**CIS 700 Differential Privacy in Game Theory and Mechanism Design** April 18, 2014

Lecture 11

*Lecturer: Aaron Roth* *Scribe: Aaron Roth*

# Using Differential Privacy to Analyze the Equilibria of Repeated Games in Settings of Imperfect Monitoring

## 1 Introduction

So far in this class we have mostly been studying one-shot games, in which players simultaneously decide on an action to take and then experience the resulting payoff. When studying ascending price auctions, we did have the opportunity to think about an extensive form game, in which players interact with one another in rounds – but even then, we were discussing a game that had a finite time horizon.

In this lecture, we will use the tools of differential privacy to think about the equilibria of *infinitely repeated games*, in which players repeatedly play the same game against one another, and observe some signal about what their opponents have done each day.

One common feature of repeated games is that they can have a much richer set of equilibria than single-shot games. Informally, this is because players can threaten to punish each other for deviating from proscribed behavior, since they receive signals about one-another's actions over the course of play. This can be both good news and bad news. On the one hand, equilibria of the repeated game can have higher payoff than any equilibrium of the corresponding stage game; however, they can also have *lower* payoff than any equilibrium of the stage game. In general, in repeated games, there are a huge multiplicity of equilibria, which is *good news* for a mechanism designer who wants to implement a high welfare equilibrium, but *bad news* for a game theorist who wants equilibria to be meaningful predictions of behavior.

In this class we will see that the complicated structure of the equilibria of a repeated game collapses to the simpler structure of the equilibria of the stage game whenever the signals that agents see about one another are differentially private – and we give examples showing how differentially private signals arise naturally in repeated large games.

## 2 Repeated Games

A repeated game is played over an infinite series of discrete time steps, which we index by $t \in \{0, 1, 2, \ldots\}$. Every day, players choose an action in a fixed *stage game* $\mathcal{G}$. $\mathcal{G}$ is defined by an action set $A$, together with a utility function for each player $i$: $u_i : A^n \to [0, 1]$. As usual, on any given day players can also play mixed strategies which are distributions over actions, in which case they receive the expected utility for the action they play.

These are just the usual games we always talk about, and so we are already familiar with *stage game equilibria.*

**Definition 1** *An $\eta$-approximate Nash equilibrium of the stage game $\mathcal{G}$ is a vector of mixed strategies $\alpha \in (\Delta A)^n$ such that for all players $i$ and for all actions $a'_i \in A$:*

$$u_i(\alpha) \geq u_i(a'_i, \alpha_{-i}) - \eta$$

Equilibria of the repeated game, can however, be more complex. This is because *strategies* are more complex: they are no longer simply distributions over actions, but rather functions from histories of observations to distributions over actions. To define them, we must first define the monitoring structure of a game.

In these notes, we describe only games of *public monitoring*, but analogous results hold also for games of private monitoring. At the end of every day $t$, after players play some set of actions $a \in A^n$, they all commonly observe a *signal* $s \in S$, which is drawn according to some probability distribution parameterized by $a$: $s \sim P_a \in \Delta S$. In a game of *perfect monitoring* we would have $S = A^n$ players would deterministically observe exactly the actions played every day, but more generally, they might receive only some noisy signal which is a function of the plays made that day.

Players have to reason about how to play the game without necessarily observing the actions of their opponents, which means that they do not have direct access to their own payoff. In order to make it tractable for them to reason about how to play the game, we make the assumption that the signal structure is enough for players to reason about their expected payoff. Formally, we assume:

**Assumption 2** *For each player $i$ exists a function $U_i : A \times S \to [0,1]$ so that for all $a \in A^n$:*

$$u_i(a_i, a_{-i}) = \sum_{s \in S} P_a(s) \cdot U_i(a_i, s)$$

*In other words, for every action profile $a$, $\mathrm{E}_{s \sim P_a}[U_i(a_i, s)] = u_i(a)$.*

Since risk neutral players can equivalently play so as to maximize $U$ rather than $u$, this allows players to understand their utility given only their own action together with the observed signal.

Play proceeds in a series of rounds, in which players play vectors of actions $a^0, a^1, \ldots, a^t, \ldots$. Following each round $t$, a signal is generated, resulting in a series of signals $s^1, s^2, \ldots, s^{t+1}, \ldots$. Public *histories* are records of the signals that have been observed in any finite prefix of the game: the history at time $t$ is $h^t = (s^1, \ldots, s^{t-1})$, and the set of all possible public histories is:

$$\mathcal{H} = \bigcup_{t=1}^{\infty} S^t$$

The private history of player $i$ is the record of public signals $h^t$, together with the actions $(a^0, \ldots, a^{t-1})$ that player $i$ has played so far. A *strategy* for player $i$ is a mapping from the set of histories to distributions over actions (i.e. a rule telling him how to play given any sequence of observations he might have made in the game). In general, strategies can depend on players private histories, but for simplicity in this lecture, we will focus on public strategies:

**Definition 3** *A public strategy for player $i$ is a function $\sigma_i : \mathcal{H} \to \Delta A$.*

Given a strategy $\sigma_i$ for every player, play proceeds as follows: At stage $t$, every player $i$ draws an action $a_i$ from $\sigma_i(h^t)$. Then, a signal $s^t$ is drawn from $P_a$, which induces a new history $h^{t+1} = (h^t, s^t)$. Play continues recursively.

Since play is infinitely repeated, players evaluate their payoff with respect to a daily *discount factor* $\delta \in [0,1)$: in other words, a payoff of 1 tomorrow is worth only $\delta$, and a payoff of 1 in 2 days is worth only $\delta^2$. Since this allows the total sum payoff over time to be as large as $\sum_{t=0}^{\infty} 1 \cdot \delta^t = \frac{1}{1-\delta}$, it is standard to *normalize* payoffs to again lie in $[0,1]$ by scaling the sum total payoff by $(1-\delta)$. Formally, we can recursively define the expected payoff of a player $i$ given strategies $\sigma$ at a history $h^t$ as:

$$\begin{aligned} V_{i,\sigma}(h^t) &= (1-\delta)\left(u_i(\sigma(h^t)) + \delta \sum_{s \in S} P_{\sigma(h^t)}(s) \cdot V_{i,\sigma}((h^t, s))\right) \\ &= (1-\delta)u_i(\sigma(h^t)) + \delta \sum_{s \in S} P_{\sigma(h^t)}(s) \cdot V_{i,\sigma}((h^t, s)) \end{aligned}$$

We can now define the simplest solution concept we might want to think about in repeated games:

**Definition 4** *A set of public strategies $\sigma_1, \ldots, \sigma_n$ form a* public Nash equilibrium *if for every player $i$ and strategy $\sigma'_i$:*

$$V_{i,\sigma}(\emptyset) \geq V_{i,(\sigma'_i, \sigma_{-i})}(\emptyset)$$

We could also consider equilibria that do not consist only of public strategies, and insist that our equilibria be sub-game perfect (which would require that the above equilibrium condition hold for all possible histories $h^t$, not only the empty history). Everything we do in this lecture can be done for these more complicated equilibrium sets, but the main ideas are all captured in our discussion of public Nash equilibria.

## 3 The Folk Theorem

We here sketch the simplest version of a large collection of theorems known as "Folk Theorems", that characterize a multiplicity of equilibria in repeated games. In the simplest variant, we will consider only games of perfect monitoring (i.e. $S = A^n$, and when players play $a \in A^n$ they deterministically view the signal $s = a$), and will not require that equilibria be subgame perfect – but variants of the folk theorem hold also in settings of imperfect monitoring, for subgame perfect equilibrium, and for private monitoring settings – see e.g. [FLM94] and [Sug11].

We start by defining the min-max payoff for each player in the stage game $\mathcal{G}$:

**Definition 5** *For each player $i$, the min-max payoff $\theta_i$ for player $i$ is defined to be:*

$$\theta_i = \min_{\alpha_{-i} \in (\Delta A)^{n-1}} \max_{a_i \in A} u_i(a_i, \alpha_{-i})$$

*In other words, it is the best payoff that player $i$ can guarantee for himself when all other players are colluding to try and minimize his payoff.*

*For a player $i$, let $\alpha_{j \to i}$ be the mixed strategy that player $j$ plays that achieves this min. (i.e. let $\alpha_{j \to i}$ be the punishment strategy that player $j$ would use against player $i$).*

**Definition 6** *Say that an action profile $a \in A^n$ is individually rational if for every player $i$, $u_i(a) > \theta_i$.*

**Theorem 7 (The (simplest) Folk Theorem)** *For any individually rational action profile $a$, there is a sufficiently large discount factor $\delta$ such that there exists a Nash equilibrium $\sigma$ in which every player deterministically plays $a$ every round.*

**Proof** [Proof Sketch] Each player $j$ plays according to the following strategy $\sigma_j$: if $a^1 = \ldots = a^{t-1} = a$, then $\sigma_j(a^1, \ldots, a^{t-1}) = a$. Otherwise, let $i$ be the index of the first player who deviated from playing $a$. In this case, $\sigma_j(a^1, \ldots, a^{t-1}) = \alpha_{j \to i}$.

Under strategy profile $\sigma$, no player ever deviates from playing $a$, so it suffices to show that $\sigma$ is a Nash equilibrium. By assumption, $a$ is individually rational, so $\Delta_i \overset{\text{def}}{=} u_i(a) - \theta_i > 0$. Suppose player $i$ decides to deviate and play $a'_i$ on some round $t$: he increases his utility on day $t$ by $u_i(a'_i, a_{-i}) - u_i$, however he triggers punishment from the other players, and so looses (discounted) future utility equal to at least $\Delta_i \cdot \frac{\delta}{1-\delta}$ Since $\Delta_i > 0$, for $\delta$ sufficiently close to 1, we have:

$$\Delta_i \cdot \frac{\delta}{1-\delta} \geq u_i(a'_i, a_{-i}) - u_i$$

and so the deviation is not beneficial. ■

There is a large literature in economics proving versions of this "folk theorem" under different settings, including many settings in which only a noisy signal is observed. However, these theorems tend to have the following form: Fixing a signal distribution, there exists a large enough $\delta$ such that there exists a huge multiplicity of equilibria in the repeated game. However, the structure of equilibria in repeated games for fixed values of $\delta$ bounded away from 1 is much less well understood. In particular, since $\delta$ might be viewed as an individual (rather than population level) parameter, it might make sense to study values of $\delta$ that may be arbitrarily close to 1, but remain fixed as the population size of the game $n \to \infty$. This is what we will do in the remainder of the lecture.

# 4 Differentially Private Signal Distributions

We can ask what the structure of the equilibria of the repeated game is when the distribution from which *signals* are drawn satisfies differential privacy.

**Definition 8** *A public signaling structure satisfies $(\epsilon, \gamma)$-differential privacy if for all $a \in A^n$, for all $i$, for all $a_i' \in A$, and for all $E \subseteq S$:*

$$P_a(E) \leq \exp(\epsilon) P_{(a_i', a_{-i})}(E) + \gamma$$

We now argue that if the signal structure satisfies differential privacy, then the only equilibria of the repeated game play (approximate) equilibria of the stage game every day.

The intuition is essentially encapsulated as follows: in general, the only thing that prevents an agent at some stage $t$ from deviating from his equilibrium strategy, and playing a (stage-game) best response to the distribution over his opponent's actions is fear of punishment: if his opponents can detect this deviation, then they can change their behavior to lower his expected *future* payoff. Differential privacy provides us a simple, worst-case way of quantifying the decrease in expected future payoff that can result from a single player's one-stage unilateral deviation. If this decrease can be made small enough, then it cannot serve as an incentive to prevent any player from playing anything other than an (approximate) stage-game best response to his opponents. Thus, every day, all players must be playing an approximate equilibrium of the stage game. We formalize this intuition in the next theorem:

**Theorem 9** *Fix any repeated game with discount factor $\delta$, with public signals that satisfy $(\epsilon, \gamma)$-differential privacy. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ denote a public equilibrium. Then for every history $h^t$ that occurs with positive probability when players play according to $\sigma$, the distribution on actions at stage t $(\sigma_1(h^t), \ldots, \sigma_n(h^t))$ forms an $\eta$-approximate Nash equilibrium of the stage game, for*

$$\eta = \frac{\delta}{1-\delta}(\epsilon + \gamma)$$

**Proof** We will write $\Pr_\sigma[h^t]$ to denote the probability that a given public history $h^t$ arises when players use strategies $\sigma$. For each $j \leq t$, write $h^{t,\leq j}$ to denote the sub-history of $h^t$ consisting of the first $j$ periods, and $h^{t,j}$ to be the $j^{\text{th}}$ period signal. Then:

$$\Pr_\sigma[h^t] = \prod_{j=1}^{t} P_{\sigma(h^{t,\leq j})}(h^{t,j})$$

Now fix any $T$. We can write:

$$V_{i,\sigma}(\emptyset) = (1-\delta)\left(\sum_{h^T \in S^T} \Pr_\sigma[h^T]\left(\left(\sum_{t=0}^{T} \delta^t u_i(\sigma(h^{T,\leq t})) + \frac{\delta^{T+1}}{1-\delta}\sum_{s \in S} V_{i,\sigma}(h^T, s) P_{\sigma(h^t)}(s)\right)\right)\right)$$

Consider any history $h^T$ such that $\Pr_\sigma[h^T] > 0$, and consider the deviation of player $i$, $\sigma_i'$ that is identical to $\sigma_i$, except that on history $h^T$ player $i$ plays a stage-game best response to his opponents. $\sigma_i'(h^T) = \arg\max_{a \in A_i} u_i(a, \sigma_{-i}(h^T)) \equiv a_i^*$. Since $\sigma$ is a Nash equilibrium of the repeated game, we know that $V_{i,\sigma}(\emptyset) - V_{i,(\sigma'_i,\sigma_{-i})}(\emptyset) \geq 0$. Since $\Pr_\sigma[h^T] > 0$, we can divide and write this difference as:

$$\begin{aligned}
0 &\leq \frac{1}{(1-\delta)\Pr_\sigma[h^T]}\left(V_{i,\sigma}(\emptyset) - V_{i,(\sigma'_i,\sigma_{-i})}(\emptyset)\right) \\
&= \left(\delta^T\left(u_i(\sigma(h^T)) - u_i(a_i^*,\sigma_{-i}(h^T))\right) + \frac{\delta^{T+1}}{1-\delta}\sum_{s\in S} V_{i,\sigma}(h^T,s)\left(P_{\sigma(h^t)}(s) - P_{(a_i^*,\sigma(h^T))}(s)\right)\right) \\
&\leq \left(\delta^T\left(u_i(\sigma(h^T)) - u_i(a_i^*,\sigma_{-i}(h^T))\right) + \frac{\delta^{T+1}}{1-\delta}\sum_{s\in S}\left|P_{\sigma(h^t)}(s) - P_{(a_i^*,\sigma(h^T))}(s)\right|\right) \\
&\leq \left(\delta^T\left(u_i(\sigma(h^T)) - u_i(a_i^*,\sigma_{-i}(h^T))\right) + \frac{\delta^{T+1}}{1-\delta}(\epsilon+\gamma)\right)
\end{aligned}$$

where the last inequality follows from $(\epsilon,\gamma)$-differential privacy, and the fact that $\exp(-\epsilon) \geq 1-\epsilon$. Dividing through by $\delta^T$ and rearranging, we find:

$$\left(u_i(a_i^*,\sigma_{-i}(h^T)) - u_i(\sigma(h^T))\right) \leq \frac{\delta}{1-\delta}(\epsilon+\gamma)$$

which completes the proof. ■

Similar theorems hold for non public strategies as well, as well as for games with private signalling structures. We can also prove similar theorems restricted to subgame perfect equilibria, which imply that even action profiles outside of the path of play will be approximate Nash equilibria of the stage game.

## 5 How Private Signals Can Naturally Arise

We here give a couple of simple examples in which private signals can naturally arise in large games, where the parameters $(\epsilon+\gamma)$ tends to 0 as the game grows large. Note that in such cases, for a fixed discount value $\delta$, asymptotically, the equilibrium structure of the repeated game collapses to the equilibrium structure of the stage game.

In general, private signals can arise because generically, adding noise to low sensitivity functions yields differential privacy. We have already seen this for Laplace noise, but the same holds for other natural noise distributions: for example, for Gaussian noise. Say that the $\ell_2$ sensitivity of a function $f : T^n \to \mathbb{R}^k$ is:

$$\Delta_2(f) = \max_{t\in T^n, t'_i\in T} ||f(t) - f(t'_i,t_{-i})||_2$$

**Theorem 10 ([DKM$^+$06])** *Suppose a function $f : T^n \to \mathbb{R}^k$ has $\ell_2$ sensitivity $\Delta_2(f)$. Then the algorithm that computes $f(t)+Z$, where $Z \in \mathbb{R}^k$ is a random vector where each coordinate is drawn i.i.d. from the normal distribution $N(0,\sigma^2)$ is $(\epsilon,\gamma)$-differentially private, for*

$$\sigma = \frac{\Delta_2(f)}{\epsilon}\sqrt{\log\left(\frac{1.25}{\gamma}\right)}.$$

A simple corollary of this is:

**Corollary 11** *Let $\sigma$ be any constant, and let $f : T^n \to \mathbb{R}^k$ have $\ell_2$ sensitivity $O(1/n)$. Then the algorithm that computes $f(t)+Z$, where $Z \in \mathbb{R}^k$ is a random vector where each coordinate is drawn i.i.d. from the normal distribution $N(0,\sigma^2)$ is $(\epsilon,\gamma)$-differentially private, for*

$$(\epsilon+\gamma) = O\left(\frac{\sqrt{\log(n)}}{n}\right)$$

### 5.1 Large Anonymous Games

A natural condition in a large game is that it be *anonymous* – i.e. that players do not care precisely *who* is playing which actions, but only what fraction of the population is playing each action. (Equivalently, player $i$'s utility is invariant to permuting the action choices of his opponents). In other words, if $|A| = k$, player $i$'s utility can be written only in terms of his own action $a_i$, as well as a *histogram* $s \in [0,1]^k$ of the actions of the other players, defined as $s_\ell = \frac{1}{n}|\{j : a_j = \ell\}|$.

Suppose also that for each action $a_i$ that player $i$ might play, his utility can be written as a linear function of this histogram: i.e. $u_i(a_i, s) = \langle v^{a_i}, s\rangle$ for some $v^{a_i} \in [0,1]^k$.

A natural noisy signal for each player to observe is a perturbation of $s$: i.e. they see not *exactly* what fraction of players play each action, but the noisy estimate of these fractions: Suppose when players play an action profile $a \in A^n$ that corresponds to a histogram $s$, they observe a signal $\hat{s} + Z$, where $Z \in \mathbb{R}^k$ consists of i.i.d. entries drawn from $N(0, \sigma^2)$, where $\sigma$ can be some arbitrarily small constant. (Other noise distributions work work as well – for example, the histogram could be generated via subsampling). Since the noise is unbiased, we can take $U_i(a_i, \hat{s}) = \langle v^{a_i}, \hat{s}\rangle$, which is easily seen to satisfy Assumption 2, since for any $a \in A^n$,

$$\mathrm{E}[U_i(a_i, \hat{s})] = \langle v^{a_i}, s\rangle + \mathrm{E}[\langle v^{a_i}, Z\rangle] = \langle v^{a_i}, s\rangle = u_i(a_i, s)$$

Since the function that computes the histogram $s$ from $a$ is clearly $O(1/n)$ sensitive, we can apply corollary 11 together with our theorem to obtain:

**Theorem 12** *Fix any anonymous game with $n$ players and with discount factor $\delta$, and let the signaling structure consist of the histogram of actions played, perturbed with i.i.d. Gaussian noise of any constant variance. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ denote a public equilibrium. Then for every history $h^t$ that occurs with positive probability when players play according to $\sigma$, the distribution on actions at stage $t$ $(\sigma_1(h^t), \ldots, \sigma_n(h^t))$ forms an $\eta$-approximate Nash equilibrium of the stage game, for*

$$\eta = O\left(\left(\frac{\sqrt{\log n}}{n}\right) \cdot \left(\frac{\delta}{1-\delta}\right)\right)$$

### 5.2 Noisy Cournot Games

In an $n$ player Cournot competition, each firm $i$ simultaneously chooses quantity $q_i \in [0,1]$ of some good to produce, at a cost of $c(q_i)$. Given the productions of all firms, the market arrives at a (noisy) price:

$$p = \theta \cdot P\left(\frac{1}{n}\sum_{i=1}^{n} q_i\right)$$

where $Z$ is a random variable drawn from some noise distribution with mean 1. Player $i$'s utility is then $u_i(q_i, p) = p \cdot q_i - c(q_i)$. Once again, taking $U_i = u_i$ clearly satisfies Assumption 2. Let us suppose that $\theta$ is distributed according to a log-normal distribution with mean 1 and any (arbitrarily small) constant variance. To argue that the signal $p$ is differentially private, it will be sufficient to argue that $\log p$ is differentially private (because $p$ can be deterministically computed from $\log p$). We have:

$$\log p = \log\theta + \log\left(P\left(\frac{1}{n}\sum_{i=1}^{n} q_i\right)\right)$$

The $\ell_2$ sensitivity of $\log(P(\cdot))$ is at most:

$$\Delta_2(\log P) \leq \sup_{x\in[0,1]} \frac{1}{n}\frac{P'(x)}{P(x)}$$

For any fixed pricing function $P$ such that $P'(x)/P(x)$ is bounded, we therefore have $\Delta_2(\log p) = O(1/n)$. Since $\log(\theta)$ is normally distributed with constant variance, we can again apply corollary 11 together with our theorem to obtain:

**Theorem 13** *Fix any Cournot game with $n$ players and with discount factor $\delta$, with pricing function $P$ such that $P'(x)/P(x)$ is bounded. Let $\sigma = (\sigma_1, \ldots, \sigma_n)$ denote a public equilibrium. Then for every history $h^t$ that occurs with positive probability when players play according to $\sigma$, the distribution on actions at stage $t$ $(\sigma_1(h^t), \ldots, \sigma_n(h^t))$ forms an $\eta$-approximate Nash equilibrium of the stage game, for*

$$\eta = O\left(\left(\frac{\sqrt{\log n}}{n}\right) \cdot \left(\frac{\delta}{1-\delta}\right)\right)$$

## 5.3 Other games

We have given two examples here, but games with differentially private signaling structures are not hard to come by. For example, we could have proven a similar theorem about repeated games with private signaling (in which each player observes his own signal). A natural signaling structure here is one in which every player observes a noisy estimate of the payoff he would have received had he played each of his $k$ actions. Again, this will be differentially private if the payoffs are perturbed with any (arbitrarily small) constant variance, and the game is large.

More generally, differential privacy is a quantity that arises naturally in noisy settings, and is naturally resilient to composition and post-processing. So, for example, if the signal is any combination or function of combinations of differentially private values, the signaling structure remains differentially private. In large games, it will then often be the case that the privacy parameters tend to 0 as the size of the game gets large. In any such game, our theorems apply to quantify what the structure of the equilibria of the repeated game can look like for fixed values of $\delta$ bounded away from 1. For any such game, the equilibria of the repeated game must involve only repeated play of approximate equilibria of the stage game, where the approximation factor tends to 0 at a rate controlled by $\delta$ and the size of the game.

**Bibliographic Information** The results from this lecture are taken from Pai, Roth, and Ullman [PRU14], in which a variety of similar theorems are proven.

# References

[DKM+06] Cynthia Dwork, Krishnaram Kenthapadi, Frank McSherry, Ilya Mironov, and Moni Naor. Our data, ourselves: Privacy via distributed noise generation. In *EUROCRYPT*, pages 486–503, 2006.

[FLM94] Drew Fudenberg, David Levine, and Eric Maskin. The folk theorem with imperfect public information. *Econometrica: Journal of the Econometric Society*, pages 997–1039, 1994.

[PRU14] Mallesh M Pai, Aaron Roth, and Jonathan Ullman. An anti-folk theorem for large repeated games with imperfect monitoring. *arXiv preprint arXiv:1402.2801*, 2014.

[Sug11] Takuo Sugaya. Folk theorem in repeated games with private monitoring. *Economic Theory Center Working Paper*, (011-2011), 2011.