Web-based Supplementary Materials for "Bias-Corrected Diagonal Discriminant Rules for High-dimensional Data Classification"

Song Huang
Program of Computational Biology and Bioinformatics,
Yale University, New Haven, Connecticut, U.S.A.

Tiejun Tong
Department of Applied Mathematics,
University of Colorado, Boulder, Colorado, U.S.A.

Hongyu Zhao
Department of Epidemiology and Public Health,
Department of Genetics,
Yale University, New Haven, Connecticut, U.S.A.
hongyu.zhao@yale.edu

Web Appendix A: $E(\hat{L}_{k1})$, $E(\hat{L}_{k2})$ and $E(\hat{L}_k)$

Note that $\hat{\mu}_{ki} \sim N(\mu_{ki}, \sigma_{ki}^2/n_k)$, $\hat{\sigma}_{ki}^2 \sim \sigma_{ki}^2 \chi_{n_k-1}^2/(n_k-1)$, and $\hat{\mu}_{ki}$ and $\hat{\sigma}_{ki}^2$ are independent of each other. We have

$$
\begin{aligned}
E(\hat{L}_{k1}) &= \sum_{i=1}^{p} E\Big(\frac{(y_i - \hat{\mu}_{ki})^2}{\hat{\sigma}_{ki}^2}\Big) \\
&= \sum_{i=1}^{p} E(y_i - \mu_{ki} + \mu_{ki} - \hat{\mu}_{ki})^2 E\frac{1}{\hat{\sigma}_{ki}^2} \\
&= \sum_{i=1}^{p} \Big((y_i - \mu_{ki})^2 + \frac{\sigma_{ki}^2}{n_k}\Big) \frac{n_k - 1}{(n_k - 3)\sigma_{ki}^2} \\
&= \frac{n_k - 1}{n_k - 3} L_{k1} + \frac{(n_k - 1)p}{n_k(n_k - 3)}.
\end{aligned}
$$

For $\hat{L}_{k2}$, note that $E(\ln \chi_\nu^2) = \Psi(\nu/2) + \ln(2)$ (Tong and Wang, 2007), we have

$$
\begin{aligned}
E(\hat{L}_{k2}) &= \sum_{i=1}^{p} E(\ln \hat{\sigma}_{ki}^2) \\
&= \sum_{i=1}^{p} E\Big(\ln(\frac{(n_k - 1)\hat{\sigma}_{ki}^2}{\sigma_{ki}^2}) + \ln(\frac{\sigma_{ki}^2}{n_k - 1})\Big) \\
&= \sum_{i=1}^{p} \Big(\Psi(\frac{n_k - 1}{2}) + \ln(2) + \ln(\frac{\sigma_{ki}^2}{n_k - 1})\Big) \\
&= L_{k2} + p\Big(\Psi(\frac{n_k - 1}{2}) - \ln(\frac{n_k - 1}{2})\Big).
\end{aligned}
$$

For $\hat{L}_k$, note that $\hat{\sigma}_i^2 \sim \sigma_i^2 \chi_{n-K}^2/(n-K)$, we have

$$
\begin{aligned}
E(\hat{L}_k) &= \sum_{i=1}^{p} E\Big(\frac{(y_i - \hat{\mu}_{ki})^2}{\hat{\sigma}_i^2}\Big) \\
&= \sum_{i=1}^{p} E(y_i - \mu_{ki} + \mu_{ki} - \hat{\mu}_{ki})^2 E\frac{1}{\hat{\sigma}_i^2} \\
&= \sum_{i=1}^{p} \Big((y_i - \mu_{ki})^2 + \frac{\sigma_{ki}^2}{n_k}\Big) \frac{n - K}{(n - K - 2)\sigma_i^2} \\
&= \frac{n - K}{n - K - 2} L_{k1} + \frac{(n - K)p}{(n - K - 2)n_k}.
\end{aligned}
$$

Web Figure 1: Comparison of methods with $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.3$ (sample version)

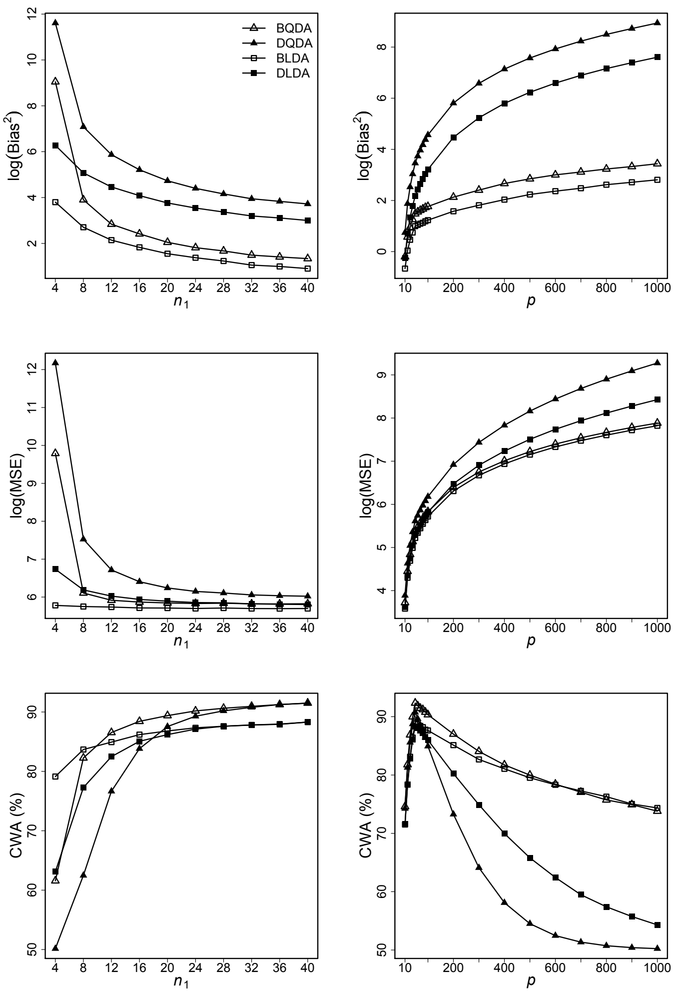

Figure 1: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.3$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 2: Comparison of methods with $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.3$ (MLE-based)

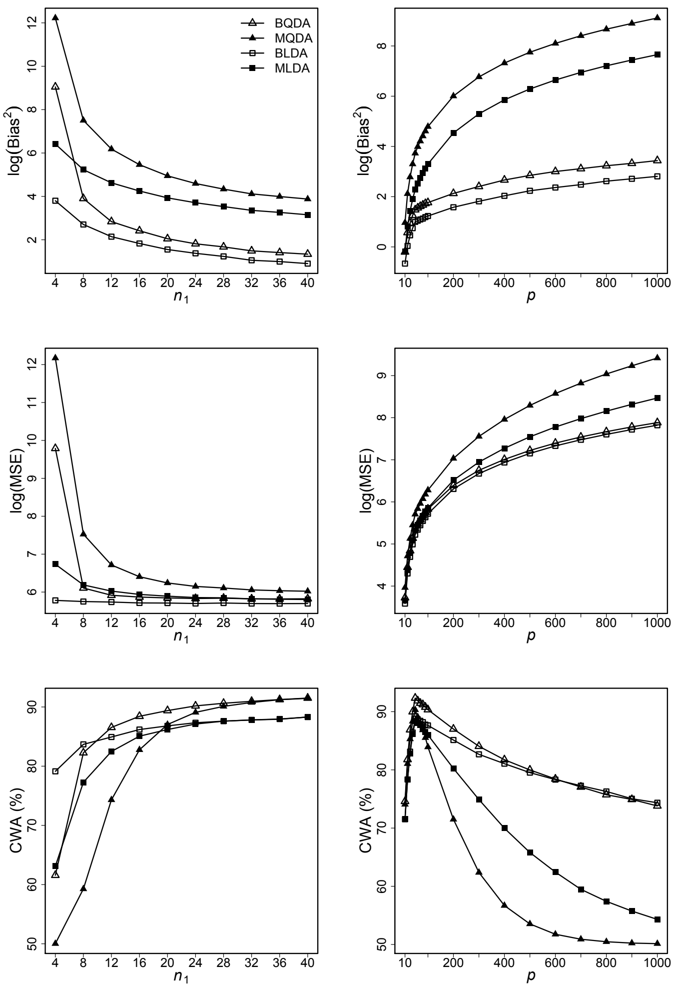

Figure 2: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.3$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 3: Comparison of methods with $\Sigma_1 = \Sigma_2$ and $\rho = 0.3$ (MLE-based)

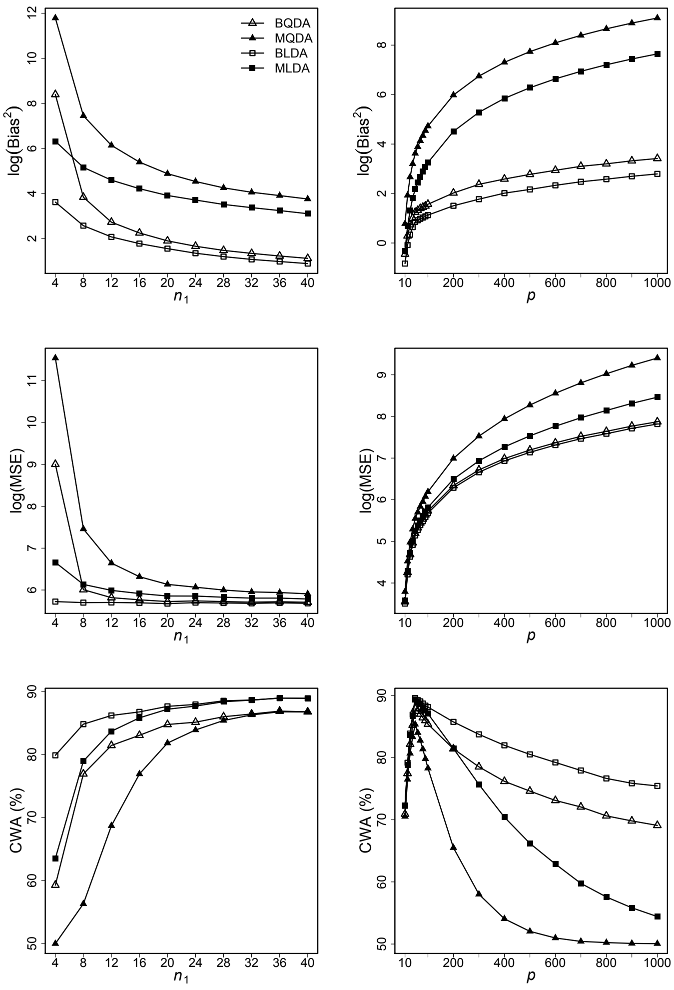

Figure 3: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.3$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 4: Comparison of methods with $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$ (sample version)

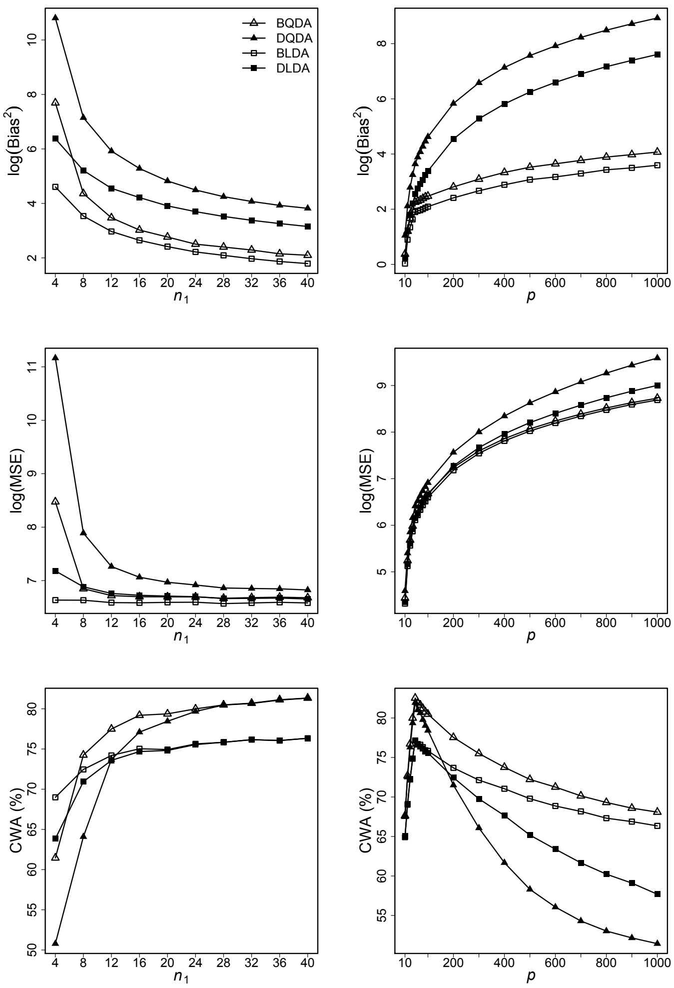

Figure 4: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 5: Comparison of methods with $\Sigma_1 = \Sigma_2$ and $\rho = 0.7$ (sample version)

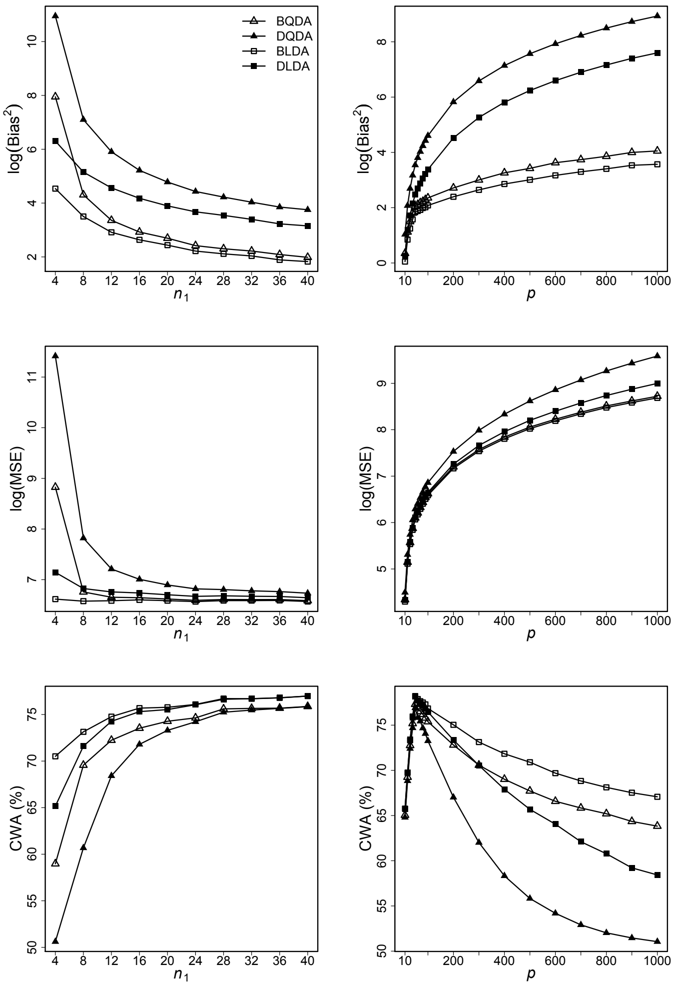

Figure 5: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 6: Comparison of methods with $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$ (MLE-based)

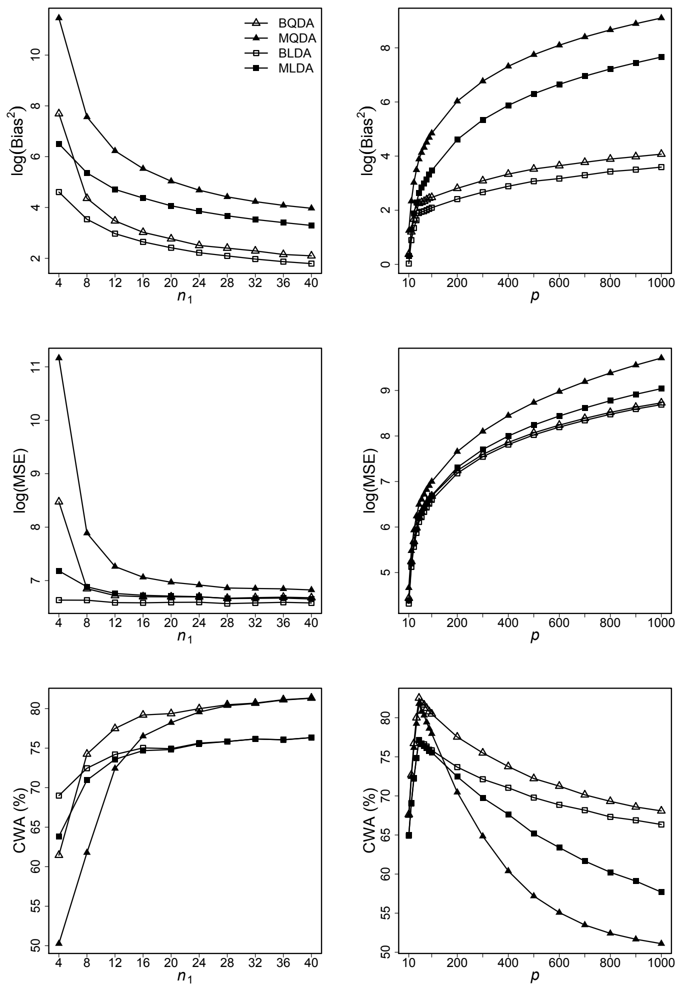

Figure 6: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 7: Comparison of methods with $\Sigma_1 = \Sigma_2$ and $\rho = 0.7$ (MLE-based)

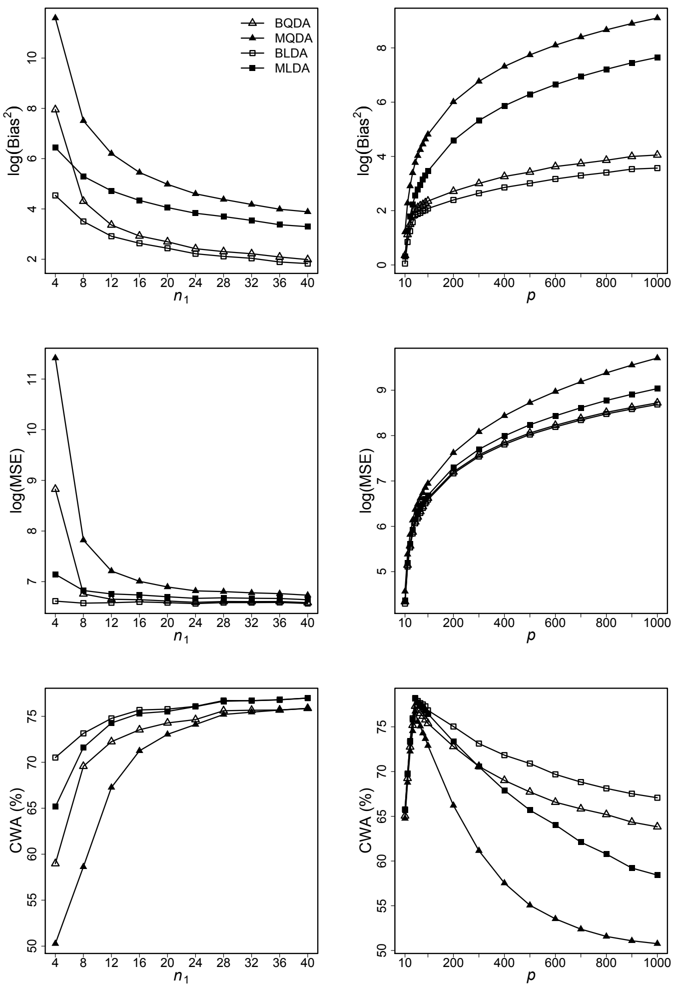

Figure 7: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2$ and $\rho = 0.7$. Left column: $p = 100$. Right column: $n_1 = 20$ and $n_2 = 100$.

Web Figure 8: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 1$ (sample version)

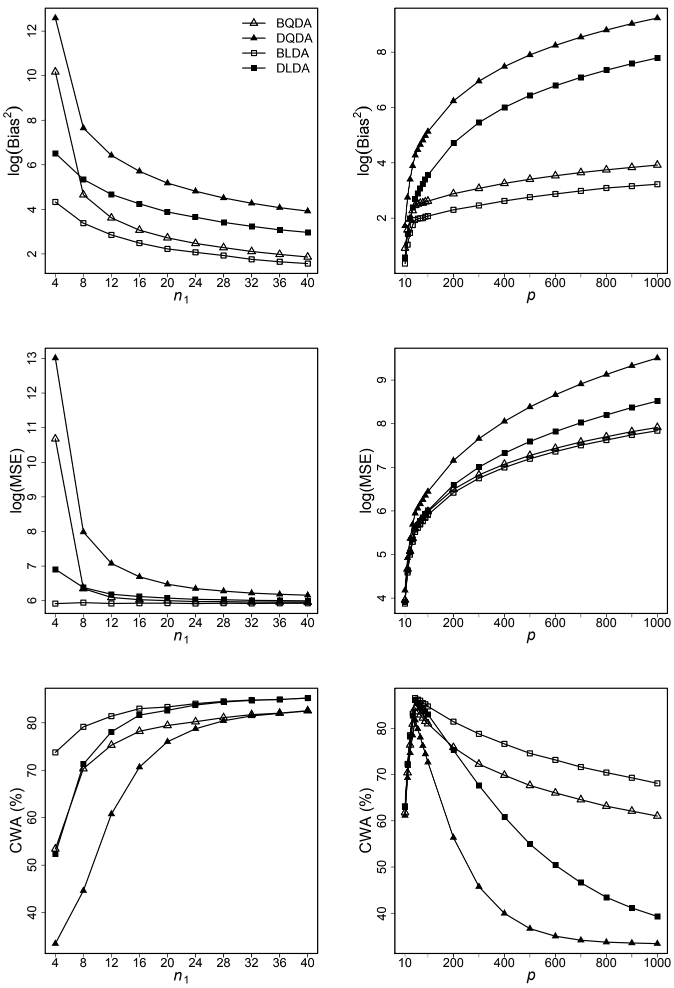

Figure 8: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 1$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 20$.

Web Figure 9: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 1$ (MLE-based)

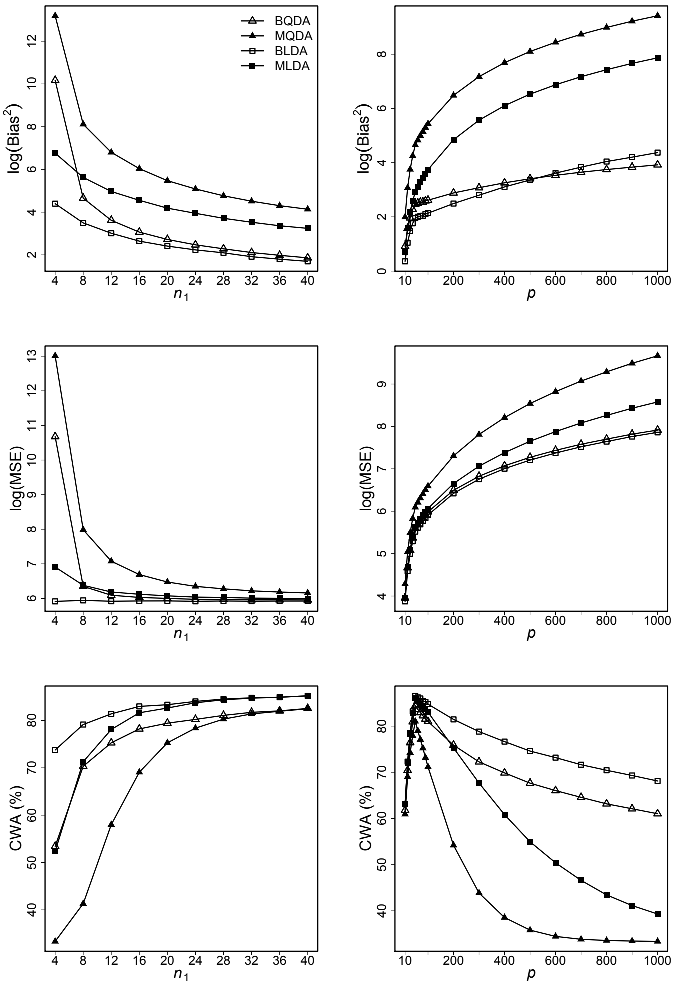

Figure 9: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 1$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 20$.

Web Figure 10: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 1$ (sample version)

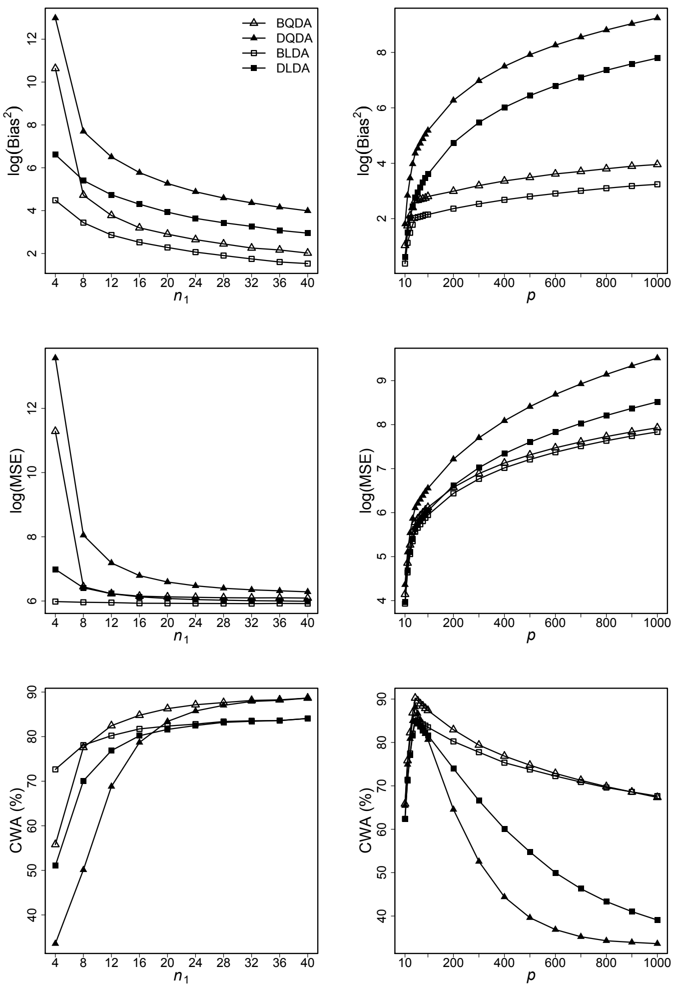

Figure 10: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 1$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 20$.

Web Figure 11: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 1$ (MLE-based)

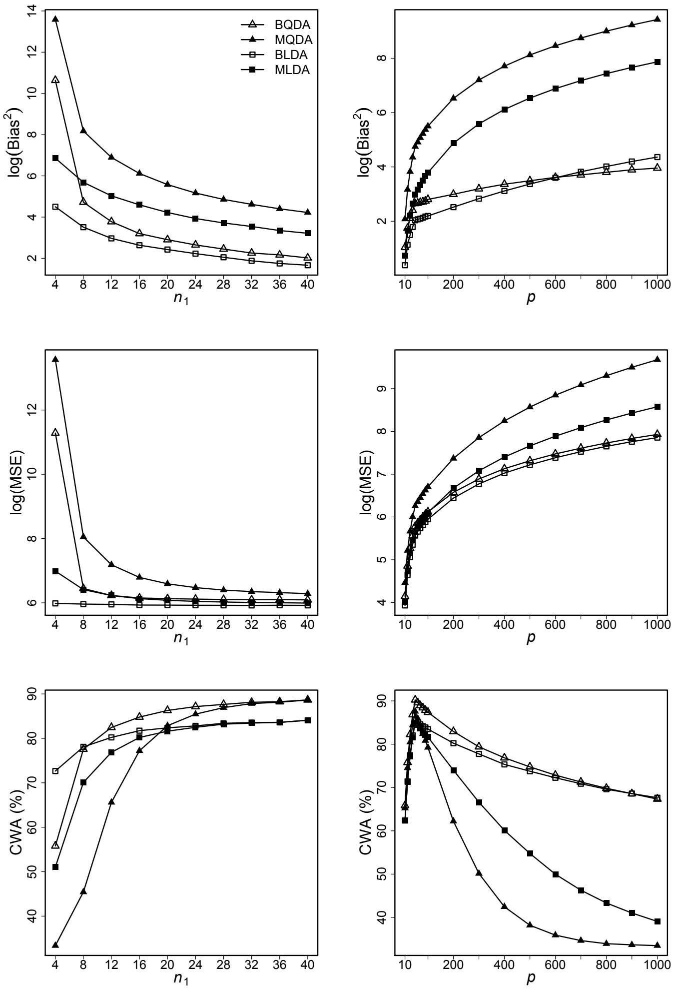

Figure 11: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 1$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 20$.

Web Figure 12: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 5$ (sample version)

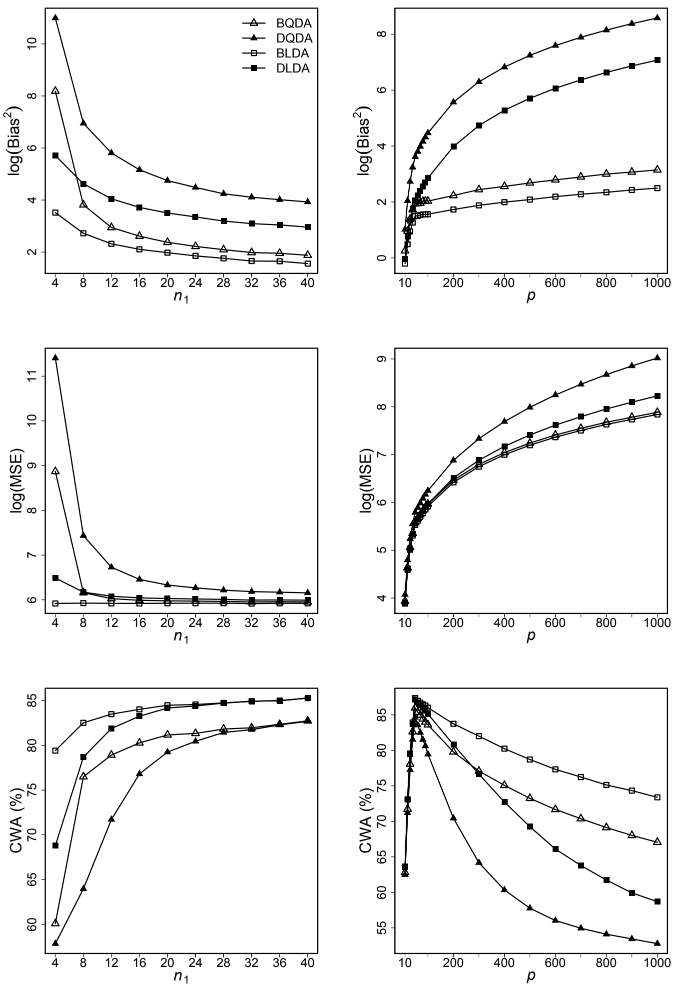

Figure 12: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 5$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 100$.

Web Figure 13: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 5$ (MLE-based)

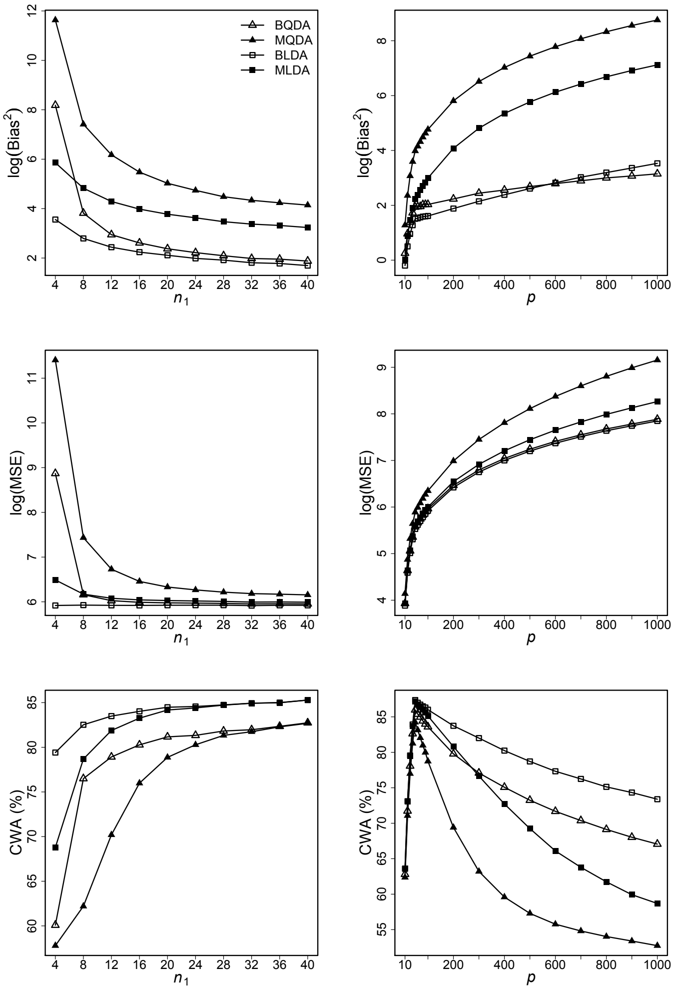

Figure 13: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 5$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 100$.

Web Figure 14: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 5$ (sample version)

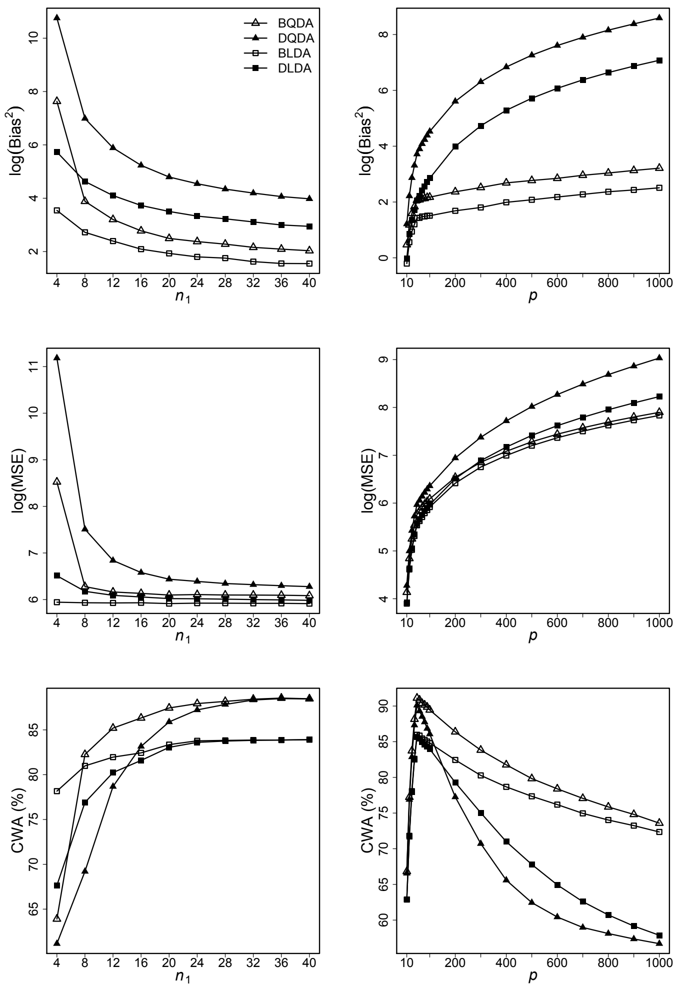

Figure 14: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 5$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 100$.

Web Figure 15: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 5$ (MLE-based)

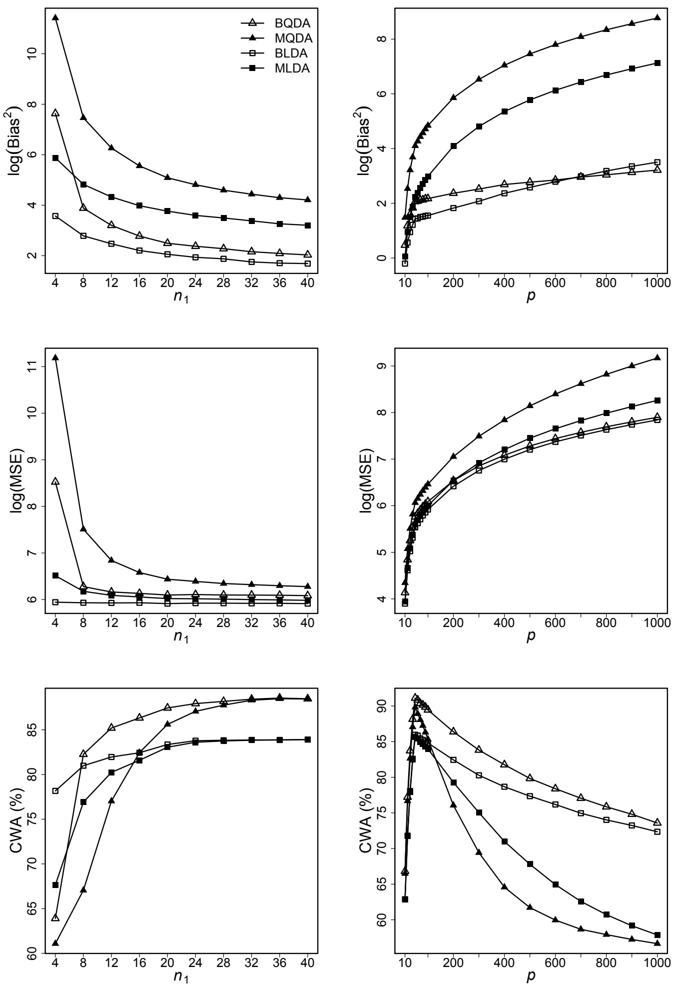

Figure 15: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 5$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 100$.

Web Figure 16: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 10$ (sample version)

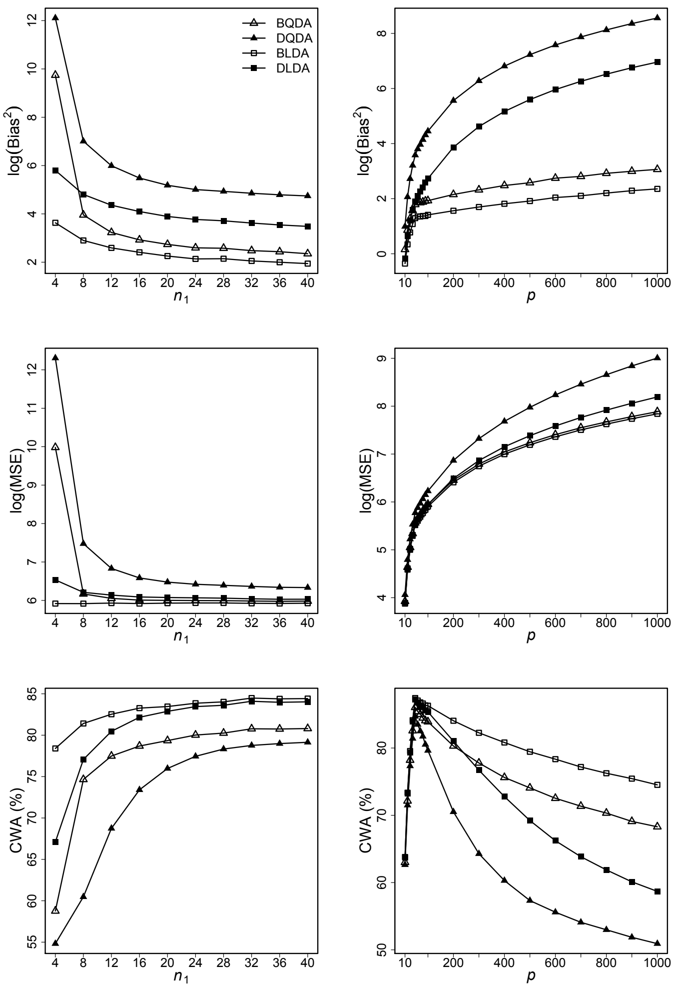

Figure 16: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 10$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 200$.

Web Figure 17: Comparison of methods with $\Sigma_1 = \Sigma_2 = \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 10$ (MLE-based)

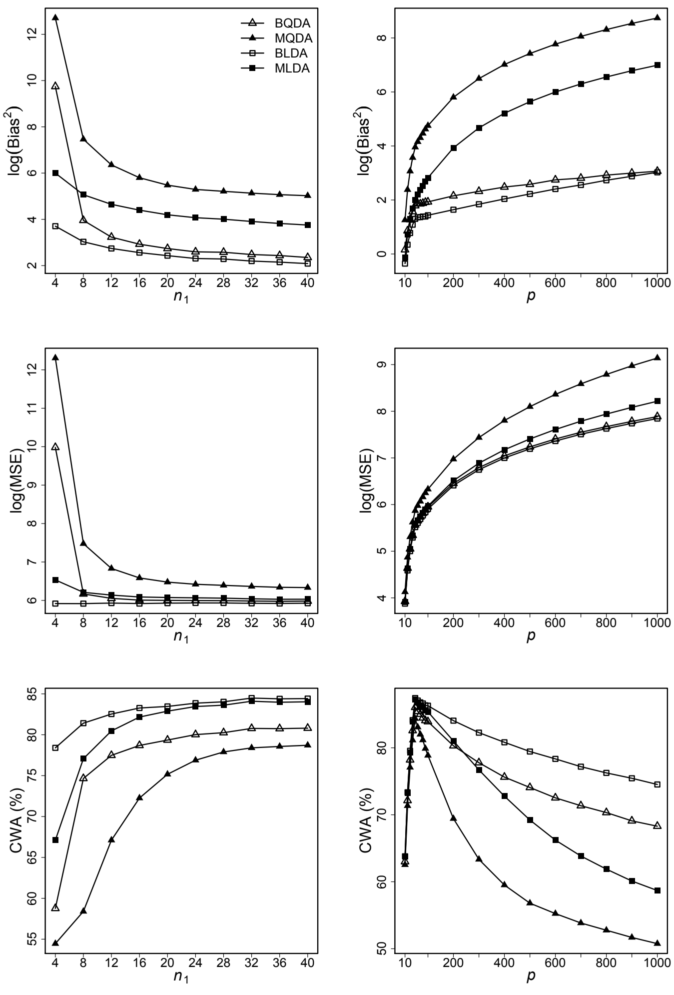

Figure 17: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 = \Sigma_2 = \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 10$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 200$.

Web Figure 18: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 10$ (sample version)

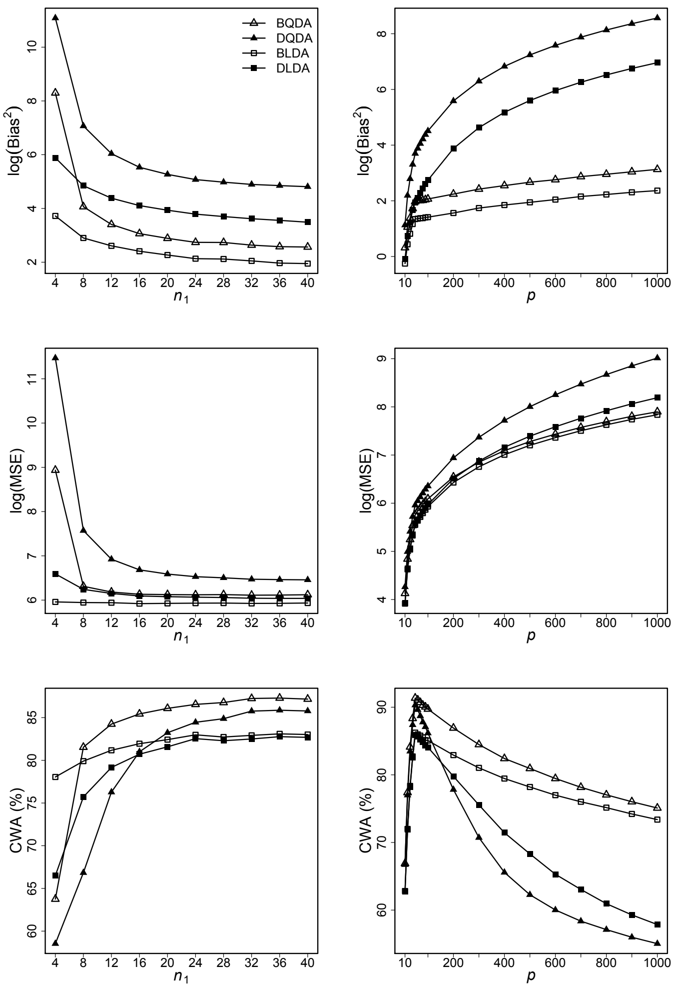

Figure 18: Comparison of bias-corrected discriminant scores with the original ones (sample version) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 10$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 200$.

Web Figure 19: Comparison of methods with $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$, $\rho = 0.3$, and $n_1 : n_2 : n_3 = 1 : 5 : 10$ (MLE-based)

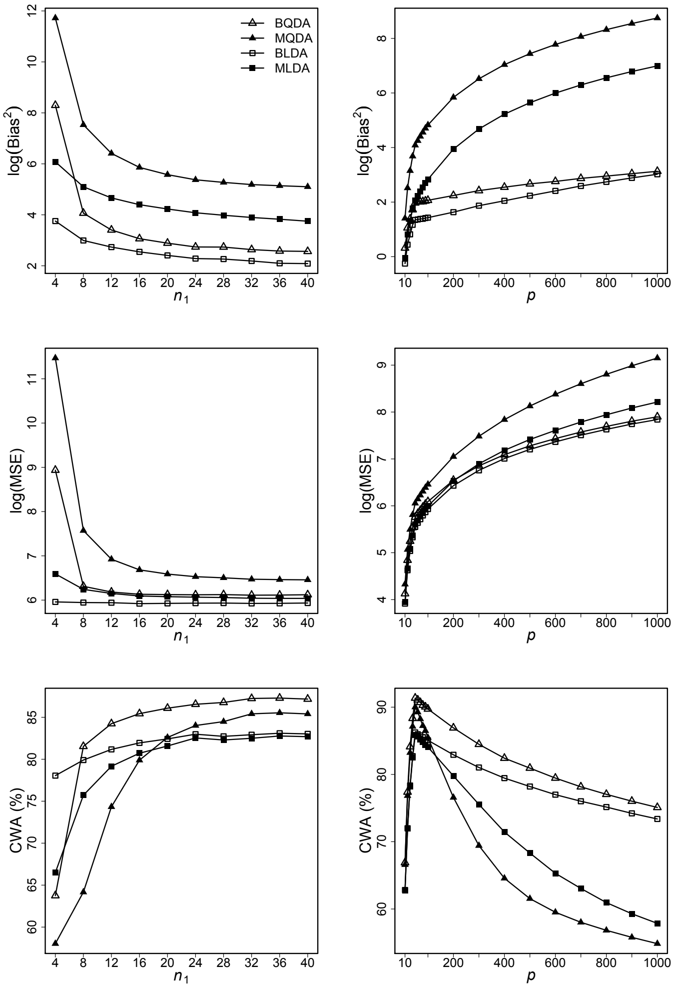

Figure 19: Comparison of bias-corrected discriminant scores with the original ones (MLE-based) when $\Sigma_1 \neq \Sigma_2 \neq \Sigma_3$ and $\rho = 0.3$. Left column: $p = 100$ and $n_1 : n_2 : n_3 = 1 : 5 : 10$. Right column: $n_1 = 20$, $n_2 = 100$ and $n_3 = 200$.

References

[1] Tong, T. and Wang, Y. (2007). Optimal Shrinkage Estimation of Variances With Applications to Microarray Data Analysis. *Journal of the American Statistical Association* **102,** 113-122.