A General Statistical Framework for Designing
Strategy-proof Assignment Mechanisms

Appendix

A Generalization Bounds

For a rule class $\mathcal{F}_i$ with finite Natarajan dimension of at most $D$, the following result relates the empirical and population 0-1 errors of any rule in $\mathcal{F}_i$: w.p. at least $1 - \delta$ (over draw of $S$), for all $f_i \in \mathcal{F}_i$,

$$\left| E_{\theta \sim \mathcal{D}}[1(g_i(\theta) \neq f_i(\theta))] - \frac{1}{N} \sum_{k=1}^{N} 1(y^k \neq f_i(\theta^k)) \right| \leq O\left( \sqrt{\frac{D \ln(m) + \ln(1/\delta)}{N}} \right).$$  

(6)

The proof involves a reduction to binary classification, and an application of a VC dimension based generalization bound (see for example proof of Theorem 4 in [21]; also see Eq. (6) in [21]). It is straightforward to extend the above result to a similar bound on the Hamming error metric of an outcome rule $f \in \mathcal{F}$.

**Lemma 10.** With probability at least $1 - \delta$ (over draw of $S \sim \mathcal{D}^N$), for all $f \in \mathcal{F}$,

$$\left| E_{\theta \sim \mathcal{D}}[\ell(g(\theta), f(\theta))] - \frac{1}{N} \sum_{k=1}^{N} \ell(y^k, f(\theta^k)) \right| \leq O\left( \frac{D \ln(m) + \ln(n/\delta)}{N} \right).$$

Proof. We would like to bound:

$$\sup_{f \in \mathcal{F}} \left| E_{\theta \sim \mathcal{D}}[\ell(g(\theta), f(\theta))] - \frac{1}{N} \sum_{k=1}^{N} \ell(y^k, f(\theta^k)) \right| \leq \frac{1}{n} \sum_{i=1}^{n} \sup_{f \in \mathcal{F}_i} \left| E_{1}[1(g_i(\theta) \neq f_i(\theta))] - \frac{1}{N} \sum_{k=1}^{N} 1(y^k \neq f_i(\theta^k)) \right|.$$  

Applying (6) to the above expression, along with a union bound over all $i$, gives us the desired result.

B Proofs

B.1 Complete Proof of Lemma 5

Proof. For any $f : \Theta \rightarrow \Omega$, define a binary function $G_f : \Theta \rightarrow \{0, 1\}$ as $G_f(\theta) = 1(f_1(\theta) \neq \ldots \neq f_n(\theta))$. Clearly, $f$ is feasible on $S$ iff $G_f$ evaluates to 1 on all type profiles in $S$, and feasible on all type profiles iff $G_f$ evaluates to 1 on all type profiles.

Treating $G_f$ as a binary classifier, the desired result can be derived using standard VC dimension based learnability results for binary classification [22], with the loss function being the 0-1 loss against a labeling of 1 on all profiles. Let $\mathcal{G} = \{G_f : \Theta \rightarrow \{0, 1\} : f \in \mathcal{F}\}$ be the set of all such binary classifiers. Also, $\epsilon_{\text{infeasible}} = E_{\theta \sim \mathcal{D}}[1(G_f(\theta) \neq 1)]$. We then wish to bound the expected 0-1 error of a classifier $G_f$ from $\mathcal{G}$ that outputs 1 on all type profiles in $S$.

We first bound the VC dimension of $\mathcal{G}$. Since each $\mathcal{F}_i$ has a Natarajan dimension of at most $D$, we have from Lemma 11 in [21] that the maximum number of ways a set of $N$ profiles can be labeled by $\mathcal{F}_i$ with labels $[m]$ is at most $N^D m^{2D}$. Since each $G_f$ is a function solely of the outputs of $f_1, \ldots, f_n$, the number of ways a set of $N$ profiles can be labeled by $\mathcal{G}$ with labels $[0, 1]$ is at most $(N^D m^{2D})^n$.

The VC dimension of $\mathcal{G}$ is then given by the maximum value of $N$ for which $2^N \leq (Nm^2)^n$. We thus have that the VC dimension is at most $O(nD \ln(nmD))$.

Since $\mathcal{F}_\emptyset \neq \emptyset$, there always exists a function $G_f$ consistent with a labeling of 1 on all profiles. A standard VC dimension based argument then gives us the following guarantee for the outcome rule $\hat{f}$ that is feasible on sample $S$: w.p. at least $1 - \delta$ (over draw of $S$),

$$\epsilon_{\text{infeasible}} = E_{\theta \sim \mathcal{D}}[1(G_f(\theta) \neq 1)] \leq O\left( \frac{nD \ln(nmD) \ln(n) + \ln(1/\delta)}{N} \right),$$

which implies the statement of the lemma.
B.2 Proof of Theorem 7

Proof. Let \( w_i = \left\{ 1, 1, \ldots, 1, -1, -1, \ldots, -1 \right\} \). We first show that the corresponding payments are non-negative.

\[
\hat{t}_i^w(\theta_{-i}, o) = w_i^T \hat{\Psi}_i(\theta_{-i}, o) = \sum_{j \neq i} \sum_{o' = 1}^m v_j(\theta_j, o') - \sum_{j \neq i} v_j(\theta_j, y_j^i) = \sum_{j \neq i} \sum_{o' \neq y_j^i} v_j(\theta_j, o') \geq 0.
\]

We next show that the outcome rule \( f^w \) is feasible, and in particular, outputs a welfare-maximizing assignment. Note that \( f^w_i(\theta) \) can output any one of the following items:

\[
\mathcal{I}_i = \arg\max_{o \in [m]} \left\{ v_i(\theta_i, o) - w_i^T \hat{\Psi}_i(\theta_{-i}, o) \right\} = \arg\max_{o \in [m]} \left\{ v_i(\theta_i, o) + \sum_{j \neq i} v_j(\theta_j, y_j^i) - \sum_{j \neq i} \sum_{o' = 1}^m v_j(\theta_j, o') \right\}
\]

where \( T_{-i} \) is a term independent of agent \( i \)'s valuations and the item \( o \) over which the argmax is taken. If the above max is achieved by more than one item, then the individual functions \( f^w_i \) may not pick distinct items. However, in each of the following feasible assignments, agent \( i \) is assigned an optimal item from \( \mathcal{I}_i \): \( \arg\max_{y \in [m]} \left\{ \sum_{i=1}^m v_i(\theta_i, y_i) \right\} \). Thus \( \hat{f} \) is feasible as long as it uses a tie-breaking scheme that picks an assignment from this set. Such a tie-breaking scheme will not violate the agent-independence condition, as the agents continue to receive an optimal item based on their agent-independent prices.

B.3 Proof for Theorem 8

Proof. For ease of presentation, we omit the subscript \( i \) whenever clear from context. Let \( A \subseteq \Theta \) be a set of \( N \) profiles \( N \)-shattered by \( \hat{F}^\Psi \). Then there exists labelings \( L_1, L_2 : A \rightarrow [m] \) that disagree on all profiles in \( A \) such that for all \( B \subseteq A \), there is a \( w \) with \( f^w(\theta) = L_1(\theta), \forall \theta \in B \) and \( f^w(\theta) = L_2(\theta), \forall \theta \in A \setminus B \).

To bound the Natarajan dimension of \( \hat{F}^\Psi \), define \( \xi^w : \Theta \rightarrow \{0, 1\} \) that for any \( \theta \in \Theta \) outputs 1 if \( f^w(\theta) = L_1(\theta) \) and 0 otherwise. Then for all subsets \( B \) of a \( N \)-shattered set \( A \), there is a \( w \) with \( \xi^w(\theta) = 1, \forall \theta \in B \) and \( \xi^w(\theta) = 0, \forall \theta \in A \setminus B \). This implies that if a set is \( N \)-shattered by \( \hat{F}^\Psi \), it is (binary) shattered by the class \( \{ \xi^w : w \in \mathbb{R}^d \} = \Xi \) (say). Thus the size of the largest set \( N \)-shattered by \( \hat{F}^\Psi \) is no larger than the size of the largest set (binary) shattered by \( \Xi \). The Natarajan dimension of \( \hat{F}^\Psi \) is therefore upper bounded by the VC dimension of \( \Xi \).

What remains is to bound the VC dimension of \( \Xi \). Note that \( \xi^w(\theta) = 1 \) only when \( w^T \Psi_i(\theta_{-i}, L_1(\theta)) \leq 1 \) and \( L_1(\theta) \geq l_o \, \forall o \in \{o' \in [m] : w^T \Psi_i(\theta_{-i}, o' \leq 1) \} \). Also note that when \( \theta \in \Theta \) is fixed, the output of \( \xi^w(\theta) \) for different \( w \in \mathbb{R}^d \) is solely determined by the value of the binary vector \( [1(w^T \Psi_i(\theta_{-i}, o \leq 1))_{o=1}^m \in \{0, 1\}^m \). Thus the number of ways a fixed set \( A \subseteq \Theta \) can be labeled by \( \Xi \) cannot be larger than the number of ways \( A \) can be labeled with the binary vectors \( [1(w^T \Psi_i(\theta_{-i}, o \leq 1))_{o=1}^m \in \{0, 1\}^m \) for different \( w \in \mathbb{R}^d \).

Each entry of the above binary vector can be seen as a linear separator. Given that the VC dimension of linear separators in \( \mathbb{R}^d \) (with a constant bias term) is \( d \), by Sauer’s lemma, the number of ways a set of \( N \) profiles can be labeled by a single entry \( 1(w^T \Psi_i(\theta_{-i}, o \leq 1) \) is at most \( (Ne)^d \). The total number of ways the set can be labeled with binary vectors of the above form is at most \( (Ne)^{md} \). The VC dimension of \( \Xi \) is then the largest \( N \) for which \( 2^N \leq (Ne)^{md} \). We thus get that the VC dimension of \( \Xi \) is at most \( O((md) \ln(md)) \), as desired.
B.4 Proof of Theorem 9

Proof. Fix a priority \( \pi : [n] \rightarrow [n] \) over the agents, where \( \pi(i) \) denotes the priority to agent \( i \) (with 1 indicating the lowest priority, and \( n \) indicating the highest). Define \( w_i \in \mathbb{R}^{n \times m} \) as follows: for \( j \in [n], k \in [m] \),

\[
 w_i[j,k] = \begin{cases} 
 2 & \pi(j) > \pi(i), \ k \geq m - n + \pi(j) \\
 0 & \text{otherwise}.
\end{cases}
\]

We show that the resulting outcome rule is a feasible serial dictator style mechanism where the agents are served according to the priority ordering \( \pi \). We show this for the case when \( m = n \). The proof easily extends to the case where this is not true.

Recall that the entry \((j,k)\) for \( j \neq i \) in the feature map \( \hat{\Psi}_i(\theta_{-i}, o) \) is 1 when agent \( j \) assigns a rank of \( k \) to item \( o \), i.e. \( \text{rank}_j(\theta_j, o) = k \). One can then observe that virtual price function, \( t^\text{vir.}_i(\theta_{-i}, o) = w^\top_i \hat{\Psi}_i(\theta_{-i}, o) \geq 2 \) whenever an agent with a higher priority assigns item \( o \) a rank greater or equal to its priority level, i.e. whenever \( \text{rank}_j(\theta_j, o) \geq \pi(j) \) for some \( j \) with \( \pi(j) > \pi(i) \). The item \( o \) is then not affordable to agent \( i \), as the virtual price exceeds a budget of 1.

The resulting outcome rule is similar to a serial dictatorship mechanism and serves the agents according to the priorities \( \pi \): agent \( \pi^{-1}(1) \) affords all items; agent \( \pi^{-1}(2) \) affords all but the item most-preferred by agent \( \pi^{-1}(1) \); agent \( \pi^{-1}(3) \) affords all items except the most-preferred item by agent \( \pi^{-1}(1) \), and the first- and second-most preferred items by agent \( \pi^{-1}(2) \); and so on. Thus the most-preferred affordable item for a given agent is always unaffordable for lower priority agents. Since each agent receives its most-preferred affordable item (and is unassigned if it cannot afford any), there are no conflicts in assignments. \( \square \)