

Multidimensional Dynamic Pricing for Welfare Maximization

AARON ROTH, University of Pennsylvania

ALEKSANDRS SLIVKINS, Microsoft Research

JONATHAN ULLMAN, Northeastern University

ZHIWEI STEVEN WU, University of Pennsylvania

We study the problem of a seller dynamically pricing d distinct types of indivisible goods, when faced with the online arrival of unit-demand buyers drawn independently from an unknown distribution. The goods are not in limited supply, but can only be produced at a limited rate and are costly to produce. The seller observes only the bundle of goods purchased at each day, but nothing else about the buyer's valuation function. Our main result is a dynamic pricing algorithm for optimizing welfare (including the seller's cost of production) that runs in time and a number of rounds that are polynomial in d and the approximation parameter. We are able to do this despite the fact that (i) the price-response function is not continuous, and even its fractional relaxation is a non-concave function of the prices, and (ii) the welfare is not observable to the seller.

We derive this result as an application of a general technique for optimizing welfare over *divisible* goods, which is of independent interest. When buyers have strongly concave, Hölder continuous valuation functions over d divisible goods, we give a general polynomial time dynamic pricing technique. We are able to apply this technique to the setting of unit demand buyers despite the fact that in that setting the goods are not divisible, and the natural fractional relaxation of a unit demand valuation is not strongly concave. In order to apply our general technique, we introduce a novel price randomization procedure which has the effect of implicitly inducing buyers to “regularize” their valuations with a strongly concave function. Finally, we also extend our results to a limited-supply setting in which the number of copies of each good cannot be replenished.

Additional Key Words and Phrases: Dynamic Pricing, Welfare Maximization, Learning, Revealed Preferences

1 INTRODUCTION

Consider the problem of an online retailer who sells a large variety of goods. The seller can in principle produce or procure more copies of each good as needed, but only at a limited rate, and at some per-unit production /procurement cost that varies by good. In each round, the seller can dynamically set the price for each type of good. Each buyer has an unknown valuation function defined (in general) over bundles of goods, and quasi-linear utility for money. Each buyer chooses which item to buy to optimize his utility function given the prices. The seller observes the purchased bundle—i.e. the *revealed preferences of the buyer*—but not the buyer's valuation of the purchased bundle (or of any other bundle). The buyer's valuation function is drawn independently from a fixed but *unknown* distribution, called the *buyer distribution*. The seller's objective is to optimize social welfare: the expected buyer valuation of the purchased good minus its production cost.

This work was partially supported by NSF grant CNS-1253345, a Sloan Foundation Fellowship, and a DARPA grant. A full version of the paper can be found on arXiv: <https://arxiv.org/abs/1607.05397>.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

EC'17, June 26–30, 2017, Cambridge, Massachusetts, USA. ACM ISBN 978-1-4503-4527-9/17/06...\$15.00.

© 2017 Copyright is held by the owner/author(s). Publication rights licensed to ACM.

DOI: <http://dx.doi.org/10.1145/3033274.3085106>

Social welfare, like profit, is a natural objective for the seller: in particular, sellers attempting to grow their market (rather than exploit an existing monopoly position) might prefer to optimize social welfare rather than profit in the short term.

A tempting first attempt at solving this problem would be to simply set the price for each good to be equal to its cost of production, which would indeed maximize social welfare if there were no other constraints on the bundles of items purchased by buyers. However, this solution is unsatisfactory when additional supply can only be generated at a bounded rate, because the cost of production bears no relationship to the buyers' values for a good. Because of this, setting prices equal to costs can result, for example, in every buyer demanding the largest possible quantity of the same good, which the seller may not be able to accommodate. In a more realistic setting, there will be constraints on the rate of production and resupply for each good. Hence, we study the welfare maximization problem in which we impose the additional constraint that the *expected bundle* purchased (in expectation over the draw of the buyer) lies in a bounded set. Because constraints of this sort bind across buyers, setting prices equal to costs fails, and the problem requires a nontrivial solution.

Since the buyer distribution is unknown, the seller cannot directly compute the prices that optimize social welfare. Instead, she faces a learning problem: she can try different prices over time and observe the responses from random buyers drawn from the distribution, and try to learn the optimal price vector. More formally, the goal is to use a small number of rounds to learn a price vector that nearly optimizes expected social welfare. We want the algorithm's guarantees hold in the worst case over the choice of distributions over buyer valuation functions.

Essentially, we are studying a welfare-optimization version of the well-known *dynamic pricing* problem, also known as *learn-and-earn*, with $d > 1$ goods for sale. (Prior work on dynamic pricing focused on profit maximization.) At a very high level, the main challenge presented is to learn the price response function—i.e. the function mapping prices to expected bundles purchased—and then optimize it with respect to welfare. Moreover, this is a high dimensional function (for large d), and so one must overcome the curse of dimensionality. Prior work on non-Bayesian dynamic pricing (e.g., [4, 10, 11, 14, 19, 26, 33]) dealt with this challenge by making strong assumptions on the price response function itself. Typical assumptions include Lipschitzness [10, 11, 33] (which allows for discretization in low-dimensional problems), and particularly for high dimensional problems, linearity [19, 26] or concavity [4, 10, 33].¹ However, assumptions of this sort are not well supported by a micro-economic foundation. In fact, natural assumptions on the buyer valuations do not necessarily result in price-response function with these properties.

In this paper, we pursue a different approach which stands on stronger microeconomic foundations: we make assumptions on the form of the valuation functions directly (and no assumptions on the distribution over valuation functions), and show that we can work with the price response function that results. This is the case *despite* the fact that our problem is high dimensional, and the price response function that results from our assumptions is not concave. We also face an additional challenge: unlike profit, welfare is not observable, and we can observe the purchased bundle but not the buyer's valuation for that bundle. Nevertheless, we design algorithms that find a near-optimal price vector with respect to welfare, in a number of rounds that is polynomial in d and the accuracy parameter. (Whereas, for example, a naive solution based on discretization and Lipschitzness of the price-response function requires a number of rounds that is exponential in d .) Our results also extend to the limited-supply setting.

¹Prior work that does *not* make assumptions on the shape of valuations or demand curves is either restricted to selling a single good ($d = 1$) [4, 27], or suffers from the curse of dimensionality *and* comes with performance guarantees relative to the discretized prices rather than all prices [5].

1.1 Our Contributions

Our main result solves the problem in the setting of *indivisible goods* when buyers are unit-demand or, alternatively, when the seller only allows each customer to buy a single item. Surprisingly, no other assumptions are needed! Further, we give a general result for the setting of *divisible goods* under certain assumptions on the valuation functions. In fact, we show how this general result can be leveraged to yield the result for unit demands.

Both settings work as follows. There are d goods. In each round, prices are set and one buyer arrives and purchases her most preferred bundle from the set of feasible bundles. The seller incurs production/procurement costs for each sale, which are linear in the sold bundle.

We give a computationally efficient and round efficient algorithm for finding a nearly welfare-maximizing price vector subject to a constraint on the expected consumption. Let $\text{SW}(p)$ be the expected social welfare that results from setting prices p . The seller would like to set prices to ensure that the expected per-round purchase of each good j , denoted $x_j(p)$, is bounded above by some supply s_j . This models a realistic scenario in which the seller's inventory can be replenished, but only at a limited rate. For example, perhaps at most one truckload of goods can be stocked per day. Approximating a restocking period constraint with a constraint on the expected per-customer purchase is reasonable if the restocking period corresponds to a large number of rounds, because then the realized consumption over these rounds concentrates around its expectation. In the following, we will write $x(p) = (x_1(p), \dots, x_d(p))$ to denote the bundle *induced* by prices p .

Divisible goods. Departing from previous work, instead of making assumptions about the functional form of the price-response function, which depends on the buyers' valuations in aggregate, we make assumptions on the individual buyers' valuations themselves. Specifically, we assume the buyers' valuations are strongly concave and Hölder continuous. (These assumptions are satisfied by a large class of well-studied valuation, including *CES* and *Cobb-Douglas* as shown in [29]). The sold bundles are constrained to lie in the bounded set $\mathcal{F} \subset \mathbb{R}_+^d$ (e.g. $\mathcal{F} = [0, 1]^d$ means at most one unit of each good can be purchased).

THEOREM 1.1 (DIVISIBLE GOODS). *Assume divisible goods, and buyers with strongly concave and Hölder-continuous valuations. There is an algorithm that takes as input parameters $d, \alpha, \delta > 0$ and a supply vector $s \in \mathbb{R}_{>0}^d$, such that with probability at least $1 - \delta$, the algorithm outputs a price vector $p \in \mathbb{R}_+^d$ such that*

$$x(p) \leq s \quad \text{and} \quad \text{SW}(p) \geq \max_{p \in \mathbb{R}_+^d: x(p) \leq s} \text{SW}(p) - \alpha. \quad (1)$$

The number of rounds and the total computation time are polynomial in $d, \frac{1}{\alpha}$ and $\log \frac{1}{\delta}$.

Unit-demand buyers and indivisible goods. We use our result for divisible goods to give a polynomial time dynamic pricing algorithm for welfare maximization in the *indivisible goods* setting when buyers have unit demand valuations. Here we consider *distributions* \mathcal{D} over price vectors $p \in \mathbb{R}_+^d$, rather than fixed price vectors p . To extend our notation, let $x(\mathcal{D}) = \mathbb{E}_{p \sim \mathcal{D}} [x(p)]$ and $\text{SW}(\mathcal{D}) = \mathbb{E}_{p \sim \mathcal{D}} [\text{SW}(p)]$. We prove:

THEOREM 1.2 (INDIVISIBLE GOODS). *Assume indivisible goods, and buyers with unit-demand valuations. There is an algorithm that takes as input parameters $d, \alpha, \delta > 0$ and a supply vector $s \in \mathbb{R}_+^d$, such that with probability at least $1 - \delta$ the algorithm outputs a distribution \mathcal{D} over price vectors $p \in \mathbb{R}_+^d$ such that*

$$x(\mathcal{D}) \leq s \quad \text{and} \quad \text{SW}(\mathcal{D}) \geq \max_{\text{distributions } \mathcal{D}': x(\mathcal{D}') \leq s} \text{SW}(\mathcal{D}') - \alpha. \quad (2)$$

The number of rounds and the total computation time are polynomial in d , $\frac{1}{\alpha}$ and $\log \frac{1}{\delta}$.

In the generality that we state our theorem, using a *distribution* over prices rather than a fixed price vector is unavoidable. The reason has to do with how the buyers break ties when they are indifferent between goods. As shown in [25], without further genericity assumptions, it can be that *no* fixed pricing can induce optimal (or even feasible) allocations if buyers use uncoordinated tie breaking rules. Instead, tie-breaking needs to be *coordinated* amongst the buyers: the tie-breaking rule needs to be different for different buyers, and essentially needs to be specified by the mechanism. The randomness in our pricing scheme serves as a coordination mechanism amongst buyers (since each buyer faces a different realization of prices).

Remark. In both settings, we prove more general theorems in which we express our results in terms of a stronger benchmark—the welfare of the optimal lottery over allocations, without restriction to those that can be induced by posted pricing.²

The above theorems can be reformulated in terms of cumulative regret for a given time horizon T . Then the execution of the algorithm in the respective theorem corresponds to an *exploration phase* of bounded length. The price vector p computed by the algorithm is used in an *exploitation phase* consisting of all subsequent rounds. Theorem 1.1 guarantees that OWEL completes in $\text{poly}(d, \log \frac{1}{\delta}) \cdot \alpha^{-m}$ rounds, for some constant m . Expected regret relative to the best fixed price vector can be upper-bounded by 1 for every round of exploration, and $\alpha + \delta$ per round of exploitation. Optimizing the choice of α and δ , we obtain regret $\text{poly}(d, \log T) \cdot T^{m/(m+1)}$. Theorem 1.2 implies a similar corollary for the unit-demands setting.

Extension to limited supply. We extend our results to a limited-supply setting. In our model, there is a fixed horizon of T rounds and the seller has a non-replenishable supply of Ts_j units of each good j . T and $s \in [0, 1]^d$ are known in advance. Each day, the seller will set prices and a random buyer will purchase their preferred bundle until either the time horizon or the sellers' supply is exhausted, whichever comes first.

For a pricing policy π , we use $\text{SW}_{\text{tot}}(\pi)$ to denote its expected total welfare.³ A “fixed-vector” pricing policy uses the same price vector p in all rounds. Likewise, “fixed-distribution” pricing policy always draws the price vector independently from the same fixed distribution \mathcal{D} . The expected total welfare of these policies is denoted, resp., $\text{SW}_{\text{tot}}(p)$ and $\text{SW}_{\text{tot}}(\mathcal{D})$.

In the setting of divisible goods, we simply use the algorithm from Theorem 1.1 with the same constraint vector s . The price vector p computed by this algorithm achieves high expected total welfare for a given problem instance: we prove that it is nearly optimal compared to the best fixed-vector pricing policy. Further, it is nearly optimal compared to *any* pricing policy.

Likewise, in the setting of unit demands, we use the algorithm from Theorem 1.2 with the same constraint vector s . The distribution \mathcal{D} computed by this algorithm is nearly optimal compared to the best fixed-distribution pricing policy with $x(\mathcal{D}) \leq s$.

THEOREM 1.3. *Consider dynamic pricing with limited supply. Fix constraint vector $s \in \mathbb{R}_+^d$ and time horizon $T > 32 \log(T)/s^*$, where $s^* = \min_j s_j$.*

²A posted price vector (resp., distribution over them) computed by an algorithm can only hope to compete with the best posted price vector (resp., distribution). Thus, the mathematical statement behind (ii) is that the posted price benchmarks used in the theorems are in fact equivalent to the stronger benchmark.

³Considering expected welfare per round is not enough, as one pricing policy may halt sooner than another.

- (a) Consider the setting of divisible goods. When the algorithm from Theorem 1.1 is given as input d, s, α, δ , with probability $1 - \delta$ it outputs a price vector $p \in \mathbb{R}_+^d$ such that

$$SW_{\text{tot}}(p) \geq \sup_{\text{pricing policies } \pi} SW_{\text{tot}}(\pi) - \alpha T - O\left(\sqrt{T \log(T)/s^*}\right).$$

- (b) Assume indivisible goods and unit demands. When the algorithm from Theorem 1.2 is given as input d, s, α, δ , with probability $1 - \delta$ it outputs distribution \mathcal{D} over price vectors such that

$$SW_{\text{tot}}(\mathcal{D}) \geq \sup_{\text{distributions } \mathcal{D}' \text{ with } x(\mathcal{D}') \leq s} SW_{\text{tot}}(\mathcal{D}') - \alpha T - O\left(\sqrt{T \log(T)/s^*}\right).$$

The number of rounds and the total computation time are polynomial in $d, \frac{1}{\alpha}$ and $\log \frac{1}{\delta}$.

1.2 Our Techniques

Our general results for divisible goods build on a crucial structural property: even though the expected welfare of the induced bundle $x(p)$ is not concave in the price vector p , it becomes concave if we treat the bundle itself as the decision variable. We illustrate this via a simple 1-dimensional example, adapted from [29]:

Example 1.4. There is a single good ($d = 1$), and a single buyer with valuation $v(x) = \sqrt{x}$. he seller's cost function is $c(x) = x$. If price p is posted, the buyer's utility for x units is $\sqrt{x} - p \cdot x$, so she would purchase $x^*(p) = \frac{1}{4p^2}$ units of the good. Consequently, the welfare is

$$SW(p) = v(x^*(p)) - c(x^*(p)) = \frac{1}{2p} - \frac{1}{4p^2}.$$

Note that welfare is not a concave function of the price. However, if we write the welfare as a function of $x = x^*(p)$, the purchased amount of good, this function is concave:

$$SW(x) = v(x) - c(x) = \sqrt{x} - x.$$

Thus, we would like to optimize expected welfare as a function of the induced bundle. However, we only control prices and not induced bundles. To address this, our algorithm has two “layers,” where the outer layer optimizes over induced bundles, and the inner layer finds a price vector which approximately induces a given bundle. Another challenge is that welfare is not observed, since we do not observe buyer valuations. Instead, we find a way to approximate the *subgradients* of welfare, and use noise-tolerant subgradient descent to optimize over the bundles.

We build on and extend the result of [29] for the special case of a single buyer and unlimited supply (which focuses on profit rather than welfare). The main distinction is single buyer vs. distributions over buyers; in other words, [29] assume that for a given price vector the outcome is deterministic, whereas in our paper it is drawn from a fixed but unknown distribution over the possible outcomes.

The “inner layer” of our algorithm extends the algorithm in [29] from a single buyer to distributions over buyers. This extension presents several technical challenges, and answers one of their main open questions. In particular, we analyze a generalization of the convex programming technique used in [29] to accommodate a distribution over (arbitrarily many) buyers. We cannot use the “outer layer” from [29] because it requires direct observations of the objective function to feed into a procedure for zeroth-order optimization, and our seller cannot directly observe the buyers' welfare (unlike profit, which is observable). Instead, we develop a new technique to obtain the subgradient for the welfare function so as to enable first-order optimization. Also, we remove a major assumption of homogenous buyer valuations.

As stated, our general result does not apply to unit demand buyers over indivisible goods. In order to cast this problem as a divisible goods problem, we view buyers as having *linear* valuations over

divisible goods, optimizing over the set of bundles that have at most unit ℓ_1 norm. The bundle that maximizes a linear function is always at a vertex of the feasible region, and hence is integral. That is, it is the bundle purchased by a unit-demand buyer in the indivisible goods setting. However, there is a substantial difficulty: our general technique relies on buyer valuation functions being *strongly* concave, a condition not satisfied by linear functions. A standard way to obtain strong convexity in the convex optimization literature is to add a strongly convex *regularizer* to the objective function. However, we do not get to modify the buyer's objective function in this way. Instead, we perturb the price vectors proposed by our general dynamic pricing algorithm with Gumbel noise. Doing so has the property that the expected bundle purchased by each buyer (where expectation is taken over the price perturbation) is the bundle that maximizes the buyer's linear valuation function, plus an entropy regularization term [34]. Thus, in expectation over our perturbations, we can view buyers as optimizing valuation functions which are strongly concave over the ℓ_1 norm—even though for every fixed perturbation, buyers are maximizing some linear function and thus buy a unit bundle of indivisible goods. By perturbing the price vectors used over the run of our algorithm for divisible goods, therefore, we can optimize welfare over these “regularized” buyers. By reducing the noise rate (and hence the implicit regularization parameter), we approach the optimal welfare of the actual, unit-demand buyers.

The extension to limited supply (Theorem 1.3(a)) relies on a structural result about *bandits with knapsacks* [5], a general framework of which dynamic pricing with limited supply is a special case. We use a non-standard “embedding” of dynamic pricing into this framework, and a concentration inequality for total welfare that requires a somewhat delicate proof.

1.3 Related Work

Our setting is related to several lines of work. First, dynamic pricing, a.k.a. *learn-and-earn*, focuses on a seller with a large inventory of each good, facing a stream of buyers with unknown valuations. This is a large line of work, mainly in operations research — see [13] for a review. Most related are non-Bayesian approaches. As mentioned above, the main distinction is that we make assumptions on the buyer valuations rather than on the price response function. Also, the learn-and-earn literature does not consider welfare-optimization, to the best of our knowledge.

Second, our problem can be viewed as an instance of the *multi-armed bandits* problem [16, 23], a well-studied abstract framework in which an algorithm repeatedly chooses actions (e.g., price vectors) and receives rewards (e.g., revenue from a sale). The main issue is the tension between acquisition and usage of information, a.k.a. the *exploration-exploitation tradeoff*. Bandit algorithms are directly applicable to dynamic pricing either via discretization [4, 5, 27] or via assumptions on expected revenue.⁴ The main distinction is (again) that solutions to bandit problems tend to make assumptions directly on the rewards, in part because they do not model the finer structure behind the rewards (such as valuation functions).

Third, there are several papers on welfare-optimizing posted pricing in combinatorial auctions [6, 12, 18, 21]. These papers tackle more difficult scenarios with non-divisible goods and non-IID valuations, and accordingly obtain weaker, multiplicative guarantees. Also, the pricing is either static [6, 21] (not changing over time), or changing over time but not adapting to the observed purchases [12, 18]. This research is mainly motivated by connections to mechanism design for combinatorial auctions.

Fourth, there is a large literature on *revealed preferences*, starting from Samuelson [31], see [28, 30, 32] for background. Most work in economics has focused on the construction of utility functions that explain or *rationalize* a given sequence of price/bundle observations, e.g. [1]. A

⁴E.g., if expected revenue is concave in prices, one can apply bandit algorithms for concave rewards [2, 17, 22, 24].

recent literature studies the problem of predicting purchase decisions given past observations at different price points [8, 9, 35]. More related to our paper is Amin et al. [3] who study the problem of iteratively setting prices to maximize the profit obtained from a single budgeted buyer (who repeatedly makes purchase decisions) with a linear utility function. The most related paper in this line is [29], as discussed in the previous subsection. Ours is the first paper in this line of work able to handle indivisible goods. A related, but distinct literature focuses on *learning valuation functions* from example evaluations of those functions [7], rather than from example *maximizations* of those functions (as in the revealed preference literature).

1.4 Map of the Paper

Section 3 contains our general result for divisible goods (a generalization of Theorem 1.1). Our application to unit demand valuations over indivisible goods, and the perturbation techniques that go into deriving this result are presented in Section 4. The limited supply setting and some technical details are deferred to the full version.

2 MODEL AND PRELIMINARIES

There is a *seller* selling d different types of goods to a sequence of *buyers* arriving one after another in rounds. Each buyer's valuation v is drawn independently from an unknown distribution ψ over a finite class \mathcal{V} of valuation functions over the goods, where $|\mathcal{V}| = n$.⁵ Both ψ and \mathcal{V} are unknown to the seller. Throughout, we will use i to index the buyer's types in \mathcal{V} , and write v_i for the valuation function for a buyer of type i , and $\psi(v_i)$ for the probability mass on buyers of type i .

At each round t , the seller posts a price vector $p = p^t \in \mathbb{R}_+^d$, and the t -th buyer with valuation $v \sim \psi$ makes a purchase to maximize his utility under these prices. In particular, we consider two different settings: one with divisible goods and the other with indivisible goods.

Divisible goods. Each valuation $v: \mathbb{R}_+^d \rightarrow \mathbb{R}_+$ is a function from (fractional) bundles of goods to values. Under prices p , the buyer with valuation v will purchase the utility-maximizing bundle

$$x_v^*(p) \equiv \operatorname{argmax}_{x \in \mathcal{F}} [v(x) - \langle x, p \rangle]$$

where $\mathcal{F} \subset \mathbb{R}_+^d$ denotes the set of feasible bundles available for purchase.

Indivisible goods. We consider unit-demand buyers, who will either purchase exactly 1 unit of some good or nothing. Each buyer's valuation is defined by a value vector $v \in \mathbb{R}_{>0}^d$ such that v_j denotes her value for 1 unit of the j -th good. At round t , a buyer with valuation v purchases

$$x_v^*(p) \equiv \operatorname{argmax}_{j \in [d] \cup \{\perp\}} [v_j - p_j] .$$

where \perp denotes the choice of buying nothing (we define $v_\perp = p_\perp = 0$), and we allow arbitrary tie-breaking rules.

The seller has a (*known*) cost vector $c \in \mathbb{R}_+^d$ such that the cost of producing a unit of good j is c_j . The seller wishes to set prices so as to optimize the expected *social welfare* — the expected valuation of the buyer's purchased bundle or item minus its production cost. In particular, if the seller posts a price vector p over the goods, the expected social welfare is

$$SW(p) = \mathbb{E}_{v \sim \psi} [v(x_v^*(p)) - c(x_v^*(p))] . \quad (3)$$

⁵We take \mathcal{V} to be finite only for convenience. Our results do not depend on $n = |\mathcal{V}|$, so it can be arbitrarily large.

⁶Throughout, $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$ and $\mathbb{R}_{>0} = \{x \in \mathbb{R} \mid x > 0\}$ denote non-negative reals and positive reals, resp.

where we write $c(x_v^*(p))$ to denote the production cost for the purchase $x_v^*(p)$. For any distribution \mathcal{D} over prices, the expected welfare is defined as $\text{SW}(\mathcal{D}) = \mathbb{E}_{p \sim \mathcal{D}} [\text{SW}(p)]$.

Computational model. We will think of the algorithm as having access to a *revealed preference* oracle $\text{ReP}(\psi)$: given any input price vector $p \in \mathbb{R}_+^d$, it will draw a random valuation v from ψ , and return the purchase decision $x_v^*(p)$. Our goal is to design computationally efficient algorithms to compute optimal prices using only polynomially many queries to ReP . Notably, the expected or realized social welfare is not observable to the algorithm, since it cannot observe $v(x_v^*(p))$.

2.1 Noisy Subgradient Descent

A key ingredient in our algorithms is the ability to minimize a convex function (or maximize a concave function), given access only to noisy sub-gradients of the function. We accomplish this using the gradient descent algorithm. Below we recap some necessary background.

Let $C \subseteq \mathbb{R}^d$ be a compact and convex set of diameter at most D (w.r.t. ℓ_2 norm). A *subgradient* of a function $f : C \rightarrow \mathbb{R}$ at point $x \in C$ is any vector $g \in \mathbb{R}^d$ that satisfies the inequality $f(y) \geq f(x) + \langle g, y - x \rangle$ for any point $y \in C$. The set of all subgradients at x is denoted $\partial f(x)$. If f is differentiable, the only subgradient g is the gradient $\nabla f(x)$.

The basic subgradient descent method is an iterative algorithm that starts at some point $x_1 \in C$ and iterates the following equations

$$y_{t+1} = x_t - \eta g_t, \quad \text{and} \quad x_{t+1} = \Pi_C(y_{t+1})$$

where η is the learning rate and $g_t \in \partial f(x_t)$ is a subgradient of the function f at point x_t , and $\Pi_C(x) = \operatorname{argmin}_{y \in C} \|x - y\|$ denote the projection operator onto C .

Now, we will assume that g_t and/or x_t are subject to noise. We will use two variants of the algorithm, which operate under two different models of noise. In the first model, the algorithm only has access to *unbiased estimates* of the subgradient.

THEOREM 2.1 ([36]). *Suppose that f is convex, and for some constant D, G , the estimates of the subgradients satisfy $\mathbb{E}[g_t] \in \partial f(x_t)$ and $\|g_t\| \leq G$ for all steps t , and the diameter of the set satisfies $\|C\| \leq D$. Then if we run the subgradient descent method with step size $\eta = D/(G\sqrt{T})$, then for any T and any initial point $x_1 \in C$, the point $z = \frac{1}{T} \sum_{t=1}^T x_t$ satisfies*

$$\mathbb{E}[f(z)] \leq \min_{x \in C} f(x) + 2DG/\sqrt{T}.$$

In the second model, the algorithm has access to the noiseless subgradients, but the points x^t are adversarially perturbed after the projection.

THEOREM 2.2. *Suppose that f is convex, fix constants D, E , and G . Suppose that the gradient descent algorithm performs the following update in each iteration*

$$y_{t+1} = x_t - \eta g_t \quad \text{and} \quad x_{t+1} = \Pi_C(y_{t+1}) + \xi_t$$

such that $g_t \in \partial f(x_t)$ and $\xi_t \in \mathbb{R}^d$ is a noise vector. Suppose that $\|g_t\| \leq G$ and $\|\xi_t\| \leq E$, $x_t \in C$ for all steps t , and the diameter of the set satisfies $\|C\| \leq D$. Then if we run the subgradient descent method with step size $\eta = D/(G\sqrt{T})$, for any T and any initial point $x_1 \in C$, the point $z = \frac{1}{T} \sum_{t=1}^T x_t$ satisfies

$$f(z) \leq \min_{x \in C} f(x) + DG/\sqrt{T} + GE\sqrt{T}.$$

The proof is similar to the standard analysis of gradient descent (e.g., see Theorem 3.1 in [15]). For the sake of completeness, we provide a self-contained proof of this result in the full version.

3 A GENERAL ALGORITHM IN THE DIVISIBLE GOODS SETTING

This section is dedicated to the divisible good setting: we give a computationally efficient algorithm for finding a price vector that approximately optimizes social welfare subject to the constraint that the expected per-round demand of each good j is no more than s_j . Specifically, let $x_\psi^*(p)$ denote the expected bundle purchased by a random buyer under the prices p (or the *induced bundle* by p), that is

$$x_\psi^*(p) = \mathbb{E}_{v \sim \psi} [x_v^*(p)].$$

Given access to the revealed preference oracle ReP, the algorithm finds an approximately optimal price vector p using polynomially queries to ReP and guarantees that $x_\psi^*(p) \leq s$. Our algorithm consists of two layers, and we present it in three main steps.

- (1) First, we analyze a pertinent convex program and derive several structural results. In particular, we show that the expected social welfare can be expressed as a concave function of the induced bundle.
- (2) Next, we present the inner layer of the algorithm: given any target bundle \hat{x} , we can iteratively find price vectors p^t such that the induced bundle $x_\psi^*(p^t)$ converges to \hat{x} over time.
- (3) Finally, we show how to derive subgradients of the expected social welfare function from information available. The outer layer of the algorithm will then use (noisy) subgradient descent to optimize the welfare function over the bundle space.

We make the following assumptions on the feasible set \mathcal{F} and each valuation function $v \in \mathcal{V}$.

ASSUMPTION 3.1 (FEASIBLE SET). *For the set $\mathcal{F} \subset \mathbb{R}_+^d$ of feasible bundles, a canonical example is $\mathcal{F} = [0, 1]^d$: each buyer can simultaneously buy up to one unit of each good. More generally, we assume that \mathcal{F} is convex, closed, has a non-empty interior.⁷*

ASSUMPTION 3.2 (VALUATIONS). *Each valuation function v in \mathcal{V} satisfies:*

- (1) v is monotonically increasing in each coordinate. (This can be relaxed to Assumption 3.4.)
- (2) v is (λ, β) -Hölder continuous with respect to the ℓ_1 norm over \mathcal{F} , for some $\lambda \geq 1$ and some absolute constant $\beta \in (0, 1]$. Namely: $|v(x) - v(x')| \leq \lambda \cdot \|x - x'\|_1^\beta$ for all $x, x' \in \mathcal{F}$.
- (3) v is σ -strongly concave over \mathcal{F} — for all $x, x' \in \mathcal{F}$, $v(x') \leq v(x) + \langle \nabla v(x), x' - x \rangle - (\sigma/2) \cdot \|x - x'\|_2^2$.

These assumptions on the valuations are satisfied by a large class of well-studied valuation functions, including *Constant Elasticity of Substitution (CES)* and *Cobb-Douglas* (See [29] for a proof). We crucially rely on a use property of strongly concave functions: any point in the domain that is close to the minimum in objective value is also close to the minimum in Euclidean distance.

3.1 A Stochastic Convex Program

Let us say that a bundle $\hat{x} \in \mathbb{R}_+^d$ is *inducible* if there exists a price vector $p \in \mathbb{R}_+^d$ such that $x_\psi^*(p) = \hat{x}$. Note that each inducible bundle \hat{x} is a convex combination of n bundles (purchased by all the buyers) in \mathcal{F} , so it must lie in the set \mathcal{F} .

A centerpiece in our analysis is the following welfare maximization convex program that characterizes the relation between the posted prices and inducible bundles.

⁷For a set $C \subset \mathbb{R}^d$ and a norm $\|\cdot\|$, we write $\|C\| = \sup_{x \in C} \|x\|$. When the norm is unspecified, it is assumed to be ℓ_2 .

Definition 3.3. For any bundle $\hat{x} \in \mathcal{F}$, let convex program $\text{SCP}(\hat{x})$ be the following

$$\max_{x \in \mathcal{F}^n} \sum_{v_i \in \mathcal{V}} \psi(v_i) v_i(x_i) \quad (4)$$

$$\text{such that } \sum_{v_i \in \mathcal{V}} \psi(v_i) x_{ij} \leq \hat{x}_j \quad \text{for every } j \in [d] \quad (5)$$

$$x_i \in \mathcal{F} \quad \text{for every } v_i \in \mathcal{V} \quad (6)$$

Let $\text{VAL}(\hat{x})$ be the optimal value of the convex program $\text{SCP}(\hat{x})$. We also say that $\text{SCP}(\hat{x})$ is *supply-saturating* if its optimal solution x^\bullet saturates all of the supply constraints defined by eq. (5), that is $\sum_i \psi(v_i) x_{ij}^\bullet = \hat{x}_j$ for all j .

To interpret the above as a stochastic welfare maximization program, consider a market in which there are d types of goods and each good j has supply \hat{x}_j . For each valuation function $v_i \in \mathcal{V}$, we introduce a buyer i with this valuation, who shows up to the market with probability $\psi(v_i)$. We use a vector $x_i = (x_{i1}, \dots, x_{id}) \in \mathcal{F}$ to represent the bundle of goods allocated to a buyer i if he shows up. Then the program is precisely computing an allocation over all buyers to maximize the expected welfare subject to the constraint that the expected demand is no more than the supply given by \hat{x} .⁸

ASSUMPTION 3.4 (RELAXING MONOTONICITY IN VALUATIONS). *In fact, the assumption that each valuation in the class \mathcal{V} is increasing can be relaxed. Our algorithm works as long as the class \mathcal{V} and the feasible set \mathcal{F} guarantees that $\text{SCP}(\hat{x})$ is supply-saturating for any $\hat{x} \in \mathcal{F}$. For the sake of generality, our analysis will rely on the supply saturation condition instead of the monotonicity of the valuations. This will be useful for applying the algorithm to the indivisible goods setting.*

If the valuations in the class \mathcal{V} are increasing functions, the optimal solution of $\text{SCP}(\hat{x})$ will saturate all of the supply constraints in eq. (5).

Claim 3.5. *Suppose that each valuation $v \in \mathcal{V}$ is monotonically increasing in each coordinate. Then for any $\hat{x} \in \mathcal{F}$, the convex program $\text{SCP}(\hat{x})$ is supply-saturating.*

For each of the supply constraints in eq. (5), we can introduce a dual (price) variable p_j and write down the following *partial Lagrangian*

$$\mathcal{L}_{\hat{x}}(x, p) = \sum_{v_i \in \mathcal{V}} \psi(v_i) v_i(x_i) - \sum_{j=1}^d p_j \left(\sum_{v_i \in \mathcal{V}} \psi(v_i) x_{ij} - \hat{x}_j \right) \quad (7)$$

We can also consider the *Lagrange dual function* of the convex program $g_{\hat{x}}: \mathbb{R}^d \rightarrow \mathbb{R}$:

$$g_{\hat{x}}(p) = \max_{x \in \mathcal{F}^n} \mathcal{L}_{\hat{x}}(x, p). \quad (8)$$

We will mostly focus on the case where $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$, which we can show is a sufficient condition for inducibility.⁹

LEMMA 3.6. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$ be a bundle, then \hat{x} is inducible.*

⁸Similar construction of such stochastic convex programs also appeared in [20].

⁹The restriction that the bundle be positive in each coordinate is necessary – a bundle with zero in some coordinate may not be inducible. Consider the same simple setting in Theorem 1.4 where $d = 1$ and there is a single buyer with valuation $v(x) = \sqrt{x}$. Because the marginal valuation at 0 is infinity, there is no bounded price to induce the buyer to purchase 0 units of the good.

PROOF. Consider the convex program $\text{SCP}(\hat{x})$. Since the convex program satisfies the *Slater's condition*, strong duality gives

$$\max_{x \in \mathcal{F}^n} \min_{p \in \mathbb{R}_+^d} \mathcal{L}_{\hat{x}}(x, p) = \min_{p \in \mathbb{R}_+^d} \max_{x \in \mathcal{F}^n} \mathcal{L}_{\hat{x}}(x, p) = \text{VAL}(\hat{x})$$

Furthermore, since $\text{SCP}(\hat{x})$ is supply-saturating, the optimal solution satisfies $\mathbb{E}_{v_i \sim \psi} [x_i^\bullet] = \hat{x}$. Let p^\bullet be the optimal dual solution. It follows that

$$\begin{aligned} x^\bullet &= \operatorname{argmax}_{x \in \mathcal{F}^n} \mathcal{L}(x, p^\bullet) = \operatorname{argmax}_{x \in \mathcal{F}^n} \sum_i \psi(v_i) v_i(x_i) - \sum_{j=1}^d p_j^\bullet \left(\sum_{v_i \in \mathcal{V}} \psi(v_i) x_{ij} - \hat{x}_j \right) \\ &= \operatorname{argmax}_{x \in \mathcal{F}^n} \sum_i \psi(v_i) (v_i(x_i) - \langle p_j^\bullet, x_i \rangle + \langle p^\bullet, \hat{x} \rangle) \end{aligned}$$

Note that the expression inside the argmax is linearly separable across i . Therefore,

$$x_i^\bullet = \operatorname{argmax}_{x_i \in \mathcal{F}} [v_i(x_i) - \langle p^\bullet, x_i \rangle + \langle p^\bullet, \hat{x} \rangle] = \operatorname{argmax}_{x_i \in \mathcal{F}} [v_i(x_i) - \langle p^\bullet, x_i \rangle] \quad (9)$$

It follows that $x_i^\bullet = x_{v_i}^*(p^\bullet)$ for each i , and hence the price vector p^\bullet induces the bundle \hat{x} . \square

Next, we show that the prices that induce the bundle \hat{x} are an optimal solution of the Lagrangian dual, and the bundles purchased by each buyer in response to these prices form the unique primal optimal solution.

LEMMA 3.7. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{\geq 0}^d)$ be a bundle, and let $\hat{p} \in \mathbb{R}_+^d$ be a price vector such that $x_{\psi}^*(\hat{p}) = \hat{x}$. Then*

- the price vector \hat{p} is an optimal dual solution for $\text{SCP}(\hat{x})$, and
- the vector $x^\bullet \in \mathcal{F}^n$ such that $x_i^\bullet = x_{v_i}^*(\hat{p})$ for each i is the unique optimal primal solution.

A very nice consequence of Lemma 3.7 is that whenever the induced bundle \hat{x} is fixed, the realized bundles purchased by buyers of each type are also fixed. This allows us to express the expected social welfare as a function only of the induced bundle. In particular, the expected valuation for inducing \hat{x} in expectation is exactly $\text{VAL}(\hat{x})$. This suggests a different way to express the welfare: as a function of the induced bundle (as opposed to a function of the price vector defined in eq. (3)). For each $\hat{x} \in \mathcal{F}$, we can define

$$\text{SW}(\hat{x}) = \text{VAL}(\hat{x}) - \langle c, \hat{x} \rangle. \quad (10)$$

We can show that the expected social welfare for inducing \hat{x} in expectation is exactly $\text{SW}(\hat{x})$. More importantly, by rewriting the welfare as a function of the bundle, we obtain a *concave* objective function. This is crucial for us to obtain an efficient algorithm later.

LEMMA 3.8. *The expected social welfare function $\text{SW}: \mathcal{F} \rightarrow \mathbb{R}$ as defined in eq. (10) is concave.*

With all of structural results above, we are ready to give our two-layered algorithm for finding the welfare-maximizing prices.

3.2 Inner Layer: Converting Target Bundles to Prices

Even though we can express the expected welfare as a concave function of the induced bundle, we still cannot directly optimize the function because the seller only controls the prices of the goods instead of the expected induced bundle itself. To optimize over the bundle space, we give an algorithm that finds a price vector that approximately induces any target expected bundle \hat{x} . Specifically, suppose that the seller has some target bundle \hat{x} in mind, we can learn a price vector \hat{p} such that the expected induced bundle is close to the target bundle: $\|\hat{x} - x_{\psi}^*(\hat{p})\| \leq \varepsilon$.

In Lemma 3.7, we show that the prices that exactly induce the target bundle \hat{x} are the optimal dual solution for the convex program $\text{SCP}(\hat{x})$, which is the price vector p that minimizes the Lagrangian dual function $g_{\hat{x}}$. We will show that if we can find an approximate minimizer for $g_{\hat{x}}$, we can then approximately induce the target expected bundle \hat{x} . In particular, we will apply the noisy gradient descent method (Theorem 2.1) to minimize the function $g_{\hat{x}}$, and for the sake of convergence of the algorithm, we will restrict the search space for the price vector to be

$$\mathcal{P}(\varepsilon) = \left\{ p \in \mathbb{R}_+^d \mid \|p\|_2 \leq \sqrt{d}\lambda^{(1/\beta)} \left(\frac{4d}{\varepsilon^2\sigma} \right)^{(1-\beta)/\beta} \right\} \quad (11)$$

where ε is the target accuracy parameter. First, we will show that the minimax value of the Lagrangian remains close to $\text{VAL}(\hat{x})$ even when we restrict the dual variables/prices to be in $\mathcal{P}(\varepsilon)$.

LEMMA 3.9. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$. There exists a value R-OPT such that*

$$\max_{x \in \mathcal{F}^n} \min_{p \in \mathcal{P}(\varepsilon)} \mathcal{L}_{\hat{x}}(x, p) = \min_{p \in \mathcal{P}(\varepsilon)} \max_{x \in \mathcal{F}^n} \mathcal{L}_{\hat{x}}(x, p) = \text{R-OPT}$$

Moreover, $\text{VAL}(\hat{x}) \leq \text{R-OPT} \leq \text{VAL}(\hat{x}) + \frac{\varepsilon^2\sigma}{4}$.

The next result translates the approximation error in minimizing the function $g_{\hat{x}}$ to the error in inducing the target bundle \hat{x} by making use of the strong concavity of the valuations in \mathcal{V} .

LEMMA 3.10. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$ and p' be a price vector in $\mathcal{P}(\varepsilon)$ such that $g_{\hat{x}}(p') \leq \min_{p \in \mathcal{P}(\varepsilon)} g_{\hat{x}}(p) + \alpha$ for some $\alpha > 0$. Let $x' = x_{\psi}^*(p')$ be the expected bundle induced by prices p' . Then $\|x' - \hat{x}\|_2 \leq 2\sqrt{\alpha/\sigma}$.*

Therefore, in order to (approximately) induce a target bundle in expectation, we just need to compute an (approximate) minimizer for the Lagrangian dual function $g_{\hat{x}}$. We first show that we can compute an unbiased estimate of the gradient of $g_{\hat{x}}$ by using the observed bundle purchased by a random buyer.

LEMMA 3.11. *Let $p \in \mathbb{R}_+^d$ be any price vector, and $x_v^*(p)$ be bundle purchased by a buyer with valuation function v under prices p . Then*

$$\mathbb{E}_{v \sim \psi} [\hat{x} - x_v^*(p)] = \hat{x} - x_{\psi}^*(p) = \nabla g_{\hat{x}}(p).$$

The result of Theorem 3.11 shows that we can obtain unbiased estimates of the gradients of the function $g_{\hat{x}}$ at different prices, as long as we can obtain unbiased estimates for the expected demand $x_{\psi}^*(p)$. In the next section, we will give another technique to obtain unbiased estimates for the gradients. Given access to unbiased estimate of the gradients of $g_{\hat{x}}$, we can rely on the noisy subgradient descent method (and its guarantee in Theorem 2.1) to minimize the function $g_{\hat{x}}$. Note that the algorithm will only find a point that approximately minimizes the function in expectation, but we can get an approximate minimizer with high probability using a standard amplification technique — running the subgradient descent method for logarithmically many times, so that one of the output price vectors is guaranteed to be accurate with high probability. More formally:

LEMMA 3.12. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$ be any target bundle. There exists an algorithm that given any target accuracy ε and confidence parameter δ as input, outputs a list P of $\log(1/\delta)$ price vectors such that with probability at least $1 - \delta$, there exists a price vector $\hat{p} \in P$ that satisfies $\|x_{\psi}^*(\hat{p}) - \hat{x}\| \leq \varepsilon$. Furthermore, the running time, the length of the list and the number of queries to ReP is bounded by $\text{poly}(d, 1/\varepsilon, \log(1/\delta))$.*

Lastly, we have one remaining technical problem to solve: given a set of price vectors P in which at least one price vector can approximately induce the target expected bundle \hat{x} , we need to identify one such price vector. To accomplish this, we will simply post each price vector $p \in P$ repeatedly, to obtain polynomially many observations from the buyers and compute the empirical average bundles over these polynomially many rounds. We select the price vector whose empirical average purchased bundle is closest to the target bundle \hat{x} . Putting all the pieces together, we obtain our full algorithm BUNTOPRICE (formal description in the full version).

THEOREM 3.13. *Let $\hat{x} \in (\mathcal{F} \cap \mathbb{R}_{>0}^d)$ be any target bundle. For any target accuracy parameter ε and confidence parameter δ , the instantiation $\text{BUNTOPRICE}(\hat{x}, \varepsilon, \delta)$ outputs a price vector \hat{p} that with probability at least $1 - \delta$ satisfies $\|\hat{x} - x_{\psi}^*(\hat{p})\| \leq \varepsilon$. Furthermore, the number of queries to ReP is bounded by $\text{poly}(d, 1/\varepsilon, \log(1/\delta))$.*

3.3 Outer Layer: Welfare Maximization

Finally, we combine the subroutine BUNTOPRICE with subgradient descent to find the welfare maximizing prices. At a high level, we will use subgradient descent to optimize the function SW over the bundle space, and along the way use the algorithm BUNTOPRICE to obtain prices which induce each target bundle that arises along subgradient descent's optimization path. To ensure that the per-round expected demand for each good j is bounded by some supply s_j , the algorithm will optimize over bundles in the set $S = \{x \in \mathcal{F} \mid x_j \leq s_j \text{ for each } j \in [d]\}$.

There are several technical challenges remaining. First, in order to optimize the concave function SW using subgradient descent, we need to compute a subgradient for each bundle the subgradient descent method chooses at intermediate steps. The following result establishes a very nice property that the price vector that induces each target bundle \hat{x} gives us a simple way to compute a subgradient in $\partial \text{SW}(\hat{x})$. In particular, this means we can obtain a subgradient of the function SW using our subroutine BUNTOPRICE.

LEMMA 3.14. *Let $\hat{x} \in (\mathbb{R}_{>0}^d \cap \mathcal{F})$, and \hat{p} be the price vector that induces \hat{x} . Then $(\hat{p} - c) \in \partial \text{SW}(\hat{x})$.*

Second, at each iteration t , subgradient descent may require a subgradient at some bundle x^t , but because of the error in BUNTOPRICE, we only find prices to *approximately* induce the target bundle. To overcome this issue, we will rely on the analysis of subgradient descent under *adversarial noise* (given in Theorem 2.2).

Lastly, instead of optimizing over the entire set S , we will optimize over a slightly smaller set

$$S_{\xi} = \{x \in \mathcal{F} \mid \xi \leq x_j \leq s_j - \xi\}.$$

This allows us to settle two issues: (1) we can guarantee that all of the induced bundles lie in the set S despite the error of BUNTOPRICE and (2) each bundle in S_{ξ} is guaranteed to be inducible since it is strictly positive in every coordinate (as required by Theorem 3.13).

LEMMA 3.15. *For any $\xi \in (0, 1)$, $\max_{x \in S} \text{SW}(x) - \max_{x' \in S_{\xi}} \text{SW}(x') \leq \lambda(d\xi)^{\beta} + \sqrt{d\xi}\|c\|$.*

Putting all the pieces together, we obtain our main algorithm OWEL (full description presented in the full version). To establish the welfare guarantee of $\hat{\mathcal{D}}$, we will compare to an even stronger benchmark—the welfare of the optimal lottery over allocations. In particular, given any constraint vector $s \in \mathbb{R}_+^{d+1}$, a feasible lottery over allocations is a randomized mapping $\pi: [n] \rightarrow \mathcal{F}$ that assigns each buyer to a randomized bundle such that $\mathbb{E}_{v_i \sim \psi} [\pi(i)] \leq s$. Let OPT^{lot} be the optimal social welfare achieved by a lottery over allocations. The following is the formal guarantee of OWEL (corresponding to Theorem 1.1).

THEOREM 3.16. *For any accuracy parameter $\alpha > 0$, confidence parameter $\delta > 0$, and subset $S = \{x \in \mathcal{F} \mid x_j \leq s_j \text{ for each } j \in [d]\}$ given by a supply vector s . Given query access to ReP , the instantiation $\text{OWEL}(\alpha, \delta, s)$ outputs a price vector p' that with probability at least $1 - \delta$ satisfies $x_{\psi}^*(p') \leq s$ and*

$$\text{SW}(p') \geq \text{OPT}^{\text{lot}} - \alpha.$$

Furthermore, both the run-time of the algorithm and the number of queries to ReP is bounded by $\text{poly}(d, 1/\alpha, 1/\sigma, 1/\lambda, \log(1/\delta))$.

Remark. The only part of the algorithm that interacts with the oracle ReP (or the buyers) is BUNTOPRICE of the inner layer. Since the BUNTOPRICE only requires a bounded and unbiased estimate of $x_{\psi}^*(p)$ for each price vector p it queries, we can replace ReP by any procedure that can compute such unbiased estimate. This is crucial for solving the unit-demand problem.

4 UNIT-DEMAND BUYERS WITH INDIVISIBLE GOODS

We now switch to the setting of indivisible goods and unit-demand buyers. Our goal is to develop a computationally and query efficient algorithm to find an approximately optimal *distribution* over prices subject to the constraint that the per-round expected demand of each good j is bounded by s_j . In particular, we will use the algorithm OWEL in Section 3 as a main tool for our solution. Throughout, we impose the following mild boundedness assumption on the values.

Assumption on valuations. There exists a constant upper bound V_{\max} such that for any buyer $i \in [n]$ and item $j \in [d]$, $0 < v_{ij} \leq V_{\max}$.

Overview: relax and regularize. A natural starting point for solving our problem with OWEL is to consider the *linear relaxation* of unit demand valuations: that is we can view the buyers as having linear valuation functions over divisible goods, and optimizing over a feasible set of bundles that is the non-negative orthant of the ℓ_1 ball. This relaxation maintains the property that buyers buy integral quantities of each good. However, this approach runs into a substantial difficulty, because linear valuation functions are not *strongly* concave, and strong concavity was an important ingredient in our analysis of OWEL . Instead, we will imagine that we have access to a regularized version of the linear relaxation of our original problem: that is, we imagine that each buyer has a regularized valuation function of the form $\langle v, x \rangle + \eta H(x)$, where H is the entropy function (of course, in reality, we cannot modify the valuations of the buyers). We show that we can solve the (imagined) regularized version of the problem, and also that we can induce buyers to behave (in expectation) as if their valuation functions were regularized by appropriately perturbing the price vectors we present to them. Our solution then consists of the following three steps:

- (1) We show that the algorithm OWEL can compute an approximately optimal price vector for the regularized version of problem, as long as the algorithm has access to an unbiased estimate of the expected demand of a random regularized buyer.
- (2) Next, we show that we can obtain such unbiased estimates given access to the revealed preference oracle $\text{ReP}(\psi)$ in the original un-regularized instance. The key ingredient is a novel price perturbation technique.
- (3) Finally, we show how to construct an approximately optimal price distribution based on the price vector output by OWEL .

To facilitate the discussion, we will introduce a dummy good (indexed as $(d + 1)$) to represent the buyer's option of buying nothing. The seller's price and each buyer's value for this item is always 0, and the per-round demand upper bound is simply $s_{d+1} = 1$. Moreover, we will write Δ_{d+1} to denote the set of all probability distributions over the $(d + 1)$ items (or simply the simplex over

the items). For any $x \in \Delta_{d+1}$, the entropy of x is defined as $H(x) = \sum_{j \in [d+1]} x_j \log \frac{1}{x_j}$. The function H is strongly concave with respect to the ℓ_1 norm over the simplex.

4.1 Solving the regularized problem

Given a probability distribution ψ over value vectors and a parameter η , we imagine a corresponding η -regularized problem with a distribution $\tilde{\psi}$ over valuation functions: for each valuation vector v_i in the support of ψ , create a regularized valuation function $\tilde{v}_i: \Delta_{d+1} \rightarrow \mathbb{R}_{>0}$ with the same probability mass $\tilde{\psi}_{v_i} = \psi_{v_i}$ such that $\tilde{v}_i(x) = \langle v_i, x \rangle + \eta H(x)$.

The regularized problem is an instance of the divisible good setting, with feasible set $\mathcal{F} = \Delta_{d+1}$. Suppose that we have access to an unbiased estimate for $x_{\tilde{\psi}}^*(p)$ for any price vector p : then we can apply OWEL from Section 3 to compute approximately optimal prices. There is a small obstacle in the analysis—the regularized valuations defined above are not monotonically increasing in each coordinate (because of the entropy term). However, recall that we were able to substitute monotonicity for Assumption 3.4—that the convex program $\text{SCP}(\hat{x})$ is supply-saturating for each \hat{x} in our feasible region $\mathcal{F} = \Delta_{d+1}$. This is indeed satisfied in our setting:

LEMMA 4.1. *Let $\mathcal{F} = \Delta_{d+1}$ and $\tilde{\psi}$ be a distribution over the regularized valuation functions of the form $[\langle v, x \rangle + \eta H(x)]$. Then the convex program $\text{SCP}(\hat{x})$ defined as*

$$\max_{x \in \Delta_{d+1}} \sum_{v_i \in \mathcal{V}} \tilde{\psi}(v_i) [\langle v_i, x_i \rangle + \eta H(x_i)] \quad \text{such that} \quad \sum_{v_i \in \mathcal{V}} \tilde{\psi}(v_i) x_{ij} \leq \hat{x}_j \quad \text{for every } j \in [d+1]$$

is supply-saturating for any $\hat{x} \in \Delta_{d+1}$.

PROOF. Let x^\bullet be an optimal solution to $\text{SCP}(\hat{x})$. The convex combination $\mathbb{E}_{\tilde{\psi}} [x_i^\bullet] \in \Delta_{d+1}$, so we must have $\sum_{j \in [d+1]} \mathbb{E}_{\tilde{\psi}} [x_{ij}^\bullet] = 1$. Note that \hat{x} also lies in Δ_{d+1} (these are the only inducible average bundles for unit demand buyers), so $\sum_{j \in [d+1]} \hat{x}_j = 1$. It follows that all of the constraints $\mathbb{E}_{\tilde{\psi}} [x_{ij}^\bullet] \leq \hat{x}_j$ are saturated. \square

Moreover, the feasible set Δ_{d+1} is convex, closed, has a non-empty interior, and each of the regularized valuations is η -strongly concave with respect to the ℓ_2 norm,¹⁰ and $((\sqrt{d+1} + V_{\max}, 1/2)$ -Hölder continuous (see the full version for a proof). Thus we satisfy all conditions needed to apply OWEL, so long as we have access to unbiased estimates of $x_{\tilde{\psi}}^*(p)$ for any price vector $p \in \mathbb{R}_+^{d+1}$.

4.2 From price perturbation to value regularization

Solving the regularized problem using OWEL requires an unbiased estimate for $x_{\tilde{\psi}}(p)$ for any price vector p that the algorithm queries, but in our problem instance the valuations are actually drawn from ψ (without the entropy term). To obtain such an estimate, we give a price perturbation technique that allows us to simulate the response for the regularized buyers: given any price vector p , we will perturb each coordinate to obtain a noisy price vector p' , with the effect that the random item purchased by the unit-demand buyer in expectation over the perturbation equals the bundle purchased by her regularized counterpart.¹¹

¹⁰This follows from the fact that the ℓ_1 norm of any vector is bigger than its ℓ_2 norm.

¹¹The technique of simulating regularization through perturbation was also used to establish the equivalence between “Follow the Regularized Leader” and “Follow the Perturbed Leader” [34]. In our setting, it is important that we can obtain this effect by perturbing the price vector, rather than the valuation vector, because we do not have control over buyer valuations.

LEMMA 4.2 ([34]). Fix any $\eta > 0$ and any vector $u \in \mathbb{R}_{>0}^{d+1}$. For each $j \in [d+1]$, let G_j be a random number drawn independently and uniformly at random from $[0, 1]$, then

$$x^* = \operatorname{argmax}_{x \in \Delta_{d+1}} [\langle u, x \rangle + \eta H(x)] = \mathbb{E} \left[\operatorname{argmax}_{x \in \Delta_{d+1}} \left[\langle u, x \rangle + \eta \sum_{j \in [d+1]} x_j \ln(\ln(1/G_j)) \right] \right].$$

The random variable $\ln \ln(G_j)$ is distributed according to the *Gumbel* distribution. One immediate technical issue we have is that the Gumbel distribution is unbounded, so the perturbed prices might be negative. We also need to guarantee that the price on the dummy good \perp is 0. To overcome this issue, we will translate the perturbed price vector p into a non-negative price vector p' with the following procedure: set $p'_{d+1} = 0$ and $p'_j = (p_j - p_{d+1}) - \min\{0, \min_{j' \in [d]} (p_{j'} - p_{d+1})\}$. We will refer to this procedure as *Convert*, and show that the choice made by any unit-demand buyer remains the same under the new price vector p' .

LEMMA 4.3. Let $p \in \mathbb{R}^{d+1}$ any real-valued vector and let $p' = \text{Convert}(p)$. Then for any vector $v \in \mathbb{R}_{>0}^{d+1}$ such that $v_{d+1} = 0$,

$$\operatorname{argmax}_{j \in [d+1]} [v_j - p_j] = \operatorname{argmax}_{j \in [d+1]} [v_j - p'_j]$$

Furthermore, $p'_j \geq 0$ for all $j \in [d]$.

Combining the Gumbel noise addition and the procedure *Convert*, we can now obtain unbiased estimates for $x_{\psi}^*(p)$ using feedback from $\text{ReP}(\psi)$, the revealed preference oracle for the original problem instance. More formally, for any fixed price vector p , consider the following distribution $\mathcal{D}(p)$ of random prices:

- (1) For each $j \in [d+1]$, let G_j be a random number drawn independently and uniformly at random from $[0, 1]$, and let $\tilde{p}_j = p_j + \ln \ln(G_j)$ for each $j \in [d+1]$.
- (2) Output $p' = \text{Convert}(\tilde{p})$.

Given this subroutine for generating random prices, we have a procedure *SIM* (presented in Algorithm 1) for obtaining an unbiased estimate of $x_{\psi}^*(p)$ for any price vector p . We will establish the correctness of *SIM* in Theorem 4.4.

Algorithm 1 Simulate the response of a regularized buyer $\text{SIM}(p, \eta)$

Input: A price vector p and regularization parameter η

Let p' a price vector drawn from $\mathcal{D}(p)$

Query the oracle $\text{ReP}(\psi)$ with p' (i.e. post p' to a random consumer) and obtain feedback x'

Output: x' as an unbiased estimate for $x_{\psi}^*(p)$

LEMMA 4.4. Fix any price vector $p \in \mathbb{R}_+^{d+1}$, parameter $\eta > 0$, and any distribution ψ over value vectors in $\mathbb{R}_{>0}^{d+1}$. Let x' be the estimate output by $\text{SIM}(p, \eta)$, then $\mathbb{E}_{\psi, \text{SIM}} [x'] = x_{\psi}^*(p)$.

Using the subroutine *SIM* to obtain unbiased estimate of $x_{\psi}^*(p)$ for different price vectors p , we can now instantiate *OWEL* to solve the η -regularized problem.

4.3 Wrap-up: An approximately optimal distribution over prices

Let \hat{p} be the price vector output by OWEL when solving the η -regularized problem, and consider the distribution over prices $\hat{\mathcal{D}} = \mathcal{D}(\hat{p})$. Similar to the divisible goods setting, we will again compare to the optimal lottery over allocations. In particular, the optimal lottery is given by

$$\max_{x \in \Delta_{d+1}^n} \sum_{i \in [n]} \psi(v_i) [\langle v_i - c, x_i \rangle] \quad \text{such that} \quad \sum_{i \in [n]} \psi(v_i) x_{ij} \leq s_j \quad \text{for each } j \in [d+1]$$

Let x^* and OPT^{lot} be the optimal solution and value for the program defined above. Since any distribution of prices is just inducing a lottery over allocations, we know that

$$\max_{\mathcal{D} \text{ s.t. } x_{\psi}^*(\mathcal{D}) \leq s} \text{SW}(\mathcal{D}) \leq \text{OPT}^{\text{lot}}$$

where we write $x_{\psi}^*(\mathcal{D}) = \mathbb{E}_{p \sim \mathcal{D}, v \sim \psi} [x_v^*(p)]$ to denote expected demand over the goods.

The following lemma bounds the sub-optimality of $\hat{\mathcal{D}}$ compared to OPT^{lot} in terms of the regularization parameter η and accuracy guarantee of the price vector \hat{p} .

LEMMA 4.5. *Fix any regularization parameter η . Suppose that \hat{p} is an ε -approximately optimal price vector for the η -regularized problem. Then the distribution of prices $\hat{\mathcal{D}} = \mathcal{D}(\hat{p})$ satisfies*

$$\text{SW}(\hat{\mathcal{D}}) \geq \text{OPT}^{\text{lot}} - \varepsilon - \eta \log(d+1).$$

Therefore, to achieve a target accuracy of α , it suffices to instantiate OWEL with accuracy parameter $\varepsilon = \alpha/2$ to solve the η -regularized problem with $\eta = \alpha/(2 \log(d+1))$. Putting all the pieces together, we have our algorithm OWEL-UD (formally presented in the version) that achieves the following main result, which recovers Theorem 1.2.

THEOREM 4.6. *For any accuracy parameter $\alpha > 0$, confidence parameter $\delta > 0$, and subset $S = \{x \in \mathcal{F} \mid x_j \leq s_j \text{ for each } j \in [d]\}$ given by a supply vector s . Given query access to ReP, the instantiation OWEL-UD(α, δ, s) outputs a distribution $\hat{\mathcal{D}}$ over prices such that with probability at least $1 - \delta$ satisfies $x_{\psi}^*(\hat{\mathcal{D}}) \leq s$ and*

$$\text{SW}(\hat{\mathcal{D}}) \geq \text{OPT}^{\text{lot}} - \alpha.$$

Furthermore, both the run-time of the algorithm and the number of queries to ReP is bounded by $\text{poly}(d, 1/\alpha, V_{\max}, \log(1/\delta))$.

REFERENCES

- [1] Sydney Afriat. 1967. The Construction of Utility Functions from Expenditure Data. *International Economic Review* 8, 1 (1967), 67–77.
- [2] Alekh Agarwal, Dean P. Foster, Daniel J. Hsu, Sham M. Kakade, and Alexander Rakhlin. 2013. Stochastic Convex Optimization with Bandit Feedback. *SIAM Journal on Optimization* 23, 1 (2013), 213–240.
- [3] Kareem Amin, Rachel Cummings, Lili Dworkin, Michael Kearns, and Aaron Roth. 2015. Online Learning and Profit Maximization from Revealed Preferences. In *Twenty-Ninth AAAI Conference on Artificial Intelligence (AAAI-15)*.
- [4] Moshe Babaioff, Shaddin Dughmi, Robert D. Kleinberg, and Aleksandrs Slivkins. 2015. Dynamic Pricing with Limited Supply. *ACM Trans. on Economics and Computation* 3, 1 (2015), 4. Special issue for *13th ACM EC*, 2012.
- [5] Ashwinkumar Badanidiyuru, Robert Kleinberg, and Aleksandrs Slivkins. 2013. Bandits with Knapsacks. In *54th IEEE Symp. on Foundations of Computer Science (FOCS)*.
- [6] Maria-Florina Balcan, Avrim Blum, and Yishay Mansour. 2008. Item pricing for revenue maximization. In *9th ACM Conf. on Electronic Commerce (EC)*. 50–59.
- [7] Maria-Florina Balcan, Florin Constantin, Satoru Iwata, and Lei Wang. 2012. Learning Valuation Functions.. In *23rd*.
- [8] Maria-Florina Balcan, Amit Daniely, Ruta Mehta, Ruth Urner, and Vijay V Vazirani. 2014. Learning Economic Parameters from Revealed Preferences. In *Web and Internet Economics*. Springer, 338–353.

- [9] Eyal Beigman and Rakesh Vohra. 2006. Learning from revealed preference. In *Proceedings of the 7th ACM Conference on Electronic Commerce*. ACM, 36–42.
- [10] Omar Besbes and Assaf Zeevi. 2009. Dynamic Pricing Without Knowing the Demand Function: Risk Bounds and Near-Optimal Algorithms. *Operations Research* 57 (2009), 1407–1420. Issue 6.
- [11] Omar Besbes and Assaf J. Zeevi. 2012. Blind Network Revenue Management. *Operations Research* 60, 6 (2012), 1537–1550.
- [12] Avrim Blum, Anupam Gupta, Yishay Mansour, and Ankit Sharma. 2011. Welfare and Profit Maximization with Production Costs. In *52nd IEEE Symp. on Foundations of Computer Science (FOCS)*. 77–86.
- [13] Arnoud V. Den Boer. 2015. Dynamic pricing and learning: Historical origins, current research, and new directions. *Surveys in Operations Research and Management Science* 20, 1 (June 2015).
- [14] Josef Broder and Paat Rusmevichientong. 2012. Dynamic Pricing Under a General Parametric Choice Model. *Operations Research* 60, 4 (2012), 965–980.
- [15] Sébastien Bubeck. 2015. Convex Optimization: Algorithms and Complexity. *Foundations and Trends in Machine Learning* 8, 3-4 (2015), 231–357.
- [16] Sébastien Bubeck and Nicolo Cesa-Bianchi. 2012. Regret Analysis of Stochastic and Nonstochastic Multi-armed Bandit Problems. *Foundations and Trends in Machine Learning* 5, 1 (2012).
- [17] Sébastien Bubeck, Ofer Dekel, Tomer Koren, and Yuval Peres. 2015. Bandit Convex Optimization: \sqrt{T} Regret in One Dimension. In *28th Conf. on Learning Theory (COLT)*. 266–278.
- [18] Tanmoy Chakraborty, Zhiyi Huang, and Sanjeev Khanna. 2013. Dynamic and Nonuniform Pricing Strategies for Revenue Maximization. *SIAM J. on Computing (SICOMP)* 42, 6 (2013), 2424–2451. Preliminary version in *IEEE FOCS 2009*.
- [19] Arnoud V. den Boer and Bert Zwart. 2014. Simultaneously Learning and Optimizing Using Controlled Variance Pricing. *Management Science* 60, 3 (2014), 770–783.
- [20] Nikhil R. Devanur, Balasubramanian Sivan, and Yossi Azar. 2012. Asymptotically optimal algorithm for stochastic adwords. In *ACM Conference on Electronic Commerce, EC '12, Valencia, Spain, June 4-8, 2012*. 388–404.
- [21] Michal Feldman, Nick Gravin, and Brendan Lucier. 2015. Combinatorial Auctions via Posted Prices. In *26th ACM-SIAM Symp. on Discrete Algorithms (SODA)*. 123–135.
- [22] Abraham Flaxman, Adam Kalai, and H. Brendan McMahan. 2005. Online Convex Optimization in the Bandit Setting: Gradient Descent without a Gradient. In *16th ACM-SIAM Symp. on Discrete Algorithms (SODA)*. 385–394.
- [23] J. C. Gittins. 1989. *Multi-Armed Bandit Allocation Indices*. John Wiley & Sons.
- [24] Elad Hazan and Kfir Y. Levy. 2014. Bandit Convex Optimization: Towards Tight Bounds. In *27th Advances in Neural Information Processing Systems (NIPS)*. 784–792.
- [25] Justin Hsu, Jamie Morgenstern, Ryan Rogers, Aaron Roth, and Rakesh Vohra. 2016. Do prices coordinate markets?. In *Proceedings of the 48th Annual ACM SIGACT Symposium on Theory of Computing*. ACM, 440–453.
- [26] N. Bora Keskin and Assaf J. Zeevi. 2014. Dynamic Pricing with an Unknown Demand Model: Asymptotically Optimal Semi-Myopic Policies. *Operations Research* 62, 5 (2014), 1142–1167.
- [27] Robert Kleinberg and Tom Leighton. 2003. The Value of Knowing a Demand Curve: Bounds on Regret for Online Posted-Price Auctions.. In *44th IEEE Symp. on Foundations of Computer Science (FOCS)*. 594–605.
- [28] Andreu Mas-Colell, Michael D. Whinston, and Jerry R. Green. 1995. *Microeconomic Theory*. Oxford University Press.
- [29] Aaron Roth, Jonathan Ullman, and Zhiwei Steven Wu. 2016. Watch and learn: optimizing from revealed preferences feedback. In *Proc. of the 48th Annual ACM SIGACT Symp. on Theory of Computing, STOC 2016*.
- [30] Ariel Rubinstein. 2012. *Lecture notes in microeconomic theory: the economic agent*. Princeton University Press.
- [31] Paul A. Samuelson. 1938. A Note on the Pure Theory of Consumers' Behavior. *Economica* 5, 17 (1938), 61–71.
- [32] Hal R. Varian. 2006. Revealed preference. In *Samuelsonian economics and the twenty-first century*, Michael Szenberg, Lall Ramrattan, and Aron A. Gottesman (Eds.). Oxford University Press, 99–115.
- [33] Zizhuo Wang, Shiming Deng, and Yinyu Ye. 2014. Close the Gaps: A Learning-While-Doing Algorithm for Single-Product Revenue Management Problems. *Operations Research* 62, 2 (2014), 318–331.
- [34] Manfred Warmuth. 2009. A perturbation that makes "Follow the leader" equivalent to "Randomized weighted majority". (2009).
- [35] Morteza Zadimoghaddam and Aaron Roth. 2012. Efficiently Learning from Revealed Preference. In *Internet and Network Economics - 8th International Workshop, WINE 2012 (Lecture Notes in Computer Science)*, Vol. 7695. Springer, 114–127.
- [36] Martin Zinkevich. 2003. Online Convex Programming and Generalized Infinitesimal Gradient Ascent. In *Machine Learning, Proceedings of the Twentieth International Conference (ICML 2003), August 21-24, 2003, Washington, DC, USA*. 928–936.