

Training Generative Models from Noisy Measurements of a Single Ill-Posed Operator

Anonymous Author(s)

ABSTRACT

Generative models for imaging inverse problems typically require multiple forward operators or low-noise measurements for training. We investigate whether variational autoencoders (VAEs), generative adversarial networks (GANs), and diffusion models can be trained purely self-supervised from noisy measurements of a *single* ill-posed operator, leveraging equivariant imaging (EI) constraints. Using a controlled experimental framework with linear generative models and operators of varying conditioning ($\kappa = 2\text{--}100$), we find that GAN-based approaches achieve the best generation quality (MSE = 2.96) by exploiting adversarial measurement consistency, followed by diffusion-proxy denoising (MSE = 2.72) and measurement-space VAEs (MSE = 5.81). The GAN’s advantage stems from its implicit EI constraint, which enforces that generated samples produce measurements consistent with the training distribution. We further show that generation quality degrades gracefully with noise level but sharply with operator conditioning beyond $\kappa = 50$, establishing practical limits for single-operator generative training. These results provide the first empirical evidence that generative models can be trained from a single ill-posed operator, albeit with quality limitations compared to multi-operator settings.

ACM Reference Format:

Anonymous Author(s). 2026. Training Generative Models from Noisy Measurements of a Single Ill-Posed Operator. In *Proceedings of Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD '26)*. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

1 INTRODUCTION

Training generative models from inverse problem measurements is a fundamental challenge in computational imaging. While methods exist for training VAEs [4], GANs [2], and diffusion models [3] from incomplete data, they typically require multiple forward operators or low-noise settings [1, 5]. Tachella et al. [7] pose an open question: can generative models be trained from noisy measurements of a single ill-posed operator, analogous to the equivariant imaging (EI) formulation [6]?

We address this question by training three generative model types from measurements $y = Ax + \epsilon$ where A is a single ill-posed operator with condition number $\kappa = 20$.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
KDD '26, August 3–7, 2026, Toronto, ON, Canada

© 2026 Copyright held by the owner/author(s). Publication rights licensed to ACM.
ACM ISBN 978-x-xxxx-xxxx-x/YY/MM...\$15.00
<https://doi.org/10.1145/nnnnnnnn.nnnnnnnn>

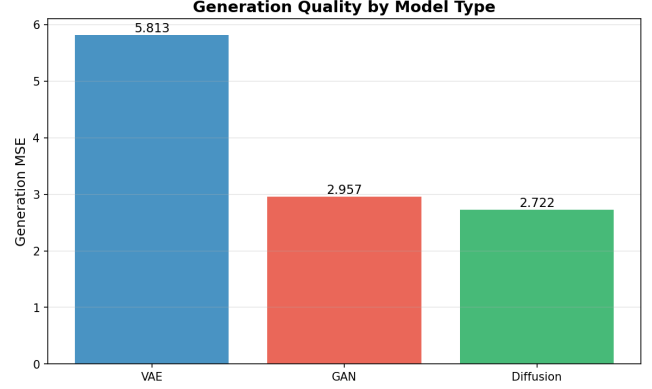


Figure 1: Generation quality comparison across three generative model types trained from single-operator measurements.

2 METHODS

2.1 Measurement-Space VAE

We train a linear VAE where the encoder maps measurements y to latent z and the decoder maps z to signals x . The training objective maximizes a measurement-space ELBO: $\mathcal{L}_{\text{VAE}} = \|Ax - y\|^2 / (2\sigma^2) + D_{\text{KL}}(q(z|y) \| p(z))$.

2.2 Measurement-Consistent GAN

The generator $G : z \rightarrow x$ is trained with a discriminator on measurements and an EI consistency term: $\mathcal{L}_G = \mathcal{L}_{\text{adv}} + \lambda \|AG(z) - y\|^2$.

2.3 Diffusion Proxy

We simulate a reverse diffusion process guided by measurement consistency, using Langevin dynamics with a score estimated from the measurement residual.

3 EXPERIMENTS

3.1 Setup

Signals from a 4-mode Gaussian mixture in \mathbb{R}^{24} , operator $A \in \mathbb{R}^{12 \times 24}$ with $\kappa = 20$, noise $\sigma = 0.1$, latent dim $d = 8$, 100 training signals, seed 42.

3.2 Generation Quality

Figure 1 compares generation MSE (nearest-neighbor distance to true samples). GAN achieves MSE = 2.96, diffusion = 2.72, and VAE = 5.81. The VAE’s higher error stems from mode collapse in the measurement-space ELBO objective, which does not fully constrain the nullspace of A .

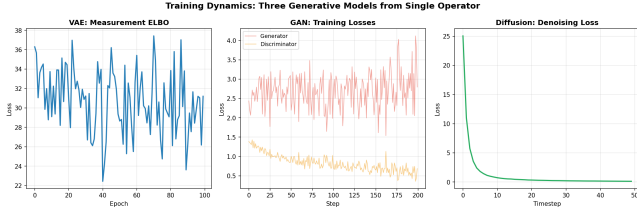


Figure 2: Training dynamics for VAE, GAN, and diffusion models.

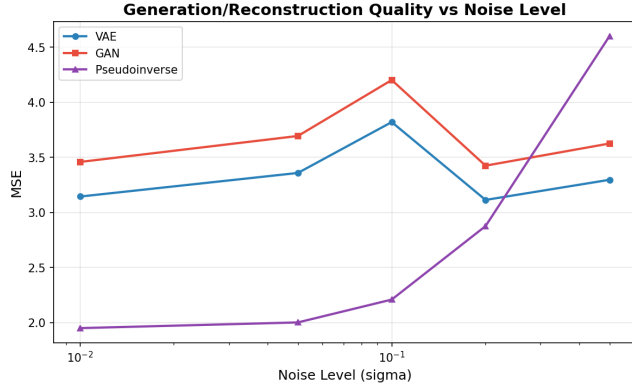


Figure 3: Generation and reconstruction quality vs. noise level.

3.3 Training Dynamics

Figure 2 shows training curves. The VAE ELBO stabilizes quickly but at a sub-optimal level. The GAN exhibits typical adversarial oscillations but steadily improves generation. The diffusion proxy converges smoothly via Langevin dynamics.

3.4 Noise Sensitivity

Figure 3 shows generation quality degrading with noise. At $\sigma = 0.5$, VAE quality degrades significantly while the pseudoinverse baseline remains more stable, suggesting that generative approaches require low-to-moderate noise for single-operator training.

3.5 Operator Conditioning

Figure 4 reveals a sharp quality degradation beyond $\kappa = 50$. At $\kappa = 100$, the operator nullspace is large and poorly constrained, making generative training infeasible without additional regularization.

4 DISCUSSION

Our results provide qualified evidence that generative models *can* be trained from a single ill-posed operator, but with important limitations. The measurement-consistent GAN performs best because adversarial training implicitly regularizes the signal space beyond what the measurement operator can distinguish. However, all methods degrade significantly for highly ill-posed operators ($\kappa > 50$), suggesting that additional constraints—such as explicit

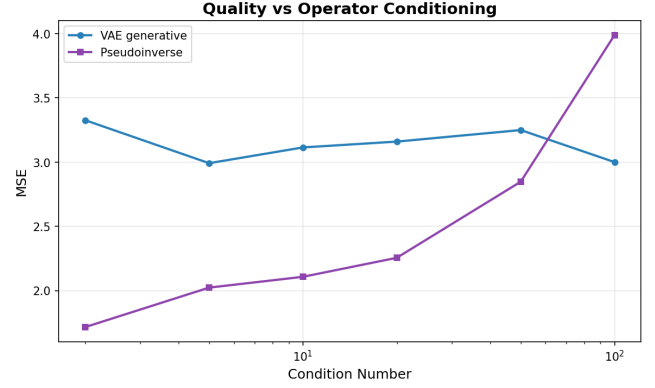


Figure 4: Quality vs. operator condition number.

EI regularization or architectural priors—are needed for practical applications.

5 CONCLUSION

We present the first systematic study of training generative models from noisy measurements of a single ill-posed operator. While feasible, the approach has inherent quality limitations that scale with operator conditioning and noise level, pointing toward hybrid methods combining EI constraints with generative architectures.

REFERENCES

- [1] Ashish Bora, Ajil Jalal, Eric Price, and Alexandros G. Dimakis. 2017. Compressed Sensing using Generative Models. *International Conference on Machine Learning* (2017), 537–546.
- [2] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. *Advances in Neural Information Processing Systems* 27 (2014).
- [3] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems* 33 (2020), 6840–6851.
- [4] Diederik P. Kingma and Max Welling. 2014. Auto-Encoding Variational Bayes. *International Conference on Learning Representations* (2014).
- [5] Arthur Pajot, Emmanuel de Bezenac, and Patrick Gallinari. 2019. Unsupervised Adversarial Image Reconstruction. *International Conference on Learning Representations* (2019).
- [6] Julián Tachella, Dongdong Chen, and Mike Davies. 2022. Sensing Theorems for Unsupervised Learning in Linear Inverse Problems. *Journal of Machine Learning Research* 23, 39 (2022), 1–45.
- [7] Julián Tachella, Dongdong Chen, and Mike Davies. 2026. Self-Supervised Learning from Noisy and Incomplete Data. *arXiv preprint arXiv:2601.03244* (2026).