

1 ADDITIONAL EXPERIMENTS FOR CENSUS1990 DATASET

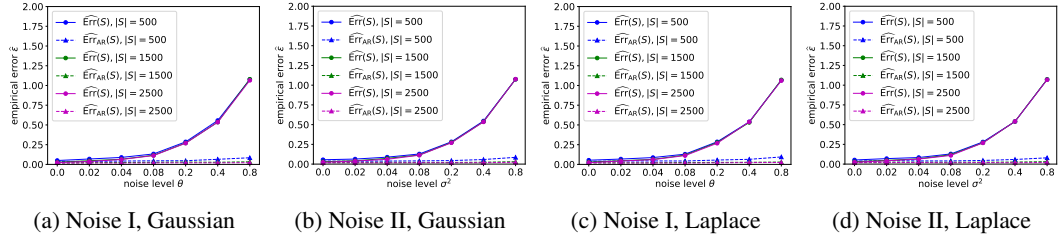


Figure 1: The empirical errors $\widehat{\text{Err}}(S)$ and $\widehat{\text{Err}}_{\text{AR}}(S)$ versus the noise level (θ or σ^2) plot for coreset S with different sizes. The solid lines in each figure represent the coreset measure $\widehat{\text{Err}}(S)$, and the dashed lines represent the approximation-ratio measure $\widehat{\text{Err}}_{\text{AR}}(S)$. We show the results from Census1990 dataset, and Figures 1a to 1d denote the noise models I, II with Gaussian and Laplace noise respectively.

Table 1: Empirical errors $\widehat{\text{Err}}(S)$ and $\widehat{\text{Err}}_{\text{AR}}(S)$ under different noise models and coreset sizes for Census1990 dataset. We fix the noise level $\theta = \sigma^2 = 0.2$. We repeat each setting 20 times, and provide the mean and the standard deviation (in the subscript) of the empirical errors.

Size	Model I, Gaussian		Model II, Gaussian		Model I, Laplace		Model II, Laplace	
	$\widehat{\text{Err}}(S)$	$\widehat{\text{Err}}_{\text{AR}}(S)$	$\widehat{\text{Err}}(S)$	$\widehat{\text{Err}}_{\text{AR}}(S)$	$\widehat{\text{Err}}(S)$	$\widehat{\text{Err}}_{\text{AR}}(S)$	$\widehat{\text{Err}}(S)$	$\widehat{\text{Err}}_{\text{AR}}(S)$
500	0.286 _{0.02}	0.047 _{0.01}	0.282 _{0.01}	0.044 _{0.01}	0.286 _{0.02}	0.055 _{0.01}	0.280 _{0.02}	0.046 _{0.01}
1000	0.275 _{0.02}	0.030 _{0.01}	0.275 _{0.01}	0.029 _{0.01}	0.275 _{0.01}	0.034 _{0.01}	0.276 _{0.01}	0.029 _{0.01}
2000	0.272 _{0.01}	0.022 _{0.01}	0.270 _{0.01}	0.019 _{0.01}	0.269 _{0.01}	0.021 _{0.01}	0.272 _{0.01}	0.022 _{0.01}
3000	0.262 _{0.01}	0.016 _{0.01}	0.271 _{0.01}	0.016 _{0.00}	0.271 _{0.01}	0.019 _{0.01}	0.268 _{0.01}	0.015 _{0.01}
4000	0.268 _{0.01}	0.017 _{0.01}	0.269 _{0.01}	0.015 _{0.00}	0.264 _{0.01}	0.017 _{0.01}	0.266 _{0.00}	0.012 _{0.01}
5000	0.263 _{0.01}	0.017 _{0.01}	0.265 _{0.01}	0.011 _{0.01}	0.266 _{0.01}	0.015 _{0.01}	0.267 _{0.01}	0.014 _{0.01}