Sufficient and Necessary Conditions for the Identifiability of the Q-matrix

Yuqi Gu and Gongjun Xu

University of Michigan

Supplementary Material

In this supplementary material, we present proofs of the theorems in the main text and also provide simulation results to support them. We first introduce some notations and a useful technical lemma. We then give proofs of Proposition 1, Theorems 1–5, and Proposition 4 in Sections S1-S7, respectively. We perform various simulation studies to verify the proposed theoretical conditions in Section S8.

Before presenting the proofs of the theoretical results, we introduce a useful notation, the T-matrix $T(Q,\mathbf{\Theta})$ of size $2^J \times 2^K$. The rows of $T(Q,\mathbf{\Theta})$ are indexed by the 2^J different response patterns $\boldsymbol{r} = (r_1,\ldots,r_J)^\top \in \{0,1\}^J$, and columns by attribute patterns $\boldsymbol{\alpha} \in \{0,1\}^K$, while the $(\boldsymbol{r},\boldsymbol{\alpha})$th entry of $T(Q,\mathbf{\Theta})$, denoted by $T_{\boldsymbol{r},\boldsymbol{\alpha}}(Q,\mathbf{\Theta})$, represents the marginal probability that subjects in latent class $\boldsymbol{\alpha}$ provide positive responses to the set of items $\{j : r_j = 1\}$, namely

$$T_{\boldsymbol{r},\boldsymbol{\alpha}}(Q,\mathbf{\Theta}) = P(\boldsymbol{R} \succeq \boldsymbol{r} \mid Q,\mathbf{\Theta},\boldsymbol{\alpha}) = \prod_{j=1}^{J} \theta_{j,\boldsymbol{\alpha}}^{r_j}.$$

We denote the $\boldsymbol{\alpha}$th column vector and the $\boldsymbol{r}$th row vector of the T-matrix by $T_{\bullet,\boldsymbol{\alpha}}(Q,\mathbf{\Theta})$

and $T_{\boldsymbol{r},\bullet}(Q,\boldsymbol{\Theta})$ respectively. Let $\boldsymbol{e}_j$ denote the J-dimensional unit vector with the jth element being one and all the other elements being zero, then any response pattern $\boldsymbol{r}$ can be written as a sum of some $\boldsymbol{e}$-vectors, namely $\boldsymbol{r}=\sum_{j:r_j=1}\boldsymbol{e}_j$. The $\boldsymbol{r}$th element of the 2^J-dimensional vector $T(Q,\boldsymbol{\Theta})\boldsymbol{p}$ is

$$T_{\boldsymbol{r},\bullet}(Q,\boldsymbol{\Theta})\boldsymbol{p}=\sum_{\boldsymbol{\alpha}\in\{0,1\}^K}T_{\boldsymbol{r},\boldsymbol{\alpha}}(Q,\boldsymbol{\Theta})p_{\boldsymbol{\alpha}}=P(\boldsymbol{R}\succeq\boldsymbol{r}\mid Q,\boldsymbol{\Theta},\boldsymbol{p}).$$

Based on the T-matrix, there is an equivalent definition of identifiability for $(Q,\boldsymbol{\Theta},\boldsymbol{p})$. The T-matrix also has a nice property that will be useful in proving the identifiability results. These are summarized in the following lemma, whose proof can be found in Xu (2017).

Lemma 1. *Under a restricted latent class model, $(Q,\boldsymbol{\Theta},\boldsymbol{p})$ are identifiable if and only if for any $(Q,\boldsymbol{\Theta},\boldsymbol{p})$ and $(\bar{Q},\bar{\boldsymbol{\Theta}},\bar{\boldsymbol{p}})$,*

$$T(Q,\boldsymbol{\Theta})\boldsymbol{p}=T(\bar{Q},\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}} \tag{S0.1}$$

implies $(Q,\boldsymbol{\Theta},\boldsymbol{p})=(\bar{Q},\bar{\boldsymbol{\Theta}},\bar{\boldsymbol{p}})$. For any $\boldsymbol{\theta}^*=(\theta_1,\ldots,\theta_J)^\top\in\mathbb{R}^J$, there exists an invertible matrix $D(\boldsymbol{\theta}^*)$ depending only on $\boldsymbol{\theta}^*$, such that

$$T(Q,\boldsymbol{\Theta}-\boldsymbol{\theta}^*\mathbf{1}^\top)=D(\boldsymbol{\theta}^*)T(Q,\boldsymbol{\Theta}). \tag{S0.2}$$

We introduce some additional notations. For a submatrix Q_1 of Q that has size $J_1\times K$, we denote the item parameter matrix corresponding to these J_1 items by $\boldsymbol{\Theta}_{Q_1}$, then $\boldsymbol{\Theta}_{Q_1}$ is a $J_1\times K$ submatrix of $\boldsymbol{\Theta}$. Denote Q_1's corresponding T-matrix by $T(Q_1, \mathbf{\Theta}_{Q_1})$, then $T(Q_1, \mathbf{\Theta}_{Q_1})$ has size $2^{J_1} \times 2^K$. For notational simplicity, in the following we denote $\boldsymbol{c} \equiv \mathbf{1} - \boldsymbol{s}$ under the DINA model, then $\mathbf{\Theta} = (\mathbf{1} - \boldsymbol{s}, \boldsymbol{g}) = (\boldsymbol{c}, \boldsymbol{g})$ under DINA.

We add some remarks on Lemma 1. First, Equation (S0.1) can be written as that, for any response pattern $\boldsymbol{r} \in \{0,1\}^J$, $T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q, \mathbf{\Theta})\boldsymbol{p} = T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q}, \bar{\mathbf{\Theta}})\bar{\boldsymbol{p}}$. Second, thanks to (S0.2), for any $\boldsymbol{\theta}^* = (\theta_1, \ldots, \theta_J)^\top \in \mathbb{R}^J$, equality (S0.1) leads to

$$T(Q, \mathbf{\Theta} - \boldsymbol{\theta}^* \mathbf{1}^\top)\boldsymbol{p} = T(\bar{Q}, \bar{\mathbf{\Theta}} - \boldsymbol{\theta}^* \mathbf{1}^\top)\bar{\boldsymbol{p}},$$

and further $T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q, \mathbf{\Theta} - \boldsymbol{\theta}^* \mathbf{1}^\top)\boldsymbol{p} = T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q}, \bar{\mathbf{\Theta}} - \boldsymbol{\theta}^* \mathbf{1}^\top)\bar{\boldsymbol{p}}$ for any $\boldsymbol{r} \in \{0,1\}^J$. Besides, If (S0.1) holds, then for any submatrix Q_1 of Q, equality $T(Q_1, \mathbf{\Theta}_{Q_1})\boldsymbol{p} = T(\bar{Q}_1, \bar{\mathbf{\Theta}}_{\bar{Q}_1})\bar{\boldsymbol{p}}$ also holds.

S1. Proof of Proposition 1

Consider a Q-matrix of size $J \times K$ in the form

$$Q = \begin{pmatrix} Q' \\ \mathbf{0} \end{pmatrix},$$

where Q' is of size $J' \times K$ and contains those nonzero $\boldsymbol{q}$-vectors of Q. For any item $j \in \{J'+1, \ldots, J\}$ which has $\boldsymbol{q}_j = \mathbf{0}$, all the attribute profiles $\boldsymbol{\alpha}$ satisfy $\boldsymbol{\alpha} \succeq \boldsymbol{q}_j$, so there is only one item parameter associated with j under Q, and we denote it by θ_j. Denote the first J' rows of $\mathbf{\Theta}$ by $\mathbf{\Theta}'$. Denote the $2^{J'} \times 2^K$ T-matrix associated with matrix Q' by $T'(Q', \mathbf{\Theta}')$.

First consider the case where $(Q', \mathbf{\Theta}', \boldsymbol{p})$ are strictly (or generically) identifiable, and we will show $(Q, \mathbf{\Theta}, \boldsymbol{p})$ are also strictly (or generically) identifiable. Assume there is a $J \times K$ matrix $\bar{Q}$ and associated parameters $(\bar{\mathbf{\Theta}}, \bar{\boldsymbol{p}})$ such that (S0.1) holds. Denote the submatrix of $\bar{Q}$ containing its first J' rows by $\bar{Q}'$, and the submatrix of $\bar{\mathbf{\Theta}}$ containing its first J' rows by $\bar{\mathbf{\Theta}}'$. Then (S0.1) implies $T(Q', \mathbf{\Theta}')\boldsymbol{p}' = T(\bar{Q}', \bar{\mathbf{\Theta}}')\bar{\boldsymbol{p}}'$, and the strict (or generic) joint identifiability of $(Q', \mathbf{\Theta}', \boldsymbol{p})$ gives that $\bar{Q}' \sim Q'$ and $(\bar{\mathbf{\Theta}}', \bar{\boldsymbol{p}}) = (\mathbf{\Theta}', \boldsymbol{p})$. For an arbitrary RLCM, the strict (or generic) identifiability of $(Q', \mathbf{\Theta}', \boldsymbol{p})$ implies that $T(Q', \mathbf{\Theta}')$ has full rank 2^K strictly (or generically). This is because if not so, then the proportion parameters $\boldsymbol{p}$ can not be strictly (or generically) identifiable, in the sense that there exist multiple different $\boldsymbol{p}$ such that $T(Q', \mathbf{\Theta}')\boldsymbol{p}$ are all equal. This would contradict the assumption that $(Q', \mathbf{\Theta}', \boldsymbol{p})$ are strictly (or generically) identifiable. Therefore $T(Q', \mathbf{\Theta}')$ is strictly (or generically) full-rank. Then for each $\boldsymbol{\alpha} \in \{0,1\}^K$ there must exist a 2^K-dimensional vector $\boldsymbol{v}_{\boldsymbol{\alpha}}$ such that

$$\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot} T(Q', \mathbf{\Theta}') = \boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot} T(\bar{Q}', \bar{\mathbf{\Theta}}') = (\mathbf{0}, \underbrace{x_{\boldsymbol{\alpha}}}_{\text{column } \boldsymbol{\alpha}}, \mathbf{0}), \quad x_{\boldsymbol{\alpha}} \neq 0,$$

and $\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot} T(Q', \mathbf{\Theta}')\boldsymbol{p} = \boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot} T(\bar{Q}', \bar{\mathbf{\Theta}}')\bar{\boldsymbol{p}} = x_{\boldsymbol{\alpha}} p_{\boldsymbol{\alpha}} \neq 0$. Then again use the property (S0.2) and we have the following equality for any $j \in \{J'+1, \dots, J\}$,

$$\begin{aligned} \theta_{j,\boldsymbol{\alpha}} =& \frac{\{T_{\boldsymbol{e}_j, \boldsymbol{\cdot}}(Q, \mathbf{\Theta}) \odot [\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot}\ T(Q', \mathbf{\Theta}')]\}\boldsymbol{p}}{\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot}\ T(Q', \mathbf{\Theta}')\boldsymbol{p}} \\ =& \frac{\{T_{\boldsymbol{e}_j, \boldsymbol{\cdot}}(Q, \mathbf{\Theta}) \odot [\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot}\ T(\bar{Q}', \bar{\mathbf{\Theta}}')]\}\bar{\boldsymbol{p}}}{\boldsymbol{v}_{\boldsymbol{\alpha}}^\top \boldsymbol{\cdot}\ T(\bar{Q}', \bar{\mathbf{\Theta}}')\bar{\boldsymbol{p}}} = \bar{\theta}_{j,\boldsymbol{\alpha}}, \end{aligned}$$

where "$\odot$" represents the element-wise product of two vectors. This proves $\mathbf{\Theta} = \bar{\mathbf{\Theta}}$ and

$Q \sim \bar{Q}$. So $(Q, \mathbf{\Theta}, \boldsymbol{p})$ are strictly (or generically) identifiable.

Next consider the case where $(Q', \mathbf{\Theta}', \boldsymbol{p})$ are **not** strictly (or generically) identifiable, so there exist $(\bar{Q}', \bar{\mathbf{\Theta}}', \bar{\boldsymbol{p}}) \nsim (Q', \mathbf{\Theta}', \boldsymbol{p})$ such that $T'(\bar{Q}', \bar{\mathbf{\Theta}}')\bar{\boldsymbol{p}} = T'(Q', \mathbf{\Theta}')\boldsymbol{p}$. Now extend $\bar{Q}'$ to $\bar{Q}$ of size $J \times K$ by adding $J - J'$ all-zero $\boldsymbol{q}$-vectors, i.e.,

$$\bar{Q} = \begin{pmatrix} \bar{Q}' \\ \\ \mathbf{0} \end{pmatrix},$$

and set $\bar{\theta}_j = \theta_j$ for $j \in \{J'+1, \ldots, J\}$. Then for any $\boldsymbol{r} = (r_1, \ldots, r_{J'}, r_{J'+1}, \ldots, r_J) \in \{0,1\}^J$ and the corresponding $\boldsymbol{r}' = (r_1, \ldots, r_{J'})$,

$$T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q, \mathbf{\Theta})\boldsymbol{p} = \Big\{T'_{\boldsymbol{r}',\boldsymbol{\cdot}}(Q', \mathbf{\Theta}')\boldsymbol{p}\Big\} \prod_{j>J'} \theta_j^{r_j};$$

$$T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q}, \bar{\mathbf{\Theta}})\bar{\boldsymbol{p}} = \Big\{T'_{\boldsymbol{r}',\boldsymbol{\cdot}}(\bar{Q}', \bar{\mathbf{\Theta}}')\boldsymbol{p}\Big\} \prod_{j>J'} \theta_j^{r_j}.$$

Now that $T(Q, \mathbf{\Theta})\boldsymbol{p} = T(\bar{Q}, \bar{\mathbf{\Theta}})\bar{\boldsymbol{p}}$ but $(\bar{Q}, \bar{\mathbf{\Theta}}, \bar{\boldsymbol{p}}) \nsim (Q, \mathbf{\Theta}, \boldsymbol{p})$, we obtain that $(Q, \mathbf{\Theta}, \boldsymbol{p})$ are not strictly (or generically) identifiable. The proof of the proposition is complete.

S2. Proof of Theorem 1

We first prove the sufficiency, and then show the necessity of the conditions. Under DINA, (S0.1) can be equivalently written as that for any $\boldsymbol{r} \in \{0,1\}^J$,

$$T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q, \boldsymbol{c}, \boldsymbol{g})\boldsymbol{p} = T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}}, \bar{\boldsymbol{g}})\bar{\boldsymbol{p}}. \tag{S2.3}$$

We first introduce some notations. In the following discussion, for an integer M, we denote $[M] = \{1, \ldots, M\}$. For an item set $S \subseteq [J]$, denote $\boldsymbol{q}_S = \vee_{j \in S} \boldsymbol{q}_j = (\max_{j\in S} q_{j,1},$

$\max_{j\in S} q_{j,2},\ldots,\max_{j\in S} q_{j,K})$, then $\boldsymbol{q}_S$ is also a K-dimensional binary vector, and we denote its k element by $q_{S,k}$. Recall

$$Q = \begin{pmatrix} I_K \\ Q^\star \end{pmatrix},$$

and we denote the submatrix of $\bar{Q}$ consisting of its first K row vectors by $\bar{Q}_{1:K,\bullet}$. We next show in five steps that if (S2.3) holds, then $\bar{Q} \sim Q$, and also $\boldsymbol{c} = \bar{\boldsymbol{c}}$, $\bar{\boldsymbol{g}} = \boldsymbol{g}$, $\bar{\boldsymbol{p}} = \boldsymbol{p}$.

Step 1. After some column rearrangement, $\bar{Q}_{1:K,\bullet}$ is an upper-triangular matrix with all the diagonal elements being ones.

Step 2. $\bar{c}_j = c_j$ for all $j \in \{K+1,\ldots,J\}$.

Step 3. $\bar{g}_k = g_k$ for all $k \in \{1,\ldots,K\}$.

Step 4. $\bar{Q}_{1:K,\bullet} \sim I_K$

Step 5. $\bar{Q} \sim Q$, $\boldsymbol{c} = \bar{\boldsymbol{c}}$, $\bar{\boldsymbol{g}} = \boldsymbol{g}$, $\bar{\boldsymbol{p}} = \boldsymbol{p}$.

For any item set $S \subseteq \{1,\ldots,J\}$, denote $\boldsymbol{c}_S = \sum_{j\in S} c_j \boldsymbol{e}_j$, and denote $\boldsymbol{g}_S$, $\bar{\boldsymbol{c}}_S$, and $\bar{\boldsymbol{g}}_S$ similarly. Consider the response pattern $\boldsymbol{r}^\star = \sum_{j\in S} \boldsymbol{e}_j$ and any $\boldsymbol{\theta}^\star = \sum_{j\in S} \theta_j^\star \boldsymbol{e}_j$, then Equation (S2.3) together with Lemma 1 imply that

$$T_{\boldsymbol{r}^\star,\bullet}(Q, \boldsymbol{c}_S - \boldsymbol{\theta}^\star, \boldsymbol{g}_S - \boldsymbol{\theta}^\star)\boldsymbol{p} = T_{\boldsymbol{r}^\star,\bullet}(\bar{Q}, \bar{\boldsymbol{c}}_S - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}}_S - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}}. \tag{S2.4}$$

We will frequently use (S2.4) in the following proof. And when the item set S and response pattern $\boldsymbol{r}^\star$ are clearly implied by the definition of $\boldsymbol{\theta}^\star$, we will omit the subscript S in the above (S2.4). We also frequently use the fact that when (S2.4) holds, $c_j \neq \bar{g}_j$

and $g_j \neq \bar{c}_j$ for any item j. This is true because if $c_j = \bar{g}_j$, we would have

$$T_{\boldsymbol{e}_j,\bullet}(Q,\boldsymbol{c},\boldsymbol{g})\boldsymbol{p} = c_j(\sum_{\boldsymbol{\alpha}\succeq\boldsymbol{q}_j} p_{\boldsymbol{\alpha}}) + g_j(\sum_{\boldsymbol{\alpha}\nsucceq\boldsymbol{q}_j} p_{\boldsymbol{\alpha}}) < c_j = \bar{g}_j$$
$$\leq \bar{c}_j(\sum_{\boldsymbol{\alpha}\succeq\boldsymbol{q}_j} \bar{p}_{\boldsymbol{\alpha}}) + \bar{g}_j(\sum_{\boldsymbol{\alpha}\nsucceq\boldsymbol{q}_j} \bar{p}_{\boldsymbol{\alpha}}) = T_{\boldsymbol{e}_j,\bullet}(\bar{Q},\bar{\boldsymbol{c}},\bar{\boldsymbol{g}})\bar{\boldsymbol{p}},$$

which contradicts (S2.3). So $c_j \neq \bar{g}_j$ and similarly $g_j \neq \bar{c}_j$ for each j. As stated in the main text, we assume without loss of generality that there is no all-zero row vector in true Q-matrix. If, however, the jth row vector of $\bar{Q}$ equals $\mathbf{0}$, then $\bar{c}_j$ would equal $\bar{g}_j$, and we denote this value by $\bar{\theta}_j$. Equation (S2.3) gives

$$\bar{\theta}_j = c_j\Big(\sum_{\boldsymbol{\alpha}:\,\boldsymbol{\alpha}\succeq\boldsymbol{q}_j} p_{\boldsymbol{\alpha}}\Big) + g_j\Big(\sum_{\boldsymbol{\alpha}:\,\boldsymbol{\alpha}\nsucceq\boldsymbol{q}_j} p_{\boldsymbol{\alpha}}\Big),$$

and hence $g_j < \bar{\theta}_j < c_j$ holds for this j.

Step 1. In this step we prove that $\bar{Q}_{1:K,\bullet}$ must take the following form after some column rearrangement,

$$\bar{Q}_{1:K,\bullet} \sim \begin{pmatrix} 1 & * & \dots & * \\ 0 & 1 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}. \tag{S2.5}$$

Namely, after properly rearranging the columns of $\bar{Q}_{1:K,\bullet}$, we have $\bar{Q}_{k,k} = 1$ and $\bar{Q}_{k,h} = 0$ for any $k > h$.

We first introduce the following useful lemma.

Lemma 2. *Suppose the true Q satisfies Condition A that $Q_{1:K} = I_K$. If there exists an item set $S \subseteq \{K+1, \ldots, J\}$ such that*

$$\max_{m\in S} q_{m,h} = 0, \quad \max_{m\in S} q_{m,j} = 1 \;\; \forall j \in \mathcal{J}$$

for some attributes $h \in [K]$ and a set of attributes $\mathcal{J} \subseteq [K] \setminus \{h\}$, then

$$\vee_{j\in\mathcal{J}}\, \bar{\boldsymbol{q}}_j \not\succeq \bar{\boldsymbol{q}}_h.$$

Proof of Lemma 2. We prove by contradiction. Assume there exist attribute $h \in [K]$ and a set of attributes $\mathcal{J} \subseteq [K] \setminus \{h\}$, such that $\vee_{j\in\mathcal{J}}\, \bar{\boldsymbol{q}}_j \not\succeq \bar{\boldsymbol{q}}_h$; and that there exists $S \subseteq \{K+1, \ldots, J\}$ such that $\max_{m\in S} q_{m,h} = 0$ and $\max_{m\in S} q_{m,j} = 1$. Define

$$\boldsymbol{\theta}^\star = \bar{c}_h \boldsymbol{e}_h + \sum_{j\in\mathcal{J}} \bar{g}_j \boldsymbol{e}_j + \sum_{m=K+1}^{J} g_m \boldsymbol{e}_m, \quad \boldsymbol{r}^\star = \boldsymbol{e}_h + \sum_{j\in\mathcal{J}} \boldsymbol{e}_j + \sum_{m=K+1}^{J} \boldsymbol{e}_m,$$

and we claim that $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)$ is an all-zero vector. This is because for any $\boldsymbol{\alpha} \in \{0,1\}^K$, the corresponding element in $T_{\boldsymbol{r}^\star,\boldsymbol{\alpha}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)$ contains a factor $F_{\boldsymbol{\alpha}} = (\bar{\theta}_{h,\boldsymbol{\alpha}} - \bar{c}_h)\prod_{j\in\mathcal{J}}(\bar{\theta}_{j,\boldsymbol{\alpha}} - \bar{g}_j)$. While this factor $F_{\boldsymbol{\alpha}} \neq 0$ only if $\bar{\theta}_{h,\boldsymbol{\alpha}} = \bar{g}_h$ and $\bar{\theta}_{j,\boldsymbol{\alpha}} = \bar{c}_j$ for all $j \in \mathcal{J}$, which happens if and only if $\boldsymbol{\alpha} \not\succeq \bar{\boldsymbol{q}}_h$ and $\boldsymbol{\alpha} \succeq \bar{\boldsymbol{q}}_j$ for all $j \in \mathcal{J}$, which is impossible because $\vee_{j\in\mathcal{J}}\bar{\boldsymbol{q}}_j \succeq \bar{\boldsymbol{q}}_h$ by our assumption. So the claim $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star) = \boldsymbol{0}$ is proved, and further $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = 0$. Equality (S2.4) becomes $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^\star, \boldsymbol{g} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = 0$, which leads to

$$0 = T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^\star, \boldsymbol{g} - \boldsymbol{\theta}^\star)\boldsymbol{p} = p_{\boldsymbol{1}}(c_h - \bar{c}_h)\prod_{j\in\mathcal{J}}(c_j - \bar{g}_j)\prod_{m>K}(c_m - g_m),$$

which is because for any $\boldsymbol{\alpha} \neq \mathbf{1}$, we must have $\boldsymbol{\alpha} \nsucceq \boldsymbol{q}_m$ for some $m > K$ under Condition C, and hence the element $T_{\boldsymbol{r}^\star,\boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^\star, \boldsymbol{g}-\boldsymbol{\theta}^\star)$ contains a factor $(g_m - g_m) = 0$. Since $c_m - g_m > 0$ for $m > K$ and $c_j - \bar{g}_j \neq 0$, we obtain $c_h = \bar{c}_h$.

We remark here that $c_h = \bar{c}_h$ also implies $\bar{\boldsymbol{q}}_h \neq \mathbf{0}$, because otherwise we would have $\bar{\theta}_h = \bar{c}_h = c_h$, which contradicts the $g_h < \bar{\theta}_h < c_h$ proved before the current Step 1. This indicates the $\bar{Q}_{1:K,\boldsymbol{\cdot}}$ can not contain any all-zero row vector, because otherwise $\bar{\boldsymbol{q}}_j \succeq \bar{\boldsymbol{q}}_h$ for the all-zero row vector $\bar{\boldsymbol{q}}_h$, which we showed is impossible.

Consider the item set S in the lemma that satisfies $S \subseteq \{K+1, \ldots, J\}$ such that $\max_{m\in S} q_{m,h} = 0$ and $\max_{m\in S} q_{m,j} = 1$ for all $j \in \mathcal{J}$. Define

$$\boldsymbol{\theta}^\star = \bar{c}_h \boldsymbol{e}_h + \sum_{j\in\mathcal{J}} \bar{g}_j \boldsymbol{e}_j + \sum_{m\in S} g_m \boldsymbol{e}_m.$$

Note that $c_h = \bar{c}_h$. The RHS of (S2.4) is zero, and so is the LHS of it. The row vector $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^\star, \boldsymbol{g}-\boldsymbol{\theta}^\star)$ has the following property

$$\begin{aligned}
&T_{\boldsymbol{r}^\star,\boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^\star, \boldsymbol{g}-\boldsymbol{\theta}^\star) \\
&= \begin{cases} (g_h - \bar{c}_h)\prod_{j\in\mathcal{J}}(c_j - \bar{g}_j)\prod_{m\in S}(c_m - g_m), & \boldsymbol{\alpha} \nsucceq \boldsymbol{q}_h,\ \boldsymbol{\alpha} \succeq \boldsymbol{q}_{\mathcal{J}},\ \boldsymbol{\alpha} \succeq \boldsymbol{q}_S; \\ 0, & \text{otherwise.} \end{cases}
\end{aligned}$$

An important observation is that $\{\boldsymbol{\alpha} \in \{0,1\}^K : \boldsymbol{\alpha} \nsucceq \boldsymbol{q}_h,\ \boldsymbol{\alpha} \succeq \boldsymbol{q}_{\mathcal{J}},\ \boldsymbol{\alpha} \succeq \boldsymbol{q}_S\} = \mathcal{A} \neq \varnothing$. This is because $q_{S,h} = 0$ and $q_{S,j} = 1$ for all $j \in \mathcal{J}$ hold, and we can just choose $\boldsymbol{\alpha}$ for which $\alpha_h = 0$ and $\alpha_k = 1$ for all $q_{S,k} = 1$, then such $\boldsymbol{\alpha}$ belongs to the set $\mathcal{A}$. Therefore we have

$$
\begin{aligned}
&T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q,\boldsymbol{c}-\boldsymbol{\theta}^\star,\boldsymbol{g}-\boldsymbol{\theta}^\star)\boldsymbol{p} \\
&=(g_h-\bar{c}_h)\prod_{j\in\mathcal{J}}(c_j-\bar{g}_j)\prod_{m\in S}(c_m-g_m)\Big(\sum_{\boldsymbol{\alpha}\in\mathcal{A}}p_{\boldsymbol{\alpha}}\Big)=0,
\end{aligned}
$$

which leads to a contradiction since $g_h-\bar{c}_h\neq 0$, $c_j-\bar{g}_j\neq 0$, $c_m-g_m\neq 0$ and $\sum_{\boldsymbol{\alpha}\in\mathcal{A}}p_{\boldsymbol{\alpha}}>0$, i.e., every factor in the above product is nonzero. This completes the proof of Lemma 2. $\square$

We now proceed with the proof of Step 1 using an induction argument. We first introduce the definition of *lexicographic order* between two binary vectors of the same length. Specifically, for two binary vectors $\boldsymbol{a}=(a_1,\ldots,a_L)^\top$ and $\boldsymbol{b}=(b_1,\ldots,b_L)^\top$ both of length L, we say $\boldsymbol{a}$ is of smaller lexicographic order than $\boldsymbol{b}$ and denote $\boldsymbol{a}\prec_{\text{lex}}\boldsymbol{b}$, if either $a_1<b_1$, or there exists a integer $l\in\{2,\ldots,L\}$ such that $a_l<b_l$ and $a_m=b_m$ for all $m=1,\ldots,l-1$. It is not hard to see when Condition B that $Q^\star$ contains K distinct column vectors is satisfied, the K columns of $Q^\star$ can be arranged in an increasing lexicographic order. Namely, under Condition B, there exists a permutation map $\sigma(\cdot):[K]\to[K]$ such that

$$
Q^\star_{,\sigma(1)}\prec_{\text{lex}}Q^\star_{,\sigma(2)}\prec_{\text{lex}}\cdots\prec_{\text{lex}}Q^\star_{,\sigma(K)}. \tag{S2.6}
$$

Without loss of generality, next we consider the case where $\sigma(\cdot)$ is the identity map, i.e., $\sigma(k)=k$ for all $k\in[K]$.

We first consider attribute 1. Since $Q^\star_{,1}$ has the smallest lexicographic order among the columns of $Q^\star$, we have the conclusion that there must exist an item set $S \subseteq \{K+1,\ldots,J\}$ such that

$$q_{S,1}=0, \quad q_{S,\ell}=1 \ \ \forall \ell=2,\ldots,K.$$

We apply Lemma 2 to obtain $\vee_{\ell\in\{2,\ldots,K\}}\bar{\boldsymbol{q}}_\ell \nsucceq \bar{\boldsymbol{q}}_1$, which means

$$\begin{aligned}&(\max_{m\in\{2,\ldots,K\}} \bar{q}_{\ell,1}, \max_{m\in\{2,\ldots,K\}} \bar{q}_{\ell,2},\ldots, \max_{m\in\{2,\ldots,K\}} \bar{q}_{\ell,K})\\ \nsucceq\ &(\bar{q}_{1,1},\ \ldots,\ \bar{q}_{1,K}).\end{aligned}$$

This implies there must exist an attribute $m_1\in[K]$ such that

$$\max_{k\in[K]\setminus\{1\}} \bar{q}_{k,m_1}=0, \quad \bar{q}_{1,m_1}=1, \tag{S2.7}$$

which exactly says the m_1-th column vector of $\bar{Q}_{1:K,\bullet}$ must equal the basis vector $(\underbrace{1}_{\text{column 1}},\mathbf{0})^\top=\boldsymbol{e}_1$, i.e., we have $\bar{Q}_{1:K,m_1}=\boldsymbol{e}_1$.

Now we assume as the inductive hypothesis that for $h\in[K]$ and $h>1$, we have a distinct set of attributes $\{m_1,\ldots,m_{h-1}\}\subseteq[K]$ such that their corresponding column vectors in $\bar{Q}_{1:K,\bullet}$ satisfy

$$\forall i=1,\ldots,h-1, \quad \bar{Q}_{1:K,m_i}=(*,\ldots,*,\underbrace{1}_{\text{column } i},0,\ldots,0)^\top. \tag{S2.8}$$

Now we focus on attribute h. By (S2.6), the column vector $Q^\star_{,h}$ has the smallest lexicographic order among the $K-h-1$ columns in $\{Q^\star_{\bullet,h}, Q^\star_{\bullet,h+1},\ldots,Q^\star_{\bullet,K}\}$, therefore similar to the argument in the previous paragraph, there must exist an item set

$S \subseteq \{K+1, \ldots, J\}$ such that

$$q_{S,h} = 0, \quad q_{S,\ell} = 1 \;\; \forall \ell = h+1, \ldots, K. \tag{S2.9}$$

Therefore Lemma 2 implies $\vee_{\ell \in \{h+1,\ldots,K\}} \bar{\boldsymbol{q}}_\ell \nsucceq \bar{\boldsymbol{q}}_1$, and further leads to

$$\max_{\ell \in \{h+1,\ldots,K\}} \bar{q}_{\ell,m_h} = 0, \quad \bar{q}_{h,m_h} = 1. \tag{S2.10}$$

We point out that $m_h \notin \{m_1, \ldots, m_{h-1}\}$, because by the induction hypothesis (S2.8) we have $\bar{q}_{h,m_i} = 0$ for $i = 1, \ldots, h-1$. So $\{m_1, \ldots, m_{h-1}, m_h\}$ contains h distinct attributes. Furthermore, (S2.10) gives that $\bar{Q}_{\bullet,m_h} = (*, \ldots, *, \underbrace{1}_{\text{column } h}, 0, \ldots, 0)^\top$, which generalizes (S2.8) by extending $h-1$ there to h. Therefore, we use the induction argument to obtain

$$\forall k \in \{1, \ldots, K-1\}, \quad \bar{Q}_{1:K,m_k} = (*, \ldots, *, \underbrace{1}_{\text{column } k}, 0, \ldots, 0)^\top.$$

Furthermore, when considering the last attribute K, the Kth item must have $\boldsymbol{q}$-vector taking the form of $\bar{\boldsymbol{q}}_K = (0, \ldots, 0, \underbrace{*}_{\text{column } m_K}, 0, \ldots, 0)$, where the "$*$" in $\bar{\boldsymbol{q}}_K$ is the only element unspecified. Since previously we have shown in the proof of Lemma 2 that $\bar{\boldsymbol{q}}_j = \mathbf{0}$ can not happen for any item j, there must be $\bar{\boldsymbol{q}}_K = (0, \ldots, 0, \underbrace{1}_{\text{column } m_K}, 0, \ldots, 0)$. Now we have essentially obtained

$$\bar{Q}_{1:K,\,(m_1,\ldots,m_K)} = \begin{pmatrix} 1 & * & \ldots & * \\ 0 & 1 & \ldots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & 1 \end{pmatrix}, \tag{S2.11}$$

and the conclusion of Step 1 in (S2.5) is proved.

Step 2. In this step we prove $c_j = \bar{c}_j$ for $j = K+1, \ldots, J$. For an arbitrary item $j \in \{K+1, \ldots, J\}$, define a response vector $\boldsymbol{r}^* = \sum_{h:\, h\neq j} \boldsymbol{e}_j$ and

$$\boldsymbol{\theta}^* = \sum_{h=1}^{K} \bar{g}_h \boldsymbol{e}_h + \sum_{h>K,\, h\neq j} g_h \boldsymbol{e}_h.$$

We claim that $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^*, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^*)$ contains only one nonzero element corresponding to the all-one attribute pattern $\boldsymbol{\alpha} = \mathbf{1}$. The reasoning is as follows. Under the conclusion of Step 1, $\bar{Q}_{1:K,\boldsymbol{\cdot}}$ takes the form of (S2.5), which means each attribute is required by at least one item in $\{\bar{\boldsymbol{q}}_1, \ldots, \bar{\boldsymbol{q}}_K\}$. Then for any $\boldsymbol{\alpha} \neq \mathbf{1}$, there must exist some attribute $k \in [K]$ such that $\boldsymbol{\alpha} \not\succeq \bar{\boldsymbol{q}}_k$, which implies for this particular $\boldsymbol{\alpha}$ the element $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^*, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^*)$ contains a factor $(\bar{g}_h - \bar{g}_h) = 0$. Therefore $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) \neq 0$ only if $\boldsymbol{\alpha} = \mathbf{1}$. Next consider $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)$. Under Condition A, in the true Q each attribute is required by at least three items, so the row vector corresponding to response pattern $\boldsymbol{r}^*$ in $T(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)$ only contains one nonzero element, in column $\boldsymbol{\alpha} = \mathbf{1}_K^\top$, representing the attribute profile mastering all the K attributes. This is because for any other attribute profile $\boldsymbol{\alpha}'$ that lacks at least one attribute k, there must be some item $h > K$, $h \neq j$ requiring attribute k so that $\boldsymbol{\alpha}' \not\succeq \boldsymbol{q}_h$; and this results in $\theta_{\boldsymbol{e}_h,\boldsymbol{\alpha}'} = g_h$ and $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}'}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = 0$. In summary,

$$T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = \prod_{h=1}^{K} (\theta_{h,\boldsymbol{\alpha}} - \bar{g}_h) \prod_{\substack{h>K:\\ h\neq j}} (\theta_{h,\boldsymbol{\alpha}} - g_h) \neq 0 \quad \text{iff} \quad \boldsymbol{\alpha} = \mathbf{1};$$

$$T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) = \prod_{h=1}^{K}(\bar{\theta}_{h,\boldsymbol{\alpha}} - \bar{g}_h) \prod_{\substack{h>K:\\ h\neq j}} (\bar{\theta}_{h,\boldsymbol{\alpha}} - g_h) \neq 0 \quad \text{iff} \quad \boldsymbol{\alpha} = \mathbf{1}.$$

Now further consider item j. Since $\mathbf{1}_K^\top \succeq \boldsymbol{q}_j$ and $\mathbf{1}_K^\top \succeq \bar{\boldsymbol{q}}_j$, one must have $\theta_{j,\mathbf{1}_K^\top} = c_j$ and $\bar{\theta}_{j,\mathbf{1}_K^\top} = \bar{c}_j$. Since we assume $p_{\boldsymbol{\alpha}} > 0$ for each $\boldsymbol{\alpha}$, we have $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p} = T_{\boldsymbol{r}^*,\mathbf{1}_K^\top}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)p_{\mathbf{1}_K^\top} \neq 0$. So (S0.2) in Lemma 1 implies that

$$c_j = \frac{T_{\boldsymbol{r}^*+\boldsymbol{e}_j,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p}}{T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p}} = \frac{T_{\boldsymbol{r}^*+\boldsymbol{e}_j,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}}}{T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}}} = \bar{c}_j.$$

In the above argument j is arbitrary, so $c_j = \bar{c}_j$ for any $j = K+1, \ldots, J$.

Step 3. In this step we prove $g_k = \bar{g}_k$ for $k = 1, \ldots, K$. Recall that in Step 1 we showed that (S2.6) about the lexicographic order holds and assumed $\sigma(k) = k$ for $k \in [K]$ without loss of generality. We now prove $g_1 = \bar{g}_1$. Define

$$\boldsymbol{\theta}^* = \sum_{h=1}^{K} \bar{g}_h \boldsymbol{e}_h + \sum_{\substack{h>K:\\ q_{h,1}=0}} g_h \boldsymbol{e}_h + \sum_{\substack{h>K:\\ q_{h,1}=1}} c_h \boldsymbol{e}_h, \tag{S2.12}$$

then

$$T_{\sum_h \boldsymbol{e}_h,\boldsymbol{\alpha}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = \prod_{h=1}^{K}(\theta_{h,\boldsymbol{\alpha}} - \bar{g}_h) \prod_{\substack{h>K:\\ q_{h,1}=0}} (\theta_{h,\boldsymbol{\alpha}} - g_h) \prod_{\substack{h>K:\\ q_{h,1}=1}} (\theta_{h,\boldsymbol{\alpha}} - c_h);$$

$$T_{\sum_h \boldsymbol{e}_h,\boldsymbol{\alpha}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) = \prod_{h=1}^{K}(\bar{\theta}_{h,\boldsymbol{\alpha}} - \bar{g}_h) \prod_{\substack{h>K:\\ q_{h,1}=0}} (\bar{\theta}_{h,\boldsymbol{\alpha}} - g_h) \prod_{\substack{h>K:\\ q_{h,1}=1}} (\bar{\theta}_{h,\boldsymbol{\alpha}} - c_h).$$

First, the row vector $T_{\sum_{h=1}^J \boldsymbol{e}_h,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)$ equals the zero vector. This is because $\bar{Q}_{1:K,\boldsymbol{\cdot}}$ takes the form in (S2.5) by Step 1, and any attribute profile $\boldsymbol{\alpha} \neq \mathbf{1}_K^\top$ would have $\bar{\theta}_{h,\boldsymbol{\alpha}} = \bar{g}_h$ for some $h \in \{1, \ldots, K\}$, which makes the corresponding element in the above row vector zero. Furthermore, $T_{\sum_{h=1}^J \boldsymbol{e}_h,\mathbf{1}_K^\top}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)$ is also zero, because

$\bar{\theta}_{h,\boldsymbol{\alpha}} = \bar{c}_h = c_h$ for those $h > K$ such that $q_{h,1} = 1$. Since $Q^\star_{\bullet,1}$ has the smallest lexicographic order among the columns of $Q^\star$, for any $k \in \{2, \ldots, K\}$, there must exist some item $h \in \{K+1, \ldots, J\}$ that requires attribute 1, as a result

$$\vee_{h>K:\, q_{h,1}=0}\, \boldsymbol{q}_h = (0, 1, \ldots, 1).$$

This ensures $T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)$ would equal zero if $\boldsymbol{\alpha}$ lacks any attribute other than the first one. So the nonzero elements in the row vector $T_{\sum_{h=1}^J \boldsymbol{e}_h, \bullet}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)$ can only correspond to columns $\boldsymbol{\alpha}^1 = (0, 1, \ldots, 1)$ or $\boldsymbol{\alpha}^2 = \mathbf{1}_K^\top$. Further, we claim $T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}^2}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = 0$, this is because $\theta_{h,\boldsymbol{\alpha}} = c_h$ for those h such that $q_{h,1} = 1$. So the row vector $T_{\sum_{h=1}^J \boldsymbol{e}_h, \bullet}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)$ only contains one potentially nonzero element in column $\boldsymbol{\alpha}_1 = (0, 1, \ldots, 1)$ as follows

$$T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}_1}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*) = (g_1 - \bar{g}_1) \prod_{h=2}^{K} (c_h - \bar{g}_h) \prod_{\substack{h>K:\\ q_{h,1}=0}} (c_h - g_h) \prod_{\substack{h>K:\\ q_{h,1}=1}} (g_h - c_h). \tag{S2.13}$$

Using the fact $T_{\sum_{h=1}^J \boldsymbol{e}_h, \bullet}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) = \mathbf{0}_{2^K}$, the equality

$$T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}^1}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p} = T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}^1}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}} = 0$$

implies the element in (S2.13) must also be zero. As shown earlier, $c_h - \bar{g}_h \neq 0$ for any h, so $g_1 = \bar{g}_1$ must hold.

Next we use an induction argument to prove that for $k = 2, \ldots, K$, $g_k = \bar{g}_k$. In particular, suppose for any $1 \leq m \leq k-1$, we already have $g_m = \bar{g}_m$. Define

$$\boldsymbol{\theta}^* = \sum_{h=1}^{K} \bar{g}_h \boldsymbol{e}_h + \sum_{h>K:\, q_{h,k}=0} g_h \boldsymbol{e}_h + \sum_{h>K:\, q_{h,k}=1} c_h \boldsymbol{e}_h. \tag{S2.14}$$

For the similar reason as stated before, $T_{\sum_{h=1}^J \boldsymbol{e}_h, \bullet}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)$ equals the zero vector. We claim that the row vector $T_{\sum_{h=1}^J \boldsymbol{e}_h, \bullet}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)$ only contains one potentially nonzero element in column $\boldsymbol{\alpha}' := (1, \dots, 1, \underbrace{0}_{\text{column } k}, 1, \dots, 1)$. The reason is as follows. On the one hand, for any attribute profile $\boldsymbol{\alpha}$ that lacks some attribute $l \in \{k+1, \dots, K\}$, due to the assumption in (S2.6) that $Q^*_{\bullet,k} \prec_{\text{lex}} Q^*_{\bullet,l}$, there must exist some item $h > K$ such that $q_{h,k} = 0$, $q_{h,l} = 1$. So for this particular $\boldsymbol{\alpha}$ we have $\boldsymbol{\alpha} \nsucceq \boldsymbol{q}_h$, $\theta_{h,\boldsymbol{\alpha}} = g_h$, which makes $T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = 0$. On the other hand, for any attribute profile $\boldsymbol{\alpha}'$ that lacks some attribute $m \in \{1, \dots, k-1\}$, one has $\boldsymbol{\alpha}' \nsucceq \boldsymbol{q}_m = \boldsymbol{e}_m$ and $\theta_{m,\boldsymbol{\alpha}'} = g_m = \bar{g}_m$, where the last equality $g_m = \bar{g}_m$ comes from the induction assumption. This results in $T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}'}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = 0$ for all such $\boldsymbol{\alpha}'$. In conclusion, the nonzero elements in this transformed row vector can only be in columns $\boldsymbol{\alpha}'$ or $\boldsymbol{\alpha}_2 = \mathbf{1}_K^\top$. For similar reason as in proving $g_1 = \bar{g}_1$, $T_{\sum_{h=1}^J \boldsymbol{e}_h, \boldsymbol{\alpha}_2}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) = 0$. So the transformed row vector only contains one potentially nonzero entry corresponding to $\boldsymbol{\alpha}'$:

$$
\begin{aligned}
&T_{\sum_h \boldsymbol{e}_h, \boldsymbol{\alpha}'}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*) \\
&= (g_k - \bar{g}_k) \prod_{\substack{1 \le h \le K: \\ h \neq k}} (c_h - \bar{g}_h) \prod_{\substack{h > K: \\ q_{h,k} = 0}} (c_h - g_h) \prod_{\substack{h > K: \\ q_{h,k} = 1}} (g_h - c_h).
\end{aligned}
$$

The same argument after (S2.13) gives $g_k = \bar{g}_k$. In conclusion, the induction method yields $g_k = \bar{g}_k$ for $k = 1, \dots, K$.

Step 4. In this step we show that $\bar{Q}_{1:K,\bullet} \sim I_K$. Recall that in Step 1 we already obtained (S2.11), and now we aim to show that the $\bar{Q}_{1:K,(m_1,\dots,m_K)}$ in (S2.11) can be further written as

$$\bar{Q}_{1:K,(m_1,\dots,m_K)} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

We now claim that $\bar{\boldsymbol{q}}_j \nsucceq \bar{\boldsymbol{q}}_h$ for any $1 \le j < h \le K$. If this claim is true, then $\bar{Q}_{1:K,(m_1,\dots,m_K)} = I_K$ must hold and the conclusion $\bar{Q}_{1:K,\boldsymbol{\cdot}} \sim I_K$ is reached. We next prove that claim by contradiction. If there exist some $1 \le j < h \le K$ such that $\bar{\boldsymbol{q}}_j \succeq \bar{\boldsymbol{q}}_h$, then define

$$\boldsymbol{\theta}^\star = \bar{c}_h \boldsymbol{e}_h + \bar{g}_j \boldsymbol{e}_j + \sum_{m=K+1}^{J} g_m \boldsymbol{e}_m,$$

we have

$$\begin{aligned} 0 &= T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} \\ &= T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^\star, \boldsymbol{g} - \boldsymbol{\theta}^\star)\boldsymbol{p} \\ &= p_{\boldsymbol{1}}(c_h - \bar{c}_h)(c_j - \bar{g}_j) \prod_{m=K+1}^{J} (c_m - g_m), \end{aligned}$$

which implies $c_h = \bar{c}_h$. Note that we have obtained $g_j = \bar{g}_j$ in Step 3, and we next define $\boldsymbol{\theta}^\star = \bar{c}_h \boldsymbol{e}_h + \bar{g}_j \boldsymbol{e}_j$. The equality $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = 0$ still holds and (S2.4) gives

$$0 = T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^\star, \boldsymbol{g} - \boldsymbol{\theta}^\star)\boldsymbol{p} \tag{S2.15}$$

$$= (g_h - \bar{c}_h)(c_j - \bar{g}_j)\Big(\sum_{\boldsymbol{\alpha}:\, \boldsymbol{\alpha}\not\succeq \boldsymbol{q}_h, \boldsymbol{\alpha}\succeq \boldsymbol{q}_j} p_{\boldsymbol{\alpha}}\Big)$$

$$= (g_h - c_h)(c_j - g_j)\Big(\sum_{\boldsymbol{\alpha}:\, \boldsymbol{\alpha}\not\succeq \boldsymbol{q}_h, \boldsymbol{\alpha}\succeq \boldsymbol{q}_j} p_{\boldsymbol{\alpha}}\Big).$$

Since $Q_{1:K,\bullet} = I_K$, we have that $\boldsymbol{q}_j$ and $\boldsymbol{q}_h$ in the true Q are distinct basis vectors, therefore $\Big(\sum_{\boldsymbol{\alpha}:\, \boldsymbol{\alpha}\not\succeq \boldsymbol{q}_h, \boldsymbol{\alpha}\succeq \boldsymbol{q}_j} p_{\boldsymbol{\alpha}}\Big) > 0$. Therefore (S2.15) leads to a contradiction, and we have proved the claim that $\bar{\boldsymbol{q}}_j \not\succeq \bar{\boldsymbol{q}}_h$ for any $1 \le j < h \le K$. As stated earlier, this claim naturally leads to the conclusion of Step 3 that $\bar{Q}_{1:K,\bullet} \sim I_K$.

Step 5. In this step we prove that after reordering the columns in $\bar{Q}$ such that $\bar{Q}_{1:K} = I_K$, we must have $\boldsymbol{q}_j = \bar{\boldsymbol{q}}_j$ for $j = K+1, \ldots, J$. In the following two parts, we first prove $\bar{\boldsymbol{q}}_j \succeq \boldsymbol{q}_j$ for all $j \in \{K+1, \ldots, J\}$ in part (a); and then prove $\bar{\boldsymbol{q}}_j = \boldsymbol{q}_j$ for all $j \in \{K+1, \ldots, J\}$ in part (b).

(a) We next show $\bar{\boldsymbol{q}}_j \succeq \boldsymbol{q}_j$ for all $j \in \{K+1, \ldots, J\}$. We use proof by contradiction, and assume $\bar{\boldsymbol{q}}_j \not\succeq \boldsymbol{q}_j$ for some $j \in \{K+1, \ldots, J\}$. Then $\{\boldsymbol{\alpha} : \boldsymbol{\alpha} \succeq \bar{\boldsymbol{q}}_j, \boldsymbol{\alpha} \not\succeq \boldsymbol{q}_j\} = \mathcal{A} \neq \varnothing$ and $\sum_{\boldsymbol{\alpha}\in\mathcal{A}} p_{\boldsymbol{\alpha}} \neq 0$. Define

$$\boldsymbol{\theta}^* = \sum_{k\in[K]:\, \bar{q}_{j,k}=1} g_k \boldsymbol{e}_k + c_j \boldsymbol{e}_j, \tag{S2.16}$$

then $T_{\boldsymbol{r}^*,\bullet}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) = \boldsymbol{0}$ and $T_{\boldsymbol{r}^*,\bullet}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}} = 0$. However, for any $\boldsymbol{\alpha} \in \mathcal{A}$, one has $\theta_{j,\boldsymbol{\alpha}} = g_j$ and $\theta_{k,\boldsymbol{\alpha}} = c_k$ for any k s.t. $\bar{q}_{j,k} = 1$, so for any

$\boldsymbol{\alpha} \in \mathcal{A}$ we have

$$T_{\boldsymbol{r}^*,\boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*) = \prod_{\substack{1\le k\le K:\\ q_{j,k}=1}} (\theta_{k,\boldsymbol{\alpha}} - g_k)(\theta_{j,\boldsymbol{\alpha}} - c_j)$$

$$= \prod_{\substack{1\le k\le K:\\ q_{j,k}=1}} (c_k - g_k)(g_j - c_j) \neq 0,$$

and hence

$$T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)\boldsymbol{p} = \prod_{\substack{1\le k\le K:\\ q_{j,k}=1}} (c_k - g_k)(g_j - c_j) \sum_{\boldsymbol{\alpha}\in\mathcal{A}} p_{\boldsymbol{\alpha}}$$

$$\neq 0 = T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^*, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^*)\bar{\boldsymbol{p}},$$

which contradicts (S2.3).

(b) Based on (a), we next show $\bar{\boldsymbol{q}}_j = \boldsymbol{q}_j$ for all $j \in \{K+1, \dots, J\}$ using proof by contradiction. Since part (a) gives $\bar{\boldsymbol{q}}_j \succeq \boldsymbol{q}_j$, if $\bar{\boldsymbol{q}}_j \neq \boldsymbol{q}_j$, then there must exist some attribute $k \in [K]$ such that $\bar{q}_{j,k} = 1$ and $q_{j,k} = 0$. This implies $\bar{\boldsymbol{q}}_j \succeq \bar{\boldsymbol{q}}_k$. Define

$$\boldsymbol{\theta}^\star = \bar{c}_k \boldsymbol{e}_k + \bar{g}_j \boldsymbol{e}_j + \sum_{m>K:\, m\neq j} g_m \boldsymbol{e}_m,$$

then $T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^\star, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = 0$. Since Condition C holds, each attribute is required by at least one item in the set $\{m > K : m \neq j\}$, which implies $T_{\boldsymbol{r}^\star,\boldsymbol{\alpha}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^\star, \boldsymbol{g}-\boldsymbol{\theta}^\star) \neq 0$ only if $\boldsymbol{\alpha} = \mathbf{1}$. Therefore (S2.4) gives that

$$0 = T_{\boldsymbol{r}^\star,\boldsymbol{\cdot}}(Q, \boldsymbol{c}-\boldsymbol{\theta}^\star, \boldsymbol{g}-\boldsymbol{\theta}^\star)\boldsymbol{p}$$

$$= (c_k - \bar{c}_k)(c_j - \bar{g}_j) \prod_{m>K:\, m\neq j} (c_m - g_m) p_{\mathbf{1}},$$

so $c_k = \bar{c}_k$. Now we further define

$$\boldsymbol{\theta}^\star = \bar{c}_k \boldsymbol{e}_k + \bar{g}_j \boldsymbol{e}_j + \sum_{h \in [K] \setminus \{k\}} g_m \boldsymbol{e}_m,$$

then $T_{\boldsymbol{r}^\star, \boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^\star, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^\star)\bar{\boldsymbol{p}} = 0$. However, $\boldsymbol{q}_j \nsucceq \boldsymbol{q}_k$ under the true Q, and (S2.4) gives

$$T_{\boldsymbol{r}^\star, \boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^\star, \boldsymbol{g} - \boldsymbol{\theta}^\star)\boldsymbol{p} = (g_k - \bar{c}_k) \prod_{h \in [K] \setminus \{k\}} (c_h - g_h)(c_j - \bar{g}_j) p_{\boldsymbol{\alpha} - \boldsymbol{e}_k},$$

where $\boldsymbol{\alpha} - \boldsymbol{e}_k = (\mathbf{1}, \underbrace{0}_{\text{column } k}, \mathbf{1})$, so the above display is nonzero. This contradicts (S2.4), and this means $\bar{\boldsymbol{q}}_j \neq \boldsymbol{q}_j$ can not happen. So we have $\bar{\boldsymbol{q}}_j = \boldsymbol{q}_j$ for $j \in \{K+1, \ldots, J\}$.

Now we have proved $Q \sim \bar{Q}$. Now that $Q \sim \bar{Q}$, Theorem 1 in Gu and Xu (2018b) gives that Conditions A and B ensure the identifiability of the model parameters ($\boldsymbol{s} := \mathbf{1} - \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}$). This concludes the proof of the sufficiency of the conditions.

In the end we show the necessity of the conditions. By Theorem 1 in Gu and Xu (2018b), Conditions A and B are necessary for identifiability of the model parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ given a known Q, so they are also necessary for identifiability of $(Q, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$.

S3. Proof of Theorem 2

Proof of the necessity of each attribute required by ≥ 2 items. Suppose Q takes the form of

$$Q = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & Q^\star \end{pmatrix},$$

then for any valid $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ associated with Q, we next construct $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}}) \neq (\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ such that $T(Q, \boldsymbol{c}, \boldsymbol{g})\boldsymbol{p} = T(Q, \bar{\boldsymbol{c}}, \bar{\boldsymbol{g}})\bar{\boldsymbol{p}}$ holds. In particular, we arbitrarily choose $\bar{c}_1$ that is not equal to $c_1 = 1 - s_1$ and set

$$\bar{p}_{\boldsymbol{\alpha}} = \begin{cases} (c_1/\bar{c}_1)p_{\boldsymbol{\alpha}}, & \text{if } \alpha_1 = 1, \\ p_{\boldsymbol{\alpha}} + (1 - c_1/\bar{c}_1)p_{\boldsymbol{\alpha}+\boldsymbol{e}_1}, & \text{if } \alpha_1 = 0. \end{cases}$$

Then set $\bar{g}_1 = g_1$, and $\bar{c}_j = c_j$, $\bar{g}_j = g_j$ for $j = 2, \ldots J$. Then it is not hard to check that $T(Q, \boldsymbol{c}, \boldsymbol{g})\boldsymbol{p} = T(Q, \bar{\boldsymbol{c}}, \bar{\boldsymbol{g}})\bar{\boldsymbol{p}}$. Since $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ are arbitrary, we have shown the non-identifiability set spans the entire parameter space and $(Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ are not generically identifiable. Therefore, this proves that $(Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ are not generically identifiable if some attribute is required by only one item.

In the following we prove part (a), (b), and (c) when some attribute is required by only two items.

Proof of Part (a). Under the assumption of part (a), Q takes the form

$$Q = \begin{pmatrix} 1 & \mathbf{0}^\top \\ 1 & \mathbf{1}^\top \\ \mathbf{0} & Q^\star \end{pmatrix}.$$

Given arbitrary DINA model parameters $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ under this Q, we next construct another different set of DINA parameters $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}}) \neq (\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ also associated with this Q, such that

$$T(Q, \boldsymbol{c}, \boldsymbol{g})\boldsymbol{p} = T(Q, \bar{\boldsymbol{c}}, \bar{\boldsymbol{g}})\bar{\boldsymbol{p}}. \tag{S3.17}$$

In particular, we set $\bar{c}_j = c_j$ and $\bar{g}_j = g_j$ for all $j = 3, \ldots, J$. Under this construction, (S3.17) simplifies to the following two sets of equations

$$\forall \boldsymbol{\alpha}' \in \{0,1\}^{K-1},\ \boldsymbol{\alpha}' \neq \mathbf{1}, \quad \begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}, \\ g_1 p_{(0,\boldsymbol{\alpha}')} + c_1 p_{(1,\boldsymbol{\alpha}')} = \bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{c}_1 \bar{p}_{(1,\boldsymbol{\alpha}')}, \\ g_2[p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')}] = \bar{g}_2[\bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}], \\ g_2[g_1 p_{(0,\boldsymbol{\alpha}')} + c_2 p_{(1,\boldsymbol{\alpha}')}] = \bar{g}_2[\bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{c}_1 \bar{p}_{(1,\boldsymbol{\alpha}')}]; \end{cases} \tag{S3.18}$$

and for $\boldsymbol{\alpha}' = \mathbf{1}$,

$$\begin{cases} p_{(0,\mathbf{1})} + p_{(1,\mathbf{1})} = \bar{p}_{(0,\mathbf{1})} + \bar{p}_{(1,\mathbf{1})}, \\ g_1 p_{(0,\mathbf{1})} + c_1 p_{(1,\mathbf{1})} = \bar{g}_1 \bar{p}_{(0,\mathbf{1})} + \bar{c}_1 \bar{p}_{(1,\mathbf{1})}, \\ g_2 p_{(0,\mathbf{1})} + c_2 p_{(1,\mathbf{1})} = \bar{g}_2 \bar{p}_{(0,\mathbf{1})} + \bar{c}_2 \bar{p}_{(1,\mathbf{1})}, \\ g_1 g_2 p_{(0,\mathbf{1})} + c_1 c_2 p_{(1,\mathbf{1})} = \bar{g}_1 \bar{g}_2 \bar{p}_{(0,\mathbf{1})} + \bar{c}_1 \bar{c}_2 \bar{p}_{(1,\mathbf{1})}. \end{cases} \tag{S3.19}$$

The above (S3.18) obviously leads to $\bar{g}_2 = g_2$, and the last two equations of (S3.18) are automatically satisfied if the first two of (S3.18) are satisfied. Then the last two equations of (S3.19) can be transformed to

$$\begin{cases} (c_2 - g_2) p_{(1,\mathbf{1})} = (\bar{c}_2 - g_2) \bar{p}_{(1,\mathbf{1})}, \\ c_1 (c_2 - g_2) p_{(1,\mathbf{1})} = \bar{c}_1 (\bar{c}_2 - g_2) \bar{p}_{(1,\mathbf{1})}; \end{cases}$$

which gives $\bar{c}_1 = c_1$. Additionally, when $\bar{c}_1 = c_1$, we also have that the last equality of (S3.19) holds as long as the first three equalities of (S3.19) hold. In summary, now there are $2^K + 2$ parameters to be determined, which are $\{\bar{g}_1, \bar{c}_2\} \cup \{\bar{p}_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \{0,1\}^K\}$, while they only have to satisfy the following $2 \times (2^{K-1} - 1) + 3 = 2^K + 1$ constraints,

$$\forall \boldsymbol{\alpha}' \in \{0,1\}^{K-1}, \text{ for } \boldsymbol{\alpha}' \neq \mathbf{1}, \quad \begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}, \\ g_1 p_{(0,\boldsymbol{\alpha}')} + c_1 p_{(1,\boldsymbol{\alpha}')} = \bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}')} + c_1 \bar{p}_{(1,\boldsymbol{\alpha}')}; \end{cases}$$

$$\text{and for } \boldsymbol{\alpha}' = \mathbf{1}, \quad \begin{cases} p_{(0,\mathbf{1})} + p_{(1,\mathbf{1})} = \bar{p}_{(0,\mathbf{1})} + \bar{p}_{(1,\mathbf{1})}, \\ g_1 p_{(0,\mathbf{1})} + c_1 p_{(1,\mathbf{1})} = \bar{g}_1 \bar{p}_{(0,\mathbf{1})} + c_1 \bar{p}_{(1,\mathbf{1})}, \\ g_2 p_{(0,\mathbf{1})} + c_2 p_{(1,\mathbf{1})} = g_2 \bar{p}_{(0,\mathbf{1})} + \bar{c}_2 \bar{p}_{(1,\mathbf{1})}. \end{cases}$$

Since the number of free variables $2^K + 2$ is greater than the number of constraints $2^K + 1$, there exist infinitely many different solutions to the above system of equations. This means that the $(Q, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ are not generically identifiable. In particular, one can arbitrarily choose $\bar{g}_1$ close to but not equal to g_1, then solve for the remaining parameters $\{\bar{p}_{\boldsymbol{\alpha}},\ \boldsymbol{\alpha} \in \{0,1\}^K\}$ and $\bar{c}_2$ as follows,

$$\forall \boldsymbol{\alpha}' \in \{0,1\}^{K-1}, \quad \begin{cases} \bar{p}_{(0,\boldsymbol{\alpha}')} = p_{(0,\boldsymbol{\alpha}')}(g_1 - c_1)/(\bar{g}_1 - c_1), \\ \bar{p}_{(1,\boldsymbol{\alpha}')} = p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} - \bar{p}_{(0,\boldsymbol{\alpha}')}; \end{cases}$$

$$\bar{c}_2 = \frac{g_2[p_{(0,\mathbf{1})} - \bar{p}_{(0,\mathbf{1})}] + c_2 p_{(1,\mathbf{1})}}{\bar{p}_{(1,\mathbf{1})}}.$$

This concludes the proof of part (a) of the theorem.

Next we first prove (b.2), i.e. when $Q^\star$ has two submatrices $\mathcal{I}_{K-1}$. In this case, the Q contains a submatrix of the form $(I_K, I_K)^\top$. The proof of (b.1), i.e. when $Q^\star$ satisfies Conditions A, B and C, is combined with the proof of part (c) later.

Proof of Part (b.2). We first give the proof when Q only consists of two I_K's, namely $Q = (I_K, I_K)^\top$. In this case, we first prove that $\bar{Q} \sim Q$ must hold, using an argument similar to Step 1 of the proof of Theorem 1. Suppose $T(Q, \boldsymbol{c}, \boldsymbol{g})\boldsymbol{p} = T(\bar{Q}, \bar{\boldsymbol{c}}, \bar{\boldsymbol{g}})\bar{\boldsymbol{p}}$. Since

$Q_{(K+1):(2K),\bullet} = I_K$, we have that for each attribute $h \in [K]$, there is

$$\max_{\substack{m\in\{K+1,\dots,2K\},\\ m\neq K+h}} q_{m,h} = 0, \qquad \max_{m\in\{K+1,\dots,2K\}} q_{m,k} = 1\ \forall k \in [K]\setminus\{h\}.$$

Therefore we can apply Lemma 2 with $S = \{K+1,\dots,2K\}\setminus\{K+h\}$ and $\mathcal{J} = [K]\setminus\{h\}$ to obtain

$$\max_{k\in\mathcal{J}} \bar{\boldsymbol{q}}_k \nsucceq \bar{\boldsymbol{q}}_h.$$

This essentially implies that for an arbitrary $h \in [K]$, there must be a $m_h \in [K]$ such that $\bar{q}_{h,m_h} = 0$ and $\bar{q}_{k,m_h} = 0$ for all $k \in [K]\setminus\{h\}$. Moreover, the K integers $m_1, m_2, \dots, m_K$ must all be distinct, otherwise it is easy to see $\max_{k\in\mathcal{J}} \bar{\boldsymbol{q}}_k \nsucceq \bar{\boldsymbol{q}}_h$ would fail to hold for some $h \in [K]$. So $(m_1, m_2, \dots, m_K)$ is a permutation of $(1, 2, \dots, K)$. Now we have obtained that $\bar{Q}_{1:K,(m_1,\dots,m_K)}$ must be an identity matrix, i.e., $\bar{Q}_{1:K,\bullet} \sim Q_{1:K,\bullet}$. Reasoning in exactly the same way gives $\bar{Q}_{(K+1):(2K),\bullet} \sim Q_{(K+1):(2K),\bullet}$, and we have $\bar{Q} \sim Q$. Now for an arbitrary $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_K) \equiv (\alpha_1, \boldsymbol{\alpha}')$, define

$$\begin{aligned}\boldsymbol{\theta}^* =& \bar{g}_1\boldsymbol{e}_1 + \bar{c}_{K+1}\boldsymbol{e}_{K+1} + \sum_{k>1:\,\alpha_k=1} g_k\boldsymbol{e}_k + \sum_{k>1:\,\alpha_k=0} c_k\boldsymbol{e}_k\\ \equiv& \bar{g}_1\boldsymbol{e}_1 + \bar{c}_{K+1}\boldsymbol{e}_{K+1} + \boldsymbol{\theta}^{\boldsymbol{\alpha}}\end{aligned}$$

then $T_{\boldsymbol{e}_1+\boldsymbol{e}_{K+1}}(Q, \bar{\boldsymbol{s}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*) = \mathbf{0}$, so

$$\begin{aligned}0 =& T_{\boldsymbol{e}_1+\boldsymbol{e}_{K+1}}(Q, \bar{\boldsymbol{s}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}} = T_{\boldsymbol{e}_1+\boldsymbol{e}_{K+1}}(Q, \boldsymbol{s} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p}\\ =& \prod_{k>1:\,\alpha_k=1}(c_k - g_k)\times\prod_{k>1:\,\alpha_k=0}(g_k - c_k)\times\\ &\Big[(g_1 - \bar{g}_1)(g_{K+1} - \bar{c}_{K+1})p_{(0,\alpha_2,\dots,\alpha_K)} + (c_1 - \bar{g}_1)(c_{K+1} - \bar{c}_{K+1})p_{(1,\alpha_2,\dots,\alpha_K)}\Big].\end{aligned}$$

This implies that for any $\boldsymbol{\alpha}' = (\alpha_2, \dots, \alpha_K) \in \{0,1\}^{K-1}$, we have

$$(g_1 - \bar{g}_1)(g_{K+1} - \bar{c}_{K+1})p_{(0,\alpha_2,\dots,\alpha_K)} + (c_1 - \bar{g}_1)(c_{K+1} - \bar{c}_{K+1})p_{(1,\alpha_2,\dots,\alpha_K)} = 0.$$

Since $g_{K+1} - \bar{c}_{K+1} \neq 0$, we have that

$$g_1 - \bar{g}_1 = \frac{(c_1 - \bar{g}_1)(c_{K+1} - \bar{c}_{K+1})p_{(1,\boldsymbol{\alpha}')}}{(\bar{c}_{K+1} - g_{K+1})p_{(0,\boldsymbol{\alpha}')}}, \quad \text{for any } \boldsymbol{\alpha}' \in \{0,1\}^{K-1}.$$

This equality indicates that if there exists $\boldsymbol{\alpha}'_1 \neq \boldsymbol{\alpha}'_2$ such that

$$\frac{p_{(1,\boldsymbol{\alpha}'_1)}}{p_{(0,\boldsymbol{\alpha}'_1)}} \neq \frac{p_{(1,\boldsymbol{\alpha}'_2)}}{p_{(0,\boldsymbol{\alpha}'_2)}}, \tag{S3.20}$$

then one must have

$$c_{K+1} - \bar{c}_{K=1} = 0, \quad g_1 - \bar{g}_1 = 0.$$

Redefine $\boldsymbol{\theta}^* = \bar{c}_1 \boldsymbol{e}_1 + \bar{g}_{K+1}\boldsymbol{e}_{K+1} + \boldsymbol{\theta}^{\boldsymbol{\alpha}}$, then following the same procedure as above one gets that if $\boldsymbol{p}$ satisfy (S3.20), then $g_{K+1} - \bar{g}_{K=1} = 0$ and $c_1 - \bar{c}_1 = 0$.

Similarly as the above procedure for $k = 1$, we have that if for any attribute $k \in \{1, \dots, K\}$, there exist two attribute profiles $\boldsymbol{\alpha}^{k,1}, \boldsymbol{\alpha}^{k,2} \in \{0,1\}^{k-1} \times \{0\} \times \{0,1\}^{K-k-1}$ such that

$$\frac{p_{\boldsymbol{\alpha}^{k,1}+\boldsymbol{e}_k}}{p_{\boldsymbol{\alpha}^{k,1}}} \neq \frac{p_{\boldsymbol{\alpha}^{k,2}+\boldsymbol{e}_k}}{p_{\boldsymbol{\alpha}^{k,2}}}, \tag{S3.21}$$

then

$$\bar{g}_k = g_k, \quad \bar{c}_k = c_k\ , \quad \bar{g}_{K+k} = g_{K+k}, \quad \bar{c}_{K+k} = c_{K+k} \quad \text{for every } k \in \{1, \dots, K\}.$$

Now that all the item parameters are identified under (S3.21), Equation (S3.22) gives $\bar{\boldsymbol{p}} = \boldsymbol{p}$. Therefore other than the measure zero set of the parameter space specified by constraints (S3.21), $(Q, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ are identifiable. This means $(Q, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ are generically identifiable.

In particular, if Q takes form of the $Q_{2\times 4}$ in (3.5),

$$Q_{2\times 4} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

then constraints (S3.21) just simplify to

$$\frac{p_{(10)}}{p_{(00)}} \neq \frac{p_{(11)}}{p_{(01)}} \quad \text{and} \quad \frac{p_{(01)}}{p_{(00)}} \neq \frac{p_{(11)}}{p_{(10)}},$$

which can be equivalently written as inequality (3.6) that $p_{(01)}p_{(10)} \neq p_{(00)}p_{(11)}$ in the main text.

Next we prove the conclusion when Q contains other rows besides the two identity submatrices, namely $Q = (I_K, I_K, (Q^\star)^\top)^\top$. Using exactly the same arguments as previously we have that generically, all the item parameters of the first $2K$ items as well as all the proportion parameters are satisfied. Now for any $J > 2K$ and $\boldsymbol{\alpha} \in \{0,1\}^K$ define $\boldsymbol{r}^* = \sum_{k=1}^K \boldsymbol{e}_j$ and

$$\boldsymbol{\theta}^* = \sum_{1\leq k\leq K:\, \boldsymbol{\alpha}\succeq \boldsymbol{q}_j} g_j \boldsymbol{e}_j + \sum_{1\leq k\leq K:\, \boldsymbol{\alpha}\not\succeq \boldsymbol{q}_j} c_j \boldsymbol{e}_j,$$

then (S0.2) implies that

$$\theta_{j,\boldsymbol{\alpha}} = \frac{T_{\boldsymbol{r}^*+\boldsymbol{e}_j}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)\boldsymbol{p}}{T_{\boldsymbol{r}^*}(Q, \boldsymbol{c}-\boldsymbol{\theta}^*, \boldsymbol{g}-\boldsymbol{\theta}^*)\boldsymbol{p}} = \frac{T_{\boldsymbol{r}^*+\boldsymbol{e}_j}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^*, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^*)\bar{\boldsymbol{p}}}{T_{\boldsymbol{r}^*}(\bar{Q}, \bar{\boldsymbol{c}}-\boldsymbol{\theta}^*, \bar{\boldsymbol{g}}-\boldsymbol{\theta}^*)\bar{\boldsymbol{p}}} = \bar{\theta}_{j,\boldsymbol{\alpha}}.$$

This proves that any slipping or guessing parameter associated with item $j > 2K$ is identifiable under the generic constraints (S3.21), and this completes the proof of part (b.2) of the theorem.

Next we prove (b.1) and (c) in Theorem 2 in four steps.

Proof of Part (b.1) and Part (c).

Step 1. In this step, we aim to show that if

$$T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q,\boldsymbol{s},\boldsymbol{g})\boldsymbol{p} = T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q},\bar{\boldsymbol{s}},\bar{\boldsymbol{g}})\bar{\boldsymbol{p}} \quad \text{for every } \boldsymbol{r} \in \{0,1\}^J, \tag{S3.22}$$

then $\bar{Q}$ must take the following form up to column permutation

$$\bar{Q} = \begin{pmatrix} 1 & \mathbf{0} \\ \bar{u} & \bar{\boldsymbol{v}} \\ \mathbf{0} & Q^{\star} \end{pmatrix}. \tag{S3.23}$$

Here $(\bar{u}, \bar{\boldsymbol{v}})$ is a K dimensional binary vector. The structure of $(\bar{u}, \bar{\boldsymbol{v}})$ will be studied in Steps 2 and 3.

Since the submatrix $Q^{\star}$ of Q satisfies Conditions A, B and C, the matrix Q can be written as

$$Q = \begin{pmatrix} 1 & \mathbf{0}^{\top} \\ 1 & \boldsymbol{v}^{\top} \\ \mathbf{0} & \mathcal{I}_{K-1} \\ \mathbf{0} & Q^{\star\star} \end{pmatrix},$$

then follow the same procedure as Step 1 in the proof of Theorem 1 one has that, up to some column permutation, $\bar{Q}$ takes the form

$$\bar{Q} = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \bar{u} & \bar{\boldsymbol{v}}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ \bar{\boldsymbol{b}} & \bar{Q}^{\star\star} \end{pmatrix}.$$

For notational convenience and without loss of generality, in the following proof we rearrange the order of the row vectors of Q (and $\bar{Q}$) and rewrite them as follows

$$Q = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ 1 & \boldsymbol{v}^\top \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}, \qquad \bar{Q} = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ \bar{u} & \bar{\boldsymbol{v}}^\top \\ \bar{\boldsymbol{b}} & \bar{Q}^{\star\star} \end{pmatrix}. \tag{S3.24}$$

Now that each column of $Q^{\star\star}$ contains at least two entries of "1" from the assumption of scenarios (b.1) and (c), following the same procedure as Step 2 in the proof of Theorem 1 we can obtain

$$c_j = \bar{c}_j, \quad \text{for } j = K+2, \ldots, J.$$

Note that slightly different from Step 2 in the proof of Theorem 1, here we *do not have* $c_{K+1} = \bar{c}_{K+1}$ due to the fact that the first attribute is required by only two items.

Now denote the $(J-K) \times (K-1)$ bottom-right submatrix of Q by Q^s and the

$(J-K)\times K$ bottom submatrix of Q by Q^l, i.e.,

$$Q^s = \begin{pmatrix} \boldsymbol{v}^\top \\ \\ Q^{\star\star} \end{pmatrix}, \quad Q^l = \begin{pmatrix} 1 & \boldsymbol{v}^\top \\ \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}$$

and assume without loss of generality that the $K-1$ column vectors of Q^s are arranged in the lexicographic order. Specifically, for any $1 \leq k_1 < k_2 \leq K-1$, assume $Q^s_{\bullet,k_1} \prec_{\text{lex}} Q^s_{\bullet,k_2}$. This implies that the vector $\boldsymbol{v}$ can be written as

$$\boldsymbol{v} = (0, \dots, 0, 1, \dots, 1)$$

Note that in scenario (b.1), $\boldsymbol{v} = \mathbf{0}$ and $k_0 = K-1$. where its first k_0 elements are zero and the remain $K-1-k_0$ elements are one. So $\boldsymbol{q}_2 = (1, \boldsymbol{v}) = (1, 0, \dots, 0, 1, \dots, 1)$. We now use an induction method to prove that

$$g_k = \bar{g}_k, \quad \forall k = 2,\, \dots,\, 1+k_0. \tag{S3.25}$$

A key observation is that if considering the order of the columns of the larger submatrix Q^l instead of Q^s, then the first column of Q^l, i.e. $Q^l_{\bullet,1}$ is of larger lexicographic order of $Q^l_{\bullet,k}$ for any $k = 2, \dots, 1+k_0$. This indicates that we can follow a similar induction argument as Step 3 in the proof of Theorem 1 by defining $\boldsymbol{\theta}^*_k$ as (the same form as (S2.14))

$$\boldsymbol{\theta}^*_k = \sum_{h=1}^{K} \bar{g}_h \boldsymbol{e}_h + \sum_{h>K:\, q_{h,k}=0} g_h \boldsymbol{e}_h + \sum_{h>K:\, q_{h,k}=1} c_h \boldsymbol{e}_h, \tag{S3.26}$$

for $k = 2, \dots, 1+k_0$ one after another, to obtain (S3.25).

We emphasize here that if $\boldsymbol{v} = \mathbf{0}$, i.e. in scenario (b.1) of the theorem, then $k_0 = K-1$ and by far we have already obtained $\bar{g}_k = g_k$ for all $k = 2, \dots, K$. So we can directly go to the next step, Step 2 of the proof, without the local condition to appear in (S3.29) later. That is why in scenario (b.1) of the theorem, we have *global* generic identifiability of $(Q, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$.

Next we consider the case $\boldsymbol{v} \neq \mathbf{0}$, i.e. in scenario (c) of the theorem, then $k_0 < K-1$. We will use another induction argument to show $\bar{g}_k = g_k$ for $k = k_0+2, \dots, K$, under an additional local condition. First we consider $\bar{g}_k$ and g_k for $k = k_0 + 2$. Note that $Q_{\boldsymbol{\cdot},k} \succ_{\text{lex}} Q_{\boldsymbol{\cdot},1}$, and $Q_{\boldsymbol{\cdot},k} \prec_{\text{lex}} Q_{\boldsymbol{\cdot},m}$ for any $m = k+1, \dots, K$. Define $\boldsymbol{\theta}^*_k$ the same as in (S3.26), then $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*_k, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*_k) = \mathbf{0}$ and $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*_k, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*_k)\bar{\boldsymbol{p}} = 0$, so $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*_k, \boldsymbol{g} - \boldsymbol{\theta}^*_k)\boldsymbol{p} = 0$. We claim that in the the vector $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*_k, \boldsymbol{g} - \boldsymbol{\theta}^*_k)$, denoted by $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}$ afterwards for notational simplicity, only contains two potentially nonzero elements corresponding to attribute profiles $\boldsymbol{\alpha}_{1k} = \sum_{m=1}^K \boldsymbol{e}_m - \boldsymbol{e}_k = (1, \dots, 1, \alpha_k = 0, 1, \dots, 1)$ and $\boldsymbol{\alpha}_{0k} = \boldsymbol{\alpha}_1 - \boldsymbol{e}_1 = (\alpha_1 = 0, 1, \dots, 1, \alpha_k = 0, 1, \dots, 1)$. This is because on the one hand, for any attribute profile $\boldsymbol{\alpha}$ that lacks some attribute $m \in \{k+1, \dots, K\}$, $\theta_{h,\boldsymbol{\alpha}} = g_h$ for some item $h > K$ with $q_{h,k} = 0$, which makes $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}} = 0$; and on the other hand, for any attribute profile that lacks some attribute $m \in \{2, \dots, k-1\}$, since we already have (S3.25), $\theta_{h,m} = g_h = \bar{g}_h$ for some $h \in \{2, \dots, K\}$, which makes $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}} = 0$. Now $T_{\boldsymbol{r}^*,\boldsymbol{\alpha}} \neq 0$ would only happen if $\boldsymbol{\alpha} = (\alpha_1, 1, \dots, 1, \alpha_k, 1, \dots, 1)$. However, if $\alpha_k = 1$ and $\boldsymbol{\alpha} = (\alpha_1, 1, \dots, 1)$, then $\theta_{h,\boldsymbol{\alpha}} = c_h$ for some item $h > K$ with $q_{h,k} = 1$, which also makes

$T_{\boldsymbol{r}^*,\boldsymbol{\alpha}} = 0$. Now we have proven the claim that $T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}$ has only two potentially nonzero elements corresponding to $\boldsymbol{\alpha}_{1k}$ and $\boldsymbol{\alpha}_{0k}$. Therefore we have for $k = k_0 + 2$,

$$\begin{aligned}
0 =& T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}_k^*, \boldsymbol{g} - \boldsymbol{\theta}_k^*)\boldsymbol{p} \\
=& \prod_{h=2}^{K}(c_h - \bar{g}_h) \prod_{\substack{h>K:\\ q_{h,k}=0}} (c_h - g_h) \prod_{\substack{h>K:\\ q_{h,k}=1}} (g_h - c_h) \\
& \times \Big[(g_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{0k}} + (c_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{1k}}\Big](g_k - \bar{g}_k),
\end{aligned}$$

which further gives

$$\Big[(g_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{0k}} + (c_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{1k}}\Big](g_k - \bar{g}_k) = 0 \text{ for } k = k_0 + 2. \tag{S3.27}$$

Note that if $\bar{g}_1 = g_1$, then the part in the bracket in the above display becomes $(c_1 - g_1)p_{\boldsymbol{\alpha}_1}$, which is nonzero. Therefore, when $\bar{g}_1$ is sufficiently close to the true parameter g_1, the part in the bracket in (S3.27) would be nonzero. We formally write it as

$$\begin{aligned}
&\text{for } k = k_0 + 2, \quad \forall \bar{g}_1 \in \mathcal{N}_k, \quad (g_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{0k}} + (c_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{1k}} \neq 0, \\
&\qquad \text{where } \mathcal{N}_k = \{x : 0 < x < \frac{g_1 p_{\boldsymbol{\alpha}_{0k}} + c_1 p_{\boldsymbol{\alpha}_{1k}}}{p_{\boldsymbol{\alpha}_{0k}} + p_{\boldsymbol{\alpha}_{1k}}}\}.
\end{aligned} \tag{S3.28}$$

This indicates that in the neighborhood $\mathcal{N}_k$ of g_1, (S3.27) leads to $g_k = \bar{g}_k$ for $k = k_0 + 2$.

Then we use induction to prove $g_k = \bar{g}_k$ for all $k = k_0 + 3, \ldots, K$. As the induction assumption, assume that when $\bar{g}_1 \in \bigcap_{m=k_0+2}^{k-1} \mathcal{N}_m$ holds, we have $g_m = \bar{g}_m$ for all $m = 2, \ldots, k-1$. Then define $\boldsymbol{\theta}^*$ the same as in (S3.26), and deduce in the same way as in proving $g_{k_0+2} = \bar{g}_{k_0+2}$, we have

$$\Big[(g_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{0k}} + (c_1 - \bar{g}_1)p_{\boldsymbol{\alpha}_{1k}}\Big](g_k - \bar{g}_k) = 0,$$

and further for any $\bar{g}_1 \in \mathcal{N}_k$ (more accurately any $\bar{g}_1 \in \left[\cap_{m=k_0+2}^{k-1} \mathcal{N}_m\right] \cap \mathcal{N}_k$), we must have $\bar{g}_k = g_k$. Here $\mathcal{N}_k$ takes the same form as that in (S3.28). Now by induction, we have that if

$$\bar{g}_1 \in \bigcap_{m=k_0+2}^{K} \mathcal{N}_m, \tag{S3.29}$$

then $g_k = \bar{g}_k$ for $k = k_0 + 2, \ldots, K$. Combined with the previous results shown in (S3.25), now we have proven that in scenario (c) of the theorem, if the local condition (S3.29) is satisfied, then $\bar{g}_k = g_k$ for $k = 2, \ldots, K$.

In summary, we have shown $\bar{g}_k = g_k$ for $k = 2, \ldots, K$ (under (S3.29) if in scenario (c)) and $\bar{c}_j = c_j$ for $j = K + 2, \ldots, J$. Based on these, following similar procedures as in Step 5 of the proof of Theorem 1, we obtain that

$$\bar{\boldsymbol{q}}_j = \boldsymbol{q}_j, \quad \forall j = K + 2, \ldots, J.$$

Step 2. In this step we show $\bar{u} = 1$ in (S3.23). If $\bar{u} = 0$, set

$$\boldsymbol{\theta}^* = c_1 \boldsymbol{e}_1 + \bar{c}_2 \boldsymbol{e}_2 + \sum_{j=3}^{K+3} g_k \boldsymbol{e}_k, \quad \boldsymbol{r}^* = \sum_{j=1}^{K+3} \boldsymbol{e}_j,$$

then

$$T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(\bar{Q}, \bar{\boldsymbol{c}} - \boldsymbol{\theta}^*, \bar{\boldsymbol{g}} - \boldsymbol{\theta}^*)\bar{\boldsymbol{p}} = \boldsymbol{0}^\top \cdot \bar{\boldsymbol{p}} = 0,$$

$$T_{\boldsymbol{r}^*,\boldsymbol{\cdot}}(Q, \boldsymbol{c} - \boldsymbol{\theta}^*, \boldsymbol{g} - \boldsymbol{\theta}^*)\boldsymbol{p} = (g_1 - c_1)(g_2 - \bar{c}_2) \prod_{j=3}^{K+3} (c_j - g_j) p_{(0,1,\ldots,1)} \neq 0,$$

which contradicts Equation (S2.3). So $\bar{u} = 1$. Now we have obtained

$$Q = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ 1 & \boldsymbol{v}^\top \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}, \qquad \bar{Q} = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ 1 & \bar{\boldsymbol{v}}^\top \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}. \tag{S3.30}$$

Step 3. In this step we show $\bar{\boldsymbol{v}} = \boldsymbol{v}$. For notational simplicity in the following proof, we rearrange the order of the row vectors in Q and $\bar{Q}$ in (S3.30) again to the following forms

$$Q = \begin{pmatrix} 1 & \mathbf{0}^\top \\ 1 & \boldsymbol{v}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}, \qquad \bar{Q} = \begin{pmatrix} 1 & \mathbf{0}^\top \\ \bar{u} & \bar{\boldsymbol{v}}^\top \\ \mathbf{0} & \mathcal{I}_{K-1} \\ \mathbf{0} & Q^{\star\star} \end{pmatrix}, \tag{S3.31}$$

and our conclusions proved so far are $\bar{g}_k = g_k$ for $k = 3, \ldots, K+1$ and $\bar{c}_j = c_j$ for $j = K+2, \ldots, J$ (under the local condition (S3.29) if in scenario (b.1)). Given that the last $J-2$ rows of Q and $\bar{Q}$ are equal, we claim that (S3.22) for response pattern $\boldsymbol{r}$ can be equivalently written as

$$\begin{aligned} &\sum_{\substack{\boldsymbol{\alpha}'\in \\ \{0,1\}^{K-1}}} \prod_{\substack{j>2 \\ r_j=1}} \theta_{j,(0,\boldsymbol{\alpha}')} \cdot \mathbb{P}(R_1 \geq r_1,\, R_2 \geq r_2,\, \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid Q, \boldsymbol{\Theta}, \boldsymbol{p}) \\ = &\sum_{\substack{\boldsymbol{\alpha}'\in \\ \{0,1\}^{K-1}}} \prod_{\substack{j>2 \\ r_j=1}} \bar{\theta}_{j,(0,\boldsymbol{\alpha}')} \cdot \mathbb{P}(R_1 \geq r_1,\, R_2 \geq r_2,\, \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid \bar{Q}, \bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}}). \end{aligned} \tag{S3.32}$$

Here $\boldsymbol{A} = (A_1, \ldots, A_K)$ denotes a random attribute profile following a categorical distribution with proportion parameters $\boldsymbol{p}$, and $\boldsymbol{A}_{2:K}$ denotes the vector consisting of the last $K-1$ elements of $\boldsymbol{A}$. The reason for the equivalence of (S3.32) and (S3.22) is stated as follows. Since all items other than the first two do not require the first attribute, we have that for any $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$, the two attribute profiles $(0,\boldsymbol{\alpha}')$ and $(1,\boldsymbol{\alpha}')$ always have the same response probability $\theta_{j,(0,\boldsymbol{\alpha}')}$ to any item $j > 2$. This indicates that the left hand side of (S3.22) can be written as

$$T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q,\boldsymbol{s},\boldsymbol{g})\boldsymbol{p} = \sum_{\substack{\boldsymbol{\alpha}'\in\\ \{0,1\}^{K-1}}} \prod_{\substack{j>2\\ r_j=1}} \theta_{j,(0,\boldsymbol{\alpha}')} \cdot \mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid Q, \boldsymbol{\Theta}, \boldsymbol{p}),$$

and this further leads to the equivalence between (S3.22) and (S3.32). In particular, when $(r_1,r_2) = (0,0)$, we have $\mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid Q, \boldsymbol{\Theta}, \boldsymbol{p}) = p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')}$. Now for any J-dimensional response pattern $\boldsymbol{r}$ with $(r_1,r_2) = (0,0)$, then the constraint $T_{\boldsymbol{r},\boldsymbol{\cdot}}(Q,\boldsymbol{c},\boldsymbol{g})\boldsymbol{p} = T_{\boldsymbol{r},\boldsymbol{\cdot}}(\bar{Q},\bar{\boldsymbol{c}},\bar{\boldsymbol{g}})\bar{\boldsymbol{p}}$ simply becomes

$$\sum_{\substack{\boldsymbol{\alpha}'\in\\ \{0,1\}^{K-1}}} \prod_{\substack{j>2\\ r_j=1}} \theta_{j,(0,\boldsymbol{\alpha}')} \cdot (p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')}) = \sum_{\substack{\boldsymbol{\alpha}'\in\\ \{0,1\}^{K-1}}} \prod_{\substack{j>2\\ r_j=1}} \bar{\theta}_{j,(0,\boldsymbol{\alpha}')} \cdot (\bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}).$$

Since the above equality holds for any $(r_3, r_4, \ldots, r_J) \in \{0,1\}^{J-2}$, we claim that, parameters $\theta_{j,(0,\boldsymbol{\alpha}')}$ and $\bar{\theta}_{j,(0,\boldsymbol{\alpha}')}$ for $j = 3,\ldots,J$ can be equivalently viewed as all the item parameters (slipping or guessing) associated with the submatrix $Q^\star$, while grouped proportion parameters $p_{(0,\boldsymbol{\alpha}')}+p_{(1,\boldsymbol{\alpha}')}$ and $\bar{p}_{(0,\boldsymbol{\alpha}')}+\bar{p}_{(1,\boldsymbol{\alpha}')}$ can be viewed as all the "proportion parameters" associated with $Q^\star$. Since $Q^\star$ satisfy the sufficient conditions A, B, C in Theorem 1 for identifiability, by Theorem 1 we conclude that $\theta_{j,(0,\boldsymbol{\alpha}')} = \bar{\theta}_{j,(0,\boldsymbol{\alpha}')}$ for any $j \in \{3,\ldots,J\}$ and any $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$. This indicates $\bar{c}_k = c_k$ for $k = 3,\ldots,K+1$ and

$\bar{g}_j = g_j$ for $j = K+2, \ldots, J$.

Then an important observation is that, fix any particular pair of $(r_1, r_2) \in \{0,1\}^2$, quantities in (S3.32) can be viewed parameters associated with the $(J-2) \times (K-1)$ matrix $Q^\star$, just similar to the argument in the previous paragraph. Specifically, $\theta_{j,(0,\boldsymbol{\alpha}')}$ and $\bar{\theta}_{j,(0,\boldsymbol{\alpha}')}$ for $j = 3, \ldots, J$ are item parameters (slipping or guessing) associated with the $Q^\star$, and $\mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid Q, \boldsymbol{\Theta}, \boldsymbol{p})$ and $\mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid \bar{Q}, \bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}})$ for each $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$ can be viewed as the "proportion parameters" associated with $Q^\star$. Now because the submatrix $Q^\star$ satisfy the identifiability conditions A, B, C; and $\bar{Q}_{3:J,\bullet} = Q_{3:J,\bullet} = Q^\star$ and $\bar{c}_j = c_j$, $\bar{g}_j = g_j$ for $j = 3, \ldots, J$, we must have

$$\begin{aligned}\forall \boldsymbol{\alpha}' \in \{0,1\}^{K-1}, \qquad & \mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid Q, \boldsymbol{\Theta}, \boldsymbol{p}) \\ = \ & \mathbb{P}(R_1 \geq r_1,\ R_2 \geq r_2,\ \boldsymbol{A}_{2:K} = \boldsymbol{\alpha}' \mid \bar{Q}, \bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}}).\end{aligned} \tag{S3.33}$$

Now take (r_1, r_2) to be $(0,0)$, $(0,1)$, $(1,0)$, $(1,1)$ in the above (S3.33) respectively, we obtain

$$\begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{1,\,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{1,\,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} = \bar{\theta}_{1,\,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{1,\,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{2,\,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{2,\,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} = \bar{\theta}_{2,\,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{2,\,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{1,\,(0,\boldsymbol{\alpha}')}\theta_{2,\,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{1,\,(1,\boldsymbol{\alpha}')}\theta_{2,\,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} \\ \qquad = \bar{\theta}_{1,\,(0,\boldsymbol{\alpha}')}\bar{\theta}_{2,\,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{1,\,(1,\boldsymbol{\alpha}')}\bar{\theta}_{2,\,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}. \end{cases} \tag{S3.34}$$

Next we show $\boldsymbol{v} = \bar{\boldsymbol{v}}$. (S3.34) implies that,

$$\forall \boldsymbol{\alpha}' \geq \boldsymbol{v},\ \boldsymbol{\alpha}' \ngeq \bar{\boldsymbol{v}}, \quad \begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')} \\ g_1 p_{(0,\boldsymbol{\alpha}')} + c_1 p_{(1,\boldsymbol{\alpha}')} = \bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{c}_1 \bar{p}_{(1,\boldsymbol{\alpha}')} \\ g_2 p_{(0,\boldsymbol{\alpha}')} + c_2 p_{(1,\boldsymbol{\alpha}')} = \bar{g}_2 [\bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}] \\ g_1 g_2 p_{(0,\boldsymbol{\alpha}')} + c_1 c_2 p_{(1,\boldsymbol{\alpha}')} = \bar{g}_2 [\bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{c}_1 \bar{p}_{(1,\boldsymbol{\alpha}')}] \end{cases}$$

If $\bar{\boldsymbol{v}} \ngeq \boldsymbol{v}$, then taking $\boldsymbol{\alpha}' = \boldsymbol{v}$ in the above equation and doing some transformation gives

$$\begin{cases} (g_2 - \bar{g}_2) p_{(0,\boldsymbol{\alpha}')} + (c_2 - \bar{g}_2) p_{(1,\boldsymbol{\alpha}')} = 0, \\ (g_1 - c_1)(g_2 - \bar{g}_2) p_{(0,\boldsymbol{\alpha}')} = 0. \end{cases}$$

Since $g_1 \neq c_1$, we have $g_2 - \bar{g}_2 = 0$, which further gives $c_2 - \bar{g}_2 = 0$. This contradicts $c_h > \bar{g}_h$ for any item h, so $\bar{\boldsymbol{v}} \ngeq \boldsymbol{v}$ can not happen. Similarly $\bar{\boldsymbol{v}} \nleq \boldsymbol{v}$ also can not happen, so $\bar{\boldsymbol{v}} = \boldsymbol{v}$.

Step 4. In the final step we show c_1, c_2, g_1, g_2 and $\boldsymbol{p}$ are generically identifiable if $\boldsymbol{v} \neq \mathbf{1}$. First we show that if there exist $\boldsymbol{\alpha}_1'$, $\boldsymbol{\alpha}_2' \in \{0,1\}^{K-1}$, $\boldsymbol{\alpha}_1' \neq \boldsymbol{\alpha}_2'$ such that

$$p_{(1,\boldsymbol{\alpha}_1')} p_{(0,\boldsymbol{\alpha}_2')} \neq p_{(1,\boldsymbol{\alpha}_2')} p_{(0,\boldsymbol{\alpha}_1')}, \tag{S3.35}$$

then one must have

$$c_i = \bar{c}_i,\ g_i = \bar{g}_i, \quad i = 1, 2. \tag{S3.36}$$

After some transformations, the system of equations (S4.40) yields

$$\begin{cases} (g_1 - c_1) \cdot (g_2 - \bar{c}_2) \cdot p_{(0,\boldsymbol{\alpha}')} = (\bar{g}_1 - c_1) \cdot (\bar{g}_2 - \bar{c}_2) \cdot \bar{p}_{(0,\boldsymbol{\alpha}')}, \\ (g_2 - \bar{c}_2) \cdot p_{(0,\boldsymbol{\alpha}')} + (c_2 - \bar{c}_2) \cdot \bar{p}_{(1,\boldsymbol{\alpha}')} = (\bar{g}_2 - \bar{c}_2) \cdot \bar{p}_{(0,\boldsymbol{\alpha}')}. \end{cases}$$

Since we have $\bar{g}_1 \neq c_1$, the left hand side of the first equation above is nonzero. And obviously the right hand side of the second equation above is nonzero. Taking the ratio of the above two equations gives

$$\frac{(g_1 - c_1) \cdot (g_2 - \bar{c}_2)}{(g_2 - \bar{c}_2) + (c_2 - \bar{c}_2) \cdot p_{(1,\boldsymbol{\alpha}')}/p_{(0,\boldsymbol{\alpha}')}} = (\bar{g}_1 - c_1) \equiv f(\boldsymbol{\alpha}').$$

The right hand side of the above display does not involve any proportion parameter $\boldsymbol{p}$ or $\bar{\boldsymbol{p}}$. So for $\boldsymbol{\alpha}'_1$, $\boldsymbol{\alpha}'_2$ satisfying (S3.35), $f(\boldsymbol{\alpha}'_1) = f(\boldsymbol{\alpha}'_2)$. Note that the left hand side of the above equation involves a ratio $p_{(1,\boldsymbol{\alpha}')}/p_{(0,\boldsymbol{\alpha}')}$ depending on $\boldsymbol{\alpha}'$. Equality $f(\boldsymbol{\alpha}'_1) = f(\boldsymbol{\alpha}'_2)$ along with (S3.35) imply

$$(c_2 - \bar{c}_2) \cdot \frac{p_{(1,\boldsymbol{\alpha}'_1)}}{p_{(0,\boldsymbol{\alpha}'_1)}} = (c_2 - \bar{c}_2) \cdot \frac{p_{(1,\boldsymbol{\alpha}'_2)}}{p_{(0,\boldsymbol{\alpha}'_2)}}$$

$$(c_2 - \bar{c}_2) \cdot \left(\frac{p_{(1,\boldsymbol{\alpha}'_1)}}{p_{(0,\boldsymbol{\alpha}'_1)}} - \frac{p_{(1,\boldsymbol{\alpha}'_2)}}{p_{(0,\boldsymbol{\alpha}'_2)}} \right) = 0$$

then since $p_{(1,\boldsymbol{\alpha}'_1)} p_{(0,\boldsymbol{\alpha}'_2)} \neq p_{(1,\boldsymbol{\alpha}'_2)} p_{(0,\boldsymbol{\alpha}'_1)}$ by assumption (S3.35), one must have $c_2 = \bar{c}_2$. By symmetry of the four item parameters g_1, c_1, g_2 and c_2 in (S4.40), equalities (S3.36) hold as claimed following similar arguments. Now that all the item parameters are identified, $\boldsymbol{p} = \bar{\boldsymbol{p}}$. This completes the proof of part (b.1) and part (c) of the theorem. The proof of Theorem 2 is now complete.

S4. Proof of Theorem 3

When Condition C fails and some attribute is required by less than three items, there are two possible scenarios: some attribute is required by only one item, or only two items. We consider them separately, and in both cases prove that $(Q, \boldsymbol{\Theta}, \boldsymbol{p})$ are not generically identifiable.

(a) If some attribute is required by only one item. Then Q must take the following form in (S4.37) up to column and row permutations, where $\boldsymbol{v}_1$ is a binary vector of length $K-1$.

$$Q = \begin{pmatrix} 1 & \boldsymbol{v}_1^\top \\ \mathbf{0} & Q^\star \end{pmatrix}; \qquad \bar{Q} = \begin{pmatrix} 1 & \mathbf{1}^\top \\ \mathbf{0} & Q^\star \end{pmatrix}. \tag{S4.37}$$

Now for arbitrary model parameters $(\boldsymbol{\Theta}, \boldsymbol{p})$ associated with Q, we also construct $(\bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}})$ associated with the $\bar{Q}$ in (S4.37), such that (S0.1) holds. Firstly, for any item $j \geq 2$, set $\bar{\theta}_{j,\boldsymbol{\alpha}} = \theta_{j,\boldsymbol{\alpha}}$ for all $\boldsymbol{\alpha} \in \{0,1\}^K$, then following a similar argument as in Step 3 of the proof of Theorem 2 (b.1) and (c), we have that (S0.1) hold as long as the following constraints are satisfied: for any $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$,

$$\begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{1,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{1,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} = \bar{\theta}_{1,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{1,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}. \end{cases} \tag{S4.38}$$

For each $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$, we now still arbitrarily set the value of $\bar{\theta}_{1,(0,\boldsymbol{\alpha}')}$ and

$\bar\theta_{1,(1,\boldsymbol{\alpha}')}$, and set the proportions parameters to be

$$\bar p_{(1,\boldsymbol{\alpha}')} = \frac{(\theta_{1,(0,\boldsymbol{\alpha}')} - \bar\theta_{1,(0,\boldsymbol{\alpha}')})p_{(0,\boldsymbol{\alpha}')} + (\theta_{1,(1,\boldsymbol{\alpha}')} - \bar\theta_{1,(0,\boldsymbol{\alpha}')})p_{(1,\boldsymbol{\alpha}')}}{\bar\theta_{1,(1,\boldsymbol{\alpha}')} - \bar\theta_{1,(0,\boldsymbol{\alpha}')}}$$

$$\bar p_{(0,\boldsymbol{\alpha}')} = p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} - \bar p_{(1,\boldsymbol{\alpha}')},$$

for each $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$. Then (S4.38) holds and further (S0.1) holds. Since the choice of the 2^K item parameters $\{\theta_{1,(0,\boldsymbol{\alpha}')},\, \theta_{1,(1,\boldsymbol{\alpha}')} : \boldsymbol{\alpha}' \in \{0,1\}^{K-1}\}$ are arbitrary, the original Q and associated parameters are not generically identifiable.

(b) If some attribute is required by only two items, then Q takes the form in (S4.39) up to column/row permutations, where $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are vectors of length $K-1$ and $Q^\star$ is a submatrix of size $(J-2)\times(K-1)$.

$$Q = \begin{pmatrix} 1 & \boldsymbol{v}_1^\top \\ 1 & \boldsymbol{v}_2^\top \\ \mathbf{0} & Q^\star \end{pmatrix}; \qquad \bar Q = \begin{pmatrix} 1 & \mathbf{1}^\top \\ 1 & \mathbf{1}^\top \\ \mathbf{0} & Q^\star \end{pmatrix}, \tag{S4.39}$$

Then for arbitrary model parameters $(\boldsymbol{\Theta}, \boldsymbol{p})$ associated with Q, we next carefully construct $(\bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}})$ associated with the $\bar Q$ in (S4.39), such that (S0.1) holds. This would prove the conclusion that joint generic identifiability fails. Firstly, for any item $j \geq 3$, set $\bar\theta_{j,\boldsymbol{\alpha}} = \theta_{j,\boldsymbol{\alpha}}$ for all $\boldsymbol{\alpha} \in \{0,1\}^K$, then following the same argument as in Step 3 of the proof of Theorem 2 (b.1) and (c), we have that (S0.1) hold as long as the following constraints are satisfied for every $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$,

$$\begin{cases} p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} = \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{1,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{1,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} = \bar{\theta}_{1,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{1,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{2,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{2,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} = \bar{\theta}_{2,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{2,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}; \\ \theta_{1,(0,\boldsymbol{\alpha}')}\theta_{2,(0,\boldsymbol{\alpha}')} \cdot p_{(0,\boldsymbol{\alpha}')} + \theta_{1,(1,\boldsymbol{\alpha}')}\theta_{2,(1,\boldsymbol{\alpha}')} \cdot p_{(1,\boldsymbol{\alpha}')} \\ \qquad = \bar{\theta}_{1,(0,\boldsymbol{\alpha}')}\bar{\theta}_{2,(0,\boldsymbol{\alpha}')} \cdot \bar{p}_{(0,\boldsymbol{\alpha}')} + \bar{\theta}_{1,(1,\boldsymbol{\alpha}')}\bar{\theta}_{2,(1,\boldsymbol{\alpha}')} \cdot \bar{p}_{(1,\boldsymbol{\alpha}')}. \end{cases} \tag{S4.40}$$

For each $\boldsymbol{\alpha}' \in \{0,1\}^{K-1}$, arbitrarily choose $\bar{\theta}_{1,(0,\boldsymbol{\alpha}')}$ and $\bar{\theta}_{2,(0,\boldsymbol{\alpha}')}$ from the neighborhood of the true parameter values $\theta_{1,(0,\boldsymbol{\alpha}')}$ and $\theta_{2,(1,\boldsymbol{\alpha}')}$ respectively. Then set

$$\begin{cases} \bar{\theta}_{1,(1,\boldsymbol{\alpha}')} = \theta_{1,(0,\boldsymbol{\alpha}')} + \frac{([\theta_{1,(1,\boldsymbol{\alpha}')} - \theta_{1,(0,\boldsymbol{\alpha}')}][\theta_{2,(1,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}]p_{(1,\boldsymbol{\alpha}')}}{[\theta_{2,(0,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}]p_{(0,\boldsymbol{\alpha}')} + [\theta_{2,(1,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}]p_{(1,\boldsymbol{\alpha}')}}, \\ \bar{\theta}_{2,(1,\boldsymbol{\alpha}')} = \theta_{2,(0,\boldsymbol{\alpha}')} + \frac{[\theta_{2,(1,\boldsymbol{\alpha}')} - \theta_{2,(0,\boldsymbol{\alpha}')}][\theta_{1,(1,\boldsymbol{\alpha}')} - \bar{\theta}_{1,(0,\boldsymbol{\alpha}')}]p_{(1,\boldsymbol{\alpha}')}}{[\theta_{1,(0,\boldsymbol{\alpha}')} - \bar{\theta}_{1,(0,\boldsymbol{\alpha}')}]p_{(0,\boldsymbol{\alpha}')} + [\theta_{1,(1,\boldsymbol{\alpha}')} - \theta_{1,(0,\boldsymbol{\alpha}')}]p_{(1,\boldsymbol{\alpha}')}}, \\ \bar{p}_{(1,\boldsymbol{\alpha}')} = \frac{[\theta_{2,(0,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}]p_{(0,\boldsymbol{\alpha}')} + [\theta_{2,(1,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}]p_{(1,\boldsymbol{\alpha}')}}{\bar{\theta}_{2,(1,\boldsymbol{\alpha}')} - \bar{\theta}_{2,(0,\boldsymbol{\alpha}')}}, \\ \bar{p}_{(0,\boldsymbol{\alpha}')} = p_{(0,\boldsymbol{\alpha}')} + p_{(1,\boldsymbol{\alpha}')} - \bar{p}_{(1,\boldsymbol{\alpha}')}. \end{cases} \tag{S4.41}$$

Then one can check that (S4.40) holds and further (S0.1) holds. Since in the above construction the choice of the 2^K item parameters $\{\theta_{1,(0,\boldsymbol{\alpha}')}, \theta_{2,(0,\boldsymbol{\alpha}')} : \boldsymbol{\alpha}' \in \{0,1\}^{K-1}\}$ are arbitrary, we have proved that the Q and associated model parameters are not generically identifiable.

S5. Proof of Theorem 4

We prove this theorem following a similar argument as the proof of Theorem 7 in Gu and Xu (2018a). Assume Q takes the form $Q = (Q_1^\top, Q_2^\top, (Q^\star)^\top)^\top$, where Q_1 and Q_2 have all diagonal elements being 1. Assume

$$\theta_{j,\boldsymbol{\alpha}} = f\Big(\beta_{j,0} + \sum_{k=1}^{K}\beta_{j,k}q_{j,k}\alpha_k + \sum_{k'=k+1}^{K}\sum_{k=1}^{K-1}\beta_{j,kk'}(q_{j,k}\alpha_k)(q_{j,k'}\alpha_{k'}) + \cdots + \beta_{j,12\cdots K}\prod_{k}(q_{j,k}\alpha_k)\Big),$$

where $f(\cdot)$ is some link function and when $f(\cdot)$ is the identify function, the model is the GDINA model. We first show that under Condition D, the $2^K \times 2^K$ matrices $T(Q_1, \boldsymbol{\Theta}_{Q_1})$ and $T(Q_2, \boldsymbol{\Theta}_{Q_2})$ both have full rank 2^K generically. It suffices to find some valid $\boldsymbol{\Theta}$ (i.e., $\boldsymbol{\Theta}_Q$) that gives

$$\det(T(Q_1, \boldsymbol{\Theta}_{Q_1})) \neq 0, \quad \det(T(Q_2, \boldsymbol{\Theta}_{Q_2})) \neq 0. \tag{S5.42}$$

The reason is as follows. (S5.42) would imply the polynomials defining the two matrix determinants are not zero polynomials in the Q-restricted parameter space. Therefore for almost all parameters, $T(Q_1, \boldsymbol{\Theta}_{Q_1})$ and $T(Q_2, \boldsymbol{\Theta}_{Q_2})$ would have full rank. Next we only focus on $T(Q_1, \boldsymbol{\Theta}_{Q_1})$. For every item $k = 1, \ldots, K$, we set $\beta_{k,k} = 1$, $\beta_{k,k'} = 0$ for any $k' \neq k$, and set all the interaction effects to zero. Then $T(Q_1, \boldsymbol{\Theta}_{Q_1})$ becomes identical to $T(I_K, \widehat{\boldsymbol{\Theta}}_{I_K})$ under a Q-matrix I_K with associated item parameters $\widehat{\boldsymbol{\Theta}}_{I_K}$ defined as follows: $\hat{\theta}_{\boldsymbol{e}_k,\boldsymbol{0}} = \beta_{k,0}$, and $\hat{\theta}_{\boldsymbol{e}_k,\boldsymbol{e}_k} = \hat{\theta}_{\boldsymbol{e}_k,\boldsymbol{1}} = \beta_{k,0} + \beta_{k,k}$ for $k \in \{1, \ldots, K\}$. It is not hard to see that $T(I_K, \widehat{\boldsymbol{\Theta}}_{I_K})$ can be viewed as a T-matrix under the DINA model with the Q-matrix equal to I_K, and guessing parameters $\beta_{k,0}$, slipping parameters $1 - \beta_{k,0} - \beta_{k,k}$

for $k = 1, \dots, K$. Therefore $T(I_K, \widehat{\mathbf{\Theta}}_{I_K})$ has full rank as argued in Step 1 of the proof of Theorem 1. So $T(Q_1, \mathbf{\Theta}_{Q_1})$ has full rank generically.

We next prove that if Condition E holds in addition, then any two different columns of $T(Q^\star, \mathbf{\Theta}_{Q^\star})$ are distinct generically. For $\boldsymbol{\alpha}$, $\boldsymbol{\alpha}' \in \{0,1\}^K$ and $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$, they at least differ in one element. Assume without loss of generality that $\alpha_k = 1 > 0 = \alpha'_k$. Then Condition E ensures that there is some item $j > 2K$ with $q_{j,k} = 1$. Under the general RLCM, this implies $\theta_{j,\boldsymbol{\alpha}} \neq \theta_{j,\boldsymbol{\alpha}'}$ generically. By Kruskal (1977), a matrix's Kruskal rank is the largest number I such that every set of I columns of the matrix are independent. When a matrix has full rank, its Kruskal rank equals its rank. By this definition, $T(Q^\star,$ $\mathbf{\Theta}_{Q^\star})$ has Kruskal rank at least 2 generically, and $T(Q_1, \mathbf{\Theta}_{Q_1})$, $T(Q_2, \mathbf{\Theta}_{Q_2})$ have Kruskal rank 2^K generically. Then for generic $\mathbf{\Theta}_Q$, we have

$$\operatorname{rank}_K\{T(Q_1, \mathbf{\Theta}_{Q_1})\} + \operatorname{rank}_K\{T(Q_2, \mathbf{\Theta}_{Q_2})\} + \operatorname{rank}_K\{T(Q^\star, \mathbf{\Theta}_{Q^\star})\} \geq 2 \times 2^K + 2. \tag{S5.43}$$

Applying Corollary 2 of Rhodes (2010) to this 2^K-class latent class model, we get $T(Q, \mathbf{\Theta}) = T(Q, \bar{\mathbf{\Theta}})$ and $\boldsymbol{p} = \bar{\boldsymbol{p}}$ up to column permutation. This proves generic identifiability of $(Q, \mathbf{\Theta}, \boldsymbol{p})$ in the model. Moreover, we also have the following form of the identifiable set

$$\begin{aligned} \boldsymbol{\vartheta}_Q \setminus \boldsymbol{\vartheta}_{non} = \{(\mathbf{\Theta}_Q, \boldsymbol{p}) : \det(T(Q_1, \mathbf{\Theta}_{Q_1})) \neq 0, \det(T(Q_2, \mathbf{\Theta}_{Q_2})) \neq 0, \\ T(Q^\star, \mathbf{\Theta}_{Q^\star}) \cdot \operatorname{Diag}(\boldsymbol{p}) \text{ has column vectors different from each other}\}. \end{aligned}$$

This is because when $(\mathbf{\Theta}_Q, \boldsymbol{p}) \in \boldsymbol{\vartheta}_Q \setminus \boldsymbol{\vartheta}_{non}$, the rank condition (S5.43) is satisfied and joint identifiability of $(Q, \mathbf{\Theta}_Q, \boldsymbol{p})$ follows.

S6. Proof of Theorem 5.

We prove the theorem in two steps. In the first step, we show that if Q is not generically complete, than it must take the following form (up to column/row permutations) for some $k > m$,

$$Q = \left(\begin{array}{ccc|ccc} q_{1,1} & \cdots & q_{1,k} & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ q_{m,1} & \cdots & q_{m,k} & * & \cdots & * \\ \hline 0 & \cdots & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & * & \cdots & * \end{array}\right) = \left(\begin{array}{c|c} Q_{11} & Q_{12} \\ \hline Q_{21} & Q_{22} \end{array}\right) = \begin{pmatrix} Q_1 \\ Q_2 \end{pmatrix}. \tag{S6.44}$$

The bottom-left submatrix $Q_{21} = \mathbf{0}_{(J-m)\times k}$. Any entry not in Q_{21} can be either 0 or 1. We introduce some definitions first. Given a Q-matrix Q, define a family S_Q of K finite sets $S_Q = \{\mathcal{A}_1, \mathcal{A}_1, \ldots, \mathcal{A}_K\}$, where $\mathcal{A}_k = \{1 \leq j \leq J : \ q_{j,k} = 1\}$ for each k. Then $\mathcal{A}_k$ denotes the set of items that require attribute k. For the family S_Q, a *transversal* is a system of distinct representatives from each of its elements $\mathcal{A}_1, \ldots, \mathcal{A}_K$. For example, for

$$Q = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix},$$

we have $S_Q = \{\mathcal{A}_1 = \{1,3\},\ \mathcal{A}_2 = \{1,2\},\ \mathcal{A}_3 = \{2,3\}\}$. Then $(1,2,3)$ is a valid transversal of S_Q, and so as $(3,1,2)$; but $(1,1,2)$ is not a transversal. Now it is not hard to see that, the assumption that Q is not generically complete is equivalent to the following statement $H^\star$,

$H^\star$. Given Q, the family S_Q does not have a valid transversal.

Then by Hall's Marriage Theorem (Hall, 1967), the nonexistence of a transversal indicates the failure of the marriage condition. So there must exist a subfamily $W \subseteq S_Q$ such that $|W| > |\bigcup_{\mathcal{A}\in W} \mathcal{A}|$. More specifically, this means there exist some $l_1, l_2, \ldots, l_k \in \{1, \ldots, K\}$ and $W = \{\mathcal{A}_{l_1}, \ldots, \mathcal{A}_{l_k}\}$ such that

$$|W| = k > |\mathcal{A}_{l_1} \cup \cdots \cup \mathcal{A}_{l_k}| \overset{\text{def}}{=} m.$$

In other words, we have shown that there exist some attributes, the number of which (e.g., k) exceeds the number of items that require any of these attributes (e.g., m). This is exactly saying that Q has to take the form of (S6.44) with $k > m$ after some column/row permutation.

In the second step, we show that if Q takes the form of (S6.44) with $k > m$, then $(Q, \boldsymbol{\Theta}, \boldsymbol{p})$ under general RLCMs are not generically identifiable. Now we define another potentially different $\bar{Q}$ as

$$\bar{Q} = \left(\begin{array}{c|c} Q_{11} & \bar{Q}_{12} \\ \hline Q_{21} & Q_{22} \end{array}\right) = \begin{pmatrix} \bar{Q}_1 \\ Q_2 \end{pmatrix}, \quad \text{where } \bar{Q}_{12} = \mathbf{1}_{m\times(K-k)}.$$

Then given arbitrary $(\mathbf{\Theta}, \boldsymbol{p})$ associated with Q, we set $\bar{\theta}_{j,\boldsymbol{\alpha}} = \theta_{j,\boldsymbol{\alpha}}$ for every $j = m + 1, \ldots, J$ and every $\boldsymbol{\alpha} \in \{0,1\}^K$. Because Q_{21} is a $(J-m)\times k$ zero matrix, we claim that under the current construction, the original 2^J constraints in (S0.1) are satisfied as long as the following constraints are satisfied

$$\forall \boldsymbol{\alpha}' = (\alpha_{k+1}, \ldots, \alpha_K) \in \{0,1\}^{K-k}, \quad \forall \boldsymbol{r}' = (r_1, \ldots, r_m) \in \{0,1\}^m,$$

$$\sum_{\boldsymbol{\alpha}^\star \in \{0,1\}^k} T_{\boldsymbol{r}', (\boldsymbol{\alpha}^\star, \boldsymbol{\alpha}')}(Q_1, \mathbf{\Theta}_{Q_1}) \cdot p_{(\boldsymbol{\alpha}^\star, \boldsymbol{\alpha}')} = \sum_{\boldsymbol{\alpha}^\star \in \{0,1\}^k} T_{\boldsymbol{r}', (\boldsymbol{\alpha}^\star, \boldsymbol{\alpha}')}(\bar{Q}_1, \bar{\mathbf{\Theta}}_{\bar{Q}_1}) \cdot \bar{p}_{(\boldsymbol{\alpha}^\star, \boldsymbol{\alpha}')}.$$

This claim can be shown following a similar argument as that in Step 3 of the proof of Theorem 2 (b.1) and (c). Then the above system of equations contain $2^{K-k} \times 2^m$ constraints, while under the general RLCMs the number of free variables in $(\bar{\mathbf{\Theta}}, \bar{\boldsymbol{p}})$ involved is

$$\begin{aligned} &\Big|\{\bar{p}_{\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \{0,1\}^K\} \bigcup \{\bar{\theta}_{j,\boldsymbol{\alpha}} : j \in \{1,\ldots,m\}, \boldsymbol{\alpha} \in \{0,1\}^K\}\Big| \\ &= 2^K + 2^{K-k} \times \Big(\sum_{j=1}^m 2^{q_{j,1}+\cdots+q_{j,k}}\Big) \geq 2^K + 2^{K-k} \times m. \end{aligned}$$

Under the assumption $m < k$, we have that the number of constraints $2^{K-k} \times 2^m$ is smaller than the number of variables to solve (which is lower bounded by $2^{K-k} \times (2^k + m)$), because $2^m < 2^k + m$. So there exist infinitely many different sets of solutions of $(\bar{\mathbf{\Theta}}, \bar{\boldsymbol{p}})$ associated with $\bar{Q}$ such that $T(Q, \mathbf{\Theta})\boldsymbol{p} = T(\bar{Q}, \bar{\mathbf{\Theta}})\bar{\boldsymbol{p}}$. Therefore $(Q, \mathbf{\Theta}, \boldsymbol{p})$ are not generically identifiable and the proof of the theorem is complete.

S7. Proof of Proposition 4

We show the conclusion following a similar argument as the proof of Proposition 1 in Xu and Shang (2018). To establish the bound (6.11) in the proposition, we check the technical conditions in Theorem 1 in Shen et al. (2012). We first define some notations. For a family of probability mass functions $\mathcal{F}$, define $H(\cdot, \mathcal{F})$ to be the bracketing Hellinger metric entropy of $\mathcal{F}$. We call a finite set of function pairs $S(\epsilon, n) = \{(f_1^l, f_1^u), \ldots, (f_n^l, f_n^u)\}$ a *Hellinger ϵ-bracketing* of $\mathcal{F}$ if the L_2 norm $\left\|\sqrt{f_i^l} - \sqrt{f_i^u}\right\| \leq \epsilon$ for all $i = 1, \ldots, n$; and further fur any $f \in \mathcal{F}$, there is an i such that $f_i^l \leq f \leq f_i^u$. The *bracketing Hellinger metric entropy* is defined to be the logarithm of the cardinality of the ϵ-bracketing with the smallest size, namely $H(\cdot, \mathcal{F}) = \log\min\{n : S(\epsilon, n)\}$. We next argue that the size of the parameter space of $(\boldsymbol{\Theta}, \boldsymbol{p})$ is well controlled under the Hellinger metric. Recall S is defined in the main text before Proposition 4, and we define $\mathcal{B}_S = \mathcal{F}_S \cap \{h(\boldsymbol{\eta}, \boldsymbol{\eta}^0) \leq 2\epsilon\}$ as the local parameter space with $\boldsymbol{\eta} = (\boldsymbol{B}, \boldsymbol{p})$ denoting general model parameters and $\boldsymbol{\eta}^0 = (\boldsymbol{B}^0, \boldsymbol{p}^0)$ denoting the true model parameters. According to the argument in the proof of Proposition 1 in Xu and Shang (2018), in the considered scenario with fixed J and K, for any $\epsilon < 1$ and any $t \in (\epsilon/2^4, \epsilon)$, there is $H(t, \mathcal{B}_S) \leq c\log(J2^K)|S|\log(2\epsilon/t)$; indeed, there is $H(t, \mathcal{B}_S) = O(\log(2\epsilon/t))$ uniformly for any S, ϵ and t.

With this upper bound on the Hellinger bracketing entropy, we can apply Theorem 1 in Shen et al. (2012) to obtain

$$\mathbb{P}(\widehat{Q} \neq Q^0) \leq \mathbb{P}(\widehat{\boldsymbol{\eta}} \neq \widehat{\boldsymbol{\eta}}^0) \leq c_2 \exp\{-c_1 N C_{\min}(\boldsymbol{\Theta}^0, \boldsymbol{p}^0)\},$$

where $C_{\min}(\boldsymbol{\Theta}^0, \boldsymbol{p}^0) := \inf_{\boldsymbol{\eta}:\,|S|\leq m, S\neq S_0} h^2(\boldsymbol{\eta}, \boldsymbol{\eta}^0)$. The above display is the desired (6.11) in the proposition.

Next we show that when the proposed sufficient conditions for joint strict identifiability hold, the $C_{\min}(\boldsymbol{\Theta}^0, \boldsymbol{p}^0)$ in (6.11) is bounded away from zero by some positive constant. When the proposed conditions for joint strict identifiability (such as Conditions A, B and C under DINA model are satisfied), the $(\boldsymbol{B}^0, \boldsymbol{p}^0)$ here are strictly identifiable. The consequence is that there exists a constant $\delta > 0$ such that $h^2(\boldsymbol{\eta}, \boldsymbol{\eta}^0) \geq \delta$, where the m denotes the number of free parameters under the Q^0 and the RLCM specification. Therefore,

$$C_{\min}(\boldsymbol{\Theta}^0, \boldsymbol{p}^0) \geq \inf_{\boldsymbol{\eta}:\,|S|\leq m, S\neq S_0} \frac{h^2(\boldsymbol{\eta}, \boldsymbol{\eta}^0)}{2m} \geq \frac{\delta}{2m} > 0,$$

so $C_{\min}(\boldsymbol{\Theta}^0, \boldsymbol{p}^0) \geq c_0$ for some positive constant c_0 holds. This proves the conclusion that under the proposed strict identifiability conditions, the finite sample error bound $\mathbb{P}(\widehat{Q} \neq Q^0)$ has an exponential rate. This completes the proof of the proposition.

S8. Simulation Studies

In this section, we provide more simulation results to verify the developed identifiability theory. In Section S8.1, we perform simulation studies to verify Theorems 1 and 2 for the DINA model. In Section S8.2, we perform simulation studies to verify Theorems 3 and 4 for the GDINA model. The Matlab code for performing the simulation studies are available at `https://github.com/yuqigu/Identify_Q`.

To better illustrate the identifiability or non-identifiability phenomena of Q-matrix, in some of the following simulation studies, we conduct exhaustive search of all possible Q-matrices of a certain size 5×2. Specifically, consider the set of all the 5×2 binary Q-matrices other than those containing some all-zero row vectors. If treating two Q-matrices that are identical up to permuting the two columns as equivalent (because they are indeed equivalent in terms of model identifiability), then there are in total 121 types of Q-matrices. We denote such a set of Q-matrices by $\text{Exhaus}(Q_{5\times 2})$, and denote its elements by $Q^1, Q^2, \ldots, Q^{121}$. For example, the first three and the last three Q-matrices in $\text{Exhaus}(Q_{5\times 2})$ are

$$Q^1 = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}; \quad Q^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}; \quad Q^3 = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}; \quad \cdots\cdots$$

$$Q^{119} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad Q^{120} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}; \quad Q^{121} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

The complete list of the 121 Q-matrices in the set Exhaus($Q_{5\times 2}$) is available in the Matlab file `Q_aa.mat` at `https://github.com/yuqigu/Identify_Q`.

In the exhaustive-search scenario, to illustrate the identifiability/non-identifiability phenomenon, we will generate data using some particular Q-matrix, and fit the dataset using all the 121 candidate Q-matrices in Exhaus($Q_{5\times 2}$) and plot the log-likelihood values corresponding to all these 121 Q-matrices. Investigating whether the true data-generating Q-matrix achieves the maximum of the likelihood would help gain insight into whether this true Q-matrix is identifiable in the considered practical setting. We will see from these simulations how the developed identifiability theory matches the practice.

S8.1 Two-Parameter RLCM: DINA Model

In this section, we carry out four simulation studies.

Study I: When Q-matrix satisfies the necessary and sufficient conditions A, B and C for strict identifiability.

In this simulation study, we choose those Q-matrices from Exhaus($Q_{5\times 2}$) that satisfies the proposed necessary and sufficient identifiability conditions A, B and C in

Theorem 1 of the main text. In particular, after rearranging rows, there are exactly two forms the 5×2 Q-matrix that satisfies A, B and C. Their representatives are Q^{18} and Q^{15} as follows,

$$Q^{18} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}; \qquad Q^{15} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

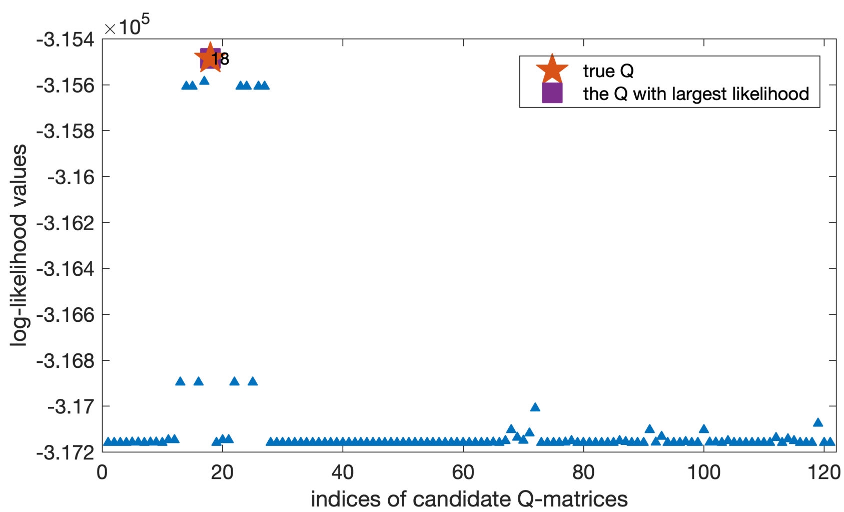
Note that Q^{18} contains only on identity submatrix I_2, while Q^{15} contains two copies of submatrix I_2. As introduced prior to this section S8.1, we generate datasets with sample size $N = 10^5$ with true Q-matrix being Q^{18} and Q^{15}, respectively; and for each case, we run EM algorithm with several random initializations to fit the dataset with all the 121 Q-matrixes in Exhaus($Q_{5\times2}$) and obtain their log-likelihood values.

Figure 1a and 1b present the log-likelihood plots, with x-axis denoting the indices of the 121 candidate Q-matrices in Exhaus($Q_{5\times2}$), and y-axis denoting the log-likelihood values. Each blue triangle denotes a candidate Q-matrix; the red star denotes the true data-generating Q-matrix, and the purple square denotes the Q-matrix that achieves the largest likelihood.

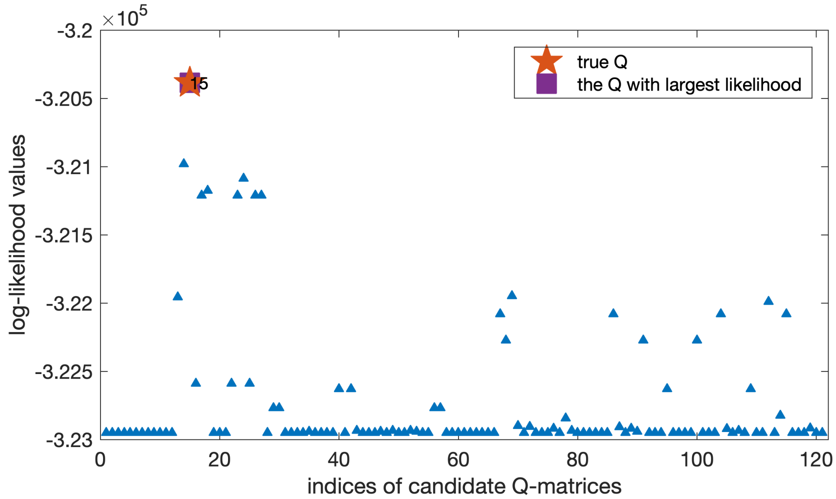
We can see from these two plots in Figure 1 that when the true data-generating Q-matrix (Q^{15} and Q^{18}) satisfies our proposed conditions A, B and C, it indeed achieves the largest likelihood compared to all other possible candidate Q-matrices. Therefore for any algorithm seeking the maximum likelihood estimator of $(Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$, the true

Q-matrix can be identified and any other Q-matrix will not be confused with the true Q. Another observation from Figure 1a and 1b is that, for Q^{15} that contains one more identity submatrix I_2 than Q^{18}, the true Q can be relatively better distinguished from the other Q's due to the larger gap in the likelihood values. This phenomenon might imply that the more identity submatrices the true data-generating Q-matrix contain, the easier the estimation for the true structure would be.

Figure 1: DINA: exhaustive search in the set of 5×2 Q-matrices with a true Q-matrix satisfying Conditions A, B and C in Theorem 1.



(a) true Q containing one I_2: $Q^{18} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}^\top$



(b) true Q containing two I_2's: $Q^{15} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}^\top$

Study II: When Q-matrix does not satisfy all of Conditions A, B, C but satisfies conditions in Theorem 2 for generic identifiability.

In this simulation study, we take the data-generating Q-matrix from Exhaus($Q_{5\times 2}$) that do NOT satisfy some of Conditions A, B and C, but satisfy the conditions in Theorem 2 for joint generic identifiability of $(Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$. In particular, for the considered case of $K = 2$, the only possibility for (global) generic identifiability is scenario (b.2) described in Theorem 2, where Condition C is violated and some column of Q contains only two entries of "1". After rearranging the rows of Q, it is not hard to see that there is only one possible case of the form of Q leading to generic identifiability, and the following Q^5 is a representative,

$$Q^5 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}. \tag{S8.45}$$

The log-likelihood value plot is presented in Figure 2. One can see in this generically identifiable scenario, with randomly generated true parameters, the true Q-matrix Q^5 achieves the largest likelihood and hence can be identified from data. We point out that although only the result of one simulated dataset is presented here, the generically identifiable Q-matrix (as the true Q-matrix) generally can achieve the largest likelihood among all the candidate Q-matrices, based on our experience in various simulations.

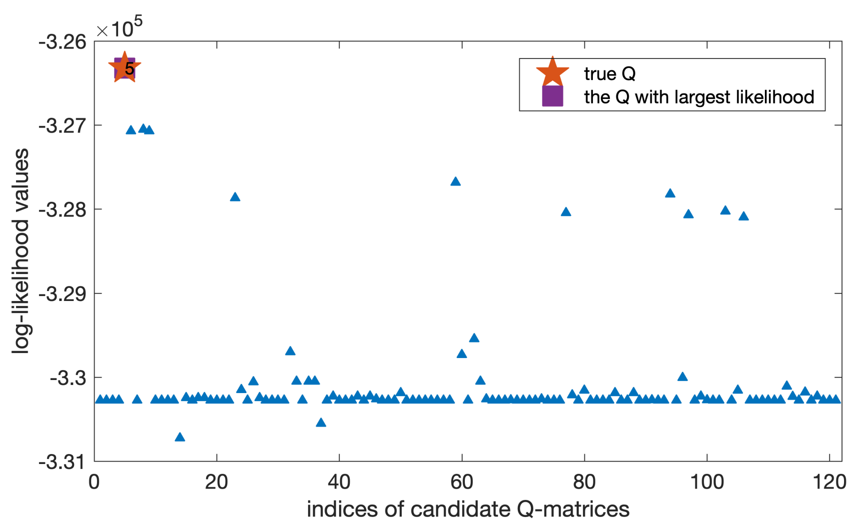


Figure 2: DINA: exhaustive search in 5×2 Q-matrices with a true Q-matrix Q^5 in (S8.45) generically identifiable, corresponding to scenario (b.2) in Theorem 2.

Study III: When Q-matrix does not even lead to local identifiability.

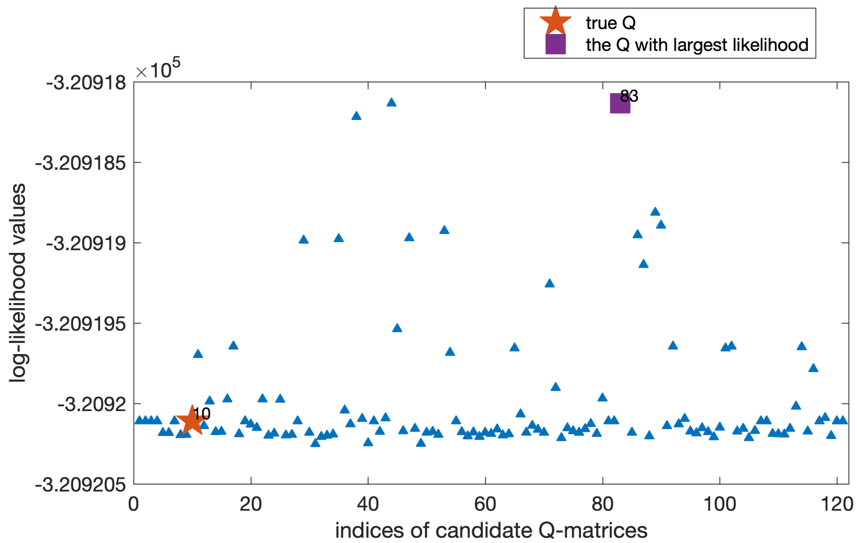
In this simulation study, we take the data-generating Q-matrix from Exhaus($Q_{5\times 2}$) that *do not even lead to local identifiability.* That is, under such true Q-matrix, even in a small neighborhood of the true parameters, there exist infinitely many different alternative parameters that are not distinguishable from the true one.

Consider the following three different forms of Q-matrices from the set Exhaus($Q_{5\times 2}$),

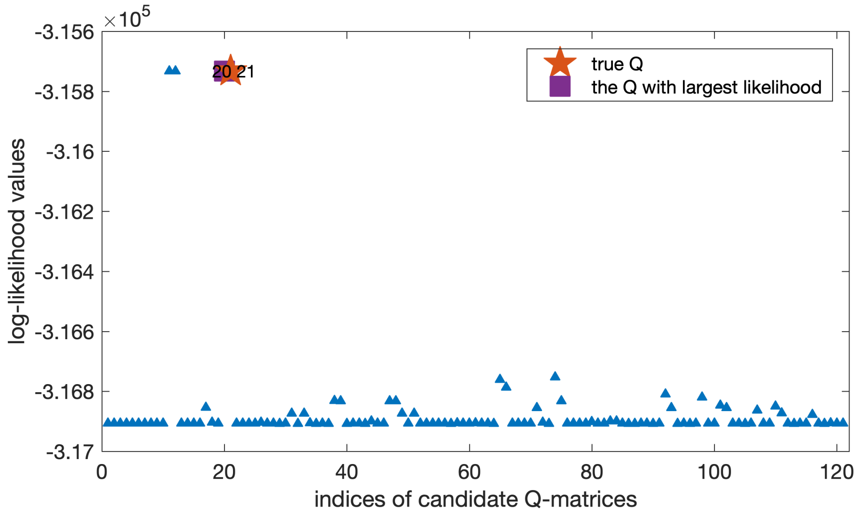
$$Q^{10} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}; \quad Q^{21} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 0 & 1 \end{pmatrix}; \quad Q^{55} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix},$$

where Q^{10} contains only one entry of "1" in one column, Q^{21} is incomplete (i.e., lacks I_2), and Q^{55} contains an all-zero column. Their corresponding log-likelihood plots in the exhaustive-search scenario are presented in Figure 3a, 3b and 3c. One can see from these plots that in these no even locally identifiable settings, the true data-generating Q-matrix does not achieve the maximum of the likelihood. Instead, many other alternative Q-matrices would have larger likelihood, and a wrong Q-matrix will be selected as the maximum likelihood estimator.

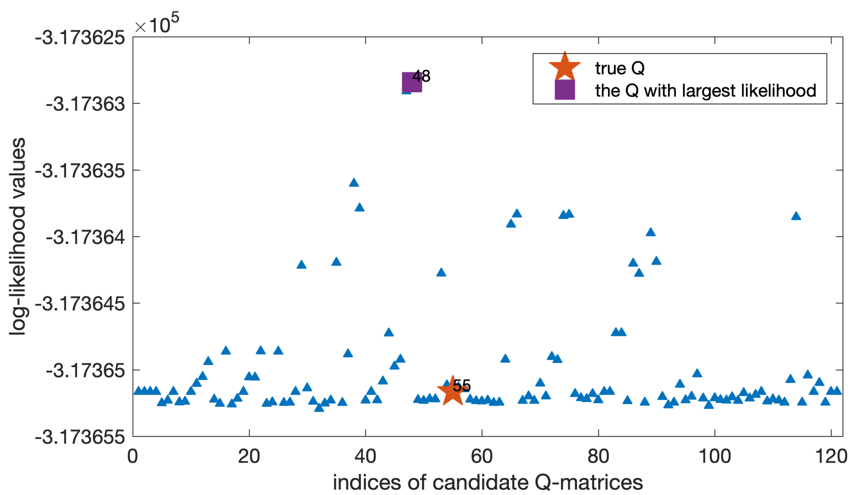
Figure 3: DINA: exhaustive search in 5×2 Q-matrices with a true Q-matrix *not even locally* identifiable, corresponding to scenario (b.1) in Theorem 2.



(a) true Q not even locally identifiable: $Q^{10} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}^\top$



(b) true Q not even locally identifiable: $Q^{21} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$



(c) true Q not even locally identifiable: $Q^{55} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$

Study IV: Verifying necessity of Condition A "completeness".

We verify the necessity of Condition A "*completeness*" of the Q-matrix for identifiability. Consider two settings of incomplete Q-matrices, Q_1 with $(K, J) = (3, 20)$ and Q_2 with $(K, J) = (5, 20)$. For $i = 1, 2$, for the matrix $Q = Q_i$ and arbitrary DINA model parameters $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$, we follow our theoretical derivations to construct two alternative Q-matrices $Q' = Q'_i$ and $Q'' = Q''_i$ and corresponding parameters $(\boldsymbol{c}^{'}, \boldsymbol{g}^{'}, \boldsymbol{p}^{'})$ and $(\boldsymbol{c}^{''}, \boldsymbol{g}^{''}, \boldsymbol{p}^{''})$. Then we compute the marginal probabilities for all the possible $2^{20} \approx 10^6$ response patterns under each of the Q, Q' and Q'', which characterize the distribution of the 20-dimensional binary vector $\boldsymbol{R}$. We give visualization plots to show how these different Q-matrices and different model parameters lead to exactly the same distribution of the observed responses $\boldsymbol{R}$.

$$
Q_1 = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}_{20\times 3}
\qquad
Q_1' = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{1} & \mathbf{1} \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}_{20\times 3}
\qquad
Q_1'' = \begin{pmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 \\
1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 1 & 1 \\
1 & 1 & 1
\end{pmatrix}_{20\times 3}
\tag{S8.46}
$$

$$
Q_2 = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
\mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1
\end{pmatrix}_{20\times 5}
\qquad
Q_2' = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1
\end{pmatrix}_{20\times 5}
\qquad
Q_2'' = \begin{pmatrix}
1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
\mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 0 \\
1 & 1 & 1 & 1 & 1
\end{pmatrix}_{20\times 5}
\tag{S8.47}
$$

First, consider the following Q_1 with $(K, J) = (3, 20)$ in (S8.46), which is incomplete because its row vectors does not contain the unit vector $(0, 0, 1)$. For arbitrarily generated parameters $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$, we set $\boldsymbol{c}^{''} = \boldsymbol{c}^{'} = \boldsymbol{c}$ and $\boldsymbol{g}^{''} = \boldsymbol{g}^{'} = \boldsymbol{g}$ and set the proportion parameters as follows,

$$\begin{cases} p'_{(011)} = 0, \\ p'_{(010)} = p_{(010)} + p_{(011)}, \\ p'_{\boldsymbol{\alpha}} = p_{\boldsymbol{\alpha}}, \ \forall \boldsymbol{\alpha} \neq (011), (010); \end{cases} \qquad \begin{cases} p''_{(001)} = p^2_{(011)} = p^2_{(111)} = 0, \\ p''_{(000)} = p_{(000)} + p_{(001)}, \\ p''_{(010)} = p_{(010)} + p_{(011)}, \\ p''_{(110)} = p_{(110)} + p_{(111)}, \\ p''_{\boldsymbol{\alpha}} = p_{\boldsymbol{\alpha}}, \ \forall \boldsymbol{\alpha} = (100), (101). \end{cases} \tag{S8.48}$$

We define a notation $\Gamma(Q)$ to briefly explain the rationale behind the above constructions. The $\Gamma(Q)$ is a $J \times 2^K$ binary matrix defined based on Q. The columns and rows of $\Gamma(Q)$ are indexed by the J items and the 2^K possible attribute patterns, respectively; and the $(j, \boldsymbol{\alpha})$th entry of it is defined to be $\Gamma_{j,\boldsymbol{\alpha}}(Q) = I(\boldsymbol{\alpha} \succeq \boldsymbol{q}_j)$. An important observation is that, due to the forms of Q, Q' and Q'', the unique column vectors in $\Gamma(Q)$ form a subset of those of $\Gamma(Q')$; and further the unique column vectors of $\Gamma(Q')$ form a subset of those of $\Gamma(Q'')$. Therefore, to construct $\boldsymbol{p}'$ such that $(Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$ and $(Q', \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}')$ that are non-distinguishable, we only need to set $p'_{\boldsymbol{\alpha}} = 0$ for those $\boldsymbol{\alpha}$ whose corresponding column vector in $\Gamma(Q')$ does not appear as the column vector of $\Gamma(Q)$; and let the proportions (in vector $\boldsymbol{p}'$) of other attribute patterns to absorb the proportions of these $\boldsymbol{\alpha}$'s in the vector $\boldsymbol{p}'$. The proportions $\boldsymbol{p}''$ under Q'' are constructed similarly. This is exactly how Equation (S8.48) are derived. For the Q_2, Q_2' and Q_2'' defined in (S8.47), we construct the proportion parameters $\boldsymbol{p}'$ under Q_2' and $\boldsymbol{p}''$ under Q_2'' following the same rationale; the details of defining them are omitted but their values are later revealed in Figure 5(c).

In Figure 4, we visualize the non-identifiability phenomenon of Q_1. In Figure 4(a), we plot the differences of proportions parameters under the alternative models and the true model with Q_1. The red dotted line with "×" plots the values $\boldsymbol{p}' - \boldsymbol{p} = (p_{000}' - p_{000},\ p_{001}' - p_{001},\ p_{010}' - p_{010},\ p_{011}' - p_{011},\ p_{100}' - p_{100},\ p_{101}' - p_{101},\ p_{110}' - p_{110},\ p_{111}' - p_{111})$ correspondent to the 8 attribute patterns; and the green dotted line with "+" plots $\boldsymbol{p}'' - \boldsymbol{p}$. Despite these three sets of parameters are quite different, the 2^{20}-dimensional vector of marginal probabilities of $\boldsymbol{R}$ are exactly the same, as shown in plots (b) and (c) of Figure 4. In particular, in plot (b), the x-axis presents the indices of the response patterns in $\boldsymbol{r} \in \{0,1\}^J$, the y-axis presents the values of $\mathbb{P}(\boldsymbol{R} = \boldsymbol{r} \mid Q, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$, where the blue circles denote those under $(Q_1, \boldsymbol{p})$, red "×" for $(Q_1', \boldsymbol{p}')$, and green "+" for $(Q_1'', \boldsymbol{p}'')$. Plot (c) of Figure 4 is a zoomed-in version of plot (b), by only showing those marginal probabilities in $[0.2 \times 10^{-4}, 2 \times 10^{-4}]$, which contains around 7×10^3 response patterns. One can roughly see from both plots (b) and (c) that the three underlying parameters yield identical distribution of the response vector. Indeed, the computation carried out using Matlab yields

$$\max_{\boldsymbol{r}\in\{0,1\}^{20}} |\mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_1, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_1', \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}')| = 2.17 \times 10^{-19},$$

$$\max_{\boldsymbol{r}\in\{0,1\}^{20}} |\mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_1, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_1'', \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}'')| = 4.34 \times 10^{-19},$$

which are both smaller than the machine epsilon (machine error) of Matlab 2.22×10^{-16}. This confirms that Q_1 defined in (S8.46) is not identifiable.

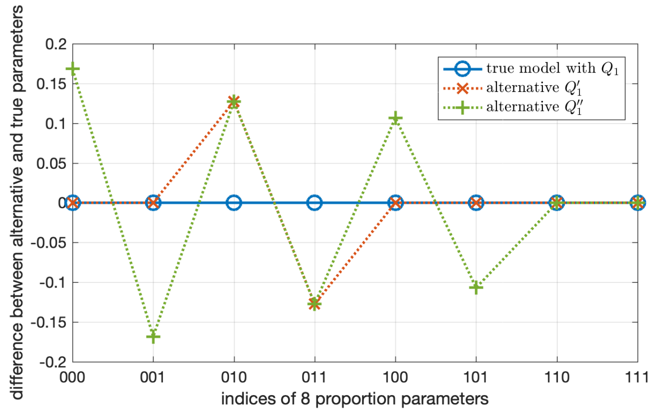
Figure 5 shows the analogous results for Q_2 of size 20×5. Plot (a) in Figure 5 shows the difference of the $2^5 = 32$-dimensional proportion parameters under alternative and true Q-matrices, and plots (b) and (c) give marginal probabilities of $\boldsymbol{R}$. The computation using Matlab gives

$$\max_{\boldsymbol{r}\in\{0,1\}^{20}} |\mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_2, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_2', \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}')| = 2.17 \times 10^{-19},$$

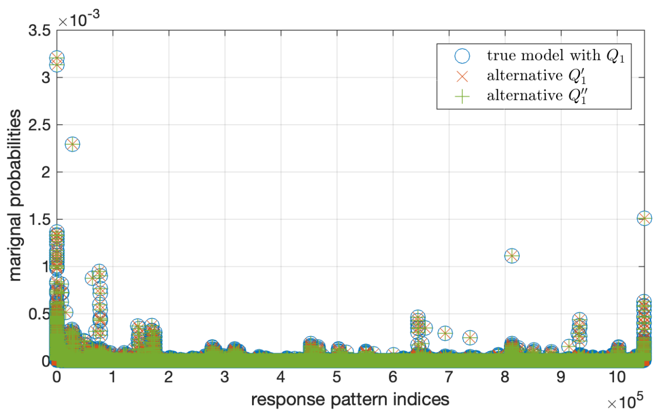
$$\max_{\boldsymbol{r}\in\{0,1\}^{20}} |\mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_2, \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_2'', \boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p}'')| = 6.51 \times 10^{-19},$$

which are also both smaller than the machine error 2.22×10^{-16} of Matlab. This verifies the non-identifiability of Q_2 defined in (S8.47).

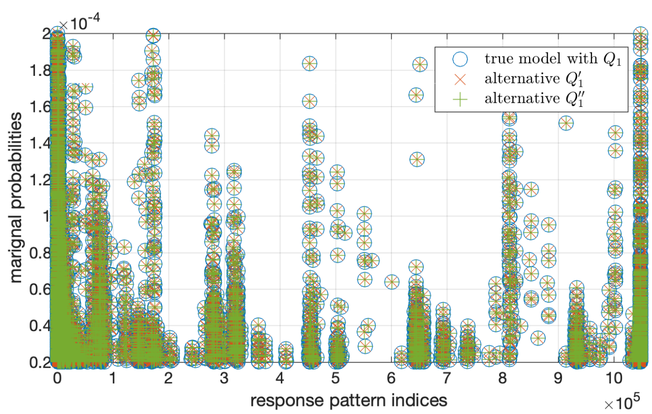
Figure 4: DINA: true Q-matrix of size 20×3 is not complete and hence not identifiable.



(a) $K = 3$ and $J = 20$, three sets of parameters

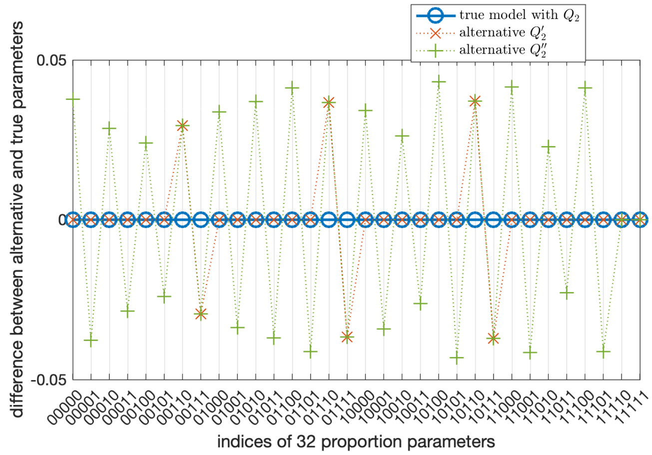


(b) $K = 3$ and $J = 20$, $|\{0,1\}^{20}| = 2^{20}$ response probabilities

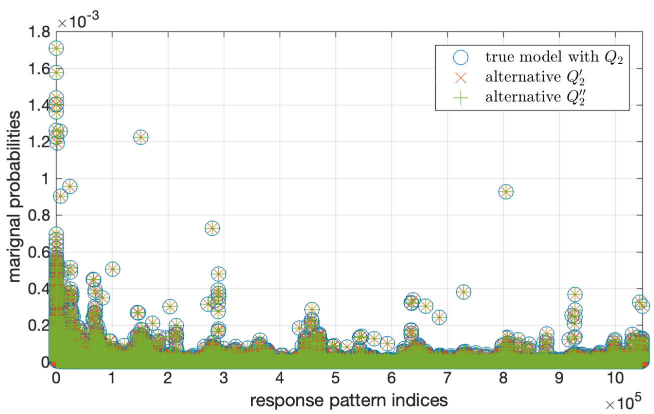


(c) $K = 3$ and $J = 20$, response probabilities zoomed in

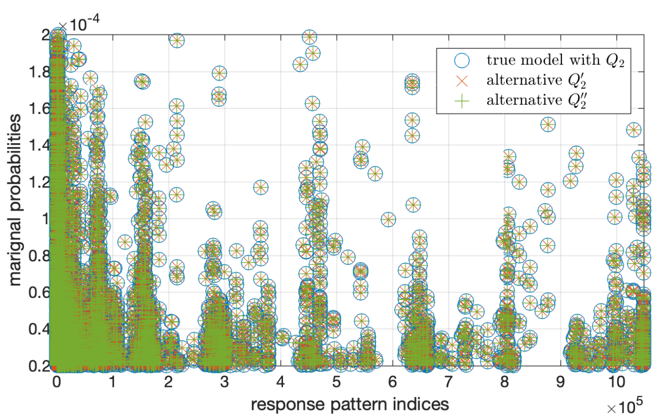
Figure 5: DINA: true Q-matrix of size 20×5 is not complete and hence not identifiable.



(a) $K = 5$ and $J = 20$, three sets of parameters



(b) $K = 5$ and $J = 20$, $|\{0,1\}^{20}| = 2^{20}$ response probabilities



(c) $K = 5$ and $J = 20$, response probabilities zoomed in

S8.2 General RLCM: GDINA Model

In this section, we design simulation studies to verify the proposed identifiability conditions under the GDINA model introduced in Example 2. In Study V, we use exhaustive search within 5×2 Q-matrices to verify the sufficient conditions in Theorem 4. In Study VI and Study VII, we verify the necessary conditions in Theorem 3.

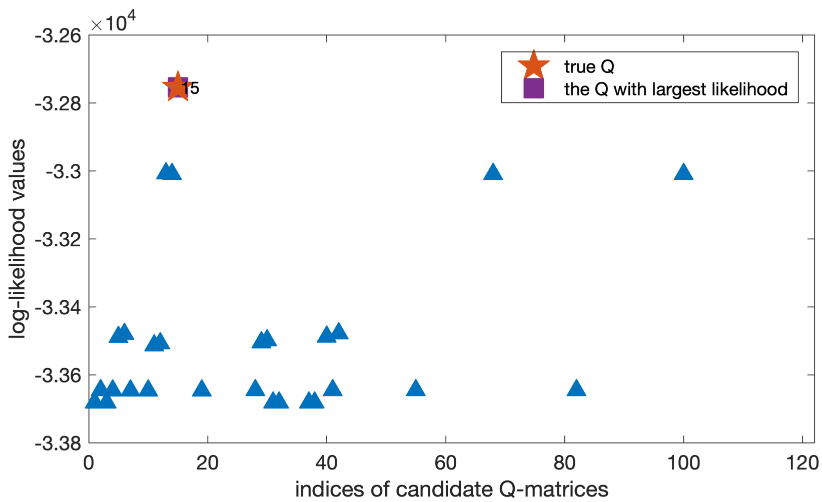
Study V: When Q-matrix satisfies Conditions D, E for generic identifiability.

Within the set of 5×2 Q-matrices Exhaus($Q_{5\times 2}$), if Q satisfies the sufficient conditions D and E for generic identifiability under the GDINA model, then other than the all-one Q-matrix Q^{121} which corresponds to the unrestricted latent class model, Q can take the forms of Q^{15}, Q^{18}, Q^{27}, Q^{54}, and Q^{81} (up to rearrangement of rows and columns). When using some Q-matrix to generate data, we also set the sample size to $N = 10^5$ and randomly set the true parameters which satisfy the monotonicity constraint (2.1) in the main text. In plots (a), (b), (c), (d) and (e) in Figure 6, we present the exhaustive search results when the true data-generating Q-matrix is Q^{15}, Q^{18}, Q^{27}, Q^{54}, or Q^{81}. We point out that for GDINA model, in each scenario, we did not plot all the 121 Q-matrices' log-likelihood values, although we fit all the 121 ones to the simulated data. Instead, we only plot those Q-matrices under which the estimated parameters satisfies the stringent monotonicity constraint

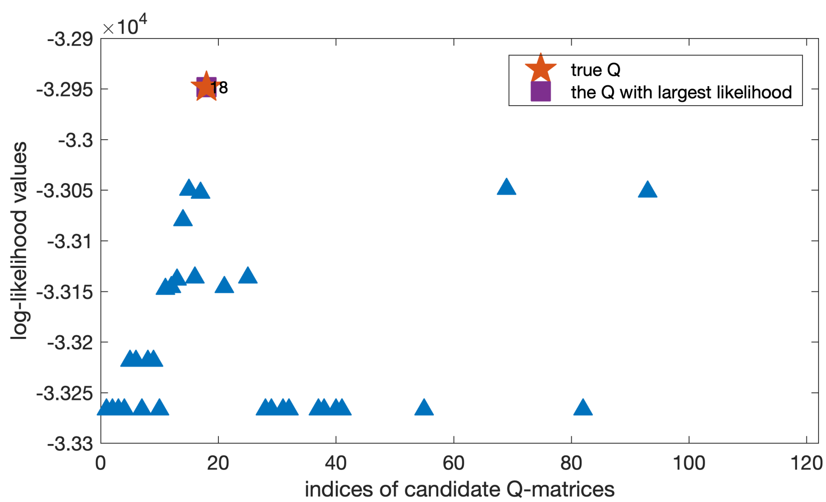
$$\theta_{j,\boldsymbol{\alpha}} > \theta_{j,\boldsymbol{\alpha}'} \quad \text{if} \quad \boldsymbol{\alpha} \odot \boldsymbol{q}_j \succ \boldsymbol{\alpha}' \odot \boldsymbol{q}_j. \tag{S8.49}$$

This constraint is stronger than requiring merely (2.1), and it is often imposed in practice when fitting the general RLCM that models the main and interaction effects of the latent attributes; for example, see the LCDM proposed in Henson et al. (2009). So each blue triangle in each plot of Figure 6 corresponds to a Q-matrix with estimated $\boldsymbol{\Theta}$ satisfying (S8.49). We can see from the five plots in Figure 6 that when the generic identifiability conditions D and E are satisfied, the true data-generating Q-matrix achieves the maximum of the data likelihood compared to all the candidate Q-matrices of the same size.

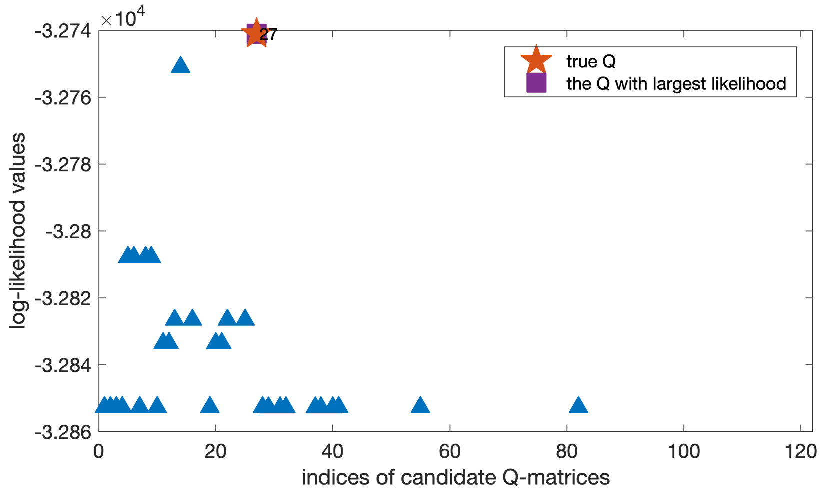
Figure 6: GDINA: exhaustive search in 5×2 Q-matrices with a true Q satisfying Conditions D and E.



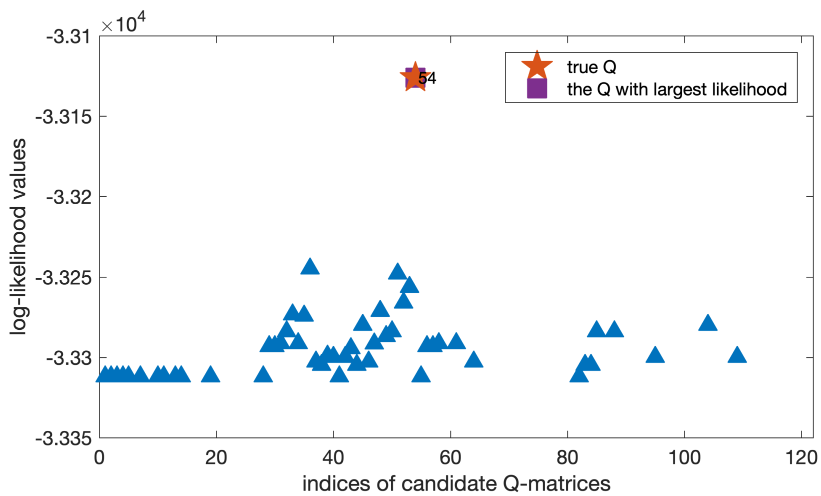
(a) GDINA: generically identifiable: $Q^{15} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 1 \end{pmatrix}^{\top}$



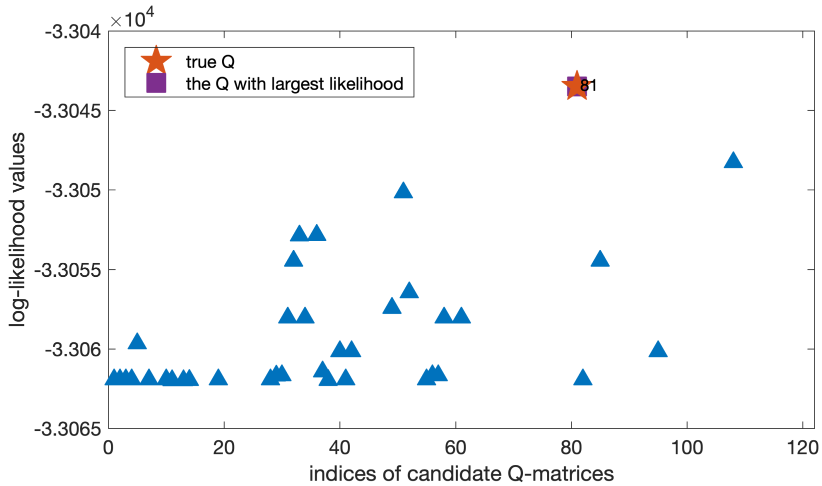
(b) GDINA: generically identifiable: $Q^{18} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \end{pmatrix}^{\top}$



(c) GDINA: generically identifiable: $Q^{27} = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$



(d) GDINA: generically identifiable: $Q^{54} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}^\top$



(e) GDINA: generically identifiable: $Q^{81} = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$

Study VI: When Q-matrix does not even lead to local generic identifiability.

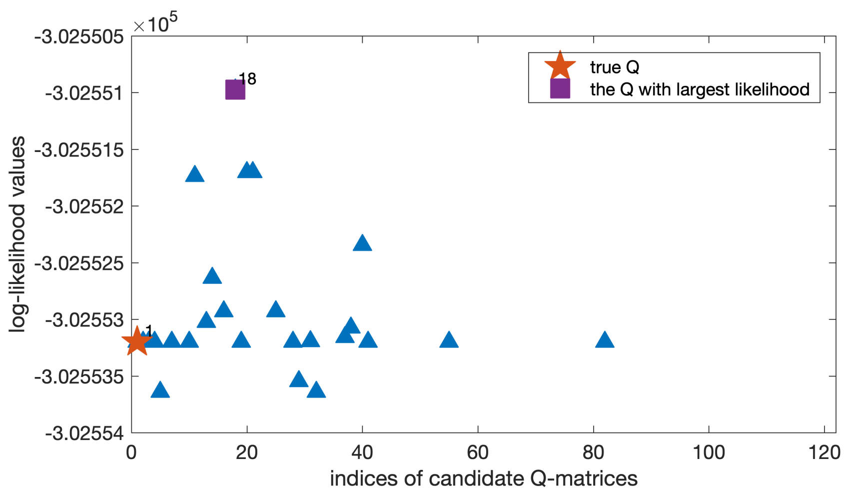
We now use the not even locally generically identifiable Q-matrices Q^1, Q^2, or Q^3 to generate the data, and perform the exhaustive search among Exhaus($Q_{5\times 2}$). The log-likelihood plots along with the forms of the data generating matrices Q^1, Q^2, Q^3 are presented in Figure 7. Similar to the previous Study V, here in each scenario we only plot those Q-matrix whose estimated $\mathbf{\Theta}$ parameters satisfy the stringent monotonicity constraint (S8.49). One can see from the plots in Figure 7 that these Q^1, Q^2, Q^3 do not maximize the data likelihood, implying severe non-identifiability. Note that for Figure 7(b) corresponding to Q^2, there are only two Q-matrices satisfying the constraint (S8.49) among the 121 Q-matrices fitted to the data; these two Q-matrices are the true

Q-matrix Q^2 and another Q-matrix Q^{56},

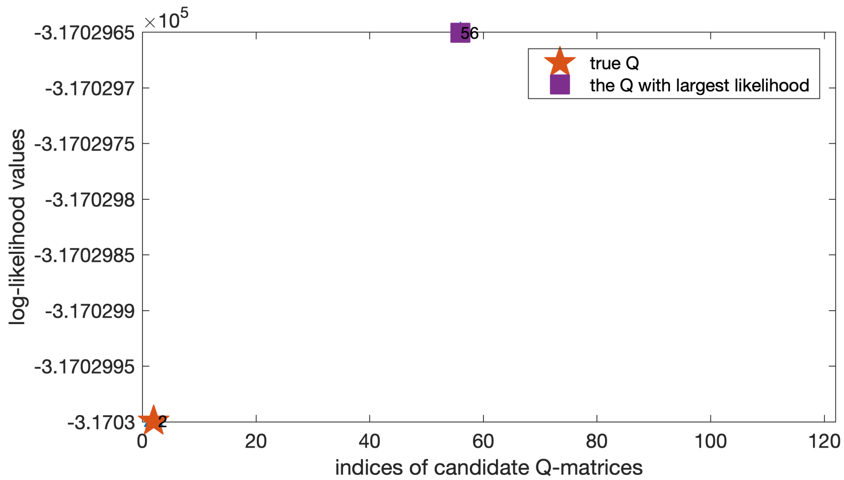
$$Q^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{pmatrix}, \qquad Q^{56} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Note that even there are only two Q-matrices satisfying the monotonicity constraint (S8.49), the true Q^2 used to generate the data is not the one that has the larger likelihood, according to Figure 7(b). This illustrates the non-identifiability of Q_2.

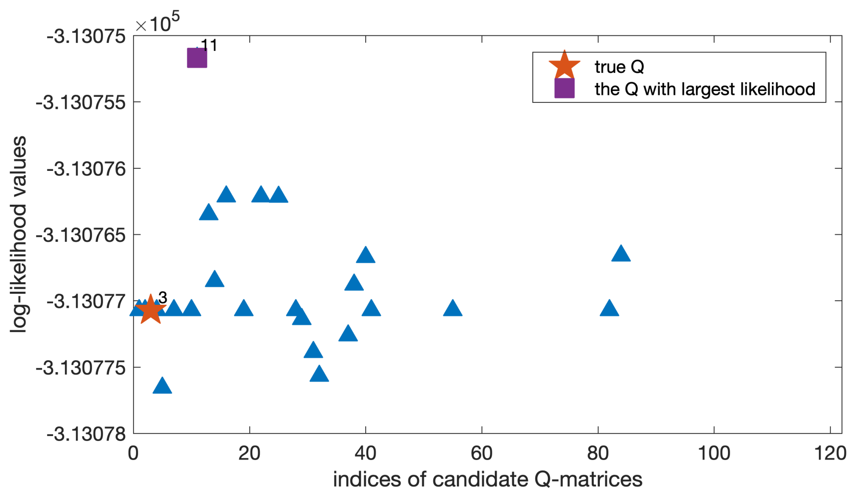
Figure 7: GDINA: exhaustive search in 5×2 Q-matrices with a true Q-matrix which leads to a not even locally generically identifiable model.



(a) GDINA: true Q not even locally identifiable: $Q^1 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$



(b) GDINA: true Q not even locally identifiable: $Q^2 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \end{pmatrix}^\top$



(c) GDINA: true Q not even locally identifiable: $Q^3 = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}^\top$

Study VII: Construction of many alternative sets of parameters when true Q-matrix violates the necessary condition for generic identifiability.

In this study, we verify Theorem 3, i.e., verify the necessity of Condition C that *each attribute is required by at least two items in the Q-matrix* for joint generic identifiability. We consider two cases with $(K, J) = (3, 20)$ and $(K, J) = (5, 20)$.

First, for $(K, J) = (3, 20)$, consider the following Q-matrix Q_3 and an alternative $\bar{Q}_3$.

$$Q_3 = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}_{20\times 3} \qquad \bar{Q}_3 = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}_{20\times 3} \tag{S8.50}$$

We first construct true parameters $(\boldsymbol{\Theta}, \boldsymbol{p})$ under Q_3. For each attribute pattern $\boldsymbol{\alpha}$, we set its population proportion $p_{\boldsymbol{\alpha}}$ to be $1/2^K$. For each item, set the baseline probability, the positive response probability of the all-zero attribute profile $\boldsymbol{\alpha} = \mathbf{0}^\top$, to be 0.2 and the positive response probability of $\boldsymbol{\alpha} = \mathbf{1}^\top$ to be 0.8. And we take all the main effects and interaction effects parameters to be equal.

For the defined true parameters $(\mathbf{\Theta}, \boldsymbol{p})$ under Q_3, we next construct 70 alternative sets of parameters $(\bar{\mathbf{\Theta}}^\ell, \bar{\boldsymbol{p}}^\ell)$ for $\ell = 1, 2, \ldots, 70$, all under the alternative Q-matrix $\bar{Q}_3$, that are non-distinguishable from the true parameters. Following the proof of Theorem 3, we first set $\bar{\theta}_{j,\boldsymbol{\alpha}} = \theta_{j,\boldsymbol{\alpha}}$ for any $j > 2$ and any $\boldsymbol{\alpha}$. Then we *randomly generate* the values of the $\bar{\mathbf{\Theta}}_{1:2,\,1:4}$ (the first four elements of the first two rows of $\bar{\mathbf{\Theta}}$) from the neighborhood of their true values, and enforce the monotonicity constraint (2.1). Specifically, for each alternative set (the ℓ-th set) of parameters, there is

$$\bar{\mathbf{\Theta}}^\ell_{i,j} = \mathbf{\Theta}_{i,j} + \mathcal{U}(-0.1, 0.1), \quad i = 1, 2;\ j = 1, 2, 3, 4;\ \ell = 1, 2, \ldots, 70.$$

where $\mathcal{U}(-0.1, 0.1)$ denotes a uniformly distributed random variable between -0.1 and 0.1. Next we just use Equation (S4.41) to get the remaining item parameters $\bar{\mathbf{\Theta}}^\ell_{1:2,\,5:8}$ and $\bar{\boldsymbol{p}}^\ell$.

Figure 8 presents the constructed 70 other parameters sets $(\bar{\mathbf{\Theta}}^\ell, \bar{\boldsymbol{p}}^\ell)$ under the alternative $\bar{Q}_3$, by plotting the values of difference between the alternative parameters and the true parameters. In particular, In Figure 8(a), the black solid line with dots is the reference line at zero, and each of the 70 colored dotted line with "+"'s represents one particular set of alternative parameters. For each colored line corresponding to the ℓth set of parameters, the following 16-dimensional vector of parameter difference is plotted,

$$(\bar{\theta}^\ell_{1,000} - \theta_{1,000},\ \bar{\theta}^\ell_{1,001} - \theta_{1,010},\ \bar{\theta}^\ell_{1,010} - \theta_{1,010},\ \bar{\theta}^\ell_{1,011} - \theta_{1,011},$$

$$\bar{\theta}^{\ell}_{1,100} - \theta_{1,100},\ \bar{\theta}^{\ell}_{1,101} - \theta_{1,110},\ \bar{\theta}^{\ell}_{1,110} - \theta_{1,110},\ \bar{\theta}^{\ell}_{1,111} - \theta_{1,111},$$

$$\bar{\theta}^{\ell}_{2,000} - \theta_{2,000},\ \bar{\theta}^{\ell}_{2,001} - \theta_{2,010},\ \bar{\theta}^{\ell}_{2,010} - \theta_{2,010},\ \bar{\theta}^{\ell}_{2,011} - \theta_{2,011},$$

$$\bar{\theta}^{\ell}_{2,100} - \theta_{2,100},\ \bar{\theta}^{\ell}_{2,101} - \theta_{2,110},\ \bar{\theta}^{\ell}_{2,110} - \theta_{2,110},\ \bar{\theta}^{\ell}_{2,111} - \theta_{2,111}).$$

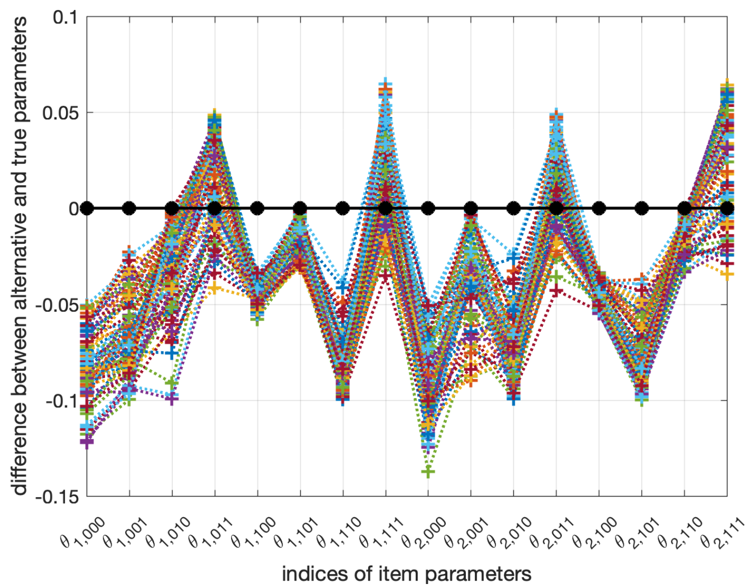
Similarly, in Figure 8(b), for each colored line corresponding to the ℓth set of parameters, the following 8-dimensional vector of parameter difference is plotted, $(\bar{p}^{\ell}_{000} - p_{000},\ \bar{p}^{\ell}_{001} - p_{010},\ \bar{p}^{\ell}_{010} - p_{010},\ \bar{p}^{\ell}_{011} - p_{011},\ \bar{p}^{\ell}_{100} - p_{100},\ \bar{p}^{\ell}_{101} - p_{110},\ \bar{p}^{\ell}_{110} - p_{110},\ \bar{p}^{\ell}_{111} - p_{111})$. In summary, a total number of 70 colored lines corresponding to 70 alternative sets of parameters are plotted in Figure 8.

The $(\boldsymbol{\Theta}, \boldsymbol{p})$ and all the $(\bar{\boldsymbol{\Theta}}^{\ell}, \bar{\boldsymbol{p}}^{\ell})$, $\ell = 1, \ldots, 70$ give the identical distribution of $\boldsymbol{R}$. Specifically, from the computation in Matlab, we have

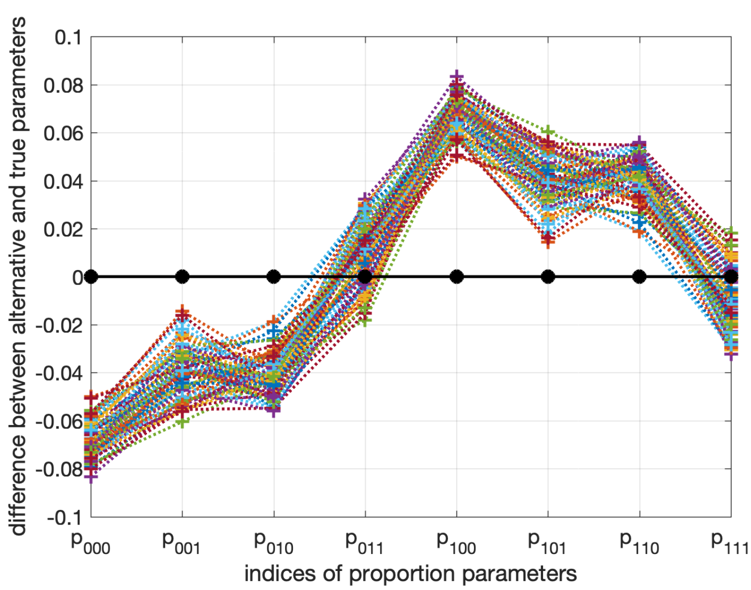
$$\max_{1 \le \ell \le 70} \max_{\boldsymbol{r} \in \{0,1\}^{20}} \left| \mathbb{P}(\boldsymbol{R} = \boldsymbol{r} \mid Q_3, \boldsymbol{\Theta}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R} = \boldsymbol{r} \mid \bar{Q}_3, \boldsymbol{\Theta}^{\ell}, \boldsymbol{p}^{\ell}) \right| = 1.30 \times 10^{-18},$$

which is smaller than the Matlab machine error 2.22×10^{-16}. This verifies that despite the underlying parameters are different from the truth, they all lead to the identical distribution of responses. So $(Q_3, \boldsymbol{\Theta}, \boldsymbol{p})$ are not identifiable. We emphasize that under this Q_3, for any true parameters, one can similarly construct arbitrarily many such alternative parameter sets under $\bar{Q}_3$.

Figure 8: GDINA: true Q is Q_3 with $(K, J) = (3, 20)$; each of the 70 colored line corresponds to one set of alternative parameters under $\bar{Q}_3$; all sets non-distinguishable.



(a) $K = 3$ and $J = 20$, 70 alternative sets of parameters



(b) $K = 3$ and $J = 20$, 70 sets of parameters

For $(K, J) = (5, 20)$, consider the following Q_4 and an alternative $\bar{Q}_4$,

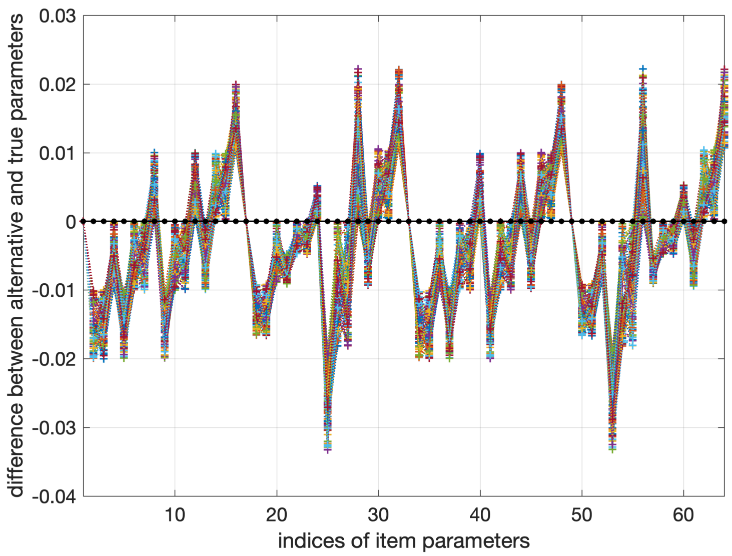
$$Q_4 = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}_{20\times 5} \quad \bar{Q}_4 = \begin{pmatrix} \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix}_{20\times 5}. \tag{S8.51}$$

We set the true parameters under Q_4 similarly as those under Q_3, and also use (S4.41) in the proof of Theorem 3 to randomly construct 70 sets of parameters under the $\bar{Q}_4$. Figure 9 (a) and (b) plot the values of difference between alternative and true item parameters (of the first two items), and that between alternative and true proportion parameters, respectively. Despite the differences in parameter values, our computation in Matlab shows the maximum difference between marginal response probabilities is

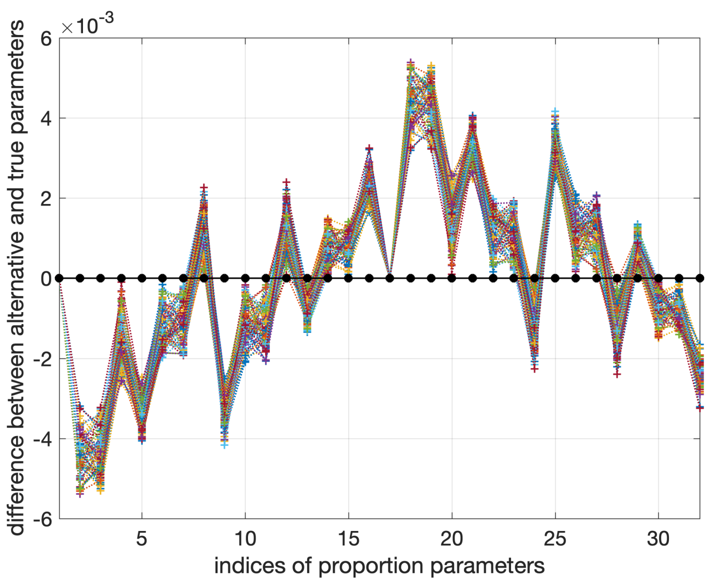
$$\max_{1\leq\ell\leq 70} \max_{\boldsymbol{r}\in\{0,1\}^{20}} \left|\mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid Q_4, \boldsymbol{\Theta}, \boldsymbol{p}) - \mathbb{P}(\boldsymbol{R}=\boldsymbol{r} \mid \bar{Q}_4, \boldsymbol{\Theta}^\ell, \boldsymbol{p}^\ell)\right| = 5.42\times 10^{-19},$$

also smaller than the Matlab machine error 2.22×10^{-16}. This illustrates the non-identifiability of Q_4.

Figure 9: GDINA: true Q is Q_4 with $(K, J) = (5, 20)$; each of the 70 colored line corresponds to one set of alternative parameters under $\bar{Q}_4$; all sets non-distinguishable.



(a) $K = 5$ and $J = 20$, 70 alternative sets of parameters



(b) $K = 5$ and $J = 20$, 70 alternative sets of parameters

References

Gu, Y. and Xu, G. (2018a). Partial identifiability of restricted latent class models. *arXiv preprint arXiv:1803.04353*.

Gu, Y. and Xu, G. (2018b). The sufficient and necessary condition for the identifiability and estimability of the DINA model. *Psychometrika*. to appear.

Hall, M. (1967). *Combinatorial theory*, volume 71. John Wiley & Sons.

Henson, R. A., Templin, J. L., and Willse, J. T. (2009). Defining a family of cognitive diagnosis models using log-linear models with latent variables. *Psychometrika*, 74:191–210.

Kruskal, J. B. (1977). Three-way arrays: rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics. *Linear algebra and its applications*, 18(2):95–138.

Rhodes, J. A. (2010). A concise proof of Kruskal's theorem on tensor decomposition. *Linear Algebra and its Applications*, 432(7):1818–1824.

Shen, X., Pan, W., and Zhu, Y. (2012). Likelihood-based selection and sharp parameter estimation. *Journal of the American Statistical Association*, 107(497):223–232.

Xu, G. (2017). Identifiability of restricted latent class models with binary responses. *The Annals of Statistics*, 45:675–707.

Xu, G. and Shang, Z. (2018). Identifying latent structures in restricted latent class models. *Journal of the American Statistical Association*, 113(523):1284–1295.