Roadmap to the appendix

Appendix A includes omitted material from the Model section from the main text. Appendix B includes the proofs omitted from Section 3 (Offline Setting). Appendix C includes the proofs omitted from Section 4 (Online Setting). Appendix D includes the equilibrium analysis. A few results from prior work that we invoke are in Appendix E.

A Appendix: Examples for the Model

Example 1. *Let $K = 3$ and $n = 2$. Suppose the valuations are $\mathbf{v}_1 = (5, 2)$ and $\mathbf{v}_2 = (4, 1)$. If the players submit bids $\mathbf{b}_1 = (2, 1)$ and $\mathbf{b}_2 = (3, 2)$, the bid are sorted in the order $(b_{2,1}, b_{1,1}, b_{2,2}, b_{1,2})$. Then the allocation is $x_1 = 1$ and $x_2 = 2$.*

Under the K-th price auction, the price is set to $p = b_{2,2} = 2$. Then the utilities of the players are $u_1(\mathbf{b}) = V_1(1) - p = 5 - 2 = 3$ and $u_2(\mathbf{b}) = V_2(2) - 2 \cdot p = (4 + 1) - 2 \cdot 2 = 1$.

Under the $(K + 1)$-st price auction, the price is set to $p = b_{1,2} = 1$. Then the utilities of the players are $u_1(\mathbf{b}) = V_1(1) - p = 5 - 1 = 4$ and $u_2(\mathbf{b}) = V_2(2) - 2 \cdot p = (4 + 1) - 2 \cdot 1 = 3$.

B Appendix: Offline Setting

In this section we include the proofs omitted from the main text for the offline setting.

Observation 1 (restated). *Player i has an optimum bid vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathbb{D}^K$ with $\beta_j \in \mathcal{S}_i$ for all $j \in [K]$.*

Proof. Let $\mathbf{c} = (c_1, \ldots, c_K)$ be an arbitrary optimum strategy for player i. Suppose $\mathbf{c} \notin \mathcal{S}_i^K$.

We use $\mathbf{c}$ to construct an optimum bid vector $\boldsymbol{\beta} \in \mathcal{S}_i^K$ as follows. For each $j \in [K]$:

- If $c_j \in \mathcal{S}_i$, then set $\beta_j = c_j$.
- Else, set $\beta_j = \max\{y \in \mathcal{S}_i \mid y \leq c_j\}$.

Then in each round t, player i gets the same allocation when playing $\boldsymbol{\beta}$ as it does when playing $\mathbf{c}$ and the others play $\mathbf{b}_{-i}^t$ since the ordering of the owners of the bids is the same under $(\mathbf{c}, \mathbf{b}_{-i}^t)$ as it is under $(\boldsymbol{\beta}, \mathbf{b}_{-i}^t)$; moreover, the price weakly decreases in each round t. Thus player i's utility weakly improves. Since $\mathbf{c}$ was an optimal strategy for player i, it follows that $\boldsymbol{\beta}$ is also an optimal strategy for player i and, moreover, $\boldsymbol{\beta} \in \mathcal{S}_i^K$ as required. □

Next we show how to find an optimal bid vector in polynomial time. The proof uses several lemmas, which are proved after the theorem.

Theorem 1 (restated, formal). *Suppose we are given a number n of players, number K of units, valuation $\mathbf{v}_i$ of player i, discretization level $\varepsilon > 0$, and bid history $H_{-i} = (\mathbf{b}_{-i}^1, \ldots, \mathbf{b}_{-i}^T)$ by players other than i. Then an optimum bid vector for player i can be computed in polynomial time in the input parameters.*

Proof of Theorem 1. Compute the set $\mathcal{S}_i$ given by equation (1) and the graph G_i from Definition 1. The proof has several steps as follows.

G_i is a DAG. All the edges in G_i flow from the source z_- to the nodes from layer 1, then from the nodes in layer j to those in layer $j+1$ (i.e. of the form $(z_{r,j}, z_{s,j+1})$ for all $j \in [K-1]$ and $r, s \in \mathcal{S}_i$ with $r \geq s$), and finally from all the nodes in layer K to the sink z_+. A cycle would require at least one back edge, but such edges do not exist.

Bijective map between bid vectors and paths from source to sink in G_i. To each bid vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathcal{S}_i^K$, associate the following path in G_i: $P(\boldsymbol{\beta}) = \left(z_-, z_{\beta_1,1}, \ldots, z_{\beta_K,K}, z_+\right)$.

We show next the map P is a bijection from the set $\mathcal{S}_i$ of candidate bid vectors for player i to the set of paths from source to sink in G_i. Consider arbitrary bid vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathcal{S}_i^K$. By definition of a bid vector, we have $\beta_1 \geq \ldots \geq \beta_K$. Then $P(\boldsymbol{\beta}) = (z_-, z_{\beta_1,1}, \ldots, z_{\beta_K,K}, z_+)$ is a valid path in G_i.

Conversely, since G_i has an edge $(z_{r,j}, z_{s,j+1})$ if and only if $r \geq s$, each path from source to sink in G_i has the form $Q = (z_-, z_{\beta_1,1}, \ldots, z_{\beta_K,K}, z_+)$ for some numbers $\beta_1, \ldots, \beta_K \in \mathcal{S}_i$, and so it can be mapped to bid vector $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$. Thus $Q = P(\boldsymbol{\beta})$, and so P is a bijective map.

Utility of player i and weight of a path of length K in G_i. Let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathcal{S}_i^K$. Additionally, let $\beta_{K+1} = 0$. The total utility of player i when bidding $\boldsymbol{\beta}$ in each round t while the others bid $\mathbf{b}_{-i}^t$ is $U_i(\boldsymbol{\beta}) = \sum_{t=1}^T u_i(\mathbf{h}^t)$, where $\mathbf{h}^t = (\boldsymbol{\beta}, \mathbf{b}_{-i}^t)\ \forall t \in [T]$.

Let $P(\boldsymbol{\beta}) = (z_-, z_{\beta_1,1}, \ldots, z_{\beta_K,K}, z_+)$ be the path in G_i corresponding to $\boldsymbol{\beta}$. The edges of the path $P(\boldsymbol{\beta})$ are $(z_-, z_{\beta_1,1})$, $(z_{\beta_1,1}, z_{\beta_2,2}), \ldots, (z_{\beta_{K-1},K-1}, z_{\beta_K,K})$, $(z_{\beta_K,K}, z_+)$, while the weight of the path is equal to the sum of its edges. Summing equation (2), which gives the weight of an edge, across all edges of $P(\boldsymbol{\beta})$ implies that the weight of path $P(\boldsymbol{\beta})$ is equal to

$$w(P(\boldsymbol{\beta})) = \sum_{j=1}^{K}\sum_{t=1}^{T}\Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j}-\beta_j) + j\left(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j-\beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t))\right)\Big]\,. \tag{7}$$

We claim that $U_i(\boldsymbol{\beta}) = w(P(\boldsymbol{\beta}))$. The high level idea is to rewrite the utility to "spread it" across the edges of the path corresponding to bid profile $\boldsymbol{\beta}$. Towards this end, recall the utility of player i at strategy profile $\mathbf{h}^t = (\boldsymbol{\beta}, \mathbf{b}_{-i}^t)$ is

$$u_i(\mathbf{h}^t) = \sum_{j=1}^{K}\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - p(\mathbf{h}^t))\,. \tag{8}$$

By Lemma 1, equation (8) is equivalent to

$$u_i(\mathbf{h}^t) = \sum_{j=1}^{K}\Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j}-\beta_j) + j\left(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j-\beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t))\right)\Big]\,. \tag{9}$$

Summing $u_i(\mathbf{h}^t)$ over all rounds t gives

$$\begin{aligned}
U_i(\boldsymbol{\beta}) &= \sum_{t=1}^{T} u_i(\mathbf{h}^t)\\
&= \sum_{t=1}^{T}\sum_{j=1}^{K}\Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j}-\beta_j) + j\left(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j-\beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t))\right)\Big] && \text{(By equation (9))}\\
&= \sum_{j=1}^{K}\sum_{t=1}^{T}\Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j}-\beta_j) + j\left(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j-\beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t))\right)\Big] && \text{(Swapping the order of summation)}\\
&= w(P(\boldsymbol{\beta}))\,. && \text{(By equation (7))}
\end{aligned}$$

Thus the weight of the path $P(\boldsymbol{\beta})$ is equal to the utility of player i from bidding $\boldsymbol{\beta}$. The implication is that to find an optimum bid vector, it suffices to compute a maximum weight path in G_i.

Computing a maximum weight path in G_i. The graph G_i is a DAG with a number of vertices of G_i that is polynomial in the input parameters. By Lemma 2, the edge weights of G_i can be computed in polynomial time.

A maximum weight path in a DAG can be computed in polynomial time by changing every weight to its negation to obtain a graph $-G$. Since G has no cycles, the graph $-G$ has no negative cycles. Thus running a shortest path algorithm on $-G$ will yield a longest (i.e. maximum weight) path on G in polynomial time, as required.

The unique bid vector corresponding to the maximum weight path found can then be recovered in time $O(K)$ and it represents an optimum bid vector for player i. $\square$

Lemma 1. *In the setting of Theorem 1, for each bid profile $\boldsymbol{\beta} \in \mathcal{S}_i^K$ and round $t \in [T]$, define $\mathbf{h}^t = (\boldsymbol{\beta}, \mathbf{b}_{-i}^t)$. Then we have*

$$u_i(\mathbf{h}^t) = \sum_{j=1}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + j \left(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j - \beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t)) \right) \Big] . \tag{10}$$

Proof. Given bid profile $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K)$, we also define $\beta_{K+1} = 0$.

If $x_i(\mathbf{h}^t) = 0$ then both sides of equation (10) are zero, so the statement holds. Thus it remains to prove equation (10) when $x_i(\mathbf{h}^t) > 0$. Let $u_{i,j}(\mathbf{h}^t) = \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - p(\mathbf{h}^t))$ for each $j \in [K]$. The term $u_{i,j}(\mathbf{h}^t)$ represents the amount of utility obtained from the j-th unit acquired by player i at price $p(\mathbf{h}^t)$, so $u_i(\mathbf{h}^t) = \sum_{j=1}^{K} u_{i,j}(\mathbf{h}^t)$.

We first show that for each $j \in [K]$:

$$u_{i,j}(\mathbf{h}^t) = \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + \sum_{k=j}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] . \tag{11}$$

To prove (11), we will rewrite $u_{i,j}(\mathbf{h}^t)$ by considering three cases and writing a unified expression for all of them.

1. $x_i(\mathbf{h}^t) = j$**.** Then

$$\begin{aligned} u_{i,j}(\mathbf{h}^t) &= v_{i,j} - p(\mathbf{h}^t) \\ &= (v_{i,j} - \beta_j) + (\beta_j - p(\mathbf{h}^t)) \\ &= \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t) \\ &= \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + \sum_{k=j}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] . \end{aligned}$$

(Since $\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}} = 0 \; \forall k \geq j$ and $\mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}} = 1$ if and only if $k = j$.)

2. $x_i(\mathbf{h}^t) > j$**.** Then $x_i(\mathbf{h}^t) = j + \ell$, for some $\ell \in \{1, \ldots, K - j\}$. We have

$$\begin{aligned} u_{i,j}(\mathbf{h}^t) &= v_{i,j} - p(\mathbf{h}^t) \\ &= (v_{i,j} - \beta_j) + (\beta_j - \beta_{j+\ell}) + (\beta_{j+\ell} - p(\mathbf{h}^t)) \\ &= (v_{i,j} - \beta_j) + \left(\sum_{k=j}^{j+\ell-1} \beta_k - \beta_{k+1} \right) + (\beta_{j+\ell} - p(\mathbf{h}^t)) . \end{aligned} \tag{12}$$

We are in the case where $\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}} = 1$, $\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}} = 1$ if and only if $k \in \{j, \ldots, j + \ell - 1\}$, and $\mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}} = 1$ if and only if $k = j + \ell$. Adding indicators to the terms in (12) gives

$$u_{i,j}(\mathbf{h}^t) = \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + \sum_{k=j}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] .$$

3. $x_i(\mathbf{h}^t) < j$. Then $\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}} = 0$, $\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}} = 0$ for all $k \in \{j, \ldots, K\}$, and $\mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}} = 0$ for all $k \in \{j, \ldots, K\}$. Thus we trivially have the identity required by the lemma statement since $u_{i,j}(\mathbf{h}^t) = 0$ and the right hand side of (11) is zero as well.

Thus equation (11) holds in all three cases as required.

For each $i \in [n]$ and $j \in [K]$, define

$$A_{i,j} = \sum_{k=j}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] . \tag{13}$$

Then equation (11) is equivalent to

$$u_{i,j}(\mathbf{h}^t) = \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + A_{i,j}, \tag{14}$$

Summing equation (13) over all $j \in [K]$ gives

$$\sum_{j=1}^{K} A_{i,j} = \sum_{j=1}^{K}\sum_{k=j}^{K} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] \tag{15}$$

$$\overset{a}{=} \sum_{k=1}^{K}\sum_{j=1}^{k} \Big[\mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}}(\beta_k - \beta_{k+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}}(\beta_k - p(\mathbf{h}^t)) \Big] \tag{16}$$

$$= \sum_{k=1}^{K}(\beta_k - \beta_{k+1}) \sum_{j=1}^{k} \mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}} + \sum_{k=1}^{K}(\beta_k - p(\mathbf{h}^t)) \sum_{j=1}^{k} \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}} \tag{17}$$

$$= \sum_{k=1}^{K}(\beta_k - \beta_{k+1}) \cdot k \cdot \mathbb{1}_{\{x_i(\mathbf{h}^t)>k\}} + \sum_{k=1}^{K}(\beta_k - p(\mathbf{h}^t)) \cdot k \cdot \mathbb{1}_{\{x_i(\mathbf{h}^t)=k\}} \tag{18}$$

where equation (a) holds because we change the order of the double summations and in the double summations, we consider any (integer) pair of (j, k) such that $1 \leq k \leq j \leq K$.

Then we can rewrite the utility $u_i(\mathbf{h}^t)$ as

$$\begin{aligned}
u_i(\mathbf{h}^t) &= \sum_{j=1}^{K} u_{i,j}(\mathbf{h}^t) \\
&= \sum_{j=1}^{K} \left[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + A_{i,j} \right] && \text{(By equation (11))} \\
&= \sum_{j=1}^{K} \mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + \sum_{j=1}^{K}(\beta_j - \beta_{j+1}) \cdot j \cdot \mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}} + \sum_{k=1}^{K}(\beta_j - p(\mathbf{h}^t)) \cdot j \cdot \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}} \\
& && \text{(By equation (18))} \\
&= \sum_{j=1}^{K} \left[\mathbb{1}_{\{x_i(\mathbf{h}^t)\geq j\}}(v_{i,j} - \beta_j) + j\Big(\mathbb{1}_{\{x_i(\mathbf{h}^t)>j\}}(\beta_j - \beta_{j+1}) + \mathbb{1}_{\{x_i(\mathbf{h}^t)=j\}}(\beta_j - p(\mathbf{h}^t)) \Big) \right],
\end{aligned}$$

as required by the lemma statement. □

Lemma 2. *In the setting of Theorem 1, the edge weights of the graph G_i can be computed in polynomial time.*

Proof. Recall the graph G_i is constructed given as parameters the number n of players, the number of units K, a player i with valuation $\mathbf{v}_i$, discretization level $\varepsilon > 0$, and a bid history $H_{-i} = (\mathbf{b}^1_{-i}, \ldots, \mathbf{b}^T_{-i})$ by players other than i. Thus the goal is to show the edge weights of G_i can be computed in polynomial time in the bit length of these parameters.

Towards this end, we will show there exist efficiently computable values $I^t_{>j}, I^t_{\geq j}, I^t_j \in \{0,1\}$ and $q^t \in \mathcal{S}_i$ such that the weight w_e of edge $e = (z_{r,j}, z_{s,j+1})$ is equal to

$$w_e = \sum_{t=1}^{T} I^t_{\geq j}\left(v_{i,j} - r\right) + j\Big(I^t_{>j}\left(r - s\right) + I^t_j\left(r - q^t\right)\Big)\,. \tag{19}$$

Roughly, q^t will correspond to the price and $I^t_{>j}, I^t_{\geq j}, I^t_j$ will tell whether player i gets more than j units, at least j units, or exactly j units, respectively, at some profile $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathcal{S}_i$ with $\beta_j = r$ and $\beta_{j+1} = s$. The choice of $\boldsymbol{\beta}$ in will not matter, as long as $\beta_j = r$ and $\beta_{j+1} = s$.

We prove equation (19) in several steps:

Step (i). For each $t \in [T]$, define $\Gamma^t : \mathbb{R} \to \mathbb{R}$, where

$$\Gamma^t(x) = \Big|\Big\{(\ell, j) \mid \left(b^t_{\ell,j} > x \text{ and } \ell \in [n] \setminus \{i\}, j \in [K]\right) \text{ or } \left(b^t_{\ell,j} = x \text{ and } \ell \in [n], \ell < i, j \in [K]\right)\Big\}\Big|\,.$$

Thus $\Gamma^t(x)$ counts the bids in profile $\mathbf{b}^t_{-i}$ that would have priority to a bid of value x submitted by player i (i.e. it counts bids strictly higher than x and submitted by players other than i, as well as bids equal to x but submitted by players lexicographically before i).

Step (ii). Recall the edge is denoted $e = (z_{r,j}, z_{s,j+1})$. Let $\boldsymbol{\beta} \in \mathcal{S}_i$ be an arbitrary bid profile with $\beta_j = r$ and $\beta_j = s$. Also define $\beta_{K+1} = 0$.

If $\Gamma^t(s) < K - j$, then at $\mathbf{h}^t = (\boldsymbol{\beta}, \mathbf{b}^t_{-i})$ player i receives one unit for each of the bids $\beta_1, \ldots, \beta_{j+1}$, since there are at most $K - j - 1$ bids of other players that have higher priority than player i's highest $j+1$ bids. Else, player i does not get more than j units. Thus $I^t_{>j} = \mathbb{1}_{\{x_i(\mathbf{h}^t) > j\}}$.

Step (iii). If $\Gamma^t(r) \leq K - j$, then player i receives at least j units at $(\boldsymbol{\beta}, \mathbf{b}^t_{-i})$. The corresponding indicator is

$$I^t_{\geq j} = \mathbb{1}_{\{\Gamma^t(s) \leq K - j\}} = \mathbb{1}_{\{x_i(\mathbf{h}^t) \geq j\}}\,.$$

Step (iv). If $\Gamma^t(r) \leq K - j$ and $\Gamma^t(s) > K - j$, then player i receives exactly j units at $(\boldsymbol{\beta}, \mathbf{b}^t_{-i})$. The indicator is

$$I^t_j = \mathbb{1}_{\{\Gamma^t(r) \leq K - j\}} \cdot \mathbb{1}_{\{\Gamma^t(s) > K - j\}} = \mathbb{1}_{\{x_i(\mathbf{h}^t) = j\}}\,.$$

Now suppose $I^t_j = 1$. Then we show the price $q^t = p(\mathbf{h}^t)$ can be computed precisely without knowing the whole bid $\boldsymbol{\beta}$.

Consider the multiset of bids $B = \mathbf{b}^t_{-i} \cup \{\beta_1, \ldots, \beta_{j+1}\}$, recalling β_{K+1} was defined as zero. Sort B in descending order. We know the top $j - 1$ bids of player i are winning, so the price is not determined by any of them. Thus the price is determined by $\beta_j = r$, $\beta_{j+1} = s$, or by one of the bids in $\mathbf{b}^t_{-i}$. Remove elements $\beta_1, \ldots, \beta_{j-1}$ from B and set the price as follows:

- ***Case (iv.a).*** For $(K+1)$-st price auction: set q^t to the $(K + 2 - j)^{\text{th}}$ highest value in B.
- ***Case (iv.b).*** For K-th price auction: set q^t to the $(K + 1 - j)^{\text{th}}$ highest value in B.

Step (v). Combining steps (i-iv), the weight of edge $e = (z_{r,j}, z_{s,j+1})$ can be rewritten as

$$\begin{aligned} w_e &= \sum_{t=1}^{T} \mathbb{1}_{\{x_i(\mathbf{h}^t) \geq j\}}\left(v_{i,j} - r\right) + j\Big[\mathbb{1}_{\{x_i(\mathbf{h}^t) > j\}}\left(r - s\right) + \mathbb{1}_{\{x_i(\mathbf{h}^t) = j\}}\left(r - p(\mathbf{h}^t)\right)\Big] && \text{(By Definition 1)} \\ &= \sum_{t=1}^{T} I^t_{\geq j}\left(v_{i,j} - r\right) + j\Big[I^t_{>j}\left(r - s\right) + I^t_j\left(r - q^t\right)\Big]\,. && \text{(By steps (i-iv))} \end{aligned}$$

Thus the weight of each edge can be computed in polynomial time as required. □

C Appendix: Online Setting

In this section we include the material omitted from the main text for the online setting.

Before studying the full information and bandit feedback models in greater detail, we replace the set $\mathcal{S}_i$ of candidate bids for player i from equation (1) of the offline section by a coarser set

$$\mathcal{S}_\varepsilon = \{\varepsilon, 2\varepsilon, \cdots, \lceil v_{i,1}/\varepsilon \rceil \varepsilon\} .$$

The regret, which was defined in equation (3), depends on whether the model is the K-th or $(K+1)$-st price auction; however the next lemma holds for both variants of the auction.

Lemma 3. *For all $\varepsilon > 0$, let $\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon)$ be the counterpart of (3) where $\mathcal{S}_i$ is replaced by the set $\mathcal{S}_\varepsilon$. Then $\mathrm{Reg}_i(\pi_i, H_{-i}^T) \leq \mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) + TK\varepsilon$.*

Proof. First observe that it is not beneficial to bid above $v_{i,1}$, so we can assume without loss of generality that the maximizer $\boldsymbol{\beta}$ in (3) satisfies $\beta_j \leq v_{i,1}$ for all $j \in [K]$. Now we convert $\boldsymbol{\beta}$ into another bid vector $\boldsymbol{\beta}^\varepsilon \in \mathcal{S}_\varepsilon^K$ as follows: for each $j \in [K]$, let

$$\beta_j^\varepsilon = \min\{y \in \mathcal{S}_\varepsilon \mid y \geq \beta_j\} .$$

Clearly $\beta_j \leq \beta_j^\varepsilon \leq \beta_j + \varepsilon$.

Then we claim that the next inequalities hold:

$$p(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) \leq p(\boldsymbol{\beta}, \mathbf{b}_{-i}^t) + \varepsilon; \tag{20}$$

$$\mathbb{1}_{\{x_i(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) \geq j\}} \geq \mathbb{1}_{\{x_i(\boldsymbol{\beta}, \mathbf{b}_{-i}^t) \geq j\}} . \tag{21}$$

Inequality (20) follows since

- $p(\mathbf{b})$ is either the K-th or the $(K+1)$-st largest element of $\mathbf{b}$, and
- increasing each entry of $\mathbf{b}$ by at most ε can only increase $p(\mathbf{b})$ by no more than ε.

Inequality (21) is due to the fact that bidding a higher price can only help to win the unit.

Consequently, we have

$$\begin{aligned}
\sum_{t=1}^{T}\sum_{j=1}^{K} \big(v_{i,j} - p(\boldsymbol{\beta}, \mathbf{b}_{-i}^t)\big) \cdot \mathbb{1}_{\{x_i(\boldsymbol{\beta}, \mathbf{b}_{-i}^t) \geq j\}} &\leq \sum_{t=1}^{T}\sum_{j=1}^{K} \big(v_{i,j} - p(\boldsymbol{\beta}, \mathbf{b}_{-i}^t)\big) \cdot \mathbb{1}_{\{x_i(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) \geq j\}} \\
&\leq \sum_{t=1}^{T}\sum_{j=1}^{K} \big(v_{i,j} - p(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) - \varepsilon\big) \cdot \mathbb{1}_{\{x_i(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) \geq j\}} \\
&\leq \sum_{t=1}^{T}\sum_{j=1}^{K} \big(v_{i,j} - p(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t)\big) \cdot \mathbb{1}_{\{x_i(\boldsymbol{\beta}^\varepsilon, \mathbf{b}_{-i}^t) \geq j\}} + TK\varepsilon .
\end{aligned}$$

This gives the desired statement of the lemma. □

C.1 Full Information Feedback

In this section we include the omitted details for the full information setting.

We begin with the formal definition of the graph $G^t = (V, E, w^t)$ used by the online learning algorithm with full information feedback.

Definition 2 (The graph G^t)**.** *Given valuation $\mathbf{v}_i$ of player i, bid profile $\mathbf{b}_{-i}^t$ of the players at round t, and $\varepsilon > 0$, construct a graph $G^t = (V, E, w^t)$ as follows.*

- ***Vertices.*** *Create a vertex $z_{s,j}$ for each $s \in \mathcal{S}_\varepsilon$ and index $j \in [K]$. We say vertex $z_{s,j}$ is in layer j. Add source z_- and sink z_+.*
- ***Edges.*** *For each index $j \in [K-1]$ and pair of bids $r, s \in \mathcal{S}_\varepsilon$ with $r \geq s$, create a directed edge from vertex $z_{r,j}$ to vertex $z_{s,j+1}$. Moreover, add edges from source z_- to each node in layer 1 and from each node in layer K to the sink z_+.*

- ***Edge weights.*** *For each edge* $e = (z_{r,j}, z_{s,j+1})$ *or* $e = (z_{r,K}, z_+)$, *let* $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_K) \in \mathcal{S}_\varepsilon^K$ *be a bid vector with* $\beta_j = r$ *and* $\beta_{j+1} = s$ *(we define* $s = 0$ *if* $j = K$*). For each* t*, let* $\mathbf{h}^t = (\boldsymbol{\beta}, \mathbf{b}_{-i}^t)$*. Define the weight of edge* e *as*

$$w^t(e) = \mathbb{1}_{\{x_i(\mathbf{h}^t) \geq j\}} (v_{i,j} - r) + j \left[\mathbb{1}_{\{x_i(\mathbf{h}^t) > j\}} (r - s) + \mathbb{1}_{\{x_i(\mathbf{h}^t) = j\}} \left(r - p(\mathbf{h}^t) \right) \right].$$

The edges incoming from z_- *have weight zero.*

The next observation follows immediately from the definition.

Observation 2. *The next properties hold:*

- *The weight* $w^t(e)$ *of each edge* e *of* G^t *can be computed efficiently (see Lemma 2).*
- *There is a bijective map between bid vectors* $\boldsymbol{\beta} \in \mathcal{S}_\varepsilon^K$ *of player* i *and paths from the source to the sink of* G^t *(see proof of Theorem 1).*

Next we include the main theorem with its proof for the online learning algorithm under full information feedback.

Theorem 2 (restated). *For each player* i *and time horizon* T*, under full-information feedback, Algorithm 2 runs in time* $O(T^2)$ *and guarantees the player's regret is at most* $O\left(v_{i,1}\sqrt{TK^3 \log T}\right)$.

Proof. The detailed algorithm description is presented as Algorithm 2. At a high level, the proof has three parts: (i) showing that Algorithm 2 is a correct implementation of the Hedge algorithm where each expert is a path from source to sink in the DAG G^t for time t, (ii) bounding its regret for an appropriate choice of the learning rate, and (iii) bounding the runtime.

Step I: Algorithm 2 is a correct implementation of Hedge. We show that Algorithm 2 is the same as the Hedge algorithm in which every path in the DAG is equivalent to an expert in the Hedge algorithm.

To do so, for each vertex u in the DAG, let $\phi^t(u, \cdot)$ be a probability distribution over the outneighbors of u. Then, given the recursive sampling of the bids based on the probability distribution ϕ^t's, the probability that a path $\mathfrak{p}$ is chosen in Algorithm 2 is equal to

$$P^t(\mathfrak{p}) = \prod_{e \in \mathfrak{p}} \phi^t(e), \tag{22}$$

where ϕ^t's are updated in equation (5), which is restated below:

$$\phi^t(e) = \phi^{t-1}(e) \cdot \exp\left(\eta \cdot w^{t-1}(e)\right) \cdot \frac{\Gamma^{t-1}(v)}{\Gamma^{t-1}(u)}, \quad \text{for all } e = (u,v) \in E(G^t). \qquad \text{(5 revisited .)}$$

On the other hand, in the Hedge algorithm, the path probabilities are updated as follows: Given a learning rate η, define $P_h^1(\mathfrak{p}) = \prod_{e \in \mathfrak{p}} \phi^1(e)$, and for $t \geq 2$,

$$P_h^t(\mathfrak{p}) = \frac{P_h^{t-1}(\mathfrak{p}) \exp(\eta \sum_{e \in \mathfrak{p}} w^{t-1}(e))}{\sum_{\mathfrak{q}} P_h^{t-1}(\mathfrak{q}) \exp(\eta \sum_{e \in \mathfrak{q}} w^{t-1}(e))}. \tag{23}$$

The subscript 'h' in $P_h^t(\mathfrak{p})$ stands for Hedge. To show that the update rule in Algorithm 2 is equivalent to the one in Hedge, we first prove the following statement:

(†) For any vertex u in the graph, let $\mathcal{P}(u)$ be the set of all paths from u to z_+. Then

$$\Gamma^{t-1}(u) = \sum_{\mathfrak{q} \in \mathcal{P}(u)} \prod_{e \in \mathfrak{q}} \left[\phi^{t-1}(e) \exp(\eta w^{t-1}(e)) \right]. \tag{24}$$

We prove equation (24) by induction on u, from the bottom layer to the top layer. If $u = z_+$, by definition $\Gamma^{t-1}(z_+) = 1$, and (24) holds. Now suppose that (24) holds for all u in the $(k+1)$-st layer, for some $0 \le k \le K$. Then if u is in the k-th layer, the recursion of Γ^{t-1} gives that

$$\begin{aligned}\Gamma^{t-1}(u) &= \sum_{v:(u,v)\in E} \phi^{t-1}((u,v)) \exp(\eta w^{t-1}((u,v)) \cdot \Gamma^{t-1}(v) \\ &= \sum_{v:(u,v)\in E} \phi^{t-1}((u,v)) \exp(\eta w^{t-1}((u,v))) \cdot \sum_{\mathfrak{q}\in\mathcal{P}(v)} \prod_{e\in\mathfrak{q}} \left[\phi^{t-1}(e)\exp(\eta w^{t-1}(e))\right] \\ &= \sum_{\mathfrak{q}\in\mathcal{P}(u)} \prod_{e\in\mathfrak{q}} \left[\phi^{t-1}(e)\exp(\eta w^{t-1}(e))\right],\end{aligned}$$

and therefore (24) holds.

Having shown equation (24), we prove by induction on t our claim that Algorithm 2 is a correct implementation of Hedge, i.e. that $P^t(\mathfrak{p}) = P_h^t(\mathfrak{p})$ for all paths $\mathfrak{p}$ and rounds t. For $t = 1$, by definition of our initialization we have $P_h^1(\mathfrak{p}) = P^1(\mathfrak{p})$. Suppose that at time $t-1$, for every path $\mathfrak{p}$ we have $P_h^{t-1}(\mathfrak{p}) = P^{t-1}(\mathfrak{p}) = \prod_{e\in\mathfrak{p}} \phi^{t-1}(e)$. Then at time t, it holds that

$$\begin{aligned}P^t(\mathfrak{p}) = \prod_{e\in\mathfrak{p}} \phi^t(e) &\overset{\text{(a)}}{=} \prod_{e=(u,v)\in\mathfrak{p}} \left[\phi^{t-1}(e) \cdot \exp\left(\eta \cdot w^{t-1}(e)\right) \cdot \frac{\Gamma^{t-1}(v)}{\Gamma^{t-1}(u)}\right] \\ &\overset{\text{(b)}}{=} P_h^{t-1}(\mathfrak{p}) \exp\left(\eta \sum_{e\in\mathfrak{p}} w^{t-1}(e)\right) \cdot \frac{\Gamma^{t-1}(z_+)}{\Gamma^{t-1}(z_-)},\end{aligned}$$

where step (a) is due to equation (5), and step (b) follows using telescoping and the induction hypothesis $P_h^{t-1}(\mathfrak{p}) = \prod_{e\in\mathfrak{p}} \phi^{t-1}(e)$.

Given equation (23), by applying equation (24) to $u \in \{z_-, z_+\}$ and the induction hypothesis $P_h^{t-1}(\mathfrak{p}) = \prod_{e\in\mathfrak{p}} \phi^{t-1}(e)$, the induction is complete.

Thus Algorithm 2 is a correct implementation of Hedge.

Step II: Regret upper bound. Let

$$\varepsilon = v_{i,1}\sqrt{K/T}\,. \tag{25}$$

Applying Lemma 3 yields

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T) \le \mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) + TK\varepsilon = \mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) + v_{i,1}\sqrt{TK^3}, \tag{26}$$

where $\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon)$ is the counterpart of (3) where $\mathcal{S}_i$ is replaced by the set

$$\mathcal{S}_\varepsilon = \{\varepsilon, 2\varepsilon, \cdots, \lceil v_{i,1}/\varepsilon\rceil\varepsilon\}\,.$$

We also claim that $\sum_{e\in\mathfrak{p}} w^t(e) \le Kv_{i,1}$ for each path $\mathfrak{p}$ from source to sink in G^t. To see this, let $\boldsymbol{\beta}$ be the bid vector corresponding to path $\mathfrak{p}$. As shown in Lemma 1 and using the fact that edges outgoing from z_- have weight zero, we have $u_i(\boldsymbol{\beta}, b_{-i}^t) = \sum_{e\in\mathfrak{p}} w^t(e)$. Since player i's utility satisfies $u_i(\boldsymbol{\beta}, \mathbf{b}_{-i}^t) \le V_i(x_i(\boldsymbol{\beta}, \mathbf{b}_{-i}^t)) \le Kv_{i,1}$, we obtain $\sum_{e\in\mathfrak{p}} w^t(e) \le Kv_{i,1}$.

By Step I, Algorithm 2 is a correct implementation of Hedge. To bound the regret of the algorithm, we will invoke Corollary 1—which is a slight variant of [CBL06, Theorem 2.2]—with the following parameters:

- N experts, where each expert is a path from source to sink in G^t;
- learning rate η, time horizon T, and maximum reward $L = Kv_{i,1}$;
- initial distribution σ on the experts, where $\sigma_{\mathfrak{p}} = P^1(\mathfrak{p})$ for each expert (path from source to sink) $\mathfrak{p}$.

By Corollary 1, we obtain

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) \le \frac{1}{\eta} \max_{\mathfrak{p}\in[N]} \log\left(\frac{1}{\sigma_{\mathfrak{p}}}\right) + \frac{TL^2\eta}{8}\,. \tag{27}$$

Recall Algorithm 2 initially selects a path by starting at the source z_- and then performing an unbiased random walk in the layered DAG until reaching the sink z_+. Since the number of vertices in each layer is $\lceil v_{i,1}/\varepsilon\rceil$, the initial probability of selecting a particular expert (i.e. path from source to sink) $\mathfrak{p}$ is

$$\sigma_{\mathfrak{p}} = P^1(\mathfrak{p}) \ge \frac{1}{\lceil v_{i,1}/\varepsilon\rceil^K}, \quad \forall \mathfrak{p}\in[N]\,. \tag{28}$$

Substituting (28) in (27) yields

$$\begin{aligned}\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) &\le \frac{1}{\eta} \max_{\mathfrak{p}\in[N]} \log\left(\frac{1}{\sigma_{\mathfrak{p}}}\right) + \frac{TL^2\eta}{8} \le \frac{K\log\left(\lceil v_{i,1}/\varepsilon\rceil\right)}{\eta} + \frac{TL^2\eta}{8} \\ &= \frac{K\log\left(\left\lceil\sqrt{\frac{T}{K}}\right\rceil\right)}{\eta} + \frac{T\,(Kv_{i,1})^2\,\eta}{8} \qquad (\text{Since } \epsilon = v_{i,1}\sqrt{\tfrac{K}{T}} \text{ and } L = Kv_{i,1}.) \\ &\le \frac{K\log T}{\eta} + \frac{T\,(Kv_{i,1})^2\,\eta}{8}\,. \end{aligned} \tag{29}$$

For $\eta = \sqrt{\log T}/(v_{i,1}\sqrt{KT})$, inequality (29) gives

$$\begin{aligned}\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) &\le \frac{K\cdot\log T}{\sqrt{\log T}/(v_{i,1}\sqrt{KT})} + \frac{T\sqrt{\log T}}{8(v_{i,1}\sqrt{KT})}\cdot(Kv_{i,1})^2 \\ &= \frac{9\,v_{i,1}}{8}\sqrt{TK^3\log T}\,. \end{aligned} \tag{30}$$

Combining inequalities (26) and (30), we get that the regret of player i when running Algorithm 2 is

$$\begin{aligned}\mathrm{Reg}_i(\pi_i, H_{-i}^T) &\le \mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) + v_{i,1}\sqrt{TK^3} \\ &\le \frac{9\,v_{i,1}}{8}\sqrt{TK^3\log T} + v_{i,1}\sqrt{TK^3} \in O\left(v_{i,1}\sqrt{TK^3\log T}\right)\,. \end{aligned} \tag{31}$$

This completes the proof of the regret upper bound.

Polynomial time implementation. Finally we analyze the running time of the above algorithm. At every step, the computation of path kernels traverses all edges (note that each edge only appears in the sum once) and could be done in $O(|E|) = O(Kv_{i,1}^2/\varepsilon^2)$ time. Similarly, the update of edge probabilities ϕ^t for all edges also takes $O(|E|) = O(Kv_{i,1}^2/\varepsilon^2)$ time.

Therefore, the overall computational complexity is $O(TKv_{i,1}^2/\varepsilon^2) = O(T^2)$, where we used the choice of $\varepsilon = v_{i,1}\sqrt{K/T}$. This completes the proof of the theorem. □

C.2 Bandit Feedback

In this section we include the main theorem and proof for the bandit setting. Recall that our algorithm for the bandit setting is the same as Algorithm 2, only with $w^t(e)$ replaced by $\widehat{w}^t(e)$:

$$\widehat{w}^t(e) = \overline{w}(e) - \frac{\overline{w}(e) - w^t(e)}{p^t(e)}\mathbb{1}_{\{e\in\mathfrak{p}^t\}}, \qquad \text{with } p^t(e) = \sum_{\mathfrak{p}:e\in\mathfrak{p}} P^t(\mathfrak{p}),$$

with P^t given in (22), and

$$\overline{w}(e) = \begin{cases} v_{i,1} - r + j(r-s) & \text{if } e = (z_{r,j}, z_{s,j+1}), \\ v_{i,1} - r + Kr & \text{if } e = (z_{r,K}, z_+), \\ 0 & \text{if } e = (z_-, z_{r,1}). \end{cases}$$

The resolution parameter and learning rate are chosen to be

$$\varepsilon = v_{i,1} \min\{(K^3 \log T/T)^{1/4}, 1\}, \quad \eta = \min\left\{\varepsilon\sqrt{\log(v_{i,1}/\varepsilon)/(TK^3 v_{i,1}^4)}, 1/(Kv_{i,1})\right\}. \quad (32)$$

Theorem 3 (restated). *For each player i and time horizon T, under the bandit feedback, there is an algorithm for bidding that runs in time $O(TK + K^{-5/4}T^{7/4})$ and guarantees the player's regret is at most $O(\min\{v_{i,1}(T^3K^7 \log T)^{1/4}, v_{i,1}KT\})$.*

Proof. By Lemma 3 and the choice of ε in (32), it suffices to show that the above algorithm π_i runs in time $O(TKv_{i,1}^3/\varepsilon^3)$ and achieves

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) = O\left(v_{i,1}^2\sqrt{TK^5 \log(v_{i,1}/\varepsilon)}/\varepsilon + v_{i,1}K^2 \log(v_{i,1}/\varepsilon)\right).$$

The claimed result then follows from the fact that the regret is always upper bounded by $O(v_{i,1}KT)$.

Before we proceed to the proof, we first comment on the choice of estimator $\widehat{w}^t(e)$. First of all, this is an unbiased estimator of $w^t(e)$, i.e. $\mathbb{E}_{\mathfrak{p}^t\sim P^t}[\widehat{w}^t(e)] = w^t(e)$ for every edge e in G^t. Second, instead of using the natural importance-weighted estimator $\widehat{w}^t(e) = w^t(e)\mathbb{1}_{\{e\in\mathfrak{p}^t\}}/p^t(e)$, the current form in (6) is the loss-based importance-weighted estimator used for technical reasons, similar to [LS20, Eqn. (11.6)]. Third, by exploiting our DAG structure and the definition of $w^t(e)$ in (4), we construct an edge-specific quantity $\overline{w}(e)$ which always upper bounds $w^t(e)$.

We now analyze the regret of the algorithm with $\widehat{w}^t(e)$ given by (6). The standard EXP3 analysis (see, e.g. [LS20, Chapter 11]) gives that

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) \le \frac{1}{\eta}\max_{\mathfrak{p}} \log\frac{1}{P^1(\mathfrak{p})} + \sum_{t=1}^T \mathbb{E}\left[\frac{1}{\eta}\log\left(\sum_{\mathfrak{p}} P^t(\mathfrak{p})e^{\eta\widehat{w}^t(\mathfrak{p})}\right) - \sum_{\mathfrak{p}} P^t(\mathfrak{p})\widehat{w}^t(\mathfrak{p})\right], \quad (33)$$

where $\widehat{w}^t(\mathfrak{p}) = \sum_{e\in\mathfrak{p}} \widehat{w}^t(e)$ is the estimated total weight of path $\mathfrak{p}$, and the expectation is with respect to the randomness in the estimator $\widehat{w}^t(e)$. For every path $\mathfrak{p} = (z_-, z_{r_1,1}, \cdots, z_{r_K,K}, z_+)$ from the source to the sink, we have (by convention $r_{K+1} = 0$):

$$\widehat{w}^t(\mathfrak{p}) = \sum_{e\in\mathfrak{p}} w^t(e) \le \sum_{e\in\mathfrak{p}} \overline{w}(e) = \sum_{j=1}^K (v_{i,1} - (j-1)r_j + jr_{j+1}) = Kv_{i,1}.$$

Since $e^x \le 1 + x + x^2$ whenever $x \le 1$ and $\log(1+y) \le y$ whenever $y > -1$, if $\eta \le 1/(Kv_{i,1})$, inequality (33) gives

$$\begin{aligned}\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) &\le \frac{1}{\eta}\max_{\mathfrak{p}} \log\frac{1}{P^1(\mathfrak{p})} + \eta\sum_{t=1}^T\sum_{\mathfrak{p}} P^t(\mathfrak{p})\,\mathbb{E}[\widehat{w}^t(\mathfrak{p})^2] \\ &\le \frac{K}{\eta}\log\left\lceil\frac{v_{i,1}}{\varepsilon}\right\rceil + \eta\sum_{t=1}^T\sum_{\mathfrak{p}} P^t(\mathfrak{p})\cdot(K+1)\sum_{e\in\mathfrak{p}}\mathbb{E}[\widehat{w}^t(e)^2],\end{aligned}$$

where the last equality uses $(\sum_{i=1}^n x_i)^2 \le n\sum_{i=1}^n x_i^2$, and that each path $\mathfrak{p}$ from the source to the sink has length $K+1$. To proceed, note that

$$\begin{aligned}\mathbb{E}[\widehat{w}^t(e)^2] &= \overline{w}(e)^2\cdot(1-p^t(e)) + \left(\overline{w}(e) - \frac{\overline{w}(e)-w^t(e)}{p^t(e)}\right)^2\cdot p^t(e) \\ &= w^t(e)^2 + (\overline{w}(e)-w^t(e))^2\cdot\left(\frac{1}{p^t(e)}-1\right),\end{aligned}$$

and therefore

$$\begin{aligned}\sum_{\mathfrak{p}} P^t(\mathfrak{p})\sum_{e\in\mathfrak{p}}\mathbb{E}[\widehat{w}^t(e)^2] &= \sum_{\mathfrak{p}} P^t(\mathfrak{p})\sum_{e\in\mathfrak{p}}\left[w^t(e)^2 + (\overline{w}(e)-w^t(e))^2\cdot\left(\frac{1}{p^t(e)}-1\right)\right] \\ &= \sum_e\left[w^t(e)^2 + (\overline{w}(e)-w^t(e))^2\cdot\left(\frac{1}{p^t(e)}-1\right)\right]\sum_{\mathfrak{p}:e\in\mathfrak{p}} P^t(\mathfrak{p}) \\ &= \sum_e\left[w^t(e)^2p^t(e) + (\overline{w}(e)-w^t(e))^2(1-p^t(e))\right] \le \sum_e \overline{w}(e)^2,\end{aligned}$$

where in the middle we have used the definition $\sum_{\mathfrak{p}:e\in\mathfrak{p}} P^t(\mathfrak{p}) = p^t(e)$ in (6). To further upper bound the above quantity, note that

$$\begin{aligned}\sum_e \overline{w}(e)^2 &\le \sum_{j=1}^K \sum_{1\le s\le r\le\lceil v_{i,1}/\varepsilon\rceil} (v_{i,1}+(j-1)r\varepsilon)^2 \\ &\le \left\lceil \frac{v_{i,1}}{\varepsilon}\right\rceil \sum_{j=1}^K \sum_{r=1}^{\lceil v_{i,1}/\varepsilon\rceil} (V+(j-1)r\varepsilon)^2 \\ &\le 2\left\lceil \frac{v_{i,1}}{\varepsilon}\right\rceil \sum_{j=1}^K \sum_{r=1}^{\lceil v_{i,1}/\varepsilon\rceil} (v_{i,1}^2+(j-1)^2r^2\varepsilon^2) = O\left(\frac{K^3v_{i,1}^4}{\varepsilon^2}\right).\end{aligned}$$

A combination of the above inequalities shows that as long as $\eta \le 1/(Kv_{i,1})$,

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) = O\left(\frac{K}{\eta}\log\frac{v_{i,1}}{\varepsilon} + \frac{\eta T K^4 v_{i,1}^4}{\varepsilon^2}\right).$$

Next, by the choice of η in (32), we have

$$\mathrm{Reg}_i(\pi_i, H_{-i}^T, \varepsilon) = O\left(\frac{v_{i,1}^2\sqrt{TK^5\log(v_{i,1}/\varepsilon)}}{\varepsilon} + K^2 v_{i,1}\log\frac{v_{i,1}}{\varepsilon}\right).$$

As for the computational complexity, the only difference from the Algorithm 2 is the additional computation of the estimated weights $\widehat{w}^t(e)$ in (6), or equivalently, the marginal probability $p^t(e)$. As P^t has a product structure in (22), the celebrated message passing algorithm in graphical models [WJ08, Section 2.5.1] takes $O(|E|v_{i,1}/\varepsilon) = O(Kv_{i,1}^3/\varepsilon^3)$ time to compute all edge marginals $\{p^t(e)\}_{e\in E}$. Therefore the overall computational complexity is $O(TKv_{i,1}^3/\varepsilon^3)$. □

C.3 Regret Lower Bound

In this section we include the theorem and proof for the regret lower bound. The construction uses the $(K+1)$-st highest price, but is similar for the K-th highest price.

Theorem 4 (restated, formal). *Let $K \ge 2$. For any policy π_i used by player i, there exists a bid sequence $\{\mathbf{b}_{-i}^t\}_{t=1}^T$ for the other players such that the expected regret in (3) satisfies $\mathbb{E}[\mathrm{Reg}_i(\pi_i, \{\mathbf{b}_{-i}^t\}_{t=1}^T)] \ge cv_{i,1}K\sqrt{T}$, where $c > 0$ is an absolute constant.*

Proof. Without loss of generality assume that $K = 2k$ is an even integer, and by scaling we may assume that $v_{i,1} = 1$. Consider the following two scenarios:

- the utility of the bidder i is $v_{i,j} \equiv 1$, for all $j \in [K]$;
- at scenario 1, for every $t \in [T]$, the other bidders' bids are

$$\mathbf{b}_{-i}^t = \begin{cases} \left(\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3},0,\cdots,0\right) & \text{with probability } 0.5+\delta, \\ \left(\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3}\right) & \text{with probability } 0.5-\delta, \end{cases}$$

  where the number of non-zero entries is k in the first line and $2k$ in the second line, and $\delta \in (0, 1/4)$ is a parameter to be determined later. The randomness used at different times is independent.
- at scenario 2, for every $t \in [T]$, the other bidders' bids are

$$\mathbf{b}_{-i}^t = \begin{cases} \left(\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3},0,\cdots,0\right) & \text{with probability } 0.5-\delta, \\ \left(\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3},\frac{2}{3},\cdots,\frac{2}{3}\right) & \text{with probability } 0.5+\delta, \end{cases}$$

  where the number of non-zero entries is k in the first line and $2k$ in the second line, and $\delta \in (0, 1/4)$ is a parameter to be determined later. The randomness used at different times is independent.

We denote by P and Q the distributions of $\{\mathbf{b}_{-i}^t\}_{t=1}^T$ under scenarios 1 and 2, respectively, then

$$\begin{aligned} D_{\text{KL}}(P\|Q) &= T \cdot D_{\text{KL}}(\text{Bern}(0.5+\delta)\|\text{Bern}(0.5-\delta)) \\ &\le T \cdot \chi^2(\text{Bern}(0.5+\delta)\|\text{Bern}(0.5-\delta)) \\ &= T \cdot \frac{(2\delta)^2}{(0.5+\delta)(0.5-\delta)} \le \frac{64}{3} T\delta^2. \end{aligned}$$

Consequently, by [Tsy09, Lemma 2.6],

$$1 - \text{TV}(P,Q) \ge \frac{1}{2}\exp\left(-D_{\text{KL}}(P\|Q)\right) \ge \frac{1}{2}\exp\left(-\frac{64T\delta^2}{3}\right).$$

Next we investigate the separation between these two scenarios. It is clear that

$$\begin{aligned} \max_{\mathbf{b}_i} \mathbb{E}_P\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] &\ge \mathbb{E}_P\left[u_i((1,1,\cdots,1,0,\cdots,0);\mathbf{b}_{-i}^t)\right] \\ &= \left(\frac{1}{2}+\delta\right)\cdot k + \left(\frac{1}{2}-\delta\right)\cdot\frac{k}{3} = \frac{2+2\delta}{3}\cdot k; \\ \max_{\mathbf{b}_i} \mathbb{E}_Q\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] &\ge \mathbb{E}_Q\left[u_i((1,1,\cdots,1,1,\cdots,1);\mathbf{b}_{-i}^t)\right] \\ &= \left(\frac{1}{2}-\delta\right)\cdot\frac{2k}{3} + \left(\frac{1}{2}+\delta\right)\cdot\frac{2k}{3} = \frac{2k}{3}. \end{aligned}$$

Moreover, under $(P+Q)/2$ (i.e. $\mathbf{b}_{-i}^t$ follows a Bern$(1/2)$ distribution), suppose that the vector $\mathbf{b}_i$ has k' components smaller than $2/3$. Distinguish into two scenarios:

- if $k' < k$, then

$$\mathbb{E}_{(P+Q)/2}\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] \le \frac{1}{2}\cdot\frac{2k-k'}{3} + \frac{1}{2}\cdot\frac{2k-k'}{3} \le \frac{2k}{3};$$

- if $k' \ge k$, then

$$\mathbb{E}_{(P+Q)/2}\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] \le \frac{1}{2}\cdot(2k-k') + \frac{1}{2}\cdot\frac{2k-k'}{3} \le \frac{2k}{3}.$$

Therefore, it always holds that

$$\max_{\mathbf{b}_i} \mathbb{E}_{(P+Q)/2}\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] \le \frac{2k}{3}.$$

Consequently, for each $\mathbf{b}_i$,

$$\begin{aligned} &\max_{\mathbf{b}_i^\star} \mathbb{E}_P\left[u_i(\mathbf{b}_i^\star;\mathbf{b}_{-i}^t) - u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] + \max_{\mathbf{b}_i^\star} \mathbb{E}_Q\left[u_i(\mathbf{b}_i^\star;\mathbf{b}_{-i}^t) - u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] \\ &\ge \max_{\mathbf{b}_i^\star} \mathbb{E}_P\left[u_i(\mathbf{b}_i^\star;\mathbf{b}_{-i}^t)\right] + \max_{\mathbf{b}_i^\star} \mathbb{E}_Q\left[u_i(\mathbf{b}_i^\star;\mathbf{b}_{-i}^t)\right] - 2\max_{\mathbf{b}_i} \mathbb{E}_{(P+Q)/2}\left[u_i(\mathbf{b}_i;\mathbf{b}_{-i}^t)\right] \\ &\ge \frac{2+2\delta}{3}k + \frac{2k}{3} - 2\cdot\frac{2k}{3} = \frac{2\delta k}{3}. \end{aligned}$$

In other words, any bid vector $\mathbf{b}_i$ either incurs a total regret $(\delta kT)/3$ under P, or incurs a total regret $(\delta kT)/3$ under Q.

Now the classical two-point method (see, e.g. [Tsy09, Theorem 2.2]) gives

$$\mathbb{E}_{(P+Q)/2}[\text{Reg}_i(\pi_i, \{\mathbf{b}_{-i}^t\}_{t=1}^T)] \ge \frac{\delta kT}{3}\cdot(1-\text{TV}(P,Q)) \ge \frac{\delta kT}{6}\exp\left(-\frac{64T\delta^2}{3}\right)$$

for every $\delta \in (0, 1/4)$. Choosing $\delta = 1/(8\sqrt{T})$ gives that

$$\mathbb{E}_{(P+Q)/2}[\text{Reg}_i(\pi_i, \{\mathbf{b}_{-i}^t\}_{t=1}^T)] \ge \frac{K\sqrt{T}}{96e^{1/3}},$$

i.e. the claimed lower bound holds with $c = 1/(96e^{1/3})$. □

D Appendix: Equilibrium Analysis

In this section we show that the pure Nash equilibria with price zero of the $(K+1)$-st auction are the only ones that are robust to deviations by groups of players, captured through the notion of the core in the game among the bidders. The core of a game was formulated by Edgeworth [Edg81] and brought into game theory by Gillies [Gil59]. There is an extensive body of literature on the core of various games, including for auctions; see, e.g., analysis of collusion (cartel behavior) in first price auctions in [Pes00].

Our focus here is the core of the game among the bidders, where the auctioneer first sets the auction format and then the bidders can strategize and collude among themselves, without the auctioneer. Roughly speaking, a strategy profile is core-stable if no group C of players can deviate simultaneously (i.e. each player $i \in C$ deviates to some alternative strategy profile), such that each player in C weakly improves and the improvement is strict for at least one player. For a deviation to take place, the players in C coordinate and switch their strategies simultaneously, while the players outside C keep their previous strategies. Every core-stable strategy profile is also a Nash equilibrium, since a strategy profile that is stable against deviations by groups of players is also stable against deviations by individuals.

We consider two variants of the core, with and without monetary transfers [SS69, Bon63, Sha67]. In the case with transfers, the players can make monetary payments to each other, and so a strategy profile consists of a tuple of bids and transfers. In the case without transfers, the players cannot make such transfers and their strategy is the bid vector; thus the only agreement they can make in this case is to coordinate their bids.

D.1 Core with transfers

A strategy profile in this setting is described by a tuple $(\mathbf{b}, \mathbf{t})$, where $\mathbf{b}$ is a bid profile and $\mathbf{t}$ is a profile of payments (aka monetary transfers), such that $t_{i,j} \geq 0$ is the monetary payment of player i to player j. At this strategy profile, the auctioneer runs the auction with bids $\mathbf{b}$ and returns the outcome (price and allocation), while the players make the monetary transfers $\mathbf{t}$ to each other.

For each profile of monetary transfers $\mathbf{t}$, let $m_i(\mathbf{t})$ be the net amount of money that player i gets after all the transfers are made:

$$m_i(\mathbf{t}) = \sum_{j=1}^{n} t_{j,i} - \sum_{j=1}^{n} t_{i,j} \, .$$

The utility of player i at profile $(\mathbf{b}, \mathbf{t})$ is

$$u_i(\mathbf{b}, \mathbf{t}) = m_i(\mathbf{t}) + \left(\sum_{j=1}^{x_i(\mathbf{b})} v_{i,j} \right) - p \cdot x_i(\mathbf{b}) \, .$$

Deviations. Since in the case of auctions the actions (e.g. bids) of a group of players can affect the utility of the players outside of the group, it is necessary to model how the players outside S react to the deviation. Such reactions have been studied in the literature on the core with externalities (see, e.g., [Koc07, Koc09]).

We consider neutral reactions, where non-deviators (i.e. players outside S) have a mild reaction to the deviation: they maintain the same bids as before the deviation and the monetary transfer of each player $i \in [n] \setminus S$ to each player $j \in S$ is non-negative. The core where the deviators assume the non-deviators will have neutral reactions to the deviation is known as the neutral core [SMR+13]. Several other variants of the core exist, such as pessimistic core, where each deviator assumes that they will be punished in the worst possible way by the non-deviators. Such variants are also interesting to study, but next we focus on the basic case of neutral reactions.

A group of players will alternatively be called a *coalition*. The set of all players, $[n]$, will also be called sometimes the grand coalition.

A group of players that agree on a deviation are known as a blocking coalition, formally defined next.

Definition 3 (Blocking coalition, with transfers)**.** *Let* $(\mathbf{b}, \mathbf{t})$ *be a tuple of bids and monetary transfers. A group* $S \subseteq [n]$ *of players is a* blocking coalition *if there exists a profile* $(\widetilde{\mathbf{b}}, \widetilde{\mathbf{t}})$*, at which each player* $i \in S$ *weakly improves their utility, the improvement is strict for at least one player in* S*, and*

- $\widetilde{b}_{i,j} = b_{i,j}$ *if* $i \in [n] \setminus S, j \in [K]$.
- $\widetilde{t}_{i,j} = 0$ *if* $i \in [n] \setminus S$ *and* $j \in S$.

In other words, the blocking coalition S needs to agree on their bids and transfers to each other, such that they improve their utility when the players outside S maintain their existing bids but stop payments to players in S. In fact our characterization holds even if the players outside S make any non-negative transfers to the players in S; the case where the transfers are zero is the extreme case. If a coalition S deviates with zero transfers from players outside S, it also deviates for any transfers that are non-negative.

Definition 4 (The core with transfers)**.** *The* core with transfers *consists of profiles* $(\mathbf{b}, \mathbf{t})$ *at which there are no blocking coalitions. Such profiles are* core-stable.

Theorem 6 (Core with transfers; restated). *Consider* K *units and* $n > K$ *hungry players. The core with transfers of the* $(K+1)$*-st auction can be characterized as follows:*

- *Let* $(\mathbf{b}, \mathbf{t})$ *be an arbitrary tuple of bids and transfers that is core stable. Then the allocation* $\mathbf{x}(\mathbf{b})$ *maximizes social welfare, the price is zero (i.e.* $p(\mathbf{b}) = 0$*), and there are no transfers between the players (i.e.* $\mathbf{t} = 0$*).*

Proof. The proof has three steps as follows.

Step I: core transfers are zero. Assume towards a contradiction that $(\mathbf{b}, \mathbf{t})$ is core stable and $\mathbf{t} \neq \mathbf{0}$. Consider the directed weighted graph $G = ([n], E, \mathbf{t})$, where E consists of all the directed edges (i, j) and the weight of each edge is $t_{i,j}$. The net amount of money that each player i gets from transfers is $m_i(\mathbf{t}) = \sum_{j=1}^n t_{j,i} - \sum_{j=1}^n t_{i,j}$.

If there is a cycle $C = (i_1, \ldots, i_k)$ such that the payments along the cycle are strictly positive: $t_{i_1,i_2} > 0, \ldots, t_{i_{k-1},i_k} > 0$, and $t_{i_k,i_1} > 0$, then some cancellations take place. That is, by subtracting $\min\{t_{i_1,i_2}, \ldots, t_{i_{k-1},i_k}, t_{i_k,i_1}\}$ from the weight of each edge $(i, j) \in C$, we obtain a weighted directed graph without cycle C and where each player has the same net amount of money as in the original graph. Iterating the operation of removing cycles, we obtain a directed acyclic graph where the players have the same net amount of money as in the original graph.

Thus we can in fact assume the transfers $\mathbf{t}$ are such that G is acyclic.

Consider a topological ordering $(i_1, \ldots, i_n)$ of the vertices of G. Let j be the minimum index for which vertex i_j has no incoming edges and at least one outgoing edge. Then $m_{i_j}(\mathbf{t}) < 0$. The utility of player i_j can be upper bounded as follows:

$$u_{i_j}(\mathbf{b}, \mathbf{t}) = m_{i_j}(\mathbf{t}) + \left(\sum_{k=1}^{x_{i_j}(\mathbf{b})} v_{i_j,k}\right) - p(\mathbf{b}) \cdot x_{i_j}(\mathbf{b}) < \left(\sum_{k=1}^{x_{i_j}(\mathbf{b})} v_{i_j,k}\right) - p(\mathbf{b}) \cdot x_{i_j}(\mathbf{b})\,.$$

We claim that player i_j has an improving deviation by keeping its bid vector $\mathbf{b}_{i_j}$ and stopping all payments to other players. Since the other players have neutral reactions to the deviations, we obtain an outcome $(\widetilde{\mathbf{b}}, \widetilde{\mathbf{t}})$ such that $\widetilde{\mathbf{b}} = \mathbf{b}$, $\widetilde{t}_{i_j,k} = 0$ for all $k \in [n]$, and $\widetilde{t}_{k,i_j} \leq t_{k,i_j}$ for all $k \neq i_j$. The gain of player i_j from the deviation can be bounded by:

$$\begin{aligned}
u_{i_j}(\widetilde{\mathbf{b}}, \widetilde{\mathbf{t}}) - u_{i_j}(\mathbf{b}, \mathbf{t}) &= u_{i_j}(\mathbf{b}, \widetilde{\mathbf{t}}) - u_{i_j}(\mathbf{b}, \mathbf{t}) && \text{(Since } \widetilde{\mathbf{b}} = \mathbf{b}.\text{)}\\
&= \left[m_{i_j}(\widetilde{\mathbf{t}}) + \left(\sum_{k=1}^{x_{i_j}(\mathbf{b})} v_{i_j,k}\right) - p(\mathbf{b}) \cdot x_{i_j}(\mathbf{b})\right] - \left[m_{i_j}(\mathbf{t}) + \left(\sum_{k=1}^{x_{i_j}(\mathbf{b})} v_{i_j,k}\right) - p(\mathbf{b}) \cdot x_{i_j}(\mathbf{b})\right]\\
&= m_{i_j}(\widetilde{\mathbf{t}}) - m_{i_j}(\mathbf{t})\\
&= -m_{i_j}(\mathbf{t}) && \text{(Since } m_{i_j}(\widetilde{\mathbf{t}}) = 0.\text{)}\\
&> 0\,. && \text{(Since } m_{i_j}(\mathbf{t}) < 0 \text{ by choice of } i_j.\text{)}
\end{aligned}$$

Thus player i_j has a strictly improving deviation, which contradicts the choice of $(\mathbf{b}, \mathbf{t})$ as core-stable with $\mathbf{t} \neq \mathbf{0}$. Thus the assumption must have been false. It follows that the only core stable profiles (if any) have $\mathbf{t} = \mathbf{0}$.

Step II: the social welfare is maximized. Next we show that if $(\mathbf{b}, \mathbf{t})$ is a core-stable outcome, then the allocation induced by $\mathbf{b}$ is welfare maximizing. Assume towards a contradiction this is not the case. By Step I, we have $\mathbf{t} = \mathbf{0}$.

Let $w_1 \geq \ldots \geq w_{n \cdot K}$ be the bids sorted in decreasing order (breaking ties lexicographically) at the truth-telling bid profile $\mathbf{v}$. For each $j \in [n \cdot K]$, let π_j be the player that submitted bid w_j in this ordering.

Let $\widetilde{w}_1 \geq \ldots \geq \widetilde{w}_{n \cdot K}$ be the bids sorted in decreasing order (breaking ties lexicographically) at the bid profile $\mathbf{b}$. Let $\widetilde{\pi}_j$ be the player that submitted bid $\widetilde{w}_j$ in this ordering.

Consider an undirected bipartite graph $G = (L, R, E)$, where L is the left part, R the right part, and E the set of edges. Define

- $L = \{(i, j) \mid i \in [n], j \in [K], \text{ and } x_i(\mathbf{v}) \geq j\}$. For example, if $x_i(\mathbf{v}) = 2$, then L has nodes $(i, 1)$ and $(i, 2)$. If on the other hand $x_i(\mathbf{v}) = 0$, then L has no nodes (i, j), for any j.
- $R = \{(i, j) \mid i \in [n], j \in [K], \text{ and } x_i(\mathbf{b}) \geq j\}$.
- $E = (s_1, s_2)$, for all $s_1 \in L$ and $s_2 \in R$.

Since both allocations $\mathbf{x}(\mathbf{v})$ and $\mathbf{x}(\mathbf{b})$ allocate exactly K units, we have $|L| = |R| = K$. Consider now a graph $G_1 = (L_1, R_1, E_1)$ obtained from G as follows. Set $G = G_1$. Then for each node $(i, j) \in L$: if the node also appears in R, then delete both copies of the node, together with any edges containing them.

Thus in G_1, the left side L_1 consists of nodes (i, j) with $x_i(\mathbf{v}) \geq j$ but $x_i(\mathbf{b}) < j$. For each such node (i, j), let k be the rank of the valuation $v_{i,j}$ in the ordering π. By definition of G_1, we have that R_1 does not have any node of the form (i, s) for any s. To see this, observe that

- nodes (i, s) with $s < j$ were deleted when constructing G_1 from G, and
- nodes (i, s) with $s \geq j$ do not exist even in G (if they did, then (i, j) would exist in both L and R and so would have been deleted when constructing G_1).

Let $d(i, j)$ be the player that displaces player i's bid for the j-th unit at the bid profile $\mathbf{b}$. Formally, $d(i, j)$ is the owner of bid $\widetilde{w}_k$ when considering the bids $\mathbf{b}$ in descending order. Let $\omega(i, j) \in \mathbb{N}$ be such that player $d(i, j)$ obtains an $\omega(i, j)$-th unit in their bundle at $\mathbf{b}$.

Since $x_i(\mathbf{b}) < j$, we have $i \neq d(i, j)$. Since at the truth-telling profile player i gets a j-th unit but player $d(i, j)$ does not get an $\omega(i, j)$-th unit, we have $v_{i,j} \geq v_{d(i,j),\omega(i,j)}$. Moreover, since the allocation $\mathbf{x}(\mathbf{v})$ maximizes welfare but $\mathbf{x}(\mathbf{b})$ does not, the inequality is strict for some $(i, j) \in L_1$.

We claim that $[n]$ is a blocking coalition. To show this, we will argue there is a bid profile $\mathbf{b}^*$ and vector of transfers $\mathbf{t}^*$ such that at $(\mathbf{b}^*, \mathbf{t}^*)$ the utility of each player $i \in [n]$ is weakly improved and the improvement is strict for at least one player. For each $i \in [n], k \in [K]$, let

$$b^*_{i,j} = \begin{cases} v_{i,j} & \text{if } x_i(\mathbf{v}) \geq j \\ 0 & \text{otherwise.} \end{cases}$$

Then $\mathbf{x}(\mathbf{b}^*) = \mathbf{x}(\mathbf{v})$ and $p(\mathbf{b}^*) = 0$. Also define monetary transfers $\mathbf{t}^*$ as follows:

- Initialize $\mathbf{t}^* = 0$. Let $\varepsilon = \min_{(i,j) \in L_1} \left(v_{i,j} - v_{d(i,j),\omega(i,j)}\right)/2$. Then $v_{i,j} \geq v_{d(i,j),\omega(i,j)} + \varepsilon$ for each $(i, j) \in L_1$.
- For each $(i, j) \in L_1$, let $t^*_{i,d(i,j)} := t^*_{i,d(i,j)} + v_{d(i,j),\omega(i,j)} + \varepsilon$.

Thus each player i that got a j-th unit in their bundle at the truth-telling profile did so because their bid for the j-th unit had rank $k \leq K$. Since i does not get the j-th unit at bid profile $\mathbf{b}$, there is a player $d(i,j)$ whose bid for the j-th unit had rank k and who received this way a $\omega(i,j)$-th unit in their bundle.

We argue that all the players weakly improve their utility at $(\mathbf{b}^*, \mathbf{t}^*)$, and the improvement is strict for at least one of them.

- For each pair $(i,j) \in L_1$, under the bid profile $\mathbf{b}^*$ player i receives the j-th unit at a cost of zero and transfers an amount of $v_{d(i,j),\omega(i,j)} + \varepsilon$ to player $d(i,j)$.

  The component of the utility that player i gets from unit j, counting the value, price, and transfer related to unit j, is $v_{i,j} - 0 - (v_{d(i,j),\omega(i,j)} + \varepsilon) \geq 0$, where the inequality holds by choice of ε. This is a weak improvement compared to the utility that player i gets from unit j at profile $(\mathbf{b}, \mathbf{0})$, which is zero. Moreover, the improvement is strict for at least one pair $(i,j) \in L_1$, since the bid profile $\mathbf{b}$ does not induce a welfare maximizing allocation.

- For each pair $(i,j) \in L \setminus L_1$, at the profile $(\mathbf{b}, \mathbf{0})$ player i gets utility $v_{i,j} - p(\mathbf{b})$ from unit j, since it makes no transfers. At profile $\mathbf{b}^*$, player i gets unit j at a price of zero and makes no transfers (towards other players) related to unit j. Thus the component of the utility related to unit j is $v_{i,j} - 0 - 0 = v_{i,j}$. Since $p(\mathbf{b}) \geq 0$, we have $v_{i,j} - p(\mathbf{b}) \leq v_{i,j}$, a weak improvement for player i with respect to unit j.

- For each pair $(i,j) \notin L$: if $(i,j) \notin R$, then player i does not get a j-th unit under either $\mathbf{b}$ or $\mathbf{b}^*$, so its utility from unit j is zero at both profiles. If on the other hand $(i,j) \in R$, since $(i,j) \notin L$, it must be that $(i,j) \in R_1$. At profile $(\mathbf{b}, \mathbf{t})$ the utility of player i from unit j is $v_{i,j} - p(\mathbf{b})$ since it gets the unit and receives no transfers. At profile $(\mathbf{b}^*, \mathbf{t}^*)$ player i does not receive the unit but receives a transfer of $v_{i,j} + \varepsilon$ from the player that gets the unit instead. This is again a weak improvement.

Thus all players weakly improve their utility at $(\mathbf{b}^*, \mathbf{t}^*)$ and the improvement is strict for at least one player, so the profile $(\mathbf{b}, \mathbf{0})$ is not stable. This is a contradiction, thus the assumption that $\mathbf{x}(\mathbf{b})$ is not welfare maximizing must have been false.

Step III: the price is zero. Next we show that if profile $(\mathbf{b}, \mathbf{0})$ is core-stable, then $p(\mathbf{b}) = 0$. By Step II, the bid profile $\mathbf{b}$ induces a welfare maximizing allocation.

Suppose towards a contradiction that $p(\mathbf{b}) > 0$. Then we show there exists a blocking coalition. Let $w_1 \geq \ldots \geq w_{n \cdot K}$ be the bids sorted in decreasing order (breaking ties lexicographically) at bid profile $\mathbf{b}$. For each $i \in [n \cdot K]$, let π_i be the owner of bid w_i.

We match the players as follows. Create a bipartite graph with left part $L = (\pi_1, \ldots, \pi_K)$ (allowing repetitions) and right part $R = (\pi_{K+1}, \ldots, \pi_{2K})$ (allowing repetitions). For each $i \in [K]$, create edge (π_i, π_{K+i}). Consider the profile $(\mathbf{b}^*, \mathbf{t}^*)$, where

$$b^*_{i,j} = \begin{cases} v_{i,j} & \text{if } x_i \geq j \\ 0 & \text{otherwise.} \end{cases}$$

Define the transfers as follows:

- Initialize $\mathbf{t}^* = \mathbf{0}$. Set $\varepsilon = p/(2K)$. For each $i \in [K]$, let player π_i pay an additional amount of ε to player π_{K+i}: $t^*_{\pi_i, \pi_{K+i}} = t^*_{\pi_i, \pi_{K+i}} + \varepsilon$.

Then each player i with $x_i(\mathbf{b}) > 0$ gets the same allocation at $\mathbf{b}^*$ as at $\mathbf{b}$, but pays a price of zero for the units and makes a transfer of at most $p/2$ to other players, resulting in improved utility compared to the utility at profile $(\mathbf{b}, \mathbf{0})$.

On the other hand, each player i with $x_i(\mathbf{b}) = 0$ on the other hand gets the same allocation at $\mathbf{b}^*$ as at $\mathbf{b}$ (i.e. no units), but receives a non-negative amount of money from other players, and the net amount of money received is strictly positive for all the players $\pi_{K+1}, \ldots, \pi_{2K}$.

Thus there is an improving deviation, which contradicts the choice of $(\mathbf{b}, \mathbf{t})$ as core-stable. Thus the assumption must have been false, and $p(\mathbf{b}) = 0$. □

D.2 Core without transfers

A strategy profile in this setting is described by a bid profile $\mathbf{b}$, where $\mathbf{b}_i = (b_{i,1}, \ldots, b_{i,K})$ is the bid vector of player i. In the setting without transfers, given a profile $\mathbf{b}$ of bids, a blocking coalition S needs to agree on simultaneously changing their bids, such that when the players outside S still bid according to $\mathbf{b}$, the players in S weakly improve their utility and the improvement is strict for at least one player in S.

The core stable profiles are those that have no such blocking coalitions. Formally, we have:

Definition 5 (Blocking coalition, without transfers)**.** *Let* $\mathbf{b}$ *be a bid profile. A group* $S \subseteq [n]$ *of players is a* blocking coalition *if there exists a bid profile* $\widetilde{\mathbf{b}}$*, at which each player* $i \in S$ *weakly improves their utility, the improvement is strict for at least one player in* S*, and* $\widetilde{b}_{i,j} = b_{i,j}$ *for all* $i \in [n] \setminus S, j \in [K]$.

Definition 6 (The core with transfers)**.** *The* core with transfers *consists of bid profiles* $\mathbf{b}$ *at which there are no blocking coalitions. Such profiles are* core-stable*.*

The non-transferable utility core can be characterized as follows.

Theorem 5 (Core without transfers; restated). *Consider* K *units and* $n > K$ *hungry players. The core without transfers of the* $(K+1)$*-st auction can be characterized as follows:*

- *every bid profile* $\mathbf{b}$ *that is core stable has price zero (i.e.* $p(\mathbf{b}) = 0$*);*
- *each allocation* $\mathbf{z}$ *where all the units are allocated can be supported in a core stable bid profile* $\mathbf{b}$ *with price zero (i.e.* $\mathbf{x}(\mathbf{b}) = \mathbf{z}$ *and* $p(\mathbf{b}) = 0$*).*

Proof. The proof is in two parts.

Part I: the price in the core is zero. Consider a bid profile $\mathbf{b}$ that is core-stable. Suppose towards a contradiction that $p(\mathbf{b}) > 0$.

Let $M = \sum_{\ell_1=1}^{n} \sum_{\ell_2=1}^{K} v_{\ell_1,\ell_2}$. Define a bid profile $\widetilde{\mathbf{b}}$ such that for all $i \in [n], j \in [K]$:

$$\widetilde{b}_{i,j} = \begin{cases} M & \text{if } j \leq x_i(\mathbf{b}) \\ \varepsilon & \text{otherwise.} \end{cases} \tag{34}$$

At $\widetilde{\mathbf{b}}$, the players only submit bids equal to $M > 0$ for the units they are supposed to get at allocation $\mathbf{x}(\mathbf{b})$, and moreover, there are exactly K strictly positive bids. Thus $\mathbf{x}(\widetilde{\mathbf{b}}) = \mathbf{x}(\mathbf{b})$. Moreover, $p(\mathbf{b}) = 0$ since the $(K+1)$-st highest bid is 0.

Then the grand coalition $C = [n]$ is blocking with the profile $\widetilde{\mathbf{b}}$, in contradiction with $\mathbf{b}$ being core stable. Thus the assumption must have been false, so $p(\mathbf{b}) = 0$.

Part II: every allocation can be implemented at a core-stable bid profile. Let $\mathbf{z}$ be an arbitrary allocation at which all the units are allocated.

Define $\mathbf{b}$ such that for all $i \in [n], j \in [K]$, we have $b_{i,j} = M$ if $j \leq z_i$ and $b_{i,j} = 0$ otherwise. Then $\mathbf{x}(\mathbf{b}) = \mathbf{z}$ and $p(\mathbf{b}) = 0$. Let $W = \{i \in [n] \mid z_i > 0\}$ be the set of "winners" at $\mathbf{b}$.

Assume towards a contradiction that $\mathbf{b}$ is not stable. Then there is a blocking coalition $C = \{i_1, \ldots, i_k\} \subseteq [n]$ with alternative bid profile $\mathbf{d} = (\mathbf{d}_{i_1}, \ldots, \mathbf{d}_{i_k})$. Denote $\widetilde{\mathbf{b}} = (\mathbf{d}, \mathbf{b}_{-C})$ the profile where each player $i \in C$ bids $\widetilde{\mathbf{b}}_i = \mathbf{d}_i$ and each player $i \notin C$ bids $\widetilde{\mathbf{b}}_i = \mathbf{b}_i$. We must have $u_i(\widetilde{\mathbf{b}}) \geq u_i(\mathbf{b})$ for all $i \in C$, with strict inequality for some player $i \in C$.

If $C \cap W = \emptyset$, then the only way for at least one of the players in C to change their allocation is to make some of their bids at least H. But this increases the price from zero to at least H, which yields negative utility for everyone. Thus $C \cap W \neq \emptyset$. We consider two cases:

1. Case $p(\widetilde{\mathbf{b}}) > 0$. Each player $i \in C \cap W$ requires strictly more units at $\widetilde{\mathbf{b}}$ than at $\mathbf{b}$, to compensate for the higher price at $\widetilde{\mathbf{b}}$. Thus $\mathbf{x}_i(\widetilde{\mathbf{b}}) > z_i$ for all $i \in C \cap W$ (†). If $x_i(\widetilde{\mathbf{b}}_i) < z_i$ for some player $i \in W \setminus C$, there would have to exist at least $K+1$ bids with value at least H, and so $p(\widetilde{\mathbf{b}}) \geq H$, which would give negative utility to all the players including the deviators. Thus $x_i(\widetilde{\mathbf{b}}_i) \geq z_i$ for all $i \in W \setminus C$ (‡).

   Combining (†) and (‡) gives a contradiction:

$$\sum_{i \in C \cup (W \setminus C)} x_i(\widetilde{\mathbf{b}}) > \sum_{i \in C \cup (W \setminus C)} z_i = K\,.$$

   Thus $p(\widetilde{\mathbf{b}}) > 0$ cannot hold.

2. Case $p(\mathbf{b}) = 0$. Then each player $i \in C \cap W$ requires $u_i(\widetilde{\mathbf{b}}) \geq z_i$. Since $p(\widetilde{\mathbf{b}}) = 0$, the top K bids are strictly positive and the remaining bids are zero. Then each player $i \in C \cap W$ submits exactly z_i strictly positive bids. It follows that $\mathbf{x}_i(\widetilde{\mathbf{b}}) = \mathbf{x}_i(\mathbf{b})$ and $p(\widetilde{\mathbf{b}}) = p(\mathbf{b})$ for each $i \in [n]$, which means no player in C strictly improves. Thus C cannot be blocking.

In both cases 1 and 2 we obtained a contradiction, so $\mathbf{b}$ is core-stable, $p(\mathbf{b}) = 0$, and $\mathbf{x}(\mathbf{b}) = \mathbf{z}$ as required. □

E Theorems from prior work

In this section we include the theorem from [CBL06] that we use. In the problem of prediction under expert advice, there are N experts in total, and at each round $t \in [T]$:

- learner chooses a probability distribution p_t over $[N]$;
- nature reveals the losses $\{\ell_{t,i}\}_{i \in [N]}$ of all experts at time t, where $\ell_{t,i} \in [0, L]$.

For a given sequence of probability distributions $(p_1, \cdots, p_T)$, the learner's regret is defined to be

$$\mathrm{Reg}(T, (p_1, \cdots, p_T)) = \sum_{t=1}^{T} \sum_{i=1}^{N} p_t(i)\ell_{t,i} - \min_{i^\star \in [N]} \sum_{t=1}^{T} \ell_{t,i^\star}.$$

For this problem, the exponentially weighted average forecaster, also known as the Hedge algorithm, is defined as follows. The algorithm initializes $p_1 = \sigma$, an arbitrary prior distribution over the experts $[N]$ such that each expert i is selected with probability $\sigma_i > 0$. For $t \geq 2$, the forecaster updates

$$p_t(i) = \frac{p_{t-1}(i)\exp(-\eta \ell_{t-1,i})}{\sum_{j=1}^{N} p_{t-1}(j)\exp(-\eta \ell_{t-1,j})}, \quad \forall i \in [N],$$

where $\eta > 0$ is a learning rate.

Next we include the statement of Theorem 2.2 of [CBL06] in our notation.

Theorem 7 (Theorem 2.2 of [CBL06]). *Consider the exponentially weighted average forecaster with N experts, learning rate $\eta > 0$, time horizon T, and rewards in $[0, 1]$. Suppose the initial distribution σ on the experts is uniform, that is, $\sigma = (1/N, \ldots, 1/N)$. The regret of the forecaster is*

$$\mathrm{Reg}(T, (p_1, \cdots, p_T)) \leq \frac{\log N}{\eta} + \frac{T\eta}{8}\,.$$

The following corollary is a well known variant of the above theorem, to allow rewards in an interval $[0, L]$ and an arbitrary initial distribution σ over the experts.

Corollary 1. *Consider the exponentially weighted average forecaster with N experts, learning rate $\eta > 0$, time horizon T, and rewards in $[0, L]$. Suppose the initial distribution on the experts is σ. The regret of the forecaster is*

$$\mathrm{Reg}(T, (p_1, \cdots, p_T)) \leq \frac{1}{\eta} \max_{i \in [N]} \log\left(\frac{1}{\sigma_i}\right) + \frac{TL^2\eta}{8}\,.$$

The proof of Corollary 1 is identical to that of Theorem 7, except for the following two differences:

- Instead of $W_t = \sum_{i=1}^{N} \exp(-\eta \sum_{s \le t} \ell_{s,i})$, we define $W_t = \sum_{i=1}^{N} \sigma_i \exp(-\eta \sum_{s \le t} \ell_{s,i})$. In this way,

$$\log \frac{W_T}{W_0} = \log W_T \ge -\eta \min_{i^\star \in [N]} \sum_{t=1}^{T} \ell_{t,i^\star} - \max_{i \in [N]} \log \frac{1}{\sigma_i}.$$

- When applying [CBL06, Lemma 2.2], we use the interval for the rewards as $[0, L]$ instead of $[0, 1]$.