Supplementary Materials for "A Graphical Comparison of Screening Designs using Support Recovery Probabilities"

Kade Young[1], Maria L. Weese[2], Jonathan W. Stallrich[1], Byran J. Smucker[3], and David J. Edwards[4]

[1]Department of Statistics, North Carolina State University, Raleigh, NC
[2]Department of Information Systems & Analytics, Miami University, Oxford, OH
[3]Department of Statistics, Miami University, Oxford, OH
[4]Department of Statistical Sciences and Operations Research, Virginia Commonwealth University, Richmond, VA

1 Inconsistencies in SSDs

This section provides the remainder of the simulations scenarios based on Marley and Woods (2010) from Section 2.1 in the main paper.

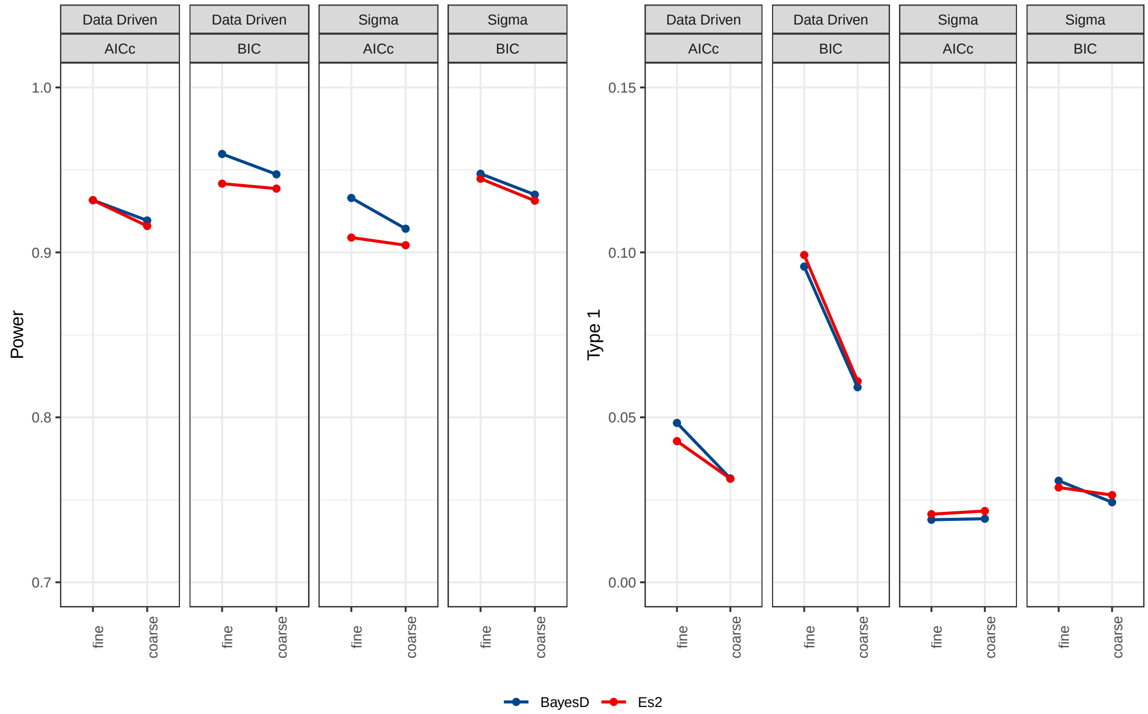

Figure 1: Comparison of (12, 26) Bayesian D-optimal and balanced $E(s^2)$ SSD for 3 active effects with an effect magnitude of 5 for unknown signs.

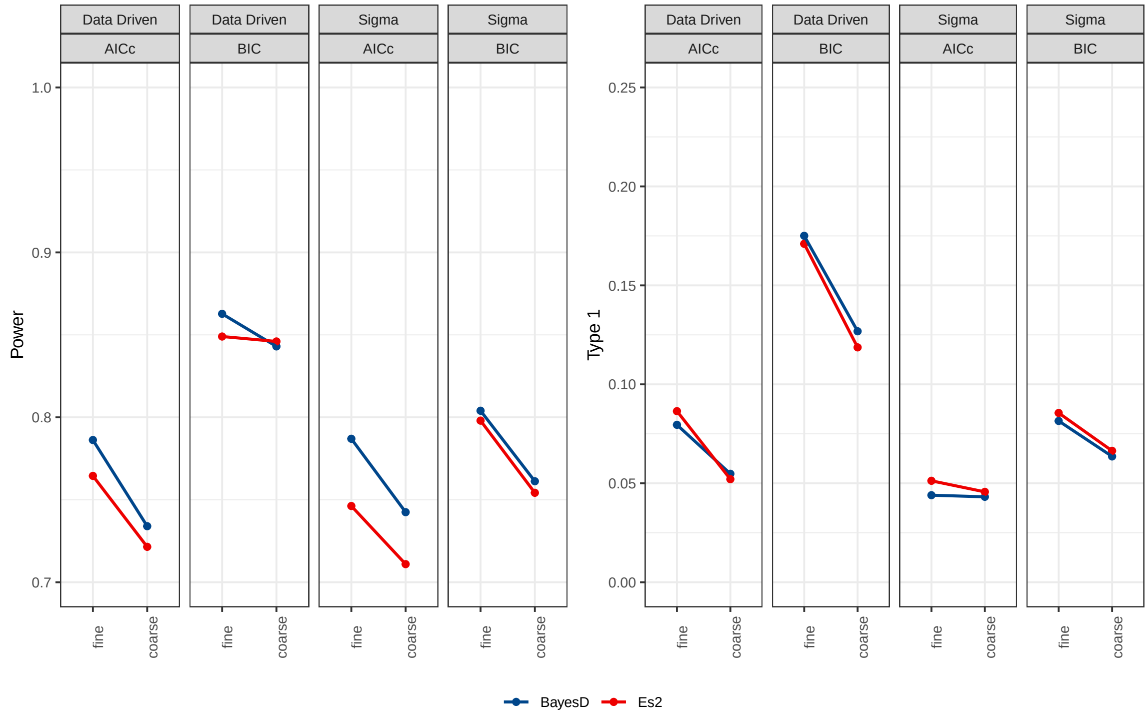

Figure 2: Comparison of (12, 26) Bayesian D-optimal and balanced $E(s^2)$ SSD for 4 active effects with an effect magnitude of 4 for unknown signs.

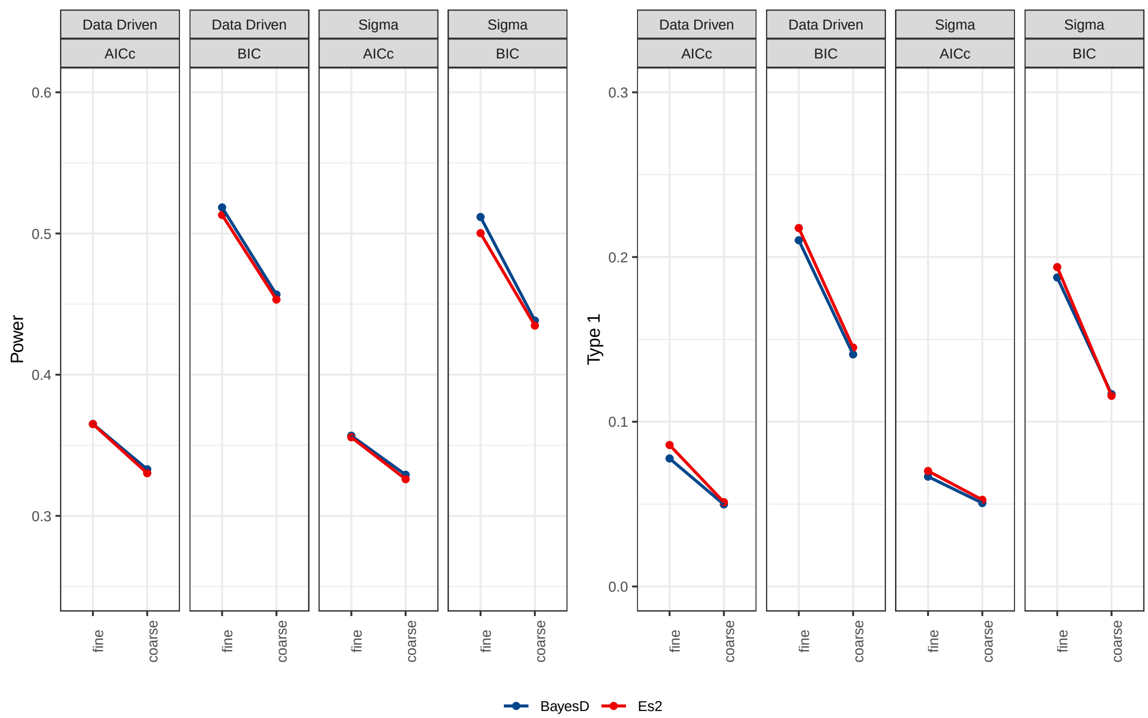

Figure 3: Comparison of (12, 26) Bayesian D-optimal and balanced $E(s^2)$ SSD for 9 active effects with an effect magnitude of the following values: 10,8,5,3 and the rest 2 for unknown signs.

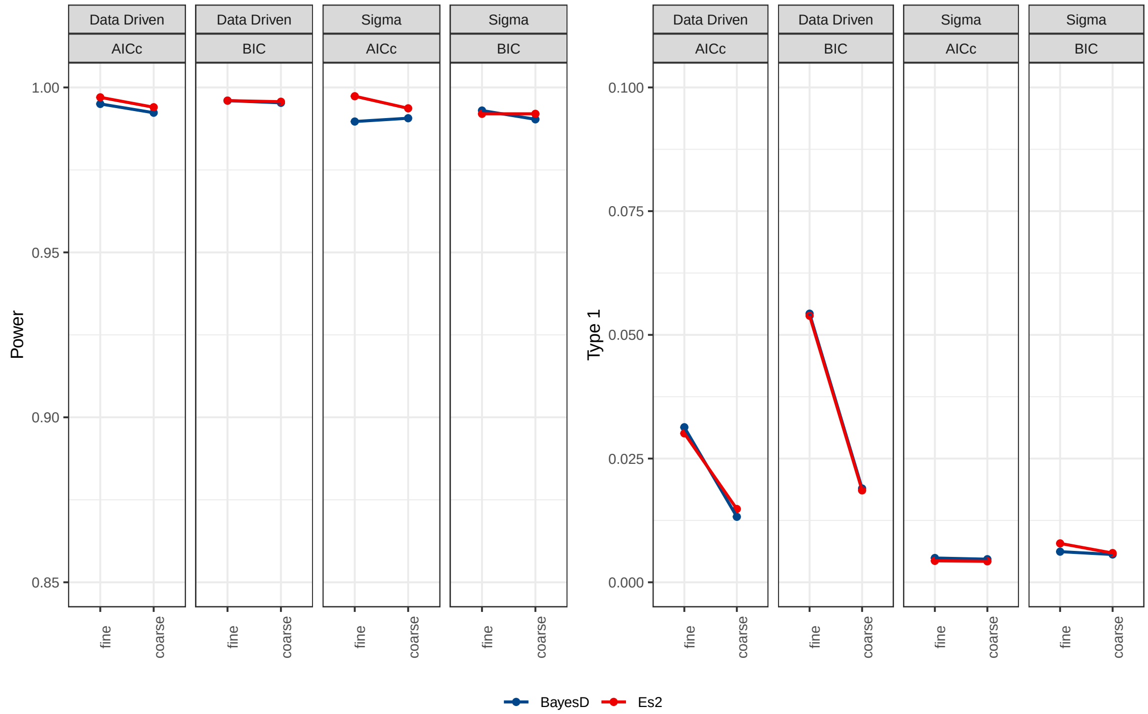

Figure 4: Comparison of (14, 24) Bayesian D-optimal and balanced $E(s^2)$ SSD for 3 active effects with an effect magnitude of 5 for both known and unknown signs.

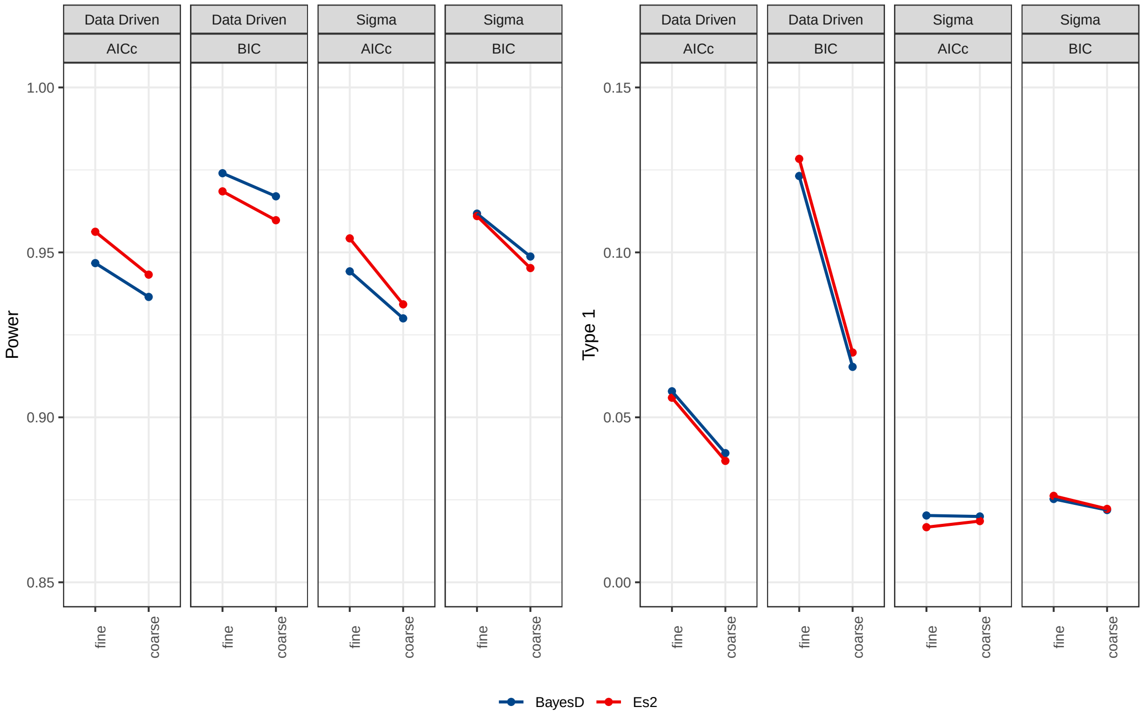

Figure 5: Comparison of (14, 24) Bayesian D-optimal and balanced $E(s^2)$ SSD for 4 active effects with an effect magnitude of 4 for both known and unknown signs.

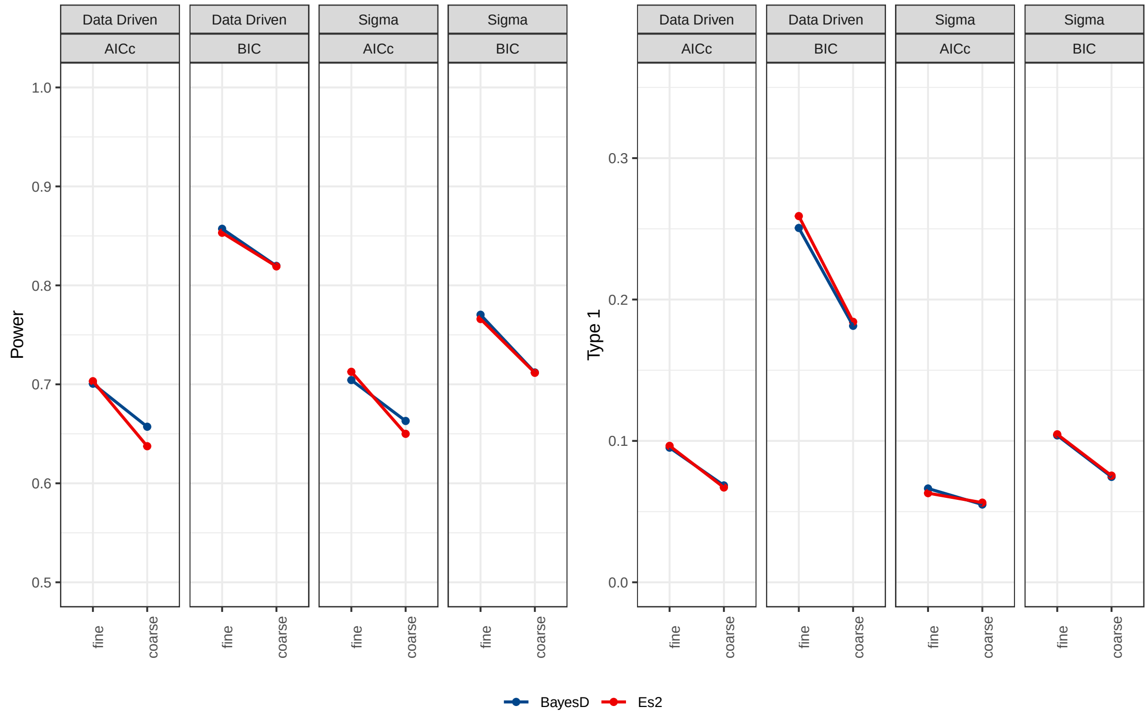

Figure 6: Comparison of (14, 24) Bayesian D-optimal and balanced $E(s^2)$ SSD for 6 active effects with an effect magnitude of 3 for both known and unknown signs.

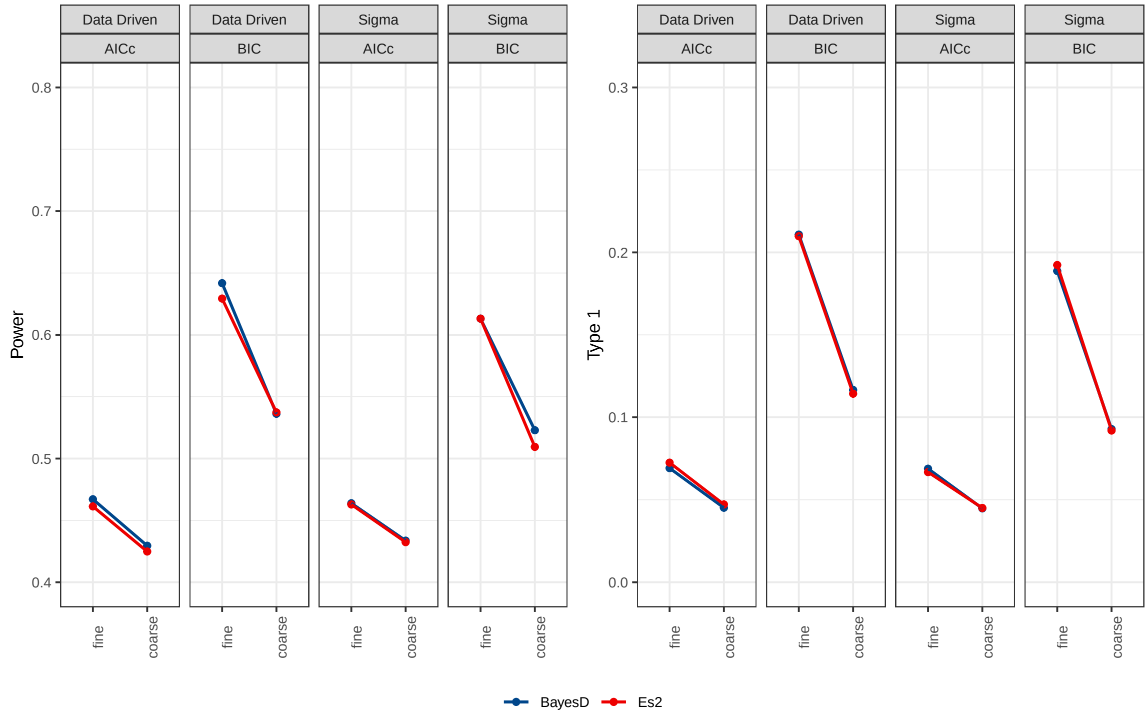

Figure 7: Comparison of (14, 24) Bayesian D-optimal and balanced $E(s^2)$ SSD for 9 active effects with an effect magnitude of the following values: 10,8,5,3 and the rest 2 for unknown signs.

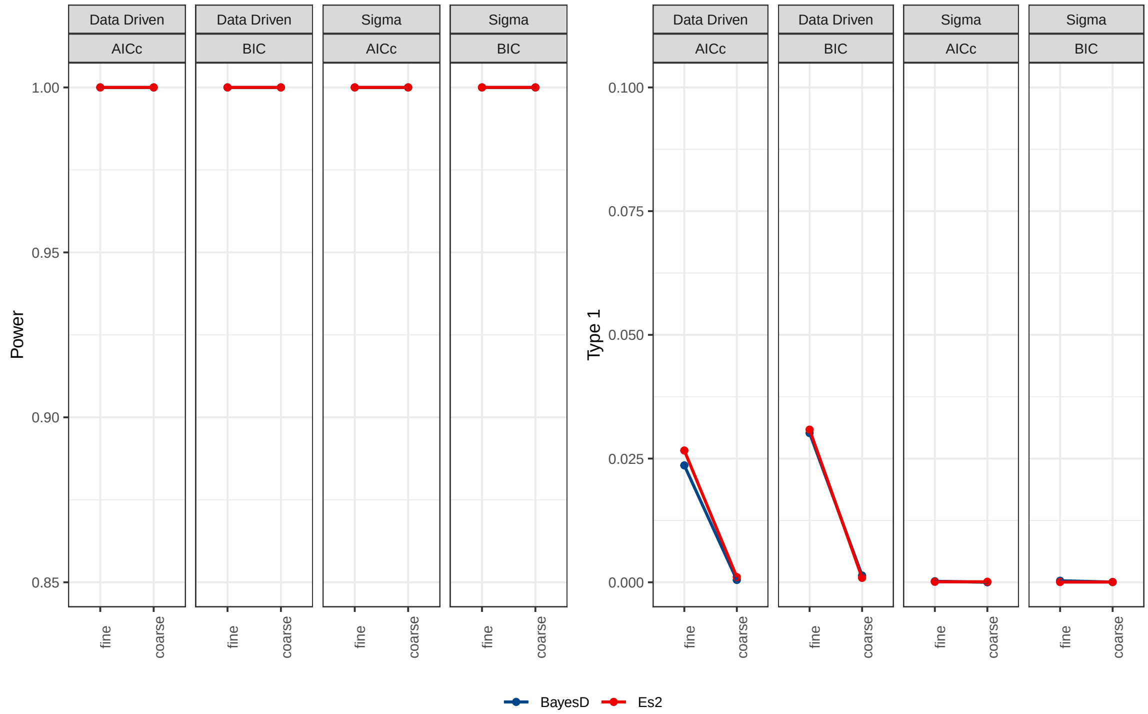

Figure 8: Comparison of (18, 22) Bayesian D-optimal and balanced $E(s^2)$ SSD for 3 active effects with an effect magnitude of 5 for unknown signs.

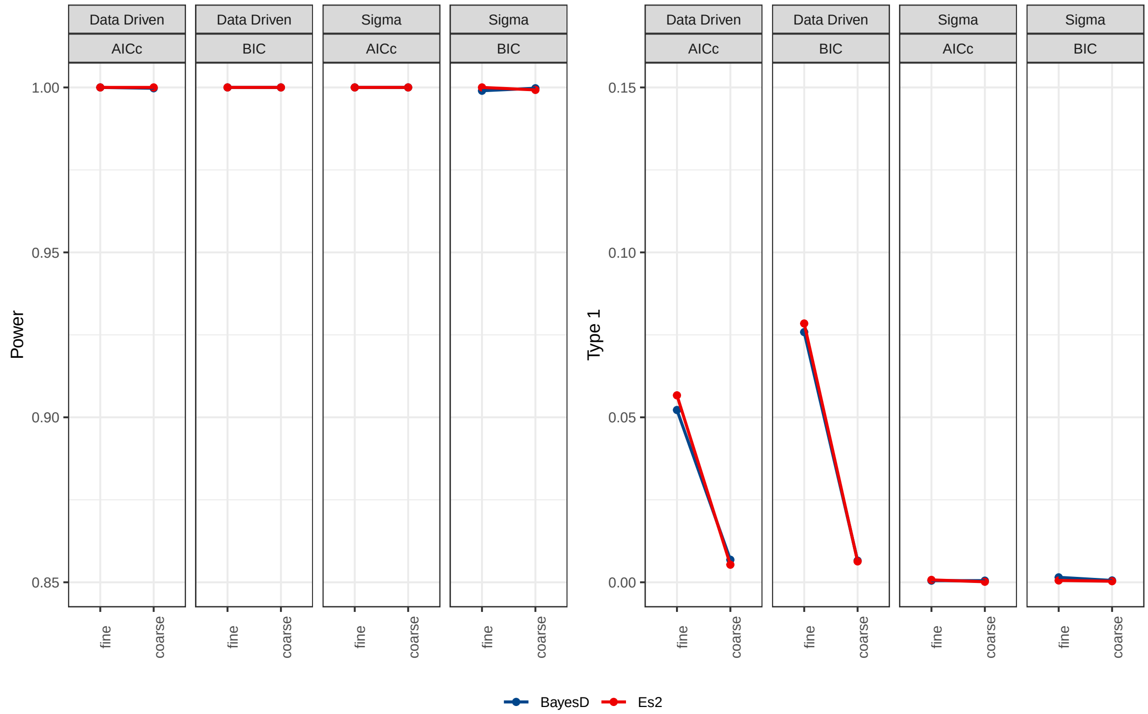

Figure 9: Comparison of (18, 22) Bayesian D-optimal and balanced $E(s^2)$ SSD for 4 active effects with an effect magnitude of 4 for both unknown signs.

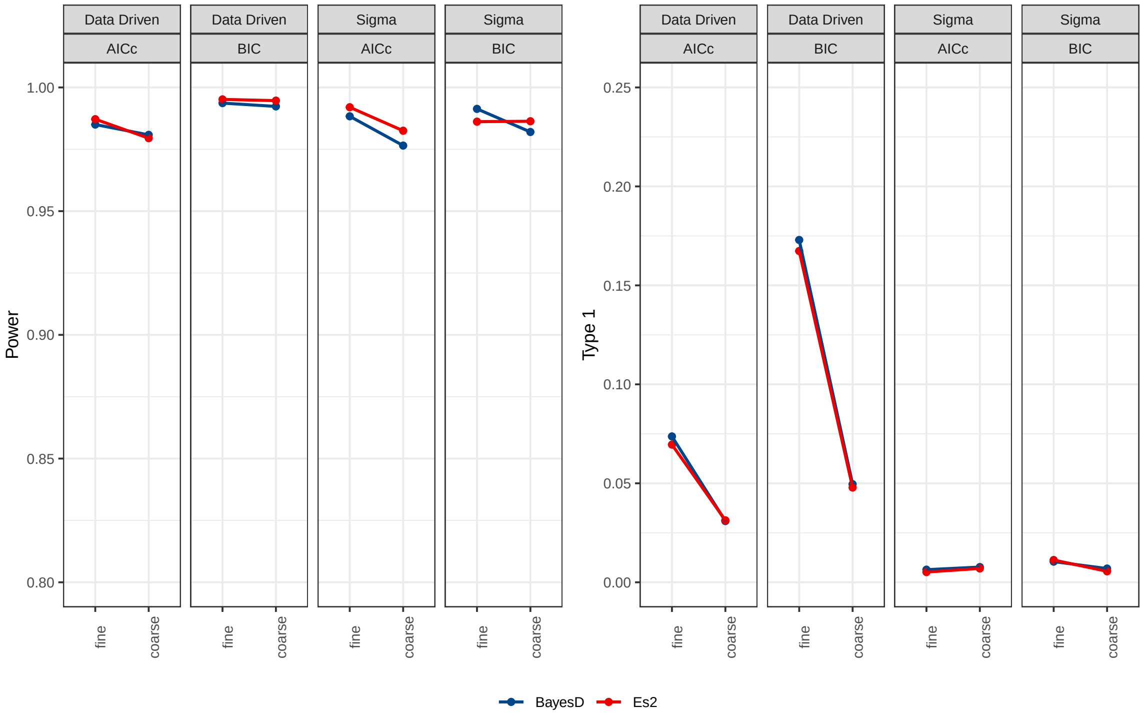

Figure 10: Comparison of (18, 22) Bayesian D-optimal and balanced $E(s^2)$ SSD for 6 active effects with an effect magnitude of 3 for both unknown signs.

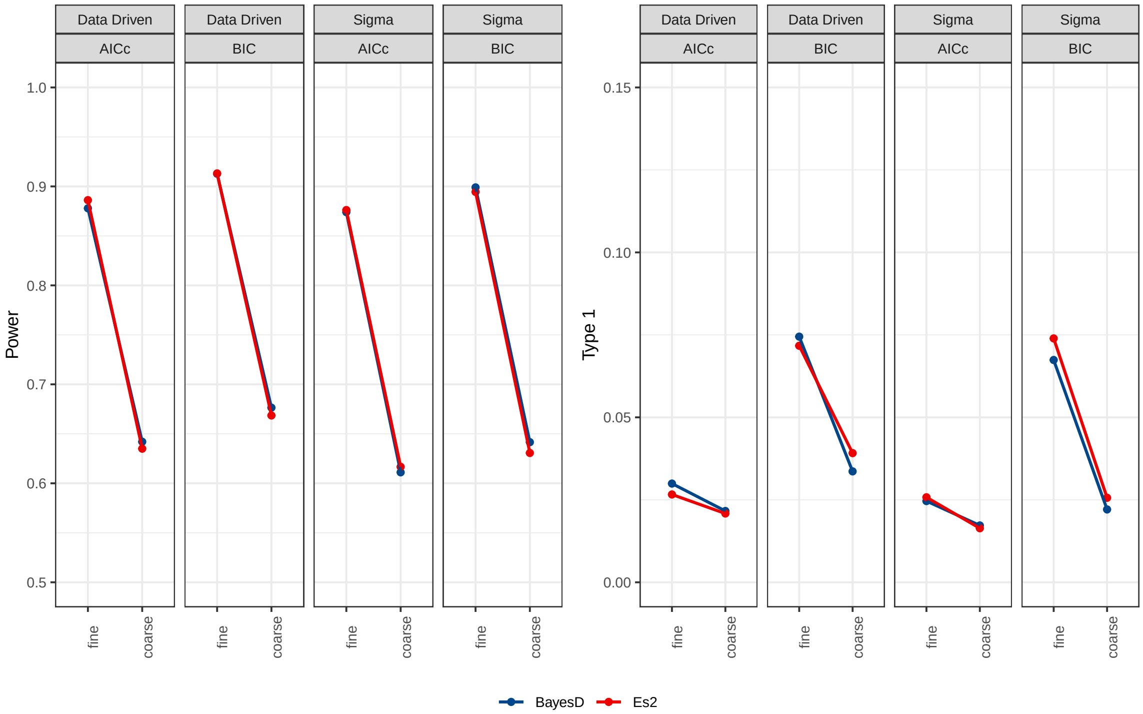

Figure 11: Comparison of (18, 22) Bayesian D-optimal and balanced $E(s^2)$ SSD for 9 active effects with an effect magnitude of the following values: 10,8,5,3 and the rest 2 for unknown signs.

2 Inconsistencies in Larger Screening Designs

This section contains the remainder of the simulations from Mee et al. (2017) not shown in Section 2.2 of the main paper. These are the complete simulation results for the $n = 20$, $k = 7$ designs.

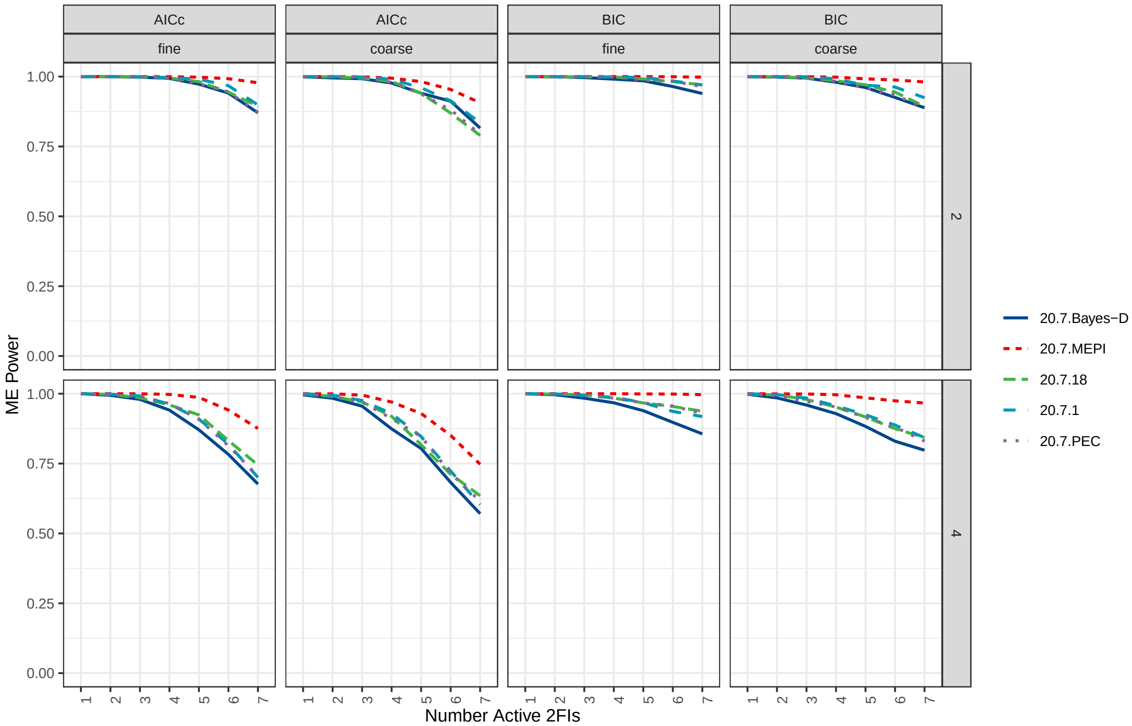

Figure 12: Power of the main effects for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects equivalent in magnitude to the main effects.

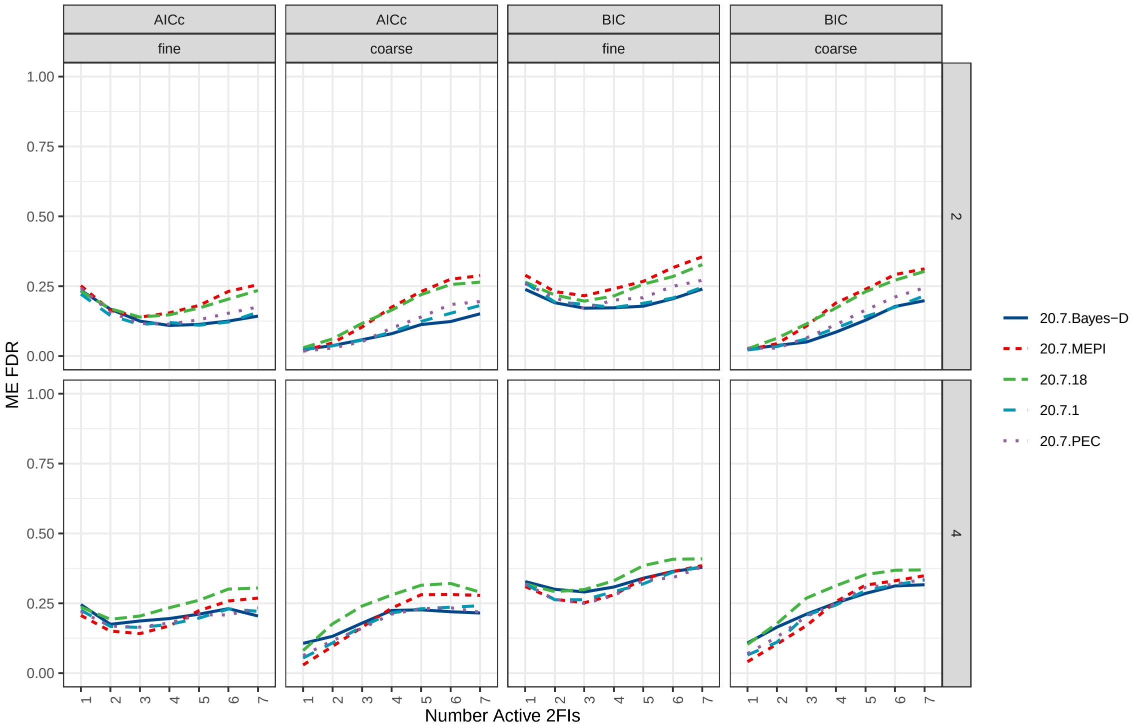

Figure 13: FDR of the main effects for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects equivalent in magnitude to the main effects. This figure corresponds to Figure 6 in the main document.

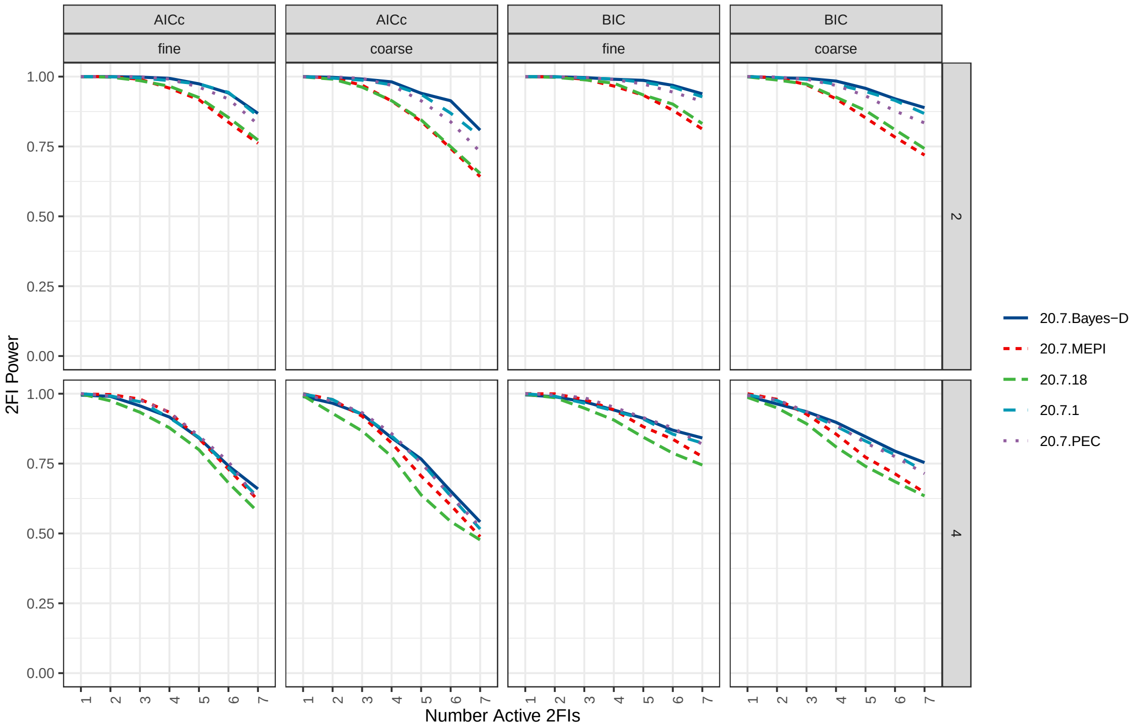

Figure 14: FDR of the main effects for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects equivalent in magnitude to the main effects. This figure corresponds to Figure 6 in the main document.

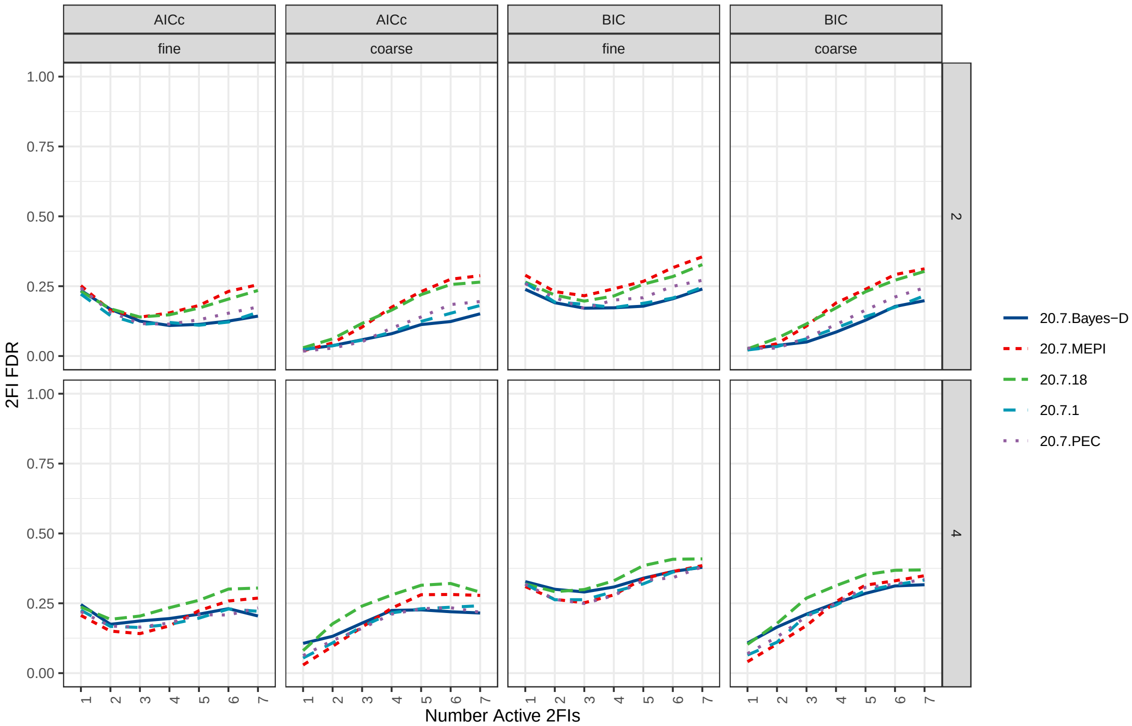

Figure 15: FDR of the two-factor interactions for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects equivalent in magnitude to the main effects. This figure corresponds to Figure 5 in the main document.

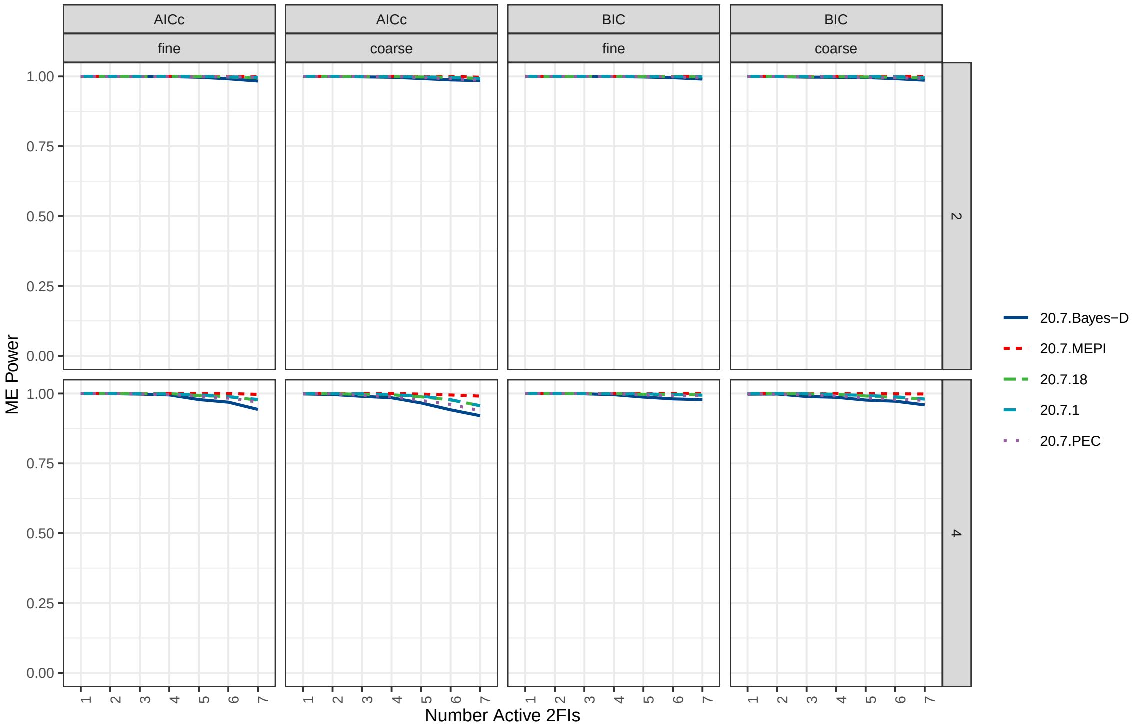

Figure 16: Power of the main effects for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects smaller in magnitude to the main effects.

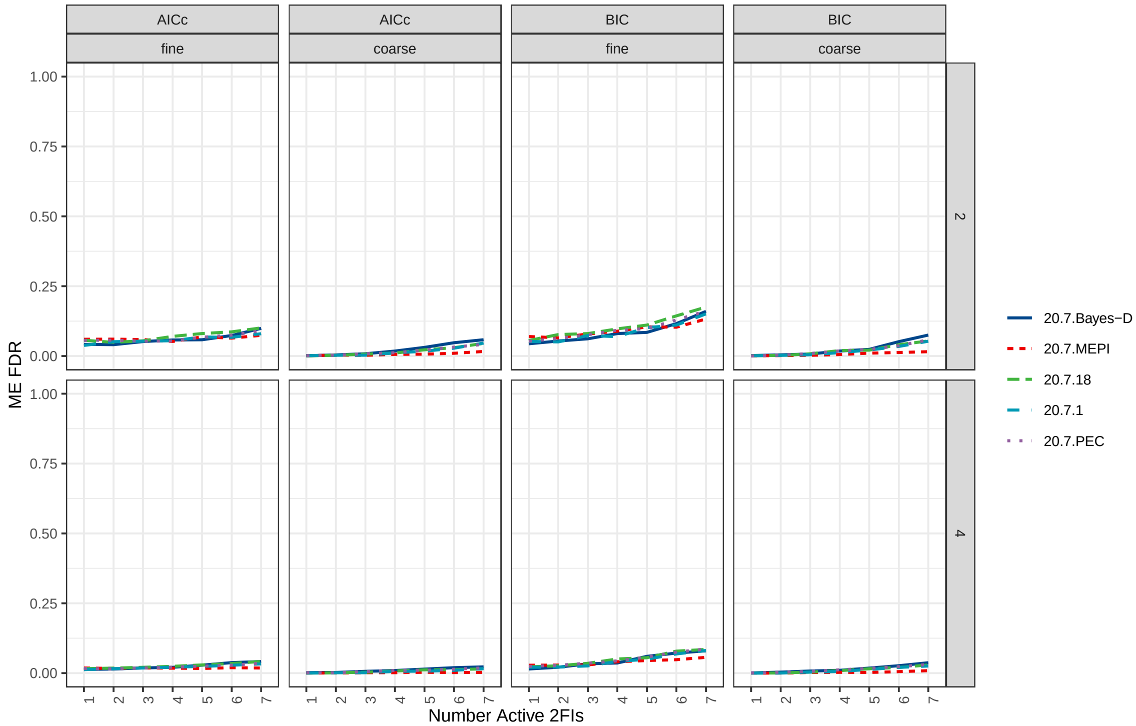

Figure 17: FDR of the main effects for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects smaller in magnitude to the main effects.

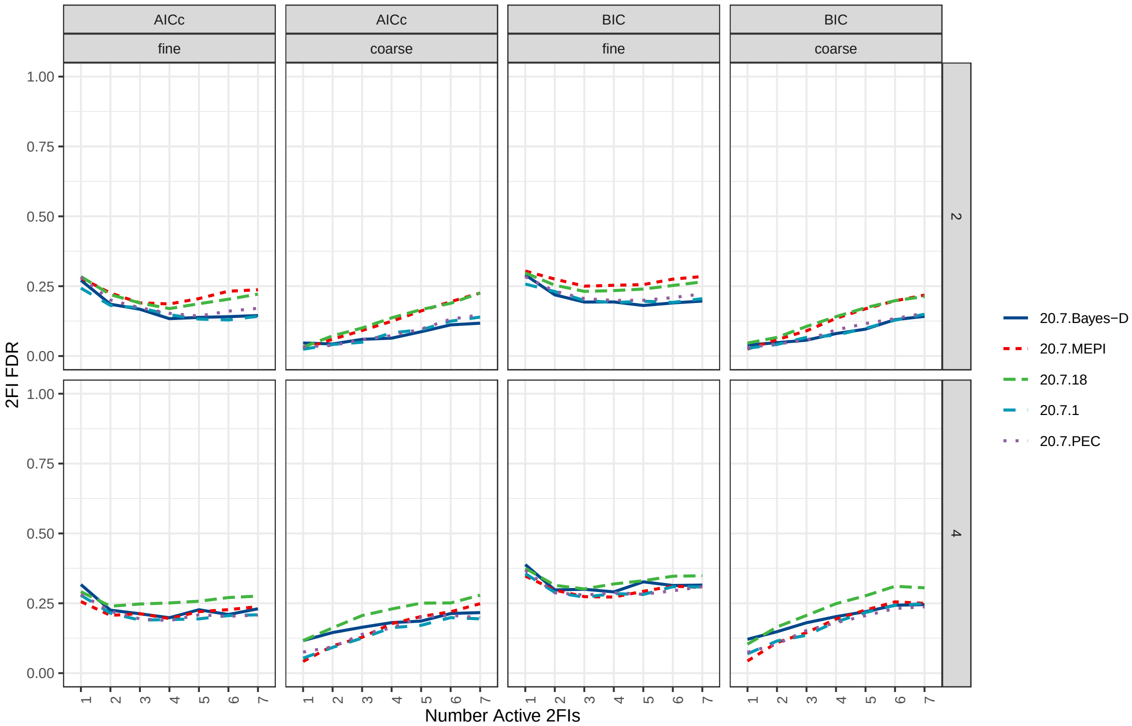

Figure 18: FDR of the two-factor interactions for 2 or 4 (top row and bottom row, respectively) active two-factor interactions with true effects smaller in magnitude to the main effects.

3 Lasso Sign Recovery Probabilities

Sign recovery at a fixed λ involves two events of normal random variables, denoted $\mathcal{I}_\lambda$ and $\mathcal{S}_\lambda$. The $\mathcal{I}_\lambda$ event represents the event of estimating all inactive effects as 0. The $\mathcal{S}_\lambda$ event represents the event of selecting all truly active factors and estimating them with the correct sign. Both events must occur for sign recovery under the lasso. These two events are independent and therefore the probability of sign recovery under the lasso for a fixed $\mathbf{X}$ and $\boldsymbol{\beta}$ with sign vector $\boldsymbol{z}$ is defined as:

$$\phi_\lambda(\mathbf{X}|\boldsymbol{\beta}) = P(\hat{\boldsymbol{z}} = \boldsymbol{z}|\mathbf{X}, \boldsymbol{\beta}) = P[\mathcal{I}_\lambda \cap \mathcal{S}_\lambda|\mathbf{X}, \boldsymbol{\beta}] = P[\mathcal{I}_\lambda|\mathbf{X}, \boldsymbol{z}] \times P[\mathcal{S}_\lambda|\mathbf{X}, \boldsymbol{\beta}], \quad (1)$$

where $\hat{\boldsymbol{z}}$ is the estimated sign vector. The computation of this probability requires knowledge of λ and $\boldsymbol{\beta}$ which also implicitly requires knowledge of the set of active effects, $\mathcal{A}$. To handle these issues, Stallrich et al. (2023) evaluate ϕ_λ over all possible active sets of size a (or, in some cases, a subset of all possible active sets) and sign vectors for a range of λ values. To make the computations even faster, they assumed the active effects had a constant absolute magnitude of β, being a user-specified minimum signal-to-noise ratio to detect.

If the sign vector of the active effects is assumed to be known, then without loss of generality we may assume the signs are all positive and $\Phi_\lambda(\mathbf{X}|a,\beta)$ is defined to be the average of $\phi_\lambda(\mathbf{X}|\beta)$ over all possible active sets of size a. Alternatively, when one is unwilling to make the assumption of known effect signs, $\Phi^{\pm}_\lambda(\mathbf{X}|a,\beta)$ is calculated as the average of $\phi_\lambda(\mathbf{X}|\beta)$ over all possible active sets of size a and all 2^a possible sign vectors for the active factors. Stallrich et al. (2023) do provide some results that halve the number of sign vectors (2^{a-1})) that need ot be considered in the computation.

4 Comparing Larger Screening Designs via Sign Probabilities

The graphs below show the sign recovery probabilities for the entire set of possible two-factor interactions from the scenarios presented in Mee et al. (2017).

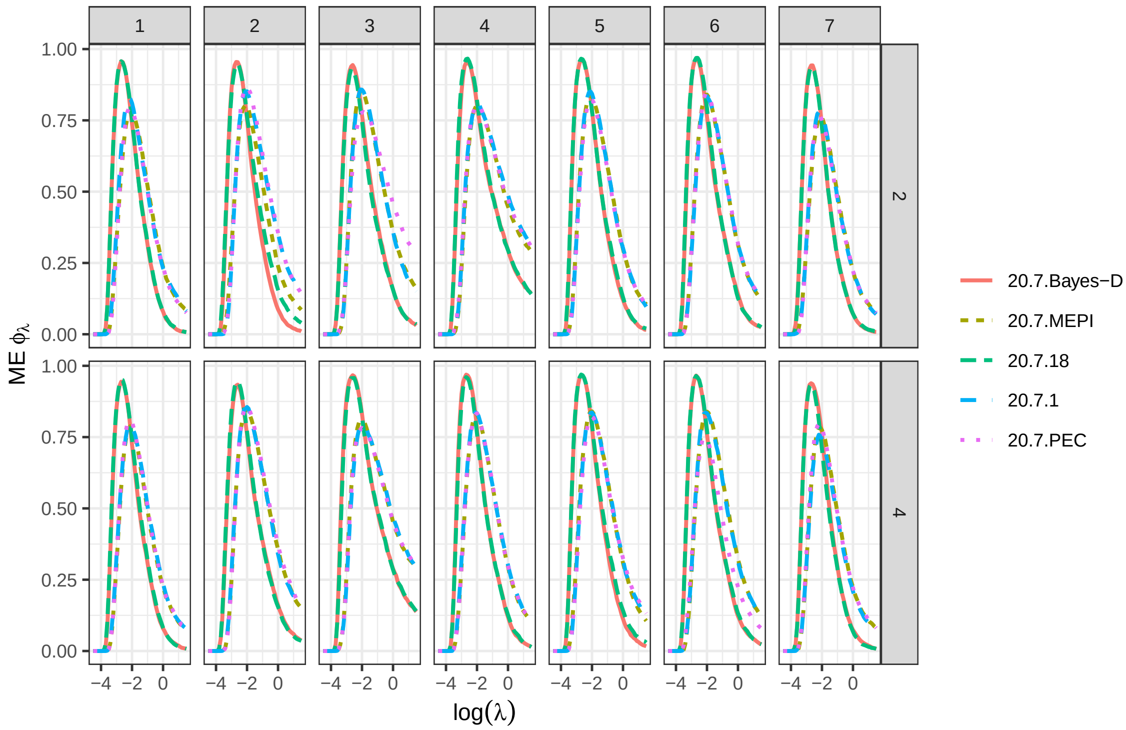

Figure 19: Comparison of $n = 20$, $k = 7$ designs using the simulated lasso sign recovery probability for the main effects only. The sign recovery probability for main effects is compared for 2 and 4 active main effects and 1-7 active two-factor interactions (columns in the figure).

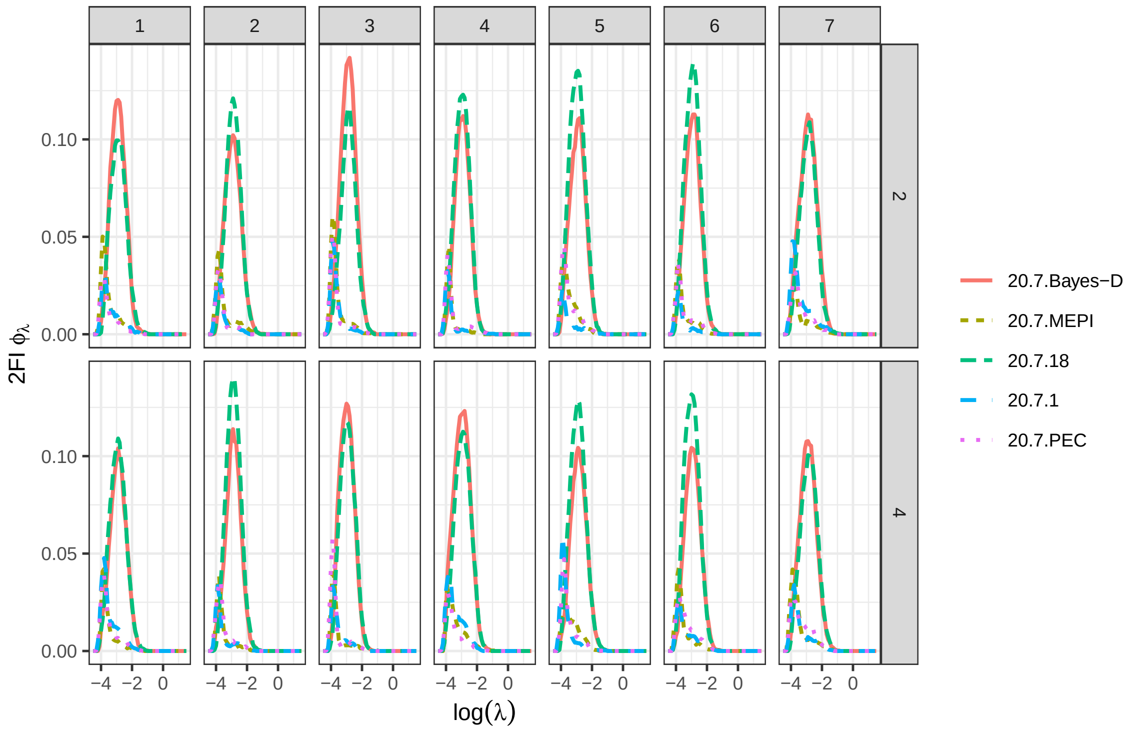

Figure 20: Comparison of $n = 20$, $k = 7$ designs using the simulated lasso sign recovery probability for the two factor interactions only. The sign recovery probability for two factor interactions is compared for 2 and 4 active main effects (rows in the figure) and 1-7 active two-factor interactions (columns in the figure).

References

Marley, C. J. and Woods, D. C. (2010), "A comparison of design and model selection methods for supersaturated experiments," *Computational Statistics and Data Analysis*, 54, 3158–3167.

Mee, R. W., Schoen, E. D., and Edwards, D. J. (2017), "Selecting an orthogonal or nonorthogonal two-level design for screening," *Technometrics*, 59, 305–318.

Stallrich, J. W., Young, K., Weese, M. L., Smucker, B. J., and Edwards, D. J. (2023), "Optimal Supersaturated Designs for Lasso Sign Recovery," .