**Online Appendix**

Evaluating the Fit of Sequential G-DINA Model using Limited-information Measures

## Appendix I: Model dissimilarity analysis

This appendix presents some details about the model dissimilarity analysis in the article. Note that in this paper, the q-vector for each category in the simulated Q-matrix required at most two attributes. When only one attribute is needed, all processing functions are identical. Therefore, when investigating the dissimilarity of different processing functions, we only considered the q-vector with two required attributes.

For a category that involves two attributes, there are four latent groups ($\alpha_1^* = 00$, $\alpha_2^* = 10$, $\alpha_3^* = 01$, and $\alpha_4^* = 11$). The corresponding processing functions based on the generating model are denoted by $s_T(x|00)$, $s_T(x|10)$, $s_T(x|01)$, and $s_T(x|11)$, where subscript T represents the *True* model. Similar notations with subscript $A$ are used for approximating models. Note that $x$ could be 1 or 0, representing success or failure. The dissimilarity between a true model and an approximating model is defined as

$$\text{Dissimilarity} = \min_{\phi_A} \sum_{l=1}^{4} \left[ \sum_{x=0}^{1} \log \left( \frac{s_T(x|\alpha_l^*)}{s_A(x|\alpha_l^*)} \right) s_T(x|\alpha_l^*) \right].$$

Five hundred sets of $s_T(00)$, $s_T(10)$, $s_T(01)$, and $s_T(11)$ were generated based on each of the sDINA, sDINO, sA-CDM and sG-DINA models under the high, moderate and low quality conditions given in the article. For example, when the true model was the sDINA model and items were of high quality, $s_T(00) = s_T(10) = s_T(01) \sim U(0.05, 0.15)$ and $s_T(11) \sim U(0.85, 0.95)$. Note that no item responses were generated in this process. When the processing functions of the true model were generated, the following function was minimized

$$\sum_{l=1}^{4} \left[ \sum_{x=0}^{1} \log \left( \frac{s_T(x|\alpha_l^*)}{s_A(x|\alpha_l^*)} \right) s_T(x|\alpha_l^*) \right]$$

based on an approximating model A (which could be sDINA, sDINO, sA-CDM or sG-DINA) with respect to the parameters in model A. The minimization was performed using the Nelder and Mead's method in R, which is believed to be robust, though relatively slow.

## Appendix II: Scatter plots of PCV and SRMSR

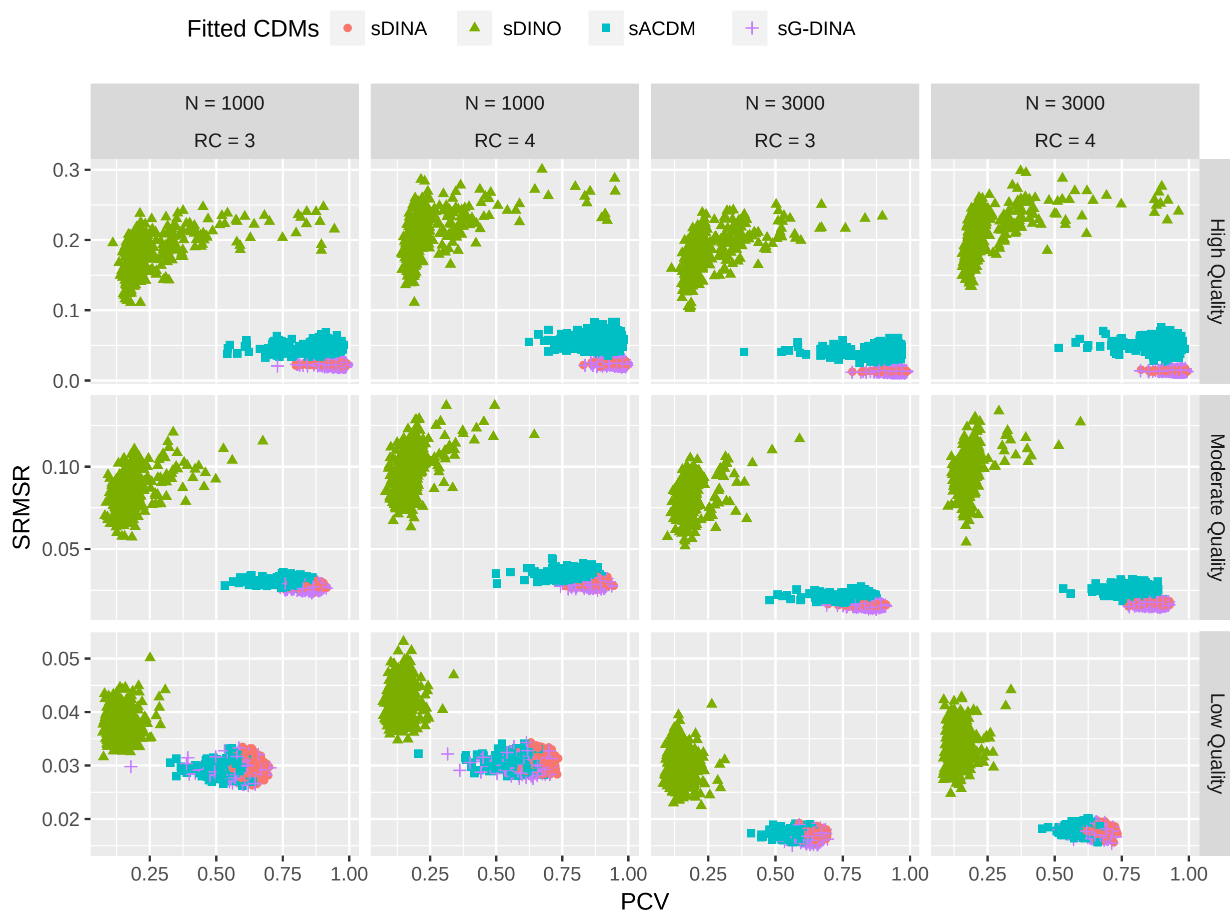


*Figure 1*. Relation between SRMSR and PCV for sDINA model generated data

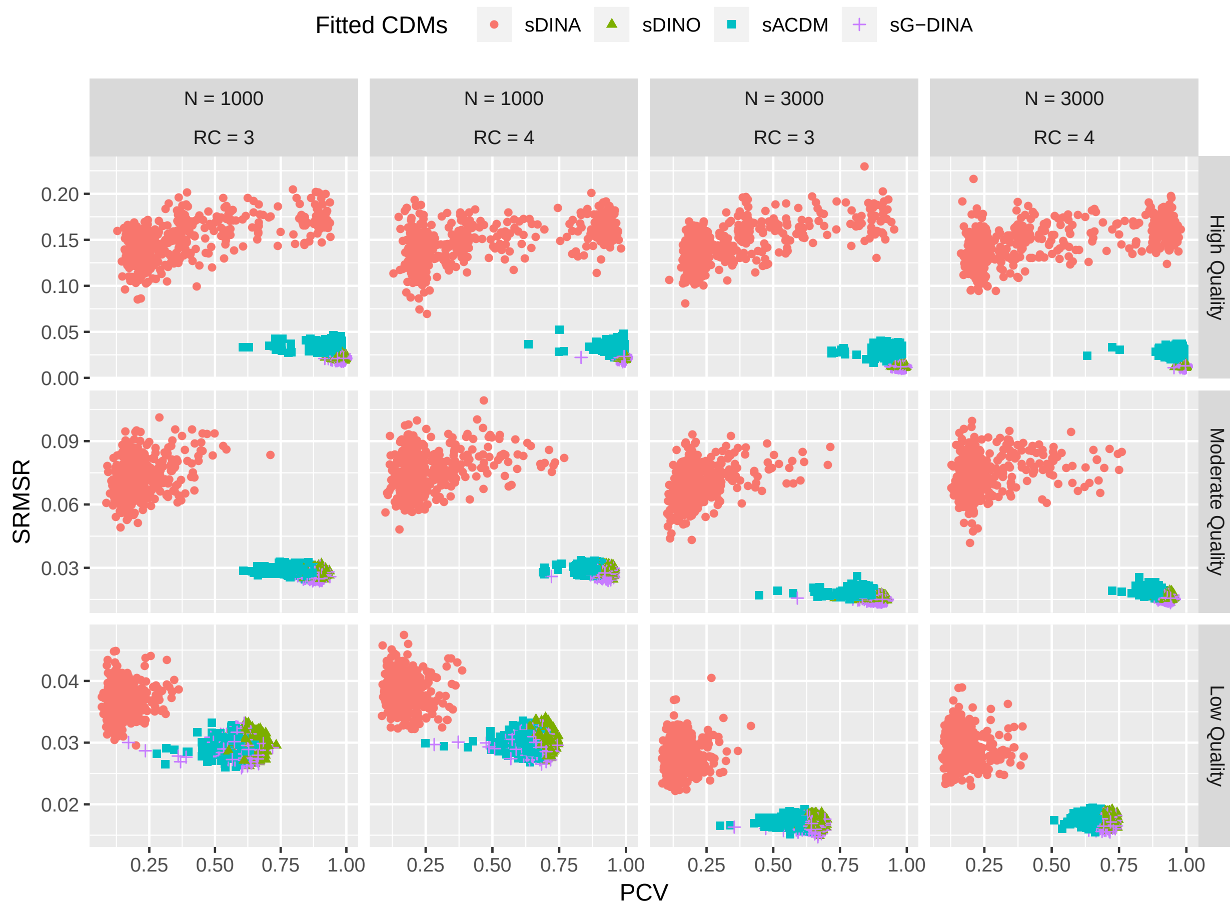


*Figure 2*. Relation between SRMSR and PCV for sDINO model generated data

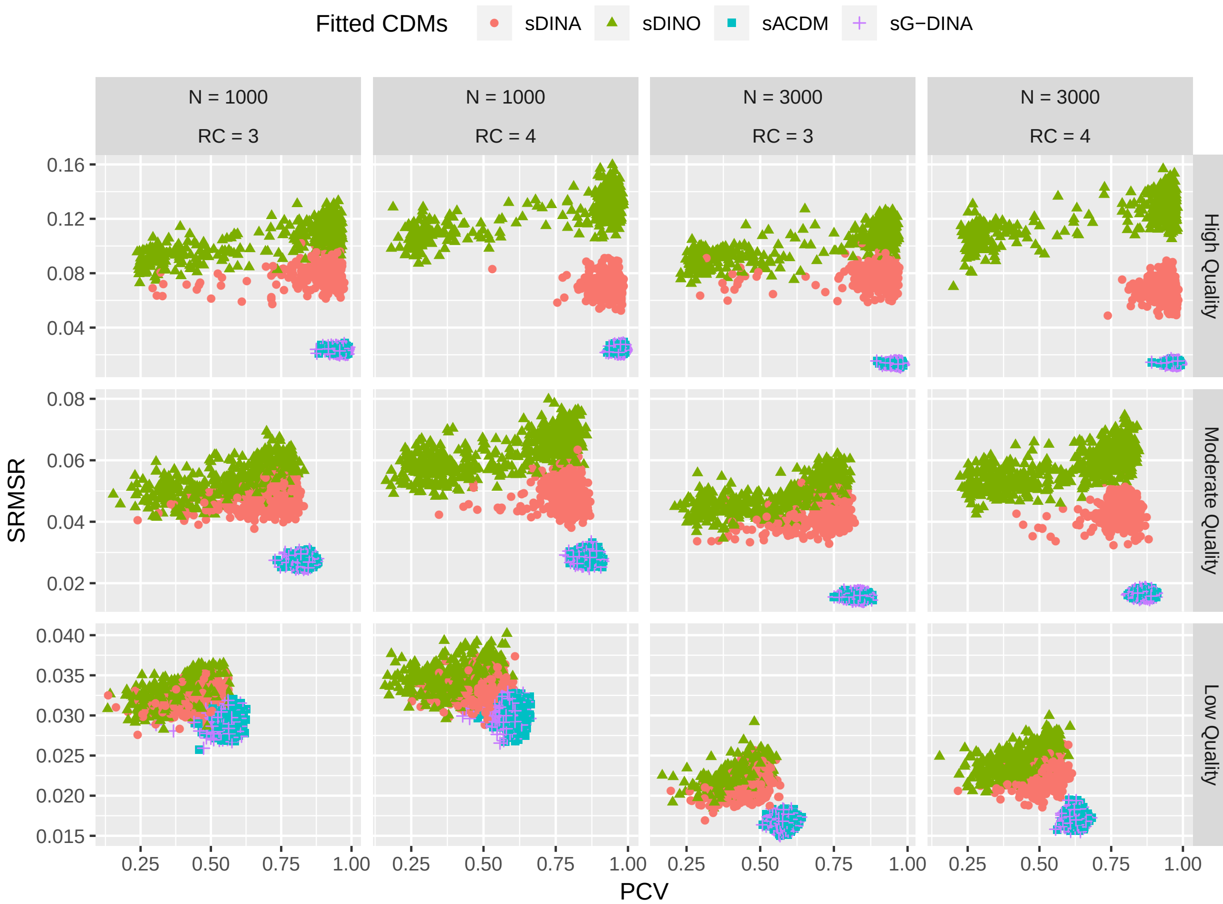


*Figure 3*. Relation between SRMSR and PCV for s*A*-CDM generated data

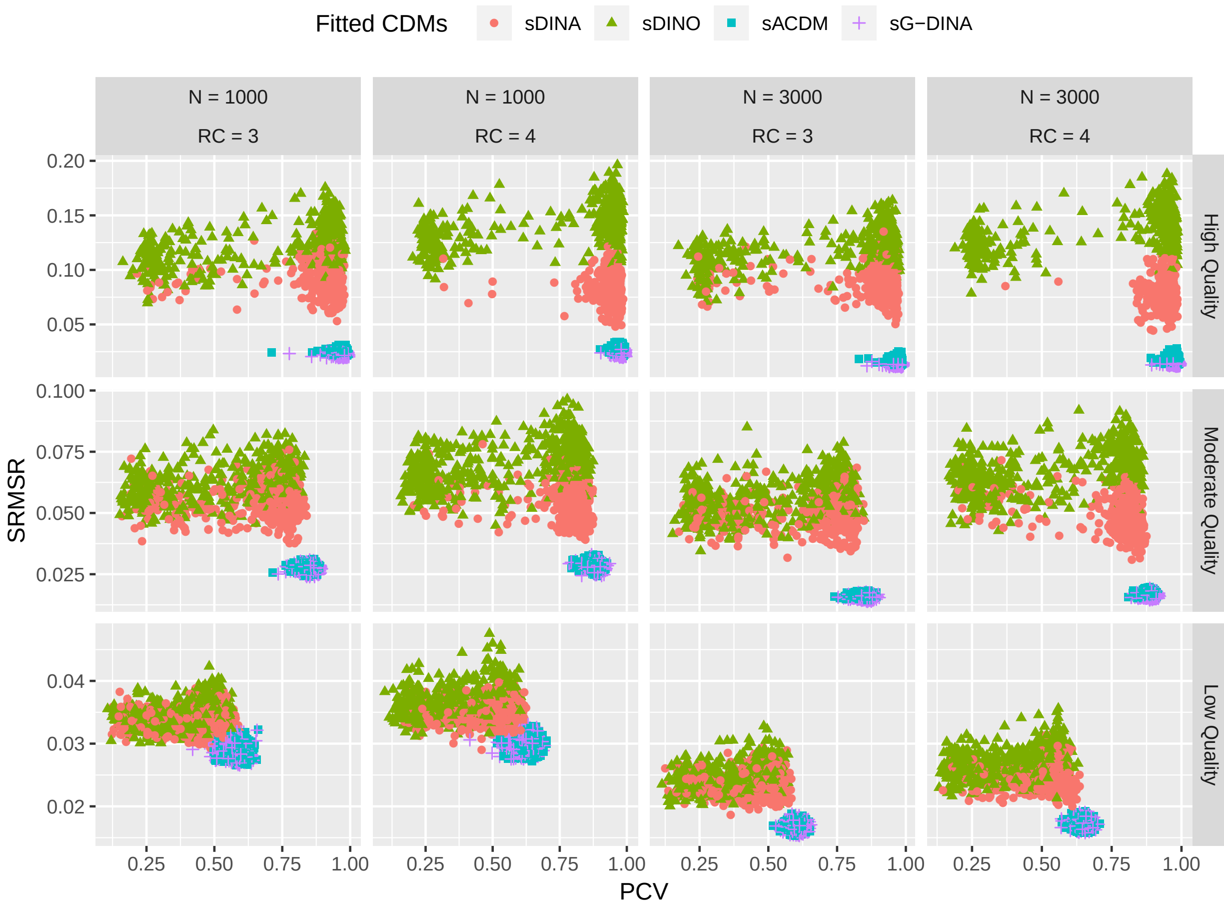


*Figure 4*. Relation between SRMSR and PCV for sG-DINA model generated data

## Appendix III: Real Data Analysis

To illustrate, responses of 1328 students from the United States to 17 items from Block 4 of the Trends in International Mathematics and Science Study (TIMSS) 2007 eigth-grade mathematics assessment were analyzed in this study. Out of 17 items, three are polytomously scored, each with three response categories. The data were previously analyzed by L. Ma (2014) and W. Ma (2017). The Q-matrix originally developed by L. Ma (2014) was later modified to be suitable for the sG-DINA model by W. Ma (2017), which involves seven content attributes, namely, ($\alpha_1$) whole numbers and integers, ($\alpha_2$) fractions, decimals, ratio proportion, and percent, ($\alpha_3$) algebraic expressions and equations/formulas functions, ($\alpha_4$) geometric shapes, ($\alpha_5$) geometric measurement and location movement, ($\alpha_6$) data organization and representation, and ($\alpha_7$) data interpretation and chance. The Q-matrix is given in Table 1, where the suffix "-1" and "-2" were added to item numbers to indicate category 1 and 2, respectively, for polytomously scored items.

The sG-DINA model was fitted to the data. The total number of parameters is 237, including 110 item parameters and 127 population proportion parameters. Given that there are only 17 items, the $M_{\text{ord}}^{\text{All}}$ is not calculable due to the lack of degrees of freedom. The $M_{\text{ord}}^{\text{Item}}$ is equal to 114.89 with 43 degrees of freedom, and the corresponding $p$-value is less than .001, suggesting an inadequate model-data fit. The model-data misfit may be attributed to factors such as the misspecifications in the Q-matrix. W. Ma (2017) suggested some modifications to the Q-matrix based on a stepwise validation procedure. By fitting the sG-DINA model along with this suggested Q-matrix, it can be observed that $M_{\text{ord}}^{\text{Item}} = 89.71, df = 65, p = .023$. This indicates an adequate model-data fit under .01 alpha level. By using the modified Q-matrix, the SRMSR changed from .041 to .026, which provides additional empirical evidence to support the modification of the Q-matrix.

Table 1

*Q-matrix for the TIMSS 2007 data*

| Item No. | TIMSS Item ID | $\alpha_1$ | $\alpha_2$ | $\alpha_3$ | $\alpha_4$ | $\alpha_5$ | $\alpha_6$ | $\alpha_7$ |
|---|---|---|---|---|---|---|---|---|
| 1 | M042001 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | M042022 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 3 | M042082 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 4 | M042088 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| 5 | M042304A | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6-1 | M042304B-1 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 6-2 | M042304B-2 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 7 | M042304C | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| 8-1 | M042304D-1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 8-2 | M042304D-2 | 1 | 0 | 0 | 0 | 0 | 0 | 0 |
| 9 | M042267 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 10 | M042239 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| 11 | M042238 | 1 | 0 | 1 | 0 | 1 | 0 | 0 |
| 12 | M042279 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| 13 | M042036 | 1 | 0 | 0 | 1 | 1 | 0 | 0 |
| 14 | M042130 | 1 | 0 | 0 | 1 | 1 | 1 | 0 |
| 15 | M042303A | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 16-1 | M042303B-1 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 16-2 | M042303B-2 | 1 | 0 | 0 | 0 | 1 | 1 | 1 |
| 17 | M042222 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |

## Appendix IV: Calculation of $M_{\text{ord}}$ statistics for the sequential G-DINA model

To calculate the $M_{\text{ord}}$ statistic, we need to calculate the $w \times w$ matrix $\Gamma_\kappa$, which is the asymptotic covariance of $\sqrt{N}(\boldsymbol{m} - \boldsymbol{\kappa})$, where $\boldsymbol{\kappa} = \boldsymbol{L}\boldsymbol{\pi}$. $\boldsymbol{\Gamma}_\kappa$ can be partitioned into four blocks:

$$\boldsymbol{\Gamma}_\kappa = \left[\begin{array}{c|c} \boldsymbol{\Gamma}_\kappa^{11} & \boldsymbol{\Gamma}_\kappa^{12} \\ \hline \boldsymbol{\Gamma}_\kappa^{21} & \boldsymbol{\Gamma}_\kappa^{22} \end{array}\right], \tag{1}$$

where, with Acov() denoting an asymptotic covariance matrix, $\boldsymbol{\Gamma}_\kappa^{11} = \sqrt{N}\text{Acov}(\boldsymbol{m_1})$, $\boldsymbol{\Gamma}_\kappa^{12} = \sqrt{N}\text{Acov}(\boldsymbol{m_1}, \boldsymbol{m_2})$ and $\boldsymbol{\Gamma}_\kappa^{22} = \sqrt{N}\text{Acov}(\boldsymbol{m_2})$. In addition, $\boldsymbol{\Gamma}_\kappa^{11}$, $\boldsymbol{\Gamma}_\kappa^{21}$ and $\boldsymbol{\Gamma}_\kappa^{22}$ have elements

$$\sqrt{N}\text{Acov}(m_a, m_b) = E[X_a X_b] - E[X_a]E[X_b], \tag{2}$$

$$\sqrt{N}\text{Acov}(m_{a,b}, m_r) = E[X_a X_b X_r] - E[X_a X_b]E[X_r], a < b, \text{and} \tag{3}$$

$$\sqrt{N}\text{Acov}(m_{a,b}, m_{r,t}) = E[X_a X_b X_r X_t] - E[X_a X_b]E[X_r X_t], a < b, r < t, \tag{4}$$

respectively. Under the sG-DINA model, we have

$$E[X_a|\boldsymbol{\alpha}_c] = \sum_{x_a=0}^{H_a} x_a P(x_a|\boldsymbol{\alpha}_c) = \sum_{x_a=1}^{H_a} \prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c). \tag{5}$$

As a result,

$$\begin{aligned} E[X_a] &= \sum_{c}^{2^K} p(\boldsymbol{\alpha}_c) E\left[X_a|\boldsymbol{\alpha}_c\right] && (6)\\ &= \sum_{c}^{2^K} p(\boldsymbol{\alpha}_c) \left[\sum_{x_a=1}^{H_a} \prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c)\right], && (7) \end{aligned}$$

and

$$E[X_a X_b] = \begin{cases} \sum_{c}^{2^K} p(\boldsymbol{\alpha}_c) \left[\sum_{x_a=1}^{H_a} (2x_a - 1) \prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c)\right] & \text{if } a = b, \\ \sum_{c}^{2^K} p(\boldsymbol{\alpha}_c) \left[\sum_{x_a=1}^{H_a} \prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c)\right] \left[\sum_{x_b=1}^{H_b} \prod_{h=1}^{x_b} s_{bh}(\boldsymbol{\alpha}_c)\right] & \text{otherwise.} \end{cases} \tag{8}$$

$E[X_a X_b X_r]$ and $E[X_a X_b X_r X_t]$ can be calculated similarly.

In addition to $\Gamma_\kappa$, we also need to calculate the $w \times v$ Jacobian matrix $\boldsymbol{\Delta}_\kappa = \partial\hat{\boldsymbol{\kappa}}/\partial\boldsymbol{\gamma}$. Recall that for $M_{\text{ord}}^{\text{All}}$ statistic, $\boldsymbol{\gamma} = (\boldsymbol{\phi}^T, \boldsymbol{\rho}^T)^T$; and thus, $\boldsymbol{\Delta}_\kappa$ can be partitioned as in

$$\boldsymbol{\Delta}_\kappa = \left[\begin{array}{c|c} \boldsymbol{\Delta}_\kappa^{11} & \boldsymbol{\Delta}_\kappa^{12} \\ \hline \boldsymbol{\Delta}_\kappa^{21} & \boldsymbol{\Delta}_\kappa^{22} \end{array}\right] = \left[\begin{array}{c|c} \frac{\partial\hat{\boldsymbol{\kappa}}_{\mathbf{1}}}{\partial\boldsymbol{\phi}} & \frac{\partial\hat{\boldsymbol{\kappa}}_{\mathbf{1}}}{\partial\boldsymbol{\rho}} \\ \hline \frac{\partial\hat{\boldsymbol{\kappa}}_{\mathbf{2}}}{\partial\boldsymbol{\phi}} & \frac{\partial\hat{\boldsymbol{\kappa}}_{\mathbf{2}}}{\partial\boldsymbol{\rho}} \end{array}\right].$$

Under the sG-DINA model, submatrix $\boldsymbol{\Delta}_\kappa^{11}$ has element

$$\frac{\partial\kappa_a}{\partial\phi_{jh'm}} = \frac{\partial E[X_a]}{\partial\phi_{jh'm}} \tag{9}$$

$$= \begin{cases} 0 & \text{if } j \neq a, \\ \sum_c^{2^K} p(\boldsymbol{\alpha}_c)\left[\frac{E[X_a|\boldsymbol{\alpha}_c] - I(h' \neq 1)\sum_{x_a=1}^{h'-1}\prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c)}{s_{ah'}(\boldsymbol{\alpha}_c)}\right]\frac{\partial s_{ah'}(\boldsymbol{\alpha}_c)}{\partial\phi_{jh'm}} & \text{otherwise,} \end{cases} \tag{10}$$

and submatrix $\boldsymbol{\Delta}_\kappa^{21}$ has element

$$\frac{\partial\kappa_{a,b}}{\partial\phi_{jh'm}} = \frac{\partial E[X_aX_b]}{\partial\phi_{jh'm}} \tag{11}$$

$$= \begin{cases} 0 & \text{if } j \neq a \;\&\; j \neq b, \\ \sum_c^{2^K} p(\boldsymbol{\alpha}_c)E[X_b|\boldsymbol{\alpha}_c]\left[\frac{E[X_a|\boldsymbol{\alpha}_c] - I(h' \neq 1)\sum_{x_a=1}^{h'-1}\prod_{h=1}^{x_a} s_{ah}(\boldsymbol{\alpha}_c)}{s_{ah'}(\boldsymbol{\alpha}_c)}\right]\frac{\partial s_{ah}(\boldsymbol{\alpha}_c)}{\partial\phi_{jh'm}} & \text{if } j = a, \\ \sum_c^{2^K} p(\boldsymbol{\alpha}_c)E[X_a|\boldsymbol{\alpha}_c]\left[\frac{E[X_b|\boldsymbol{\alpha}_c] - I(h' \neq 1)\sum_{x_b=1}^{h'-1}\prod_{h=1}^{x_b} s_{bh}(\boldsymbol{\alpha}_c)}{s_{bh'}(\boldsymbol{\alpha}_c)}\right]\frac{\partial s_{bh}(\boldsymbol{\alpha}_c)}{\partial\phi_{jh'm}} & \text{if } j = b, \end{cases} \tag{12}$$

In addition, for $M_{\text{ord}}^{\text{All}}$ statistic, the submatrices $\boldsymbol{\Delta}_\kappa^{12}$ and $\boldsymbol{\Delta}_\kappa^{22}$ have elements

$$\frac{\partial\kappa_a}{\partial p(\boldsymbol{\alpha}_l)} = \frac{\partial E[X_a]}{\partial p(\boldsymbol{\alpha}_l)} = E[X_a|\boldsymbol{\alpha}_l] - E[X_a|\boldsymbol{\alpha}_{2^K}], \text{and} \tag{13}$$

$$\frac{\partial\kappa_{a,b}}{\partial p(\boldsymbol{\alpha}_l)} = \frac{\partial E[X_aX_b]}{\partial p(\boldsymbol{\alpha}_l)} = E[X_a|\boldsymbol{\alpha}_l]E[X_b|\boldsymbol{\alpha}_l] - E[X_a|\boldsymbol{\alpha}_{2^K}]E[X_b|\boldsymbol{\alpha}_{2^K}], \tag{14}$$

respectively. For the $M_{\text{ord}}^{\text{Item}}$ statistic, $\boldsymbol{\gamma} = \boldsymbol{\phi}$. Hence, the calculations of $\boldsymbol{\Delta}_\kappa^{12}$ and $\boldsymbol{\Delta}_\kappa^{22}$ are not needed.

References

Ma, L. (2014). *Validation of the item-attribute matrix in TIMSS: Mathematics using multiple regression and the LSDM* (Unpublished doctoral dissertation). University of Denver.

Ma, W. (2017). *A sequential cognitive diagnosis model for graded response: Model development, Q-matrix validation, and model comparison* (Unpublished doctoral dissertation). Rutgers, The State University of New Jersey.