Jul 17–23, 2022

Uncertainty Modeling in Generative Compressed Sensing

Yilang Zhang

School of Data Science, Fudan University

Acknowledgements: Mengchu Xu, Prof. Xiaojun Mao, and Prof. Jian Wang*

NSFC grants 12001109, 92046021, 61971146

Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \ (m \ll n)$$

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \; (m \ll n)$$

Applications: data compression, MRI, super-resolution, …

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \ (m \ll n)$$

Applications: data compression, MRI, super-resolution, …

underdetermined system ➢ need additional priors to ensure unique solution

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \ (m \ll n)$$

Applications: data compression, MRI, super-resolution, …

underdetermined system ➢ need additional priors to ensure unique solution

- Compressed sensing with sparsity prior
  - Sparsity prior $\|\mathbf{x}\|_0 \leq K$
    - ✓ natural signals are nearly sparse in some transform domains, e.g. Fourier, wavelet

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \ (m \ll n)$$

Applications: data compression, MRI, super-resolution, …

underdetermined system ➢ need additional priors to ensure unique solution

- Compressed sensing with sparsity prior
  - Sparsity prior $\|\mathbf{x}\|_0 \leq K$

✓ natural signals are nearly sparse in some transform domains, e.g. Fourier, wavelet

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$$

➢ solve via greedy algorithms or convex relaxation

Sparsity-based Compressed Sensing

- Inverse problems

Goal: recover signal $\mathbf{x} \in \mathbb{R}^n$ from compressed linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n} \qquad \mathbf{A} \in \mathbb{R}^{m \times n} \ (m \ll n)$$

Applications: data compression, MRI, super-resolution, …

underdetermined system ➤ need additional priors to ensure unique solution

- Compressed sensing with sparsity prior
  - Sparsity prior $\|\mathbf{x}\|_0 \leq K$
    - ✓ natural signals are nearly sparse in some transform domains, e.g. Fourier, wavelet

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0, \quad \text{s.t. } \mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$$

➤ solve via greedy algorithms or convex relaxation

Limitation: $\mathbf{x}$ is not strictly sparse ➤ inaccurate recovery results

Generative Compressed Sensing

- Compressed sensing with generative models (CSGM)

Key idea: learn more accurate priors in a data-driven way

A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed Sensing using Generative Models," *Proceedings of the 34th International Conference on Machine Learning,* 2017.

Generative Compressed Sensing

- Compressed sensing with generative models (CSGM)

Key idea: learn more accurate priors in a data-driven way

S1. Pre-train a generator (GAN) / decoder (VAE) using training signals $\mathbf{X}_{\mathrm{tr}} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$

$$\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) : \mathbb{R}^k \mapsto \mathbb{R}^n$$

latent variables $\mathbf{z} \in \mathbb{R}^k$ $(k \ll n)$
with prior $\mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I}_k)$

- output signals that resemble the training ones

A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed Sensing using Generative Models," *Proceedings of the 34th International Conference on Machine Learning,* 2017.

Generative Compressed Sensing

- Compressed sensing with generative models (CSGM)

Key idea: learn more accurate priors in a data-driven way

S1. Pre-train a generator (GAN) / decoder (VAE) using training signals $\mathbf{X}_{\text{tr}} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$

$$\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) : \mathbb{R}^k \mapsto \mathbb{R}^n$$

latent variables $\mathbf{z} \in \mathbb{R}^k$ $(k \ll n)$
with prior $\mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I}_k)$

- output signals that resemble the training ones

S2. Optimize latent variable to minimize the loss

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathbf{A}\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2^2 + \lambda_{\mathbf{z}} \|\mathbf{z}\|_2^2$$

relative weight $\lambda_{\mathbf{z}}$

and then reconstruct

$$\hat{\mathbf{x}} = \mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}})$$

A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed Sensing using Generative Models," *Proceedings of the 34th International Conference on Machine Learning,* 2017.

Generative Compressed Sensing

- Compressed sensing with generative models (CSGM)

Key idea: learn more accurate priors in a data-driven way

S1. Pre-train a generator (GAN) / decoder (VAE) using training signals $\mathbf{X}_{\text{tr}} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$

$$\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) : \mathbb{R}^k \mapsto \mathbb{R}^n$$

latent variables $\mathbf{z} \in \mathbb{R}^k \ (k \ll n)$ with prior $\mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I}_k)$

- output signals that resemble the training ones

S2. Optimize latent variable to minimize the loss

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathbf{A}\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2^2 + \lambda_{\mathbf{z}} \|\mathbf{z}\|_2^2$$

relative weight $\lambda_{\mathbf{z}}$

and then reconstruct

$$\hat{\mathbf{x}} = \mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}})$$

✓ Bounded reconstruction error for $\mathbf{x}$ inside the range $\mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}})) := \{\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) | \mathbf{z} \in \mathbb{R}^k\}$

$$\gamma \|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

constants γ, ϵ

A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed Sensing using Generative Models," *Proceedings of the 34th International Conference on Machine Learning*, 2017.

Generative Compressed Sensing

- Compressed sensing with generative models (CSGM)

Key idea: learn more accurate priors in a data-driven way

S1. Pre-train a generator (GAN) / decoder (VAE) using training signals $\mathbf{X}_{\text{tr}} = [\mathbf{x}_1, \ldots, \mathbf{x}_N]$

$$\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) : \mathbb{R}^k \mapsto \mathbb{R}^n$$

latent variables $\mathbf{z} \in \mathbb{R}^k$ $(k \ll n)$
with prior $\mathcal{N}(\mathbf{z}; \mathbf{0}, \mathbf{I}_k)$

- output signals that resemble the training ones

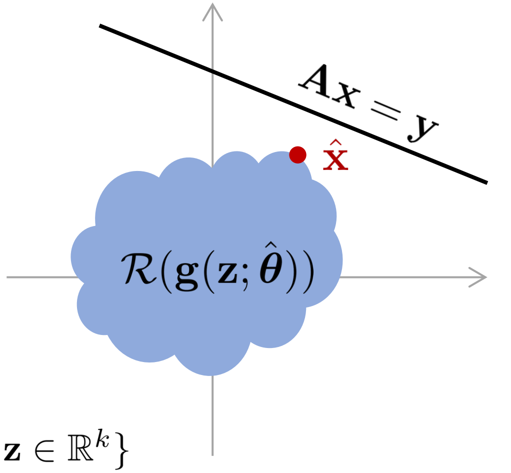
S2. Optimize latent variable to minimize the loss

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathbf{A}\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2^2 + \lambda_{\mathbf{z}}\|\mathbf{z}\|_2^2$$

relative weight $\lambda_{\mathbf{z}}$

and then reconstruct

$$\hat{\mathbf{x}} = \mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}})$$



✓ Bounded reconstruction error for $\mathbf{x}$ inside the range $\mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}})) := \{\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}) | \mathbf{z} \in \mathbb{R}^k\}$

$$\gamma\|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

constants γ, ϵ

❖ Inferior performance when $\mathbf{x} \notin \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$

A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed Sensing using Generative Models," *Proceedings of the 34th International Conference on Machine Learning*, 2017.

Modeling Uncertainties in CSGM

Key idea: deterministic $\hat{\boldsymbol{\theta}}$ leads to fixed range $\mathcal{R}(\mathbf{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ ➢ model uncertainties in $\boldsymbol{\theta}$

Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Modeling Uncertainties in CSGM

Key idea: deterministic $\hat{\boldsymbol{\theta}}$ leads to fixed range $\mathcal{R}(\mathbf{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ ➢ model uncertainties in $\boldsymbol{\theta}$

✓ Uncertainties come from distributional difference between $\mathbf{X}_{\mathrm{tr}}$ and $\mathbf{x}$

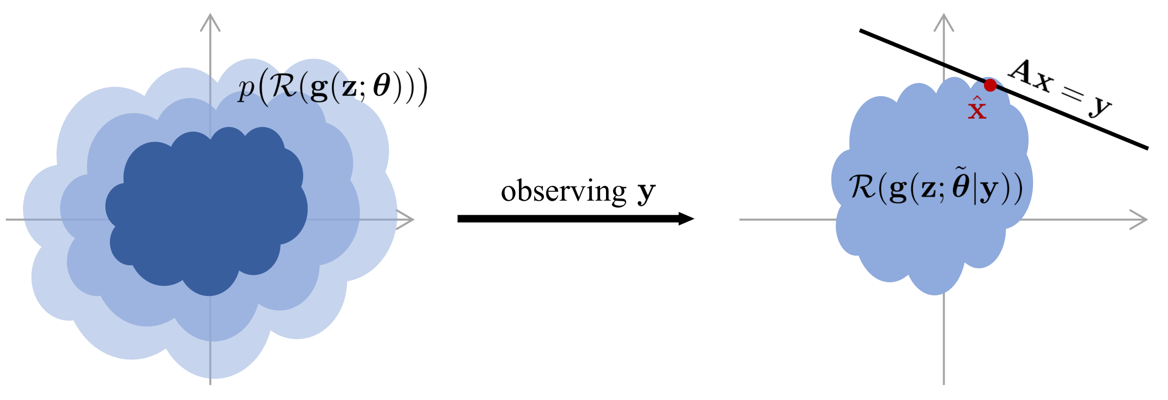
Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Modeling Uncertainties in CSGM

Key idea: deterministic $\hat{\boldsymbol{\theta}}$ leads to fixed range $\mathcal{R}(\mathbf{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ ➢ model uncertainties in $\boldsymbol{\theta}$

✓ Uncertainties come from distributional difference between $\mathbf{X}_{\mathrm{tr}}$ and $\mathbf{x}$

- Compressed sensing with Bayesian generative models (CS-BGM)
  - Model $p(\boldsymbol{\theta};\mathbf{X}_{\mathrm{tr}}) = \mathcal{N}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}},\lambda\mathbf{I})$ ➢ allow for slight adjustment of $\boldsymbol{\theta}$



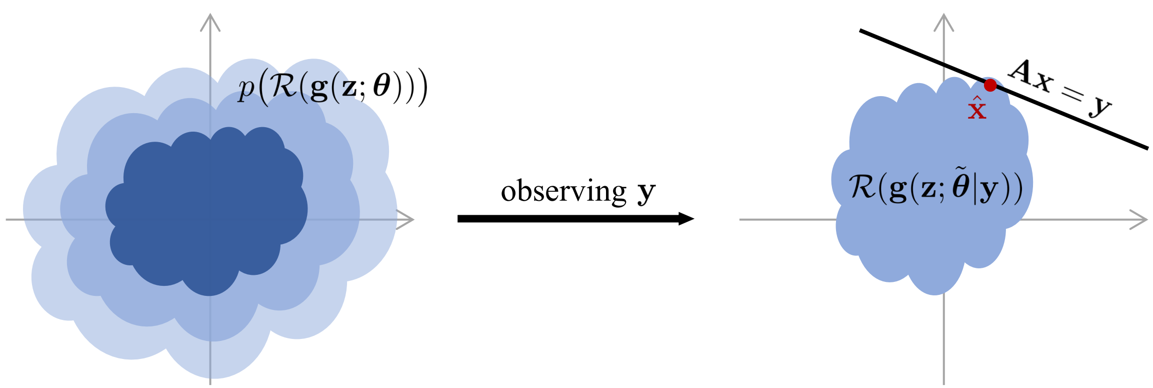
Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Modeling Uncertainties in CSGM

Key idea: deterministic $\hat{\theta}$ leads to fixed range $\mathcal{R}(\mathbf{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ ➢ model uncertainties in $\boldsymbol{\theta}$

✓ Uncertainties come from distributional difference between $\mathbf{X}_{\text{tr}}$ and $\mathbf{x}$

- Compressed sensing with Bayesian generative models (CS-BGM)
  - Model $p(\boldsymbol{\theta};\mathbf{X}_{\text{tr}}) = \mathcal{N}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}},\lambda\mathbf{I})$ ➢ allow for slight adjustment of $\boldsymbol{\theta}$
  - Prediction $\mathbb{E}[\mathbf{x}|\mathbf{y};\mathbf{X}_{\text{tr}}] = \mathbb{E}_{p(\mathbf{z},\boldsymbol{\theta}|\mathbf{y};\mathbf{X}_{\text{tr}})}[\mathbf{g}(\mathbf{z};\boldsymbol{\theta})]$



Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

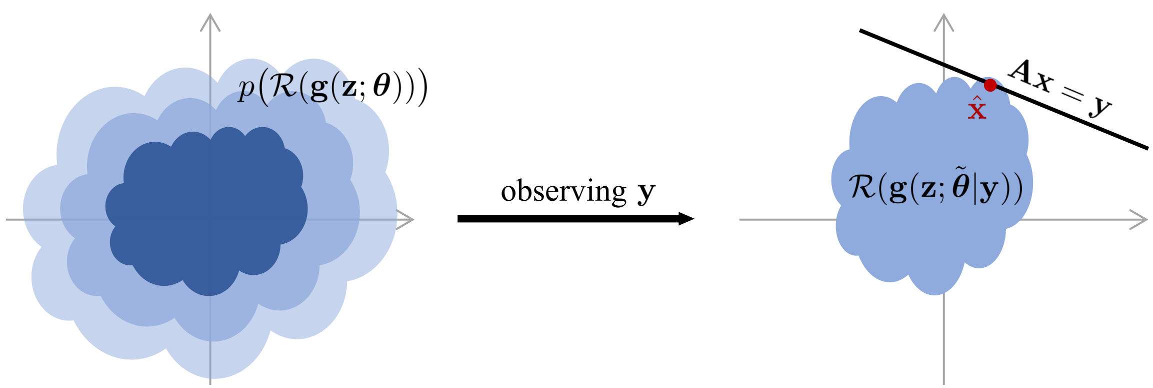
Modeling Uncertainties in CSGM

Key idea: deterministic $\hat{\boldsymbol{\theta}}$ leads to fixed range $\mathcal{R}(\mathbf{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ ➢ model uncertainties in $\boldsymbol{\theta}$

✓ Uncertainties come from distributional difference between $\mathbf{X}_{\text{tr}}$ and $\mathbf{x}$

❑ Compressed sensing with Bayesian generative models (CS-BGM)

- Model $p(\boldsymbol{\theta};\mathbf{X}_{\text{tr}}) = \mathcal{N}(\boldsymbol{\theta};\hat{\boldsymbol{\theta}},\lambda\mathbf{I})$ ➢ allow for slight adjustment of $\boldsymbol{\theta}$
- Prediction $\mathbb{E}[\mathbf{x}|\mathbf{y};\mathbf{X}_{\text{tr}}] = \mathbb{E}_{p(\mathbf{z},\boldsymbol{\theta}|\mathbf{y};\mathbf{X}_{\text{tr}})}[\mathbf{g}(\mathbf{z};\boldsymbol{\theta})]$
- Solve for $p(\mathbf{z},\boldsymbol{\theta}|\mathbf{y};\mathbf{X}_{\text{tr}})$ via alternating optimization
  - maximum a posteriori (MAP) for high-dimensional $\boldsymbol{\theta}$
  - variational inference (VI) for low-dimensional $\mathbf{z}$



Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Theoretical Justification of CS-BGM

Theoretical Justification of CS-BGM

- Necessary condition for generator range
  - ✓ (Bora et al'17) If $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$, small reconstruction error can be achieved

$$\text{S-REC: } \gamma \|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

Theoretical Justification of CS-BGM

- Necessary condition for generator range
  - ✓ (Bora et al'17) If $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$, small reconstruction error can be achieved

$$\text{S-REC: } \gamma \|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

Q. But… how can we know when $\mathbf{x} \notin \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$?

Theoretical Justification of CS-BGM

- Necessary condition for generator range
  - ✓ (Bora et al'17) If $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$, small reconstruction error can be achieved

$$\text{S-REC: } \gamma \|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

Q. But… how can we know when $\mathbf{x} \notin \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$?

- Sufficient condition for generator range
  - ✓ (Ours) $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$ only if small measurement error can be achieved

$$\text{S-RIP: } \sqrt{1-\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2 \leq \sqrt{1+\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon$$

constants δ, ϵ

Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Theoretical Justification of CS-BGM

- Necessary condition for generator range
  - ✓ (Bora et al'17) If $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$, small reconstruction error can be achieved

$$\text{S-REC: } \gamma \|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

Q. But… how can we know when $\mathbf{x} \notin \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$?

- Sufficient condition for generator range
  - ✓ (Ours) $\mathbf{x} \in \mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$ only if small measurement error can be achieved

$$\text{S-RIP: } \sqrt{1-\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon \leq \|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}}; \hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2 \leq \sqrt{1+\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon$$

constants δ, ϵ

  - ➢ Otherwise, adjust $\mathcal{R}(\mathbf{g}(\mathbf{z}; \hat{\boldsymbol{\theta}}))$ to include $\mathbf{x}$

Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Theoretical Justification of CS-BGM

- Necessary condition for generator range
  - ✓ (Bora et al'17) If $\mathbf{x} \in \mathcal{R}(\mathrm{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$, small reconstruction error can be achieved

$$\text{S-REC: } \gamma\|\hat{\mathbf{x}} - \mathbf{x}\|_2 - \epsilon \leq \|\mathbf{A}\mathrm{g}(\hat{\mathbf{z}};\hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2$$

Q. But… how can we know when $\mathbf{x} \notin \mathcal{R}(\mathrm{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$?

- Sufficient condition for generator range
  - ✓ (Ours) $\mathbf{x} \in \mathcal{R}(\mathrm{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ only if small measurement error can be achieved

$$\text{S-RIP: } \sqrt{1-\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon \leq \|\mathbf{A}\mathrm{g}(\hat{\mathbf{z}};\hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2 \leq \sqrt{1+\delta}\|\hat{\mathbf{x}} - \mathbf{x}\|_2 + \epsilon$$

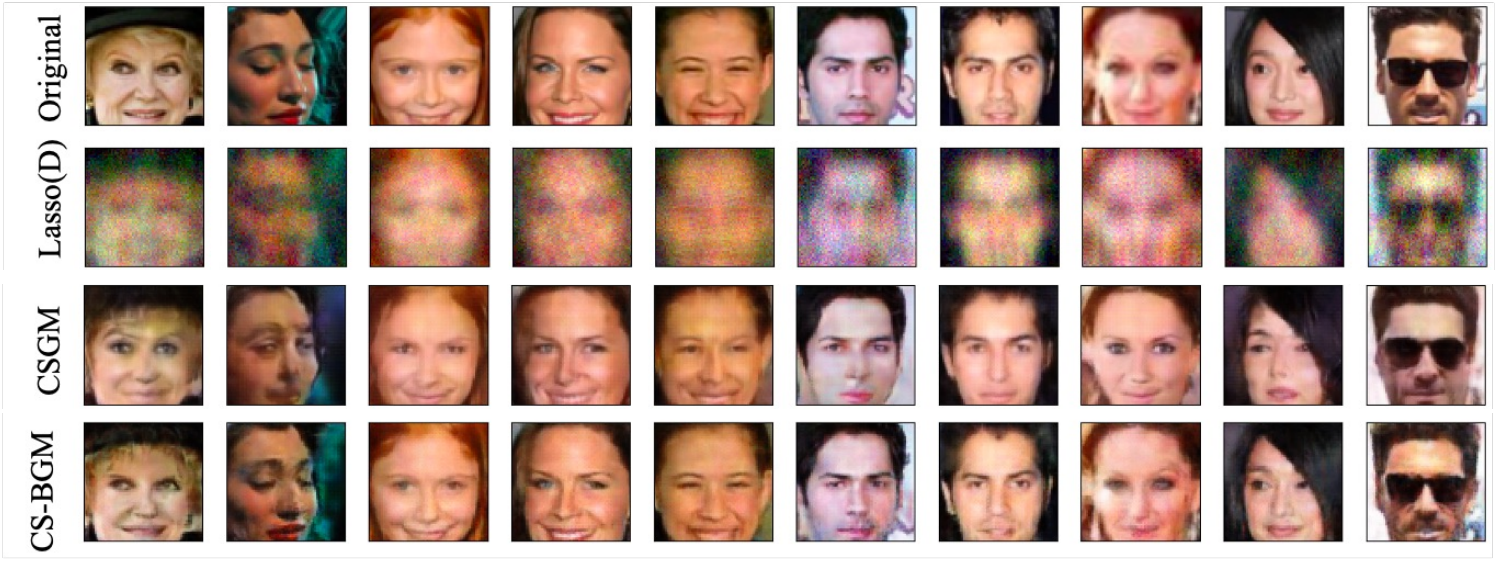
constants δ, ϵ

- ➢ Otherwise, adjust $\mathcal{R}(\mathrm{g}(\mathbf{z};\hat{\boldsymbol{\theta}}))$ to include $\mathbf{x}$

	CSGM	CS-BGM
Measurement error $\frac{1}{n}\|\mathbf{A}\mathbf{g}(\hat{\mathbf{z}};\hat{\boldsymbol{\theta}}) - \mathbf{y}\|_2^2$:	1.073	0.047
Reconstruction error $\frac{1}{n}\|\mathbf{g}(\hat{\mathbf{z}};\hat{\boldsymbol{\theta}}) - \mathbf{x}\|_2^2$:	0.0137	0.0044

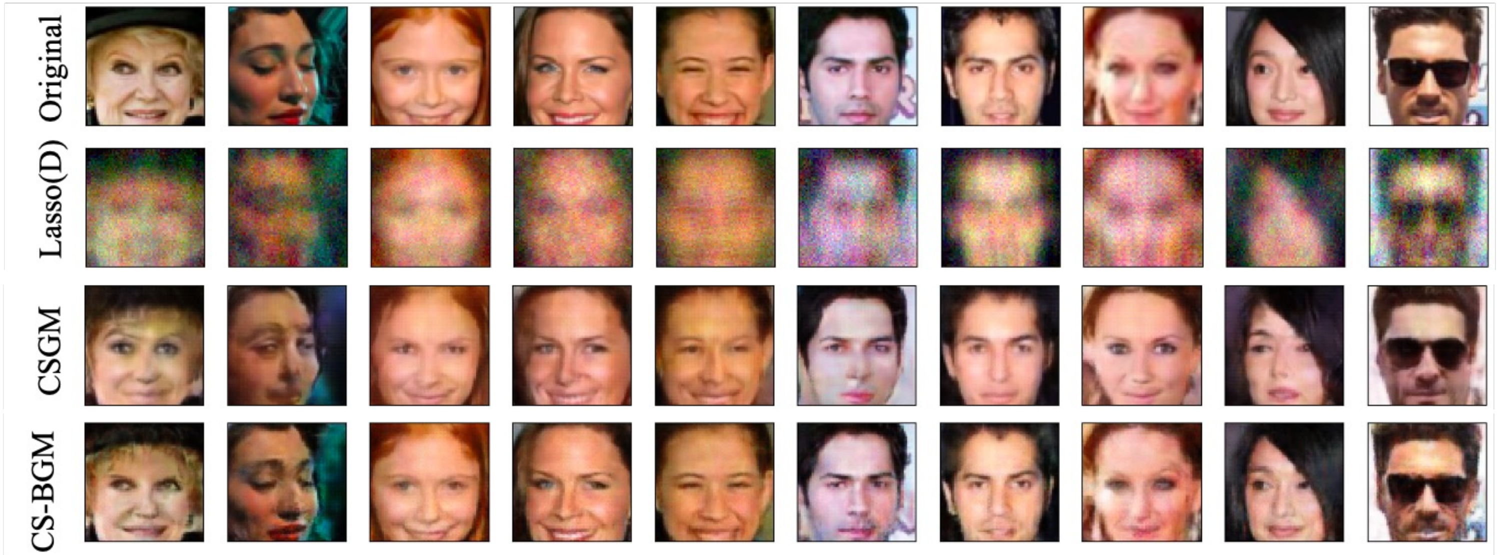
Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

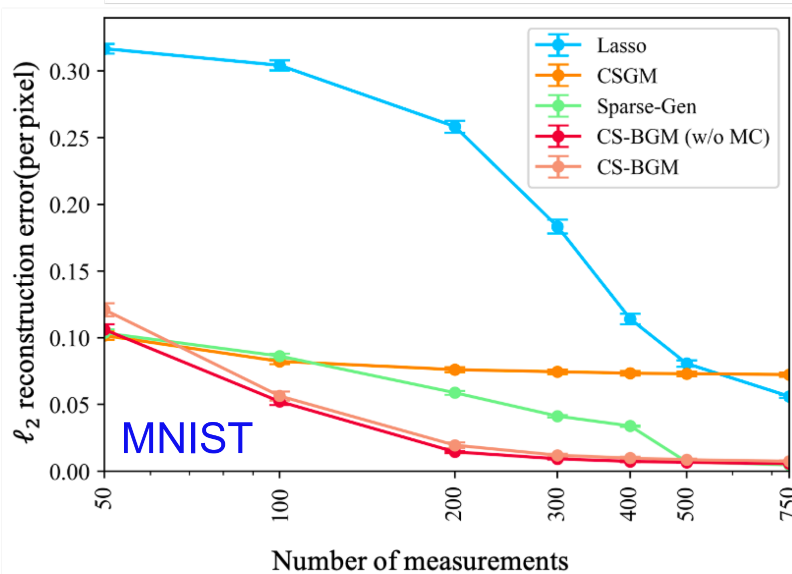
Experiments

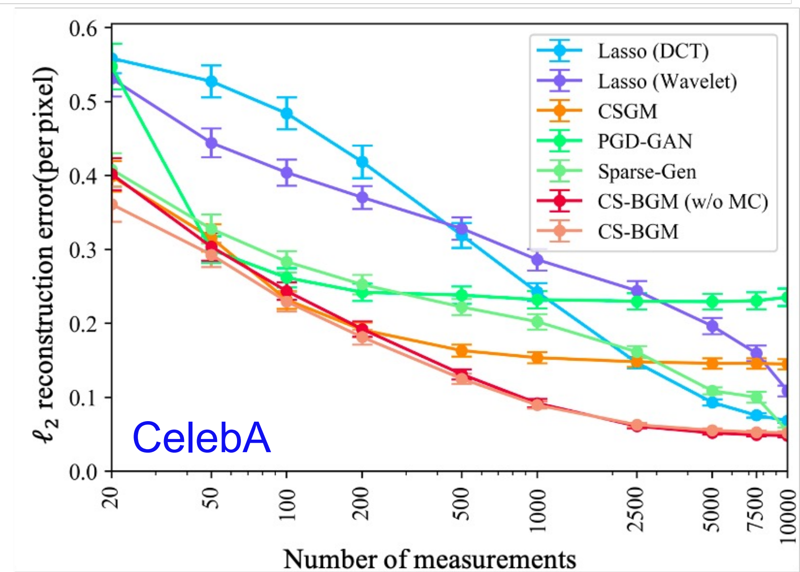


Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Experiments

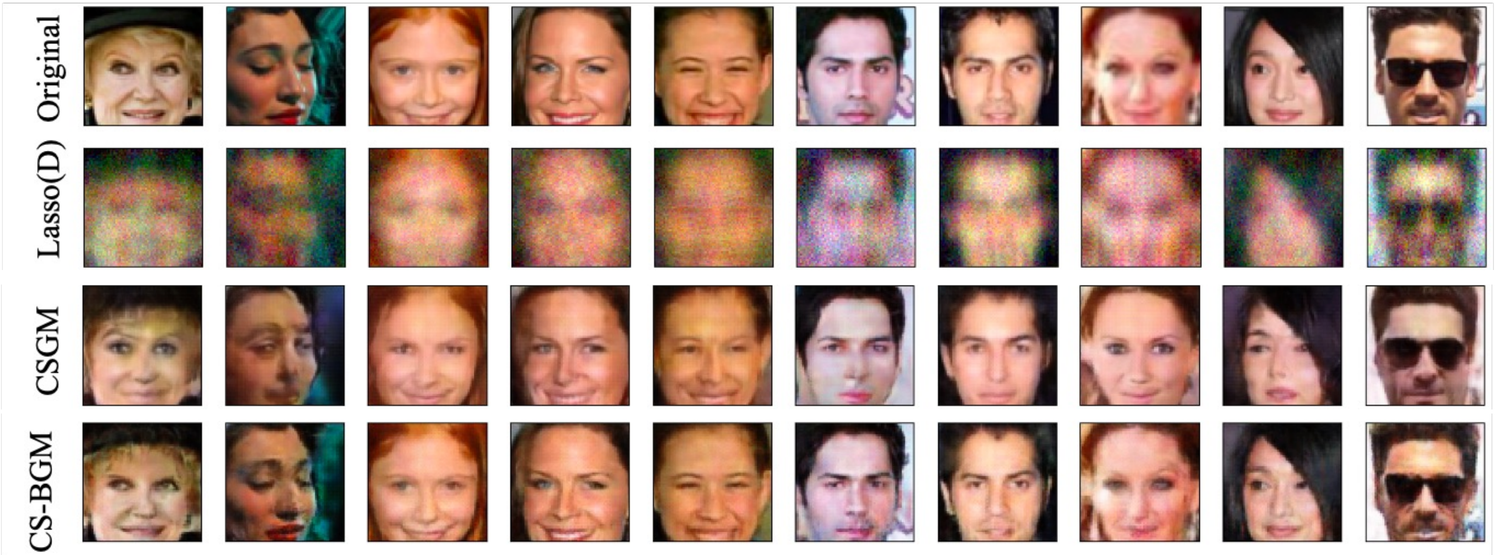


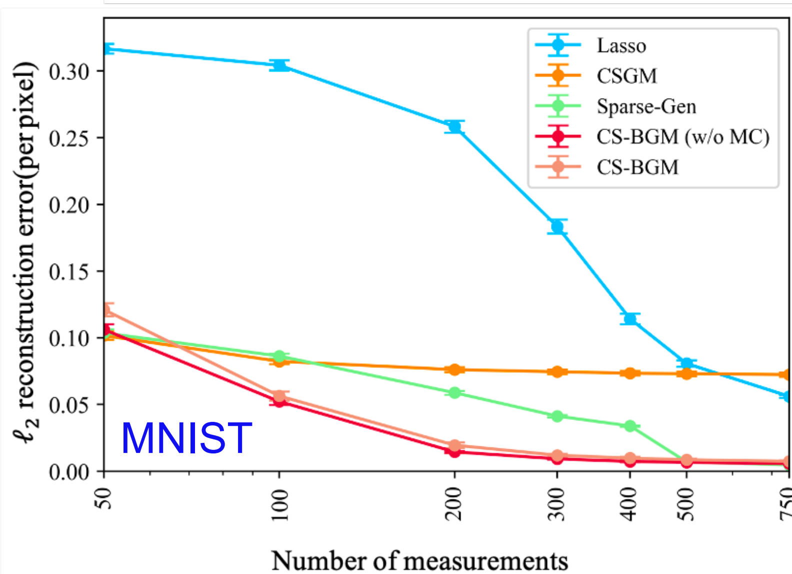


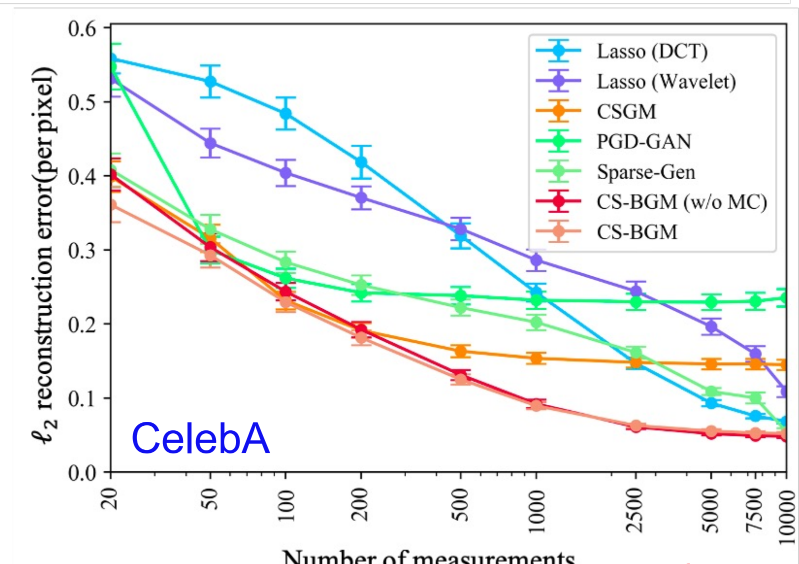


Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.

Experiments







Thank you!

Y. Zhang, M. Xu, X. Mao, and J. Wang, "Uncertainty Modeling in Generative Compressed Sensing," *Proceedings of the 39th International Conference on Machine Learning*, 2022.