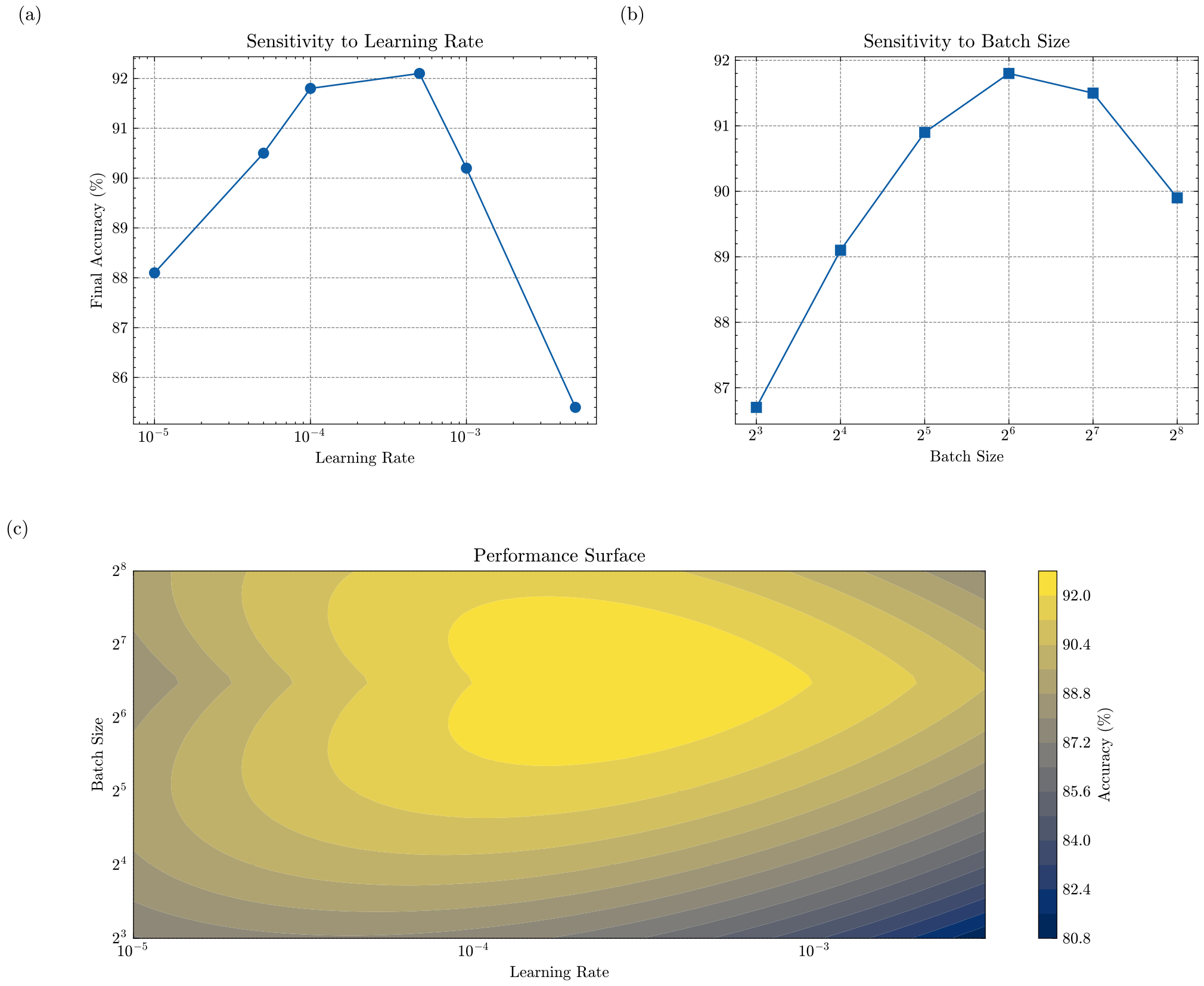
(a)
Sensitivity to Learning Rate
Final Accuracy (%)
92
91
90
89
88
87
86
10^{-5}
10^{-4}
10^{-3}
Learning Rate
(b)
Sensitivity to Batch Size
92
91
90
89
88
87
2^3
2^4
2^5
2^6
2^7
2^8
Batch Size
(c)
Performance Surface
Batch Size
2^8
2^7
2^6
2^5
2^4
2^3
10^{-5}
10^{-4}
10^{-3}
Learning Rate
Accuracy (%)
92.0
90.4
88.8
87.2
85.6
84.0
82.4
80.8