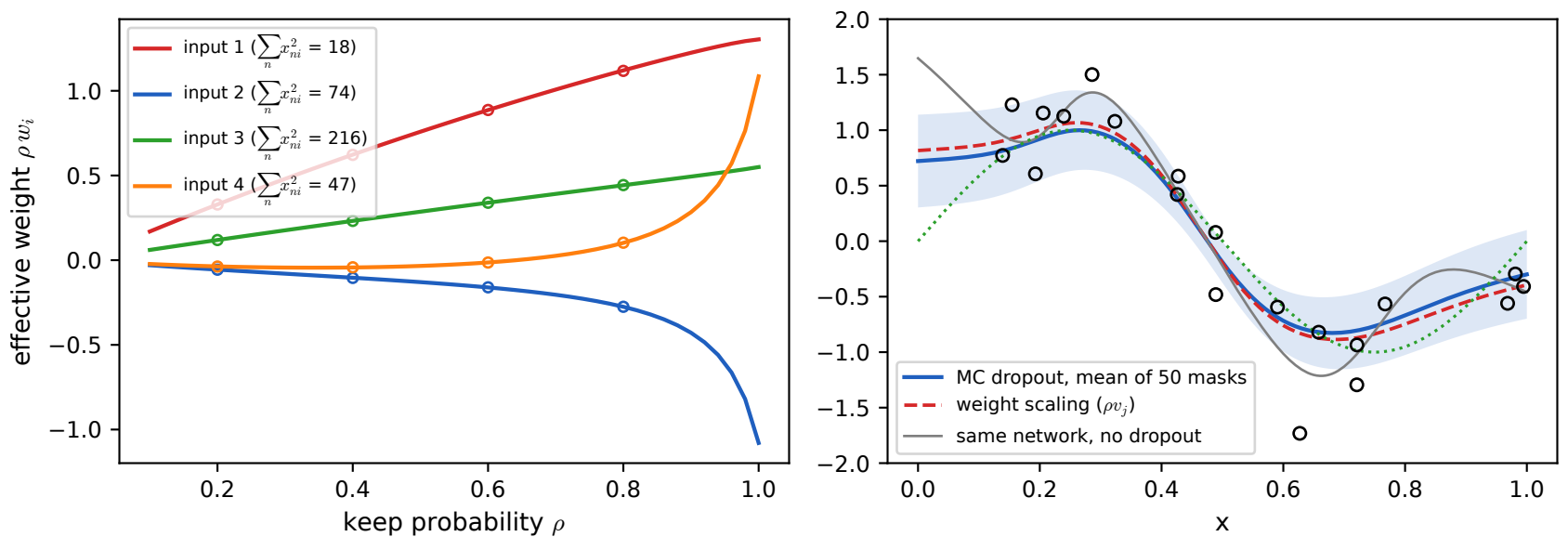

input 1 ($\sum_n x_{ni}^2$ = 18)
input 2 ($\sum_n x_{ni}^2$ = 74)
input 3 ($\sum_n x_{ni}^2$ = 216)
input 4 ($\sum_n x_{ni}^2$ = 47)
effective weight $\rho\, w_i$
keep probability ρ
MC dropout, mean of 50 masks
weight scaling (ρv_j)
same network, no dropout
x