS13]. Our methodology leverages the universal proof technique for monotone variational inequalities by [EN22]. Notably, our proof technique deviates from classic simplex case due to the absence of a Farkas lemma and the non-metric subregularity in spectraplexes. In constrast with the commutable inner products of vectors found in classical normal form games, the quantum operators in quantum games are inherently non-commutative. However, our proposed approach (Algorithm 2) handles these quantum non-Abelian operators and we prove their monotonicity via techniques from quantum analysis of Boolean functions (as originally proposed by [MO10]) to manipulate Pauli decompositions of the involved operators. Our analysis also encompasses a novel variant of the Optimistic Matrix Multiplicative Weights Update (OMMWU) algorithm, that we define below similar to [DP18], circumventing the issue of non-commutativity in quantum games:

$$\begin{pmatrix}\alpha_{t+1}\\ \beta_{t+1}\end{pmatrix} := \begin{pmatrix}\Lambda_{\mathcal{X}}(\log\alpha_t - 2\eta\text{Feedback}_t^{[Alice]} + \eta\text{Feedback}_{t-1}^{[Alice]})\\ \Lambda_{\mathcal{X}}(\log\beta_t - 2\eta\text{Feedback}_t^{[Bob]} + \eta\text{Feedback}_{t-1}^{[Bob]})\end{pmatrix}$$

We conclude with an interesting open problem/conjecture:

Conjecture 1. *Theorem 2's rates are tight. In other words, there exists a quantum game such that 1-MMP methods need in $\Omega(1/\epsilon)$ steps for computing an ϵ-Nash Equilibrium.*

ALGORITHM 1: Matrix Dual Averaging

Input: Accuracy Parameter: ϵ, Convex Regularizer $h_{\mathcal{A},\mathcal{B}}(\cdot) : \{\mathcal{A},\mathcal{B}\} \to \mathbb{R}$,
Feedback$^{[Alice,Bob]} := (\nabla_{\alpha^\top}\phi(\Psi), \nabla_{\beta^\top}\phi(\Psi))$// Roger's Feedback.
Define $\eta := \epsilon/100$ and $N := \lceil 100\log(nm)/\epsilon^2\rceil$,$(A_0, B_0) :=$ RandomSDP$_n$, RandomSDP$_m$,
$(\alpha_0, \beta_0) := (A_0/\text{Tr}(A_0), B_0/\text{Tr}(B_0))$;
Let $\text{Breg}\Pi_{\mathcal{X}}^h(\cdot) = \arg\max_{x\in\mathcal{X}}\{\langle\cdot,\alpha\rangle - h_{\mathcal{X}}(x)\}$ // Bregman Projection at $\mathcal{X}$;
for $t \in [0 : N-1]$ **do**
▹ Alice and Bob **play** their quantum mixed-state strategies (α_t, β_t) to Roger
▹ Roger **sends** back its feedback to Alice and Bob.
▹ Alice and Bob **aggregate approximately their payoff** based on Roger's feedback.
$\text{CumPayoff}_t^{[Alice]} = \sum_{k=0}^{t-1}\text{Feedback}^{[Alice]}(\Psi_k)$, $\text{CumPayoff}_t^{[Bob]} = \sum_{k=0}^{t-1}\text{Feedback}^{[Bob]}(\Psi_k)$;
▹ Alice and Bob **prepare their new mixed states** based on Best Regularized Response.
$\alpha_{t+1} = \text{Breg}\Pi_{\mathcal{A}}^{h_{\mathcal{A}}}(\eta\text{CumPayoff}_t^{[Alice]})$, $\beta_{t+1} = \text{Breg}\Pi_{\mathcal{B}}^{h_{\mathcal{B}}}(\eta\text{CumPayoff}_t^{[Bob]})$;
end
Output: $\bar\alpha = \frac{1}{N}\sum_{i=0}^{N-1}\alpha_i, \bar\beta = \frac{1}{N}\sum_{i=0}^{N-1}\beta_i$

ALGORITHM 2: Single-Call Matrix Mirror Prox (1-MMP)

Input: Same as Algorithm 1.
Define $\eta := 1/100$ and $N := \lceil 100\log(nm)/\epsilon^2\rceil$, $(A_0, B_0) :=$ RandomSDP$_n$, RandomSDP$_m$,
$(\alpha_0, \beta_0) = (\hat\alpha_0, \hat\beta_0) := (A_0/\text{Tr}(A_0), B_0/\text{Tr}(B_0))$;
Let $\text{Mirror}\Pi_{\mathcal{X}}^h(t, y) = \arg\max_{x\in\mathcal{X}}\{\langle t,\alpha\rangle - (h_{\mathcal{X}}(x) - h_{\mathcal{X}}(y) - \partial h_{\mathcal{X}}(y)(x-y))\}$;
▹ Alice and Bob **request** the feedback of (α_0, β_0) from Roger.
for $t \in [0 : N-1]$ **do**
▹ Alice and Bob **prepare their new mixed states** based on Best Regularized Response.
$\alpha_{t+1} = \text{Mirror}\Pi_{\mathcal{A}}^{h_{\mathcal{A}}}(\eta\text{Feedback}_t^{[Alice]}, \hat\alpha_t)$, $\beta_{t+1} = \text{Mirror}\Pi_{\mathcal{B}}^{h_{\mathcal{B}}}(\eta\text{Feedback}_t^{[Bob]}, \hat\beta_t)$;
▹ Alice and Bob **play** their quantum mixed-state strategies $(\alpha_{t+1}, \beta_{t+1})$ from Roger.
▹ Roger **sends** back its feedback to Alice and Bob.
▹ Alice and Bob **prepare their optimistic-momentum** mixed states Response.
$\hat\alpha_{t+1} = \text{Mirror}\Pi_{\mathcal{A}}^{h_{\mathcal{A}}}(\eta\text{Feedback}_{t+1}^{[Alice]}, \hat\alpha_t)$, $\hat\beta_{t+1} = \text{Mirror}\Pi_{\mathcal{B}}^{h_{\mathcal{B}}}(\eta\text{Feedback}_{t+1}^{[Bob]}, \hat\beta_t)$;
end
Output: $\bar\alpha = \frac{1}{N}\sum_{i=0}^{N-1}\alpha_i, \bar\beta = \frac{1}{N}\sum_{i=0}^{N-1}\beta_i$

REFERENCES

[AHK12] Sanjeev Arora, Elad Hazan, and Satyen Kale. The Multiplicative Weights Update Method: a Meta-Algorithm and Applications. *Theory of Computing*, 8(6):121–164, 2012.

[AK07] Sanjeev Arora and Satyen Kale. A Combinatorial, Primal-Dual Approach to Semidefinite Programs. In *Proceedings of the Thirty-Ninth Annual ACM Symposium on Theory of Computing*, STOC '07, page 227–236. ACM, 2007.

[Amb01] Andris Ambainis. A New Protocol and Lower Bounds for Quantum Coin Flipping. In *Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing*, STOC '01, page 134–142, New York, NY, USA, 2001. Association for Computing Machinery.

[AZLO15] Zeyuan Allen-Zhu, Zhenyu Liao, and Lorenzo Orecchia. Spectral Sparsification and Regret Minimization Beyond Matrix Multiplicative Updates. In *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, STOC '15, page 237–245, New York, NY, USA, 2015. Association for Computing Machinery.

[Bra75] Steven J. Brams. *Game theory and politics*. Free Press, 1975.

[CHL+19] Shouvanik Chakrabarti, Yiming Huang, Tongyang Li, Soheil Feizi, and Xiaodi Wu. Quantum wasserstein GANs. In *Proceedings of the 33rd International Conference on Neural Information Processing Systems*, NIPS '19, pages 6781 – 6792, 2019.

[CHTW04] Richard Cleve, Peter Høyer, Benjamin Toner, and John Watrous. Consequences and Limits of Nonlocal Strategies. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 236–249. IEEE Computer Society, 2004.

[CKGS23] Li Chen, Rasmus Kyng, Maximilian Probst Gutenberg, and Sushant Sachdeva. A Simple Framework for Finding Balanced Sparse Cuts via APSP. In *Symposium on Simplicity in Algorithms (SOSA)*, pages 42–55. SIAM, 2023.

[DP18] Constantinos Daskalakis and Ioannis Panageas. Last-Iterate Convergence: Zero-Sum Games and Constrained Min-Max Optimization. In Avrim Blum, editor, *10th Innovations in Theoretical Computer Science Conference (ITCS 2019)*, volume 124 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 27:1–27:18, Dagstuhl, Germany, 2018. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[EN22] Alina Ene and Huy Lê Nguyên. Adaptive and Universal Algorithms for Variational Inequalities with Optimal Convergence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(6):6559–6567, Jun. 2022.

[ITT22] Constantin Ickstadt, Thorsten Theobald, and Elias Tsigaridas. Semidefinite games. *arXiv preprint arXiv:2202.12035*, 2022.

[JJUW11] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. QIP = PSPACE. *Journal of the ACM*, 58(6):30:1–30:27, 2011.

[JNT11] Anatoli Juditsky, Arkadi Nemirovski, and Claire Tauvel. Solving Variational Inequalities with Stochastic Mirror-Prox Algorithm. *Stochastic Systems*, 1(1):17–58, 2011.

[JPS22] Rahul Jain, Georgios Piliouras, and Ryann Sim. Matrix Multiplicative Weights Updates in Quantum Zero-Sum Games: Conservation Laws & Recurrence. In *Proceedings of the 36th International Conference on Neural Information Processing Systems*, NIPS '22, 2022.

[JW09] Rahul Jain and John Watrous. Parallel Approximation of Non-interactive Zero-sum Quantum Games. In *2009 24th Annual IEEE Conference on Computational Complexity*, pages 243–253, 2009.

[KLVY23] Yael Kalai, Alex Lombardi, Vinod Vaikuntanathan, and Lisa Yang. Quantum Advantage from Any Non-local Game. In Barna Saha and Rocco A. Servedio, editors, *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023, Orlando, FL, USA, June 20-23, 2023*, pages 1617–1628. ACM, 2023.

[Kor76] Galina M Korpelevich. The extragradient method for finding saddle points and other problems. *Matecon*, 12:747–756, 1976.

[LMB23] Kyriakos Lotidis, Panayotis Mertikopoulos, and Nicholas Bambos. Learning in quantum games, 2023.

[Mil22] Carl A Miller. The Mathematics of Quantum Coin-Flipping. *Notices of the American Mathematical Society*, 2022.

[MO10] Ashley Montanaro and Tobias J. Osborne. Quantum Boolean Functions . *Chicago Journal of Theoretical Computer Science*, (1), January 2010.

[Nas51] John Nash. Non-Cooperative Games. *Annals of Mathematics*, 54(2):286–295, 1951.

[Pop80] Leonid Denisovich Popov. A modification of the Arrow-Hurwicz method for search of saddle points. *Mathematical notes of the Academy of Sciences of the USSR*, 28(5):845–848, November 1980.

[PR16] Parvathy S. Pillai and Shrisha Rao. Resource Allocation in Cloud Computing Using the Uncertainty Principle of Game Theory. *IEEE Systems Journal*, 10(2):637–648, 2016.

[RML14] Patrick Rebentrost, Masoud Mohseni, and Seth Lloyd. Quantum support vector machine for big data classification. *Physical Review Letters*, 113(13):130503, 2014.

[RS13] Alexander Rakhlin and Karthik Sridharan. Online Learning with Predictable Sequences. In Shai Shalev-Shwartz and Ingo Steinwart, editors, *Proceedings of the 26th Annual Conference on Learning Theory*, volume 30 of *Proceedings of Machine Learning Research*, pages 993–1019, Princeton, NJ, USA, 12–14 Jun 2013. PMLR.

[SHM+16] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016.

[Ter15] Barbara M. Terhal. Quantum error correction for quantum memories. *Reviews of Modern Physics*, 87:307–346, Apr 2015.

[TRW05] Koji Tsuda, Gunnar Rätsch, and Manfred K. Warmuth. Matrix Exponentiated Gradient Updates for On-line Learning and Bregman Projection. *Journal of Machine Learning Research*, 6:995–1018, December 2005.

[VN28] John Von Neumann. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, 100(1):295–320, December 1928.