Regularization by Denoising via Fixed-Point Projection (RED-PRO)
Regev Cohen∗, Michael Elad†, and Peyman Milanfar‡

Abstract. Inverse problems in image processing are typically cast as optimization tasks, consisting of data fidelity and stabilizing regularization terms. A recent regularization strategy of great interest utilizes the power of denoising engines. Two such methods are the Plug-and-Play Prior (PnP) and Regularization by Denoising (RED). While both have shown state-of-the-art results in various recovery tasks, their theoretical justification is incomplete. In this paper, we aim to enrich the understanding of RED and its connection to PnP. Towards that end, we reformulate RED as a convex optimization problem utilizing a projection (RED-PRO) onto the fixed-point set of demicontractive denoisers. We offer a simple iterative solution to this problem, and establish a novel unification of RED-PRO and PnP, while providing guarantees for their convergence to the globally optimal solution. We also present several relaxations of RED-PRO that allow for handling denoisers with limited fixed-point sets. Finally, we demonstrate RED-PRO for the tasks of image deblurring and super-resolution, showing improved results with respect to the original RED framework.

Key words. Inverse problems, Image denoising, Plug and Play Prior (PnP), Regularization by Denoising (RED), Demicontractive mappings, Fixed-point set.

AMS subject classifications. 62H35, 68U10, 94A08, 65F10, 65F22, 47A52.

1. Introduction. Inverse problems arise in numerous fields, ranging from astrophysics and optics to signal processing, computer vision and medical imaging [26, 25, 24, 69, 22]. Specifically to computational imaging, inverse problems relate to the task of inferring an unknown image from its corrupted measurements. Assuming a known degradation model (i.e., the forward operator), one may offer a formulation of the recovery process as an optimization problem comprising a data fidelity term. Unfortunately, this by itself is typically insufficient, as it leads to an ill-posed problem with a non-stable solution. To overcome this difficulty, regularization methods have been widely developed and used. A proper regularization enables a robust recovery by leveraging prior information on the unknown image, incorporated into the optimization formulation. Therefore, formulating and solving inverse problems often reduces to determining the appropriate regularization, depending on the specific application and the underlying signals in mind.

Many regularization schemes for natural images are available, and it is beyond the scope of this paper to review this rich literature. The focus of this paper is on the fascinating recent idea that image denoisers could be used as the mechanism behind the regularization term [65, 80]. Image denoising is the simplest known inverse problem, concerned with obtaining a clean image from its noisy measurements contaminated with additive noise, typically assumed to be zero-mean Gaussian distributed. Over the past decade, image denoising has been studied extensively, leading to a vast line of works (e.g., [75, 36, 27, 73, 53, 29, 32, 20, 14, 48, 47, 44, 74, 39, 30, 61, 87, 21, 85, 40, 45, 68]). This has resulted in the availability of extremely
efficient and effective algorithms, achieving nearly optimal performance [19, 49, 50].

The first to propose leveraging the power of denoising for regularization were Venkatakrishnan et al., presenting their Plug-and-Play Prior (PnP) framework [80, 18, 70]. PnP relies on the alternating direction method of multipliers (ADMM) for solving general inverse problems, and their method amounts to an iterative technique that decomposes the optimization problem into a sequence of simple denoising operations, known as proximal algorithms [59, 7]. The PnP approach replaces the proximal operators with state-of-the-art denoisers, thus, embodying implicit priors for regularizing inverse problems. This framework has gained great interest due to its success in various applications [70, 81, 63, 43, 88, 1, 82], and it has been extended to other proximal algorithms such as proximal gradient method (PGM) [8, 59], approximate message passing (AMP) [58, 33, 6] and half quadratic splitting [86]. Schemes similar to PnP has been proposed in [28] and [34], where the former is based on the augmented Lagrangian method and the latter relies on the notion of Nash equilibrium.

Despite its empirical success, when arbitrary denoising engines are used, PnP loses its interpretability as an optimization problem, making its convergence behavior difficult to analyse. Several studies have proven the fixed-point convergence of variants of PnP under various assumptions. Chan et al. [18] proved the convergence of PnP-ADMM with increasing penalty parameter under the assumption of bounded denoisers. A similar condition was used in [78] for the analysis of a variant of PnP. In [15], PnP has been explained via the framework of consensus equilibrium, proving the convergence of PnP for nonexpansive denoisers. The latter has been the core assumption in multiple works [70, 72, 76, 17, 77], which proved the convergence of several PnP techniques by reformulating them as fixed-point iterations of (averaged) nonexpansive mappings. In [67], the authors relaxed this assumption to a certain Lipschitz condition. Yet, none of the above proved the convergence of PnP to global minima of an explicit objective function.

As a partial remedy and an enrichment to the above, Romano et al. introduced an alternative approach called Regularization by Denoising (RED) [65]. Here, a denoiser engine is utilized to form an explicit regularizer, consisting of the inner product between the unknown image and its denoising residual. Interestingly, under certain conditions, the RED prior defines a convex function whose gradient is simply given by the denoising residual itself. Thus, this gradient can be used by first-order optimization methods such as steepest decent (SD), fixed-point (FP) iteration and ADMM, leading to a convergence to the globally optimal solution of the regularized inverse problem.

The RED framework relies on a clear and well-defined objective function, and has demonstrated state-of-the-art results in image deblurring and super-resolution, drawing a broad attention [62, 71, 56, 42, 81, 60, 86, 87]. However, RED formulation requires the denoiser to be differentiable, obey a local-homogeneity property, and have a symmetric Jacobian, conditions which are not met by various powerful denoisers [62], such as non-local means (NLM) [13], block-matching and 3D filtering (BM3D) [27], trainable nonlinear reaction-diffusion (TNRD) [21], and denoising convolutional neural network (DnCNN) [85]. In such cases, RED algorithms cease to act as optimization solvers and are deprived from their convergence guarantees to global optimum. To overcome this, Reehorst and Schniter introduced a framework called score-matching by denoising (SMD) [62], which offers a new interpretation of RED algorithms and proves their converge to a fixed-point assuming nonexpansive denois-
ers. In [71], a block coordinate RED algorithm was presented, including an analysis of its fixed-point convergence under the condition that the denoiser is block-nonexpansive. These works, however, did not show convergence to global minima of an explicit energy function.

As becomes clear from the above, the convergence justifications for both PnP and RED frameworks are incomplete. In this paper, our main goal is to enrich the theoretical understanding of RED by offering a new interpretation for its regularization strategy. Another objective we pose is the formation of a theoretical bridge between RED and PnP. The main contributions of this paper are two-fold:

• We re-introduce RED via the Fixed-Point Projection (RED-PRO) strategy. We present a framework where the denoiser is assumed to be a demicontractive function with non-empty fixed-point set. This allows us to formulate a convex minimization problem where we use the inclusion of the solution in the fixed-point set as our regularization. We present a simple provably-convergent iterative technique to solve the resulting problem, which is reminiscent of the PnP algorithm. Moreover, we relate the family of demicontractive operators to other previously made assumptions such as nonexpansiveness and boundedness, showing that the condition of demicontractivity covers a broader range of functions. The presented RED-PRO framework unifies the PnP and RED approaches, offering an increased flexibility in choosing the denoiser, while preserving global convergence guarantees.

• Furthermore, we propose relaxations of RED-PRO that enable the use of denoisers with a narrow fixed-point set at the expense of higher computational load. Results of this strategy on the image deblurring and super-resolution problems show improved performance in comparison to the RED framework.

In addition to the above, we complement the work in [62] by considering the original RED framework where the denoiser is non-differentiable. We formulate the regularization as the Rockafellar function [64], and prove that under certain monotonicity condition on the denoiser, the proposed objective is a convex function, minimized by the RED algorithms. Thus, we provide guarantees for convergence of the RED framework to the globally optimal solution, while relieving the earlier conditions on differentiability, symmetry and homogeneity of the denoiser. As this part deviates from the main theme of this paper, it is brought in Appendix C.

The remainder of the paper is organized as follows. Section 2 details preliminaries of inverse problems and briefly overviews the PnP and RED approaches. In Section 3 we review important definitions and theorems of fixed-point theory that are used throughout the paper. Section 4 serves as the central part of this work. We introduce the RED-PRO framework, which exploits the fixed-point set of denoisers to solve general structured inverse problems. We discuss the relation between the proposed scheme and the PnP approach and show its relation to previous studies. Relaxed variants of RED-PRO are presented, so as to allow handling denoisers with a narrow fixed-point set. Section 5 brings experimental results on image deblurring and super-resolution, demonstrating an improvement of the relaxed RED-PRO framework over the original RED algorithms. Finally, we conclude the paper in Section 6.

2. Preliminaries. In this section we provide the groundwork for this study. First, we describe the framework of inverse problems in image processing, formulated as optimization problems where the challenge is in determining the appropriate regularizer. Then, we discuss PnP and RED, which utilize denoising engines for regularization.
2.1. Inverse Problems. We consider the task of recovering an unknown image $\mathbf{x}$ from its corrupted measurements $\mathbf{y}$. The Bayesian maximum a posteriori (MAP) estimator seeks for a solution that maximizes the posterior conditional probability

\[(2.1) \quad \hat{x}_{\text{MAP}} \triangleq \arg \max_{\mathbf{x} \in \mathbb{R}^n} P(\mathbf{x}|\mathbf{y}). \]

The latter can be simplified using Bayes’s rule as follows:

\[(2.2) \quad \hat{x}_{\text{MAP}} = \arg \max_{\mathbf{x} \in \mathbb{R}^n} \frac{P(\mathbf{y}|\mathbf{x})P(\mathbf{x})}{P(\mathbf{y})} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} -\log P(\mathbf{y}|\mathbf{x}) - \log P(\mathbf{x}), \]

where we omit $P(\mathbf{y})$ since it is not a function of $\mathbf{x}$, and we exploit the monotonic decreasing property of $-\log(\cdot)$ to recast the estimation as a minimization problem. The log-likelihood term, denoted as $\ell(\mathbf{x}; \mathbf{y}) \triangleq -\log P(\mathbf{y}|\mathbf{x})$, describes the probabilistic relationship between the measurements $\mathbf{y}$ and the desired image $\mathbf{x}$, assumed to be known. Typically, the likelihood alone is not sufficient and leads to an ill-posed problem for which the solution is not unique or stable. The probability distribution $P(\mathbf{x})$ leads to a prior, denoted by $\lambda \rho(\mathbf{x}) \triangleq -\log P(\mathbf{x})$, incorporating the statistical nature of the unknown. This regularization term stabilizes and better-conditions the optimization problem. Thus, we can rewrite (2.2) as

\[(2.3) \quad \hat{x}_{\text{MAP}} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} \ell(\mathbf{x}; \mathbf{y}) + \lambda \rho(\mathbf{x}), \]

where $\lambda \geq 0$ represents the level of confidence in the prior. We assume hereafter that the log-likelihood $\ell(\mathbf{x}; \mathbf{y})$ is a convex, differentiable, lower semicontinuous and proper function. A classic model for the measurements, considered throughout this paper, is given by

\[(2.4) \quad \mathbf{y} = \mathbf{Hx} + \mathbf{e}, \]

where $\mathbf{H}$ is a linear degradation operator and $\mathbf{e}$ is a white Gaussian noise (WGN) of variance $\sigma^2$. This leads to

\[(2.5) \quad \ell(\mathbf{x}; \mathbf{y}) = \frac{1}{2\sigma^2} \| \mathbf{Hx} - \mathbf{y} \|_2^2. \]

Note that the noise distribution might be different, e.g., Laplacian, Gamma-distributed, Poisson, and other noise models. In these cases, the expression for the log-likelihood above changes accordingly, differing from the $L_2$-norm.

A challenge that remains is determining the prior to incorporate into the problem formulation. This task of choosing the appropriate $\rho(\mathbf{x})$ has been the center of numerous studies, and various terms have been proposed, ranging from Tikhonov smoothness [37] and the classic Laplacian [46], through wavelet sparsity [54] and total variation [66], to patch-based GMM [89, 84], sparse-representation modeling [12, 35] and recent deep-learning techniques [79]. A possible answer to this challenge may lay in the solution of a special inverse problem, which is image denoising:

\[(2.6) \quad \hat{x}_{\text{Denoise}} = \arg \min_{\mathbf{x} \in \mathbb{R}^n} \frac{1}{2\sigma^2} \| \mathbf{x} - \mathbf{y} \|_2^2 + \lambda \rho(\mathbf{x}). \]
To a large extent, the removal of an additive white Gaussian noise from an image is considered in computational imaging as a solved problem [65]. In the last decade, many extremely effective image denoising algorithms have been proposed, yielding impressive results. In general, the image denoising engine is a function \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) that maps an image \( y \) to a different image of the same size, \( \hat{x} = f(y) \), where ideally \( \hat{x} = x \) – the original noiseless image. Denoising solutions may be based on the MAP estimation (2.6), minimum mean-square-error, collaborative filtering, supervised learning, and more. The success in noise removal has led researchers to exploit the power of denoising engines for solving other problems. In the following, we describe two approaches, PnP [80] and RED [65], which utilize denoisers as implicit and explicit regularization terms respectively, while showing state-of-the-art performance in tasks such as unsupervised image deblurring and image super-resolution.

2.2. Plug and Play Prior (PnP). Here we briefly review the PnP approach [80] by Venkatakrishnan et al., who suggested the use of denoisers as an implicit regularization. We start with problem (2.3) where we assume that \( \rho(\cdot) \) is convex and non-differentiable. The MAP solution can be obtained by two common optimization solvers, PGM and ADMM, summarized in Algorithm 2.1 and Algorithm 2.2 respectively. Note that when \( \alpha_k \equiv 0 \), Algorithm 2.1 reduces to the standard PGM method, while the following update rule

\[
t_0 = 1, \quad t_{k+1} = \frac{1}{2} \left( 1 + \sqrt{1 + 4t_k^2} \right) \text{, } \alpha_k = \frac{t_k - 1}{t_{k+1}}
\]

leads to the accelerated PGM [8].

Algorithm 2.1 PGM/APGM

1: Input: \( z_0 = x_0 \in \mathbb{R}^n \), \( \mu > 0 \) and \( \{\alpha_k\}_{k \in \mathbb{N}} \).
2: for \( k = 0, 1, 2, \ldots \) do:
   • \( v_{k+1} = x_k - \mu \nabla \ell(x_k; y) \)
   • \( z_{k+1} = \arg \min_{z \in \mathbb{R}^n} \frac{1}{2} \| z - v_{k+1} \|_2^2 + \mu \lambda \rho(z) \)
   • \( x_{k+1} = z_{k+1} + \alpha_k (z_{k+1} - z_k) \)
3: Output: \( x_{k+1} \).

Algorithm 2.2 ADMM

1: Input: \( z_0 = x_0 \in \mathbb{R}^n \), \( \beta > 0 \) and \( u_0 \in \mathbb{R}^n \).
2: for \( k = 0, 1, 2, \ldots \) do:
   • \( x_{k+1} = \arg \min_{x \in \mathbb{R}^n} \frac{\beta}{2} \| x - z_k + u_k \|_2^2 + \ell(x; y) \)
   • \( z_{k+1} = \arg \min_{z \in \mathbb{R}^n} \frac{\beta}{2} \| z - x_{k+1} - u_k \|_2^2 + \lambda \rho(z) \)
   • \( u_{k+1} = u_k + x_{k+1} - z_{k+1} \)
3: Output: \( x_{k+1} \).

As can be seen, both techniques make use of the proximal operator, defined as

\[
P_g(x) \triangleq \arg \min_{v \in \mathbb{R}^n} \frac{1}{2} \| x - v \|_2^2 + g(v),
\]
for a closed, proper and convex function $g : \mathbb{R}^n \rightarrow \mathbb{R}$. Notice that (2.7) resembles (2.6), implying that proximal operators are a specific family of denoisers. Thus, the PnP framework offers to extend Algorithms 2.1 and 2.2 by replacing the proximal operators with a general denoiser,\(^1\) $f(\cdot)$, not necessarily variational in nature, i.e., a denoiser which is not originated from any explicit regularizer. The resultant methods are described in Algorithms 2.3 and 2.4 respectively. The original PnP formulation relied on ADMM [80]. However, while the PnP-PGM requires the gradient of $\ell(x; y)$, the PnP-ADMM evaluates the proximal operator of $\ell(x; y)$, which is typically more computationally expensive [72].

**Algorithm 2.3 PnP-PGM/APGM**

1: **Input:** $z_0 = x_0 \in \mathbb{R}^n$, $\mu > 0$, $\{\alpha_k\}_{k \in \mathbb{N}}$ and $f(\cdot)$.
2: **for** $k = 0, 1, 2, \ldots$ **do**:
   - $v_{k+1} = x_k - \mu \nabla \ell(x_k; y)$
   - $z_{k+1} = f(v_{k+1})$
   - $x_{k+1} = z_{k+1} + \alpha_k (z_{k+1} - z_k)$
3: **Output:** $x_{k+1}$.

**Algorithm 2.4 PnP-ADMM**

1: **Input:** $z_0 = x_0 \in \mathbb{R}^n$, $\beta > 0$, $u_0 \in \mathbb{R}^n$ and $f(\cdot)$.
2: **for** $k = 0, 1, 2, \ldots$ **do**:
   - $x_{k+1} = \arg\min_{x \in \mathbb{R}^n} \frac{\beta}{2} \| x - z_k + u_k \|_2^2 + \ell(x; y)$
   - $z_{k+1} = f(x_{k+1} + u_k)$
   - $u_{k+1} = u_k + x_{k+1} - v_{k+1}$
3: **Output:** $x_{k+1}$.

Empirically, incorporating powerful denoisers (such as BM3D, TNRD and DnCNN) into the PnP framework has led to state-of-the-art results in various inverse problems. However, when the denoiser used does not correspond to an explicit regularization $\rho(x)$, the PnP methods cannot be interpreted as optimization solvers in general, making it difficult to theoretically investigate the stability and uniqueness of their solutions. Several studies[72, 18, 15, 76, 17, 77, 31, 78, 67] proved the convergence of PnP methods to a fixed-point under different conditions, while the work reported in [70] states clear conditions for a global convergence of PnP. Yet, none of these provide an explicit expression of an objective function which is minimized. In Section 4.1 we remedy this by introducing optimization techniques reminiscent of the PnP methods, this way offering a novel theoretical explanation for Algorithms 2.1 and 2.2.

2.3. Regularization by Denoising (RED). As discussed above, the PnP framework has been the first to exploit denoisers for an implicit regularization, while lacking an underlying objective function. To overcome this shortcoming, Romano et al. [65] introduced a different,\(^1\)

\(^1\)We use hereafter the function $f(\cdot)$ to denote a general denoiser, and we omit its dependency on the noise-level (or the denoising strength to employ). However, we should note that this parameter should be chosen carefully, as it can be critical for the performance of any of the algorithms described.
yet related, strategy that harnesses image denoisers, called *RE*gularization *by Denoising*. The RED framework defines the following regularization term:

\[
\rho_{\text{RED}}(x) \triangleq \frac{1}{2} \langle x, x - f(x) \rangle.
\]

This prior is an image-adaptive Laplacian whose definition is based on the denoiser of choice, \( f(\cdot) \). Thus, the overall optimization problem to solve is

\[
\hat{x}_{\text{RED}} = \arg\min_{x \in \mathbb{R}^n} \ell(x; y) + \frac{\lambda}{2} \langle x, x - f(x) \rangle.
\]

The denoiser \( f(\cdot) \) is assumed to obey the following assumptions, which we refer to hereafter as the RED conditions

(C1) **Local Homogeneity**: For all \( x \in \mathbb{R}^n \) it holds that

\[
f((1 + \epsilon)x) = (1 + \epsilon)f(x),
\]

for sufficiently small \( \epsilon \in \mathbb{R} \setminus 0 \).

(C2) **Differentiability**: The denoiser \( f(\cdot) \) is differentiable where \( \nabla f \) denotes its Jacobian.

(C3) **Jacobian Symmetry**: The Jacobian \( \nabla f(x) \) is symmetric

\[
\nabla f(x)^T = \nabla f(x), \ \forall x \in \mathbb{R}^n.
\]

(C4) **Stability/Strong Passivity**: The Jacobian \( \nabla f(x) \) satisfies

\[
\eta(\nabla f(x)) \leq 1,
\]

where \( \eta(A) \) denotes the spectral radius of the matrix \( A \).

Interestingly, under these conditions, the regularization term (2.8) is differentiable, convex, and its gradient is given by the denoising residual \( x - f(x) \). Furthermore, denoting by \( E_{\text{RED}}(x) \) the objective function of (2.9), \( E_{\text{RED}}(x) \) is convex whenever the log-likelihood is convex, and it gradient is given simply as

\[
\nabla E_{\text{RED}}(x) = \nabla \ell(x; y) + \lambda \left( x - f(x) \right).
\]

Based on this expression, Romano et al. have proposed several iterative algorithms – steepest descent, fixed-point iteration and ADMM, which are guaranteed to converge to the global optimum of (2.9). Hence, in that aspect, RED framework seems to be superior to the PnP approach.

In a later work [62], the authors have argued that many popular denoisers lack symmetric Jacobian (C3), making the gradient expression (2.10) invalid. Moreover, they have proven that when the denoiser \( f(x) \) fails to satisfy condition (C3), there is no regularizer \( \rho(x) \) whose gradient is the denoising residual \( x - f(x) \). Yet, in practice, the RED algorithms converge and have demonstrated state-of-the-art results in super-resolution and image deblurring even when used with denoisers that are not differentiable, let alone exhibit symmetric Jacobian. Therefore, the theoretical justification of the RED approach remains an open question, which we aim to address in Section C.

\[\text{2Indeed, [62] was the first work to draw attention to the need for symmetry of the Jacobian.}]


3. Review of Fixed-Point Theory. Here we review key concepts of fixed-point theory on which we base our contributions. We start by considering a nonlinear mapping \( T : \mathbb{R}^n \to \mathbb{R}^n \), and define the point \( x \in \mathbb{R}^n \) to be a fixed-point of \( T \) iff \( T(x) = x \). Denote by \( \text{Fix}(T) \) the set of all fixed-points of \( T \), defined as

\[
\text{Fix}(T) \triangleq \{ x \in \mathbb{R}^n : T(x) = x \}.
\]

Throughout the paper we assume that \( \text{Fix}(T) \) is nonempty. Our study focuses on the set \( \text{Fix}(T) \) and its favorable properties for a certain family of mappings defined next.

**Definition 3.1 (Demiclosedness).** The mapping \( T \) is said to be demiclosed at 0, if for any sequence \( \{ x_i \} \) that satisfies \( x_i \to x \) and \( x_i - T(x_i) \to 0 \), we have \( x = T(x) \).

**Definition 3.2 (Demicontractive).** The mapping \( T \) is demicontractive with a constant \( d \in [0,1) \) (or \( d \)-demicontractive) if for any \( x \in \mathbb{R}^n \) and \( z \in \text{Fix}(T) \) it holds that

\[
\| T(x) - z \| \leq \| x - z \|^2 + d \| T(x) - x \|^2,
\]

or equivalently

\[
\frac{1 - d}{2} \| x - T(x) \|^2 \leq \langle x - T(x), x - z \rangle.
\]

**Definition 3.3.** The mapping \( T \) is quasi-nonexpansive if

\[
\| T(x) - z \| \leq \| x - z \|, \forall x \in \mathbb{R}^n, z \in \text{Fix}(T).
\]

We say \( T \) is \( \gamma \)-strongly quasi-nonexpansive with \( \gamma \geq 0 \) if for all \( x \in \mathbb{R}^n \) and all \( z \in \text{Fix}(T) \)

\[
\| T(x) - z \|^2 \leq \| x - z \|^2 - \gamma \| T(x) - x \|^2.
\]

**Definition 3.4.** The mapping \( T \) is called nonexpansive if

\[
\| T(x) - T(z) \| \leq \| x - z \|, \forall x, z \in \mathbb{R}^n.
\]

We say \( T \) is cyclically firmly nonexpansive if for every set of points \( \{ x_1, ..., x_m \} \) in \( \mathbb{R}^n \) and \( m \geq 2 \) it holds that

\[
\sum_{i=1}^{m} \langle x_i - T(x_i), T(x_i) - T(x_{i+1}) \rangle \geq 0,
\]

where \( x_{m+1} = x_1 \).

**Definition 3.5.** The mapping \( T \) is called Lipschitz continuous with a constant \( L > 0 \) (or \( L \)-Lipschitzian) if

\[
\| T(x) - T(z) \| \leq L \| x - z \|, \forall x, z \in \mathbb{R}^n.
\]

When \( L < 1 \), \( T \) is called a contraction.
Figure 3.1. The family of demicontractive mappings and its subclasses.

Notice that any Lipschitzian mapping admits a nonexpansive function by an appropriate scaling. In addition, as shown in Fig. 3.1, the class of demicontractive mappings includes the class of quasi-nonexpansive mappings, which in turn contains the widely studied class of nonexpansive operators. Therefore, demicontractive operators cover a large extent of functions, explaining their centrality in this work.

Below we study the properties of this family of functions where we are particularly interested in the closedness and convexity of the fixed-point set. To that end, we present the following definitions and theorems, starting with the next useful lemma.

Lemma 3.6. Consider a sequence \( \{x_i\} \) in \( \mathbb{R}^n \) and let \( \{\alpha_i\} \) be a sequence in \( \mathbb{R} \) such that \( \sum_i \alpha_i = 1 \). Then, we have

\[
\left\| \sum_i \alpha_i x_i \right\|_2^2 = \sum_i \alpha_i \|x_i\|_2^2 - \frac{1}{2} \sum_i \sum_j \alpha_i \alpha_j \|x_i - x_j\|_2^2.
\]

In particular, for any \( x, z \in \mathbb{R}^n \) and \( \alpha \in \mathbb{R} \) we get

\[
\|\alpha x + (1 - \alpha) z\|_2^2 = \alpha \|x\|_2^2 + (1 - \alpha) \|z\|_2^2 - \alpha(1 - \alpha) \|x - z\|_2^2.
\]

Proof. See Lemma 2.13 in [5].

The above lemma plays a major role in the proofs of the following theorems related to important properties of demicontractive mappings.

Definition 3.7. Consider a mapping \( T \) and let \( \alpha \in (0, 1) \). We define a relaxed (or averaged [4]) operator as \( T_\alpha = \alpha T + (1 - \alpha) \text{Id} \) where \( \text{Id} \) is the identity operator. Notice it holds that \( \text{Fix}(T_\alpha) = \text{Fix}(T) \).

Theorem 3.8. Consider a \( d \)-demicontractive mapping \( T \). Let \( \alpha \in (0, 1 - d] \) and define the relaxed operator \( T_\alpha = \alpha T + (1 - \alpha) \text{Id} \). Then, \( T_\alpha \) is \( \gamma \)-strongly quasi-nonexpansive where \( \gamma = \frac{(1 - d - \alpha)}{\alpha} \).

\(^3\)A reasonable assumption is that \( f(0) = 0 \).
Proof. Given arbitrary \( x \in \mathbb{R}^n \) and \( z \in \text{Fix}(T) \), we have

\[
\| T_\alpha(x) - z \|^2 = \| (x - z) + \alpha(T(x) - x) \|^2 \\
= \| x - z \|^2 - 2\alpha \left( \langle x - z, x - T(x) \rangle + \alpha^2 \| x - T(x) \|^2 \right).
\]

From (3.3), we get

\[
\| T_\alpha(x) - z \|^2 \leq \| (x - z) \|^2 - \alpha(1 - \alpha) \| T(x) - x \|^2.
\]

Note that \( \| T_\alpha(x) - x \| = \alpha \| T(x) - x \| \), implying that

\[
\| T_\alpha(x) - z \|^2 \leq \| (x - z) \|^2 - \left( \frac{1 - \alpha}{\alpha} \right) \| T_\alpha(x) - x \|^2,
\]

which completes the proof.

Theorem 3.9. Let \( T \) be a \( \gamma \)-strongly quasi-nonexpansive mapping with \( \gamma \geq 1 \). Then, we have

\[
(3.11) \quad \| x - T(x) \| \leq \| x - P_{\text{Fix}(T)}(x) \|,
\]

where \( P_{\text{Fix}(T)} \) represents the projection onto the fixed-point set.

Proof. See [16].

Theorem 3.8 provides a simple way to construct a strong quasi-nonexpansive mapping from a demicontractive one, which we utilize in the next section.

Now, we present the main theorem of this part.

Theorem 3.10. Consider a \( d \)-demicontractive mapping \( T \). Then, the fixed-point set \( \text{Fix}(T) \) is closed and convex [23].

Proof. See Appendix A.

As the class of demicontractive mappings includes quasi-nonexpansive and nonexpansive operators, the result of Theorem 3.10 holds for a broad range of functions. This result is the foundation on which we shall base our problem formulation, as introduced in the following section.

Before we conclude this review, we briefly bring preliminaries of monotone operators and their relation to nonexpansive mappings. This connection shall allow us to establish the convergence of RED algorithms later on.

Definition 3.11. The mapping \( A \) is called monotone when

\[
(3.12) \quad \langle x - z, A(x) - A(z) \rangle \geq 0, \ \forall x, z \in \mathbb{R}^n.
\]

We say that \( A \) is maximal monotone if it is monotone and it is contained by no other monotone operator.
Definition 3.12. The mapping $A$ is called $\tau$-strongly monotone if it satisfies
\[
\langle x - z, A(x) - A(z) \rangle \geq \tau \|x - z\|^2, \quad \forall x, z \in \mathbb{R}^n.
\]

Definition 3.13. The mapping $A$ is cyclically monotone if for every set of points $\{x_1, \ldots, x_m\}$ in $\mathbb{R}^n$ and arbitrary $m \geq 2$ it holds that
\[
\sum_{i=1}^{m} \langle A(x_i), x_{i+1} - x_i \rangle \leq 0,
\]
where $x_{m+1} = x_1$. We say that $A$ is maximal cyclically monotone if it is cyclically monotone and its graph cannot be enlarged without destroying this property.

Equipped with the above definitions, we now describe the connection between nonexpansiveness and monotonicity as given by the next theorems.

**Theorem 3.14.** Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be nonexpansive and define the corresponding displacement mapping by $A = \text{Id} - T$. Then, $A$ is maximal monotone.

**Proof.** See [64]

**Theorem 3.15.** Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be cyclically firmly nonexpansive and consider the displacement mapping $A = \text{Id} - T$. Then, $A$ is maximal cyclically monotone.

**Proof.** See Appendix B.

### 4. RED-PRO: Fixed-Point Projection.

As discussed in Section 2, the PnP and RED frameworks have achieved state-of-the-art performance in solving inverse problems by utilizing denoisers as regularization. However, their theoretical analysis is incomplete since it is unclear which objective functions are minimized by PnP and RED, and whether such cost functions exist. As a partial answer, we consider in Appendix C non-differentiable denoisers and we provide convergence guarantees for the RED algorithms. In this section we address the matter of the underlying objective function for PnP and RED. To that end, we reformulate RED as a convex minimization problem regularized using the fixed-point set of a demicontractive denoiser. We provide simple solutions for the proposed problem, similar to PnP-PGM and PnP-ADMM. As such, we offer a theoretical explanation for the PnP approach and establish its connection to RED. Finally, we relax our problem by replacing the hard constraint with a distance function and by considering broader and richer domains than the fixed-point set. These modifications offer larger flexibility, allowing for a broader group of denoisers to be applicable.

#### 4.1. Fixed-Point Projection.

As discussed earlier, several previous works rely on the nonexpansivity of the denoiser for proving the fixed-point convergence of the RED algorithms. However, the latter requirement may be too restrictive since several denoisers, e.g. NLM, are expansive [18]. To overcome the above limitation, we require henceforth a considerably weaker assumption where the denoiser $f(\cdot)$ is a $d$-demicontractive mapping for some $d \in [0, 1)$.

We start with the following observation which serves as the motivation for our regularization strategy shown later.

---

4 We are referring to the case where the denoiser is non-differentiable or having a non-symmetric Jacobian.
Proposition 4.1. Consider a demicontractive denoiser \( f(\cdot) \) and assume \( f(0) = 0 \). Then,
\[
\rho_{\text{RED}}(x) = \frac{1}{2} \langle x, x - f(x) \rangle = 0 \iff x \in \text{Fix}(f).
\]

Proof. It is clear that for any \( x \in \text{Fix}(f) \), we get \( \rho_{\text{RED}}(x) = 0 \). For the other direction, we recall that by the definition of demicontractive mapping for any \( x \in \mathbb{R}^n \)
\[
0 \leq \frac{1}{2} - d \|x - f(x)\|^2 \leq \langle x, x - f(x) \rangle,
\]
where we use the assumption that \( x = 0 \) is a fixed-point. Therefore, when \( \rho_{\text{RED}}(x) = 0 \), then \( \|x - f(x)\|^2 = 0 \), suggesting that \( x \in \text{Fix}(f) \).

Inspired by the RED framework and the theorem above, we introduce the following minimization problem
\[
\hat{x}_{\text{RED-PRO}} = \arg \min_{x \in \mathbb{R}^n} \ell(x; y) \quad \text{s.t.} \quad x \in \text{Fix}(f).
\]

We refer to the above as the RED via Fixed-Point Projection (RED-PRO) paradigm, where we utilize the fixed-point set of a denoising engine as a regularization for our inverse problem. The optimization task (4.2) can be interpreted as searching for a minimizer of \( \ell(x; y) \) over the set of ”clean” images. Surprisingly, the RED-PRO approach admits a convex optimization problem as stated by the next theorem.

Theorem 4.2. Assume the denoiser \( f(\cdot) \) is a \( d \)-demicontractive mapping. Then, (4.2) defines a convex minimization problem.

Proof. By Theorem 3.10, the set \( \text{Fix}(f) \) is closed and convex, hence, the convexity of log-likelihood implies problem (4.2) is a minimization of a convex function over a convex domain.

Note that when \( \ell(x; y) \) is strongly convex, there exist a unique solution for problem (4.2). In addition, as shown in Section 3, for any \( \alpha \in (0, 1) \) it holds that \( \text{Fix}(f_\alpha) = \text{Fix}(f) \) where \( f_\alpha = \alpha f + (1 - \alpha)\text{Id} \) is the relaxed operator. Hence, we can rewrite (4.2) equivalently as
\[
\hat{x}_{\text{RED-PRO}} = \arg \min_{x \in \mathbb{R}^n} \ell(x; y) \quad \text{s.t.} \quad x \in \text{Fix}(f_\alpha).
\]

To find a solution for our proposed problem (4.2), one may apply projected gradient descent [59, 9] with the update rule
\[
x_{k+1} = P_{\text{Fix}(f)} \left( x_k - \mu \nabla \ell(x; y) \right),
\]
where \( \mu > 0 \) is the gradient step size and \( P_{\text{Fix}(f)} \) denotes the projection onto the fixed-point set \( \text{Fix}(f) \). However, we offer a simpler solution which performs one activation of the denoising per iteration. The proposed technique, detailed in Algorithm 4.1, is based on the hybrid steepest descent method (HSD [4, 83]) which we modify for denoisers.

The next theorem provides the conditions for convergence of Algorithm 4.1 to an optimal solution of (4.2).
Algorithm 4.1 RED-PRO HSD

1: Input: $x_0 \in \mathbb{R}^n$, $\{\mu_k\}_{k \in \mathbb{N}}$, $\alpha \in (0, 1)$ and $f(\cdot)$.
2: for $k = 0, 1, 2, \ldots$ do:
   • $v_{k+1} = x_k - \mu_k \nabla \ell(x_k; y)$
   • $z_{k+1} = f(v_{k+1})$
   • $x_{k+1} = \alpha z_{k+1} + (1 - \alpha) v_{k+1}$
3: Output: $x_{k+1}$.

Theorem 4.3. Let $f(\cdot)$ be a $d$-demicontractive denoiser and $\ell(\cdot; y)$ be a convex, differentiable and lower-bounded function. Assume the following:
1. $f(\cdot)$ is demiclosed at 0.
2. $\nabla \ell(\cdot; y)$ is $L$-Lipschitzian and $\tau$-strongly monotone for some $L, \tau > 0$.
3. $\alpha \in (0, \frac{1 - d}{2}]$.
4. $\{\mu_k\}_{k \in \mathbb{N}}$ is non-increasing in $[0, 1)$ sequence such that
   \[
   \lim_{k \to \infty} \mu_k = \infty, \sum_{k \in \mathbb{N}} \mu_k = \infty.
   \]

Then, there is a unique solution $x^*$ for problem (4.2) and the sequence $\{x_k\}_{k \in \mathbb{N}}$ generated by Algorithm 4.1 converges to it.

Proof. See Theorem 3.4 in [52].

Notice that the update rules of Algorithm 4.1 can be simplified to the following form
\[
(4.5) \quad x_{k+1} = f_\alpha \left(x_k - \mu_k \nabla \ell(x_k; y)\right).
\]

Thus, Algorithm 4.1 can be seen as the PnP-PGM where we use a diminishing gradient step size and the denoiser is $d$-demicontractive function. Hence, Theorem 4.3 leads to the important conclusion that the RED-PRO formulation given by (4.2) is the underlying optimization problem on which PnP-PGM is based. Furthermore, it has been shown [57, 72, 67] that other PnP variants, e.g., PnP-ADMM and PnP primal-dual hybrid gradient method (PnP-PDHG), satisfy the same fixed-point equation as PnP-PGM
\[
(4.6) \quad x^* = f_\alpha \left(x^* - \mu \nabla \ell(x^*; y)\right).
\]

Therefore, under the assumptions stated in Theorem 4.3, any algorithm that converges to a fixed-point satisfying (4.6), leads in principle to a solution of problem (4.2). However, in contrast to ADMM-based PnP techniques, the proposed algorithm (4.1) relies on the gradient $\nabla \ell(x; y)$ of $\ell(x; y)$ rather than the proximal operator $P_\ell(x)$ which is typically more computationally expensive.

4.1.1. Demicontractive Functions. A question that remains is whether PnP algorithms, other than PnP-PGM, converge under the conditions of Theorem 4.3. Thus, we now discuss in detail the conditions on the denoiser for facilitating such a result.
Any denoiser, reasonably assumed to be lower semicontinuous function, is demiclosed at 0. As for demicontractivity, various studies [70, 15, 72, 77, 76, 17, 62] have proven the convergence of variants of PnP-ADMM for nonexpansive or averaged nonexpansive denoisers, which are in particular demicontractive. Recently, the authors of [67] have shown that PnP-ADMM and PnP forward-backward splitting converge under weaker conditions where \( f \) is assumed to be an averaged operator satisfying

\[
\| f(x) - f(z) \| \leq (1 + \epsilon) \| x - z \|, \quad \forall x, z \in \mathbb{R}^n,
\]

for some \( \epsilon > 0 \). With respect to the fixed-points \( z \in \text{Fix}(f) \), the above condition translates to the following assumption

\[
\| f(x) - z \| \leq (1 + \epsilon) \| x - z \|, \quad \forall x \in \mathbb{R}^n, z \in \text{Fix}(f).
\]

Notice that by the definition of demicontractivity (3.3) in conjunction with the Cauchy-Schwartz inequality we obtain that any \( d \)-demicontractive function \( f(\cdot) \) satisfies

\[
\| x - f(x) \| \leq \frac{2}{1 - d} \| x - z \|,
\]

which it turn implies that

\[
\| f(x) - z \| \leq \left( 1 + \frac{2d}{1 - d} \right) \| x - z \|, \quad \forall x \in \mathbb{R}^n, z \in \text{Fix}(f).
\]

Thus, it is clear that any \( d \)-demicontractive function meets condition (4.8) with \( \epsilon \triangleq \frac{2d}{1 - d} \).

In [72], Chan et al. have proposed a version of PnP-ADMM for bounded denoisers satisfying

\[
\frac{1}{n} \| f_\sigma(x) - x \|^2 \leq \sigma^2 c,
\]

where we assume any point \( x \) is bounded in some interval \( x \in [a, b]^n \) \( (a < b) \), \( \sigma > 0 \) is a parameter controlling the strength of the denoiser and \( c > 0 \) is a constant independent of \( n \) and \( \sigma \). Bounded denoisers are asymptotically invariant in the sense that \( f_\sigma \to \text{Id} \) as \( \sigma \to 0 \). Under the boundedness assumption (4.10), a fixed-point convergence of the modified PnP-ADMM has been proven where a diminishing step size \( \mu_k \) has been used and \( \sigma_k^2 = \lambda \mu_k \) for some predefined \( \lambda > 0 \). However, this approach requires the denoiser to have an internal parameter \( \sigma^2 \) controlling its strength which may not be available. We offer an external control which relies only on the demicontractivity of the denoiser, as given the in the following theorem.

**Theorem 4.4.** Consider a \( d \)-demicontractive denoiser \( f(\cdot) \) and define the relaxed operator \( f_\alpha(\cdot) \) for some \( \alpha \in (0, 1 - d) \). Then, it holds that

\[
\frac{1}{n} \| f_\alpha(x) - x \|^2 \leq \sigma^2(\alpha) c,
\]

where \( \sigma^2(\alpha) \triangleq \frac{\alpha}{1 - d - \alpha} \) and \( c \triangleq (b - a)^2 \).

5 The original formulation in [72] defines an increasing penalty parameter \( \beta_k \) which satisfies \( \beta_k = \frac{1}{\mu_k} \).
Proof. First, for any bounded pair of points $x, z \in [a, b]^n$, we have

$$\frac{1}{n} \|x - z\|^2 \leq (b - a)^2.$$ 

In addition, by Theorem 3.8 we obtain

$$\|f_\alpha(x) - x\|^2 \leq \frac{\alpha}{1 - d - \alpha} \|x - z\|^2, \forall x \in [a, b]^n, z \in \text{Fix}(f).$$

Hence, we get

$$\frac{1}{n} \|f_\alpha(x) - x\|^2 \leq \frac{1}{n} \cdot \frac{\alpha}{1 - d - \alpha} \|x - z\|^2 \leq \frac{\alpha}{1 - d - \alpha} (b - a)^2,$$

which completes the proof.

The above theorem allows to externally control the strength of a demicontractive denoiser by performing appropriate averaging. Note that $\sigma^2(\alpha) \to 0$ and $f_\alpha(x) \to \text{Id}$ as $\alpha \to 0$. Moreover, by updating $\alpha_k = \frac{\lambda_k}{1 + \lambda_k}(1 - d) \in (0, 1 - d)$ at each iteration we obtain $\sigma^2_k \triangleq \sigma^2(\alpha_k) = \lambda_k$, which ensures the convergence of the PnP-ADMM presented in [72] for demicontractive denoisers.

Next, recalling the original RED framework, notice that any denoiser $f(\cdot)$, which fulfills the RED conditions given in Section 2.3, satisfies

$$\|f(x) - f(z)\| = \left\| \int_0^1 \nabla f\left(z + t(x - z)\right) dt (x - z) \right\|$$

$$\leq \left\| \int_0^1 \nabla f\left(z + t(x - z)\right) dt \right\| \cdot \|(x - z)\|$$

$$\leq \int_0^1 \|\nabla f\left(z + t(x - z)\right)\| dt \cdot \|(x - z)\|$$

$$\leq \int_0^1 1 dt \cdot \|(x - z)\| = \|x - z\|, \forall x, z \in \mathbb{R}^n.$$ 

where $\nabla f$ is the Jacobian of the denoiser and the first equality is obtained by the generalized fundamental theorem of calculus. This implies that under the RED conditions, the denoiser $f(\cdot)$ is nonexpansive, hence, it is demicontractive.

As a final note, we remark that in general there is no efficient way for verifying our assumption of demicontractivity nor evaluating $d$ since it is unclear what is the fixed-point set of arbitrary denoisers. Yet, as shown, the family of demicontractive mappings covers a broad area of functions denoisers and it can be directly related to previously made assumptions. Therefore, we conclude the RED-PRO framework offers a unified approach for PnP and RED which can utilize various denoisers and is guaranteed to converge to global solutions of convex optimization.
4.2. Relaxed RED-PRO. In the previous part we introduced the RED-PRO framework which utilizes the fixed-point set of a denoiser as a regularization, and we established a connection to the PnP and the original RED approaches, including convergence guarantees. However, depending on the denoiser in use, the fixed-point set might be small, thus, making the regularization term in (4.2) too restrictive. To overcome the above limitation, we present here two relaxations of problem (4.2).

4.2.1. Approximate via Distance to the Fixed Set. First, we weaken our hard constraint by replacing it with a squared distance function. For a closed and convex set $C$, the Euclidean distance \cite{11, 7} of a point $x \in \mathbb{R}^n$ from $C$ is defined as

\begin{equation}
\label{eq:4.12}
d_C(x) \triangleq \min_{v \in C} \|x - v\| = \|x - P_C(x)\|,
\end{equation}

where $P_C(x)$ is the projection of a point $x \in \mathbb{R}^n$ onto $C$. The squared distance function $D_C(\cdot) \triangleq \frac{1}{2}d^2_C(\cdot)$ is known to be convex and differentiable function whose gradient is

\begin{equation}
\label{eq:4.13}
\nabla D_C(\cdot) = Id - P_C(\cdot).
\end{equation}

Following the above, we propose the relaxed RED-PRO (RRP) minimization problem

\begin{equation}
\label{eq:4.14}
\hat{x}_{RRP} = \arg \min_{x \in \mathbb{R}^n} \ell(x; y) + \lambda \left\| x - P_{\text{Fix}(f)}(x) \right\|^2,
\end{equation}

where $P_{\text{Fix}(f)}(x)$ denotes the projection onto the fixed-point set of the denoiser $f(x)$. When the denoiser is demicontractive, the fixed-point set is closed and convex, hence, problem (4.14) is a well-defined convex minimization. Here $\lambda > 0$ balances between the log-likelihood term and the regularization and allows us to control the distance of the optimal solution from the fixed-point set. Notice that when $\lambda \to \infty$ (or it sufficiently large), problem (4.14) reduces to problem (4.2).

To solve (4.14), we exploit the fact that the gradient of the objective function is simply

\begin{equation}
\label{eq:4.15}
\nabla E_{RRP}(x) = \nabla \ell(x; y) + \lambda \left(x - P_{\text{Fix}(f)}(x)\right).
\end{equation}

Thus, any global minimizer $x^*$ of (4.14) should satisfy the first-order optimality condition

\begin{equation}
\label{eq:4.16}
\nabla \ell(x^*; y) + \lambda \left(x^* - P_{\text{Fix}(f)}(x^*)\right) = 0.
\end{equation}

Notice that the latter equation is similar to the original RED optimality condition (see equation (C.2)), where we replace the denoiser with the projection onto its fixed-point set. Therefore, the solution of (4.14) can be found by various optimization solvers, e.g., the SD method, ADMM and FP strategy used in \cite{65}.

As can be seen from (4.16), the gradient expression requires computing the projection onto the fixed-point set, which can be written as

\begin{equation}
\label{eq:4.17}
P_{\text{Fix}(f)}(x) = \arg \min_{v \in \mathbb{R}^n} \frac{1}{2} \|x - v\|^2 \text{ s.t. } v \in \text{Fix}(f).
\end{equation}
The above problem is a special case of problem (4.2) where the objective is
\[ h(x) = \frac{1}{2} \| x - v \|^2. \]
The latter is a differentiable function whose gradient \( \nabla h(x) = x - v \) is clearly \( L \)-Lipschitzian
and \( \tau \)-strongly monotone with \( L = \tau = 1 \). Therefore, according to Theorem 4.3, we can
compute the projection \( P_{\text{Fix}(f)}(x_0) \) of a given point \( x_0 \) using the following iterative scheme
\[ x_{j+1} = f_\alpha \left( t_j x_0 + (1 - t_j) x_j \right), \]
where \( \alpha \in (0, \frac{1-d}{2}) \) and \( \{t_k\}_{k \in \mathbb{N}} \) is a non-increasing in \([0, 1)\) sequence such that \( t_k \to 0 \)
and \( \sum_{k \in \mathbb{N}} t_k = \infty \). Notice the above update rule can be rewritten equivalently as
\[ x_{j+1} = t_j x_0 + (1 - t_j) f_\alpha(x_j), \]
which is the well-known Halpern iteration \([41, 23, 55, 51]\). This of course might require
numerous activations of the denoising engine in each iteration, hence, in practice we perform
only a small number of iterations to reduce computationally load, yielding an approximation
of the projection. As an example, we outline in Algorithm 4.2 the overall method for solving
(4.14) via SD.

**Algorithm 4.2 RED-PRO via SD**

1: Input: \( x_0 \in \mathbb{R}^n \), \( \{t_k\}_{k \in \mathbb{N}} \), \( \alpha \in (0, 1) \), \( \lambda, \mu, J > 0 \), \( f(\cdot) \).
2: for \( k = 0, 1, 2, ... \) do:
   for \( j = 0, 1, 2, ... J \) do:
      \[ x_{k,j+1} = f_\alpha \left( t_j x_k + (1 - t_j) x_{k,j} \right) \]
      \[ x_{k+1} = x_k - \mu \left( \nabla \ell(x_k; y) + \lambda (x_k - x_{k,J+1}) \right) \]
3: Output: \( x_{k+1} \).

Notice that RED-PRO via SD requires \( J \) activation of the denoising engine per gradient
evaluation due to projection operation. Therefore, the RED-PRO approach given by (4.14)
offers a broader search domain at the expense of increased computational load.

**4.2.2. Approximate Fixed-Points.** Here, we further relax our problem by considering
richer sets than the fixed-point one. To that end, we define for \( \epsilon > 0 \) the \( \epsilon \)-approximate
fixed-points as
\[ \text{Fix}_\epsilon(f) \triangleq \{ x \in \mathbb{R}^n : \| x - f(x) \| \leq \epsilon \}. \]
Note that \( \text{Fix}_\epsilon(f) \) can be seen as the set of ”almost clean” images, thus, it can act as a valid
and wide domain for searching restored solutions. A possible approach to exploit \( \text{Fix}_\epsilon(f) \) is
defining an adaptive averaged denoiser
\[ f_\epsilon(x) \triangleq \alpha_\epsilon(x) x + \left( 1 - \alpha_\epsilon(x) \right) f(x), \]
where \( \alpha(x) = \frac{\epsilon}{\max(\epsilon, \|x - f(x)\|)} \). Now notice that \( \text{Fix}(f_{\epsilon}) = \text{Fix}_{\epsilon}(f) \), hence, the modified denoiser (4.21) can be utilized via the algorithms derived earlier to find inverse solutions constrained by \( \text{Fix}_{\epsilon}(f) \). While empirically this may lead to satisfactory results, the approximate fixed-point set is generally not convex for demicontractive or even nonexpansive functions. Consequently, the projection onto \( \text{Fix}_{\epsilon}(f) \) is not well-defined and no convergence guarantees can be given. One option standing ahead of us is to restrict the valid denoisers to those for which this set if necessarily convex.

Another option for circumventing the above difficulty is to consider a dilated fixed-point set, defined for some \( \delta > 0 \) as

\[
B_{\delta}(f) \triangleq \{ x \in \mathbb{R}^n : \| x - P_{\text{Fix}(f)}(x) \| \leq \delta \}.
\]

Note that \( B_{\delta}(f) \) is closed and convex for any demicontractive denoiser. Moreover, as illustrated in Fig. 4.1, for an appropriate choice of \( \delta \) we have that \( B_{\delta}(f) \subseteq \text{Fix}_{\epsilon}(f) \) as stated by the following.

**Theorem 4.5.** Let \( f(\cdot) \) be \( d \)-demicontractive function and consider \( \delta \in [0, \alpha \epsilon] \) for some \( \epsilon > 0 \) and \( \alpha \in (0, \frac{1-d^2}{2}] \). Then,

\[
B_{\delta}(f) \subseteq \text{Fix}_{\epsilon}(f).
\]

**Proof.** See Appendix D.

We exploit the result of the latter theorem to formulate the next convex problem

\[
\hat{x}_{\epsilon} = \arg \min_{x \in \mathbb{R}^n} \ell(x; y) + \frac{\lambda}{2} \| x - P_{B_{\delta}(f)}(x) \|^2,
\]

where \( P_{B_{\delta}(f)} \) is the projection onto \( B_{\delta}(f) \) for \( \delta = \frac{1-d^2}{2} \epsilon \). The latter minimization problem consists of another parameter \( \epsilon \) which allows to control the distance of the solution from...
the fixed-point set, thus, increasing our search domain. When $\epsilon$ is set to zero, we return to problem (4.14).

Similar to (4.15), the gradient of the objective function is

\[
\nabla \ell(x; y) + \lambda \left( x - P_{B_\delta(f)}(x) \right),
\]

and it can be used to find the solution of (4.23) via various solvers. To that end, we need to compute the projection $P_{B_\delta(f)}$ whose expression for any $x \notin B_\delta(f)$ is given by [2]

\[
P_{B_\delta(f)}(x) = \frac{x - \delta}{\|x - P_{Fix(f)}(x)\|} P_{Fix(f)}(x).
\]

Thus, we can apply the Halpern iteration to compute the gradient and minimize problem (4.23) using different gradient-based solvers. In Algorithm 4.3 we detail the overall solution via SD as an example.

**Algorithm 4.3 Approximate RED-PRO with via SD**

1: **Input:** $x_0 \in \mathbb{R}^n$, $\{t_k\}_{k \in \mathbb{N}}$, $\alpha \in (0, 1)$, $\lambda, \mu, J > 0$, $f(\cdot)$ $d$-demiconttractive, $\epsilon > 0$ and $\delta = \frac{1 - d^2}{2\epsilon}$.

2: **for** $k = 0, 1, 2, \ldots$ **do:**
   **for** $j = 0, 1, 2, \ldots J$ **do:**
   \[
x_{k+1} = f_{\alpha} \left( t_j x_k + (1 - t_j) x_{k,j} \right)
   \]
   \[
v_k = \frac{x_k - \delta}{\|x_k - P_{Fix(f)}(x)\|} x_k + \left( 1 - \frac{x_k - \delta}{\|x_k - P_{Fix(f)}(x)\|} \right) x_{k,j+1}.
   \]
   \[
x_{k+1} = x_k - \mu \left( \nabla \ell(x_k; y) + \lambda (x_k - v_k) \right).
   \]

3: **Output:** $x_{k+1}$.

To summarize this part, we described here relaxed versions of the RED-PRO framework where we use the distance function as regularization and we extend our search to certain approximate fixed-points. This allows us to obtain stable inverse solutions even when the denoiser exhibits a limited fixed-point set.

**5. Experiments.** In this section we evaluate the performance of the proposed RED-PRO framework. Since the goal of this study is to unify RED and PnP and to provide theoretical justifications rather than achieving state-of-the-art results, we compare ourselves to the original RED approach alone, for the tasks of image deblurring and super-resolution. We follow the same line of experiments performed in [65], where we employ NLM and TNRD denoisers, rather than the median filter and TNRD. We use NLM to show the application of the RED-PRO with a denoiser that does not meet the original RED conditions. The TNRD denoiser, pretrained to remove WGN whose standard deviation equals 5, is brought to demonstrate the full extent of the proposed framework.

---

\[\text{We use TNRD to address any arbitrary noise level } \sigma \text{ by exploiting the relation } f_\sigma(x) = \frac{\sigma}{5} f_{\sigma/5}(\frac{x}{5}).\]
For fair comparisons, we set the parameters of the RED algorithms according to the values given in [65].

The same values are used for the RED-PRO methods that are in common with RED, while other additional parameters of RED-PRO variants are tuned manually. In addition, we assume throughout the experiments that the denoiser is 0-demicontractive, i.e., it is quasi-nonexpansive.

5.1. Image Deblurring. As done in [65], we carry out here the synthetic non-blind deblurring experiments first described in [32]. To that end, we perform the following process on the RGB images provided by the authors of [65]. Each image is convolved either with a \( 9 \times 9 \) uniform point spread function (PSF) or a 2D Gaussian function with a standard deviation of 1.6. Then, we contaminate the resulted blurred images with an additive WGN with \( \sigma = \sqrt{2} \). The inverse operation is performed by first converting the RGB image to YCbCr color-space, applying the deblurring technique only on the luminance channel and then converting the result back to RGB to obtain the final image. The complete details of the parameter values are given in Table 5.1 and Table 5.4.

We start with providing in Fig. 5.1 an illustration of the convergence of the RED-PRO (via HSD) for various denoisers. In Fig. 5.1(a) we plot the evolution of the fidelity term throughout the iterations, while Fig. 5.1(b) displays a measure/penalty of the solution proximity to the fixed-point set over iterations. As can be seen, we exhibit same rapid decline the fidelity term for all denoisers. Moreover, similar behavior is observed for all the denoisers shown in Fig. 5.1(b) where we first see an increase in the fixed-point penalty, followed by a dramatic decrease and then a slow decay, depending on the denoiser in use. Note that while these measures demonstrate the convergence of the proposed approach, they might not accurately indicate the quality of the resultant images.

Table 5.2 and Table 5.5 detail the recovery results of RED and RED-PRO algorithms when NLM is used, while Table 5.3 and Table 5.6 provides the results when TNRD is employed. We evaluate the performance of the different techniques using the peak signal to noise ratio (PSNR) measure, where higher is better, computed on the luminance channel of the ground truth and the restored images. As can be seen, when NLM is used within the discussed frameworks, the RED-PRO approach displays an apparent improvement, in particular for the case of uniform deblurring. When TNRD is incorporated, the RED-PRO framework still leads to enhanced results in comparison to RED, but the performance gap is small. Furthermore, we notice that SD-based methods require more iterations which is consistent with the observations given in [65].

We present visual evidence for the performance of RED-PRO in Fig. 5.2 and Fig. 5.3 for uniform and Gaussian blur kernels, respectively. These images support the previous results where RED-PRO offer an improvement over the RED framework when NLM is utilized. In the other case, both approaches achieve similar performance with a clear advantage for applying TNRD rather than employing NLM.

5.2. Super-Resolution. Similar to [65], we continue with super-resolution experiments where ground-truth high-resolution images are blurred with \( 7 \times 7 \) Gaussian kernel with stan-
**Figure 5.1.** An illustration of the convergence of RED-PRO via HSD using various denoisers. Here we degrade the image *Plants*, degraded by a uniform PSF and restore it with Algorithm 4.1 where $\alpha = 0.2$. (a) The evolution of the fidelity term during the iterations. (b) A measure of the solution proximity to the fixed-point set throughout the iterations.

| Parameter | RED | RED-PRO | Approximate RED-PRO |
|-----------|-----|---------|---------------------|
|           | FP  | ADMM   | SD                  | FP  | ADMM   | SD                  |
| $N$       | 200 | 200    | 1500               | 400 | 200    | 200                |
| $\sigma_f$| 3.25| 3.25   | 3.25               | 3.25| 3.25   | 3.25               |
| $\lambda$ | 0.02| 0.02   | 0.02               | 0.02| 0.02   | 0.02               |
| $\beta$   | 0.001|—       |—                   | 0.001|—       |—                   |
| $m_1$     | *CF | *CF    |—                   | *CF | *CF    |—                   |
| $m_2$     | 1   | 1      |—                   | 1   | 1      |—                   |
| $\alpha$  | 0.055| 1      | 1                  | 1   | 1      | 1                  |
| $J$       | 3   | 3      | 3                  | 3   | 3      | 3                  |
| $\epsilon$| —   | —      |—                   | —   | —      |—                   |
| $\delta$  | —   | 0.0002 | 0.0002             | 0.0002| 0.0002| 0.0002             |

*Closed-form using FFT

**Table 5.1**

The set of parameters used in RED and RED-PRO frameworks for the task of deblurring images corrupted by a uniform PSF. The description of ADMM inner parameters $m_1$ and $m_2$ is given in [65].

Standard deviation of 1.6, downsampled by a factor of three at each axis and finally they are corrupted by an additive WGN with $\sigma = 5$. To restore the images from their low-resolution versions, we convert them to YCbCr color-space and perform super-resolution where the chroma channels are up-scaled by bicubic interpolation and the luminance channel is processed by the discussed algorithms. The results are then transformed back to RGB, yielding the final super-resolved images. The parameter values of this case are given in Table 5.7.

We present the restoration results of the different variants of RED and RED-PRO frame-
Table 5.2
Deblurring with NLM: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with uniform blur kernel. Performance is measured in PSNR [dB] where highest result are highlighted.

| Image/Algorithm | RED          | RED-PRO       | Approximate RED-PRO |
|-----------------|--------------|---------------|---------------------|
|                 | FP  | ADMM | SD  | FP | ADMM | SD  | FP  | ADMM | SD  |
| Bike            | 23.07 | 23.12 | 24.41 | 24.95 | 23.09 | 23.16 | 24.46 | 23.08 | 23.13 | 24.51 |
| Butterfly       | 24.93 | 25.01 | 26.93 | 27.47 | 24.93 | 25.00 | 26.96 | 24.83 | 24.87 | 26.89 |
| Flower          | 26.21 | 26.25 | 27.89 | 29.26 | 26.27 | 26.33 | 27.97 | 26.25 | 26.27 | 28.01 |
| Girl            | 31.29 | 31.32 | 31.86 | 32.68 | 31.35 | 31.42 | 31.93 | 31.40 | 31.43 | 31.98 |
| Hat             | 29.04 | 29.06 | 30.40 | 31.42 | 29.15 | 29.13 | 30.49 | 29.08 | 29.05 | 30.46 |
| Leaves          | 24.40 | 24.45 | 26.47 | 27.07 | 24.36 | 24.44 | 26.44 | 24.21 | 24.29 | 26.34 |
| Parrots         | 28.82 | 28.92 | 30.06 | 30.03 | 28.82 | 28.90 | 30.08 | 28.77 | 28.77 | 30.06 |
| Parthenon       | 25.14 | 25.20 | 26.98 | 29.35 | 25.22 | 25.27 | 27.10 | 25.18 | 25.26 | 27.14 |
| Plants          | 30.52 | 30.47 | 32.14 | 33.55 | 30.55 | 30.56 | 32.20 | 30.48 | 30.53 | 32.22 |
| Racoon          | 26.86 | 26.83 | 27.60 | 28.62 | 26.98 | 26.98 | 27.75 | 26.90 | 26.87 | 27.76 |
| Starfish        | 24.98 | 25.03 | 26.99 | 28.96 | 25.02 | 25.05 | 27.08 | 24.95 | 24.99 | 27.06 |
| Average         | 26.84 | 26.88 | 28.34 | 29.40 | 26.89 | 26.93 | 28.40 | 26.83 | 26.86 | 28.40 |

Table 5.3
Deblurring with TNRD: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with uniform blur kernel. Performance is measured in PSNR [dB] where highest result are highlighted.

| Image/Algorithm | RED          | RED-PRO       | Approximate RED-PRO |
|-----------------|--------------|---------------|---------------------|
|                 | FP  | ADMM | SD  | FP | ADMM | SD  | FP  | ADMM | SD  |
| Bike            | 26.10 | 26.09 | 25.95 | 24.95 | 26.21 | 26.20 | 26.05 | 26.48 | 26.47 | 26.28 |
| Butterfly       | 30.41 | 30.40 | 30.20 | 27.24 | 30.48 | 30.47 | 30.24 | 30.64 | 30.63 | 30.35 |
| Flower          | 30.18 | 30.18 | 30.13 | 29.38 | 30.26 | 30.26 | 30.21 | 30.46 | 30.46 | 30.39 |
| Girl            | 33.11 | 33.11 | 33.10 | 32.99 | 33.14 | 33.14 | 33.13 | 33.19 | 33.19 | 33.19 |
| Hat             | 32.16 | 32.16 | 32.15 | 31.55 | 32.17 | 32.18 | 32.17 | 32.25 | 32.25 | 32.24 |
| Leaves          | 30.13 | 30.12 | 29.72 | 26.92 | 30.20 | 30.20 | 29.77 | 30.39 | 30.38 | 29.88 |
| Parrots         | 31.83 | 31.86 | 31.62 | 30.12 | 31.92 | 31.93 | 31.68 | 32.13 | 32.12 | 31.82 |
| Parthenon       | 29.85 | 29.85 | 29.85 | 29.82 | 29.97 | 29.97 | 29.86 | 30.23 | 30.23 | 30.22 |
| Plants          | 34.25 | 34.25 | 34.19 | 33.59 | 34.27 | 34.27 | 34.21 | 34.40 | 34.39 | 34.33 |
| Racoon          | 28.83 | 28.84 | 28.81 | 28.99 | 28.94 | 28.94 | 28.91 | 29.12 | 29.12 | 29.08 |
| Starfish        | 30.57 | 30.57 | 30.47 | 29.41 | 30.65 | 30.65 | 30.55 | 30.84 | 30.84 | 30.72 |
| Average         | 30.67 | 30.67 | 30.56 | 29.54 | 30.75 | 30.75 | 30.63 | 30.92 | 30.92 | 30.77 |

works when NLM and TNRD are used in Table 5.8 and Table 5.9 respectively. As observed, the RED-PRO approach leads to slightly better results than RED. This is visually supported by the recovered images shown in Fig 5.4 where we compare our approach to naive bicubic interpolation in addition to RED. Both quantitative and qualitative results point out the clear advantage of using TNRD over NLM.
Figure 5.2. Visual comparison of deblurring the image *Butterfly*, degraded by a uniform PSF, along with the corresponding PSNR [dB] score.

| Parameter | RED | RED-PRO | Approximate RED-PRO |
|-----------|-----|---------|---------------------|
|           | FP  | ADMM    | SD                  |
| $N$       | 200 | 200     | 1500                |
| $\sigma_f$| 4.1 | 4.1     | 4.1                 |
| $\lambda$ | 0.01| 0.01    | 0.01                |
| $m_1$     | *CF | *CF     | *CF                 |
| $m_2$     | 1   | 1       | 1                   |
| $\alpha$  | 0.035| 1       | 1                   |
| $J$       | 3   | 3       | 3                   |
| $\epsilon$|     |         | 0.0002              |
| $\delta$  |     |         | 0.0001              |

*Closed-form using FFT

Table 5.4

The set of parameters used in RED and RED-PRO frameworks for the task of deblurring images corrupted by a Gaussian PSF. The description of ADMM inner parameters $m_1$ and $m_2$ is given in [65].

6. Conclusion. The massive advancement in image denoising has led in recent years to the development of PnP and RED frameworks, which leverage the power of denoisers to solve other challenging inverse problems. These approaches have shown remarkable performance, leading to state-of-the-art results in unsupervised tasks such as image deblurring and super-resolution. However, PnP lacks an underlying global objective function, while RED demands restrictive conditions that are not met by many popular denoisers.

In this paper, we address the above limitations. The main contribution of this work is the introduction of the RED-PRO framework, which relies on the fixed-point set of demicon-
| Image/Algorithm | RED | RED-PRO | Approximate RED-PRO |
|-----------------|-----|---------|-------------------|
|                 | FP  | ADMM  | SD    | FP  | ADMM  | SD    | FP  | ADMM  | SD    |
| Bike            | 25.30 | 25.45 | 26.97 | **27.13** | 25.38 | 25.54 | 27.03 | 25.42 | 25.57 | 27.08 |
| Butterfly       | 27.14 | 27.28 | 29.61 | **29.87** | 27.18 | 27.36 | 29.64 | 27.21 | 27.37 | 29.66 |
| Flower          | 28.47 | 28.62 | 30.75 | **31.64** | 28.58 | 28.73 | 30.85 | 28.61 | 28.77 | 30.92 |
| Girl            | 33.32 | 33.45 | 33.33 | 32.97 | 33.33 | 33.41 | 33.29 | 33.29 | 33.36 | 33.27 |
| Hat             | 30.84 | 30.95 | 32.39 | **32.74** | 30.92 | 31.04 | 32.45 | 30.96 | 31.08 | 32.47 |
| Leaves          | 27.50 | 27.66 | 29.98 | **30.26** | 27.53 | 27.70 | 30.00 | 27.51 | 27.68 | 30.00 |
| Parrots         | 30.76 | 30.89 | 32.30 | **32.34** | 30.82 | 30.97 | 32.30 | 30.81 | 30.96 | 32.30 |
| Parthenon       | 26.93 | 27.06 | 28.45 | **28.86** | 27.05 | 27.17 | 28.54 | 27.10 | 27.22 | 28.59 |
| Plants          | 33.30 | 33.50 | 35.01 | **35.47** | 33.36 | 33.56 | 34.98 | 33.39 | 33.55 | 34.97 |
| Raccoon         | 29.16 | 29.22 | 30.20 | **30.52** | 29.25 | 29.40 | 30.34 | 29.33 | 29.41 | 30.39 |
| Starfish        | 28.05 | 28.19 | 30.72 | **31.47** | 28.11 | 28.28 | 30.78 | 28.13 | 28.31 | 30.82 |
| Average         | 29.16 | 29.30 | 30.88 | **31.21** | 29.24 | 29.38 | 30.93 | 29.25 | 29.39 | 30.95 |

Table 5.5

Deblurring with NLM: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with Gaussian blur kernel. Performance is measured in PSNR [dB] where highest result are highlighted.

| Image/Algorithm | RED | RED-PRO | Approximate RED-PRO |
|-----------------|-----|---------|-------------------|
|                 | FP  | ADMM  | SD    | FP  | ADMM  | SD    | FP  | ADMM  | SD    |
| Bike            | 27.90 | 27.90 | 27.88 | 27.36 | 27.95 | 27.94 | **27.92** | 28.02 | 28.02 | 28.00 |
| Butterfly       | **31.66** | 31.66 | 31.57 | 30.55 | 31.65 | 31.64 | 31.54 | **31.66** | 31.65 | 31.52 |
| Flower          | 32.05 | 32.05 | 32.05 | 31.81 | 32.05 | 32.05 | 32.04 | 32.08 | **32.08** | 32.07 |
| Girl            | **34.44** | 34.44 | 34.10 | 33.23 | 34.40 | 34.40 | 34.00 | 34.38 | 34.38 | 33.88 |
| Hat             | 33.29 | 33.29 | **33.30** | 33.07 | 33.25 | 33.25 | 33.26 | 33.26 | 33.26 | 33.27 |
| Leaves          | 31.93 | 31.93 | **31.95** | 30.79 | 31.91 | 31.91 | 31.91 | 31.92 | 31.92 | 31.87 |
| Parrots         | **33.33** | 33.32 | 33.19 | 32.59 | 33.29 | 33.29 | 33.15 | 33.31 | 33.30 | 33.15 |
| Parthenon       | 29.15 | 29.15 | 29.14 | 28.96 | 29.17 | 29.17 | 29.16 | **29.23** | 29.23 | 29.20 |
| Plants          | 36.38 | 36.38 | **36.39** | 35.69 | 36.26 | 36.27 | 36.25 | 36.22 | 36.22 | 36.16 |
| Raccoon         | 31.07 | 31.08 | 31.03 | 30.82 | 31.13 | 31.14 | 31.08 | **31.20** | 31.20 | 31.11 |
| Starfish        | 32.49 | 32.49 | 32.46 | 31.92 | 32.49 | 32.49 | 32.44 | **32.52** | 32.52 | 32.45 |
| Average         | 32.15 | 32.15 | 32.09 | 31.53 | 32.14 | 32.14 | 32.07 | **32.16** | 32.16 | 32.06 |

Table 5.6

Deblurring with TNRD: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with Gaussian blur kernel. Performance is measured in PSNR [dB] where highest result are highlighted.

tractive denoisers. We formulate a general inverse problem that exploits the convexity of the fixed-point set and utilize it as a regularization. Then, we derive several iterative methods for solving the proposed problems and prove their convergence to global optimum. We discuss the relation of the presented framework to PnP and RED, including the connection of demicontractivity to previous well-known assumptions such nonexpansiveness and boundedness. This show that RED-PRO joins ideas from both PnP and RED to create a unified approach which
Figure 5.3. Visual comparison of deblurring the image Starfish, degraded by a Gaussian PSF, along with the corresponding PSNR [dB] score.

Table 5.7
The set of parameters used in RED and RED-PRO frameworks for the task of single-image super-resolution. The description of ADMM inner parameters $m_1$ and $m_2$ is given in [65].

| Parameter | RED      | RED-PRO | Approximate RED-PRO |
|-----------|----------|---------|---------------------|
|           | FP       | ADMM    | SD                  | FP       | ADMM    | SD                  |
| $N$       | 200      | 200     | 1500                | 400      | 200     | 200                 | 1500      | 200     | 200     |
| $\sigma_f$| 3        | 3       | 3                   | 3        | 3       | 3                   | 3         | 3       | 3       |
| $\lambda$ | 0.008    | 0.008   | 0.008               | 0.008    | 0.008   | 0.008               | 0.008     | 0.008   | 0.008   |
| $\beta$   | —        | 0.001   | —                   | 0.001    | —       | —                   | 0.001     | —       | —       |
| $m_1$     | 200      | 200     | —                   | 200      | 200     | —                   | 200       | 200     | —       |
| $m_2$     | —        | 1       | —                   | 1        | —       | 1                   | 1         | 1       | 1       |
| $\alpha$  | —        | —       | —                   | 0.035    | 1       | 1                   | 1         | 1       | 1       |
| $J$       | —        | —       | —                   | 3        | 3       | 3                   | 3         | 3       | 3       |
| $\epsilon$| —        | —       | —                   | —        | —       | —                   | 0.0002    | 0.0002  | 0.0002  |
| $\delta$  | —        | —       | —                   | —        | —       | —                   | 0.0001    | 0.0001  | 0.0001  |

provides flexibility in choosing the denoiser and is guaranteed to converge to a stable solution. Furthermore, we present variants of RED-PRO based on the dilated fixed-point set that allow robustness when the fixed-point set of the denoiser is narrow. As proven, the dilated fixed-point set is contained in the approximated fixed-point set, which is a much broader domain but not convex in general. Therefore, an open research direction is to provide the conditions under which the latter set is convex and derive different means to exploit it as a regularization. Finally, we perform experiments in image deblurring and super-resolution to demonstrate the validity of the regularization strategy, showing competitive and even improved performance.
We conclude this section by briefly discussing key questions for future research. This study relies on the condition that the denoiser is demicontractive. Currently, determining whether an arbitrary denoiser satisfies this condition remains an open challenge. Certainly, any nonexpansive denoiser is also demicontractive, yet, it is unclear how to verify the nonexpansiveness of a general function. Therefore, it is of utmost practical importance to derive an efficient

Table 5.8
Super-Resolution with NLM: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with Gaussian blur kernel and downsampled by a factor of 3 in each axis. Performance is measured in PSNR [dB] where highest result are highlighted.

| Image/Algorithm | RED | RED-PRO | Approximate RED-PRO |
|-----------------|-----|---------|---------------------|
|                 | FP  | ADMM    | SD                  | FP  | ADMM    | SD                  |
| Bike            | 22.97 | 23.04 | 22.27              | 23.10 | 23.02 | 23.08              | 22.34 | 23.05 | 23.11 | 22.41 |
| Butterfly       | 25.27 | 25.37 | 24.23              | 24.85 | 25.30 | 25.39              | 24.29 | 25.32 | 25.40 | 24.37 |
| Flower          | 26.70 | 26.82 | 25.42              | 27.21 | 26.79 | 26.91              | 25.48 | 26.89 | 27.02 | 25.64 |
| Girl            | 30.75 | 30.78 | 30.03              | 30.50 | 30.84 | 30.90              | 30.13 | 30.96 | 31.02 | 30.19 |
| Hat             | 29.19 | 29.27 | 28.19              | 28.73 | 29.25 | 29.34              | 28.29 | 29.28 | 29.38 | 28.31 |
| Leaves          | 24.60 | 24.68 | 23.77              | 24.32 | 24.64 | 24.73              | 23.79 | 24.69 | 24.77 | 23.85 |
| Parrots         | 28.35 | 28.39 | 27.84              | 27.98 | 28.42 | 28.46              | 27.87 | 28.49 | 28.54 | 27.92 |
| Parthenon       | 25.27 | 25.64 | 24.79              | 25.71 | 25.62 | 25.71              | 24.87 | 25.73 | 25.78 | 24.97 |
| Plants          | 30.38 | 30.49 | 28.96              | 29.88 | 30.47 | 30.57              | 28.95 | 30.47 | 30.63 | 28.89 |
| Raccoon         | 27.14 | 27.20 | 26.52              | 27.30 | 27.23 | 27.29              | 26.63 | 27.28 | 27.33 | 26.62 |
| Starfish        | 25.90 | 26.01 | 24.88              | 26.69 | 25.98 | 26.09              | 24.98 | 26.09 | 26.19 | 25.13 |
| Average         | 26.98 | 27.06 | 26.08              | 26.94 | 27.05 | 27.13              | 26.15 | 27.11 | 27.20 | 26.21 |

Table 5.9
Super-Resolution with TNRD: Recovery results obtained by RED and RED-PRO evaluated on the set of images, provided by the authors of [65], which corrupted with Gaussian blur kernel and downsampled by a factor of 3 in each axis. Performance is measured in PSNR [dB] where highest result are highlighted.

| Image/Algorithm | RED | RED-PRO | Approximate RED-PRO |
|-----------------|-----|---------|---------------------|
|                 | FP  | ADMM    | SD                  | FP  | ADMM    | SD                  |
| Bike            | 23.97 | 23.96 | 24.04              | 23.22 | 24.00 | 23.98              | 24.06 | 24.03 | 24.01 | 24.09 |
| Butterfly       | 27.26 | 27.22 | 27.37              | 24.95 | 27.23 | 27.19              | 27.37 | 27.11 | 27.06 | 27.32 |
| Flower          | 28.24 | 28.24 | 28.23              | 27.11 | 28.26 | 28.26              | 28.25 | 28.30 | 28.30 | 28.28 |
| Girl            | 32.08 | 32.08 | 32.08              | 29.93 | 32.09 | 32.09              | 32.08 | 32.05 | 32.06 | 32.05 |
| Hat             | 30.35 | 30.35 | 30.36              | 28.21 | 30.37 | 30.36              | 30.37 | 30.36 | 30.35 | 30.37 |
| Leaves          | 26.12 | 26.10 | 26.17              | 24.47 | 26.11 | 26.09              | 26.22 | 26.06 | 26.02 | 26.25 |
| Parrots         | 29.42 | 29.41 | 29.43              | 27.72 | 29.44 | 29.44              | 29.50 | 29.44 | 29.43 | 29.46 |
| Parthenon       | 26.52 | 26.51 | 26.54              | 25.58 | 26.52 | 26.52              | 26.54 | 26.48 | 26.48 | 26.51 |
| Plants          | 31.77 | 31.77 | 31.79              | 29.44 | 31.79 | 31.79              | 31.82 | 31.76 | 31.75 | 31.77 |
| Raccoon         | 27.97 | 27.97 | 27.98              | 27.20 | 28.01 | 28.01              | 28.01 | 28.02 | 28.02 | 28.03 |
| Starfish        | 27.94 | 27.93 | 27.96              | 26.83 | 27.96 | 27.95              | 27.99 | 27.98 | 27.97 | 28.04 |
| Average         | 28.33 | 28.32 | 28.36              | 26.79 | 28.34 | 28.33              | 28.38 | 28.33 | 28.31 | 28.38 |
technique for testing the demicontractivity of a given mapping. An alternative approach can be using learned mappings enforced to satisfy this condition. Similar concept is applied in the field of deep learning where neural networks are constrained by a given Lipschitz constant \cite{38} to obtain regularization during training and robustness to adversarial attacks.

In addition, as in \cite{65}, the RED-PRO framework exploits denoisers that are designed for optimized distortion measure, such as PSNR. However, this might lead to a low perceptual quality of the resultant images \cite{10}. Hence, an interesting research direction is the deployment of denoisers based on some perceptual loss. Employing such denoisers in the RED-PRO scheme might result in solutions that better address the perception-distortion tradeoff \cite{10}. Finally, the family of demicontractive mappings covers a broad range of functions, hence, a promising study direction is incorporating within the RED-PRO framework powerful mechanisms, different from denoisers, which encapsulate prior knowledge on the unknown image.

**Appendix A. Proof of Theorem 3.10.** We first show that $\text{Fix}(T)$ is closed. Let $\{x_i\}$ be a sequence in $\text{Fix}(T)$ such that $x_i \to x$. By the definition of demicontractive mappings, we get that

$$\|T(x) - x_i\|^2 \leq \|x_i - x\|^2 + d \|x - T(x)\|^2, \forall i \in \mathbb{N}.$$  

By taking the limit as $i \to \infty$ of both sides the above inequality we obtain

$$\|T(x) - x\|^2 \leq d \|T(x) - x\|^2.$$  

Since $d < 1$, it implies that $T(x) = x$, i.e., $x \in \text{Fix}(T)$. 

**Figure 5.4.** Visual comparison of super-resolution by a factor of 3 at each axis for the image *Parrots*, along with the corresponding PSNR [dB] score.
We proceed to prove the convexity of Fix($T$). Let $x, z \in \text{Fix}(T)$ and define $u \triangleq \lambda x + (1 - \lambda)z$ for arbitrary $0 \leq \lambda \leq 1$. From (3.10), we get

$$
\|u - T(u)\|^2 = \|\lambda(x - T(u)) + (1 - \lambda)(z - T(u))\|^2 \\
= \lambda \|x - T(u)\|^2 + (1 - \lambda) \|z - T(u)\|^2 \\
- \lambda(1 - \lambda) \|x - z\|^2.
$$

Since $T$ is $d$-demicontactive, we obtain

$$
\|u - T(u)\|^2 \leq \lambda \left( \|x - u\|^2 + d \|u - T(u)\|^2 \right) \\
+ (1 - \lambda) \left( \|z - u\|^2 + d \|u - T(u)\|^2 \right) \\
- \lambda(1 - \lambda) \|x - z\|^2 \\
= d \|u - T(u)\|^2 + \lambda \|x - u\|^2 \\
+ (1 - \lambda) \|z - u\|^2 - \lambda(1 - \lambda) \|x - z\|^2.
$$

By using (3.10) again, we have

$$
\lambda \|x - u\|^2 + (1 - \lambda) \|z - u\|^2 - \lambda(1 - \lambda) \|x - z\|^2 = 0.
$$

Hence,

$$
\|u - T(u)\|^2 \leq d \|u - T(u)\|^2,
$$

which implies that $T(u) = u$ since $d < 1$.

**Appendix B. Proof of Theorem 3.15.** Since $T$ is cyclically firmly nonexpansive it is nonexpansive in particular, implying that $A$ is maximal monotone. Therefore, it is left to prove that $A$ is also cyclically monotone. First, consider any set of points $\{x_1, ..., x_m\}$ in $\mathbb{R}^n$ and arbitrary $m \geq 2$. Then, the following holds

$$
\sum_{i=1}^m \langle A(x_i), A(x_i) - A(x_{i+1}) \rangle = \frac{1}{2} \sum_{i=1}^m \|A(x_i) - A(x_{i+1})\|^2 \geq 0,
$$

where $x_{m+1} = x_1$. In addition, by Definition 3.4 we have that

$$
\sum_{i=1}^m \left\langle x_i - T(x_i), T(x_{i+1}) - T(x_i) \right\rangle \leq 0,
$$
where \( x_{m+1} = x_1 \). Hence,

\[
\sum_{i=1}^{m} \langle x_i - T(x_i), T(x_{i+1}) - T(x_i) \rangle = \sum_{i=1}^{m} \langle A(x_i), x_{i+1} - A(x_{i+1}) - x_i + A(x_i) \rangle \\
= \sum_{i=1}^{m} \langle A(x_i), x_{i+1} - x_i + A(x_i) - A(x_{i+1}) \rangle \\
= \sum_{i=1}^{m} \langle A(x_i), x_{i+1} - x_i \rangle + \sum_{i=1}^{m} \langle A(x_i), A(x_i) - A(x_{i+1}) \rangle \\
= \sum_{i=1}^{m} \langle A(x_i), x_{i+1} - x_i \rangle + \frac{1}{2} \sum_{i=1}^{m} \|A(x_i) - A(x_{i+1})\|^2 \leq 0.
\]

The latter suggests that

\[
\sum_{i=1}^{m} \langle A(x_i), x_{i+1} - x_i \rangle \leq 0,
\]
i.e., \( A \) is also cyclically monotone and thus maximal cyclically monotone.

**Appendix C. RED Revisited.** Here, we complement the work in [62] by considering the RED framework and studying the case in which the denoiser is non-differentiable. We do so by reformulating the RED optimization problem as follows.

Let \( f : \mathbb{R}^n \to \mathbb{R}^n \) be a chosen denoiser and define the denoiser residual (or displacement operator) \( A \triangleq \text{Id} - f \). For arbitrary \( u \in \mathbb{R}^n \), we consider the following minimization problem

(C.1) \[
\min_{x \in \mathbb{R}^n} E_{\text{Rock}}(x) = \ell(x; y) + \lambda R_A(x; u)
\]

where \( \lambda \geq 0 \) and \( R_A(x; u) \) is the Rockafellar function, defined below.

**Definition C.1 (Rockafellar function [64]).** Consider a mapping \( A : \mathbb{R}^n \to \mathbb{R}^n \) and let \( u \in \mathbb{R}^n \) be an arbitrary point. For \( m = 2 \), we define the following function with parameter \( u \)

\[
C^2_A(x; u) \triangleq \langle x, A(u) \rangle - \langle u, A(u) \rangle, \quad m = 2.
\]

Otherwise, suppose \( m > 2 \), then \( C^m_A(x; u) \) is defined by

\[
C^m_A(x; u) \triangleq \sup_{(x_2, \ldots, x_{m-1})} \left( \langle x - x_{m-1}, A(x_{m-1}) \rangle + \sum_{i=1}^{m-2} \langle x_{i+1} - x_i, A(x_i) \rangle \right),
\]

where \( x_1 = u \). The Rockafellar function is defined as

\[
R_A(x; u) \triangleq \sup_{m \in \{2, 3, \ldots\}} C^m_A(x; u).
\]

The Rockafellar function in a well-known function in the area of variational analysis [3]. The next theorem and the subsequent corollary provide the conditions for the convexity of (C.1).
Theorem C.2. Let $A$ be maximal cyclically monotone and let $u \in \mathbb{R}^n$. Then $R_A(x; u)$ is convex, lower semicontinuous, proper with $R_A(u; u) = 0$ and for any $x \in \mathbb{R}^n$ we have

$$A(x) = \nabla R_A(x; u).$$

Moreover, for any convex, lower semicontinuous and proper function $h : \mathbb{R}^n \to \mathbb{R}$ for which $A = \nabla h$, it holds that

$$h(x) = h(u) + R_A(x; u), \; \forall x \in \mathbb{R}^n,$$

i.e., $h(\cdot)$ is determined by $A$ uniquely up to an additive constant.

Proof. See [3].

Corollary C.3. Assume that $f(\cdot)$ is cyclically firmly nonexpansive. Then, problem (C.1) is a convex optimization task where the gradient of the objective function is readily available by (2.10).

Proof. By Theorem 3.15, the displacement mapping $A = Id - f$ is maximal cyclically monotone, hence, $R_A(x; u)$ is a convex function whose gradient is $\nabla R_A(x; u) = A(x) = x - f(x)$. The latter in turn implies that

$$\nabla E_{Rock}(x) = \nabla \ell(x; y) + \lambda \nabla R_A(x; u) = \nabla \ell(x; y) + \lambda \left( x - f(x) \right).$$

The RED framework seeks a point $x^* \in \mathbb{R}^n$ for which

$$(C.2) \quad \nabla \ell(x^*; y) + \lambda \left( x^* - f(x^*) \right) = 0.$$

Thus, Theorem C.2 and Corollary C.3 provide the conditions under which the RED approach minimizes a convex minimization problem given by (C.1), and thus they prove the convergence of RED algorithms. This holds even for the case where the denoiser is not differentiable.

Note that the demand of $f(\cdot)$ to be cyclically firmly nonexpansive is a sufficient condition, whereas the requirement of $A = Id - f$ to be maximal cyclically monotone is a necessary condition which might be fulfilled regardless to the nonexpansiveness of the denoiser.

Appendix D. Proof of Theorem 4.5. Consider $\alpha \in (0, \frac{1 - d}{2})$ and define the averaged operator $f_\alpha$. By Theorem 3.8, $f_\alpha$ is $\gamma$-strongly quasi-nonexpansive with $\gamma > 1$. Hence, according to Theorem 3.9 we have $\|x - f_\alpha(x)\| \leq \|x - P_{Fix(f)}(x)\|$. In addition, it holds that

$$\alpha \|x - f(x)\| = \|x - f_\alpha(x)\|.$$

Therefore, for any $x \in B_\delta(f)$ where $\delta \in [0, \alpha \epsilon]$, we have

$$\|x - f(x)\| = \frac{1}{\alpha} \|x - f_\alpha(x)\| \leq \frac{1}{\alpha} \|x - P_{Fix(f)}(x)\| \leq \frac{\delta}{\alpha} \leq \epsilon,$$

which implies that $x \in Fix_\epsilon(f)$. 

REFERENCES

[1] R. Ahmad, C. A. Bouman, G. T. Buzzard, S. Chan, E. T. Reehorst, and P. Schnitter, Plug and play methods for magnetic resonance imaging, arXiv preprint arXiv:1903.08616, (2019).

[2] C. Barguet, S. Reich, and R. Zalas, Convergence properties of dynamic string-averaging projection methods in the presence of perturbations, Numerical Algorithms, 77 (2018), pp. 185–209.

[3] S. Bartz, H. H. Bauschke, J. M. Borwein, S. Reich, and X. Wang, Fitzpatrick functions, cyclic monotonicity and Rockafellar’s antiderivative, Nonlinear Analysis: Theory, Methods & Applications, 66 (2007), pp. 1198–1223.

[4] H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, Fixed-point algorithms for inverse problems in science and engineering, vol. 49, Springer Science & Business Media, 2011.

[5] H. H. Bauschke, P. L. Combettes, et al., Convex analysis and monotone operator theory in Hilbert spaces, vol. 408, Springer, 2011.

[6] M. Bayati and A. Montanari, The dynamics of message passing on dense graphs, with applications to compressed sensing, IEEE Transactions on Information Theory, 57 (2011), pp. 764–785.

[7] A. Beck, First-order methods in optimization, vol. 25, SIAM, 2017.

[8] A. Beck and M. Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, IEEE Transactions on image processing, 18 (2009), pp. 2419–2434.

[9] D. P. Bertsekas, Nonlinear programming, Journal of the Operational Research Society, 48 (1997), pp. 334–334.

[10] Y. Blau and T. Michaeli, The perception-distortion tradeoff, in Proceedings of the IEEE CVPR, 2018, pp. 6228–6237.

[11] S. Boyd and L. Vandenberghe, Convex optimization, Cambridge university press, 2004.

[12] A. M. Bruckstein, D. L. Donoho, and M. Elad, From sparse solutions of systems of equations to sparse modeling of signals and images, SIAM review, 51 (2009), pp. 34–81.

[13] A. Buades, B. Coll, and J.-M. Morel, A non-local algorithm for image denoising, in IEEE CVPR, vol. 2, 2005, pp. 60–65.

[14] H. C. Burger, C. J. Schuler, and S. Harmeling, Image denoising: Can plain neural networks compete with BM3D?, in IEEE CVPR, 2012, pp. 2392–2399.

[15] G. T. Buzzard, S. H. Chan, S. Sreehari, and C. A. Bouman, Plug-and-play unplugged: Optimization-free reconstruction using consensus equilibrium, SIAM Journal on Imaging Sciences, 11 (2018), pp. 2001–2020.

[16] A. Cegielski, S. Reich, and R. Zalas, Regular sequences of quasi-nonexpansive operators and their applications, SIAM Journal on Optimization, 28 (2018), pp. 1508–1532.

[17] S. H. Chan, Performance analysis of plug-and-play admm: A graph signal processing perspective, IEEE Transactions on Computational Imaging, 5 (2019), pp. 274–286.

[18] S. H. Chan, X. Wang, and O. A. Elgendy, Plug-and-play ADMM for image restoration: Fixed-point convergence and applications, IEEE Transactions on Computational Imaging, 3 (2016), pp. 84–98.

[19] P. Chatterjee and P. Milanfar, Is denoising dead?, IEEE Transactions on Image Processing, 19 (2009), pp. 895–911.

[20] P. Chatterjee and P. Milanfar, Patch-based near-optimal image denoising, IEEE Transactions on Image Processing, 21 (2011), pp. 1635–1649.

[21] Y. Chen and T. Pock, Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2016), pp. 1256–1272.

[22] T. Chernyakova, R. Cohen, R. Mulayoff, Y. Sde-Chen, C. Fraschini, J. Bercoff, and Y. C. Eldar, Fourier-domain beamforming and structure-based reconstruction for plane-wave imaging, IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 65 (2018), pp. 1810–1821.

[23] C. E. Chidume and Ş. Măruşter, Iterative methods for the computation of fixed points of demicontractive mappings, Journal of Computational and Applied Mathematics, 234 (2010), pp. 861–882.

[24] R. Cohen and Y. C. Eldar, Optimized sparse array design based on the sum covariance, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2018, pp. 3340–3344.

[25] R. Cohen and Y. C. Eldar, Sparse convolutional beamforming for ultrasound imaging, IEEE Transac-
[50] A. Levin, B. Nadler, F. Durand, and W. T. Freeman, *Patch complexity, finite pixel correlations and optimal denoising*, in European Conference on Computer Vision, Springer, 2012, pp. 73–86.

[51] F. Lieder, *On the convergence rate of the Halpern-iteration*, Optimization Online preprint, (2017), pp. 11–6336.

[52] P.-E. Maingé, *Convex minimization over the fixed point set of demicontractive mappings*, Positivity, 12 (2008), pp. 269–280.

[53] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, *Non-local sparse models for image restoration*, in IEEE 12th international conference on computer vision, 2009, pp. 2272–2279.

[54] S. Mallat, *A wavelet tour of signal processing*, Elsevier, 1999.

[55] P.-E. Maingé, *Convex minimization over the fixed point set of demicontractive mappings*, Positivity, 12 (2008), pp. 269–280.

[56] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, *Non-local sparse models for image restoration*, in IEEE 12th international conference on computer vision, 2009, pp. 2272–2279.

[57] S. Mallat, *A wavelet tour of signal processing*, Elsevier, 1999.

[58] L. Mǎruster and S. Mǎruster, *Strong convergence of the Mann iteration for Γ-demicontractive mappings*, Mathematical and Computer Modelling, 54 (2011), pp. 2486–2492.

[59] N. Parikh, S. Boyd, et al., *Proximal algorithms*, Foundations and Trends® in Optimization, 1 (2014), pp. 127–239.

[60] A. Perelli and M. S. Andersen, *Regularized Newton sketch by denoising score matching for computed tomography reconstruction*, in Workshop on Signal Processing with Adaptive Sparse Structured Representations, 2019.

[61] N. Parikh, S. Boyd, et al., *Proximal algorithms*, Foundations and Trends® in Optimization, 1 (2014), pp. 127–239.

[62] R. T. Rockafellar and R. J.-B. Wets, *Variational analysis*, vol. 317, Springer Science & Business Media, 2009.

[63] J. Rick Chang, C.-L. Li, B. Poczos, B. Vijaya Kumar, and A. C. Sankaranarayanan, *One network to solve them all: Solving linear inverse problems using deep projection models*, in Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 5888–5897.

[64] Y. Sun, B. Wohlberg, and U. S. Kamilov, *An online plug-and-play algorithm for regularized image reconstruction*, IEEE Transactions on Image Processing, 23 (2013), pp. 755–768.
[75] T. Tasdizen, *Principal neighborhood dictionaries for nonlocal means image denoising*, IEEE Transactions on Image Processing, 18 (2009), pp. 2649–2660.

[76] A. M. Teodoro, J. M. Bioucas-Dias, and M. A. Figueiredo, *Scene-adapted plug-and-play algorithm with convergence guarantees*, in IEEE 27th International Workshop on Machine Learning for Signal Processing (MLSP), 2017, pp. 1–6.

[77] A. M. Teodoro, J. M. Bioucas-Dias, and M. A. Figueiredo, *A convergent image fusion algorithm using scene-adapted Gaussian-mixture-based denoising*, IEEE Transactions on Image Processing, 28 (2018), pp. 451–463.

[78] T. Tirer and R. Giryes, *Image restoration by iterative denoising and backward projections*, IEEE Transactions on Image Processing, 28 (2018), pp. 1220–1234.

[79] D. Ulyanov, A. Vedaldi, and V. Lempitsky, *Deep image prior*, in Proceedings of the IEEE CVPR, 2018, pp. 9446–9454.

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

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

[82] J. Xing, U. Kamilov, W. Wu, Y. Wang, and M. Zhang, *Plug-and-play priors for reconstruction-based placental image registration*, in Smart Ultrasound Imaging and Perinatal, Preterm and Paediatric Image Analysis, Springer, 2019, pp. 133–142.

[83] I. Yamada, *Hybrid steepest descent method for variational inequality problem over the fixed point set of certain quasi-nonexpansive mappings*, Numerical Functional Analysis and Optimization, 25 (2004), pp. 619–655.

[84] G. Yu, G. Sapiro, and S. Mallat, *Solving inverse problems with piecewise linear estimators: From Gaussian mixture models to structured sparsity*, IEEE Transactions on Image Processing, 21 (2011), pp. 2481–2499.

[85] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, *Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising*, IEEE Transactions on Image Processing, 26 (2017), pp. 3142–3155.

[86] K. Zhang, W. Zuo, S. Gu, and L. Zhang, *Learning deep CNN denoiser prior for image restoration*, in Proceedings of the IEEE CVPR, 2017, pp. 3929–3938.

[87] K. Zhang, W. Zuo, and L. Zhang, *FFDNet: Toward a fast and flexible solution for CNN-based image denoising*, IEEE Transactions on Image Processing, 27 (2018), pp. 4608–4622.

[88] K. Zhang, W. Zuo, and L. Zhang, *Deep plug-and-play super-resolution for arbitrary blur kernels*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019, pp. 1671–1681.

[89] D. Zoran and Y. Weiss, *From learning models of natural image patches to whole image restoration*, in IEEE International Conference on Computer Vision, 2011, pp. 479–486.