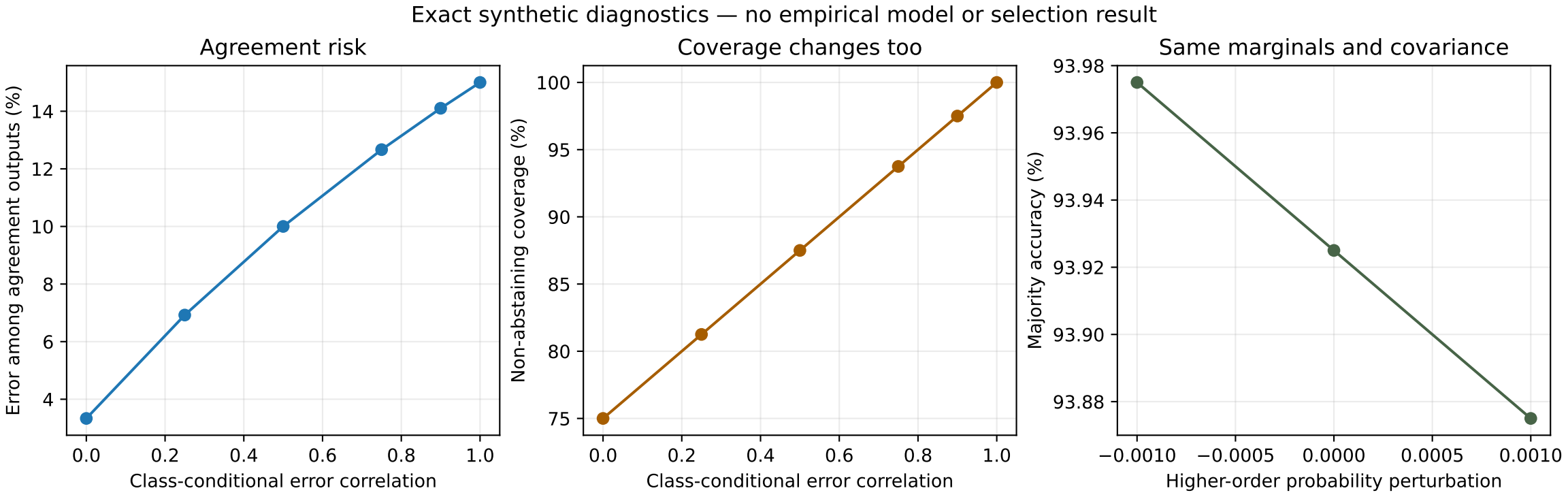

Exact synthetic diagnostics — no empirical model or selection result
Agreement risk
Error among agreement outputs (%)
4
6
8
10
12
14
0.0
0.2
0.4
0.6
0.8
1.0
Class-conditional error correlation
Coverage changes too
Non-abstaining coverage (%)
75
80
85
90
95
100
0.0
0.2
0.4
0.6
0.8
1.0
Class-conditional error correlation
Same marginals and covariance
Majority accuracy (%)
93.88
93.90
93.92
93.94
93.96
93.98
−0.0010
−0.0005
0.0000
0.0005
0.0010
Higher-order probability perturbation