
Improving data interpretability with new differential variance gene set tests (Additional File 1)

Yasir Rahmatallah ¹ and Galina Glazko ¹
¹ Department of Biomedical Informatics,
University of Arkansas for Medical Sciences,
Little Rock, AR 72205

August 8, 2024

List of Tables

S1	Type I error rate at sample size $N = 20$ and gene set size $p = 20$ ($\alpha = 0.05$)	5
S2	Type I error rate at sample size $N = 20$ and gene set size $p = 60$ ($\alpha = 0.05$)	5
S3	Type I error rate at sample size $N = 20$ and gene set size $p = 100$ ($\alpha = 0.05$)	5
S4	Type I error rate at sample size $N = 40$ and gene set size $p = 20$ ($\alpha = 0.05$)	6
S5	Type I error rate at sample size $N = 40$ and gene set size $p = 100$ ($\alpha = 0.05$)	6
S6	Type I error rate at sample size $N = 60$ and gene set size $p = 20$ ($\alpha = 0.05$)	6
S7	Type I error rate at sample size $N = 60$ and gene set size $p = 60$ ($\alpha = 0.05$)	7
S8	Type I error rate at sample size $N = 60$ and gene set size $p = 100$ ($\alpha = 0.05$)	7

List of Figures

S1	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with true differential variance in a gene set (γ) and inter-gene correlation (r).	8
S2	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	9
S3	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	10
S4	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	11
S5	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	12
S6	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	13
S7	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	14
S8	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	15
S9	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).	16

S10	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	17
S11	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	18
S12	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	19
S13	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	20
S14	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	21
S15	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	22
S16	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	23
S17	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	24
S18	Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	25
S19	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 20$ when the alternative hypothesis $\sigma_X \neq \sigma_Y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	26
S20	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	27
S21	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	28
S22	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	29

S23	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	30
S24	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	31
S25	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	32
S26	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	33
S27	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	34
S28	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 20$ when the alternative hypothesis $\mu_X \neq \mu_Y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	35
S29	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	36
S30	Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	37
S31	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	38
S32	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	39
S33	Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	40
S34	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	41
S35	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	42
S36	Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).	43

Table S1: Type I error rate at sample size $N = 20$ and gene set size $p = 20$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.039	0.049	0.049	0.049
0.043	0.052	0.051	0.051
0.043	0.050	0.052	0.052
0.041	0.043	0.052	0.051

Table S2: Type I error rate at sample size $N = 20$ and gene set size $p = 60$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.040	0.045	0.046	0.048
0.038	0.047	0.049	0.050
0.037	0.048	0.046	0.049
0.037	0.038	0.048	0.049

Table S3: Type I error rate at sample size $N = 20$ and gene set size $p = 100$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.041	0.047	0.050	0.049
0.040	0.049	0.048	0.047
0.042	0.051	0.049	0.050
0.037	0.037	0.046	0.046

Table S4: Type I error rate at sample size $N = 40$ and gene set size $p = 20$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.034	0.047	0.048	0.048
0.034	0.051	0.051	0.050
0.035	0.051	0.051	0.051
0.035	0.034	0.048	0.050

Table S5: Type I error rate at sample size $N = 40$ and gene set size $p = 100$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.035	0.047	0.049	0.049
0.035	0.048	0.052	0.053
0.032	0.049	0.051	0.048
0.031	0.031	0.047	0.046

Table S6: Type I error rate at sample size $N = 60$ and gene set size $p = 20$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.040	0.049	0.051	0.053
0.041	0.049	0.050	0.050
0.040	0.050	0.050	0.050
0.038	0.038	0.048	0.049

Table S7: Type I error rate at sample size $N = 60$ and gene set size $p = 60$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.039	0.046	0.048	0.048
0.036	0.044	0.044	0.045
0.037	0.046	0.048	0.047
0.037	0.037	0.045	0.047

Table S8: Type I error rate at sample size $N = 60$ and gene set size $p = 100$ ($\alpha = 0.05$)

Method Placement			
RKS1	RMD1	RCVM1	RAD1
RKS2	RMD2	RCVM2	RAD2
RKS3	RMD3	RCVM3	RAD3
KS	MD	CVM	AD
Type I error rate			
0.041	0.049	0.047	0.047
0.044	0.052	0.054	0.054
0.042	0.050	0.053	0.053
0.044	0.044	0.054	0.054

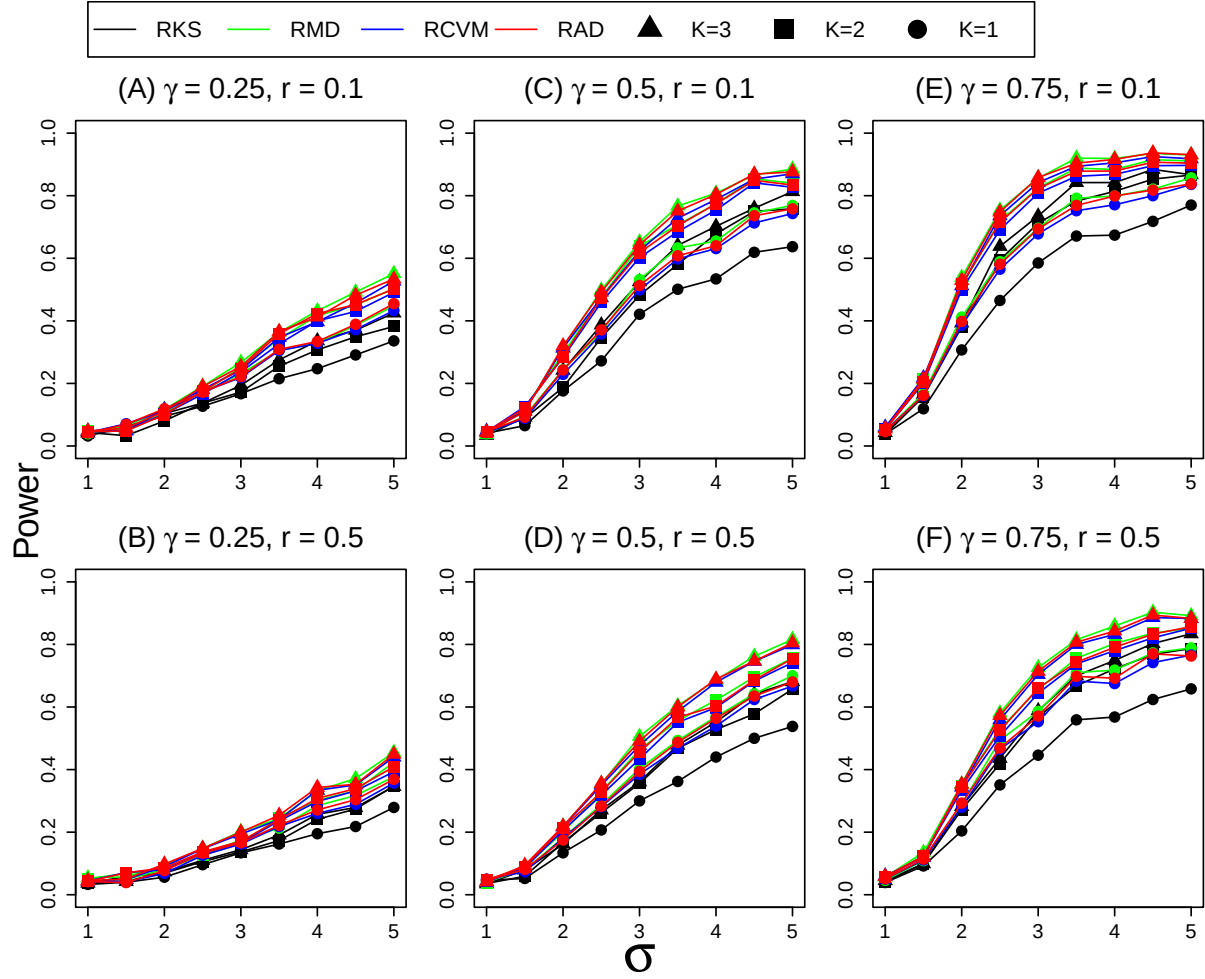


Figure S1: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with true differential variance in a gene set (γ) and inter-gene correlation (r).

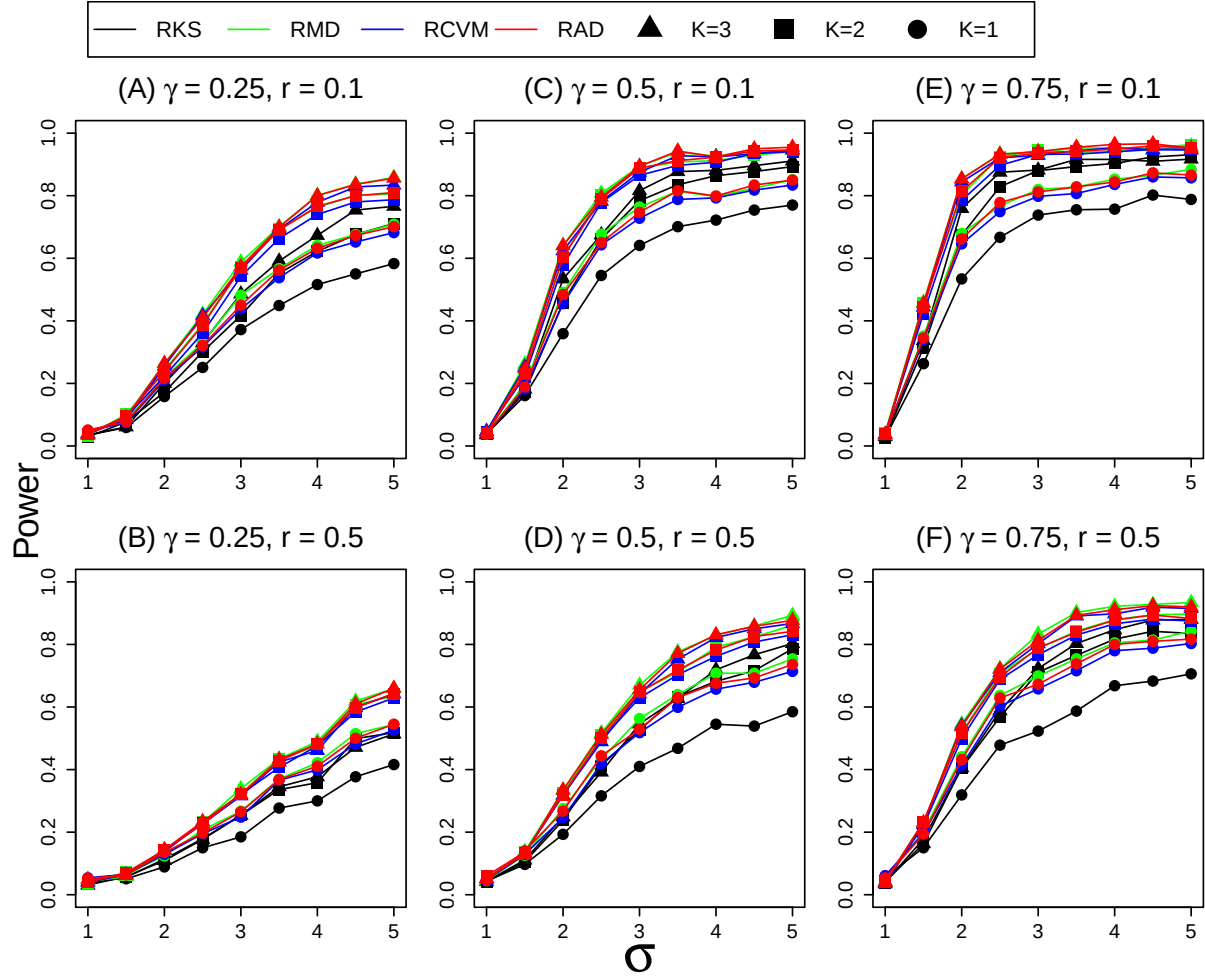


Figure S2: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

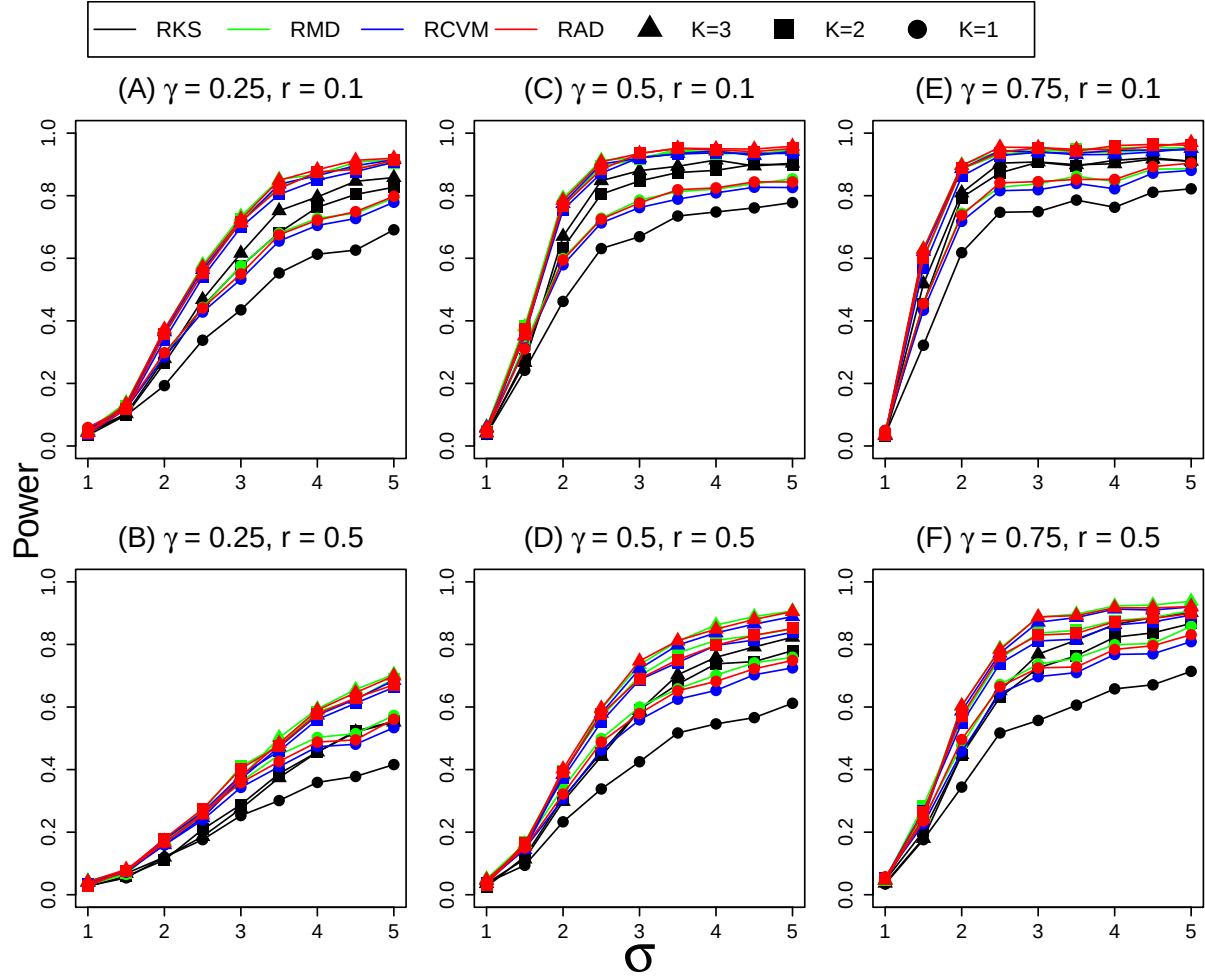


Figure S3: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

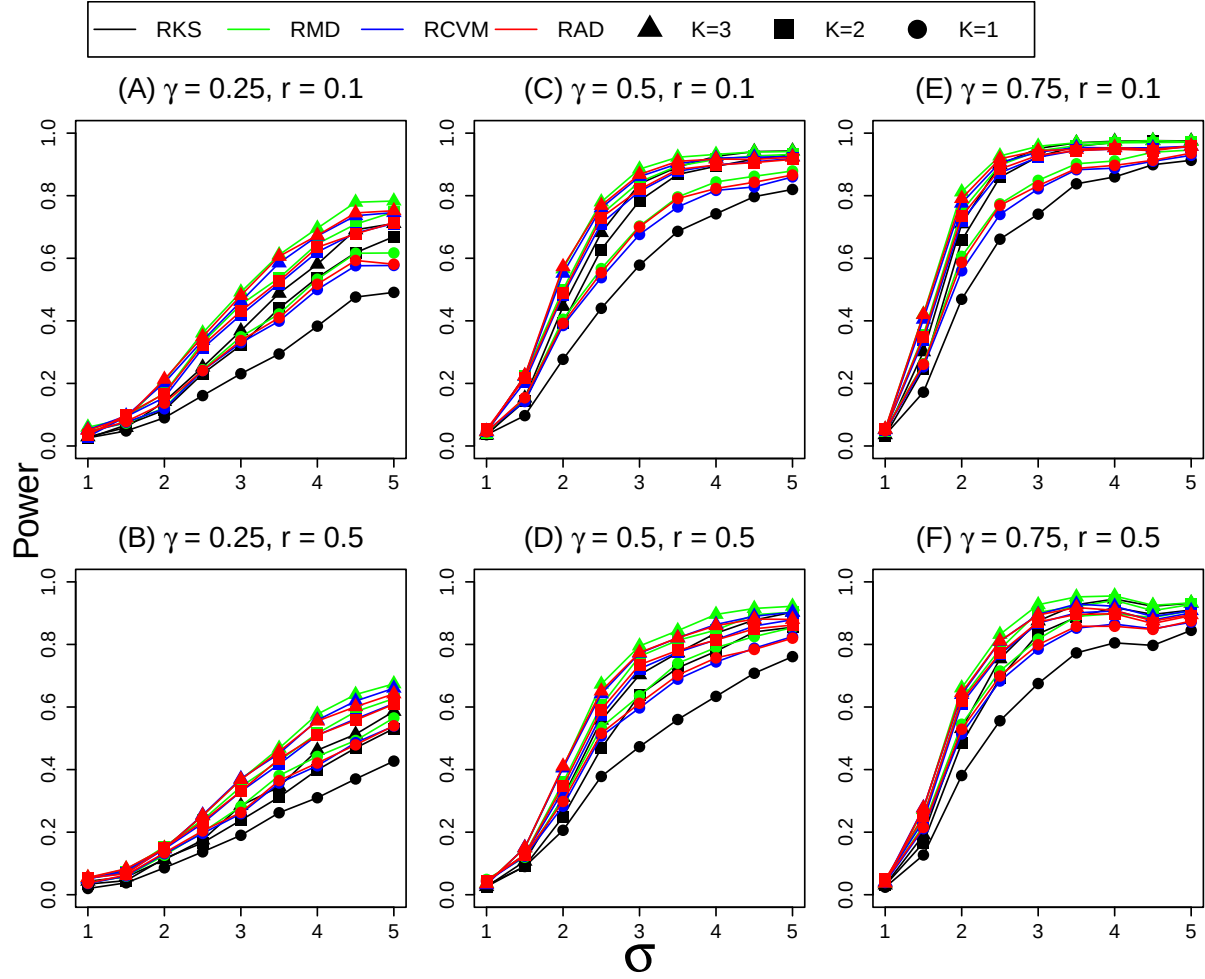


Figure S4: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

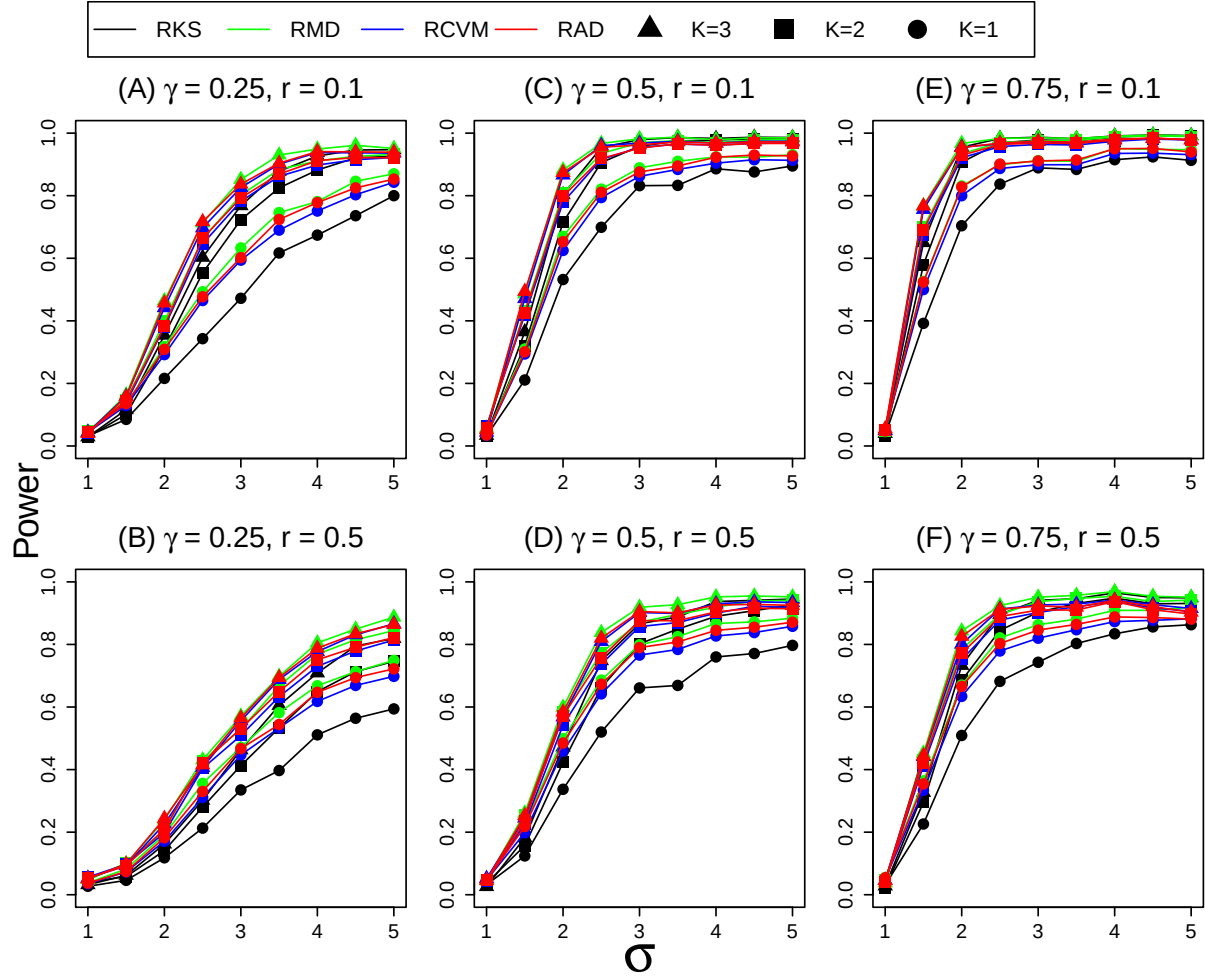


Figure S5: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

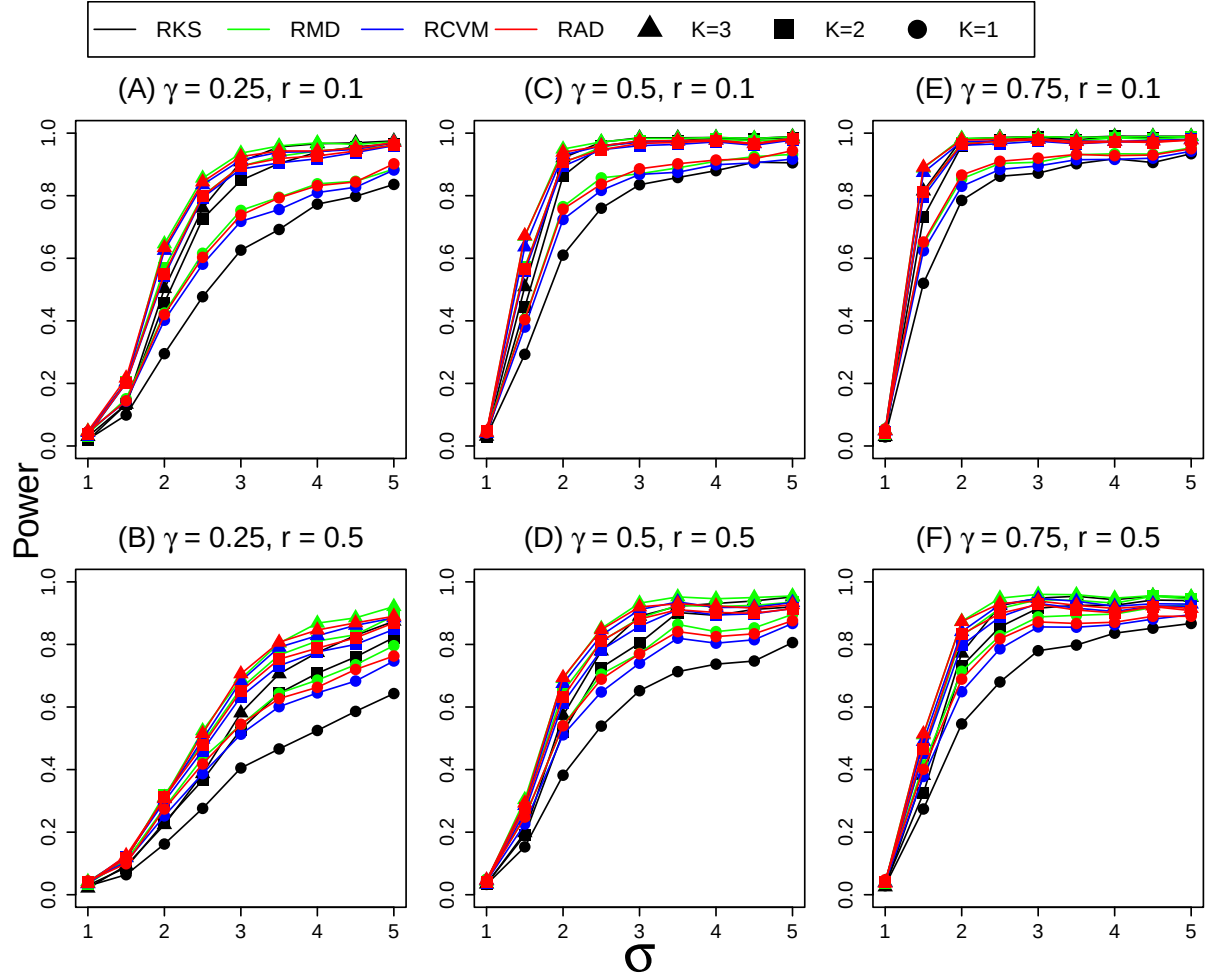


Figure S6: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

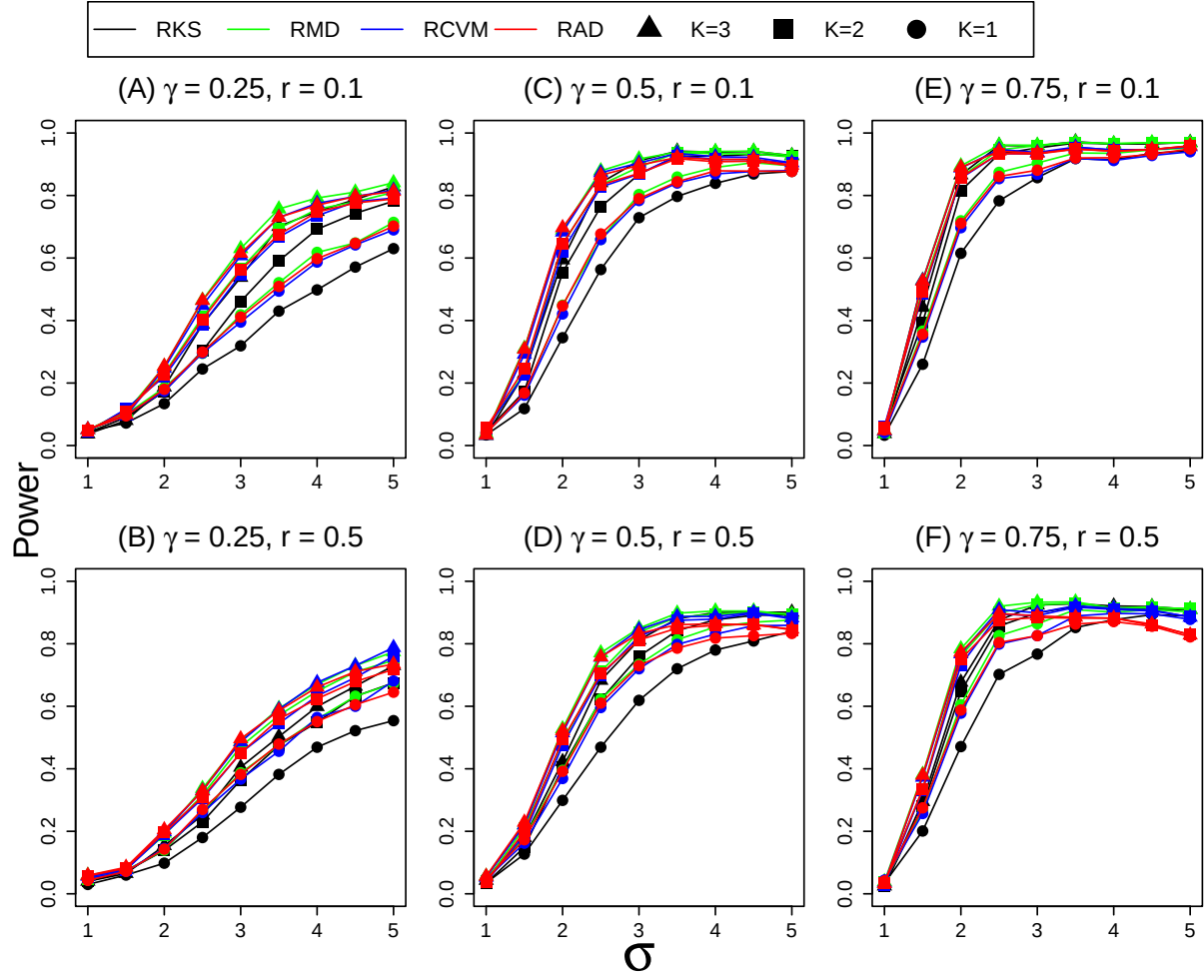


Figure S7: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

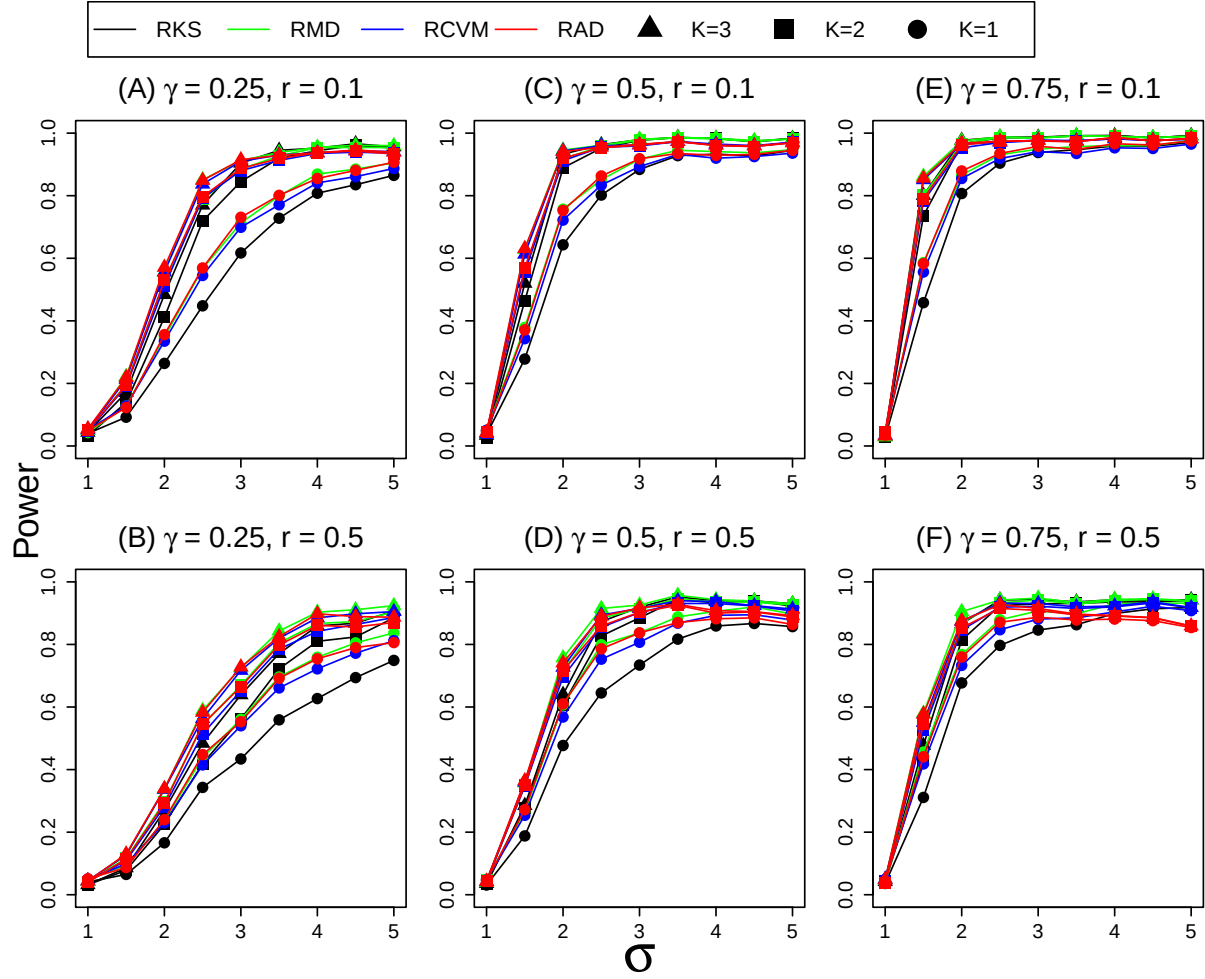


Figure S8: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

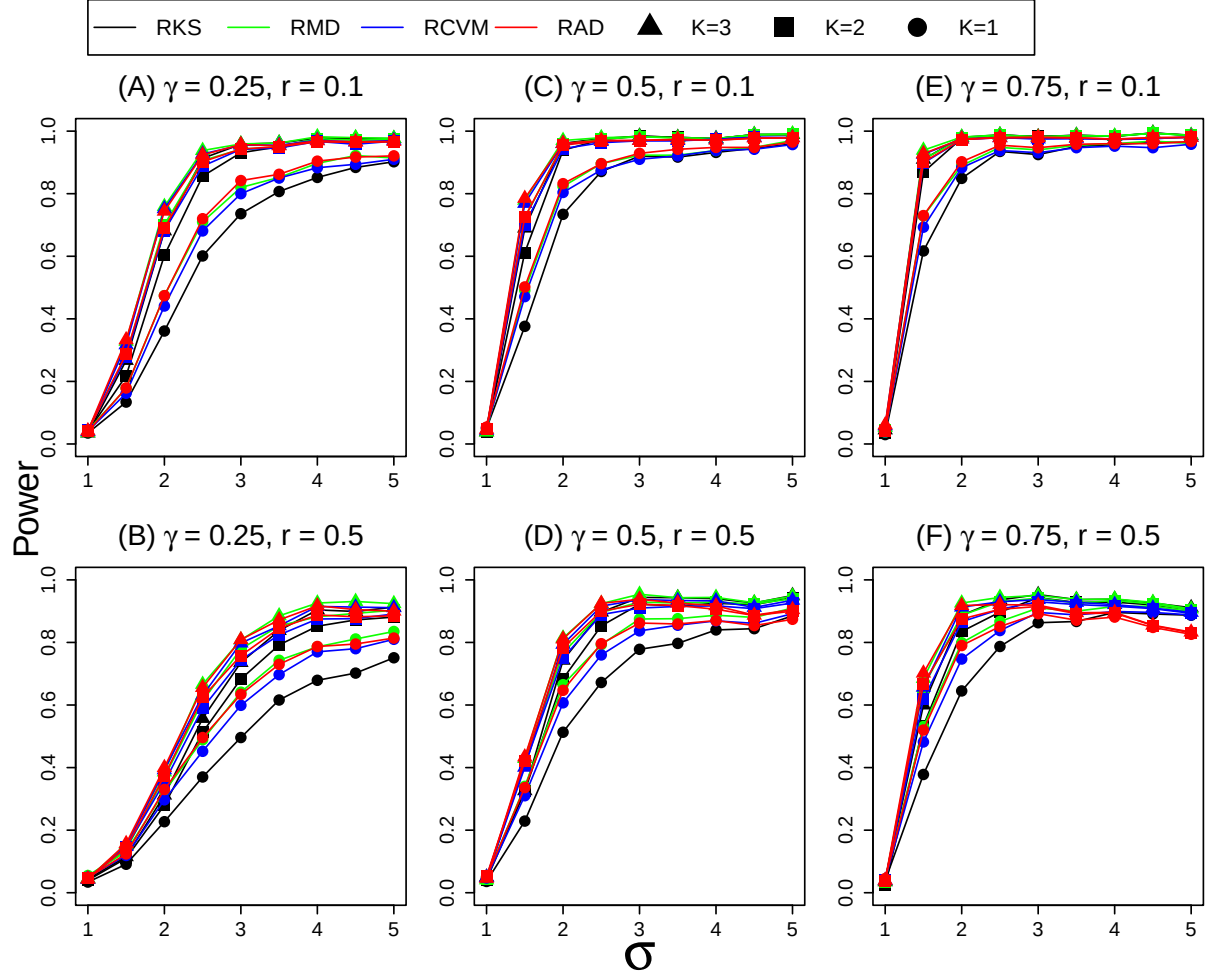


Figure S9: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential variance in a gene set (γ) and inter-gene correlation (r).

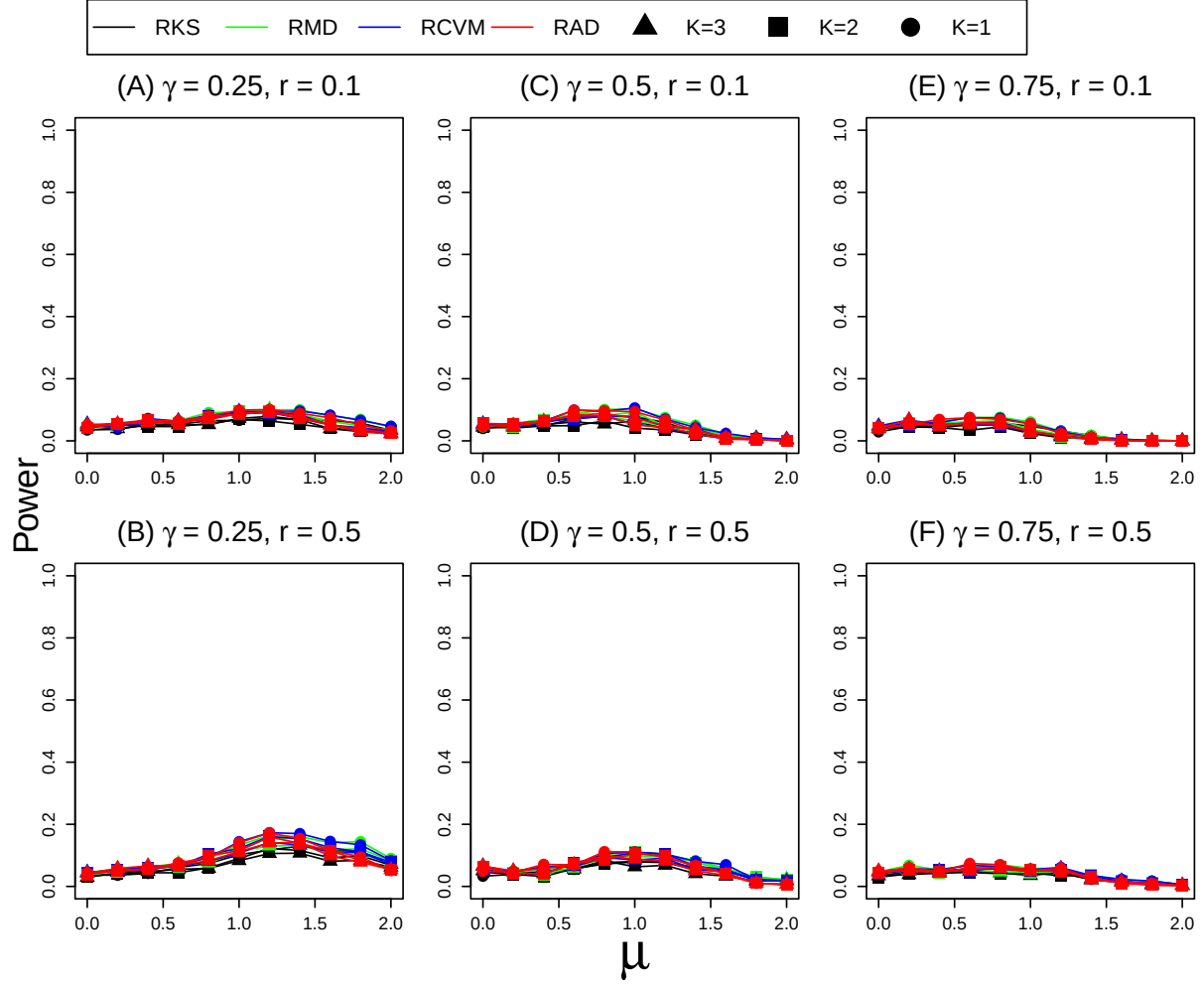


Figure S10: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

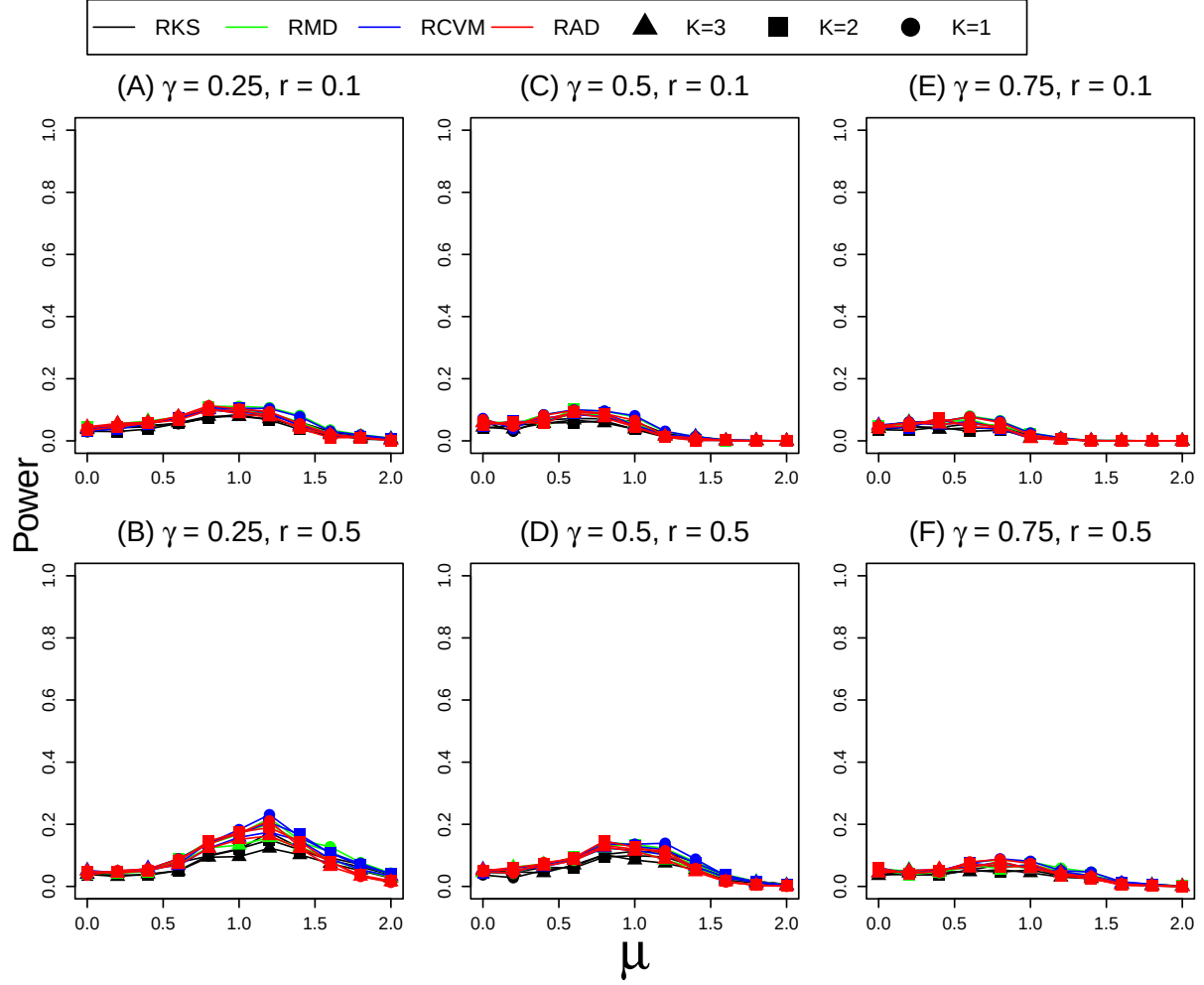


Figure S11: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

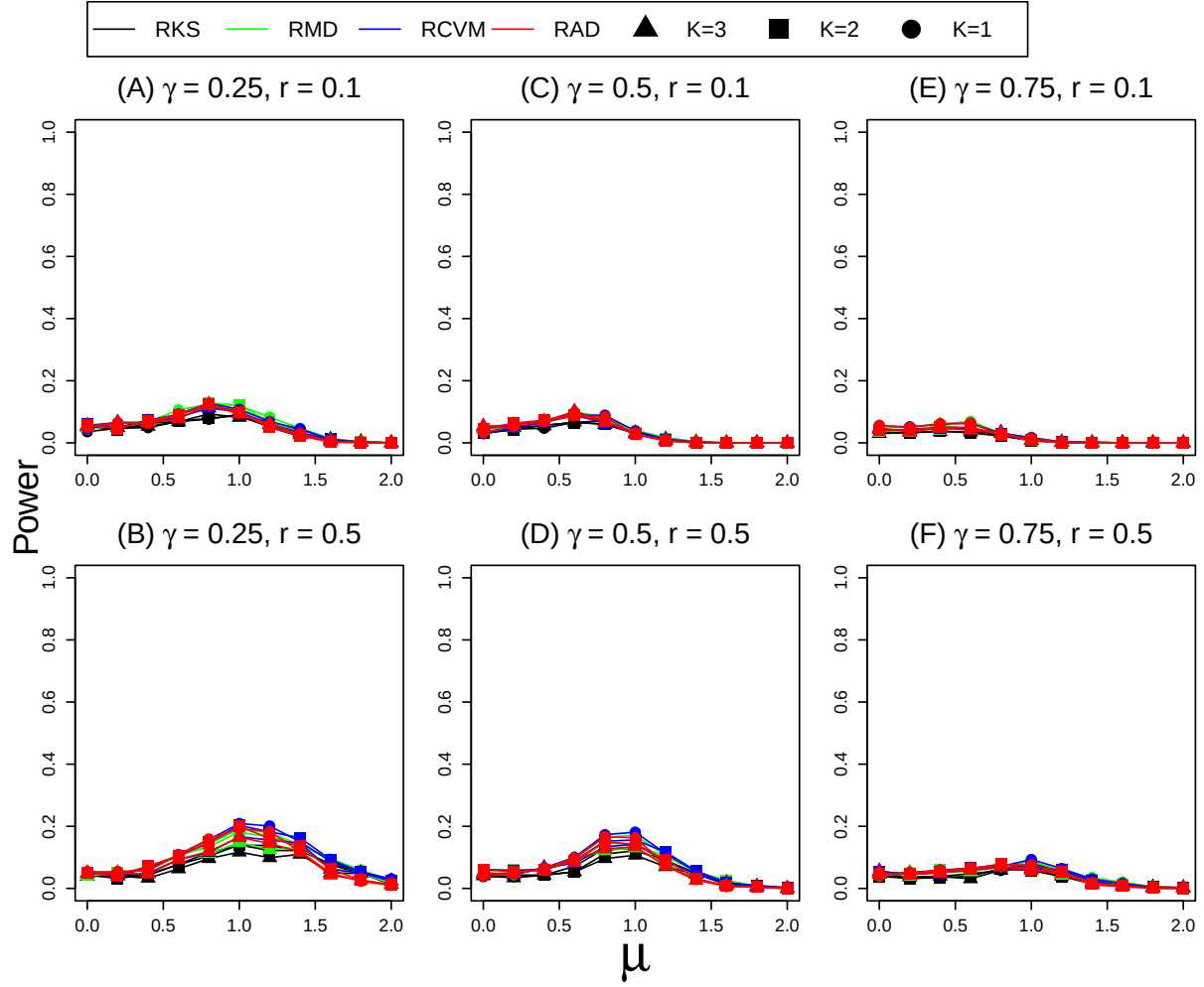


Figure S12: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 20$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

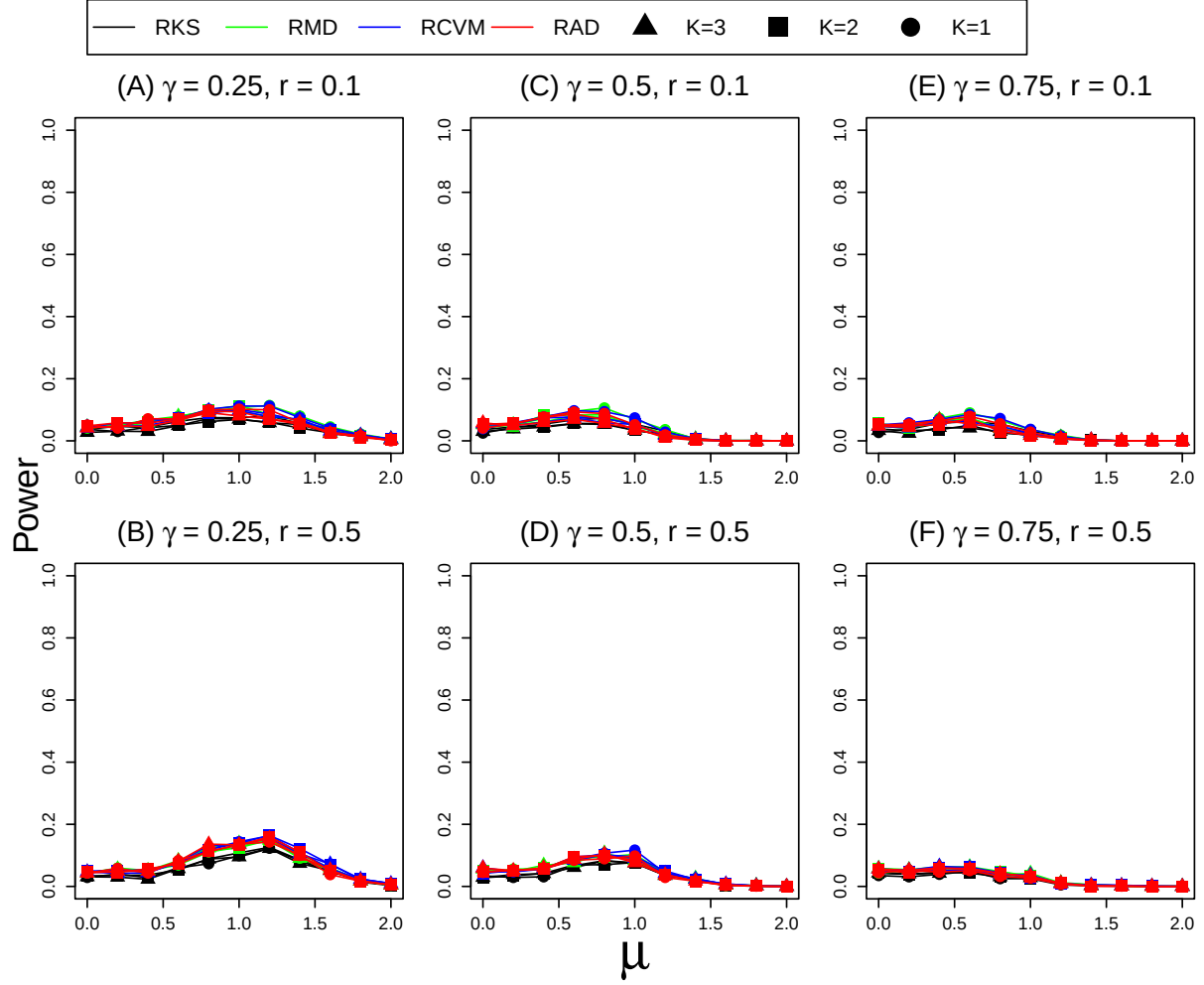


Figure S13: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

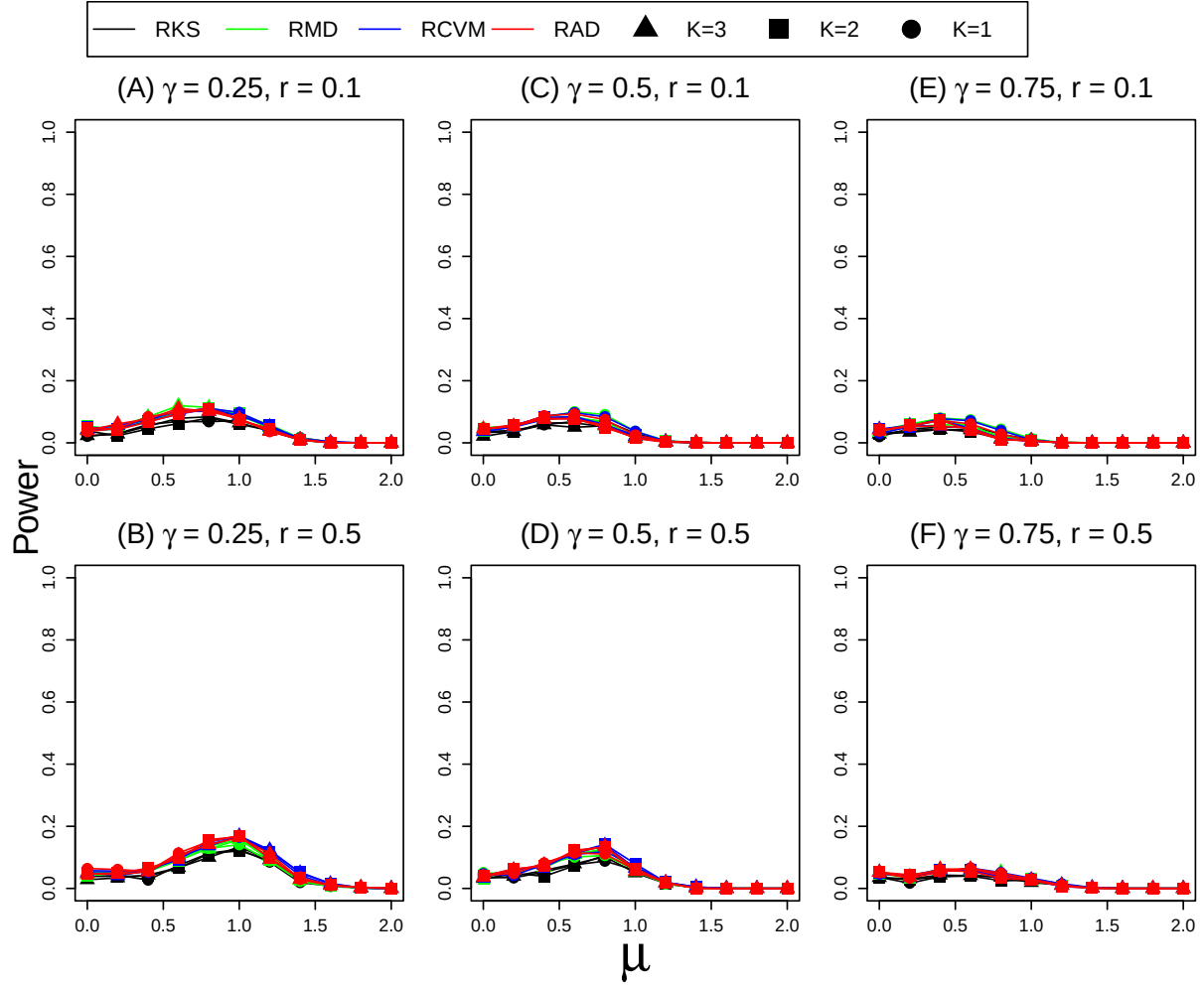


Figure S14: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

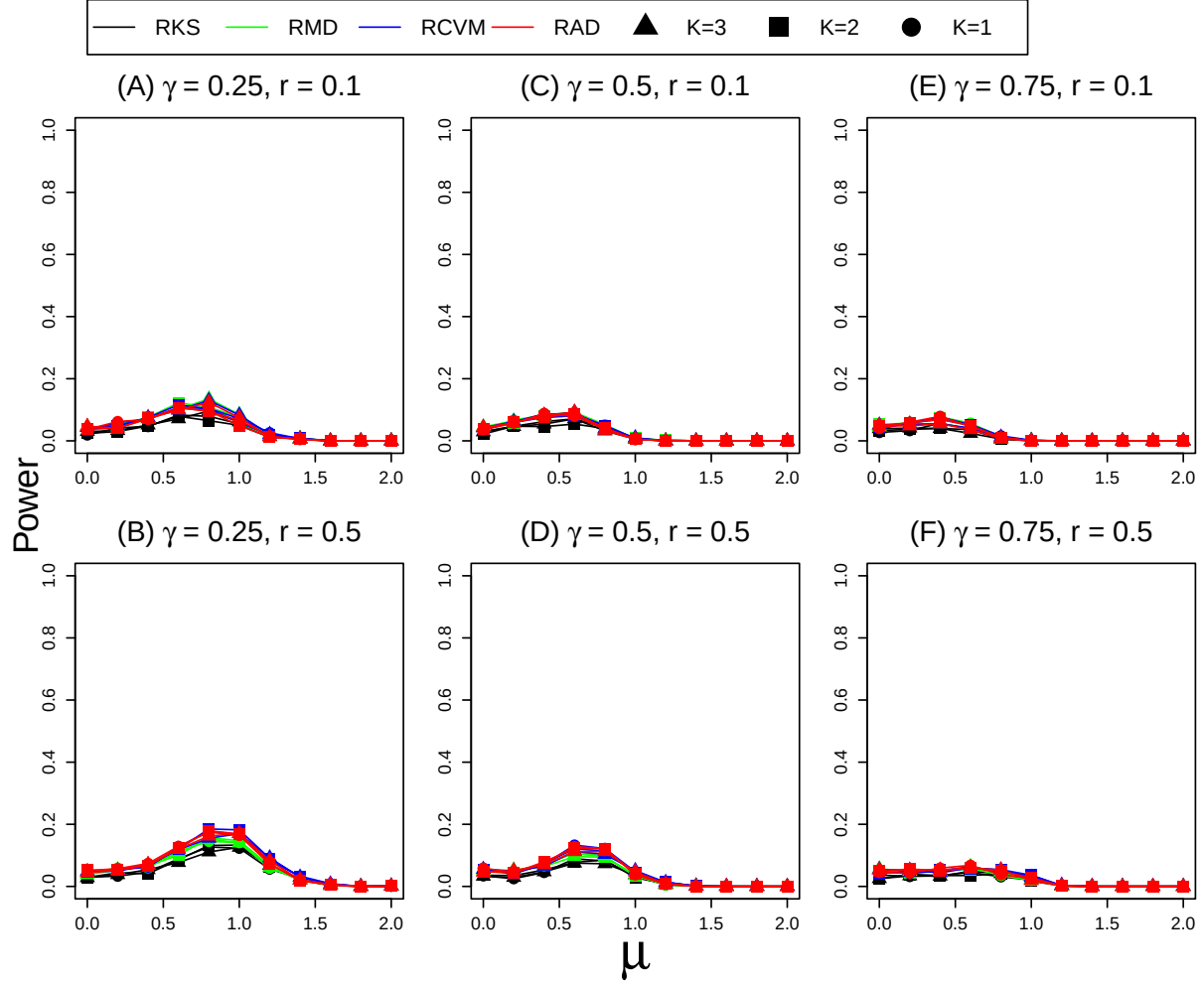


Figure S15: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 40$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

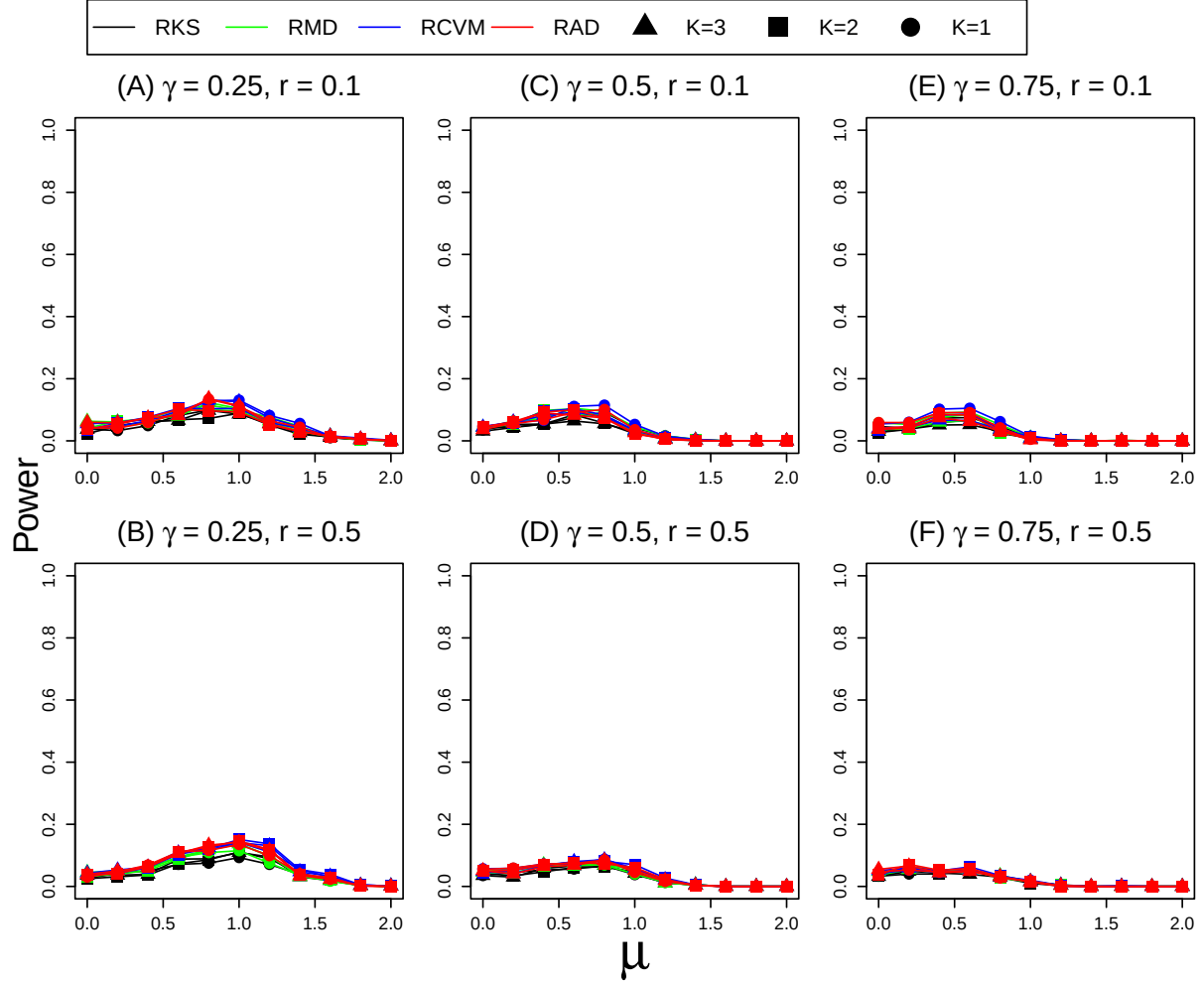


Figure S16: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

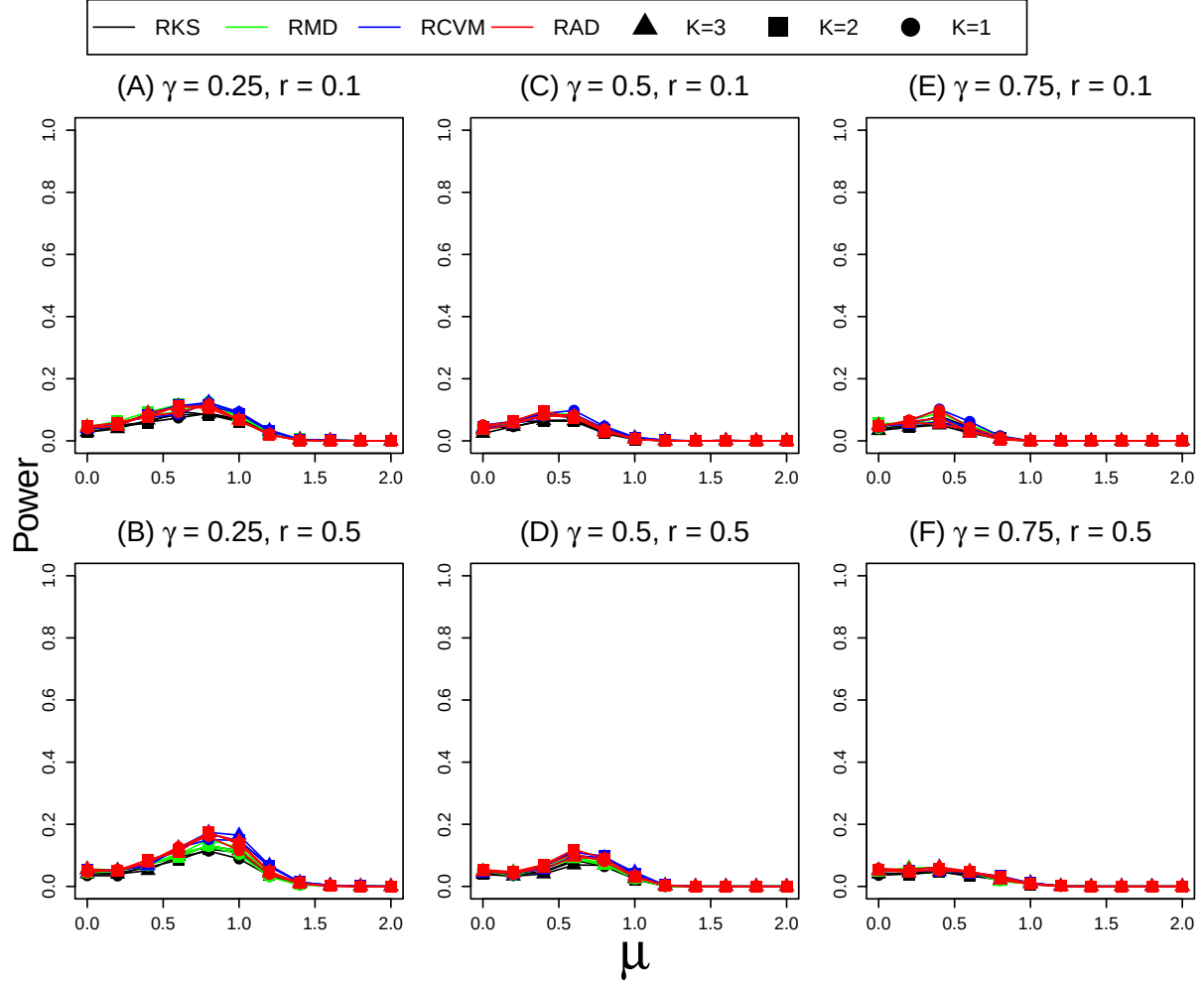


Figure S17: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

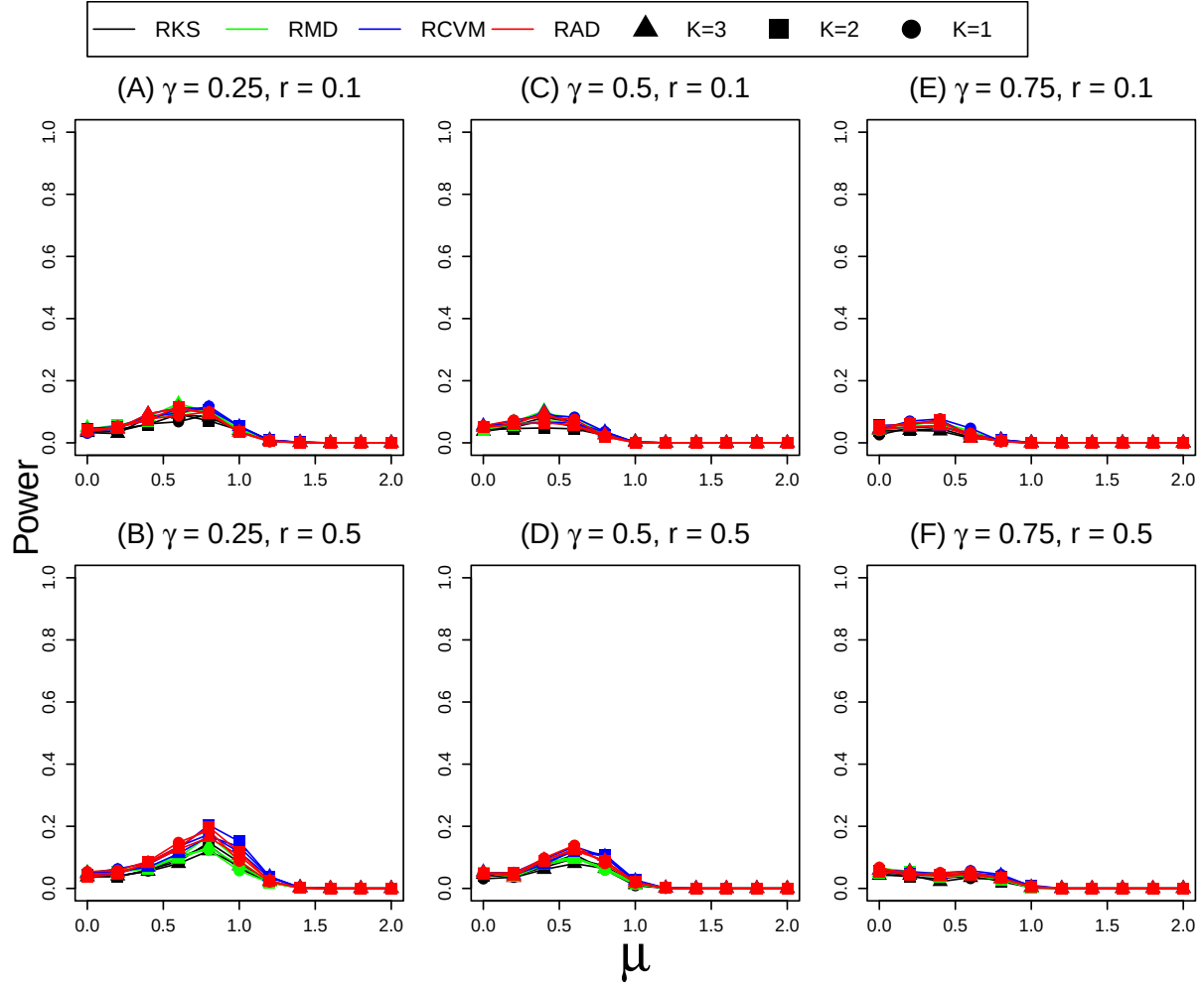


Figure S18: Detection power of different radial ranking-based methods with MST order $K = 1, 2, 3$ ($\alpha = 0.05$) at sample size $N = 60$ and gene set size $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

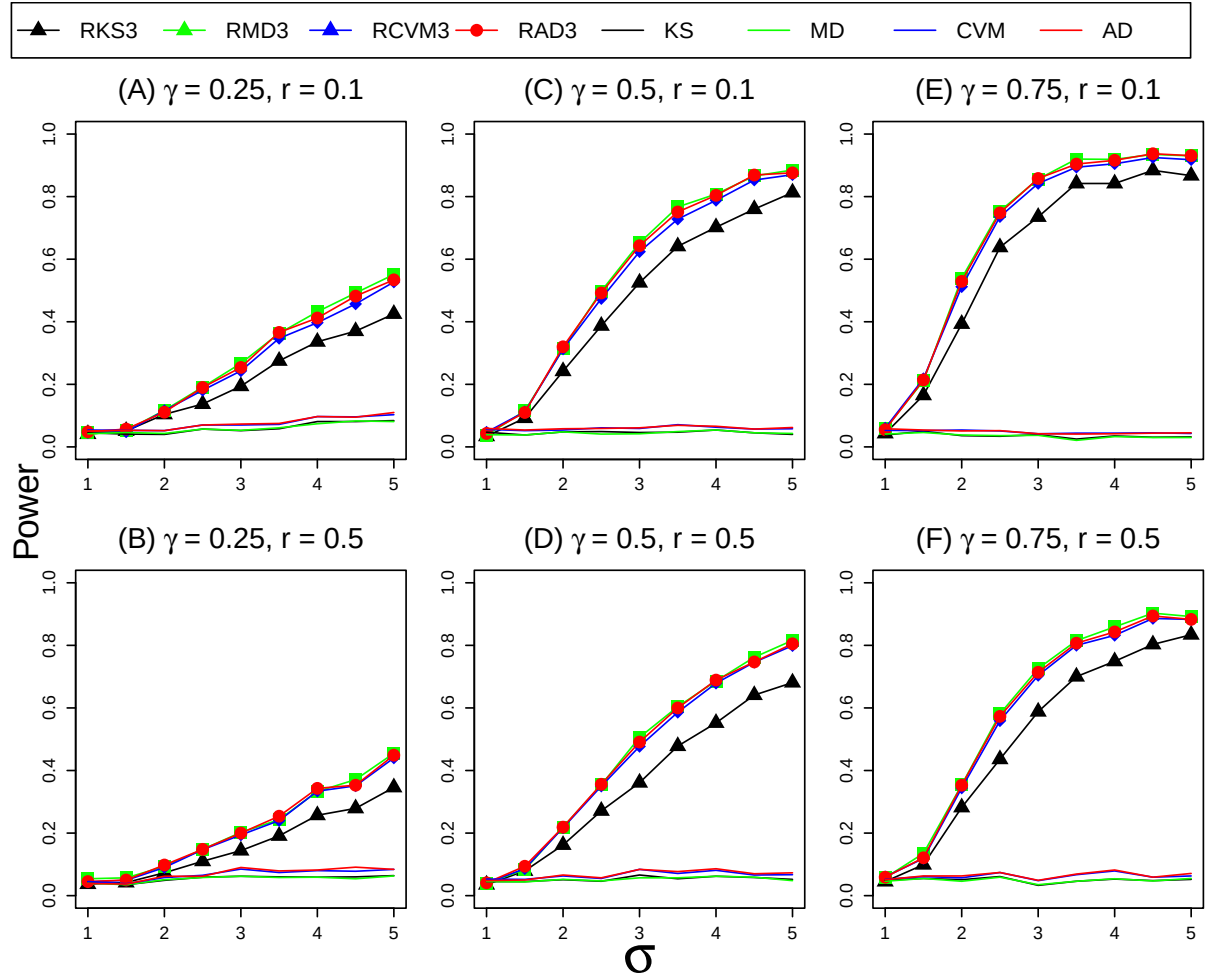


Figure S19: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 20$ when the alternative hypothesis $\sigma_X \neq \sigma_Y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

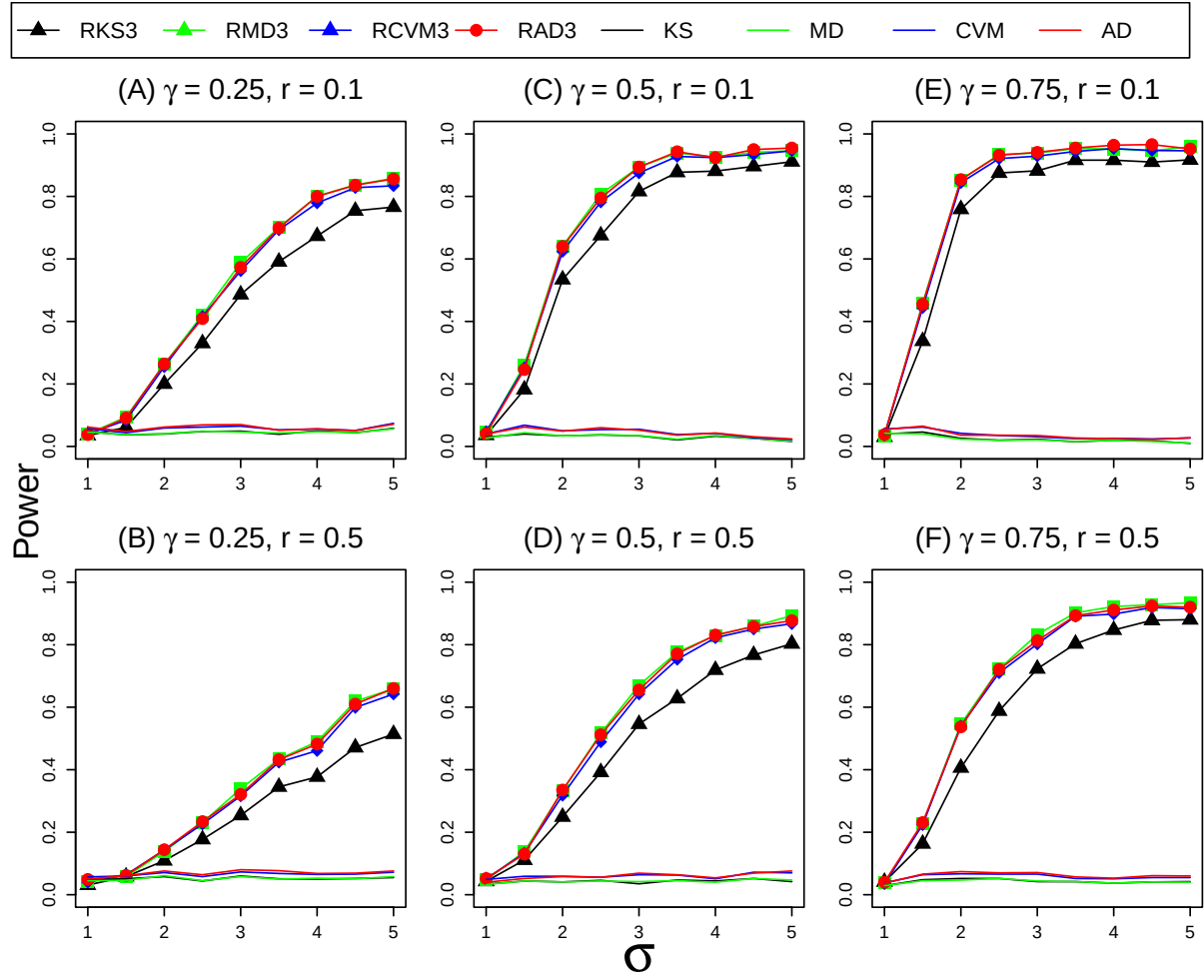


Figure S20: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

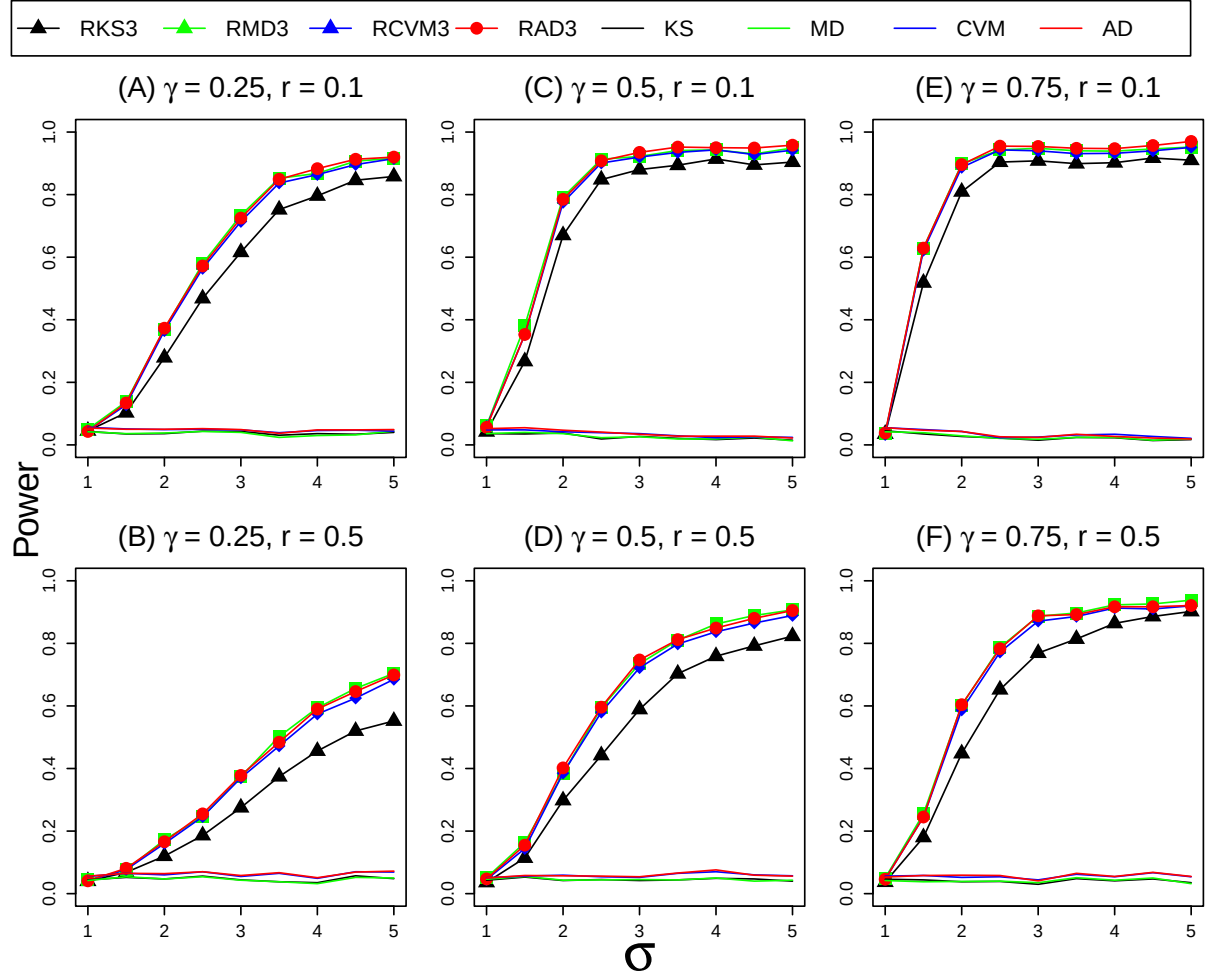


Figure S21: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

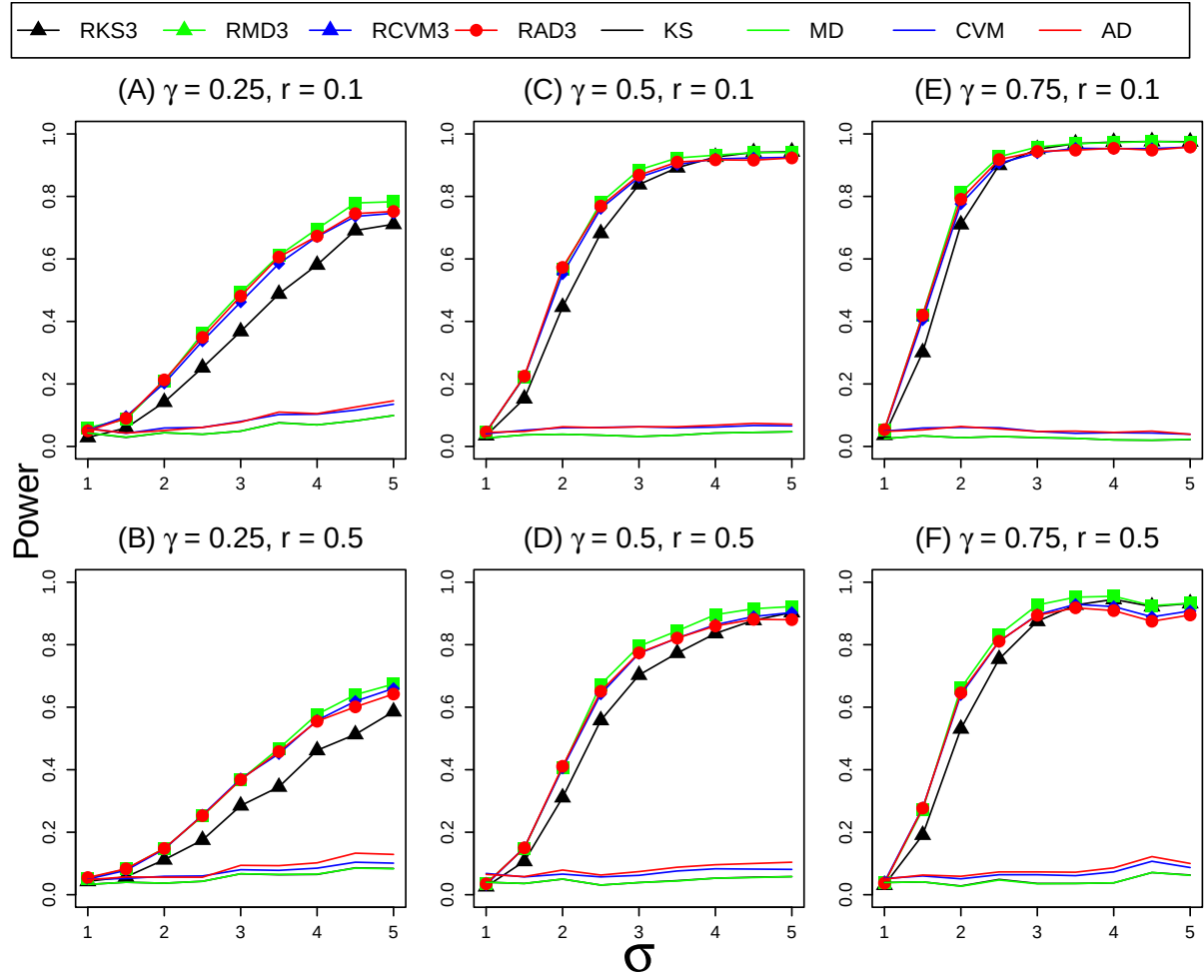


Figure S22: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

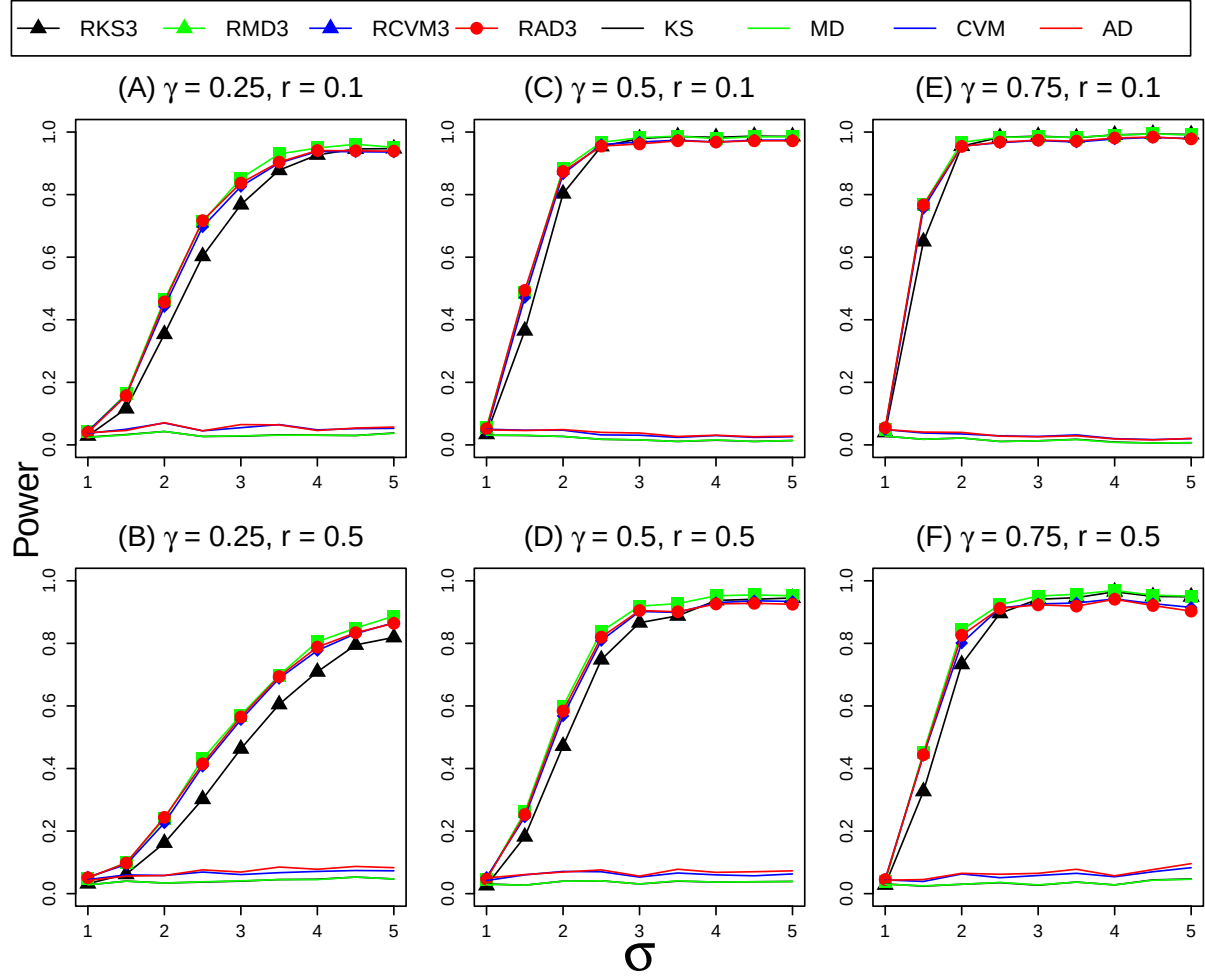


Figure S23: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

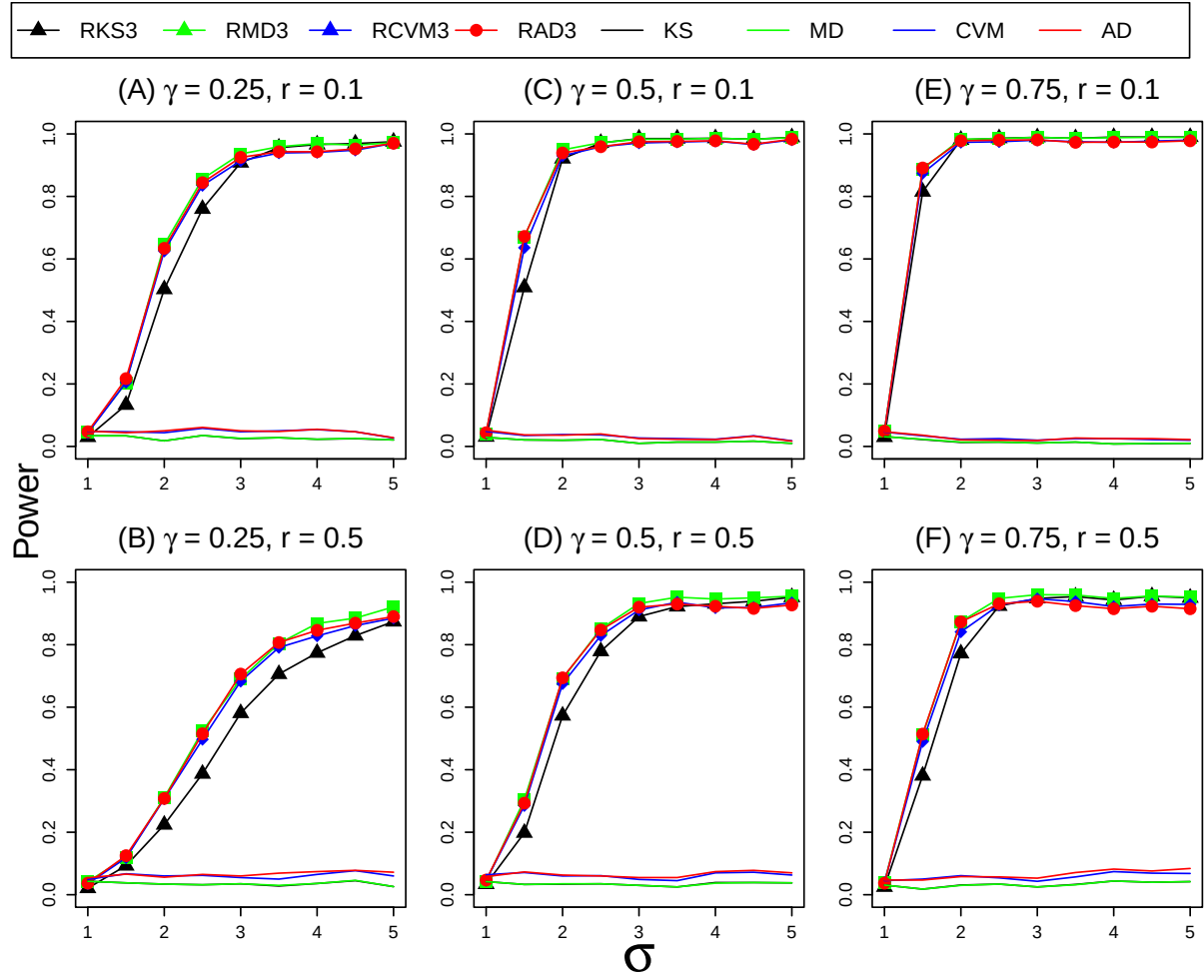


Figure S24: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

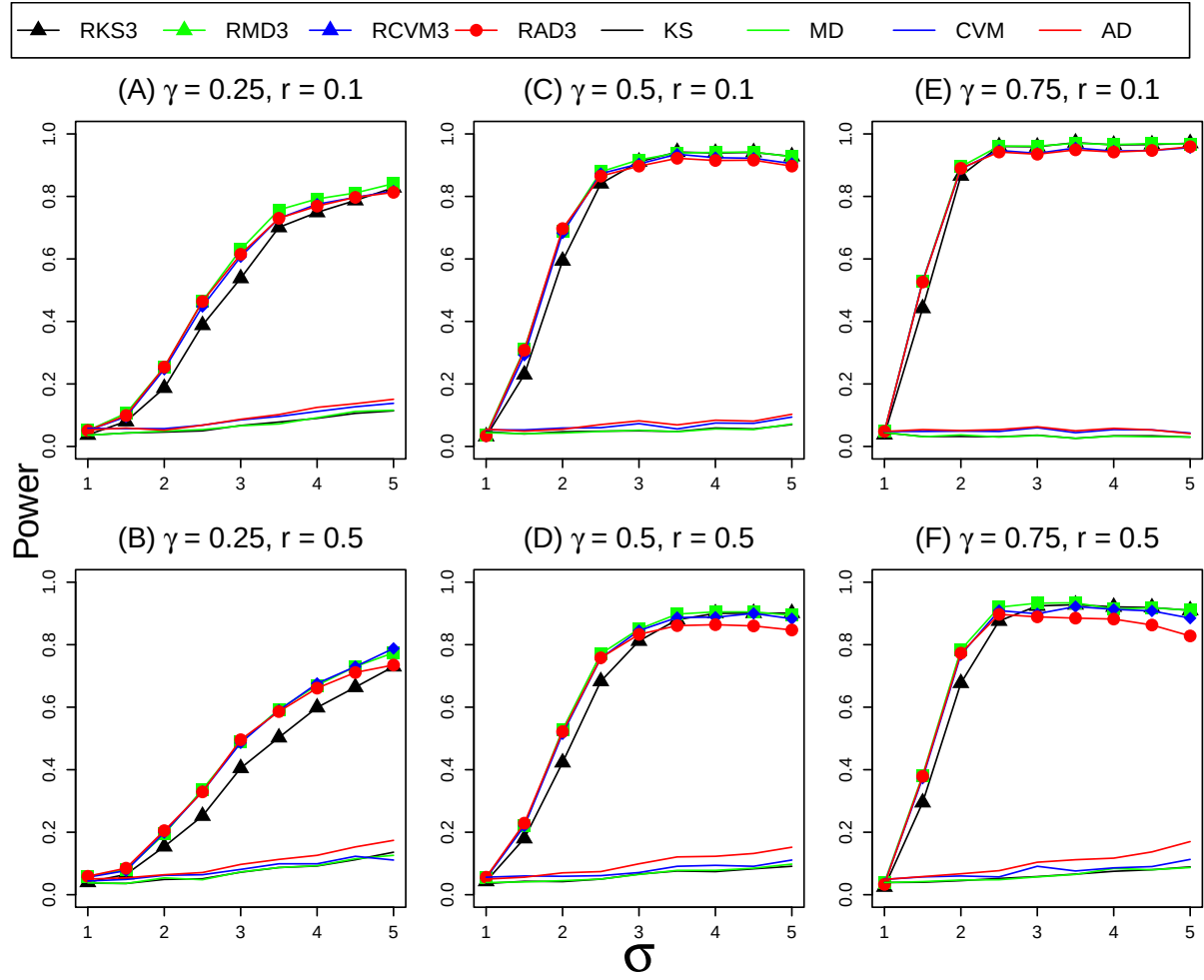


Figure S25: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 20$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

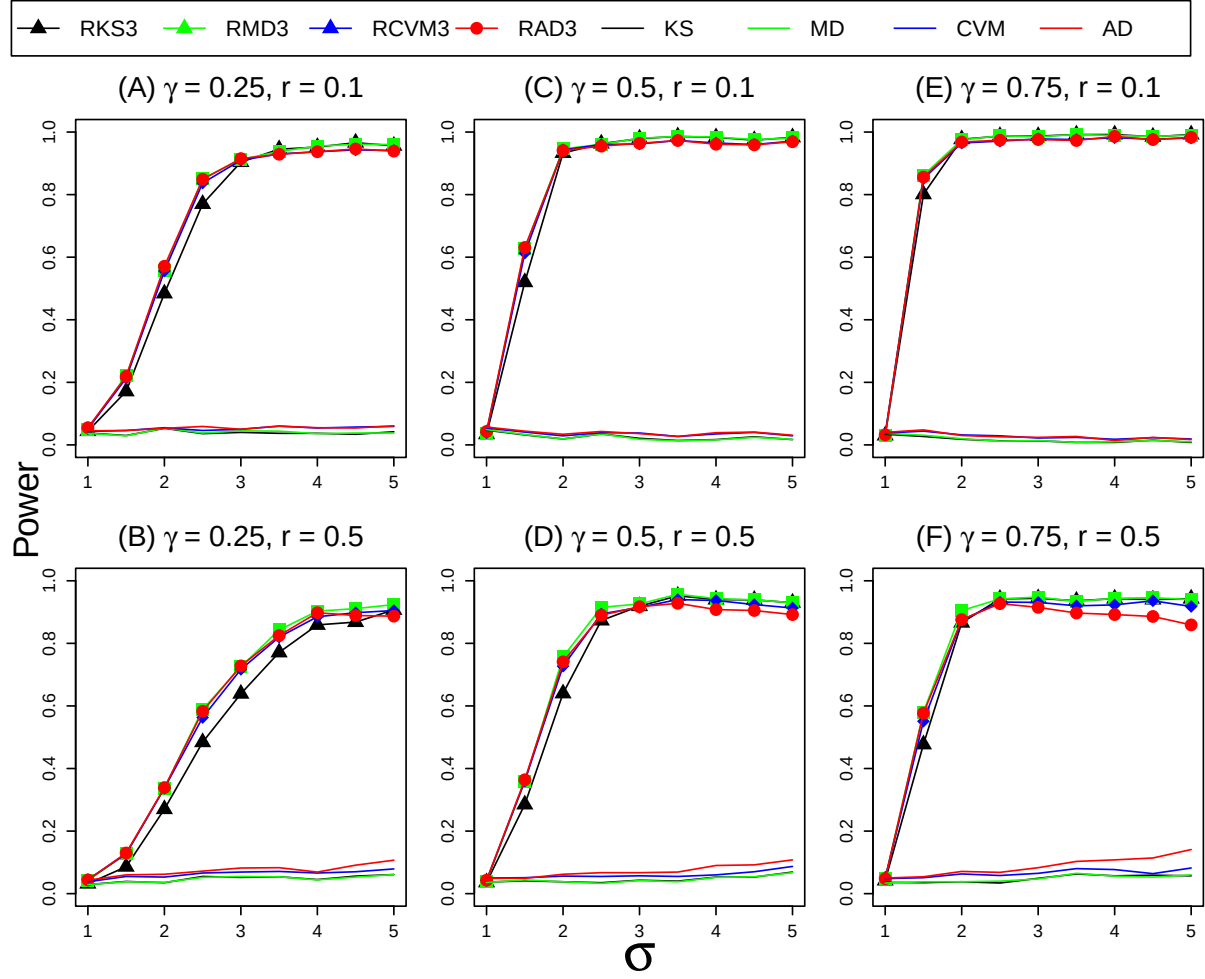


Figure S26: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 60$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

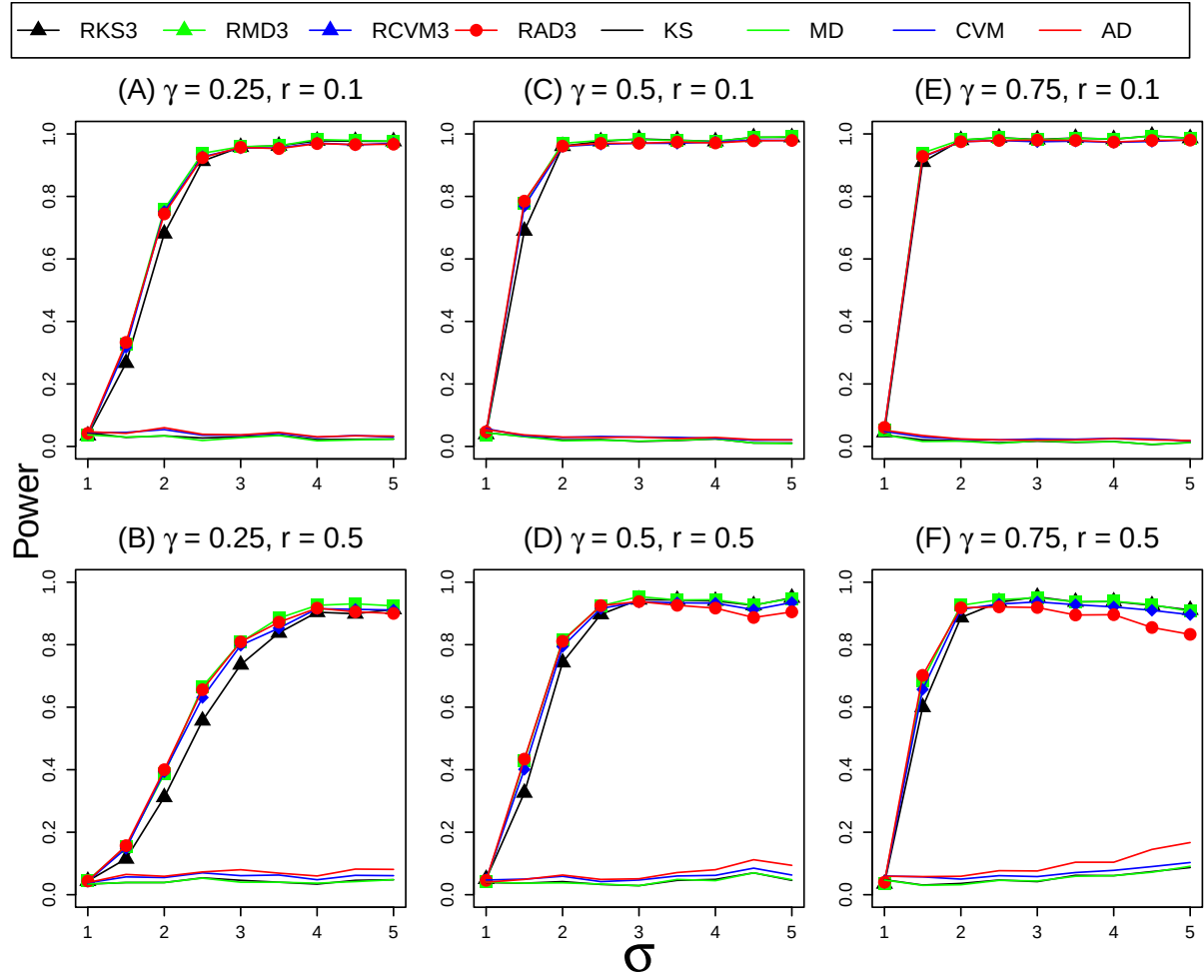


Figure S27: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 100$ when the alternative hypothesis $\sigma_x \neq \sigma_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

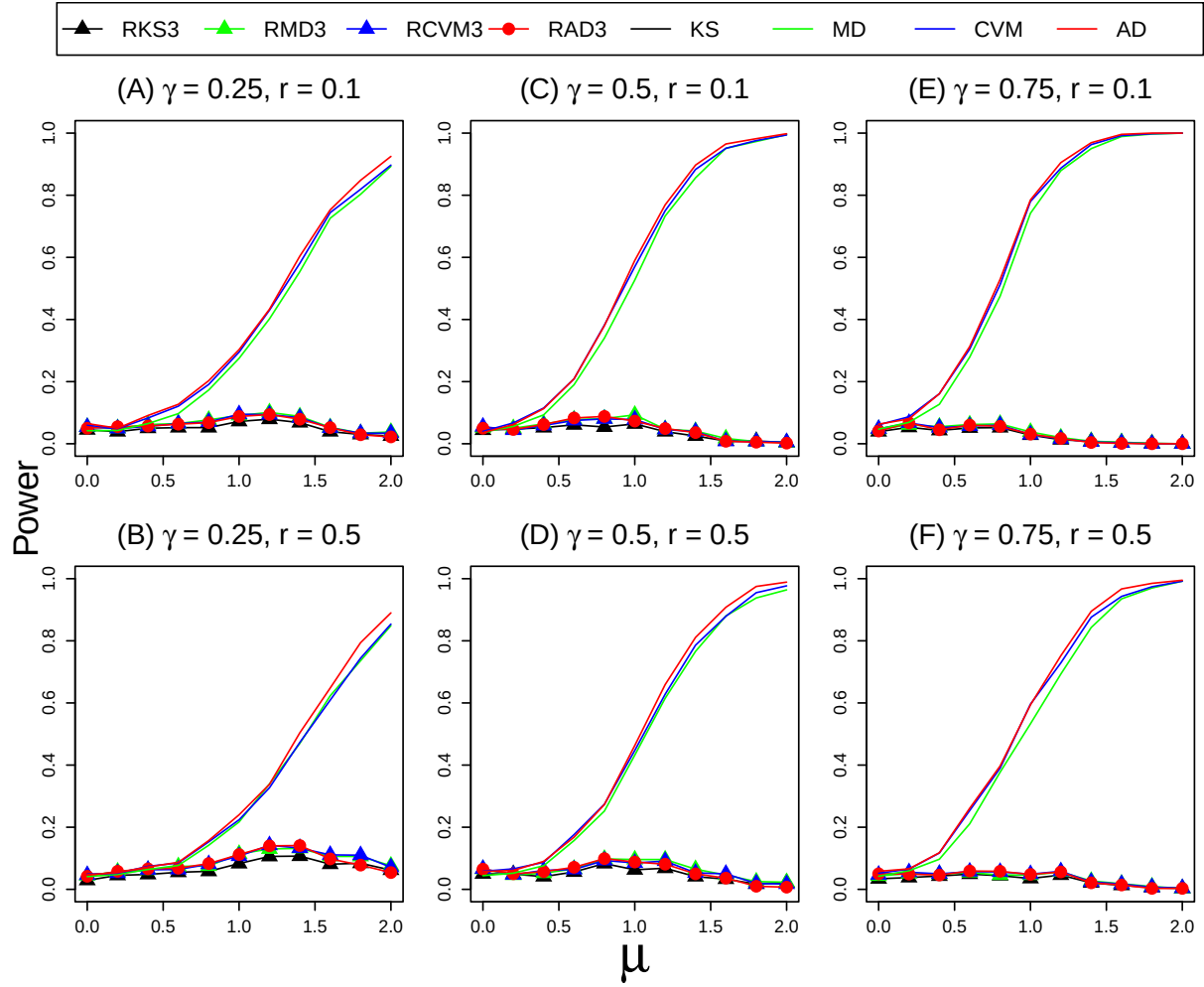


Figure S28: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 20$ when the alternative hypothesis $\mu_X \neq \mu_Y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

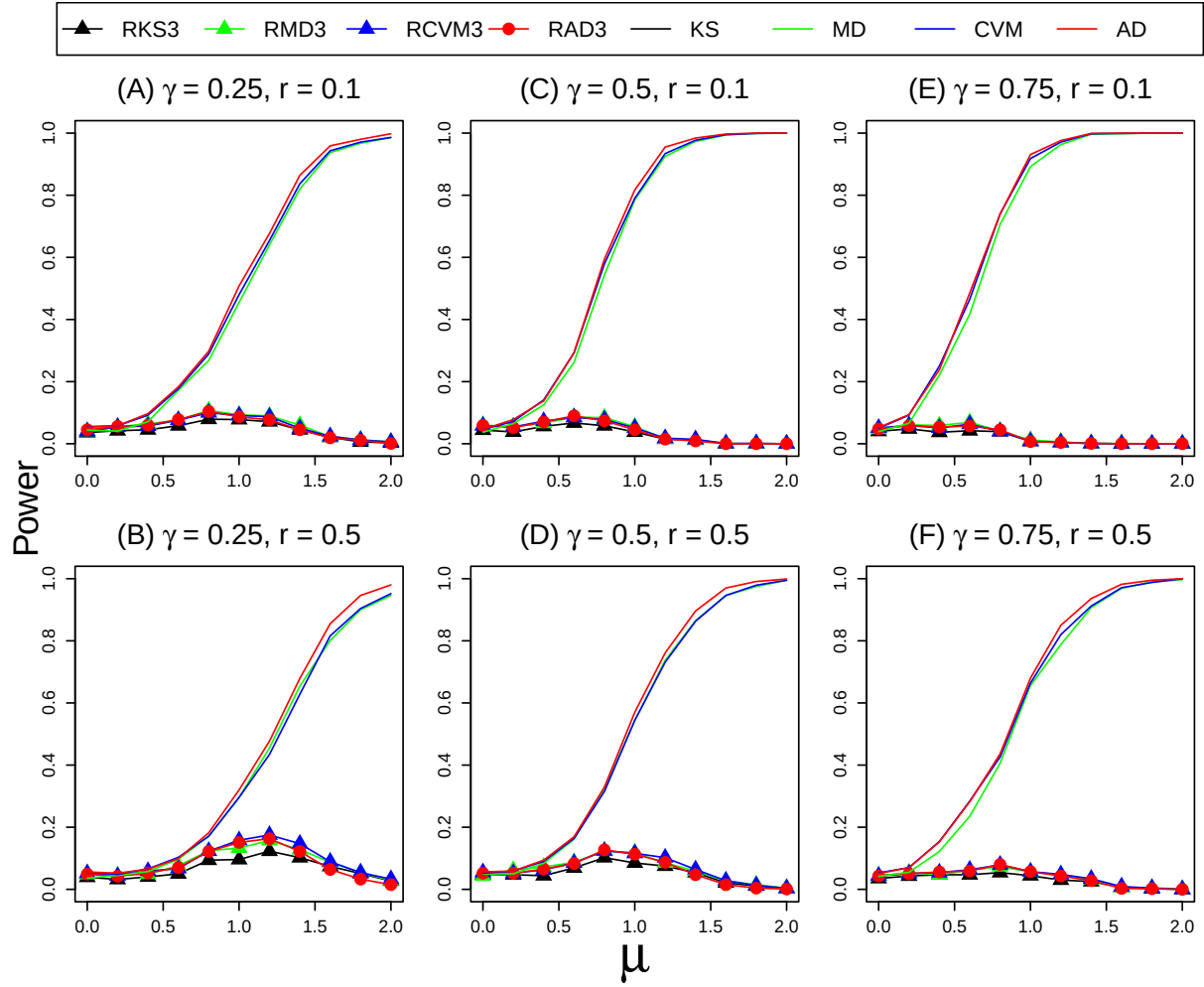


Figure S29: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

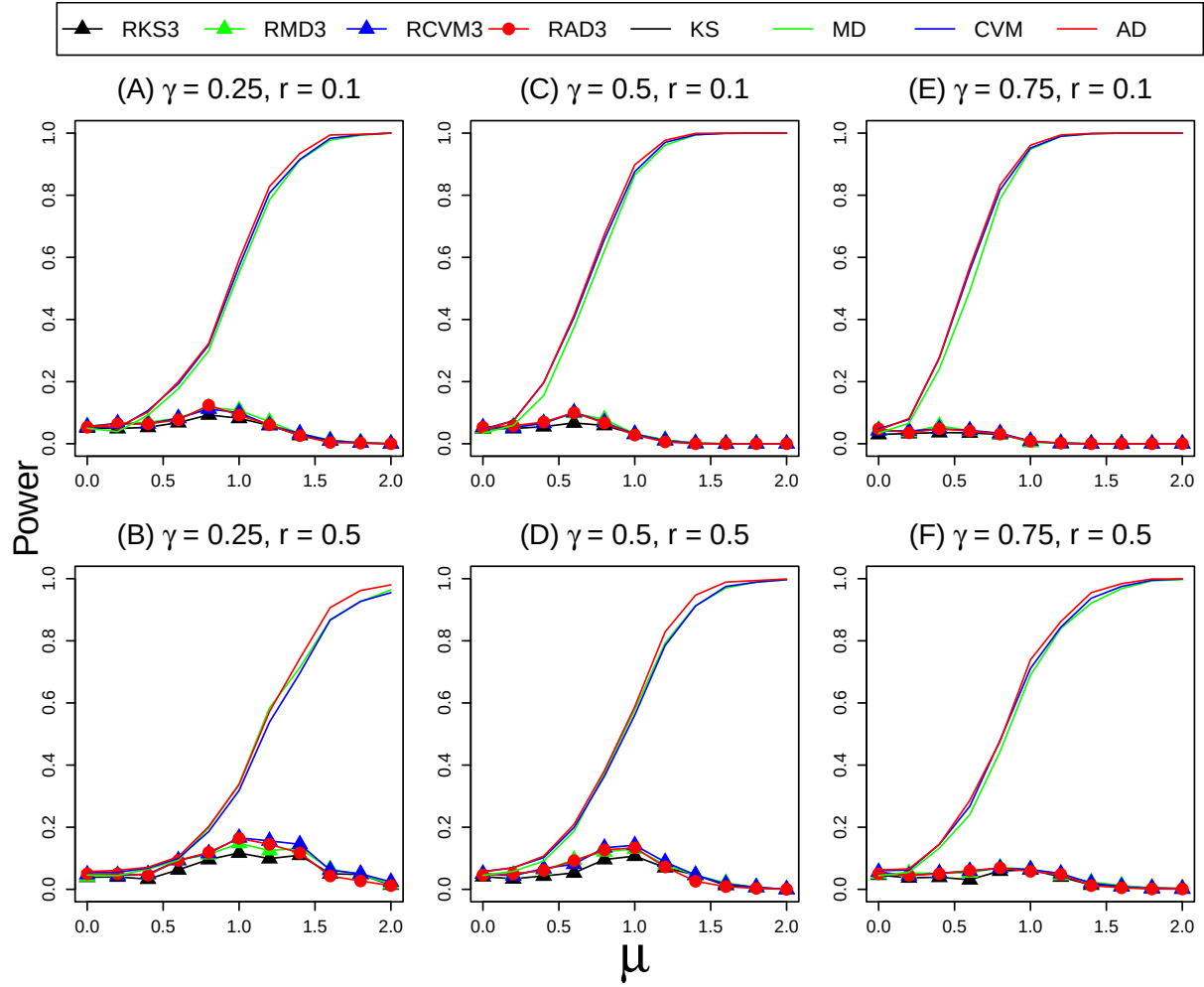


Figure S30: Detection power of different methods ($\alpha = 0.05$) at $N = 20$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

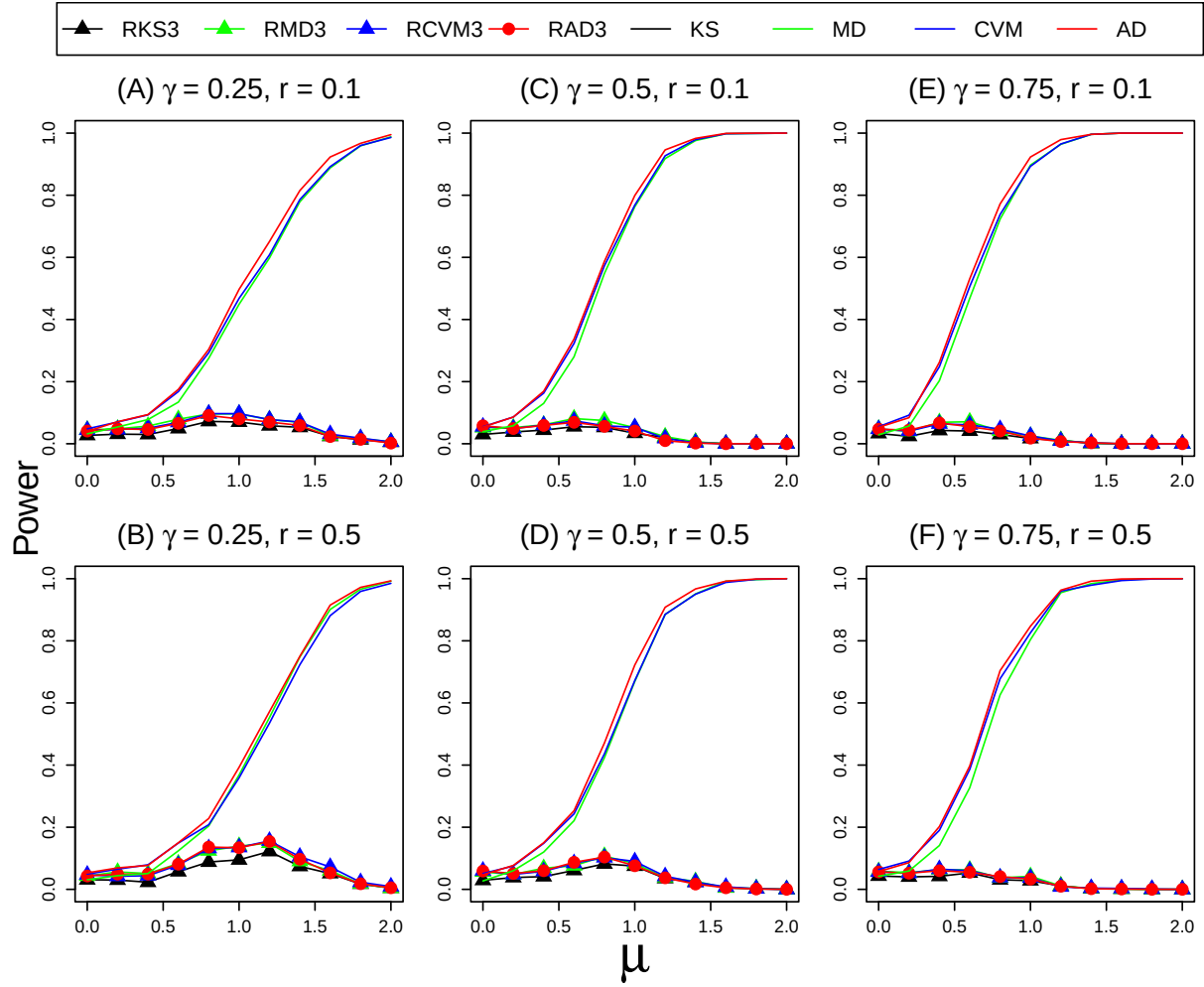


Figure S31: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

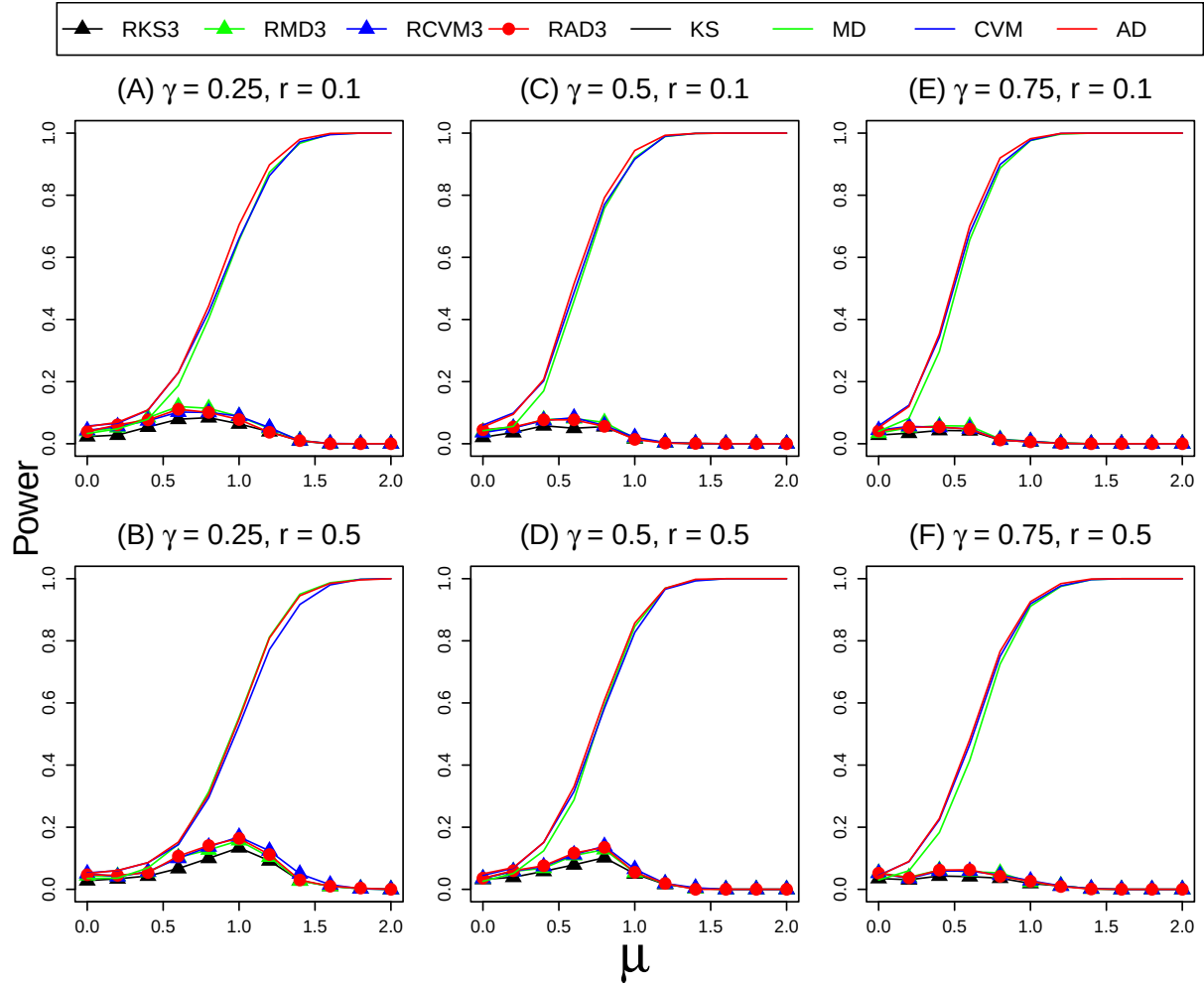


Figure S32: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

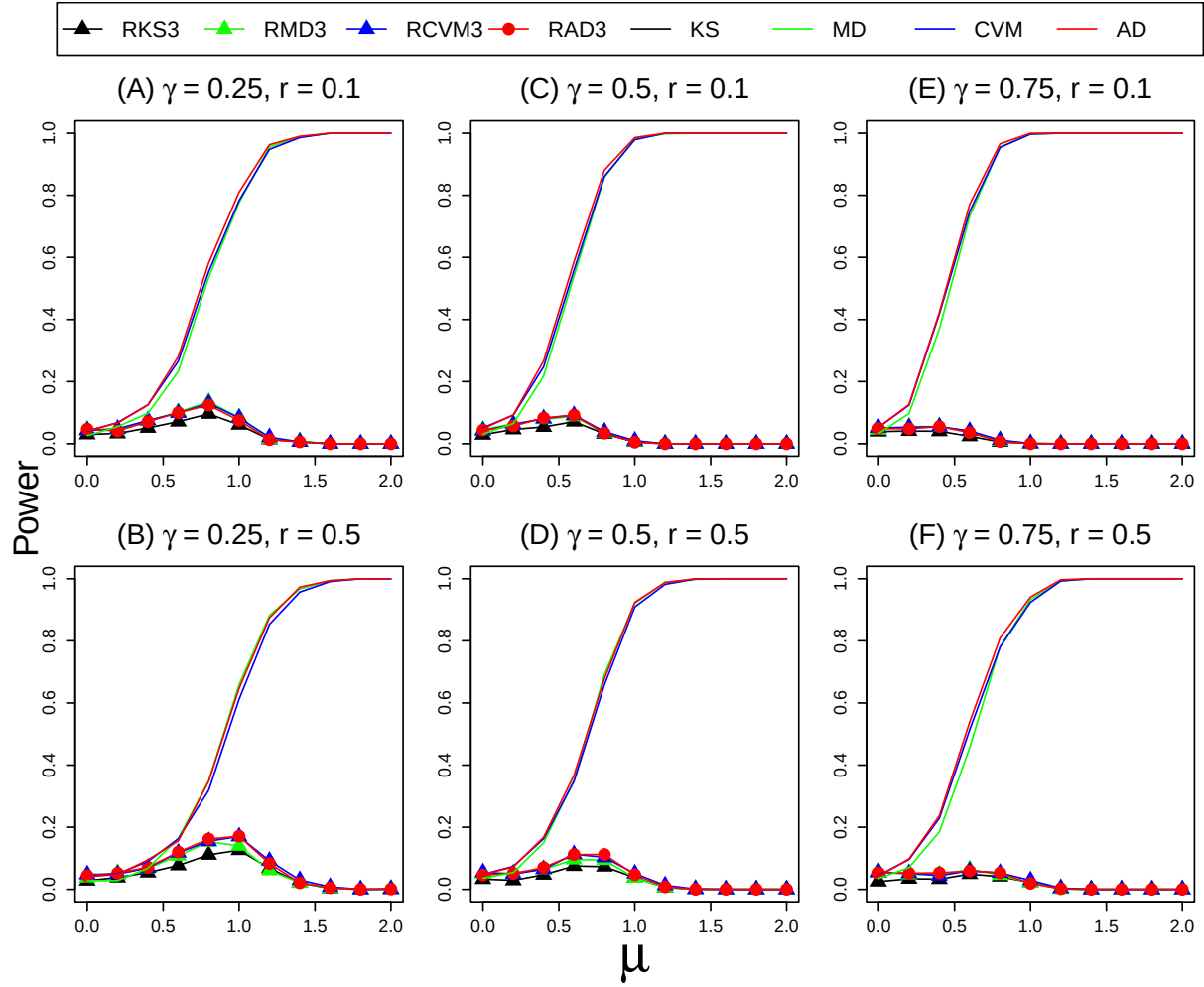


Figure S33: Detection power of different methods ($\alpha = 0.05$) at $N = 40$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

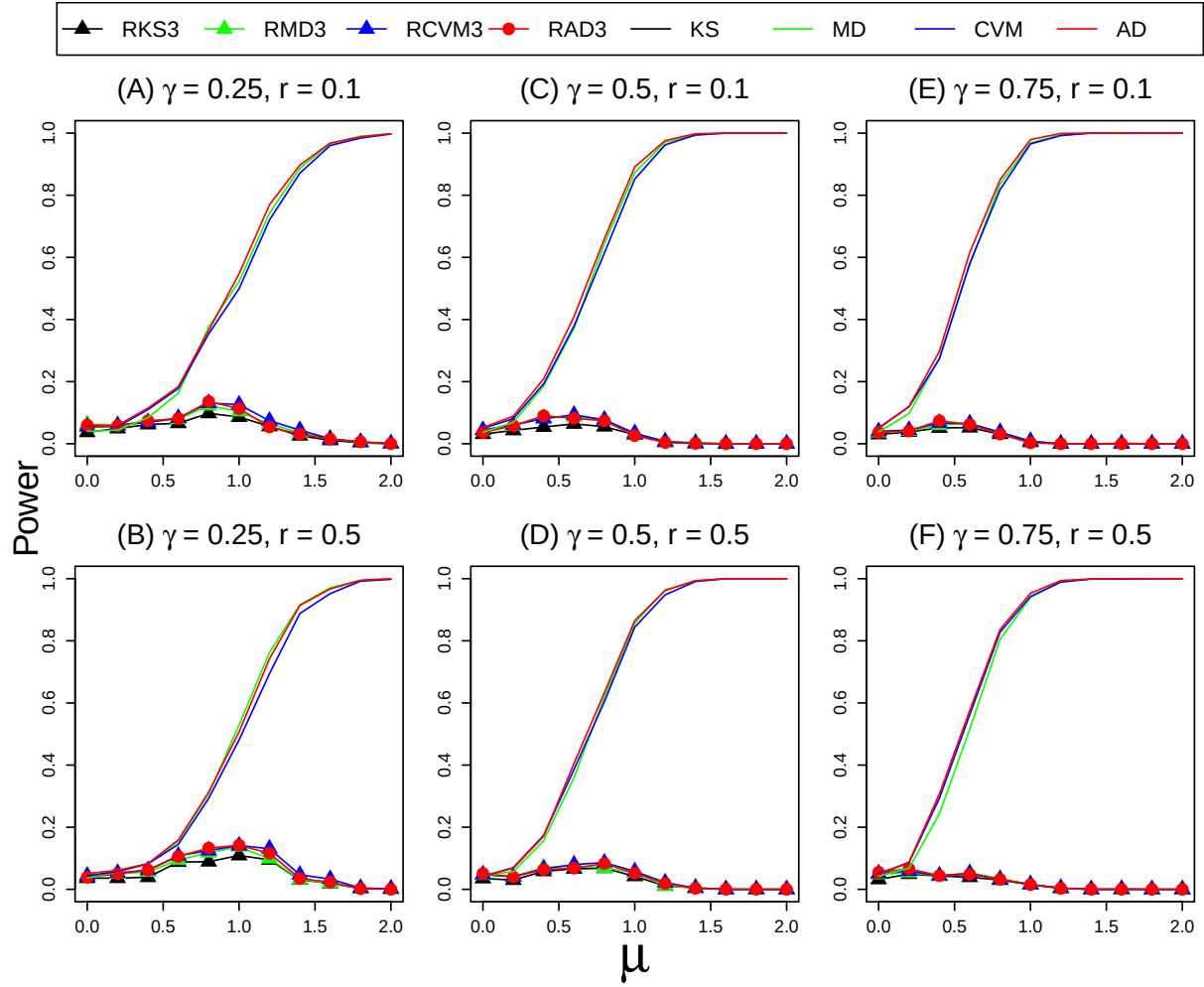


Figure S34: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 20$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

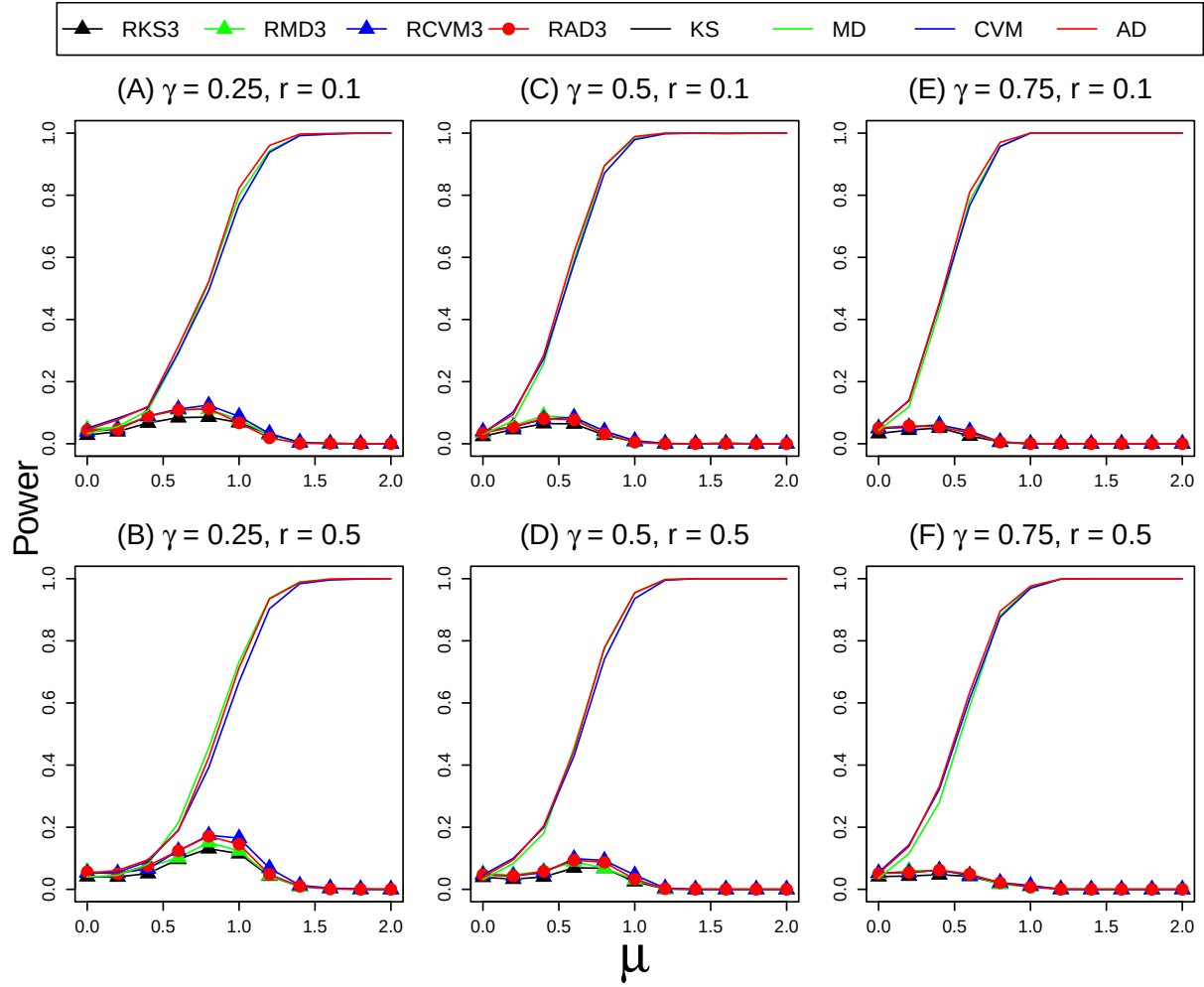


Figure S35: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 60$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).

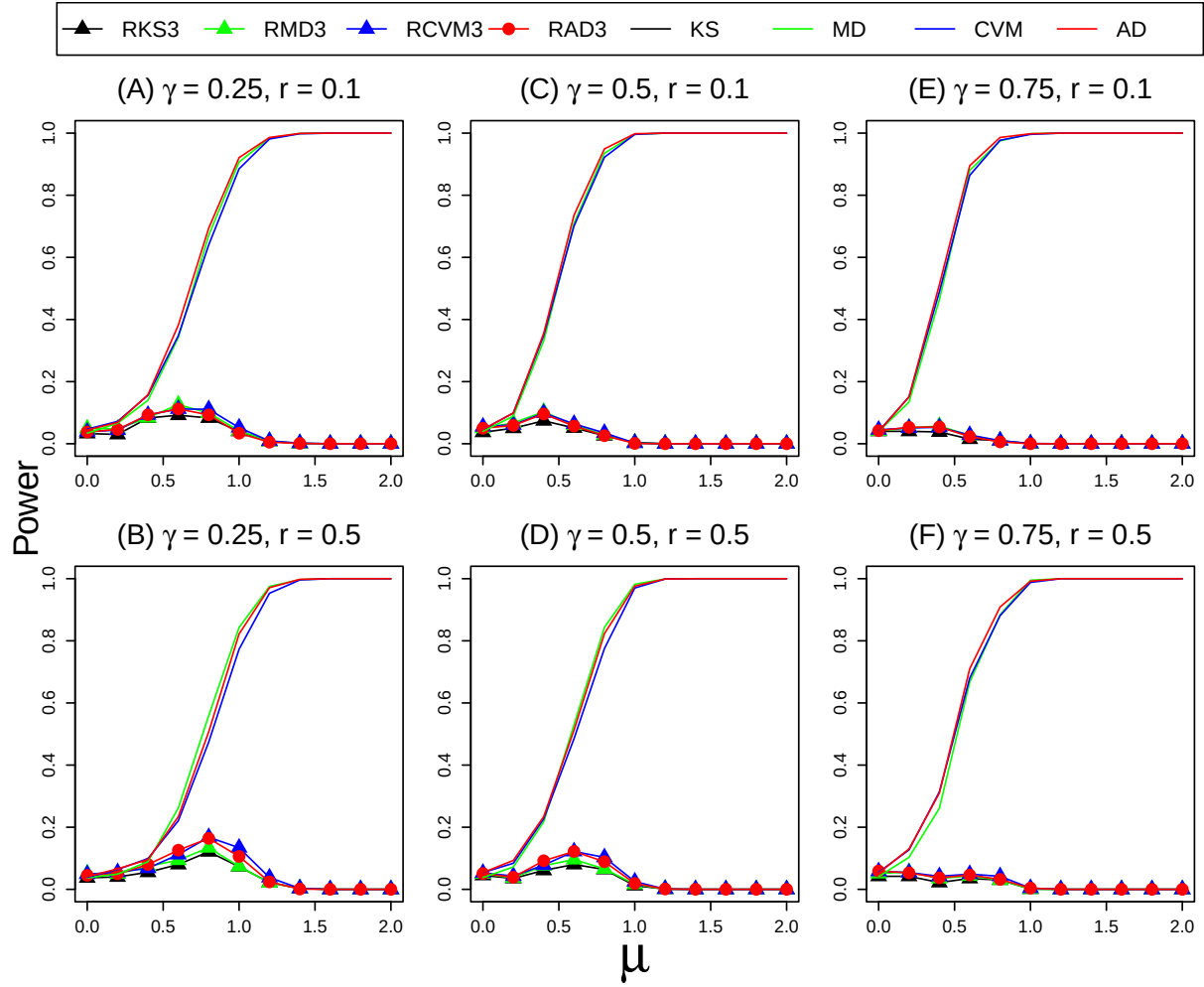


Figure S36: Detection power of different methods ($\alpha = 0.05$) at $N = 60$ and $p = 100$ when the alternative hypothesis $\mu_x \neq \mu_y$ is true with different proportions of genes with truly differential mean in a gene set (γ) and inter-gene correlation (r).