

The Exponential Mechanism for Social Welfare: Private, Truthful, and Nearly Optimal

Zhiyi Huang* Sampath Kannan†

January 19, 2012

Abstract

In this paper, we show that for *any* mechanism design problem, the exponential mechanism can be implemented as a *truthful* mechanism while still preserving differential privacy, if the objective is to maximize social welfare. Our instantiation of the exponential mechanism can be interpreted as a generalization of the VCG mechanism in the sense that the VCG mechanism is the extreme case when the privacy parameter goes to infinity. To our knowledge, this is the first general tool for designing mechanisms that are both truthful and differentially private.

*Computer and Information Science, University of Pennsylvania. Email: hziyi@cis.upenn.edu. Supported in part by ONR MURI Grant N000140710907.

†Computer and Information Science, University of Pennsylvania. Email: kannan@cis.upenn.edu. Supported in part by an EAGER grant, NSF CCF 1137084.

1 Introduction

In mechanism design a central entity seeks to allocate resources among a set of selfish agents in order to optimize a specific objective function such as revenue or social welfare. Each agent has a private valuation for the resources being allocated, which is commonly referred to as her *type*. A major challenge in designing mechanisms for problems of resource allocation among selfish agents is getting them to reveal their true types. While in principle mechanisms can be designed to optimize some objective function even when agents are not truthful, the analysis of such mechanisms is complicated and the vast majority of mechanisms are designed to incentivize agents to be truthful.

One reason that an agent might not want to be truthful is that lying gives her a better payoff. Research in algorithmic mechanism design has mostly focused on this possibility and has successfully designed computationally-efficient mechanisms for many problems that are *incentive compatible*, i.e., where each agent achieves optimal payoff by bidding truthfully (See [17] for a survey of results). However, a second reason that an agent might not bid truthfully is that the *privacy* of her type might itself be of value to her. In most traditional mechanism, bidding truthfully almost surely results in an allocation that reveals the private type of the agent.

Consider for example, a matching market in which n oil companies are bidding for n oil fields. Each company may have done extensive research in figuring out their valuations for each field. It may regard this information as giving it competitive advantage and seek to protect the privacy of the information. If it participates in a traditional incentive compatible mechanism, say, the VCG mechanism, it has two choices – 1) bid truthfully, get the optimum payoff but potentially lose information privacy or 2) introduce random noise into its bid to (almost) preserve privacy, but settle for a suboptimal payoff. In this and more generally in multi-agent settings where each agent’s type is multidimensional, we aim to answer the following question:

Can we design mechanisms that simultaneously achieve nearly optimal social welfare, are incentive compatible, and protect the privacy of each agent?

The notion of privacy we will consider is *differential privacy*, which is a paradigm for private data analysis developed in the past decade, aiming to reveal information about the population as a whole, while protecting the privacy of each individual (See surveys [8, 7] and the reference therein). Roughly speaking, a differentially private mechanism is one that behaves almost identically on any two data sets that are almost identical. Here, by behaving almost identically we mean that the probability of any event happening changes by at most a small multiplicative factor. As an important tool in the literature, the exponential mechanism of McSherry and Talwar [16] is a general mechanism that produces differentially private output for a large family of problems. For each problem, a quality value is associated with each possible answer. The exponential mechanism then outputs an answer with probability proportional to the exponent of its quality scaled by the desired differential privacy and the sensitivity of the answer.

Related Works. McSherry and Talwar [16] first proposed using differentially private mechanisms to design auction by pointing out that differential privacy implies approximate incentive compatibility and further resilience to collusion. In particular, they study the problem of revenue maximization in digital auctions and attribute auctions. They propose the exponential mechanism as a solution for these problems. McSherry and Talwar also suggest using the exponential mechanism to solve mechanism design problems with different objective, such as social welfare. Their instantiation of the exponential mechanism is differentially private, but only approximately truthful. Nissim et al. [18] show how to convert differentially private mechanisms into exactly truthful mechanism in some settings. However, the mechanism loses its privacy property after such

conversion. Xiao [21] seeks to design mechanisms that are both differentially private and perfectly truthful and proposes a method to convert any truthful mechanism into a differentially private and truthful one when the type space is small. Unfortunately, it does not seem possible to extend the results in [18, 21] to more general mechanism design problems. Finally, Ghosh and Roth [9] study the problem of selling privacy in auctions, which can be viewed as an orthogonal approach to combining mechanism design and differential privacy.

Our Results and Techniques. Our main contribution is a novel instantiation of the exponential mechanism for *any* mechanism design problem with payments, that aims to maximize social welfare. We show that our version of the exponential mechanism is incentive compatible, individually rational, and has no positive transfer, while preserving differential privacy. In fact, we show that the exponential mechanism can be interpreted as a natural generalization of the VCG mechanism in the sense that the VCG mechanism is the special case when the privacy parameter goes to infinity. To our knowledge, this is the first general tool for designing truthful and differentially private mechanism.

We provide two proofs of the incentive compatibility of the exponential mechanism. The first uses the classical characterization of when an allocation mechanism can be associated with prices to make it incentive-compatible. Rochet [19] showed that this is possible exactly in the case that the mechanism is cyclic monotone. In Section 3, we prove that the exponential mechanism is cyclic monotone and derive the payments according to Rochet’s characterization. We also provide another very different proof in Section 4 by connecting the exponential mechanism to the Gibbs measure and free energy in statistical mechanics. We exploit this connection to provide a simple proof of the incentive compatibility of the mechanism.

While we do not have a computationally efficient way for computing the allocation and prices of the exponential mechanism in general (this is also not known for VCG), we do show that in special cases such as multi-item auctions and procurement auctions for spanning tree, we can efficiently implement the exponential mechanism either exactly or approximately. Further, we show that the trade-off between privacy and social welfare in the exponential mechanism is asymptotically optimal in these two cases, even if we compare to mechanisms that need not be truthful.

Interestingly, our implementation of the exponential mechanism for the multi-item auction has further implications in the recent work on blackbox reductions in Bayesian mechanism design [10, 3]. Combining our exponential mechanism for the matching market with the blackbox reduction procedure in [10, 3], we can get a blackbox reduction that converts any algorithm into BIC, differentially private mechanisms without hurting the social welfare too much. We will leave further discussions to the related section.

2 Preliminaries

Model. A mechanism design problem is defined by a set of n agents and a range R of feasible outcomes. Each agent i has a private valuation $v_i : R \mapsto [0, 1]$. A central principal chooses one of the outcomes based on the agents’ valuations. We will let $\mathbf{0}$ denote the all-zero valuation and let v_{-i} denote the valuations of every agent except i .

A *mechanism* M consists of an allocation rule $x(\cdot)$ and a payment rule $p(\cdot)$. The mechanism first lets the agents submit their valuations. However, an agent may strategically submit a fake valuation if that is beneficial to her. We will let $b_1, \dots, b_n : R \mapsto [0, 1]$ denote the *reported valuations* from the agents and let \mathbf{b} denote the vector of these valuations. After the agents submit their bids, the allocation rule $x(\cdot)$ chooses a feasible outcome $r = x(\mathbf{b}) \in R$ and the payment rule $p(\cdot)$ chooses a vector of payments $p(\mathbf{b}) \in \mathbb{R}^n$. We will let $p_i(\mathbf{b})$ denote the payment for agent i . Note that both

$x(\cdot)$ and $p(\cdot)$ may be randomized. We will consider the standard setting of quasi-linear utility: given the allocation rule, the payment rule, and the reported valuations \mathbf{b} , for each $i \in [n]$, the utility of agent i is $u_i(v_i, x(\mathbf{b}), p_i(\mathbf{b})) = v_i(x(\mathbf{b})) - p_i(\mathbf{b})$.

The goal is to design polynomial time mechanisms M that satisfy various objectives. In this paper, we will focus on the problem of maximizing the expected *social welfare*, which is defined to be the sum of the agents' valuations: $\mathbf{E}[\sum_{i=1}^n v_i(x(\mathbf{b}))]$.

Besides the expected social welfare, we take into consideration the strategic play of utility-maximizing agents and their concern about the mechanism leaking non-trivial information about their private data. Thus, we will restrict our attention to mechanisms that satisfy several game-theoretic requirements and have a privacy guarantee that we will define in the rest of this section.

Game-Theoretical Solution Concepts. A mechanism is *incentive compatible* (IC) if truth-telling is a dominant strategy, that is, by reporting the true values an agent always maximizes her expected utility regardless of what other agents do, that is, $v_i \in \arg \max_{b_i} \mathbf{E}[v_i(x(b_i, b_{-i})) - p_i(b_i, b_{-i})]$. We will also consider an approximate notion of truthfulness. A mechanism is γ -incentive compatible (γ -IC) if no agent could get more than γ extra utility by lying. Further, a mechanism is *individually rational* (IR) if the expected utility of each agent is always non-negative, assuming this agent reports truthfully: $\mathbf{E}[v_i(x(v_i, b_{-i})) - p_i(v_i, b_{-i})] \geq 0$. Finally, a mechanism has *no positive transfer* if the payments are always non-negative: $\forall b_1, \dots, b_n, \forall i \in [n], p(\mathbf{b})_i \geq 0$. We seek to design mechanisms that are incentive compatible, individually rational, and without positive transfer.

An allocation rule $x(\cdot)$ is *rationalizable* if there exists a payment rule $p(\cdot)$, such that (x, p) is an IC mechanism. In his seminal work, Rochet [19] gave a characterization of rationalizable rules.

Theorem 2.1 (Rochet's Characterization [19]). *An allocation rule $x(\cdot)$ is rationalizable if and only if it is cyclically monotone: for any agent i , any valuation profile v_{-i} of the other agents any $t \in \mathbb{N}$, and any sequence of possible valuations v_i^1, \dots, v_i^t of agent i ,*

$$\sum_{k=1}^t \mathbf{E}[v_i^k(x(v_i^k, v_{-i}))] \geq \sum_{k=1}^t \mathbf{E}[v_i^{k+1}(x(v_i^k, v_{-i}))] .$$

Moreover, the payment rule of a cyclically monotone allocation rule $x(\cdot)$ can be computed as

$$p_i(v_i, v_{-i}) = \mathbf{E}[v_i(x(\mathbf{v}))] - \sup_{\substack{\text{all chains } (v_i^0=0, \\ v_i^1, \dots, v_i^t, v_i^{t+1}=v_i)}} \sum_{k=0}^t \left(\mathbf{E}[v_i^{k+1}(x(v_i^k, v_{-i}))] - \mathbf{E}[v_i^k(x(v_i^k, v_{-i}))] \right) .$$

Differential Privacy and Approximate Differential Privacy. Differential privacy is a notion of privacy that has received much attention in the past decade. It requires the distribution of outcomes to be nearly identical when the agent profiles are nearly identical. Formally,

Definition 1. *A mechanism is ϵ -differentially private if for any two valuation profiles $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{v}' = (v'_1, \dots, v'_n)$ such that only one agent has different valuations in the two profiles, and for any set of outcomes $S \subseteq R$, we have $\Pr[x(\mathbf{v}) \in S] \leq \exp(\epsilon) \cdot \Pr[x(\mathbf{v}') \in S]$.*

This definition of privacy has many appealing theoretical properties. The readers are referred to [8, 7] for excellent surveys on the subject.

Note that in this definition we are implicitly assuming that the adversary can only observe the chosen outcome $x(\cdot)$, but not the payments. We want to stress that this assumption is w.l.o.g. for, by adding arbitrary noise with zero mean we can obtain a payment scheme that is almost perfectly private without affecting our objective or any of the game-theoretic requirements.

We will also consider a standard variant that defines a more relaxed notion of privacy.

1. Choose outcome $r \in R$ with probability $\Pr[r] \propto \exp\left(\frac{\epsilon}{2} \sum_i v_i(r)\right)$.
2. For $1 \leq i \leq n$, charge agent i price

$$p_i = \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(\mathbf{v})} [v_i(r)] - \frac{2}{\epsilon} \ln \left(\sum_{r \in R} \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r) \right) \right) + \frac{2}{\epsilon} \ln \left(\sum_{r \in R} \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r) \right) \right) .$$

Figure 1: EXP_ϵ^R : the incentive compatible instantiation of the exponential mechanism.

Definition 2. A mechanism is (ϵ, δ) -differentially private if for any two valuation profile $\mathbf{v} = (v_1, \dots, v_n)$ and $\mathbf{v}' = (v'_1, \dots, v'_n)$ such that only one agent has different valuations in the two profiles, and for any set of outcomes $S \subseteq R$, $\Pr[x(\mathbf{v}) \in S] \leq \exp(\epsilon) \cdot \Pr[x(\mathbf{v}') \in S] + \delta$.

Typically, we will consider very small values of δ , say, $\delta = \exp(-n)$. This relaxed notion of differential privacy states that the probability of some event may be sensitive to the change of a single agent's valuation, but that could only happen for very low probability events.

The Exponential Mechanism. One particularly useful tool in the differential privacy literature is the exponential mechanism of McSherry and Talwar [16]. The exponential mechanism is a general tool for constructing differentially private algorithms over an arbitrary range R of outcomes and any objective function $Q(D, r)$ (often referred to as the quality function in the differential privacy literature) that maps a pair consisting of a data set D and a feasible outcome $r \in R$ to a real-valued score. In our setting, D is a valuation profile and the quality function $Q(\mathbf{v}, r) = \sum_{i=1}^n v_i(r)$ is the social welfare.

Given a range R , a data set D , a quality function Q , and a privacy parameter ϵ , the *exponential mechanism* $\text{EXP}(R, D, Q, \epsilon)$ chooses an outcome r from the range R with probability

$$\Pr [\text{EXP}(R, D, Q, \epsilon) = r] \propto \exp \left(\frac{\epsilon}{2\Delta} Q(D, r) \right) ,$$

where Δ is the Lipschitz constant of the quality function Q , that is, for any two adjacent data set D_1 and D_2 , and for any outcome r , the score $Q(D_1, r)$ and $Q(D_2, r)$ differs by at most Δ . In our setting, the Lipschitz constant of the social welfare function is 1. We will use the following theorem of the exponential mechanism. Readers are referred to [16, 20] for the proof of this theorem.

Theorem 2.2. *The exponential mechanism is ϵ -differentially private and ensures that*

$$\Pr \left[Q(D, \text{EXP}(R, D, Q, \epsilon)) < \max_{r \in R} Q(D, r) - \frac{\ln |R|}{\epsilon} - \frac{t}{\epsilon} \right] \leq \exp(-t) .$$

3 The Exponential Mechanism is Incentive Compatible

In this section, we will show that if we choose the social welfare to be the quality function, then the exponential mechanism can be implemented in a truthful-in-expectation, individually rational, and no-positive-transfer manner. Formally, for any range R and any privacy parameter $\epsilon > 0$, our instantiation of the exponential mechanism EXP_ϵ^R with its pricing scheme is presented in Figure 1.

Theorem 3.1. *The exponential mechanism with our pricing scheme is incentive compatible, individually rational, and has no positive transfer.*

We will give two proofs for the truthfulness of the mechanism. The first proof goes over the procedure developed by Rochet [19]: we will start with a proof of cyclic monotonicity of the exponential allocation rule, which is known to be the necessary and sufficient condition for being the allocation rule of a truthful mechanism (Section 3.1); then we will derive the pricing scheme that rationalizes the exponential allocation rule via Rochet's characterization (Section 3.2). The second proof is a short proof via an interesting connection between the exponential mechanism and the Gibbs measure. We will defer the discussion of the second proof to Section 4. The proof of individual rationality and that EXP_ϵ^R has no positive transfer is deferred to Appendix A.

3.1 The Exponential Allocation Rule is Cyclically Monotone

We now show that the exponential allocation rule satisfies the cyclically monotone condition in Rochet's characterization for rationalizable allocation.

Lemma 3.2. *The exponential allocation rule is cyclically monotone.*

Proof. By symmetry, it suffices to show cyclic monotonicity for a particular agent i . Fix the valuation of the other agents to be v_{-i} . Consider any $t > 0$ and any valuation profiles v_i^1, \dots, v_i^t . For notational convenience, we will let $v_i^{t+1} = v_i^1$ and $v_{-i}(r) \stackrel{\text{def}}{=} \sum_{j \neq i} v_j(r)$. To show cyclic monotonicity is to prove the following inequality:

$$\sum_{k=1}^t \sum_{r \in R} \frac{\exp(\frac{\epsilon}{2} v_i^k(r) + \frac{\epsilon}{2} v_{-i}(r))}{\sum_{r' \in R} \exp(\frac{\epsilon}{2} v_i^k(r') + \frac{\epsilon}{2} v_{-i}(r'))} v_i^k(r) \geq \sum_{k=1}^t \sum_{r \in R} \frac{\exp(\frac{\epsilon}{2} v_i^k(r) + \frac{\epsilon}{2} v_{-i}(r))}{\sum_{r' \in R} \exp(\frac{\epsilon}{2} v_i^k(r') + \frac{\epsilon}{2} v_{-i}(r'))} v_i^{k+1}(r) . \quad (1)$$

We first note that we can assume $\frac{\epsilon}{2} = 1$ w.l.o.g. by rescaling the valuations v_i^1, \dots, v_i^t to simplify notation. Consider t distributions D_1, \dots, D_t over range R where the probability mass function of the k^{th} distribution is given by $\mathbf{Pr}_k[r] \propto \exp(v_i^k(r) + v_{-i}(r))$. By the non-negativity of KL-divergence, we have for all $k \in [t]$,

$$D_{\text{KL}}(D_k || D_{k+1}) = \sum_{r \in R} \mathbf{Pr}_k[r] \ln(\mathbf{Pr}_k[r]) - \sum_{r \in R} \mathbf{Pr}_k[r] \ln(\mathbf{Pr}_{k+1}[r]) \geq 0 .$$

Summing up this inequality for all $k \in [t]$, we have

$$\sum_{k=1}^t \sum_{r \in R} \mathbf{Pr}_k[r] \ln(\mathbf{Pr}_k[r]) \geq \sum_{k=1}^t \sum_{r \in R} \mathbf{Pr}_k[r] \ln(\mathbf{Pr}_{k+1}[r]) . \quad (2)$$

By the definition of $\mathbf{Pr}_k[r]$, the left-hand-side of (2) is

$$\sum_{k=1}^t \left(\sum_{r \in R} \mathbf{Pr}_k[r] (v_i^k(r) + v_{-i}(r)) - \ln \left(\sum_{r' \in R} \exp(v_i^k(r') + v_{-i}(r')) \right) \right) .$$

Similarly, the right-hand-side of (2) can be transformed into

$$\sum_{k=1}^t \left(\sum_{r \in R} \mathbf{Pr}_k[r] (v_i^{k+1}(r) + v_{-i}(r)) - \ln \left(\sum_{r' \in R} \exp(v_i^{k+1}(r') + v_{-i}(r')) \right) \right) .$$

Comparing these two formulas, note that $\sum_{k=1}^t \sum_{r \in R} \mathbf{Pr}_k[r] v_{-i}(r)$ is a common term on both sides. Moreover, $\sum_{k=1}^t \ln(\sum_{r' \in R} \exp(v_i^k(r') + v_{-i}(r')))$ and $\sum_{k=1}^t \ln(\sum_{r' \in R} \exp(v_i^{k+1}(r') + v_{-i}(r')))$ are also common terms on both sides as well. Therefore, by canceling the common terms, we can deduce from (2) that $\sum_{k=1}^t \sum_{r \in R} \mathbf{Pr}_k[r] v_i^k(r) \geq \sum_{k=1}^t \sum_{r \in R} \mathbf{Pr}_k[r] v_i^{k+1}(r)$, which is exactly (1) (recall we have assumed $\frac{\epsilon}{2} = 1$). So we have proved Lemma 3.2. \square

3.2 Deriving the Pricing Scheme

In this section, we will use Rochet's characterization to compute the prices that rationalize the exponential mechanism. First we define

$$\phi_i(v_i, v_{-i}) \stackrel{\text{def}}{=} \sup_{\substack{\text{all chains } (v_i^0 = \mathbf{0}, \\ v_i^1, \dots, v_i^t, v_i^{t+1} = v_i)} } \sum_{k=0}^t \left(\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^{k+1}(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^k(r)] \right). \quad (3)$$

Rochet's characterization asserts the prices $p_i(\mathbf{v}) = \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i, v_{-i})} [v_i(r)] - \phi_i(v_i, v_{-i})$, $i \in [n]$, rationalize the exponential allocation rule. So it remains to derive a close form of $\phi_i(v_i, v_{-i})$.

Lemma 3.3. *Suppose $\mu(\mathbf{v}) \stackrel{\text{def}}{=} \frac{2}{\epsilon} \ln \left(\sum_{r \in R} \exp \left(\frac{\epsilon}{2} \sum_{j=1}^n v_j(r) \right) \right)$ for any value profile $\mathbf{v} = (v_1, \dots, v_n)$. Then we have $\phi_i(v_i, v_{-i}) = \mu(v_i, v_{-i}) - \mu(\mathbf{0}, v_{-i})$.*

Proof. We first observe that if we view each the valuation of agent i as a $|R|$ -dimensional vector, then the probability that EXP_ϵ^R choose outcome $r \in R$ is

$$\frac{\exp \left(\frac{\epsilon}{2} \sum_{j=1}^n v_j(r) \right)}{\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} \sum_{j=1}^n v_j(r') \right)} = \frac{\partial \mu}{\partial v_i(r)}.$$

Hence, if we fixed the valuations v_j for $j \neq i$ and view μ as a function v_i , then for any v_i and v'_i

$$\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i, v_{-i})} [v'_i(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i, v_{-i})} [v_i(r)] = \nabla \mu(v_i, v_{-i}) \cdot (v'_i - v_i).$$

Here $\nabla \mu(v_i, v_{-i}) \cdot (v'_i - v_i)$ denote the dot product of the gradient of $\mu(\cdot)$ and the $|R|$ -dimensional vector $v'_i - v_i$. If we choose $v_i^k = \frac{k}{t+1} \mathbf{0} + \frac{t-k+1}{t+1} v_i$ for $k = 1, \dots, t$, and let P be the straight path from $\mathbf{0}$ to v_i , then we have

$$\begin{aligned} \phi_i(v_i, v_{-i}) &\geq \lim_{t \rightarrow \infty} \sum_{k=0}^t \left(\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^{k+1}(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^k(r)] \right) \\ &= \lim_{t \rightarrow \infty} \sum_{k=0}^t \nabla \mu(v_i^k, v_{-i}) \cdot (v_i^{k+1} - v_i^k) = \int_P \nabla \mu(u_i, v_{-i}) \cdot du_i = \mu(v_i, v_{-i}) - \mu(\mathbf{0}, v_{-i}). \end{aligned}$$

Next, we will show the reverse direction: $\mu(v_i, v_{-i}) - \mu(\mathbf{0}, v_{-i}) \geq \phi_i(v_i, v_{-i})$. In order to do so, it suffices to show that for any choice $v_i^0 = \mathbf{0}, v_i^1, \dots, v_i^t, v_i^{t+1} = v_i$, we have

$$\mu(v_i, v_{-i}) - \mu(\mathbf{0}, v_{-i}) \geq \sum_{k=0}^t \left(\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^{k+1}(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(v_i^k, v_{-i})} [v_i^k(r)] \right).$$

Note that $\mu(v_i, v_{-i}) - \mu(\mathbf{0}, v_{-i}) = \sum_{k=0}^t \left(\mu(v_i^{k+1}, v_{-i}) - \mu(v_i^k, v_{-i}) \right)$. So it suffices to show that

$$\begin{aligned} \mu(v_i^{k+1}, v_{-i}) - \mu(v_i^k, v_{-i}) &\geq \mathbf{E}_{r \sim \text{EXP}_\epsilon(v_i^k, v_{-i})} [v_i^{k+1}(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon(v_i^k, v_{-i})} [v_i^k(r)] \\ &= \mathbf{E}_{r \sim \text{EXP}_\epsilon(v_i^k, v_{-i})} [v_i^{k+1}(r) + v_{-i}(r)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon(v_i^k, v_{-i})} [v_i^k(r) + v_{-i}(r)] \quad (4) \end{aligned}$$

We let $v^k(r) = v_i^k(r) + v_{-i}(r)$ for notational convenience, and let D_k and D_{k+1} denote the distribution given by probability mass function $\mathbf{Pr}_k[r] \propto \exp(\frac{\epsilon}{2}v^k(r))$ and $\mathbf{Pr}_{k+1}[r] \propto \exp(\frac{\epsilon}{2}v^{k+1}(r))$ respectively. By the definition of $\mu(\cdot)$ and EXP_ϵ^R , we can transform (4) into the following

$$\frac{2}{\epsilon} \ln \left(\frac{\epsilon}{2} \sum_{r \in R} \exp(v^{k+1}(r)) \right) - \frac{2}{\epsilon} \ln \left(\frac{\epsilon}{2} \sum_{r \in R} \exp(v^k(r)) \right) \geq \sum_{r \in R} \mathbf{Pr}_k[r] v^{k+1}(r) - \sum_{r \in R} \mathbf{Pr}_k[r] v^k(r) .$$

Finally, we conclude that this is equivalent to the KL-divergence $D_{KL}(D_k||D_{k+1})$ being non-negative because we have $\sum_{r \in R} \mathbf{Pr}_k[r] v^k(r) - \ln \left(\sum_{r' \in R} \exp(v^k(r')) \right) = \sum_{r \in R} \mathbf{Pr}_k[r] \ln \mathbf{Pr}_k[r]$ and $\sum_{r \in R} \mathbf{Pr}_k[r] v^{k+1}(r) - \ln \left(\sum_{r' \in R} \exp(v^{k+1}(r')) \right) = \sum_{r \in R} \mathbf{Pr}_k[r] \ln \mathbf{Pr}_{k+1}[r]$. \square

4 A Direct Proof via Connection to the Gibbs Measure

In this section, we will present a direct proof of truthfulness of the exponential mechanisms via an interesting connection to the Gibbs measure and free energy. We believe this intriguing connection is of independent interest and may lead to new ways of understanding the exponential mechanism and differential privacy.

Gibbs measure is a probability measure widely used in probability and statistical mechanics. In chemistry and physics, it is also known as the Boltzmann distribution. Formally,

Definition 3 (Gibbs measure). *Suppose we have a system consisting of particles of a gas. If the particles have k states $1, \dots, k$, possessing energy E_1, \dots, E_k respectively, then the probability that a random particle in the system has state i follows the Gibbs measure: $\mathbf{Pr}[\text{state} = i] \propto \exp\left(-\frac{1}{k_B T} E_i\right)$, where T is the temperature, and k_B is the Boltzmann constant.*

Note that the Gibbs measure asserts that nature prefers states with lower energy level. Indeed, if $T \rightarrow 0$, then almost surely we will see a particle with lowest-energy state. On the other hand, if $T \rightarrow +\infty$, then all states are equally likely to appear. In this sense, the temperature T is a measure of uncertainty in the system: the lower the temperature is the less uncertainty we have in the system, and vice versa.

Gibbs Measure vs. Exponential Mechanism. It is not difficult to see a connection between the Gibbs measure and the exponential mechanism. Firstly, the quality $Q(r)$ of an outcome $r \in R$ (in our instantiation, $Q(r)$ is the social welfare $\sum_i v_i(r)$) is an analogue of the energy (more precisely, the inverse of the energy) of a state i . In the exponential mechanism the goal is to maximize the expected quality of the outcome, while in physics nature tries to minimize the expected energy. Second, the privacy parameter ϵ is an analogue of the inverse temperature T^{-1} , both measuring the level of uncertainties in the system. The more privacy we want in the mechanism, the more uncertainty we need to impose in the distribution of outcomes¹. Finally, the Lipschitz constant Δ and Boltzmann constant k_B are both scaling factors that come from the environment. Table 1 summarize this connection between the Gibbs measure and the exponential mechanism.

Gibbs Measure Minimizes Free Energy. It is well-known that the Gibbs measure maximizes entropy given the expected energy. In fact, a slightly stronger claim (e.g. see [15]) states that the

¹We note that the privacy guarantee ϵ is not necessarily a monotone function of the entropy of the outcome distribution. So the statement above is only for the purpose of establishing a high-level connection between the Gibbs measure and the exponential mechanism.

Table 1: A high-level comparison between the Gibbs measure and the exponential mechanism

	Gibbs measure	Exponential mechanism
Probability mass function	$\Pr[\text{state} = i] \propto \exp\left(-\frac{1}{k_B T} E_i\right)$	$\Pr[\text{outcome} = r] \propto \exp\left(\frac{\epsilon}{2\Delta} Q(r)\right)$
Objective function	$-E_i$	$Q(r)$
Measure of uncertainty	temperature T	privacy parameter ϵ
Environment parameter	Boltzmann constant k_B	Lipschitz constant Δ

Gibbs measure minimizes free energy. Precisely, suppose T is the temperature, D is a distribution over the states, and $S(D)$ is the Shannon entropy of D . Then the *free energy* of the system is

$$F(D, T) = \mathbf{E}_{i \sim D} [E_i] - k_B T \cdot S(D) .$$

Moreover, the free energy is minimized when D is the Gibbs measure. The proof of this claim is not difficult (e.g. see [15]). We will omit the details in this extended abstract.

By the connection between Gibbs measure and exponential mechanism, we have the following analogue for our instantiation of the exponential mechanism.

Lemma 4.1. $\mathbf{E}_{r \sim D} [\sum_i v_i(r)] + \frac{2}{\epsilon} \cdot S(D)$ is maximized when $D = \text{EXP}_\epsilon^R(v_1, \dots, v_n)$.

Equipped with this lemma, we are ready to present a direct proof of the incentive compatibility of the exponential mechanism. We will prove that EXP_ϵ^R is incentive compatible by showing that our pricing scheme encodes the inverse of the “free energy” into the agents’ utilities. As a result, when the agents report their value truthfully, the exponential mechanism will choose an outcome distribution that minimizes the “free energy”, and thus maximizes the agents’ utilities.

Proof of the incentive compatibility of EXP_ϵ^R . Let us consider a particular agent i , and fix the bids b_{-i} of the other agents. Suppose agent i has value v_i and bids b_i . We will let $h_i(b_{-i})$ denote the function $\frac{2}{\epsilon} \ln(\sum_{r \in R} \exp(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r)))$, $b(r) = \sum_{k=1}^n b_k(r)$. We have

Lemma 4.2. *The payment scheme of the exponential mechanism EXP_ϵ^R can be written as for all $i \in [n]$, $p_i = -\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [\sum_{k \neq i} b_k(r)] - \frac{2}{\epsilon} \cdot S(\text{EXP}_\epsilon^R(b_i, b_{-i})) + h_i(b_{-i})$.*

The proof of Lemma 4.2 follows easily from the definition of p_i so we defer the tedious calculation to Appendix B. By Lemma 4.2 we get that agent i ’s utility is $\mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [v_i(r) + \sum_{k \neq i} b_k(r)] + \frac{2}{\epsilon} \cdot S(\text{EXP}_\epsilon^R(b_i, b_{-i})) - h_i(b_{-i})$. By Lemma 4.1 this quantity is maximized when $\text{EXP}_\epsilon^R(b_i, b_{-i}) = \text{EXP}_\epsilon^R(v_i, b_{-i})$. So truthful bidding is a utility-maximizing strategy for agent i . \square

5 Applications

Our result in Theorem 3.1 applies to a large family of problems. In fact, it can be used to derive truthful and differentially private mechanisms for any problem in mechanism design (with payments) that aims for social welfare maximization.

In this section, we will consider two examples – the multi-item auction and the procurement auction for a spanning tree. We will analyze the computational efficiency issue, including how to efficiently choose an outcome from the desired distribution and how to efficiently generate the payment. We will also consider the trade-off between social welfare and privacy of our instantiation of the exponential mechanism.

5.1 Multi-Item Auction and Implication in BIC Blackbox Reduction

The first application we will consider is the multi-item auction. In the multi-item auction, the auctioneer has n heterogeneous items (one copy of each item) that she wishes to allocate to n different agents². Each agent i has a private valuation $\mathbf{v}_i = (v_{i1}, \dots, v_{ik})$, where v_{ij} is agent i 's value for item j . We will assume the agents are unit-demand, that is, each agent wants at most one item. It is easy to see that each feasible allocation of the multi-item auction is a matching between agents and items. We will let the R_M denote the range of multi-item auction, that is, the set Π_n of all permutations on $[n]$.

The multi-item auction and related problems are very well-studied in the algorithmic game theory literature (e.g. [6, 4]). It captures the motivating scenario of allocating oil fields and many other problems that arise from allocating public resources. The VCG mechanism can be implemented in polynomial time to maximize social welfare in this problem since max-matching can be solved in polynomial time. The new twist in our setting is to design mechanisms that are *both truthful and differentially private* and have good social welfare guarantee.

Approximate Implementation of the Exponential Mechanism. Unfortunately, exactly sampling matchings according to the distribution specified in the exponential mechanism seems hard due to its connection to the problem of computing the permanent of non-negative matrices (e.g. see [11]), which is $\#P$ -complete. Instead, we will sample from the desired distribution approximately. Moreover, we show that there is an efficient approximate implementation of the payment scheme. As a result of the non-exact implementation, we only get γ -IC instead of perfect IC, (ϵ, γ) -differential privacy instead of ϵ -differential privacy, and loses an additional $n\gamma$ additive factor in social welfare. Here, γ will be inverse polynomially small. The discussion of this approximate implementation of the exponential mechanism is deferred to [Appendix C](#).

Note that the size of the range of feasible outcomes of multi-item auction is the number of different matching between n agents and n items, which equals $n!$. By [Theorem 2.2](#) we have:

Theorem 5.1. *For any $\delta \in (0, 1)$, $\epsilon > 0$, $\gamma > 0$, there is a polynomial time (in n , ϵ^{-1} , γ^{-1} , and $\log(\delta^{-1})$) approximate implementation of the exponential mechanism, $\widehat{\text{EXP}}_\epsilon^{R_M}$ that is γ -IC, (ϵ, δ) -differentially private, and ensures that*

$$\Pr \left[\sum_{i=1}^n v_i \left(\widehat{\text{EXP}}_\epsilon^{R_M} \right) < \text{opt} - \gamma n - \frac{\ln(n!)}{\epsilon} - \frac{t}{\epsilon} \right] \leq \exp(-t) .$$

The trade-off between privacy and social welfare in [Theorem 5.1](#) can be interpreted as the following: if we want to achieve social welfare that is worse than optimal by at most an $O(n)$ additive term, then we need to choose $\epsilon = \Omega(\log n)$. We note that this is tight. The proof of the next theorem is deferred to [Section D](#).

Theorem 5.2. *Suppose M is an ϵ -differentially private mechanism for the multi-item auction problem and the expected welfare achieve by M is at least $\text{opt} - \frac{n}{10}$. Then $\epsilon = \Omega(\log n)$.*

Note that in this above theorem, we do not restrict M to be incentive compatible. In other word, this lower bound holds for arbitrary differentially private mechanisms. So there is no extra cost for imposing the truthfulness constraint.

Implication in BIC Blackbox Reduction. Recently, Hartline et al. [10] and Bei and Huang [3] introduce blackbox reductions that convert any algorithm into nearly Bayesian incentive compatible

²The case when the number of items is not the same as the number of agents can be reduced to this case by adding dummy items or dummy agents. So our setting is w.l.o.g.

mechanisms with only a marginal loss in the social welfare. Both approach essentially create a virtual interface for each agent which has the structure of a matching market and then run VCG in the virtual matching markets. By running the exponential mechanism instead of the VCG mechanism, we can obtain a blackbox reduction that converts any algorithm into a nearly Bayesian incentive compatible and differentially private mechanism. We will defer more details to the full version of this paper.

5.2 Procurement Auction for Spanning Trees

Another interesting application is the procurement auction for spanning tree (e.g. see [5]). In this problem, $n = \binom{k}{2}$ selfish agents own edges in a publicly known network of k nodes. We shall imagine the nodes as cities and each edge as a potential highway connecting the cities at its two endpoints. Each agent i has a non-negative cost c_i for building a highway along the corresponding edge. The principal wants to purchase a spanning tree from the network so that she can build highways to connect the cities. The goal is to design incentive compatible and differentially private mechanisms that provide good social welfare (minimizing total cost).

Although this is a reverse auction in which agents have costs instead of having values and the payments are from the principal to the agents, by interpreting the costs as the inverse of the valuations (i.e. $v_i = -c_i$ if the edge is purchased and $v_i = 0$ otherwise), we can show that our instantiation of the exponential mechanism with the same payment scheme is truthful for this problem via almost identical proofs. We will omit the details in this extended abstract.

Next, we will discuss how to efficiently implement the exponential mechanism.

Sampling Spanning Trees. There has been a large body of literature on sampling spanning tree (e.g. see [14] and the reference therein). Recently, Asadpour et al. [1] develop a polynomial time algorithm for sampling *entropy-maximizing* distributions, which is exactly the distribution used by the exponential mechanism. Therefore, the allocation rule of the exponential mechanism can be implemented in polynomial time for spanning tree auction.

Implicit Payment Scheme by Babaioff, Kleinberg, and Slivkins [2]. Although we can efficiently generate samples from the desired distribution, it is not clear how to compute the exact payment explicitly. Fortunately, Babaioff et al. [2, 13] provide a general method of computing an unbiased estimator for the payment given any rationalizable allocation rule³. Hence, we can use the implicit payment method in [2, 13] to generate the payments in polynomial time.

Note that the size of the range of feasible outcomes of spanning tree auction is the number of different spanning tree in a complete graph with k vertices, which equals k^{k-2} . By [Theorem 2.2](#) we have the following:

Theorem 5.3. *For any $\epsilon > 0$, the exponential mechanism $\text{EXP}_\epsilon^{\text{tree}}$ runs in polynomial time, is IC, ϵ -differentially private, and ensures that $\Pr \left[\sum_{i=1}^n c_i \left(\widehat{\text{EXP}}_\epsilon^{\text{tree}} \right) > \text{opt} + \frac{(k-2)\log k}{\epsilon} + \frac{t}{\epsilon} \right] \leq \exp(-t)$.*

This trade-off between privacy and social welfare in [Theorem 5.3](#) essentially means that we need $\epsilon = \Omega(\log k)$ in order to get $\text{opt} + O(k)$ guarantee on expected total cost. This trade-off is also tight. The proof of the next theorem is deferred to [Appendix E](#).

Theorem 5.4. *Suppose M is an ϵ -differentially private mechanism for the procurement auction for spanning tree and the expected total cost by M is at most $\text{opt} + \frac{k}{24}$. Then $\epsilon = \Omega(\log k)$.*

³Although the result in [2] only applies to single-parameter problems, Kleinberg [13] pointed out the same approach can be extended to multi-parameter problems if the type space is convex.

Acknowledgement

The authors would like to thank Aaron Roth for many useful comments and helpful discussions.

References

- [1] A. Asadpour, M.X. Goemans, A. Madry, S.O. Gharan, and A. Saberi. An $O(\log n / \log \log n)$ -approximation algorithm for the asymmetric traveling salesman problem. In *SODA*, pages 379–389. ACM-SIAM, 2010. [10](#)
- [2] M. Babaioff, R.D. Kleinberg, and A. Slivkins. Truthful mechanisms with implicit payment computation. In *EC*, pages 43–52. ACM, 2010. [10](#)
- [3] X. Bei and Z. Huang. Bayesian incentive compatibility via fractional assignments. In *SODA*. ACM-SIAM, 2011. [2](#), [9](#)
- [4] S. Bhattacharya, G. Goel, S. Gollapudi, and K. Munagala. Budget constrained auctions with heterogeneous items. In *STOC*, pages 379–388. ACM, 2010. [9](#)
- [5] M.C. Cary, A.D. Flaxman, J.D. Hartline, and A.R. Karlin. Auctions for structured procurement. In *SODA*, pages 304–313. ACM-SIAM, 2008. [10](#)
- [6] S. Chawla, J. Hartline, D. Malec, and B. Sivan. Sequential posted pricing and multi-parameter mechanism design. In *STOC*, pages 311–320. ACM, 2010. [9](#)
- [7] C. Dwork. A firm foundation for private data analysis. *Communications of the ACM*, to appear. [1](#), [3](#)
- [8] C. Dwork. Differential privacy: A survey of results. In *TAMC*, pages 1–19, 2008. [1](#), [3](#)
- [9] A. Ghosh and A. Roth. Selling privacy at auction. In *EC*, pages 199–208. ACM, 2011. [2](#)
- [10] J. Hartline, R. Kleinberg, and A. Malekian. Bayesian incentive compatibility via matchings. In *SODA*. ACM-SIAM, 2011. [2](#), [9](#)
- [11] M. Jerrum and A. Sinclair. Approximating the permanent. *SIAM Journal on Computing*, 18:1149, 1989. [9](#)
- [12] M. Jerrum, A. Sinclair, and E. Vigoda. A polynomial-time approximation algorithm for the permanent of a matrix with nonnegative entries. *Journal of the ACM (JACM)*, 51(4):671–697, 2004. [13](#)
- [13] R.D. Kleinberg. Personal communication. [10](#)
- [14] V.G. Kulkarni. Generating random combinatorial objects. *Journal of Algorithms*, 11(2):185–207, 1990. [10](#)
- [15] A. Le Ny. Introduction to (generalized) Gibbs measures. *Ensaïos Matemáticos*, 15:1–126, 2008. [7](#), [8](#)
- [16] F. McSherry and K. Talwar. Mechanism design via differential privacy. In *FOCS*, 2007. [1](#), [4](#)
- [17] N. Nisan, T. Roughgarden, E. Tardos, and V.V. Vazirani. *Algorithmic game theory*. Cambridge University Press, 2007. [1](#)

- [18] K. Nissim, R. Smorodinsky, and M. Tennenholtz. Approximately optimal mechanism design via differential privacy. *ITCS, to appear*, 2012. 1, 2
- [19] J.C. Rochet. A necessary and sufficient condition for rationalizability in a quasi-linear context. *Journal of Mathematical Economics*, 16(2):191–200, 1987. 2, 3, 5
- [20] K. Talwar, A. Gupta, K. Ligett, F. McSherry, and A. Roth. Differentially private combinatorial optimization. In *SODA*. ACM-SIAM, 2010. 4
- [21] D. Xiao. Is privacy compatible with truthfulness? In *Cryptology ePrint Technical Report, 2011/005*, 2011. 2

A Omitted Proofs in Section 3

In this section, we will finish the proof of [Theorem 3.1](#) by showing the exponential mechanism EXP_ϵ^R is individually rational and has no positive transfer.

Individual Rationality. We first note that for any agent i , if $v_i = \mathbf{0}$, then by the definition of our pricing scheme we always have $p_i = 0$ regardless of bidding valuations of other agents. Therefore, by bidding $\mathbf{0}$ agent i could always guarantee non-negative expected utility. Since we have showed that the exponential mechanism is truthful-in-expectation, we get that the utility of agent i when she truthful reports her valuation is always non-negative.

No-Positive-Transfer. Let us turn to the second part: showing the payments are always non-negative in the exponential mechanism. Recall the payment for agent i is

$$p_i = \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(\mathbf{v})} [v_i(r)] - \frac{2}{\epsilon} \ln \left(\sum_{r \in R} \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r) \right) \right) + \frac{2}{\epsilon} \ln \left(\sum_{r \in R} \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r) \right) \right) .$$

Let us consider two distributions P and Q , such that the probability mass functions of P and Q are given by

$$\Pr_{r \sim P}[r] \propto \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r) \right) , \quad \Pr_{r \sim Q}[r] \propto \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r) \right)$$

By the non-negativity of KL-divergence, we have $D_{KL}(P||Q) \geq 0$, that is

$$\sum_{r \in R} \Pr_{r \sim P}[r] \ln \left(\frac{\Pr_{r \sim P}[r]}{\Pr_{r \sim Q}[r]} \right) \geq \sum_{r \in R} \Pr_{r \sim P}[r] \ln \left(\frac{\Pr_{r \sim P}[r]}{\Pr_{r \sim Q}[r]} \right) . \quad (5)$$

The left-hand-side is

$$\sum_{r \in R} \Pr_{r \sim P}[r] \log \left(\frac{\Pr_{r \sim P}[r]}{\Pr_{r \sim P}[r]} \right) = \sum_{r \in R} \Pr_{r \sim P}[r] \cdot \frac{\epsilon}{2} \sum_{k=1}^n v_k(r) - \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r') \right) \right) .$$

The right-hand-side is

$$\sum_{r \in R} \Pr_{r \sim Q}[r] \log \left(\frac{\Pr_{r \sim P}[r]}{\Pr_{r \sim P}[r]} \right) = \sum_{r \in R} \Pr_{r \sim P}[r] \cdot \frac{\epsilon}{2} \sum_{k \neq i} v_k(r) - \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r') \right) \right) .$$

So we have

$$\begin{aligned}
0 &\leq \sum_{r \in R} \Pr_{r \sim P}[r] \ln \left(\Pr_{r \sim P}[r] \right) - \sum_{r \in R} \Pr_{r \sim P}[r] \ln \left(\Pr_{r \sim Q}[r] \right) \\
&= \frac{\epsilon}{2} \sum_{r \in R} \Pr_{r \sim P}[r] v_i(r) - \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r') \right) \right) + \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r') \right) \right) \\
&= \frac{\epsilon}{2} p_i .
\end{aligned}$$

Hence, we conclude that the payments are non-negative.

B Omitted Proofs in Section 4

Proof of Lemma 4.2. We let $\Pr[r] \propto \exp(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r))$ be the probability $\text{EXP}_\epsilon^R(b_i, b_{-i})$ chooses r . The payment for agent i is

$$\begin{aligned}
p_i &= \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [b_i(r)] - \frac{2}{\epsilon} \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} b(r') \right) \right) + h_i(b_{-i}) \\
&= \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [b(r)] - \frac{2}{\epsilon} \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} b(r') \right) \right) - \sum_{k \neq i} \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [b_k(r)] + h_i(b_{-i}) \\
&= \sum_{r \in R} \Pr[r] b(r) - \frac{2}{\epsilon} \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} b(r') \right) \right) - \sum_{k \neq i} \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [b_k(r)] + h_i(b_{-i}) \\
&= \frac{2}{\epsilon} \sum_{r \in R} \Pr[r] \left(\frac{\epsilon}{2} b(r) - \ln \left(\sum_{r' \in R} \exp \left(\frac{\epsilon}{2} b(r') \right) \right) \right) - \sum_{k \neq i} \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} [b_k(r)] + h_i(b_{-i}) \\
&= \frac{2}{\epsilon} \sum_{r \in R} \Pr[r] \ln (\Pr[r]) - \mathbf{E}_{r \sim \text{EXP}_\epsilon^R(b_i, b_{-i})} \left[\sum_{k \neq i} b_k(r) \right] + h_i(b_{-i}) .
\end{aligned}$$

□

C Approximate Implementation for Multi-Item Auction

In this section, we will explain how to approximately implement the exponential mechanism in the multi-item auction setting. The main technical tool in this section is the seminal work of Jerrem, Sinclair, and Vigoda [12] on approximating the permanent of non-negative matrices, which can be phrased as follows:

Lemma C.1 (FPRAS for permanent of non-negative matrices [12]). *For any $\gamma > 0$ and any $\delta \in (0, 1)$, there is an algorithm that computes the permanent of an arbitrary $n \times n$ matrix $A = \{a_{ij}\}_{i,j \in [n]}$ up to a multiplicative factor of $\exp(\gamma)$ with probability at least $1 - \delta$. The running time is polynomial in n , γ^{-1} , $\log(\delta^{-1})$, and $\log(\max_{i,j \in [n]} a_{ij} / \min_{i,j \in [n]} a_{ij})$.*

To see the connection between the permanent of non-negative matrices and implementation of the exponential mechanism in the multi-item auction setting, we point out that the normalization

factor in the outcome distribution of the exponential mechanism is the permanent of a non-negative matrix:

$$\sum_{r \in R_M} \exp\left(\frac{\epsilon}{2} \sum_{i=1}^n v_i(r)\right) = \sum_{\pi \in \Pi_n} \prod_{i=1}^n \exp\left(\frac{\epsilon}{2} v_{i\pi[i]}\right) = \text{perm}\left(\left\{\exp\left(\frac{\epsilon}{2} v_{ij}\right)\right\}_{i,j \in [n]}\right).$$

We will let $A(\mathbf{v})$ denote the matrix $\{\exp(\frac{\epsilon}{2} v_{ij})\}_{i,j \in [n]}$. Moreover, we let $A_{-i,-j}(\mathbf{v})$ denote the $(n-1) \times (n-1)$ matrix obtained by removing the i^{th} row and the j^{th} column of $A(\mathbf{v})$.

C.1 Approximate Sampler

Now we are ready to introduce the approximate sampler for the multi-item auction.

Lemma C.2. *For any $\delta \in (0, 1)$ and $\gamma > 0$, there is a sampling algorithm whose running time is polynomial in n , $\epsilon^{-1} \gamma^{-1}$, and $\log \delta^{-1}$, such that with probability at least $1 - \delta$, the sampling algorithm choose an outcome r with probability*

$$\Pr[r] \in [\exp(-\gamma), \exp(\gamma)] \Pr[\text{EXP}_\epsilon^{R_M} = r].$$

Proof. We will recursively decide which item we will allocate to agent i for $i = 1, 2, \dots, n$ by repeatedly computing an accurate estimation of the marginal distribution. Concretely, the algorithm is given as follows:

1. Use the FPRAS in [Lemma C.1](#) to compute $\text{perm}(A_{-1,-j}(\mathbf{v}))$ up to a multiplicative factor of $\exp(\frac{\gamma}{2n})$ with success probability at least $1 - \frac{\delta}{n^2}$. Let x_j denote the approximate value.
2. Sample an item j with probability $\Pr[j] \propto x_j$.
3. Allocate item j to agent 1 and recurse on the remaining $n-1$ agents and $n-1$ items.

First we note that for each allocation $\pi \in \Pi_n$, the probability that π is chosen as the outcome can be decomposed into n stages by Bayes' rule:

$$\Pr[\text{EXP}_\epsilon^{R_M}(\mathbf{v}) = \pi] = \Pr[\text{agent 1 gets } \pi[1]] \cdot \Pr[\text{agent 2 gets } \pi[2] \mid \pi[1]] \cdots \Pr[\text{agent } n \text{ gets } \pi[n] \mid \pi[1], \dots, \pi[n-1]].$$

In the first recursion of our algorithm, we use the distribution

$$\Pr[\text{agent 1 gets item } j] \propto x_j \approx \text{perm}(A_{-1,-j}(\mathbf{v})).$$

Further, in the exponential mechanism

$$\begin{aligned} \Pr[\text{agent 1 gets item } j \text{ in } \text{EXP}_\epsilon^{R_M}] &\propto \sum_{\pi: \pi[1]=j} \exp\left(\frac{\epsilon}{2} \sum_{k=1}^n v_{k\pi[k]}\right) \\ &= \exp\left(\frac{\epsilon}{2} v_{1j}\right) \text{perm}(A_{-1,-j}(\mathbf{v})). \end{aligned}$$

Since x_j approximate $\text{perm}(A_{-1,-j}(\mathbf{v}))$ up to an $\exp(\frac{\gamma}{2n})$ factor, we know the probability that item j is allocated to agent 1 in our algorithm approximate the correct marginal up to an $\exp(\frac{\gamma}{n})$ multiplicative factor.

Similar claim holds for the rest of the $n-1$ stages as well. So the probability that we samples a permutation $\pi \in R_M$ differs from the correct distribution by at most a $\exp(\frac{\gamma}{n})^n = \exp(\gamma)$ factor. Moreover, by union bound the failure probability is at most δ . \square

C.2 Approximate Payments

Next, we will turn to approximate implementation of the payment scheme. First, recall that the payment for agent i is

$$\begin{aligned} p_i &= \mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)] - \frac{2}{\epsilon} \ln \left(\sum_{r \in R_M} \exp \left(\frac{\epsilon}{2} \sum_{k=1}^n v_k(r) \right) \right) + \frac{2}{\epsilon} \ln \left(\sum_{r \in R_M} \exp \left(\frac{\epsilon}{2} \sum_{k \neq i} v_k(r) \right) \right) \\ &= \mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)] - \frac{2}{\epsilon} \ln (\text{perm}(A(v_i, v_{-i}))) + \frac{2}{\epsilon} \ln (\text{perm}(A(\mathbf{0}, v_{-i}))) . \end{aligned}$$

The next lemma states that we can efficiently compute an estimator for the payment p_i with inverse polynomially small bias.

Lemma C.3. *For any $\delta \in (0, 1)$ and $\gamma \in (0, 1)$, we can compute in polynomial time (in n , ϵ^{-1} , and γ^{-1}) a random estimator \hat{p}_i for p_i such that the bias is small: $|\mathbf{E}[\hat{p}_i] - p_i| \leq \gamma$.*

Proof. By [Lemma C.1](#), we can efficiently estimate $\text{perm}(A(v_i, v_{-i}))$ and $\text{perm}(A(\mathbf{0}, v_{-i}))$ up to a multiplicative factor of $\exp(\frac{\gamma}{6})$ with success probability at least $1 - \frac{\gamma}{6}$. Hence, we can compute $\ln(\text{perm}(A(v_i, v_{-i})))$ and $\ln(\text{perm}(A(\mathbf{0}, v_{-i})))$ up to additive bias of $\frac{\gamma}{6}$ with probability $1 - \frac{\gamma}{6}$. Note that the total bias introduced if the FPRAS fails is at most 1 and that could happen with probability at most $\frac{\gamma}{6}$. So the total bias from estimating $\ln(\text{perm}(A(v_i, v_{-i})))$ and $\ln(\text{perm}(A(\mathbf{0}, v_{-i})))$ is at most $\frac{\gamma}{2}$.

It remains to compute an estimator for $\mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)]$ with bias less than $\frac{\gamma}{2}$. In order to do so, we will use the algorithm in [Lemma C.2](#) to sample an outcome r^* from a distribution whose probability mass function differs from that of $\text{EXP}_\epsilon^{RM}(\mathbf{v})$ by at most a $\exp(\frac{\gamma}{6})$ factor point-wise, with success probability at least $1 - \frac{\gamma}{6}$. Then we will use $v_i(r^*)$ as our estimator. Note that conditioned on the sampler runs correctly, we have

$$\left| \mathbf{E}[v_i(r^*)] - \mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)] \right| \leq \left(\exp \left(\frac{\gamma}{6} \right) - 1 \right) \mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)] \leq \left(\exp \left(\frac{\gamma}{6} \right) - 1 \right) \leq \frac{\gamma}{3} .$$

Moreover, the maximum bias conditioned on the failure of the sampler is at most 1, which happens with probability at most $\frac{\gamma}{6}$. So the total bias from the estimator for $\mathbf{E}_{r \sim \text{EXP}_\epsilon^{RM}(\mathbf{v})} [v_i(r)]$ is at most $\frac{\gamma}{2}$. \square

D Lower Bound for Multi-Item Auction

Proof of [Theorem 5.2](#). Let us first define some notations. For any $j^* \in [n]$, we will let e^{j^*} denote the valuation profile such that $e_j^{j^*} = 1$ if $j = j^*$ and $e_j^{j^*} = 0$ if $j \neq j^*$. That is, an agent with valuation e^{j^*} is single-minded who only value getting item j^* (with value 1) and has no interest in getting any other item. We will say j^* is the *critical item* for this agent.

Suppose M is an ϵ -differentially private mechanism such that M always obtain at least $\text{opt} - \frac{n}{10}$ expected social welfare. Let us consider the following randomly chosen instance: each agent's valuation is chosen from e^1, \dots, e^n independently and uniformly at random. Let us consider the social welfare we get by running mechanism M on this randomly constructed instance. We first note that $\mathbf{E}_{\mathbf{v}}[\text{opt}(\mathbf{v})] = (1 - e^{-1})n$ for that each item has probability $1 - e^{-1}$ of being the critical

item of at least one of the agents. By our assumption, the expected welfare obtained by M shall be at least $(1 - e^{-1})n - \frac{n}{10} > \frac{n}{2}$. Therefore, we have

$$\sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is critical for } i] \Pr[j \text{ is critical for } i] \geq \frac{n}{2} .$$

Note that $\Pr[j \text{ is critical for } i] = \frac{1}{n}$ for all $i, j \in [n]$, we get that the average probability that a critical item-agent pair is allocated is at least half:

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is critical for } i] \geq \frac{1}{2} . \quad (6)$$

Similarly, we have

$$\sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is not critical for } i] \Pr[j \text{ is not critical for } i] \leq \frac{n}{2} .$$

Note that $\Pr[j \text{ is not critical for } i] = \frac{n-1}{n}$ for all $i, j \in [n]$, we get that the average probability that the average probability that a non-critical item-agent pair is chosen in the allocation is very small:

$$\frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is not critical for } i] \leq \frac{1}{2(n-1)} . \quad (7)$$

By (6) and (7), we have

$$\frac{\sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is critical for } i]}{\sum_{i=1}^n \sum_{j=1}^n \Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is not critical for } i]} \geq n - 1 .$$

In particular, we know there exists a (i, j) pair such that

$$\frac{\Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is critical for } i]}{\Pr[M \text{ allocate } j \text{ to agent } i \mid j \text{ is not critical for } i]} \geq n - 1 .$$

Since M is ϵ -differentially private, we get that $\exp(\epsilon) \geq n - 1$, and thus $\epsilon = \Omega(\log n)$. \square

E Lower Bound for Procurement Auction for Spanning Trees

Proof of Theorem 5.4. Suppose M is an ϵ -differentially private mechanism whose expected total cost is at most $\text{opt} + \frac{k}{24}$.

We will consider the following randomly generated instance. Each agent i 's cost value c_i is independently chosen as

$$c_i = \begin{cases} 1 & , \text{ w.p. } 1 - \frac{1}{2k} \\ 0 & , \text{ w.p. } \frac{1}{2k} \end{cases}$$

If an agent has cost 0, we say this agent and the corresponding edge are *critical*. Let us first analyze the expected value of opt for such randomly generated instances. Intuitively, we want to pick as many critical edges as possible. In particular, when there are no cycles consists of only critical edges, the minimum spanning tree shall pick all critical edges, which comprise a forest in the graph, and then pick some more edges to complete the spanning tree.

Lemma E.1. *With probability at least $\frac{1}{2}$, there are no cycle consists of only critical edges.*

Proof of Lemma E.1. For each cycle of length t , the probability that all edges on this cycle are critical is $(2k)^{-t}$. Note that the number of cycles of length t is at most $\binom{k}{t}(t-1)! \leq k^t$. Here $\binom{k}{t}$ is the number of subsets of t vertices and $(t-1)!$ is the number of different Hamiltonian cycles among t vertices. Hence, by union bound, the probability that there is any cycle consists of only critical edges is at most $\sum_{t=2}^k (2k)^{-t} \cdot k^t = \sum_{t=2}^k 2^{-t} < \frac{1}{2}$. \square

Moreover, by Chernoff-Hoeffding bound, we have that the number of critical edges is at least $\frac{k}{3}$ with probability at least $\frac{3}{4}$.

Therefore, by union bound, with probability at least $\frac{1}{4}$, we have that there are at least $\frac{k}{3}$ critical edges and there are no cycle consists of only critical edges. So in this case, we have $\text{opt} \leq k - \frac{k}{3} = \frac{2k}{3}$. Therefore, the expectation of the optimal total cost is at most $\mathbf{E}[\text{opt}] \leq \frac{3}{4}k + \frac{1}{4}\frac{2k}{3} = \frac{11k}{12}$.

By our assumption on M , we get that the expected total cost of the outcome chosen by M is at most $\frac{11k}{12} + \frac{k}{24} = \frac{23k}{24}$. In other words, the expected number of critical edges chosen by M is at least $\frac{k}{24}$. That is,

$$\sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is critical}] \Pr[\text{edge } i \text{ is critical}] \geq \frac{k}{24} .$$

Note that $\Pr[\text{edge } i \text{ is critical}] = \frac{1}{2k}$ for all $i \in [n]$ and $n = \binom{k}{2} = \frac{k(k-1)}{2}$, we get that on average a critical edge is chosen with at least constant probability

$$\frac{1}{n} \sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is critical}] \geq \frac{1}{6} .$$

On the other hand, it is easy to see

$$\sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is not critical}] \Pr[\text{edge } i \text{ is not critical}] \leq k .$$

By $\Pr[\text{edge } i \text{ is not critical}] = 1 - \frac{1}{2k}$ and $n = \binom{k}{2}$, we get that on average a non-critical edge is chosen with very small probability

$$\frac{1}{n} \sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is not critical}] \leq \frac{2k^2}{(2k-1)n} = \frac{4k}{(k-1)(2k-1)} \leq \frac{8}{2k-1} .$$

Therefore, we have

$$\frac{\sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ critical}]}{\sum_{i=1}^n \Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is not critical}]} \geq \frac{2k-1}{48} .$$

In particular, there exists an agent i , such that

$$\frac{\Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ critical}]}{\Pr[\text{edge } i \text{ is chosen} \mid \text{edge } i \text{ is not critical}]} \geq \frac{2k-1}{48} .$$

However, the above amount is upper bounded by $\exp(\epsilon)$ since M is ϵ -differentially private. So we conclude that $\epsilon = \Omega(k)$. \square