Aggregating Preferences in Multi-Issue Domains by Using Maximum Likelihood Estimators

Lirong Xia **Vincent Conitzer** **Jérôme Lang**

Abstract

In this paper, we study a maximum likelihood estimation (MLE) approach to preference aggregation and voting when the set of alternatives has a multi-issue structure, and the voters' preferences are represented by CP-nets.

We first consider multi-issue domains in which each issue is binary; for these, we propose a general family of *distance-based noise models*, of which give an axiomatic characterization. We then propose a more specific family of natural distance-based noise models that are parameterized by a threshold. We show that computing the winner for the corresponding MLE voting rule is NP-hard when the threshold is 1, but can be done in polynomial time when the threshold is equal to the number of issues.

Next, we consider general multi-issue domains, and study whether and how issue-by-issue voting rules and sequential voting rules can be represented by MLEs. We first show that issue-by-issue voting rules in which each local rule is itself an MLE (resp. a ranking scoring rule) can be represented by MLEs with a weak (resp. strong) decomposability property. Then, we prove two theorems that state that if the noise model satisfies a very weak decomposability property, then no sequential voting rule that satisfies unanimity can be represented by an MLE, unless the number of voters is bounded.

Finally, we propose and study the MLE approach for CP-net aggregators, which take CP-nets as input, and output one or more aggregate CP-nets.

1 Introduction

A natural way for agents to make a joint decision when they have possibly conflicting preferences over a set of alternatives is by *voting*. Each agent (voter) is asked to report her preferences, and then a *voting rule* (or *voting correspondence*) selects the winning alternative (or multiple winning alternatives). Mathematically, a voting rule or correspondence is defined as a mapping from the set of possible preference *profiles* to the set of alternatives. Here, a profile is a vector of all the agents' preferences.

In some sense, this means that the agents' preferences are the "causes" of the joint decision. However, there is a different (and almost reversed) point of view: there is a "correct" joint decision, but the agents may have different perceptions (estimates) of what this correct decision is. Thus, the agents' preferences can be viewed as noisy reports on the correct joint decision. Even in this framework, the agents still need to make a joint decision based on their preferences, and it makes sense to choose their best estimate of the correct decision. Given a noise model, one natural approach is to choose the maximum likelihood estimate of the correct decision. The maximum likelihood estimator is a function from profiles to alternatives (more accurately, subsets of alternatives, since there may be ties), and as such is a voting rule (more accurately, a correspondence).

This maximum likelihood approach was first studied by Condorcet [5] for the case of two and three alternatives. Much later, Young [12] showed that for arbitrary numbers of alternatives, the MLE rule derived from Condorcet's noise model coincides with the Kemeny rule [7]. The approach was further pursued by Drissi and Truchon [6]. More recently, Conitzer and Sandholm [4] studied whether and how common voting rules and *preference functions* (that is, mappings that take agents' preferences as input, and output one or more aggregate rankings of the alternatives) can be represented as maximum likelihood estimators. More recently, the maximum likelihood approach for preference functions has been investigated in more detail [3].

All of the above work does not assume any structure on the set of alternatives. However, in real life, the set of alternatives often has a multi-issue structure: there are multiple *issues* (or *attributes*), each taking values in its respective domain, and an alternative is characterized by the values that the issues take. For example, consider a situation where the citizens of a country vote to directly determine a government plan, composed of multiple sub-plans for several interrelated issues, such as transportation, environment, and health [2]. Clearly, a voter's preferences for one issue in general depend on the decision taken on the other issues: for example, if a new highway is constructed through a forest, a voter may prefer a nature reserve to be established; but if the highway is not constructed, the voter may prefer that no nature reserve is established.

The number of alternatives in a multi-issue domain is exponential in the number of issues, which makes commonly studied voting methods impractical (for one, they require the agents to rank all the alternatives). One straightforward way to aggregate preferences in multi-issue domains is *issue-by-issue* (a.k.a. *seat-by-seat*) voting, which requires that the voters explicitly express their preferences over each issue separately, after which each issue is decided by applying local (issue-wise) voting rules independently. This makes sense if voters' preferences are *separable*, that is, if the preferences of every voter over any issue are independent of the values taken by the other issues. However, if a voter has nonseparable preferences, it is not clear how she should vote in such an issue-by-issue election. Indeed, it is known that natural strategies for voting in such a context can lead to very undesirable results [2, 8].

While in general, a voter's preferences for one issue depend on the decisions taken on other issues, on the other hand, one would not necessarily expect the preferences for one issue to depend on *all* other issues. CP-nets [1] were developed as a natural representation language for capturing such limited dependence among the preferences over multiple issues; they have some obvious similarities to Bayesian networks. Recent work has started to investigate using CP-nets to represent preferences in voting contexts with multiple issues. If there is an order over issues such that every voter's preferences for "later" issues depends only on the decisions made on "earlier" issues, then the voters' CP-nets are acyclic, and a natural approach is to apply issue-wise voting rules *sequentially* [9]. This sequential voting process has a low communication cost, and a low computational cost if each of the local voting rules is easy to compute. While the assumption that such an order exists is still restrictive, it is much less restrictive than assuming that preferences are separable (for one, the resulting preference domain is exponentially larger [9]). Recent extensions of sequential voting rules include order-independent sequential voting rules [11], as well as a framework for voting when preferences are modeled by general (that is, not necessarily acyclic) CP-nets [10].

In this paper, we combine the two research directions: we take a maximum likelihood estimation approach to preference aggregation in multi-issue domains, when the voters' preferences are represented by (not necessarily acyclic) CP-nets. Considering the structure of CP-nets, we focus on probabilistic models that are *very weakly decomposable*. That is, given the "correct" winner, a voter's local preferences over an issue are independent from her local preferences over other issues, as well as from her local preferences over the same issue given a different setting of (at least some of) the other issues.

After reviewing some background, we start with the special case in which each issue has only two possible values. For such domains, we introduce *distance-based noise models*, in which the local distribution over any issue i under some setting of the other issues depends only on the Hamming distance from this setting to the restriction of the "correct" winner to the issues other than i. We axiomatically characterize distance-based noise models by very weak decomposability and *inter-issue neutrality*. Then we focus on *distance-based threshold noise models* in which there is a threshold such that if the distance is smaller than the threshold, then a fixed nonuniform local distribution is used, whereas if the distance is at least as large as the threshold, then a uniform local distribution is used. We study the computational complexity of the two extreme cases of this model: for the case where the threshold is one, we prove that it is NP-hard to compute the winner; but for the case where the threshold is equal to the number of issues, we prove that the winner can be computed in polynomial time.

Then, we move to the general case in which the issues are not necessarily binary. The goal here is to investigate when issue-by-issue or sequential voting rules can be modeled as maximum likelihood estimators. When the input profile is separable, we completely characterize the set of all voting correspondences that can be modeled as an MLE for a noise model satisfying a weak decomposability (resp. strong decomposability) property. Lastly, when the input profile of CP-nets is consistent with a common order over issues, we prove that no sequential voting rule satisfying unanimity can be represented by an MLE, provided the noise model satisfies very weak decomposability. We show that this impossibility result no longer holds if the number of voters is bounded above by a constant.

Finally, we generalize the idea to define MLEs that aggregate CP-nets to a single CP-net or multiple CP-nets (in contrast to a single winner or multiple winners). We show that such MLEs correspond to a family of natural CP-net aggregators that are composed of local MLEs.

2 Technical background

2.1 Basics of voting

Let $\mathcal{X}$ be a finite set of *alternatives* (or *candidates*). A *vote* V is a linear order on $\mathcal{X}$, *i.e.*, a transitive, antisymmetric, and total relation on $\mathcal{X}$. For any $k \leq |\mathcal{X}|$, $(V)_k$ denotes the alternative ranked in the kth position in V; $top(V) = (V)_1$ denotes the alternative that is ranked in the top position in V. The set of all linear orders on $\mathcal{X}$ is denoted by $L(\mathcal{X})$. An n-voter profile P is a collection of n votes, that is, $P = (V_1, \ldots, V_n)$, where $V_j \in L(\mathcal{X})$ for every $j \leq n$. The set of all profiles on $\mathcal{X}$ is denoted by $P(\mathcal{X})$. A *(voting) rule* $r : P(\mathcal{X}) \to \mathcal{X}$ maps any profile to a single candidate (the winner). A *(voting) correspondence* $c : P(\mathcal{X}) \to 2^{\mathcal{X}}$ maps any profile to a subset of candidates. A *preference function* $f : P(\mathcal{X}) \to 2^{L(\mathcal{X})}$ maps any profile to a set of linear orders over $\mathcal{X}$.

2.2 The maximum likelihood approach to voting

In the maximum likelihood approach to voting rules, it is assumed that there is a correct winner $d \in \mathcal{X}$, and each vote V is drawn conditionally independently given d, according to a conditional probability distribution $Pr(V|d)$. The independence structure of the noise model is illustrated in Figure 1. The use of this independence structure is standard. Moreover, if conditional independence among votes is not required, then any voting rule can be represented by an MLE for some noise model [4], which trivializes the question.

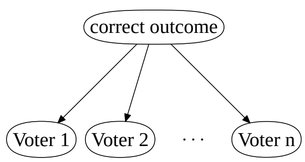


Figure 1: The noise model.

Under this independence assumption, the probability of a profile $P = (V_1, \ldots, V_n)$ given the correct winner d is $Pr(P|d) = \prod_{i=1}^{n} Pr(V_i|d)$ Then, the maximum likelihood estimate of the correct winner is

$$MLE_{Pr}(P) = \arg\max_{d \in \mathcal{X}} Pr(P|d)$$

MLE_{Pr} is a voting correspondence, as there may be several alternatives d that maximize $Pr(P|d)$. Another model that has been studied assumes that there is a correct *ranking* of the alternatives. Here, the model is defined similarly: given the correct linear order V^*, each vote V is drawn conditionally independently according to $Pr(V|V^*)$. The maximum likelihood estimate is defined as follows.

$$MLE_{Pr}(P) = \arg\max_{V^* \in L(\mathcal{X})} \prod_{V \in P} Pr(V|V^*)$$

Definition 1 ([4]) *A voting rule (correspondence) r is a* maximum likelihood estimator for winners under i.i.d. votes (MLEWIV) *if there exists a noise model Pr such that for any profile P, we have that $MLE_{Pr}(P) = r(P)$.*

Definition 2 ([4]) *A preference function f is a* maximum likelihood estimator for rankings under i.i.d. votes (MLERIV) *if there exists a noise model Pr such that for any profile P, we have that $MLE_{Pr}(P) = f(P)$.*

Conitzer and Sandholm studied which common voting rules/preference functions are MLEWIVs/MLERIVs [4]. A *ranking scoring correspondence* c is a correspondence defined by a scoring function $s : L(\mathcal{X}) \times \mathcal{X} \to \mathbb{R}$ in the following way: for any profile P, $c(P) = \arg\max_{\vec{d} \in \mathcal{X}} \sum_{V \in P} s(V, d)$.

2.3 Voting in multi-issue domains

In this paper, the set of all alternatives $\mathcal{X}$ is a *multi-issue domain*. That is, let $\mathfrak{A} = \{\mathbf{x}_1, \ldots, \mathbf{x}_p\}$ ($p \geq 2$) be a set of *issues*, where each issue $\mathbf{x}_i$ takes values in a finite *local domain* D_i. The set of alternatives is $\mathcal{X} = D_1 \times \ldots \times D_p$, that is, an alternative is uniquely identified by its values on all issues. A multi-issue domain is *binary* if for every i we have $D_i = \{0_i, 1_i\}$. For any alternative $\vec{d} = (d_1, \ldots, d_p)$ and any issue $\mathbf{x}_i$, we let $\vec{d}|_{\mathbf{x}_i} = d_i$ and $\vec{d}_{-i} = (d_1, \ldots, d_{i-1}, d_{i+1}, \ldots, d_p)$. For any $I \subseteq \{1, \ldots, p\}$, we let $D_I = \prod_{i \in I} D_i$, and $D_{-i} = D_{\{1,\ldots,i-1,i+1,\ldots,p\}}$.

Example 1 *A group of people must make a joint decision on the menu for dinner (the caterer can only serve the same menu to everyone). The menu is composed of two issues: the main course (**M**) and the wine (**W**). There are three choices for the main course: beef (b), fish (f), or salad (s). The wine can be either red wine (r), white wine (w), or pink wine (p). The set of alternatives is a multi-issue domain: $\mathcal{X} = \{b, f, s\} \times \{r, w, p\}$.*

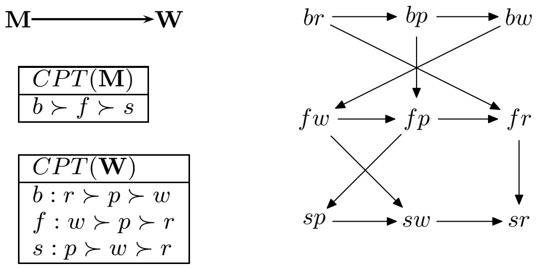
CP-nets [1] constitute a useful language for compactly expressing preferences over multi-issue domains. A CP-net $\mathcal{N}$ over $\mathcal{X}$ consists of two components: (a) a directed graph $G = (\mathfrak{A}, E)$ and (b) a set of conditional linear preferences $\succeq^i_{\vec{u}}$ over D_i, for any setting $\vec{u}$ of the parents of $\mathbf{x}_i$ in G (denoted by $Par_G(\mathbf{x}_i)$). These conditional linear preferences $\succeq^i_{\vec{u}}$ over D_i form the *conditional preference table* for issue $\mathbf{x}_i$, denoted by $CPT(\mathbf{x}_i)$. When G is acyclic, $\mathcal{N}$ is said to be an *acyclic CP-net*.

A CP-net $\mathcal{N}$ induces the partial preorder $\succeq_{\mathcal{N}}$, defined as the transitive closure of $\{(a_i, \vec{u}, \vec{z}) \succeq (b_i, \vec{u}, \vec{z})) \,|\, i \leq p; a_i, b_i \in D_i; \vec{u} \in D_{Par_G(\mathbf{x}_i)}\}$.

It is known [1] that when $\mathcal{N}$ is acyclic, $\succeq_{\mathcal{N}}$ is transitive and asymmetric, that is, a strict partial order. (This is not necessarily the case if $\mathcal{N}$ is not acyclic.) For any graph G' on $\mathfrak{A}$, a CP-net $\mathcal{N}$ is *compatible* with G' if its graph G is a subgraph of G'.

Example 2 *Let $\mathcal{X}$ be the multi-issue domain defined in Example 1. We define a CP-net $\mathcal{N}$ as follows: **M** is the parent of **W**, and the CPTs consist of the following conditional preferences: $CPT(\mathbf{M}) = \{b \succ f \succ s\}, CPT(\mathbf{W}) = \{b : r \succ p \succ w, f : w \succ p \succ r, s : p \succ w \succ r\}$, where $b : r \succ p \succ w$ is interpreted as follows: "when **M** is b, then, r is the most preferred value for **W**, p is the second most preferred value, and w is the least preferred value." $\mathcal{N}$ and its induced partial order $\succeq_{\mathcal{N}}$ (without edges implied by transitivity) are illustrated in Figure 2.*

When all issues are binary, a CP-net $\mathcal{N}$ can be visualized as a hypercube with directed edges in p-dimensional



(a) CP-net $\mathcal{N}$. (b) The partial order induced by $\mathcal{N}$.

Figure 2: An acyclic CP-net $\mathcal{N}$ and its induced partial order.

space, in the following way: each vertex is an alternative, any two adjacent vertices differ in only one component (issue). That is, for any $i \leq p$, any $\vec{d}_{-i} \in D_{-i}$, there is a directed edge connecting $(0_i, \vec{d}_{-i})$ and $(1_i, \vec{d}_{-i})$, and the direction of the edge is from $(0_i, \vec{d}_{-i})$ to $(1_i, \vec{d}_{-i})$ if and only if $(0_i, \vec{d}_{-i}) \succ_{\mathcal{N}} (1_i, \vec{d}_{-i})$.

Example 3 *Let $p = 3$ and let $\mathcal{N}$ be a CP-net defined as follows: the directed graph of $\mathcal{N}$ has an edge from $\mathbf{x}_1$ to $\mathbf{x}_2$ and an edge from $\mathbf{x}_2$ to $\mathbf{x}_3$; the CPTs are $CPT(\mathbf{x}_1) = \{0_1 \succ 1_1\}$, $CPT(\mathbf{x}_2) = \{0_1 : 0_2 \succ 1_2, 1_1 : 1_2 \succ 0_2\}$, $CPT(\mathbf{x}_3) = \{0_2 : 0_3 \succ 1_3, 1_2 : 1_3 \succ 0_3\}$. $\mathcal{N}$ is illustrated in Figure 3 (for simplicity, in the figure, a vertex abc represents the alternative $a_1b_2c_3$, for example, the vertex 000 represents the alternative $0_10_20_3$).*

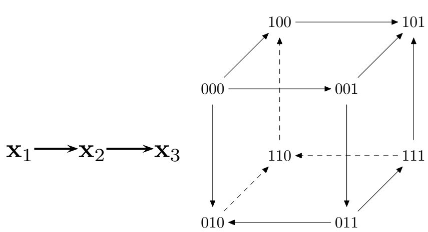


Figure 3: The hypercube representation of the CP-net.

A linear order V *extends* a CP-net $\mathcal{N}$, denoted by $V \sim \mathcal{N}$, if it extends $\succeq_{\mathcal{N}}$. For any setting $\vec{u}$ of $Par_G(\mathbf{x}_i)$, let $V|_{\mathbf{x}_i:\vec{u}}$ and $\mathcal{N}|_{\mathbf{x}_i:\vec{u}}$ denote the the restriction of V (or equivalently, $\mathcal{N}$) to $\mathbf{x}_i$, given $\vec{u}$. That is, $V|_{\mathbf{x}_i:\vec{u}}$ (or $\mathcal{N}|_{\mathbf{x}_i:\vec{u}}$) is the linear order $\succeq^i_{\vec{u}}$.

For any graph G on $\mathfrak{A}$, V is *compatible* with G if there exists a CP-net $\mathcal{N}$ such that $V \sim \mathcal{N}$ and $\mathcal{N}$ is compatible with G. If V is compatible with G, we also say that V is *G-legal*; we say V is *legal* if it is G-legal for some acyclic graph G. A profile is G-legal if all of its votes are G-legal. For any linear order $\mathcal{O}$ on $\mathfrak{A}$, we let $G_{\mathcal{O}}$ be the *graph induced by* $\mathcal{O}$—that is, there is an edge $(\mathbf{x}_i, \mathbf{x}_j)$ in $G_{\mathcal{O}}$ if and only if $\mathbf{x}_i >_{\mathcal{O}} \mathbf{x}_j$. For any directed acyclic graph G, a linear order $\mathcal{O}$ can be found such that $G \subseteq G_{\mathcal{O}}$, which means that any G-legal profile is also $G_{\mathcal{O}}$-legal (which we abbreviate as $\mathcal{O}$-legal). For example, let $\mathcal{N}$ be the CP-net defined in Example 2. Any linear order over $\mathcal{X}$ that extends $\succeq_{\mathcal{N}}$ is $G_{(\mathbf{M}>\mathbf{W})}$-legal (or, equivalently, $(\mathbf{M} > \mathbf{W})$-legal). V is *separable* if and only if it extends a CP-net in which there is no edge. Therefore, any separable vote is $\mathcal{O}$-legal for any ordering $\mathcal{O}$ of issues. We emphasize that votes are not always required to be separable or legal in this paper.

In this paper, we fix $\mathcal{O}$ to be $\mathbf{x}_1 > \ldots > \mathbf{x}_p$. Given a collection of *local rules* $(r_1, \ldots, r_p)$ (where for any $i \leq p$, r_i is a voting rule on D_i), the *sequential composition* of $r_1, \ldots, r_p$ w.r.t. $\mathcal{O}$, denoted by $Seq(r_1, \ldots, r_p)$, is defined for all $\mathcal{O}$-legal profiles as follows: $Seq(r_1, \ldots, r_p)(P) = (d_1, \ldots, d_p) \in \mathcal{X}$, where for any $i \leq p$, $d_i = r_i(P|_{\mathbf{x}_i:d_1\ldots d_{i-1}})$. Thus, the winner is selected in p steps, one for each issue, in the following way: in step i, d_i is selected by applying the local rule r_i to the preferences of voters over D_i, conditioned on the values $d_1, \ldots, d_{i-1}$ that have already been determined for issues that precede $\mathbf{x}_i$. $Seq(r_1, \ldots, r_p)$ is well-defined, because for any G-legal profile, the set of winners is the same for all $\mathcal{O}'$ such that $G \subseteq G_{\mathcal{O}'}$ (see [9]). When G has no edges, $Seq(r_1, \ldots, r_p)$ becomes an *issue-by-issue* voting rule. *Sequential composition* of *local correspondences* $c_1, \ldots, c_p$, denoted by $Seq(c_1, \ldots, c_p)$ is defined in a similar way: for any $\mathcal{O}$-legal profile P, $\vec{d} \in Seq(c_1, \ldots, c_p)(P)$ if and only if for any $i \leq p$, we have that $d_i \in c_i(P|_{\mathbf{x}_i:d_1\ldots d_{i-1}})$.

We will focus on voting methods that only use information about voters' preferences that is represented in the CP-nets that those preferences extend. Therefore, we can consider an input profile to be composed of CP-nets instead of linear orders.

3 Noise models in multi-issue domains

In this section, we extend the maximum likelihood estimation approach to multi-issue domains (where $\mathcal{X} = D_1 \times \ldots \times D_p$). For now, we consider the case where there is a correct winner, $\vec{d} \in \mathcal{X}$. Votes are given by CP-nets and are conditionally independent, given $\vec{d}$. The probability of drawing CP-net $\mathcal{N}$ given that the correct winner is $\vec{d}$ is $\pi(\mathcal{N}|\vec{d})$, where π is some noise model. Given this noise model, for any profile of CP-nets $P = (\mathcal{N}_1, \ldots, \mathcal{N}_n)$, the maximum likelihood estimate of the correct winner is

$$MLE_\pi(P) = \arg\max_{\vec{a}\in\mathcal{X}} \prod_{j=1}^n \pi(\mathcal{N}_j|\vec{a})$$

Again, MLE_π is a voting correspondence.

Even if for all i, $|D_i| = 2$, the number of CP-nets (including cyclic ones) is $2^{p2^{p-1}}$. Hence, to specify a probability distribution over CP-nets, we will assume some structure in this distribution so that it can be compactly represented. Throughout the paper, we will assume that the local preferences for individual issues (given the setting of the other issues) are drawn conditionally independently, both across issues and across settings of the other issues, given the correct winner. More precisely:

Definition 3 *A noise model is* very weakly decomposable *if for every $\vec{d} \in \mathcal{X}$, every $i \leq p$, and every $\vec{a}_{-i} \in D_{-i}$,* *there is a probability distribution $\pi_{\vec{d}}^{\vec{a}_{-i}}$ over $L(D_i)$, so that for every $\vec{d} \in \mathcal{X}$ and every $\mathcal{N} \in$ CPnet($\mathcal{X}$),*

$$\pi(\mathcal{N}|\vec{d}) = \prod_{i \leq p, \vec{a}_{-i} \in D_{-i}} \pi_{\vec{d}}^{\vec{a}_{-i}}(\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}})$$

For instance, if $D_i = \{0_i, 1_i, 2_i\}$, $\pi_{\vec{d}}^{\vec{a}_{-i}}(0_i \succ 2_i \succ 1_i)$ is the probability that the CP-net of a given voter contains $\vec{a}_{-i} : 0_i \succ 2_i \succ 1_i$, given that the correct winner is $\vec{d}$. Then, the probability of CP-net $\mathcal{N}$ is the product of the probabilities of all its local preferences $\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}}$ over specific $\mathbf{x}_i$ given specific $\vec{a}_{-i}$ (which contains the setting for $\mathbf{x}_i$'s parents as a sub-vector), when the winner is $\vec{d}$. (We will introduce stronger decomposability notions soon.)

Assuming very weak decomposability is reasonable in the sense that a voter's preferences for one issue are not directly linked to her *preferences* for another issue. We note that this is completely different from saying that the voter's preferences for an issue do not depend on the *values* of the other issues. Indeed, the voter's preferences for an issue can, at least in principle, change drastically depending on the values of the other issues.

However, we do not want to argue that such a distribution always generates realistic preferences. In fact, with some probability, such a distribution generates cyclic preferences. This is not a problem, in the sense that the purpose of the maximum likelihood approach is to find a natural voting rule that maps profiles to outcomes. The fact that this rule is also defined for cyclic preferences does not hinder its application to acyclic preferences. Similarly, Condorcet's original noise model for the single-issue setting also generates cyclic preferences with some probability, but this does not prevent us from applying the corresponding (Kemeny) rule [7] to acyclic preferences.

Even assuming very weak decomposability, we still need to define exponentially many probabilities. We will now introduce some successive strengthenings of the decomposability notion. First, we introduce *weak decomposability*, which removes the dependence of an issue's local distribution on the settings of the other issues *in the correct winner*.

Definition 4 *A very weakly decomposable noise model π is* weakly decomposable *if for any $i \leq p$, any $\vec{d}_1, \vec{d}_2 \in \mathcal{X}$ such that $\vec{d}_1|_{\mathbf{x}_i} = \vec{d}_2|_{\mathbf{x}_i}$, we must have that for any $\vec{a}_{-i} \in D_{-i}$, $\pi_{\vec{d}_1}^{\vec{a}_{-i}} = \pi_{\vec{d}_2}^{\vec{a}_{-i}}$. Let $WD(\mathcal{X})$ denote the set of correspondences that are the MLE for some weakly decomposable noise model.*

Next, we introduce an even stronger notion, namely *strong decomposability*, which removes all dependence of an issue's distribution on the settings of the other issues. That is, the local distribution only depends on the value of that issue in the correct winner.

Definition 5 *A very weakly decomposable noise model π is* strongly decomposable *if it is weakly decomposable, and for any $i \leq p$, any $\vec{a}_{-i}, \vec{b}_{-i} \in D_{-i}$, any $\vec{d} \in \mathcal{X}$, we must have that $\pi_{\vec{d}}^{\vec{a}_{-i}} = \pi_{\vec{d}}^{\vec{b}_{-i}}$. Let $SD(\mathcal{X})$ denote the set of correspondences that are the MLE for some strongly decomposable noise model.*

4 Distance-based models

In this section, we focus on maximum likelihood estimators that are based on noise models defined over binary multi-issue domains (domains composed of binary issues). We recall that a CP-net on a binary multi-issue domain corresponds to a directed hypercube in which each edge has a direction representing the local preferences. A very weakly decomposable noise model π can be represented by multiple weighted directed hypercubes, one for each correct winner, in which the weight of each directed edge is the probability of the local preferences represented by the directed edge. For example, when the correct winner is $0_1 0_2 0_3$, the weight on the directed edge $(0_1 1_2 0_3, 0_1 1_2 1_3)$ is the probability $\pi_{0_1 0_2 0_3}^{0_1 1_2}(0_3 \succ 1_3)$. We now propose and study very weakly decomposable noise models in which the weight of each edge depends only on the Hamming distance between the edge and the correct winner. First we define the Hamming distance between two alternatives and between an alternative and an edge in the hypercube.

For any pair of alternatives $\vec{d}, \vec{d}' \in \mathcal{X}$, the *Hamming distance* between $\vec{d}$ and $\vec{d}'$, denoted by $|\vec{d} - \vec{d}'|$, is the number of components in which $\vec{d}$ is different from $\vec{d}'$, that is, $|\vec{d} - \vec{d}'| = \#\{i \leq p : d_i \neq d'_i\}$. Let $e = (\vec{d}_1, \vec{d}_2)$ be a pair of alternatives such that $|\vec{d}_1 - \vec{d}_2| = 1$ (equivalently, an edge in the hypercube representation of $\mathcal{X}$). The distance between e and an alternative $\vec{d} \in \mathcal{X}$, denoted by $|e - \vec{d}|$, is the smaller Hamming distance between $\vec{d}$ and the two ends of e, that is, $|e - \vec{d}| = \min\{|\vec{d}_1 - \vec{d}|, |\vec{d}_2 - \vec{d}|\}$. For example, $|0_1 1_2 0_3 - 0_1 0_2 0_3| = 1$, $|0_1 1_2 1_3 - 0_1 0_2 0_3| = 2$, and $|(0_1 1_2 0_3, 0_1 1_2 1_3) - 0_1 0_2 0_3| = 1$.

We next introduce *distance-based noise models* in which the probability distribution $\pi_{\vec{d}}^{\vec{a}_{-i}}$ only depends on d_i and the Hamming distance between $\vec{a}_{-i}$ and $\vec{d}_{-i}$.

Definition 6 *Let $\mathcal{X}$ be a binary multi-issue domain. For any $\vec{q} = (q_0, \ldots, q_{p-1})$ such that $1 > q_0, \ldots, q_{p-1} > 0$, a* distance-based (noise) model *$\pi_{\vec{q}}$ is a very weakly decomposable noise model such that for any $\vec{d} \in \mathcal{X}$, any $i \leq p$, and any $\vec{a}_{-i} \in D_{-i}$ with $|\vec{a}_{-i} - \vec{d}_{-i}| = k \leq p-1$, we have that $\pi_{\vec{d}}^{\vec{a}_{-i}}(d_i \succ \bar{d}_i) = q_k$.*

Given the correct winner $\vec{d}$, a distance-based model $\pi_{\vec{q}}$ can be visualized by the following weighted directed graph built on the hypercube:

- For any undirected edge $e = (\vec{d}_1, \vec{d}_2)$ in the hypercube, where $\vec{d}_1, \vec{d}_2$ differ only on the value assigned to $\mathbf{x}_i$ for some $i \leq p$, if $\vec{d}_1|_{\mathbf{x}_i} = d_i$, then the direction of e is from $\vec{d}_1$ to $\vec{d}_2$; if $\vec{d}_2|_{\mathbf{x}_i} = d_i$, then the direction of e is from $\vec{d}_2$ to $\vec{d}_1$. That is, the direction of the edge is always from the alternative whose $\mathbf{x}_i$ component is the same as the $\mathbf{x}_i$ com-

ponent of the correct winner to the other end of the edge.
• For any edge e with $|e - \vec{d}| = l$, the weight of e is q_l.

For example, given that $0_1 0_2 0_3$ is the correct winner, the distance-based model is illustrated in Figure 4. To char-

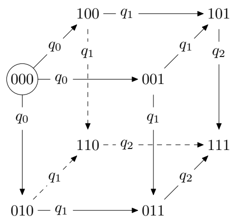


Figure 4: The distance-based model $\pi_{(q_0,q_1,q_2)}$ when the correct winner is 000.

acterize distance-based models, we first define *inter-issue permutations*. Intuitively, an inter-issue permutation is a permutation that exchanges two issues.

Definition 7 *Let $i, j \leq p$. An* inter-issue permutation *is a permutation $m_{i,j}$ on $D_1 \cup \ldots \cup D_p$ satisfying: (1) for any $k \neq i, j$ and any $d_k \in D_k$, $m_{i,j}(d_k) = d_k$ and (2) for any $d_i \in D_i$, $m_{i,j}(d_i) \in D_j$; for any $d_j \in D_j$, $m_{i,j}(d_j) \in D_i$; and for any $k \in \{i, j\}$ and any $d_k \in D_k$, $m_{i,j}(d_k) = (m_{i,j})^{-1}(d_k)$.*

$m_{i,j}$ induces a permutation $M_{i,j}$ on the set of all subvectors of any $\vec{d} \in \mathcal{X}$ as follows: for any $I \subseteq \mathfrak{A}$ and $\vec{d}_I = (d_{i_1}, \ldots, d_{i_{|I|}}) \in D_I$, $M_{i,j}(d_I) = (m_{i,j}(d_{i_1}), \ldots, m_{i,j}(d_{i_{|I|}}))$. For example, let $p = 3$, and $m_{1,2}$ be an inter-issue permutation such that $m_{1,2}(0_1) = 1_2$. Then we have $M_{1,2}(1_1) = 0_2$, $M_{1,2}(0_2) = 1_1$, $M_{1,2}(1_2) = 0_1$; $M_{1,2}(0_1 0_2 0_3) = 1_1 1_2 0_3$, $M_{1,2}(1_1 1_3) = 0_2 1_3$.

We note that since each issue is binary, there are exactly two ways of exchanging issue $\mathbf{x}_i$ and $\mathbf{x}_j$: either map 0_i to 0_j (and 1_i to 1_j), or map 0_i to 1_j (and 1_i to 0_j).

Definition 8 *A very weakly decomposable noise model π satisfies* inter-issue neutrality *if for any $i, j \leq p$, any inter-issue permutation $m_{i,j}$ (which induces $M_{i,j}$), any $i' \leq p$, any $\vec{d} \in \mathcal{X}$, and any $\vec{a}_{-i'} \in D_{-i'}$, we have that $\pi_{\vec{d}}^{\vec{a}_{-i'}}(0_{i'} \succ 1_{i'}) = \pi_{M_{i,j}(\vec{d})}^{M_{i,j}(\vec{a}_{-i'})}(m_{i,j}(0_{i'}) \succ m_{i,j}(1_{i'}))$.*

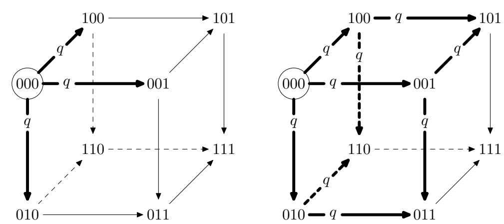
Thus, the noise model π satisfies inter-issue neutrality if after exchanging any two issues, the resulting noise model is still π. Or equivalently, π is indifferent to the names of the issues as well as the names of the values they take. We next show that the class of distance-based models can be completely characterized as the class of noise models that satisfy very weak decomposability and inter-issue neutrality.

Theorem 1 *Let $\mathcal{X}$ be a binary multi-issue domain. A very weakly decomposable noise model π is a distance-based noise model if and only if it satisfies inter-issue neutrality.*

The proofs of all theorems are in the appendix, which is uploaded separately as the supplementary material.

We are especially interested in a special type of distance-based model in which there exists a threshold $1 \leq k \leq p$ and $q > \frac{1}{2}$, such that for any $i < k$, we have that $q_i = q$, and for any $k \leq i \leq p - 1$, we have that $q_i = \frac{1}{2}$. Such a model is denoted by $\pi_{k,q}$. We call $\pi_{k,q}$ a *distance-based threshold noise model* with threshold k. We say that a noise model π has threshold $k \leq p$ if and only if there exists $q > \frac{1}{2}$ such that $\pi = \pi_{k,q}$. The MLE for a distance-based threshold model $\pi_{k,q}$ is denoted by $MLE_{\pi_{k,q}}$.

Example 4 *Let $p = 3$. $\pi_{1,q}$ and $\pi_{2,q}$ are illustrated in Figure 5 (when the correct winner is 000).*



(a) The threshold is 1. (b) The threshold is 2.

Figure 5: Distance-based threshold models. The weight of the bold edges is $q > \frac{1}{2}$; the weight of all other edges is $\frac{1}{2}$.

The following theorem provides an axiomatic characterization of the set of all noise models that have threshold p, which is the number of issues. This axiomatization is similar to the one in Theorem 1.

Theorem 2 *Let $\mathcal{X}$ be a binary multi-issue domain. A noise model π is a distance-based threshold noise model with threshold p if and only if π is strongly decomposable and satisfies inter-issue neutrality.*

We next present a direct method for computing winners under the MLE correspondences of distance-based threshold models. For any $1 \leq k \leq p$, any $\vec{d} \in \mathcal{X}$, and any CP-net $\mathcal{N}$, we define the *consistency of degree* k between $\vec{d}$ and $\mathcal{N}$, denoted by $N_k(\vec{d}, \mathcal{N})$, as follows. $N_k(\vec{d}, \mathcal{N})$ is the number of triples $(\vec{a}, \vec{b}, i)$ such that $\vec{a}_{-i} = \vec{b}_{-i}$, $a_i = d_i$, $b_i = \bar{d}_i$, $|(a_i, b_i) - \vec{d}| \leq k - 1$, and $\mathcal{N}$ contains $a_{-i} : d_i \succ \bar{d}_i$. That is, $N_k(\vec{d}, \mathcal{N})$ is the number of local preferences (over any issue $\mathbf{x}_i$, given any $\vec{a}_{-i} \in D_{-i}$) in $\mathcal{N}$ that are $d_i \succ \bar{d}_i$, and the distance between $\vec{d}$ and the edge $((d_i, \vec{a}_{-i}), (\bar{d}_i, \vec{a}_{-i}))$ is at most $k - 1$. For any profile P_{CP} of CP-nets, we let $N_k(\vec{d}, P_{CP}) = \sum_{\mathcal{N} \in P_{CP}} N_k(\vec{d}, \mathcal{N})$.

Proposition 1 *For any $k \leq p$, any $q > \frac{1}{2}$, and any profile P_{CP} of CP-nets, we have that $MLE_{\pi_{k,q}}(P_{CP}) = \arg\max_{\vec{d}} N_k(\vec{d}, P_{CP})$.*

That is, the winner for any profile of CP-nets under any MLE for a distance-based threshold model $\pi_{k,q}$ maximizes the sum of the consistencies of degree k between the winning alternative and all CP-nets in the profile. Therefore, we have the following corollary, which states that the winners for any profile under $MLE_{\pi_{k,q}}$ do not depend on q, provided that $q > \frac{1}{2}$.

Corollary 1 *For any $k \leq p$, any $q_1 > \frac{1}{2}, q_2 > \frac{1}{2}$, and any profile P_{CP} of CP-nets, we have $MLE_{\pi_{k,q_1}}(P_{CP}) = MLE_{\pi_{k,q_2}}(P_{CP})$.*

We next investigate the computational complexity of applying MLE rules with distance-based threshold models. First, we present a polynomial-time algorithm that computes the winners and outputs the winners in a compact way, under $MLE_{\pi_{p,q}}$, where p is the number of issues. This algorithm computes the correct value(s) of each issue separately: for any issue $\mathbf{x}_i$, the algorithm counts the number of tuples $(\vec{a}_{-i}, \mathcal{N})$, where $\vec{a}_{-i} \in D_{-i}$ and $\mathcal{N}$ is a CP-net in the input profile P_{CP}, such that $\mathcal{N}$ contains $a_{-i} : 0_i \succ 1_i$. If there are more tuples $(\vec{a}_{-i}, \mathcal{N})$ in which $\mathcal{N}$ contains $a_{-i} : 0_i \succ 1_i$ than there are tuples in which $\mathcal{N}$ contains $a_{-i} : 1_i \succ 0_i$, then we select 0_i to be the ith component of the winning alternative, and vice versa. We note that the time required to count tuples $(\vec{a}_{-i}, \mathcal{N})$ depends on the size of $\mathcal{N}$. Therefore, even though computing the value for $\mathbf{x}_i$ takes time that is exponential in $|Par_G(\mathbf{x}_i)|$ (the number of parents of $\mathbf{x}_i$ in the directed graph of $\mathcal{N}$), the CPT of $\mathbf{x}_i$ in $\mathcal{N}$ itself is also exponential in $|Par_G(\mathbf{x}_i)|$ (for each setting of $Par_G(\mathbf{x}_i)$, there is an entry in $CPT(\mathbf{x}_i)$). This explains why the algorithm runs in polynomial time.

Algorithm 1 **INPUT:** $p \in \mathbb{N}$, $1 > q > \frac{1}{2}$, and a profile of CP-nets P_{CP}.

1. For each $i \leq p$:

- **1a.** Let $S_i = 0$, $W_i = \emptyset$.
- **1b.** For each CP-net $\mathcal{N} \in P_{CP}$: let $Par_G(\mathbf{x}_i) = \{\mathbf{x}_{i_1}, \ldots, \mathbf{x}_{i_{p'}}\}$ be the parents of $\mathbf{x}_i$ in the directed graph of $\mathcal{N}$. Let l be the number of settings $\vec{y}$ of $Par_G(\mathbf{x}_i)$ for which $\mathcal{N}|_{\mathbf{x}_i:\vec{y}} = 0_i \succ 1_i$. Let $S_i \leftarrow S_i + l2^{p-p'} - 2^{p-1}$. Here $l2^{p-p'} - 2^{p-1}$ is the number of edges in the CP-net where $0_i \succ 1_i$, minus the number of edges where $1_i \succ 0_i$.
- **1c.** At this point, let $W_i = \begin{cases} \{0_i\} & \text{if } S_i > 0 \\ \{1_i\} & \text{if } S_i < 0 \\ \{0_i, 1_i\} & \text{if } S_i = 0 \end{cases}$

2. Output $W_1 \times \ldots \times W_p$.

Proposition 2 *The output of Algorithm 1 is $MLE_{\pi_{p,q}}(P_{CP})$, and the algorithm runs in polynomial time.*

However, when the threshold is one, computing the winners is NP-hard, and the associated decision problem, namely checking whether there exists an alternative $\vec{d}$ such that $N_1(\vec{d}, P_{CP}) \geq T$, is NP-complete.

Theorem 3 *It is NP-hard to find a winner under $MLE_{\pi_{1,q}}$. More precisely, it is NP-complete to decide whether there exists an alternative $\vec{d}$ such that $N_1(\vec{d}, P_{CP}) \geq T$.*

5 Characterizations of MLE correspondences

The voting rules studied in Section 4 are quite different from the voting rules that have previously been studied in the context of multi-issue domains, such as issue-by-issue voting and sequential voting. This illustrates that the maximum likelihood approach can generate sensible new rules for multi-issue domains. Nevertheless, we may wonder whether previously studied rules also fit under the MLE framework.

In this section, we study whether or not issue-by-issue and sequential voting correspondences can be modeled as the MLEs for very weakly decomposable noise models. We note that voting rules (which always output a unique winner) are a special case of voting correspondences. Therefore, our results easily extend to the case of voting rules. First, we restrict the domain to separable profiles, and characterize the set of all correspondences that can be modeled as the MLEs for strongly/weakly decomposable noise models.

Theorem 4 *Over the domain of separable profiles, a voting correspondence c can be modeled as the MLE for a strongly decomposable noise model if and only if c is an issue-by-issue voting correspondence composed of MLEWIVs.*

Theorem 5 *Over the domain of separable profiles, a voting correspondence c can be modeled as the MLE for a weakly decomposable noise model if and only if c is an issue-by-issue voting correspondence composed of ranking scoring correspondences.*

However, for sequential voting correspondences, we have the following negative result. A voting correspondence c satisfies *unanimity* if for any profile P in which each vote ranks an alternative $\vec{d}$ first, we have $r(P) = \{\vec{d}\}$.

Theorem 6 *Let $Seq(c_1, \ldots, c_p)$ be a sequential voting correspondence that satisfies unanimity. Over the domain of $\mathcal{O}$-legal profiles, there is no very weakly decomposable noise model such that $Seq(c_1, \ldots, c_p)$ is the MLE.*

However, a positive result can be obtained if there is an upper bound on the number of voters. The next theorem states that for any natural number n and any sequential composition of MLEWIVs, there exists a very weakly decomposable noise model such that for any profile of no more than n $\mathcal{O}$-legal votes, the set of winners under the MLE for that noise model is always a subset of the set of winners under the sequential correspondence.

Theorem 7 *For any $n \in \mathbb{N}$ and any sequential voting correspondence $Seq(c_1, \ldots, c_p)$ where for each $i \leq p$, c_i is an MLEWIV, there exists a very weakly decomposable noise model π such that for any $\mathcal{O}$-legal profile P composed of no more than n votes, we have that $MLE_\pi(P) \subseteq Seq(c_1, \ldots, c_p)(P)$.*

6 Noise models for CP-net aggregators

So far, we have only considered using voting to obtain a winning alternative. Generating a full ranking of all alternatives is impractical in multi-issue domains. However, we can try to generate an aggregate CP-net that represents the aggregate preferences. We call such a mapping a *CP-net aggregator*. Formally, a CP-net aggregator R is defined as a mapping from the set of all profiles of CP-nets to the set of CP-nets.

To model CP-net aggregators as MLEs, we define noise models δ similarly as in the case of voting correspondences. We focus on weakly decomposable models: for any $i \leq p$, any $\vec{a}_{-i} \in D_{-i}$, and any $W^i \in L(D_i)$, there is a probability distribution $\delta_{W^i}^{\vec{a}_{-i}}$ over $L(D_i)$. For any profile P_{CP} composed of CP-nets, we define the maximum likelihood estimate as follows:

$$MLE_\delta(P_{CP}) = \arg\max_{\mathcal{N}} \prod_{\mathcal{N}' \in P_{CP}} \prod_{i \leq p, \vec{a}_{-i} \in D_{-i}} \delta_{\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}}}^{\vec{a}_{-i}} (\mathcal{N}'|_{\mathbf{x}_i:\vec{a}_{-i}})$$

δ is strongly decomposable if (a) it is weakly decomposable, and (b) for every issue $\mathbf{x}_i$, the probability distribution over D_i does not depend on the choice of $\vec{a}_{-i}$: formally, for any $i \leq p$, any $\vec{a}_{-i}, \vec{b}_{-i} \in D_{-i}$, and any $V^i, W^i \in L(D_i)$, we must have that $\delta_{W^i}^{\vec{a}_{-i}}(V^i) = \delta_{W^i}^{\vec{b}_{-i}}(V^i)$.

Let $f_1, \ldots, f_p$ be local preference functions. We define the *issue-by-issue composition* of $f_1, \ldots, f_p$, denoted by $Com(f_1, \ldots, f_p)$, as follows: for any profile P, any $i \leq p$, and any $\vec{d}_{-i} \in D_{-i}$, we have that $Com(f_1, \ldots, f_p)(P)|_{\mathbf{x}_i:\vec{d}_{-i}} = f_i(P|_{\mathbf{x}_i:\vec{d}_{-i}})$.

Proposition 3 *If P is $\mathcal{O}$-legal, then $Com(f_1, \ldots, f_p)(P)$ is $\mathcal{O}$-legal.*

The next theorem characterizes all CP-net aggregators that can be modeled as the MLE for a strongly decomposable noise model.

Theorem 8 *Over the domain of profiles of CP-nets, a CP-net aggregator R can be modeled as an MLE for a strongly decomposable noise model if and only if it is an issue-by-issue CP-net aggregator composed of local MLERIVs.*

7 Conclusion

The central problems in preference aggregation in multi-issue domains are to find practical ways for voters to represent and report their preferences, as well as to find natural and computationally feasible ways of aggregating these reported preferences. In this paper, we considered the maximum likelihood estimation (MLE) approach to voting, and generalized it to multi-issue domains, assuming that the voters' preferences are expressed by CP-nets. In the case where all issues are binary, we proposed and axiomatized a class of distance-based noise models; then, we focused on a specific subclass of such models, parameterized by a threshold. We identified the computational complexity of winner determination for the two most relevant values of the threshold (it is NP-hard for one, but doable in polynomial time for the other; we also gave an axiomatization of the latter). We then considered the case where issues are not necessarily binary. For separable input profiles, we characterized MLEs of strongly/weakly decomposable models as issue-by-issue voting correspondences composed of local MLEWIVs/ranking scoring correspondences. Although we showed that no sequential voting correspondence can be represented as the MLE for a very weakly decomposable model, we did obtain a positive result here under the assumption that the number of voters is bounded above by a constant. Finally, we studied the MLE approach for CP-net aggregators, and characterized all CP-net aggregators that can be modeled as the MLE for a strongly decomposable noise model.

We note that, whereas Section 5 has a non-constructive flavor because we studied existing voting mechanisms and Theorem 6 is an impossibility theorem, quite the opposite is the case for Section 4. Indeed, the MLE principle led us to define genuinely new families of voting rules and correspondences for multi-issue domains. These rules are radically different from the rules that had previously been proposed and studied for these domains. Unlike sequential or issue-by-issue rules, they do not require any domain restriction, and yet their computational complexity is not that bad (NP-complete at worst, and sometimes polynomial in the size of the CP-nets). We believe that this new family of rules is very original and promising.

Future research could further investigate the computational aspects of determining the winners for MLE correspondences. For example, in this paper, we characterized the complexity of computing winners under MLEs of distance-based threshold models with thresholds 1 and p (the number of issues). It would be interesting to identify the complexity for other thresholds (however, we conjecture that it is at least NP-hard). More generally, the study of voting in multi-issue domains is still in its infancy. Unlike in the standard (single-issue) case, relatively few rules have been proposed and relatively little is known about social-choice-theoretic properties. We believe that this paper has demonstrated the potential of the maximum likelihood approach to build a theory of social choice in multi-issue domains.

References

[1] Craig Boutilier, Ronen Brafman, Carmel Domshlak, Holger Hoos, and David Poole. CP-nets: a tool for representing and reasoning with conditional ceteris paribus statements. *Journal of Artificial Intelligence Research*, 21:135–191, 2004.

[2] Steven Brams, D. Marc Kilgour, and William Zwicker. The paradox of multiple elections. *Social Choice and Welfare*, 15(2):211–236, 1998.

[3] Vincent Conitzer, Matthew Rognlie, and Lirong Xia. Preference functions that score rankings and maximum likelihood estimation. In *COMSOC-08*, 2008.

[4] Vincent Conitzer and Tuomas Sandholm. Common voting rules as maximum likelihood estimators. In *Proceedings of UAI-05*, pages 145–152, 2005.

[5] Marie Jean Antoine Nicolas de Caritat (Marquis de Condorcet). Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix. 1785. Paris: L'Imprimerie Royale.

[6] Mohamed Drissi and Michel Truchon. Maximum likelihood approach to vote aggregation with variable probabilities. Technical Report 0211, Departement d'economique, Universite Laval, 2002.

[7] John Kemeny. Mathematics without numbers. *Daedalus*, 88:575–591, 1959.

[8] Dean Lacy and Emerson M.S. Niou. A problem with referendums. *Journal of Theoretical Politics*, 12(1):5–31, 2000.

[9] Jérôme Lang and Lirong Xia. Sequential composition of voting rules in multi-issue domains. *Mathematical Social Sciences*, 57(3):304–324, 2009.

[10] Lirong Xia, Vincent Conitzer, and Jérôme Lang. Voting on multiattribute domains with cyclic preferential dependencies. In *Proceedings of AAAI-08*, pages 202–207, 2008.

[11] Lirong Xia, Jérôme Lang, and Mingsheng Ying. Strongly decomposable voting rules on multiattribute domains. In *Proceedings of AAAI-07*, pages 776–781, 2007.

[12] H. Peyton Young. Optimal voting rules. *Journal of Economic Perspectives*, 9(1):51–64, 1995.

The appendix: proofs

Proof of Theorem 1: We define *intra-issue permutation*, denoted by $\gamma = (\gamma_1, \ldots, \gamma_p)$, as follows: for any $i \le p$, γ_i is a permutation over D_i (that is, either γ_i exchanges 0_i and 1_i, or γ_i is the identity permutaition), and for any $\vec{d} \in \mathcal{X}$, we have that $\gamma(\vec{d}) = (\gamma_1(d_1), \ldots, \gamma_p(d_p))$.

It is easy to check that every distance-based noise model is very weakly decomposable and satisfies inter-issue neutrality. We then prove that if a noise model π is very weakly decomposable and satisfies inter-issue neutrality, then it must be a distance-based noise model. It suffices to prove the following two claims.

1. For any $i, i' \le p$, any pair $\vec{a}_{-i} \in D_{-i}, \vec{b}_{-i'} \in D_{-i'}$ in which the numbers of issues that take 1 are the same, we have that $\pi^{\vec{a}_{-i}}_{0_1\ldots 0_p}(0_i \succ 1_i) = \pi^{\vec{b}_{-i'}}_{0_1\ldots 0_p}(0_{i'} \succ 1_{i'})$.

2. π satisfies intra-issue neutrality. That is, for every intra-issue permutation $\gamma = (\gamma_1, \ldots, \gamma_p)$, every $i \le p$, every $\vec{a}_{-i} \in D_{-i}$ and every $\vec{d} \in \mathcal{X}$, we have that $\pi^{\vec{a}_{-i}}_{\vec{d}}(d_i \succ \bar{d}_i) = \pi^{\gamma(\vec{a}_{-i})}_{\gamma(\vec{d})}(\gamma_i(d_i) \succ \gamma_i(\bar{d}_i))$.

We first prove 1. Let $I = \{i_1, \ldots, i_k\} \subseteq \{1, \ldots, p\}$ be the set of components such that for any $l \in I$, the $\mathbf{x}_l$ component of $\vec{a}_{-i}$ is 0_l and the $\mathbf{x}_l$ component of $\vec{b}_{-j}$ is 1_l; let $I' = \{i'_1, \ldots, i'_k\} \subseteq \{1, \ldots, p\}$ be the set of components such that for any $l \in I$, the $\mathbf{x}_l$ component of $\vec{a}_{-i}$ is 1_l and the $\mathbf{x}_l$ component of $\vec{b}_{-j}$ is 0_l. We consider the following inter-issue permutations: $M_{i_1,i'_1}, \ldots, M_{i_k,i'_k}, M_{i,i'}$ such that for any $l \le k$, $M_{i_1,i'_1}(0_{i_k}) = 0_{i'_k}$, $M_{i,i'}(0_i) = 0_{i'}$. Let $M = M_{i,i'} \circ M_{i_1,i'_1} \circ \ldots \circ M_{i_k,i'_k}$, where for any two functions f_1 and f_2, $f_1 \circ f_2(x) = f_1(f_2(x))$. We note that $M(\vec{a}_{-i}) = b_{-i'}$, $M(0_1 \ldots 0_p) = 0_1 \ldots 0_p$, and $M(0_i) = 0_{i'}$. Therefore, from the inter-issue neutrality we have that $\pi^{\vec{a}_{-i}}_{0_1\ldots 0_p}(0_i \succ 1_i) = \pi^{M(\vec{a}_{-i})}_{M(0_1\ldots 0_p)}(M(0_i) \succ M(1_i)) = \pi^{\vec{b}_{-i'}}_{0_1\ldots 0_p}(0_{i'} \succ 1_{i'})$.

Next, we prove 2. For any $i \le p$, we let $m'_{i,i}$ be the inter-issue permutation in which $m'_{i,i}(0_i) = 1_i$ (and $m'_{i,1}(1_i) = 0_i$). We note that any intra-issue permutation is equivalent to the composition of multiple inter-issue permutation in the following way: suppose for any $i \in \{i_1, \ldots, i_k\}$, we have that γ_i exchanges 0_i and 1_i; for any $i \not\in \{i_1, \ldots, i_k\}$, we have that γ_i is the identity permutation. Then, $\gamma = M'_{i_1,i_1} \circ \ldots \circ M'_{i_k,i_k}$. Because π satisfies inter-issue neutrality, it also satisfies intra-issue neutrality.

(End of the proof of Theorem 1.) □

Proof of Proposition 1: For any $k \le p$, any $\vec{d} \in \mathcal{X}$, we let $L_k = \#\{e : |e - \vec{d}| \le k\}$. That is, L_k is the number of edges in the hypercube whose distance from a given alternative $\vec{d}$ is no more than k. For any $\vec{d} \in \mathcal{X}$ and any CP-net $\mathcal{N}$, we have that

$$
\begin{aligned}
&\ln \pi(\mathcal{N}|\vec{d}) \\
&= \ln \prod_{i, \vec{a}_{-i} \in D_{-i}} \pi^{\vec{a}_{-i}}_{d_i}(\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}}) \\
&= \sum_{\mathcal{N} \in P_{CP}} (N_k(\vec{d}, \mathcal{N}) \ln q + (L_k - N_k(\vec{d}, \mathcal{N})) \ln(1-q)) \\
&= \sum_{\mathcal{N} \in P_{CP}} (N_k(\vec{d}, \mathcal{N}) \ln \frac{q}{1-q} + L_k \ln(1-q))
\end{aligned}
$$

Therefore, $MLE_{\pi_{k,q}}(P_{CP}) = \arg\max_{\vec{d}} \pi(P_{CP}|\vec{d}) = \arg\max_{\vec{d}} \sum_{\mathcal{N} \in P_{CP}} (N_k(\vec{d}, \mathcal{N}) \ln \frac{q}{1-q} + L_k \ln(1-q)) = \arg\max_{\vec{d}} N_k(\vec{d}, P_{CP})$. □

Proof of Theorem 2: It is easy to check that any $\pi_{p,q}$ satisfies strong decomposability and inter-issue neutrality. We next prove that any noise model π that satisfies strong decomposability and inter-neutrality must be a distance-based threshold noise model.

By Theorem 1, π is a distance-based model, denoted by $\pi_{\vec{q}}$. From strong decomposability, we have that for any $k \le p-1$, $\pi^{0_1\ldots 0_{p-1}}_{0_1\ldots 0_p}(0_p \succ 1_p) = \pi^{1_1\ldots 1_k 0_{k+1}\ldots 0_{p-1}}_{0_1\ldots 0_p}(0_p \succ 1_p)$, which means that $q_0 = q_k$. Hence, we have that π is a distance-based threshold noise model. □

Proof of Proposition 2: First we prove that the output of Algorithm 1 is $MLE_{\pi_{p,q}}(P_{CP})$. For any $\vec{d} \in \mathcal{X}$, $N_p(\vec{d}, P_{CP}) = \sum_{i \le p} \#\{\vec{a}_{-1} \in D_{-1} : (d_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (\bar{d}_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\}$. We note that $d_i \in W_i$ if and only if $\#\{\vec{a}_{-1} \in D_{-1} : (d_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (\bar{d}_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\} \succ \#\{\vec{a}_{-1} \in D_{-1} : (\bar{d}_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (d_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\}$. Therefore, $\vec{d} \in MLE_{\pi_{p,q}}(P_{CP})$ if and only if for all $i \le p$, we have that $d_i \in W_i$.

Next we prove that the algorithm runs in polynomial time. We note that in step 1b, the complexity of computing l is $O(2^{|Par_G(\mathbf{x}_i)|})$, and $CPT(\mathbf{x}_i)$ of the CP-net $\mathcal{N}$ has exactly $2^{|Par_G(\mathbf{x}_i)|}$ entries, which means that the complexity of computing l is in polynomial of the size of $CPT(\mathbf{x}_i)$ of the input. Therefore, Algorithm 1 is a polynomial-time algorithm. □

Proof of Theorem 3: By Proposition 1, the decision problem of finding a winner under $MLE_{\pi_{1,q}}$ is the following: for any profile P that consists of n CP-nets, and any $T \le pn$, we are asked whether or not there exists $\vec{d} \in \mathcal{X}$ such that $N_1(\vec{d}, P) \ge T$.

We prove the NP-hardness by reduction from the decision problem of `MAX-2-SAT`. The inputs of an instance of the decision problem of `MAX-2-SAT` are:

- A set of t atomic propositions $x_1, \ldots, x_t$.
- $T \le t$.

- A formula $F = c_1 \wedge \ldots \wedge c_m$ represented in *conjunctive normal form*, in which for any $i \leq m$, $c_i = l_{i_1} \vee l_{i_2}$, and there exists $j_1, j_2 \leq t$ such that l_{i_1} is x_{j_1} or $\neg x_{j_1}$, and l_{i_2} is x_{j_2} or $\neg x_{j_2}$.

We are asked whether or not there exists a valuation $\vec{x}$ for the atomic propositions $x_1, \ldots, x_t$ such that at least T clauses are satisfied under $\vec{x}$.

Given any instance of MAX-2-SAT, we construct a decision problem of computing a winner under $MLE_{\pi_{1,q}}$ as follows.

- Let $\mathcal{X}$ be composed of t issues $\mathbf{x}_1, \ldots, x_t$.
- Let $T' = 16T - 12m$.
- For any $i \leq m$, we let v_{i_1} be the valuation of x_{i_1} under which l_{i_1} is true; let v_{i_2} be the valuation of x_{i_2} under which l_{i_2} is true. For any $j \leq t$, we let 0_j corresponds to $\mathbf{x}_j$ being false, and 1_j corresponds to $\mathbf{x}_j$ being true. Then, any valuation of the atomic propositions is uniquely identified by an alternative. We next define six CP-net as follows:
  - $\mathcal{N}_{i,1}$ and $\mathcal{N}'_{i,1}$: the DAG of $\mathcal{N}_{i,1}$ has only one directed edge $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$. In $\mathcal{N}_{i,1}$, $v_{i_1} \succ \bar{v}_{i_1}$, $v_{i_1} : v_{i_2} \succ \bar{v}_{i_2}$, $\bar{v}_{i_1} : v_{i_2} \succ \bar{v}_{i_2}$, and for any $j \neq i_1$ and $j \neq i_2$, we have that $0_j \succ 1_j$. $\mathcal{N}'_{i,1}$ is the same as $\mathcal{N}_{i,1}$, except that for any $j \neq i_1$ and $j \neq i_2$, we have that $1_j \succ 0_j$.
  - $\mathcal{N}_{i,2}$ and $\mathcal{N}'_{i,2}$: the DAG of $\mathcal{N}_{i,2}$ has only one directed edge $(\mathbf{x}_{i_1}, \mathbf{x}_{i_2})$. In $\mathcal{N}_{i,2}$, $v_{i_1} \succ \bar{v}_{i_1}$, $v_{i_1} : \bar{v}_{i_2} \succ v_{i_2}$, $\bar{v}_{i_1} : v_{i_2} \succ \bar{v}_{i_2}$, and for any $j \neq i_1$ and $j \neq i_2$, we have that $0_j \succ 1_j$. $\mathcal{N}'_{i,2}$ is the same as $\mathcal{N}_{i,2}$, except that for any $j \neq i_1$ and $j \neq i_2$, we have that $1_j \succ 0_j$.
  - $\mathcal{N}_{i,3}$ and $\mathcal{N}'_{i,3}$: the DAG of $\mathcal{N}_{i,3}$ has only one directed edge $(\mathbf{x}_{i_2}, \mathbf{x}_{i_1})$. In $\mathcal{N}_{i,1}$, $v_{i_2} \succ \bar{v}_{i_2}$, $v_{i_2} : \bar{v}_{i_1} \succ v_{i_1}$, $\bar{v}_{i_2} : v_{i_1} \succ \bar{v}_{i_1}$, and for any $j \neq i_1$ and $j \neq i_2$, we have that $0_j \succ 1_j$. $\mathcal{N}'_{i,3}$ is the same as $\mathcal{N}_{i,3}$, except that for any $j \neq i_1$ and $j \neq i_2$, we have that $1_j \succ 0_j$.

  Let $\vec{\mathcal{N}}_i = (\mathcal{N}_{i,1}, \mathcal{N}'_{i,1}, \mathcal{N}_{i,2}, \mathcal{N}'_{i,2}, \mathcal{N}_{i,3}, \mathcal{N}'_{i,3})$. We let the profile of CP-net be $P_{CP} = (\vec{\mathcal{N}}_1, \ldots, \vec{\mathcal{N}}_m)$.

Let P be a profile of CP-nets. The aggregated majority graph of P is a weighted directed graph in which the vertices are $\mathcal{X}$, and the edges are defined in the following way: for any neighboring vertices $\vec{d}_1, \vec{d}_2$ in the hypercube, if $D_P(\vec{d}_1, \vec{d}_2) > 0$, then there is a directed edge from $\vec{d}_1$ to $\vec{d}_2$ with weight $D_P(\vec{d}_1, \vec{d}_2)$; if $D_P(\vec{d}_2, \vec{d}_1) > 0$, the direction and weight of the edge is defined similarly; if $D_P(\vec{d}_1, \vec{d}_2) = 0$, then there is no edge between $\vec{d}_1$ and $\vec{d}_2$). The aggregated majority graph of $\mathcal{N}_i$ is illustrated in Figure 6 (the vertices that are not connected to any other vertex is not drawn in the figure). We make the following claim on the number of consistent edges between an alternative $\vec{d}$ and $\vec{\mathcal{N}}_i$.

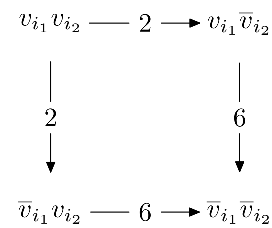


Figure 6: The aggregated majority graph of $\vec{\mathcal{N}}_i$.

Claim 1 *For any $\vec{d} \in \mathcal{X}$ and any $i \leq m$,* $N_1(\vec{d}, \vec{\mathcal{N}}_i) = \begin{cases} 4 & \text{if } \vec{d}_{i_1} = v_{i_1} \text{ or } d_{i_2} = v_{i_2} \\ -12 & \text{if } \vec{d}_{i_1} = \bar{v}_{i_1} \text{ and } d_{i_2} = \bar{v}_{i_2} \end{cases}$

Claim 1 states that the number of consistent edges between $\vec{d}$ and $\vec{\mathcal{N}}_i$ within distance 1 is 4 if the clause l_i is true under the valuation represented by $\vec{d}$; otherwise it is -12. For any $\vec{d} \in \mathcal{X}$, we let $T_{\vec{d}}$ denote the number of clauses in $c_1, \ldots, c_m$ that are true under $\vec{d}$. Then, we have that $N_1(\vec{d}, P_{CP}) = 4T_{\vec{d}} - 12(m - T_{\vec{d}}) = 16T_{\vec{d}} - 12m$. It follows from Proposition 1 that for any $q > \frac{1}{2}$, $MLE_{\pi_{1,q}}(P_{CP}) = \arg\max_{\vec{d}} N_1(\vec{d}, P_{CP}) = \arg\max_{\vec{d}} T_{\vec{d}}$. Therefore, a winner of P_{CP} under $MLE_{\pi_{1,q}}$ corresponds to a valuation under which the number of satisfied clauses is maximized; and any valuation that maximizes the number of satisfied clauses corresponds to a winner of P_{CP} under $MLE_{\pi_{1,q}}$. We note that the size of P_{CP} is $O(mt)$. It follows that computing a winner under $MLE_{\pi_{1,q}}$ is NP-hard.

Clearly the problem is in NP. Therefore, it is NP-complete to compute a winner under $MLE_{\pi_{1,q}}$. $\Box$

Proof of Theorem 4: First we prove the "if" part. Let c be an issue-by-issue voting correspondence that is composed of $c_1, \ldots, c_p$, in which for any $i \leq p$, c_i is an MLEWIV over D_i of the noise model $Pr(V^i|d_i)$, where $V^i \in L(D_i)$ and $d_i \in D_i$. Let π be a noise model over $\mathcal{X}$ defined as follows: for any $i \leq p$, any $\vec{d} \in \mathcal{X}$, any $\vec{a}_{-i} \in D_{-i}$ and any $V^i \in L(D_i)$, we have that $\pi_{\vec{d}}^{\vec{a}}(V_i) = Pr(V^i|d_i)$. We next prove that for any separable profile P, we must have that $MLE_\pi(P) = c(P)$.

$$MLE_\pi(P) = \arg\max_{\vec{d}} \prod_{i \leq p, \vec{a}_{-i} \in D_{-i}} \prod_{j=1}^{n} \pi_{\vec{d}}^{\vec{a}_{-i}}(V_j)$$
$$= \arg\max_{\vec{d}} \prod_{i \leq p} \prod_{j=1}^{n} Pr((V_i|_{\mathbf{x}_i})|d_i)^{|D_{-i}|}$$

Therefore, $\vec{b} \in MLE_\pi(P)$ if and only if for any $i \leq p$, we have $b_i \in \arg\max_{d_i} \prod_{j=1}^{n} Pr((V_i|_{\mathbf{x}_i})|b_i)$. We note that for any $\vec{d'} \in r(P)$, we must have that

$d'_i = \arg\max_{d_i} \prod_{j=1}^{n} Pr((V_i|_{\mathbf{x}_i})|d_i)$. Therefore, $\vec{d'} \in MLE_\pi(P)$.

Next, we prove the "only if" part. For any $MLE_\pi \in SD(\mathcal{X})$, we define an issue-by-issue voting rule as follows: for any $i \leq p$, let c_i be the MLEWIV that corresponds to the noise model in which for any $d_i \in D_i$, we have that $Pr(V^i|d_i) = \pi_{\vec{d}}^{\vec{a}_{-i}}(V^i)$. Similar to the proof for the "if" part, we have that c and MLE_π are equivalent over the domain of separable profiles. □

Proof of Theorem 5: First we prove the "if" part. Let c be an issue-by-issue voting correspondence in which the issue-wise correspondence over D_i is c_{s_i}, which has scoring function s_i. Let $\pi_{d_i}^{\vec{a}_{-i}}$ denote $\pi_{\vec{d}}^{\vec{a}_{-i}}$, where the ith component of $\vec{d}$ is d_i. Because r is strongly decomposable, $\pi_{d_i}^{\vec{a}_{-i}}$ is well-defined. For any $i \leq p$, we claim that there exists a set of probability distributions $\pi_{\vec{d}}^{\vec{a}_{-i}}$, $\vec{d} \in \mathcal{X}, \vec{a}_{-i} \in D_{-i}$ over $L(D_i)$ such that for any $d_i \in D_i$, $d_i \in \arg\max_{b_i \in D_i} \prod_{j=1}^{n} \prod_{\vec{a}_{-i} \in D_{-i}} \pi_{b_i}^{\vec{a}_{-i}}(V_j|_{\mathbf{x}_i})$ if and only if $d_i \in c_{s_i}(P|_{\mathbf{x}_i})$.

We note that for any scoring function s and any constant t, the ranking scoring rule that corresponds to s is equivalent to the ranking scoring rule that corresponds to $s + t$. Therefore, without loss of generality we let $s_i(V^i, d_i) < 0$ for any $i \leq p$, any $V^i \in L(D_i)$, and any $d_i \in D_i$. Let $K_i = |D_i|$, $L(D_i) = \{l_1, \ldots, l_{K_i!}\}$.

Claim 2 *There exist $k_i, t_i \in \mathbb{R}$ with $k_i > 0$, such that for any $V^i \in L(D_i)$ and any $d_i \in D_i$, we have that $\ln(\prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V^i)) = k_i s_i(V^i, d_i) + t_i$.*

Proof of Claim 2: We let k_i be a real number such that for any $d_i \in D_i$, we have that $\sum_{j=1}^{K_i!} (\exp(s_i(l_j, d_i)))^{k_i} < 1$; let $\hat{p}_{d_i}^j = \exp(s_i(l_j, d_i))$. For any $d_i \in D_i$, any $1 \leq \alpha < \dfrac{K_i!}{K_i! - 1}$, we let

$$f_{d_i}(\alpha) = \ln((1 - \sum_{j=1}^{K_i!-1} \frac{\hat{p}_{d_i}^j}{\alpha})(1 - (K_i - 1)\frac{\alpha}{K_i!}))$$

Because $\sum_{j=1}^{K_i!} \hat{p}_{d_i}^j < 1$, we have that $\ln(1 - \sum_{j=1}^{K_i!-1} \hat{p}_{d_i}^j) > \ln \hat{p}_{d_i}^{K_i!} = k_i s_i(l_{K_i!}, d_i)$. Therefore, $f_{d_i}(1) \geq k_i s_i(l_{K_i!}, d_i) - \ln(K_i!)$. We note that $\lim_{\alpha \to \frac{K_i!}{K_i!-1}} f_{d_i}(\alpha) = -\infty$. It follows that there exists $1 \leq \alpha_{d_i} \leq \dfrac{K_i!}{K_i! - 1}$ such that $f_{d_i}(\alpha_{d_i}) = k_i s_i(l_{K_i!}, d_i) - \ln(K_i!)$.

For any $i \leq p$, any $d_i \in D_i$, we let $\vec{a}'_{-i}, \vec{a}^*_{-i} \in D_{-i}$ such that $\vec{a}'_{-i} \neq \vec{a}^*_{-i}$. We define $\pi_{d_i}^{\vec{d}_{-i}}$ as follows.

- for any $j \leq K_i! - 1$, $\pi_{d_i}^{\vec{a}'_{-i}}(l_j) = \dfrac{1}{\alpha_{d_i}}(\exp(s_i(l_j, d_i)))^{k_i}$, $\pi_{d_i}^{\vec{a}^*_{-i}}(l_j) = \dfrac{\alpha_{d_i}}{K_i!}$.
- for any $j \leq K_i!$, any $\vec{d}_{-i} \in D_{-i}$ such that $\vec{d}_{-i} \neq \vec{a}'_{-i}$ and $\vec{d}_{-i} \neq \vec{a}^*_{-i}$, we have that $\pi_{d_i}^{\vec{d}_{-i}}(l_j) = \frac{1}{K_i!}$.

For any $\vec{d}_i \in D_i$ and any $j \leq K_i! - 1$, we have that

$$\begin{aligned}
&\ln(\prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(l_j)) \\
=&\ln(\pi_{d_i}^{\vec{a}'_{-i}}(l_j) \cdot \pi_{d_i}^{\vec{a}^*_{-i}}(l_j)) + (|D_{-i}| - 2)\ln(\frac{1}{K_i!}) \\
=&\ln(\frac{1}{\alpha_{d_i}}(\exp(s_i(l_j, d_i)))^{k_i} \cdot \frac{\alpha_{d_i}}{K_i!}) - (|D_{-i}| - 2)\ln(K_i!) \\
=&k_i s_i(l_j, d_i) - (|D_{-i}| - 1)\ln(K_i!)
\end{aligned}$$

For $j = K_i!$, we have the following calculation.

$$\begin{aligned}
&\ln(\prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(l_{K_i!})) \\
=&\ln(\pi_{d_i}^{\vec{a}'_{-i}}(l_{K_i!}) \cdot \pi_{d_i}^{\vec{a}^*_{-i}}(l_{K_i!})) + (|D_{-i}| - 2)\ln(\frac{1}{K_i!}) \\
=&f_{d_i}(\alpha_i) - (|D_{-i}| - 2)\ln(K_i!) \\
=&k_i s_i(l_{K_i!}, d_i) - (|D_{-i}| - 1)\ln(K_i!)
\end{aligned}$$

Therefore, let $t_i = -(|D_{-i}| - 1)\ln(K_i!)$. It follows that for any $V^i \in L(D_i)$, and any $d_i \in D_i$, we must have that $\ln(\prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V^i)) = k_i s_i(V^i, d_i) + t_i$. □

Next, we show that for any separable profile P, $c(P) = MLE_\pi(P)$. Similar to in the proof of Theorem 4, it suffices to prove that for any $i \leq p$, $\arg\max_{d_i \in D_i} \prod_{j \leq n} \prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{\mathbf{x}_i}) = c_{s_i}(P|_{\mathbf{x}_i})$.

$$\begin{aligned}
&\arg\max_{d_i \in D_i} \prod_{j \leq n} \prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{\mathbf{x}_i}) \\
=&\arg\max_{d_i \in D_i} \ln(\prod_{j \leq n} \prod_{\vec{d}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{\mathbf{x}_i})) \\
=&\arg\max_{d_i \in D_i} \sum_{j \leq n} \sum_{\vec{d}_{-i} \in D_{-i}} \ln(\pi_{d_i}^{\vec{d}_{-i}}(V_j|_{\mathbf{x}_i})) \\
=&\arg\max_{d_i \in D_i} \sum_{j \leq n} \sum_{\vec{d}_{-i} \in D_{-i}} (k_i s_i(V_j|_{\mathbf{x}_i}, d_i) + t_i) \\
=&\arg\max_{d_i \in D_i} \sum_{j \leq n} \sum_{\vec{d}_{-i} \in D_{-i}} s_i(V_j|_{\mathbf{x}_i}, d_i) \\
=&c_{s_i}(P|_{\mathbf{x}_i})
\end{aligned}$$

Next, we prove the "only if" part. Let π be a weakly decomposable noise model. For any $i \leq p$, any $d_i \in D_i$, and any $V^i \in L(D_i)$, we let $s_i(V^i, d_i) = \ln \prod_{\vec{a}_{-i} \in D_{-i}} \pi_{d_i}^{\vec{a}_{-i}}(V^i)$. Then, we have that

d_i maximizes $s_i(P|_{\mathbf{x}_i}, d_i)$ if and only if d_i maximizes $\prod_{\mathcal{N}\in P}\prod_{\vec{a}_{-i}\in D_{-i}}\pi_{d_i}^{\vec{a}_{-i}}(\mathcal{N}|_{\mathbf{x}_i})$, which means that $c(P) = MLE_\pi(P)$.

(End of the proof of Theorem 5). □

Proof of Theorem 6: For the sake of contradiction, we let $Seq(c_1,\ldots,c_p)$ be a sequential voting correspondence and MLE_π be an MLE model equivalent to it. A voting correspondence c satisfies *consistency*, if for any profiles P_1, P_2, if $c(P_1) = c(P_2)$, then $c(P_1\cup P_2) = c(P_1)$; c satisfies anonymity, if it is indifferent with the name of the voters. Because MLE_π satisfies consistency and anonymity, we have the following claim.

Claim 3 *For any $i \le p$, c_i satisfies consistency, anonymity (see [9]) and unanimity.*

For any $\vec{d}\in\mathcal{X}$, any $\mathcal{O}$-legal CP-net $\mathcal{N}$, we let

$$\pi_{\vec{d}}^{\mathbf{x}_1}(\mathcal{N}) = \prod_{\vec{a}_{-1}\in D_{-1}}\pi_{\vec{d}}^{\vec{a}_{-1}}(\mathcal{N}|_{\mathbf{x}_1})$$

$$\pi_{\vec{d}}^{\mathbf{x}_{-1}}(\mathcal{N}) = \prod_{2\le i\le p, \vec{a}_{-i}\in D_{-i}}\pi_{\vec{d}}^{\vec{a}_{-i}}(\mathcal{N}|_{\mathbf{x}_i:a_1\ldots a_{i-1}})$$

Let $\mathcal{N}_1, \mathcal{N}_2$ be CP-nets. We note that if $\mathcal{N}_1|_{\mathbf{x}_1} = \mathcal{N}_2|_{\mathbf{x}_1}$, then $\pi_{\vec{d}}^{\mathbf{x}_1}(\mathcal{N}_1) = \pi_{\vec{d}}^{\mathbf{x}_1}(\mathcal{N}_2)$; if for any $d_1 \in D_1$, $\mathcal{N}_1|_{\mathbf{x}_{-1}:d_1} = \mathcal{N}_2|_{\mathbf{x}_{-1}:d_1}$, then we must have that $\pi_{\vec{d}}^{\mathbf{x}_{-1}}(\mathcal{N}_1) = \pi_{\vec{d}}^{\mathbf{x}_{-1}}(\mathcal{N}_2)$, where $\mathcal{N}_1|_{\mathbf{x}_{-1}:d_1}$ is the sub-CP-net of $\mathcal{N}_1$ given $\mathbf{x}_1 = d_1$. For any $\mathcal{O}$-legal vote V that extends a CP-net $\mathcal{N}$, we write $\pi_{\vec{d}}^{\mathbf{x}_1}(V) = \pi_{\vec{d}}^{\mathbf{x}_1}(\mathcal{N})$ and $\pi_{\vec{d}}^{\mathbf{x}_{-1}}(V) = \pi_{\vec{d}}^{\mathbf{x}_{-1}}(\mathcal{N})$; for any $\mathcal{O}$-legal profile P, we write $\pi_{\vec{d}}^{\mathbf{x}_1}(P) = \prod_{V\in P}\pi_{\vec{d}}^{\mathbf{x}_1}(V)$ and $\pi_{\vec{d}}^{\mathbf{x}_{-1}}(P) = \prod_{V\in P}\pi_{\vec{d}}^{\mathbf{x}_{-1}}(V)$. It follows that for any $\mathcal{O}$-legal profile P, we have that

$$MLE_\pi(P) = \arg\max_{\vec{d}\in\mathcal{X}}[\pi_{\vec{d}}^{\mathbf{x}_1}(P)\cdot\pi_{\vec{d}}^{\mathbf{x}_{-1}}(P)]$$

For any $V_1^1, V_2^1 \in L(D_1)$ with $top(V_1^1) \ne top(V_2^1)$, and any $n\in\mathbb{N}$, we let $P_{1,n}^1$ be the profile that is composed of n copies of V_1^1; let $P_{2,n}^1$ be the profile that is composed of n copies of V_2^1. Because c_1 satisfies unanimity, we must have that $c_1(P_{1,n}^1) = \{top(V_1^1)\}$ and $c_1(P_{2,n}^1) = \{top(V_2^1)\}$. For any $j \le n$, we let $Q_{j,n}$ be the profile in which the preferences of the first j voters are V_1^1, and the preferences of the remaining $n-j$ voters are V_2^1. We have that $Q_{1,n} = P_{1,n}^1$ and $Q_{n,n} = P_{2,n}^1$. Therefore, there exists $j \le n-1$ and $b_1 \in D_1$ with $b_1 \ne top(V_1^1)$, such that $top(V_1^1) \in c_1(Q_{j,n})$ and $b_1 \in c_1(Q_{j+1,n})$. For any $n\in\mathbb{N}$, we let C_n denote the set of pairs (a_1, b_1) such that

- $a_1, b_1 \in D_1$, $a_1 \ne b_1$.
- There exists two profiles W_1^1, W_2^1 over D_1 such that $a_1 \in c_1(W_1^1)$, $b_1 \in c_1(W_2^1)$, and W_1^1 differs from W_2^1 only on one vote.

That is, C_n is composed of the pairs (a_1, b_1) such that there exists a profile Q over D_1 that consists of n votes, $a_1 \in c_1(Q)$, and by changing one vote of Q, there is another alternative b_1 who is one of the winners. We note that for any $n\in\mathbb{N}$, $(a_1,b_1)\in C_n$ if and only if $(b_1,a_1)\in C_n$. It follows that for any $n\in\mathbb{N}$, $C_n \ne \emptyset$. Because $|D_1| < \infty$, there exists $(a_1,b_1)\in(D_1)^2$ such that for any $k\in\mathbb{N}$, there exists $n\ge k$ such that $(a_1,b_1)\in C_n$.

Claim 4 *For any $\vec{a}_{-1}, \vec{b}_{-1} \in D_{-1}$, and any pair of CP-nets $\mathcal{N}', \mathcal{N}^*$, we must have that $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}')} = \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}$, where $\vec{a} = (a_1, \vec{a}_{-1})$, $\vec{b} = (b_1, \vec{b}_{-1})$.*

Proof of Claim 4: Suppose for the sake of contradiction there exist $\vec{a}_{-1}, \vec{b}_{-1}$, and $\mathcal{N}', \mathcal{N}^*$ so that $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}')} \ne \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}$. Without loss of generality we let $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}')} > \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}$. We next claim that there exits a natural number k such that for any $i\le p$ and any profile P^i composed of k votes, if at least $k-1$ votes in P^i rank the same alternative d_i in the top position, then $c_i(P^i) = \{d_i\}$.

Claim 5 *There exists $k\in\mathbb{N}$ such that for any $i\le p$, any $d_i \in D_i$, and any profile $P^i = (V_1^i,\ldots,V_k^i)$ with $d_i = top(V_1^i) = \ldots = top(V_{k-1}^i)$, we have that $c_i(P) = \{d_i\}$.*

Proof of Claim 5: Let $U = \max_{\vec{d}_1,\vec{d}_2,\mathcal{N}}\frac{Pr(\mathcal{N}|\vec{d}_1)}{Pr(\mathcal{N}|\vec{d}_2)}$. Let $u = \min_{\vec{d}_1\ne\vec{d}_2,\mathcal{N}:top(\mathcal{N})=\vec{d}_1}\frac{Pr(\mathcal{N}|\vec{d}_1)}{Pr(\mathcal{N}|\vec{d}_2)}$. Because $MLE_\pi(\mathcal{N})$ satisfies unanimity, for any $\vec{d}_1$ and $\mathcal{N}$ such that $top(\mathcal{N}) = \vec{d}_1$, we must have that $MLE_\pi(\mathcal{N}) = \{\vec{d}_1\}$, which means that $u > 1$. Let k be a natural number such that $u^{k-1} > U$. We arbitrarily choose $\vec{d}_{-i}\in D_{-i}$, and let $\vec{d} = (d_i, \vec{d}_{-i})$. We define k CP-nets $\mathcal{N}_1,\ldots,\mathcal{N}_k$ as follows.

- For any $j\le k$, $top(\mathcal{N}_j) = (\vec{d}_{-i}, top(V^i))$.
- For any $j\le k$, $\mathcal{N}_j|_{\mathbf{x}_i:d_1,\ldots,d_{i-1}} = V^i$.
- Other conditional preferences are defined arbitrarily.

Because $Seq(c_1,\ldots,c_p)$ satisfies unanimity, we have that $Seq(c_1,\ldots,c_p)(\mathcal{N}_1,\ldots,\mathcal{N}_{k-1}) = \{\vec{d}\}$. Therefore, for any $\vec{d}'\in\mathcal{X}$ and any CP-net $\mathcal{N}$, we have the following calculation:

$$\frac{Pr((\mathcal{N}_1,\ldots,\mathcal{N}_k)|\vec{d})}{Pr((\mathcal{N}_1,\ldots,\mathcal{N}_k)|\vec{d}')} = \frac{\prod_{j=1}^{k-1}Pr(\mathcal{N}_j|\vec{d})}{\prod_{j=1}^{k-1}Pr(\mathcal{N}_j|\vec{d}')}\cdot\frac{Pr(\mathcal{N}_k|\vec{d})}{Pr(\mathcal{N}_k|\vec{d}')}$$
$$\ge u^{(k-1)}\frac{1}{U} > 1$$

Therefore $c_i(V^1, \ldots, V^k) = \{d_i\}$.

(End of proof of Claim 5.) □

Let $\mathcal{N}_{\vec{a}}$ be a CP-net such that $top(\mathcal{N}_{\vec{a}}) = \vec{a}$ and $top(\mathcal{N}|_{\mathbf{x}_{-1}:b_1}) = \vec{b}_{-1}$. That is, $\mathcal{N}_{\vec{a}}$ is a CP-net in which $\vec{a}$ is ranked in the top position, and given $\mathbf{x}_1 = b_1, \vec{b}_{-1}$ is ranked in the top position. Next, we show that for any CP-net $\mathcal{N}$, $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} = \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}$. Suppose for the sake of contradiction, there exists $\mathcal{N}$ such that $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} \neq \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}$. We next show contradiction in the case $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} > \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}$. Let $U_{\mathbf{x}_1} = \max_{\vec{d}_1, \vec{d}_2, \mathcal{N}} \frac{\pi_{\vec{d}_1}^{\mathbf{x}_1}(\mathcal{N})}{\pi_{\vec{d}_2}^{\mathbf{x}_1}(\mathcal{N})}$. Let K be a natural number such that $(\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} / \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})})^K > U_{\mathbf{x}_1}^2$. Let $n \in \mathbb{N}$ be such that $n > kK$ and $(a_1, b_1) \in C_n$. It follows that there exist $(V_1^1, \ldots, V_n^1)$ and W_1^1 such that $a_1 \in c_1(V_1^1, \ldots, V_n^1)$ and $b_1 \in c_1(W_1^1, V_2^1, \ldots, V_n^1)$. We define $2n+1$ CP-nets $\mathcal{N}_1', \mathcal{N}_1, \mathcal{N}_2, \ldots, \mathcal{N}_n, \hat{\mathcal{N}}_1, \hat{\mathcal{N}}_2, \ldots, \hat{\mathcal{N}}_n$ as follows.

- For any $j \leq n, \mathcal{N}_j|_{\mathbf{x}_1} = \hat{\mathcal{N}}_j|_{\mathbf{x}_1} = V_j^1; \mathcal{N}_1'|_{\mathbf{x}_1} = W_1^1$.
- For any $j_1 \leq K$, $1 \leq j_2 \leq k-1$, and any $d_1 \in D_1$, $\mathcal{N}_{(j_1-1)k+j_2}|_{\mathbf{x}_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{\mathbf{x}_{-1}:d_1}$ and $\mathcal{N}_{j_1 k}|_{\mathbf{x}_{-1}:d_1} = \mathcal{N}|_{\mathbf{x}_{-1}:d_1}$; for any $j \leq n$ and any $d_1 \in D_1, \mathcal{N}_j|_{\mathbf{x}_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{\mathbf{x}_{-1}:d_1}$.
- For any $kK+1 \leq j \leq n, \mathcal{N}_j = \hat{\mathcal{N}}_j = \mathcal{N}_{\vec{a}}$.
- For any $d_1 \in D_1, \mathcal{N}_1'|_{\mathbf{x}_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{\mathbf{x}_{-1}:d_1}$.

For any $j \leq n$, we let V_j ($\hat{V}_j$) be an arbitrary linear order that extends $\mathcal{N}_j$ ($\hat{\mathcal{N}}_j$); let V_1' be an arbitrary linear order that extends $\mathcal{N}_1'$; let $P = (V_1, \ldots, V_n)$, $P' = (V_1', V_2, \ldots, V_n)$, $\hat{P} = (\hat{V}_1, \ldots, \hat{V}_n)$, $\hat{P}' = (\hat{V}_1', \hat{V}_2, \ldots, \hat{V}_n)$. We make the following observations.

- $a_1 \in c_1(P|_{\mathbf{x}_1}), a_1 \in c_1(\hat{P}|_{\mathbf{x}_1}), b_1 \in c_1(P'|_{\mathbf{x}_1}), b_1 \in c_1(\hat{P}'|_{\mathbf{x}_1})$.
- For any $1 \leq i \leq p-1$, $P|_{\mathbf{x}_i:a_1\ldots a_{i-1}} = K((k-1)\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:a_1\ldots a_{i-1}} \cup \mathcal{N}'|_{\mathbf{x}_i:a_1\ldots a_{i-1}}) \cup (n-kK)\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:a_1\ldots a_{i-1}}$. From Claim 5 we have that $c_i((k-1)\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:a_1\ldots a_{i-1}} \cup \mathcal{N}'|_{\mathbf{x}_i:a_1\ldots a_{i-1}}) = \{a_i\}$. Because c_i satisfies unanimity and consistency, and for any $i \leq p$, $top(\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:a_1\ldots a_{i-1}}) = a_i$, we have that for any $i \leq p$, $c_i(P|_{\mathbf{x}_i:a_1\ldots a_{i-1}}) = \{a_i\}$. Similarly for any $i \leq p$, $c_i(\hat{P}|_{\mathbf{x}_i:a_1\ldots a_{i-1}}) = \{a_i\}$.
- For any $1 \leq i \leq p-1$, $P|_{\mathbf{x}_i:b_1\ldots b_{i-1}} = K((k-1)\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:b_1\ldots b_{i-1}} \cup \mathcal{N}'|_{\mathbf{x}_i:b_1\ldots b_{i-1}}) \cup (n-kK)\mathcal{N}_{\vec{a}}|_{\mathbf{x}_i:b_1\ldots b_{i-1}}$. Similarly, we have that for any $1 \leq i \leq p$, $c_i(P'|_{\mathbf{x}_i:b_1\ldots b_{i-1}}) = c_i(\hat{P}'|_{\mathbf{x}_i:b_1\ldots b_{i-1}}) = \{b_i\}$.

Therefore, we have that $\vec{a} \in Seq(c_1, \ldots, c_p)(P), \vec{a} \in Seq(c_1, \ldots, c_p)(\hat{P})$, and $\vec{b} \in Seq(c_1, \ldots, c_p)(P'), \vec{b} \in Seq(c_1, \ldots, c_p)(\hat{P}')$. That is, $\frac{Pr(P'|\vec{b})}{Pr(P'|\vec{a})} \geq 1$, $\frac{Pr(\hat{P}'|\vec{b})}{Pr(\hat{P}'|\vec{a})} \geq 1$. We note that P and P' differ only on the first vote. Therefore, we have the following calculation.

$$\begin{aligned}
1 \leq & \frac{Pr(P'|\vec{b})}{Pr(P'|\vec{a})} \\
= & \frac{\pi_{\vec{b}}^{\mathbf{x}_1}(V_1') \cdot \pi_{\vec{b}}^{\mathbf{x}_{-1}}(V_1') \prod_{2\leq j\leq n}(\pi_{\vec{b}}^{\mathbf{x}_1}(V_j) \cdot \pi_{\vec{b}}^{\mathbf{x}_{-1}}(V_j))}{\pi_{\vec{a}}^{\mathbf{x}_1}(V_1') \cdot \pi_{\vec{a}}^{\mathbf{x}_{-1}}(V_1') \prod_{2\leq j\leq n}(\pi_{\vec{a}}^{\mathbf{x}_1}(V_j) \cdot \pi_{\vec{a}}^{\mathbf{x}_{-1}}(V_j))} \\
= & \frac{\pi_{\vec{b}}^{\mathbf{x}_1}(V_1')}{\pi_{\vec{a}}^{\mathbf{x}_1}(V_1')} \cdot \frac{\pi_{\vec{a}}^{\mathbf{x}_1}(V_1)}{\pi_{\vec{b}}^{\mathbf{x}_1}(V_1)} \cdot \frac{Pr(P|\vec{b})}{Pr(P|\vec{a})} \\
\leq & U_{\mathbf{x}_1}^2 \frac{Pr(P|\vec{b})}{Pr(P|\vec{a})}
\end{aligned}$$

Therefore, $\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})} \leq U_{\mathbf{x}_1}^2$. We note that P and P' differ on K votes.

$$\begin{aligned}
& (\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})})/(\frac{Pr(\hat{P}|\vec{a})}{Pr(\hat{P}|\vec{b})}) \\
= & (\prod_{j=1}^{K} \frac{\pi_{\vec{a}}^{\mathbf{x}_1}(V_{jk}) \cdot \pi_{\vec{a}}^{\mathbf{x}_{-1}}(V_{jk})}{\pi_{\vec{b}}^{\mathbf{x}_1}(V_{jk}) \cdot \pi_{\vec{b}}^{\mathbf{x}_{-1}}(V_{jk})})/(\prod_{j=1}^{K} \frac{\pi_{\vec{a}}^{\mathbf{x}_1}(\hat{V}_{jk}) \cdot \pi_{\vec{a}}^{\mathbf{x}_{-1}}(\hat{V}_{jk})}{\pi_{\vec{b}}^{\mathbf{x}_1}(\hat{V}_{jk}) \cdot \pi_{\vec{b}}^{\mathbf{x}_{-1}}(\hat{V}_{jk})}) \\
= & (\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} / \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})})^K \\
> & U_{\mathbf{x}_1}^2
\end{aligned}$$

We note that $(\frac{Pr(\hat{P}|\vec{a})}{Pr(\hat{P}|\vec{b})}) \geq 1$. Therefore, $\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})} > U_{\mathbf{x}_1}^2$, which is a contradiction.

Similarly, for the case of $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} < \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}$ we still have a contradiction. Hence, $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N})} = \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})}$ for all $\mathcal{N}$, which means that for any $\mathcal{N}'$ and $\mathcal{N}^*$, we must have that $\frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}')} = \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{\mathbf{x}_{-1}}(\mathcal{N}^*)}$.

(End of proof of Claim 4.) □

By Claim 4, for any CP-net $\mathcal{N}$, any $\vec{b}_{-1}, \vec{b}'_{-1} \in D_{-1}$, we must have that $\frac{\pi_{(b_1,\vec{b}_{-1})}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{(b_1,\vec{b}_{-1})}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})} = \frac{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N})}{\pi_{\vec{a}}^{\mathbf{x}_{-1}}(\mathcal{N}_{\vec{a}})} =$

$\frac{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N})}{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}$, which means that $\frac{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N})}{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N})} = \frac{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_{\vec{a}})}{\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}$. Let $\mathcal{N}_1$ be a CP-net such that $top(\mathcal{N}_1) = (b_1, \vec{b}'_{-1})$, $\mathcal{N}_2$ be a CP-net such that $top(\mathcal{N}_2) = (b_1, \vec{b}_{-1})$ and $\mathcal{N}_1|_{\mathbf{x}_1} = \mathcal{N}_2|_{\mathbf{x}_1}$. Because $Seq(c_1, \ldots, c_p)$ satisfies unanimity, we have that $\frac{Pr(\mathcal{N}_1|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_1|(b_1,\vec{b}_{-1}))} > 1$ and $\frac{Pr(\mathcal{N}_2|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_2|(b_1,\vec{b}_{-1}))} < 1$. However, we have the following calculation.

$$\begin{aligned}
1 <& \frac{Pr(\mathcal{N}_1|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_1|(b_1,\vec{b}_{-1}))} \\
=& \frac{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_1)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_1)}{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_1)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_1)} \\
=& \frac{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_{\vec{a}})} \quad \text{(Because } \mathcal{N}_1|_{\mathbf{x}_1} = \mathcal{N}_2|_{\mathbf{x}_1}\text{)} \\
=& \frac{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)}{\pi^{\mathbf{x}_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)\cdot\pi^{\mathbf{x}_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)} \\
=& \frac{Pr(\mathcal{N}_2|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_2|(b_1,\vec{b}_{-1}))} \\
<& 1
\end{aligned}$$

Therefore, we have a contradiction. **(End of proof of Theorem 6.)** □

Proof of Theorem 7: Let r_i be the MLEWIV with the conditional probabilistic distribution $Pr_i(V^i|d_i)$, where $V^i \in L(D_i)$, $d_i \in D_i$. For any $i \le p$, we let $R^{i,n}_{max} = \max_{P_i,P'_i,d_i,d'_i}\{\frac{Pr_i(P|d_i)}{Pr_i(P'|d'_i)}\}$, where $d_i, d'_i \in D_i$, P_i and P'_i are profiles with the same number (but no more than n) of linear orders over D_i. We let $R^{i,n}_{min} = 1$ if r_i is the trivial correspondence that always output the whole domain; and $R^{i,n}_{min} = \min_{P_i,\vec{d}_i,\vec{d}'_i}\{\frac{Pr_i(P_i|d_i)}{Pr_i(P_i|d'_i)} : \frac{Pr_i(P_i|d_i)}{Pr_i(P_i|d'_i)} > 1\}$, where $d_i, d'_i \in D_i$, and P_i is a profile of no more than n linear orders over D_i. We note that for any $i \le p$, any $n \in \mathbb{N}$, we have that $R^{i,n}_{max} \ge R^{i,n}_{min} \ge 1$.

For any $V^i \in L(D_i)$, any $\vec{d} \in \mathcal{X}$, and any $\vec{a}_{-i} \in D_{-1}$, we let

$$\pi^{\vec{a}_{-i}}_{\vec{d}} = \begin{cases} Pr_i(V^i|d_i)^{k_i}/Z_i & \text{if } \vec{a}_{-i} = \vec{d}_{-i} \\ \frac{1}{N_i!} & \text{otherwise} \end{cases},$$

where $Z_i = \sum_{V^i \in L(D_i)} Pr_i(V^i|d_i)^{k_i}$ is the normalizing factor, $1 = k_1 > k_2 > \ldots > k_p > 0$ are chosen in the following way: for any $i' < i \le p$, any $V^i, W^i \in L(D_i)$, and any $d_i, d'_i \in D_i$, if $R^{i,n}_{min} > 1$, then we must have that $(R^{i,n}_{max})^{k_i} < (R^{i',n}_{min})^{k_{i'}/2^{i-i'}}$.

We next prove that for any profile P_{CP} of no more than n CP-nets, we must have that $MLE_\pi(P_{CP}) \subseteq Seq(r_1, \ldots, r_p)(P_{CP})$. For the sake of contradiction, let P_{CP} be a profile of no more than n CP-nets with $MLE_\pi(P_{CP}) \not\subseteq Seq(r_1, \ldots, r_p)(P_{CP})$. Let $\vec{d} \in MLE_\pi(P_{CP})$, and i^* be the number such that there exists $\vec{d}^* \in Seq(r_1, \ldots, r_p)(P_{CP})$ such that for all $i' < i^*$, $d_{i'} = d^*_{i'}$, and $d_{i^*} \notin r_{i^*}(P_{CP}|_{\mathbf{x}_{i^*}:d_1\ldots d_{i^*-1}})$. Because $r_{i^*}(P_{CP}|_{\mathbf{x}_{i^*}:d_1\ldots d_{i^*-1}}) \ne D_{i^*}$, we must have that $R^{i^*,n}_{min} > 1$. Because $\vec{d} \in MLE_\pi(P_{CP})$, we must have that $\frac{\pi(P_{CP}|\vec{d})}{\pi(P_{CP}|\vec{d}^*)} \ge 1$. However, we have the following calculation that leads to a contradiction.

$$\begin{aligned}
1 \le& \frac{\pi(P_{CP}|\vec{d})}{\pi(P_{CP}|\vec{d}^*)} \\
=& \frac{\prod_{i=1}^p Pr_i(P_{CP}|_{\mathbf{x}_i:d_1\ldots d_{i-1}}|d_i)}{\prod_{i=1}^p Pr_i(P_{CP}|_{\mathbf{x}_i:d^*_1\ldots d^*_{i-1}}|d^*_i)} \\
=& \frac{\prod_{i=i^*}^p Pr_i(P_{CP}|_{\mathbf{x}_i:d_1\ldots d_{i-1}}|d_i)}{\prod_{i=i^*}^p Pr_i(P_{CP}|_{\mathbf{x}_i:d^*_1\ldots d^*_{i-1}}|d^*_i)} \\
\le& \frac{1}{(R^{i^*,n}_{min})^{k_{i^*}}} \cdot \prod_{i=i^*+1}^p (R^{i,n}_{max})^{k_i} \\
<& \frac{1}{(R^{i^*,n}_{min})^{k_{i^*}}} \cdot \prod_{i=i^*+1}^p (R^{i^*,n}_{min})^{k_{i^*}/2^{i-i^*}} < 1
\end{aligned}$$

Therefore, we must have that $MLE_\pi(P) \subseteq Seq(r_1, \ldots, r_p)(P)$ for all profiles P that consists of no more than n CP-nets. □

Proof of Theorem 8: We note that $\mathcal{N}$ that maximizes $\prod_{\mathcal{N}' \in P_{CP}} \prod_{i \le p, \vec{a}_{-i} \in D_{-i}} \delta^{\vec{a}_{-i}}_{\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}}}(\mathcal{N}'|_{\mathbf{x}_i:\vec{a}_{-i}})$ if and only if for any $i \le p$ and any $\vec{a}_{-i} \in D_{-i}$, the restriction of $\mathcal{N}$ on $\mathbf{x}_i$ given $\vec{a}_{-1}$ maximizes $\prod_{\mathcal{N}' \in P_{CP}} \delta^{\vec{a}_{-i}}_{\mathcal{N}|_{\mathbf{x}_i:\vec{a}_{-i}}}(\mathcal{N}'|_{\mathbf{x}_i:\vec{a}_{-i}})$.

The proof is then similar as the proof of Theorem 4. For the "if" part, for any $i \le p$, any $\vec{a}_{-i} \in D_{-i}$, and any $V^i, W^i \in L(D_i)$, we let $\delta^{\vec{a}_{-i}}_{W^i}(V^i) = Pr_i(V^i|W^i)$, where Pr_i is the probabilistic distribution that m_i corresponds to. For the "only if" part, we let $Pr_i(V^i|W^i) = \delta^{\vec{a}_{-i}}_{W^i}(V^i)$. □