1-2018

Identifying and computing the exact core-determining class

Ye LUO

Hai WANG
Singapore Management University, haiwang@smu.edu.sg

Identifying and Computing the Exact Core-determining Class

Ye Luo and Hai Wang

Faculty of Business and Economics
University of Hong Kong
Hong Kong
e-mail: kurtluo@gmail.com

School of Information Systems
Singapore Management University
Singapore
e-mail: wanghaimit@gmail.com

Abstract: The indeterministic relations between unobservable events and observed outcomes in partially identified models can be characterized by a bipartite graph. Given a probability measure on observed outcomes, the set of feasible probability measures on unobservable events can be defined by a set of linear inequality constraints, according to Artstein's Theorem. This set of inequalities is called the "core-determining class". However, the number of inequalities defined by Artstein's Theorem is exponentially increasing with the number of unobservable events, and many inequalities may in fact be redundant. In this paper, we show that the "exact core-determining class", i.e., the smallest possible core-determining class, can be characterized by a set of combinatorial rules of the bipartite graph. We prove that if the bipartite graph and the measure on observed outcomes are non-degenerate, the exact core-determining class is unique and it only depends on the structure of the bipartite graph. We then propose an algorithm that explores the structure of the bipartite graph to construct the exact core-determining class. We design and implement the model and algorithm in a set of examples to show that our methodology could efficiently discard the redundant inequalities that are not useful to identify the parameter of interest. We also demonstrate that, by using the inequalities corresponding to the exact core-determining class to perform set inference, the power of test statistics against local alternatives can be improved.

Keywords and phrases: Core-determining Class, Inequality Selection, Linear Programming, Partially Identified Models, Set Inference.

1. Introduction

Correlations and causalities between "facts" are wide-spread. Suppose we define hidden facts as "unobservable events" and observable facts as "observed

*The authors gratefully thank all participants in MIT Econometrics seminar and the session Machine Learning in Econometrics at the AEA 2017 annual meeting in Chicago for invaluable comments. We thank Victor Chernozhukov, Alfred Galichon, Adam Rosen, Amedeo Odoni and Xiaoxia Shi for providing valuable comments.

outcomes"; in reality, we could observe the outcomes, but not the events that cause the outcomes. In many situations, relations between events and outcomes are indeterministic—i.e., a single event may lead to different outcomes, and an outcome may be the result of several events. It is important to infer information about unobservable events by using information about observed outcomes. One example, from game theory, is to infer individual players' private information (unobservable events), given observations of their strategies, under multiple Nash equilibria (observed outcomes); another is to infer customers' hidden characteristics (unobservable events) given their purchase histories and sales data (observed outcomes).

The relations between unobservable events and observed outcomes can be characterized by a bipartite graph. Given a bipartite graph and an observed measure on outcomes, our paper's key interest is to estimate bounds on the probability measure on unobservable events. The feasible set of probability measures on unobservable events is defined by a set of linear inequality constraints. In practice, the number of inequality constraints needed from a standard method (e.g., Artstein's Theorem, which is introduced in Section 2) could grow very quickly, even exponentially with the number of unobservable events, while some or even many of them are redundant. However, including many inequalities and especially many redundant inequalities may impose two difficulties in inference on the measure on unobservable events: (1) for the inference procedures, such as those described by Andrews and Shi (2014) and others may lose power; and (2) these inference procedures are computationally intractable when the number of unobservable events is too large.

In this paper, we propose an algorithm (an inequality selection procedure) that may dramatically reduce the number of inequalities that define the feasible set of probability measures on unobservable events. A subset of all inequalities described in Artstein's Theorem that characterizes the identified region is referred to as the core-determining class by Galichon and Henry (2011). Obviously, there can be more than one core-determining class. The goal of our algorithm is to find the smallest possible core-determining class, which we formally define as the "exact core-determining class". Such an algorithm for constructing the exact core-determining class allows us to identify which inequalities are truely binding. We show that the "exact core-determining class" only depends on the structure of the bipartite graph, not the probability measure on the observed outcomes. This result is important because usually the inequalities defined in Artstein's Theorem are estimated with noise, which seems to affect inference. However, our results show that the identity of binding inequalities in the core-determining class would not be affected by the noise. More specifically, (1) we show that, the smallest set of irredundant inequalities, denoted as "exact core-determining class", can be characterized by a set of combinatorial rules. We prove that, under certain mild conditions, the exact core-determining class only depends on the structure of the bipartite graph, not on the observed probability measure on the outcomes. (2) we then propose an algorithm for constructing the exact core-determining class.

The closest studies on our topic are by Galichon and Henry (2006, 2011), Chesher and Rosen (2012), Andrews and Soares (2010), etc. Specifically, Galichon and Henry (2006, 2011) propose the core-determining class problem, i.e., finding the minimum set of inequalities to describe the feasible region of probability measures on the unobservable events. Chesher and Rosen (2012) provide an inequality selection algorithm, but may still contain some redundant inequalities in its selected set. Andrews and Soares (2010) propose a moment inequality selection procedure using criterions such as BIC.

There are many studies on performing inference of partially identified models. Chernozhukov, Hong and Tamer (2007) propose a general inference procedure with moment inequality constraints. Romano and Shaikh (2010) provide improvements for CHT (2007). Beresteanu, Molchanov, and Molinari (2011) use random set theory to perform inference with convex inequality restrictions. Andrews and Shi (2014) construct set inference in conditional moment inequalities settings. In the many moment inequality environment, Chernozhukov, Chetverikov and Kato (2014) proposes tests statistics based on fixed critical value and bootstrap critical value for testing whether one point is in the identified region or not. Andrews and Shi (2016) proposes tests for identified region in the setting of many unconditional moment inequalities. Other procedures include Romano and Shaikh (2010), etc.. For all of these method metioned, the computation of the test-statistics requires us to use of all moment inequalities, regardless of their redundancy. Though this paper is not about performing inference on partially identified models, by selecting the set of "irredundant" inequalities and then utilizing the selected inequalities for inference, practitioners could potentially reduce the computational cost when performing these inference methods. We implement the proposed algorithm and demonstrate its good performance through two large scale examples. In these examples, we find substantial reduction of the inequalities and hence, reduction in computational time when performing inference procedures using the inequalities that correspond to the "exact core-determining class". In addition, there are improvements of power of tests compared to local alternatives, by only using the reduced set of inequalities. For related empirical studies, see Manski and Tamer (2002); Bajari, Benkard, and Levin (2007); Bajari, Hong, and Ryan (2010); etc.

There is also a wide literature on detection and elimination of redundant constraints. For example, Telgen (1983) develops two methods to identify redundant constraints and implicit equalities. Caron, McDonald, and Ponic (1989) present a degenerate extreme point strategy that classifies linear constraints as either redundant or necessary. Paulraj, Chellappan, and Natesan (2006) propose a heuristic approach using an intercept matrix to identify redundant constraints. Since these papers do not look into irredundant constraints in the set of linear inequalities defined by Artstein's theorem, none of them focuses on the special structure of the core-determining class studied in our paper,

The paper is organized as follows: In Section 2, we introduce the definition of "exact core-determining class" and preview of main results. In Section 3, we study the relationship between the structure of the bipartite graph and identity of irredundant inequalities. We prove that the irredundant inequalities that characterize the identified region can be described by a set of combinatorial rules and then propose a fast algorithm for computing the exact core-determining class. In Section 4, we implement the algorithm in a set of examples with large bipartite graphs and demonstrate its good performance. In Section 5, we evaluate the performance of inference procedures using the exact core-determining class, compared to the performance of the same procedures using the full set of inequalities. Section 6 concludes.

2. Exact Core-determining Class and Preview of Results

2.1. core-determining Class and Exact core-determining Class

The relations between unobservable events and observed outcomes can be characterized by a bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$, where $\mathcal{U}$ is a finite set of vertices representing unobservable events, $\mathcal{Y}$ is a finite set of vertices representing observed outcomes, and $\varphi : 2^{\mathcal{U}} \mapsto 2^{\mathcal{Y}}$ is a correspondence mapping from $\mathcal{U}$ to $\mathcal{Y}$ such that $\varphi(u) \subset \mathcal{Y}$ is the set of all possible outcomes that could be caused by the event $u \in \mathcal{U}$, and $\varphi(A) := \cup_{u \in A} \varphi(u)$ for any set $A \subset \mathcal{U}$. The inverse of φ, denoted as φ^{-1} is defined as $\varphi^{-1} : 2^{\mathcal{Y}} \mapsto 2^{\mathcal{U}}$, $\varphi^{-1}(B) = \{u \in \mathcal{U} | \varphi(u) \cap B \neq \emptyset\}$, $\forall B \subset \mathcal{Y}$. Let v be the probability measure on $\mathcal{U}$ and μ be the probability measure on $\mathcal{Y}$. Essentially, we aim to identify the set of feasible probability measures v on $\mathcal{U}$ given an observed measure μ on $\mathcal{Y}$.

For a graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$, we say G is connected if (1) $\forall A_1, A_2 \subset \mathcal{U}$ such that $A_1 \cap A_2 = \emptyset$ and $A_1 \cup A_2 = \mathcal{U}$, it holds that $\varphi(A_1) \cap \varphi(A_2) \neq \emptyset$, and (2) the$\varphi(\mathcal{U}) = \mathcal{Y}$. For any non-empty set $A \subset \mathcal{U}$, we say A is self-connected (in the graph G) if the subgraph $(A, \varphi(A), \varphi)$ is connected.

Definition 1 (Non-Degeneracy of G and μ)**.** (1) We say that G is non-degenerate if G is a connected graph.

(2) We say that μ is non-degenerate if $\mu(y) > 0$ for any $y \in \mathcal{Y}$.

The non-degeneracy of the graph G can break down when G have more than one mutually disconnected branches, in which we call G is degenerate. Later we will discuss both the situations when G is non-degenerate and when G is degenerate. The non-degeneracy of μ is a very mild condition and we assume that it holds through out the paper.

Denote $d_u = |\mathcal{U}|$ and $d_y = |\mathcal{Y}|$. The parameter of interest in this paper is the $d_u \times 1$ vector v, which is the probability measure that generates the events $u \in \mathcal{U}$. In general, we are unable to obtain a point estimation of v given only the bipartite graph and the observed measure μ on $\mathcal{Y}$, unless additional information is provided. Instead, we can obtain a set of inequalities on v given the bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$ and the measure μ on $\mathcal{Y}$. More specifically, for any set of events $A \subset \mathcal{U}$, the outcome should fall into the set $\varphi(A)$. Thus, for any $A \subset \mathcal{U}$, we can obtain an inequality $v(A) := \sum_{u \in A} v(u) \leqslant \mu(\varphi(A)) := \sum_{y \in \varphi(A)} \mu(y)$.

Artstein's theorem in Artstein (1983) states that all information of v in the bipartite graph model $G = (\mathcal{U}, \mathcal{Y}, \varphi)$, given a probability measure μ on $\mathcal{Y}$, is characterized by the set of constraints described below.

Lemma 1 (Artstein's Theorem)**.** *The following set of inequalities and equalities contains all information on $v \geqslant 0$:*

1. For any $A \subset \mathcal{U}$,

$$v(A) := \sum_{u \in A} v(u) \leqslant \mu(\varphi(A)) \tag{2.1}$$

where $\mu(\varphi(A)) := \sum_{y \in \varphi(A)} \mu(y)$;

2.

$$\sum_{u \in \mathcal{U}} v(u) = \sum_{y \in \mathcal{Y}} \mu(y). \tag{2.2}$$

By Artstein's Theorem, we can define the identified region Q of the parameter of interest, v, as follows:

Definition 2 (Identified Region Q)**.** *The set of all feasible probability measures v on $\mathcal{U}$ with bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$ and probability measure μ on $\mathcal{Y}$ such that:*

(1) For any $A \subset \mathcal{U}$, $v(A) \leqslant \mu(\varphi(A))$;

(2) $\sum_{u \in \mathcal{U}} v(u) = 1$.

Or equivalently, $Q := \{v \geqslant 0 | \sum_{u \in \mathcal{U}} v(u) = 1, v(A) \leqslant \mu(\varphi(A))$ for all $A \subset \mathcal{U}, A \neq \emptyset\}$.

The number of inequalities defined by Artstein's Theorem is $2^{d_u} - 2$, which could be an extremely large number when d_u is large. In general, many of the inequalities in Definition 2 are redundant. For any subset $A \subset \mathcal{U}$ such that $A \neq \mathcal{U}$ and $A \neq \emptyset$, we say the inequality $v(A) \leqslant \mu(\varphi(A))$ is the inequality corresponding to A. Consider $\mathcal{S}$ as a collection of non-empty subsets $A \subset \mathcal{U}, A \neq \mathcal{U}$. We define $Q(\mathcal{S}) := \{v \geqslant 0 | \sum_{u \in \mathcal{U}} v(u) = 1, v(A) \leqslant \mu(\varphi(A))$ for all $A \in \mathcal{S}\}$. If $Q(\mathcal{S}) = Q$ as defined in Definition 2, to identify the region Q, it is suffice to drop all other inequalities that correspond to subsets of $\mathcal{U}$ which are not in $\mathcal{S}$. In other words, if we are able to find a collection of subsets $\mathcal{S}$ such that $Q(\mathcal{S}) = Q$ with $|\mathcal{S}|$ being much smaller than $2^{d_u} - 1$, we will be able to describe the identified region Q with much less number of inequality constraints. Galichon and Henry (2011) propose the concept of the core-determining class, which fits into the idea of inequality selection. Below, we define our definition of ("$\mathcal{U}$") core-determining class.

Definition 3 (("$\mathcal{U}$") Core-determining Class)**.** *A subset $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{U\}$ is a core-determining class if $\{v \geqslant 0 | v(A) \leqslant \mu(A)$ for all $A \in \mathcal{S}, v(\mathcal{U}) = \mu(\mathcal{Y})\}$ is identical to Q.*

In Galichon and Henry (2011), the authors define the core-determining class in a slightly different way: A class of subsets $\mathcal{S}'$ of $\mathcal{Y}$ is core-determining if $\{v \geqslant 0 | v(\varphi^{-1}(B)) \geqslant \mu(B)$ for all $B \in \mathcal{S}', v(\mathcal{U}) = \mu(\mathcal{Y})\} = Q$. In the biparitite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$, since the events in A would lead to outcome $y \in \varphi(A)$, we have a constraint $v(A) \leqslant \mu(\varphi(A))$. However, if we consider a graph $G^{-1} = (\mathcal{Y}, \mathcal{U}, \varphi^{-1})$, any set of outcomes B implies that the event u must be in the set $\varphi^{-1}(B)$, and therefore resulting in a constraint $v(\varphi^{-1}(B)) \geqslant \mu(B)$. The two different ways characterize the same information on v, so the identified region Q remains unchanged regardless of considering G or G^{-1}.

We call the core-determining class defined in Galichon and Henry as "$\mathcal{Y}$ core-determining class" and core-determining class defined in this paper as "$\mathcal{U}$ core-determining class". Later in Comment 2, we point out that there is a very simple one-to-one relationship between the smallest "$\mathcal{U}$ core-determining class" and the smallest "$\mathcal{Y}$ core-determining class". In Definition 4, we define the smallest "$\mathcal{U}$ core-determining class" as the "exact core-determining class". Therefore, for simplicity, we abbreviate the "$\mathcal{U}$ core-determining class" as "core-determining class" in the rest of this paper.

For any collection of sets $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{\emptyset\}$, we can define $Q(\mathcal{S}) := \{v \geqslant 0 | v(\mathcal{U}) = \mu(\mathcal{Y}), v(A) \leqslant \mu(\varphi(A))$ for all $A \in \mathcal{S}\}$.

By Definition 3, for any core-determining class $\mathcal{S}$, it must be that $Q(\mathcal{S}) = Q = Q(2^{\mathcal{U}} \backslash \{\emptyset\})$, which is the identified region defined by all possible inequalities described in the Artstein's Theorem, along with the equality $v(\mathcal{U}) = 1 = \mu(\mathcal{Y})$. By Definition 3, any collection of inequalities of the Artstein's theorem, amid the equality that $v(\mathcal{U}) = \mu(\mathcal{Y})$, contain the full information if and only if the collections of the corresponding subsets is a core-determining class.

In many situations, a core-determining Class could be way smaller than the full set of non-empty subsets of $\mathcal{U}$. The motivation of finding a Core Determing Class are two folds: first, the full set of inequalities may be too many to effectively perform estimation and inference procedures in affordable time; second, as we demonstrate in Section 5, using redundant inequalities may lead to loss of power when testing whether one point belongs to the identified region Q. Therefore, it is worth to find the smallest core-determining class, defined as "exact core-determining class" below, in order to enhance performance of existing set inference procedures in both computational aspect and statistical aspect.

Definition 4 (Exact Core-determining Class)**.** *We say a collection of non-empty subsets of $\mathcal{U}$, denoted as $\mathcal{S}$, is an exact core-determining class if and only if $\mathcal{S} \in argmin_{\mathcal{S}' \subset 2^{\mathcal{U}}, Q(\mathcal{S}')=Q} |\mathcal{S}'|$, i.e., $\mathcal{S}$ is the smallest core-determining class such that, along with the equality $v(\mathcal{U}) = \mu(\mathcal{Y})$, they together characterize the identified region Q described by (2.1) and (2.2).*

2.2. Preview of Main Results

There is a list of papers that discuss computing a core-determining class, e.g., Galichon and Henry (2009), Chesher and Rosen (2012) , etc.. To our best knowledge, there is no existing results that, in general, characterize the Exact core-determining Class, nor there exist any algorithms that compute the Exact core-determining Class.

In this paper, we characterize the property of Exact core-determining Class using a small set of combinatorial rules. With these rules, one can easily tell whether a subset of $\mathcal{U}$ is exact core-determining or not. With a series of lemmas and propositions in the Appendix, when the graph G and the measure μ are non-degenerate, we prove that such characterization is necessary and sufficient.

To be more specific, we show that the exact core-determining class is unique and it consists all the subset $A \subset \mathcal{U}$ such that the following properties hold:

(1) A is self-connected, i.e., $\forall A_1, A_2 \subset A$ such that $A_1, A_2 \neq \emptyset$ and $A_1 \cup A_2 = A$, it holds that $\varphi(A_1) \cap \varphi(A_2) \neq \emptyset$;

(2) There exists no $u \in \mathcal{U}$ such that $u \notin A$ and $\varphi(u) \subset \varphi(A)$.

(3) $\varphi(A)^c$ is non-empty and self-connected in graph G^{-1}.

The conditions required for the results above are mild. In practice, the μ is often considered as non-degenerate; otherwise, we could simple delete the outcome states that corresponds to 0 probability in $\mathcal{Y}$. Connectedness of G is important, and it garantees that the exact core-determining class is unique. The results said, in the core-determining class problem, the identity of redundant inequality does not depend on the probability measure μ, but depends on the structure of the graph G only. That means, given a non-degenerate measure $\widehat{\mu}$ as an estimator of a non-degenerate true measure μ_0, the set of redundant inequalities is exactly the same as the set of redundant inequalities given μ_0.

When the graph G is not connected, it could be written as a set of disjoint subgraphs that are connected by themselves. These subgraphs are often referred as connected branches of G. We can decompose G as $G_j = (\mathcal{U}_j, \mathcal{Y}_j, \varphi)$, $j = 1, 2, ..., k$, $k \geqslant 2$, where G_j is connected, $\mathcal{U}_j$, $j = 1, 2, ..., k$ form a partition of $\mathcal{U}$ and $\mathcal{Y}_j = \varphi(\mathcal{U}_j)$, $j = 1, 2, ..., k$, form a partition of $\mathcal{Y}$. In this case, the exact core-determining class could be non-unique, as we demonstrate later in Example 3. We propose a specific way to construct one of the exact core-determining classes:

Denote $\mathcal{S}_j^*$ as the exact core-determining class in the subgraph G_j, $j = 1, 2, ..., k$. Then $\mathcal{S}^* := \cup_{j=1}^k (\mathcal{S}_j^* \cup \{\mathcal{U}_j\})$ is an exact core-determining class with resepect to G, i.e., a class of subsets with the minimum possible cardinalty that is core-determining.

When the graph G is connected, we propose an algorithm that computes the exact core-determining class using the combinatorial properties above, which are formally stated in Theorem 2 in Section 3.3. The algorithm is presented in Algorithm 1. When the graph G is non-connected, we could apply Algorithm 1 on each connected branch G_j, $j = 1, 2, ..., k$, and then construct an exact core-determining class accordingly.

2.3. Examples

There exist many economic examples that lead to a partitially identified model characterized by the inequalities/equality in (2.1) and (2.2). One of the leading examples is the the two players entry game, as discussed in Galichon and Henry (2011).

Example 1. [Two Players Entry Game] *Suppose there are two firms, firm 1 and firm 2, in a market. The costs for firm 1 and firm 2 are $c + r_1$ and $c + r_2$, respectively, where c is a constant and r_1 and r_2 are random shocks that are observable only by the corresponding firm (and hence are unobservable to the public).*

The two firms face a total demand $D = a_1 - a_2 p$, where a_1 and a_2 are known parameters and p is the price of their product. If both firms (players) are in the market, they will play a Cournot Nash equilibrium. If there is only one firm (player), this firm will reach a monopolist's equilibrium. If the costs are so high for both firms (players) that even a monopolist is unprofitable, there will be no firm in the market. Therefore, there are 4 possible equilibria: $(0,0)$, $(1,0)$, $(0,1)$, and $(1,1)$:

(1) if $\frac{a_1}{a_2} - c \geqslant 2/3r_1 - 1/3r_2$ and $\frac{a_1}{a_2} - c \geqslant 2/3r_2 - 1/3r_1$, the equilibrium is $(1,1)$;

(2) if $\frac{a_1}{a_2} - c < 2/3r_1 - 1/3r_2$ and $\frac{a_1}{a_2} - c \geqslant 2/3r_2 - 1/3r_1$, the equilibrium is $(0,1)$;

(3) if $\frac{a_1}{a_2} - c \geqslant 2/3r_1 - 1/3r_2$ and $\frac{a_1}{a_2} - c < 2/3r_2 - 1/3r_1$, the equilibrium is $(1,0)$;

Otherwise, if $\frac{a_1}{a_2} - c < 2/3r_1 - 1/3r_2$ and $\frac{a_1}{a_2} - c < 2/3r_2 - 1/3r_1$:

(4) if $c + r_1 \leqslant \frac{a_1}{a_2}$ and $c + r_2 \leqslant \frac{a_1}{a_2}$, there are two equilibria: $(1,0)$ and $(0,1)$;

(5) if $c + r_1 \leqslant \frac{a_1}{a_2}$ and $c + r_2 > \frac{a_1}{a_2}$, the equilibrium is $(1,0)$;

(6) if $c + r_1 > \frac{a_1}{a_2}$ and $c + r_2 \leqslant \frac{a_1}{a_2}$, the equilibrium is $(0,1)$;

(7) if $c + r_1 > \frac{a_1}{a_2}$ and $c + r_2 > \frac{a_1}{a_2}$, the equilibrium is $(0,0)$.

Let $\mathcal{U} = \{u_1, u_2, u_3, u_4, u_7\}$, where u_i is the event representing case (i), with the exceptions that u_2 represents cases (2) and (6), and u_3 represents cases (3) and (5). Let $Y := \{y_1, y_2, y_3, y_4\}$, where $y_1 = (1,1)$, $y_2 = (0,1)$, $y_3 = (1,0)$, and $y_4 = (0,0)$. We have $d_u = |\mathcal{U}| = 5$ and $d_y = |\mathcal{Y}| = 4$. The correspondence mapping φ between $\mathcal{U}$ and $\mathcal{Y}$ is:

$\varphi(u_1) = \{y_1\}$, $\varphi(u_2) = \{y_2\}$, $\varphi(u_3) = \{y_3\}$, $\varphi(u_4) = \{y_2, y_3\}$, and $\varphi(u_7) = \{y_4\}$.

The correspondence mapping for Example 1 is illustrated in Figure 1.

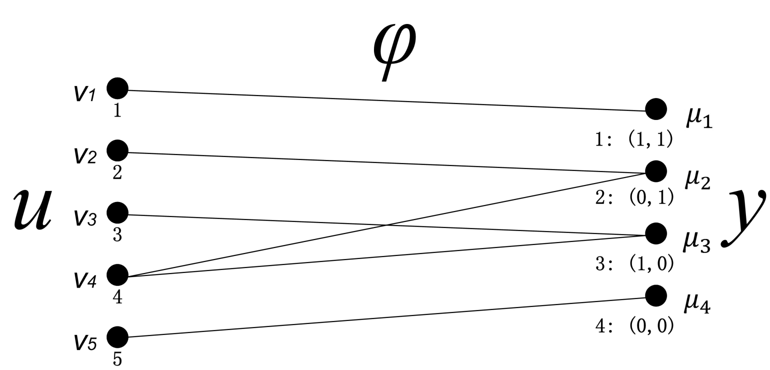

FIG 1. *Correspondence Mapping for Example 1.*

Given an observed probability measure μ on $\mathcal{Y}$, the bounds of probability measures v on $\mathcal{U}$ are given by the inequalities stated in Artstein's Theorem. According to Artstein's Theorem statement 2.1, there are $2^5 - 2 = 30$ inequalities. However, *it is obvious that only the inequalities corresponding to the following 5 subsets are irredundant:* $\{u_1\}$, $\{u_2\}$, $\{u_3\}$, $\{u_2, u_3, u_4\}$ *and* $\{u_5\}$.

In many cases, there exists a parametric model for v, with $v(u_i) = F_i(\theta, Z)$, $i = 1, 2, ..., d_u$. The function F_i can be a nonlinear transformation of θ and the observable characteristic Z, and it is usually derived from random errors with a jointly continuous probability distribution function. The parameter of interest is θ, not the v_i. Therefore, the identifies set of parameters is characterized as:

$\Theta := \{\theta | v(A) \leqslant \mu_{Y|Z}(\varphi(A)),\text{ for any } A \subset \mathcal{U}\},$, where $v(A) = \sum_{i\in A} F_i(\theta, Z)$, and $\mu_{Y|Z}$ is the probability measure of $Y|Z$.

Computationally, it is worth to divide the estimation and inference on Θ into two steps.

(1) First, find a core-determining class for the inequality system:

$v(A) \geqslant 0, v(A) \leqslant \mu_{Y|Z}(\varphi(A))$, for all $A \subset \mathcal{U}$, and $v(\mathcal{U}) = \mu(\mathcal{Y}) = 1$,

To the best situation, we can try to find the smallest core-determining class, which is later defined as the "exact core-determining class" in Definition 4. Denote this collection of subsets in $\mathcal{U}$ as $\mathcal{S}_Z^*$. The step (1) can be further simplified in many applications since $\mathcal{S}_Z^*$ does depend on Z as shown in Theorem 2.

(2) Perform any estimation or inference procedure on the set of inequalities: $\sum_{u_i\in A} F_i(\theta, Z) \leqslant \mu_{Y|Z}(\varphi(A))$ for any $A \in \mathcal{S}_Z^*$ and $Z \in \mathcal{Z}$.

Example 2. [Simultaneous Equation Models for Binary Outcomes] This example is previously considered in Chesher and Rosen (2012). The study of this model can be refered to Heckman (1978), Bresnahan and Reiss (1990, 1991), Tamer (2003) and etc. The model is described by the following equations: $Y_1 = 1(Z_1\beta_1 + Y_2\delta_2 + U_1 > 0)$, $Y_2 = 1(Z_2\beta_2 + Y_1\delta_1 + U_2 > 0)$, where $Y = (Y_1, Y_2) \in \{0, 1\}^2$, (Y, Z) are observables, and $U = (U_1, U_2)$ is an observable 2-dimension vector in $\mathbb{R}^2$. In the previous literature, the vector U is often assumed to have a multi-dimensional normal distribution with mean zero and an unknown covariance matrix Σ, with the variance of U_1 being restricted to 1, i.e., $\Sigma_{11} = 1$. The parameter of interest is $\theta = (\beta_1, \beta_2, \delta_1, \delta_2, \Sigma)$, where $\Sigma_{11} = 1$ and $\Sigma_{12} = \Sigma_{21}$.

To formulate the problem into our framework, we can define $\mathcal{Y} := \{y_1 := (0, 0), y_2 := (1, 0), y_3 := (0, 1), y_4 := (1, 1)\}$. We can also define a collection of measurable sets in $\mathbb{R}^2$ which may lead to the events in $\mathcal{Y}$. Then, we name these sets as the set of events, denoted as $\mathcal{U}$. The likelihood of the events in $\mathcal{U}$ depends on the parameter θ.

(a) If $\delta_1 > 0, \delta_2 > 0$, we can define the following 5 events on the space of $\mathcal{R}^2$:

$u_1 := \{U | U \in (-z_1\beta_1 - \delta_1, -z_1\beta_1] \times (-z_2\beta_2 - \delta_2, -z_2\beta_2]\}$. In this case, both $(0, 0)$ and $(1, 1)$ are possible outcomes, so $\varphi(u_1) = \{y_1, y_4\}$.

$u_2 := \{U | U_1 \leqslant -z_1\beta_1 - \delta_1, U_2 \leqslant -z_2\beta_2 - \delta_2\} \backslash u_1$. The only possible outcome of u_2 is $(0, 0)$, so $\varphi(u_2) = \{y_1\}$.

$u_3 := \{U | U_1 > -z_1\beta_1, U_2 \leqslant -z_2\beta_2 - \delta_2\}$. The only possible outcome of u_3 is $(1, 0)$, so $\varphi(u_3) = \{y_2\}$.

$u_4 := \{U | U_1 \leqslant -z_1\beta_1 - \delta_1, U_2 > -z_2\beta_2\}$. The only possible outcome of u_2 is $(0, 1)$, so $\varphi(u_4) = \{y_3\}$.

$u_5 := \{U|U_1 > -z_1\beta_1 - \delta_1, U_2 > -z_2\beta_2 - \delta_2\}\backslash u_1$. The only possible outcome of u_5 is $(1,1)$, so $\varphi(u_5) = \{y_4\}$.

The correspondence mapping for Example 2 case (a) is illustrated in Figure 2.

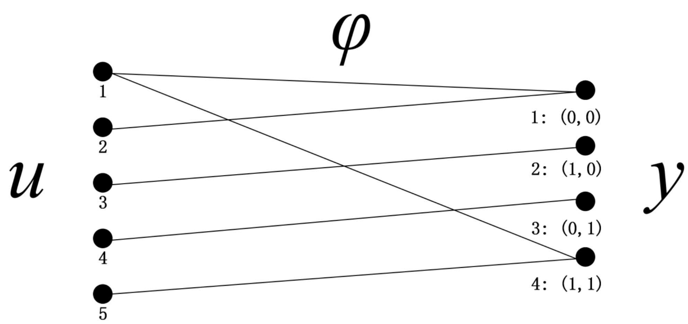

FIG 2. *Correspondence Mapping for Example 2 Case (a).*

(b) If $\delta_1 < 0, \delta_2 < 0$, we can define the following 5 events on the space of $\mathcal{R}^2$:

$u_1 := \{U|U \in (-z_1\beta_1, -z_1\beta_1 - \delta_1] \times (-z_2\beta_2 - \delta_2, -z_2\beta_2]\}$. In this case, both $(0,1)$ and $(1,0)$ are possible outcomes, so $\varphi(u_1) = \{y_2, y_3\}$.

$u_2 := \{U|U_1 \leqslant -z_1\beta_1, U_2 \leqslant -z_2\beta_2\}$. The only possible outcome of u_2 is $(0,0)$, so $\varphi(u_2) = \{y_1\}$.

$u_3 := \{U|U_1 > -z_1\beta_1, U_2 \leqslant -z_2\beta_2 - \delta_2\}\backslash u_1$. The only possible outcome of u_3 is $(1,0)$, so $\varphi(u_3) = \{y_2\}$.

$u_4 := \{U|U_1 \leqslant -z_1\beta_1 - \delta_1, U_2 \geqslant -z_2\beta_2\}\backslash u_1$. The only possible outcome of u_2 is $(0,1)$, so $\varphi(u_4) = \{y_3\}$.

$u_5 := \{U|U_1 > -z_1\beta_1 - \delta_1, U_2 > -z_2\beta_2 - \delta_2\}$. The only possible outcome of u_5 is $(1,1)$, so $\varphi(u_5) = \{y_4\}$.

The correspondence mapping for Example 2 case (b) is illustrated in Figure 3.

(c) $\delta_1 > 0, \delta_2 < 0$ or $\delta_1 < 0, \delta_2 > 0$.

Without loss of generality, we can assume that $\delta_1 > 0, \delta_2 < 0$. There is an event $u_1 := \{U|U_1 \in (-z_1\beta_1 - \delta_1, -z_1\beta_1] \times (-z_2\beta_2, -z_2\beta_2 - \delta_2]\}$ which leads to no feasible value of $(Y_1, Y_2) \in \{0,1\}^2$, i.e., the set of values in $\{0,1\}^2$ that corresponds to u_1 is $\emptyset$. Chesher and Rosen (2012) discusses a few different strategies to deal with such a situation, which they refer as the "incoherent" problem.

One of the methodologies is the "anything goes" approach, which is first considered in Beresteanu, Molchanov, and Molinari (2011). In such approach, correspondence mapping φ is constructed such that $\varphi(u_1) = \mathcal{Y}$, i.e., all values of $y \in \mathcal{Y}$ are allowed.

We can then define the rest of the events as follows:

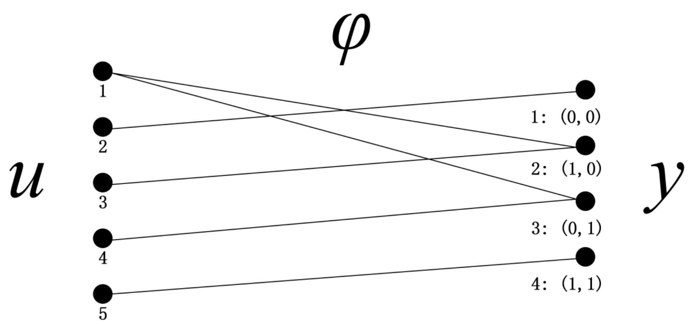

FIG 3. *Correspondence Mapping for Example 2 Case (b).*

$u_2 := \{U|U_1 \leqslant z_1\beta_1 - \delta_1, U_2 \leqslant z_2\beta_2\}$. The only possible outcome for u_2 is $(0,0)$, so $\varphi(u_2) = \{y_1\}$.

$u_3 := \{U|U_1 > z_1\beta_1, U_2 \leqslant z_2\beta_2 - \delta_2\}$. The only possible outcome for u_3 is $(1,0)$, so $\varphi(u_3) = \{y_2\}$.

$u_4 := \{U|U_1 \leqslant -z_1\beta_1 - \delta_1, U_2 > z_2\beta_2\}$. The only possible outcome for u_4 is $(0,1)$, so $\varphi(u_4) = \{y_3\}$.

$u_5 := \{U|U_1 > -z_1\beta_1, U_2 > z_2\beta_2 - \delta_2\}$. The only possible outcome for u_5 is $(1,1)$, so $\varphi(u_5) = \{y_4\}$.

The correspondence mapping for Example 2 case (c) is illustrated in Figure 4.

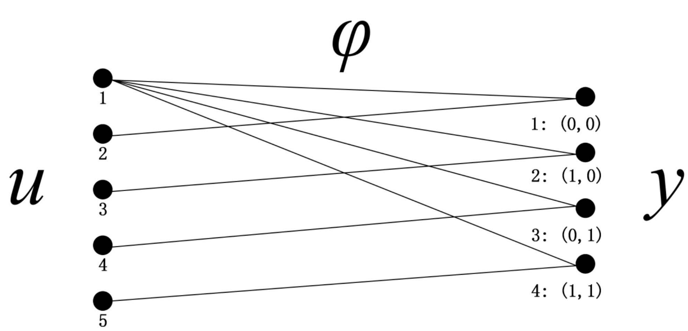

FIG 4. *Correspondence Mapping for Example 2 Case (c).*

In cases (a), (b) or (c), we can always find a correspondence mapping φ that represents the relationship between the events and outcomes. Although the random error $U \in \mathbb{R}^2$ has a continuous probability density funciton, we can divide the space $\mathbb{R}^2$ into a partition, namely $u_1, u_2, ..., u_5$, with the probability of each u_i, $i = 1, 2, ..., 5$, is represented by a function $F_i(Z, \theta)$.

The graph G in the examples above are all connected. In general, when G is unconnected, the exact core-determining class, e.g., the set of inequalities that characterizes the identified region with smallest possible cardinality, is non-unique. We provide an example in Example 3.

Example 3. Consider a bipartite graph G with set $\mathcal{U} = \{u_1, u_2, u_3\}$ and set $\mathcal{Y} = \{y_1, y_2, y_3\}$. The correspondence mapping φ between $\mathcal{U}$ and $\mathcal{Y}$ is illustrated in Figure 5. $\varphi(u_1) = \{y_1\}$, $\varphi(u_2) = \{y_2\}$, and $\varphi(u_3) = \{y_2, y_3\}$. Apparently, G is unconnected. We can easily check $\mathcal{S} = \{\{u_1\}, \{u_2\}, \{u_2, u_3\}\}$ is an exact core-determning class, and $\mathcal{S} = \{\{u_1\}, \{u_1, u_2\}, \{u_2, u_3\}\}$ is also an exact core-determing class.

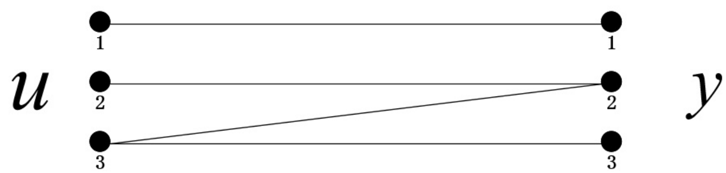

FIG 5. *Correspondence Mapping of Example 3.*

Large scale examples are provided in Section 4.

2.4. Computation, Inference and Power of Tests

According to Artstein's Theorem, the identified region Q is described by $2^{d_u} - 2$ inequalities—a very large number that exponentially increases with d_u. The numerous inequalities lead to both computational difficulties and undesirable statistical properties; some or even most of the inequalities stated in Artstein's Theorem may be redundant.

Galichon and Henry (2011) analyze the monotonic structure of the graph G and claim that there are at most $2d_u - 2$ irredundant inequalities under a special structure. Chesher and Rosen (2012) provide an algorithm that could get rid of some, but not necessarily all, redundant inequalities.

One good feature of the exact core-determining class is that the computational costs when performing inference and other tasks can be significantly reduced. We propose a combinatorial algorithm in Subsection 3.4 to compute the exact core-determining class. In Section 4, we demonstrate that the algorithm could perform very effectly inequality selection in the core-determining class problems. The computational time of finding the exact core-determining class takes less than 1 second in the large examples in Section 4 and can be ingored compared to the time cost of any state of art inference procedures. In Section 5, we demonstrate that, by using the exact core-determining class to perform inference, one could reduce the time of performing inference procedures substantially.

Another interesting property of utilizing the exact core-determining class to perform set inference is that the power of tests against local alternatives can be improved. We perform the inference procedure in Andrews and Shi (2014) on Example 5. The methodologies and results are formally stated in Section 5. We consider the two test statistics listed in Andrews and Shi (2014), denoted as S_1 and S_3. These test statistics are calculated using either the set of inequalities that correspond to the exact core-determining class, or the full set of inequalities stated in the Artstein's Theorem. Starting with a randomly selected point on the boundary of the identified region, we use simulation to obtain the power of tests on the local alternatives that lies on a radial originating from the selected point. The results show that the power of tests using the inequalities corresponding to the exact core-determining class is better than the power of tests using the full set of inequalities.

A quick view of results are demonstrated in Figure 6. For more detailed discussion, we refer to Section 5.

3. Characterization of the Exact Core-determining Class

In this section, we present our discovery of the combinatorial structure of the core-determining class, along with an algorithm to generate the exact core-determining class, which we will define formally later. In Galichon and Henry (2011), whether an inequality is redundant is determined by numerical computation using the probability measure μ. This is also true in Chesher and Rosen (2012). Our results in this section show that under mild conditions, the measure μ does not affect the identity of redundant inequalities.

3.1. Characterizing Redundancy of Inequalities

Given the correspondence mapping φ of the bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$, we can identify some redundant inequalities without any observations of the outcomes in $\mathcal{Y}$. We list a set of rules that could allow us to conclude that some of the inequalities are redundant given the others.

Rule 1 *For any subset $A_1 \subset \mathcal{U}$ and $A_2 \subset \mathcal{U}$, if $A_1 \cap A_2 = \emptyset$ and $\varphi(A_1) \cap \varphi(A_2) = \emptyset$, then the two inequalities, $v(A_1) \leqslant \mu(\varphi(A_1))$ and $v(A_2) \leqslant \mu(\varphi(A_2))$, can generate the inequality $v(A_1 \cup A_2) = v(A_1) + v(A_2) \leqslant \mu(\varphi(A_1)) + \mu(\varphi(A_2)) = \mu(\varphi(A_1) \cup \varphi(A_2)) = \mu(\varphi(A_1 \cup A_2))$, which is exactly the inequality corresponding to the subset $A = A_1 \cup A_2$. In other words, the inequality $v(A) \leqslant \mu(A)$ is redundant given $v(A_1) \leqslant \mu(\varphi(A_1))$ and $v(A_2) \leqslant \mu(\varphi(A_2))$.*

Rule 2 *If $u \notin A$ satisfies $\varphi(\{u\}) \subset \varphi(A)$, then the inequality $v(\{u\} \cup A) \leqslant \mu(\varphi(\{u\} \cup A))$ will imply a redundant inequality $v(A) \leqslant \mu(\varphi(A))$.*

Rule 3 *Suppose $\varphi(\mathcal{U}) = \mathcal{Y}$. For any sets $A_1 \subset \mathcal{U}$ and $A_2 \subset \mathcal{U}$, $A_1 \cup A_2 = \mathcal{U}$, we can define $A_1^c := \mathcal{U} \backslash A_1$ and $A_2^c := \mathcal{U} \backslash A_2$. Denote $A_0 := A_1 \cap A_2$. Then, A_0, A_1^c, A_2^c form a partition of $\mathcal{U}$. Define $B_0 := \varphi(A_0)$, $B_1 := \varphi(A_1) \backslash B_0$, $B_2 := \varphi(A_2) \backslash B_0$. If $B_1 \cap B_2 = \emptyset$, then B_0, B_1, B_2 form a partition of $\mathcal{Y}$.*

The inequalities $v(A_1) \leqslant \mu(\varphi(A_1))$, $v(A_2) \leqslant \mu(\varphi(A_2))$ and $v(\mathcal{U}) = \mu(\mathcal{Y})$ imply that:
$v(A_1) + v(A_2) \leqslant \mu(\varphi(A_1)) + \mu(\varphi(A_2))$, *i.e.,*
$v(A_0) + v(A_2^c) + v(A_0) + v(A_1^c) \leqslant \mu(B_0) + \mu(B_1) + \mu(B_0) + \mu(B_2)$.
Since $v(\mathcal{U}) = v(A_0) + v(A_2^c) + v(A_1^c)$ and $\mu(\mathcal{Y}) = \mu(B_0) + \mu(B_1) + \mu(B_2)$, and $v(\mathcal{U}) = \mu(\mathcal{Y})$, it follows that $v(A_0) \leqslant \mu(B_0) = \mu(\varphi(A_0))$.
That said, if $B_1 \cap B_2 = \emptyset$, then $v(A_1 \cap A_2) \leqslant \mu(\varphi(A_1 \cap A_2))$ is implied by $v(A_1) \leqslant \mu(\varphi(A_1))$, $v(A_2) \leqslant \mu(\varphi(A_2))$ and $v(\mathcal{U}) = \mu(\mathcal{Y})$.

To find the exact core-determining class, the combinatorial rules stated above are helpful. The key questions are: whether the set of rules in the above are necessary and sufficient? Are there any other rules that would imply redundancy of inequalities? Would one inequality that is redundant defined by the combinatorial rules become irredundant once we remove a certain number of inequalities? To answer these questions, we utilize another way that rigorously defines redundancy with techniques in linearing programming.

Definition 5 (Set $\mathcal{S}^*$)**.** *For any non-empty set $A \subset \mathcal{U}$, define $\mathcal{V}_A^* := \{v | v \in \mathbb{R}^{d_u}, v \geqslant 0; v(A') \leqslant \mu(\varphi(A'))$, for all $A' \subset \mathcal{U}, A' \neq A; v(\mathcal{U}) = \mu(\mathcal{Y})\}$. Define $v_A^{M^*} \in argmax_{v \in \mathcal{V}_A^*} v(A)$ as a probability measure on $\mathcal{U}$ that is maximizing $v(A)$ subject to the constraint that $v \in \mathcal{V}_A^*$.*

Define the set $\mathcal{S}^*$ as the collection of all subsets $A \subset \mathcal{U}$ and $A \neq \mathcal{U}$ such that

$$v_A^{M^*}(A) > \mu(\varphi(A)).$$

Each element A in $\mathcal{S}^*$ is characterized by linear programming that checks the redundancy of the inequality that correspond to A with respect to all other inequalities/equality in (2.1) and (2.2). For any collection of subsets of $\mathcal{U}$, denoted as $\mathcal{S}$, we can define $Q(\mathcal{S}) := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A)),\ \text{for all} A \in \mathcal{S},\ \text{and } v(\mathcal{U}) = \mu(\mathcal{Y})\}$. The following theorem shows that $\mathcal{S}^*$ defined in Definition 5 is the only exact core-determining class under the conditions that $\mu(\cdot)$ and G are non-degenerate.

Theorem 1 (Sharpness of $\mathcal{S}^*$)**.** *Suppose both G and μ are non-degenerate. Then,*

(1) $Q(\mathcal{S}^*) = Q$.

(2) For any $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{\emptyset\}$ that does not contain $\mathcal{S}^*$, $Q(\mathcal{S}) \supsetneq Q$, i.e., $\mathcal{S}$ is not exact core-determining; or equivalently, any core-determining class must contain $\mathcal{S}^*$.

(3) $\mathcal{S}^*$ is the unique exact core-determining class.

Theorem 1 says that for any inequality $v(A) \leqslant \mu(\varphi(A))$, as long as it is redundant given all other inequalities/equality, we could remove this inequality without losing any information on the identified region Q. Theorem 1 proves that $\mathcal{S}^*$ is the smallest possible core-determining class under non-degeneracy conditions of G and $\mathcal{Y}$. However, it is not conveniet enough for us to apply Theorem 1 in order to construct the exact core-determining class, since running $2^{d_u} - 1$ linear programming problems can be computational costly when d_u is large. We would like to characterize $\mathcal{S}^*$ using combintorial rules like those in

Rules 1-3 stated in the beginning of this subsection, which could potentially allow us to obtain exact core-determining class within a short time. The main results on the combinatorial properties of exact core-determining class are formally stated in Subsection 3.3. Before jumping to the main theoretical results, we require a set of definitions and lemmas that are presented in Subsection 3.2.

3.2. Intermediate Results

In this subsection, we present a list of results that are important for establishing our main results in Subsection 3.3.

We start by considering the following collection of subsets $\mathcal{S}_u$.

Definition 6 (v_A^M)**.** *Define* $\mathcal{V}_A := \{v | v \geqslant 0, v \in \mathbb{R}^{|\mathcal{U}|},$ *and* $v(A') \leqslant \mu(\varphi(A'))$, *for any* $A' \neq A, A' \subset \mathcal{U}, A' \neq \emptyset\}$. *Define* $v_A^M \in argmax_{v \in \mathcal{V}_A} v(A)$. *Since the true probability distribution* $v_0 \in \mathcal{V}_A$, v_A^M *is always well-defined when* $|\mathcal{U}| \geqslant 2$.

Definition 7 (Set $\mathcal{S}_u$)**.** $\mathcal{S}_u \subset 2^{\mathcal{U}}$ *is the collection of all non-empty subsets* $A \subset \mathcal{U}$ *such that*

$$v_A^M(A) > \mu(\varphi(A)).$$

Set $\mathcal{S}_u$ is defined from a numerical perspective given the observed probability measure μ. The difference between $\mathcal{S}_u$ and $\mathcal{S}^*$ is that $\mathcal{S}_u$ ignores the equality that $v(\mathcal{U}) = \mu(\mathcal{Y})$. Define $Q_1 := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A))$ for all non-empty subset $A \subset \mathcal{U}\}$. Q_1 is the identified region when we only consider inequalities in (2.1).

$\mathcal{S}_u$ is unique and well-defined. Figure 7 show example of subset $A \in \mathcal{S}_u$ and $A \notin \mathcal{S}_u$.

In Figure 7, the white polygon region represents the set $\widetilde{Q}_1 := \{v \geqslant 0 | v(A') \leqslant \mu(\varphi(A'))$ for all $A' \neq A\}$. When $A \in \mathcal{S}_u$, by Definition 7, the half space, colored in purple, $\{v | v(A) \leqslant \mu(\varphi(A))\}$ should interect $\widetilde{Q}_1$ and therefore the inequality $v(A) \leqslant \mu(\varphi(A))$ is irredundant given all other inequalities in (2.1). When $A \notin \mathcal{S}_u$, the half space $\{v | v(A) \leqslant \mu(\varphi(A))\}$ contains $\widetilde{Q}_1$ and therefore the inequality $v(A) \leqslant \mu(\varphi(A))$ is redundant given all other inequalities in (2.1).

The definition of Q_1 ingores the important equality $v(\mathcal{U}) = \mu(\mathcal{Y})$, and therefore Q_1 is not our final target of interest. However, it is essential to identify irredundant inequalities for Q_1 which is defined by $2^{d_u} - 1$ inequalities corresponding to all non-empty subsets of $\mathcal{U}$. Below we present a lemma to formally show that the inequalities corresponding to the subsets in $\mathcal{S}_u$ contain all information defined by the system of inequalities described in (2.1). At the same time, dropping any inequality that corresponds to a subset in $\mathcal{S}_u$ would lead to loss of information, i.e., the identified region will be strictly larger than Q_1 if any inequality corresponding to a subset in $\mathcal{S}_u$ is eliminated.

Lemma 2 (Sharpness of $\mathcal{S}_u$)**.** *Suppose* G, μ *are non-degenerate. For any collection of non-empty subsets* $\mathcal{S} \subset 2^{\mathcal{U}}$, *define* $Q_1(\mathcal{S}) := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A))$ *for all* $A \in \mathcal{S}\}$. *Then,*

(1) $Q_1(\mathcal{S}_u) = Q_1$.

(2) $Q_1(\mathcal{S}) \supsetneq Q_1$ *if* $\mathcal{S}$ *does not contain* $\mathcal{S}_u$.

In a system of inequalities (equalities can be allowed as well), we say a subset of inequalities is sharp, if and only if this subset of inequalities identifies the same region as the entire system of inequalilities.

Lemma 2 provides a simple way to exactly declare which inequalities can be eliminated from the system described in (2.1) without losing information on the identified region defined by all inequalities in (2.1). For each non-empty subset $A \subset \mathcal{U}$, we can solve the following linear programming problem to obtain the value of v_A^M, and then check whether A is an element of $\mathcal{S}_u$ according to Definition 7. However, it is very computationally costly since there are $2^{d_u} - 2$ elements A in total.

$$v_A^M := \underset{v}{\text{maximize}}\ v(A)$$

subject to

$$v(A') \leqslant \mu(\varphi(A')), \forall A' \neq A, A' \subset \mathcal{U}, A' \neq \emptyset$$
$$v \geqslant 0, v \in \mathbb{R}^{|\mathcal{U}|}$$

Lemma 2 also shows that any inequality corresponding to a subset $A \in \mathcal{S}_u$ can not be eliminated without losing information on the identified region Q_1. Also, in the linear programming problem shown above, it is unclear that whether μ would have an effect on $\mathcal{S}_u$. Next, we present a major result which allows us to construct $\mathcal{S}_u$ only via the combinatorial structure of the bipartite graph G under some mild conditions.

Definition 8 (Set $\mathcal{S}'_u$)**.** *$\mathcal{S}'_u \subset 2^{\mathcal{U}}$ is the collection of all non-empty subsets $A \subset \mathcal{U}$ such that:*

(1) A is self-connected, i.e., $\forall A_1, A_2 \subset A$ such that $A_1, A_2 \neq \emptyset$ and $A_1 \cup A_2 = A$, it holds that $\varphi(A_1) \cap \varphi(A_2) \neq \emptyset$;

(2) There exists no $u \in \mathcal{U}$ such that $u \notin A$ and $\varphi(u) \subset \varphi(A)$.

Set $\mathcal{S}'_u$ is defined from a combinatorial perspective using the structure of the bipartite graph G. In the lemma below, we show that the set $\mathcal{S}_u$ equals to $\mathcal{S}'_u$.

Lemma 3 (Combinatorial Description of Q_1)**.** *If both G and μ are non-degenerate, then the collection of subsets defined in Definitions 7 and 8 are identical, i.e., $\mathcal{S}_u = \mathcal{S}'_u$.*

The Lemma 3 says, $\mathcal{S}'_u$, whose corresponding inequalities is sharp amongst all inequalities in (2.1). By Definition 8, we know that $\mathcal{S}_u$ covers rules 1-2 stated in Subsection 3.1.

Comment 1. Under the conditions of Lemma 3, Theorem 5 of Chesher and Rosen (2012) selects a subset of inequalities that is equivalent to those corresponding to S'_u, which is a core-determining class, but not necessarily the exact core-determining class. Lemma 3 shows that we can find all irredundant inequalities using combinatorial rules when the following constraint is not considered

$$v(\mathcal{U}) := \sum_{u \in \mathcal{U}} v(u) \geqslant \sum_{y \in \mathcal{Y}} \mu(y) = \mu(\mathcal{Y}).$$

Under the mild regularity conditions in Lemma 7, the set of irredudant inequalities defined in Definition 7 can be characterized by the set of combinatorial rules defined in Defintion 8. It implies that the collection of irredudant inequalities defined in Definition 7 is unique and only depends on the structure of the Graph G, not the probability measure $\mu(\cdot)$ conditional on the assumption that both G and $\mu(\cdot)$ are non-degenerate.

To find the minimum set of irredudant inequalities in the entire system described in (2.1) and (2.2), (the exact core-determining class in Definition 4), we next consider a bipartite graph G^{-1} with correspondence φ^{-1} mapping from $2^{\mathcal{Y}} \mapsto 2^{\mathcal{U}}$, i.e., $G^{-1} := (\mathcal{Y}, \mathcal{U}, \varphi^{-1})$. For any non-degenerate probability measure $\widetilde{v}$ on $\mathcal{U}$, we define $\mathcal{S}_y$ as the following:

Definition 9. *Assume that $\mu \geqslant 0$ is a vector defined on the space $\mathbb{R}^{d_y}$. Suppose there exists a probability measure $\widetilde{v} > 0$ defined on $\mathcal{U}$. Define $\mu_B^M := argmax_{\mu \geqslant 0}\{\mu(B) | \mu(B') \leqslant \widetilde{v}(\varphi^{-1}(B')),$ for all $B' \subset \mathcal{Y}, B' \neq B, B' \neq \emptyset\}$. Such a $\widetilde{v}$ can be the unknown true probability distribution v_0 on $\mathcal{U}$.*

Definition 10 (Set $\mathcal{S}_y$)**.** *Given a non-degenerate probability measure $\widetilde{v}$ on $\mathcal{U}$, $\mathcal{S}_y \subset 2^{\mathcal{Y}}$ is the collection of all subsets $B \subset \mathcal{Y}$ such that*

$$\mu_B^M(B) > \widetilde{v}(\varphi^{-1}(B)).$$

By Lemma 2, when replacing G with G^{-1}, assuming that $\widetilde{v}$ is non-degenerate, the inequalities corresponding to $\mathcal{S}_y$ contain sharp information in the system of inequalities: $\{\mu \geqslant 0 | \mu(B) \leqslant \varphi^{-1}(B),$ for all non-empty $B \subset \mathcal{Y}\}$. Set $\mathcal{S}_y$ is defined from a numerical perspective using any non-degenerate probability measure $\widetilde{v}$. Similar to the Definition 7, the set $\mathcal{S}_y$ consists the collection of irredudant inequalities amongst all $\mu(B) \geqslant \widetilde{v}(\varphi^{-1}(B))$, $B \subset \mathcal{Y}, B \neq \emptyset$.

Definition 11 (Set $\mathcal{S}'_y$)**.** *$\mathcal{S}'_y \subset 2^{\mathcal{Y}}$ is the collection of all subsets $B \subset \mathcal{Y}$ such that:*

(1) B is self-connected, i.e., $\forall B_1, B_2 \subset B$ such that $B_1, B_2 \neq \emptyset$ and $B_1 \cup B_2 = B$, it holds that $\varphi^{-1}(B_1) \cap \varphi^{-1}(B_2) \neq \emptyset$;

(2) There exists no $y \in \mathcal{Y}$ such that $y \notin B$ and $\varphi^{-1}(y) \subset \varphi^{-1}(B)$.

The set of $\mathcal{S}_y$ is defined as the "irredundant inequalities" in $G^{-1}(\mathcal{Y}, \mathcal{U}, \varphi^{-1})$ when the equality $v(\mathcal{U}) = \mu(\mathcal{Y})$ is not taken into consideration. It provide additional information on the identity of redundant inequalities compared to $\mathcal{S}_u$. Later in the main Theorem 2, we show that $\mathcal{S}_y$ and $\mathcal{S}_u$ together determines the exact identity of irredundant inequalities.

Similar to $\mathcal{S}'_u$, the set $\mathcal{S}'_y$ is defined from a combinatorial perspective using the structure of the bipartite graph G.

Lemma 4. *Suppose both the graph G^{-1} and $\widetilde{v}$ are non-degenerate. The collection of subsets defined in Definition 10 and 11 are identical, i.e., $\mathcal{S}_y = \mathcal{S}'_y$.*

The proof of Lemma 4 can be obtained simply by replacing G and μ with G^{-1} and $\widetilde{v}$ in Lemma 3. Therefore, we abbreviate this proof.

Definition 12 (Set $\mathcal{S}_y^{-1}$). *The Set $\mathcal{S}_y^{-1}$ is the collection of $A \subset \mathcal{U}$ and $A \neq \mathcal{U}$ such that there exists $B \subset \mathcal{S}'_y$ that $A = \varphi^{-1}(B)^c$.*

The reason that we would like to introduce $\mathcal{S}_y^{-1}$ is that the combinatorial rules defined in Definition 11 covers Rule 3 stated in Subsection 3.1. The Definitions of $\mathcal{S}'_u$ and $\mathcal{S}'_y$ and their properties allow us to obtain the main theoretical results on exact core-determining class in Subsection 3.3.

3.3. Main Theoretical Results

In the lemma below, we show that, by combining combinatorial rules defined for $\mathcal{S}_u$ and $\mathcal{S}_y$, we are able to obtain the exact core-determining class.

Lemma 5. *Assume that both G and μ are non-degenerate. Then, the exact core-determining class $\mathcal{S}^*$ is characterized by the following equation:*

$$\mathcal{S}^* = \left(\mathcal{S}_u \cap \mathcal{S}_y^{-1}\right) \backslash \{\mathcal{U}\}.$$

The Example 4 demonstrate the importance of considering $\mathcal{S}_y^{-1}$ in the construction of exact core-determining class. The collection of sets could be substantially reduced in some cases when intersecting $\mathcal{S}_y^{-1}$ with $\mathcal{S}_u$.

Example 4. *Consider a bipartite graph G with set $\mathcal{U} = \{u_1, ..., u_5\}$ and set $\mathcal{Y} = \{y_1, ..., y_4\}$. φ is the correspondence mapping between $\mathcal{U}$ and $\mathcal{Y}$ such that $\varphi(u_j) = \{y_j\}$ for all $1 \leqslant j \leqslant 4$ and $\varphi(u_5) = \{y_1, y_2, y_3, y_4\}$. The correspondence mapping for Example 4 is illustrated in Figure 8.*

In this example, $\mathcal{S}_y^{-1} \backslash \{\mathcal{U}\} = \{\mathcal{U}_1 | \mathcal{U}_1 \subset \{u_1, u_2, u_3, u_4\}, \mathcal{U}_1 \neq \emptyset\}$, which consists of $2^4 - 2$ subsets. $\mathcal{S}_u \backslash \{\mathcal{U}\} = \{u_j | 1 \leqslant j \leqslant 4\}$. The exact core-determining class $\mathcal{S}^*$ is $\{u_j | 1 \leqslant j \leqslant 4\}$, which equals to $\mathcal{S}_u \cap \mathcal{S}_y^{-1} \backslash \{\mathcal{U}\}$.

Since both $\mathcal{S}_u$ and $\mathcal{S}_y^{-1}$ are defined via combinatorial rules, we can construct the exact core-determining class using a combinatorial algorithm. The set $\mathcal{U}$ is automatically ruled out in computing the exact core-determining class because the equality in (2.2) implies $v(\mathcal{U}) \leqslant \mu(\mathcal{Y})$. The exact core-determining class is also independent of μ if μ and G are both non-degenerate. Also, since $\mathcal{S}_u$ and $\mathcal{S}_y^{-1}$ are uniquely defined, when the assumptions in Theorem 2 holds, any core-determining class should contain $(\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$. Therefore, in such a case, the exact core-determining class is unique and characterized by the formula $(\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$.

With Lemma 5, we establish our main theorem.

Theorem 2 (Combinatorial Rules For Exact Core-determining). *Suppose both G and μ are non-degenerate. Then a non-empty set $A \subsetneq \mathcal{U}$ is exact core-determining, i.e., $A \in \mathcal{S}^*$, if and only if:*

(1) A is self-connected, i.e., $\forall A_1, A_2 \subset A$ such that $A_1, A_2 \neq \emptyset$ and $A_1 \cup A_2 = A$, it holds that $\varphi(A_1) \cap \varphi(A_2) \neq \emptyset$;

(2) There exists no $u \in \mathcal{U}$ such that $u \notin A$ and $\varphi(u) \subset \varphi(A)$.

(3) $\varphi(A)^c$ is self-connected in the graph G^{-1}.

Comment 2. *Given the conditions of Theorem 2, it is easy to obtain a core-determining class in Galichon and Henry (2011) framework. Define* $\mathcal{S}^{-1*} := \{B|B = \varphi(A)^c, A \in \mathcal{S}^*\}$.

For any $A \in \mathcal{S}^*$, $v(A) \leqslant \mu(\varphi(A))$ *holds if and only if* $v(\varphi^{-1}(B)) \geqslant \mu(B)$, *since* $v(A) + v(\varphi^{-1}(B)) = v(\mathcal{U}) = 1$, *and* $\mu(\varphi(A)) + \mu(B) = \mu(\mathcal{Y}) = 1$. *Then, the identified region* Q *can also be characterized as* $Q = \{v \geqslant 0|v(\varphi^{-1}(B)) \geqslant \mu(B), v(\mathcal{U}) = \mu(\mathcal{Y})\}$.

Then given $v(\mathcal{U}) = \mu(\mathcal{Y})$, $\mathcal{S}^{-1*} := \{B|B = \varphi(A)^c$ *for any* $A \in \mathcal{S}^*\}$ *is a core-determining class (and the smallest) in the sense of Definition 7 of Galichon and Henry (2011).*

Theorem 2 offers theoretical garantee that the exact core-determining class $\mathcal{S}^*$ is unique when both G and μ are non-degenerate, and $\mathcal{S}^*$ can be identified via combinatorial rules. In practice, the graph G can be unconnected, which violates the non-degeneracy assumption. In such a case, the exact core-determining class may be non-unique. We can assume that the connected branches of G include $G_1, ..., G_k$, with $G_j := (\mathcal{U}_j, \mathcal{Y}_j, \varphi)$, where $\varphi(\mathcal{U}_j) = \mathcal{Y}_j$, $j = 1, 2, ..., k$. These $\mathcal{U}_j$, $1 \leqslant j \leqslant k$ form a partition of $\mathcal{U}$, i.e., $\cup_{1 \leqslant j \leqslant k} \mathcal{U}_j = \mathcal{U}$, and $\mathcal{U}_{j_1} \cap \mathcal{U}_{j_2} = \emptyset$ for any $j_1 \neq j_2, j_1, j_2 \in \{1, 2, ..., k\}$. Similary, $\mathcal{Y}_j$, $1 \leqslant j \leqslant k$ also form a partition of $\mathcal{U}$. By defintion, each G_j is connected for $j = 1, 2, ..., k$. We can define the set $\mathcal{S}_j^*$ as the set $\mathcal{S}^*$ described in Theorem 2, with G replaced by G_j, $j = 1, 2, ..., k$.

Theorem 3. *Suppose* μ *is non-degenerate, and* $\varphi(\mathcal{U}) = \mathcal{Y}$. *Assume that* $G_1, ..., G_k$ *are the connected branches of* G, $k \geqslant 2$, *with each* $G_j = (\mathcal{U}_j, \mathcal{Y}_j, \varphi)$, *where* $\mathcal{U}_1, ..., \mathcal{U}_k$ *form a partition of* $\mathcal{U}$ *and* $\mathcal{Y}_1, ..., \mathcal{Y}_k$ *form a partition of* $\mathcal{Y}$. *Denote* $\mathcal{S}_j^*$ *as the set* $\mathcal{S}^*$ *described in Theorem 1 with* G *being replaced by* G_j, $j = 1, 2, ..., k$. *Then* $\mathcal{S}^* := \cup_{j=1}^k (\mathcal{S}_j^* \cup \{\mathcal{U}_j\})$ *is an exact core-determining class with resepect to* G, *i.e., a class of subsets with the minimum possible cardinalty that is core-determining.*

3.4. Algorithm that Computes the Exact core-determining Class

To be able to compute the exact core-determining class, we propose an algorithm for those graphs G that are connected. When the graph G is unconnected, we could perform the Algorithm 1 for each connected branch of G and construct an exact core-determining class according to Theorem 3.

Based on results in Lemma 5, we only need to know how to compute $\mathcal{S}_u$, since the construction of $\mathcal{S}_y$ could be performed by applying the same algorithm on the graph G^{-1}. The Algorithm 1 constructs $\mathcal{S}_u$.

The complexity of Algorithm 1 is $o(2^{\max(d'_u, d'_y)} \cdot d'^2_u \cdot d'^2_y)$, where d'_u is defined as

$$d'_u := \max_A |A|$$

$$s.t. A \subset \mathcal{U}$$

$$\varphi(A) = \mathcal{Y}$$

$$\varphi(A/u) \subsetneq \mathcal{Y}, \forall u \in A$$

input : Bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$
output: Set $\mathcal{S}'_u$
Initiation: $\mathcal{S}'_u = \{\emptyset\}$
for $i \leftarrow 0$ **to** $|\mathcal{U}| - 1$ **do**
    Identify additional $A' \in \mathcal{S}'_u$ *as the union of* $u \in \mathcal{U}$ *and* $A \in \mathcal{S}'_u$ *with* $|A| = i$,
    foreach $A \in \mathcal{S}'_u$ *with* $|A| = i$ **do**
        foreach $u \notin A$ *and* $(\varphi(u) \cap \varphi(A) \neq \emptyset$ *or* $\varphi(A) = \emptyset)$ **do**
            $A' \leftarrow A \cup \{u\}$,
            if $\varphi(A') < 1$ **then**
                foreach $u' \notin A'$ **do**
                    if $\varphi(u') \subset \varphi(A')$ **then** $A' \leftarrow A' \cup \{u'\}$;
                end
                if $A' \notin \mathcal{S}'_u$ **then** $\mathcal{S}'_u \leftarrow \mathcal{S}'_u \cup \{A'\}$;
            end
        end
    end
end
Termination: $\mathcal{S}'_u \leftarrow \mathcal{S}'_u - \{\emptyset\}$

Algorithm 1: Construct Set $\mathcal{S}'_u$

d'_y is defined as

$$\begin{gathered} d'_y := \max_B |B| \\ s.t. B \subset \mathcal{Y} \\ \varphi^{-1}(B) = \mathcal{U} \\ \varphi^{-1}(B/y) \subsetneq \mathcal{U}, \forall y \in B \end{gathered}$$

Intuitively, the first foreach loop tries to add one more element which keeps the sets being self-connected. That said, the property (1) of the Definition 8 holds for all sets that are found by the algorithm. In the second foreach loop, when an additional element is added to a set A, the algorithm will include all u' such that $\varphi(u')$ is subset of $\varphi(A \cup \{u\})$, which (1) substantially reduce the search cost (2) garantees that the set found always statisfies property (2) of the Definition 8 (3) the new set is still self-connected. It is also easy to see that any set in $\mathcal{S}'_u$ must be constructed during the process, since Algorithm 1 scans all possible subgraphs of $\mathcal{G}$ that satisfies the two properties stated in Definition 8. Under the assumption of non-degeneracy of G and μ, for the bipartite graph in most practical examples, d'_u and d'_y will be much smaller than d_u and d_y respectively, so in practice the algorithm could be very fast.

4. Large Scale Examples on Constructing Exact core-determining Class

As illustrations of the models and algorithms described in Section 3, we now implement the models and algorithm in two numerical examples. For each example, we construct the exact core-determining class and demonstrate the performance of the algorithm.

The first numerical example is an illustrative example described by Galichon and Henry (2011).

Example 5. [Two-type Oligopoly Entry Model] Consider a bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$ with 9 events in $\mathcal{U}$ and 14 outcomes in $\mathcal{Y}$. The corresponding mapping is illustrated in Figure 9. Graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$ is unconnected in this example. Specifically, u_1 and u_9 are unconnected with other u in $\mathcal{U}$, and y_1 and y_{14} are unconnected with other y in $\mathcal{Y}$. Therefore, we can decompose G to three bipartite graphs: The first graph, G_1, contains a single event u_1 and a single outcome y_1, with the correspondence mapping $\varphi(u_1) = \{y_1\}$; the second graph, G_2, contains a single event u_9 and a single outcome y_{14}, with the correspondence mapping $\varphi(u_9) = \{y_{14}\}$; the third graph, G_3, contains 7 events—i.e., u_2 to u_8—and 12 outcomes—i.e., y_2 to y_{13}—with the correspondence mapping illustrated in the other parts of Figure 9.

We combine the inequalities generated from the three bipartite graphs G_1, G_2, and G_3 to obtain the exact core-determining class for the original bipartite graph G. Apparently, the first graph, G_1, generates a single equality, $v\{u_1\} = \mu\{y_1\}$, and the second graph, G_2, generates another single equality $v\{u_9\} = \mu\{y_{14}\}$. If we implement the algorithm described in Section 3 on the third graph, G_3, we can obtain corresponding results (shown in Table 1) within one second using a normal computer. Specifically, the bipartite graph G_3 indicates $d_u = 7$ unobservable events and $d_y = 12$ observed outcomes. According to Artstein's Theorem, the number of inequalities defining the feasible set of probability measures on $\mathcal{U}$ is $2^{d_u} - 2 = 2^9 - 2 = 126$. Then, Algorithm 1 generates a set $\mathcal{S}_u$ with $|\mathcal{S}_u| = 22$, corresponding to 22 inequalities. A similar algorithm generates a set $\mathcal{S}_y^{-1}$ with $|\mathcal{S}_y^{-1}| = 30$, corresponding to 30 inequalities. Our target, the exact core-determining class for G_3, can be obtained as $\mathcal{S}^* = \mathcal{S}_u \cap \mathcal{S}_y^{-1}$ with $|\mathcal{S}^*| = 11$, corresponding to 11 irredundant inequalities. Compared to the inequalities generated for G_3, according to Artstein's Theorem, we can reduce 91.26% of inequalities.

From statement 2.2 in Artstein's Theorem, we can obtain the equalities $v\{u_1\} = \mu\{y_1\}$, $v\{u_9\} = \mu\{y_{14}\}$, and $v\{\cup_{2\leqslant i\leqslant 8}\{u_i\}\} = \mu\{\cup_{2\leqslant j\leqslant 13}\{y_j\}\}$ for G_1, G_2, and G_3, respectively. For the original graph G, because of the existence of the equality $v\{\cup_{1\leqslant i\leqslant 9}\{u_i\}\} = \mu\{\cup_{1\leqslant j\leqslant 14}\{y_j\}\}$ from statement 2.2 in Artstein's Theorem, the three equalities from the three subgraphs can be relaxed as inequalities. Therefore, the exact core-determining class for G contains 1 inequality from G_1, 1 inequality from G_2, and 12 inequalities from G_3 (11 inequalities obtained from algorithm and 1 inequality with all elements in G_3). The corresponding subsets are: $\{u_1\}$, $\{u_2\}$, $\{u_9\}$,$\{u_5\}$, $\{u_8\}$, $\{u_2, u_3\}$, $\{u_7, u_8\}$, $\{u_2, u_3, u_4, u_5\}$, $\{u_5, u_6, u_7, u_8\}$, $\{u_2, u_3, u_4, u_5, u_6\}$, $\{u_4, u_5, u_6, u_7, u_8\}$, $\{u_2, u_3, u_4, u_5, u_6, u_7\}$, $\{u_3, u_4, u_5, u_6, u_7, u_8\}$, and $\{u_2, u_3, u_4, u_5, u_6, u_7, u_8\}$.

Next, we design a second numerical exeperiment with a larger dimension. Assuming that many marginal firms are facing a volatile market, let u be a random variable representing the cost of a firm, and is an unobservable event for us. The firm is aware of certain private information, denoted by $\theta \in \{H, L\}$, which is also unknown to us. Let y be the action of a firm based on the private information θ and the cost u—e.g., action y is the price set by the firm. We

Number of events × Number of outcomes ($d_u \times d_y$)	7×12
Number of inequalities according to Artstein's Theorem	$2^7 - 2 = 126$
Number of inequalities defined by $\mathcal{S}_u$	22
Number of inequalities defined by $\mathcal{S}_y^{-1}$	30
Number of inequalities defined by $\mathcal{S}^* = \mathcal{S}_u \cap \mathcal{S}_y^{-1}$	11
Percentage of redundant inequalities	91.26%

Table 1
Results of G_3 on Example 5

now consider a simple decision-making problem for the firms. If the objective of any firm is to maximize its profit $\pi(y, u, \theta)$, it might adopt different actions (set different prices y) when facing $\theta = H$ or $\theta = L$.

Assume that the profit function is

$$\pi(y, u, H) = (y - u)(C - y),$$

and

$$\pi(y, u, L) = (y - u)(C/2 - y),$$

where C is a constant.

If the firm considers any price $y^* \in \{y|\pi(y, u, \theta) \geqslant \max_y \pi(y, u, \theta) - w, a_1 \leqslant y \leqslant a_2\}$ as a robust pricing strategy, where w is a constant and a_1, a_2 are bounds on price y, then $\varphi(u) := \{y|\pi(y, u, \theta) \geqslant \max_y \pi(y, u, \theta) - w, \theta \in \{H, L\}, a_1 \leqslant y \leqslant a_2\}$ is the correspondence mapping from the set of costs (unobservable events) $\mathcal{U}$ to the set of prices (observed outcomes) $\mathcal{Y}$.

The prices y set by different firms are public information (observed outcomes). Since we could obtain the probability measure on the price, we can implement the models and algorithms described in Section 3 to obtain the distribution (probability measures) of the costs (unobservable events) of these firms. Specifically, we observe the measure on the set of prices $\mathcal{Y}$. The objective is to find the feasible set of probability measures on the set of costs $\mathcal{U}$, assuming that y is i.i.d. across observations.

Example 6. [I.I.D Pricing] In the contexts described above, assuming the set of costs $\mathcal{U} = [0, 3]$, the set of prices $\mathcal{Y} = [1, 3.5]$, constant $C = 4$ and $w = 0.01$, then the correspondence mapping φ from $\mathcal{U}$ to $\mathcal{Y}$ is $\varphi(u) = [(1.9 + u/2), (2.1 + u/2) \wedge 3.5] \cup [(0.9 + u/2 \vee 1), (1.1 + u/2)]$.

Assuming the probability measure on prices $\mathcal{Y}$ is observed, to estimate the probability measure on costs $\mathcal{U}$ we discretize the continuous set of costs (unobservable events) $\mathcal{U}$ and prices (observed outcomes) $\mathcal{Y}$. Let $d_u = 15$ and $d_y = 25$ be the number of discretized segments of costs and prices, respectively. Then we define $u_i = ((i-1)/5, i/5)$ and $y_j = ((j-1)/10 + 1, j/10 + 1)$ for $i = 1, 2, ..., 15$ and $j = 1, 2, ..., 25$. The correspondence mapping φ_{Dsc} from the discretized set $\mathcal{U}_{Dsc} = \{u_i | i = 1, 2, ..., 15\}$ to the discretized set $\mathcal{Y}_{Dsc} = \{y_j | j = 1, 2, ..., 25\}$ is generated by:

$$\varphi_{Dsc}(u_i) = \{y_j | y_j \cap \varphi(u_i) \neq \emptyset\}$$

Number of events × Number of outcomes ($d_u \times d_y$)	15×25
Number of inequalities according to Artstein's Theorem	$2^{15} - 2 = 32766$
Number of inequalities defined by $\mathcal{S}_u$	937
Number of inequalities defined by $\mathcal{S}_y^{-1}$	585
Number of inequalities defined by $\mathcal{S}^* = \mathcal{S}_u \cap \mathcal{S}_y^{-1}$	471
Percentage of redundant inequalities	98.56%

TABLE 2
Results of Example 6

Therefore, $\varphi_{Dsc}(u_1) = \{y_1, y_2, y_{10}, y_{11}, y_{12}\}$, $\varphi_{Dsc}(u_{15}) = \{y_{14}, y_{15}, y_{16}, y_{24}, y_{25}\}$, and for any $2 \leqslant i \leqslant 14$, $\varphi_{Dsc}(u_i) = \{y_{i-1}, y_i, y_{i+1}, y_{i+9}, y_{i+10}, y_{i+11}\}$. Figure 10 illustrates the bipartite graph $G_{Dsc} = (\mathcal{U}_{Dsc}, \mathcal{Y}_{Dsc}, \varphi_{Dsc})$ constructed in this manner for Example 6. Apparently, the graph G_{Dsc} is non-degenerate. Without loss of generality, we also assume that the observed probability measure μ on $\mathcal{Y}_{Dsc}$ is non-degenerate.

We implement the algorithm described in Section 3 to explore the structure of the bipartite graph $G_{Dsc} = (\mathcal{U}_{Dsc}, \mathcal{Y}_{Dsc}, \varphi_{Dsc})$ for Example 6. We can obtain corresponding results (shown in Table 2) within one second using a normal computer. . The bipartite graph G_{Dsc} indicates $d_u = 15$ unobservable events and $d_y = 25$ observed outcomes. According to Artstein's Theorem, the number of inequalities that define the feasible set of probability measures on $\mathcal{U}_{Dsc}$ is $2^{d_u} - 2 = 2^{15} - 2 = 32766$. Then, Algorithm 1 generates a set $\mathcal{S}_u$ with $|\mathcal{S}_u| = 937$, corresponding to 937 inequalities. A similar algorithm generates a set $\mathcal{S}_y^{-1}$ with $|\mathcal{S}_y^{-1}| = 585$, corresponding to 585 inequalities. Our final target, the exact core-determining class for Example 6, can be obtained as $\mathcal{S}^* = \mathcal{S}_u \cap \mathcal{S}_y^{-1}$ with $|\mathcal{S}^*| = 471$, corresponding to 471 irredundant inequalities. Compared to the inequalities generated according to Artstein's Theorem, we can reduce 98.56% of inequalities.

5. Impact on Inference Procedures and Power of Tests

We consider the test-statistics described in Andrew and Shi (2014) to evaluate the impact of the exact core-determining class on inference procedures and power of statistical tests. In this section, we use $\widehat{\mu}$ to replace μ, where $\widehat{\mu}(y) = \frac{1}{n}\sum_{i=1}^{n} 1(y_i = y)$ and $n =$ sample size.

Take a vector v, the test-statistics accepts the null hypothesis at probability at least $1 - \alpha$ that $v \in Q$ if and only if $T(v) \leqslant c_v(1-\alpha)$, where $T(v)$ is the test-statistics in Andrew and Shi (2014). The cut-off value $c_v(1-\alpha)$ can be calculated using the full set of inequalities, denoted as $c_v^{full}(1-\alpha)$, or can be calculated using just the exact core-determining class $\mathcal{S}^*$ combined with the equality $v(\mathcal{U}) = \mu(\mathcal{Y})$, denoted as $c_v^{\mathcal{S}^*}(1-\alpha)$. We implement the following steps in the inference procedures:

(1) Compute $c_v^{full}(1-\alpha)$ and $c_v^{\mathcal{S}^*}(1-\alpha)$. v can be chosen as a point on the boundary of region Q.

(2) Report the time of the inference procedures, and report the time of calculating the core-determining class $\mathcal{S}^*$.

(3) Consider a local alternative $v'(t)$ to v, i.e., $v'(t) = v + t\widetilde{v}/\sqrt{n}$ where $\widetilde{v}$ is a unit measure on $\mathcal{U}$ such that $v'(t) \notin Q$ for $\forall t > 0$ and $\widetilde{v}(\mathcal{U} = 0)$, report the power of test as $P^{full}(t) = Pr(T^{full}(v'(t)) > c^{full}_{v'(t)}(1-\alpha))$ and $P^{\mathcal{S}^*}(t) = Pr(T^{\mathcal{S}^*}(v'(t)) > c^{\mathcal{S}^*}_{v'(t)}(1-\alpha))$. Plot the curve of the $P^{full}(t)$ and $P^{\mathcal{S}^*}(t)$ for $t \in [0, a]$ where a is a positive number. We should expect the curve of $P^{full}(t)$ to lay below the curve of $P^{\mathcal{S}^*}$.

5.1. Test Statistics

Suppose, in the core-determining class problem, we can observe a i.i.d. sample of outcomes $y_k \in \mathcal{Y}$, $k = 1, 2, ..., n$. Assume that this i.i.d. sample is generated by a distribution measure μ on $\mathcal{Y}$. The probability measure v is not observable.

We follow the procedures suggested in Andrews and Shi (2014). WLOG., suppose both the graph G and μ are non-degenerate, by Artstein's Theorem, we have these moment inequalities:

$v(A) - \widehat{\mu}(\varphi(A)) \leqslant 0$ for all $A \subset \mathcal{U}$, $A \notin \{\emptyset, \mathcal{U}\}$, where $\widehat{\mu}(\varphi(A)) = \mathbb{E}_n[1(y_i \in \varphi(A))] = \frac{1}{n}\sum_{i=1}^n 1(y_i \in \varphi(A))$,

and an equality $v(\mathcal{U}) = 1$.

The equality has no stochastic error and we will simply assume that it holds for v. Then, we can rewrite the moment function as:

$m_k(v|A) := v(A) - 1(y_i \in \varphi(A))$. According to the Artstein's Theorem, $\mathbb{E}[m_k(v|A)] \leqslant 0$ holds for all $A \subset \mathcal{U}$, $A \neq \emptyset, \mathcal{U}$. The empirical moment inequalities are:

$\mathbb{E}_n[m_k(v|A)] \leqslant 0$.

We can stack the $m_k(v|A)$ as a vector, denoted as $m_k(v)$. The sample variance-covariance matrix of $n^{\frac{1}{2}}\mathbb{E}_n[m_k(v)]$ is:

$\widehat{\Sigma}_n(v) = \frac{1}{n}\sum_{k=1}^n (m_k(v) - \mathbb{E}_n[m_k(v)])(m_k(v) - \mathbb{E}_n[m_k(v)])'$.

To avoid singularity of $\widehat{\Sigma}_n(v)$, Andrews and Shi (2014) suggests to use

$\bar{\Sigma}_n(v) = \widehat{\Sigma}_n(v) + \epsilon Diag(\widehat{\Sigma}_n(v))$ for some fixed $\epsilon > 0$.

Also, to avoid singularity of the diagnomal matrix of $\bar{\Sigma}_n(v)$, for the diagnomal element that equals to 0 in $\bar{\Sigma}_n(v)$, we make it equals to ϵ.

Andrew and Shi (2014) suggests test-statistics for the scenario where conditional inequalities exist. In our case, the test-statistics collapse into the following form:

$$T_n(v) = S(\widehat{m}(v), \bar{\Sigma}_n(v)), \tag{5.3}$$

where S is a function to be picked, and $\widehat{m}(v) = n^{\frac{1}{2}}\mathbb{E}_n[m_i(v)]$.

Denote p as the length of $\widehat{m}(v)$, i.e., $p = 2^{d_u} - 2$ when we use all the inequalities suggested in Artstein's Theorem, and $p = |\mathcal{S}^*|$ when we only use the inequalities selected in the exact core-determining class $\mathcal{S}^*$.

In Andrews and Shi (2014), three functions S are recommended:

(1) $S_1(\widehat{m}, \bar{\Sigma}) = \sum_{j=1}^{p} [\widehat{m}_j/\sigma_j]_+^2$, where $\widehat{m}_j$ is the j^{th} element of vector $\widehat{m}(v)$ and σ_j is the square root of the j^{th} diagonal element of matrix $\bar{\Sigma}$, where $x_+ = \max(0, x)$.

(2) $S_2(\widehat{m}, \bar{\Sigma}) = \inf_{t \in [0,\infty]^p} (\widehat{m}(v) - t)' \Sigma^{-1} (\widehat{m}(v) - t)$.

(3) $S_3(\widehat{m}, \bar{\Sigma}) = \max_{1 \leqslant j \leqslant p} [\widehat{m}_j/\sigma_j]_+^2$.

To make computation simple, we consider the $S(\cdot, \cdot)$ function stated in (1) and (3) in the following tests.

5.2. Computing the critical value $c(1-\alpha, v)$

Andrews and Shi (2014) suggests to use bootstrap to approximate the $(1-\alpha)$ quantile of $S(\widehat{m}(v), \bar{\Sigma}_n)$ as c_v, also denoted as $c(1-\alpha, v)$ in this section. Specifically, the procedure described in the Section 9 of Andrews and Shi (2014) could be reduced to the following steps in our context:

Step 1: Compute the test stastistics $T(v) := S(\widehat{m}(v), \bar{\Sigma}_n)$.

Step 1_{boot}: Bootstrap $y_1, ..., y_n$. Assume the bootstrapped sample is $y_1^{*,b}, ..., y_n^{*,b}$, $b = 1, 2, ..., M$, where M is the total number of bootstraps.

Compute $\widehat{m}(v)^{*,b}$ and $\bar{\Sigma}_n^{*,b}$, which are replacements of $\widehat{m}(v)$ and $\bar{\Sigma}_n$ by substituting $y_1, ..., y_n$ with $y_1^{*,b}, ..., y_n^{*,b}$.

Then, compute $\widetilde{m}^{*,b}(v) := \bar{D}_n(v)^{\frac{-1}{2}} (\widehat{m}(v)^{*,b} - \widehat{m}(v)) + \varphi_n(v)$, where $\bar{D}_n(v) := Diag(\bar{\Sigma}_n(v))$, and $\varphi_n(v)$ is defined as:

$\varphi_{n,j} = B_n 1(\zeta_{n,j}(v) > 1)$, where $\zeta_n := \kappa_n^{-1} \bar{D}_n^{-1/2} \widehat{m}(v)$ and $\zeta_{n,j}$ is the j^{th} entry of ζ_n. B_n and κ_n are chosen as two sequences of constants such that:

(1) $\kappa_n \to \infty$ (2) $B_n/\kappa_n \to 0$, as $n \to \infty$. Andrews and Shi (2014) suggests using $\kappa_n = (0.3ln(n))^{\frac{1}{2}}$ and $B_n = (0.4ln(n)/lnln(n))^{\frac{1}{2}}$.

Compute $\widetilde{\Sigma}_n^{*,b} := \bar{D}_n(v)^{\frac{-1}{2}} \bar{\Sigma}_n^{*,b}(v) \bar{D}_n(v)^{\frac{-1}{2}}$.

Compute the test stastistics $T^{*,b}(v) = S(\widetilde{m}^{*,b}(v), \widetilde{\Sigma}_n^{*,b})$.

Step 2: Compute the $1-\alpha+\eta$ quantile of the statistics $T^{*,1}(v)$,...,$T^{*,M}(v)$, denoted as $c_v(1-\alpha)$, where η is an arbitrarily small constant. In practice, we follow the suggestion from Andrews and Shi (2014) that $\epsilon = 0.01$ and $\eta = 10^{-6}$.

Step 3: Reject the hypothesis that $v \in Q$ if $T(v) > c_v(1-\alpha)$.

Andrews and Shi (2014) shows that asymptotically, the probability of $T(v) > c_v(1-\alpha)$ is bounded by α from the above.

5.3. Numerical Experiments of Power of Tests Against Local Alternatives

To investigate the benefit of performing statistical tests on the identified set of parameters, we consider local alternatives and compare the performance of the test stastistics proposed in Andrews and Shi (2014), based on both the full set of inequalities and the inequalities that correspond to the exact core-determining class. The power test procedure is implemented based on Example 5. To make things simple, we exclude the two disconnected events u_1 and u_9 by assuming $\mu\{y_1\} = 0$ and $\mu\{y_{14}\} = 0$. Essentially, we explore the non-degenerate subgraph

G_3 with 7 events and 12 outcomes. It takes less than 1 second to compute the exact core-determing class for the subgraph.

We choose a point v on the boundary of Q by solving a linear optimization problem on Q. Specifically, we get the optimal solution as v by maximizing the objective function $c'v$ in Q, where $c = \{1, 2, 3, 4, 5, 6, 7\}$. Then, we choose $\tilde{v} = 1/\sqrt{6}\{-1, -1, -1, 1, 1, 1, 0\}$, such that $\tilde{v}(\mathcal{U}) = 0$ and $v' = v + t\tilde{v}/\sqrt{n} \notin Q$ for all $t > 0$. We assume t is from 0 to 3, sample size $n = 1000$, boostrap size $M = 500$, and $\alpha = 0.05$. We replicate 1000 instances for each t. To avoid huge computational requrements of inverse matrix in function S_2, we test functions S_1 and S_3 in the inference experiment.

In a set of Monte-Carlo experiments, we denote $P_{S_j}^{\mathcal{S}^*}(t)$ as the probablity of $T(v + t\tilde{v}/\sqrt{n}) > c_v(1 - \alpha)$ where T and c are calculated using the set of inequalities in the exact core-determining class, and the function $S = S_j$, $j = 1, 3$. Similarly, $P_{S_j}^{full}(t)$ as the probablity of $T(v + t\tilde{v}/\sqrt{n}) > c_v(1 - \alpha)$ where T and c are calculated using the set of inequalities stated in the Artstein's Theorem, and the function $S = S_j$, $j = 1, 3$. The corresponding results are illustrated in Table 3 and Figure 6.

In Table 3, we observe that computing the test-statistics using the exact core-determining class are much faster than using the full set of inequalities. This result is not surprising because the exact core-determining class contains much less inequalities compared to the full set of inequalities. In Figure 6, the red curve represents the power of the statistical test (denoted as $T^{S^*}(t)$) suggested by Andrews and Shi (2014) with moment inequalities corresponding to the exact core-determining class, while the blue curve represents the power of the statistical test (denoted as $T^{full}(t)$) suggested by Andrews and Shi (2014) using the full set of inequalities.

We can find that, for function $S = S_1$, the power of T^{S^*} is much better than T^{full} for local alternatives $v' = v + \frac{t}{\sqrt{n}}\tilde{v}$; for function $S = S_3$, the power of T^{S^*} is also better than T^{full}, though the discrepancy is much smaller. This is because that, hueristically, the binding inequalities, i.e., those inequalities that correspond to the exact core-determining class, are more likely to generate the value of the S_3 function, compared to the unbinding inequalities. Or equivalently, if the unbinding inequalities are truely uninformative about the identified region, then with high probability, the maximum of $(\mathbb{E}_n[m_k(v|A)]/\sigma_A)_+^2$ over all non-empty set $A \subsetneq \mathcal{U}$ can be very close to the maximum of $(\mathbb{E}_n[m_k(v|A)]/\sigma_A)_+^2$ over all A in the exact core-determining class.

6. Conclusion

In this paper, we study the indeterministic relations between unobservable events and observed outcomes. A bipartite graph $G = (\mathcal{U}, \mathcal{Y}, \varphi)$ represents the relations. Given a probability measure on observed outcomes, the set of feasible probability measures on unobservable events can be defined by a set of linear inequality constraints, according to Artstein's Theorem. This set of inequalities is called the core-determining class.

	S_1		S_3	
t	$P_{S_1}^{\mathcal{S}^*}(t)$	$P_{S_1}^{full}(t)$	$P_{S_3}^{\mathcal{S}^*}(t)$	$P_{S_3}^{full}(t)$
0.0	0.1%	0.0%	1.4%	0.5%
0.1	0.4%	0.0%	1.7%	0.6%
0.2	0.4%	0.0%	1.5%	1.1%
0.3	0.7%	0.0%	3.2%	1.8%
0.4	1.2%	0.0%	3.4%	2.1%
0.5	2.2%	0.0%	4.5%	3.4%
0.6	3.6%	0.0%	5.9%	4.5%
0.7	5.1%	0.0%	8.6%	6.5%
0.8	8.4%	0.0%	13.7%	11.3%
0.9	12.4%	0.0%	20.2%	15.5%
1.0	20.5%	0.0%	25.7%	20.5%
1.1	25.1%	0.0%	34.3%	30.5%
1.2	34.6%	0.0%	40.6%	36.0%
1.3	45.8%	0.0%	52.7%	48.4%
1.4	50.6%	0.0%	64.7%	60.6%
1.5	64.2%	0.1%	72.9%	68.1%
1.6	72.8%	0.0%	82.0%	78.4%
1.7	80.6%	0.0%	85.5%	82.2%
1.8	86.3%	0.6%	92.4%	90.7%
1.9	92.9%	0.4%	95.3%	94.1%
2.0	95.1%	1.7%	96.9%	96.2%
2.1	97.4%	1.1%	97.9%	97.6%
2.2	98.6%	2.4%	99.2%	99.1%
2.3	98.9%	2.6%	99.6%	99.3%
2.4	99.6%	5.4%	99.5%	99.3%
2.5	99.8%	7.1%	99.9%	99.9%
2.6	100%	10.9%	100%	100%
2.7	100%	20.4%	100%	100%
2.8	100%	23.4%	100%	100%
2.9	100%	29.4%	100%	100%
3.0	100%	39.0%	100%	100%
Avg. Computational Time (min)	97	1147	96	1239

TABLE 3
Results of Power of Tests for Example 5

We aim to select the minimum set of irredundant inequalities, which is defined as the exact core-determining class. Specifically, we show that the exact core-determining class can be characterized by a set of combinatorial rules of the graph G. We then propose an algorithm that explores the structure of the bipartite graph to construct the exact core-determining class. We prove that, if the bipartite graph and the measure on observed outcomes are non-degenerate, the exact core-determining class does not depend on the observed probability measure on outcomes, but only on the structure of the bipartite graph.

We design a set of examples to implement the algorithm for the construction of the exact core-determining class. We demonstrate the good performance of the algorithm on the reduction of inequalities. We demonstrate in simulation that, on the one hand, finding the exact core-determining class can reduce the computational time of performing the inference procedure proposed by Andrews and Shi (2014). On the other hand, the power of the statistical tests based on the selected set of inequalities, i.e., the exact core-determining class, are better than the power of the statistical tests based on the full set of inequalities. That said, our approach could improve performance of the existing inference procedure, both computationally and statistically.

We require a set of supporting lemmas to prove the Theorem 1, 2, 3.

Appendix A: Proof of Lemmas and Auxliary Lemmas

Proof of Lemma 2.

When $|\mathcal{U}| = 1$, there is only one inequality $v(\mathcal{U}) \leqslant \mu(\mathcal{Y})$, so this inequality contains all the information on v. That said, the conclusion of Lemma 2 holds.

Now we assume that $|\mathcal{U}| \geqslant 2$.

(1) We prove that $Q_1(\mathcal{S}_u) = Q_1$.

First, for any $\mathcal{S} \subset 2^{\mathcal{U}}$, because the inequality restrictions that describe $Q_1(\mathcal{S})$ is a subset of those that describe Q_1, it must be that $Q_1(\mathcal{S}) \supset Q_1$. Consequently, $Q_1(\mathcal{S}_u) \supset Q_1$.

We need to show that $Q_1 \supset Q_1(\mathcal{S}_u)$. It is easy to see that $Q_1 := \cap_{A \subset \mathcal{U}, A \neq \emptyset}\{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A))\} = \cap_{A \subset \mathcal{U}, A \neq \emptyset} Q_1(\{A\})$. And for any $\mathcal{S}$, $Q_1(\mathcal{S}) = \cap_{A \in \mathcal{S}} Q_1(\{A\})$. Each $Q_1(\{A\}) := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A))\}$ is a convex polyhedron.

Denote $R_u := 2^{\mathcal{U}} \backslash (\mathcal{S}_u \cup \{\emptyset\})$ as the collection of non-empty subsets of $\mathcal{U}$ that are not in $\mathcal{S}_u$. Denote R_u as $\{\widetilde{A}_1, ..., \widetilde{A}_s\}$, where $s := |R_u| = 2^{d_u} - 1 - |\mathcal{S}_u|$, and each $\widetilde{A}_i$, $i = 1, 2, ..., s$, is a non-empty subset of $\mathcal{U}$.

For $k = 1, 2, ..., s$, consider $\mathcal{S}^k = \mathcal{S}_u \cup R_u^k$, where R_u^k can be any subset of R_u with $|R_u^k| = s - k$. We would like to show that $Q_1(\mathcal{S}^k) = Q_1$, for any possible R_u^k and for all $k = 1, 2, ..., s$.

When $k = 1$: For any R_u^1 with $R_u^1 \subset R_u$, $|R_u^1| = s - 1$, there must exists a set $\widetilde{A} \in R_u \backslash R_u^1$. Since $|R_u^1| = s - 1$, $\widetilde{A}$ is the only non-empty subset of $\mathcal{U}$ that is not in $\mathcal{S}_u \cup R_u^1$. Therefore,

$Q_1(\mathcal{S}^1) = \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A)), \text{ for all non-empty } A \subset \mathcal{U}, A \neq \widetilde{A}\}$.

By Definition 7, since $\widetilde{A} \notin \mathcal{S}_u$, for any $v \in Q_1(\mathcal{S}^1)$, it must be that $v(\widetilde{A}) \leqslant \mu(\varphi(\widetilde{A}))$; or equivalently, $v \in Q_1(\{\widetilde{A}\})$. Consequently, $Q_1(\mathcal{S}^1) \subset Q_1(\{\widetilde{A}\})$, and $Q_1 = Q_1(\mathcal{S}^1) \cap Q_1(\{\widetilde{A}\}) = Q_1(\mathcal{S}^1)$.

Now suppose the statement that $Q_1(\mathcal{S}^k) = Q_1$ holds for $k = k_0$ and for any $\mathcal{S}^{k_0} = \mathcal{S}_u \cup R_u^{k_0}$ with $|R_u^{k_0}| = s - k_0$, where $1 \leqslant k_0 \leqslant s - 1$. For $k = k_0 + 1$:

Consider $\mathcal{S}^{k_0+1} = \mathcal{S}_u \cup R_u^{k_0+1}$, where $R_u^{k_0+1}$ is a subset of R_u with cardinality $s - k_0 - 1$. Since $k_0 + 1 \geqslant 2$, there must exists two different sets $A_1, A_2 \in R_u$, and $A_1, A_2 \notin R_u^{k_0+1}$.

Therefore, $\mathcal{S}^{k_0+1} \cup \{A_1\} = \mathcal{S}_u \cup (R_u^{k_0+1} \cup \{A_1\})$, where $(R_u^{k_0+1} \cup \{A_1\})$ has cardinality $s - k_0$. Thus, $Q_1(\mathcal{S}^{k_0+1}) \cap Q_1(\{A_1\}) = Q_1(\mathcal{S}^{k_0+1} \cup \{A_1\}) = Q_1$.

Similarly, $Q_1(\mathcal{S}^{k_0+1}) \cap Q_1(\{A_2\}) = Q_1(\mathcal{S}^{k_0+1} \cup \{A_2\}) = Q_1$.

We claim the following proposition must be true.

Proposition 1. *Denote $W = Q_1(\mathcal{S}^{k_0+1}) \backslash Q_1$. Then, W must be empty.*

Proof of Proposition 1.

By definition, $Q_1(\mathcal{S}^{k_0+1}) = Q_1 \cup W$, $Q_1 \cap W = \emptyset$.

Since $Q_1 = Q_1(\mathcal{S}^{k_0+1}) \cap Q_1(A_1) = Q(\mathcal{S}^{k_0+1}) \cap Q_1(A_2)$, it easy to see that $W \cap Q_1(\{A_1\}) = (Q_1(\mathcal{S}^{k_0+1}) \backslash Q_1) \cap Q_1(\{A_1\}) = (Q_1(\mathcal{S}^{k_0+1}) \cap Q_1(A_1)) \backslash Q_1 = \emptyset$. Similarly, $W \cap Q_1(\{A_2\}) = \emptyset$. Therefore, $W \cap (Q_1(\{A_1\}) \cup Q_1(\{A_2\})) = \emptyset$.

If W is non-empty, then there exist $v_W \geqslant 0, v_W \in W$. It follows that $v_W \notin Q_1(\{A_1\})$, $v_W \notin Q_1(\{A_2\})$. By defintion of $Q_1(\{A_1\})$ and $Q_1(\{A_2\})$, it must be that $v_W(A_1) > \mu(\varphi(A_1))$ and $v_W(A_2) > \mu(\varphi(A_2))$.

Since μ is non-degenerate, for any non-empty $A \subset \mathcal{U}$, the inequality $v(A) \leqslant \mu(\varphi(A))$ can be re-written as $\sum_{i=1}^{d_u} 1(u_i \in A)v(u_i) \leqslant \mu(\varphi(A))$, where $\mu(\varphi(A)) = \sum_{y\in\varphi(A)} \mu(y) > 0$. In these inequalities, the coefficients on the vector $(v(u_1), ..., v(u_{d_u}))$ are non-negative and bounded from the above. Therefore, for $\epsilon > 0$ small enough, there exists an open box $B_\epsilon := \{v|v(u) \in (0, 2\epsilon),\ \text{for any } u \in \mathcal{U}\}$ such that $B_\epsilon \subset Q_1$. The center of this box is $v_\epsilon = (\epsilon, \epsilon, ..., \epsilon)$. Since $v_\epsilon \in B_\epsilon \subset Q_1$, it must be that $v_\epsilon(A_1) < \mu(\varphi(A_1))$, and $v_\epsilon(A_2) \leqslant \mu(\varphi(A_2))$, i.e., $v_\epsilon \in Q_1(\{A_1\}) \cap Q_1(\{A_2\})$.

Consider the segment line $L = \{v|v = tv_\epsilon + (1-t)v_W, t \in [0, 1]\}$. Since the set $Q_1(\mathcal{S}^{k_0+1})$ is convex and $v_W, v_\epsilon \in Q_1(\mathcal{S}^{k_0+1})$, any point on the segment L must be in $Q_1(\mathcal{S}^{k_0+1})$, i.e., $L \subset Q_1(\mathcal{S}^{k_0+1})$.

Because $v_W \notin Q_1(\{A_1\}) \cup Q_1(\{A_2\})$, $v_\epsilon \in Q_1 \subset Q_1(\{A_1\}) \cap Q_1(\{A_2\})$, and both $Q_1(\{A_1\})$ and $Q_1(\{A_2\})$ are closed and convex sets, there must exist $t_1 \in (0, 1]$ and $t_2 \in (0, 1]$ such that $v_t = tv_\epsilon + (1-t)v_W \in Q_1(\{A_1\})$ if and only if $t \in [t_1, 1]$, and $v = tv_\epsilon + (1-t)v_W \in Q_1(\{A_2\})$ if any only if $t \geqslant [t_2, 1]$. WLOG., we can assume that $t_1 \leqslant t_2$. For $t = t_1$, we have $v_{t_1}(A_1) = \mu(\varphi(A_1))$, and for $t = t_2$, we have $v_{t_2}(A_2) = \mu(\varphi(A_2))$.

Therefore, $v_{t_1} := t_1v_\epsilon + (1-t_1)v_W \in Q_1(\{A_1\}) \cap Q_1(\mathcal{S}^{k_0+1}) = Q_1$. It implies that $v \in Q_2(\{A_2\})$ since $Q_1 = Q_2(\{A_2\}) \cap Q_1(\mathcal{S}^{k_0+1})$. Thus, $t_1 \geqslant t_2$. Consequently, t_1 must equal to t_2.

Thus, $\widetilde{v} := v_{t_1} = v_{t_2} \in Q_1$ satisfies that: $\widetilde{v}(A_1) = \mu(\varphi(A_1))$, and $\widetilde{v}(A_2) = \mu(\varphi(A_2))$.

By construction of B_ϵ, we know that $B_\epsilon \subset Q_1 \subset Q_1(\mathcal{S}^{k_0+1})$. $\widetilde{B}_\epsilon := \{v|v = t_1v_1 + (1-t_1)v_W,\ \text{for any } v_1 \in B_\epsilon\}$ is an open set centered at $\widetilde{v}$. Since any elements in $\widetilde{B}_\epsilon$ is a linear combination of $v_1 \in B_\epsilon \subset Q_1(\mathcal{S}^{k_0+1})$ and $v_W \in Q_1(\mathcal{S}^{k_0+1})$ with $t_1 \in (0, 1]$, by convexity of $Q_1(\mathcal{S}^{k_0+1})$, $\widetilde{B}_\epsilon \subset Q_1(\mathcal{S}^{k_0+1})$.

Since $A_1 \neq A_2$ and $d_u = |\mathcal{U}| \geqslant 2$, there must exists a unit vector $v_3 \in \mathbb{R}^{d_u}$ such that $v_3(A_1) = 0$ and $v_3(A_2) > 0$. Since $\widetilde{B}_\epsilon$ is an open set in $\mathbb{R}^{d_u}$, for $t_3 > 0$ small enough, $v_4 = \widetilde{v} + t_3v_3 \in \widetilde{B}_\epsilon \subset Q_1(\mathcal{S}^{k_0+1})$. It is easy to see that $v_4(A_1) = \widetilde{v}(A_1) + t_3v_3(A_1) = \mu(\varphi(A_1))$ and $v_4(A_2) = \widetilde{v}(A_2) + t_3v_3(A_2) = \mu(\varphi(A_2)) + t_3v_3(A_2) > \mu(\varphi(A_2))$. By definition of $Q_1(\{A_1\})$ and $Q_1(\{A_2\})$, $v_4 \in Q_1(\mathcal{S}^{k_0+1}) \cap Q_1(\{A_1\}) = Q_1$, but $v_4 \notin Q_1(\{A_2\})$.

However, we know that $Q_1 \subset Q_1(\{A_2\})$, so the existence of v_4 leads to a contradiction. Thus, $W = \emptyset$.

□

By Proposition 1, $W = \emptyset$, so $Q_1(\mathcal{S}^{k_0+1}) = Q_1 \cup W = Q_1$.

By induction, $Q_1(\mathcal{S}^k) = Q_1$ holds for all $k = 1, 2, ..., s$ and $\mathcal{S}^k = \mathcal{S}_u \cup R_u^k$, where R_u^k can be any possible subset of R_u with cardinality $s - k$. When $k = s$, $R_u^k = \emptyset$. Therefore, $Q_1(\mathcal{S}_u) = Q_1(\mathcal{S}^s) = Q_1$.

(2) It is obviouse that $Q_1(\mathcal{S}) \supset Q_1$, since $Q_1(\mathcal{S})$ is defined by a subset of constraints that define Q_1.

For any $\mathcal{S}$ such that it does not contain $\mathcal{S}_u$, there must exist $A_0 \in \mathcal{S}_u$ such that

$A_0 \notin \mathcal{S}$. By Definition 7, $v^M_{A_0}(A_0) > \mu(\varphi(A_0))$, which implies that $v^M_{A_0} \notin Q_1$. By Definition 6, $v^M_{A_0} \in Q_1(2^{\mathcal{U}} \backslash \{\emptyset, A_0\})$. Since $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{\emptyset, A_0\}$, it holds that $Q_1(\mathcal{S}) \supset Q_1(2^{\mathcal{U}} \backslash \{\emptyset, A_0\})$. Therefore, $v^M_{A_0} \in Q_1(\mathcal{S})$, so $v^M_{A_0}$ is an element in $Q_1(\mathcal{S})$ but not in Q_1, i.e., $Q_1(\mathcal{S}) \supsetneq Q_1$.

□

Proof of Lemma 3.

By Definition 8, for any $A \notin \mathcal{S}_u^{'}$, one of the following two statements must be true:

(1) $\exists A_1, A_2 \subset A, A_1, A_2 \neq \emptyset, A_1 \cup A_2 = A$ such that $\varphi(A_1) \cap \varphi(A_2) = \emptyset$;

(2) $\exists u \in \mathcal{U}$ such that $u \notin A$, and $\varphi(u) \subset \varphi(A)$.

If (1) is true, by Definition 7, for every $A' \neq A$, $v^M_A(A') \leqslant \mu(\varphi(A'))$. Therefore, $v^M_A(A_i) \leqslant \mu(\varphi(A_i)), i = 1, 2$. Then, $v^M_A(A) = v^M_A(A_1 \cup A_2) = v^M_A(A_1) + v^M_A(A_2) \leqslant \mu(\varphi(A_1)) + \mu(\varphi(A_2)) = \mu(\varphi(A_1) \cup \varphi(A_2)) = \mu(\varphi(A_1 \cup A_2)) = \mu(\varphi(A))$. By Definition 7, $A \notin \mathcal{S}_u$.

If (2) is true, then $\varphi(A \cup \{u\}) = \varphi(A)$ and it follows that $\mu(\varphi(A \cup \{u\})) = \mu(\varphi(A))$. Thus, by Definition 6, $v^M_A(A \cup \{u\}) \leqslant \mu(\varphi(A \cup \{u\})) = \mu(\varphi(A))$. Therefore, $v^M_A(A) \leqslant v^M_A(A \cup \{u\}) \leqslant \mu(\varphi(A \cup \{u\})) = \mu(\varphi(A))$. By Definition 7, $A \notin \mathcal{S}_u$.

Consequently, combining the discussions above, we have:

$$\mathcal{S}_u \subset \mathcal{S}'_u. \tag{A.4}$$

Based on (A.4), to prove the statement of the Lemma that $\mathcal{S}_u = \mathcal{S}'_u$, we need to show that $\mathcal{S}'_u \subset \mathcal{S}_u$.

Consider an arbitrary non-empty set $A \subset \mathcal{U}, A \neq \mathcal{U}$ such that $A \notin \mathcal{S}_u$. It is sufficient to prove that $A \notin \mathcal{S}'_u$ to show $\mathcal{S}'_u \subset \mathcal{S}_u$. We prove the above statement by contradiction.

By contradiction, we assume that there exists a set $A \in \mathcal{S}'_u \backslash \mathcal{S}_u$. Therefore, A satisfies the two properties defined in Definition 8.

Denote $\mathcal{S}_u := \{A_i | 1 \leqslant i \leqslant r := |\mathcal{S}_u|\}$. For every set $A \subset \mathcal{U}$, we can define a vector $w(A) \in \{0, 1\}^{d_u}$ such that the i^{th} entry of $w(A)$ equals to 1 if and only if $u_i \in A$. Therefore, $w : A \to \{0, 1\}^{d_u}$ is a mapping from the $2^{\mathcal{U}}$ to $\{0, 1\}^{d_u}$. Such $w(A)$ is often refered as the characteristic vector of the set A.

By Lemma 2, the inequalities $v(A_1) \leqslant \mu(\varphi(A_1)), ..., v(A_r) \leqslant \mu(\varphi(A_r))$ are irredundant and $Q_1(\mathcal{S}_u) := \{v | v(A_i) \leqslant \mu(\varphi(A_i))\}$ is the same as $Q_1 := \{v | v(A) \leqslant \mu(\varphi(A)), A \subset \mathcal{U}, A \neq \emptyset\}$.

By Farkas Lemma, there exists a $1 \times r$ vector $\pi \geqslant 0$ such that

$$(a) \sum_{i=1}^{r} \pi_i w(A_i) \geqslant w(A), (b) \sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) \leqslant \mu(\varphi(A)). \tag{A.5}$$

Since the vector $w(A_i)$ has some entries which equal to one, at least one π_i must be positive, $i \in \{1, 2, ..., r\}$. WLOG., assume $\pi_i > 0$, $i = 1, 2, ..., r'$, where $1 \leqslant r' \leqslant r$, and $\pi_i = 0$ for $r' < i \leqslant r$; We would simply omit the A_i that corresponds to $\pi_i = 0$ in the statements above.

Proposition 2. *The vector $\pi > 0$ satisfies:*

(1) For any $y \in \mathcal{Y}$, $\sum_{i=1}^{r'} 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$ and for any $u \in \mathcal{U}$, $\sum_{i=1}^{r'} 1(u \in A_i) = 1(u \in A)$.

(2) For any $y \in \mathcal{Y}$, either $(\varphi^{-1}(y) \cap A_i) \cap A = \emptyset$; or $(\varphi^{-1}(y) \cap A_i) \subset A$, $i = 1, 2, ..., r'$.

Proof of Proposition 2. (1) Proof of statement (1).

Since $\sum_{i=1}^{r'} \pi_i w(A_i) \geqslant w(A)$, it implies that for any $u \in \mathcal{U}$, $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) \geqslant 1(u \in A)$.

If $A \cap \varphi^{-1}(y) \neq \emptyset$, i.e., $y \in \varphi(A)$, pick any $u \in A \cap \varphi^{-1}(y)$, we have:

$$\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) \geqslant \sum_{i=1}^{r'} \pi_i 1(u \in A_i) \geqslant 1(u \in A) = 1 = 1(A \cap \varphi^{-1}(y) \neq \emptyset). \tag{A.6}$$

If $A \cap \varphi^{-1}(y) = \emptyset$, i.e., $y \notin \varphi(A)$, then $\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) \geqslant 0 = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$. Therefore, the inequality (A.6) holds for any $y \in \mathcal{Y}$.

By applying inequality (A.6) on every $y \in \mathcal{Y}$, we have:

$$\sum \pi_i \mu(\varphi(A_i)) = \sum_{y \in \mathcal{Y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) \geqslant \sum_{y \in \mathcal{Y}} \mu(y) 1(A \cap \varphi^{-1}(y) \neq \emptyset) = \mu(\varphi(A)). \tag{A.7}$$

By assumption, $\mu(\cdot)$ is non-degenerate, i.e., $\mu(y) > 0$ for any $y \in \mathcal{Y}$. Combining (A.7) with statement (b) of (A.5), the inequality (A.7) must hold as an equality, i.e., for any $y \in \mathcal{Y}$, $\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) = 1(A \cap \varphi^{-1}(y) \neq \emptyset)$. Therefore, the first part of Statement (1) is proven.

We claim that the inequality $\sum_{i=1}^{r'} \pi_i w(A_i) \geqslant w(A)$ in statement (a) of (A.5) must hold as an equality, i.e.,

$$\sum_{i=1}^{r'} \pi_i w(A_i) = w(A). \tag{A.8}$$

Or equivalently, for any $u \in \mathcal{U}$, $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) = 1(u \in A)$.

To prove the equation (A.8), let $\widetilde{w} = \sum_{i=1}^{r'} w(A_i)$. If, by contradiction, $\widetilde{w} \neq w(A)$, there can only be two cases: (a) there exists $u \notin A$, such that $u \in A_{i_0}$ for some $i_0 \in \{1, 2, ..., r'\}$. (b) there exists $u \in A$, such that $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) > 1$.

For case (a), by property (2) of Definition 8, $\varphi(u)$ is not a subset of $\varphi(A)$. Therefore, there exists $\widetilde{y}$ such that $\widetilde{y} \in \varphi(u)$ but $y \notin \varphi(A)$.

Let us recall inequality (A.7).

$$\sum \pi_i \mu(\varphi(A_i)) = \sum_{y \in \mathcal{Y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \mu(\widetilde{y}) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(\widetilde{y}) \neq \emptyset) + \sum_{y \in \varphi(A)} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \pi_{i_0} \mu(\widetilde{y}) + \sum_{y \in \varphi(A)} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \pi_{i_0} \mu(\widetilde{y}) + \sum_{y \in \varphi(A)} \mu(y) 1(\varphi^{-1}(y) \cap A \neq \emptyset)$$

$$= \pi_{i_0} \mu(\widetilde{y}) + \mu(\varphi(A)) > \mu(\varphi(A)),$$

which contradicts to (A.5).

For case (b), pick a $\widetilde{y} \in \varphi(u) \subset \varphi(A)$. Again let us recall inequality (A.7).

$$\sum \pi_i \mu(\varphi(A_i)) = \sum_{y \in \mathcal{Y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \mu(\widetilde{y}) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) + \sum_{y \in \varphi(A), y \neq \widetilde{y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \mu(\widetilde{y}) \sum_{i=1}^{r'} \pi_i 1(u \in A_i) + \sum_{y \in \varphi(A), y \neq \widetilde{y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$> \mu(\widetilde{y}) \cdot 1 + \sum_{y \in \varphi(A), y \neq \widetilde{y}} \mu(y) \sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)$$

$$\geqslant \mu(\widetilde{y}) + \sum_{y \in \varphi(A), y \neq \widetilde{y}} \mu(y)$$

$$= \mu(A),$$

which contradicts to (A.5).

Thus, equation (A.8) must hold. Therefore, for any $u \in \mathcal{U}$, $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) = 1(u \in A)$, i.e., the second part of Statement (1) is true.

(2) Proof of Statement (2).

We claim that for any $y \in \mathcal{Y}$, either $(\varphi^{-1}(y) \cap A) \subset A_i$ or $(\varphi^{-1}(y) \cap A) \cap A_i = \emptyset$ for all i. We prove this argument by contradiction.

Assume that there exists a $y \in \mathcal{Y}$ and $1 \leqslant i_1 \leqslant r'$ such that $(\varphi^{-1}(y) \cap A) \cap A_{i_1} \neq \emptyset$, and $(\varphi^{-1}(y) \cap A)$ is not a subset of A_{i_1}. Therefore, there exists $u \neq u'$ such that $u, u' \in \varphi^{-1}(y)$, $u \in A \cap A_{i_1}$, $u' \in A$ but $u' \notin A_{i_1}$. Thus,

$\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) = \pi_{i_1} + \sum_{j \neq i_1} \pi_j 1(A_j \cap \varphi^{-1}(y) \neq \emptyset) \geqslant \pi_{i_1} + \sum_{j \neq i_1} \pi_j 1(u' \in A_j) = \pi_{i_1} - \pi_{i_1} 1(u' \in A_i) + \sum_{j=1}^{r'} \pi_j 1(u' \in A_j) \geqslant \pi_{i_1} - 0 + 1(u' \in A) = \pi_{i_1} + 1 > 1 = 1(A \cap \varphi^{-1}(y))$, which is contradictory with Statement (1) of this proposition.

Thus, for any $y \in \mathcal{Y}$,

$$\text{either } (\varphi^{-1}(y) \cap A) \subset A_i \text{ or } (\varphi^{-1}(y) \cap A) \cap A_i = \emptyset, \tag{A.9}$$

for all $i = 1, 2, ..., r'$.

□

By Proposition 2, given the equation (A.8), since $\pi_i > 0$ for all $i = 1, 2, ..., r'$, it must be that $A_i \subset A$ for any $i = 1, 2, ..., r'$.

Therefore, for any $y \in \varphi(A_i)$, $\varphi^{-1}(y) \cap A_i \neq \emptyset$, and $\varphi^{-1}(y) \cap A_i \subset A$. Then, since $\varphi^{-1}(y) \cap A \cap A_i \neq \emptyset$, it must be that $\varphi^{-1}(y) \cap A_i = \varphi^{-1}(y) \cap A \cap A_i \subset A_i$. It implies that $\varphi^{-1}(y) \subset A_i$. Thus, $\varphi^{-1}(\varphi(A_i)) = \cup_{y \in \varphi(A_i)} \varphi^{-1}(y) \subset A_i$. On the other hand, $\varphi^{-1}(\varphi(A_i)) \supset A_i$ by definition of φ and φ^{-1}, so

$$\varphi^{-1}(\varphi(A_i)) = A_i. \tag{A.10}$$

By assumption, $A \notin \mathcal{S}_u$. Therefore, $A \neq A_i$ for any $i = 1, 2, ..., r'$. Since $A \in \mathcal{S}'_u$, A is self-connected. By self-connectivity of A, $\varphi(A \backslash A_i) \cap \varphi(A_i) \neq \emptyset$. Consequently, there must exist $u \in A \backslash A_i$ such that $\varphi(u) \cap \varphi(A_i) \neq \emptyset$. Thus, $u \in \varphi^{-1}(\varphi(A_i))$ but $u \notin A_i$, which contradicts to (A.10).

Therefore, there exists no $A \in \mathcal{S}'_u \backslash \mathcal{S}_u$. It implies that $\mathcal{S}'_u \subset \mathcal{S}_u$. Combining this statement with (A.4), we have $\mathcal{S}_u = \mathcal{S}'_u$. □

Appendix B: Proof of Main Theorems

B.1. Proof of Theorem 1

The following proposition characterize the dimensionality of Q.

Lemma 6. *Suppose both G and μ are non-degenerate. Then, Q contains an open set in a $(d_u - 1)$ dimensional linear subspace of $\mathbb{R}^{d_u}$.*

Proof of Lemma 6.

Consider $Q_2 := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A)), \text{ for any } A \subset \mathcal{U}, A \neq \mathcal{U}, A \neq \emptyset\}$.

Let us enumerate the elements of $\mathcal{U}$ as $u_1, ..., u_{d_u}$, and let us enumerate the elements of $2^{\mathcal{U}} \backslash \{\emptyset, \mathcal{U}\}$ as $A'_1, A'_2, ..., A'_r$, with $r := 2^{d_u} - 2$. For any $A \subset \mathcal{U}$, we can define $w(A) \in \{0, 1\}^{d_u}$ as the characteristic vector of A, i.e., the i^{th} entry of $w(A)$ equals to 1 if and only $u_i \in A$.

We claim that the following proposition must be true.

Proposition 3. *Suppose both G and μ are non-degenerate. Then, there exists a vector $v \in Q_2$ such that $v(\mathcal{U}) > \mu(\mathcal{Y})$.*

Proof of Proposition 3.. We prove the statement of this proposition by contradiction.

Suppose the statement of Proposition 3 is untrue, i.e., $Q_2 \subset \{v \geqslant 0 | v(\mathcal{U}) \leqslant \mu(\mathcal{Y})\}$.

By Farkas Lemma, there must exists a vector $\pi \in \mathbb{R}^r$, $\pi \geqslant 0$ such that:

$$\sum_{i=1}^{r} \pi_i w(A_i') \geqslant w(\mathcal{U}) = (1, 1, ..., 1), \tag{B.11}$$

and

$$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i')) \leqslant \mu(\mathcal{Y}). \tag{B.12}$$

Since $w(\mathcal{U})$ has some entries that equal to one, at least one π_i must be positive, $i \in \{1, 2, ..., r\}$. WLOG., we can assume that all $\pi_i > 0$, $i = 1, 2, ..., r'$, with $1 \leqslant r' \leqslant r$, and $\pi_i = 0$ for $i > r'$; for those $\pi_i = 0$, we can just drop those indices i that corresponds to $\pi_i = 0$.

(B.11) implies that for any $u \in \mathcal{U}$,

$$\sum_{i=1}^{r'} \pi_i 1(u \in A_i') \geqslant 1. \tag{B.13}$$

(B.12) can be re-rewritten as:

$$\sum_{y \in \mathcal{Y}} \mu(y) (\sum_{i=1}^{r'} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1) \leqslant 0. \tag{B.14}$$

For any $y \in \mathcal{Y}$, we can pick a $u \in \varphi^{-1}(y)$. It is easy to see that $1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) \geqslant 1(u \in A_i')$ holds for all $i = 1, 2, ..., r'$. Therefore, by (B.13),

$$\sum_{i=1}^{r'} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1 \geqslant \sum_{i=1}^{r'} \pi_i 1(u \in A_i') - 1 \geqslant 0. \tag{B.15}$$

Plug in (B.15) in the left hand side of (B.14), we have:
$\sum_{y \in \mathcal{Y}} \mu(y) (\sum_{i=1}^{r'} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1) \geqslant 0$. Therefore, (B.14) must hold as an equality. Since μ is non-degenerate, i.e., $\mu(y) > 0$ for all $y \in \mathcal{Y}$, it must be that:

$$\sum_{i=1}^{r'} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1 = 0. \tag{B.16}$$

For any $y \in \mathcal{Y}$, if there exists A_{i_0}' for some $i_0 \in \{1, 2, ..., r'\}$, such that $\varphi^{-1}(y) \cap A_{i_0}' \neq \emptyset$, but $\varphi^{-1}(y)$ is not a subset of A_{i_0}, then there must exist $u_1, u_2 \in \varphi^{-1}(y)$, such that $u_1 \in A_{i_0}$, and $u_2 \notin A_{i_0}$.

Therefore, $\sum_{i=1}^{r'} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1$
$= \pi_{i_0} + \sum_{1 \leqslant i \leqslant r', i \neq i_0} \pi_i 1(A_i' \cap \varphi^{-1}(y) \neq \emptyset) - 1$
$\geqslant \pi_{i_0} - \pi_{i_0} 1(u_2 \in A_{i_0}) + \sum_{i=1}^{r'} \pi_i 1(u_2 \in A_i') - 1$
$\geqslant \pi_{i_0} - 0 + 1 - 1 > 0$, which contradicts to (B.16). Therefore, for any $y \in \mathcal{Y}$ and any $i = 1, 2, ..., r'$, either

$$\varphi^{-1}(y) \subset A_i' \text{ or } \varphi^{-1}(y) \cap A_i' = \emptyset. \tag{B.17}$$

It follows that, for any $i = 1, 2, ..., r'$, $\varphi(A'_i) \cap \varphi(\mathcal{U} \backslash A'_i) = \emptyset$, because otherwise there exists a y such that $\varphi^{-1}(y) \cap A'_i \neq \emptyset$ and $\varphi^{-1}(y) \cap (\mathcal{U} \backslash A'_i) \neq \emptyset$, which violates (B.17). However, by assumption G is connected, it must be that $\varphi(A'_i) \cap \varphi(\mathcal{U} \backslash A'_i) \neq \emptyset$, contradiction!

Therefore, Q_2 must not be a subset of $\{v \geqslant 0 | v(\mathcal{U}) \leqslant \mu(\mathcal{Y})\}$; or equivalently, there must exists a $v \in Q_2$ such that $v(\mathcal{U}) > \mu(\mathcal{Y})$. □

By Proposition 3, there exists a $v_3 \in Q_2$, such that $v_3(\mathcal{U}) > \mu(\mathcal{Y})$.

Since μ is non-degenerate, so for any non-empty $A \subset \mathcal{U}$, the inequality $v(A) \leqslant \mu(\varphi(A))$ can be re-written as $\sum_{i=1}^{d_u} 1(u \in A) v(u_i) \leqslant \mu(\varphi(A))$, where $\mu(\varphi(A)) > 0$. All of these inequalities have non-negative and bounded coefficients on the vector $(v(u_1), ..., v(u_{d_u}))$. Therefore, there must exist an $\epsilon > 0$ small enough such that the open box $B_\epsilon := \{v | v(u) \in (0, 2\epsilon) \text{ for all } u \in \mathcal{U}\}$ satisifies that $B_\epsilon \subset Q_2 \cap \{v | v(\mathcal{U}) < \mu(\mathcal{Y})\}$. By convexity of Q_2, for any $t \in [0, 1]$ and $v_\epsilon \in B_\epsilon$, we have $tv_3 + (1-t)v_\epsilon \in Q_2$. Define the set $\mathcal{C} := \{v_4 | v_4 := tv_3 + (1-t)v_\epsilon, v_\epsilon \in B_\epsilon, t \in (0, 1)\}$, so $\mathcal{C}$ is an open set in $\mathbb{R}^{d_u}$. Since $B_\epsilon \subset \{v \geqslant 0 | v(\mathcal{U}) < \mu(\mathcal{Y})\}$, and $v_3 \notin \{v \geqslant 0 | v(\mathcal{U}) \leqslant \mu(\mathcal{Y})\}$, and B_ϵ is an open box in $\mathbb{R}^{d_u}$, then $\mathcal{C} \cap \{v | v(\mathcal{U}) = \mu(\mathcal{Y})\}$ contains an open set in the linear subspace $\{v | v(\mathcal{U}) = \mu(\mathcal{Y})\}$.

Since $Q = Q_2 \cap \{v | v(\mathcal{U}) = \mu(\mathcal{Y})\}$, and $Q_2 \supset \mathcal{C}$, it follows that $Q \supset (\mathcal{C} \cap \{v | v(\mathcal{U}) = \mu(\mathcal{Y})\})$. Therefore, Q contains an open set in a $d_u - 1$ dimensional linear subspace of $\mathbb{R}^{d_u}$. □

With Lemma 6 and Proposition 3, we can start to prove Theorem 1.

Proof of Theorem 1.

(1) For all $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{\emptyset\}$, because $Q(\mathcal{S})$ is defined by a subset of inequality restrictions that define Q, it is easy to see that $Q(\mathcal{S}) \supset Q$.

Obvioully, $\mathcal{U} \notin \mathcal{S}^*$, since $v(\mathcal{U}) \leqslant \mu(\mathcal{Y})$ is implied by $v(\mathcal{U}) = \mu(\mathcal{Y})$.f

When $d_u = 1$, there is only one inequality $v(\mathcal{U}) \leqslant \mu(\mathcal{Y})$, and it is implied by $v(\mathcal{U}) = \mu(\mathcal{Y})$. Therefore, $\mathcal{S}^*$ is $\emptyset$ and obviously it is the exact core-determining class.

Now we assume that $d_u = |\mathcal{U}| \geqslant 2$.

Denote $R_u := 2^{\mathcal{U}} \backslash (\mathcal{S}^* \cup \{\emptyset, \mathcal{U}\})$ as the collection of non-empty subsets of $\mathcal{U}$ that are not in $\mathcal{S}^* \cup \{\mathcal{U}\}$. Denote R_u as $\{A_1, ..., A_s\}$, where $s := |R_u| = 2^{d_u} - 2 - |\mathcal{S}^*|$, and each A_i is a subset of $\mathcal{U}$, $i = 1, 2, ..., s$.

Define $\mathcal{S}^{*k} := \mathcal{S}^* \cup R_u^k$, where R_u^k can be any subset of R_u with $|R_u^k| = s - k$. For $k = 1, 2, ..., s$, we claim that $Q(\mathcal{S}^{*k}) = Q$, We prove this claim by performing induction on k.

For $k = 1$, for any $R_u^1 \subset R_u$, $|R_u^1| = s - 1$, denote $\widetilde{A}$ as the only element that is in R_u but not in R_u^1.

Since $\widetilde{A} \notin \mathcal{S}^*$, by Definition 5, $v(\widetilde{A}) \leqslant \mu(\varphi(\widetilde{A}))$ for all $v \in Q(\mathcal{S}^{*1})$, where $\mathcal{S}^{*1} = \mathcal{S}^* \cup R_u^1$. It implies that $Q(\mathcal{S}^{*1}) \subset Q(\{\widetilde{A}\})$. Therefore, $Q = Q(\mathcal{S}^{*1}) \cap Q(\{\widetilde{A}\}) = Q(\mathcal{S}^{*1})$; or equivalently, $Q(\mathcal{S}^* \cup R_u^k) = Q$ holds for $k = 1$ and any possible R_u^k.

Now, suppose $Q(\mathcal{S}^* \cup R_u^k) = Q$ holds for all possible R_u^k with $k = k_0$, where $1 \leqslant k_0 \leqslant s - 1$. For $k = k_0 + 1$:

Consider $\mathcal{S}^{*(k_0+1)} = \mathcal{S}^* \cup R_u^{k_0+1}$, where $R_u^{k_0+1}$ can be any subset of R_u with $|R_u^{k_0+1}| = s - k_0 - 1$. Since $k_0 + 1 \geqslant 2$, there must exist two different sets $\in R_u \backslash R_u^{k_0+1}$. WLOG., let these two sets be A_1 and A_2.

It is easy to see that $R_u^{k_0+1} \cup \{A_1\}$ and $R_u^{k_0+1} \cup \{A_2\}$ have cardinality $s - k_0$. Therefore, $Q(\mathcal{S}^{*(k_0+1)}) \cap Q(\{A_1\}) = Q(\mathcal{S}^{*(k_0+1)} \cup \{A_1\}) = Q(\mathcal{S}^* \cup (R_u^{k_0+1} \cup \{A_1\})) = Q$, and $Q(\mathcal{S}^{*(k_0+1)}) \cap Q(\{A_2\}) := Q(\mathcal{S}^{*(k_0+1)} \cup \{A_2\}) = Q(\mathcal{S}^* \cup (R_u^{k_0+1} \cup \{A_2\})) = Q$.

Let us decompose $Q(\mathcal{S}^{*(k_0+1)})$ as $Q(\mathcal{S}^{*(k_0+1)}) = Q \cup W$, where $Q \cap W = \emptyset$. Since $Q(\mathcal{S}^{*(k_0+1)}) \cap Q(\{A_1\}) = Q(\mathcal{S}^{*(k_0+1)}) \cap Q(\{A_2\}) = Q$, it must be that $W \cap (Q(\{A_1\}) \cup Q(\{A_2\})) = \emptyset$.

We claim that the following proposition must be true.

Proposition 4. *If W is non-empty, then Q must be embedded in a $d_u - 2$ dimensional linear space, i.e., Q is a subset of a $d_u - 2$ dimensional linear space.*

Proof of Proposition 4.

Suppose W is non-empty.

Then, there exists $v_W \in W \subset Q(\mathcal{S}^{*(k_0+1)})$. Since $v_W \notin Q(\{A_1\})$, $v_W \notin Q(\{A_2\})$, it must be that $v_W(A_1) > \mu(\varphi(A_1))$ and $v_W(A_2) > \mu(\varphi(A_2))$. Since $Q(\mathcal{S}^{*(k_0+1)})$ is convex, then for any $v \in Q$ and $t \in [0, 1]$, $v_t := tv + (1 - t)v_W \in Q(\mathcal{S}^{*(k_0+1)})$. Therefore, there must exist unique $t_1 \in (0, 1]$ and $t_2 \in (0, 1]$ such that $v_{t_1}(A_1) = \mu(\varphi(A_1))$ and $v_{t_2}(A_2) = \mu(\varphi(A_2))$.

By definition of t_1 and t_2, we have: $v_{t_1} \in Q(\{A_1\}) \cap Q(\mathcal{S}^{*(k_0+1)}) = Q$, and $v_{t_2} \in Q(\{A_2\}) \cap Q(\mathcal{S}^{*(k_0+1)}) = Q$. Thus, $v_{t_1} \in Q \subset Q(\{A_2\})$, which implies that $t_1 \geqslant t_2$. Similarly, $t_2 \geqslant t_1$. Therefore, it must be that $t_1 = t_2$.

Consequently, for any $v \in Q$, there exists a $t = t_1 = t_2 \in (0, 1]$ such that $v_t := tv + (1 - t)v_W$ satisfies: $v_t(A_1) = \mu(\varphi(A_1))$ and $v_t(A_2) = \mu(\varphi(A_2))$. By convexity of $Q(\mathcal{S}^{*(k_0+1)})$, it must be that $v_t \in Q(\mathcal{S}^{*(k_0+1)})$, and $v_t(\mathcal{U}) = tv(\mathcal{U}) + (1 - t)v_W(\mathcal{U}) = t\mu(\mathcal{Y}) + (1 - t)\mu(\mathcal{Y}) = \mu(\mathcal{Y})$. The three equations mentioned above can be viewed as a system of linear equations of the vector $v_t = (v_t(u), u \in \mathcal{U})$. There can be two cases:

Case (a): A_1, A_2 is a partition of $\mathcal{U}$, i.e., $A_1 \cup A_2 = \mathcal{U}$, $A_1 \cap A_2 = \emptyset$. Since $t = t_1 = t_2$, by definition of t_1 and t_2, $v_t(A_1) = \mu(\varphi(A_1))$ and $v_t(A_2) = \mu(\varphi(A_2))$. By assumption that G is non-degenerate, we have $\varphi(A_1) \cap \varphi(A_2) \neq \emptyset$, and $\varphi(A_1) \cup \varphi(A_2) = \mathcal{Y}$. Therefore, $v_t(\mathcal{U}) = v_t(A_1) + v_t(A_2) = \mu(\varphi(A_1)) + \mu(\varphi(A_2)) = \sum_{y \in \varphi(A_1)} \mu(y) + \sum_{y \in \varphi(A_2)} \mu(y) = \sum_{y \in \mathcal{Y}} \mu(y) + \sum_{y \in \varphi(A_1) \cap \varphi(A_2)} \mu(y) > \mu(\mathcal{Y})$. This is contradictory to $v_t(\mathcal{U}) = \mu(\varphi(\mathcal{Y}))$. Thus, case (a) can not happen.

Case (b): A_1, A_2 is not a partition of $\mathcal{U}$: the vector $v_t = (v_t(u), u \in \mathcal{U})$ satisfies three linear equations with independent coefficients: $v_t \in \{v' | v'(A_1) = \mu(\varphi(A_1)), v'(A_2) = \mu(\varphi(A_2)), v'(\mathcal{U}) = \mu(\mathcal{Y})\}$.

In this case, $\mathcal{U}$ must be larger than 2: if $|\mathcal{U}| = 2$ it must be that $A_1 = \{u_1\}$ and $A_2 = \{u_2\}$ or $A_1 = \{u_2\}$ and $A_2 = \{u_1\}$. In any situations, A_1, A_2 form a partition of $\mathcal{U}$.

Hence, $|\mathcal{U}| \geqslant 3$, and v_t satifies three linearly independent equations. Therefore, v_t must live in a $d_u - 3$ dimensional linear space.

Since $v_t := tv + (1-t)v_W$, for $t' = \frac{1}{t} \in \mathbb{R}$, we have $v = t'v_t + (1-t')v_W$. Because v_W is a fixed point, $t' \in \mathbb{R}$, and v_t lives in a $d_u - 3$ dimensional linear space, v must live in a $d_u - 2$ dimensional linear space. That said Q is embedded in a $d_u - 2$ dimensional linear space. □

By Lemma 6, Q contains an open set in a $d_u - 1$ dimensional linear subspace of $\mathbb{R}^{d_u}$. If W is non-empty, by Proposition 4, Q is contained in a $d_u - 2$ dimensional linear space. We know that an open set in $d_u - 1$ dimensional linear space can not be a subset of a $d_u - 2$ dimensional linear space, therefore, W must be empty!

Thus, $Q(\mathcal{S}^{*(k_0+1)}) = Q$, then the induction is complete.

By induction, let $k = s$, so $Q(\mathcal{S}^*) = Q(\mathcal{S}^{*s}) = Q$.

(2) For any $\mathcal{S}$ that does not contain $\mathcal{S}^*$, there exists a set $A \neq \emptyset$ such that $A \in \mathcal{S}^* \backslash \mathcal{S}$. It is easy to see that $Q(\mathcal{S}) \supset Q(2^{\mathcal{U}} \backslash \{\emptyset, A\})$, because $\mathcal{S} \subset 2^{\mathcal{U}} \backslash \{\emptyset, A\}$. Since $A \in \mathcal{S}^*$, by Definition 5, there exists $v_A^{M^*} \in Q(2^{\mathcal{U}} \backslash \{\emptyset, A\})$ such that $v_A^{M^*}(A) > \mu(\varphi(A))$, i.e., $v_A^{M^*} \notin Q$. Therefore, $Q(\mathcal{S}) \supset Q(2^{\mathcal{U}} \backslash \{\emptyset, A\}) \supsetneq Q$; or equivalently, for any $\mathcal{S}$ such that $Q(\mathcal{S}) = Q$, it must be that $\mathcal{S} \supset \mathcal{S}^*$.

(3) By statement (2) of Theorem 1, any exact core-determining class must contain $\mathcal{S}^*$. By statement (1) of Theorem 1, $Q(\mathcal{S}^*) = Q$. Therefore, $\mathcal{S}^*$ is the only exact core-determining class due to the minimum cardinality property in Definition 4. □

B.2. Proof of Theorem 2

Proof of Lemma 5.

We divide the proof of the theorem into two steps.

• Step 1: we prove that $\mathcal{S}^* \subset (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$.

By definition, $\mathcal{S}^*$ corresponds to the set of inequalities that are irredundant amongst all inequalities $v(A) \leqslant \mu(\varphi(A))$, $A \subset \mathcal{U}, A \neq \emptyset, A \neq \mathcal{U}$, under the constraint that $v(\mathcal{U}) = \mu(\mathcal{Y}) > 0$. Obviously, $\mathcal{U} \notin \mathcal{S}^*$ because $v(\mathcal{U}) \leqslant \mu(\mathcal{Y})$ is implied by the equation $v(\mathcal{U}) = \mu(\mathcal{Y})$.

For any $A \notin \mathcal{S}_u$ and $A \neq \mathcal{U}$, by Definition , we know that either (a) there exists a $u \in \mathcal{U} \backslash A$ such that $\varphi(u) \subset A$ or (b) A is not self-connected.

For case (a), by Definition 5, for every $A' \neq A$, $v_A^M(A') \leqslant \mu(\varphi(A'))$. Therefore, $v_A^{M^*}(A) \leqslant v_A^{M^*}(A \cup \{u\}) \leqslant \mu(\varphi((A \cup \{u\})) = \varphi(A)$. Thus, $A \notin \mathcal{S}^*$.

For case (b), since A is not self-connected, there exists two non-empty sets A_1 and A_2 such that $A_1 \cup A_2 = A$, $A_1 \cap A_2 = \emptyset$, $\varphi(A_1) \cap \varphi(A_2) = \emptyset$, $\varphi(A_1) \cup \varphi(A_2) = A$. Again, by Definition 5, we have: $v_A^{M^*}(A_1) \leqslant \mu(\varphi(A_1))$ and $v_A^{M^*}(A_2) \leqslant \mu(\varphi(A_2))$. Therefore, $v_A^{M^*}(A) = v_A^{M^*}(A_1) + v_A^{M^*}(A_2) \leqslant \mu(\varphi(A_1)) + \mu(\varphi(A_2)) = \mu(\varphi(A))$. Thus, $A \notin S^*$.

Case (a) and (b) together imply that for any $A \notin \mathcal{S}_u$ and $A \neq \mathcal{U}$, so it must be that $A \notin \mathcal{S}^*$. Consequently,

$$\mathcal{S}^* \subset \mathcal{S}_u \backslash \{\mathcal{U}\} \tag{B.18}$$

.

Now we prove that $\mathcal{S}^* \subset \mathcal{S}_y^{-1} \cap (S_u \backslash \{\mathcal{U}\})$.

For any set A such that $A \in S_u \backslash \{\mathcal{U}\}$ and $A \notin \mathcal{S}^{-1}(y)$, we need to show that $A \notin \mathcal{S}^*$.

Since such a set $A \in \mathcal{S}_u \backslash \mathcal{U}$, define $B = \varphi(A)^c := \mathcal{Y} \backslash \varphi(A)$. $\varphi(A)$ must not equal to $\mathcal{Y}$, because otherwise, by statement (2) of Definition 8, $A = \mathcal{U} \notin \mathcal{S}_u \backslash \mathcal{U}$. Therefore, B is non-empty and $B^c = \varphi(A)$. For any $u \notin A$, since $A \in \mathcal{S}_u$, $\varphi(u)$ must not be a subset of $\varphi(A)$. It implies that $\varphi(u) \cap B \neq \emptyset$, i.e., $u \in \varphi^{-1}(B)$. Thus, $A^c \subset \varphi^{-1}(B)$. On the other hand, since $B = \varphi(A)^c$, so $\varphi^{-1}(B) \cap A = \emptyset$, i.e., $\varphi^{-1}(B) \subset A^c$. Consequently, $\varphi^{-1}(B) = A^c$.

By assumption, $A \notin \mathcal{S}_y^{-1}$ implies that $B \notin \mathcal{S}_y$, i.e., either (a) there exist $y \in B^c$ such that $\varphi^{-1}(y) \subset \varphi^{-1}(B) = A^c$ or (b) B is not self-connected, i.e., there is a partition B_1, B_2 of B, $B_1 \cup B_2 = B$, $B_1 \cap B_2 = \emptyset$ such that $\varphi^{-1}(B_1) \cap \varphi^{-1}(B_2) = \emptyset$.

For case (a), since $y \in B^c := \varphi(A)$, so $\varphi^{-1}(y) \cap A \neq \emptyset$, which contradicts with the statement in (a) that $\varphi^{-1}(y) \subset A^c$. Thus, case (a) can not happen.

For case (b), Define $A_1 := \varphi^{-1}(B_1) \cup A$, and $A_2 := \varphi^{-1}(B_2) \cup A$. Therefore, $\varphi(A_1) = B_1 \cup B^c$, and $\varphi(A_2) = B_2 \cup B^c$. Notice that $\varphi^{-1}(B_1) \cup \varphi^{-1}(B_2) \cup A = \mathcal{U}$ and $B_1 \cup B_2 \cup \varphi(A) = \mathcal{Y}$, we have $v(\varphi^{-1}(B_1)) + v(\varphi^{-1}(B_2)) + v(A) = v(\mathcal{U})$, and $\mu(B_1) + \mu(B_2) + \mu(\varphi(A)) = \mu(\mathcal{Y})$.

For any $v \geqslant 0$ such that $v(A') \leqslant \mu(\varphi(A'))$ for all $A' \neq A$, and $v(\mathcal{U}) = \mu(\mathcal{Y})$, it follows that:

$v(A) = v(A) + \big(v(A) + v(\varphi^{-1}(B_1)) + v(\varphi^{-1}(B_2))\big) - v(\mathcal{U})$
$= v(A_1) + v(A_2) - v(\mathcal{U})$
$\leqslant \mu(\varphi(A_1)) + \mu(\varphi(A_2)) - \mu(\mathcal{Y})$
$= \mu(B_1) + \mu(B^c) + \mu(B_2) + \mu(B^c) - \mu(\mathcal{Y})$
$= \mu(B^c) = \mu(\varphi(A))$. Thus, $v_A^{M^*} \leqslant \mu(\varphi(A))$, i.e., $A \notin \mathcal{S}^*$.

Therefore, for any $A \in \mathcal{S}_u \backslash \{\mathcal{U}\}$, if $A \notin \mathcal{S}^{-1}(y)$, then $A \notin \mathcal{S}^*$. It implies that $\mathcal{S}^* \subset (\mathcal{S}_u \cap \mathcal{S}^{-1}(y)) \backslash \{\mathcal{U}\}$.

• Step 2. We prove that $(\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\} \subset \mathcal{S}^*$.

We prove this statement by contradiction.

Suppose $(\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$ is not a subset of $\mathcal{S}^*$.

Suppose there exists a set A such that $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$, but $A \notin \mathcal{S}^*$. By Theorem 1, $\mathcal{S}^*$ is core-determining. Denote $\mathcal{S}^* = \{A_1, ..., A_r\}$, where $r = |\mathcal{S}^*|$.

Let us enumerate the elements of $\mathcal{U}$ as $u_1, ..., u_{d_u}$. Denote $w(A) \in \{0,1\}^{d_u}$ as the characteristic vector of the set $A \subset \mathcal{U}$, i.e., the i^{th} entry of $w(A)$ equals to 1 if and only if $u_i \in A$.

By Farkas Lemma, there exists $\pi_i \geqslant 0$ and $A_i \in \mathcal{S}^*$, $1 \leqslant i \leqslant r$, with $r = |\mathcal{S}^*|$, and $\pi_0 \in \mathbb{R}$, such that:

$$\sum_{i=1}^{r} \pi_i w(A_i) - \pi_0 w(\mathcal{U}) \geqslant w(A). \tag{B.19}$$

$$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) - \pi_0 \leqslant \mu(\varphi(A)). \tag{B.20}$$

First, there must exist at least one π_i, $i \in \{1,2,...,r\}$ such that $\pi_i > 0$; Otherwise we assume that $\pi_0 = 0$ for all $i = 1,2,...,r$, since A is non-empty, there must exist some i such that the i^{th} entry of $w(A)$ equals to one. By (B.19), it must be that $-\pi_0 \geqslant 1$. By (B.20), we have $\mu(\varphi(A)) \geqslant 1 = \mu(\mathcal{Y})$, but this can not happen for any $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1})\backslash\{\mathcal{U}\}$, which leads to a contradiction!

WLOG., we assume $\pi_i > 0$ holds for all $i = 1,2,...,r'$, with $1 \leqslant r' \leqslant r$, and $\pi_i = 0$ for all $i > r'$; for those i such that $\pi_i = 0$, we can simply ignore the component related to A_i in (B.19) and (B.20).

Based on (B.19) and (B.20), we claim that the following proposition must be true.

Proposition 5. *Suppose $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1})\backslash\{\mathcal{U}\}$, but $A \notin \mathcal{S}^*$. In (B.19) and (B.20), Suppose $\pi_i > 0$ holds for $i = 1,2,...,r'$, and $\pi_i = 0$ for $i > r'$. Then,*

(1) For any $y \in \mathcal{Y}$, $\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0 = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$, and $\sum_{i=1}^{r'} \pi_i w(A_i) - \pi_0 w(\mathcal{U}) = w(A)$.

(2) $\pi_0 \geqslant 0$.

(3) For any $i = 1,2,...,r'$, either $A_i \supset A^c$ or $A_i \cap A^c = \emptyset$.

(4) For any $i = 1,2,...,r'$, either $A_i \supset A$ or $A_i \cap A = \emptyset$.

(5) For each $i = 1,2,...,r'$, $A_i = A$.

Proof of Proposition 5.

(1) Proof of Statement (1).

By inequality (B.19), for any $u \in \mathcal{U}$, $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0 \geqslant 1(u \in A)$.

For any $y \in \varphi(A)$, we can pick a $u \in A \cap \varphi^{-1}(y)$. It is easy to see that for any A_i, $1(\varphi^{-1}(y) \cap A_i \neq \emptyset) \geqslant 1(u \in A_i)$.

Thus,

$(\sum_{1\leqslant i\leqslant r'} \pi_i 1(A_i \cap \varphi^{-1}(y)) \neq \emptyset) - \pi_0 \geqslant (\sum_{1\leqslant i\leqslant r'} \pi_i 1(u \in A_i)) - \pi_0 \geqslant 1(u \in A) = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$.

For any $y \notin \varphi(A)$, we can pick a $u \notin A$ such that $u \in \varphi^{-1}(y)$. Again, for any A_i, $1(\varphi^{-1}(y) \cap A_i \neq \emptyset) \geqslant 1(u \in A_i)$

Therefore,

$(\sum_{1\leqslant i\leqslant r'} \pi_i 1(A_i \cap \varphi^{-1}(y)) \neq \emptyset) - \pi_0 \geqslant (\sum_{1\leqslant i\leqslant r'} \pi_i 1(u \in A_i) - \pi_0 \geqslant 1(u \in A) = 0 = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$.

Consequently, for any $y \in \mathcal{Y}$, it must be that

$$(\sum_{1\leqslant i\leqslant r'} \pi_i 1(A_i \cap \varphi^{-1}(y)) \neq \emptyset) - \pi_0 \geqslant 1(\varphi^{-1}(y) \cap A \neq \emptyset) \tag{B.21}$$

The left side of the inequality (B.20) can be written as:

$\sum_{1\leqslant i\leqslant r'} \pi_i \mu(\varphi(A_i)) - \pi_0$

$= \sum_{y\in\mathcal{Y}} \mu(y)(\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0)$

Applying inequality (B.21) to the quantity above, we have:

$\sum_{y\in\mathcal{Y}} \mu(y)(\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0)$

$\geqslant \sum_{y\in\mathcal{Y}} \mu(y) 1(\varphi^{-1}(y) \cap A \neq \emptyset) = \mu(\varphi(A))$.

Therefore, (B.20) must hold as an equality. By assumption that μ is non-degenerate, $\mu(y) > 0$ for all $y \in \mathcal{Y}$. Thus, the equality of (B.20) also implies that for all $y \in \mathcal{Y}$,

$\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0 = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$. Therefore, the first part of Statement (1) is proved.

Now we prove the second part of Statement (1).

If $\sum_{i=1}^{r'} \pi_i w(A_i) - \pi_0 w(\mathcal{U}) \neq w(A)$, by (B.19), one of the following situations must be true:

Case (a): there exists a $u \notin A$ such that $\sum_{i=1}^{r'} 1(u \in A_i) - \pi_0 > 0$.

Case (b): there exists a $u \in A$ such that $\sum_{i=1}^{r'} 1(u \in A_i) - \pi_0 > 1$.

For case (a), since $A \in \mathcal{S}_u$, so $\varphi(u)$ must not be a subset of $\varphi(A)$. Thus, there exists a $y \in \varphi(u) \backslash \varphi(A)$.

Therefore, $\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) - \pi_0$
$\geqslant \sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0 > 0$.

However, by the first part of Statement (1) of this proposition,
$\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) - \pi_0$
$= 1(\varphi^{-1}(y) \cap A \neq \emptyset) = 0$, which leads to a contradiction.

For case (b), since $u \in A$, we can find a $y \in \varphi(u) \subset \varphi(A)$. Therefore, $\sum_{i=1}^{r'} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) - \pi_0$
$\geqslant \sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0$
$> 1 = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$, which leads to a contradiction to the first part of the Statement (1).

Therefore, it must be that $\sum_{i=1}^{r'} \pi_i w(A_i) - \pi_0 w(\mathcal{U}) = w(A)$.

(2) Proof of Statement (2).

Since $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$, there must exists a $u \in A^c$.

By the second part of statement (1), $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0 = 1(u \in A) = 0$ holds for any $u \in \mathcal{U}$. Since $\pi_i > 0$ for $i = 1, 2, ..., r'$, it follows that $\pi_0 = \sum_{i=1}^{r'} \pi_i 1(u \in A_i) \geqslant 0$.

(3) Proof of Statement (3).

Since $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$, there must exist a non-empty set $B \subset \mathcal{Y}$ such that:

1. B satifies the conditions stated in Definition 11.
2. $\varphi^{-1}(B) = A^c$. It follows that $B \subset \varphi(A)^c$.

For any $y \in \varphi(A)^c$, $\varphi^{-1}(y) \subset A^c = \varphi^{-1}(B)$. By Definition 11, $y \in B$. Therefore, $B = \varphi(A)^c$.

By Statement (1) of this proposition, for any $y \in B$,

$$\Big(\sum_{1 \leqslant i \leqslant r} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)\Big) - \pi_0 = 0. \tag{B.22}$$

And for all $y \notin B$,

$$\Big(\sum_{1 \leqslant i \leqslant r} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)\Big) - \pi_0 = 1. \tag{B.23}$$

For a $y \in B$ and $i_0 \in \{1, 2, ..., r'\}$, if $\varphi^{-1}(y) \cap A_{i_0} \neq \emptyset$, we claim that $\varphi^{-1}(y) \subset A_{i_0}$.

By contradiction, we assume that $\varphi^{-1}(y)$ is not a subset of A_{i_0}. Then, we can find a $u_2 \in \varphi^{-1}(y) \backslash A_{i_0}$. Because $\varphi^{-1}(y) \subset \varphi^{-1}(B) = A^c$, it follows that $u_2 \notin A$.

By equation (B.22),

$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0 = 1(\varphi^{-1}(y) \cap A) = 0$.

On the other hand,

$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0$

$= \pi_{i_0} - \pi_0 + \sum_{1 \leqslant i \leqslant r, i \neq i_0} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset)$

$\geqslant \pi_{i_0} - \pi_0 + \sum_{1 \leqslant i \leqslant r, i \neq i_0} \pi_i 1(u_2 \in A_i)$

$= \pi_{i_0} - \pi_{i_0} 1(u_2 \in A_{i_0}) + \sum_{i=1}^{r} \pi_i 1(u_2 \in A_i) - \pi_0$

$= \pi_{i_0} - 0 + 0 > 0$, which leads to a contradiction to (B.22)!

Therefore, for any $y \in B$ and $i = 1, 2, ..., r'$, either $\varphi^{-1}(y) \subset A_i$ or $\varphi^{-1}(y) \cap A_i = \emptyset$.

Now, for any A_i, $i = 1, 2, ..., r'$, we can define $K_1^i := \{y \in B | \varphi^{-1}(y) \subset A_i\}$ and $K_2^i := \{y \in B | \varphi^{-1}(y) \cap A_i = \emptyset\}$, so K_1^i and K_2^i form a partition of B. However, if K_1^i and K_2^i are both non-empty, then $\varphi^{-1}(K_1^i) \cap \varphi^{-1}(K_2^i) = \emptyset$, i.e., B is not self-connected, which is untrue by assumption that $B \in \mathcal{S}_y$. Then, either K_1^i or K_2^i must be empty set.

It implies that either

$$A_i \cap A^c = \emptyset \text{ or } A_i \supset A^c \tag{B.24}$$

, for any $i = 1, 2, ..., r'$.

(4) Proof of Statement (4).

We prove this statement by contradiction: assume that there exists a $i_1 \in \{1, 2, ..., r'\}$ such that $A_{i_1} \cap A \neq \emptyset$ and A_{i_1} does not contain A.

Since A is self-connected, it must be that $\varphi(A_{i_1} \cap A) \cap \varphi(A \backslash A_{i_1}) \neq \emptyset$. Thus, there must exists $u_1 \in A_{i_1} \cap A$ and $u_2 \in A \backslash A_{i_1}$ such that there exists $y \in \varphi(A)$, $y \in \varphi(u_1) \cap \varphi(u_2)$.

The left hand side of (B.23) can be written as:

$(\sum_{1 \leqslant i \leqslant r} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)) - \pi_0$

$= \pi_{i_1} + \sum_{1 \leqslant i \leqslant r, i \neq i_1} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset) - \pi_0$

$\geqslant \pi_{i_1} + \sum_{1 \leqslant i \leqslant r, i \neq i_1} \pi_i 1(u_2 \in A_i) - \pi_0$.

By (B.19) and the assumption that $u_2 \notin A_{i_1}$,

$\sum_{1 \leqslant i \leqslant r, i \neq i_1} \pi_i 1(u_2 \in A_i) - \pi_0$

$= \sum_{1 \leqslant i \leqslant r} \pi_i 1(u_2 \in A_i) - \pi_0 \geqslant 1$.

Therefore, $(\sum_{1 \leqslant i \leqslant r} \pi_i 1(A_i \cap \varphi^{-1}(y) \neq \emptyset)) - \pi_0 \geqslant \pi_{i_1} + 1 > 1$, which contradicts with (B.23).

Thus, such i_1 does not exist. It follows that for any A_i, $i = 1, 2, ..., r$, either $A_i \cap A = \emptyset$ or $A_i \supset A$.

(5) Proof of Statement (5).

By Statement (3) and (4) of this proposition, for $i = 1, 2, ..., r'$, $A_i \cap A \in \{\emptyset, A\}$ and $A_i \cap A^c \in \{\emptyset, A^c\}$. Since A_i is neither $\emptyset$ or $\mathcal{U}$, it must be that $A_i \in \{A, A^c\}$.

To prove the Statement (5) of this proposition, we only need to show that $A_i \neq A^c$ for all $i = 1, 2, ..., r'$.

By contradiction, suppose there exists a $i_0 \in \{1, 2, ..., r'\}$ such that $A_{i_0} = A^c$. Since G is connected and $\varphi(A_{i_0}) \cap \varphi(A) \neq \emptyset$, there exists $y \in \mathcal{Y}$ such that $y \in \varphi(A_{i_0}) \cap \varphi(A)$, and there exists a $u \in \mathcal{U}$ such that $u \in \varphi^{-1}(y) \cap A$.

Therefore, $\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0$
$= \pi_{i_0} + \sum_{1 \leqslant i \leqslant r', i \neq i_0} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0$
$\geqslant \pi_{i_0} + \sum_{1 \leqslant i \leqslant r', i \neq i_0} \pi_i 1(u \in A_i) - \pi_0$
$= \pi_{i_0} - \pi_{i_0} 1(u \in A_{i_0}) + \sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0$
$= \pi_{i_0} - 0 + \sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0$,

By the second part of Statement (1) of this proposition,
$\sum_{i=1}^{r'} \pi_i 1(u \in A_i) - \pi_0 = 1(u \in A) = 1$.

Thus, $\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) - \pi_0$
$\geqslant \pi_{i_0} - 0 + 1 > 1(\varphi^{-1}(y) \cap A \neq \emptyset)$, which is contradictory to the first part of Statement (1) of this proposition. Consequently, there exist no A_{i_0} such that $A_{i_0} = A^c$, i.e., $A_i = A$ for all $i = 1, 2, ..., r'$. □

Statement (5) of Proposition 5 implies that $A = A_1 \in \mathcal{S}^*$, which obviously leads to a contradiction to the assumption that $A \notin \mathcal{S}^*$.

Thus, there does not exist any A such that $A \in (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$ and $A \notin \mathcal{S}^*$, i.e., $\mathcal{S}^* \supset (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$. Together with (B.18), we have:
$\mathcal{S}^* = (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$.

□

Proof of Theorem 2.

By Lemma 5, when both G and μ are non-degenerate, the exact core-determining class $\mathcal{S}^* = (\mathcal{S}_u \cap \mathcal{S}_y^{-1}) \backslash \{\mathcal{U}\}$.

Therefore, for any non-empty set $A \subsetneq \mathcal{U}$, $A \in \mathcal{S}^*$ if and only if A satisfies the conditions stated in Definition 8 and $A \in \mathcal{S}_y^{-1}$.

Statement (1) and (2) in Theorem 2 cover the conditions stated in Definition 8. Therefore, we only need to prove that, if $A \in \mathcal{S}_u$, $A \neq \mathcal{U}$, $A \in \mathcal{S}_y^{-1}$ if and only if Statement (3) in Theorem 2 holds.

Denote $B := \varphi(A)^c$. Given that $A \in \mathcal{S}_u, A \neq \mathcal{U}$, for any $u \in A^c$, it must be that $\varphi(u)$ is not a subset of $\varphi(A)$. That said, $u \in \varphi^{-1}(B)$. Thus, $\varphi^{-1}(B) \supset A^c$. On the other hand, since $B = \varphi(A)^c$, it must be that $\varphi^{-1}(B) \subset A^c$. Therefore, $\varphi^{-1}(B) = A^c$.

If $A \in \mathcal{S}_y^{-1}$, it implies that $B = \varphi(A)^c \in \mathcal{S}_y$. By Definition 11, statement (3) of Theorem 2 holds.

On the other hand, if statement (3) of Theorem 2 holds, then $B = \varphi(A)^c$ is connected in G^{-1}. For any $y \in \mathcal{Y} \backslash B = \varphi(A)$, $\varphi^{-1}(y) \cap A \neq \emptyset$. Therefore, $\varphi^{-1}(y)$ is not a subset of $\varphi^{-1}(B) = A^c$. Thus, by Definition 11, $B \in \mathcal{S}_y$. Consequently, $A = \varphi^{-1}(B)^c \in \mathcal{S}_y^{-1}$. Therefore, $A \in \mathcal{S}^*$. □

B.3. Proof of Theorem 3

Proof of Theorem 3.

We divide our proof into two steps.

• Step 1: We show that $\mathcal{S}^*$ is a core-determining class, i.e., $Q(\mathcal{S}^*) := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A)),$ for all $A \in \mathcal{S}^*, v(\mathcal{U}) = \mu(\mathcal{Y}) = 1\} = Q$, where $Q := \{v \geqslant 0 | v(A) \leqslant \mu(\varphi(A)),$ for all non-empty $A \subset \mathcal{U}, v(\mathcal{U}) = \mu(\mathcal{Y}) = 1\}$.

First of all, since v and μ are probability measures, $\sum_{j=1}^k v(\mathcal{U}_j) = 1 = \sum_{j=1}^k \mu(\varphi(\mathcal{U}_j))$. Since $v(\mathcal{U}_j) \leqslant \mu(\varphi(\mathcal{U}_j)) = \mu(\mathcal{Y}_j)$ holds for all $j = 1, 2, ..., k$, it must be that:

$v(\mathcal{U}_j) = \mu(\varphi(\mathcal{U}_j)) = \mu(\mathcal{Y}_j) > 0$, $j = 1, 2, ..., k$.

Therefore, for any $v \in Q(\mathcal{S}^*)$, we have $v(\mathcal{U}_j) = \mu(\mathcal{Y}_j) > 0$.

By Theorem 1, $\mathcal{S}_j^*$ is an exact core-determining class of G_j, for any $j = 1, 2, ..., k$.

Therefore, for any non-empty $A \subset \mathcal{U}$ and any $v \in Q(\mathcal{S}^*)$, it must be that $v(A \cap \mathcal{U}_j) \leqslant \mu(\varphi(A \cap \mathcal{U}_j))$, for any $j = 1, 2, ..., k$. These inequalities also hold when $A \cap \mathcal{U}_j = \emptyset$. Because $G_1, ..., G_k$ are distinct connected branches, $v(A) = \sum_{j=1}^k v(A \cap \mathcal{U}_j)$, and $\mu(\varphi(A)) = \sum_{j=1}^k \mu(\varphi(A \cap \mathcal{U}_j))$.

Therefore, $v(A) = \sum_{j=1}^k v(A \cap \mathcal{U}_j) \leqslant \sum_{j=1}^k \mu(\varphi(A \cap \mathcal{U}_j)) = \mu(\varphi(A))$.

Since A can be selected as any arbitrary non-empty subset of $\mathcal{U}$, we have $Q(\mathcal{S}^*) = Q$.

• Step 2: we show that $\mathcal{S}^*$ is an exact core-determining class, i.e., among all core-determining classes, $\mathcal{S}^*$ has the smallest possible cardinality.

For a core-determining class $\widetilde{\mathcal{S}}$, we can denote $\widetilde{\mathcal{S}}$ as $\{\widetilde{U}_1, ..., \widetilde{U}_q, A_1, ..., A_r\}$, such that:

(a) For each $l = 1, 2, ..., q$, $\widetilde{U}_l$ is a union of some sets in $\{\mathcal{U}_1, ..., \mathcal{U}_k\}$.

(b) For each $i = 1, 2, ..., r$, A_i is not a union of some sets in $\{\mathcal{U}_1, ..., \mathcal{U}_k\}$; or equivalently, for each $i = 1, 2, ..., r$, there exists a $j \in \{1, 2, ..., k\}$ such that $A_i \cap \mathcal{U}_j \neq \emptyset$ and $A_i \cap \mathcal{U}_j \neq \mathcal{U}_j$.

For any subset $\widetilde{A}$ of $\mathcal{U}$, we can define the characteristic vector $w(\widetilde{A})$ as a vector in $\{0, 1\}^{d_u}$ such that the i^{th} entry of $w(\widetilde{A})$ equals to 1 if and only if $u_i \in \widetilde{A}$.

By the proof of Step 1, we know that for any $v \in Q$, $v(\mathcal{U}_j) = \mu(\varphi(\mathcal{U}_j))$, $j = 1, 2, ..., k$. Therefore, $v(\widetilde{U}_l) = \mu(\varphi(\widetilde{U}_l))$ holds for $l = 1, 2, ..., q$.

We claim that the following proposition must be true.

Proposition 6. *(1) For any $\widetilde{\xi}_1^*, ..., \widetilde{\xi}_q^* \in \mathbb{R}$, there exist $\xi_1, ..., \xi_k \in \mathbb{R}$ such that:*
$\sum_{l=1}^q \widetilde{\xi}_l^* w(\widetilde{U}_l) = \sum_{j=1}^k \xi_j w(\mathcal{U}_j)$,
and
$\sum_{l=1}^q \widetilde{\xi}_l^* \mu(\varphi(\widetilde{U}_l)) = \sum_{j=1}^k \xi_j \mu(\varphi(\mathcal{U}_j))$.
(2) q must be greater than or equal to k.

Proof of Proposition 6.

(1) Proof of Statement (1).

Since for any $l = 1, 2, ..., q$, $\widetilde{U}_l$ is a union of some $\mathcal{U}_j$s. Therefore, $w(\widetilde{U}_l) = \sum_{j=1}^{k} \delta_{lj} w(\mathcal{U}_j)$, and $\mu(\varphi(\widetilde{U}_l)) = \sum_{j=1}^{k} \delta_{lj}\mu(\varphi(\mathcal{U}_j))$, where $\delta_{lj} = 1(\mathcal{U}_j \subset \widetilde{U}_l)$, $j = 1, 2, ..., k$.

Thus, for any $\widetilde{\xi}_1^*, ..., \widetilde{\xi}_q^* \in \mathbb{R}$, define $\xi_j := \sum_{l=1}^{q} \delta_{lj}\widetilde{\xi}_l^*$. We have:
$\sum_{l=1}^{q} \widetilde{\xi}_l^* w(\widetilde{U}_l) = \sum_{l=1}^{q}\sum_{j=1}^{k} \delta_{lj}\widetilde{\xi}_l^* w(\mathcal{U}_j) = \sum_{j=1}^{k} \xi_j w(\mathcal{U}_j)$,
and
$\sum_{l=1}^{q} \widetilde{\xi}_l^* \mu(\varphi(\widetilde{U}_l)) = \sum_{l=1}^{q}\sum_{j=1}^{k} \delta_{lj}\widetilde{\xi}_l^* \mu(\varphi(\mathcal{U}_j)) = \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j))$.
(2) Proof of Statement (2).

For each $j_1 = 1, 2, ..., k$, we know that $v(\mathcal{U}_{j_1}) = \mu(\varphi(\mathcal{U}_{j_1}))$ for all $v \in Q$. By assumption, $\widetilde{\mathcal{S}}$ is exact core-determining, so $Q(\widetilde{\mathcal{S}}) = Q$. Therefore, $-v(\mathcal{U}_{j_1}) \leqslant -\mu(\varphi(\mathcal{U}_{j_1}))$ is true under the constraints that $v(A) \leqslant \mu(\varphi(A))$ for all $A \in \widetilde{\mathcal{S}}$ and $v(\mathcal{U}) = \mu(\varphi(\mathcal{U}))$.

By Farkas Lemma, there must exist $\pi_1, ..., \pi_r \geqslant 0$ and $\widetilde{\xi}_1, ..., \widetilde{\xi}_q \in \mathbb{R}$ such that:

$$\sum_{i=1}^{r} \pi_i w(A_i) + \sum_{l=1}^{q} \widetilde{\xi}_l w(\widetilde{U}_l) \geqslant -w(\mathcal{U}_{j_1}), \tag{B.25}$$

and

$$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{l=1}^{q} \widetilde{\xi}_l \mu(\varphi(\widetilde{U}_l)) \leqslant -\mu(\varphi(\mathcal{U}_{j_1})). \tag{B.26}$$

By Statement (1) of this proposition, there exists $\xi_1, ..., \xi_k \in \mathbb{R}$ such that

$$\sum_{i=1}^{r} \pi_i w(A_i) + \sum_{j=1}^{k} \xi_j w(\mathcal{U}_j) \geqslant -w(\mathcal{U}_{j_1}), \tag{B.27}$$

and

$$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j)) \leqslant -\mu(\varphi(\mathcal{U}_{j_1})). \tag{B.28}$$

(B.27) implies that for any $u \in \mathcal{U}$,

$$\sum_{i=1}^{r} \pi_i 1(u \in A_i) + \sum_{j=1}^{k} \xi_j 1(u \in \mathcal{U}_j) \geqslant -1(u \in \mathcal{U}_{j_1}). \tag{B.29}$$

It is easy to see that for any $y \in \mathcal{Y}$, $\widetilde{A} \subset \mathcal{U}$ and $u \in \varphi^{-1}(y)$, $1(\varphi^{-1}(y) \cap \widetilde{A} \neq \emptyset) \geqslant 1(u \in \widetilde{A})$.

For $y \in \mathcal{Y}_{j_1}$, we can pick a $u \in \varphi^{-1}(y) \subset \mathcal{U}_{j_1}$, so by utilizing (B.29), we have:

$$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) \tag{B.30}$$

$$\geqslant \sum_{i=1}^{r} \pi_i 1(u \in A_i) + \sum_{j=1}^{k} \xi_j 1(u \in \mathcal{U}_j) \geqslant -1(u \in \mathcal{U}_{j_1}) = -1 = -1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1} \neq \emptyset).$$

For any $y \notin \mathcal{Y}_{j_1}$, there exists a $j_2 \neq j_1$ such that $y \in \mathcal{Y}_{j_2}$. Therefore, $\varphi^{-1}(y) \subset \mathcal{U}_{j_2}$. We can pick a u such that $u \in \varphi^{-1}(y)$ and $u \notin \mathcal{U}_{j_1}$.

Then, by utilizing (B.29), we have:

$$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) \tag{B.31}$$

$$\sum_{i=1}^{r} \pi_i 1(u \in A_i) + \sum_{j=1}^{k} \xi_j 1(u \in \mathcal{U}_j) \geqslant -1(u \in \mathcal{U}_{j_1}) = 0 = -1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1} \neq \emptyset).$$

Therefore,

$$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) \geqslant -1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1} \neq \emptyset) \tag{B.32}$$

holds for all $y \in \mathcal{Y}$.

Now, the left hand side of (B.28) can be written as:
$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j))$
$= \sum_{y \in \mathcal{Y}} \mu(y) \left(\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) \right)$.

By applying (B.32) on all $y \in \mathcal{Y}$, we have:
$\sum_{y \in \mathcal{Y}} \mu(y) \left(\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) \right)$
$\geqslant \sum_{y \in \mathcal{Y}} -\mu(y) 1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1} \neq \emptyset)$
$= -\mu(\varphi(\mathcal{U}_j))$.

It implies that (B.25), (B.26), (B.27), (B.28) and (B.29) must hold as an equality. Since $\mu(y) > 0$ for all $y \in \mathcal{Y}$, it must be that:
$\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset) = -1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1} \neq \emptyset)$ holds for all $y \in \mathcal{Y}$.

We claim that π_i must be 0 for all $i = 1, 2, ..., r$.

By contradiction, if there exists $i_2 \in \{1, 2, ..., r\}$ such that $\pi_{i_2} > 0$, then there must exist a $j_2 \in \{1, 2, ..., k\}$ such that $A_{i_2} \cap \mathcal{U}_{j_2} \neq \emptyset$, $A_{i_2} \cap \mathcal{U}_{j_2} \neq \mathcal{U}_{j_2}$.

Since G_{j_2} is connected, there must exist a $y \in \mathcal{Y}_{j_2}$ such that $y \in \varphi(A_{i_2} \cap \mathcal{U}_{j_2})$ and $y \in \varphi(\mathcal{U}_{j_2} \backslash A_{i_2})$. There must exist a $u \in \varphi^{-1}(y) \cap (\mathcal{U}_{j_2} \backslash A_{i_2})$. It is easy to see that $u \notin A_{i_2}$.

Therefore, $\sum_{i=1}^{r} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset)$
$= \pi_{i_2} + \sum_{1 \leqslant i \leqslant r, i \neq i_2} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \sum_{j=1}^{k} \xi_j 1(\varphi^{-1}(y) \cap \mathcal{U}_j \neq \emptyset)$
$\geqslant \pi_{i_2} + \sum_{1 \leqslant i \leqslant r, i \neq i_2} \pi_i 1(u \in A_i) + \sum_{j=1}^{k} \xi_j 1(u \in \mathcal{U}_j)$
$= \pi_{i_2} - \pi_{i_2} 1(u \in A_{i_2}) + \sum_{1 \leqslant i \leqslant r} \pi_i 1(u \in A_i) + \sum_{j=1}^{k} \xi_j 1(u \in \mathcal{U}_j)$
$=_{(\text{ by (B.29) being an equation})} \pi_{i_2} - 0 - 1(u \in \mathcal{U}_{j_1})$
$= \pi_{i_2} - 1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1}) > -1(\varphi^{-1}(y) \cap \mathcal{U}_{j_1})$, which leads to a contradiction to (B.32). Therefore, $\pi_i = 0$ for all $i = 1, 2, ..., r$.

Now, equation (B.25) becomes: for each $j_1 = 1, 2, ..., k$, there exist $\widetilde{\xi}_1, ..., \widetilde{\xi}_q \in \mathbb{R}$ such that

$\sum_{l=1}^{q} \widetilde{\xi}_l w(\widetilde{U}_l) = -w(\mathcal{U}_{j_1})$.

Define $L := \{z | z = \sum_{l=1}^{q} \widetilde{\xi}_l w(\widetilde{U}_l), \widetilde{\xi}_1, ..., \widetilde{\xi}_q \in \mathbb{R}\}$, then $L \subset \mathbb{R}^{d_u}$ is a linear space with dimensionality of at most q.

From the analysis above, we know that $-w(\mathcal{U}_{j_1}) \in L$, $j_1 = 1, 2, ..., k$. We also know that $-w(\mathcal{U}_{j_1})$ are linearly independent for $j_1 = 1, 2...., k$. It implies that L contains k linearly independent vectors in $\mathbb{R}^{d_u}$. Therefore, $q \geqslant dim(L) \geqslant k$.

□

For any $A \in \mathcal{S}^* \backslash \{\mathcal{U}_1, ..., \mathcal{U}_k\}$, since $\mathcal{S}^* = \cup_{j=1}^{k} \mathcal{S}_j^*$, there must exist a $j_0 \in \{1, 2, ..., k\}$ such that $A \subset \mathcal{U}_{j_0}$.

Since $\widetilde{\mathcal{S}}$ is core-determining, by Farkas Lemma, there exist $\pi_1, ..., \pi_r \geqslant 0$ and $\widetilde{\xi}_1, ..., \widetilde{\xi}_q \in \mathbb{R}$ such that:

$\sum_{i=1}^{r} \pi_i w(A_i) + \sum_{l=1}^{q} \widetilde{\xi}_l w(\widetilde{U}_l) \geqslant w(A)$,

and

$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{l=1}^{q} \widetilde{\xi}_l \mu(\varphi(\widetilde{U}_l)) \leqslant \varphi(A)$.

By Proposition (6), there exist $\xi_1, ..., \xi_k \in \mathbb{R}$ such that:

$\sum_{l=1}^{q} \widetilde{\xi}_l w(\widetilde{U}_l) = \sum_{j=1}^{k} \xi_j w(\mathcal{U}_j)$,

and

$\sum_{l=1}^{q} \widetilde{\xi}_l \mu(\varphi(\widetilde{U}_l)) = \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j))$.

Therefore, we have:

$$\sum_{i=1}^{r} \pi_i w(A_i) + \sum_{j=1}^{k} \xi_j w(\mathcal{U}_j) \geqslant w(A), \tag{B.33}$$

$$\sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j)) \leqslant \varphi(A). \tag{B.34}$$

There must be at least one π_i that is strictly positive. Otherwise let us assume that all $\pi_i = 0$, $i = 1, 2, ..., r$. Since A is a non-empty subset of $\mathcal{U}_{j_0}$, there must be some entries in $w(A)$ that equal to one. Such entries correspond to certain elements in $\mathcal{U}_{j_0}$. Then, by (B.33), it must be that $\widetilde{xi}_{j_0} \geqslant 1$ and $\widetilde{xi}_j \geqslant 0$ for all $j \neq j_0$. Then, by (refeq:coromain2), we have: $\mu(\varphi(\mathcal{U}_{j_0})) \leqslant \sum_{i=1}^{r} \pi_i \mu(\varphi(A_i)) + \sum_{j=1}^{k} \xi_j \mu(\varphi(\mathcal{U}_j)) \leqslant \varphi(A)$. Since $A \in \mathcal{S}_{j_0}^*$, we know that $A \neq \mathcal{U}_{j_0}$ and $\varphi(A) \neq \varphi(\mathcal{U}_{j_0}) = \mathcal{Y}_{j_0}$. Then, $\varphi(A) < \varphi(\mathcal{U}_j)$, which leads to a contradiction! Thus, there must be at least one $\pi_i > 0$, $i \in \{1, 2, ..., r\}$.

WLOG., we can assume that $\pi_i > 0$ holds for all $i = 1, 2, ..., r'$ with $1 \leqslant r' \leqslant r$, and $\pi_i = 0$ for all $i > r'$; if $\pi_i = 0$, we can simple drop the corresponding A_i in the two inequalities stated above.

Proposition 7. *(1) For each $j \in \{1, 2, ..., k\}$, we have: $\sum_{i=1}^{r'} \pi_i w(A_i \cap \mathcal{U}_j) + \xi_j w(\mathcal{U}_j) = w(A \cap \mathcal{U}_j)$ and for any $y \in \mathcal{Y}_j$, $\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap A_i \neq \emptyset) + \xi_j = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$.*

(2) There must exists $i \in \{1, 2, ..., r'\}$ such that $A_i \cap \mathcal{U}_{j_0} = A$, and $A_i \cap \mathcal{U}_j = \emptyset$ or $\mathcal{U}_j$, for all $j \neq j_0, j \in \{1, 2, ..., k\}$; or equivalently, each A_i is the union of A and some sets $\in (\{\mathcal{U}_1, \mathcal{U}_2, ..., \mathcal{U}_k\} \backslash \{\mathcal{U}_{j_0}\})$.

Proof of Proposition 7.

(1) Proof of Statement (1).

For every $u \in \mathcal{U}$, there exist a $j_1 \in \{1, 2, ..., k\}$ such that $u \in \mathcal{U}_{j_1}$.

By inequality (B.33), we have $\sum_{i=1}^{r'} \pi_i 1(u \in A_i) + \xi_{j_1} \geqslant 1(u \in A)$.

Or equivalently,

$$\sum_{i=1}^{r'} \pi_i 1(u \in A_i \cap \mathcal{U}_{j_1}) + \xi_{j_1} \geqslant 1(u \in A \cap \mathcal{U}_{j_1}). \tag{B.35}$$

For any $j = 1, 2, ..., k$ and $y \in \mathcal{Y}_j$, it must be that $\varphi^{-1}(y) \subset \mathcal{U}_j$.

If $y \in \varphi(A)$, we can pick a u such that $u \in \varphi^{-1}(y) \cap A$ and $u \in \mathcal{U}_j$.

It is easy to see that $1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) \geqslant 1(u \in (A_i \cap \mathcal{U}_j))$ for any $i = 1, 2, ..., r'$.

Therefore,

$$\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) + \xi_j \geqslant \sum_{i=1}^{r'} \pi_i 1(u \in (A_i \cap \mathcal{U}_j)) + \xi_j \geqslant 1(u \in A \cap \mathcal{U}_j) = 1 = 1(\varphi^{-1}(y) \cap (A \cap \mathcal{U}_j) \neq \emptyset). \tag{B.36}$$

If $y \notin \varphi(A)$, then pick any $u \in \varphi^{-1}(y)$, we have:

$\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) - \xi_j \geqslant \sum_{i=1}^{r'} \pi_i 1(u \in (A_i \cap \mathcal{U}_j)) - \xi_j \geqslant 1(u \in A \cap \mathcal{U}_j) = 0 = 1(\varphi^{-1}(y) \cap (A \cap \mathcal{U}_j) \neq \emptyset)$.

Therefore, inequality (B.36) holds for any $j = 1, 2, ..., k$ and $y \in \mathcal{Y}_j$.

The left hand side of (B.34) can be written as:

$\sum_{j=1}^{k} \left(\sum_{y \in \mathcal{U}_j} \mu(y) \left(\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) + \xi_j \right) \right)$

By plugging (B.36) in the above expression, we have:

$\sum_{j=1}^{k} \left(\sum_{y \in \mathcal{U}_j} \mu(y) \left(\sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) + \xi_j \right) \right)$

$\geqslant \sum_{j=1}^{k} \left(\sum_{y \in \mathcal{U}_j} \mu(y) 1(\varphi^{-1}(y) \cap (A \cap \mathcal{U}_j) \neq \emptyset) \right)$

$= \mu(\varphi(A))$.

It implies that (B.34) must hold as an equality. Consequently,

$$\text{For every } y \in \mathcal{Y}_{j_0}, \sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_{j_0}) \neq \emptyset) + \xi_{j_0} = 1(\varphi^{-1}(y) \cap A \neq \emptyset), \tag{B.37}$$

and

$$\sum_{i=1}^{r'} \pi_i w(A_i \cap \mathcal{U}_{j_0}) + \xi_{j_0} w(\mathcal{U}_{j_0}) = w(A). \tag{B.38}$$

Similarly, for every $j \neq j_0$, it must be that

$$\text{For every } y \in \mathcal{Y}_j, \sum_{i=1}^{r'} \pi_i 1(\varphi^{-1}(y) \cap (A_i \cap \mathcal{U}_j) \neq \emptyset) + \xi_j = 0, \tag{B.39}$$

and

$$\sum_{i=1}^{r'} \pi_i w(A_i \cap \mathcal{U}_j) + \xi_j w(\mathcal{U}_j) = 0, \tag{B.40}$$

Therefore, Statement (1) is proved.

(2) Proof of Statement (2).

For each $j = 1, 2, ..., k$, there may exist some i such that $A_i \cap \mathcal{U}_j \in \{\emptyset, \mathcal{U}_j\}$.

If $A_i \cap \mathcal{U}_j = \emptyset$, then A_i can be automatically eliminated from the equations (B.37) - (B.40) for the corresponding j.

If $A_i \cap \mathcal{U}_j = \mathcal{U}_j$, then we can let ξ_j to obsorbe π_i and therefore, A_i is also eliminated from equations (B.37) - (B.40). We can assume that for each $j = 1, 2, ..., k$, there exists an integer $r_j \geqslant 0, r_j \leqslant r'$ such that $\widetilde{\mathcal{S}}_j := \{A_{i_{j_1}}, ..., A_{i_{j_{r_j}}}\}$ is the collection of those sets in $\{A_1, ..., A_{r'}\}$ whose intersection with $\mathcal{U}_j$ is neither $\emptyset$ nor $\mathcal{U}_j$.

Part (A): Now, let us prove that $r_j = 0$ for all $j \neq j_0$.

By contradiction, we assume that $r_j > 0$ for some $j \in \{1, 2, ..., k\}, j \neq j_0$.

Equations (B.39) and (B.40) can be rewritten as: for each $y \in \mathcal{Y}_j$,
$\sum_{m=1}^{r_j} \pi_{i_{j_m}} 1(\varphi^{-1}(y) \cap (A_{i_{j_m}} \cap \mathcal{U}_j) \neq \emptyset) + \widetilde{\xi}_j = 0$,
and

$$\sum_{m=1}^{r_j} \pi_{i_{j_m}} w(A_{i_{j_m}} \cap \mathcal{U}_j) + \widetilde{\xi}_j w(\mathcal{U}_j) = 0, \tag{B.41}$$

where $\widetilde{\xi}_j = \xi_j + \sum_{i=1}^{r'} \pi_i 1(A_i \cap \mathcal{U}_j = \mathcal{U}_j)$.

Denote $C_{i_{j_m}} = A_{i_{j_m}} \cap \mathcal{U}_j$ for $m = 1, 2, ..., r_j$. Then, each $C_{i_{j_m}}$ is a subset of $\mathcal{U}_j$, and $C_{i_{j_m}} \notin \{\emptyset, \mathcal{U}_j\}$.

Since G_j is connected, there must exist a $y \in \mathcal{Y}_j$ such that $y \in \varphi(C_{i_{j_1}}) \cap \varphi(\mathcal{U}_j \backslash C_{i_{j_1}})$. Therefore, there exists $u \in \mathcal{U}_j \backslash C_{i_1}^j$ such that $u \in \varphi^{-1}(y)$.

Therefore, $\sum_{m=1}^{r_j} \pi_{i_{j_m}} 1(\varphi^{-1}(y) \cap (A_{i_{j_m}} \cap \mathcal{U}_j) \neq \emptyset) + \widetilde{\xi}_j$
$= \pi_{i_{j_1}} + \sum_{m=2}^{r_j} \pi_{i_{j_m}} 1(\varphi^{-1}(y) \cap (A_{i_{j_m}} \cap \mathcal{U}_j) \neq \emptyset) + \widetilde{\xi}_j$
$\geqslant \pi_{i_{j_1}} + \sum_{m=2}^{r_j} \pi_{i_{j_m}} 1(u \in (A_{i_{j_m}} \cap \mathcal{U}_j)) + \widetilde{\xi}_j$
$= \pi_{i_{j_1}} - \pi_{i_{j_1}} 1(u \in (A_{i_{j_1}} \cap \mathcal{U}_j)) + \sum_{m=1}^{r_j} \pi_{i_{j_m}} 1(u \in (A_{i_{j_m}} \cap \mathcal{U}_j)) + \widetilde{\xi}_j$
$= \pi_{i_{j_1}} + 0 > 0$, which leads to a contradiction to (B.41).

Thus, there is no $j \neq j_0$, such that $r_j > 0$. That said, $r_j = 0$ for all $j \neq j_0$; or equivalently, $A_i \cap \mathcal{U}_j \in \{\emptyset, \mathcal{U}_j\}$ for all $i = 1, 2, ..., r$ and $j \neq j_0$.

Part (B): For $j = j_0$, we would like to prove that for all $i = 1, 2, ..., r$, $A_i \cap \mathcal{U}_{j_0} \in \{A, \emptyset, \mathcal{U}_{j_0}\}$.

Equations (B.39) and (B.40) can be rewritten as: for each $y \in \mathcal{Y}_{j_0}$,
$\sum_{m=1}^{r_{j_0}} \pi_{i_{j_m}} 1(\varphi^{-1}(y) \cap (A_{i_{j_m}} \cap \mathcal{U}_{j_0}) \neq \emptyset) + \widetilde{\xi}_{j_0} = 1(\varphi^{-1}(y) \cap A \neq \emptyset)$,
and

$$\sum_{m=1}^{r_{j_0}} \pi_{i_{j_m}} w(A_{i_{j_m}} \cap \mathcal{U}_{j_0}) + \widetilde{\xi}_{j_0} w(\mathcal{U}_{j_0}) = w(A), \tag{B.42}$$

where $\widetilde{\xi}_{j_0} = \xi_{j_0} + \sum_{i=1}^{r} \pi_i 1(A_i \cap \mathcal{U}_{j_0} = \mathcal{U}_{j_0})$.

Since $w(A) \neq 0$ and $w(A)$ is not proportional to $w(\mathcal{U}_{j_0})$, by (B.42), it must be that $r_{j_0} > 0$.

Therefore, we can apply the statement (5) of Proposition 5 to G_{j_0}, i.e., $A_{i_{j_m}} \cap \mathcal{U}_{j_0} = A$ for all $m = 1, 2, ..., r_{j_0}$. It follows that for all $i = 1, 2, ..., r'$, $A \cap \mathcal{U}_{j_0} \in \{\emptyset, A, \mathcal{U}_{j_0}\}$, and there exists at least one $i \in \{1, 2, ..., r'\}$ such that $A_i \cap \mathcal{U}_{j_0} = A$. □

For each $A \in \mathcal{S}^* \backslash \{\mathcal{U}_1, ..., \mathcal{U}_k\}$, we know that there must exist a j_0 such that $A \subset \mathcal{U}_{j_0}$.

By Proposition 7, there must exist a $i(A) \in \{1, 2, ..., r\}$ such that $A_{i(A)} \in \widetilde{\mathcal{S}}$, $A_{i(A)} \cap \mathcal{U}_{j_0} = A$ and $A_{i(A)} \cap \mathcal{U}_j \in \{\emptyset, \mathcal{U}_j\}$ for all $j \neq j_0$.

By construction, $A_{i(A_1^*)} \neq A_{i(A_2^*)}$ for any $A_1^*, A_2^* \in \mathcal{S}^* \backslash \{\mathcal{U}_1, ..., \mathcal{U}_k\}$, $A_1^* \neq A_2^*$; or equivalently, $A_{i(A)}$ are all different from each other for all $A \in \mathcal{S}^* \backslash \{\mathcal{U}_1, ..., \mathcal{U}_k\}$.

Let $\widetilde{\mathcal{S}}' := \{A_{i(A)} | A \in \mathcal{S}^*, A \notin \{\mathcal{U}_1, ..., \mathcal{U}_k\}\}$. Then, $\widetilde{\mathcal{S}}' \subset \{A_1, ..., A_r\}$, and therefore $r \geqslant \left|\widetilde{\mathcal{S}}'\right| = |\{A | A \in \mathcal{S}^*, A \notin \{\mathcal{U}_1, ..., \mathcal{U}_k\}| = |\mathcal{S}^*| - k$.

Since $|\widetilde{\mathcal{S}}| = r + q \geqslant (|\mathcal{S}^*| - k) + k = |\mathcal{S}^*|$, it implies that any core-determining class $\widetilde{\mathcal{S}}$ has at least as many elements as $\mathcal{S}^*$.

Therefore, $\mathcal{S}^*$ is an exact core-determining class. □

References

[1] Andrews, D. W., Soares, G. (2010) Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica* 78(1): 119-157.

[2] Andrews, D. W., Shi, X. (2014) Inference based on conditional moment inequalities. *Econometrica* 81(2), 609-666.

[3] Andrews, D. W., Shi, X. (2016) Inference based on many conditional moment inequalities. *Journal of Econometrics* 196(2), 275-287 .

[4] Artstein, Z. (1983) Distributions of random sets and random selections. *Israel Journal of Mathematics* 46(4): 313-324.

[5] Bajari, P., Benkard, C. L., Levin, J. (2007) Estimating dynamic models of imperfect competition. *Econometrica* 75(5): 1331-1370.

[6] Bajari, P., Hong, H., Ryan, S. P. (2010) Identification and estimation of a discrete game of complete information. *Econometrica* 78(5): 1529-1568.

[7] Beresteanu, A., Molchanov, I., Molinari, F. (2011) Sharp identification regions in models with convex moment predictions. *Econometrica* 79(6): 1785-1821.

[8] Bresnahan, T.F., and R.J. Smith (1990) Coherency and EStimation in Simulataneous Models with censored or qualitative dependent variables. *Journal of Econometrics* 64: 355-373.

[9] Bresnahan, T.F., and R.J. Smith (1991) Empirical Models of Discrete Games. *Journal of Econometrics* 57: 531-553.

[10] Caron, R. J., McDonald, J. F., Ponic, C. M. (1989) A degenerate extreme point strategy for the classification of linear constraints as redundant or necessary. *Journal of Optimization Theory and Applications* 62(2): 225-237.

[11] Chernozhukov, V., D. Chetverikov, K. Kato (2014) Testing many moment inequalities. *Department of Economics, MIT, unpublished manuscript.*

[12] Chernozhukov, V., Hong, H., Tamer, E. (2007) Estimation and confidence regions for parameter sets in econometric models. *Econometrica* 75(5): 1243-1284.

[13] Chesher, A. (2010) Instrumental Variable Models for Discrete Outcomes, *Econometrica*, Volume 78, Issue 2, Pages 575601.

[14] Chesher, A., Rosen, A. (2012) Simultaneous equations models for discrete outcomes: coherence, completeness, and identification. http://eprints.ncrm.ac.uk/2912/1/cwp211212.pdf.

[15] Chesher, A., Rosen, A., Smolinski, K. (2013) An Instrumental Variable Model of Multiple Discrete Choice, Quantitative Economics, Volume 4, Issue 2, Pages 157196.

[16] Galichon, A., Henry, M. (2006) Inference in incomplete models. *Available at SSRN 886907.*

[17] Galichon, A., Henry, M. (2011) Set identification in models with multiple equilibria. *The Review of Economic Studies* 78(4): 1264-1298.

[18] Dummy Endogenous Variables in a simultaneous equation system. *Econometrica*, 46, 931-959.

[19] Manski, C. F., Tamer, E. (2002) Inference on regressions with interval data on a regressor or outcome. *Econometrica* 70(2): 519-546.

[20] Paulraj, S., Chellappan, C., Natesan, T. R. (2006) A heuristic approach for identification of redundant constraints in linear programming models. *International Journal of Computer Mathematics* 83(8-9): 675-683.

[21] Romano, J. P., Shaikh, A. M. (2010) Inference for the identified set in partially identified econometric models. *Econometrica* 78(1): 169-211.

[22] Telgen, J. (1983) Identifying redundant constraints and implicit equalities in systems of linear constraints. *Management Science* 29(10): 1209-1222.

[23] TE Tamer, E., (2003) Incomplete Simulatenous Discrete Response Models with Multiple Equilibria. *Review of Economic Studies* 70(1): 147-167.

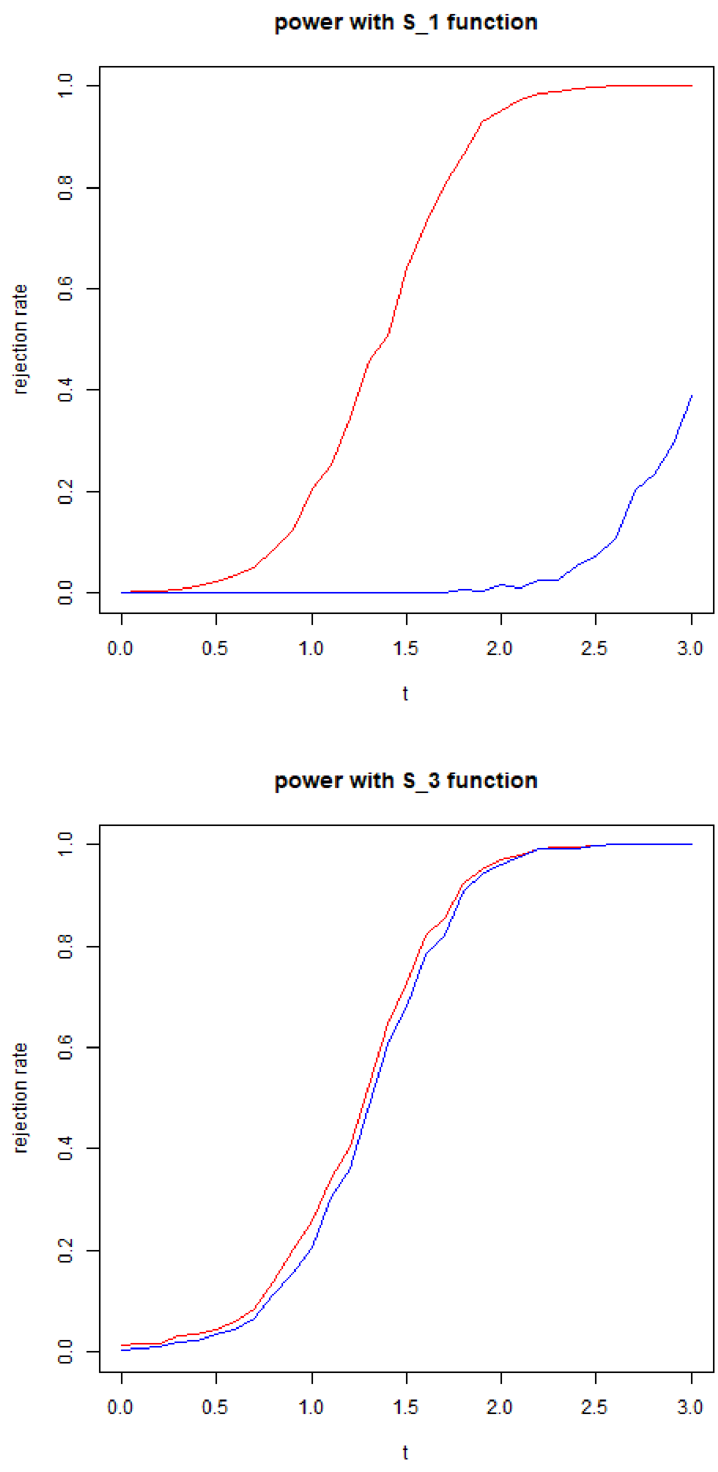

FIG 6. *Power Curve Comparison: Power of Test with Full Set of Inequalities (Blue) vs. Power of Test with Inequalities Corresponding to the Exact Core-determining Class (Red)*

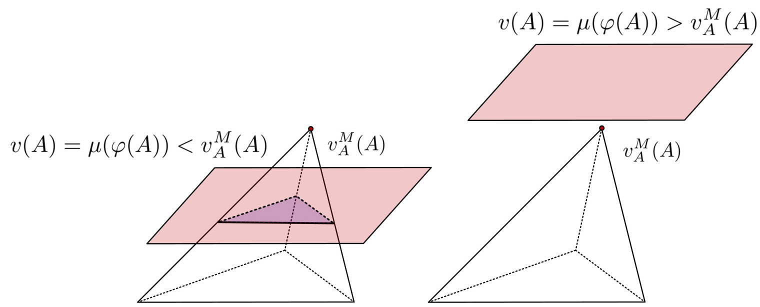

FIG 7. *Example of subset $A \in \mathcal{S}_u$ (left) and $A \notin \mathcal{S}_u$ (right) .*

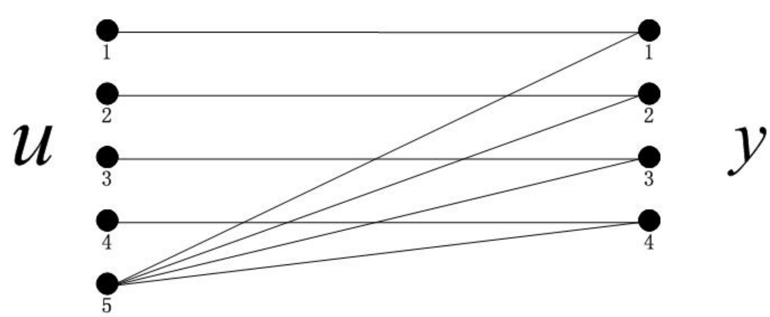

FIG 8. *Correspondence Mapping of Example 4.*

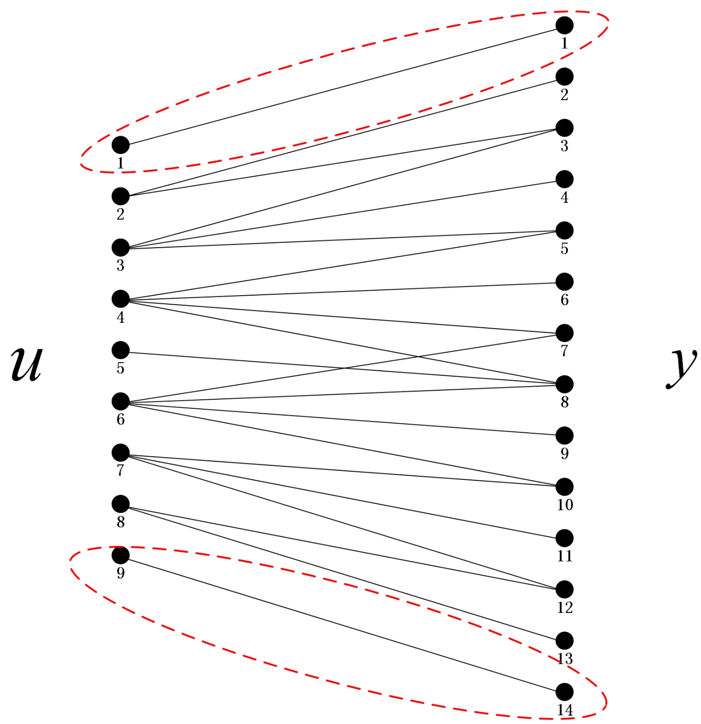

Fig 9. *Correspondence Mapping for Example 5.*

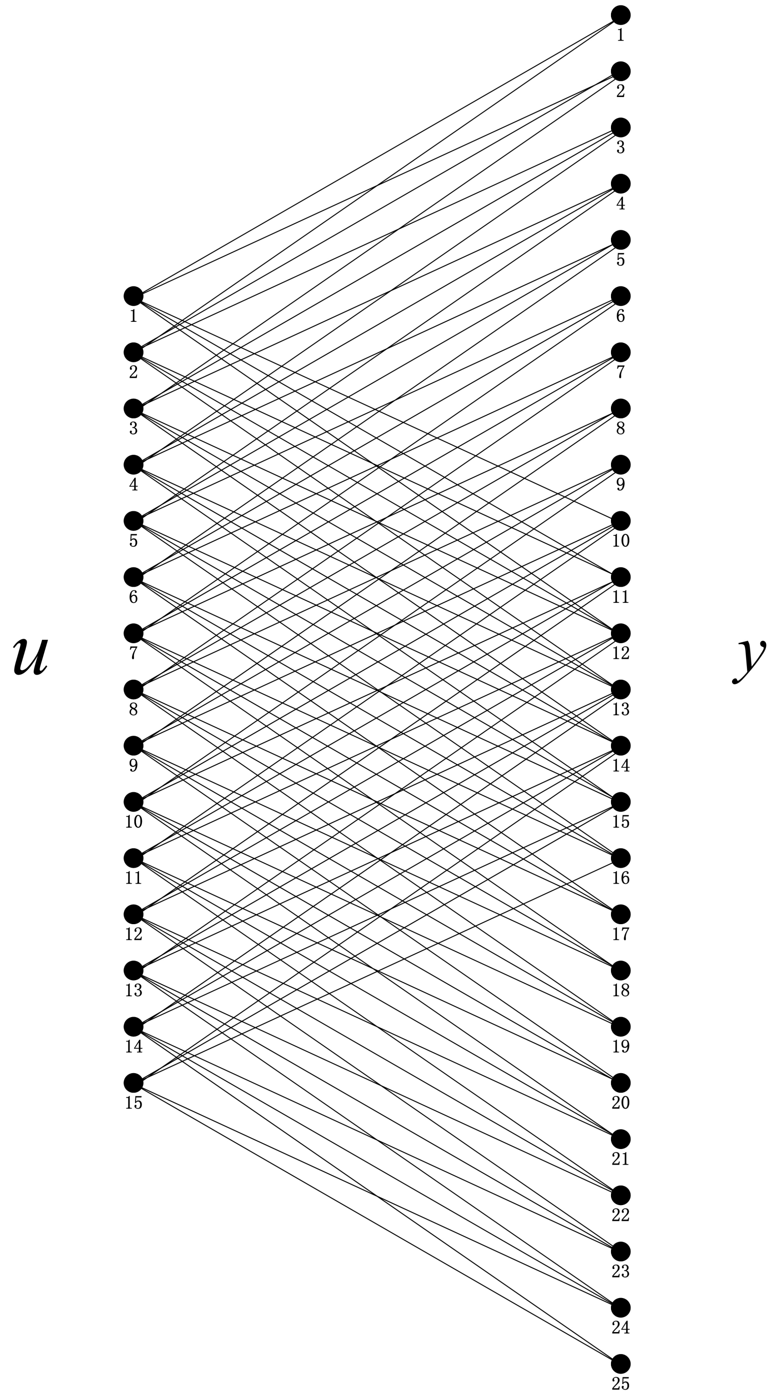

FIG 10. *Correspondence Mapping for Example 6.*