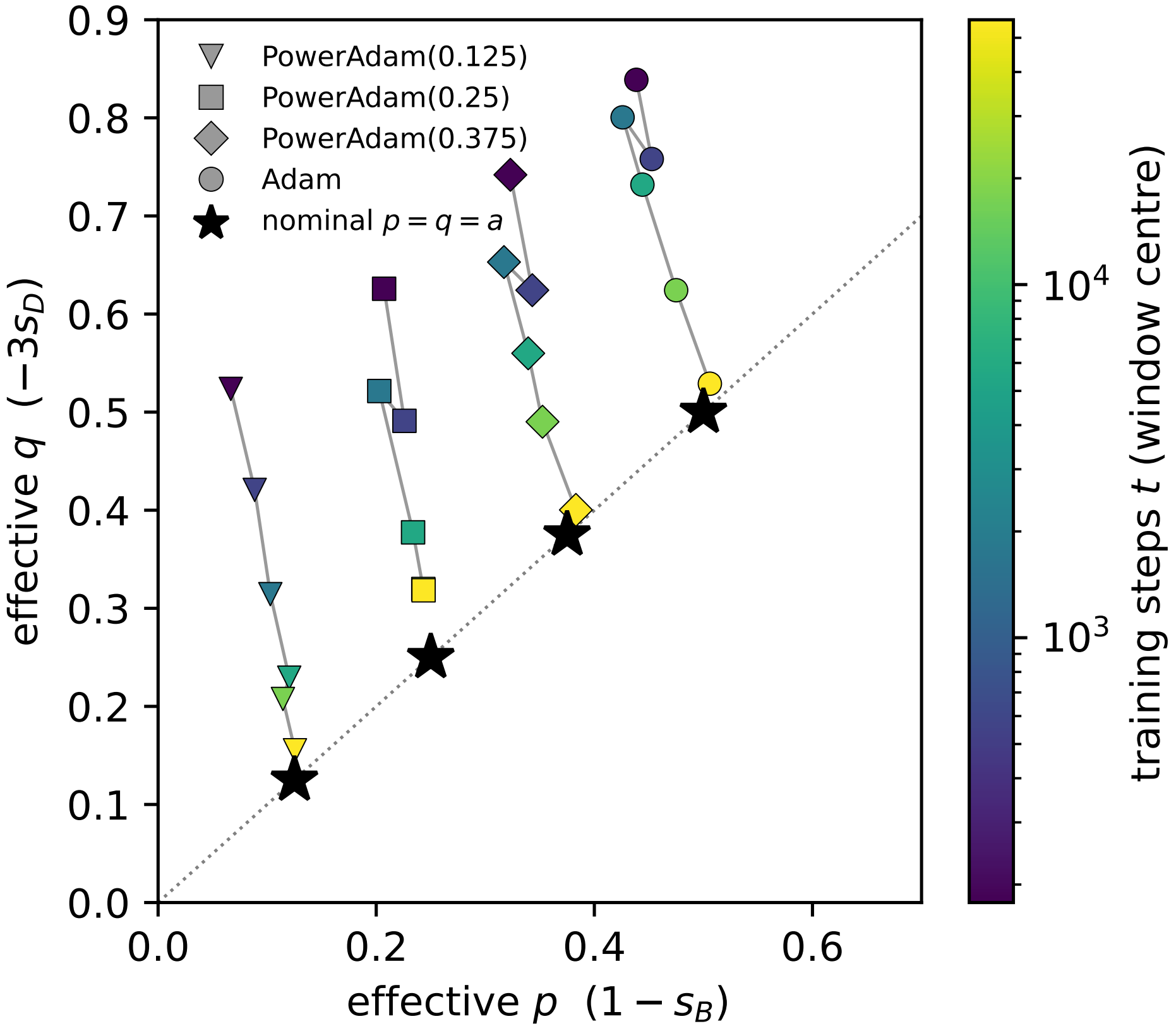

PowerAdam(0.125)
PowerAdam(0.25)
PowerAdam(0.375)
Adam
nominal $p = q = a$
effective q $(-3s_D)$
effective p $(1 - s_B)$
training steps t (window centre)
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
0.0
0.2
0.4
0.6
10^3
10^4