

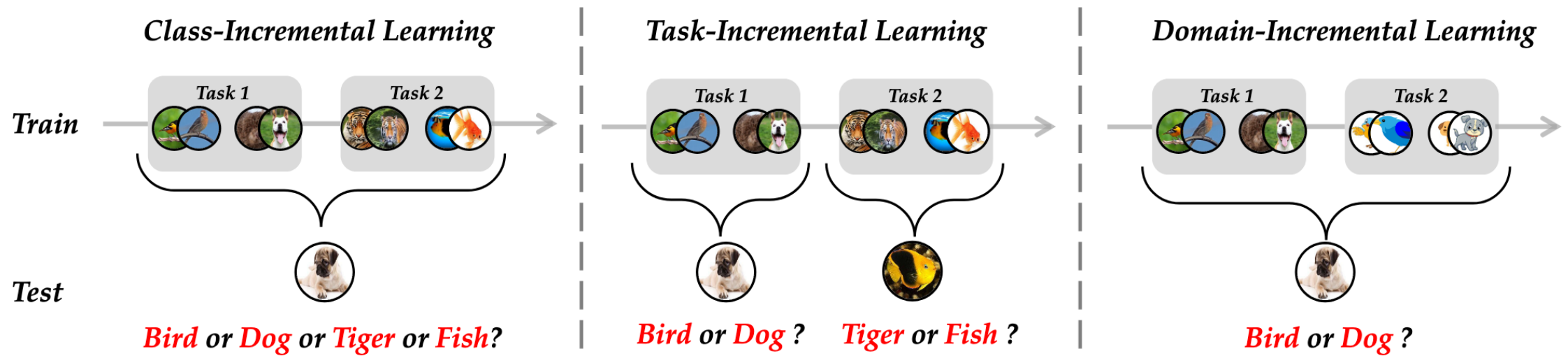
GFedCL: Graph-Based Federated Continual Learning with Spatial and Temporal Awareness

Qingyang Yu¹, Yang Hua², Qizhen Zhang³, Hao Wang¹

Stevens Institute of Technology¹, Queen's University
Belfast², University of Toronto³

Recall: Federated Continual Learning (FCL)

FCL categories: CIL, TIL, and DIL

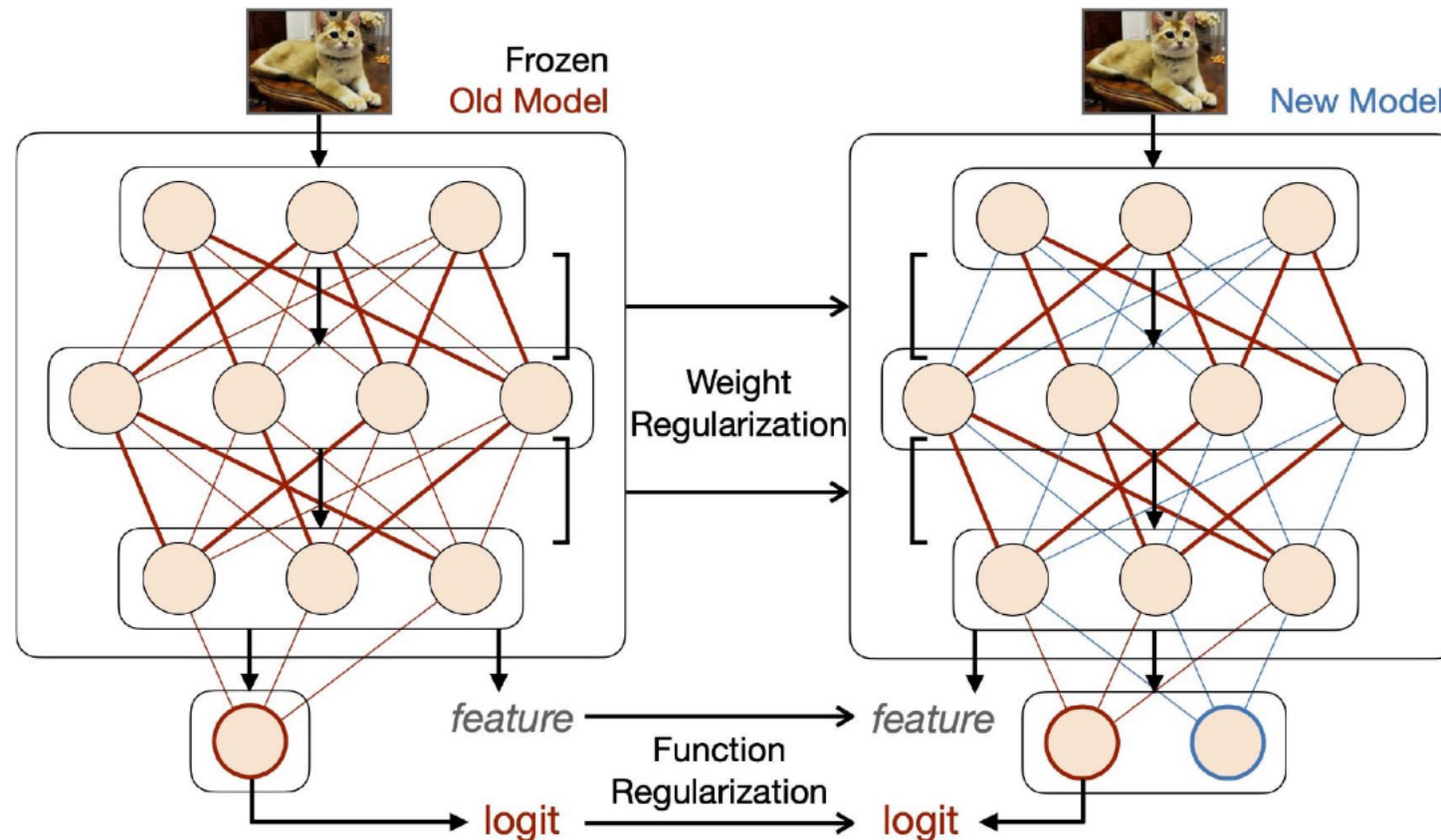


Images are from [CIL survey](#)



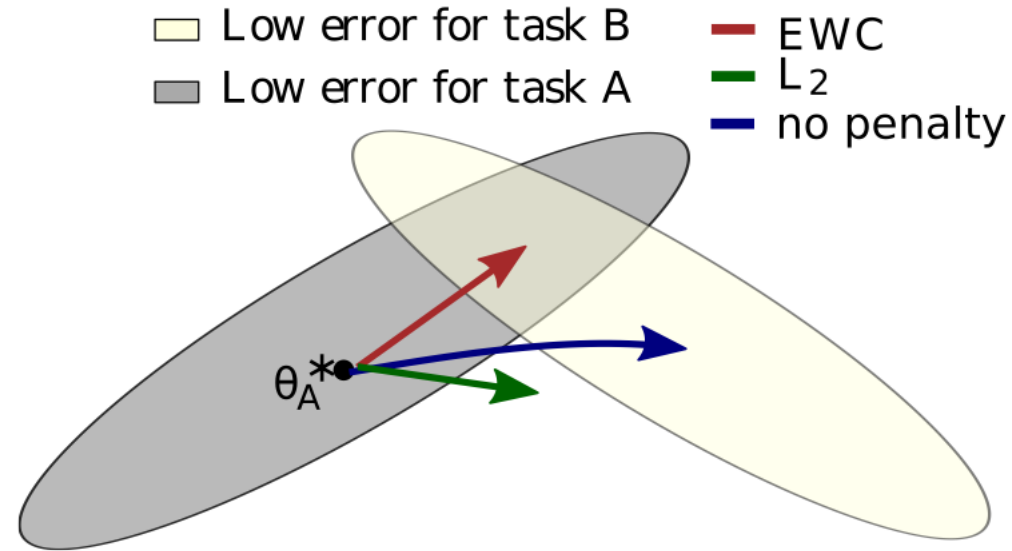
Recall: Regularization-Based FCL

Constrain the divergence



Images are from [survey](#)

Recall: Regularization-Based FCL

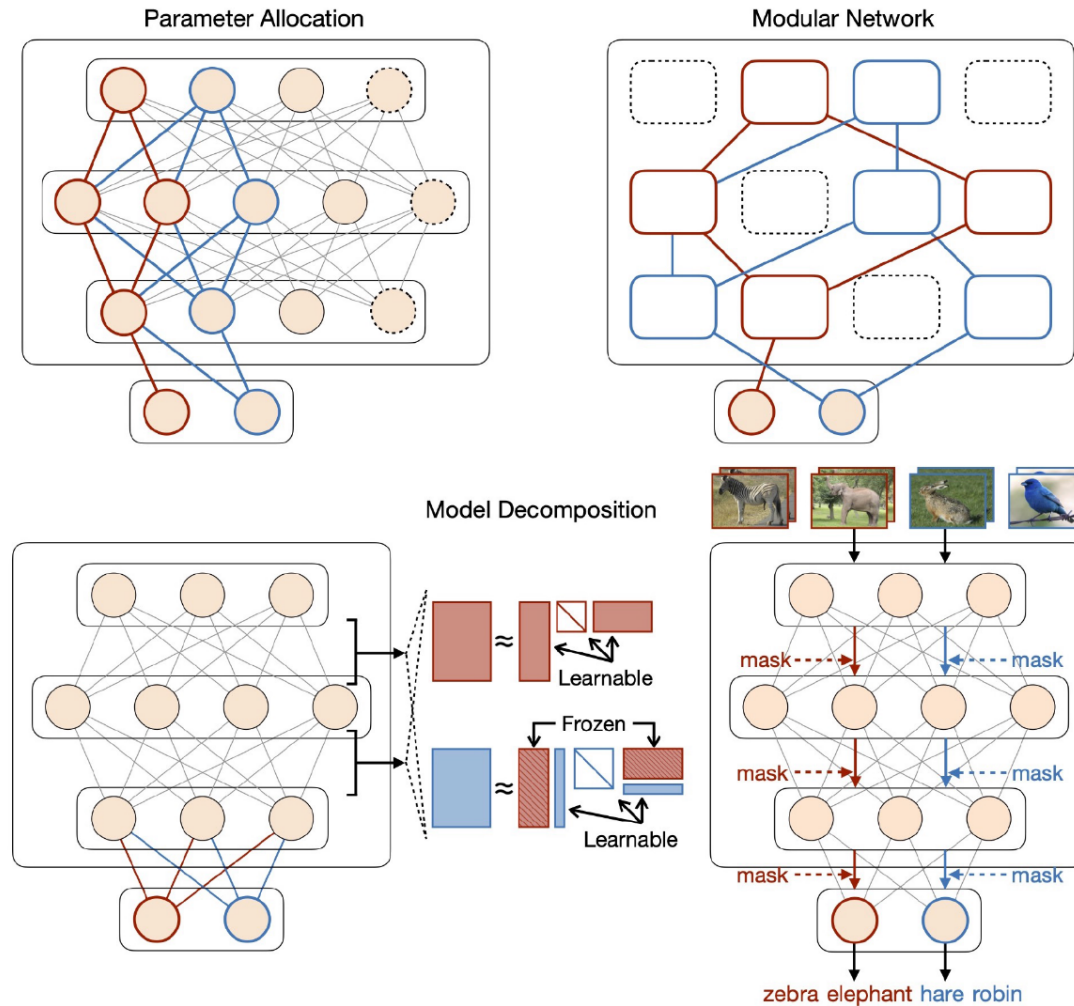


$$\mathcal{L}(\theta) = \mathcal{L}_B(\theta) + \sum_i \frac{\lambda}{2} F_i(\theta_i - \theta_{A,i}^*)$$

Images are from [EWC](#)

Recall: Architecture-Based FCL

Split the model into several parts for different roles.

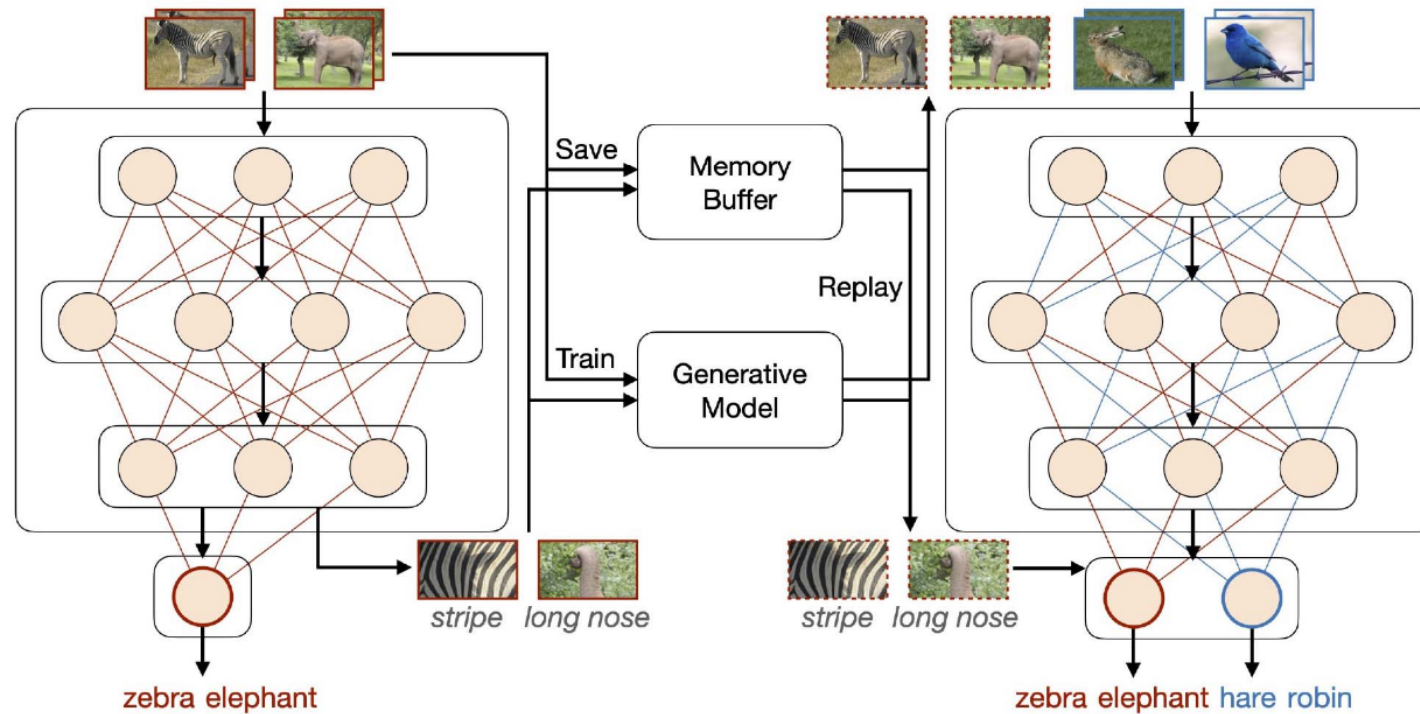


Images are from [HAT](#)



Recall: Replay-based FCL

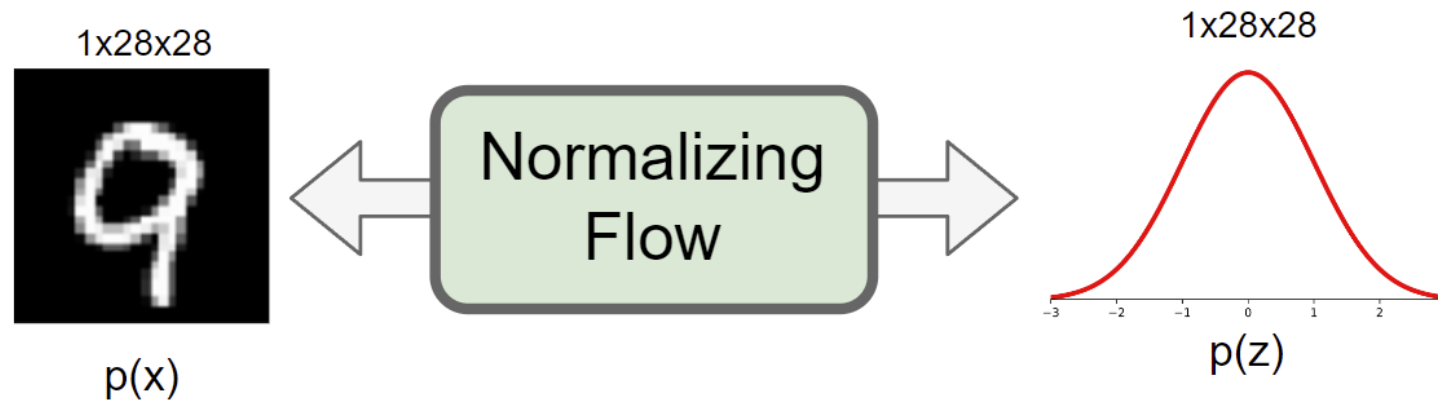
Replay synthetic data samples



Images are from [survey](#)

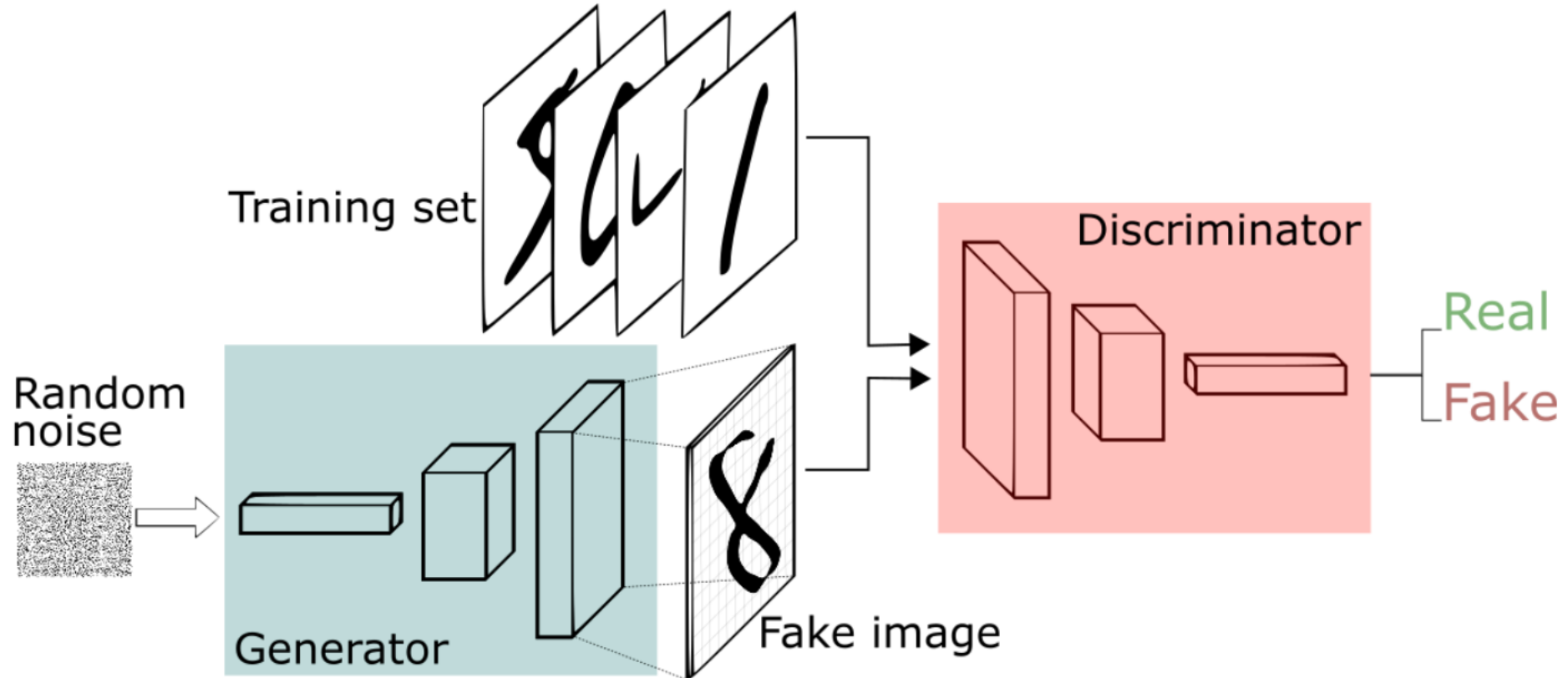
Recall: Normalizing Flow

Map samples into a distribution, e.g., Gaussian, and then generate



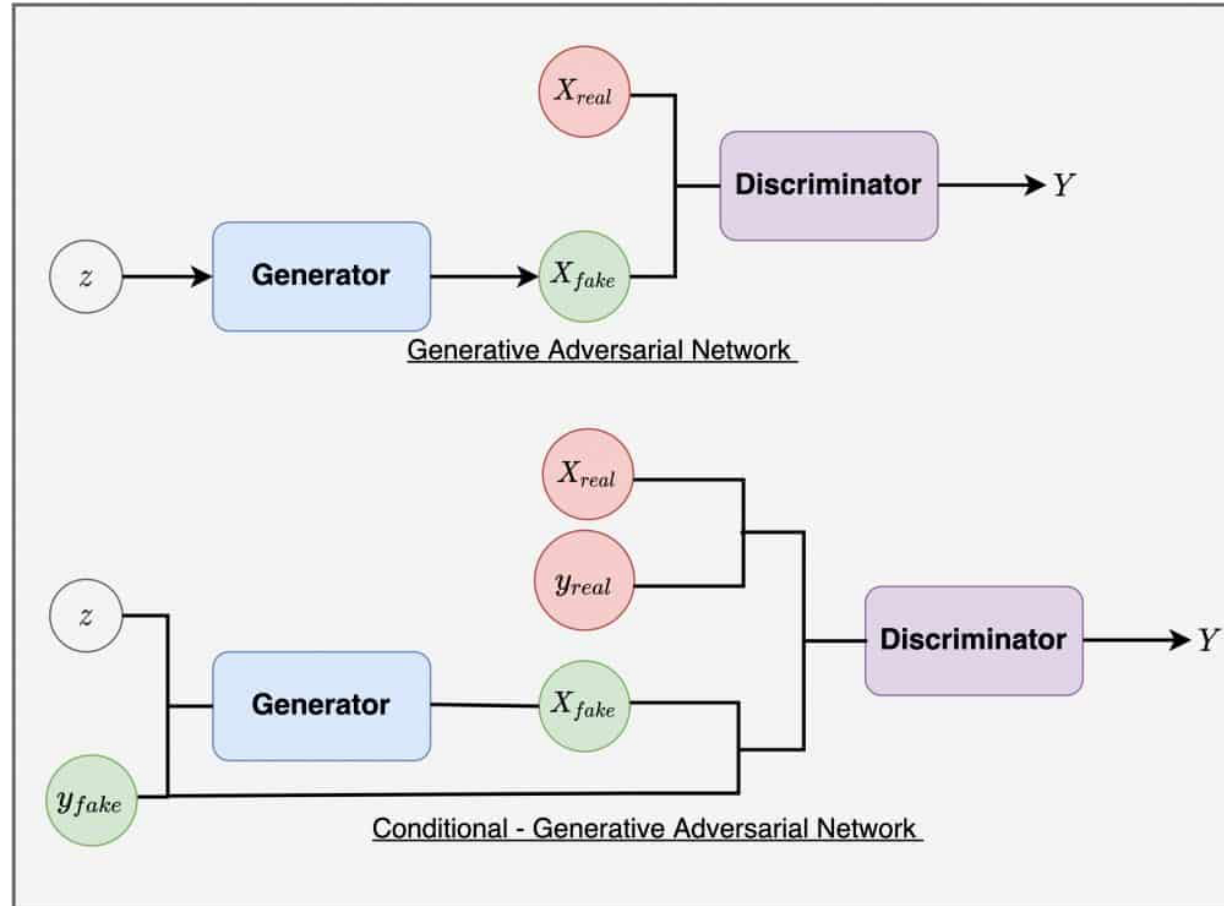
Images are from [survey](#)

Recall: GAN



Images source: [GAN](#)

Conditioning Determines Synthetic Data Quality



Images source: [GAN](#)



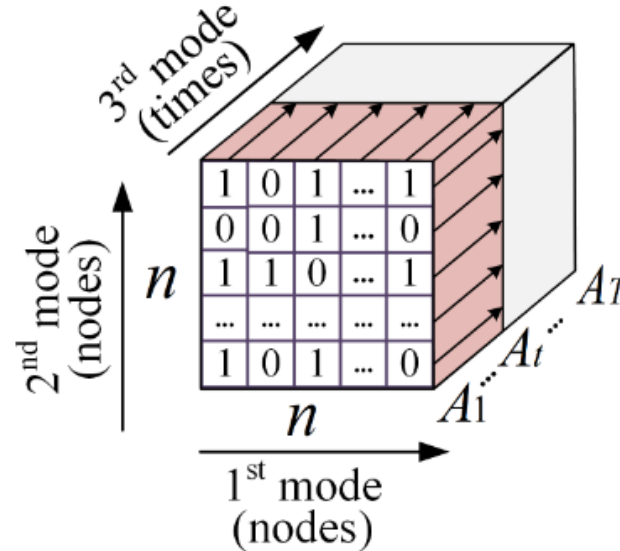
What Conditioning Does FCL Have?

Labels vs. ?

Spatial and Temporal Variation Patterns

Spatial variational patterns: Relationships between clients within one task

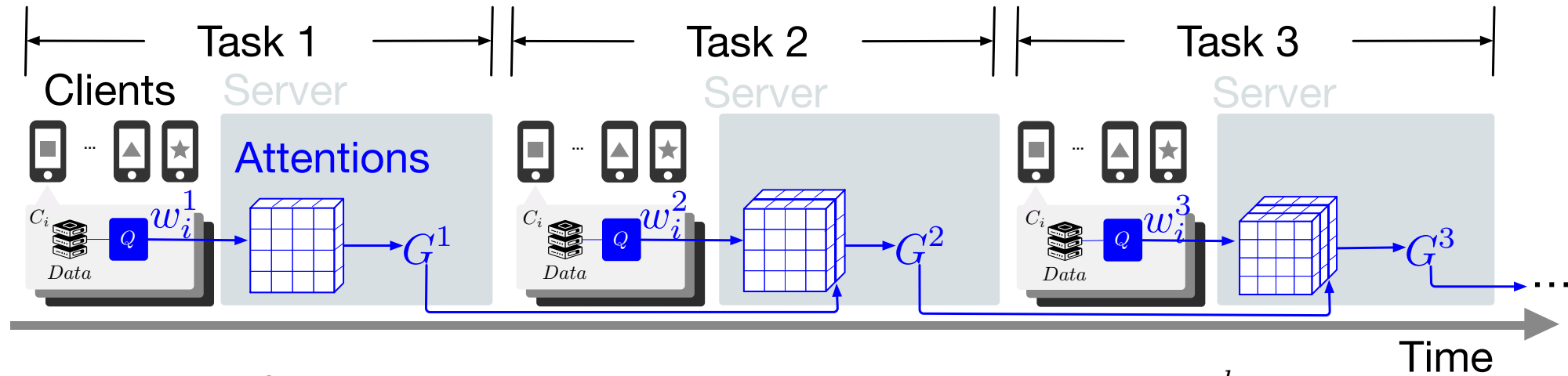
Temporal variational patterns: Evolution patterns of spatial variational patterns



Images from [GAEN](#)

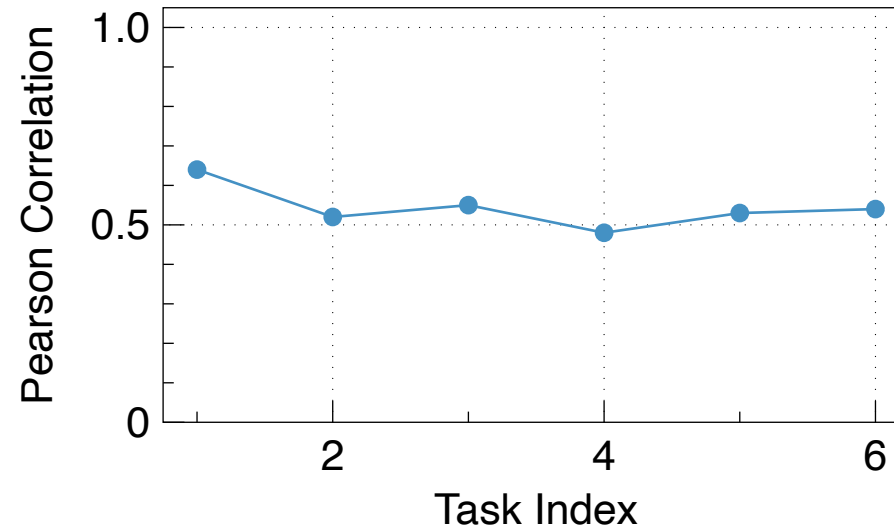
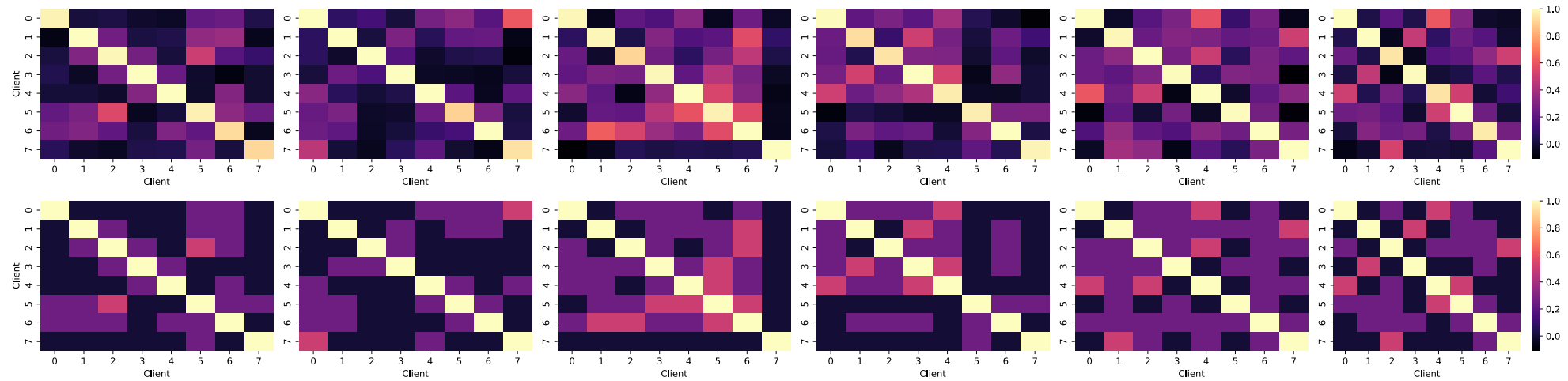


Relational-Graph Construction

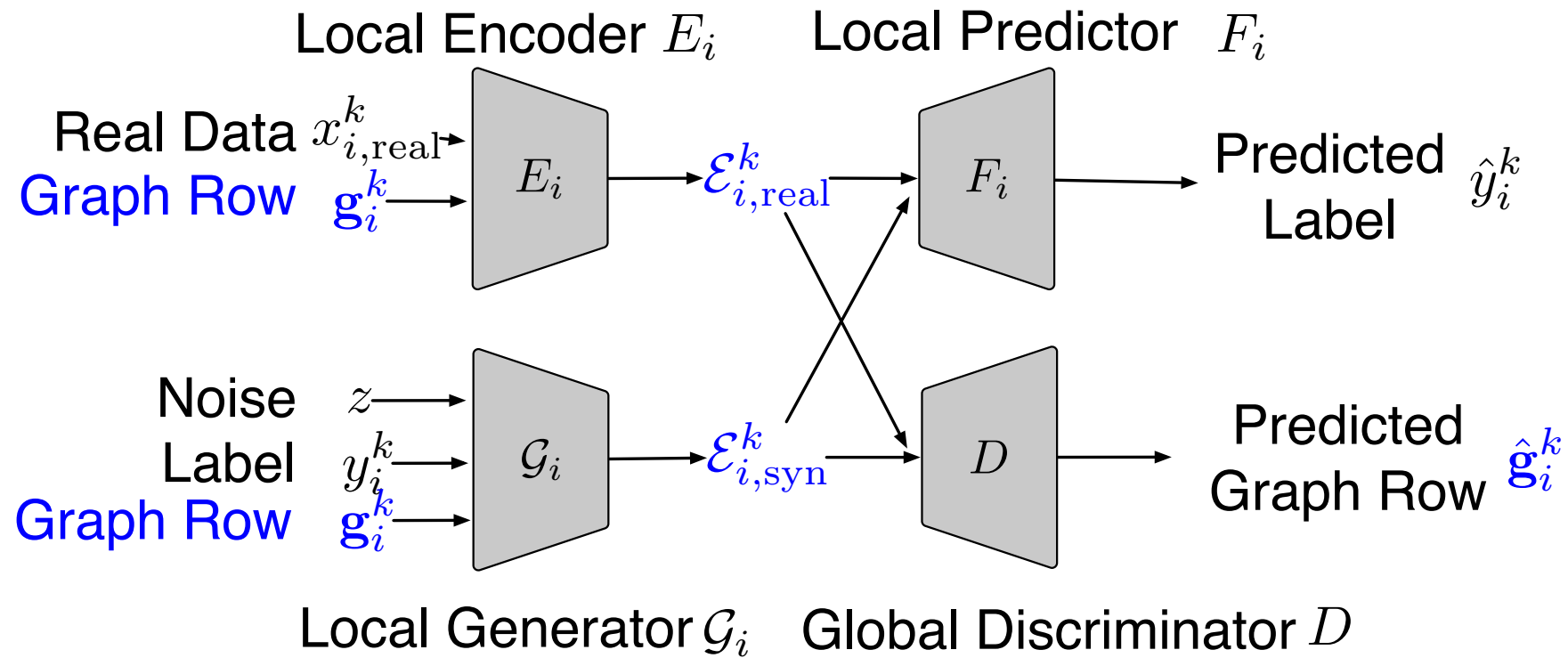


- Classifier: Q .
- Classifier updates of client i with task k : w_i^k .
- Spatial attention: $\alpha_{ij}^k = \text{softmax}_j (a(w_i^k, w_j^k)) \in (0, 1)$.
- Temporal attention: $\beta_{ij}^k = \text{softmax}_j (b(H_i^k, H_j^k, m)) \in (0, 1)$, where $H_i^k = \{a_{i,;}^t\}_{t=1}^k$.

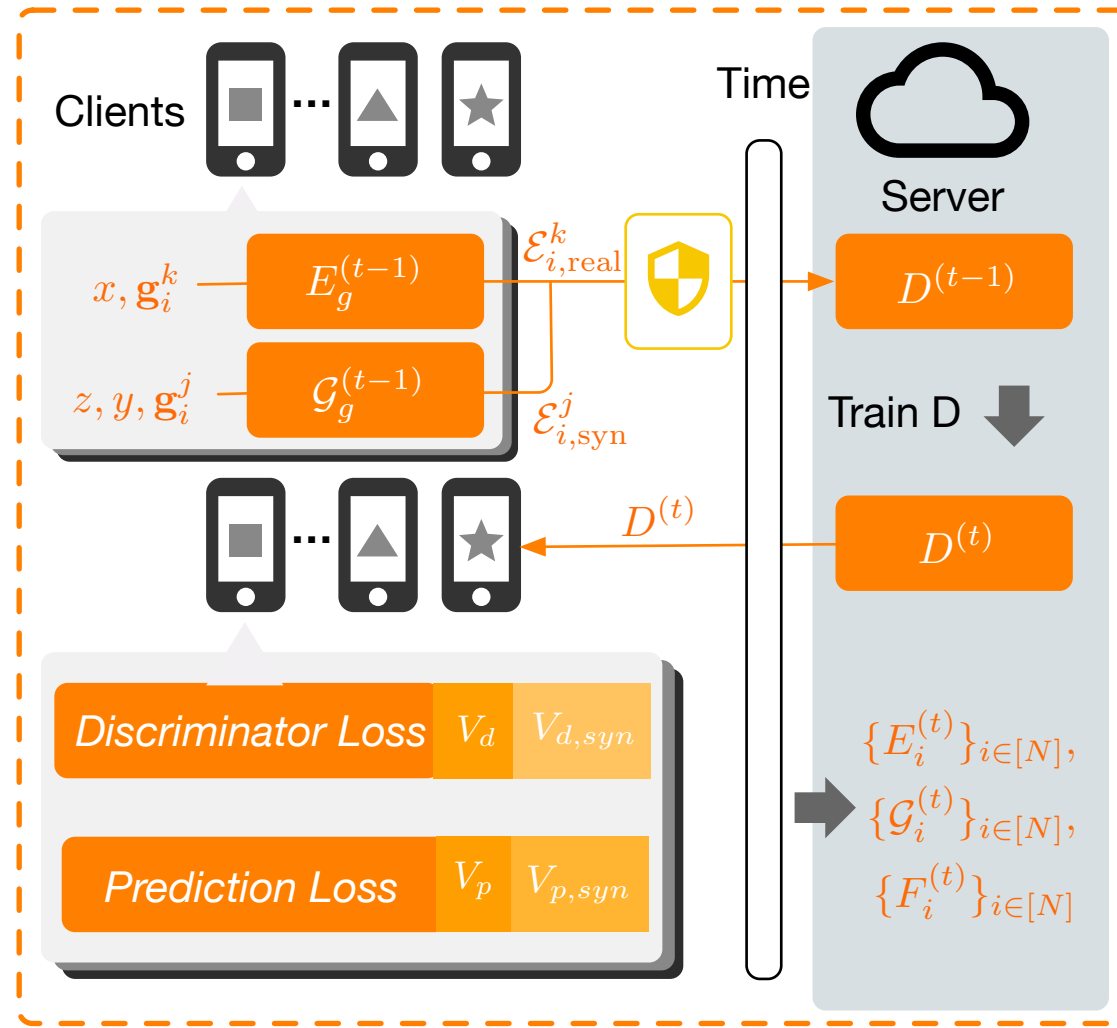
Meaning of Relational Graph



Utilize Relational Graphs as Conditioning

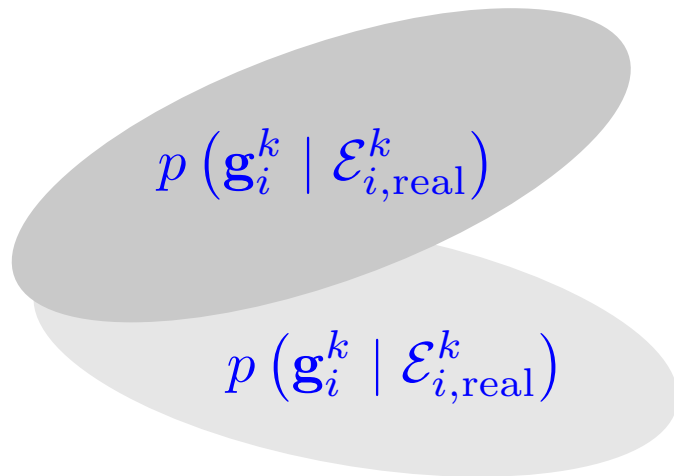


Utilize Relational Graphs as Conditioning



Representation Alignment at Optimum

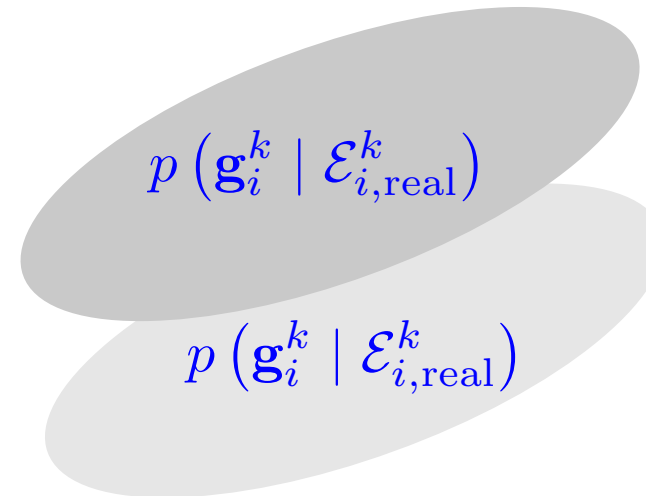
Before Training



Training →

After Training

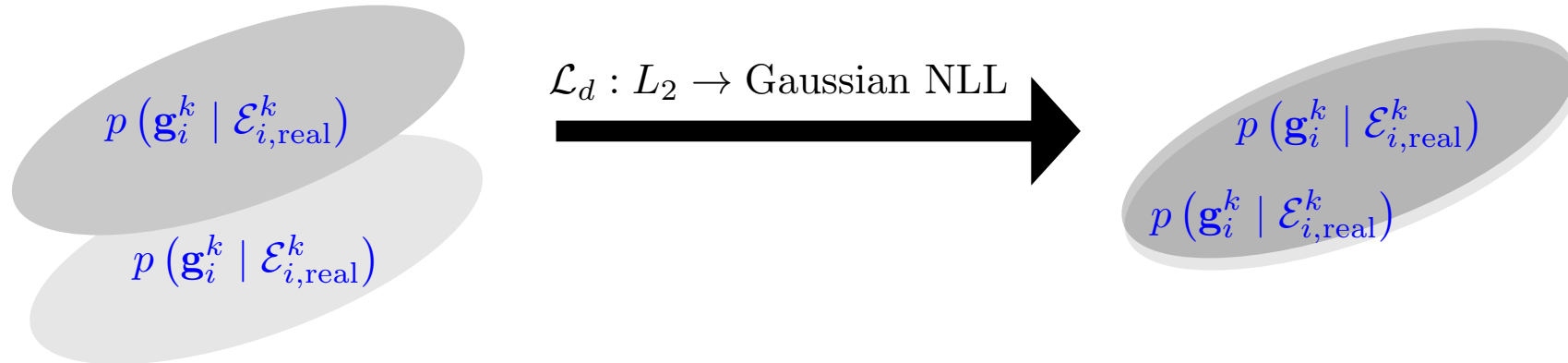
Mean Alignment: $\forall i \in [N], \forall k \in [K],$
 $\mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\text{real}}^k] = \mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\text{syn}}^k] = \mathbb{E}[\mathbf{g}_i^k] = \mathbf{g}^*.$



Higher-Order Alignment

Mean Alignment: $\forall i \in [N], \forall k \in [K],$
 $\mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\text{real}}^k] = \mathbb{E}[\mathbf{g}_i^k | \mathcal{E}_{i,\text{syn}}^k] = \mathbb{E}[\mathbf{g}_i^k] = \mathbf{g}^*.$

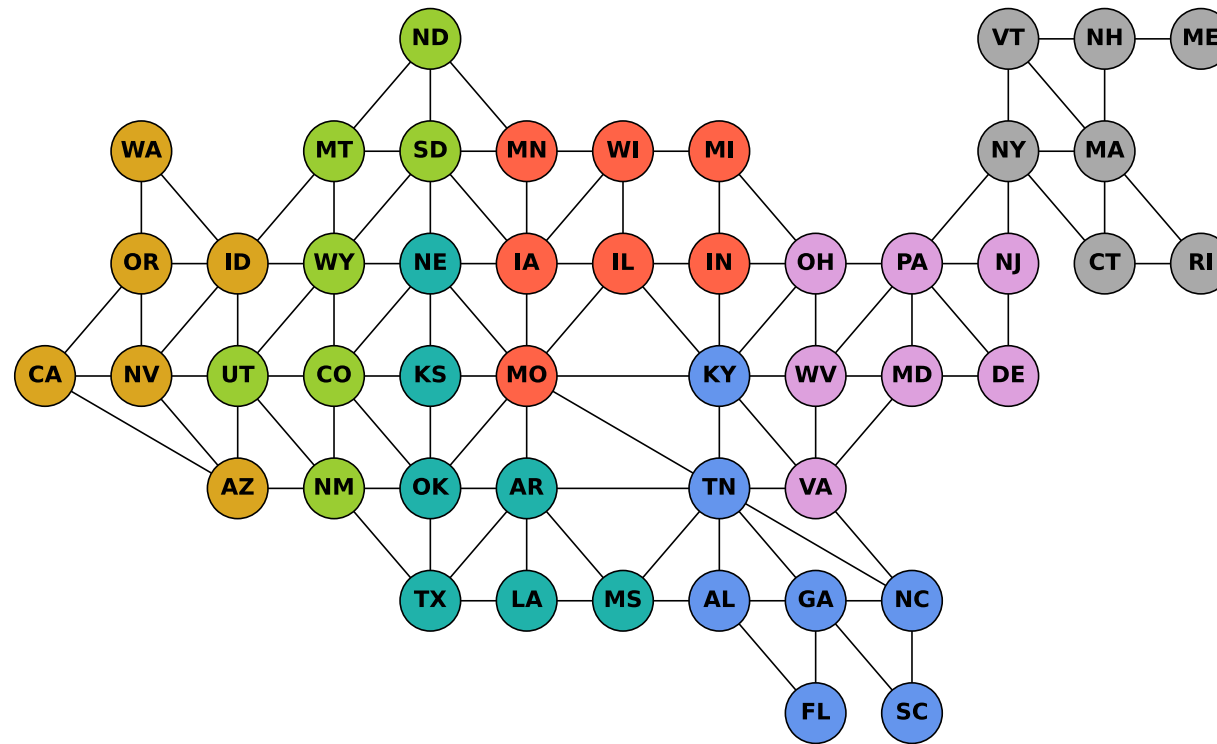
Mean and Variance Alignment



Main Results

Method	EMNIST-Letters		CIFAR-100		TinyImageNet		TCGA-BRCA		US-States MSE ($\times 10^{-2}$) $\downarrow$
	Acc (%) $\uparrow$	Forgetting (%) $\downarrow$	Acc (%) $\uparrow$	Forgetting (%) $\downarrow$	Acc (%) $\uparrow$	Forgetting (%) $\downarrow$	Acc (%) $\uparrow$	Forgetting (%) $\downarrow$	
GLFC	43.47 $\pm$ 2.14	8.74 $\pm$ 2.78	46.08 $\pm$ 2.70	5.66 $\pm$ 1.74	42.93 $\pm$ 2.88	7.64 $\pm$ 2.91	59.74 $\pm$ 1.97	5.08 $\pm$ 1.17	3.40 $\pm$ 0.22
FedAvg-Adv	25.67 $\pm$ 4.49	9.77 $\pm$ 3.17	32.05 $\pm$ 4.51	8.92 $\pm$ 2.01	25.41 $\pm$ 6.55	9.74 $\pm$ 3.29	32.70 $\pm$ 5.20	7.92 $\pm$ 4.61	3.95 $\pm$ 0.20
FedCIL	38.95 $\pm$ 4.53	8.87 $\pm$ 2.31	39.47 $\pm$ 3.24	6.54 $\pm$ 1.72	45.75 $\pm$ 5.26	7.51 $\pm$ 1.33	56.59 $\pm$ 4.75	8.13 $\pm$ 2.40	3.14 $\pm$ 0.31
MFCL	52.66 $\pm$ 3.24	6.76 $\pm$ 1.24	37.09 $\pm$ 3.76	5.26 $\pm$ 1.87	47.82 $\pm$ 3.09	8.11 $\pm$ 1.07	54.81 $\pm$ 2.89	8.24 $\pm$ 1.84	2.74 $\pm$ 0.24
AF-FCL	41.22 $\pm$ 4.51	6.58 $\pm$ 1.44	43.27 $\pm$ 4.77	5.87 $\pm$ 2.35	54.91 $\pm$ 3.72	7.84 $\pm$ 2.02	67.44 $\pm$ 2.59	<u>2.92</u> $\pm$ 1.04	3.84 $\pm$ 0.19
C ² Prompt	68.47 $\pm$ 3.81	7.42 $\pm$ 1.80	<u>67.93</u> $\pm$ 5.19	6.94 $\pm$ 2.77	<u>72.41</u> $\pm$ 4.53	6.72 $\pm$ 1.84	<u>72.48</u> $\pm$ 3.11	4.28 $\pm$ 1.95	–
AFL	<u>75.17</u> $\pm$ 1.31	2.94 $\pm$ 0.79	55.33 $\pm$ 1.55	3.47 $\pm$ 0.85	49.44 $\pm$ 1.35	3.95 $\pm$ 1.59	68.48 $\pm$ 1.49	5.91 $\pm$ 0.76	<u>2.62</u> $\pm$ 0.11
GFedCL	77.23 $\pm$ 2.76	<u>3.34</u> $\pm$ 1.21	73.67 $\pm$ 3.42	<u>4.24</u> $\pm$ 1.45	77.39 $\pm$ 3.02	<u>5.57</u> $\pm$ 1.88	75.32 $\pm$ 2.19	3.93 $\pm$ 1.25	2.41 $\pm$ 0.18

Visualization of Groups in US-States

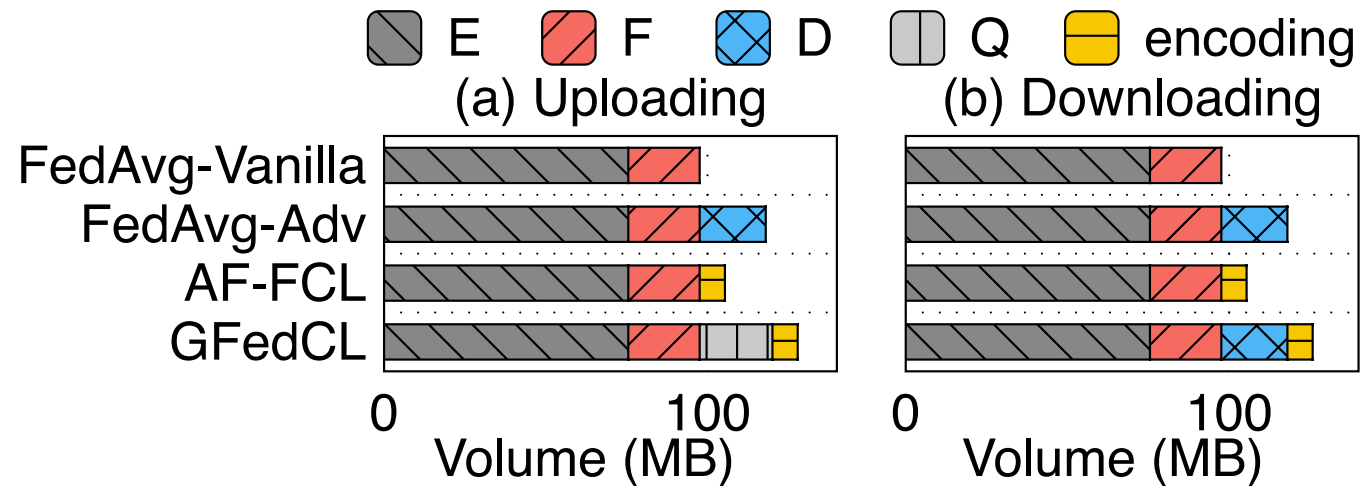


Analysis: Replay Budget

Replay budget influences model performance

Ratio	Model	EMNIST-Letters		CIFAR-100		TinyImageNet		TCGA-BRCA		US-States	
		Acc. (%) $\uparrow$	Time (s) $\downarrow$	Acc. (%) $\uparrow$	Time (s) $\downarrow$	Acc. (%) $\uparrow$	Time (s) $\downarrow$	Acc. (%) $\uparrow$	Time (s) $\downarrow$	MSE ($\times 10^{-2}$) $\downarrow$	Time (s) $\downarrow$
100%	GLFC	43.47 $\pm$ 2.14	9.33	46.08 $\pm$ 2.70	14.61	42.93 $\pm$ 2.88	31.25	59.74 $\pm$ 1.97	11.45	3.40 $\pm$ 0.22	4.94
	FedAvg-Adv	25.67 $\pm$ 4.49	7.33	32.05 $\pm$ 4.51	12.11	25.41 $\pm$ 6.55	27.25	32.70 $\pm$ 5.20	9.89	3.95 $\pm$ 0.20	4.88
	FedCIL	38.95 $\pm$ 4.53	9.67	39.47 $\pm$ 3.24	14.61	45.75 $\pm$ 5.26	33.25	56.59 $\pm$ 4.75	10.07	3.14 $\pm$ 0.31	5.88
	MFCL	52.66 $\pm$ 3.24	12.17	37.09 $\pm$ 3.76	20.53	47.82 $\pm$ 3.09	40.70	54.81 $\pm$ 2.89	15.77	2.74 $\pm$ 0.24	5.10
	AF-FCL	41.22 $\pm$ 4.51	9.17	43.27 $\pm$ 4.77	16.22	54.91 $\pm$ 3.72	35.14	67.44 $\pm$ 2.59	13.25	3.84 $\pm$ 0.19	6.12
	GFedCL	77.23 $\pm$ 2.76	11.41	73.67 $\pm$ 3.42	17.34	77.39 $\pm$ 3.02	36.11	75.32 $\pm$ 2.19	12.37	2.41 $\pm$ 0.18	5.78
50%	GLFC	37.19 $\pm$ 2.51	8.11	38.90 $\pm$ 3.01	12.93	36.60 $\pm$ 3.15	28.45	54.33 $\pm$ 2.25	9.94	3.67 $\pm$ 0.28	4.41
	FedAvg-Adv	22.13 $\pm$ 4.88	6.52	28.34 $\pm$ 4.90	10.94	21.78 $\pm$ 6.90	23.66	29.80 $\pm$ 5.40	8.60	4.24 $\pm$ 0.25	4.36
	FedCIL	33.40 $\pm$ 4.95	8.35	33.20 $\pm$ 3.62	12.73	40.84 $\pm$ 5.70	30.10	52.01 $\pm$ 5.01	8.76	3.45 $\pm$ 0.35	5.23
	MFCL	46.38 $\pm$ 3.66	10.28	32.33 $\pm$ 4.10	17.92	42.18 $\pm$ 3.44	36.98	49.66 $\pm$ 3.20	13.72	2.98 $\pm$ 0.28	4.79
	AF-FCL	37.35 $\pm$ 4.90	8.43	37.38 $\pm$ 5.15	14.35	49.68 $\pm$ 4.11	32.91	61.20 $\pm$ 2.90	11.53	4.10 $\pm$ 0.25	5.48
	GFedCL	69.27 $\pm$ 2.85	10.02	66.34 $\pm$ 3.61	15.86	71.16 $\pm$ 3.20	33.22	69.80 $\pm$ 2.45	10.76	2.66 $\pm$ 0.21	5.13
0%	GLFC	29.84 $\pm$ 2.80	6.89	30.85 $\pm$ 3.44	11.03	28.56 $\pm$ 3.50	24.33	43.12 $\pm$ 2.60	8.46	4.07 $\pm$ 0.35	3.87
	FedAvg-Adv	17.97 $\pm$ 5.20	5.89	22.52 $\pm$ 5.33	9.82	17.61 $\pm$ 7.40	20.51	20.73 $\pm$ 5.80	7.32	4.74 $\pm$ 0.33	3.91
	FedCIL	26.91 $\pm$ 5.30	7.34	26.24 $\pm$ 4.11	11.12	34.54 $\pm$ 6.10	26.45	41.22 $\pm$ 5.40	7.45	3.89 $\pm$ 0.40	4.87
	MFCL	40.16 $\pm$ 4.01	9.46	26.90 $\pm$ 4.55	15.53	35.43 $\pm$ 3.88	31.90	40.01 $\pm$ 3.55	11.67	3.24 $\pm$ 0.32	4.34
	AF-FCL	33.16 $\pm$ 5.25	7.10	32.84 $\pm$ 5.60	12.18	44.52 $\pm$ 4.45	30.25	54.33 $\pm$ 3.22	9.80	4.41 $\pm$ 0.30	4.85
	GFedCL	59.82 $\pm$ 3.12	8.94	41.46 $\pm$ 3.98	14.22	63.42 $\pm$ 3.44	31.87	61.50 $\pm$ 2.80	9.15	2.43 $\pm$ 0.22	4.49

Analysis: Communication Overhead



Limitation and Discussions

Limitations

- No quantitative way to support the priority of using graphs as conditioning
- Graph is not general enough

However, it shows that

- FCL has many **auxiliary conditioning** beyond labels
- The **information structure** in CL matters



THANK YOU

Stevens Institute of Technology
1 Castle Point Terrace, Hoboken, NJ 07030