pcaGAN: Improving posterior-sampling cGANs via principal component regularization

Matthew Bendel, Rizwan Ahmad, and Philip Schniter

THE OHIO STATE UNIVERSITY

Supported by NIH grant R01EB029957

NEURAL INFORMATION PROCESSING SYSTEMS

Image inverse problems

Goal: Recover image $\boldsymbol{x}$ from measurements $\boldsymbol{y} = \mathcal{M}(\boldsymbol{x})$:

- $\mathcal{M}(\cdot)$ masks, distorts, and/or corrupts $\boldsymbol{x}$ with noise.
- Solution typically posed as finding single best recovery $\widehat{\boldsymbol{x}}$, known as "point-estimation"

Challenges with point-estimation:

- Inability to navigate the perception-distortion tradeoff [1]
- Inability to quantify reconstruction uncertainty

Solution: Sample from posterior distribution $p_{\mathsf{x}|\mathsf{y}}(\boldsymbol{x}|\boldsymbol{y}) = \dfrac{p_{\mathsf{y}|\mathsf{x}}(\boldsymbol{y}|\boldsymbol{x})p_{\mathsf{x}}(\boldsymbol{x})}{\int p_{\mathsf{y}|\mathsf{x}}(\boldsymbol{y}|\boldsymbol{x})p_{\mathsf{x}}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x}}$

Existing approaches:

- Conditional VAEs, conditional NFs, conditional GANs
- Langevin/Diffusion methods

Our contribution

Our approach: We build on rcGAN [2], a type of Wasserstein cGAN:

- Generator $G_{\boldsymbol{\theta}}$: outputs $\widehat{\boldsymbol{x}}_i = G_{\boldsymbol{\theta}}(\boldsymbol{z}_i, \boldsymbol{y})$ for code realization $\boldsymbol{z}_i \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$
- Discriminator $D_{\boldsymbol{\phi}}$: aims to distinguish true $(\boldsymbol{x}, \boldsymbol{y})$ from fake $(\widehat{\boldsymbol{x}}_i, \boldsymbol{y})$
- Training:

$$\min_{\boldsymbol{\theta}} \max_{\boldsymbol{\phi}} \left\{ \mathbb{E}_{\mathsf{x},\mathsf{z},\mathsf{y}}\{D_{\boldsymbol{\phi}}(\boldsymbol{x},\boldsymbol{y}) - D_{\boldsymbol{\phi}}(G_{\boldsymbol{\theta}}(\boldsymbol{z},\boldsymbol{y}),\boldsymbol{y})\} + \mathcal{R}(\boldsymbol{\theta}) - \mathcal{L}_{\mathsf{gp}}(\boldsymbol{\phi}) \right\}$$

- rcGAN's regularization $\mathcal{R}(\boldsymbol{\theta})$ rewards correctness in conditional mean and conditional trace-covariance

Our method: A new $\mathcal{R}(\boldsymbol{\theta})$ that **also** enforces correctness in the K principal components of the conditional covariance matrix, given by

$$\mathcal{R}(\boldsymbol{\theta}) = \mathcal{R}_{\mathsf{rc}}(\boldsymbol{\theta}) + \beta_{\mathsf{pca}}\mathcal{L}_{\mathsf{evec}}(\boldsymbol{\theta}) + \beta_{\mathsf{pca}}\mathcal{L}_{\mathsf{eval}}(\boldsymbol{\theta}),$$

where $\mathcal{R}_{\mathsf{rc}}(\boldsymbol{\theta}) \triangleq \mathbb{E}_{\mathsf{x},\mathsf{z}_1,\dots,\mathsf{z}_P,\mathsf{y}}\left\{\|\boldsymbol{x} - \widehat{\boldsymbol{x}}_{(P)}\|_1\right\} - \beta_{\mathsf{std}}\sum_{i=1}^{P}\mathbb{E}_{\mathsf{z}_1,\dots,\mathsf{z}_P,\mathsf{y}}\left\{\|\widehat{\boldsymbol{x}}_i - \widehat{\boldsymbol{x}}_{(P)}\|_1\right\}$ and $\widehat{\boldsymbol{x}}_{(P)} = \sum_{i=1}^{P}\widehat{\boldsymbol{x}}_i$.

- $\{(\widehat{\boldsymbol{v}}_k, \widehat{\lambda}_k)\}_{k=1}^K$ are the principal evecs/evals of $\operatorname{Cov}\{\widehat{\boldsymbol{x}}_i|\boldsymbol{y}\}$
- $\{(\boldsymbol{v}_k, \lambda_k)\}_{k=1}^K$ are the principal evecs/evals of $\operatorname{Cov}\{\boldsymbol{x}|\boldsymbol{y}\}$
- We dub our approach pcaGAN

Eigenvector regularization:

$$\mathcal{L}_{\mathsf{evec}}(\boldsymbol{\theta}) = -\mathbb{E}_{\mathsf{y}}\left\{\mathbb{E}_{\mathsf{x},\mathsf{z}_1,\dots,\mathsf{z}_P|\mathsf{y}}\left\{\textstyle\sum_{k=1}^K[\widehat{\boldsymbol{v}}_k^{\mathsf{T}}(\boldsymbol{x} - \boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}})]^2|\boldsymbol{y}\right\}\right\}$$

- If $\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}} = \mathbb{E}\{\boldsymbol{x}|\boldsymbol{y}\}$ was known, minimizing over $\boldsymbol{\theta}$ would force $\{\widehat{\boldsymbol{v}}_k = \boldsymbol{v}_k\}_{k=1}^K$
- We set $\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}} \approx \widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}} = \mathtt{StopGrad}(\widehat{\boldsymbol{x}}_{(P_{\mathsf{pca}})})$ after $\widehat{\boldsymbol{x}}_{(P_{\mathsf{pca}})}$ stabilizes
- We compute $\{\widehat{\boldsymbol{v}}_i\}_{k=1}^K$ using an SVD of $\{\widehat{\boldsymbol{x}}_i\}_{i=1}^{P_{\mathsf{pca}}}$, where $P_{\mathsf{pca}} \approx 10K$

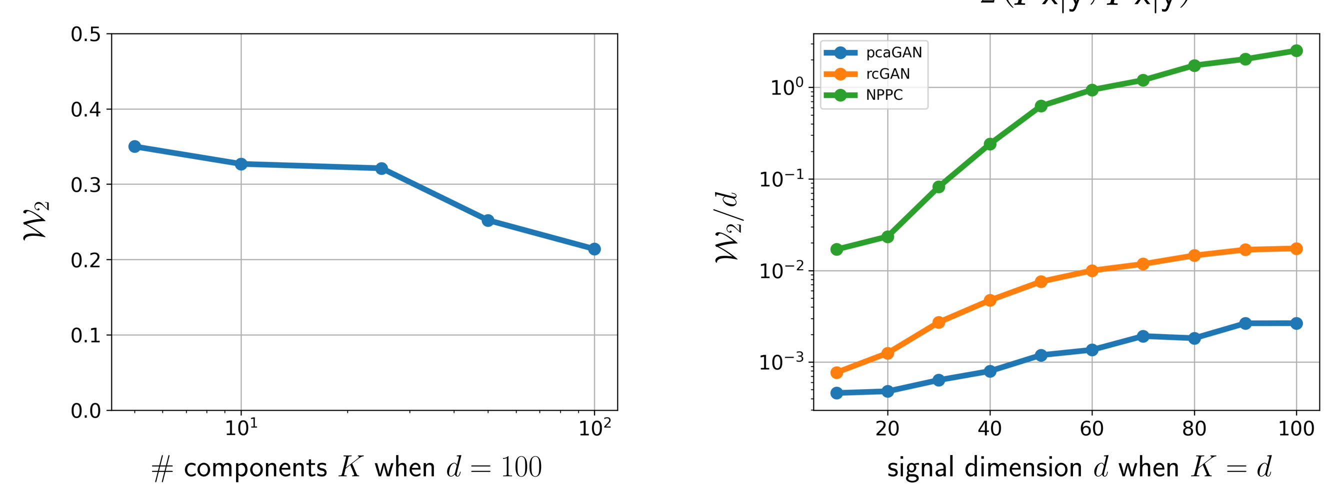
Eigenvalue regularization:

$$\mathcal{L}_{\mathsf{eval}}(\boldsymbol{\theta}) = \mathbb{E}_{\mathsf{y}}\left\{\mathbb{E}_{\mathsf{x},\mathsf{z}_1,\dots,\mathsf{z}_P|\mathsf{y}}\left\{\textstyle\sum_{k=1}^K\left(1 - \lambda_k/\widehat{\lambda}_k\right)^2|\boldsymbol{y}\right\}\right\}$$

- If λ_k was known, minimizing over $\boldsymbol{\theta}$ would force $\{\widehat{\lambda}_k = \lambda_k\}_{k=1}^K$
- We set $\lambda_k \approx \mathtt{StopGrad}(\frac{1}{P_{\mathsf{pca}}+1}\|\widehat{\boldsymbol{v}}_k^{\mathsf{T}}[\boldsymbol{x} - \widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}}, \widehat{\boldsymbol{x}}_1 - \widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}}, \dots, \widehat{\boldsymbol{x}}_{P_{\mathsf{pca}}} - \widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}}]\|_2^2)$ after $\{\widehat{\boldsymbol{v}}_k\}$ stabilize, which is correct in expectation when $\widehat{\boldsymbol{v}}_k$ and $\widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}}$ are correct
- We compute $\{\widehat{\lambda}_i\}_{k=1}^K$ using the same SVD as above

Toy Gaussian experiment

- Here we recover $\boldsymbol{x} \sim \mathcal{N}(\boldsymbol{\mu}_{\mathsf{x}}, \boldsymbol{\Sigma}_{\mathsf{x}})$ from $\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{w}$ with inpainting $\boldsymbol{A} \in \mathbb{R}^{\frac{d}{2}\times d}$ and $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \sigma^2\boldsymbol{I})$. Both $\boldsymbol{\mu}_{\mathsf{x}}$ and $\boldsymbol{\Sigma}_{\mathsf{x}}$ are random.
- Performance measured via Wasserstein-2 distance $\mathcal{W}_2(p_{\mathsf{x}|\mathsf{y}}, p_{\widehat{\mathsf{x}}|\mathsf{y}})$



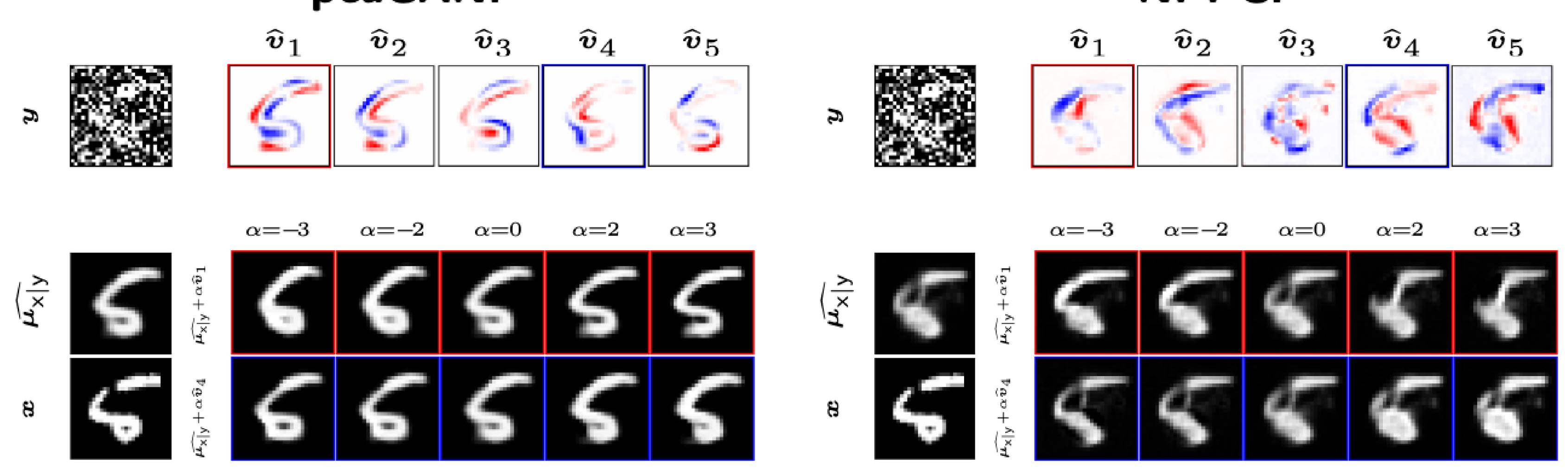
- $\mathcal{W}_2$ decreases with K, and pcaGAN beats both rcGAN and NPPC [3] in $\mathcal{W}_2$
- NPPC directly estimates the posterior principal components

MNIST denoising results

We recovered MNIST digits $\boldsymbol{x}$ from measurements $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{w}$ with $\boldsymbol{w} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{I})$

Model	rMSE↓	REM$_5$↓	CFID↓	Time (128 samples)↓
NPPC ($K = 5$)	**3.94**	3.63	–	**112 ms**
rcGAN	4.04	3.41	63.44	118 ms
pcaGAN (ours, $K = 5$)	4.02	**3.31**	**61.48**	118 ms

- $\mathrm{REM}_5 \triangleq \mathbb{E}_{\mathsf{x},\mathsf{y}}\left\{\|(\boldsymbol{I} - \widehat{\boldsymbol{V}}_5\widehat{\boldsymbol{V}}_5^{\mathsf{T}})(\boldsymbol{x} - \widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}})\|_2\right\}$ where $\widehat{\boldsymbol{V}}_5 \triangleq [\widehat{\boldsymbol{v}}_1, \dots, \widehat{\boldsymbol{v}}_5]$
- Conditional FID (CFID) [4] like FID but for conditional distributions
- pcaGAN won in both REM and CFID (note NPPC is not generative)



- Principal eigenvectors $\{\boldsymbol{v}_k\}_{k=1}^K$ are shown above for $K = 5$
- Also shown are $\widehat{\boldsymbol{\mu}_{\mathsf{x}|\mathsf{y}}} \pm \alpha\widehat{\boldsymbol{v}}_k$ for $k \in \{1, 4\}$ and $\alpha \in \{-3, -2, 0, 2, 3\}$

Large-scale image completion results

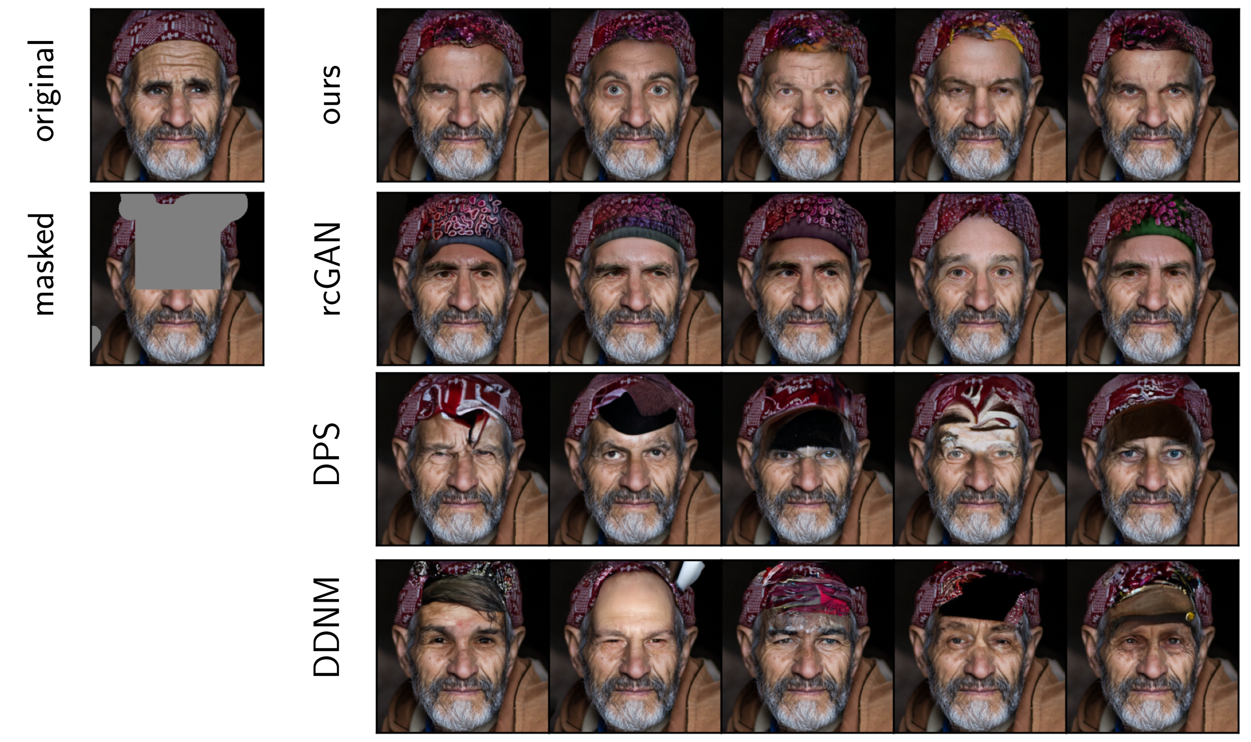
We inpainted random masks on 256x256 FFHQ faces

Model	CFID↓	FID↓	LPIPS↓	Time (40 samples)↓
DPS [5] (1000 NFEs)	7.26	2.00	0.1245	14 min
DDNM [6] (100 NFEs)	11.30	3.63	0.1409	30 s
DDRM [7] (20 NFEs)	13.17	5.36	0.1587	5 s
pscGAN [8]	18.44	8.40	0.1716	**325 ms**
CoModGAN [9]	7.85	2.23	0.1290	**325 ms**
rcGAN [2]	7.51	2.12	0.1262	**325 ms**
pcaGAN (ours, $K = 2$)	**7.08**	**1.98**	**0.1230**	**325 ms**

- pcaGAN outperformed all diffusion *and* cGAN competitors!
- cGANs are 15x to 2500x faster than the diffusion methods

Large-scale image completion example

pcaGAN generates samples that are both high quality and diverse

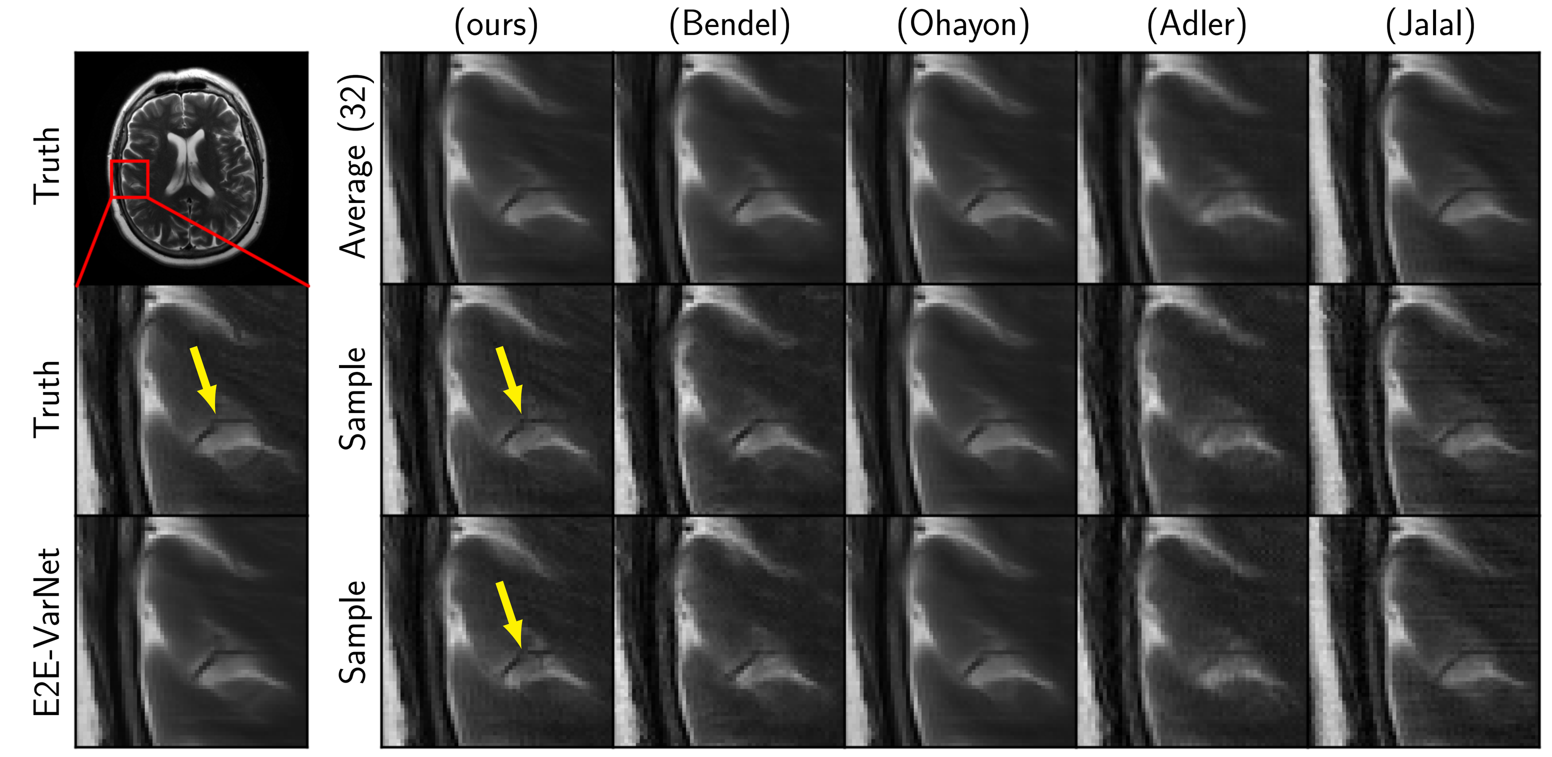


Accelerated MRI reconstruction results

We recovered fastMRI multicoil T2 brain images at acceleration $R = 8$

Model	CFID↓	FID↓	PSNR↑	SSIM↑	LPIPS↓	DISTS↓	Time (4 samples)↓
E2E-VarNet [10]	36.86	44.04	**36.49**	0.9220	0.0575	0.1253	316ms
Langevin (Jalal [11])	48.59	52.62	33.90	0.9137	0.0579	0.1086	14 min
cGAN (Adler [12])	59.94	31.81	33.51	0.9111	0.0614	0.1252	**217 ms**
pscGAN [8]	39.67	43.39	34.92	0.9222	0.0532	0.1128	**217 ms**
rcGAN [2]	24.04	28.43	35.42	0.9257	0.0379	0.0877	**217 ms**
pcaGAN (ours, $K = 1$)	**21.65**	**28.35**	35.94	**0.9283**	**0.0344**	**0.0799**	**217 ms**

- Metrics with * are reported for the optimal averaging constant P



- posterior samples from pcaGAN show meaningful variations without artifacts
- posterior samples from cGAN (Adler), Langevin (Jalal) show unwanted artifacts

Conclusion

We proposed pcaGAN, a novel regularized posterior sampling cGAN and showed that it outperforms cGAN and diffusion competitors in several different inverse problems.

References

[1] Y. Blau and T. Michaeli, "The perception-distortion tradeoff," CVPR, 2018.
[2] M. Bendel, R. Ahmad, and P. Schniter, "A Regularized Conditional GAN for Posterior Sampling in Inverse Problems," NeurIPS 2023.
[3] E. Nehme, O. Yair, and T. Michaeli, "Uncertainty Quantification via Neural Posterior Principal Components," NeurIPS 2023.
[4] M. Soloveitchik, T. Diskin, E. Morin, and A. Wiesel, "Conditional Frechet inception distance," *arXiv:2103.11521*, 2021.
[5] H. Chung, J. Kim, M. T. Mccann, M. L. Klasky, and J.-C. Ye, "Diffusion posterior sampling for general noisy inverse problems," ICLR 2023.
[6] Y. Wang, J. Yu, and J. Zhang, "Zero-Shot Image Restoration Using Denoising Diffusion Null-Space Model," ICLR 2023.
[7] B. Kawar, M. Elad, S. Ermon, and J. Song, "Bahjat Kawar and Michael Elad and Stefano Ermon and Jiaming Song," NeurIPS 2022.
[8] G. Ohayon, T. Adrai, G. Vaksman, M. Elad, and P. Milanfar, "High perceptual quality image denoising with a posterior sampling CGAN," ICCVW 2021.
[9] S.Zhao, J.Cui, Y.Sheng, Y.Dong, X.Liang, E.I.Chang, Y.Xu, "Large scale image completion via co-modulated generative adversarial networks," ICLR 2021.
[10] A. Sriram *et al.*, "End-to-end variational networks for accelerated MRI reconstruction," MICCAI 2020.
[11] A. Jalal, M. Arvinte, G. Daras, E. Price, A. Dimakis, and J. Tamir, "Robust compressed sensing MRI with deep generative priors," NeurIPS 2021.
[12] J. Adler and O. Öktem, "Deep Bayesian inversion," *arXiv:1811.05910*, 2018.