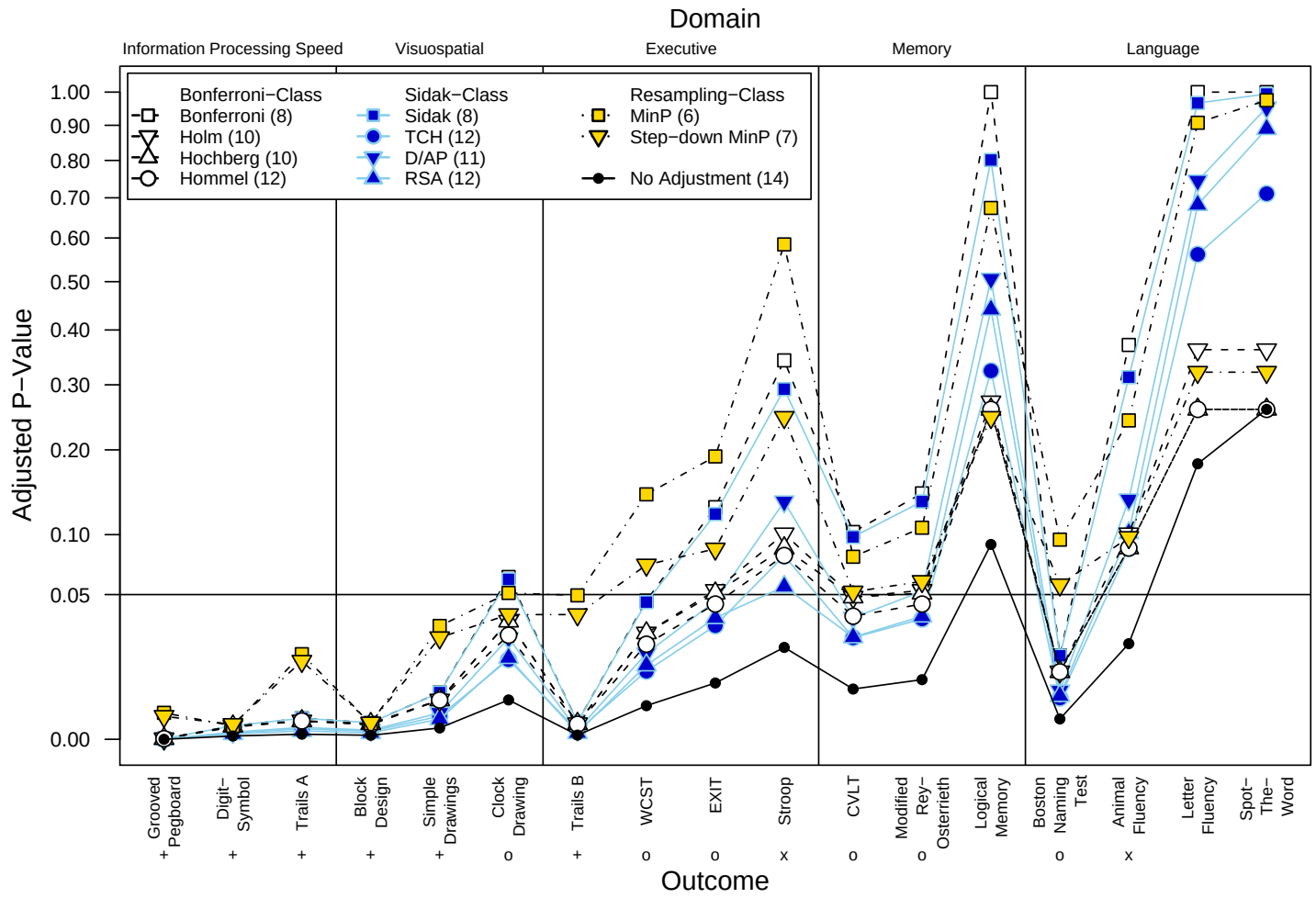

Figure 1: Adjusted *P*-Values by Method across Neuropsychological Outcomes

Seventeen observed p-values for a set of 17 neuropsychological measures, and adjusted p-values per each method. A square-root scale is used to reduce overlapping points. Numbers in parentheses in the legend indicate the number of rejected hypotheses for that method. Symbols for outcomes with a null hypothesis rejected without adjustment indicate the following:

+ null hypothesis rejected using each adjustment method
× null hypothesis not rejected using any adjustment method
o null hypothesis rejected by some adjustment methods

Note. WCST = Wisconsin Card Sorting Test, EXIT = Executive Interview, CVLT = California Verbal Learning Test. Adapted from Table 2 of Butters et al. (2004), *Archives of General Psychiatry*, *61*(6), 587–595. *Copyright ©(2004), American Medical Association.* All rights reserved.

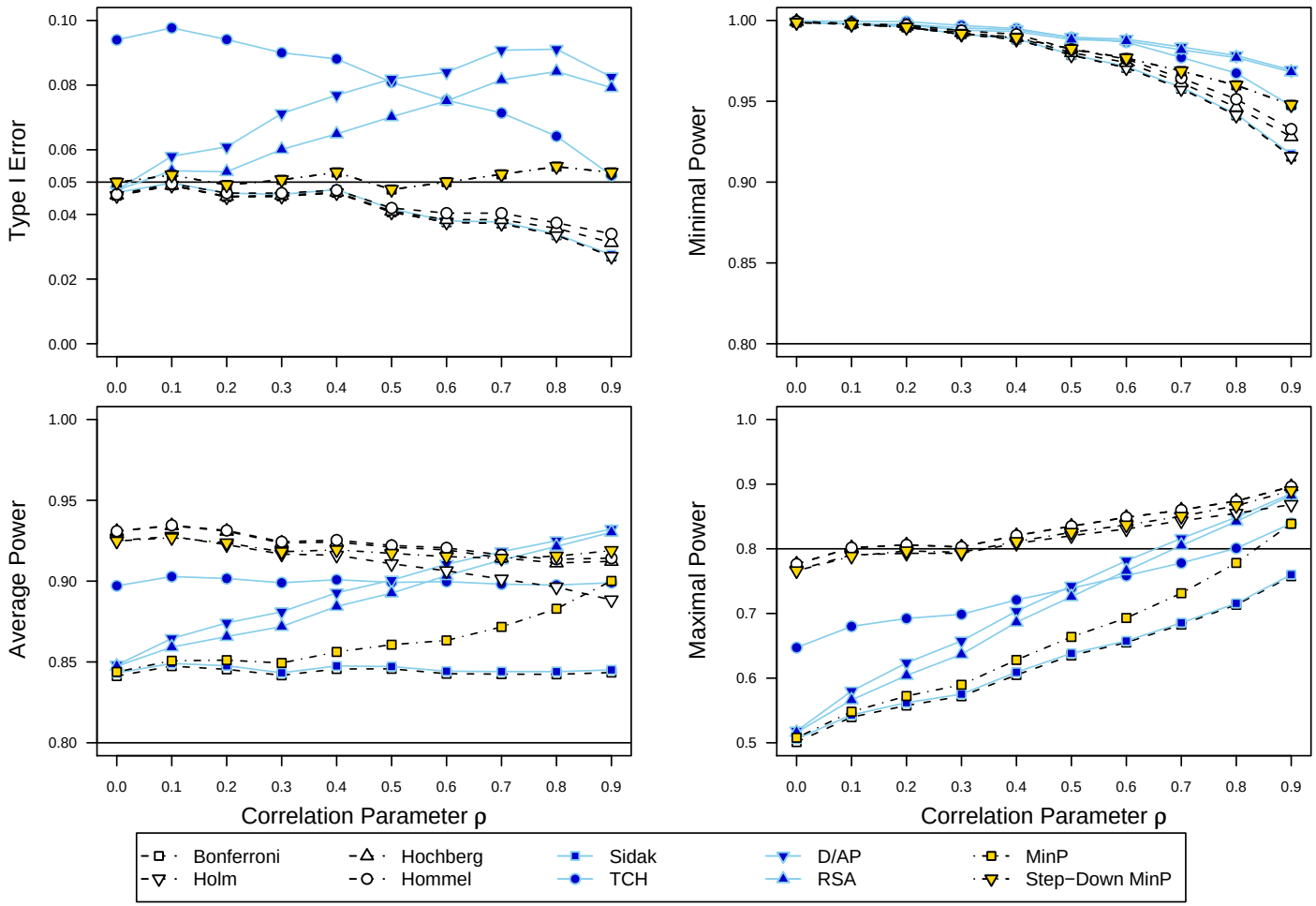

Figure 2: *P*-Value Adjustment Method Performance across Compound-Symmetry Correlation Structures
Type I Error and Power Estimates for Uniform Hypothesis Set

The upper-left panel shows Type I error rates of the p-value adjustment methods across increasing values of the CS correlation parameter ρ. In this case, all $M = 4$ hypotheses are simulated to be true. Values near $\alpha = .05$ are optimal. Values well above $\alpha = .05$ indicate failure to protect Type I error at α. The remaining panels show different measures of power, where the 4 hypotheses are simulated to be false. Higher power is optimal, conditional upon Type I error not exceeding α.

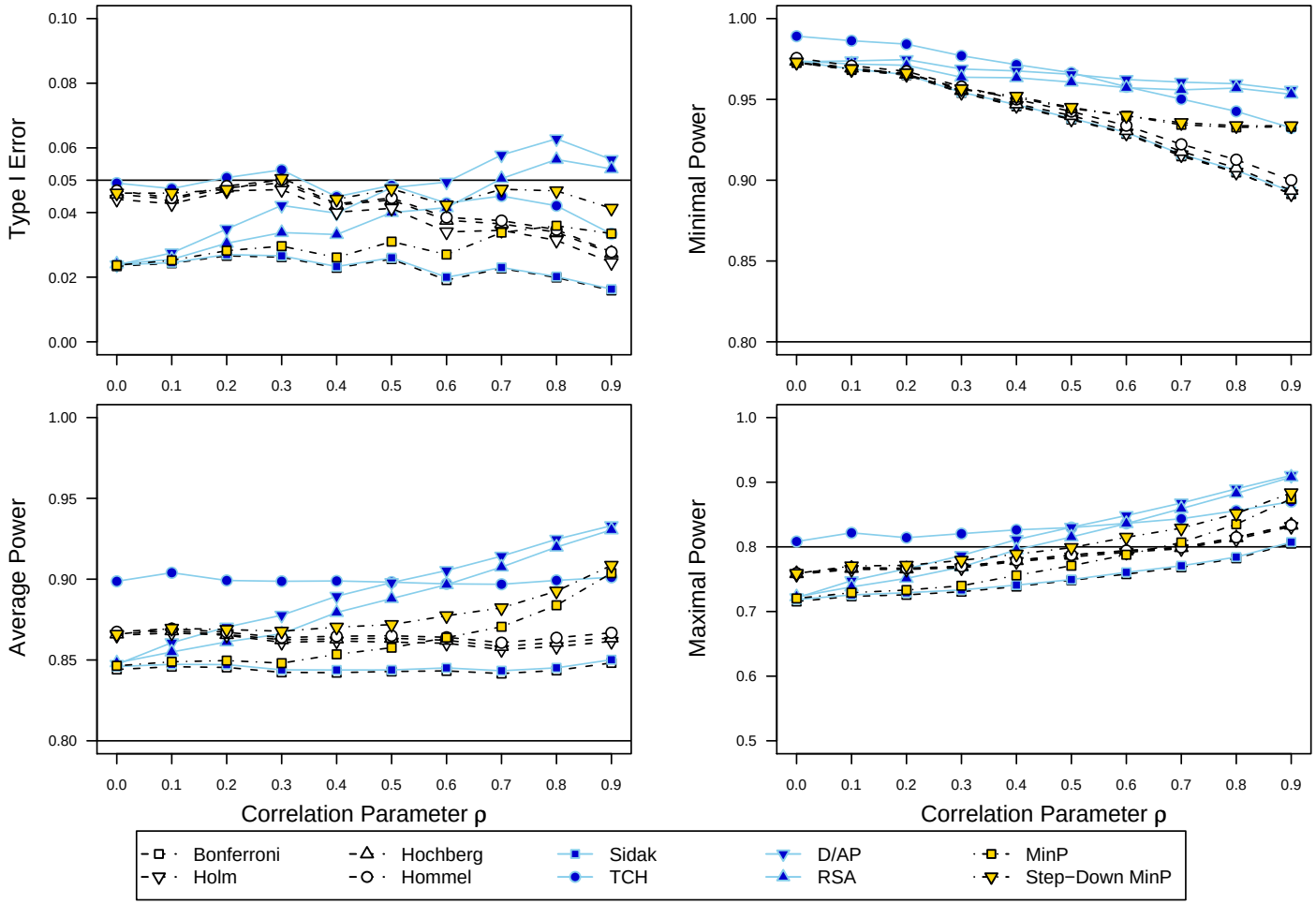

Figure 3: *P*-Value Adjustment Method Performance across Compound-Symmetry Correlation Structures
Type I Error and Power Estimates for Split Hypothesis Set

The upper-left panel shows Type I error rates of the p-value adjustment methods across increasing values of the CS correlation parameter ρ. In this case, all only 2 of the $M = 4$ hypotheses are simulated to be true. Values near $\alpha = .05$ are optimal. Values well above $\alpha = .05$ indicate failure to protect Type I error at α. The remaining panels show different measures of power, using the two hypotheses simulated to be false. Higher power is optimal, conditional upon Type I error not exceeding α.

Table S1: Adjusted P-Values by Method across Neuropsychological Outcomes

Domain	Outcome	No Adjustment	Bonferroni	Holm	Hochberg	Hommel	Sidak	TCH	D/AP	RSA	minP	sd.minP
Information Processing Speed	Grooved Pegboard	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	0.0017	0.0013
	Digit-Symbol	<0.0001	0.0004	0.0004	0.0004	0.0004	0.0004	0.0001	0.0001	<0.0001	0.0004	0.0005
	Trails A	<0.0001	0.0011	0.0008	0.0008	0.0008	0.0011	0.0003	0.0003	0.0002	0.0173	0.0146
Visuospatial	Block Design	<0.0001	0.0007	0.0006	0.0006	0.0005	0.0007	0.0002	0.0002	<0.0001	0.0006	0.0007
	Simple Drawings	0.0003	0.0052	0.0037	0.0037	0.0037	0.0052	0.0013	0.0017	0.0010	0.0307	0.0247
	Clock Drawing	0.0037	0.0629	0.0333	0.0333	0.0259	0.0611	0.0152	0.0245	0.0159	0.0511	0.0371
Executive	Trails B	<0.0001	0.0007	0.0006	0.0006	0.0005	0.0007	0.0002	0.0002	<0.0001	0.0495	0.0371
	WCST	0.0027	0.0459	0.0270	0.0270	0.0216	0.0449	0.0111	0.0176	0.0131	0.1432	0.0726
	EXIT	0.0076	0.1286	0.0530	0.0510	0.0437	0.1211	0.0308	0.0456	0.0352	0.1910	0.0866
	Stroop	0.0202	0.3428	0.1008	0.0874	0.0807	0.2927	0.0806	0.1348	0.0557	0.5847	0.2475
Memory	CVLT	0.0060	0.1026	0.0483	0.0483	0.0362	0.0978	0.0246	0.0353	0.0249	0.0796	0.0519
	Modified Rey-Osterrieth	0.0085	0.1444	0.0530	0.0510	0.0437	0.1350	0.0346	0.0522	0.0361	0.1069	0.0596
	Logical Memory	0.0906	>0.9999	0.2719	0.2599	0.2599	0.8012	0.3241	0.5059	0.4410	0.6740	0.2475
Language	Boston Naming Test	0.0010	0.0168	0.0109	0.0109	0.0109	0.0167	0.0041	0.0056	0.0045	0.0952	0.0570
	Animal Fluency	0.0218	0.3713	0.1008	0.0874	0.0874	0.3130	0.0870	0.1371	0.1016	0.2428	0.0974
	Language Fluency	0.1812	>0.9999	0.3624	0.2599	0.2599	0.9666	0.5615	0.7446	0.6822	0.9076	0.3218
	Spot-The-Word	0.2599	>0.9999	0.3624	0.2599	0.2599	0.9940	0.7108	0.9528	0.8896	0.9750	0.3218

Note. WCST = Wisconsin Card Sorting Test, EXIT = Executive Interview, CVLT = California Verbal Learning Test. Adapted from Table 2 of Butters et al. (2004), *Archives of General Psychiatry*, *61*(6), 587–595.

Table S2: BS Simulation Series Parameters

	Outcome Types			
	Block 1		Block 2	
Hypothesis Sets	V1	V2	V3	V4
Uniform - True	TN	TN	TN	TN
Uniform - False	FN	FN	FN	FN
Split - Uniform	TN	TN	FN	FN
Split - Split	TN	FN	FN	TN

The M outcomes of the r^{th} replicate in a given trial are simulated according to the choice of hypothesis set. Outcomes $V1-V4$ may be one of two types. True null (TN) outcomes are simulated with effect size 0.0, and are used to estimate Type I error. False null (FN) outcomes are simulated with effect size 0.5, and are used to estimate power.

Correlation Structure	V1	V2	V3	V4
V1	1	W	B	B
V2	W	1	B	B
V3	B	B	1	W
V4	B	B	W	1

Data may be simulated with a block symmetry (BS) correlation structure, where all outcomes within a block are equicorrelated with parameter W, and outcomes from different blocks are equicorrelated with parameter B, where $W > B$. The W and B parameters take on values of $\{0.2, 0.5, 0.8\}$ and $\{0.0, 0.2, 0.5\}$, respectively.

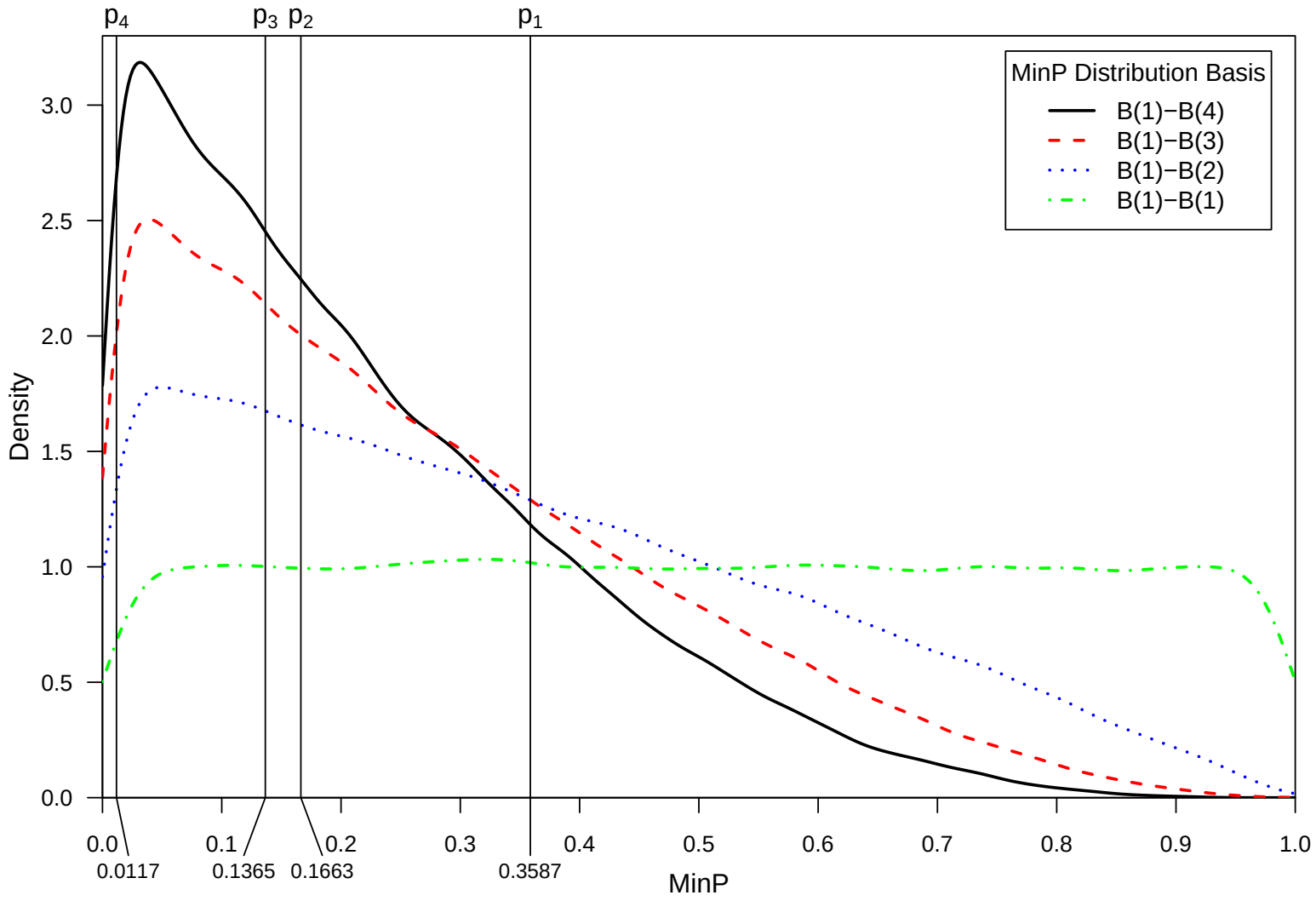

Figure S1: Bootstrap Empirical MinP Null Distributions for the Illustrative Example

For the minP method, an adjusted p-value p_{aj}, $j = 1$ to 4, is calculated by the area left of p_j and below the distribution curve based on all bootstrap outcomes, $B(1)$-$B(4)$. For the sd.minP method, p_{aj}, $j = 1$ to 4, is calculated by the area left of p_j and below the distribution curve based on outcomes $B(1)$-$B(j)$, and adjusted to ensure the same order of the observed p_j's.

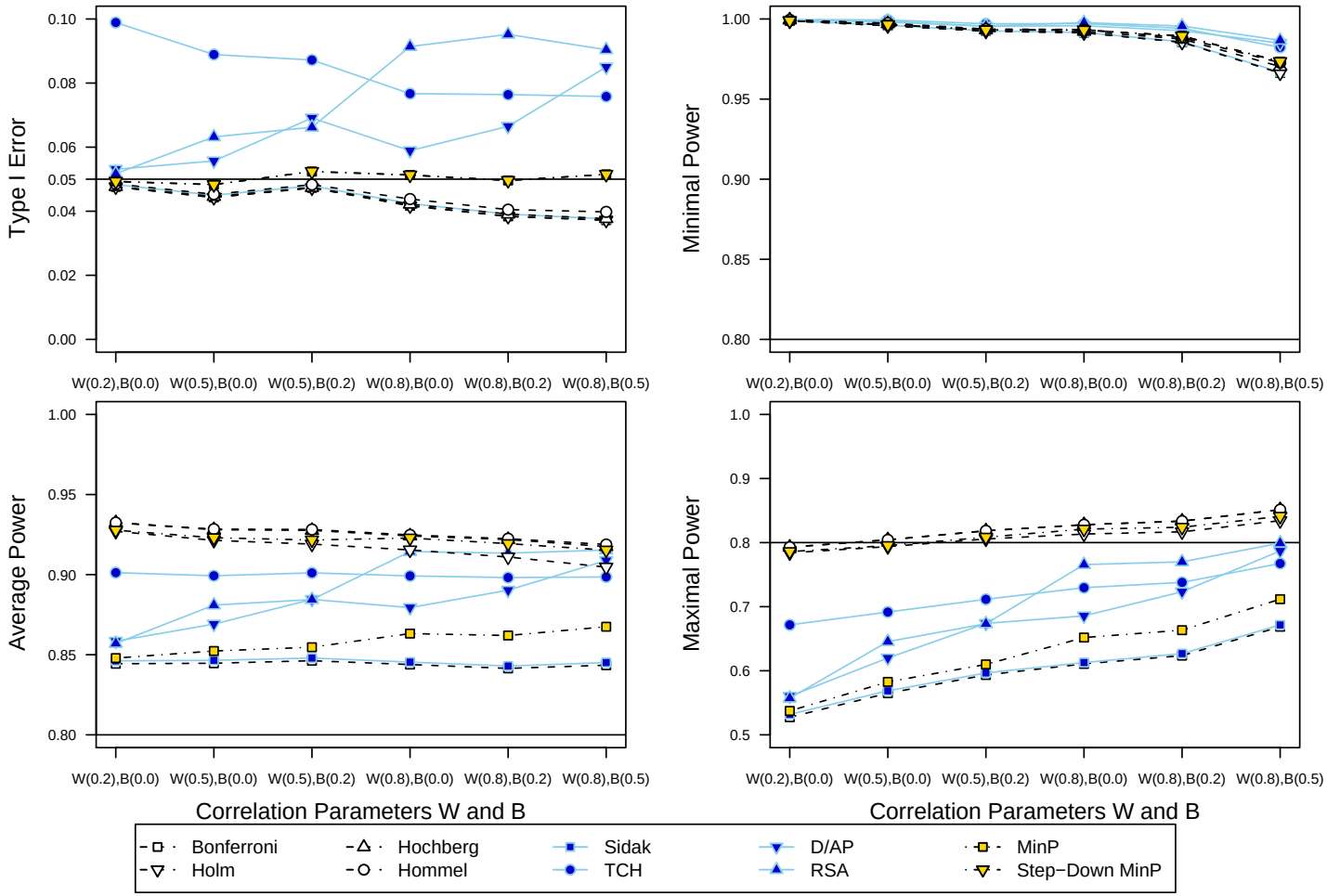

Figure S2: P-Value Adjustment Method Performance across Block-Symmetry Correlation Structures
Type I Error and Power Estimates for Uniform Hypothesis Set

Each figure represents a different hypothesis set. The upper-left panel of each figure shows Type I error rates of the p-value adjustment methods across increasing values of the block-symmetry correlation parameters B and W. Values near $\alpha = 0.05$ are optimal. Values well above $\alpha = 0.05$ indicate failure to protect Type I error at α. Higher power is optimal, conditional upon Type I error not exceeding α.

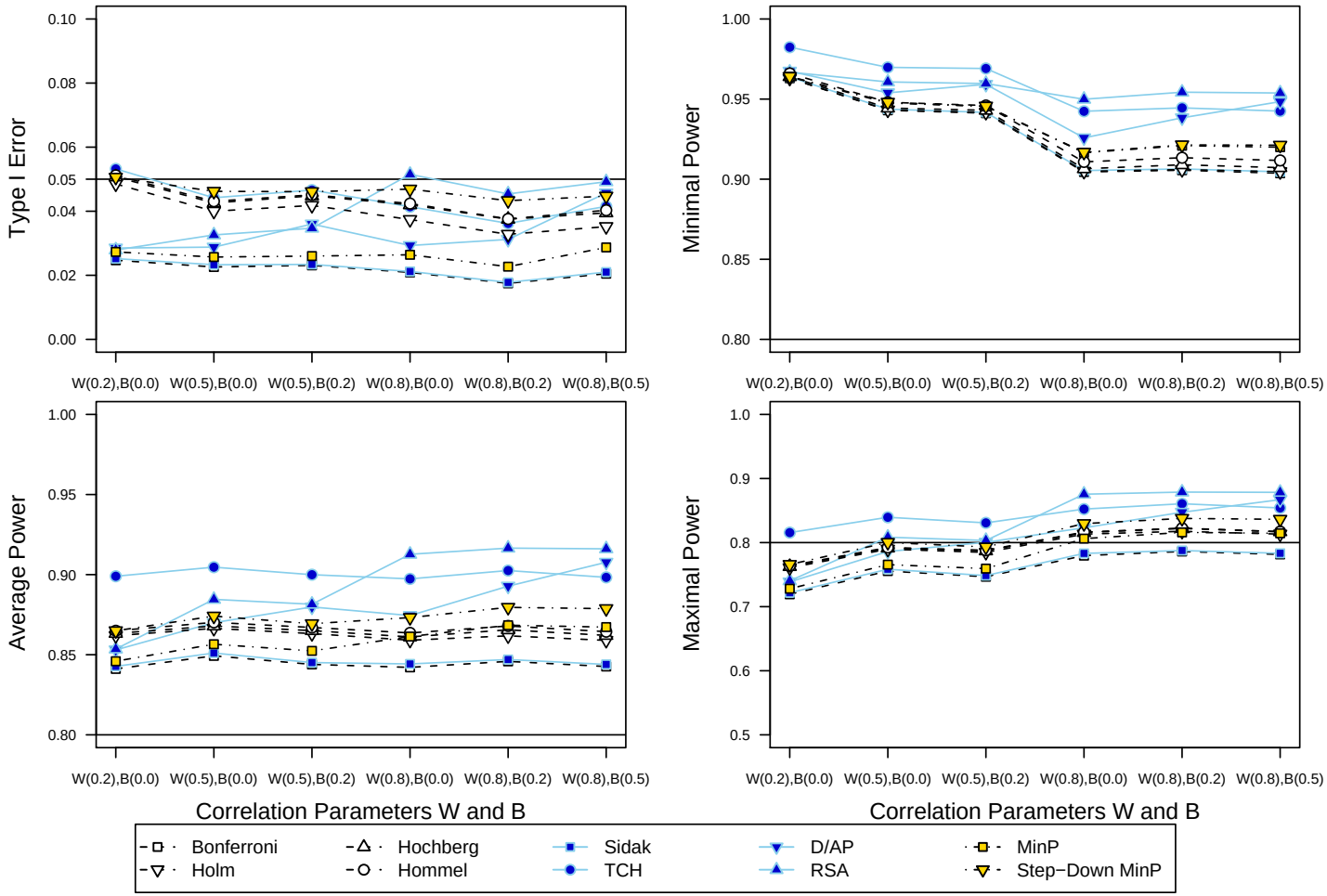

Figure S3: P-Value Adjustment Method Performance across Block-Symmetry Correlation Structures
Type I Error and Power Estimates for Split - Uniform Hypothesis Set

Each figure represents a different hypothesis set. The upper-left panel of each figure shows Type I error rates of the p-value adjustment methods across increasing values of the block-symmetry correlation parameters B and W. Values near $\alpha = 0.05$ are optimal. Values well above $\alpha = 0.05$ indicate failure to protect Type I error at α. Higher power is optimal, conditional upon Type I error not exceeding α.

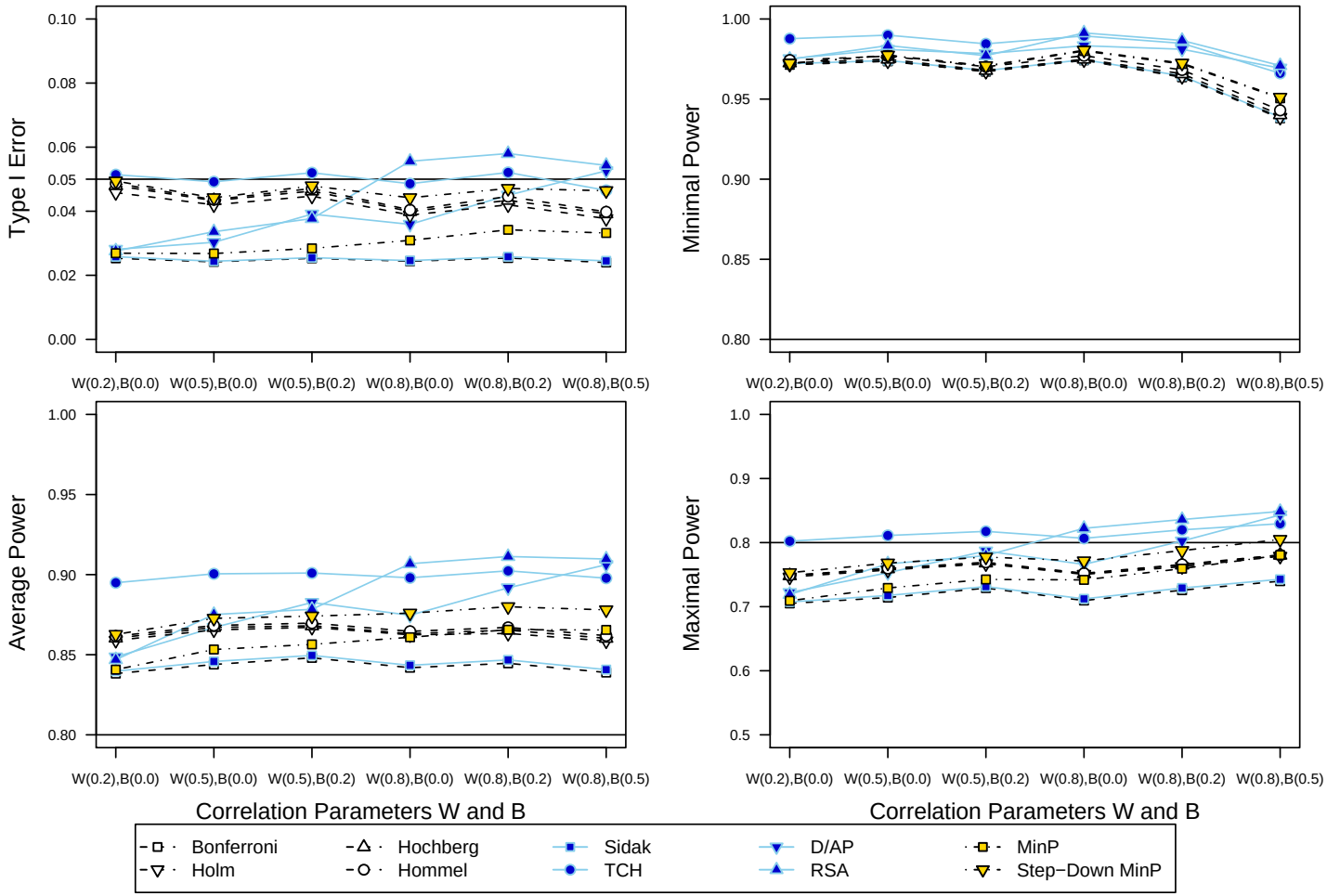

Figure S4: P-Value Adjustment Method Performance across Block-Symmetry Correlation Structures
Type I Error and Power Estimates for Split - Split Hypothesis Set

Each figure represents a different hypothesis set. The upper-left panel of each figure shows Type I error rates of the p-value adjustment methods across increasing values of the block-symmetry correlation parameters B and W. Values near $\alpha = 0.05$ are optimal. Values well above $\alpha = 0.05$ indicate failure to protect Type I error at α. Higher power is optimal, conditional upon Type I error not exceeding α.