EVALUATION WITH T_A = 100 μS

This document presents the evaluation results with T_A set to 100 μs. In the following results, T_R is configured to 5,300 μs, and the bandwidth $\widehat{BW_i}$ ranges from 3,000 MB/s to 9,000 MB/s, as determined by Eq. (3) and Eq. (5). Other experimental configurations are identical to those in the evaluation section of the paper.

A. Memory Access Regulation

The evaluation results of memory access regulation exhibited overall trends similar to those observed with T_A = 200 μs. However, due to the additional overhead introduced by T_A = 100 μs, both the bandwidth usage and user-perceived measures are observed to be lower compared to the original experiments conducted with T_A = 200 μs. Note that in figure 3, we set BW_1 = 9,000 MB/s and BW_2 = 3,000 MB/s, whereas in figure 4, the values are reversed.

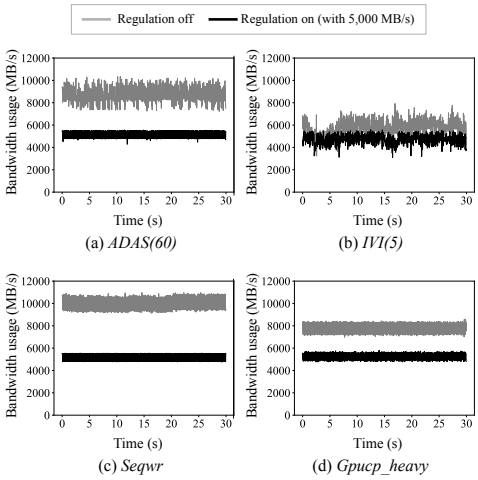

Fig. 1. Time plot of memory bandwidth usage

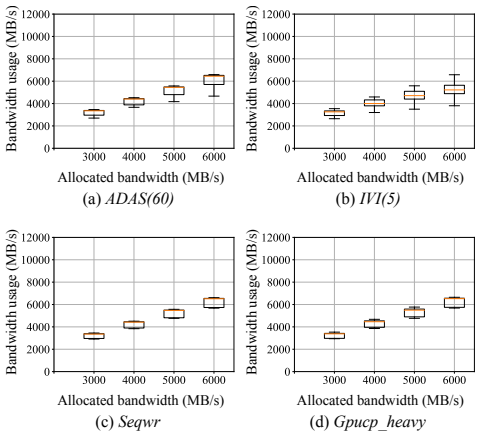

Fig. 2. Statistics of memory bandwidth usage by regulation

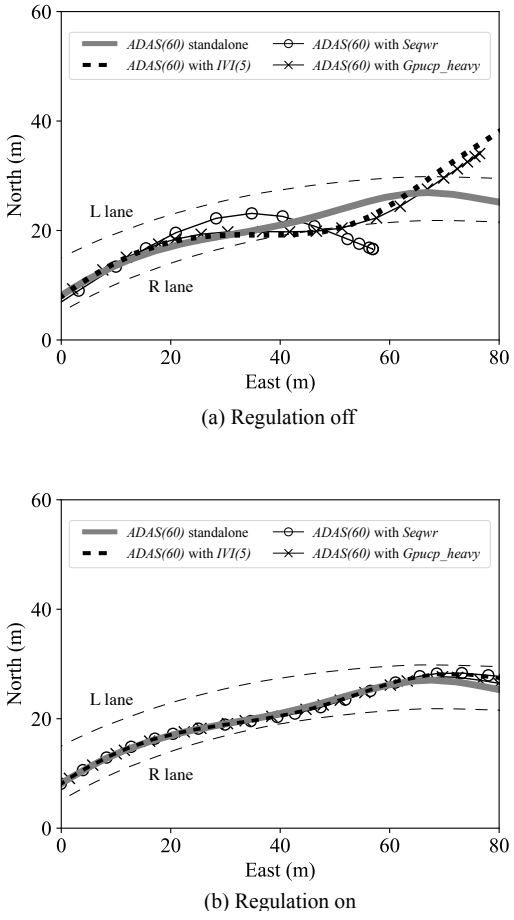

Fig. 3. Lane-keeping behaviors

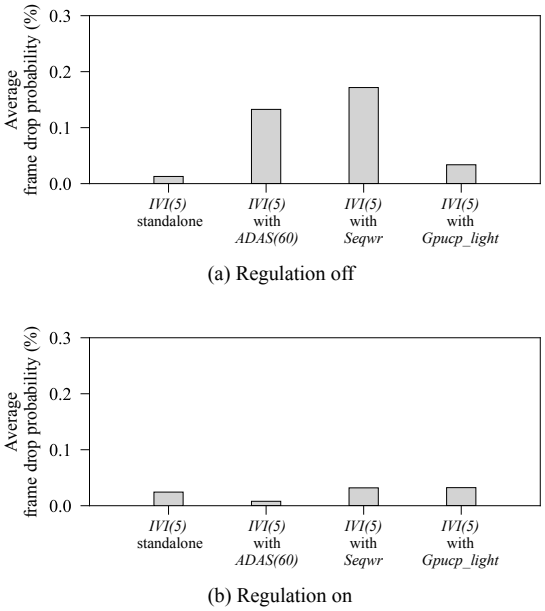

Fig. 4. Average frame drop probability

B. Few-Shot Measurement based Optimal Memory Bandwidth Allocation

In the evaluation, the baseline allocates the total bandwidth among applications in proportion to their average bandwidth usage. This becomes possible because a T_A = μs expands the range of bandwidth $\widehat{BW}_i$ from (4,800 MB/s, 7,200 MB/s) to (3,000 MB/s, 9,000 MB/s).

Figure 5 (a) shows the global utility of the complete curve, the baseline, and the estimated curve when T_A = 100 μs. Unlike the case with T_A = 200 μs, there are some instances (e.g., *ADAS(20) + IVI(4)*) where the baseline slightly outperforms our method. However, the difference is marginal, and in scenarios with large utility drops (e.g., *ADAS(60) + IVI(2)*), the estimated curve effectively converges to a near-optimal allocation. In terms of the geometric mean, the estimated curve also outperforms the baseline, achieving 5% compared to 12% from the baseline. In Figure 5 (b), the estimated curve achieves near-optimal utility with only 2 initial and up to 2 additional measurements. These results show that our mechanism can still achieve near-optimal global utility with a very small number of measurements when T_A = 100 μs.

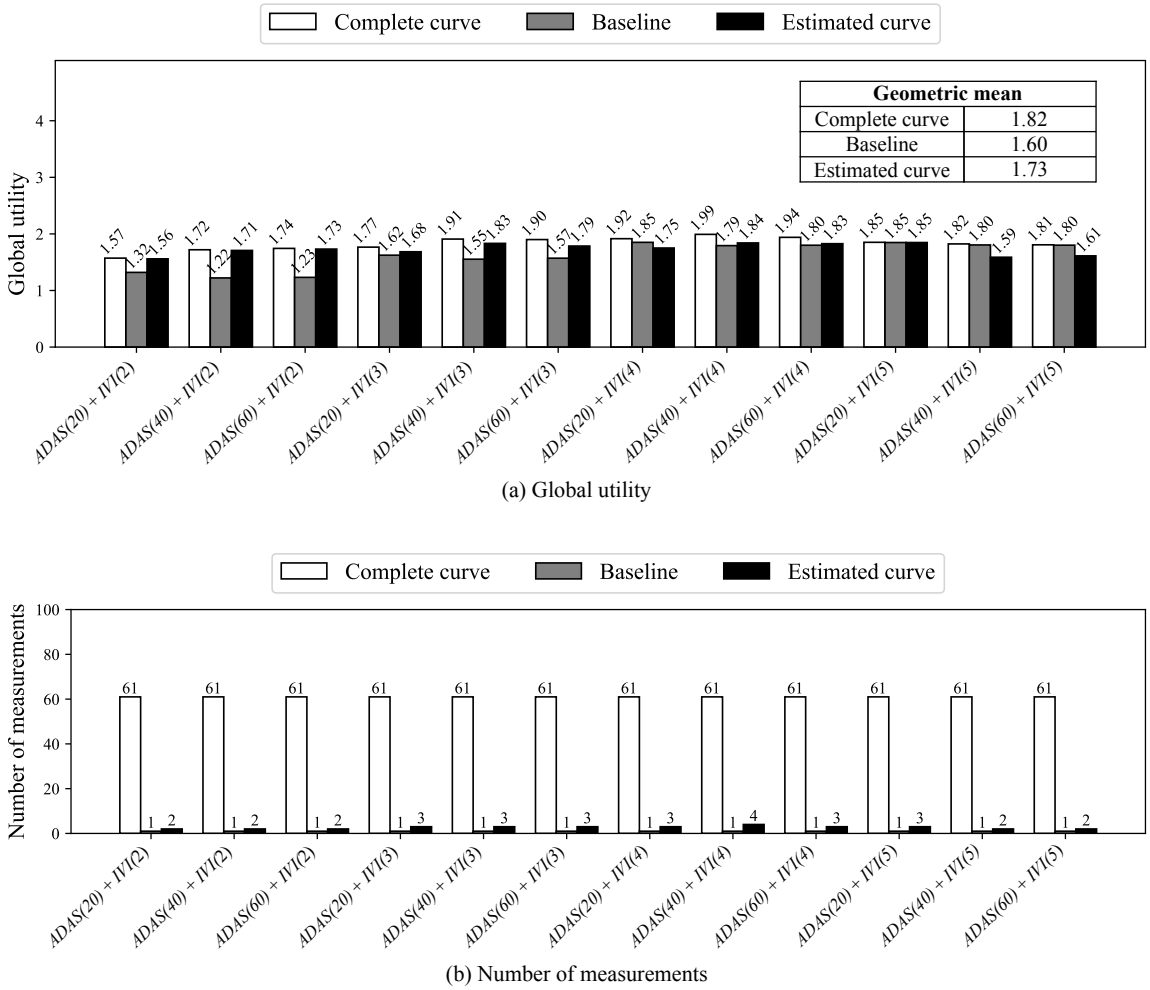

Fig. 5. Performance of few-shot measurement based optimal memory bandwidth allocation