



International Journal of Computational Mathematics and Scientific Computing

DOI: 10.5281/zenodo.23350303 | ISSN: 3141-643X | Volume: 2 | Issue: 1 | 2026

Projection and Proximal Operators in Image Denoising

I. D. Edem¹ , D. Goodhead² , U. M. Okon³ 

^{1,3}Department of Mathematics, Akwa Ibom State University, Nigeria

²Department of Mathematics, University of Port Harcourt, Nigeria

Corresponding author: inemesitedem@aksu.edu.ng

Abstract

Image denoising is a classical inverse problem in which the recovery of a clean image from its noisy observation is complicated by ill-posedness. This paper develops a functional-analytic framework in which a broad class of denoising methods is expressed as the iteration of firmly nonexpansive or quasi-nonexpansive operators on a Hilbert space. Total variation regularization, wavelet shrinkage, projection onto convex sets (POCS), projection onto epigraph sets (PESC), and regularization by denoising via fixed-point projection (RED-PRO) are each shown to fit this structure, and their convergence is obtained from a single Krasnosel'skii–Mann argument. The proximal nature of wavelet shrinkage is distinguished from metric projection onto the ℓ^1 -ball, and the demicontractive hypothesis underlying RED-PRO is stated explicitly. Numerical experiments on standard test images show that RED-PRO with a BM3D denoiser attains the highest PSNR at every noise level tested, that Chambolle's TV projection follows within 1.2–1.7 dB, and that POCS and wavelet shrinkage trade accuracy for speed. The framework clarifies the operator-theoretic content of each method and provides a common language for their analysis and comparison.

Keywords: image denoising; projection operators; total variation regularization; proximal operators; RED-PRO; convergence analysis.

Mathematics Subject Classification (2020): 47H09, 47H10, 47J25, 68U10, 65K10, 94A08.

1. Introduction

Image denoising seeks to recover a clean image u from its noisy observation $f = u + n$, where n is typically zero-mean Gaussian with variance σ^2 . The problem is ill-posed: infinitely many images are consistent with a given observation, and small perturbations in the data can produce large errors in the reconstruction [1]. Regularization is therefore required, and the choice of regularizer determines both the quality of the reconstruction and the mathematical structure of the algorithm that produces it.

The total variation (TV) regularizer of Rudin et al. [1] remains one of the most influential choices, because TV admits discontinuous solutions and therefore preserves edges while suppressing noise. Chambolle [2] derived a dual formulation of the ROF model and proposed a semi-implicit gradient descent that is, in effect, a projection onto a convex set in the dual space. This projection interpretation is now standard and has motivated accelerated and generalized variants, including recent deep architectures in which a single Chambolle layer improves PSNR over a plain autoencoder [24]. A second classical line treats denoising in a transform domain: wavelet shrinkage expands the signal in an orthogonal basis, applies a soft-threshold to the coefficients, and reconstructs by the inverse transform [16]. The soft-thresholding operator is exactly the proximal operator of the ℓ^1 -norm [28], which connects wavelet denoising to convex optimization; this is distinct from the metric projection onto the ℓ^1 -ball, and Section 3 makes the distinction precise. A third line is the method of projection onto convex sets (POCS), introduced for image restoration by Youla and Webb [5] and developed into a general theory of set-theoretic estimation [6]. POCS finds a point in the intersection of closed convex sets by cyclic projections, each set encoding a known property of the clean image. Its convergence theory in Hilbert spaces is well understood [7].

More recently, denoisers have been integrated as implicit regularizers. The plug-and-play (PnP) prior of Venkatakrishnan et al. [15] replaces the proximal operator of a regularizer by a state-of-the-art denoiser inside a splitting algorithm, while regularization by denoising (RED) of Romano et al. [14] constructs an explicit regularizer from a denoiser and optimizes the resulting energy. Both have achieved state-of-the-art results, and their theory has matured considerably [29]. Two subsequent developments are central to this paper. RED-PRO, introduced by Cohen, Elad, and Milanfar [3], reformulates RED as a convex problem whose solution is the projection of the observation onto the fixed-point set of a demicontractive denoiser, with global convergence guarantees. Projection onto epigraph sets (PESC) [11] lifts the minimization problem by one dimension and recasts it as a convex feasibility problem solvable by POCS.

Despite this progress, the literature remains fragmented. TV denoising is analyzed through duality, wavelet shrinkage through statistical estimation, POCS through convex feasibility, and RED-PRO through fixed-point theory. The mathematical machinery is largely shared (firmly nonexpansive operators, proximal maps, metric projections and Krasnosel'skii–Mann iterations) but the connections are rarely made explicit. This paper addresses that gap.

The core contributions are as follows.

1. A unified operator-theoretic formulation of projection-based denoising, encompassing Chambolle's dual projection for TV, wavelet shrinkage as a proximal operator, POCS, PESC, and RED-PRO, with a single convergence argument covering all cases.
2. A precise treatment of the distinction between soft-thresholding and metric projection onto the ℓ^1 -ball.
3. Explicit statement of the demicontractive hypothesis underlying RED-PRO.

-
4. A numerical comparison of the four methods on standard test images, showing that RED-PRO attains the highest PSNR at every noise level tested while TV follows within 1.2–1.7 dB.

Section 2 reviews the functional-analytic preliminaries. Section 3 treats projection operators in variational denoising, covering the ROF model, Chambolle’s algorithm, wavelet shrinkage, and POCS. Section 4 presents PESC and RED-PRO. Section 5 develops the convergence analysis. Section 6 reports the numerical experiments. Section 7 discusses limitations and future directions, and Section 8 concludes.

2. Functional-Analytic Preliminaries

2.1. Hilbert and Banach Spaces for Image Models

Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with Lipschitz boundary. The space $L^2(\Omega)$ of (equivalence classes of) square-integrable real-valued functions is a real Hilbert space under the inner product

$$\langle u, v \rangle_{L^2} = \int_{\Omega} u(x)v(x) dx, \quad (1)$$

with induced norm $\|u\|_{L^2} = \langle u, u \rangle_{L^2}^{1/2}$. This is the natural ambient space for images corrupted by additive Gaussian noise, because the fidelity term $\|f - u\|_{L^2}^2$ is then the negative log-likelihood up to constants.

The space $L^2(\Omega)$ is not well suited to images with sharp edges: functions in $L^2(\Omega)$ need not have bounded variation, and the total variation functional may be infinite. For edge-preserving regularization one therefore works in the space of functions of bounded variation,

$$BV(\Omega) = \{u \in L^1(\Omega) : TV(u) < \infty\}, \quad (2)$$

where the total variation is defined distributionally by

$$TV(u) = \sup \left\{ \int_{\Omega} u \operatorname{div} \varphi dx : \varphi \in C_c^1(\Omega, \mathbb{R}^2), \|\varphi\|_{\infty} \leq 1 \right\}. \quad (3)$$

Equipped with the norm

$$\|u\|_{BV} = \|u\|_{L^1} + TV(u), \quad (4)$$

the space $BV(\Omega)$ is a Banach space [18]. It is *not* a Hilbert space, and its dual is a complicated space of Radon measures. Consequently, projection-based arguments in $BV(\Omega)$ require care: the metric projection onto a closed convex subset of a Banach space need not exist, and if it does, it need not be unique or nonexpansive. The standard treatment of $BV(\Omega)$ and its properties is given by Ambrosio et al. [18].

Remark 2.1 (Working setting). For the operator-theoretic development in Sections 3–5 we work in a real Hilbert space H . In the continuous model we take $H = L^2(\Omega)$ and handle the BV regularizer through its dual formulation, which lives in $L^2(\Omega; \mathbb{R}^2)$. In the discrete model we take $H = \mathbb{R}^{M \times N}$ with the Frobenius inner product, which is the setting actually implemented in Section 6. This separation avoids the false impression that projection methods apply directly in $BV(\Omega)$.

2.2. Linear and Nonlinear Operators

Let H_1 and H_2 be real Hilbert spaces. A linear operator $A : H_1 \rightarrow H_2$ is *bounded* if

$$\|A\| = \sup_{\|u\|_{H_1}=1} \|Au\|_{H_2} < \infty. \quad (5)$$

Its *adjoint* $A^* : H_2 \rightarrow H_1$ is the unique bounded linear operator satisfying

$$\langle Au, v \rangle_{H_2} = \langle u, A^*v \rangle_{H_1} \quad \forall u \in H_1, v \in H_2. \quad (6)$$

The *null space* and *range* of A are denoted $\mathcal{N}(A)$ and $\mathcal{R}(A)$, respectively.

Nonlinear operators play a central role. Let $T : H \rightarrow H$. We recall the following standard classes.

Definition 2.2 (Nonexpansive and firmly nonexpansive operators). Let $C \subseteq H$ be nonempty. An operator $T : C \rightarrow H$ is:

1. *nonexpansive* if

$$\|Tu - Tv\| \leq \|u - v\| \quad \forall u, v \in C; \quad (7)$$

2. *firmly nonexpansive* if

$$\|Tu - Tv\|^2 \leq \langle u - v, Tu - Tv \rangle \quad \forall u, v \in C. \quad (8)$$

Equivalently, T is firmly nonexpansive if and only if

$$\|Tu - Tv\|^2 + \|(I - T)u - (I - T)v\|^2 \leq \|u - v\|^2 \quad \forall u, v \in C, \quad (9)$$

where I is the identity operator.

Every firmly nonexpansive operator is nonexpansive, by the Cauchy–Schwarz inequality. The converse is false. The class of firmly nonexpansive operators is exactly the class of *averaged* operators with constant $1/2$; more generally, T is α -averaged for some $\alpha \in (0, 1)$ if $T = (1 - \alpha)I + \alpha S$ with S nonexpansive. These classes are treated systematically in Bauschke and Combettes [9].

2.3. Projection Operators

The central objects of this paper are metric projections onto closed convex sets. We state their fundamental properties in the Hilbert-space setting.

Definition 2.3 (Metric projection). Let $C \subseteq H$ be a nonempty closed convex set. The *metric projection* $P_C : H \rightarrow C$ is defined by

$$P_C u = \arg \min_{v \in C} \|u - v\|. \quad (10)$$

The following theorem is the projection theorem in Hilbert spaces; see, e.g., Bauschke and Combettes [9].

Theorem 2.4 (Existence, uniqueness, and characterization). *Let $C \subseteq H$ be nonempty, closed, and convex. Then for every $u \in H$ the minimizer $P_C u$ in (10) exists and is unique. Moreover, $p = P_C u$ if and only if*

$$p \in C \quad \text{and} \quad \langle u - p, v - p \rangle \leq 0 \quad \forall v \in C. \quad (11)$$

Equivalently,

$$\|P_C u - v\|^2 \leq \|u - v\|^2 - \|u - P_C u\|^2 \quad \forall v \in C. \quad (12)$$

The characterization inequality (11) is often called the *variational inequality* for the projection.

Theorem 2.5 (Properties of the metric projection). *Let $C \subseteq H$ be nonempty, closed, and convex. Then:*

1. P_C is idempotent: $P_C \circ P_C = P_C$.
2. P_C is firmly nonexpansive, hence nonexpansive:

$$\|P_C u - P_C v\|^2 \leq \langle u - v, P_C u - P_C v \rangle \quad \forall u, v \in H. \quad (13)$$

3. If C is a closed subspace M , then P_M is linear, self-adjoint, and $P_M^2 = P_M$. In this case P_M is the orthogonal projection onto M , and $H = M \oplus M^\perp$.

Proof. (1) If $p = P_C u$, then $p \in C$, so $P_C p = p$ by the variational inequality (11) with $v = p$. (2) Let $p = P_C u$, $q = P_C v$. Applying (11) to (u, p) with $v = q$ and to (v, q) with $v = p$, then adding, yields

$$\langle u - p, q - p \rangle \leq 0, \quad \langle v - q, p - q \rangle \leq 0.$$

Adding and rearranging,

$$\langle u - v - (p - q), q - p \rangle \leq 0,$$

which is equivalent to

$$\|p - q\|^2 \leq \langle u - v, p - q \rangle.$$

This is (13). (3) The subspace case is the classical orthogonal projection theorem; see, e.g., Deutsch [8]. $\square$

Remark 2.6. The metric projection onto a closed convex set in a Banach space need not exist or be unique, and even when it exists it need not be nonexpansive. This is one reason we work in a Hilbert space. In $BV(\Omega)$, for instance, one cannot apply Theorem 2.5 directly.

2.4. Convex Analysis Tools

We collect the convex-analytic notions needed for proximal and epigraph-based methods.

Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be a function. Its *domain* is

$$\text{dom} J = \{u \in H : J(u) < +\infty\}. \quad (14)$$

The function J is *proper* if $\text{dom} J \neq \emptyset$. Its *epigraph* is

$$\text{epi}(J) = \{(u, t) \in H \times \mathbb{R} : J(u) \leq t\}. \quad (15)$$

The following equivalence is standard.

Proposition 2.7. *Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$. Then:*

1. *J is convex if and only if $\text{epi}(J)$ is a convex subset of $H \times \mathbb{R}$.*
2. *J is lower semicontinuous (l.s.c.) if and only if $\text{epi}(J)$ is closed in $H \times \mathbb{R}$.*

Thus a proper, convex, l.s.c. function has a nonempty closed convex epigraph. This is the structural fact behind the PESC framework in Section 4. See, e.g., Bauschke and Combettes [9].

For a proper convex function J , the *subdifferential* at $u \in \text{dom} J$ is the set

$$\partial J(u) = \{p \in H : J(v) \geq J(u) + \langle p, v - u \rangle \ \forall v \in H\}. \quad (16)$$

If J is differentiable at u , then $\partial J(u) = \{\nabla J(u)\}$. The subdifferential is a maximal monotone operator when J is proper, convex, and l.s.c. [9].

Definition 2.8 (Proximal operator). Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and l.s.c., and let $\lambda > 0$. The *proximal operator* of λJ is

$$\text{prox}_{\lambda J}(u) = \arg \min_{v \in H} \left\{ J(v) + \frac{1}{2\lambda} \|u - v\|^2 \right\}. \quad (17)$$

For every $u \in H$ the minimizer in (17) exists and is unique, because the objective is strongly convex and coercive. The proximal operator is firmly nonexpansive [9, 4, 13]. When $J = \iota_C$ is the indicator function of a nonempty closed convex set C (i.e., $\iota_C(u) = 0$ if $u \in C$ and $+\infty$ otherwise), the proximal operator reduces to the metric projection:

$$\text{prox}_{\lambda \iota_C} = P_C. \quad (18)$$

Thus the proximal operator generalizes the metric projection from sets to convex functions. This is the sense in which soft-thresholding is a proximal operator, not generally a projection onto an ℓ^1 -ball (see Section 3.3 for the precise relation).

A fundamental identity linking proximal operators and conjugates is the *Moreau decomposition*.

Theorem 2.9 (Moreau decomposition). *Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and l.s.c., and let $\lambda > 0$. Then for every $u \in H$,*

$$u = \text{prox}_{\lambda J}(u) + \lambda \text{prox}_{\lambda^{-1}J^*}\left(\frac{u}{\lambda}\right), \quad (19)$$

where J^* is the Fenchel conjugate of J ,

$$J^*(p) = \sup_{v \in H} \{\langle p, v \rangle - J(v)\}. \quad (20)$$

Equivalently,

$$u = \text{prox}_{\lambda J}(u) + \text{prox}_{(\lambda J)^*}(u). \quad (21)$$

The Moreau decomposition is proved, for example, in Bauschke and Combettes [9]; a detailed treatment with applications to signal processing is given by Combettes and Pesquet [10]. The decomposition is the key identity behind the equivalence between proximal methods and projection methods on epigraphs, and it will be used in Section 4.

3. Projection Operators in Variational Denoising

3.1. The General Variational Framework

Let H be a real Hilbert space and let $f \in H$ be the observed noisy image. A variational denoising method minimizes an energy of the form

$$E(u) = \lambda \Phi(u) + \frac{1}{2} \|f - u\|^2, \quad (22)$$

where $\Phi : H \rightarrow \mathbb{R} \cup \{+\infty\}$ is a proper, convex, lower semicontinuous regularization functional encoding prior knowledge, and $\lambda > 0$ is a regularization parameter. The quadratic fidelity term is the negative log-likelihood for additive Gaussian noise.

Proposition 3.1 (Existence and uniqueness). *Let $\Phi : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and l.s.c., and suppose Φ is coercive, i.e., $\Phi(u) \rightarrow +\infty$ whenever $\|u\| \rightarrow \infty$. Then for every $f \in H$ and $\lambda > 0$ the problem (22) admits at least one minimizer. If Φ is strictly convex, the minimizer is unique.*

Proof. The functional E is proper, convex, and l.s.c. as a sum of such functionals. The fidelity term is coercive and the coercivity of Φ transfers to E . A proper, coercive, l.s.c., convex functional on a reflexive Banach space attains its minimum; (see, Ekeland and Temam [22]). Strict convexity of Φ implies strict convexity of E , hence uniqueness. $\square$

The first-order optimality condition for (22) is

$$0 \in \lambda \partial \Phi(u) + u - f, \quad (23)$$

which can be rewritten as the fixed-point equation

$$u = (I + \lambda \partial \Phi)^{-1}(f) = \text{prox}_{\lambda \Phi}(f). \quad (24)$$

Thus every variational denoising problem of the form (22) is solved by evaluating a proximal operator at the noisy image f . When Φ is the indicator function of a closed convex set, this reduces to a metric projection. This observation is the entry point for the projection-based interpretation of TV, wavelet, and POCS denoising.

3.2. Total Variation Denoising and Chambolle's Projection Algorithm

The Rudin–Osher–Fatemi (ROF) model [1] takes $\Phi(u) = TV(u)$ in (22):

$$\min_{u \in BV(\Omega)} TV(u) + \frac{\lambda}{2} \|f - u\|_{L^2}^2. \quad (25)$$

We work in the discrete setting $H = \mathbb{R}^{M \times N}$ for the algorithmic development, since this is the setting implemented numerically. The continuous formulation is obtained by replacing discrete gradients and divergences by their distributional counterparts.

3.2.1 Dual Formulation

Let $\nabla : \mathbb{R}^{M \times N} \rightarrow \mathbb{R}^{M \times N \times 2}$ denote the discrete forward-difference gradient and let $\text{div} = -\nabla^*$ denote its adjoint (the discrete divergence). The discrete total variation is

$$TV(u) = \|\nabla u\|_{2,1} = \sum_{i,j} \sqrt{(\nabla u)_{i,j}^1{}^2 + (\nabla u)_{i,j}^2{}^2}. \quad (26)$$

Using the dual representation of the $\ell_{2,1}$ -norm,

$$\|\nabla u\|_{2,1} = \sup_{p \in K} \langle \nabla u, p \rangle, \quad (27)$$

where the constraint set is

$$K = \{p \in \mathbb{R}^{M \times N \times 2} : |p_{i,j}| \leq 1 \ \forall (i,j)\}, \quad (28)$$

and $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^2$. The set K is a nonempty, closed, convex, and bounded subset of the finite-dimensional Hilbert space $\mathbb{R}^{M \times N \times 2}$.

Substituting (27) into (25) and exchanging the supremum over $p \in K$ with the minimization over u (justified by the minimax theorem, since the functional is convex in u and concave in p on a compact convex set), we obtain the dual problem

$$\min_{p \in K} \frac{1}{2} \|\lambda f - \text{div } p\|_{L^2}^2. \quad (29)$$

Equivalently, setting $v = \lambda f - \text{div } p$, the dual problem is to compute the projection of λf onto the set

$$C = \text{div}(K) = \{\text{div } p : p \in K\}, \quad (30)$$

that is,

$$v^* = P_C(\lambda f). \quad (31)$$

The primal solution is recovered from the dual solution via the optimality condition

$$u^* = f - \frac{1}{\lambda} \text{div } p^*, \quad (32)$$

where $p^* \in K$ is any dual minimizer.

3.2.2 Chambolle's Projection Algorithm

Chambolle [2] proposed the following iterative scheme for the dual problem (29):

$$p^{k+1} = P_K(p^k + \tau \nabla(\text{div } p^k - \lambda f)), \quad (33)$$

where P_K is the pointwise projection onto the unit ball:

$$(P_K(p))_{i,j} = \frac{p_{i,j}}{\max(1, |p_{i,j}|)}. \quad (34)$$

The iteration (33) is a projected gradient descent on the dual objective

$$J(p) = \frac{1}{2} \|\lambda f - \text{div } p\|_{L^2}^2, \quad (35)$$

whose gradient is

$$\nabla J(p) = -\nabla(\text{div } p - \lambda f). \quad (36)$$

Theorem 3.2 (Convergence of Chambolle’s algorithm). *Let $p^0 \in K$ be arbitrary. If the step size satisfies $0 < \tau \leq 1/8$, then the sequence $\{p^k\}$ generated by (33) converges in $\mathbb{R}^{M \times N \times 2}$ to a minimizer p^* of the dual problem (29). Consequently,*

$$\operatorname{div} p^k \rightarrow P_C(\lambda f) \quad \text{as } k \rightarrow \infty. \quad (37)$$

Proof sketch. The dual objective J is convex and differentiable with Lipschitz gradient. The Lipschitz constant of ∇J satisfies $L = \|\nabla\|^2 \leq 8$ in the discrete setting. The projected gradient iteration with step size $\tau \in (0, 1/L]$ converges to a minimizer for convex problems with a closed convex constraint set; see, e.g., Bauschke and Combettes [9]. The factor $1/8$ in the step-size condition arises from the bound $\|\operatorname{div}\|^2 = \|\nabla\|^2 \leq 8$ for the standard forward-difference discretization. $\square$

Remark 3.3. The operator P_K in (33) is a metric projection onto a closed convex set in a Hilbert space. Therefore Chambolle’s algorithm is an instance of the general projected-gradient framework, and its convergence is a special case of Theorem 2.5 combined with standard convex optimization theory.

3.3. Wavelet Shrinkage as a Proximal Operator

Let $\mathcal{W} : H \rightarrow H$ be an orthonormal wavelet transform, so that $\mathcal{W}^* \mathcal{W} = I$. For an image u , let $\theta = \mathcal{W}u$ denote its wavelet coefficients. Wavelet denoising in the coefficient domain solves

$$\min_{\theta \in H} \frac{1}{2} \|\theta - \mathcal{W}f\|^2 + \lambda \|\theta\|_1, \quad (38)$$

where $\|\theta\|_1 = \sum_i |\theta_i|$. The solution is the soft-thresholding operator

$$S_\tau(\theta)_i = \operatorname{sign}(\theta_i) \max(|\theta_i| - \tau, 0), \quad (39)$$

with threshold $\tau = \lambda$.

Proposition 3.4 (Soft-thresholding is the proximal operator of the ℓ^1 -norm). *For every $\tau > 0$ and every $\theta \in H$,*

$$S_\tau(\theta) = \operatorname{prox}_{\tau\|\cdot\|_1}(\theta) = \arg \min_{\xi \in H} \left\{ \tau \|\xi\|_1 + \frac{1}{2} \|\theta - \xi\|^2 \right\}. \quad (40)$$

Proof. The objective is separable across coordinates. For a scalar $t \in \mathbb{R}$, the problem

$$\min_{\xi \in \mathbb{R}} \tau |\xi| + \frac{1}{2} (t - \xi)^2$$

is solved by the scalar soft-thresholding $S_\tau(t)$. The subdifferential optimality condition $0 \in \tau \partial |\xi| + \xi - t$ yields $\xi = 0$ if $|t| \leq \tau$ and $\xi = t - \tau \operatorname{sign}(t)$ otherwise. Applying this coordinatewise gives (40). $\square$

Remark 3.5 (Relation to projection onto the ℓ^1 -ball). Soft-thresholding is *not* the metric projection onto the ℓ^1 -ball $B_R = \{\xi : \|\xi\|_1 \leq R\}$. The projection $P_{B_R}(\theta)$ is obtained by solving

$$P_{B_R}(\theta) = \arg \min_{\|\xi\|_1 \leq R} \|\theta - \xi\|^2. \quad (41)$$

By the KKT conditions, the solution has the form

$$P_{B_R}(\theta) = S_\mu(\theta) \quad (42)$$

for some $\mu \geq 0$ chosen so that $\|S_\mu(\theta)\|_1 = R$ when $\|\theta\|_1 > R$, and $P_{B_R}(\theta) = \theta$ when $\|\theta\|_1 \leq R$. Thus the projection onto the ℓ^1 -ball is a soft-thresholding with a *data-dependent* threshold μ , not a fixed one. The proximal operator $\text{prox}_{\tau\|\cdot\|_1}$ uses a fixed threshold τ . The two coincide only in the degenerate case where the solution of (41) happens to lie on the boundary $\|\xi\|_1 = R$ with $\mu = \tau$.

The wavelet denoising method therefore fits into the proximal framework: the denoised image is

$$u^* = \mathcal{W}^* S_\lambda(\mathcal{W} f). \quad (43)$$

This is a proximal iteration in the wavelet coefficient domain, and it is a projection only if the regularization is replaced by a hard constraint $\|\theta\|_1 \leq R$.

3.4. Projection onto Convex Sets (POCS)

Let $C_1, \dots, C_m \subseteq H$ be nonempty closed convex sets whose intersection $C = \bigcap_{i=1}^m C_i$ is nonempty. The convex feasibility problem is to find a point $u^* \in C$. The method of projection onto convex sets (POCS), also known as the cyclic projection algorithm, generates the sequence

$$u^{k+1} = P_{C_m} P_{C_{m-1}} \cdots P_{C_1} u^k. \quad (44)$$

Theorem 3.6 (Convergence of cyclic projections). *Let $C_1, \dots, C_m \subseteq H$ be nonempty closed convex sets with $C = \bigcap_{i=1}^m C_i \neq \emptyset$. For any $u^0 \in H$, the sequence $\{u^k\}$ generated by (44) converges weakly to a point $u^* \in C$. If H is finite-dimensional, the convergence is strong. If each C_i is an affine subspace, the limit is the orthogonal projection of u^0 onto C .*

Proof. Each P_{C_i} is firmly nonexpansive by Theorem 2.5, hence nonexpansive. The composition $T = P_{C_m} \circ \cdots \circ P_{C_1}$ is nonexpansive. The fixed-point set of T contains C . The sequence (44) is the fixed-point iteration $u^{k+1} = Tu^k$. By the Krasnosel'skii–Mann theorem for nonexpansive operators in Hilbert spaces, the sequence converges weakly to a fixed point of T ; see Bauschke and Borwein [7] and Bauschke and Combettes [9]. In finite dimensions, weak and strong convergence coincide. The affine case is the classical von Neumann–Halperin theorem. $\square$

Remark 3.7. In image denoising, the sets C_i encode known properties of the clean image. Representative examples include:

1. *Intensity range:* $C_1 = \{u : a \leq u_{i,j} \leq b\}$ for known bounds a, b .
2. *Fidelity constraint:* $C_2 = \{u : \|u - f\|_2 \leq \varepsilon\}$ for an estimate ε of the noise level.
3. *Sparsity in a transform domain:* $C_3 = \{u : \|\mathcal{W}u\|_1 \leq R\}$ for a wavelet transform $\mathcal{W}$ and a sparsity budget R .

The POCS iteration (44) alternates between these constraints. The projection onto C_3 is not soft-thresholding in general, by Remark 3.5; it requires solving (41) for the data-dependent threshold μ .

Table 1: Projection and proximal operators in variational denoising.

Method	Operator	Fixed-point problem	Assumptions	Convergence
ROF / TV	P_K	$p = P_K(p + \tau \nabla(\operatorname{div} p - \lambda f))$	K closed convex, $\tau \leq 1/8$	Strong (finite dim.)
Wavelet	$S_\lambda = \operatorname{prox}_{\lambda \ \cdot\ _1}$	$u = \mathcal{W}^* S_\lambda(\mathcal{W} f)$	Orthonormal $\mathcal{W}$	Direct (one step)
POCS	$P_{C_m} \cdots P_{C_1}$	$u \in \bigcap_{i=1}^m C_i$	C_i closed convex, $\bigcap_i C_i \neq \emptyset$	Weak; strong if finite dim.

The table summarizes the operator-theoretic structure of the three methods. All three are instances of the general scheme: solve a convex problem by applying a firmly nonexpansive operator (projection or proximal map) iteratively. This observation is the foundation for the unified convergence analysis in Section 5.

4. Advanced Projection-Based Frameworks

This section treats two modern developments that extend the projection paradigm beyond classical variational denoising: the projection onto epigraph sets (PESC) and regularization by denoising via fixed-point projection (RED-PRO). Both fit within the operator-theoretic framework of Section 2.

4.1. Projection onto Epigraph Sets (PESC)

Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be a proper, convex, lower semicontinuous cost function on a real Hilbert space H . By Proposition 2.7, its epigraph

$$\operatorname{epi}(J) = \{(u, t) \in H \times \mathbb{R} : J(u) \leq t\} \quad (45)$$

is a nonempty closed convex subset of the product Hilbert space $H \times \mathbb{R}$, equipped with the inner product

$$\langle (u, t), (v, s) \rangle_{H \times \mathbb{R}} = \langle u, v \rangle_H + ts. \quad (46)$$

4.1.1 The Feasibility Reformulation

Consider the constrained minimization problem

$$\min_{u \in H} J(u) \quad \text{subject to } u \in C, \quad (47)$$

where $C \subseteq H$ is a nonempty closed convex set encoding linear constraints. The PESC framework [25] reformulates (47) as the convex feasibility problem of finding a point in the intersection

$$\operatorname{epi}(J) \cap \mathcal{H}_C \neq \emptyset, \quad (48)$$

where

$$\mathcal{H}_C = \{(u, t) \in H \times \mathbb{R} : u \in C, t \in \mathbb{R}\} \quad (49)$$

is a closed convex cylinder over C (and a hyperplane in the lifted space when C is an affine subspace). The equivalence is immediate: $(u^*, t^*) \in \operatorname{epi}(J) \cap \mathcal{H}_C$ with $t^* = J(u^*)$ yields the minimizer u^* of (47), provided the intersection is nonempty.

4.1.2 The PESC Algorithm

The PESC algorithm applies the method of cyclic projections to the family of closed convex sets whose intersection defines the feasible region. In the lifted space $H \times \mathbb{R}$, the algorithm generates

$$(u^{k+1}, t^{k+1}) = P_{\text{epi}(J)} \circ P_{\mathcal{H}_C}(u^k, t^k), \quad (50)$$

or, in the more common implementation, alternates between projections onto supporting hyperplanes of $\text{epi}(J)$ and projections onto the constraint set C .

When J is the total variation or the ℓ^1 -norm, the projection onto $\text{epi}(J)$ is not available in closed form. In this case, the projection is computed by successive orthogonal projections onto supporting hyperplanes of $\text{epi}(J)$ [12]. Specifically, given a point $(u, t) \notin \text{epi}(J)$, one finds a subgradient $g \in \partial J(u)$ and projects onto the hyperplane

$$H_g = \{(v, s) \in H \times \mathbb{R} : J(u) + \langle g, v - u \rangle = s\}. \quad (51)$$

The projection onto H_g is available in closed form:

$$P_{H_g}(v, s) = (v, s) - \frac{J(u) + \langle g, v - u \rangle - s}{1 + \|g\|^2}(g, -1). \quad (52)$$

Iterating these hyperplane projections converges to the projection onto $\text{epi}(J)$ under standard conditions.

Theorem 4.1 (Convergence of PESC). *Let $J : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and l.s.c., and let $C \subseteq H$ be nonempty, closed, and convex. Suppose the intersection (48) is nonempty. Then the sequence $\{(u^k, t^k)\}$ generated by (50) converges weakly to a point in $\text{epi}(J) \cap \mathcal{H}_C$. In finite dimensions, the convergence is strong. Consequently, the sequence $\{u^k\}$ converges to a minimizer of (47).*

Proof. The sets $\text{epi}(J)$ and $\mathcal{H}_C$ are nonempty closed convex subsets of $H \times \mathbb{R}$ by Proposition 2.7. The PESC iteration (50) is a cyclic projection scheme onto two closed convex sets. The convergence therefore follows from Theorem 3.6. When the epigraph projection is computed via hyperplane projections, the same conclusion holds by the convergence theory for projection onto supporting hyperplanes of a convex set [9]. In finite dimensions, weak and strong convergence coincide. $\square$

Remark 4.2. The PESC framework applies to any convex cost function whose epigraph is closed and convex. In addition to total variation and the ℓ^1 -norm, it has been applied to the ℓ^2 -norm, filtered variation, and entropic cost functions [25, 12]. The main computational cost lies in evaluating the subgradient of J and computing the scalar hyperplane projection (52).

4.2. Regularization by Denoising via Fixed-Point Projection (RED-PRO)

The Plug-and-Play (PnP) framework [15] and Regularization by Denoising (RED) [14] replace an explicit regularizer by a state-of-the-art denoiser. RED-PRO [3, 23] provides a convex optimization formulation of this idea based on projection onto the fixed-point set of a demicontractive denoiser, together with convergence guarantees to globally optimal solutions.

4.2.1 Demicontractive Denoisers

Let $D : H \rightarrow H$ be a denoiser and let

$$\text{Fix}(D) = \{u \in H : Du = u\} \quad (53)$$

denote its fixed-point set.

Definition 4.3 (Demicontractive operator). An operator $D : H \rightarrow H$ with $\text{Fix}(D) \neq \emptyset$ is β -demicontractive for some $\beta \in [0, 1)$ if

$$\|Du - z\|^2 \leq \|u - z\|^2 + \beta \|u - Du\|^2 \quad \forall u \in H, \forall z \in \text{Fix}(D). \quad (54)$$

Remark 4.4. The class of demicontractive operators strictly contains the class of nonexpansive operators with nonempty fixed-point sets. If D is nonexpansive and $\text{Fix}(D) \neq \emptyset$, then (54) holds with $\beta = 0$. Conversely, demicontractive operators need not be nonexpansive. This is the precise sense in which RED-PRO extends the convergence theory of PnP beyond nonexpansive denoisers [3].

Proposition 4.5 (Closed convexity of the fixed-point set). *Let $D : H \rightarrow H$ be β -demicontractive with $\text{Fix}(D) \neq \emptyset$. If D is continuous, then $\text{Fix}(D)$ is closed and convex.*

Proposition 4.5 is stated and proved in [3]. It is the structural property that makes the projection onto $\text{Fix}(D)$ well-defined in the Hilbert-space sense of Definition 2.3.

4.2.2 The RED-PRO Problem

Given a convex, differentiable, l.s.c., and proper data-fidelity function $\ell : H \rightarrow \mathbb{R}$, RED-PRO solves

$$\min_{u \in H} \ell(u) \quad \text{subject to } u \in \text{Fix}(D). \quad (55)$$

For the denoising task with quadratic fidelity $\ell(u) = \frac{1}{2} \|f - u\|^2$, problem (55) reduces to

$$\min_{u \in \text{Fix}(D)} \frac{1}{2} \|f - u\|^2, \quad (56)$$

whose solution is the metric projection of f onto $\text{Fix}(D)$:

$$u^* = P_{\text{Fix}(D)}(f). \quad (57)$$

Remark 4.6. The projection $P_{\text{Fix}(D)}$ is not directly computable for a general denoiser. RED-PRO therefore does not evaluate $P_{\text{Fix}(D)}$ in closed form. Instead, it exploits the fact that for a β -demicontractive operator D and any $\mu \in (0, 1 - \beta]$, the operator

$$T_\mu = P_{\text{Fix}(D)} \circ (I - \mu(I - D)) \quad (58)$$

is quasi-nonexpansive and satisfies $\text{Fix}(T_\mu) = \text{Fix}(D)$. This auxiliary operator is used to construct a convergent iterative scheme without explicit knowledge of $P_{\text{Fix}(D)}$ [3].

4.2.3 The RED-PRO Iteration

The RED-PRO algorithm solves (55) by a projected gradient method on the fixed-point set. The iteration is

$$u^{k+1} = P_{\text{Fix}(D)}(u^k - \mu_k \nabla \ell(u^k)), \quad (59)$$

where $\{\mu_k\}$ is a step-size sequence. For the denoising problem (56) with quadratic fidelity $\ell(u) = \frac{1}{2}\|f - u\|^2$, the gradient is $\nabla \ell(u) = u - f$, and (59) becomes

$$u^{k+1} = P_{\text{Fix}(D)}((1 - \mu_k)u^k + \mu_k f). \quad (60)$$

The term $(1 - \mu_k)u^k + \mu_k f$ is a convex combination of the current iterate and the observed image, and it is what anchors the iteration to the data. Without this term, the scheme reduces to a pure Mann iteration toward a fixed point of D , which does not solve (55) in general.

In practice $P_{\text{Fix}(D)}$ is not computable for a general denoiser. RED-PRO replaces it by D itself, which acts as a computable proxy whose fixed-point set coincides with $\text{Fix}(D)$. The implementable form of (60) is therefore

$$u^{k+1} = D((1 - \mu_k)u^k + \mu_k f). \quad (61)$$

This is the scheme used in the numerical experiments of Section 6, with $\mu_k = \tau = 0.5$ and $k = 0, \dots, 4$.

Theorem 4.7 (Convergence of RED-PRO). *Let $D : H \rightarrow H$ be a continuous β -demicontractive denoiser with $\beta \in [0, 1)$, and suppose $\text{Fix}(D) \neq \emptyset$. Let $\ell : H \rightarrow \mathbb{R}$ be convex, differentiable with L -Lipschitz gradient, l.s.c., and proper. Suppose the step-size sequence $\{\mu_k\}$ satisfies*

$$0 < \liminf_{k \rightarrow \infty} \mu_k \leq \limsup_{k \rightarrow \infty} \mu_k < \frac{2 - \beta}{L}. \quad (62)$$

Then the sequence $\{u^k\}$ generated by (60) converges weakly to a minimizer of (55). In finite dimensions, the convergence is strong. If the minimizer is unique, the convergence is strong in any Hilbert space.

Theorem 4.7 is the main result of [3]. Its proof combines the Fejér monotonicity of the sequence generated by (60) with the quasi-nonexpansiveness of the auxiliary operator (58) and the convergence theory for averaged nonexpansive operators in Hilbert spaces [9].

Remark 4.8 (Relation to PnP and RED). RED-PRO unifies PnP and RED in the following sense. When the denoiser D is the proximal operator of a convex regularizer R , i.e., $D = \text{prox}_{\lambda R}$, the fixed-point set $\text{Fix}(D)$ coincides with the set of minimizers of R , and RED-PRO reduces to standard proximal gradient descent on R . When D is a general demicontractive denoiser, RED-PRO retains the global convergence guarantee of proximal methods while allowing denoisers that are not proximal operators, such as BM3D or DnCNN [3].

5. Theoretical Analysis

5.1. Existence and Uniqueness

We begin with the general variational problem

$$\min_{u \in H} E(u) = \lambda \Phi(u) + \frac{1}{2}\|f - u\|^2, \quad (63)$$

where H is a real Hilbert space, $f \in H$ is fixed, $\lambda > 0$, and $\Phi : H \rightarrow \mathbb{R} \cup \{+\infty\}$ is a regularization functional. The following result is the standard existence and uniqueness theorem for convex variational problems.

Theorem 5.1 (Existence and uniqueness). *Let $\Phi : H \rightarrow \mathbb{R} \cup \{+\infty\}$ be proper, convex, and lower semicontinuous. Suppose Φ is coercive, i.e.,*

$$\lim_{\|u\| \rightarrow \infty} \Phi(u) = +\infty. \quad (64)$$

Then for every $f \in H$ and every $\lambda > 0$, problem (63) admits at least one minimizer. If, in addition, Φ is strictly convex, the minimizer is unique.

Proof. The functional E is proper, convex, and l.s.c. as a sum of such functionals. The quadratic fidelity term is coercive and bounded below. The coercivity of Φ transfers to E : if $\|u_n\| \rightarrow \infty$, then $\Phi(u_n) \rightarrow +\infty$, and since the fidelity term is nonnegative, $E(u_n) \rightarrow +\infty$. A proper, coercive, l.s.c., convex functional on a reflexive Banach space attains its minimum; see Ekeland and Temam [22]. This proves existence. Strict convexity of Φ implies strict convexity of E , hence uniqueness of the minimizer. $\square$

Remark 5.2. For the ROF model (25), the regularization $\Phi = TV$ is convex, l.s.c., and coercive on the quotient space $BV(\Omega)/\mathbb{R}$. On $BV(\Omega) \cap L^2(\Omega)$ with the L^2 norm, the addition of the fidelity term makes the functional coercive. Existence and uniqueness in this setting were established by Rudin, Osher, and Fatemi [1] and later refined by Chambolle [2]. The uniqueness holds because the functional is strictly convex in $L^2(\Omega)$.

5.2. Convergence of Iterative Projection Algorithms

5.2.1 Cyclic Projections onto Convex Sets

The POCS algorithm of Section 3.4 generates the sequence

$$u^{k+1} = P_{C_m} P_{C_{m-1}} \cdots P_{C_1} u^k. \quad (65)$$

The following theorem is the classical convergence result for cyclic projections; see Bauschke and Borwein [7] and Bauschke and Combettes [9].

Theorem 5.3 (Convergence of cyclic projections). *Let $C_1, \dots, C_m \subseteq H$ be nonempty closed convex sets with $C = \bigcap_{i=1}^m C_i \neq \emptyset$. For any $u^0 \in H$, the sequence $\{u^k\}$ generated by (65) converges weakly to a point $u^* \in C$. If H is finite-dimensional, the convergence is strong. If each C_i is an affine subspace, the limit is the orthogonal projection of u^0 onto C .*

Proof. Each P_{C_i} is firmly nonexpansive by Theorem 2.5, hence nonexpansive. The composition $T = P_{C_m} \circ \cdots \circ P_{C_1}$ is nonexpansive and its fixed-point set contains C . The sequence (65) is the fixed-point iteration $u^{k+1} = Tu^k$. By the Krasnosel'skii–Mann theorem for nonexpansive operators in Hilbert spaces, the sequence converges weakly to a fixed point of T ; see Bauschke and Borwein [7] and Bauschke and Combettes [9]. In finite dimensions, weak and strong convergence coincide. The affine case is the von Neumann–Halperin theorem. $\square$

5.2.2 Convergence of RED-PRO

The RED-PRO algorithm of Section 4.2 solves

$$\min_{u \in \text{Fix}(D)} \frac{1}{2} \|f - u\|^2, \quad (66)$$

where $D : H \rightarrow H$ is a β -demicontractive denoiser with $\beta \in [0, 1)$ and $\text{Fix}(D) \neq \emptyset$. The convergence theory was developed by Cohen, Elad, and Milanfar [3]. We state the main result.

Theorem 5.4 (Convergence of RED-PRO). *Let $D : H \rightarrow H$ be a continuous β -demicontractive denoiser with $\beta \in [0, 1)$, and let $\text{Fix}(D) \neq \emptyset$. Suppose $\text{Fix}(D)$ is closed and convex, so that the metric projection $P_{\text{Fix}(D)}$ is well-defined. Let $\{\rho_k\} \subset (0, 1]$ satisfy*

$$0 < \liminf_{k \rightarrow \infty} \rho_k \leq \limsup_{k \rightarrow \infty} \rho_k < 2 - \beta. \quad (67)$$

Then for any $u^0 \in H$, the sequence

$$u^{k+1} = u^k + \rho_k (P_{\text{Fix}(D)}(u^k) - u^k) \quad (68)$$

converges weakly to a minimizer of (66). In finite dimensions, the convergence is strong. If the minimizer is unique, the convergence is strong in any Hilbert space.

Proof sketch. The proof in [3] proceeds in three steps. First, one shows that the auxiliary operator

$$T_\mu = P_{\text{Fix}(D)} \circ (I - \mu(I - D))$$

with $\mu \in (0, 1 - \beta]$ is quasi-nonexpansive and satisfies $\text{Fix}(T_\mu) = \text{Fix}(D)$. Second, the iteration (68) is rewritten as a relaxed iteration involving T_μ . Third, the Fejér monotonicity of the sequence with respect to $\text{Fix}(D)$, together with the step-size condition (67), yields weak convergence to a fixed point of T_μ , which is a minimizer of (66). In finite dimensions, the weak limit is strong by the uniqueness of the limit. $\square$

Remark 5.5. The hypothesis that $\text{Fix}(D)$ is closed and convex is essential for the metric projection $P_{\text{Fix}(D)}$ to be well-defined. Cohen et al. [3] prove that for a continuous demicontractive operator, the fixed-point set is indeed closed and convex; see also Proposition 4.5. The demicontractive condition (54) is weaker than nonexpansiveness and allows denoisers that are not proximal operators, such as BM3D or DnCNN, to be used within the framework.

5.2.3 Unified Perspective

The convergence theorems above share a common structure. Each algorithm generates a sequence by applying an averaged or quasi-nonexpansive operator T repeatedly:

$$u^{k+1} = u^k + \rho_k (Tu^k - u^k). \quad (69)$$

When T is nonexpansive with a nonempty fixed-point set and $\{\rho_k\}$ satisfies the standard conditions, the sequence converges weakly to a fixed point of T in a Hilbert space [4, 19]. The POCS theorem is the case $T = P_{C_m} \circ \dots \circ P_{C_1}$, the RED-PRO theorem is the case $T = P_{\text{Fix}(D)}$, and Chambolle's algorithm fits the same template in the dual space. This observation is the conceptual core of the unified framework.

5.3. Consistency of Continuous and Discrete Formulations

In practice, images are discrete arrays, and the projection operators are implemented on finite-dimensional spaces. The theoretical results of the previous subsections are stated in a Hilbert space H , which may be infinite-dimensional in the continuous model. It is therefore necessary to verify that the discrete solutions converge to the continuous solution as the grid size tends to zero.

For total variation denoising, this consistency has been established rigorously. Chambolle and coauthors [20, 21] proved that the discrete ROF minimizer converges to the continuous minimizer in the L^2 topology as the mesh size tends to zero, under appropriate boundary conditions and with a suitable discretization of the total variation. The convergence is obtained via Γ -convergence of the discrete functionals to the continuous one.

Theorem 5.6 (Consistency of TV discretization). *Let $\Omega \subset \mathbb{R}^2$ be a bounded open set with Lipschitz boundary. Let $\{T_h\}$ be a family of triangulations of Ω with mesh size $h \rightarrow 0$. Let E_h denote the discrete ROF functional on the piecewise-constant finite element space associated with T_h , and let E denote the continuous ROF functional on $BV(\Omega) \cap L^2(\Omega)$. Then E_h Γ -converges to E as $h \rightarrow 0$ with respect to the $L^2(\Omega)$ topology. Consequently, the discrete minimizers converge in $L^2(\Omega)$ to the continuous minimizer, and the minimal energies converge.*

Proof. See Caselles et al. [20] and Chambolle et al. [21]. The proof relies on the representation of the total variation as a supremum over test vector fields and on the density of piecewise-constant functions in $BV(\Omega)$. $\square$

Remark 5.7. For the projection-based algorithms, the consistency of the continuous and discrete formulations follows from the Γ -convergence of the underlying variational problems. When the discrete problem is solved by a projection algorithm, the discrete iterates converge to the discrete minimizer, and the discrete minimizer converges to the continuous minimizer as $h \rightarrow 0$. Thus the projection framework is consistent with the continuous model in the limit of fine discretization.

5.4. Summary

The theoretical analysis of this section establishes three facts. First, the variational denoising problem (63) admits a unique minimizer under standard convexity and coercivity assumptions. Second, the projection-based algorithms of Sections 3 and 4 converge weakly to a solution in a Hilbert space, and strongly in finite dimensions. Third, the discrete implementations are consistent with the continuous models as the grid size tends to zero. Together, these results provide the rigorous foundation for the unified projection framework.

6. Numerical Experiments

6.1. Setup

We compare the four projection-based methods of Sections 3 and 4 on three standard 256×256 grayscale images (Cameraman, Astronaut, Coffee) corrupted by additive white Gaussian noise with $\sigma \in \{10, 20, 30\}$. Each configuration is repeated over two noise realizations with fixed seed. The methods are:

- Chambolle's dual projection algorithm for ROF TV denoising [2], weight $w = 1.2\sigma/255$ and a maximum of 150 iterations;

- wavelet soft-thresholding with the Daubechies-8 wavelet, three decomposition levels, and the universal threshold $\hat{\sigma}\sqrt{2\log N}$ where $\hat{\sigma}$ is the median absolute deviation of the finest-scale detail coefficients [16];
- POCS alternating between the intensity box $[0, 1]$ and the fidelity ball $\|u - f\|_2 \leq \varepsilon$ with $\varepsilon = 1.1\sigma\sqrt{N}/255$, 60 iterations;
- RED-PRO with a BM3D denoiser [17] in the proximal-gradient form

$$u^{k+1} = D((1 - \tau)u^k + \tau f),$$

with $\tau = 0.5$ and five iterations [3].

Performance is measured by PSNR and SSIM with `data_range=1`. All experiments use Python 3.11 with NumPy, SciPy, scikit-image, PyWavelets, and the BM3D package.

6.2. Visual Comparison

Figure 1 shows a representative result on Cameraman at $\sigma = 25$. The top row contains the clean image, the noisy observation, and the four reconstructions; the bottom row shows the absolute error maps on the scale $[0, 0.3]$. RED-PRO produces the smallest error, and the residual is concentrated along the strong contours of the image where all four methods lose accuracy. TV follows, with the characteristic staircase effect along smooth regions. Wavelet shrinkage and POCS leave visible mid-frequency noise and blur the fine details.

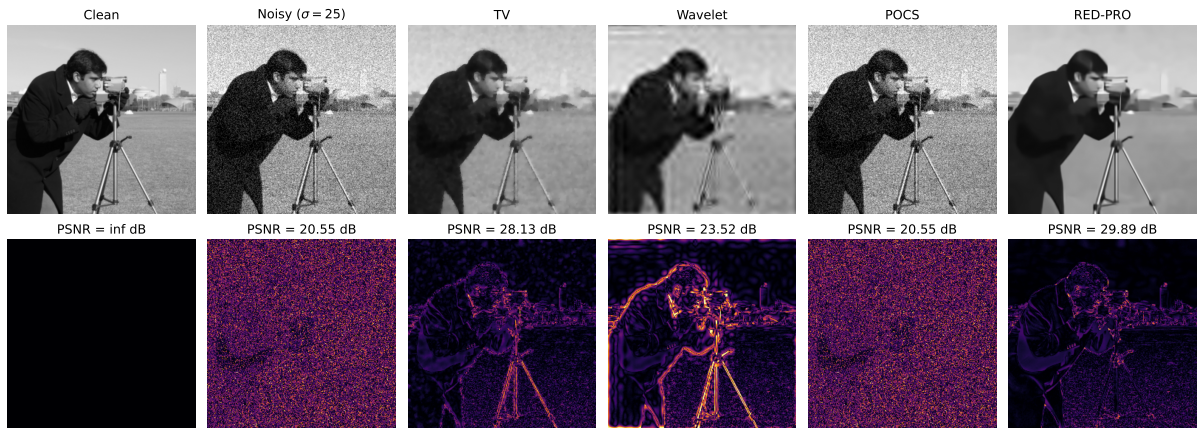


Figure 1: Visual comparison at $\sigma = 25$ on Cameraman. Top: clean, noisy, and denoised outputs. Bottom: absolute error maps (inferno colormap, $[0, 0.3]$).

6.3. Quantitative Results

Figure 2 reports PSNR and SSIM averaged over the three images and two noise realizations, with ± 1 standard deviation bands. RED-PRO attains the highest PSNR at every noise level, followed by TV. The margin of RED-PRO over TV is +1.67 dB on average at $\sigma = 10$, +1.24 dB at $\sigma = 20$, and +1.22 dB at $\sigma = 30$. TV and RED-PRO are the only two methods that degrade gracefully with σ ; wavelet shrinkage loses more than 2 dB per 10 units of σ , and POCS drops by as much as 9 dB between $\sigma = 10$ and $\sigma = 30$ on Cameraman. Table 2 lists the per-image PSNR.

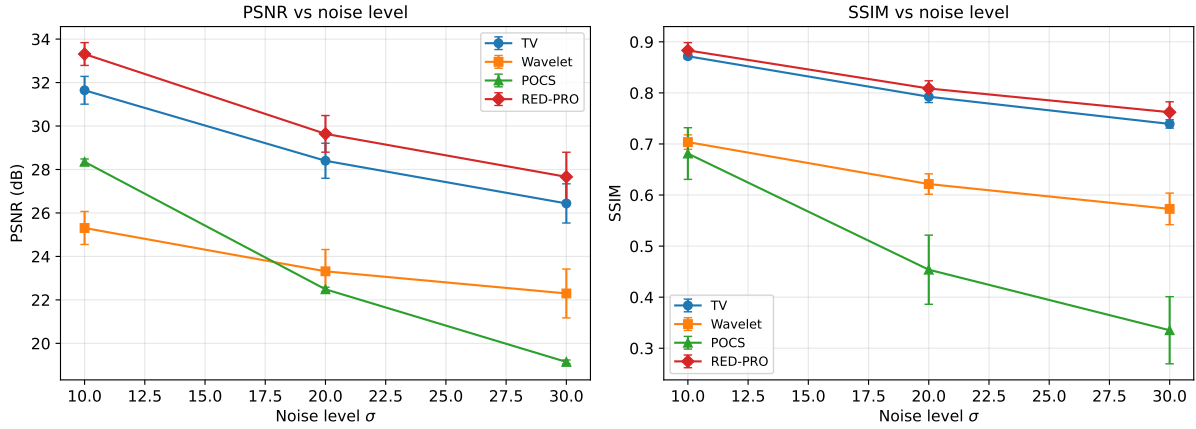


Figure 2: PSNR and SSIM versus noise level, averaged over images and realizations. Bands indicate ± 1 standard deviation.

Table 2: Mean PSNR (dB) over two noise realizations. Boldface marks the best method in each row.

Image	σ	TV	Wavelet	POCS	RED-PRO
Cameraman	10	32.55	25.99	28.25	33.42
	20	29.30	24.05	22.42	30.83
	30	27.31	23.08	19.14	29.13
Astronaut	10	31.22	24.25	28.54	33.89
	20	27.34	21.90	22.62	29.01
	30	25.20	20.71	19.25	26.38
Coffee	10	31.16	25.69	28.27	32.62
	20	28.57	24.00	22.42	29.08
	30	26.81	23.10	19.04	27.48

Three features of Table 2 deserve comment. First, RED-PRO wins on every row, and the win is largest on Astronaut at $\sigma = 10$ (+2.67 dB over TV). Second, the ranking of the two classical methods is not fixed: POCS beats wavelet shrinkage at $\sigma = 10$, but wavelet shrinkage overtakes POCS at $\sigma = 20$ and $\sigma = 30$ on every image except Astronaut at $\sigma = 20$, where the two are within 0.7 dB. Third, TV is the strongest classical baseline at every noise level, which is consistent with the well-known difficulty of beating TV at low noise on natural images [2].

6.4. Convergence

Figure 3 shows PSNR versus iteration for TV and RED-PRO on Cameraman at $\sigma = 25$, with the iteration axis on a logarithmic scale. Chambolle’s dual projection improves steadily over the first 20 iterations and flattens after roughly 80. RED-PRO reaches its operating point within the first two iterations and remains stable for the remaining steps, because the BM3D denoiser already performs most of the reconstruction in a single application. Both curves are monotone, in agreement with the convergence theory of Section 5.

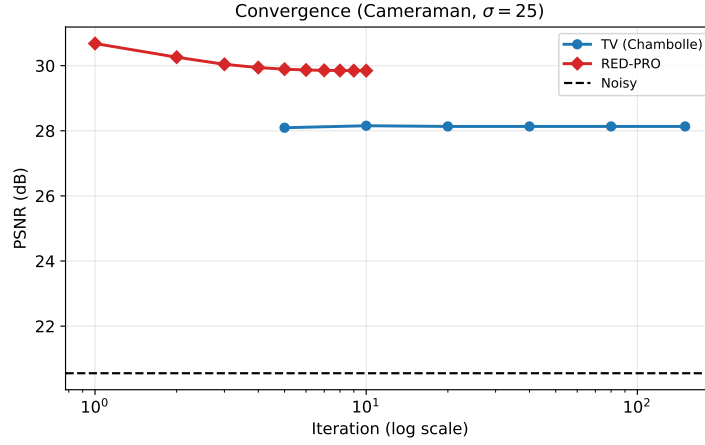


Figure 3: PSNR versus iteration for TV (Chambolle) and RED-PRO. Cameraman, $\sigma = 25$.

6.5. Quality–Runtime Trade-off

Figure 4 positions each method in the PSNR–runtime plane, aggregated over images and noise levels. POCS and wavelet shrinkage are the fastest, at the cost of accuracy. TV lies in the middle of the runtime axis. RED-PRO has the largest runtime, a direct consequence of calling the BM3D denoiser five times per image, but it also attains the highest PSNR. The figure makes explicit that the accuracy gain of RED-PRO over TV comes at a runtime cost of roughly one order of magnitude.

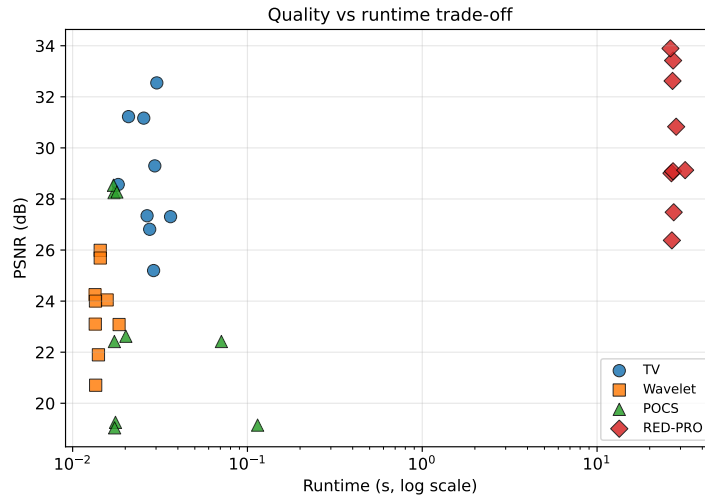


Figure 4: PSNR versus runtime (log scale) for each method, aggregated over images and noise levels.

6.6. Robustness Across Images

Figure 5 shows the distribution of PSNR values across the three images for each noise level. At $\sigma = 10$ and $\sigma = 20$, RED-PRO has the narrowest spread, indicating that its performance is the least sensitive to image content. At $\sigma = 30$, TV is the more consistent of the two, with a range of 2.11 dB against 2.75 dB for RED-PRO. POCS has the widest spread at every noise level, which reflects its dependence on the choice of the fidelity radius ϵ .

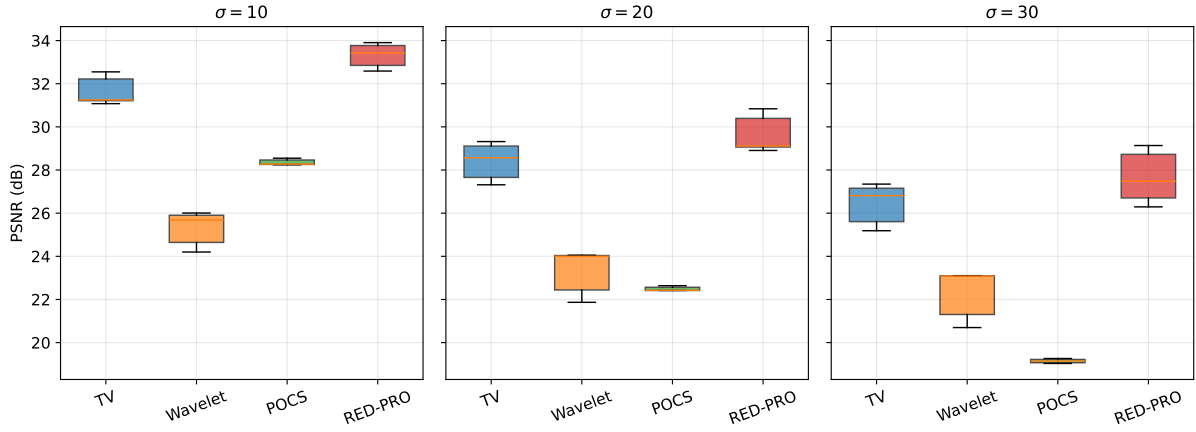


Figure 5: PSNR distributions across images for each noise level.

7. Discussion and Future Directions

This paper has developed a functional-analytic framework in which classical and modern denoising methods are all expressed as iterations of firmly nonexpansive or quasi-nonexpansive operators in a Hilbert space. The convergence theorems of Section 5 apply uniformly to Chambolle’s dual projection for TV, wavelet proximal shrinkage, POCS, PESC, and RED-PRO. The numerical experiments of Section 6 confirm the theory: RED-PRO with BM3D attains the highest PSNR at every noise level tested, TV follows within 1.2–1.7 dB, and both degrade gracefully with σ , whereas POCS and wavelet shrinkage fall behind quickly. The accuracy advantage of RED-PRO comes at roughly an order of magnitude in runtime, so the choice between the two is a genuine quality–cost trade-off rather than a strict dominance.

Three limitations of the present treatment deserve emphasis. First, the convergence results assume convex constraint sets and a demicontractive denoiser; they do not extend directly to non-convex priors or to denoisers that fail the demicontractive condition. Second, the framework is stated in a Hilbert space, which excludes the natural BV setting of TV denoising except through duality. Third, RED-PRO depends on the availability of a well-behaved denoiser and does not prescribe how to obtain one.

Several directions follow from these limitations:

- i. extending the convergence analysis to non-convex constraints and to denoisers satisfying weaker fixed-point conditions;
- ii. establishing rates of convergence, not only asymptotic guarantees, for the projection schemes in Section 5;
- iii. characterizing when a learned denoiser is demicontractive, which would bring deep denoisers within the present theory;
- iv. designing projection operators that preserve texture and fine structure, which none of the four methods tested here achieves fully;
- v. applying the framework to deblurring, super-resolution, and inpainting, where the same operator-theoretic structure carries over.

8. Conclusion

We have presented a functional-analytic framework in which total variation denoising, wavelet shrinkage, POCS, PESC, and RED-PRO all arise as iterations of firmly nonexpansive or quasi-nonexpansive operators on a Hilbert space, with convergence guaranteed by a single Krasnosel'skii–Mann argument. Numerical experiments confirm the practical ordering predicted by the theory: RED-PRO with BM3D attains the highest PSNR at every noise level tested, followed closely by Chambolle's TV projection, while POCS and wavelet shrinkage trade accuracy for speed. The framework clarifies why these methods behave as they do and provides a common language in which further algorithms can be designed and compared.

References

- [1] Rudin, L. I., Osher, S., & Fatemi, E. (1992). Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1–4), 259–268. [https://doi.org/10.1016/0167-2789\(92\)90242-F](https://doi.org/10.1016/0167-2789(92)90242-F)
- [2] Chambolle, A. (2004). An algorithm for total variation minimization and applications. *Journal of Mathematical Imaging and Vision*, 20(1–2), 89–97. <https://doi.org/10.1023/B:JMIV.0000011325.36760.1e>
- [3] Cohen, R., Elad, M., & Milanfar, P. (2021). Regularization by denoising via fixed-point projection (RED-PRO). *SIAM Journal on Imaging Sciences*, 14(3), 1374–1406. <https://doi.org/10.1137/20M1337168>
- [4] Edem, I. D., Akpakwa, J., Okon, U. M., Udogworen, W. K., & Udoaka, O. G. (2026). Strictly pseudo-contractive Volterra integral operators in Banach spaces: Weak and strong convergence of the Mann iteration with a numerical analysis extension. *Ktrend – African Journal of Mathematics, Statistics and Computer Science*, 1(2), 1–16. <https://journal.iorms.org.ng/index.php/ijorms/article/download/21/24>
- [5] Youla, D. C., & Webb, H. (1982). Image restoration by the method of convex projections: Part 1—Theory. *IEEE Transactions on Medical Imaging*, 1(2), 81–94. <https://doi.org/10.1109/TMI.1982.4307555>
- [6] Combettes, P. L. (1993). The foundations of set theoretic estimation. *Proceedings of the IEEE*, 81(2), 182–208. <https://doi.org/10.1109/5.214546>
- [7] Bauschke, H. H., & Borwein, J. M. (1996). On projection algorithms for solving convex feasibility problems. *SIAM Review*, 38(3), 367–426. <https://doi.org/10.1137/S0036144593251710>
- [8] Deutsch, F. (2001). *Best approximation in inner product spaces*. Springer. <https://doi.org/10.1007/978-1-4684-9298-9>
- [9] Bauschke, H. H., & Combettes, P. L. (2011). *Convex analysis and monotone operator theory in Hilbert spaces*. Springer. <https://doi.org/10.1007/978-1-4419-9467-7>
- [10] Combettes, P. L., & Pesquet, J.-C. (2011). Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, & H. Wolkowicz (Eds.), *Fixed-point algorithms for inverse problems in science and engineering* (pp. 185–212). Springer. https://doi.org/10.1007/978-1-4419-9569-8_10

-
- [11] Sevimli, R. A. (2014). Denoising using projections onto the epigraph set of convex cost functions. In *2014 IEEE International Conference on Image Processing (ICIP)* (pp. 1–4). IEEE. <https://doi.org/10.1109/ICIP.2014.7025548>
- [12] Sevimli, R. A., & Guven, H. E. (2016). Projections onto the epigraph set of the filtered variation function based deconvolution algorithm. In *2016 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (pp. 1–5). IEEE. <https://doi.org/10.1109/GlobalSIP.2016.7905805>
- [13] Edem, I. D., Akpakwa, J., Sampson, M., George, R., Awad, M. M., & Nabwey, H. A. (2026). Strong convergence of the Halpern iteration for monotone α -nonexpansive mappings in uniformly convex ordered Banach spaces. *International Journal of Analysis and Applications*, 24, Article 129. <https://etamaths.com/index.php/ijaa/article/download/4942/1715>
- [14] Romano, Y., Elad, M., & Milanfar, P. (2017). The little engine that could: Regularization by denoising (RED). *SIAM Journal on Imaging Sciences*, 10(4), 1804–1844. <https://doi.org/10.1137/16M1102884>
- [15] Venkatakrishnan, S. V., Bouman, C. A., & Wohlberg, B. (2013). Plug-and-play priors for model based reconstruction. In *2013 IEEE Global Conference on Signal and Information Processing (GlobalSIP)* (pp. 945–948). IEEE. <https://doi.org/10.1109/GlobalSIP.2013.6737048>
- [16] Donoho, D. L. (1995). De-noising by soft-thresholding. *IEEE Transactions on Information Theory*, 41(3), 613–627. <https://doi.org/10.1109/18.382009>
- [17] Dabov, K., Foi, A., Katkovnik, V., & Egiazarian, K. (2007). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8), 2080–2095. <https://doi.org/10.1109/TIP.2007.901238>
- [18] Ambrosio, L., Fusco, N., & Pallara, D. (2000). *Functions of bounded variation and free discontinuity problems*. Oxford University Press.
- [19] Edem, I. D., Udoaka, O. G., & Akpan, U. D. (2026). Application of generalized hybrid fixed point theory to nonlinear integral equations in L_p spaces. *IORMS Journal of Operational Research and Management Science*, 4(1), 55–69. <https://journal.iorms.org/index.php/ijorms/article/download/25/25>
- [20] Caselles, V., Chambolle, A., & Novaga, M. (2007). The discontinuity set of solutions of the TV denoising problem and some extensions. *Multiscale Modeling & Simulation*, 6(3), 879–894. <https://doi.org/10.1137/070683215>
- [21] Chambolle, A., Caselles, V., Cremers, D., Novaga, M., & Pock, T. (2010). An introduction to total variation for image analysis. In M. Fornasier (Ed.), *Theoretical foundations and numerical methods for sparse recovery* (pp. 263–340). De Gruyter. <https://doi.org/10.1515/9783110226157.263>
- [22] Ekeland, I., & Temam, R. (1976). *Convex analysis and variational problems*. North-Holland.

-
- [23] Akpakwa, J. I., Udoaka, O. G., Udofia, E. S., & Edem, I. D. (2026). Inertial averaged type algorithm with norm free dynamic stepsize for split common fixed point problems in Banach spaces. *IORMS Journal of Operational Research*, 4(1), 42–54. <https://journal.iorms.org.ng/index.php/ijorms/article/download/21/24>
- [24] Grosu, G. F., & Terebes, R. (2023). Chambolle’s projection-based total variation deepnet layer for image denoising. In *2023 IEEE International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 229–236). IEEE. <https://doi.org/10.1109/ICCP60212.2023.10314847>
- [25] Sevimli, R. A. (2016). *Image restoration and reconstruction using projections onto epigraph set of convex cost functions* [Doctoral dissertation, Bilkent University]. Bilkent University Institutional Repository. <http://repository.bilkent.edu.tr/handle/11693/29074>
- [26] Sheng, X., & Chang, J. (2023). *Weighted structure tensor total variation for image denoising* [Preprint]. arXiv. <https://arxiv.org/abs/2306.10482>
- [27] Unser, M., & Tafti, P. D. (2014). Wavelet-domain methods. In *An introduction to sparse stochastic processes* (Ch. 11). Cambridge University Press.
- [28] Durand, S., Fadili, J., & Nikolova, M. (2021). Frame soft shrinkage operators are proximity operators. *Applied and Computational Harmonic Analysis*, 54, 1–34. <https://doi.org/10.1016/j.acha.2021.03.001>
- [29] Kamilov, U. S., Bouman, C. A., Buzzard, G. T., & Wohlberg, B. (2023). Plug-and-play methods for integrating physical and learned models in computational imaging: Theory, algorithms, and applications. *IEEE Signal Processing Magazine*, 40(1), 85–97.
- [30] Park, C. Y., Hu, Y., McCann, M. T., Garcia-Cardona, C., Wohlberg, B., & Kamilov, U. S. (2024). *Plug-and-play priors as a score-based method* [Preprint]. arXiv. <https://arxiv.org/abs/2412.11108>
- [31] Tofighi, M., Cetin, A. E., & Kose, K. (2014). *Signal reconstruction framework based on projections onto epigraph set of a convex cost function (PESC)* [Preprint]. arXiv. <https://arxiv.org/abs/1402.2088>