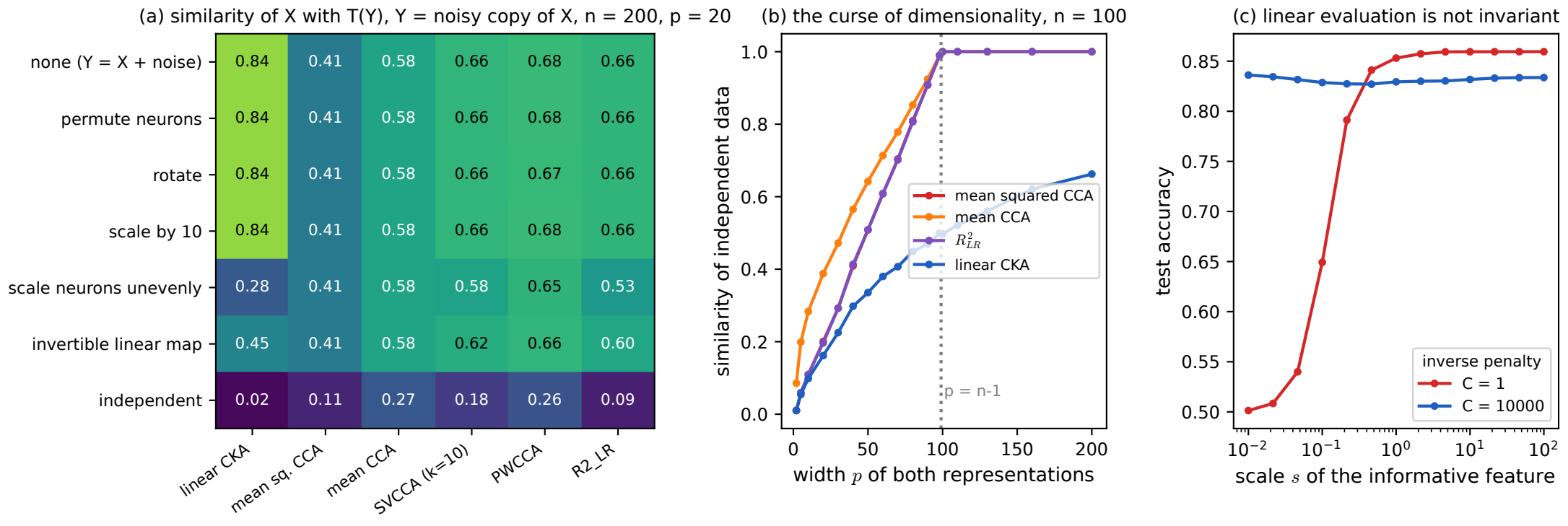

(a) similarity of X with T(Y), Y = noisy copy of X, n = 200, p = 20
none (Y = X + noise)
permute neurons
rotate
scale by 10
scale neurons unevenly
invertible linear map
independent
0.84 0.41 0.58 0.66 0.68 0.66
0.84 0.41 0.58 0.66 0.68 0.66
0.84 0.41 0.58 0.66 0.67 0.66
0.84 0.41 0.58 0.66 0.68 0.66
0.28 0.41 0.58 0.58 0.65 0.53
0.45 0.41 0.58 0.62 0.66 0.60
0.02 0.11 0.27 0.18 0.26 0.09
linear CKA
mean sq. CCA
mean CCA
SVCCA (k=10)
PWCCA
R2_LR
(b) the curse of dimensionality, n = 100
similarity of independent data
width p of both representations
p = n-1
mean squared CCA
mean CCA
R^2_{LR}
linear CKA
(c) linear evaluation is not invariant
test accuracy
scale s of the informative feature
inverse penalty
C = 1
C = 10000