Private Set Generation with Discriminative Information

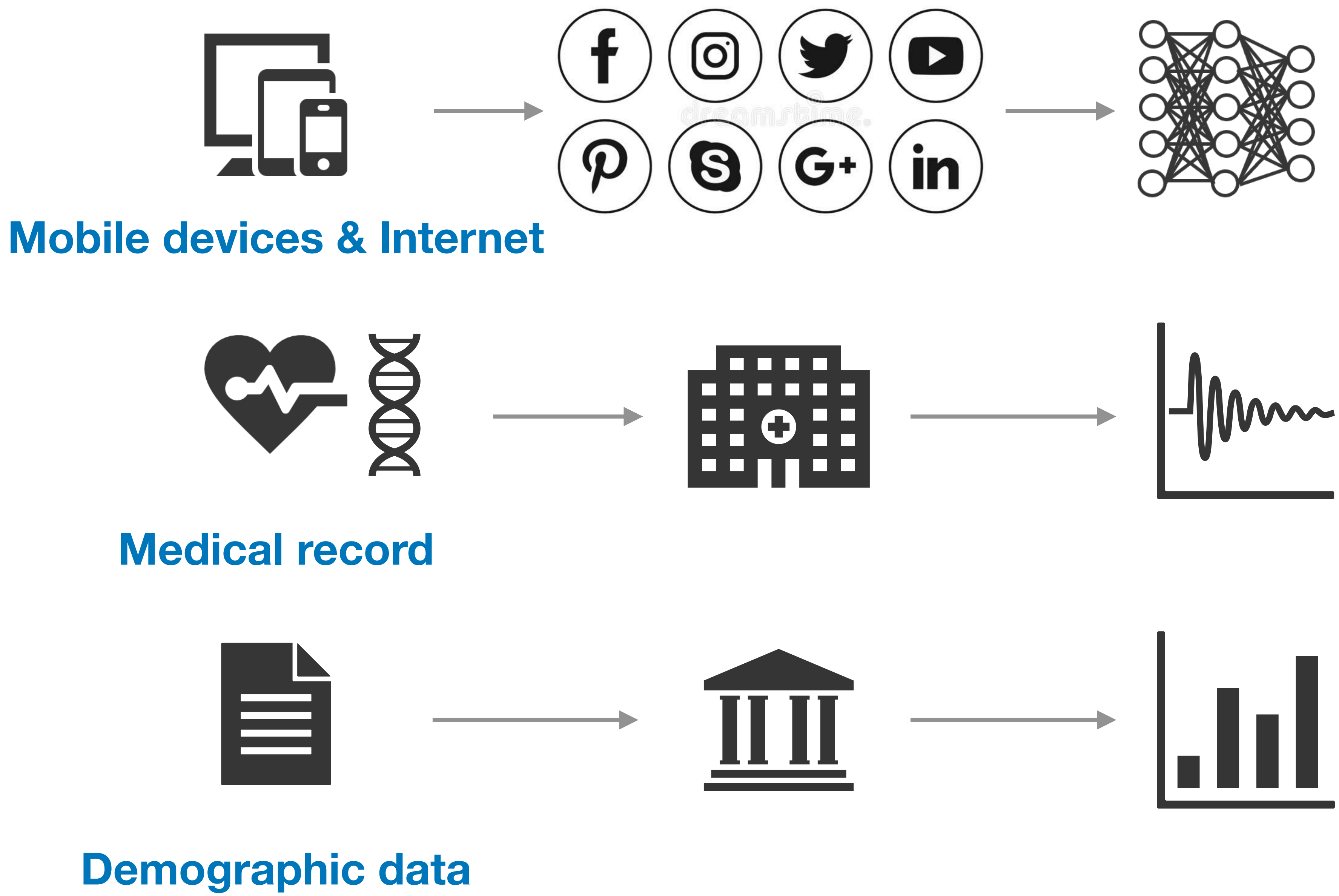
Dingfan Chen   Raouf Kerkouche   Mario Fritz

CISPA Helmholtz Center for Information Security

Data Privacy in ML:

NEURAL INFORMATION PROCESSING SYSTEMS

- Sensitive data is **ubiquitous**

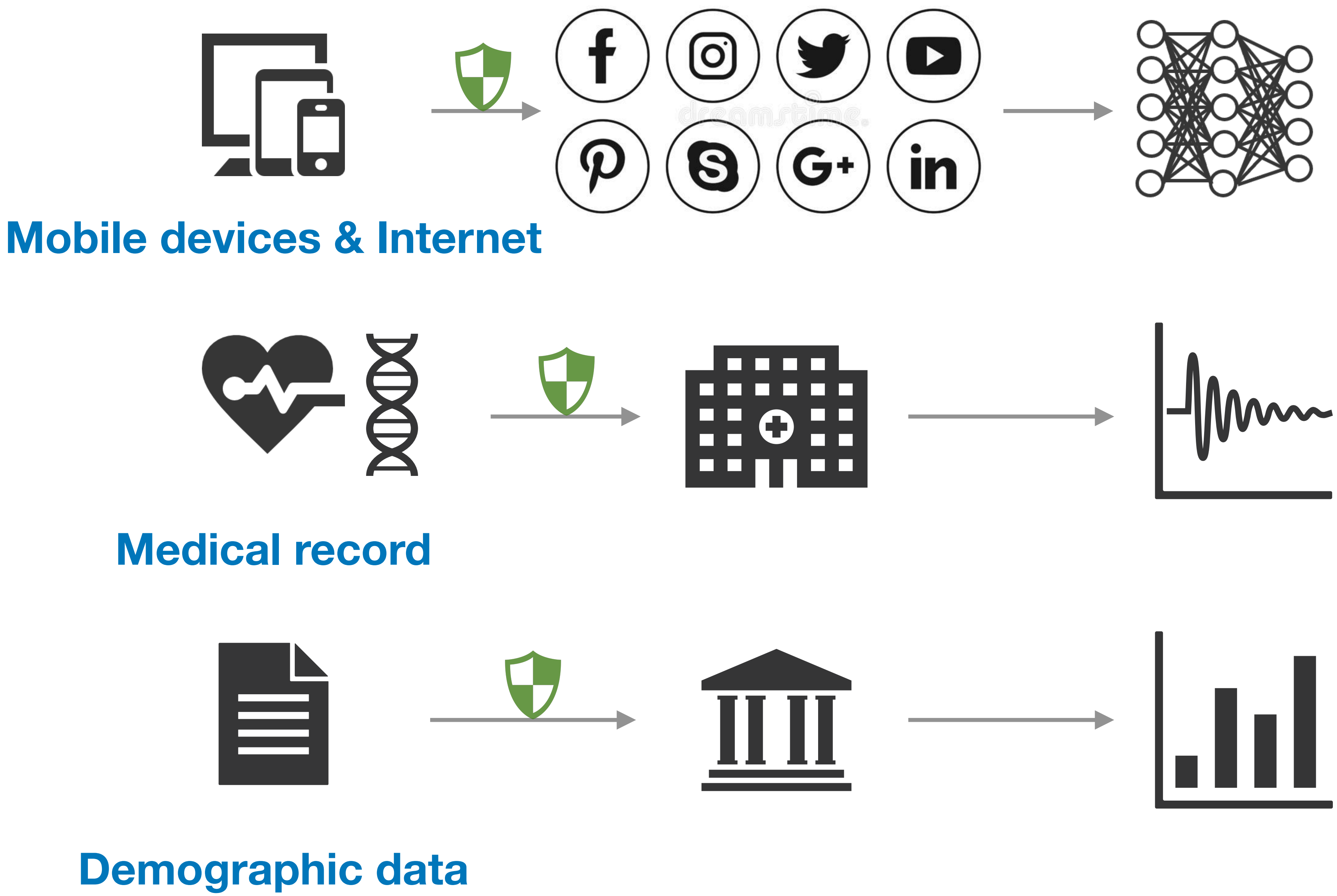


Data Privacy in ML:

NEURAL INFORMATION PROCESSING SYSTEMS

- Sensitive data is **ubiquitous**

Our task:
Data sanitization



Problem

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)

- **Existing approaches**

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)

- **Existing approaches**

[1] Jordon, James, et al., "PATE-GAN: Generating synthetic data with differential privacy guarantees.", *ICLR,* 2018.
[2] Torkzadehmahani, Reihaneh et al., "Dp-cgan: Differentially private synthetic data and label generation.", *CVPR Workshops,* 2019.
[3] Chen, Dingfan, et al.,"Gs-wgan: A gradient-sanitized approach for learning differentially private generators.", *NeurIPS*, 2020.
[4] Harder, Frederik, et al., "Dp-merf: Differentially private mean embeddings with randomfeatures for practical privacy-preserving data generation.", *AISTAT*, 2021.
[5] Rakotomamonjy, Alain, et al., "Differentially private sliced wasserstein distance.", *ICML,* 2021.
[6] Long, Yunhui, et al., "G-PATE: Scalable Differentially Private Data Generator via Private Aggregation of Teacher Discriminators.", *NeurIPS,* 2021.
[7] Cao, Tianshi, et al., "Don't Generate Me: Training Differentially Private Generative Models with Sinkhorn Divergence.", *NeurIPS,* 2021.
[8] Wang, Boxin et al., "Datalens: Scalable privacy preserving training via gradient compression and aggregation.", *CCS,* 2021.
[9] Lee, Jaewoo, et al., "Differentially Private Normalizing Flows for Synthetic Tabular Data Generation.", *AAAI*, 2022.
[10] Liew, Seng Pei, et al., "PEARL: Data Synthesis via Private Embeddings and Adversarial Reconstruction Learning.", *ICLR*, 2022.

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)

- **Existing approaches**
  - Aim at fitting the complete data distribution

- **Our approach**

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)
- **Existing approaches**
  - Aim at fitting the complete data distribution
- **Our approach**
  - Target at common downstream tasks (e.g., classification)

Generally easier

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)

- **Existing approaches**
  - Aim at fitting the complete data distribution
  - Optimize deep generative models

- **Our approach**
  - Target at common downstream tasks (e.g., classification) **Generally easier**

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)

- **Existing approaches**
  - Aim at fitting the complete data distribution
  - Optimize deep generative models

- **Our approach**
  - Target at common downstream tasks (e.g., classification)
  - Directly optimize a set of representative samples

Generally easier

Better convergence

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)
- **Existing approaches**
  - Aim at fitting the complete data distribution
  - Optimize deep generative models
  - Suboptimal utility: <85% for MNIST with $(\varepsilon, \delta)=(10,10^{-5})$
- **Our approach**
  - Target at common downstream tasks (e.g., classification)
  - Directly optimize a set of representative samples

Generally easier

Better convergence

Problem

NEURAL INFORMATION PROCESSING SYSTEMS

- **Privacy-preserving data generation**
  - High-dimensional data → Deep Neural Network (NN)
  - Rigorous privacy guarantee → Differential Privacy (DP)
- **Existing approaches**
  - Aim at fitting the complete data distribution
  - Optimize deep generative models
  - Suboptimal utility: <85% for MNIST with $(\varepsilon, \delta)=(10, 10^{-5})$
- **Our approach**
  - Target at common downstream tasks (e.g., classification)
  - Directly optimize a set of representative samples
  - ~10% downstream test accuracy improvement over SOTA

Generally easier

Better convergence

Useful samples

Approach

NEURAL INFORMATION PROCESSING SYSTEMS

- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)

[1] Zhao, Bo, et al., “Dataset condensation with gradient matching.”, *ICLR,* 2021.
[2] Zhao, Bo, et al., “Dataset condensation with differentiable siamese augmentation.”, *ICML*, 2021

Approach

- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)

fake $\mathcal{S}$

real $\boldsymbol{x}, y \sim P_{\mathcal{D}}$

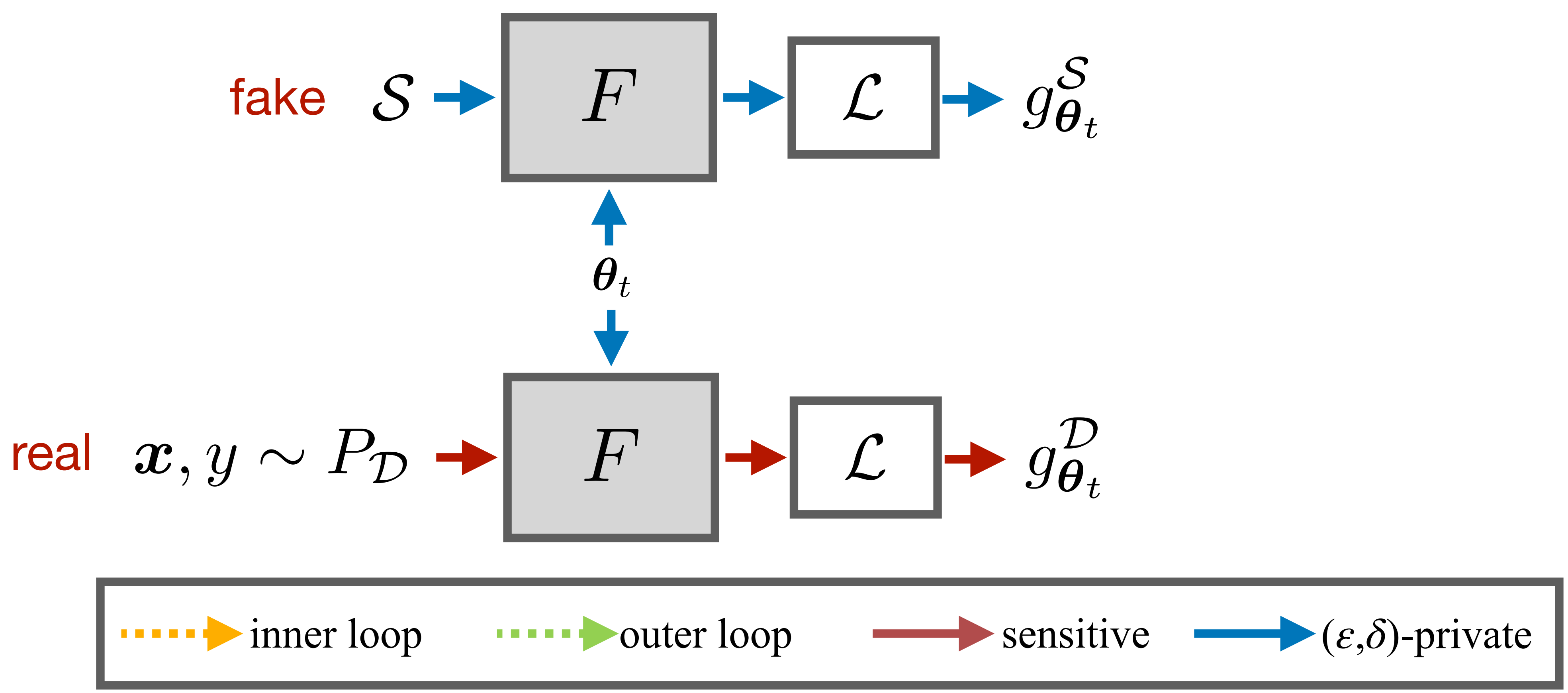
inner loop outer loop sensitive (ε,δ)-private

[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

- **Target:**
  - Optimize for training downstream Neural Network classifier

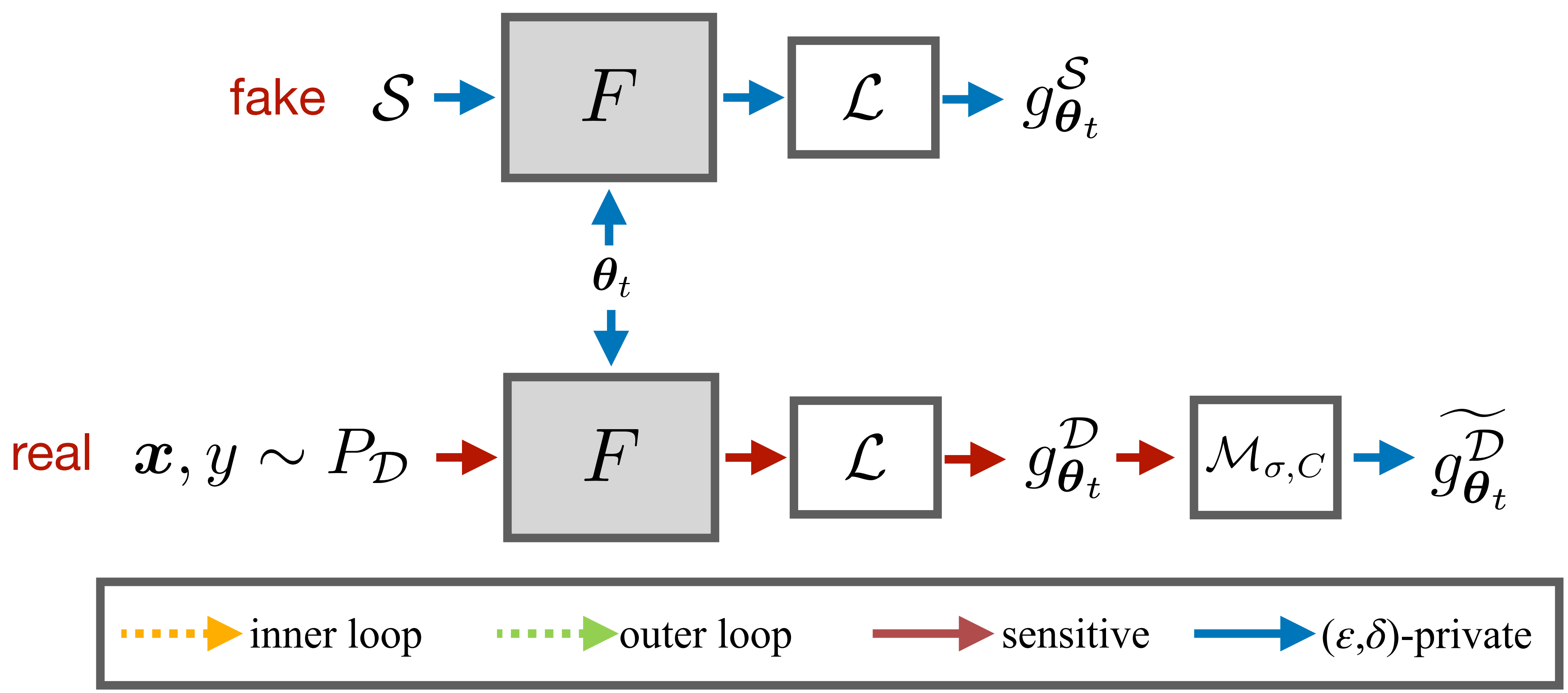
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)



[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)

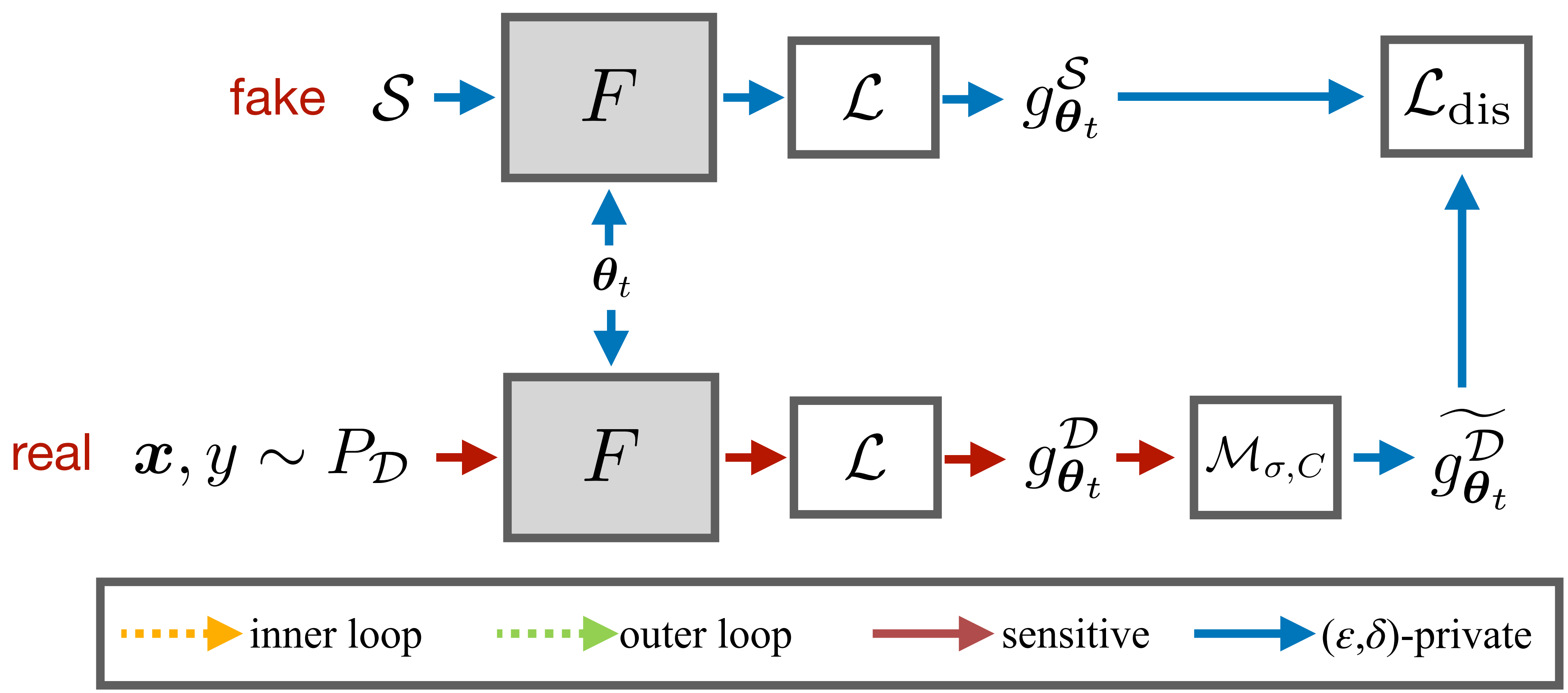


[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

- **Target:**
  - Optimize for training downstream Neural Network classifier

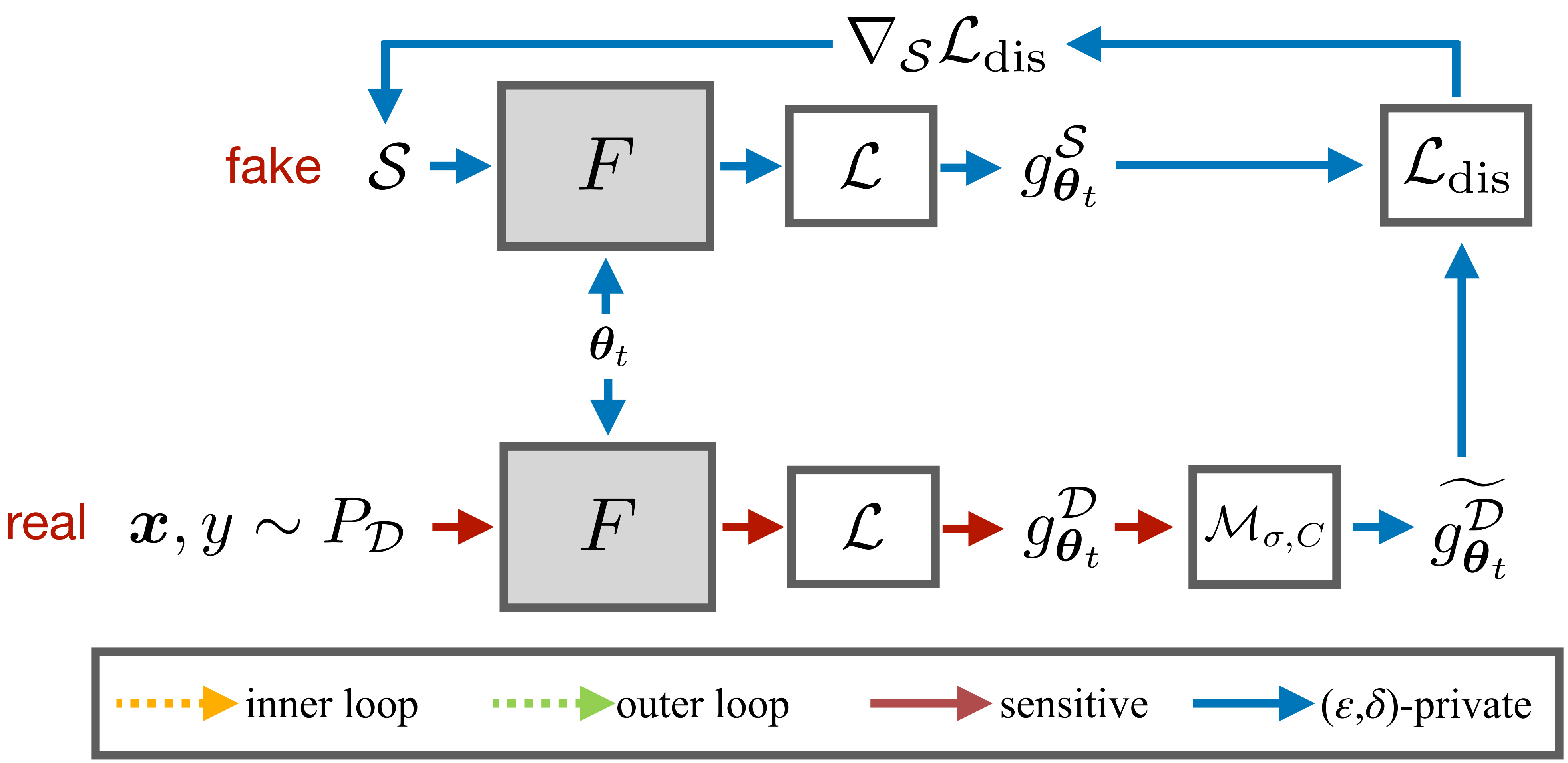
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)



[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

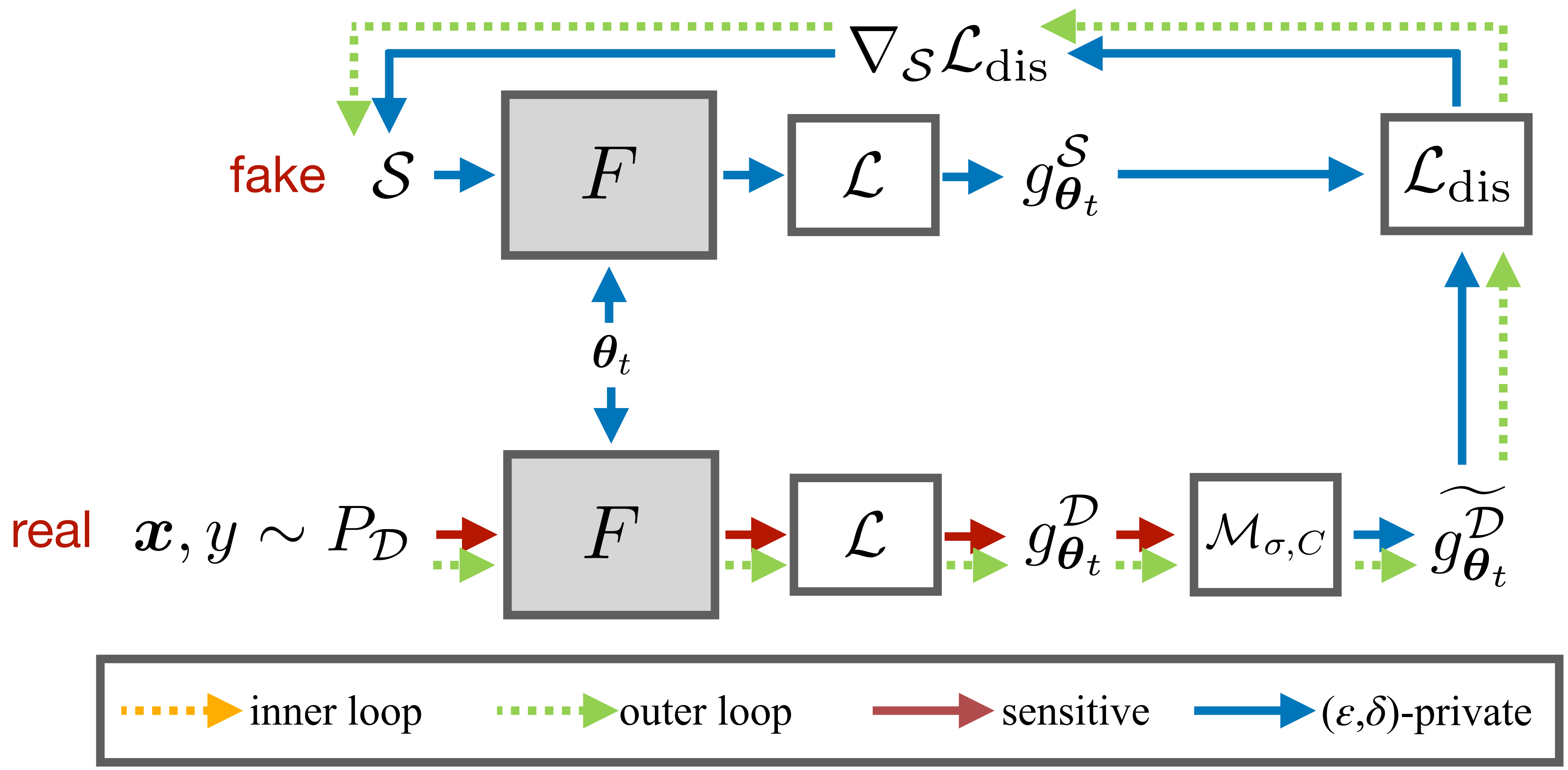
- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)



[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)

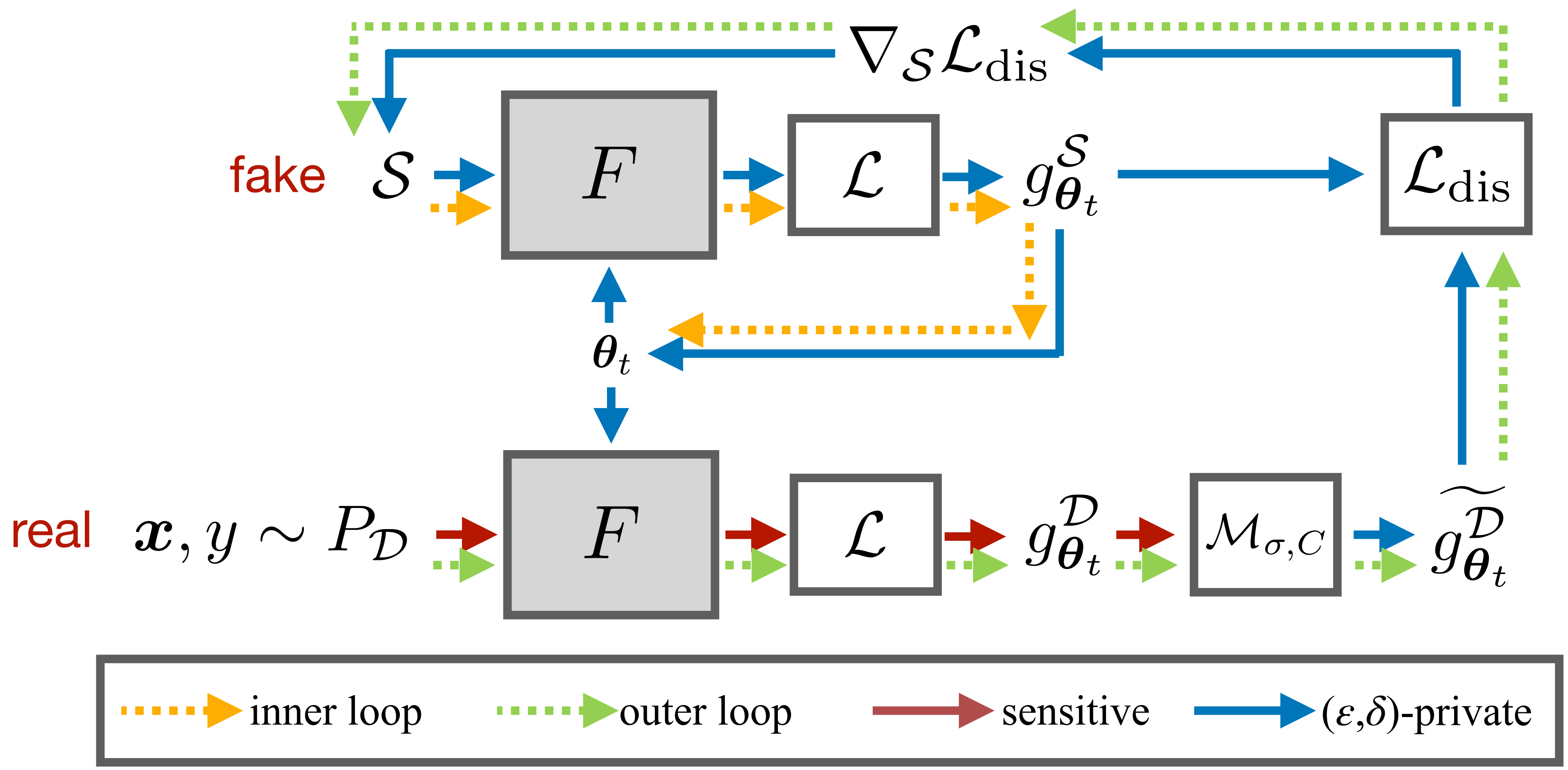


[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Approach

NEURAL INFORMATION PROCESSING SYSTEMS

- **Target:**
  - Optimize for training downstream Neural Network classifier
- **Basic idea:**
  - Gradient-based **coreset generation**[1,2]
  - DP stochastic gradient descent (DP-SGD)



[1] Zhao, Bo, et al., "Dataset condensation with gradient matching.", *ICLR,* 2021.
[2] Zhao, Bo, et al., "Dataset condensation with differentiable siamese augmentation.", *ICML*, 2021

Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

- **Comparison to SOTA**
  - Utility for downstream classification task (train on fake; test on real)

Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

- **Comparison to SOTA**
  - Utility for downstream classification task (train on synthetic; test on real)

	MNIST		FashionMNIST	
	ε=1	ε=10	ε=1	ε=10
DP-CGAN	-	52.5	-	50.2
G-PATE	58.8	80.9	58.1	69.3
DataLens	71.2	80.7	64.8	70.6
GS-WGAN	-	84.9	-	63.1
DP-Merf	72.7	85.7	61.2	72.4
DP-Sinkhorn	-	83.2	-	71.1
Ours (spc=20)	**80.9**	**95.6**	**70.2**	**77.7**

Evaluation

- **Comparison to SOTA**
  - Utility for downstream classification task (train on synthetic; test on real)

	MNIST		FashionMNIST	
	ε=1	ε=10	ε=1	ε=10
DP-CGAN	-	52.5	-	50.2
G-PATE	58.8	80.9	58.1	69.3
DataLens	71.2	80.7	64.8	70.6
GS-WGAN	-	84.9	-	63.1
DP-Merf	72.7	85.7	61.2	72.4
DP-Sinkhorn	-	83.2	-	71.1
Ours (spc=20)	**80.9**	**95.6**	**70.2**	**77.7**

Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

- **Comparison to SOTA**
  - Generalization across model architecture

	MNIST						FashionMNIST					
	ConvNet	LeNet	AlexNet	VGG11	ResNet18	MLP	ConvNet	LeNet	AlexNet	VGG11	ResNet18	MLP
Real	99.6	99.2	99.5	99.6	99.7	98.3	93.5	88.9	91.5	93.8	94.5	86.9
DP-CGAN	50.2	52.6	52.1	54.7	51.8	54.3	50.2	52.6	52.1	54.7	51.8	54.3
GS-WGAN	84.9	83.2	80.5	87.9	89.3	74.7	54.7	62.7	55.1	57.3	58.9	65.4
DP-Merf	85.7	87.2	84.4	81.7	81.3	85.0	72.4	67.9	64.9	70.1	66.7	**73.1**
Ours (spc=10)	94.9	91.3	90.3	93.6	**94.3**	86.1	75.6	**68.0**	**66.2**	74.7	**72.1**	62.8
Ours (spc=20)	**95.6**	**93.0**	**92.3**	**94.5**	94.1	**87.1**	**77.7**	**68.0**	59.1	**76.8**	70.8	62.2

Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

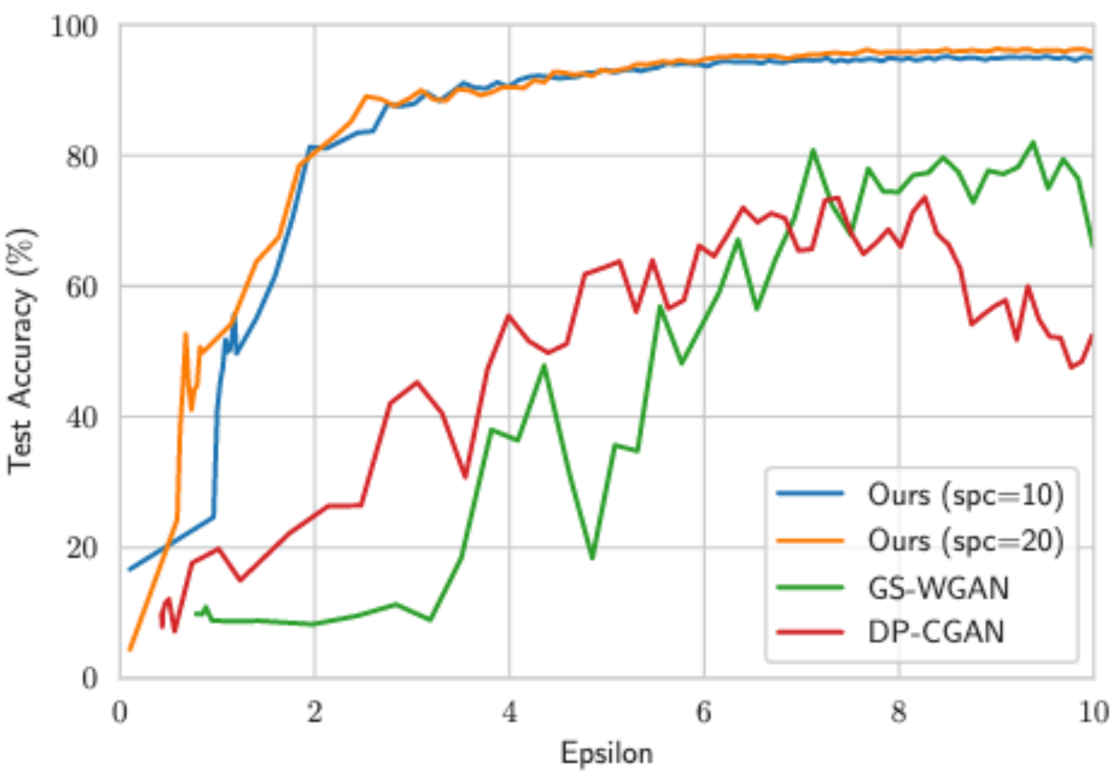
- **Comparison to SOTA**
  - Generalization across model architecture

	MNIST						FashionMNIST					
	ConvNet	LeNet	AlexNet	VGG11	ResNet18	MLP	ConvNet	LeNet	AlexNet	VGG11	ResNet18	MLP
Real	99.6	99.2	99.5	99.6	99.7	98.3	93.5	88.9	91.5	93.8	94.5	86.9
DP-CGAN	50.2	52.6	52.1	54.7	51.8	54.3	50.2	52.6	52.1	54.7	51.8	54.3
GS-WGAN	84.9	83.2	80.5	87.9	89.3	74.7	54.7	62.7	55.1	57.3	58.9	65.4
DP-Merf	85.7	87.2	84.4	81.7	81.3	85.0	72.4	67.9	64.9	70.1	66.7	**73.1**
Ours (spc=10)	94.9	91.3	90.3	93.6	**94.3**	86.1	75.6	**68.0**	**66.2**	74.7	**72.1**	62.8
Ours (spc=20)	**95.6**	**93.0**	**92.3**	**94.5**	94.1	**87.1**	**77.7**	**68.0**	59.1	**76.8**	70.8	62.2

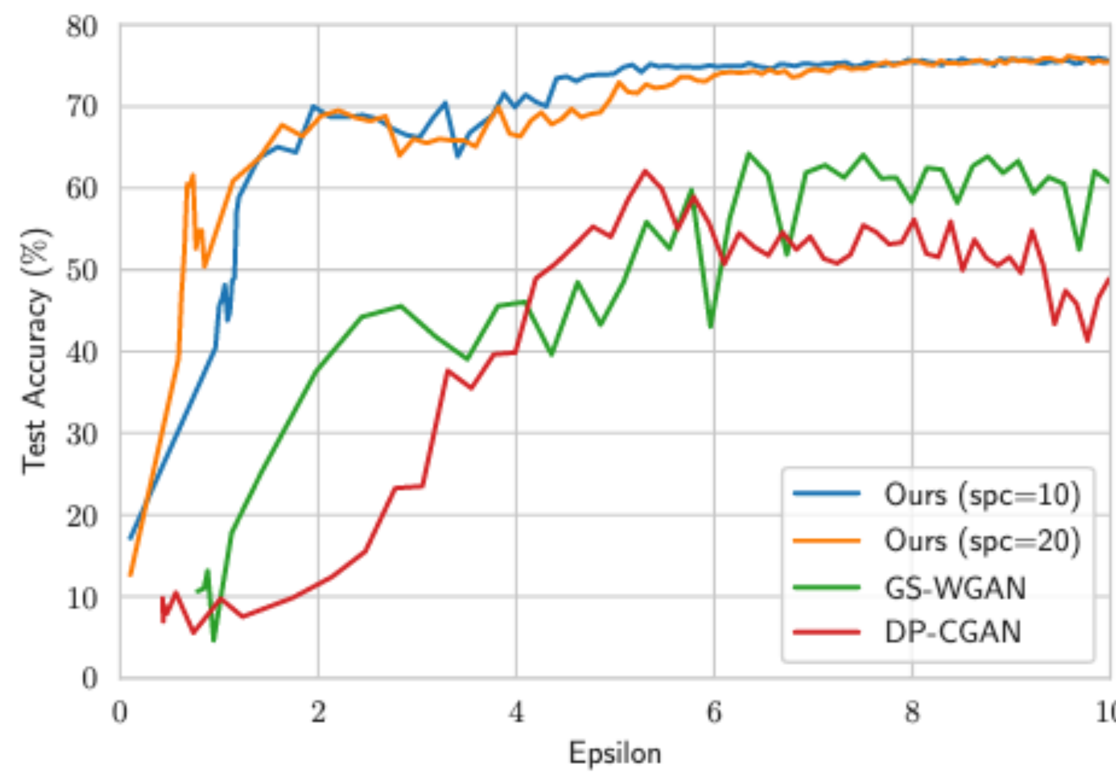
Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

- **Comparison to SOTA**
  - Convergence rate



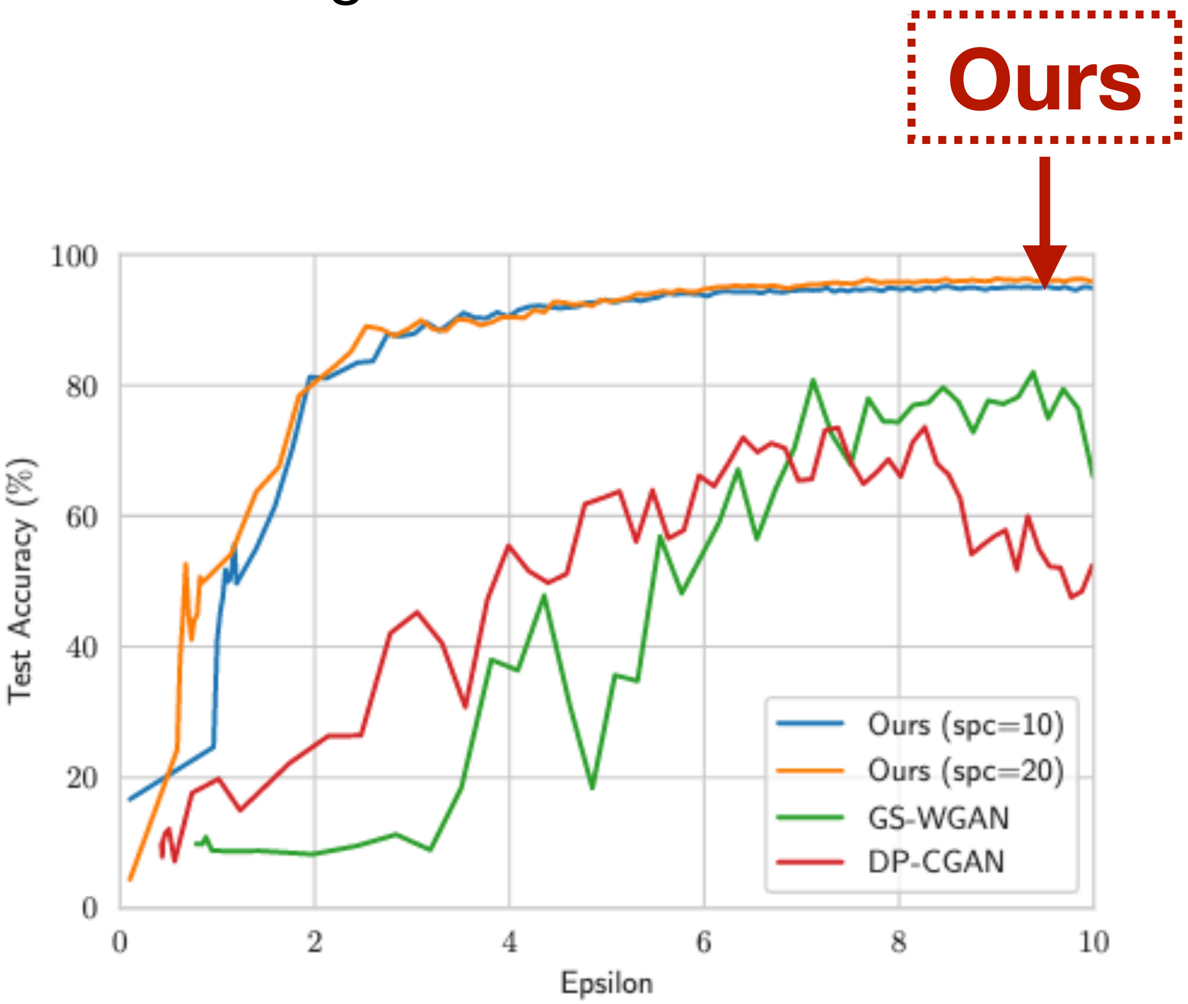
(a) MNIST



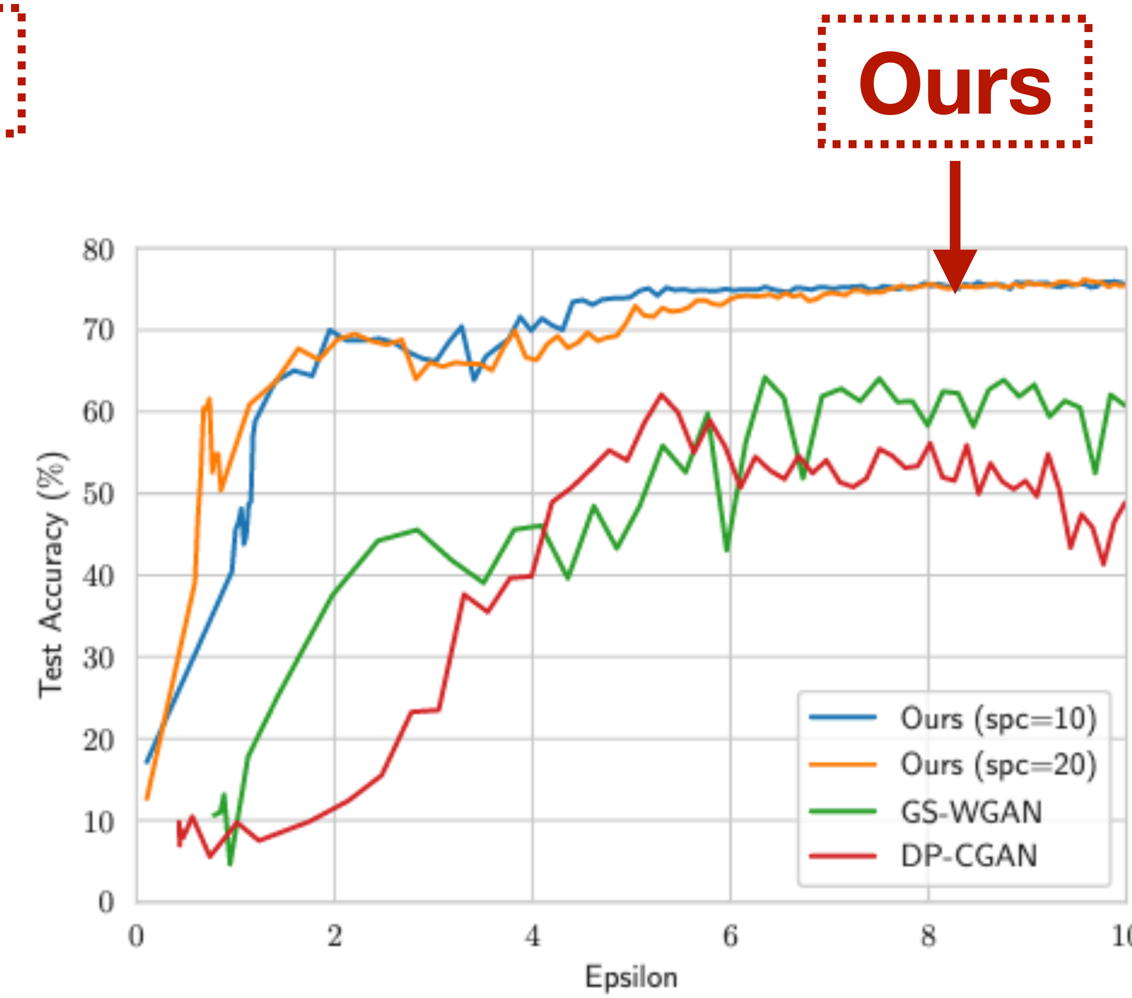
(b) FashionMNIST

Evaluation

- **Comparison to SOTA**
  - Convergence rate



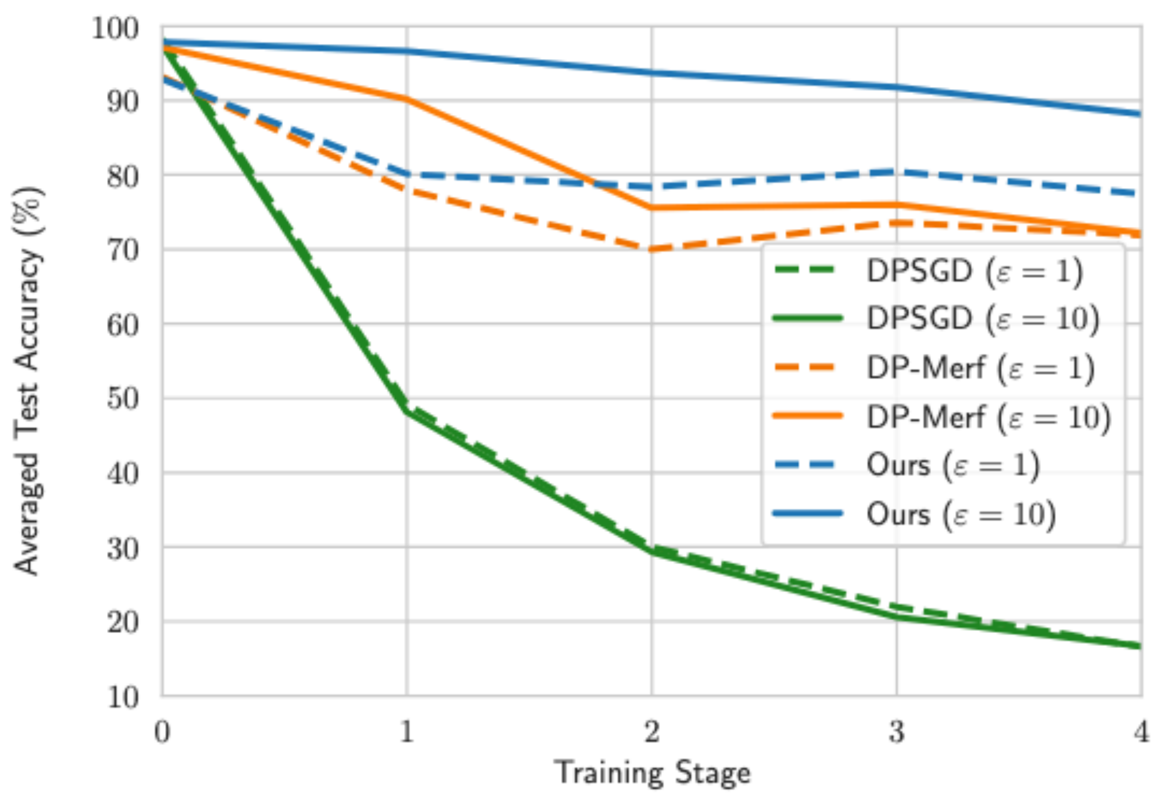
(a) MNIST



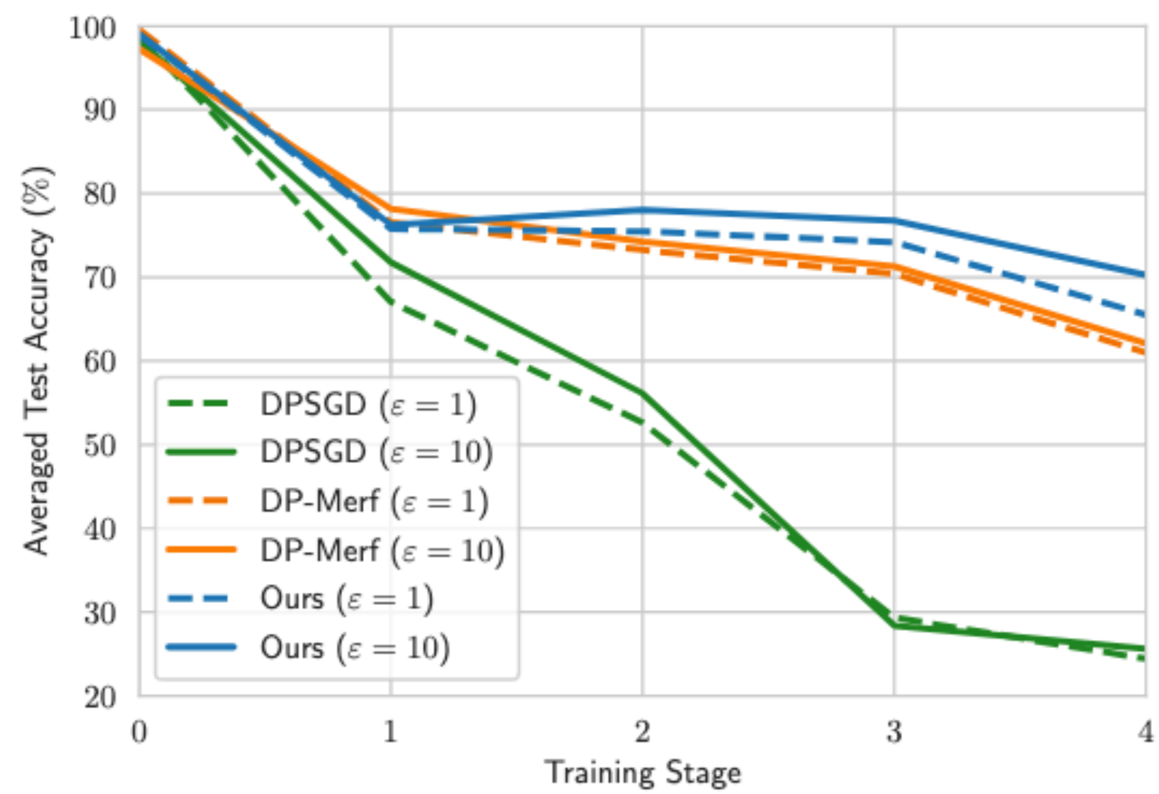
(b) FashionMNIST

Evaluation

- **Application:** Continual learning with DP



(a) SplitMNIST

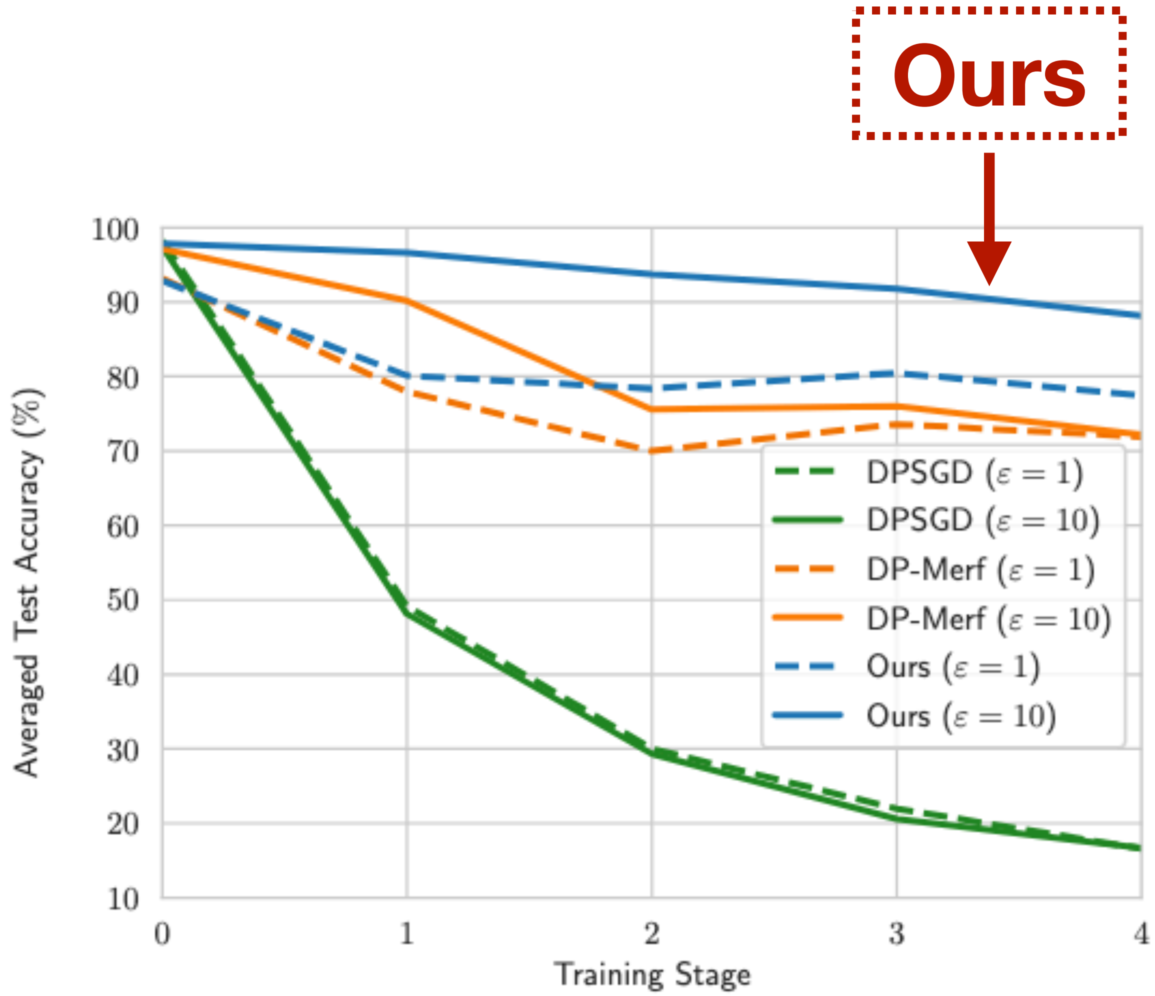


(b) SplitFashionMNIST

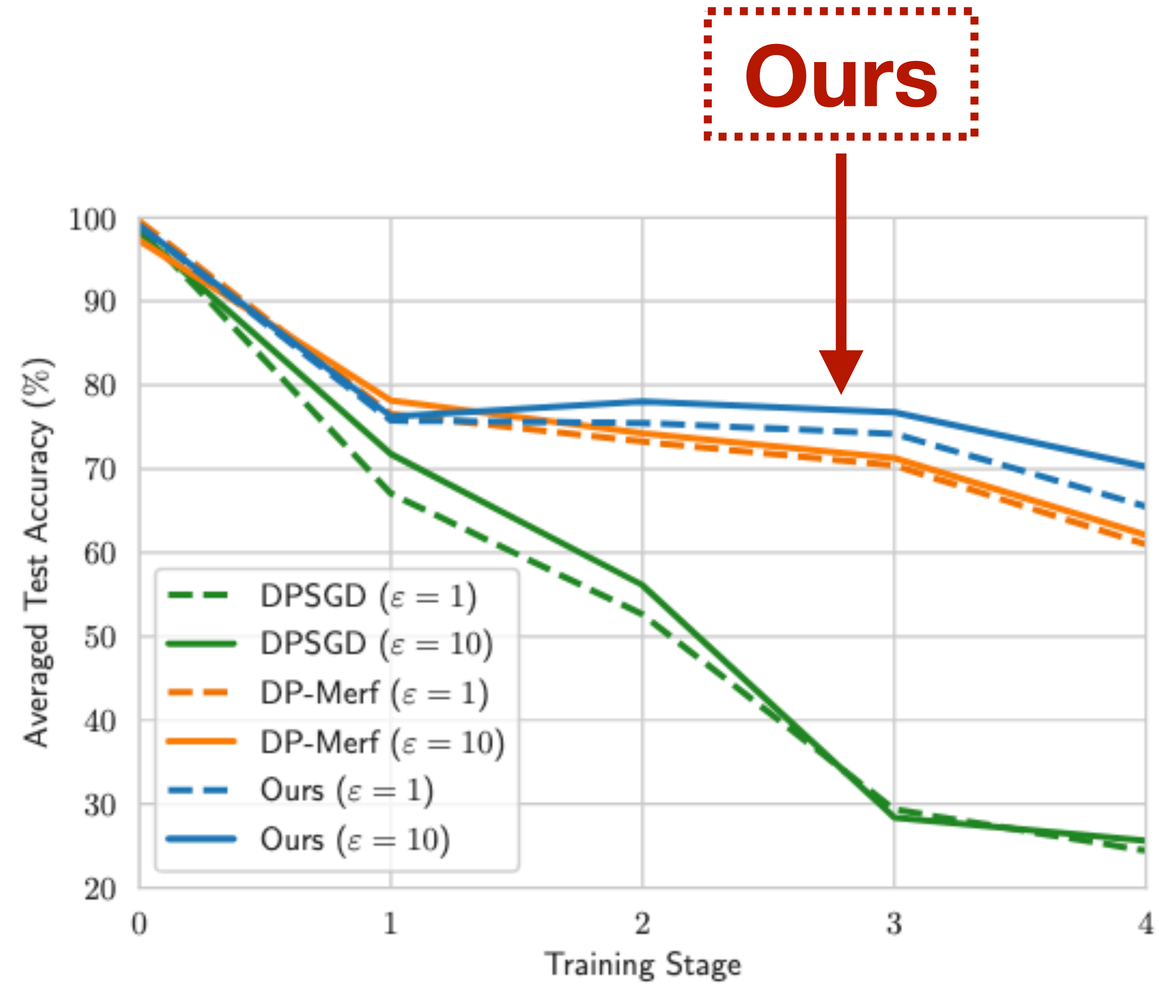
Evaluation

NEURAL INFORMATION PROCESSING SYSTEMS

- **Application:** Continual learning with DP



(a) SplitMNIST



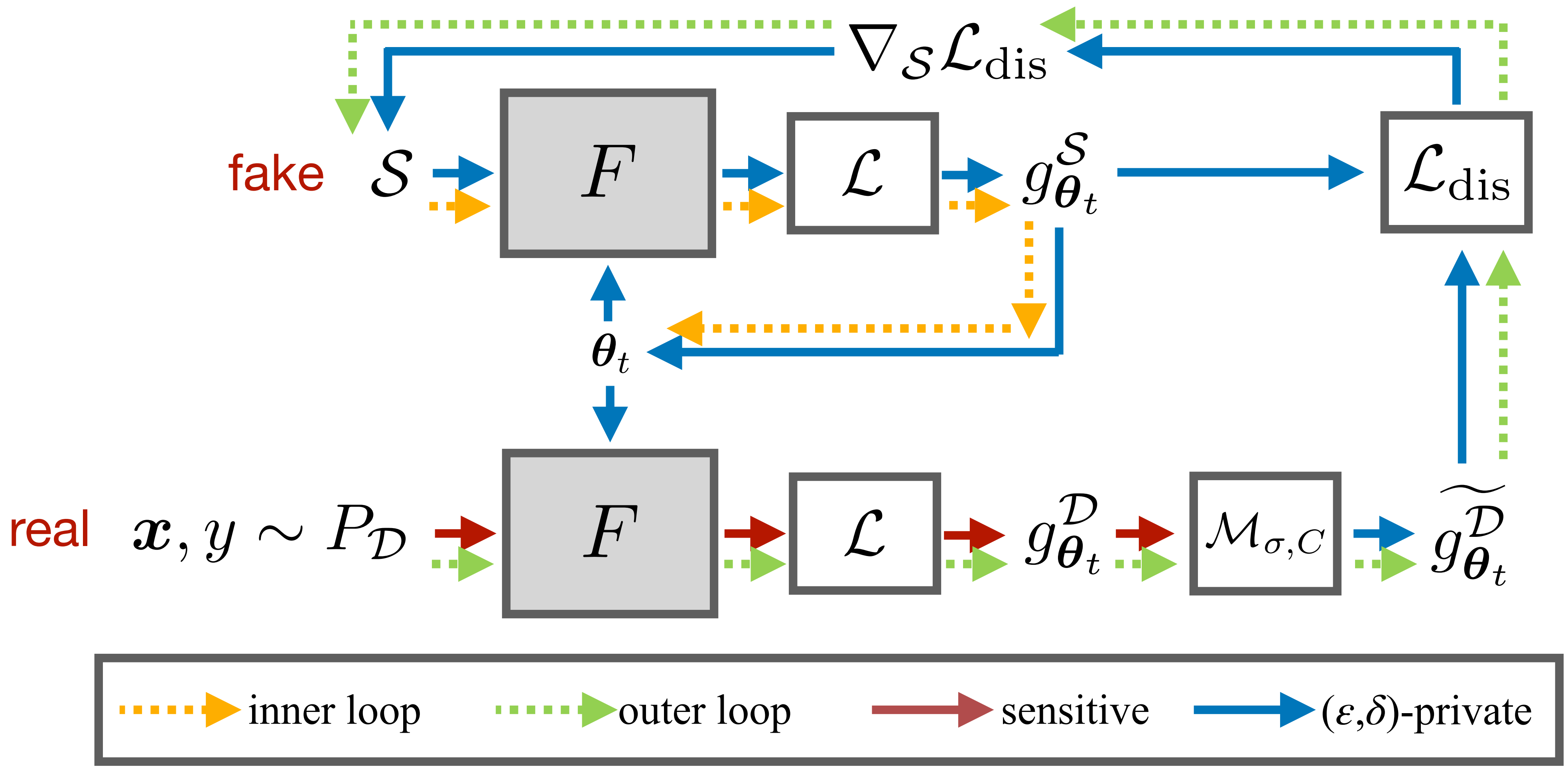
(b) SplitFashionMNIST

Discussion

- **Are deep generative models the best option for this task?**

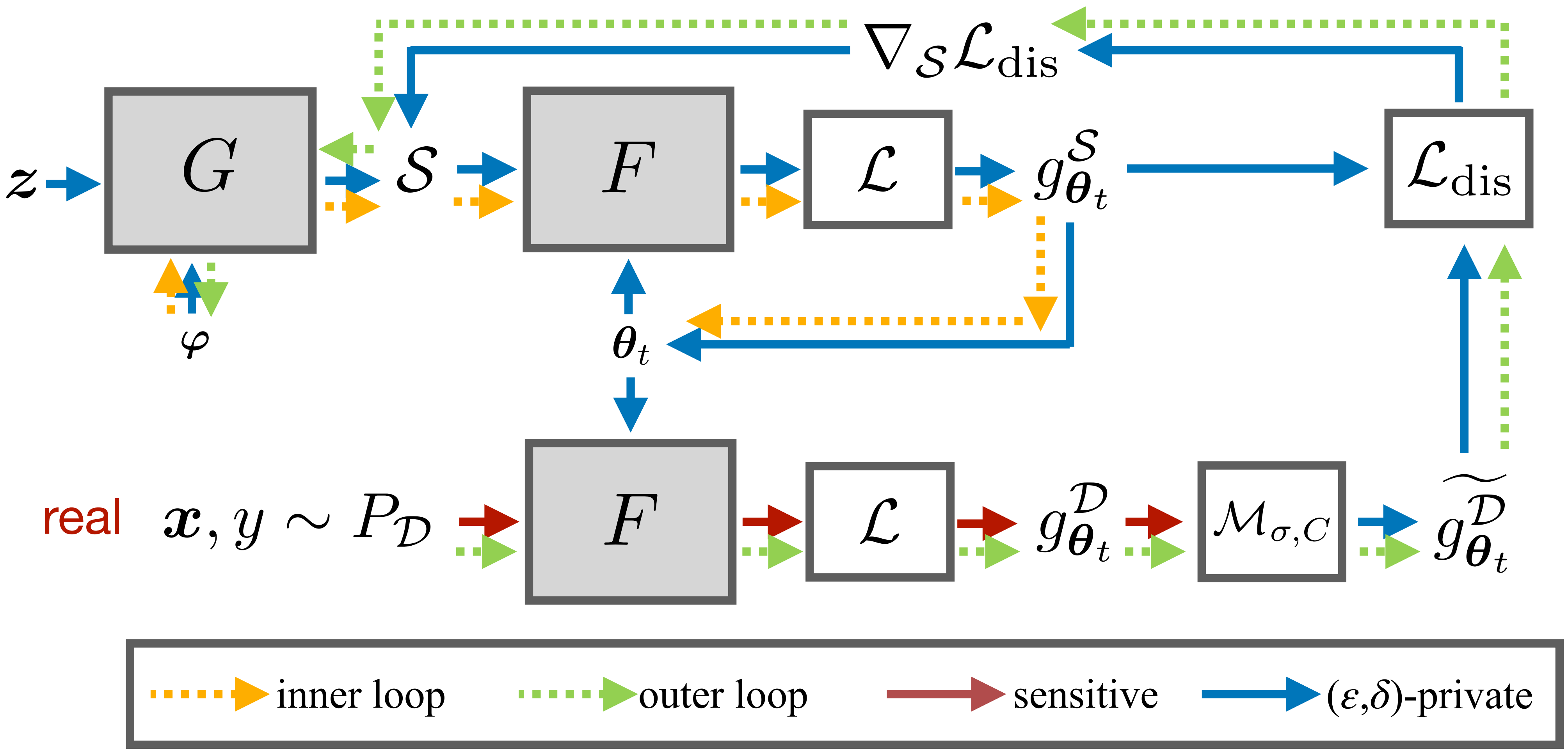
Discussion

- **Are deep generative models the best option for this task?**



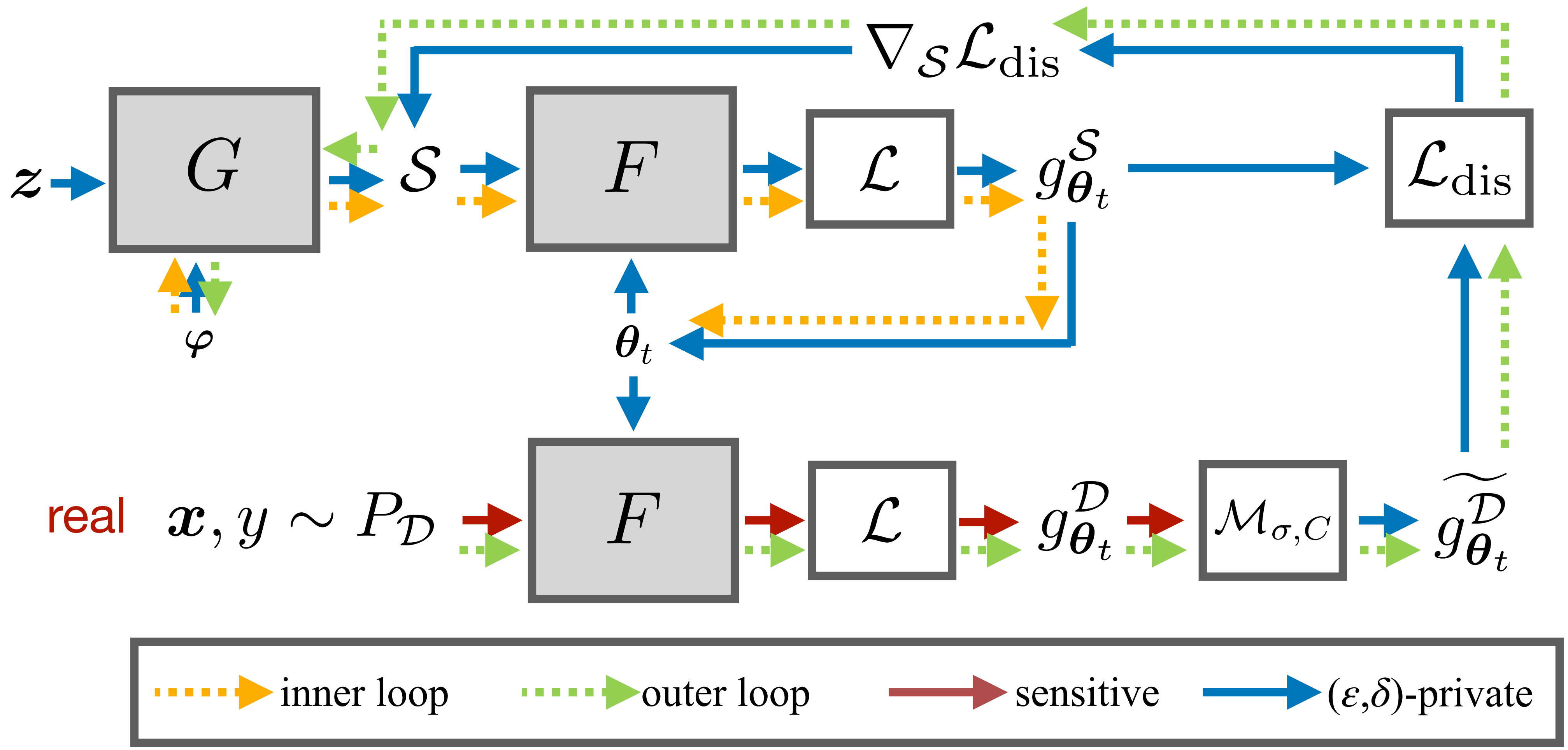
Discussion

- **Are deep generative models the best option for this task?**



Discussion

- **Are deep generative models the best option for this task?**



Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**

Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**

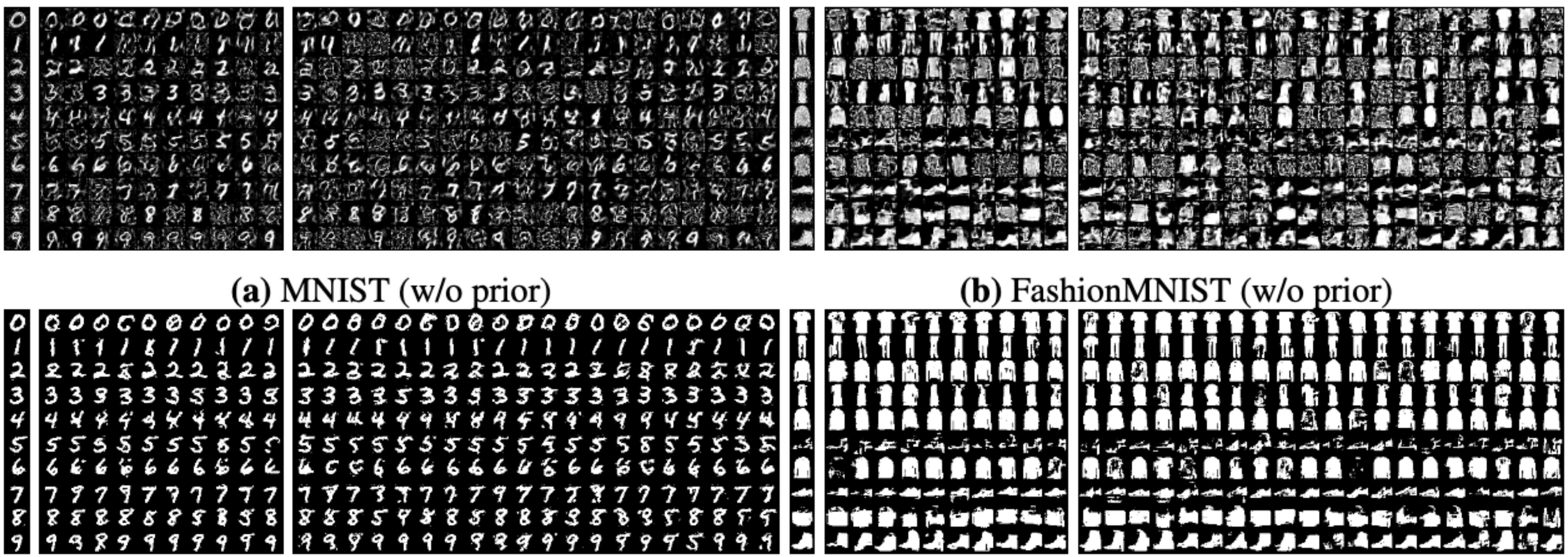
Most probably not!

Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Sub-optimal downstream utility

Most probably not!



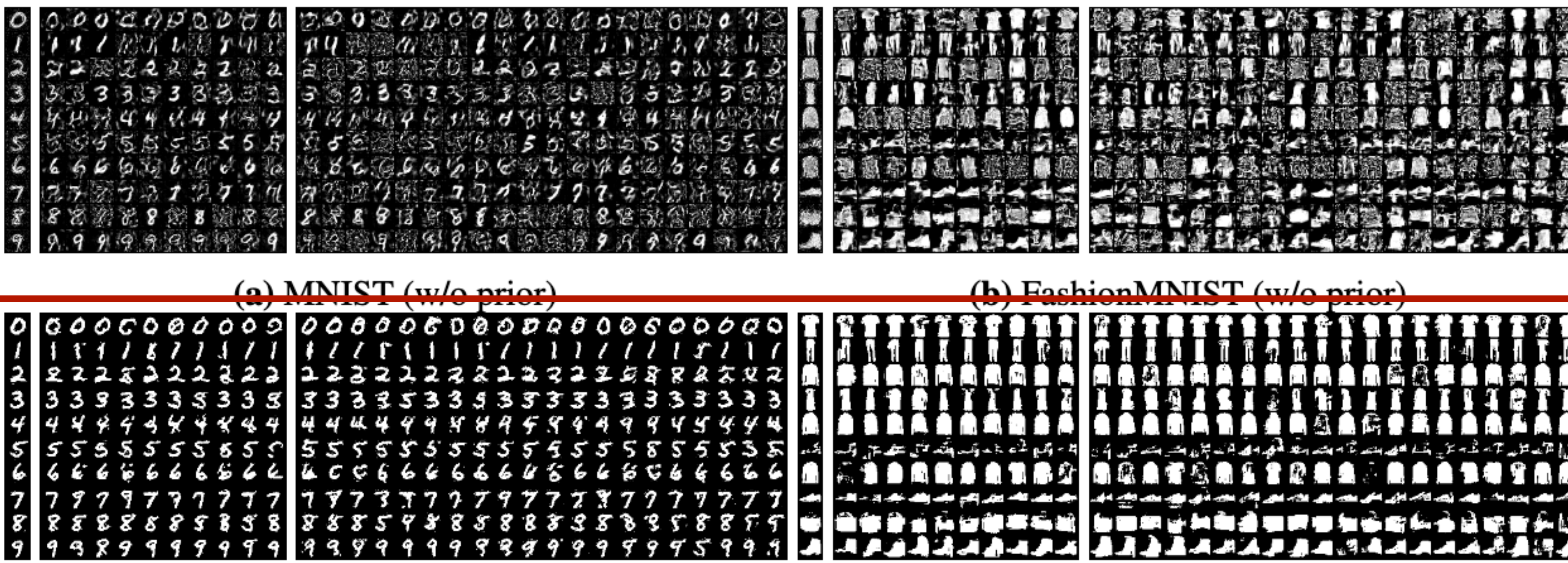
(a) MNIST (w/o prior) (b) FashionMNIST (w/o prior)

(c) MNIST (with prior) (d) FashionMNIST (with prior)

Discussion

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Sub-optimal downstream utility

Most probably not!



(a) MNIST (w/o prior)

(b) FashionMNIST (w/o prior)

(c) MNIST (with prior)

(d) FashionMNIST (with prior)

with generative model

Discussion

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Slow convergence

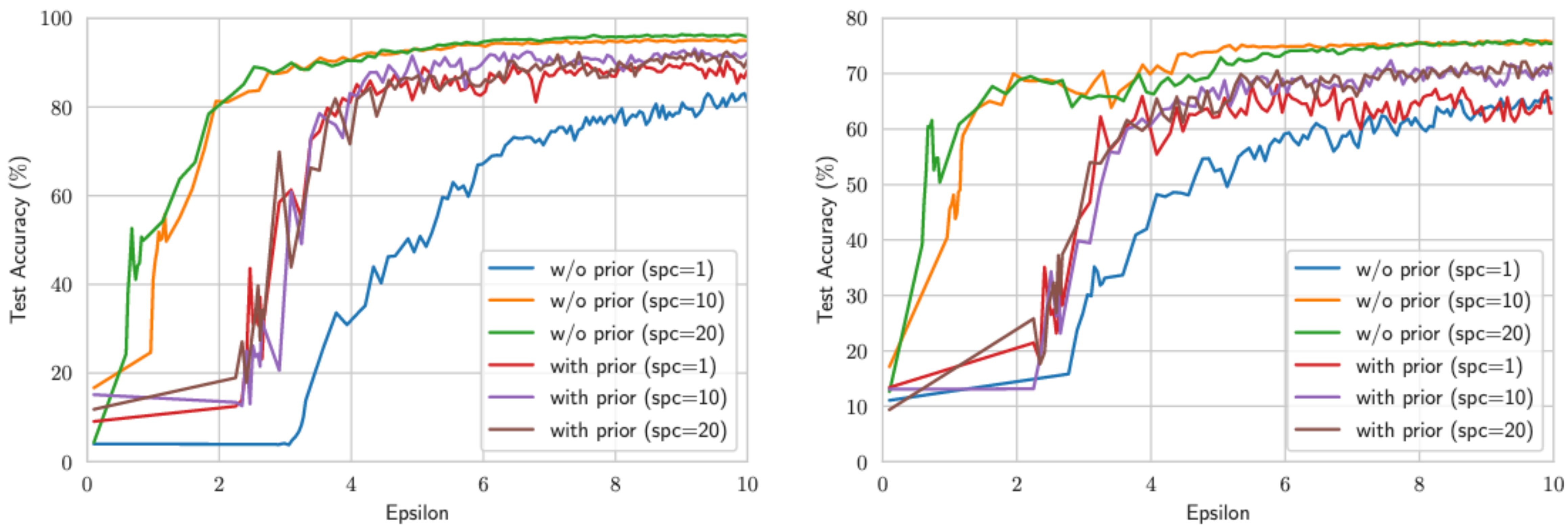
Most probably not!

Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Slow convergence

Most probably not!



(a) MNIST **(b)** FashionMNIST

Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Slow convergence
    - Sub-optimal downstream utility

Most probably not!

	MNIST			FashionMNIST		
	1	10	20	1	10	20
w/o prior	81.4	**94.9**	**95.6**	**66.7**	**75.6**	**77.7**
with prior	**88.2**	92.2	90.6	63.0	70.2	70.7

Discussion

NEURAL INFORMATION PROCESSING SYSTEMS

- **Are deep generative models the best option for this task?**
  - Deep generative models result in:
    - Better visual quality
    - Slow convergence
    - Sub-optimal downstream utility

Most probably not!

	MNIST			FashionMNIST		
	1	10	20	1	10	20
w/o prior	81.4	**94.9**	**95.6**	**66.7**	**75.6**	**77.7**
with prior	**88.2**	92.2	90.6	63.0	70.2	70.7

with generative model

More details in the paper

NEURAL INFORMATION PROCESSING SYSTEMS

Private Set Generation with Discriminative Information

Dingfan Chen Raouf Kerkouche Mario Fritz

Source code available on Github:

https://github.com/DingfanChen/Private-Set

Contact:
Dingfan Chen, dingfan.chen@cispa.de