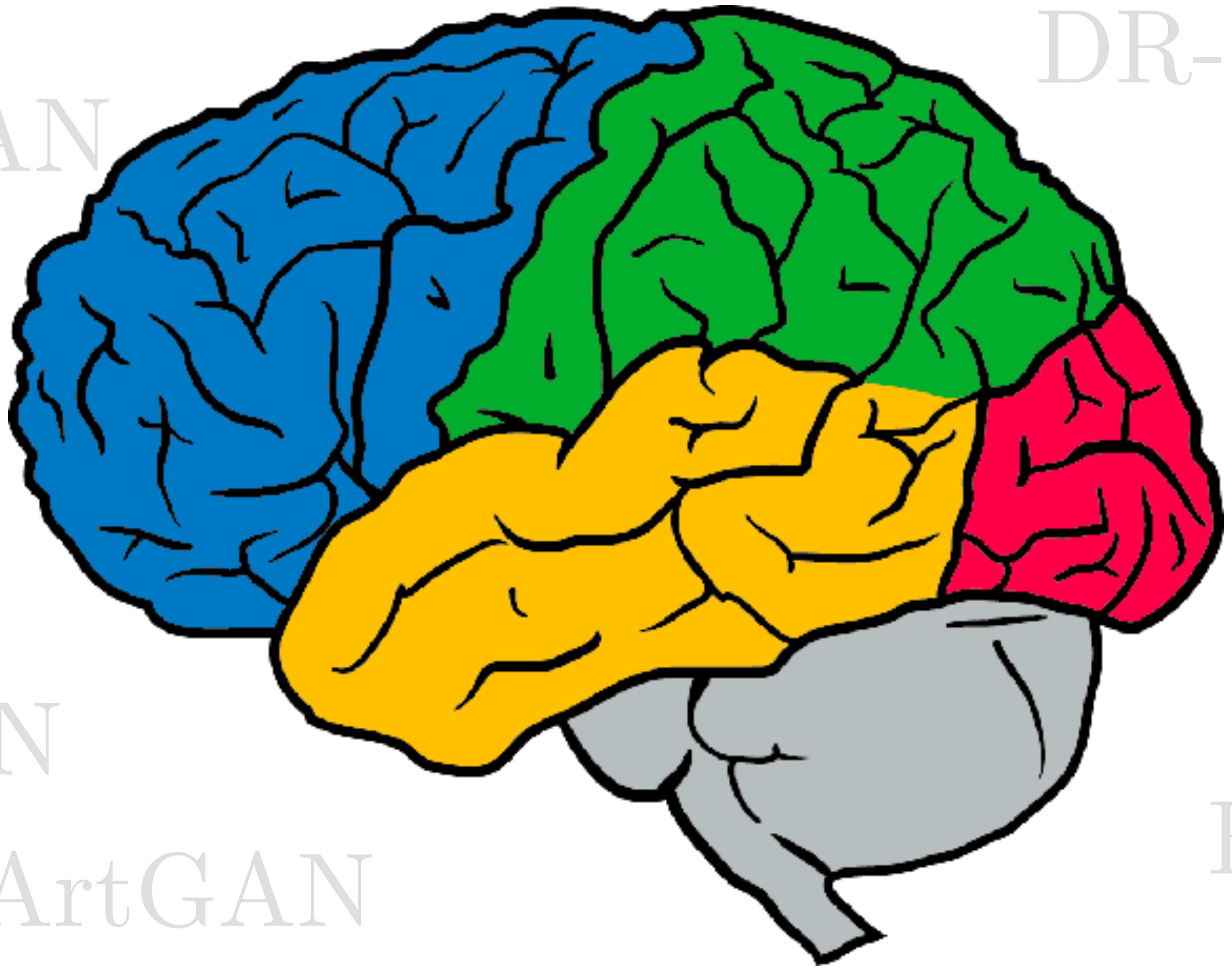
Open Challenges for Improving GANs

Ian Goodfellow, Staff Research Scientist, Google Brain

Seminar at CERN

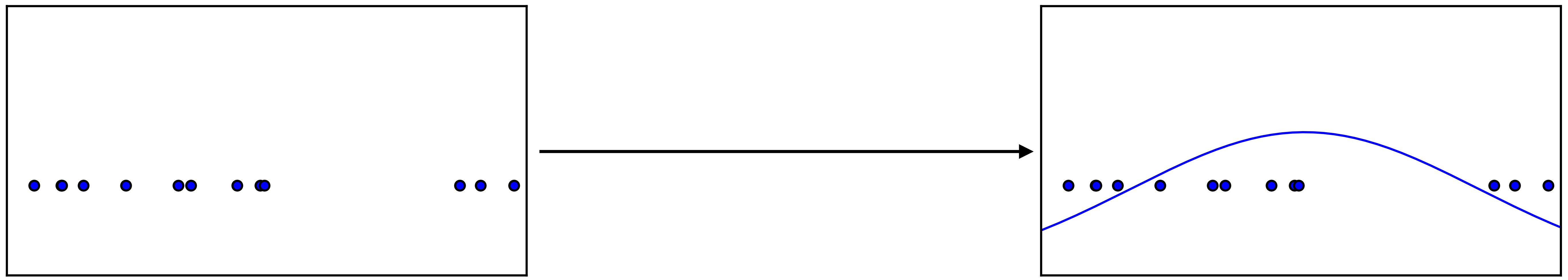
Geneva, 2017-10-29



MedGAN ID-CGAN CoGAN LR-GAN CGAN IcGAN b-GAN LS-GAN AffGAN LAPGAN DiscoGAN MPM-GAN AdaGAN LSGAN InfoGAN CatGAN AMGAN iGAN IAN McGAN MIX+GAN alpha-GAN GMAN MGAN FF-GAN GoGAN BS-GAN C-VAE-GAN C-RNN-GAN DR-GAN DCGAN CCGAN AC-GAN MAGAN 3D-GAN BiGAN GAWWN DualGAN CycleGAN GP-GAN Bayesian GAN AnoGAN WGAN-GP EBGAN DTN MAD-GAN Context-RNN-GAN ALI BEGAN AL-CGAN f-GAN MARTA-GAN ArtGAN MalGAN

Generative Modeling

- Density estimation



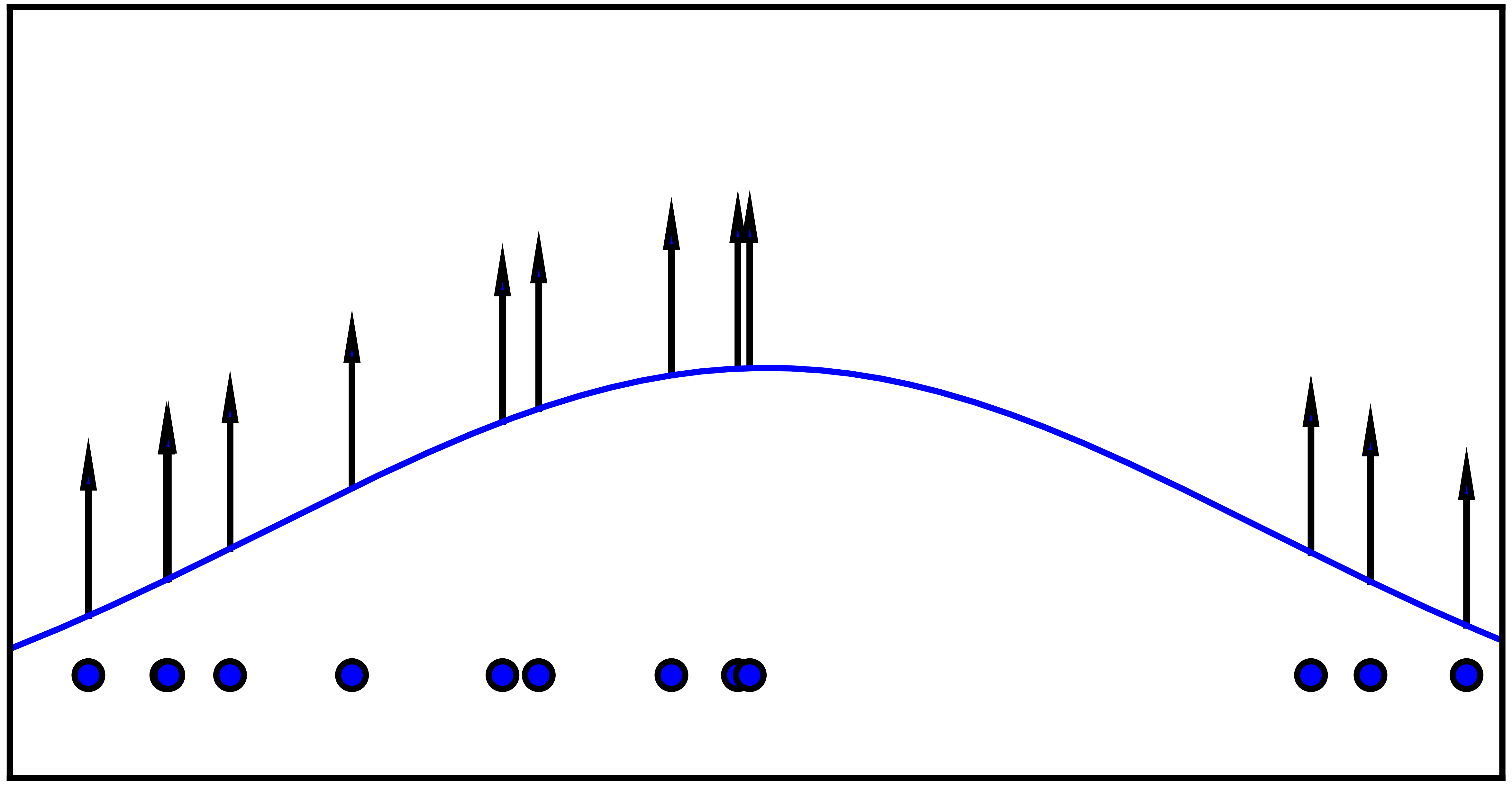
- Sample generation

Training examples

Model samples

(Goodfellow 2017)

Maximum Likelihood



$$\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \mathbb{E}_{x\sim p_{\text{data}}} \log p_{\text{model}}(\boldsymbol{x} \mid \boldsymbol{\theta})$$

Adversarial Nets Framework

$D(\mathrm{x})$ tries to be near 1

↑

Differentiable function D

↑

x sampled from data

D tries to make $D(G(z))$ near 0, G tries to make $D(G(z))$ near 1

↑

D

↑

x sampled from model

↑

Differentiable function G

↑

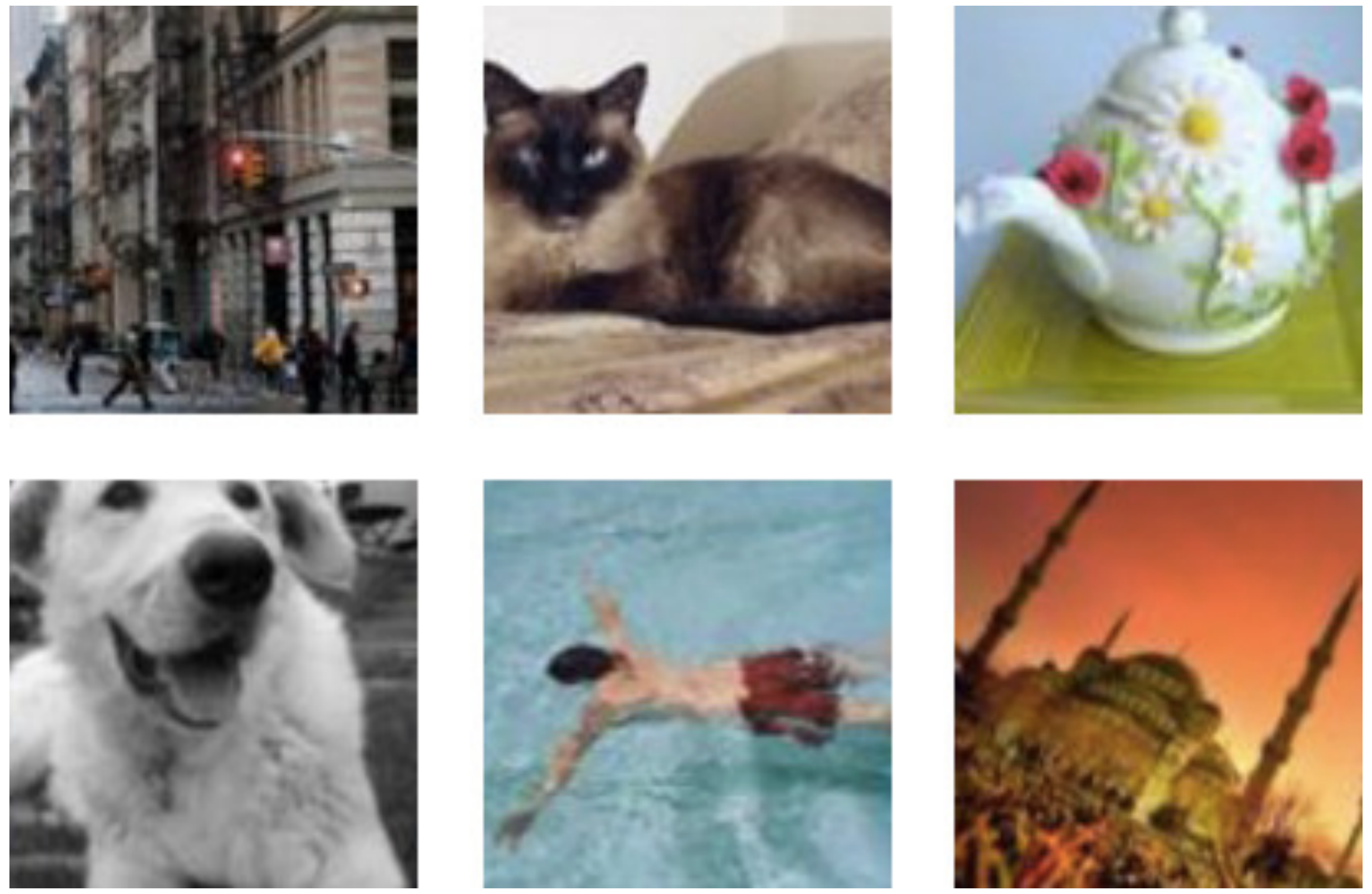
Input noise z

(Goodfellow et al., 2014)

(Goodfellow 2017)

How long until GANs can do this?

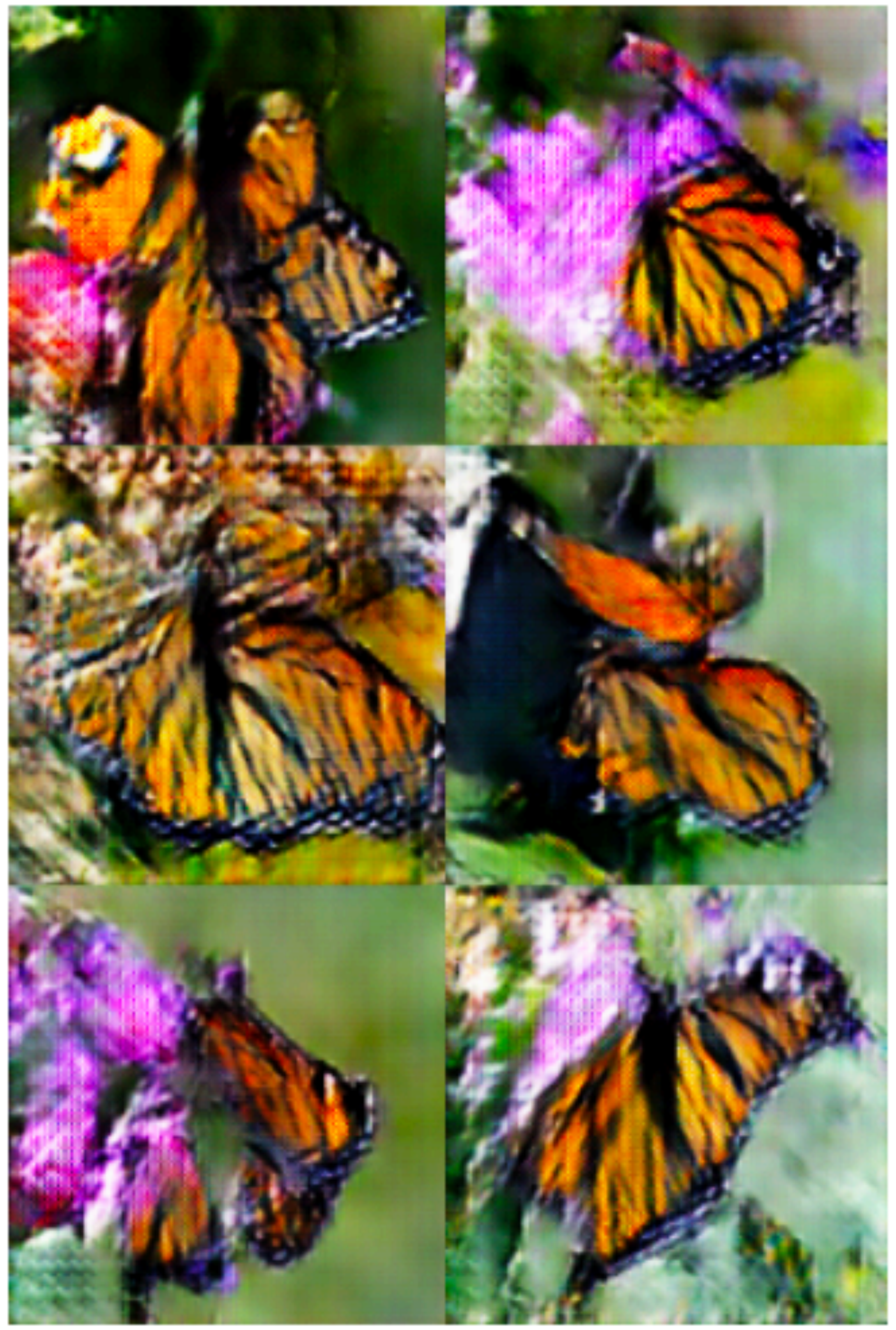
Training examples



Model samples

(Goodfellow 2017)

AC-GANs



monarch butterfly

goldfinch

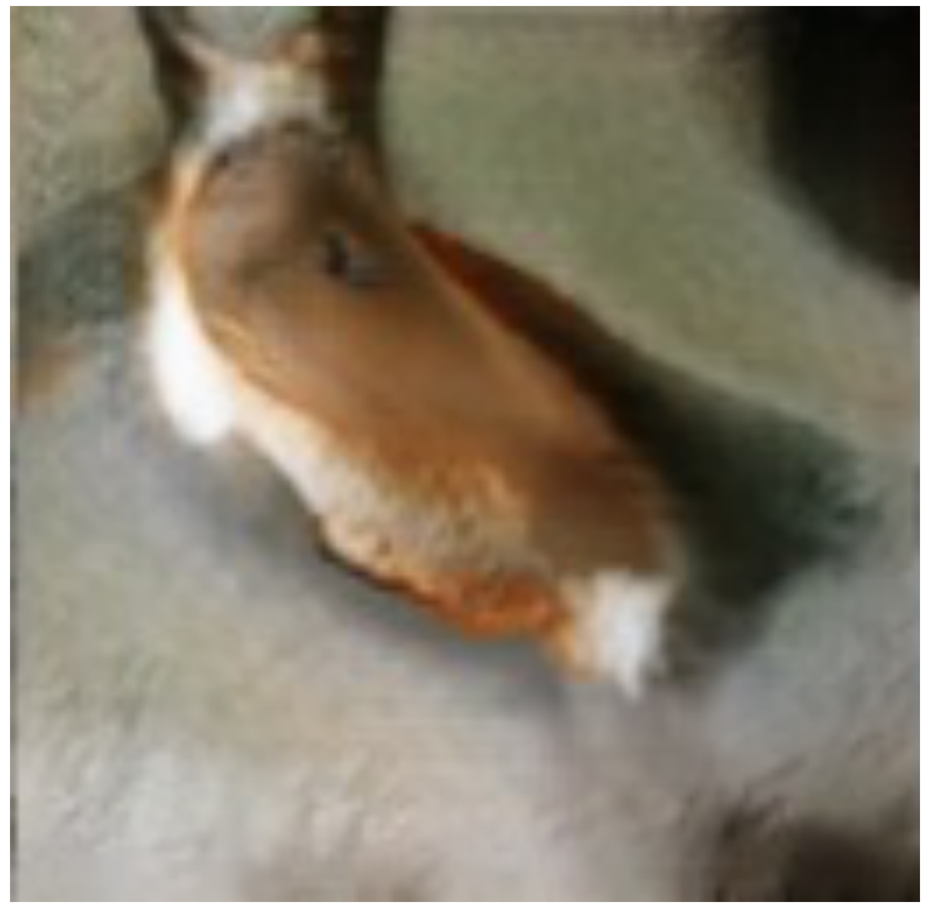
daisy

(Odena et al., 2016)

Minibatch GANs







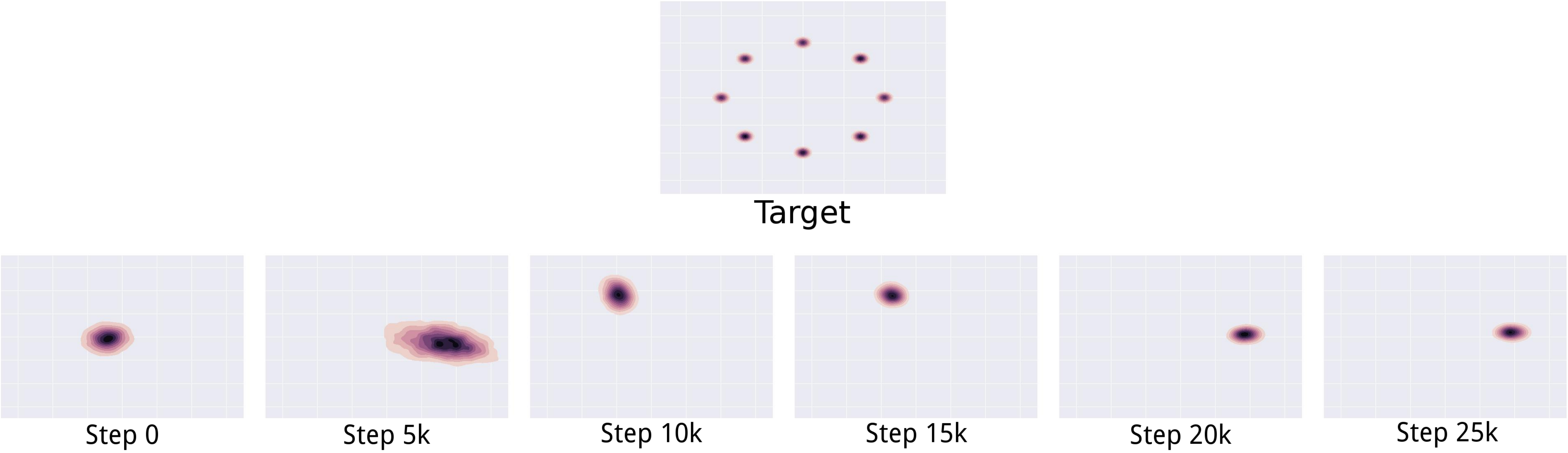
(Salimans et al, 2016)

(Goodfellow 2017)

Mode Collapse

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

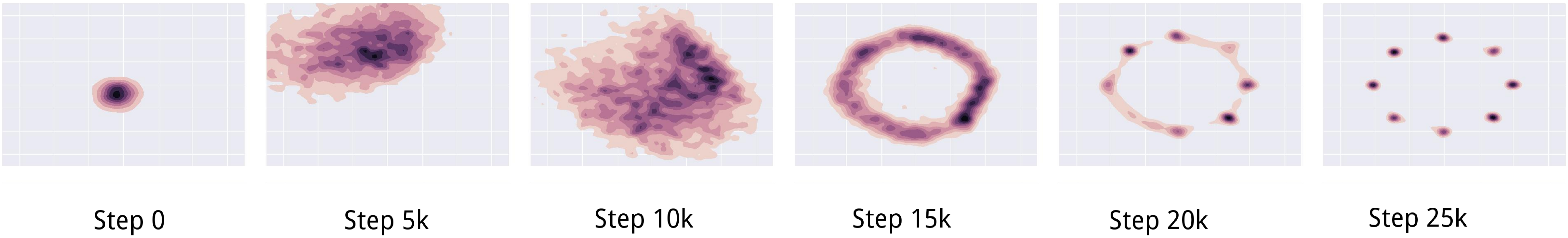
- D in inner loop: convergence to correct distribution
- G in inner loop: place all mass on most likely point



(Metz et al 2016)

Unrolled GANs

- Backprop through k updates of the discriminator to prevent mode collapse:



(Metz et al 2016)

Simple Non-convergence Example

- For scalar x and y, consider the value function:

$$V(x, y) = xy$$

- Does this game have an equilibrium? Where is it?

- Consider the learning dynamics of simultaneous gradient descent with infinitesimal learning rate (continuous time). Solve for the trajectory followed by these dynamics.

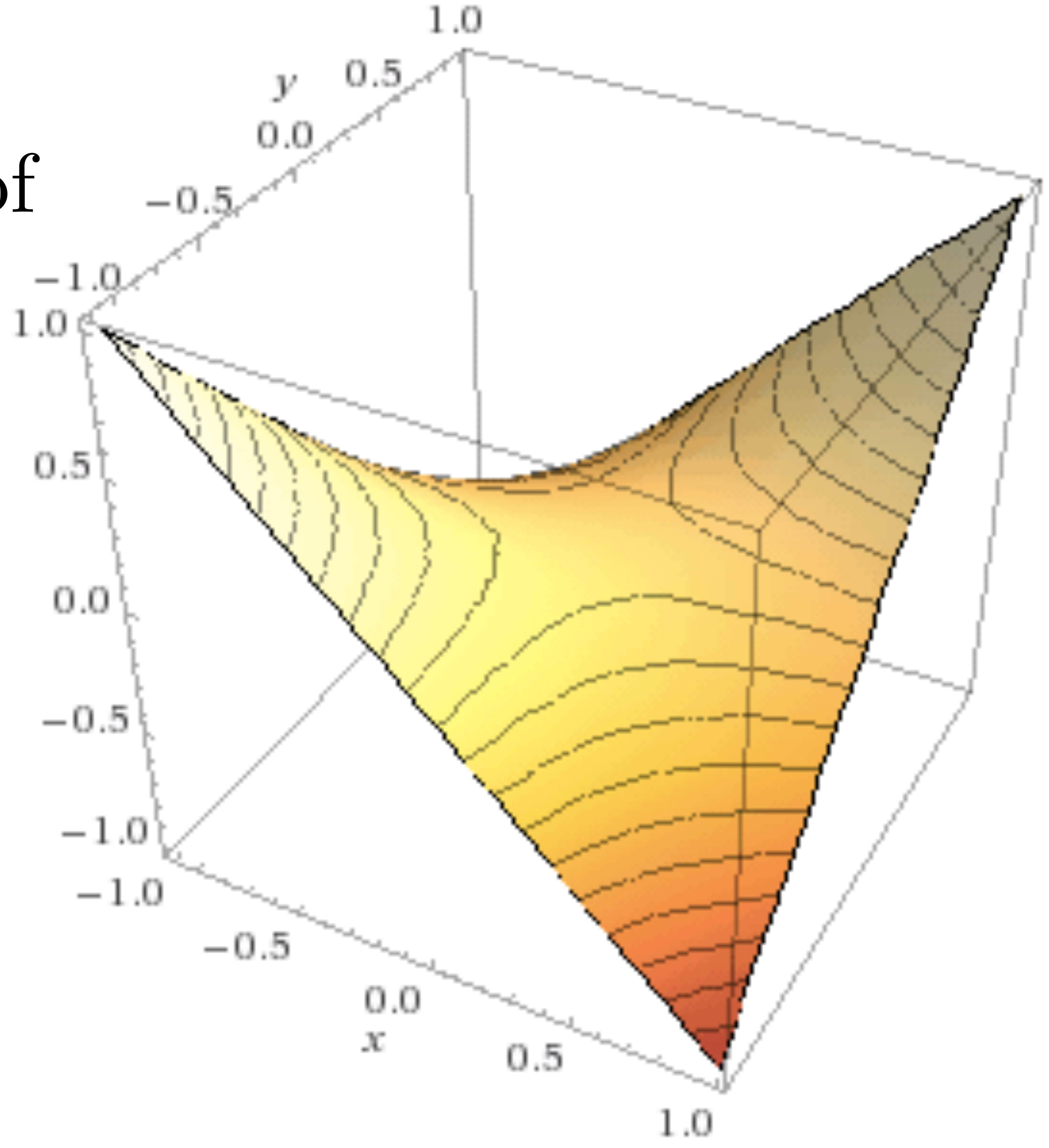
$$\frac{\partial x}{\partial t} = -\frac{\partial}{\partial x} V(x(t), y(t))$$
$$\frac{\partial y}{\partial t} = \frac{\partial}{\partial y} V(x(t), y(t))$$

(Goodfellow 2016)

Solution

This is the canonical example of a saddle point.

There is an equilibrium, at $x = 0$, $y = 0$.



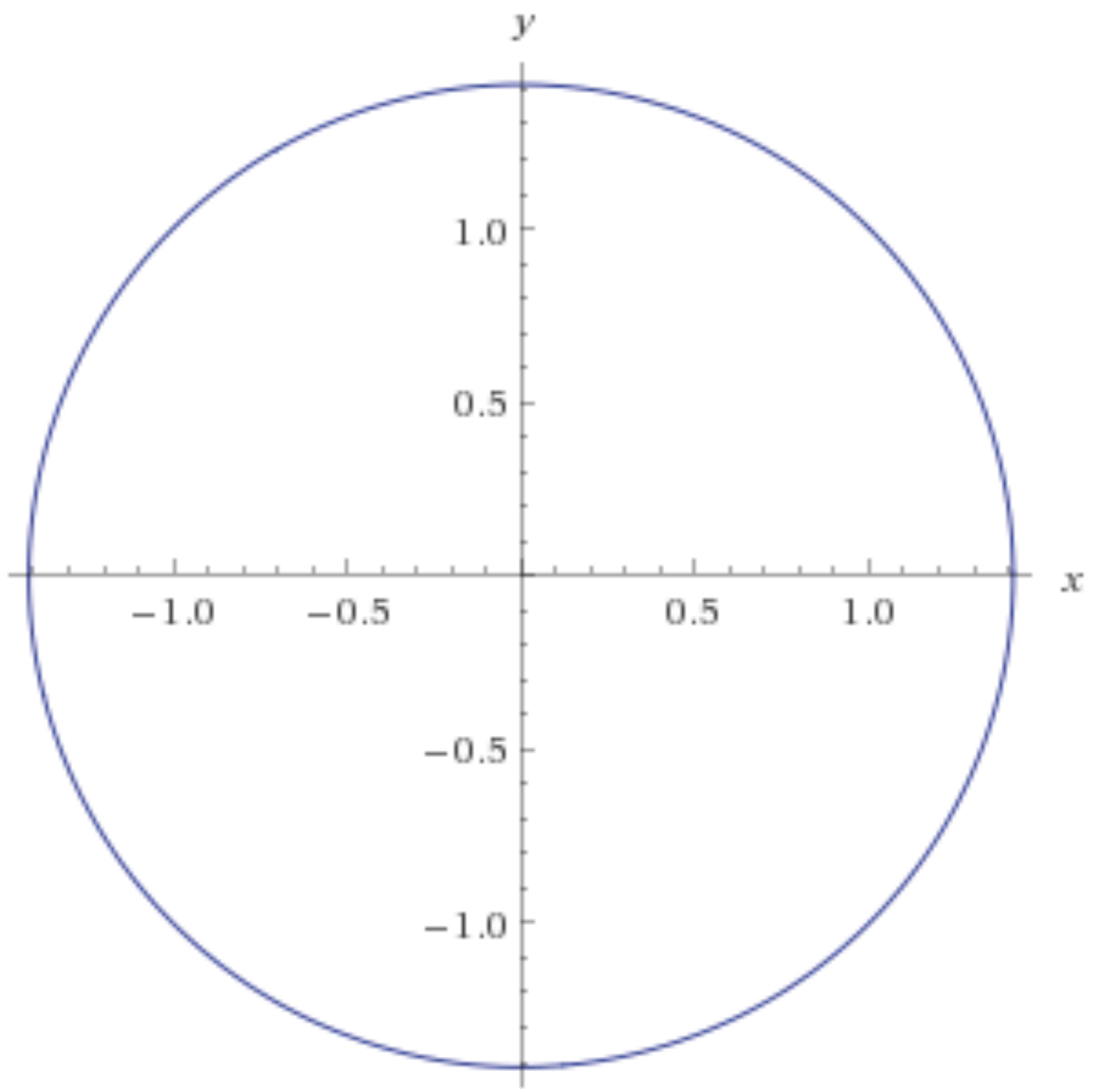
(Goodfellow 2016)

Solution

- The dynamics are a circular orbit:

$$x(t) = x(0)\cos(t) - y(0)\sin(t)$$
$$y(t) = x(0)\sin(t) + y(0)\cos(t)$$

Discrete time gradient descent can spiral outward for large step sizes



(Goodfellow 2016)

Questions

- Does a Nash equilibrium exist for GANs and is it at the right location?

- Does GAN training converge?

- If yes, how quickly?

(Goodfellow 2016)

Does a Nash equilibrium exist, in the right place?

- For some GANs, it's trivial
    - Let z be random one-hot code with m values for task with m training samples

(Goodfellow 2016)

Does a Nash equilibrium exist, in the right place?

- For some GANs, it's trivial
- Original GAN paper: yes... in function space

(Goodfellow 2016)

Does a Nash equilibrium exist, in the right place?

- For some GANs, it's trivial
- Original GAN paper: yes... in function space
- What about for neural nets with a finite number of finite-precision parameters?

(Goodfellow 2016)

Does a Nash equilibrium exist, in the right place?

- For some GANs, it's trivial
- Original GAN paper: yes... in function space
- What about for neural nets with a finite number of finite-precision parameters?
  - Arora et al, 2017: yes... for mixtures
    - Infinite mixture
    - Approximate an infinite mixture with a finite mixture

(Goodfellow 2016)

Does a Nash equilibrium exist, in the right place?

- Still an open question for single net vs single net as used most often in practice

(Goodfellow 2016)

Questions

- Does a Nash equilibrium exist for GANs and is it at the right location?
  - In many cases, yes
  - Still open for most practical cases
- Does GAN training converge?
- If yes, how quickly?

(Goodfellow 2016)

Does GAN training converge?

- Goodfellow et al, 2014:
  - Yes, if...
    - You optimize in function space
    - You optimize D fully in the inner loop, take small steps on G in the outer loop

Does GAN training converge?

- Goodfellow et al, 2014:
  - Yes, if...
    - You optimize in function space
    - You optimize D fully in the inner loop, take small steps on G in the outer loop
  - But in practice, *simultaneous SGD* works much better!

(Goodfellow 2016)

Does GAN training converge?

- Goodfellow et al, 2014: yes, in function space, fully optimizing D
- Kodali et al 2017: yes, in function space, simultaneous SGD converges

Does GAN training converge?

- Goodfellow et al, 2014: yes, in function space, fully optimizing D
- Kodali et al 2017: yes, in function space, simultaneous SGD converges
- but... we can't optimize over function space in a digital computer

Continuous games

- Ratliff et al, 2013
- Two players, with two sets of parameters

$$\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}$$

- Two costs

$$J^{(1)}\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\right), J^{(2)}\left(\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}\right)$$

- Each player controls *only their own cost*

Does simultaneous SGD converge?

- Regard SGD as a nonlinear dynamical system
- Examine the Jacobian of the concatenated gradients with respect to the concatenated parameters:

$$\boldsymbol{M} = \begin{array}{c|c|c} & \boldsymbol{g}^{(1)} & \boldsymbol{g}^{(2)} \\ \hline \boldsymbol{\theta}^{(1)} & \boldsymbol{H}^{(1)} & \nabla_{\boldsymbol{\theta}^{(1)}} g^{(2)} \\ \hline \boldsymbol{\theta}^{(2)} & \nabla_{\boldsymbol{\theta}^{(2)}} g^{(1)} & \boldsymbol{H}^{(2)} \end{array}$$

(Ratliff et al 2013)

Does simultaneous SGD converge?

- Ratliff et al 2013:
  - If the real part of each eigenvalue of M is positive, then yes (exponentially stable fixed point)

Intuition for the Jacobian

How firmly does player 1 want to stay in place?

How much can player 1 dislodge player 2?

	$\boldsymbol{g}^{(1)}$	$\boldsymbol{g}^{(2)}$
$\boldsymbol{\theta}^{(1)}$	$\boldsymbol{H}^{(1)}$	$\nabla_{\boldsymbol{\theta}^{(1)}} g^{(2)}$
$\boldsymbol{\theta}^{(2)}$	$\nabla_{\boldsymbol{\theta}^{(2)}} g^{(1)}$	$\boldsymbol{H}^{(2)}$

How much can player 2 dislodge player 1?

How firmly does player 2 want to stay in place?

(Goodfellow 2016)

What happens for GANs?

$$
\begin{array}{c|c|c|c}
 & & D & G \\
 & & \boldsymbol{g}^{(1)} & \boldsymbol{g}^{(2)} \\
\hline
D & \boldsymbol{\theta}^{(1)} & \boldsymbol{H}^{(1)} & \nabla_{\boldsymbol{\theta}^{(1)}} g^{(2)} \\
\hline
G & \boldsymbol{\theta}^{(2)} & \nabla_{\boldsymbol{\theta}^{(2)}} g^{(1)} & \boldsymbol{H}^{(2)}
\end{array}
$$

All zeros!

The optimal discriminator is constant.

Locally, the generator does not have any "retaining force"

(Goodfellow 2016)

What happens for GANs?

- The real part of some eigenvalues of M can be exactly 0

- The result from Ratliff et al 2013 does not apply

(Goodfellow 2016)

What happens for GANs?

- The real part of some eigenvalues of M can be exactly 0
- The result from Ratliff et al 2013 does not apply
- So do they fail to converge?

(Goodfellow 2016)

What happens for GANs?

- The real part of some eigenvalues of M can be exactly 0
- The result from Ratliff et al 2013 does not apply
- Nagarajan et al 2017:
  - Extend the analysis to handle the 0!
  - The equilibrium is stable!

Does GAN training converge?

- Goodfellow et al, 2014: yes, in function space, fully optimizing D
- Kodali et al 2017: yes, in function space, simultaneous SGD converges
- Nagarajan et al 2017: yes, even in parameter space of neural nets, if you start near the equilibrium

Does GAN training converge?

- Goodfellow et al, 2014: yes, in function space, fully optimizing D
- Kodali et al 2017: yes, in function space, simultaneous SGD converges
- Nagarajan et al 2017: yes, even in parameter space of neural nets, if you start near the equilibrium
- Heusel et al 2017: yes, if the generator learning rate decays relative to the discriminator learning rate

Questions

- Does a Nash equilibrium exist for GANs and is it at the right location?
  - In many cases, yes
  - Still open for most practical cases
- Does GAN training converge?
  - In many *quite realistic* cases, yes
- If yes, how quickly?

(Goodfellow 2016)

Numerics of GANs

- Mescheder et al, 2013
- The zero real part of the eigenvalues of M did not go away
- SGD may converge, but it converges slowly
- Imaginary part of eigenvalues is large too

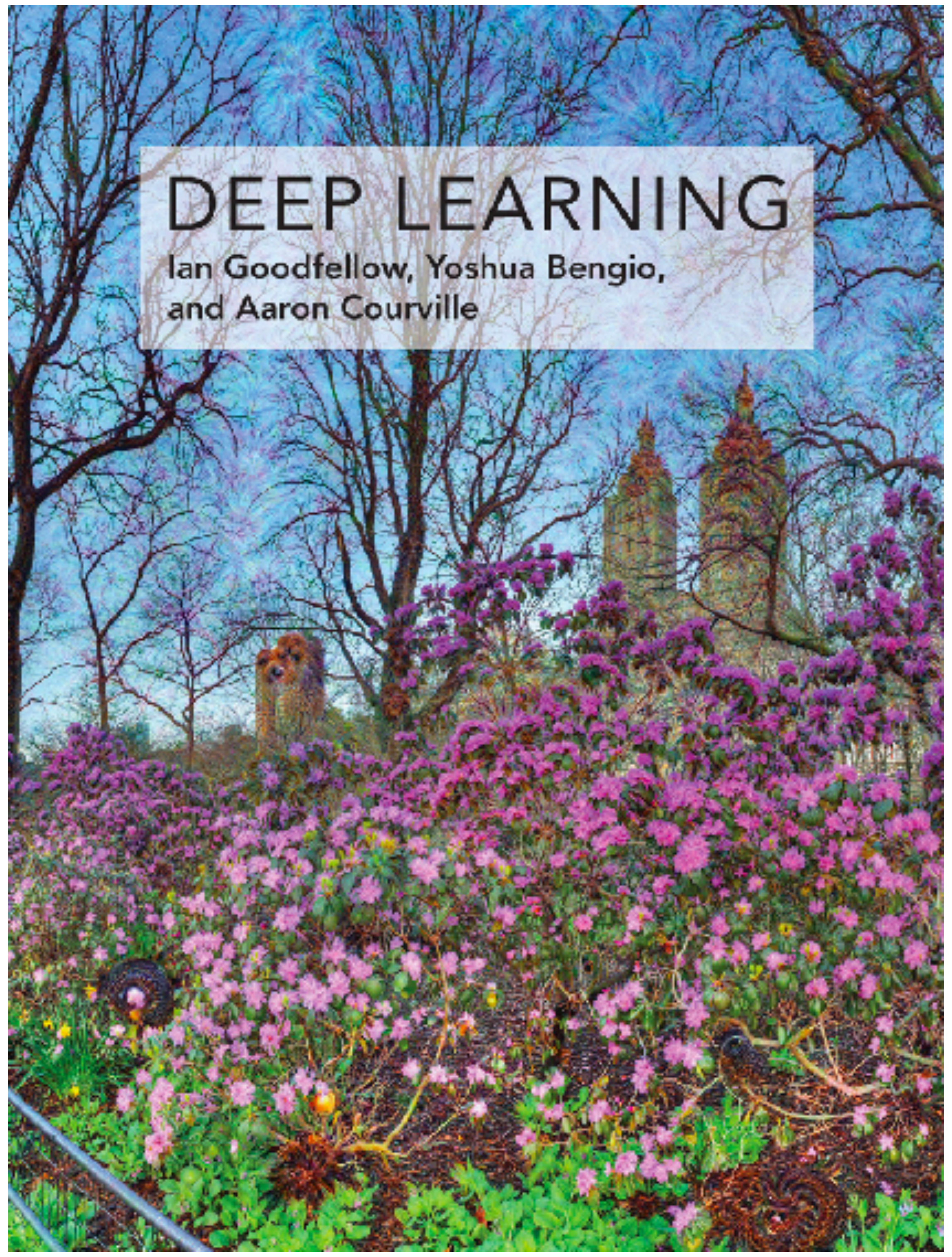
How to speed up convergence?

- Penalize the gradient?
  - Nagaran et al and Mescheder et al both suggest it
- Improves numerics
- Original equilibrium remains an equilibrium
- Non-equilibrium saddle points of original game become equilibria
  - (For example: zero out some layer of the discriminator... now the gradient is gone)

Questions

- Does a Nash equilibrium exist for GANs and is it at the right location?
  - In many cases, yes
  - Still open for most practical cases
- Does GAN training converge?
  - In many *quite realistic* cases, yes
- If yes, how quickly?
  - Slowly
  - Open problem: how to make convergence *fast*, without adding equilibria in the wrong place

Further Resources



www.deeplearningbook.org

(Goodfellow 2017)