# Bayesian Optimization with Clustering and Rollback for CNN Auto Pruning (Supplementary Material)

Hanwei Fan[1,2], Jiandong Mu[2], and Wei Zhang[2]

[1] The Hong Kong University of Science and Technology (Guangzhou)
[2] The Hong Kong University of Science and Technology
{hfanah,jmu}@connect.ust.hk, wei.zhang@ust.hk

## 1 Analysis of the Choice of $C^*$

This document support our paper by providing a detailed analysis of the choice of the bridge stage $C^*$ on MobileNetV2.

In this experiment, we divide the layers into 6 clusters and then perform the gradual rollback with different $C^*$. Table 1 shows the top-1 mean and variance the results. Our proposed gradual rollback method consistently outperforms the direct rollback method, showing its superiority. Although the results of different choices of $C^*$ are slightly different, the performance is fairly robust when the choice of $C^*$ is reasonable. Based on our experience, we recommend choosing $C^*$ to be 2.5 times the cluster number $C$. Note that if the recommended $C^*$ is larger than the original dimensionality $N$, there should be no need to use gradual rollback as N is not high enough.

| $C^*$ | top-1 $m$ | top-1 $\sigma$ |
|---|---|---|
| 12 | 52.89 | 0.77 |
| 13 | 52.89 | 0.75 |
| 14 | 53.02 | 0.94 |
| 15 | 53.01 | 0.88 |
| 16 | 52.97 | 0.97 |
| 17 | 52.96 | 0.89 |
| Direct | 52.86 | 0.61 |

Table 1: Performance of different choices of $C^*$.