Supplementary material to article ‘Selecting the correct RI-CLPM using Chi square-type tests and AIC-type criteria’

1 Simulation results

In this supplement, we present additional results for RI-CLPM data with 4, 5, and 6 waves, in addition to the 3 waves of data discussed in the main article. The structure of this supplement follows the same subsections as in the main text.

1.1 Selection between CLPM and RI-CLPM

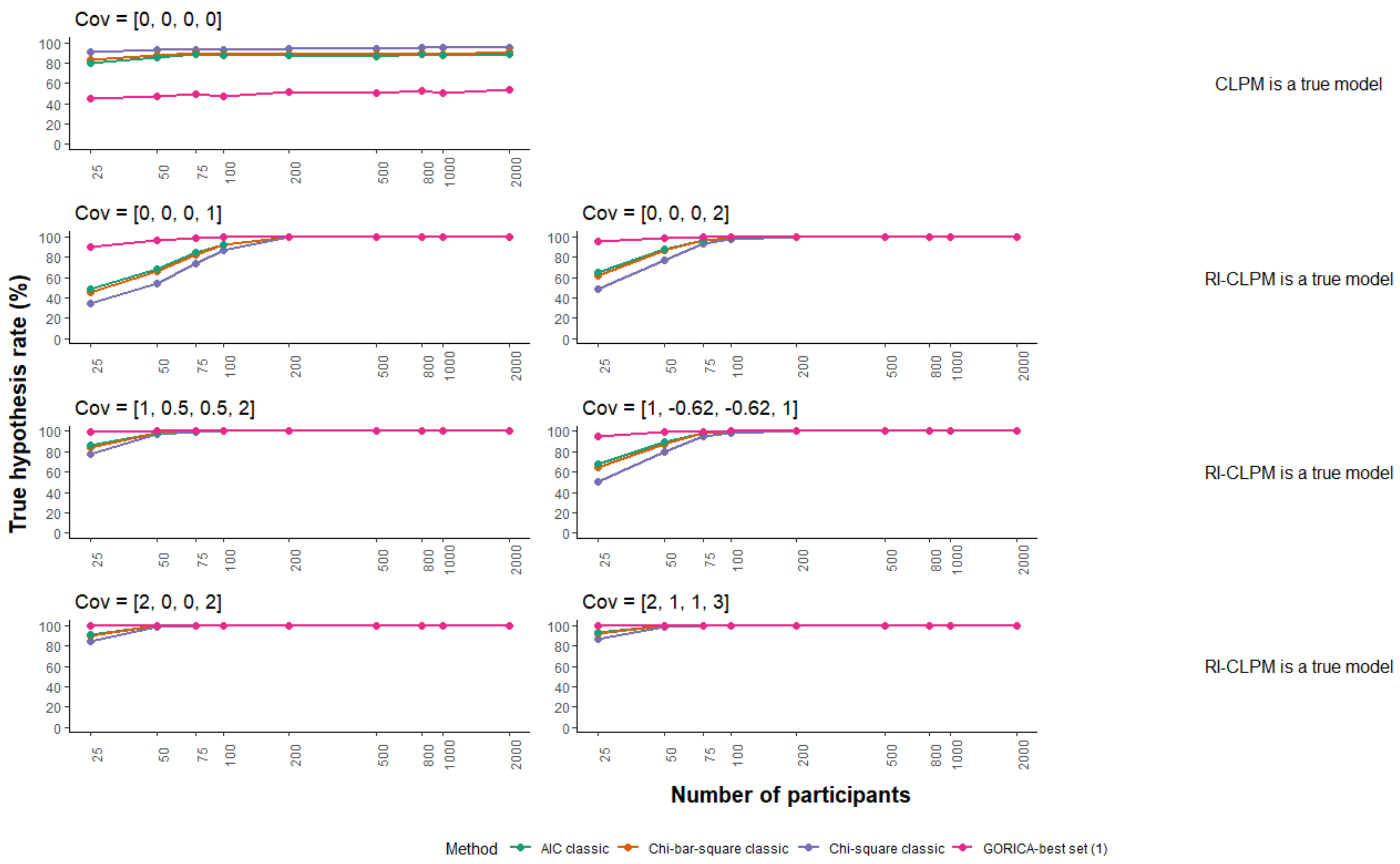

Figure 1: The performance of Chi(bar)square classic, AIC classic, and GORICA-best set (1) to select the true model for 4 waves of data. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

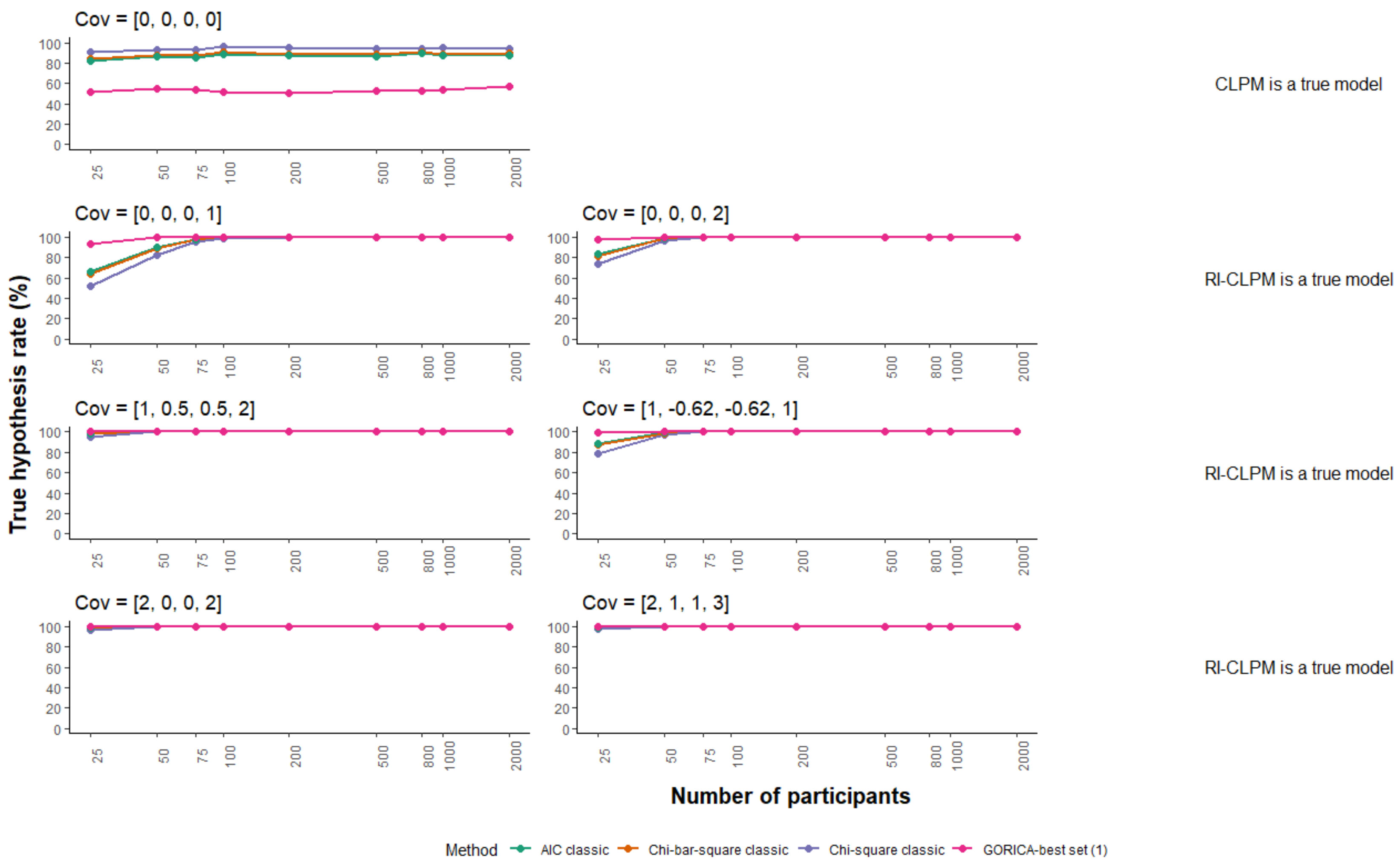

Figure 2: The performance of Chi(bar)square classic, AIC classic, and GORICA-best set (1) to select the true model for 5 waves of data. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

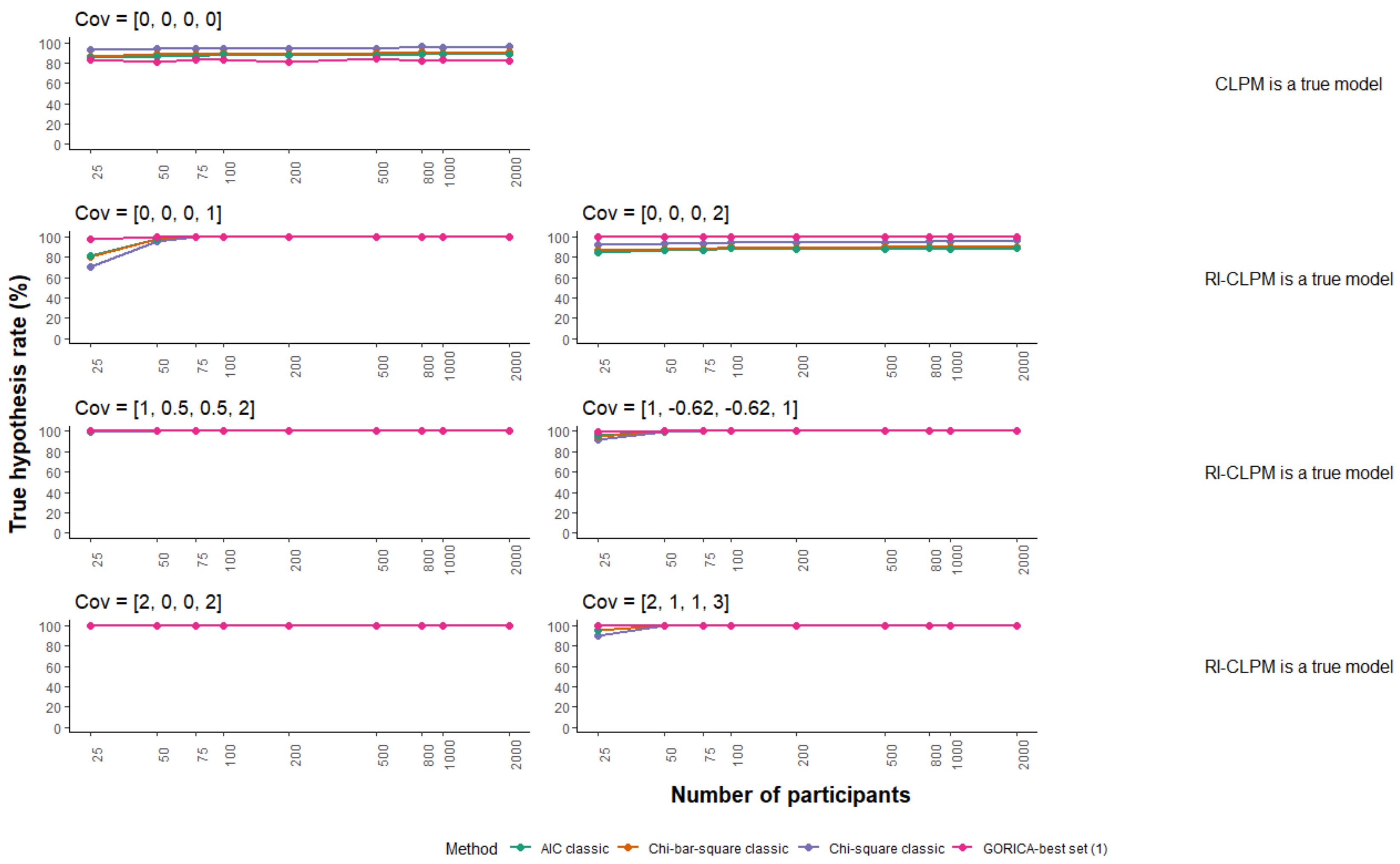

Figure 3: The performance of Chi(bar)square classic, AIC classic, and GORICA-best set (1) to select the true model for 6 waves of data. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

1.2 Selection number of RIs in RI-CLPM

1.2.1 Comparison of Chi-square and Chi-bar-square difference tests

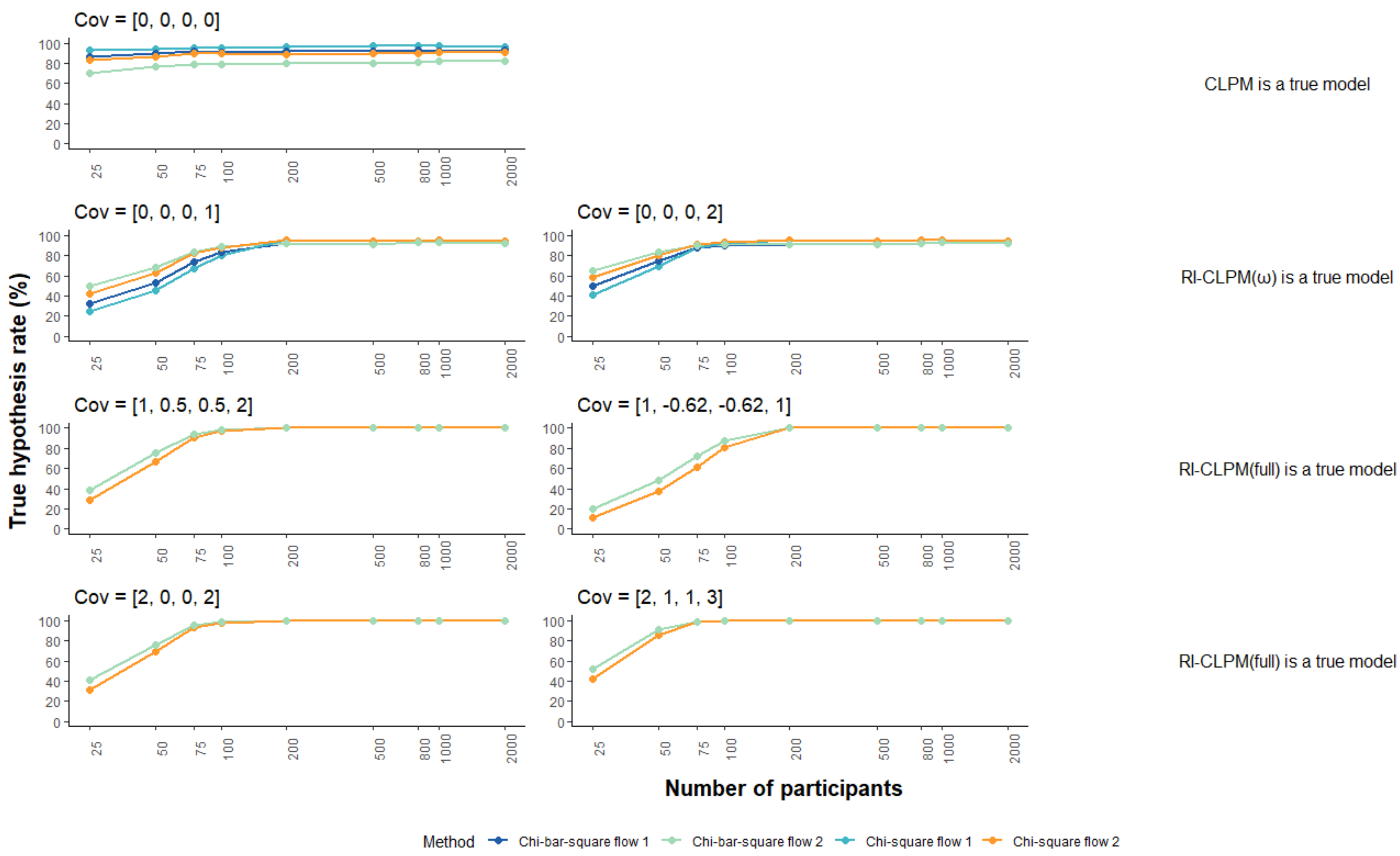

Figure 4: The performance of Chi(bar)square flow 1 and Chi(bar)square flow 2 in selecting the true model for 4 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

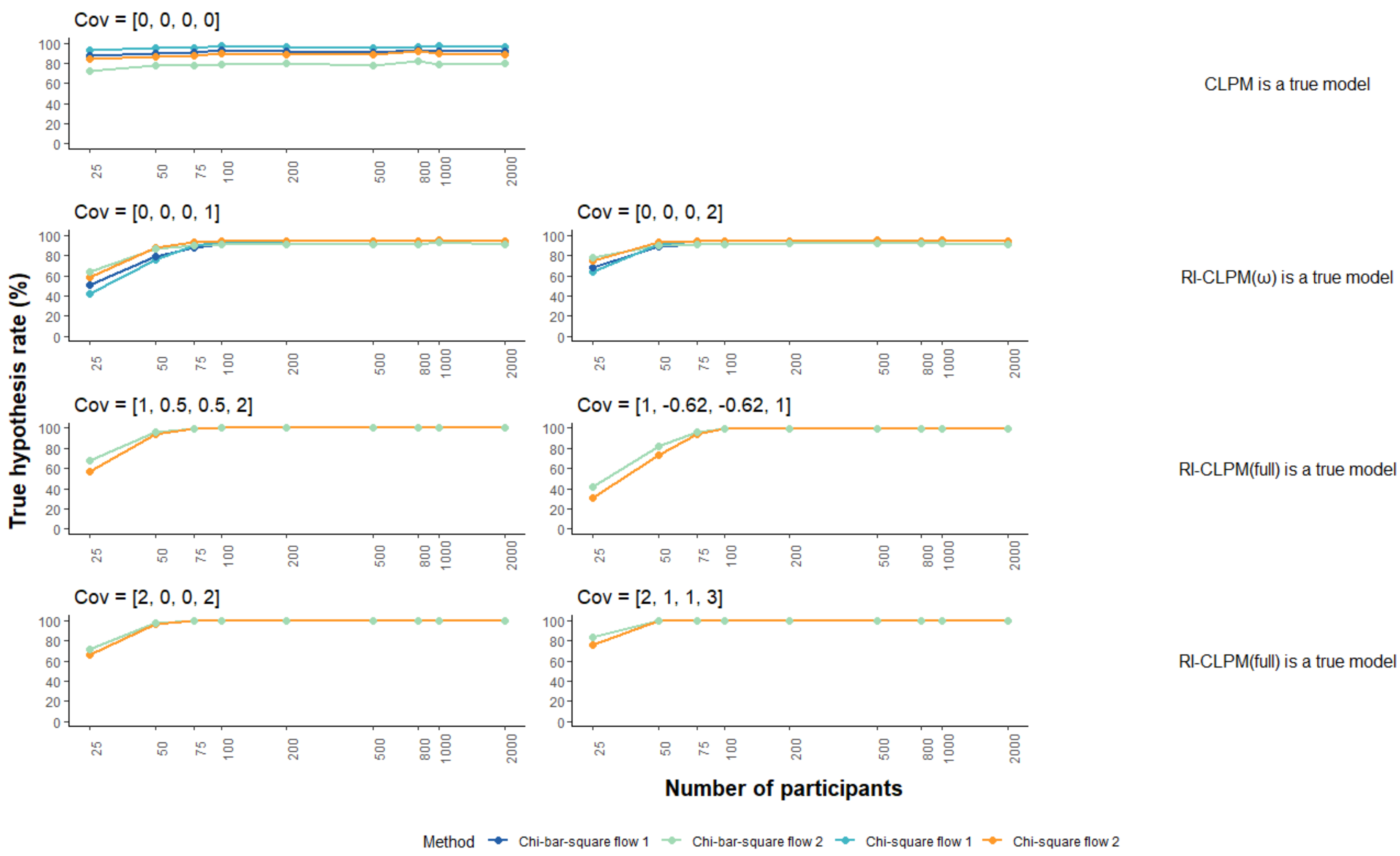

Figure 5: The performance of Chi(bar)square flow 1 and Chi(bar)square flow 2 in selecting the true model for 5 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

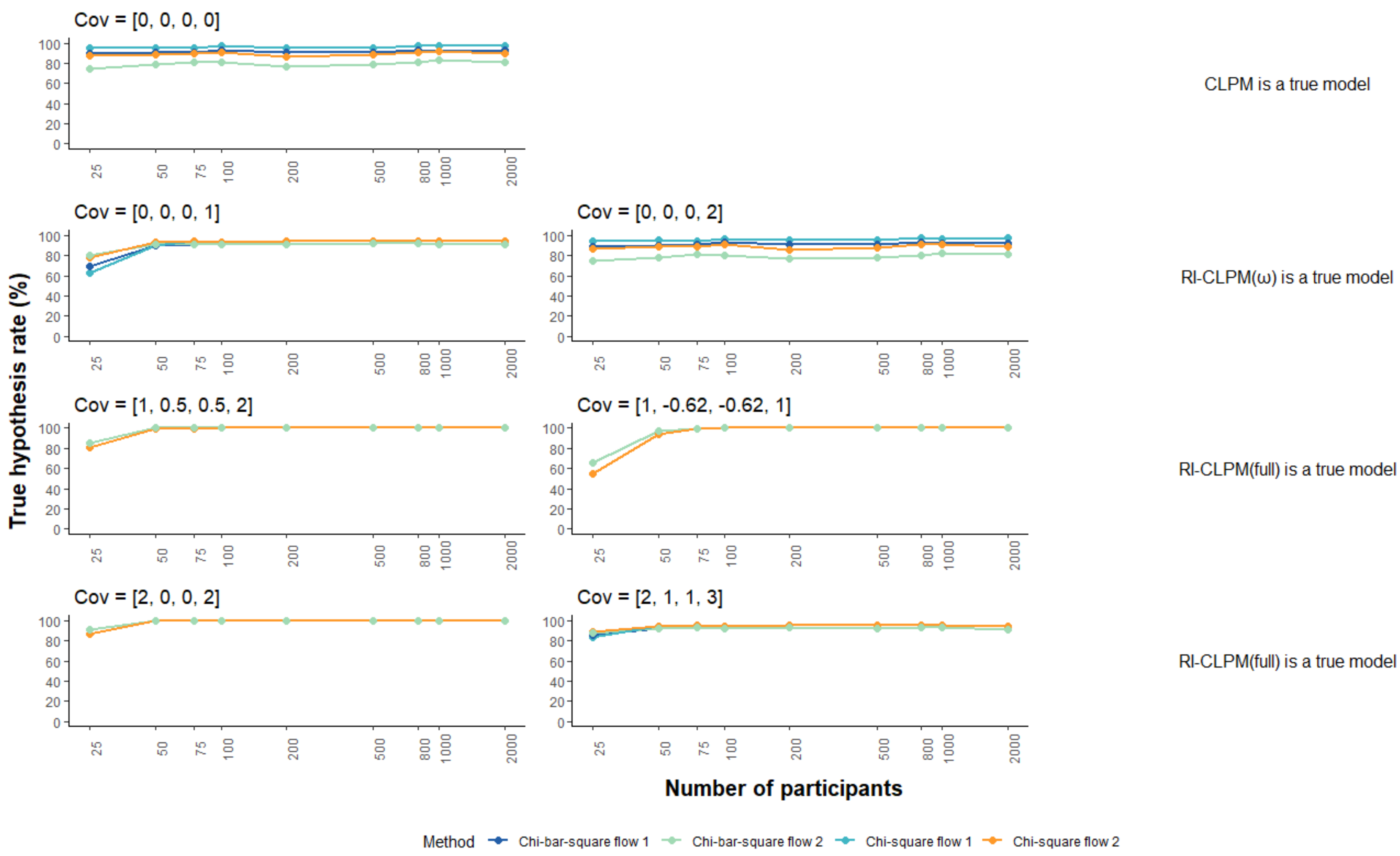

Figure 6: The performance of Chi(bar)square flow 1 and Chi(bar)square flow 2 in selecting the true model for 6 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

1.2.2 Comparison of AIC Methods

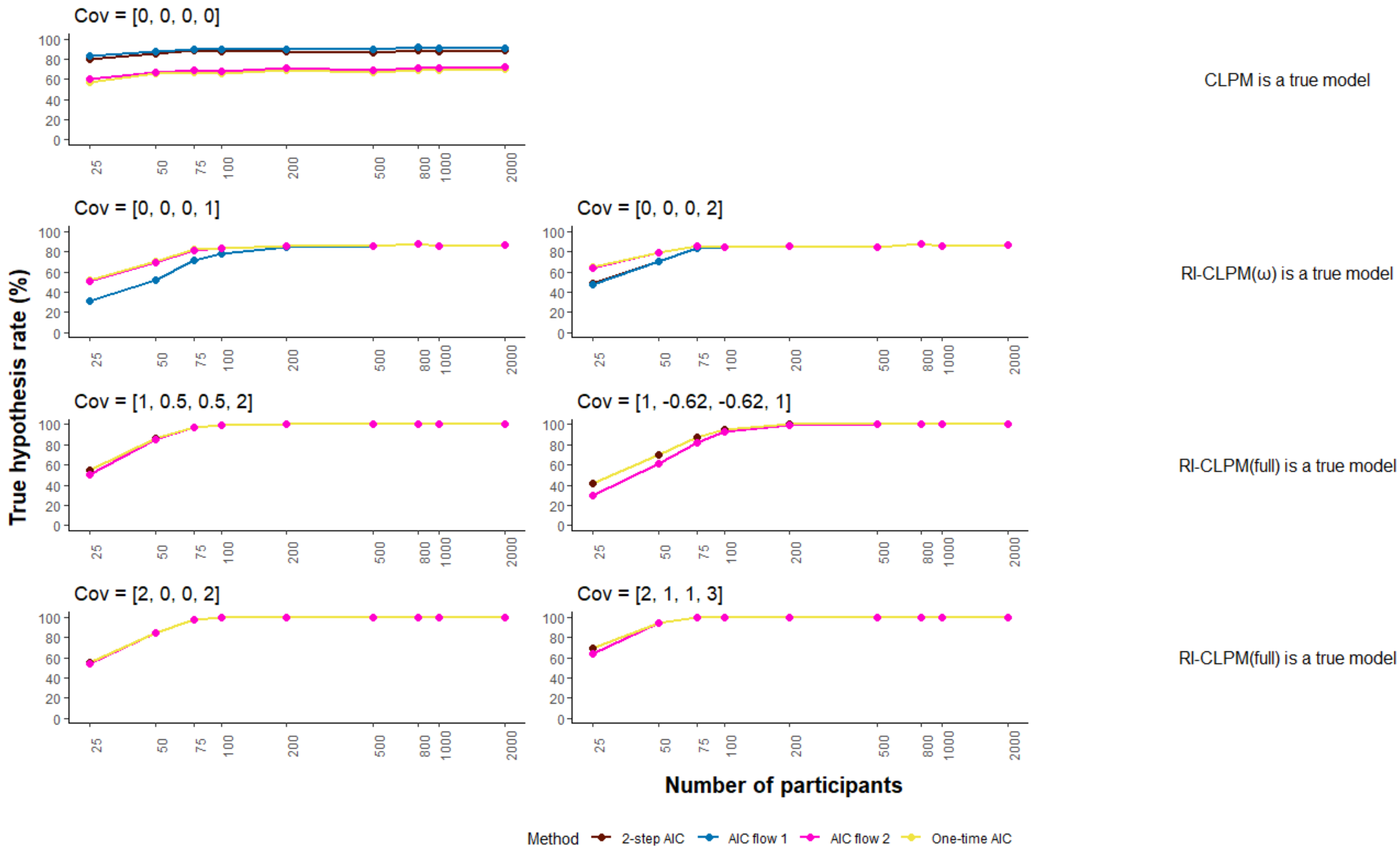

Figure 7: The performance of One-time AIC, 2-step AIC, AIC flow 1, and AIC flow 2 in selecting the true model for 4 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

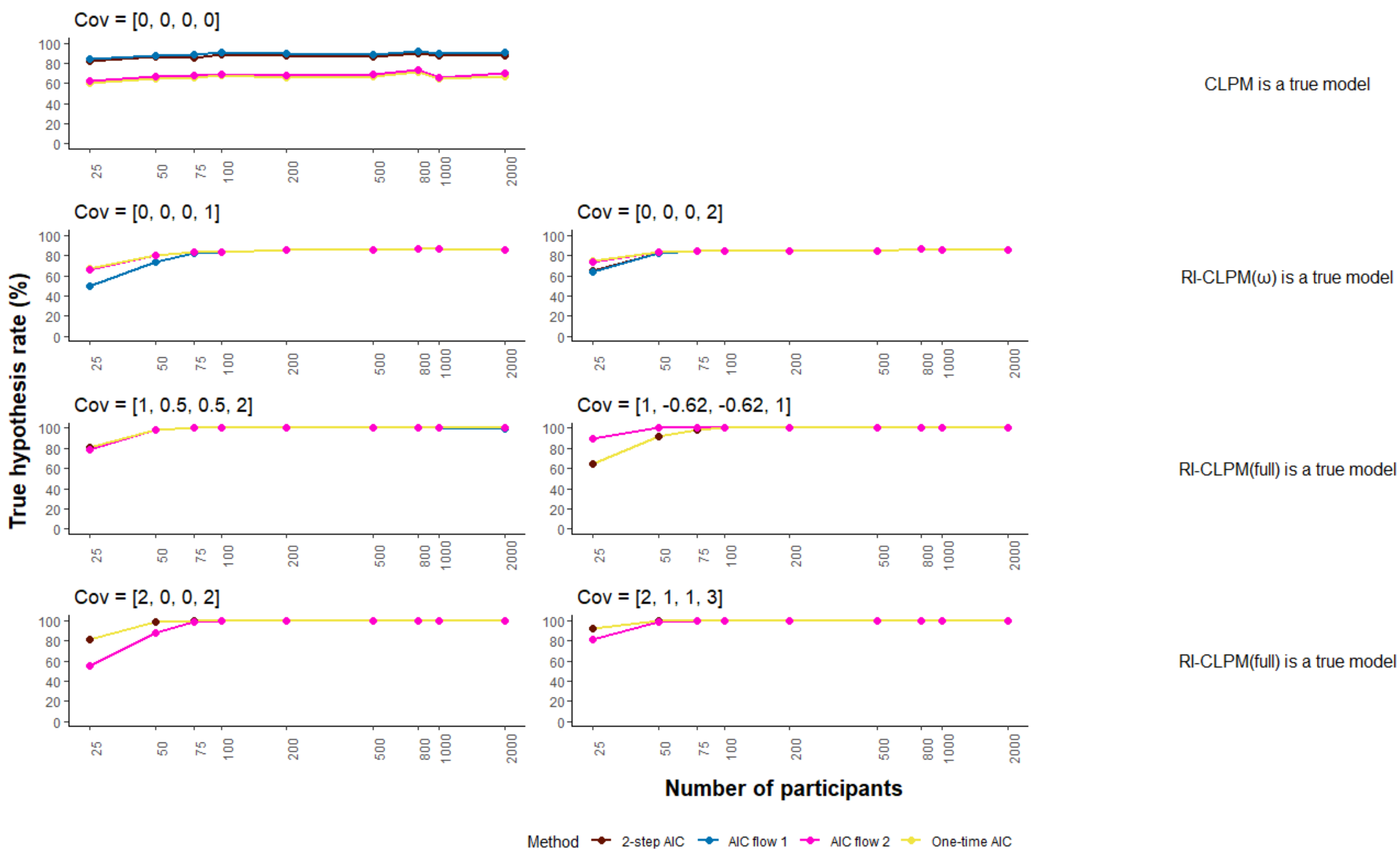

Figure 8: The performance of One-time AIC, 2-step AIC, AIC flow 1, and AIC flow 2 in selecting the true model for 5 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

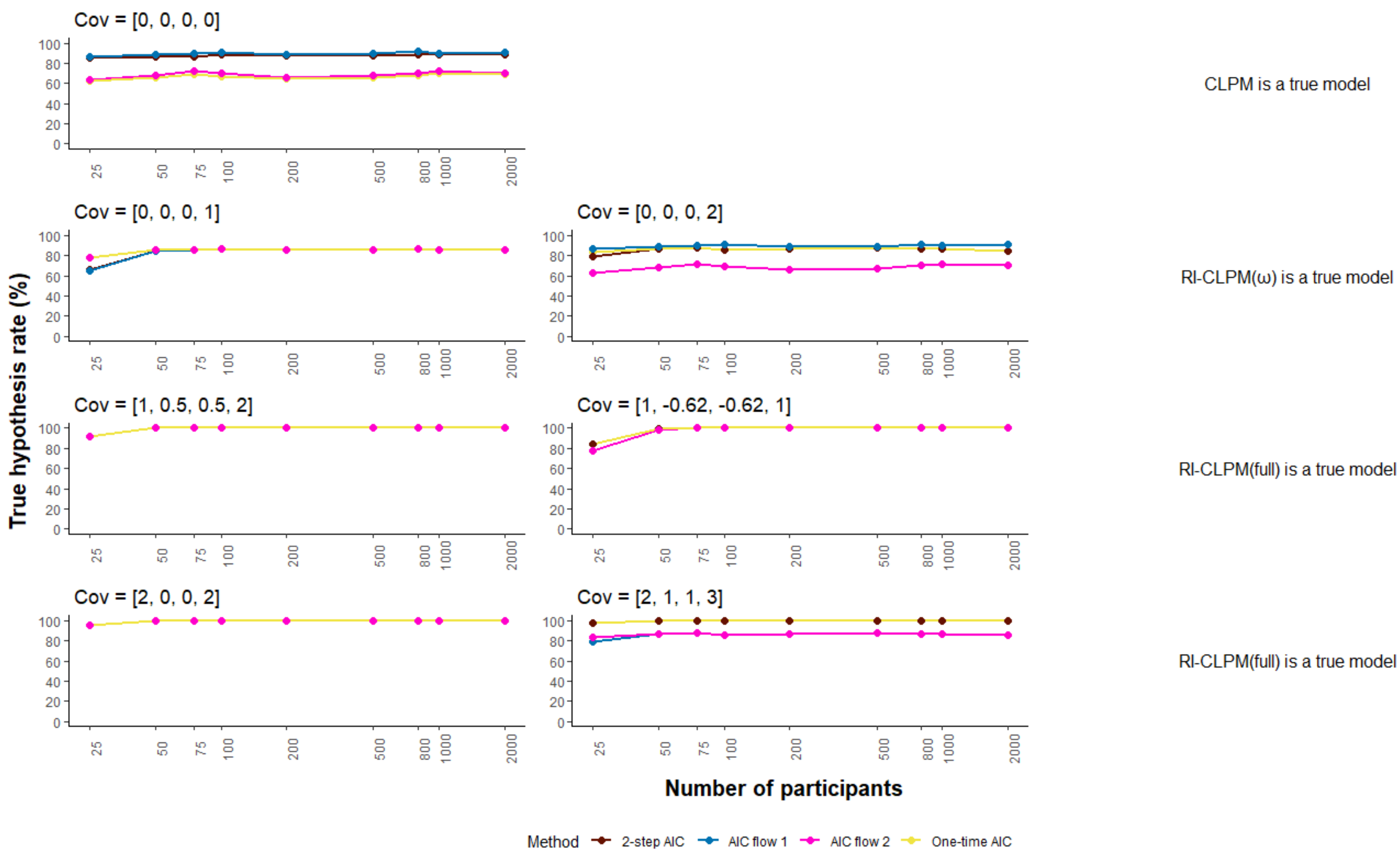

Figure 9: The performance of One-time AIC, 2-step AIC, AIC flow 1, and AIC flow 2 in selecting the true model for 6 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

1.2.3 Comparison of GORICA Methods

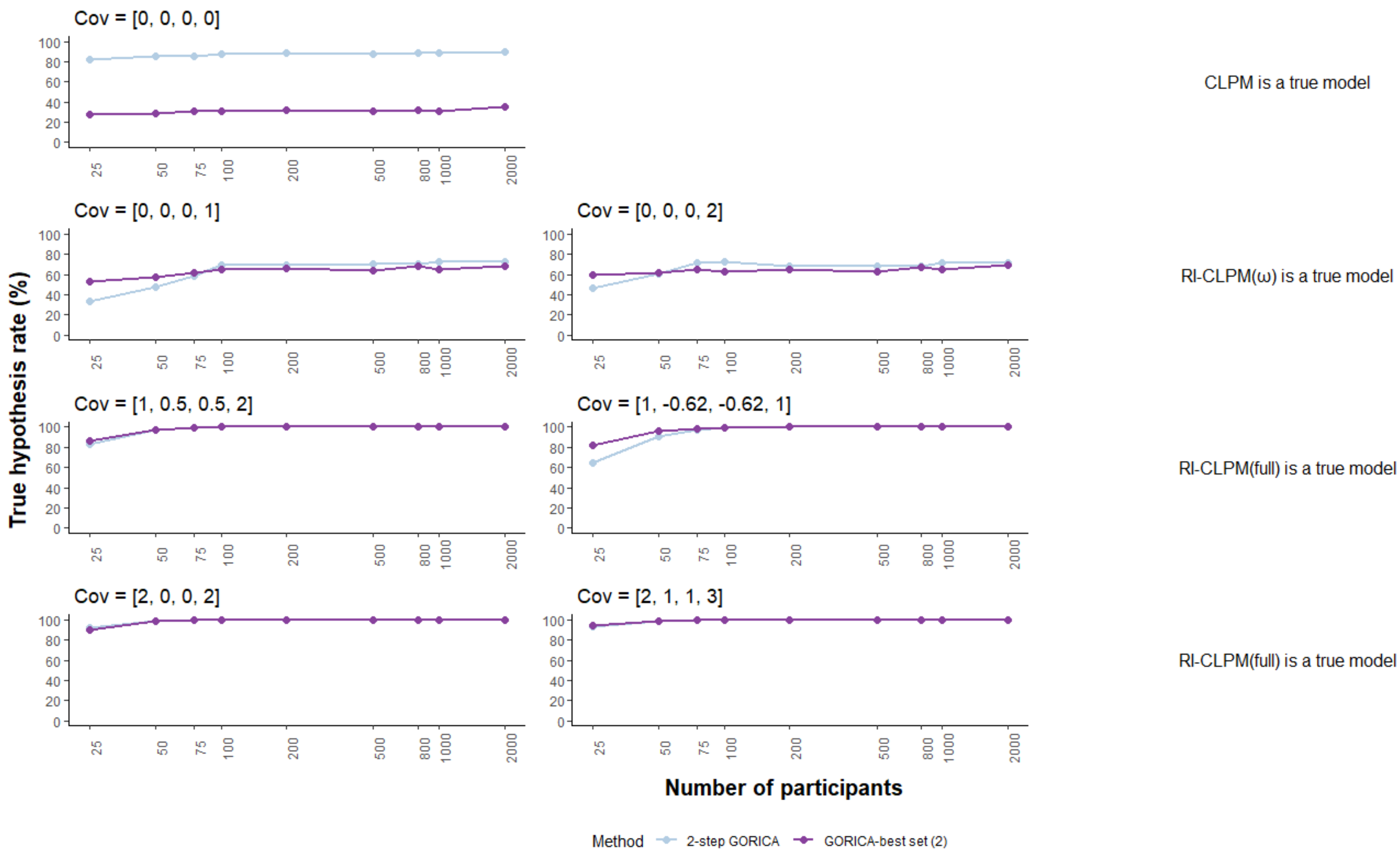

Figure 10: The performance of 2-step GORICA and GORICA-best set (2) in selecting the true model for 4 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

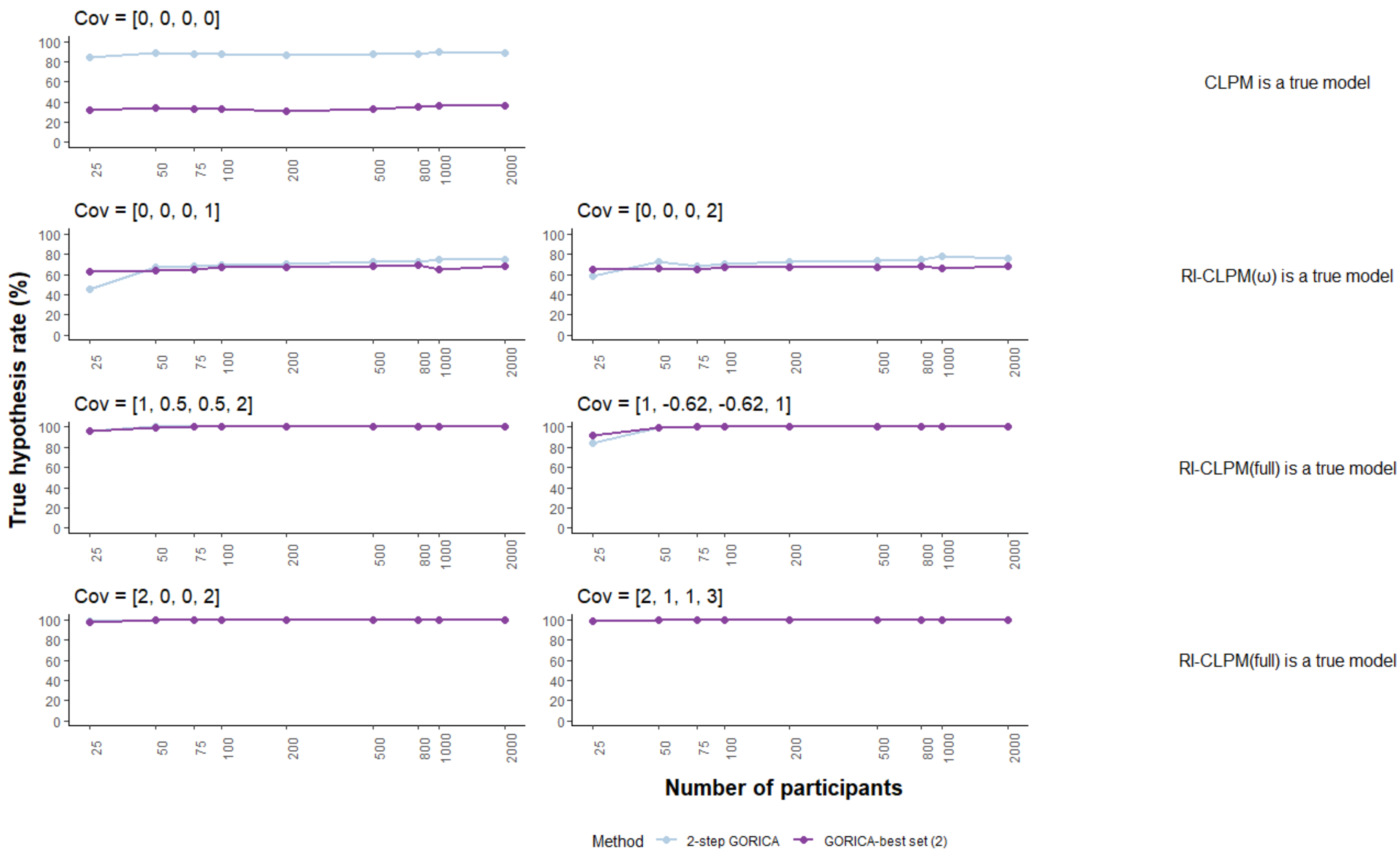

Figure 11: The performance of 2-step GORICA and GORICA-best set (2) in selecting the true model for 5 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

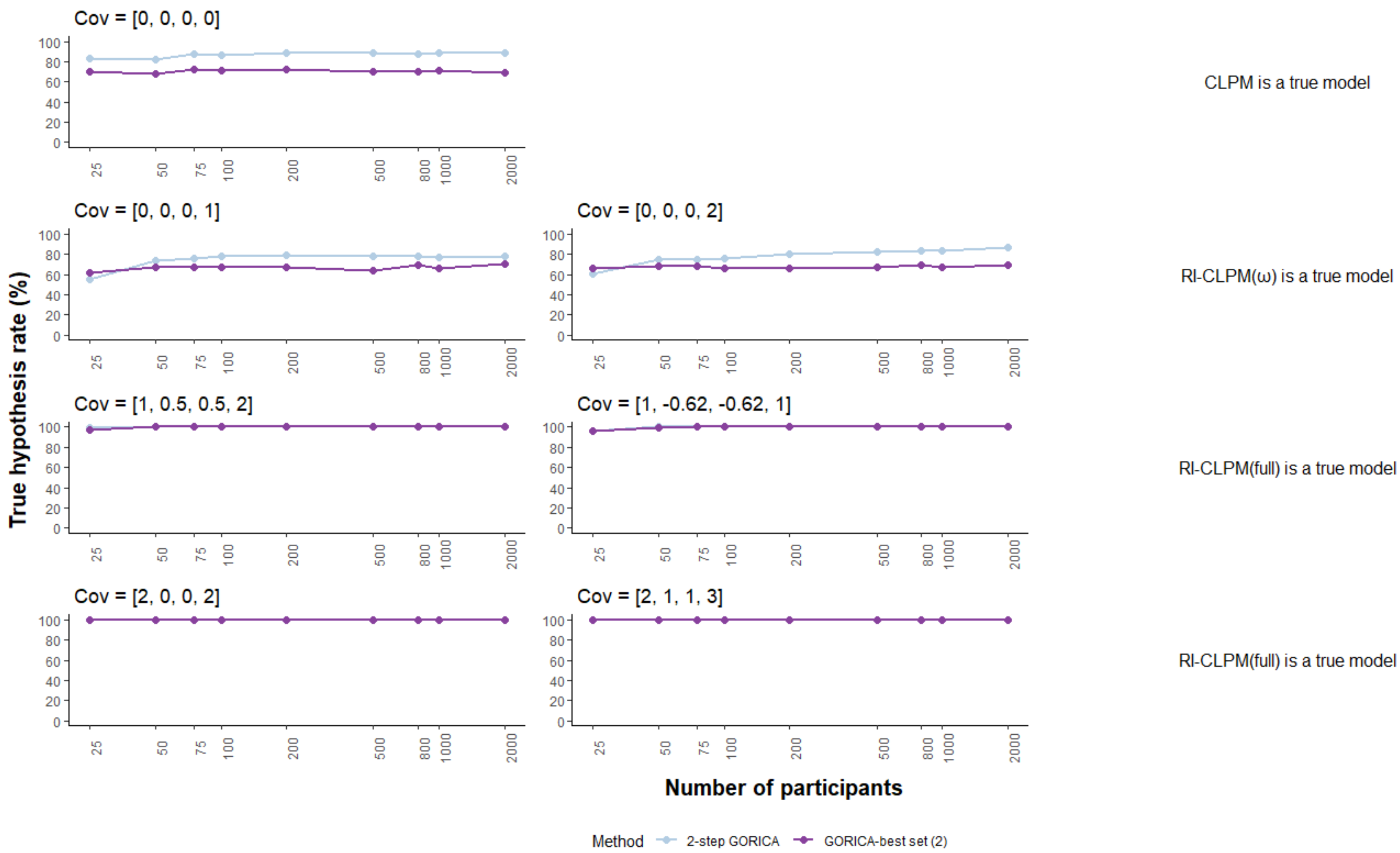

Figure 12: The performance of 2-step GORICA and GORICA-best set (2) in selecting the true model for 6 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

1.2.4 Selecting the best method among the Chi-square and Chi-bar-square difference tests, AIC, and GORICA

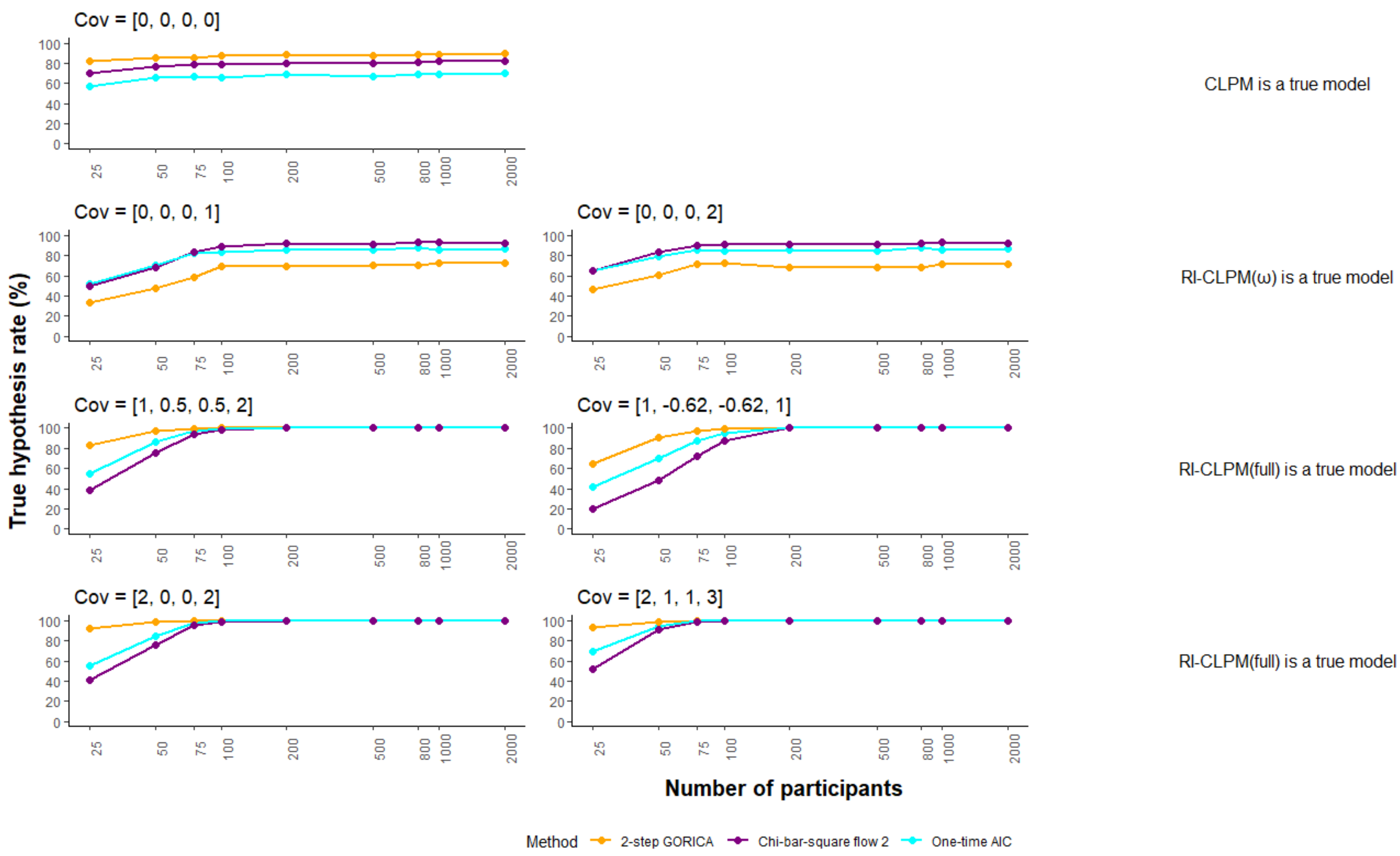

Figure 13: The performance of Chi-bar-square flow 2, 2-step AIC and 2-step GORICA in selecting the true model for 4 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix}0 & 0\\0 & 0\end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}$ or $\begin{bmatrix}0 & 0\\0 & 2\end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

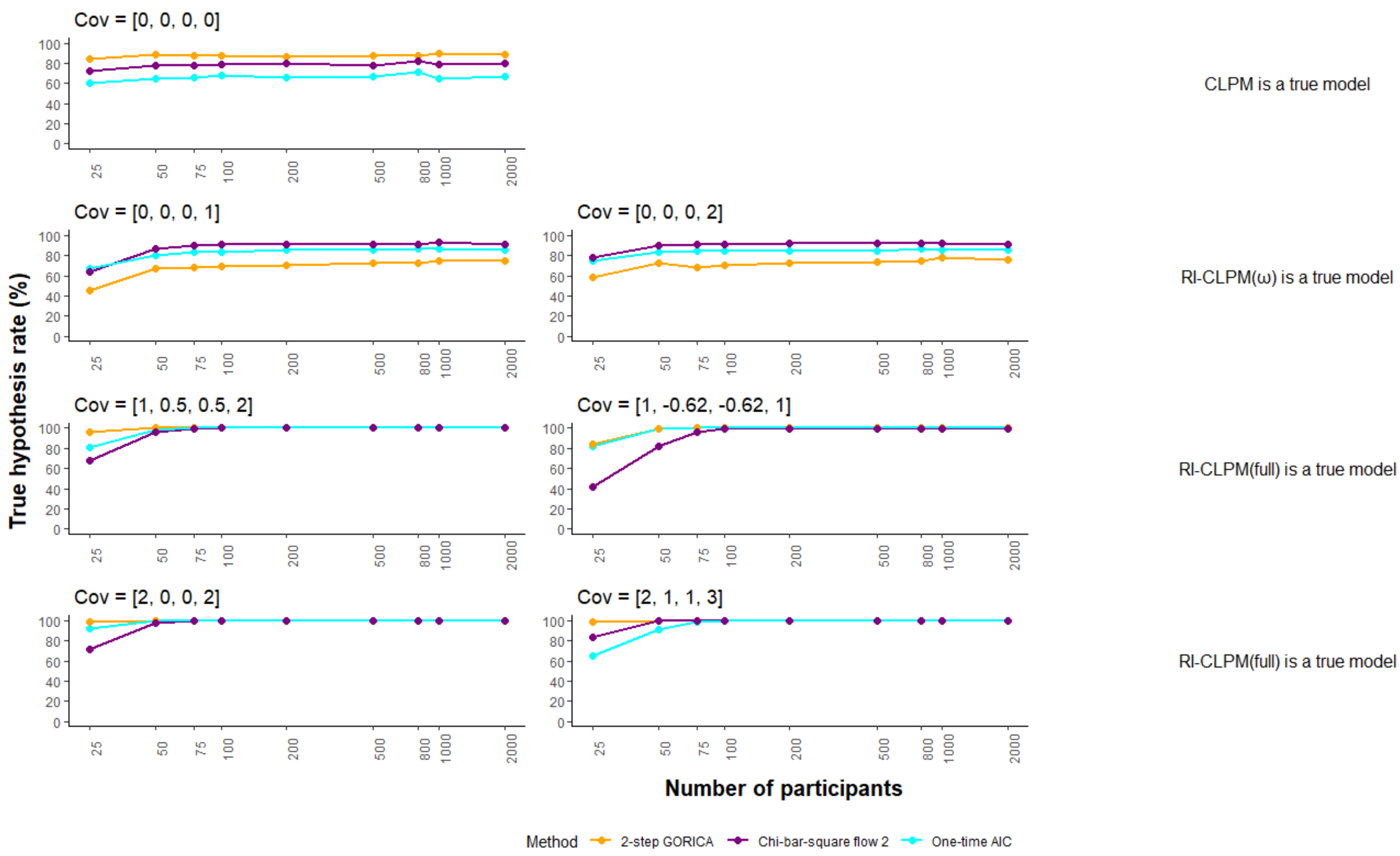

Figure 14: The performance of Chi-bar-square flow 2, 2-step AIC and 2-step GORICA in selecting the true model for 5 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

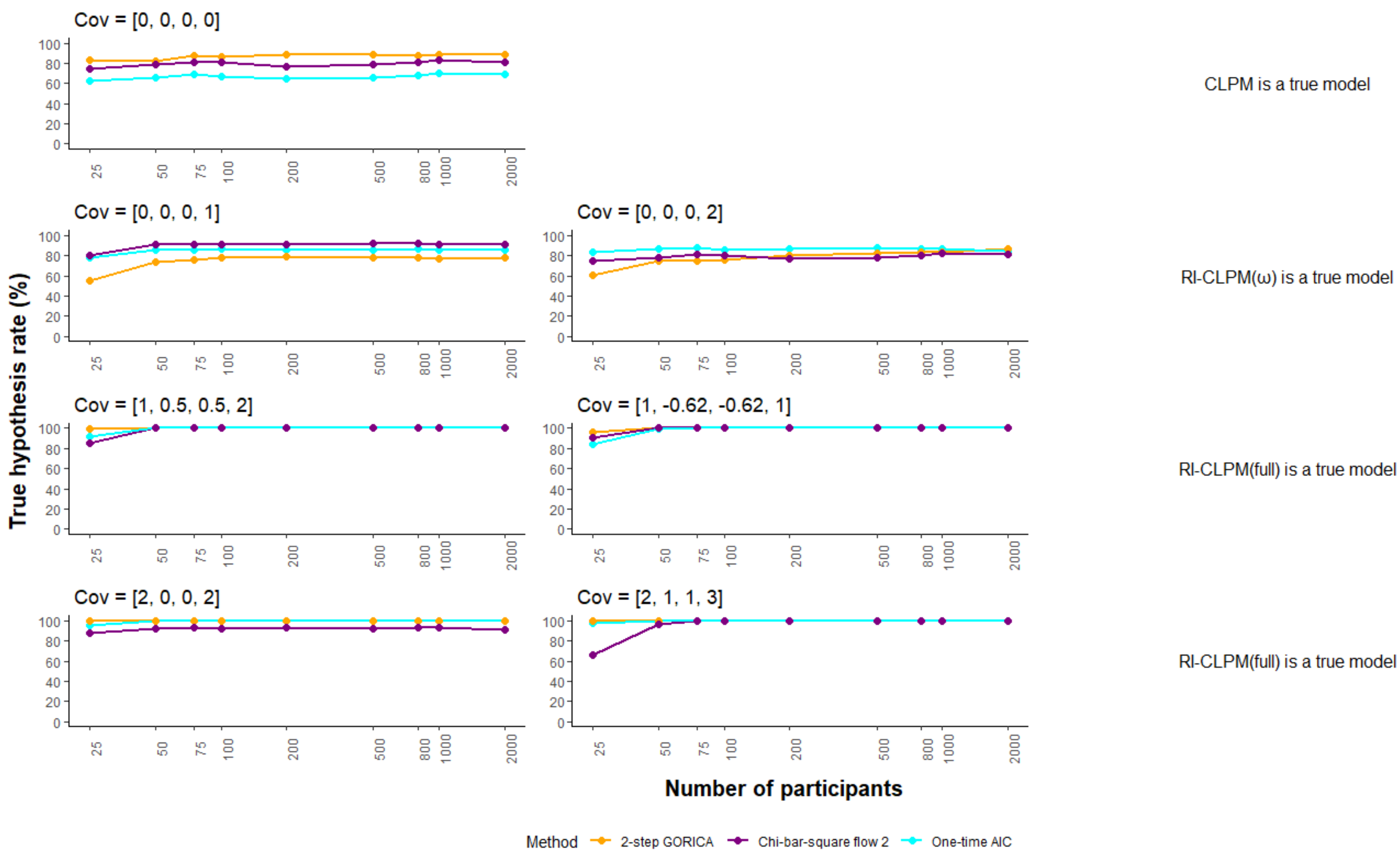

Figure 15: The performance of Chi-bar-square flow 2, 2-step AIC and 2-step GORICA in selecting the true model for 6 waves of data, varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ or $\begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all other instances involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM (full), which includes two random intercepts.

2 Simulation results for GORICA when using only the mean as boundaries and when using both the mean and percentiles as boundaries

As discussed in the article titled "Selecting the correct RI-CLPM using Chi square-type tests and AIC-type criteria" when employing GORICA to select the true model, it becomes necessary to establish boundaries. This is because the simulated data does not result in zero variances for random intercepts in the CLPM. In our exploration, we considered alternative boundaries, and the hypotheses are presented as follows:

$$\begin{aligned}
&H_1 : \text{Var}_{[\kappa]} > \text{its mean}, \text{Var}_{[\omega]} > \text{its mean} && \text{(RI-CLPM (full))} \\
&H_5 : \text{Var}_{[\kappa]} > \text{its mean}, \text{25th percentile of Var}_{[\omega]} < \text{Var}_{[\omega]} < \text{75th percentile of Var}_{[\omega]} && \text{(RI-CLPM}(\kappa)) \\
&H_6 : \text{25th percentile of Var}_{[\kappa]} < \text{Var}_{[\kappa]} < \text{75th percentile of Var}_{[\kappa]}, \text{Var}_{[\omega]} > \text{its mean} && \text{(RI-CLPM}(\omega))
\end{aligned} \tag{1}$$

The results of the comparison between the use of GORICA with only the mean as boundaries (referred to as the "Old" approach) and the combination of mean and percentiles (referred to as the "New" approach) are presented in the following figures. Results in Figures 16, 17, 18, and 19 for selecting between CLPM and RI-CLPM, and in Figures 20, 21, 22, and 23 for selecting the number of RIs in RI-CLPM.

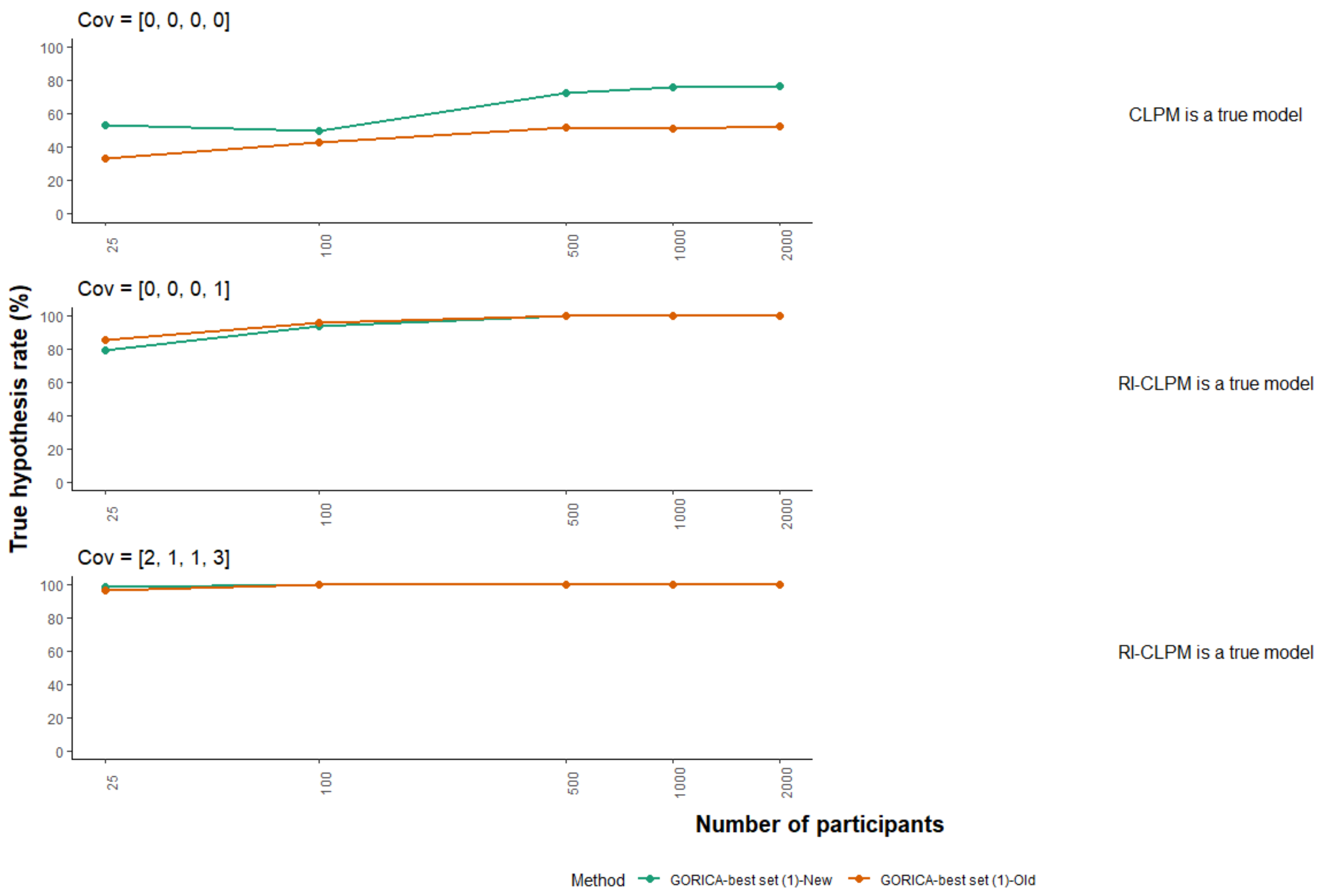

Figure 16: The performance of GORICA-best set (1)–New and –Old to select the true model for 3 waves of data. When the variances and covariance matrix of the random intercepts are $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

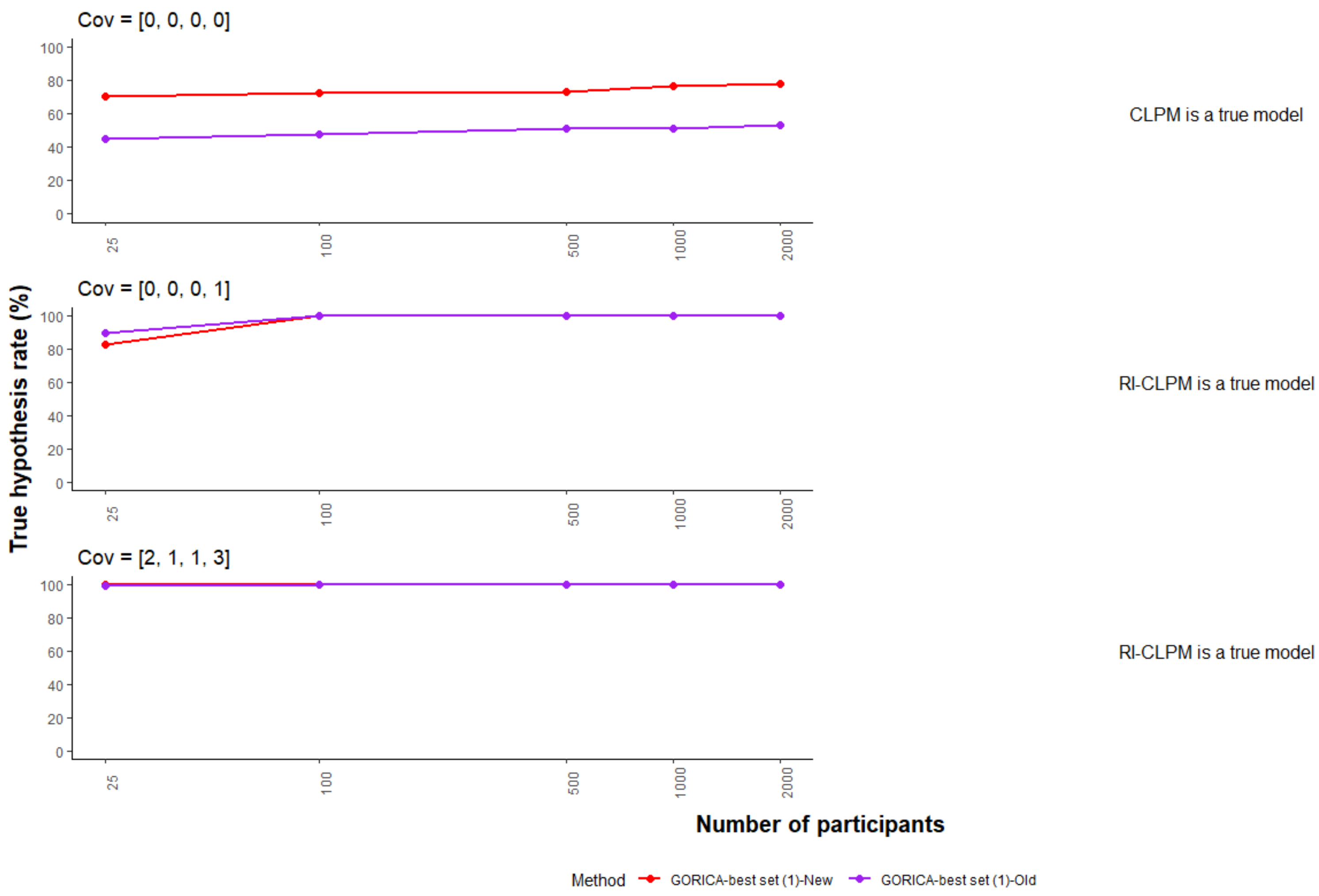

Figure 17: The performance of GORICA-best set (1)–New and –Old to select the true model for 4 waves of data. When the variances and covariance matrix of the random intercepts are $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

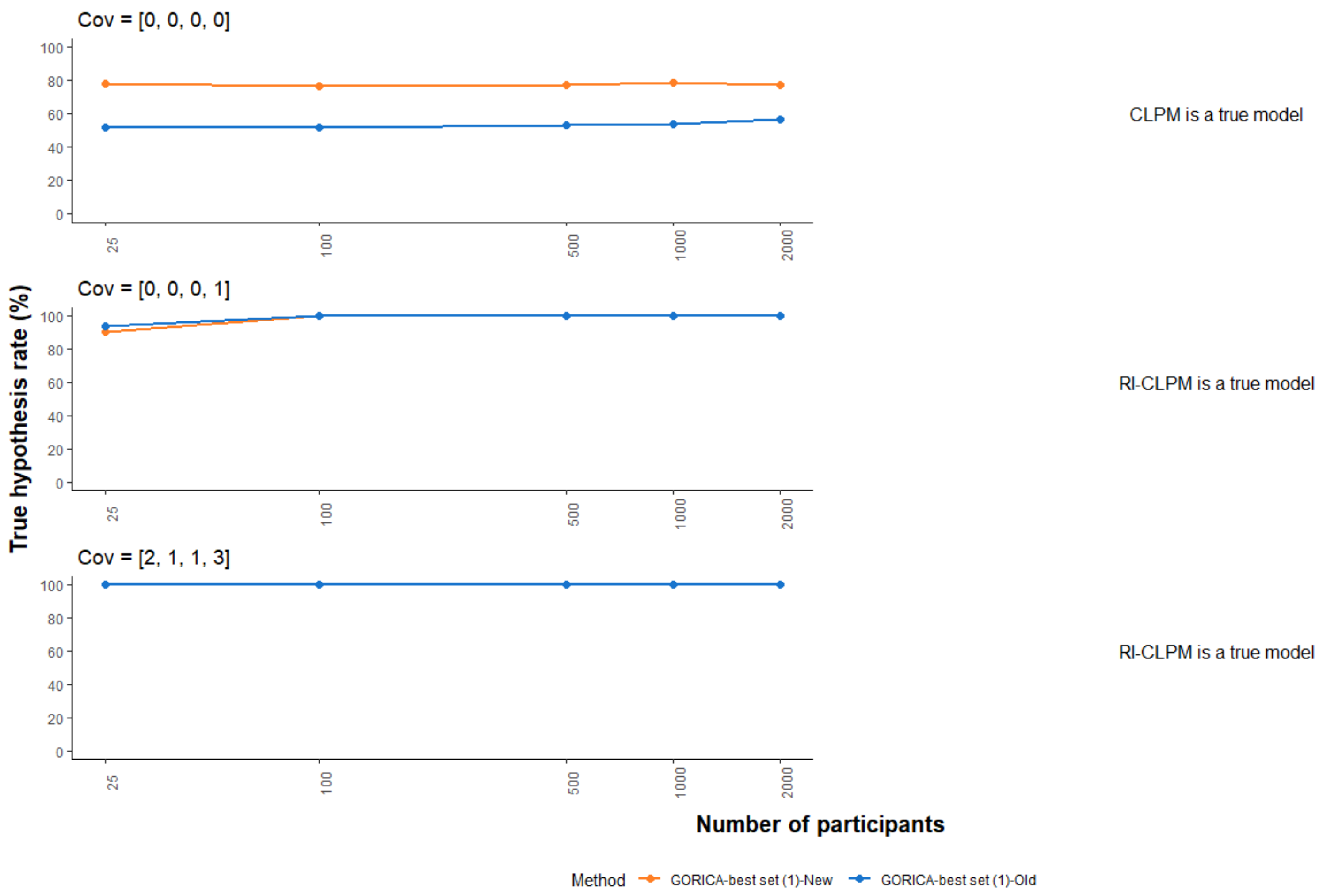

Figure 18: The performance of GORICA-best set (1)–New and –Old to select the true model for 5 waves of data. When the variances and covariance matrix of the random intercepts are $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

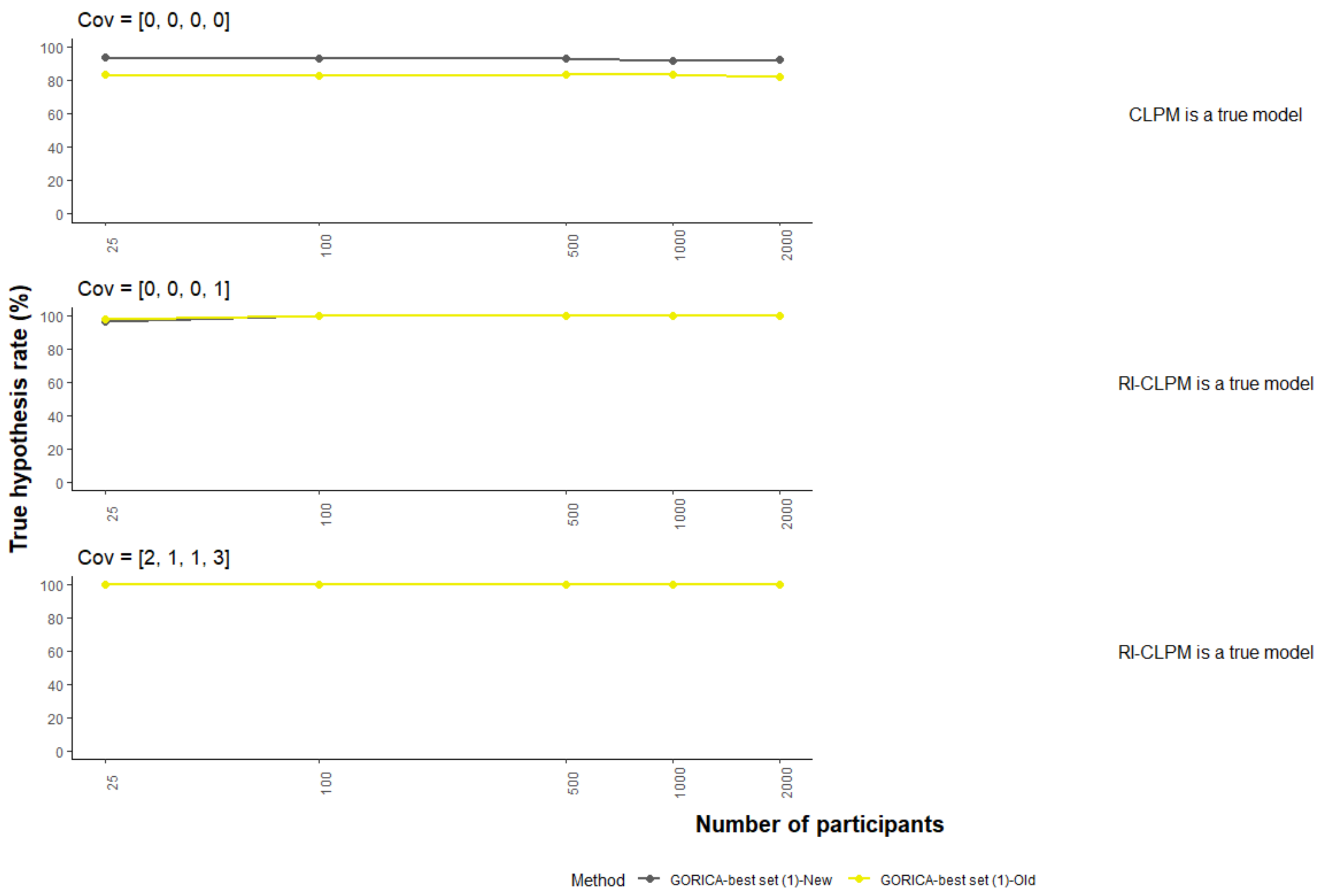

Figure 19: The performance of GORICA-best set (1)–New and –Old to select the true model for 6 waves of data. When the variances and covariance matrix of the random intercepts are $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM. For all other cases involving different variances and covariance matrices of the random intercepts, the true model is determined as RI-CLPM.

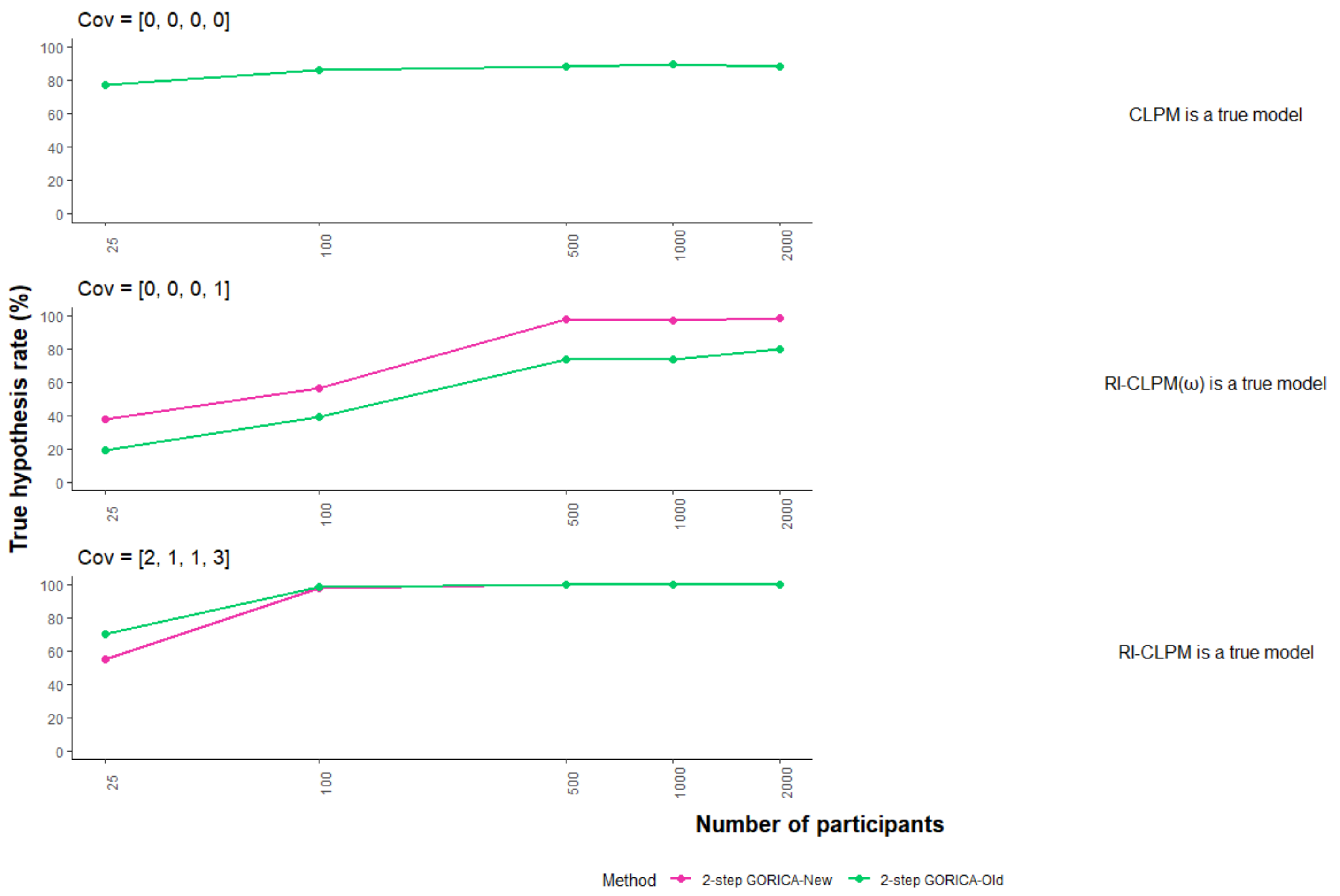

Figure 20: The performance of 2-step GORICA in selecting the true model for 3 waves of data varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all another case where the variances and covariance matrix are $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, the true model is determined as RI-CLPM (full), which includes two random intercepts.

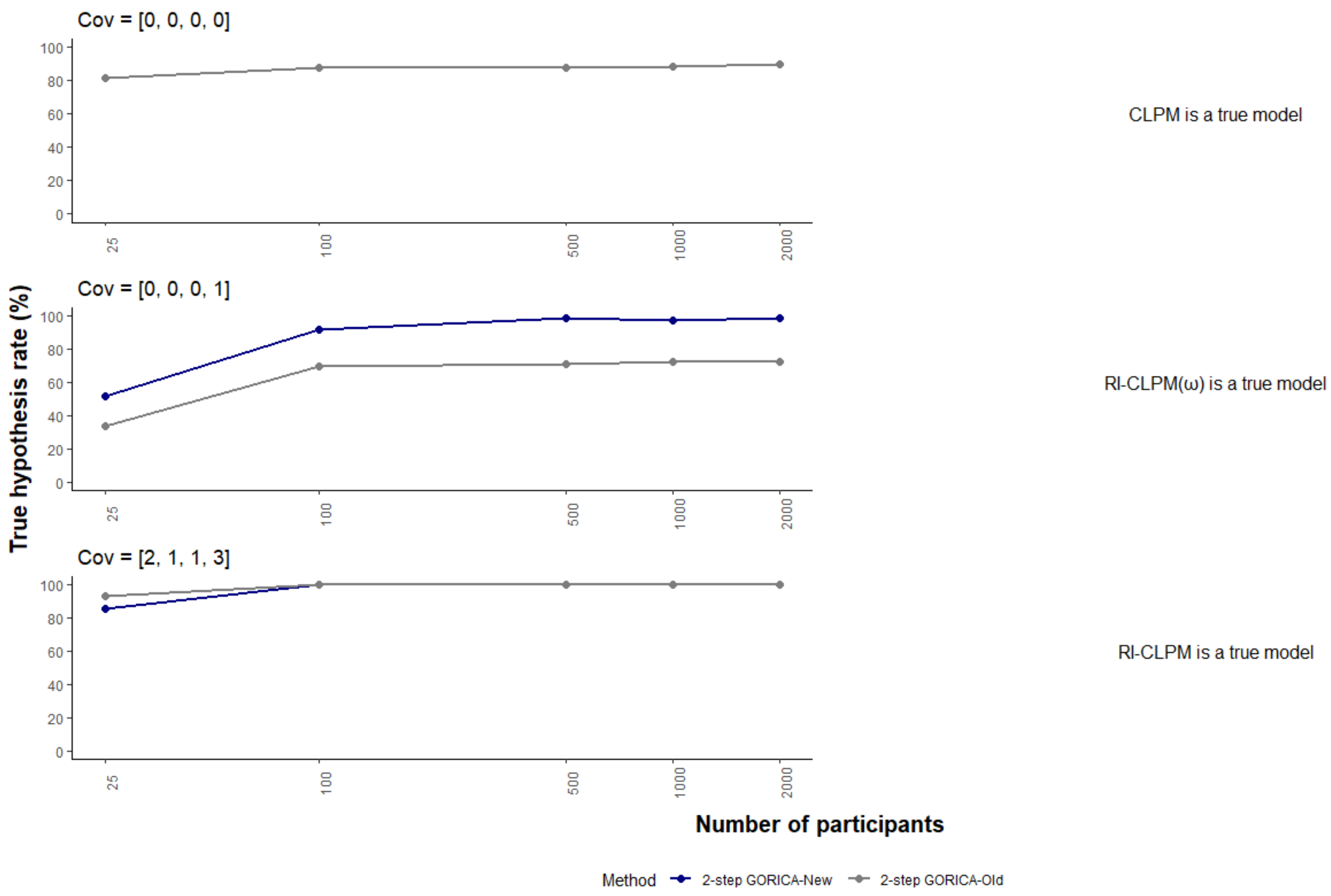

Figure 21: The performance of 2-step GORICA in selecting the true model for 4 waves of data varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all another case where the variances and covariance matrix are $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, the true model is determined as RI-CLPM (full), which includes two random intercepts.

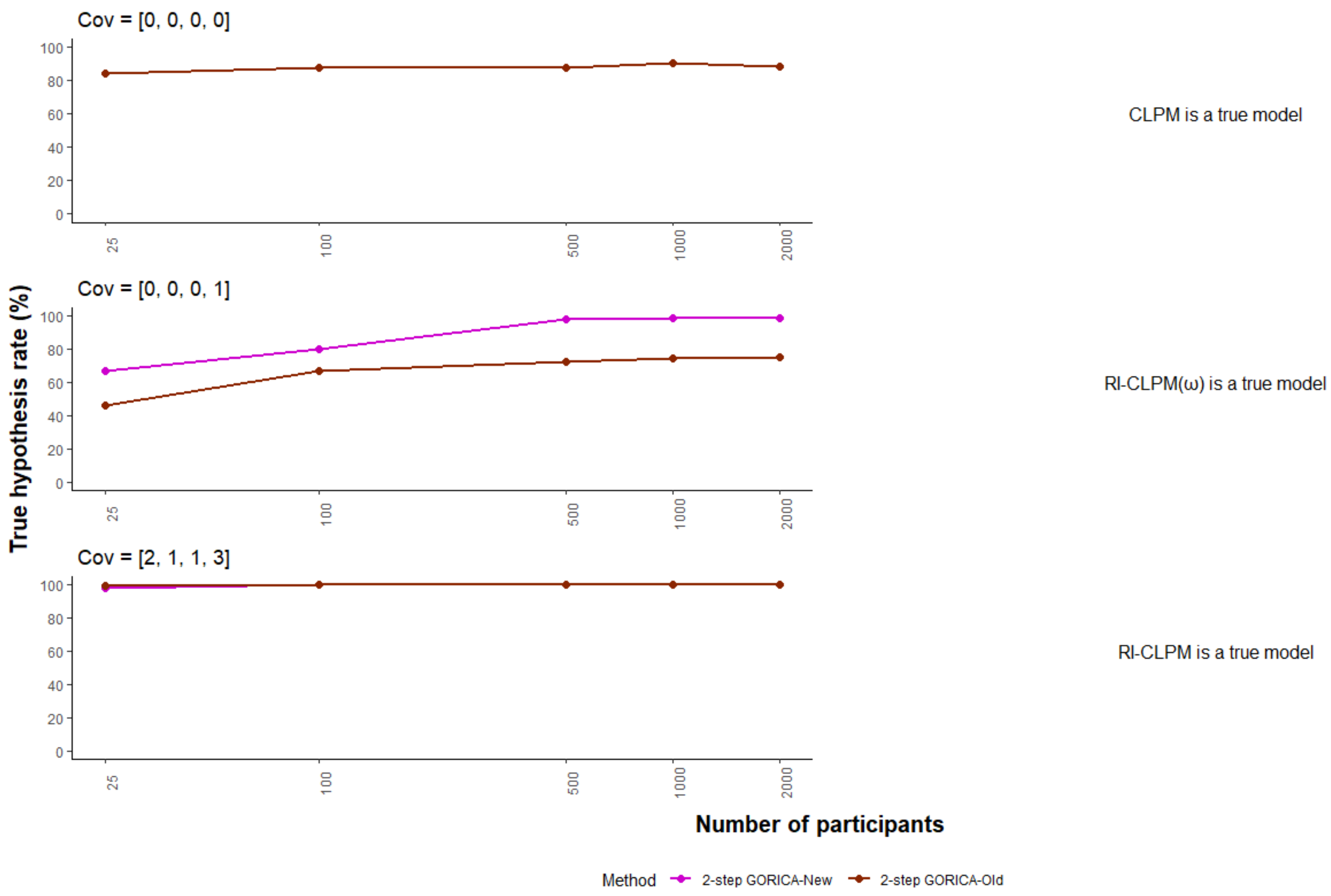

Figure 22: The performance of 2-step GORICA in selecting the true model for 5 waves of data varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all another case where the variances and covariance matrix are $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, the true model is determined as RI-CLPM (full), which includes two random intercepts.

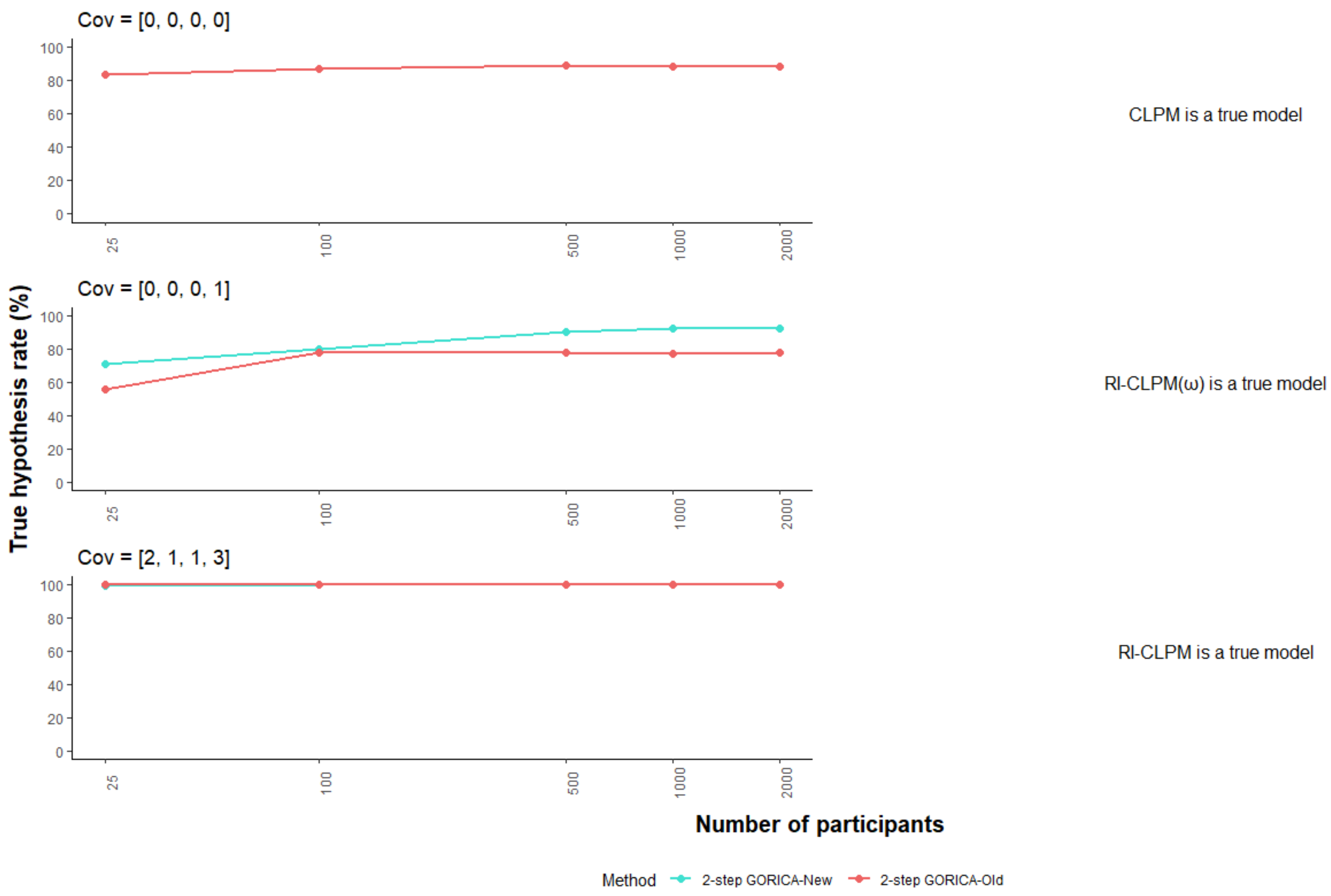

Figure 23: The performance of 2-step GORICA in selecting the true model for 6 waves of data varies based on different scenarios. When the variances and covariance matrix of the random intercepts is $\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the true model is identified as CLPM (RI-CLPM with zero random intercept). For cases where the variances and covariance matrix are $\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$, the true model is identified as RI-CLPM (ω) with one random intercept. In all another case where the variances and covariance matrix are $\begin{bmatrix} 2 & 1 \\ 1 & 3 \end{bmatrix}$, the true model is determined as RI-CLPM (full), which includes two random intercepts.