This Appendix first provides details of the three baseline models in Appendix A. Comparisons of different incremental learning approaches with SPL are provided in Appendix B. In Appendix C, we provide additional experimental results on CIFAR100 and CUB200.

A. Details of Baseline Models

We choose three prevalent baseline models, i.e., the naive CE-trained model, fantasy-based model FACT [49], and SAVC as the baseline models [29].

- **CE**: The base model is trained simply using cross-entropy loss in the base session. For the incremental sessions, the feature extractor is frozen, and only the classifier is updated.
- **FACT** [49]: During the base session, virtual new classes are synthesized by manifold mixup [36] to assist the base training, intending to save feature space for new classes. For incremental sessions, the model is updated by adding new prototypes to the classifier.
- **SAVC** [29]: Contrastive learning [13] is adopted in base session to learn compact representations. During the incremental sessions, multiple prototypes from each new class are ensembled as new classifier parameters.

B. Comparison of Different Incremental Learning Approaches

We conduct experiments on comparing different incremental learning approaches on fine-grained dataset CUB200 to verify the effectiveness when learning new classes. As shown in Tab. A3, we compare the proposed SPL with two commonly used incremental learning approaches, i.e., prototype-based update in Sec. 3.1.2, finetuning the last layer of the model by CE using few-shot data [29, 44]. We use the same base model, followed by 10 incremental sessions. The prototype-based model obtains the lowest new class performance, average performance of ten sessions, and harmonic accuracy since it cannot separate the new classes from other classes efficiently. The prototype-based model does obtain the highest old class performance and the least drop in PD, as it does not involve any update of feature space. The finetune approach boosts the new class performance by updating the feature space, hence obtaining higher average performance and harmonic accuracy. Compared to the finetuning approach, our SPL expands the new class feature distributions and facilitates a wider margin between classes. Therefore, SPL can retrain higher base class performance while learning new classes more effectively, achieving the highest performance of new classes, average performance and harmonic accuracy.

Table A3. Comparison of different incremental Methods. "Prototype-based" refers to the approach that simply updates the new class prototypes during incremental learning. "Finetune by CE" denotes using CE to finetune the last layer of the model with few-shot data.

Incremental Method	CUB200					
	Base↑	**Old↑**	**New↑**	**Avg↑**	**PD↓**	**H.↑**
Prototype-based [49]	81.31	**76.96**	47.00	68.88	**4.35**	58.35
Finetune by CE [29]	81.31	76.54	47.71	68.93	4.77	58.78
SPL	81.31	76.68	**47.85**	**69.12**	4.63	**58.93**

C. More Benchmark Results

We also present the performance of our method on the CIFAR100 and CUB200 datasets, as shown in Tab. A4 and Tab. A5, respectively. On CIFAR100, our approach boosts the performance of baseline methods in all sessions. Our method improves the final performance of three baselines by at least 0.65% and boosts the average performance on all incremental sessions. The improvement is attributed to the covariance constraint loss and semantic perturbation learning, which promote effective class separation and few-shot new class learning. On the fine-grained dataset CUB200, which includes 200 classes, our method achieves a final performance of 62.70%, demonstrating the effectiveness of our approaches. We obtain an improvement in final accuracy of 2.61% by applying our approach to the CE baseline model. In session 1 and session 2, our method yields lower performance on the CE model due to the imbalance of base class and new class in the testing data, but in the following incremental sessions, our approach is able to boost the overall performance.

Table A4. Incremental learning performance on CIFAR100 under 5-way 5-shot setup. "Avg Acc." represents the average accuracy of all sessions. "Final Improv." calculates the improvement of our method after learning in the final session. **Bold** represents best performance. $*$ indicates that we reproduce the results using public open-source code

Methods	Accuracy in each session (%) ↑									Avg Acc.	Final Improv.
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**		
iCaRL [26]	64.10	53.28	41.69	34.13	27.93	25.06	20.41	15.48	13.73	32.87	+39.41
NCM [11]	64.10	53.05	43.96	36.97	31.61	26.73	21.23	16.78	13.54	34.22	+39.60
Data-free Replay [22]	74.40	70.20	66.54	62.51	59.71	56.58	54.52	52.39	50.14	60.78	+3.00
Self-promoted [54]	64.10	65.86	61.36	57.45	53.69	50.75	48.58	45.66	43.25	54.52	+9.89
CEC [18]	73.07	68.88	65.26	61.19	58.09	55.57	53.22	51.34	49.14	59.53	+4.00
MetaFSCIL [5]	74.50	70.10	66.84	62.77	59.48	56.52	54.36	52.56	49.97	60.79	+8.05
C-FSCIL [10]	77.47	72.40	67.47	63.25	59.84	56.95	54.42	52.47	50.47	61.64	+3.17
LIMIT [51]	72.32	68.47	64.30	60.78	57.95	55.07	52.70	50.72	49.19	59.06	+3.95
CE	76.87	72.38	68.06	63.83	60.52	57.76	55.47	53.25	50.94	62.12	+2.20
CE-Ours	78.27	73.80	69.69	65.53	62.07	59.33	57.22	54.75	52.30	62.21	+0.84
FACT* [49]	78.38	71.86	67.87	64.10	60.70	57.75	55.83	53.6	51.34	62.17	+1.80
FACT*-Ours	**79.12**	72.62	68.49	64.31	61.51	58.64	56.38	54.22	52.34	62.82	+0.80
SAVC* [29]	78.98	73.02	68.69	64.49	60.91	58.08	55.79	53.61	51.75	62.81	+1.39
SAVC*-Ours	79.00	**73.29**	**68.84**	**64.75**	**61.60**	**58.74**	**56.84**	**55.12**	**53.14**	**63.48**	

Table A5. Performance of FSCIL in each session on CUB200 under 10-way 5-shot setup and comparison with other studies. "Average Acc." is the average accuracy of all sessions. "Final Improv." calculates the improvement of our method in the last session. **Bold** represents best performance. $*$ indicates that we reproduce the results using public open-source code.

Methods	Accuracy in each session (%) ↑											Avg Acc.	Final Improv.
	0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**		
iCaRL [26]	68.68	52.65	48.61	44.16	36.62	29.52	27.83	26.26	24.01	23.89	21.16	36.67	+41.54
Data-free Replay [22]	75.90	72.14	68.64	63.76	62.58	59.11	57.82	55.89	54.92	53.58	52.39	61.52	+10.31
LDC [21]	77.89	76.93	74.64	70.06	68.88	67.15	64.83	64.16	63.03	62.39	61.58	68.32	+1.12
CEC [18]	75.85	71.94	68.50	63.50	62.43	58.27	57.73	55.81	54.83	53.52	52.28	61.33	+10.42
LIMIT [51]	76.32	74.18	72.68	69.19	68.79	65.64	63.57	62.69	61.47	60.44	58.45	66.67	+4.25
MetaFSCIL [5]	75.90	72.41	68.78	64.78	62.96	59.99	58.3	56.85	54.78	53.82	52.64	61.93	+10.06
CE	79.32	75.67	72.56	67.42	66.46	62.00	60.85	59.31	57.78	56.88	55.73	64.91	+6.97
CE-Ours	79.59	75.32	72.31	67.46	66.68	63.61	62.68	61.07	59.09	59.20	58.34	65.71	+4.36
FACT*	77.28	73.67	70.19	65.59	64.77	61.60	60.68	58.89	57.38	57.26	56.11	63.87	+6.59
FACT*-Ours	77.78	74.23	70.42	65.97	65.31	61.58	61.42	59.61	57.42	57.26	56.49	65.15	+6.21
SAVC*	81.31	77.35	74.49	69.65	69.78	67.10	66.48	64.09	63.16	62.48	61.81	68.88	+0.89
SAVC*-Ours	**82.67**	**78.58**	**75.66**	**70.83**	**70.37**	**67.30**	**66.80**	**65.57**	**64.01**	**63.45**	**62.70**	**69.81**	