Noise Robust Core-stable Coalitions of Hedonic Games

Prashant Trivedi TRIVEDI.PRASHANT15@IITB.AC.IN
IEOR, Indian Institute of Technology Bombay

Nandyala Hemachandra NH@IITB.AC.IN
IEOR, Indian Institute of Technology Bombay

Editors: Emtiyaz Khan and Mehmet Gönen

Abstract

In this work, we consider the coalition formation games with an additional component, 'noisy preferences'. Moreover, such noisy preferences are available only for a sample of coalitions. We propose a multiplicative noise model (equivalent to an additive noise model) and obtain the prediction probability, defined as the probability that the estimated PAC core-stable partition of the *noisy* game is also PAC core-stable for the *unknown noise-free* game. This prediction probability depends on the probability of a combinatorial construct called an 'agreement event'. We explicitly obtain the agreement probability for n agent noisy game with $l \geq 2$ support noise distribution. For a user-given satisfaction value on this probability, we identify the noise regimes for which an estimated partition is noise robust; that is, it is PAC core-stable in both noisy and noise-free games. We obtain similar robustness results when the estimated partition is not PAC core-stable. These noise regimes correspond to the level sets of the agreement probability function and are non-convex sets. Moreover, an important fact is that the prediction probability can be high even if high noise values occur with a high probability. Further, for a class of top-responsive hedonic games, we obtain the bounds on the extra noisy samples required to get noise robustness with a user-given satisfaction value.

We completely solve the noise robustness problem of a 2 agent hedonic game. In particular, we obtain the prediction probability function for $l = 2$ and $l = 3$ noise support cases. For $l = 2$, the prediction probability is convex in noise probability, but the noise robust regime is non-convex. Its minimum value, called the safety value, is 0.62; so, below 0.62, the noise robust regime is the entire probability simplex. However, for $l \geq 3$, the prediction probability is non-convex; so, the safety value is the global minima of a non-convex function and is computationally hard.

Keywords: Prediction probability; noise regimes; combinatorial events; safety value; non-convex optimisation; global minima; weak supervision; PAC stability; multiplicative noise

1. Introduction

Coalition formation games are of great interest to researchers because they model natural interactions among multi-agent societies. The coalition formation process can be formalized using the framework of hedonic games. In these games, each agent has a preference over the coalitions they form with the other agents. An outcome of a hedonic game consists of dividing the agent set into disjoint coalitions called partition. Such a partition is referred to as *coalition structure*. A desirable property in hedonic games is the formation of a stable coalition structure. However, any stability notion (Bogomolnaia and Jackson, 2002; Aziz

and Savani, 2016) assumes the complete information of each agent's preferences, i.e., the entire ranking of coalitions by each agent is known. This is one of the strong assumptions in hedonic games. Nonetheless, there is significant work in finding a stable partition of the agent set, if it exists (Brandt et al., 2016).

Authors in Sliwinski and Zick (2017) relax the assumption of complete information and assume that the preferences over only some coalitions are available; they introduce the notion of ϵ-Probably Approximately Correct (ϵ-PAC) stability to learn the stable outcome of the hedonic game. Apart from the assumption about the complete information, we can have preferences corrupted by noise, i.e., the exact preferences of agents are not available; instead, the preferences with errors are observed. We call such observed erroneous preferences, *noisy preferences*. A consequence of these noisy preferences is that a partition that is not stable in a noisy game can be stable in a noise-free game with non-trivial probability or vice-versa.

In this work, one of our goals is to find the probability that a stable partition learned from the observed noisy sample is the same as that of the stable partition of the *unknown noise-free* game (Sec. 2). We obtain similar results when one starts with a particular partition that is not PAC stable for the noisy game. In such a case, we are interested in the probability that the estimated partition is also not core-stable for a noise-free game (Sec. 3). We call these probabilities the *prediction probabilities*. These prediction probabilities depend on a probability of an event called the 'agreement event'. We also obtain the noise regimes where the agreement probabilities are more than a user-given threshold.

As a motivation, let us consider a stylized model of a market for a specific product that three manufacturers $N = \{1, 2, 3\}$ serve. Each manufacturer has preferences, denoted by $\succ_i$, $\forall\ i \in N$, over the coalitions they want to form with other manufacturers. Based on their preferences, a market analyst would like to predict the coalition structure that these three manufacturers form. However, these preferences being private to manufacturers, the market analyst collects them through a noisy channel (or estimates them based on the market's history). For simplicity, assume that the analyst has noisy preferences, denoted by $\succ'_i$, $\forall\ i \in N$, of all the agents (a complete information model) as in the game (1) below:

$$\begin{aligned} &\{12\} \succ'_1 \{1\} \succ'_1 \{123\} \succ'_1 \{13\} \\ &\{12\} \succ'_2 \{2\} \succ'_2 \{123\} \succ'_2 \{23\} \\ &\{123\} \succ'_3 \{23\} \succ'_3 \{13\} \succ'_3 \{3\} \end{aligned} \quad (1) \qquad \begin{aligned} &\{1\} \succ_1 \{12\} \succ_1 \{123\} \succ_1 \{13\} \\ &\{12\} \succ_2 \{2\} \succ_2 \{123\} \succ_2 \{23\} \\ &\{123\} \succ_3 \{23\} \succ_3 \{13\} \succ_3 \{3\} \end{aligned} \quad (2)$$

The noisy core-stable partition as predicted by the market analyst is $\tilde{\pi} = \{\{12\}, \{3\}\}$. However, suppose the noise-free preferences are as in game (2) (these are not known to market analyst). Based on these noise-free preferences the unique core-stable partition is $\pi = \{\{1\}, \{2\}, \{3\}\}$. So, while the market analyst concludes that the manufacturers form a coalition based on the available noisy preferences, they will not. Thus, the market analyst needs to know the prediction probability, the probability that the predicted partition based on the available noisy preferences is the same as the partition of the unknown noise-free game in (2). An interesting phenomenon in the noisy hedonic game is that even if the market analyst misses identifying a core-stable partition, the market has one with non-trivial probability. We consider this in Sec. 3, via their complimentary event.

As a generalization to the above three manufacturers' model, we assume that a learner has preferences over some coalitions collected via a noisy channel. Based on these noisy preferences, the learner's task is to predict the core-stable partition for an unknown noise-

free game based on these partial and noisy preferences. We propose a noise model to investigate noise regimes where the predicted noisy partition is the same as an unknown noise-free partition. It addresses an important aspect of noise-robustness, meaning these partitions are the same with high probability. Specifically, our major contributions are:
(a) In Sec. 2, we propose a multiplicative noise model and obtain the prediction probability. This probability depends on a combinatorial construct called 'agreement event'. For a user-given value on agreement probability, we obtain the noise regimes where an estimated partition of the noisy game is noise-robust.
(b) In Subsec. 2.1, we obtain the lower and upper bounds on the number of noisy samples required to get PAC stable partition for the top-responsive class of hedonic games.
(c) In Sec. 3, we obtain the prediction probability function that a partition $\tilde{\pi}$ is not PAC stable for the noise-free game given that it is not PAC stable for the noisy game.
(d) In Sec. 4, we consider a noisy game with 2 agents with complete information on each agent's preferences. The allowable noise regimes for noise-robustness are non-convex, even though the agreement probability is a convex function.
(e) We now mention some observations of 2 agent game. The prediction probability is non-convex when the noise distribution has l (≥ 3) support. Thus, computing the safety value, i.e., the minimum prediction probability, is computationally hard. So, for user satisfaction values below this safety value, the prediction probability is 1, regardless of the noise values and their probabilities. Also, the noise values that render a user given minimum prediction probability form non-contiguous regions (superlevel sets). A counter-intuitive fact is that the prediction probabilities can be high for some high noise values occurring with high probability. A simple illustration is in the 2 support case, where the prediction probability is 1 even when the value of both agents is inflated with probability 1.

1.1. Notations and preliminaries

This Sec. provides some notations, definitions, and other related backgrounds that we use in the paper subsequently. Let $N = \{1, 2, \ldots, n\}$ be the set of agents and for each agent $i \in N$, let $\mathcal{C}_i = \{S \subseteq N \mid i \in S\}$ be the set of coalitions containing agent i. A hedonic game is a pair $(N, \succeq)$, where $\succeq = (\succeq_1, \succeq_2, \ldots, \succeq_n)$. Here $\succeq_i$ is a reflexive, transitive, and complete preference ranking of agent $i \in N$ over the set $\mathcal{C}_i$. The preference $\succeq_i$ of agent $i \in N$ represents its willingness to form a coalition with other agents.

For any two distinct coalitions $S, T \subseteq \mathcal{C}_i$ we say $S \succ_i T$ if agent $i \in N$ prefers coalition S over T. Also, $S \sim_i T$ iff $S \succeq_i T$ and $T \succeq_i S$, that is agent i is indifferent to coalition S and T. Since the preferences are reflexive, transitive and complete there exists a value function $\boldsymbol{v} : S \subseteq N \mapsto \mathbb{R}^{|S|}$ such that $\boldsymbol{v}(S) = (v_i(S))_{i \in S}$ [1], where $v_i(S) \in \mathbb{R}^+$ is the valuation of an agent i in coalition S. For any coalitions $S, T \in \mathcal{C}_i$, it satisfies that $S \succeq_i T \iff v_i(S) \geq v_i(T)$ (Mas-Colell et al., 1995; Narahari, 2014). The valuation $v_i(S)$ often depends on value $v_i(j) \in \mathbb{R}^+$ of agent j in the eyes of agent i, here $i, j \in S$. We use $(N, \boldsymbol{v})$ to denote the hedonic game.

A typical partition of the agent set in the hedonic game $(N, \boldsymbol{v})$ is denoted by π. Let the coalition containing $i \in N$ in partition π be $\pi(i)$. The hedonic game's outcome is finding a 'stable' partition according to some stability criterion. A partition is 'stable' if no agent

1. Note that $(v_i(S))_{i \in S}$ is a vector of size $|S|$ with each element $v_i(S)$ for agent $i \in S$.

or a group of agents can deviate from it to reach a subjectively better outcome. Various stability criteria are introduced in Bogomolnaia and Jackson (2002) and are nicely reviewed by Aziz and Savani (2016). However, in this paper, we use core, one of the popular stability criteria. A coalition S *core blocks* a partition π, if every agent i in coalition S strictly prefers S to $\pi(i)$, i.e., $S \succ_i \pi(i),\ \forall\ i \in S$. Further, a coalition structure π is said to be *core-stable* if there is no coalition that core blocks π, meaning there is at least one agent $i \in S$ who prefers $\pi(i)$ over S, i.e., $\pi(i) \succeq_i S$.

Recently, for a partial information hedonic game, authors in Sliwinski and Zick (2017) have proposed the PAC learning framework to find a ϵ-PAC stable outcome for several classes of hedonic games. We briefly describe the ϵ-PAC stability framework here (Sliwinski and Zick, 2017). Given a sample $\mathcal{S} = \{(S_1, \boldsymbol{v}(S_1)), \ldots, (S_m, \boldsymbol{v}(S_m))\}$, where $S_1, S_2, \ldots, S_m$ are drawn *i.i.d.* from a distribution over 2^N and the corresponding values $\boldsymbol{v}$'s are obtained from $\mathcal{D}$. An algorithm $\mathcal{A}$ is said to PAC stabilize a class $\mathcal{H}$ of hedonic games if for any hedonic game $(N, \boldsymbol{v}) \in \mathcal{H}$, after seeing examples in $\mathcal{S}$ it can propose a partition π that is unlikely to be core blocked by a coalition sampled from $\mathcal{D}$ with high probability. Formally, for any error and the confidence parameter $\epsilon, \delta > 0$, a partition π is ϵ-PAC stable under $\mathcal{D}$ if $\mathcal{A}$ outputs a ϵ-PAC stable coalition structure or reports that the core is empty, i.e.,

$$\mathbb{P}_{\mathcal{S}}[\mathbb{P}_{T\sim\mathcal{D}}[T \text{ core blocks } \pi \text{ in noise-free game } (N, \boldsymbol{v})] < \epsilon] \geq 1 - \delta, \tag{3}$$

here, the number of samples m are required to be polynomial in $n, \frac{1}{\epsilon}$ and $\log \frac{1}{\delta}$.

As mentioned above, the ϵ-PAC stability notion assumes the correct preferences over the sample of coalitions. However, it is not the case in most realistic scenarios. Often the preferences are erroneous, i.e., corrupted by noise. In this work, we relax both the assumptions of correct and complete knowledge of the preferences. Let the value of each agent in any coalition be corrupted by an unknown noise distribution, $\mathcal{N}$. We denote the complete, reflexive and transitive noisy preferences by $\succeq' = (\succeq'_1, \succeq'_2, \ldots, \succeq'_n)$. The noisy hedonic game is therefore represented by $(N, \succeq')$ or equivalently $(N, \tilde{\boldsymbol{v}})$, where $\tilde{\boldsymbol{v}}(S) = (\tilde{v}_i(S))_{i\in S}$ is such that $\tilde{v}_i(S) \in \mathbb{R}^+$. Formally, we are given a sample $\tilde{\mathcal{S}} = \{(S_1, \tilde{\boldsymbol{v}}(S_1)), \ldots, (S_{\tilde{m}}, \tilde{\boldsymbol{v}}(S_{\tilde{m}}))\}$ from the noisy hedonic game $(N, \tilde{\boldsymbol{v}})$. Here $S_1, S_2, \ldots, S_m$ are drawn *i.i.d.* from a distribution over 2^N and the corresponding values $\tilde{\boldsymbol{v}}$'s are obtained from $\widetilde{\mathcal{D}}$. So, we can find a $\tilde{\epsilon}$-PAC stable partition (this can be done by using an algorithm similar to one given in say, Sliwinski and Zick (2017); Alcalde and Revilla (2004)) if it exists. Let $\tilde{\pi}$ be an $\tilde{\epsilon}$-PAC stable partition of the noisy hedonic game, i.e., with probability at least $1 - \delta$, we have,

$$\mathbb{P}_{\tilde{\mathcal{S}}}[\mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[T \text{ core blocks } \tilde{\pi} \text{ in noisy game } (N, \tilde{\boldsymbol{v}})] < \tilde{\epsilon}] \geq 1 - \delta. \tag{4}$$

Again, the number of samples required are $\tilde{m}$ which is polynomial in $n, \frac{1}{\tilde{\epsilon}}$, and $\log \frac{1}{\delta}$. The entire paper uses the inner probability given in Equation (3) for the noisy game. However, for the noise-free game, we are interested in the probability given in Equation (5) below. This is because we only have samples from the noisy game; hence, the outer probability is taken on noisy samples for noisy and noise-free games.

Let $\alpha_i(S) \sim \mathcal{N}$ be the noise realized to an agent $i \in S \subseteq N$. We assume that for each agent $i \in S$, the noise is the same, i.e., $\alpha_i(S) = \alpha(S) \in \mathbb{R}^+,\ \forall\ i \in S$. To ensure noise distribution support, $\mathcal{N}_{sp}$ is non-empty we assume that it contains 1 and other noise values. So, the noisy value is $\tilde{v}_i(S) := \alpha(S) \cdot v_i(S),\ \forall\ i \in S \subseteq N$. We call this noise model

the *multiplicative noise* model; this is equivalent to the additive noise model as given in Remark 3 below. In Sections 2.2, and 3.1, we also consider scaling at various levels by taking $l \geq 2$ support on the noise distribution (the same level for all members of a given coalition). Another motivation for the same noise level scaling for each agent in the coalition is the following: A learner is collecting the valuation of each coalition via a noisy channel. So, we assume that a noisy channel affects the value of the entire coalition by the same amount. Hence, each agent in a coalition will have the same noise impact, irrespective of its identity. However, suppose an agent i is a member of two coalitions $S, T \in \mathcal{C}_i$. The noise valuation of agent i in coalition S is $\alpha(S)v_i(S)$, and $\alpha(T)v_i(T)$ in coalition T, so, we also have different noise values for the same agent depending on the coalition. Moreover, the assumption of common scaling $\alpha(S)$ is necessary to carry out the Probably Approximately Correct (PAC) analysis. The PAC stability definition uses a hypothesis class, in our case, the class of hedonic games. The common $\alpha(S)$ preserves the class of hedonic games under noise, which need not be the case when we scale the value of each coalition at the individual agent level. For example, if the noise-free game belongs to the class of additively separable hedonic games (ASHGs), the noisy game with agent dependent noise scaling $\alpha_i(S)$, $\forall\, i \in S$ may not be ASHG, but it is within ASHG class with common noise scaling $\alpha(S)$. So, we use a common scaling that restricts noisy and noise-free games to the same class. We believe this assumption can be relaxed by taking the larger class of hedonic games; however, we might need additional conditions to ensure the class-preserving property.

It is important to note that our noisy hedonic game setup can be reduced to the noise-free setup in a very specialized setting, i.e., only if $\alpha(S) = 1$ for all coalitions S.

Suppose $\tilde{\pi}$ is any partition of the noisy game $(N, \tilde{\boldsymbol{v}})$. We aim to find the probability that any $T \sim \widetilde{\mathcal{D}}$ core blocks $\tilde{\pi}$ in the noise-free game $(N, \boldsymbol{v})$, i.e.,

$$\mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[T \textit{ core blocks } \tilde{\pi} \textit{ in noise-free game } (N, \boldsymbol{v})]. \tag{5}$$

We call the above probability, *prediction probability*. In each case, i.e., when $\tilde{\pi}$ is $\tilde{\epsilon}$-PAC stable partition of the noisy game or not, we bound these prediction probability in Sections 2 and 3, respectively. Prediction probability is a performance measure associated with noise robustness, as defined below:

Definition 1 (ζ noise-robust core-stable partition $\tilde{\pi}$) *A partition $\tilde{\pi}$ is ζ noise-robust core-stable partition if (a) $\tilde{\pi}$ is $\tilde{\epsilon}$-PAC stable partition of noisy game $(N, \tilde{\boldsymbol{v}})$, and (b) prediction probability in Equation (5) is less than ϵ, where $\epsilon = 1 - (1 - \tilde{\epsilon})\zeta$ with $\zeta \in (0, 1]$.*

Definition 2 (η noise-robust non core-stable partition $\tilde{\pi}$) *A partition $\tilde{\pi}$ is η noise-robust non core-stable partition if (a) $\tilde{\pi}$ is not $\tilde{\epsilon}$-PAC stable partition of noisy game $(N, \tilde{\boldsymbol{v}})$ and (b) prediction probability in Equation (5) is more than $1 - \epsilon$, where $\epsilon = 1 - (1 - \tilde{\epsilon})\eta$ with $\eta \in (0, 1]$.*

Remark 3 *Additive noise model: Our noise model is a fairly generic one. For example, if the noise is additive, i.e., $\tilde{v}_i(S) = \alpha(S) + v_i(S)$ then, taking exponential on both sides, we have $e^{\tilde{v}_i(S)} = e^{\alpha(S)+v_i(S)} = e^{\alpha(S)} \cdot e^{v_i(S)}$. With $\tilde{V}_i(S) = e^{\tilde{v}_i(S)}$, $\Gamma(S) = e^{\alpha(S)}$, and $V_i(S) = e^{v_i(S)}$, we have $\tilde{V}_i(S) = \Gamma(S)V_i(S)$. Hence, for robustness to an additive noise model, one can reduce it to a noisy hedonic game $(N, \widetilde{\boldsymbol{V}})$ with multiplicative noise $\Gamma(S)$.*

Remark 4 *Note that in Equation (5) we use noise-free values $\boldsymbol{v}$'s to check whether a coalition can potentially block a given noisy core-stable partition $\tilde{\pi}$. However, we only have samples from the noisy game, so we use $T \sim \widetilde{\mathcal{D}}$ instead of $T \sim \mathcal{D}$.*

2. Partial information noisy game with $\tilde{\pi}$ as $\tilde{\epsilon}$-PAC stable partition

Let $\tilde{\mathcal{S}} = \{(S_1, \tilde{\boldsymbol{v}}(S_1), \ldots, (S_{\tilde{m}}, \tilde{\boldsymbol{v}}(S_{\tilde{m}}))\}$ be a sample of coalitions drawn *i.i.d* from the distribution $\widetilde{\mathcal{D}} = \mathcal{D} \times \mathcal{N}$ over 2^N. Let $\tilde{\pi}$ be $\tilde{\epsilon}$-PAC stable outcome of noisy game $(N, \tilde{\boldsymbol{v}})$. Therefore, with probability at least $1-\delta$, $\forall\ \tilde{\epsilon} > 0$, we have

$$\begin{aligned} \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[T \textit{ core blocks } \tilde{\pi}] < \tilde{\epsilon}, \quad \textit{or} \quad & \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i)), \quad \forall\ i \in T] < \tilde{\epsilon}, \\ & \textit{or} \quad \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{i\in T} \quad \tilde{v}_i(\tilde{\pi}(i)) \geq \tilde{v}_i(T)] \geq 1-\tilde{\epsilon}. \end{aligned} \tag{6}$$

For an unknown noise-free hedonic game $(N, \boldsymbol{v})$, we now find the prediction probability given in Equation (5). To this end, we first define set $\mathcal{R}(T)$ for any coalition T as $\mathcal{R}(T) \coloneqq \{\tilde{\pi}(i) \in \tilde{\pi} \mid i \in T\}$, i.e., for all agents $i \in T$, it is the set of all coalitions containing agent i in the partition $\tilde{\pi}$. Moreover, for any coalition T, and partition $\tilde{\pi}$, we define an agreement event $M(\tilde{\pi}, T)$ containing the set of all noise levels $\alpha(\tilde{\pi}(i))$ and $\alpha(T)$ such that all the coalitions $\tilde{\pi}(i) \in \mathcal{R}(T)$ are preferred over coalition T by every agent $i \in T$ in both noisy and noise-free game. Formally, it is defined as

$$M(\tilde{\pi}, T) \coloneqq \{(\{\alpha(\tilde{\pi}(i))\}_{\tilde{\pi}(i)\in\mathcal{R}(T)}, \alpha(T)) : \cap_{i\in T}\{v_i(\tilde{\pi}(i)) \geq v_i(T) \ \cap\ \alpha(\tilde{\pi}(i))v_i(\tilde{\pi}(i)) \geq \alpha(T)v_i(T)\}\}.$$

Let $f_T(\boldsymbol{p}, \boldsymbol{\alpha}) \coloneqq \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[M(\tilde{\pi}, T)]$ be the probability of agreement event $M(\tilde{\pi}, T)$, where $\boldsymbol{p}$ is the probability mass function of noise values $\boldsymbol{\alpha}$ [2]. Note that $M(\tilde{\pi}, T)$ is not known, since the noise-free values $v_i(T)$ and $v_i(\tilde{\pi}(i))$ are not known. However, for $l \geq 2$ support noise distribution we obtain explicit expressions for $f_T(\boldsymbol{p}, \boldsymbol{\alpha})$ in Sec. 2.2. We also use $f_T(\boldsymbol{p}, \boldsymbol{\alpha})$ later as user satisfaction value. The following Theorem gives probability that unknown noise-free game $(N, \boldsymbol{v})$ has $\tilde{\pi}$ as ϵ-PAC stable partition (ϵ is identified in the Theorem 5 below in terms of $\boldsymbol{p}, \boldsymbol{\alpha}$ and $\tilde{\epsilon}$) if noisy game $(N, \tilde{\boldsymbol{v}})$ has $\tilde{\pi}$ as $\tilde{\epsilon}$-PAC stable partition.

Theorem 5 *Let $\tilde{\pi}$ be $\tilde{\epsilon}$-PAC stable outcome of the noisy game $(N, \tilde{\boldsymbol{v}})$. Then, $\tilde{\pi}$ is ϵ-PAC stable for noise-free game $(N, \boldsymbol{v})$, i.e., $\mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{i\in T} \ v_i(\tilde{\pi}(i)) \geq v_i(T)] \geq 1-\epsilon$, where $\epsilon > 0$ satisfies $(1-\tilde{\epsilon})f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = 1-\epsilon$ with $f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \mathbb{P}[M(\tilde{\pi}, T)]$.*

Proof Consider the following probability

$$\begin{aligned} \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{i\in T}\ v_i(\tilde{\pi}(i)) \geq v_i(T)] &\geq \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{i\in T} v_i(\tilde{\pi}(i)) \geq v_i(T) | \cup_{j\in T}\ \tilde{v}_j(\tilde{\pi}(j)) \geq \tilde{v}_j(T)] \\ &\quad \times\ \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{j\in T}\tilde{v}_j(\tilde{\pi}(j)) \geq \tilde{v}_j(T)] \\ &\geq (1-\tilde{\epsilon})\mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cup_{i\in T} v_i(\tilde{\pi}(i)) \geq v_i(T) | \cup_{j\in T}\ \tilde{v}_j(\tilde{\pi}(j)) \geq \tilde{v}_j(T)] \\ &\geq (1-\tilde{\epsilon})\mathbb{P}[(\cup_{i\in T} v_i(\tilde{\pi}(i)) \geq v_i(T)) \cap (\cup_{j\in T}\tilde{v}_j(\tilde{\pi}(j)) \geq \tilde{v}_j(T))] \\ &\quad (\because \mathbb{P}(A|B) \geq \mathbb{P}(A \cap B)) \\ &= (1-\tilde{\epsilon})\mathbb{P}[\cup_{j\in T} \cup_{i\in T} \{v_i(\tilde{\pi}(i)) \geq v_i(T) \cap \tilde{v}_j(\tilde{\pi}(j)) \geq \tilde{v}_j(T)\}] \\ &\geq (1-\tilde{\epsilon})\mathbb{P}[\cap_{i\in T}\{v_i(\tilde{\pi}(i)) \geq v_i(T) \cap \tilde{v}_i(\tilde{\pi}(i)) \geq \tilde{v}_i(T)\}] \\ &= (1-\tilde{\epsilon})\mathbb{P}[M(\tilde{\pi}, T)] = (1-\tilde{\epsilon})f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = 1-\epsilon. \end{aligned}$$

2. Here $\boldsymbol{\alpha}$ contains all possible noise values, and $\boldsymbol{p}$ is the probability mass function of noise values in $\boldsymbol{\alpha}$.

This ends the proof. ■

In Theorem 5, we have $(1-\tilde{\epsilon})f_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1-\epsilon$ for any $\tilde{\epsilon} > 0$. This implies $\epsilon = 1-(1-\tilde{\epsilon})f_T(\boldsymbol{p},\boldsymbol{\alpha}) = \tilde{\epsilon}$ if $f_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1$. So, for an arbitrary $\epsilon > 0$, we have arbitrary $\tilde{\epsilon} > 0$ if $f_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1$. However, it is not true even in the $l = 2$ support noise model. For example, as we see in Sec. 1 of the supplementary material (SM) that we have $f_T(p,\alpha) = 1$ iff $p = 0$ or $p = 1$, i.e., when values of all the coalitions are either scaled by some scalar $\alpha > 1$, or they are retained. Therefore, we relax the requirement of $f_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1$, and ask for $f_T(\boldsymbol{p},\boldsymbol{\alpha}) = \zeta$ for user-given ζ. In some situations, the ζ captures the satisfaction value of an external agent trying to predict the partition of a noise-free game without having its knowledge. That is, a higher ζ is preferred. In particular, if $\zeta = 1$, we have $\epsilon = \tilde{\epsilon}$.

Theorem 6 *If a partition $\tilde{\pi}$ is $\tilde{\epsilon}$-PAC stable for the noisy game $(N,\tilde{\boldsymbol{v}})$ and for $\epsilon = 1-(1-\tilde{\epsilon})\zeta$ it is ϵ-PAC stable for the noise-free game $(N,\boldsymbol{v})$, then it is also ζ noise-robust core-stable.*

Proof Recall, from Theorem 5, we have $\mathbb{P}[T \text{ core blocks } \tilde{\pi} \text{ for } (N,\boldsymbol{v})] \leq \epsilon$. Here $\epsilon = 1-(1-\tilde{\epsilon})f_T(\boldsymbol{p},\boldsymbol{\alpha})$. Setting $f_T(\boldsymbol{p},\boldsymbol{\alpha}) = \zeta$, we have $\epsilon = 1-(1-\tilde{\epsilon})\zeta = 1-\zeta+\tilde{\epsilon}\zeta$. So, for this ϵ, the partition $\tilde{\pi}$ is ζ noise-robust core-stable from Definition 1. ■

The agreement probability $f_T(\boldsymbol{p},\boldsymbol{\alpha})$ being the same as user-given satisfaction value ζ identifies the noise regimes $I^\star(T,\zeta)$ for which at least the ζ fraction of preferences are preserved. The following Theorem shows that the noise-regime for which a partition $\tilde{\pi}$ is core-stable in both noisy and noise-free games with a user-given satisfaction value ζ is indeed non-empty.

Theorem 7 *Let $\tilde{\pi}$ be $\tilde{\epsilon}$-PAC stable partition of noisy game $(N,\tilde{\boldsymbol{v}})$ and it is ϵ-PAC stable for noise-free game $(N,\boldsymbol{v})$. Then, for a sample $\mathcal{S}_t = \{T_1,\ldots,T_{m_t}\}$ drawn i.i.d. from $\widetilde{\mathcal{D}}$ we obtain a non-empty noise regime $I^\star(\mathcal{S}_t,\zeta) = \cap_{i=1}^{m_t} I^\star(T_i,\zeta)$ for which $\tilde{\pi}$ is ζ noise-robust core-stable partition. Moreover, $\tilde{\pi}$ is ζ noise-robust core-stable partition for the noise regime $I^\star(\zeta) = \cap_{T\subseteq N} I^\star(T,\zeta)$.*

Proof For any coalition T, we first note that $I^\star(T,\zeta) \neq \emptyset$, because $\alpha(\tilde{\pi}(i)) = 1,\ \forall \tilde{\pi}(i) \in \mathcal{R}(T)$; $\alpha(T) = 1$ is always an element of $M(\tilde{\pi},T)$. So, for $\mathcal{S}_t = \{T_1,\ldots,T_{m_t}\}$ we have non-empty noise regimes $I^\star(T_1,\zeta),\cdots,I^\star(T_{\tilde{m}_t},\zeta)$. Also, $\alpha(S) = 1,\ \forall\ S \subseteq N$ is a common element of each $I^\star(T,\zeta),\ \forall\ T \in \mathcal{S}_t$. Therefore, $I^\star(\mathcal{S}_t,\zeta) = \cap_{i=1}^{m_t} I^\star(T_i,\zeta) \neq \emptyset$, i.e., is non-empty. Hence, partition $\tilde{\pi}$ is ζ noise robust on the sample $\mathcal{S}_t$ with noise regime $I^*(\mathcal{S}_t,\zeta)$ in accordance to Theorem 5 and Definition 1. Moreover, $I^\star(\zeta) \neq \emptyset$ because of the same reason as mentioned above. The ζ noise-robustness follows from Theorem 5 and 6. ■

In the next Section, we provide the relation between m, and $\tilde{m}$, i.e., the number of samples used to get ϵ and $\tilde{\epsilon}$-PAC stable partition $\tilde{\pi}$ in noise-free and noisy game, respectively for top-responsive hedonic games (Alcalde and Revilla, 2004) and other hedonic games.

2.1. Sample size for top-responsive and other games

In a top-responsive game, the value of each agent in a given coalition depends on the most preferred sub-coalition. Formally, the top-responsive games are described via choice sets $Ch(i,S)$, defined as $Ch(i,S) \coloneqq \{X \subseteq S : \forall\, Y \subseteq S, i \in Y : X \succeq_i Y\}$. The game satisfies the top-responsiveness if (a) $\forall\ i \in N$, and $S \in \mathcal{C}_i, |Ch(i,S)| = 1$, and (b) $\forall\ i \in N$, and $S,T \in \mathcal{C}_i$ if $Ch(i,S) \succ_i Ch(i,T)$ then $S \succ_i T$ or if $Ch(i,S) = Ch(i,T)$, and $S \subset T$, then $S \succ_i T$.

Theorem 8 *For a top-responsive game, let $\tilde{m}$ be the number of samples required to get $\tilde{\epsilon}$-PAC stable partition in noisy game $(N, \tilde{\boldsymbol{v}})$, and m be the samples required for $\tilde{\pi}$ to be ϵ-PAC partition in unknown noise-free game $(N, \boldsymbol{v})$. Then $m\zeta \leq \tilde{m} \leq m + (2n^3 + 2n^4)\left(\frac{(1-\tilde{\epsilon})+\tilde{\epsilon}\zeta}{\tilde{\epsilon}(1+\tilde{\epsilon}\zeta)} \log \frac{2n^3}{\delta}\right)$.*

Proof Recall, to get $\tilde{\epsilon}$-PAC stable partition in the noisy top-responsive games authors in (Sliwinski and Zick, 2017) provide $\tilde{m}$ for top-responsive games as $\tilde{m} = (2n^3+2n^4)\left(\frac{1}{\tilde{\epsilon}} \log \frac{2n^3}{\delta}\right)$. However, from Theorem 5 we have $(1-\tilde{\epsilon})\zeta = 1-\epsilon$, this implies $\epsilon = (1-\zeta) + \zeta\tilde{\epsilon} \geq \zeta\tilde{\epsilon}$. Thus, for ϵ-PAC stability of partition $\tilde{\pi}$ in a top-responsive noise-free game the number of samples m are given by $m = (2n^3 + 2n^4)\left(\frac{1}{\epsilon} \log \frac{2n^3}{\delta}\right) \leq (2n^3 + 2n^4)\left(\frac{1}{\zeta\tilde{\epsilon}} \log \frac{2n^3}{\delta}\right) = \frac{\tilde{m}}{\zeta}$. This gives an upper bound. For a lower bound, again consider $(1-\tilde{\epsilon})\zeta = 1-\epsilon$, therefore we have $\epsilon = (1-\zeta) + \zeta\tilde{\epsilon} \leq 1 + \zeta\tilde{\epsilon}$. That is $\frac{1}{\epsilon} \geq \frac{1}{1+\zeta\tilde{\epsilon}}$. Thus, we have

$$m = (2n^3 + 2n^4)\left(\frac{1}{\epsilon} \log \frac{2n^3}{\delta}\right) \geq (2n^3 + 2n^4)\left(\frac{1}{1+\zeta\tilde{\epsilon}} \log \frac{2n^3}{\delta}\right)$$
$$= (2n^3 + 2n^4)\left(\left\{\frac{1}{\tilde{\epsilon}} - \frac{(1-\tilde{\epsilon})+\tilde{\epsilon}\zeta}{\tilde{\epsilon}(1+\tilde{\epsilon}\zeta)}\right\} \log \frac{2n^3}{\delta}\right) = \tilde{m} - (2n^3 + 2n^4)\left(\frac{(1-\tilde{\epsilon})+\tilde{\epsilon}\zeta}{\tilde{\epsilon}(1+\tilde{\epsilon}\zeta)} \log \frac{2n^3}{\delta}\right).$$

From the lower and upper bounds, we have the result. ■

The above Theorem gives a bound on the extra samples required to get ϵ-PAC stable partition of the unknown noise-free game given $\tilde{\epsilon}$-PAC stable partition of the noisy game. Again the number of samples to get $\epsilon = (1-(1-\tilde{\epsilon})\zeta)$-PAC stable outcome in an *unknown* noise-free game are bounded by the number of samples $\tilde{m}$, the satisfaction value ζ, and the confidence parameter δ. In particular, the number of samples m are polynomial in $n, \frac{1}{\epsilon}, \log\left(\frac{1}{\delta}\right)$, but its upper bound is non-linear in ζ.

We next relate the number of samples and errors in noisy and *unknown* noise-free games. Let $\tilde{\pi}$ be $\tilde{\epsilon}$-PAC stable partition of noisy game when $\tilde{m}$ samples are used. Suppose, we get $(\tilde{\epsilon} - \tilde{\epsilon}')$-PAC partition of the noisy game on increasing the noisy samples to $\tilde{m} + \tilde{m}'$. Let $\tilde{\pi}$ be $(\tilde{\epsilon} - \tilde{\epsilon}')$-PAC stable for the noisy game that uses $\tilde{m} + \tilde{m}'$ samples. Let ϵ_{new} be the error incurred to get $\tilde{\pi}$ partition with $\tilde{m} + \tilde{m}'$ samples in a given noise-free game, then $\epsilon_{new} = 1 - (1 - (\tilde{\epsilon} - \tilde{\epsilon}'))f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = 1 - (1-\tilde{\epsilon})f_T(\boldsymbol{p}, \boldsymbol{\alpha}) - \tilde{\epsilon}' f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \epsilon - \tilde{\epsilon}' f_T(\boldsymbol{p}, \boldsymbol{\alpha}) \leq \epsilon$.

Theorem 9 *For an unknown noise-free game, let $\tilde{\pi}$ be ϵ_{new}-PAC stable partition with $\tilde{m} + \tilde{m}'$ noisy samples, and it is ϵ-PAC stable partition with $\tilde{m}$ noisy samples, then $\epsilon_{new} \leq \epsilon$.*

Remark 10 *The results of Theorem 8 and Theorem 9 can be generalized to any class of hedonic games by suitably obtaining the sample complexity of that class. This is because the number of samples required in noise-free game is function of $n, \frac{1}{\epsilon}, \log\left(\frac{1}{\delta}\right)$.*

To get some more insights we next identify the agreement probability $f_T(\boldsymbol{p}, \boldsymbol{\alpha})$ defined for partial information noise model with $l \geq 2$ noise support in the following subsection. We use the base case of $l = 2$ noise support case in the proofs of results in the next Section. these are deferred to Sec. 1 of the SM due to space considerations.

2.2. n agent l support partial information noisy game

We now consider the $l \geq 2$ support case, i.e., $\mathcal{N}_{sp} = \{\alpha_1, \alpha_2, \ldots, \alpha_l\}$ with respective probabilities $p_1, p_2, \ldots, p_l$, and $\sum_{j\in[l]} p_j = 1$. Here $p_j = \mathbb{P}(\alpha(S) = \alpha_j)$ and $\alpha_j > 0,\ \forall\ j \in [l]$. Moreover, without loss of generality we assume that $\alpha_i < \alpha_j,\ \forall\ i < j$. For above noise support the following Theorem give expression of $f_T(\boldsymbol{p}, \boldsymbol{\alpha})$. To this end, for any coalition T and for all r, s such that $\alpha_r > \alpha_s$, define $\mathcal{I}(\alpha_r, \alpha_s, T) = \left\{ \tilde{\pi}(i) \in \mathcal{R}(T) \;\middle|\; \frac{\tilde{v}_i(\tilde{\pi}(i))}{\tilde{v}_i(T)} \geq \frac{\alpha_r}{\alpha_s} \right\}$.

Theorem 11 *Let $\tilde{\pi}$ be a $\tilde{\epsilon}$-PAC stable outcome of the noisy game $(N, \tilde{\boldsymbol{v}})$ and let $\tilde{\pi}$ be a ϵ-PAC stable outcome of noise-free game $(N, \tilde{\boldsymbol{v}})$, where ϵ is identified as in Theorem 5. Then for noise support $\mathcal{N}_{sp} = \{\alpha_1, \alpha_2, \ldots, \alpha_l\}$, the $f_T(\boldsymbol{p}, \boldsymbol{\alpha})$ is given by:*

$$f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \begin{cases} 1, & if\ \tilde{\pi}(i) = T,\ \forall\ i \in T, \\ \sum_{r,s\in[l]:\alpha_r>\alpha_s} p_s^{|\mathcal{R}(T)|-|\mathcal{I}(\alpha_r,\alpha_s,T)|+1} \times \{(p_r + p_s)^{|\mathcal{I}(\alpha_r,\alpha_s,T)|} - p_s^{|\mathcal{I}(\alpha_r,\alpha_s,T)|}\} & \\ + \sum_{a=1}^{l} p_a \left(\sum_{b=1}^{a} p_b\right)^{|\mathcal{R}(T)|}, & otherwise. \end{cases}$$

The proof uses the principle of Mathematical induction on noise support $l \geq 2$ with base case of $l = 2$ support (Lemma 1 of the SM). The detailed proof is available in Appendix A.

Remark 12 *If we allow $f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \zeta$ for all coalitions $T \subseteq N$ for some user-given satisfaction value ζ, we have noise set $I^\star(\zeta)$ in accordance to Theorem 7. This noise set corresponds to the superlevel sets of the prediction probability function. For this super level set the partition $\tilde{\pi}$ is ζ noise-robust core-stable. Later, Sec. 4 shows that these superlevel sets are non-convex by explicitly deriving the prediction probability.*

3. Partial information noisy game when $\tilde{\pi}$ is not $\tilde{\epsilon}$-PAC stable partition

So far we have assumed that $\tilde{\pi}$ is $\tilde{\epsilon}$-PAC stable partition of the noisy game $(N, \tilde{\boldsymbol{v}})$; however, that is not always the case. For example $\tilde{\pi} = \{\{1\}, \{23\}\}$ is not core stable for the game in (1). In this section, we consider the other case where an estimated partition $\tilde{\pi}$ is not $\tilde{\epsilon}$-PAC stable for the noisy game $(N, \tilde{\boldsymbol{v}})$. Note that $\tilde{\pi}$ not being $\tilde{\epsilon}$-PAC stable doesn't mean that the noisy game $(N, \tilde{\boldsymbol{v}})$ has no stable partition. Given a sample $\tilde{\mathcal{S}}$, we say $\tilde{\pi}$ is not $\tilde{\epsilon}$-PAC stable partition of noisy game $(N, \tilde{\boldsymbol{v}})$ if there is a coalition T that core blocks it with probability at least $1 - \tilde{\epsilon}$. Formally, $\forall\ \tilde{\epsilon} > 0,\ \exists\ T \sim \widetilde{\mathcal{D}}$, such that

$$\mathbb{P}[\cap_{i\in T}\ \ \tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i))] \geq 1 - \tilde{\epsilon}. \tag{7}$$

Our interest is in finding the prediction probability (Equation (5)) that a noise-free game does not have $\tilde{\pi}$ as ϵ-PAC stable outcome (ϵ to be identified in terms of $\tilde{\epsilon}$) when the noisy game does not have $\tilde{\pi}$ as $\tilde{\epsilon}$-PAC stable partition. To this end, for any coalition T, we again define an agreement event $F(T, \tilde{\pi})$ [3]. It contains all the noise values $(\alpha(T), \{\alpha(\tilde{\pi}(i))\}_{\tilde{\pi}\in\mathcal{R}(T)})$ such that coalition T is preferred over all the coalitions $\tilde{\pi}(i) \in \mathcal{R}(T)$ by every agent $i \in T$ in both the noisy and noise-free games. Formally,

$$F(T, \tilde{\pi}) \coloneqq \{(\alpha(T), \{\alpha(\tilde{\pi}(i))\}_{\tilde{\pi}(i)\in\mathcal{R}(T)}) : \cap_{i\in T}\{v_i(T) \geq v_i(\tilde{\pi}(i)) \cap \alpha(T)v_i(T) \geq \alpha(\tilde{\pi}(i))v_i(\tilde{\pi}(i))\}\}.$$

3. Though we use the same names, the agreement event in Sec. 2 is different from this agreement event.

For probability mass $\boldsymbol{p}$ and noise value set $\boldsymbol{\alpha}$, let $h_T(\boldsymbol{p},\boldsymbol{\alpha}) \coloneqq \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[F(T,\tilde{\pi})]$ be the agreement probability. Note that $F(T,\tilde{\pi})$ and hence $h_T(\boldsymbol{p},\boldsymbol{\alpha})$ are not known since the noise-free values $v_i(T)$ and $v_i(\tilde{\pi}(i))$ are not known. However, for $l \geq 2$ support noise distribution $\mathcal{N}$ we obtain $h_T(\boldsymbol{p},\boldsymbol{\alpha})$ explicitly in Sec. 3.1.

Theorem 13 *Suppose the noisy game $(N,\tilde{\boldsymbol{v}})$ does not have $\tilde{\pi}$ as $\tilde{\epsilon}$-PAC stable outcome, i.e., equation (7) is satisfied. Then the prediction probability given in Equation (5) is given by: $\mathbb{P}[\cap_{i\in T}(v_i(T) > v_i(\tilde{\pi}(i)))] \geq (1-\tilde{\epsilon})h_T(\boldsymbol{p},\boldsymbol{\alpha})$, where $\epsilon > 0$ satisfy $(1-\tilde{\epsilon})h_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1-\epsilon$.*

Proof Consider the following: $\mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cap_{i\in T}\ (v_i(T) > v_i(\tilde{\pi}(i)))]$

$$
\begin{aligned}
&\geq \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cap_{i\in T}\ (v_i(T) > v_i(\tilde{\pi}(i)))\ |\ \cap_{i\in T}\ (\tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i)))] \times \mathbb{P}_{T\sim\widetilde{\mathcal{D}}}[\cap_{i\in T}(\tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i)))] \\
&\geq (1-\tilde{\epsilon})\ \mathbb{P}[\cap_{i\in T}(v_i(T) > v_i(\tilde{\pi}(i)))\ |\ \cap_{i\in T}\ (\tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i)))] \\
&\geq (1-\tilde{\epsilon})\ \mathbb{P}[\cap_{i\in T}\{v_i(T) > v_i(\tilde{\pi}(i)) \cap \tilde{v}_i(T) > \tilde{v}_i(\tilde{\pi}(i))\}] \quad (\because \mathbb{P}(A|B) \geq \mathbb{P}(A\cap B)) \\
&= (1-\tilde{\epsilon})\ h_T(\boldsymbol{p},\boldsymbol{\alpha}) = 1-\epsilon.
\end{aligned}
$$

This ends the proof. ■

Let η be the probability of noise agreement event for which coalition T core blocks $\tilde{\pi}$, i.e., $\eta \coloneqq h_T(\boldsymbol{p},\boldsymbol{\alpha})$. Thus $\epsilon = (1-\eta)+\eta\tilde{\epsilon}$ and hence partition $\tilde{\pi}$ is η noise-robust non core-stable in accordance to Definition 2. Moreover, if $\eta = 1$ then $\tilde{\epsilon} = \epsilon$ so, for arbitrary $\tilde{\epsilon} > 0$, we also have arbitrary $\epsilon > 0$.

Remark 14 *Similar to Sec. 2, for a user-given η, we get a noise set $I^{\star}(T,\eta)$ on $\boldsymbol{p}$ for coalition T, i.e., the noise set in which the coalition T core blocks $\tilde{\pi}$ with error more than $1-\epsilon$. This is obtained by setting $\mathbb{P}[F(T,\tilde{\pi})] = h_T(\boldsymbol{(p)},\boldsymbol{\alpha}) = \eta$; that is, $I^{\star}(T,\eta)$ is η level set of agreement probability function $\mathbb{P}[F(T,\tilde{\pi})]$; in other words, it is a super level set of the prediction probability. Hence, $\tilde{\pi}$ is η noise-robust non core-stable in this noise set $I^{\star}(T,\eta)$.*

To better understand the noise robustness, we provide the expression of $h_T(\boldsymbol{p},\boldsymbol{\alpha})$ for $l \geq 2$ support noise models in the following subsection. For $l = 2$ support noise model, we refer the readers to Lemma 3 of the SM. The detailed analysis of the 2 support model gives many more insights and also serves as the base case in the proof of results in the next Section.

3.1. n agents l support partial information noisy game without core

In this section, we obtain the expression of the agreement probability $h_T(\boldsymbol{p},\boldsymbol{\alpha})$ for $l \geq 2$ support noise model, $\mathcal{N}_{sp} = \{\alpha_1,\alpha_2,\ldots,\alpha_l\}$. To this end, for all r,s such that $\alpha_r > \alpha_s$ define $\mathcal{J}(\alpha_r,\alpha_s,T) = \left\{\tilde{\pi}(i) \in \mathcal{R}(T) \;\middle|\; \frac{\tilde{v}_i(\tilde{\pi}(i))}{\tilde{v}_i(T)} \geq \frac{\alpha_s}{\alpha_r}\right\}$. It contains the set of all coalitions in the set $\mathcal{R}(T)$, such that $\alpha_r > \alpha_s$, and $\frac{\tilde{v}_i(\tilde{\pi}(i))}{\tilde{v}_i(T)} \geq \frac{\alpha_s}{\alpha_r}$. The following Theorem provides the expression of $h_T(\boldsymbol{p},\boldsymbol{\alpha})$. For proof refer to Sec. 3 of the SM.

Theorem 15 *For n agent noisy hedonic game $(N,\tilde{\boldsymbol{v}})$ with $\mathcal{N}_{sp} = \{\alpha_1,\alpha_2,\ldots,\alpha_l\}$, the agreement probability $h_T(\boldsymbol{p},\boldsymbol{\alpha})$ is given by:*

$$
h_T(\boldsymbol{p},\boldsymbol{\alpha}) = \begin{cases} 1, & \text{if } \tilde{\pi}(i) = T,\ \forall\ i \in T, \\ \sum_{r,s\in[l]:\alpha_r>\alpha_s} p_r^{|\mathcal{R}(T)|-|\mathcal{J}(\alpha_r,\alpha_s,T)|+1} \times \{(p_s+p_r)^{|\mathcal{J}(\alpha_r,\alpha_s,T)|} - p_r^{|\mathcal{J}(\alpha_r,\alpha_s,T)|}\} & \\ + \sum_{a=1}^{l} p_a \left(\sum_{b=a}^{l} p_b\right)^{|\mathcal{R}(T)|}, & \text{otherwise.} \end{cases}
$$

Remark 16 *Let $h_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \eta$ for some user-given satisfaction value η, we get a set of noise values in accordance to the Remark 14. In this case, the noise set depends on $|\mathcal{R}(T)|$, and $|\mathcal{J}(\alpha_r, \alpha_s, T)|$, $\forall\ \alpha_r > \alpha_s$ for coalition T. Also, the partition $\tilde{\pi}$ is η noise-robust non core-stable in the noise set $I^\star(T, \eta)$.*

Remark 17 *Theorem 13 provides the probability that $\tilde{\pi}$ is not ϵ-PAC stable outcome in noise-free game $(N, \boldsymbol{v})$, when it is not $\tilde{\epsilon}$-PAC stable outcome in noisy game. Therefore, the probability that a noise-free game has $\tilde{\pi}$ as ϵ-PAC stable outcome, given that the noisy game does not have $\tilde{\pi}$ as $\tilde{\epsilon}$-PAC stable outcome is compliment of the probability in Theorem 13.*

4. 2 agent full information model

This Section considers the complete information game with 2 agents. So, valuations on all coalitions are known in the noisy game; hence, a noisy core-stable partition is also known. Even though this Section is a particular case of previous sections, we get many valuable insights that enhance our understanding regarding noise robustness in noisy hedonic games. For example, we have this counter-intuitive fact that the prediction probability can be high if the noise value occurs with a high probability. A concrete illustration is that the prediction probability turns out to be 1 when values of all agents are inflated by $\alpha > 1$ with probability 1; in fact, both games have the same preferences and hence identical partitions.

4.1. 2 support noise distribution

Let the noise support be $\mathcal{N}_{sp} \in \{1, \alpha\}$ with $\alpha > 1$, such that $\mathbb{P}[\mathcal{N}_{sp} = \alpha] = p = 1 - \mathbb{P}[\mathcal{N}_{sp} = 1]$. Note that $\alpha > 1$ is not a restrictive condition; even if we allow $\alpha < 1$, we will get results similar to the ones presented below. Given a noisy game and its corresponding core-stable partition, we aim to find the prediction probability that a core-stable partition of the *unknown* noise-free game is the same as a core-stable partition of a noisy game. Formally, for a user-given $\zeta \in (0, 1]$, we find $\mathbb{P}[\pi = \tilde{\pi} \mid \textit{noisy game}] \geq \zeta$.

We want to emphasize that the above prediction probability is the same as the one given in Equation (5). Since, in a 2 agent complete information game, the noisy core-stable partition $\tilde{\pi}$ always exists; hence $\tilde{\epsilon} = 0$ in Theorem 5. So, with $f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \zeta$ we have $\mathbb{P}_{T \sim \widetilde{\mathcal{D}}}[T \text{ does not core blocks } \tilde{\pi} \textit{ in } (N, \boldsymbol{v})] \geq \zeta$. Consider the following 2 agents' noisy game.

$$\tilde{v}_1(12) > \tilde{v}_1(1);\ \ \tilde{v}_2(12) > \tilde{v}_2(2). \qquad \text{(game 1)}$$

Clearly, $\tilde{\pi} = \{12\} = N$ is the core-stable outcome of the above noisy game. The following Lemma gives the prediction probability for the above game.

Lemma 18 *For noisy game 1 with complete information on $\tilde{\boldsymbol{v}}$, and $\mathcal{N}_{sp} = \{1, \alpha\}$, we have*

$$\mathbb{P}[\pi = \tilde{\pi} \mid \textit{game 1}] = \begin{cases} 1 - p(1 - p^2), & \textit{if } \alpha \geq \overline{r} \\ 1 - p(1 - p), & \textit{if } \underline{r} \leq \alpha < \overline{r} \\ 1, & \textit{if } \alpha < \underline{r}, \end{cases} \tag{8}$$

where $\overline{r} = \max\left\{\frac{\tilde{v}_1(12)}{\tilde{v}_1(1)}, \frac{\tilde{v}_2(12)}{\tilde{v}_2(2)}\right\}$, *and* $\underline{r} = \min\left\{\frac{\tilde{v}_1(12)}{\tilde{v}_1(1)}, \frac{\tilde{v}_2(12)}{\tilde{v}_2(2)}\right\}$.

Also, this prediction probability $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1]$ *is convex in* p*. So, while the minimal value for* $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1]$ *occurs for noise probabilities around* $p = 0.5$ *(depending on* $\alpha, \overline{r}$ *and* $\underline{r}$*), the maximal value of it is* 1 *at* $p = 0$ *and* $p = 1$*.*

The proof is deferred to Sec. 4.1 of the SM. The above lemma has following insights: the prediction probability depends on three factors, p, α, and $\tilde{\boldsymbol{v}}$. Note that p is not known because the noise distribution is unknown. We know $\tilde{\boldsymbol{v}}$'s hence $\underline{r}$, and $\overline{r}$ are known. If p is close to 0.5, then the prediction probability $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1]$ is close to 0.62, for $\alpha \geq \overline{r}$, and close to 0.75 for $\underline{r} \leq \alpha < \overline{r}$. So, the prediction probability is the least when noise is random. We call this minimum prediction probability the *safety value*.

Suppose we allow the user-given satisfaction value on the prediction probability, i.e., we relax the condition $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1] = 1$, and allow a user-given satisfaction value, $\zeta \in (0, 1]$ on the prediction probability, i.e., $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1] = \zeta$. So, for different ranges of α, we get an interval of the noise probabilities p allowable to attain a user-given probability ζ. For example, if $\zeta = 0.9$, then the noise regime, $I^{\star}(\zeta = 0.9) = [0, 0.101] \cup [0.946, 1]$, if $\alpha \geq \overline{r}$; it is $[0, 0.113] \cup [0.887, 1]$, if $\underline{r} \leq \alpha < \overline{r}$; and it is $[0, 1]$, if $\alpha < \underline{r}$. So, for the noise set $I^{\star}(\zeta = 0.9)$, the core-stable partition of the noise-free game is the same as the core-stable partition of the noisy game 1 with probability 0.9. Thus, the noise regime achieving a user given satisfaction value can be non-convex. Figure 1 illustrates these observations. Note

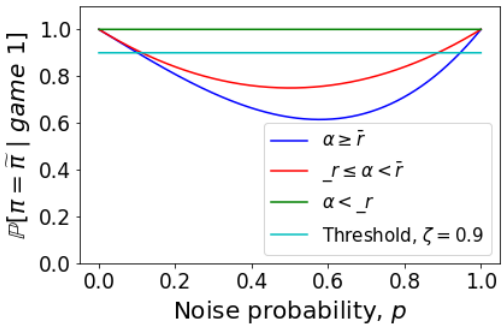

Figure 1: In two agent hedonic game 1 with 2 support noise model, we plot the prediction probability $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1]$ for different ranges of α.

that in Equation (8) we have obtained the conditional probability $\mathbb{P}[\pi = \tilde{\pi} \mid game\ 1]$. One can obtain a similar prediction probability for other noisy games, which we call game 2, game 3, and game 4, whose details are available in Sec. 4.2 of the SM. We summarize the main observations of 2 agents 2 support noise in the Theorem below:

Theorem 19 *Consider 2 agent noisy hedonic game with 2 support noise model, then the prediction probability that* $\pi = \tilde{\pi}$ *given any noisy game* k*,* $k = 1, 2, 3, 4$ *is:*

$$\mathbb{P}[\pi = \tilde{\pi} \mid game\ k] = \begin{cases} 1, & \text{under any condition in } A, \\ q(p, \tilde{v}_1(\cdot), \tilde{v}_2(\cdot), \alpha), & \text{otherwise}, \end{cases} \tag{9}$$

for some function $q(p, \tilde{v}_1(\cdot), \tilde{v}_2(\cdot), \alpha) < 1$, *depending on* $\tilde{v}_1(\cdot), \tilde{v}_2(\cdot)$ *and* α. *The conditions in A are (a)* $k = 1$, *and* $\alpha < \underline{r}$; *(b)* $k = 2$ *and* $\frac{1}{\alpha} \geq \overline{r}$; *(c)* $k = 3$ *and* $\frac{1}{\alpha} \geq \frac{\tilde{v}_1(12)}{\tilde{v}_1(1)}$; *and (d)* $k = 4$ *and* $\frac{1}{\alpha} \geq \frac{\tilde{v}_2(12)}{\tilde{v}_2(2)}$.

Moreover, if $p = 0$ *or* $p = 1$, *we have* $\mathbb{P}[\pi = \tilde{\pi} \mid \text{game } k] = 1$, *for all* $k = 1, 2, 3, 4$.

For a 2 support noise model, the probability that a noise-free game has the same core-stable partition as the noisy game is 1 in many cases, including $p = 1$, i.e., when all values are inflated by α. Thus, the allowable noise regimes for high prediction probabilities can include high noise values. Moreover, $q(p, \tilde{v}_1(\cdot), \tilde{v}_2(\cdot), \alpha) \geq 0.62$ is the safety value. So, we have a lower bound on the prediction probability.

4.2. 3 support noise distribution

Next, consider the 3 support noise, $\mathcal{N}_{sp} = \{1, \alpha_1, \alpha_2\}$, where $\alpha_1 > 1$, and $0 < \alpha_2 < 1$. Let $\mathbb{P}[\alpha(S) = \alpha_1] = p_1$; $\mathbb{P}[\alpha(S) = \alpha_2] = p_2$; *and* $\mathbb{P}[\alpha(S) = 1] = 1 - p_1 - p_2$. Given the noisy game 1, the prediction probability for 3 support noise is given in the Lemma below.

Lemma 20 *For the 3 support noise model the prediction probability* $\mathbb{P}[\pi = \tilde{\pi} \mid \text{game } 1]$ *is*

$$\mathbb{P}[\pi = \tilde{\pi} \mid \text{game } 1] = \begin{cases} g(p_1, p_2), & \text{if } \alpha_1 \geq \overline{r}\ ;\ \frac{1}{\alpha_2} \geq \overline{r}\ ;\ \frac{\alpha_1}{\alpha_2} \geq \overline{r} \\ 1, & \text{if } \alpha_1 < \underline{r}\ ;\ \frac{1}{\alpha_2} < \underline{r}\ ;\ \frac{\alpha_1}{\alpha_2} < \underline{r} \end{cases} \tag{10}$$

where $g(p_1, p_2) = p_1^3 + p_2^3 + 2(p_1(1 - p_1 - p_2)^2 + p_2^2(1 - p_1 - p_2) + p_1 p_2(1 - p_1 - p_2) + p_1 p_2^2) + p_1^2 p_2 + p_1^2(1 - p_1 - p_2) + p_2(1 - p_1 - p_2)^2 + (1 - p_1 - p_2)^3$.

There are **106** more cases in the above Lemma, where in each case, the prediction probability is strictly less than 1 (Sec. 5.1 of SM). Unlike 2 support model (Sec. 4.1) in this case they are *non-convex*; a counter-example is available in Sec. 5.2 of SM.

5. Related work

Stability notions in hedonic games: Researchers have extensively studied hedonic games in the computational social choice community. Some early works in coalition formation games describing the economic situations include that of Dreze and Greenberg (1980); Elkind and Wooldridge (2009). The agents collaborate and have personal preferences on different coalitions. Based on these preferences, agents seek a partition of the agent set. However, which partition to form led to various notions of stability (Bogomolnaia and Jackson, 2002; Banerjee et al., 2001; Aziz and Brandl, 2012). Some of them are core stability, Nash stability, and perfect. In this work, we use core stability. In particular, the core for the simple hedonic games is available in Banerjee et al. (2001).

representation of hedonic games: In many real-life scenarios, there are multiple agents, so storing the hedonic game in a machine takes exponential space. In literature, various concise representations are used because they (often) only require polynomial space. So, apart from various stability notions, much literature is on representing the hedonic games. Some of them includes individually rational lists of coalitions (IRLC) (Ballester, 2004), hedonic coalition nets (HCNs) (Elkind and Wooldridge, 2009), additively separable

games, fractional hedonic games, $\mathcal{B}$-games, $\mathcal{W}$-games, top-responsive games (Alcalde and Revilla, 2004). A detailed survey of the hedonic games is available in Aziz and Savani (2016); Aziz et al. (2019); Cechlárová and Hajduková (2004). In our work, we only use partial information and ϵ-PAC stable notion for the existence of partition $\tilde{\pi}$. However, our work is valid for any class of hedonic games as long as both noise-free and noisy values have the same representation.

Existence of solution concepts: Another line of literature focuses on the algorithmic aspects of solution concepts of hedonic games. Regarding solution concepts like core stability and nature of partitions, there are two questions: does there exist a partition π satisfying the solution concept's properties; if there is such a π, find one. To this end, for different classes of hedonic games, there are various algorithms and hardness results such as Sung and Dimitrov (2010); Rahwan et al. (2009); Woeginger (2013).

PAC learning in hedonic games: Uncertainty in the agents' preferences in the cooperative games has been carefully analyzed by Balcan et al. (2015). The authors used the PAC learning model to learn an underlying game. A new connection is established between PAC learnability and core stability for various classes of TU cooperative games. It turned out that only a few classes of TU games are learnable and stable. Sliwinski and Zick (2017) extended the PAC learning approach to the premise of hedonic games, where complete information about individual preferences is unavailable. We incorporate noise in the preferences and use PAC bounds to obtain the prediction probabilities.

6. Discussion and looking ahead

This work considers the noisy hedonic game with partial information on preferences. Given a PAC stable partition of the noisy game, we find the prediction probability that *unknown noise-free* game has PAC stable partition. This requires a combinatorial construct called agreement event and its probability. For $l \geq 2$ noise support, we obtain the agreement probability as a function of noise probabilities. For a user-given satisfaction value on agreement probability, we obtain the noise set such that a given partition is noise-robust. An interesting observation is that the prediction probability can be high for some high noise values with high probabilities. In particular, for a 2 agent game with 2 noise support, we obtain the noise set for which the prediction probability is more than a user-given satisfaction value. We have noise robustness for the entire noise probability simplex for the prediction probability below 0.62, i.e., the safety value. However, if the prediction probability function exceeds this safety value, the noise robust regime is non-convex. For the case of 3 noise support, finding a safety value is difficult as it is a global minimum of a non-convex prediction probability function. We obtain the bounds on the extra noisy samples required to get the PAC stable partition in a noise-free game. These extra samples are polynomial in the number of agents and the user-given satisfaction value on agreement probability.

The aspects we investigated offer many other rich possibilities; we mention some of them here. Firstly, since the prediction probability function for 3 support noise distribution is non-convex, which renders the computation of the fundamental limit of noise robustness hard, one may investigate suitable approximations. Another possibility is to consider other noise models where the value of each coalition is perturbed at the individual player level.

References

José Alcalde and Pablo Revilla. Researching with whom? Stability and manipulation. *Journal of Mathematical Economics*, 40(8):869–887, 2004.

Haris Aziz and Florian Brandl. Existence of stability in hedonic coalition formation games. In *Proceedings of the 11th AAMAS*, pages 763–770, 2012.

Haris Aziz and Rahul Savani. Hedonic Games (Chapter 15). In F. Brandt, V. Conitzer, J. Lang U. Endriss, and A.D. Procaccia, editors, *Handbook of Computational Social Choice*. Cambridge University Press, Cambridge, 2016.

Haris Aziz, Florian Brandl, Felix Brandt, Paul Harrenstein, Martin Olsen, and Dominik Peters. Fractional hedonic games. *ACM TEAC*, 7(2):1–29, 2019.

Maria-Florina Balcan, Ariel D Procaccia, and Yair Zick. Learning cooperative games. In *Proceedings of the 24th IJCAI*, pages 475–481, 2015.

Coralio Ballester. NP-completeness in hedonic games. *Games and Economic Behavior*, 49 (1):1–30, 2004.

Suryapratim Banerjee, Hideo Konishi, and Tayfun Sönmez. Core in a simple coalition formation game. *Social Choice and Welfare*, 18(1):135–153, 2001.

Anna Bogomolnaia and Matthew O Jackson. The stability of hedonic coalition structures. *Games and Economic Behavior*, 38(2):201–230, 2002.

Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D Procaccia. *Handbook of computational social choice*. Cambridge University Press, 2016.

Katarína Cechlárová and Jana Hajduková. Stability of partitions under $\mathcal{BW}$-prefernces and $\mathcal{WB}$- preferences. *International Journal of Information Technology & Decision Making*, 3(04):605–618, 2004.

Jacques H Dreze and Joseph Greenberg. Hedonic coalitions: Optimality and stability. *Econometrica (pre-1986)*, 48(4):987, 1980.

Edith Elkind and Michael J Wooldridge. Hedonic coalition nets. In *AAMAS (1)*, pages 417–424, 2009.

Andreu Mas-Colell, Michael Dennis Whinston, and Jerry R Green. *Microeconomic theory*, volume 1. Oxford University Press New York, 1995.

Yadati Narahari. *Game theory and mechanism design*, volume 4. World Scientific, 2014.

Talal Rahwan, Sarvapali D Ramchurn, Nicholas R Jennings, and Andrea Giovannucci. An anytime algorithm for optimal coalition structure generation. *Journal of Artificial Intelligence Research*, 34:521–567, 2009.

Jakub Sliwinski and Yair Zick. Learning Hedonic Games. In *IJCAI*, pages 2730–2736, 2017.

Shao-Chin Sung and Dinko Dimitrov. Computational complexity in additive hedonic games. *European Journal of Operational Research*, 203(3):635–639, 2010.

Gerhard J Woeginger. Core stability in hedonic coalition formation. In *International Conference on Current Trends in Theory and Practice of Computer Science*, pages 33–50. Springer, 2013.

Appendix A. Proof of Theorem 11

Proof We prove this via induction on noise support $l \geq 2$. The base case with $l = 2$ support is available in Lemma 1 of the SM. Let us assume that it is true for $l = k$, i.e., there are sets $\mathcal{I}(\alpha_r, \alpha_s, T) = \left\{ \tilde{\pi}(i) \in \mathcal{R}(T) \;\middle|\; \frac{\tilde{v}_i(\tilde{\pi}(i))}{\tilde{v}_i(T)} \geq \frac{\alpha_r}{\alpha_s} \right\}$, such that $\alpha_s < \alpha_r,\ \forall\ 1 \leq s < r \leq k$.

For this k we have $f_T(p_j, \alpha_j; j \in [k]) =: f_T(\boldsymbol{p}, \boldsymbol{\alpha})$ (by assumption), here $[k] = \{1, 2, \ldots, k\}$

$$f_T(\boldsymbol{p}, \boldsymbol{\alpha}) = \sum_{a=1}^{k} p_a \left(\sum_{b=1}^{a} p_b \right)^{|\mathcal{R}(T)|} + \sum_{r,s \in [k]: \alpha_r > \alpha_s} p_s^{|\mathcal{R}(T)| - |\mathcal{I}(\alpha_r, \alpha_s, T)| + 1} ((p_r + p_s)^{|\mathcal{I}(\alpha_r, \alpha_s, T)|} - p_s^{|\mathcal{I}(\alpha_r, \alpha_s, T)|}).$$

We will now show that this is true for $l = k + 1$. To this end, for all $s \in [k]$ such that for $\alpha_{k+1} > \alpha_s$ we define $\mathcal{I}(\alpha_{k+1}, \alpha_s, T) = \left\{ \tilde{\pi}(i) \in \mathcal{R}(T) \;\middle|\; \frac{\tilde{v}_i(\tilde{\pi}(i))}{\tilde{v}_i(T)} \geq \frac{\alpha_{k+1}}{\alpha_s} \right\}$. Now, there are two cases, $\mathcal{I}(\alpha_{k+1}, \alpha_s, T) = \emptyset,\ \forall\ \alpha_s,\ s \in [k]$, or $\mathcal{I}(\alpha_{k+1}, \alpha_s, T) \neq \emptyset$ for at least for one $s \in [k]$.

Case 01: $[\mathcal{I}(\alpha_{k+1}, \alpha_s, T) = \emptyset,\ \forall\ \alpha_s,\ s \in [k]]$. With one more element in noise support, apart from the existing $\{\alpha(\tilde{\pi}(i)\}_{\tilde{\pi}(i) \in \mathcal{R}(T)}$, and $\alpha(T)$ for k support case it will also have $\alpha(T) = \alpha_{k+1}$, and $\alpha(\tilde{\pi}(i)) \in \{\alpha_1, \alpha_2, \ldots, \alpha_{k+1}\},\ \forall\ \tilde{\pi}(i) \in \mathcal{R}(T)$. The probability of such α's is $p_{k+1} \left(\sum_{b=1}^{k+1} p_b \right)^{|\mathcal{R}(T)|}$. Therefore, the overall probability is

$$\sum_{a=1}^{k} p_a \left(\sum_{b=1}^{a} p_b \right)^{|\mathcal{R}(T)|} + p_{k+1} \left(\sum_{b=1}^{k+1} p_b \right)^{|\mathcal{R}(T)|} = \sum_{a=1}^{k+1} p_a \left(\sum_{b=1}^{a} p_b \right)^{|\mathcal{R}(T)|}.$$

Case 02: $[\mathcal{I}(\alpha_{k+1}, \alpha_s, T) \neq \emptyset$ for at least for one $s \in [k]]$. In this case, apart from the existing $\{\alpha(\tilde{\pi}(i))\}_{\tilde{\pi}(i) \in \mathcal{I}(\alpha_r, \alpha_s, T)}$, and $\alpha(T)$ for k support, we also have $\{\alpha(\tilde{\pi}(i))\}_{\tilde{\pi}(i) \in \mathcal{I}(\alpha_{k+1}, \alpha_s, T)}$, $\alpha(T)$ such that $\alpha(\tilde{\pi}(i)) = \alpha_s,\ \forall\ \tilde{\pi}(i) \in \mathcal{R}(T) \setminus \mathcal{I}(\alpha_{k+1}, \alpha_s, T)$, and $\alpha(T) = \alpha_{k+1}$. Thus, for $k + 1$ support the probability is:

$$\sum_{r,s \in [k]: \alpha_r > \alpha_s} p_s^{|\mathcal{R}(T)| - |\mathcal{I}(\alpha_r, \alpha_s, T)| + 1} ((p_r + p_s)^{|\mathcal{I}(\alpha_r, \alpha_s, T)|} - p_s^{|\mathcal{I}(\alpha_r, \alpha_s, T)|}) + p_s^{|\mathcal{R}(T)| - |\mathcal{I}(\alpha_{k+1}, \alpha_s, T)| + 1} ((p_{k+1} + p_s)^{|\mathcal{I}(\alpha_{k+1}, \alpha_s, T)|} - p_s^{|\mathcal{I}(\alpha_{k+1}, \alpha_s, T)|}).$$

From case 01 and case 02 above, for $k + 1$ support we have,

$$f_T(p_j, \alpha_j; j \in [k+1]) = \sum_{r,s \in [k+1]: \alpha_r > \alpha_s} p_s^{|\mathcal{R}(T)| - |\mathcal{I}(\alpha_r, \alpha_s, T)| + 1} \left((p_r + p_s)^{|\mathcal{I}(\alpha_r, \alpha_s, T)|} - p_s^{|\mathcal{I}(\alpha_r, \alpha_s, T)|} \right) + \sum_{a=1}^{k+1} p_a \left(\sum_{b=1}^{a} p_b \right)^{|\mathcal{R}(T)|}$$

Therefore, from the principle of Mathematical induction, this is true for any $l \geq 2$. ■