Annals of Operations Research
https://doi.org/10.1007/s10479-023-05700-z

ORIGINAL RESEARCH

Calculating Nash equilibrium on quantum annealers

Faisal Shah Khan[1] · **Olga Okrut**[1] · **Keith Cannon**[1] · **Kareem H. El-Safty**[1] · **Nada Elsokkary**[1,2]

Received: 21 December 2022 / Accepted: 3 November 2023

Abstract
Adiabatic quantum computing is implemented on specialized hardware using the heuristics of the quantum annealing algorithm. To solve a problem using quantum annealing, the problem requires formatting as a discrete quadratic function without constraints. The problem of finding Nash equilibrium in two-player, non-cooperative games is a two-fold quadratic optimization problem with constraints. This problem was formatted as a single, constrained quadratic optimization in 1964 by Mangasarian and Stone. Here, we show that adding penalty terms to the quadratic function formulation of Nash equilibrium gives a quadratic unconstrained binary optimization (QUBO) formulation of this problem that can be executed on quantum annealers. Three examples are discussed to highlight the success of the formulation, and an overall, time-to-solution (hardware + software processing) speed up by seven to ten times is reported on quantum annealers developed by D-Wave System.

Keywords Nash equilibrium · Quantum annealing · Quantum computing · Quadratic unconstrained binary optimization · Adiabatic quantum computing

1 Introduction

Arguably, one of the most practically successful near-term quantum computing solutions is quantum annealing, such as that developed by the Canadian company D-Wave Systems and utilized for many years now by several industry leaders. While there is abundant literature demonstrating the utility of D-Wave's quantum annealer for solving practical problems, there are also concerns about its quasi-quantum architecture and whether it will eventually capture quantum computing's full potential in terms of the breadth of problems it can solve, given the persistence of noise and decoherence problems affecting it (Denchev et al., 2016). Nevertheless, quantum annealing has shown both promising performance and progress over the years (Tabi et al., 2020, 2021; Venturelli, 2015; Fernández-Campoamor et al., 2021; Ayanzadeh et al., 2020; King et al., 2022).

✉ Faisal Shah Khan
faisal@darkstarquantumlab.com

1 Dark Star Quantum Lab, Apex, NC, USA

2 Center on Cyber-Physical Systems, Khalifa University, Abu Dhabi, UAE

Here, we will use the rich literature on optimization techniques and make necessary modifications so that they are in the proper format for a quantum annealer to calculate Nash equilibrium, a solution concept from the theory of non-cooperative games. Nash equilibrium can be viewed as a special, stronger case of multi-objective optimization where one wants to simultaneously maximize (or minimize) a collection of objective functions under several conflicting criteria. In symbols,

$$\max\left[f_1(x),\ f_2(x),\ \ldots,\ f_n(x)\right]. \tag{1}$$

For example, in finance, investment portfolios require optimization of appropriate functions under the conflicting criteria of minimizing risk and maximizing expected return. Maximizing each coordinate function in (1),

$$\left[\max f_1(x), \max f_2(x), \ldots, \max f_n(x)\right], \tag{2}$$

and adding further constraints by considering the objective functions as payoff functions to notional players of an underlying game produces Nash equilibrium. As a special case of multi-objective optimization, the calculation of Nash equilibrium on quantum annealers opens up the potential of utilizing quantum computing technology for solving a large class of optimization problems.

We believe that the tremendous impact of formal game theory on politics and warfare (De Mesquita, 2010), socioeconomics (Arrow & Debreu, 1954), and scientific development (evolutionary biology (Maynard Smith, 1982)), are sufficiently motivating. Not to mention game theory as a fundamental historical force that has carved events in social history, as evident from ancient classics like *Art of War* by Sun Tzu, *Arthashastra* by Chanakya, and more recently, *The Prince* by Niccolò Machiavelli. Finally, it is shown in Gottlob et al. (2005) that the problem of finding all the Nash equilibria in a strategic game is NP-complete, making this problem worthy of execution on a quantum annealer.

However, due to the current limitations of D-wave's quantum annealing platform, it is not yet feasible to find "mixed strategy" Nash equilibria that manifest as (real-valued) probability distributions. Therefore, our discussion will be limited to finding the deterministic "pure strategy" Nash equilibria. And while there exists a polynomial time algorithm for finding these equilibria, the use of a quantum annealer is motivated by the fact that there exist games in which the number of pure strategies can be as high as $2^{349,525}$ (Binmore, 2007). In such big data situations, finding pure strategy Nash equilibria can become infeasible on conventional computers despite the polynomial time algorithm for finding them. A similar situation occurs in the case of linear regression for which a quantum annealing implementation has also been developed recently (Date & Potok, 2021).

2 Quantum annealing

The central physical idea behind quantum annealing is that of the adiabatic theorem, whereby a system beginning in the lowest energy solution of an initial Hamiltonian will remain in the lowest energy state of the final Hamiltonian after some sufficiently small perturbation - thereby avoiding the need to calculate the solution for the more complex Hamiltonian. It is this final Hamiltonian that is sometimes called the "problem Hamiltonian." In our case, it is the physical twin of the game theoretic optimization problem. For further reading on quantum annealing, we refer the reader to the work of Yarkoni et al. (2022) which provides a clear survey. We will draw on only the mathematical parts here.

The model system for a quantum annealer is a graph $G(V, E)$ of V qubits at locations i and j connected by an edge E_{ij} of strength J_{ij} to which we will apply a transverse field. Mathematically, as time t progresses, we perturb the system $\mathcal{H}_i$ to $\mathcal{H}_f$

$$\mathcal{H}(t) = A(t)\mathcal{H}_i + B(t)\mathcal{H}_f \tag{3}$$

where $A(t)$ and $B(t)$ are interpolating parameters and

$$\mathcal{H}_i = \sum_{i \in V} \sigma_i^x, \tag{4}$$

$$\mathcal{H}_f = \sum_{i \in V} h_i \sigma_i^z + \sum_{i,j \in E} J_{ij} \sigma_i^z \sigma_j^z \tag{5}$$

with σ^x and σ^z being Pauli spin matrices. Upon measurement, the qubits in superposed spin states take definite values $\{-1, 1\}$ and we are left with a simplified equation, the Ising model:

$$E_{Ising} = \sum_{i \in V} h_i s_i + \sum_{i,j \in E} J_{ij} s_i s_j \tag{6}$$

As is also shown in Yarkoni et al. (2022), with a change of variable $s \mapsto 2x - 1$, we can transform the above into the quadratic unconstrained binary optimization problem (QUBO) below and vice versa, since that is the format of the problem derived from applications (in our case, game theory) whose minima we seek:

$$\text{minimize} : \sum_i a_i x_i + \sum_{j>i} b_{i,j} x_i x_j \tag{7}$$

with $a_i, b_{i,j} \in \mathbb{R}$, $x_i \in \{0, 1\}$, and no constraints. However, realistic quadratic binary optimization problems arise under constraints. One removes the constraints by adding penalty terms to the quadratic formulation. For example, consider the quadratic formulation of the binary Markowitz portfolio optimization problem Elsokkary et al. (2017) of investing in an asset i in the portfolio, or not:

$$\begin{aligned} \text{minimize} : & \sum_i [-E(R_i)x_i] + \sum_{j>i} \text{Cov}(R_i, R_j) x_i x_j \\ & + \left(\sum_i A_i x_i - B \right)^2 \end{aligned} \tag{8}$$

where R_i denotes the random variable representing the return from asset i, $E(R_i)$ represents its expected value, and $\text{Cov}(R_i, R_j)$ the co-variance of R_i and R_j. Finally, the constraints A_i and B representing the maximum amount of money that can be invested in asset i and the total budget, respectively, are subsumed into the quadratic expression as $\left(\sum_i A_i x_i - B\right)^2$.

This paper is organized as follows: Sect. 3 gives a concise introduction to the idea of Nash equilibrium and its presentation as a doubly quadratic optimization problem in two player games. Section 4 formulates the Nash equilibrium problem as a single quadratic optimization problem. In Sect. 4.1, we transform this quadratic expression into a QUBO expression for execution on a quantum annealer by adding penalty terms to it. Finally, in Sect. 5, we give examples of our work, using as examples the two player, two strategy toy game *Battle of the Sexes*, a two player, three strategy “bird game” from evolutionary biology, and finally an eight strategy example based on finite automata interactions. Results of executing our problems on D-Wave’s quantum annealer, accessed through Amazon Web Service, are also presented.

3 Nash equilibrium

Nash equilibrium (Nash, 1950; Binmore, 2007) is the solution concept for non-cooperative games. It is a play of the game, that is, a collection of strategies, one per player, such that no player is motivated to unilaterally deviate from his particular strategic choice in the collection. Strategies can take on any form in general, but when the players randomize between their strategies, the probability distributions over the strategies used to randomize are referred to as *mixed strategies*, with the original set of strategies now referred to as *pure strategies*.

In the case of two-player games, Nash equilibrium is a pair of mixed strategies (p^*, q^*) such that the expected payoff functions Π_1 and Π_2 of player 1 and player 2 respectively, satisfy

$$\Pi_1(p^*, q^*) \geq \Pi_1(p, q^*) \quad \text{and} \quad \Pi_2(p^*, q^*) \geq \Pi_2(p^*, q), \tag{9}$$

for all mixed strategies p and q. Let M be the $n \times m$ matrix whose entries are the payoff to player 1 and N be the $n \times m$ matrix whose entries are payoffs to player 2 when pure strategies are in play, with the (i, j) entry of either matrix equaling the payoff to the player when player 1 employs his ith pure strategy and player 2 employs her jth pure strategy. The payoffs to the players can be calculated as

$$\Pi_1(p, q) = p^T Mq \quad \text{and} \quad \Pi_2(p, q) = p^T Nq \tag{10}$$

with $p = (p_1, \ldots p_n)$ and $q = (q_1, \ldots, q_m)$ being the probability distributions so that $\sum p_i = \sum q_j = 1$.

We note that finding Nash equilibrium in two player games involves calculating two quadratic expressions in (10) which are then optimized by comparing them as in (9). However, to execute this problem on a quantum annealer, it has to be formatted as a QUBO which has the general form given in (7). To express the Nash equilibrium problem as a QUBO, it needs to be addressed as a single quadratic optimization problem. We note that the authors of Roch et al. (2020) address the question of finding Nash equilibrium in "graphical" games.

To avoid possible confusion, we clarify here that there is a difference between calculating a Nash equilibrium on a quantum computer and implementing or playing a game on a quantum computer. The consequences of the latter are studied in the theory of *quantum games*, where the main interests are either to find better quality Nash equilibria in games that are played using a quantum computer, or to study quantum physical features like phase as non-cooperative games or game in a Bayesian setting. The reader can explore this subject further in the following sample references (Meyer, 1999; Eisert et al., 1999; Khan et al., 2018; Phoenix et al., 2020; Teklu et al., 2009; Brivio et al., 2010; Teklu et al., 2010).

4 Nash equilibrium as quadratic optimization

Using the transformations in Mangasarian et al. (1964), the statement of Nash equilibrium in (10) can be restated as a quadratic optimization problem:

$$\text{maximize}: p^T(M + N)q - \alpha - \beta \tag{11}$$

where, by taking e and l to respectively be the $n \times 1$ and $m \times 1$ vectors of ones, we get the constraints

$$Mq - \alpha e \leq 0 \tag{12}$$

$$N^T p - \beta l \leq 0 \tag{13}$$

$$e^T p - 1 = 0 \tag{14}$$

$$l^T q - 1 = 0 \tag{15}$$

$$p, q \geq 0 \tag{16}$$

with α and β scalars whose maximum values, α^* and β^*, equal to the expected payoffs to players I and II respectively, and for which

$$p^{*T}(M + N)q^* - \alpha^* - \beta^* = 0. \tag{17}$$

4.1 Formatting for quantum annealers - QUBO

Quantum annealers require that problems to be executed on them be formatted as QUBO problems. The QUBO format requires that the variables be binary-valued, that is, $p_i, q_i \in \{0, 1\}$ for all i and j. This restricts Nash equilibrium solutions to be calculated in terms of pure strategies only, with $p_i = q_j = 1$ for appropriate i, j. Being an NP-complete problem, it is worthy of execution on a quantum annealer. On the other hand, it is problematic for two reasons: first, because the larger class of proper mixed strategy Nash equilibria, for which the values of p_i and q_i are rational or real numbers, is missed; second, because the *guarantee* of the existence of Nash equilibria in games is only available with respect to mixed strategies.

To identify mixed strategy Nash equilibria, real-valued variables require encoding as binary variables in quantum annealers. This is an active topic of study in the field (Karimi et al., 2019; Rogers et al., 2020), but no standards exist yet, and no implementations in actual quantum annealing hardware are available to date. Therefore, we do not consider mixed strategy Nash equilibria here and remain focused on the problem of finding, or not, pure strategy Nash equilibria.

QUBO formulation also assumes that the matrix representation of the problem is such that the matrix is symmetric or upper-triangular. For convenience, we will assume that the matrix $M + N$ is square of size $n \times n$, meaning that both players have n pure strategies and that it is symmetric. To remove the constraints in Eqs. (12–16) and add the corresponding penalties into the objective function in (11), note that $Mq - \alpha e$ is the sequence of n inequalities

$$\sum_j m_{i,j} q_j - \alpha \leq 0, \tag{18}$$

with $m_{i,j}$ the ith row and j^{th} column element of the matrix M, and q_j the j^{th} element of the vector q. Similarly, $N^T p - \beta l$ is the sequence of n inequalities

$$\sum_j n_{j,i} p_i - \beta \leq 0, \tag{19}$$

with $n_{j,i}$ the jth row and ith column element of the matrix N^T, and p_i the ith element of the vector p.

Inequality constraints of the form of $Ax - b \leq 0$ can be transformed to an equality constraint $Ax - b + \zeta = 0$ with an added non-negative slack variable $\zeta \geq 0$. Hence, one can obtain the following sequence of equality constraints corresponding to the inequalities (18) and (19):

$$\sum_i m_{i,j} q_j - \alpha + \zeta = 0 \tag{20}$$

$$\sum_j n_{j,i} p_i - \beta + \eta = 0. \tag{21}$$

Motivated by the heuristics described in Glover et al. (https://arxiv.org/ftp/arxiv/papers/1811/1811.11538.pdf), Kia (http://solmaz.eng.uci.edu/Teaching/MAE206/Lecture14.pdf), Asghari et al. (2022), we add these $2n$ equations into the objective function, call it F, as penalties, together with penalties for Eqs. (14) and (15), to get:

$$\begin{aligned} F = {} & p^T(M+N)q - \alpha - \beta + \theta_1 \left(\sum_i p_i - 1 \right)^2 \\ & + \theta_2 \left(\sum_j q_j - 1 \right)^2 + \lambda_i \left(\sum_j m_{i,j} q_j - \alpha + \zeta \right)^2 \\ & + \phi_j \left(\sum_j n_{j,i} p_i - \beta + \eta \right)^2 \end{aligned} \tag{22}$$

where the penalty term multipliers θ_1, θ_2, and λ_i and ϕ_i, for $1 \leq i, j \leq n$, are appropriate real-valued weights. Finally, since quantum annealers only minimize objective functions, we negate (22) and

$$\text{minimize} : -F. \tag{23}$$

5 Examples

We give three examples, one where both players have two pure strategies, one where they have three strategies each, and one where they both have eight pure strategies. More details on examples implementation and explanations can be found in the GitHub repository.

5.1 Battle of the sexes

This example was considered in Mangasarian et al. (1964) at a time of rapid development of classical computing hardware in 1964:

$$\text{maximize} : p^T(M+N)q - \alpha - \beta \tag{24}$$

where

$$M = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}, \quad N = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \tag{25}$$

are the payoff matrices of the player I and player II, respectively, arising from the strategic form of the game Battle of the Sexes appearing in Fig. 1 in which player I is the row player and player II is the column player and both players choose between making the strategic choice of watching an opera or a football game. The payoffs to the players are represented by ordered pairs of numbers in which the first number is the payoff to the row player and the second number is the payoff to the column player.

Each player can make the strategic choice of watching Opera versus watching Football. As a trivial case of mixed strategies, these pure strategies are respectively represented by the elements of the set $\left\{(1,0)^T, (0,1)^T\right\}$ for each player. Direct analysis shows that the two pure

Fig. 1 Strategic form of the two-player game Battle of the Sexes. The two pure strategy Nash equilibria are highlighted

	Opera	Football
Opera	(2,1)	(-1,-1)
Football	(-1,-1)	(1,2)

strategy Nash equilibria in this game are the highlighted plays (Opera, Opera) and (Football, Football) of Fig. 1.

Expanding the expression in (24) using matrices M and N from (25) gives:

$$\text{maximize} : 3p_1q_1 - 2p_1q_2 - 2p_2q_1 + 3p_2q_2 - \alpha - \beta \tag{26}$$

subject to

$$2q_1 - q_2 - \alpha \leq 0 \tag{27}$$

$$-q_1 + q_2 - \alpha \leq 0 \tag{28}$$

$$p_1 - p_2 - \beta \leq 0 \tag{29}$$

$$-p_1 + 2p_2 - \beta \leq 0 \tag{30}$$

$$p_1 + p_2 - 1 = 0 \tag{31}$$

$$q_1 + q_2 - 1 = 0 \tag{32}$$

Removing the constraints by adding penalties to F from (22), setting all penalty term multipliers equal to 1, and then negating give:

$$\begin{aligned} -F = &-3p_1q_1 + 2p_1q_2 + 2p_2q_1 - 3p_2q_2 + \alpha + \beta \\ &- (p_1 + p_2 - 1)^2 - (q_1 + q_2 - 1)^2 - (2q_1 - q_2 - \alpha + \zeta)^2 \\ &- (-q_1 + q_2 - \alpha + \zeta)^2 - (p_1 - p_2 - \beta + \eta)^2 \\ &- (-p_1 + 2p_2 - \beta + \eta)^2 , \end{aligned} \tag{33}$$

which we need to minimize using quantum annealing. The constraints outlined in (16) are satisfied since p and q are binary vectors.

Note that while QUBO formulation involves only one vector variable, $x = (x_1, \cdots, x_n)^T$, the formulation of the Nash equilibrium as a QUBO problem in (22) and (33) involves two, $p = (p_1, \cdots, p_n)^T$ and $q = (q_1, \cdots, q_n)^T$. This issue is resolved by D-Wave's software development kit, which handles the transformation of multiple vector variables into an appropriate single, higher-dimensional vector variable. For the current example, this means that the quadratic expression

$$-3p_1q_1 + 2p_1q_2 + 2p_2q_1 - 3q_2p_2$$

appearing in (33) becomes

$$-3x_0x_1 + 2x_0x_2 + 0 \cdot x_0x_3 + 0 \cdot x_1x_2 + 2x_1x_3 - 3x_2x_3,$$

with $p_1 = x_0, q_1 = x_1, q_2 = x_2, p_2 = x_3$, which is consistent with the condition $j > i$ in the quadratic part of the QUBO formulation (7). Details about this transformation and the procedure of introducing slack variables are available in D-Wave (https://docs.dwavesys.com/docs/latest/doc_getting_started.html) and Condello et al. (https://github.com/dwavesystems/dimod). Additionally, the slack variables are represented as binary variables using binary expansion of $p, q \in \{0, 1\}$.

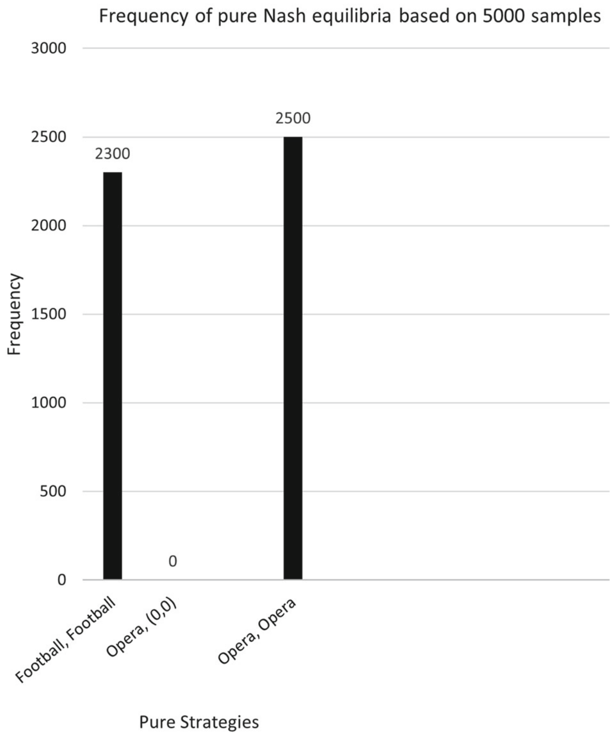


Fig. 2 Approximate frequency of pure Nash equilibrium points based on 5000 samples for two-strategy game executed on D-Wave Advantage 4.1 QPU. D-Wave 2000 Q6 QPU was also utilized and produced similar results, though these are not shown here for brevity. The Nash equilibrium points of high frequency identified by both QPUs are consistent with those identified with direct analysis. The QPUs also show (Opera, $(0,0)^T$) as a potential Nash equilibrium, albeit with frequency 0. Because $(0,0)^T$ is not a valid pure strategy, this result is considered to be noise. Raw frequency graphs can be referred to at the following github repository: https://github.com/DarkStarQuantumLab/Nash-Equilibrium-quantum-annealing/blob/main/Results/Frequency_2strategiesAdvantage.png

Running the experiment on D-Wave quantum annealer simulator as a binary quadratic modle (BQM), we obtain the two pure strategy Nash equilibria at $x = (1, 0)^T = y = (1, 0)^T$, which corresponds to both players watching Opera, and at $x = (0, 1)^T$, $y = (0, 1)^T$, which corresponds to both players watching Football. This is consistent with a direct analysis of the strategic form of the game in Fig. 1. We next submitted the experiment to D-Wave 2000 Q6 and D-Wave Advantage System 4.1 quantum processing unit (QPU) available via Amazon Braket. Minimizing the objective function in (33) over 5000 samples using the sampler *DWaveSampler*, together with the minor embedding *EmbeddingComposite* available in the D-Wave software development kit, we obtain results consisted with those produced by simulated annealing. The results from the QPUs are presented in Fig. 2.

The Nash equilibrium points were determined from the results of the QPU based on the frequencies of occurrence. That is, the points with the high frequency of occurrences correspond to the minimum of the quadratic function. Therefore, following the mathematical formulation of the problem, the points with the high frequency of occurrence correspond to the global minima (maxima) and therefore to the Nash equilibria. The necessity to sample many times (around 1000–5000) is dictated by the probabilistic nature of the quantum annealers

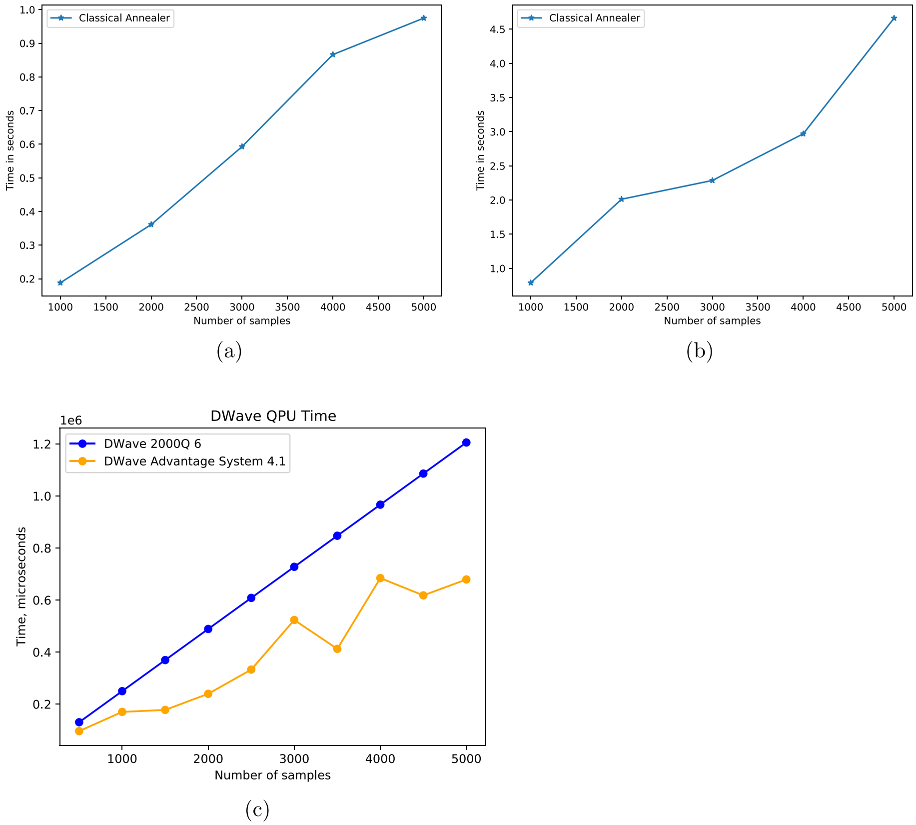


Fig. 3 **a** CPU access time, in seconds, for solving two-strategy (Battle of the Sexes) game using classical annealing simulator. **b** CPU access time for solving three-strategy (Bird game) game using classical annealing simulator. **c** QPU access time, in seconds (labeled on the vertical axis as microseconds $\times 1e^6$), for *both* the two and three-strategy games on D-Wave 2000 Q6 and D-Wave Advantage System 4.1 with the default annealing time of $20\mu s$

and quantum computers in general. The frequencies with low number of counts (less than 50) are considered as noise in the quantum hardware and are not treated as solutions.

The solutions produced by the D-Wave QPU are influenced by the change in penalty multipliers θ_1 and θ_2 in the quadratic formulation of the problem in Eq. 22. We note that the variation of the penalty multipliers allows us to fine-tune the lowest energy state and avoid getting stuck in local minima. On the other hand, the variation of the penalty multipliers can potentially lead to unstable points. In this work, the values of θ_1 and θ_2 were selected automatically by D-Wave QPU's software development kit.

We investigated the performance of classical annealing simulation executed on a CPU versus quantum annealing on D-Wave's QPU for the two and three-strategy games in terms of the length of time respective processing units were engaged. The results are summarised in Fig. 3 where Fig. 3a and b show the CPU access time for 5000 samples was close to 1 and 11 s, respectively, for an Intel Core i-5, 1.6 GHz CPU running Windows 10 operating system. On the other hand, Fig. 3c shows that D-Wave 2000Q 6 QPU access time for both the two and three-strategy problem was 1.2 seconds for 5000 samples while D-Wave Advantage 4.1 access time was around 0.5 seconds. The QPU access time includes the time spent on sampling

Fig. 4 Strategic form of the two player, three strategy game Bird Game. The three pure strategy Nash equilibria are highlighted

	Hawk	Dove	Bourgeois
Hawk	(-5,-5)	(10,0)	(2.5,-2.5)
Dove	(0,10)	(2,2)	(1,6)
Bourgeois	(-2.5,2.5)	(6,1)	(5,5)

and the time spent to program D-Wave hardware as described in D-Wave documentations (D-Wave, https://docs.dwavesys.com/docs/latest/c_qpu_timing.html#id21). An investigation of the performance and quality of solutions depending on the type of minor embedding to D-Wave hardware topologies is a topic worthy of further research.

5.2 A bird game: an example with three strategies

A three strategy example of a two-player game comes from evolutionary biology. We have sourced this example from table 5 in Davis and Brams (https://www.britannica.com/science/game-theory/The-prisoners-dilemma) and reproduced it in strategic form in Fig. 4. The pay-off matrices for this game are

$$M = \begin{pmatrix} -5 & 10 & 2.5 \\ 0 & 2 & 1 \\ -2.5 & 6 & 5 \end{pmatrix}, N = \begin{pmatrix} -5 & 0 & -2.5 \\ 10 & 2 & 6 \\ 2.5 & 1 & 5 \end{pmatrix} \tag{34}$$

and the pure strategic choices, as a trivial case of mixed strategies, are represented by the elements of the set

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Direct analysis shows that the highlighted entries (Hawk, Dove), (Dove, Hawk), and (Bourgeois, Bourgeois) of Fig. 4 are the pure strategy Nash equilibria of this game.

Following the procedure of formulating the objective function F and constraints gives:

$$\begin{aligned} \text{maximize} : -10p_1q_1 + 10p_1q_2 + 0p_1q_3 + 10p_2q_1 + 4p_2q_2 + \\ 7p_2q_3 + 0p_3q_1 + 7p_3q_2 + 10p_3q_3 - \alpha - \beta \end{aligned} \tag{35}$$

subject to

$$-5q_1 + 10q_2 + 2.5q_3 - \alpha \leq 0 \tag{36}$$
$$2q_2 + q_3 - \alpha \leq 0 \tag{37}$$
$$-2.5q_1 + 6q_2 + 5q_3 - \alpha \leq 0 \tag{38}$$
$$-5p_1 - 2.5p_3 - \beta \leq 0 \tag{39}$$
$$10p_1 + 2p_2 + 6p_3 - \beta \leq 0 \tag{40}$$
$$2.5p_1 + p_2 + 5p_3 - \beta \leq 0 \tag{41}$$
$$p_1 + p_2 + p_3 - 1 = 0 \tag{42}$$
$$q_1 + q_2 + q_3 - 1 = 0. \tag{43}$$

Since D-Wave's BQM model does not allow the submission of fractional coefficients as constraints, the inequalities (36), (38), and (41) are multiplied by 2. Additionally, inequality

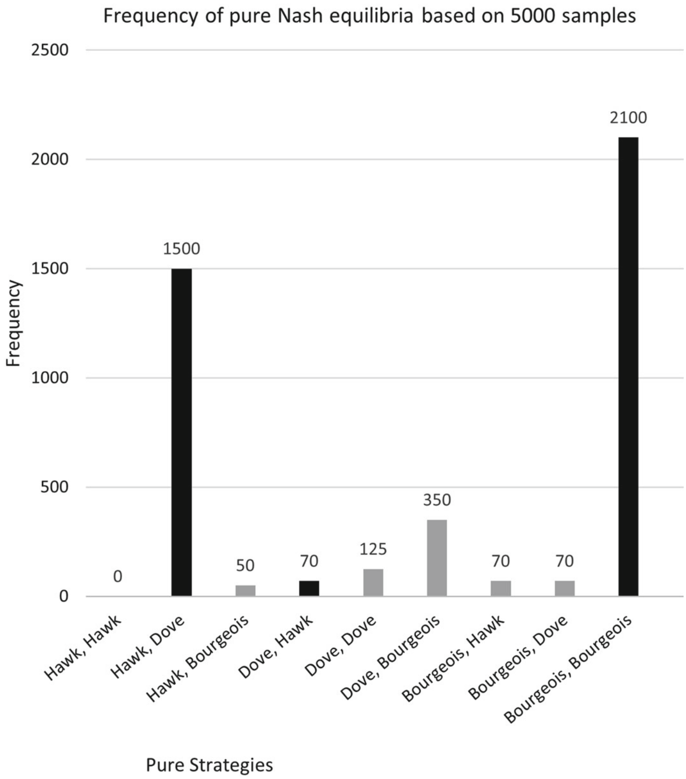


Fig. 5 Approximate frequency of pure Nash equilibrium points (black bars) based on 5000 samples for three-strategy game executed on D-Wave Advantage 4.1 QPU. D-Wave 2000 Q6 QPU was also utilized and produced similar results, though these are not shown here for brevity. Two Nash equilibria of high frequency identified by both QPUs are consistent with those identified with direct analysis, namely, (Hawk, Dove) and (Bourgeois, Bourgeois). Note that the third Nash equilibrium point (Dove, Hawk) appears with a very low frequency, barely meeting the cut off frequency of 50 between solution and noise. See text for a discussion of this feature. Raw frequency graphs can be referred to at the following github repository: https://github.com/DarkStarQuantumLab/Nash-Equilibrium-quantum-annealing/blob/main/Results/Frequency_3strategiesAdvantage(1)(1).png

(40) can be reduced by 2. Removing the constraints by adding penalties to F from (22), setting all penalty term multipliers equal to 1, and then negating give:

$$\begin{aligned}
-F =&10p_1q_1 - 10p_1q_2 - 0p_1q_3 - 10p_2q_1 - 4p_2q_2 \\
&- 7p_2q_3 - 0p_3q_1 - 7p_3q_2 - 10p_3q_3 + \alpha + \beta \\
&- (p_1 + p_2 + p_3 - 1)^2 - (q_1 + q_2 + q_3 - 1)^2 \\
&- (2q_2 + q_3 - \alpha + \zeta)^2 \\
&- (-2q_1 + 4q_2 + q_3 - \alpha + \zeta)^2 \\
&- (-5q_1 + 12q_2 + 10q_3 - \beta + \zeta)^2 \\
&- (-10p_1 - 5p_3 - \beta + \eta)^2 \\
&- (5p_1 + p_2 + 3p_3 - \beta + \eta)^2 - (5p_1 + 2p_2 + 10p_3 - \beta + \eta)^2 ,
\end{aligned} \tag{44}$$

which we need to minimize.

	Dove	Hawk	Grim	Tit-for-Tat	Tat-for-Tit	Tweedledum	Tweedledee	Tweetypie
Dove	(2,2)	(-1,3)	(2,2)	(2,2)	(-1,3)	(2,2)	(2,2)	(-1,3)
Hawk	(3,-1)	(0,0)	(0,0)	(0,0)	$\left(\frac{3}{2},-\frac{1}{2}\right)$	$\left(\frac{3}{2},-\frac{1}{2}\right)$	$\left(\frac{3}{2},-\frac{1}{2}\right)$	(3, -1)
Grim	(2,2)	(0,0)	(2,2)	(2,2)	(0,0)	(2,2)	(2,2)	(3,-1)
Tit-for-Tat	(2,2)	(0,0)	(2,2)	(2,2)	$\left(\frac{2}{3},\frac{2}{3}\right)$	(2,2)	(2,2)	(2,2)
Tat-for-Tit	(3,-1)	$\left(-\frac{1}{2},\frac{3}{2}\right)$	(0,0)	$\left(\frac{2}{3},\frac{2}{3}\right)$	(2,2)	(2,2)	(2,2)	(2,2)
Tweedledum	(2,2)	$\left(-\frac{1}{2},\frac{3}{2}\right)$	(2,2)	(2,2)	(2,2)	(2,2)	(2,2)	(2,2)
Tweedledee	(2,2)	$\left(-\frac{1}{2},\frac{3}{2}\right)$	(2,2)	(2,2)	(2,2)	(2,2)	(2,2)	(2,2)
Tweetypie	(3,-1)	(-1, 3)	(-1, 3)	(2,2)	(2,2)	(2,2)	(2,2)	(2,2)

Fig. 6 A two (row versus column) player game in which each player has eight pure strategies. The payoffs to the players are given by ordered pair of numbers where the first number is the payoff to the row player and the second number is the payoff to the column player. Shaded plays are the twenty two Nash equilibria of this game

By executing the experiment on D-Wave Quantum Annealer Simulator, we obtained two pure Nash equilibria at $p = (0, 0, 1)^T$, $q = (0, 0, 1)^T$, which corresponds to both players using the strategy Bourgeois, and at $p = (1, 0, 0)^T$, $q = (0, 1, 0)^T$, which corresponds to player I using the strategy Hawk and player II using the strategy Dove. Executing the problem on the D-Wave QPUs with over 5000 samples produced the solutions appearing in Fig. 5 where, in addition to the Nash equilibrium solutions, we also observed the solution points $p = (0, 1, 0)^T$, $q = (0, 0, 1)^T$ ((Dove, Bourgeois)) and $p = (0, 1, 0)^T$, $q = (0, 1, 0)^T$ ((Dove, Dove)). These plays of the game are not Nash equilibria. Rather, these points appear as local solutions to the quadratic minimization problem that the Nash equilibrium problem was formatted as. Further, keeping in mind that a QUBO formulation requires that symmetric indices be considered identical (so that $x_{ij} = x_{ji}$), the results are consistent with a direct calculation which shows the symmetric Nash equilibrium $p = (0, 1, 0)^T$, $q = (1, 0, 0)^T$ corresponding to player I using the strategy Dove and player II using the strategy Hawk. Finally, note that this equilibrium appears separately in Fig. 5 with a very small frequency. We attribute this to the noisy nature of first-generation QPUs.

We investigated the QPU access times and qualities of the pure strategy Nash equilibrium points for both D-Wave QPU topologies for the different number of samples. We did not observe any order of magnitude increase in the QPU access time for solving for Nash equilibria in this game compared to the case of solving for Nash equilibria in *Battle of the Sexes* example (see Fig. 3). On the other hand, it took on average around 3.2 seconds to solve the same problem on a classical machine with an Intel Core i-5, 1.6 GHz CPU running Windows 10 operating system.

5.3 An example with eight strategies

Consider an example of a game from Binmore (2007) in which each player has eight pure strategies, generated from a dynamic, finite automaton version of the game Prisoner's Dilemma in which the notion of long term stability of a Nash equilibrium is of interest. This notion is referred to as *evolutionary stable strategy* (ESS) because as the game evolves over time, an ESS Nash equilibrium will be robust against invasion by mutant strategies. This game is described by the payoff table in Fig. 6. The payoff matrices for the row and column players can be extracted from the payoff table to be, respectively,

$$M = \begin{pmatrix} 2 & -1 & 2 & 2 & -1 & 2 & 2 & -1 \\ 3 & 0 & 0 & 0 & 3/2 & 3/2 & 3/2 & 3 \\ 2 & 0 & 2 & 2 & 0 & 2 & 2 & 3 \\ 2 & 0 & 2 & 2 & 2/3 & 2 & 2 & 2 \\ 3 & -1/2 & 0 & 2/3 & 2 & 2 & 2 & 2 \\ 2 & -1/2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & -1/2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 3 & -1 & -1 & 2 & 2 & 2 & 2 & 2 \end{pmatrix}, \tag{45}$$

$$N = \begin{pmatrix} 2 & 3 & 2 & 2 & 3 & 2 & 2 & 3 \\ -1 & 0 & 0 & 0 & -1/2 & -1/2 & -1/2 & -1 \\ 2 & 0 & 2 & 2 & 0 & 2 & 2 & -1 \\ 2 & 0 & 2 & 2 & 2/3 & 2 & 2 & 2 \\ -1 & 3/2 & 0 & 2/3 & 2 & 2 & 2 & 2 \\ 2 & 3/2 & 2 & 2 & 2 & 2 & 2 & 2 \\ 2 & 3/2 & 2 & 2 & 2 & 2 & 2 & 2 \\ -1 & 3 & 3 & 2 & 2 & 2 & 2 & 2 \end{pmatrix} \tag{46}$$

with pure strategy choices $p = (p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8)^T$ and $q = (q_1, q_2, q_3, q_4, q_5, q_6, q_7, q_8)^T$ such that only one of the p_i and q_i are equal to 1. Following procedures similar to those in formulating the objective function in the first two examples gives:

$$\begin{aligned}
-F = &- 12p_1q_1 - 6p_1q_2 - 12p_1q_3 - 12p_1q_4 - 6p_1q_5 \\
&- 12p_1q_6 - 12p_1q_7 - 6p_1q_8 - 3p_2q_5 - 3p_2q_6 \\
&- 3p_2q_7 - 3p_2q_8 - 12p_3q_3 - 12p_3q_4 - 12p_3q_6 \\
&- 12p_3q_7 - 6p_3q_8 - 12p_4q_4 - 4p_4q_5 \\
&- 12p_4q_7 - 12p_4q_8 - 12p_5q_5 - 12p_5q_6 \\
&- 12p_5q_7 - 12p_5q_8 - 12p_6q_6 - 12p_6q_7 - 12p_6q_8 \\
&- 12p_7q_7 - 12p_7q_8 - 12p_8q_8 + \alpha + \beta \\
&- (10(p_1 + p_2 + p_3 + p_4 + p_5 + p_6 + p_7 + p_8 - 1))^2 \\
&- (10(q_1 + q_2 + q_3 + q_4 + q_5 + q_6 + q_7 + q_8 - 1))^2 \\
&- (2q_1 + 3q_2 + 2q_3 + 2q_4 + 3q_5 + 2q_6 + 2q_7 + 3q_8 - \alpha + \zeta)^2 \\
&- (-2q_1 - q_5 - q_6 - q_7 - 2q_8 - \alpha + \zeta)^2 \\
&- (2q_1 + 2q_3 + 2q_4 + 2q_6 + 2q_7 - q_8 - \beta + \zeta)^2 \\
&- (3q_1 + 3q_3 + 3q_4 + q_5 + 3q_6 + 3q_7 + 3q_8 - \beta + \zeta)^2 \\
&- (-6q_1 + 9q_3 + 4q_4 + 12q_5 + 12q_6 + 12q_7 + 12q_8 - \beta + \zeta)^2 \\
&- (4q_1 + 3q_2 + 4q_3 + 4q_4 + 4q_5 + 4q_6 + 4q_7 + 4q_8 - \beta + \zeta)^2 \\
&- (4q_1 + 3q_2 + 4q_3 + 4q_4 + 4q_5 + 4q_6 + 4q_7 + 4q_8 - \beta + \zeta)^2 \\
&- (-1q_1 + 3q_2 + 3q_3 + 2q_4 + 2q_5 + 2q_6 + 2q_7 + 2q_8 - \beta + \zeta)^2 \\
&- (2p_1 - p_2 + 2p_3 + 2p_4 - p_5 + 2p_6 + 2p_7 - p_8 - \beta + \eta)^2 \\
&- (2p_1 + p_5 + p_6 + 1p_7 + 2p_8 - \beta + \eta)^2 \\
&- (2p_1 + 2p_3 + 2p_4 + 2p_6 + 2p_7 + 3p_8 - \beta + \eta)^2 \\
&- (2p_1 + 2p_3 + 2p_4 + p_5 + 2p_6 + 2p_7 + 2p_8 - \beta + \eta)^2 \\
&- (18p_1 - 3p_2 + 2p_4 + 12p_5 + 12p_6 + 12p_7 + 12p_8 - \beta + \eta)^2
\end{aligned}$$

$$
\begin{aligned}
&- (4p_1 - p_2 + 4p_3 + 4p_4 + 4p_5 + 4p_6 + 4p_7 + 4p_8 - \beta + \eta)^2 \\
&- (4p_1 - p_2 + 4p_3 + 4p_4 + 4p_5 + 4p_6 + 4p_7 + 4p_8 - \beta + \eta)^2 \\
&- (3p_1 - p_2 - p_3 + 2p_4 + 2p_5 + 2p_6 + 2p_7 + 2p_8 - \beta + \eta)^2
\end{aligned} \tag{47}
$$

where payoffs with fractional coefficients were multiplied by appropriate factors to get integer coefficients.

In Fig. 7, we give truncated QPU results containing only those plays of the game that have a reasonably high frequency (50 or higher) of occurrence. A direct analysis of the table in Fig. 6 shows that this game has twenty two Nash equilibria in the form of the highlighted entries. The QPU calculation misses many of these, identifying only eight Nash equilibria and proposing many non-equilibrium outcomes as solutions. For instance, the first six solutions in Fig. 7 are not Nash equilibria. Also, the QPU proposes twenty three potential solutions. The likely reason for this noisy result is that more qubits were used for processing, and because these are relatively less interconnected than a smaller number of qubits, the solutions returned were local minima of the Hamiltonian. This hypothesis is consistent with the bench marking results in Willsch et al. (2022). The current limitations to solve larger problems are related to the qubits connectivity in quantum hardware. In case of the quantum annealers, the connectivity between different qubits may be broken during the annealing, making it challenging to obtain exact solutions with high fidelity for larger problems.

Possible approaches to improve solution fidelity include more post-processing and custom embedding. On the other hand, the solutions where both players use their second, third, and fourth pure strategies (Hawk, Grim, and Tit-for-Tat) are in fact Nash equilibria and appear as solutions with high frequency in Fig. 7, a feature also consistent with the results from simulated annealing.

In Fig. 8, we show the performance of classical annealing simulation executed on a CPU versus quantum annealing on D-Wave's QPU for the eight-strategy game in terms of the length of time respective processing units were engaged. As Fig. 8a shows, the CPU access time for 5000 samples was close to 11 s for an Intel Core i-5, 1.6 GHz CPU running Windows 10 operating system. On the other hand, Fig. 8b shows that D-Wave Advantage QPU access time for 5000 samples was close to 0.679,271 s. For all three games considered, we observed a time delay of around $12 - 13$ seconds in receiving results from D-Wave QPU which is caused by Internet latency, Solver API time, and QPU queue wait time. Although it is a significant time delay, the expected quantum speed-up in solving larger problems is expected to outweigh third-party requests time. For all three experiments, D-Wave Advantage system was able to produce outputs twice as fast as D-Wave 2000 Q6. This is consistent with observations made by the authors of Willsch et al. (2022).

We again note that the variation of the penalty multiplier terms like θ_1, θ_2 allows us to fine-tune the lowest energy state and avoid getting stuck in local minima. This is useful for D-Wave's QPU since for large problems the gap between the lowest energy state and the first excited state becomes very small making it possible to get stuck in a local minimum. This impact of the variations in penalty multiplier terms is a topic for future research. In particular, what is the relationship between the change in these parameters and the frequency of equilibrium points with less total payoffs?

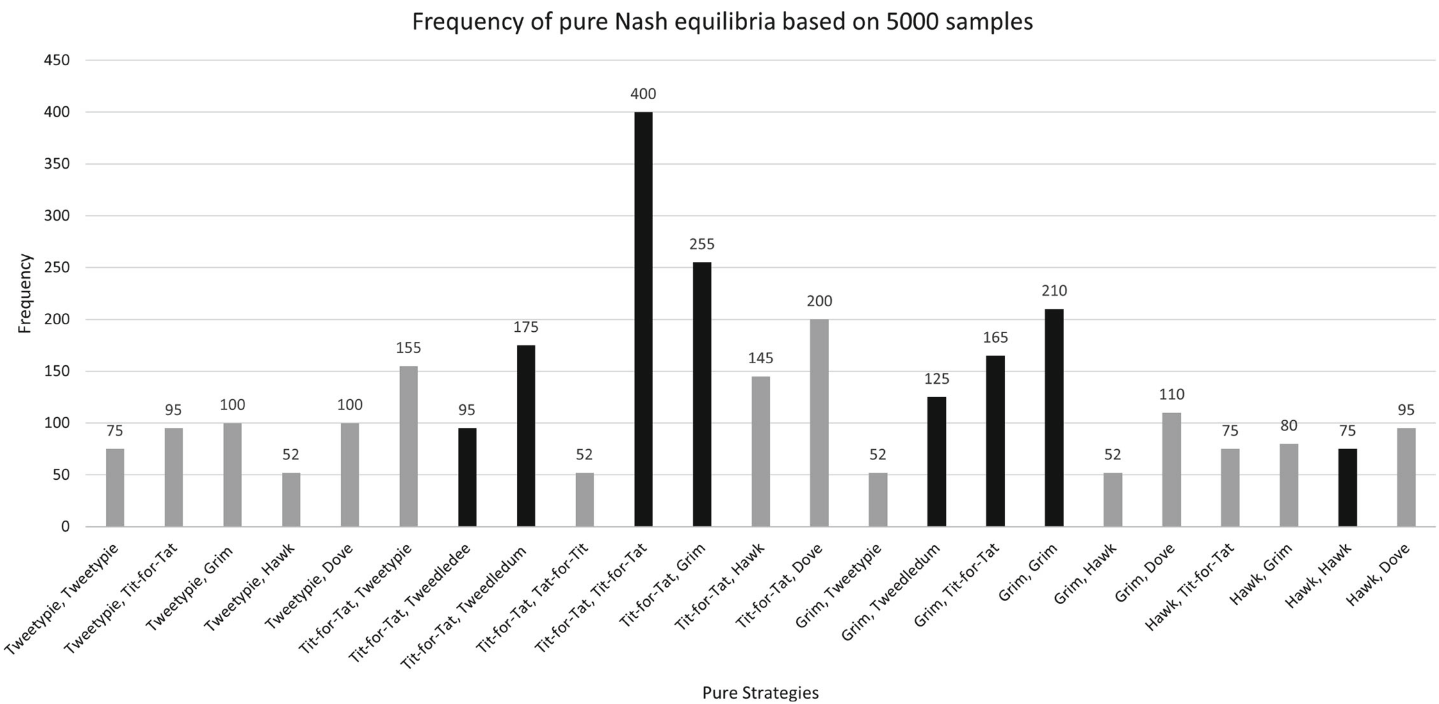
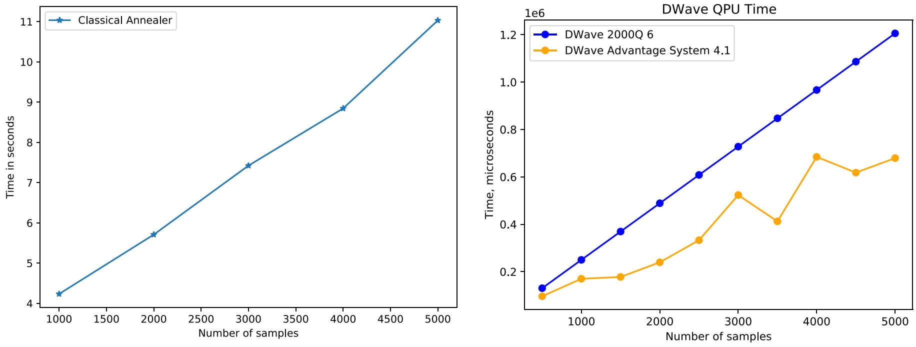


Fig. 7 A truncated, approximate frequency chart shows 23 out of the 64 possible plays of the game. Only those plays of the game that appear with a frequency of 50 or higher are considered for inclusion. Based on 5000 samples, the D-Wave Advantage 4.1 QPU identifies only eight of the twenty two Nash equilibria in this game, represented here by black bars. The raw graph can be referred to at the github repository: https://github.com/DarkStarQuantumLab/Nash-Equilibrium-quantum-annealing/blob/main/Results/Frequency_8strategies_Advantage.pdf



(a) (b)

Fig. 8 **a** CPU access times, in seconds, for solving the eight-strategy problem using classical annealing simulator. **b** QPU access time, in seconds (labeled on the vertical axis as microseconds $\times 1e^6$), for the eight-strategy game for D-Wave 2000 Q6 and D-Wave Advantage System 4.1 with an annealing time of $100\mu s$

6 Conclusion

We format the problem of calculating Nash equilibrium in two player competitive games for execution on a quantum annealer. To do this, we use the result of Mangasarian et al. (1964). This first allows us to express the two traditionally used quadratic optimization problems (9) for describing Nash equilibrium as a single quadratic optimization problem. Next, by adding penalty terms, we remove constraints from this quadratic optimization problem to formulate Nash equilibrium as a QUBO problem, and execute three examples of this formulation on a classical computer (laptop), D-Wave's Quantum Annealer Simulator, and their 2000Q and Advantage QPUs, observing a time-to-solution (hardware + software processing) speed up by seven to ten times in comparison to the classical machine. Python code for this work is found at the following link: https://github.com/DarkStarQuantumLab/Nash-Equilibrium-quantum-annealing.

The values of the penalty multipliers like θ_1 and θ_2 in the unconstrained quadratic formulation define the degree to which a penalty should be applied for violating the corresponding constraint in the original quadratic formulation. Thus, the solutions to the unconstrained quadratic formulation will depend on the choice of multiplier values in relation to the number of pure strategies, as well as the payoffs to the players. This is observed in the experiments we conducted in this work where D-wave's software development kit selected the values of these parameter. This suggests that these parameters require tuning for each individual QUBO problem formulated from a constrained quadratic problem.

Non-cooperative game theory has proven its mettle as an accurate model of real world, two-player competitive scenarios. One example is political landscapes where two powers are dominant (the continuing cold war between the West and the Soviet Union/Russia is an example), and another is economic and financial decision making with respect to multinational, multi-trillion dollar development projects (the ongoing Chinese led Belt and Road Initiative). Each of these large-scale applications of competitive game theory is high stake game in which rapid calculation of accurate best-response strategic choices is paramount for success.

Given that the future developments in science, technology, and socioeconomics will be more complex, our future work will explore ways to solve for Nash equilibrium on quantum annealers in more realistic mixed strategies. Finally, it is our goal to develop quantum computational solutions to game-theoretic models of fundamental problems of modern society. For example, the Nash bargaining solution which accurately models the problem of value determination of commodities.

Acknowledgements FSK thanks Travis Humble for useful discussions.

Declarations

Conflict of interest This research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. FSK has courtesy affiliation with Center on Cyber-Physical Systems, Khalifa University, Abu Dhabi.

References

Arrow, K., & Debreu, G. (1954). Existence of an equilibrium for a competitive economy. *Econometrica, 22*(3), 265–290.

Asghari, M., et al. (2022). Transformation and linearization techniques in optimization: A state-of-the-art survey. *Mathematics, 10*, 283.

Ayanzadeh, R., Halem, M., & Finin, T. (2020). Reinforcement quantum annealing: A hybrid quantum learning automata. *Scientific Reports, 10*, 7952.

Binmore, K. (2007). *Player for real* (Illustrated). Oxford: Oxford University Press.

Brivio, D., et al. (2010). Experimental estimation of one-parameter qubit gates in the presence of phase diffusion. *Phys. Rev. A, 81*, 012305.

Condello, A., et.al. Dimod: A shared API for QUBO/Ising samplers. GitHub repository https://github.com/dwavesystems/dimod.

Date, P., & Potok, T. (2021). Adiabatic quantum linear regression. *Scientific Reports, 11*, 21905.

Davis, M. D., & Brams, S. J. The prisoner's dilemma, Britannica, https://www.britannica.com/science/game-theory/The-prisoners-dilemma

De Mesquita, B. (2010). The Predictioneer's game: Using the logic of brazen self-interest to see and shape the future, Random House Trade Paperbacks; Illustrated edition, October 12.

Denchev, V., Boixo, S., Isakov, S., Ding, N., Babbush, R., Smelyanskiy, V., Martinis, J., & Neven, H. (2016). What is the computational value of finite-range tunneling? *Physical Review X, 6*, 031015.

D-Wave. The Quantum Computing Company, D-Wave QPU Architecture: Topologies https://docs.dwavesys.com/docs/latest/c_gs_4.html?highlight=advantage%20system#pegasus-graph

D-Wave. The quantum computing company, getting started with D-wave solvers, user manual, 2021-10-20. https://docs.dwavesys.com/docs/latest/doc_getting_started.html

D-Wave. The Quantum Computing Company, Operation and Timing https://docs.dwavesys.com/docs/latest/c_qpu_timing.html#id21

Eisert, J., et al. (1999). Quantum games and quantum strategies. *Physical Review Letters, 83*, 3077.

Elsokkary, N., et al. (2017). Financial portfolio management using adiabatic quantum optimization: The case of Abu Dhabi securities exchange. In *Proceedings of IEEE high performance extreme computing conference*.

Fernández-Campoamor, M., O'Meara, C., Cortiana, G., Peric, V. & Bernabé-Moreno, J. (2021). Community detection in electrical grids using quantum annealing. ArXiv Preprint arXiv:2112.08300

Glover, F., et al. Quantum Bridge Analytics I: A tutorial on formulating and using QUBO models, pre-print https://arxiv.org/ftp/arxiv/papers/1811/1811.11538.pdf.

Gottlob, G., et al. (2005). Pure Nash equilibria: Hard and easy games. *Journal of Artificial Intelligence Research, 24*, 357–406.

Karimi, S., et al. (2019). Practical integer-to-binary mapping for quantum annealers. *Quantum Information Processing, 18*(4), 94.

Khan, F. S., et al. (2018). Quantum games: A review of the history, current state, and interpretation. *Quantum Information Processing, 17*, 309.

Kia, S. Optimization Method. http://solmaz.eng.uci.edu/Teaching/MAE206/Lecture14.pdf

King, A. D., Suzuki, S., Raymond, J., et al. (2022). Coherent quantum annealing in a programmable 2,000-qubit Ising chain. *Nature Physics, 18*, 1324–1328.

Mangasarian, O. L., et al. (1964). Two person non-zero sum games and quadratic programming. *Journal of Mathematical Analysis and Applications, 9*, 348–355.
Maynard Smith, J. (1982). *Evolution and the theory of games*. Cambridge: Cambridge University Press.
Meyer, D. A. (1999). Quantum strategies. *Physical Review Letters, 82*, 1052.
Nash, J. (1950). Equilibrium points in n-person games. *Proceedings of the National Academy of Sciences, 36*, 1.
Phoenix, S. J. D., et al. (2020). Preferences in quantum games. *Physics Letters A, 384*(15), 29.
Roch, C., et al. (2020). A quantum annealing algorithm for finding pure Nash equilibria in graphical games. *Lecture Notes in Computer Science*, Vol. 12142.
Rogers, M. L., et al. (2020). Floating-point calculations on a quantum annealer: Division and matrix inversion. *Frontiers in Physics, 8*, 265.
Tabi, Z., El-Safty, K., Kallus, Z., Hága, P., Kozsik, T., Glos, A., & Zimborás, Z. (2020). Quantum optimization for the graph coloring problem with space-efficient embedding. In *2020 IEEE international conference on quantum computing and engineering (QCE)* (pp. 56–62).
Tabi, Z., Marosits, Á., Kallus, Z., Vaderna, P., Gódor, I., & Zimborás, Z. (2021). Evaluation of quantum annealer performance via the massive MIMO problem. *IEEE Access, 9*, 131658–131671.
Teklu, B., et al. (2009). Bayesian estimation of one-parameter qubit gates. *Journal of Physics B: Atomic, Molecular and Optical Physics, 42*, 035502.
Teklu, B., et al. (2010). Phase estimation in the presence of phase diffusion: The qubit case. *Physica Scripta, T140*, 014062.
Venturelli, D. (2015). Quantum optimization of fully connected spin glasses. *Physical Review, 5*(3), 031040.
Willsch, D., Willsch, M., Gonzalez Calaza, C. D., et al. (2022). Benchmarking advantage and D-Wave 2000Q quantum annealers with exact cover problems. *Quantum Inf Process, 21*, 141.
Yarkoni, S., et al. (2022). Quantum annealing for industry applications: Introduction and review. *Reports on Progress in Physics, 85*, 104001.