Socratic Game Theory: Playing Games in Many Possible Worlds

Seth Gilbert* April Rasala Lehman† Matt Lepinski* David Liben-Nowell‡

Abstract. In traditional game theory, players are typically endowed with exogenously given knowledge of the structure of the game—either full omniscient knowledge or partial but fixed information. In real life, however, people are often unaware of the utility of taking a particular action until they perform research into its consequences. In this paper, we introduce *Socratic game theory* to model this phenomenon. (We imagine a player engaged in a question-and-answer session, asking questions both about his or her own preferences and about the state of reality; thus we call this setting "Socratic" game theory.) In a Socratic game, players begin with an *a priori* probability distribution over many *possible worlds*, with a different utility function for each world. Players can make *queries*, at some cost, to learn information about which of the possible worlds is the actual world, before choosing an action. We consider two query models: (1) an observable query model in which each player knows which query the other players made and (2) an unobservable query model in which each player learns only the response to his or her own query.

The results in this paper consider cases in which the underlying worlds of a two-player Socratic game are either *constant-sum games* or *strategically zero-sum games*, a class that generalizes constant-sum games to include all games in which the sum of payoffs depends linearly on the interaction between the players. When the underlying worlds are constant sum, we give a polynomial-time algorithm to find Nash equilibria in both the observable- and unobservable-query models. When the worlds are strategically zero sum, we give efficient algorithms to find Nash equilibria in unobservable-query Socratic games and correlated equilibria in observable-query Socratic games.

1 Introduction

Late October 1960. A smoky room. Democratic Party strategists huddle around a map. How should the Kennedy campaign allocate its remaining advertising budget? Should it focus on, say, California or New York? The Nixon campaign faces the same dilemma. Of course, neither campaign knows the effectiveness of its advertising in each state. Perhaps Californians are susceptible to Nixon's advertising, but are unresponsive to Kennedy's. In light of this uncertainty, the Kennedy campaign may conduct a survey, at some cost, to estimate the effectiveness of its advertising. Moreover, the larger—and more expensive—the survey, the more accurate it will be. Is the cost of a survey worth the information that it provides? How should one balance the cost of acquiring more information against the risk of playing a game with higher uncertainty?

We introduce *Socratic game theory* as a general formal model for situations of this type. As in traditional game theory, the players in a Socratic game choose actions to maximize their payoffs, but we model players with incomplete information who can make costly queries to reduce their uncertainty about the state of the world before they choose their actions. This approach contrasts with traditional game theory, in which players are usually modeled as having fixed, exogenously given information about the structure of the game and its payoffs. (In traditional games of incomplete and imperfect information, there is information that the players do not have; in Socratic games, unlike in these games, the players have a chance to acquire the missing information, at some cost.)

A Socratic game proceeds as follows. A *real world* is chosen randomly from a set of *possible worlds* according to a common prior distribution. Each player then selects an arbitrary *query* from a set of available costly queries and receives a corresponding piece of information about the real world. Finally each player selects an action and receives a payoff—a function of the players' selected actions and the identity of the real world—less the cost of the query that he or she made.

The novelty of our model is the introduction of explicit costs to the players for learning arbitrary partial information about which of the many possible worlds is the real world. Like all games, Socratic games can be viewed as a special case of *extensive-form games*, which represent games by trees in which internal nodes represent choices made by chance or by the players, and the leaves are labeled with a vector of payoffs to the players. Algorithmically, the generality of extensive-form games makes them difficult to solve efficiently, and the special cases that are known to be efficiently solvable do not include even simple Socratic games.

*Computer Science and Artificial Intelligence Lab, Massachusetts Institute of Technology. †Google, Inc. ‡Department of Mathematics and Computer Science, Carleton College. {sethg,lepinski,arasala}@theory.lcs.mit.edu, dlibenno@carleton.edu.

Thanks to Erik Demaine, Natalia Hernandez Gardiol, Claire Monteleoni, Jason Rennie, Madhu Sudan, and Katherine White.

Our research was initially inspired by recent results in psychology on decision making, but it soon became clear that Socratic game theory is also a general tool for understanding the "exploitation versus exploration" tradeoff, well-studied in machine learning, in a strategic multiplayer environment. This tension between the risk arising from uncertainty and the cost of acquiring information is ubiquitous in political science, economics, and beyond.

Our results. We consider Socratic games under two models: an *unobservable* model where players learn only the response to their own queries and an *observable* model where players also learn which queries their opponents made. We give efficient algorithms to find Nash equilibria—i.e., tuples of strategies from which no player has unilateral incentive to deviate—in broad classes of two-player Socratic games in both models. Our first result is an efficient algorithm to find Nash equilibria in unobservable-query Socratic games with *constant-sum* worlds, in which the sum of the players' payoffs is independent of their actions. Our techniques also yield Nash equilibria in unobservable-query Socratic games with *strategically zero-sum* worlds. Strategically zero-sum games generalize constant-sum games by allowing the sum of the players' payoffs to depend on individual players' choices of strategy, but not on any interaction of their choices. Our second result is an efficient algorithm to find Nash equilibria in observable-query Socratic games with constant-sum worlds. Finally, we give an efficient algorithm to find *correlated equilibria*—a weaker but increasingly well-studied solution concept for games [2, 3, 21, 41, 42]—in observable-query Socratic games with strategically zero-sum worlds.

Every (complete-information) classical game is a trivial Socratic game (with no uncertainty and a single trivial query), and efficiently finding Nash equilibria in classical games is believed to be hard [8, 39, 40]. Therefore we would not expect to find a straightforward polynomial-time algorithm to compute Nash equilibria in general Socratic games. However, it is well known that Nash equilibria can be found efficiently via an LP for two-player constant-sum games [34, 52] (and strategically zero-sum games [36]). A Socratic game is itself a classical game, so one might hope that these results can be applied to Socratic games with contant-sum (or strategically zero-sum) worlds. We face two major obstacles in extending these classical results to Socratic games. First, a Socratic game with constant-sum worlds is not itself a constant-sum classical game—rather, the resulting classical game is only strategically zero sum. Further, a Socratic game with strategically zero-sum worlds is not itself classical strategically zero sum—indeed, there are no known efficient algorithmic techniques to compute Nash equilibria in the resulting class of classical games. (Exponential-time algorithms like Lemke/Howson, of course, can be used [31].) Thus even when it is easy to find Nash equilibria in each of the worlds of a Socratic game, we require new techniques to solve the Socratic game itself. Second, even when the Socratic game itself is strategically zero sum, the number of possible strategies available to each player is exponential in the natural representation of the game. As a result, the standard linear programs for computing equilibria have an exponential number of variables *and* an exponential number of constraints.

For unobservable-query Socratic games with strategically zero-sum worlds, we address these obstacles by formulating a new LP that uses only polynomially many variables (though still an exponential number of constraints) and then use ellipsoid-based techniques to solve it. For observable-query Socratic games, we handle the exponentiality by decomposing the game into stages, solving the stages separately, and showing how to efficiently reassemble the solutions. To solve the stages, it is necessary to find Nash equilibria in Bayesian strategically zero-sum games, and we give an explicit polynomial-time algorithm to do so.

This paper contains three main contributions: (1) the definition of Socratic game theory, a new and interesting game-theoretic model; (2) efficient algorithms to find Nash equilibria in two-player unobservable-query Socratic games with strategically zero-sum worlds and observable-query Socratic games with constant-sum worlds; and (3) an efficient algorithm to find correlated equilibria in two-player observable-query Socratic games with strategically zero-sum worlds.

2 Socratic Game Theory

We review background on game theory and formally introduce Socratic games. Boldface variables will be used to denote a pair of variables ($\mathbf{a} = \langle a_{\mathrm{I}}, a_{\mathrm{II}} \rangle$). Let $\mathtt{Pr}[x \leftarrow \pi]$ denote the probability that a particular value x is drawn from the distribution π, and let $\mathtt{E}_{x\sim\pi}[g(x)]$ denote the expectation of $g(x)$ when x is drawn from π.

2.1 Background on Game Theory

Consider two players, Player I and Player II, each of whom is attempting to maximize his or her utility (or payoff). A *(two-player) game* is a pair $\langle \mathbf{A}, \mathbf{u} \rangle$, where, for $i \in \{\mathrm{I},\mathrm{II}\}$,

- A_i is the set of *pure strategies* for Player i, and $\mathbf{A} = \langle A_{\mathrm{I}}, A_{\mathrm{II}} \rangle$; and
- $u_i : \mathbf{A} \rightarrow \mathbb{R}$ is the *utility function* for Player i, and $\mathbf{u} = \langle u_{\mathrm{I}}, u_{\mathrm{II}} \rangle$.

We require that $\mathbf{A}$ and $\mathbf{u}$ be common knowledge. If each Player i chooses strategy $a_i \in A_i$, then the payoffs to Players I and II are $u_{\text{I}}(\mathbf{a})$ and $u_{\text{II}}(\mathbf{a})$, respectively. A game is *constant sum* if, for all $\mathbf{a} \in \mathbf{A}$, we have that $u_{\text{I}}(\mathbf{a}) + u_{\text{II}}(\mathbf{a}) = c$ for some fixed c independent of $\mathbf{a}$.

Player i can also play a *mixed strategy* $\alpha_i \in \mathcal{A}_i$, where $\mathcal{A}_i$ is the space of probability measures over the set A_i. Payoff functions are generalized as $u_i(\boldsymbol{\alpha}) = u_i(\alpha_{\text{I}}, \alpha_{\text{II}}) := \mathtt{E}_{\mathbf{a}\sim\boldsymbol{\alpha}}[u_i(\mathbf{a})] = \sum_{\mathbf{a}\in\mathbf{A}} \boldsymbol{\alpha}(\mathbf{a}) u_i(\mathbf{a})$, where $\boldsymbol{\alpha}(\mathbf{a}) = \alpha_{\text{I}}(a_{\text{I}})\alpha_{\text{II}}(a_{\text{II}})$ denotes the joint probability of the independent events that each Player i chooses action a_i from the distribution α_i. This generalization is known as *von Neumann/Morgenstern utility* [51], in which players are indifferent between a guaranteed payoff x and an expected payoff of x.

A *Nash equilibrium* is a pair $\boldsymbol{\alpha}$ of mixed strategies so that neither player has an incentive to unilaterally change his or her strategy. Formally, the strategy pair $\boldsymbol{\alpha}$ is a Nash equilibrium if and only if both $u_{\text{I}}(\alpha_{\text{I}}, \alpha_{\text{II}}) = \max_{\alpha'_{\text{I}} \in \mathcal{A}_{\text{I}}} u_{\text{I}}(\alpha'_{\text{I}}, \alpha_{\text{II}})$ and $u_{\text{II}}(\alpha_{\text{I}}, \alpha_{\text{II}}) = \max_{\alpha'_{\text{II}} \in \mathcal{A}_{\text{II}}} u_{\text{II}}(\alpha_{\text{I}}, \alpha_{\text{II}})$; that is, the strategies α_{I} and α_{II} are *mutual best responses*. A *correlated equilibrium* is a distribution ψ over $\mathbf{A}$ that obeys the following: if $\mathbf{a} \in \mathbf{A}$ is drawn randomly according to ψ and Player i learns a_i, then no Player i has incentive to unilaterally deviate from playing a_i. (A Nash equilibrium is a correlated equilibrium in which $\psi(\mathbf{a}) = \alpha_{\text{I}}(a_{\text{I}}) \cdot \alpha_{\text{II}}(a_{\text{II}})$ is a product distribution.) Formally, in a correlated equilibrium, for every $\mathbf{a} \in \mathbf{A}$ we must have that a_{I} is a best response to a randomly chosen $\hat{a}_{\text{II}} \in A_{\text{II}}$ drawn according to $\psi(a_{\text{I}}, \hat{a}_{\text{II}})$, and analogously for Player II.

2.2 Socratic Game Theory

In this section, we formally define Socratic games. We present our model in the context of two-player games, but of course the multiplayer case fits naturally into this framework. A *Socratic game* is a 6-tuple $\langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$, where, for $i \in \{\text{I}, \text{II}\}$:

- A_i is, as before, the set of pure strategies for Player i.
- W is a set of *possible worlds*, one of which is the *real world* w_{real}.
- $\vec{u_i} = \{u_i^w : \mathbf{A} \to \mathbb{R} |\ w \in W\}$ is a set of payoff functions for Player i, one for each possible world.
- Q_i is a set of available *queries* for Player i. When Player i makes query $q_i : W \to \mathscr{P}(W)$, he or she learns the value of $q_i(w_{\mathsf{real}})$, i.e., the set of possible worlds from which query q_i cannot distinguish w_{real}. We require that (i) $\forall w \in W : w \in q_i(w)$ and (ii) $\forall w, w' \in W : w \in q_i(w') \iff w' \in q_i(w)$.
- $p : W \to [0, 1]$ is a probability distribution over the possible worlds.
- $\delta_i : Q_i \to \mathbb{R}^{\geq 0}$ gives the *query cost* for each available query for Player i.

Initially, the world w_{real} is chosen according to the probability distribution p, but the identity of w_{real} remains unknown to the players. That is, it is as if the players are playing the game $\langle \mathbf{A}, \mathbf{u}^{w_{\mathsf{real}}} \rangle$ but do not know w_{real}. The players make queries $\mathbf{q} \in \mathbf{Q}$; Player i learns a subset $q_i(w_{\mathsf{real}})$ of the possible worlds, one of which is the real world. We consider both *observable* queries and *unobservable* queries. When queries are observable, each player learns *which* query was made by the other player, and the *results* of his or her own query—that is, each Player i learns q_{I}, q_{II}, and $q_i(w_{\mathsf{real}})$. For unobservable queries, Player i learns only q_i and $q_i(w_{\mathsf{real}})$. After learning the results of the queries, the players select strategies $\mathbf{a} \in \mathbf{A}$ and receive as payoffs $u_i^{w_{\mathsf{real}}}(\mathbf{a}) - \delta_i(q_i)$.

In the Socratic game, a pure strategy for Player i consists of a query $q_i \in Q_i$ and a *response function* mapping any result of the query q_i to a strategy $a_i \in A_i$ to play. A player's state of knowledge after a query is a point in $\mathcal{R} := \mathbf{Q} \times \mathscr{P}(W)$ or $\overline{\mathcal{R}}_i := Q_i \times \mathscr{P}(W)$ for observable or unobservable queries, respectively. (Note that there are at most $|\mathbf{Q}| \cdot |W|$ elements of $\mathcal{R}$, and similarly $\overline{\mathcal{R}}_i$, that are consistent with a Socratic game. When we define functions on $\mathcal{R}$, we are content to define them only on this relevant subset.) Thus Player i's response function maps $\mathcal{R}$ or $\overline{\mathcal{R}}_i$ to A_i. Note that the number of pure strategies is exponential, as there are exponentially many response functions. A mixed strategy involves both randomly choosing a query $q_i \in Q_i$ and randomly choosing an action $a_i \in A_i$ in response to the results of the query. Formally, we will consider a *mixed-strategy function profile* $\mathbf{f} = \langle \mathbf{f}^{\mathsf{query}}, \mathbf{f}^{\mathsf{resp}} \rangle$ to have two parts:

- a function $f_i^{\mathsf{query}} : Q_i \to [0, 1]$, where $f_i^{\mathsf{query}}(q_i)$ is the probability that Player i makes query q_i.
- a function f_i^{resp} that maps $\mathcal{R}$ (or $\overline{\mathcal{R}}_i$) to a probability distribution over actions. Player i chooses an action $a_i \in A_i$ according to the probability distribution $f_i^{\mathsf{resp}}(\mathbf{q}, q_i(w))$ for observable queries, and $f_i^{\mathsf{resp}}(q_i, q_i(w))$ for unobservable queries. (With unobservable queries, for example, the probability that Player I plays action a_{I} after making query q_{I} in world w is given by $\mathtt{Pr}[a_{\text{I}} \leftarrow f_{\text{I}}^{\mathsf{resp}}(q_{\text{I}}, q_{\text{I}}(w))]$.)

Mixed strategies are typically defined as probability distributions over the pure strategies, but here we represent a mixed strategy by $\langle \mathbf{f}^{\mathsf{query}}, \mathbf{f}^{\mathsf{resp}} \rangle$. One can easily map a mixture of pure strategies to an $\mathbf{f} = \langle \mathbf{f}^{\mathsf{query}}, \mathbf{f}^{\mathsf{resp}} \rangle$ which induces the same probability of making a particular query q_i or playing a particular action after q_i in a particular world. Thus it suffices to consider this representation of mixed strategies. For a strategy function profile $\mathbf{f}$ for observable queries, the (expected) payoff to Player i is given by

$$\sum_{\mathbf{q}\in\mathbf{Q}, w\in W, \mathbf{a}\in\mathbf{A}} \left[\begin{array}{c} f_{\mathrm{I}}^{\mathsf{query}}(q_{\mathrm{I}}) \cdot f_{\mathrm{II}}^{\mathsf{query}}(q_{\mathrm{II}}) \cdot p(w) \\ \cdot \mathtt{Pr}[a_{\mathrm{I}} \leftarrow f_{\mathrm{I}}^{\mathsf{resp}}(\mathbf{q}, q_{\mathrm{I}}(w))] \\ \cdot \mathtt{Pr}[a_{\mathrm{II}} \leftarrow f_{\mathrm{II}}^{\mathsf{resp}}(\mathbf{q}, q_{\mathrm{II}}(w))] \\ \cdot (u_i^w(\mathbf{a}) - \delta_i(q_i)) \end{array} \right].$$

The payoffs for unobservable queries are analogous, with $f_j^{\mathsf{resp}}(q_j, q_j(w))$ in place of $f_j^{\mathsf{resp}}(\mathbf{q}, q_j(w))$.

3 Strategically Zero-Sum Games

We can view a Socratic game G with constant-sum worlds as an exponentially large classical game, with pure strategies "make query q_i and respond according to f_i." However, this classical game is not constant sum. The sum of the players' payoffs varies depending upon their strategies, because different queries incur different costs. However, this game still has significant structure: the sum of payoffs varies *only* because of varying query costs. Thus the sum of payoffs does depend on players' choice of strategies, but not on the interaction of their choices—i.e., for fixed functions g_{I} and g_{II}, we have $u_{\mathrm{I}}(\mathbf{q},\mathbf{f}) + u_{\mathrm{II}}(\mathbf{q},\mathbf{f}) = g_{\mathrm{I}}(q_{\mathrm{I}}, f_{\mathrm{I}}) + g_{\mathrm{II}}(q_{\mathrm{II}}, f_{\mathrm{II}})$ for all strategies $\langle \mathbf{q}, \mathbf{f}\rangle$. Such games are called *strategically zero sum* and were introduced by Moulin and Vial [36], who introduce a notion of strategic equivalence and define strategically zero-sum games as those strategically equivalent to zero-sum games. It is interesting to note that two Socratic games with the same queries and strategically equivalent worlds are not necessarily strategically equivalent.

A game $\langle \mathbf{A}, \mathbf{u}\rangle$ is *strategically zero sum* if there exist labels $\ell(i, a_i)$ for every Player i and every pure strategy $a_i \in A_i$ such that, for all mixed-strategy profiles $\boldsymbol{\alpha}$, we have that the sum of the utilities satisfies $u_{\mathrm{I}}(\boldsymbol{\alpha}) + u_{\mathrm{II}}(\boldsymbol{\alpha}) = \sum_{a_{\mathrm{I}}\in A_{\mathrm{I}}} \alpha_{\mathrm{I}}(a_{\mathrm{I}}) \cdot \ell(\mathrm{I}, a_{\mathrm{I}}) + \sum_{a_{\mathrm{II}}\in A_{\mathrm{II}}} \alpha_{\mathrm{II}}(a_{\mathrm{II}}) \cdot \ell(\mathrm{II}, a_{\mathrm{II}})$. Note that any constant-sum game is also strategically zero sum.

It is not immediately obvious that one can efficiently decide if a given game is strategically zero sum. For completeness, we give a characterization of classical strategically zero-sum games in terms of the rank of a simple matrix derived from the game's payoffs, allowing us to efficiently decide if a given game is strategically zero sum and, if it is, to compute the labels $\ell(i, a_i)$.

Theorem 3.1. *For a game $G = \langle \mathbf{A}, \mathbf{u}\rangle$ with $A_i = \{a_i^1, \ldots, a_i^{n_i}\}$, let M^G be the n_{I}-by-n_{II} matrix whose $\langle i, j\rangle$th component $M^G_{(i,j)}$ satisfies $\log M^G_{(i,j)} = u_{\mathrm{I}}(a_{\mathrm{I}}^i, a_{\mathrm{II}}^j) + u_{\mathrm{II}}(a_{\mathrm{I}}^i, a_{\mathrm{II}}^j)$. The following are equivalent: (i) G is strategically zero sum; (ii) there exist labels $\ell(i, a_i)$ for every player $i \in \{\mathrm{I}, \mathrm{II}\}$ and every pure strategy $a_i \in A_i$ such that, for all* pure *strategies $\mathbf{a} \in \mathbf{A}$, we have $u_{\mathrm{I}}(\mathbf{a}) + u_{\mathrm{II}}(\mathbf{a}) = \ell(\mathrm{I}, a_{\mathrm{I}}) + \ell(\mathrm{II}, a_{\mathrm{II}})$; and (iii)* $\mathsf{rank}(M^G) = 1$. □

Proof sketch. *(i ⇒ii)* is immediate; every pure strategy is a trivially mixed strategy. For *(ii ⇒iii)*, let $\vec{c_i}$ be the n-element column vector with jth component $2^{\ell(i, a_i^j)}$; then $\vec{c_{\mathrm{I}}} \cdot \vec{c_{\mathrm{II}}}^{\mathsf{T}} = M^G$. For *(iii ⇒i)*, if $\mathsf{rank}(M^G) = 1$, then $M^G = u \cdot v^{\mathsf{T}}$. We can prove that G is strategically zero sum by choosing labels $\ell(\mathrm{I}, a_{\mathrm{I}}^j) := \log_2 u_j$ and $\ell(\mathrm{II}, a_{\mathrm{II}}^j) := \log_2 v_j$. □

4 Socratic Games with Unobservable Queries

We begin with *unobservable* queries, where a player's choice of query is not revealed to her opponent. We give an efficient algorithm to solve unobservable-query Socratic games with strategically zero-sum worlds. Our algorithm is based upon the LP shown in Fig. 1, whose feasible points are Nash equilibria for the game. The LP has polynomially many variables but exponentially many constraints. We give an efficient separation oracle for the LP, implying that the ellipsoid method [17, 27] yields an efficient algorithm. This approach extends the techniques of Koller and Megiddo [28] (see also [29]) to solve constant-sum games represented in *extensive form*, which is similar to a multiplayer decision tree. (Note that their result does not directly apply in our case; even a Socratic game with constant-sum worlds is *not* a constant-sum classical game.)

Lemma 4.1. *Let $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta}\rangle$ be an unobservable-query Socratic game with strategically zero-sum worlds. Any feasible point for the LP in Figure 1 can be efficiently mapped to a Nash equilibrium for G, and any Nash equilibrium for G can be mapped to a feasible point for the program.*

Proof sketch. We describe the correspondence between feasible points for the LP and Nash equilibria for G. First, suppose that strategy profile $\mathbf{f} = \langle \mathbf{f}^{\mathsf{query}}, \mathbf{f}^{\mathsf{resp}}\rangle$ forms a Nash equilibrium for G. Then the following setting for the LP variables is feasible. (We omit the straightforward calculations that verify feasibility.)

$$y^i_{q_i} = \mathbf{f}_i^{\mathsf{query}}(q_i)$$
$$x^i_{a_i, q_i, w} = \mathtt{Pr}[a_i \leftarrow f_i^{\mathsf{resp}}(q_i, q_i(w))] \cdot y^i_{q_i}$$
$$\rho_i = \sum_{w, \mathbf{q}\in\mathbf{Q}, \mathbf{a}\in\mathbf{A}} p(w) \cdot x^{\mathrm{I}}_{a_{\mathrm{I}}, q_{\mathrm{I}}, w} \cdot x^{\mathrm{II}}_{a_{\mathrm{II}}, q_{\mathrm{II}}, w} \cdot [u_i^w(\mathbf{a}) - \delta_i(q_i)].$$

Next, suppose $\langle x^i_{a_i, q_i, w}, y^i_{q_i}, \rho_i\rangle$ is feasible for the LP. Let $\mathbf{f}$ be the strategy function profile defined as

$$f_i^{\mathsf{query}} \;:\; q_i \mapsto y^i_{q_i}$$
$$f_i^{\mathsf{resp}}(q_i, q_i(w)) \;:\; a_i \mapsto x^i_{a_i, q_i, w} / y^i_{q_i}.$$

("Player i does not prefer 'make query q_i, then play according to the function f_i' ")

$\forall q_{\mathrm{I}} \in Q_{\mathrm{I}}, f_{\mathrm{I}} : \overline{\mathcal{R}}_{\mathrm{I}} \to A_{\mathrm{I}}$:

$$\rho_{\mathrm{I}} \geq \sum_{w \in W, a_{\mathrm{II}} \in A_{\mathrm{II}}, q_{\mathrm{II}} \in Q_{\mathrm{II}}, a_{\mathrm{I}} = f_{\mathrm{I}}(q_{\mathrm{I}}, q_{\mathrm{I}}(w))} \left(p(w) \cdot x^{\mathrm{II}}_{a_{\mathrm{II}}, q_{\mathrm{II}}, w} \cdot [u^w_{\mathrm{I}}(\mathbf{a}) - \delta_{\mathrm{I}}(q_{\mathrm{I}})] \right) \quad \text{(I)}$$

$\forall q_{\mathrm{II}} \in Q_{\mathrm{II}}, f_{\mathrm{II}} : \overline{\mathcal{R}}_{\mathrm{II}} \to A_{\mathrm{II}}$:

$$\rho_{\mathrm{II}} \geq \sum_{w \in W, a_{\mathrm{I}} \in A_{\mathrm{I}}, q_{\mathrm{I}} \in Q_{\mathrm{I}}, a_{\mathrm{II}} = f_{\mathrm{II}}(q_{\mathrm{II}}, q_{\mathrm{II}}(w))} \left(p(w) \cdot x^{\mathrm{I}}_{a_{\mathrm{I}}, q_{\mathrm{I}}, w} \cdot [u^w_{\mathrm{II}}(\mathbf{a}) - \delta_{\mathrm{II}}(q_{\mathrm{II}})] \right) \quad \text{(II)}$$

("For every player and every world, it's really a probability distribution")

$$\forall i \in \{\mathrm{I}, \mathrm{II}\}, w \in W : \qquad 1 = \sum_{a_i \in A_i, q_i \in Q_i} x^i_{a_i, q_i, w} \qquad 0 \leq x^i_{a_i, q_i, w} \qquad \text{(III,IV)}$$

("Queries are independent of the world; actions depend only on query output")

$\forall i \in \{\mathrm{I}, \mathrm{II}\}, q_i \in Q_i, w \in W, w' \in W$ s.t. $q_i(w) = q_i(w')$:

$$y^i_{q_i} = \sum_{a_i \in A_i} x^i_{a_i, q_i, w} \qquad x^i_{a_i, q_i, w} = x^i_{a_i, q_i, w'} \qquad \text{(V,VI)}$$

("The payoffs are consistent with the labels $\ell(i, a_i, w)$")

$$\rho_{\mathrm{I}} + \rho_{\mathrm{II}} = \sum_{i \in \{\mathrm{I}, \mathrm{II}\}} \sum_{w \in W, q_i \in Q_i, a_i \in A_i} \left(p(w) \cdot x^i_{a_i, q_i, w} \cdot [\ell(i, a_i, w) - \delta_i(q_i)] \right) \qquad \text{(VII)}$$

Figure 1: An LP to find Nash equilibria in unobservable-query Socratic games with strategically zero-sum worlds. The input is a Socratic game $\langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$ so that world w is strategically zero sum with labels $\ell(i, a_i, w)$. Player i makes query $q_i \in Q_i$ with probability $y^i_{q_i}$ and, when the actual world is $w \in W$, makes query q_i and plays action a_i with probability $x^i_{a_i, q_i, w}$. The expected payoff to Player i is given by ρ_i.

Verifying that this strategy profile is a Nash equilibrium requires checking that $f_i^{\mathsf{resp}}(q_i, q_i(w))$ is a well-defined function (from constraint VI), that f_i^{query} and $f_i^{\mathsf{resp}}(q_i, q_i(w))$ are probability distributions (from III,IV), and that each player is playing a best response to his or her opponent's strategy (from I, II). Finally, from (I,II), the expected payoff to Player i is at most ρ_i. Because the right-hand side of (VII) is equal to the expected sum of the payoffs from $\mathbf{f}$ and is at most $\rho_{\mathrm{I}} + \rho_{\mathrm{II}}$, the payoffs are correct and imply the lemma. □

Recall that a separation oracle is a function that, given a setting for the variables in the LP, either returns "feasible" or returns a particular constraint of the LP that is violated by that setting of the variables. An efficient, correct separation oracle allows us to solve the LP efficiently via the ellipsoid method.

Lemma 4.2. *The separation oracle* SP *(on p. 6) is correct for the LP in Fig. 1 and runs in polynomial time.*

Proof. We first argue that the separation oracle runs in polynomial time and then prove its correctness. Steps 1 and 4 are clearly polynomial. For Step 2, given f_{II} and the result $q_{\mathrm{I}}(w_{\mathsf{real}})$ of the query q_{I}, it is straightforward to compute the probability that, conditioned on the fact that the result of query q_{I} is $q_{\mathrm{I}}(w)$, the world is w and Player II will play action $a_{\mathrm{II}} \in A_{\mathrm{II}}$. Therefore, for each query q_{I} and response $q_{\mathrm{I}}(w)$, Player I can compute the expected utility of each pure response a_{I} to the induced mixed strategy over A_{II} for Player II. Player I can then select the a_{I} maximizing this expected payoff. There are only polynomially many queries, worlds, query results, and pure actions, so the running time of Steps 2 and 3 is thus polynomial.

We now show that the separation oracle works correctly. The main challenge is to show that if any constraint (I-q'_{I}-f'_{I}) is violated then (I-$\hat{q}_{\mathrm{I}}$-$\hat{f}_{\mathrm{I}}$) is violated in Step 4. First, we observe that, by construction, the function $\hat{f}_{\mathrm{I}}$ computed in Step 3 must be a best response to Player II playing f_{II}, no matter what query Player I makes. Therefore the strategy "make query $\hat{q}_{\mathrm{I}}$, then play response function $\hat{f}_{\mathrm{I}}$" must be a best response to Player II playing f_{II}, by definition of $\hat{q}_{\mathrm{I}}$. The right-hand side of each constraint (I-q'_{I}-f'_{I}) is equal to the expected payoff that Player I receives when playing the pure strategy "make query q'_{I} and then play response function f'_{I}" against Player II's strategy of f_{II}. Therefore, because the pure strategy "make query $\hat{q}_{\mathrm{I}}$ and then play response function $\hat{f}_{\mathrm{I}}$" is a best response to Player II playing f_{II}, the right-hand side of constraint (I-$\hat{q}_{\mathrm{I}}$-$\hat{f}_{\mathrm{I}}$) is at least as large as the right hand side of any constraint (I-$\hat{q}_{\mathrm{I}}$-f'_{I}). Therefore, if any constraint (I-q'_{I}-f'_{I}) is violated, constraint (I-$\hat{q}_{\mathrm{I}}$-$\hat{f}_{\mathrm{I}}$) is also violated. An analogous argument holds for Player II. □

These lemmas and the well-known fact that Nash equilibria always exist [37] imply the following theorem:

Separation Oracle SP. On input $\langle x^i_{a_i,q_i,w}, y^i_{q_i}, \rho_i\rangle$:

1. Check each constraint (III, IV, V, VI, VII). If any constraint is violated, return it.
2. Define the strategy profile $\mathbf{f}$ as $f_i^{\mathsf{query}} : q_i \mapsto y^i_{q_i}$ and $f_i^{\mathsf{resp}}(q_i, q_i(w)) : a_i \mapsto x^i_{a_i,q_i,w}/y^i_{q_i}$.
   For each query q_{I}, compute a pure best-response function $\hat{f}_{\text{I}}^{q_{\text{I}}}$ for Player I to strategy f_{II} after making query q_{I}. Let $\hat{f}_{\text{I}}$ be the response function such that $\hat{f}_{\text{I}}(q_{\text{I}}, q_{\text{I}}(w)) = \hat{f}_{\text{I}}^{q_{\text{I}}}(q_{\text{I}}(w))$ for all $q_{\text{I}} \in Q_{\text{I}}$. Similarly, compute $\hat{f}_{\text{II}}$.
3. Let $\hat{\rho}_{\text{I}}^{q_{\text{I}}}$ be the expected payoff to Player I using strategy "make query q_{I} and play response function $\hat{f}_{\text{I}}$" if Player II plays according to f_{II}. Let $\hat{\rho}_{\text{I}} = \max_{q_{\text{I}} \in Q_q} \hat{\rho}_{\text{I}}^{q_{\text{I}}}$ and let $\hat{q}_{\text{I}} = \arg\max_{q_{\text{I}} \in Q_q} \hat{\rho}_{\text{I}}^{q_{\text{I}}}$. Similarly, define $\hat{\rho}_{\text{II}}^{q_{\text{II}}}$, $\hat{\rho}_{\text{II}}$, and $\hat{q}_{\text{II}}$.
4. For the $\hat{f}_i$ and $\hat{q}_i$ defined in Step 3, return constraint (I-$\hat{q}_{\text{I}}$-$\hat{f}_{\text{I}}$) or (II-$\hat{q}_{\text{II}}$-$\hat{f}_{\text{II}}$) if either is violated. If both are satisfied, then return "feasible."

Theorem 4.3. *Nash equilibria can be found in polynomial time for any two-player unobservable-query Socratic game with strategically zero-sum worlds.* □

5 Socratic Games with Observable Queries

In this section, we give efficient algorithms to find (1) a Nash equilibrium for observable-query Socratic games with constant-sum worlds and (2) a correlated equilibrium in the broader class of Socratic games with strategically zero-sum worlds. Recall that a Socratic game $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta}\rangle$ with observable queries proceeds in two stages:

Stage 1: The players simultaneously choose queries $\mathbf{q} \in \mathbf{Q}$. Player i receives as output q_{I}, q_{II}, and $q_i(w_{\mathsf{real}})$.

Stage 2: The players simultaneously choose strategies $\mathbf{a} \in \mathbf{A}$. The payoff to Player i is $u_i^{w_{\mathsf{real}}}(\mathbf{a}) - \delta_i(q_i)$.

Using backward induction, we first solve Stage 2 and then proceed to the Stage 1 game.

For a query $\mathbf{q} \in \mathbf{Q}$, we would like to analyze the Stage 2 "game" $\hat{G}_{\mathbf{q}}$ resulting from the players making queries $\mathbf{q}$ in Stage 1. Technically, however, $\hat{G}_{\mathbf{q}}$ is not actually a game, because at the beginning of Stage 2 the players have different information about the world: Player I knows $q_{\text{I}}(w_{\mathsf{real}})$, and Player II knows $q_{\text{II}}(w_{\mathsf{real}})$. Fortunately, the situation in which players have asymmetric private knowledge has been well studied in the game-theory literature. A *Bayesian game* is a quadruple $\langle \mathbf{A}, \mathbf{T}, r, \mathbf{u}\rangle$, where:

- A_i is the set of *pure strategies* for Player i.
- T_i is the set of *types* for Player i.
- r is a probability distribution over $\mathbf{T}$; $r(\mathbf{t})$ denotes the probability that Player i has type t_i for all i.
- $u_i : \mathbf{A} \times \mathbf{T} \to \mathbb{R}$ is the payoff function for Player i. If the players have types $\mathbf{t}$ and play pure strategies $\mathbf{a}$, then $u_i(\mathbf{a}, \mathbf{t})$ denotes the payoff for Player i.

Initially, a type $\mathbf{t}$ is drawn randomly from $\mathbf{T}$ according to the distribution r. Player i learns his type t_i, but does not learn any other player's type. Player i then plays a mixed strategy $\alpha_i \in \mathcal{A}_i$—that is, a probability distribution over A_i—and receives payoff $u_i(\boldsymbol{\alpha}, \mathbf{t})$. A *strategy function* is a function $h_i : T_i \to \mathcal{A}_i$; Player i plays the mixed strategy $h_i(t_i) \in \mathcal{A}_i$ when her type is t_i. A strategy-function profile $\mathbf{h}$ is a *Bayesian Nash equilibrium* if and only if no Player i has unilateral incentive to deviate from h_i if the other players play according to $\mathbf{h}$. For a two-player Bayesian game, if $\boldsymbol{\alpha} = \mathbf{h}(\mathbf{t})$, then the profile $\mathbf{h}$ is a Bayesian Nash equilibrium exactly when the following condition and its analogue for Player II hold: $\mathtt{E}_{\mathbf{t}\sim r}[u_{\text{I}}(\boldsymbol{\alpha}, \mathbf{t})] = \max_{h'_{\text{I}}} \mathtt{E}_{\mathbf{t}\sim r}[u_{\text{I}}(\langle h'_{\text{I}}(t_{\text{I}}), \alpha_{\text{II}}\rangle, \mathbf{t})]$. These conditions hold exactly if, for all $t_i \in T_i$ occurring with positive probability, Player i's expected utility conditioned on his type being t_i is maximized by $h_i(t_i)$. A Bayesian game is *constant sum* if for all $\mathbf{a} \in \mathbf{A}$ and all $\mathbf{t} \in \mathbf{T}$, we have $u_{\text{I}}(\mathbf{a}, \mathbf{t}) + u_{\text{II}}(\mathbf{a}, \mathbf{t}) = c_{\mathbf{t}}$, for some constant $c_{\mathbf{t}}$ independent of $\mathbf{a}$. A Bayesian game is *strategically zero sum* if the classical game $\langle \mathbf{A}, \mathbf{u}(\cdot, \mathbf{t})\rangle$ is strategically zero sum for every $\mathbf{t} \in \mathbf{T}$. Whether a Bayesian game is strategically zero sum can be determined as in Theorem 3.1. (For further discussion of Bayesian games, see [15, 20].)

We now formally define the Stage 2 "game" as a Bayesian game. Given Socratic game $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta}\rangle$ and a query profile $\mathbf{q} \in \mathbf{Q}$, we define $G_{\mathsf{stage2}}(\mathbf{q}) := \langle \mathbf{A}, \mathbf{T}^{\mathbf{q}}, p^{\mathsf{stage2}(\mathbf{q})}, \mathbf{u}^{\mathsf{stage2}(\mathbf{q})}\rangle$, where:

- A_i, the set of pure strategies for Player i, is the same as in the original Socratic game;
- $T_i^{\mathbf{q}} = \{q_i(w) : w \in W\}$, the set of types for Player i, is the set of possible outcomes of query q_i;
- $p^{\mathsf{stage2}(\mathbf{q})}(\mathbf{t}) = \mathtt{Pr}[\mathbf{q}(w) = \mathbf{t} \mid w \leftarrow p]$; and
- $u_i^{\mathsf{stage2}(\mathbf{q})}(\mathbf{a}, \mathbf{t}) = \sum_{w \in W} \mathtt{Pr}[w \leftarrow p \mid \mathbf{q}(w) = \mathbf{t}] \cdot u_i^w(\mathbf{a})$.

We now define the Stage 1 game in terms of the payoffs for the Stage 2 games. Fix any algorithm alg that finds a Bayesian Nash equilibrium $\mathbf{h}^{\mathbf{q},\mathsf{alg}} := \mathsf{alg}(G_{\mathsf{stage2}}(\mathbf{q}))$ for each Stage 2 game. Define $\mathsf{value}_i^{\mathsf{alg}}(G_{\mathsf{stage2}}(\mathbf{q}))$ to be the expected payoff received by Player i in the Bayesian game $G_{\mathsf{stage2}}(\mathbf{q})$ if each player plays according to $\mathbf{h}^{\mathbf{q},\mathsf{alg}}$, that is,

$$\begin{aligned}&\mathsf{value}_i^{\mathsf{alg}}(G_{\mathsf{stage2}}(\mathbf{q}))\\ &\quad := \textstyle\sum_{w \in W} p(w) \cdot u_i^{\mathsf{stage2}(\mathbf{q})}(\mathbf{h}^{\mathbf{q},\mathsf{alg}}(\mathbf{q}(w)), \mathbf{q}(w)).\end{aligned}$$

Define the game $G_{\mathsf{stage1}}^{\mathsf{alg}} := \langle \mathbf{A}^{\mathsf{stage1}}, \mathbf{u}^{\mathsf{stage1}(\mathsf{alg})} \rangle$, where:

- $\mathbf{A}^{\mathsf{stage1}} := \mathbf{Q}$, the set of available queries in the Socratic game; and
- $u_i^{\mathsf{stage1}(\mathsf{alg})}(\mathbf{q}) := \mathsf{value}_i^{\mathsf{alg}}(G_{\mathsf{stage2}}(\mathbf{q})) - \delta_i(q_i)$.

I.e., players choose queries $\mathbf{q}$ and receive payoffs corresponding to $\mathsf{value}^{\mathsf{alg}}(G_{\mathsf{stage2}}(\mathbf{q}))$, less query costs.

Lemma 5.1. *Let $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$ be an observable-query Socratic game, let $G_{\mathsf{stage2}}(\mathbf{q})$ be the Stage 2 games for all $\mathbf{q} \in \mathbf{Q}$, let alg be an algorithm finding a Bayesian Nash equilibrium in each $G_{\mathsf{stage2}}(\mathbf{q})$, and let $G_{\mathsf{stage1}}^{\mathsf{alg}}$ be the Stage 1 game. Let $\boldsymbol{\alpha}$ be a Nash equilibrium for $G_{\mathsf{stage1}}^{\mathsf{alg}}$, and let $\mathbf{h}^{\mathbf{q},\mathsf{alg}} := \mathsf{alg}(G_{\mathsf{stage2}}(\mathbf{q}))$ be a Bayesian Nash equilibrium for each $G_{\mathsf{stage2}}(\mathbf{q})$. Then the following strategy is a Nash equilibrium for G:*

- *In Stage 1, Player i makes query q_i with probability $\alpha_i(q_i)$. (That is, set $\mathbf{f}^{\mathsf{query}}(\mathbf{q}) := \boldsymbol{\alpha}(\mathbf{q})$.)*
- *In Stage 2, if $\mathbf{q}$ is the query in Stage 1, then Player i chooses action a_i with probability $h_i^{\mathbf{q},\mathsf{alg}}(q_i(w_{\mathsf{real}}))$, where $q_i(w_{\mathsf{real}})$ is the response to the query. (That is, set $f_i^{\mathsf{resp}}(\mathbf{q}, q_i(w)) := h_i^{\mathbf{q},\mathsf{alg}}(q_i(w))$.)* □

We now find equilibria in the stage games for Socratic games with constant- or strategically zero-sum worlds.

Lemma 5.2. *If $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$ is an observable-query Socratic game with constant-sum worlds, then the Stage 1 game $G_{\mathsf{stage1}}^{\mathsf{alg}}$ is strategically zero sum for every algorithm alg, and every Stage 2 game $G_{\mathsf{stage2}}(\mathbf{q})$ is Bayesian constant sum. If the worlds of G are strategically zero sum, then every $G_{\mathsf{stage2}}(\mathbf{q})$ is Bayesian strategically zero sum.* □

Theorem 5.3. *There is a polynomial-time algorithm BNE finding Bayesian Nash equilibria in strategically zero-sum Bayesian (and thus classical strategically zero-sum or Bayesian constant-sum) two-player games.*

Proof sketch. Let $G = \langle \mathbf{A}, \mathbf{T}, r, \mathbf{u} \rangle$ be a strategically zero-sum Bayesian game. Define an unobservable-query Socratic game G^* with one possible world for each $\mathbf{t} \in \mathbf{T}$, one available zero-cost query q_i for each Player i so that q_i reveals t_i, and all else as in G. Bayesian Nash equilibria in G correspond directly to Nash equilibria in G^*, and the worlds of G^* are strategically zero sum. Thus by Thm. 4.3 we can compute Nash equilibria for G^*, and thus we can compute Bayesian Nash equilibria for G. □

(LP's for zero-sum two-player Bayesian games have been previously developed and studied [44].)

Theorem 5.4. *In time $\mathrm{poly}(|\mathbf{A}|, |W|, |\mathbf{Q}|)$, we can compute a Nash equilibrium for any two-player observable-query Socratic game $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$ with constant-sum worlds.*

Proof. Because the worlds of G are constant sum, by Lemma 5.2 we know that the induced Stage 2 games $G_{\mathsf{stage2}}(\mathbf{q})$ are Bayesian constant sum. Thus we can use algorithm BNE to compute a Bayesian Nash equilibrium $\mathbf{h}^{\mathbf{q},\mathsf{BNE}} := \mathsf{BNE}(G_{\mathsf{stage2}}(\mathbf{q}))$ for each $\mathbf{q} \in \mathbf{Q}$, by Theorem 5.3. Furthermore, again by Lemma 5.2, the induced Stage 1 game $G_{\mathsf{stage1}}^{\mathsf{BNE}}$ is classical strategically zero sum. Therefore we can again use algorithm BNE to compute a Nash equilibrium $\boldsymbol{\alpha} := \mathsf{BNE}(G_{\mathsf{stage1}}^{\mathsf{BNE}})$, again by Theorem 5.3. Therefore, by Lemma 5.1, we can assemble $\boldsymbol{\alpha}$ and the $\mathbf{h}^{\mathbf{q},\mathsf{BNE}}$'s into a Nash equilibrium for the Socratic game G. □

We would like to extend our results on observable-query Socratic games to Socratic games with strategically zero-sum worlds. While we can still find Nash equilibria in the Stage 2 games, the resulting Stage 1 game is not in general strategically zero sum. Thus, finding Nash equilibria in observable-query Socratic games with strategically zero-sum worlds seems to require substantially new techniques. However, our techniques for decomposing observable-query Socratic games do allow us to find correlated equilibria in this case.

Lemma 5.5. *Let $G = \langle \mathbf{A}, W, \vec{\mathbf{u}}, \mathbf{Q}, p, \boldsymbol{\delta} \rangle$ be an observable-query Socratic game, let alg be an algorithm finding a Bayesian Nash equilibrium in each of the derived Stage 2 games $G_{\mathsf{stage2}}(\mathbf{q})$, and let $G_{\mathsf{stage1}}^{\mathsf{alg}}$ be the derived Stage 1 game. Let ϕ be a correlated equilibrium for $G_{\mathsf{stage1}}^{\mathsf{alg}}$, and let $\mathbf{h}^{\mathbf{q},\mathsf{alg}} := \mathsf{alg}(G_{\mathsf{stage2}}(\mathbf{q}))$ be a Bayesian Nash equilibrium for each $G_{\mathsf{stage2}}(\mathbf{q})$. Then the following distribution over pure strategies is a correlated equilibrium for G: $\psi(\mathbf{q}, \mathbf{f}) := \phi(\mathbf{q}) \prod_{\langle \mathbf{q}, \mathbf{S} \rangle \in \mathcal{R}} \mathtt{Pr}\left[\mathbf{f}(\mathbf{q}, \mathbf{S}) \leftarrow \mathbf{h}^{\mathbf{q},\mathsf{alg}}(\mathbf{S})\right]$.* □

Thus to find a correlated equilibrium in an observable-query Socratic game with strategically zero-sum worlds, we need only our algorithm BNE from Theorem 5.3 along with an efficient algorithm for finding a correlated equilibrium in a general game. Such an algorithm exists (the definition of correlated equilibria can be directly translated into an LP [3]), and therefore we have the following theorem:

Theorem 5.6. *In polynomial time, we can find a correlated equilibrium for any observable-query two-player Socratic game with strategically zero-sum worlds.* □

By Lemma 5.5 we can also compute correlated equilibria in any observable-query Socratic game for which Nash equilibria are computable in the induced $G_{\mathsf{stage2}}(\mathbf{q})$ games (e.g., when $G_{\mathsf{stage2}}(\mathbf{q})$ is of constant size).

Another potentially interesting model of queries in Socratic games is what one might call *public queries*, in which both the choice *and* outcome of a player's query is observable by all players in the game. (This model might be most appropriate in the presence of corporate espionage or media leaks, or in a setting in which the queries—and thus their results—are done in plain view.) The techniques that we have developed in this section also yield exactly the same results as for observable queries. The proof is actually simpler: with public queries, the players' payoffs are common knowledge when Stage 2 begins, and thus Stage 2 really is a complete-information game. (There may still be uncertainty about the real world, but all players have exactly the same set of possible worlds in which w_{real} may lie; thus they are playing a complete-information game against each other.) Thus we have the same results as in Theorems 5.4 and 5.6 more simply, by solving Stage 2 using a (non-Bayesian) Nash-equilibrium finder and solving Stage 1 as before.

Our results for observable queries are weaker than for unobservable: in Socratic games with worlds that are strategically zero sum but not constant sum, we find only a correlated equilibrium in the observable case, whereas we find a Nash equilibrium in the unobservable case. We might hope to extend our unobservable-query techniques to observable queries, but there is no obvious way to do so. The fundamental obstacle is that the LP's payoff constraint becomes nonlinear if there is any dependence on the probability that the other player made a particular query. This dependence arises in observable queries, suggesting that observable Socratic games with strategically zero-sum worlds may be harder to solve.

6 Related Work

Our work was initially motivated by research in the social sciences indicating that real people seem (irrationally) paralyzed when they are presented with additional options. In this section, we briefly review some of these social-science experiments and then discuss technical approaches related to Socratic game theory.

Prima facie, a rational agent's happiness given an added option can only increase. However, recent research has found that more choices tend to decrease happiness: for example, students choosing among extra-credit options are *more* likely to do extra credit if given a small subset of the choices and, moreover, produce higher-quality work [24]. The psychology literature explores a number of explanations: people may miscalculate their *opportunity cost* by comparing their choice to a "component-wise maximum" of all other options instead of the single best alternative [46], a new option may draw undue attention to aspects of the other options [48], and so on. The present work explores an economic explanation of this phenomenon: *information is not free.* When there are more options, a decision-maker must spend more time to achieve a satisfactory outcome. See, for example, the work of Skyrms [49] for a philosophical perspective on the role of deliberation in strategic situations. Finally, we note the connection between Socratic games and modal logic [23], a formalism for the logic of possibility and necessity.

The observation that human players typically do not play "rational" strategies has inspired some attempts to model "partially" rational players. The typical model of this *bounded rationality* [25, 45, 47] is to postulate bounds on computational power in computing the consequences of a strategy. The work on bounded rationality [13, 14, 38, 43] differs from the models that we consider here in that instead of putting hard limitations on the computational power of the agents, we instead restrict their *a priori* knowledge of the state of the world, requiring them to spend time (and therefore money/utility) to learn about it.

Partially observable stochastic games (POSGs) are a general framework used in AI to model situations of multi-agent planning in an evolving, unknown environment, but the generality of POSGs seems to make them very difficult [4]. Recent work has been done in developing algorithms for restricted classes of POSGs, most notably classes of cooperative POSGs—e.g., [11, 19]—which are very different from the competitive strategically zero-sum games we address in this paper.

The fundamental question in Socratic game theory is deciding on the comparative value of making a more costly but more informative query, or concluding the data-gathering phase and picking the best option, given current information. This tradeoff has been explored in a variety of other contexts; a sampling of these contexts includes aggregating results from information sources that may be slow to respond [6], doing approximate reasoning in intelligent systems [53], deciding when to take the current best guess of disease diagnosis from a belief-propagation network and when to let it continue inference [22], among many others.

This issue can also be viewed as another perspective on the general question of *exploration versus exploitation* that arises often in AI: when is it better to actively seek additional information instead of exploiting the knowledge one already has? (See, e.g., [50].) Most of this work differs significantly from our own in that it considers single-agent planning as opposed to the game-theoretic setting. A notable exception is the work of Larson and Sandholm (see [30]) on mechanism design for interacting agents whose computation is costly and limited. They present a model in which players must solve a computationally intractable valuation problem, using costly computation to learn some hidden parameters, and results for auctions and bargaining games in this model.

7 FUTURE DIRECTIONS

Efficiently finding Nash equilibria in general Socratic games (i.e, with non-strategically zero sum worlds) is probably difficult because such an algorithm is not known for classical games [8, 39, 40]. There has, however, been some algorithmic success in finding Nash equilibria in restricted classical settings (e.g., [12, 32, 33, 42]); we might hope to extend our results to analogous Socratic games.

An efficient algorithm to find correlated equilibria in general Socratic games seems more attainable. Suppose the players receive recommended queries and responses. The difficulty is that when a player considers a deviation from his recommended query, he already knows his recommended response in each of the Stage 2 games. In a correlated equilibrium, a player's expected payoff generally depends on his recommended strategy, and thus a player may deviate in Stage 1 so as to land in a Stage 2 game where he has been given a "better than average" recommended response. (Socratic games are "succinct games of superpolynomial type," so Papadimitriou's results [41] do not imply correlated equilibria for them.)

Socratic games can be extended to allow players to make *adaptive* queries, choosing subsequent queries based on previous results. Our techniques carry over to $O(1)$ rounds of unobservable queries, but it would be interesting to compute equilibria in Socratic games with adaptive observable queries or with $\omega(1)$ rounds of unobservable queries. Special cases of adaptive Socratic games are closely related to single-agent problems like minimum latency [1, 5, 16], determining strategies for using priced information [7, 18, 26], and an online version of minimum test cover [10, 35]. Although there are important technical distinctions between adaptive Socratic games and these problems, approximation techniques from this literature may apply to Socratic games. The question of approximation raises interesting questions even in non-adaptive Socratic games. An *ε-approximate Nash equilibrium* is a strategy profile $\boldsymbol{\alpha}$ so that no player can increase her payoff by an additive ε by deviating from $\boldsymbol{\alpha}$. Finding approximate Nash equilibria in both adaptive and non-adaptive Socratic games is an interesting direction to pursue.

A natural scenario for Socratic games is when $Q = \mathscr{P}(S)$—i.e., each player chooses to make a *set* $q \in \mathscr{P}(S)$ of queries from a specified groundset S of queries. Here we take the query cost to be a linear function, so that $\delta(q) = \sum_{s \in q} \delta(\{s\})$. Natural groundsets include *comparison queries* ("if my opponent is playing strategy a_{II}, would I prefer to play a_{I} or $\hat{a}_{\text{I}}$?"), *strategy queries* ("what is my vector of payoffs if I play strategy a_{I}?"), and *world-identity queries* ("is the world $w \in W$ the real world?"). When one can infer a polynomial bound on the number of queries made by a rational player, then our results yield efficient solutions. (For example, we can efficiently solve games in which every groundset element $s \in S$ has $\delta(s) = \Omega(\overline{M} - \underline{M})$, where $\overline{M}$ and $\underline{M}$ denote the maximum and minimum payoffs to any player in any world.) Conversely, it is NP-hard to compute a Nash equilibrium for such a game when every $\delta(s) \leq 1/|W|^2$, even when the worlds are constant sum and Player II has only a single available strategy. Thus even computing a best response for Player I is hard. (This proof proceeds by reduction from set cover; intuitively, for sufficiently low query costs, Player I must fully identify the actual world through his queries. Selecting a minimum-sized set of these queries is hard.) Computing Player I's best response can be viewed as maximizing a submodular function, and thus a best response can be $(1 - 1/e) \approx 0.63$-approximated greedily [9]. An interesting open question is whether this approximate best-response calculation can be leveraged to find an approximate Nash equilibrium.

[1] A. Archer, D. Williamson. Faster approximation algorithms for the minimum latency problem. *SODA*'03.

[2] R. Aumann. Subjectivity and correlation in randomized strategies. *J. Math. Econ.*, 1:67–96, 1974.

[3] R. Aumann. Correlated equilibrium as an expression of Bayesian rationality. *Econometrica*, 55(1):1–18, 1987.

[4] D. Bernstein, S. Zilberstein, N. Immerman. The complexity of decentralized control of Markov Decision Processes. *Math. of OR*, pp. 819–840, 2002.

[5] A. Blum, P. Chalasani, D. Coppersmith, B. Pulleyblank, P. Raghavan, M. Sudan. The minimum latency problem. *STOC*'94.

[6] A. Broder, M. Mitzenmacher. Optimal plans for aggregation. *PODC*'02.

[7] M. Charikar, R. Fagin, V. Guruswami, J. Kleinberg, P. Raghavan, A. Sahai. Query strategies for priced information. *JCSS*, 64(4):785–819, 2002.

[8] V. Conitzer, T. Sandholm. Complexity results about Nash equilibria. *IJCAI*'03.

[9] G. Cornuejols, M. Fisher, G. Nemhauser. Location of bank accounts to optimize float. *Manag. Sci.*, 23(8):789–810, 1977.

[10] K. De Bontridder, B. Halldórsson, M. Halldórsson, C. Hurkens, J. Lenstra, R. Ravi, L. Stougie. Approximation algorithms for the test cover problem. *Math. Prog.*, 98(1–3):477–491, 2003.

[11] R. Emery-Montemerlo, G. Gordon, J. Schneider, S. Thrun. Approximate solutions for partially observable stochastic games with common payoffs. *Autonomous Agents and Multi-Agent Systems*'04.

[12] A. Fabrikant, C. Papadimitriou, K. Talwar. The complexity of pure Nash equilibria. *STOC*'04.

[13] L. Fortnow, D. Whang. Optimality and domination in repeated games with bounded players. *STOC*'94.

[14] Y. Freund, M. Kearns, Y. Mansour, D. Ron, R. Rubinfeld, R. Schapire. Efficient algorithms for learning to play repeated games against computationally bounded adversaries. *FOCS*'95.

[15] D. Fudenberg, J. Tirole. *Game Theory*. MIT, Cambridge, Massachusetts, 1991.

[16] M. Goemans, J. Kleinberg. An improved approximation ratio for the minimum latency problem. *Math. Prog.*, 82:111–124, 1998.

[17] M. Grotschel, L. Lovasz, A. Schrijver. The ellipsoid method and its consequences in combinatorial optimization. *Combinatorica*, 1:70–89, 1981.

[18] A. Gupta, A. Kumar. Sorting and selection with structured costs. *FOCS*'01.

[19] E. Hansen, D. Bernstein, S. Zilberstein. Dynamic programming for partially observable stochastic games. *AAAI*'04.

[20] J. Harsanyi. Games with incomplete information played by "Bayesian" players. *Manag. Sci.*, 14(3,5,7):159–182, 320–334, 486–502, 1967–1968.

[21] S. Hart, D. Schmeidler. Existence of correlated equilibria. *Math. of OR*, 14(1):18–25, 1989.

[22] E. Horvitz, G. Rutledge. Time-dependent utility and action under uncertainty. *Uncertainty in AI*'91.

[23] G. Hughes, M. Cresswell. *A New Introduction to Modal Logic*. Routledge, 1996.

[24] S. Iyengar, M. Lepper. When choice is demotivating: Can one desire too much of a good thing? *J. Personality and Social Psychology*, 79(6):995–1006, 2000.

[25] E. Kalai. Bounded rationality and strategic complexity in repeated games. *Game Th. App.*, pp. 131–157, 1990.

[26] S. Kannan, S. Khanna. Selection with monotone comparison costs. *SODA*'03.

[27] L. Khachiyan. A polynomial algorithm in linear programming. *Dokklady Akademiia Nauk SSSR*, 244, 1979.

[28] D. Koller, N. Megiddo. The complexity of two-person zero-sum games in extensive form. *Games and Economic Behavior*, 4:528–552, 1992.

[29] D. Koller, N. Megiddo, B. von Stengel. Efficient computation of equilibria for extensive two-person games. *Games and Economic Behavior*, 14:247–259, 1996.

[30] K. Larson. *Mechanism Design for Computationally Limited Agents*. PhD thesis, CMU, 2004.

[31] C. Lemke, J. Howson, Jr. Equilibrium points of bimatrix games. *J. SIAM*, 12, 1964.

[32] R. Lipton, E. Markakis, A. Mehta. Playing large games using simple strategies. *Electronic Commerce*'03.

[33] M. Littman, M. Kearns, S. Singh. An efficient exact algorithm for singly connected graphical games. *NIPS*'01.

[34] R. McKelvey, A. McLennan. Computation of equilibria in finite games. H. Amman, D. Kendrick, J. Rust, Eds., *Handbook of Comput. Econ.* Elsevier, 1996.

[35] B. Moret, H. Shapiro. On minimizing a set of tests. *SIAM J. Scientific Statistical Computing*, 6:983–1003, 1985.

[36] H. Moulin, J. Vial. Strategically zero-sum games: The class of games whose completely mixed equilibria cannot be improved upon. *Intl. J. Game Th.*, 7(3/4), 1978.

[37] J. Nash, Jr. Equilibrium points in n-person games. *Proc. Natl. Acad. Sci.*, 36:48–49, 1950.

[38] A. Neyman. Finitely repeated games with finite automata. *Math. of OR*, 23(3):513–552, 1998.

[39] C. Papadimitriou. On the complexity of the parity argument and other inefficient proofs of existence. *JCSS*, 48:498–532, 1994.

[40] C. Papadimitriou. Algorithms, games, and the internet. *STOC*'01.

[41] C. Papadimitriou. Computing correlated equilibria in multi-player games. *STOC*'05.

[42] C. Papadimitriou, T. Roughgarden. Computing equilibria in multiplayer games. *SODA*'05.

[43] C. Papadimitriou, M. Yannakakis. On bounded rationality and computational complexity. *STOC*'94.

[44] J. Ponssard, S. Sorin. The lp formulation of finite zero-sum games with incomplete information. *Intl. J. Game Th.*, 9(2):99–105, 1980.

[45] A. Rubinstein. *Modeling Bounded Rationality*. MIT, 1988.

[46] B. Schwartz. *The Paradox of Choice*. Ecco, 2004.

[47] H. Simon. *Models of Bounded Rationality*. MIT, 1982.

[48] I. Simonson, A. Tversky. Choice in context: Tradeoff contrast and extremeness aversion. *J. Marketing Research*, 29:281–295, 1992.

[49] B. Skyrms. Dynamic models of deliberation and the theory of games. *TARK*'90.

[50] R. Sutton, A. Barto. *Reinforcement Learning: An Introduction*. MIT, 1998.

[51] J. von Neumann, O. Morgenstern. *Theory of Games and Economic Behavior*. Princeton, 1957.

[52] B. von Stengel. Computing equilibria for two-person games. R. Aumann, S. Hart, Eds., *Handbook of Game Th. with Econ. App.*, vol. 3, pp. 1723–1759. Elsevier, 2002.

[53] S. Zilberstein, S. Russell. Approximate reasoning using anytime algorithms. S. Natarajan Ed., *Imprecise and Approximate Computation*. Kluwer, 1995.