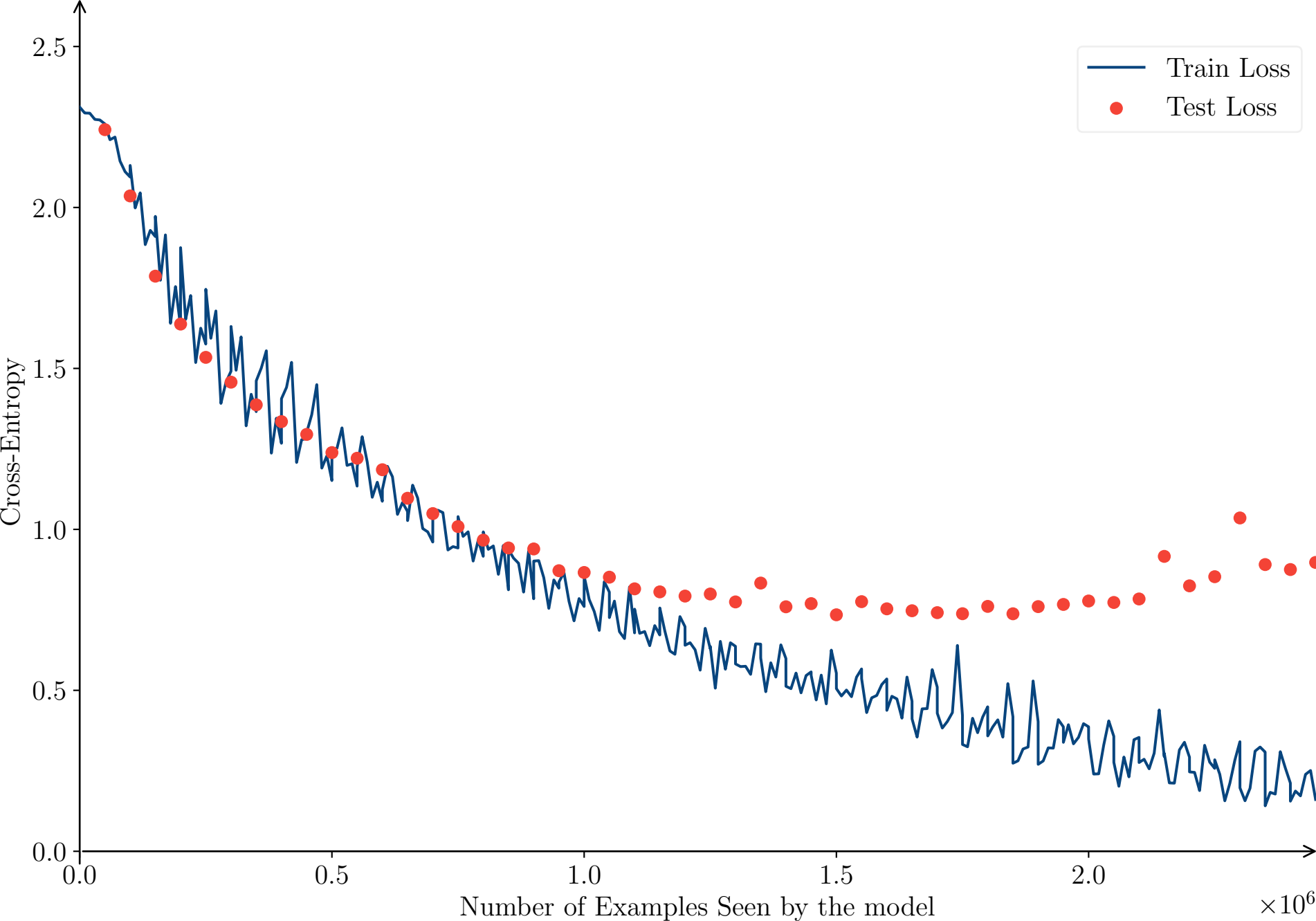
Train Loss
Test Loss
Cross-Entropy
2.5
2.0
1.5
1.0
0.5
0.0
0.0
0.5
1.0
1.5
2.0
Number of Examples Seen by the model
$\times 10^6$