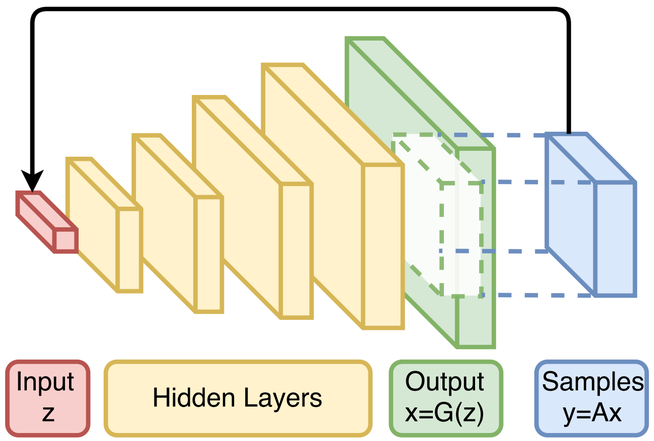
Invertibility of Convolutional Generative Networks from Partial Measurements

Fangchang Ma*, Ulas Ayaz*, Sertac Karaman {* Indicates equal contribution}

I. Problem



Find the latent code z from y=Ax (partial measurements of the network output x)

II. Background and Motivation

- Application in **image inpainting**, **depth completion** problems: given only partial measurements, recover the entire image
- Theoretical understanding of the mapping between input latent space and output pixel space for generative networks (**GAN**)

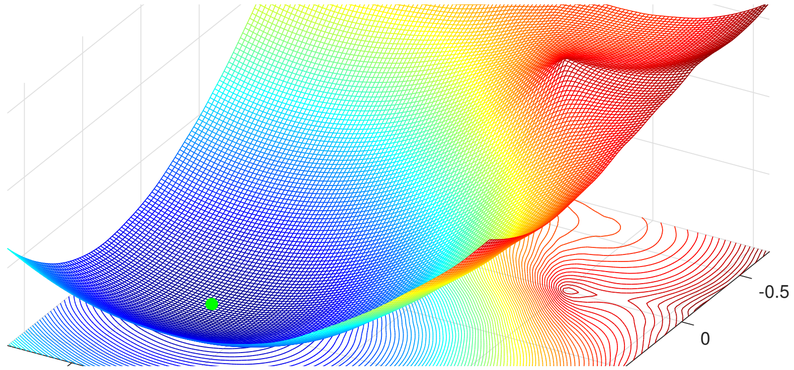
III. Algorithm

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \frac{1}{2}\|\mathbf{y} - \mathbf{A} \cdot \mathbf{G}(\mathbf{z})\|^2$$

Solve the **highly non-convex** minimization problem with fast and simple gradient descent algorithms, with guarantees

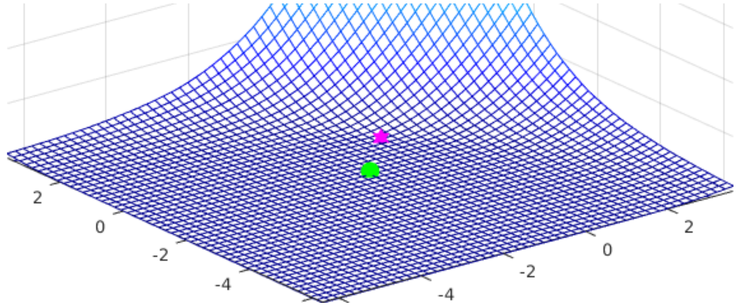
IV. Main Result (for 2-layer networks)

Assume the transposed convolutional-ReLU network is sufficiently expansive, and the kernel weights follow a zero-mean Gaussian distribution. Then with high probability there exists a gradient direction everywhere except for two neighborhoods.



V. Implication for "Mode Collapse"

The assumptions serve as a sufficient condition for **one-to-one mapping**: avoid mode collapse in the training of GAN



VI. Experiments (multi-layer networks)