Table 1: *AdaGauss* results without or with frozen batch-norm.

	CIFAR100	ImageNetSubset	CUB200
Resnet18 (no BN)	44.6±0.7 | 58.1±0.7	49.1±0.8 | 61.7±0.9	54.2±0.5 | 66.1±0.7
Resnet18 (freezed BN)	45.3±0.7 | 58.7±0.9	49.4±0.8 | 62.0±1.0	55.2±0.4 | 65.7±0.6
Resnet18	46.1±0.8 | 60.2±0.9	51.1±1.0 | 65.0±1.2	55.8±0.5 | 66.2±0.8

Table 2: Time complexity of EFCIL methods measured on CIFAR100.

	LwF	FeTrIL	FeCAM	SSRE	EFC	AdaGauss	AdaGauss (lin. head)
Inference time (sec)	66.3±2.7	68.3±3.4	74.2±4.0	274.7±12.3	67.4±3.1	72.8±3.2	65.3±2.2
Training time (min)	53.8±2.9	5.3±0.3	5.9±0.4	548.1±17.4	94.±2.9	86.3±3.2	103.3±4.2

Table 3: *AdaGauss* results with different backbone architectures.

	CUB200		FGVCAircraft	
	T=10	T=20	T=10	T=20
Resnet18	55.8 | 66.2	47.4 | 60.6	47.5 | 58.5	34.8 | 48.6
ConvNext (small)	63.4 | 73.1	47.9 | 64.1	49.3 | 62.9	37.3 | 51.4
ViT (small)	68.2 | 77.5	48.9 | 67.5	48.0 | 60.6	35.7 | 50.2

Table 4: Half dataset in the first task results.

	CIFAR100		ImageNetSubset	
	T = 5	T = 10	T = 5	T = 10
PASS	54.5 | 61.8	53.8 | 60.9	57.9 | 64.4	58.2 | 61.8
SSRE	55.7 | 63.9	54.9 | 63.2	58.3 | 65.2	61.4 | 67.7
FeTrIL	58.3 | 65.1	56.2 | 64.6	65.6 | 72.8	65.3 | 72.1
FeCAM	60.2 | 67.2	59.9 | 66.9	67.3 | 75.3	67.6 | 74.9
ACIL	57.8 | 66.3	57.7 | 66.0	67.0 | 74.8	67.2 | 74.6
DS-AL	61.4 | 68.4	61.4 | 68.4	68.0 | 75.2	67.7 | 75.1
AdaGauss	58.9 | 65.7	55.4 | 63.7	66.8 | 74.1	62.8 | 68.0

Table 5: Additional ablation study.

Classifier	Cov. Matrix	Adapt mean	Adapt covariance	L_{AC}	CIFAR100	ImagenetSubset
Bayes	Full	✗	✗	✗	22.9 | 42.8	22.5 | 43.4
Bayes	Full	✓	✓	✓	46.1 | 60.2	51.1 | 65.0

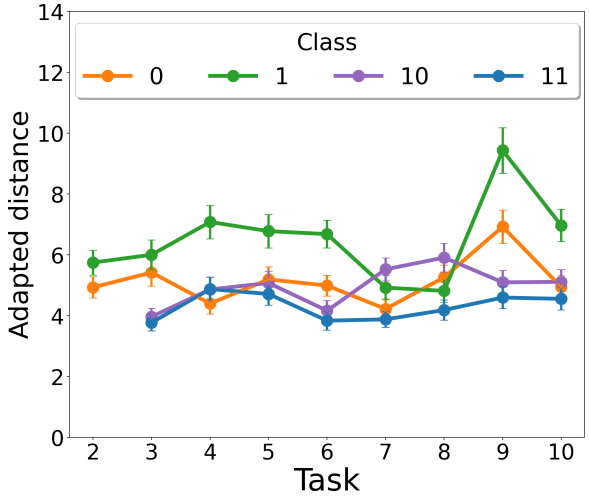

Figure 1: Predicted shift for different classes on CIFAR100 (10 equal tasks) by *AdaGauss*.