Unrolled Optimization With Deep Learning-Based Priors for Phaseless Inverse Scattering Problems

Samruddhi Deshmukh, Student Member, IEEE, Amartansh Dubey, Member, IEEE, and Ross Murch, Fellow, IEEE

Abstract—Inverse scattering problems (ISPs), such as those in electromagnetic imaging using phaseless data (PD-ISPs), involve imaging objects using phaseless measurements of wave scattering. Such inverse problems can be highly nonlinear and ill-posed under extremely strong scattering conditions as such when the objects have very high permittivity or are large in size. In this work, we propose an end-to-end reconstruction framework using unrolled optimization with deep priors to solve PD-ISPs under very strong scattering conditions. We incorporate an approximate linear-physics-based model into our optimization framework along with a deep learning (DL)-based prior and solve the resulting problem using an iterative algorithm which is unfolded into a deep network. This network not only learns data-driven regularization but also overcomes the shortcomings of approximate linear models and learns nonlinear features. More important, unlike the existing PD-ISP methods, the proposed framework learns optimum values of all the tunable parameters (including multiple regularization parameters) as a part of the framework. Results from simulations and experiments are shown for the use case of indoor imaging using 2.4-GHz phaseless Wi-Fi measurements, where the objects exhibit extremely strong scattering and low absorption. Results show that the proposed framework outperforms the existing model-driven and data-driven techniques by a significant margin and provides up to 20 times higher validity range.

Index Terms—Data driven regularization, deep learning (DL), indoor imaging, inverse scattering, RF sensing, unrolled optimization.

I. INTRODUCTION

INVERSE scattering problems (ISPs) in electromagnetic imaging can be highly nonlinear and ill-posed under strong scattering conditions and have therefore found limited use in fields such as large-scale microwave imaging and indoor radio imaging. The nonlinearity and ill-posedness of ISPs can vary based on factors such as the size and relative permittivity ($\varepsilon_r$) of objects, the incident wavelength ($\lambda$), and the number and type of measurements available. They also increase significantly when the measurements are phaseless [1], [2], [3]. Thus, phaseless data ISPs (PD-ISPs) are more challenging to solve than full data ISPs (FD-ISPs).

A. ISP With Phaseless Data

Practical solutions of PD-ISPs would be extremely useful, especially for applications such as indoor imaging using microwave frequencies where it is difficult to collect accurate phase measurements. Removing the need for phase can significantly reduce the cost and complexity of the measurement system and remove the need for synchronization between measurement nodes.

There has been limited research in solving PD-ISPs, mainly because they are highly nonlinear and ill-posed and there is no exact formulation for its forward model [2], [3]. Several linear and nonlinear techniques have been proposed, and since PD-ISPs are inherently nonlinear, the nonlinear techniques achieve state-of-the-art performance in terms of validity range. However, these techniques are computationally expensive and are sensitive to experimental noise and errors, and therefore need controlled environments such as anechoic chambers to minimize noise and multipath reflections from clutter (such as nearby objects, walls, ceiling, and floor). On the other hand, linear models such as the phaseless Rytov approximation (RA) have a lower validity range, but they provide more stable reconstructions even with imperfect experimental data. Thus, there is a trade-off between the linear and nonlinear models when it comes to handling nonideal experimental data and achieving a better validity range.

In our recent work [4], we propose a linear model denoted as the extended phaseless RA in lossy media (xPRA-LM) which extends this trade-off. In terms of shape reconstruction, xPRA-LM can handle objects with extremely large permittivity and sizes. However, in terms of estimating permittivity, it fails under very strong scattering conditions as it relies on a linear approximation.

Motivated by the practical feasibility and the higher validity range of linear models such as xPRA-LM, this work aims to develop a deep learning (DL) framework that incorporates the physics-based xPRA-LM model to achieve a higher validity range. This can be achieved by learning the nonlinear relationship between the measurements and the object permittivity profile, while maintaining the ability to handle imperfectexperimental data.

In Section I-B, we provide a brief overview of the existing DL techniques to solve ill-posed inverse problems.
B. DL for Inverse Problems

Consider a general inverse problem of the form

$$\min_x f(y - A(x))$$  \hspace{1cm} (1)$$

where $y$ is an $M \times 1$ measurement vector, $A$ is the $M \times N$ forward model that can operate linearly or nonlinearly on the $N \times 1$ variable $x$ (which we wish to retrieve), and $f$ is the penalty function. The straightforward looking problem (1) can be difficult to solve due to three reasons: 1) the ill-posedness arising from scarcity of measurements ($M \ll N$) and noisy measurements; 2) the nonlinear relationship between $A$ and $x$; and 3) the unavailability or inadequate knowledge of the underlying forward model $A$. Inverse problems of the form (1) in different fields can have one or more of these problems. For example, inverse problems in medical imaging are ill-posed, but the underlying model is usually known and is linear. In classical inverse scattering, FD-ISPs are ill-posed and nonlinear, but the underlying nonlinear model is known. However, PD-ISPs are more difficult to solve since they have all the three problems—they are ill-posed, nonlinear, and the exact underlying model is unknown.

The existing DL techniques solve inverse problems by tackling one or more of the aforementioned issues. The approaches can be broadly categorized as follows.

1) Direct Inversion: Direct inversion (DI) techniques are fully data-driven/black-box approaches that attempt to directly map measurements $y$ to reconstructions $x$ using a DL network. The measurements, which are typically 1-D, often need to be mapped to 2-D or even 3-D reconstructions. Also, since these techniques do not explicitly use any model information, the network needs to learn the forward model $A$ entirely from the training data. Since these techniques completely rely on data to learn to tackle the ill-posedness and nonlinearity and to estimate the forward model, they are highly data-intensive and lead to poor generalization.

2) Model-Based DL: These techniques use a fully/partially known forward model $A$ to solve the inverse problem which is then enhanced using DL. They can be further classified into two categories.

1) Model-Assisted DL (MADL): In MADL techniques, a regularized pseudoinverse of a linearized forward model is used to provide an approximate reconstruction, which is then processed using the DL networks to obtain an enhanced final reconstruction [2], [5], [6], [7]. Therefore, the DL network in MADL learns to tackle ill-posedness and nonlinearity, while the forward model is known. These techniques are less data-intensive when compared with DI and provide better generalization due to the use of the forward model.

2) Model-Guided DL (MGDL): These techniques rely on the forward model $A$ more strongly by incorporating it into an optimization framework or network architecture. A recent MGDL approach uses data-driven DL priors on the reconstructions while minimizing $f(y - A(x))$ as opposed to using conventional handcrafted priors such as Tikhonov, least absolute shrinkage and selection operator (LASSO), or total variation (TV). These data-driven priors can be further categorized based on their training mechanism as follows.

a) Plug and Play priors [8], [9], [10] where a DL network is first trained to learn structure in $x$, following which it is plugged into the objective in (1). The resulting objective is solved using the iterative optimization techniques.

b) Priors Trained During Optimization [11], [12], [13] where the deep prior is a part of the objective function from the beginning and it is trained while the optimization problem is solved. This is usually performed using the unrolled optimization techniques. Learning deep priors through unrolled optimization leads to better generalization when compared with using plug and play priors and also removes the need to tune parameters of the optimization algorithm.

Based on the characteristics of the DL approaches described above, we use the deep prior approach to provide enhanced solutions to our PD-ISP.

C. Motivation and Contributions

Most of the research on deep priors has been performed in the fields of medical imaging and image denoising or enhancement. Since the underlying forward model in these problems is linear in most cases, the key focus of most MGDAL techniques is to reduce the error caused by modeling errors, noise, and ill-posedness. PD-ISPs, on the other hand, are different and in this work we attempt to solve PD-ISP of the form

$$\min_x f(|y| - \tilde{A}_p x)$$  \hspace{1cm} (2)$$

where we wish to reconstruct the $N \times 1$ variable $x$ with an $M \times 1$ magnitude-only measurement vector $|y|$. The key difference is that the $M \times N$ matrix $\tilde{A}_p$ here is a linear forward model that approximates an unknown underlying forward model $A$ which is highly nonlinear. As mentioned in Section I-A, the extent of the nonlinearity of the actual forward model also depends on values of the elements of $x$ which are a function of the relative permittivity $\epsilon_r$ of the object. As $\epsilon_r$ increases, the nonlinearity of the PD-ISP increases and the accuracy with which $\tilde{A}_p$ approximates $A$ decreases.

To the best of our knowledge, there have been two attempts to solve PD-ISPs using DL [2], [6], both of which use the MADL approach to process noisy reconstructions obtained from a regularized pseudoinverse of the forward model. However, there are a few issues that these techniques do not address. Xu et al. [2] use a nonlinear iterative forward model to generate the initial reconstructions, which is computationally expensive and limits experimental utility since the nonlinear models can be sensitive to experimental errors. Also, the results are only demonstrated for moderate scattering conditions ($\epsilon_r \leq 3$), whereas typical objects around us have $2 < \epsilon_r < 77$ (at frequencies around 2.4 GHz), making it a useful contribution but not yet widely useful. [6], on the other hand, uses a linear model to generate initial reconstructions and demonstrates results for strong scattering, but is not formulated
for the lossy medium. Furthermore, both [2] and [6] need a large amount of data to train their DL networks to efficiently denoise the initial reconstructions. Also, both these techniques need a trial-and-error-based tuning of a critical hyperparameter to regularize the inversion of the ill-conditioned forward model and generate the initial reconstruction. This is not practical since this parameter also needs to be tuned during inference and the results depend heavily on it. The MADL approaches are also not highly interpretable since they typically use black-box DL networks to process the initial reconstruction.

In this work, we present an end-to-end unrolled optimization framework with deep priors which is equipped to tackle the severe nonlinearity and ill-posedness of PD-ISP and handle the aforementioned problems. The key contributions can be summarized as follows.

1) **Forward Model:** We use our recently proposed xPRA-LM [4] xPRA-LM as a phaseless linear forward model \( A_p \) in our framework. xPRA-LM is shown to have the highest validity range in handling strong scattering among all the PD-ISP models and can also handle experimental errors due to its linear formulation. It is also the only PD-ISP model which can be formulated to include background subtraction to remove multipath reflections (from clutter, walls, ceiling, floor) from the experimental data. We incorporate xPRA-LM into our proposed end-to-end reconstruction framework.

2) **End-to-End Unrolled Optimization Framework With Deep Priors** which:
   a) uses data-driven regularization in the form of DL-based priors for the xPRA-LM inverse problem while also learning the underlying nonlinearity which xPRA-LM cannot model;
   b) uses conventional multiparameter regularization in the objective in addition to the data-driven regularization to precondition xPRA-LM, thus tackling the ill-posedness of the xPRA-LM inverse problem; and
   c) learns all the optimum parameter values (including multiple regularization parameters and learning rate of the optimization algorithm) during training, removing the need to tune parameters manually.

We demonstrate the performance of the proposed framework using the simulation examples and experimental examples where measurement data are collected in a realistic indoor environment using Wi-Fi devices.

**D. Notation**

We use \( \mathbf{X} \) and \( \mathbf{x} \) to denote the matrix and vector form of discretized parameter \( X \), respectively. Lower case boldfaced letters represent the position vectors and italic letters represent the scalar parameters.

**II. Problem Formulation**

Consider a domain of interest (DoI) \( D \subset \mathbb{R}^2 \) situated in an indoor region of a building as shown in Fig. 1. It has dimensions \( d_x \times d_y \) m and has 2.4-GHz Wi-Fi transceiver nodes placed at its boundary (denoted as \( B \subset \mathbb{R}^2 \)). There are in total \( M \) transceiver nodes, each of which transmit and receive signals to acquire received signal strength (RSS) measurements of links between the nodes. These nodes cannot transmit and receive at the same time. Therefore, the total number of wireless links is equal to \( L = M(M - 1)/2 \) excluding the reciprocal links and self-measurement. The \( m \)th node is located at \( r_{m} \in B \), and \( l_{m,m'} \) denotes the link between the transmitting node \( m \) (at \( r_{m} \)) and the receiving node \( m' \) (at \( r_{m'} \)). For the remainder of this work, we use the subscripts \( m \) and \( m' \) to refer to the transmitter and the receiver, respectively, for all the relevant quantities.

We approximate the environment in Fig. 1 as a 2-D electromagnetic problem where we consider the DoI to be a planar cross section parallel to the floor at a height \( d_h \). This 2-D approximation is valid since we use directional antennas and background subtraction [4] to reduce the effect of scattering from the floor, ceiling, and other clutter outside the DoI. We discretize this 2-D DoI into \( N = n_x \times n_y \) rectangular cells, each of size \( \Delta d_x \times \Delta d_y \), where \( r_n \) denotes the location of the \( n \)th cell (see Fig. 1).

Let us consider the \( i \)th wireless link with the source at \( r_{m_i} \) and the receiver at \( r_{m_i} \) as shown in Fig. 1. The total field \( E_{m_i}(r) \) at the receiver \( r = r_{m_i} \) or inside any cell in the DoI \( r = r_n \) is

\[
E_{m_i}(r) = E_{m_i}^i(r) + E_{m_i}^s(r)
\]

where \( E_{m_i}^i(r) \) is the free-space incident field, and \( E_{m_i}^s(r) \) is the scattered field in the presence of scattering objects. The total field at the receiver \( E_{m_i}(r_{m_i}) \) can be expressed in an exact form in terms of the total field inside the DoI \( E_{m_i}(r_n) \) and the permittivity profile of the DoI \( \epsilon_r(r_n) \) using a volume source
integral (VSI) formulation as
\[ E_m(r_m) = E_m^i(r_m) + k_0^2 \int_D g(r_m, r_n) \chi(r_n) E(r_n) dr_n \]
where \( \chi(r_n) = \epsilon(r_n) - 1 \) is the permittivity contrast, and \( g \) is the homogeneous Green’s function. The VSI equation in (4) is also known as the Lipmann–Schwinger equation and provides an exact description of wave scattering [1], [14].

Solving VSI as an inverse problem implies solving it to estimate the contrast profile \( \chi(r_n) \) of the DoI given measurements of \( E \) (on \( B \)). This is known to be a nonlinear and ill-posed problem since there are two unknowns inside the integral \([\chi(r_n) \text{ and } E(r_n)]\), both of which need to be solved for a 2-D DoI with only 1-D measurements of the total field (along the measurement boundary \( B \)). Also, VSI in (4) needs both the phase and magnitude and is therefore an FD-ISP. There is no exact reformulation of VSI as a PD-ISP. Therefore, in the next section, we provide a brief description of an approximate linear PD-ISP model derived from FD-ISP in (4).

III. PROPOSED PHASELESS APPROXIMATE MODEL FOR MGDL

In this section, we approximate VSI (4) using our recently proposed extended phaseless RA (xPRA) model [4] derived by correcting the well-known RA and reformulate it as a linear approximate PD-ISP where the contrast is linearly related to RSS values.

In RA, instead of defining the total field as \( E = E_i + E_s \), it is defined as
\[ E_m(r) = E_m^i(r) e^{\phi_m(r)} \]
where the complex phase \( \phi_m(r) \) represents the phase and log amplitude deviations of the total field from the incident field as the field wavefront passes through the scattering media. However, RA is only valid for extremely weak scattering (\(|\epsilon| \approx 1\)) with limited practical applications. Our recent work [4] provides corrections to RA using a high-frequency approximation theory in lossy media to also handle strong scattering conditions. Combining this new corrected contrast with RA leads to the extended RA for lossy media (xRA-LM) [4] which can be written as
\[ E_m(r_m) = E_m^i(r_m) \]
\[ \cdot \exp\left( k_0^2 \int_D g(r_m, r_n) \chi_m(r_n) E_m^i(r_n) dr_n \right) \]
where \( \chi_m \) is the contrast given as
\[ \chi_m(r_n) = 2 \left( \sqrt{\epsilon_R(r_n)} \cos \theta_m^s(r_n) - 1 \right) \]
\[ + j \frac{\epsilon_I(r_n)}{\sqrt{\epsilon_R(r_n) - \sin^2 \theta_m^s(r_n)}} \cos \theta_m^i(r_n) \]
where \( \epsilon_R(r_n) \) and \( \epsilon_I(r_n) \), respectively, are the real and imaginary parts of relative permittivity \( \epsilon(r_n) = \epsilon_R(r_n) + j \epsilon_I(r_n) \). \( \theta_m^s(r_n) \) and \( \theta_m^i(r_n) \) are the scattering angle and the incident angle, respectively.¹ Unlike VSI, (6) is now a linear inverse problem where the unknown contrast \( \chi_m(r_n) \) is linearly related to measurements \( E_m(r_m) \). Note that \( \chi_m(r_n) \) is also a function of the location of the transmitting node due to its dependence on \( \theta_m^s(r_n) \) and \( \theta_m^i(r_n) \), where \( \theta_m^s(r_n) \) itself depends on \( \theta_m^i(r_n) \).

A. Phaseless Form

One of the biggest advantages of the exponential forms of FD-ISP in (6) is that it can be transformed into a phaseless form (PD-ISP). This can be done by multiplying (6) by its conjugate and taking logarithm on both sides to obtain \( \Delta P_m(r_m) \)
\[ \Delta P_m(r_m) = C_0 \cdot \Re \left( \frac{k_0^2}{E_m^i(r_m)} \int_D g(r_m, r_n) \chi_m(r_n) E_m^i(r_n) dr_n \right) \]
where the constant \( C_0 = 20 \log_{10} e \), and \( \Delta P(r_m) \) is the change in RSS values due to presence of the scatterer and can be written as
\[ \Delta P_m(r_m) [\text{dB}] = P_m(r_m) [\text{dB}] - P_m^0(r_m) [\text{dB}] \]
\[ = 20 \log_{10} \left( \frac{E_m(r_m)}{E_m^i(r_m)} \right) \]
where \( P_m \) and \( P_m^0 \) are the total received power and the free-space incident power, respectively. This means that to estimate the contrast profile \( \chi_m(r_n) \), we only need phaseless measurements of the total field. Therefore, unlike (4) and (6) which are FD-ISPs, (8) is a PD-ISP. Also, (8) is a linear inverse problem where the unknown \( \chi \) is linearly related to the change in RSS values \((P - P)\). We refer to this as the xPRA model in the remainder of the article. To the best of our knowledge, there are no other noniterative linear PD-ISP models.

Unlike exact models such as VSI where the contrast \( \chi \) is a linear function of permittivity \( \epsilon_r \), the new contrast \( \chi_m \) in (7) is now a nonlinear function of \( \epsilon_r \) and includes new distortion terms that depend on the scattering angle \( \theta_m^s \) and the incident angles \( \theta_m^i \). Therefore, even if we obtain the contrast profile by solving (8), it will contain distortions that are not trivial to remove. Since \( \theta_m^s(r_n) \) itself depends on both \( \theta_m^i(r_n) \) and \( \epsilon_r(r_n) \), it causes significant distortion in \( \Re(\chi_m(r_n)) \) and this distortion is different for different \( \epsilon_r \) values. On the other hand, distortion in \( \Im(\chi_m(r_n)) \) due to \( \theta_m^i(r_n) \) terms does not depend on \( \epsilon_r(r_n) \) since \( \theta_m^i(r_n) \) does not depend on \( \epsilon_r(r_n) \) and instead only depends on the transmitting node location. Hence, any distortions in \( \Im(\chi_m(r_n)) \) will be the same no matter what the permittivity of the object is. The effect of this is shown in our recent work [4] where the reconstruction of \( \Re(\chi_m(r_n)) \) using xPRA is highly distorted due to \( \theta_m^i(r_n) \) term. On the other hand, reconstruction of \( \Im(\chi_m(r_n)) \) using xPRA provides accurate shape reconstruction for very strong, low-loss scatterers and helps distinguish between objects with different permittivity values. However, even reconstructing

¹The incident angle is the angle made by incident wave at the cell \( r_s \) of the DoI and the scattering angle represents how much this incident wave gets scattered (refracted/reflectected) from its original path.
Im(χrn) using xPRA has several limitations. Distortion occurs in the reconstructions because θrn is not known or estimated. Also, xPRA cannot estimate Im(χrn) and distinguish between objects when the relative permittivity becomes very large (εn(rn) > 10). Finally, it cannot accurately image objects with nonconvex shapes.

The limitations of xPRA are a result of using a linear approximation and the distortions due to the θrn terms. In this work, we aim to tackle these limitations of xPRA to enhance its validity. The idea is to combine xPRA with a DL framework that learns to reconstruct the spatial structure (such as nonconvex shapes) and large contrast amplitude values of scatterers which cannot be reconstructed using xPRA alone, while also removing distortions in the reconstructions due to the θrn and δrn terms.

Our previous results [4] showed that Im(χrn) is closely related to the reconstruction of the ratio εr(rn)/εεg(rn)1/2 = δ(rn)(εεg(rn))1/2 where δ(rn) is the loss tangent. While it cannot separately estimate εεg and εr profiles of the DoI, the estimation of the δ(rn)1/2 profile provides an estimate of both the scattering ability (εεg) and the absorbing ability of the object under test. It can be thought of as a type of extinction coefficient. It is sufficient to obtain accurate shape reconstructions of extremely strong scatterers (even of nonconvex shapes) and can also efficiently distinguish between objects of different permittivity values.

For the reasons described above, we propose to learn the δ(rn)(εεg(rn))1/2 ratio from Im(χrn)) and Re(χrn)). We show later in the results section that our technique can provide extremely accurate reconstructions of δ(rn)(εεg(rn))1/2.

B. Imaging Model in Matrix-Vector Form

Since the DoI is divided into N = NtNc cells, we can replace the integral in the aforementioned equations with summation. Also, the xPRA model in (8) is for a single wireless link. We can stack it for all L = M(M − 1)/2 measurement links to obtain a linear system of equations

\[ \Delta P \approx \text{Re}(\overline{G}^T \overline{T}) \]

where the contrast vector \( \overline{T} \in \mathbb{C}^{N \times 1} \) contains the elements \( T(r_n) \) for all N cells inside the DoI \( (n = 1, 2, \ldots, N) \). Note that \( T(r_n) \) in (8) is a function of a single transmitting node \( m \), whereas the vector \( \overline{T} \) in (10) is a single \( N \times 1 \) vector obtained using L measurements due to illumination of DoI by all the transmitting nodes in the measurement setup. Therefore, \( \overline{T} \) can be considered to be a weighted linear combination of \( T(r_n) \) vectors for all the measurement links and should only be a function of permittivity. It is from this term \( \overline{T} \) that we learn δ(rn)(εεg(rn))1/2. We do this using data-driven regularization modeled as deep priors as shown later. Each element of the measurement vector \( \Delta P \in \mathbb{R}^{L \times 1} \) is

\[ p_l = \Delta P(r_m) = P(r_m) - P(r_m) \] (in dB).

The position vector \( r_m \) gives the location of the receiver in the lth wireless link where the transmitter is at \( r_m \). The xPRA kernel matrix \( \overline{G} \in \mathbb{C}^{L \times N} \) contains entries in \( \overline{G}_{l,n} \) given as

\[ \overline{G}_{l,n} = \frac{C_0 k_0^2}{E^l(r_m) E^l(r_n)} \sum_{v} g(r_m, r_n) \Delta a \]

where \( \Delta a \) is the area of each cell.

We can further simplify (10) by manipulating the operator Re to rewrite (10) as

\[ \Delta P = \left[ \text{Re}(\overline{G}) - \text{Im}(\overline{G}) \right] \left[ \text{Re}(\overline{T}) \text{Im}(\overline{T}) \right] \]

We can define our inverse problem as estimating Re(\( \overline{T} \)) and Im(\( \overline{T} \)) given the measurements \( \Delta P \). Therefore, the model-based objective function (which will be used in the next section) can be defined as

\[ \min_{\overline{T}} \left\| \Delta P - \overline{G} \left[ \begin{array}{c} \text{Re}(\overline{T}) \\ \text{Im}(\overline{T}) \end{array} \right] \right\|_2^2 \]

where

\[ \overline{G} = \left[ \begin{array}{c} \text{Re}(\overline{G}) - \text{Im}(\overline{G}) \end{array} \right] \in \mathbb{R}^{L \times 2N} \]

\[ \overline{T} = \left[ \begin{array}{c} \text{Re}(\overline{T}) \\ \text{Im}(\overline{T}) \end{array} \right] \in \mathbb{R}^{N \times 1} \]

The matrix \( \overline{G} \in \mathbb{R}^{L \times 2N} \) is an approximation to the underlying nonlinear forward model.

As explained in Section III-A, Im(\( \overline{T} \)) is more relevant than Re(\( \overline{T} \)) for reconstructing the \( \delta(\epsilon_{\text{r}})^{1/2} \) profile since Re(\( \overline{T} \)) is distorted due to the presence of the unknown \( \delta_{\epsilon_{\text{r}}} \) terms. In the next section, we use our MGDL framework to reconstruct the \( \delta(\epsilon_{\text{r}})^{1/2} \) profile of the DoI using both Im(\( \overline{T} \)) and Re(\( \overline{T} \)) so that all useful information is exploited. More specifically, since Im(\( \overline{T} \)) provides approximate reconstructions of the \( \delta(\epsilon_{\text{r}})^{1/2} \) profile, we use Im(\( \overline{T} \)) as the primary input and Re(\( \overline{T} \)) as the secondary input in our MGDL framework.

IV. END-TO-END RECONSTRUCTION FRAMEWORK

In this section, we propose an MGDL framework to achieve better estimates of shape and electrical properties \( \delta(\epsilon_{\text{r}})^{1/2} \) of objects inside the DoI by overcoming the following key drawbacks of the xPRA model.

1) Errors due to inversion of the severely ill-posed xPRA inverse problem, and the need to tune regularization parameters during inference.
2) Errors caused by the model itself, since it is a linear approximation to a nonlinear problem.
3) Distortions due to the presence of \( \delta_{\text{r}} \) and \( \delta_{\text{I}} \) terms in the contrast.

We show that overcoming these drawbacks can allow us to image objects with very high permittivity, nonconvex shapes, and also improve the accuracy of overall shape reconstruction without the need for hyperparameter tuning.

A. Objective

The goal of our proposed framework is to incorporate the state-of-the-art xPRA model into the DL framework to mitigate the aforementioned drawbacks. We do this by combining the xPRA inverse problem objective with the traditional and DL-based priors and solving the resulting optimization problem using an unrolled iterative algorithm.

We start with the model-based objective function (24) and reformulate it as a variational regularization problem with
multiple penalties on the reconstruction as

\[
\min_{\mathbf{x}, \mathbf{r}} \left\| \Delta \mathbf{P} - \mathbf{Q} \mathbf{R} \mathbf{L} \right\|_2^2 + \lambda_1 \left\| \mathbf{Q}_1 \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \right\|_2^2 + \lambda_2 \left\| \mathbf{Q}_2 \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \right\|_2^2 + R \left( \frac{\mathbf{R}}{\mathbf{L}} \right)
\]  

(15)

where \( \mathbf{Q}_1 \) and \( \mathbf{Q}_2 \) represent the Tikhonov prior matrices of the classical multiparameter regularization, and \( R \) represents the data-driven regularizer which we model as a deep prior. The need for each of these priors is explained in Sections IV-B and IV-D that follow.

B. Preconditioning With Multiparameter Tikhonov Regularization

Unlike inverse problems in other domains such as medical imaging, PD-ISPs are highly ill-posed due to the scarcity of measurements, missing phase information, and the highly nonlinear underlying model. This can lead to large errors in the inversion of the poorly conditioned matrix \( \mathbf{P} \in \mathbb{R}^{L \times N} \), where \( L \ll N \). Hence, a least-squares solution provides extremely distorted reconstructions that are highly sensitive to the changes in the measurements. Therefore, we add classical Tikhonov priors to precondition this highly ill-posed problem and stabilize the inversion.

For our proposed framework, we add two Tikhonov priors to our PD-ISP inverse problem as shown in (15). \( \mathbf{Q}_1 \) here represents the identity matrix and the resulting term minimizes the \( L_2 \)-norm of the coefficients in \( \left[ \frac{\mathbf{R}}{\mathbf{L}} \right] \). \( \mathbf{Q}_2 \) represents the difference matrix approximating the derivative operator, thus minimizing the \( L_2 \)-norm of the difference between successive coefficient terms. This is also called the \( H \) regularization. Since we obtain reconstructions of a 2-D DoI, we apply the derivative operator in both the dimensions.

The problem in (15) can now be written as

\[
\min_{\mathbf{x}, \mathbf{r}} \left\| \Delta \mathbf{P} - \mathbf{Q} \mathbf{R} \mathbf{L} \right\|_2^2 + \lambda_1 \left\| \mathbf{Q}_1 \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \right\|_2^2 + \lambda_2 \left\| \mathbf{Q}_2 \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \right\|_2^2 + R \left( \frac{\mathbf{R}}{\mathbf{L}} \right)
\]  

(16)

where \( \mathbf{D}_x \in \mathbb{R}^{2N \times 2N} \) and \( \mathbf{D}_y \in \mathbb{R}^{2N \times 2N} \) are the approximate differentiation operators in the horizontal and vertical directions, respectively.

The Tikhonov prior terms in (16) enforce both sparsity and smoothness in the solutions obtained, thus not only estimating important pixels in the solutions but also capturing the spatial continuity in the object of interest. This is especially desirable when we consider large piecewise homogeneous objects, which are commonly found in indoor regions. This idea of the prior enforcing both sparsity and smoothness is inspired by Fused Lasso [15] which is combination of LASSO and TV regularization where an \( L_1 \)-norm penalty is used on both the solution coefficients and their successive differences. However, we restrict ourselves to the use of Tikhonov priors due to their convexity and smoothness properties which can be leveraged to obtain a better convergence rate for algorithms used to solve the optimization problem. Also, the existence of analytical solutions for these priors is useful for initializing the framework, as will be explained in Section IV-D.

C. Optimization Framework

We can rewrite (16) in a compact form as

\[
\min_{\mathbf{x}, \mathbf{r}} f \left( \frac{\mathbf{R}}{\mathbf{L}} \right) + R \left( \frac{\mathbf{R}}{\mathbf{L}} \right)
\]  

(17)

where \( f \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \) is a smooth and differentiable function consisting of the datafit and Tikhonov prior terms, and \( R \left( \frac{\mathbf{R}}{\mathbf{L}} \right) \) is the deep prior term modeled using convolutional neural network (CNNs) and is therefore nonconvex and nondifferentiable. This results in (17) being a nondifferentiable objective function, making gradient-based methods such as gradient descent or Newton method unsuitable. A better approach is to use proximal algorithms such as proximal gradient method (PGM) [16] or alternating direction method of multipliers (ADMM) [17] designed for a nonsmooth objective where a proximal step replaces the gradient step for the nonsmooth terms. For our proposed framework, we use the PGM since it is usually considered to be a good default choice. However, the use of other proximal algorithms can also be explored.

A single step of the PGM for the objective in (17) can be given as

\[
\mathbf{x}^i = \mathbf{x}^{i-1} - \eta^i \nabla f \left( \mathbf{x}^{i-1} \right)
\]

\[
\mathbf{x}^i = \text{prox}_{\eta^i R}^{f} \left( \mathbf{z}^i \right)
\]

where \( \text{prox}_{\eta^i R}^{f} \left( \mathbf{z}^i \right) = \arg \min_{\mathbf{x}} \left( R(\mathbf{x}) + \frac{1}{2\eta^i} \| \mathbf{x} - \mathbf{z}^i \|_2^2 \right) \)  

(18)

where \( \eta^i \) is the step size at the \( i \)-th iteration of PGM, \( \mathbf{x}^{i-1} = \left[ \frac{\mathbf{R}}{\mathbf{L}} \right]^{-1} \left[ \frac{\mathbf{R}}{\mathbf{L}} \right]^{-1} \) is the input at the \( i \)-th iteration, \( \mathbf{z}^i \) is the gradient step on the smooth function \( f \), and \( \mathbf{x}^i \) is the output of the proximal operator of \( R \) applied to \( \mathbf{z}^i \). Here, since \( R \) represents the data-driven regularizer in our framework, we directly parameterize the proximal operator of \( R \) in the form of a deep CNN, defining the actual regularizer \( R \) implicitly in the process.

It is important to note again that our goal is to estimate the \( \delta(r_\eta)^{1/2} \) profile of the DoI, and as mentioned in the previous section, this reconstruction ratio is more closely related to \( \mathbf{T} \) than to \( \frac{\mathbf{R}}{\mathbf{L}} \). Therefore, we consider \( \mathbf{T} \) to be the primary variable of our optimization framework from which we can learn the \( \delta(r_\eta)(\epsilon_\eta(r_\eta))^{1/2} \) values of objects, while also using \( \frac{\mathbf{R}}{\mathbf{L}} \) as an aiding variable to improve it since it provides a good reconstruction of the scatterer boundaries [4].
We now incorporate this discussion mathematically into the optimization algorithm in (18). We first obtain an initial guess \( \{\mathbf{Z}_R, \mathbf{Z}_I\}^0 \) from the measurements \( \Delta \mathbf{F} \) by analytically solving an objective containing only the datafit and Tikhonov prior terms to obtain

\[
\begin{align*}
\mathbf{Z}_R^0 &= \mathbf{G}^T \mathbf{G} + \lambda_1^0 \mathbf{I} + \lambda_2^0 \left( \mathbf{D}_x^T \mathbf{D}_x + \mathbf{D}_y^T \mathbf{D}_y \right)^{-1} \mathbf{G}^T \Delta \mathbf{F} \\
\mathbf{Z}_I^0 &= \Pi_{\mathbf{Z}_R} \Delta \mathbf{F}
\end{align*}
\]

where \( \Pi \) is the regularized pseudoinverse parameterized by two hyperparameters \( \lambda_1^0 \) and \( \lambda_2^0 \), controlling sparsity and smoothness in the initial reconstructions, respectively. The values of \( \lambda_1^0 \) and \( \lambda_2^0 \) are learned during training similar to the traditional counterparts \([18]\).

With this initialization, we formulate the PGM algorithm in a way that we can optimize \( \mathbf{Z}_I \) to bring it close to ground-truth \( \delta(r_n)(\epsilon(r_n))^{1/2} \) values, while also using information from the aiding variable \( \mathbf{Z}_R \). To this end, we write the \( i \)th iteration of the PGM algorithm as

\[
\begin{align*}
\mathbf{Z}_R^i &= \mathbf{Z}_R^{i-1} - \eta^i \nabla f \left( \mathbf{Z}_R^{i-1} \right) \\
\mathbf{Z}_I^i &= \text{prox}_{\eta R} \left( \mathbf{Z}_R^i, \mathbf{Z}_I^i \right)
\end{align*}
\]

where the gradient of \( f \) is obtained from (16) as

\[
\nabla f \left( \mathbf{Z}_R^{i-1} \right) = \mathbf{G}^T \left( \mathbf{G} \mathbf{Z}_R^{i-1} - \Delta \mathbf{F} \right) + \left( \lambda_1^i I + \lambda_2^i \left( \mathbf{D}_x^T \mathbf{D}_x + \mathbf{D}_y^T \mathbf{D}_y \right) \right) \mathbf{Z}_R^{i-1}.
\]

Here, \( \lambda_1^i \) and \( \lambda_2^i \) are the regularization parameters for each iteration of PGM and are learned during training similar to \( \lambda_1^0 \) and \( \lambda_2^0 \). Note that \( \lambda_1^0 \) and \( \lambda_2^0 \) can be different for different iterations depending on the residual error between the ground-truth \( \delta(\epsilon)^{1/2} \) profile and the input \( \mathbf{Z}_R \) for that iteration. Having this flexibility in regularization parameter values for different iterations is often ignored, but it is crucial since as the output improves with iterations, the regularization parameter needs to change.

\( \mathbf{Z}_R \) and \( \mathbf{Z}_I \) in (20a) are the outputs of the gradient update step corresponding to \( \mathbf{Z}_R^{i-1} \) and \( \mathbf{Z}_I^{i-1} \), respectively. Both \( \mathbf{Z}_R \) and \( \mathbf{Z}_I \) are given as inputs to the \text{prox} operator modeled as a CNN to learn the ground-truth \( \delta(\epsilon)^{1/2} \) profile from \( \mathbf{Z}_I \) while also using the boundary information contained in \( \mathbf{Z}_R \). However, it is important to note that CNN does not operate on \( \mathbf{Z}_R \); it is directly passed to the next iteration after the gradient update step as \( \mathbf{Z}_R \) and we only receive \( \mathbf{Z}_I \) as the CNN output. This is akin to solving the optimization problem

\[
\min_{\mathbf{Z}_I} \text{prox}_{\eta R} \left( \mathbf{Z}_R^i, \mathbf{Z}_I^i \right) - \eta^i \nabla f \left( \mathbf{Z}_R^i \right)
\]

The output of the CNN is thus an improved \( \mathbf{Z}_I \) that is closer to the ground-truth \( \delta(\epsilon)^{1/2} \) profile. As we increase the number of iterations, the output of the \text{prox} operator becomes closer and closer to the ground-truth \( \delta(\epsilon(\epsilon_n))^{1/2} \) values, eventually giving us accurate reconstructions of the \( \delta(\epsilon)^{1/2} \) profile in the last iteration.

D. End-to-End Unrolled Optimization Algorithm

To construct an end-to-end reconstruction framework, we first choose a fixed number of iterations \( T \) for (20) and unroll the resulting iterative algorithm into a network, with each layer representing one iteration \([18]\). This is illustrated in Fig. 2. Each layer of this unrolled network consists of a gradient step that is a function of the measurement vector \( \Delta \mathbf{F} \), the xPRA matrix \( \mathbf{G} \), and the Tikhonov prior terms, followed by a \text{prox} operator of the data-driven regularizer parameterized by a CNN. Passing through this network is equivalent to executing the iterative algorithm in (20) \( T \) times. Also, algorithm parameters such as the step size \( \eta^i \), the Tikhonov regularization parameters \( (\lambda_1^i, \lambda_2^i) \), and the deep prior network parameters all transfer to the unrolled network parameters. This unrolled network can then be trained to learn the optimal values of these parameters using training data. Thus, the trained network can be interpreted as a parameter optimized iterative algorithm.

Another advantage of unrolled iterative algorithms is that they converge in much fewer iterations when compared with their traditional counterparts \([18]\).

The final MGDL framework is described in Fig. 2. The key components of the framework are described below.

1) Initialization Layer: The network obtained by unrolling the iterative algorithm in (20) is preceded by an initialization layer which takes as input the measurement vector \( \Delta \mathbf{F} \) and generates an initial reconstruction \( \{\mathbf{Z}_R^0, \mathbf{Z}_I^0\}^T \) using (19).
regularized solution lets us initialize the unrolled network with a noisy but stable reconstruction as opposed to initializing it with random values or the least-squares solution of a highly ill-posed system, thus leading to faster convergence of the unrolled network. We add a rectified linear unit (ReLU) activation function at the end of this layer to prevent negative values, since they are not physically possible. Also, the regularization parameters \( \lambda_0 \) and \( \lambda_1 \) do not need to be tuned manually and are learned from data.

2) Deep Prior: For the proximal operator of \( R \) parameterized by a deep network, we use the U-Net architecture [19]. An important feature of this architecture is its encoder–decoder structure that leads to a bottleneck in the network, which helps it learn structure in data that cannot be captured by the traditional CNNs [6], [20].

The U-Net prior in each optimization layer contains four encoder layers and four decoder layers, each of which contain 64 \( 3 \times 3 \) convolutional filters followed by batch normalization and ReLU activation. It also contains skip connections from the encoder layers to the decoder layers which lead to faster training of the network. Specifically, we add skip connections between each encoder layer \( i \) and decoder layer \( n-i \) where \( n \) is the total number of layers in the encoder or the decoder. Each skip connection concatenates all the channels at the output of the decoder layer \( n-i \) with those at the output of the encoder layer \( i \).

In each optimization layer (described in (20) and in Fig. 2), the outputs \( \mathbf{Z}_R \in \mathbb{R}^{N \times 1} \) and \( \mathbf{Z}_T \in \mathbb{R}^{N \times 1} \) of the gradient update step in (20a) are converted into 2-D images of dimensions \( n_x \times n_y \) to map them to the 2-D DoI in Fig. 1. These 2-D images \( \mathbf{Z}_R, \mathbf{Z}_T \in \mathbb{R}^{n_x \times n_y} \) are then given as inputs to the U-Net as two channels. The output of the U-Net is a 2-D image representing \( \mathbf{Z}_I \in \mathbb{R}^{n_x \times n_y} \). This is unrolled into a vector \( \mathbf{Z}_I \in \mathbb{R}^{N \times 1} \), concatenated to \( \mathbf{Z}_R \), and propagated to the next layer to repeat the same process.

In addition to using the conventional regularization in the form of Tikhonov priors that enforce sparsity and smoothness, using data-driven regularization in the form of a deep prior such as U-Net in each layer is useful for learning prior information such as shapes and contrast values of the reconstruction while also learning to remove the distortion caused in \( \mathbf{Z}_I \) by the \( \theta^i \) term. The deep prior also extracts useful information from \( \mathbf{Z}_R \) to aid the reconstruction. As we move through the optimization layers, \( \mathbf{Z}_I \) becomes closer and closer to the ground-truth profile, and the U-Net prior in the final optimization layer provides reconstruction of the \( \delta(\epsilon_R) \) profile of the DoI as the output.

3) Learning Hyperparameters and Network Parameters: All the hyperparameters involved in the optimization process including the regularization parameters in the initialization layer \( \lambda_0 \) and \( \lambda_1 \) (in (19)) and the regularization parameters \( \lambda_0, \lambda_1 \) and the learning rates \( \eta^i \) in the \( i \)-th optimization layer in (20) are learned during the training of the unrolled network. Also, as mentioned before, \( \lambda_0 \) and \( \lambda_1 \) can vary across layers depending on the residual error between the ground-truth \( \delta(\epsilon_R) \) profile and \( \mathbf{Z}_I \). This leads to different intensities of regularization being applied at different iterations/layers, and as the reconstruction gets better in later layers, the regularization parameter can adapt to the residual noise and distortion. Also, even though the same U-Net architecture is used as the deep prior in all the optimization layers, we parameterize them as separate CNNs. This offers tremendous flexibility by allowing the prior to learn a specialized function at every layer based on the noise and distortions in reconstructions in that layer.

To build our end-to-end reconstruction framework, we unroll the PGM algorithm for \( T = 5 \) iterations and precede it by the initialization layer. This complete network is then trained using a mean squared error loss and Adam optimizer. The key issue in training this framework is that the scale at which regularization parameters vary can be very different from the scale at which other network parameters such as the U-Net weights change. Regularization parameters for a layer can typically vary on a logarithmic scale and take different values depending on distortions in reconstructions in that layer. Therefore, we cannot use an optimizer with the same learning rate for both these types of parameters. To tackle this, we use a learning rate of \( 10^{-2} \) for the regularization parameters in all the layers \( (\lambda_0, \lambda_1, \lambda_2) \) and \( 10^{-4} \) for all other parameters throughout the network. The higher learning rate of regularization parameters offers greater flexibility in tuning these parameter values correctly since very small changes to these do not have a significant effect on the reconstructions.

V. SIMULATION RESULTS

This section provides simulation examples to evaluate the performance of our proposed framework. The setup used for the simulations is compatible with that used for the experiments (experimental results are presented in the next section). In the simulations, the values of the size and relative permittivity \( \varepsilon_r = \varepsilon_R + j\varepsilon_i \) of scatterers are selected such that these represent objects found in a typical indoor environment.

The peak signal-to-noise ratio (PSNR) results for each reconstruction are also provided for quantitative evaluation.

A. Simulation Setup

We consider a \( 3 \times 3 \) m indoor region for both our simulation and experiments. It contains a \( 1.5 \times 1.5 \) m DoI to be imaged as shown in Fig. 3(a). We simulate the transceiver nodes operating at 2.4 GHz with incident wavelength \( \lambda_0 = 12.5 \) cm (in the experiments these are Wi-Fi transceivers). The size of the DoI is \( 12 \lambda_0 \times 12 \lambda_0 \) and there are 40 equidistant transceivers \( (M = 40) \) placed at the DoI boundary to collect measurements. Therefore, the total number of wireless links is \( L = M(M-1)/2 = 780 \). We discretize the DoI into \( 160000 \) cells \((400 \times 400)\) for the forward problem and \( N = 2500 \) cells \((50 \times 50)\) for the inverse problem. To generate the forward measurement data (RSS values) for training, we use the method of moments approach detailed in [1].

B. Training Data

For training the network, we generate a dataset containing circular- and square-shaped objects placed in the DoI (see...
that the origin lies at the center of the DoI, the model used as a deep prior is translationally invariant, we only sample object profiles in Fig. 4). A single object can be a circle or a square with equal probability. Since the U-Net model used as a deep prior is translationally invariant, we only place objects in the upper half of the DoI. Therefore, assuming that the origin lies at the center of the DoI, the x and y coordinates of the object centers are sampled uniformly from 

\[ c_x = [-0.6, 0.6] \]

and

\[ c_y = [0.15, 0.6] \]

respectively.

Since we are demonstrating the effectiveness of our framework for indoor imaging, we select the real component of each object. This forms the ground-truth image which is used for training, 150 samples for validation, and the remaining 500 samples for testing the trained framework for generalization to samples not present in the training data. We call this test data the in-sample test data, since the objects in these samples have the same parameters as the ones used in training.

To better test the proposed framework, we also generate an out-of-sample test data in which object parameters differ from the ones mentioned in this section, i.e., objects have shapes other than circles and squares and have permittivity values different from the ones used in training. The performance of our proposed framework on this out-of-sample dataset is also demonstrated in this section.

For brevity, from here on, we refer to our proposed MGDL framework using the xPRA model, Tikhonov priors, and the deep prior as xPRA-TK-DPrior.

C. Reconstruction Results

To demonstrate the comparative performance of our proposed xPRA-TK-DPrior framework, we compare it with other state-of-the-art end-to-end frameworks including the unrolled ADMM TV regularized approach, the DI approach, and the existing deep prior approach. Specifically, we compare the proposed framework to the following:

1) xPRA Model With TV Regularization (xPRA-TV):

To analyze the effect of the deep prior on the reconstructions obtained, we remove it and instead solve the optimization problem in (24) with a widely used TV prior as

\[
\min_{\mathbf{Z}_I} \left[ \Delta \mathbf{F} - \mathbf{F}_I \right] \leq \lambda \left[ \mathbf{F}_I \right]_1^2 + \lambda_{TV} \left( \left\| \mathbf{D}_x \mathbf{Z}_I \right\|_1 + \left\| \mathbf{D}_y \mathbf{Z}_I \right\|_1 \right) \quad (23)
\]

where \( \mathbf{D}_x \in \mathbb{R}^{2N \times 2N} \) and \( \mathbf{D}_y \in \mathbb{R}^{2N \times 2N} \) are the derivative operators in the horizontal and vertical directions, respectively. We solve (23) using the ADMM algorithm which we unroll for \( T = 5 \) iterations and learn the value of \( \lambda_{TV} \) while training the unrolled network, similar to the regularization parameters learned in our framework.

2) DI: For the sake of providing a complete picture, we also observe the effect of including the xPRA forward model in our framework. A framework that does not use the xPRA model only consists of the initialization layer followed by a cascade of \( T \) deep priors. This is an example of the DI technique discussed in Section I. In this case, the initialization layer consists of a fully connected layer that directly translates measurements \( \Delta \mathbf{F} \in \mathbb{R}^{780 \times 1} \) to the reconstruction domain \( [\mathbf{F}_R^T \mathbf{F}_I^T]^T \in \mathbb{R}^{5000 \times 1} \).

3) xPRA Model With Only Deep Prior (xPRA-DPrior):

To understand how much of the reconstruction performance can be attributed to preconditioning the highly ill-conditioned xPRA model matrix using multiparameter regularization, we also initialize and train the framework by omitting the Tikhonov prior terms. The resulting framework is an xPRA-based
Fig. 4. Reconstruction results for in-sample test data where the reconstructed $\delta(\mathbf{r}_n)(\epsilon_{\mathbf{r}_n})^{1/2}$ values are indicated by color and the exact permittivities of the ground truth objects are listed in (a)–(e). The first column shows the ground truth, and the columns from second to last show reconstructions obtained using xPRA-TV, DI, xPRA-DPrior, and xPRA-TK-DPrior, respectively. Dimensions are shown on the $x$- and $y$-axis of the ground-truth results in the first column and are in units of meters. Also note that the color scales in the second, third, and fourth columns of images are different from those in the first and final columns.

In terms of performance evaluation of xPRA model, our recent work [4] is used as a reference since it significantly outperforms the state-of-the-art approach known as subspace-based optimization method using phaseless data (PD-SOM) [1], [2], [3]. Therefore, we do not provide comparison with PD-SOM in this work.

We train all these frameworks (xPRA-TV, DI, xPRA-DPrior, and xPRA-TK-DPrior) for 20 epochs on a training data of 1500 samples as explained in Section V-B.

Fig. 4 shows the results for samples contained in the in-sample test data. The first column shows the ground-truth $\delta(\mathbf{r}_n)(\epsilon_{\mathbf{r}_n})^{1/2}$ profiles and the other columns show reconstructions using xPRA-TV, DI, xPRA-DPrior, and xPRA-TK-DPrior frameworks, respectively. The complex-valued relative deep prior technique that is similar to the state-of-the-art deep prior-based methods used to solve inverse problems in the field of medical imaging and image enhancement [11]. We call this xPRA-DPrior. However, as we have mentioned before, since PD-ISPs are much more ill-posed when compared to problems in these domains, preconditioning the model matrix becomes important. We can see the importance of this preconditioning for PD-ISPs by comparing this framework with our proposed framework that contains additional Tikhonov priors.

The summary and average PSNR values for all these frameworks over the in-sample test data are provided in Table I.
permittivity values for the samples are listed in the respective subfigure captions. These values of relative permittivity are considered extremely high in the inverse scattering community and can create extremely strong scattering, especially since the object sizes are comparable to or larger than \( \lambda_0 \).

It can be seen that for all the five test profiles shown in Fig. 4, our proposed xPRA-TK-DPrior framework significantly outperforms all other frameworks. The average PSNR values of reconstructions on the in-sample test data for all these frameworks are listed in Table I.

The DI technique performs the worst as shown in Fig. 4 and Table I. This is because it contains no model information to aid reconstructions and therefore needs to learn the nonlinear relationship between the measurements \( \Delta \mathbf{F} \) and the scatterer profile entirely from data. Doing so is a highly data-intensive task for which a dataset of 1500 samples is insufficient. Therefore, the DI framework is only able to localize objects to the upper half of the DoI, but it cannot localize them further without additional training data. Fig. 4 shows that unlike the DI approach, the xPRA-TV method is able to correctly estimate the location of objects. But it cannot accurately estimate their shape and reconstruction amplitude. This is because TV regularization in xPRA-TV makes it a fully linear method. Therefore, it is not able to correct errors caused by the linear approximation (xPRA) of the underlying nonlinear forward model. The xPRA-DPrior framework that contains the xPRA model along with deep priors provides much better results when compared with both the fully linear xPRA-TV and the highly nonlinear DI frameworks. It does so on account of the framework using the xPRA model to aid reconstructions and the deep prior handling the ill-posedness and learning the intricate spatial structures caused by the underlying nonlinearity in the forward model. However, as can be seen in the fourth column of Fig. 4, even this technique does not provide highly accurate reconstructions of shape and \( \delta(r_n)(\varepsilon_F(r_n))^{1/2} \) values of objects. It is also important to note that in the xPRA-DPrior framework, we use a U-Net architecture with skip connections, similar to our proposed xPRA-TK-DPrior framework. However the original work related to deep priors [11] uses a much simpler CNN architecture. Therefore, the xPRA-DPrior results in Fig. 4 are better than the existing deep-prior-based methods due to our choice of the U-Net architecture and xPRA model.

The last column in Fig. 4 shows the reconstructions using our proposed xPRA-TK-DPrior framework. Similar to xPRA-DPrior, our framework uses the xPRA model to aid reconstructions and deep priors to learn prior information about reconstructions and correct errors in xPRA due to its linear nature. However, unlike xPRA-DPrior in which the highly ill-conditioned and approximate xPRA model causes high distortions in the initialization and subsequent iterations, preconditioning it using the multiparameter regularization stabilizes the solution obtained and results in highly accurate reconstructions of the shape and \( \delta(r_n)(\varepsilon_F(r_n))^{1/2} \) values of all the objects.

We can also see the superior performance of our proposed framework in Fig. 5 where it converges to a higher average PSNR value on the out-of-sample test data, performing much better than all other frameworks in comparison.

### D. Generalization Tests

The results shown in Fig. 4 are for the in-sample test data. To test the generalization of our proposed framework, we also evaluate it on the out-of-sample test data in which the sample properties differ from those used in training. The results of some of the generalization tests are shown in Fig. 6. Fig. 6(a) and (b) contains a single object having shapes different from squares or circles, whereas the proposed framework is only trained on profiles that always consist of circular- or square-shaped scatterers. It can be seen that our proposed framework provides accurate shape and \( \delta(r_n)(\varepsilon_F(r_n))^{1/2} \) value reconstructions for such samples, showing the high generalization ability of our proposed framework in terms of reconstructing objects with different shapes.

Fig. 7 shows generalization tests for \( \varepsilon_r \) values different from the ones used in training. Fig. 7(a) and (b) shows results for a single object that has \( \varepsilon_r = 30 + 3j \) and \( 40 + 4j \), respectively. These permittivity values are not present in our training dataset. It can be seen that our proposed framework reconstructs the \( \delta(\varepsilon_F)^{1/2} \) profile for these samples with high accuracy. Therefore, our proposed framework can generalize to both object shapes and reconstruction amplitudes not present in the training data. However, the generalization to other reconstruction amplitudes holds as long as \( \delta = 0.1 \) as in the training.
electromagnetic problem. The DoI is a 1.5 × 1.5 m area that lies in a 3 × 3 m 2-D planar cross section parallel to the floor. The 1.5 × 1.5 m DoI is discretized into 2500 cells (50 × 50) for the inverse problem. Similar to the setup described in Section V-A, the experiment also uses \( M = 40 \) Wi-Fi transceiver nodes which are placed at the edge of the 2-D cross section at a height of 1.2 m above the floor, thus creating a total of \( L = M(M - 1)/2 = 780 \) measurement links. Each node consists of a SparkFun ESP32 Thing board with an integrated Wi-Fi transceiver operating at 2.4 GHz, with the inbuilt omni-directional antennas of the SparkFun ESP32 boards replaced by directional Yagi antennas of 6.6 dBi to minimize the multipath reflections from the floor, ceiling and other clutter placed outside the DoI. More information about the setup can be found in [24].

B. Background Subtraction

The DoI for the experimental setup as shown in Fig. 3(b) is a 2-D cross section of a 3-D region. Therefore, data collected for this DoI can contain multipath reflections due to the floor, ceiling, and other clutter around the DoI. In addition to using directional antennas to minimize these multipath reflections, we also incorporate background subtraction into xPRA.

At some initial time instant \( t_0 \) when the object to be imaged is not present in the DoI, the contrast profile of the DoI is \( \mathbf{I}^0 \), and the total received power vector is given as \( \mathbf{P}^0 \). This total power measured contains reflections from the clutter inside and outside the DoI. We can write the xPRA model for this time instant as

\[
\mathbf{P}^0 - \mathbf{P} = \mathbf{y} \left[ \frac{\mathbf{I}_R^0}{\mathbf{I}^0} \right] (24)
\]

where \( \mathbf{P} \) is the free-space incident power. Objects are placed inside the DoI at the instant \( t_0 + \Delta t_i \), and the contrast and total measured power at this instant are given as \( \mathbf{I}_R^{\Delta t_i} \) and \( \mathbf{P}^{\Delta t_i} \), respectively. For this instant, we can write the xPRA model as

\[
\mathbf{P}^{t_0 + \Delta t_i} - \mathbf{P} = \mathbf{y} \left[ \frac{\mathbf{I}_R^{t_0 + \Delta t_i}}{\mathbf{I}^{t_0 + \Delta t_i}} \right] (25)
\]

Note that the multipath reflections due to clutter are present even at the time instant \( t_0 + \Delta t_i \), but now there is additional information in the total power vector due to the presence of the objects to be imaged in the DoI. We can then subtract the xPRA equation for the time instant \( t_0 \) from that of the instant \( t_0 + \Delta t_i \) to obtain the background subtraction form as

\[
\Delta \mathbf{P}^{\Delta t_i} = \mathbf{y} \left[ \frac{\Delta \mathbf{I}_R^{\Delta t_i}}{\Delta \mathbf{I}^{\Delta t_i}} \right] (26)
\]

where \( \Delta \mathbf{P}^{\Delta t_i} = \mathbf{P}^{t_0 + \Delta t_i} - \mathbf{P}^0 \) and \( \Delta \mathbf{I} = \mathbf{I}^{t_0 + \Delta t_i} - \mathbf{I}^0 \) are, respectively, the change in measured total power (in dB) and the change in contrast profile of the DoI due to the presence of objects of interest. From this form of xPRA, it can be seen that the change in measured total power is linearly related to the change in contrast. This removes multipath reflections that are the same across the two time instants.

The background subtraction form of xPRA also provides calibration. This is because subtracting the total measured data, and the reconstruction PSNR decreases as we change the \( \delta \) value of the test samples. This is because unlike for \( \epsilon_R \), the training data use only one value of \( \delta \), so the network fails to generalize on any other value. This can be addressed in future work by including multiple values of \( \delta \) in the training data.

An important point to note for examples in Figs. 4, 6, and 7 is that all the objects considered are piecewise homogeneous and solid/nonhollow. This is because of an important constraint of our proposed framework due to which it cannot provide good reconstructions for hollow objects. This constraint of the objects being piecewise homogeneous and solid is imposed by the xPRA model and its use in the optimization layers of the proposed unrolled optimization framework. This remains an important limitation of our work and needs to be explored further in future.

VI. EXPERIMENT RESULTS

A. Experiment Setup

The experiment setup we use for indoor imaging is shown in Fig. 3(b) where the environment is approximated as a 2-D cross section at a height of 1 m above the floor. The 2-D cross section at a height of 1 m above the floor, thus creating a total of \( L = M(M - 1)/2 = 80 \) measurement links. Each node consists of a SparkFun ESP32 Thing board with an integrated Wi-Fi transceiver operating at 2.4 GHz, with the inbuilt omni-directional antennas of the SparkFun ESP32 boards replaced by directional Yagi antennas of 6.6 dBi to minimize the multipath reflections from the floor, ceiling, and other clutter placed outside the DoI. More information about the setup can be found in [24].
placed on the Styrofoam platforms in the DoI. We obtain the incident power $P_i$ by only placing the Styrofoam in the DoI and then measure the total power $P$ after placing the objects on the Styrofoam. We then obtain the measurement vector $\Delta P = P - P_i$ (dB), which contains the change in measurements due to the object being placed in the DoI, thereby canceling the scattering effects due to the floor, ceiling, and all the other clutter inside and outside the DoI. We then use these measurements to obtain reconstructions. Note that this straightforward background subtraction is only possible with the xPRA model due to its linear relationship between the change in RSS values and contrast.

C. Experiment Tests

To obtain reconstructions for measurement data obtained from experiments, we use the same end-to-end network of our proposed framework that is trained on the simulation data described in Section V-B. We test our proposed framework on the measurement data obtained using some of the objects that can be commonly found in an indoor region.

Fig. 8 shows results for various such objects placed in the DoI. For all these results, the first column shows the experiment ground truth and the second column shows the reconstruction using xPRA-TK-DPrior. The respective PSNR values of the reconstruction are (a) 36.42 dB, (b) 24.92 dB, and (c) 20.57 dB.

The background subtraction form of xPRA is possible since it is linear in both the measured power and the unknown contrast. For this reason, the nonlinear techniques cannot be extended to include background subtraction.

Fig. 8. Reconstruction results for various objects placed in the DoI where the reconstructed $\delta(r_n)\epsilon_R(r_n)^{1/2}$ values are indicated by color. First column shows the experiment ground truth and the second column shows the reconstruction using xPRA-TK-DPrior. The respective PSNR values of the reconstruction are (a) 36.42 dB, (b) 24.92 dB, and (c) 20.57 dB.

power values in dB for the two time instants implies that we are dividing these values in a linear scale. This helps remove the experimental errors which are present across both the time frames.

The background subtraction form of xPRA is possible since it is linear in both the measured power and the unknown contrast. For this reason, the nonlinear techniques cannot be extended to include background subtraction.

VII. Conclusion

An unrolled optimization framework with deep learning-based priors is proposed to solve highly nonlinear and ill-posed PD-ISPs. The framework uses an approximate linear physics-based model along with data-driven regularization in the form of deep priors, which help extend the validity of the physics-based model to very strong scattering conditions. The optimum values of all the tunable parameters are learned
from data, thus removing the need to tune them manually. The results are demonstrated for both the simulated and experimental data and show that the framework provides accurate reconstructions for piecewise homogeneous objects exhibiting extremely strong scattering and surpasses the validity range of the existing methods by a significant margin.

Future work can include extending this framework to reconstruct hollow objects. The proposed framework can also be extended to 3-D imaging using 6G technologies and reconfigurable intelligent surfaces (RIS). Being able to image perfect electrical conductors (PECs) is also a useful goal.

References

[1] X. Chen, Computational Methods for Electromagnetic Inverse Scattering. Hoboken, NJ, USA: Wiley, 2018.

[2] K. Xu, L. Wu, X. Ye, and X. Chen, “Deep learning-based inversion methods for solving inverse scattering problems with phaseless data,” IEEE Trans. Antennas Propag., vol. 68, no. 11, pp. 7457–7470, Nov. 2020.

[3] L. Pan, Y. Zhong, X. Chen, and S. P. Yeo, “Subspace-based optimization method for inverse scattering problems utilizing phaseless data,” IEEE Trans. Geosci. Remote Sens., vol. 49, no. 3, pp. 981–987, Mar. 2011.

[4] A. Dubey, S. Deshmukh, L. Pan, X. Chen, and R. Murch, “A phaseless extended Rytov approximation for strongly scattering low-loss media and its application to indoor imaging,” IEEE Trans. Geosci. Remote Sens., vol. 60, pp. 1–17, 2022.

[5] K. H. Jin, M. T. Mccann, E. Froustey, and M. Unser, “Deep convolutional neural network for inverse problems in imaging,” IEEE Trans. Image Process., vol. 26, no. 5, pp. 4509–4522, Jun. 2017.

[6] S. Deshmukh, A. Dubey, D. Ma, Q. Chen, and R. Murch, “Physics assisted deep learning for indoor imaging using phaseless Wi-Fi measurements,” IEEE Trans. Antennas Propag., early access, May 30, 2022, doi: 10.1109/TAP.2022.3177533.

[7] Y. Sanghvi, Y. Kalepu, and U. Khankhoje, “Embedding deep learning in inverse scattering problems,” IEEE Trans. Comput. Imag., vol. 6, pp. 46–56, 2020.

[8] K. Zhang, W. Zuo, S. Gu, and L. Zhang, “Learning deep CNN denoiser prior for image restoration,” in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 3929–3938.

[9] J. H. R. Chang, C.-L. Li, B. Poczos, B. V. K. V. Kumar, and A. C. Sankaranarayanan, “One network to solve them all—Solving linear inverse problems using deep projection models,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 5888–5897.

[10] T. Meinhardt, M. Moeller, C. Hazirasas, and D. Cremers, “Learning proximal operators: Using denoising networks for regularizing inverse imaging problems,” in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 1781–1790.

[11] S. Diamond, V. Sitzmann, F. Heide, and G. Wetzstein, “Unrolled optimization with deep priors,” 2017, arXiv:1705.08041.

[12] J. Adler and O. Öktem, “Solving ill-posed inverse problems using iterative deep neural networks,” Inverse Problems, vol. 33, no. 12, 2017, Art. no. 124007.

[13] J. Schlemper, J. Caballero, J. V. Hajnal, A. N. Price, and D. Rueckert, “A deep cascade of convolutional neural networks for dynamic MR image reconstruction,” IEEE Trans. Med. Imag., vol. 37, no. 2, pp. 491–503, Feb. 2018.

[14] A. F. Peterson, S. L. Ray, and R. Mittra, Computational Methods for Electromagnetics, vol. 351. New York, NY, USA: IEEE Press, 1998.

[15] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, “Sparsity and smoothness via the fused lasso,” J. Roy. Stat. Soc. B, Stat. Methodol., vol. 67, pp. 91–108, Feb. 2005.

[16] N. Parikh and S. Boyd, “Proximal algorithms,” Found. Trends Optim., vol. 1, no. 3, pp. 127–239, Jan. 2014.

[17] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” Found. Trends Mach. Learn., vol. 3, no. 1, pp. 1–122, Jan. 2010.

[18] V. Monga, Y. Li, and Y. C. Eldar, “Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing,” IEEE Signal Process. Mag., vol. 38, no. 2, pp. 18–44, Mar. 2021.

[19] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation,” in Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2015, pp. 234–241.

[20] V. Lempitsky, A. Vedaldi, and D. Ulyanov, “Deep image prior,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2018, pp. 9486–9494.

[21] Y. Pinhasi, A. Yahalom, and S. Petnev, “Propagation of ultra wide-band signals in lossy dispersive media,” in Proc. IEEE Int. Conf. Microw., Commun., Antennas Electron. Syst., May 2008, pp. 1–10.

[22] F. Ahmad, M. G. Amin, and T. Dogaru, “Partially sparse imaging of stationary indoor scenes,” EURASIP J. Adv. Signal Process., vol. 2014, no. 1, pp. 1–15, Dec. 2014.

[23] K. C. Yaw, “Measurement of dielectric material properties,” Rohde & Schwarz, Munich, Germany, Appl. Note, 2012, pp. 1–35.

[24] A. Dubey, P. Sood, J. Santos, D. Ma, C.-Y. Chiu, and R. Murch, “An enhanced approach to imaging the indoor environment using WiFi RSSI measurements,” IEEE Trans. Veh. Technol., vol. 70, no. 9, pp. 8415–8430, Sep. 2021.

Amartansh Dubey (Member, IEEE) received the bachelor’s degree in electronics and communication engineering from the Visvesvaraya National Institute of Technology, Nagpur, India, in 2016, and the M.Phil. and Ph.D. degrees in electronic and computer engineering from the Hong Kong University of Science and Technology (HKUST), Hong Kong, in 2018 and 2022, respectively.

He is currently a Post-Doctoral Researcher with HKUST. His research interests include inverse scattering, indoor imaging, machine learning, statistical signal processing, and computational electromagnetics.

Ross Murch (Fellow, IEEE) received the bachelor’s degree in electrical and electronic engineering from the University of Canterbury, Christchurch, New Zealand, in 1987 and 1990, respectively.

He was the Department Head of the Department of Electronic and Computer Engineering, Hong Kong University of Science and Technology (HKUST), Hong Kong, for two three-year terms from 2009 to 2015. He is currently a Chair Professor with the Department of Electronic and Computer Engineering and an IAS Fellow with the Institute of Advanced Study, HKUST. His unique expertise lies in his intradisciplinary knowledge of both wireless communication systems and electromagnetics. He publishes in both disciplines and research highlights from this intradisciplinary approach include being one of the first to propose multiuser multiple input multiple output (MU-MIMO) systems and pioneering compact MIMO antenna designs. In total, his research contributions include nearly 200 journal publications and 20 patents, and he has successfully supervised over 50 research students. His research focuses on creating new RF wave technology for making a better world and this includes areas such as the Internet-of-Things, RF imaging, RF sensing, RF navigation, ambient RF systems, energy harvesting, electromagnetic information theory, 6G, multiprot antenna systems, and reconfigurable intelligent surfaces.

Prof. Murch is a fellow of Institute of Engineering and Technology (IET) and Hong Kong Institution of Engineers (HKIE). He has been involved in IEEE activities including being an Area Editor of IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS and the Chair of the IEEE Technology Committee on Wireless Communications. He has won several prizes including three teaching awards.