PtyGenography: using generative models for regularization of the phase retrieval problem

Selin Aslan
Department of Mathematics
Koç University
saslan@ku.edu.tr

Tristan van Leeuwen
Computational Imaging
Centrum Wiskunde & Informatica (CWI)
t.van.Leeuwen@cwi.nl

Allard Mosk
Debye Institute
& Department of Physics
Utrecht University
a.p.mosk@uu.nl

Palina Salanevich
Mathematical Institute
Utrecht University
p.salanevich@uu.nl

***Abstract*—In phase retrieval and similar inverse problems, the stability of solutions across different noise levels is crucial for applications. One approach to promote it is using signal priors in a form of a generative model as a regularization, at the expense of introducing a bias in the reconstruction. In this paper, we explore and compare the reconstruction properties of classical and generative inverse problem formulations. We propose a new unified reconstruction approach that mitigates overfitting to the generative model for varying noise levels.**

***Index Terms*—phase retrieval, inverse problems, regularization, generative priors**

I. Introduction

In many signal processing applications in imaging, acoustics, and optics, we aim to reconstruct an unknown signal from its intensity measurements. This inverse problem is known as the *phase retrieval problem*. In some applications, such as diffraction imaging [5], [18] and ptychography [22], it arises due to the limitations of the measurement process. Namely, the detectors are unable to capture the phases of the measurements, but only their magnitudes. In other applications, including audio processing [2], [20], [23], the phases of the measurements may be too noisy to be used in the reconstruction process.

Current technological advances and the pressing need for fast, high-fidelity methods in imaging and audio processing inspired active interest in phase retrieval. In recent years, substantial progress has been made in the mathematical foundations of this problem [8], [10]. However, there is still a significant gap between the available theoretical reconstruction guarantees and the application restrictions and requirements. As such, it is often desired to achieve stable reconstruction while facing high measurement noise and a limited measurement budget that is smaller than the minimal number of measurements theoretically required for reconstruction [12].

More formally, we consider a non-linear inverse problem of retrieving $\mathbf{f} \in \mathbb{C}^n$ from noisy measurements $\mathbf{y} = \mathcal{A}(\mathbf{f}) + \boldsymbol{\varepsilon}$, where $\mathcal{A}\colon \mathbb{C}^n \to \mathbb{C}^m$ is the measurement map, and $\boldsymbol{\varepsilon} \in \mathbb{C}^m$ is the noise term. In the case of the phase retrieval problem, the measurement map is defined as $\mathcal{A}(\mathbf{f}) = |A\mathbf{f}|^2$, where $A \in \mathbb{C}^{m\times n}$ is a given measurement matrix and $|\cdot|^2$ denotes the element-wise squared absolute value of a vector. Note that in this paper, we only consider the additive noise model, which is typical for detector noise and does not cover Poisson shot noise. Assuming $\mathcal{A}$ is injective, one can formulate the reconstruction problem as

$$\min_{\mathbf{f}} \|\mathcal{A}(\mathbf{f}) - \mathbf{y}\|_2^2. \tag{1}$$

To account for instabilities in the presence of noise, one often adds a regularization term. The most common classical regularization approach is Tikhonov regularization, which is theoretically well-understood [26]. However, it tends to smooth high-frequency components, which can be problematic when trying to capture detailed signal features.

An alternative budding approach to enforce stability in the presence of noise is incorporating the available prior information on the signal into the reconstruction process. This reduces the set of admissible signals and thus can drastically improve reconstruction accuracy and the number of phaseless measurements required for reconstruction.

Inspired by the advances in the field of compressive sensing, various works have explored sparsity priors on the signal [9], [10], [14]. However, while in many applications, the set of admissible signals has a lower-dimensional structure, it is not always the case that it can be approximated by a sparse representation. Following the recent remarkable success of neural networks in phase retrieval applications [7], [21], [25], a new framework for modeling the set of admissible signals as the range $\{\mathbf{f} = \mathcal{G}(\mathbf{z})\colon \mathbf{z} \in \mathbb{R}^k\}$ of a (deep) generative neural network $\mathcal{G}\colon \mathbb{C}^k \to \mathbb{C}^n$ with $k \ll n$ was proposed in [11]. In this paper, the authors studied the case of a generic measurement frame and an untrained neural network $\mathcal{G}$ with random parameters. In the case of structured, application-relevant measurements, the injectivity of phase retrieval from power spectra (squared absolute values of the oversampled Fourier measurements) was recently established in a more general setting of (generic) semi-algebraic generative models [4].

Expressing the unknown signal as $\mathbf{f} = \mathcal{G}(\mathbf{z})$, the problem (1) becomes

$$\min_{\mathbf{z}} \|\mathcal{A} \circ \mathcal{G}(\mathbf{z}) - \mathbf{y}\|_2^2. \tag{2}$$

The authors would like to thank the Lorentz Center for their support during PRiMA workshop, where this project has been initiated. PS is supported by NWO Talent program Veni ENW grant, file number VI.Veni.212.176.

The rationale here is that the conditioning of $\mathcal{A} \circ \mathcal{G}$ is more favorable than that of $\mathcal{A}$, at the expense of introducing a bias in the reconstruction. It has indeed been observed in numerical experiments that for high signal-to-noise ratio levels formulation (1) performs better, while in the case of low signal-to-noise ratio formulation (2) outperforms it [24].

In this paper, we aim to characterize the reconstruction error of (1) and (2) in terms of the bias $\min_{\mathbf{z}} \|\mathcal{G}(\mathbf{z}) - \mathbf{f}\|_2$ and variance $\|\boldsymbol{\varepsilon}\|_2$ and propose a practical way to detect if a significant bias has been introduced by the generative model. We furthermore propose a unified variational approach that bridges the two formulations (1) and (2) and performs well for different noise levels.

The remaining part of this paper is organized as follows. In Section II-A, we study the recovery properties of the two inverse problem formulations (1) and (2) and discuss the role of the bias introduced by a generative model in Section II-B. In Section II-C we propose and evaluate a unified approach and present numerical results supporting our findings in Section III. Finally, Section IV contains the discussion of further research directions.

II. Reconstruction approach

Let us assume that both the measurement map $\mathcal{A}$ and the generative model $\mathcal{G}$ are injective and bi-Lipschitz with constants $\alpha, \beta \geq 1$. That is, for any $\mathbf{f}, \mathbf{f}' \in \mathbb{C}^n$,

$$\alpha^{-1}\|\mathbf{f} - \mathbf{f}'\|_2 \leq \|\mathcal{A}(\mathbf{f}) - \mathcal{A}(\mathbf{f}')\|_2 \leq \alpha\|\mathbf{f} - \mathbf{f}'\|_2,$$

and likewise for $\mathcal{G}$, with β in place of α.

Remark 1. *Note that in the case of the phase retrieval problem, that is, when $\mathcal{A}(\mathbf{f}) = |A\mathbf{f}|^2$, for a matrix $A \in \mathbb{C}^{m\times n}$, the measurement map $\mathcal{A}$ satisfies the (bi-)Lipschitz property whenever it is injective [3]. However, for most matrices A, no bound on the Lipschitz constant α is known. One exception to this is phase retrieval from locally supported measurements [13].*

We furthermore assume that the bi-Lipschitz constant γ of $\mathcal{A} \circ \mathcal{G}$ is more favorable than that of $\mathcal{A}$, that is, $\gamma < \alpha$. In other words, $\mathcal{G}$ effectively regularities the inverse problem. Moreover, we assume that $\mathcal{G}$ is well-conditioned, in the sense that $0 < \beta - 1 \ll 1$.

A. Characterizing the reconstruction error

We denote by $\mathbf{f_0}$ the ground-truth signal we aim to recover from the measurements

$$\mathbf{y} = \mathcal{A}(\mathbf{f_0}) + \boldsymbol{\varepsilon},$$

where $\boldsymbol{\varepsilon}$ is a bounded noise term. We can now formulate the following results regarding the reconstructing errors resulting from (1) and (2).

Lemma 1. *Let $\tilde{\mathbf{f}} = \arg\min_{\mathbf{f}} \|\mathcal{A}(\mathbf{f}) - \mathbf{y}\|_2$ with $\mathbf{y} = \mathcal{A}(\mathbf{f_0}) + \boldsymbol{\varepsilon}$. Then the reconstruction error is given by*

$$\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq 2\alpha\|\boldsymbol{\varepsilon}\|_2.$$

Proof. First, note that for any $\mathbf{f} \in \mathbb{C}^n$,

$$\|\mathcal{A}(\tilde{\mathbf{f}}) - \mathbf{y}\|_2 \leq \|\mathcal{A}(\mathbf{f}) - \mathbf{y}\|_2 \leq \|\mathcal{A}(\mathbf{f}) - \mathcal{A}(\mathbf{f_0})\|_2 + \|\boldsymbol{\varepsilon}\|_2.$$

By letting $\mathbf{f} = \mathbf{f_0}$, we get

$$\|\mathcal{A}(\tilde{\mathbf{f}}) - \mathbf{y}\|_2 \leq \|\boldsymbol{\varepsilon}\|_2.$$

Now, using the reverse triangle inequality and the bi-Lipschitz property of $\mathcal{A}$, we obtain

$$\begin{aligned}\|\mathcal{A}(\tilde{\mathbf{f}}) - \mathbf{y}\|_2 &= \|\mathcal{A}(\tilde{\mathbf{f}}) - \mathcal{A}(\mathbf{f_0}) - \boldsymbol{\varepsilon}\|_2 \\ &\geq \|\mathcal{A}(\tilde{\mathbf{f}}) - \mathcal{A}(\mathbf{f_0})\|_2 - \|\boldsymbol{\varepsilon}\|_2 \\ &\geq \alpha^{-1}\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 - \|\boldsymbol{\varepsilon}\|_2.\end{aligned}$$

Combining the two inequalities yields

$$\alpha^{-1}\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq 2\|\boldsymbol{\varepsilon}\|_2,$$

which gives the desired result. □

Lemma 2. *Define $\tilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathcal{A} \circ \mathcal{G}(\mathbf{z}) - \mathbf{y}\|_2$ with $\mathbf{y} = \mathcal{A}(\mathbf{f_0}) + \boldsymbol{\varepsilon}$, and $\tilde{\mathbf{f}} = \mathcal{G}(\tilde{\mathbf{z}})$. Then the reconstruction error is bounded by*

$$\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq (1 + 2\alpha\beta\gamma)\|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2 + 2\beta\gamma\|\boldsymbol{\varepsilon}\|_2,$$

where $\mathbf{z_0} = \arg\min_{\mathbf{z}} \|\mathcal{G}(\mathbf{z}) - \mathbf{f_0}\|_2$.

Proof. The first step is to upper bound the residual in terms of the bias and variance. By definition of $\mathbf{y}$ and using the triangular inequality, we have that for any $\mathbf{z} \in \mathbb{C}^k$,

$$\|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{y}\|_2 \leq \|\mathcal{A} \circ \mathcal{G}(\mathbf{z}) - \mathcal{A}(\mathbf{f_0})\|_2 + \|\boldsymbol{\varepsilon}\|_2.$$

Picking $\mathbf{z} = \mathbf{z_0}$ and utilizing the bi-Lipschitz property of $\mathcal{A}$ then yields

$$\|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{y}\|_2 \leq \alpha\|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2 + \|\boldsymbol{\varepsilon}\|_2.$$

The second step is to lower-bound the residual in terms of the error $\|\tilde{\mathbf{z}} - \mathbf{z}\|_2$. We split the residual $\|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{y}\|_2$ as

$$\|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathcal{A} \circ \mathcal{G}(\mathbf{z_0}) + \mathcal{A} \circ \mathcal{G}(\mathbf{z_0}) - \mathcal{A}(\mathbf{f_0}) - \boldsymbol{\varepsilon}\|_2,$$

and apply the reverse triangle inequality to obtain

$$\begin{aligned}\|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{y}\|_2 \geq\ & \|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathcal{A} \circ \mathcal{G}(\mathbf{z_0})\|_2 \\ & - \|\mathcal{A} \circ \mathcal{G}(\mathbf{z_0}) - \mathcal{A}(\mathbf{f_0})\|_2 - \|\boldsymbol{\varepsilon}\|_2.\end{aligned}$$

Using bi-Lipschitz property of $\mathcal{A} \circ \mathcal{G}$ and $\mathcal{A}$ and moving terms around we find

$$\gamma^{-1}\|\tilde{\mathbf{z}} - \mathbf{z_0}\|_2 \leq \|\mathcal{A} \circ \mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{y}\|_2 + \alpha\|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2 + \|\boldsymbol{\varepsilon}\|_2.$$

The third step is to split the error $\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2$ and apply the triangle inequality

$$\begin{aligned}\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 =& \|\mathcal{G}(\tilde{\mathbf{z}}) - \mathcal{G}(\mathbf{z_0}) + \mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2 \\ \leq& \|\mathcal{G}(\tilde{\mathbf{z}}) - \mathcal{G}(\mathbf{z_0})\|_2 + \|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2.\end{aligned}$$

The bi-Lipschitz property of $\mathcal{G}$ then yields

$$\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq \beta\|\tilde{\mathbf{z}} - \mathbf{z_0}\|_2 + \|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2.$$

The desired result follows by combing the three steps. □

Remark 2. *Under the assumptions stated above, it is not unreasonable to assume $\beta\gamma < \alpha$, so that the regularized problem indeed leads to less amplification of noise, at the expense of a bias which mainly depends on the expressiveness of the generative model. More specifically, there is a trade-off between the introduced bias and the noise amplification. More restrictive generative model would have smaller β, leading to the smaller variance, but at the same time it would lead to large bias term in out-of-distribution scenarios (see Figure 3).*

B. Detecting bias

Even though using regularization in the form of a generative model allows to significantly reduce the noise amplification in the scenario when the ground truth signal $\mathbf{f_0} = \mathcal{G}(\mathbf{z_0})$ fits the generative model perfectly, in practice this is often not the case. In many applications, such as material science [1], [16], lithography [19], and circuits board manufacturing [15], the ground truth signal has the form $\mathbf{f_0} = \mathcal{G}(\mathbf{z_0}) + \boldsymbol{\eta}$, where $\mathcal{G}$ models all the "perfect" signals and $\boldsymbol{\eta}$ represents the signal imperfections and manufacturing defects. In such a scenario, detecting these defects is an important part of the reconstruction problem. Therefore, given a solution $\tilde{\mathbf{z}} = \arg\min_{\mathbf{z}} \|\mathcal{A} \circ \mathcal{G}(\mathbf{z}) - \mathbf{y}\|_2$ to the optimization problem (2), we would like to be able to detect and ideally also quantify the bias $\|\mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{f_0}\|_2$.

Using the result from the previous section, we directly get bounds on the bias in terms of the residual $\tilde{\rho} = \|\mathcal{A}\circ\mathcal{G}(\tilde{\mathbf{z}})-\mathbf{y}\|_2$ and the noise level $\sigma = \|\boldsymbol{\varepsilon}\|_2$:

$$\alpha^{-1}\left(\tilde{\rho} - \sigma\right) \leq \|\mathcal{G}(\tilde{\mathbf{z}}) - \mathbf{f_0}\|_2 \leq \alpha\left(\tilde{\rho} + \sigma\right).$$

However, we can expect to get more refined results by exploiting further properties of the measurement map $\mathcal{A}$.

C. Towards a unified approach

To minimize the influence of the generative model bias on the reconstruction accuracy, we combine the classical formulation (1) and the generative formulation (2), and propose the following unified approach to inference using a generative model

$$\min_{\mathbf{z},\mathbf{f}} \|\mathcal{A}(\mathbf{f}) - \mathbf{y}\|_2^2 + \lambda^2\|\mathcal{G}(\mathbf{z}) - \mathbf{f}\|_2^2. \tag{3}$$

Obviously, as $\lambda = 0$, we retrieve (1), and as $\lambda \to \infty$, we retrieve (2).

Lemma 3. *The estimate $\tilde{\mathbf{f}}$ resulting from* (3) *has a reconstruction error bounded as*

$$\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq \lambda\alpha\|G(\mathbf{z_0}) - \mathbf{f_0}\|_2 + 2\alpha\|\boldsymbol{\varepsilon}\|_2$$

Proof. As in the previous results, the bi-Lipschitz property of the map $\mathcal{A}$ yields

$$\alpha^{-1}\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq \|\mathcal{A}(\tilde{\mathbf{f}}) - \mathbf{y}\|_2 + \|\boldsymbol{\varepsilon}\|_2.$$

Using the optimality of $\tilde{\mathbf{f}}$, for any $\mathbf{f} \in \mathbb{C}^n$, $\mathbf{z} \in \mathbb{C}^k$, we trivially have

$$\|\mathcal{A}(\tilde{\mathbf{f}})-\mathbf{y}\|_2 \leq \sqrt{\|\mathcal{A}(\mathbf{f}) - \mathcal{A}(\mathbf{f_0})\|_2^2 + \lambda^2\|\mathcal{G}(\mathbf{z}) - \mathbf{f}\|_2^2 + \|\boldsymbol{\varepsilon}\|_2^2}.$$

Now by picking $\mathbf{f} = \mathbf{f_0}$ and $\mathbf{z} = \mathbf{z_0} \equiv \min_{\mathbf{z}} \|\mathcal{G}(\mathbf{z}) - \mathbf{f_0}\|$ and bounding the expression above using $\sqrt{a^2+b^2} \leq |a|+|b|$, we get the desired result. □

Remark 3. *To ensure optimal reconstruction, the regularization parameter λ should be adaptively selected based on the noise intensity $\|\boldsymbol{\varepsilon}\|_2$. If we let $\lambda \propto \|\boldsymbol{\varepsilon}\|_2$, the reconstruction error of the combined method is bounded as*

$$\|\tilde{\mathbf{f}} - \mathbf{f_0}\|_2 \leq C\|\boldsymbol{\varepsilon}\|_2,$$

where the constant C may depend on the bias $\|\mathcal{G}(\mathbf{z_0}) - \mathbf{f_0}\|_2$. In the regime when $\|\boldsymbol{\varepsilon}\|_2 \to 0$, this gives us the desired asymptotic result. It is not clear if this is also the case in the high noise intensity regime $\|\boldsymbol{\varepsilon}\|_2 \to \infty$, that is, if for lower signal-to-noise ratio, $\lambda \propto \|\boldsymbol{\varepsilon}\|_2$ increases fast enough to activate the generative model.

In practice, however, the noise intensity is not known in advance. In this case, one possible solution is to let the parameter λ vary throughout the iterative reconstruction process, depending on the size of the data fitting term $\|\mathcal{A}(\mathbf{f}) - \mathbf{y}\|_2^2$.

III. Numerical results

One can view (1), (2), and (3) as partial cases of the generic non-linear optimization problem of the form

$$\min_{\mathbf{x}\in\mathbb{C}^d} \|\mathcal{A} \circ \mathcal{B}(\mathbf{x}) - \mathbf{y}\|_2^2 + \lambda^2\|\mathbf{w} \odot \mathbf{x}\|_2^2. \tag{4}$$

We retrieve the specific instances as follows:

- (1) by letting $d = n$, $\mathbf{x} := \mathbf{f}$, $\mathcal{B} = \mathcal{I}$, and $\lambda = 0$;
- (2) by letting $d = k$, $\mathbf{x} := \mathbf{z}$, $\mathcal{B} = \mathcal{G}$, and $\lambda = 0$;
- (3) by letting $d = k + n$ with $\mathbf{x} = (\mathbf{x_1}, \mathbf{x_2})$, where $\mathbf{x_1} \in \mathbb{C}^k$ and $\mathbf{x_2} \in \mathbb{C}^n$. We define $\mathcal{B} : \mathbb{C}^{n+k} \to \mathbb{C}^n$ as $\mathcal{B}(\mathbf{x_1}, \mathbf{x_2}) = \mathcal{G}(\mathbf{x_1}) + \mathbf{x_2}$, and $w \in \mathbb{C}^{k+n}$ is defined by $w(t) = 0$ for $t \in [k]$ and $w(t) = 1$ for $t \in [k+n]\backslash[k]$. The optimization problem (3) then follows by taking $\mathbf{x_1} := \mathbf{z}$ and $\mathbf{x_2} := \mathbf{f} - \mathcal{G}(\mathbf{z})$.

To obtain an approximation of the object reconstruction $\tilde{\mathbf{f}} = \mathcal{B}(\tilde{\mathbf{z}})$, we numerically solve (4) using a Quasi-Newton method, such as limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [17]. The reconstruction algorithm we employ thus has parameters $\mathcal{B}, \lambda, \mathbf{w}$, and a stopping tolerance for L-BFGS.

A. Experiment

For our numerical experiment, we define the measurement map as

$$\mathcal{A}(\mathbf{f}) = |A\mathbf{f}|^2,$$

where matrix A represents a masked Fourier transform, corresponding to the measurements with ℓ different probes

$$A = \begin{pmatrix} F\mathrm{diag}(\mathbf{a}_1) \\ F\mathrm{diag}(\mathbf{a}_2) \\ \vdots \\ F\mathrm{diag}(\mathbf{a}_\ell) \end{pmatrix}.$$

Here, $F \in \mathbb{C}^{n\times n}$ is the discrete Fourier transform matrix and $\mathbf{a}_i \in \mathbb{R}^n$ are random binary probes, that is, their entries are

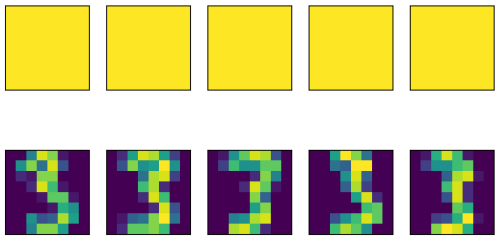

Fig. 1: Samples of the data set on which the generative model was trained ($n = 64$). The top row displays the real part while the bottom row displays the imaginary part.

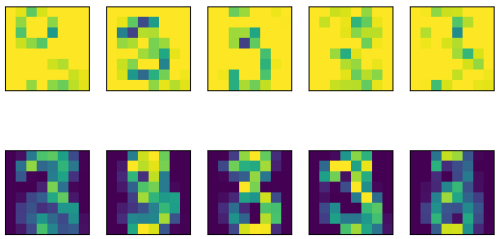

Fig. 2: Samples generated by the generative model ($k = 30$) The top row displays the real part while the bottom row displays the imaginary part.

independent identically distributed Bernoulli random variables. The number of measurements is thus $m = n \cdot \ell$.

We define the generative model as

$$\mathcal{G}(\mathbf{z}) = G\mathbf{z} + \mathbf{b},$$

where $\mathbf{b} \in \mathbb{C}^n$ and $G \in \mathbb{R}^{n \times k}$ are obtained by principle component analysis of a data set of handwritten digits [6]. The elements of the data set are 8×8 images, so that $n = 64$. For our experiment, we choose $k = 30$.

Examples of the elements of the data set on which the generative model $\mathcal{G}$ is trained are shown in Figure 1. Figure 2 shows examples of the signals $\mathcal{G}(\mathbf{z})$, $\mathbf{z} \in \mathbb{C}^{30}$, obtained from the trained generative model. To define the measurement map $\mathcal{A}$, we use $\ell = 100$ randomly generated binary probes $\mathbf{a_j}$, $j \in [\ell]$. For the numerical experiment, we generate measurements with additive Gaussian noise $\boldsymbol{\varepsilon} \sim \mathcal{N}(0, \sigma^2 I_m)$.

To make the reconstruction methods (1) and (2) more stable to the measurement noise, we introduced additional Tikhonov regularization term by setting $\lambda = \sigma^2$ and $\mathbf{w} = \mathbf{1}$ in (4). For the reconstruction method (3), we set $\lambda = 10 \cdot \sigma^2$ and $\mathbf{w} = \mathbf{1}$, which also introduces additional Tikhonov regularization term for $\mathbf{z}$. The reconstruction errors for the three methods, tested on both in-distribution (that is, the ground truth is generated by the generative model) and out-of-distribution (where we use samples from the original data set as ground truth) data are shown in Figure 3.

We see that for high signal-to-noise ratio levels, all methods perform well on in-distribution data, and that for low signal-to-noise ratio levels the generative model shows a slight advantage. On out-of-distribution data, the generative approach (2) clearly shows the bias in the error for high signal-to-noise ratio regime. The combined method (3) achieves the best result for both low and high signal-to-noise ratio levels.

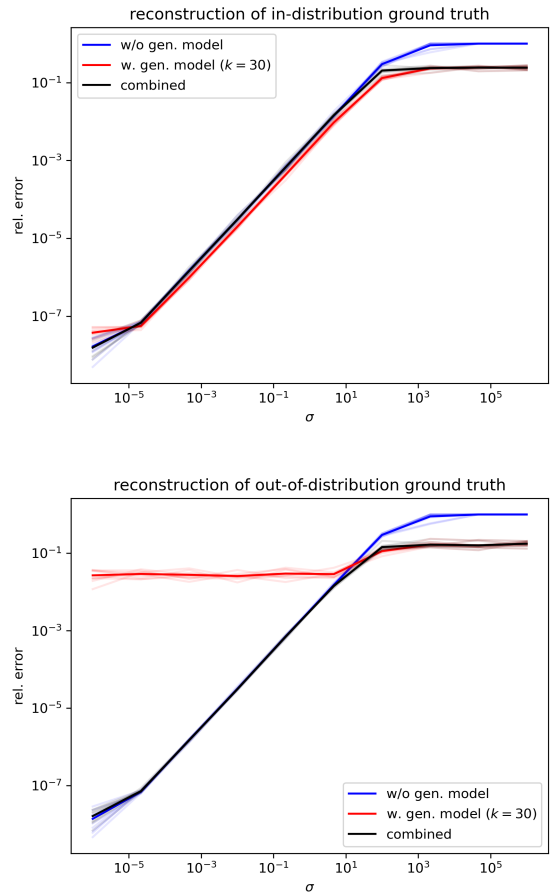

Fig. 3: Relative reconstruction error for the three methods for varying signal-to-noise ratio levels, tested on both in-distribution and out-of-distribution scenarios.

IV. Conclusion and Discussion

In this paper, we explored the use of generative models to regularize certain inverse problems, such as phase retrieval. These preliminary results we showed indicate that generative priors can indeed improve the robustness of the inverse problem solution to measurement noise, at the expense of introducing a bias in the reconstruction. To mitigate this issue, we propose a method that aims to combine the best characteristics of both conventional and regularized methods by interpolating between them. Numerical results on phase retrieval from masked Fourier measurements show that the combined method can indeed achieve the best results. However, the presented error bounds are rather crude and can probably be improved with more careful analysis methods. Further research is needed to solidify our understanding of the combined method, to refine it, and to make it feasible for high-dimensional problems.

REFERENCES

[1] Mostafa Agour, Claas Falldorf, and Ralf B Bergmann. Investigation of composite materials using slm-based phase retrieval. *Optics Letters*, 38(13):2203–2205, 2013.

[2] Radu Balan, Pete Casazza, and Dan Edidin. On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3):345–356, 2006.

[3] Radu Balan and Dongmian Zou. On lipschitz analysis and lipschitz synthesis for the phase retrieval problem. *Linear Algebra and its Applications*, 496:152–181, 2016.

[4] Tamir Bendory, Nadav Dym, Dan Edidin, and Arun Suresh. Phase retrieval with semi-algebraic and relu neural network priors. *arXiv preprint arXiv:2311.08833*, 2023.

[5] Oliver Bunk, Ana Diaz, Franz Pfeiffer, Christian David, Bernd Schmitt, Dillip K Satapathy, and J Friso Van Der Veen. Diffractive imaging for periodic samples: retrieving one-dimensional concentration profiles across microfluidic channels. *Acta Crystallographica Section A: Foundations of Crystallography*, 63(4):306–314, 2007.

[6] Li Deng. The mnist database of handwritten digit images for machine learning research. *IEEE Signal Processing Magazine*, 29(6):141–142, 2012.

[7] Mo Deng, Shuai Li, Alexandre Goy, Iksung Kang, and George Barbastathis. Learning to synthesize: robust phase retrieval at low photon counts. *Light: Science & Applications*, 9(1):36, 2020.

[8] Jonathan Dong, Lorenzo Valzania, Antoine Maillard, Thanh An Pham, Sylvain Gigan, and Michael Unser. Phase retrieval: From computational imaging to machine learning: A tutorial. *IEEE Signal Processing Magazine*, 40:45–57, 2023.

[9] Yonina C Eldar, Pavel Sidorenko, Dustin G Mixon, Shaby Barel, and Oren Cohen. Sparse phase retrieval from short-time fourier measurements. *IEEE Signal Processing Letters*, 22(5):638–642, 2014.

[10] Philipp Grohs, Sarah Koppensteiner, and Martin Rathmair. Phase retrieval: Uniqueness and stability. *SIAM Review*, 62:301–350, 2020.

[11] Paul Hand, Oscar Leong, and Vlad Voroninski. Phase retrieval under a generative prior. *Advances in Neural Information Processing Systems*, 31, 2018.

[12] Teiko Heinosaari, Luca Mazzarella, and Michael M Wolf. Quantum tomography under prior information. *Communications in Mathematical Physics*, 318(2):355–374, 2013.

[13] Mark A Iwen, Sami Merhi, and Michael Perlmutter. Lower lipschitz bounds for phase retrieval from locally supported measurements. *Applied and Computational Harmonic Analysis*, 47(2):526–538, 2019.

[14] Kishore Jaganathan, Samet Oymak, and Babak Hassibi. Sparse phase retrieval: Uniqueness guarantees and recovery algorithms. *IEEE Transactions on Signal Processing*, 65(9):2402–2410, 2017.

[15] Gulhan Ustabas Kaya. Development of hybrid optical sensor based on deep learning to detect and classify the micro-size defects in printed circuit board. *Measurement*, 206:112247, 2023.

[16] Lucien Lehmann, Stéphane Bouillet, Christophe Leymarie, Christel Ameil-Schuh, Melusine Benoit, and Claude Rouyer. Characterization of surface defects using a phase retrieval technique in a high-power laser system. *Applied Optics*, 61(6):1545–1551, 2022.

[17] Dong C Liu and Jorge Nocedal. On the limited memory bfgs method for large scale optimization. *Mathematical programming*, 45(1):503–528, 1989.

[18] Rick P Millane. Phase retrieval in crystallography and optics. *JOSA A*, 7(3):394–411, 1990.

[19] Iacopo Mochi, Kenneth A Goldberg, and Sungmin Huh. Actinic imaging and evaluation of phase structures on extreme ultraviolet lithography masks. *Journal of Vacuum Science & Technology B*, 28(6):C6E11–C6E16, 2010.

[20] Lawrence R Rabiner and Biing-Hwang Juang. *Fundamentals of speech recognition*. Tsinghua University Press, 1999.

[21] Yair Rivenson, Yibo Zhang, Harun Günaydın, Da Teng, and Aydogan Ozcan. Phase recovery and holographic image reconstruction using deep learning in neural networks. *Light: Science & Applications*, 7(2):17141–17141, 2018.

[22] John M Rodenburg. Ptychography and related diffractive imaging methods. *Advances in imaging and electron physics*, 150:87–184, 2008.

[23] Sören Schulze and Emily J King. Sparse pursuit and dictionary learning for blind source separation in polyphonic music recordings. *EURASIP Journal on Audio, Speech, and Music Processing*, 2021(1):1–25, 2021.

[24] Jacob Seifert, Yifeng Shao, and Allard P. Mosk. Noise-robust latent vector reconstruction in ptychography using deep generative models. *Optics Express*, 32:1020, 1 2024.

[25] Ayan Sinha, Justin Lee, Shuai Li, and George Barbastathis. Lensless computational imaging through deep learning. *Optica*, 4(9):1117–1125, 2017.

[26] VV Vasin. Some tendencies in the tikhonov regularization of ill-posed problems. *Journal of Inverse & Ill-Posed Problems*, 14(8), 2006.