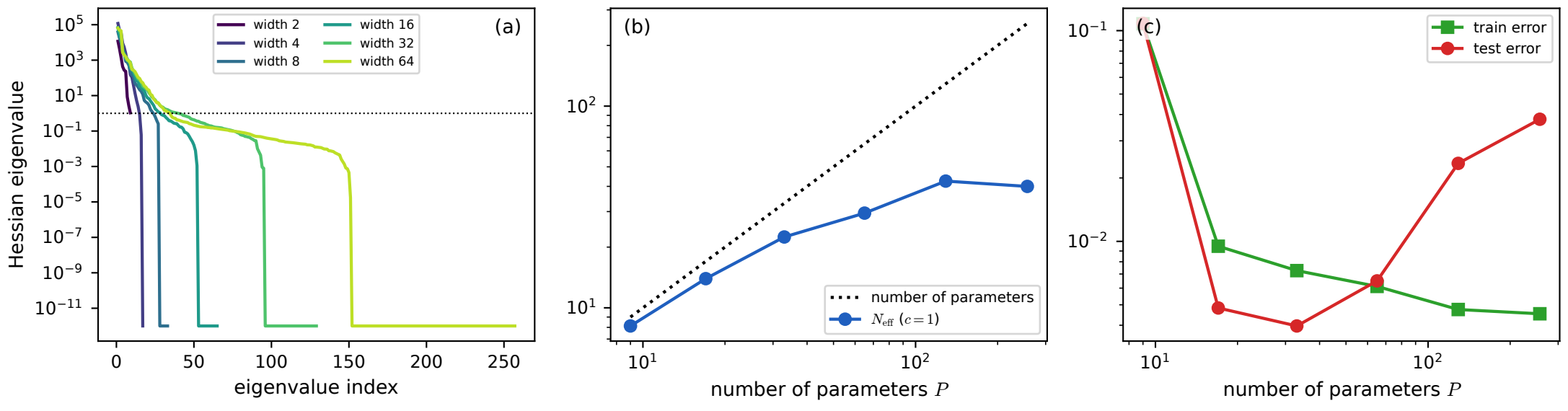

(a)
width 2
width 4
width 8
width 16
width 32
width 64
Hessian eigenvalue
eigenvalue index
(b)
number of parameters
N_{eff} $(c=1)$
number of parameters P
(c)
train error
test error
number of parameters P