

# Numerical Experiments for “OPTIMAL BATCHING UNDER COMPUTATION BUDGET”

We set  $\psi(P)$  as the 0.7 quantile of  $X \sim P$  where the true distribution is standard log-normal. We set the sample size as  $n = 3000$ , nominal level  $\alpha = 90\%$ , and try the following methods under different computational budgets ( $K = 6, 12, 17$ ) measured by the number of times we can compute  $\psi$ :

1. Standard batching:  $CI_B$  as introduced in Section 3.
2. Batching with general batch sizes:  $CI_B^{(\gamma)}$  as introduced in Theorem 2. We use  $\gamma_j = \frac{j}{K(K+1)/2}, 1 \leq j \leq K$ .
3. Overlapping batching:  $CI_{OB}$  as introduced in Theorem 4. The  $K$  batched estimates  $\psi_1, \dots, \psi_K$  are constructed as follows.  $\psi_1 = \psi(X_1, \dots, X_n)$  is the estimation using the entire empirical distribution.

$$\psi_j = \psi(X_{(j-2)n(1-\gamma)/(K-2)+1}, \dots, X_{(j-2)n(1-\gamma)/(K-2)+\gamma n}), 2 \leq j \leq K.$$

Here  $\gamma$  is set as 0.3. Note that  $\gamma$  measures the batch size (the size of each batch is  $\gamma n$ ).

4. Cheap bootstrap:  $CI_{CB}$  introduced in Section 6 with  $K - 1$  resampled estimates (subtracting one since the estimation using the entire empirical distribution already used one budget)
5. Overlapping batching (OB-I) as in Su et al. [2023]. The batched estimates are the same as our overlapping batching outlined in item 3 above. The CI in their construction is

$$CI_{OB-I} := \psi_1 \pm c_{\beta, b, \alpha_1} \sqrt{\frac{\beta}{1-\beta} \frac{1}{K-1} \sum_{j=2}^K (\psi_j - \psi_1)^2}.$$

Here, the critical values  $c_{\beta, b, \alpha_1}$  can be checked from Fig 3 from Su et al. [2023] and in correspondance with our settings in 3, we use  $\beta = 0.3, b = K - 1, \alpha_1 = 0.95$ .

For each method, we replicate  $10^5$  times to find its empirical coverages and half lengths. The results are shown in Table 1. We can check that the empirical coverages are all close to the nominal coverage: the largest gap is observed for  $CI_{OB}$  when  $K = 17$  where there is an overcoverage of about 1.3%. This suggests that the sample size is large enough for our asymptotic argument to be true. If we compare the half lengths of different methods, we find that 1) the half lengths of  $CI_B, CI_B^{(\gamma)}$ , and  $CI_{CB}$  are very close, which agrees with our theoretical findings. 2) the half length of  $CI_{OB}$  is longer than  $CI_B, CI_B^{(\gamma)}$ , and  $CI_{CB}$ . For example, when  $K = 6$ , the half length of  $CI_{OB}$  is 0.082 where the half lengths of  $CI_B, CI_B^{(\gamma)}$ , and  $CI_{CB}$  are all 0.078. This difference can be explained by its higher coverage, which is a result of finite-sample errors as we discussed. 3) The half length of  $CI_{OB-I}$  is even longer than  $CI_{OB}$ . For example, when  $K = 12$ , the half length of  $CI_{OB-I}$  is longer than  $CI_{OB}$  even though its coverage is smaller than  $CI_{OB}$ . This is expected as the construction in Su et al. [2023] does not have the optimality guarantee in terms of CI accuracy as we do.

method	$K = 6$		$K = 12$		$K = 17$	
	coverage	half length	coverage	half length	coverage	half length
$CI_B$	90.0%	0.078	90.0%	0.071	90.0%	0.070
$CI_B^{(\gamma)}$	90.0%	0.078	89.9%	0.071	90.0%	0.070
$CI_{OB}$	90.5%	0.082	91.1%	0.076	91.3%	0.075
$CI_{CB}$	89.4%	0.078	89.6%	0.072	89.9%	0.071
$CI_{OB-I}$	90.7%	0.086	90.3%	0.080	90.3%	0.079

Table 1: Empirical coverages and half lengths for each batching method. The uncertainty of these empirical estimates (measured by the 95% CI half length) is less than 0.2% for the empirical coverage and is less than  $3 \times 10^{-4}$  for the estimation of the half length.

## References

Ziwei Su, Raghu Pasupathy, Yingchieh Yeh, and Peter W. Glynn. Overlapping batch confidence intervals on statistical functionals constructed from time series: Application to quantiles, optimization, and estimation, 2023.