Comprehensive performance evaluation of interaction detection methods

Step 1: assess the accuracy of $P$-value based significance assessment

Simulation 1: no ground-truth SNP

Simulation 2: interacting SNPs only

Simulation 3: main-effect SNPs only

Compare family-wise error rate to $P$ value threshold

Different conservativeness level

Compare power based on statistical significance

Compare power based on SNP ranking

Compare the number of false positive interactions

Power definition 1: Interaction detection power

Power definition 2: Exact interaction detection power

Power definition 3: Partial interaction detection power

Power definition 4: Individual SNP detection power

Analytical results on the relation between detection power and the factors (Penetrance, MAF, LD)