

Supplementary Material to the Paper Weight Weaving: Parameter Pooling for Data-Free Model Merging

The supplementary material includes the following sections:

- **A1** Extra information about the code base, checkpoints, and infrastructure used to run all experiments.
- **A2** Shows the performances for all merging methods evaluated in this work, including our proposed method in a data-free setting.
- **A3** Showcases the task vector’s weight correlations between sequential and individual fine-tuning.

A1 Extra implementation details

Model checkpoints & Computing resources. We used following open-source code to obtain model checkpoints 1) vision multi-task learning and domain generalization: https://github.com/mlfoundations/task_vectors; 2) continual learning: <https://github.com/danielm1405/magmax>. All experiments run in a single NVIDIA-A100. We will make the source code available.

A2 Performances by dataset

We compare the accuracy of state-of-the-art methods using the optimal scaling factor found by the authors (when available), and the data-free setting ($\lambda = 1$) with the Weight Weaving applied to these methods. Each row presents results for a specific model merging method, while the columns show the dataset, including whether it uses weight augmentation (‘augment’), and the scaling factor (λ) value. Tables 4 to 6 show the performances for multi-task learning, continual learning results appear in Tables 7 to 9, and domain generalization in Tables 10 to 12.

Table 4: ViT-B-32 results for experimental setup Vision Multi-Task learning 8 Datasets.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	53.46	96.79	76.93	79.54	66.27	54.24	49.84	69.92	68.37
		1.00	0.62	63.05	20.04	41.86	6.22	0.27	8.03	18.46	19.82
	Weight Weaving	1.00	45.77	97.42	70.67	80.06	59.97	43.25	45.05	68.20	63.80
DARE	No	0.30	53.19	96.72	77.19	79.35	66.13	53.97	49.63	69.60	68.22
		1.00	0.68	61.93	17.63	40.58	5.97	0.26	6.60	17.37	18.88
	Weight Weaving	1.00	45.40	97.37	71.04	79.74	59.89	42.84	45.05	68.00	63.67
TIES	No	1.00	58.09	98.04	80.48	86.53	71.44	58.72	53.72	71.50	72.32
	Weight Weaving	1.00	62.94	96.69	80.22	82.45	74.22	64.42	55.80	70.88	73.45
Breadcrumbs	No	0.30	62.95	83.24	70.56	60.25	70.73	65.56	50.32	50.24	64.23
		1.00	49.47	96.84	75.33	79.66	65.19	51.07	50.05	70.73	67.29
	Weight Weaving	1.00	63.25	85.96	73.30	64.31	71.92	65.70	51.76	54.12	66.29
MagMax	No	0.50	65.20	93.14	72.33	76.45	73.92	67.60	53.88	61.42	70.49
		1.00	49.83	98.08	70.89	88.01	64.52	49.18	48.35	63.49	66.55
	Weight Weaving	1.00	64.44	91.65	73.59	73.11	73.30	66.89	53.24	60.09	69.54
PCB	No	1.00	64.89	97.72	81.67	85.16	77.87	65.31	58.51	76.38	75.94
	Weight Weaving	1.00	63.18	98.08	82.07	86.08	76.71	63.36	57.82	77.13	75.55
ISO-C	No	1.00	74.24	97.83	90.56	82.84	88.13	71.87	68.30	87.72	82.69
	Weight Weaving	1.00	71.72	95.66	86.52	76.86	83.92	70.52	62.93	78.48	78.33
TSV	No	1.00	71.78	99.26	93.74	91.65	84.78	66.92	68.40	92.26	83.60
	Weight Weaving	1.00	64.95	99.39	92.30	92.69	80.90	59.15	64.89	91.97	80.78

Table 5: ViT-B-16 results for experimental setup Vision Multi-Task learning 8 Datasets.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	66.98	98.60	80.22	87.37	74.16	63.36	52.98	75.22	74.86
		1.00	0.90	49.51	12.22	49.88	4.13	0.97	7.13	11.24	17.00
DARE	No	1.00	61.92	98.76	75.48	88.54	69.17	55.32	50.69	73.25	71.64
		0.30	66.78	98.59	79.85	87.16	74.22	63.11	53.35	75.57	74.83
TIES	No	1.00	0.71	47.86	11.93	49.50	3.73	0.98	7.98	9.59	16.53
		1.00	61.48	98.76	75.41	88.46	68.70	55.00	51.01	73.56	71.55
Breadcrumbs	No	1.00	72.47	98.87	86.48	89.21	79.89	67.34	58.14	75.07	78.43
		1.00	73.47	98.42	86.15	86.13	80.43	69.91	56.91	75.03	78.31
MagMax	No	0.30	69.83	92.22	75.81	71.62	75.32	68.85	50.16	57.91	70.21
		1.00	66.14	98.63	81.22	87.80	72.11	61.09	54.41	76.87	74.78
PCB	No	1.00	70.63	94.15	79.41	74.10	76.19	69.06	51.06	61.27	71.98
		0.50	73.14	96.64	77.33	80.53	79.10	71.41	53.99	61.71	74.23
ISO-C	No	1.00	67.35	98.73	74.96	90.21	75.25	61.29	54.26	60.62	72.84
		1.00	72.32	96.30	79.15	79.04	78.30	70.64	53.24	63.06	74.01
TSV	No	1.00	75.05	98.68	91.96	86.22	82.87	69.75	59.79	82.77	80.89
		1.00	74.12	98.87	91.67	87.38	82.14	68.54	60.11	83.35	80.77
TSV	No	1.00	83.66	98.86	95.44	87.55	91.63	75.35	71.22	92.49	87.03
		1.00	82.76	98.74	94.56	87.38	90.14	74.70	68.88	91.08	86.03
TSV	No	1.00	82.14	99.35	95.59	93.47	88.83	72.26	72.02	94.17	87.23
		1.00	81.58	99.28	94.96	92.68	87.68	71.75	69.41	92.75	86.26

Table 6: ViT-L-14 results for experimental setup Vision Multi-Task learning 8 Datasets.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	82.56	99.21	88.19	80.14	84.21	70.82	65.05	93.08	82.91
		1.00	0.61	28.04	11.26	26.55	4.05	0.35	3.40	16.99	11.41
DARE	No	1.00	79.37	99.38	79.11	80.76	80.83	67.14	62.55	93.99	80.39
		0.30	82.56	99.22	88.44	79.92	84.43	70.98	64.79	93.04	82.92
TIES	No	1.00	0.60	19.45	15.56	25.29	4.17	0.42	3.03	16.48	10.62
		1.00	79.32	99.35	79.93	80.47	80.76	67.15	62.87	94.09	80.49
Breadcrumbs	No	1.00	86.10	99.57	85.19	73.77	86.25	73.29	69.26	96.80	83.78
		1.00	86.33	99.40	89.67	74.05	86.24	74.25	67.82	95.17	84.12
MagMax	No	0.30	81.21	94.14	81.78	70.71	78.54	70.93	60.37	67.71	75.67
		1.00	82.85	99.13	85.30	78.27	85.03	71.95	66.01	92.49	82.63
PCB	No	1.00	81.71	95.70	84.78	72.09	79.95	71.45	61.06	72.09	77.35
		0.50	85.84	99.33	82.56	71.43	82.60	73.54	64.79	96.10	82.02
ISO-C	No	1.00	75.65	99.66	36.70	70.72	69.68	61.52	60.37	98.06	71.55
		1.00	85.08	99.00	87.15	72.89	82.76	73.11	63.99	93.29	82.16
TSV	No	1.00	86.67	99.49	95.26	84.29	88.94	74.40	71.91	93.54	86.81
		1.00	85.90	99.49	94.22	84.10	88.24	73.41	71.86	94.18	86.43
TSV	No	1.00	91.84	99.33	97.74	87.01	94.32	79.64	79.10	95.83	90.60
		1.00	90.70	99.04	97.30	84.55	93.24	78.34	75.37	94.43	89.12
TSV	No	1.00	90.65	99.71	97.48	87.00	93.37	77.85	80.00	97.66	90.46
		1.00	90.37	99.67	97.11	86.08	92.65	77.32	78.88	97.21	89.91

Table 7: ViT-B-32 results for experimental setup 4 Vision Continual Learning.

Method	Augment	Dataset n_splits Factor	CIFAR100				CUB200				ImageNetR				Cars			Avg
			5	10	20	50	5	10	20		5	10	20	50	5	10	20	
Task Arithmetic	No	0.30	78.68	48.16	3.63	1.73	45.34	1.66	0.47		73.33	39.35	11.77	2.33	44.85	0.68	0.53	25.18
		1.00	14.22	1.02	1.01	1.03	0.79	0.33	0.52		3.23	0.42	0.68	1.13	0.44	0.53	0.51	1.85
DARE	No	1.00	69.51	42.63	12.70	11.87	4.52	1.12	1.05		62.87	31.38	33.42	53.97	10.53	0.56	0.44	24.04
		0.30	78.65	48.18	3.65	1.76	45.50	1.64	0.41		73.23	39.28	11.85	2.35	44.37	0.60	0.55	25.14
TIES	No	1.00	14.27	1.03	1.20	1.04	0.79	0.33	0.52		3.18	0.40	0.67	1.17	0.44	0.56	0.51	1.86
		1.00	69.39	42.81	12.70	11.85	4.44	1.07	1.04		63.13	31.50	33.50	54.02	10.06	0.55	0.44	24.03
Breadcrumbs	No	1.00	81.57	76.72	75.47	73.76	53.56	53.18	52.97		75.80	74.97	74.67	72.85	57.59	52.78	48.65	66.04
		1.00	80.87	78.87	75.99	73.48	56.63	55.49	54.85		76.25	75.85	74.53	72.62	63.08	59.86	56.14	68.18
MagMax	No	0.30	78.25	78.05	68.42	23.03	56.80	55.21	45.53		74.55	75.67	72.90	64.78	63.24	57.27	18.00	59.41
		1.00	79.19	44.73	3.36	2.36	41.37	1.12	0.38		73.77	35.10	11.33	2.55	44.42	0.56	0.55	24.34
PCB	No	1.00	80.16	77.74	74.30	69.95	56.73	54.57	52.11		75.72	75.38	74.28	72.53	63.05	55.58	42.66	66.05
		0.50	81.45	79.15	76.98	74.73	56.83	55.23	54.99		76.42	75.93	74.52	72.58	63.40	60.55	56.71	68.53
ISO-C	No	1.00	79.34	71.91	68.54	66.77	46.46	43.67	45.10		73.13	71.98	73.05	72.83	43.79	32.16	25.57	58.16
		1.00	81.65	78.44	75.87	73.74	56.13	54.75	54.21		76.60	75.90	74.40	72.73	61.56	58.04	54.04	67.72
TSV	No	1.00	81.43	77.72	75.28	73.41	55.76	54.94	53.73		75.98	75.50	74.07	72.30	62.55	58.02	51.57	67.30
		1.00	80.41	78.77	76.13	73.60	57.01	55.57	54.73		75.98	75.85	74.42	72.52	63.96	60.84	56.49	68.31
TSV	No	1.00	80.82	75.41	67.64	58.12	59.04	57.84	52.64		82.20	81.13	77.03	72.92	67.62	61.29	39.25	66.64
		1.00	80.66	79.56	75.84	73.63	57.97	56.82	55.47		79.37	78.77	76.07	72.90	66.06	62.54	56.41	69.43
TSV	No	1.00	82.32	79.20	75.77	74.05	55.76	54.78	54.31		80.75	78.27	75.98	72.98	60.94	57.01	54.43	68.33
		1.00	81.16	79.13	76.04	73.47	57.04	55.59	54.99		79.45	77.18	74.95	72.20	63.61	60.65	56.92	68.74

Table 8: ViT-B-16 results for experimental setup 4 Vision Continual Learning.

Method	Augment	Dataset n_splits Factor	CIFAR100				CUB200			ImageNetR				Cars			Avg
			5	10	20	50	5	10	20	5	10	20	50	5	10	20	
Task Arithmetic	No	0.30	79.38	51.00	5.61	1.55	54.19	0.90	0.45	80.28	55.77	5.88	0.77	63.25	0.92	0.55	28.61
		1.00	9.63	1.02	0.91	0.99	0.36	0.33	0.45	7.55	0.55	0.55	0.67	0.61	0.49	0.56	1.76
DARE	Weight Weaving	1.00	71.13	45.23	28.50	27.05	22.68	0.35	0.45	71.48	48.98	28.25	27.03	38.70	0.62	0.62	29.36
	No	0.30	79.47	50.95	5.75	1.54	54.14	0.90	0.47	80.32	55.70	5.78	0.78	63.45	0.92	0.53	28.62
TIES		1.00	9.90	1.12	0.95	1.05	0.36	0.33	0.45	7.50	0.60	0.47	0.67	0.57	0.47	0.53	1.78
	Weight Weaving	1.00	71.12	45.30	28.58	27.00	22.70	0.35	0.43	71.32	48.87	28.17	26.92	38.94	0.58	0.53	29.34
Breadcrumbs	No	1.00	82.34	78.51	75.19	75.22	61.48	56.63	54.66	83.20	82.35	82.05	81.93	73.24	63.50	55.75	71.86
	Weight Weaving	1.00	83.23	79.72	77.70	76.51	62.44	60.51	58.53	84.02	83.12	82.73	81.63	74.12	68.86	64.28	74.10
MagMax	No	0.30	82.33	79.00	67.89	30.75	61.56	58.41	34.45	83.30	83.07	74.40	29.52	72.96	64.73	17.49	59.99
	Weight Weaving	1.00	76.63	42.21	3.25	1.23	44.70	0.31	0.50	77.53	39.25	2.13	0.62	55.86	0.55	0.53	24.67
PCB	No	0.50	83.02	78.43	74.15	71.70	62.32	56.99	51.26	84.03	82.62	80.40	79.70	73.10	62.74	45.09	70.40
	Weight Weaving	1.00	84.09	80.47	78.55	77.02	62.65	60.91	59.15	84.15	83.60	82.73	81.83	74.48	69.43	65.90	74.64
ISO-C	No	1.00	80.18	74.59	72.43	71.57	53.62	47.50	45.60	80.27	79.62	79.50	80.23	63.95	48.35	38.54	65.42
	Weight Weaving	1.00	83.74	79.78	77.80	76.59	62.91	59.92	58.20	83.92	83.53	82.87	81.82	73.68	67.53	62.92	73.94
TSV	No	1.00	82.02	79.64	76.57	74.46	61.72	60.80	58.41	83.60	82.83	82.37	81.47	74.10	70.43	66.47	73.92
	Weight Weaving	1.00	82.32	79.83	77.65	76.52	61.84	60.99	58.68	83.57	83.02	82.57	81.63	74.11	70.44	66.15	74.24
TSV	No	1.00	81.90	68.19	53.66	36.45	66.40	62.96	50.19	88.73	83.85	71.73	37.32	77.64	60.10	25.28	61.74
	Weight Weaving	1.00	83.88	79.24	76.61	75.57	64.03	62.25	59.70	88.77	87.15	85.10	83.65	76.93	71.56	63.90	75.60
TSV	No	1.00	84.79	80.55	78.81	76.99	63.26	60.11	58.09	87.85	86.03	84.62	83.50	73.98	66.63	61.91	74.79
	Weight Weaving	1.00	84.18	80.51	78.31	76.63	62.82	61.03	58.82	87.45	85.52	84.18	82.67	74.63	69.53	65.14	75.10

Table 9: ViT-L-14 results for experimental setup 4 Vision Continual Learning.

Method	Augment	Dataset n_splits Factor	CIFAR100				CUB200			ImageNetR				Cars			Avg
			5	10	20	50	5	10	20	5	10	20	50	5	10	20	
Task Arithmetic	No	0.30	87.37	55.79	2.98	1.38	58.91	2.36	2.49	90.35	78.13	25.22	13.70	71.74	0.51	0.47	35.10
		1.00	12.11	1.00	1.04	1.02	0.38	0.45	0.35	0.33	0.53	0.88	0.97	0.51	0.51	0.56	1.47
DARE	Weight Weaving	1.00	83.36	46.09	14.11	13.23	5.52	0.85	16.85	84.43	71.85	59.28	89.17	26.02	0.50	0.70	36.57
	No	0.30	87.46	55.77	2.94	1.40	59.06	2.30	2.38	90.30	78.10	25.00	13.78	71.86	0.56	0.46	35.10
TIES		1.00	12.24	1.00	1.05	1.02	0.43	0.45	0.33	0.35	0.52	0.90	0.95	0.49	0.51	0.56	1.48
	Weight Weaving	1.00	83.46	45.70	14.18	13.35	5.35	0.83	16.72	84.40	71.70	59.25	89.23	26.08	0.52	0.76	36.54
Breadcrumbs	No	1.00	88.89	86.61	85.92	84.17	69.99	66.76	66.76	91.40	91.28	90.35	89.30	80.72	72.29	68.97	80.96
	Weight Weaving	1.00	88.42	86.85	85.74	84.02	70.52	67.64	66.43	91.77	91.05	90.03	89.02	82.91	78.91	75.53	82.06
MagMax	No	0.30	86.30	86.76	83.12	56.79	68.29	67.45	63.76	91.07	91.15	90.28	89.98	82.17	77.08	23.89	75.58
	Weight Weaving	1.00	87.27	59.95	3.71	1.27	58.60	3.61	2.95	90.40	80.68	38.60	23.67	69.92	0.56	0.50	37.26
PCB	No	0.50	87.69	86.51	85.40	83.57	69.74	67.64	65.93	91.63	91.10	90.18	89.35	82.43	75.69	62.64	80.68
	Weight Weaving	1.00	88.72	87.48	86.27	84.54	71.07	67.95	66.67	91.88	91.13	90.12	89.03	83.24	79.34	75.84	82.38
ISO-C	No	1.00	87.29	83.44	81.78	82.29	51.67	55.56	62.60	90.05	90.52	90.52	89.78	67.73	31.86	25.63	70.77
	Weight Weaving	1.00	88.84	87.21	86.03	83.97	71.30	67.88	66.52	91.72	91.20	90.22	89.08	82.40	78.91	73.86	81.97
TSV	No	1.00	89.00	87.24	86.07	83.92	71.44	67.92	66.78	91.78	91.15	89.85	88.92	82.81	77.42	72.96	81.95
	Weight Weaving	1.00	88.00	86.74	85.61	83.99	70.06	67.54	66.45	91.48	91.10	89.83	89.00	83.05	79.38	76.04	82.02
TSV	No	1.00	88.60	83.48	79.45	72.51	71.97	70.25	68.38	94.72	94.02	92.57	91.78	86.64	81.05	64.71	81.04
	Weight Weaving	1.00	87.95	87.02	86.00	84.26	70.33	68.69	67.14	94.07	93.62	91.32	89.98	84.32	81.47	76.81	83.47
TSV	No	1.00	89.17	87.53	86.79	84.90	71.99	68.59	67.09	94.12	93.42	91.40	89.80	82.81	77.40	74.73	82.84
	Weight Weaving	1.00	88.52	87.06	85.89	84.10	70.64	68.02	66.53	93.67	92.97	90.58	89.27	83.32	79.63	76.36	82.61

Table 10: ViT-B-32 results for experimental setup 8 Vision Domain Generalization.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	45.18	72.01	36.15	45.57	45.25	49.09	34.84	29.54	44.70
		1.00	0.49	20.78	11.70	14.27	3.95	0.44	5.74	6.65	8.00
DARE	No	0.30	45.18	72.01	36.15	45.57	45.27	49.11	34.84	29.52	44.71
		1.00	0.49	20.78	11.70	14.27	3.95	0.44	5.74	6.65	8.00
TIES	No	1.00	43.34	76.08	36.74	50.40	43.51	46.63	34.52	30.24	45.18
	Weight Weaving	1.00	56.72	72.54	46.56	46.50	58.51	60.80	41.76	36.09	52.43
Breadcrumbs	No	0.30	58.56	65.11	48.93	43.02	61.73	62.48	43.35	35.52	52.34
		1.00	41.18	70.77	31.26	45.43	42.60	45.59	34.63	29.25	42.59
MagMax	Weight Weaving	1.00	57.44	67.21	48.19	44.73	60.76	61.82	43.09	35.35	52.32
	No	0.50	54.78	74.15	47.30	46.93	56.40	59.28	41.01	34.27	51.76
PCB		1.00	33.16	76.42	34.19	47.55	32.86	33.79	29.89	22.99	38.86
	Weight Weaving	1.00	55.38	73.02	47.11	46.61	56.98	59.84	41.12	34.49	51.82
ISO-C	No	1.00	54.22	76.04	40.52	53.30	55.25	57.49	40.00	35.66	51.56
	Weight Weaving	1.00	58.23	71.26	46.44	45.83	60.57	61.92	42.71	36.08	52.88
TSV	No	1.00	53.67	78.26	49.93	53.56	53.25	60.20	40.37	36.32	53.20
	Weight Weaving	1.00	57.98	71.75	52.41	46.15	60.38	62.53	42.45	36.03	53.71
TSV	No	1.00	48.81	82.08	41.44	57.64	43.68	52.47	37.29	34.33	49.72
	Weight Weaving	1.00	57.12	75.28	50.04	49.97	57.59	61.34	42.29	36.89	53.81

Table 11: ViT-B-16 results for experimental setup 8 Vision Domain Generalization.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	53.65	86.82	35.81	64.96	51.71	56.91	38.88	39.54	53.54
		1.00	0.86	28.03	10.26	18.93	2.86	1.36	5.69	4.99	9.12
DARE	No	0.30	53.65	86.82	35.81	64.96	51.71	56.91	38.88	39.54	53.54
		1.00	0.86	28.03	10.26	18.93	2.86	1.36	5.69	4.99	9.12
TIES	No	1.00	53.56	87.88	39.56	65.34	50.62	55.72	38.56	38.98	53.78
		1.00	63.16	84.49	51.26	62.50	63.79	64.56	44.47	44.95	59.90
Breadcrumbs	No	0.30	64.88	77.88	54.04	60.01	66.06	65.60	44.73	45.38	59.82
		1.00	50.80	86.42	32.67	64.19	47.87	54.95	37.93	38.92	51.72
	Weight Weaving	1.00	64.42	81.48	52.74	61.42	65.35	65.36	44.47	45.33	60.07
		1.00	61.09	85.89	48.70	60.73	61.29	63.53	43.24	43.59	58.51
MagMax	No	1.00	44.50	88.44	32.85	59.67	40.51	46.29	33.67	31.60	47.19
		1.00	61.88	85.41	48.81	61.44	62.41	63.97	43.40	44.01	58.92
PCB	No	1.00	59.58	87.18	49.70	67.26	58.73	62.22	42.34	42.50	58.69
		1.00	64.03	82.78	53.89	62.48	64.84	65.27	44.20	44.85	60.29
ISO-C	No	1.00	60.60	87.27	51.96	66.36	57.27	63.96	41.91	44.04	59.17
		1.00	64.41	80.87	54.04	62.17	65.44	65.44	44.31	45.10	60.22
TSV	No	1.00	55.28	90.47	37.48	68.95	45.35	58.63	38.78	40.80	54.47
		1.00	63.51	86.06	51.00	64.16	62.67	64.94	43.46	45.47	60.16

Table 12: ViT-L-14 results for experimental setup 8 Vision Domain Generalization.

Method	Augment	Dataset Factor	Cars	MNIST	EuroSAT	SVHN	RESISC45	SUN397	DTD	GTSRB	Avg
Task Arithmetic	No	0.30	71.17	87.39	57.41	66.84	67.27	65.92	53.09	47.50	64.57
		1.00	0.86	8.83	9.70	18.39	3.87	0.66	3.67	2.39	6.05
DARE	No	0.30	71.21	87.39	57.44	66.83	67.30	65.94	53.09	47.49	64.59
		1.00	0.86	8.83	9.70	18.39	3.87	0.66	3.67	2.41	6.05
TIES	No	1.00	69.73	81.56	54.30	69.10	65.32	65.29	52.66	47.28	63.16
		1.00	76.88	84.73	70.15	67.56	70.79	68.70	56.06	51.96	68.35
Breadcrumbs	No	0.30	77.83	81.13	67.89	64.84	71.79	68.70	55.80	52.79	67.60
		1.00	71.63	87.25	52.96	65.45	67.25	66.55	53.67	49.33	64.26
	Weight Weaving	1.00	77.59	82.36	68.22	65.48	71.67	68.75	55.90	52.84	67.85
		0.50	74.95	85.20	66.07	69.14	69.27	67.80	54.63	49.70	67.10
MagMax	No	1.00	53.34	71.11	20.07	67.28	49.32	53.36	46.17	38.58	49.90
		1.00	75.56	85.95	68.81	68.53	69.90	68.25	55.43	50.51	67.87
PCB	No	1.00	72.42	89.16	60.70	67.00	65.87	66.36	54.52	49.02	65.63
		1.00	77.12	85.28	69.63	66.72	71.00	68.77	56.28	52.54	68.42
ISO-C	No	1.00	76.81	88.08	72.37	68.25	68.03	68.59	56.70	52.75	68.95
		1.00	77.69	83.64	70.41	66.73	71.40	68.90	56.22	53.11	68.51
TSV	No	1.00	73.35	90.21	62.00	70.97	64.22	66.96	54.63	49.95	66.54
		1.00	77.03	85.39	70.74	68.38	70.65	68.75	56.01	52.62	68.70

A3 Weight correlation for sequential and individual fine-tuning

We analyze how the fine-tuning strategy affects the correlation between task vectors by comparing two distinct approaches. The first, sequential fine-tuning, where training on a new task T_i initializes from the weights θ_{t-1} , previously fine-tuned on tasks T_1, T_2, \dots, T_{n-1} . In contrast, the second strategy, which we term individual fine-tuning, trains a model for each task T_i starts from the weights θ_{pre} , the pre-trained model, which we name this setting in individual fine-tuning. Figure 3 shows the resulting task vector correlations for the sequential (left) and individual (right) approaches for CIFAR100 dataset and $T = 10$ tasks. While the precise implications of this increased correlation for model merging remain an open question, we hypothesize that it indicates greater parameter redundancy across tasks. Consequently, effectively merging models trained in a sequential approach may require specialized methods tailored to address this high level of redundancy.

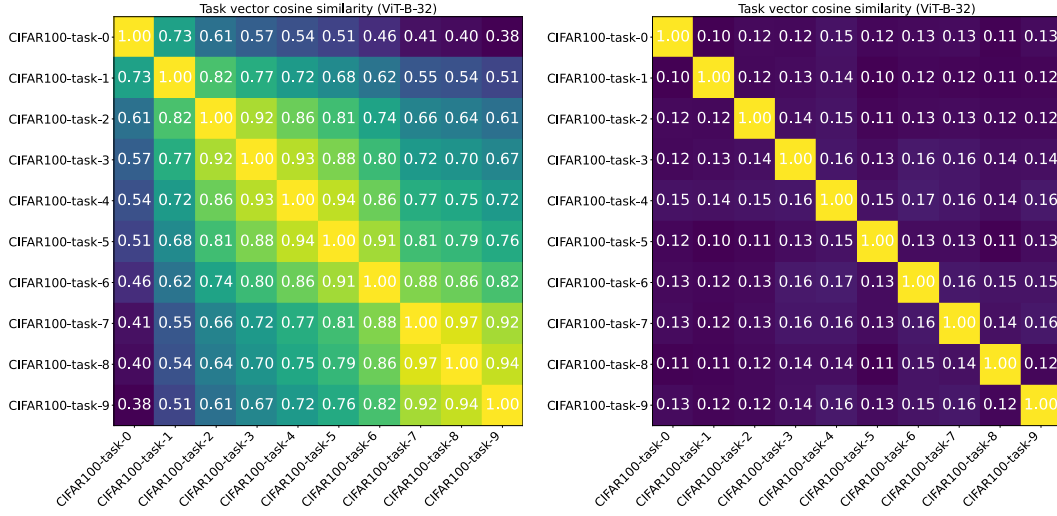


Figure 3: Pair-wise cosine similarity between task vectors for two fine-tuning strategies using the CIFAR-100 dataset split into 10 tasks. The left panel shows the results of sequential fine-tuning, where each task initializes from the weights of the previous one, leading to highly correlated task vectors (high cosine similarity). In contrast, the right panel shows individual fine-tuning, where each task initializes independently from the original pre-trained model. The latter results in low cosine similarity, indicating that the task vectors are substantially less correlated.