A Impossibility results – missing proofs

Hereafter, we report all the proofs omitted from Section 4. We start by proving the following corollary of Proposition 4.1 on simple fractional HGs.

Corollary 4.2. *Given a parameter λ there exists a bounded distribution with parameter λ such that in simple fractional HGs no ε-core exists for $\varepsilon < \frac{\lambda}{2^{40}(\lambda-1)+2^n}$.*

Proof. Consider the instance $\mathcal{I}'$ in the proof of Proposition 4.1. We have shown that there exists a collection of coalitions $\mathcal{F} = 2^N \cup \{\{41, \dots, n\}\} \setminus \{\emptyset\}$ of size 2^{40} that always contains a core-blocking coalition for any possible partition π.

Let $\mathcal{D}$ be a bounded distribution of parameter λ such that, for some $p > 0$, $\Pr_{C \sim \mathcal{D}}[\, C \,] = p$ if $C \in \mathcal{F}$ and p/λ, otherwise. For such a distribution the probability of sampling a specific coalition in $\mathcal{F}$ is therefore given by $p = \frac{\lambda}{2^{40}(\lambda-1)+2^n}$. In conclusion,

$$\Pr_{C\sim\mathcal{D}}[\, C \text{ core-blocks } \pi \,] \geq \Pr_{C\sim\mathcal{D}}[\, C \text{ core-blocks } \pi \,\wedge\, C \in \mathcal{F} \,] \geq \frac{\lambda}{2^{40}(\lambda-1)+2^n}$$

and the thesis follows. □

In the following, we will prove the results regarding anonymous HGs. The approach is similar to the one used for fractional HGs, and it is based on a single-peaked instance with an empty core shown in [6].

Proposition 4.3. *There exists a distribution $\mathcal{D}$ and an anonymous (single-peaked) HG instance such that for every $\varepsilon \leq 1/2^7$ there is no ε-fractional core w.r.t. $\mathcal{D}$.*

Proof. Anonymous hedonic games have been shown to have an empty core even under single-peaked preferences [6]. Such an example considers an instance $\mathcal{I}$ with seven agents whose preferences are single-peaked with respect to the natural ordering $1 \dots 7$. For our purposes, it is not important how the instance $\mathcal{I}$ looks like but only the aforementioned properties it has.

Starting from $\mathcal{I}$, we can define an instance $\mathcal{I}'$, with $N' = [n]$ being the set of agents, which still has an empty core. Let us denote by $N = \{1, \dots, 7\}$ the set of the seven agents in $\mathcal{I}$. Their preferences for coalitions of size 1 to 7 remain the same while any other coalition size is strictly less preferred. In particular, we can extend their preferences in such a way that the instance is still singled-peaked with respect to the natural ordering $1 \dots n$; namely, if $h \in [7]$ is the less preferred size according to agent $i \in N$, then, we set $h \succ_i 8 \succ_i 9 \cdots \succ_i n$. For the remaining agents, we assume preferences to be decreasing for decreasing coalition size, i.e., $n \succ_i n-1 \succ_i \cdots \succ_i 1$ for $i \in N' \setminus N$, which are clearly single-peaked in the natural ordering.

We next show this instance has an empty core and that for any coalition structure π there exists a core blocking coalition in $2^N \cup \{\{8, \dots, n\}\} \setminus \{\emptyset\}$.

Given a partition π, any agent $i \in N$ if not in a coalition of size at most 7 would deviate and form the singleton coalition. Assume now that in π agents in N are in coalitions of size at most 7, as a consequence, agents in $N' \setminus N = \{8, \dots, n\}$ are in a coalition of size at most $n-7$; if such agents are in a coalition of size strictly smaller than $n-7$ then the coalition $\{8, \dots, n\}$ is a core blocking coalition. Let us finally assume that agents $8, \dots, n$ form a coalition together of size $n-7$ in π. No matter how the agents of N are partitioned in π we know that any partition of N can always be core blocked by a coalition in 2^N since $\mathcal{I}$ has an empty core. Therefore, the new instance $\mathcal{I}'$ has an empty core and it is always possible to find a blocking coalition in $2^N \cup \{\{8, \dots, n\}\} \setminus \{\emptyset\}$.

We are now ready to show our claim. Let $\mathcal{D}$ be the uniform distribution over $2^N \cup \{\{8, \dots, n\}\} \setminus \emptyset$. For any coalition structure π it holds that $\Pr_{C\sim\mathcal{D}}[C \text{ blocking for } \pi] > 1/2^7$ completing the proof. □

As for fractional HGs, the following corollary focuses specifically on λ-bounded distributions.

Corollary 4.4. *Given a parameter λ, there exists a λ-bounded distribution such that for $\varepsilon < \frac{\lambda}{2^7(\lambda-1)+2^n}$ no ε-fractional core-stable solution exists in anonymous single-peaked HGs.*

Proof. Consider the instance $\mathcal{I}'$ in the proof of Proposition 4.3. We have shown that there exists a collection of coalitions $\mathcal{F} = 2^N \cup \{\{8, \ldots, n\}\} \setminus \{\emptyset\}$ of size 2^7 that always contains a core-blocking coalition for any possible partition π.

Let $\mathcal{D}$ be a bounded distribution of parameter λ such that, for some $p > 0$, $\Pr_{C\sim\mathcal{D}}[\,C\,] = p$ if $C \in \mathcal{F}$ and p/λ, otherwise. For such a distribution the probability of sampling a specific coalition in $\mathcal{F}$ is therefore given by $p = \frac{\lambda}{2^7(\lambda-1)+2^n}$. In conclusion,

$$\Pr_{C\sim\mathcal{D}}[\,C \text{ core-blocks } \pi\,] \geq \Pr_{C\sim\mathcal{D}}[\,C \text{ core-blocks } \pi \wedge C \in \mathcal{F}\,] \geq \frac{\lambda}{2^7(\lambda-1)+2^n}$$

and the thesis follows. □

B Simple fractional Hedonic Games – missing proofs

In the following, we will report the proofs of the technical lemmas omitted from Section 5.

The first result concerns the sample complexity of exactly learning agents' preferences in simple FHGs.

Lemma 5.2. *By sampling $m \geq 16 \log \frac{n}{\delta} + 4n$ sets from $U(2^N)$ it is possible to learn exactly the valuation functions $v_1, \ldots, v_n$ with confidence $1 - \delta$.*

Proof. Let us first focus on a fixed agent i. Every sampled set $S \in \mathcal{S}$ that contains i corresponds to a linear equation

$$\alpha_{S,1} v_i(\{1\}) + \cdots + \alpha_{S,i-1} v_i(\{i-1\}) + \alpha_{S,i+1} v_i(\{i+1\}) + \cdots + \alpha_{S,n} v_i(\{n\}) = v_i(S)/|S|,$$

where the $v_i(\{j\})$s are the unknowns and

$$\alpha_{S,j} = \begin{cases} 1 & \text{if } j \in S; \\ 0 & \text{otherwise.} \end{cases}$$

Now, let us denote by $A^i = [\alpha_{S_q,j}]_{q\in[m'],j\in N\setminus\{i\}}$ the $m' \times (n-1)$ binary matrix corresponding to a sample $\mathcal{S} = \langle (S_1, v(S_1)), \ldots, (S_{m'}, v(S_{m'})) \rangle$ in which every set $S_q, q \in [m']$, contains agent i. Learning the exact valuation function v_i can be accomplished when the sets $S_1, \ldots, S_{m'}$ are such that the matrix A^i has full rank, i.e., rank $n-1$.

Let us denote by $\alpha_1, \ldots, \alpha_{n-1}$ the columns of A^i. Having in mind that each entry in A^i is equal to 0 and 1 with the same probability of $1/2$, we can inductively compute the probability that a set of columns $\{\alpha_1, \ldots, \alpha_{k+1}\}$ is linearly independent as a function of m'. In particular, if we denote by p_k the probability that the set $\{\alpha_1, \ldots, \alpha_k\}$ is linearly independent, then it is easy to see that $p_{k+1} = p_k \cdot (1 - 2^{-(m'-k)})$ and $p_1 = 1 - 2^{-m'}$. Therefore, $p_k = \Pi_{i=1}^k \left(1 - 2^{-(m'-i+1)}\right)$, so that

$$p_{n-1} = \Pi_{i=1}^{n-1}\left(1 - \frac{1}{2^{m'-i+1}}\right) \geq \left(1 - \frac{1}{2^{m'-n+2}}\right)^{n-1} \geq 1 - \frac{n-1}{2^{m'-n+2}}\ .$$

It follows that $p_{n-1} \geq 1 - \delta/(2n)$ when

$$m' = 3\log\frac{n}{\delta} + n - 1 \geq 2\log_2\frac{n}{\delta} + n - 1 \geq \log_2\frac{2n(n-1)}{\delta} + n - 2\ .$$

A sampled set S contains agent i with probability $1/2$. Let E_i be the number of sets containing i in a sample of size m. Then, by the Chernoff's inequality, taking $\beta = 1 - 2m'/m$,

$$\Pr[\,E_i < m'\,] = \Pr[\,E_i < (1-\beta)m/2\,] \leq e^{-\frac{\left(1-\frac{2m'}{m}\right)^2 \frac{m}{2}}{2}} = e^{-\frac{(m-2m')^2}{4m}}\ ,$$

which is less then $\delta/(2n)$ if and only if $\frac{(m-2m')^2}{4m} \geq \log\frac{2n}{\delta}$. Since $\frac{(m-2m')^2}{4m} = \frac{m}{4} - m' + \frac{m'^2}{m}$, it suffices to require that $\frac{m}{4} - m' \geq \log\frac{2n}{\delta}$, i.e., $m \geq 4\log\frac{2n}{\delta} + 4m'$. This holds letting $m = 16\log\frac{n}{\delta} + 4n \geq 4\log\frac{2n}{\delta} + 12\log\frac{n}{\delta} + 4n - 4 = 4\log\frac{2n}{\delta} + 4m'$.

Summarizing, the probability of learning exactly the valuation function v_i is at least the probability that with m examples $E_i \geq m'$ times the probability that A^i has rank $n-1$ given m' examples containing i, that is at least $(1-\delta/(2n)) \cdot (1-\delta/(2n)) \geq 1-\delta/n$.

By a direct union bound argument, the probability of not learning exactly one of the n valuation functions $v_1, \dots, v_n$ is at most δ, hence the claim. $\square$

Hereafter, we will prove the main technical lemma of the section:

Lemma 5.5. *Let $\phi \geq \frac{n^{1/3}}{62}$ and let π be the partition returned by Algorithm 1 in this case. Then the following statements hold:*

(i) H is never empty when executing line 11 of the while loop of Algorithm 1;

(ii) each agent $i \in \mathrm{Gr}$ is green.

The proof of Lemma 5.5 is quite involved and we will divide it in several parts. Throughout the proof, N_i will denote the neighborhood of agent i in the graph induced by the game. We start by proving the first statement i.e. that there is always at least one agent in H in each iteration of the while loop (line 10).

Lemma B.1. *Let $\phi \geq \frac{n^{1/3}}{62}$. Then, H is never empty when executing line 11 of the while loop of Algorithm 1.*

Proof. Consider an iteration t of the while loop, and assume that in line 11 a given $i \in H$ is selected. Then, if in line 13 it results $F_i \subseteq N \setminus H$, in line 16 only agent i is removed from H.

We now show that such a property is guaranteed whenever $d_i \geq \frac{n^{1/3}}{15}$.

If $\frac{n^{1/3}}{15} \leq d_i \leq \frac{n}{3}$, then $\lceil \frac{2d_i}{n-d_i} \rceil = 1$. We prove that agent i, in line 13, selects in F_i exactly one singleton neighbor in $N \setminus H$. In fact, the number of agents in $N \setminus H$ that have been already grouped in some coalitions during the previous iterations is less than $\frac{n^{1/3}}{124}$, as by the ordering of H the property $|F_j| = 1$ holds for all agents j selected in line 11 during the previous iterations. As a consequence, since $H = |\frac{n^{1/3}}{62}|$, the number of singleton neighbors of i left in $N \setminus H$ is at least $d_i - \left(\frac{n^{1/3}}{62} + \frac{n^{1/3}}{124}\right) \geq \frac{n^{1/3}}{15} - \left(\frac{n^{1/3}}{62} + \frac{n^{1/3}}{124}\right) \geq 1$.

If $d_i > \frac{n}{3}$, again F_i will contain exactly $\lceil \frac{2d_i}{n-d_i} \rceil$ singleton neighbors in $N \setminus H$. In fact, denoted as Gr^t the set of agents selected in line 11 in the previous iterations, the number of available singleton neighbors of i outside H at iteration t is more than

$$\begin{aligned} d_i - \sum_{j \in Gr^t} \lceil \frac{2d_j}{n-d_j} \rceil - \frac{n^{1/3}}{62} &\geq d_i - \frac{n^{1/3}}{124} \lceil \frac{2d_i}{n-d_i} \rceil - \frac{n^{1/3}}{62} \\ &\geq d_i - \frac{n^{2/3}}{62 \cdot 31} + \frac{n^{1/3}}{124} - \frac{n^{1/3}}{62} > d_i - \frac{n^{2/3}}{62 \cdot 31} - \frac{n^{1/3}}{62} \\ &\geq \frac{n}{3} - \frac{n^{2/3}}{62 \cdot 31} - \frac{n^{1/3}}{62} \geq \frac{3n^{2/3}}{10} \geq \frac{2n^{1/3}}{31} - 1 \\ &\geq \lceil \frac{2d_i}{n-d_i} \rceil , \end{aligned}$$

where we have used the fact that, being $i \in H$, $d_i \leq n - 31n^{2/3}$, so that

$$\lceil \frac{2d_i}{n-d_i} \rceil \leq \frac{2(n-31n^{2/3})}{31n^{2/3}} + 1 = \frac{2n^{1/3}}{31} - 1 .$$

In order to prove the claim, we finally observe that if $d_i < \frac{n^{1/3}}{15}$, again $|F_i| = 1$ and in line 13 at most two agents are removed from H. Therefore, if s is the total number of agents of degree less than $\frac{n^{1/3}}{15}$ selected in line 11 during the execution of Algorithm 1, at the beginning of iteration t the number of agents left in H is at least $\frac{n^{1/3}}{62} - 2s - (t-1-s) = \frac{n^{1/3}}{62} - s - t + 1 \geq 1$, as $s \leq \frac{n^{1/3}}{124}$ and $t \leq \frac{n^{1/3}}{124}$. $\square$

To prove the second statement of Lemma 5.5, we will differentiate two cases, based on the degree of the agent $i \in H$ picked at a certain iteration. First we will consider agents with degree $d_i \geq \frac{n^{1/3}}{15}$.

Lemma B.2. *Let $\phi \geq \frac{n^{1/3}}{62}$ and π be the partition returned by Algorithm 1. Then, each agent $i \in Gr$ with degree $d_i \geq \frac{n^{1/3}}{15}$ is green.*

Proof. As shown in the proof of the previous lemma, if i has degree $d_i \geq \frac{n^{1/3}}{15}$, then coalition $\pi(i)$ will contain agent i and exactly $\lceil \frac{2d_i}{n-d_i} \rceil$ neighbors of i in $N \setminus H$. This allows us to determine exactly the utility of i in π.

In particular, if $d_i \leq \frac{n}{3}$, being $\lceil \frac{2d_i}{n-d_i} \rceil = 1$, $v_i(\pi) = 1/2$, as agent i in her coalition forms a matching with one neighbor in $N \setminus H$. Consider then a coalition C sampled according to the uniform distribution and containing i. We now show that the probability that $v_i(C) > 1/2$ is low enough to respect the definition of greeness. In fact, the following Chernoff bounds hold:

$$\Pr_{C \sim U(2^N)} \left[|C \cap N_i| > \frac{3d_i}{4} \right] \leq e^{-\frac{d_i}{24}} \leq 2^{-n^{1/3}/15 \cdot 24} ;$$

$$\Pr_{C \sim U(2^N)} \left[|C| < \frac{n}{4} \right] \leq e^{-\frac{n}{16}} < 2^{-n^{1/3}/16} .$$

By applying the union bound, the probability that one of the above two events holds is at most $2^{-n^{1/3}/15 \cdot 24} + 2^{-n^{1/3}/16} < 2^{-n^{1/3}}$. Therefore, with probability at least $1 - 2^{-n^{1/3}}$, none of the two events hold, and recalling that by hypothesis $d_i \leq n/3$, $v_i(C) \leq \frac{3d_i}{4n} < \frac{1}{2}$, hence i is green.

If $d_i > \frac{n}{3}$, being i in a coalition with exactly $\lceil \frac{2d_i}{n-d_i} \rceil$ neighbors in $N \setminus H$, it is possible to lower bound $v_i(\pi)$ as

$$v_i(\pi) = \frac{\lceil \frac{2d_i}{n-d_i} \rceil}{\lceil \frac{2d_i}{n-d_i} \rceil + 1} \geq \frac{2d_i}{n+d_i} .$$

We now show that the probability that for a sampled coalition C containing i it results $v_i(C) \geq \frac{2d_i}{n+d_i}$ is again low enough to satisfy the definition of greeness. We will make use of the following Chernoff's bounds. Given α and β, positive constants, it holds:

$$\Pr_{C \sim U(2^N)} \left[|C \cap N_i| > (1+\alpha) \frac{d_i}{2} \right] \leq e^{-\frac{\alpha^2 d_i}{6}} ;$$

$$\Pr_{C \sim U(2^N)} \left[|C| < (1-\beta) \frac{n}{2} \right] \leq e^{-\frac{\beta^2 n}{4}} .$$

Similarly as above, with probability a least $1 - (e^{-\frac{\alpha^2 d_i}{6}} + e^{-\frac{\beta^2 n}{4}})$, it holds that $v_i(C) \leq \frac{(1+\alpha)d_i}{(1-\beta)n}$. We want to show that, for suitable α and β, this quantity is lower than $\frac{2d_i}{n+d_i}$ giving us the desired result. To this aim, choosing $\beta = \frac{1}{2n^{1/3}}$ and $\alpha = \frac{1}{\frac{n^{1/3}}{15} - 1}$,

$$\begin{aligned} v_i(C) \leq \frac{(1+\alpha)d_i}{(1-\beta)n} &= \frac{\left(1 + \frac{1}{\frac{n^{1/3}}{15} - 1}\right) d_i}{\left(1 - \frac{1}{2n^{1/3}}\right) n} \\ &= \frac{2d_i \left(\frac{n^{1/3}}{15}\right)}{\left(\frac{n^{1/3}}{15} - 1\right)(2n - n^{2/3})} \\ &= \frac{2d_i}{15\left(\frac{n^{1/3}}{15} - 1\right)(2n^{2/3} - n^{1/3})} \\ &= \frac{2d_i}{2n - 31n^{2/3} + 15n^{1/3}} \leq \frac{2d_i}{n+d_i} . \end{aligned}$$

It is straightforward to see that:

$$e^{\frac{-\alpha^2 d_i}{6}} + e^{\frac{-\beta^2 n}{4}} \le e^{-\frac{25n^{1/3}}{2}} + e^{-\frac{n^{1/3}}{16}} \le 2 \cdot 2^{-\frac{n^{1/3}}{16}} = 2^{-\frac{n^{1/3}}{16}+1} .$$

□

The following lemma deals with the agents with lowest degrees, that can be regrouped by Algorithm 1.

Lemma B.3. *Let $\phi \ge \frac{n^{1/3}}{62}$ and π be the partition returned by Algorithm 1. Then, each agent $i \in Gr$ with degree $d_i < \frac{n^{1/3}}{15}$ is green.*

Proof. Differently from the case addressed in the previous lemma, agents $i \in H$ with degree $d_i < \frac{n^{1/3}}{15}$ are possibly regrouped by the procedure. This in particular happens when they are not able to select a singleton neighbor in $N \setminus H$, and thus they must select a neighbor already included in some non-singleton coalition formed in the previous iterations. Since the algorithm stops after at most $\frac{n^{1/3}}{124}$ steps, the worst possible scenario is when all these agents fall together in a single coalition of size $\frac{n^{1/3}}{124} + 1$, which implies that $v_i(\pi) \ge \frac{1}{1+n^{1/3}/124}$.

Consider an agent $i \in \mathrm{Gr}$ with $d_i < \frac{n^{1/3}}{15}$. By a Chernoff's bound it holds that:

$$\Pr_{C\sim U(2^N)}\left[|C| \le \frac{n}{2}\left(1 - \frac{1}{n^{1/3}}\right)\right] \le e^{-\frac{n^{1/3}}{4}} = 2^{-\Omega(n^{1/3})} .$$

Observe that, for $|C| > \frac{n}{2}\left(1 - \frac{1}{n^{1/3}}\right) = \frac{n-n^{2/3}}{2}$:

$$v_i(C) \le \frac{d_i}{|C|} \le \frac{2n^{1/3}}{15\left(n - n^{2/3}\right)} = \frac{2}{15n^{1/3}\left(n^{1/3} - 1\right)} < \frac{1}{n^{1/3}} < \frac{1}{1 + n^{1/3}/124} .$$

Thus, with probability greater than $1 - 2^{-\Omega(n^{1/3})}$, $v_i(C) < v_i(\pi)$, which concludes the proof. □

C Anonymous Hedonic Games – missing proofs

In the following, we will report the missing proofs of Section 6.

First, we will prove the Chernoff bounds over the size of sampled coalitions from a bounded distribution:

Lemma 6.3. *Let X be the random variable representing the size of $C \sim \mathcal{D}$, with $\mathcal{D}$ λ-bounded, and let $\mu := E[X]$. Then,*

$$\frac{n}{\lambda+1} \le \mu \le \frac{\lambda n}{\lambda+1} \qquad \textit{and} \qquad \Pr_{\mathcal{D}}\left[\,|X-\mu| \ge \Delta\cdot\mu\,\right] \le \frac{\varepsilon}{2} ,$$

where ε is such that $0 < \varepsilon < 1$ and Δ is the quantity $\sqrt{\frac{3(\lambda+1)\log\frac{4}{\varepsilon}}{n}}$.

Proof. Observe that X can be seen as the sum of n Poisson trials $X_i, i \in N$ where $X_i = 1$ if i belongs to the sampled coalition and 0 otherwise. Let us denote by p_i the probability of the event $X_i = 1$. From Lemma 3.5, we know that $\frac{1}{\lambda+1} \le p_i \le \frac{\lambda}{\lambda+1}$, as the size of the family of subsets that we focus on is 2^{n-1}. Since the mean μ of X is equal to $\sum_{i\in N} p_i$, by summing the above expression over $i \in N$ we obtain exactly:

$$\frac{n}{\lambda+1} \le \mu \le \frac{\lambda n}{\lambda+1} . \tag{4}$$

To show the bound instead, we will use Equation 1, i.e., for $b \in (0,1)$ being constant:

$$\Pr_{\mathcal{D}}\left[\,|X-\mu| \ge b\mu\,\right] \le 2e^{-b^2\mu/3} \le 2e^{-b^2 n/3(\lambda+1)} ,$$

where the last inequality follows by Equation 4. Setting $\alpha = \sqrt{\frac{3(\lambda+1)\log\frac{4}{\varepsilon}}{n}}$ yields $e^{-b^2 n/3(\lambda+1)} = \frac{\varepsilon}{4}$ and concludes the proof. □

The following two results concern the estimation of the interval $I_{\mathcal{D}}(\varepsilon)$. First, we will prove that, with a certain number of samples, it is possible to learn a confidence interval for the mean coalition size μ.

Lemma 6.4. *Given any two constants $\alpha > 0, \delta < 1$, if $m > \frac{n^2 \log 2/\delta}{2\varepsilon^2}$, then:*

$$\Pr_{\mathcal{S}\sim\mathcal{D}^m} [\, |\bar{\mu} - \mu| < \alpha \,] \geq 1 - \delta \, . \tag{3}$$

Proof. By the Hoeffding bound (see Theorem 4.13 in [23]):

$$\Pr_{\mathcal{S}\sim\mathcal{D}^m} [\, |\bar{\mu} - \mu| \geq \alpha \,] \leq 2e^{-\frac{2\varepsilon^2 m}{n^2}} < \delta \, ,$$

where the last inequality holds by hypothesis on m. □

We are now able to prove that $I_{\mathcal{D}}(\varepsilon)$ can indeed be learned from samples.

Lemma 6.5. *By sampling $m = \frac{2\lambda(1+\lambda)n^2 \log n^2/\delta}{\varepsilon}$ sets from $\mathcal{D}$ it is possible to learn exactly the valuations in $I_{\mathcal{D}}(\varepsilon)$, with confidence $1-\delta$.*

Proof. Let us start observing that, for any $s \in I_{\mathcal{D}}(\varepsilon)$,

$$\frac{\varepsilon}{2} \leq \left(1 - \frac{\varepsilon}{2}\right) \leq \sum_{t \in I_{\mathcal{D}}(\varepsilon)} \Pr_{C\sim\mathcal{D}} [\, |C| = t \,] \leq \lambda \cdot |I_{\mathcal{D}}(\varepsilon)| \cdot \Pr_{C\sim\mathcal{D}} [\, |C| = s \,] \, ,$$

where the second inequality follows by Lemma 6.3 while the third by the definition of λ-bounded. Therefore, $\Pr_{C\sim\mathcal{D}} [\, |C| = s \,] \geq \frac{\varepsilon}{2\lambda n}$. In order to learn exactly $v_i(s)$ for a certain s, it is sufficient to sample a coalition of size s containing agent i. By the definition of conditional probability we have $\Pr_{C\sim\mathcal{D}} [\, |C| = s \wedge i \in C \,] = \Pr_{C\sim\mathcal{D}} [\, i \in C \mid |C| = s \,] \cdot \Pr_{C\sim\mathcal{D}} [\, |C| = s \,]$. It is not hard to see that the conditional probability remains a λ-bounded distribution if we restrict the probability space on coalitions of size s. We can therefore apply Lemma 3.5 and conclude that

$$\Pr_{C\sim\mathcal{D}} [\, |C| = s \wedge i \in C \,] \geq \frac{s}{s + \lambda(n-s)} \cdot \frac{\varepsilon}{2\lambda n} \geq \frac{\varepsilon}{2\lambda(1+\lambda)n^2} \, .$$

As a consequence, the probability that, sampled m coalitions, none of them is of size s and contains i is upper bounded by:

$$\left(1 - \frac{\varepsilon}{2\lambda(1+\lambda)n^2}\right)^m \leq e^{-\frac{m\varepsilon}{2\lambda(1+\lambda)n^2}} \, .$$

By setting $m = \frac{2\lambda(1+\lambda)n^2 \log n^2/\delta}{\varepsilon}$, we finally obtain that the probability that we have not learned the valuation of i for the size s after m samples is less than δ/n^2.

In conclusion, applying the union bound twice, the probability that for at least one agent we have not learned the valuation for some size $s \in I_{\mathcal{D}}(\varepsilon)$ is at most δ. This concludes our proof. □

We will turn our attention now to single-peaked anonymous HGs. In this case, we are able to show a different procedure, that refines the result already obtained in the general case.

Theorem 6.2. *Given a λ-bounded distribution $\mathcal{D}$ and a parameter $\delta \in (0,1)$, for any single-peaked anonymous HG instance and with confidence $1-\delta$, we can efficiently compute an ε-fractional core-stable partition for every $\varepsilon \geq 4 \cdot \frac{\lambda}{2^{n/4}}$.*

Also in this case, we will have at first a learning phase which is the very same learning phase as in the previous section. Let us define I as in the proof of Theorem x; therefore we know that with confidence $1-\delta$, $I_{\mathcal{D}}(\varepsilon) \subseteq I$. A crucial observation is that if we restrict single-peaked preferences on I preferences remain single-peaked. Let $\{s_1, \ldots, s_k\}$ be the sizes in I according to the single-peaked ordering. Giovanna

We define p_i the peak of i in I, and for each $h \in [k]$:

$$\begin{aligned} L_h &= \{i \in N \,|\, p_i = s_\ell \wedge \ell < h\} \\ E_h &= \{i \in N \,|\, p_i = s_h\} \\ G_h &= \{i \in N \,|\, p_i = s_\ell \wedge \ell > h\} \, . \end{aligned}$$

Let $h^* \in [k]$ be the highest index such that $|L_{h^*}| \leq n/2$. For the sake of simplicity, from now on we will omit h^* and use the notation L, E, G in place of $L_{h^*}, E_{h^*}, G_{h^*}$. Notice that, by definition, $|L| + |E| \geq n/2$, and hence $|G| \leq n/2$.

The computation phase works as follows.

Let $s^* = s_{h^*}$ and let $r = n \mod s^*$. We create $(n-r)/s^*$ coalitions of size s^* and a coalition of size r (if $r > 0$). In forming coalitions of size s^*, we give priority to agents in E, that is, in the coalition of size r there is an agent of E if and only if any other coalition contains only agents of E.

Let π be the resulting coalition structure and $N' = \{i \in N \, s.t. \, |\pi(i)| = s^*\}$. We denote by $X' = X \cap N'$ and by $x' = |X'|$ for $X \in \{L, G, E\}$.

Lemma C.1. *For any C core-blocking of π, $E' \cap C = \emptyset$.*

Proof. Each agent in E' is in the most preferred coalition. □

Lemma C.2. *For any core-blocking coalition C for π having size $c \in I$,it must hold $L' \cap C = \emptyset \vee G' \cap C = \emptyset$.*

Proof. Consider i, i' in L' and G', respectively. The sizes that are better than s^* for i in I are the ones that are worse than s^* for i'. Therefore, i and i' cannot be both in C. □

Proof of Theorem 6.2. First we will show that $l' + e' \geq n/4 \vee g' + e' \geq n/4$ must hold. Indeed, by definition, $l' + g' + e' = n - r \geq n/2$. If $l' + e' \leq n/4$, then $g' + e' \geq g' \geq n/4$; similarly, if $g' + e' \leq n/4$, then $l' + e' \geq l' \geq n/4$.

Let us estimate how many core-blocking coalitions C with sizes $c \in I$ do exist. By the previous lemmas, all C of this kind do not contain agents from E', and contain agents either from L' or from G' but not from both. Therefore, the number of possible core-blocking coalitions with size in I satisfies:

$$\begin{aligned} |\{C \text{ core-blocks } \pi \wedge c \in I\}| &\leq 2^{n-e'-l'} + 2^{n-e'-g'} \\ &\leq 2 \cdot \max\{2^{n-e'-l'}, 2^{n-e'-g'}\} \\ &\leq 2^{\frac{3n}{4}+1} \, , \end{aligned}$$

where the last inequality holds by the observation above. Let us set $a = 2^{\frac{3n}{4}+1}/2^n = 2^{1-\frac{n}{4}}$.

In conclusion, the probability of sampling a core blocking coalition is given by

$$\begin{aligned} \Pr_{C \sim \mathcal{D}}[\, C \text{ core-blocks } \pi \,] &= \Pr_{C \sim \mathcal{D}}[\, C \text{ core-blocks } \pi \wedge c \in I \,] + \Pr_{C \sim \mathcal{D}}[\, C \text{ core-blocks } \pi \wedge c \notin I \,] \\ &\leq \Pr_{C \sim \mathcal{D}}[\, C \text{ core-blocks } \pi \wedge c \in I \,] + \Pr_{C \sim \mathcal{D}}[\, c \notin I \,] \\ &\leq \frac{\lambda a}{\lambda a + 1 - a} + \varepsilon/2 \leq \varepsilon \, , \end{aligned}$$

where the second to last inequality holds true, with confidence $1 - \delta$, because of Lemma 3.5 and the last inequality holds for $\varepsilon \geq 4 \cdot \frac{\lambda}{2^{\frac{n}{4}}} \geq \frac{4\lambda}{2^{\frac{n}{4}} + 2(\lambda - 1)} = \frac{2\lambda a}{\lambda a + 1 - a}$. □