# Pricing Loss Leaders can be Hard

Yi Wu
IBM Almaden Research, San Jose 95120, CA, U.S.A
wuyi@us.ibm.com

**Abstract:** Consider the problem of pricing $n$ items under an unlimited supply with $m$ single minded buyers, each of which is interested in at most $k$ of the items. The goal is to price each item with profit margin $p_1, p_2, \ldots, p_n$ so as to maximize the overall profit. There is an $O(k)$-approximation algorithm by [BB06] when the price on each item must be above its margin cost; i.e., each $p_i > 0$.

We investigate the above problem when the seller is allowed to price some of the items below their margin cost. It was shown in [BB06, BBCH08] that by pricing some of the items below cost, the seller could possibly increase the maximum profit by $\Omega(\log n)$ times. These items sold at low prices to stimulate other profitable sales are usually called "*loss leader*". It is unclear what kind of approximation guarantees are achievable when some of the items can be priced below cost. Understanding this question is posed as an open problem in [BB06].

In this paper, we give a strong negative result for the problem of pricing loss leaders . We prove that assuming the Unique Games Conjecture (UGC) [Kho02], there is no constant approximation algorithm for item pricing with prices below cost allowed even when each customer is interested in at most 3 items.

Conceptually, our result indicates that *although it is possible to make more money by selling some items below their margin cost, it can be computationally intractable to do so.*

**Keywords:** Complexity Theory; Game Theory; Approximation Algorithm; and Unique Games Conjecture

## 1 Introduction

We study the following item pricing problem. A seller has an infinite supply of $n$ different items. There are $m$ buyers, each of which are interested in a subset of the items with certain budget limit. These buyers are all *single minded*; i.e., they either buy all the items they are interested in if the overall cost is within their budget or they will buy none of them. The algorithmic task is to price each item $i$ with a profit margin $p_i$ to maximize the overall profit of the seller.

Serval results were known when the profit margin $p_i$ on each item is required to be *positive*. A $O(\log n + \log m)$ approximation for the general problems is given by Guruswami *et al.* [5]. If we assume that each customer is only interested in a constant number $k$ of the items, a $O(k^2)$-approximation algorithm was given in [3] by Briest and Krysta. Later in [1], Balcan and Blum improved the approximation ratio to $O(k)$. In particular, when $k = 2$ (such a problem is also called *graph vertex pricing*), their algorithm gave an $4$-approximation. On the hardness side, an APX-hardness result was obtained for the general problem in [5]. Later, Demaine, Feige, Hajiaghayi, and Salavatipour obtained a poly-logarithmic hardness [4]. As for the case that each customer is only interested in at most $2$ of the items, a $2$-hardness result was obtained in [6] assuming the Unique Games Conjecture (UGC).

Much less is known when the seller is allowed to assign negative profit margin $p_i$ for some of the items. The motivation behind selling some items below the margin cost is to increase the overall profit by stimulating the sales of other products. These items sold below the cost are usually referred as the "*loss leaders*". One example of the loss leaders is that in the market of digital book reader (such as the Kindle and IPad), the seller may price the reading device at a low price so as to make more money on the sales of the digital books.

Studying the problem of pricing loss leaders is formulated as an open problem in [1]; the authors asked: " what kind of approximation guarantees are achievable if one allows the seller to price some items below their margin cost?" Interestingly, the authors found that by optimally pricing some of the items below cost, one could possibly achieve a profit that is $\Omega(\log n)$ times of the maximum profit under the positive price model. The problem of pricing loss leaders is further studied by Balcan *et al.* in [2]. They introduced two new models: the *coupon* and *discount* model. Roughly speaking, the discount model is the item pricing problem with negative profit margin allowed; the coupon model adds an additional assumption that a seller's profit is at least $0$ for the entire transactions with each customer. The same $\Omega(\log n)$ "profitability gap" was shown under these models.

In this paper, we give a negative result for pricing loss leaders. In particular, we show that obtain a *constant approximation* for item pricing, under either the coupon or discount model, is NP-hard assuming the Unique Games Conjectures; our hardness result holds even for the very simple case that each customer is only interested at most $k = 3$ items. Our result should be compared with the case when only positive prices is assigned, there is an $\frac{1}{3e}$-approximation for such a problem. In addition, given a item pricing instances, we show that it is hard to distinguish whether it has a big profitability gap. Formally assuming the UGC and for $k = 3$, we show that finding out whether the profitability gap (either under the coupon or the discount model) is above $\alpha$ for any $\alpha > 6$ is NP-hard. Therefore, our result also indicates the hardness of finding a loss-leader pricing strategy, not necessary optimal, that is substantially larger than the maximum profit using positive prices only.

Conceptually, our results convey the following message: *although it is possible to make more money by selling items below their cost, it can be computationally intractable to do so.*

### 1.1 Problem definitions

The item pricing problem is also called the VERTEX-PRICING problem; it can be defined on a graph where each customer is corresponding to a hyperedge and each item to price is corresponding to a vertex. Let us start by formally define the following VERTEX-PRICING problem.

**Definition 1.1.** *(*VERTEX-PRICING*) A vertex pricing problem is specified by the tuple*

$$(G(V,E), \{b_e \mid e \in E\})$$

*Here $G(V,E)$ is a multigraph where each vertex $v_i \in V$ represents an item. Each hyperedge $e \in E$ represents a set of items (vertices) that a particular customer is interested with the budget $b_e$ ($b_e > 0$).*

*For the purpose of normalization, let us also assume that the minimum budget is 1.*

When the corresponding graph is $k$-hypergraph (i.e., each customer is interested in at most $k$ items), we call the problem VERTEX-PRICING$_k$.

**Definition 1.2.** *Given a* VERTEX-PRICING *instance $\mathcal{I}$, and a price function $p : V \to \mathbb{R}$, the profit is defined as follows:*

$$\mathbf{profit}_{\mathcal{I}}(p) = \sum_{b_e \geq price(e)} \mathbf{price}(e)$$

*where* $\mathbf{price}(e) = \sum_{v \in e} p(v)$.

When we restrict the range of the price function $p$, we get the positive price model, as well as the discount model, coupon model and $B$-bounded model that is introduced in [2]

**Definition 1.3.** *Given a instance $\mathcal{I}$ of* VERTEX-PRICING*:*

*For the positive price model, the objective function is*

$$\text{Opt}_{pos} = \max_{p:V \to R^+} \mathbf{profit}_{\mathcal{I}}(p).$$

*For the discount model, the objective function is*

$$\text{Opt}_{disc} = \max_{p:V \to R} \mathbf{profit}_{\mathcal{I}}(p)$$

*For the $B$-bounded coupon model, the objective function is*

$$\text{Opt}_{B} = \max_{p:V \to [-B,\infty)} \mathbf{profit}_{\mathcal{I}}(p)$$

The $B$-bounded model applies to the case that each item has the same margin cost $B$ and the seller could not price the profit margin below $-B$. The authors in [2] also defined the coupon model which assumes that the profit is at least 0 for each sale with the customer.

**Definition 1.4.** *Given a instance $\mathcal{I}$ of* VERTEX-PRICING*, the profit under coupon model is defined as*

$$\mathbf{profit}^+_{\mathcal{I}}(p) = \sum_{b_e \geq price(e)} \max(\mathbf{price}(e), 0)$$

*and the objective function is the following:*

$$\mathrm{Opt}_{coup} = \max_{p:V\to\mathbb{R}} \mathbf{profit}^+(p)$$

It is easy to see the following relationship among these models.

**Fact 1.5.** *For any $B > 0$ and a* VERTEX-PRICING *instance $\mathcal{I}$,*

$$\mathrm{Opt}_{pos} \leq \mathrm{Opt}_B \leq \mathrm{Opt}_{disc} \leq \mathrm{Opt}_{coup}.$$

We also define the profitability gap as the ratio between the optimum profit under these negative profit models and positive profit model.

**Definition 1.6.** *We define the profitability gaps as follows:*

- $\mathrm{Gap}_B = \frac{\mathrm{Opt}_B}{\mathrm{Opt}_{pos}}$.
- $\mathrm{Gap}_{disc} = \frac{\mathrm{Opt}_{disc}}{\mathrm{Opt}_{pos}}$.
- $\mathrm{Gap}_{coup} = \frac{\mathrm{Opt}_{disc}}{\mathrm{Opt}_{pos}}$.

### 1.2 Main result

Our main result is the following theorem:

**Theorem 1.7.** *Assuming the UGC, given a* VERTEX-PRICING$_3$ *instance. Then for any positive integer $B$, it is NP-hard to distinguish the following two cases:*

- $\mathrm{Opt}_B \geq \Omega(\log B)$*;*
- $\mathrm{Opt}_{coup} \leq 6 + o(1)$.

*Above decision problem is is hard even under the additional assumption that $3 \geq \mathrm{Opt}_{pos} \geq 1$.*

Using fact (1.5) and taking $B = 2^{\Omega(\alpha)}$, we get the following corollaries:

**Corollary 1.8.** *(Hardness for coupon model) Assuming the Unique Games Conjecture, for any constant $\alpha > 0$,* VERTEX-PRICING$_3$ *under the coupon model is NP-hard to $\alpha$-approximate.*

**Corollary 1.9.** *(Hardness for discount model) Assuming the Unique Games Conjecture, for any constant $\alpha > 0$,* VERTEX-PRICING$_3$ *under the discount model is NP-hard to $\alpha$-approximate.*

**Corollary 1.10.** *(Hardness for B-bounded model) Assuming the Unique Games Conjecture,* VERTEX-PRICING$_3$ *under the B-bounded model is NP-hard to $\Omega(\log B)$-approximate.*

**Corollary 1.11.** *(Hardness for deciding profitability gap) Assuming the Unique Games Conjecture and for any $\alpha > 0$, it is NP-hard to tell whether the profitability gap (under either the coupon or the discount modes) is above $\alpha$ or below $6 + o(1)$.*

## 2 Preliminaries

### 2.1 Mathematical tools

**Notations:** For $m \in R^+$, we use $[m]$ to denote the set $\{1, 2, \ldots, \lfloor m \rfloor\}$ (This is slightly non-standard as we usually use $[m]$ to denote$\{0, 1, \ldots, m-1\}$ for $m \in Z^+$). For $q$ being an integer, we use the notation $\oplus_q$ to denote the addition of two numbers (or vectors) modulo $q$. We use $\mathbf{1}(\cdot)$ to denote the indicator function.

Our proof relies on tools from Harmonic analysis of Discrete functions. Here we make a quick review. For a complete introduction, one can check [7, 12]. We will be considering functions of the form $f : [q]^n \to \mathbb{R}^t$ where $q, n, t \in \mathbb{N}$. We denote $f = (f^1, f^2, \ldots, f^t)$ where $f^i$ is the $i$-th coordinate of $f$. The set of all functions $f : [q]^n \to \mathbb{R}^t$ forms an inner product space with inner product

$$\langle f, g \rangle = \mathop{\mathbf{E}}_{x\sim[q]^n}[f(x) \cdot g(x)];$$

Here $x$ is uniformly random over $[q]^n$ and $f(x) \cdot g(x)$ is the usual vector inner product. We also write $\|f\| = \sqrt{\langle f, f \rangle}$.

For a given $x$, we say $y$ is $\rho$-correlated with $x$ if $y$ is generated by setting each $y_i = x_i$ with probability $\rho$ and a random number from $[q]$ with probability $1 - \rho$.

For $0 \le \rho \le 1$, we define $T_\rho$ to be the linear operator on this inner product space given by

$$T_\rho f(x) = \mathop{\mathbf{E}}_y[f(y)],$$

where $y$ is a random string in $[q]^n$ which is $\rho$-correlated to $x$.

For $i \in [n]$, we define the influence of $i$ on $f : [q]^n \to \mathbb{R}$ to be

$$\mathrm{Inf}_i[f] = \mathop{\mathbf{E}}_{x_1,\dots,x_{i-1},x_{i+1},\dots,x_n \sim [q]} \left[\mathbf{Var}_{x_i \sim [q]}[f(x)]\right],$$

where $\mathbf{Var}[f]$ is defined to be $\mathbf{E}[\|f\|^2] - \|\mathbf{E}[f]\|^2$. More generally, for $0 \le \eta \le 1$ we define the $\eta$-*noisy-influence of* $i$ *on* $f$ to be

$$\mathrm{Inf}_i^{(1-\eta)}[f] = \mathrm{Inf}_i[T_{1-\eta}f].$$

One may observe that $\sum_{j=1}^n \mathrm{Inf}_i^{1-\eta} f^j = \mathrm{Inf}_i^{1-\eta} f$.

Following facts are well known:

**Fact 2.1.** *For any* $\eta$*,*

$$\sum_{i=1}^n \mathrm{Inf}_i^{1-\eta}(f) \le \frac{\mathbf{Var}(f)}{2e\eta}$$

We also need the following "convexity of noisy-influences" fact:

**Fact 2.2.** *Let* $f^1, \dots, f^t$ *be a collection of functions* $[q]^n \to \mathbb{R}^k$*. Then*

$$\mathrm{Inf}_i^{(1-\eta)}\left[\mathop{\mathit{avg}}_{m\in[t]}\left\{f^{(m)}\right\}\right] \le \mathop{\mathit{avg}}_{m\in[t]}\left\{\mathrm{Inf}_i^{(1-\eta)}[f^{(m)}]\right\}$$

Here for any $c_1, c_2, \dots c_m \in \mathbb{R}$ (or $\mathbb{R}^k$), we use the notation avg$(c_1, \dots, c_m)$ to denote their average:

$$\frac{\sum_{m=1}^t c_t}{t}.$$

One major advanced tool we need in our analysis is the following theorem that is essentially from [9]. Here we use one of its variants appeared in [12].

**Theorem 2.3.** *Let* $(\Omega = [q]^t, \mu)$ *be a finite probability spaces with the following properties:*

- $a = (a_1, a_2, \dots, a_t) \sim \mu$ *are* pairwise independent.
- $\alpha = \min_{a\in\Omega} \mu(a) > 0$.

*For* $\eta > 0$ *and* $f = (f^1, \dots f^t) : \Omega^n \to [0,1]^t$ *be function satisfying that for any* $i \in [n], j \in [k]$ *and some constant* $\tau > 0$*,*

$$\mathrm{Inf}_i^{1-\eta} f^j \le \tau$$

*Then*

$$\mathbf{E}[\prod_{i=1}^t T_{1-\eta}f^{(i)}] - \prod_{i=1}^t \mathbf{E}[f^{(i)}] \le \tau^{C_0\eta/\log(1/\alpha))}$$

*Here* $C_0$ *is a constant that only dependent on* $t$*. The expectation is taken with respect to the product distribution* $(\Omega, \mu)^n$*.*

Roughly speaking, above theorem states that for calculating the product of $t$ different functions, if these functions do not have big noisy influence on each coordinate, then the product of them is the essentially the same under any pairwise independent distribution or the fully independent distribution.

### 2.2 Relationship between Dictator Test and hardness of approximation

Let us start by thinking of the VERTEX-PRICING as defined on a weighted multigraph graph; i.e. each edge $e$ in the hypergraph has a certain positive weight and the objective function is the weighted sum of the profit obtained on each edge. As we shall show later in section 4.3, the hardness result we obtain for the weighted vertex pricing problem also hold for the unweighted version VERTEX-PRICING.

The weighted VERTEX-PRICING$_3$ problem can be viewed as a 3-CSP over a set of variables $p_1, p_2, \dots p_n$ and a set of constraints specified by the budget $b_{ijk}$ with weight $w_{ijk}$[1]. For example for the VERTEX-PRICING$_3$ problem under the discount model, the payoff function on $b_{ijk}$ is

$$\begin{aligned}&\mathbf{revenue}(p_i, p_j, p_k, w_{ijk}) \\ &\quad = \mathbf{1}(p_i + p_j + p_k \le b_{ijk})(p_i + p_j + p_k).\end{aligned}$$

The goal is to find $p_1, p_2, \dots p_n$ to maximize the overall profit:

$$\sum_{i,j,k} w_{ijk} \cdot \mathbf{revenue}(p_i, p_j, p_k, b_{ijk})$$

[1] strictly speaking, for each $(i, j, k)$, there can be different $b_{ijk}$ with different weights $w_{ijk}$.

The work of Khot, Kindler, Mossel, and O'Donnell [7] introduced a now-standard methodology for proving hardness results for weighted CSPs based on the Unique Games Conjecture: namely, the construction of Dictator vs. Small Noisy-Influences Tests.

Formally speaking, a test for functions f with domain $[q]^n$ is an explicit instance $\mathcal{T}$ of a VERTEX-PRICING$_3$ with variable set $[q]^n$. It is given in the form of a probability distribution over $(x, y, z, w) \sim [q]^n \times [q]^n \times [q]^n \times \mathbb{R}^+$ ,where the probability here can be thought of the weight put on the constraint associated with $(x, y, z, w)$. For a given price function $f : [q]^n \to \mathbb{R}$ , we define its profit to be

$$\begin{aligned}&\mathbf{profit}_{\mathcal{T}}(f)\\&= \mathbf{E}_{x,y,z,w}[\mathbf{revenue}(f(x), f(y), f(z), w)].\end{aligned}$$

We may now informally state what a *Dictator vs. Small Noisy-Influences Test* is. It is a test for functions $f : [q]^n \to \mathbb{R}$ with the following two properties: (i) *Dictator functions* — i.e., functions of the form $h(x_i)$ for a particular function $h : [q] \to \mathbb{R}$ and each $i \in [n]$ — pass the test with high $\mathbf{profit}_{\mathcal{T}}(f) = c$;[2] (ii) Functions $f$ that is of "low noisy influence" on each coordinate pass the test with low $\mathbf{profit}_{\mathcal{T}}(f) = s$. Then roughly speaking, by the technique of [7], we can show that assuming the UGC, it is NP-hard to distinguish whether a VERTEX-PRICING$_3$ instance with profit above $c$ or below $s$ (which directly implies a hardness of approximation ratio $s/c$).

Above is the description of the Dictator Test for the discount model. As for the coupon model, the Dictator Test is essentially of the same except the pay off function is defined as

$$\begin{aligned}&\mathbf{revenue}^+(p_i, p_j, p_k, w_{ijk})\\&= \mathbf{1}(p_i+p_j+p_k \le w_{ijk})\cdot \max(p_i+p_j+p_k, 0).\end{aligned}$$

and the profit of a function $f$ is defined as

$$\begin{aligned}&\mathbf{profit}^+_{\mathcal{T}}(f)\\&= \mathbf{E}_{x,y,z,w}[\mathbf{revenue}^+(f(x)+f(y)+f(z), w)].\end{aligned}$$

In the rest of the paper, we first design and analyze a proper Dictator Test for VERTEX-PRICING$_3$. With such a test, we then apply the reduction of [7]. We want to emphasize here that we can not directly use results from [7] as the variables in VERTEX-PRICING is unbounded. The same problem also occurs in [11] on proving a hardness result for Unique Games over integer domain; the authors handle it through a improved analysis with the "hyper-contractive inequality". In comparison, the proof in this paper can be viewed improved as an analysis with the "invariance principle" [10] over unbounded functions.

[2] usually $h(t) = t$ for most of the previous work.

# 3 Dictator Test for VERTEX-PRICING

## 3.1 Description of the Dictator Test

To introduce our Dictator Test as well as analyzing it, first let us define the following distributions $\mathcal{D}_0, \mathcal{D}_1, \mathcal{D}_2$ on $(x, y, z) \in [q]^n \times [q]^n \times [q]^n$. We also assume here that $\sqrt{q}$ is an integer.

**Definition 3.1.** *(Distribution $\mathcal{D}_0$) Choose $x, y$ uniform randomly and independently from $[q]^n$; for each $i$, we have that*

- $z_i = q - (x_i + y_i)$ *if* $x_i + y_i < q$.
- $z_i = 2q - (x_i + y_i)$ *if* $q \le x_i + y_i \le 2q$.

By definition, we know that $x_i + y_i + z_i = 0 \mod q$ for each $i$. One important property of above distribution is that $(x_i, y_i, z_i)$ for each $i$ are *pairwise independent*.

**Definition 3.2.** *(Distribution $\mathcal{D}_1$) For $x, y, z \sim \mathcal{D}_0$,, Let $x', y', z'$ be $1 - \epsilon$ correlated with $x, y, z$. We call the corresponding distribution on $x', y', z'$ as $\mathcal{D}_1$*

**Definition 3.3.** *(Distribution $\mathcal{D}_2$) Choose $x, y, z$ uniform randomly and independently from $[q]^n$.*

Following is the Dictator Test for vertex pricing. Here, we use $\vec{1}$ to indicate the all "1" vector: $(1, 1, \ldots, 1) \in \mathbb{R}^n$.

**Definition 3.4.** *(Dictator Test $\mathcal{T}$) For $x', y', z'$ generated from $\mathcal{D}_1$ with $\epsilon$ set to be $1/q$ and an integer $k$ randomly chosen from $[\sqrt{q}]$, we generate a VERTEX-PRICING constraint among $f(x'), f(y'), f(z' \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1})$ with budget $\lfloor \sqrt{q}/k \rfloor$. We define*

$$\begin{aligned}\mathbf{profit}_{\mathcal{T}}(f) = \mathbf{E}\big[&\mathbf{revenue}((f(x'), f(y'),\\&f(z' \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1}), \lfloor \sqrt{q}/k \rfloor)\big].\end{aligned}$$

*and*

$$\mathbf{profit}^{+}_{\mathcal{T}}(f) = \mathbf{E}[\mathbf{revenue}^{+}((f(x'), f(y'), f(z' \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}), \lfloor\sqrt{q}/k\rfloor)].$$

For the purpose of analyzing $\mathcal{T}$, we also define the following Test $\mathcal{T}'$.

**Definition 3.5.** *(Test $\mathcal{T}'$) For $x, y, z$ generated from $\mathcal{D}_2$ and randomly choose $k \in [\sqrt{q}]$, we generate a* VERTEX-PRICING *constraint among $f(x), f(y), f(z)$ with budget $\lfloor\sqrt{q}/k\rfloor$).*

*We define*

$$\mathbf{profit}_{\mathcal{T}'}(f) = \mathbf{E}[\mathbf{revenue}\left((f(x), f(y), f(z), \lfloor\sqrt{q}/k\rfloor\right)].$$

*and*

$$\mathbf{profit}^{+}_{\mathcal{T}'}(f) = \mathbf{E}[\mathbf{revenue}^{+}\left((f(x), f(y), f(z), \lfloor\sqrt{q}/k\rfloor\right)].$$

We claim that for $\mathcal{T}'$, it has the following property:

**Proposition 3.6.** *For any function $f : [q]^n \to \mathbb{R}$, $\mathbf{profit}^{+}_{\mathcal{T}'}(f) \le 1$.*

*Proof.* Notice that for each triple $(x, y, z)$ , if there exists $k'$ such that $\sqrt{q}/(k'+1) < f(x) + f(y) + f(z) \le \lfloor\sqrt{q}/k'\rfloor$. Then the profit on edges associated with $(x, y, z)$ is nonzero only when $k$ is set to be $1, 2, ...k'$. Therefore, the profits on edges associated with $(x, y, z)$ are bounded by:

$$\frac{k'(f(x) + f(y) + f(z))}{\sqrt{q}} \le \frac{\lfloor\sqrt{q}/k'\rfloor k'}{\sqrt{q}} \le \frac{\sqrt{q}}{\sqrt{q}} = 1.$$

The last step here uses the fact that $\sqrt{q}$ is an integer.

If $f(x) + f(y) + f(z) \le 0$ or $f(x) + f(y) + f(z) \ge \sqrt{q}$, then the profit on edges associated with $x, y, z$ is below $0$.

Condition on every triple $(x', y', z')$, the expect profit associated with $f(x'), f(y'), f(z')$ is at most 1, therefore the overall profit is also at most 1. □

## 3.2 Analysis of the Dictator Test $\mathcal{T}$

We prove the completeness (Theorem 3.7) and soundness (Theorem 3.8) for $\mathcal{T}$ in this section.

**Theorem 3.7.** *(Completeness of $\mathcal{T}$) For function $f(x) = x_i - q/3$ for $x \in [q]^n$, $\mathbf{profit}_{\mathcal{T}}(f) \ge \Omega(\log q)$.*

*Proof.* Suppose $x', y', z' \sim \mathcal{D}_1$ is $1 - 1/q$ correlated of $x, y, z \sim \mathcal{D}_0$.

Since $x_i, y_i$ are randomly generated from $[q]$, we know that $\sqrt{q} \le x_i + y_i \le q$ with probability at least $1/3$. When this happens, $x_i + y_i + z_i = q$ and $z_i \le q - \sqrt{q}$. Also as each of the $x_i, y_i, z_i$ is reset to a random number with probability $\epsilon = 1/q$, we know that with probability $1/3 - 3/q$, $x'_i = x_i, y'_i = y_i, z'_i = z_i$ and we have that $x'_i + y'_i + z'_i = q$ and $z'_i \le q - \sqrt{q}$. We call these $(x', y', z')$ "good".

Then for "good" $(x', y', z')$, if we choose $f(t) = t_i - q/3$, we know that $f(x') + f(y') + f(z' \oplus_q \lfloor\sqrt{q}/k\rfloor) = x_i + y_i + z_i \oplus_q \lfloor\sqrt{q}/k\rfloor - q = \lfloor\sqrt{q}/k\rfloor$. Therefore,

$$\mathbf{revenue}\left((f(x'), f(y'), f(z' \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}), \lfloor\sqrt{q}/k\rfloor\right) = \lfloor\sqrt{q}/k\rfloor.$$

Therefore for "good" $(x', y', z')$, the associate profit is at least

$$\begin{aligned}&(1/3 - 3/q) \cdot \frac{\sum_{k=1}^{\sqrt{q}} \lfloor\sqrt{q}/k\rfloor}{\sqrt{q}} \\ &\ge (1/3 - 3/q) \cdot \frac{\sum_{k=1}^{\sqrt{q}} \left(\lfloor\sqrt{q}/k\rfloor - 1\right)}{\sqrt{q}} \\ &\ge (1/3 - 3/q) \cdot (\log\sqrt{q} - 1) \ge 1/8 \log q\end{aligned}$$

for large enough $q$.

Since the profit could be negative; we also need to show bound the profit loss on those "bad" $x', y', z'$ such that for some $k$

$$f(x') + f(y') + f(z' \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}) < 0.$$

This could happen for $x', y', z'$ generated from the following two cases:

1. At least one of the $x'_i, y'_i, z'_i$ is reset, this happens with probability at most $3/q$.
2. None of the $x'_i, y'_i, z'_i$ is reset. Since $x_i + y_i + z_i = q$ or $2q$, to make $f(x') + f(y') + f(z' \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}) < 0$, we know that we must have $x_i + y_i + z_i = q$ and $z_i > q - \lfloor\sqrt{q}/k\rfloor$. We must then have $x_i + y_i \le \lfloor\sqrt{q}/k\rfloor$ . We know that $\mathbf{Pr}(x_i + y_i \le \lfloor\sqrt{q}/k\rfloor \le \mathbf{Pr}(x_i, y_i \le \lfloor\sqrt{q}/k\rfloor) = \frac{1}{qk^2}$.

Therefore, we can have negative profit on $(x', y', z')$ occur with probability at most $4/q$. As we know that $f(x) = x_i - q/3 \geq -q/3$,therefore, $f(x') + f(y') + f(z' \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1}) \geq -q$, overall, we lose at most $4/q \cdot q = -4$ on those "bad" $(x', y', z')$.

Overall, for $f(x) = x_i - q/3$, we must have that $\mathbf{profit}_{\mathcal{T}}(f) \geq 1/8 \cdot \log q - 4 = \Omega(\log q)$ for sufficient large $q$. □

Now we state the soundness statement. The high-level idea is to show that for low influence function, the profit is about the same under either $\mathcal{T}$ or $\mathcal{T}'$ (of which the profit is bounded by 1). As $f : [q]^n \to R$ is not bounded, we define its influence on a transformation of $f$ as follows. We define $\tilde{f}$ be the integral part of $f$, being $\lfloor f \rfloor$. We then define $f' \in [q]$ and is uniquely determined by $f' = \tilde{f} \mod q$. By abuse of the notation, we also write $f' : [q]^n \to \{0,1\}^q$ with $f'^{(i)}$ being the indicator function $\mathbf{1}(\tilde{f} = i \mod q)$. The influence of $f'$ is defined with respect to its vector form.

**Theorem 3.8.** *(Soundness of $\mathcal{T}$) For $\tau^{C_0 1/q \log(p)} \leq 1/q^5$ and any function $f : [q]^n \to \mathbb{R}$ such that*

$$\max_i \mathrm{Inf}_i^{1-\epsilon} f' \leq \tau,$$

*we have that $\mathbf{profit}^{\pm}_{\mathcal{T}}(f) < 6 + O(1/q)$.*

*Proof.* Notice that the soundness statement is proved for the coupon model which automatically gives an upper bound for $\mathbf{profit}_{\mathcal{T}}(f)$.

First let us prove a stronger bound under the assumption that the output value of $f$ is an integer in $[q]$.

**Lemma 3.9.** *When $f \in [q]$, $\mathbf{profit}^{\pm}_{\mathcal{T}}(f) \leq 1 + O(1/q)$.*

*Proof.* We know that by definition $f'^{(i)} = \mathbf{1}(f = i)$. We also use $\mu_a$ to denote $\mathbf{E}_{x \in [q]^n}[f'^a(x)]$. We can arithmetize and bound the objective function $\mathbf{profit}^{\pm}_{\mathcal{T}}(f)$ in terms of $f'^{(i)}$ as follows:

$$\mathbf{profit}^{\pm}_{\mathcal{T}}(f) = \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mathop{\mathbf{E}}_{x',y',z' \sim \mathcal{D}_{1,k}}$$
$$\left[f'^a(x') f'^b(y') f'^c(z' \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1})(a+b+c)\right]$$
$$= \mathop{\mathbf{E}}_{x,y,z \sim \mathcal{D}_{0,k}} \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} T_{1-\epsilon} f'^a(x)$$
$$\cdot T_{1-\epsilon} f'^b(y) \cdot T_{1-\epsilon} f'^c(z \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1})(a+b+c)\Big]$$
$$= \mathop{\mathbf{E}}_k \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mathop{\mathbf{E}}_{x,y,z \sim \mathcal{D}_0} T_{1-\epsilon} f'^a(x) T_{1-\epsilon}$$
$$\cdot f'^b(y) T_{1-\epsilon} f'^c(z \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1})(a+b+c)\Big].$$

Notice that $\mathrm{Inf}_i^{1-\epsilon} f'^a \leq \mathrm{Inf}_i^{1-\epsilon} f' \leq \tau$ for $i \in [n], a \in [q]$. Also $x, y, z \sim \mathcal{D}_0$ are pairwise independent, by Theorem 2.3 (with minimum probability $\alpha = 1/q$), we can plug in independent $x, y, z \sim \mathcal{D}_2$ with additive error bounded by $\tau^{C_0 1/(q \log q)} \leq 1/q^5$. That is

$$\mathbf{profit}^{\pm}_{\mathcal{T}}(f) < \mathop{\mathbf{E}}_k \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mathop{\mathbf{E}}_{x,y,z \sim \mathcal{D}_2} (f'^a(x)$$
$$\cdot f'^b(y) \cdot f'^c(z \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1}) + 1/q^5)(a+b+c)\Big]$$
$$\leq \mathbf{E}_k \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mu_a \mu_b f'^c(z \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1})$$
$$\cdot (a+b+c)\Big] + O(1/q).$$

The last inequality uses the fact that $a+b+c \leq \sqrt{q}$ and there are at most $q^3$ terms in the summation.

A important observation is that for any fixed $k$, the random vector variable $z \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1}$ always follows the same uniform distribution over $[q]^n$ and therefore is independent of the choice of $k$. Therefore, we can further bound $\mathbf{profit}^{\pm}_{\mathcal{T}}(f)$ by

$$\mathbf{E}_k \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mu_a \mu_b \mu_c (a+b+c)\Big] + O(1/q).$$

As for the term

$$\mathbf{E}_k \Big[ \sum_{0 \leq a+b+c \leq \lfloor \sqrt{q}/k \rfloor, a,b,c \in [q]} \mu_a \mu_b \mu_c (a+b+c)\Big].$$

It is equal to $\mathbf{profit}^{\pm}_{\mathcal{T}'}(f)$ and by proposition 3.6, we know that $\mathbf{profit}^{\pm}_{\mathcal{T}'}(f) \leq 1$. Overall, we bound $\mathbf{profit}^{\pm}_{\mathcal{T}}(f)$ by $1 + O(1/q)$ when $f \in [q]$.

Following two observation are useful in our analysis of the more general case of $f$.

**Observation 3.10.** *Above proof also works even for random function $f(x) \in [q]$ specified by $f^i$ in the following way: for each $x$, with probability $f^i(x)$, $f$ outputs $i$. Here $\sum f^i(x) = 1$ for any $x \in [q]^n$.*

**Observation 3.11.** *For any $\theta \in \mathbb{R}^+, f \in [q]$, we can also obtain the same bound on the profit of function $f-\theta$; i.e., $\mathbf{profit}^+_{\mathcal{T}}(f-\theta) \le 1+O(1/q)$.*

To see this, simply notice that

$$\begin{aligned}&\mathbf{profit}^+_{\mathcal{T}}(f-\theta)\\ =&\sum_{3\theta\le a+b+c\le 3\theta+\lfloor\sqrt{q}/k\rfloor, a,b,c\in[q]} \mathop{\mathbf{E}}_{x',y',z'\sim\mathcal{D}_{1,k}}[f^a(x')\\ &\cdot f^b(y')f^c(z'\oplus_q\lfloor\sqrt{q}/k\rfloor\cdot\vec{1})(a+b+c-3\theta)]\end{aligned}$$

and then we use the same proof and show that $\mathbf{profit}^+_{\mathcal{T}}(f-\theta) \le \mathbf{profit}^+_{\mathcal{T}'}(f-\theta)+O(1/q) \le 1+O(1/q)$. □

Now we are ready to analyze real-valued function $f$. Recall that $\tilde{f} = \lfloor f \rfloor$. First use the fact that $f \le \tilde{f}+1$, we have

$$\begin{aligned}&\mathbf{revenue}^+((f(x'), f(y'), f(z'\oplus_q\lfloor\sqrt{q}/k\rfloor\cdot\vec{1}),\\ &\lfloor\sqrt{q}/k\rfloor) \le \mathbf{revenue}^+((\tilde{f}(x'),\tilde{f}(y'),\\ &\tilde{f}(z'\oplus_q\lfloor\sqrt{q}/k\rfloor\cdot\vec{1}),\lfloor\sqrt{q}/k\rfloor)+3. \quad (1)\end{aligned}$$

Therefore,

$$\begin{aligned}\mathbf{profit}^+_{\mathcal{T}}(f) &= \mathbf{E}_{x',y',z',k}[\mathbf{revenue}^+((f(x'),f(y'),\\ &\quad f(z'\oplus_q\lfloor\sqrt{q}/k\rfloor\cdot\vec{1}),\lfloor\sqrt{q}/k\rfloor)]\\ &< \mathbf{E}_{x',y',z',k}[\mathbf{revenue}^+((\tilde{f}(x'),\tilde{f}(y'),\\ &\quad \tilde{f}(z'\oplus_q\lfloor\sqrt{q}/k\rfloor\cdot\vec{1}),\lfloor\sqrt{q}/k\rfloor)+3]\\ &\le \mathbf{profit}^+_{\mathcal{T}}(\tilde{f})+3.\end{aligned}$$

Recall that $f' = \tilde{f} \mod q$. In the next step we show that

$$\begin{aligned}\mathbf{profit}^+_{\mathcal{T}}(\tilde{f}) \le \mathbf{profit}^+_{\mathcal{T}}(f')&+\mathbf{profit}^+_{\mathcal{T}}(f'-q/3)\\ &+\mathbf{profit}^+_{\mathcal{T}}(f'-2q/3) \quad (2)\end{aligned}$$

By definition of $f'$, we know that

$$\tilde{f}(x)+\tilde{f}(y)+\tilde{f}(z) = f'(x)+f'(y)+f'(z) \mod q.$$

Therefore, if $\tilde{f}(x)+\tilde{f}(y)+\tilde{f}(z) \le \lfloor\sqrt{q}/k\rfloor$ for some $k$, it must be the case that

$$\begin{aligned}f'(x)+f'(y)+f'(z) \in [0,\lfloor\sqrt{q}/k\rfloor], [q, q+\lfloor\sqrt{q}/k\rfloor]\\ \text{or } [2q, 2q+\lfloor\sqrt{q}/k\rfloor].\end{aligned}$$

Then we have that,

$$\begin{aligned}&\mathbf{revenue}^+(\tilde{f}(x),\tilde{f}(y),\tilde{f}(z),\lfloor\sqrt{q}/k\rfloor)\\ \le&\ \mathbf{revenue}^+(f'(x),f'(y),f'(z),\lfloor\sqrt{q}/k\rfloor))\\ &+\mathbf{revenue}^+(f'(x)-q/3, f'(y)-q/3,\\ &f'(z)-q/3,\lfloor\sqrt{q}/k\rfloor))+\mathbf{revenue}^+(f'(x)-2q/3,\\ &f'(y)-2q/3, f'(z)-2q/3,\lfloor\sqrt{q}/k\rfloor)).\end{aligned}$$

This proves (2). And by Observation 3.11, we have that

$$\mathbf{profit}^+_{\mathcal{T}}(f) < \mathbf{profit}^+_{\mathcal{T}}(\tilde{f})+3 \le 3\cdot 1+3 \le 6+O(1/q).$$

□

## 4 The reduction from the UNIQUE-GAMES

In this section we show how to use our Dictator Test $\mathcal{T}$ to obtain our main result, Theorem 1.7. First let us recall the definition of the UNIQUE-GAMES.

**Definition 4.1.** *For $L \in \mathbb{N}$, a* UNIQUE-GAMES$_L(U,V,E,(\pi^{u,v})_{(u,v)\in E})$ *instance consists of a bipartite graph having vertex sets $U$, $V$ and edge set $E$, together with a* bijective constraint $\pi^{v,u} : [L] \to [L]$ *for each $(u,v) \in E$. .*

The following equivalent version of the UGC is due to Khot and Regev [8, Lemma 3.6]:

**Theorem 4.2.** *Assume the UGC. For all small $\zeta,\gamma > 0$, there exists $L \in \mathbb{N}$ such given an* UNIQUE-GAMES$_L$ *instance $\mathcal{G} = (U,V,E,(\pi^{u,v})_{(u,v)\in E})$ which is $U$-regular, it is* NP*-hard to distinguish the following two cases:*

1. *There is an assignment $A : (U \cup V) \to [L]$ and a subset $U' \subseteq U$ with $|U'|/|U| \ge 1-\zeta$ such that $A$ satisfies all constraints incident on $U'$.*
2. *There is no assignment $A$ that satisfies more than $\gamma$ fraction of the constraints.*

1. Choose $u \in U$ randomly.
2. Choose 3 of $u$'s neighbor $v_1, v_2, v_3$ randomly (with replacement).
3. Generate $(x, y, z) \sim \mathcal{D}_2$ and $k$ randomly from $[\sqrt{q}]$.
4. Add a constraint among $f_{v_1}(\pi^{v_1,u}(x)), f_{v_2}(\pi^{v_2,u}(y))$ and $f_{v_3}(\pi^{v_3,u}(z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1})$ with budget $\lfloor\sqrt{q}/k\rfloor$.

Figure 1: Reduction to Unique Games

We make the reduction from a UNIQUE-GAMES instance $\mathcal{G}$ to a VERTEX-PRICING$_3$ instance $\mathcal{I}$. Given the UNIQUE-GAMES$_L$ instance $\mathcal{G} = (U, V, E, \{\pi^{uv}\})$, the reduction produces a weighted VERTEX-PRICING instance $\mathcal{I}$ with variable set $V \times [q]^L$. We think of a price assignment $F$ to these variables as a collection of functions $F = \{f_v : [q]^L \to \mathbb{R}\}$, one for each $v \in V$.

For $x \in [q]^L$ and mapping $\pi : [L] \to [L]$, we also denote $\pi(x) \in [q]^L$ as the permutation of $x$'s coordinate according to $i$; i.e., $\pi(x)_i = x_{\pi(i)}$. The specific steps of the reduction is described in Figure 1.

We claim that such a reduction have the following properties.

**Theorem 4.3.** *For $\zeta = 1/q$, $\tau$ satisfies that $\tau^{C_0 q \log q} \leq 1/q^5$ and $\gamma = \tau^2/q^5$, above reduction satisfies that*

- *(Completeness.) If statement 1 in Theorem 4.2 holds for $\mathcal{G}$, then there is a price assignment $F$ such that $\mathbf{profit}_{\mathcal{I}}(F) = \Omega(\log p)$. In addition, the price assigned on each variable is $q$-bounded, i.e., with value $\geq -q$.*
- *(Soundness.) If there is non assignment for $\mathcal{G}$ that satisfies more than $\gamma$ fraction of the edges, then for every price assignment $F$ such that $\mathbf{profit}^+_{\mathcal{I}}(F) \leq 6 + 1/\sqrt{q}$*
- $1 \leq \mathrm{Opt}_{pos} \leq 3$.

By combining Theorem 4.3 with Theorem 4.2, and setting $q = B$, we immediately prove Theorem 1.7.

We prove the completeness and soundness properties in Section 4.1 and bound the value of $\mathrm{Opt}_{pos}$ in Section 4.2.

### 4.1 Completeness and soundness proof

*Proof.* (Completeness) To prove the completeness part of Theorem 4.3, suppose that assignment $A : V \to [L]$ and subset $U' \subseteq U$ are as in statement 1 of Theorem 4.2. Define an price assignment $F$ for $\mathcal{I}$ by taking $f_v(x) = x_{A(v)} - q/3$. Then by definition and the property of $A$, for $u' \in U'$, $f_{v_i}(\pi^{v_i,u'}(x)) = x_{A(u')} - q/3$ for $i = 1, 2, 3$. Thus by the completeness of the Dictator Test (Theorem 3.7), assignment $F$ will have profit at least $\Omega(\log p)$ conditioned on $u' \in U$ is picked. As for the case that $u \notin U'$ is picked, we lose a negative profit bounded by $-q$ . Overall, we have that $\mathbf{profit}_{\mathcal{I}}(F) \geq (1-\zeta)\Omega(\log q) + \zeta q$. Notice that we choose $\zeta = 1/q$, therefore, $\mathbf{profit}_{\mathcal{I}}(F) \geq \Omega(\log p)$. In addition, we know that the assignment on each $f_v$ is above $-q/3$.

(Soundness) We prove the soundness statement by contradiction. Suppose that some assignment $F$ have $\mathbf{profit}^+_{\mathcal{I}}(F) \geq 6 + 1/\sqrt{q}$, we will exhibit a assignment to the Unique Games instance $\mathcal{G}$ that satisfies at least a $\gamma$ fraction of the edges. Notice that the maximum profit on each constraint is at most $\sqrt{q}$, then by an average argument, we must have for at least $1/q$ fraction of the vertices $u \in U$ picked in the first step, such that the expected profit on these $u$ is above $6 + \frac{1}{2\sqrt{q}}$.

Let us call these $u$ "good". Write $N(u)$ as the neighbor of $u$. By definition, for a fixed "good" $u$, let us denote the expected cost given $u$ is picked as $\mathbf{profit}^+_{\mathcal{I},u}(F) =$

$$\mathbf{E}_{v_1,v_2,v_3 \in N(u),x,y,z,k}[\ \mathbf{revenue}^+(f_{v_1}(\pi^{u,v_1}x), f_{v_2}(\pi^{u,v_2}(y)), f_{v_3}(\pi^{u,v_3}(z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}), \lfloor\sqrt{q}/k\rfloor)]. \quad (3)$$

Then we know that for good $u$,

$$\mathbf{profit}^+_{\mathcal{I},u}(F) \geq 6 + \frac{1}{2\sqrt{q}}.$$

Similar to the analysis of Theorem 3.8 , we define $\tilde{f}_v = \lfloor f_v \rfloor$ and introduce $f'_v \in [q]$ such that $f'_v = \tilde{f}_v \mod q$, although we also write $f'_v$ as $[q]^n \to \{0,1\}^q$ with its $i$-th coordinate indicate whether $f'_v$ is $i$. We call the assignment corresponding to $\{\tilde{f}_v\}_{v \in V}$ as $\tilde{F}$ and the assignment corresponding to $\{f'_v\}_{v \in V}$ as $F'$.

By the proof of (1), we know that

$$\mathbf{profit}^+_{\mathcal{I},u}(\tilde{F}) \geq \mathbf{profit}^+_{\mathcal{I},u}(F) - 3 \geq 3 + \frac{1}{2\sqrt{q}}$$

and by the proof of (2), we have that

$$\mathbf{profit}^+_{\mathcal{I},u}(F') + \mathbf{profit}^+_{\mathcal{I},u}(F' - q/3) + \mathbf{profit}^+_{\mathcal{I},u}(F' - 2q/3) \geq \mathbf{profit}^+_{\mathcal{I}}(\tilde{F}).$$

Therefore, one of $\mathbf{profit}^+_{\mathcal{I},u}(F'), \mathbf{profit}^+_{\mathcal{I},u}(F' - q/3), \mathbf{profit}^+_{\mathcal{I},u}(F' - 2q/3)$ should be above $1 + \frac{1}{6\sqrt{q}}$.

Let us assume that $\mathbf{profit}^+_{\mathcal{I},u}(F' - q/3) \geq 1 + \frac{1}{6\sqrt{q}}$ and we will show that we can decode $f_u$ into a few influential coordinates. (The other 2 cases are very similar).

We know then

$$\mathbf{profit}^+_{\mathcal{I},u}(F' - q/3) = \mathbf{E}_{x,y,z,k,v_1,v_2,v_3}[\sum_{q<a+b+c\leq q+\lfloor\sqrt{q}/k\rfloor} [f^a_{v_1}(\pi^{v_1,u}(x)) \cdot f^b_{v_2}(\pi^{v_2,u}(y)) f^c_{v_3}(\pi^{v_3,u}(z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}) \cdot (a+b+c-q)]$$

$$= \mathbf{E}_{x,y,z,k}[\sum_{q<a+b+c\leq q+\lfloor\sqrt{q}/k\rfloor} \mathbf{E}_{v_1\in N(u)}[f^a_{v_1}(\pi^{v_1,u}(x))] \cdot \mathbf{E}_{v_2\in N(u)}[f^b_{v_2}(\pi^{v_2,u}(y))] \cdot \mathbf{E}_{v_3\in N(u)}[f^c_{v_3}(\pi^{v_3,u}(z\oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1}))](a+b+c-q)]$$

If we define $f^i_u = \mathbf{E}_{v\in N(u)}[f^i_v(\pi^{v,u}(x))]$ for $i \in [q]$, then we have that

$$\mathbf{profit}^+_{\mathcal{I}}(F' - q/3) = \mathbf{E}_{x,y,z,k}[\sum_{q<a+b+c\leq q\oplus_q\lfloor\sqrt{q}/k\rfloor} f^i_u(x) f^i_u(y) f^i_u((z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1})(a+b+c-q)] \geq 1 + \frac{1}{6\sqrt{q}}. \quad (4)$$

Denote $f_u(x) = (f^1_u(x), f^2_u(x), \ldots, f^q_u(x))$. It is easy to check that $\sum f^i_u(x) = 1$ for every $x$. Then $f_u$ can be viewed as a random function that on a particular $x$ such that it outputs $i$ with probability $f^i_u(x)$. Then (4) is equal to the profit of the Dictator Test $\mathcal{T}$ on $f_u - q/3$ and we have that $\mathbf{profit}^+_{\mathcal{T}}(f_u - q/3) \geq 1 + \frac{1}{6\sqrt{q}}$.

We know then by a contrapositive statement of Lemma 3.9 along with Observation 3.11 and Observation 3.10, there must be some $i$ such that $\mathrm{Inf}^{1-\epsilon}_i f_u \geq \tau$.

Then by Fact 2.2, we know that

$$\tau \leq \mathrm{Inf}^{1-\epsilon}_i f_u \leq \mathbf{E}_{v\in N(u)}[\mathrm{Inf}^{1-\epsilon}_i f_v(\pi^{v,u}(x))]$$

By an averaging argument, since $\mathrm{Inf}_i f_v(\pi^{v,u}(x)) = \sum_{j\in[q]} \inf_i f^j_v \leq q$, for $\frac{\tau}{2q}$ fraction of the $v \in N(u)$, we have that $\mathrm{Inf}^{1-\epsilon}_i f_v(\pi^{v,u}(x)) = \mathrm{Inf}^{1-\epsilon}_{j=(\pi^{v,u})^{-1}(i)} f_v \geq \tau/2$.

Now consider choosing the following randomized assignment to the Unique Games instance. Let $S_u$ be $\{i \mid \mathrm{Inf}^{1-\epsilon}_i f_u \geq \tau\}$ and $S_v$ be $\{i \mid \mathrm{Inf}^{1-\epsilon}_i f_u \geq \tau/2\}$. By fact 2.1, we know that $|S_v| \leq q^2/\tau$.

The assignment would be randomly set a label in $S_u$ for $u$ and a label in $S_v$ for $v$. Then it is easy to see for good vertex $u$ and any of its coordinate $i \in S_u$, $\tau/2q$ fraction of its neighbor will have a matching coordinate $j = (\pi^{v,u})^{-1}(i)$ in $S_v$. Therefore above assignment satisfies at least $1/|S_v| \cdot \tau/2q$ fraction of the edges for "good" $u$. We know that there is at least a $1/\sqrt{q}$ fraction of the $u$ is good. By the regularity of the graph at the $U$ side, we know that such a labeling strategy satisfies at least $1/\sqrt{q} \cdot (\tau/q^2)\tau/2q \geq \tau^2/q^5 = \gamma$ fraction of the edges. □

## 4.2 Bounding $\mathrm{Opt}(\mathcal{I})$

*Proof.* It is easy to verify that $opt(\mathcal{I}) \geq 1$ by assigning a constant function $f_v = 1/3$ for every $v \in V$.

Notice that the reduction add an edge among $f_{v_1}(\pi^{v_1,u}(x)), f_{v_2}(\pi^{v_2,u}(y))$ and $f_{v_3}(\pi^{v_3,u}(z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1})$ with budget $\sqrt{q}/k$.

To upper bound $\mathrm{Opt}_{pos}(\mathcal{I})$, for the purpose of analysis,we can imagine that along with the generation process of $\mathcal{I}$, we also construct the following three instance $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$. The instance $\mathcal{I}_k$ for $k = 1, 2$ is constructed according to Figure 1 except in the last step we add an edge only on $f_{v_k}(\pi^{v_1,u}(x))$ with budget $\lfloor\sqrt{q}/k\rfloor$. And for $\mathcal{I}_3$ is constructed that in the last step a constraint is added to $f_{v_3}(\pi^{v_3,u}(z) \oplus_q \lfloor\sqrt{q}/k\rfloor \cdot \vec{1})$ with budget $\lfloor\sqrt{q}/k\rfloor$ . First we claim that

$$\mathrm{Opt}(\mathcal{I}_k) = 1 \text{ for } k = 1, 2, 3.$$

Let us just first look at the most complicated case when $k = 3$. Notice that for any $u$ and fixed $k$, $\pi^{v_3,u}(z) \oplus_q \lfloor \sqrt{q}/k \rfloor \cdot \vec{1}$ follows the uniform distribution over $[q]^L$. Therefore, we can view the constraint as added on $f_{v_3}(x)$ with budget $\lfloor \sqrt{q}/k \rfloor$ for $x$ randomly chosen from $[q]^n$. Similar analysis can be made to $\mathcal{I}_1, \mathcal{I}_2$ and essentially $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$ can all be viewed as generated by the following process:

1. pick a random vertex $u$ in $U$.
2. pick a random neighbour $v$ in $N(u)$
3. generate $x$ randomly from $[q]^L$.
4. pick $k$ from $[\sqrt{q}]$
5. add a budget over $\lfloor \sqrt{q}/k \rfloor$ over $f_v(x)$.

By the analysis of Proposition 3.6, we know that $\mathrm{Opt}_{pos}$ is 1 for above instance.

For the remaining proof, we show that $\mathrm{Opt}(\mathcal{I}) \le \mathrm{Opt}(\mathcal{I}_1) + \mathrm{Opt}(\mathcal{I}_2) + \mathrm{Opt}(\mathcal{I}_3) = 3$.

This can be obtained from the observation that for any constraint on variables $(p_1, p_2, p_3)$ with budget $b$, it is easy to verify that

$$\mathbf{revenue}(p_1, p_2, p_3, b) \le \sum_{i=1}^{3} p_i \cdot (p_i < b) \quad (5)$$

when $p_1, p_2, p_3 > 0$.

Therefore, for any *positive* pricing function $f$ for $\mathcal{I}$, if use the same price function on $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$, the sum of the profit of $f$ over these three instance is more than the profit of $f$ over $\mathcal{I}$. □

### 4.3 The hardness of unweighted vertex pricing problem

Since the unweighted vertex pricing problem is also defined over graph with parallel edges (as many customers may be interested in the same set of items). These parallel edges can be viewed as assigning an integer weight to each edge in the graph. Therefore, the only difference between weighted and unweighted vertex pricing instance is that weighted instance may take non-integer weights. However, if we look at the reduction, all the weights assigned to each edge is a constant that only depends on $L$, which is the label size of the UNIQUE-GAMES; therefore, we can properly scale them up to get an instance with integer weights. [3]

[3] In fact, it is possible to obtain the same hardness results even assuming that the graph is unweighted and without parallel edges by following by following essentially the same reduction as in [6], Lemma 2.2.

## 5 Conclusion and open problems

We believe the Dictator Test as well as the techniques developed in this paper is also useful in proving hardness results for the VERTEX-PRICING problem under other settings. For example, for the problem of VERTEX-PRICING$_k$ with positive price, one should be able to obtain an improved hardness results by designing a proper Dictator Test.

One of the obvious open problem is to prove the same hardness result under the assumption that $\mathrm{P} \neq \mathrm{NP}$. This is feasible as our Dictator test is similar to Hastad's 3Lin Test which is used to prove the NP-hardness results of MAX 3LIN$_q$.

Technically, it would be challenging to generalize the above hardness results to VERTEX-PRICING$_2$ (i.e., the graph vertex pricing problem), although we feel that VERTEX-PRICING$_3$ is already simple enough to demonstrate the inherent intractability of the problem of pricing loss leaders.

## 6 Acknowledgement

The author would like to thank Nina Balcan for suggesting me to work this problem and anonymous reviewer for suggesting me to understand the approximability of the profitable gap.

## References

[1] Nina Balcan and Avrim Blum. Approximation algorithms and online mechanisms for item pricing. In *Proc. 7th ACM Conference on Electronic commerce*, pages 29–35, New York, NY, USA, 2006. ACM.

[2] Nina Balcan, Avrim Blum, Hubert Chan, and Mohammadtaghi Hajiaghayi. Hajiaghayi: A theory of loss-leaders: Making money by pricing below cost. In *Proc. 3rd Intern. Workshop on Internet and Network Economics, Lecture Notes in Computer Science*. Springer, 2007.

[3] Patrick Briest and Piotr Krysta. Single-minded unlimited supply pricing on sparse instances. In *Proc. 17th ACM-SIAM Symp. on Discrete Algorithms*, pages 1093–1102, New York, NY, USA, 2006. ACM.

[4] Erik D. Demaine, Uriel Feige, Mohammadtaghi Hajiaghayi, and Mohammad R. Salavatipour. Combination can be hard: Approximability of the

unique coverage problem. In *Proc. 17th ACM-SIAM Symp. on Discrete Algorithms*, pages 162–171, 2006.

[5] Venkatesan Guruswami, Jason D. Hartline, Anna R. Karlin, David Kempe, Claire Kenyon, and Frank McSherry. On profit-maximizing envy-free pricing. In *Proc. 15th ACM-SIAM Symp. on Discrete Algorithms*, pages 1164–1173, Philadelphia, PA, USA, 2005. Society for Industrial and Applied Mathematics.

[6] Rohit Khandekar, Tracy Kimbrel, Konstantin Makarychev, and Maxim Sviridenko. On hardness of pricing items for single-minded bidders. In *APPROX-RANDOM*, pages 202–216, 2009.

[7] Subhash Khot, Guy Kindler, Elchanan Mossel, and Ryan O'Donnell. Optimal inapproximability results for MAX-CUT and other 2-variable CSPs? *SIAM Journal on Computing*, 37(1):319–357, 2007.

[8] Subhash Khot and Oded Regev. Vertex Cover might be hard to approximate to within $2 - \epsilon$. In *Proc. 18th IEEE Conference on Computational Complexity*, pages 379–386, 2003.

[9] Elchanan Mossel. Gaussian bounds for noise correlation of functions and tight analysis of long codes. In *Proc. 49th IEEE Symp. on Foundations of Comp. Sci.*, pages 156–165, 2008.

[10] Elchanan Mossel, Ryan O'Donnell, and Krzysztof Oleszkiewicz. Noise stability of functions with low influences: invariance and optimality. In *Proc. 46th IEEE Symp. on Foundations of Comp. Sci.*, pages 21–30, 2005. To appear, *Annals of Mathematics*.

[11] Ryan O'Donnell, Yi Wu, and Yuan Zhou. Unique Games over Integers. Manuscript, 2010.

[12] Prasad Raghavendra. *Approximating NP-hard problems: efficient algorithms and their limits*. PhD thesis, University of Washington, 2009.