Deep Learning based Phaseless SAR without Born Approximation

Samia Kazemi
Department of Electrical, Computer and Systems Engineering,
Rensselaer Polytechnic Institute
110 8th Street, Troy, NY 12180 USA
Email: kazems@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*—In this paper, we present a phase retrieval approach from intensity measurements using a Deep Learning (DL) based Wirtinger Flow (WF) algorithm for the case where the measurement model is non-linear, and this non-linearity depends on the unknown signal. In the context of synthetic aperture radar (SAR), this is relevant to the image reconstruction problem for the scenario where the Born approximation is no longer valid which results in multi-scattering effect within the extended target being imaged. Since we are adopting WF for DL based imaging, the underlying optimization problem is non-convex. However, unlike the WF algorithm, the unknown image is estimated from the measurement intensities in a learned encoding space with the goal of achieving effective reconstruction performance. The overall DL network is composed of an encoding network for determining a suitable initial value in the transformed space, a recurrent neural network (RNN) that models the steps of a gradient descent algorithm for an optimization problem, and a decoding network that can incorporate the generative image prior and transforms the encoded estimation from the RNN output to the original image space. Numerical results are included to verify feasibility of the proposed approach.**

I. Introduction

In recent years, DL has been utilized in a number of SAR related tasks including image reconstruction, segmentation, estimation of the unknown parameters of the forward model, automatic target recognition (ATR) problem [1]–[4] etc. for the case where the phase information of the received signal is available. In recent years, there have been significant interest in the phase retrieval problem which refers to the task of reconstructing an unknown signal from a set of phaseless measurements [5]–[7]. In the context of SAR, this is relevant for the image reconstruction problem for the case when the phases of the reflected signal are unavailable, or they significantly deviate from the correct values due to random fluctuations in the sensor locations, variations in transmission signal speed etc.

A number of different algorithms have been proposed in the literature for the phase retrieval problem which can be broadly categorized into two groups. In the first category, by algebraic manipulation, phase retrieval is performed by convex optimization, either through lifting and convex relaxation [8]–[10], or by the PhaseMax approach [11], [12] without lifting. For the second category, the unknown signal is estimated via non-convex optimization in the original unknown space, and it includes alternative minimization [13], [14], and WF [5] and its many variants [15]–[19]. Out of these methods, WF and its variants are the most prominent since they offer convergence guarantee alongside better scaling property and computational complexity compared to most of the convex optimization based techniques. These algorithms are characterized by two major stages: a sophisticated initialization stage and an iterative update stage for estimating the unknown.

Previous implementations of WF algorithm for radar applications focused mainly on the interferometric imaging problem. This is due to the fact that the cross-correlation of the measurements from a pair of receivers has the characteristic generalized Wirtinger flow [6] measurement model in the frequency domain. In [20], this approach has been applied for the multi-static passive radar imaging problem with Born approximation for the case when knowledge about only the transmitter look direction is available. In [21], a deterministic lifted forward model is designed for interferometric multi-static radar in terms of limits on pixel spacing and sample complexity for exact image reconstruction. In [22], generalized WF approach was applied for simultaneous reconstruction of unknown image and target polarization states from passive interferometric measurements. In this paper on the other hand, we are assuming complete knowledge of the transmitter location, and instead consider the effect of phase errors in the measurements and the multiple scattering effect within the target begin imaged.

An iterative image reconstruction approach for the measurement model with a single transmitter and multiple receivers and target multiple scattering scenario is presented in [23] that simultaneously recovers the unknown image vector and the unknown phase of measurement. A total variation (TV) regularization, non-negativity of the image vector components and unit magnitude elements of the unknown phase vector are incorporated as well to limit the solution space for faster convergence. However, an all zero initial value for the unknown is not rigorously justified and has room to be improved upon along with the arbitrarily selection of the learning rate. For the same measurement scenario but with the measurement phases available, a DL based approach is

proposed in [24], where starting with an approximate inversion via back-projection, a deep network is applied for further refinement. Another notable DL based imaging method for this scenario include [25] where a plug-and-play (PnP) approach is employed for implementing DL based projection steps.

In this paper, we propose an approach that combines a learned non-linear initialization step with the generative prior [26]–[29] based phase retrieval approach in order to achieve effective reconstruction from the phaseless measurements. The goal of the learning phase is to find an appropriate non-linear transformation from the spectral initialization [5] output, which is employed in the WF algorithm, to an encoded domain with reduced dimensionality. Once an estimated solution is found in the encoded space, it is applied to a DL based decoding network and the solution in the original image domain is determined.

Rest of the paper is organized as follows: in Section II, we present the relevant notations and details of the measurement model for the SAR image reconstruction task; in Section III, we briefly summarize the steps of the WF algorithm; in Section IV, we elaborate on the iterative DL based reconstruction approach for the inverse and the forward problem; in Section V, we present the overall deep network; in Section VI, we present example numerical reconstruction output; and finally, Section VII includes our conclusion.

II. Problem statement

Let $\mathbf{x}$ and $\boldsymbol{x}$ denote the position vectors in $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively, and x_1, x_2 and x_3 are the components of $\mathbf{x}$ along the three dimensions. If $\phi(\boldsymbol{x}) \in \mathbb{R}$ denotes the height at $\boldsymbol{x}$, the location of a point on the ground topography at $\boldsymbol{x}$ is denoted by $\mathbf{x} = [x_1, x_2, \phi(\boldsymbol{x})]$. The speed of the electromagnetic wave in the homogeneous background medium and inside the heterogeneous scattering medium at $\mathbf{x}$ are denoted by c_b and $c(\mathbf{x})$, respectively. Scattering potential $V(\mathbf{x})$ is the reciprocal of the perturbation in the background wave speed caused by the scattering object, and it is related to c_b and $c(\mathbf{x})$ as $c^2(\mathbf{x}) = \left(\frac{1}{c_b^2} + V(\mathbf{x})\right)^{-1}$. $\rho(\boldsymbol{x})$ is a function of $\boldsymbol{x}$, and it is defined such that $V(\mathbf{x}) = \rho(\boldsymbol{x})\delta(x_3 - \phi(\boldsymbol{x}))$.

For simplicity, we restrict our proposed image reconstruction approach to the passive bi-static and multi-static SAR for which the fixed transmitter location and the transmission signal is fully known. Modifications required for extending our approach to the mono-static SAR operation with the changing transmitter locations is quite straight-forward.

Suppose $G(\mathbf{x}, \mathbf{y}, \omega)$ and $S(\mathbf{x}, \omega)$ are the Fourier transformations of the Green's function and the source wave, respectively, and $K_T(\mathbf{x}, \omega)$ models the transmission antenna beampattern. For the homogeneous background with constant speed c_b, $G(\mathbf{x}, \mathbf{y}, \omega)$ is defined as $G(\mathbf{x}, \mathbf{y}, \omega) = \frac{\mathrm{e}^{-\mathrm{i}\frac{\omega}{c_b}|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|}$. Suppose the transmission antenna is located at $\boldsymbol{\gamma}_T \in \mathbb{R}^3$ and it transmits a signal with center frequency ω_c. $S(\mathbf{x}, \omega)$ is approximated as $S(\mathbf{x}, \omega) \approx P(\omega)\delta(\mathbf{x} - \boldsymbol{\gamma}_T)$, and the incident field $U^{in}(\mathbf{x}, \omega)$ can be expressed as

$$U^{in}(\mathbf{x}, \omega) = G(\mathbf{x}, \boldsymbol{\gamma}_T, \omega) K_T(\boldsymbol{\gamma}_T, \omega) P(\omega). \tag{1}$$

The scattered field $U^{sc}(\mathbf{x}, \omega)$ and the total field $U(\mathbf{x}, \omega)$ are related as $U(\mathbf{x}, \omega) = U^{in}(\mathbf{x}, \omega) + U^{sc}(\mathbf{x}, \omega)$. Since the Born approximation is not valid, unknown total wave $U(\mathbf{x}, \omega)$ is required to be estimated. Using the background Green's function, $U^{sc}(\mathbf{x}, \omega)$ can be expressed in the integral format as

$$U^{sc}(\mathbf{x}, \omega) = \omega^2 \int G(\mathbf{x}, \mathbf{y}, \omega) K_R(\mathbf{y}, \omega) V(\mathbf{y}) U(\mathbf{y}, \omega) d\mathbf{y}, \tag{2}$$

which is commonly referred to as the Lippman-Schwinger equation and it captures the non-linearity in the measurement model. The $K_R(\mathbf{x}, \omega)$ term in (2) is associated with the receiving antenna beampattern. For simplicity, we consider both K_T and K_R to be equal to 1.

Suppose the heterogeneous target being imaged is confined within a set of position vectors denoted by Ω. Similarly, the receiver locations at $\boldsymbol{\gamma}_R(s) \in \mathbb{R}^3$ for $s \in \{1, \cdots, S\}$, and the transmitter location $\boldsymbol{\gamma}_T$ form the set Γ. For this setting,

$$U(\mathbf{x}, \omega) = U^{in}(\mathbf{x}, \omega) + \omega^2 \int_{\Omega} G(\mathbf{x}, \boldsymbol{y}, \omega) \rho(\boldsymbol{y}) U(\boldsymbol{y}, \omega) d\boldsymbol{y}, \tag{3}$$

where $\mathbf{x} \in \Omega$. The scattered wave at the s^{th} receiver location with phase error $\mathrm{e}^{\mathrm{i}\theta(s,\omega)}$ relates to $U(\mathbf{x}, \omega)$ and $\boldsymbol{\rho}(\boldsymbol{x})$ as

$$\begin{aligned} &d(s, \omega) \\ &\quad \approx \mathrm{e}^{\mathrm{i}\theta(s,\omega)} \omega^2 \int_{\Omega} G(\boldsymbol{\gamma}_R(s), \boldsymbol{y}, \omega) \rho(\boldsymbol{y}) U(\boldsymbol{y}, \omega) d\boldsymbol{y}. \end{aligned} \tag{4}$$

To circumvent the adversarial effect of phase error on the imaging output, we only consider the intensity values of the measurements and denote this by $\hat{d}(s, \omega)$, i.e.,

$$\hat{d}(s, \omega) = |d(s, \omega)|^2. \tag{5}$$

Suppose the region Ω is composed of N discrete points at $\{\mathbf{x}_n\}_{n=1}^N$. From (1), (3), (4) and (5), the discretized version of the phaseless data model is described by the following set of equations:

$$\mathbf{u} = \mathbf{u}^{in} + \mathbf{H}(\mathbf{u} \odot \boldsymbol{\rho}_1), \tag{6}$$

$$\mathbf{c} = \mathbf{G}(\mathbf{u} \odot \boldsymbol{\rho}_1), \tag{7}$$

$$\mathbf{d} = \mathrm{diag}(\mathbf{c})^H \mathbf{c}. \tag{8}$$

Component-wise product of two vectors or matrices is denoted by $\odot$. We use ω_k for $k \in \{1, \cdots, K\}$ to denote the fast-time frequencies and the total number of measurements $M = SK$. Incident wave vector $\mathbf{u}^{in} \in \mathbb{C}^{NK}$ is fully known, and it is composed of its corresponding function values at the N locations in Ω and repeated for each of the K frequencies as $\mathbf{u}^{in} = \begin{bmatrix} \mathbf{g}_1^T & \cdots & \mathbf{g}_K^T \end{bmatrix}^T$ where $\mathbf{g}_k \in \mathbb{C}^N = \begin{bmatrix} U^{in}(\boldsymbol{x}_1, \omega_k) & \cdots & U^{in}(\boldsymbol{x}_N, \omega_k) \end{bmatrix}^T$. Similarly, the fully known matrices $\mathbf{H} \in \mathbb{C}^{NK \times NK}$ and $\mathbf{G} \in \mathbb{C}^{M \times NK}$ are related to the Green's function of the background medium

transmission. $\mathbf{H}$ is a block diagonal matrix with square matrices $\mathbf{H}_k \in \mathbb{C}^{N\times N}$ for $k \in \{1,\cdots,K\}$ along its diagonal while the remaining elements equal to zeros. $\mathbf{H}_k$ is defined as

$$\mathbf{H}_k = \omega_k^2 \begin{bmatrix} G(\boldsymbol{x}_1,\boldsymbol{x}_1,\omega_k) & \cdots & G(\boldsymbol{x}_1,\boldsymbol{x}_N,\omega_k) \\ \vdots & \vdots & \vdots \\ G(\boldsymbol{x}_N,\boldsymbol{x}_1,\omega_k) & \cdots & G(\boldsymbol{x}_N,\boldsymbol{x}_N,\omega_k) \end{bmatrix}. \quad (9)$$

$\mathbf{G}$ on the other hand is composed of $\mathbf{G}_k \in \mathbb{C}^{S\times N}$ for $k \in \{1,\cdots,K\}$ as

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_K \end{bmatrix} \quad (10)$$

where $\mathbf{0}$ is a $S \times N$ matrix of all zeros and

$$\mathbf{G}_k = \omega_k^2 \begin{bmatrix} G(\boldsymbol{\gamma}_R(1),\boldsymbol{x}_1,\omega_k) & \cdots & G(\boldsymbol{\gamma}_R(1),\boldsymbol{x}_N,\omega_k) \\ \vdots & \vdots & \vdots \\ G(\boldsymbol{\gamma}_R(S),\boldsymbol{x}_1,\omega_k) & \cdots & G(\boldsymbol{\gamma}_R(S),\boldsymbol{x}_N,\omega_k) \end{bmatrix}. \quad (11)$$

Unknown image vector $\boldsymbol{\rho} \in \mathbb{R}^N$ is defined as $\boldsymbol{\rho} = \begin{bmatrix}\rho(\boldsymbol{x}_1) & \cdots & \rho(\boldsymbol{x}_N)\end{bmatrix}^T$ and $\boldsymbol{\rho}_1 \in \mathbb{R}^{NK}$ in (6) and (7) is a column vector defined as $\boldsymbol{\rho}_1 = \begin{bmatrix}\boldsymbol{\rho}^T & \cdots & \boldsymbol{\rho}^T\end{bmatrix}^T$. Total wave vector $\mathbf{u} \in \mathbb{C}^{NK}$ relates to its corresponding function values at the N locations in Ω as $\mathbf{u} = \begin{bmatrix}U(\boldsymbol{x}_1,\omega_1) & \cdots & U(\boldsymbol{x}_N,\omega_1) & \cdots & U(\boldsymbol{x}_N,\omega_K)\end{bmatrix}^T$. Scattered signal vector $\mathbf{c} \in \mathbb{C}^M$ is defined as $\mathbf{c} = \begin{bmatrix}d(1,\omega_1) & d(2,\omega_1) & \cdots & d(S,\omega_K)\end{bmatrix}^T$. The vector comprising of the intensity measurements is $\mathbf{d} \in \mathbb{R}^M$ where $\mathbf{d} = \begin{bmatrix}\hat{d}(1,\omega_1) & \hat{d}(2,\omega_1) & \cdots & \hat{d}(S,\omega_K)\end{bmatrix}$. Our goal is to implement a DL based approach to recover the unknown image vector $\boldsymbol{\rho}$ from the measured intensity vector $\mathbf{d}$.

III. WF ALGORITHM FOR PHASE RETRIEVAL

WF algorithm for phase retrieval is characterized by the direct optimization of an un-regularized objective function via gradient descent approach using a carefully designed initial estimation [5]. The original algorithm and its variants utilize the complex Wirtinger derivatives for calculating the gradient values used in the update steps. A direct adaptation of the WF algorithm for the phaseless reconstruction problem specified in Section II considers the following optimization problem:

$$\boldsymbol{\rho}^* = \underset{\boldsymbol{\rho}}{\operatorname{argmin}} \frac{1}{2M} \sum_{m=1}^{M} \left[\mathbf{d}(m) - \mathbf{L}_m^H \boldsymbol{\rho}\boldsymbol{\rho}^H \mathbf{L}_m\right]^2, \quad (12)$$

where $\mathbf{L}_m^H \in \mathbb{C}^N$ denotes the m^{th} measurement vector, and it is defined as $\mathbf{L}_m^H = \mathbf{G}(m,:)\text{diag}(\mathbf{u})$.

Since $\mathbf{u}$ is unknown, unlike the existing WF based phaseless recovery methods, the measurement matrix is only partially known and $\mathbf{u}$ needs to be estimated simultaneously with unknown $\boldsymbol{\rho}$. We know that the convergence to the correct unknown for the WF algorithm is guaranteed only if the forward map satisfies specific condition as shown in [7]. This condition also implies a sufficiently accurate initial value from the spectral initialization step. Due to partially known forward map in our problem, any inaccuracy in the initial value of the total field adds additional inaccuracy to the initial image estimation along with the limitation imposed by insufficient number of measurements.

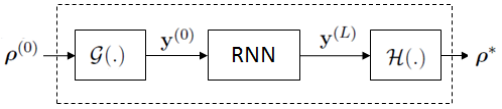

Fig. 1: Network schematic for the proposed DL based reconstruction approach.

IV. DL BASED IMAGE RECONSTRUCTION APPROACH

The goal of adapting DL for the SAR image reconstruction problem in Section II is to utilize datasets composed of phaseless measurements and corresponding reconstructed images for improving upon a basic adaptation of the WF type algorithm. This is done by estimating the unknown in a suitable DL based transformed space as well as by determining optimal values of the learning rates for the iterative updates. We denote the DL based encoding and decoding networks by $\mathcal{G}(.)$ and $\mathcal{H}(.)$, respectively. The iterative updates of a gradient descent approach for minimization in the transformed space are modelled into an RNN network similarly to the methods in [1], [3], [4]. Schematic diagram of the overall network is shown in Figure 1 where $\mathbf{y}^{(0)} \in \mathbb{R}^{N_y}$, $\mathbf{y}^{(L)} \in \mathbb{R}^{N_y}$ and $\boldsymbol{\rho}^* \in \mathbb{R}^N$ denote the encoded initial value, encoded estimated value and the decoded original unknown, respectively. Unlike the case with the Born approximation assumption, we have to consider both the inverse and the forward problem for the gradient descent updates.

A. Inverse problem

Since we are solving the inverse problem in the transformed space, by setting $\tilde{\mathcal{H}}(\mathbf{y}) = \mathcal{H}(\mathbf{y})\mathcal{H}(\mathbf{y})^T$, we consider the following optimization problem:

$$\mathbf{y}^* = \underset{\mathbf{y}}{\operatorname{argmin}} \frac{1}{2M} \sum_{m=1}^{M} \left[\mathbf{d}(m) - \mathbf{d}_y(m)\right]^2, \quad (13)$$

where $\mathbf{d}_y(m) = \mathbf{L}_m^H \tilde{\mathcal{H}}(\mathbf{y}) \mathbf{L}_m$. The term being minimized in (13) is denoted by $\mathcal{J}(\mathbf{y})$, and it is commonly referred to as the data fitting term that imposes consistency with the measured intensity values. Gradient descent approach involves calculating the gradient of $\mathcal{J}(\mathbf{y})$ with respect to unknown $\mathbf{y}$ at each iteration step. Output value at the k^{th} step is denoted by $\mathbf{y}^{(k)}$, and it is calculated as

$$\mathbf{y}^{(k)} = \mathbf{y}^{(k-1)} - \frac{\lambda_y^{k-1}}{\|\mathbf{y}^{(0)}\|^2} \boldsymbol{\nabla}_{\mathbf{y}} \mathcal{J}(\mathbf{y})\big|_{\mathbf{y}=\mathbf{y}^{(k-1)}}, \quad (14)$$

where $\lambda_y^{k-1} \in \mathbb{R}^+$ is a positive-valued constant referred to as the learning rate. The gradient of the data fitting term is calculated as

$\boldsymbol{\nabla}_{\mathbf{y}} \mathcal{J}(\mathbf{y})$

$$= \frac{2}{M} \sum_{m=1}^{M} \boldsymbol{\nabla}_{\mathbf{y}} \mathcal{H}(\mathbf{y}) \mathrm{Re}\left[\mathrm{diag}(\mathbf{u})^H (\mathbf{I} + \mathbf{H}^H \mathbf{W}_y) \mathbf{b}_m\right] e_m, \tag{15}$$

where $\mathbf{b}_m$ and e_m are defined as

$$\mathbf{b}_m = \mathbf{G}(m,:)^H \mathbf{L}_m^H \mathcal{H}(\mathbf{y}), \tag{16}$$

$$e_m = \mathbf{L}_m^H \tilde{\mathcal{H}}(\mathbf{y}) \mathbf{L}_m - \mathbf{d}(m), \tag{17}$$

for $m \in \{1, \cdots, M\}$. $\mathbf{W}_y$ is related to $\mathcal{H}(\mathbf{y})$ and $\mathbf{H}$ as $\mathbf{A}_1 \mathbf{W}_y = \mathrm{diag}(\mathcal{H}(\mathbf{y}))$ where $\mathbf{A}_1 = \mathbf{I} - \mathrm{diag}(\mathcal{H}(\mathbf{y}))\mathbf{H}^H$. The expression of $\boldsymbol{\nabla}_{\mathbf{y}} \mathcal{H}(\mathbf{y})$ depends on the architecture of the decoding network $\mathcal{H}(.)$. For calculating $\boldsymbol{\nabla}_{\mathbf{y}} \mathcal{J}(\mathbf{y})$ during each update, the values of $\mathbf{u}$ and $\mathbf{W}_y$ are required to be determined for the current value of the encoded unknown $\mathbf{y}$.

Suppose $\boldsymbol{\rho}_c$ denotes the correct unknown for the measured intensity $\mathbf{d}$, i.e. $\mathbf{d}(m) = \mathbf{L}_m^H \tilde{\boldsymbol{\rho}}_c \mathbf{L}_m$. Ideally, if we can start with the initial estimation $\boldsymbol{\rho}^{(0)}$ within an ϵ-neighbourhood of $\boldsymbol{\rho}_c$, the transformation network $\mathcal{G}(.)$ should be such that the corresponding initial value in the transformed domain, i.e. $\mathbf{y}^{(0)} = \mathcal{G}(\boldsymbol{\rho}^{(0)})$, is within a neighborhood of $\mathbf{y}_c$, where $\boldsymbol{\rho}_c = \mathcal{H}(\mathbf{y}_c)$, that is less than or equal to ϵ. For initializing $\boldsymbol{\rho}$, first the eigenvector $\mathbf{z}_{\rho_0}$ corresponding to the maximum eigenvalue of $\hat{\mathbf{X}}$ is calculated where

$$\hat{\mathbf{X}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{d}(m) \mathbf{M}_m \mathbf{M}_m^H, \tag{18}$$

for $\mathbf{M}_m^H = \mathbf{G}(m,:)\mathrm{diag}(\mathbf{u}^{in})$. Initial value $\boldsymbol{\rho}^{(0)}$ is set equal to $a_{\rho_0}\mathrm{Re}(\mathbf{z}_{\rho_0})$ where $a_{\rho_0} = \sqrt{\frac{1}{\sqrt{2}M}\|\mathbf{d}\|}$ and the initial value for the transformed domain is therefore $\mathbf{y}^{(0)} = \mathcal{G}(\boldsymbol{\rho}^{(0)})$.

B. Forward problem and gradient calculation

Since the total field deviates from the incident field due to the target multiple scattering effect, and depends on the encoded image vector as $\mathbf{u} = (\mathbf{I} - \mathbf{H}\mathrm{diag}(\mathcal{H}(\mathbf{y})))^{-1}\mathbf{u}^{in}$, its value can be estimated solving the following minimization problem:

$$\mathbf{u}^* = \underset{\mathbf{u}}{\mathrm{argmin}} \|\mathbf{A}_2 \mathbf{u} - \mathbf{u}^{in}\|^2, \tag{19}$$

where $\mathbf{A}_2 = \mathbf{I} - \mathbf{H}\mathrm{diag}(\mathcal{H}(\mathbf{y}))$. Another term required for the gradient calculation is $\mathbf{W}_y$, and its value is determined by the solutions of the following set of minimization problems:

$$\mathbf{W}_y^*(:,n) = \underset{\mathbf{W}_y(:,n)}{\mathrm{argmin}} \|\mathbf{A}_1 \mathbf{W}_y(:,n) - \mathbf{h}_{y,n}\|^2, \tag{20}$$

for $n \in \{1, \cdots, N\}$ where the n^{th} component of $\mathbf{h}_{y,n} \in \mathbb{R}^N$ equals to $\mathcal{H}(\mathbf{y})(n)$ and the rest of its elements are equal to 0's. We can implement Nesterov's method [30] and solve $N+1$ vector minimization problems parallelly to estimate $\mathbf{W}_y$ and $\mathbf{u}$. The initial value of the total field and $\mathbf{W}_y$ are set equal to $\mathbf{u}^{in}$ and an all zero $N \times N$ matrix, respectively, for the first iteration of the phase retrieval algorithm. For each additional iteration, these initial values are set equal to their corresponding values determined during the previous step. The updates can be continued until the norms of the derivatives are smaller than a tolerance value or until a fixed number of steps has been reached. Overall SAR image reconstruction algorithm is summarized in Algorithm 1.

Algorithm 1 Phase retrieval using DL based WF

1: Set $\boldsymbol{\rho}^{(0)} = a_{\rho_0}\mathrm{Re}(\mathbf{z}_{\rho_0})$ where $\mathbf{z}_{\rho_0}$ is the eigenvector for the maximum eigenvalue of $\hat{\mathbf{X}}$ and $a_{\rho_0} = \frac{1}{(2M)^{1/4}}\sqrt{\|\mathbf{d}\|}$
2: Set $\mathbf{y}^{(0)} = \mathcal{G}(\boldsymbol{\rho}^{(0)})$
3: Set $k = 1$
4: **while** $k \leq L$ **do**
5: Estimate $\mathbf{W}_{y,k}^*$, $\mathbf{u}_k^*$ using $\mathbf{y} = \mathbf{y}^{(k-1)}$
6: $\mathbf{y}^{(k)} = \mathbf{y}^{(k-1)} - \frac{\lambda_y^{k-1}}{\|\mathbf{y}^{(0)}\|^2} \boldsymbol{\nabla}_{\mathbf{y}} \mathcal{J}(\mathbf{y})|_{\mathbf{y}=\mathbf{y}^{(k-1)}}$
7: $k = k + 1$
8: $\mathbf{y}^* = \mathbf{y}^L$

V. DL NETWORK ARCHITECTURE AND TRAINING

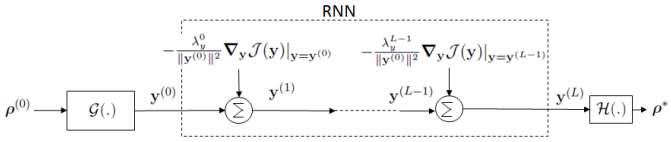

Fig. 2: Overall DL network.

$\mathcal{G}(.)$ and $\mathcal{H}(.)$ can be designed explicitly as convolutional neural networks (CNN) or by using the artificial neural network (ANN) framework. Overall network diagram is depicted in Figure 2. The parameters of $\mathcal{G}(.)$, as well as the learning rate for the iterative updates are learned by end-to-end training. Parameters of $\mathcal{H}(.)$ can be learned simultaneously with $\mathcal{G}(.)$ or separately as part of a generative adversarial network (GAN) using a set of training images. Suppose $\mathcal{P}$ denotes the set of all parameters and $\mathcal{T}$ denotes the set of all T training samples. For the t^{th} sample, we have the measured intensity vector $\mathbf{d}_t$ and the corresponding correct unknown $\boldsymbol{\rho}_t$. Assuming known fixed transmission signal and transmitter location for all samples, the incident field vector $\mathbf{u}^{in}$ for the points in Ω is fully known Training loss function $\mathcal{C}(\mathcal{P})$ is defined as

$$\mathcal{C}(\mathcal{P}) = \frac{1}{T} \sum_{t=1}^{T} \|\boldsymbol{\rho}_t^* - \boldsymbol{\rho}_t\|^2, \tag{21}$$

and it is minimized with respect to the variable set $\mathcal{P}$ via stochastic gradient descent updates during training. The number of RNN stages L, and the number of layers and the dimensions of the weight matrices and bias vectors of $\mathcal{G}(.)$ and $\mathcal{H}(.)$ are adjusted as tuning parameters.

VI. NUMERICAL RESULTS

In order to demonstrate the feasibility of our proposed approach, we present image reconstruction result on a simulated dataset, and compare with the reconstruction output from the WF algorithm [5], [7] with Born approximation. We start by considering simple scenes with a single extended rectangular object of arbitrary dimensions and locations. The task is to

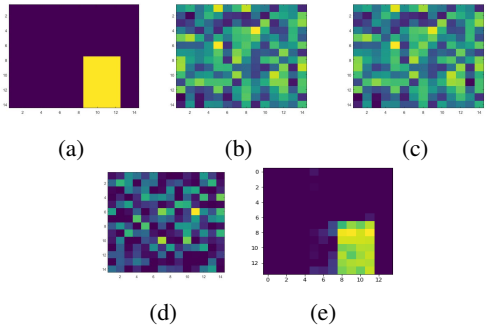

Fig. 3: (a) is an example original unknown image. For $M = 0.5N$, (b), (c), (d) and (e) denote the spectral initialization output, the reconstructed image using the WF [5] algorithm, the reconstructed image using the approach in [23] and the estimated image by the proposed method with 5 RNN stages, respectively.

image the scene located within a $280 \times 280m^2$ area into a 14×14 pixel format. The transmission signal is monotone, which is similar to the scenario considered in [23], [24] and valid for narrow-band assumption, with single frequency equal to 2GHz. A total of $M = 98$ receiving sensors around a circle of 10km radius with the center of the scene as its origin and at a height of 0.5km are considered. The transmitter sensor is located at the same fixed location for all samples. With $N = 196$ and $M = 98$, we have an under-determined system with measurements to unknown ratio of 0.5.

For the two networks $\mathcal{G}(.)$ and $\mathcal{H}(.)$, we have used CNN and ANN architectures, respectively, with the number of network layers equal to 5. The activation functions used for $\mathcal{G}(.)$ and $\mathcal{H}(.)$ are *relu* and *softplus*, both of which operate component-wise on their inputs. Nonlinear *softplus* function is defined as $softplus(\mathbf{t}) = \begin{bmatrix}\log(\mathrm{e}^{\mathbf{t}(1)}+1) & \cdots & \log(\mathrm{e}^{\mathbf{t}(N_t)}+1)\end{bmatrix}^T$ for $\mathbf{t} \in \mathbb{R}^{N_t}$, and its derivative with respect to its argument is $\mathrm{diag}(\sigma(\mathbf{t}))$ where $\sigma(.)$ is the widely used *sigmoid* function in the DL literature. Output vector length at the last stage of $\mathcal{G}(.)$ is 81, while for the five consecutive layers of $\mathcal{H}(.)$, output vector lengths are 81, 81, 81, 100 and 196, respectively. Additionally, we make the simplifying assumption that the maximum scattering potential for all the scenes are equal and known as a prior for our DL based approach as well as the WF algorithm under Born approximation. Example original image and the reconstructed image by our proposed method along with the reconstruction output using the WF approach with 100000 updates are shown in Figure 3. We also show the corresponding reconstruction output using the phaseless multiple scattering inversion method in [23] with 100000 iterations in Figure 3(d). Compared to this non-DL based approach, our proposed method results in better image quality for this low sample complexity regime.

VII. Conclusion

In this paper, we have presented a DL based phase retrieval approach for the non-linear measurement model which performs the reconstruction task in an encoded domain where the DL based encoding and decoding networks enables efficient initialization and incorporation of a generative prior, respectively. Starting with an efficient initialization, aside from solving the inverse problem in the encoded space, proposed approach additionally solves the forward problem which is necessary due to the non-linear nature of the measurement model. An important open problem is the identification of the range of forward maps and the types of scenes where the proposed method guarantees convergence to the correct unknown which we will explore in our future work.

VIII. Acknowledgements

This work was supported by the Air Force Office of Scientific Research (AFOSR) under the agreement FA9550-16-1-0234, Office of Naval Research (ONR) under the agreement N00014-18-1-2068 and by the National Science Foundation (NSF) under Grant No ECCS-1809234.

References

[1] B. Yonel, E. Mason, and B. Yazıcı, "Deep learning for passive synthetic aperture radar," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 90–103, 2017.

[2] S. Kazemi and B. Yazici, "Deep learning for joint image reconstruction and segmentation for sar," in *2020 IEEE International Radar Conference (RADAR)*, 2020, pp. 839–843.

[3] B. Yonel, E. Mason, and B. Yazici, "Deep learning for waveform estimation and imaging in passive radar," *IET Radar, Sonar & Navigation*, vol. 13, no. 6, pp. 915–926, 2019.

[4] S. Kazemi, B. Yonel, and B. Yazici, "Deep learning for direct automatic target recognition from SAR data," in *2019 IEEE Radar Conference*, 2019, pp. 1–6.

[5] 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.

[6] 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.

[7] 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.

[8] E. J. Candes, Y. C. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM review*, vol. 57, no. 2, pp. 225–251, 2015.

[9] 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.

[10] I. Waldspurger, A. d'Aspremont, and S. Mallat, "Phase recovery, maxcut and complex semidefinite programming," *Mathematical Programming*, vol. 149, no. 1-2, pp. 47–81, 2015.

[11] S. Bahmani and J. Romberg, "Phase retrieval meets statistical learning theory: A flexible convex relaxation," ser. Proceedings of Machine Learning Research, A. Singh and J. Zhu, Eds., vol. 54. Fort Lauderdale, FL, USA: PMLR, 20–22 Apr 2017, pp. 252–260.

[12] T. Goldstein and C. Studer, "Phasemax: Convex phase retrieval via basis pursuit," *IEEE Transactions on Information Theory*, vol. 64, no. 4, pp. 2675–2689, 2018.

[13] R. W. Gerchberg, "A practical algorithm for the determination of phase from image and diffraction plane pictures," *Optik*, vol. 35, pp. 237–246, 1972.

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

[15] 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.

[16] J. Chen, L. Wang, X. Zhang, and Q. Gu, "Robust wirtinger flow for phase retrieval with arbitrary corruption," *arXiv preprint arXiv:1704.06256*, 2017.
[17] H. Zhang, Y. Zhou, Y. Liang, and Y. Chi, "A nonconvex approach for phase retrieval: Reshaped wirtinger flow and incremental algorithms," *J. Mach. Learn. Res.*, vol. 18, no. 1, p. 5164–5198, Jan. 2017.
[18] 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.
[19] F. Wu and P. Rebeschini, "Hadamard wirtinger flow for sparse phase retrieval," *arXiv preprint arXiv:2006.01065*, 2020.
[20] I.-Y. Son, B. Yonel, and B. Yazici, "Multistatic passive radar imaging using generalized wirtinger flow," in *2019 20th International Radar Symposium (IRS)*. IEEE, 2019, pp. 1–7.
[21] B. Yonel, I.-Y. Son, and B. Yazici, "Exact multistatic interferometric imaging via generalized wirtinger flow," *IEEE Transactions on Computational Imaging*, vol. 6, pp. 711–726, 2020.
[22] ——, "Generalized wirtinger flow for passive polarimetric reconstruction of extended dipole targets," in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 1353–1358.
[23] M. A. Lodhi, Y. Ma, H. Mansour, P. T. Boufounos, and D. Liu, "Inverse multiple scattering with phaseless measurements," in *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2020, pp. 1519–1523.
[24] Y. Sun, Z. Xia, and U. S. Kamilov, "Efficient and accurate inversion of multiple scattering with deep learning," *Optics express*, vol. 26, no. 11, pp. 14 678–14 688, 2018.
[25] U. S. Kamilov, H. Mansour, and B. Wohlberg, "A plug-and-play priors approach for solving nonlinear imaging inverse problems," *IEEE Signal Processing Letters*, vol. 24, no. 12, pp. 1872–1876, 2017.
[26] P. Hand, O. Leong, and V. Voroninski, "Phase retrieval under a generative prior," in *Advances in Neural Information Processing Systems*, 2018, pp. 9136–9146.
[27] ——, "Compressive phase retrieval: Optimal sample complexity with deep generative priors," *arXiv preprint arXiv:2008.10579*, 2020.
[28] F. Shamshad and A. Ahmed, "Robust compressive phase retrieval via deep generative priors," *arXiv preprint arXiv:1808.05854*, 2018.
[29] F. Shamshad and A. Ahmed, "Compressed sensing based robust phase retrieval via deep generative priors," *IEEE Sensors Journal*, pp. 1–1, 2020.
[30] Y. Nesterov, "A method for solving the convex programming problem with convergence rate $\mathcal{O}(1/k^2)$," *Dokl. Akad. Nauk SSSR*, vol. 269, pp. 543–547, 1983.