

Supplementary materials

Table S1. Empirical type I error of quantitative traits with t distribution error terms using different distance functions in microbiome genetic association.

α	CI_lower	CI_upper	Kw_SKAT	Kw_aSKAT	Ku_SKAT	Ku_aSKAT	KBC_SKAT	KBC_aSKAT
df = 1								
5.00E-02	4.86E-02	5.14E-02	2.84E-02	5.61E-02	4.07E-03	4.24E-02	9.34E-03	4.77E-02
1.00E-02	9.38E-03	1.06E-02	3.53E-03	1.16E-02	7.00E-05	6.78E-03	4.30E-04	8.13E-03
1.00E-03	8.04E-04	1.20E-03	2.30E-04	1.00E-03	0.00E+00	4.10E-04	0.00E+00	6.90E-04
df = 5								
5.00E-02	4.86E-02	5.14E-02	2.59E-02	5.11E-02	6.44E-03	4.96E-02	1.15E-02	5.06E-02
1.00E-02	9.38E-03	1.06E-02	3.32E-03	1.03E-02	2.00E-04	1.02E-02	7.30E-04	1.01E-02
1.00E-03	8.04E-04	1.20E-03	1.60E-04	1.06E-03	0.00E+00	9.50E-04	1.00E-05	1.06E-03

The total sample size is 100, α is the nominal significance level. CI_lower and CI_upper are the lower and upper endpoint of the 95% confidence interval assuming the true type I error is α based on 10^5 replicates. Values in bold indicate type I errors above the upper endpoint of the confidence interval. The error term was simulated using the t distribution with 1 and 5 degrees of freedom (df). Kw denotes the kernel from the weighted phylogeny-based UniFrac distances, Ku denotes the kernel from the unweighted phylogeny-based UniFrac distances, and KBC denotes the kernel from the nonphylogeny-based Bray-Curtis distance. uSKAT does not use small sample adjustment for quantitative trait. aSKAT is the proposed adjusted SKAT.

Table S2. Empirical type I error of quantitative traits with non-linear covariates using different distance functions in microbiome genetic association.

α	CI_lower	CI_upper	Kw_SKAT	Kw_aSKAT	Ku_SKAT	Ku_aSKAT	KBC_SKAT	KBC_aSKAT
5.00E-02	4.86E-02	5.14E-02	2.57E-02	5.10E-02	6.26E-03	4.97E-02	1.19E-02	4.98E-02
1.00E-02	9.38E-03	1.06E-02	3.23E-03	1.00E-02	2.00E-04	9.51E-03	6.00E-04	1.02E-02
1.00E-03	8.04E-04	1.20E-03	1.80E-04	9.00E-04	1.00E-05	1.08E-03	0.00E+00	9.60E-04

The total sample size is 100, α is the nominal significance level. CI_lower and CI_upper are the lower and upper endpoint of the 95% confidence interval assuming the true type I error is α based on 10^5 replicates. The non-linear term was a squared covariate but not accounted for in the test. Kw denotes the kernel from the weighted phylogeny-based UniFrac distances, Ku denotes the kernel from the unweighted phylogeny-based UniFrac distances, and KBC denotes the kernel from the nonphylogeny-based Bray-Curtis distance. uSKAT does not use small sample adjustment for quantitative trait. aSKAT is the proposed adjusted SKAT.

Figure S1

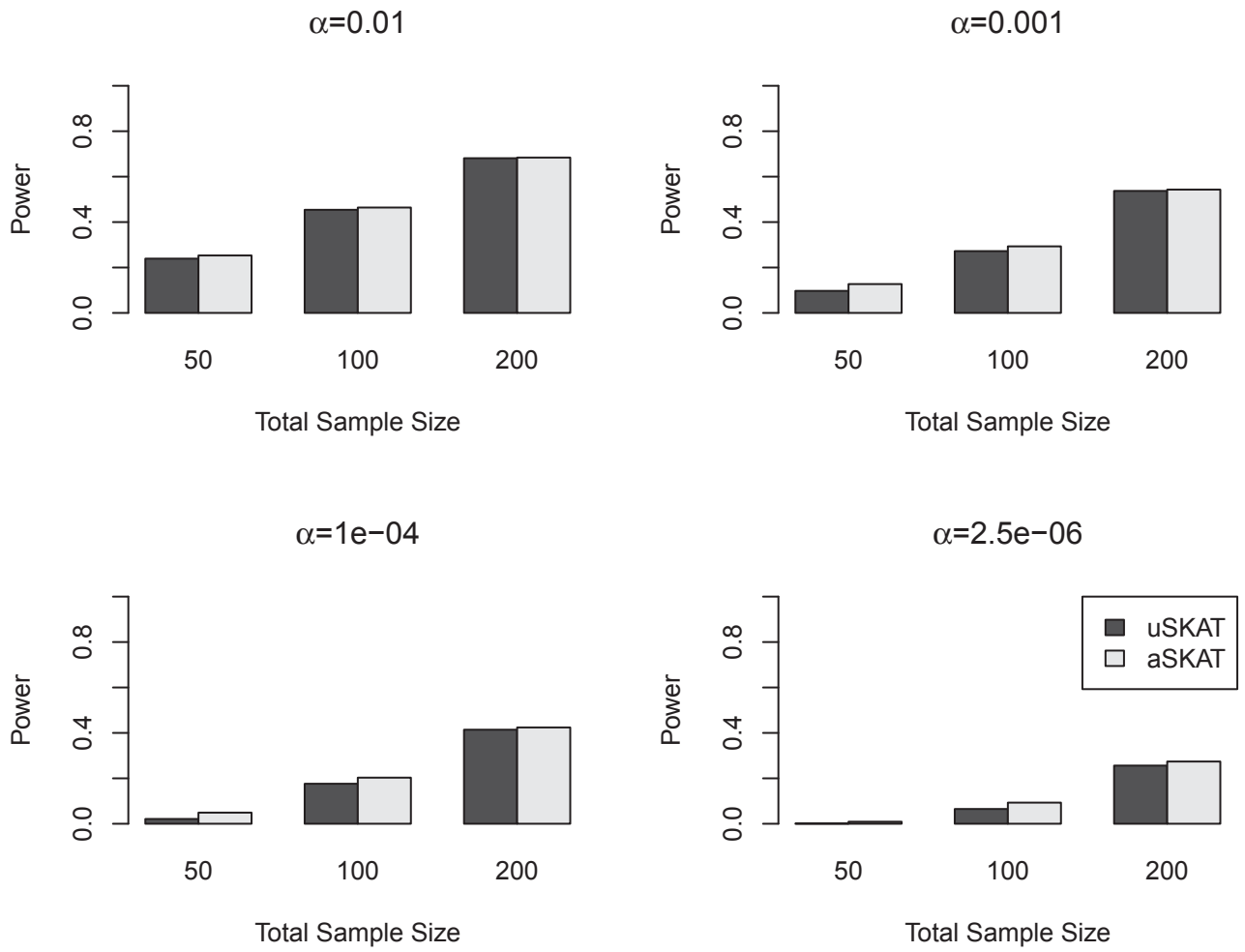


Figure S2

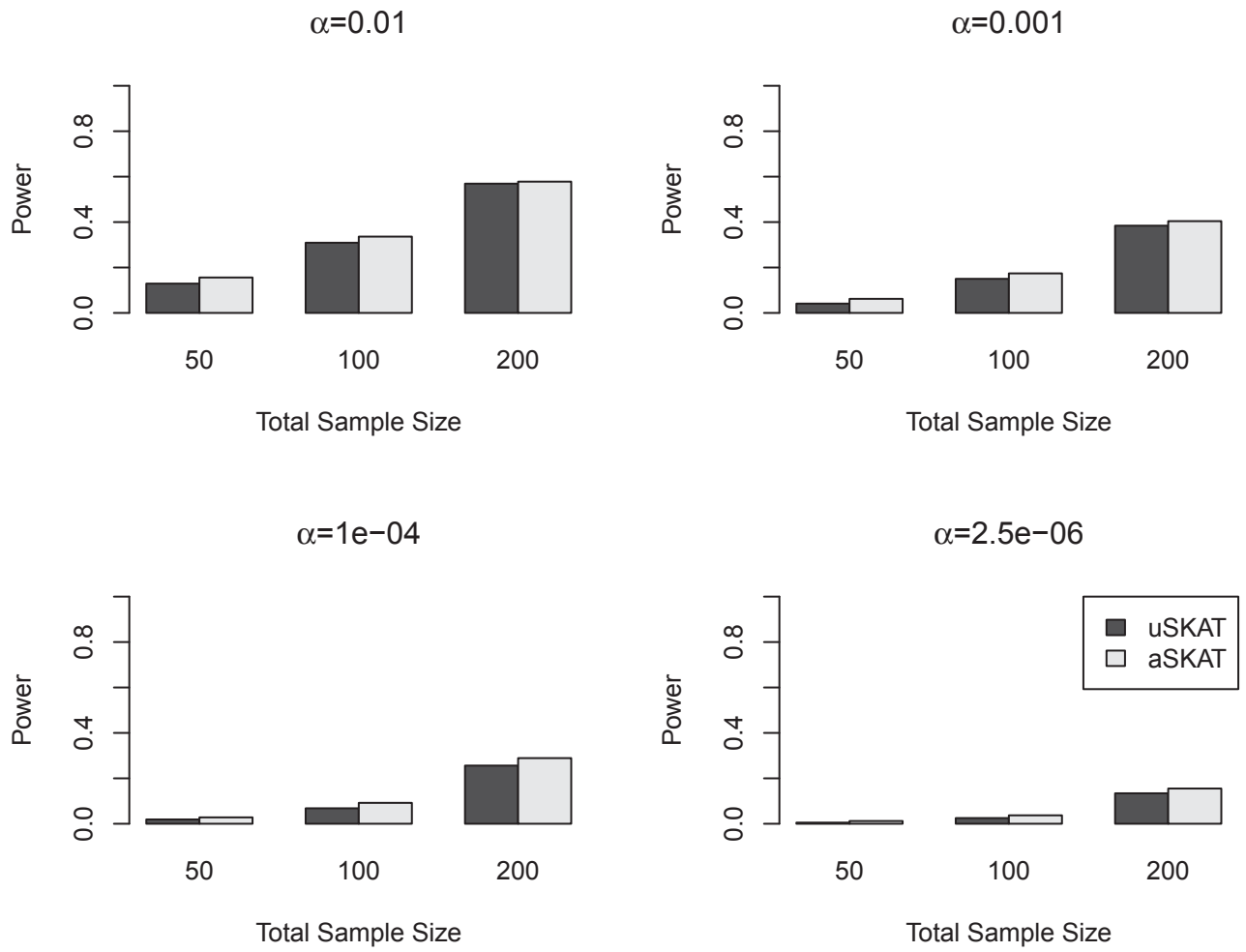


Figure S3

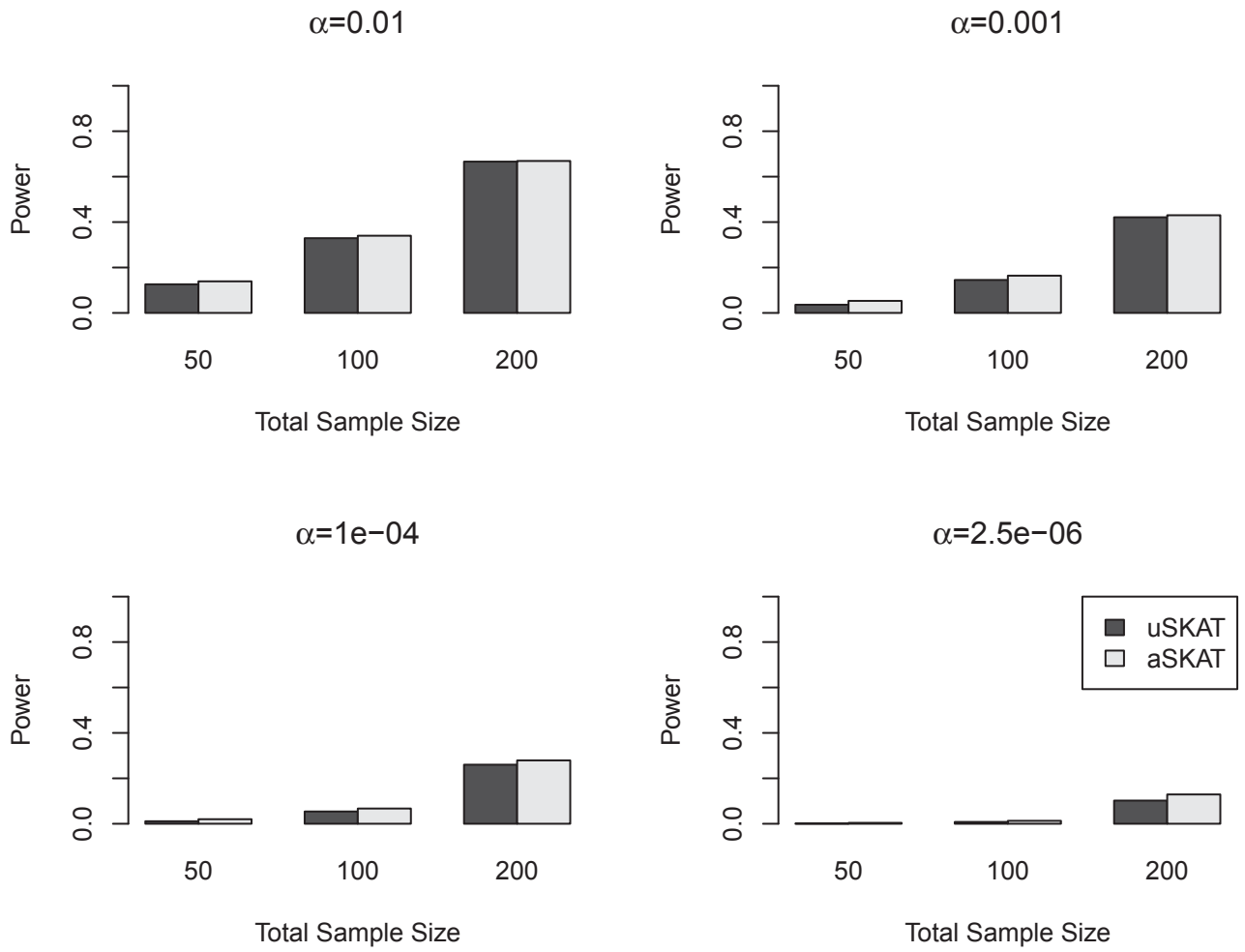


Figure S4

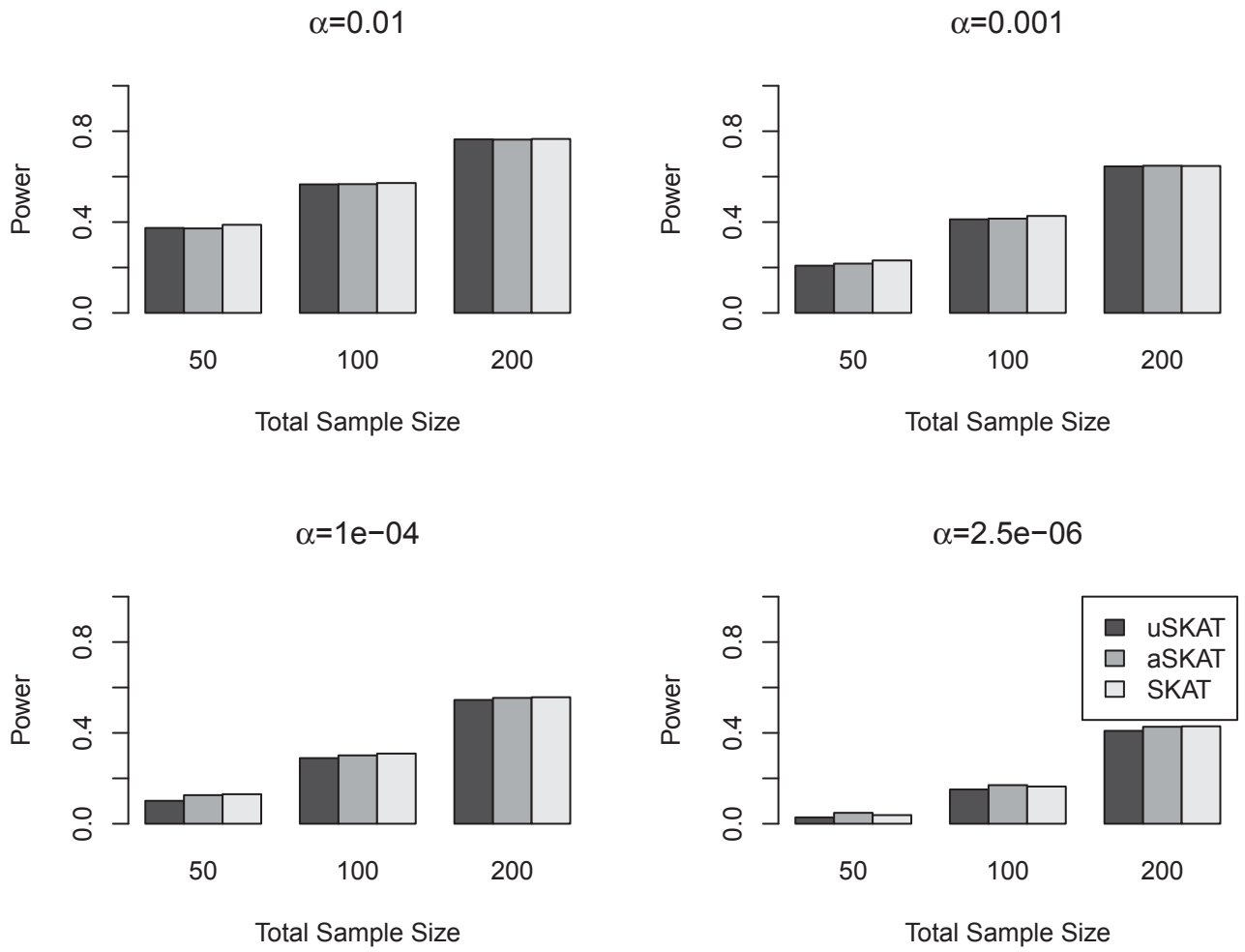


Figure S5

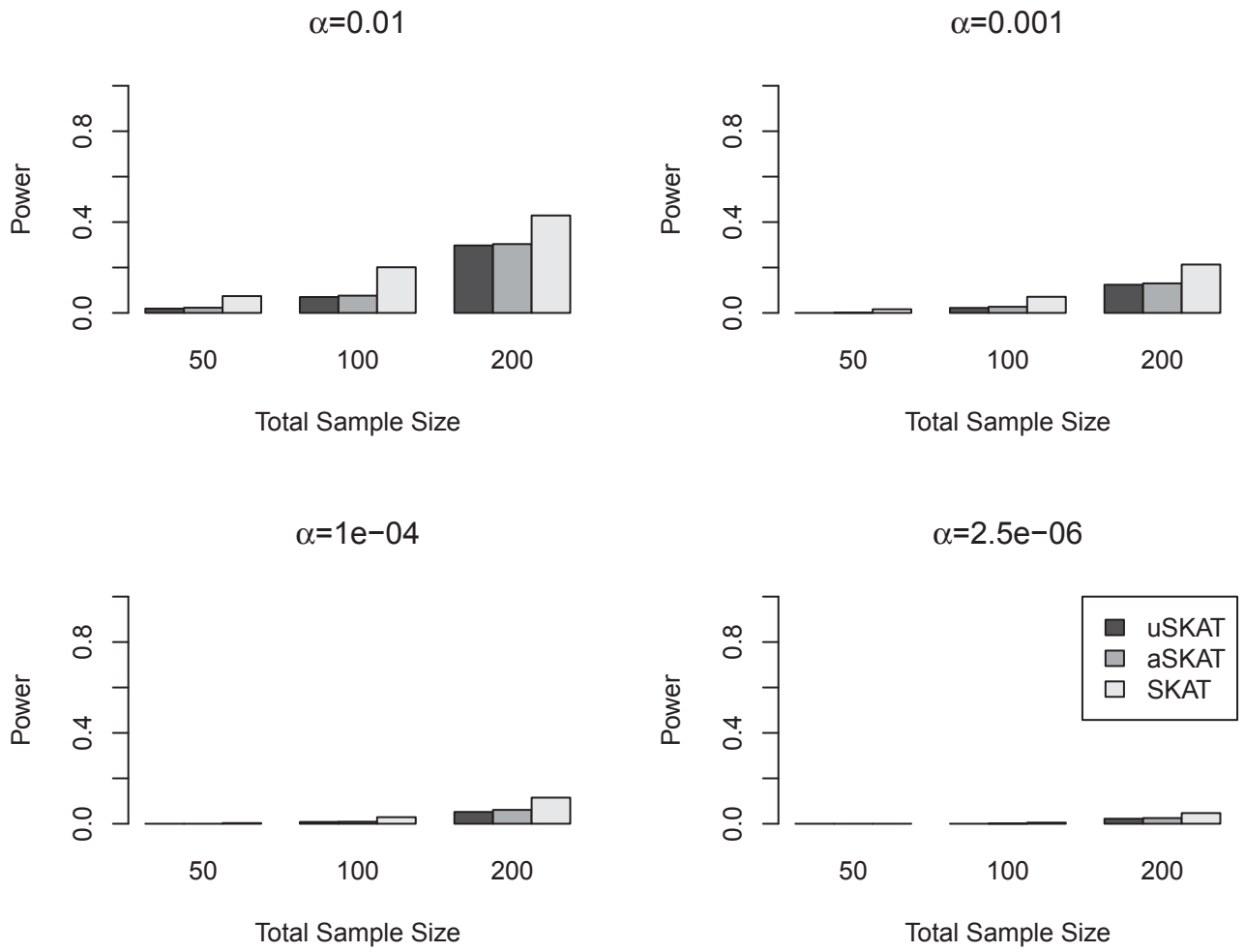


Figure S6

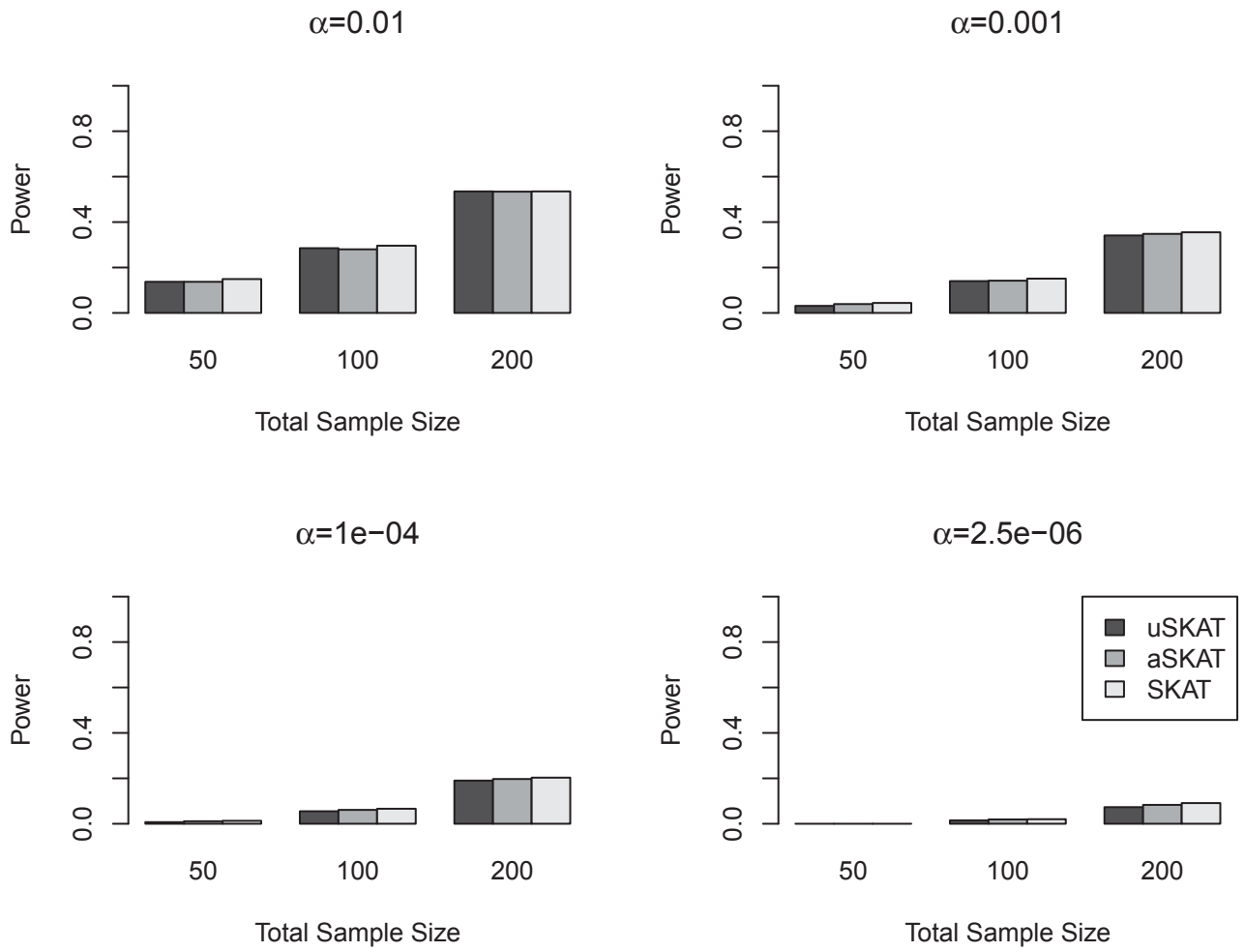


Figure S7

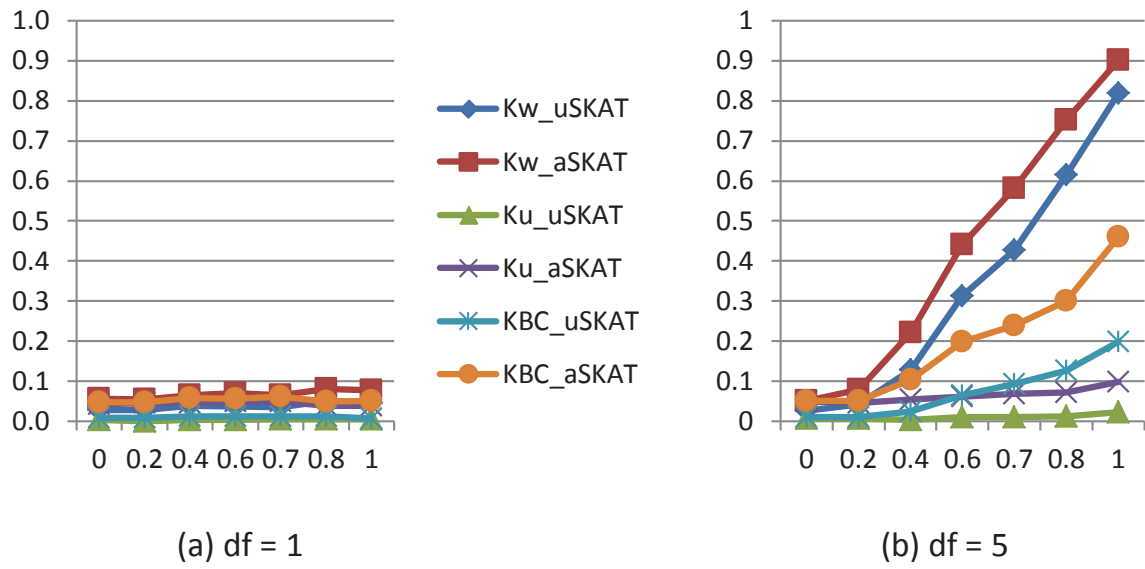


Figure S8

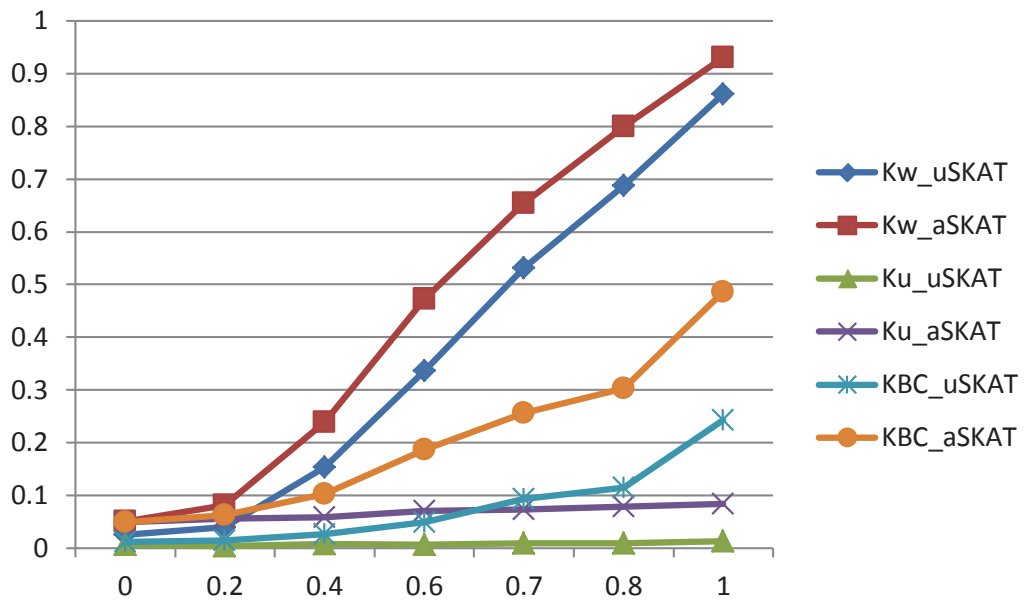


Figure S9

