We let $\mu_{ij} = 0 \cdot (1 - h_{ij}) + \mu_i \cdot h_{ij}$, where $i \in \{1, 2\}$, and $\mu_i \in \{0.5, 1, \ldots, 5\}$. We set $m = 1000$, and $f_{ij} = |I_{ij}|/m$ for $i, j \in \{0, 1\}$ as follows: $f_{00} = 0.9$, $f_{11} = 0.1$; $f_{00} = 0.9$, $f_{01} = f_{10} = 0.025$, $f_{11} = 0.05$; $f_{01} = f_{10} = 0.5$; $f_{00} = 0.8$, $f_{01} = f_{10} = 0.1$. The standard deviations σ_1 and σ_2 were either fixed values $\sigma_i \in \{0.3, 1\}$, $i \in \{1, 2\}$, or reflected the fraction of sample size allocated to the first study: $\sigma_1 = \sigma/\sqrt{\zeta N}$, $\sigma_2 = \sigma/\sqrt{(1-\zeta)N}$, $\sigma = 10$, $\zeta \in \{0.1, 0.2, \ldots, 0.9\}$, $N = 1000$.

The simulation results were based on 1000 repetitions. The FDR was estimated by averaging the FDP. The average power was estimated by the average number of rejected false no replicability null hypotheses, divided by mf_{11}.

6.1 Simulation results

As expected from our theoretical results, in all the settings considered the estimated FDR was below 0.05 for all procedures but the naive (BH-i, BH-j) procedure. The SE of the estimated FDR and power were of the order of 10^{-3} for all procedures under all configurations considered.

Figure 1 compares the power of the BH procedure on maximum p-values, (1) above, and Procedure 4.1 with $w_1 \in \{0, 0.5, 1\}$, $q_1 \in \{0.01, 0.025, 0.04\}$, (2) above, in a configuration with parameters $\sigma_1 = 0.3, \sigma_2 = 1, f_{00} = 0.9, f_{01} = f_{10} = 0.025, f_{11} = 0.05$. The oracle Procedure 3.2, where the primary study is study one with $\sigma_1 = 0.3$, is also examined. For each procedure the estimated power and FDR is shown as a function of the common expectation under the alternative, $\mu = \mu_1 = \mu_2$. Procedure 4.1 with $w_1 = 1$ is more powerful than with $w_1 = 0.5$ or $w_1 = 0$, while the choice $w_1 = 0$ is the worst in terms of power of Procedure 4.1. Moreover, Procedure 4.1 with $w_1 \in \{0.5, 1\}$ is more powerful than the BH procedure on maximum p-values. These findings were consistent across all configurations of $f_{00}, f_{10}, f_{01}, f_{11}$ examined,