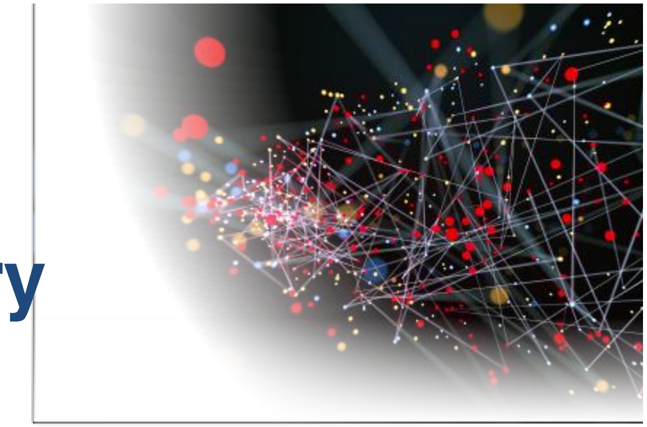
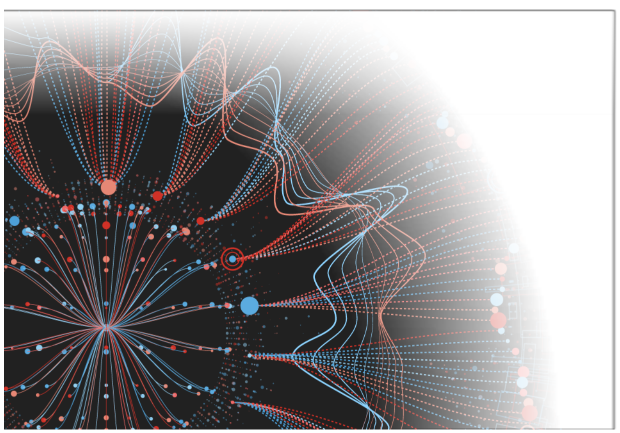
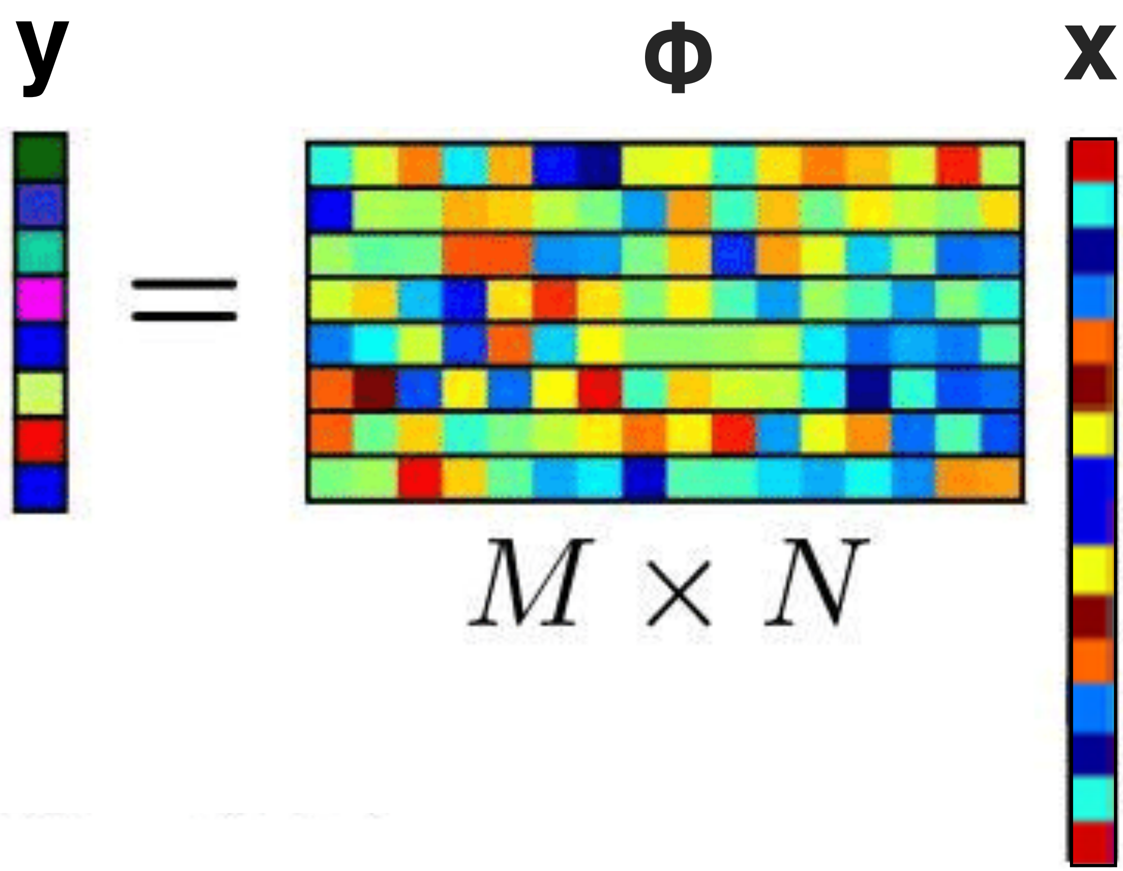


Fast Compressive Sensing Recovery Using Generative Models with Structured Latent Variables



Shaojie (Kyle) Xu, Sihan Zeng
Prof. Justin Romberg
Georgia Institute of Technology

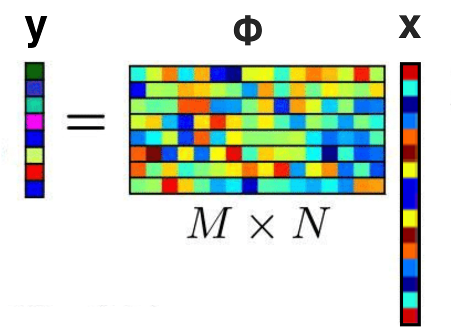
Georgia Tech


y
=
Φ
x
M × N

Agenda

- Solve inverse problems with priors
- Incorporate generative models into alternating direction method of multipliers (ADMM) framework to inverse problem solving
- Exploit the structure in InfoGAN's latent variable space
- Demonstrate the improvement of the proposed algorithm on MNIST and Celeb-A datasets

Solving Inverse Problems with Learned Prior



$$\min_x ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2$$

+

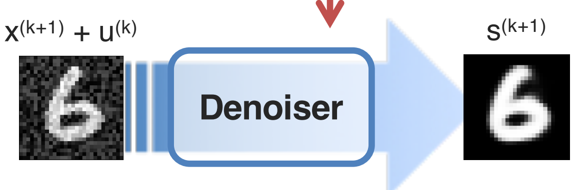
- x is sparse → $\min_x ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda||x||_1$
- x is locally smooth → $\min_x ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda||x||_{TV}$
- Capturing the property of x by J(x) → $\min_x ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda J(\boldsymbol{x})$
- x can be generated by a latent variable z and a mapping function $G_{gen}(\cdot)$ and the property of z can be captured by H(z) → $\min_{x,z} \quad ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda H(\boldsymbol{z}) \quad s.t. \quad \boldsymbol{x} = G_{gen}(\boldsymbol{z})$

Applying ADMM to Inverse Problems

$$\min_{x,s} \quad ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda J(\boldsymbol{s})$$
$$s.t. \quad \boldsymbol{x} = \boldsymbol{s}$$

$$\boldsymbol{x}^{(k+1)} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \rho\mathbf{I}\right)^{-1}\left(\boldsymbol{\Phi}^T\boldsymbol{y} + \rho(\boldsymbol{s}^{(k)} - \boldsymbol{\mu}^{(k)})\right)$$
$$\boldsymbol{s}^{(k+1)} = \arg\min_{\boldsymbol{s}} \ \lambda J(\boldsymbol{s}) + \frac{\rho}{2}||\boldsymbol{s} - (\boldsymbol{x}^{(k+1)} + \boldsymbol{\mu}^{(k)})||_2^2$$
$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \boldsymbol{x}^{(k+1)} - \boldsymbol{s}^{(k)}$$

Denoising step:



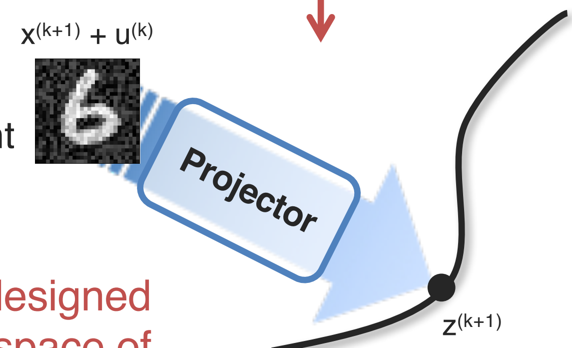
J(s) is implicitly formed during the training of the denoiser

$$\min_{x,s} \quad ||\boldsymbol{y} - \boldsymbol{\Phi x}||_2^2 + \lambda H(\boldsymbol{z})$$
$$s.t. \quad \boldsymbol{x} = G_{gen}(\boldsymbol{z})$$

Explicitly trained previously

$$\boldsymbol{x}^{(k+1)} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \rho\mathbf{I}\right)^{-1}\left(\boldsymbol{\Phi}^T\boldsymbol{y} + \rho(G_{gen}(\boldsymbol{z}^{(k)}) - \boldsymbol{\mu}^{(k)})\right)$$
$$\boldsymbol{z}^{(k+1)} = \arg\min_{\boldsymbol{z}} \ \lambda H(\boldsymbol{z}) + \frac{\rho}{2}||G_{gen}(\boldsymbol{z}) - (\boldsymbol{x}^{(k+1)} + \boldsymbol{\mu}^{(k)})||_2^2$$
$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \boldsymbol{x}^{(k+1)} - G_{gen}(\boldsymbol{z}^{(k)})$$

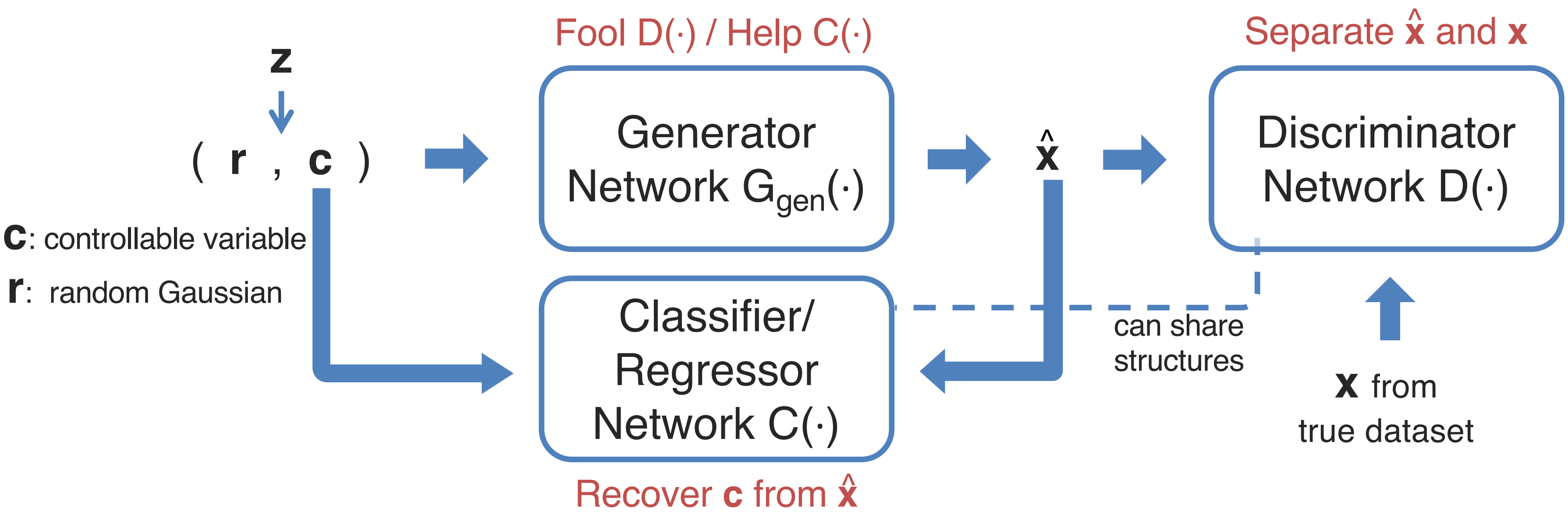
Projection step:
(project to the latent variable space)



H(z) is formed by designed the latent variable space of the generative model

Generative Model with Structured Latent Variable Space

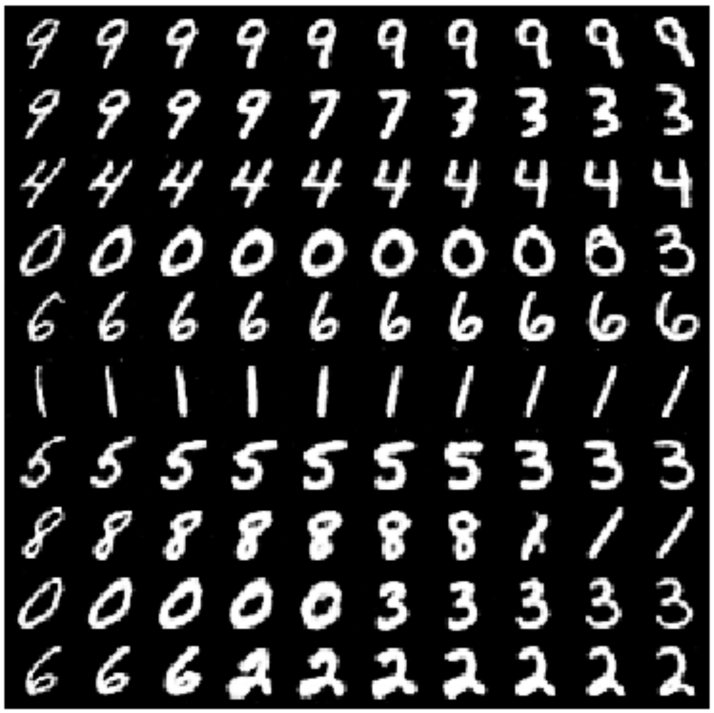
InfoGAN:



Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel, "Infogan: Interpretable representation learning by information maximizing generative adversarial nets," in *Advances in neural information processing systems*, 2016, pp. 2172–2180.

InfoGAN MNIST Examples

InfoGAN Generated Images



- **c** contains one categorical variable and two continuous variables
- Each row corresponds to one categorical code
- Each column corresponds to one 2D continuous code

Fast CS recovery using generative models

Train a generative model $G_{gen}(\cdot)$ on the dataset

Method I:

Generate random latent variable $\boldsymbol{z}$s following its distribution.

Generate random noise $\boldsymbol{\epsilon}$ according to some distribution.

Construct noisy signals $\tilde{\boldsymbol{x}}$s such that $\tilde{\boldsymbol{x}} = G_{gen}(\boldsymbol{z}) + \boldsymbol{\epsilon}$

Train a projector network $G_{proj}(\cdot)$ that maps $\tilde{\boldsymbol{x}}$ to $\boldsymbol{z}$

Method II:

Draw samples $\boldsymbol{x}$ from the training set

Generate random noise $\boldsymbol{\epsilon}$ according to some distribution.

Construct noisy signals $\tilde{\boldsymbol{x}}$s such that $\tilde{\boldsymbol{x}} = \boldsymbol{x} + \boldsymbol{\epsilon}$

Train a projector network $G_{proj}(\cdot)$ such that $G_{gen}(G_{proj}(\cdot))$ maps $\tilde{\boldsymbol{x}}$ to $\boldsymbol{x}$. (G_{gen} is fixed)

For signal recovery:

Given compression matrix $\boldsymbol{\Phi}$, compressed measurements $\boldsymbol{y}$

while Stopping criteria not met **do**

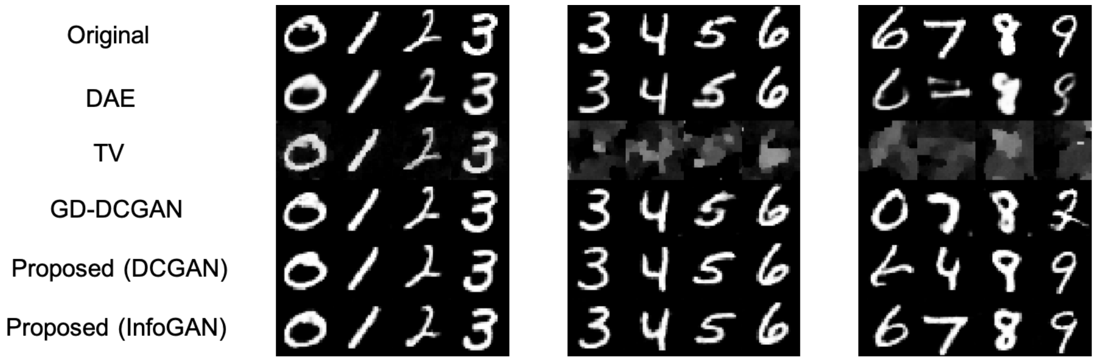
$$\boldsymbol{x}^{(k+1)} = \left(\boldsymbol{\Phi}^T\boldsymbol{\Phi} + \rho\mathbf{I}\right)^{-1}\left(\boldsymbol{\Phi}^T\boldsymbol{y} + \rho(G_{gen}(\boldsymbol{z}^{(k)}) - \boldsymbol{\mu}^{(k)})\right)$$

$$\boldsymbol{z}^{(k+1)} = G_{proj}(\boldsymbol{x}^{(k+1)})$$

$$\boldsymbol{\mu}^{(k+1)} = \boldsymbol{\mu}^{(k)} + \boldsymbol{x}^{(k+1)} - G_{gen}(\boldsymbol{z}^{(k)})$$

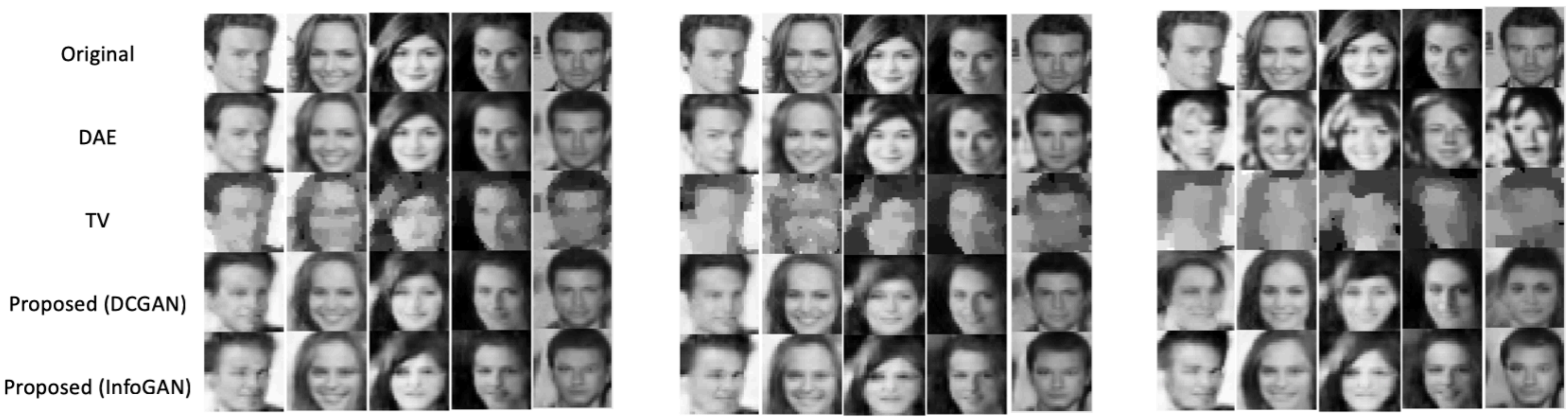
end while

Testing on MNIST Digits and Celeb-A Database



4x Compression 16x Compression 32x Compression

Compression Ratio	DAE	F-CSRG with DCGAN	F-CSRG with InfoGAN
4x	2.20 / 98.2%	2.25 / 98.3%	2.68 / 97.7%
8x	2.54 / 97.8%	2.72 / 97.3%	3.06 / 97.2%
16x	3.23 / 94.8%	3.70 / 91.7%	3.79 / 93.8%
32x	5.13 / 73.5%	5.86 / 66.4%	5.37 / 77.4%
64x	7.33 / 41.8%	7.91 / 36.2%	7.43 / 48.0%



4x Compression 8x Compression 16x Compression

Theoretical Background

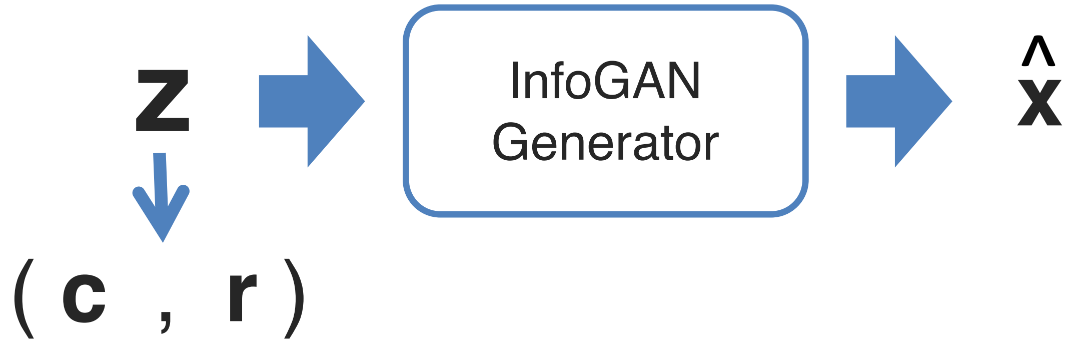
Theorem 1.1. Let $G : \mathbb{R}^k \to \mathbb{R}^n$ be a generative model from a d-layer neural network using ReLU activations. Let $A \in \mathbb{R}^{m \times n}$ be a random Gaussian matrix for $m = O(kd \log n)$, scaled so $A_{i,j} \sim N(0, 1/m)$. For any $x^* \in \mathbb{R}^n$ and any observation $y = Ax^* + \eta$, let $\widehat{z}$ minimize $\|y - AG(z)\|_2$ to within additive ϵ of the optimum. Then with $1 - e^{-\Omega(m)}$ probability,

$$\|G(\widehat{z}) - x^*\|_2 \leq 6 \min_{z^* \in \mathbb{R}^k} \|G(z^*) - x^*\|_2 + 3\|\eta\|_2 + 2\epsilon.$$

- Solving in the compressed domain (LHS) has comparable performance as solving in the uncompressed domain (RHS).
- Number of measurements required is linear to the dimension of the latent variable space (k).

Ashish Bora, Ajil Jalal, Eric Price, andAlexandros G Dimakis, “Compressed sensing using generative models,” in International Conference on Machine Learning, 2017, pp. 537–546.

Theoretical Background



c: controllable variable of size k_c ; controls major variations in the generated images

r: random Gaussian of size k_r ; adds fine details in the generated images

$$k_c \ll k_c + k_r = k$$

Lose details when the number of compressed measurements is extremely limited.

Conclusion

- **Strong prior knowledge** captured by the **generative models** allows higher reduction in the number of required compressed measurements.
- Apply **ADMM** framework to inverse problem solving leads to more freedom in algorithm design.
- Train a **projector network** that maps signals to the latent variable space to **accelerate** the recovery.
- **Structures in the latent variable space** play an important role in increasing the robustness of the recovery algorithm.