Compressive Sensing with Augmented Measurements via Generative Self-Distillation

Emmanuel Martinez*, Roman Jacome*, Romario Gualdrón-Hurtado*, Iñaki Esnaola[†] & Henry Arguello*

* Universidad Industrial de Santander, Bucaramanga, Colombia

[†] University of Sheffield, Sheffield, United Kingdom & Princeton University, Princeton, NJ, USA

Abstract-Signal reconstruction from compressed, noisy observations is a ubiquitous challenge in various applications. To address its ill-posed nature, a suitable prior of the underlying signal is required. Generative adversarial networks (GANs) emerge as a natural prior, enabling realistic reconstructions. However, existing approaches either optimize a GAN conditioned on the measurements from scratch or use pre-trained GANs to find images that best fit real measurements. We propose an alternative GAN-based method that, instead of sampling directly from the signal distribution, generates low-dimensional synthetic observations from the real ones. An adversarial selfdistillation strategy optimizes the GAN, extracting meaningful signal information for synthetic measurement generation. These samples form an augmented measurement set, improving the conditioning of compressed sensing solvers, including model-based and deep learning-based methods. We validate our approach on MNIST with a binary sensing matrix for the single-pixel camera, achieving significant improvements in ADMM, PnP-ADMM, and Unrolled ADMM using generated measurements.

Index Terms—Compressive sensing, generative adversarial networks, self-distillation, inverse problems.

I. INTRODUCTION

Compressive sensing (CS) recovers a high-dimensional signal from a few linear random or structured projections thus reducing acquisition time and costs [1]. To recover the underlying signal, some structural assumption (prior) is included to solve the undetermined system of linear equations. Particularly, this theory has been popularized in a wide range of applications including wireless communications [2], [3], radar [4], and imaging [5], [6]. The literature has focused on developing accurate signal priors where the most common choice is that the signal is compressible on a given basis [1]. Several priors have been imposed depending on the signal and applications [7], [8]. Recent advances in deep learning (DL) have revolutionized CS by enabling data-driven approaches that can infer complex signal distributions, where deep neural networks (DNN) have been designed to implicitly learn the signal prior [9], [10] outperforming traditional hand-crafted priors. With the rise of generative models like Generative Adversarial Networks (GANs) [11], these have become effective for recovery by accurately learning complex data distributions,

impacting fields such as signal processing and computer vision [12]-[15]. In CS, they provide implicit priors for recovering high-dimensional signals x from limited measurements y [12], [13]. In the context of CS, signals $\mathbf{x} \in \mathbb{R}^n$ are acquired through linear, coded, and noisy projections usually modeled as $\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\eta}$, where $\mathbf{H} \in \mathbb{R}^{m \times n}$ represents the sensing matrix of the CS acquisition system, $\eta \in \mathbb{R}^m$ represents additive Gaussian noise, and m denotes the number of real measurements $(m \ll n)$. GAN-based approaches in CS typically fall into two categories. One method employs the range of a pre-trained GAN as a natural prior for the signal [12], [13], [16], while another trains the generator and discriminator from scratch by conditioning the generation on the real measurements [14], [15]. Both methods leverage the GAN's ability to capture the underlying data distribution and produce realistic reconstructions, although they often struggle in low-measurement scenarios or when the sensing matrix is poorly conditioned [17].

In this work, we propose an alternative GAN-based recovery method that, instead of directly generating the highdimensional signal x from real measurements y, estimates a novel low-dimensional projection of x to guide recovery methods and enhance performance. Our approach starts with a baseline GAN recovery, producing an initial estimate of the signal, which is then passed through a generator network that maps it to a low-dimensional space. During training, the same network projects the ground truth (GT) into this space, enabling it to learn accurate projections while extracting additional information from the GT, following a self-distillation strategy [18]. Unlike conventional teacher-student setups [19], self-distillation allows a model to refine its own outputs, with techniques like multi-crop training helping capture both global and local features [20]. Applied to CS recovery, this process improves the learned prior, leading to better reconstruction under limited measurements. We define the real and estimated projections as augmented measurements, with GT projections derived using a synthetic sensing matrix S, incoherent with the real sensing matrix H, ensuring diversity and complementary information about x. This approach acts as a plug-in enhancement for recovery methods. For example, in modelbased schemes, it improves the conditioning of the fidelity term. Our contributions are the following:

 A GAN-based method that estimates a low-dimensional representation of x (synthetic measurements), providing a plug-in enhancement for various CS solvers.

This work was supported by ICETEX and MINCIENCIAS through the CTO 2022-0716 Sistema óptico-computacional tipo pushbroom en el rango visible e infrarrojo cercano (VNIR), para la clasificación de frutos cítricos sobre bandas transportadoras mediante aprendizaje profundo, desarrollado en alianza con citricultores de Santander, under Grant 8284. This work was supported by the Vicerrectoría de Investigación y Extensión from Universidad Industrial de Santander under Project 3925.

- A self-distillation adversarial scheme to accurately estimate synthetic measurements by learning a prior from both ground truth and GAN-based estimations.
- Experimental validation of synthetic measurements using three reconstruction methods: i) model-based ADMM [21], ii) plug-and-play (PnP) ADMM [22], and iii) data-driven Unrolling ADMM [23].
 - II. COMPRESSED SENSING BACKGROUND
- A. Model-based recovery

Classical CS recovery solves

$$\mathbf{x}^* \in \operatorname*{arg\,min}_{\mathbf{x}} f(\mathbf{x}) + \lambda h(\mathbf{x}),\tag{1}$$

where $f(\mathbf{x}) = \|\mathbf{y} - \mathbf{H}\mathbf{x}\|_2^2$ measures data fidelity and λ weights the regularizer $h(\mathbf{x})$. We examine ADMM with a sparsity prior $h(\mathbf{x}) = \|\mathbf{x}\|$ [21], [24], and Plug-and-Play [22]. ADMM decouples $f(\mathbf{x})$ and $h(\mathbf{x})$ via an auxiliary variable and dual updates, but degrades when **H** is ill-conditioned.

B. Learning-based recovery

Data-driven recovery leverages large image datasets and deep networks to learn a mapping $\mathcal{M}_{\theta} : \mathbb{R}^m \to \mathbb{R}^n$. Its parameters θ minimize the loss

$$\boldsymbol{\theta}^* \in \operatorname*{arg\,min}_{\boldsymbol{\theta}} \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})}[\mathcal{L}(\mathcal{M}_{\boldsymbol{\theta}}(\mathbf{H}\mathbf{x}), \mathbf{x})], \tag{2}$$

capturing complex image priors. Here \mathcal{M}_{θ} is implemented as an unrolled network [23], [25], where each layer mimics one iteration of a model-based solver with a trainable denoiser.

C. GAN-based recovery

GAN-based recovery approximates $p(\mathbf{x})$ by a learned distribution $p_{\mathcal{R}}$. A generator \mathcal{R} and discriminator $\mathcal{D}_{\mathbf{x}}$ play a min-max game: $\mathcal{D}_{\mathbf{x}}$ must distinguish if an input image comes from the real distribution $\mathbf{x} \sim p(\mathbf{x})$ or the estimated distribution $\hat{\mathbf{x}} \sim p_{\mathcal{R}}$. The adversarial loss is

$$\mathcal{L}_{gan}^{\mathbf{x}}(\mathcal{R}, \mathcal{D}_{\mathbf{x}}) = \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim p(\mathbf{x}, \mathbf{y})} \left[\log \mathcal{D}_{\mathbf{x}}(\mathbf{x}, \mathbf{H}^{\dagger} \mathbf{y}) \right] + \\ \mathbb{E}_{\mathbf{y} \sim p(\mathbf{y})} \left[\log(1 - \mathcal{D}_{\mathbf{x}}(\mathcal{R}(\mathbf{H}^{\dagger} \mathbf{y}), \mathbf{H}^{\dagger} \mathbf{y})) \right].$$
(3)

where $\mathbf{H}^{\dagger} = (\mathbf{H}^{\top}\mathbf{H})^{-1}\mathbf{H}^{\top}$ is the pseudoinverse of \mathbf{H} , and $(\mathbf{x}, \mathbf{y}) \sim p(\mathbf{x}, \mathbf{y})$ with $\mathbf{y} = \mathbf{H}\mathbf{x}$. We add a consistency term

$$\mathcal{L}^{\mathbf{x}}(\mathcal{R}) = \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim p(\mathbf{x}, \mathbf{y})} \left[\|\mathbf{x} - \mathcal{R}(\mathbf{H}^{\dagger} \mathbf{y})\|_{1} \right]$$
(4)

for recovering \mathbf{x} from \mathbf{y} with \mathcal{R} . Thus, the problem becomes

$$\{\mathcal{R}^*, \mathcal{D}^*_{\mathbf{x}}\} \in \underset{\mathcal{R}}{\operatorname{arg\,min}} \underset{\mathcal{D}_{\mathbf{x}}}{\operatorname{arg\,max}} - \mathcal{L}_{gan}^{\mathbf{x}}(\mathcal{R}, \mathcal{D}_{\mathbf{x}}) + \lambda_{\mathbf{x}} \mathcal{L}^{\mathbf{x}}(\mathcal{R}),$$
(5)

where $\lambda_{\mathbf{x}}$ trades off adversarial realism and recovery fidelity.

III. AUGMENTED MEASUREMENTS WITH ADVERSARIAL SELF-DISTILLATION

Our key insight is a GAN-based method to generate an augmented measurement set for any CS solver. This set follows the synthetic forward model

$$\mathbf{g} = \mathbf{S}\mathbf{x},\tag{6}$$

where $\mathbf{g} \in \mathbb{R}^d$ represents the synthetic measurements, and $\mathbf{S} \in \mathbb{R}^{d \times n}$ is the corresponding synthetic sensing matrix.



Fig. 1. **Proposed method.** The framework consists of two stages: *Initial Image Recovery* and *Synthetic Measurements Generation*. Real measurements **y** are mapped to a target signal $\mathbf{x}_0 = \mathcal{R}(\mathbf{H}^{\dagger}\mathbf{y})$ and encoded to obtain synthetic measurements $\tilde{\mathbf{g}}$, while the GT is encoded with the same \mathcal{E} to obtain $\hat{\mathbf{g}}$. An adversarial self-distillation scheme optimizes the models, refining synthetic measurements for enhanced recovery. Discriminators $\mathcal{D}_{\mathbf{x}}$ and $\mathcal{D}_{\mathbf{g}}$ evaluate whether *reconstructed images* and *synthetic measurements* are real or generated, guiding the learning process toward more realistic representations.

By integrating both known and synthetic measurements, we construct an augmented sensing matrix

$$\mathbf{A} = \begin{bmatrix} \mathbf{H} \\ \mathbf{S} \end{bmatrix} \in \mathbb{R}^{(m+d) \times n},\tag{7}$$

where $m + d \le n$ and $rank(\mathbf{A}) \le min(m + d, n)$ ensuring highly incoherent synthetic sampling.

The first stage of our method builds upon traditional GANbased image restoration, as described in Subsection II-C. We start by using a rough estimate $\mathbf{x} \approx \mathbf{x}_0 = \mathcal{R}(\mathbf{H}^{\dagger}\mathbf{y})$. Next, we design an encoder network $\mathcal{E} : \mathbb{R}^n \to \mathbb{R}^d$ that maps information from the image space to the synthetic measurement space as $\tilde{\mathbf{g}} = \mathcal{E}(\mathbf{x}_0)$. The encoder is also used to compute $\hat{\mathbf{g}} = \mathcal{E}(\mathbf{x})$ using the GT \mathbf{x} during training. To ensure the generated samples align with real distributions, we employ two discriminator models: $\mathcal{D}_{\mathbf{x}}$ and $\mathcal{D}_{\mathbf{g}}$. These discriminators assess whether an image \mathbf{x} or synthetic measurement \mathbf{g} is real or synthetic. The proposed framework integrates two adversarial losses to capture information from the image distribution $p(\mathbf{x})$ and the synthetic measurement distribution $p(\mathbf{g})$. The adversarial loss associated with $p(\mathbf{x})$ is given in Eq. (3), while the adversarial loss for $p(\mathbf{g})$ is formulated as

$$\mathcal{L}_{gan}^{\mathbf{g}}(\mathcal{R}, \mathcal{E}, \mathcal{D}_{\mathbf{g}}) = \mathbb{E}_{\mathbf{x}, \mathbf{g} \sim p(\mathbf{x}, \mathbf{g})} \left[\log \mathcal{D}_{\mathbf{g}}(\mathbf{S}^{\dagger} \mathbf{g}, \mathbf{x}) \right] + \\ \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim p(\mathbf{x}, \mathbf{y})} \left[\log(1 - \mathcal{D}_{\mathbf{g}}(\mathbf{S}^{\dagger} \mathcal{E}(\mathcal{R}(\mathbf{H}^{\dagger} \mathbf{y})), \mathbf{x})) \right] + \\ \mathbb{E}_{\mathbf{x} \sim p(\mathbf{x})} \left[\log(1 - \mathcal{D}_{\mathbf{g}}(\mathbf{S}^{\dagger} \mathcal{E}(\mathbf{x}), \mathbf{x})) \right],$$
(8)

where $p(\mathbf{x}, \mathbf{g})$ represents the joint distribution for $\mathbf{g} = \mathbf{S}\mathbf{x}$. This adversarial loss applies self-distillation over $p(\mathbf{g})$, allowing \mathcal{E} to refine its mapping using both estimated and GT samples, improving \mathbf{g} estimation. The recovery model, encoder, and discriminator are optimized to approximate $p(\mathbf{g})$, ensuring realistic and diverse generated samples. To further refine the model, consistency terms (CTs) are introduced. The signal CT from Eq. (4) enforces \mathcal{R} to reconstruct \mathbf{x} from synthetic measurements \mathbf{g} . Additionally, the measurement CT

$$\mathcal{L}^{\mathbf{g}}(\mathcal{R}, \mathcal{E}) = \mathbb{E}_{\mathbf{g}, \mathbf{y} \sim p(\mathbf{g}, \mathbf{y})} \left[\| \mathbf{g} - \mathcal{E}(\mathcal{R}(\mathbf{H}^{\dagger}\mathbf{y})) \|_{1} \right] + \\ \mathbb{E}_{\mathbf{x}, \mathbf{g} \sim p(\mathbf{x}, \mathbf{g})} \left[\| \mathbf{g} - \mathcal{E}(\mathbf{x}) \|_{1} \right] + \\ \mathbb{E}_{\mathbf{x}, \mathbf{y} \sim p(\mathbf{x}, \mathbf{y})} \left[\| \mathcal{E}(\mathbf{x}) - \mathcal{E}(\mathcal{R}(\mathbf{H}^{\dagger}\mathbf{y})) \|_{1} \right],$$
(9)

ensures accurate synthetic measurement estimation while maintaining consistency with the GT signal.

TABLE I

Recovery results comparing the quantity of synthetic measurements for GAN-based recovery. The best PSNR, SSIM, FID, and LPIPS results are highlighted in <u>Green underline</u>, **BOLD BLUE**, and <u>Teal underline</u>, **BOLD orange**, respectively.

Compression ratio	PSNR ↑			SSIM ↑			FID ↓			LPIPS ↓			
m/n d/n	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	
0 (Baseline)	14.82	<u>24.17</u>	27.18	0.7430	0.9545	0.9768	10.6	5.49	4.46	0.0399	0.0078	0.0046	
0.25	14.76	24.14	26.82	0.7402	0.9534	0.9746	9.51	<u>5.45</u>	4.91	0.0402	0.0077	0.0048	
0.5	<u>14.83</u>	24.12	27.07	0.7432	0.9539	0.9758	10.2	6.02	<u>3.22</u>	0.0401	0.0080	0.0043	
0.75	14.68	23.89	26.94	0.7392	0.9483	0.9751	<u>8.64</u>	8.11	3.90	0.0399	0.0082	0.0045	
1-m/n	14.71	24.03	26.96	0.7394	0.9536	0.9757	10.5	8.67	4.71	0.0397	0.0082	0.0047	

Finally, the overall optimization problem is formulated as

$$\{\mathcal{R}^*, \mathcal{E}^*, \mathcal{D}^*_{\mathbf{x}}, \mathcal{D}^*_{\mathbf{g}}\} \in \underset{\mathcal{R}, \mathcal{E}}{\arg\min} \underset{\mathcal{D}_{\mathbf{x}}, \mathcal{D}_{\mathbf{g}}}{\arg\max} \left[-\mathcal{L}^{\mathbf{x}}_{gan}(\mathcal{R}, \mathcal{D}_{\mathbf{x}}) -\mathcal{L}^{\mathbf{g}}_{gan}(\mathcal{R}, \mathcal{E}, \mathcal{D}_{\mathbf{g}}) + \lambda_{\mathbf{x}} \mathcal{L}^{\mathbf{x}}(\mathcal{R}) + \lambda_{\mathbf{g}} \mathcal{L}^{\mathbf{g}}(\mathcal{R}, \mathcal{E}) \right], \quad (10)$$

where $\lambda_{\mathbf{x}}$ and $\lambda_{\mathbf{g}}$ balance the consistency terms.

The proposed method has several design criteria to take into account. In this work, we employed random binary sensing matrices, i.e., $\mathbf{H} \in \{0, 1\}^{m \times n}$ where each entry of the matrix follows a Bernoulli distribution with p = 0.5. We chose a synthetic matrix **S** from the same distribution.

The selection is based on the properties of this kind of matrix, which has a high probability of full-row rank $\mathbb{E}[\operatorname{rank}(\mathbf{A})] \approx \min(m+d,n)$ [26]. Nevertheless, more robust synthetic matrix design methods can be used, which can be adapted from existing sensing matrix optimization [27]–[29] with the facility that this matrix is not limited by any physical constraint of a real acquisition system. Additionally, one crucial aspect is the choice of d as we can freely select this value since this is a virtual sensing matrix. In the following section, we show that this hyperparameter is crucial for obtaining a better recovery of x with the augmented measurement set.

IV. EXPERIMENTS & RESULTS

The proposed method is evaluated using the Single-Pixel Camera (SPC) system [30], which employs a random binary sensing matrix $\mathbf{A} = [\mathbf{H}^{\top}, \mathbf{S}^{\top}]^{\top} \in \{0, 1\}^{(m+d) \times n}$. The evaluation is conducted on the MNIST dataset, split into 48000 training, 12000 validation, and 10000 testing images, all resized to 32×32 , thus n = 1024. The method is tested under various real

Algorithm 1 ADMM Algorith	
Require: A, y, $\hat{\mathbf{g}}$, $\rho > 0$, K, μ ,	$\frac{1}{\mathcal{P}(\cdot)}$
1: $\mathbf{b} \leftarrow [\mathbf{y}^{\top}, \hat{\mathbf{g}}^{\top}]^{\top}$	
2: $\mathbf{x}^{(0)} \leftarrow \mathbf{A}^{\top} \mathbf{b}$	> Initialization from backprojection
3: $\mathbf{z}^{(0)}, \ \mathbf{u}^{(0)} \leftarrow 0$	
4: for $k = 1$ to K_{do}	
5: $\tilde{\mathbf{x}}^{(k)} \leftarrow (\mathbf{A}^{\top}\mathbf{A} + \rho \mathbf{I})^{-}$	×
6: $ \begin{bmatrix} \mathbf{A}^{T}\mathbf{b} + \mu \left(\mathbf{z}^{(k-1)} - \mathbf{u}\right) \end{bmatrix} $	$\mathbf{u}^{(k-1)}$]. \triangleright Fidelity step
7: $\mathbf{z}^{(k)} = \mathcal{P}(\mathbf{x}^{(k)} + \mathbf{u}^{(k-1)})$) \triangleright Proximal step
8: $\mathbf{u}^{(k)} = \mathbf{u}^{(k-1)} + (\mathbf{x}^{(k)})$	$-\mathbf{z}^{(k)}$ \triangleright Dual update
9: return $\mathbf{z}^{(K)}$	

compression ratios (CRs) $m/n = \{0.01, 0.05, 0.1\}$ and synthetic CRs $d/n = \{0, 0.25, 0.5, 0.25, 1 - m/n\}$ to analyze the trade-off between measurement reduction and reconstruction quality. The proposed GAN framework is inspired by [31], where \mathcal{R} and \mathcal{E} serve as U-Net-based generators [32]. The key difference is that \mathcal{E} includes an additional linear layer that projects the penultimate feature map onto the synthetic measurement space. Discriminators $\mathcal{D}_{\mathbf{x}}$ and $\mathcal{D}_{\mathbf{g}}$ employ a patch-based architecture, producing a small feature map that penalizes structural inconsistencies at the patch level [31]. Once trained, real and synthetic measurements are merged into the augmented measurement set. For the GAN optimization, we use the Adam optimizer [33], with a learning rate for the generator and discriminator of 0.001 and 0.0008, respectively. We apply a batch size of 256 samples and train for 200 epochs. The method was implemented using the PyTorch Lightning framework [34]. This set is validated through ADMMbased methods. Specifically, classical ADMM ($\mathcal{P}(\cdot)$ is a soft-thresholding operator in the DCT domain), PnP-ADMM $(\mathcal{P}(\cdot))$ is a pre-trained denoiser, specifically the Restormer [35] implemented in the DeepInv library [36]). For these methods, the number of iterations was set to 100 and the parameter $\mu = 1$. Unrolled ADMM, where $\mathcal{P}(\cdot)$ is a convolutional network learned for each iteration, was set to K = 5 stages and trained for 200 epochs using Adam optimizer with a learning rate of 0.001 and a ℓ_2 loss function with optimization problem in 2. We summarize the adaptation of those ADMM families in the Alg. 1. Finally, we assess the proposed method using consistency and generative metrics. PSNR and SSIM [37] evaluate pixel-level fidelity and structural similarity, while FID and LPIPS [38] measure distributional differences and perceptual similarity in deep feature space.

A. Generative Self-Distilled Augmented Measurements results

We evaluate image recovery and synthetic measurement estimation to assess the proposed method. Table I reports recovery performance in PSNR, SSIM, FID, and LPIPS. The baseline corresponds to the initial recovery method (Section II-C) with d/n = 0. The proposed method achieves similar PSNR and SSIM but outperforms the baseline in FID and LPIPS, especially at the lowest CR (m/n = 0.01), indicating a better approximation of $p(\mathbf{x})$ in highly constrained scenarios.

Additionally, we compare synthetic measurements $S\mathcal{R}^*(\mathbf{y})$ from the baseline with estimated $\hat{\mathbf{g}}$ using the squared ℓ_2 -

TABLE II

Recovery results comparing the quantity of synthetic measurements for ADMM, PnP-ADMM, and Unrolled ADMM. The best PSNR results are highlighted in <u>Bold Blue</u>.

Compression ratio	ADMM						PnP-ADMM						Unrolled ADMM					
	PSNR ↑			SSIM ↑		PSNR ↑			SSIM ↑			PSNR ↑			SSIM ↑			
m/n d/n	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1	0.01	0.05	0.1
0 (Baseline)	10.75	10.44	9.81	0.02	0.03	0.06	10.71	11.78	13.67	0.03	0.17	0.43	14.85	19.88	24.43	0.46	0.64	0.74
0.25	9.87	11.58	13.02	0.10	0.22	0.31	14.72	20.53	22.07	0.40	0.52	0.55	15.37	<u>23.26</u>	25.03	0.57	0.72	0.71
0.5	14.32	20.28	21.98	0.38	0.58	0.61	15.26	<u>22.33</u>	23.77	0.47	0.61	0.61	15.47	23.19	<u>25.06</u>	0.60	0.70	0.71
0.75	15.30	22.29	23.92	0.46	0.69	0.70	15.40	22.10	23.14	0.49	0.62	0.62	15.51	22.80	24.52	0.53	0.67	0.69
1-m/n	<u>15.44</u>	<u>22.71</u>	<u>24.19</u>	0.55	0.73	0.72	<u>15.40</u>	22.07	22.81	0.52	0.65	0.62	<u>15.51</u>	22.78	24.22	0.56	0.68	0.67



Fig. 2. Synthetic measurement consistency in terms of squared ℓ_2 -norm. The comparison between the baseline described in Subsection II-C and the proposed method is shown in the same color, with the same colors in clear and dark modes representing the baseline and proposed method, respectively.

norm (Fig. 2). Results show improved consistency across all scenarios, highlighting the potential of synthetic measurements to enhance ADMM-based recovery with the *augmented measurement set*, discussed in the next section.

B. ADMM-based recovery method results

We assess the effectiveness of the synthetic measurements by evaluating ADMM algorithms, with results reported in Tab. II. The baseline corresponds to the recovery performance of each method using only the real measurements (d/n = 0). Results are presented in terms of PSNR and SSIM, as these are the most relevant metrics for assessing image recovery consistency. The proposed method achieves higher performance in highly constrained scenarios, particularly with classical and PnP-ADMM recovery methods. The most challenging case is Unrolled ADMM, where a slight improvement is observed. This suggests that further enhancement in synthetic measurement estimation is possible by refining the model design and imposed priors, which will be explored in future work. Quantitative results for two handwritten samples (4 and 6) are shown in Fig. 3. The proposed method outperforms the baseline both quantitatively and qualitatively, as the baseline fails to reconstruct the images due to insufficient real measurements. The limitation in highly constrained scenarios is effectively mitigated using the proposed approach.

V. DISCUSSION AND FUTURE WORK

Our encoder \mathcal{E} , trained via adversarial self-distillation, produces synthetic measurements that outperform direct projections by leveraging learned priors beyond the limited-fidelity



Fig. 3. Two test samples of the MNIST dataset and their reconstruction with ADMM-based recovery methods. d/n = 0 for the baseline.

signal x_0 . Future work should explore mutual-information maximization to align synthetic and real measurement distributions better, reduce artifacts, and improve robustness, along with regularization strategies and formal guarantees across CS scenarios. Although developed for CS, this framework also applies to tasks like medical imaging, hyperspectral reconstruction, and wireless channel estimation, where measurement limitations degrade quality. Future research should investigate synthetic measurement generation, domain-specific priors in the encoder, and the transferability of synthetic encoders across applications.

VI. CONCLUSIONS

We propose a generative adversarial network with selfdistillation for synthetic measurement augmentation. Unlike traditional recovery methods that rely solely on real measurements, our approach estimates additional synthetic measurements and integrates them into the reconstruction process, improving recovery algorithms. This augmentation mitigates the limitations of low compression ratios, leading to more accurate image reconstruction. Extensive experiments show that our method produces synthetic measurements with a lower squared ℓ_2 -norm error than the baseline. Incorporating these into ADMM-based recovery methods significantly improves performance, especially when real measurements are limited. These findings highlight the potential of synthetic measurements to enhance reconstruction quality.

REFERENCES

- E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 21– 30, 2008.
- [2] Z. Qin, H. Ye, G. Y. Li, and B.-H. F. Juang, "Deep learning in physical layer communications," *IEEE Wireless Communications*, vol. 26, no. 2, pp. 93–99, 2019.
- [3] H. He, S. Jin, C.-K. Wen, F. Gao, G. Y. Li, and Z. Xu, "Modeldriven deep learning for physical layer communications," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 77–83, 2019.
- [4] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin, "Sparsity and compressed sensing in radar imaging," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1006–1020, 2010.
- [5] R. M. Willett, R. F. Marcia, and J. M. Nichols, "Compressed sensing for practical optical imaging systems: a tutorial," *Optical Engineering*, vol. 50, no. 7, pp. 072 601–072 601, 2011.
- [6] G. R. Arce, D. J. Brady, L. Carin, H. Arguello, and D. S. Kittle, "Compressive coded aperture spectral imaging: An introduction," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 105–115, 2013.
- [7] B. Jin, P. Maaß, and O. Scherzer, "Sparsity regularization in inverse problems," *Inverse Problems*, vol. 33, no. 6, 2017.
- [8] W. Dong, G. Shi, X. Li, Y. Ma, and F. Huang, "Compressive sensing via nonlocal low-rank regularization," *IEEE transactions on image* processing, vol. 23, no. 8, pp. 3618–3632, 2014.
- [9] J. Scarlett, R. Heckel, M. R. Rodrigues, P. Hand, and Y. C. Eldar, "Theoretical perspectives on deep learning methods in inverse problems," *IEEE journal on selected areas in information theory*, vol. 3, no. 3, pp. 433–453, 2022.
- [10] G. Ongie, A. Jalal, C. A. Metzler, R. G. Baraniuk, A. G. Dimakis, and R. Willett, "Deep learning techniques for inverse problems in imaging," *IEEE Journal on Selected Areas in Information Theory*, vol. 1, no. 1, pp. 39–56, 2020.
- [11] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.
- [12] V. Shah and C. Hegde, "Solving linear inverse problems using gan priors: An algorithm with provable guarantees," in 2018 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, 2018, pp. 4609–4613.
- [13] D. V. Patel, D. Ray, and A. A. Oberai, "Solution of physics-based bayesian inverse problems with deep generative priors," *Computer Methods in Applied Mechanics and Engineering*, vol. 400, p. 115428, 2022.
- [14] X. Ouyang, Y. Chen, K. Zhu, and G. Agam, "Image restoration refinement with uformer gan," in *Proceedings of the IEEE/CVF Conference* on Computer Vision and Pattern Recognition, 2024, pp. 5919–5928.
- [15] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, Y. Qiao, and C. Change Loy, "Esrgan: Enhanced super-resolution generative adversarial networks," in *Proceedings of the European conference on computer* vision (ECCV) workshops, 2018, pp. 0–0.
- [16] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models," in *International conference on machine learning*. PMLR, 2017, pp. 537–546.
- [17] A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, "An estimate for the condition number of a matrix," *SIAM Journal on Numerical Analysis*, vol. 16, no. 2, pp. 368–375, 1979.
- [18] L. Zhang, C. Bao, and K. Ma, "Self-distillation: Towards efficient and compact neural networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 8, pp. 4388–4403, 2021.
- [19] J. Gou, B. Yu, S. J. Maybank, and D. Tao, "Knowledge distillation: A survey," *International Journal of Computer Vision*, vol. 129, no. 6, pp. 1789–1819, 2021.
- [20] M. Caron, H. Touvron, I. Misra, H. Jégou, J. Mairal, P. Bojanowski, and A. Joulin, "Emerging properties in self-supervised vision transformers," in *Proceedings of the International Conference on Computer Vision* (ICCV), 2021.
- [21] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein *et al.*, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends*® *in Machine learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [22] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play admm for image restoration: Fixed-point convergence and applications," *IEEE Transactions on Computational Imaging*, vol. 3, no. 1, pp. 84–98, 2016.

- [23] J. Sun, H. Li, Z. Xu et al., "Deep admm-net for compressive sensing mri," Advances in neural information processing systems, vol. 29, 2016.
- [24] R. Tibshirani, "Regression shrinkage and selection via the lasso," Journal of the Royal Statistical Society Series B: Statistical Methodology, vol. 58, no. 1, pp. 267–288, 1996.
- [25] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Processing Magazine*, vol. 38, no. 2, pp. 18–44, 2021.
- [26] C. Cooper, A. Frieze, and W. Pegden, "On the rank of a random binary matrix," in *Proceedings of the Thirtieth Annual ACM-SIAM Symposium* on Discrete Algorithms. SIAM, 2019, pp. 946–955.
- [27] H. Arguello and G. R. Arce, "Colored coded aperture design by concentration of measure in compressive spectral imaging," *IEEE Transactions* on *Image Processing*, vol. 23, no. 4, pp. 1896–1908, 2014.
- [28] L. Zelnik-Manor, K. Rosenblum, and Y. C. Eldar, "Sensing matrix optimization for block-sparse decoding," *IEEE Transactions on Signal Processing*, vol. 59, no. 9, pp. 4300–4312, 2011.
- [29] G. Li, Z. Zhu, D. Yang, L. Chang, and H. Bai, "On projection matrix optimization for compressive sensing systems," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2887–2898, 2013.
- [30] M. F. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk, "Single-pixel imaging via compressive sampling," *IEEE signal processing magazine*, vol. 25, no. 2, pp. 83–91, 2008.
- [31] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *Proceedings of the IEEE* conference on computer vision and pattern recognition, 2017, pp. 1125– 1134.
- [32] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18.* Springer, 2015, pp. 234–241.
- [33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings, Y. Bengio and Y. LeCun, Eds., 2015.
- [34] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga *et al.*, "Pytorch: An imperative style, high-performance deep learning library," *Advances in neural information processing systems*, vol. 32, 2019.
- [35] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang, "Restormer: Efficient transformer for high-resolution image restoration," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 5728–5739.
- [36] J. Tachella, D. Chen, S. Hurault, M. Terris, and A. Wang, "DeepInverse: A deep learning framework for inverse problems in imaging," Jun. 2023. [Online]. Available: https://github.com/deepinv/deepinv
- [37] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions* on *Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [38] J. An, Z. Yang, J. Wang, L. Li, Z. Liu, L. Wang, and J. Luo, "Bring metric functions into diffusion models," arXiv preprint arXiv:2401.02414, 2024.