Supplement to "Generic Identifiability of the DINA Model and Blessing of Latent Dependence"

Yuqi Gu
Department of Statistics, Columbia University

Before presenting the proofs of the identifiability results, we introduce a useful technical tool, the T-matrix of marginal response probabilities. This technical tool was proposed by Xu and Zhang (2016) and also used in Gu and Xu (2019) to study the identifiability of the DINA model. First, consider a general notation $\mathbf{\Theta} = (\theta_{j,\boldsymbol{\alpha}})_{J\times 2^K}$ collecting all of the item parameters under the DINA model. The $J \times 2^K$ matrix $\mathbf{\Theta}$ has rows indexed by the J items and rows by all of the $|\{0,1\}^K| = 2^K$ configurations of the binary latent attribute pattern, where the $(j, \boldsymbol{\alpha})$th entry $\theta_{j,\boldsymbol{\alpha}} = \mathbb{P}(R_j = 1 \mid \mathbf{A} = \boldsymbol{\alpha})$ denotes the probability of a positive response to the jth item given the latent attribute pattern $\boldsymbol{\alpha}$. Then under the conjunctive assumption of DINA, we can write $\theta_{j,\boldsymbol{\alpha}}$ as

$$\theta_{j,\boldsymbol{\alpha}} = \begin{cases} 1 - s_j, & \text{if } \xi_{j,\boldsymbol{\alpha}} = \prod_{k=1}^K \alpha_k^{q_{j,k}} = 1; \\ g_j, & \text{otherwise.} \end{cases}$$

Note that given a $\mathbf{Q}$-matrix, there is a one-to-one mapping between the matrix $\mathbf{\Theta}$ and the item parameters $(\boldsymbol{s}, \boldsymbol{g})$. We next define a $2^J \times 2^K$ matrix $T(\mathbf{\Theta})$ based on $\mathbf{\Theta}$. The rows of $T(\mathbf{\Theta})$ 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 $T(\mathbf{\Theta})$, denoted by $T_{\boldsymbol{r},\boldsymbol{\alpha}}(\mathbf{\Theta})$, represents the marginal probability that subjects with latent pattern $\boldsymbol{\alpha}$ provide

Email: yuqi.gu@columbia.edu. Address: Department of Statistics, Columbia University, Room 928 SSW, 1255 Amsterdam Avenue, New York, NY 10027, USA.

positive responses to the set of items $\{j : r_j = 1\}$, namely

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

We denote the $\boldsymbol{\alpha}$th column vector and the $\boldsymbol{r}$th row vector of the T-matrix by $T_{:,\boldsymbol{\alpha}}(\boldsymbol{\Theta})$ and $T_{\boldsymbol{r},:}(\boldsymbol{\Theta})$, respectively. The $\boldsymbol{r}$th element of the 2^J-dimensional vector $T(\boldsymbol{\Theta})\boldsymbol{p}$ is

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

Based on the T-matrix, there is an equivalent definition of identifiability of $(\boldsymbol{\Theta}, \boldsymbol{p})$ (equivalently, identifiability of $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$). Further, the T-matrix has a nice property that will facilitate proving the identifiability results. We summarize them in the following lemma, whose proof can be found in Xu (2017).

Lemma 1. *Consider the DINA model defined in* (1)*.*

(a) The parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ are identifiable if and only if there does not exist $(\bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}}) \neq (\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ such that

$$T(\boldsymbol{\Theta})\boldsymbol{p} = T(\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}}.$$

(b) For any vector $\boldsymbol{\theta}^* = (\theta_1^*, \dots, \theta_J^*)^\top \in \mathbb{R}^J$, there exists an $2^J \times 2^J$ invertible matrix $D(\boldsymbol{\theta}^*)$ which depends only on $\boldsymbol{\theta}^*$ such that

$$T(\boldsymbol{\Theta} - \boldsymbol{\theta}^* \cdot \mathbf{1}_{2^K}^\top) = D(\boldsymbol{\theta}^*) \cdot T(\boldsymbol{\Theta}).$$

Lemma 1 (a) and (b) imply that for any vector $\boldsymbol{\theta}^* = (\theta_1^*, \dots, \theta_J^*)^\top$, there holds

$$T(\boldsymbol{\Theta} - \boldsymbol{\theta}^* \cdot \mathbf{1}_{2^K}^\top)\boldsymbol{p} = D(\boldsymbol{\theta}^*)T(\boldsymbol{\Theta})\boldsymbol{p} = D(\boldsymbol{\theta}^*)T(\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}} = T(\bar{\boldsymbol{\Theta}} - \boldsymbol{\theta}^* \cdot \mathbf{1}_{2^K}^\top)\bar{\boldsymbol{p}} \qquad \text{(S.1)}$$

The above equality will be frequently used throughout the proof of our identifiability results. In the following proofs, we sometimes will denote $\boldsymbol{c} := \mathbf{1}_J - \boldsymbol{s} = (1-s_1, \ldots, 1-s_J)^\top$ for notational convenience. Using this notation, the DINA model parameters can be equivalently expressed as $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$.

S.1 Proof of Proposition 1

We rewrite Eq. (4) in the main text below,

$$\begin{aligned}
\mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p}) &= \sum_{\boldsymbol{\alpha}\in\{0,1\}^K} p_{\boldsymbol{\alpha}} \cdot \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \mathbf{A} = \boldsymbol{\alpha}, \boldsymbol{s}, \boldsymbol{g}) \\
&= \sum_{\boldsymbol{\alpha}\in\{0,1\}^K} p_{\boldsymbol{\alpha}} \cdot \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \boldsymbol{\xi}_{:,\mathbf{A}} = \boldsymbol{\xi}_{:,\boldsymbol{\alpha}}, \boldsymbol{s}, \boldsymbol{g}) \\
&= \sum_{\boldsymbol{\alpha}\in\mathcal{R}} \Big(\sum_{\substack{\boldsymbol{\beta}\in\{0,1\}^K: \\ \boldsymbol{\xi}_{:,\boldsymbol{\beta}} = \boldsymbol{\xi}_{:,\boldsymbol{\alpha}}}} p_{\boldsymbol{\alpha}} \Big) \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \boldsymbol{\xi}_{:,\mathbf{A}} = \boldsymbol{\xi}_{:,\boldsymbol{\alpha}}, \boldsymbol{s}, \boldsymbol{g}),
\end{aligned}$$

where the notation $\mathcal{R} \subseteq \{0,1\}^K$ denotes a collection of representative latent attribute patterns, such that $\{\boldsymbol{\xi}_{:,\boldsymbol{\alpha}} : \boldsymbol{\alpha} \in \mathcal{R}\}$ contains mutually distinct ideal response vectors and also covers all the possible ideal response vectors under the $\mathbf{Q}$-matrix. Because of (4), for any $\boldsymbol{\alpha} \in \mathcal{R}$, those patterns $\boldsymbol{\beta} \in \{0,1\}^K$ with $\boldsymbol{\xi}_{:,\boldsymbol{\beta}} = \boldsymbol{\xi}_{:,\boldsymbol{\alpha}}$ can be considered to be equivalent to $\boldsymbol{\alpha}$ under the DINA model with the considered $\mathbf{Q}$-matrix. For $\boldsymbol{\alpha} \in \mathcal{R}$, define the equivalence class of latent attribute patterns by

$$[\boldsymbol{\alpha}] := \{\boldsymbol{\beta} \in \{0,1\}^K : \boldsymbol{\xi}_{:,\boldsymbol{\beta}} = \boldsymbol{\xi}_{:,\boldsymbol{\alpha}}\}.$$

We next show that if for some $\boldsymbol{\alpha} \in \{0,1\}^K$, the set $[\boldsymbol{\alpha}]$ contains multiple elements, say $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}' \in [\boldsymbol{\alpha}]$ with $\boldsymbol{\alpha} \neq \boldsymbol{\alpha}'$, then their corresponding proportion parameters $p_{\boldsymbol{\alpha}}$ and $p_{\boldsymbol{\alpha}'}$ will always be unidentifiable, no matter what values $p_{\boldsymbol{\alpha}}$ and $p_{\boldsymbol{\alpha}'}$ take. Specifically, if two sets of parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ and $(\bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$ satisfy that $\mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p}) = \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$ for all

$\boldsymbol{r} \in \{0,1\}^J$ under a same $\mathbf{Q}$-matrix, then (4) gives

$$\sum_{\boldsymbol{\alpha}\in\mathcal{R}}\Big(\sum_{\substack{\boldsymbol{\beta}\in\{0,1\}^K:\\ \boldsymbol{\xi}_{:,\boldsymbol{\beta}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}}}} p_{\boldsymbol{\alpha}}\Big)\mathbb{P}(\mathbf{R}=\boldsymbol{r}\mid \boldsymbol{\xi}_{:,\mathbf{A}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}},\boldsymbol{s},\boldsymbol{g}) = \sum_{\boldsymbol{\alpha}\in\mathcal{R}}\Big(\sum_{\substack{\boldsymbol{\beta}\in\{0,1\}^K:\\ \boldsymbol{\xi}_{:,\boldsymbol{\beta}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}}}} \bar{p}_{\boldsymbol{\alpha}}\Big)\mathbb{P}(\mathbf{R}=\boldsymbol{r}\mid \boldsymbol{\xi}_{:,\mathbf{A}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}},\bar{\boldsymbol{s}},\bar{\boldsymbol{g}});$$

and even if $(\boldsymbol{s},\boldsymbol{g}) = (\bar{\boldsymbol{s}},\bar{\boldsymbol{g}})$, the identifiability equations $\mathbb{P}(\mathbf{R}\mid \boldsymbol{s},\boldsymbol{g},\boldsymbol{p}) = \mathbb{P}(\mathbf{R}\mid \bar{\boldsymbol{s}},\bar{\boldsymbol{g}},\bar{\boldsymbol{p}})$ only give the following,

$$\sum_{\boldsymbol{\alpha}\in\mathcal{R}}\Big(\sum_{\substack{\boldsymbol{\beta}\in\{0,1\}^K:\\ \boldsymbol{\xi}_{:,\boldsymbol{\beta}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}}}} p_{\boldsymbol{\alpha}} - \sum_{\substack{\boldsymbol{\beta}\in\{0,1\}^K:\\ \boldsymbol{\xi}_{:,\boldsymbol{\beta}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}}}} \bar{p}_{\boldsymbol{\alpha}}\Big)\mathbb{P}(\mathbf{R}=\boldsymbol{r}\mid \boldsymbol{\xi}_{:,\mathbf{A}}=\boldsymbol{\xi}_{:,\boldsymbol{\alpha}},\boldsymbol{s},\boldsymbol{g}) = 0, \quad \forall \boldsymbol{r}\in\{0,1\}^J.$$

From the above equations, one can not identify individual parameters $p_{\boldsymbol{\beta}}$ for those $\boldsymbol{\beta}$ belonging to a same equivalence class $[\boldsymbol{\alpha}]$. Next we will show that if $\mathbf{Q}$ violates the Completeness Condition (C), then some equivalence class $[\boldsymbol{\alpha}]$ will contain multiple elements, leading to the aforementioned non-identifiability consequence.

According to Gu and Xu (2020), the set of representative patterns $\mathcal{R}$ in (4) can be obtained using the row vectors of the $\mathbf{Q}$-matrix as follows,

$$\mathcal{R} = \left\{\bigvee_{j\in S}\boldsymbol{q}_j:\ S\subseteq\{1,\dots,J\}\text{ is an arbitrary subset of item indices}\right\}, \tag{S.2}$$

where $\bigvee_{j\in S}\boldsymbol{q}_j =: \boldsymbol{\alpha}$ denotes the elementwise maximum of the set of vectors $\{\boldsymbol{q}_j:\ j\in S\}$ and the kth entry of the resultant vector $\boldsymbol{\alpha}$ is $\alpha_k = \max_{j\in S}\{q_{j,k}\}$. So $\bigvee_{j\in S}\boldsymbol{q}_j$ is also a K-dimensional binary vector and hence $\mathcal{R}\succeq\{0,1\}^K$. In fact, $\mathcal{R}=\{0,1\}^K$ if and only if $\mathbf{Q}$ contains a submatrix $\mathbf{I}_K$ after some row permutation. To see this, consider if the row vectors of $\mathbf{Q}$ do not include a certain standard basis vector $\boldsymbol{e}_k$ (which has a "1" in the kth entry and "0" otherwise), then $\boldsymbol{e}_k$ does not belong to $\mathcal{R}$ defined in (S.2) because $\boldsymbol{e}_k$ cannot be written in the form of $\bigvee_{j\in S}\boldsymbol{q}_j$ for any subset $S\subseteq[J]$. Therefore, if $\mathbf{Q}$ violates the Completeness Condition (C), then $\mathcal{R}$ is a proper subset of $\{0,1\}^K$, which implies certain attribute patterns become equivalent under such a $\mathbf{Q}$-matrix. In summary, if a $\mathbf{Q}$-matrix does not contain a submatrix $\mathbf{I}_K$, certain proportion parameters $p_{\boldsymbol{\alpha}}$'s will always be unidentifiable regardless of the values of these $p_{\boldsymbol{\alpha}}$'s. This implies the failure of generic identifiability of the DINA model parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ according to Definition 2 and proves Proposition 1. $\square$

S.2 Proof of Proposition 2

The construction for non-identifiable parameters in this setting is the same as that in the proof of Theorem 1 in Xu and Zhang (2016). We next elaborate on this construction to make clear the failure of generic identifiability. Since $\mathbf{Q}$ satisfies Condition (C), we can write the form of $\mathbf{Q}$ as follows without loss of generality,

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

where the first attribute A_1 is required by only one item, the first item. Next construct two different sets of DINA model parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ and $(\bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$ which lead to the same distribution of $\mathbf{R}$. In particular, if setting $s_j = \bar{s}_j$ and $g_j = \bar{g}_j$ for all $j \geq 2$, then the identifiability equations $\mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p}) = \mathbb{P}(\mathbf{R} = \boldsymbol{r} \mid \bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$ for all $\boldsymbol{r} \in \{0,1\}^J$ will exactly reduce to the following set of equations,

$$\forall \boldsymbol{\alpha}^* \in \{0,1\}^{K-1}, \quad \begin{cases} p_{(0,\boldsymbol{\alpha}^*)} + p_{(1,\boldsymbol{\alpha}^*)} = \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,\boldsymbol{\alpha}^*)}; \\ g_1 p_{(0,\boldsymbol{\alpha}^*)} + (1-s_1) p_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_1 \bar{p}_{(0,\boldsymbol{\alpha}^*)} + (1-\bar{s}_1) \bar{p}_{(1,\boldsymbol{\alpha}^*)}. \end{cases}$$

The above system of equations involve $|\{\bar{g}_1, \bar{s}_1\} \cup \{\bar{p}_{\boldsymbol{\alpha}};\ \boldsymbol{\alpha} \in \{0,1\}^K\}| = 2^K + 2$ free unknown variables regarding $(\bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$, while there are only 2^K equations, so there exist infinitely many different solutions to $(\bar{\boldsymbol{s}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$. In particular, we can let $\bar{g}_1 = g_1$ and arbitrarily set $\bar{s}_1$ in a small neighborhood of s_1 with $\bar{s}_1 \neq s_1$. Then correspondingly solve for the proportion parameters $\bar{\boldsymbol{p}}$ as

$$\forall \boldsymbol{\alpha}^* \in \{0,1\}^{K-1}, \quad \bar{p}_{(1,\boldsymbol{\alpha}^*)} = \frac{1-s_1}{1-\bar{s}_1} p_{(1,\boldsymbol{\alpha}^*)}, \quad \bar{p}_{(0,\boldsymbol{\alpha}^*)} = p_{(0,\boldsymbol{\alpha}^*)} + \left(1 - \frac{1-s_1}{1-\bar{s}_1}\right) p_{(1,\boldsymbol{\alpha}^*)}.$$

Since $\bar{s}_1$ can vary arbitrarily in the neighborhood of s_1 without changing the distribution of $\mathbf{R}$, we have shown that the parameter s_1 is always unidentifiable in the parameter space. The parameter g_1 can be similarly shown to be always unidentifiable. The fact that item parameters (s_1, g_1) are always unidentifiable whatever their values are indicates the failure of generic identifiability. This proves the conclusion of Proposition 2. □

S.3 Proof of Theorem 1 and Theorem 4

Proof of Theorem 1. Below we rewrite the form of the $\mathbf{Q}$-matrix stated in the theorem,

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

By Lemma 1, if parameters $(\mathbf{\Theta}, \boldsymbol{p})$ and $(\bar{\mathbf{\Theta}}, \bar{\boldsymbol{p}})$ give rise to the same distribution of the observed responses, then the following equality holds,

$$T_{\boldsymbol{r},:}(\mathbf{\Theta})\boldsymbol{p} = T_{\boldsymbol{r},:}(\bar{\mathbf{\Theta}})\bar{\boldsymbol{p}} \quad \text{for all} \quad \boldsymbol{r} \in \{0,1\}^J, \tag{S.3}$$

Note that the last $J-2$ rows of $\mathbf{Q}$ has the first column being an all-zero column, and has the other $K-1$ columns forming a sub-matrix $\mathbf{Q}^\star$ which satisfies the C-R-D conditions. Since the C-R-D conditions are sufficient for identifiability of DINA model parameters by Gu and Xu (2019), the last $J-2$ rows of the $\mathbf{Q}$-matrix implies a nice identifiability result for a subset of the model parameters $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$. We next elaborate on this observation.

For notational convenience, denote by $\mathbb{P}(\cdot)$ the probability under the true parameters $(\boldsymbol{c}, \boldsymbol{g}, \boldsymbol{p})$, and denote by $\overline{\mathbb{P}}(\cdot)$ the probability under the alternative parameters $(\bar{\boldsymbol{c}}, \bar{\boldsymbol{g}}, \bar{\boldsymbol{p}})$. For a $\boldsymbol{\alpha}^* \in \{0,1\}^{K-1}$, let $(0, \boldsymbol{\alpha}^*)$, $(1, \boldsymbol{\alpha}^*) \in \{0,1\}^K$ denote two K-dimensional binary vectors. Since $\mathbf{Q}_{1,3:J}$ is an all-zero vector, it is always true that $\theta_{j,(1,\boldsymbol{\alpha}^*)} = \theta_{j,(0,\boldsymbol{\alpha}^*)}$ for $j \geq 3$ and $\boldsymbol{\alpha}^* \in \{0,1\}^{K-1}$. Therefore, for any response pattern $\boldsymbol{r} = (r_1, r_2, \boldsymbol{r}^*) \in \{0,1\}^J$, Eq. (S.3) for

$\boldsymbol{r}$ implies the following,

$$\sum_{(z,\boldsymbol{\alpha}^*)\in\{0,1\}^K}\prod_{j>2:\,r_j=1}\theta_{j,\,(0,\boldsymbol{\alpha}^*)}\cdot\mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,A_1=z,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)$$

$$\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-1}}\prod_{j>2:\,r_j=1}\theta_{j,\,(0,\boldsymbol{\alpha}^*)}\cdot[\mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,A_1=1,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)$$
$$+\,\mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,A_1=0,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)]$$
$$=\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-1}}\prod_{j>2:\,r_j=1}\bar{\theta}_{j,\,(0,\boldsymbol{\alpha}^*)}\cdot[\overline{\mathbb{P}}(R_1\geq r_1,\,R_2\geq r_2,\,A_1=1,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)$$
$$+\,\overline{\mathbb{P}}(R_1\geq r_1,\,R_2\geq r_2,\,A_1=0,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)];$$

which can be further simplified to be

$$\begin{aligned}&\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-1}}\prod_{j>2:\,r_j=1}\theta_{j,\,(0,\boldsymbol{\alpha}^*)}\cdot\mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)\\ =&\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-1}}\prod_{j>2:\,r_j=1}\bar{\theta}_{j,\,(0,\boldsymbol{\alpha}^*)}\cdot\overline{\mathbb{P}}(R_1\geq r_1,\,R_2\geq r_2,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*).\end{aligned}\tag{S.4}$$

Note that fixing an arbitrary (r_1, r_2) and varying $\boldsymbol{r}^*\in\{0,1\}^{J-1}$, the above systems of equations (S.4) can be viewed as surrogate identifiability equations $T(\boldsymbol{\Theta}^*)\boldsymbol{p}^*=T(\bar{\boldsymbol{\Theta}}^*)\bar{\boldsymbol{p}}^*$ for the last $J-2$ items in the test, where those $\theta_{j,(0,\boldsymbol{\alpha}^*)}=:\theta^*_{j,\boldsymbol{\alpha}^*}$ serve as surrogate item parameters $\boldsymbol{\Theta}^*=\{\theta^*_{j,\boldsymbol{\alpha}^*}:\ j=3,\ldots,J;\ \boldsymbol{\alpha}^*\in\{0,1\}^{K-1}\}$; and those $\mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)=:p^*_{\boldsymbol{\alpha}^*}$ serve as surrogate proportion parameters $\boldsymbol{p}^*=\{p^*_{\boldsymbol{\alpha}^*}:\ \boldsymbol{\alpha}^*\in\{0,1\}^{K-1}\}$. An important observation is that the parameters $(\boldsymbol{\Theta}^*,\boldsymbol{p}^*)$ can be viewed as associated with the matrix $\mathbf{Q}^\star$ under a DINA model with $J-2$ items and $K-1$ latent attributes. Now that $\mathbf{Q}^\star$ satisfies the C-R-D conditions (which are sufficient for identifiability), we obtain the following "identifiability conclusions" for the parameters $(\boldsymbol{\Theta}^*,\boldsymbol{p}^*)$,

$$\begin{cases}\theta_{j,(0,\boldsymbol{\alpha}^*)}=\bar{\theta}_{j,(0,\boldsymbol{\alpha}^*)};\\ \mathbb{P}(R_1\geq r_1,\,R_2\geq r_2,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*)=\overline{\mathbb{P}}(R_1\geq r_1,\,R_2\geq r_2,\,\mathbf{A}_{2:K}=\boldsymbol{\alpha}^*);\end{cases}\tag{S.5}$$

which hold for any $j \in \{3, \ldots, J\}$ and $\boldsymbol{\alpha}^* \in \{0,1\}^{K-1}$. Recall that for any item $j \geq 3$, the parameter $\theta_{j,(0,\boldsymbol{\alpha}^*)}$ ranges over both item parameters c_j and g_j) when $\boldsymbol{\alpha}^*$ ranges in $\{0,1\}^{K-1}$, so the first part of (S.5) implies

$$c_j = \bar{c}_j, \quad g_j = \bar{g}_j, \quad \forall j \in \{3, \ldots, J\}. \tag{S.6}$$

Recall the form of $\mathbf{Q}$ and the vector $\boldsymbol{u}$ stated in the theorem, for any $\boldsymbol{\alpha}^* \in \{0,1\}^{K-1}$ and $\boldsymbol{\alpha}^* \succeq \boldsymbol{u}$ (i.e. vector $\boldsymbol{\alpha}$ is elementwisely greater than or equal to vector $\boldsymbol{u}$), the second part of (S.5) implies the following must hold,

$$(r_1, r_2) = \begin{cases} (0,0) \implies p_{(0,\boldsymbol{\alpha}^*)} + p_{(1,\boldsymbol{\alpha}^*)} = \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,\boldsymbol{\alpha}^*)}; \\ (1,0) \implies g_1 \cdot p_{(0,\boldsymbol{\alpha}^*)} + c_1 \cdot p_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_1 \cdot \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_1 \cdot \bar{p}_{(1,\boldsymbol{\alpha}^*)}; \\ (0,1) \implies g_2 \cdot p_{(0,\boldsymbol{\alpha}^*)} + c_2 \cdot p_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_2 \cdot \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_2 \cdot \bar{p}_{(1,\boldsymbol{\alpha}^*)}; \\ (1,1) \implies g_1 g_2 \cdot p_{(0,\boldsymbol{\alpha}^*)} + c_1 c_2 \cdot p_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_1 \bar{g}_2 \cdot \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_1 \bar{c}_2 \cdot \bar{p}_{(1,\boldsymbol{\alpha}^*)}. \end{cases} \tag{S.7}$$

First, we transform the system of equations (S.7) to obtain

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

Note that the right hand sides of both the above equations are nonzero. So we can take the ratio of the two equations to obtain

$$f_1(\boldsymbol{\alpha}^*) := \frac{(g_1 - c_1) \cdot (g_2 - \bar{c}_2)}{(g_2 - \bar{c}_2) + (c_2 - \bar{c}_2) \cdot p_{(1,\boldsymbol{\alpha}^*)}/p_{(0,\boldsymbol{\alpha}^*)}} = \bar{g}_1 - c_1.$$

So for two arbitrary patterns $\boldsymbol{\alpha}_1^*$, $\boldsymbol{\alpha}_2^* \in \{0,1\}^{K-1}$ with $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}$, our above deduction gives $f_1(\boldsymbol{\alpha}_1^*) = f_1(\boldsymbol{\alpha}_2^*) = \bar{g}_1 - c_1$. This equality of $f_1(\boldsymbol{\alpha}_1^*)$ and $f_1(\boldsymbol{\alpha}_2^*)$ implies

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

$$\Longrightarrow \quad (c_2-\bar{c}_2)\cdot\left(\frac{p_{(1,\boldsymbol{\alpha}_1^*)}}{p_{(0,\boldsymbol{\alpha}_1^*)}}-\frac{p_{(1,\boldsymbol{\alpha}_2^*)}}{p_{(0,\boldsymbol{\alpha}_2^*)}}\right)=0. \tag{S.8}$$

The above equation indicates that as long as there exist one pair of patterns $\boldsymbol{\alpha}_1^*$, $\boldsymbol{\alpha}_2^* \in \{0,1\}^{K-1}$ with $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}$ and $\boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^*$ such that

$$p_{(1,\boldsymbol{\alpha}_1^*)}p_{(0,\boldsymbol{\alpha}_2^*)}-p_{(0,\boldsymbol{\alpha}_1^*)}p_{(1,\boldsymbol{\alpha}_2^*)}\neq 0, \tag{S.9}$$

then $p_{(1,\boldsymbol{\alpha}_1^*)}/p_{(0,\boldsymbol{\alpha}_1^*)} \neq p_{(1,\boldsymbol{\alpha}_2^*)}/p_{(0,\boldsymbol{\alpha}_2^*)}$ and we must have $c_2=\bar{c}_2$ from (S.8). Under the assumption stated in Theorem 1 that $\boldsymbol{u}\neq \mathbf{1}_{K-1}$, there indeed exist such two distinct vectors $\boldsymbol{\alpha}_1^*$, $\boldsymbol{\alpha}_2^*$ satisfying $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}$. Therefore, $c_2=\bar{c}_2$ (i.e., c_2 is identifiable) as long as $\boldsymbol{p}\notin\mathcal{N}_{R,1}$, where the set $\mathcal{N}_{R,1}$ is defined in the statement of Theorem 4:

$$\mathcal{N}_{R,1}=\{\boldsymbol{p} \text{ satisfies } p_{(1,\boldsymbol{\alpha}_1^*)}p_{(0,\boldsymbol{\alpha}_2^*)}-p_{(0,\boldsymbol{\alpha}_1^*)}p_{(1,\boldsymbol{\alpha}_2^*)}=0 \text{ for any } \boldsymbol{\alpha}_1^*\neq\boldsymbol{\alpha}_2^* \text{ with } \boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*\succeq\boldsymbol{u}\}.$$

Next, we transform the system of equations (S.7) in another way to obtain

$$\begin{cases}(c_1-g_1)\cdot(c_2-\bar{g}_2)\cdot p_{(1,\boldsymbol{\alpha}^*)}=(\bar{c}_1-g_1)\cdot(\bar{c}_2-\bar{g}_2)\cdot\bar{p}_{(1,\boldsymbol{\alpha}^*)};\\[2ex](g_2-\bar{g}_2)\cdot p_{(0,\boldsymbol{\alpha}^*)}+(c_2-\bar{g}_2)\cdot p_{(1,\boldsymbol{\alpha}^*)}=(\bar{c}_2-\bar{g}_2)\cdot\bar{p}_{(1,\boldsymbol{\alpha}^*)}.\end{cases}$$

The ratio of the above two equations gives

$$f_2(\boldsymbol{\alpha}^*):=\frac{(c_1-g_1)\cdot(c_2-\bar{g}_2)}{(g_2-\bar{g}_2)\cdot p_{(0,\boldsymbol{\alpha}^*)}/p_{(1,\boldsymbol{\alpha}^*)}+(c_2-\bar{g}_2)}=\bar{c}_1-g_1.$$

Again we have $f_2(\boldsymbol{\alpha}_1^*)=f_2(\boldsymbol{\alpha}_2^*)$ for any $\boldsymbol{\alpha}_1^*,\boldsymbol{\alpha}_2^*\succeq\boldsymbol{u}$ with $\boldsymbol{\alpha}_1^*\neq\boldsymbol{\alpha}_2^*$. Such an equality implies

$$(g_2-\bar{g}_2)\cdot\frac{p_{(0,\boldsymbol{\alpha}_1^*)}}{p_{(1,\boldsymbol{\alpha}_1^*)}}=(g_2-\bar{g}_2)\cdot\frac{p_{(0,\boldsymbol{\alpha}_2^*)}}{p_{(1,\boldsymbol{\alpha}_2^*)}}, \quad \Longrightarrow \quad (g_2-\bar{g}_2)\cdot\left(\frac{p_{(0,\boldsymbol{\alpha}_1^*)}}{p_{(1,\boldsymbol{\alpha}_1^*)}}-\frac{p_{(0,\boldsymbol{\alpha}_2^*)}}{p_{(1,\boldsymbol{\alpha}_2^*)}}\right)=0.$$

Therefore, as long as $\boldsymbol{p}\notin\mathcal{N}_{R,1}$, we also have $g_2=\bar{g}_2$ and g_2 is identifiable.

Now note that the systems of equations (S.7) are symmetric about (c_1,g_1) and (c_2,g_2). Since we have already obtained $c_2=\bar{c}_2$ and $g_2=\bar{g}_2$ if $\boldsymbol{p}\notin\mathcal{N}_{R,1}$, we also have $c_1=\bar{c}_1$ and

$g_1 = \bar{g}_1$ if $\boldsymbol{p} \notin \mathcal{N}_{R,1}$ following the same argument. Thus far we have already established $\boldsymbol{c} = \bar{\boldsymbol{c}}$ and $\boldsymbol{g} = \bar{\boldsymbol{g}}$, i.e., have shown the identifiability of all the item parameters in $\boldsymbol{\Theta}$.

Since the item parameters $(\boldsymbol{c}, \boldsymbol{g})$ (equivalently, $\boldsymbol{\Theta}$) are already identified, and we have $T(\boldsymbol{\Theta})\boldsymbol{p} = T(\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}} = T(\boldsymbol{\Theta})\bar{\boldsymbol{p}}$. Since $\mathbf{Q}$ contains a submatrix $\mathbf{I}_K$, the matrix $T(\boldsymbol{\Theta})$ has full column rank from a statement in Xu and Zhang (2016), and hence we obtain $\boldsymbol{p} = \bar{\boldsymbol{p}}$. This means all the parameters $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p})$ are identifiable as long as $\boldsymbol{p}$ satisfies (S.9). More precisely, we have that the DINA model parameters are identifiable if $(\boldsymbol{s}, \boldsymbol{g}, \boldsymbol{p}) \in \mathcal{T} \setminus \mathcal{N}_{R,1}$ where the set $\mathcal{N}_{R,1}$ is defined by (8) in the main text in Theorem 4. We rewrite the definition of $\mathcal{N}_{R,1}$, The above set $\mathcal{N}_{R,1}$ has measure zero with respect to the Lebesgue measure defined on the parameter space $\mathcal{T}$. This is because $\mathcal{N}_{R,1}$ is characterized by the zero set of a polynomial equation about entries of $\boldsymbol{p}$, and by basic algebraic geometry, $\mathcal{N}_{R,1}$ necessarily has measure zero in the parameter space of $\boldsymbol{p}$. This completes the proof of Theorem 1.

Proof of Theorem 4. We next examine the statistical interpretation of the null set $\mathcal{N}_{R,1}$ defined in (8) where identifiability breaks down. Recall the definition of the population proportion parameter $p_{\boldsymbol{\alpha}} = \mathbb{P}(\mathbf{A} = \boldsymbol{\alpha})$, where $\mathbf{A} = (A_1, \dots, A_K)$ denotes a random attribute profile. For an arbitrary attribute pattern $\boldsymbol{\alpha} = (\alpha_1, \boldsymbol{\alpha}^*)$ where the subvector satisfies $\boldsymbol{\alpha}^* \in \{0,1\}^{K-1}$ and $\boldsymbol{\alpha}^* \succeq \boldsymbol{u}$, we have

$$
\begin{aligned}
&\mathbb{P}(A_1 = \alpha_1)\mathbb{P}(\mathbf{A}_{2:K} = \boldsymbol{\alpha}^*) \\
&= \Big(\sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(\alpha_1, \boldsymbol{\beta})} \Big) \big(p_{(\alpha_1, \boldsymbol{\alpha}^*)} + p_{(1-\alpha_1, \boldsymbol{\alpha}^*)} \big) \\
&= \sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(\alpha_1, \boldsymbol{\beta})} p_{(\alpha_1, \boldsymbol{\alpha}^*)} + \sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(\alpha_1, \boldsymbol{\beta})} p_{(1-\alpha_1, \boldsymbol{\alpha}^*)} \\
&= \sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(\alpha_1, \boldsymbol{\beta})} p_{(\alpha_1, \boldsymbol{\alpha}^*)} + \sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(1-\alpha_1, \boldsymbol{\beta})} p_{(\alpha_1, \boldsymbol{\alpha}^*)} \quad (\text{because } \boldsymbol{p} \in \mathcal{N}_{R,1}) \\
&= \Big(\sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(\alpha_1, \boldsymbol{\beta})} + \sum_{\boldsymbol{\beta} \in \{0,1\}^{K-1}} p_{(1-\alpha_1, \boldsymbol{\beta})} \Big) p_{(\alpha_1, \boldsymbol{\alpha}^*)} \\
&= p_{(\alpha_1, \boldsymbol{\alpha}^*)} = \mathbb{P}(\mathbf{A} = \boldsymbol{\alpha}).
\end{aligned}
$$

The third equality above follows from the fact that for $\boldsymbol{p} \in \mathcal{N}_{R,1}$, the $p_{(\alpha_1,\boldsymbol{\beta})}p_{(1-\alpha_1,\boldsymbol{\alpha}^*)} = p_{(1-\alpha_1,\boldsymbol{\beta})}p_{(\alpha_1,\boldsymbol{\alpha}^*)}$ holds for any $\alpha_1 \in \{0,1\}$ and $\boldsymbol{\alpha}^*, \boldsymbol{\beta} \in \{0,1\}^{K-1}$. Now we obtain that if $\boldsymbol{p} \in \mathcal{N}_{R,1}$, then $\mathbb{P}(\mathbf{A} = (\alpha_1, \boldsymbol{\alpha}^*)) = \mathbb{P}(A_1 = \alpha_1)\mathbb{P}(\mathbf{A}_{2:K} = \boldsymbol{\alpha}^*)$ for any $\alpha_1 \in \{0,1\}$ and $\boldsymbol{\alpha}^* \succeq \boldsymbol{u}$. This implies if $\boldsymbol{p} \in \mathcal{N}_{R,1}$, then latent attribute A_1 is conditionally independent of latent attributes $\mathbf{A}_{2:K}$ given $\mathbf{A}_{2:K} \succeq \boldsymbol{u}$.

On the other hand, if latent variables A_1 and $\mathbf{A}_{2:K}$ are conditionally independent given $\mathbf{A}_{2:K} \succeq \boldsymbol{u}$, then for any $\boldsymbol{\alpha}^* \succeq \boldsymbol{u}$ we have

$$\frac{p_{(1,\boldsymbol{\alpha}^*)}}{p_{(0,\boldsymbol{\alpha}^*)}} = \frac{\mathbb{P}(\mathbf{A} = (1, \boldsymbol{\alpha}^*))}{\mathbb{P}(\mathbf{A} = (0, \boldsymbol{\alpha}^*))} = \frac{\mathbb{P}(A_1 = 1)\mathbb{P}(\mathbf{A}_{2:K} = \boldsymbol{\alpha}^*)}{\mathbb{P}(A_1 = 0)\mathbb{P}(\mathbf{A}_{2:K} = \boldsymbol{\alpha}^*)} = \frac{\mathbb{P}(A_1 = 1)}{\mathbb{P}(A_1 = 0)} =: \rho.$$

This means for any $\boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^*$ with $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}$, the equality $p_{(1,\boldsymbol{\alpha}_1^*)}/p_{(0,\boldsymbol{\alpha}_1^*)} - p_{(1,\boldsymbol{\alpha}_2^*)}/p_{(0,\boldsymbol{\alpha}_2^*)} = \rho - \rho = 0$ must hold, which is equivalent to $p_{(1,\boldsymbol{\alpha}_1^*)}p_{(0,\boldsymbol{\alpha}_2^*)} - p_{(0,\boldsymbol{\alpha}_1^*)}p_{(1,\boldsymbol{\alpha}_2^*)} = 0$ for any $\boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^*$ with $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}$. This means if $A_1 \perp\!\!\!\perp \mathbf{A}_{2:K} \mid \mathbf{A}_{2:K} \succeq \boldsymbol{u}$ holds, then we must have $\boldsymbol{p} \in \mathcal{N}_{R,1}$ with $\mathcal{N}_{R,1}$ defined in (8) in Theorem 4.

Now we have proved the statement that

$$A_1 \perp\!\!\!\perp \mathbf{A}_{2:K} \mid \mathbf{A}_{2:K} \succeq \boldsymbol{u},$$

is exactly equivalent to the statement that

$$\boldsymbol{p} \in \mathcal{N}_{R,1} = \{p_{(1,\boldsymbol{\alpha}_1^*)}p_{(0,\boldsymbol{\alpha}_2^*)} - p_{(0,\boldsymbol{\alpha}_1^*)}p_{(1,\boldsymbol{\alpha}_2^*)} = 0 \text{ holds for any } \boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^* \text{ with } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}\}.$$

This completes the proof of Theorem 4. □

S.4 Proof of Theorem 2 and Theorem 5

Proof of Theorem 2. We rewrite the form of $\mathbf{Q}$ in (6) below,

$$\mathbf{Q} = \begin{pmatrix} 1 & \mathbf{0} \\ 1 & \boldsymbol{u}_1 \\ \hline \mathbf{0} & \mathbf{Q}^{(1)} \end{pmatrix}, \quad \mathbf{Q}^{(1)} = \begin{pmatrix} 1 & \mathbf{0} \\ 1 & \boldsymbol{u}_2 \\ \hline \mathbf{0} & \mathbf{Q}^{(2)} \end{pmatrix}, \quad \cdots, \quad \mathbf{Q}^{(m)} = \begin{pmatrix} 1 & \mathbf{0} \\ 1 & \boldsymbol{u}_{m+1} \\ \hline \mathbf{0} & \mathbf{Q}^{(m+1)} \end{pmatrix}.$$

Under the assumption that the first $m+1$ latent attributes are each required by only two items, we know $\boldsymbol{u}_{1,\,1:m} = \mathbf{0}$, $\boldsymbol{u}_{2,\,1:(m-1)} = \mathbf{0}$, $\ldots$, $u_{m,\,1} = 0$. First consider the last $J-m-2$ items corresponding to the bottom $(J-m-2)\times K$ submatrix of $\mathbf{Q}$,

$$(\mathbf{0}, \quad \mathbf{Q}^{(m+1)}) =: \widetilde{\mathbf{Q}}^{(m+1)}$$

The $(J-m-2)\times(K-m-1)$ matrix $\mathbf{Q}^{(m+1)}$ satisfies the C-R-D conditions under the assumption stated in the corollary, and that the first $m+1$ columns of the $\widetilde{\mathbf{Q}}^{(m+1)}$ are all-zero columns. Next we use an argument similar to the proof of Theorem 1. Consider a true set of parameters $(\boldsymbol{\Theta}, \boldsymbol{p})$ and an alternative set $(\bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}})$ with $T(\boldsymbol{\Theta})\boldsymbol{p} = T(\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}}$. Then the following equations must hold for an arbitrary fixed response pattern $\boldsymbol{r} = (r_1, \ldots, r_{m+2}, \boldsymbol{r}^*)$,

$$\begin{aligned} &\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-m-2}} \prod_{j>m+2:\, r_j=1} \theta_{j,\,(\mathbf{0},\boldsymbol{\alpha}^*)} \cdot \mathbb{P}(\mathbf{R}_{1:(m+2)} \geq \boldsymbol{r}_{1:(m+2)},\, \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*) \\ = &\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-m-2}} \prod_{j>m+2:\, r_j=1} \bar{\theta}_{j,\,(\mathbf{0},\boldsymbol{\alpha}^*)} \cdot \overline{\mathbb{P}}(\mathbf{R}_{1:(m+2)} \geq \boldsymbol{r}_{1:(m+2)},\, \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*). \end{aligned}$$

Similar to the argument in the proof of Theorem 1, the fact that $\mathbf{Q}^{(m)}$ satisfies the C-R-D conditions imply $\boldsymbol{c}_{(J-m-1):J} = \bar{\boldsymbol{c}}_{(J-m-1):J}$ and $\boldsymbol{g}_{(J-m-1):J} = \bar{\boldsymbol{g}}_{(J-m-1):J}$, and also imply the following for all $\boldsymbol{\alpha}^* \in \{0,1\}^{K-m-2}$,

$$\mathbb{P}(\mathbf{R}_{1:(m+2)} \geq \boldsymbol{r}_{1:(m+2)},\, \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*) = \overline{\mathbb{P}}(\mathbf{R}_{1:(m+2)} \geq \boldsymbol{r}_{1:(m+2)},\, \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*). \quad \text{(S.10)}$$

Define surrogate (grouped) proportion parameters to be

$$p^{(m)}_{(z,\boldsymbol{\alpha}^*)} = \mathbb{P}(A_{m+1} = z,\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*), \quad z = 0, 1; \tag{S.11}$$

and define $\bar{p}^{(m)}_{(z,\boldsymbol{\alpha}^*)}$ similarly based on the alternative set of parameters $(\bar{\boldsymbol{\Theta}}, \bar{\boldsymbol{p}})$. Now fixing $(r_1, \ldots, r_m)^\top = \mathbf{0}$ and varying $(r_{m+1}, r_{m+2}) \in \{0,1\}^2$, the equality in (S.10) becomes

$$\begin{aligned} &\mathbb{P}((R_{m+1}, R_{m+2}) \geq (r_{m+1}, r_{m+2}),\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*) \\ =\ &\overline{\mathbb{P}}((R_{m+1}, R_{m+2}) \geq (r_{m+1}, r_{m+2}),\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*). \end{aligned}$$

This implies the following equations for any fixed $\boldsymbol{\alpha}^* \succeq \boldsymbol{u}^{(m+1)}$ when (r_{m+1}, r_{m+2}) vary,

$$(r_{m+1}, r_{m+2}) = \begin{cases} (0,0) \implies p^{(m)}_{(0,\boldsymbol{\alpha}^*)} + p^{(m)}_{(1,\boldsymbol{\alpha}^*)} = \bar{p}^{(m)}_{(0,\boldsymbol{\alpha}^*)} + \bar{p}^{(m)}_{(1,\boldsymbol{\alpha}^*)}; \\ (1,0) \implies g_{m+1} \cdot p^{(m)}_{(0,\boldsymbol{\alpha}^*)} + c_{m+1} \cdot p^{(m)}_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_{m+1} \cdot \bar{p}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_{m+1} \cdot \bar{p}^{(m)}_{(1,\boldsymbol{\alpha}^*)}; \\ (0,1) \implies g_{m+2} \cdot p^{(m)}_{(0,\boldsymbol{\alpha}^*)} + c_{m+2} \cdot p^{(m)}_{(1,\boldsymbol{\alpha}^*)} = \bar{g}_{m+2} \cdot \bar{p}^{(m)}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_{m+2} \cdot \bar{p}^{(m)}_{(1,\boldsymbol{\alpha}^*)}; \\ (1,1) \implies g_{m+1}g_{m+2} \cdot p^{(m)}_{(0,\boldsymbol{\alpha}^*)} + c_{m+1}c_{m+2} \cdot p^{(m)}_{(1,\boldsymbol{\alpha}^*)} \\ \qquad\qquad = \bar{g}_{m+1}\bar{g}_{m+2} \cdot \bar{p}^{(m)}_{(0,\boldsymbol{\alpha}^*)} + \bar{c}_{m+1}\bar{c}_{m+2} \cdot \bar{p}^{(m)}_{(1,\boldsymbol{\alpha}^*)}. \end{cases} \tag{S.12}$$

The above system of four equations are similar in form to Eq. (S.7) in the proof of Theorem 1. So following a similar argument as before, we obtain that (c_{m+1}, c_{m+2}) and (g_{m+1}, g_{m+2}) and all the $p^{(m)}_{(z,\boldsymbol{\alpha}^*)}$'s are identifiable as long as $\boldsymbol{p} \in \mathcal{T} \setminus \mathcal{N}_m$ where

$$\mathcal{N}_m = \{p^{(m)}_{(1,\boldsymbol{\alpha}_1^*)} p^{(m)}_{(0,\boldsymbol{\alpha}_2^*)} - p^{(m)}_{(0,\boldsymbol{\alpha}_1^*)} p^{(m)}_{(1,\boldsymbol{\alpha}_2^*)} = 0 \text{ for any } \boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^* \text{ with } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}^{(m+1)}\}. \tag{S.13}$$

Note the definition (S.11) implies that each surrogate proportion $p^{(m)}_{(z,\boldsymbol{\alpha}^*)}$ is a sum of certain individual proportion parameters in that

$$p^{(m)}_{(z,\boldsymbol{\alpha}^*)} = \sum_{\boldsymbol{\beta}\in\{0,1\}^m} p_{(\boldsymbol{\beta},z,\boldsymbol{\alpha}^*)}.$$

Note that the $p^{(m)}_{(z,\boldsymbol{\alpha}^*)}$ defined above exactly characterizes the joint distribution of latent attributes A_m and $\mathbf{A}_{(m+1):K}$. Now we have that the set $\mathcal{N}_m$ defined in (S.13) corresponds to the zero set of certain polynomials about the proportion parameters $\boldsymbol{p}$, so $\mathcal{N}_m$ has Lebesgue measure zero in the parameter space $\mathcal{T}$. Therefore we have shown (c_{m+1}, c_{m+2}), (g_{m+1}, g_{m+2}), and $\boldsymbol{p}^{(m)} := (p^{(m)}_{(z,\boldsymbol{\alpha}^*)};\ (z,\boldsymbol{\alpha}^*) \in \{0,1\}^{K-m})$ are generically identifiable.

Moreover, we go back to the equality in (S.10) and define surrogate proportions alternatively as

$$p^{(m),\boldsymbol{r}}_{(z,\boldsymbol{\alpha}^*)} = \mathbb{P}(\mathbf{R}_{1:m} \succeq \boldsymbol{r}_{1:m},\ A_{m+1} = z,\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*), \quad x = 0, 1;$$

and define $\bar{p}^{(m),\boldsymbol{r}}_{(z,\boldsymbol{\alpha}^*)}$ similarly. Fixing $\boldsymbol{r}_{1:m}$ and varying $(r_{m+1}, r_{m+2}) \in \{0,1\}^2$, Eq. (S.10) can be written in a similar form as the four equations in (S.12), with $p^{(m)}_{(z,\boldsymbol{\alpha}^*)}$ there replaced by $p^{(m),\boldsymbol{r}}_{(z,\boldsymbol{\alpha}^*)}$ now. Since when $\boldsymbol{p} \in \mathcal{T} \setminus \mathcal{N}_m$, we already have the item parameters (c_{m+1}, c_{m+2}) and (g_{m+1}, g_{m+2}) are identifiable, based on the equations about (c_{m+1}, c_{m+2}), (g_{m+1}, g_{m+2}), and $\boldsymbol{p}^{(m),\boldsymbol{r}}$, the parameters $\boldsymbol{p}^{(m),\boldsymbol{r}}$ are also identifiable. Now we write out the equality $\boldsymbol{p}^{(m),\boldsymbol{r}} = \bar{\boldsymbol{p}}^{(m),\boldsymbol{r}}$ by their definitions as

$$\mathbb{P}(\mathbf{R}_{1:m} \geq \boldsymbol{r}_{1:m},\ A_{m+1} = z,\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*) = \overline{\mathbb{P}}(\mathbf{R}_{1:m} \geq \boldsymbol{r}_{1:m},\ A_{m+1} = z,\ \mathbf{A}_{(m+2):K} = \boldsymbol{\alpha}^*),$$

where $(z,\boldsymbol{\alpha}^*) \in \{0,1\}^{K-m}$. Therefore the above equation can be equivalently written as follows, with the new $\boldsymbol{\alpha}^*$ defined to be $(K-m)$-dimensional,

$$\mathbb{P}(\mathbf{R}_{1:m} \geq \boldsymbol{r}_{1:m},\ \mathbf{A}_{(m+1):K} = \boldsymbol{\alpha}^*) = \overline{\mathbb{P}}(\mathbf{R}_{1:m} \geq \boldsymbol{r}_{1:m},\ \mathbf{A}_{(m+1):K} = \boldsymbol{\alpha}^*). \tag{S.14}$$

Comparing the above (S.14) to the previous (S.10) give an immediate similarity, with the difference being only the changes of subscripts of $\mathbf{R}$ and $\mathbf{A}$. Therefore, we can proceed in the same way as before, and show the identifiability of (c_{m-1}, c_m) and (g_{m-1}, g_m) and all the $p^{(m-1)}_{(z,\boldsymbol{\alpha}^*)}$ when $\boldsymbol{p}$ satisfies $\boldsymbol{p} \in \mathcal{T} \setminus (\mathcal{N}_m \cup \mathcal{N}_{m-1})$, where

$$\mathcal{N}_{m-1} = \{p^{(m-1)}_{(1,\boldsymbol{\alpha}_1^*)} p^{(m-1)}_{(0,\boldsymbol{\alpha}_2^*)} - p^{(m-1)}_{(0,\boldsymbol{\alpha}_1^*)} p^{(m-1)}_{(1,\boldsymbol{\alpha}_2^*)} = 0 \text{ for any } \boldsymbol{\alpha}_1^* \neq \boldsymbol{\alpha}_2^* \text{ with } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}^{(m)} \vee (0, \boldsymbol{u}^{(m+1)})\}.$$

In the definition of $\mathcal{N}_{m-1}$, we have $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}^{(m)} \vee (0, \boldsymbol{u}^{(m+1)}) = \widetilde{\boldsymbol{u}}^{(m)} \vee \widetilde{\boldsymbol{u}}^{(m+1)}$ because the $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*$ first need to satisfy the previous requirement before (S.12) and hence $\boldsymbol{\alpha}_{1,-1}^*, \boldsymbol{\alpha}_{2,-1}^* \succeq \boldsymbol{u}^{(m+1)}$ (equivalently, $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq (0, \boldsymbol{u}^{(m+1)})$); and additionally they also need to satisfy the new requirement $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \boldsymbol{u}^{(m)}$.

Recall the definition that $\widetilde{\boldsymbol{u}}^{(\ell)} = (\mathbf{0}, \boldsymbol{u}^{(\ell)})$ is a $(K-1)$-dimensional vector for $\ell = 2, \ldots, m+1$, and $\widetilde{\boldsymbol{u}}^{(1)} = \boldsymbol{u}^{(1)}$ is also a $(K-1)$-dimensional vector. Proceeding in an iterative manner as done in the previous paragraphs, we obtain that as long as $\boldsymbol{p}$ satisfies the following condition, then all the item parameters $\boldsymbol{c}$, $\boldsymbol{g}$ and all the proportion parameters $\boldsymbol{p}$ are identifiable.

$$\boldsymbol{p} \in \mathcal{T} \setminus \left\{ \bigcup_{\ell=0}^{m} \mathcal{N}_\ell \right\},$$

$$\mathcal{N}_\ell = \left\{ p^{(\ell)}_{(1,\boldsymbol{\alpha}_1^*)} p^{(\ell)}_{(0,\boldsymbol{\alpha}_2^*)} - p^{(\ell)}_{(0,\boldsymbol{\alpha}_1^*)} p^{(\ell)}_{(1,\boldsymbol{\alpha}_2^*)} = 0 \text{ for any } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \succeq \bigvee_{t=\ell+1}^{m+1} \widetilde{\boldsymbol{u}}^{(t)} \right\};$$

$$\text{with the definition } \; p^{(\ell)}_{(z,\boldsymbol{\alpha}^*)} = \mathbb{P}(A_{\ell+1} = z,\, \mathbf{A}_{(\ell+2):K} = \boldsymbol{\alpha}^*),$$

Because of the assumption

$$\bigvee_{t=1}^{m+1} \widetilde{\boldsymbol{u}}^{(t)} \neq \mathbf{1}_{K-1}^\top \tag{S.15}$$

stated in the theorem, we claim that the set $\mathcal{T} \setminus \{\bigcup_{\ell=0}^m \mathcal{N}_\ell\}$ is nonempty. To see this, note that $\bigvee_{t=\ell+1}^{m+1} \widetilde{\boldsymbol{u}}^{(t)} \neq \mathbf{1}_{K-1}^\top$ for each $\ell = 0, \ldots, m$ follows from (S.15). This means there must exist two distinct patterns $\boldsymbol{\alpha}_{1,\ell}^* \neq \boldsymbol{\alpha}_{2,\ell}^*$ with $\boldsymbol{\alpha}_{1,\ell}^*,\, \boldsymbol{\alpha}_{2,\ell}^* \succeq \bigvee_{t=\ell+1}^{m+1} \widetilde{\boldsymbol{u}}^{(t)}$. Therefore as long as $\boldsymbol{p}$ satisfies $p^{(\ell)}_{(1,\boldsymbol{\alpha}_{1,\ell}^*)} p^{(\ell)}_{(0,\boldsymbol{\alpha}_{2,\ell}^*)} - p^{(\ell)}_{(0,\boldsymbol{\alpha}_{1,\ell}^*)} p^{(\ell)}_{(1,\boldsymbol{\alpha}_{2,\ell}^*)} \neq 0$ for each $\ell = 0, \ldots, m$, such $\boldsymbol{p}$ does not belong to

$\bigcup_{\ell=0}^{m}\mathcal{N}_\ell$ and hence $\boldsymbol{p}\in\mathcal{T}\setminus\{\bigcup_{\ell=0}^{m}\mathcal{N}_\ell\}$. This proves the earlier claim that the subset of the identifiable parameters $\mathcal{T}\setminus\{\bigcup_{\ell=0}^{m}\mathcal{N}_\ell\}$ is nonempty.

Now note that the subset of the parameter space where identifiability may break down $\bigcup_{\ell=0}^{m}\mathcal{N}_\ell$ is a finite union of several zero sets of polynomial equations about entries of $\boldsymbol{p}$, so it necessarily has Lebesgue measure zero in $\mathcal{T}$. This proves the generic identifiability of parameters $(\boldsymbol{c},\boldsymbol{g},\boldsymbol{p})$ and completes the proof of Theorem 2. Furthermore, note that the $\mathcal{N}_\ell$ in the last paragraph gives the form of the non-identifiable null sets in Theorem 5. Recall that the notation $p^{(\ell)}_{(z,\boldsymbol{\alpha}^*)}$ exactly corresponds to the marginal distribution of the $K-\ell$ latent attributes $A_{\ell+1},\ldots,A_K$. So each set $\mathcal{N}_\ell$ can be equivalently written as

$$\mathcal{N}_\ell=\Big\{A_\ell \perp\!\!\!\perp \mathbf{A}_{(\ell+1):K}\;\Big|\;\Big\{\mathbf{A}_{(\ell+1):K}\succeq\bigvee_{t=\ell+1}^{m+1}\widetilde{\boldsymbol{u}}^{(t)}\Big\}\Big\}.$$

The above set $\mathcal{N}_\ell$ carries the statistical interpretation of latent conditional independence. This completes the proof Theorem 5. □

S.5 Proof of Theorem 3 and Theorem 6

We rewrite the form of the $\mathbf{Q}$-matrix in the theorem below,

$$\mathbf{Q}=\begin{pmatrix}1 & 0 & \mathbf{0}\\ 0 & 1 & \mathbf{0}\\ \hline \boldsymbol{v} & \boldsymbol{v} & \mathbf{Q}^\star\end{pmatrix}=\begin{pmatrix}1 & 0 & \mathbf{0}\\ 0 & 1 & \mathbf{0}\\ \hline \mathbf{0} & \mathbf{0} & \mathbf{Q}'\\ \mathbf{1} & \mathbf{1} & \mathbf{Q}''\end{pmatrix}.$$

Denote the size of the above submatrix $\mathbf{Q}'$ by $J_1\times(K-2)$, then $\mathbf{Q}''$ has size $(J-2-J_1)\times(K-2)$. By Remark 4, we have $J-2-J_1\geq 2$. Consider two sets of DINA model parameters $(\boldsymbol{c},\boldsymbol{g},\boldsymbol{p})$ and $(\bar{\boldsymbol{c}},\bar{\boldsymbol{g}},\bar{\boldsymbol{p}})$ that lead to the same distribution of $\mathbf{R}$ so we have $T(\boldsymbol{\Theta})\boldsymbol{p}=T(\bar{\boldsymbol{\Theta}})\bar{\boldsymbol{p}}$. Theorem 4 in Xu and Zhang (2016) established that if $\mathbf{Q}$ satisfies Conditions (C) and (R), then the guessing parameters associated with those items requiring more than one attribute (i.e., $\{g_j:\ \sum_{k=1}^K q_{j,k}>1\}$) and all the slipping parameters (i.e., $\{c_1,\ldots,c_J\}$) are identifiable. Since the considered $\mathbf{Q}$-matrix satisfies Conditions (C) and (R) by the assumption in the theorem, we have $\boldsymbol{c}=\bar{\boldsymbol{c}}$ and $\boldsymbol{g}_{(3+J_1):J}=\bar{\boldsymbol{g}}_{(3+J_1):J}$.

Next consider an arbitrary $\boldsymbol{\alpha}^*\in\{0,1\}^{K-2}$. The form of the $\mathbf{Q}$-matrix implies

$$\theta_{j,(0,0,\boldsymbol{\alpha}^*)}=\theta_{j,(0,1,\boldsymbol{\alpha}^*)}=\theta_{j,(1,0,\boldsymbol{\alpha}^*)}=\theta_{j,(1,1,\boldsymbol{\alpha}^*)},\quad \forall j\in\{2,\ldots,2+J_1\}.$$

So for a response pattern $\boldsymbol{r}$ with $\boldsymbol{r}_{(3+J_1):J}=\mathbf{0}$, we can write $T_{\boldsymbol{r},:}(\boldsymbol{\Theta})\boldsymbol{p}$ as follows,

$$\begin{aligned}
&T_{\boldsymbol{r},:}(\boldsymbol{\Theta})\boldsymbol{p}\\
=&\sum_{\substack{\boldsymbol{\alpha}\in\{0,1\}^K\\ \boldsymbol{\alpha}=(\alpha_1,\alpha_2,\boldsymbol{\alpha}^*)}} p_{\boldsymbol{\alpha}}\cdot\mathbb{P}(R_1\geq r_1,R_2\geq r_2\mid \mathbf{A}=\boldsymbol{\alpha})\prod_{j=3}^{2+J_1}\theta_{j,(0,0,\boldsymbol{\alpha}^*)}\\
=&\sum_{\boldsymbol{\alpha}^*\in\{0,1\}^{K-2}}\underbrace{\left[\sum_{(\alpha_1,\alpha_2)\in\{0,1\}^2} p_{(\alpha_1,\alpha_2,\boldsymbol{\alpha}^*)}\cdot\mathbb{P}(R_1\geq r_1,R_2\geq r_2\mid \mathbf{A}_{1:2}=(\alpha_1,\alpha_2))\right]}_{\text{define this to be } p^{(r_1,r_2)}_{\boldsymbol{\alpha}^*}}\prod_{j=3}^{2+J_1}\theta_{j,(0,0,\boldsymbol{\alpha}^*)}.
\end{aligned}$$

Now define surrogate DINA model parameters: surrogate proportions $\boldsymbol{p}^{(r_1,r_2)}=(p^{(r_1,r_2)}_{\boldsymbol{\alpha}^*}:\ \boldsymbol{\alpha}^*\in\{0,1\}^{K-2})$ and surrogate item parameters $\boldsymbol{\Theta}^*=\{\theta_{j,(0,0,\boldsymbol{\alpha}^*)}:\ j=3,\ldots,2+J_1;\ \boldsymbol{\alpha}^*\in\{0,1\}^{K-2}\}$. These surrogate parameters $\boldsymbol{p}^{(r_1,r_2)}$ and $\boldsymbol{\Theta}^*$ can be viewed as associated with the $J_1\times(K-2)$ matrix $\mathbf{Q}'$. Since $\mathbf{Q}'$ satisfies the C-R-D conditions, we obtain the identifiability of $\boldsymbol{p}^{(r_1,r_2)}$ and $\boldsymbol{\Theta}^*$. Note that $\boldsymbol{\Theta}^*$ includes all the item parameters associated with items with indices $3,\ldots,J$; i.e., we have established the identifiability of $\{c_3,\ldots,c_{2+J_1},g_3,\ldots,g_{2+J_1}\}$. So far we have obtained $\boldsymbol{c}=\bar{\boldsymbol{c}}$ and $\boldsymbol{g}_{3:J}=\bar{\boldsymbol{g}}_{3:J}$. It only remains to identify $\boldsymbol{p}$ and (g_1,g_2).

The identifiability of $\boldsymbol{p}^{(r_1,r_2)}$ means $\boldsymbol{p}^{(r_1,r_2)} = \bar{\boldsymbol{p}}^{(r_1,r_2)}$ for $(r_1, r_2) \in \{0,1\}^2$, which gives

$$(r_1, r_2) = \begin{cases} (0,0): \; p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)} + p_{(1,1,\boldsymbol{\alpha}^*)} \\ \qquad = \bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(0,1,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}; \\ (1,0): \; g_1[p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}] + c_1[p_{(1,0,\boldsymbol{\alpha}^*)} + p_{(1,1,\boldsymbol{\alpha}^*)}] \\ \qquad = \bar{g}_1[\bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(0,1,\boldsymbol{\alpha}^*)}] + c_1[\bar{p}_{(1,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}]; \\ (0,1): \; g_2[p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)}] + c_2[p_{(0,1,\boldsymbol{\alpha}^*)} + p_{(1,1,\boldsymbol{\alpha}^*)}] \\ \qquad = \bar{g}_2[\bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,0,\boldsymbol{\alpha}^*)}] + c_2[\bar{p}_{(0,1,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}]; \\ (1,1): \; g_1 g_2 p_{(0,0,\boldsymbol{\alpha}^*)} + c_1 g_2 p_{(1,0,\boldsymbol{\alpha}^*)} + g_1 c_2 p_{(0,1,\boldsymbol{\alpha}^*)} + c_1 c_2 p_{(1,1,\boldsymbol{\alpha}^*)} \\ \qquad = \bar{g}_1 \bar{g}_2 \bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + c_1 \bar{g}_2 \bar{p}_{(1,0,\boldsymbol{\alpha}^*)} + \bar{g}_1 c_2 \bar{p}_{(0,1,\boldsymbol{\alpha}^*)} + c_1 c_2 \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}. \end{cases} \tag{S.16}$$

Since $\mathbf{Q}'$ satisfies Condition (C) and contains a submatrix $\mathbf{I}_{K-2}$, we can assume without loss of generality that the first $K-2$ rows of $\mathbf{Q}'$ form $\mathbf{I}_{K-2}$; namely, the first K rows of $\mathbf{Q}$ forms an identity matrix $\mathbf{I}_K$. According to the form of $\mathbf{Q}$, let $\boldsymbol{q}_m = (1,1,0,\dots,0)$ for some $m \in \{3 + J_1, \dots, J\}$. Given an arbitrary pattern $\boldsymbol{\alpha}^* = (\alpha_3, \dots, \alpha_K) \in \{0,1\}^{K-2}$, define

$$\boldsymbol{\theta}^* = \sum_{\substack{3 \le k \le K: \\ \alpha_k = 1}} g_k \boldsymbol{e}_k + \sum_{\substack{3 \le k \le K: \\ \alpha_k = 0}} c_k \boldsymbol{e}_k + g_m \boldsymbol{e}_m.$$

Then $T_{\boldsymbol{r},:}(\boldsymbol{\Theta} - \boldsymbol{\theta}^* \cdot \mathbf{1}_{2^K})\boldsymbol{p} = T_{\boldsymbol{r},:}(\bar{\boldsymbol{\Theta}} - \boldsymbol{\theta}^* \cdot \mathbf{1}_{2^K})\bar{\boldsymbol{p}}$ gives

$$\begin{aligned} & p_{(1,1,\boldsymbol{\alpha}^*)} \prod_{\substack{3 \le k \le K: \\ \alpha_k = 1}} (c_k - g_k) \prod_{\substack{3 \le k \le K: \\ \alpha_k = 0}} (g_k - c_k)(c_m - g_m) \\ = {} & \bar{p}_{(1,1,\boldsymbol{\alpha}^*)} \prod_{\substack{3 \le k \le K: \\ \alpha_k = 1}} (c_k - g_k) \prod_{\substack{3 \le k \le K: \\ \alpha_k = 0}} (g_k - c_k)(c_m - g_m), \end{aligned}$$

which implies $p_{(1,1,\boldsymbol{\alpha}^*)} = \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}$. Note that this identifiability conclusion holds for any $\boldsymbol{\alpha}^* \in \{0,1\}^K$. Plugging the $p_{(1,1,\boldsymbol{\alpha}^*)} = \bar{p}_{(1,1,\boldsymbol{\alpha}^*)}$ into (S.16) gives the following equations about undetermined parameters $\bar{g}_1$, $\bar{g}_2$, and $\{p_{(0,0,\boldsymbol{\alpha}^*)}, p_{(0,1,\boldsymbol{\alpha}^*)}, p_{(1,0,\boldsymbol{\alpha}^*)} : \boldsymbol{\alpha}^* \in \{0,1\}^{K-2}\}$,

$$(r_1, r_2) = \begin{cases} (0,0) \implies p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)} = \bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(0,1,\boldsymbol{\alpha}^*)}; \\ (1,0) \implies g_1[p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}] + c_1 p_{(1,0,\boldsymbol{\alpha}^*)} = \bar{g}_1[\bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(0,1,\boldsymbol{\alpha}^*)}] + c_1 \bar{p}_{(1,0,\boldsymbol{\alpha}^*)}; \\ (0,1) \implies g_2[p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)}] + c_2 p_{(0,1,\boldsymbol{\alpha}^*)} = \bar{g}_2[\bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + \bar{p}_{(1,0,\boldsymbol{\alpha}^*)}] + c_2 \bar{p}_{(0,1,\boldsymbol{\alpha}^*)}; \\ (1,1) \implies g_1 g_2 p_{(0,0,\boldsymbol{\alpha}^*)} + c_1 g_2 p_{(1,0,\boldsymbol{\alpha}^*)} + g_1 c_2 p_{(0,1,\boldsymbol{\alpha}^*)} \\ \qquad\qquad = \bar{g}_1 \bar{g}_2 \bar{p}_{(0,0,\boldsymbol{\alpha}^*)} + c_1 \bar{g}_2 \bar{p}_{(1,0,\boldsymbol{\alpha}^*)} + \bar{g}_1 c_2 \bar{p}_{(0,1,\boldsymbol{\alpha}^*)}. \end{cases} \tag{S.17}$$

After some transformation, (S.17) yields

$$\begin{cases} (g_1 - \bar{g}_1)(p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}) + (c_1 - \bar{g}_1) p_{(1,0,\boldsymbol{\alpha}^*)} = (c_1 - \bar{g}_1) \bar{p}_{(1,0,\boldsymbol{\alpha}^*)}, \\ (g_1 - \bar{g}_1)(g_2 - c_2) p_{(0,0,\boldsymbol{\alpha}^*)} + (c_1 - \bar{g}_1)(g_2 - c_2) p_{(1,0,\boldsymbol{\alpha}^*)} = (c_1 - \bar{g}_1)(\bar{g}_2 - c_2) \bar{p}_{(1,0,\boldsymbol{\alpha}^*)}. \end{cases} \tag{S.18}$$

The right hand sides of both of the above equations are nonzero. So we can take the ratio of these two equations to obtain

$$\frac{(g_1 - \bar{g}_1) p_{(0,0,\boldsymbol{\alpha}^*)} / p_{(1,0,\boldsymbol{\alpha}^*)} + (c_1 - \bar{g}_1)}{(g_1 - \bar{g}_1)[p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}] / p_{(1,0,\boldsymbol{\alpha}^*)} + (c_1 - \bar{g}_1)} (g_2 - c_2) = \bar{g}_2 - c_2.$$

Define $f(\boldsymbol{\alpha}^*) = p_{(0,0,\boldsymbol{\alpha}^*)} / p_{(1,0,\boldsymbol{\alpha}^*)}$, $g(\boldsymbol{\alpha}^*) = [p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}] / p_{(1,0,\boldsymbol{\alpha}^*)}$ as functions of $\boldsymbol{\alpha}^*$, then the above equation can be written as

$$\frac{A \cdot f(\boldsymbol{\alpha}^*) + B}{A \cdot g(\boldsymbol{\alpha}^*) + B} = C,$$

where $A = g_1 - \bar{g}_1$, $B = c_1 - \bar{g}_1$, and $C = \bar{g}_2 - c_2$. So we have

$$A \cdot (f(\boldsymbol{\alpha}^*) - C \cdot g(\boldsymbol{\alpha}^*)) = BC - B,$$

which is equivalent to

$$(g_1 - \bar{g}_1) \cdot \left[\frac{p_{(0,0,\boldsymbol{\alpha}^*)}}{p_{(1,0,\boldsymbol{\alpha}^*)}} - (\bar{g}_2 - c_2) \frac{p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}}{p_{(1,0,\boldsymbol{\alpha}^*)}} \right] = (c_1 - \bar{g}_1)(\bar{g}_2 - c_2) - (c_1 - \bar{g}_1).$$

Consider $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*$, we further obtain the following function $h(\boldsymbol{\alpha}^*)$ does not depend on $\boldsymbol{\alpha}^*$,

$$\begin{aligned} h(\boldsymbol{\alpha}^*) &:= (g_1 - \bar{g}_1) \cdot \left[\frac{p_{(0,0,\boldsymbol{\alpha}^*)}}{p_{(1,0,\boldsymbol{\alpha}^*)}} - (\bar{g}_2 - c_2) \frac{p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)}}{p_{(1,0,\boldsymbol{\alpha}^*)}} \right] \\ &= (g_1 - \bar{g}_1) \cdot \frac{p_{(0,0,\boldsymbol{\alpha}^*)} + (c_2 - \bar{g}_2)(p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)})}{p_{(1,0,\boldsymbol{\alpha}^*)}}; \end{aligned}$$

therefore we have

$$\begin{aligned} 0 &= h(\boldsymbol{\alpha}_1^*) - h(\boldsymbol{\alpha}_2^*) \\ &= (g_1 - \bar{g}_1) \cdot \Big[\frac{p_{(0,0,\boldsymbol{\alpha}_1^*)} + (c_2 - \bar{g}_2)(p_{(0,0,\boldsymbol{\alpha}_1^*)} + p_{(0,1,\boldsymbol{\alpha}_1^*)})}{p_{(1,0,\boldsymbol{\alpha}_1^*)}} \\ &\qquad - \frac{p_{(0,0,\boldsymbol{\alpha}_2^*)} + (c_2 - \bar{g}_2)(p_{(0,0,\boldsymbol{\alpha}_2^*)} + p_{(0,1,\boldsymbol{\alpha}_2^*)})}{p_{(1,0,\boldsymbol{\alpha}_2^*)}} \Big] \\ &= (g_1 - \bar{g}_1) \frac{1}{p_{(1,0,\boldsymbol{\alpha}_1^*)} p_{(1,0,\boldsymbol{\alpha}_2^*)}} \Big\{ [p_{(0,0,\boldsymbol{\alpha}_1^*)} + (c_2 - \bar{g}_2)(p_{(0,0,\boldsymbol{\alpha}_1^*)} + p_{(0,1,\boldsymbol{\alpha}_1^*)})] p_{(1,0,\boldsymbol{\alpha}_2^*)} \\ &\qquad - [p_{(0,0,\boldsymbol{\alpha}_2^*)} + (c_2 - \bar{g}_2)(p_{(0,0,\boldsymbol{\alpha}_2^*)} + p_{(0,1,\boldsymbol{\alpha}_2^*)})] p_{(1,0,\boldsymbol{\alpha}_1^*)} \Big\}. \end{aligned}$$

According to the above equality, if $g_1 - \bar{g}_1 \neq 0$, then $h(\boldsymbol{\alpha}_1^*) - h(\boldsymbol{\alpha}_2^*) = 0$ gives

$$\begin{aligned} & p_{(0,0,\boldsymbol{\alpha}_1^*)} p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)} p_{(1,0,\boldsymbol{\alpha}_1^*)} \\ & + (c_2 - \bar{g}_2)[(p_{(0,0,\boldsymbol{\alpha}_1^*)} + p_{(0,1,\boldsymbol{\alpha}_1^*)}) p_{(1,0,\boldsymbol{\alpha}_2^*)} - (p_{(0,0,\boldsymbol{\alpha}_2^*)} + p_{(0,1,\boldsymbol{\alpha}_2^*)}) p_{(1,0,\boldsymbol{\alpha}_1^*)}] = 0. \end{aligned} \tag{S.19}$$

We rewrite below the definitions of the functions m_1, m_2, m_3 stated in (11) in the theorem,

$$\begin{cases} m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,1,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,1,\boldsymbol{\alpha}_2^*)}p_{(1,0,\boldsymbol{\alpha}_1^*)}, \\ m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,0,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)}p_{(1,0,\boldsymbol{\alpha}_1^*)}, \\ m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,0,\boldsymbol{\alpha}_1^*)}p_{(0,1,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)}p_{(0,1,\boldsymbol{\alpha}_1^*)}. \end{cases}$$

Then (S.19) can be written as

$$m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) + (c_2 - \bar{g}_2)[m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) + m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)] = 0. \tag{S.20}$$

Note that $c_2 - \bar{g}_2 \neq 0$. If $m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) \neq 0$ holds for some $\boldsymbol{\alpha}_1^*$ and $\boldsymbol{\alpha}_2^*$, then we can obtain the following from (S.20),

$$\frac{m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)}{m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)} := \frac{p_{(0,1,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,1,\boldsymbol{\alpha}_2^*)}p_{(1,0,\boldsymbol{\alpha}_1^*)}}{p_{(0,0,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)}p_{(1,0,\boldsymbol{\alpha}_1^*)}} = \frac{1}{\bar{g}_2 - c_2} - 1. \tag{S.21}$$

Therefore, as long as there exist $\boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*,\ \boldsymbol{\beta}_1^*,\ \boldsymbol{\beta}_2^* \in \{0,1\}^{K-2}$ such that $\boldsymbol{p}$ satisfies

$$\frac{m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)}{m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)} \neq \frac{m_1(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*)}{m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*)}, \quad m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) \neq 0, \quad m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) \neq 0,$$

then (S.21) cannot hold true; such a contradiction implies the earlier assumption $g_1 - \bar{g}_1 \neq 0$ is incorrect, and we should have $g_1 = \bar{g}_1$. Equivalently, we have shown that if there exist $\boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*,\ \boldsymbol{\beta}_1^*,\ \boldsymbol{\beta}_2^* \in \{0,1\}^{K-2}$ such that

$$m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) - m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)m_1(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) \neq 0, \quad m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) \neq 0, \quad m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) \neq 0,$$

then $g_1 = \bar{g}_1$ and hence parameter g_1 is identifiable.

Define a subset $\mathcal{N}_{D,1}$ of the parameter space $\mathcal{T}$ to be

$$\begin{aligned} \mathcal{N}_{D,1} = \{&\text{For all } \boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*,\ \boldsymbol{\beta}_1^*,\ \boldsymbol{\beta}_2^* \in \{0,1\}^{K-2}, \\ &\qquad \text{Either } \ m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) - m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)m_1(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) = 0, \end{aligned}$$

$$\text{Or}\ \ m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = 0,\ \text{Or}\ \ m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) = 0.\}$$

$$= \{\text{For all } \boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*,\ \boldsymbol{\beta}_1^*,\ \boldsymbol{\beta}_2^* \in \{0,1\}^{K-2},$$

$$m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) \cdot m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) \cdot [m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) m_2(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) - m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) m_1(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*)] = 0.\}.$$

Then we have established that as long as $\boldsymbol{p} \in \mathcal{T} \setminus \mathcal{N}_{D,1}$, then $g_1 = \bar{g}_1$ and parameter g_1 is identifiable. By the symmetry between g_1 and g_2, we similarly obtain that if $\boldsymbol{p} \in \mathcal{T} \setminus \mathcal{N}_{D,2}$, then $g_2 = \bar{g}_2$ and parameter g_2 is identifiable, where $\mathcal{N}_{D,2}$ takes the following form,

$$\mathcal{N}_{D,2} = \{\text{For all } \boldsymbol{\alpha}_1^*,\ \boldsymbol{\alpha}_2^*,\ \boldsymbol{\beta}_1^*,\ \boldsymbol{\beta}_2^* \in \{0,1\}^{K-2},$$

$$m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) \cdot m_3(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) \cdot [m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) m_3(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*) - m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) m_1(\boldsymbol{\beta}_1^*, \boldsymbol{\beta}_2^*)] = 0.\}.$$

The function $m_3(\cdot,\cdot)$ has been defined earlier together with $m_1(\cdot,\cdot)$ and $m_2(\cdot,\cdot)$. In summary, if $\boldsymbol{p} \in \mathcal{T} \setminus (\mathcal{N}_{D,1} \cup \mathcal{N}_{D,2})$, then g_1 and g_2 are identifiable.

Recall that we previously have already proved the identifiability of all the other item parameters and also identifiability of $\{p_{(1,1,\boldsymbol{\alpha}^*)} :\ \boldsymbol{\alpha}^* \in \{0,1\}^{K-2}\}$. Now we can replace $\bar{g}_1$ by g_1 in the first equation in (S.18) and obtain $\bar{p}_{(1,0,\boldsymbol{\alpha}^*)} = p_{(1,0,\boldsymbol{\alpha}^*)}$; similarly, replacing $\bar{g}_2$ by g_2 in (S.17) gives $\bar{p}_{(0,1,\boldsymbol{\alpha}^*)} = p_{(0,1,\boldsymbol{\alpha}^*)}$. With $\bar{p}_{(1,0,\boldsymbol{\alpha}^*)}$ and $\bar{p}_{(0,1,\boldsymbol{\alpha}^*)}$ both determined, (S.17) finally gives $\bar{p}_{(1,1,\boldsymbol{\alpha}^*)} = p_{(1,1,\boldsymbol{\alpha}^*)}$. Noting that the above argument holds for an arbitrary $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$, we have established the identifiability of all the parameters under the DINA model under the condition that the true proportion parameters $\boldsymbol{p}$ satisfies $\boldsymbol{p} \in \mathcal{T} \setminus (\mathcal{N}_{D,1} \cup \mathcal{N}_{D,2})$. Note that the set $\mathcal{N}_{D,1} \cup \mathcal{N}_{D,2}$ where identifiability potentially breaks down is characterized by the zero sets of certain nontrivial polynomial equations about the entries of $\boldsymbol{p}$, and hence necessarily has Lebesgue measure zero in the parameter space $\mathcal{T}$. This proves the conclusion of generic identifiability and concludes the proof of Theorem 3. Further note that the forms of $\mathcal{N}_{D,1}$ and $\mathcal{N}_{D,2}$ defined in the last paragraph are exactly the same as those stated in Theorem 6, so we have also proved Theorem 6. □

S.6 Proof of Proposition 3

We introduce some new notation to facilitate understanding the null sets $\mathcal{N}_{D,1}$ and $\mathcal{N}_{D,2}$. Consider the joint distribution of two discrete random variables $Z_1 := (A_1, A_2)$ and $Z_2 := (A_3, \ldots, A_K)$, each concatenated from the latent attributes. That is, Z_1 concatenates two variables A_1 and A_2 and takes $|\{0,1\}^2| = 4$ possible values, and Z_2 concatenates $K-2$ binary variables and takes $|\{0,1\}^{K-2}| = 2^{K-2}$ possible values. The joint distribution of Z_1 and Z_2 can be written in the form of a $4 \times 2^{K-2}$ contingency table, whose rows are indexed by the possible values Z_1 can take and columns by the possible values Z_2 can take. Each entry in this table represents the probability of a specific configuration of (Z_1, Z_2). We write out this $4 \times 2^{K-2}$ table below and denote it by $\mathcal{B}$,

$$\begin{array}{c} \\ (00) \\ (10) \\ (01) \\ (11) \end{array} \begin{array}{c} \begin{array}{cccc} (10\cdots0) & (01\cdots0) & \cdots & (11\cdots1) \end{array} \\ \begin{pmatrix} p_{(00,10\cdots0)} & p_{(00,01\cdots0)} & \cdots & p_{(00,11\cdots1)} \\ p_{(10,10\cdots0)} & p_{(10,01\cdots0)} & \cdots & p_{(10,11\cdots1)} \\ p_{(01,10\cdots0)} & p_{(01,01\cdots0)} & \cdots & p_{(01,11\cdots1)} \\ p_{(11,10\cdots0)} & p_{(11,01\cdots0)} & \cdots & p_{(11,11\cdots1)} \end{pmatrix} \end{array} \tag{S.22}$$

Note that when the previously used notation $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$ can indicate the configurations of Z_2, so the above matrix $\mathcal{B}$ have columns indexed by $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$. The definition of $m_i(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)$, $i = 1,2,3$ can be understood as certain 2×2 minor of the matrix $\mathcal{B}$. Denote the determinant of a matrix $\mathbf{C}$ by $|\mathbf{C}|$. In particular, we have the following equalities,

$$m_1(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,1,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(1,0,\boldsymbol{\alpha}_1^*)}p_{(0,1,\boldsymbol{\alpha}_2^*)} = \begin{vmatrix} p_{(0,1,\boldsymbol{\alpha}_1^*)} & p_{(0,1,\boldsymbol{\alpha}_2^*)} \\ p_{(1,0,\boldsymbol{\alpha}_1^*)} & p_{(1,0,\boldsymbol{\alpha}_2^*)} \end{vmatrix} = |\mathcal{B}(\{2,3\},\ \{\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*\})|,$$

$$m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,0,\boldsymbol{\alpha}_1^*)}p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(1,0,\boldsymbol{\alpha}_1^*)}p_{(0,0,\boldsymbol{\alpha}_2^*)} = \begin{vmatrix} p_{(0,0,\boldsymbol{\alpha}_1^*)} & p_{(0,0,\boldsymbol{\alpha}_2^*)} \\ p_{(1,0,\boldsymbol{\alpha}_1^*)} & p_{(1,0,\boldsymbol{\alpha}_2^*)} \end{vmatrix} = |\mathcal{B}(\{1,2\},\ \{\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*\})|,$$

$$m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = p_{(0,0,\boldsymbol{\alpha}_1^*)}p_{(0,1,\boldsymbol{\alpha}_2^*)} - p_{(0,1,\boldsymbol{\alpha}_1^*)}p_{(0,0,\boldsymbol{\alpha}_2^*)} = \begin{vmatrix} p_{(0,0,\boldsymbol{\alpha}_1^*)} & p_{(0,0,\boldsymbol{\alpha}_2^*)} \\ p_{(0,1,\boldsymbol{\alpha}_1^*)} & p_{(0,1,\boldsymbol{\alpha}_2^*)} \end{vmatrix} = |\mathcal{B}(\{1,3\},\ \{\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*\})|.$$

In the above display, the $\mathcal{B}(\{1,2\}, \{\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*\})$ denotes the 2×2 submatrix of $\mathcal{B}$ containing the entries in rows with indices 1, 2 and columns $\boldsymbol{\alpha}_1^*$, $\boldsymbol{\alpha}_2^*$.

We can use the technical machinery in the last paragraph to discover some meaningful subsets of the non-identifiable null set $\mathcal{N}_{D,1} \cup \mathcal{N}_{D,2}$. First, define

$$\mathcal{N}_{1,\text{sub}} = \{m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = 0 \text{ for all } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \in \{0,1\}^{K-2}\}, \tag{S.23}$$

$$\mathcal{N}_{2,\text{sub}} = \{m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = 0 \text{ for all } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \in \{0,1\}^{K-2}\}. \tag{S.24}$$

According to the definition of $\mathcal{N}_{D,1}$ and $\mathcal{N}_{D,2}$, it is clear that the two sets defined above satisfy $\mathcal{N}_{1,\text{sub}} \subseteq \mathcal{N}_{D,1}$ and $\mathcal{N}_{2,\text{sub}} \subseteq \mathcal{N}_{D,2}$. First consider the statistical implication of $\mathcal{N}_{1,\text{sub}}$. Since $m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = |\mathcal{B}(\{1,2\}, \{\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*\})|$, when $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*$ range over all the possible patterns in $\{0,1\}^{K-2}$, the $m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*)$ will take on values of all the possible 2×2 minors of the $2 \times 2^{(K-2)}$ matrix $\mathcal{B}(\{1,2\}, :)$ (i.e., the submatrix of $\mathcal{B}$ consisting of its first two rows). The assertion in $\mathcal{N}_{1,\text{sub}}$ that all these determinants equal zero essentially implies the submatrix $\mathcal{B}(\{1,2\}, :)$ has rank one, i.e., has the two rows proportional to each other. This means for all $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$, the ratio $p_{(1,0,\boldsymbol{\alpha}^*)}/p_{(0,0,\boldsymbol{\alpha}^*)}$ is a constant δ, which further implies the ratio $p_{(1,0,\boldsymbol{\alpha}^*)}/(p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)})$ is also a constant equal to $1/(1+1/\delta)$, which we denote by ρ:

$$\begin{aligned}\rho = \frac{p_{(1,0,\boldsymbol{\alpha}^*)}}{p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)}} &= \frac{\mathbb{P}(A_1 = 1, A_2 = 0, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)}{\mathbb{P}(A_2 = 0, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)} \\ &= \frac{\mathbb{P}(A_1 = 1, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0)}{\mathbb{P}(\mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0)}, \quad \forall \boldsymbol{\alpha}^* \in \{0,1\}^{K-2}.\end{aligned}$$

So we have the following

$$\mathbb{P}(A_1 = 1, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0) = \rho \cdot \mathbb{P}(\mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0). \tag{S.25}$$

Now summing over the above equation for all $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$, we obtain

$$\sum_{\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}} \mathbb{P}(A_1 = 1, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0) = \rho \cdot \sum_{\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}} \mathbb{P}(\mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0),$$

$$\implies \quad \mathbb{P}(A_1 = 1 \mid A_2 = 0) = \rho.$$

Plugging back $\rho = \mathbb{P}(A_1 = 1 \mid A_2 = 0)$ into (S.25) gives the following for all $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$,

$$\mathbb{P}(A_1 = 1, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0) = \mathbb{P}(A_1 = 1 \mid A_2 = 0) \cdot \mathbb{P}(\mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0);$$

in a very similar fashion we can also obtain $\mathbb{P}(A_1 = 0, \mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0) = \mathbb{P}(A_1 = 0 \mid A_2 = 0) \cdot \mathbb{P}(\mathbf{A}_{3:K} = \boldsymbol{\alpha}^* \mid A_2 = 0)$ for all $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$. This essentially means attribute A_1 and attributes $\mathbf{A}_{3:K}$ are conditionally independent given $A_2 = 0$. So we have obtained that $\boldsymbol{p} \in \mathcal{N}_{1,\text{sub}}$ implies A_1 and $\mathbf{A}_{3:K}$ are conditionally independent given $A_2 = 0$. By symmetry, we similarly have that $\boldsymbol{p} \in \mathcal{N}_{2,\text{sub}}$ implies A_2 and $\mathbf{A}_{3:K}$ are conditionally independent given $A_1 = 0$. In summary, we have proved that $\mathcal{N}_{1,\text{sub}}$ and $\mathcal{N}_{2,\text{sub}}$ defined in (S.23)-(S.24) correspond to the following conditional independence statements,

$$\mathcal{N}_{1,\text{sub}} = \{\boldsymbol{p} \text{ satisfies } (A_1 \perp\!\!\!\perp \mathbf{A}_{3:K} \mid A_2 = 0)\} \subseteq \mathcal{N}_{D,1};$$
$$\mathcal{N}_{2,\text{sub}} = \{\boldsymbol{p} \text{ satisfies } (A_2 \perp\!\!\!\perp \mathbf{A}_{3:K} \mid A_1 = 0)\} \subseteq \mathcal{N}_{D,2}.$$

Additionally, by the basic property of marginal independence and conditional independence, if $\boldsymbol{p}$ satisfies the marginal independence statement such as "$A_1 \perp\!\!\!\perp \mathbf{A}_{3:K}$", then it necessarily also satisfies the conditional independence statement "$A_1 \perp\!\!\!\perp \mathbf{A}_{3:K} \mid A_2 = 0$". Therefore we have we also have

$$\mathcal{N}_{1,\text{sub}} = \{\boldsymbol{p} \text{ satisfies } (A_1 \perp\!\!\!\perp \mathbf{A}_{3:K} \mid A_2 = 0)\} \supseteq \{\boldsymbol{p} \text{ satisfies } (A_1 \perp\!\!\!\perp \mathbf{A}_{3:K})\};$$
$$\mathcal{N}_{2,\text{sub}} = \{\boldsymbol{p} \text{ satisfies } (A_2 \perp\!\!\!\perp \mathbf{A}_{3:K} \mid A_1 = 0)\} \supseteq \{\boldsymbol{p} \text{ satisfies } (A_2 \perp\!\!\!\perp \mathbf{A}_{3:K})\}.$$

Combining the two conclusions above, we have proved the first two conclusions in (12) in Proposition 3.

Next we prove the third conclusion in (12) in Proposition 3. Define

$$\mathcal{N}_{\text{both}} = \{m_2(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = m_3(\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^*) = 0 \text{ holds for all } \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \in \{0,1\}^{K-2}.\} \tag{S.26}$$

First note that $\mathcal{N}_{\text{both}} \subseteq \mathcal{N}_{D,1} \cap \mathcal{N}_{D,2}$ obviously holds according to definition of $\mathcal{N}_{D,1}$ and $\mathcal{N}_{D,2}$. We next examine the statistical implication the set $\mathcal{N}_{\text{both}}$. If $\boldsymbol{p} \in \mathcal{N}_{\text{both}}$, then we have the following for all $\boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \in \{0,1\}^{K-2}$,

$$\begin{aligned}
& p_{(0,0,\boldsymbol{\alpha}_1^*)} p_{(1,0,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)} p_{(1,0,\boldsymbol{\alpha}_1^*)} = p_{(0,0,\boldsymbol{\alpha}_1^*)} p_{(0,1,\boldsymbol{\alpha}_2^*)} - p_{(0,0,\boldsymbol{\alpha}_2^*)} p_{(0,1,\boldsymbol{\alpha}_1^*)} = 0; \\
\Longrightarrow \quad & \frac{p_{(1,0,\boldsymbol{\alpha}_1^*)}}{p_{(0,0,\boldsymbol{\alpha}_1^*)}} = \frac{p_{(1,0,\boldsymbol{\alpha}_2^*)}}{p_{(0,0,\boldsymbol{\alpha}_2^*)}}, \quad \frac{p_{(0,1,\boldsymbol{\alpha}_1^*)}}{p_{(0,0,\boldsymbol{\alpha}_1^*)}} = \frac{p_{(0,1,\boldsymbol{\alpha}_2^*)}}{p_{(0,0,\boldsymbol{\alpha}_2^*)}}, \quad \forall \boldsymbol{\alpha}_1^*, \boldsymbol{\alpha}_2^* \in \{0,1\}^{K-2}.
\end{aligned}$$

This implies there exist some constants ρ_1, ρ_2 such that

$$\frac{p_{(1,0,\boldsymbol{\alpha}^*)}}{p_{(0,0,\boldsymbol{\alpha}^*)}} = \rho_1, \quad \frac{p_{(0,1,\boldsymbol{\alpha}^*)}}{p_{(0,0,\boldsymbol{\alpha}^*)}} = \rho_2, \quad \forall \boldsymbol{\alpha}^* \in \{0,1\}^{K-2}. \tag{S.27}$$

Then for arbitrary $(x,y) \in \{(0,0),\ (0,1),\ (1,0)\}$ and $\boldsymbol{\alpha}^* \in \{0,1\}^{K-2}$, we will have

$$\begin{aligned}
& \mathbb{P}(\mathbf{A}_{1:2} = (x,y) \mid \mathbf{A}_{1:2} \neq (1,1),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*) \\
= & \frac{\mathbb{P}(\mathbf{A}_{1:2} = (x,y),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)}{\mathbb{P}(\mathbf{A}_{1:2} \neq (1,1),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)} \\
= & \frac{p_{(x,y,\boldsymbol{\alpha}^*)}}{p_{(0,0,\boldsymbol{\alpha}^*)} + p_{(0,1,\boldsymbol{\alpha}^*)} + p_{(1,0,\boldsymbol{\alpha}^*)}} = \frac{\frac{p_{(x,y,\boldsymbol{\alpha}^*)}}{p_{(0,1,\boldsymbol{\alpha}^*)}}}{1 + \frac{p_{(0,1,\boldsymbol{\alpha}^*)}}{p_{(0,1,\boldsymbol{\alpha}^*)}} + \frac{p_{(1,0,\boldsymbol{\alpha}^*)}}{p_{(0,1,\boldsymbol{\alpha}^*)}}} \\
= & \begin{cases} \dfrac{1}{1+\rho_1+\rho_2}, & \text{if } (x,y) = (0,0); \\ \dfrac{\rho_1}{1+\rho_1+\rho_2}, & \text{if } (x,y) = (1,0); \\ \dfrac{\rho_2}{1+\rho_1+\rho_2}, & \text{if } (x,y) = (0,1). \end{cases}
\end{aligned} \tag{S.28}$$

The above deduction implies that the conditional distribution $\mathbb{P}(\mathbf{A}_{1:2} = (x,y) \mid \mathbf{A}_{1:2} \neq$

$(1,1),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)$ does not depend on $\mathbf{A}_{3:K}$ and hence can be indeed written as

$$\mathbb{P}(\mathbf{A}_{1:2} = (x,y) \mid \mathbf{A}_{1:2} \neq (1,1),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*) = \mathbb{P}(\mathbf{A}_{1:2} = (x,y) \mid \mathbf{A}_{1:2} \neq (1,1)).$$

Statistically, the above observation means the conditional independence $(\mathbf{A}_{1:2} \perp\!\!\!\perp \mathbf{A}_{3:K} \mid \mathbf{A}_{1:2} \neq (1,1))$ holds. Also, note that in order for $\mathbb{P}(\mathbf{A}_{1:2} = (x,y) \mid \mathbf{A}_{1:2} \neq (1,1),\ \mathbf{A}_{3:K} = \boldsymbol{\alpha}^*)$ in (S.28) to not depend on $\boldsymbol{\alpha}^*$, we must have (S.27) holds for some constants ρ_1, ρ_2. In summary, we have shown that $\boldsymbol{p} \in \mathcal{N}_{\text{both}}$ if and only if $(\mathbf{A}_{1:2} \perp\!\!\!\perp \mathbf{A}_{3:K} \mid \mathbf{A}_{1:2} \neq (1,1))$ holds. Namely, the $\mathcal{N}_{\text{both}}$ defined in (S.26) can be equivalently written as

$$\mathcal{N}_{\text{both}} = \{\boldsymbol{p} \text{ satisfies } (\mathbf{A}_{1:2} \perp\!\!\!\perp \mathbf{A}_{3:K} \mid \mathbf{A}_{1:2} \neq (1,1))\}.$$

Finally, recall that we have $\mathcal{N}_{\text{both}} \subseteq \mathcal{N}_{D,1} \cap \mathcal{N}_{D,2}$, so

$$\mathcal{N}_{D,1} \cap \mathcal{N}_{D,2} \supseteq \{\boldsymbol{p} \text{ satisfies } (\mathbf{A}_{1:2} \perp\!\!\!\perp \mathbf{A}_{3:K} \mid \mathbf{A}_{1:2} \neq (1,1))\} \supseteq \{\boldsymbol{p} \text{ satisfies } (\mathbf{A}_{1:2} \perp\!\!\!\perp \mathbf{A}_{3:K})\}.$$

This completes the proof of Proposition 3. □

References

Gu, Y. and Xu, G. (2019). The sufficient and necessary condition for the identifiability and estimability of the DINA model. *Psychometrika*, 84(2):468–483.

Gu, Y. and Xu, G. (2020). Partial identifiability of restricted latent class models. *Annals of Statistics*, 48(4):2082–2107.

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

Xu, G. and Zhang, S. (2016). Identifiability of diagnostic classification models. *Psychometrika*, 81(3):625–649.