Statistical prediction of the outcome of a game

David H. Wolpert DAVID.H.WOLPERT@NASA.GOV
NASA Ames Research Center
MS 269-1
Moffet Field, CA 94035, USA

Editor:

Abstract

Many machine learning problems involve predicting the joint strategy choice of some goal-directed "players" engaged in a noncooperative game. Conventional game theory predicts that that joint strategy satisfies an "equilibrium concept". The relative probabilities of the joint strategies satisfying that concept are not given, and all other joint strategies are given probability zero. As an alternative, I view this prediction problem as one of statistical inference, where the "data" includes the game specification. This replaces the game theory issue of how to specify a set of equilibrium joint strategies with the issue of how to specify a density function over joint strategies.

I explore a Bayesian version of such a Predictive Game Theory (PGT) using the entropic prior and a likelihood that quantifies the rationalities of the players. A popular game theory equilibrium concept parameterized by player rationalities is the Quantal Response Equilibrium concept (QRE). I show that for some games the local peaks of the posterior density over joint strategies approximate the associated QRE's, and derive the associated correction terms. I also discuss how to estimate parameters of the likelihood from observational data. I end by discussing how PGT can be used to define an equilibrium concept, thereby solving a long-standing problem of conventional game theory.

Keywords: Multi-agent systems, Noncooperative Games, Quantal Response Equilibrium, Bayesian Statistics, Statistical Physics

1. Introduction

Say we have a system of interest that contains some goal-seeking agents. The specification of the utility functions of those agents is potentially very informative data concerning that system's state. Accordingly, it would be useful if we could incorporate such utility function data in statistical modeling of the system using the same kinds of techniques used in machine learning to incorporate other, more conventional kinds of data concerning the system. In this paper I present one way to do this.

1.1 Background

Many scenarios of interest in machine learning involve a set of goal-directed agents. Typically those agents have differing utility functions. In some of these scenarios all the goal-seeking agents are artificial. Example include distributed adaptive control, distributed reinforcement learning (e.g., such systems involving multiple autonomous adaptive rovers on Mars or multiple adaptive telecommunications routers), and more generally multi-agent systems involving adaptive agents (Ferber (1996); Shamma and Arslan (2004); Schaerf et al. (1995); Bieniawski et al. (2005); Tesfatsion and Judd (2006); Kalyanakrishnan et al. (2007); Greenwald et al. (2003); Greenwald and Littman (2007); Brafman and Tennenholtz (2003); Shoham et al. (2007); Mannor and Shamma (2007)). In other instances some of the agents are human beings. Examples here include air-traffic management (Hwang et al. (2007), multi-disciplinary optimization (Cramer et al. (1994); Choi and Alonso (2004)), and in a certain sense, much of mechanism design, and in particular design of auctions (Fudenberg and Tirole (1991); Myerson (1991); Nisan and Ronen (2001)).

Often in such scenarios we can quantify the goals of the agents as utility functions. Traditional machine learning analysis of such scenarios has taken one of three approaches to exploiting knowledge of those utility functions. The first approach is to simply ignore them, and conduct the analysis as if one didn't know them. In the second approach one tries to exploit knowledge of the utility functions by identifying the agents as players in a non-cooperative game defined by the utility functions, and then assuming that the actions of the agents/players are at a Nash Equilibrium (NE) of that game (Myerson (1991); von Neuman and Morgenstern (1944); Luce and Raiffa (1985)). (See Sec. 1.3 for the formal definition of Nash equilibrium.) The third approach explored in machine learning has been to model the agents as automata whose behavior involves their utility functions (Fudenberg and Levine (1998); Jennings et al. (1998); Rubinstein (1998)).

The first approach throws away a major piece of data concerning the system. Indeed, in the scenarios traditionally studied in behavioral game theory (Camerer (2003); Starmer (2000); Hey and Orme (1994); Kahneman (2003b); Tversky and Kahneman (1992); Loomes et al. (1998)), that data is *all* of our information about the system. While the second approach doesn't have this shortcoming, it conflicts with the extensive experimental data that has established that even human beings do not play NE, never mind artificial agents. (See below and Sec. 4 for other problems with using NE.)

While avoiding the shortcomings of the first two approaches, the instances of the third approach explored to date have modeled the agents as simple learning algorithms. However this introduces a dynamics to the analysis even if there is no such dynamics in the scenario under study. Temporal quantities like initial conditions, asymptotic convergence properties, time since initialization, etc. are all crucial issues with systems of interacting learning algorithms. However time may not even be a variable in one's data set. Another difficulty with modeling the agents as learning algorithms is that typically closed-form calculations are impossible and elaborate computer simulations of the entire system are necessary to get results.

In the next subsection I introduce a fourth approach to this problem. In this approach, loosely speaking, I try to exploit knowledge of the utility functions of the players the same way one would exploit other, more conventional machine learning data.

1.2 Approach of this paper

Say one wishes to predict some characteristic of interest y concerning some physical system, based on some information $\mathscr{I}$ concerning the system. Machine learning provides many ways to convert such a $\mathscr{I}$ into a probability distribution over y.[1] Such a distribution is far more informative than a single "best prediction". However if needed we can synopsize the distribution with a single prediction. One way to do that is to use the mode of the distribution as the prediction. When the distribution is a Bayesian posterior probability, $P(y \mid \mathscr{I})$, this mode is called the Maximum A Posterior (MAP) prediction. Alternatively, say there is a real-valued loss function, $L(y, y')$ that quantifies the penalty we will incur if we predict y' and the true value is y. Then Bayesian decision theory counsels us to predict the "Bayes optimal" value, which is the y' that minimizes the posterior expected loss, $\int dy L(y, y') P(y \mid \mathscr{I})$ (Jaynes and Bretthorst (2003); Gull (1988); Loredo (1990); Bernardo and Smith (2000); Berger (1985); Zellner (2004); Paris (1994); Horn (2003)).

A priori, there is no reason that this standard approach to predicting the behavior of physical systems is not appropriate when the physical system in question is some human beings playing a game. To do this we would identify y with the joint choice made by the players in the game. For example, if the players are engaged in a conventional strategic form game, the choice of each player i is a probability distribution over her possible moves in the game (Fudenberg and Tirole (1991); Myerson (1991); Aumann and Hart (1992); Basar and Olsder (1999)). In game theory, this distribution is called the "mixed strategy" of i (or just "strategy" for short). In such strategic form games the moves of the players are independent, so the joint choice of the players — y — is the product of their mixed strategies, which I write as $q \equiv \prod_i q_i$. In this example, $\mathscr{I}$ is the details of the game (e.g., the utility functions of the players), perhaps in conjunction with other information, like quantifications of how rational each player is. So the Bayesian posterior is a distribution over joint strategies, $P(q \mid \mathscr{I})$.

I use the term **Predictive Game Theory** (PGT) to refer to any application of statistical inference (Bayesian or otherwise) *to* games, in contrast to the use of statistical inference by some players *within* a game. The ultimate goal of PGT is to use the same kinds of statistical tools to exploit all information about a system being predicted, whether that information is the utility functions of some players in the system, or some more conventional kind of statistical data concerning the system.

In this paper I focus on PGT for non-cooperative strategic form games, as in the example sketched above (although PGT is also applicable to cooperative games, unstructured bargaining, etc.). Perhaps the primary alternative to this type of PGT is conventional non-cooperative strategic form game theory. That theory predicts that the joint strategy q of a game specified in $\mathscr{I}$ necessarily falls in a set $E(\mathscr{I})$ that is given by applying an "equilibrium concept" (e.g., the NE) to $\mathscr{I}$. The relative probabilities of the strategies $q \in E(\mathscr{I})$ are not specified. Moreover, all joint strategies not in $E(\mathscr{I})$ are assigned probability zero. PGT replaces the game theory issue of how to specify such a set of equilibrium joint strategies for a specified game, $E(\mathscr{I})$, with the issue of how to specify a density function over all possible joint strategies of that game, e.g., a Bayesian posterior $P(q \mid \mathscr{I})$.

There are several important distinctions arising from this difference between the approach of game theory and of PGT. Using the kind of statistical techniques common in machine learning, PGT typically assigns non-zero probability density to a set of joint strategies

[1] In this paper I will sometimes be loose in distinguishing between probability distributions, probability density functions, etc., and will generically write any of them as "$P(\ldots)$" with the context making the meaning clear.

with non-zero measure. In contrast, the sets satisfying equilibrium concepts typically have measure zero. Indeed, historically the motivation behind the use of equilibrium concepts was a desire to make a "point prediction" of a unique joint strategy for any provided game.

In the usual way, a loss function can be used to distill PGT's density function over joint strategies into a single strategy, via decision theory. This mapping of a game to a single Bayes optimal joint strategy can be viewed as an "equilibrium concept". This equilibrium concept depends on the loss function though (which is not specified in the game), unlike conventional equilibrium concepts. Furthermore, this concept typically *does* produce a single joint strategy. This contrasts with conventional equilibrium concepts, which typically need some sort of "refinement" to produce a single strategy.

In addition, the Bayes optimal joint strategy typically is not one under which the players are statistically independent. This is true even though the support of the density function is restricted to joint strategies under which the players *are* statistically independent. Furthermore, often under the Bayes optimal strategy no player's strategy is best-response to the strategies of the other players. Assuming there is more than one NE of the game, this is true even if the players are all fully rational, i.e., if the support of the density over joint strategies is restricted to the NE. In this sense, bounded rationality is automatic under PGT, in contrast to conventional game theory.

To keep this paper tractable I have restricted the type of PGT issues explored here. The first such restriction is to focus on non-repeated strategic form games. The presumption, implicit in much of conventional game theory, is that there is some form of coupling among the players that allows each player i to (partially) account for the other players' utility functions when choosing her moves. However that coupling is not explicitly considered.[2]

The second restriction is to only consider the Bayesian approach to inferring a distribution over q's (Zellner (2004); Loredo (1990)). Thirdly, I restrict attention to $\mathscr{I}$ that consists solely of specification of the game the players are engaged in, together with the rationalities of the players, suitably quantified. This restriction is manifested in what kinds of likelihood are considered. More sophisticated analysis might consider $\mathscr{I}$ that also includes samples of distributions (e.g., of the mixed strategy of the players), that includes the kind of data found in decision analysis, user modeling (Train (2003); Heckerman (1999); Horvitz (2005)), etc. Such analysis would involve different kinds of likelihood from the one considered here, likelihoods designed to integrate knowledge of the utility functions of the agents in a system with other, more conventional probability-based data concerning the agents. In contrast, conventional game theory does not concern posterior distributions over the possible states of a system, and therefore cannot exploit such likelihoods. So it cannot be used to integrate knowledge of utility functions with more more conventional data for prediction purposes.

A fourth restriction is to only explore the Bayesian posterior for an entropic prior over mixed strategies (Loredo (1990)). More sophisticated analysis could consider alternative priors. Finally, I am only considering a likelihood that says, in essence, that the logit Quantal Response Equilibria (QRE) (J. K. Goeree (1999); McKelvey and Palfrey (1995, 1998); Chen and Friedman (1997))) — a very popular model of bounded rationality in

[2]For work' that explicitly does consider such coupling as "learning processes" in repeated form games, see (Fudenberg and Levine (1998); Shoham et al. (2007); Stone (2007); Erev and Roth (2007); Hu and Wellman (2003)).

conventional game theory — is consistent with human behavior in games against Nature. As elaborated below, this is a very weak choice for the likelihood function.

These choices for the prior and likelihood over q's specify the posterior, in the usual way. The QRE q's turn out to be the local peaks of that posterior, for games against Nature. In particular the MAP distribution is a QRE for such a game. However with the same choices of prior and likelihood, the QRE's are only approximations to the local peaks of the posterior for strategic games involving other goal-seeking players. Below I derive conditions under which those approximations are accurate. I also derive correction terms to those approximations.

The likelihood considered in this paper is parameterized by the usual QRE parameters (the exponents in the logit distributions of the players). I call these parameters the "rationalities" of the players. Some joint mixed strategies q (typically almost all of them) are not a QRE for any choice of the player rationalities. For example, this can be the case for certain NE. This can be a major problem for the QRE as a predictive model, since it means that certain data is incompatible with that model. However I show here that every joint strategy — including every NE — has non-zero posterior for an appropriate set of rationalities of the players. So this issue is not a problem for PGT. I also discuss how to estimate rationalities of players from observational data.

In the remainder of this introduction I present some notation. (A formal definition of the NE is presented there.) In the next section I present background on the QRE, the entropic prior, and logit (Boltzmann) distributions. In the following section I present a likelihood function based on the QRE for a game against Nature. I then discuss the posterior over joint mixed strategies given by combining that likelihood with the entropic prior. It is here that I derive sufficient conditions for the QRE's of an N-player game to be the MAP's of the posterior over joint mixed strategies. In this section I also discuss how to estimate the parameters of logit-based distributions in noncooperative games, which is important for actually using PGT. (There seems to be misunderstanding in the experimental game theory literature on valid methods for estimating such parameters, e.g., in the context of the QRE.)

In the following section I step back and discuss in broad terms why the equilibrium concepts of conventional game theory are insufficient for the purposes of PGT, i.e., why conventional game theory is deficient under its positive interpretation. In particular I discuss how and why probability theory, as manifested in PGT, demands (sic) bounded rationality, in contradiction of must of conventional game theory. (This section can be read without first reading the earlier sections.) In the final section I briefly survey some applications of PGT not presented in this paper.

1.3 Notation

Consider a general noncooperative game that has N independent **players**, indicated by the natural numbers $\{1, 2, \ldots, N\}$. Each player i has the finite set of allowed pure strategies $x_i \in X_i$, where $|X_i|$ is the (finite) cardinality of X_i. The set of all possible joint strategies is $X \triangleq X_1 \times X_2 \times \ldots \times X_N$ with cardinality $|X| \triangleq \prod_{i=1}^{N} |X_i|$, a generic element of X being written as x. $u^i : X \to \mathbb{R}$ is player i's utility function, the mixed strategy of i is the distribution $q_i(x_i)$, and $q(x) \triangleq \prod_{i=1}^{N} q_i(x_i)$.

$\Delta_{\mathcal{X}}$ is the Cartesian product of the simplices Δ_{X_i} (implicitly imbued with the standard product topology over simplicial complexes). So mixed joint strategies (i.e., product densities) are elements of $\Delta_{\mathcal{X}}$. The expected utility of player i is written as $\mathbb{E}_q(u^i) = \sum_x \prod_j q_j(x_j) u^i(x)$. I define each player i's **environment** function, often with the associated random variable q_{-i} implicit, as $U^i_{q_{-i}}(x_i) \triangleq \mathbb{E}_{q_{-i}}(u^i \mid x_i)$. I will sometimes write $\mathbb{E}_q(u^i) = q_i \cdot U^i$. As an example, given a set of utility functions, the NE of the associated game is any $q \in \Delta_{\mathcal{X}}$ such that for no $i \in \{1, \ldots N\}$ is there a $q'_i \in \Delta_{X_i}$ where $q'_i \cdot U^i_{q_{-i}} > q_i \cdot U^i_{q_{-i}}$.

Cov is the covariance operator, defined for any countable set of variables $\{y\}$ and associated distribution $p \in \Delta_Y$ by

$$\mathrm{Cov}_p[a(y), b(y)] \;\; \triangleq \;\; \sum_{y \in Y} p(y)a(y)b(y) \; - \; \sum_y p(y)a(y) \sum_y p(y)b(y).$$

(For added clarity, I will sometimes write this as $\mathrm{Cov}_{p(y)}[a(y), b(y)]$.) Given any player i, I will use $-i$ to refer to the set of all $N-1$ other players. In particular, I will sometimes write $q_{-i} \times q_i$ to indicate the $p \in \Delta_{\mathcal{X}}$ with components $p(x) = p(x_i, x_{-i}) \triangleq q_i(x_i)q_{-i}(x_{-i})$.

Curly braces indicate an entire set and vertical bars the cardinality of a finite set, e.g., $\{\beta_i\}$ is the set of all values of β_i for all i, and $|\{\beta_i\}|$ the number of such i. Bold letters, e.g., $\vec{a}$, mean a finite-dimensional vector over the extended real numbers $\mathbb{R}^*$ (i.e., the reals together with positive and negative infinity (Aliprantis and Border (2006))). $\vec{a} \succeq \vec{b}$ indicates the generalized inequality that $\forall i$, $a_i \geq b_i$. $I(.)$ is the indicator function that equals 1 if the equation that is its argument is true and 0 otherwise. Just as "$P(.)$" means a distribution or density function as appropriate, so "$\delta(.)$" indicates the Dirac or Kronecker delta function, as appropriate.

Often no distinction will be made in the notation between finite and infinite spaces, with measures and the like being implicitly matched to the type of space. In particular, sometimes the symbol "$\int$" will be used with the associated measure implicit. As an example, for finite spaces the point-mass measure is presumed, so that $\int$ is equivalent to a sum. Similarly, sometimes for expository simplicity, the term "distribution" will be taken to mean either a distribution or a density, with the context making the precise meaning clear.

The Shannon entropy of a density q is written as $S(q) = -\sum_y q(y)\ln[\frac{q(y)}{\mu(y)}]$. As usual, μ is an *a priori* measure over y, often interpreted as a prior probability distribution. Unless explicitly stated otherwise, here we will always assume it is uniform, and not write it explicitly. (See Jaynes (1957); Jaynes and Bretthorst (2003); Cover and Thomas (1991)).)

To distinguish it from densities like q that we wish to predict from outside of a game, a distribution P that describes our prediction concerning a variable in a game is called a **predictive distribution**.[3] So for example, $P(q \mid \mathscr{I})$, $P(x \mid \mathscr{I}) = \int dq \; P(q \mid \mathscr{I})q(x)$, and $P(q \mid \mathscr{I}, x_j)$ are all predictive distributions. Predictive distributions reflect *our* knowledge/insight/ignorance concerning the game. This contrasts with distributions like q, which reflect the "physical" distributions of the players in the game.

[3]This use of the term "predictive distribution" should not be confused with the one arising in Bayesian statistics.

2. Mathematical Background

Readers familiar with the QRE, entropic prior, and logit distributions are highly encouraged to skip this section, going straight to the "meat" of this paper.

2.1 The QRE

In this paper I consider behavioral formalizations of the limited rationality of humans rather than algorithmic ones. (See Fudenberg and Levine (1998); Hart (2005); Rubinstein (1998); Russell and Subramanian (1995); Georgeff et al. (1999) for examples of formalizations of limited rationality that instead model human thought.) There is an extensive literature on such formalizations. Examples are stochastic preference theory, non-expected utility theory, behavioral game theory in general and prospect theory in particular, etc. (See Starmer (2000); Camerer (2003); Kahneman (2003a); Kurzban and Houser (2005); Fudenberg and Levine (1998); List and Haigh (2005) for excellent overviews of this work.)

Of particular interest here is the behavioral formalization of limited rationality given by the (logit) Quantal Response Equilibrium (QRE). This is a modification of the conventional Nash equilibrium concept where one simultaneously models every players i in a game as playing a mixed strategy $q_i(x_i)$ that is a logit (i.e., Boltzmann) distribution in her expected utilities. More precisely, one predicts that the outcome of the game is a solution to the simultaneous set of equations

$$q_i(x_i) \propto e^{\beta_i \mathbb{E}_q(u^i | x_i)} \; \forall i \tag{1}$$

where the joint distribution $q(x) = \prod_i q_i(x_i)$.

Not all q can be cast as a QRE for some appropriate $\{\beta_i\}$ (see Sec. 3.3 below). So in particular, a q that occurs in the real world will in general differ, even if only slightly, from every possible QRE. This means that with enough experimental data to tightly constrain what q could be, in general we will be able to rule out every QRE as inconsistent with the data. This is a major shortcoming of the QRE (a shortcoming of all equilibrium concepts with a small number of parameters). Another shortcoming is that many games and associated sets $\{\beta_i\}$ have multiple QRE's and the QRE formalism provides no way to assign them relative probabilities (just like with the NE).

I use the notation that $q^*_{\{\beta_i\}}(x)$ means a QRE, where the parameters $\{\beta_i\}$ are often implicit. In general, for any particular game and (non-negative) $\{\beta_i\}$, there is at least one (and may be more than one) associated q^*. This follows from Brouwer's fixed point theorem (McKelvey and Palfrey (1995); Wolpert (2004a)).

At a NE q, simultaneously each player i sets her strategy q_i to maximize her expected utility $\mathbb{E}_{q_i, q_{-i}}(u^i) = \mathbb{E}_{q_i}(U^i_{q_{-i}})$ for the given q_{-i}. Consider modifying this by having q_i instead maximize the associated **free utility**, given by

$$\mathscr{F}_{U^i_{q_{-i}}, T_i}(q_i) \;\; \triangleq \;\; \mathbb{E}_{q_i}(U^i_{q_{-i}}) + T_i S(q_i) \tag{2}$$

for some fixed T_i. For all $T_i \to 0$ the q that simultaneously maximizes $\mathscr{F}_{U^i_{q_{-i}}, T_i} \; \forall i$ is a NE (Wolpert (2004a); McKelvey and Palfrey (1995); Meginniss (1976); Fudenberg and Kreps (1993); Fudenberg and Levine (1993); Luce (1959)). For $T_i > 0$ one instead gets bounded rationality. Indeed, under the identity $T_i \triangleq \beta_i^{-1} \; \forall i$ the solution to this extension of the

Nash equilibrium concept is a QRE.[4] In this, T_i can be viewed as a quantification of the rationality of player i.

In the context of game theory, the free utility Lagrangian has been investigated in Fudenberg and Kreps (1993); Fudenberg and Levine (1993); Shamma and Arslan (2004). The first attempt to derive it from first principles in that game theory context was in Meginniss (1976).

Historically, the QRE was not motivated in terms of free utilities but by modeling payoff uncertainty (McKelvey and Palfrey (1995)). It can also be motivated as the equilibrium of a learning process by the players, a process that is closely related to replicator dynamics (Wolpert (2004b); Anderson et al. (2002); Goeree and Holt (1999)). In addition, in a non-game-theory context, the QRE can be derived from first principles as a way to do distributed control (Wolpert and Rajnarayan (2007); Wolpert et al. (2006)).

Finally, there has been a large body of work relating economics and statistical physics (Brock and Durlauf (2001); Durlauf (1999); Dragulescu and Yakovenko (2000); Aoki (2004); Farmer et al.). (Indeed, there is now an entire field of "econophysics".) Since the logit distribution is the cornerstone of statistical physics (where it occurs in the "canonical ensemble" and the "grand canonical ensemble"), the QRE is also connected to statistical physics. In particular, consider a team game, in which all u^i are the same. Say that all players in such a game have the same rationality, i.e., T_i is independent of i. As discussed in Wolpert (2004a, 2005), for such a game the (shared) free utility essentially becomes what in statistical physics is known as a "mean field approximation" to the "free energy" of a system (hence the terminology "free utility").

More generally, the QRE is akin to the canonical ensemble of statistical physics, with three modifications. First, the single Hamiltonian of a statistical physics system is replaced by the multiple utility functions of an N player game. This means we have multiple Boltzmann distributions rather than one, with move-conditioned expected utility functions playing the role in each such distribution played by the energy spectrum in statistical physics systems. Second, the exponent of those Boltzmann distributions is positive rather than negative, since the bias in an N player game is towards higher utility values, rather than towards lower energy values. Finally, under the QRE the players do not need to have the same "temperature". (Interestingly though, many of the experimental studies of the QRE have assumed that they do share the same temperature.)

2.2 The entropic prior

Shannon was the first person to realize that based on any of several separate sets of very simple desiderata, there is a unique real-valued quantification of the amount of syntactic information in a distribution $P(y)$ (Cover and Thomas (1991); Mackay (2003); Grunwald and Dawid (2004); Topsoe (1979)). He showed that this amount of information is (the negative of) the *Shannon entropy* of that distribution, $S(P)$. Note that for a product distribution $P(y) = \prod_i P_i(y_i)$, entropy is additive: $S(P) = \sum_i S(P_i)$. So for example, the distribution with minimal information is the one that doesn't distinguish at all between the

[4] In McKelvey and Palfrey (1995), $U^i_{q_{-i}}$ is called "a statistical reaction function", and the set of coupled equations giving that solution is called the "logit equilibrium correspondence".

various y, i.e., the uniform distribution. Conversely, the most informative distribution is the one that specifies a single possible y.

Say that the possible values of the underlying variable y in some particular probabilistic inference problem have no known *a priori* stochastic relationship with one another. For example, y may not be numeric, but rather consist of the three symbolic values, {red, dog, Republican}. Then simple desiderata-based counting arguments can be used to conclude that the prior probability of any distribution $p(y)$ is proportional to the *entropic prior*, $\exp(\alpha S(p))$, for some associated finite constant $\alpha \geq 0$ (Cover and Thomas (1991); Mackay (2003); Gull (1988); Loredo (1990); Jaynes and Bretthorst (2003)). [5]

Intuitively, this prior says that absent any other information concerning a particular distribution p, then the larger its entropy the more *a priori* likely it is. Independent of the entropic prior's desideratum-based motivations, it has proven extremely successful in applications ranging from image reconstruction to density estimation to signal processing (Mackay (2003); Gull (1988); Jaynes and Bretthorst (2003)). Indeed, it can be used to derive statistical physics, whose predictions are arguably the best tested in science (Jaynes (1957)). Nonetheless, I do not claim that it is the best possible choice for prior over mixed strategies. Here it serves as a reasonable starting point for exploring PGT.

Under the entropic prior the posterior probability of p given information $\mathscr{I}$ concerning p is

$$P(p \mid \mathscr{I}) \propto \exp(\alpha S(p)) P(\mathscr{I} \mid p). \tag{3}$$

The associated MAP prediction of p is $\text{argmax}_p P(p \mid \mathscr{I})$. As an example, say that $\mathscr{I}$ is a particular element of a partition on the space of possible p's, i.e., a restriction of p to some particular set. Then for any $\alpha > 0$, the MAP p is the one that maximizes $S(p)$, subject to being one of the p's delineated by $\mathscr{I}$.

Intuitively, Eq. 3 pushes us to be conservative in our inference. Of all hypotheses p equally consistent (probabilistically) with our provided information, we are led to view as more *a priori* likely those p that contain minimal extra information beyond that provided in $\mathscr{I}$.[6] For this reason, the entropic prior has been proposed as a formalization of Occam's razor.

Note that the entropic prior evaluated for a product distribution is itself a product, i.e., if $q(x) = \prod_i q_i(x_i)$, then $e^{\alpha S(q)} = \prod_i e^{\alpha S(q_i)}$. As a result, by symmetry the associated marginal over x,

$$\sum_x q(x) P(q) \propto \sum_x \prod_i q_i(x_i) e^{\alpha S(q_i)}, \tag{4}$$

must be uniform over x.

2.3 Miscellaneous properties of logit distributions

Certain simple identities and associated definitions concerning logit distributions will prove useful below. First, given any function $f : Y \to \mathbb{R}$ and $c \in \mathbb{R}$, as in statistical physics I

[5]The issue of how to choose α for a particular application — or better yet integrate over it — is subtle, with a long history. See work on ML-II (Berger (1985)) and the "evidence procedure" (Strauss et al. (1994)).

[6]This is different from saying that the larger the entropy s is, the more *a priori* likely it is that the system has that s: $P_S(s) = \int dp\, \delta(S(p) - s) P(p) = \frac{\int dp\, \delta(S(p)-s) \exp(\alpha S(p))}{\int dp\, \exp(\alpha S(p))}$ which may actually decrease with increasing s, depending on the nature of $\frac{dS}{dp}$.

define the associated **partition function**

$$Z_f(c) \triangleq \sum_y e^{cf(y)} \tag{5}$$

where I implicitly assume that f is bounded. For finite c, the logit (Boltzmann) distribution in values of $f(y)$ having exponent c is defined by

$$\mathcal{L}_{f,c}(y) \triangleq e^{cf(y)}/Z_f(c) \tag{6}$$

for finite c, and for infinite c it is defined by $\mathcal{L}_{f,\infty}(y) \triangleq \delta(y, \text{argmax}f(.))$, $\mathcal{L}_{f,-\infty}(y) \triangleq \delta(y, \text{argmin}f(.))$. Note that for any c and f, $\mathcal{L}_{f,c}(y)$ is uniform over any set of y sharing the same value for $f(y)$.

I define the **Boltzmann utility** as the first moment of f under the logit distribution:

$$K(f,c) \triangleq \sum_y f(y)\mathcal{L}_{f,c}(y). \tag{7}$$

$K(f,c)$ is the expected value of f under the logit distribution in values of f having (potentially infinite) exponent c. The function $K(f,.) : \mathbb{R} \to \mathbb{R}$ is C^∞. Moreover, for any $c \in \mathbb{R}^*$, $K(.,c) : \mathbb{R}^{|Y|} \to \mathbb{R}$ is continuous. (It can be nondifferentiable for infinite c at the point where $f(y) = f(y')$ for some pair $(y, y' \neq y)$.)

A crucial identity in statistical physics which I will use here gives the Boltzmann utility (first moment of f) in terms of the partition function:

$$K(f,c) = \frac{d\ln[Z_f(c)]}{dc}. \tag{8}$$

Similarly, the variance of f under the logit distribution over $f(y)$ values equals the second derivative of $\ln[Z_f(c)]$ with respect to c. This variance is strictly positive for finite c and non-constant f. So for such f, $K(f,.)$ is a nowhere decreasing bijection from $\mathbb{R}^* \to [\min_y f(y)), \max_y f(y))]$. In particular, say we know that $K(f,c) = q \cdot f$ for some given distribution q. Then c is unique.

We can solve for this c satisfying $K(f,c) = q \cdot f$ by solving for the distribution $\mathcal{L}_{f,c}$ that has minimal Kullback-Leibler distance (Mackay (2003); Cover and Thomas (1991)) from q.[7] Similarly, we will often want to find the $p \in \Delta_Y$ that maximizes $S(p)$ subject to the constraint that $p \cdot f = k$. The (unique) solution is the logit distribution $\mathcal{L}_{f,c}$ where c is a Lagrange parameter set to enforce the constraint, i.e., set so that $K(f,c) = k$. For example, consider maximizing the entropy of a player with distribution p in a game against Nature subject to a provided expected value of that player's utility function, $k = \mathbb{E}_p(f)$. The Lagrangian for this problem is the free utility of the player, $\mathscr{F}_{f,c}(p)$. As mentioned at the end of Sec. 2.1, the associated solution for p is the QRE. In this game-against-Nature context, that is just the logit distribution $\mathcal{L}_{f,c}$.

Since for the appropriate value of c, $\mathcal{L}_{f,c}$ is the maximizer over $p \in \Delta_Y$ of $S(p)$ subject to the constraint $f \cdot p = k$, it is also the maximizer of $S(p)$ subject to the constraint

[7] To see this expand the logit distribution inside the integral that defines that Kullback-Leibler distance, $-\sum_y q(y)\ln[\mathcal{L}_{f,c}(y)/q(y)]$, differentiate with respect to c, and then use Eq. 8.

$f \cdot p = K(f, c)$. This can be used to show that the entropy of the logit distribution $\mathcal{L}_{f,c}$ cannot increase as c rises.[8] So the picture that emerges is that as c increases, the logit distribution gets more peaked, with lower entropy. At the same time, it also gets higher associated expected value of f. See Grunwald and Dawid (2004); Topsoe (1979) for other relationships between game theory and entropy.

3. The two kinds of randomness

In PGT we are considering two kinds of randomness. The first is the intrinsic randomness in how the players physically choose their moves in any single play of the game. This randomness is encapsulated in the q_i of each player i. As an example of this kind of randomness, for strategic reasons, a player i might consciously choose her move x_i by randomly sampling an explicitly chosen mixed strategy distribution q_i. (Historically, such strategic considerations were one of the first reasons that game theoreticians considered mixed strategies in addition to pure strategies.) More generally, q_i might not be explicitly chosen by the player, but instead reflect stochasticity in the physical move-choosing process in the player's brain.

The second kind of randomness embodies ignorance that we, the external statistician, have concerning how any particular real human players choose their mixed strategies. This is the usual kind of randomness that underlies the use of probability distributions in Bayesian statistics. It is encapsulated in P, and reflects the fact that we only know $\mathscr{I}$, and from that wish to infer something different, namely q. This randomness can be viewed (though does not need to be) as a "degree of belief" we have in the various q's. Note that $q_i(x_i)$ can differ from $P(x_i \mid \mathscr{I})$ in general.

In this section I present one way to combine those two kinds of randomness, into a posterior distribution $P(q \mid \mathscr{I})$. That posterior over possible joint mixed strategies q is given by the prior and the likelihood. For pedagogical simplicity I have adopted the entropic prior. This means that if we know nothing about the players in a game (so in particular we do not know their utility functions, their rationalities, etc.), then we view a particular almost uniform joint mixed strategy q as *a priori* more likely than a particular highly peaked joint mixed strategy q. α quantifies how much more likely we find such relatively uniform q.

Given a choice for the prior, our next task is to choose the likelihood, i.e., to formalize what we know about the human players *a priori*.

[8] To see this say we replace the invariant $p \cdot f = K(f, c)$ with $p \cdot f \geq K(f, c)$. Entropy is a concave function of its argument, as is this inequality constraint, so our new optimization problem is concave. Therefore the critical point of the associated Lagrangian is the optimizing p. Now if we increase c, and therefore increase $K(f, c)$, the feasible region for our new invariant decreases. This means that when we do that the maximal feasible value of S cannot increase. So the entropy of the critical point of the Lagrangian for our new invariant cannot increase as c does. However that critical point is just the logit distribution $p = \mathcal{L}_{f,c}$, i.e., it is the optimizing p for the original equality invariant, $p \cdot f = K(f, c)$. So the property that increasing c cannot increase the entropy under the new invariant must also hold for the original equality invariant.

3.1 The likelihood

The first thing we know about the players is that under their joint mixed strategy their moves are statistically independent (since we are restricting attention to normal form games). Beyond that, all of the insights of behavioral game theory, psychology, and human modeling (Camerer (2003); Starmer (2000); Allais (1953); List and Haigh (2005); Kurzban and Houser (2005)) could be brought to bear on the task of determining the likelihood.

Here though I will not try to formalize those insights. Instead I will simply assume that the expected utility of any player is uniquely fixed (in a bounded rational way) by her environment, i.e., by the expected utility values of each of her possible moves. I don't even assume that her precise mixed strategy is uniquely fixed by her environment. I only assume that her mixed strategy is in the equivalence class of all her strategies that give some particular (environment-dependent) value for her expected utility.

To formalize this minimal assumption, first consider just those instances in which player i is confronted with some single environment $U^i_{q_{-i}}$. I assume that on average, the move i chooses results in the same utility in all those instances: $q_i \cdot U^i_{q_{-i}}$ has the same (potentially unknown) value in all of them. I write that value as $\epsilon_i(U^i_{q_{-i}})$. $\mathscr{I}$ is the restriction that q is a product distribution and that simultaneously for all players i, $q_i \cdot U^i_{q_{-i}} = \epsilon_i(U^i_{q_{-i}})$. As an example, at a NE $\epsilon_i(U^i_{q_{-i}}) = \max_{x_i} U^i_{q_{-i}}(x_i) \; \forall i$.

This likelihood amounts to saying that as far as player i is concerned when she chooses her move, there is only one salient aspect of q_{-i}. That salient aspect is the effect of q_{-i} on the utility values for i's possible moves, i.e., its effect on $U^i_{q_{-i}}$. The likelihood embodies this aspect of q_{-i} and ignores all other (non-salient) aspects of q_{-i}. In stipulating that only the effects of q_{-i} on her utility are salient to any player i, this likelihood follows the spirit of the axioms of utility theory.

Our next step is to specify the function ϵ_i. To do this I consider how player i would behave in a counterfactual "game against Nature". In that new problem I focus on just one player i, fixing the mixed strategies of the others, so that there are no common knowledge issues, no reasoning about the reasoning of others. The presumption is simply that any player i's expected utility in such a game against Nature is consistent with what it would be if — as in a QRE — she were using a logit mixed strategy for some associated exponent b_i. In essence, I assume that at the very least, the QRE is consistent with a player's expected utility in games against Nature, i.e., that its likelihood is non-zero in such a game.[9]

There is only one QRE for a game against Nature, namely $q^*_i = \mathcal{L}_{U^i_{q_{-i}}, b_i}$ for some appropriate constant b_i, which I call i's **rationality**. Since ϵ_i must give the expected value of $U^i_{q_{-i}}$ under this distribution,

$$\epsilon_i(U^i_{q_{-i}}) \;\; \triangleq \;\; K(U^i_{q_{-i}}, b_i). \tag{9}$$

Our likelihood for this game against Nature is that (q is a product distribution and that) $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$ with q_{-i} being fixed.

[9] Arguably, the rationales offered in McKelvey and Palfrey (1995) for the QRE, involving learning errors, computational errors, etc. are most compelling for such a scenario where the environment is fixed.

As mentioned above, from a statistical physics perspective each U^i is akin to the energy spectrum of a statistical physics system. Adopting this perspective, our QRE-based likelihood is akin to assuming that "expected energy" for each player i is a single-valued function of that player's energy spectrum. For any single such spectrum, many probability distributions across the energy levels will have the same expected energy; our stipulation concerning ϵ_i is only that the actual distribution have the same expected energy as would a Boltzmann distribution with inverse temperature β_i.

Another perspective on the QRE-based likelihood arises by rewriting it as

$$\begin{aligned}[\max_{x_i} U^i_{q_{-i}}(x_i)] - q_i \cdot U^i_{q_{-i}} &= [\max_{x_i} U^i_{q_{-i}}(x_i)] - K(U^i_{q_{-i}}, b_i) \\ &= K(U^i_{q_{-i}}, \infty) - K(U^i_{q_{-i}}, b_i). \end{aligned} \tag{10}$$

The left-hand-side of Eq. 10 can be viewed as the "expected regret" of player i for playing mixed strategy q_i (Shoham et al. (2007)). The right-hand side is independent of q_i, but depends on q_{-i}. This suggests that we simplify Eq. 10, by replacing its right-hand side with some constant $\rho_i > 0$ that is independent of q_{-i}. This would replace the likelihood given by Eq. 10 with one saying that the expected regret of player i must equal ρ_i. Unfortunately though, for any constant ρ_i, if the function $U^i_{q_{-i}}$ is close enough to uniform, then there is no q_i that satisfies $[\max_{x_i} U^i_{q_{-i}}(x_i)] - q_i \cdot U^i_{q_{-i}} = \rho_i$. For such a ρ_i and $U^i_{q_{-i}}$, the likelihood function (and therefore the posterior distribution) would be undefined. The likelihood of Eq. 10 can be viewed as a way to modify such a regret-based likelihood, to ensure that there are always some q_i's with non-zero likelihood, no matter what $U^i_{q_{-i}}$ is.

A nice aspect of this QRE-based likelihood is that it provides a single number for each player that characterizes that player's rationality, namely b_i. For fixed distributions of all the players, that rationality for player i will not change if the utility function of player i is changed by an additive constant, i.e., the likelihood is unchanged if a constant is added to u^i but q does not change. Nor is the rationality of player i sensitive to changes in the lower extreme of the possible values of the player's utility. (Propagated through the likelihood, changes to that lower bound typically force modifications to q_i that don't correspond to significant changes in b_i.) Both of those properties are desirable in any behavioral quantification of a player's rationality.

As in previous theoretical work on the QRE, one can assume that the rationality of each player is provided exogenously, as prior knowledge. Alternatively, like in any other parameterized statistical inference problem, it can be estimated from empirical data. Such estimation has sometimes been done in experimental tests involving the QRE. Some formal approaches to such estimation are briefly recounted below.

It should be emphasized that no claim is being made that the bounded rationality of real human beings is perfectly captured in some single rationality number, and that the associated likelihood of joint mixed strategies exactly obeys Eq. 9. (Particularly objectionable are the facts that for any set $\{b_i\}$, the associated likelihood in Eq. 9 forbids certain regions of Δ_χ, and that it has such a pronounced discontinuity at the border of those forbidden regions.) Rather the QRE-based likelihood is suggested as an approximation to human behavior, one that involves a small number of parameters (the rationalities of the players). While this approximation is primarily intended to illustrate PGT, it is certainly more accurate than the "approximations" of conventional game theory. (See Sec. 4.)

3.2 The posterior

Given the entropic prior and the likelihood of Eq. 9 the posterior for a single player is

$$P(q \mid \mathscr{I}) \quad \propto \quad e^{\alpha S(q)} I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i)) \prod_{j \neq i} \delta(q_j - q'_j) \tag{11}$$

where $\{q'_j : j \neq i\}$ are the pre-fixed (Nature) distributions of all players other than i. The MAP for this game against Nature equals the associated QRE:

Proposition 1 *In a game against Nature with $\alpha > 0$, there is a single local peak of the posterior over q_i for rationality b_i. This peak equals the QRE q_i for the game. If $\alpha = 0$, then there is no q_i with higher posterior than the QRE q_i.*

Proof For $\alpha > 0$ any local peak of the posterior is a distribution $q = (q_i, q'_{-i})$ that maximizes $S(q)$ subject to the constraint that $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$. Since S is additive for product distributions, this q is given by the q_i(s) that maximize $S(q_i)$ subject to the constraint that $q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$. As described at the end of Sec. 2.3, there is a unique such local peak q_i, given by the logit distribution $\mathcal{L}_{U^i_{q_{-i}}, b_i}$. This proves the claim for $\alpha > 0$. The validity of the claim for $\alpha = 0$ is immediate. **QED.** ■

The likelihood for more general games is given by requiring that q be a product distribution and that Eq. 9 hold simultaneously for all players i (other than Nature players). Combining, our full posterior involving all the players is

$$P(q \mid \mathscr{I}) \quad \propto \quad e^{\alpha S(q)} I(q \in \Delta_{\mathcal{X}}) \prod_{i=1}^{N} I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i)). \tag{12}$$

Note that for any QRE q with logit exponents set to the $\{b_i\}$ specified in $\mathscr{I}$, the associated likelihood $P(q \mid \mathscr{I})$ equals 1.

Unlike motivations of the QRE, to motivate this choice of ϵ_i I do *not* say that each q_i must be a logit distribution. The probability density over possible q_i is not assumed to be a delta function about a logit q_i. This reflects an unambiguous fact, ignored in most work on the QRE: in the real world human beings almost never play mixed strategies that are exactly logit distributions in move-conditioned expected utilities. (Often the strategies they play aren't even well approximated by such logit distributions.) Rather I make the weaker presumption that QRE distribution has non-zero likelihood in the single-player inference problem. That presumption motivates a $\vec{b}$-parameterized likelihood that can then be applied in the multi-player scenario. (An even weaker assumption — beyond the scope of this paper — would have $\vec{b}$ be a random variable that is sampled before the game is played.)

Define $\mathscr{I}_{\vec{b}}$ as the set of q such that $\forall i, q_i \cdot U^i_{q_{-i}} = K(U^i_{q_{-i}}, b_i)$, where it is implicitly assumed that $\vec{b} \succeq \vec{0}$. For any such $\vec{b}$ there is always at least one $q \in \mathscr{I}_{\vec{b}}$; every QRE for the set of logit exponents $\vec{b}$ is a member of $\mathscr{I}_{\vec{b}}$. Since the support of the entropic prior is all $\Delta_{\mathcal{X}}$, this means that for any $\vec{b} \succeq \vec{0}$, the posterior conditioned on $q \in \mathscr{I}_{\vec{b}}$ is always non-zero at every $q^*_{\vec{b}}$. Accordingly, the posterior is well-defined (in the sense of being non-zero somewhere). Therefore so are its local peaks, and in particular the associated MAP.

On the other hand, for any finite $\vec{b}$, in general the set of q's that (in addition to the QRE's) satisfy $\mathscr{I}_{\vec{b}}$ has non-zero measure. Indeed, in general $P(q \mid \mathscr{I}_{\vec{b}})$ is non-zero for q's that are not products of logit distributions.

Given the posterior of Eq. 12, one can do many things not possible using the QRE equilibrium concept alone. For example, say one is interested in predicting a single q as the outcome of a provided game, as in conventional game theory. One way to do this is to predict the QRE. However as an alternative, if one is given a loss function over $\Delta_{\mathcal{X}}$, then one can generate such a single q from the posterior $P(q \mid \mathscr{I})$ as the associated Bayes optimal prediction for q. In general, this Bayes optimal prediction will vary with the loss function of the statistician (who is external to the game) who is making the prediction. This contrasts with the QRE, which ignores the concerns of the external statistician when telling that statistician what prediction to make. Moreover, the posterior of Eq. 12 allows us to calculate quantities like posterior variances (in Δ_X) about this Bayes optimal prediction, something not provided by the QRE equilibrium concept.

Alternatively, one can produce a single distribution over x's by marginalizing $P(q \mid \mathscr{I})$ down to a posterior distribution over x's. This distribution does not depend on the statistician's loss function:

$$\begin{aligned} P(x \mid \mathscr{I}) &= \int dq \; P(x \mid q, \mathscr{I}) P(q \mid \mathscr{I}) \\ &= \int dq \; q(x) P(q \mid \mathscr{I}). \end{aligned} \tag{13}$$

In general, under the posterior $P(q \mid \mathscr{I})$ the distributions $\{q_i\}$ are statistically coupled. (Recall that q reflects the players, and P reflects our inference concerning them.) Now for the entropic prior $P(q)$, there is no statistical coupling between x_i and x_j in the prior distribution $P(x)$ (cf. Eq. 4). However the potential coupling between the $\{q_i\}$ means that under $P(x \mid \mathscr{I})$, the moves typically are *not* statistically independent. In such a situation, *to us*, x_i and x_j are statistically coupled.

This is true even if the support of $P(q)$ is restricted to NE of the game, so long as there is more than one such NE. Intuitively, if we observe one player's move, that tells us something about which NE the players have jointly adopted, which in turn tells us something about the other players' likely moves. Hence, the moves are statistically coupled.

This phenomenon means that in some situations the joint mixed strategy $P(x \mid \mathscr{I})$ cannot equal a NE of the underlying game, no matter what the player rationalities are; a NE is impossible. Similarly, so long as more than NE exists for the game, often the Bayes optimal q cannot be one of those NE. These conclusions about the impossibility of a NE do not depend on our choice of ϵ_i, or even on our encapsulating $\mathscr{I}$ in terms of ϵ_i's. (N.b., the phenomenon holds even if $P(q \mid \mathscr{I})$ is restricted to NE.) Rather they arise from the fact that our prior allows non-zero probability for all of the NE.

Example 1 *Consider a common payoff symmetric game involving two players, each with move space* $\{A, B\}$*. Let the shared utility function be* $u(A, A) = 2, u(A, B) = u(B, A) = 0, u(B, B) = 1$*. Say the players are fully rational, so the support of* $P(q \mid \mathscr{I})$ *is restricted to the NE of the game. The game has three NE:* $(A, A), (B, B)$*, and the mixed strategy where each player makes move A with probability 1/3. The first two of those q have entropy 0* *(they are delta functions). The associated value of the entropic prior, $exp(\alpha S(q))/Z(\alpha)$, is just $[Z(\alpha)]^{-1}$. The last NE has entropy ln[3] - 2/3ln[2].*

If we define $w(\alpha) \triangleq exp(\alpha\{ln[3] - 2/3ln[2]\})$, then the prior probability of the first two (pure strategy) NE are $1/[2 + w(\alpha)]$, and the prior probability of the last (mixed strategy) NE is $w(\alpha)/[2 + w(\alpha)]$. Since all three equilibria have the same likelihood (namely, 1), these prior probabilities of the equilibria are also their posterior probabilities, $P(q \mid \mathscr{I})$. Accordingly, by Eq. 13,

$$
\begin{aligned}
P(x = (A, A) \mid \mathscr{I}) &= \frac{1}{2 + w(\alpha)} + \frac{w}{(2 + w)}[\frac{1}{3}]^2 = \frac{9 + w(\alpha)}{9(2 + w(\alpha))}, \\
P(x = (B, B) \mid \mathscr{I}) &= \frac{1}{2 + w(\alpha)} + \frac{w}{(2 + w)}[\frac{2}{3}]^2 = \frac{9 + 4w(\alpha)}{9(2 + w(\alpha))}, \\
P(x = (A, B) \mid \mathscr{I}) &= P(x = (B, A) \mid \mathscr{I}) = \frac{2w}{9(2 + w(\alpha))} \qquad (14)
\end{aligned}
$$

This distribution $P(x \mid \mathscr{I})$ not a NE; under this distribution neither player i plays best-response to $P(x_{-i} \mid \mathscr{I})$. In fact, $P(x \mid \mathscr{I})$ is not even a product distribution.

A third way to produce a single distribution from a posterior is to find its MAP. This is discussed at length for the posterior of Eq. 12 in Sec. 3.4 below. In general, that MAP, the distribution in Eq. 13, and the Bayes optimal q will all differ from the QRE q. See Sec. 4 below for general discussion of how to use PGT to predict single q's and how such predictions are related to equilibrium concepts.

3.3 The posterior q covers all NE

Let q be a NE where for some player i, $R_{i,q} \triangleq \text{supp}[q_i]$ includes multiple $x_i \in X_i$ and q_i is not uniform over $R_{i,q}$. Since q is a NE, $U^i_{q_{-i}}(x_i)$ is uniform over $x_i \in R_{i,q}$. This means that any logit distribution $\mathcal{L}_{U^i_{q_{-i}}, b_i}(x_i)$ must be uniform across all $x_i \in R_{i,q}$. Since by hypothesis q_i is not uniform over $R_{i,q}$, this means that q_i cannot be described by a logit distribution. So such a NE q is not a QRE for any vector of rationalities $\vec{b}$, even one including infinite components. This complicates consideration of NE in terms of QRE's, leading to the analysis of limits of QRE's as $\vec{b} \rightarrow \vec{\infty}$.

Similarly, consider any q where for some player i, for some $\Xi_i \subseteq X_i$ consisting of more than one element, $U^i_{q_{-i}}(x_i)$ is uniform across $x_i \in \Xi_i$, but $q_i(x_i)$ is not. Such a q is not a QRE for any $\vec{b}$. More generally, consider the case where $|X_i| > 2$ for some i. For that case, unless $\ln[q_i(x_i)]$ is a linear function of $U^i_{q_{-i}}(x_i)$, q is not a QRE for any $\vec{b}$.

The QRE concept cannot be used to analyze these types of q's directly (i.e., without resorting to limiting procedures or the like). In particular, if one has experimental data that mandates such a q, then one cannot analyze that data directly with the QRE concept. Note that the difficulty is not the use of a logit distribution by the QRE; any equilibrium concept using a distribution that is purely a function of $U^i_{q_{-i}}(x_i)$ has this difficulty.

Such complications are greatly reduced in PGT. To see this, for every i define $\bar{\Delta}^i_{\mathcal{X}} \subseteq \Delta_{\mathcal{X}}$ as the set of all joint mixed strategies q such that $U^i_{q_{-i}}(x_i)$ is not uniform across all (!) $x_i \in X_i$. Next define $\bar{\Delta}_{\mathcal{X}} \equiv \cap_i \bar{\Delta}^i_{\mathcal{X}}$. So $\bar{\Delta}_{\mathcal{X}}$ is all q where for no i is $U^i_{q_{-i}}(x_i)$ constant-valued. In particular, a NE q is in $\bar{\Delta}_{\mathcal{X}}$ so long as $R_{i,q}$ is a proper subset of X_i for all i (since that means that $\forall\, i$, $\exists\, x_i : q_i(x_i) = 0$, while q_i is best-response to $U^i_{q_{-i}}$). This holds even if for some i's, $R_{i,q}$ contains multiple elements. Then we have the following result:

Proposition 2 *For any i, $\forall q \in \Delta_{\mathcal{X}}$, $\exists\, b_i$ such that $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}$. If $q \in \bar{\Delta}^i_{\mathcal{X}}$, then that b_i is unique. Define $B_i : \bar{\Delta}^i_{\mathcal{X}} \to \mathbb{R}$ as that function taking $q \in \bar{\Delta}^i_{\mathcal{X}}$ to the associated unique rationality of i. Then if $\bar{\Delta}^i_{\mathcal{X}}$ is an open set, B_i is differentiable everywhere in $\bar{\Delta}^i_{\mathcal{X}}$ that it is finite.*

Proof First consider any $q \in \Delta_{\mathcal{X}}$ for which $\max[U^i_{q_{-i}}(x_i)] = \min[U^i_{q_{-i}}(x_i)]$. For this q, for any b_i, $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}$.

Next consider the remaining type of q's, for which $\max[U^i_{q_{-i}}(x_i)] \neq \min[U^i_{q_{-i}}(x_i)]$. For such a q, if $q^i \cdot U^i_{q_{-i}} = \max[U^i_{q_{-i}}(x_i)]$ then we have $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}\ \forall i$ iff $b_i = \infty$. Similarly $K(U^i_{q_{-i}}, b_i) = \min[U^i_{q_{-i}}(x_i)]$ iff $b_i = -\infty$. Now consider the remaining cases, where $q^i \cdot U^i_{q_{-i}} \in (\min[U^i_{q_{-i}}(x_i)], \max[U^i_{q_{-i}}(x_i)])$. Due to the bijectivity of $K(U^i_{q_{-i}}, .)$ with that codomain, we again see that there is a unique b_i such that $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}\ \forall i$. This completes the first claim.

To establish the second claim, simply note that by definition, all elements of $\bar{\Delta}_{\mathcal{X}}$ are of this second type where $\max[U^i_{q_{-i}}(x_i)] \neq \min[U^i_{q_{-i}}(x_i)]$.

To establish the third claim, evaluate the derivative of B_i and show that it is finite. I do this by applying the chain rule to $K(U^i_{q_{-i}}, B_i(q)) - q^i \cdot U^i_{q_{-i}} = 0$. The result for the components $q_i(x_i)$ and $q_{-i}(x_{-i})$ of the argument list of B_i are

$$\frac{\partial B_i(q_i, q_{-i})}{\partial q_i(x_i)} = \frac{U^i_{q_{-i}}(x_i)}{\frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial B_i}}$$

$$\frac{\partial B_i(q_i, q_{-i})}{\partial q_{-i}(x_{-i})} = \frac{\left(q_i(x_i) \cdot \frac{\partial U^i_{q_{-i}}(x_i)}{\partial q_{-i}(x_{-i})}\right) - \frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial q_{-i}(x_i)}}{\frac{\partial K(U^i_{q_{-i}}, B_i)}{\partial B_i}}$$

where the shared denominator is non-zero since B_i is finite by hypothesis and since $K(U^i_{q_{-i}}, .)$ is an increasing function of its second argument. **QED.** ■

As an immediate corollary of Prop. 2, $\forall q \in \bar{\Delta}_{\mathcal{X}}$ there is one and only one $\vec{b}$ such that $K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}}\ \forall i$, and the associated function $B : \bar{\Delta}_{\mathcal{X}} \to \mathbb{R}^N$ is a differentiable function everywhere in $\bar{\Delta}_{\mathcal{X}}$ that it is finite, assuming that $\bar{\Delta}_{\mathcal{X}}$ is an open set.

Typically, the set of points in $\Delta_{\mathcal{X}} \setminus \bar{\Delta}^i_{\mathcal{X}}$ (i.e., all $q : U^i_{q_{-i}}$ is constant-valued) are isolated from one another. This is because changing q_{-i} infinitesimally typically will break a uniformity of $U^i_{q_{-i}}(x_i)$ across $x_i \in X_i$. In this case, the domain of B_i is all of $\Delta_{\mathcal{X}}$ except for a set of points of measure 0, and B_i is differentiable at all q that are outside that measure-0 set and such that $B_i(q)$ is finite.

Every element of $\bar{\Delta}_{\mathcal{X}}$ has a well-defined value of B, by Prop. 2. In contrast, as elaborated above, many of the elements of $\bar{\Delta}_{\mathcal{X}}$ cannot be QRE's for any $\vec{b}$. In particular, even NE in $\bar{\Delta}_{\mathcal{X}}$ cannot be QRE's. In this, the "problem q's" for PGT are a subset of the "problem q's" for the QRE.

Moreover, in many regards the elements of $\Delta_{\mathcal{X}} \setminus \bar{\Delta}_{\mathcal{X}}$ are not as much of a problem for a PGT analysis as they are for a QRE analysis. For any such q, there is at least one i for which $U^i_{q_{-i}}$ is uniform. For that q and that i, there is not a well-defined value of $B_i(q)$; any b_i is consistent with q. Analogously, if for that q the component q_i is uniform, and if q is a QRE for some rationality vector (b_{-i}, b_i), then q is a QRE for (b_i, b'_i) for any b'_i.

On the other hand, if the q_i component is *not* uniform, then the uniformity of $U^i_{q_{-i}}$ means that q could not be a QRE for any b_i. In contrast, even if q_i is non-uniform, the likelihood $I(K(U^i_{q_{-i}}, b_i) = q^i \cdot U^i_{q_{-i}})$ equals 1 for any b_i. So such a q can be analyzed using PGT, whereas it cannot be analyzed (directly at least) using the QRE concept. Indeed, Prop. 2 gives the following result:

Proposition 3 *For any $q \in \bar{\Delta}_{\mathcal{X}}$ there is one and only one $\vec{b}$ such that the posterior $P(q \mid \mathscr{I}_{\vec{b}}) \neq 0$, namely $B(q)$. For any $q \in \Delta_{\mathcal{X}}$, let $\vec{b}^*$ be any rationality vector where $b^*_i \triangleq B_i(q)$ if $q \in \bar{\Delta}^i_{\mathcal{X}}$. Then for all $q' \in \Delta_{\mathcal{X}}$,*

$$\frac{P(q' \mid \mathscr{I}_{\vec{b}^*})}{P(q \mid \mathscr{I}_{\vec{b}^*})} \;\leq\; |X|^{\alpha}$$

where α is the exponent of the entropic prior.

Proof Prop. 2 means that for every $q \in \bar{\Delta}_{\mathcal{X}}$, there is one (and only one) $\vec{b}$ such that the likelihood $P(\mathscr{I}_{\vec{b}} \mid q)$ is non-zero. Since the entropic prior is non-zero for all q, this means that every q has non-zero posterior $P(q \mid \mathscr{I}_{\vec{b}})$ under exactly one $\vec{b}$, as claimed.

By definition, $P(\mathscr{I}_{\vec{b}^*} \mid q) = 1$. Now $P(\mathscr{I}_{\vec{b}^*} \mid q') \leq 1$ for any q'. Accordingly, the ratio in the proposition is bounded above by the ratio of the exponential prior at q to that at q'. However the ratio of $e^{\alpha S(q'')}$ between any two points q'' is bounded below by $\frac{exp(\alpha \cdot 0)}{exp(\alpha \ln(|X|))}$.
QED. ■

In particular, the first part of Prop. 3 holds for any Nash equilibrium $q \in \bar{\Delta}_{\mathcal{X}}$; such equilibria arise for $\vec{b} = \vec{\infty}$. The relative probabilities of those Nash q are given by the ratios of the associated prior probabilities. The second part of Prop. 3 is a bound on how much greater the posterior can be at some $q' \neq q$, given that the rationality vector is consistent with q.

The picture that emerges is that $\forall \vec{b}$, $\exists$ non-empty proper submanifold of $\bar{\Delta}_{\mathcal{X}}$ that is the support of the associated posterior. (Given $\vec{b}$, the associated submanifold is those q's such that $B(q) = \vec{b}$, which we know is non-empty since every $\vec{b}$ has at least one associated QRE.) By the single-valuedness of B, there is no overlap between those submanifolds (one for each $\vec{b}$). Since B is defined over all $\bar{\Delta}_{\mathcal{X}}$, the union of those submanifolds over all $\vec{b}$ is all of $\bar{\Delta}_{\mathcal{X}}$, including any Nash equilibria q's lying in $\bar{\Delta}_{\mathcal{X}}$ (for which $\vec{b} = \vec{\infty}$). All q within a single submanifold have the same value (namely 1) of their likelihoods. Accordingly, the ratios of the posteriors of the q's within the submanifold is given by the ratios of (the exponentials of) the entropies of those q's. This means that within any single one of the submanifolds no q has too small a posterior (cf. Prop. 3).

3.4 The MAP q

Naively, one might presume that a QRE is the MAP of our posterior. After all, this is the case when a single player plays against Nature. Furthermore, when there are multiple players, every QRE q obeys our constraints that $\mathbb{E}_q(u^i) = \epsilon_i(U^i_q)\ \forall i$, and it maximizes the entropy of each player's strategy considered in isolation of the others. However in general a QRE will not maximize the entropy of the joint mixed strategy subject to our constraints when there are multiple players. In other words, while MAP for each individual player's strategy, in general it is not MAP for the joint strategy of all the players. The reason is that setting each separate q_i to maximize the associated entropy (subject to having q obey our invariant), in a sequence, one after the other, will not in general result in a q that maximizes the sum of those entropies. So it will not in general result in a q that maximizes the entropy of the joint system.

Proceeding more carefully, call a local maximum of the posterior that is in the interior of $\Delta_{\mathcal{X}}$ a "local peak" of the posterior. As shorthand, introduce the following notation:

Definition 4 *For all* $j, \vec{b}, \phi \in \mathbb{R}^N$,

$$
\begin{aligned}
q^\dagger_j(x_j) &\triangleq \mathcal{L}_{U^j_{q_{-j}}, b_j}(x_j), \\
r_j(q, x_i) &\triangleq \sum_{x_j} q^\dagger_j(x_j)\mathbb{E}_q(u^j \mid x_i, x_j)\,[1 \;+\; b_j\{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j}\times q^\dagger_j}(u^j)\}], \\
s_i(\phi, x_i) &\triangleq \sum_{j \neq i} \phi_j \left[\mathbb{E}_{q_{-i}}(u^j \mid x_i) - r_j(q, x_i)\right]
\end{aligned}
$$

then we have the following lemma:

Lemma 5 *For a given* $\vec{b}$*, any local peak of the posterior is given by the* q_i *members of a set of pairs* $\{q_i \in \Delta_{X_i}, \lambda_i \in \mathbb{R}\}$ *that simultaneously solves the following equations for all* i*:*

$$
\begin{aligned}
q_i(x_i) &\propto e^{\lambda_i U^i_{q_{-i}}(x_i) + s_i(\lambda, x_i)}, \\
\mathbb{E}_q(u^i) &= K(U^i_{q_{-i}}, b_i).
\end{aligned}
$$

Proof By examination of the posterior, its maxima are q's in $\Delta_{\mathcal{X}}$ that maximize $S(q)$ subject to the constraints in Eq. 9. (Recall that there always exist $q \in \Delta_{\mathcal{X}}$ obeying those constraints.) So the local peaks of the posterior are the critical points of the Lagrangian $\mathscr{L}(q, \{\lambda_i\}) \;=\; S(q) + \sum_i \lambda_i(q_i \cdot U^i - \epsilon_i(U^i)) + \sum_i \gamma_i(\sum_{x_i} q_i(x_i) - 1)$ that obey $q_i(x_i) > 0\ \forall i, x_i$, where the λ_i are Lagrange parameters enforcing the constraints in Eq. 9 and the γ_i are Lagrange parameters forcing each q_i to be normalized. At any such critical point, $\forall\ i, x_i \in X_i$,

$$
\begin{aligned}
0 &= \frac{\partial \mathscr{L}}{\partial q_i(x_i)} \;=\; -1 - \gamma_i - \ln[q_i(x_i)] + \lambda_i \mathbb{E}(u^i \mid x_i) + \sum_{j \neq i} \lambda_j [\mathbb{E}(u^j \mid x_i) - \frac{\partial \epsilon_j(U^j)}{\partial q_i(x_i)}] \\
&= -1 - \ln[q_i(x_i)] + \lambda_i \mathbb{E}(u^i \mid x_i) + \\
&\qquad \sum_{j \neq i} \lambda_j \left[\mathbb{E}_{q_{-i}}(u^j \mid x_i) - \sum_{x_j} \frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)} \mathbb{E}_{q_{-i,-j}}(u^j \mid x_i, x_j)\right].
\end{aligned}
$$

Accordingly, at such q's, for all players i,

$$q_i(x_i) \quad \propto \quad e^{\lambda_i \mathbb{E}_{q_{-i}}(u^i|x_i) + \sum_{j \neq i} \lambda_j \left[\mathbb{E}_{q_{-i}}(u^j|x_i) - \sum_{x_j} \frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)} \mathbb{E}_{q_{-i,-j}}(u^j|x_i,x_j)\right]}$$

where the proportionality constant enforces normalization. By inspection, for any real-valued Lagrange parameters, each such q_i does obey $q_i(x_i) > 0 \; \forall x_i$, as required.

To proceed plug in Eq. 9 and then Eq. 8 to evaluate $\frac{\partial \epsilon_j(U^j)}{\partial U^j(x_j)}$. Then interchange the order of the two differentiations, to differentiate with respect to $U^j(x_j)$ before differentiating with respect to b_j The result is

$$\frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} \quad = \quad q^\dagger_j(x_j)[1 \; + \; b_j\{U^j_{q_{-j}}(x_j) - \mathbb{E}_{q^\dagger_j}(U^j_{q_{-j}})\}]$$

where I have made explicit the dependence of each U^j on q_{-j}. Next use the definition of $U^j_{q_{-j}}$ and the fact that q is a product distribution to expand this result as

$$\frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} \quad = \quad q^\dagger_j(x_j)[1 \; + \; b_j\{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times \mathcal{L}_{U^j_{q_{-j}}, b_j}}(u^j)\}].$$

Now plug this result into the outer summands in our equation above for each $q_i(x_i)$, getting

$$\sum_{x_j} \frac{\partial \epsilon_j(U^j_{q_{-j}})}{\partial U^j_{q_{-j}}(x_j)} \mathbb{E}_q(u^j \mid x_i, x_j)$$

$$= \sum_{x_j} q^\dagger_j(x_j) \mathbb{E}_q(u^j \mid x_i, x_j) \; [1 \; + \; b_j\{\mathbb{E}_{q_{-j}}(u^j \mid x_j) - \mathbb{E}_{q_{-j} \times q^\dagger_j}(u^j)\}].$$

Plugging in the definition of r_j completes the proof. **QED.** ■

In particular, the MAP is a local peak of the posterior. Therefore if the MAP is interior to $\Delta_{\mathcal{X}}$ it must solve the coupled set of equations given in Lemma 5.

3.5 The modes of $P(q \mid \mathscr{I})$ and the QRE's

It is illuminating to compare the conditions of Lemma 5 for q to be a local peak of the posterior to conditions for it to be a QRE: any QRE is given by the q_i members of a set of values $\{q_i \in \Delta_{X_i}, \lambda'_i \in \mathbb{R}\}$ that simultaneously solves the following equations for all i:

$$\begin{aligned} q_i(x_i) \quad &\propto \quad e^{\lambda'_i U^i_{q_{-i}}(x_i)}, \\ \mathbb{E}_q(u^i) \quad &= \quad K(U^i_{q_{-i}}, b_i) \end{aligned} \tag{15}$$

where the second equation forces $\lambda'_i = b_i \; \forall i$.[10]

[10]To see this, use the first equation to write $\mathbb{E}_q(u^i) = K(U^i_{q_{-i}}, \lambda'_i)$, and recall that $K(.,.)$ is monotonically increasing in its second argument.

This comparison suggests that in some circumstances the QRE is an approximation of the local peaks of the posterior $P(q \mid \mathscr{I})$. To confirm this, first note that, ultimately, the only free parameter in our solution for the local peak q's is $\vec{b}$. In addition, any QRE q^* is a solution to a set of coupled nonlinear equations parameterized by $\vec{b}$. In general there is a very complicated relation between the the local peak q's and the q^*'s, one that varies with $\vec{b}$ (as well as with the $\{u^j\}$, of course).

Intuitively, the reason for the difference between the two solutions is that each player i does not operate in a fixed environment, but rather in one containing intelligent players trying to adapt their moves to take into account i's moves. This is embodied in the likelihood of Eq. 9. In contrast to that likelihood, the likelihoods of the QRE each implicitly assume that the associated player i operates in a fixed environment.

Formally, the difference arises because each q_i not only appears in the first term in the argument of $I(q_i \cdot U_{q_{-i}} = K(U_{q_{-i}}, b_i))$ (which is the case in the game against Nature). It also occurs in the second arguments of $I(q_j \cdot U_{q_{-j}} = K(U_{q_{-j}}, b_j))$ for the players $j \neq i$. This means that if we change q_i, then the likelihood of Eq. 9 induces a change to q_{-i}, to have the invariant for the players other than i still be satisfied. This change to q_{-i} then induces a "second order" change to q_i, to satisfy the invariant for player i.

This second-order effect will not arise in a game against Nature, which treats the other players as fixed. This reflects the fact that such a game against Nature is an instance of decision theory, lacking the common knowledge aspect of games with multiple conflicting players.

Now in general it is not the case that for every i, $q_i(x_i)$ equals $q_i^\dagger(x_i)$ on an x_i-by-x_i basis. Indeed, if this were the case then q would be a QRE. However as an approximation we can impose the weaker condition that the differences between those distributions approximately cancel out inside the appropriate sum from Lemma 5:

$$\begin{aligned}
\mathbb{E}_{q_j^\dagger \times q_{-j,-i}}(u^j \mid x_i) &= \sum_{x_j} q_j^\dagger(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j) \\
&\approxeq \sum_{x_j} q_j(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j) \\
&= \mathbb{E}_{q_{-i}}(u^j \mid x_i). \qquad (16)
\end{aligned}$$

(In particular, any QRE obeys this approximation exactly.) Making this approximation inside all r_j, for any $\phi \in \mathbb{R}^N$,

$$s_i(\phi, x_i) \quad = \quad -\sum_{j \neq i} \phi_j b_j \sum_{x_j} q_j^\dagger(x_j)\mathbb{E}_{q_{-j,-i}}(u^j \mid x_i, x_j) \left[\mathbb{E}_{q_{-j}}(u^j \mid x_j) \; - \; \mathbb{E}_{q_{-j} \times q_j^\dagger}(u^j)\right] \qquad (17)$$

which we can write as

$$s_i(\phi, x_i) \quad = \quad -\sum_{j \neq i} \phi_j b_j \mathrm{Cov}_{q_j^\dagger(x_j)} \left[\mathbb{E}_q(u^j \mid x_i, x_j), \; \mathbb{E}_q(u^j \mid x_j)\right]. \qquad (18)$$

Combined with Lemma 5 this provides the following result:

Theorem 6 *Let q be a joint mixed strategy where $\exists\ \mu \in \mathbb{R}^N$ and $t \in \mathbb{R}$ such that simultaneously $\forall i, x_i$,*

$$\sum_{j \neq i} \mu_j b_j \, Cov_{q_j(x_j)}[\mathbb{E}_q(u^j \mid x_j, x_i), \mathbb{E}_q(u^j \mid x_j)] \quad = \quad (\mu_i - b_i)\mathbb{E}_q(u^i \mid x_i) + t.$$

Then the following two conditions are equivalent:
i) q is a QRE.
ii) q is a local peak of the posterior and obeys Eq. 16 exactly.

Proof It is immediate that if q is a QRE then it obeys Eq. 16 exactly. This means that $s_i(\mu, x_i)$ equals the expression in Eq. 18 for $\phi = \mu$. Accordingly, the condition in the theorem involving a sum of covariances means that the exponent in Lemma 5 reduces to $(b_i + \lambda_i - \mu_i)\mathbb{E}_{q_{-i}}(u^i \mid x_i) - t$ for all $i, x_i \in X_i$. So by that lemma, for our q to be a local peak of the posterior it suffices for there to be a $\lambda \in \mathbb{R}^N$ such that $\mathbb{E}_q(u^i) = K(U^i_{q_{-i}}, b_i)$ and $q_i(x_i) \propto e^{(b_i + \lambda_i - \mu_i)\mathbb{E}_{q_{-i}}(u^i | x_i)}$ for all i, x_i. Since q is a QRE with exponent b_i, both of these conditions are met for $\lambda = \mu$. Therefore q is a local peak of the posterior, as claimed.

To prove the converse, plug the condition in the theorem involving a sum of covariances with $\phi = \mu$ into the expression in Eq. 18 for $s_i(\phi, x_i)$. Identifying $\lambda' = \lambda - \mu + b$, this reduces Lemma 5 to the conditions in Eq.'s 15 sufficient for q to be a QRE. **QED.** ∎

In particular, say that $\sum_{j \neq i}(b_j)^2 \text{Cov}_{q^*_j(x_j)}[\mathbb{E}_{q^*}(u^j \mid x_j, x_i), \mathbb{E}_{q^*}(u^j \mid x_j)]$ is independent of x_i $\forall i$ at some QRE q^*. Then the condition in Thm. 1 holds, with $\mu = \vec{b}$. So any such QRE is a local peak of the posterior.

Particularly for very large systems (e.g., a human economy), it may be that at some QRE q^*, $\mathbb{E}_{q^*}(u^j \mid x_j, x_i) = \mathbb{E}_{q^*}(u^j \mid x_j)$ for almost any i, j and associated moves x_i, x_j. In this situation, at the QRE the move of almost any player i has no effect on how the expected payoff to player j depends on j's move. If this is in fact the case for player i and all other players j, then the covariance for each j, x_i that occurs in Thm. 6 reduces to the variance of $\mathbb{E}_{q^*}(u^j \mid x_j)$ as one varies x_j according to q^*_j. By the discussion in Sec. 2.3 this variance is given by the partition function:

$$\text{Var}_{q^*_j}(\mathbb{E}_{q^*}(u^j \mid x_j)) \quad = \quad \text{Var}_{q^*_j}(U^j_{q^*}) \;=\; \frac{\partial^2 \text{ln}(Z_{U^j_{q^*}}(b'_j))}{\partial (b'_j)^2}|_{b'_j = b_j}. \tag{19}$$

In particular, for $b_j \to \infty$ — perfectly rational behavior on the part of agent j — the variance goes to 0. So assume that Eq. refeq:approx holds to a very good approximation. Then if every player i is "decoupled" from all other players, in the limit that all players become perfectly rational the condition in Thm. 6 generically is met for $\mu = \vec{b}$. (The b_j-dependence in the covariance occurs in an exponent, and therefore generically overpowers the $(b_j)^2$ multiplicative factor.) So the QRE's approach the local peaks of the posterior in that situation.

On the other hand, if the players have bounded rationality, their variances are nonzero. In this case the expression in Thm. 1 is nonzero for each i, j, x_i. Typically for fixed i the precise nonzero value of that variance will vary with x_i. In this case, Thm. 1 suggests that the QRE differs from the local peaks of the posterior, and in particular differ from the MAP.

There are many ways that these results can be extended. For example say a particular QRE is a local peak of the posterior for some $\vec{b}$. Then we can use a Laplace expansion to approximate the posterior in the vicinity of that QRE as a Gaussian projected onto the submanifold of joint mixed strategies that obey Eq. 15 (Robert and Casella (2004)). Say that that QRE is close to the mean of the posterior over q's (e.g., this would be the case if that QRE is the MAP and the posterior is sharply peaked). Then our Gaussian approximation could be used to approximate the variance of any function of q under our posterior.

3.6 Estimating player rationalities

Say we do not know the player rationalities, but have to estimate them from data. There are several such scenarios to distinguish. In one, the data consists of a single q generated by sampling $P(q \mid \mathscr{I})$. (For example, we would have this scenario if all the players report a mixed strategy, and we believe what they report to be true.) Say that for this q, for no player i is $U^i_{q_i}$ uniform across X_i. Then for every i, $K(U_{q_{\underline{i}}}, .)$ is a bijection from $\mathbb{R} \rightarrow \mathbb{R}$ (see Sec. 2.3). Accordingly, in this case we can solve for the unique b_i such that $K(U_{q_{\underline{i}}}, b_i) = q_i \cdot U_{q_{\underline{i}}}$. This gives the single $\vec{b}$ with non-zero likelihood, and therefore the single $\vec{b}$ that has non-zero posterior given our data.

Alternatively, say our data is a set D that consists of m IID samples of q. Then the problem of estimating $\vec{b}$ from D is the problem of estimating a hyperparameter from sample data. We can do this with many conventional techniques. As an example, let $D(j)$ refer to our j'th sample of q, and let γ be our other information in addition to D, namely the game specification. So the likelihood of $\vec{b}$ is given by

$$
\begin{aligned}
P(D \mid \vec{b}, \gamma) &= \int dq \; P(D \mid \vec{b}, q, \gamma) P(q \mid \vec{b}, \gamma) \\
&= \int dq \; \left[\prod_{j=1}^{m} P(D(j) \mid q) \right] P(q \mid \vec{b}, \gamma). \qquad (20)
\end{aligned}
$$

If we have a prior over $\vec{b}$, we can convert this likelihood into a posterior over $\vec{b}$. Such a prior also provides us with a posterior over q:

$$
\begin{aligned}
P(q \mid D, \gamma) &\propto \int d\vec{b} \; P(D \mid \vec{b}, q, \gamma) P(q \mid \vec{b}, \gamma) P(\vec{b} \mid \gamma) \\
&= \int d\vec{b} \; \left[\prod_{j=1}^{m} P(D(j) \mid q) \right] P(q \mid \vec{b}, \gamma) P(\vec{b} \mid \gamma)
\end{aligned}
$$

where $P(q \mid \vec{b}, \gamma)$ is the posterior considered in the earlier part of this paper.

If one does not have a prior over $\vec{b}$, one can still estimate it from D, for example by using ML-II (Berger (1985); Bernardo and Smith (2000)). Under that technique one estimates $\vec{b}$ as the value that maximizes the likelihood given in Eq. 20. To illustrate this, say that m

is large enough so that we can approximate $\prod_{j=1}^{m} P(D(j) \mid q) \approxeq \delta(q - \nu(D))$, where ν is the frequency counts function that maps its argument to the associated (perhaps Laplace-corrected) normalized histogram over X. Then under formal conditions that are often met in practice, we can approximate the integral in Eq. 20 by

$$P(D \mid \vec{b}, \gamma) \approxeq P(\nu(D) \mid \vec{b}, \gamma). \tag{21}$$

The right-hand side of this equation is the posterior $P(q \mid \mathscr{I})$ of Eq. 12, evaluated for $q = \nu(D)$ and for $\mathscr{I}$ having the rationality values $\vec{b}$. So in this limit of large enough m, using ML-II to estimate $\vec{b}$ is identical to the $\vec{b}$-estimation technique discussed just above where our provided data is a single q, with that single q set to $\nu(D)$. Accordingly, for such large enough m, we can invert K to solve for $\vec{b}$, in the manner discussed just above. For m that are not sufficiently large, to solve for the ML-II estimate of $\vec{b}$ we must calculate correction terms to this $\vec{b}$ found by inverting K.

As an alternative approximation, say that we replace $P(q \mid \vec{b}, \gamma)$ with $\delta(q - q^*_{\vec{b}})$, i.e., assume that q is a QRE.[11,12] Under this approximation,

$$\begin{aligned} P(D \mid \vec{b}, \gamma) &\approxeq P(D \mid q^*_{\vec{b}}) \\ &= \prod_{j=1}^{m} P(D(j) \mid q^*_{\vec{b}}) \end{aligned} \tag{22}$$

where for simplicity γ is implicit on the right hand side. So ML-II in this approximation reduces to solving for the set of rationalities $\vec{b}$ that maximizes the likelihood of D given the QRE for $\vec{b}$. Assume for simplicity that the correspondence $\vec{b} \to q^*_{\vec{b}}$ produces a singleton for all $|bfb$. Then the maximizing $\vec{b}$ can be found by doing gradient ascent over the function $\vec{b} \to P(D \mid q^*_{\vec{b}})$. That gradient is

$$\frac{\partial P(D \mid q^*_{\vec{b}})}{\partial \vec{b}} = \sum_{k=1}^{m} \frac{\partial P(D(k) \mid q^*_{\vec{b}})}{\partial \vec{b}} \left[\prod_{j \neq k} P(D(j) \mid q^*_{\vec{b}}) \right]. \tag{23}$$

To calculate this gradient for a current $\vec{b}$ we must evaluate $P(D(n) \mid q^*_{\vec{b}})$ for all $n \in 1, \ldots m$. To that end simply note that

$$P(D(n) \mid q^*_{\vec{b}}) = \prod_{i=1}^{N} q^*_{i,\vec{b}}(D^i(n)) \tag{24}$$

where $D(n) \triangleq (D^1(n), D^2(n), \ldots, D^N(n))$. Accordingly, to evaluate each $P(D(n) \mid q^*_{\vec{b}})$ for a current $\vec{b}$, we only need use a numerical fixed-point solution algorithm to calculate each term $q^*_{i,\vec{b}}(D^i(n))$.

[11] Formally, this approximation is equivalent to replacing the entropic prior over $\Delta_{\mathcal{X}}$ with a prior $P(q)$ whose support is restricted to the q's that are QRE's for some associated set of player rationalities.

[12] For simplicity, in the current discussion I ignore the possibility that there might be multiple QRE for some $\vec{b}$.

To complete the calculation of $\frac{\partial P(D|q^*_{\vec{b}})}{\partial \vec{b}}$ given in Eq. 24 we also need to evaluate the partial derivatives $\frac{\partial P(D(k)|q^*_{\vec{b}})}{\partial \vec{b}}$. The formula for those partial derivatives is given by using implicit differentiation on the fixed point set of coupled equations that defines $q^*_{\vec{b}}$. This formula involves some matrix inversions and $q^*_{\vec{b}}(D(n))$, which can be evaluated numerically as discussed just above. (See Wolpert and Kulkarni (2008)).

It is important to realize that in general the i'th component of the $\vec{b}'$ that maximizes the likelihood $P(D \mid q^*_{\vec{b}'})$ differs from the b_i that maximizes the likelihood of i's data given the non-i components of $q^*_{\vec{b}'}$. Formally, defining $\vec{b}'$ as the maximizer of $P(D \mid q^*_{\vec{b}'})$ and recalling that player i's logit distribution for environment $U^i_{q^*_{\vec{b}'_{-i}}}$ and rationality b_i is written as $\mathcal{L}_{U^i_{q^*_{\vec{b}'_{-i}}}, b_i}$,

$$\vec{b}'_i \quad \neq \quad \text{armgax}_{b_i} P(D^i(1), D^i(2), \ldots D^i(m) \mid q^*_{\vec{b}'_{-i}}, \mathcal{L}_{U^i_{q^*_{\vec{b}'_{-i}}}, b_i}). \tag{25}$$

In other words, the maximum likelihood QRE assigns a different rationality to each player i from the one that maximizes the likelihood of player i's data sample considered by itself. Intuitively, in finding the maximum likelihood QRE, we must distort our estimate of b_i to account for the effects of b_i on the likelihoods of data samples of the *other* players' mixed strategies. Similarly, in general

$$\vec{b}'_i \quad \neq \quad \text{armgax}_{b_i} P(D^i(1), D^i(2), \ldots D^i(m) \mid \nu(D)_{-i}, \mathcal{L}_{U^i_{q^*_{\vec{b}'_{-i}}}, b_i}). \tag{26}$$

These effects are not accounted for in much of the behavioral game theory literature in which QRE's are estimated from empirical data.

As a final comment, it is worth noting that in practice sampling of a human being's mixed strategy is almost never a stationary process. Like almost all of the behavioral game theory literature, the analysis above ignores this and related deep problems in estimating the mixed strategy of a human purely from her sample behavior. See Wolpert et al. (2008) for a discussion of these problems.

4. Equilibrium concepts, bounded rationality and PGT

In this section I compare Bayesian PGT for noncooperative games with conventional noncooperative game theory from a broad perspective, not restricted to any particular prior or likelihood. In particular I highlight what shortcomings of conventional noncooperative game theory are overcome by PGT.

4.1 The two equilibrium concepts of PGT

Say we have information $\mathscr{I}$ about a game involving a set of human players. We want to predict what mixed joint strategy q those humans will play. Adopting the role of a Bayesian statistician external to the physical system of those humans, to make this prediction means determining the posterior $P(q \mid \mathscr{I})$. This contrasts with what conventional game-theoretic equilibrium concepts provide, which is a subset of all possible q's with no associated probability values (except in the degenerate sense that if that set contains a single element we can interpret it as having probability 1.0). Due to this difference, PGT allows more sophisticated tests comparing experiment and theory than do conventional equilibrium concepts, e.g., tests of theoretical predictions concerning the variances of various attributes of the players' behavior.

In practice sometimes one must produce a single joint strategy as one's "prediction" or "estimate" of the joint strategy. To do that with PGT, say that we have a loss function $L(q', q)$ that quantifies the penalty we will incur if we predict the joint mixed strategy q' and the true joint mixed strategy is q. Then decision theory counsels us to set our single prediction to the "Bayes-optimal" joint mixed strategy, i.e., to the q' that minimizes expected $L(q', q)$ under the posterior density over q. By mapping a game to a single predicted joint strategy this way, decision theoretic PGT provides an "equilibrium concept". Unlike typical equilibrium concepts (Fudenberg and Tirole (1991); Aumann and Hart (1992); Basar and Olsder (1999); Binmore (1992); Luce and Raiffa (1985)), this one does not require refinements; typically the Bayes-optimal prediction is unique.

Note that the Bayes-optimal equilibrium concept depends on the loss function of the external statistician. It is *not* specified within the game, for example as a utility function. So in particular, it varies with the person making the prediction.

Under special circumstances, the equilibria of conventional game theory arise as a special case of Bayes optimality. Let T be a set of (perhaps refined) NE of the game. Say that we have some reason to assign equal probability mass to every $q \in T$ and zero (or infinitesimal) probability density to all other q. Say furthermore that we have some reason to use the L_1 loss function. Then there are multiple Bayes optimal predictions — the elements of T — and we have no basis for choosing among them.

Of course, if the external statistician making the prediction were to change, then so would the loss function, and in general the elements of T would no longer be the Bayes optimal predictions. In addition, strictly speaking, the equilibrium concepts of game theory (and so in particular whatever the one is that defines T) do not assign relative probabilities to the elements of T, a shortcoming rectified in PGT. More generally, it is hard to construct physical scenarios involving real human beings that have such a posterior distribution that "justifies" some equilibrium concept: infinite probability density at some q's (the ones in T), with *exactly* equal probability mass at all of those q's, and infinitesimal density at joint distributions that are arbitrarily close to those q's.

Recall from Sec. 3.2 that an alternative "equilibrium concept" to the Bayes-optimal equilibrium concept is given by marginalizing the posterior over densities, $P(q \mid \mathscr{I})$, down to a (single) posterior over joint moves, $P(x \mid \mathscr{I})$. Typically the moves of the players are stochastically coupled under that posterior over joint moves, even for priors and likelihoods different from the ones considered in Sec. 3.2. Moreover, often the marginal distribution over a particular player's moves, x_i, is not utility-maximizing against the marginal distribution over the other players' moves. This can occur even if the support of the posterior over joint mixed strategies is restricted to Nash Equilibria (NE). In this sense, NE may be impossible, and bounded rationality is unavoidable.

The decision-theoretic equilibrium concept (i.e., Bayes optimal q) varies with the loss function, unlike the $P(x \mid \mathscr{I})$ equilibrium concept. However under both equilibrium concepts X_i and X_j may be statistically dependent. This is true even though the support of $P(q \mid \mathscr{I})$ is restricted to distributions where X_i and X_j are independent (a linear combination of product distributions typically is not a product distribution). In addition, say that $P(q \mid \mathscr{I})$ is restricted to NE q. Typically, if there are multiple such equilibria, then $P(x_i \mid \mathscr{I})$ is not an optimal response to $P(x_{-i} \mid \mathscr{I})$. Even if we know that all the players are perfectly rational, *our prediction* of their moves has "cross-talk" among the multiple equilibria, which prevents perfect rationality. This is one sense in which PGT has built-in bounded rationality. Note that this phenomenon has nothing to do with the use of an entropic prior, a Boltzmann-based likelihood, or the like, as is illustrated in the following example.

Example 2 *Consider a two player game in which both players have two possible moves, L and R. Indicate any (product distribution) q by two numbers, $q_1(x_1 = L)$ and $q_2(x_2 = L)$. Suppose that we happen to have a likelihood and/or prior such that*

$$P(q \mid \mathscr{I}) \quad = \quad \frac{\delta(q - (3/4, 3/4)) + \delta(q - (1/4, 1/4))}{2} \tag{27}$$

where "$\delta(.)$" is the Dirac delta function. Suppose also that we have quadratic loss. For that loss function, as is easy to verify, the Bayes-optimal q is the average q, $\int dq\; q(x)P(q \mid \mathscr{I})$. Viewed as a function of x, that particular Bayes-optimal q is the same as $P(x \mid \mathscr{I})$. Here that Bayes-optimal q is the distribution $P(L, L) = P(R, R) = 5/16, P(R, L) = P(L, R) = 3/16$. Indicate that distribution as p. p is not a product distribution, so $P(q = p \mid \mathscr{I}) = 0$. In other words, $q = P(x \mid \mathscr{I}) = p$, this game's "equilibrium", is a joint mixed strategy that cannot arise.

To help distinguish when one should use one or the other of our equilibrium concepts, consider a frequentist scenario, where we first give our prediction $q' \in \Delta_{\mathcal{X}}$ for the outcome of a game, and after that $P(q \mid \mathscr{I})$ is IID sampled an infinite number of times. If our reward for making prediction p is the average value of $L(q', q)$ over that infinite number of samples, then to maximize our reward we should use the Bayes-optimality equilibrium concept.

Say that instead, each time $P(q \mid \mathscr{I})$ is sampled to produce a q, that that q is itself sampled, to produce an x. This means that the IID samples of $P(q \mid \mathscr{I})$ provide an empirical distribution of the frequency with which each x occurs. With probability 1.0, the uniform metric distance between this empirical distribution and $P(x \mid \mathscr{I}) = \int dq\; P(q \mid \mathscr{I})q(x)$ is zero. But that integral is just the second equilibrium concept discussed above. So if the reward is how accurately we guess the empirical distribution over x's, then we should use this second equilibrium concept instead of the Bayes-optimality equilibrium concept.

Example 3 *Say we want to compare two theories that both make predictions for the equilibrium outcome of a pair of goal-seeking agents engaged in a game, and that can make their predictions with no knowledge concerning the two algorithms except that at the equilibrium of the game they are both perfectly rational. To do this we first collect many Reinforcing Learning (RL) algorithms from the literature. Next we repeatedly and randomly choose pairs from that set of RL algorithms. Then we have each such pair play the game in Ex. 1 many times, producing a sequence of game outcomes. Finally, we remove all such sequences of repeated play that don't converge to a NE.*

We can now randomly this set of equilibrium joint mixed strategies and then randomly sample the resultant mixed strategy to get a joint move x. We wish to compare our theories by asking what each one predicts for that distribution, $P(x)$. Standard noncooperative game theory would only say that each sequence converges to one of the three q's given by the NE product distributions. To use noncooperative game theory to make a prediction about $P(x)$ we would need to reduce that set of three distributions to a single one. To do that we would argue that one of the many NE refinements that have been studied Myerson (1991); Fudenberg and Tirole (1991) applies to our scenario, a refinement that it so happens selects a single one of those three NE. Noncooperative game theory would then predict that $P(x)$ is the product distribution of that selected NE.

In contrast, PGT would predict that $P(x)$ is not even a product distribution. Rather it is a weighted average of three product distributions. The weighting factors are biased in favor of joint distributions q that have high entropy. Note that these are the kind of q's that one might expect to arise from RL algorithms that engage in exploration as well as exploitation.

PGT would also provide other predictions that conventional noncooperate game theory cannot. For example, $P(x)$ equals the average q, $\int dq\ P(q)q(x)$. We could use PGT to predict the covariance matrix of q's,

$$\left[\int dqdq'\ P(q)P(q')q^T q'\right] - [P(x)]^2\,\mathbf{1} \tag{28}$$

(In this equation, $\mathbf{1}$ indicates the identity matrix, and the superscript "T" indicates transpose). We could then compare that prediction to the empirical covariance matrix given by the set of sequences. This cannot be done using conventional game theory.

Finally, suppose that our information $\mathscr{I}$ concerning a game does not *explicitly* tell us that the players in the game are all fully rational. Then the rationalities of the (human) players are random variables, and we must average over them to get the posterior over joint mixed strategies. This generically means that $P(q \mid \mathscr{I})$ is non-zero for joint mixed strategies q that are not perfectly rational. This is another way that PGT provides built-in bounded rationality.

4.2 PGT as a meta-game

The first type of equilibrium concept in PGT can also be motivated as a "meta-game" played against Nature. To formalize this meta-game, say we have a set of possible games G, differing in their utility functions, their players, etc. For each such game $\gamma \in G$, let $\Delta_{\mathcal{X}}(\gamma)$ indicate the a finite subset of all possible joint mixed strategies in γ, with $x \in X(\gamma)$ the possible joint moves in that game.[13]. Now consider a two-stage "meta-game" Γ that consists of a statistician ($\mathscr{S}$) playing against Nature (N). In this meta-game N's set of possible moves is $\{(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))\}$, i.e., the set of all possible games γ, and for each such game, the set of all possible joint mixed strategies q over the joint moves in γ. The mixed strategy of player N is a distribution over this space, $P(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))$. At the end of the first stage of the meta-game, N's mixed strategy is fairly sampled, producing an outcome (γ', q'). $\mathscr{S}$ knows that mixed strategy of N, $P(\gamma \in G, q \in \Delta_{\mathcal{X}}(\gamma))$, *ex ante*.

[13] We usually have in mind that $\Delta_{\mathcal{X}}(\gamma)$;is the set of all mixed strategies in γ, measured to some precision $\epsilon > 0$.

As an example, much work in conventional game theory presumes that a unique equilibrium of any underlying game γ' is highly desirable. So if there are multiple NE of the game, one has the problem of how to "refine" that set of multiple NE into a singleton. This problem has led to a huge body of work exploring different refinements like trembling hand perfection, weak dominance, etc. (Myerson (1991); Fudenberg and Tirole (1991); Aumann and Hart (1992)). In contrast, consider a PGT approach, where it is known that the players are perfectly rational. So the support of $P(q' \mid \gamma')$ is the set of multiple NE of γ'. In PGT there is no *a priori* need to decide among those multiple NE. Instead one allows all of them to occur, with relative probabilities given by $P(q' \mid \gamma')$.

However in the real world sometimes a statistician must produce a single prediction rather than a full posterior density over predictions. This provides a second stage for our meta-game. In this second stage $\mathscr{S}$ is told γ'. Together with the (known) mixed strategy of N, this gives $\mathscr{S}$ a posterior over what q' is, $P(q' \mid \gamma')$. In the second stage, $\mathscr{S}$ makes a move in $\Delta_{\mathcal{X}}(\gamma')$, i.e., picks a joint mixed strategy for the game γ'. We interpret that move of the statistician $\mathscr{S}$ as a prediction of what $q' \in \Delta_{\mathcal{X}}(\gamma')$ was produced by the sampling of player N.

As usual in games against Nature, N has no utility function. However $\mathscr{S}$ may have a utility function, given by the negative of a loss function $L(q, q')$ that quantifies how accurate her move q is as a prediction of N's move q'. In this case, to maximize her expected utility the statistician chooses her move — her prediction of the joint mixed strategy that governs the game γ' — to minimize her expected loss under the posterior $P(q' \mid \gamma')$. Formally, she should guess a distribution whose support is $\text{argmin}_p \mathbb{E}_{P(q'|\gamma')} L(p, q')$. This mapping of a game γ' to a single predicted joint strategy comprises the first equilibrium concept described above. (See Grunwald and Dawid (2004) for some work related to this meta-game.)

5. Other applications of PGT

Under the QRE-based likelihood introduced above, if q_{-i} changes, then $U^i_{q_{-i}}$ changes, and therefore q_i may have to change. So this likelihood implicitly presumes the players have had some form of interaction to couple them (just as do conventional equilibrium concepts when they have multiple solutions). In Wolpert (2005) a different likelihood is introduced that involves no such coupling. This likelihood can be viewed as a novel formulation of common knowledge (Aumann (1999); Aumann and Brandenburger (1995); Fudenberg and Tirole (1991)).

Interestingly, as it arises with this likelihood, bounded rationality is identical to an information-theoretic cost of computation. In this sense, under this likelihood cost of computation is *derived* as a cause of bounded rationality. It is not simply imputed, as an explanation of experimentally observed bounded rationality.

While various models of bounded rationality have been found to have some experimental validity (e.g., QRE's), no model with a small number of parameters will ever hold *exactly*. This means that to analyze the rationality of human behavior in experimental settings we need a way to quantify the rationality of *any* mixed strategy in *any* environment. As elaborated in Wolpert (2005), PGT provides such a rationality measure, one that can be derived from first-principles involving the Kullbach-Leibler distance.

PGT is applicable to many domains beyond those considered in this paper. In particular, in work in progress, PGT has been used to derive power law distributions over the possible outcomes in unstructured bargaining. Those distributions have the Nash bargaining solution as their mode.

Extending further, PGT should provide an extension of mechanism design to allow for bounded rational players. Such an extension would also allow design of the mechanism to depend on variances (and higher moments) of the mechanism's utility function. In this way risk aversion of the mechanism designer can be accommodated in the design of the mechanism.

More speculatively still, PGT might be applicable to cooperative game theory. Just as PGT obviates the issue of how to define a noncooperative game's "equilibrium", it might do the same thing for cooperative games. This would remove one of the major stumbling blocks to progress in cooperative game theory, which is notorious for having many competing equilibrium concepts.

Acknowledgments

I would like to thank Julian Jamison, George Judge, Duncan Luce, Herb Gintis, and Larry Karp for helpful discussions.

References

C.D. Aliprantis and K. C. Border. *Infinite Dimensional Analysis*. Springer Verlag, 2006.

M. Allais. *Econometrica*, 21:503–546, 1953.

S. P. Anderson, J.K. Goeree, and C. A. Holt. The logit equilibrium: A perspective on intuitive behavioral anomalies. *Southern Economic Journal*, 69(1):21–47, 2002.

M. Aoki. *Modeling Aggregate Behavior and Fluctuations in Economics : Stochastic Views of Interacting Agents*. Cambridge University Press, 2004.

R. J. Aumann. Interactive epistemology ii: Probability. *Int. J. Game Theory*, 28:301–314, 1999.

R. J. Aumann and A. Brandenburger. Epistemic conditions for nash equilibrium. *Econometrica*, 63(5):1161–1180, 1995.

R.J. Aumann and S. Hart. *Handbook of Game Theory with Economic Applications*. North-Holland Press, 1992.

T. Basar and G.J. Olsder. *Dynamic Noncooperative Game Theory*. Siam, Philadelphia, PA, 1999. Second Edition.

J. M. Berger. *Statistical Decision theory and Bayesian Analysis*. Springer-Verlag, 1985.

J. Bernardo and A. Smith. *Bayesian Theory*. Wiley and Sons, 2000.

S. Bieniawski, I. Kroo, and D. H. Wolpert. Flight Control with Distributed Effectors. In *Proceedings of 2005 AIAA Guidance, Navigation, and Control Conference, San Francisco, CA*, 2005. AIAA Paper 2005-6074.

K. Binmore. *Fun and Games: A Text on Game Theory.* D. C. Heath and Company, Lexington, MA, 1992.

R. Brafman and M. Tennenholtz. Learning to coordinate efficiently: A model-based approach. *Journal of Artificial Intelligence Research*, 19, 2003.

W. A. Brock and S. N. Durlauf. Interaction-based models. In *Handbook of Econometrics: Volume 5*, pages 3297–3380. Elsevier, 2001. Chapter 54.

C.F. Camerer. *Behavioral Game theory: experiments in strategic interaction.* Princeton University Press, 2003.

H. C. Chen and J. W. Friedman. *Games and Economic Behavior*, 18:32–54, 1997.

S. Choi and J.J. Alonso. Multi-fidelity design optimization of low-boom supersonic business jet. In *Proceedings of 10th AIAA/ISSMO Multidisciplany Analysis and Optimization Conference*, 2004. AIAA Paper 2004-4371.

T. Cover and J. Thomas. *Elements of Information Theory.* Wiley-Interscience, New York, 1991.

E.J. Cramer, J.E. Dennis, and et alia. Problem formulation for multidisciplinary optimization. *SIAM J. of Optimization*, 4, 1994.

A. Dragulescu and V.M. Yakovenko. Statistical mechanics of money. *Eur. Phys. J. B*, 17: 723–729, 2000.

S. Durlauf. How can statistical mechanics contribute to social science? *Proc. Natl. Acad. Sci. USA*, 96:10582–10584, 1999.

I. Erev and A. Roth. Multi-agent learning and the descriptive value of simple models. *Artificial Intelligence*, pages 423–428, 2007.

J.D. Farmer, M. Shubik, and D. E. Smith. Economics: The next physical science? SFI working paper 05-06-027.

J. Ferber. Reactive distributed artificial intelligence: Principles and applications. In G. O-Hare and N. Jennings, editors, *Foundations of Distributed Artificial Intelligence*, pages 287–314. John Wiley and Sons, 1996.

D. Fudenberg and D. Kreps. Learning mixed equilibria. *Game and Economic Behavior*, 5: 320–367, 1993.

D. Fudenberg and D. K. Levine. Steady state learning and Nash equilibrium. *Econometrica*, 61(3):547–573, 1993.

D. Fudenberg and D. K. Levine. *The Theory of Learning in Games*. MIT Press, Cambridge, MA, 1998.

D. Fudenberg and J. Tirole. *Game Theory*. MIT Press, Cambridge, MA, 1991.

M. Georgeff, B. Pell, M. Pollack, M. Tambe, and M Wooldridge. In *Intelligent Agents V, Springer-Verlag Lecture Notes in AI*. 1999. Volume 1365, March 1999.

J. K. Goeree and C. A. Holt. Stochastic game theory: for playing games, not just doing theory. *Proceedings National Academy of Sciences*, 96:10564–10567, 1999.

A. Greenwald and M. Littman. Introduction to the special issue on learning and computational game theory. *Machine Learning Journal*, 67, 2007.

A. Greenwald, N. Jennings, and P. Stone. Guest editors' introduction: Agents and markets. *IEEE Intelligent Systems*, 18, 2003.

P. D. Grunwald and A. P. Dawid. Game theory, maximum entorpy, minimum discrpenacy and robust bayesian decision theory. *Annals of Statistics*, 2004.

S. F. Gull. Bayesian inductive inference and maximum entropy. In *Maximum Entropy and Bayesian Methods*, pages 53–74. Kluwer Academic Publishers, 1988.

S. Hart. Adaptive heuristics. *Econometrica*, 73:14011430, 2005.

D. Heckerman. A tutorial on learning with bayesian networks. In M. Jordan, editor, *Learning in Graphical Models*. MIT Press, 1999.

J.D. Hey and C. Orme. Investigating generalizations of expected utilty theory using experimental data. *Econometrica*, 62:1291–1326, 1994.

K. S. Van Horn. Constructing a logic of plausible inference: a guide to cox's theorem. *International Journal of Approximate Reasoning*, 34(1):3–24, 2003.

E. Horvitz. From the editor: special issue on graph-based representations. *Decision Analysis*, pages 125–126, 2005.

J. Hu and M. P. Wellman. Nash q-learning for general-sum stochastic games. *Journal of Machine Learning Research*, pages 1039–1069, 2003.

H. Hwang, J. Kim, and C. Tomlin. Protocol-based conflict resolution for air traffic control. *Air Traffic Control Quarterly*, 2007. in press.

T. R. Palfrey J. K. Goeree, C. A. Holt. Quantal response equilibrium and overbidding in private-value auctions. 1999.

E. T. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106:620, 1957.

E. T. Jaynes and G. Larry Bretthorst. *Probability Theory : The Logic of Science*. Cambridge University Press, 2003.

N. R. Jennings, K. Sycara, and M. Wooldridge. A roadmap of agent research and development. *Autonomous Agents and Multi-Agent Systems*, 1:7–38, 1998.

D. Kahneman. A psychological perspective on economics. *American Economic Review (Proceedings)*, 93:2:162–168, 2003a.

D. Kahneman. Maps of bounded rationality: Psychology of behavioral economics. *American Economic Review*, 93:1449–1475, 2003b.

S. Kalyanakrishnan, Y. Liu, and P. Stone. Offense in robocup soccer: A multiagent reinforcement learning case study. In *RoboCup-2006: Robot Soccer World Cup X*, pages 72–85, 2007.

R. Kurzban and D. Houser. Experiments investigating cooperative types in humans. *Proceedings of the National Academy of Sciences*, 102(5):1803–1807, 2005.

J. A. List and M. S. Haigh. A simple test of expected utility theory using professional traders. *Proceedings of the National Academy of Sciences*, 102:945–948, 2005.

G. Loomes, P. Moffat, and R. Sugden. A microeconomic test of alternative stochastic theories of risk choice. ERC discussion paper 9806, 1998.

T.J. Loredo. From laplace to sn 1987a: Bayesian inference in astrophysics. In *Maximum Entropy and Bayesian Methods*, pages 81–142. Kluwer Academic Publishers, 1990.

D. Luce. *Individual Choice Behavior*. Wesley, 1959.

R. D. Luce and H. Raiffa. *Games and Decisions*. Dover Press, 1985.

D. Mackay. *Information theory, inference, and learning algorithms*. Cambridge University Press, 2003.

S. Mannor and J. Shamma. Multi-agent learning for engineers. *Artificial Intelligence*, pages 417–422, 2007.

R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for normal form games. *Games and Economic Behavior*, 10:6–38, 1995.

R. D. McKelvey and T. R. Palfrey. Quantal response equilibria for extensive form games. *Experimental Economics*, 1:9–41, 1998.

J. R. Meginniss. A new class of symmetric utility rules for gambles, subjective marginal probability functions, and a generalized Bayes' rule. *Proc. of the American Statisticical Association, Business and Economics Statistics Section*, pages 471 – 476, 1976.

Roger B. Myerson. *Game theory: Analysis of Conflict*. Harvard University Press, 1991.

N. Nisan and A. Ronen. Algorithmic mechanism design. *Games and Economic Behavior*, 35:166–196, 2001.

J. B. Paris. *The Uncertain Reasoner's Companion: A Mathematical Perspective*. Cambridge University Press, 1994.

C. P. Robert and G. Casella. *Monte Carlo Statistical Methods*. Springer-Verlag, New York, 2004.

Ariel Rubinstein. *Modeling Bounded Rationality*. MIT press, 1998.

S. Russell and D. Subramanian. Provably bounded-optimal agents. *Journal of AI Research*, page 575, 1995.

A. Schaerf, Y. Shoham, and M. Tennenholtz. Adaptive load balancing: A study in multi-agent learning. *Journal of Artificial Intelligence Research*, 162:475–500, 1995.

J.S. Shamma and G. Arslan. Dynamic fictitious play, dynamic gradient play, and distributed convergence to nash equilibria. *IEEE Trans. on Automatic Control*, 50(3):312–327, 2004.

Y. Shoham, R. Powers, and T. Grenager. If multi-agent learning is the answer, what is the question? *Artificial Intelligence*, pages 365–377, 2007.

C. Starmer. Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature*, 38:332–382, 2000.

P. Stone. Multiagent learning is not the answer. it is the question. *Artificial Intelligence*, pages 402–405, 2007.

C.E Strauss, D.H. D.H. Wolpert, and D.R. Wolf. Alpha, evidence, and the entropic prior. In A. Mohammed-Djafari, editor, *Maximum Entropy and Bayesian Methods 1992*. Kluwer, 1994.

L. Tesfatsion and K.L. Judd. *Handbook of Computational Economics: Volume 2, Agent-Based Computational Economics*. Elsevier, North-Holland Imprint, 2006. Proceedings, in Handbooks in Economics Series.

Flemming Topsoe. Information-theoretical optimization techniques. *Kybernetika*, pages 8–27, 1979.

K. E. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, 2003.

A. Tversky and D. Kahneman. Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297–323, 1992.

J. von Neuman and O. Morgenstern. *Theory of Games and Economics Behavior*. Princeton university Press, 1944.

D. H. Wolpert. Information theory - the bridge connecting bounded rational game theory and statistical physics. In D. Braha, A. Minai, and Y. Bar-Yam, editors, *Complex Engineered Systems: Science meets technology*, pages 262–290. Springer, 2004a.

D. H. Wolpert. What Information theory says about best response, binding contracts, and Collective Intelligence. In A. Namatame, editor, *Proceedings of WEHIA04*. Springer Verlag, 2004b.

D. H. Wolpert. A predictive theory of games. http://lanl.arxiv.org/abs/nlin/0512015, 2005.

D. H. Wolpert, C. E. M. Strauss, and D. Rajnayaran. Advances in distributed optimization using probability collectives. *Advances in Complex Systems*, 2006. in press.

D.H. Wolpert and N. Kulkarni. Using game theory to manage multiple interacting systems. submitted, 2008.

D.H. Wolpert and D. Rajnarayan. Parametric learning and monte carlo optimization. http://lanl.arxiv.org/abs/0704.1274, 2007.

D.H. Wolpert, J. Jamison, and W. Cho. Experimental measurement of the fundamental variables of game theory. In preparation, 2008.

A. Zellner. Some aspects of the history of bayesian information processing. *Journal of Econometrics*, 2004.