10

A Maximum-Likelihood Approach

In voting, the joint decision is made based on the agents' preferences. Therefore, 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 (1785) for the cases of two and three alternatives. Much later, Young (1995) and Young (1988) showed that for arbitrary numbers of alternatives, the MLE rule derived from Condorcet's noise model coincides with Kemeny's rule (Kemeny, 1959). The approach was further pursued by Drissi-Bakhkhat and Truchon (2004). More recently, Conitzer and Sandholm (2005a) studied whether and how common voting correspondences can be represented as maximum likelihood estimators. Truchon (2008) studied a different way of viewing Borda as an MLE. We studied the relationship between MLEs and *ranking scoring rules* (Conitzer et al., 2009b). Conitzer (2011) took an MLE approach towards voting in social networks. We studied an MLE approach towards voting with partial orders Xia and Conitzer (2011b). The related notion of *distance rationalizability* has also received attention in the computational social choice community recently (Elkind et al., 2009a).

All of the above work does not assume any structure on the set of alternatives. In this chapter, we take an MLE 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, and 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 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 (respectively, strong decomposability) property. Then, when the input profile of CP-nets is compatible 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.

Then, we move to 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 characterize distance-based noise models axiomatically. 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 show that when the threshold is one, it is NP-hard to compute the winner, but that when it is equal to the number of issues, the winner can be computed in polynomial time.

10.1 Maximum-Likelihood Approach to Voting in Unstructured Domains

In the maximum likelihood approach to voting, it is assumed that there is a correct winner $d \in \mathcal{C}$, and each vote V is drawn conditionally independently given d, according to a conditional probability distribution $\pi(V|d)$. The independence structure of the noise model is illustrated in Figure 10.1. 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 (Conitzer and Sandholm, 2005a), which trivializes the question.

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

MLE_π is a voting correspondence, as there may be several alternatives d that maximize $\pi(P|d)$. Of course we can turn it into a voting rule by using a tie-breaking

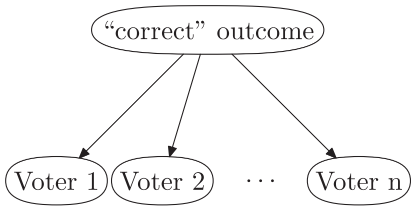


FIGURE 10.1: The noise model.

mechanism, but for most part of this chapter, we will study the properties of MLE correspondences. 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 $\pi(V|V^*)$. The maximum likelihood estimate is defined as follows.

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

In this chapter, we require that all such conditional probabilities to be positive for technical reasons.

Definition 10.1.1. *(Conitzer and Sandholm, 2005a). A voting rule (correspondence) r is a* maximum likelihood estimator for winners under i.i.d. votes (MLEWIV) *if there exists a noise model π such that for any profile P, we have that $MLE_\pi(P) = r(P)$.*

Conitzer and Sandholm (2005a) studied which common voting rules/correspondences are MLEWIVs.

10.2 Multi-Issue Domain Noise Models

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, $\vec{d} \in \mathcal{X}$. We let the voting language to be the set of all (possibly cyclic) CP-nets, that is, votes are given by CP-nets and are conditionally independent, given $\vec{d}$. Let CPnet($\mathcal{X}$) denote the set of all (possibly cyclic) CP-nets over $\mathcal{X}$. 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. We note that π is a conditional probability distribution over all CP-nets (in contrast to all linear orders in previous studies). Given this noise model, for any profile of CP-nets $P_{CP} = (\mathcal{N}_1, \ldots, \mathcal{N}_n)$, the maximum likelihood estimate of the correct winner is

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

Again, MLE_π is a voting correspondence.

Even if for all i, $|D_i| = 2$, the number of CP-nets (including cyclic ones) is $2^{p\cdot 2^{p-1}}$ (2 options for each entry of each CPT, and the CPT of any issue i has 2^{p-1} entries, one for each setting of the issues other than i). 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 chapter, 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 10.2.1. *A noise model is* very weakly decomposable *if for every $\vec{d}\in\mathcal{X}$, every $i \leqslant p$, 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($\mathcal{X}$), $\pi(\mathcal{N}|\vec{d}) = \prod_{i\leqslant p, \vec{a}_{-i}\in D_{-i}} \pi_{\vec{d}}^{\vec{a}_{-i}}(\mathcal{N}|_{X_i:\vec{a}_{-i}})$*

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 $\mathcal{N}|_{X_i:\vec{a}_{-i}}$ over specific X_i given specific $\vec{a}_{-i}$ (which contains the setting for X_i's parents as a sub-vector), when the correct 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. For instance, in Example 8.2.2, the event "the voter prefers white to pink to red wine when the main course is fish" is probabilistically independent (conditional on the correct outcome) of the event "the voter prefers beef to salad to fish when the wine is red."

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 (Kemeny, 1959) 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 10.2.2. *A very weakly decomposable noise model π is* weakly decomposable *if for any* $i \leqslant p$*, any* $\vec{d}_1, \vec{d}_2 \in \mathcal{X}$ *such that* $\vec{d}_1|_{X_i} = \vec{d}_2|_{X_i}$*, we must have that for any* $\vec{a}_{-i} \in D_{-i}$*,* $\pi_{\vec{d}_1}^{\vec{a}_{-i}} = \pi_{\vec{d}_2}^{\vec{a}_{-i}}$*. Here* $\vec{d}_1|_{X_i}$ *is the* X_i*-component of* $\vec{d}_1$*.*

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 10.2.3. *A very weakly decomposable noise model* π *is* strongly decomposable *if it is weakly decomposable, and for any* $i \leqslant p$*, any* $\vec{a}_{-i}, \vec{b}_{-i} \in D_{-i}$*, any* $\vec{d} \in \mathcal{X}$*, we must have that* $\pi_{\vec{d}}^{\vec{a}_{-i}} = \pi_{\vec{d}}^{\vec{b}_{-i}}$*.*

10.3 Characterizations of MLE correspondences

It seems that the MLE approaches 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 may imply 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 even though MLEs for very weakly decomposable noise models are defined over profiles of CP-nets, they can be easily extended to deal with profiles of linear orders in the following way. For each linear order V_j in the input profile P, let $\mathcal{N}_j$ denote the CP-net (possibly cyclic) that V_j extends. Then, we apply the MLE rule to select winner(s) from $(\mathcal{N}_1, \ldots, \mathcal{N}_n)$. We recall 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 10.3.1. *Over the domain of separable profiles, a voting correspondence r^c can be modeled as the MLE for a strongly decomposable noise model if and only if r^c is an issue-by-issue voting correspondence composed of MLEWIVs.*

Proof of Theorem 10.3.1: First we prove the "if" part. Let r^c be an issue-by-issue voting correspondence that is composed of $r_1^c, \ldots, r_p^c$, in which for any $i \leqslant p$, r_i^c 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 $i \leqslant p$, 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 $MLE_\pi(P) = r^c(P)$.

$$\begin{aligned} MLE_\pi(P) &= \arg\max_{\vec{d}} \prod_{i \leqslant p, \vec{a}_{-i} \in D_{-i}} \prod_{j=1}^{n} \pi_{\vec{d}}^{\vec{a}_{-i}}(V_j) \\ &= \arg\max_{\vec{d}} \prod_{i \leqslant p} \prod_{j=1}^{n} Pr((V_i|_{X_i})|d_i)^{|D_{-i}|} \end{aligned}$$

Therefore, $\vec{b} \in MLE_\pi(P)$ if and only if for any $i \leqslant p$, we have

$$b_i \in \arg\max_{d_i} \prod_{j=1}^{n} Pr((V_i|_{X_i})|b_i)$$

We note that for any $\vec{d'} \in r(P)$, we must have that $d'_i = \arg\max_{d_i} \prod_{j=1}^{n} Pr((V_i|_{X_i})|d_i)$. Therefore, $\vec{d'} \in MLE_\pi(P)$.

Next, we prove the "only if" part. For any MLE_π where π is strongly decomposable, we define an issue-by-issue voting rule as follows: for any $i \leqslant p$, let r_i^c 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 r^c and MLE_π are equivalent over the domain of separable profiles. □

A *candidate 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, $c(P) = \arg\max_{d \in \mathcal{X}} \sum_{V \in P} s(V, d)$.

Theorem 10.3.2. *Over the domain of separable profiles, a voting correspondence r^c can be modeled as the MLE for a weakly decomposable noise model if and only if r^c is an issue-by-issue voting correspondence composed of candidate scoring correspondences.*

Proof of Theorem 10.3.2: First we prove the "if" part. Let r^c be an issue-by-issue voting correspondence in which the issue-wise correspondence over D_i is $r^c_{s_i}$, which has scoring function s_i. Let $\pi^{\vec{a}_{-i}}_{d_i}$ denote $\pi^{\vec{a}_{-i}}_{\vec{d}}$, where the ith component of $\vec{d}$ is d_i. Because r is strongly decomposable, $\pi^{\vec{a}_{-i}}_{d_i}$ is well-defined. For any $i \leqslant p$, we claim that there exists a set of probability distributions $\pi^{\vec{a}_{-i}}_{\vec{d}}$, $\vec{d} \in \mathcal{X}, \vec{a}_{-i} \in D_{-i}$ over $L(D_i)$ such that for any $d_i \in D_i$, $d_i \in \arg\max_{b_i \in D_i} \prod_{j=1}^{n} \prod_{\vec{a}_{-i} \in D_{-i}} \pi^{\vec{a}_{-i}}_{b_i}(V_j|_{X_i})$ if and only if $d_i \in r^c_{s_i}(P|_{X_i})$.

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 $i \leqslant p$, 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 10.3.1. *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 $\ln(\prod_{\vec{d}_{-i} \in D_{-i}} \pi^{\vec{d}_{-i}}_{d_i}(V^i)) = k_i s_i(V^i, d_i) + t_i$.*

Proof of Claim 10.3.1: 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 $\hat{p}^j_{d_i} = \exp(s_i(l_j, d_i))$. For any $d_i \in D_i$, any

$1 \leqslant \alpha < \dfrac{K_i!}{K_i! - 1}$, 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, $f_{d_i}(1) \geqslant k_i s_i(l_{K_i!}, d_i) - \ln(K_i!)$. We note that $\lim_{\alpha \to \frac{K_i!}{K_i!-1}} f_{d_i}(\alpha) = -\infty$. It follows that there exists $1 \leqslant \alpha_{d_i} \leqslant \dfrac{K_i!}{K_i! - 1}$ such that $f_{d_i}(\alpha_{d_i}) = k_i s_i(l_{K_i!}, d_i) - \ln(K_i!)$.

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

- for any $j \leqslant K_i! - 1$, $\pi_{d_i}^{\vec{a}'_{-i}}(l_j) = \dfrac{1}{\alpha_{d_i}}(\exp(s_i(l_j, d_i)))^{k_i}$, $\pi_{d_i}^{\vec{a}^*_{-i}}(l_j) = \dfrac{\alpha_{d_i}}{K_i!}$.
- for any $j \leqslant K_i!$, 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 $j \leqslant K_i! - 1$, 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.

$$\ln(\prod_{\vec{d}_{-i}\in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(l_{K_i!}))$$

$$= \ln(\pi_{d_i}^{\vec{a}'_{-i}}(l_{K_i!}) \cdot \pi_{d_i}^{\vec{a}^*_{-i}}(l_{K_i!})) + (|D_{-i}| - 2)\ln(\frac{1}{K_i!})$$

$$= f_{d_i}(\alpha_i) - (|D_{-i}| - 2)\ln(K_i!)$$

$$= k_i s_i(l_{K_i!}, d_i) - (|D_{-i}| - 1)\ln(K_i!)$$

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 $\ln(\prod_{\vec{d}_{-i}\in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V^i)) = k_i s_i(V^i, d_i) + t_i$. $\square$

Next, we show that for any separable profile P, $r^c(P) = MLE_\pi(P)$. Similar to in the proof of Theorem 10.3.1, it suffices to prove that for any $i \leqslant p$, $\arg\max_{d_i\in D_i} \prod_{j\leqslant n} \prod_{\vec{d}_{-i}\in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{X_i}) = r^c_{s_i}(P|_{X_i})$.

$$\arg\max_{d_i\in D_i} \prod_{j\leqslant n} \prod_{\vec{d}_{-i}\in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{X_i})$$

$$= \arg\max_{d_i\in D_i} \ln(\prod_{j\leqslant n} \prod_{\vec{d}_{-i}\in D_{-i}} \pi_{d_i}^{\vec{d}_{-i}}(V_j|_{X_i}))$$

$$= \arg\max_{d_i\in D_i} \sum_{j\leqslant n} \sum_{\vec{d}_{-i}\in D_{-i}} \ln(\pi_{d_i}^{\vec{d}_{-i}}(V_j|_{X_i}))$$

$$= \arg\max_{d_i\in D_i} \sum_{j\leqslant n} \sum_{\vec{d}_{-i}\in D_{-i}} (k_i s_i(V_j|_{X_i}, d_i) + t_i)$$

$$= \arg\max_{d_i\in D_i} \sum_{j\leqslant n} \sum_{\vec{d}_{-i}\in D_{-i}} s_i(V_j|_{X_i}, d_i)$$

$$= r^c_{s_i}(P|_{X_i})$$

Next, we prove the "only if" part. Let π be a weakly decomposable noise model. For any $i \leqslant p$, 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 $s_i(P|_{X_i}, d_i)$ if and only if d_i

maximizes $\prod_{\mathcal{N}\in P}\prod_{\vec{a}_{-i}\in D_{-i}}\pi_{d_i}^{\vec{a}_{-i}}(\mathcal{N}|_{X_i})$, which means that $r^c(P) = MLE_\pi(P)$.

(End of the proof of Theorem 10.3.2). □

However, for sequential voting correspondences, we have the following negative result. A voting correspondence r^c satisfies *unanimity* if for any profile P in which each vote ranks an alternative $\vec{d}$ first, we have $r(P) = \{\vec{d}\}$. In the remainder of this section, w.l.o.g. we let $\mathcal{O} = X_1 > \cdots > X_p$.

Theorem 10.3.3. *Let $Seq(r_1^c, \ldots, r_p^c)$ 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 $Seq(r_1^c, \ldots, r_p^c)$ is the MLE.*

This theorem tells us that even assuming the weakest conditional independence of the noise model, the voting correspondence defined by the MLE of that noise model is different from any sequential voting correspondence satisfying unanimity. This suggests that the MLE approach gives us new voting rules/correspondences.

Proof of Theorem 10.3.3: For the sake of contradiction, we let $Seq(r_1^c, \ldots, r_p^c)$ 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 $r^c(P_1) = r^c(P_2)$, then $r^c(P_1 \cup P_2) = r^c(P_1)$; 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 10.3.2. *For any $i \leqslant p$, r_i^c satisfies consistency, anonymity (see Lang and Xia (2009)) and unanimity.*

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

$$\pi_{\vec{d}}^{X_1}(\mathcal{N}) = \prod_{\vec{a}_{-1}\in D_{-1}} \pi_{\vec{d}}^{\vec{a}_{-1}}(\mathcal{N}|_{X_1})$$

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

Let $\mathcal{N}_1, \mathcal{N}_2$ be CP-nets. We note that if $\mathcal{N}_1|_{X_1} = \mathcal{N}_2|_{X_1}$, then $\pi_{\vec{d}}^{X_1}(\mathcal{N}_1) = \pi_{\vec{d}}^{X_1}(\mathcal{N}_2)$; if for any $d_1 \in D_1$, $\mathcal{N}_1|_{X_{-1}:d_1} = \mathcal{N}_2|_{X_{-1}:d_1}$, then we must have that $\pi_{\vec{d}}^{X_{-1}}(\mathcal{N}_1) = \pi_{\vec{d}}^{X_{-1}}(\mathcal{N}_2)$, where $\mathcal{N}_1|_{X_{-1}:d_1}$ is the sub-CP-net of $\mathcal{N}_1$ given $X_1 = d_1$. For any $\mathcal{O}$-legal vote V that extends a CP-net $\mathcal{N}$, we write $\pi_{\vec{d}}^{X_1}(V) = \pi_{\vec{d}}^{X_1}(\mathcal{N})$ and $\pi_{\vec{d}}^{X_{-1}}(V) = \pi_{\vec{d}}^{X_{-1}}(\mathcal{N})$; for any $\mathcal{O}$-legal profile P, we write $\pi_{\vec{d}}^{X_1}(P) = \prod_{V \in P} \pi_{\vec{d}}^{X_1}(V)$ and $\pi_{\vec{d}}^{X_{-1}}(P) = \prod_{V \in P} \pi_{\vec{d}}^{X_{-1}}(V)$. It follows that for any $\mathcal{O}$-legal profile P, we have that

$$MLE_\pi(P) = \arg\max_{\vec{d} \in \mathcal{X}} [\pi_{\vec{d}}^{X_1}(P) \cdot \pi_{\vec{d}}^{X_{-1}}(P)]$$

For any linear order V, let $top(V) = \text{Alt}(V, 1)$. That is, $top(V)$ is the alternative that is ranked in the top position of V. For any $V_1^1, V_2^1 \in L(D_1)$ with $top(V_1^1) \neq top(V_2^1)$, 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 r_1^c satisfies unanimity, we must have that $r_1^c(P_{1,n}^1) = \{top(V_1^1)\}$ and $r_1^c(P_{2,n}^1) = \{top(V_2^1)\}$. For any $j \leqslant n$, 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 $j \leqslant n-1$ and $b_1 \in D_1$ with $b_1 \neq top(V_1^1)$, such that $top(V_1^1) \in r_1^c(Q_{j,n})$ and $b_1 \in r_1^c(Q_{j+1,n})$. For any $n \in \mathbb{N}$, we let C_n denote the set of pairs (a_1, b_1) such that

- $a_1, b_1 \in D_1$, $a_1 \neq b_1$.
- There exists two profiles W_1^1, W_2^1 over D_1 such that $a_1 \in r_1^c(W_1^1)$, $b_1 \in r_1^c(W_2^1)$, 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, $a_1 \in r_1^c(Q)$, 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 $n \in \mathbb{N}$, $C_n \neq \varnothing$. Because $|D_1| < \infty$, there exists $(a_1, b_1) \in (D_1)^2$ such that for any $k \in \mathbb{N}$, there exists $n \geqslant k$ such that $(a_1, b_1) \in C_n$.

Claim 10.3.3. *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* $\dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}')} = \dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}^*)}$, *where* $\vec{a} = (a_1, \vec{a}_{-1})$, $\vec{b} = (b_1, \vec{b}_{-1})$.

Proof of Claim 10.3.3: Suppose for the sake of contradiction there exist $\vec{a}_{-1}, \vec{b}_{-1}$, and $\mathcal{N}', \mathcal{N}^*$ so that $\dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}')} \neq \dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}^*)}$. Without loss of generality we let $\dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}')} > \dfrac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}^*)}$. We next claim that there exits a natural number k such that for any $i \leqslant p$ 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 $r_i^c(P^i) = \{d_i\}$.

Claim 10.3.4. *There exists $k \in \mathbb{N}$ such that for any $i \leqslant p$, 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 $r_i^c(P) = \{d_i\}$.*

Proof of Claim 10.3.4: Let $U = \max_{\vec{d}_1, \vec{d}_2, \mathcal{N}} \dfrac{Pr(\mathcal{N}|\vec{d}_1)}{Pr(\mathcal{N}|\vec{d}_2)}$. Let $u = \min_{\vec{d}_1 \neq \vec{d}_2, \mathcal{N}: top(\mathcal{N}) = \vec{d}_1} \dfrac{Pr(\mathcal{N}|\vec{d}_1)}{Pr(\mathcal{N}|\vec{d}_2)}$. 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 $j \leqslant k$, $top(\mathcal{N}_j) = (\vec{d}_{-i}, top(V^i))$.
- For any $j \leqslant k$, $\mathcal{N}_j|_{X_i: d_1, \ldots, d_{i-1}} = V^i$.
- Other conditional preferences are defined arbitrarily.

Because $Seq(r_1^c, \ldots, r_p^c)$ satisfies unanimity, we have that $Seq(r_1^c, \ldots, r_p^c)(\mathcal{N}_1, \ldots, \mathcal{N}_{k-1}) = \{\vec{d}\}$. Therefore, for any $\vec{d'} \in \mathcal{X}$ and any CP-net $\mathcal{N}$, we have the following calculation:

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

$$\geqslant u^{(k-1)} \frac{1}{U} > 1$$

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

(End of proof of Claim 10.3.4.) □

Let $\mathcal{N}_{\vec{a}}$ be a CP-net such that $top(\mathcal{N}_{\vec{a}}) = \vec{a}$ and $top(\mathcal{N}|_{X_{-1}:b_1}) = \vec{b}_{-1}$. That is, $\mathcal{N}_{\vec{a}}$ is a CP-net in which $\vec{a}$ is ranked in the top position, and given $X_1 = b_1$, $\vec{b}_{-1}$ is ranked in the top position. Next, we show that for any CP-net $\mathcal{N}$, $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} = \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}$. Suppose for the sake of contradiction, there exists $\mathcal{N}$ such that $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} \neq \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}$. We next show contradiction in the case $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} > \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}$. Let $U_{X_1} = \max_{\vec{d}_1, \vec{d}_2, \mathcal{N}} \frac{\pi_{\vec{d}_1}^{X_1}(\mathcal{N})}{\pi_{\vec{d}_2}^{X_1}(\mathcal{N})}$. Let K be a natural number such that $(\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} / \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})})^K > U_{X_1}^2$. 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 $a_1 \in r_1^c(V_1^1, \ldots, V_n^1)$ and $b_1 \in r_1^c(W_1^1, V_2^1, \ldots, V_n^1)$. 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 $j \leqslant n$, $\mathcal{N}_j|_{X_1} = \hat{\mathcal{N}}_j|_{X_1} = V_j^1$; $\mathcal{N}_1'|_{X_1} = W_1^1$.
- For any $j_1 \leqslant K$, $1 \leqslant j_2 \leqslant k-1$, and any $d_1 \in D_1$, $\mathcal{N}_{(j_1-1)k+j_2}|_{X_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{X_{-1}:d_1}$ and $\mathcal{N}_{j_1 k}|_{X_{-1}:d_1} = \mathcal{N}|_{X_{-1}:d_1}$; for any $j \leqslant n$ and any $d_1 \in D_1$, $\mathcal{N}_j|_{X_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{X_{-1}:d_1}$.

- For any $kK+1 \leqslant j \leqslant n$, $\mathcal{N}_j = \hat{\mathcal{N}}_j = \mathcal{N}_{\vec{a}}$.
- For any $d_1 \in D_1$, $\mathcal{N}'_1|_{X_{-1}:d_1} = \mathcal{N}_{\vec{a}}|_{X_{-1}:d_1}$.

For any $j \leqslant n$, 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.

- $a_1 \in r_1^c(P|_{X_1}), a_1 \in r_1^c(\hat{P}|_{X_1})$, $b_1 \in r_1^c(P'|_{X_1}), b_1 \in r_1^c(\hat{P}'|_{X_1})$.
- For any $1 \leqslant i \leqslant p-1$, $P|_{X_i:a_1\ldots a_{i-1}} = K((k-1)\mathcal{N}_{\vec{a}}|_{X_i:a_1\ldots a_{i-1}} \cup \mathcal{N}'|_{X_i:a_1\ldots a_{i-1}}) \cup (n-kK)\mathcal{N}_{\vec{a}}|_{X_i:a_1\ldots a_{i-1}}$. From Claim 10.3.4 we have that $r_i^c((k-1)\mathcal{N}_{\vec{a}}|_{X_i:a_1\ldots a_{i-1}} \cup \mathcal{N}'|_{X_i:a_1\ldots a_{i-1}}) = \{a_i\}$. Because r_i^c satisfies unanimity and consistency, and for any $i \leqslant p$, $top(\mathcal{N}_{\vec{a}}|_{X_i:a_1\ldots a_{i-1}}) = a_i$, we have that for any $i \leqslant p$, $r_i^c(P|_{X_i:a_1\ldots a_{i-1}}) = \{a_i\}$. Similarly for any $i \leqslant p$, $r_i^c(\hat{P}|_{X_i:a_1\ldots a_{i-1}}) = \{a_i\}$.
- For any $1 \leqslant i \leqslant p-1$, $P|_{X_i:b_1\ldots b_{i-1}} = K((k-1)\mathcal{N}_{\vec{a}}|_{X_i:b_1\ldots b_{i-1}} \cup \mathcal{N}'|_{X_i:b_1\ldots b_{i-1}}) \cup (n-kK)\mathcal{N}_{\vec{a}}|_{X_i:b_1\ldots b_{i-1}}$. Similarly, we have that for any $1 \leqslant i \leqslant p$, $r_i^c(P'|_{X_i:b_1\ldots b_{i-1}}) = r_i^c(\hat{P}'|_{X_i:b_1\ldots b_{i-1}}) = \{b_i\}$.

Therefore, we have that $\vec{a} \in Seq(r_1^c, \ldots, r_p^c)(P), \vec{a} \in Seq(r_1^c, \ldots, r_p^c)(\hat{P})$, and $\vec{b} \in Seq(r_1^c, \ldots, r_p^c)(P'), \vec{b} \in Seq(r_1^c, \ldots, r_p^c)(\hat{P}')$. That is, $\dfrac{Pr(P'|\vec{b})}{Pr(P'|\vec{a})} \geqslant 1$, $\dfrac{Pr(\hat{P}'|\vec{b})}{Pr(\hat{P}'|\vec{a})} \geqslant 1$. We note that P and P' differ only on the first vote. Therefore, we have the following calculation.

$$
\begin{aligned}
1 \leqslant & \frac{Pr(P'|\vec{b})}{Pr(P'|\vec{a})} \\
= & \frac{\pi_{\vec{b}}^{X_1}(V_1') \cdot \pi_{\vec{b}}^{X_{-1}}(V_1') \prod_{2\leqslant j\leqslant n}(\pi_{\vec{b}}^{X_1}(V_j) \cdot \pi_{\vec{b}}^{X_{-1}}(V_j))}{\pi_{\vec{a}}^{X_1}(V_1') \cdot \pi_{\vec{a}}^{X_{-1}}(V_1') \prod_{2\leqslant j\leqslant n}(\pi_{\vec{a}}^{X_1}(V_j) \cdot \pi_{\vec{a}}^{X_{-1}}(V_j))} \\
= & \frac{\pi_{\vec{b}}^{X_1}(V_1')}{\pi_{\vec{a}}^{X_1}(V_1')} \cdot \frac{\pi_{\vec{a}}^{X_1}(V_1)}{\pi_{\vec{b}}^{X_1}(V_1)} \cdot \frac{Pr(P|\vec{b})}{Pr(P|\vec{a})} \\
\leqslant & U_{X_1}^2 \frac{Pr(P|\vec{b})}{Pr(P|\vec{a})}
\end{aligned}
$$

Therefore, $\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})} \leqslant U_{X_1}^2$. We note that P and P' differ on K votes.

$$
\begin{aligned}
& (\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})})/(\frac{Pr(\hat{P}|\vec{a})}{Pr(\hat{P}|\vec{b})}) \\
= & (\prod_{j=1}^{K} \frac{\pi_{\vec{a}}^{X_1}(V_{jk}) \cdot \pi_{\vec{a}}^{X_{-1}}(V_{jk})}{\pi_{\vec{b}}^{X_1}(V_{jk}) \cdot \pi_{\vec{b}}^{X_{-1}}(V_{jk})})/(\prod_{j=1}^{K} \frac{\pi_{\vec{a}}^{X_1}(\hat{V}_{jk}) \cdot \pi_{\vec{a}}^{X_{-1}}(\hat{V}_{jk})}{\pi_{\vec{b}}^{X_1}(\hat{V}_{jk}) \cdot \pi_{\vec{b}}^{X_{-1}}(\hat{V}_{jk})}) \\
= & (\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})}/\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})})^K \\
> & U_{X_1}^2
\end{aligned}
$$

We note that $(\frac{Pr(\hat{P}|\vec{a})}{Pr(\hat{P}|\vec{b})}) \geqslant 1$. Therefore, $\frac{Pr(P|\vec{a})}{Pr(P|\vec{b})} > U_{X_1}^2$, which is a contradiction.

Similarly, for the case of $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} < \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}$ we still have a contradiction.

Hence, $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N})} = \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}_{\vec{a}})}$ for all $\mathcal{N}$, which means that for any $\mathcal{N}'$ and $\mathcal{N}^*$, we must have that $\frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}')}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}')} = \frac{\pi_{\vec{a}}^{X_{-1}}(\mathcal{N}^*)}{\pi_{\vec{b}}^{X_{-1}}(\mathcal{N}^*)}$.

(End of proof of Claim 10.3.3.) □

By Claim 10.3.3, for any CP-net $\mathcal{N}$, any $\vec{b}_{-1}, \vec{b}'_{-1} \in D_{-1}$, we must have that $\frac{\pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N})}{\pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_{\vec{a}})} = \frac{\pi^{X_{-1}}_{\vec{a}}(\mathcal{N})}{\pi^{X_{-1}}_{\vec{a}}(\mathcal{N}_{\vec{a}})} = \frac{\pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N})}{\pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}$, which means that $\frac{\pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N})}{\pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N})} = \frac{\pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_{\vec{a}})}{\pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}$. 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 $\mathcal{N}_1|_{X_1} = \mathcal{N}_2|_{X_1}$. Because $Seq(r^c_1, \ldots, r^c_p)$ 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.

$$
\begin{aligned}
1 <& \frac{Pr(\mathcal{N}_1|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_1|(b_1,\vec{b}_{-1}))} \\
=& \frac{\pi^{X_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_1)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_1)}{\pi^{X_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_1)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_1)} \\
=& \frac{\pi^{X_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_{\vec{a}})}{\pi^{X_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_{\vec{a}})} && \text{(Because } \mathcal{N}_1|_{X_1} = \mathcal{N}_2|_{X_1}\text{)} \\
=& \frac{\pi^{X_1}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}'_{-1})}(\mathcal{N}_2)}{\pi^{X_1}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)\cdot \pi^{X_{-1}}_{(b_1,\vec{b}_{-1})}(\mathcal{N}_2)} \\
=& \frac{Pr(\mathcal{N}_2|(b_1,\vec{b}'_{-1}))}{Pr(\mathcal{N}_2|(b_1,\vec{b}_{-1}))} \\
<& 1
\end{aligned}
$$

Therefore, we have a contradiction. **(End of proof of Theorem 10.3.3.)** □

However, a connection between MLEs for very weakly decomposable noise models and sequential voting correspondences 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. That is, if the local correspondences can be justified by a noise model, then, to some extent, so can the sequential voting correspondence that uses these local rules.

Theorem 10.3.4. *For any $n \in \mathbb{N}$ and any sequential voting correspondence $Seq(r_1^c, \ldots, r_p^c)$ where for each $i \leqslant p$, r_i^c 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 $MLE_\pi(P) \subseteq Seq(r_1^c, \ldots, r_p^c)(P)$.*

Proof of Theorem 10.3.4: Let r_i be the MLEWIV with the conditional probability distribution $Pr_i(V^i|d_i)$, where $V^i \in L(D_i)$, $d_i \in D_i$. For any $i \leqslant p$, we let $R_{max}^{i,n} = \max_{P_i, P_i', d_i, d_i'} \left\{ \dfrac{Pr_i(P_i|d_i)}{Pr_i(P_i'|d_i')} \right\}$, where $d_i, d_i' \in D_i$, and P_i and P_i' are profiles with the same number (but no more than n) of linear orders over D_i. We let $R_{min}^{i,n} = 1$ if r_i is the trivial correspondence that always outputs the whole domain; and $R_{min}^{i,n} = \min_{P_i, \vec{d}_i, \vec{d}_i'} \left\{ \dfrac{Pr_i(P_i|d_i)}{Pr_i(P_i|d_i')} : \dfrac{Pr_i(P_i|d_i)}{Pr_i(P_i|d_i')} > 1 \right\}$, 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 $i \leqslant p$, any $n \in \mathbb{N}$, we have that $R_{max}^{i,n} \geqslant R_{min}^{i,n} \geqslant 1$.

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

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

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

We next prove that for any profile P_{CP} of no more than n CP-nets, we must have that $MLE_\pi(P_{CP}) \subseteq Seq(r_1^c, \ldots, r_p^c)(P_{CP})$. For the sake of contradiction, let P_{CP} be a profile of no more than n CP-nets with $MLE_\pi(P_{CP}) \nsubseteq Seq(r_1^c, \ldots, r_p^c)(P_{CP})$. Let $\vec{d} \in$

$MLE_\pi(P_{CP})$, and i^* be the number such that there exists $\vec{d^*} \in Seq(r_1^c, \ldots, r_p^c)(P_{CP})$ such that for all $i' < i^*$, $d_{i'} = d^*_{i'}$, and $d_{i*} \notin r^c_{i*}(P_{CP}|_{X_{i*}:d_1\ldots d_{i*-1}})$. Because $r^c_{i*}(P_{CP}|_{X_{i*}:d_1\ldots d_{i*-1}}) \neq D_{i*}$, we must have that $R^{i*,n}_{min} > 1$. Because $\vec{d} \in MLE_\pi(P_{CP})$, we must have that $\frac{\pi(P_{CP}|\vec{d})}{\pi(P_{CP}|\vec{d^*})} \geqslant 1$. However, we have the following calculation that leads to a contradiction.

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

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

10.4 Distance-Based Models

We have shown in the previous section that the MLE approach may give us new voting rules in multi-issue domains. However, assuming very weak decomposability, there are too many (exponentially many) parameters in the noise model, which makes it very hard to implement a rule based on the MLE approach. In this section, we focus on a family of maximum likelihood estimators that are based on noise models defined over multi-binary-issue domains (domains composed of binary issues), and that need only a few parameters to be specified. We recall that a CP-net on a multi-binary-issue domain corresponds to a directed hypercube in which each edge has a direction representing the local preference. A very weakly decomposable noise model π can be represented by a collection of weighted directed hypercubes, one for each correct winner, in which the weight of each directed edge is the probability of the local preference represented by the directed edge. For any outcome $\vec{d} \in \mathcal{X}$, any issue X_i, any $\vec{e}_{-i} \in D_{-i}$, and any $d_i \neq d'_i \in D_i$, the weight on the directed edge $((\vec{e}_{-i}, d_i), (\vec{e}_{-i}, d'_i))$ of the weighted hypercube corresponding to the correct winner $\vec{d}$ is denoted by $\pi_{\vec{d}}^{\vec{e}_{-i}}(d_i \succ d'_i)$, and represents the probability that a given voter reports the preference $\vec{e}_{-i} : d_i \succ d'_i$ in her CP-net, given that the correct winner is $\vec{d}$.[1] 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.

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, $|\vec{d} - \vec{d'}| = \#\{i \leqslant p : d_i \neq d'_i\}$. 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). 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 10.4.1. *Let $\mathcal{X}$ be a multi-binary-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 $i \leqslant p$, and any $\vec{a}_{-i} \in D_{-i}$*

[1] For every pair of alternatives differing on exactly one issue, there is exactly one weighted edge between them; the direction of the edge only says that we are going further from the correct winner. This will be made more precise after Definition 10.4.1.

with $|\vec{a}_{-i} - \vec{d}_{-i}| = k \leqslant p-1$*, we have that* $\pi_{\vec{d}}^{\vec{a}_{-i}}(d_i \succ \bar{d}_i) = q_k$.

The intuition behind the notion of a distance-based model is as follows. First, it is plausible to assume that the "closer" two alternatives are to the correct alternative, the more likely a given voter will order them in the "correct" way, that is, will prefer the one which is closer to the correct alternative. The family of distance-based voting rules is actually more general than this, because we do not impose $q_1 \geqslant \ldots \geqslant q_{p-1}$, but we may of course add this restriction if we wish to. Moreover, the choice of the Hamming distance is not necessary, and other intuitive distance-based models can be defined, using other distances – for instance, domain-dependent distances. But, the Hamming distance is a natural starting point (most works in distance-based belief base merging and distance-based belief revision also focus on the Hamming distance).

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 X_i for some $i \leqslant p$, if $\vec{d_1}|_{X_i} = d_i$, then the direction of e is from $\vec{d_1}$ to $\vec{d_2}$; if $\vec{d_2}|_{X_i} = d_i$, 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 X_i component is the same as the X_i component 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 10.2.

We are especially interested in a special type of distance-based models in which there exists a threshold $1 \leqslant k \leqslant p$ and $q > \frac{1}{2}$, such that for any $i < k$, we have that $q_i = q$, and for any $k \leqslant i \leqslant p-1$, we have that $q_i = \frac{1}{2}$. Such a model is denoted

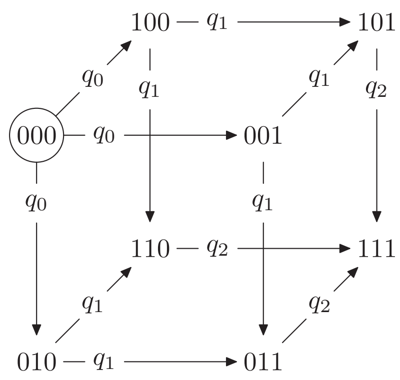
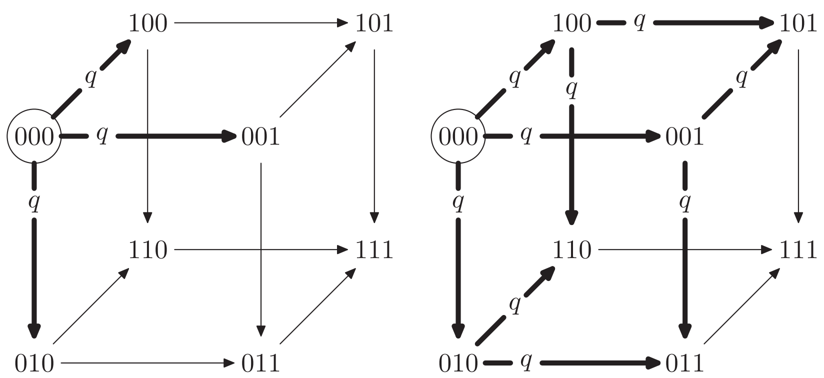


FIGURE 10.2: The distance-based model $\pi_{(q_0,q_1,q_2)}$ when the correct winner is 000.

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 $k \leqslant p$ 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 10.4.2. *Let* $p = 3$. $\pi_{1,q}$ *and* $\pi_{2,q}$ *are illustrated in Figure 10.3 (when the correct winner is* 000*).*



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

FIGURE 10.3: Distance-based threshold models. The weight of the bold edges is $q > \frac{1}{2}$; the weight of all other edges is $\frac{1}{2}$.

We next present a direct method for computing winners under the MLE correspondences of distance-based threshold models. For any $1 \leqslant k \leqslant p$, 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 $\vec{a}_{-i}=\vec{b}_{-i}$, $a_i=d_i$, $b_i=\bar{d}_i$, $|(a_i,b_i)-\vec{d}|\leqslant k-1$, 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 X_i, given any $\vec{a}_{-i}\in D_{-i}$) in $\mathcal{N}$ that are $d_i\succ\bar{d}_i$, where 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 $N_k(\vec{d},P_{CP})=\sum_{\mathcal{N}\in P_{CP}}N_k(\vec{d},\mathcal{N})$.

Theorem 10.4.3. *For any $k\leqslant p$, 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.

Proof of Theorem 10.4.3: For any $k\leqslant p$, any $\vec{d}\in\mathcal{X}$, we let $L_k=\#\{e:|e-\vec{d}|\leqslant k-1\}$. That is, L_k is the number of edges in the hypercube whose distance from a given alternative $\vec{d}$ is no more than $k-1$. For any $\vec{d}\in\mathcal{X}$ and any CP-net $\mathcal{N}$, we have that

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

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

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

Example 10.4.5. *Consider two binary issues X_1, X_2, and three voters, who report the following CP-nets:*

- *$\mathcal{N}_1$ has an edge from X_1 to X_2, and the following local preferences: $\{0_1 \succ 1_1, 0_1 : 0_2 \succ 1_2, 1_1 : 1_2 \succ 0_2\}$.*
- *$\mathcal{N}_2$ has an edge from X_1 to X_2 and an edge from X_2 to X_1, and the following local preferences: $\{0_2 : 1_1 \succ 0_1, 1_2 : 0_1 \succ 1_1, 0_1 : 1_2 \succ 0_2, 1_1 : 0_2 \succ 1_2\}$.*
- *$\mathcal{N}_3$ has no edge, and the following local preferences: $\{1_1 \succ 0_1, 1_2 \succ 0_2\}$.*

Let $P_{CP} = (\mathcal{N}_1, \mathcal{N}_2, \mathcal{N}_3)$.

First, consider $k = 1$. Let us compute $N_1(1_1 1_2, \mathcal{N}_1)$. There are two edges whose distance to $1_1 1_2$ is 0: one from $1_1 1_2$ to $1_1 0_2$ and one from $1_1 1_2$ to $0_1 1_2$. The first one is in the preference relation induced from $\mathcal{N}_1$; the second one is not. Therefore, $N_1(1_1 1_2, \mathcal{N}_1) = 1$. Similarly, we get $N_1(1_1 1_2, \mathcal{N}_2) = 0$ and $N_1(1_1 1_2, \mathcal{N}_3) = 2$, henceforth, $N_1(1_1 1_2, P_{CP}) = 3$. Similar calculations lead to $N_1(1_1 0_2, P_{CP}) = 3$, $N_1(0_1 1_2, P_{CP}) = 4$ and $N_1(0_1 0_2, P_{CP}) = 2$, hence $MLE_{\pi_{1,q}}(P_{CP}) = \{0_1 1_2\}$ (for any value of $q > \frac{1}{2}$).

Now, consider $k = 2$. Let us compute $N_1(1_1 1_2, \mathcal{N}_1)$. Now, we have to consider all four edges, since all of them are at a distance 0 or 1 to $1_1 1_2$. The two edges not considered for the case $k = 1$ are the edge from $0_1 1_2$ to $0_1 0_2$ and one from $1_1 0_2$ to $0_1 1_2$. In both cases, voter 1 prefers the alternative which is further from $1_1 1_2$, therefore, $N_2(1_1 1_2, \mathcal{N}_1) = 1$. Similarly, we get $N_2(1_1 1_2, \mathcal{N}_2) = 2$ and $N_2(1_1 1_2, \mathcal{N}_3) = 4$, henceforth, $N_2(1_1 1_2, P_{CP}) = 7$. Similar calculations lead to $N_2(1_1 0_2, P_{CP}) = 5$, $N_2(0_1 1_2, P_{CP}) = 7$ and $N_1(0_1 0_2, P_{CP}) = 5$, hence $MLE_{\pi_{2,q}}(P_{CP}) = \{0_1 1_2, 1_1 1_2\}$.

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 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 X_i takes time that is exponential in $|Par_G(X_i)|$ (the number of parents of X_i in the directed graph of $\mathcal{N}$), the CPT of X_i in $\mathcal{N}$ itself is also exponential in $|Par_G(X_i)|$ (for each setting of $Par_G(X_i)$, there is an entry in $CPT(X_i)$). This explains why the algorithm runs in polynomial time.

Algorithm 10.4.1. INPUT: $p \in \mathbb{N}$, $\frac{1}{2} < q < 1$, and a profile of CP-nets P_{CP} over a binary domain consisting of p issues.

1. For each $i \leqslant p$:

1a. Let $S_i = 0$, $W_i = \emptyset$.

1b. For each CP-net $\mathcal{N} \in P_{CP}$: let $Par_G(X_i) = \{X_{i_1}, \ldots, X_{i_{p'}}\}$ be the parents of X_i in the directed graph of $\mathcal{N}$. Let l be the number of settings $\vec{y}$ of $Par_G(X_i)$ for which $\mathcal{N}|_{X_i:\vec{y}} = 0_i \succ 1_i$. Let $S_i \leftarrow S_i + l2^{p-p'} - 2^{p-1}$. Here, p' is the number of parents of X_i, and $l2^{p-p'} - 2^{p-1}$ 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 10.4.6. *The output of Algorithm 10.4.1 is $MLE_{\pi_{p,q}}(P_{CP})$, and the algorithm runs in polynomial time.*

Proof of Proposition 10.4.6: First we prove that the output of Algorithm 10.4.1 is $MLE_{\pi_{p,q}}(P_{CP})$. For any $\vec{d} \in \mathcal{X}$, $N_p(\vec{d}, P_{CP}) = \sum_{i \leqslant p,} \#\{\vec{a}_{-1} \in D_{-1} : (d_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (\bar{d}_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\}$. We note that $d_i \in W_i$ if and only if $\#\{\vec{a}_{-1} \in D_{-1} : (d_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (\bar{d}_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\} > \#\{\vec{a}_{-1} \in D_{-1} : (\bar{d}_i, \vec{a}_{-i}) \succ_{\mathcal{N}} (d_i, \vec{a}_{-i}), \mathcal{N} \in P_{CP}\}$. Therefore, $\vec{d} \in MLE_{\pi_{p,q}}(P_{CP})$ if and only if for all $i \leqslant p$, 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 $O(2^{|Par_G(X_i)|})$, and $CPT(X_i)$ of the CP-net $\mathcal{N}$ has exactly $2^{|Par_G(X_i)|}$ entries, which means that the complexity of computing l is in polynomial of the size of $CPT(X_i)$ of the input. Therefore, Algorithm 10.4.1 is a polynomial-time algorithm. □

The next example shows how to compute the winners under $MLE_{\pi_{p,q}}$ for the profile defined in Example 10.4.5.

Example 10.4.5, continued *Let us first compute S_1. In $\mathcal{N}_1$ (respectively, $\mathcal{N}_1$ and $\mathcal{N}_3$), the table for x_1 contributes to 2 edges (respectively, one edge and no edge) from 0_1 to 1_1, and to no edge (respectively, one edge and two edge) from 1_1 to 0_1, therefore $S_1 = (+2) + 0 + (-2) = 0$. Similarly, $S_2 = 0 + 0 + (-2) = -2$. Therefore, $W_1 = \{0_1, 1_1\}$ and $W_2 = \{1_2\}$, which gives us $MLE_{\pi_{2,q}}(P_{CP}) = \{0_1 1_2, 1_1 1_2\}$.*

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 $N_1(\vec{d}, P_{CP}) \geqslant T$, is NP-complete.

Theorem 10.4.7. *It is* NP*-complete 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 $N_1(\vec{d}, P_{CP}) \geqslant T$.*

Proof of Theorem 10.4.7: By Theorem 10.4.3, 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 $T \leqslant pn$, we are asked whether or not there exists $\vec{d} \in \mathcal{X}$ such that

$N_1(\vec{d}, P) \geqslant T$.

We prove the NP-hardness by reduction from the decision problem of MAX2SAT. The inputs of an instance of the decision problem of MAX2SAT consists of (1) a set of t atomic propositions $x_1, \ldots, x_t$; (2) a formula $F = C_1 \wedge \ldots \wedge C_m$ represented in *conjunctive normal form*, in which for any $i \leqslant m$, $C_i = l_{i_1} \vee l_{i_2}$, and there exists $j_1, j_2 \leqslant t$ such that l_{i_1} is x_{j_1} or $\neg x_{j_1}$, and l_{i_2} is x_{j_2} or $\neg x_{j_2}$; (3) $T \leqslant m$. 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 MAX2SAT, we construct a decision problem instance of computing a winner under $MLE_{\pi_{1,q}}$ as follows.

• Let $\mathcal{X}$ be composed of t issues $X_1, \ldots, X_t$.

• Let $T' = 16T - 12m$.

• For any $i \leqslant m$, 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 $j \leqslant t$, we let 0_j corresponds to X_j being false, and 1_j corresponds to X_j being true. Then, any valuation of the atomic propositions is uniquely identified by an alternative. We next define six CP-nets as follows:

– $\mathcal{N}_{i,1}$: the DAG of $\mathcal{N}_{i,1}$ has only one directed edge (X_{i_1}, 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 $0_j \succ 1_j$.

– $\mathcal{N}_{i,2}$: the DAG of $\mathcal{N}_{i,2}$ has only one directed edge (X_{i_1}, 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 $0_j \succ 1_j$.

– $\mathcal{N}_{i,3}$: the DAG of $\mathcal{N}_{i,3}$ has only one directed edge (X_{i_2}, 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 $0_j \succ 1_j$.

We next obtain $\mathcal{N}'_{i,1}$, $\mathcal{N}'_{i,2}$, and $\mathcal{N}'_{i,3}$ from $\mathcal{N}_{i,1}$, $\mathcal{N}_{i,2}$, and $\mathcal{N}_{i,3}$, respectively, by letting $1_j \succ 0_j$ for any j with $j \neq i_1$ and $j \neq i_2$. 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-nets be $P_{CP} = (\vec{\mathcal{N}_1}, \ldots, \vec{\mathcal{N}_m})$.

We make the following claim about the number of consistent edges between an alternative $\vec{d}$ and $\vec{\mathcal{N}}_i$.

Claim 10.4.1. *For any $\vec{d} \in \mathcal{X}$ and any $i \leqslant m$,*

$$N_1(\vec{d}, \vec{\mathcal{N}}_i) = \begin{cases} 4 & \text{if } \vec{d}_{i_1} = v_{i_1} \text{ or } d_{i_2} = v_{i_2} \\ -12 & \text{if } \vec{d}_{i_1} = \bar{v}_{i_1} \text{ and } d_{i_2} = \bar{v}_{i_2} \end{cases}$$

Claim 10.4.1 states that the number of consistent edges between $\vec{d}$ and $\vec{\mathcal{N}}_i$ within distance 1 is 4 if the clause C_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 Theorem 10.4.3 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 decision problem is in NP. Therefore, the decision problem is NP-complete to compute a winner under $MLE_{\pi_{1,q}}$. □

As we have seen (cf. Corollary 10.4.4), for a given multi-issue domain composed of p binary issues, there are *exactly* p voting correspondences defined by distance-based threshold models. As far as we know, these voting correspondences are entirely novel, and are tailored especially for multi-issue domains. Now, among these p voting correspondences, two are even more natural and interesting: $MLE_{\pi_{1,q}}$ and $MLE_{\pi_{p,q}}$. $MLE_{\pi_{1,q}}$ proceeds by electing the alternatives which maximize the sum, over all voters, of the number of neighboring alternatives in the voter's hypercube to which she prefers $\vec{x}$. Now, recall that the Borda correspondence can be characterized as the correspondence where candidate x is a winner if it maximizes the sum, over all voters, of the number of candidates the voter prefers to x. Therefore, $MLE_{\pi_{1,q}}$ is somewhat reminiscent of Borda—except, of course, that we do not count all alternatives defeated by $\vec{x}$ but only defeated alternatives that are one of its neighbors in the hypercube. $MLE_{\pi_{p,q}}$ is even more intuitive: for each issue X_i, the winning value maximizes the number of edges (summing over all voters) that are in favor of it, that is, it is somewhat reminiscent of Kemeny.

So, $MLE_{\pi_{1,q}}$ and $MLE_{\pi_{p,q}}$ are genuinely new voting correspondences for multi-issue binary domains, which can be characterized in terms of maximum likelihood estimators and are quite intuitive; lastly, $MLE_{\pi_{p,q}}$ can be computed in polynomial time. We conjecture that for any $2 \leqslant k \leqslant p-1$, winner determination for $MLE_{\pi_{k,q}}$ is NP-hard.

10.5 Summary

In this chapter, 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. We first studied whether issue-by-issue voting rules and sequential voting rules can be represented by the MLE of some noise model. For separable input profiles, we characterized MLEs of strongly/weakly decomposable models as issue-by-issue voting correspondences composed of local MLEWIVs/candidate 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.

In the case where all issues are binary, we proposed 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.

We note that, whereas Section 10.3 has a non-constructive flavor because we studied existing voting mechanisms and Theorem 10.3.3 is an impossibility theorem, quite the opposite is the case for Section 10.4. 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 (the decision problem is NP-complete at worst, and sometimes polynomial in the size of the CP-nets). We believe that these new rules are promising.