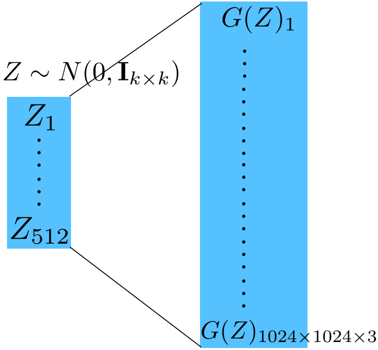
The power of two samples for Generative Adversarial Networks

Sewoong Oh

Department of Industrial and Enterprise Systems Engineering
University of Illinois at Urbana-Champaign

joint work with
Zinan Lin(CMU), Ashish Khetan(UIUC), Giulia Fanti(CMU)

Generative models

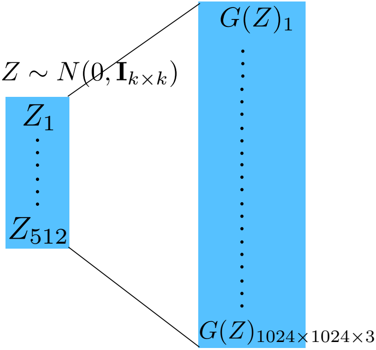


$G(Z) \in \mathbb{R}^{1024 \times 1024 \times 3}$

- A generative model is a black box that takes a random vector $Z \in \mathbb{R}^k$ and produces a sample vector $G(Z) \in \mathbb{R}^n$

[“Progressive Growing of GANs for Improved Quality, Stability, and Variation”, T. Karras, T. Aila, S. Laine, J. Lehtinen 2017]

Generative models

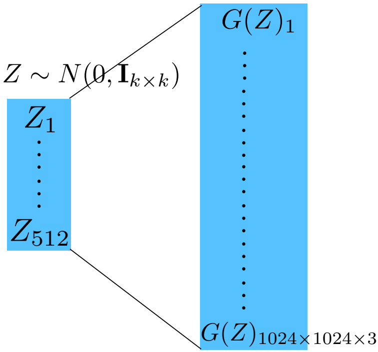


$G(Z) \in \mathbb{R}^{1024 \times 1024 \times 3}$

- A generative model is a black box that takes a random vector $Z \in \mathbb{R}^k$ and produces a sample vector $G(Z) \in \mathbb{R}^n$

[“Progressive Growing of GANs for Improved Quality, Stability, and Variation”, T. Karras, T. Aila, S. Laine, J. Lehtinen 2017]

Generative models

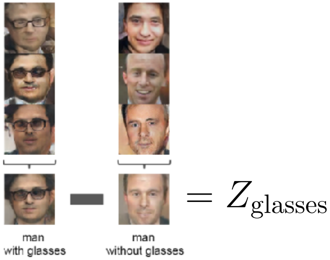


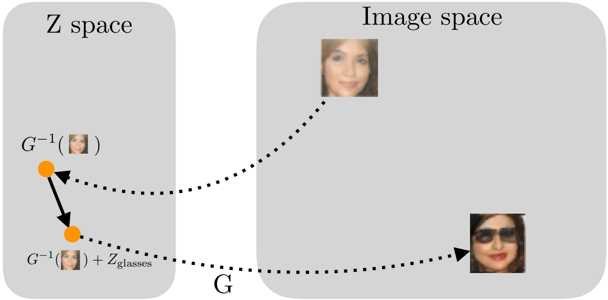
$G(Z) \in \mathbb{R}^{1024\times1024\times3}$

- A generative model is a black box that takes a random vector $Z \in \mathbb{R}^k$ and produces a sample vector $G(Z) \in \mathbb{R}^n$

["Progressive Growing of GANs for Improved Quality, Stability, and Variation", T. Karras, T. Aila, S. Laine, J. Lehtinen 2017]

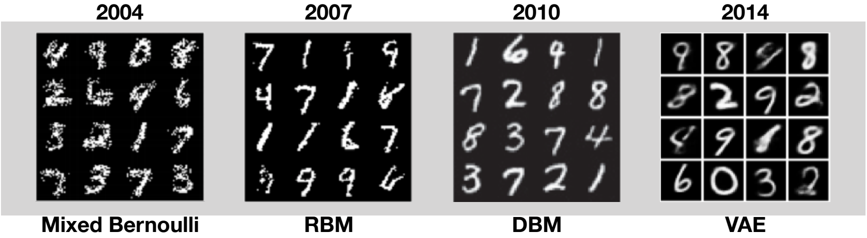
Generative models learn fundamental representations



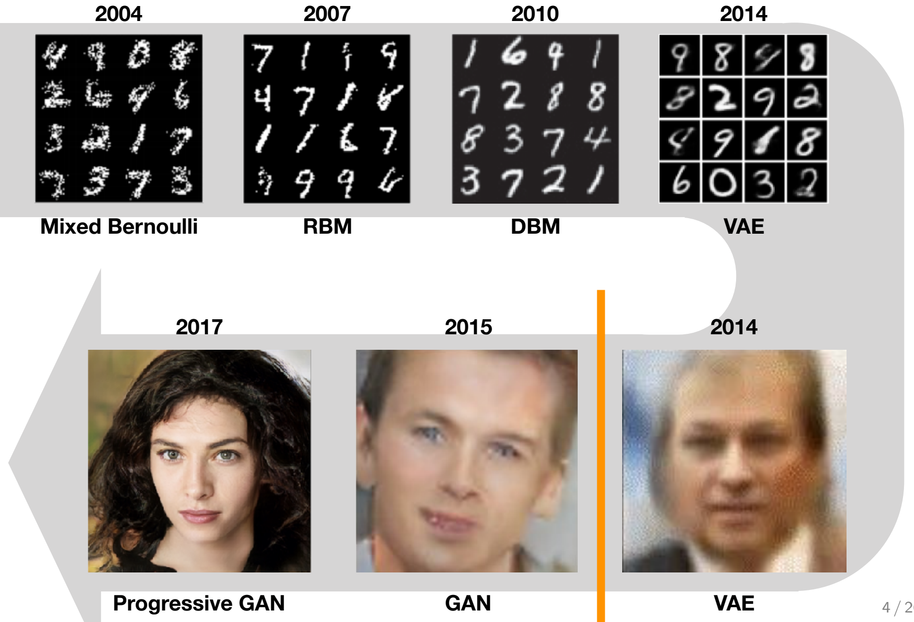


[DCGAN, Radford et al. 2015]

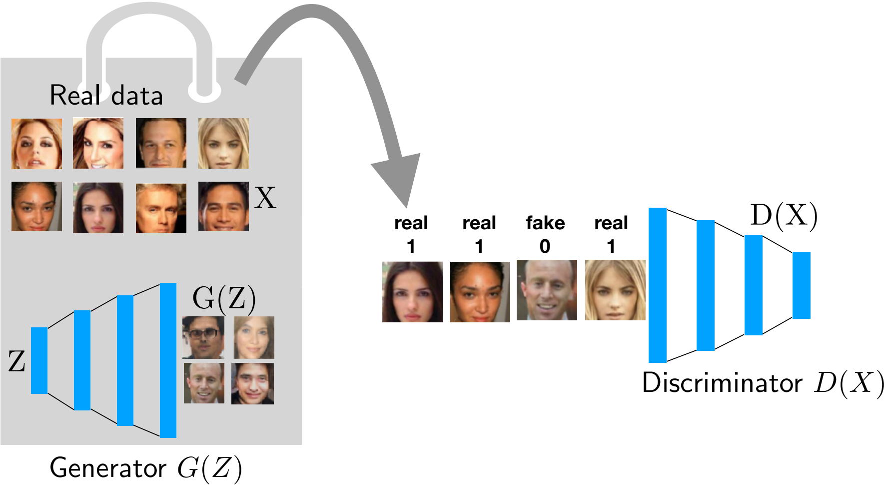
GAN: a breakthrough in training generative models



GAN: a breakthrough in training generative models

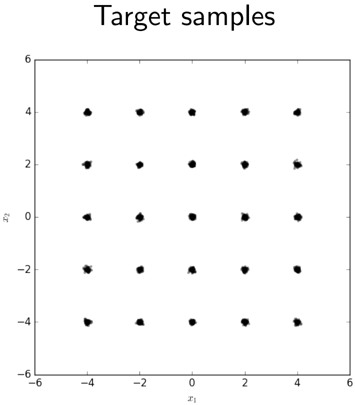


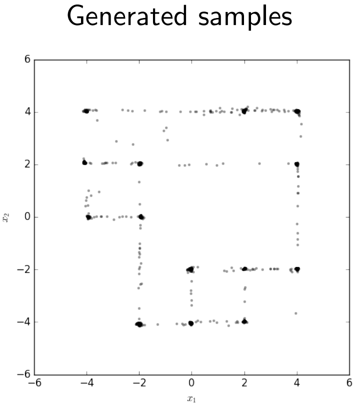
Generative Adversarial Networks (GAN)



$$\min_G \max_D V(G, D)$$

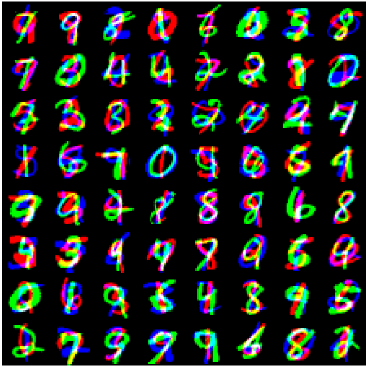
"Mode collapse" is a main challenge



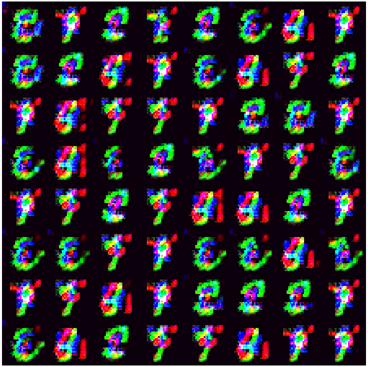


“Mode collapse” is a main challenge

Target samples



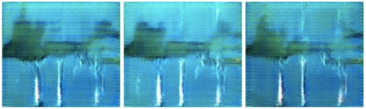
Generated samples



“Mode collapse” is a main challenge

- “A man in a orange jacket with sunglasses and a hat ski down a hill.”

- “This guy is in black trunks and swimming underwater.”



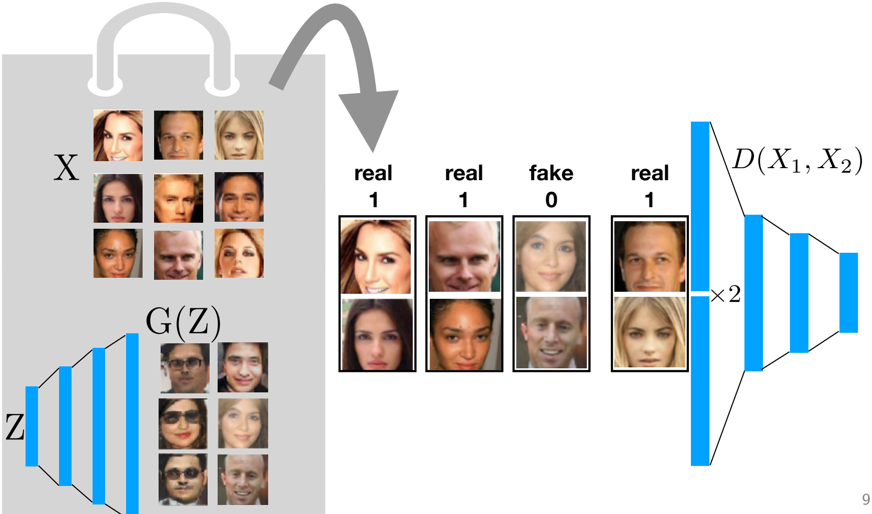
- “A tennis player in a blue polo shirt is looking down at the green court.”

[“Generating interpretable images with controllable structure”, by Reed et al., 2016]

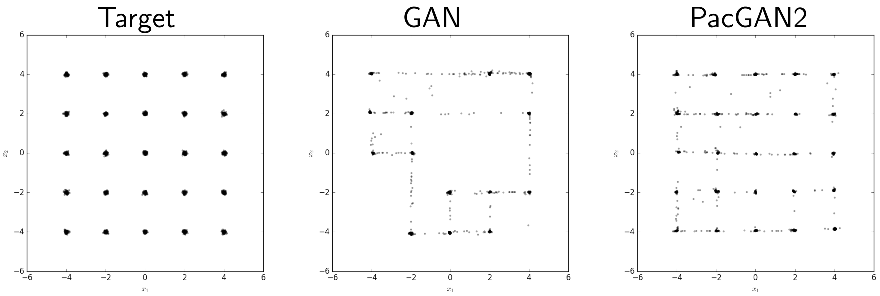
Lack of diversity is easier to detect
if the discriminator sees multiple sample jointly

New framework: PacGAN

- lightweight overhead
- experimental results
- principled



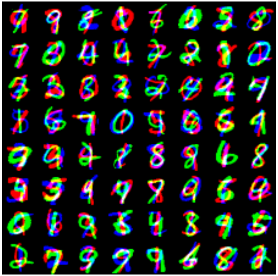
Benchmark tests



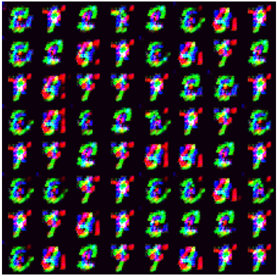
	Modes (Max 25)
GAN	17.3
PacGAN2	23.8
PacGAN3	24.6
PacGAN4	24.8

Benchmark datasets from VeeGAN paper

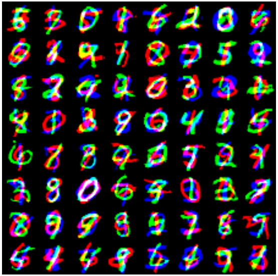
Real data



DCGAN



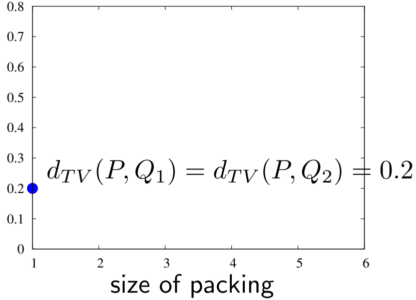
PacDCGAN2



	Modes (Max 1000)
DCGAN	99.0
ALI	16.0
Unrolled GAN	48.7
VeeGAN	150.0
PacDCGAN2	1000.0
PacDCGAN3	1000.0
PacDCGAN4	1000.0

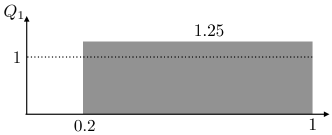
Intuition behind packing via toy example

Target distribution P

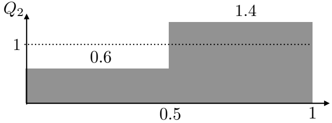


Generator Q_1
with mode collapse

Generator Q_2
without mode collapse

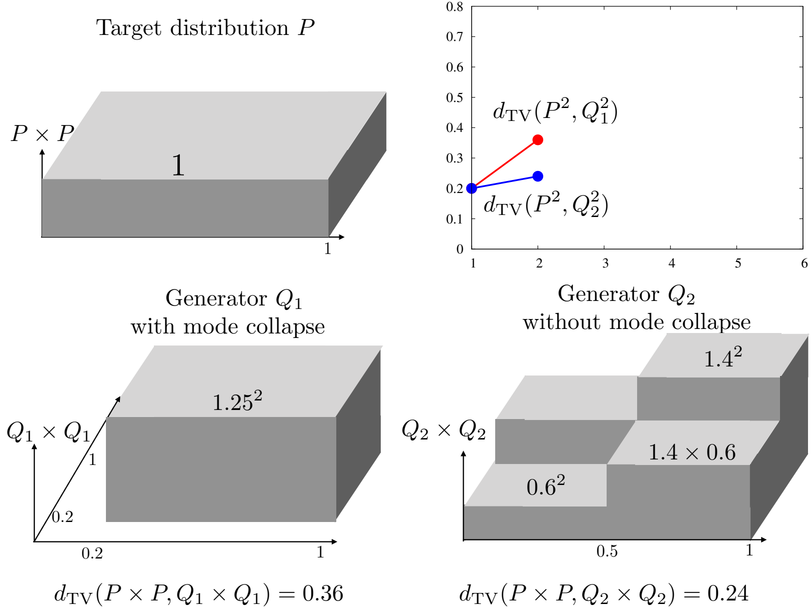


$d_{\mathrm{TV}}(P, Q_1) = 0.2$



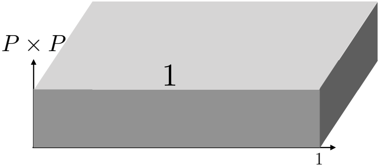
$d_{\mathrm{TV}}(P, Q_2) = 0.2$

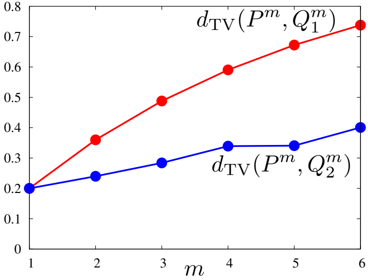
Intuition behind packing via toy example



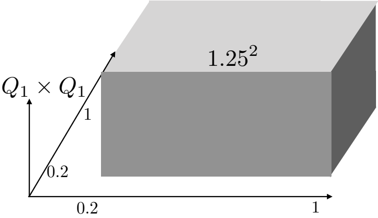
Intuition behind packing via toy example

Target distribution P



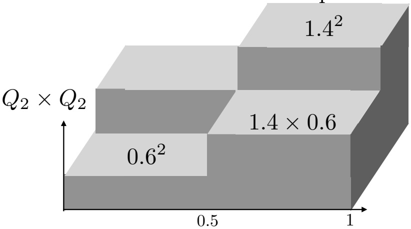


Generator Q_1
with mode collapse



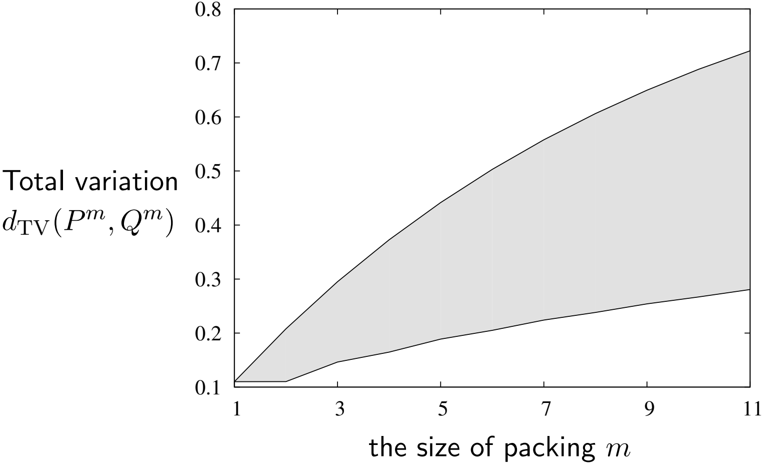
$$d_{\mathrm{TV}}(P \times P, Q_1 \times Q_1) = 0.36$$

Generator Q_2
without mode collapse

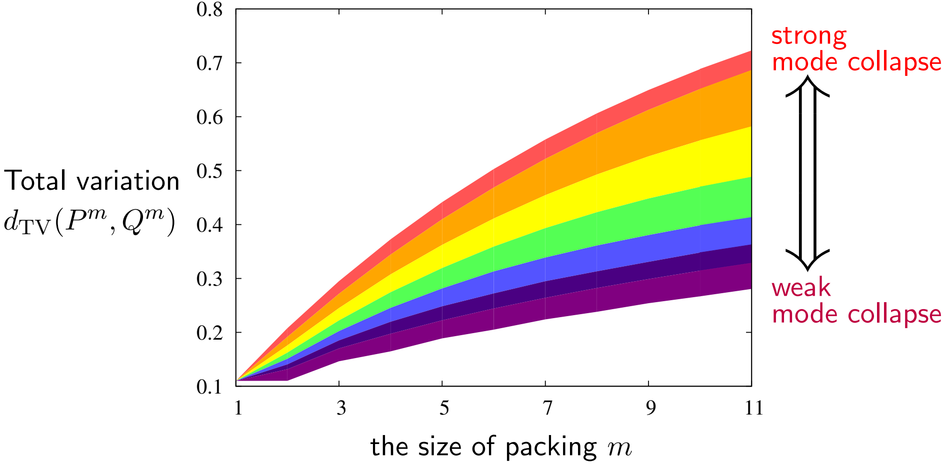


$$d_{\mathrm{TV}}(P \times P, Q_2 \times Q_2) = 0.24$$

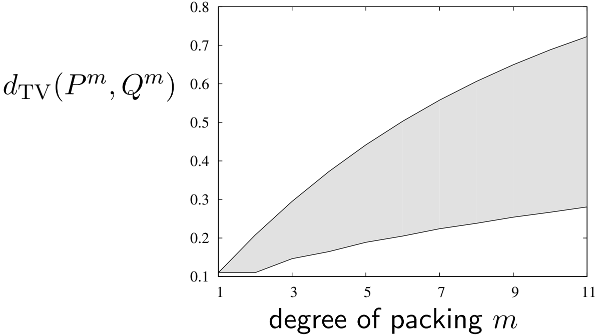
Evolution of TV distances



Evolution of TV distances through the prism of packing



Through packing, the target-generator pairs are expanded over the strengths of the mode collapse



$$\begin{array}{ll} \max_{P,Q} / \min_{P,Q} & d_{\text{TV}}(P^2, Q^2) \\ \text{subject to} & d_{\text{TV}}(P, Q) = \tau \end{array}$$

- we focus on $m = 2$ for this talk

Intuition from Blackwell

Definition [mode collapse region]

We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

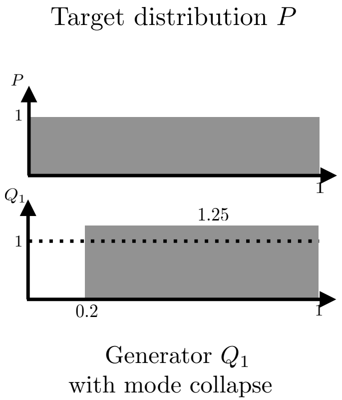
$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$

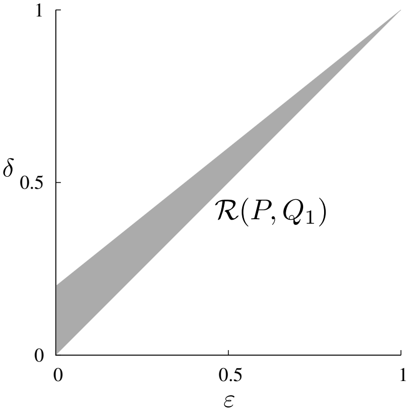
Intuition from Blackwell

Definition [mode collapse region]

We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \;, \quad \text{and} \quad Q(S) \leq \varepsilon \,.$$





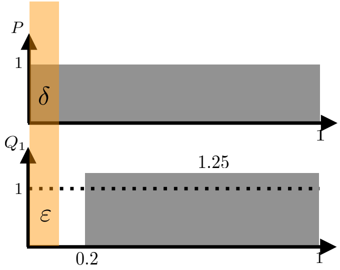
Intuition from Blackwell

Definition [mode collapse region]

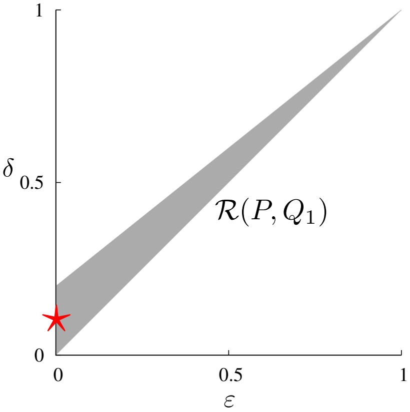
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \qquad \text{and} \qquad Q(S) \leq \varepsilon \ .$$

Target distribution P



Generator Q_1
with mode collapse



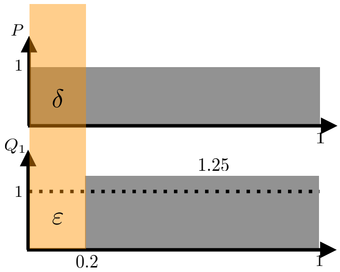
Intuition from Blackwell

Definition [mode collapse region]

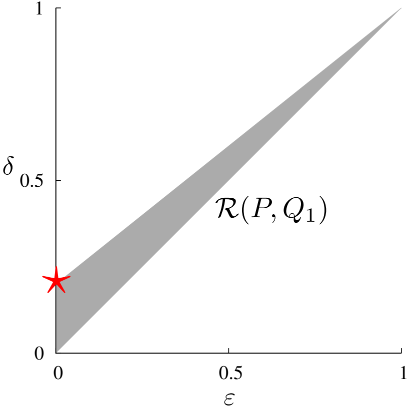
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$

Target distribution P



Generator Q_1
with mode collapse



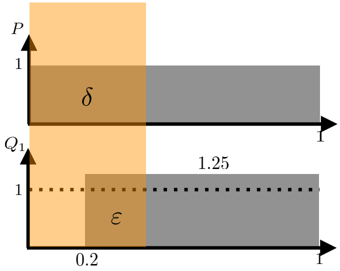
Intuition from Blackwell

Definition [mode collapse region]

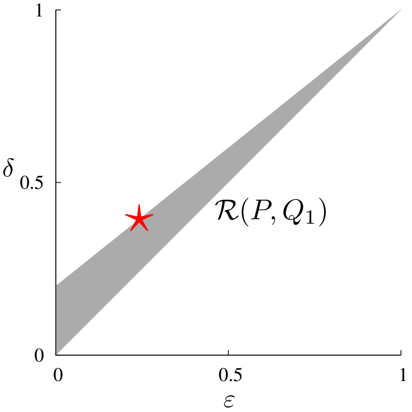
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \;, \quad \text{and} \quad Q(S) \leq \varepsilon \;.$$

Target distribution P



Generator Q_1
with mode collapse



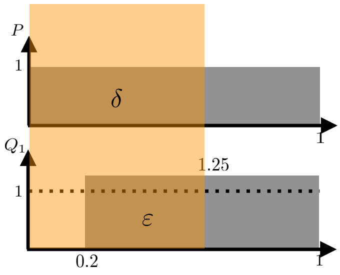
Intuition from Blackwell

Definition [mode collapse region]

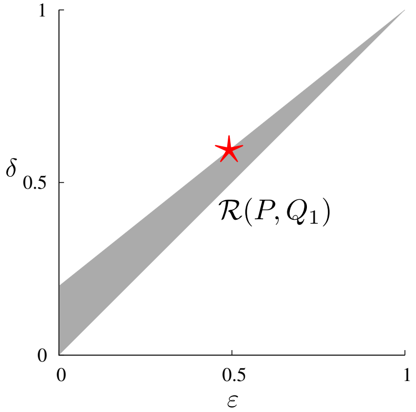
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$

Target distribution P



Generator Q_1
with mode collapse



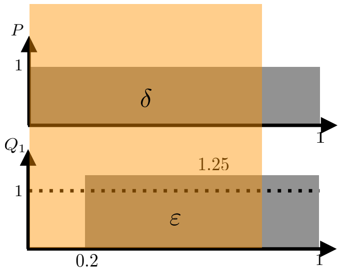
Intuition from Blackwell

Definition [mode collapse region]

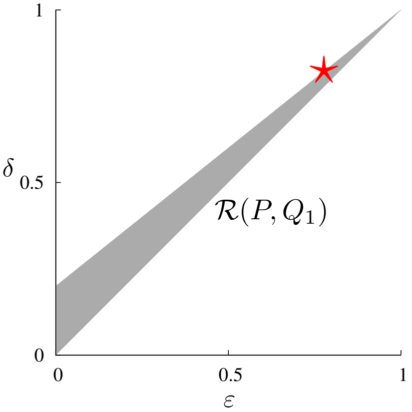
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \qquad \text{and} \qquad Q(S) \leq \varepsilon \ .$$

Target distribution P



Generator Q_1
with mode collapse

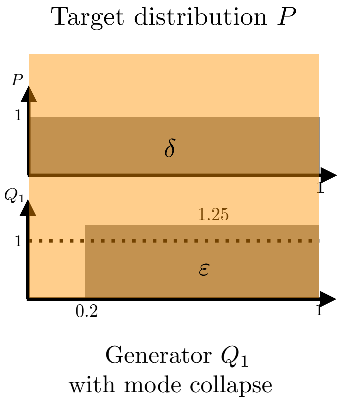


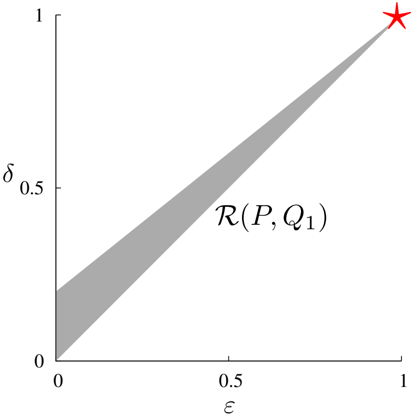
Intuition from Blackwell

Definition [mode collapse region]

We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$





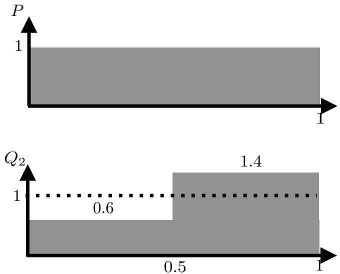
Intuition from Blackwell

Definition [mode collapse region]

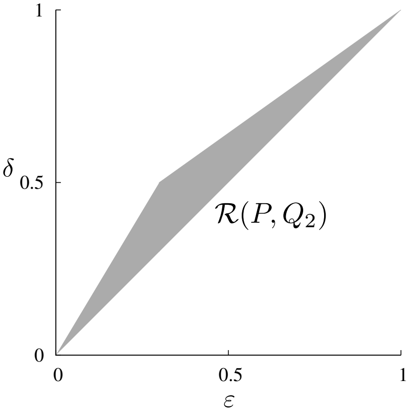
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$

Target distribution P



Generator Q_2
without mode collapse

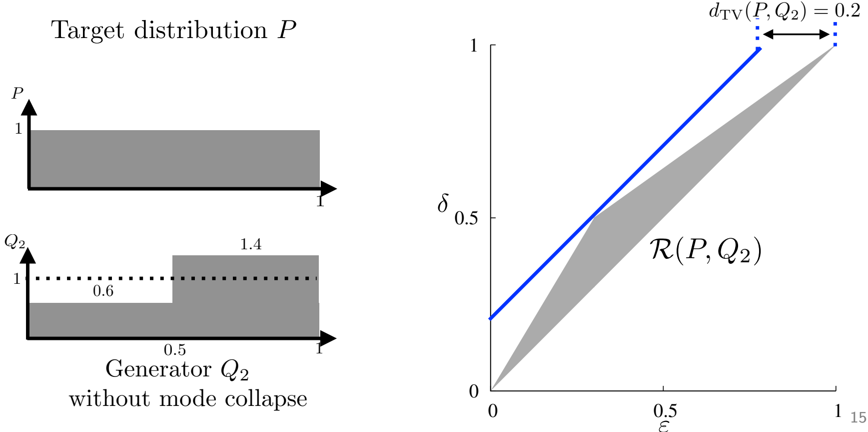


Intuition from Blackwell

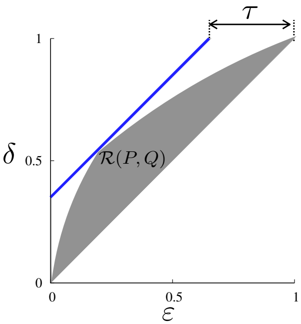
Definition [mode collapse region]

We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

$$P(S) \geq \delta \ , \quad \text{and} \quad Q(S) \leq \varepsilon \ .$$

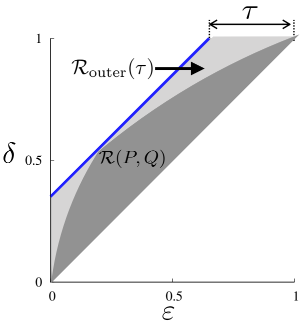


Upper bound



$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$
$$\text{subject to} \quad d_{\mathrm{TV}}(P, Q) = \tau$$

Upper bound

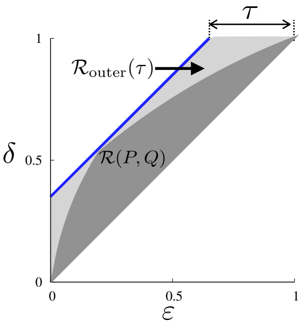


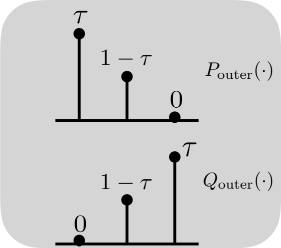
$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

subject to $\quad d_{\mathrm{TV}}(P, Q) = \tau$

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}_{\text{outer}}(\tau)$$

Upper bound

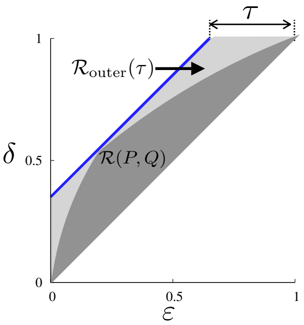


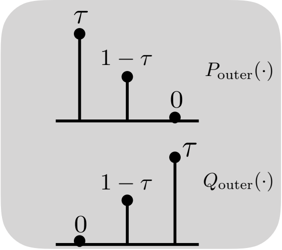


$$\begin{aligned} \max_{P,Q} \quad & d_{\text{TV}}(P^2, Q^2) \\ \text{subject to} \quad & d_{\text{TV}}(P,Q) = \tau \end{aligned}$$

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}_{\text{outer}}(\tau)$$

Upper bound





$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

subject to $\quad d_{\text{TV}}(P,Q) = \tau$

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}_{\text{outer}}(\tau)$$

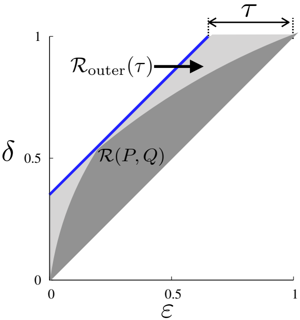
$$\mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P^2_{\text{outer}}, Q^2_{\text{outer}})$$

Blackwell's theorem

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P',Q')$$

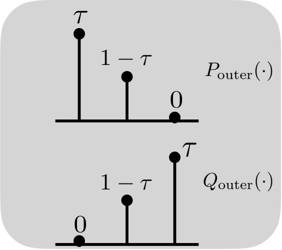
$$\Rightarrow \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2,Q'^2)$$

Upper bound



$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$
$$\text{subject to} \quad d_{\text{TV}}(P, Q) = \tau$$

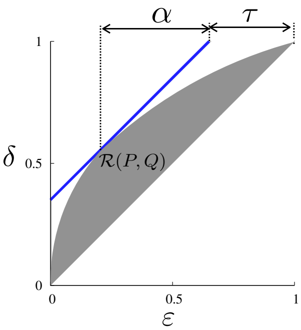
$$\begin{aligned} \mathcal{R}(P,Q) &\subseteq \mathcal{R}_{\text{outer}}(\tau) \\ \mathcal{R}(P^2,Q^2) &\subseteq \mathcal{R}(P^2_{\text{outer}}, Q^2_{\text{outer}}) \\ d_{\text{TV}}(P^2,Q^2) &\leq \underbrace{d_{\text{TV}}(P^2_{\text{outer}}, Q^2_{\text{outer}})}_{1-(1-\tau)^2} \end{aligned}$$



Blackwell's theorem

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P',Q')$$
$$\Rightarrow \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2,Q'^2)$$

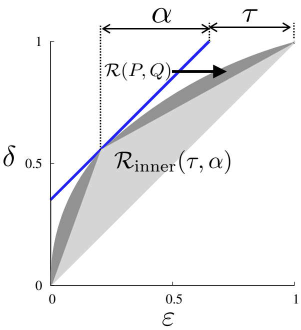
Lower bound



$$\min_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

$$\text{subject to} \quad d_{\mathrm{TV}}(P, Q) = \tau$$

Lower bound

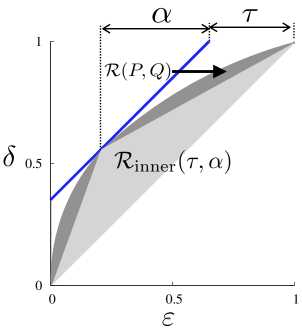


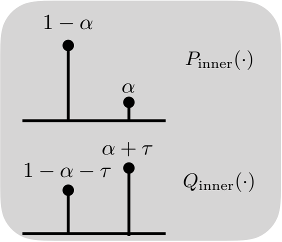
$$\min_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

$$\text{subject to} \quad d_{\text{TV}}(P,Q) = \tau$$

$$\mathcal{R}_{\text{inner}}(\tau,\alpha) \quad \subseteq \quad \mathcal{R}(P,Q)$$

Lower bound



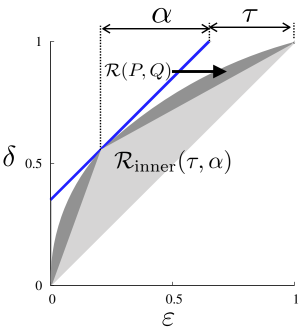


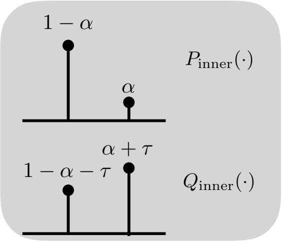
$$\min_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

$$\text{subject to} \quad d_{\text{TV}}(P,Q) = \tau$$

$$\mathcal{R}_{\text{inner}}(\tau,\alpha) \subseteq \mathcal{R}(P,Q)$$

Lower bound





$$\min_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

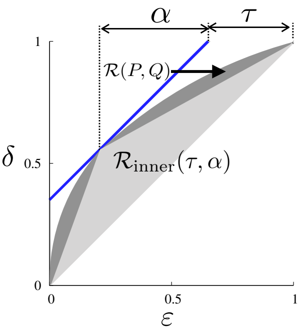
subject to $\quad d_{\text{TV}}(P,Q) = \tau$

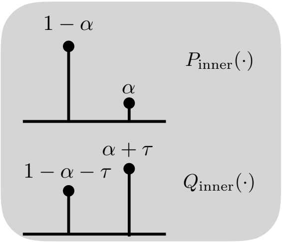
$$\begin{aligned} \mathcal{R}_{\text{inner}}(\tau,\alpha) &\subseteq \mathcal{R}(P,Q) \\ \mathcal{R}(P^2_{\text{inner}}, Q^2_{\text{inner}}) &\subseteq \mathcal{R}(P^2,Q^2) \end{aligned}$$

Blackwell's theorem

$$\begin{aligned} &\mathcal{R}(P,Q) \subseteq \mathcal{R}(P',Q') \\ \Rightarrow\ &\mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2,Q'^2) \end{aligned}$$

Lower bound





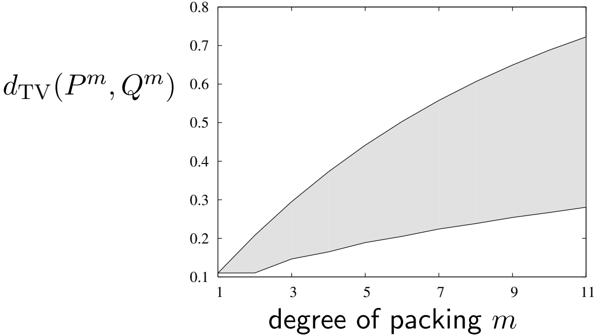
$$\min_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

$$\text{subject to} \quad d_{\mathrm{TV}}(P,Q) = \tau$$

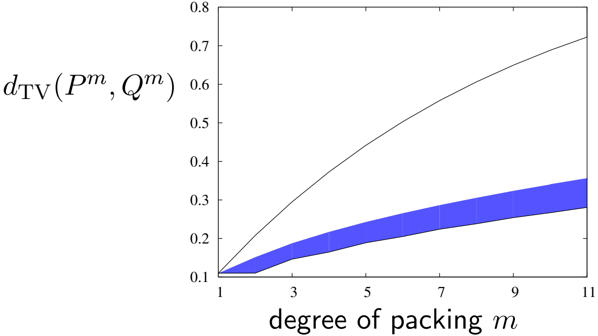
$$\begin{aligned}
\mathcal{R}_{\text{inner}}(\tau,\alpha) &\subseteq \mathcal{R}(P,Q) \\
\mathcal{R}(P_{\text{inner}}^2, Q_{\text{inner}}^2) &\subseteq \mathcal{R}(P^2,Q^2) \\
\underbrace{\min_{\alpha}\ d_{\mathrm{TV}}(P_{\text{inner}}^2, Q_{\text{inner}}^2)}_{\tau} &\leq d_{\mathrm{TV}}(P^2,Q^2)
\end{aligned}$$

Blackwell's theorem

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P',Q')$$
$$\Rightarrow \quad \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2,Q'^2)$$

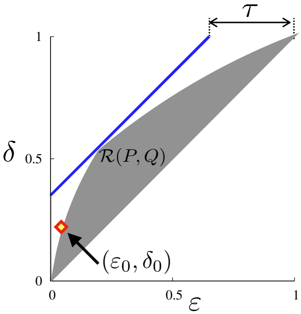


$$\max_{P,Q} / \min_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$
$$\text{subject to} \quad d_{\mathrm{TV}}(P, Q) = \tau$$



$$\begin{array}{rl} \max\limits_{P,Q} & d_{\mathrm{TV}}(P^2, Q^2) \\ \text{subject to} & d_{\mathrm{TV}}(P, Q) = \tau \\ & \text{no } (\varepsilon_0, \delta_0)\text{-mode collapse} \end{array}$$

Upper bound without $(\varepsilon_0, \delta_0)$-mode collapse

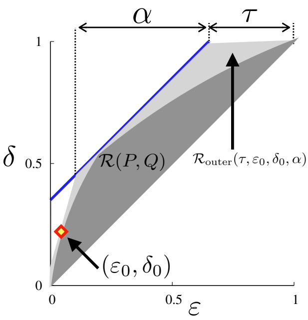


$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

subject to $\quad d_{\mathrm{TV}}(P, Q) = \tau$

no $(\varepsilon_0, \delta_0)$-mode collapse

Upper bound without $(\varepsilon_0, \delta_0)$-mode collapse



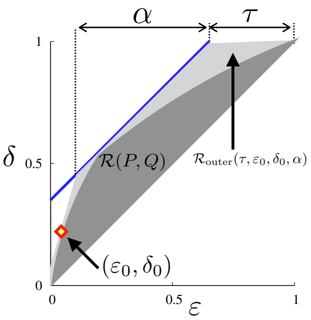
$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

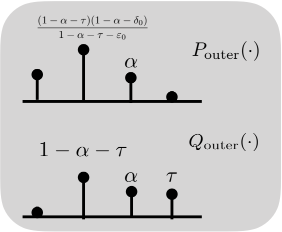
subject to $\quad d_{\mathrm{TV}}(P,Q) = \tau$

no $(\varepsilon_0, \delta_0)$-mode collapse

$$\mathcal{R}(P,Q) \quad \subseteq \quad \mathcal{R}_{\text{outer}}(\tau, \varepsilon_0, \delta_0, \alpha)$$

Upper bound without $(\varepsilon_0, \delta_0)$-mode collapse





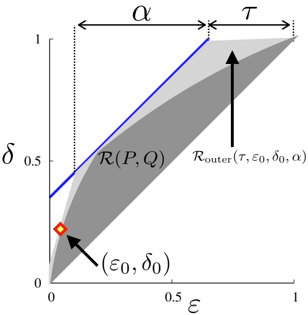
$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

subject to $\quad d_{\mathrm{TV}}(P, Q) = \tau$

no $(\varepsilon_0, \delta_0)$-mode collapse

$$\mathcal{R}(P,Q) \quad \subseteq \quad \mathcal{R}_{\text{outer}}(\tau, \varepsilon_0, \delta_0, \alpha)$$

Upper bound without $(\varepsilon_0, \delta_0)$-mode collapse

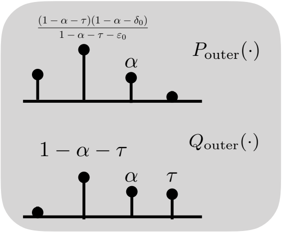


$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

subject to $\quad d_{\text{TV}}(P,Q) = \tau$

no $(\varepsilon_0, \delta_0)$-mode collapse

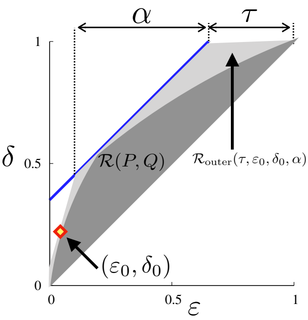
$$\mathcal{R}(P,Q) \subseteq \mathcal{R}_{\text{outer}}(\tau, \varepsilon_0, \delta_0, \alpha)$$
$$\mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P^2_{\text{outer}}, Q^2_{\text{outer}})$$

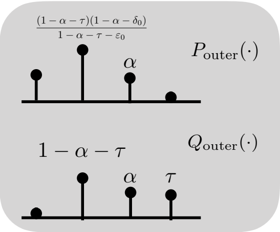


Blackwell's theorem

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P',Q')$$
$$\Rightarrow \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2,Q'^2)$$

Upper bound without $(\varepsilon_0, \delta_0)$-mode collapse





$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

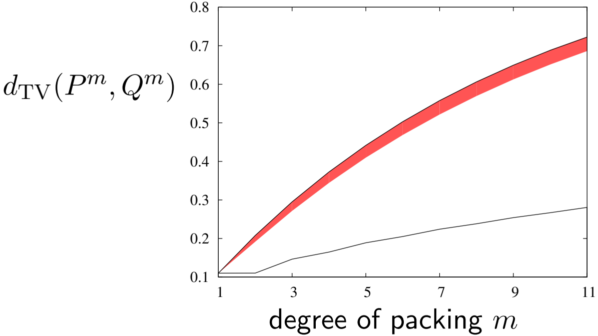
subject to $\quad d_{\text{TV}}(P,Q) = \tau$

no $(\varepsilon_0, \delta_0)$-mode collapse

$$\begin{aligned}
\mathcal{R}(P,Q) &\subseteq \mathcal{R}_{\text{outer}}(\tau, \varepsilon_0, \delta_0, \alpha) \\
\mathcal{R}(P^2,Q^2) &\subseteq \mathcal{R}(P^2_{\text{outer}}, Q^2_{\text{outer}}) \\
d_{\text{TV}}(P^2,Q^2) &\leq \underbrace{\max_{\alpha} d_{\text{TV}}(P^2_{\text{outer}}, Q^2_{\text{outer}})}_{\text{simple to evaluate}}
\end{aligned}$$

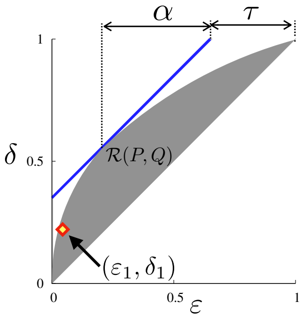
Blackwell's theorem

$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P', Q')$$
$$\Rightarrow \quad \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2, Q'^2)$$



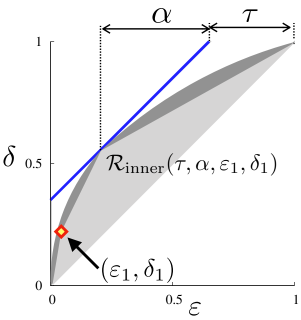
$$\begin{array}{rl} \min\limits_{P,Q} & d_{\mathrm{TV}}(P^2, Q^2) \\ \text{subject to} & d_{\mathrm{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{array}$$

Lower bound with $(\varepsilon_1, \delta_1)$-mode collapse



$$
\begin{aligned}
\min_{P,Q} \quad & d_{\mathrm{TV}}(P^2, Q^2) \\
\text{subject to} \quad & d_{\mathrm{TV}}(P, Q) = \tau \\
& (\varepsilon_1, \delta_1)\text{-mode collapse}
\end{aligned}
$$

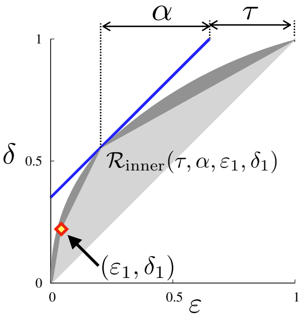
Lower bound with $(\varepsilon_1, \delta_1)$-mode collapse

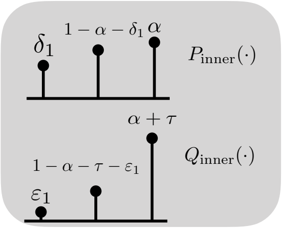


$$\begin{aligned} \min_{P,Q} \quad & d_{\mathrm{TV}}(P^2, Q^2) \\ \text{subject to} \quad & d_{\mathrm{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{aligned}$$

$$\mathcal{R}_{\text{inner}}(\tau, \alpha, \varepsilon_1, \delta_1) \subseteq \mathcal{R}(P, Q)$$

Lower bound with $(\varepsilon_1, \delta_1)$-mode collapse

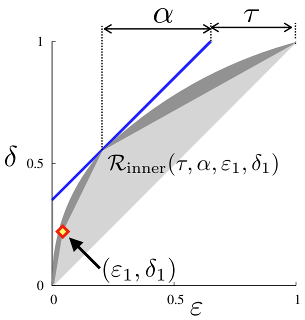


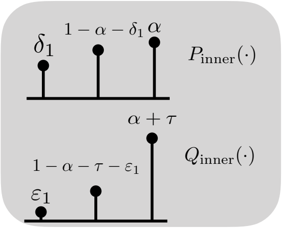


$$\begin{aligned} \min_{P,Q} \quad & d_{\text{TV}}(P^2, Q^2) \\ \text{subject to} \quad & d_{\text{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{aligned}$$

$$\mathcal{R}_{\text{inner}}(\tau, \alpha, \varepsilon_1, \delta_1) \subseteq \mathcal{R}(P, Q)$$

Lower bound with $(\varepsilon_1, \delta_1)$-mode collapse





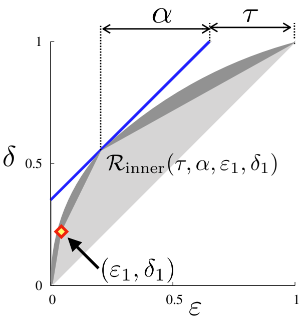
$$\begin{aligned} \min_{P,Q} \quad & d_{\text{TV}}(P^2, Q^2) \\ \text{subject to} \quad & d_{\text{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{aligned}$$

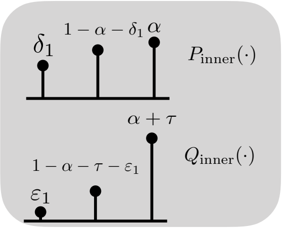
$$\begin{aligned} \mathcal{R}_{\text{inner}}(\tau, \alpha, \varepsilon_1, \delta_1) & \subseteq \mathcal{R}(P, Q) \\ \mathcal{R}(P^2_{\text{inner}}, Q^2_{\text{inner}}) & \subseteq \mathcal{R}(P^2, Q^2) \end{aligned}$$

Blackwell's theorem

$$\begin{aligned} & \mathcal{R}(P, Q) \subseteq \mathcal{R}(P', Q') \\ \Rightarrow \quad & \mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P'^2, Q'^2) \end{aligned}$$

Lower bound with $(\varepsilon_1, \delta_1)$-mode collapse





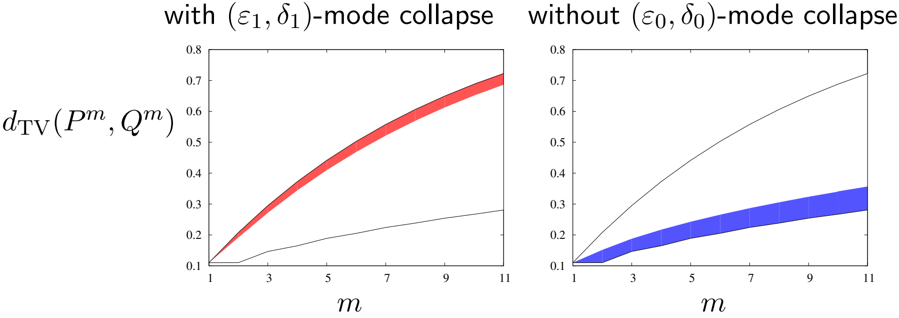
$$\begin{aligned} \min_{P,Q} \quad & d_{\text{TV}}(P^2, Q^2) \\ \text{subject to} \quad & d_{\text{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{aligned}$$

$$\begin{aligned} \mathcal{R}_{\text{inner}}(\tau, \alpha, \varepsilon_1, \delta_1) &\subseteq \mathcal{R}(P, Q) \\ \mathcal{R}(P^2_{\text{inner}}, Q^2_{\text{inner}}) &\subseteq \mathcal{R}(P^2, Q^2) \\ \underbrace{\min_{\alpha} \; d_{\text{TV}}(P^2_{\text{inner}}, Q^2_{\text{inner}})}_{\text{simple to evaluate}} &\leq d_{\text{TV}}(P^2, Q^2) \end{aligned}$$

Blackwell's theorem

$$\begin{aligned} & \mathcal{R}(P, Q) \subseteq \mathcal{R}(P', Q') \\ \Rightarrow \quad & \mathcal{R}(P^2, Q^2) \subseteq \mathcal{R}(P'^2, Q'^2) \end{aligned}$$

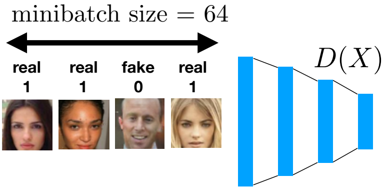
Achievable TV distances for distributions



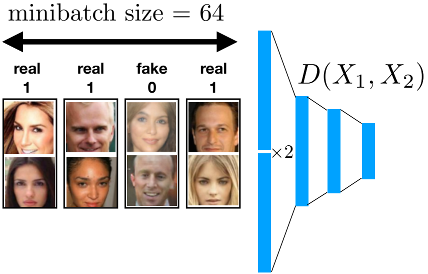
with packing, the discriminator naturally penalizes (P, Q) with severe mode collapses

Could we be cheating (hyper-parameter tuning)?

1. Discriminator size



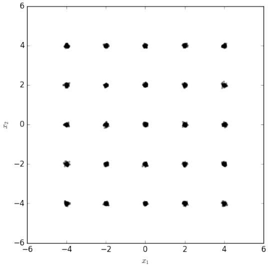
GAN



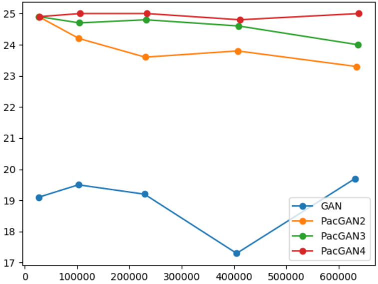
PacGAN2

Could we be cheating (hyper-parameter tuning)?

1. Discriminator size



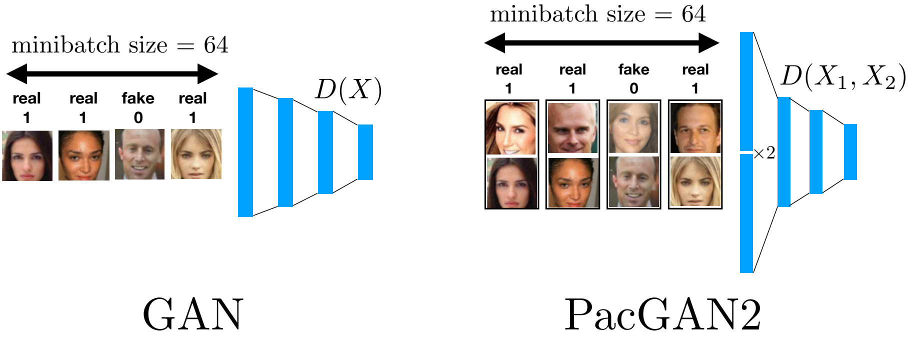
\# modes captured



\# of parameters in $D(\cdot)$

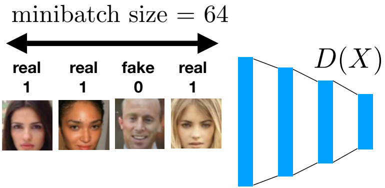
Could we be cheating (hyper-parameter tuning)?

2. Minibatch size

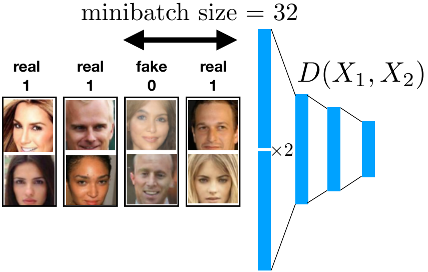


Could we be cheating (hyper-parameter tuning)?

2. Minibatch size



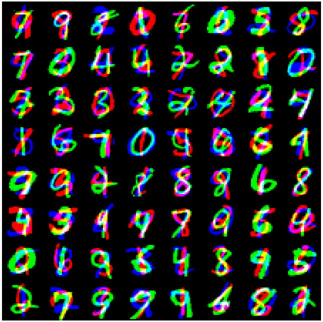
GAN



PacGAN2

Could we be cheating (hyper-parameter tuning)?

2. Minibatch size



	Modes
DCGAN	99.0
PacDCGAN2	1000.0

Theoretical challenges in GAN

Designing Loss	
$D_{\mathrm{JS}}(P, Q)$	Jansen-Shannon
$D_f(P, Q)$	f-divergence
$D_{\mathrm{W}}(P, Q)$	Wasserstein

[FeiziSuhXiaTse 2017]

Theoretical challenges in GAN

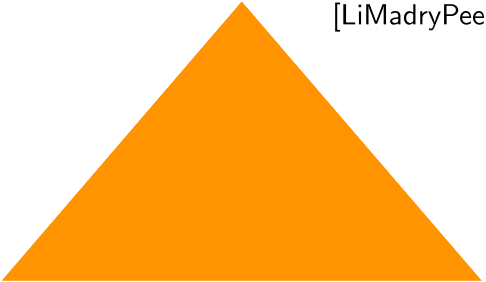
Designing Loss		Evaluation
$D_{\mathrm{JS}}(P,Q)$	Jansen-Shannon	
$D_f(P,Q)$	f-divergence	
$D_{\mathrm{W}}(P,Q)$	Wasserstein	

$D(P^m,Q^m) \succ D(P,Q)$ (ε,δ)-mode collapse

Theoretical challenges in GAN

Training Dynamics

[LiMadryPeeblesSchmidt 2017]



Designing Loss

Evaluation

$D_{\mathrm{JS}}(P, Q)$ Jansen-Shannon

$D_f(P, Q)$ f-divergence

$D_{\mathrm{W}}(P, Q)$ Wasserstein

$D(P^m, Q^m) \succ D(P, Q)$

(ε, δ)-mode collapse

Our paper is available at:
https://arxiv.org/abs/1712.04086

All codes for the experiments at:
https://github.com/fjxmlzn/PacGAN



Zinan Lin (CMU) Ashish Khetan (UIUC) Giulia Fanti (CMU)