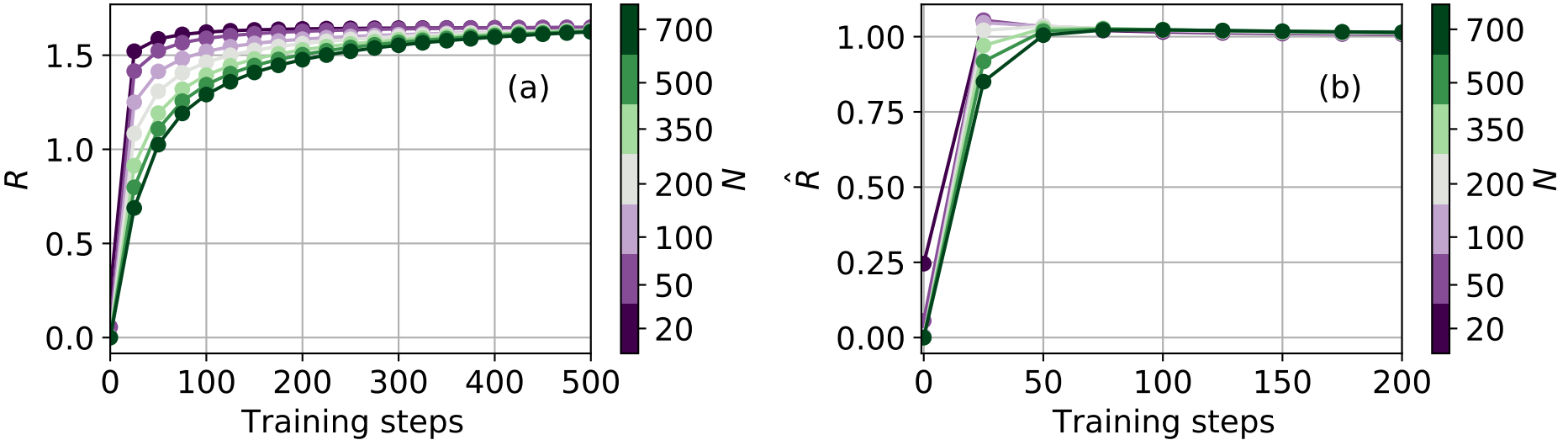
(a)
R
Training steps
0.0
0.5
1.0
1.5
0
100
200
300
400
500
N
20
50
100
200
350
500
700
(b)
$\hat{R}$
Training steps
0.00
0.25
0.50
0.75
1.00
0
50
100
150
200
N
20
50
100
200
350
500
700