Quantum Theoretic Values of Collaborative and Self-organizing Agents

Ying Zhao, and Charles Zhou

Abstract—When multi-agents collaborate to perform distributed operations, they can be modeled as cooperative games. Considering a network of agents work together and they can only communicate in a limited way (e.g., only to neighbor peers), the goal is to maximize the cooperation success globally, or maximize the total value and social welfare of the whole network. The type of cooperation is challenging since the game is not zero-sum. There are not any outside agents to serve as referees. The objective functions may be non-stationary and non-convex. In this paper, each agent is modeled as a content supplier or consumer. Each agent optimizes its own objective locally. We show that each agent self-organizes or converges to its "value" via the principles of quantum computing and game theories. We prove two theorems that can optimize an agent's own objective and simultaneously optimize the global social welfare of its peer network. The quantum intelligence game algorithms are unsupervised and self-organizing, where the weights expressed in quantum neural networks can be computed from a natural mechanism known as a quantum adiabatic evolution.

Index Terms—collaborative learning agents, self-organizing, lexical link analysis, LLA, quantum machine learning, LLA quantum intelligence game, LLAQIG, quantum adiabatic evolution, QAE, social welfare

1 Introduction

A collaborative learning agent(CLA) [19] is the basic architecture focused in this paper. A single agent represents a single system capable of ingesting data, indexing, cataloging information, and performing knowledge and pattern discovery, machine learning (ML) from data, and separating patterns and anomalies from data.

Multiple CLAs work collaboratively in a peer-to-peer network. Each agent has a peer list. A CLA first mines the data (structured and unstructured data sources) locally and then fuses the models from its peers. Therefore, the models and indexes are available for the whole peer network.

Considering ranking a node using traditional network theory, the importance of a node in a network can be ranked using established hyperlinks, citation networks, social networks, or other collective intelligence marked by humans. However, few or no hyperlinks are available for private or proprietary data. Furthermore, high-value information can be different from applications. Current methods mainly score patterned information, which are useful for certain types of applications, e.g., marketing applications. Anomalous information is important for some other types of applications, e.g., intelligence analysis and novelty search.

- *Y. Zhao is with the Department of Information Sciences, Naval Postgraduate School, Monterey, CA. E-mail: yzhao@nps.edu*
- *C. Zhou is with Quantum Intelligence, Inc. E-mail: charles.zhou@quantumii.com*

ASONAM '23, November 6–9, 2023, Kusadasi, Turkiye

ACM ISBN 979-8-4007-0409-3/23/11. . . 15.00
https : //*doi.org*/10.1145/3625007.3627509

When a peer CLA is a content provider and its peers are the content consumers in a peer-to-peer network, their interaction can be modeled as a strategic cooperation game of two players. The content provider's search for the best value of itself as a Nash equilibrium may not achieve a full Pareto efficiency or the so-called optimal social welfare for the whole peer network, referred as the Prisoner's dilemma.

In this paper, we show it is essential to apply quantum computing and game theory to discover a CLA's true value. We prove that the Prisoner's dilemma is naturally escaped via a quantum adiabatic process, which allows a CLA in a peer-to-peer network reach both the Nash equilibrium and optimal social welfare [23].

Fig. 1 shows a schematic network of CLAs used in a peer-to-peer network. Each unit or node is represented as a single CLA. A unit can be a content supplier, consumer, or broker. When a unit receives a request/demand of its capabilities, it searches its peer network and finds the best match to fulfil the request. The major contributions of this paper are summarized as follows:

1) We model a network of CLAs as a demand and supply network in a cooperative game. We show each agent can self-organize or converge to its maximum "value" and meanwhile maximize social welfare via quantum principles.
2) We integrate CLAs and lexical link analysis (LLA), quantum computing, and quantum game, into the LLA quantum intelligence game (LLAQIG) algorithms. The LLAQIG reaches the Nash Equilibrium for each individual agent or unit to maximize its value, meanwhile, achieves the total social welfare for the whole peer-to-peer system in an unsupervised and self-organizing learning.
3) We prove two theorems that the LLAQIG algorithm is unsupervised or self-organizing, where the value of

an agent (to its peer network) with new information is formed over time and modeled as a changing Hamiltonian in a Schrodinger's equation. Such unsupervised learning is about optimizing energy or entropy which can be realized by natural mechanisms such as laws of physics and dynamics.

The rest of this paper is organized as follows. Section 2 reviews relevant literature. Section 3 describes the main methods used in our system. Section 4 describes results in terms of preliminaries and proofs of two theorems. Section 5 discusses the results while Section 6 concludes the paper.

2 Literature Review

2.1 Quantum Game (QG) and Quantum Advantages

We consider a typical game with a self-player and opponent that is set up so each player may take mixed strategies. In a classical game, a mixed strategy is a probabilistic convex combination of pure strategies; while in a quantum game, a quantum mixed state is represented by a density matrix. A quantum game is not a "tensorial" extension of a finite game [1]. A quantum game is a continuous game where each player can generate a quantum density matrix.

The quantum advantage has a profound impact on computer science and machine learning, such as quantum generative adversarial networks (QGANs) and adversarial learning [2], [3]. Quantum players can have distinct advantages over classical players, achieving higher payoffs at equilibrium than classical counterpart due to the unique intricate interactions that the classical counterpart does not have. For example, in classical games, a mixed strategy is a probabilistic convex combination of pure strategies; while in a quantum game, a mixed state is a superposition of pure quantum states which can result in payoffs that lie outside the convex hull of classical mixed strategies [1].

2.2 Nash Equilibrium (NE)

In a classic or quantum game, Nash equilibrium characterizes a strategy ψ^* discourages unilateral deviations such that

$$u_k(\psi^*) \geq u_k(\psi_k; \psi^*_{-k}) \quad (1)$$

for all ψ_k and k. $(\psi_k; \psi_{-k})$ is the choice of player k relative to all other players $-k$. The existence of Nash equilibrium follows the seminal theorem [4].

2.3 Correlated Equilibrium (CE)

In a classic game theory, a correlated equilibrium is a solution concept generalization of Nash equilibrium. A mixed Nash equilibrium in the classic game theory is a distribution on the strategy space that is "uncorrelated" (i.e., the product of independent distributions, one for each player), while a correlated equilibrium is a general distribution over strategy profiles, where changing the p-th component strategy profile can affect other components [5]. A mixed Nash equilibrium is a special case correlated equilibrium that happens to be a product distribution. Firstly, a quantum game with a mixed state is a superposition of pure quantum states which could result in better payoffs than a CE's of classical mixed strategies. Secondly, a quantum game is to encode players' actions in qubits and then let a "referee" – a natural mechanism, i.e., an unsupervised learning or self-organizing algorithm such as a quantum adiabatic evolution (QAE, see Section 4.1.5) to determine the payoffs of a quantum game.

2.4 Quantum Neural Networks (QNN) and Fourier Transformation [10]

An emerging trend in machine learning and artificial intelligence (ML/AI) is to investigate if quantum computing can offer quantum advantages, giving rise to an interdisciplinary area called quantum machine learning [6]. For example, QNN are quantum analogs of artificial neural networks, which are extended classes of functions that can approximate or as learning representation of the underlying structures. Fourier series are used to justify multi-qubit QNN models that can realize universal function approximators (UFA) [7]. It is suggested that a single-qubit native QNN can express any Fourier series, which is a universal approximator for any square-integrable univariate function. Furthermore, the theorems and proofs in this research result in a QNN model that is a quantum circuit that consists of interleaved data encoding blocks $U(C, \gamma)$ and trainable coefficient $U(B, \beta)$. $U(C, \gamma)$ and $U(B, \beta)$ are both

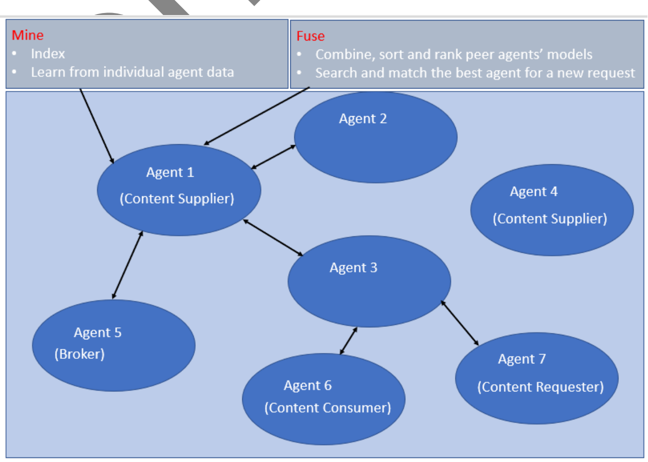

Fig. 1. Collaborative Learning Agents (CLAs): Each agent can be a content supplier, requester, consumer, or broker. The patterns of content, data, links in each unit are indexed and learned from historical data. When an agent receives a new request, it searches its peer network and finds the best match to fulfil the request.

unitary matrices (also see Section 4.1.5). For a measurable H, the output of the QNN is the optimization $\langle\psi| H |\psi\rangle$, where ψ is a wavefunction is or a quantum state Although, the research gives an explanation why layered QNN is a good UFA, however, the parameters in the trainable circuit blocks are learned in a typical supervised ML process.

3 Methods

We apply CLA, LLA, and LLAQIG as a whole process for discovering quantum game-theoretic value of collaborative and self-organizing agents. We briefly describe CLA and LLA in this paper and focus on the LLAQIG.

3.1 Mining and Fusing in a Single CLA

As shown in Fig. EABO, a single CLA represents a single agent capable of ingesting data, analyzing data, and separating patterns and anomalies from data.

Multiple CLAs work collaboratively in a peer-to-peer network. Each agent has a peer list. In a more detail, a CLA first indexes structured and unstructured data locally using unsupervised ML and data mining algorithms, and then fuses the local models with the models of its peers. Therefore, the models and indexes are available for the whole peer network. The collaboration of a network of CLAs is achieved through a peer list defined within each agent, through which each agent passes shared information only to its peers.

Each CLA j includes an analytic engine with two algorithms, i.e., a ML (Mine) and fusion (Fuse) algorithm, which can be customized externally.

- The Mine algorithm integrates the local knowledge base $k(t, j)$ and global knowledge base $K(t-1, j)$, which is only global to its peer network, into a new knowledge base $K(t, j)$.
- The Fuse algorithm assesses the total value of Agent j by separating the total knowledge base into the categories of patterns, emerging, and anomalies, and predict a total value $V(t, j, c)$ for each piece c of the content in Agent j.

The whole process is illustrated in Algorithm 1, where $p(j)$ represents the peer list of Agent j. The total value $V(t, j, c)$ is used in the global sorting and ranking of relevant content c in Agent j when it publishes the content to its peer network. Algorithm 1 can run continuously over time and parallel for each Agent j.

Algorithm 1 Mine and Fuse in a Single CLA j

while $t \geq 0$ **do**
  for each i in Agent j's peer list $p(j)$ **do**
    $K(t, j) \Leftarrow Mine(K(t-1, i), k(t, j))$
  end for
  for each content $c(j)$ in Agent j **do**
    $V(t, j, c(j)) \Leftarrow Fuse(K(t, j), c(j))$
  end for
end while=0

The output of the Mine algorithm is the statistically significant association matrix using LLA (see Section 3.2) for all word feature pairs $w_{lm}, l = 1, ..., W; m = 1, ..., W$, where W is the total number of unique word features in Agent j's peer network in Eq. (2):

$$K(t,j) = \begin{bmatrix} k_{11}(t,j) & ... & \\ ... & k_{lm}(t,j) & ... \end{bmatrix} \quad (2)$$

where k_{lm} is a z-score [21] and $k_{lm} > 1.96$. The link between word feature w_l and w_m has a statistically significant and causal link $p - value < 0.05$.

The Fuse algorithm for Agent j is then used to sort and rank the content or compute the value $V(t, j, c)$ for each content c based on its knowledge base $K(t, j)$ at time t, and then aggregate the scores to rank Agent j.

The $K(t, j)$ is a primitive matrix related to the Perron-Frobenius theorem [27]: If a matrix is non-negative (i.e., all its elements are non-negative real numbers) and its mth power is positive (i.e., all its elements are positive) for some natural number m and the same m works for all pairs of indices, then eigenvalue with the maximum magnitude of $B(t, j)$ or $B(t, j)$'s spectrum radius is positive, i.e., $\lambda_{max} > 0$.

3.2 Lexical Link Analysis (LLA)

In an LLA, a complex system can be expressed in a list of attributes or features with specific vocabularies or lexicon terms to describe its characteristics. LLA automatically discovers word features, clusters of word features, and displays them as word feature networks. LLA is related to but significantly different from so called bag-of-words (BOW) methods such as Latent Dirichlet Allocation (LDA) [20].

Bi-gram allows LLA to be extended to numerical, categorical, or time series data. For example, for structured data such as attributes from databases, LLA discretizes and then categorizes attributes and their values to word-like features. LLA computes the counterfactual proportion difference as the strength of two associated items for Agent j. The word features in LLA are then clustered into groups or themes using the community detection algorithms [28].

- Patterned (P) themes: A patterned theme is more likely to be shared across multiple diversified domains, which are already in the public consensus and awareness and can be authoritative.
- Anomalous (A) themes: These themes may not seem to belong to the data domain as compared to others. They can be interesting, unique, innovative to specific entities and may be high-value and need further investigation.

Let the value of a content c computed from patterned (P) themes be $P(t, j, c)$ and from anomalous (A) themes be $A(t, j, c)$ of LLA for Agent j at time t, respectively. The total value $V(t, j, c)$ for c is a function of $P(t, j, c)$ and $A(t, j, c)$ as shown in Eq. (3).

$$V(t, j, c) = f(P(t, j, c), A(t, j, c)) \quad (3)$$

Assume we can compute P, A for each piece of information c in a CLA, we show how to apply the principle of quantum computing and quantum game to discover the function f or the value of the content c in Eq. (3) using the natural mechanism and process of QAE. In other words, although f may be approximated using the single-qubit QNN function approximator theory as a supervised learning, our method learns the function in a unsupervised fashion. We further show that the method allows each agent reach the Nash Equilibrium to maximize its own value, meanwhile, achieve the optimal total social welfare.

3.3 Game-theoretic Properties of LLA

In a traditional game theory, a player's search for its own best value is modeled as to reach a Nash equilibrium. However, with multiple cooperative players, the whole system may not achieve a full Pareto efficiency or optimal social welfare for all other players, referred as the Prisoner's dilemma. When comparing one Nash equilibrium to another, the one that has a higher social welfare is a Pareto superior solution.

When a self-player acts as a content provider to seek value from a cooperative opponent, the self-player has to be also Pareto efficient or superior in order to obtain the right response from the opponent, i.e., the content consumers or crowdsourcing audience. Being Pareto efficient means the system can not make at least one player better off without making any other player worse off, i.e., achieving the so-called optimal social welfare.

The value of a content is defined or naturally realized to achieve Nash equilibrium, i.e., maximizes its own payoff, meanwhile, is Pareto efficient. $K(t, j)$ is computed for an agent's value with its opponent value, i.e., the overlapping word features an agent has with respect to the population. For example, the opponent of a content provider is the population of content consumers or the crowdsourcing audience. Assume the total set of word features for a peer network of CLAs is categorized into patterned and anomalous themes. The number of word features for a CLA in the two categories is P and A, respectively. We show a connection between LLA, quantum computing, and quantum game theory so both requirements are satisfied. We show that it is necessary for a content to realize its value (i.e., to combine P and A) through a quantum process which can meet the two requirements simultaneously.

4 Results

4.1 Preliminaries

4.1.1 *Probability Amplitudes and Quantum Wavefunction*

In a classical situation of uncertainty, a weighted combinations of event A and B, i.e., $w \times A$ plus $v \times B$ where w, v are probabilities and $w + v = 1$. In quantum mechanics, since p and q are complex numbers, they are probability amplitudes. However, $(p+q)^\dagger(p+q) \neq p^\dagger p + q^\dagger q$. And [12]

$$(p+q)^\dagger(p+q) = p^\dagger p + q^\dagger q + q^\dagger p + p^\dagger q \tag{4}$$

A wavefunction is $\psi(x)$ or a quantum state. $\psi(x)^\dagger\psi(x)$ is a probability amplitude of a quantum particle at a location x. The wavefunction ψ evolves in time in the Schrodinger's equation. ψ can be decomposed into pure states or Fourier bases. ψ is governed by the deterministic Schrodinger evolution, and the measurement process introduces uncertainty.

4.1.2 *Quantum Density Matrix*

In quantum mechanics, a density matrix (or density operator) is a matrix that describes the quantum state of a physical system, which allows to represent quantum mixed states.

A system of qubits has a mixed state $|\psi\rangle$ can be modeled as an uniform superposition over pure states z_j:

$$|\psi\rangle = \sum_z a_z |z\rangle \tag{5}$$

where a_z is a complex number. If pure states z_j occurs with probability amplitude p_j, then the corresponding density operator ρ in Eq.(6):

$$\rho = \sum_j p_j |z_j\rangle \langle z_j| \tag{6}$$

A density matrix is a positive semi-definite, or Hermitian operator of trace one acting on the Hilbert space of a quantum system. The trace one requirement is due to the probability amplitudes needed to add up to 1. Density matrix elements represent values of the eigenstate coeffcients for a quantum mixed state:

- Diagonal elements ($n = m$) give the probability amplitude of occupying a pure quantum state.
- Off-diagonal elements ($n \neq m$) are complex numbers and have a time-dependent phase factor that describes the evolution of coherent superpositions.

There is the difference between a probabilistic mixture of quantum states and their superpositions. For example, if a physical system is to be in two qubits quantum states $|\psi_1\rangle$ or $|\psi_2\rangle$ with probability λ and $1-\lambda$, it can be described by a density matrix:

$$\rho = \begin{bmatrix} \lambda & 0 \\ 0 & 1-\lambda \end{bmatrix} \tag{7}$$

On the other hand, a quantum superposition of the same pure states with probability amplitudes λ and $1-\lambda$ results in a new pure state $\lambda|\psi_1\rangle + (1-\lambda)|\psi_2\rangle$ with a density matrix:

$$\rho = \begin{bmatrix} \lambda & 1-\lambda \\ \lambda & 1-\lambda \end{bmatrix} \tag{8}$$

Unlike the probabilistic mixture, this superposition can display quantum interference [3] even the two pure states are orthogonal. For example, a convex superposition of two pure states may give rise to quantum interference terms of the form $|\psi_1\rangle\langle\psi_2|$ and $|\psi_2\rangle\langle\psi_1|$ in the induced payoff. These cross-terms have no analogue in traditional probability or game theory.

4.1.3 *Measurement*

Let H be an observable of the system, and suppose the ensemble is in a quantum mixed state $|\psi\rangle$ such that each of the pure states ψ_j occurs with probability p_j. Then the corresponding density operator equals Eq.(9)

$$\rho = \sum_j p_j |\psi_j\rangle \langle\psi_j| \tag{9}$$

The expectation value of the measurement can be calculated:

$$\langle\psi| H |\psi\rangle = tr(\rho H) \tag{10}$$

4.1.4 Evolution of Quantum States

The time evolution of a quantum state is described by the Schrödinger equation:

$$i\hbar \frac{d}{dt}\psi(t) = H\psi(t) \tag{11}$$

H is the Hamiltonian, whose observable corresponding to the total energy of the system, $\hbar$ is the reduced Planck constant. The Hamiltonian is an observable of particular importance, which not only defines the total energy of a system, but also determines how a system evolves in time, as in the time dependent Schrodinger equation.

4.1.5 Quantum Adiabatic Evolution (QAE)

In physics, a diabatic process can rapidly change conditions to prevent a system from adapting its configuration during the process, while an adiabatic process can gradually change conditions, which allow the system to adapt its configuration, hence the probability density is modified by the process. If the system starts in an eigenstate of the initial Hamiltonian, it will end in the corresponding eigenstate of the final Hamiltonian.

The QAE [14] is designed to study a slow evolution for a time dependent Hamiltonian $H(t) = (1-t/T)H_B + (t/T)H_C$. Starting with an initial energy H_B, if the run time T were long enough we would find the highest energy H_C. Because H_B has only non-negative off-diagonal elements, the Perron-Frobenius theorem implies that the difference in energies between the top state and the one below is greater than 0 for all $t < T$, so for sufficiently large T, success is assured. A Trotterized approximation to the evolution consists of an alternation of the operators $U(C, \gamma)$ and $U(B, \beta)$, where the sum of the angles is the total run time. For a good approximation, each γ and β have to be small and a long run time T.

A quantum circuit or QNN with $2p$ unitary operations $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_p)$ and $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$ can approximate the natural mechanism of QAE a $p - step$ Trotter expansion of the adiabatic evolution.

According to the adiabatic evolution theorem [14],one can slowly change the system's Hamiltonian from H_B to H_C and obtain the maximum energy state of H_C with high probability [15]. The changing process is exampled in below equation:

$$H(t) = [1 - s(t)]H_B + s(t)H_C, \tag{12}$$

where $s(t)$ is a smooth function, $0 \le s(t) \le 1$, $s(0) = 0$ and $s(T) = 1$.

$\langle\psi| H_C |\psi\rangle$ is the expectation value of system energy for a quantum system described by Hamiltonian H_C for a quantum state ψ. The best quantum state is a superposition of pure states which gives the ground state of H_C [14].

This can be realized by discretizing the total time interval $[0, T]$ into intervals $[j\Delta t, (j+1)\Delta t]$ with small enough Δt. Over the jth interval and $H(t) = H((j+1)\Delta t)$. Therefore, the total time evolution operator $U(T, 0)$ can be approximately discretized into $2p$ implementable operators as shown in Eq.(14). The approximation will improve as Δt gets smaller or, equivalently, as p increases.

$$U(T,0) \approx \prod_{j=1}^{p} e^{-iH(j\Delta t)\Delta t} \tag{13}$$

$$\approx \prod_{j=1}^{p} U_B^j U_C^j \tag{14}$$

where

$$U_C^j = e^{-i\gamma_j H_C} \tag{15}$$

$$= I\cos\gamma_j - iH_C \sin\gamma_j \tag{16}$$

$$U_B^j = e^{-i\beta_j H_B} \tag{17}$$

$$= I\cos\beta_j - iH_B \sin\beta_j \tag{18}$$

In the evolution, $|\psi\rangle$ represents the quantum state. Through applying U_B^j and U_C^j to the initial state alternately, one can compute the final state $|s(T)\rangle$, which is expected to collapse to the maximum energy state after measurement.

$$|\psi(T,\gamma,\beta)\rangle = \prod_{j=1}^{p} U_B^j U_C^j |\psi(0)\rangle \tag{19}$$

4.2 LLA Quantum Intelligence Game (LLAQIG)

Quantum Intelligence Game (QIG) focus on a form of QG in a cooperative game. LLAQIG is a catergory of ranking algorithms to determine the quality, value, or impact of a content from its LLA decomposition of P and A and quantum theories.

Theorem 1 (Superposition Value)**.** *A quantum game-theoretic value is computed as two pure strategies of P and A superpositioned with an initial entanglement θ, reaching the maximum value via an angle ϕ.*

Proof. A content provider (self-player) and its crowdsouring audience (opponent) starts with an initial degree of entanglement since their knowledge bases are correlated and overlapped that constitute the content propagation and value delivery. We first show the quantum relation geometrically in Fig. 2, similiar to a Fourier transformation,

Since $|P\rangle$ and $|A\rangle$ are entangled initially with θ and superpositioned with ϕ, among the four points (x_1, y_1), (x_2, y_2), (x_3, y_3), and (x_4, y_4) in the superpositioned two circles. (x_3, y_3) has the highest payoff for the self-player agent (the projection in $\mathbf{x}$), however, this is not a Nash equilibrium since the opponent's payoff is zero. The opponent tries to change to a position that its payoff is not zero. Should this happen, the self-player agent needs to move as well. (x_1, y_1) is the unique Nash equilibrium for all given $\cos\theta$, neither player can unilaterally improve its reward. To prove the relation more formally, for two pure strategy systems with probability amplitude of P and A are in a single-bit superposition is shown in Eq. (20):

$$|\psi\rangle = c_0 |P\rangle + c_1 |A\rangle \tag{20}$$

where the coefficients c_0, c_1 are complex numbers. Let

$$\begin{aligned} c_0 &= Pe^{i\theta} \\ c_1 &= Ae^{i\phi} \end{aligned} \tag{21}$$

2023 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

According to Eq. (8) and $\lambda = \frac{1}{2}$, the probability amplitude of the superposition as a total reward, i.e., total social welfare r, for the LLAQIG is

$$r = \frac{1}{2}\begin{bmatrix} Pe^{-i\theta} & Ae^{-i\phi}\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}Pe^{i\theta}\\ Ae^{i\phi}\end{bmatrix} \tag{22}$$

$$= \frac{1}{2}(P^2 + A^2 + 2PA\cos(\phi - \theta)) \tag{23}$$

According to Eq. (23), the social welfare r is maximized when $\phi = \theta$. The off-diagonal elements in the density matrix contributes the extra payoff value of the social welfare due to the quantum superposition.

The agent value is:

$$V = \frac{1}{2}(P + A)^2\cos^2\theta \tag{24}$$

The opponent value is:

$$O = \frac{1}{2}(P + A)^2\sin^2\theta \tag{25}$$

□

Theorem 1 is used to score a new content c from a content provider for its value, and assume $\mathbf{P}(t, j)$ and $\mathbf{A}(t, j)$ are the number of patterned and anomalous word features for Agent j and its peer network as content consumers, $P(t, j, c)$ and $A(t, j, c)$ are the number of patterns and anomalies for the content c. The value of c from LLAQIG is computed in Algorithm 2.

Algorithm 2 LLAQIG Algorithm

while $t \geq 0$ **do**
  $[\mathbf{P}(t, j), \mathbf{A}(t, j)] \Leftarrow LLA[B(t, j)]$
  for each new content c of Agent j **do**
  $V(t, j, c) \Leftarrow [\frac{P(t,j,c)}{L} + \frac{A(t,j,c)}{L}]^2 \frac{P(t,j,c)}{L}$
  end for
end while=0

$P(t, j, c)$ and $A(t, j, c)$ are the numbers of patterned and anomalous word features for the content c, respectively. Content c depends on j, omitted for brevity. L is a normalization factor, e.g., the length of a sentence or message of a content. $\frac{P(t,j,c)}{L}$ represents the degree of initial entanglement content c with other content in Agent j's peer network.

Intuitively, ranking and sorting the nodes using $V(t, j, c)$ would give a content, which represents a capability or service, more weight if it contains more anomalous features when matched with a search in a demand. Therefore, this mechanism balances the load of the naturally high-connected nodes with the nodes of fewer connections.

Theorem 2 (QAE Value)**.** *The quantum game-theoretic value of a CLA, defined in Eq. (3), converges to a higher value along with a higher social welfare of its peer network than the one defined in Eq. (23) via a QAE process.*

Proof. We set the Hamiltonian as follows:

$$H_B = \frac{1}{2}\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \tag{26}$$

$$H_C = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix} \tag{27}$$

$$H_C = (H_C)^\dagger H_C \tag{28}$$

$$H(t) = (1 - s(t))H_B + s(t)H_C \tag{29}$$

where $0 <= s(t) <= 1$ is the variable that changes with time. When $s(t) = 1$, $H(t) = H_C$. H_B evolves to H_C, since in the beginning, two systems are independent and in the end, they are entangled into one, which is a new system. In the new single-bit quantum game, the value of an agent is the Nash equilibrium $\langle\psi| H_C |\psi\rangle$ measured at $|0\rangle$.

$$|\psi(0)\rangle = \begin{bmatrix}P\\ A\end{bmatrix} \tag{30}$$

where P and A are real and computed from LLA.

According to Eq.(14) and $p = 1$:

$$U_B^1 = e^{-i\beta_1 H_B} \tag{31}$$

$$= I\cos\frac{\beta_1}{2} - i2H_B\sin\frac{\beta_1}{2} \tag{32}$$

$$= \begin{bmatrix}e^{-i\frac{\beta_1}{2}} & 0\\ 0 & e^{-i\frac{\beta_1}{2}}\end{bmatrix} \tag{33}$$

$$U_C^1 = e^{-i\gamma_1 H_C} \tag{34}$$

$$= I\cos\frac{\gamma_1}{2} - i2H_C\sin\frac{\gamma_1}{2} \tag{35}$$

$$= \begin{bmatrix}e^{-i\frac{\gamma_1}{2}} & -i\sin\frac{\gamma_1}{2}\\ -i\sin\frac{\gamma_1}{2} & e^{-i\frac{\gamma_1}{2}}\end{bmatrix} \tag{36}$$

$$U^1(T, 0) = U_B^1 U_C^1 \tag{37}$$

$$= \begin{bmatrix}e^{-i\frac{\beta_1}{2}} & 0\\ 0 & e^{-i\frac{\beta_1}{2}}\end{bmatrix}\begin{bmatrix}e^{-i\frac{\gamma_1}{2}} & -i\sin\frac{\gamma_1}{2}\\ -i\sin\frac{\gamma_1}{2} & e^{-i\frac{\gamma_1}{2}}\end{bmatrix} \tag{38}$$

According to Eq. (19) and $p = 1$:

$$\langle\psi(T, \gamma_1, \beta_1)| = U^1(T, 0)\,|\psi(0)\rangle \tag{39}$$

$$= \begin{bmatrix}e^{-i\frac{\beta_1}{2}} & 0\\ 0 & e^{-i\frac{\beta_1}{2}}\end{bmatrix}\begin{bmatrix}e^{-i\frac{\gamma_1}{2}} & -i\sin\frac{\gamma_1}{2}\\ -i\sin\frac{\gamma_1}{2} & e^{-i\frac{\gamma_1}{2}}\end{bmatrix}\begin{bmatrix}P\\ A\end{bmatrix} \tag{40}$$

$$= \begin{bmatrix}Pe^{-i(\frac{\beta_1}{2}+\frac{\gamma_1}{2})} - iA\sin\frac{\gamma_1}{2}e^{-i\frac{\beta_1}{2}}\\ -iP\sin\frac{\gamma_1}{2}e^{-i\frac{\beta_1}{2}} + Ae^{-i(\frac{\beta_1}{2}+\frac{\gamma_1}{2})}\end{bmatrix} \tag{41}$$

$$= \begin{bmatrix}P'\\ A'\end{bmatrix} \tag{42}$$

$$H_C\,|\psi(T)\rangle = \frac{1}{2}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}P'\\ A'\end{bmatrix} \tag{43}$$

$$= \frac{1}{2}\begin{bmatrix}P' + A'\\ P' + A'\end{bmatrix} \tag{44}$$

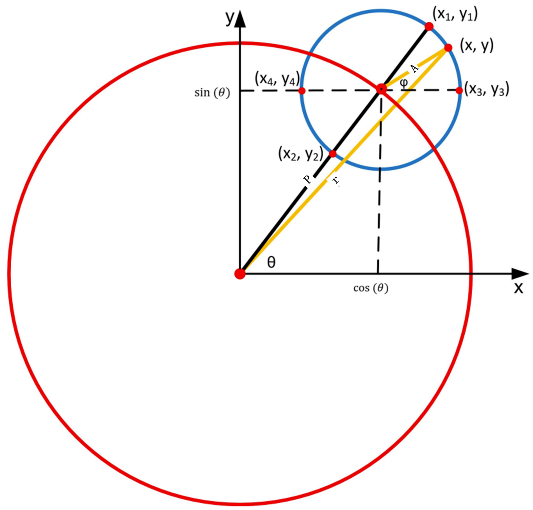

Fig. 2. Superpositioned patterned (P) and anomalous (A) themes. The highest-value of a self-player is achieved both at the Nash equilibrium and maximal total social welfare. Each CLA is a self-player, its peer network is the opponent.

$$\langle\psi(T)| H_C |\psi(T)\rangle = \langle\psi(T)| (H_C)^\dagger H_C |\psi(T)\rangle \tag{45}$$
$$= \frac{1}{2}\begin{bmatrix}P'^\dagger & A'^\dagger\end{bmatrix}\begin{bmatrix}1 & 1\\ 1 & 1\end{bmatrix}\begin{bmatrix}P'\\ A'\end{bmatrix} \tag{46}$$
$$= \frac{1}{2}(P'^\dagger(P'+A') + A'^\dagger(P'+A')) \tag{47}$$
$$= \frac{1}{2}((P'+A')^\dagger(P'+A')) \tag{48}$$
$$= \frac{1}{2}((P+A)^2(1+3\sin^2\frac{\gamma_1}{2})) \tag{49}$$

From Eq. (49), we have the following inequality:

$$\frac{1}{2}(P+A)^2 \le \langle\psi(T)| H_C |\psi(T)\rangle \le 2(P+A)^2 \tag{50}$$

Eq. (50) shows a better social welfare is obtained from the QAE process than the one from Theorem 1.

Agent value is Eq. (44) measured at $|0\rangle$:

$$V = \frac{1}{4}(P'+A')^\dagger(P'+A') \tag{51}$$
$$= \frac{1}{4}((P+A)^2(1+3\sin^2\frac{\gamma_1}{2})) \tag{52}$$

The cooperating opponent value is Eq. (44) measured at $|1\rangle$:

$$O = \frac{1}{4}(P'+A')^\dagger(P'+A') \tag{53}$$
$$= \frac{1}{4}((P+A)^2(1+3\sin^2\frac{\gamma_1}{2})) \tag{54}$$

□

5 Discussions

Both measurements or values of the self-player agent and opponent as in Eq. (52) and Eq. (54) reach the maxima when the whole system's social welfare Eq. (50) is maximized. However, both maxima do not depend on β, only depend on γ, where $\sin^2\frac{\gamma_1}{2}$ represents a small time step in the evolution and the maxima can not be reached in one step $p = 1$. When $p \to \infty$ or $\gamma_1 \to 0$, both values of agents and social welfare reach the maxima during a finite time T simultaneously. The maximum of the social welfare in Eq. (50) or QAE value even when $p = 1$ is better than the one shown in Eq. (23) that is derived from a simple superposition.

The QAE value is computed from the unsupervised and self-organizing process in the sense it does not need labeled data from supervised ML. The QAE value of an agent evolves in a natural quantum mechanism. Fig. 3 shows neural networks of layered and interleaved quantum operators that lead to the true value of a CLA without the need of feedback from the outside environment. In real-life, we might not be able to directly observe a QAE process, however, one can use the theoretic optimized values or their approximations for finite p to compare content and information derived from a practical algorithm such as LLA and LLAQIG.

In literature, several related research have been proposed [22], [23], [24]. Based on these frameworks, in a quantum Prisoner's dilemma game where each player has a qubit and can manipulate it independently. Each player can decide to play classical pure strategy C (cooperate), D (defect), or a quantum strategy Q, i.e., an unitary positive definite complex matrix. It has been shown that there exists such quantum operation with the correct amount initial degree of entanglement of players, the Q-strategy is both at Nash equilibrium and Pareto optimal.

These quantum games can be played in classic devices, where a system characterized by a graph in which the nodes are players' quantum actions and edges are quantum interactions. Cooperation of nodes can emerge from such evolutionary games to escape the Prisoner's dilemma. Li et al. [26] shows that Q can become the dominant strategy which favors emergence of cooperation for both random networks and scale-free networks. Li & Yong [24] show that initial quantum entanglement of players is the key mechanism of guaranteed emergence of cooperation. Our results are different from these approaches in the sense that two different systems or two

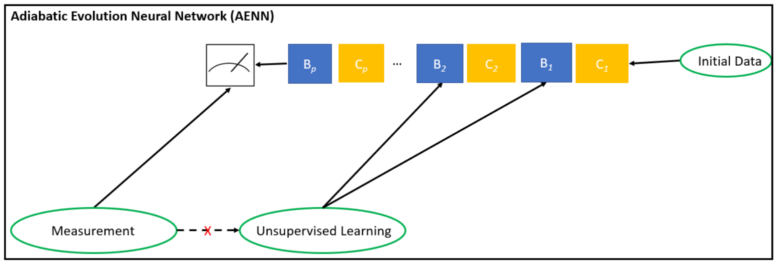

Fig. 3. QAE Neural Networks (QAENN): QAENN is a layered and alternative quantum operations which are proved to converge and evolve to both individual agent's Nash equilibrium, meanwhile the peer network is Pareto efficient, i.e., reaches the maximum social welfare. QAENN computes the QAE value of a CLA from the unsupervised and self-organizing process in the sense it does not need labeled data from supervised ML. The QAE value of an agent evolves in a natural quantum mechanism without the outside feedback.

players (i.e., P and A) can cooperate to a degree that they are totally merged into one, i.e., H_B evolves to H_C. In other words, the entanglement of the two players increases slowly in an adiabatic process.

Our results do not need quantum computers. For example, there is an efficient classical calculation that determines γ and β [11] in Eq. (19) and related to quantum approximate optimization algorithms (QAOA) for a wide range of NP-complete combinatorial optimization problems. The adiabatic quantum optimization for QAOA is an alternative to the approaches such as Max-Cut [17] or correlation clustering [18] algorithms, which usually does not require quantum computers to perform the approximation, since γ and β have a physical meaning of representing small steps to naturally change a Hamiltonian in an adiabatic process.

Some quantum adiabatic applications do use quantum computers, which is different from our approach. For example, when exploiting the computational advantage of quantum devices, such D-wave and adiabatic process, Li et al. [25] shows a better optimization for selecting actions for a deep reinforcement learning system. Our application only needs to show there are theoretic and totally entangled equilibrium optimized values for both individual agents and the peer network. We do not need to compute the actual quantum optimization parameters such as RL actions. Our results are unsupervised and self-organizing.

6 Conclusions

We show and prove two quantum theoretic value scoring methods based on both superposition and quantum adiabatic evolution to compute Nash equilibrium values of a peer system of collaborative learning agents. QAE method is unsupervised and self-organizing in the sense that one can set a small step to let the system to evolve to a higher value in a natural mechanism.

Acknowledgment

Authors would like to thank the Office of Naval Research (ONR) to support the research. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied of the U.S. Government.

References

[1] Lotidis, K., Mertikopoulos, P., and Bambos, N. (2023). Learning in Quantum Games.
[2] Chakrabarti, S., Huang, Y., Li, T., Feizi, S., & Wu, X. ((2019) Quantum Wasserstein generative adversarial networks. In NeurIPS '19: Proceedings of the 33rd International Conference on Neural Information Processing Systems, 2019.
[3] Chuang, I. & Nielsen, M. (2010). Quantum Computation and Quantum Information. Cambridge University Press, 2 edition, 2010.
[4] Debreu, G. (1952). A social equilibrium existence theorem. Proceedings of the National Academy of Sciences of the USA, 38(10):886–893.
[5] Papadimitriou, C. H. & Roughgarden, T. (2008). Computing correlated equilibria in multi-player games. J. ACM. 55 (3): 14:1–14:29. CiteSeerX 10.1.1.335.2634. doi:10.1145/1379759.1379762. S2CID 53224027.
[6] Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. Quantum machine learning. Nature, 549(7671):195–202, September 2017.
[7] Schuld, M., Sweke, R., & Johannes Meyer, J. (2021). Effect of data encoding on the expressive power of variational quantum-machine-learning models. Physical Review A, 103(3):032430, 2021.
[8] Goto, T., Tran, Q., & Nakajima, T. Universal approximation property of quantum machine learning models in quantum-enhanced feature spaces. Physical Review Letters, 127(9):090506, 2021.
[9] Pérez-Salinas, A., López-Núñez, D., García-Sáez, A., Forn-Díaz, P., & Latorre, J. (2021). One qubit as a universal approximant. Physical Review A, 104(1):012405, 2021.
[10] Yu, Z., Yao, H., Li, M., & Wang, X. (2022). Power and limitations of single-qubit native quantum neural networks. 36th Conference on Neural Information Processing Systems (NeurIPS 2022).
[11] Farhi, E., Goldstone, J., & Gutmann,S. (2014). A Quantum Approximate Optimization Algorithm, arXiv Prepr. arXiv1411.4028, pages 1-16, 2014.
[12] Penrose, R. (1990). Empire's New Mind.
[13] Jing, H.,Wang, Y. & Li, Y. (2023). Max-Cut quantum approximate optimization algorithm for power systems. Communications Engineering volume 2, Article number: 12 (2023).
[14] Farhi, E., Goldstone, J.,Gutmann, S., & Sipser, M. (2000). Quantum Computation by Adiabatic Evolution. MIT CTP # 2936 quant-ph/000110
[15] Farhi, E., Goldstone, J. & Gutmann, S. (2014). A quantum approximate optimization algorithm. arXiv https://doi.org/10.48550/ARXIV.1411.4028.
[16] IBM (2023). Intro to QAOA and MaxCut: Practical Quantum Computing: Asynchronous Workshop.
[17] Farhi, E., Gamarnik,D. & Gutmann, S. (2020). The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: A Typical Case. arXiv e-prints, April 2020. DOI: 10.48550/arXiv.2004.09002.
[18] Weggemans, J. R. et al. (2022). Solving correlation clustering with QAOA and a Rydberg qudit system: a full-stack approach. Quantum 6, 687. https://arxiv.org/pdf/2106.11672.pdf

[19] Zhou, C., Zhao, Y., Kotak, C. (2009). The Collaborative Learning agent (CLA) in Trident Warrior 08 exercise. In Proceedings of the International Conference on Knowledge Discovery and Information Retrieval - Volume 1: KDIR, IC3K (pp. 323-328) DOI: 10.5220/0002332903230328. Madeira, Portugal.
[20] Blei, D., Ng, A. and Jordan, M.: Latent Dirichlet allocation. Journal of Machine Learning Research, 3:993-1022 (2003). Retrieved from http://jmlr.csail.mit.edu/papers/volume3/blei03a/blei03a.pdf
[21] Penn State University (PSU), (2021). Online Statistics: Normal Approximation Method Formulas. https://online.stat.psu.edu/stat200/lesson/9/9.1/9.1.2/9.1.2.1
[22] Eisert,J. and Wilkins, M. and Lewenstein, M.(1999).Phys. Rev. Lett. 83. p.3077.
[23] Sun, Z.W. (2009). The rule for evolution of cooperation in quantum games. ACTA PHYSICA POLONICA A. 116(2).
[24] Li, A., Yong, X. (2014). Entanglement Guarantees Emergence of Cooperation in Quantum Prisoner's Dilemma Games on Networks. Sci Rep. 4, 6286.
[25] Lin, J., Lai, Z. and Li, X. (2020). Quantum adiabatic algorithm design using reinforcement learning. Physical Review A 101.5 (2020): 052327.
[26] Li, Q., Chen, M., Perc, M., Iqbal, A. Abbott, D. (2013). Effects of adaptive degrees of trust on coevolution of quantum strategies on scale-free networks. Sci. Rep. 3, 2949; DOI:10.1038/srep02949.
[27] http://www.math.umd.edu/˜mboyle/courses/475sp05/spec.pdf
[28] Newman, M. E. J.: Finding community structure in networks using the eigenvectors of matrices. Phys. Rev. E 74, 036104 (2006).