Supplementary Material: Distribution-Free Calibration of Statistical Confidence Sets

November 28, 2024

In this supplementary material, we present a comprehensive comparison of methods across all combinations of sample sizes and simulation budgets, covering various statistics and model/simulator configurations in both tractable likelihood and likelihood-free experiments. Specifically, we provide the MAE curves for each method as a function of the sample size B, evaluated across different sample sizes n. For clarity, the results are divided into two sections: Section 1 focuses on the MAE curves for tractable likelihood experiments, while Section 2 covers the likelihood-free experiments. Within each section, results are further divided into subsections, each corresponding to a fixed statistic, for easier interpretation.

1 Tractable Likelihood MAE Plots
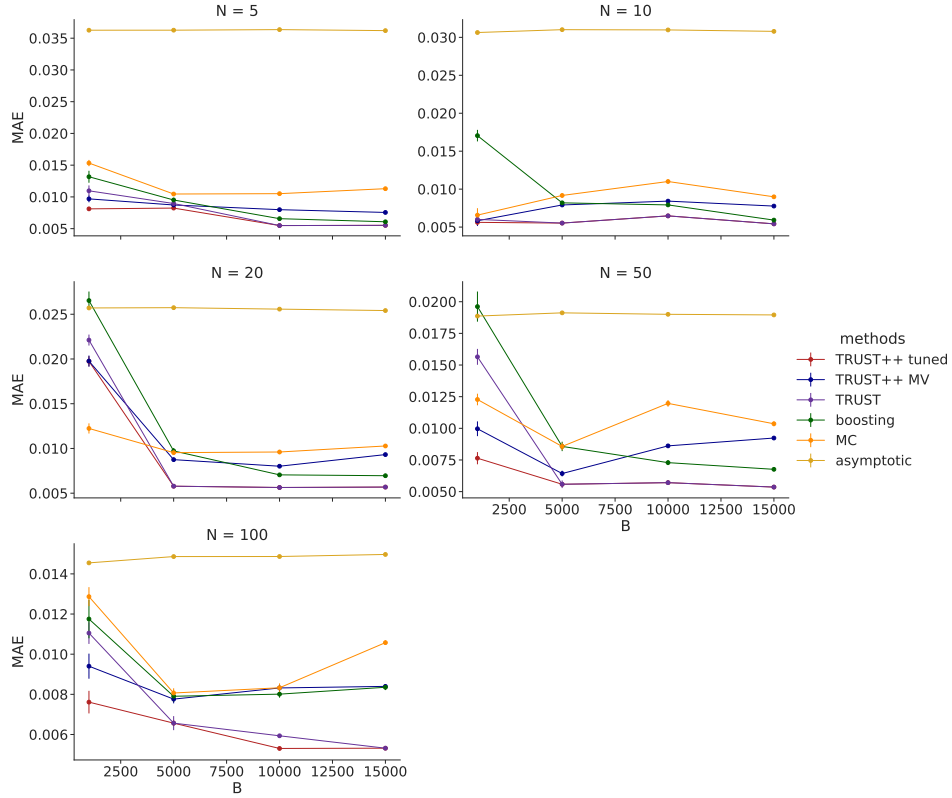
1.1 All KS MAE plots

Figure 1: MAE curve for 1d Normal model and KS statistic
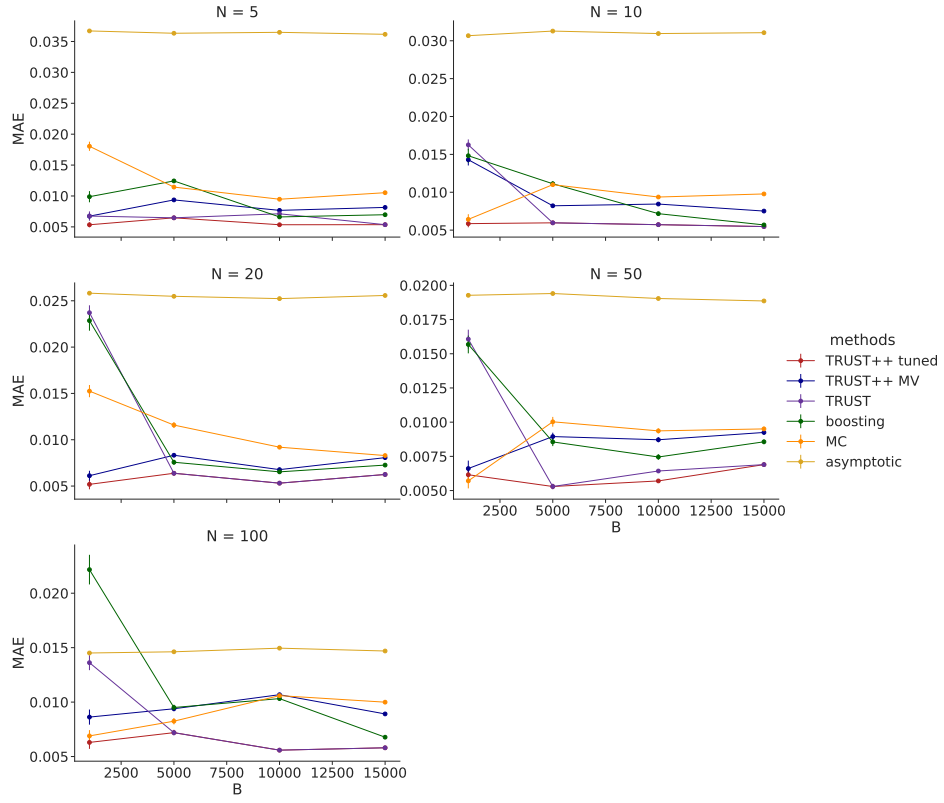
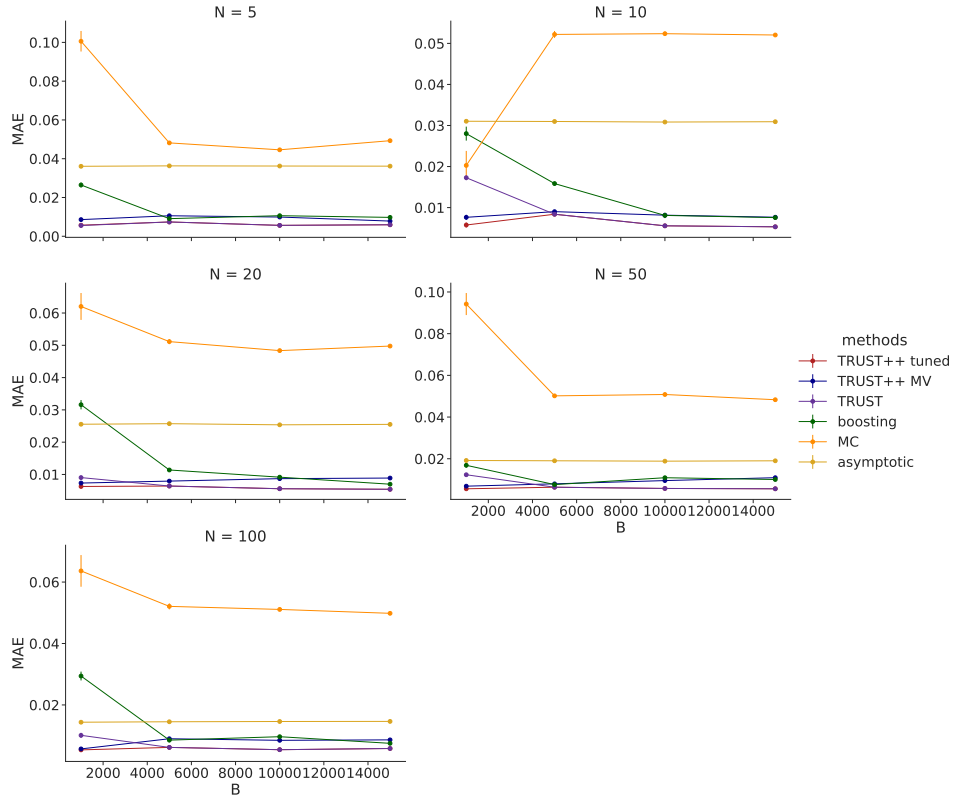Figure 2: MAE curve for gmm model and KS statistic

Figure 3: MAE curve for lognormal model and KS statistic
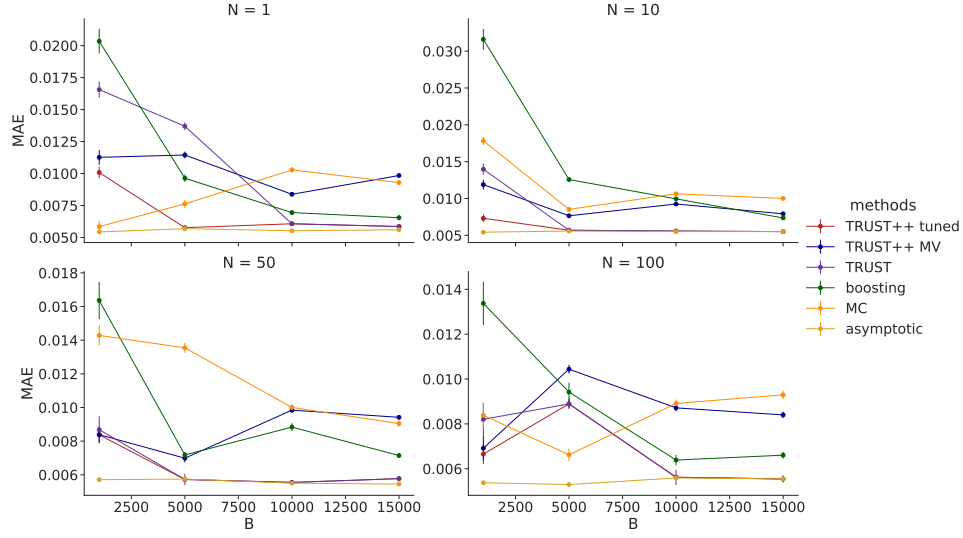
1.2 All LR MAE curves

Figure 4: MAE curve for 1d normal model and Likelihood Ratio statistic
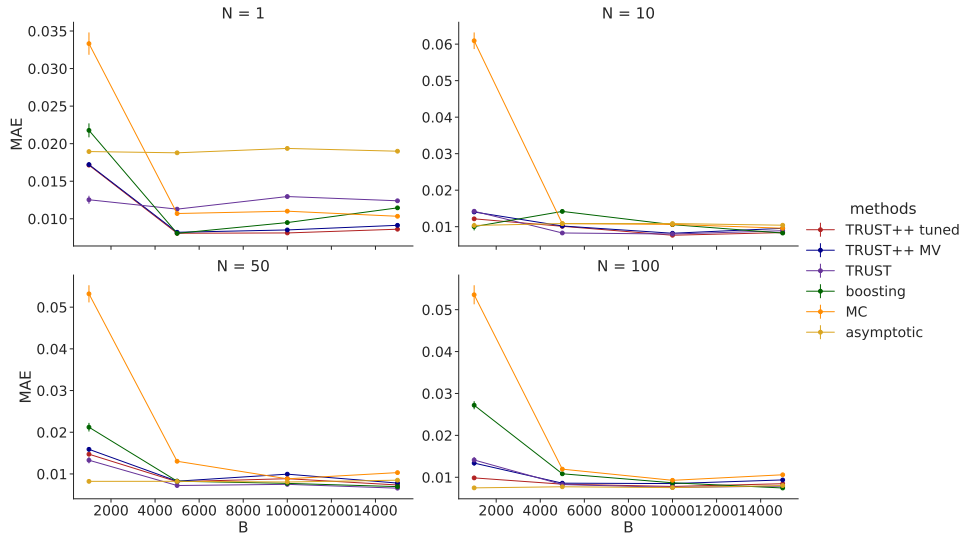

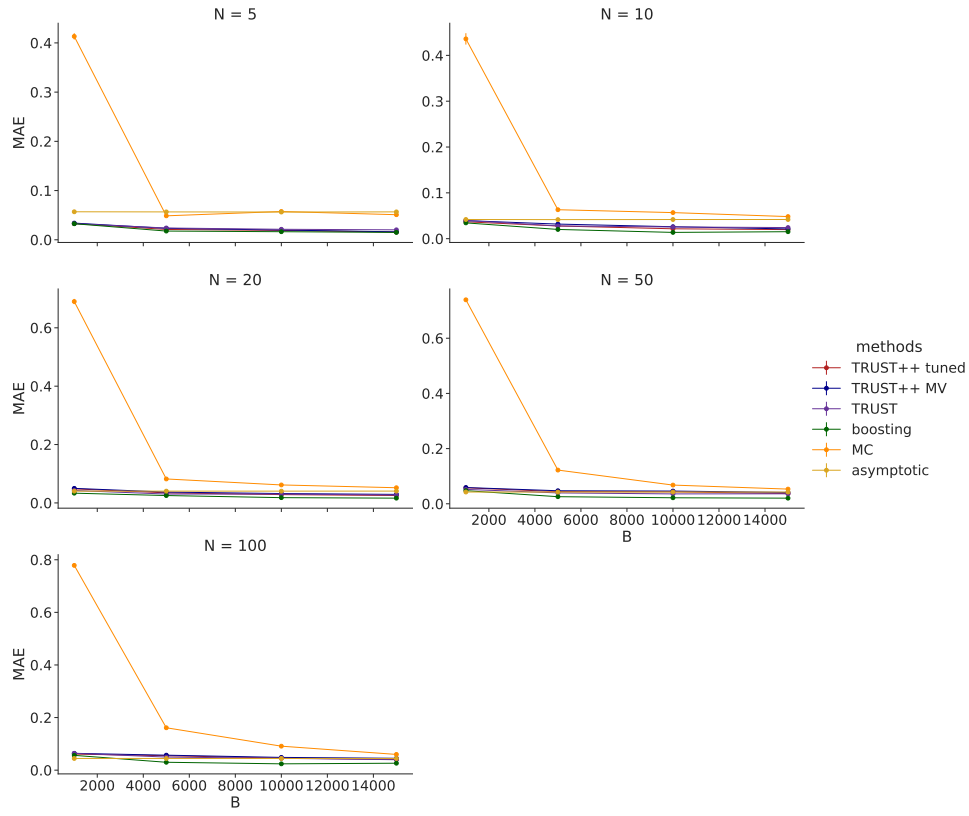Figure 5: MAE curve for gmm model and Likelihood Ratio statistic

Figure 6: MAE curve for lognormal model and Likelihood Ratio statistic
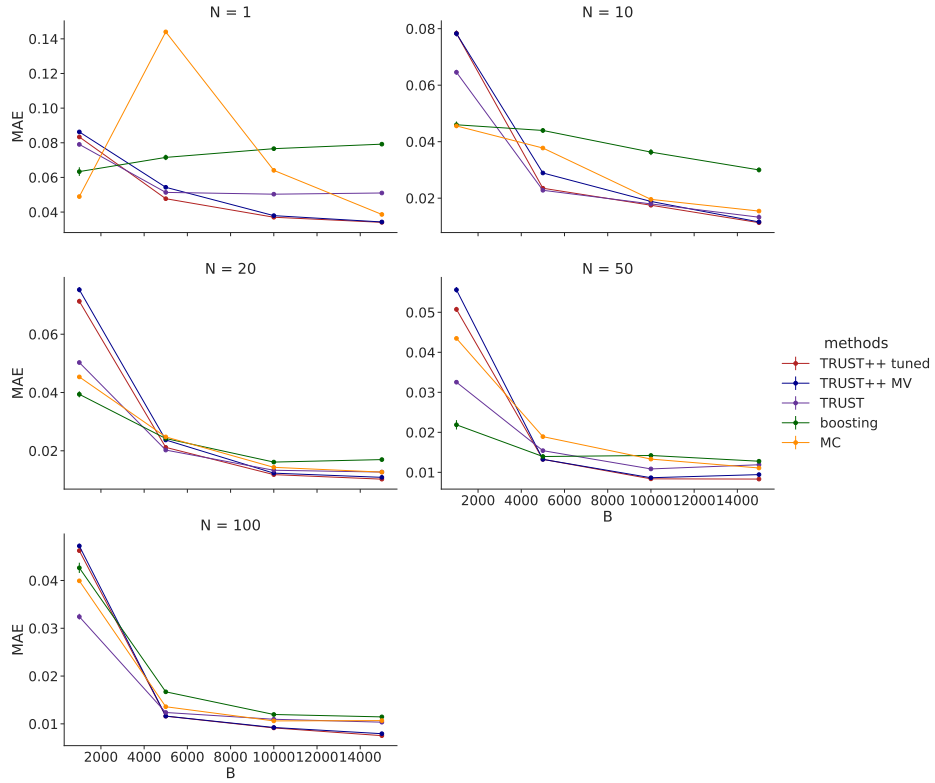
1.3 All BFF MAE curves

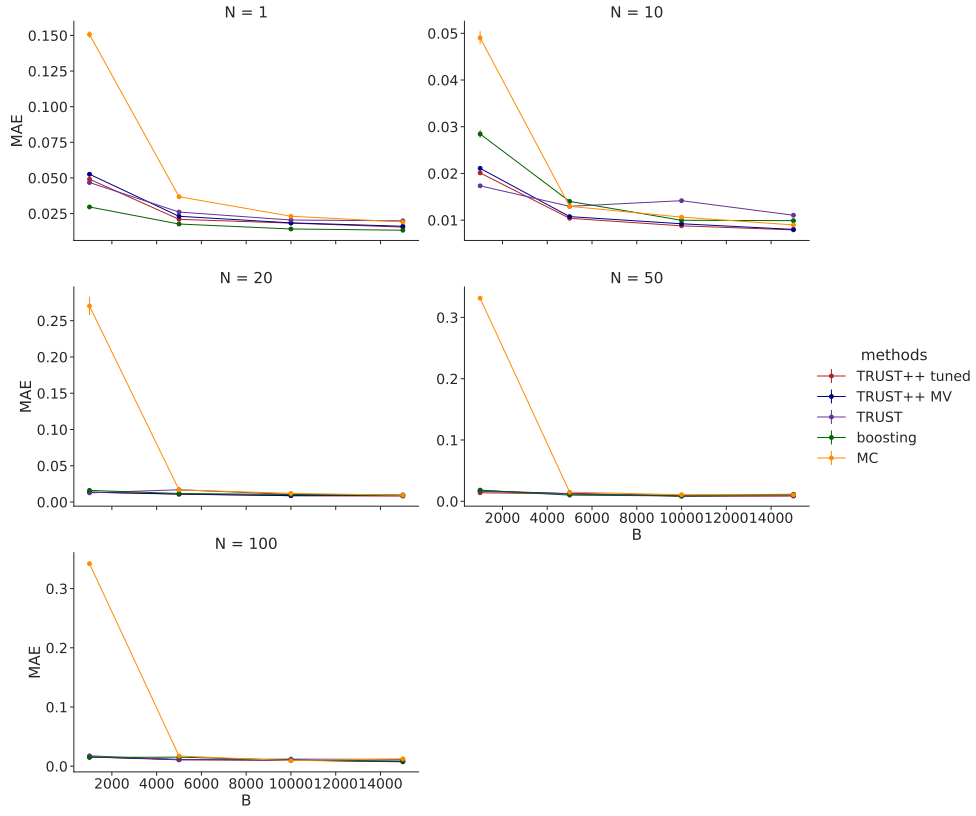Figure 7: MAE curve for 1d normal model and Bayes Frequentist Factor statistic

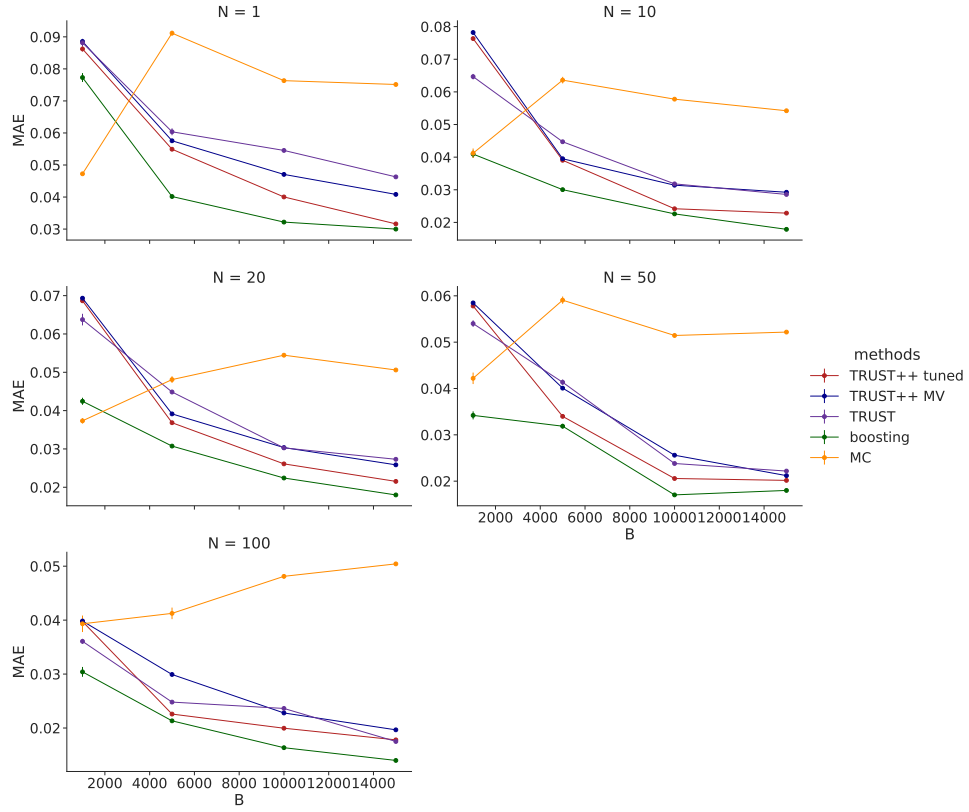Figure 8: MAE curve for 1d normal model and Bayes Frequentist Factor statistic

Figure 9: MAE curve for log-normal model and Bayes Frequentist Factor statistic
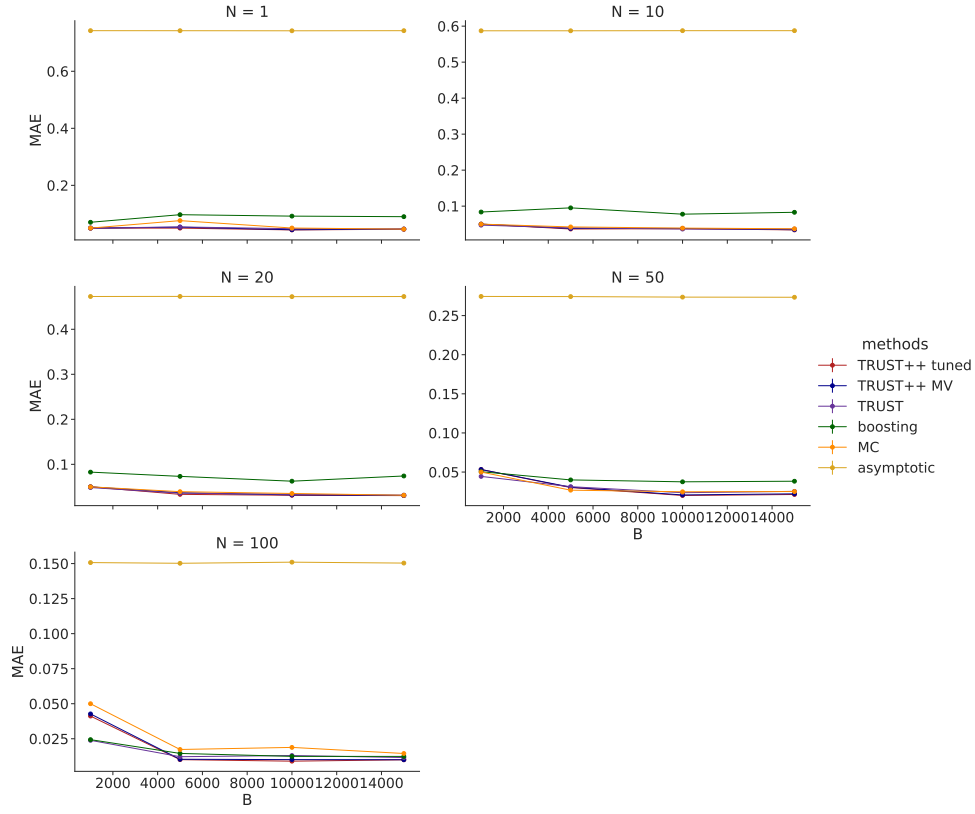
1.4 All e-value MAE curves

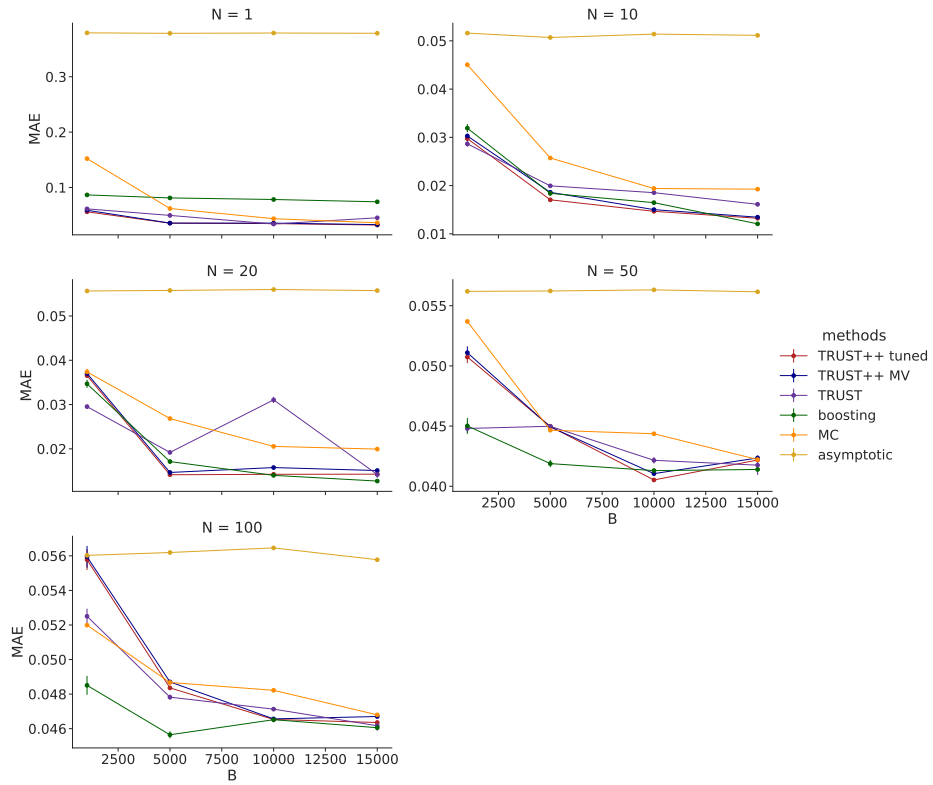Figure 10: MAE curve for 1d normal model and e-value statistic

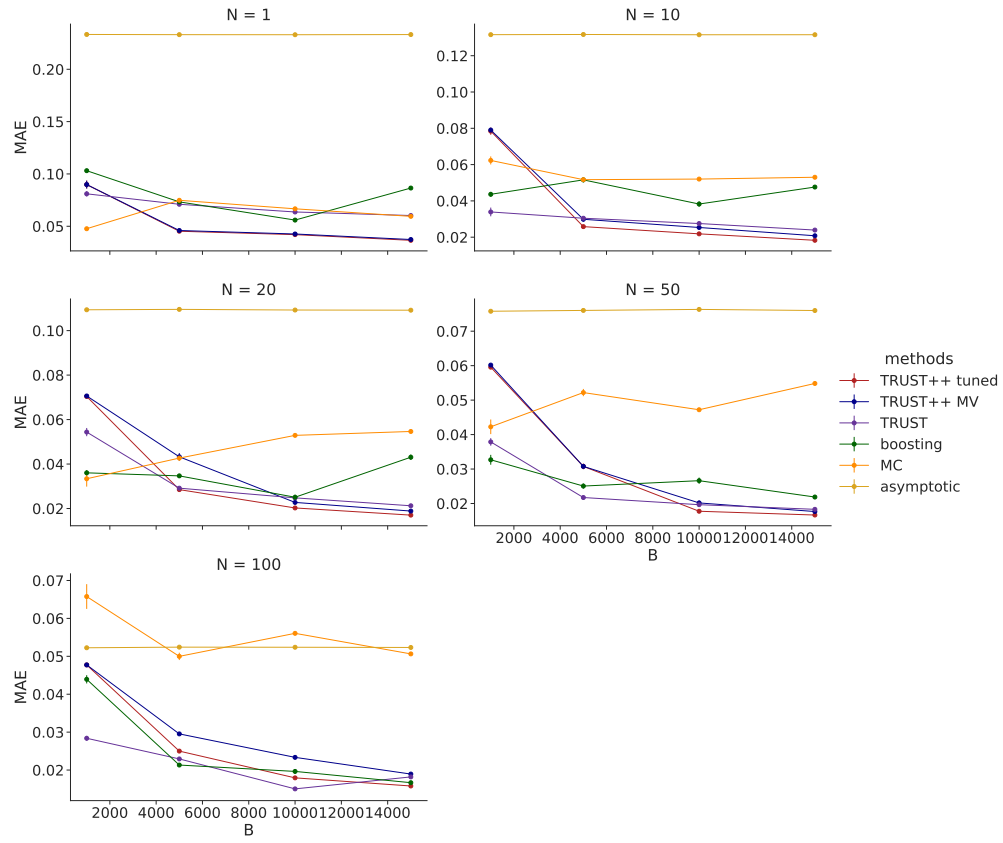

Figure 11: MAE curve for gmm model and e-value statistic

Figure 12: MAE curve for log-normal and e-value

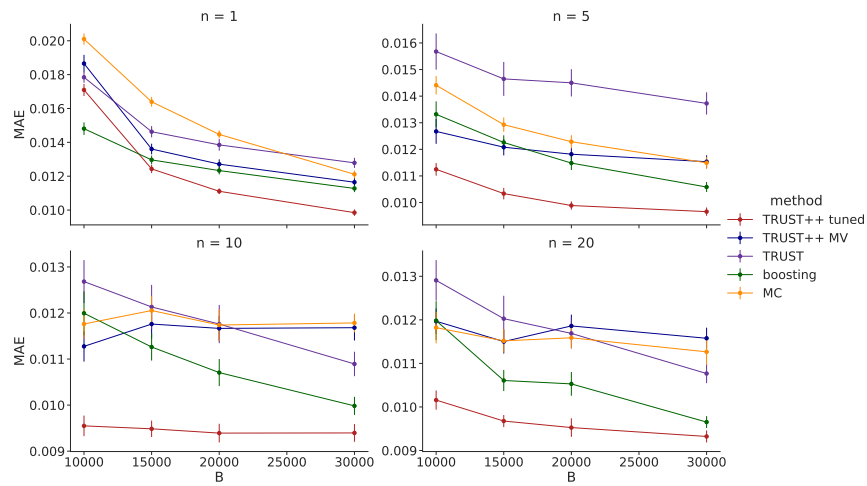# 2 Likelihood-free MAE plots

2.1 All Waldo MAE curve

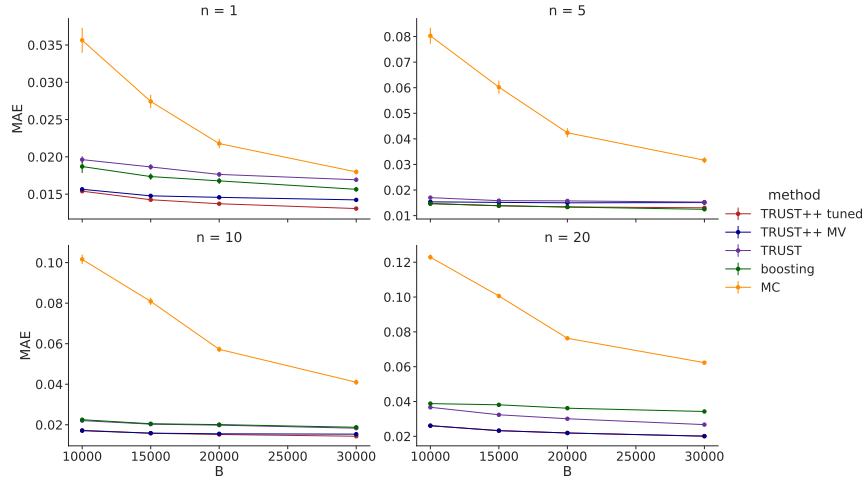Figure 13: MAE curve for weinberg simulator and waldo statistic

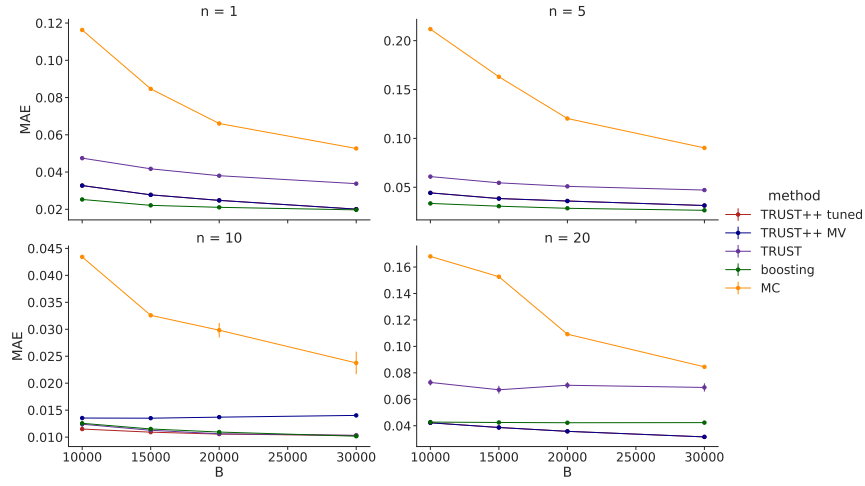Figure 14: MAE curve for two moons model and waldo statistic

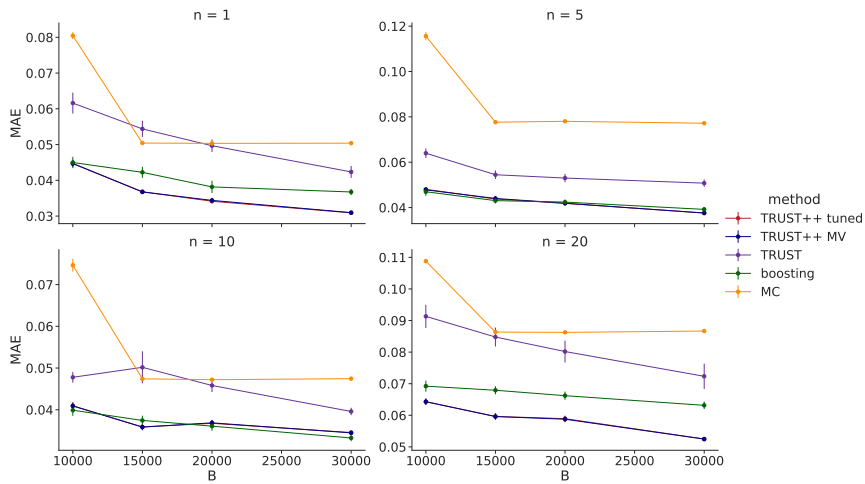Figure 15: MAE curves for SIR model and waldo statistic

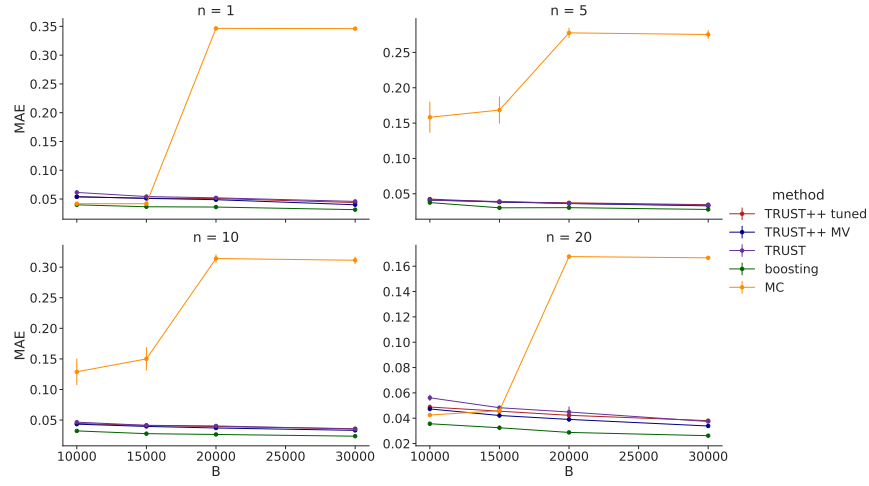Figure 16: MAE curves for MG1 model and waldo statistic

Figure 17: MAE curves for tractable simulator and waldo statistic
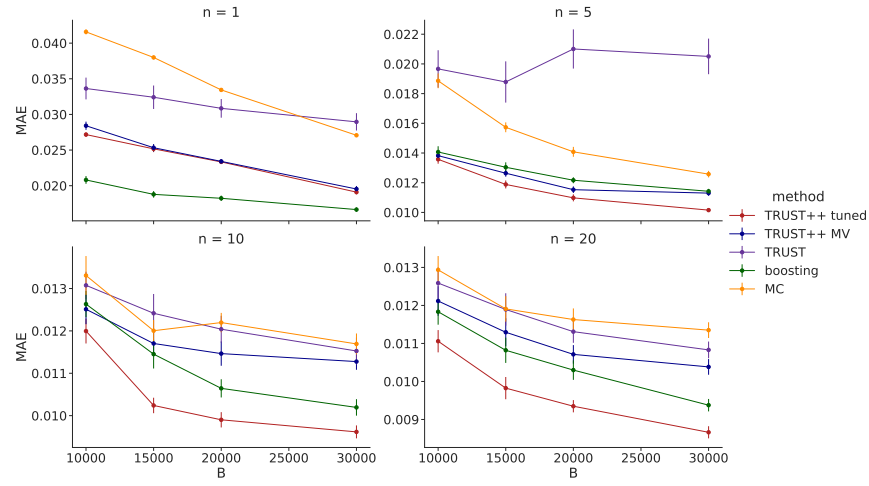
2.2 All BFF curve

Figure 18: MAE curves for weinberg simulator and Bayes Frequentist Factor statistic
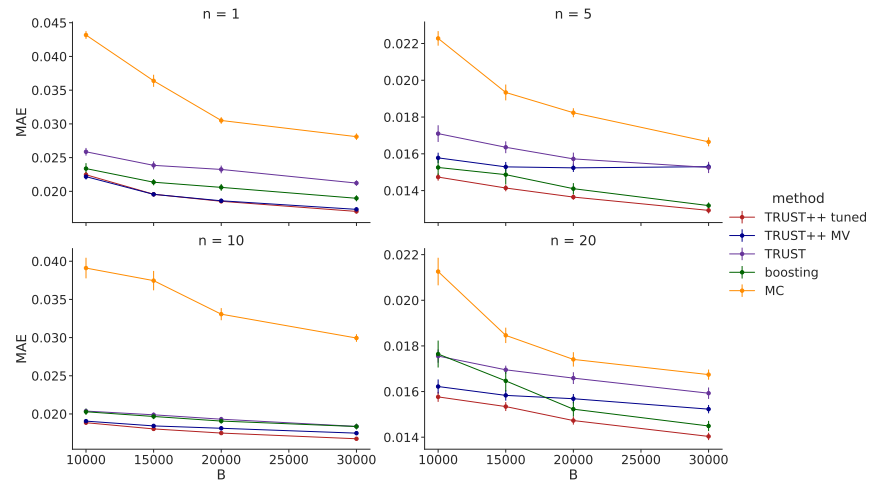

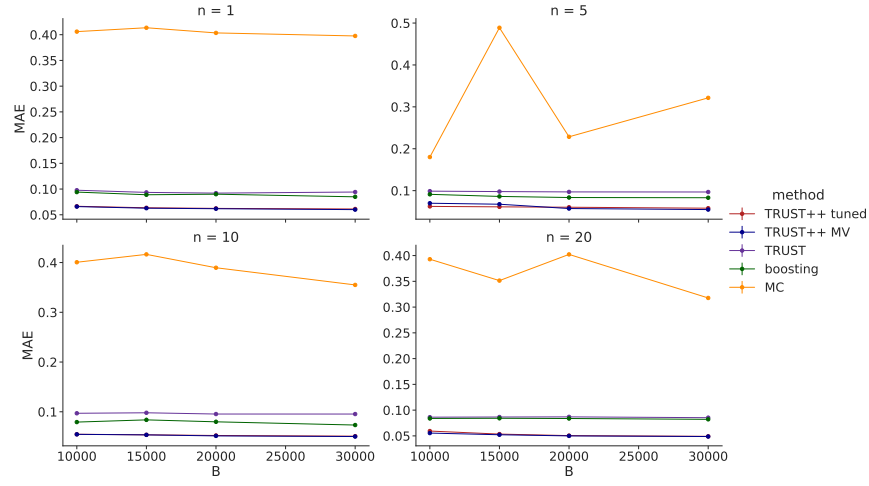Figure 19: MAE curves for two moons model and two moons statistic

Figure 20: MAE curves for SIR simulator and bayes frequentist factor statistic
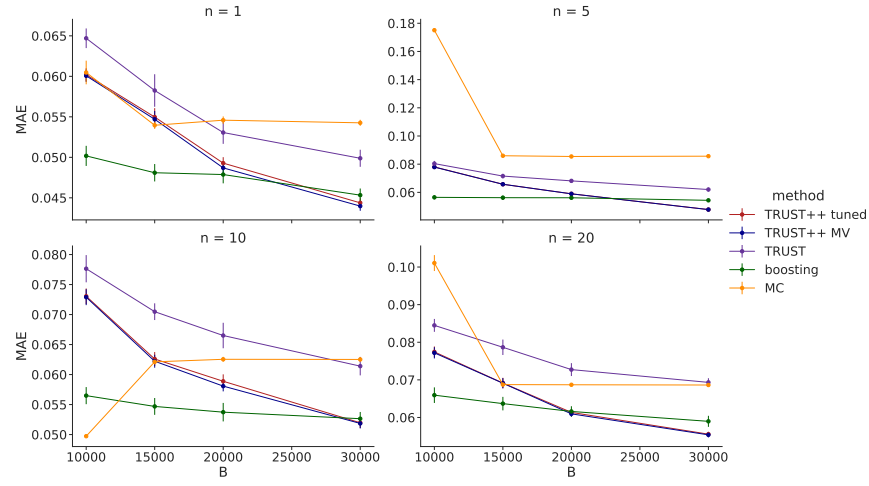

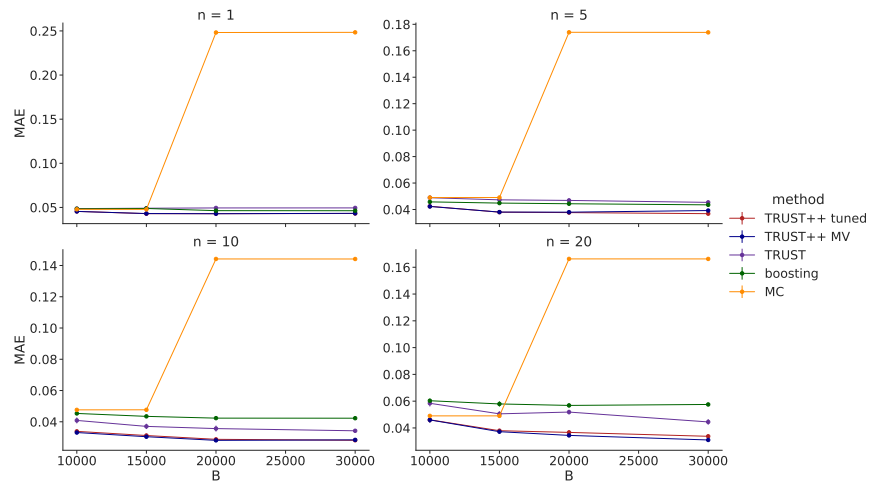Figure 21: MAE curves for mg1 model and bayes frequentist factor statistic

Figure 22: MAE curves for tractable simulator and bayes frequentist factor statistic
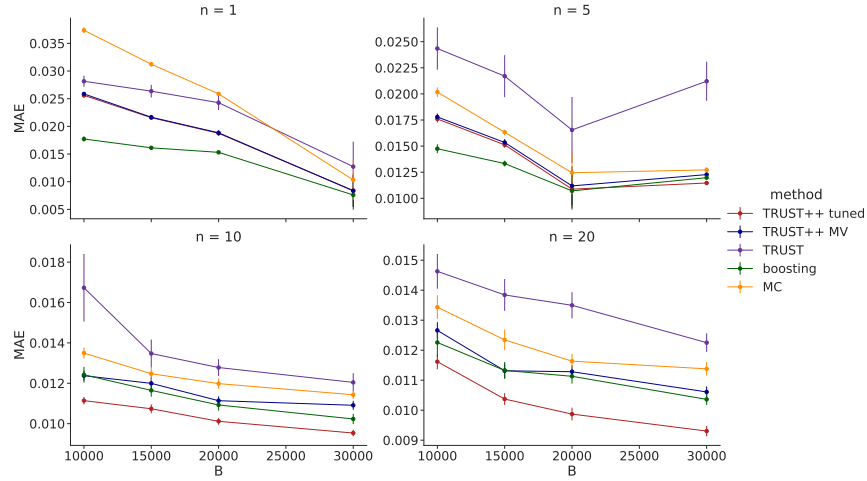
2.3 All e-value MAE curves

Figure 23: MAE curves for weinberg simulator and e-value statistic
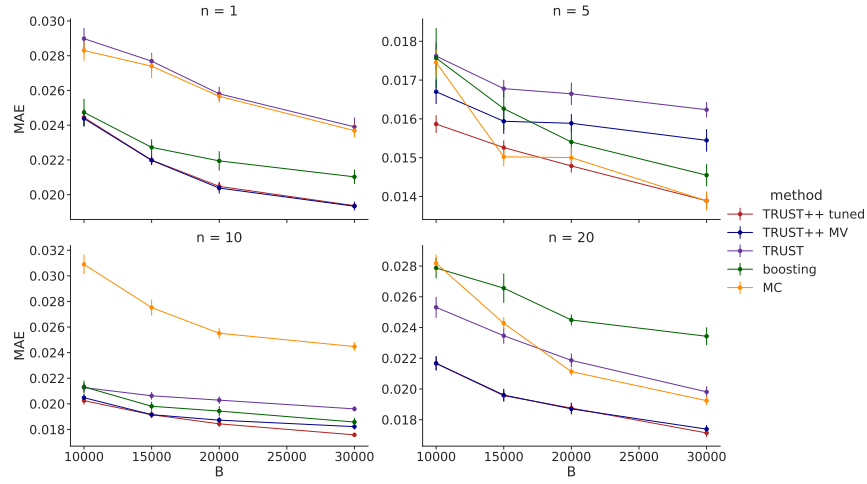

Figure 24: MAE curves for two-moons model and e-value statistic
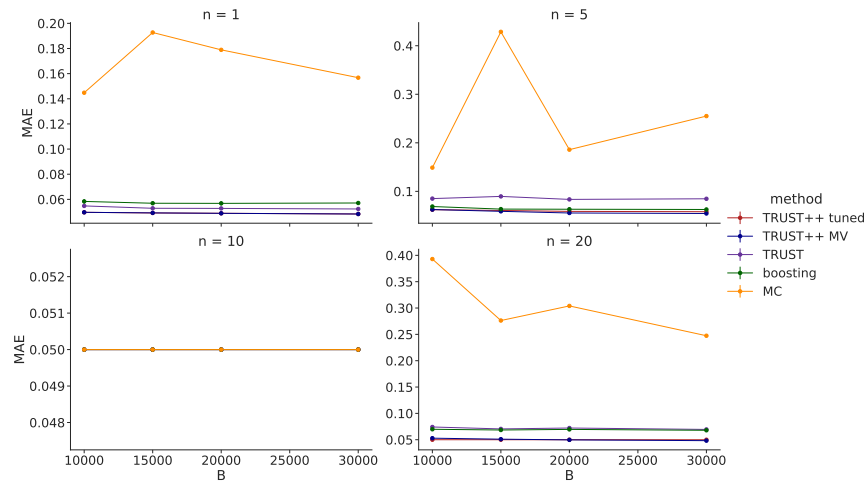

Figure 25: MAE curves for SIR simulator and e-value statistic
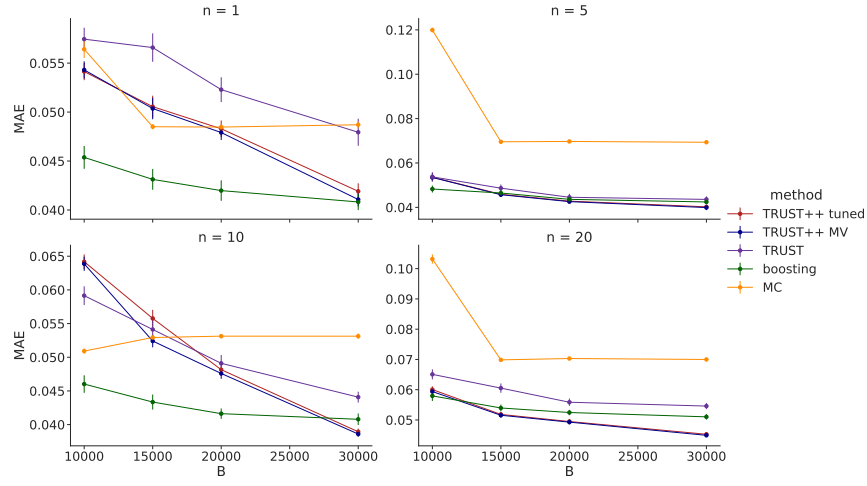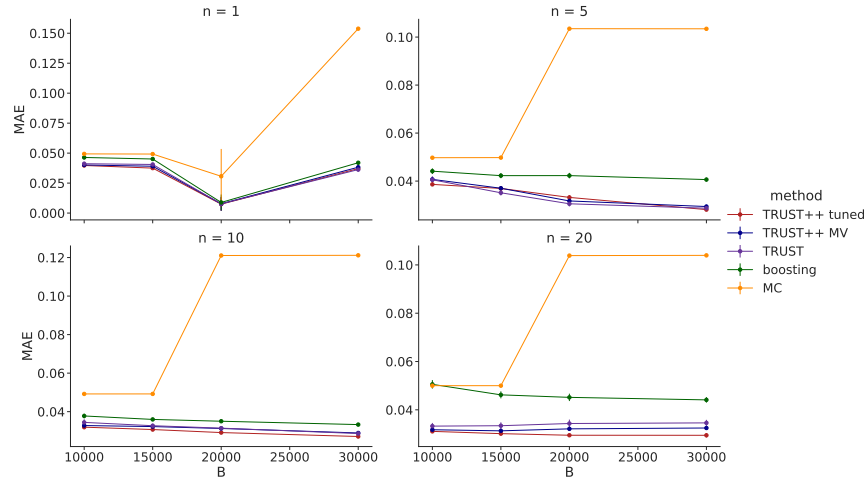
Figure 26: MAE curves for mg1 model and e-value statistic

Figure 27: MAE curves for tractable simulator and e-value statistic