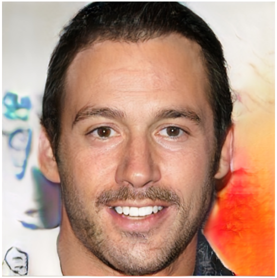
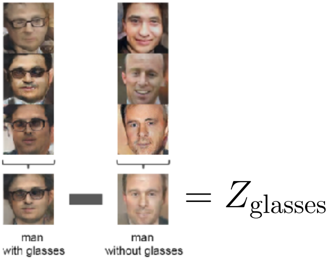
The power of two samples in generative adversarial networks

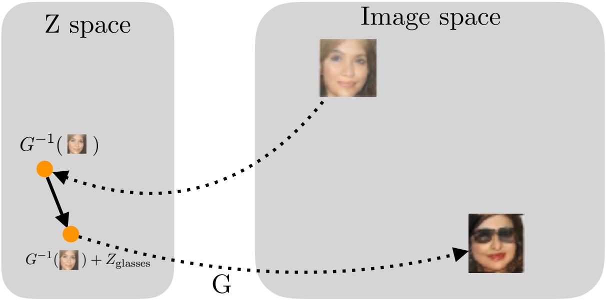
Sewoong Oh

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

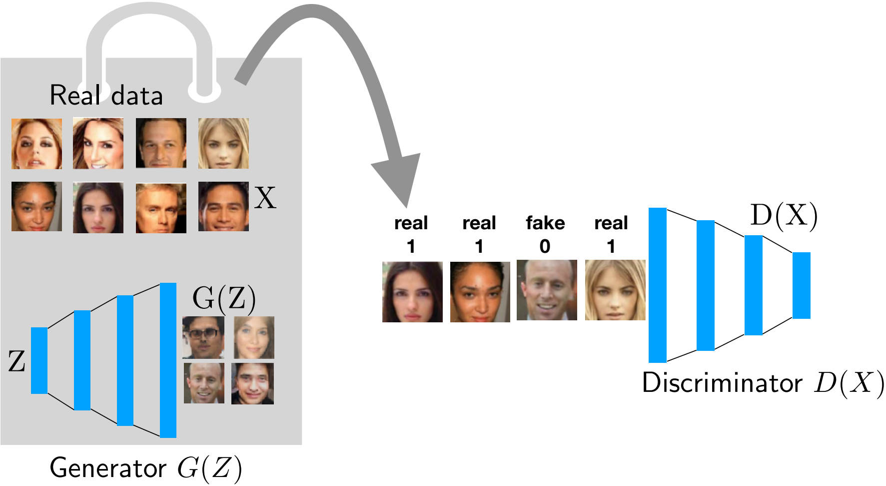

Generative models learn fundamental representations





[DCGAN, Radford et al. 2015]

Generative Adversarial Networks (GAN)



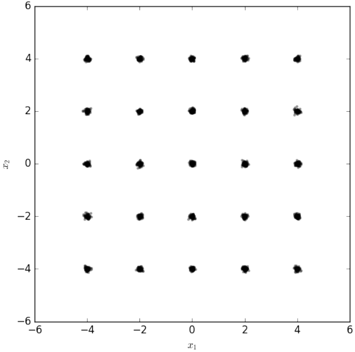
$$\min_{G} \max_{D} V(G, D)$$

Challenges in training GAN

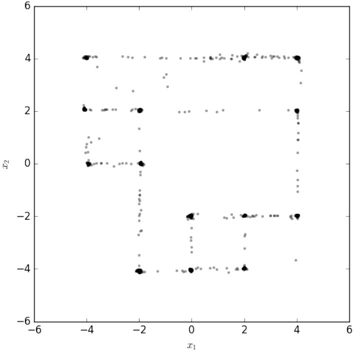
1. Instability: non-convergence of training loss
2. Evaluation: likelihood is not available
3. Mode collapse: loss of diversity

"Mode collapse" is a main challenge

Target samples

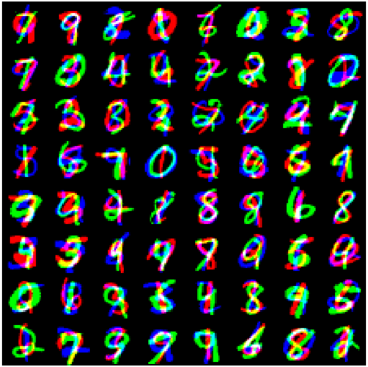


Generated samples

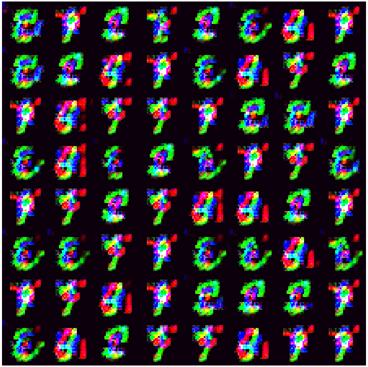


“Mode collapse” is a main challenge

Target samples



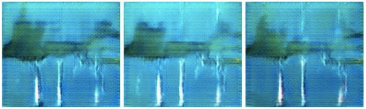
Generated samples



“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.”

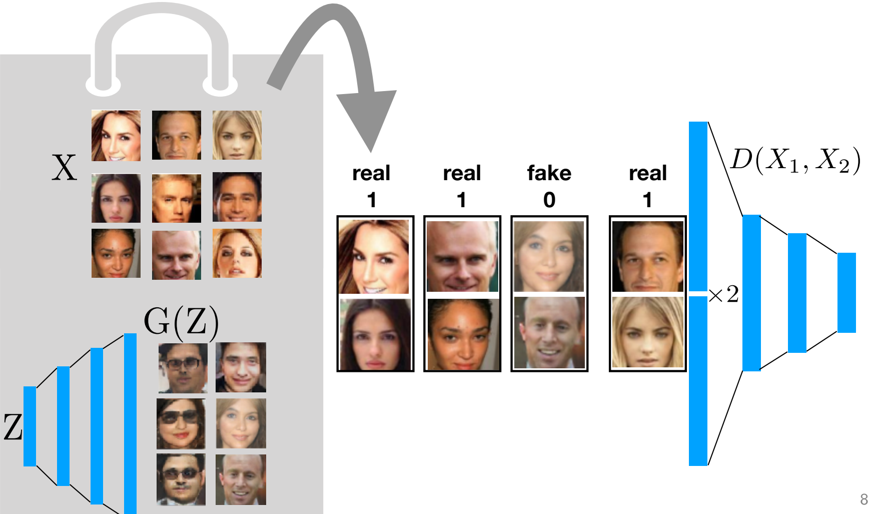
[“Generating interpretable images with controllable structure”, by Reed et al., 2016]

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

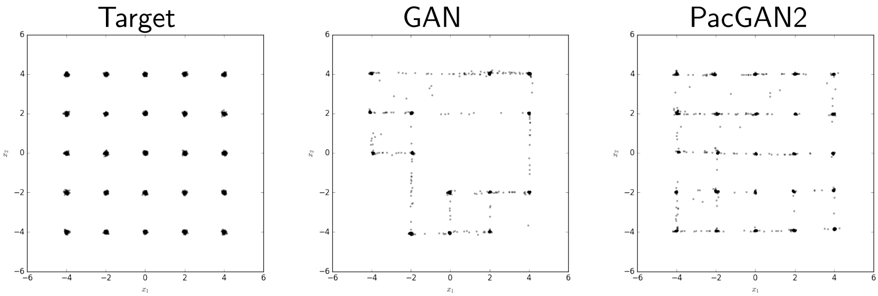
- "Improved Techniques for Training GANs", Salimans, Goodfellow, Zaremba, Cheung, Radford, Chen, 2016
- "Progressive Growing of GANs for Improved Quality, Stability, and Variation", Karras, Aila, Laine, Lehtinen, 2017
- "Distributional Adversarial Networks", Li, Alvarez-Melis, Xu, Jegelka, Sra, 2017

New framework: PacGAN

- lightweight overhead
- experimental results
- principled



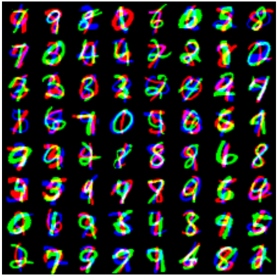
Benchmark tests



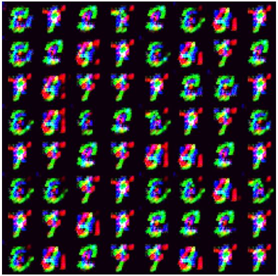
	Modes (Max 25)
GAN	17.3
PacGAN2	23.8
PacGAN3	24.6
PacGAN4	24.8

Benchmark datasets from VEEGAN paper

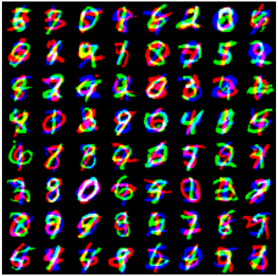
Real data



DCGAN



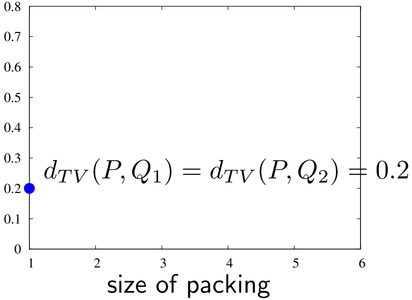
PacDCGAN2



	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

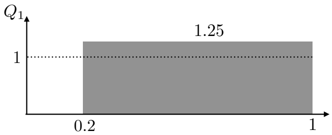
Intuition behind packing via toy example

Target distribution P

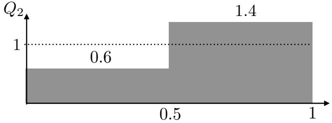


Generator Q_1
with mode collapse

Generator Q_2
without mode collapse



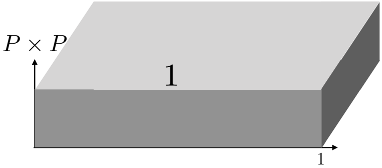
$d_{\mathrm{TV}}(P, Q_1) = 0.2$

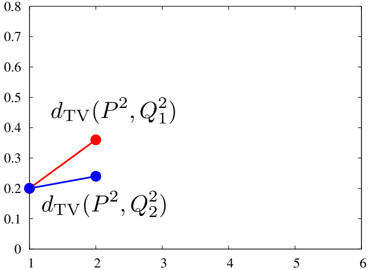


$d_{\mathrm{TV}}(P, Q_2) = 0.2$

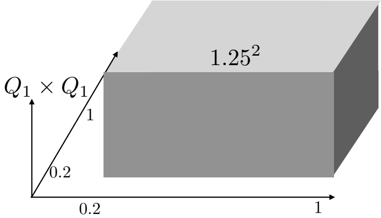
Intuition behind packing via toy example

Target distribution P

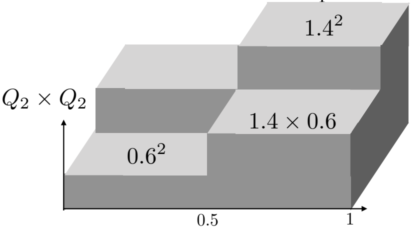




Generator Q_1
with mode collapse



Generator Q_2
without mode collapse

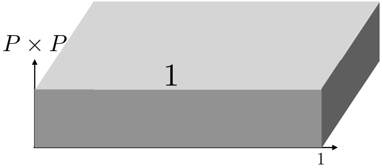


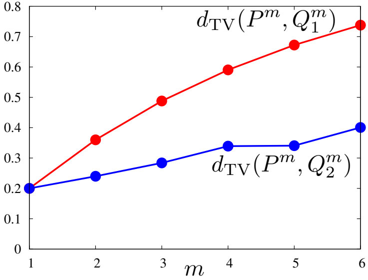
$d_{\mathrm{TV}}(P \times P, Q_1 \times Q_1) = 0.36$

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

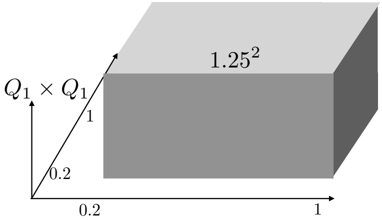
Intuition behind packing via toy example

Target distribution P



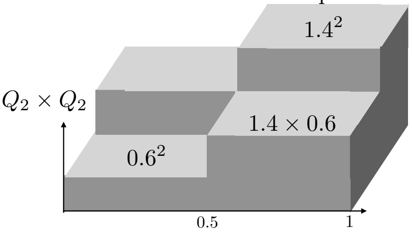


Generator Q_1
with mode collapse



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

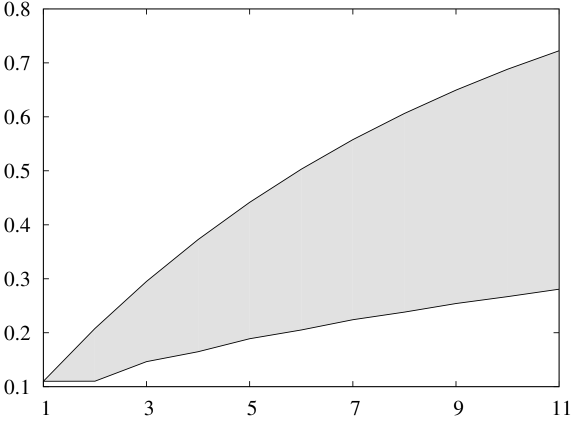
Generator Q_2
without mode collapse



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

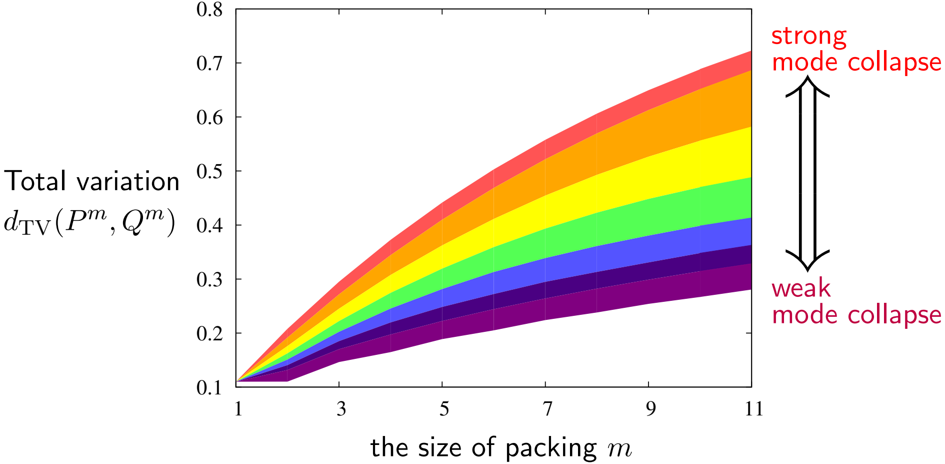
Evolution of TV distances

Total variation
$d_{\mathrm{TV}}(P^m, Q^m)$

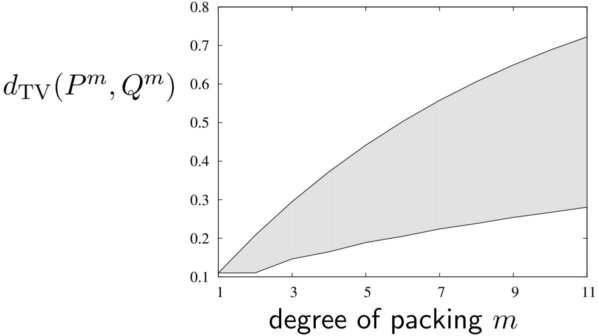


the size of packing m

Evolution of TV distances through the prism of packing



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



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

- we focus on $m = 2$ for this talk
- this is easy, but we have a new proof technique
- nothing to do with mode collapse, but we use it as proof technique

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

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

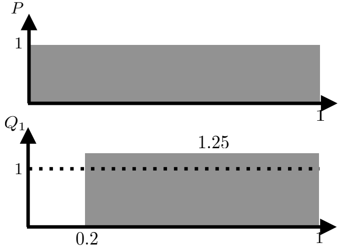
Intuition from Blackwell

Definition [mode collapse region]

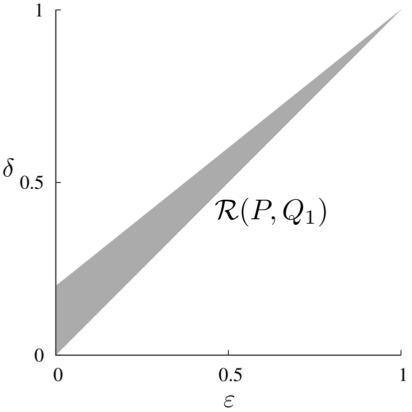
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

Target distribution P



Generator Q_1
with mode collapse

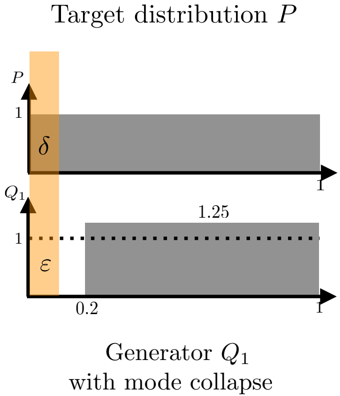


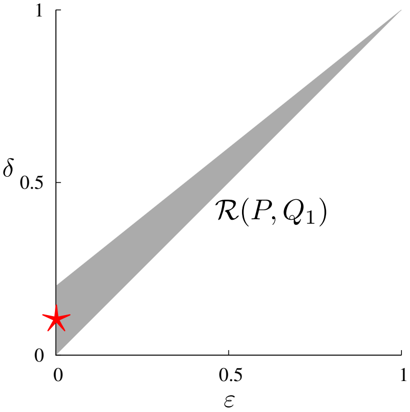
Intuition from Blackwell

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

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





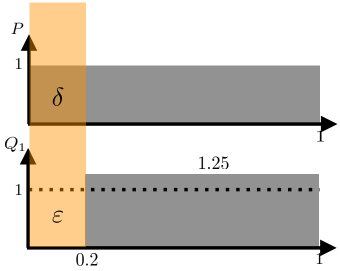
Intuition from Blackwell

Definition [mode collapse region]

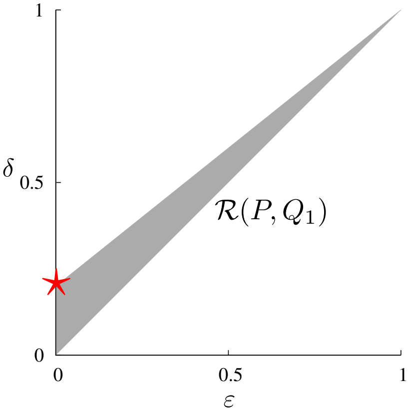
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

Target distribution P



Generator Q_1
with mode collapse

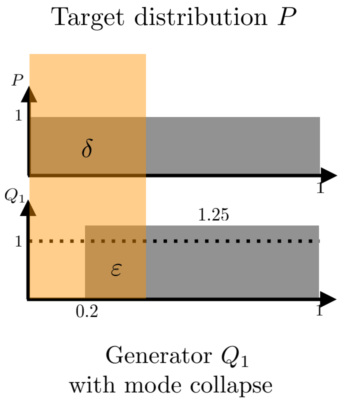


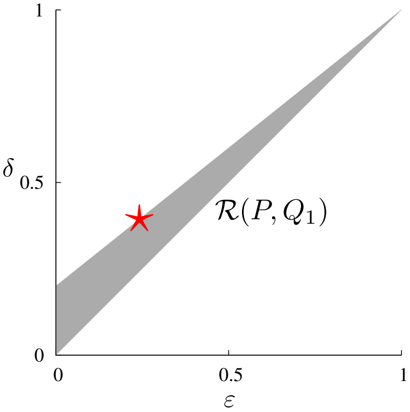
Intuition from Blackwell

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

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



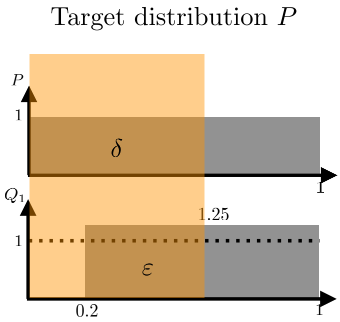


Intuition from Blackwell

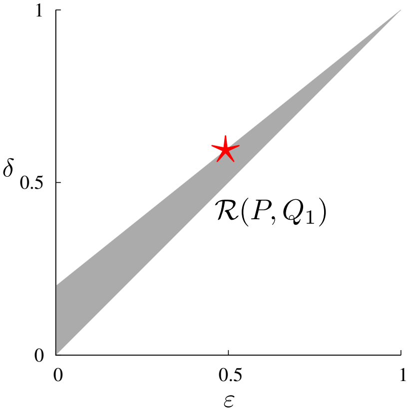
Definition [mode collapse region]

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

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



Generator Q_1
with mode collapse



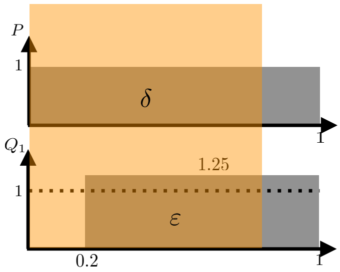
Intuition from Blackwell

Definition [mode collapse region]

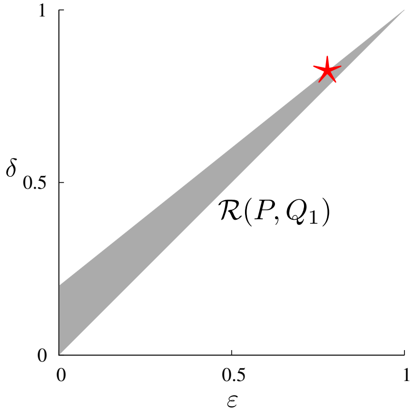
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

Target distribution P



Generator Q_1
with mode collapse



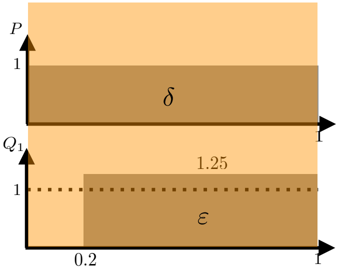
Intuition from Blackwell

Definition [mode collapse region]

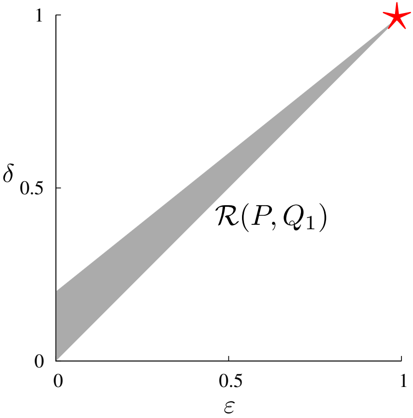
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

Target distribution P



Generator Q_1
with mode collapse



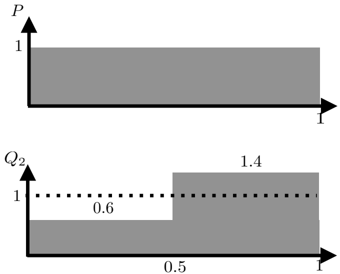
Intuition from Blackwell

Definition [mode collapse region]

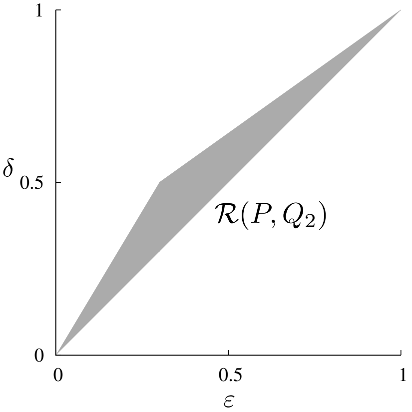
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

Target distribution P



Generator Q_2
without mode collapse



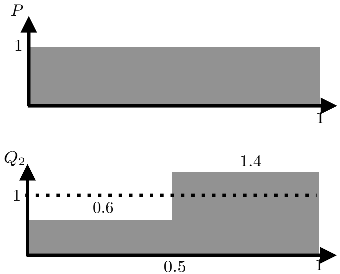
Intuition from Blackwell

Definition [mode collapse region]

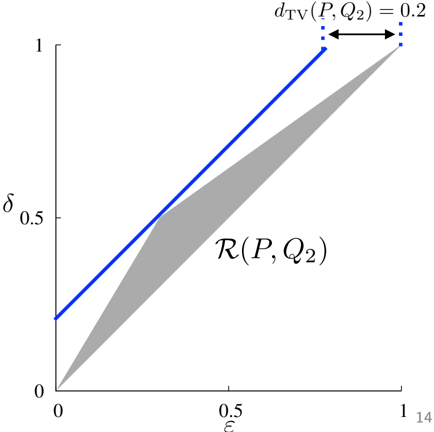
We say a pair (P, Q) of a target distribution P and a generator distribution Q has (ε, δ)-**mode collapse** if there exists a set S such that

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

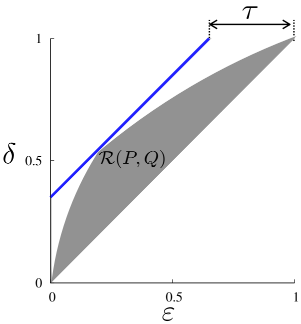
Target distribution P



Generator Q_2
without mode collapse



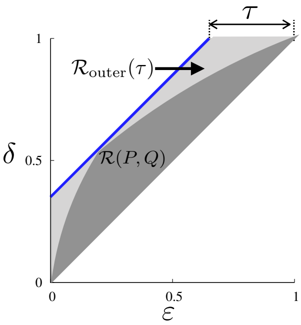
Upper bound



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

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

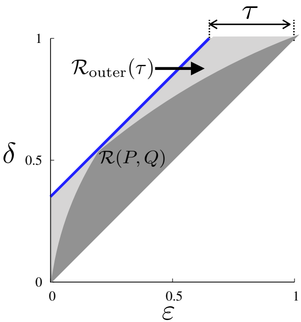
Upper bound

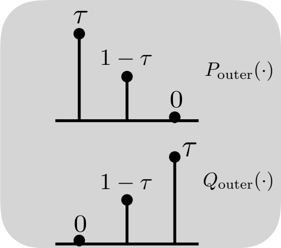


$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$
$$\text{subject to} \quad d_{\text{TV}}(P,Q) = \tau$$

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

Upper bound



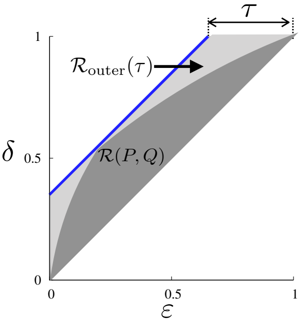


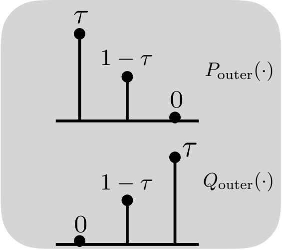
$$\max_{P,Q} \quad d_{\text{TV}}(P^2, Q^2)$$

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

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

Upper bound





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

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

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

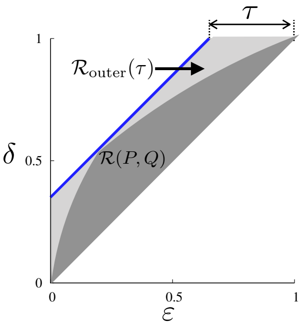
$$\mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P^2_{\rm outer}, Q^2_{\rm outer})$$

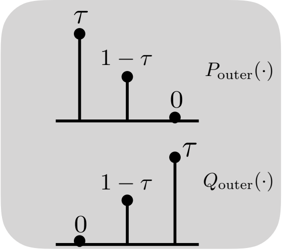
Blackwell's theorem

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

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

Upper bound



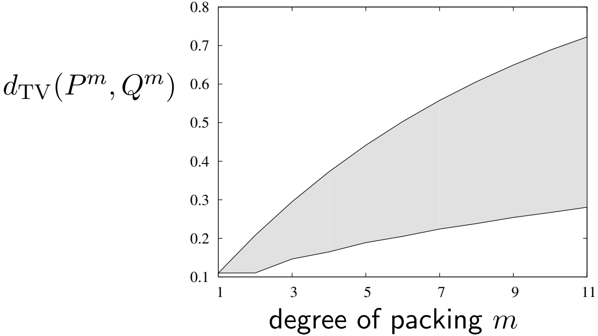


$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$
$$\text{subject to} \quad d_{\mathrm{TV}}(P, Q) = \tau$$

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

Blackwell's theorem

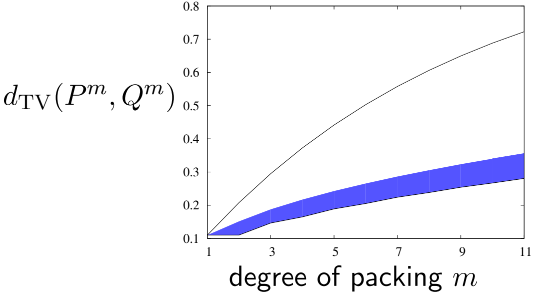
$$\mathcal{R}(P,Q) \subseteq \mathcal{R}(P', Q')$$
$$\Rightarrow \quad \mathcal{R}(P^2,Q^2) \subseteq \mathcal{R}(P'^2, Q'^2)$$



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

PacGAN naturally penalizes mode collapse

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

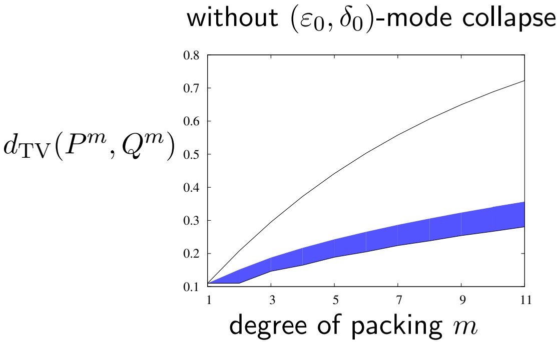


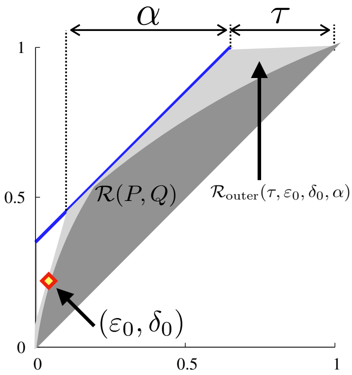
$$\max_{P,Q} \quad d_{\mathrm{TV}}(P^2, Q^2)$$

$$\text{s.t.} \quad d_{\mathrm{TV}}(P, Q) = \tau$$

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

PacGAN naturally penalizes mode collapse





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

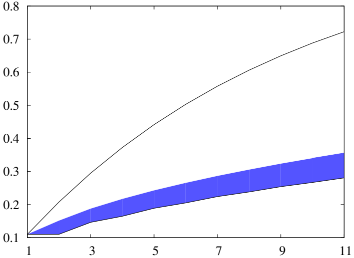
s.t. $\quad d_{\mathrm{TV}}(P, Q) = \tau$

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

PacGAN naturally penalizes mode collapse

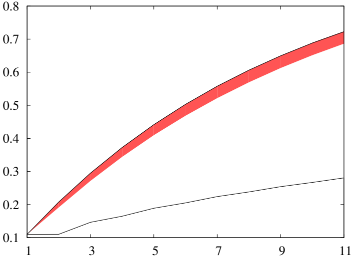
$d_{\mathrm{TV}}(P^m, Q^m)$

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



degree of packing m

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

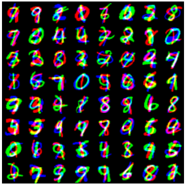


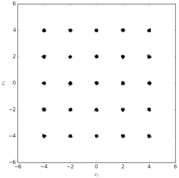
degree of packing m

$$\begin{aligned} \max_{P,Q} \quad & d_{\mathrm{TV}}(P^2, Q^2) \\ \text{s.t.} \quad & d_{\mathrm{TV}}(P, Q) = \tau \\ & \text{no } (\varepsilon_0, \delta_0)\text{-mode collapse} \end{aligned}$$

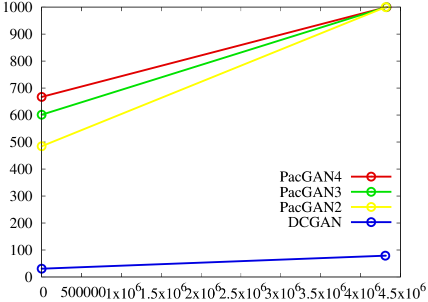
$$\begin{aligned} \min_{P,Q} \quad & d_{\mathrm{TV}}(P^2, Q^2) \\ \text{s.t.} \quad & d_{\mathrm{TV}}(P, Q) = \tau \\ & (\varepsilon_1, \delta_1)\text{-mode collapse} \end{aligned}$$

Size of the discriminator



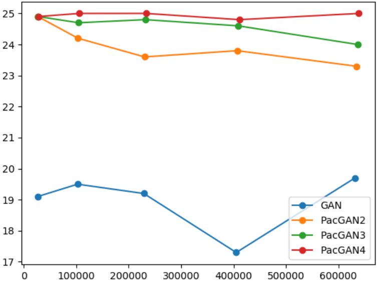


$\#$ modes captured



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

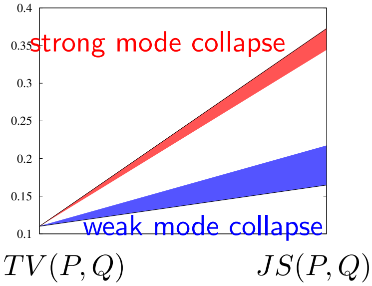
$\#$ modes captured

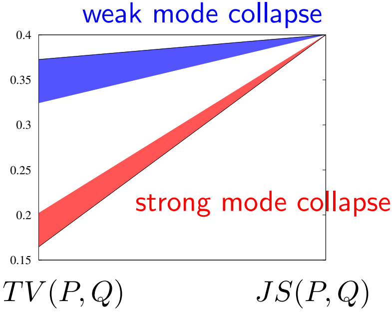


$\#$ of parameters in $D(\cdot)$

Mini-batch discrimination requires +38,748,557, PacGAN2 requires +54

0-1 loss (Total Variation) vs. cross entropy loss (Jensen-Shannon Divergence)





Jensen-Shannon is better measure for detecting 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)