

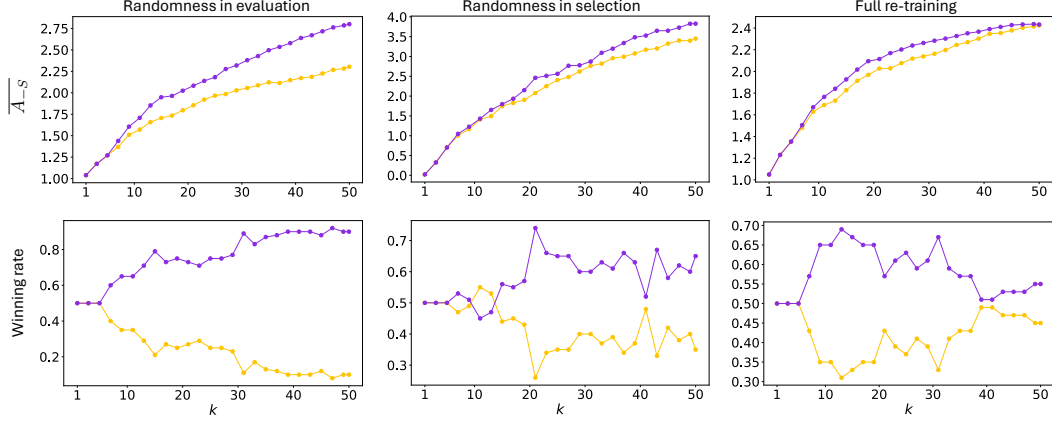


Figure 1: **Additional experiments on MLP.** We examined the randomness in both the evaluation and the selection steps, and experimented with full-retraining. We report the actual effect and the winning rate.

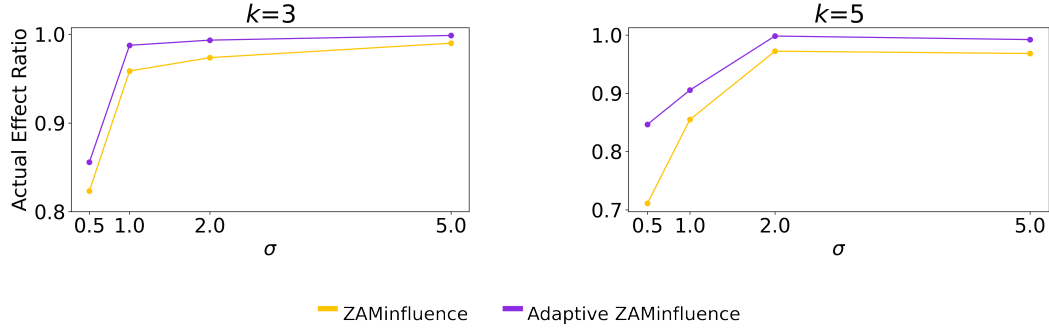


Figure 2: **Comparison with ground truth influence.** Experiments on synthetic datasets of a mixture of Gaussians with 50 data points. σ denotes the standard deviation of Gaussians. We report the *actual effect ratio*, the ratio of the actual effect achieved by the set selected by the two methods over the actual effect achieved by the ground truth subset (of maximum influence).

Table 1: **Cross-validation performance.** for MLP Model on MNIST. Width stands for the width of the hidden layer of the MLP, lr stands for the learning rate, and β stands for the momentum.

width	lr	β	epochs	Accuracy	width	lr	β	epochs	Accuracy
64	0.01	0.9	30	91.96%	64	0.1	0.9	30	39.36%
128	0.01	0.9	30	93.44%	128	0.1	0.9	30	40.08%
64	0.01	0.9	50	92.88%	64	0.1	0.9	50	41.64%
128	0.01	0.9	50	93.40%	128	0.1	0.9	50	45.36%
64	0.01	0.95	30	92.48%	64	0.1	0.95	30	13.48%
128	0.01	0.95	30	93.12%	128	0.1	0.95	30	13.36%
64	0.01	0.95	50	93.48%	64	0.1	0.95	50	10.8%
128	0.01	0.95	50	94.68%	128	0.1	0.95	50	15.71%
64	0.05	0.9	30	88.44%	64	0.5	0.9	30	11.68%
128	0.05	0.9	30	87.64%	128	0.5	0.9	30	11.68%
64	0.05	0.9	50	86.64%	64	0.5	0.9	50	11.68%
128	0.05	0.9	50	89.60%	128	0.5	0.9	50	11.68%
64	0.05	0.95	30	45.80%	64	0.5	0.95	30	11.04%
128	0.05	0.95	30	41.24%	128	0.5	0.95	30	11.20%
64	0.05	0.95	50	53.32%	64	0.5	0.95	50	11.20%
128	0.05	0.95	50	54.60%	128	0.5	0.95	50	11.20%