Do GANs always have Nash equilibria?

Farzan Farnia, Asu Ozdaglar

Massachusetts Institute of Technology

Generative Adversarial Networks

- GANs learn the distribution of data via a zero-sum game between:
  - Generator G mimicking the data distribution,
  - Discriminator D distinguishing G's samples from real data.
- GANs are commonly formulated through a minimax problem:

$$\min_{G \in \mathcal{G}} \max_{D \in \mathcal{D}} \underbrace{V(G, D) = \mathbb{E}\left[\log\left(D(\mathbf{X})\right)\right] + \mathbb{E}\left[\log\left((1 - D(G(\mathbf{Z}))\right)\right]}_{d(P_{\mathbf{X}}, P_{G(\mathbf{Z})})\text{ : distance between } P_{\mathbf{X}} \text{ and } P_{G(\mathbf{Z})}}$$

Optimality in GAN Minimax Optimization

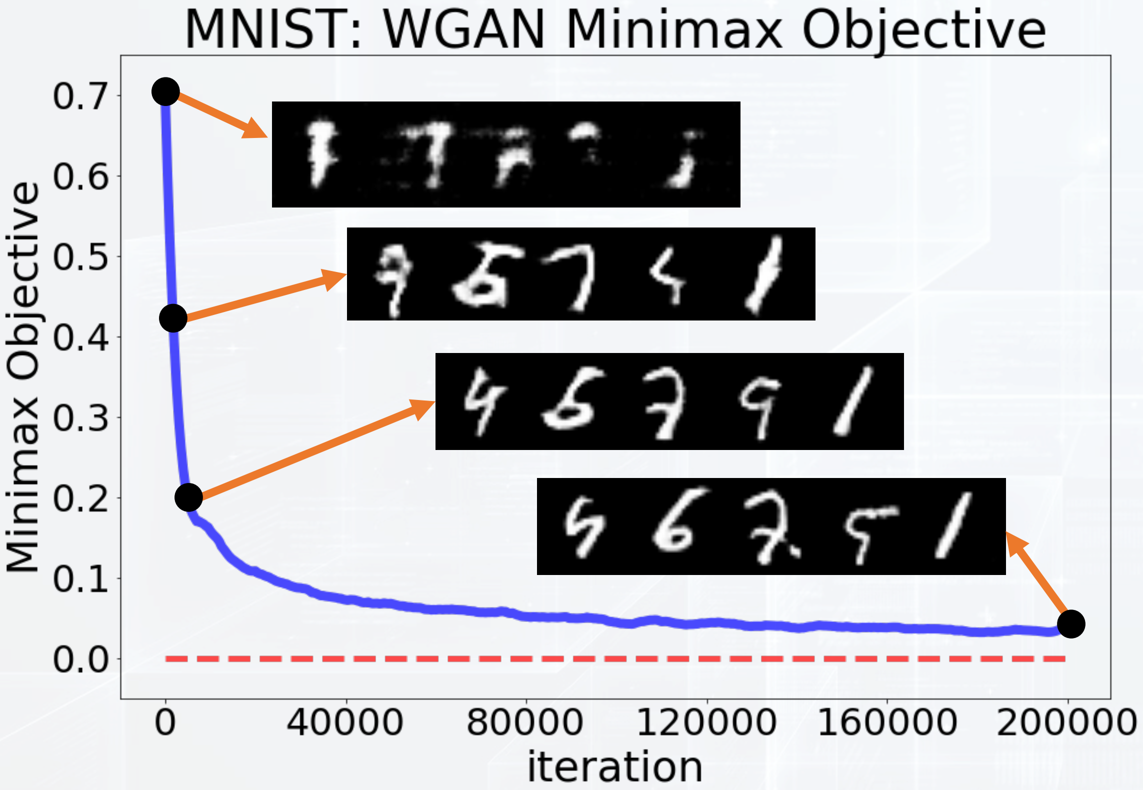
- What is the proper notion of optimality in GAN minimax problems?
  - Nash equilibrium (NE) of the underlying game:

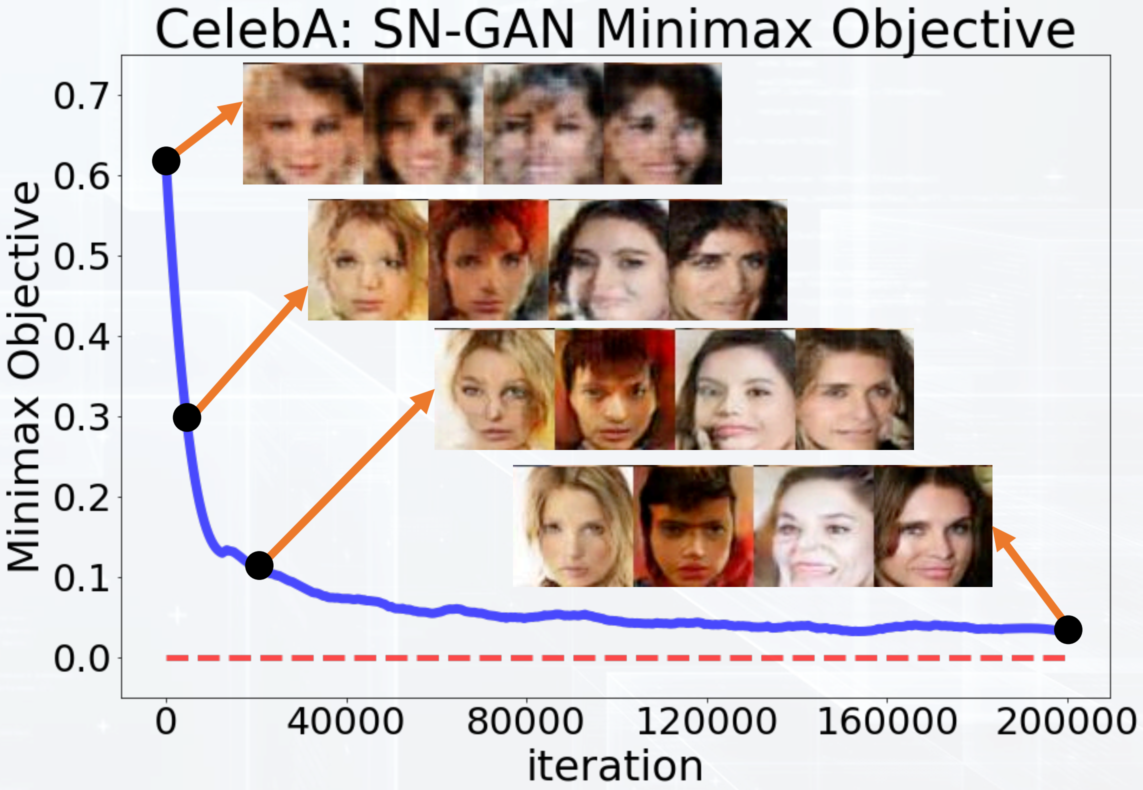
$$\forall G, D: \;\; V(G^*, D) \leq V(G^*, D^*) \leq V(G, D^*)$$

- Does Nash equilibrium exist for GANs?
  - Yes under the realizability assumption: $P_{G^*(\mathbf{Z})} = P_{\mathbf{X}}$
  - G^* paired with a constant D gives a NE.

Realizability in Standard GANs

- Do standard GANs produce the exact data distribution?
    - No, the minimax objective does not usually reach zero.





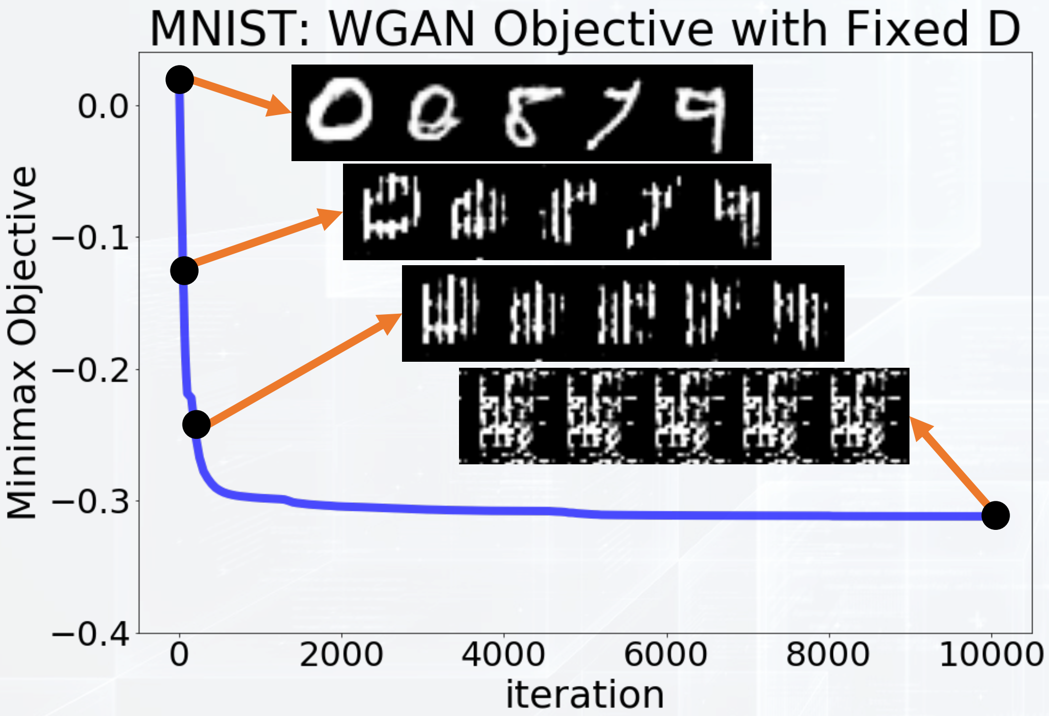
Realizability in Standard GANs

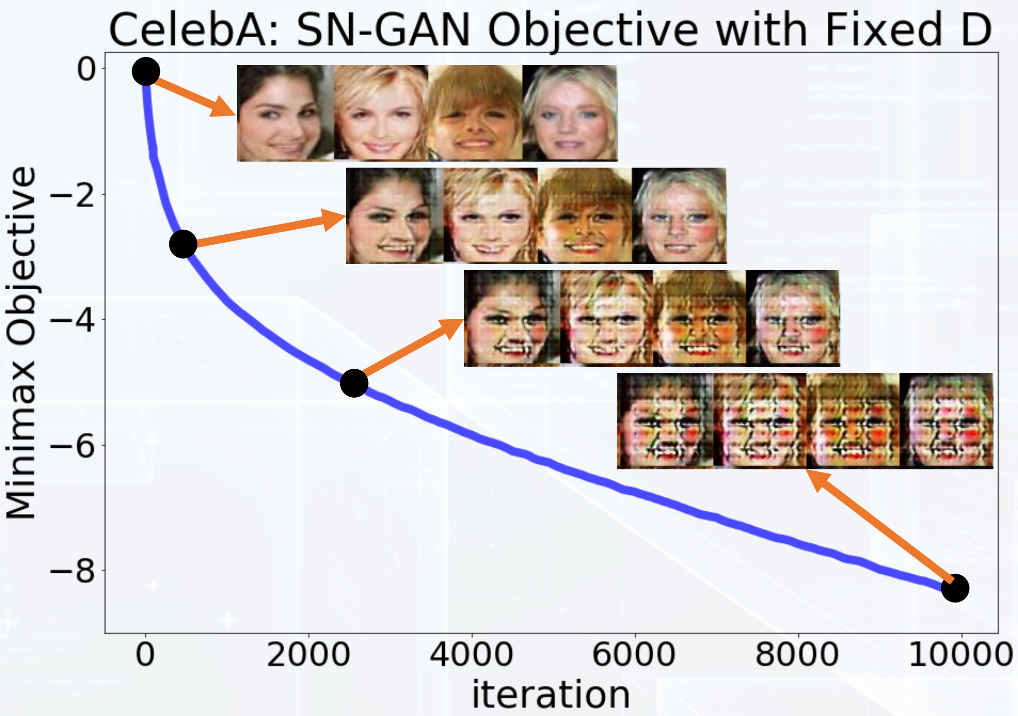
- Then, are the solutions found (local) Nash equilibria?
  - **Experimen**t: Fix the trained D and keep optimizing G

- More empirical evidence in recent related works:
  - Berard et al., *"A closer look at the Optimization Landscapes of GANs"*, ICLR 2020
  - Schafer et al., *"Implicit competitive regularization in GANs"*, ICML 2020

Realizability in Standard GANs

- Then, are the solutions found (local) Nash equilibria?
  - **Experiment**: Fix the trained D and keep optimizing G





Nash Equilibrium in Non-realizable GANs

- Do Nash equilibria exist in non-realizable GAN problems?

- **Theorem:** Suppose $\mathbf{X} \sim \mathcal{N}(\mathbf{0}, \Sigma)$, $\sigma_{\max}(\Sigma) > 1$. Consider a regularized linear $G(\mathbf{Z}) = A\mathbf{Z} + \mathbf{b}$, $\mathbf{Z} \sim \mathcal{N}(\mathbf{0}, I)$, $\sigma_{\max}(A) \leq 1$, $\|\mathbf{b}\|_2 \leq 1$. Then,
    - Vanilla GAN and f-GANs with **unconstrained *D*** have no NEs.
    - Wasserstein GAN with **1-Lipschitz *D*** has no NEs.
    - 2-Wasserstein GAN has no NEs with **c-concave *D*** and no **local** NEs with **quadratic** $D(\mathbf{x}) = \mathbf{x}^T \Lambda \mathbf{x} + \boldsymbol{\gamma}^T \mathbf{x}$.

Stackelberg Equilibrium Exists in GAN games

- Consider the equilibria of Stackelberg GAN game:

$$G^* \in \operatorname*{argmin}_{G \in \mathcal{G}} \left\{ \max_{D \in \mathcal{D}} V(G, D) \right\}, \quad D^*(G^*) \in \operatorname*{argmax}_{D \in \mathcal{D}} V(G^*, D)$$

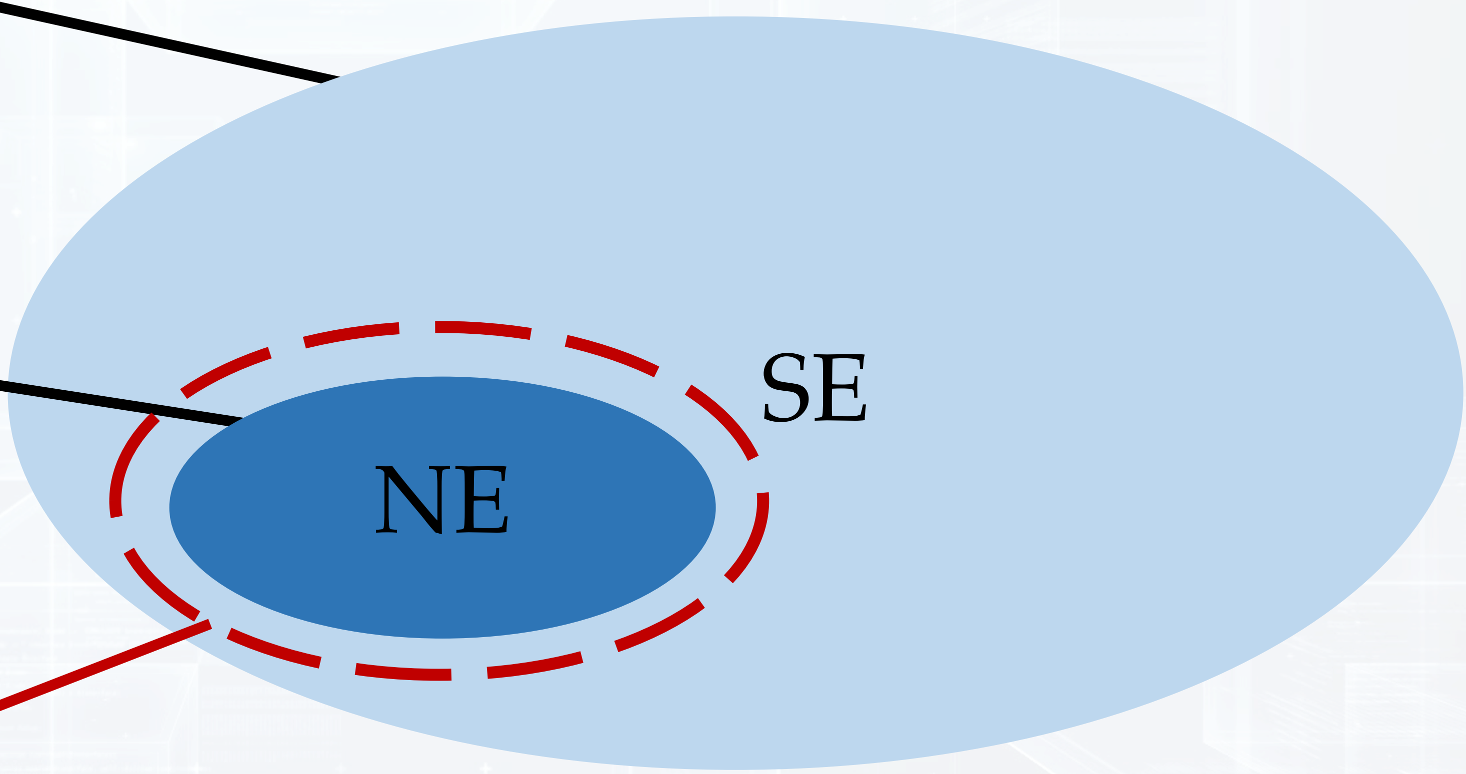
- Stackelberg equilibrium will exist under mild assumptions but is in general less stable than a Nash equilibrium.
  - Stable limit points of $1, \infty$-gradient descent ascent (GDA) vs. 1,1-GDA.
  - Jin et al., *"What is Local Optimality in Non-convex Non-concave Minimax Optimizaion?"*, ICML 2020.

Stackelberg Equilibria vs. Nash Equilibria

Stackelberg Equilibria
guaranteed to exist 🙂
lower stability 🙁

Nash Equilibria
may not exist 🙁
higher stability 🙂

Non-empty & stable?



Proximal Equilibria: Spectrum between Nash and Stackelberg Equilibria

- For $\|D\|, \lambda$, we define the proximal objective:

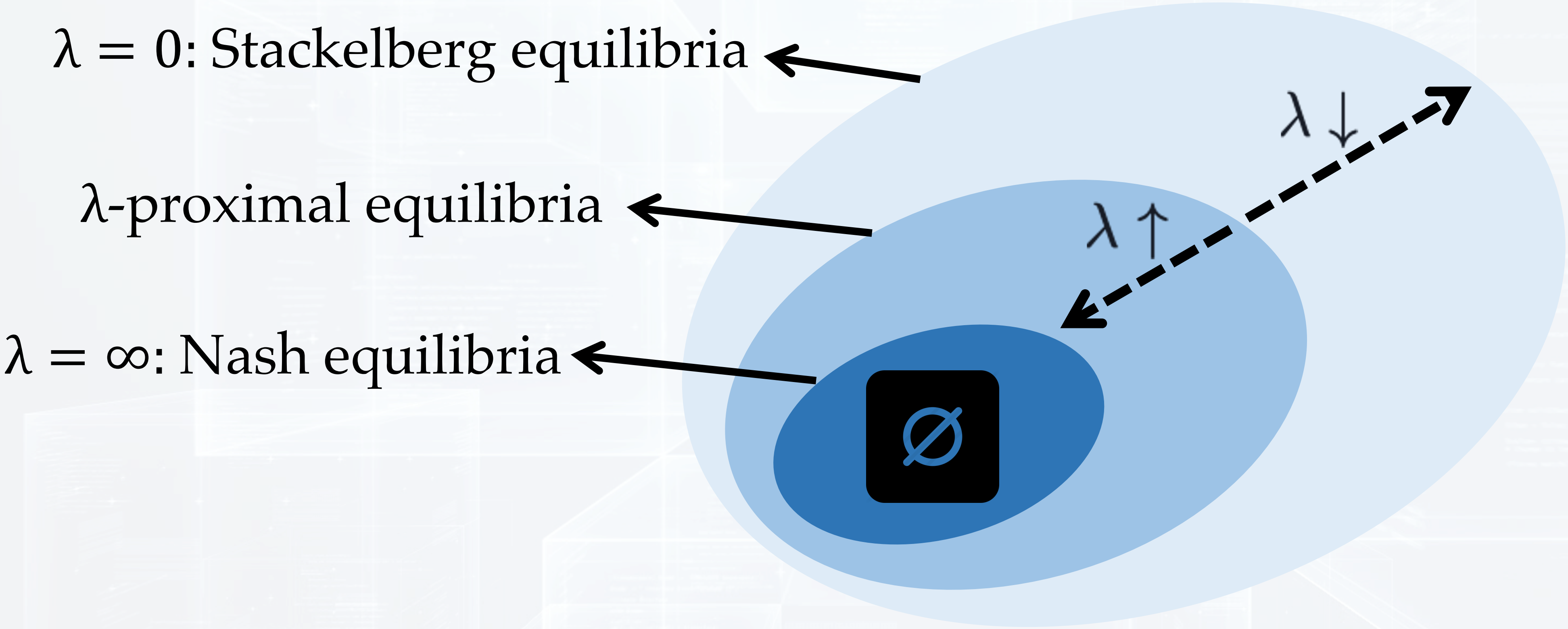
$$V_{\lambda}^{\text{prox}}(G, D) := \max_{\widetilde{D} \in \mathcal{D}} V(G, \widetilde{D}) - \lambda \|D - \widetilde{D}\|^2.$$

- We define $V_{\lambda}^{\text{prox}}(G, D)$'s Nash equilibria as *$\lambda$-proximal equilibria*.

- Nested property of proximal equilibria:

$$\lambda_1 \leq \lambda_2 \Rightarrow \text{PE}(\lambda_2) \subseteq \text{PE}(\lambda_1)$$

Proximal Equilibria: Spectrum between Nash and Stackelberg Equilibria



Question: Does a λ-proximal equilibrium exist for λ>0?

Proximal Equilibria in Wasserstein GANs

- **Theorem:** Consider the W2GAN problem minimizing the following optimal transport cost:

$$W_2\big(P_{\mathbf{X}}, P_{G(\mathbf{Z})}\big) := \min_{\Pi(P_{\mathbf{X}}, P_{G(\mathbf{Z})})} \mathbb{E}\left[\beta \|\mathbf{X} - \mathbf{X}'\|_2^2\right]$$

  Then, the W2GAN problem has a $1/\beta$-proximal equilibrium w.r.t.

$$\|D\|_{\text{Sobolev}} = \sqrt{\mathbb{E}_{P_{\mathbf{X}}}\left[\|\nabla D(\mathbf{X})\|_2^2\right]}$$

- We also prove a similar result for standard WGANs.

Proximal Equilibria in Wasserstein GANs: Proof

- Brenier's theorem from optimal transport theory implies for optimal D_G

$$W_2\big(P_{\mathbf{X}}, P_{G(\mathbf{Z})}\big) \; = \; \frac{1}{\beta}\,\mathbb{E}_{P_{\mathbf{X}}}\Big[\big\|\nabla D_G(\mathbf{X})\big\|_2^2\Big]$$

- We reformulate the W2GAN problem as

$$\min_{G\in\mathcal{G}} W_2\big(P_{\mathbf{X}}, P_{G(\mathbf{Z})}\big) \; \equiv \; \min_{D_G\in\mathcal{D}_{\mathcal{G}}} \; \frac{1}{\beta}\,\mathbb{E}_{P_{\mathbf{X}}}\Big[\big\|\nabla D_G(\mathbf{X})\big\|_2^2\Big]$$

convex set

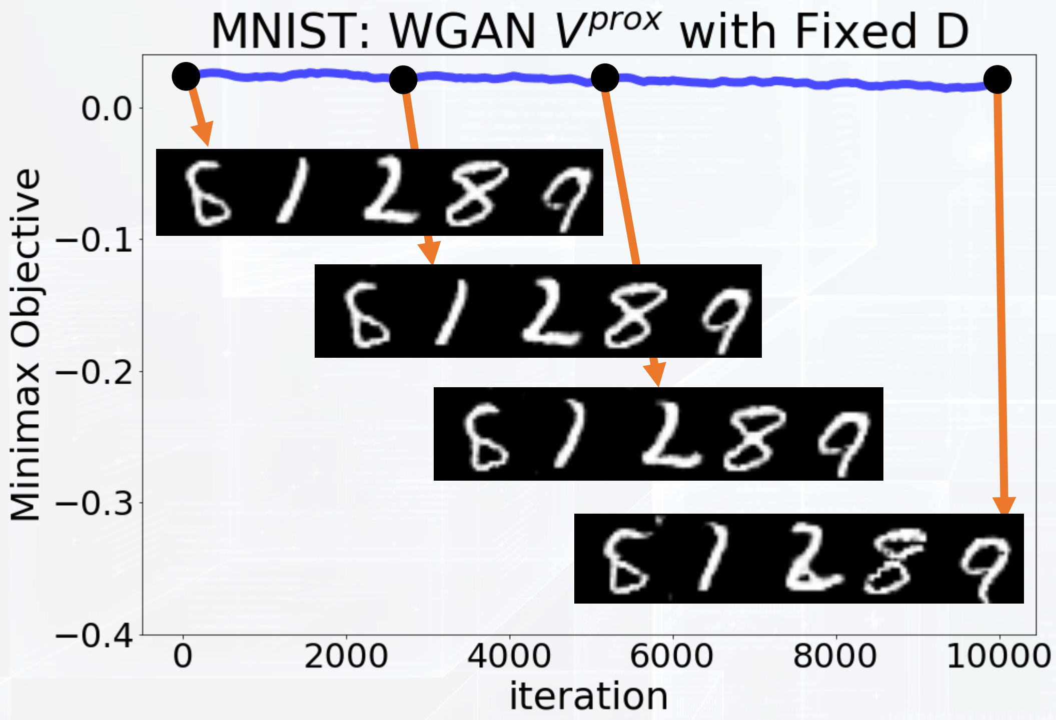
Strongly-convex w.r.t. the Sobolev norm

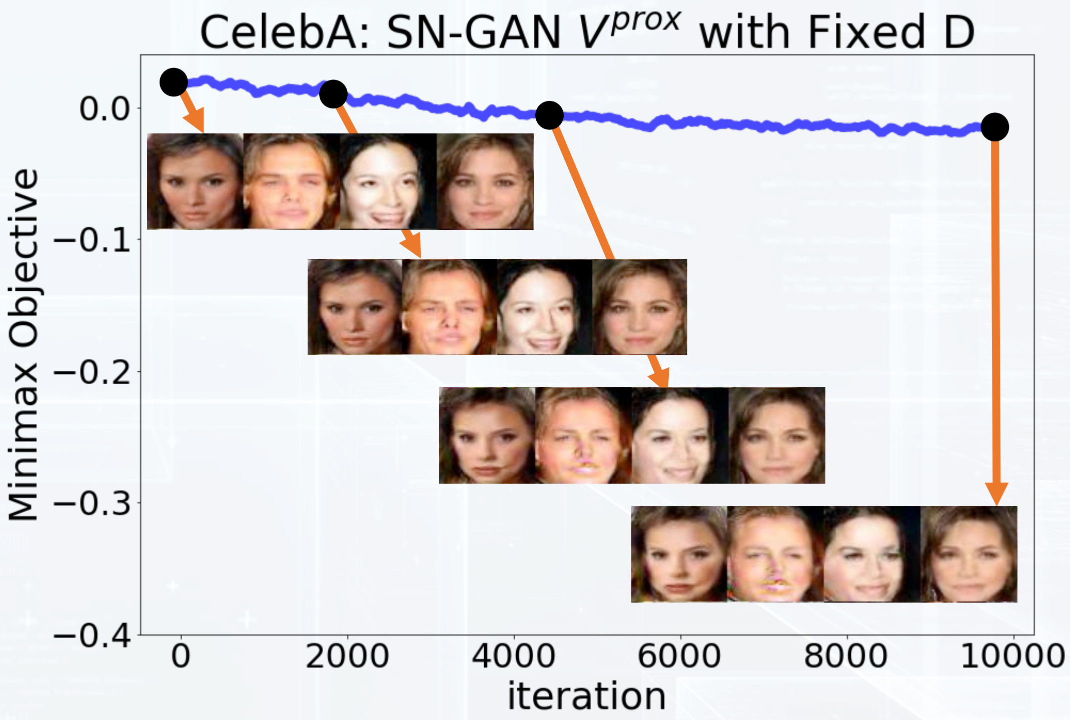
- Minimizing a strongly-convex function over a convex set implies

$$V_{1/\beta}^{\text{prox}}(G^*, D_{G^*}) = W_2\big(P_{\mathbf{X}}, P_{G^*(\mathbf{Z})}\big) \leq W_2\big(P_{\mathbf{X}}, P_{G(\mathbf{Z})}\big) - \frac{1}{\beta}\|D_G - D_{G^*}\|_{\text{Sobolev}}^2 \leq V_{1/\beta}^{\text{prox}}(G, D_{G^*})$$

Proximal Equilibrium in Standard GANs

- Are the solutions found (local) proximal equilibria?
  - **Experiment:** Fix the final trained D and optimize $V_\lambda^{\text{prox}}(G, D)$





Proximal Training via Optimizing the Proximal Objective

- **Observation:** λ-Proximal equilibria are stable limit points of every alternating gradient method in solving:

$$\min_{G \in \mathcal{G}} \max_{D \in \mathcal{D}} V_\lambda^{\text{prox}}(G, D)$$

- **Proximal training:** Apply alternating gradient methods to optimize $V_\lambda^{\text{prox}}(G, D)$ instead of the original $V(G, D)$.

Proximal Training vs. Regular Training

SN-GAN: Regular Training

SN-GAN: Proximal Training

Proximal Training vs. Regular Training

SN-GAN: Regular Training

Inception score: **5.62 ± 0.23**

SN-GAN: Proximal Training

Inception score: **6.12 ± 0.22**

Thank you for your attention!

arXiv link: arxiv.org/abs/2002.09124