# A Plug-and-Play Priors Approach for Solving Nonlinear Imaging Inverse Problems

Ulugbek S. Kamilov, *Member, IEEE*, Hassan Mansour, *Senior Member, IEEE*, and Brendt Wohlberg, *Senior Member, IEEE*

***Abstract*—In the past two decades, nonlinear image reconstruction methods have led to substantial improvements in the capabilities of numerous imaging systems. Such methods are traditionally formulated as optimization problems that are solved iteratively by simultaneously enforcing data consistency and incorporating prior models. Recently, the Plug-and-Play Priors (PPP) framework suggested that by using more sophisticated denoisers, not necessarily corresponding to an optimization objective, it is possible to improve the quality of reconstructed images. In this letter, we show that the PPP approach is applicable beyond linear inverse problems. In particular, we develop the fast iterative shrinkage/thresholding algorithm (FISTA) variant of PPP for model-based nonlinear inverse scattering. The key advantage of the proposed formulation over the original ADMM-based one is that it does not need to perform an inversion on the forward model. We show that the proposed method produces high quality images using both simulated and experimentally measured data.**

***Index Terms*—Image reconstruction, plug-and-play priors, FISTA, inverse scattering, nonlinear inverse problems.**

## I. Introduction

THE problem of estimating an unknown image from noisy measured data is fundamental in computational imaging. The estimation task is often formulated as an inverse problem that minimizes a cost functional. The functional typically includes two terms: a data-fidelity term that ensures that the final image is consistent with measured data and a regularizer that mitigates the ill-posedness of the problem by promoting solutions with desirable properties [1]. For example, the $\ell_1$-penalty promoting spatial sparsity and total variation (TV) penalty promoting sparsity in the image gradient are popular image regularizers [2]–[4].

Optimization problems in imaging are often nontrivial to solve. The challenging aspects are the non-smooth nature of the popular regularizers and the large amount of data that needs to be processed in a typical application. Proximal methods [5] such as fast iterative shrinkage/thresholding algorithm (FISTA) [6] and alternating direction method of multipliers (ADMM) [7] are standard approaches that circumvent the non-smoothness of the regularizers while being able to handle large-scale optimization problems. The key ingredient in both of these methods is the proximal operator, which in itself is an optimization problem corresponding to a simple denoising of intermediate solutions.

Recently, Venkatakrishnan *et al.* [8] introduced the idea of Plug-and-Play Priors (PPP), which simply replaces the proximal operator by a suitable image denoising method. Plug-and-play methods generally lose interpretability as optimization problems, as most image denoising algorithms such as BM3D [9] do not have a known optimization formulation. Nonetheless, the framework has gained in popularity due to its effectiveness in a range of applications, in particular, in the context of linear inverse problems [10]–[13].

In this letter, we develop a new algorithm based on PPP and show its effectiveness in solving nonlinear imaging inverse problems. Specifically, we consider the problem of reconstructing the spatial distribution of dielectric permittivity from the measurements of its scattered light. This problem is common in numerous imaging applications such as tomographic microscopy, digital holography, and radar imaging [14]–[18]. When scattering is strong, multiple scattering of light leads to a nonlinear forward model [19], for which the challenge lies in finding computationally tractable methods that can account for the nonlinearity while also accommodating model-based priors for imaging. To that end, we propose a new FISTA variant of PPP for model-based nonlinear inverse scattering. The nonlinearity of the forward model is handled by a recently introduced technique—called Series Expansion with Accelerated Gradient Descent on Lippmann-Schwinger Equation (SEAGLE)—based on a series expansion of the scattered light with an accelerated-gradient method [20], [21]. The proposed method is validated on both simulated and experimentally measured data.

## II. Proposed Approach

In this section, we describe Plug-and-Play FISTA for solving nonlinear inverse problems. As a practically relevant nonlinear inverse problem, we also introduce the problem of estimating the spatial permittivity distribution of an unknown object from the measurements of the scattered waves.

### A. Nonlinear Inverse Problem

We consider an imaging inverse problem

$$\mathbf{y} = \mathbf{H}(\mathbf{x}) + \mathbf{e}\,, \tag{1}$$

where the goal is to recover the unknown image $\mathbf{x} \in \mathbb{R}^N$ from the noisy measurements $\mathbf{y} \in \mathbb{C}^M$. The measurement operator $\mathbf{H} : \mathbb{R}^N \to \mathbb{C}^M$ models the response of the imaging system

U. S. Kamilov (email: kamilov@wustl.edu) is with Computational Imaging Group (CIG), Washington University in St. Louis, St. Louis, MO 63130, USA.

H. Mansour (email: mansour@merl.com) is with Mitsubishi Electric Research Laboratories (MERL), 201 Broadway, Cambridge, MA 02139, USA.

B. Wohlberg (email: brendt@lanl.gov) is with Los Alamos National Laboratory, Theoretical Division, Los Alamos, NM 87545 USA.

and the vector $\mathbf{e} \in \mathbb{C}^M$ represents the measurement noise, which is often assumed to be independent and identically distributed (i.i.d.) Gaussian. When the inverse problem is linear, the measurement operator is represented as a measurement matrix $\mathbf{H} \in \mathbb{C}^{M \times N}$.

In practice, problems such as (1) are often ill-posed; the standard approach for solving them is by formulating an optimization problem

$$\widehat{\mathbf{x}} = \underset{\mathbf{x} \in \mathbb{R}^N}{\arg\min} \left\{ \mathcal{D}(\mathbf{x}) + \tau \mathcal{R}(\mathbf{x}) \right\}, \tag{2}$$

where the data-fidelity term $\mathcal{D}$ ensures that the final image is consistent with measured data, the regularizer $\mathcal{R}$ promotes solutions with desirable properties, and $\tau \in \mathbb{R}_+$ is a parameter that controls the strength of the regularization. For example, a popular data-fidelity term is least-squares

$$\mathcal{D}(\mathbf{x}) \triangleq \frac{1}{2} \|\mathbf{y} - \mathbf{H}(\mathbf{x})\|_{\ell_2}^2 . \tag{3}$$

Two common regularizers for images include the spatial sparsity-promoting penalty $\mathcal{R}(\mathbf{x}) \triangleq \|\mathbf{x}\|_{\ell_1}$ and total variation (TV) penalty $\mathcal{R}(\mathbf{x}) \triangleq \|\mathbf{D}\mathbf{x}\|_{\ell_1}$, where $\mathbf{D}$ is the discrete gradient operator [2]–[4].

FISTA is a well-studied iterative method, effective in large-scale setting, for solving (2) when $\mathcal{R}$ is non-smooth. It can be expressed as

$$\mathbf{z}^t \leftarrow \mathbf{s}^{t-1} - \gamma \nabla \mathcal{D}(\mathbf{s}^{t-1}) \tag{4a}$$

$$\mathbf{x}^t \leftarrow \mathrm{prox}_{\gamma\tau\mathcal{R}}\left(\mathbf{z}^t\right) \tag{4b}$$

$$q_t \leftarrow \frac{1}{2}\left(1 + \sqrt{1 + 4q_{t-1}^2}\right) \tag{4c}$$

$$\mathbf{s}^t \leftarrow \mathbf{x}^t + ((q_{t-1} - 1)/q_t)(\mathbf{x}^t - \mathbf{x}^{t-1}) , \tag{4d}$$

for $t = 1, 2, \ldots$, where $q_0 = 1$ and $\mathbf{x}^0 = \mathbf{s}^0 = \mathbf{x}_{\text{init}} \in \mathbb{R}^N$. Here, $\gamma \in \mathbb{R}_+$ is a step-size and $\{q_t\}_{t \in \mathbb{N}}$ are the parameters for improving the global convergence rate of the algorithm [6]. Note that different strategies for selecting $\gamma$ for nonconvex $\mathcal{D}$ and convex $\mathcal{R}$ were described in [22]. The proximal operator in (4) is defined as

$$\mathrm{prox}_{\tau\mathcal{R}}(\mathbf{z}) \triangleq \underset{\mathbf{x} \in \mathbb{R}^N}{\arg\min} \left\{ \frac{1}{2}\|\mathbf{x} - \mathbf{z}\|_{\ell_2}^2 + \tau \mathcal{R}(\mathbf{x}) \right\}, \tag{5}$$

and corresponds to an image denoising problem with a given regularizer $\mathcal{R}$. Note how (4a) only depends on the forward model via the gradient of the data-fidelity term, while (4b) only depends on the prior via the proximal operator of the regularizer.

### B. Plug-and-Play Priors with FISTA

FISTA enjoys a modular structure similar to ADMM, where the prior model on the image is only imposed via the proximal operator [8]. In the spirit of plug-and-play, one can then consider the Plug-and-Play FISTA

$$\mathbf{z}^t \leftarrow \mathbf{s}^{t-1} - \gamma \nabla \mathcal{D}(\mathbf{s}^{t-1}) \tag{6a}$$

$$\mathbf{x}^t \leftarrow \mathrm{D}_\sigma\left(\mathbf{z}^t\right) \tag{6b}$$

$$q_t \leftarrow \frac{1}{2}\left(1 + \sqrt{1 + 4q_{t-1}^2}\right) \tag{6c}$$

$$\mathbf{s}^t \leftarrow \mathbf{x}^t + ((q_{t-1} - 1)/q_t)(\mathbf{x}^t - \mathbf{x}^{t-1}) , \tag{6d}$$

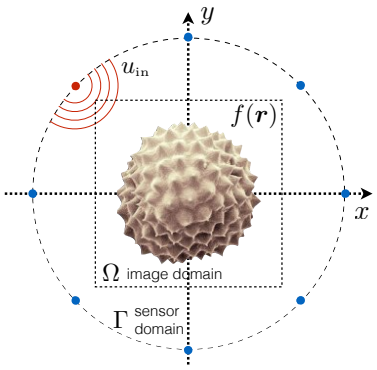


Fig. 1. Schematic representation of scattering scenarios considered in this letter. An object of a scattering potential $f(\boldsymbol{r})$ is illuminated with an input wave $u_{\text{in}}$, which interacts with the object and leads to the scattered field $u_{\text{sc}}$ at the sensors.

where $\mathrm{D}_\sigma$ is a general denoiser of strength $\sigma \triangleq \sqrt{\gamma\tau} \in \mathbb{R}_+$, which might not necessarily conform to equation (5). For example, we will later see that the denoising operators BM3D [9] and WNNM [23], which are not easily formulated as an optimization problem, can be used as a powerful regularizer for nonlinear inverse scattering of images consisting of similar, spatially replicated features.

The main advantage of Plug-and-Play FISTA (6) over the original ADMM-based formulation is that there is no need to perform an inversion on the forward model. This advantage becomes especially important when dealing with nonlinear inverse problems, as such inversions become computationally prohibitive. FISTA only requires an efficient evaluation of the gradient $\nabla\mathcal{D}$ without an explicit inversion of the forward model. This gradient can be easily obtained for a range of nonlinear inverse problems, including for nonlinear inverse scattering, described in Section II-C.

Convergence of Plug-and-Play ADMM to the global minimum of some implicitly defined MAP cost function was recently studied by Sreehari *et al.* [10]. In particular, the work builds on the result by Moreau [24] that establishes that $\mathrm{D}_\sigma$ is a valid proximal mapping of some implicit regularizer if and only if it is non-expansive and it is the sub-gradient of a convex function on $\mathbb{R}^N$. Note that when $\mathcal{D}$ is convex and the sufficient conditions outlined in [10] are met, the global convergence of our Plug-and-Play FISTA (6) can be established in a similar fashion to Plug-and-Play ADMM. However, as we shall see, similarly to Plug-and-Play ADMM, Plug-and-Play FISTA exhibits reliably converging behavior even for a possibly nonconvex $\mathcal{D}$ and a general denoiser $\mathrm{D}_\sigma$.

### C. Nonlinear Inverse Scattering

Consider the scattering problem (Fig. 1) where an object of the permittivity distribution $\epsilon(\boldsymbol{r})$ in the bounded domain $\Omega \subseteq \mathbb{R}^2$ is immersed into a background medium of permittivity $\epsilon_b$ and illuminated with the incident electric field $u_{\text{in}}(\boldsymbol{r})$. We assume that the incident field is monochromatic and coherent, and it is known inside $\Omega$ and at the locations of the sensors. The result of object-light interaction is measured at the location of the sensors as a scattered field $u_{\text{sc}}(\boldsymbol{r})$. The

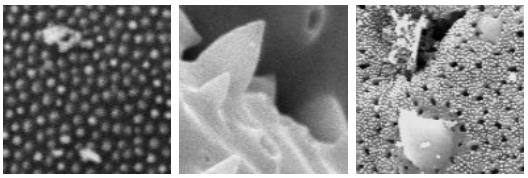

Fig. 2. Three scanning electron micrographs. From left to right: *Zea mays* pollen, *Schlumbergera truncata* pollen, and *Malva alcea* pollen.

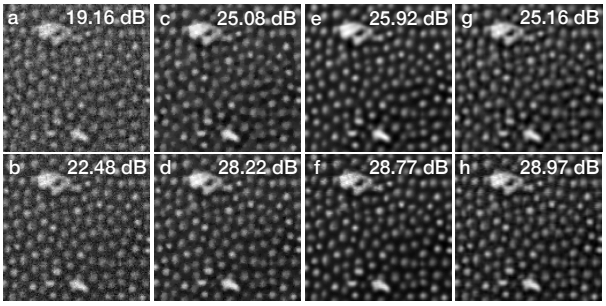


Fig. 3. Illustration of the reconstructed image for *Zea mays* at $f_{\max} = 1\mathrm{e}-4$ with 40 dB input SNR: (a) linear, non-negativity; (b) nonlinear, non-negativity; (c) linear, TV; (d) nonlinear, TV; (e) linear, WNNM; (f) nonlinear, WNNM; (g) linear, BM3D; (h) nonlinear, BM3D.

multiple scattering of light can be accurately described by the Lippmann-Schwinger equation [19] inside the image domain

$$u(\boldsymbol{r}) = u_{\text{in}}(\boldsymbol{r}) + \int_{\Omega} g(\boldsymbol{r}-\boldsymbol{r}')\, f(\boldsymbol{r}')\, u(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}', \quad (\boldsymbol{r} \in \Omega) \quad (7)$$

where $u(\boldsymbol{r}) = u_{\text{in}}(\boldsymbol{r}) + u_{\text{sc}}(\boldsymbol{r})$ is the total electric field, $f(\boldsymbol{r}) \triangleq k^2(\epsilon(\boldsymbol{r}) - \epsilon_b)$ is the scattering potential, which is assumed to be real, and $k = 2\pi/\lambda$ is the wavenumber in vacuum. The function $g(\boldsymbol{r})$ is the Green's function, defined as

$$g(\boldsymbol{r}) \triangleq \frac{\mathrm{j}}{4} H_0^{(1)}(k_b \|\boldsymbol{r}\|_{\ell_2}) \quad (8)$$

where $k_b \triangleq k\sqrt{\epsilon_b}$ is the wavenumber of the background medium and $H_0^{(1)}$ is the zero-order Hankel function of the first kind. Note that the knowledge of the total-field $u$ inside the image domain $\Omega$ enables the prediction of the scattered field at the sensor domain

$$u_{\text{sc}}(\boldsymbol{r}) = \int_{\Omega} g(\boldsymbol{r}-\boldsymbol{r}')\, f(\boldsymbol{r}')\, u(\boldsymbol{r}')\, \mathrm{d}\boldsymbol{r}'. \quad (\boldsymbol{r} \in \Gamma) \quad (9)$$

The discretization and combination of (7) and (9) leads to the following matrix-vector description of the scattering problem

$$\mathbf{y} = \mathbf{S}(\mathbf{u} \odot \mathbf{x}) + \mathbf{e} \quad (10a)$$

$$\mathbf{u} = \mathbf{u}_{\text{in}} + \mathbf{G}(\mathbf{u} \odot \mathbf{x}), \quad (10b)$$

where $\mathbf{x} \in \mathbb{R}^N$ is the discretized scattering potential $f$, $\mathbf{y} \in \mathbb{C}^M$ is the measured scattered field $u_{\text{sc}}$ at $\Gamma$, $\mathbf{u}_{\text{in}} \in \mathbb{C}^N$ is the input field $u_{\text{in}}$ inside $\Omega$, $\mathbf{S} \in \mathbb{C}^{M\times N}$ is the discretization of the Green's function at $\Gamma$, $\mathbf{G} \in \mathbb{C}^{N\times N}$ is the discretization of the Green's function inside $\Omega$, $\odot$ denotes a component-wise multiplication between two vectors, and $\mathbf{e} \in \mathbb{C}^M$ models the additive noise at the measurements. Using the shorthand

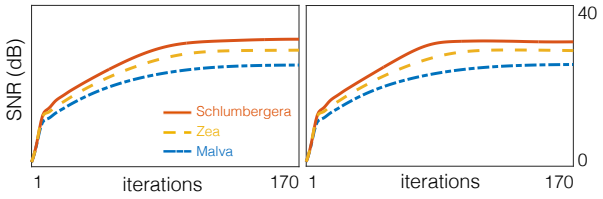


Fig. 4. Evolution of average SNR over all contrast levels with respect to iteration number. Left: nonlinear, BM3D. Right: nonlinear, WNNM.

notation $\mathbf{A} \triangleq \mathbf{I} - \mathbf{G}\mathrm{diag}(\mathbf{x})$, where $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix and $\mathrm{diag}(\cdot)$ is an operator that forms a diagonal matrix from its argument, we can formally specify the nonlinear forward model in (1) as follows

$$\mathbf{H}(\mathbf{x}) \triangleq \mathbf{S}(\widehat{\mathbf{u}}(\mathbf{x}) \odot \mathbf{x}) \quad (11a)$$

where

$$\widehat{\mathbf{u}}(\mathbf{x}) \triangleq \underset{\mathbf{u}\in\mathbb{C}^N}{\arg\min} \{\mathcal{S}(\mathbf{u})\} \quad (11b)$$

$$\text{with } \mathcal{S}(\mathbf{u}) \triangleq \frac{1}{2}\|\mathbf{A}\mathbf{u} - \mathbf{u}_{\text{in}}\|_{\ell_2}^2. \quad (11c)$$

Most remarkably, several recent works have proposed computationally tractable approaches for evaluating the gradient

$$\nabla \mathcal{D}(\mathbf{x}) = \left[\frac{\partial \mathbf{H}(\mathbf{x})}{\partial \mathbf{x}}\right]^{\mathrm{H}} (\mathbf{H}(\mathbf{x}) - \mathbf{y}) \quad (12)$$

which can be readily plugged into steps (4a) and (6a) of FISTA [20], [21], [25], [26]. In the remainder of this work, we rely on the SEAGLE algorithm [20], [21], which is obtained by expanding the forward computation using Nesterov's accelerated gradient method [27] for a fixed number of iterations. The final expression for the gradient is then obtained by running an error backpropagation algorithm on a fixed number of Nesterov's iterations [20].

## III. Performance Evaluation

We now present the results of validating our method on two distinct sets of scattering data. The first data-set was obtained using a high-fidelity solution of the forward scattering problem with the conjugate-gradient solver. The second data-set was collected and made publicly available by Institut Fresnel [28].

We validate the model-based recovery capability of Plug-and-Play FISTA by selecting three electron microscopy images[1] with high levels of spatial self-similarity (see Fig. 2). Non-local algorithms such as BM3D and WNNM are known to be effective for denoising such images [9], [23]. We consider images of physical size 18 cm $\times$ 18 cm, discretized to a $128 \times 128$ grid. The background medium is assumed to be air with $\epsilon_b = 1$ and the wavelength of the illumination is set to $\lambda = 0.15$ cm. The measurements are collected over 30 transmissions uniformly distributed along a circle of radius 1.78 m and, for each transmission, 360 measurements around the object are recorded. The simulated scattered data was

[1] All images are in public domain and can be downloaded at the Cell Image Library `http://www.cellimagelibrary.org/images`.

TABLE I
SNR (dB) PERFORMANCES OF THREE METHODS ON THE SET OF THREE IMAGES FOR VARIOUS CONTRAST LEVELS

| Image name | Contrast ($f_{\max}$) | Non-negativity Constraint | | Total Variation (TV) | | Plug-and-Play WNNM | | Plug-and-Play BM3D | |
|---|---|---|---|---|---|---|---|---|---|
| | | linear | nonlinear | linear | nonlinear | linear | nonlinear | linear | nonlinear |
| *Zea mays* | 1e−5 | 22.42 | 22.46 | 28.17 | 28.22 | 28.75 | 28.78 | 29.05 | **29.13** |
| | 1e−4 | 19.16 | 22.48 | 25.08 | 28.22 | 25.92 | 28.77 | 25.16 | **28.97** |
| | 1e−3 | 4.46 | 22.33 | 16.36 | 28.24 | 15.99 | 28.76 | 15.40 | **29.02** |
| *Schlumbergera truncata* | 1e−5 | 22.65 | 22.71 | 31.34 | 31.43 | 30.91 | 30.99 | 31.48 | **31.56** |
| | 1e−4 | 17.54 | 22.72 | 28.00 | 31.44 | 28.21 | 30.97 | 27.90 | **31.59** |
| | 1e−3 | 3.61 | 22.70 | 19.47 | 31.42 | 20.12 | 30.96 | 19.77 | **31.54** |
| *Malva alcae* | 1e−5 | 21.93 | 21.99 | 24.76 | 24.82 | 25.27 | **25.32** | 25.13 | 25.17 |
| | 1e−4 | 17.35 | 22.11 | 20.86 | 24.67 | 21.24 | **25.27** | 21.01 | 25.09 |
| | 1e−3 | 3.49 | 22.18 | 13.75 | 24.68 | 12.76 | 25.18 | 13.26 | **25.19** |

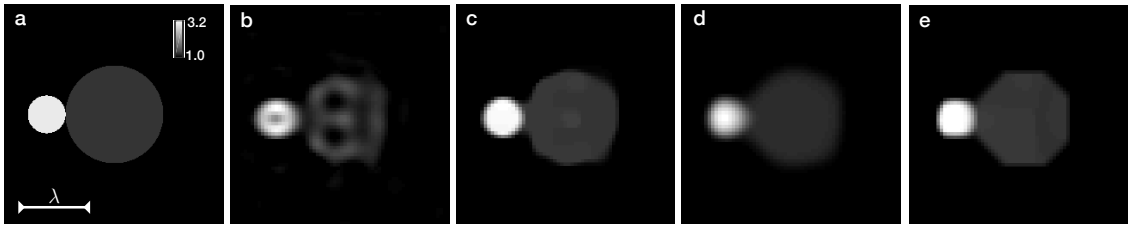


Fig. 5. Experimental validation of the proposed method on the highly scattering object *FoamDielExtTM*: (a) ground truth; (b) non-negativity; (c) TV; (d) WNNM; (e) BM3D. All results correspond to a nonlinear forward model. The scale bar length is equal to the wavelength $\lambda$ at 5 GHz.

additionally corrupted by an additive white Gaussian noise (AWGN) corresponding to 40 dB of input signal-to-noise ratio (SNR).

Table I summarizes the results of comparing four distinct priors: non-negativity constraints on the image, total variation (TV), and two plug-and-play priors, BM3D and WNNM. Note that for TV, BM3D, and WNNM non-negativity of the solution was imposed by setting the negatively valued pixels in the iterates to zero. For each prior, we consider the effects of the linearity versus nonlinearity of the measurement model. The linear measurement model is obtained using the first Born-approximation [29]. The nonlinear model takes into account multiple scattering using the SEAGLE framework. The permittivity contrast was defined as $f_{\max} \triangleq (\epsilon_{\max} - \epsilon_b)/\epsilon_b$, where $\epsilon_{\max} \triangleq \max_{\boldsymbol{r}\in\Omega}\{\epsilon(\boldsymbol{r})\}$. The permittivity contrast quantifies the degree of nonlinearity in the inverse problem, with higher $f_{\max}$ indicating stronger levels of multiple scattering. The regularization parameters for all the experiments were optimized for the best SNR performance. For each prior, the maximal number of FISTA iterations was set to $t_{\max} = 500$ with an additional tolerance parameter $\delta_{\max} = 10^{-4}$ on the relative change of the solution in two successive iterations. We relied on a fixed step-size $\gamma = 10$ that was manually selected to ensure the convergence. The results confirm that as the level of scattering increases, the performance under the linear inverse problem formulation degenerates with or without regularization. On the other hand, regularization leads to substantial improvements in quality under both linear and nonlinear inverse problem formulations. The results confirm that for highly self-similar images, the best quality results are obtained by using nonlocal priors such as BM3D or WNNM. Fig. 3 visually illustrates the result for *Zea mays* at $f_{\max} = 1\mathrm{e}-4$. Finally, Fig. 4 illustrates the convergence of Plug-and-Play FISTA to the final solution. We observed the method to be stable over all the experiments reported here.

To experimentally validate Plug-and-Play FISTA, we consider microwave measurements at 5 GHz of the object *FoamDielExtTM* from the dataset in [28]. The object consists of two cylinders on a rim of radius 1.67 m, and a measurement set-up that is identical to that in Fig. 1. The object is highly scattering due to a high permittivity contrast of $f_{\max} = 2$. The image domain has size 18 cm $\times$ 18 cm, discretized to a $64 \times 64$ grid. There are 9 transmitters and 240 detectors for each transmitter, all equally spaced on the rim. The results of image reconstruction are summarized in Fig. 5. As expected, the imaging results indicate that TV is well suited for piecewise-constant signals such as our object. On the other hand, we also note that both BM3D and WNNM substantially improve over the simple non-negativity prior by offering stronger regularization of the solution. Generally, our results indicates that Plug-and-Play FISTA can be used to impose more sophisticated priors in the context of inverse scattering on experimental data.

## IV. Conclusion

The PPP framework is a powerful approach for model-based image regularization. In this letter, we have presented a variant of the PPP based on FISTA, for nonlinear inverse problems where the data-fidelity term can be efficiently differentiated. We have illustrated the method by reconstructing images from the measurements of scattered waves and validated it on both simulated and experimentally measured data.

## REFERENCES


[1] A. Ribés and F. Schmitt, "Linear inverse problems in imaging," *IEEE Signal Process. Mag.*, vol. 25, no. 4, pp. 84–99, July 2008.

[2] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D*, vol. 60, no. 1–4, pp. 259–268, November 1992.

[3] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, February 2006.

[4] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, April 2006.

[5] H. H. Bauschke and P. L. Combettes, *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer, 2010.

[6] A. Beck and M. Teboulle, "Fast gradient-based algorithm for constrained total variation image denoising and deblurring problems," *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2419–2434, November 2009.

[7] M. V. Afonso, J. M.Bioucas-Dias, and M. A. T. Figueiredo, "Fast image recovery using variable splitting and constrained optimization," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2345–2356, September 2010.

[8] S. V. Venkatakrishnan, C. A. Bouman, and B. Wohlberg, "Plug-and-play priors for model based reconstruction," in *Proc. IEEE Global Conf. Signal Process. and INf. Process. (GlobalSIP)*, Austin, TX, USA, December 3-5, 2013, pp. 945–948.

[9] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 16, pp. 2080–2095, August 2007.

[10] S. Sreehari, S. V. Venkatakrishnan, B. Wohlberg, G. T. Buzzard, L. F. Drummy, J. P. Simmons, and C. A. Bouman, "Plug-and-play priors for bright field electron tomography and sparse interpolation," *IEEE Trans. Comp. Imag.*, vol. 2, no. 4, pp. 408–423, December 2016.

[11] S. H. Chan, X. Wang, and O. A. Elgendy, "Plug-and-play ADMM for image restoration: Fixed-point convergence and applications," *IEEE Trans. Comp. Imag.*, vol. 3, no. 1, pp. 84–98, March 2017.

[12] A. Brifman, Y. Romano, and M. Elad, "Turning a denoiser into a super-resolver using plug and play priors," in *Proc. IEEE Int. Conf. Image Proc. (ICIP 2016)*, Phoenix, AZ, USA, September 25-28, 2016, pp. 1404–1408.

[13] A. M. Teodoro, J. M. Biocas-Dias, and M. A. T. Figueiredo, "Image restoration and reconstruction using variable splitting and class-adapted image priors," in *Proc. IEEE Int. Conf. Image Proc. (ICIP 2016)*, Phoenix, AZ, USA, September 25-28, 2016, pp. 3518–3522.

[14] W. Choi, C. Fang-Yen, K. Badizadegan, S. Oh, N. Lue, R. R. Dasari, and M. S. Feld, "Tomographic phase microscopy," *Nat. Methods*, vol. 4, no. 9, pp. 717–719, September 2007.

[15] D. J. Brady, K. Choi, D. L. Marks, R. Horisaki, and S. Lim, "Compressive holography," *Opt. Express*, vol. 17, no. 15, pp. 13 040–13 049, 2009.

[16] L. Tian, N. Loomis, J. A. Dominguez-Caballero, and G. Barbastathis, "Quantitative measurement of size and three-dimensional position of fast-moving bubbles in airwater mixture flows using digital holography," *Appl. Opt.*, vol. 49, no. 9, pp. 1549–1554, March 2010.

[17] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Learning approach to optical tomography," *Optica*, vol. 2, no. 6, pp. 517–522, June 2015.

[18] D. Liu, U. S. Kamilov, and P. T. Boufounos, "Compressive tomographic radar imaging with total variation regularization," in *Proc. IEEE 4th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar, and Remote Sensing (CoSeRa 2016)*, Aachen, Germany, September 19-22, 2016, pp. 120–123.

[19] M. Born and E. Wolf, *Principles of Optics*, 7th ed. Cambridge Univ. Press, 2003, ch. Scattering from inhomogeneous media, pp. 695–734.

[20] H.-Y. Liu, U. S. Kamilov, D. Liu, H. Mansour, and P. T. Boufounos, "Compressive imaging with iterative forward models," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP 2017)*, New Orleans, LA, USA, March 5-9, 2017, pp. 6025–6029.

[21] H.-Y. Liu, D. Liu, H. Mansour, P. T. Boufounos, L. Waller, and U. S. Kamilov, "SEAGLE: Sparsity-driven image reconstruction under multiple scattering," May 2017, arXiv:1705.04281 [cs.CV].

[22] P. Ochs, Y. Chen, and T. Pock, "iPiano: Inertial proximal algorithm for nonconvex optimization," *SIAM J. Imaging Sci.*, vol. 7, no. 2, pp. 1388–1419, 2014.

[23] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Columbus, OH, USA, September 30-October 3, 2014, pp. 2862–2869.

[24] J. J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bull. Soc. Math. France*, vol. 93, pp. 273–299, 1965.

[25] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, "Optical tomographic image reconstruction based on beam propagation and sparse regularization," *IEEE Trans. Comp. Imag.*, vol. 2, no. 1, pp. 59–70,, March 2016.

[26] U. S. Kamilov, D. Liu, H. Mansour, and P. T. Boufounos, "A recursive Born approach to nonlinear inverse scattering," *IEEE Signal Process. Lett.*, vol. 23, no. 8, pp. 1052–1056, August 2016.

[27] Y. E. Nesterov, "A method for solving the convex programming problem with convergence rate $O(1/k^2)$," *Dokl. Akad. Nauk SSSR*, vol. 269, pp. 543–547, 1983, (in Russian).

[28] J.-M. Geffrin, P. Sabouroux, and C. Eyraud, "Free space experimental scattering database continuation: experimental set-up and measurement precision," *Inv. Probl.*, vol. 21, no. 6, pp. S117–S130, 2005.

[29] E. Wolf, "Three-dimensional structure determination of semi-transparent objects from holographic data," *Opt. Commun.*, vol. 1, no. 4, pp. 153–156, September/October 1969.