Interferometric Passive Radar Imaging with Deep Denoising Priors

Samia Kazemi
Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Bariscan Yonel, Member, IEEE
Rensselaer Polytechnic Institute, Troy, NY 12180, USA

Birsen Yazici, Fellow, IEEE
Rensselaer Polytechnic Institute, Troy, NY 12180, USA

***Abstract*— Passive radar has advantages over its active counterpart in terms of cost and stealth. In this paper, we address passive radar imaging problem by interferometric inversion using a spectral estimation method with a priori information within a deep learning (DL) framework. Cross-correlating the received signals from different look directions mitigates the influence of shared transmitter related phase components despite lack of a cooperative transmitter, and permits tractable inference via interferometric inversion. To this end, we leverage deep architectures for modeling a priori information and for improving sample efficiency of state-of-the-art interferometric inversion methods. Our approach comprises of an iterative algorithm based on generalizing the power method, and applies denoisers using plug-and-play (PnP) and regularization by denoising (RED) techniques. We evaluate our approach using simulated data for passive synthetic aperture radar (SAR) by using convolutional neural networks (CNN) as denoisers, and compare our results with state-of-the-art. The numerical experiment shows that our method can achieve faster reconstruction and superior image quality in sample starved regimes than the state-of-the-art passive interferometric imaging algorithms.**

***Index Terms*—Deep learning, interferometric imaging, plug-and-play, denoiser.**

I. INTRODUCTION

A. Problem Statement

THIS paper studies passive interferometric imaging, which involves the recovery of a signal of interest from the cross-correlations of its linear measurements collected in a spatially diverse sensing geometry. For such imaging geometries, let $k = 1, \cdots K$ correspond to the frequency samples over the transmission band $\omega \in [\omega_c - B/2, \omega_c + B/2]$ used in the acquisition system, whereas $p, q \in \mathbb{S}$ index the locations of the receivers, with $|\mathbb{S}| = S$. Let $\mathbf{a}_p^k, \mathbf{a}_q^k \in \mathbb{C}^N$ denote the sampling vectors corresponding to the p^{th} and q^{th} sensors at a given frequency ω_k, and $\boldsymbol{\rho}^* \in \mathbb{C}^N$ be the ground truth/signal of interest. Consider the measurement matrix $\mathbf{A}^k$ per frequency, where $\mathbf{a}_{p,q}^k$ are the two distinct columns such that, for $k = 1, \cdots, K$,

$$\mathbf{f}^k = (\mathbf{A}^k)^H \boldsymbol{\rho}^*, \text{ with } f_p^k = \langle \mathbf{a}_p^k, \boldsymbol{\rho}^* \rangle, f_q^k = \langle \mathbf{a}_q^k, \boldsymbol{\rho}^* \rangle, \quad (1)$$

as the linear measurements at each receive location. The cross-correlated measurements from each location pair (p, q) correspond to the interferometric measurements in frequency domain, as:

$$d_{pq}^k = f_p^k \overline{f_q^k} = (\mathbf{a}_p^k)^H \boldsymbol{\rho}^* (\boldsymbol{\rho}^*)^H \mathbf{a}_q^k \quad k = 1, \cdots K, \quad (2)$$

where $\overline{(\cdot)}$ denotes complex conjugation. Thus, interferometric inversion involves recovery of $\boldsymbol{\rho}^* \in \mathbb{C}^N$ from $d_{pq}^k \in \mathbb{C},\ k = 1, ..., K$ under the quadratic model in (2).

In essence, this is equivalent to recovering $\boldsymbol{\rho}^*$ from the collection of rank-1, data scatter matrices

$$\mathbf{D}^k := \mathbf{f}^k (\mathbf{f}^k)^H = (\mathbf{A}^k)^H \boldsymbol{\rho}^* (\boldsymbol{\rho}^*)^H \mathbf{A}^k, \quad (3)$$

where $\mathbf{f}^k = [f_1^k, f_2^k, \cdots f_S^k]^T$, with (2) corresponding to the upper triangular entries of $\mathbf{D}^k$ for each k. In this generic form, interferometric inversion problem arises in many applications in different disciplines. These include radar and sonar interferometry [1]–[3], passive imaging in acoustic, electromagnetic and geophysical applications [4]–[7], and beamforming and sensor localization in large area networks [8], [9] among others. In wave-based imaging, correlations were shown to provide robustness to statistical fluctuations in scattering media or incoherent sources [10], [11], and with respect to phase errors in the correlated linear transformations [12]–[14].

In this paper, we leverage interferometric inversion for the purpose of addressing the passive radar imaging problem. Passive radar systems do not use their own dedicated transmitters, and instead use scattered ambient signals originating from a source of opportunity. As a result, passive radar systems are realizable with small mobile receivers that operate with long acquisition modes, providing spatial diversity and robustness in challenging sensing environments. In the setting that illuminators are non-cooperative, precise transmitter location and waveform are unavailable at the receive end to describe the underlying forward mapping for the inversion task. To this end, cross-correlating the measurements from different receive locations mitigate the influence of transmitter related terms by removal of the shared phase components. Hence, the main motivation for interferometric processing in passive imaging applications is to instead describe a model $\{\tilde{\mathbf{A}}^k\}_{k=1}^K$ that can accurately facilitate the inversion in lieu of the ideal but only partially known $\{\mathbf{A}^k\}_{k=1}^K$, having per $k = 1, \cdots, K$,

$$(\mathbf{A}^k)^H \boldsymbol{\rho}^* \boldsymbol{\rho}^{*H} \mathbf{A}^k \approx (\tilde{\mathbf{A}}^k)^H \boldsymbol{\rho}^* \boldsymbol{\rho}^{*H} \tilde{\mathbf{A}}^k := \tilde{\mathbf{f}}^k (\tilde{\mathbf{f}}^k)^H. \quad (4)$$

Manuscript received October 1, 2022. Revised manuscript submitted on January 14, 2023. 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. *(Corresponding author: B. Yazici)*

The authors are with the Department of Electrical, Computer and Systems Engineering, Rensselaer Polytechnic Institute, Troy, NY 12180 USA (e-mail: kazems@rpi.edu; yonelb2@rpi.edu; yazici@ecse.rpi.edu).

The key consideration of solving interferometric inversion is in mitigating the partial loss of phase information. Beyond the removal of undesirable phase components within the data, correlation operation results in fundamental limitations in direct factorization of (4). Without access to the full $KS \times KS$ data scatter matrix[1], i.e., with correlations only computed per fixed frequency, retrieval of the equalized data $\tilde{\mathbf{f}}^k$ by a rank-1 decomposition results in k-dependent arbitrary factors of $\mathrm{e}^{j\phi_k}$. This is a consequence of the quadratic nature of interferometric data matrix via invariance to global phase multipliers. As a result, direct factorization demands a crucial phase synchronization step, which increases the number of unknowns to $K + N$ and requires the use of underlying common parameterization with respect to the unknown of interest. Performance of such formulation then strongly hinges on the accurate recovery of the phase factors. This is undesirable as small phase errors are known to yield drastic errors in the reconstructed imagery [14].

Ultimately, using the underlying parameterization of the scene is necessary for the feasibility of the resulting interferometric inversion problem. This motivates approaches for direct inversion from the quadratic measurement model of (2) to avoid inducing phase ambiguity over frequency samples, which form the state-of-the art.

B. Prior Art and Motivation

Conventionally, interferometric inversion in imaging applications has been approached by Fourier based techniques, such as time or frequency difference of arrival (TDOA/FDOA) backprojection [6], [15]–[20]. While these methods are practical and computationally efficient, their applicability is limited to scenes composed of well-separated point targets due to underlying assumptions. As an alternative, low rank matrix recovery (LRMR) theory has been explored for interferometric inversion [4], [21]. Notably, these solvers are inspired by the PhaseLift method [22]–[24], hence suffer from the same drawbacks in computation and memory to semi-definite programming (SDP) in practice. In [4], an iterative optimization approach to LRMR was developed for interferometric passive imaging to circumvent the poor scaling properties of SDP approaches. While this method is more efficient than the SDP solvers, it still operates by squaring the number of unknowns, hence still requires significant memory and computational resources for imaging. Additionally, these convexified lifting based solvers require stringent theoretical conditions on the measurement model, which poses a major theoretical barrier for interferometric inversion problems with deterministic forward models.

Motivated by the reduced computational complexity and memory requirements of non-convex optimization over the lifting based methods in phase retrieval literature [25], we developed the generalized Wirtinger Flow (GWF) for interferometric inversion in [26]. Namely, GWF provides deterministic exact recovery guarantees to a general class of problems that are characterized over the equivalent lifted domain by the restricted isometry property (RIP) on the set of rank-1, positive semi-definite (PSD) matrices, while operating solely on the original signal domain. In [27], we established the sufficient condition of exact recovery for passive imaging on multi-static geometry, where we determined the physical parameters of the system to ensure exact recovery. Furthermore, we introduced theoretical framework that facilitated a resolution analysis and tractable sample complexity of interferometric wave-based imaging under the far-field and small scene assumptions, and showed that the GWF algorithm achieves super-resolution in parameter regimes that commonly correspond to passive settings using $O(\sqrt{N})$ distinct look directions.

Despite its impact in theoretical outcomes, the GWF approach has certain limitations. The data-rates from correlations grow with S^2 in number of look directions, which amounts to $\mathcal{O}(N^{3/2})$ total sample complexity in [27] for super-resolution capability, and $\mathcal{O}(N^{5/4})$ for minimal feasibility. Recently, a distributed analogue of GWF was developed in [28] with tunable graph connectivity in forming local subset of correlations within the sensor geometry, hence provides control on growth of the data-rates for inference without sacrificing performance guarantees. Still, even with $\mathcal{O}(N^{5/4})$ minimal sample complexity, there exists an $\sqrt{L}$-factor growth in establishing the sufficient condition, which indicates break-down of GWF guarantees in imaging scenes beyond a critical length. Other limitations include oversampling requirements on K, poor scaling of the sufficient condition bounds with respect to the imaging aperture/field of view, and slow convergence due to the first-order nature of the algorithm updates.

Our motivation in this paper is to address the shortcomings of GWF by leveraging a priori information in the form of a constrained spectral estimation approach [25]. In recent years, synthetic aperture radar (SAR) imaging task using back-scattered measurements, whose phase information may or may not be available, have been addressed in the literature using denoising priors under both plug-and-play (PnP) [29]–[31] and regularization by denoising (RED) [32], [33] mechanisms. In [34], [35], a PnP based ADMM [36] algorithm is introduced for spot-light mode SAR imaging. When implemented with a deep denoiser, it is empirically observed to out-perform the state-of-the-art under low measurements to unknown ratios. A RED based denoising prior is applied with the ADMM algorithm for the speckle noise removal problem from SAR images in [37]. In [38], synthetic aperture imaging problem using phaseless measurements is addressed by introducing a regularized WF approach under ADMM framework that applies total variation, BM3D and untrained deep denoiser. Although apriori information is used during the iterative updates, [38] evaluates the initial

[1]Clearly, one could compute the full scatter matrix as well. However, this would not result in the tractable model used in (4), thus would not be conductive to the inverse problem at hand.

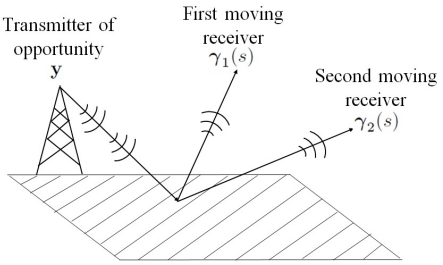

Fig. 1. Passive bistatic SAR geometry.

image by the computationally intensive spectral estimation step without taking such information into account. In general, spectral methods typically form the crucial initial step of non-convex phase retrieval techniques which uses the spectra of the back-projection estimate in the lifted domain [39]. Incorporating constraints that capture image features offers to morph the signal space during the spectral search and yield structurally sound estimates without the need of iterative updates in GWF. As a result, structural prior information provides the potential to improve the computational and acquisition efficiency of interferometric passive radar imaging by decreasing the order of required number of look directions or the oversampling factors in frequency [40]. To this end, we utilize deep architectures to learn effective representations of the signal manifold.

C. Our Approach and Contribution

Our objective in this paper is to apply deep learning (DL) for designing an interferometric imaging algorithm using a priori information. We are particularly interested in deep models to capture structure in the underlying scene with approximation capability beyond that is achievable by denoising with sparsity-based functional regularizer. For our approach, we consider regularization within the spectral estimation framework, indirectly and directly, by applying a denoising operator using PnP and RED approaches within DL frameworks, respectively.

The spectral matrix estimates in practically relevant interferometric imaging geometries may not possess the favorable characteristics such that its leading eigenvector retains important structural information about the unknown. This is especially the case when operating below the requirements identified for feasibility of the GWF theory, where the standard spectral method is not guaranteed to preserve sufficient similarity on the underlying scene of interest. Significant difficulties in denoiser training are posed in such problem settings, as it is challenging to determine the statistics for noisy training images that enables reconstruction within a few iterations to limit the computation cost, which forms the basis of PnP and RED frameworks. As a solution, we implement our denoiser based algorithms by using the unrolling technique [41], [42] with the network depth increased sequentially at each training instance. Aside from easier denoiser training, unrolling has the added benefit of reducing training data requirement, which is particularly desirable for SAR to minimize operational costs.

We implement our imaging networks for simulated passive SAR dataset by using convolutional neural networks (CNN) architectures for denoising, and compare performances to that of the spectral initialization approach, applied in the GWF algorithm [26]. Furthermore, we consider its variant using sparsity prior, as well as other state-of-the-art interferometric imaging techniques, and numerically observe the expected benefits in reduced sample complexity and faster convergence.

D. Organization

Rest of this paper is organized as follows: In Section II, we present our received signal model associated with the interferometric imaging task for passive bistatic SAR. In Section III, we present relevant background information, introduce our denoising prior and spectral estimation based imaging algorithms and discuss their DL-based implementation details. Section IV describes our observations from numerical experimentation on simulated SAR datasets. Finally, Section V concludes our paper.

II. Received Signal Model

We consider the passive bi-static SAR imaging configuration with a single stationary transmitter, and two moving airborne receivers, whose trajectories are spatially separated throughout the data-collection process[2]. This passive SAR imaging configuration is shown in Fig. 1. We assume that the area being imaged has a flat topography, and the locations within the scene are characterized by $\mathbf{x} = [\boldsymbol{x}, 0] \in \mathbb{R}^3$, where $\boldsymbol{x} \in \mathbb{R}^2$ indicates a 2-D location in ground plane. Let the stationary transmitter be located at $\mathbf{y} \in \mathbb{R}^3$, and let $\rho : \mathbb{R}^2 \mapsto \mathbb{R}$ be the ground reflectivity function. Suppose the frequency samples of the reflected wave are collected at S slow-time points by the two receivers. We represent the locations of the two receivers at slow time $s \in [S]$ by $\boldsymbol{\gamma}_1(s) \in \mathbb{R}^3$ and $\boldsymbol{\gamma}_2(s) \in \mathbb{R}^3$. For all $s \in [S]$, the trajectories of the airborne receivers are such that $\boldsymbol{\gamma}_1(s) \neq \boldsymbol{\gamma}_2(s)$. We represent the speed of transmission through the background medium and the fast-time frequency by $c_0 \in \mathbb{R}^+$ and $\omega \in \mathbb{R}^+$, respectively, where $\omega \in [\omega_c - B/2, \omega_c + B/2]$. Here, ω_c and B represent the center frequency and the transmission bandwidth, respectively.

Under the Born and start-stop approximations, the received signal model for the r^{th} receiver is [4]

$$f_r(\omega, s) = \int e^{i\frac{\omega}{c_0}\phi_r(\mathbf{x},\mathbf{y},s)} A_r(\mathbf{x}, \omega, s)\rho(\boldsymbol{x})d\boldsymbol{x}, \quad (5)$$

where $r = 1, 2$ and

$$\phi_r(\mathbf{x}, \mathbf{y}, s) = |\mathbf{x} - \boldsymbol{\gamma}_r(s)| + |\mathbf{x} - \mathbf{y}|. \quad (6)$$

[2]We note that our interferometric imaging approach similarly applies to the case with multiple stationary receivers.

The amplitude term A_r is determined by the antenna beampatterns and geometric spreading factors. The cross-correlated measurements from the two receivers evaluated at each slow-time is given as

$$d(\omega, s) = f_1(\omega, s)\overline{f_2(\omega, s)}. \tag{7}$$

Using (5), and under small-scene and far-field assumptions, the cross-correlated measurements can be modelled as [4]:

$$d(\omega, s) = \int e^{i\frac{\omega}{c_0}\phi_{12}(\mathbf{x},\mathbf{x}',\mathbf{y},s)} A_{12}(\mathbf{x}, \mathbf{x}', \omega, s)\rho(\boldsymbol{x})\bar{\rho}(\boldsymbol{x}')d\boldsymbol{x}d\boldsymbol{x}', \tag{8}$$

where

$$\phi_{12}(\mathbf{x}, \mathbf{x}', \mathbf{y}, s) = \tilde{\phi}_1(\mathbf{x}, \mathbf{y}, s) - \tilde{\phi}_2(\mathbf{x}', \mathbf{y}, s), \tag{9}$$

and

$$\tilde{\phi}_r(\mathbf{x}, \mathbf{y}, s) = |\mathbf{x} - \boldsymbol{\gamma}_r(s)| + \hat{\mathbf{y}}.\mathbf{x}, \tag{10}$$

with $r = 1, 2$ and $\hat{\mathbf{y}}$ being the unit vector in the direction of $\mathbf{y}$. $A_{12}(\mathbf{x}, \mathbf{x}', \omega, s)$ relates to the antenna beam-patterns, $J_1(\mathbf{x}, \omega)$, $J_2(\mathbf{x}', \omega)$ and $J_t(\mathbf{x}, \omega)$, as

$$A_{12}(\mathbf{x}, \mathbf{x}', \omega, s) \approx \frac{J_1(\mathbf{x}, \omega)\bar{J}_2(\mathbf{x}', \omega)C_t^2}{|\boldsymbol{\gamma}_1(s)||\boldsymbol{\gamma}_2(s)||\mathbf{y}|^2}. \tag{11}$$

$|J_t(\mathbf{x}, \omega)| \approx C_t \in \mathbb{R}^+$ under the assumption that the transmitted waveform has a flat spectrum, and -3dB beam-width encompasses the area being imaged [4], [27], [40].

Our objective is to recover ρ by using the data model in (8). Towards this objective, we proceed by first discretizing the scene into N points at $\{\boldsymbol{x}_n\}_{n=1}^N$, where $\boldsymbol{x}_n \in \mathbb{R}^2$, and define a corresponding ground truth image vector $\boldsymbol{\rho}^* \in \mathbb{C}^N$ as

$$\boldsymbol{\rho}^* = \begin{bmatrix}\rho(\boldsymbol{x}_1) & \cdots & \rho(\boldsymbol{x}_N)\end{bmatrix}^T. \tag{12}$$

Similarly, we consider K discrete fast-time frequency samples, $\{\omega_k\}_{k=1}^K$, sampled uniformly within the band $[\omega_c - B/2, \omega_c + B/2]$ to form a discretized data vector $\mathbf{d} \in \mathbb{C}^M$ with M representing the total number of measurements, i.e., $M = SK$.

Cross-correlated measurement $d(\omega_k, s)$ can be represented under this modified data model using linear sampling vector, $\mathbf{a}_r^{k,s} \in \mathbb{C}^N$ for $r = 1, 2$, $k \in [K]$ and $s \in [S]$, as

$$d(\omega_k, s) = \langle \mathbf{a}_1^{k,s}, \boldsymbol{\rho}^* \rangle \overline{\langle \mathbf{a}_2^{k,s}, \boldsymbol{\rho}^* \rangle}, \tag{13}$$

where

$$\mathbf{a}_r^{k,s} = [e^{i\frac{\omega_k}{c_0}\tilde{\phi}_r(\mathbf{x}_1,\mathbf{y},s)} A_r(\mathbf{x}_1, \omega_k, s) \quad \cdots \quad e^{i\frac{\omega_k}{c_0}\tilde{\phi}_r(\mathbf{x}_N,\mathbf{y},s)} A_r(\mathbf{x}_N, \omega_k, s)]^H,$$

with the terms from (10) and

$$A_r(\mathbf{x}_n, \omega_k, s) = \frac{J_r(\mathbf{x}_n, \omega)C_t}{|\boldsymbol{\gamma}_r(s)||\mathbf{y}|}, \tag{14}$$

for $n \in [N]$. $\mathbf{d} \in \mathbb{C}^M$ relates to $d(\omega, s)$ as

$$\mathbf{d} = \begin{bmatrix} d(\omega_1, 1) & \ldots & d(\omega_1, S) & d(\omega_2, 1) & \ldots & d(\omega_K, S)\end{bmatrix}^T. \tag{15}$$

Let $\mathcal{F} : \mathbb{C}^{N\times N} \longrightarrow \mathbb{C}^M$ be a linear lifted forward mapping operator defined such that,

$$\mathbf{d} = \mathcal{F}(\boldsymbol{\rho}^*\boldsymbol{\rho}^{*H}). \tag{16}$$

Our aim is to estimate $\boldsymbol{\rho}^*$ directly from the known cross-correlated measurement related vector, $\mathbf{d}$, and the fully-known imaging geometry related operator $\mathcal{F}$.

III. Denoising Prior and Spectral Estimation-based Interferometric Imaging Network

A. Background on Methodology

GWF for interferometric inversion is inspired by the non-convex phase retrieval algorithm in [25], [43]. It uses a two-step algorithmic approach to solve quadratic equations involving first a spectral initialization [44], and then a simple first-order iterative refinement as follows:

$$\boldsymbol{\rho}_l = \boldsymbol{\rho}_{l-1} - \frac{\mu_l}{\|\boldsymbol{\rho}_0\|^2}\boldsymbol{\nabla}\mathcal{J}(\boldsymbol{\rho})|_{\boldsymbol{\rho}=\boldsymbol{\rho}_{l-1}}. \tag{17}$$

$\mathcal{J}(\boldsymbol{\rho})$ is the quadratic objective function associated with the interferometric inversion problem given by

$$\mathcal{J}(\boldsymbol{\rho}) = \frac{1}{2M}\sum_{k,s=1}^{K,S}\left[(\mathbf{a}_1^{k,s})^H\boldsymbol{\rho}\boldsymbol{\rho}^H\mathbf{a}_2^{k,s} - d(\omega_k, s)\right]^2, \tag{18}$$

for the passive SAR problem described in Section II. The key observation of [26] is that one can guarantee sufficient accuracy of the initial *spectral estimate*, such that the simple iterations converge to the true solution, if the linear forward map in (16) satisfies the restricted isometry property over the set of rank-1, positive semi-definite (PSD) matrices with a restricted isometry constant (RIC)-$\delta \leq 0.214$.

The spectral initialization step involves setting the leading eigenvector of the following matrix $\hat{\mathbf{X}}$ as an initial estimate of $\boldsymbol{\rho}$:

$$\hat{\mathbf{X}} := \mathcal{P}_s(\mathcal{F}^H(\mathbf{d})), \tag{19}$$

where $\mathcal{P}_S(\mathbf{Z}) := 0.5(\mathbf{Z} + \mathbf{Z}^H)$ for $\mathbf{Z} \in \mathbb{C}^{N\times N}$ and

$$\mathcal{F}^H(\mathbf{d}) = \frac{1}{M}\sum_{k,s=1}^{K,S} d(\omega_k, s)\mathbf{a}_1^{k,s}(\mathbf{a}_2^{k,s})^H. \tag{20}$$

Hence, the estimate is obtained by back-projection of data on the lifted domain.

The main premise of this algorithmic framework is in the well-conditioning of the normal operator $\mathcal{F}^H\mathcal{F}$ over terms of the form $\boldsymbol{\rho}\boldsymbol{\rho}^H$, which is controlled by the RIC-δ bound used in the sufficient condition of GWF. However, this has stringent implications on the imaging geometry and the required sample complexities for the validity of theoretical arguments. Namely, the interferometric inversion by GWF becomes ill-posed at low ratios of the

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2023.3261862

number of measurements to the number of unknowns, which is quite common for high dimensional imaging problems.

One way to circumvent these shortcomings while maintaining low computational cost is to incorporate prior information about the unknown image class to the spectral estimation process. This is due to loss of information at the initialization stage since the normal operator, $\mathcal{F}^H\mathcal{F}$, does not approximate an identity map over rank-1, PSD matrices for finite number of measurements. We therefore design our approach to utilize prior information about the underlying scene of interest for attaining structurally sound estimates directly via the spectral method. As a consequence, we bypass the objective function based optimization criteria of GWF, and consider the penalty for extracting the leading eigenvector of $\hat{\mathbf{X}}$ by reformulating the power method [45], [46] with regularization.

B. Spectral Method with Prior Information

Prior information is commonly incorporated while solving an optimization problem by adding a suitable *regularization* term, $\mathcal{R}(.) : \mathbb{C}^N \mapsto \mathbb{R}$, to a *data fidelity* measure associated with the underlying data model. This leads to a modified objective function to be minimized to estimate the unknown quantity with $\mathcal{R}(\boldsymbol{\rho})$ imposing some structural prior information during the reconstruction process. In [40], we utilized a truncated power method [46] for the initialization phase of an interferometric imaging algorithm to generate an initial image under the assumption that the sparsity level k is known in advance. In general, sparse leading eigenvector estimation and the sparse principal component analysis (PCA) are well-studied problems in the literature.

In this paper, we first cast this problem as a minimization task such that the power method arises as a natural consequence of applying a proximal gradient descent (PGD) algorithm when the regularization term is not included during solution. With the regularization term included, the associated PGD algorithm becomes a realization of the power method with prior. More specifically, following optimization problem forms the basis of our DL based imaging approach:

$$\boldsymbol{\rho}^* = \underset{\boldsymbol{\rho}\in\mathbb{C}^N}{\operatorname{argmin}}\, \mathcal{J}_S(\boldsymbol{\rho}) + \mathcal{I}_n(\boldsymbol{\rho}) + \mathcal{R}(\boldsymbol{\rho}). \tag{21}$$

The data fidelity term $\mathcal{J}_S : \mathbb{C}^N \mapsto \mathbb{R}$ is defined as

$$\mathcal{J}_S(\boldsymbol{\rho}) = \frac{1}{2\gamma}\left(-\beta\boldsymbol{\rho}^H\hat{\mathbf{X}}\boldsymbol{\rho} + \boldsymbol{\rho}^H\boldsymbol{\rho}\right), \tag{22}$$

where $\gamma, \beta \in \mathbb{R}^+$. $\mathcal{I}_n : \mathbb{C}^N \mapsto \mathbb{R}^+$ is an indicator function whose output is 0 if the corresponding input vector has unit ℓ_2 norm, and it takes large values otherwise. With an explicitly defined regularization term, the proximal algorithm addressing (21) involves the following set of updates at the l^{th} iteration:

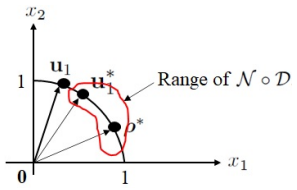

Fig. 2. Schematic diagram showing $\mathbf{u}_1$, $\mathbf{u}_1^*$ and $\boldsymbol{\rho}^*$ for $N = 2$.

$$\mathbf{w}_l = \boldsymbol{\rho}_{l-1} - \gamma\boldsymbol{\nabla}\mathcal{J}_S(\boldsymbol{\rho})|_{\boldsymbol{\rho}=\boldsymbol{\rho}_{l-1}}, \tag{23}$$

$$\mathbf{y}_l = \underset{\mathbf{x}\in\mathbb{C}^N}{\operatorname{argmin}}\, \|\mathbf{x} - \mathbf{w}_l\|^2 + \eta\mathcal{R}(\mathbf{x}), \tag{24}$$

$$\boldsymbol{\rho}_l = \mathbf{y}_l/\|\mathbf{y}_l\| = \mathcal{N}(\mathbf{y}_l), \tag{25}$$

where

$$\nabla\mathcal{J}_S(\boldsymbol{\rho}) = (-\beta\hat{\mathbf{X}}\boldsymbol{\rho} + \boldsymbol{\rho})/\gamma. \tag{26}$$

We observe that the update step in (23) simplifies to

$$\mathbf{w}_l = \beta\hat{\mathbf{X}}\boldsymbol{\rho}_{l-1}, \tag{27}$$

which is similar to the step applied during the power method updates. The normalization step in (25) accounts for the indicator function, and the proximal operator in (24) modifies $\mathbf{w}_l$ to a neighboring point that better captures the structural information imposed by $\mathcal{R}$ compared to $\mathbf{w}_l$.

However, due to the difficulty in formulating an appropriate $\mathcal{R}(.)$ in the absence of explicit prior information, and the challenges in designing an $\mathcal{R}$ that leads to preferably a closed form solution of (24), we instead design our imaging approach following the PnP and RED frameworks. For PnP, $\mathcal{R}(.)$ is not required to be explicitly defined. Instead, under the assumption that the residual noise after the update step in (23) have i.i.d. Gaussian distribution, the proximal operator in (24) can be interpreted as a denoiser for a given $\mathcal{R}$. Therefore, we can readily design a denoising prior-based power method for interferometric imaging by rewriting (24) as follows:

$$\mathbf{z}_l = \mathcal{D}(\mathbf{w}_l), \tag{28}$$

where $\mathcal{D} : \mathbb{C}^N \mapsto \mathbb{C}^N$ denotes a non-linear operator. Its output is used to calculate the image estimate $\boldsymbol{\rho}_l$ by using the normalizing operator $\mathcal{N}$, i.e.,

$$\boldsymbol{\rho}_l = \mathbf{z}_l/\|\mathbf{z}_l\| = \mathcal{N}(\mathbf{z}_l). \tag{29}$$

This PnP based formulation of our power method for interferometric imaging, presented in (23), (28) and (29), can be represented in a single step as follows:

$$\boldsymbol{\rho}_l = \mathcal{T}(\boldsymbol{\rho}_{l-1}) = \mathcal{N}\circ\mathcal{D}\left(\beta\hat{\mathbf{X}}\boldsymbol{\rho}_{l-1}\right), \tag{30}$$

where the combined operator $\mathcal{T}$ is defined as

$$\mathcal{T} = \mathcal{N}\circ\mathcal{D}\circ(\mathcal{I} - \gamma\nabla\mathcal{J}_S). \tag{31}$$

We note that $\mathcal{D}$ captures structural information about the unknown images similar to $\mathcal{R}$. Let $\mathbf{u}_1 \in \mathbb{C}^N$ denote the leading eigenvector of $\hat{\mathbf{X}}$. This algorithm attempts to recover $\mathbf{u}_1^* \in \mathbb{C}^N$, located within a small neighborhood of $\mathbf{u}_1$, such that, $\mathbf{u}_1^* \in \text{Range}(\mathcal{N}\circ\mathcal{D})$. Hence, $\mathbf{u}_1^*$ can be

interpreted as the best estimation of $\mathbf{u}_1$ that possesses the structural properties encapsulated by $\mathcal{D}$, i.e.,

$$\mathbf{u}_1^* = \underset{\mathbf{u}\in\text{Range}(\mathcal{N}\circ\mathcal{D})}{\text{argmax}} \ \mathbf{u}^H\hat{\mathbf{X}}\mathbf{u}, \tag{32}$$

and $(\mathbf{u}_1^*)^H\hat{\mathbf{X}}\mathbf{u}_1^* \leq \mathbf{u}_1^H\hat{\mathbf{X}}\mathbf{u}_1$. When $N = 2$, we show a visualization of $\mathbf{u}_1$, $\mathbf{u}_1^*$ and $\boldsymbol{\rho}^*$, with $\|\boldsymbol{\rho}^*\| = 1$, in Fig. 2. This change in the ground truth quantity from $\boldsymbol{\rho}^*$ to $\mathbf{u}_1^*$ arises from our spectral estimation based formulation, and it reveals that for all $\boldsymbol{\rho}^*$ from the image class of interest, exact recovery requires $\mathcal{D}$ to be adequately *precise* for the corresponding $\mathbf{u}_1^*$ vectors to align as closely to $\boldsymbol{\rho}^*$ as possible. For an arbitrary $\mathcal{D}$, it is difficult in general to explicitly define a corresponding regularization term for which, the PnP algorithm presented in (30) achieves the same minimum point as the one attained by the algorithm described in (23) to (25). As a consequence, $\boldsymbol{\rho}^*$ is not necessarily a minimum point of an underlying objective function anymore. Instead, we are interested in the convergence of our algorithm to a set of fixed points of the combined operator $\mathcal{T}$. Let this set be denoted by $\mathbb{F}$, i.e. $\mathbb{F} = \{\boldsymbol{\rho} \in \mathbb{C}^N : \boldsymbol{\rho} = \mathcal{T}(\boldsymbol{\rho})\}$. On the other hand, we represent the global solution set of the unconstrained interferometric inversion problem by $\mathbb{P}$, i.e., $\mathbb{P} = \{e^{i\phi}\boldsymbol{\rho}^* : \phi \in [0, 2\pi]\}$. Exact recovery for the interferometric imaging problem using our proposed algorithm in (30) therefore amounts to achieving optimal conditions on the denoiser for the given data fidelity term $\mathcal{J}_S$ in (22), such that, the iterative updates in (30) converge to an element of $\mathbb{F} \cap \mathbb{P}$.

On the other hand, under the RED framework [32], [33], the regularization term $\mathcal{R}(\boldsymbol{\rho})$ is defined explicitly as a function of the denoiser as follows:

$$\mathcal{R}(\boldsymbol{\rho}) = 0.5\boldsymbol{\rho}^H(\boldsymbol{\rho} - \mathcal{D}(\boldsymbol{\rho})). \tag{33}$$

Therefore, we can alternatively modify the power method for interferometric inversion by retaining the update steps from (23) and (25), and by using the expression of $\mathcal{R}$ from (33) in (24). Under the two conditions on the denoiser defined in [32], namely, local homogeneity and strong passivity, it is shown that the solution of the corresponding minimization problem, i.e.,

$$\mathbf{y}_l^{RED} = \underset{\mathbf{x}\in\mathbb{C}^N}{\text{argmin}} \ \|\mathbf{x} - \mathbf{w}_l\|^2 + \frac{\eta}{2}\mathbf{w}_l^T(\mathbf{w}_l - \mathcal{D}(\mathbf{w}_l)), \tag{34}$$

can be approximated as $\mathbf{r}_\infty$, where $\mathbf{r}_j$ is calculated as

$$\mathbf{r}_j = (\mathbf{r}_{j-1} + \eta\mathcal{D}(\mathbf{r}_{j-1}))/(1 + \eta), \tag{35}$$

for $j \in \{1, 2, ..., \infty\}$ and $\mathbf{r}_0$ is set equal to $\mathbf{w}_l$.

We note that our algorithm described in (30) has similarity to the projected power method presented in [47]. However, [47] implements a projection operator $\mathcal{P}$ instead of the denoising and the normalization step presented in (28) and (29), respectively. $\mathcal{P}$ is defined as $\mathcal{P}(\mathbf{z}) = \text{argmin}_{\mathbf{w}\in\text{Range}(\mathcal{H})} \|\mathbf{w} - \mathbf{z}\|^2$ with $\mathcal{H}$ being a pre-trained variational auto-encoder whose range constitutes a subset of the unit sphere. This minimization problem is addressed by iterative algorithms in [47]. For example, the Adam optimizer with 200 updates and a learning rate of 0.03 was implemented during the numerical simulations in [47]. On the other hand, for our algorithm in (30), $\mathcal{D}$ can be interpreted to be modelling a proximal operator for an unknown regularization term $\mathcal{R}$ such that $\mathcal{D}(\mathbf{z}) = \text{argmin}_{\mathbf{w}\in\mathbb{C}^N} \|\mathbf{w} - \mathbf{z}\|^2 + \mathcal{R}(\mathbf{w})$. Unlike [47], our approach in (30) does not explicitly define an associated regularization term $\mathcal{R}$ for $\mathcal{D}$, and does not apply any iterative algorithm for denoising. Our RED based formulation, on the other hand, implements the denoiser within the particular definition of $\mathcal{R}$ from (33).

C. Deep Imaging Network

We begin by noting that the two versions of our denoising prior based imaging algorithms can be implemented with or without applying DL. However, in this paper, we are aiming to introduce interferometric inversion algorithms that perform well for imaging configurations that do not necessarily satisfy the sufficient conditions for exact recovery of the state-of-the-art algorithms. In these challenging regimes, using DL can be particularly beneficial for overcoming the lack of redundancies in the measurements, as well as for potentially attaining improved reconstruction qualities, faster convergence rates, improved resolutions etc. compared to the state-of-the-art.

We use DL in two stages for implementing both of our denoising prior based algorithms, presented in (30) and in (23), (34) and (25). First, similar to the state-of-the-art PnP and RED algorithms, we adopt DNs to model the denoisers instead of using any pre-defined non-linear function for this purpose. Second, to render our algorithms suitable for the imaging configuration described in Section II, with physical parameter values such that $\mathbf{u}_1$ significantly deviates from $\boldsymbol{\rho}^*$, we adopt the unrolling technique [42]. Unrolling refers to the mapping of L number of updates of an iterative algorithm to the L stages of a recurrent neural network (RNN). Unlike existing PnP and RED based algorithms, we apply this technique instead of using pre-trained denoisers at the update steps of the algorithms implemented using specific stopping criteria. Separate denoiser training commonly proceeds using a set of noisy images generated using additive Gaussian distributed noise of different variances with the clean ground truth images. However, it is difficult in general to optimally adjust the noise levels, such that, the average number of updates necessary for convergence to the fixed points are as small as possible.

Furthermore, our proposed algorithms apply initial image vectors that are not structured or derived using any sophisticated model-based formulation. Instead, we utilize either random initialization or some pre-defined fixed normalized vector as the initial point in order to reduce the associated computation cost. Depending on the choice of the initial vector, its mapping via $\hat{\mathbf{X}}$ either may not possess the desired i.i.d. Gaussian distributed noise or its variance may be too large, such that, a denoiser of a particular capacity cannot remove it sufficiently to recover any useful structural information. In general, we can infer

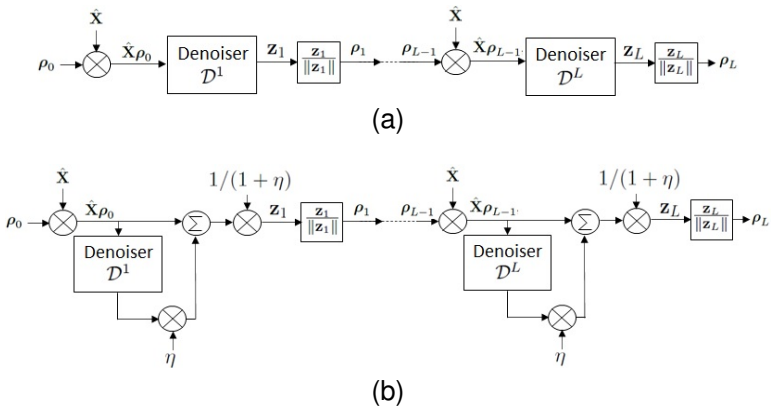

Fig. 3. Schematic diagram of imaging network designed based on (a) PnP and (b) RED algorithms.

that if the leading eigenvector $\mathbf{u}_1$ of $\hat{\mathbf{X}}$ significantly deviates from $\boldsymbol{\rho}^*$, then $\hat{\mathbf{X}}\boldsymbol{\rho}^*$ may deviate from $\boldsymbol{\rho}^*$ significantly as well, and hence may not retain much of the useful structural information present in $\boldsymbol{\rho}^*$. Furthermore, for a particular initial image and the denoiser architecture, the number of iterations required for satisfying the stopping criteria for fixed point detection can be large leading to high computation costs. Unrolling, on the other hand, can accommodate limited denoising network capacities while keeping the required number of updates as small as possible.

We unroll the L updates from (30) and derive our PnP algorithm based imaging network in Fig. 3a. It takes the spectral matrix $\hat{\mathbf{X}}$, calculated from the cross-correlated measurements, and the initial image $\boldsymbol{\rho}_0 \in \mathbb{C}^N$, with unit ℓ_2 norm, as input and generates the estimated image $\boldsymbol{\rho}_L \in \mathbb{C}^N$ at its output. The set of denoisers, $\{\mathcal{D}^l\}_{l=1}^L$, share the same DN architecture with the same set of trainable parameters. However, these parameter values are learned independently at the various update stages in order to attain an optimal set of denoisers for achieving accurate reconstruction within the fixed L updates. Similarly, for our RED based imaging algorithm in (23), (34) and (25), we unroll its L iterative update stages and derive the imaging network in Fig. 3b, by considering a single iteration of (35) for approximating (34). We set constant η as an additional trainable parameter of the imaging network aside from the set of denoiser parameters.

Both networks are trained by minimizing a loss function $c_{tr}(\mathbb{U})$, calculated as $(1/T)\sum_{t=1}^{T}\|\boldsymbol{\rho}_t^* - \boldsymbol{\rho}_{L,t}\|^2$, where $\mathbb{U}$ denotes the set of trainable parameters of the denoisers and η. Here, T refers to the number of training samples, and $\boldsymbol{\rho}_t^*$ and $\boldsymbol{\rho}_{L,t}$ are the t^{th} ground truth image and the estimated image, respectively, from the imaging networks in Fig. 3. During training, we gradually increase the number of update stages, L, to utilize as few updates as possible for imaging.

IV. Numerical Simulations

In this section, we numerically demonstrate the feasibility and performance of our deep denoising prior based imaging networks from Fig. 3, by using two simulated passive bistatic SAR datasets. For quantitatively assessing the estimated image quality, we use the normalized mean squared error (MSE) as the figure of merit. It is calculated as $\text{MSE} = (1/T_s)\sum_{t=1}^{T_s}\|\hat{\boldsymbol{\rho}}_t - \boldsymbol{\rho}_t^*\|^2/\|\boldsymbol{\rho}_t^*\|^2$, where $\hat{\boldsymbol{\rho}}_t$ refers to an estimated image for the t^{th} test sample and $\boldsymbol{\rho}_t^*$ denotes the corresponding ground truth image. Aside from feasibility verification, we have the following two major objectives for our numerical simulations:

1) Compare the performances of our PnP and RED based networks to the ones obtained using the state-of-the-art GWF algorithm, TDOA approach, as well as power method and truncated power method generated leading eigenvector estimations, where the later imposes sparsity constraint.
2) Numerically assess the performance of our approach in the presence of additive noise in the cross-correlated measurements, as well as sensor trajectory fluctuations and outliers, and empirically validate the important intuition that the deep denoising prior enables improved sample complexity and computation time compared to the state-of-the-art GWF algorithm.

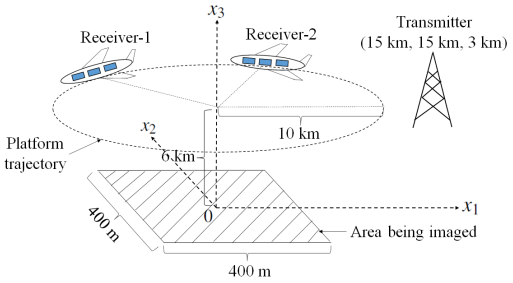

Fig. 4. Data collection model for passive bistatic SAR.

A. Dataset Description

We use two simulated image sets, each with significantly different target characteristics, and generate the corresponding sets of passive bistatic SAR data using MATLAB. We are imaging an area of dimension 400

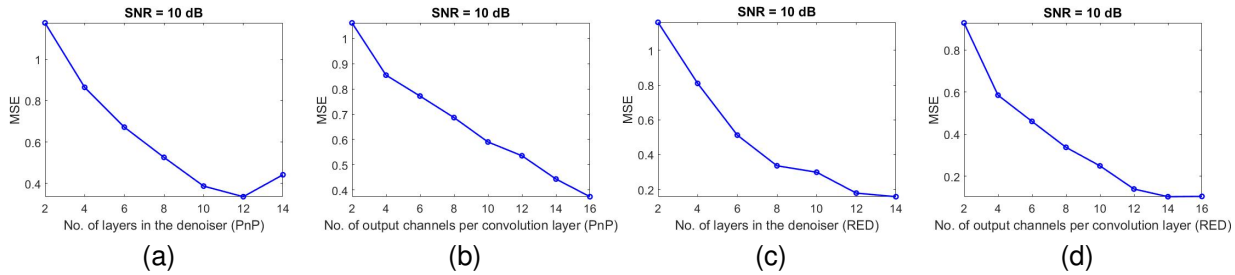

Fig. 5. Plots of MSE values versus the number of CNN layers and the number of output channels, calculated for the test dataset with rectangular targets using the PnP-based imaging network, are shown in (a) and (b), respectively. For the RED-based network and same dataset, plots of MSE versus the numbers of CNN layers and output channels per convolution layer are shown in (c) and (d), respectively. Numbers of output channels are 16 for both (a) and (c), and the numbers of convolution layers in (b) and (d) are 12 and 14, respectively.

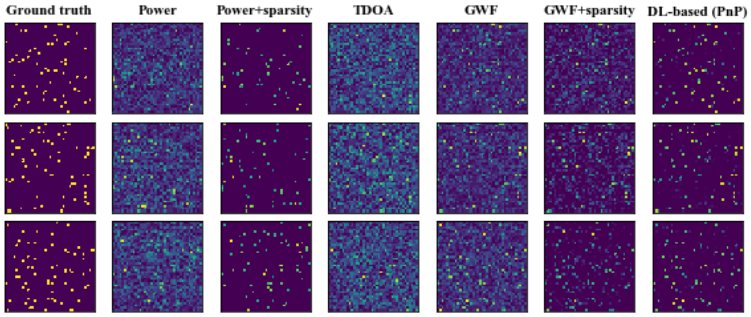

Fig. 6. For 40×40 pixel images and $M = 2N$, example reconstruction images using the power method, with and without sparsity prior, TDOA approach, GWF algorithms using 150 updates, with or without sparsity prior, and our PnP based imaging network with 8 update stages are shown in the second to seventh columns, consecutively. The first column shows the corresponding ground truth images.

m$\times$400 m, and it is being reconstructed into 40×40 pixel and 31×31 pixel images for the first and the second datasets, respectively. The first dataset includes $5,000$ training and 50 test samples, and each scene is composed of arbitrarily located random number of point targets. The second dataset, on the other hand, contains $10,000$ training and 50 test samples, and each scene contains a single randomly located rectangular target of randomly selected dimensions between 0 m to 10 m. Let the single stationary transmitter be located at $(15, 15, 3)$ km, and we assume that the two receivers are traversing the scene along a circle of radius 10 km and at 6 km height from the ground level with its origin located at the scene center. These receivers are deviated along their respective trajectories by a 45 degree angle. Imaging geometry for this passive bistatic SAR configuration is shown in Fig. 4.

B. Network Architecture and Reconstructed Images

For our imaging networks in Fig. 3, we adopt a supervised training scheme that uses the cross-correlated measurements and the corresponding ground truth images. Once an optimal set of denoiser parameters are learned, we use these values for imaging from new cross-correlated measurement vectors in the test dataset. We use CNNs as denoisers, while implementing our imaging networks for both datasets, due to the superior performance that this network architecture offers in a wide range of image processing tasks including denoising. For the first dataset, we use 16 layer CNNs with 3×3 dimensional convolution filters at the various layers. We use 16 output channels, and apply leaky_relu(.) activation functions at the hidden layers and relu(.) activation at the output layer. For a scalar input b and $\alpha \in \mathbb{R}^+$, leaky_relu(.) is defined as follows:

$$\text{leaky\_relu}(b) = \begin{cases} b & b \geq 0, \\ -\alpha b & b < 0. \end{cases} \tag{36}$$

relu(.) is an appropriate choice for the activation functions at the output layer since we are assuming that the unknown image vectors have all positive real-valued entries. On the other hand, we apply leaky_relu(.) at each hidden layer output for activation since, unlike relu(.), leaky_relu(.) allows small negative values to propagate through the layers.

For the second dataset, we implement the PnP based network from Fig. 3a by using a 12-layer CNN with 3×3 convolution filter dimensions and 16 output channels at the hidden layers. On the other hand, we implement our RED-based network from Fig. 3b by using a similar denoiser architecture, except for the number of convolution layers which is now increased to 14. For each case, we optimized our denoiser architecture by minimizing the MSE value over a range of the number of CNN layers and a range of the number of output channels or filters per layer as described for rectangular targets dataset in Fig. 5. Similar to the PnP-based network described earlier for the first dataset, we use the same activation functions

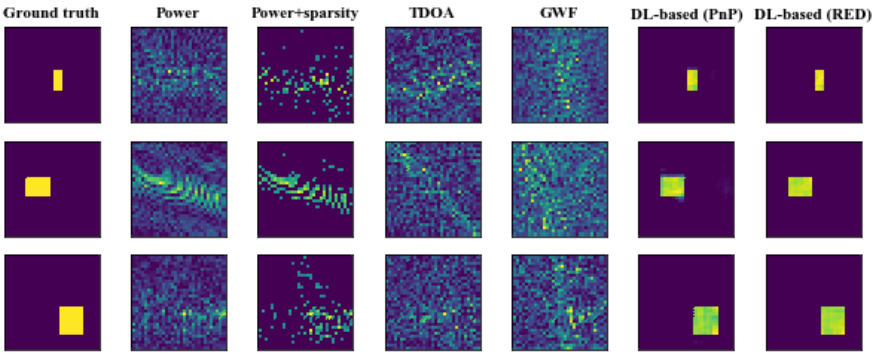

Fig. 7. For 31×31 pixel images, and using $M = 1.07N$ and 10 dB SNR, example ground truth and reconstructed images using the power method, power method with sparsity prior, TDOA approach, GWF algorithm, and our imaging networks based on PnP and RED algorithms using 4 update stages are shown in the seven columns.

at the hidden and the output layers of these two networks as well. Moreover, in all our implementations, we use the same set of denoiser parameter values for each two consecutive stages of our imaging networks, and use L equal to 8 and 4, respectively, for the point targets and the rectangular target datasets unless otherwise specified.

Example reconstruction results using our PnP based network for the two datasets are shown in the last column in Fig. 6 and the second last column in Fig. 7. The last column in Fig. 7 shows the estimated images using our RED based network. Original ground truth images are shown in the first columns in both figures. All the estimated images in Fig. 6 are obtained using 80 slow-time and 40 fast-time points. On the other hand, for the second dataset with single rectangular objects, we collected the measurements at 32 slow-time and 32 fast-time points, and considered additive Gaussian noise on the correlated measurements with 10 dB SNR. The second to fourth columns in Fig. 6 and 7 display the interferometric inversion results using the power method based spectral estimation, power method augmented by proximal operator under sparsity prior, TDOA approach and the GWF algorithm using 150 updates, respectively. The sixth column in Fig. 6 considers the GWF updates with thresholding steps. We observe that for both datasets, our proposed deep denoising prior based imaging algorithm outperforms the state-of-the-art. Aside from improved reconstruction quality, we observed improved computation time offered by our imaging network. For example, the GWF implementation in Fig. 7 using 150 updates takes 12.4968 s per test samples, whereas our PnP based imaging network takes only 0.001 s.

C. Sample Complexity

We perform training and reconstructions for the dataset with rectangular objects by applying the same PnP based imaging network using different numbers of measurements. The resulting MSE values versus the M/N ratios, using training and test datasets with the cross-correlated measurements being corrupted by additive noise of 10 dB SNR, are displayed in Fig. 8a. We observe that for attaining good reconstruction quality using our PnP based network, it is important to have sufficiently large M for a particular image dimension. However, at each M/N value, reconstruction quality obtained using our approach is significantly better compared to that of the GWF algorithm. For example, at the 5 consecutive M/N values in Fig. 8a, the MSE values obtained using the GWF algorithm are 1.6554, 1.6752, 1.6229, 1.6099 and 1.5416, respectively. Similarly, for the TDOA approach, these values are 1.7255, 1.7238, 1.6987, 1.7130 and 1.6982, respectively. This empirically shows that our deep denoising and spectral estimation based approach can overcome the strict sample complexity limitation imposed by the GWF algorithm.

Next, we consider a high additive noise scenario with -10 dB SNR. To counteract the more noisy measurements compared to the previous case, we vary the M/N ratios from 1.07 to 5 in this case. We also increase the value of L from 4 to 6 in our PnP-based imaging network implementation, and the corresponding MSE versus M/N plot is shown in Fig. 8b. We observe that although the MSE values are higher for our selected denoising network architecture relative to the ones in Fig. 8a, our approach is still better at mitigating the adverse effect of noisy measurements compared to the TDOA approach and the GWF algorithm at various M/N ratios.

D. Effect of SNR

To compare the performance of our approach with the state-of-the-art under varying levels of additive noise, at the relatively low M/N ratios that we are interested in, we set $M/N = 1.07$ and gradually increase the SNR level from 0 dB to 40 dB for the rectangular targets dataset. We applied our PnP based imaging network with $L = 4$ and the denoiser architecture described earlier for this dataset, and plotted the corresponding normalized MSE values in Fig. 8c. We observe that as expected, higher SNR values are conducive to better reconstruction quality by our imaging network. By comparing with the equivalent plot obtained for the GWF algorithm, we also observe that our approach shows better robustness to

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2023.3261862

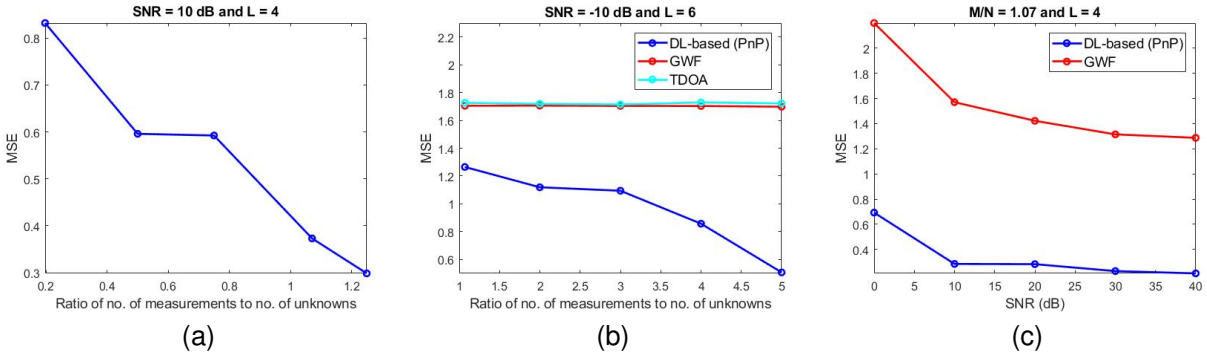

Fig. 8. Plots of MSE values versus the M/N ratios in (a) and (b) for 10 dB and -10 dB SNR values, respectively, and versus the SNR values in (c), all calculated for the test dataset with rectangular targets.

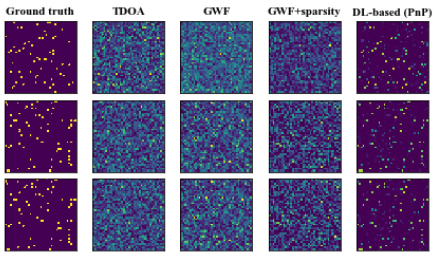

Fig. 9. For the point target dataset with additive noise of 0 dB SNR and $M = 5N$, reconstructed images using TDOA, GWF, GWF with sparsity prior and our PnP based network with $L = 8$ are shown in the second to fifth columns, respectively.

additive noise than the state-of-the-art. Next, we consider the same three example scenarios shown in Fig. 6 for the point target dataset, but introduce additive noise of 0 dB SNR. Furthermore, we increase the slow-time samples such that $M = 5N$. The resulting reconstructed images using the TDOA approach, GWF algorithm, with and without sparsity prior, and our PnP based imaging network, with $L = 8$ and the denoiser model described earlier for this dataset, are shown in Fig. 9. Similarly to the extended target dataset, we observe that our approach significantly out-performs the state-of-the-art under noisy measurements.

E. Effects of Outlier and Trajectory Fluctuations

In this subsection, aside from the random additive noise of 10 dB SNR present in the cross-correlated measurements from the training and test datasets with rectangular targets, we consider additional sources of errors for the test dataset only. Our objective is to analyze the effects of various commonly encountered sources of errors in passive setting, including trajectory error and multi-path interference. We first consider sinusoidal fluctuations of various amplitudes on the circular trajectories of the two receivers and plot the resulting MSE values in Fig. 10a by using using $L = 4$ and the denoiser model architectures described earlier for the PnP and RED based networks. Moreover, we retrained these two networks using the same training and test images, and the same additive noise level, for a passive bistatic geometry with straight line trajectories for the two receivers. We then modelled the same trajectory fluctuations for both receivers as $\mathbf{t}_e + \mathbf{v}_e s + \mathbf{a}_e s^2$, with $\mathbf{t}_e, \mathbf{v}_e, \mathbf{a}_e \in \mathbb{R}^3$. We display the corresponding MSE values in the bar graph in Fig. 10b for the following six cases of $(\mathbf{t}_e, \mathbf{v}_e, \mathbf{a}_e)$ combinations: $(\mathbf{0}, \mathbf{0}, \mathbf{0})$ m, $([0, -5, 0], \mathbf{0}, \mathbf{0})$ m, $([0, -10, 0], \mathbf{0}, \mathbf{0})$ m, $(\mathbf{0}, [0, 0.05, 0.02], \mathbf{0})$ m, $([0, -5, 0], [0, 0.05, 0.02], \mathbf{0})$ m and $(\mathbf{0}, [0, 0.05, 0.02], [0, 0.002, -0.002])$ m, with $\mathbf{0}$ denoting a length 3 vector of all-zeros. We observe that our networks are susceptible to trajectory errors due the end-to-end supervised training aspect of our implementation. Moreover, as we are introducing these errors only in the test dataset, our test data distributions differ from the ones associated with the training dataset in these cases. However, at the low M/N ratio under consideration, the corresponding MSE values using the GWF algorithm at the various entries of the plots in Fig. 10a and 10b remain significantly larger than the ones for our DL-based approach. For example, for the sinusoidal fluctuations considered in Fig. 10a, the MSE values for the GWF algorithm at the five entries are 1.4985, 1.6331, 1.6219, 1.6396 and 1.6302, respectively.

Next, we model any additional sources of errors, aside from trajectory fluctuations and additive noise, as outliers present in a random subset of the M cross-correlated measurements. For these randomly selected subset of indices, we change the magnitudes of the associated measurements to two times the maximum of the M magnitudes. We plot the resulting MSE values for different percentages of outliers in Fig. 10c, and observe that the performances of our networks deteriorate with increasing outliers levels as expected. However, the corresponding MSE values obtained using the GWF algorithm at the various entries of this plot are 1.4985, 1.6341, 1.6639, 1.6798, 1.6759 and 1.6820, respectively, which are still significantly larger, and hence indicate poorer reconstruction qualities.

We similarly analyze the effects of sinusoidal trajectory errors and outliers on the point target dataset, and present our observations in Fig. 11a and 11b. Aside from outperforming the state-of-the-art, from the example reconstructed images under 5% outliers and sinusoidal

This article has been accepted for publication in IEEE Transactions on Aerospace and Electronic Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TAES.2023.3261862

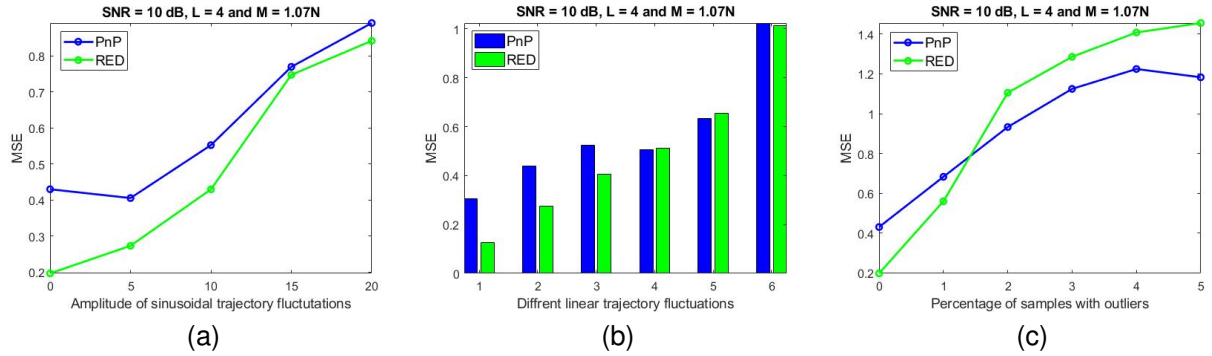

Fig. 10. Plots of MSE versus the amplitude of sinusoidal trajectory fluctuations in (a), various instances of linear trajectory fluctuations in (b) and percentage of samples with outliers in (c), each calculated for the test dataset with extended rectangular targets.

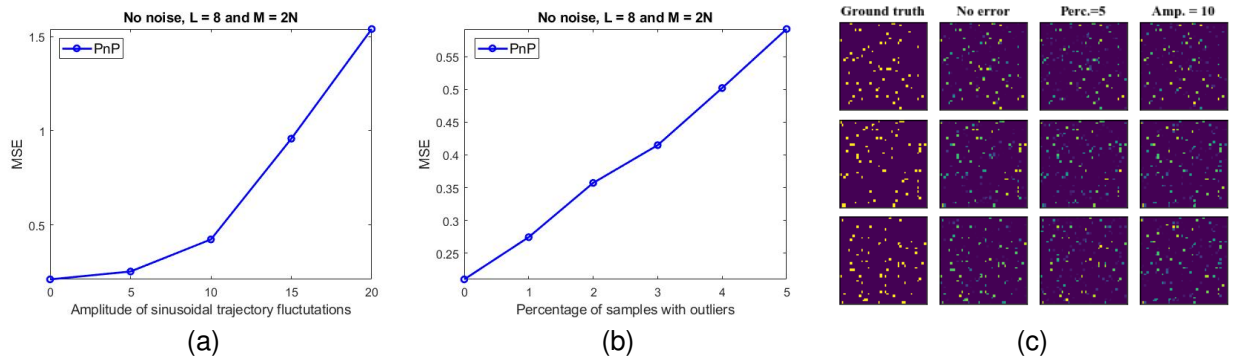

Fig. 11. For the point target dataset with $M = 2N$, no additive noise and using PnP based network with $L = 8$, plots of MSE versus amplitude of sinusoidal fluctuations on the circular trajectories and the percentages of samples with outliers are shown in (a) and (b), respectively. Three ground truth images and reconstructed images under no error, 5 percentage outliers and sinusoidal trajectory error with 10 m amplitude are shown in the four columns in (c).

trajectory fluctuations of 10 m amplitude in Fig. 11c, we observe that despite the deviations from the corresponding reconstructions in Fig. 6, our imaging network retains the point target characteristics of the images under these errors.

V. Conclusions

In this paper, we presented two deep denoising prior based interferometric imaging networks whose architectures are rooted in underlying iterative algorithms minimizing either an implicit or an explicit objective function. The data fidelity term that we used in both algorithms leads to their interpretations as expansions of the power method with learned prior knowledge imposed by the denoisers. Furthermore, we adopted the unrolling technique to facilitate easier denoiser training while applying as few iterative updates as possible and using arbitrary initializations under challenging problem regimes. Our numerical simulation results empirically demonstrate several benefits offered by our approach over the state-of-the-art including improved accuracy, computation time, sample complexity and robustness to additive noise.

As a future work, we aim to substantiate our empirical observations with theoretical justifications by studying the sufficient conditions for exact recovery guarantees of our PnP and RED based algorithms. Furthermore, determining the valid ranges of values that such conditions impose on important SAR imaging parameters for feasibility is another open question to be explored in a future work. Finally, it is important to establish the minimum requirements on the imaging network architectures for guaranteeing specific accuracy levels.

REFERENCES

[1] R. M. Goldstein and H. Zebker, "Interferometric radar measurement of ocean surface currents," *Nature*, vol. 328, no. 6132, pp. 707–709, Aug. 1987.

[2] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse Probl.*, vol. 14, no. 4, pp. R1–R54, Feb. 1998.

[3] T. O. Saebo, "Seafloor depth estimation by means of interferometric synthetic aperture sonar," Ph.D. dissertation, Dept. of Phys. and Technol., Univ. Tromso, Tromso, Norway, 2010.

[4] E. Mason, I. Y. Son, and B. Yazıcı, "Passive synthetic aperture radar imaging using low-rank matrix recovery methods," *IEEE J. Sel. Topics Sig. Process.*, vol. 9, no. 8, pp. 1570–1582, Dec 2015.

[5] L. Wang and B. Yazici, "Bistatic synthetic aperture radar imaging of moving targets using ultra-narrowband continuous waveforms," *SIAM J. Imag. Sci.*, vol. 7, no. 2, pp. 824–866, May 2014.

[6] S. Wacks and B. Yazici, "Passive synthetic aperture hitchhiker imaging of ground moving targets - Part 1: Image formation

and velocity estimation," *IEEE Trans. Image Process.*, vol. 23, no. 6, pp. 2487–2500, Jun. 2014.

[7] H. Ammari, J. Garnier, and W. Jing, "Passive array correlation-based imaging in a random waveguide," *Multiscale Model. Simul.*, vol. 11, no. 2, pp. 656–681, Jun. 2013.

[8] N. Patwari, J. N. Ash, S. Kyperountas, A. O. Hero, R. L. Moses, and N. S. Correal, "Locating the nodes: Cooperative localization in wireless sensor networks," *IEEE Signal Process. Mag.*, vol. 22, no. 4, pp. 54–69, Jul. 2005.

[9] P. Stoica, J. Li, and Y. Xie, "On probing signal design for mimo radar," *IEEE Trans. Signal Process.*, vol. 55, no. 8, pp. 4151–4161, Aug. 2007.

[10] J. Garnier, "Imaging in randomly layered media by cross-correlating noisy signals," *Multiscale Model. Simul.*, vol. 4, no. 2, pp. 610–640, Jul. 2005.

[11] O. I. Lobkis and R. L. Weaver, "On the emergence of the green's function in the correlations of a diffuse field," *J. Acoust. Soc. Am.*, vol. 110, no. 6, pp. 3011–3017, Dec. 2001.

[12] P. Blomgren, G. Papanicolaou, and H. Zhao, "Super-resolution in time-reversal acoustics," *J. Acoust. Soc. Am.*, vol. 111, no. 1, pp. 230–248, Jan. 2002.

[13] P. T. Gough and M. A. Miller, "Displaced ping imaging autofocus for a multi-hydrophone sas," *IEE Proc. - Radar, Sonar and Navigation*, vol. 151, no. 3, pp. 163–170, Jul. 2004.

[14] E. Mason and B. Yazici, "Robustness of LRMR based passive radar imaging to phase errors," in *Proc. 11th Eur. Conf. Synthetic Aperture Radar (EUSAR'16)*, Hamburg, Germany, Jun. 2016, pp. 1–4.

[15] C. E. Yarman, L. Wang, and B. Yazici, "Doppler synthetic aperture hitchhiker imaging," *Inverse Probl.*, vol. 26, no. 6, pp. 1–26, Jun. 2010.

[16] L. Wang, C. E. Yarman, and B. Yazici, "Doppler-Hitchhiker: A novel passive synthetic aperture radar using ultranarrowband sources of opportunity," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3521–3537, Oct. 2011.

[17] L. Wang and B. Yazici, "Passive imaging of moving targets exploiting multiple scattering using sparse distributed apertures," *IOP Inverse Problems J.*, vol. 28, no. 12, pp. 1–36, Dec. 2012.

[18] ——, "Ground moving target imaging using ultranarrowband continuous wave synthetic aperture radar," *IEEE Trans. Geosci. and Remote Sens.*, vol. 51, no. 9, pp. 4893–4910, Sep. 2013.

[19] S. Wacks *et al.*, "Doppler-DPCA and doppler-ATI: Novel SAR modalities for imaging of moving targets using ultra-narrowband waveforms," *IEEE Trans. Comput. Imag.*, vol. 4, no. 1, pp. 125–136, Mar. 2018.

[20] S. Wacks and B. Yazici, "Passive synthetic aperture hitchhiker imaging of ground moving targets - Part 2: Performance analysis," *IEEE Trans. Image Process.*, vol. 23, no. 9, pp. 4126–4138, Sep. 2014.

[21] L. Demanet and V. Jugnon, "Convex recovery from interferometric measurements," *IEEE Trans. Comput. Imag.*, vol. 3, no. 2, pp. 282 – 295, Jun. 2017.

[22] E. J. Candes, Y. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 199–225, Feb. 2013.

[23] E. J. Candes and T. Strohmer, "Phaselift: Exact and stable recovery from magnitude measurements via convex programming," *Commun. Pure and Appl. Math.*, vol. 66, no. 8, pp. 1241–1274, Aug. 2013.

[24] L. Demanet and P. Hand, "Stable optimizationless recovery from phaseless linear measurements," *J. Fourier Anal. and Appl.*, vol. 20, no. 1, pp. 199–221, Feb. 2014.

[25] 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.

[26] 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.

[27] 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.

[28] S. M. Farrell, A. Veeraraghavan, A. Sabharwal, and C. A. Uribe, "Distributed generalized Wirtinger flow for interferometric imaging on networks," *arXiv preprint arXiv:2206.03984*, 2022.

[29] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *2013 IEEE Global Conference on Signal and Information Processing*. IEEE, 2013, pp. 945–948.

[30] 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.

[31] Y. Sun, B. Wohlberg, and U. S. Kamilov, "An online plug-and-play algorithm for regularized image reconstruction," *IEEE Transactions on Computational Imaging*, vol. 5, no. 3, pp. 395–408, 2019.

[32] Y. Romano, M. Elad, and P. Milanfar, "The little engine that could: Regularization by denoising (red)," *SIAM Journal on Imaging Sciences*, vol. 10, no. 4, pp. 1804–1844, 2017.

[33] C. Metzler, P. Schniter, A. Veeraraghavan *et al.*, "prDeep: robust phase retrieval with a flexible deep network," in *International Conference on Machine Learning*. PMLR, 2018, pp. 3501–3510.

[34] M. B. Alver, A. Saleem, and M. Çetin, "Plug-and-play synthetic aperture radar image formation using deep priors," *IEEE Transactions on Computational Imaging*, vol. 7, pp. 43–57, 2021.

[35] M. B. Alver, A. Saleem, and M. Çetin, "A novel plug-and-play SAR reconstruction framework using deep priors," in *2019 IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–6.

[36] Y. Wang, W. Yin, and J. Zeng, "Global convergence of admm in nonconvex nonsmooth optimization," *Journal of Scientific Computing*, vol. 78, no. 1, pp. 29–63, 2019.

[37] S. Baraha and A. K. Sahoo, "Restoration of speckle noise corrupted SAR images using regularization by denoising," *Journal of Visual Communication and Image Representation*, p. 103546, 2022.

[38] K. Aditi, A. A. Kumar, A. Majumdar, T. Chakravarty, and K. Kumar, "Phaseless passive synthetic aperture imaging with regularized wirtinger flow," in *2022 30th European Signal Processing Conference (EUSIPCO)*. IEEE, 2022, pp. 1926–1930.

[39] 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.

[40] A. Akhtar, B. Yonel, and B. Yazici, "Passive multistatic radar imaging with prior information," in *2021 IEEE Radar Conference (RadarConf21)*. IEEE, 2021, pp. 1–6.

[41] 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.

[42] 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.

[43] 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.

[44] P. Netrapalli, P. Jain, and S. Sanghavi, "Phase retrieval using alternating minimization," in *Adv. in Neural Inf. Process. Syst. 26*, Lake Tahoe, NV, Dec. 2013, pp. 2796–2804.

[45] M. Journée, Y. Nesterov, P. Richtárik, and R. Sepulchre, "Generalized power method for sparse principal component analysis." *Journal of Machine Learning Research*, vol. 11, no. 2, 2010.

[46] X.-T. Yuan and T. Zhang, "Truncated power method for sparse eigenvalue problems." *Journal of Machine Learning Research*, vol. 14, no. 4, 2013.

[47] Z. Liu, J. Liu, S. Ghosh, J. Han, and J. Scarlett, "Generative principal component analysis," *arXiv preprint arXiv:2203.09693*, 2022.