Supplementary Material for Principle ERP reduction and analysis: Estimating and using *principle ERP* waveforms underlying ERPs across tasks, subjects and electrodes

Emilie Campos[a], Chad Hazlett[b], Patricia Tan[c], Holly Truong[c], Sandra Loo[c], Charlotte DiStefano[c], Shafali Jeste[c], and Damla Şentürk[d]

[a]Co-first author; Department of Biostatistics, University of California, Los Angeles
[b]Co-first author; Department of Statistics, University of California, Los Angeles
[c]Department of Psychiatry, University of California, Los Angeles
[d]Department of Biostatistics, University of California, Los Angeles

December 28, 2019

Effects of Sample Size

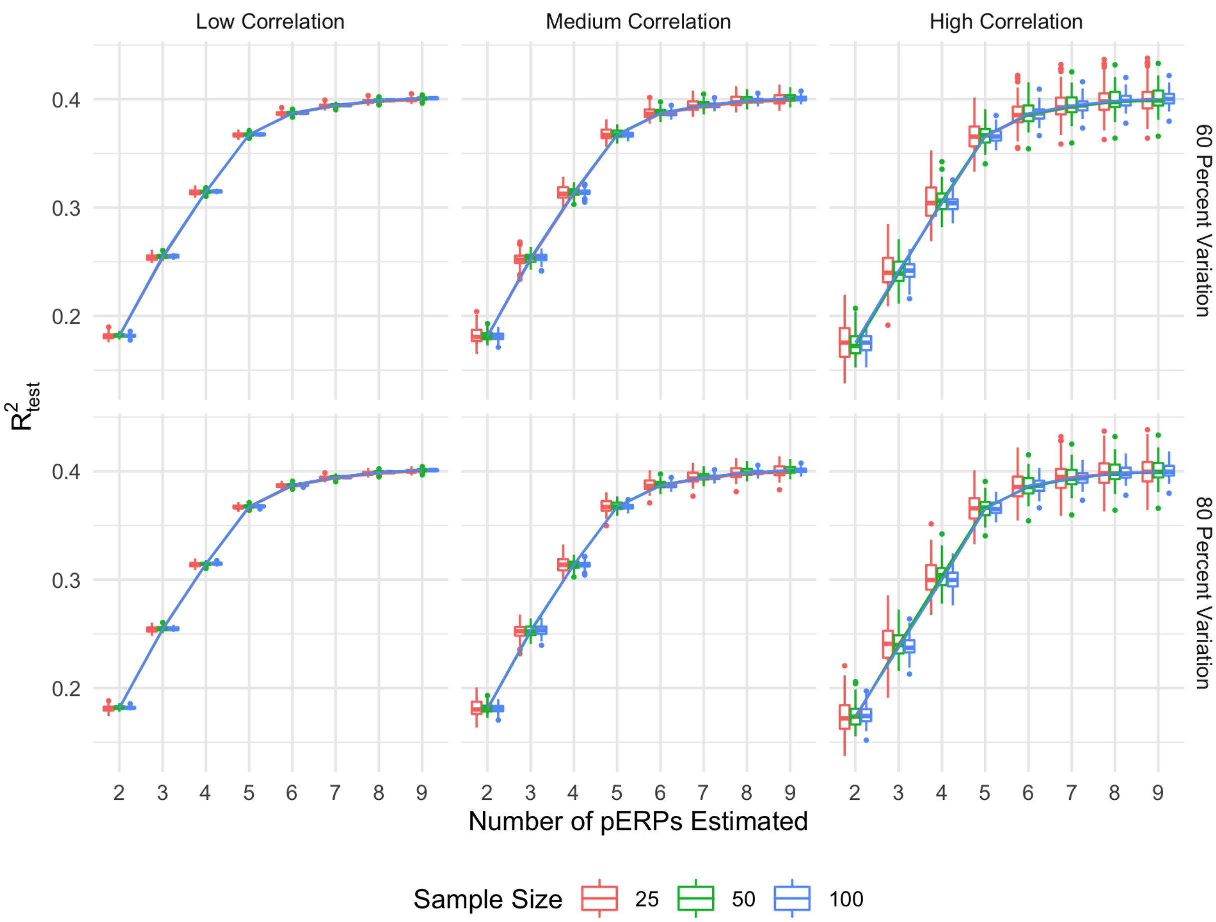

Fig. 1: The boxplots of R^2_{test} as a function of the number of pERPs estimated across 100 simulations with 60 (first row) and 80 percent variation retained (second row) and low (first column), medium (second column) and high (third column) correlation between electrodes and tasks for $N = 25$ (red), 50 (green), and 100 (blue) in the high noise set up. The variation in prediction accuracy decreases with increasing same size, as expected.

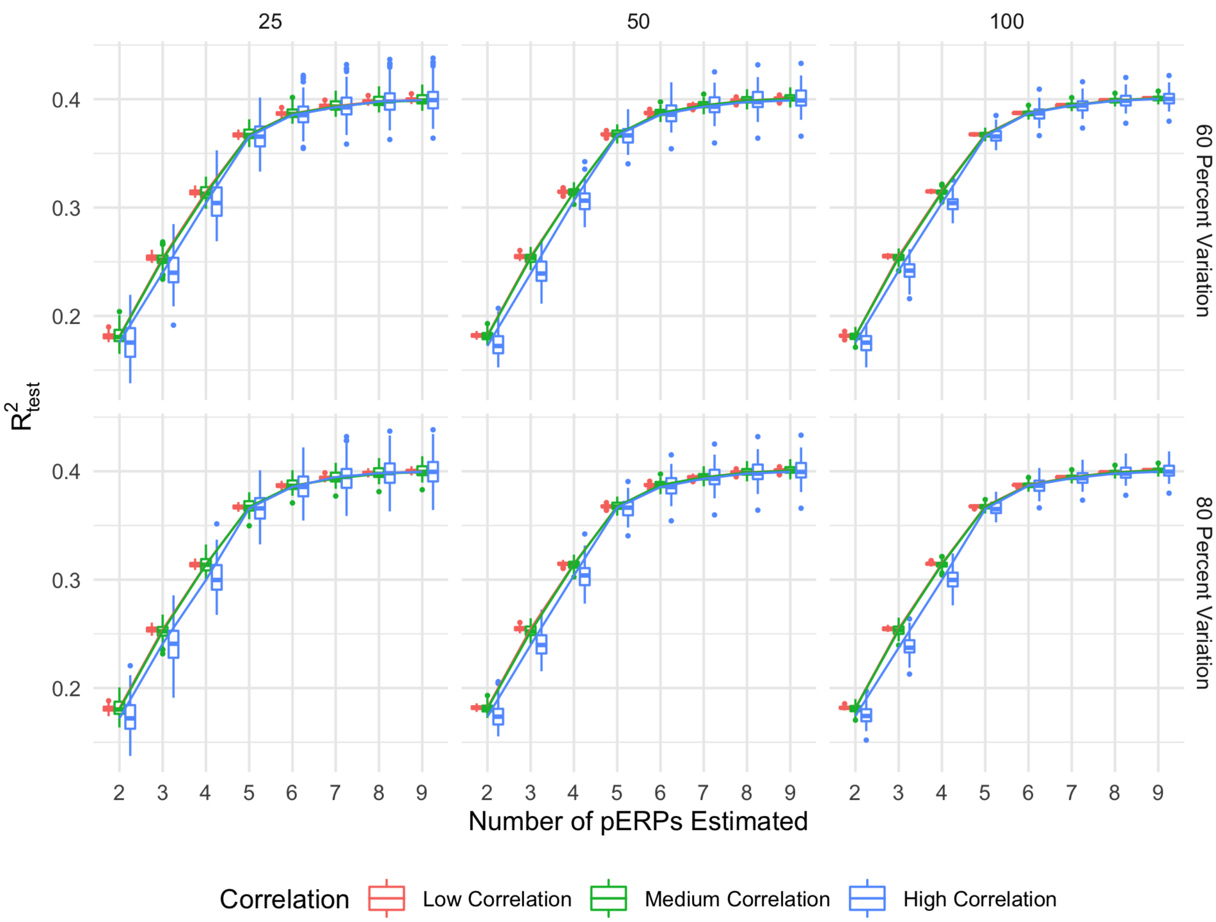

Fig. 2: The boxplots of R^2_{test} as a function of the number of pERPs estimated across 100 simulations with 60 (first row) and 80 percent variation retained (second row) and $N = 25$ (first column), 50 (second column), and 100 (third column) for low (red), medium (green) and high (blue) correlation between electrodes and tasks in the high noise set-up. With increasing correlation, the prediction accuracy gets worse (with higher variation).

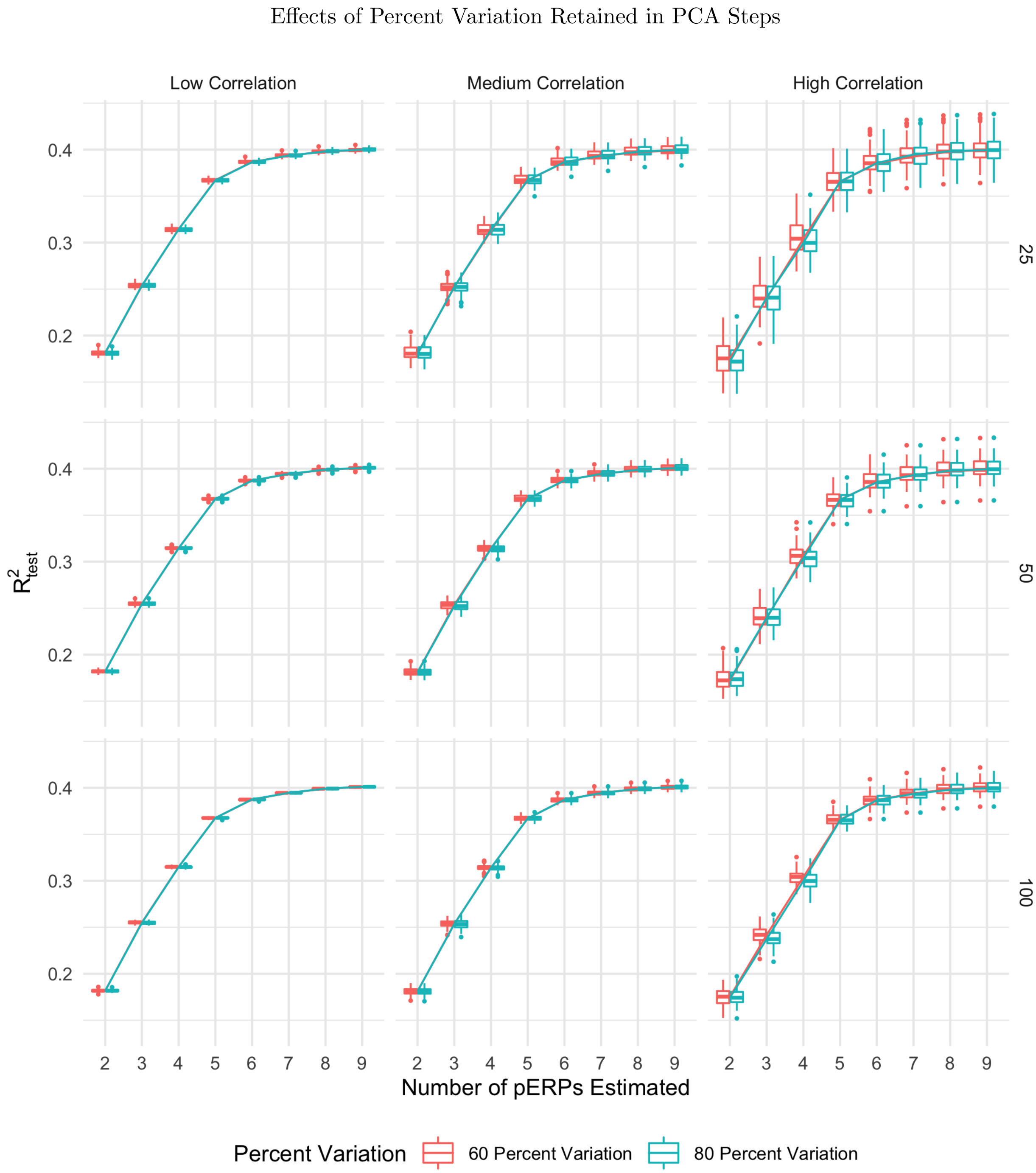

Fig. 3: The boxplots of R^2_{test} as a function of the number of pERPs estimated across 100 simulations with $N = 25$ (first row), 50 (second row), and 100 (third row) and low (first column), medium (second column) and high (third column) correlation between electrodes and tasks for 60 (red) and 80 (blue) percent variation retained in the high noise set-up. When the correlation is high, retaining more variation in PCA corresponds to retaining more noise, leading to worse prediction accuracy.

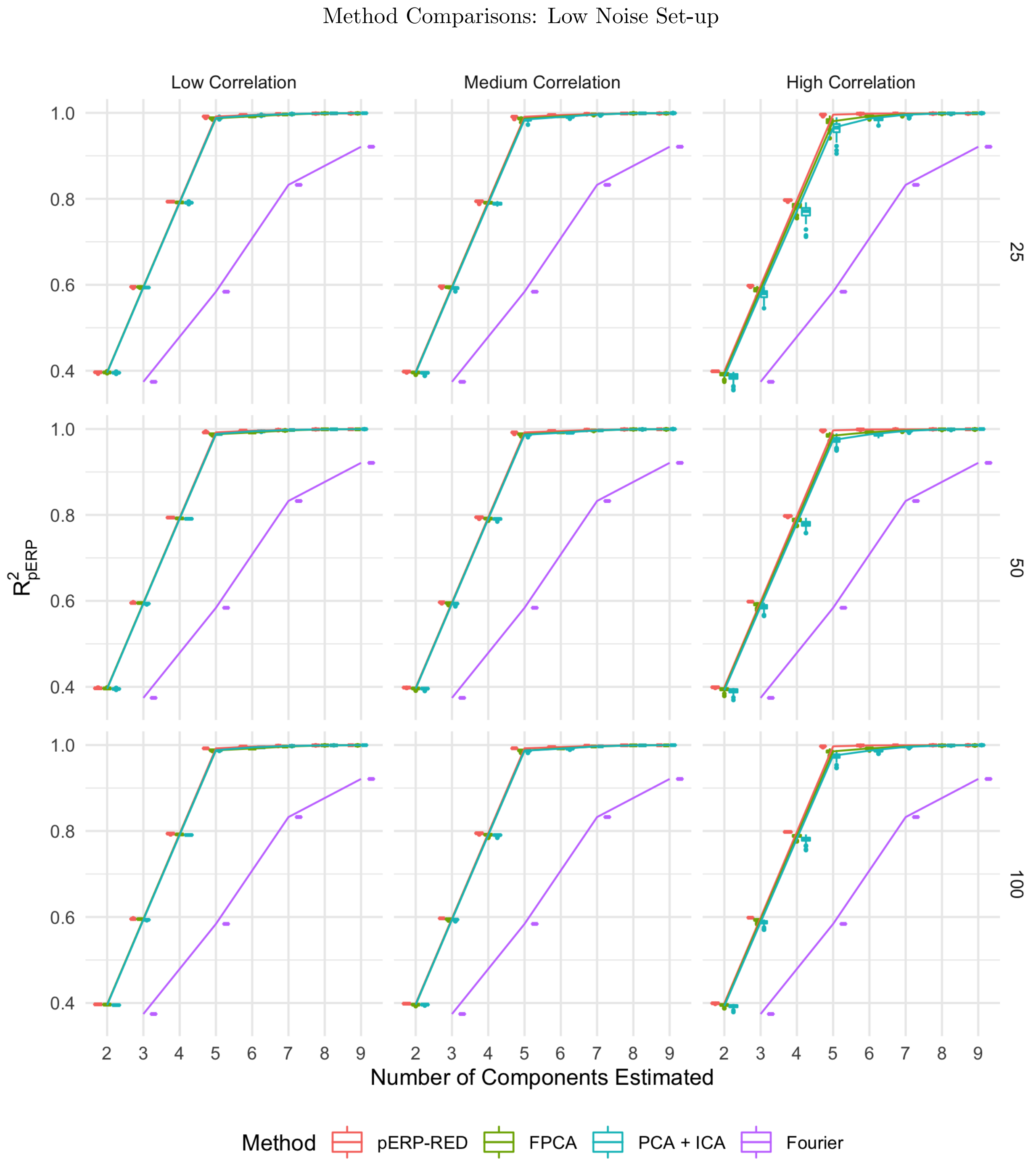

Fig. 4: The boxplots of R^2_{pERP} as a function of the number of components estimated across 100 simulation runs, using pERP-RED (red), FPCA (green), PCA + ICA (single-PCA) (blue), and Fourier (purple) with $N = 25$ (first row), 50 (second row), and 100 (third row) and low (first column), medium (second column) and high (third column) correlation (third column) in the low noise set-up. 60% variation is retained in PCA steps.

Methods Comparisons: High Noise Set-up

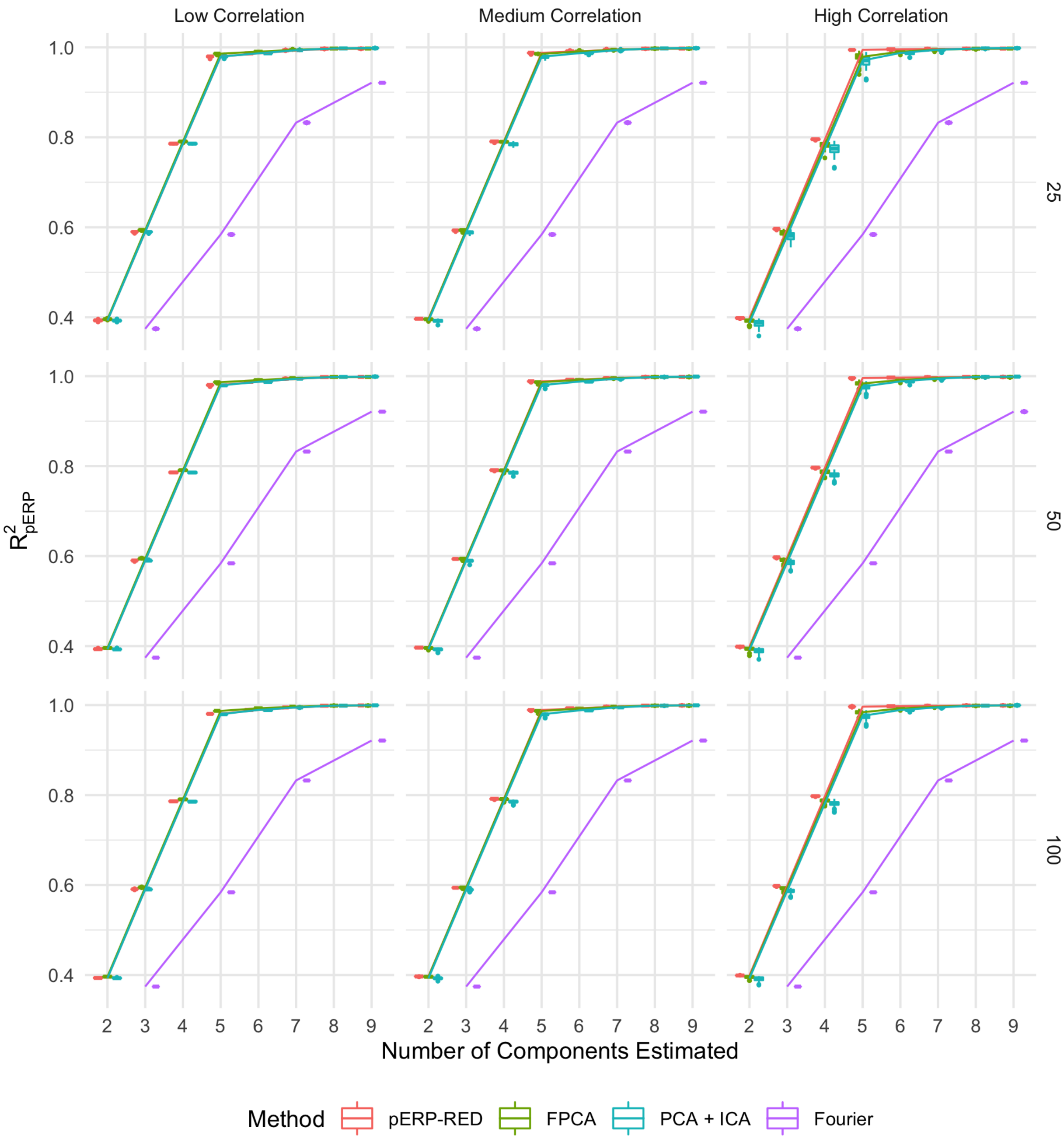

Fig. 5: The boxplots of R^2_{pERP} as a function of the number of components estimated across 100 simulation runs, using pERP-RED (red), FPCA (green), PCA + ICA (single-PCA) (blue), and Fourier (purple) with $N = 25$ (first row), 50 (second row), and 100 (third row) and low (first column), medium (second column) and high (third column) correlation in the high noise set-up. 60% variation is retained in PCA steps.

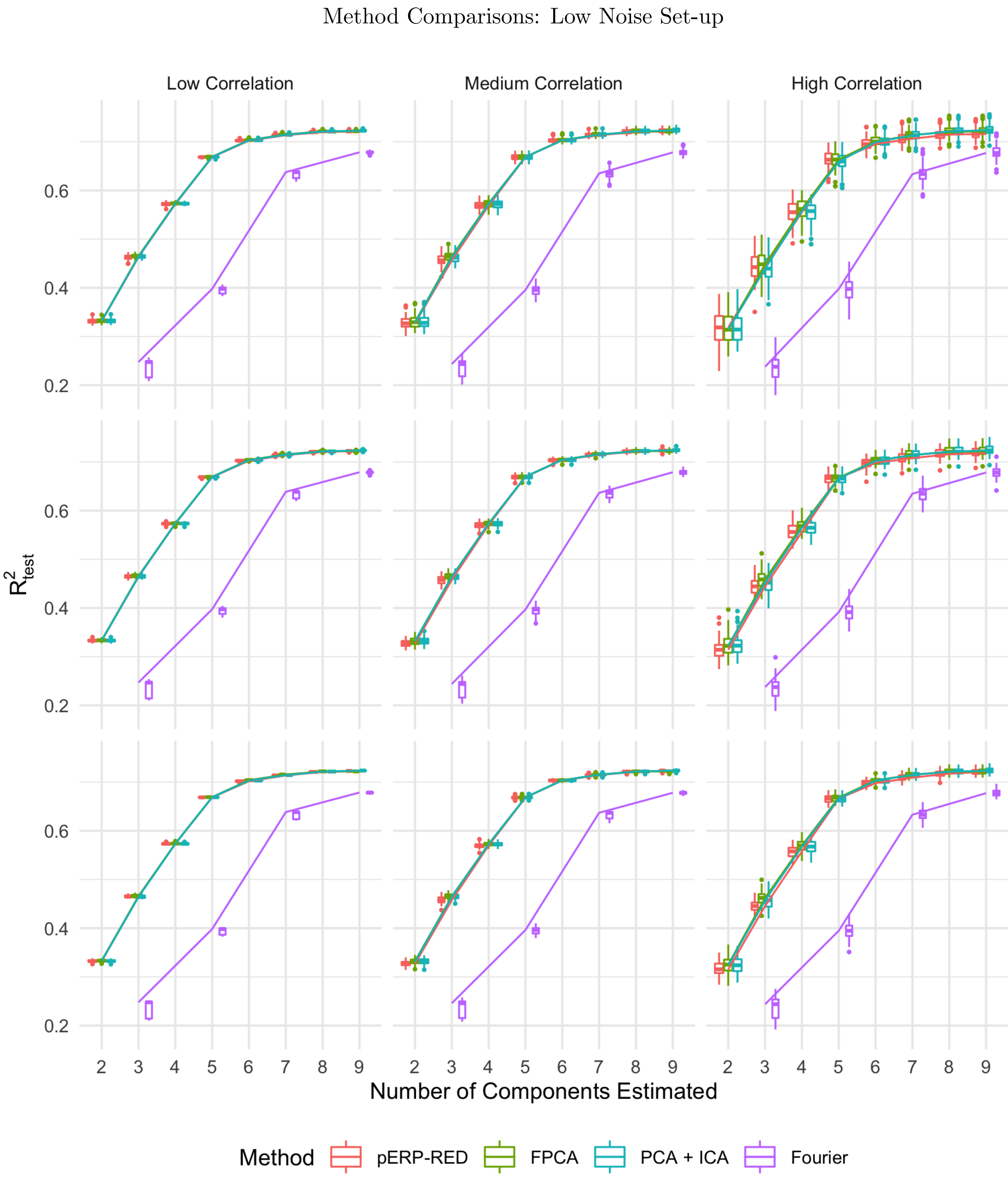

Fig. 6: The boxplots of the R^2_{test} as a function of the number of components estimated across 100 simulation runs, using pERP-RED (red), FPCA (green), PCA + ICA (single-PCA) (blue), and Fourier (purple) with $N = 25$ (first row), 50 (second row), and 100 (third row) and low (first column), medium (second column) and high (third column) correlation (third column) in the low noise set-up. 60% variation is retained in PCA steps.

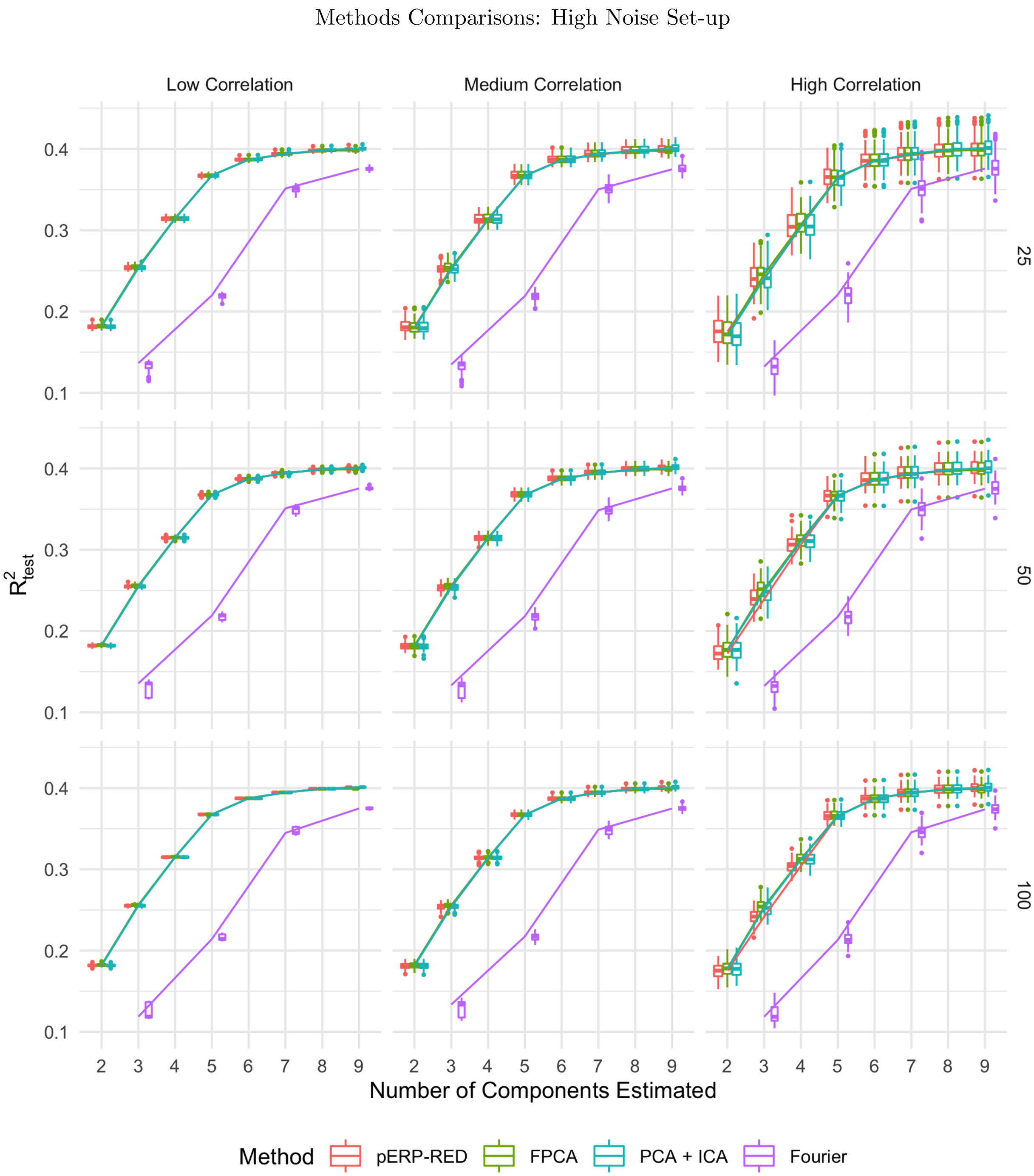

Fig. 7: The boxplots of the R^2_{test} as a function of the number of components estimated across 100 simulation runs, using pERP-RED (red), FPCA (green), PCA + ICA (single-PCA) (blue), and Fourier (purple) with $N = 25$ (first row), 50 (second row), and 100 (third row) and low (first column), medium (second column) and high (third column) correlation in the high noise set-up. 60% variation is retained in PCA steps.