Cycle-Consistent Model Merging – Rebuttal

1x	2x	4x	8x	16x
292k	1.166m	4.655m	18.600m	74.360m
n=2				
33.4s 0.24s	33.5s 0.4s	40.5s 3.4s	80.8s 8.9s	367.8s 59.4s
n=3				
32.9s 1.2s	83.18s 4.1s	91.0s 19.5s	162.0s 105.8s	715.8s 892.3s
	292k 33.4s 0.24s	292k 1.166m n= 33.4s 33.5s 0.24s 0.4s n= 32.9s 83.18s	292k 1.166m 4.655m n=2 33.4s 33.5s 40.5s 0.24s 0.4s 3.4s n=3 32.9s 83.18s 91.0s	292k 1.166m 4.655m 18.600m n=2 33.4s 33.5s 40.5s 80.8s 0.24s 0.4s 3.4s 8.9s n=3 32.9s 83.18s 91.0s 162.0s

Table 1: Wall-clock time for merging n=3 ResNet20 models with different widths.

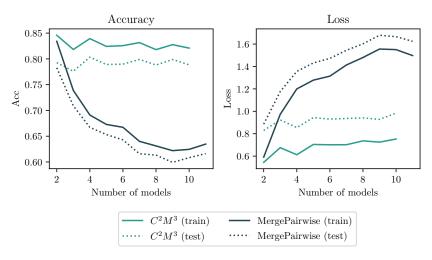


Figure 1: Accuracy and loss obtained when merging $n=2,\ldots,10$ ResNet20 models with 4x width.

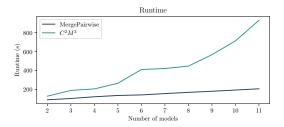


Figure 2: Matching and merging time when merging $n=2,\ldots,10$ ResNet20 models with 4x width.

Preprint. Under review.