Deep Phaseless Imaging via Wirtinger Flow with Decoding Prior

Samia Kazemi
Department of Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
110 8th Street, Troy, NY 12180 USA
Email: kazems@rpi.edu

Bariscan Yonel
Department of Electrical, Computer and Systems Engineering,
Rensselaer Polytechnic Institute
110 8th Street, Troy, NY 12180 USA
Email: yonelb2@rpi.edu

Birsen Yazici
Department of Electrical, Computer and Systems Engineering
Rensselaer Polytechnic Institute
110 8th Street, Troy, NY 12180 USA
Email: yazici@ecse.rpi.edu

***Abstract*—We introduce a deep learning (DL) based network and an associated performance guarantees for imaging from intensity-only measurements using low dimensional encoded representations. Phaseless imaging constitutes a non-convex and ill-posed problem that is relevant to a wide range of applications, where accurate measurement of phase information is challenging. State-of-the-art methods solve the original non-convex optimization problem using sophisticated initialization schemes that lead to locally benign loss function topographies. However, these are commonly contingent upon high sample complexity and restrictive conditions on the forward maps, which limit their practical applicability. To circumvent fundamental limitations, we utilize a model-based deep network for phaseless imaging that implements a fixed-step size realization of gradient descent directly on the lower dimensional encoded representation domain. Accordingly, the iterative algorithm is combined with a non-linear encoder-decoder pair that govern the mapping between low dimensional representations and the image manifold of our interest. This results in feasible regimes beyond those dictated by the standard sufficient conditions of the exact recovery theory used in non-convex optimization. We empirically demonstrate the effectiveness of our lower dimensional formulation of signal recovery through numerical simulations on a number of practical deterministic imaging geometries at reduced sample complexities.**

***Index Terms*—deep learning, phase retrieval, phaseless imaging, decoding prior, generative prior, kernel PCA**

I. Introduction

Phase retrieval problem is commonly encountered in a wide range of imaging applications including synthetic aperture imaging [1], [2], optical imaging [3], [4], x-ray crystallography [3], coded diffraction imaging [5]. In many wave-based imaging applications, accurate measurement of phase information of the back-scattered signal can be challenging due to random variations in the transmission medium properties, or maintaining phase coherence at high frequency operating regimes. Sophisticated and expensive hardware become necessary to facilitate coherent processing for reliable performance of conventional imaging algorithms in such acquisition conditions. Therefore, towards deployment of low cost, spatially diverse sensing systems in wave-based imaging it is important to design algorithms that can reconstruct the underlying scene using only the magnitude or the intensity-only measurements.

Phase retrieval is a severely ill-posed inverse problem, for which a range of approaches have been introduced in the literature involving both convex [6] and non-convex optimization [7]–[15]. Most crucially, the latter provides improved applicability to high dimensional inverse problems in imaging due to lower computational and memory requirements of directly operating on the signal domain, in lieu of the convexification in the lifted parameter space. A prominent iterative approach within this category is the Wirtinger flow (WF) algorithm [9], [13], with its performance guarantees contingent upon properties of the equivalent lifted forward map to facilitate a provably good accuracy via spectral initialization (SI) [16]. However, computationally expensive nature of this step, and stringency of the corresponding sufficient conditions for deterministic measurement maps beyond Gaussian sampling or coded diffraction models still pose critical bottlenecks for the practical applicability of these techniques for high resolution imaging problems. Notably, the underlying linear map prior to the loss of phase information in the measurements is seldom well-conditioned in practical imaging applications, or even constitutes an underdetermined system of equations corresponding to a sample-starved regime beyond the information theoretic limits for injectivity [17]. Further theoretical limitations arise in phaseless imaging with models sampled from Fourier integral operators, where ambiguity from translation invariance is unavoidable without additional redundancy from diverse illuminations [18], [19].

The aforementioned limitations ultimately necessitate the use of a-priori information for both feasibility of the problem up to trivial ambiguities, and for computationally tractable optimization for structurally meaningful image reconstruction. Prior models have been considered in phase retrieval literature using functional regularizers such as sparsity [12], [20], and low-rank models [21], [22]. However, existing theoretical guarantees for these methods pertain to i.i.d. Gaussian distributed sampling vectors [12], [20], [23], which is inadequate for prac-

This work was supported in part by the Air Force Office of Scientific Research (AFOSR) under the agreement FA9550-19-1-0284, in part by Office of Naval Research (ONR) under the agreement N00014-18-1-2068, in part by the National Science Foundation (NSF) under Grant No ECCS-1809234 and in part by the United States Naval Research Laboratory (NRL) under the agreement N00173-21-1-G007.

tical imaging applications. More recently, deep learning-based priors have been deployed using two primary mechanisms, namely, decoding or generative priors [11], [15], [24], [25] and denoising priors [26]–[28], which utilize either plug and play or regularization by denoising (RED) framework [28]. Under decoding or generative prior, an encoded representation of the unknown image is recovered from the measurements, which is then applied to a DN to generate the reconstructed imagery. If this DN is pre-trained and is incorporated as part of an iterative algorithm using random initialization for the encoded representations, then this prior is typically referred to as the generative prior [11], [15], [24]. We use the decoding prior terminology to differentiate when the encoder-decoder components are trained in cascade a recurrent neural network (RNN) architecture for an end-to-end imaging network that also incorporates an initialization step for the lower dimensional encoded representation [25]. Theoretical exact recovery guarantees have been established for both prior cases with sufficient conditions that rely on deterministic properties of the forward maps [24], [25].

Another line of research has been focused on improving the accuracy and sample efficiency of initialization step. In [29], a truncated SI step is presented that uses sparsity prior during initialization. An alternative orthogonalilty promoting initialization technique have been introduced in [30]. Another initialization approach applied to the reweighted amplitude flow algorithm is presented in [31]. Other initialization methods include SI step using a Gaussian optimal sample processing function, Bregman divergence minimizing sample processing, and a composite initialization methods introduced in [32], [33] and [34], respectively. Although these techniques result in improved initialization accuracy than the standard SI, their computational requirement are of same complexity as SI and thus still prohibitive for practice, especially in the training of an architecture implementing a decoding prior. As a result, while [25] overcomes the reliance on large amounts of training data common to generative prior approaches via its end-to-end formulation, its existing implementation is instead limited by the high computational cost of the SI step.

To address the limitations of existing phaseless imaging methods, in this paper we propose a DL based approach that applies a computationally efficient initialization step with a decoding prior based unrolled WF algorithm. Our overall imaging network is composed of two parts, namely, the initialization and imaging modules. The imaging module consists of the RNN and the decoder modules of the imaging network in in [25], and implements a fixed step realization of the WF algorithm over a lower dimensional encoded representation manifold. However, unlike [25], here we implement a model-based initialization network, in lieu of the computationally expensive SI step and an encoder, with the goal of addressing an optimization problem formulated to imitate the one associated with SI within a DL framework. While the decoding prior addresses the sample complexity limitation as in [25], our proposed DL based initialization step is designed to directly encode the intensity measurements into an initial encoded representation. The initialization module involves unrolling another optimization problem in the range of a non-linear kernel function, thus enables the lower dimensional encoded representations as the search space of similarity assessment. This is coupled with learned sample processing of the intensity measurements to induce the structure promoted by the choice of the synthesizing kernel. As a result, the unrolled initialization relieves the computational burden of the decoder gradient in generating the lower dimensional representations, and boasts lower implementation cost compared to the imaging module. We verify the feasibility of our approach by numerical evaluations using MNIST images with the numerical transmission matrix estimated through a spatial light modulator setup published in [35].

II. Problem Statement

Let the ground truth image, which is an N-length vector of either complex or real valued elements, be denoted by $\boldsymbol{\rho}^*$. We assume that the intensity values of the measurements corresponding to M sampling vectors, $\{\mathbf{a}_m\}_{m=1}^M$, where $\mathbf{a}_m \in \mathbb{C}^N$, are available during imaging. We denote the m^{th} intensity measurements by $d_m \in \mathbb{R}^+$, and it is relates to $\boldsymbol{\rho}^*$ and $\mathbf{a}_m$ as follows:

$$d_m = \mathbf{a}_m^H \boldsymbol{\rho}^* \boldsymbol{\rho}^{*H} \mathbf{a}_m. \tag{1}$$

Phaseless imaging refers to the task of estimating $\boldsymbol{\rho}^*$ from a given measurement related vector $\mathbf{d}$, where $\mathbf{d}$ defined as

$$\mathbf{d} = \begin{bmatrix} d_1 & \dots & d_M \end{bmatrix}^T. \tag{2}$$

We represent the linear mapping from the outer product of $\boldsymbol{\rho} \in \mathbb{C}^N$, commonly referred to as the lifted or Kronecker image, to the corresponding measurement vector by $\mathcal{F}$ indicating that $\mathbf{d}$ can be expressed as $\mathcal{F}(\boldsymbol{\rho}^* \boldsymbol{\rho}^{*H})$. Unless the number of measurements is large and the sampling vectors satisfy properties that are sufficient for inducing necessary redundancies in the measurements required for counteracting the loss of phase information, it is an ill-posed problem if no prior information is utilized during imaging.

Suppose $\mathbb{T}$ denotes the manifold formed by our image class of interest. Under a decoding prior framework, we assume that all $\boldsymbol{\rho} \in \mathbb{T}$ is can be expressed as the decoding of a lower dimensional representation $\mathbf{y} \in \mathbb{C}^{N_y}$ via a non-linear mapping $\mathcal{H} : \mathbb{C}^{N_y} \mapsto \mathbb{C}^N$ as $\boldsymbol{\rho} = \mathcal{H}(\mathbf{y})$. N_y denotes the encoded representation length with $N_y \leq N$. This, in effect, induces a non-linearity from the unknown component $\mathbf{y}^* \in \mathbb{C}^{N_y}$, where $\boldsymbol{\rho}^* = \mathcal{H}(\mathbf{y}^*)$, to the m^{th} measurement $\mathbf{a}_m^H \boldsymbol{\rho}^*$ even before its phase information is lost. Given this underlying model assumption, the data model from (1) can be modified as follows:

$$\mathbf{d} = \mathcal{F}(\mathcal{H}(\mathbf{y}^*)\mathcal{H}(\mathbf{y}^*)^H). \tag{3}$$

In [25], we introduced a DL and WF based phaseless imaging approach, which we refer to as DL-WF, for estimating the unknown image under the data model in (3). DL-WF models the non-linear mapping $\mathcal{H}$ using a DN, while a second DN is deployed to encode the output from the SI step to an

optimal initial encoded representation. As the reconstructed images are constrained to the range of the decoder, the model assumption in (3) and the modelling of $\mathcal{H}$ via a DN lead to the imposition of structural prior information captured by the decoding network architecture and its parameter values.

Our objective in this paper is to estimate $\boldsymbol{\rho}^*$ from the measurement related vector $\mathbf{d}$ under the data model in (3). Towards this goal, first, we address the following optimization problem similarly to DL-WF [25]:

$$\underset{\{\mathbf{y}_t\}_{t=1}^T, \mathcal{H}}{\operatorname{argmin}} \frac{1}{TM} \sum_{t=1}^{T} \sum_{m=1}^{M} \left(|\mathbf{a}_m^H \mathcal{H}(\mathbf{y}_t)|^2 - \mathbf{d}_t \right)^2, \quad s.t. \;\; \|\mathcal{H}(\mathbf{y}_t) - \boldsymbol{\rho}_t^*\| \leq \epsilon, \;\; \forall t \in [T], \tag{4}$$

where $\epsilon \in \mathbb{R}^+$ and the subscript t is used to indicate the t^{th} sample from a training set of size T. In this paper, we incorporate the WF update stage of the DL-WF network for imaging and thereby similarly address the optimization problem in (4). However, unlike [25], here we consider a separate modified data model than (3) during initialization in order to attain a more computationally efficient initialization step compared to DL-WF. This modified data model only indirectly relates to $\mathcal{H}$, but explicitly depends on the measurement vector $\mathbf{d}$ via a non-linear mapping $\mathcal{P} : \mathbb{R}_+^M \mapsto \mathbb{R}_+^M$. It can be expressed as follows:

$$\mathbf{g} := \mathcal{P}(\mathbf{d}) \approx \begin{bmatrix} |k(\psi_1, \mathbf{y}^*)|^2 & \dots & |k(\psi_M, \mathbf{y}^*)|^2 \end{bmatrix}^T, \tag{5}$$

where $\boldsymbol{\psi}_m \in \mathbb{C}^{N_y}$, for $m \in [M]$ and $k(\mathbf{v}_1, \mathbf{v}_2)$ is a non-linear function of $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{C}^{N_y}$ with positive real-valued output.

In particular, $\{\boldsymbol{\psi}_m\}_{m=1}^M$ serve as learnable parameters to form a N_y-dimensional dictionary, where the similarity to representations are assessed under a fixed kernel, k. The motivation of this model is to capture the Euclidean sense similarity in the higher $N-$dimensions that is encoded in the measurements $\mathbf{d}$ by virtue of the kernel trick, up-to a transformation $\mathcal{P}$. We refer to $\mathcal{P}$ as the sample processing operator, which is inserted for the data to reflect the structure enforced by the assumed kernel function on higher dimensional feature space. Letting $\mathbf{k} \in \mathbb{R}_+^M$ as a function of $\mathbf{v} \in \mathbb{C}^{N_y}$ defined as follows:

$$\mathbf{k}(\mathbf{v}) = \begin{bmatrix} |k(\boldsymbol{\psi}_1, \mathbf{v})|^2 & \dots & |k(\boldsymbol{\psi}_M, \mathbf{v})|^2 \end{bmatrix}^T. \tag{6}$$

During initialization, under the data model in (5), we aim to address the following optimization problem to calculate an initial encoded representation:

$$\underset{\substack{\{\mathbf{y}_t^{(0)}\}_{t=1}^T, \mathcal{H}, \\ \mathcal{P}, \{\boldsymbol{\psi}_m\}_{m=1}^M}}{\operatorname{argmin}} \frac{1}{TM} \sum_{t=1}^{T} \|\mathbf{k}(\mathbf{y}_t^{(0)}) - \mathcal{P}(\mathbf{d}_t)\|^2, \quad s.t. \;\; \|\mathcal{H}(\mathbf{y}_t^{(0)}) - \boldsymbol{\rho}_t^*)\| \leq \epsilon_0, \;\; \forall t \in [T], \tag{7}$$

where $\epsilon_0 \in \mathbb{R}^+$. Although the data model in (5) is not explicitly related to the one in (3), the two constraints in (4) and (7) contains the same decoder $\mathcal{H}$ and thereby imposes an indirect correlation during implementation.

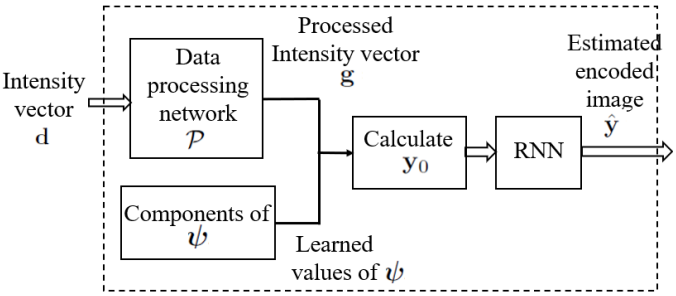


Fig. 1. Schematic diagram of the initialization module.

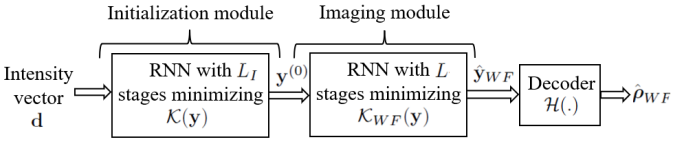


Fig. 2. Schematic diagram of the overall imaging network using WF and deep prior with efficient initialization.

Ultimately (7) amounts to a searching for a simplified structure where the decoder $\mathcal{H}$ is approximated via a fixed kernel, for the benefit of less computational burden. The initial estimate can then refined using the WF iterations under the definition of the decoder using the ideal problem formulation in (4). This is akin to the nature of standard phase retrieval and the formulation of the spectral initialization, which approximates the ideal ℓ_2-loss minimization procedure by the leading eigenvector of the backprojection estimate $\mathcal{F}^H \mathcal{F}(\boldsymbol{\rho}^* \boldsymbol{\rho}^{*H})$, where $\mathcal{F}^H$ denotes the adjoint of the lifted forward model $\mathcal{F}$. While the accuracy of such approximation in the standard setting is pursued under certain restricted isometry-type properties [33], the kernel-based formulation in (7) addresses the initialization by an analogous approximation using the training data, where the end-to-end design of our architecture is leveraged.

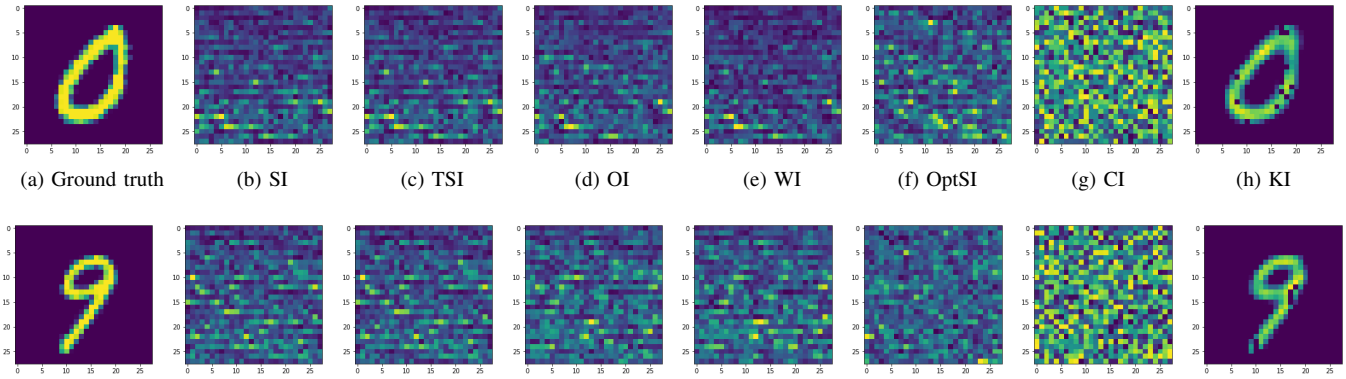
III. Our DL based Approach

We model the non-linear decoder $\mathcal{H}$ and the sample processor $\mathcal{P}$ by distinct DNs. Under this setting, we design the initialization and the imaging modules based on solving the two optimization problems in (4) and (7), respectively, by applying the gradient descent algorithm. For the initialization module, we refer to the initial encoded representation by $\mathbf{y}_0 \in \mathbb{C}^{N_y}$, and the representation after the k^{th} gradient descent update by $\mathbf{y}_k \in \mathbb{C}^{N_y}$. Let the data fidelity term associated with the initialization phase is represented by $\mathcal{K}(\mathbf{y})$, where

$$\mathcal{K}(\mathbf{y}) = \frac{1}{M} \|\mathbf{k}(\mathbf{y}) - \mathbf{g}\|^2. \tag{8}$$

Let $\lambda_k \in \mathbb{R}^+$ is the learning rate for the k^{th} step of the initialization module. The updated signal $\mathbf{y}_k$ is calculated as

$$\mathbf{y}_k = \mathbf{y}_{k-1} - \lambda_k \nabla_{\mathbf{y}} \mathcal{K}(\mathbf{y}_{k-1}), \tag{9}$$



(a) Ground truth (b) SI (c) TSI (d) OI (e) WI (f) OptSI (g) CI (h) KI

(i) Ground truth (j) SI (k) TSI (l) OI (m) WI (n) OptSI (o) CI (p) KI

Fig. 3. Two example images from the testset are shown in the first column. Column 2 to 8 includes the corresponding reconstructed images using the SI, TSI, OI, WI, OptSI, CI and our kernel PCA based methods for an under-sampled case using $M = 0.5N$. We implemented our algorithm using our initialization module with 10 RNN layers and Laplace RBF kernel function, no additional layers were used for the imaging module.

where the gradient term is calculated as follows:

$$\nabla_{\mathbf{y}}\mathcal{K}(\mathbf{y}) = \frac{4}{M}\sum_{m=1}^{M}(|k(\boldsymbol{\psi}_m,\mathbf{y})|^2 - g_m)k(\boldsymbol{\psi}_m,\mathbf{y})\frac{\partial k(\boldsymbol{\psi}_m,\mathbf{y})}{\partial \mathbf{y}}. \tag{10}$$

Estimated encoded signal from the initialization module after the $L_I{}^{th}$ update, represented by $\mathbf{y}_{L_I} \in \mathbb{C}^{N_y}$, is set as the initial value of the imaging module $\mathbf{y}^{(0)}$, i.e., $\mathbf{y}^{(0)} = \mathbf{y}_{L_I}$. As it is in [25], the cost function minimized within the imaging module, $\mathcal{K}_{WF}(\mathbf{y})$ calculated as follows:

$$\mathcal{K}_{WF}(\mathbf{y}) = \frac{1}{M}\||\mathbf{a}_m^H\mathcal{H}(\mathbf{y})|^2 - \mathbf{d}\|^2, \tag{11}$$

and the associated l^{th} update step is the following:

$$\mathbf{y}^{(l)} = \mathbf{y}^{(l-1)} - \mu_l\nabla_{\mathbf{y}}\mathcal{K}_{WF}(\mathbf{y}_{l-1}), \tag{12}$$

where $\mu_l \in \mathbb{R}^+$ is the learning rate for the l^{th} step. We map the L_I and L updates from (9) and (12), respectively, are mapped into two RNNs with L_I and L stages. Detailed diagram of the initialization part is shown in Fig. 1. We calculate initial image $\mathbf{y}_0$ for this network from the processed measurements using the following expression:

$$\mathbf{y}_0 = \frac{1}{M}\sum_{m=1}^{M} g_m\boldsymbol{\psi}_m. \tag{13}$$

This formulation can be justified by observing from the data model in (5). The objective function in (7) implies that if the ground truth encoded representation $\mathbf{y}^*$ is mostly aligned to the m^{th} frame vector $\boldsymbol{\psi}_m$, then the corresponding processed intensity value g_m is the largest element of $\mathbf{g}$.

The schematic diagram of our overall imaging network, composed of the initialization and the imaging modules, are shown in Fig. 2. The final estimated encoded representation from the imaging module, denoted by $\hat{\mathbf{y}}_{WF} \in \mathbb{C}^{N_y}$, is set equal to $\mathbf{y}^{(L)}$. The reconstructed image $\hat{\boldsymbol{\rho}}_{WF} \in \mathbb{C}^N$ is calculated from this value as $\mathcal{H}(\hat{\boldsymbol{\rho}}_{WF})$.

IV. Numerical Simulations

For the purpose of advancing our results on the image module with [25], we focus our efforts to the study of the initialization module. Accordingly, we apply our proposed kernel-based approach to a phaseless imaging problem on the simulated dataset. We considered 10000 images of dimension 28×28 from the public MNIST dataset of handwritten digits for training, where each of the 10 digit corresponds to 1000 images. Similarly, we selected another randomly drawn 100 images, 10 for each digit, constitute the ground truth images in the test set. We generate the intensity measurements using the numerical transmission matrix from in [35] to simulate an optical diffraction problem. We consider an under-sampling ratio of 0.5 with number of unknowns and measurements equal to 784 and 392, respectively.

For modelling the initialization module, we consider a 10 layer RNN with a convolutional neural network (CNN) based sample processing network. We modelled the non-linear function $k(.,.)$ by the Laplace RBF kernel, i.e.,

$$k(\mathbf{v}_1,\mathbf{v}_2) = \exp\left(-\frac{\|\mathbf{v}_1 - \mathbf{v}_2\|}{\sigma^2}\right), \tag{14}$$

where $\sigma \in \mathbb{R}$. We applied a 5-layer CNN to realize the sample processing function, with $leaky\_relu(.)$ activation functions for the first four layers and an $relu(.)$ activation function for the output layer. Three 3×3 dimensional convolution filters, with 16 output channels each, are used in this CNN with the same filters shared in the second to fourth layer. Moreover, we considered encoded representation length of 64 and applied a 5 layer feed-forward neural network with $relu(.)$ activation functions as the decoding network.

We consider a number of existing initialization approaches for performance comparison with our method. This include the spectral initialization [9] approach, truncated spectral initialization method from [29], initialization method described in [31] for the re-weighted amplitude flow approach for phase retrieval, spectral initialization using optimal sample processing function introduced in [32], orthogonality-promoting

initialization approach presented for the truncated amplitude flow algorithm in [30], and the composite initialization method for phase retrieval introduced in [34]. We refer to our proposed kernel function based initialization approach as KI, and provide the reconstructed images for the MNIST dataset in Fig. 3 when the number of measurements is half the number of the unknown, i.e., $M = 0.5N$.

V. Conclusion

We introduced a DL based phaseless imaging approach that incorporates a computationally initialization step with an RNN type network. Similarly to our DL-WF approach, proposed imaging network incorporates prior information via a learned decoding network that enables lower sampling complexity by effectively reducing the number of unknown quantities to be estimated. However, compared to our previous approach, we simultaneously attain improved initialization efficiency for the encoded signal by directly estimating it from the intensity measurements and bypassing the spectral initialization in the higher dimensional image manifold. We verified the feasibility of our proposed phaseless imaging approach by training a our network under a specific kernel function assumption using a simulated image set. For future work, we will integrate the initialization module with the refinement stage, and pursue numerical evaluations for synthetic aperture imaging.

References

[1] J. Laviada, A. Arboleya-Arboleya, Y. Alvarez-Lopez, C. Garcia-Gonzalez, and F. Las-Heras, "Phaseless synthetic aperture radar with efficient sampling for broadband near-field imaging: Theory and validation," *IEEE Transactions on Antennas and Propagation*, vol. 63, no. 2, pp. 573–584, 2014.

[2] B. Yonel, E. Mason, and B. Yazici, "Phaseless passive synthetic aperture radar imaging via wirtinger flow," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2018, pp. 1623–1627.

[3] R. P. Millane, "Phase retrieval in crystallography and optics," *JOSA A*, vol. 7, no. 3, pp. 394–411, 1990.

[4] Y. Shechtman, Y. C. Eldar, O. Cohen, H. N. Chapman, J. Miao, and M. Segev, "Phase retrieval with application to optical imaging: a contemporary overview," *IEEE signal processing magazine*, vol. 32, no. 3, pp. 87–109, 2015.

[5] D. Gross, F. Krahmer, and R. Kueng, "Improved recovery guarantees for phase retrieval from coded diffraction patterns," *Applied and Computational Harmonic Analysis*, vol. 42, no. 1, pp. 37–64, 2017.

[6] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, vol. 66, no. 8, pp. 1241–1274, 2013.

[7] J. R. Fienup, "Phase retrieval algorithms: a comparison," *Applied optics*, vol. 21, no. 15, pp. 2758–2769, 1982.

[8] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 2796–2804.

[9] E. J. Candes, X. Li, and M. Soltanolkotabi, "Phase retrieval via Wirtinger flow: Theory and algorithms," *IEEE Transactions on Information Theory*, vol. 61, no. 4, pp. 1985–2007, 2015.

[10] H. Zhang, Y. Liang, and Y. Chi, "A nonconvex approach for phase retrieval: Reshaped Wirtinger flow and incremental algorithms," *Journal of Machine Learning Research*, vol. 18, no. 141, pp. 1–35, 2017. [Online]. Available: http://jmlr.org/papers/v18/16-572.html

[11] F. Shamshad and A. Ahmed, "Compressed sensing-based robust phase retrieval via deep generative priors," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2286–2298, 2021.

[12] Z. Yuan, H. Wang, and Q. Wang, "Phase retrieval via sparse Wirtinger flow," *Journal of Computational and Applied Mathematics*, vol. 355, pp. 162–173, 2019.

[13] B. Yonel and B. Yazici, "A deterministic theory for exact non-convex phase retrieval," *IEEE Transactions on Signal Processing*, vol. 68, pp. 4612–4626, 2020.

[14] F. Wu and P. Rebeschini, "Hadamard Wirtinger flow for sparse phase retrieval," in *International Conference on Artificial Intelligence and Statistics*. PMLR, 2021, pp. 982–990.

[15] F. Shamshad and A. Ahmed, "Compressed sensing-based robust phase retrieval via deep generative priors," *IEEE Sensors Journal*, vol. 21, no. 2, pp. 2286–2298, 2020.

[16] B. Yonel and B. Yazici, "A generalization of wirtinger flow for exact interferometric inversion," *SIAM Journal on Imaging Sciences*, vol. 12, no. 4, pp. 2119–2164, 2019.

[17] A. S. Bandeira, J. Cahill, D. G. Mixon, and A. A. Nelson, "Saving phase: Injectivity and stability for phase retrieval," *Applied and Computational Harmonic Analysis*, vol. 37, no. 1, pp. 106–125, 2014.

[18] D. Kogan, Y. C. Eldar, and D. Oron, "On the 2d phase retrieval problem," *IEEE Transactions on Signal Processing*, vol. 65, no. 4, pp. 1058–1067, 2016.

[19] K. Jaganathan, S. Oymak, and B. Hassibi, "Sparse phase retrieval: Uniqueness guarantees and recovery algorithms," *IEEE Transactions on Signal Processing*, vol. 65, no. 9, pp. 2402–2410, 2017.

[20] T. T. Cai, X. Li, Z. Ma *et al.*, "Optimal rates of convergence for noisy sparse phase retrieval via thresholded Wirtinger flow," *The Annals of Statistics*, vol. 44, no. 5, pp. 2221–2251, 2016.

[21] C. A. Metzler, A. Maleki, and R. G. Baraniuk, "Bm3d-prgamp: Compressive phase retrieval based on bm3d denoising," in *2016 IEEE International Conference on Image Processing (ICIP)*. IEEE, 2016, pp. 2504–2508.

[22] N. Vaswani, S. Nayer, and Y. C. Eldar, "Low-rank phase retrieval," *IEEE Transactions on Signal Processing*, vol. 65, no. 15, pp. 4059–4074, 2017.

[23] M. Soltanolkotabi, "Structured signal recovery from quadratic measurements: Breaking sample complexity barriers via nonconvex optimization," *IEEE Transactions on Information Theory*, vol. 65, no. 4, pp. 2374–2400, 2019.

[24] P. Hand, O. Leong, and V. Voroninski, "Phase retrieval under a generative prior," in *Advances in Neural Information Processing Systems*, 2018, pp. 9136–9146.

[25] S. Kazemi, B. Yonel, and B. Yazici, "Unrolled wirtinger flow with deep decoding priors for phaseless imaging," *IEEE Transactions on Computational Imaging*, vol. 8, pp. 609–625, 2022.

[26] C. Metzler, P. Schniter, A. Veeraraghavan, and R. Baraniuk, "prDeep: Robust phase retrieval with a flexible deep network," in *Proceedings of the 35th International Conference on Machine Learning*, vol. 80. Proceedings of Machine Learning Research, 2018, pp. 3501–3510.

[27] Z. Wu, Y. Sun, J. Liu, and U. Kamilov, "Online regularization by denoising with applications to phase retrieval," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, 2019, pp. 0–0.

[28] D. Xue, Z. Zheng, W. Dai, C. Li, J. Zou, and H. Xiong, "On the convergence of non-convex phase retrieval with denoising priors," *IEEE Transactions on Signal Processing*, pp. 1–16, 2022.

[29] Y. Chen and E. Candes, "Solving random quadratic systems of equations is nearly as easy as solving linear systems," *Advances in Neural Information Processing Systems*, vol. 28, 2015.

[30] G. Wang, G. B. Giannakis, and Y. C. Eldar, "Solving systems of random quadratic equations via truncated amplitude flow," *IEEE Transactions on Information Theory*, vol. 64, no. 2, pp. 773–794, 2018.

[31] G. Wang, G. B. Giannakis, Y. Saad, and J. Chen, "Phase retrieval via reweighted amplitude flow," *IEEE Transactions on Signal Processing*, vol. 66, no. 11, pp. 2818–2833, 2018.

[32] M. Mondelli and A. Montanari, "Fundamental limits of weak recovery with applications to phase retrieval," in *Conference On Learning Theory*. PMLR, 2018, pp. 1445–1450.

[33] B. Yonel and B. Yazici, "A spectral estimation framework for phase retrieval via bregman divergence minimization," *SIAM Journal on Imaging Sciences*, vol. 15, no. 2, pp. 491–520, 2022.

[34] Q. Luo, S. Lin, and H. Wang, "A composite initialization method for phase retrieval," *Symmetry*, vol. 13, no. 11, p. 2006, 2021.

[35] C. A. Metzler, M. K. Sharma, S. Nagesh, R. G. Baraniuk, O. Cossairt, and A. Veeraraghavan, "Coherent inverse scattering via transmission matrices: Efficient phase retrieval algorithms and a public dataset," in *2017 IEEE International Conference on Computational Photography (ICCP)*. IEEE, 2017, pp. 1–16.