Approximately Optimal Mechanism Design via Differential Privacy

Kobbi Nissim[*]
Department of Computer Science
Ben-Gurion University of the Negev
kobbi@cs.bgu.ac.il

Rann Smorodinsky[†]
Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology
rann@ie.technion.ac.il

Moshe Tennenholtz
Microsoft Research
13 Shenkar, Herzlyia, *and*
Faculty of Industrial Engineering and Management
Technion – Israel Institute of Technology
moshet@microsoft.com

ABSTRACT

We study the implementation challenge in an abstract interdependent values model and an arbitrary objective function. We design a generic mechanism that allows for approximate optimal implementation of *insensitive* objective functions in ex-post Nash equilibrium. If, furthermore, values are private then the same mechanism is strategy proof. We cast our results onto two specific models: pricing and facility location. The mechanism we design is optimal up to an additive factor of the order of magnitude of one over the square root of the number of agents and involves no utility transfers.

Underlying our mechanism is a lottery between two auxiliary mechanisms — with high probability we actuate a mechanism that reduces players influence on the choice of the social alternative, while choosing the optimal outcome with high probability. This is where *differential privacy* is employed. With the complementary probability we actuate a mechanism that may be typically far from optimal but is incentive compatible. The joint mechanism inherits the desired properties from both.

Categories and Subject Descriptors

F.m [**Theory of Computation**]: Miscellaneous

General Terms

Theory

Keywords

Differential Privacy, Mechanism Design, Monopolist Pricing, Facility Location

1. INTRODUCTION

Mechanism design deals with the implementation of desired outcomes in a multi-agent system with asymmetric information, such as setting the price for a good, allocating goods to agents, and locating public facilities. The quality of the mechanism's outcome is measured by some objective function. In many instances, this function is simply the sum of the agents' valuations for an outcome, but it can take many other forms such as the revenue of a seller in an auction setting, the social inequality in a market setting and more. The challenge of mechanism design is to design mechanisms which exhibit dominant strategies for the players, such that once players play their dominant strategies the outcome of the mechanism coincides with maximizing the objective function.

As it turns out, such powerful mechanisms do not exist in general. The famous Gibbard-Satterthwaite theorem [15, 31] tells us that for non-restricted settings any non-trivial truthful mechanism is dictatorial. However, if we restrict attention to the objective function that is the sum of the agents' valuations, then this problem can be overcome by introducing monetary payments, as is the case with Vickrey-Clarke-Groves mechanisms [39, 8, 17] that guarantee that being truthful is a dominant strategy and the outcome is optimal. Unfortunately, Roberts [28] showed that a similar mechanism cannot be obtained for other objective functions. The reader is referred to [21] for a broader introduction.

This cul-de-sac induced researchers to "lower the bar" for mechanism design. One possibility is to replace the solution concept with a weaker one, as is the case with the large body of literature on Bayes-Nash implementation. Another possibility is to relax the quest from exact optimal implementation to *approximate implementation.* This research agenda turned out to be fruitful and yielded many positive results. A sequence of papers on *virtual implementation*, initiated by Matsushima [22] and Abreu and Sen [2], provides general conditions for approximate implementation where the approximation inaccuracy in a fixed model can be made arbitrarily small. On the other hand, the recent literature emerging from the *algorithmic* mechanism design commu-

[*]Part of this research was performed while the author was at Microsoft AI, Israel. Work partly supported by ISF grant 860/06, the Frankel Center for Computer Science at Ben-Gurion University and the Sapir Center for Economic Development.

[†]Research supported by Technion VPR grants, the Bernard M. Gordon Center for Systems Engineering at the Technion and the Sapir Center for Economic Development.

ITCS January 8-10, 2012, Boston, MA
Copyright 2012 ACM 978-1-4503-1115-1 ...$10.00.

nity looks at approximation inaccuracies which are a function of the *size* of the model (measured, e.g., the number of agents).

Interestingly, no *general* techniques are known for designing mechanisms that are approximately optimal for arbitrary social welfare functions. To demonstrate this consider the facility location problem, where a social planner needs to locate some facilities, based on agents' reports of their own location. This problem has received extensive attention recently, yet small changes in the model result in different techniques which seem tightly tailored to the specific model assumptions (see [5, 27, 19]).

Another line of research, initiated by Moulin [25], is that on mechanism design *without money.* Moulin, and later Schummer and Vohra [34, 35], characterized functions that are truthfully implementable without payments and studied domains in which non-dictatorial functions can be implemented. More recently, Procaccia and Tennenholtz [27] studied a relaxation of this notion – *approximate* mechanism design without money.

1.1 Our Contributions

We introduce an abstract mechanism design model where agents have interdependent values and provide a generic technique for approximate implementation of an arbitrary objective function. More precisely, we bound the worst case difference between the optimal outcome and the expected outcome of our generic mechanism by $O(\sqrt{\ln(n)/n})$, where n is the population size. Our generic construction does not involve utility transfer.

Our construction combines two very different random mechanisms:

- With high probability we deploy a mechanism that chooses social alternatives with a probability that is exponential in the objective function, assuming players are truthful. This mechanism exhibits two important properties: (1) Agents have small influence on the outcome of the mechanism and consequently have little influence on their own utility. As a result all strategies, including truthfulness, are ϵ-dominant. (2) Under the assumption that players are truthful, alternatives which are nearly optimal are most likely to be chosen. The concrete construction we use is the Exponential Mechanism presented by McSherry and Talwar [24].
- With vanishing probability we deploy a mechanism which is designed with the goal of eliciting agents' private information, while ignoring the objective function.

Our constructions are in a setting that is an extension of the classical social choice model. In the classical model agents' utilities are expressed as a function of the private types and a social alternative, and the issue of how agents exploit the social choice made is not treated explicitly. We make this choice explicit by adding an additional stage to the model, where, following the choice of the social alternative, agents take an action to exploit the social alternative and determine their utility (hereinafter *reaction*). A few examples demonstrate the prevalence of reactions in typical design problems: (1) In a Facility Location problem agents react to the mechanism's outcome by choosing one of the facilities (e.g., choose which school to attend). (2) A Monopolist posts a price based on agents input. Agents react by either buying the good or not. (3) In an exchange economy agents react to the price vector (viewed as the outcome of the invisible hand mechanism) by demanding specific bundles. (4) In a public good problem, where a set of substitutable goods is supplied, each agent must choose her favorite good. (5) Finally, consider a network design problem, where each agent must choose the path it will use along the network created by the society.

With this addendum to the model one can enrich the notion of a mechanism; in addition to determining a social choice the mechanism can also restrict the set of reactions available to an agent. For example, in the context of pricing digital goods, the mechanism can choose a price p and, in addition, allocate the good to and charges p from exactly those agents whose bid is greater or equal to p. We refer to this aspect of mechanisms as *imposition.*

1.2 Related Work

Virtual implementation.

The most closely related body of work is the literature on virtual implementation with incomplete information, derived from earlier work on virtual implementation with complete information which was initiated by Matsushima [22] and Abreu and Sen [2]. A social choice function is *virtually implementable* if for any $\epsilon > 0$ there exists a mechanism which equilibria result in outcomes that ϵ-approximate the function. Results due to Abreu and Matsushima [1], Duggan [9], and Serrano and Vohra [36, 37] provide necessary and sufficient conditions for functions to be virtually implementable in various environments with private information. A common thread throughout the results on virtual implementation under incomplete information is the incentive compatibility requirement over the social choice function, in addition to some form of type diversity. Compared with our contribution the above mentioned work provides positive results in environments with small populations, whereas we require large populations in order to have a meaningful approximation. On the other hand, the solution concepts we focus on are ex-post Nash equilibrium and strict dominance (for the private values setting), compared with iterated deletion of dominated strategies or Bayes-Nash equilibria, provided in the above mentioned papers. In addition, the virtual implementation results apply to functions that are incentive compatible from the outset, whereas our technique applies to arbitrary insensitive objective functions. In both cases the mechanisms proposed do not require utility transfers but do require some kind of type diversity.

Influence and Approximate Efficiency.

The basic driving force underlying our construction is ensuring that each agent has a vanishing influence on the outcome of the mechanism as the population grows. Intuitively, if players are non-influential, then they might as well be truthful. This idea has been used by various authors to provide mechanisms that approximate efficient outcomes when the population of players is large. Some examples of work that hinge on a similar principle for large, yet finite populations, are Roberts and Postlewaite [29], Swinkels [38] who studies auctions, Satterthwaite and Williams [32] and Rustichini, Satterthwaite and Williams [30] who study double auctions, and Al-Najjar and Smorodinsky [4] who study an exchange market. The mechanisms provided in these papers are designed for maximizing the sum of agents' valuations.

In contrast, our results hold a for a wide range of objective functions and are generic in nature. Interestingly, a similar argument, hinging on players' lack of influence, is instrumental to show inefficiency in large population models. For example, Mailath and Postlewaite [20] demonstrate 'free-riding' in the context of public goods, which eventually leads to inefficiency).

A formal statement of influence in an abstract setting appears in Levine and Pesendorfer [18] and Al-Najjar and Smorodinsky [3]. Beyond the formalization of influence these works provide bounds on aggregate measures of influence such as the average influence or on the number of influential agents. McLean and Postlewaite [23] introduce the notion of informational smallness, formalizing settings where one player's information is insignificant with respect to the aggregated information.

Differential Privacy.

The notion of *differential privacy*, recently introduced by Dwork, McSherry, Nissim and Smith [13], captures individual privacy by the limited impact of any single agent's input on the outcome of a joint computation. Differential privacy stipulates that the influence of any contributor to the computation is bounded in a very strict sense: any change in the input contributed by an individual translates to at most a near-one multiplicative factor in the probability distribution over the set of outcomes.The scope of computations that were shown to be computed in a differentially private manner has grown significantly since the introduction of the concept and the reader is referred to [11, 12] for recent surveys.

McSherry and Talwar [24] established an inspiring connection between differential privacy and mechanism design, where differential privacy is used as a tool for constructing efficient mechanisms. They observed that participants (players) that contribute private information to ϵ-differentially private computations have a limited incentive to lie, even if their utility is derived from the joint outcome. Consequently, truth-telling is approximately dominant in mechanisms that are ϵ-differentially private, regardless of the agent utility functions. McSherry and Talwar introduced a generic tool called the *Exponential Mechanism* for constructing ϵ-differentially private mechanisms.They showed that whenever agents are truthful the exponential mechanism chooses a social alternative which almost optimizes the objective function. They demonstrated the power of this mechanism in the context of Unlimited Supply Auctions, Attribute Auctions, and Constrained pricing.

The contribution of McSherry and Talwar leaves much to be desired in terms of mechanism design: (1) It is not clear how to set the value of ϵ. Lower values of ϵ imply higher compatibility with incentives, on the one hand, but deteriorate the approximation results on the other hand. The model and results of McSherry and Talwar do not provide a framework for analyzing these countervailing forces. (2) Truth telling is *approximately* dominant, but, in fact, in the mechanisms they design **all** strategies are approximately dominant, which suggests that truth telling may have no intrinsic advantage over any other strategy in their mechanism. (3) Furthermore, one can demonstrate that misreporting one's private information can actually dominate other strategies, truth-telling included. To make things worse, such dominant strategies may lead to inferior results for the social planner. This is demonstrated in Example 1 (Appendix 4.1), in the context of monopoly pricing. Following an early version of our work, Xiao has provided a second such example, in the context of facility location [40].

Facility Location.

One of the concrete examples we investigate is the optimal location of facilities. The single facility case on the line with single-peaked preferences can be solved optimally by selecting the median declaration. The 2-facility problem turns out to be more complicated. Recently, Wang et al. [19] introduced a randomized 4-(multiplicative) approximation truthful mechanism for the 2 facility problem. The techniques introduced herein provide additive $\tilde{O}(n^{-1/3})$ approximation to the average optimal distance between the agents and the facilities.

Following our formalization of *reactions* and of *imposition* and its applicability to facility location, Fotakis and Tzamos [14] provide *imposing* versions of previously known mechanisms to improve implementation accuracy. They provide constant multiplicative approximation or logarithmic multiplicative approximation, albeit with fully imposing mechanisms.

Non discriminatory Pricing of Digital Goods.

Another setting where we demonstrate our generic results is a pricing application, where a monopolist sets a single price for goods with zero marginal costs (digital goods) in order to maximize revenues. We consider environments where the potential buyers have *interdependent valuations* for the good. Pricing mechanisms for the *private values* case have been studied by Goldberg et al. [16] and Balcan et al. [7]. They consider settings where agents' valuation are not necessarily restricted to a finite set and achieve $O(n^{-1/2})$-implementation. Our mechanism provides a similar bound in a setting with finitely many possible prices. However, it is derived from general principles and therefore more robust. Furthermore, it is applicable beyond the private values setting.

2. PRELIMINARIES

Let N denote a set of n agents, S denote a *finite* set of social alternatives, and T_i, $i = 1, \ldots, n$, be a finite type space for agent i. We denote by $T = \times_{i=1}^{n} T_i$ the set of type tuples and write $T_{-i} = \times_{j \neq i} T_j$ with generic element t_{-i}. Agent i's type, $t_i \in T_i$, is her private information. Let R_i be the set of reactions available to i. Typically, once a social alternative $s \in S$ is determined agents choose a reaction $r_i \in R_i$. The utility of an agent i is therefore a function of the vector of types, the chosen social alternative and the chosen reaction. Formally, $u_i : T \times S \times R_i \to [0, 1]$.[1] A tuple (T, S, R, u), where $R = \times_{i=1}^{n} R_i$ and $u = (u_1, \ldots, u_n)$, is called an *environment.* We will use $r_i(t, s)$ to denote an arbitrary optimal reaction for agent i, i.e., $r_i(\cdot, \cdot)$ is an arbitrary function satisfying $r_i(t, s) \in \mathbf{arg\,max}_{r_i \in R_i} u_i(t, s, r_i)$.

We say that agent i has *private reactions* if the optimal reaction of i depends only on her type and the social alternative, i.e., if

$$\mathop{\mathbf{arg\,max}}_{r_i \in R_i} u_i((t_i, t_{-i}), s, r_i) = \mathop{\mathbf{arg\,max}}_{r_i \in R_i} u_i((t_i, t'_{-i}), s, r_i),$$

[1] Utilities are assumed to be bounded in the unit interval. This is without loss of generality, as long as there is some uniform bound on the utility.

for all s, i, t_i, t_{-i} and t'_{-i}. To emphasize that $r_i(t,s)$ does not depend on t_{-i} we will use in this case the notation $r_i(t_i, s)$ to denote an arbitrary optimal reaction for agent i.

We say that an agent has *private values* if she has private reactions and furthermore her utility depends only on her type, social alternative and reaction, i.e.,

$$u_i((t_i, t_{-i}), s, r_i) = u_i((t_i, t'_{-i}), s, r_i),$$

for all s, i, t_i, t_{-i} and t'_{-i}. In this case we will use the notation $u_i(t_i, s, r_i)$ to denote the agent's utility, to emphasize that it does not depend on t_{-i}. In the more general setting, where the utility u_i and the optimal reaction r_i may depend on t_{-i}, we say that agents have *interdependent values.*

An environment is *non-trivial* if for any pair of types there exists a social alternative for which the optimal reactions are distinct. Formally, $\forall i$, $t_i \neq \hat{t}_i \in T_i$ and t_{-i} there exists $s \in S$, denoted $s(t_i, \hat{t}_i, t_{-i})$, such that

$$\mathop{\mathbf{arg\,max}}_{r_i \in R_i} u_i((t_i, t_{-i}), s, r_i) \cap \mathop{\mathbf{arg\,max}}_{r_i \in R_i} u_i((\hat{t}_i, t_{-i}), s, r_i) = \emptyset.$$

We say that $s(t_i, \hat{t}_i, t_{-i})$ *separates* between t_i and $\hat{t}_i$ at t_{-i}. A set of social alternatives, $\tilde{S} \subset S$ is called *separating* if for any i and $t_i \neq \hat{t}_i$ and t_{-i}, there exists some $s(t_i, \hat{t}_i, t_{-i}) \in \tilde{S}$ that separates between t_i and $\hat{t}_i$ at t_{-i}.

The Objective Function.

A social planner, not knowing the vector of types, wants to maximize an arbitrary *objective function* (sometimes termed *social welfare function*), $F : T \times S \to [0, 1]$.[2] We focus our attention on a class of functions for which individual agents have a diminishing impact, as the population size grows:

DEFINITION 1 (SENSITIVITY [13]). *The objective function $F : T \times S \to [0, 1]$ is d-*sensitive *if $\forall i, t_i \neq \hat{t}_i, t_{-i}$ and $s \in S$,*

$$|F((t_i, t_{-i}), s) - F((\hat{t}_i, t_{-i}), s)| \leq \frac{d}{n},$$

where n is the population size.[3]

One commonly used objective function which is 1-sensitive is the average utility, $F(t, s) = \frac{1}{n}\sum_i u_i(t, s, r_i(t, s))$.

Mechanisms.

Denote by $\mathcal{R}_i = 2^{R_i} \setminus \{\emptyset\}$ the set of all subsets of R_i, except for the empty set, and let $\mathcal{R} = \times_i \mathcal{R}_i$. A (direct) mechanism randomly chooses, for any vector of inputs t a social alternative, and for each agent i a subset of available reactions. Formally:

DEFINITION 2 (MECHANISM). *A (direct)* mechanism *is a function $M : T \to \Delta(S \times \mathcal{R})$.*[4]

If some of the agents do not have private reactions, then the mechanism also discloses the vector of agents' announcements, and agents can use this information to choose a reaction.

We denote by $M_S(t)$ the marginal distribution of $M(t)$ on S and by $M_i(t)$ the marginal distribution on $\mathcal{R}_i$. We say that the mechanism M is *non-imposing* if $M_i(t)(R_i) = 1$. That is, the probability assigned to the grand set of reactions is one, for all i and $t \in T$. Put differently, the mechanism never restricts the set of available reactions. M is *ϵ-imposing* if $M_i(t)(R_i) \geq 1 - \epsilon$ for all i and $t \in T$. In words, with probability exceeding $1 - \epsilon$ the mechanism imposes no restrictions.

Strategies and Solution Concepts.

A mechanism induces the following game with incomplete information. In the first phase, agents announce their types simultaneously to the mechanism and the mechanism chooses a social alternative and a subset of reactions for each agent. In the second phase, each agent, knowing the strategy tuple of all agents, the vector of announced types, the social alternative and her available set of reactions, must choose one such reaction.

Let $W_i : T_i \to T_i$ denote the announcement of agent i, given her type and let $W = (W_i)_{i=1}^n$. Upon the announcement of the social alternative s, the vector of opponents' announcements, t_{-i} and a subset of reactions, $\hat{R}_i \subset R_i$, the rational agent will choose an arbitrary optimal reaction,

$$r_i((t_i, W_{-i}^{-1}(t_{-i})), s, \hat{R}_i),$$

where $W_{-i}^{-1}(t_{-i})$ denotes the pre-image of W_{-i} at the vector of announcements t_{-i}.[5] Therefore, we can view $(W_i)_{i=1}^n$ as the agents' strategies, without an explicit reference to the choice of reactions.

Given a vector of types, t, and a strategy tuple W, the mechanism M induces a probability distribution, $M(W(t))$ over the set of social alternatives and reaction tuples. The expected utility of i, at a vector of types t, is

$$\mathop{\mathrm{E}}_{M(W(t))} u_i(t, s, r_i),$$

where r_i is short-writing for the optimal reaction, which itself is determined by M and W. Hereinafter we suppress the reference to the reactions in our notations and write $E_{M(W(t))}u_i(t, s)$ instead of $E_{M(W(t))}u_i(t, s, r_i)$.

A strategy W_i is *dominant* for the mechanism M if for any vector of types $t \in T$, any alternative strategy $\hat{W}_i$ of i and any strategy profile $\bar{W}_{-i}$ of i's opponents, i.e.,

$$\mathop{\mathrm{E}}_{M((W_i(t_i), \bar{W}_{-i}(t_{-i})))} u_i(t, s) \geq \mathop{\mathrm{E}}_{M((\hat{W}_i(t_i), \bar{W}_{-i}(t_{-i})))} u_i(t, s).$$

In words, W_i is a strategy that maximizes the expected payoff of i for any vector of types and any strategy used by her opponents. If for all i the strategy $W_i(t_i) = t_i$ is dominant then M is called *truthful* (or *strategyproof*).A strategy W_i is *strictly dominant* if it is dominant and furthermore whenever $W(t_i) \neq \hat{W}(t_i)$ then the above inequality holds strongly. If $W_i(t_i) = t_i$ is strictly dominant for all i then M is *strictly truthful.*

[2] In fact, one can consider objective functions of the form $F : T \times S \times R \to [0, 1]$. Our results go through if for any t and s and any i and r_{-i} the functions $F(t, s, (r_{-i}, \cdot)) : R_i \to [0, 1]$ and $u_i(t, s, \cdot) : R_i \to [0, 1]$ are co-monotonic. In words, as long as the objective function's outcome (weakly) increases whenever a change in reaction increases an agent's utility.

[3] In the definition of sensitivity one can replace the constant d with a function $d = d(n)$ that depends on the population size. Our go through for the more general case as long as $d = o(n)$.

[4] The notation $\Delta(S)$ denotes the set of probability distributions over S.

[5] We slightly abuse notation as $W_{-i}^{-1}(t_{-i})$ may not be a singleton but a subset of type vectors, in which case the optimal reaction is not well defined.

A strategy tuple W is an *ex-post Nash Equilibrium* if for all i and $t \in T$ and for any strategy $\hat{W}_i$ of player i,

$$\mathop{\mathrm{E}}_{M(W(t))} u_i(t,s) \geq \mathop{\mathrm{E}}_{M((\hat{W}_i(t_i), W_{-i}(t_{-i})))} u_i(t,s).$$

If $\{W_i(t_i) = t_i\}_{i=1}^n$ is an ex-post Nash equilibrium then M is *ex-post Nash truthful.*

Implementation.

Given a vector of types, t, the expected value of the objective function, F, at the strategy tuple W is $E_{M(W(t))}[F(t,s)]$.

Definition 3 (β-implementation). *We say that the mechanism M β-implements F in (strictly) dominant strategies, for $\beta > 0$, if there exists a (strictly) dominant strategy tuple, W, and for any such tuple and for any $t \in T$,*

$$\mathop{\mathrm{E}}_{M(W(t))}[F(t,s)] \geq max_{s \in S} F(t,s) - \beta.$$

A mechanism M β-implements F in an ex-post Nash equilibrium if for some ex-post Nash equilibrium strategy tuple, W, for any $t \in T$,

$$\mathop{\mathrm{E}}_{M(W(t))}[F(t,s)] \geq max_{s \in S} F(t,s) - \beta.$$

Main Theorem (informal statement).

For any d-sensitive function F and $1 > \beta > 0$ there exists a number n_0 and a mechanism M which β-implements F in an ex-post Nash equilibrium, whenever the population has more than n_0 agents. If, in addition, reactions are private then M β-implements F in strictly dominant strategies.

3. A FRAMEWORK OF APPROXIMATE IMPLEMENTATION

In this section we present a general scheme for implementing arbitrary objective functions in large societies. The convergence rate we demonstrate is of an order of magnitude of $\sqrt{\ln(n)/n}$. Our scheme involves a lottery between two mechanisms: (1) A differentially private mechanism that we instantiate with the *Exponential Mechanism* of McSherry and Talwar [24]; and (2) The *Commitment Mechanism*, where imposition is used to commit agents to take a reaction that complies with their announced type.

3.1 Differential Privacy and the Exponential Mechanism

Dwork, McSherry, Nissim, and Smith defined differential privacy as the requirement that for any vector of announcements, a unilateral deviation changes the probabilities assigned to any social choice by a (multiplicative) factor of close to 1. Intuitively, this means that M does not disclose (almost) any information about any single contributor.

Definition 4. *[ϵ-differential privacy [13]] A mechanism, M, provides ϵ-differential privacy if it is non-imposing and for any $s \in S$, any pair of type vectors $t, \hat{t} \in T$, which differ only on a single coordinate, $M(t)(s) \leq e^{\epsilon} \cdot M(\hat{t})(s)$.*[6]

The appeal of mechanisms that provide ϵ-differential privacy is that they induce near indifference among all strategies, as stated in the following (folk) lemma which proof we include for completeness:

Lemma 1. *If M is non-imposing and provides ϵ-differential privacy, for some $\epsilon < 1$, then for any agent i, any type tuple t, any strategy tuple W, and any alternative strategy for i, $\hat{W}_i$ the following holds:*

$$\left| \mathop{\mathrm{E}}_{M(W(t))}[u_i(t,s)] - \mathop{\mathrm{E}}_{M(\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)] \right| < 2\epsilon.$$

Proof: Let W and $\hat{W}$ be two strategy vectors that differ the i'th coordinate. Then for every $t \in T$, $s \in S$, $r_i \in R_i$ and $u_i : T \times S \times R_i \to [0,1]$ we have

$$\begin{aligned}
\mathop{\mathrm{E}}_{M(W(t))}[u_i(t,s)] &= \sum_{s \in S} M(W(t))(s) \cdot u_i(t,s) \\
&\leq \sum_{s \in S} e^{\epsilon} \cdot M(\hat{W}_i(t_i), W_{-i}(t_{-i}))(s) \cdot u_i(t,s) \\
&= e^{\epsilon} \cdot \mathop{\mathrm{E}}_{\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)],
\end{aligned}$$

where the inequality follows since M provides ϵ-differential privacy, and u_i is non-negative. A similar analysis gives

$$\mathop{\mathrm{E}}_{\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)] \leq e^{\epsilon} \cdot \mathop{\mathrm{E}}_{M(W(t))}[u_i(t,s)].$$

Hence we get:

$$\begin{aligned}
&\mathop{\mathrm{E}}_{M(W(t))}[u_i(t,s)] - \mathop{\mathrm{E}}_{M(\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)] \leq \\
&\quad (e^{\epsilon} - 1) \cdot \mathop{\mathrm{E}}_{M(\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)] \leq \\
&\quad e^{\epsilon} - 1,
\end{aligned}$$

where the last inequality holds because u_i returns a values in $[0,1]$. Similarly,

$$\mathop{\mathrm{E}}_{M(\hat{W}_i(t_i), W_{-i}(t_{-i}))}[u_i(t,s)] - \mathop{\mathrm{E}}_{M(W(t))}[u_i(t,s)] \leq e^{\epsilon} - 1.$$

To conclude the proof note that $(e^{\epsilon} - 1) \leq 2\epsilon$ whenever $0 \leq \epsilon \leq 1$.

QED

We use a generic construction of differentially private mechanism by McSherry and Talwar [24]:

$$M^{\epsilon}(t)(s) = \frac{e^{n\epsilon F(t,s)}}{\sum_{\bar{s} \in S} e^{n\epsilon F(t,\bar{s})}}.$$

Lemma 2 ([24]). *If F is d-sensitive then $M^{\frac{\epsilon}{2d}}(t)$ preserves ϵ-differential privacy.*

Proof: Let t and $\hat{t}$ be or two type vectors that differ on a single coordinate. Then for any $s \in S$, $F(t,s) - \frac{d}{n} \leq F(\hat{t},s) \leq F(t,s) + \frac{d}{n}$, hence,

$$\frac{M^{\frac{\epsilon}{2d}}(t)(s)}{M^{\frac{\epsilon}{2d}}(\hat{t})(s)} = \frac{\frac{e^{\frac{n\epsilon F(t,s)}{2d}}}{\sum_{\bar{s} \in S} e^{\frac{n\epsilon F(t,\bar{s})}{2d}}}}{\frac{e^{\frac{n\epsilon F(\hat{t},s)}{2d}}}{\sum_{\bar{s} \in S} e^{\frac{n\epsilon F(\hat{t},\bar{s})}{2d}}}} \leq \frac{\frac{e^{\frac{n\epsilon F(t,s)}{2d}}}{\sum_{\bar{s} \in S} e^{\frac{n\epsilon F(t,\bar{s})}{2d}}}}{\frac{e^{\frac{n\epsilon (F(t,s) - \frac{d}{n})}{2d}}}{\sum_{\bar{s} \in S} e^{\frac{n\epsilon (F(t,\bar{s}) + \frac{d}{n})}{2d}}}} = e^{\epsilon}.$$

QED

[6] For non discrete sets of alternatives the definition requires that $\frac{M(t)(\hat{S})}{M(\hat{t})(\hat{S})} \leq e^{\epsilon} \quad \forall \hat{S} \subset S$.

McSherry and Talwar [24] note in particular that in the case of private values truthfulness is 2ϵ-dominant, which is an immediate corollary of Lemma 1. They combine this with the following observation to conclude that exponential mechanisms approximately implement F in ϵ- dominant strategies:

LEMMA 3 ([24] SIMPLIFIED). *Let $F : T^n \times S \to [0,1]$ be an arbitrary d-sensitive objective function and $n > \frac{2ed}{\epsilon|S|}$. Then for any t,*

$$\mathop{\mathrm{E}}_{M^{\frac{\epsilon}{2d}}(t)}[F(t,s)] \geq \max_s F(t,s) - \frac{4d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right).$$

Proof: Let $\delta = \frac{2d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right)$. As $n > \frac{e2d}{\epsilon|S|}$ we conclude that $\ln\left(\frac{n\epsilon|S|}{2d}\right) > \ln e > 0$ and, in particular, $\delta > 0$.

Fix a vector of types, t and denote by $\hat{S} = \{\hat{s} \in S : F(t,\hat{s}) < \max_s F(t,s) - \delta\}$. For any $\hat{s} \in \hat{S}$ the following holds:

$$M^{\frac{\epsilon}{2d}}(t)(\hat{s}) = \frac{e^{\frac{n\epsilon F(t,\hat{s})}{2d}}}{\sum_{s' \in S} e^{\frac{n\epsilon F(t,s')}{2d}}} \leq \frac{e^{\frac{n\epsilon(\max_s F(t,s)-\delta)}{2d}}}{e^{\frac{n\epsilon \max_s F(t,s)}{2d}}} = e^{-\frac{n\epsilon}{2d}\delta}.$$

Therefore, $M^{\frac{\epsilon}{2d}}(t)(\hat{S}) = \sum_{\hat{s}\in\hat{S}} M^{\frac{\epsilon}{2d}}(t)(\hat{s}) \leq |\hat{S}|e^{-\frac{n\epsilon}{2d}\delta} \leq |S|e^{-\frac{n\epsilon}{2d}\delta}$. Which, in turn, implies:

$$\begin{aligned}\mathop{\mathrm{E}}_{M^{\frac{\epsilon}{2d}}(t)}[F(t,s)] &\geq (\max_s F(t,s) - \delta)(1 - |S|e^{-\frac{n\epsilon}{2d}\delta}) \\ &\geq \max_s F(t,s) - \delta - |S|e^{-\frac{n\epsilon}{2d}\delta}.\end{aligned}$$

Substituting for δ we get that

$$\mathop{\mathrm{E}}_{M^{\frac{\epsilon}{2d}}(t)}[F(t,s)] \geq \max_s F(t,s) - \frac{2d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right) - \frac{2d}{n\epsilon}.$$

In addition, $n > \frac{e2d}{\epsilon|S|}$ which implies $\ln\left(\frac{n\epsilon|S|}{2d}\right) > \ln(e) = 1$, and hence $\frac{2d}{n\epsilon} \leq \frac{2d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right)$. Plugging this into the previous inequality yields

$$\mathop{\mathrm{E}}_{M^{\frac{\epsilon}{2d}}(t)}[F(t,s)] \geq \max_s F(t,s) - \frac{4d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right)$$

as desired.
QED

Note that $\lim_{n\to\infty} \frac{4d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right) = 0$ whenever the parameters d, ϵ and $|S|$ are held fixed.[7] Therefore, the exponential mechanism is almost optimal for a large and truthful population.

Remark:.

Other mechanisms exhibiting similar properties (i.e., 'almost indifference' and 'approximate optimality') may replace the exponential mechanism in our constructions. Such mechanisms may result, e.g., by applying techniques for converting computations into differentially private computations: addition of noise calibrated to global sensitivity [13], addition of noise calibrated to smooth sensitivity, and the sample and aggregate framework [26].

3.2 The Commitment Mechanism

We now consider a generic imposing mechanism. Our mechanism chooses $s \in S$ randomly (i.e., ignoring agents' announcements). Once s is chosen the mechanism restricts the allowable reactions for i to those that are optimal assuming all agents are truthful. Formally, if s is chosen according to the probability distribution P, let M^P denote the following mechanism:

$$M_S^P(t)(s) = P(s) \quad \text{and} \quad M_i^P(t)(r_i(t,s))|s) = 1.$$

Note that agents are (weakly) better off being truthful.

We define the *gap* of the environment, $\gamma = g(T,S,A,u)$, as:

$$\begin{aligned}&\gamma = g(T,S,A,u) = \\ &\qquad \min_{i,t_i\neq b_i,t_{-i}} \max_{s\in S}\left(u_i(t,s,r_i(t,s)) - u_i(t,s,r_i((b_i,t_{-i}),s))\right).\end{aligned}$$

In words, γ is a lower bound for the loss incurred by misreporting in case of an adversarial choice of $s \in S$. In non-trivial environments $\gamma > 0$. To provide some intuition consider a simple facility location where agents can either be in city A, B or C and the location of 2 facilities must be decided within these 3 choices. An agent that is located in A and announces B considers the outcome that the mechanism chooses to locate the 2 facilities in A and B but is assigned the optimal facility vis-a-vis her announcement. In such a case she incurs a utility loss that is proportional to the distance between A and B. Spanning this over all possible true locations and false announcements we conclude that the gap is $\min\{|A-B|, |B-C|, |C-A|\}$.

We say the a distribution P is *separating* if there exists a separating set $\tilde{S} \subset S$ such that $P(\tilde{s}) > 0$ for all $\tilde{s} \in \tilde{S}$. In this case we also say that M^P is a separating mechanism. In particular let $\tilde{p} = \min_{s\in\tilde{S}} P(s)$. Clearly one can choose P such that $\tilde{p} \geq \frac{1}{|S|}$. The following is straightforward:

LEMMA 4. *If the environment (T,S,A,u) is non-trivial and P is a separating distribution over S then $\forall b_i \neq t_i, t_{-i}$,*

$$\begin{aligned}&\mathop{\mathrm{E}}_{M^P(t_i,t_{-i})}[u_i(t,s,r_i(t,s))] \geq \\ &\qquad \mathop{\mathrm{E}}_{M^P(b_i,t_{-i})}[u_i(t,s,r_i((b_i,t_{-i}),s))] + \tilde{p}\gamma.\end{aligned}$$

If, in addition, reactions are private, then for any i, $b_i \neq t_i$, t_{-i} and b_{-i}:

$$\begin{aligned}&\mathop{\mathrm{E}}_{M^P(t_i,b_{-i})}[u_i(t,s,r_i(t_i,s))] \geq \\ &\qquad \mathop{\mathrm{E}}_{M^P(b_i,b_{-i})}[u_i(t,s,r_i(b_i,s))] + \tilde{p}\gamma.\end{aligned}$$

Proof: For any pair $b_i \neq t_i$ and for any $s \in S$

$$u_i(t,s,r_i(t_i,s)) \geq u_i(t,s,r_i(b_i,s)).$$

In addition, there exists some $\hat{s} = s(t_i,b_i)$, satisfying $P(\hat{s}) \geq \tilde{p}$, for which

$$u_i(t,\hat{s},r_i(t_i,\hat{s})) \geq u_i(t,\hat{s},r_i(b_i,\hat{s})) + \gamma.$$

Therefore, for any i, $b_i \neq t_i \in T_i$ and for any t_{-i},

$$\begin{aligned}&\mathrm{E}_{M^P(t_i,t_{-i})}[u_i(t,s,r_i(t,s))] \geq \\ &\qquad \mathrm{E}_{M^P(b_i,t_{-i})}[u_i(t,s,r_i((b_i,t_{-i}),s))] + \tilde{p}\gamma,\end{aligned}$$

as claimed.

[7] This limit also approaches zero if $d, \epsilon, |S|$ depend on n, as long as d/ϵ is sublinear in n and $|S|$ is subexponential in n.

Recall that if reactions are private then $r_i(t, s) = r_i(t_i, s)$, namely the optimal reaction of an agent, given some social alternative s, depends only on the agent's type. Therefore we derive the result for private reactions by replacing $r_i((t_i, t_{-i}), s)$ with $r_i(t_i, s)$ on the left hand side of the last inequality and $r_i((b_i, t_{-i}), s)$ with $r_i(b_i, s)$ on the right hand side.
QED

The following is an immediate corollary:

Corollary 1. *If the environment (T, S, A, u) is non-trivial and P is a separating distribution over S then*

1. *Truthfulness is an ex-post Nash equilibrium of M^P.*
2. *If agent i has private reactions then truthfulness is a strictly dominant strategy for i in M^P.*

An alternative natural imposing mechanism is that of a random dictator, where a random agent is chosen to dictate the social outcome. Similarly, agents will be truthful in such a mechanism. However, the loss from misreporting can only be bounded below by $\frac{\gamma}{n}$, whereas the commitment mechanism gives a lower bound of $\gamma\tilde{p} \geq \frac{\gamma}{|S|}$, which is independent of the population size.

3.3 A Generic and Nearly Optimal Mechanism

Fix a non-trivial environment (T, S, A, u) with a gap γ, separating set $\tilde{S}$, a d-sensitive objective function F and a commitment mechanism, M^P, with $\tilde{p} = \min_{s \in \tilde{S}} P(s)$. Set

$$\bar{M}_q^{\epsilon}(t) = (1-q)M^{\frac{\epsilon}{2d}}(t) + qM^P(t).$$

From lemmas 1 (set $W(t_i) = t_i$) and 4 we get:

Theorem 1. *If $q\tilde{p}\gamma \geq 2\epsilon$ then the mechanism $\bar{M}_q^{\epsilon}$ is ex-post Nash truthful. Furthermore, if agents have private reactions then $\bar{M}_q^{\epsilon}$ is strictly truthful.*

Let n_0 be the minimal integer satisfying

$$n_0 \geq \max\{\frac{8d}{\tilde{p}\gamma}\ln\left(\frac{\tilde{p}\gamma|S|}{2d}\right), \frac{4e^2d}{\tilde{p}\gamma|S|}\}$$

and

$$\frac{n_0}{\ln(n_0)} > \frac{8d}{\tilde{p}\gamma}.$$

Lemma 5. *Setting $\epsilon = \sqrt{\frac{\tilde{p}\gamma d}{n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}$, $q = \frac{2\epsilon}{\tilde{p}\gamma}$, we get for $n > n_0$ that (1) $q = \frac{2\epsilon}{\tilde{p}\gamma} < 1$, (2) $\epsilon < \tilde{p}\gamma$, and (3) $n > \frac{2ed}{\epsilon|S|}$.*

Proof: Part (1): $\frac{n}{\ln(n)} > \frac{n_0}{\ln(n_0)} \geq \frac{8d}{\tilde{p}\gamma}$ which implies $n > \frac{8d}{\tilde{p}\gamma}\ln(n)$. In addition, $n > n_0 > \frac{8d}{\tilde{p}\gamma}\ln\left(\frac{\tilde{p}\gamma|S|}{2d}\right)$. Therefore $n > \frac{4d}{\tilde{p}\gamma}\ln\left(\frac{\tilde{p}\gamma|S|}{2d}\right) + \frac{4d}{\tilde{p}\gamma}\ln(n) = \frac{4d}{\tilde{p}\gamma}\ln\left(\frac{\tilde{p}\gamma|S|n}{2d}\right) \implies (\tilde{p}\gamma)^2 > \frac{4\tilde{p}\gamma d}{n}\ln\left(\frac{\tilde{p}\gamma|S|n}{2d}\right)$. Taking the square root and substituting for ϵ on the right hand side yields $\tilde{p}\gamma > 2\epsilon$ and the claim follows.

Part (2) follows directly from part (1)

Part (3): $n > n_0 \geq \frac{4e^2d}{\tilde{p}\gamma|S|} \geq \frac{4e^2d}{\tilde{p}\gamma|S|^2} \implies \sqrt{n} > \frac{2ed}{\sqrt{\tilde{p}\gamma d}|S|}$. In addition $n > \frac{4e^2d}{\tilde{p}\gamma|S|} > \frac{2de}{\tilde{p}\gamma|S|}$ which implies $1 < \ln\left(\frac{\tilde{p}\gamma|S|n}{2d}\right)$. Combining these two inequalities we get:

$$\sqrt{n} > \frac{2ed}{\sqrt{\tilde{p}\gamma d}\sqrt{\ln\left(\frac{\tilde{p}\gamma|S|n}{2d}\right)}|S|}.$$

Multiplying both sides by $\sqrt{n}$ implies

$$n > \frac{2ed\sqrt{n}}{\sqrt{\tilde{p}\gamma d}\sqrt{\ln\left(\frac{\tilde{p}\gamma|S|n}{2d}\right)}|S|} = \frac{2ed}{\epsilon|S|}.$$

QED

Let $\hat{M}(t) = \bar{M}_q^{\epsilon}(t)$ where q, ϵ are as in Lemma 5 above. Our main result is:

Theorem 2. **(Main Theorem)** *$\hat{M}(t)$ is ex-post Nash truthful and, in addition, it $6\sqrt{\frac{d}{\tilde{p}\gamma n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}$-implements F in ex-post Nash equilibrium, for $n > n_0$.*

If agents have private reactions the mechanism is strictly truthful and $6\sqrt{\frac{d}{\tilde{p}\gamma n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}$-implements F in strictly dominant strategies.

Proof: Given the choice of parameters ϵ and q then, Theorem 1 guarantees that $\hat{M}(t)$ is ex-post Nash truthful (and truthful whenever reactions are private). Therefore, it is sufficient to show that for any type vector t,

$$\mathop{\mathrm{E}}_{\hat{M}(t)}(F(t,s)) \geq \max_s F(t,s) - 6\sqrt{\frac{d}{\tilde{p}\gamma n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}.$$

Note that as F is positive, $E_{M^P(t)}[F(t,s)] \geq 0$ and so $E_{\hat{M}(t)}[F(t,s)] \geq (1-q)E_{M^{\frac{\epsilon}{2d}}(t)}[F(t,s)]$. By part (3) of Lemma 5 we are guaranteed that the condition on the size of of the population of Lemma 3 holds and so we can apply Lemma 3 to conclude that:

$$\mathop{\mathrm{E}}_{\hat{M}(t)}[F(t,s)] \geq (1-q)\left(\max_s F(t,s) - \frac{4d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right)\right).$$

We substitute q with $\frac{2\epsilon}{\tilde{p}\gamma}$ and recall that $\max_s F(t,s) \leq 1$. In addition, part (1) of Lemma 5 asserts that $\frac{2\epsilon}{\tilde{p}\gamma} < 1$. Therefore,

$$\begin{aligned}\mathop{\mathrm{E}}_{\hat{M}(t)}[F(t,s)] &\geq \max_s F(t,s) - \frac{2\epsilon}{\tilde{p}\gamma} - \frac{4d}{n\epsilon}\ln\left(\frac{n\epsilon|S|}{2d}\right) \\ &\geq \max_s F(t,s) - \frac{2\epsilon}{\tilde{p}\gamma} - \frac{4d}{n\epsilon}\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right),\end{aligned}$$

where the last inequality is based on the fact $\epsilon < \tilde{p}\gamma$, which is guaranteed by part (2) of Lemma 5. Substituting ϵ for $\sqrt{\frac{\tilde{p}\gamma d}{n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}$ we obtain the desired conclusion, namely:

$$\mathop{\mathrm{E}}_{\hat{M}(t)}[F(t,s)] \geq \max_s F(t,s) \quad - \quad 6\sqrt{\frac{d}{\tilde{p}\gamma n}}\sqrt{\ln\left(\frac{n\tilde{p}\gamma|S|}{2d}\right)}.$$

QED

One particular case of interest is the commitment mechanism M^U, where U is the uniform distribution over the set S. Noting that $P = U$ implies that the minimal probability is $\tilde{p} = \frac{1}{|S|}$ we get:

COROLLARY 2. *Let n_0 be the minimal integer satisfying $n_0 \geq \max\{\frac{8\tilde{d}|S|}{\gamma}\ln\left(\frac{\gamma}{2d}\right), \frac{4e^2 d}{\gamma|S|}\}$ and $\frac{n_0}{\ln(n_0)} > \frac{8d|S|}{\gamma}$. Then the mechanism $\hat{M}^U(t)$ $6\sqrt{\frac{d|S|}{\gamma n}}\sqrt{\ln\left(\frac{n\gamma}{2d}\right)}$-implements F, in ex-post Nash equilibrium, for all $n > n_0$.*

If agents have private reactions the mechanism $\hat{M}^U(t)$ $6\sqrt{\frac{d|S|}{\gamma n}}\sqrt{\ln\left(\frac{n\gamma}{2d}\right)}$-implements F in strictly dominant strategies.

Holding the parameters of the environment $d, \gamma, |S|$ fixed the approximation inaccuracy of our mechanism converges to zero at a rate of $O(\sqrt{\ln(n)/n})$.

In summary, by concatenating the exponential mechanism, where truthfulness is ϵ-dominant with the commitment mechanism we obtain a strictly truthful mechanism. In fact, this would hold true for any mechanism where truthfulness is ϵ-dominant not only the exponential mechanism.

3.4 Applications

We show how out generic construction can be used in the setting of monopolistic pricing of goods, and discuss a multi parameter extension. We then demonstrate an application to facility location.

3.4.1 Monopolist Pricing

A monopolist producing goods for which the marginal cost of production is zero faces a set of indistinguishable buyers. Each buyer has a unit demand with a valuation in the unit interval. Agents are arranged in (mutually exclusive) cohorts and the valuations of cohort members are correlated. Each agent receives a private signal and her valuation is uniquely determined by the signals of all her cohort members. The monopolist wants to set a uniform price in order to maximize her average revenue per user.[8]

Assume there are $N \cdot D$ agents, with agents labeled (n, d) (the d^{th} agent in the n^{th} cohort). Agent (n, d) receives a signal $X_d^n \in \mathbb{R}$ and we denote a cohort's vector of signals by $X^n = \{X_d^n\}_{d=1}^D$. We assume that the valuation of an agent, V_d^n, is uniquely determined by the signals of her cohort members; $V_d^n = V_d^n(X^n)$.

We assume that each agent's signal is informative in the sense that $V_d^n(X^n) > V_d^n(\hat{X}^n)$ whenever $X^n > \hat{X}^n$ (in each coordinate a weak inequality holds and for at least one of the coordinates a strong inequality holds). That is, whenever an individual's signal increases the valuation of each of her cohort members increases.

Let $R_{(n,d)} = \{\text{`Buy'}, \text{`Not buy'}\}$ be the set of reactions for agent (n, d). The utility of (n, d), given the vector of signals $X = \{X^n\}_{n=1}^N = \{\{X_d^n\}_{d=1}^D\}_{n=1}^N$, and the price p, is

$$u_{(n,d)}(X, p, r_{(n,d)}) = \begin{cases} V_d^n(X^n) - p & \text{if } r_{(n,d)} = \text{`Buy'}, \\ 0 & \text{if } r_{(n,d)} = \text{`Not buy'}. \end{cases}$$

We assume that all valuations are restricted to the unit interval, prices are restricted to some finite grid $S = S_m = \{0, \frac{1}{m}, \frac{2}{m}, \ldots, 1\}$ (hence, $|S| = m + 1$), and X_d^n takes on only finitely many values. We assume the price grid is fine enough so that for any two vectors $X^n > \hat{X}^n$ there exists some price $p \in S$ such that $E(V^n|X^n) > p + \frac{2}{m} > p > E(V^n|\hat{X}^n)$. Therefore for vector of announcements there exists a maximal price for which optimal reaction is Buy. For that price, if an agent announces a lower value then the best reaction would be Not Buy, which will yield a loss of $\frac{1}{m}$ at least. Similarly, there exists the lowest price for which the optimal reaction is Not Buy. Announcing a higher value will result in the optimal reaction being Buy, which yields a loss of $\frac{1}{m}$ at least. We conclude that the gap is $\gamma = \frac{1}{m}$.

The monopolist wants to maximize $F(t, p) = \frac{p}{ND} \cdot |\{(n, d) : V_d^n(X^n) > p\}|$, the average revenue per buyer. Note that a unilateral change in the type of one agent may change at most the buying behavior of the D members in her cohort, resulting in a change of at most $\frac{pD}{ND} \leq \frac{D}{ND}$ in the average revenue. As the population size is ND we conclude that F is D-sensitive.

Let M_{dg} be a mechanism as in Corollary 2, where a Uniform Commitment mechanism is used:

COROLLARY 3. *For any D there exists some N_0 such that for all $N > N_0$ the mechanism M_{dg} $O(\sqrt{\frac{m^2}{N}\ln(\frac{N}{m})})$-implements F in ex-post Nash equilibrium.*

The literature on optimal pricing in this setting has so far concentrated on the private values case and has provided better approximations. For example, Balcan et al. [7], using sampling techniques from Machine Learning, provide a mechanism that $O(1/\sqrt{n})$-implements the maximal revenue without any restrictions to a grid.

A Multi Parameter Extension.

In the above setting we assumed a simple single-parameter type space. However, the technique provided does not hinge on this. In particular, it extends to more complex settings where agents have a multi-parameter type space. More concretely, consider a monopolist that produces G types of digital goods, each with zero marginal cost for production. There are N buyers, where each buyer assigns a value, in some bounded interval, to each subset of the G goods (agents want at most a singe unit of each good). The monopolist sets G prices, one for each good, and once prices are set each agent chooses his optimal bundle. The challenge of the monopolist is to maximize the average revenue per buyer. In this model types are sufficiently diverse. In fact, for any two types there exists a price vector that yields different optimal consumptions. Therefore, the scheme we provide applies just as well to this setting.

3.4.2 Application to Facility Location

Consider a population of n agents located on the unit interval. An agent's location is private information and a social planner needs to locate K similar facilities in order to minimize the average distance agents travel to the nearest facility.[9] We assume each agent wants to minimize her distance to the facility that services her. In particular, this entails that values (and reactions) are private. We furthermore assume that agent and facility locations are all restricted to a fixed finite grid on the unit interval, $L = L(m) = \{0, \frac{1}{m}, \frac{2}{m}, \ldots, 1\}$. Using the notation of previous sections, let $T_i = L$, $S = L^K$, and let $R_i = L$. The

[8] To make this more concrete one can think of the challenge of pricing a fire insurance policy to apartment owners. Each apartment building is a cohort that shares the same risk and once the risk is determined (via aggregation of agents' signals) each agent has a private valuation for the insurance.

[9] For expositional reasons we restricting attention to the unit interval and to the average travel distance. Similar results can be obtained for other sets in $\mathbb{R}^2$ and other metrics, such as distance squared.

utility of agent i is

$$u_i(t_i, s, r_i) = \begin{cases} -|t_i - r_i| & \text{if } r_i \in s, \\ -1 & \text{otherwise.} \end{cases}$$

Hence, $r_i(b_i, s)$ is the facility closest to the locations of the facility in s closest to b_i. Let $F(t, s) = \frac{1}{n} \sum_{i=1}^{n} u_i(t_i, s, r_i(t_i, s))$ be the social utility function, which is 1-sensitive (i.e., $d = 1$).

First, consider the uniform commitment mechanism $\hat{M}^U$, which is based on the uniform distribution over S for the commitment mechanism. Now consider the mechanism $\hat{M}_{LOC1}$, based on the uniform commitment mechanism, as in Corollary 2. Noting that $\gamma = \frac{1}{m}$, $|S| = (m+1)^K$ we immediately get from Theorem 2:

Corollary 4. *$\exists n_0$ such that $\forall n > n_0$ the mechanism $\hat{M}_{LOC1}$ $6\sqrt{\frac{m(m+1)^K}{n}}\sqrt{\ln\left(\frac{n}{2m}\right)}$- implements the optimal location in strictly dominant strategies.*

Now consider an alternative commitment mechanism. Consider the distribution P, over $S = L^K$, which chooses uniformly among all the following alternatives - placing one facility in location $\frac{j}{m}$ and the remaining $K-1$ facilities in location $\frac{j+1}{m}$, where $j = 0, \ldots, m-1$. Note that for any i, any pair $b_i \neq t_i$ is separated by at least one alternative in this set. For this mechanism $\tilde{p} = \frac{1}{m}$. Now consider the mechanism $\hat{M}_{LOC2}$, based on the commitment mechanism, M^P. In analogy to Theorem 2, setting $\epsilon = \frac{1}{m\sqrt{n}}\sqrt{\ln\left(\frac{n(m+1)^K}{2m^2}\right)}$ and $q = 2\epsilon m^2$ we get:

Corollary 5. *$\exists n_0$ such that $\forall n > n_0$ the mechanism $u\hat{M}_{LOC2}$ $\frac{6m}{\sqrt{n}}\sqrt{\ln\left(\frac{n(m+1)^K}{2m^2}\right)}$-implements the optimal location in strictly dominant strategies.*

For both mechanisms the approximation error converges to zero at a rate$O(\ln(n)/\sqrt{n})$ as society grows. In addition, the approximation error of both mechanisms grows as the grid size, m, grows. However in the second mechanism approximation deteriorates at a substantially slower rate.

4. DISCUSSION

4.1 Is differential privacy sufficient?

McSherry and Talwar [24] observed that differential privacy is sufficient to yield approximate implementation in ϵ-dominant strategies. However, as we show below, differential privacy does not generally imply implementation with a stronger solution concept.

Our example is a pricing mechanism that utilizes the exponential mechanism and hence yields an ϵ-dominant implementation that (assuming parties act truthfully) well approximates the optimal revenue. However, there are dominant strategies in the example that involve mis-representation and lead to a significantly inferior revenue.

Example 1. *Consider a monopolist producing an unlimited supply digital good who faces n buyers, each having a unit demand at a valuation that is either $0.5 + \mu$ or $1 + \mu$ where $0 < \mu < 0.5$. The monopolist cannot distinguish among buyers and is restricted to choosing a price in the set $\{0.5, 1\}$. Assume the monopolist is interested in maximizing the average revenue per buyer.*[10] *The optimal outcome for the auctioneer is hence*

$$OPT(\bar{t}) = \frac{\max_{s \in \{0.5,1\}} (s \cdot |\{i : t_i \geq s\}|)}{n}.$$

If the monopolist uses the appropriate exponential mechanism then it is ϵ-dominant for agents to announce their valuation truthfully, resulting in an almost optimal revenue. However, one should note that the probability that the exponential mechanism will choose the lower of the two prices increases with the number buyers that announce $0.5 + \mu$. Hence, it is **dominant** *for buyers to announce $0.5 + \mu$. This may lead to inferior results. In particular, whenever all agents value the good at $1 + \mu$ but announce $0.5 + \mu$ the mechanism will choose the price 0.5 with high probability, leading to an average revenue of 0.5 per buyer, which is half the optimal revenue per buyer.*

4.2 Is imposition sufficient?

It is tempting to think that our notion of imposition trivializes the result, i.e., that, regardless of the usage of a differentially-private mechanism, the ability to force agents to react sub-optimally, according to their announced types, already inflicts sufficient disutility that would deter untruthful announcements. The next example demonstrates that such a naive imposition is generally insufficient. Intuitively, the reason is that for inducing both truthfulness and efficiency, one needs a strong bound on an agent's benefit from mis-reporting: the utility from mis-reporting should be smaller from the disutility from being committed to a sub-optimal reaction.

Example 2. *Consider a digital goods pricing problem with n agents, where the valuation of each agent is either $\frac{1}{n}$ or $1 + \mu$, and the possible prices are $\frac{1}{n}$ and 1. In this example the optimal price is 1 whenever there exists an agent of type $1 + \mu$, $\mu < 0.5$.*

Consider the following mechanism: with high probability it implements the optimal price and with a low probability it uses an imposing mechanism. Note that the strategy to always announce a valuation of $\frac{1}{n}$ is a Nash equilibrium. This announcement is clearly optimal if an agent's valuation is indeed $\frac{1}{n}$. If an agent's valuation, on the other hand, is $1+\mu$, then complying with this strategy will result in a utility that is almost 1, whereas deviating to truthful announcement will result in a price of 1 with high probability, hence a utility of μ.

Therefore, the monopolist's average revenue from a buyer is always $\frac{1}{n}$. This is substantially inferior to the optimal outcome, which could be as high as 1, whenever all agents are of the high type.

The Nash equilibrium from example 2 survives even if we modify the mechanism to be fully imposing (i.e, it always imposes the optimal reaction). Thus, the above mentioned sub-optimality holds.

We believe that the notion of imposition is natural in many settings, and that to some extent imposition is already *implicitly* integrated into the mechanism design literature.

[10] We consider the average revenue per buyer as the objective function, instead of the total revenue, in order to comply with the requirement that the value of the objective function is restricted to the unit interval.

In fact, any mechanism that is not ex-post individually rational imposes its outcome on the players: it imposes participation and ignores the possibility players have to 'walk away' once the results are known. Moreover, models that involve transfers treat these as imposed reactions: once the social choice and transfers are determined, players must comply (consider taxation and auction payments as an example).

Acknowledgments

We thank Amos Fiat and Haim Kaplan for discussions at an early stage of this research. We thank Frank McSherry and Kunal Talwar for helping to clarify issues related to the constructions in [24]. Finally, we thank Jason Hartline, James Schummer, Roberto Serrano and Asher Wolinsky for their valuable comments.

5. REFERENCES

[1] Dilip Abreu and Hitoshi Matsushima. "Virtual Implementation in Iteratively Undominated Strategies: InComplete Information." Mimeo, Princeton University, 1992.

[2] Dilip Abreu and Arunava Sen. "Subgame perfect implementation: A Necessary and Almost Sufficient Condition." *Journal of Economic Theory*, 50:285-299, 1990.

[3] Nabil Al-Najjar and Rann Smorodinsky. "Pivotal Players and the Characterization of Influence." *Journal of Economic Theory*, Volume 92, 2, 318-342. 2000.

[4] Nabil Al-Najjar and Rann Smorodinsky. "The Efficiency of Competitive Mechanisms under Private Information." *Journal of Economic Theory*, 137:383–403, 2007.

[5] Noga Alon, Michal Feldman, Ariel D. Procaccia, and Moshe Tennenholtz. "Strategyproof Approximation of the Minimax on Networks." *Mathematics of Operations Research*, Volume 35(3): 513–526, 2010.

[6] Moshe Babaioff, Ron Lavi, Elan Pavlov. "Single-value combinatorial auctions and algorithmic implementation in undominated strategies." Journal of the ACM, Volume 56(1): , 2009.

[7] Maria-Florina Balcan, Avrim Blum, Jason D. Hartline, and Yishay Mansour. "Mechanism design via machine learning." In *FOCS*, pages 605–614. IEEE Computer Society, 2005.

[8] Edward H. Clarke. "Multipart pricing of public goods." *Public Choice*, 18:19–33, 1971.

[9] John Duggan. "Virtual Bayesian implementation." *Econometrica*, 65:1175–1199, 1997.

[10] Cynthia Dwork. "Differential privacy." In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *ICALP (2)*, volume 4052 of *Lecture Notes in Computer Science*, pages 1–12. Springer, 2006.

[11] Cynthia Dwork. "The differential privacy frontier (extended abstract)." In Omer Reingold, editor, *TCC*, volume 5444 of *Lecture Notes in Computer Science*, pages 496–502. Springer, 2009.

[12] Cynthia Dwork. "Differential Privacy in New Settings." SODA 2010: 174-183.

[13] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. "Calibrating Noise to Sensitivity in Private Data Analysis." In Shai Halevi and Tal Rabin, editors, *TCC*, volume 3876 of *Lecture Notes in Computer Science*, pages 265–284. Springer, 2006.

[14] Dimitris Fotakis and Christos Tzamos. "Winner-Imposing Strategyproof Mechanisms for Multiple Facility Location Games." *Workshop on Internet and Network Economics - WINE '10*, 2010.

[15] Allan Gibbard. "Manipulation of voting schemes: A General Result." *Econometrica*, 41:587–601, 1973.

[16] Andrew V. Goldberg, Jason D. Hartline, Anna R. Karlin, Michael Saks and Andrew Wright. "Competitive Auctions." *Games and Economic Behavior.* Volume 55(2):242–269, 2006.

[17] Theodore F. Groves. "Incentives in Teams." *Econometrica*, 41:617–631, 1973.

[18] David K. Levine and Wolfgang Pesendorfer. "When Are Agents Negligible?" *The American Economic Review*, Vol. 85(5):1160–1170, 1995.

[19] Pinyan Lu, Xiaorui Sun, Yajun Wang, and Zeyuan Zhu. "Asymptotically Optimal Strategy-Proof Mechanisms for Two-Facility Games." In *ACM Conference on Electronic Commerce*, 2010.

[20] "George J. Mailath, Andrew Postlewaite." Asymmetric Information Bargaining Problems with Many Agents *The Review of Economic Studies*, 57(3):351–367, 1990.

[21] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. *Microeconomic Theory.* Oxford University Press, 1995.

[22] Hitoshi Matsushima. "A New Approach to the Implementation Problem." *Journal of Economic Theory*, 45:128–144, 1988.

[23] Richard McLean, Andrew Postlewaite. "Informational Size and Incentive Compatibility." *Econometrica*, 70(6):2421–2453, 2002.

[24] Frank McSherry and Kunal Talwar. "Mechanism Design via Differential Privacy." In *FOCS*, pages 94–103, 2007.

[25] Herve Moulin. "On Strategy-Proofness and Single-Peakedness." *Public Choice*, 35:437–455, 1980.

[26] Kobbi Nissim, Sofya Raskhodnikova, and Adam Smith. "Smooth Sensitivity and Sampling in Private Data Analysis." In David S. Johnson and Uriel Feige, editors, *STOC*, pages 75–84. ACM, 2007.

[27] Ariel D. Procaccia and Moshe Tennenholtz. "Approximate Mechanism Design without Money." In *ACM Conference on Electronic Commerce*, pages 177–186, 2009.

[28] Kevin Roberts. "The Characterization of Implementable Choice Rules." In Jean-Jacques Laffont, editor, *Aggregation and Revelation of Preferences. Papers presented at the 1st European Summer Workshop of the Econometric Society*, pages 321–349. 1979.

[29] John Roberts and Andrew Postlewaite. "The Incentives for Price-Taking Behavior in Large Exchange Economies." *Econometrica*, 44(1):115–127, 1976.

[30] Mark A. Satterthwaite, Aldo Rustichini and Steven R. Williams. "Convergence to Efficiency in a Simple Market with Incomplete Information." *Econometrica*, 62(1):1041–1063, 1994.

[31] Mark A. Satterthwaite. "Stratey Proofness and Arrow's Conditions: Existence and Correspondence

Theorems for Voting Procedures and Social Welfare Functions." *Journal of Economic Theory*, 10:187–217, 1975.

[32] Mark A. Satterthwaite and Steven R. Williams. "The Rate of Convergence to Efficiency in the Buyer's Bid Double Auction as the Market Becomes Large." *Review of Economic Studies*, 56:477–498, 1989.

[33] James Schummer. "Almost-Dominant Strategy Implementation." *Games and Economic Behavior*, 48(1): 154-170, 2004.

[34] James Schummer and Rakesh V. Vohra. "Strategy-Proof Location on a Network." *Journal of Economic Theory*, 104(2):405–428, 2004.

[35] Janmes Schummer and Rakesh V. Vohra. "Mechanism Design without Money." In N. Nisan, T. Roughgarden, É. Tardos, and V. Vazirani, editors, *Algorithmic Game Theory*, chapter 10. Cambridge University Press, 2007.

[36] Roberto Serrano and Rajiv Vohra. "Some Limitations of Virtual Bayesian Implementation." *Econometrica*, 69:785-792, 2001.

[37] Roberto Serrano and Rajiv Vohra. "A Characterization of Virtual Bayesian Implementation." *Games and Economic Behavior*, 50:312-331, 2005.

[38] Jeroen Swinkels. "Efficiency of Large Private Value Auctions." *Econometrica*, 69(1):37–68, 2001.

[39] William S. Vickrey. "Counterspeculations, Auctions, and Competitive Sealed Tenders." *Journal of Finance*, 16:15–27, 1961.

[40] David Xiao. "Is privacy compatible with truthfulness?" Cryptology ePrint Archive, no. 2011/005, 2011.